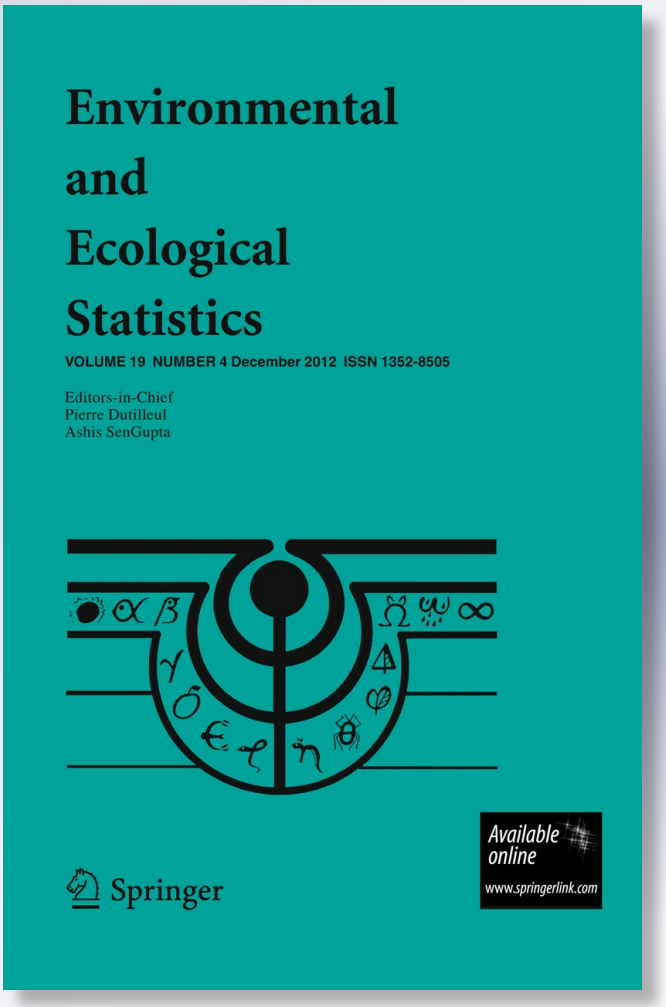
Multivariate analyses in soil microbial ecology: a new paradigm

Jean Thioulouse, Yves Prin & Robin Duponnois

Environmental and Ecological Statistics

ISSN 1352-8505
Volume 19
Number 4

Environ Ecol Stat (2012) 19:499-520
DOI 10.1007/s10651-012-0198-z



Springer

Environ Ecol Stat (2012) 19:499–520
DOI 10.1007/s10651-012-0198-z

Multivariate analyses in soil microbial ecology: a new paradigm

Jean Thioulouse · Yves Prin · Robin Duponnois

Received: 4 October 2010 / Revised: 7 March 2012 / Published online: 18 April 2012

Abstract Mycorrhizal symbiosis is a key component of a sustainable soil-plant system, governing the cycles of major plant nutrients and vegetation cover. The mycorrhizosphere includes plants roots, the mycorrhizal fungi, and a complex microbial compartment. A large number of methods have been used to characterize the genetic and functional diversity of these soil microbial communities. We present here a review of the multivariate data analysis methods that have been used in 16 research articles published in the 2005–2009 period. "Descriptive" multivariate data analysis methods have been priviledged over classical "predictive" methods and univariate statistical tests. Data sets, multivariate data analysis methods, graphical outputs and interpretation results are presented and explained in details on several examples coming from some of the 16 articles. These multivariate and graphical methods are available in the ade4 package for the R statistical software. The discussion underlines the importance of using appropriate statistical methods to obtain a good description

J. Thioulouse (✉)
Université de Lyon, 69000 Lyon; Université Lyon 1; CNRS, UMR 5558, Laboratoire de Biométrie et Biologie Evolutive, 69622 Villeurbanne Cedex, France
e-mail: jean.thioulouse@univ-lyon1.fr

Y. Prin · R. Duponnois
Laboratoire des Symbioses Tropicales et Méditerranéennes,
UMR 113 CIRAD, IRD, Université Montpellier 2, SupAgro, USC INRA,
Campus International de Baillarguet TA-A82/J,
34398 Montpellier Cedex 5, France
e-mail: yves.prin@cirad.fr

R. Duponnois
e-mail: robin.duponnois@ird.fr

of soil microbiological and biochemical indicators and a good evaluation of soil quality.

