

Modelling the environmental and soil factors that shape the niches of two common arbuscular mycorrhizal fungal families

Stavros D. Veresoglou · Tancredi Caruso · Matthias C. Rillig

Received: 22 May 2012 / Accepted: 11 November 2012 / Published online: 24 November 2012
© Springer Science+Business Media Dordrecht 2012

Abstract

Aims Arbuscular mycorrhizal (AM) fungi, a group of obligate symbionts of terrestrial plants, have a global distribution range. Yet, we lack concrete synthetic and empirical evidence that could reveal whether distinct ecological niches are distributed across Glomeromycota through determining linkages between environmental factors and the distribution of these taxa.

Methods We have modelled the probability of occurrence of Gigasporaceae and Acaulosporaceae as a function of candidate environmental factors. These families are among the most common but non-ubiquitous taxa in AM-driven ecosystems. We have constructed our database using studies with a global scope and carried out our analysis through a logistic regression approach.

Results The probability of occurrence of Acaulosporaceae increased in acidic environments and soils with high bulk density. By contrast, a key factor that affected probability of occurrence of Gigasporaceae was precipitation.

Conclusions Through the analysis of an unprecedentedly large amount of data we could infer that niche processes mediate occurrence of a group of fungi at scales broader than the local scale of the individual studies gathered in the analysed dataset. Knowledge of well-supported niche features could enhance discovery of new taxa of AM fungi, and would facilitate development of study designs with greater ecological realism.

Keywords Akaike information criterion/ environmental filter · Glomeromycota · Logistic regression · Model MaarjAM database · Model averaging

Introduction

The arbuscular mycorrhizal (AM) symbiosis is an integral part of terrestrial ecosystems (Rillig 2004) due to its ubiquity (Wang and Qiu 2006) and its potential to considerably influence plant community composition (e.g. Klironomos et al. 2011). The symbiosis constitutes an intimate association of Glomeromycotan fungi and roots of terrestrial plants (Schüßler et al. 2001). The most studied families within the Glomeromycota are the Glomeraceae, Gigasporaceae and Acaulosporaceae (Smith and Read 2008). Considerable morphological differences exist amongst isolates from these three families. These include the size of spores (spores of Gigasporaceae isolates are considerably larger than those of Glomeraceae isolates—Schenck and Pérez 1990), root

Responsible Editor: Duncan D. Cameron.

Electronic supplementary material The online version of this article (doi:10.1007/s11104-012-1531-x) contains supplementary material, which is available to authorized users.

S. D. Veresoglou (✉) · T. Caruso · M. C. Rillig
Institut für Biologie, Freie Universität Berlin,
Altensteinstr. 6,
14195 Berlin, Germany
e-mail: sveresoglou@zedat.fu-berlin.de

colonisation structures (no vesicles occur following root colonisation of Gigasporaceae isolates—Smith and Read 2008), propagation (Gigasporaceae may not propagate from colonized root sections—Klironomos and Hart 2002) and colonisation strategy (Glomeraceae preferentially colonise the root whereas Gigasporaceae grow profusely in the soil—Hart and Reader 2002; Maherali and Klironomos 2012). Powell et al. (2009) provided evidence that differences potentially exist with respect to symbiotic function across different AM fungal families, as well; Gigasporaceae could be more efficient in phosphorus (P) assimilation whereas Glomeraceae could grant better protection from pathogens.

AM fungi are obligate symbionts and thus rely on their host plant; therefore the impact of any environmental factor may be partitioned into a direct effect of the environment on the AM fungi and an indirect effect through modification of the plant community. Here we do not attempt to disentangle these direct and indirect effects, and just like in previous studies (e.g. Dumbrell et al. 2010a) only consider the overall effect. At an AM fungal community level, large differences with regards to probability of occurrence of the common AM fungal families have been recorded. Glomeraceae occur ubiquitously in nature (although ecological differences are recorded with respect to isolates within Glomeraceae, e.g. some *Glomus* group B isolates and *Glomus mosseae* predominantly occur in disturbed environments; Sýkorová et al. 2006). By contrast Gigasporaceae and Acaulosporaceae isolates are relatively less common in AM fungal communities (e.g. Öpik et al. 2010).

In principle, the ecological differences mentioned above can provide the basis for niche partitioning processes under the assumption that AM fungal populations already compete more strongly with members of the same family (Maherali and Klironomos 2007). More generally, beyond competition and in terms of the “Grinnellian niche” (Soberon 2007), it is logical to assume that these ecological differences imply differences in terms of environmental requirements at a family level. The available literature suggests that this deduction applies to Gigasporaceae and Acaulosporaceae. For example, using a nitrogen (N)-deposition gradient, Egerton-Warburton and Allen (2000) demonstrated that large-spored AM fungal species predominantly occurred at the sites that had been subjected to low N-deposition; sites receiving high N-deposition were dominated by

Glomeraceae isolates. In another study, N-fertilization over three years led to a decline of abundance of Gigasporaceae isolates at three mesic sites (Johnson et al. 2003). Gigasporaceae appear to prefer sandy rather than loamy soils (Lekberg et al. 2007) whereas both Gigasporaceae and Acaulosporaceae isolates predominantly occur at mesic rather than semiarid ecosystems in the USA (Egerton-Warburton et al. 2007). Yet, the most pronounced environmental factor that determines structure of AM fungal communities appears to be pH (e.g. Fitzsimons et al. 2008; Dumbrell et al. 2010a). Spore germination assays have shown that Acaulosporaceae and Gigasporaceae are more tolerant of acidic soil environments than the majority of Glomeraceae (e.g. Bartolomeesteban and Schenck 1994; Clark 1997). Using a long-term liming experiment Guo et al. (2012) were able to show that while most phylotypes belonging to *Glomus* group A increased in abundance following liming, *Glomus* group B and Gigasporaceae/Acaulosporaceae phylotypes declined.

