DR. YONG ZHENG (Orcid ID : 0000-0001-5239-9824)

DR. CHENG GAO (Orcid ID : 0000-0003-2522-7909)

PROF. LIANG-DONG GUO (Orcid ID : 0000-0002-6283-3548)

Received Date : 29-Jan-2016

Revised Date : 05-Jan-2017

Accepted Date : 06-Feb-2017

Article type : Original Article

Phylogenetic relatedness explains highly interconnected and nested symbiotic networks of woody plants and arbuscular mycorrhizal fungi in a Chinese subtropical forest

LIANG CHEN,*† YONG ZHENG,* CHENG GAO,* XIANG-CHENG MI,‡ KE-PING MA,‡ TESFAYE WUBET§¶ and LIANG-DONG GUO*†

**State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China; †College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China; ‡State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China; §Department of Soil Ecology, UFZ - Helmholtz Centre for Environmental Research, Halle (Saale) 06120, Germany;*

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.14061

¶The German Centre for Integrative Biodiversity Research (iDiv), University Leipzig, Leipzig 04103, Germany

Author for correspondence:

Liang-Dong Guo

Tel: +86 10 64807510

Email: guold@sun.im.ac.cn

Keywords: arbuscular mycorrhizal fungi, phylogenetic host specificity, phylogenetic niche conservatism, subtropical forest, symbiotic network properties, woody plants

Running title: Determinants of plant-AMF network structure

Abstract

Elucidating symbiotic relationships between arbuscular mycorrhizal fungi (AMF) and plants contributes to a better understanding of their reciprocally dependent co-existence and community assembly. However, the main drivers of plant and AMF community assembly remain unclear. In this study, we examined AMF communities from 166 root samples of 17 woody plant species from 10 quadrats in a Chinese subtropical forest using 454 pyrosequencing of 18S rRNA gene to describe symbiotic AMF-plant association. Our

results show the woody plant-AMF networks to be highly interconnected and nested, but in anti-modular and anti-specialized manners. The non-random pattern in the woody plant-AMF network was explained by plant and AMF phylogenies, with a tendency for a stronger phylogenetic signal by plant than AMF phylogeny. This study suggests that the phylogenetic niche conservatism in woody plants and their AMF symbionts could contribute to interdependent AMF and plant community assembly in this subtropical forest ecosystem.

Introduction

The interactions between aboveground and belowground communities influence biodiversity maintenance and ecosystem functions (Wardle *et al.* 2004; Klironomos *et al.* 2011). As key components of belowground communities, arbuscular mycorrhizal fungi (AMF) form symbiotic associations with most terrestrial plant species from which they obtain photosynthetic carbon in exchange for a variety of services, such as increased uptake of phosphorus and nitrogen (Smith & Read 2008). Plants, through selective interactions with efficient AMF partners, ultimately influence AMF communities (Bever *et al.* 2009; Kiers *et al.* 2011). In contrast, AMF can affect plant communities by increasing soil nutrient availability and/or mediating plant coexistence through the formation of common underground mycorrhizal networks (van der Heijden *et al*. 2015; Werner & Kiers 2015). Thus, by elucidating the structure and influencing factors of a plant-AMF network, we may begin to

understand mechanisms underlying species coexistence, community assembly, and ecosystem stability (Montesinos-Navarro *et al.* 2012b; Öpik & Moora 2012; van der Heijden *et al.* 2015).

Mutualistic network studies have revealed some general structures, such as high levels of nestedness and asymmetry, and have contributed substantially to our understanding of the processes behind the network patterns (e.g. Vázquez *et al.* 2009a,b; Bascompte *et al*. 2010). However, most previous mutualistic network studies have mainly focused on aboveground plant-animal mutualisms, such as plant-pollinator mutualism and plant-seed disperser mutualism (e.g. Vázquez *et al.* 2009a,b; Donatti *et al.* 2011). Recently, a growing number of studies have shown that the structure of belowground mycorrhizal mutualistic networks varies in different mycorrhizal types (Peay *et al.* 2007; Jacquemyn *et al.* 2010, 2011, 2015; Chagnon *et al.* 2012, 2015; Martos *et al.* 2012; Montesinos-Navarro *et al.* 2012a; Bahram *et al.* 2014; Toju *et al.* 2014, 2015, 2016; Beiler *et al.* 2015). For example, orchid mycorrhizal networks have been shown to be highly modular, but with non-nested or highly nested patterns in different ecosystems (Jacquemyn *et al.* 2010, 2011, 2015; Martos *et al.* 2012). Plant-ectomycorrhizal fungus networks tend to exhibit non-nested or anti-nested patterns; however, significant modularity and specialization have been detected in several of these networks in forest ecosystems (Bahram *et al.* 2014; Toju *et al.* 2014, 2015). Similarly, an ericaceous plant-fungus network has been shown to have an anti-nested, highly modular and highly specialized structure (Toju *et al.* 2016). By contrast, plant-AMF network structures have generally exhibited highly nested patterns, but with inconsistent connectance and modularity in temperate and tropical ecosystems (Chagnon *et*

al. 2012; Montesinos-Navarro *et al.* 2012a). Fortuna *et al*. (2010) proposed that the degree of connectance influences network properties: where there is high connectance, networks have a tendency to be either nested or modular, but not both; in contrast, low connectance networks that are highly nested also tend to be highly modular. In addition, most previous studies of plant-fungus association networks have been based on qualitative analysis, which ignores the frequency of association between species because the same weight is given to all links (Caruso *et al.* 2012; Öpik & Moora 2012). Furthermore, many qualitative network metrics are sensitive to variation in sampling effort (Bersier *et al.* 2002; Banašek-Richter *et al.* 2004). However, quantitative analyses reflect properties of networks more appropriately than qualitative approaches, and are less vulnerable to errors associated with variation in sampling intensity, network size, and asymmetry (Banašek-Richter *et al.* 2004; Blüthgen *et al.* 2006; Öpik & Moora 2012). Therefore, in order to draw robust conclusions, the properties of symbiotic network structures should be evaluated by both quantitative and qualitative methods.

Phylogenetic niche conservatism is a term used to describe the phenomenon that closely related plant species are more similar morphologically and functionally than distantly related plant species and tend to select similar symbiosis partners (Losos 2008). Phylogenetic niche conservatism has therefore been proposed as an important determinant of network structure (Rezende *et al.* 2007; Jacquemyn *et al.* 2011; Martos *et al.* 2012; Montesinos-Navarro *et al.* 2012b; Elias *et al.* 2013). For example, many studies have shown that phylogenetic relationships could affect aboveground mutualistic and antagonistic plant-animal networks (e.g. Rezende *et al.* 2007; Donatti *et al.* 2011; Elias *et al.* 2013). In

belowground symbiotic mycorrhizal networks, several studies have also shown that plant phylogenetic relationships affect the orchid-mycorrhizal network structures in temperate (Jacquemyn *et al.* 2011) and tropical (Martos *et al.* 2012) ecosystems and that phylogenetic relatedness of AMF influences plant-AMF network structure in a Mexican semiarid habitat (Montesinos-Navarro *et al.* 2012b). Revealing the effects of phylogenetic niche conservatism in plants and fungi on symbiotic networks would be helpful in understanding how evolution and coevolution may shape species assemblages (Maherali & Klironomos 2007; Vázquez *et al.* 2009a; Gómez *et al*. 2010). However, the effect of phylogenies of plants and AMF on symbiotic networks remains poorly understood (Montesinos-Navarro *et al.* 2012b).

