MBE Advance Access published June 20, 2010

Clonality and recombination in the life history of an asexual arbuscular mycorrhizal fungus

Research article

Henk C. den Bakker[1*], Nicholas W. VanKuren[1*], Joseph B. Morton[2] & Teresa E. Pawlowska[1]

[*] ***contributed equally***

[1] *Department of Plant Pathology and Plant Microbe-Biology, Cornell University, Ithaca, New York 14853-5904, USA.*

[2] *Division of Plant and Soil Sciences, West Virginia University, Morgantown, West Virginia 26506, USA.*

Corresponding author: Teresa E. Pawlowska
Department of Plant Pathology & Plant-Microbe Biology
Cornell University
334 Plant Sciences Building
Ithaca, NY 14853-5904
Phone: (607) 254-8726
Fax: (607) 255 4471
E-mail: tep8@cornell.edu

Key words: Glomeromycota, *Glomus etunicatum*, reproduction, population structure

Running head: *Glomus etunicatum* population structure

Arbuscular mycorrhizal (AM) fungi in the phylum Glomeromycota colonize roots of the majority of land plants and assist them in the uptake of mineral nutrients in exchange for plant-assimilated carbon. In the absence of sexual reproductive structures and with asexual spore morphology conserved since the Ordovician, Glomeromycota may be one of the oldest eukaryotic lineages that rely predominantly on asexual reproduction for gene transmission. Clonal population structure detected in the majority of AM fungi examined to date supports this hypothesis. However, evidence of recombination found in few local populations suggests that genetic exchanges may be more common in these organisms than is currently recognized. To explore the significance of clonal expansion versus genetic recombination in the life history of modern Glomeromycota, we examined the global population of a cosmopolitan fungus *Glomus etunicatum* and made inferences about the population structure and the occurrence of recombination in the history of this species. We sampled eight loci from 84 isolates. We found that, even though the global population of *G. etunicatum* showed a pattern of significant differentiation, several haplotypes had a broad geographic distribution spanning multiple continents. Molecular variation among the sampled isolates indicated an overwhelmingly clonal population structure and suggested that clonal expansion plays an important role in the ecological success of modern Glomeromycota. On the other hand, a pattern of homoplasy consistent with a history of recombination suggested that gene exchanges are not completely absent from the life history of these organisms although they are likely to be very rare.

Introduction

Arbuscular mycorrhizal (AM) fungi (phylum Glomeromycota) colonize roots of the majority of land plants and form a symbiosis known as arbuscular mycorrhiza, in which the fungi assist plants in the uptake of mineral nutrients in exchange for plant-assimilated carbon (Smith and Read 2008). Despite over a century of exploration, AM fungi remain one of the most elusive soil microorganisms with no evidence of sexual reproduction. The apparent lack of sexual reproduction in extant AM fungi combined with the Ordovician fossil record of structures that resemble modern asexual spores (Redecker, Kodner, and Graham 2000) are taken as evidence that Glomeromycota may be one of the oldest eukaryotic lineages that predominantly rely on asexual reproduction for gene transmission. Given the long evolutionary history (Redecker, Kodner, and Graham 2000; Heckman et al. 2001) and the ubiquity of AM fungi in terrestrial ecosystems (Öpik et al. 2006), the apparent absence of sexual reproduction from their life history is puzzling. Asexual lineages are usually short-lived evolutionary ephemerals (Taylor, Jacobson, and Fisher 1999). Their vulnerability to extinction is related to the limited ability of asexual populations to purge deleterious mutations (Paland and Lynch 2006) and recreate mutation-free individuals (Muller 1964) without the benefit of recombination. Asexual organisms also suffer from impediments in responding to environmental changes because advantageous mutations arising in different individuals in the same population must compete with one another for fixation instead of being recombined into one lineage (Muller 1932; Cooper 2007).

To examine whether the apparent absence of sexual structures and processes from the organism life cycle also indicates the absence of genetic exchanges, the mode of genetic transmission can be inferred from the pattern of genetic variation in a population. Asexual organisms are expected to exhibit a clonal population structure where nucleotides and alleles that make up individual genomes are linked (Maynard Smith et al. 1993), and have common genealogical histories (Maynard Smith and Smith 1998). In contrast, genetic exchanges are

likely to disrupt nucleotide and allele associations, and result in incongruent genealogical histories of individual loci in a genome. The majority of populations of AM fungi that have been examined to date showed no evidence of genetic exchanges. They included several local populations of *Glomus mosseae* and *G. caledonium* from agronomic fields in Denmark (Rosendahl and Taylor 1997; Stukenbrock and Rosendahl 2005; Rosendahl and Matzen 2008) as well as populations of *Glomus claroideum* and an unnamed *Glomus* species from agronomic and low-input sewage sludge sites in Germany (Vandenkoornhuyse, Leyval, and Bonnin 2001). Evidence for recombination was documented only in populations of *Glomus claroideum* and *Glomus* sp. from a high-input sewage sludge site in Germany (Vandenkoornhuyse, Leyval, and Bonnin 2001), and a population of *Glomus intraradices* from an agronomic site in Switzerland (Croll and Sanders 2009). These contradictory patterns indicate that understanding the roles that clonal expansion versus genetic recombination play in the life history of Glomeromycota requires further examination.

To date, all studies aimed at inferring the mode of reproduction in Glomeromycota have been conducted in small local populations. The goal of our study was to expand the sampling of AM fungi beyond the local geographic scale and make inferences about the patterns of genetic variation and the history of gene transmission at the level of an entire species. We focused on *Glomus etunicatum,* a cosmopolitan species with a life history that epitomizes characteristics of many AM fungi, including profuse formation of asexual spores, a broad range of hosts, and the ability to form symbiotic associations in a wide variety of habitats. Despite a pattern of significant differentiation in the global population of *G. etunicatum*, we found haplotypes that occurred on multiple continents. The population structure of *G. etunicatum* was overwhelmingly clonal. Nevertheless, a signature of cryptic recombination was also apparent in the history of the lineage.

Materials and Methods

Fungal material. *Glomus etunicatum* Becker & Gerdemann isolates genotyped in this study represented populations from: (i) two maize experimental fields known as Oxford Tract (CA-OT, N37°52'34" W122°16'01") and Gilman Tract (CA-GT, N37°53'12" W122°17'58") in the Bay Area of San Francisco, CA, and (ii) 22 locations around the globe available as accessions in the International Culture Collection of Arbuscular Mycorrhizal Fungi, INVAM (Table 1). INVAM accessions were started and maintained as multispore cultures to retain genetic diversity of the original environmental samples. Because the taxonomy of Glomeromycota is based at present on morphological characters, we included as outgroups isolates from accessions PL119A and ZM103 whose spores were similar to *G. etunicatum* in all morphological characters except colour, which was yellow rather than the orange to red-brown typical for *G. etunicatum*. We refer to these isolates as *Glomus* PL119A and *Glomus* ZM103.

Greenhouse cultivation. The CA-OT and CA-GT isolates of *G. etunicatum* were sampled from the field by collecting soil-encased maize root systems in November 2003, and then inducing the indigenous fungi to sporulate in greenhouse trap cultures as described by Pawlowska and Taylor (Pawlowska and Taylor 2004). INVAM accessions were cultivated using standard procedures practiced at INVAM (Morton, Bentivenga, and Wheeler 1993). After air-drying, pot culture media containing spores of AM fungi were stored at 4ºC until spores were extracted by wet sieving and sucrose centrifugation (Daniels and Skipper 1982).

***In vitro* cultivation.** Spores of *G. etunicatum* were surface-decontaminated, and used to inoculate two-week-old excised Ri T-DNA transformed carrot roots (clone DC2) grown on the M medium modified with 10 mM MES (pH 6.0) (Pawlowska, Douds, and Charvat 1999). Multiple spores representing a specific population were used to establish each monoxenic culture. From the cohort of first-generation spores formed in each of these *in vitro* cultures, individual spores were collected and analyzed genetically.

Whole-genome amplification of spore genomic DNA. To generate sufficient quantity of genomic template DNA necessary to perform numerous PCR reactions needed to retrieve sequences of several markers from individual spores, spore genomic DNAs were subjected to multiple strand displacement whole-genome (WG) amplification mediated by the ϕ29 DNA polymerase (Dean et al. 2002) available in the GenomiPhi™ DNA Amplification Kit (GE Life Sciences, Piscataway, NJ). Individual spores (monoxenic or soil-derived) were crushed in 9 µl of sample buffer supplied by the manufacturer. WG amplifications were carried out at 30ºC for 16 hrs. WG-amplified spore DNAs were diluted 1: 20 in water and stored at -20ºC.