While the phylum Glomeromycota is monophyletic (Schüßler et al. 2001), subsequent analyses revealed that polyphyly may exist inside the family of Glomeraceae (da Silva et al. 2006) and the genus *Scutellospora* in Gigasporaceae (Oehl et al. 2008). For Acaulosporaceae the possibility of a polyphyletic origin has not yet been sufficiently addressed. Clearly, the morphological adaptations that are present today should have maximized benefits for the AM isolates; the fact that there are several easily distinguishable morphological adaptations in AM fungi provides circumstantial support to the notion of “niche” based differentiation within the Glomeromycota. Therefore, despite the niche concept in the literature being typically addressed in studies with resolution down to species level, there has been evidence that phylogenetic trait conservatism in Glomeromycota and consequently niche differentiation may be detected at a Glomeromycotan family level (Lekberg et al. 2007; Maherali and Klironomos 2007; Powell et al. 2009), at least for the component of the niche which relates to environmental requirements (e.g. Chase and Leibold 2003).

In the present study we hypothesised that each of the three AM fungal families mentioned above preferentially occurs in certain environments owing to physiological and ecological constraints dictated by the general Grinnellian niche (i.e. environmental requirements) of

each family. As stated earlier, contrary to Glomeraceae isolates that possess a broad distribution range, Gigasporaceae and Acaulosporaceae isolates are only occasionally found in AM fungal community analyses. We thus engaged in a modelling exercise to test and define the niche structure for just the two families of Gigasporaceae and Acaulosporaceae. In addition to the three potential niche shaping parameters discussed so far (pH, N additions, soil texture) we considered a range of environmental and climatic variables to convincingly account for a potential Grinnelian niche of AM fungi. The three possible scenarios for the model outcome are as follows: i) We are unable to detect any factors that influence the occurrence of either Gigasporaceae and Acaulosporaceae; this result would lead to rejection of the original hypothesis. ii) The factors that affect occurrence of Gigasporaceae and Acaulosporaceae are the same; unless we observe large differences in the relative importance of the factors or the estimates for the parameters this would effectively represent a general niche for this subset of the Glomeromycota, and the hypothesis would have to be rejected. iii) The factors shaping the niches of Gigasporaceae and Acaulosporaceae differ—this result would support the original hypothesis.

Materials and methods

Sources of data

On the 20th of June 2011 the MaarjAM database (Öpik et al. 2010) was accessed. At the time when MaarjAM was accessed it contained 105 studies. Studies were screened so that they met the following inclusion criteria:

1. They presented AM fungal community analysis results that had been carried out under field conditions. Only molecular identification trials targeting root fragments were considered.
2. The community analysis used a restriction fragment length polymorphism (RFLP)—cloning approach so that it was possible to evaluate the sequencing effort in the study.
3. The number of clones sequenced was available or could be estimated from information in the article
4. Occasionally pH, altitude and/or climatic data were missing in the study. In these instances we have attempted to retrieve the data from i) former studies that were carried out at the same location; ii) in one case (Burke 2008) the pH that was used was the mean pH predicted for *Alliaria petiolata* occurrence from Grime et al. (1990); iii) for climatic and altitudinal data a range of different websites were used and are stated as part of the supplementary information. Studies for which neither pH, altitudinal or climatic data could be retrieved were excluded from the dataset. Also, when single plants were addressed, for which pH data were contained in Grime et al. (1990) but with a range of greater than 2 pH units, we excluded these from the analysis (unless pH data could be obtained from elsewhere).
5. Studies were excluded if the reported pH range of the study area was greater than 2 pH units (e.g. Dumbrell et al. 2010b) to standardise the extent of uncertainty in the pH estimates across studies. That way we have ensured that the possible error for pH estimates was below 1.

In total 35 studies were included in the database comprising 98 trials and 70 “study sites”. We used the term study site to describe unique combinations of trial locations and management treatments. A study typically contained more than one study site; however, there were also instances where more than one study pertains to the same study site (e.g. studies in Helgason et al. 1999 and Helgason et al. 2007 were both carried out at Pretty Wood, in the absence of fertilization). The global distribution of the 34 studies is presented in Fig. 1.

As stated earlier, only data on AM community occurrence that originated from molecular analysis of plant root AM communities were used. We have intentionally excluded spore-based studies since we were specifically interested in root colonizing AM fungal assemblages. We justify this on the basis of the following three arguments: (i) The correspondence between assemblages assessed via spores and root colonising structures can potentially be low (Hempel et al. 2007); (ii) Moreover, the outcome of exposure to environmental factors may be different for different components of AM fungal growth (e.g. high pH may reduce vesicle colonisation but increase the number of arbuscules—Zhu et al. 2007; simultaneous consideration of more than one measures of AM fungal abundance may consequently be problematic); (iii) consideration of the spore pool in the soil may introduce biases in community analyses as many of them are



Fig. 1 World map showing the locations of the studies included in the database. Numbers next to locations indicate the number of independent studies that have been carried out in the specific

country/state. Pie charts next to the numbers present the relative proportion of studies that have been carried out in woody (*black*), agricultural/urban (*white*) or grassland (*grey*) ecosystem types

ecologically inactive (Rosendahl 2008). Despite a re-evaluation of the classification of AM fungi (Schüßler and Walker 2010) in our analysis we have retained the older AM fungal nomenclature for reasons of compatibility with the reports assessed.—Nevertheless, because the families of Acaulosporaceae and Gigasporaceae have not been reclassified in Schüßler and Walker (2010) results drawn for the specific families still hold for the new taxonomy.

Data on Gigasporaceae and Acaulosporaceae occurrence (in the form of binary variables, 0: absent; 1: present) were collected for the following characteristics to be used as explanatory variables in the subsequent analysis.

1. Study site characterisation variables

The following categorical variables were included to describe the degree of management of the study site

- *Fertilisation*: The categorical variable received the value “1” in the case of anthropogenic nutrient inputs. This included both arable land and fertilised treatments in factorial fertilisation experiments.
- *Major anthropogenic disturbance*: The categorical variable received the value “1” in the cases

where sites were located inside villages/cities and arable fields.