Subtropical forests are widely distributed across South and East China and host a wide diversity of plant and AMF species (Zhang *et al.* 2004; Legendre *et al.* 2009). Woody plants, which constitute the dominant component of subtropical forests, contribute considerably to carbon cycling and terrestrial gross primary production (Yu *et al.* 2014). To our knowledge, most studies on plant-AMF association networks have been conducted in temperate and tropical ecosystems, using qualitative presence/absence data. Besides nestedness, plant-AMF networks were shown to differ in connectance and modularity. However, structure and factors influencing the networks of woody plants and AMF are still poorly documented.

We, therefore, used this study to characterize the AMF community of 17 woody plant species from 10 quadrats in a Chinese subtropical forest by 454 pyrosequencing 18S rRNA gene biomarkers. Connectance, nestedness, modularity, and specialization of the network

of AMF tree symbionts were quantitatively and qualitatively analyzed using the frequencies of association and co-occurrences. We addressed the following two questions: 1), what are the connected, nested, modular and specialized patterns of the woody plant-AMF network structure? 2), does the phylogenetic signal of the plants and/or AMF explain the symbiotic association network structure of the subtropical forest?

Materials and methods

Study site and root sampling

The study was conducted within a 24 ha permanent plot (29°15′6″–29°15′21″ N, 118°07′1″–118°07′24″ E; 446–714 m above sea level) of subtropical broad-leaved forest in the Gutianshan National Nature Reserve in Southeast China. The climate in the study area is dominated by subtropical monsoon with an annual mean temperature of 15.4 °C and annual mean precipitation of 1964 mm (Legendre *et al.* 2009). The 24 ha study area was subdivided into six hundred 20 m × 20 m quadrats in which all trees with diameters of ≥ 1 cm at breast height were tagged, identified, measured, and geo-referenced (Legendre *et al.* 2009). In total, 140 676 individual woody plants belonging to 159 species and 49 families are identified in the entire study area.

In this study, a total of 10 quadrats were selected from the 24 ha plot according to the criteria that the distance between any two quadrats should be more than 60 m and that these selected quadrats should contain similar dominant plant species (occurrence in > 50% quadrats) so that the same tree species could be sampled from all 10 quadrats. Seventeen

woody plant species belonging to 17 genera and 12 families were found within the 10 quadrats in the 24 ha plot. In September 2011, one individual of each tree species was randomly selected in each quadrat and root samples were collected and pooled from three directions away from the trunk. In total, 166 individual trees were sampled, except *Distylium myricoides* and *Alniphyllum fortunei* that occurred only in nine and seven quadrats, respectively; all other species could be found in each quadrat for sampling (Table 1). All roots were traced from the trunk to confirm their identity and avoid sampling roots of other tree species. The root samples were washed with sterilized distilled water and stored at −80 °C prior to DNA extraction.

Molecular analysis

Total DNA was extracted from 200 mg subsamples of the pulverized root samples, using the cetyltrimethyl-ammonium bromide method. For 454 unidirectional sequencing of the fungal 18S rRNA gene amplicons, the forward primer for emulsion PCR was constructed by combining a 454 sequencing adaptor (A adaptor; Table S1; Supporting information), a 10-base tag to discriminate the samples and primer NS31 (Simon *et al.* 1992). The reverse 454 primer was constructed by combining the annealing adaptor B for bead capturing and primer AML2 (Lee *et al.* 2008). Amplifications were performed in a Gene Amplification PCR System (East Win, Beijing, China) with initial denaturation at 94 °C for 5 min, followed by 35 cycles of 94 °C for 45 s, 52 °C for 1 min, and 72 °C for 1 min, and a final extension at 72 °C for 10 min. Amplifications were carried out in 25 μL of reaction mixture, containing 2 U *Taq*

polymerase (Takara, Japan), 2.5 µL 10× PCR buffer, 1.5 mM $MgCl_2$, 200 µM of each dNTP, 0.75 µM of each primer, and 1 µL DNA template of an approximate concentration of 100 ng μL^{-1}. PCR products of 10×, 20×, and 50× diluted templates were pooled and purified using an Axygen PCR Product Gel Purification Kit (Axygen, Inc., Union City, CA, USA). The purified PCR products were quantified using a fluorescence spectrophotometer (TBS 380, Promega, Madison, WI, USA), pooled in equimolar amounts (50 ng of DNA from each sample), and adjusted to a concentration of 10 ng μL^{-1}. The pooled products were subjected to 454 pyrosequencing on a Roche Genome Sequencer FLX Titanium (454 Life Sciences, Branford, CT, USA). The raw sequence data have been deposited in the Sequence Read Archive of the National Center for Biotechnology Information, USA (NCBI; accession no. SRP042134).

Bioinformatic analyses

Sequencing noise was removed using the 'shhh.flow' and 'trim.flow' commands in Mothur 1.31.2 (Schloss *et al.* 2009). Subsequently, sequences with no valid primer sequence or DNA tag or such containing ambiguous bases, or homopolymers (> 8 nt), or less than 300 base pairs (bp) in length, or such with an average quality score < 25 were removed using the 'trim.seqs' command in Mothur (Schloss *et al.* 2009). As the average read quality score decreases to less than 25 after 420 bp, the remaining longer sequences were trimmed to 400 bp (Schloss *et al.* 2009). Putative chimeras were discarded using the 'chimera.uchime' command in Mothur (Schloss *et al.* 2009), using both no external database and the Maarj*AM* 18S rRNA gene reference database (Öpik *et al.* 2010). In order to perform

comparisons with previous studies on AMF association networks (Montesinos-Navarro *et al.* 2012a,b), we clustered the remaining nonchimeric sequences into different operational taxonomic units (OTUs) at five nucleotide identity levels (90%, 93%, 95%, 97%, and 99%) using USEARCH v8.0 (Edgar 2013) after dereplication and discarding all singletons. The representative OTU sequences (the most abundant) were selected and blasted against the GenBank nonredundant nucleotide database; those OTUs that returned top blast results annotated as 'Glomeromycota' were assigned to AMF. Then, the AMF OTUs were confirmed by a 'blastn' search in the Maarj*AM* 18S rRNA gene database using an *E* value less than 1*e*-50 as a significant matching criterion (Öpik *et al.* 2010). The number of considered sequences per sample was rarefied to the smallest sample size by applying the 'sub.sample' command in Mothur (Schloss *et al.* 2009). Finally, we assigned the AMF OTUs to order, family, and genus levels based on the Maarj*AM* 18S rRNA gene database (Öpik *et al.* 2010). Rarefaction curves of the number of observed AMF OTUs in each tree species were calculated, using the 'specaccum' function in the VEGAN package in R 2.15 (R Core Team 2013). The representative sequences of the AMF OTUs at the 97% nucleotide identity were submitted to the International Nucleotide Sequence Database Collaboration (accession nos. LN907140–LN907224; Table S2; Supporting information).