Nuclear ribosomal RNA gene phylogeny. We examined rRNA large subunit (LSU) gene sequences from isolates of *G. etunicatum*, *Glomus* PL119A, *Glomus* ZM103 as well as from isolates of *Glomus claroideum* accessions BR147A, SF119A, SW201, and *G. luteum* accessions SA112, and SW202 from INVAM to determine: (i) whether the shared morphology of *G. etunicatum* isolates reflects a shared phylogenetic history, and (ii) whether the phylogenetic signal of anonymous genetic markers used in this study is congruent with the phylogenetic signal of genes encoding rRNA, which are widely used in phylogeny and molecular ecology studies of Glomeromycota. LSU sequences were obtained from WG-amplified spore genomic DNAs using the primers LR1 5'-GCATATCAATAAGCGGAGGA-3' and NDL22 5'-TGGTCCGTGTTTCAAGACG-3'(Van Tuinen et al. 1998). PCR reactions contained 19.5 µl ultrapure water, 1 µl template DNA, 2.5 µl 10x PCR reaction buffer (Roche, Indianapolis, IN), 1 µl GeneAmp®dNTPs (2.5 mM each, Applied Biosystems, Foster City, CA), 0.4 µl of 10 pMol per µl solution of the forward primer, 0.4 µl of 10 pMol per µl solution of the reverse primer, and 0.2 µl Taq DNA polymerase (5 U μl^{-1}, Roche). PCR cycling parameters included an initial denaturation step of 2 min at 94ºC followed by 35 cycles of 30 sec at 94ºC, 20 sec at 58ºC, and 90 sec at 72ºC as well as a final extension step of 2 min at 72ºC. Amplicons were sequenced using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and ABI 3730xl DNA Analyzer (Applied Biosystems). Sequences were edited in Sequence Navigator (Applied

Biosystems), automatically aligned in ClustalX 2.0 (Thompson et al. 1997) followed by manual adjustment in MacClade 4.08 (Maddison and Maddison 2005) yielding an alignment of 740 bp. Relationships among isolates were inferred using the methods of maximum likelihood (ML) implemented in PHYML 2.4.5 (Guindon and Gascuel 2003) and Bayesian analysis implemented in MrBayes 3.1.2 (Huelsenbeck and Ronquist 2001). GTR+Γ, the general time reversible model (Tavaré 1986) with gamma-distributed rate variation, was identified by the Akaike Information Criterion (AIC) available through ModelTest Server 1.0 (Posada 2006) as the molecular evolution model best describing the data. Statistical support for the ML tree topology was based on 1000 bootstrap replicates. Bayesian posterior probabilities were inferred based on two runs of 1,000,000 generations with a burn-in of 100,000 generations.

Genetic marker development. Genetic markers suitable for a study of population structure in *G. etunicatum* were identified by screening anonymous genomic fragments retrieved from five genomic libraries constructed individually from monoxenic spores using a protocol modified from Klein et al. (Klein et al. 1999). To construct the libraries, individual spore genomic DNA preparations of CA-OT spores generated as described by Pawlowska and Taylor (2004) were digested with 5 U of *Eco*RI restriction endonuclease (20,000 U ml^{-1}, New England BioLabs, Ipswich, MA) added in 2 μl of buffer containing 0.7 μl of 10x One-Phor-All Buffer *PLUS,* and incubated for 3 hrs at 37°C. The restriction products were then ligated to 3'-phosphorothioated HPLC-purified oligonucleotide adaptors ecolig21PT 5'-TACAGGGATTCCGCATGCTAC-3' and ecolig13PT 5'-AATTGTAGCATGC-3'. Prior to ligation, the adaptors were annealed in a temperature gradient from 65°C (maintained for 1 min) to 15°C with a rate of temperature change of 0.01°C per second. The ligation reactions were assembled by combining the spore genomic *Eco*RI digests with 5 μl of the ligation mix containing 1 μl of 10x One-Phor-All Buffer *PLUS*, 1 μl of the annealed adaptors at a concentration of 100 μM each, 0.1 μl of 100 mM ATP, and 0.18 μl T4 DNA ligase (2,000 U μl^{-1}, New England BioLabs), and incubated for 12 hrs at 15°C followed by 20 min at 65°C. The *Eco*RI genomic fragments ligated to adaptors were

globally amplified after adding 90 µl of a PCR reaction mix containing 10 µl of 10x cloned *Pfu* DNA polymerase reaction buffer (Stratagene, La Jolla, CA), 1 µl of 100 µM ecolig21PT oligonucleotide, 1 µl of 100 µM ecolig13PT oligonucleotide, 10 µl of dNTPs (2 mM each), 2 µl of *PfuTurbo*® DNA polymerase (2.5 U μl^{-1}, Stratagene). PCR cycling conditions were 10 cycles of 40 sec at 94°C, 40 sec at 57°C, 9 min at 72°C followed by 15 cycles of 40 sec at 94°C, 40 sec at 57°C, and 13 min at 72°C. Globally amplified spore genomic DNAs were purified using QIAquick® PCR Purification Kit (Qiagen, Valencia, CA), and eluted from spin columns with 30 µl of 0.1x EB buffer. Purified genomic fragments were blunt-ended by incubating them with 1.6 µl T4 DNA polymerase (3 U μl^{-1}, New England BioLabs), 4 µl of 10x T4 DNA polymerase buffer (New England BioLabs), 4 µl dNTPs (2 mM each), and 4 µl of 5 µg μl^{-1} BSA for 5 min at 37°C followed by 10 min at 75°C. Blunt-ended amplicons were cloned using Zero Blunt® TOPO® PCR Cloning Kit (Invitrogen, Carlsbad, CA). Recombinant plasmid DNAs were isolated with QIAprep® 96 Turbo Miniprep Kit (Qiagen) after recombinant *E. coli* colonies were picked and inoculated into 1.2 ml LB medium with 50 µg ml^{-1} kanamycin. To screen the libraries for candidate markers, 576 genomic clones were analyzed by sequencing with primers T7 and SP6. Sequences of 449 clones were subjected to nucleotide query BLASTN (Altschul et al. 1990) and translated nucleotide query BLASTX searches of the National Center for Biotechnology Information (NCBI) databases. Genomic sequences with no significant similarity to sequences in NCBI databases served as templates to develop marker-specific primers. A total of 58 primer pairs targeting different clones were designed using Primer Express 2.0.0 (Applied Biosystems). Primers were tested in 25 µl PCR reactions containing 19.5 µl ultrapure water, 1 µl template DNA, 2.5 µl 10x PCR reaction buffer (Roche), 1 µl GeneAmp®dNTPs (2.5 mM each, Applied Biosystems), 0.4 µl of 10 pMol per µl solution of forward primer, 0.4 µl of 10 pMol per µl solution of reverse primer, and 0.2 µl Taq DNA polymerase (5 U μl^{-1}, Roche) under the following conditions: an initial denaturation step of 2 min at 94ºC followed by 35 cycles of 30 sec at 94ºC, 20 sec at 52ºC, and 30 sec at 72ºC followed by a final extension step of 2 min. The

amplicons were sequenced using the original PCR primers as sequencing primers. Primers for the eight anonymous markers that were ultimately selected for the study are presented in Table 2.

Assessment of intraindividual marker variability. To ensure that marker sequences were monomorphic within individual spores, marker amplicons obtained from individual spores were sequenced directly as described above. Sequence chromatograms were examined using Sequencher™ 4.7 (Gene Codes Corporation, Ann Arbor, MI) for presence of double peaks at nucleotide positions known to be variable on a population scale. No overlapping peaks that would suggest intraindividual marker polymorphism were detected. To further ascertain that markers are not variable within spores, we analyzed multiple clones obtained from PCR products from WG-amplified genomic DNAs of single spores from several geographic locations (Table S1). Amplicons were cloned using the TOPO TA Cloning® Kit (Invitrogen). Recombinant plasmid DNA was isolated either with the QIAprep® 96 Turbo Miniprep Kit (Qiagen), or by plasmid amplification from positive *E. coli* colonies using TempliPhi™ Amplification Kit (GE Life Sciences). All single nucleotide variability that we detected among clones derived from single spores was attributable only to PCR or sequencing errors.

Isolate genotyping. Genetic markers were PCR-amplified and sequenced from WG-amplified genomic DNA preparations of individual monoxenic or soil-derived spores of *G. etunicatum*, *Glomus* PL119A, and *Glomus* ZM103 (Table 1) using PCR reaction and cycling conditions described in the "Genetic marker development" section. Sequences were edited as described above and concatenated yielding an alignment of 2922 bp after 317 bp were excluded due to missing data. Relationships among the isolates were inferred using the methods of maximum parsimony (MP) implemented in PAUP* 4.0b10 (Swofford 2002) and Bayesian phylogeny reconstruction described earlier. Alignment gaps were coded as binary characters using simple gap coding (Simmons and Ochoterena 2000) implemented in SeqState 12.2.0 (Müller 2006). HKY85+Γ (Hasegawa, Kishino, and Yano 1985) was identified by the hierarchical likelihood ratio tests and the Bayesian Information Criterion implemented by the ModelTest Server (Posada

2006) as the model best describing the data. Statistical support for the MP tree topology was based on 1000 bootstrap replicates; Bayesian posterior probabilities were inferred based on two runs of 4,000,000 generations with a burn-in of 400,000 generations. Phylogenetic analyses allowed us to identify isolates sharing identical marker sequences, and collapse them into haplotypes.