- *Woody vegetation*: The categorical variable received the value of “1” when either the sampling had been carried out within a wood/forest or when the root samples originated from isolated trees
- *Heavy metal contamination*: The categorical variable received the value of “1” only in cases where the soil was reported to be polluted with heavy metals.

2. Study site environmental variables and conditions at harvest

The following continuous and categorical variables were included

- *pH*: (continuous) The mean soil pH value reported for the study site or the mean pH value expected for the sampling site
- *Altitude*: (continuous) The altitude (in meters) of the sampling site. Because information on altitude was rarely provided in the articles, approximate values retrieved for the broader area close to the sampling site were used.
- *Bulk density*: (continuous) Unless reported an estimate of the bulk density of the soil was obtained

based on soil texture information (according to Saxton et al. 1986) and soil classification information (according to Brady and Weil 2008). In three studies where it was impossible to retrieve soil texture information mean values of bulk density (bulk density equalled 1.42) were used to prevent exclusion of the studies from data analysis.

- *Mean annual temperature*: (continuous) Mean annual temperature (in degrees Celsius) was taken from the article; when not present we searched internet resources for an area as close to the sampling site as possible.
- *Mean annual precipitation*: (continuous) Mean annual precipitation (in mm rain) was taken from the article; when not present we searched internet resources for an area as close to the sampling site as possible.
- *Season of harvest*: (categorical) For sampling sites in the Northern hemisphere values ranged from 0—autumn to 3—summer. For sampling sites in the Southern hemisphere values ranged from 0—spring to 3—winter. When sampling had been carried out over two consecutive months the mean was used to describe the season of harvest. If harvest was carried out over more than one non-consecutive months then the variable was left blank.
- *Temperature of month of harvest*: (continuous) In the instances when it was possible to identify the season of harvest climatic information on temperature for the respective month of harvest were retrieved from online databases.

Data analysis

1. Correcting for detection probability

The raw data we used for the analysis was the presence/absence distribution of the two targeted taxa so that, using logistic regression, we analysed the probability that a site contained the targeted AM fungi. The probability of detecting any organism in the field can be partitioned into two probabilities: the probability of occupancy or occurrence and the probability of detecting the species if the species is present (e.g. MacKenzie et al. 2002, 2003, 2006). The latter accounts for the fact that “not detecting” a species does not necessarily imply that the species is absent. In fact, detecting a species strongly depends on the sampling effort, which in our case depends on the

number of sequences available for each sampling site. Therefore, a critical step in the analysis of the database was to correct for the probability that Gigasporaceae or Acaulosporaceae sequences were present but not found in the community analysis. We assume that this probability depends on the number of sequences obtained in each sampling site. Consequently, we also made the assumption that, if present in the community, there was a fixed likelihood α (alpha) in each sequential draw of retrieving a Gigasporaceae or Acaulosporaceae sequence. The binomial probability that a sequence had been present in the sequence pool but was not sampled is $P(n') = \alpha^n$ (n being the number of sequences). Therefore the probability for the complementary event $P(n)$ of a Gigasporaceae or Acaulosporaceae sequence to have been present in the sequence pool and sampled in a trial equalled:

$$P(n) = 1 - \alpha^n \quad (1)$$

The parameter α was assessed using the least squares method by fitting of the presence-absence raw data of the AM fungal families to the realized clone-sequencing effort. In doing so, there was an implicit assumption that the sequencing effort was the sole regulator of the probability of detection of the AM fungal family. This parameter was 0.937 for Gigasporaceae and 0.957 for Acaulosporaceae.

Since we use logistic regressions, in our models the probability of detecting a Gigasporaceae or Acaulosporaceae sequence $Q(f(x))$ as a function of the number of sequences was incorporated in the form of a logistic parameter. The formula that gives the impact of the linear function $f(x)$ of any continuous parameter “x” on the logistic probability is as follows:

$$Q(f(x)) = e^{f(x)} / (1 + e^{f(x)}) \quad (2)$$

In our analysis we wanted to modify the number of sequences “n” to a linear function $f(x)$ so that the logistic function $Q(f(x))$ would give the probability of detecting a sequence. Thereby we wanted a transformation of “n” so that $Q(f(x)) = P(n)$. From (1) and (2) we deduce:

$$f(x) = \ln(\alpha^{-n} - 1) \quad (3)$$

The overall outcome of this preliminary modelling is that we obtained a term (called “detection term” below) which we were able to use in the subsequent modelling to correct for the bias of having sampling sites which were sampled with different intensity in terms of number of sequences.

2. Modelling

Modelling was carried out independently for Gigasporaceae and Acaulosporaceae based on logistic mixed modelling—only linear logistic terms were considered with the “study site” as a random effect plus the detection term parameter.

The first step of the analysis was the assessment of the importance of the study site characterisation variables. These were each tested separately for significance. When more than one study site-related factor was significant, all possible combinations of these significant factors were tested and the combination that minimised the Akaike Information Criterion (AIC) was selected to represent the study site-related terms in the models that reflected the environmental conditions best. After this first step of the analysis three groups of parameters had been screened for inclusion into all possible models that were to be examined in the next step; across-trials study site dependence, number of sequences and significant study site characterisation variables.

The next step of the modelling process included the evaluation of 16 candidate models that comprised all possible combinations of the environmental variables. When defining the 16 candidate or competing models, we used the approach advocated by Burnham and Anderson (2002) and based on the Akaike information criterion: models need to be specified *a-priori* and fitted simultaneously in order to avoid data dredging, which is likely to arise when one adopts classical approaches such as stepwise regressions (see also Johnson and Omland 2004 for a review on the topic). Only five continuous variables, pH, altitude, bulk density, mean annual temperature and mean annual precipitation, were considered and these represented fundamental parameters of what a “niche” might be in the case of the targeted taxa. Given the lack of knowledge, all possible combinations of the five environmental variables were considered.