Keywords Mycorrhizal symbiosis · Soil microbial diversity · Descriptive multivariate data analysis · BGA · Coinertia analysis · ade4

1 Introduction

Mycorrhizal fungi are a ubiquitous component of most ecosystems throughout the world (Brundrett 2002). By governing major plant nutrient cycles and sustaining the vegetation cover (Schreiner et al. 1997; Johansson et al. 2004), they are also a key component of sustainable soil-plant systems. Benefits derived by plants from mycorrhizal symbiosis include (i) increased plant uptake of low mobility minerals (i.e. phosphorus), micronutrients and nitrogen, (ii) enhanced water absorption and (iii) improved plant health by acting against some pathogens (Smith and Read 2008). Arbuscular mycorrhizas (AM) symbiosis is the most widespread mycorrhizal association and affects about 80–90 % of land plants with true roots (i.e. pteridophytes, gymnosperms and angiosperms) in natural and agricultural ecosystems (Brundrett 2002).

The trophic translocations between the host plant and the fungal symbionts results from the close relationship between each component of the symbiotic association. It has been clearly established (Wirsel 2004; Smith and Read 2008; van der Heijden et al. 1998) that, in addition to increasing the absorptive surface area of their host plant root systems, the extrametrical mycelium provide an increased area for interactions with soil microbiota. The zone influenced by both the root and the mycorrhizal fungus has been named "mycorrhizosphere" and includes one microbial compartment subjected to the dual influence of the root and the mycorrhizal symbionts (the "mycorrhizosphere" sensu stricto) and the other under the influence of mycorrhizal hyphae (the "hyphosphere") (Linderman 1988).

In the present paper, we speak about controlled mycorrhizal inoculation. As it has been reported by Ouahmane et al. (2006a) in revegetation schemes, two main reclamation strategies could be proposed: (i) inoculation of plants with selected mycosymbionts (Ouahmane et al. 2007) and/or (ii) management of the native soil mycorrhizal potential through drought-tolerant, native and highly mycotrophic plant species (Duponnois et al. 2001; Azcon-Aguilar et al. 2003). According to the characteristics of the mycorrhizal soil potential and its associated microflora (abundance, diversity), the introduction of mycorrhizal propagules into the soil could induce strong modifications in the soil microbial characteristics that could decrease the potential effect of these fungal symbionts on soil biofunctioning and plant growth (Dabire et al. 2007). Hence, the structure of mycorrhizal communities has to be evaluated in order to adopt one of these cultural strategies (controlled mycorrhization when the soil mycorrhizal potential is to low too ensure its effects on plant growth and soil functions or soil mycorrhizal management in order to increase native fungal diversity and abundance).

Beside the effects of the mycorrhizal symbiosis on soil microbiota, it has also been demonstrated that mycorrhizal symbiosis had a selective pressure on bacterial communities and favored soil microorganisms potentially beneficial to the symbiosis and

to the host plant (Frey-Klett et al. 2005; Ouahmane et al. 2009). Hence, the relationships between the fungal symbiont and the host plant have been enlarged to the soil microbiota to create a multitrophic mycorrhizal complex (Frey-Klett et al. 2005).

The mycorrhizosphere compartment is usually characterized by a high microbial activity which could contribute to the biocontrol of pathogens and improve supply of nutrients in degraded soils to maintain health and growth of plant species (Johansson et al. 2004). Hence, the positive impact of the mycorrhizal symbiosis on nutrition, N uptake, and disease resistance of host plants could be considered as an indicator of soil quality, in conjunction with the mycorrhizosphere microbiota. Soil quality has been defined as "the capacity of a soil to function, within the limits imposed by the ecosystem, to preserve the biological productivity and environmental quality, and promote plant, animal and human health" (Doran and Parkin 1994).

The quantification of the soil quality was usually based on physical and chemical indexes. It has been clearly established (Giller et al. 1997; Smith and Read 2008) that soil functioning resulted from complex interactions between soil physical, chemical and biological processes. Hence, soil quality cannot be assessed with one variable but with a combination of these factors (Barrios et al. 2006) showing the state of soil (Dumanski and Pieri 2000). Soil microbial functional diversity is a good indicator of soil quality, as it is integrative of multiple processes taking place in the soil. A large number of methods have been used to characterize the genetic and functional diversity of complex soil microbial communities. All these methods generate high volumes of data that cannot be analyzed by conventional methods, due to several problems: (i) the high number of variables measured compared to the number of samples, (ii) the potentially high correlations between these variables, and (iii) the low density of information that they contain individually. Appropriate statistical tools must be used to investigate these complex microbial interactions and to provide relevant analyses on the role of each variable involved in soil quality.