Global population differentiation. To evaluate the amount of geographic differentiation in the global population of *G. etunicatum*, we used analysis of molecular variance AMOVA (Excoffier, Smouse, and Quattro 1992) implemented in Arlequin 3.1 (Excoffier, Laval, and Schneider 2005). We tested the null hypothesis that any variation among the isolates is due to random sampling. We carried out hierarchical analysis of variability among *G. etunicatum* isolates (single spores) within each population, among populations within each continent, and among continents. To estimate variance components and Φ statistics, which are *F* statistic analogues and reflect the correlations of haplotypic diversity at different levels of hierarchical subdivsion, we used a matrix of genetic distances computed in PAUP*4.0b10 (Swofford 2002). Computations of genetic distances were based on mean character differences, which included insertion/deletions coded as binary characters using simple indel coding as described above. The specific Φ statistics were: (i) Φ_{ST}, the correlation of random haplotypes within populations relative to the correlation of random pairs of haplotypes drawn from the whole species, (ii) Φ_{CT}, the correlation of random haplotypes within a group of populations representing a continent relative to the correlation of random pairs of haplotypes drawn from the whole species, and (iii) Φ_{SC}, the correlation of the molecular diversity of random haplotypes within populations relative to the correlation of random pairs of haplotypes drawn from a continent. Statistical significance of the null hypothesis was tested by permutational analysis: 90,000 permuted matrices were generated to obtain the null distribution and to test for the significance of the variance components and the Φ statistics. We also conducted an exact test for population differentiation (Raymond and Rousset 1995) implemented in Arlequin 3.1 (Excoffier, Laval, and Schneider 2005) to examine

the null hypothesis of no differentiation among the isolates collected globally. We estimated the value of Type I error probability of rejecting the null hypothesis when it was true using a Markov chain method with 1,000,000 steps and a 100,000 burn-in.

Mode of gene transmission in *Glomus etunicatum*. To explore the relative contribution of clonal expansion versus recombination in the life history of *G. etunicatum*, we used a battery of methods designed to detect recombination in a population. First, we tested the null hypothesis that alleles at different loci are in linkage equilibrium, which is expected in a randomly mating population (Maynard Smith et al. 1993). Departures from linkage equilibrium can be expressed as the index of association I_A (Maynard Smith et al. 1993) or the standardized index of association I^S_A (Haubold and Hudson 2000). We measured I_A by computing the variance of the number of loci that differed in each pair of haplotypes in the dataset, V_D, followed by comparing V_D to the mismatch variance expected under linkage equilibrium, V_e (Maynard Smith et al. 1993). To obtain the standardized index of association I^S_A, we scaled the index of association I_A by the number of loci used in the analysis (Haubold and Hudson 2000). Both I_A and I^S_A were quantified using START2 (Jolley et al. 2001).

Second, we estimated the rates of recombination (ρ) and mutation (θ) at individual marker loci in the global population of *G. etunicatum* using LDhat (McVean, Awadalla, and Fearnhead 2002) implemented in the RDP3 package (Heath et al. 2006). LDhat utilizes a coalescent-based approach to determine the amount of recombination in the genealogical history of sampled genes under a finite-sites model of mutation.

Third, we tested the null hypothesis of phylogenetic congruence among individual marker loci in the *G. etunicatum* data set by conducting a partition homogeneity test (Farris et al. 1995) implemented in PAUP* (Swofford 2002). Under the model of clonality, genealogical histories of marker genes are expected to be congruent because they are inherited together. Under the model of recombination, genealogies of unlinked genes are likely to be in conflict because

recombination would break up allele associations. This conflict can be expressed as increased homoplasy and increased length of phylogenetic trees. The partition homogeneity test statistics reflects the difference between the numbers of steps required by individual versus combined analyses of the marker genes. The distribution of the test statistic is estimated by calculating it first for the original partition and then for a series of randomized partitions that are of the same size as the original partition but each represents a mixture of characters from all partitions (Cunningham 1997). We ran 100,000 replicates of the randomization procedure on a dataset representing concatenated marker gene sequences with invariant characters removed; alignment gaps were treated as missing data; the maximum number of trees held per replicate was set to 1000.

Fourth, we investigated patterns of recombination in the haplotype dataset using several methods implemented in the RDP3 package (Heath et al. 2006), including BOOTSCAN (Martin et al. 2005), CHIMAERA (Posada and Crandall 2001), GENECONV (Padidam, Sawyer, and Fauquet 1999), MAXCHI (Maynard Smith 1992), RDP (Martin and Rybicki 2000), SISCAN (Gibbs, Armstrong, and Gibbs 2000), and 3SEQ (Boni, Posada, and Feldman 2007). These methods infer recombination if phylogenies of different parts of an alignment yield discordant topologies (BOOTSCAN and RDP) or if a specific nucleotide substitution pattern indicative of recombination is evident (CHIMAERA, GENECONV, MAXCHI, SISCAN, and 3SEQ). All these methods, with the exception of SISCAN, scan three sequences at a time to identify the two parent sequences and the sequence that is a mosaic of the two parents. SISCAN examines triplets together with a fourth outlier sequence, which was the nearest outlier in our case. To scan the sequences, we used a window size of 60 variable nucleotide positions in RDP, MAXCHI, and CHIMAERA, and 1000 nucleotide positions with a step size of 10 nucleotides in BOOTSCAN and SISCAN. Alignment gaps in the concatenated marker dataset were excluded from the analyses. To make the Type I error experiment-wide rather than specific to the sequence triplet under examination, we applied a step-down correction for multiple comparisons.

Fifth, we carried out the single break point (SBP) screen as well as the multiple break points screen using the genetic algorithm recombination detection (GARD) method (Kosakovsky Pond et al. 2006) available through a web interface (Kosakovsky Pond and Frost 2005). These two methods search the sequence alignment for evidence of segment-specific phylogenies and assess goodness of fit using the small sample AIC (AIC_C) criterion derived from a maximum likelihood model fit to each segment. To verify whether the segment-specific topologies were significantly different, the Kishino-Hasegawa (KH) test (Kishino and Hasegawa 1989) was conducted. The KH test estimates the variance of the difference between log likelihood scores of two phylogenetic trees.

Results and Discussion

Relationships among isolates of the *G. etunicatum* lineage. To investigate the global population structure of *G. etunicatum*, we examined isolates (represented by individual spores) sampled from 24 locations around the globe (Table 1). Multiple isolates were analyzed from most of the locations. Two agricultural fields in the Bay Area of San Francisco, CA, were explored more extensively at the local scale by collecting several spatially distributed samples, and surveying multiple individuals per sample. In addition, we included as outgroups populations of *Glomus* PL119A and *Glomus* ZM103 that shared some taxonomic characteristic of *G. etunicatum* while differing from *G. etunicatum* in an important trait of spore colour.

To ensure that the morphologically defined isolates of *G. etunicatum* from diverse habitats and distant geographic locations shared taxonomic characteristics because of shared phylogenetic history rather than due to convergent morphological evolution, we examined phylogenetic relationships among them using a fragment of the rRNA LSU gene, which is commonly used in molecular ecology studies to identify field isolates of AM fungi. We used isolates of *G.*

claroideum and *G. luteum* as outgroups in these analyses. The topologies inferred in the ML and Bayesian analyses of the LSU sequences were largely congruent. The isolates of *G. etunicatum* all grouped together (Figure 1). Isolates of *Glomus* PL119A could not be separated from the isolates of *G. etunicatum*. In contrast, isolates of *Glomus* ZM103 formed a clade distinct from *G. etunicatum* and from the clade that united *G. claroideum* and *G. luteum*. The isolates of these two latter species could not be separated from each other based on the analyzed rRNA LSU gene fragment, which is an indication that the morphological characters used to separate these two putative species may not be phylogenetically informative. Similar results were also obtained from cloned LSU gene fragments (VanKuren and Pawlowska, in preparation).