As mentioned above, model selection was based on an information theoretic approach using AIC. Given the ratio between the number of predictors

and observations, we actually used the AIC corrected for small sample size (AIC_c; Burnham and Anderson 2002). The model with the minimum AIC_c was used as reference for calculating difference (Δ_i) in AIC, thus ranking all models based on their relative performance. The likelihood of a model for given data and model weights (w_i) were calculated and models within two AIC units of AIC_{min} were considered more plausible than the others; their weights were used as a measure of the strength of evidence in their favour (*sensu* Burnham and Anderson 2002).

Given the uncertainty regarding the best model (see results) and the fact that the magnitude of the effect of environmental correlates was often very modest and associated with larger uncertainty, we decided to use multimodel inference based on model averaging, which allowed a more conservative estimate of the effect of environmental variables (Burnham and Anderson 2002).

Parameter estimates for models were performed using R (v2.12.0; R Development Core Team 2008). Models were fitted with the package lme4 (Bates et al. 2011). Model averaging was conducted with package MuMin (Barton 2011), which is available at <http://www.r-project.org>.

3. Validation

A major issue when fitting ecological niche models is to assert that the parameter estimates reflect information from the entire dataset and not individual studies. To test whether individual studies may have had a high overall impact on the parameter estimates we have conducted a sensitivity analysis as follows: Following model building and parameter assessment, we have sequentially fitted parameters to the entire dataset with the exclusion of a single study at a time. The parameters were then re-assessed for the newly-fitted model. If the re-assessed means for the parameters were outside the 95% confidence intervals of the full model then the sensitivity analysis highlighted a lack of confidence in the model.

Results

Preliminary analysis revealed that the factors “season of harvest” and “temperature of month of harvest” performed poorly—this could be due to the fact that temporal data were missing from a large number of

studies, reducing the power of the analysis (Nakagawa and Freckleton 2011). Consequently, the two variables were excluded from further analysis.

Gigasporaceae

None of the study site characterization variables were significant in the evaluation of the probability of occurrence of Gigasporaceae. Thus, the models tested comprised only the across study site dependence term and the sequences term, in addition to environmental variables. Performance of individual environmental models is presented in Table 1. Model averaging revealed that the only statistical signal (CI not including 0) was for the detection term (i.e. number of sequences; Table 2a). However, another clear piece of information which emerges from AIC based model selection is that the relative importance of *Bulk density* (0.86) *Mean annual precipitation* (0.65) and *pH* (0.58) was quite high and, in fact, the model that minimized AIC contained all these factors (Table 1). Moreover, likelihood ratio tests for the relative performance of the models i) with the sole predictive variable *Bulk density* against the null model ii) with the sequential addition of *Precipitation* and *pH* tested against the respective model without the additive parameter resulted in *P* values below 0.05. We concluded that any further model of the probability of occurrence of Gigasporaceae should include data on bulk density, precipitation and pH in spite of the fact that currently available data do not have the power (CI from model averaging includes 0) to quantify their effects. The scaled Nagelkerke r^2 value (Nagekerke 1991), a pseudo R-squared, for the model that incorporated the abovementioned three variables was 0.23. A three dimensional graph on the impact of bulk density and pH on the probability of occurrence of Gigasporaceae for a high sequencing effort (predictive model) and average values for non-illustrated parameters is provided in Fig. S1.

Acaulosporaceae

The study site-related variables that were found significant were fertilisation and disturbance. Based on the AIC criterion we preferentially included fertilisation along with the other correcting factors in all models. Performance of individual environmental models is presented in Table 1. This time, model

averaging clearly detected the signal (CI not including 0) of two parameters, bulk density and pH. However, the model with the optimum AICc additionally included altitude (Table 2). The Nagelkerke r^2 value of the model that incorporated soil bulk density, pH and altitude was 0.49. The probability of detecting Acaulosporaceae isolates was low in neutral and alkaline soils (pH equal or above 7.0) whereas there were no records of Acaulosporaceae occurring in soils with a bulk density below 1.3 (Fig. 2). Probability of occurrence of Acaulosporaceae isolates when only bulk density and pH were considered and average values for non-illustrated parameters other than sequencing effort (which was assumed to be high) is depicted in Fig. S2. Sensitivity analysis revealed that parameter estimates were robust against exclusion of individual studies (Figs. S3, S4).

Discussion

Results of this study revealed that distinct environmental drivers shape the niches of Acaulosporaceae and Gigasporaceae isolates. We have based our results on an unprecedentedly large amount of data from a broad range of study sites. Acaulosporaceae are sensitive to low soil bulk densities and preferentially occur in acidic environments whereas Gigasporaceae are affected by bulk density, precipitation and pH. High dependency of Gigasporaceae isolates on precipitation is in agreement with Egerton-Warburton et al. (2007) who demonstrated that from a set of three mesic and two semiarid grasslands Gigasporaceae and Acaulosporaceae could only be retrieved in the mesic sites, and the preferential establishment of Acaulosporaceae in acidic soils is in agreement with reports that have described specific Acaulosporaceae isolates to be acidophilic (e.g. Oehl et al. 2006). We had also predicted the dependence of Gigasporaceae on soil bulk density (Lekberg et al. 2007) but it came as a surprise that N-fertilisation did not affect Gigasporaceae occurrence (e.g. Johnson et al. 2003). Given the comparable statistical power that was available for the analysis of the two AM fungal families we conclude that Acaulosporaceae isolates may exhibit a higher environmental niche specialisation than Gigasporaceae representatives.