This review is focussed on the potential benefits of using "descriptive" multivariate statistical methods such as principal component analysis, between-group analysis, and co-inertia analysis instead of other common statistical data analysis techniques (called here "predictive" methods), such as linear discriminant analysis (LDA), redundancy analysis (RDA) and canonical correspondence analysis (CCA). These "descriptive" methods should help us to get new insights into the functioning of the mycorrhizosphere and to develop indicators of soil quality and ecological resilience. The resulting outputs could be of great relevance to describe and explain biological interactions that are considered key components in the stability and productivity of terrestrial ecosystems.

The objectives of this paper are twofold. First, we want to report on some of the "descriptive" methods used by the authors, comparing them with the common "predictive" multivariate techniques (see for example the paper by Ramette (2007), who recently published a good review of these methods in the field of microbial ecology). Secondly, we also want to discuss the potential for these methods to provide new insights into the functioning of the mycorrhizosphere and help develop indicators of soil quality and ecological resilience, although this last step is still a work in progress. The data sets needed for these studies must take into account all the actors that participate in this story, namely: fungi, bacteria, plants and environment (i.e. abiotic characteristics).

To achieve these goals, we analyzed 16 research studies published during the 2005–2009 period, in which we used various multivariate data analysis methods, mainly principal component analysis (PCA), between-group analysis (BGA) (Doledec and Chessel 1987; Culhane et al. 2002), and co-inertia analysis (CoIA) (Doledec and Chessel 1994; Dray et al. 2003). We first present the kind of data tables that we analyzed in these studies, and we give a short summary of the properties of the data analysis methods, in the framework of the duality diagram (Escoufier 1987; Holmes 2006) and of the ade4 package (Chessel et al. 2004; Dray and Dufour 2007) for the R environment (R Development Core Team 2010). We also show the advantages of using these methods on several examples taken from these studies. Lastly, we present several types of graphics used with BGA and CoIA to facilitate the interpretation of results.

2 Data sets

The data needed to analyze the mycorrhizosphere effect on the structure and functioning of soil microbial communities are very diverse. They can be classified according to the subject of measure: fungi, bacteria, plant and environment (i.e. abiotic characteristics).

Fungus measures can be based on the mycorrhization rate, the mycorrhizal hyphal length, the fungal species, or the genetic fungal strain. Bacterial communities are too complex to be analyzed exhaustively, but their genetic diversity can be approached by PCR-based molecular biology methods, like for example denaturing gradient gel electrophoresis (DGGE) (Nakatsu et al. 2000), or by the use of fatty acid patterns of phospholipids and lipopolysaccharides (Zelles 1999). Their functional diversity can also be easily characterized by measurement of the patterns of in situ catabolic potential (ISCP) (Degens and Harris 1997; Degens et al. 2001). Patterns of ISCP provide a real time measure of microbial functional diversity since they assess substrate catabolism by microbial communities in soils without extracting organisms as it is required in the cultured-based methods. ISCP is based on the measurements of CO2 production of soils moistened with a range of simple organic compounds. This process is called Substrate Induced Respiration (SIR).

Plants can be described by growth variables, dry weight of the whole plant or of particular organs (shoot, root), nitrogen and phosphorus content, and many other variables. Lastly, soil samples provide numerous physico-chemical variables (particle size, pH, concentration of many chemical compounds) that determine environmental conditions.

The 16 research studies on which this review is based have been summarized in two tables: Table 1 for studies using BGA, and Table 2 for studies using CoIA. These tables give, for each paper, the bibliographic reference, the main ecological questions, and a summary of biological variables and environmental factors analysed in the study.