Network analysis

In his mycorrhizal network study, Caruso *et al.* (2012) argued that the observed links cannot represent true interactions in nature when using co-occurrence matrices. We therefore used symbiotic association frequency data to establish plant-fungal links. In the present

study, for each sequence identity level of AMF (90%, 93%, 95%, 97%, and 99%), plant-AMF associations were characterized as networks. For a given pair of plant-AMF species, association frequency was defined as the number of plant root samples in which an AMF OTU was observed (Toju *et al.* 2016), which reflects the selectivity of plant AMF association. We therefore constructed a matrix in which cells with positive integers indicate the association frequency between a pair of plant and AMF species, and cells with zeros indicate no association. We compiled the association frequency matrix for the pooled 10 quadrats as a weighted (quantitative) network matrix and also transformed it into a binary (qualitative) matrix. In addition, binary matrices of the 10 subnetworks were also constructed. These matrices were used to calculate quantitative and qualitative network metrics, such as connectance, nestedness, modularity, and specialization.

Connectance (Con) is defined as the proportion of actual links in relation to the total number of possible links in the network (Jordano 1987). Quantitative connectance (Con_q) is defined as the weighted realized proportion of possible links and calculated as the quantitative linkage density (weighted diversity of interactions per species) divided by the number of species in the network (Tylianakis *et al.* 2007). The Con and Con_q values range from 0 (no interaction) to 1 (all species connected to each other). The nestedness concept describes a particular pattern of interaction in which species with fewer interactions are connected to a subset of the partners of species with more interactions (Bascompte & Jordano 2007). We used the most common metrics, namely, nested overlap and decreasing fill (NODF) and weighted NODF (WNODF), to estimate nestedness (Almeida-Neto *et al.* 2008). The NODF and WNODF values range between 0 (random structure) and 100 (perfect

nestedness). Modularity is a measure of how much the network is structured as cohesive subgroups of nodes (modules) in which the density of interactions is higher within subgroups than among subgroups (Olesen *et al.* 2007). The value of modularity ranges between 0 (random network with no modules) and 1 (maximum modularity). We detected maximized bipartite modularity (BM) and quantitative bipartite modularity (BM_q) using the QuaBiMo algorithm (Dormann & Strauss 2014) in the BIPARTITE package (version 2.05) in R (R Core Team 2013). The unipartite modularity (UM) value was also calculated using the software Netcarto (Guimerà *et al.* 2004; Guimerà & Amaral 2005a,b) for comparison with previous mycorrhizal network studies (Martos *et al.* 2012; Montesinos-Navarro *et al.* 2012a). Given the heuristic nature of the algorithm, 10 runs were conducted for each matrix and the average value of modularity is reported. We used the operational concept of 'complementary specialization' to test for interaction specialization in the community as a whole. We estimated the degree of specialization using the index H_2', based on the deviation of the realized number of interactions of a species and that expected from the total number of interactions of each species (Blüthgen *et al.* 2006). The index H_2' ranges between 0, corresponding to no specialization, and 1, corresponding to perfect specialization, for the given total number of interactions.

To test whether the association of plants and AMF was random or not, the significance of these observed network metrics was estimated by comparison with the 95% confidence intervals of metrics calculated from 1000 iterations of the following null models: vaznull, swap.web, trial swap, mgen, and an unnamed null model in Netcarto, all of which have been commonly used in previous network analyses (Ulrich & Gotelli 2007; Martos *et al.* 2012).

The vaznull null model is more 'conservative' because it keeps the connectance constant, which might represent forbidden links (constraints to species interaction) (Dormann *et al.* 2009). Similar to the vaznull null model, the swap.web null model, which keeps the marginal totals and connectance constant, is also more conservative and realistic and offers a lower probability of Type I statistical errors (Joppa *et al.* 2010). The trial swap null model is a sequential 'swap' algorithm for binary matrices, but it tries a fixed number of times and performs zero to many swaps at one step (Miklós & Podani 2004). The mgen null model returns a random web based on a probability matrix (dividing the original network by the number of interactions) and the number of interactions in the parent network (Dormann *et al.* 2009). In order to make clear that our results on network structural properties are robust, the significance of the observed Con was evaluated against the mgen null model, the observed NODF and BM against the mgen and trial swap null models, and the observed Con_q, WNODF, BM_q, and H_2' against the vaznull and swap.web null models. These analyses were conducted using the VEGAN and BIPARTITE packages in R (R Core Team 2013). The significance of the observed UM was evaluated against the null model in the software Netcarto (Guimerà *et al.* 2004; Guimerà & Amaral 2005a,b). If the observed values of the network metrics are significantly lower than the 95% confidence intervals of null model, the network structure is characterized as anti-nested, anti-modular, or anti-specialized (Hintze & Adami 2010; Bahram *et al*. 2014; Toju *et al*. 2016). Notably, the BM and BM_q values could not be computed at the 99% nucleotide identity because of a program error.

Phylogenetic signal analysis

The phylogenetic signal of the plant-AMF association was evaluated using an estimated generalized-least squares (EGLS) analysis that fits the phylogenetic variance-covariance matrix to the plant-AMF association matrix (Ives & Godfray 2006). Using this EGLS method, we calculated the independent phylogenetic signals of the plant (d_{plant}) and AMF (d_{amf}) phylogenies on the quantitative association matrix and assessed the overall strength of the signal of both phylogenies combined by using mean squared error (MSE) values. This EGLS method is based on the Ornstein-Uhlenbeck model of evolution, which incorporates stabilizing selection and drift, and measures the presence of a phylogenetic signal by the parameter *d*. This parameter *d* determines the strength of the phylogenetic signal, with *d* = 0 indicating the lack of phylogenetic correlation (i.e. a star phylogeny) and *d* = 1 corresponding to strong phylogenetic correlation (i.e. the Brownian motion assumption). The overall strength of the signal of both plant and AMF phylogenies combined was evaluated by comparing the MSE calculated for the full model (MSE_d) with the MSE derived under the assumption of no phylogenetic signal (i.e. a star phylogeny, MSE_{star}), and with the MSE derived assuming maximum phylogenetic signal (i.e. Brownian motion evolution, MSE_b). Lower MSE values indicate a better fit of the specific model to the data, and the model with the lowest MSE leaves the smallest unexplained variance. Calculations were performed with the 'pblm' function in the PICANTE package in R (R Core Team 2013) and were conducted on the maximum likelihood (ML) phylogenies of the trees and AMF (Montesinos-Navarro *et al.* 2012b). The statistical significance of the *d* value was determined by calculating 95% bootstrap confidence intervals of 100 replicates as described

in Ives & Godfray (2006). For AMF, the representative 18S rRNA gene sequence of each AMF OTU (inferred at the five nucleotide identity levels) was used to build the phylogenetic tree using *Neurospora crassa* (accession no. X04971) as an outgroup (Schüßler *et al.* 2001). For plants, the *rbcL*, *matK*, and *trnH-psbA* genes were concatenated and used to conduct a ML phylogenetic tree, using *Pinus massoniana* as an outgroup (for the accession numbers of *rbcL*, *matK*, and *trnH-psbA* genes see Table S3; Supporting information). Multiple sequence alignments were obtained with the E-INS-I algorithm in MAFFTV6 (Katoh & Standley 2013), with 781 phylogenetically informative sites for plants and 211 phylogenetically informative sites for AMF. The best ML phylogeny was inferred in RAxML 7.3.0 using the GTR + G model with default settings (Stamatakis 2006). A total of 1000 bootstrap trees were inferred to calculate branch support values for the best-scoring ML tree for the trees and AMF (Fig. S1 & S2; Supporting information). Notably, the 'pblm' function could not be run at the 99% sequence identity, because of a problem, when scaling the covariance matrices of the AMF phylogenetic trees, which may be due to high sequence identity. The phylogenetic signals of the trees and AMF were also determined for all 10 subnetworks of the 10 quadrats and five nucleotide sequence identity levels, using the binary data.