Our analysis offers the first empirical support of this expectation in terms of the effect that key environmental parameters exert on the distribution of these

Table 1 Performance statistics and relative weights of models for Gigasporaceae and Acaulosporaceae. In both modelling attempts performance of the “global” model (that included all

candidate parameters) was low (weight < 0.05). Parameter numbers stand for: 1–Altitude, 2–Bulk density, 3–Mean annual temperature, 4–pH, 5–Mean annual precipitation

Parameters included	Gigasporaceae				Acaulosporaceae			
	Deviance	AICc	Δ AICc	Weight	Deviance	AICc	Δ AICc	Weight
(–)	94.8	101.0	8.59	0.00	100.8	109.2	26.13	0.00
(1)	94.0	102.4	9.95	0.00	94.4	105.0	21.95	0.00
(2)	88.2	96.6	4.11	0.02	87.4	98.1	14.99	0.00
(3)	92.4	100.8	8.30	0.00	100.5	111.1	28.06	0.00
(4)	90.6	99.0	6.54	0.01	86.0	96.7	13.61	0.00
(5)	87.8	96.2	3.71	0.02	99.3	109.9	26.87	0.00
(1,2)	87.6	98.2	5.72	0.01	82.5	95.4	12.33	0.00
(1,3)	90.8	101.4	8.92	0.00	94.4	107.3	24.22	0.00
(1,4)	90.4	101.0	8.52	0.00	83.0	95.9	12.84	0.00
(1,5)	87.1	97.7	5.24	0.01	93.7	106.6	23.57	0.00
(2,3)	84.3	94.9	2.44	0.04	87.3	100.2	17.15	0.00
(2,4)	84.5	95.2	2.67	0.04	70.3	83.2	0.13	0.28
(2,5)	81.9	92.6	0.08	0.14	86.2	99.2	16.11	0.00
(3,4)	86.6	97.2	4.73	0.01	86.0	98.9	15.8	0.00
(3,5)	87.8	98.4	5.92	0.01	96.4	109.3	26.25	0.00
(4,5)	85.1	95.8	3.27	0.03	85.9	98.8	15.8	0.00
(1,2,3)	82.9	95.8	3.31	0.03	82.4	97.6	14.58	0.00
(1,2,4)	84.2	97.2	4.66	0.01	67.8	83.1	0.00	0.30
(1,2,5)	81.4	94.4	1.91	0.06	81.8	97.0	14.0	0.00
(1,3,4)	85.7	88.7	6.16	0.01	82.9	98.2	15.12	0.00
(1,3,5)	87.1	100.0	7.47	0.00	93.1	108.4	25.30	0.00
(2,3,4)	80.3	93.2	0.68	0.11	70.2	85.5	2.41	0.09
(2,3,5)	81.9	94.8	2.33	0.05	83.4	98.7	15.60	0.00
(2,4,5)	79.6	92.5	0.00	0.15	70.2	85.5	2.41	0.09
(3,4,5)	84.8	97.8	5.26	0.01	85.4	100.7	17.63	0.00
(1,2,3,4)	79.4	94.6	2.11	0.05	67.8	85.4	2.32	0.09
(1,2,3,5)	81.3	95.6	4.06	0.02	81.3	98.9	15.87	0.00
(1,2,4,5)	79.4	94.6	2.11	0.05	67.8	85.4	2.36	0.09
(1,3,4,5)	84.3	99.5	7.03	0.00	82.9	100.5	17.48	0.00
(2,3,4,5)	79.2	94.5	1.97	0.06	69.8	87.5	4.41	0.03
(1,2,3,4,5)	78.7	96.3	3.85	0.02	67.7	87.8	4.71	0.03

Deviance is a measure of fit of a model based on maximum likelihood estimation. When more parameters are considered deviance typically becomes smaller. By contrast, AICc is a trade off between squared bias (declines with additional parameters) and variance (increases with additional parameters) (Burnham and Anderson 2002)

two AM fungal families. Indeed, results of our study revealed distinct environmental requirements for Acaulosporaceae and Gigasporaceae. We claim that our results are robust given the very conservative modelling approach we have employed. However, only 35 studies met the inclusion criteria and this

may raise issues of lack of statistical power. For example, due to the relatively limited number of observations, it only made sense to consider linear terms in the logistic models but ecological niches are typically described by unimodal functions. Still, in spite of this limitation, we detected clear statistical signals and

Table 2 Averaged model parameters for predicting a. Gigasporaceae and b. Acaulosporaceae occurrence with the use of logistic regression models. CI = confidence interval

	Coefficient	SE	Lower CI	Upper CI
a. Gigasporaceae				
(Intercept)	-6.19	4.07	-14.20	1.77
Altitude	-0.00017	0.00053	-0.0019	0.00086
Bulk density	3.42	2.21	-0.91	7.77
pH	-0.24	0.27	-0.78	0.30
Precipitation	0.00074	0.00073	-0.00070	0.0021
Sequencing effort	0.031	0.013	0.0057	0.056
Temperature	0.029	0.052	-0.074	0.13
b. Acaulosporaceae				
(Intercept)	-3.08	3.29	-9.53	3.36
Altitude	-0.00067	0.00098	-0.0026	0.0013
Bulk density	7.08	2.27	2.62	11.50
Fertilisation	-0.48	0.76	-1.97	1.01
pH	-1.26	0.40	-2.04	-0.49
Precipitation	-0.000026	0.00032	-0.00061	0.00066
Sequencing effort	0.039	0.018	0.0047	0.074
Temperature	-0.00067	0.029	-0.057	0.056

therefore our hypotheses received a robust support. In the supplementary materials section we include two predictive graphs on the probability of occurrence of

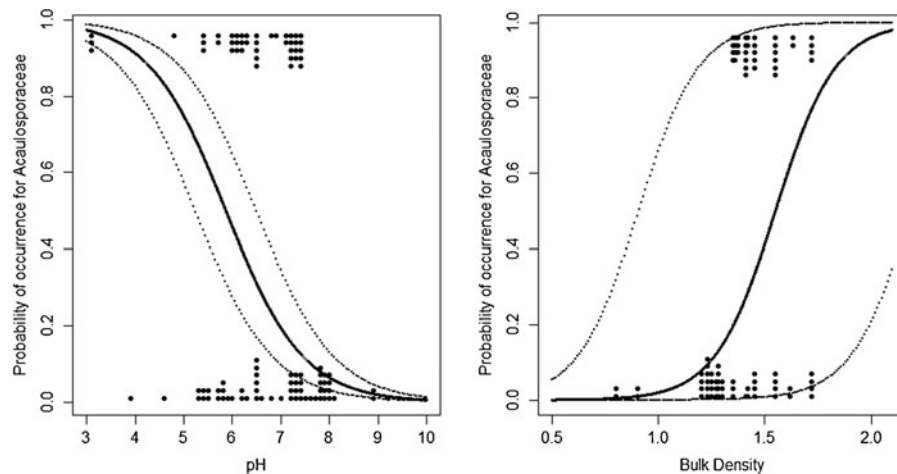


Fig. 2 Mean (continuous lines) \pm 95% CI (dotted lines) estimates following model averaging of the influence of pH and bulk density on probability of occurrence of Acaulosporaceae. Lines have been generated for the mean estimates for the other environmental variables from the dataset and for an average level of sequencing effort. Note that the probability of detecting Acaulosporaceae isolates collapses for alkaline soils and that there have been no occurrences of Acaulosporaceae in soils with