We also investigated the relationships in the *G. etunicatum* lineage using eight anonymous loci, *A1*, *A2*, *A3*, *A7*, *4A*, *4C*, *4J*, and *5F*, that were sampled from 84 isolates of *G. etunicatum*, three isolates of *Glomus* PL119A, and three isolates of *Glomus* ZM103 (Table 1). From *Glomus* ZM103, only markers *A1*, *A2*, *A3*, *4A*, and *5F* could be included in the analyses because sequences retrieved by primers designed for markers *A7*, *4C*, and *4J* were not alignable with sequences from *G. etunicatum* and *Glomus* PL119A. The marker loci analyzed in *G. etunicatum* were monomorphic within individuals (Table S1), which was a critical precondition for their utility in reconstructing population structure. Indiscriminate sampling of loci polymorphic within individuals would have confounded the analyses. Using a concatenated alignment of marker sequences, we reconstructed relationships among the sampled isolates. Tree topologies obtained from the MP and Bayesian analyses were largely congruent. We determined that all *G. etunicatum* isolates clustered in a well-supported clade distinct from the clades of *Glomus* PL119A and *Glomus* ZM103 (Figure 2). Interestingly, the relationship between the isolates of *G. etunicatum* and *Glomus* PL119A was incongruent with the relationship suggested by the rRNA LSU phylogeny (Figure 1).

Sequences sampled from the globally distributed isolates of *G. etunicatum* could be collapsed into 14 haplotypes revealing a pattern of low sequence divergence with nucleotide diversity per site $\pi = 0.0032$ (Figure 2, Table 1 and 3). Several haplotypes (I, III, IV, V, VI, VII, XI, and XII) were unique to specific geographic locations. Other haplotypes, such as X and XIII, were shared by isolates from multiple globally distributed sites. The two intensely sampled anthropogenic sites in the Bay Area of California harboured three haplotypes. One of them, haplotype X, had a broad geographic distribution whereas the two other haplotypes, XII and XIV, were specific to the local area.

Geographic structure of the *G. etunicatum* global population. To quantify the extent of geographic differentiation among the sampled isolates, we focused on the *G. etunicatum* isolates. Based on the rRNA gene and the anonymous marker sequences, we found that *Glomus* ZM103 isolates were genetically too distant to be analyzed together with the *G. etunicatum* isolates. The relationship between *Glomus* PL119A and *G. etunicatum* was less certain. However, the lack of a fair representation of isolates with close morphological and genetic similarity to *Glomus* PL119A made including them also undesirable because of their obviously biased underrepresentation relative to the isolates of *G. etunicatum sensu stricto*. Based on the findings of a recent study on population differentiation in another cosmopolitan AM fungus *G. mosseae* (Rosendahl, McGee, and Morton 2009), we expected that the global population of *G. etunicatum* would not be differentiated. However, hierarchical analysis of molecular variance among the globally sampled isolates of *G. etunicatum* revealed that only a small fraction of total variance (about 10%) was contained within local populations (Table 4). Much higher levels of divergence were apparent among populations within continents (over 50% of total variance) and among continents (nearly 40% of total variance). Furthermore, the exact test for differentiation among the isolates representing global sampling of *G. etunicatum* lead to the rejection of the null hypothesis of no differentiation ($P < 0.0001$). Through pair-wise comparisons between populations representing different continents, we found that isolates sampled in Africa were

significantly differentiated from the isolates representing North America ($P < 0.0001$), South America ($P = 0.0061$), Europe ($P = 0.0095$), and Australia ($P = 0.0091$) whereas isolates sampled in Australia showed a significant differentiation from isolates representing North America ($P = 0.0001$), and Europe ($P = 0.0286$), and a trend toward increasing differentiation from isolates sampled in South America ($P = 0.0534$). In contrast, no significant differentiation could be detected among isolates from North America, South America, and Europe. Most of these isolates originated from anthropogenic sites and several shared the same haplotype (Table 1). It is therefore likely that their distribution is attributable to human-mediated dispersal.

Mode of gene transmission in *G. etunicatum*. To make inferences about the reproductive mode in *G. etunicatum*, we first quantified the linkage of alleles at the eight loci sampled from the isolates representing the global *G. etunicatum* population using both the classical index of association I_A (Maynard Smith et al. 1993) and the standardized index of association I^S_A (Haubold and Hudson 2000) (Table 5). In a randomly mating population under complete linkage equilibrium, both I_A and I^S_A are expected to be zero. In the dataset of 84 isolates representing the global population of *G. etunicatum*, I_A was 4.9713 whereas I^S_A was 0.7102, and the null hypothesis of linkage equilibrium was rejected ($P < 0.0001$). Linkage disequilibrium strongly indicates that recombination has been absent or rare in the population history. However, under a number of conditions, linkage disequilibrium may be also apparent in populations that do experience recombination (Maynard Smith et al. 1993). These conditions include: (*i*) epistatic fitness interactions between the sampled loci, (*ii*) population subdivision limiting genetic exchanges among recombinant subpopulations (Ohta 1982), (*iii*) epidemic structure in a freely recombining population where successful individuals give rise to epidemic clones, and (*iv*) a low ratio of the rate of recombination to the rate of mutation. To elucidate whether the index of association values measured in *G. etunicatum* indicate that indeed recombination is absent or rare in its global population, we examined whether any of these scenarios may be applicable.

(*i*) *Epistatic fitness interactions among marker loci.* The marker loci examined in this study are randomly selected anonymous genomic regions that, to the best of our knowledge, have no protein-coding functions. Consequently, epistatic fitness interactions and genetic linkage are unlikely to be the underlying cause of the observed of linkage disequilibrium in the population of *G. etunicatum*.

(*ii*) *Population subdivision.* To examine the possibility that linkage disequilibrium is present in the global population of *G. etunicatum* because of population subdivision rather than because of absence of recombination, we made estimates of the indices of association (I_A and I^S_A) separately within each geographic region that showed evidence of gene flow according to the exact test of population differentiation discussed in the previous section. The null hypothesis of linkage equilibrium in each of the examined subpopulations was rejected (Table 5), which indicates that the pattern of linkage disequilibrium in the global population of *G. etunicatum* cannot be explained by population subdivision that would limit gene flow between otherwise recombinant populations.

(*iii*) *Epidemic population structure*. To explore the scenario that linkage disequilibrium in the global population of *G. etunicatum* was caused by epidemic population structure, we estimated the indices of association for clone-corrected data, where identical sequences were collapsed into distinct haplotypes. The null hypothesis of random association of alleles at the marker loci in the clone-corrected data was rejected (Table 5). This result indicates that the structure of the *G. etunicatum* global population cannot be explained by an expansion of successful clones in an otherwise freely recombining population.

(*iv*) *Low relative recombination rate.* Theoretical studies suggest that recombination must be at least 20 times more frequent than mutation for a population to approach linkage equilibrium (Hudson 1994; Maynard Smith 1994). This observation implies that linkage disequilibrium may

be apparent in populations where recombination occurs but is not frequent enough to generate random association of alleles at different loci (Haubold and Hudson 2000). We explored this possibility by estimating the rates of recombination (ρ) and mutation (θ) at marker loci *A3* and *4A* in the global population of *G. etunicatum* using LDhat (McVean, Awadalla, and Fearnhead 2002); variability at the remaining loci was not sufficient to model the population recombination rates. At both loci, the estimated rate of recombination was substantially lower than the rate of mutation (Table 6). Next, we conducted the partition homogeneity test (Farris et al. 1995) to look for phylogenetic incongruence that could be indicative of a history of recombination among the marker loci in the *G. etunicatum* population. The partition homogeneity test revealed a significant conflict among the data partitions representing individual markers ($P = 0.0194$) in the *G. etunicatum* concatenated haplotype dataset. Based on these results we hypothesised that infrequent recombination events may be an underlying cause of the conflict.

To investigate the occurrence of rare recombination events in the *G. etunicatum* population history, we used the algorithms implemented in the RDP3 package (Heath et al. 2006) to identify specific sequences that display signatures of recombination. Using the RDP3 methods with the step-down correction for multiple comparisons, statistical support was obtained for a recombination event with a footprint shared by three sets of sequences: (1) haplotypes I, IV, and XIV (BOOTSCAN $P = 0.0068$, CHIMAERA $P = 0.0045$, GENECONV $P = 0.0218$, MAXCHI $P = 0.0065$, SISCAN $P = 0.0152$, 3SEQ $P = 0.0043$), (2) haplotypes I, V, and XIV (BOOTSCAN $P = 0.0108$, CHIMAERA $P = 0.0171$, MAXCHI $P = 0.0205$, SISCAN $P = 0.0024$, 3SEQ $P = 0.0087$), and (3) haplotypes I, VI, and XIV (BOOTSCAN $P = 0.0023$, CHIMAERA $P = 0.0009$, GENECONV $P = 0.0172$, MAXCHI $P = 0.0009$, SISCAN $P = 0.0062$, 3SEQ $P = 0.0002$). In each of the sequence triplets, haplotypes I and XIV were identified as potential parents, whereas haplotypes IV, V, and VI showed sequence patterns that could be attributed to a recombinant origin.