Results

AMF community sequencing and taxon identification

After controlling for sequence quality, 412 834 nonchimeric sequences were obtained and assigned to OTUs at 90%, 93%, 95%, 97% and 99% nucleotide sequence identity levels. After excluding the non-AMF OTUs, the AMF datasets were rarefied based on the smallest read

number in samples and resulted in 11, 24, 50, 85, and 435 OTUs at the 90%, 93%, 95%, 97%, and 99% sequence identity levels, respectively (Table S4; Supporting information). The AMF OTUs at 97% sequence identity were identified to eight families (see Table S2 for an example; Supporting information). The rarefaction curve of observed AMF OTUs in each plant species reached a plateau for OTUs from clustering at the 90%, 93%, 95%, and 97% nucleotide sequence identity, but not for those from the 99% identity level (Fig. S3; Supporting information).

Structural parameters of the plant-AMF networks

One network was inferred based on the pooled data of all 10 quadrats at the 90%, 93%, 95%, 97%, and 99% sequence identity levels of AMF. Overall, the plant-AMF network properties such as high connectance, high nestedness, anti-modularity and anti-specialization were obtained by comparing with multiple null models (Table 2; Fig. 1). Namely, in all five networks at five sequence identity levels, the Con and Con_q values were significantly higher than expected based on the multiple null models, except for the Con_q value inferred for the vaznull null model at the 99% identity level (Table 2). The NODF and WNODF values of the five networks were significantly higher than expected based on the multiple null models, except for the NODF value against the trial swap null model at the 90% sequence identity level and the WNODF value against the swap.web null model at the 99% sequence identity level (Table 2). The UM values of the five networks were significantly lower than expected based on the null model in Netcarto (Table 2). The BM and BM_q values

of the four networks (at the 90%, 93%, 95%, and 97% sequence identity levels) were significantly lower than expected based on the multiple null models, except for the BM value against the trial swap null model at the 95% sequence identity level (Table 2). The H_2' values of all five networks were significantly lower than the random plant-AMF associations using the swap.web and vaznull null models (Table 2).

In all subnetworks (one for each quadrat) at the five sequence identity levels, the Con and NODF values were significantly higher than expected based on the mgen null model (except for the NODF value of one subnetwork at the 99% sequence identity level), and more than half (52%) of NODF values were significantly higher than expected based on the trial swap null model (Table S5; Supporting information). In all subnetworks at the five sequence identity levels, more than half (54%) of UM values were significantly lower than expected based on the null model in Netcarto (Table S5; Supporting information). In all subnetworks at the four sequence identity levels (at the 90%, 93%, 95%, and 97% sequence identity levels), the BM values were significantly lower than expected based on the mgen null model, but some (32.5%) BM values were significantly lower than expected based on the trial swap null model (Table S5; Supporting information).

Plant-AMF network structuring

The independent phylogenetic signal of the plant phylogeny (d_{plant}) was stronger and was included within its confidence intervals which did not overlap with zero at the 90%, 93%, 95%, and 97% sequence identity levels (Table 3). The independent phylogenetic signal of the

AMF phylogeny (d_{amf}) was included within its confidence intervals at the 93%, 95%, and 97% sequence identity levels, but this parameter was weak at the 90% sequence identity level because it is closer to zero and its confidence interval overlapped with zero (Table 3). The overall strength of the phylogenetic signal for the linear model fitted to the actual data (MSE_d) was closer to that found under the assumption of the maximum phylogenetic signal from the Brownian motion evolution model (MSE_b) than that under the assumption of no phylogenetic covariance by the 'star' phylogeny (MSE_{star}) at the 90%, 93%, 95%, and 97% sequence identity levels (Table 3), indicating that the Brownian motion evolution model best fits the actual data. Taken together, the phylogenetic signals of the plant and AMF sequences had significant effects on the plant-AMF network structure, with a tendency for a stronger signal by the plants than the AMF (Table 3; Fig. 2). Furthermore, phylogenetic signal analyses of all the subnetworks at the five sequence identity levels indicated that the plant-AMF network structures were significantly affected by plant phylogeny, but not by AMF (Table S6; Supporting information).

Discussion

The network structure of woody plants and AMF was highly interconnected and highly nested, but showed anti-modular and anti-specialized properties, based on the qualitative and quantitative analyses of the data pooled from all 10 forest quadrats tested against multiple null models. The highly interconnected and nested network patterns suggest that species coexistence could be promoted by increased robustness from random extinctions

and reduced interspecific competition (Fontaine & Thébault 2010). The anti-modular pattern of the networks may indicate that AMF from one functional group seem to preferentially interact with plant species of other functional groups, and vice versa (Hintze & Adami 2010). This anti-specialized network architecture stresses that AMF are lowly host specific and highly niche overlapped as reported by previous studies (Toju *et al.* 2014, 2015; van der Heijden *et al.* 2015). Niche overlapping among species is known to increase interspecific competition for resources and thus affects species co-existence in ecosystems (Levine & HilleRisLambers 2009). Our results also suggested that the degree of sequence identity among AMF did not alter qualitatively the interpretation of network architecture although it could affect the value of network metrics (Table 2).

As for other mutualistic networks, nestedness of plant-AMF association networks has already been observed in various ecosystems (Chagnon *et al.* 2012; Montesinos-Navarro *et al.* 2012a). However, connectance and modularity patterns of plant-AMF networks have been shown to vary among different ecosystems (Chagnon *et al.* 2012; Montesinos-Navarro *et al.* 2012a; Toju *et al.* 2014, 2015). For example, a highly connected, but not modular pattern was observed in boreonemoral and temperate forest ecosystems, based on the re-analysis of the dataset of Öpik *et al.* (2009) and Davison *et al.* (2011) by Montesinos-Navarro *et al.* (2012a), whereas a low connected, but highly modular pattern was reported in a Mexican semiarid habitat (Montesinos-Navarro *et al.* 2012a). In general, with high connectance, networks have a tendency to be either nested or modular, but not

both; in contrast, with low connectance, networks that are highly nested also tend to be highly modular together (Fortuna *et al.* 2010). These findings may explain why the plant-AMF network structures in boreonemoral, temperate, and subtropical forests differ from that in a Mexican semiarid habitat.

AMF phylogenetic diversity may also explain the different plant-AMF network structures in the different ecosystems (Montesinos-Navarro *et al.* 2012a). For example, more AMF families were found in subtropical (eight families in this study), boreonemoral (three families in Öpik *et al.* 2009), and temperate (six families in Davison *et al.* 2011) forest ecosystems than the semiarid Mexican ecosystem where only Glomeraceae was found (Montesinos-Navarro *et al.* 2012a). AMF of the Glomeraceae are thought to be host plant generalists and thus to associate with most plants in ecosystems (Öpik *et al.* 2009, 2010; Davison *et al.* 2011, 2015). On the contrary, members of the Acaulosporaceae, Gigasporaceae, and Archaeosporaceae are more specific and only associate with a subset of plant species (Montesinos-Navarro *et al.* 2012a; Öpik & Moora 2012). Therefore, in more phylogenetically diverse AMF communities, there are more opportunities for specialist AMF to co-occur with generalist host plants, leading to a highly nested symbiotic network structure. Consequently, the NODF values of plant-AMF networks tend to be higher in subtropical (65.25), boreonemoral (63.74), and temperate (65.33) forest ecosystems than semiarid (28.3) ecosystems (Öpik *et al.* 2009; Davison *et al.* 2011; Montesinos-Navarro *et al.* 2012a). Toju *et al.* (2014, 2015) found that no property of the partial plant-AMF networks was significant in temperate ecosystems. This may be because of the small size of the

plant-AMF networks (13 to 29 plant species and 10 to 58 AMF OTUs) offering few opportunities for network structuring as was suggested by Olesen *et al.* (2007).

