Appendix A: Detailed description of arbuscular mycorrhizal fungi processing

For arbuscular mycorrhizal fungi operational taxonomic units (AMF OTUs) we extracted soil microbial genomic DNA from 1 g of each composite sample using a MoBioPowerSoil DNA Isolation Kit (MoBio Laboratories Inc. Carlsbad, CA, USA), according to the manufacturer's instructions. We generated AM fungal 18S rDNA amplicon libraries using fusion primers designed with the pyrosequencing primer B, a barcode and the AM fungal specific primer AM1a or AM1b (Morris et al. 2013) as reverse primers, and the pyrosequencing primer A as well as the NS31 primer as forward primers. We used a set of 10bp MID-barcodes provided by Roche (Roche Applied Science). Amplicon libraries were produced from a pool of two dilution levels and three PCR replications. We performed the first PCR reactions in a total volume of 25 µl reaction mix containing 0.5 µl of DNA template (5-20 ng/µl), 12.5 µl GoTaq Green Mastermix 2x (Promega, USA), 1 µl of each primer (25 µM) on a Eppendorf Mastercycler DNA Engine Thermal Cycler PCR (Eppendorf, USA) with the following PCR conditions: 98°C for 30 s, 5 cycles of 94°C for 30 s, 60°C (-1°C/cycle, 4 cycles) for 30 s, 72°C for 1 min, and 25 cycles of 94°C for 30 s, 55°C for 30 s, 72°C for 1 min, and for extension 72°C for 5 min. For the semi-nested PCR, we performed two separate amplifications using 1 µl of the diluted amplified product of PCR I (1:10), 25 µl GoTaq Green Mastermix 2x (Promega), and 1 µl of each primer (25 µM), the 50 µl-reactions were run under following conditions: 98°C for 30 s, followed by 30 cycles of 94°C for 30 s, 63°C for 30 s, 72°C for 1 min., and 72°C for 5 min. Each sample in both PCR amplification steps was amplified in triplicate accompanied by a negative control. The semi-nested PCR amplified products were checked by a 1.5% agarose gel and amplified products were pooled per sample. The pooled amplicons (30 µl, each) were purified with the QIAquick Gel Extraction Kit (Qiagen GmbH, Hilden, Germany) following the manufacturer's protocol. The purified products were quantified by fluorescence spectrophotometer (Cary Eclipse, Agilent Technologies, Waldbronn, Germany) using Quant-iT^{TM} PicoGreen ® dsDNA Assay Kit (Life Technologies GmbH, Darmstadt, Germany). We

subjected an equimolar mix of the 50 amplicon libraries to unidirectional pyrosequencing from the NS31 end of the amplicons, 50 samples per lane on plate separated by four lane gasket, using a 454 titanium amplicon sequencing kit and a Genome Sequencer FLX 454 System (454 Life Sciences/Roche Applied Biosystems, Mannheim, Germany) at the Department of Soil Ecology, Helmholtz Centre for Environmental Research (UFZ, Halle/ Germany).

For the bioinformatic analysis, we performed multiple levels of sequence processing and quality filtering. We extracted the 454 fungal 18S rDNA sequences based on 100% barcode similarity. Simultaneously, we removed sequence reads with an average quality score of < 20, ambiguous bases, and homo-polymers of >8 bases and read length of < 300 bp after barcodes and primers removal using MOTHUR (Schloss et al 2009). The sequences were normalized to the smallest read number and potential chimeric sequences were checked and removed by using UCHIME (Edgar et al. 2011) algorithm as implemented in MOTHUR. Quality-filtered sequences were clustered into operational taxonomical units (OTUs) based on a CD-HIT algorithm CD-HIT-EST (Huang et al 2010) with a sequence similarity threshold of 97%. Representative sequences for each OTU were taxonomically assignment to NCBI based fungal reference data set using the default parameters in MOTHUR and all non-Glomeromycotan OTUs were removed from the data set (~11% of sequences). To avoid over clustering, representative sequences of the Glomeromycotan OTU's were aligned using MUSCLE as implemented in Seaview version 4 (Gouy et al 2010). The alignment was used to construct a phylogenetic tree using the Quick NJ tree and the OTUs were further re-clustered using the GroupCluster algorithm at a similarity threshold of 0.03 in TOPALI v2.5 (Milne et al. 2004). The taxonomic assignment of representative sequences of these Glomeromycotan OTUs were refined using the MaarjaM AMF reference database (Öpik et al 2010) at a similarity threshold of 90% using the GAST algorithm (Huse et al., 2008).

References:

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Applied and Environmental Microbiology* 75(23): 7537-7541.

Morris EK, Buscot F, Herbst C, Meiners T, Obermaier E, Waschke NW, Wubet T, Rillig MC. 2013. Land use and host neighbor identity effects on arbuscular mycorrhizal fungal community composition in focal plant rhizosphere. *Biodiversity and Conservation* 22(10): 2193-2205.

Huse SM, Dethlefsen L, Huber JA, Welch DM, Relman D, Sogin M. 2008. Exploring Microbial Diversity and Taxonomy Using SSU rRNA Hypervariable Tag Sequencing. *Plos Genetics* 4(11): e1000255.

Opik M, Vanatoa A, Vanatoa E, Moora M, Davison J, Kalwij JM, Reier U, Zobel M. 2010. The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). *New Phytol* 188(1): 223-241.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27(16): 2194-2200.

Huang Y, Niu B, Gao Y, Fu L, Li W. 2010. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* 26(5): 680-682.

Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* 27(2): 221-224.

Milne I, Lindner D, Bayer M, Husmeier D, McGuire G, Marshall DF, Wright F. 2009. TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics* 25(1): 126-127.