Collectively, the results of the partition homogeneity test, and the BOOTSCAN, CHIMAERA, GENECONV, MAXCHI, SISCAN, and 3SEQ tests provided evidence of rare cryptic recombination in the global population of *G. etunicatum*. However, this interpretation should be considered with caution. The *G. etunicatum* haplotype data set exhibits a pattern of extreme variation in nucleotide substitution rates across sites ($\alpha = 0.0162$) where most sites have very low substitution rates while some sites are substitution hotspots. Given that none of the methods that indicated recombination in *G. etunicatum* explicitly includes rate variation in the sequence evolution model, such extreme rate variation may be a cause of false positive inferences of recombination events (Posada and Crandall 2001). Moreover, none of the algorithms that incorporate site-to-site rate variation in the evolution model, such as the SBP and GARD methods (Kosakovsky Pond et al. 2006), was able to reject the null hypothesis of no recombination in our data.

The recombination analyses described above were conducted only on the isolates of *G. etunicatum sensu stricto*. However, our phylogenetic analyses indicated that *Glomus* PL119A is also a member of the *G. etunicatum* lineage. Consequently, we re-examined the occurrence of recombination in a data set that contained isolates of *G. etunicatum* as well as *Glomus* PL119A. The inclusion of *Glomus* PL119A in the analysis provided stronger support for recombination in the history of the *G. etunicatum* lineage. The SBP and GARD methods, which explicitly modelled site-to-site rate variation apparent in the data, identified one possible break point at alignment position 1110 located in marker *A3*. The AIC_C score of 8772.88 for the best-fitting GARD model allowing for different topologies of the alignment segments separated by this break point was lower than the AIC_C score of 8805.66 for a model that assumed the same topology for both segments (Figure 3). The topological incongruence between these two segments also was supported by the Kishino-Hasegawa test (Kishino and Hasegawa 1989) ($P = 0.01$). Interestingly, the RDP3 algorithms did not infer any additional recombination events in the dataset expanded by the inclusion of *Glomus* PL119A. This result suggested that the failure

of the SBP and GARD methods to detect recombination in the *G. etunicatum* dataset that excluded *Glomus* PL119A was likely caused by low sequence divergence. Nucleotide diversity in the dataset including *G. etunicatum* and *Glomus* PL119A was more than twice as high (π = 0.0083) as that in the *G. etunicatum* data without *Glomus* PL119A (π = 0.0032). The effect of the level of sequence divergence on the power of different methods to detect recombination is well established (Posada and Crandall 2001; Posada 2002; Kosakovsky Pond et al. 2006).

Overall, the reconstruction of the population structure in the cosmopolitan AM fungus *G. etunicatum* showed that, even though the population was significantly differentiated, several haplotypes associated with anthropogenic habitats had broad geographic distribution, which spanned multiple continents and represented gene flow on a worldwide scale. Inferences of the gene transmission mode revealed that the global population is overwhelmingly clonal. However, a pattern of homoplasy consistent with a history of recombination suggested that gene exchanges are not completely absent from the life history of these organisms although they are likely to be very rare. These observations are consistent with the reports of recombination in local populations of Glomeromycota (Vandenkoornhuyse, Leyval, and Bonnin 2001; Croll and Sanders 2009). Biological mechanisms that generated the signatures of recombination apparent in the *G. etunicatum* lineage are uncertain. They may involve a cryptic and yet undiscovered sexual process or a process resembling the parasexual cycle of ascomycetes (Pontecorvo 1956). In *G. intraradices*, vegetative hyphal fusions between genetically differentiated individuals were observed under laboratory conditions (Croll et al. 2009). A similar process could be responsible for gene exchanges in Glomeromycota in nature.

Clonal expansion is the predominant mode of propagation and most likely the underlying cause of the ecological success in extant populations of Glomeromycota. However, the dominant role of clonal reproduction in the life history of Glomeromycota is clearly not a recent

development. A footprint of a long history of clonality is apparent in the genomes of these organisms. Ribosomal RNA genes of Glomeromycota show a pattern of departure from the concerted evolution model (Pawlowska and Taylor 2004; Rodriguez et al. 2005). In other eukaryotes, the recombination-driven process of concerted evolution is responsible for homogeneity of rRNA gene repeats in individual genomes and populations (Arnheim et al. 1980; Dover 1982). In Glomeromycota, intraindividual and intraspecific rRNA gene polymorphisms are one of the most striking patterns of molecular evolution that set this lineage apart from other organisms (Pawlowska 2007; Rosendahl 2008).

The importance of recombination in the biology of Glomeromycota may be reduced by mechanisms associated with the mode of their asexual spore development (Jany and Pawlowska 2010). The multinucleate spores of Glomeromycota are initiated by sampling random nuclei from the surrounding mycelium. Such mode of sporogenesis, combined with selection against deleterious mutations acting on individual nuclei, may contribute to purging deleterious mutations from the populations in the absence of recombination. Nevertheless, Glomeromycota are clearly capable of exchanging genes and the significance of these cryptic exchanges in the evolution of this lineage should not be ignored.

GenBank accession numbers: rRNA LSU, HM485692 – HM485774; *A1,* EF627536 – EF627633, HM485774 – HM485801; *A2,* EF627634 – EF627721, HM485802 – HM485832; *A3,* EF627722 – EF627820, HM485833 – HM485859; *A7,* EF630637 – EF630723, HM485860 – HM485887; *4A*, EF630819 – EF630911, HM485888 – HM485917; *4C*, EF630912 – EF631023, HM485918 – HM485950; *4J*, HM485983 – HM486072; *5F*, EF641323 – EF641420, HM485951 – HM485982.

Supplementary Material

Table S1. The number of clones analyzed from different *G. etunicatum* isolates to assess intraindividual polymorphism of the eight marker genes.

Acknowledgments

We are grateful to David Douds for the gift of Ri T-DNA transformed carrot roots developed by Guillaume Bécard. We also thank Meghan den Bakker and Rima Shamieh for excellent technical assistance. This work was supported by NSF grant MCB0538363.

References

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. J Mol Biol **215**:403-410.

Arnheim, N., M. Krystal, R. Schmickel, G. Wilson, O. Ryder, and E. Zimmer. 1980. Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes. P Natl Acad Sci USA **77**:7323-7327.

Boni, M. F., D. Posada, and M. W. Feldman. 2007. An exact nonparametric method for inferring mosaic structure in sequence triplets. Genetics **176**:1035-1047.

Cooper, T. F. 2007. Recombination speeds adaptation by reducing competition between beneficial mutations in populations of *Escherichia coli*. PLoS Biology **5**:1899-1905.

Croll, D., M. Giovannetti, A. M. Koch, C. Sbrana, M. Ehinger, P. J. Lammers, and I. R. Sanders. 2009. Nonself vegetative fusion and genetic exchange in the arbuscular mycorrhizal fungus *Glomus intraradices*. New Phytol **181**:924-937.

Croll, D., and I. R. Sanders. 2009. Recombination in *Glomus intraradices*, a supposed ancient asexual arbuscular mycorrhizal fungus. BMC Evol Biol **9**:13.

Cunningham, C. W. 1997. Can three incongruence tests predict when data should be combined? Mol Biol Evol **14**:733-740.

Daniels, B. A., and H. D. Skipper. 1982. Methods for the recovery and quantitative estimation of propagules from soil. Pp. 29-35 *in* N. C. Schenck, ed. Methods and Principles of Mycorrhizal Research. The American Phytopathological Society, St. Paul, MN.

Dean, F. B., S. Hosono, L. H. Fang, X. H. Wu, A. F. Faruqi, P. Bray-Ward, Z. Y. Sun, Q. L. Zong, Y. F. Du, J. Du, M. Driscoll, W. M. Song, S. F. Kingsmore, M. Egholm, and R. S.

Lasken. 2002. Comprehensive human genome amplification using multiple displacement amplification. P Natl Acad Sci USA **99**:5261-5266.

Dover, G. 1982. Molecular drive: a cohesive mode of species evolution. Nature **299**:111-117.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. Evol Bioinform **1**:47-50.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics **131**:479-491.

Farris, J. S., M. Källersjö, A. G. Kluge, and C. Bult. 1995. Testing significance of incongruence. Cladistics:315-319.

Gibbs, M. J., J. S. Armstrong, and A. J. Gibbs. 2000. Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. Bioinformatics **16**:573-582.

Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biol **52**:696-704.

Hasegawa, M., H. Kishino, and T. A. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol **22**:160-174.

Haubold, B., and R. R. Hudson. 2000. LIAN 3.0: detecting linkage disequilibrium in multilocus data. Bioinformatics **16**:847-848.

Heath, L., E. van der Walt, A. Varsani, and D. P. Martin. 2006. Recombination patterns in aphthoviruses mirror those found in other picornaviruses. J Virol **80**:11827-11832.

Heckman, D. S., D. M. Geiser, B. R. Eidell, R. L. Stauffer, N. L. Kardos, and S. B. Hedges. 2001. Molecular evidence for the early colonization of land by fungi and plants. Science **293**:1129-1133.