Previous studies have shown that plant-AMF associations are non-random and that AMF community structures converge as the phylogenetic distance between host plants decrease (Vandenkoornhuyse *et al*. 2003; Horn *et al*. 2014; Davison *et al*. 2015). Our study showed that the phylogenies of both plants and AMF played a vital role for non-random patterns in the plant-AMF networks of the subtropical forest sites. Similarly, the effects of phylogenies of both plants and fungi on mutualistic networks have been observed in an epiphytic orchid mycorrhizal network (Martos *et al.* 2012). By contrast, several studies have reported structure in plant-AMF networks only explained by the AMF phylogeny and not that of plants (Montesinos-Navarro *et al.* 2012b). On the contrary, an orchid mycorrhizal network has been shown to be influenced by the phylogenetic relatedness of the plants, but not the mycorrhizal fungi (Jacquemyn *et al.* 2011). Such differences in the phylogenetic signals of plants and mycorrhizal fungi may be partly explained by the taxonomic diversity of the species lineages included in the analysis, as was suggested in studies on orchid mycorrhizal networks (Martos *et al.* 2012) and plant-animal networks (Rezende *et al.* 2007; Fontaine & Thebault 2015). For example, our study included eight AMF families, yet Montesinos-Navarro *et al.* (2012b) focused on just only the most frequent fungal family (Glomeraceae). Martos *et al.* (2012) included three fungal clades (Tulasnellaceae, Ceratobasidiaceae and Sebacinales), whereas Jacquemyn *et al.* (2011) restricted to one frequent clade of fungi (Tulasnellaceae) in their orchid mycorrhizal networks. When investigating more fungal clades, trait diversity among different fungal clades may amplify

the phylogenetic signal in the network structure, because plants are allowed to associate with more different fungi, which may complement each other (Kivlin *et al.* 2011; HilleRisLambers *et al.* 2012). However, when focusing on only the most frequent fungal clade whose members share similar traits because they are more closely related, competitive exclusion could have left only functionally diverging AMF (Webb *et al.* 2002; Maherali & Klironomos 2007; HilleRisLambers *et al.* 2012).

Here, we detected a much stronger phylogenetic signal at the plant level than the AMF level, which agrees well with what has been observed in a mutualistic epiphytic orchid mycorrhizal network (Martos *et al.* 2012) and several plant-animal networks (Vázquez *et al*. 2009b; Elias *et al.* 2013; Fontaine & Thebault 2015). These differences in the intensity of the phylogenetic signals in mycorrhizal networks may suggest that closely related plant species tend to associate with the same or closely related fungi, whereas the mycorrhizal fungi may be more promiscuous in host plant choice (Maherali & Klironomos 2007; HilleRisLambers *et al.* 2012; Martos *et al.* 2012; Horn *et al.* 2014). For example, one clade of trees (*Photinia glabra* and *Myrica rubra*) tends to associated with one AMF clade belonging to Glomeraceae in this study (Fig. 2). Preferential partner selection could be a mechanism explaining this pattern, since plants seem to select more with which AMF they form a symbiosis than AMF selecting host plants (Bever *et al.* 2009; Kiers *et al.* 2011). In addition, plant root traits, correlated with mycorrhizal colonization rates, have been shown to exhibit a strong phylogenetic signal across species in subtropical forests (Kong *et al.* 2014), suggesting the notion of phylogenetic niche conservatism, namely that closely related plant species are morphologically and in terms of their traits more similar than any less related species (Losos

2008). This notion was further supported by the fact that the plant-AMF networks of the individual forest quadrats showed a plant phylogenetic signal, but not AMF signal, although weaker than that revealed from the data across all study plots (Tables 3 & S6).

Delimitation of AMF OTUs is still controversial but will have a result on the phylogenetic signals measured in studies (see analyses at different percentage of nucleotide identity thresholds) (Martos *et al.* 2012; Montesinos-Navarro *et al.* 2012b). Our results have shown that the phylogenetic signal of AMF decreases as the sequence identity of AMF decreases. At lower sequence identity, AMF OTUs belonging to different genera or families are limited, which will hide any selective association between particular plants and AMF (Montesinos-Navarro *et al.* 2012b). The weak phylogenetic signal in AMF may also be due to inappropriate phylogenetic resolution in AMF, i.e. insufficient phylotaxonomic resolution with the short fragment of the 18S rRNA gene, which has less phylogenetically informative sites at AMF phylogeny (211 sites) than that at plant phylogeny (781 sites) in this study. In addition, to overcome sampling artifacts on network analysis, such as underestimation of species and abundance, there are two approaches: i) applying quantitative metrics, and ii) increasing sampling efforts (Vázquez *et al*. 2009a). In this study, we calculated the plant-AMF network metrics using the quantitative data, which are robust against variation in sampling intensity, network size, and asymmetry (Banašek-Richter *et al*. 2004; Blüthgen *et al*. 2006). To overcome possible sampling effects future studies will have to sample more plant species in a large scale.

In summary, the present study assessed the structural pattern and potential determinants of woody plant-AMF networks in a subtropical forest based on quantitative and qualitative analyses. Plant-AMF networks were found to be highly interconnected and highly nested, although they showed anti-modular and anti-specialized properties. Plant-AMF networks were found to be influenced by the relatedness of the plants and AMF, with a much stronger phylogenetic signal from plants than AMF in selective plant-AMF association. We thus conclude that symbiotic niche conservatism of woody plants in relation to AMF may exist in subtropical forest ecosystem.

Acknowledgments

We are grateful to staff of the Gutianshan Research Station of Forest Biodiversity and Climate Change for sample collection. We thank Prof. Ming-Jian Yu from Zhejiang University for identifying plant species, Prof. Xue-Jun Ge from the South China Botanical Garden, Chinese Academy of Sciences for providing DNA sequences of the *rbcL*, *matK* and *trnH-psbA* genes of plant species, and Dr. Xiao-Ling Zhang from the Institute of Microbiology, Chinese Academy of Sciences for help in bioinformatics analysis. We are thankful to Prof. Timothy Y. James from University of Michigan for English improvement. The authors also appreciate anonymous reviewers and the handling editor for their thoughtful comments. This work was supported by the National Natural Science Foundation of China (grant nos. 31210103910 and 30930005).

References

Almeida-Neto M, Guimaraes P, Guimarães PR, Loyola RD, Ulrich W (2008) A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos*, **117**, 1227–1239.

Bahram M, Harend H, Tedersoo L (2014) Network perspectives of ectomycorrhizal associations. *Fungal Ecology*, **7**, 70–77.

Banašek-Richter C, Cattin MF, Bersier LF (2004) Sampling effects and the robustness of quantitative and qualitative food-web descriptors. *Journal of Theoretical Biology*, **226**, 23–32.