Gigasporaceae and Acaulosporaceae as a function of the two most influential environmental variables. Yet, these predictive models in the absence of quadratic terms assumed monotonic taxa responses to the environmental factors which are unrealistic. We have resorted to this compromise because the amount of available information was limited. The accuracy of the inferred models is maximal for average parameter values but declines for the extremes. Nonetheless, the aim of the specific study was simply to identify the key environmental factors that may regulate occurrence of Glomeromycota.

While it is in our view quite hard to objectively define abundance, population size and even species in microbial groups such as AMF, one fact is clear: the unbalanced structure of the phylogenetic tree of the divergent lineages of Gigasporaceae and Acaulosporaceae and their general traits suggest that broad differences in niche features could be detected at a relatively high rank of the phylogeny (i.e., trait conservatism; Powell et al. 2009).

In the general ecological literature there is an ongoing debate on whether “neutral” or “niche” processes can describe diversity patterns (Adler et al. 2007; Clark 2012). This debate has recently attracted the interest of soil microbiologists and, with regards to AM fungi, the more general question has been in

bulk density below 1.3. Overlaid black bullets correspond to occurrence instances (presence for the bullets in the upper part; absence for bullets in the bottom part of the graph) of Acaulosporaceae isolates in the raw data screened according to the observed/predicted pH or bulk density of the site. We could not produce an equivalent graph for Gigasporaceae because the only identifiable parameter was the detection term (corresponding to number of sequences)

terms of the relative roles of deterministic (e.g. niche) and stochastic (e.g. demographic drift, priority effects) processes determining AM fungal community structure (Dumbrell et al. 2010a; 2010b; Lekberg et al. 2007; Lekberg et al. 2011; Caruso et al. 2012). The emerging view is that while there is evidence that deterministic factors such as key soil variables (e.g. pH: Dumbrell et al. 2010a) affect AM fungal community structure, stochastic factors can locally predominate. For example, AM fungal communities tend to be dominated by few taxa or even one taxon only but the identity of this taxon varies from site to site (Dumbrell et al. 2010b). The studies cited above used a classical community approach at a species level, or some proxy to the species level (e.g. molecular OTUs). This is in principle appropriate within the framework of debates such as the niche-neutral discussion, but the question of what a species is and how ecologists qualify and quantify species distributions in microbial communities such as AM fungal assemblages is far from being answered exhaustively (Gamper et al. 2008; Amend et al. 2010; Avis et al. 2010; Dickie and FitzJohn 2007; Morton and Msiska 2010; Powell et al. 2011).

The study is the first of its kind to define niche structure within a group of fungi following a synthesis of a considerable amount of available literature reports. “Ecological niche modelling” literature on fungi is largely focused on specific strains of human pathogens (e.g. Baptista-Rosas et al. 2007; Reed et al. 2008); availability of clinical records for disease outbreaks may have facilitated such studies. However, these studies have only addressed spatial aspects of disease spread and not the environmental factors that may have had induced it. Zalar et al. (2011) have attempted to clarify the environmental factors that could have favoured proliferation of fungal opportunistic pathogens in dishwashers throughout the world but they did not engage in any rigorous statistical procedure to describe the structure of a potential “niche” for any of the fungal pathogens. A remarkable study, although not engaging in a broader synthesis, was that of Ponce et al. (2011); the authors sampled 19 sites across Spain to model occurrence probability of the ectomycorrhizal fungus *Boletus edulis* and its common plant host *Cistus ladanifer* to assess suitability of commercial growth of the specific symbionts.

We have presented strong evidence for niche based differentiation between the two AM fungal families tested, Gigasporaceae and Acaulosporaceae. Gigasporaceae

appeared to be more sensitive to soil moisture related factors (precipitation) whereas Acaulosporaceae occurrence probability was affected by the altitudinal-related variable. Taking also into account other pieces of evidence (e.g. Lekberg et al. 2007; Dumbrell et al. 2010a), we could generalise the claim for environmentally-based niche differences for the entire phylum of Glomeromycota. This implies that any hidden AM fungal diversity could potentially be located in unexplored community types that share distinct combinations of environmental factors. It has been strongly argued that in order to overcome the “Linnean shortfall”, our ignorance of the realised diversity on earth, systematic surveys should be directed towards areas with known data deficiency (Brito 2010). Öpik et al. (2010) attempted to synthesise the existing knowledge on global distribution of Glomeromycota, concluding that a large fraction of global AM fungal richness remains unexplored. Determining the environmental factors that are likely to affect AM fungal occurrence most could result in the clear advantage of identifying unexplored sites where sampling of AM fungal communities is highly likely to uncover unknown taxa.

The fact that Glomeromycotan fungi have distinct ecological ranges suggests that their ecological optima may differ and this should be verified experimentally. Yet, the number of studies assessing AM fungal performance under non-fixed conditions is very limited (e.g. Antunes et al. 2011). This single study revealed interactive effects of temperature on performance of AM fungal isolates (Antunes et al. 2011). Thus, to realise the full potential of any AM fungal isolate to enhance fitness of their plant hosts we should try to effectively mimic the conditions of the community types where they preferentially occur. In conclusion, following an information-theoretic approach we presented strong evidence on environmental niche-based differentiation of occurrence of Gigasporaceae and Acaulosporaceae, in agreement with our hypothesis. The environmental parameters identified here represent a first step in developing predictive models of the relative performance of Gigasporaceae and Acaulosporaceae isolates under changing environmental settings and/or in agriculture.