Hudson, R. R. 1994. Analytical results concerning linkage disequilibrium in models with genetic transformation and conjugation. J Evol Biol **7**:535-548.

Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17**:754-755.

Jany, J. L., and T. E. Pawlowska. 2010. Multinucleate spores contribute to evolutionary longevity of asexual Glomeromycota. Am Nat **175**:424-435.

Jolley, K. A., E. J. Feil, M. S. Chan, and M. C. J. Maiden. 2001. Sequence type analysis and recombinational tests (START). Bioinformatics **17**:1230-1231.

Kishino, H., and M. Hasegawa. 1989. Evaluation of the maximum-likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J Mol Evol **29**:170-179.

Klein, C. A., O. Schmidt-Kittler, J. A. Schardt, K. Pantel, M. R. Speicher, and G. Riethmueller. 1999. Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. P Natl Acad Sci USA **96**:4494-4499.

Kosakovsky Pond, S. L., and S. D. W. Frost. 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics **21**:2531-2533.

Kosakovsky Pond, S. L., D. Posada, M. B. Gravenor, C. H. Woelk, and S. D. W. Frost. 2006. Automated phylogenetic detection of recombination using a genetic algorithm. Mol Biol Evol **23**:1891-1901.

Maddison, D. R., and W. P. Maddison. 2005. MacClade 4.08: Analysis of Phylogeny and Character Evolution. Sinauer Associates, Inc., Sunderland, MA.

Martin, D., and E. Rybicki. 2000. RDP: detection of recombination amongst aligned sequences. Bioinformatics **16**:562-563.

Martin, D. P., D. Posada, K. A. Crandall, and C. Williamson. 2005. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. AIDS Res Hum Retrov **21**:98-102.

Maynard Smith, J. 1994. Estimating the minimum rate of genetic transformation in bacteria. J Evol Biol **7**:525-534.

Maynard Smith, J. 1992. Analyzing the mosaic structure of genes. J Mol Evol **34**:126-129.

Maynard Smith, J., and N. H. Smith. 1998. Detecting recombination from gene trees. Mol Biol Evol **15**:590-599.

Maynard Smith, J., N. H. Smith, M. O'Rourke, and B. G. Spratt. 1993. How clonal are bacteria? P Natl Acad Sci USA **90**:4384-4388.

McVean, G., P. Awadalla, and P. Fearnhead. 2002. A coalescent-based method for detecting and estimating recombination from gene sequences. Genetics **160**:1231-1241.

Morton, J. B., S. P. Bentivenga, and W. W. Wheeler. 1993. Germ plasm in the International Collection of Arbuscular and Vesicular-Arbuscular Mycorrhizal Fungi (INVAM) and procedures for culture development, documentation and storage. Mycotaxon **48**:491-528.

Muller, H. J. 1964. The relation of recombination to mutational advance. Mutat Res **1**:2-9.

Muller, H. J. 1932. Some genetic aspects of sex. Am Nat **66**:118-138.

Müller, K. 2006. Incorporating information from length-mutational events into phylogenetic analysis. Mol Phylogenet Evol **38**:667-676.

Ohta, T. 1982. Linkage disequilibrium with the island model. Genetics **101**:139-155.

Öpik, M., M. Moora, J. Liira, and M. Zobel. 2006. Composition of root-colonizing arbuscular mycorrhizal fungal communities in different ecosystems around the globe. J Ecol **94**:778-790.

Padidam, M., S. Sawyer, and C. M. Fauquet. 1999. Possible emergence of new geminiviruses by frequent recombination. Virology **265**:218-225.

Paland, S., and M. Lynch. 2006. Transitions to asexuality result in excess amino acid substitutions. Science **311**:990-992.

Pawlowska, T. E. 2007. How the genome is organized in the Glomeromycota. Pp. 419-430 *in* J. Heitman, J. W. Kronstad, J. W. Taylor, and L. A. Casselton, eds. Sex in Fungi. Molecular Determination and Evolutionary Implications. ASM Press, Washington, D.C.

Pawlowska, T. E., D. D. Douds, and I. Charvat. 1999. *In vitro* propagation and life cycle of the arbuscular mycorrhizal fungus *Glomus etunicatum*. Mycol Res **103**:1549-1556.

Pawlowska, T. E., and J. W. Taylor. 2004. Organization of genetic variation in individuals of arbuscular mycorrhizal fungi. Nature **427**:733-737.

Pontecorvo, G. 1956. Parasexual cycle in fungi. Annual Review of Microbiology **10**:393-400.

Posada, D. 2002. Evaluation of methods for detecting recombination from DNA sequences: Empirical data. Mol Biol Evol **19**:708-717.

Posada, D. 2006. ModelTest Server: a web-based tool for the statistical selection of models of nucleotide substitution online. Nucleic Acids Res **34**:W700-W703.

Posada, D., and K. A. Crandall. 2001. Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. P Natl Acad Sci USA **98**:13757-13762.

Raymond, M., and F. Rousset. 1995. An exact test for population differentiation. Evolution **49**:1280-1283.

Redecker, D., R. Kodner, and L. E. Graham. 2000. Glomalean fungi from the Ordovician. Science **289**:1920-1921.

Rodriguez, A., J. P. Clapp, L. Robinson, and J. C. Dodd. 2005. Studies on the diversity of the distinct phylogenetic lineage encompassing *Glomus claroideum* and *Glomus etunicatum*. Mycorrhiza **15**:33-46.

Rosendahl, S. 2008. Communities, populations and individuals of arbuscular mycorrhizal fungi. New Phytol **178**:253-266.

Rosendahl, S., and H. B. Matzen. 2008. Genetic structure of arbuscular mycorrhizal populations in fallow and cultivated soils. New Phytol **179**:1154-1161.

Rosendahl, S., P. McGee, and J. B. Morton. 2009. Lack of global population genetic differentiation in the arbuscular mycorrhizal fungus *Glomus mosseae* suggests a recent range expansion which may have coincided with the spread of agriculture. Mol Ecol **18**:4316-4329.

Rosendahl, S., and J. W. Taylor. 1997. Development of multiple genetic markers for studies of genetic variation in arbuscular mycorrhizal fungi using AFLP$^{(TM)}$. Mol Ecol **6**:821-829.

Simmons, M., and H. Ochoterena. 2000. Gaps as characters in sequence-based phylogenetic analyses. Systematic Biol **49**:369-381.

Smith, S. E., and D. J. Read. 2008. Mycorrhizal Symbiosis. Academic Press, New York.

Stukenbrock, E. H., and S. Rosendahl. 2005. Clonal diversity and population genetic structure of arbuscular mycorrhizal fungi (*Glomus* spp.) studied by multilocus genotyping of single spores. Mol Ecol **14**:743-752.

Swofford, D. L. 2002. PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods). 4.0 Beta. Sinauer Associates, Inc., Sunderland, MA.

Tavaré, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Lect Math Life Sci **17**:57-86.

Taylor, J. W., D. J. Jacobson, and M. C. Fisher. 1999. The evolution of asexual fungi: Reproduction, speciation and classification. Annu Rev Phytopathol **37**:197-246.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res **25**:4876-4882.

Van Tuinen, D., E. Jacquot, B. Zhao, A. Gollotte, and V. Gianinazzi-Pearson. 1998. Characterization of root colonization profiles by a microcosm community of arbuscular mycorrhizal fungi using 25S rDNA-targeted nested PCR. Mol Ecol **7**:879-887.

Vandenkoornhuyse, P., C. Leyval, and I. Bonnin. 2001. High genetic diversity in arbuscular mycorrhizal fungi: evidence for recombination events. Heredity **87**:243-253.

Figure 1. Relationships of *G. etunicatum, Glomus* PL119A, *Glomus* ZM103, *G. luteum* and *G. claroideum* reconstructed based on rRNA LSU gene sequences. Bayesian posterior probabilities over 0.8 are shown above branches; ML bootstrap support over 70% is given below branches.

Figure 2. Relationships among the isolates of *G. etunicatum*, *Glomus* PL119A, *Glomus* ZM103 reconstructed using concatenated sequences of eight anonymous marker loci, which allows for identification of 14 haplotypes among the *G. etunicatum* isolates. Blocks of identical sequences are shaded with alternating tones of grey. Haplotypes are identified by roman numerals. Bayesian posterior probabilities over 0.8 are shown above branches; MP bootstrap support over 70% is given below branches.

Figure 3. Segment-specific ML topologies reconstructed for two incongruent segments of the *G. etunicatum* and *Glomus* PL119A haplotype sequence alignment. The breakpoint at position 1110 was identified by the SBP and GARD methods and is supported by the Kishino-Hasegawa test ($P < 0.01$). Values below branches represent ML bootstrap support over 70% (1000 replicates); Bayesian posterior probabilities above 0.8 obtained from two runs of 1,000,000 generations with a 100,000 generations burn-in are shown above branches. **A**. Topology for the segment of nucleotide positions 1 - 1110. **B**. Topology for the segment of nucleotide positions 1111 - 2914.