Bascompte J (2010) Structure and dynamics of ecological networks. *Science*, **329**, 765–766.

Bascompte J, Jordano P (2007) Plant-animal mutualistic networks: the architecture of biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, **38**, 567–593.

Beiler KJ, Simard SW, Durall DM (2015) Topology of tree-mycorrhizal fungus interaction networks in xeric and mesic Douglas-fir forests. *Journal of Ecology*, **103**, 616–628.

Bersier LF, Banašek-Richter C, Cattin MF (2002) Quantitative descriptors of food-web matrices. *Ecology*, **83**, 2394–2407.

Bever JD, Richardson SC, Lawrence BM, Holmes J, Watson M (2009) Preferential allocation to beneficial symbiont with spatial structure maintains mycorrhizal mutualism. *Ecology Letters*, **12**, 13–21.

Blüthgen N, Menzel F, Blüthgen N (2006) Measuring specialization in species interaction networks. *BMC Ecology*, **6**, 9.

Caruso T, Rillig MC, Garlaschelli D (2012) On the application of network theory to arbuscular mycorrhizal fungi-plant interactions: the importance of basic assumptions. *New Phytologist*, **194**, 891–894.

Chagnon PL, Bradley RL, Klironomos JN (2012) Using ecological network theory to evaluate the causes and consequences of arbuscular mycorrhizal community structure. *New Phytologist*, **194**, 307–312.

Chagnon PL, Bradley RL, Klironomos JN (2015) Trait-based partner selection drives mycorrhizal network assembly. *Oikos*, **124**, 1609–1616.

Davison J, Moora M, Öpik M *et al*. (2015) Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism. *Science*, **349**, 970–973.

Davison J, Öpik M, Daniell T, Moora M, Zobel M (2011) Arbuscular mycorrhizal fungal communities in plant roots are not random assemblages. *FEMS Microbiology Ecology*, **78**, 103–115.

Donatti CI, Guimarães PR, Galetti M, Pizo MA, Marquitti FMD, Dirzo R (2011) Analysis of a hyper-diverse seed dispersal network: modularity and underlying mechanisms. *Ecology Letters*, **14**, 773–781.

Dormann CF, Fründ J, Blüthgen N, Gruber B (2009) Indices, graphs and null models: analyzing bipartite ecological networks. *The Open Ecology Journal*, **2**, 7–24.

Dormann CF, Strauss R (2014) A method for detecting modules in quantitative bipartite networks. *Methods in Ecology and Evolution*, **5**, 90–98.

Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, **10**, 996–998.

Elias M, Fontaine C, van Veen FF (2013) Evolutionary history and ecological processes shape a local multilevel antagonistic network. *Current Biology*, **23**, 1355–1359.

Fontaine C, Thébault E (2015) Comparing the conservatism of ecological interactions in plant-pollinator and plant-herbivore networks. *Population Ecology*, **57**, 29–36.

Fortuna MA, Stouffer DB, Olesen JM *et al.* (2010) Nestedness versus modularity in ecological networks: two sides of the same coin? *Journal of Animal Ecology*, **79**, 811–817.

Gómez JM, Verdú M, Perfectti F (2010) Ecological interactions are evolutionary conserved across the entire tree of life. *Nature* **465**, 918–922.

Guimerà R, Amaral LAN (2005a) Cartography of complex networks: modules and universal roles. *Journal of Statistical Mechanics: Theory and Experiment*, **2005**, P02001.

Guimerà R, Amaral LAN (2005b) Functional cartography of complex metabolic networks. *Nature*, **433**, 895–900.

Guimerà R, Sales-Pardo M, Amaral LAN (2004) Modularity from fluctuations in random graphs and complex networks. *Physical Review E*, **70**, 025101.

HilleRisLambers J, Adler PB, Harpole WS, Levine JM, Mayfield MM (2012) Rethinking community assembly through the lens of coexistence theory. *Annual Review of Ecology, Evolution, and Systematics*, **43**, 227–248.

Hintze A, Adami C (2010) Modularity and anti-modularity in networks with arbitrary degree distribution. *Biology Direct*, **5**, 32.

Horn S, Caruso T, Verbruggen E, Rillig MC, Hempel S (2014) Arbuscular mycorrhizal fungal communities are phylogenetically clustered at small scales. *The ISME Journal*, **8**, 2231–2242.

Ives AR, Godfray HCJ (2006) Phylogenetic analysis of trophic associations. *The American Naturalist*, **168**, E1–E14.

Jacquemyn H, Brys R, Waud M, Busschaert P, Lievens B (2015) Mycorrhizal networks and coexistence in species-rich orchid communities. *New Phytologist*, **206**, 1127–1134.

Jacquemyn H, Honnay O, Cammue B, Brys R, Lievens B (2010) Low specificity and nested subset structure characterize mycorrhizal associations in five closely related species of the genus *Orchis*. *Molecular Ecology*, **19**, 4086–4095.

Jacquemyn H, Merckx V, Brys R *et al.* (2011) Analysis of network architecture reveals phylogenetic constraints on mycorrhizal specificity in the genus *Orchis* (Orchidaceae). *New Phytologist*, **192**, 518–528.

Joppa LN, Montoya JM, Solé R, Sanderson J, Pimm SL (2010) On nestedness in ecological networks. *Evolutionary Ecology Research*, **12**, 35–46.

Jordano P (1987) Patterns of mutualistic interactions in pollination and seed dispersal: connectance, dependence asymmetries, and coevolution. *The American Naturalist*, **129**, 657–677.

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, **30**, 772–780.

Kiers ET, Duhamel M, Beesetty Y *et al.* (2011) Reciprocal rewards stabilize cooperation in the mycorrhizal symbiosis. *Science*, **333**, 880–882.

Kivlin SN, Hawkes CV, Treseder KK (2011) Global diversity and distribution of arbuscular mycorrhizal fungi. *Soil Biology & Biochemistry*, **43**, 2294–2303.

Klironomos J, Zobel M, Tibbett M *et al*. (2011) Forces that structure plant communities: quantifying the importance of the mycorrhizal symbiosis. *New Phytologist*, **189**, 366–370.

Kong D, Ma C, Zhang Q *et al.* (2014) Leading dimensions in absorptive root trait variation across 96 subtropical forest species. *New Phytologist*, **203**, 863–872.

Lee J, Lee S, Young JPW (2008) Improved PCR primers for the detection and identification of arbuscular mycorrhizal fungi. *FEMS Microbiology Ecology*, **65**, 339–349.

Legendre P, Mi X, Ren H *et al.* (2009) Partitioning beta diversity in a subtropical broad-leaved forest of China. *Ecology*, **90**, 663–674.

Levine JM, HilleRisLambers J. (2009) The importance of niches for the maintenance of species diversity. *Nature* **461**, 254–257.

Losos JB (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters*, **11**, 995–1003.

Maherali H, Klironomos N (2007) Influence of phylogeny on fungal community assembly and ecosystem functioning. *Science*, **316**, 1746–1748.

Martos F, Munoz F, Pailler T, Kottke I, Gonneau C, Selosse MA (2012) The role of epiphytism in architecture and evolutionary constraint within mycorrhizal networks of tropical orchids. *Molecular Ecology*, **21**, 5098–5109.

Miklós I, Podani J (2004) Randomization of presence-absence matrices: comments and new algorithms. *Ecology*, **85**, 86–92.