3 Multivariate analysis methods

We have seen that five types of tables can be involved in the analysis of the mycorrhizosphere effect on the structure and functioning of soil microbial communities: fungal variable tables, plant variable tables, soil variable tables, ISCP tables, and molecular

Table 1 Bibliographic reference, ecological questions and targeted factors in the nine papers where BGA was used

Paper reference	Ecological question	Analysed factors
Duponnois et al. (2006a)	Impact of termite mound amendment on the sorghum tolerance to soil Cd content	Plant growth, fluorescent Pseudomonads abundance and functional diversity, ISCP
Duponnois et al. (2009)	Nurse plant effect in reforestation programs	Plant growth, abundance and functional diversity, ISCP
Faye et al. (2009)	Response of native *Bradyrhizobial* community diversity (structure and functional diversity) to the introduction of an exotic tree species (Australien *Acacia* species)	Genetic and functional diversity, symbiotic performance of Bradyrhizobial strains on the Acacia species *Faidherbia albida*
Kisa et al. (2007)	Response of soil microbial functions and diversity to the introduction of an exotic tree species and assessment of the role of the mycorrhizal symbiosis in plant co-existence	Plant growth, herbaceous plant species layer composition, soil microbial diversity, ISCP
Ouahmane et al. (2009)	Ectomycorrhizal impact on plant growth, rock phosphate solubilization and soil microbial functions	Plant growth, plant nutrition, mycorrhizal colonization, ISCP
Ramanankierana et al. (2006)	Effect of ectomycorrhizal symbiosis on soil microbial functions	Mycorrhizal colonization, soil microbial functional diversity, fluorescent Pseudomonads, ISCP
Ouahmane et al. (2006b)	Impact of shrub species on soil microbial and chemical characteristics and consequences on the early growth of *C. atlantica*	Plant growth, plant nutrition, mycorrhizal colonization, soil chemical characteristics, ISCP
Ramanankierana et al. (2007)	Potentiel benefits of ionoculation with mycorrhizal fungi (ectomycorrhizal and/or arbuscular mycorrhizal fungi) on plant growth and on functional diversity of soil microflora	Plant growth, plant nutrition, mycorrhizal colonization, ISCP
Remigi et al. (2008)	Response of native soil microflora functions to the introduction of an exotic tree species	ISCP

fingerprint tables. Each type of table can be analyzed separately, and further analyses can be performed according to the scientific question under study, like the examination of relationships between some of these tables, or the search for structures common to all tables.

Table 2 Bibliographic reference, ecological questions and targeted factors in the seven papers where COIA was used

Paper reference	Ecological question	Analysed factors
Andrianjaka et al. (2007)	Impact of termite mound amendment on *Striga* development	Plant growth, mycorrhizal colonization, actinomycete abundance, ISCP
Duponnois et al. (2006b)	Impact of termite mound amendment on ectomycorrhizal symbiosis between *Acacia holosericea* and *Scleroderma dictyosporum*	Plant growth, mycorrhizal colonization, rhizobial colonization, ISCP
Diallo et al. (2006)	Impact of litter amendments on plant growth, soil fertility and soil microbial biomass	Plant growth, litter chemical characteristics, soil microbial biomass, soil nitrogen content
Duponnois et al. (2005a)	Impact of termite mound amendments on plant growth, rock phophate dissolution and soil microbial characteristics	Plant growth, plant mineral nutrition, mycorrhizal colonization, ISCP
Ouahmane et al. (2006a)	Impact of *Lavandula* species on mycorrhizal soil potential, soil microbial functions and on the regeneration process of *Cupressus* spp.	Plant growth, plant nutrition, mycorrhizal soil potential, ISCP
Ouahmane et al. (2007)	Influence of native or exotic fungal symbionts on the plant growth, soil microbial functional diversity and rock phosphate alteration	Plant growth, plant nutrition, mycorrhizal colonization, ISCP
Duponnois et al. (2005b)	Arbuscular mycorrhizal effect on plant growth, soil microbial functions, rock phosphate solubilization and plant P uptake	Plant growth, plant nutrition, mycorrhizal colonization, ISCP

Multivariate analysis methods can be used to attain several distinct objectives. The simplest one is dimensionality reduction, in which the user just wants to reduce the size of the data table, without losing too much information. This is particularly useful in the analysis of DNA fingerprints (RFLP, AFLP, DGGE, TTGE, ARISA, etc). In these profiles, each individual electrophoresis band brings almost no usable information. It is only the combination of many bands that makes the profile useful to discriminate between samples.

Other objectives can be, for example, to find a sample score with maximal correlation with original variables, or finding a set of orthogonal variables in a regression problem (orthogonal regression). But what is important here, compared to univariate approaches, is that the multivariate approach allows to retain the relationships between variables and between samples. It is the correlation structure between variables (and between samples) that brings information, not the values of one variable independently from the others.