Acknowledgements TC was supported by the Alexander von Humboldt Foundation. We thank Dr Jeff Powell and Dr Federico Luebert for comments on an earlier version of this manuscript.

References

- Adler PB, HilleRisLambers J, Levine JM (2007) A niche for neutrality. *Ecol Lett* 10:95–104
- Amend AS, Seifert KA, Bruns TD (2010) Quantifying microbial communities with 454 pyrosequencing: does read abundance count? *Mol Ecol* 19:5555–5565
- Antunes PM, Koch AM, Morton JB, Rillig MC, Klironomos JN (2011) Evidence for functional divergence in arbuscular mycorrhizal fungi from contrasting climatic origins. *New Phytol* 189:507–514
- Avis PG, Branco S, Tang Y, Mueller GM (2010) Pooled samples bias fungal community descriptions. *Mol Ecol Resour* 10:135–141
- Baptista-Rosas PC, Hinojosa A, Riquelme M (2007) Ecological niche modeling of *Coccidioides* spp. in Western North American deserts. *Ann NY Acad Sci* 1111:35–46
- Bartolomeesteban H, Schenck N (1994) Spore germination and hyphal growth of arbuscular mycorrhizal fungi in relation to soil aluminum saturation. *Mycologia* 86:217–226
- Barton K (2011) MuMIn: multi-model inference. R package version 1.0.0. URL <http://CRAN.R-project.org/package=MuMIn> [accessed 10 October 2011]
- Bates D, Maechler M, Bolker B (2011) lme4: Linear mixed-effects models using Eigen and Eigen. R package version 0.999375-40. <http://CRAN.R-project.org/package=lme4> [accessed 10 October 2011]
- Brady NC, Weil RR (2008) The nature and properties of soils. Pearson Prentice Hall, New Jersey
- Brito D (2010) Overcoming the Linnean shortfall: data deficiency and biological survey priorities. *Basic Appl Ecol* 11:709–713
- Burke DJ (2008) Effects of *Alliaria petiolata* (Garlic mustard; Brassicaceae) on mycorrhizal colonisation and community structure in three herbaceous plants in a mixed deciduous forest. *Am J Bot* 95:1416–1425
- Burnham KP, Anderson DR (2002) Model selection and multi-model inference: a practical information-theoretic approach, 2nd edn. Springer, Berlin
- Caruso T, Hempel S, Powell JR, Barto EK, Rillig MC (2012) Compositional divergence and convergence in arbuscular mycorrhizal fungal communities. *Ecology* 93:1115–1124
- Chase JM, Leibold MA (2003) Ecological niches: linking classical and contemporary approaches. Chicago University Press, Chicago
- Clark RB (1997) Arbuscular mycorrhizal adaptation, spore germination, root colonization, and host plant growth and mineral acquisition at low pH. *Plant Soil* 192:15–22
- Clark JS (2012) The coherence problem with the unified neutral theory of biodiversity. *Trends Ecol Evol* 27:198–202
- da Silva GA, Lumini E, Maia LC, Bonfante P, Bianciotto V (2006) Phylogenetic analysis of Glomeromycota by partial LSU rDNA sequences. *Mycorrhiza* 16:183–189
- Dickie I, FitzJohn R (2007) Using terminal restriction fragment length polymorphism (T-RFLP) to identify mycorrhizal fungi: a methods review. *Mycorrhiza* 17:259–270
- Dumbrell AJ, Nelson M, Helgason T, Dytham C, Fitter AH (2010a) Relative roles of niche and neutral processes in structuring a soil microbial community. *ISME J* 4:337–345
- Dumbrell AJ, Nelson M, Helgason T, Dytham C, Fitter AH (2010b) Idiosyncrasy and overdominance in the structure of natural communities of arbuscular mycorrhizal fungi: is there a role for stochastic processes? *J Ecol* 98:419–428
- Egerton-Warburton LM, Allen EB (2000) Shifts in arbuscular mycorrhizal communities along an anthropogenic nitrogen deposition gradient. *Ecol Appl* 10:484–496
- Egerton-Warburton LM, Johnson NC, Allen EB (2007) Mycorrhizal community dynamics following nitrogen fertilization: a cross-site test in five grasslands. *Ecology* 77:527–544
- Fitzsimons MS, Miller RM, Jastrow JD (2008) Scale dependent niche axes of arbuscular mycorrhizal fungi. *Oecologia* 158:117–127
- Gamper HA, Young JPW, Jones DL, Hodge A (2008) Real-time PCR and microscopy: are the two methods measuring the same unit of arbuscular mycorrhizal fungal abundance? *Fungal Genet Biol* 45:581–596
- Grime JP, Hodgson JG, Hunt R (1990) The abridged comparative plant ecology. Chapman & Hall, London
- Guo YJ, Ni Y, Raman H, Wilson BAL, Ash GJ, Wang AS, Li GD (2012) Arbuscular mycorrhizal fungal diversity in perennial pastures; responses to long-term lime application. *Plant Soil* 351:389–403
- Hart MM, Reader RJ (2002) Taxonomic basis for variation in the colonization strategy of arbuscular mycorrhizal fungi. *New Phytol* 153:335–344
- Helgason T, Fitter AH, Young JPW (1999) Molecular diversity of arbuscular mycorrhizal fungi colonizing *Hyacinthoides non-scripta* (bluebell) in a seminatural woodland. *Mol Ecol* 8:659–666
- Helgason T, Merryweather JW, Young JPW, Fitter AH (2007) Specificity and resilience in the arbuscular mycorrhizal fungi of a natural woodland community. *J Ecol* 95:623–630
- Hempel S, Renker C, Buscot F (2007) Differences in the species composition of arbuscular mycorrhizal fungi in spore, root and soil communities in a grassland ecosystem. *Environ Microbiol* 9:1930–1938
- Johnson JB, Omland KS (2004) Model selection in ecology and evolution. *Trends Ecol and Evol* 19:101–108
- Johnson NC, Rowland DL, Corkidi L, Egerton-Warburton LM, Allen EB (2003) Nitrogen enrichment alters mycorrhizal allocation at five mesic to semiarid grasslands. *Ecology* 84:1895–1908
- Klironomos JN, Hart MM (2002) Colonization of roots by arbuscular mycorrhizal fungi using different sources of inoculum. *Mycorrhiza* 12:181–184
- Klironomos J, Zobel M, Tibbett M, Stock WD, Rillig MC, Parrent JL, Moora M, Koch AM, Facelli JM, Facelli E, Dickie IA, Bever JD (2011) Forces that structure plant communities: quantifying the importance of the mycorrhizal symbiosis. *New Phytol* 189:366–370
- Lekberg Y, Koide RT, Rohr JR, Aldrich-Wolf L, Morton JB (2007) Role of niche restrictions and dispersal in the composition of arbuscular mycorrhizal fungal communities. *J Ecol* 95:95–105
- Lekberg Y, Schnoor T, Kjoller R, Gibbons SM, Hansen LH, Al-Soud WA, Sørensen SJ, Rosendahl S (2011) 454-sequencing reveals stochastic local reassembly and high disturbance tolerance within arbuscular mycorrhizal fungal communities. *J Ecol* 100:151–160