Table 1. Geographic origin, habitat type, spore source, and haplotype of globally distributed isolates of *G. etunicatum, Glomus* PL119A, and *Glomus* ZM103.

Isolate	Population	Geographic origin (province, country)	Habitat	Spore source	Haplotype
AU401A-1	AU401A	Queensland, Australia	A	M	I
AU401A-2	AU401A	Queensland, Australia	A	M	I
AU401A-41	AU401A	Queensland, Australia	A	M	I
AZ201C-1	AZ201C	Arizona, US	A	M	VII
AZ201C-2	AZ201C	Arizona, US	A	M	VII
AZ201C-3	AZ201C	Arizona, US	A	M	VII
AZ201C-4	AZ201C	Arizona, US	A	M	VII
AZ201C-5	AZ201C	Arizona, US	A	M	VII
BR220-1	BR220	Santa Catarina, Brazil	A	M	X
BR220-2	BR220	Santa Catarina, Brazil	A	M	X
CA301-1	CA301	California, US	A	M	XIII
CA301-2	CA301	California, US	A	M	XIII
CA-GT4-5	CA-GT	California, US	A	S	XIV
CA-GT4-7	CA-GT	California, US	A	S	XIV
CA-GT10-6	CA-GT	California, US	A	S	XII
CA-GT10-7	CA-GT	California, US	A	S	XII
CA-GT10-8	CA-GT	California, US	A	S	XII
CA-GT10-9	CA-GT	California, US	A	S	XII
CA-GT24-1	CA-GT	California, US	A	S	XII
CA-GT24-2	CA-GT	California, US	A	S	XII
CA-GT24-3	CA-GT	California, US	A	S	XII
CA-GT24-5	CA-GT	California, US	A	S	XII
CA-GT24-6	CA-GT	California, US	A	S	XII
CA-GT39-1	CA-GT	California, US	A	M	X
CA-GT39-3	CA-GT	California, US	A	M	X
CA-OT106-10	CA-OT	California, US	A	S	X
CA-OT107-41	CA-OT	California, US	A	M	X
CA-OT107-43	CA-OT	California, US	A	M	X
CA-OT107-45	CA-OT	California, US	A	M	XIV
CA-OT112-1	CA-OT	California, US	A	S	X
CA-OT112-2	CA-OT	California, US	A	S	XIV
CA-OT112-3	CA-OT	California, US	A	S	XIV
CA-OT112-8	CA-OT	California, US	A	S	XIV
CA-OT112-9	CA-OT	California, US	A	S	X
CA-OT114-5	CA-OT	California, US	A	S	XIV
CA-OT114-6	CA-OT	California, US	A	S	XIV
CA-OT114-22	CA-OT	California, US	A	S	XIV
CA-OT122-41	CA-OT	California, US	A	M	XIV
CA-OT124-34	CA-OT	California, US	A	M	XIV
CA-OT124-35	CA-OT	California, US	A	M	XIV
CA-OT126-34	CA-OT	California, US	A	M	XIV

CA-OT134-11	CA-OT	California, US	A	M	XIV
CA-OT140-31	CA-OT	California, US	A	M	XIV
CL372-1	CL372	Colombia	A	M	XIII
CL372-2	CL372	Colombia	A	M	XIII
CU127-41	CU127	Cuba	N (grassland)	M	VIII
CU127-47	CU127	Cuba	N (grassland)	M	VIII
FL705A-4	FL705A	Florida, US	A	M	IV
FL705A-5	FL705A	Florida, US	A	M	IV
IA205-1	IA205	Iowa, US	A	M	XIII
IA205-2	IA205	Iowa, US	A	M	XIII
KE118-2	KE118	Kenya	U	M	V
KE118-13	KE118	Kenya	U	S	VI
KE118-37	KE118	Kenya	U	M	III
KS887-1	KS887	Kansas, US	N (grassland)	M	X
KS887-2	KS887	Kansas, US	N (grassland)	M	X
KS887-3	KS887	Kansas, US	N (grassland)	M	X
KS887-4	KS887	Kansas, US	N (grassland)	M	X
KS887-5	KS887	Kansas, US	N (grassland)	M	X
MA104-1	MA104	Massachusetts, US	U	M	XIII
MA104-2	MA104	Massachusetts, US	U	M	XIII
MD127-2	MD127	Maryland, US	A	M	XIII
MD127-4	MD127	Maryland, US	A	M	XIII
MR102A-1	MR102A	Morocco	A	M	XI
MR102A-2	MR102A	Morocco	A	M	XI
MR102A-3	MR102A	Morocco	A	M	XI
MR102A-4	MR102A	Morocco	A	M	XI
MX916B-2	MX916B	Zacatecas, Mexico	U	S	II
NB119-17	NB119	Namibia	N (dune)	M	XI
NB119-18	NB119	Namibia	N (dune)	M	XI
NE102-1	NE102	Nebraska, US	U	M	IX
NE102-2	NE102	Nebraska, US	U	M	IX
PL118B-1	PL118B	Poland	U	M	VIII
PL118B-2	PL118B	Poland	U	M	VIII
PL119A-1	PL119A	Poland	U	S	
PL119A-3	PL119A	Poland	U	S	
PL119A-4	PL119A	Poland	U	S	
SP108C-1	SP108C	Spain	U	M	X
SP108C-2	SP108C	Spain	U	M	X

TN101B-1	TN101B	Tennessee, US	A	M	X
TN101B-2	TN101B	Tennessee, US	A	M	X
TX104-1	TX104	Texas, US	A	M	X
TX104-2	TX104	Texas, US	A	M	X
VZ102-2	VZ102	Venezuela	N	M	II
WV579A-1	WV579A	West Virginia, US	A	M	X
WV579A-2	WV579A	West Virginia, US	A	M	X
WV579A-4	WV579A	West Virginia, US	A	M	X
ZM103-1	ZM103	Zimbabwe	U	S	
ZM103-2	ZM103	Zimbabwe	U	S	
ZM103-3	ZM103	Zimbabwe	U	S	

A, anthropogenic; N, natural; U, unknown; S, soil-based pot culture; M, monoxenic culture.

Table 2. Primer sequences for eight anonymous markers used to study the population structure of *G. etunicatum*. GenBank accession numbers identify sequences of genomic clones that served as templates to develop the anonymous markers.

Marker	Forward primer	Reverse primer	GenBank accession
A1	TGTCAAAATGAAGAAAACCCAAT	GTGCTGGTTTGAACGTTGCT	EF627525
A2	TAAACCATCGTCGTTCTTTGAA	AATTTGAATTGAATTTGTTGTCCA	EF627526
A3	TGGGTTCATCTGCATTACTTAGAA	TCTGCAGTCGAGAAAAGTTTTTAAT	EF627527
A7	TGAAGAATTGAGTGAAGAATAGAGTGA	TTTTGCTATTAAAATCTTTCGCCTA	EF627529
4A	CGCTTAAACTTTGAAAAATTGAGAA	TGAAGCATGCAGCAGTAGGA	EF627531
4C	TCGGGTATATGCAGGCTGA	AAGGCCAATCTTTTGGGACT	EF627532
4J	GAAGTGCGAATTACTTGGAAGG	GCAATGTTGATGATGGTGTTG	EF627534
5F	GAGATCATTGAGATCAAACAGCA	TGGTCCTTTTCTTCAACCTCA	EF627535

Table 3. Haplotype structure in the *G. etunicatum* global population. Dots indicate identity with the reference sequence of haplotype I.