PCA is the most basic multivariate analysis method. Several theoretical models lead to the same computational algorithm, based on eigenvalues and eigenvectors decomposition. The most simple of these models is the geometrical model

(LeRoux and Rouanet 2004), which is not based on any distributional hypothesis, and imposes no particular constraint on the data table (as opposed, for example to the multinormal adjustment model). In this geometrical model, PCA can be applied to any numeric data table, regardless of the number of variables, of their correlations, and of their distribution. Moreover, if the data table contains a mixture of quantitative and qualitative variables, then the Hill and Smith procedure can be used (Hill and Smith 1976; Kiers 1991).

Two other methods are of general interest: between-group analysis and coinertia analysis. BGA can be applied when samples belong to several groups. This is the case for example when we want to compare the effect of different treatments, like different levels of amendment, or different rates of mycorrhizal inoculation, on plant growth or on microbial communities. CoIA is useful to analyze the relationships between two tables having the same samples in rows. It can be used for example to explore the relationships between in situ catabolic potential (ISCP, representing bacterial functional diversity) and plant growth variables, or between soil variables and DNA fingerprints.

The absence of constraint on the number of samples compared to the number of variables, on the correlation between variables, and on their distribution is also true for BGA and COIA. This is very important, as the number of variables can be extremely high: several hundreds for the number of bands in DNA fingerprints, or even several thousands for the probes on a DNA chip. Less sophisticated techniques, such as ISCP, can also result in data tables that have more columns than rows. Many statistical methods cannot be used when the number of samples is lower than (or even comparable to) the number of variables, or when the number of explanatory variables is too high. This is the case for example of LDA, CCA and RDA (Ramette 2007).

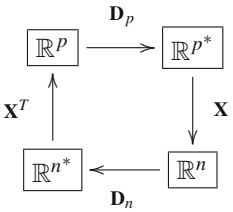
3.1 Between group analysis

BGA can be seen as a robust alternative to linear discriminant analysis (Huberty 1994) when the number of samples is too small compared to the number of variables. The aim of discriminant analysis is to separate groups, or, more precisely, to seek a linear combination of the variables that has a maximal ratio of the separation of the class means to the within-class variance (Venables and Ripley 2002). Here, the groups correspond to treatments used to analyze the mycorrhizosphere effect on soil microbial communities. For example, it can be the level of phosphorus amendment, or the rate of mycorrhizal inoculation, or the origin of soil samples. When the number of samples is high, discriminant analysis gives the coefficients of the discriminant functions that best separate groups. But when the number of samples is low, and particularly when it is lower than the number of variables, discriminant analysis cannot be used. In this case, BGA can be very useful, and provides illustrative graphical displays of between-groups differences.

BGA can also be presented as a particular case of RDA. It corresponds to the case where explanatory variables (also called "constraining variables" in the vegan package) are reduced to a single dummy variable describing the groups.

Here is a short presentation of BGA in the framework of the duality diagram (Holmes 2006). Let us first define the duality diagram of a simple PCA.

Let $\mathbf{X} = [x_{ij}]_{(n,p)}$ be the data table, with n rows (sampling sites) and p columns (variables). $\mathbf{X}^T$ is the transpose of $\mathbf{X}$. Let $\mathbf{D}_n$ be the diagonal matrix $(n \times n)$ of site weights: $\mathbf{D}_n = diag(w_1, \ldots, w_n)$, and let $\mathbf{D}_p$ be a metric on $\mathbb{R}^p$. The duality diagram of the general analysis of this data table is:



This is called a "duality diagram" because $\mathbb{R}^{p^*}$ and $\mathbb{R}^{n^*}$ are the dual spaces of $\mathbb{R}^p$ and $\mathbb{R}^n$, and because the dual operators $\mathbf{X}^T\mathbf{D}_n\mathbf{X}\mathbf{D}_p$ and $\mathbf{X}\mathbf{D}_p\mathbf{X}^T\mathbf{D}_n$ share the same spectrum. This diagram is completely defined by the "triplet notation": $(\mathbf{X}, \mathbf{D}_p, \mathbf{D}_n)$, and the total inertia of this statistical triplet is:

$$I_{\mathbf{X}} = trace(\mathbf{X}\mathbf{D}_p\mathbf{X}^T\mathbf{D}_n)$$

The generalized PCA (gPCA) of this triplet corresponds to the spectral decomposition of $\mathbf{X}^T\mathbf{D}_n\mathbf{X}\mathbf{D}_p$. When $\mathbf{D}_n$ is the matrix of uniform row weights ($w_i = 1/n$), and $\mathbf{D}_p$ is the identity (euclidean metric), then this analysis is a simple PCA, and if the variables are centered, the total inertia is the sum of their variances.