- MacKenzie DI, Nichols JD, Lachman GB, Droege S, Royle JA, Langtimm CA (2002) Estimating site occupancy rates when detection probabilities are less than one. *Ecology* 83:2248–2255
- MacKenzie DI, Nichols JD, Hines JE, Knutson MG, Franklin AB (2003) Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology* 84:2200–2207
- MacKenzie DI, Nichols JD, Royle JA, Pollock KH, Bailey LL, Hines JE (2006) Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Elsevier, San Diego
- Maherali H, Klironomos JN (2007) Influence of phylogeny on fungal community assembly and ecosystem functioning. *Science* 316:1746–1748
- Maherali H, Klironomos JN (2012) Phylogenetic and trait-based assembly of arbuscular mycorrhizal fungal communities. *PlosOne* 7:e36695
- Morton J, Msiska Z (2010) Phylogenies from genetic and morphological characters do not support a revision of Gigasporaceae (Glomeromycota) into four families and five genera. *Mycorrhiza* 20:483–496
- Nagekerke N (1991) A note on a general definition of the coefficient of determination. *Biometrika* 78:691–692
- Nakagawa S, Freckleton RP (2011) Model averaging, missing data and multiple imputation: a case study for behavioural ecology. *Behav Ecol Sociobiol* 65:103–116
- Oehl F, Sýkorová Z, Redecker D, Wiemken A, Sieverding E (2006) *Acaulospora alpina*, a new arbuscular mycorrhizal fungal species characteristic for high mountainous and alpine regions of the Swiss Alps. *Mycologia* 98:286–294
- Oehl F, de Souza FA, Sieverding E (2008) Revision of *Scutelospora* and description of five new genera and three new families in the arbuscular mycorrhiza forming Glomeromycetes. *Mycotaxon* 106:311–360
- Öpik M, Vanatoa A, Vanatoa E, Moora M, Davison J, Kalwij JM, Reier Ü, Zobel M (2010) The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). *New Phytol* 188:223–241
- Ponce RA, Águeda B, Ágreda T, Modrego MP, Aldea J, Fernández-Toirán LM, Martínez-Pena F (2011) Rockroses and *Boletus edulis* ectomycorrhizal association: realized niche and climatic suitability in Spain. *Fungal Ecol* 4:224–232
- Powell JR, Parrent JL, Hart MM, Klironomos JN, Rillig MC, Maherali H (2009) Phylogenetic trait conservatism and the evolution of functional trade-offs in arbuscular mycorrhizal fungi. *Proc Biol Sci* 276:4237–4245
- Powell JR, Monaghan MT, Öpik M, Rillig MC (2011) Evolutionary criteria outperform operational approaches in producing ecologically relevant fungal species inventories. *Mol Ecol* 20:655–666
- R Development Core Team (2008) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Reed KD, Meece JK, Archer JR, Peterson AT (2008) Ecologic niche modeling of *Blastomyces dermatitidis* in Wisconsin. *PLoS One* 3:e2034
- Rillig MC (2004) Arbuscular mycorrhizae and terrestrial ecosystem processes. *Ecol Lett* 7:740–754
- Rosendahl S (2008) Communities, populations and individuals of arbuscular mycorrhizal fungi. *New Phytol* 178:253–266
- Saxton KE, Rawls WJ, Romberger JS, Papendick RI (1986) Estimating generalized soil-water characteristics from texture. *Soil Sci Soc Am J* 50:1031–1036
- Schenck NC, Pérez Y (1990) Manual for the identification of VA mycorrhizal fungi, 3rd edn. Synergistic, Cainesville
- Schüßler A, Walker C (2010) The Glomeromycota. A species list with new families and new genera. Oregon State University, Gloucester
- Schüßler A, Schwarzott D, Walker C (2001) A new fungal phylum, the Glomeromycota: phylogeny and evolution. *Mycol Res* 105:1413–1421
- Smith SE, Read DJ (2008) *Mycorrhizal symbiosis*, 3rd Edition. Academic Press
- Soberon J (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecol Lett* 10:1115–1123
- Sýkorová Z, Ineichen K, Wiemken A, Redecker D (2006) The cultivation bias: different communities of arbuscular mycorrhizal fungi detected in roots from the field, from bait plants transplanted to the field, and from a greenhouse trap experiment. *Mycorrhiza* 18:1–14
- Wang B, Qiu Y-L (2006) Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza* 16:299–363
- Zalar P, Novak M, de Hoog GS, Gunde-Cimerman N (2011) Dishwashers—a man-made ecological niche accommodating human opportunistic fungal pathogens. *Fungal Biol* 115:997–1007
- Zhu HH, Yao Q, Sun XT, Hu YL (2007) Colonization, ALP activity and plant growth promotion of native and exotic mycorrhizal fungi at low pH. *Soil Biol Biochem* 39:942–950