	Marker																														
	A1			*A2*		*A3*						*A7*							*4A*									*4C*	*4J*	*5F*		
	Polymorphic position																														
	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2		
	0	0	1	6	6	7	7	8	9	9	1	3	3	3	3	3	4	4	5	5	6	6	6	6	7	7	8	9	4	7	7		
	0	4	6	2	5	4	6	1	9	9	0	0	1	3	3	3	3	5	1	2	4	4	4	6	7	7	0	7	7	8	9		
	0	9	1	1	9	2	4	6	4	6	6	8	1	2	1	2	3	1	7	6	4	6	7	8	3	2	3	1	8	7	5	0	
I	G	-	T	A	-	A	G	G	T	A	C	-	A	-	-	-	G	G	A	T	T	G	T	C	T	G	A	A	G	G	C
II	.	.	.	.	.	.	.	.	.	.	.	.	.	A	G	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
III	.	.	.	.	.	.	T	.	.	.	.	.	.	A	G	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
IV	.	.	.	.	.	.	T	.	.	.	.	.	.	A	G	A	A	.	T	G	C	.	C	T	C	A	-	.	.	.	.	.
V	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	G	C	.	C	T	C	A	-	.	.	.	.	.
VI	.	.	.	.	.	.	.	.	.	.	.	A	.	A	G	A	A	-	T	G	C	A	C	T	C	.	.	.	.	.	.	.
VII	.	A	G	G	.	T	.	.	C	T	T	.	.	A	G	A	A	.	T	G	C	.	C	T	C	.	.	-	A	A	T	
VIII	.	A	G	G	.	T	.	.	C	T	T	.	T	A	G	A	A	.	T	G	C	.	C	T	C	.	.	-	A	A	T	
IX	.	A	G	G	T	T	.	A	C	T	T	.	T	A	G	A	A	.	T	G	C	.	C	T	C	.	.	-	A	A	T	
X	.	A	G	G	T	T	.	.	C	T	T	.	T	A	G	A	A	.	T	G	C	.	C	T	C	.	.	-	A	A	T	
XI	T	A	G	G	T	T	.	.	C	T	T	.	.	A	G	A	A	.	T	G	C	.	C	T	C	.	.	-	A	A	T	
XII	T	A	G	G	T	T	.	.	C	T	T	.	T	A	G	A	A	.	T	G	C	.	C	T	C	.	.	-	A	A	T	
XIII	T	A	G	G	.	T	.	.	C	T	T	.	T	A	G	A	A	.	T	G	C	.	C	T	C	.	.	-	A	A	T	
XIV	T	A	G	G	.	T	.	.	C	T	T	.	T	A	G	A	A	.	T	G	C	A	C	T	C	.	.	-	A	A	T	

Table 4. Hierarchical analysis of molecular variance among the globally collected isolates of *G. etunicatum*.

Variance component	df	Variance	Percentage of variation	Φ statistics	*P*
Among continents	4	0.00044	37.27	Φ_{CT} = 0.37269	0.0321
Among populations within continents	19	0.00063	50.97	Φ_{SC} = 0.81246	<0.0001
Within populations	60	0.00014	11.76	Φ_{ST} = 0.88235	<0.0001

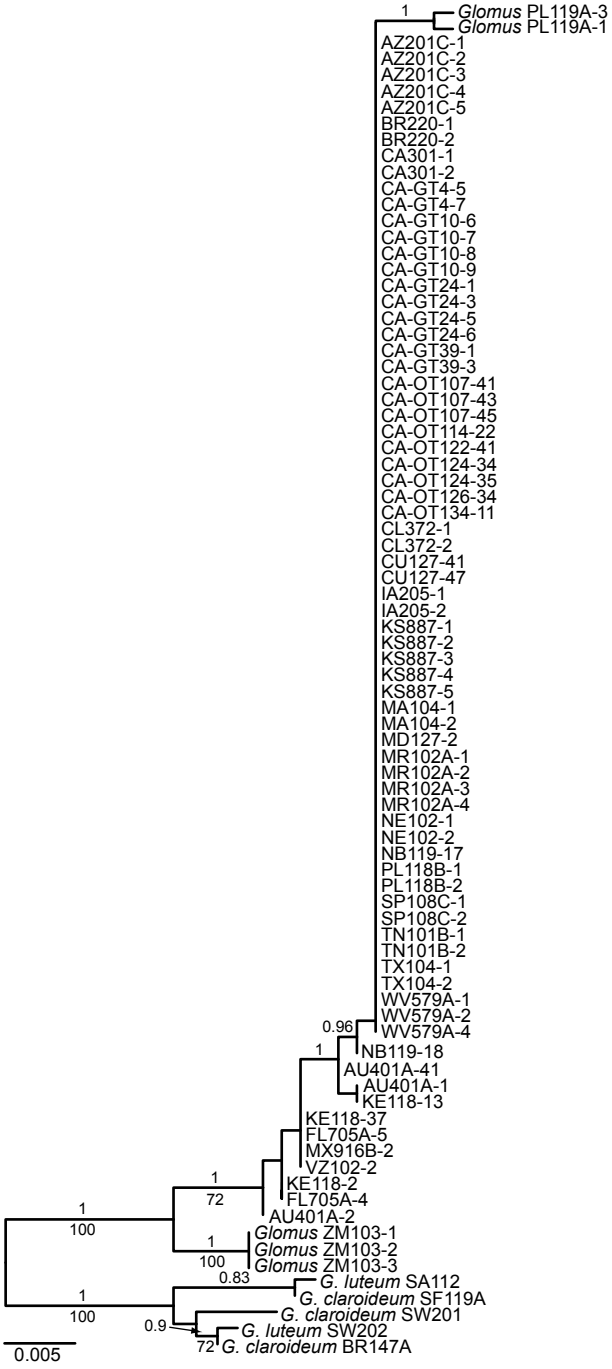
Table 5. Analysis of linkage equilibrium in the *G. etunicatum* global population.

	Global isolates	Global haplotypes	American and European isolates	Australian and South American isolates	African isolates
Number of sequences	84	14	68	8	9
Observed mismatch variance V_D	7.4395	8.8767	5.0778	13.0728	14.054
Expected mismatch variance V_e	1.2459	1.5363	0.9926	1.6518	1.7400
Index of association I_A	4.9713	4.7780	4.1159	6.9143	7.0771
P	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Standardized index of association I^S_A	0.7102	0.6826	0.5880	0.9878	1.1011
P	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

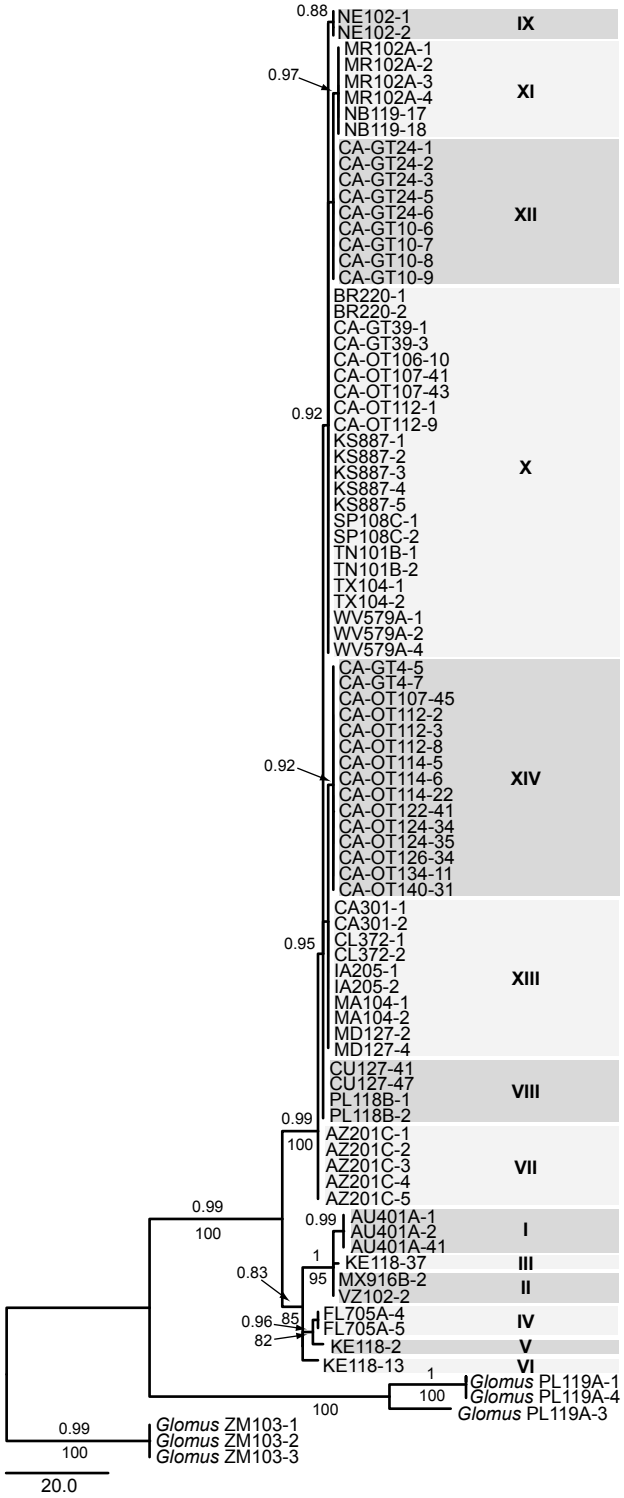
Table 6. Estimates of the ratio of the rate of recombination (ρ) to the rate of mutation (θ) in *G. etunicatum*; ρ_{lower} and ρ_{upper} represent upper and lower bounds of the 95% confidence interval.

Locus	ρ	θ	ρ_{lower}	ρ_{upper}	ρ/θ	ρ_{lower}/θ	ρ_{upper}/θ
A3	0.277	3.272	0.011	0.990	0.085	0.003	0.303
4A	0.095	4.364	0.005	0.345	0.022	0.001	0.079

den Bakker et al. Figure 1



den Bakker et al. Figure 2



den Bakker et al. Figure 3