Montesinos-Navarro A, Segarra-Moragues JG, Valiente-Banuet A, Verdú M (2012a) The network structure of plant-arbuscular mycorrhizal fungi. *New Phytologist*, **194**, 536–547.

Montesinos-Navarro A, Segarra-Moragues JG, Valiente-Banuet A, Verdú M (2012b) Plant facilitation occurs between species differing in their associated arbuscular mycorrhizal fungi. *New Phytologist*, **196**, 835–844.

Olesen JM, Bascompte J, Dupont YL, Jordano P (2007) The modularity of pollination networks. *Proceedings of the National Academy of Sciences, USA*, **104**, 19891–19896.

Öpik M, Metsis M, Daniell T, Zobel M, Moora M (2009) Large-scale parallel 454 sequencing reveals host ecological group specificity of arbuscular mycorrhizal fungi in a boreonemoral forest. *New Phytologist*, **184**, 424–437.

Öpik M, Moora M (2012) Missing nodes and links in mycorrhizal networks. *New Phytologist*, **194**, 304–306.

Öpik M, Vanatoa A, Vanatoa E *et al.* (2010) The online database Maarj*AM* reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). *New Phytologist*, **188**, 223–241.

Peay KG, Burns TD, Kennedy PG, Bergemann SE, Garbelotto M (2007) A strong species-area relationship for eukaryotic soil microbes: island size matters for ectomycorrhizal fungi. *Ecology Letters*, **10**, 470–480.

R Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rezende EL, Lavabre JE, Guimarães PR, Jordano P, Bascompte J (2007) Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature*, **448**, 925–928.

Schloss PD, Westcott SL, Ryabin T *et al.* (2009) Introducing mother: open source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, **75**, 7537–7541.

Schüßler A, Schwarzott D, Walker C (2001) A new fungal phylum, the Glomeromycota: phylogeny and evolution. *Mycological Research*, **105**, 1413–1421.

Simon L, Lalonde M, Bruns TD (1992) Specific amplification of 18S fungal ribosomal genes from vesicular-arbuscular endomycorrhizal fungi colonizing roots. *Applied and Environmental Microbiology*, **58**, 291–295.

Smith SE, Read DJ (2008) *Mycorrhizal Symbiosis*. 3rd edn. Academic Press, London, UK.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.

Toju H, Guimarães PR, Olesen JM, Thompson JN (2014) Assembly of complex plant-fungus networks. *Nature Communications*, **5**, 5273.

Toju H, Guimarães PR, Olesen JM, Thompson JN (2015) Below-ground plant-fungus network topology is not congruent with above-ground plant-animal network topology. *Science Advances*, **1**, e1500291.

Toju H, Tanabe AS, Ishii HS (2016) Ericaceous plant-fungus network in a harsh alpine-subalpine environment. *Molecular Ecology*, **25**, 3242–3257.

Tylianakis JM, Tscharntke T, Lewis OT (2007) Habitat modification alters the structure of tropical host-parasitoid food webs. *Nature*, **445**, 202–205.

Ulrich W, Gotelli NJ (2007) Null model analysis of species nestedness patterns. *Ecology*, **88**, 1824–1831.

van der Heijden MAG, Martin FM, Selosse MA, Sanders IR (2015) Mycorrhizal ecology and evolution: the past, the present, and the future. *New Phytologist*, **205**, 1406–1423.

Vandenkoornhuyse P, Ridgway KP, Watson IJ, Fitter AH, Young JPW (2003) Co-existing grass species have distinctive arbuscular mycorrhizal communities. *Molecular Ecology*, **12**, 3085–3095.

Vázquez DP, Blüthgen N, Cagnolo L, Chacoff N (2009a) Uniting pattern and process in plant-animal mutualistic networks: a review. *Annals of Botany*, **103**, 1445–1457.

Vázquez DP, Chacoff NP, Cagnolo L (2009b) Evaluating multiple determinants of the structure of plant-animal mutualistic networks. *Ecology*, **90**, 2039–2046.

Wardle DA, Bardgett RD, Klironomos JN, Setälä H, van der Putten WH, Wall DH (2004) Ecological linkages between aboveground and belowground biota. *Science*, **304**, 1629–1633.

Webb CO, Ackerly DD, Mcpeek MA, Donoghue MJ (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, **33**, 475e505.

Werner GDA, Kiers ET (2015) Partner selection in the mycorrhizal mutualism. *New Phytologist*, **205**, 1437–1442.

Yu GY, Chen Z, Piao SL *et al.* (2014) High carbon dioxide uptake by subtropical forest ecosystems in the East Asian monsoon region. *Proceedings of the National Academy of Sciences*, *USA*, **111**, 4910–4915.

Zhang Y, Guo LD, Liu RJ (2004) Survey of arbuscular mycorrhizal fungi in deforested and natural forest land in the subtropical region of Dujiangyan, southwest China. *Plant and Soil*, **261**, 257–263.

Author contributions

L.C. and L.G. designed the experiment, L.C., Y.Z., and C.G. collected the samples and performed the laboratory work, L.C., T.W., X.M. and K.M. analyzed the data, and L.C. and L.G. wrote the manuscript.

Data accessibility

The raw sequence data were deposited in the Sequence Read Archive of the National Center for Biotechnology Information, USA under accession no. SRP042134. The representative sequences of the arbuscular mycorrhizal fungal OTUs at the 97% sequence identity level have been submitted to the International Nucleotide Sequence Database Collaboration under the accession nos. LN907140–LN907224. The taxonomic affiliations of the arbuscular mycorrhizal fungal OTUs are listed in the supplementary Table S2.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Best-scoring maximum likelihood tree of the study plants based on the nucleotide sequences of their *rbcL*, *matK*, and *trnH-psbA* genes (see Materials and methods), using *Pinus massoniana* as an outgroup. Bootstrap values were calculated on the basis of 1000 data resampling (> 50% of the values are shown).

Fig. S2 Best-scoring maximum likelihood tree of arbuscular mycorrhizal fungal (AMF) operational taxonomic units (OTUs) based on 18S rRNA gene sequences (see Materials and methods), using *Neurospora crassa* as an outgroup. AMF family groupings are color-coded. Bootstrap values were calculated on the basis of 1000 data resampling (> 50% of the values are shown).

Fig. S3 Rarefaction curves of observed arbuscular mycorrhizal fungal (AMF) operational taxonomic units (OTUs) in the roots of each plant species based on sampling efforts at five nucleotide identity levels.

Table S1 The 454 sample ID, plant species, and DNA tag of each sample in each quadrat.

Table S2 Molecular identification of arbuscular mycorrhizal fungi at the 97% sequence identity level.

Table S3 GenBank accession numbers of the *rbcL*, *matK*, and *trnH-psbA* genes used in constructing the phylogenetic multigene tree for the 17 studied plant species and one outgroup species.

Table S4 The results of the bioinformatics analysis.

Table S5 Summary of the topological features of the plant-arbuscular mycorrhizal fungal subnetworks at five sequence identity levels in each quadrat.

Table S6 Phylogenetic signals of the plant-arbuscular mycorrhizal fungal associations at five sequence identity levels using presence-absence matrices in each quadrat.

Figure legends

Fig. 1 The bipartite interaction network formed by plants (lower boxes) and arbuscular mycorrhizal fungal (AMF) operational taxonomic units (OTUs, upper bars) at the 97% sequence identity level. Plant species and AMF OTUs are arranged from left to right within the network in descending order according to their total number of links. Within the network, the lower boxes are color-coded according to plant identity.

Fig. 2 Plant-arbuscular mycorrhizal fungal (AMF) association matrix combined with the phylogenetic topologies, using the 97% nucleotide identity level of AMF operational taxonomic units (OTUs). Different shades of red denote the association frequency between different plants and AMF OTUs, and white denotes an absence of association. AMF family groupings are color-coded.

Table 1 Number of sampled plant individuals and observed arbuscular mycorrhizal fungal (AMF) operational taxonomic units (OTUs) for each plant species in this study

Plant order	Plant family	Plant species	No. of root samples	No. of AMF OTUs at five nucleotide identity (%)				
				90	93	95	97	99
Ericales	Styracaceae	*Alniphyllum fortunei*	7	6	10	25	31	92
	Pentaphylacaceae	*Adinandra millettii*	10	8	18	35	43	128
		Ternstroemia gymnanthera	10	7	14	29	35	113
	Theaceae	*Camellia fraterna*	10	6	10	23	30	107
		Schima superba	10	9	17	31	41	106
Laurales	Calycanthaceae	*Chimonanthus salicifolius*	10	9	13	28	44	116
	Lauraceae	*Cinnamomum subavenium*	10	7	10	21	30	107
		Machilus thunbergii	10	8	13	27	38	95
		Neolitsea aurata	10	6	10	23	34	115
Saxifragales	Iteaceae	*Itea oblonga*	10	8	13	29	40	108
	Daphniphylaceae	*Daphniphyllum oldhamii*	10	7	14	27	43	125

	Hamamelidaceae	*Distylium myricoides*	9	5	9	23	33	89
		Loropetalum chinense	10	9	19	37	56	142
Sapindales	Anacardiaceae	*Toxicodendron succedaneum*	10	6	9	26	33	104
Rosales	Rosaceae	*Photinia glabra*	10	6	13	28	34	109
Fagales	Myricaceae	*Myrica rubra*	10	9	19	35	38	97
Myrtales	Myrtaceae	*Syzygium buxifolium*	10	7	11	27	39	102

Table 2 Summary of the topological features of the plant-AMF networks at five nucleotide identity levels

				Binary matrix (based on presence-absence matrix)									
Identity (%)	N_p	N_f	Network size	Con	Mgen null model	NODF	Mgen null model	Trial swap null model	UM	Null model in Netcarto	BM	Mgen null model	Trial swap null model
90	17	11	187	0.6578	(0.4520:0.4545)+	78.10	(53.21:53.77)+	(78.68:78.71)−	0.0909	(0.1681:0.1726)−	0.1259	(0.2568:0.2657)−	(0.1279:0.1285)−
93	17	24	408	0.5441	(0.3829:0.3845)+	81.22	(49.90:50.25)+	(81.07:81.08)+	0.1322	(0.1727:0.1761)−	0.1436	(0.2413:0.2474)−	(0.1437:0.1442)−
95	17	50	850	0.5576	(0.3965:0.3976)+	74.44	(49.78:49.98)+	(74.07:74.09)+	0.0989	(0.1540:0.1559)−	0.1093	(0.2116:0.2156)−	(0.1068:0.1074)+
97	17	85	1445	0.4443	(0.3300:0.3307)+	65.25	(43.16:43.30)+	(65.19:65.21)+	0.1354	(0.1714:0.1733)−	0.1230	(0.2076:0.2135)−	(0.1269:0.1277)−

99	17	435	7395	0.2508	(0.2088:0.2090)+	36.65	(27.39:27.43)+	(36.63:36.63)+	0.2216	(0.2347:0.2358)−	NA	NA	NA

To be continued

	Weighted matrix (based on frequency of association matrix)											
Identity (%)	Con_q	Swap.web null model	Vaznull null model	WNODF	Swap.web null model	Vaznull null model	BM_q	Swap.web null model	Vaznull null model	H_2'	Swap.web null model	Vaznull null model
90	0.3805	(0.3485:0.3497)+	(0.3697:0.3703)+	60.69	(47.04:47.44)+	(51.00:51.48)+	0.0959	(0.1334:0.1401)−	(0.1002:0.1041)−	0.0568	(0.1014:0.1034)−	(0.0633:0.0644)−
93	0.3095	(0.2865:0.2872)+	(0.3021:0.3025)+	67.98	(57.18:57.40)+	(58.55:58.80)+	0.0971	(0.1212:0.1260)−	(0.0974:0.1003)−	0.0497	(0.0784:0.0796)−	(0.0559:0.0566)−
95	0.2686	(0.2553:0.2557)+	(0.2667:0.2670)+	64.15	(62.71:62.83)+	(63.33:63.45)+	0.0907	(0.1090:0.1118)−	(0.0978:0.0996)−	0.0563	(0.0725:0.0730)−	(0.0639:0.0644)−
97	0.2052	(0.1959:0.1961)+	(0.2048:0.2050)+	58.43	(55.78:55.85)+	(55.49:55.59)+	0.1021	(0.1091:0.1130)−	(0.1035:0.1068)−	0.0717	(0.0852:0.0857)−	(0.0801:0.0806)−
99	0.1042	(0.1019:0.1020)+	(0.1053:0.1054)−	34.52	(34.56:34.58)−	(33.03:33.07)+	NA	NA	NA	0.0929	(0.0995:0.0998)−	(0.1116:0.1122)−

Identity, percentage of nucleotide identity of arbuscular mycorrhizal fungi (AMF); N_p, number of plant species; N_f, number of AMF; Network size, $N_p \times N_f$; Con, connectance; Con_q, quantitative connectance; NODF, nested overlap and decreasing fill for the overall matrix; WNODF, weighted NODF; UM, unipartite modularity; BM, bipartite modularity; BM_q, quantitative bipartite modularity; H_2', specialization. For the Con, Con_q, NODF, WNODF, UM, BM, BM_q, and H_2' parameters, the significance was tested against the null models. The numbers in parentheses refer to the 95% confidence intervals that were generated from 1000 randomization procedures; +, the observed value is significantly higher than the index of the randomized data set; −, the observed value is significantly lower than the index of randomized data set; NA, value is not available because the BM and BMq values were not computed at the 99% sequence identity level because of a program error in the running process.

Table 3 Phylogenetic signal of the plant-arbusclar mycorrhizal fungal (AMF) associations based on frequency of occurrence across ten forest quadrats

Identity (%)	MSE_d	MSE_{star}	MSE_b	d_{plant} (95% CI)	d_{amf} (95% CI)
90	0.0158	0.0442	0.0232	2.762 (1.531:4.440)	0.098 (0:0.321)
93	0.0082	0.0240	0.0112	3.078 (1.992:4.530)	0.156 (0.002:0.306)
95	0.0046	0.0107	0.0053	2.314 (1.770:2.960)	0.230 (0.112:0.374)
97	0.0036	0.0075	0.0042	1.979 (1.594:2.470)	0.183 (0.078:0.258)

MSE_d, mean squared error (MSE) calculated for the full model; MSE_{star}, MSE calculated for a 'star' phylogeny; MSE_b, MSE calculated for a Brownian evolution model; *d*, strength of the phylogenetic signals in the plants (d_{plant}) and AMF (d_{amf}).