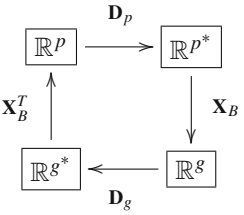
We can now define the duality diagram of Between-Group Analysis. In Between-Group Analysis, samples belong to g groups, namely $G_1, \ldots, G_g$, with group counts $n_1, \ldots, n_g$, and $\sum n_k = n$. Between-Group Analysis is the analysis of triplet $(\mathbf{X}_B, \mathbf{D}_p, \mathbf{D}_g)$, where $\mathbf{X}_B$ is the (g, p) matrix of group means:

$$\mathbf{X}_B = [\bar{x}_j^k]_{(g,p)}.$$

The term $\bar{x}_j^k = \frac{1}{n_k}\sum_{i \in G_k} x_{ij}$ is the mean of variable j in group k. In matrix notation, if $\mathbf{B}$ is the matrix of class indicators: $\mathbf{B} = [b_i^k]_{(n,g)}$, with $b_i^k = 1$ if $i \in G_k$ and $b_i^k = 0$ if $i \notin G_k$, then we have:

$$\mathbf{X}_B = \mathbf{D}_g\mathbf{B}^T\mathbf{X}.$$

Matrix $\mathbf{D}_g = Diag(\frac{1}{n_k})$ is the diagonal matrix of group weights, and $\mathbf{B}^T$ is the transpose of $\mathbf{B}$. The corresponding duality diagram is the following:



Between-Group Analysis is therefore the analysis of the table of group means, leading to the diagonalization of matrix $\mathbf{X}_B^T\mathbf{D}_g\mathbf{X}_B\mathbf{D}_p$. It's aim is to highlight the

differences between groups, and the row scores maximize the between-group variance. The statistical significance of these differences can be tested by a permutation test (Monte-Carlo test), with a criterion equal to the between/total variance ratio. Row scores of the initial data table can be computed by projecting the rows of table **X** on the principal component subspaces.

3.1.1 Using BGA

One of the advantages of BGA is the simplicity of its use: in the case of a table of quantitative variables, it is just the PCA of the table of group means, followed by a projection of the original samples as additional elements in this PCA. This second step provides sample scores that can be used to draw useful graphical displays.

But the biggest advantage of BGA is that it can be used on any type of analysis. In the ade4 package for the R software, basic one-table analysis methods include PCA, CA (correspondence analysis, for count tables), MCA (multiple correspondence analysis, for qualitative variables) and PCO (principal coordinates analyses, for distance matrices). Using BGA on these analyses leads to original methods, like between-group analysis on qualitative variables, between-group analysis on distance matrices, or between-group correspondence analysis. The underlying duality diagram framework ensures that all these methods are coherent and can be used according to the characteristics of the data. There are many other types of analyses in the ade4 package, such as FPCA and FCA (for fuzzy PCA and fuzzy CA), NSCA (non symmetric correspondence analysis), and several other variants. All these analyses are adapted to particular data sets or particular data analysis objectives, and BGA can be used on all these analyses.

3.1.2 BGA examples

Nine of the 16 research studies used BGA (see Table 1). In eight analyses, BGA was done on ISCP data, and in one analysis (Faye et al. 2009), it was on plant and fungal variables. The groups corresponded to different things: the effect of *Pisolithus sp.* inoculation compared with other factors such as phosphorus amendment (Ouahmane et al. 2009), the introduction of an exotic plant species (an Australian Acacia, *Acacia holosericea*) (Remigi et al. 2008), symbiotic bacterial inoculation (Faye et al. 2009), the nurse plants effect (Ouahmane et al. 2006a; Duponnois et al. 2009), or the effect of soil disinfection (Ramanankierana et al. 2007). The other papers were focused on the interactions between *Eucalyptus camaldulensis* seedlings, *Glomus intraradices* inoculation, and fertilizer amendment (Kisa et al. 2007), on cadmium resistance induced by termite mounts powder amendment (Duponnois et al. 2006a), and on the comparison of functional microbial diversity between rhizosphere, hyphosphere, and mycorrhizosphere soil compartments (Ramanankierana et al. 2006).

In all but three of these analyses, the number of samples was less than the number of variables (ISCP substrates), which means that we could not have used "predictive" methods like LDA to separate groups and test the statistical significance of the multivariate between-group differences. BGA allowed us to analyze these data sets and to test the significance of differences.

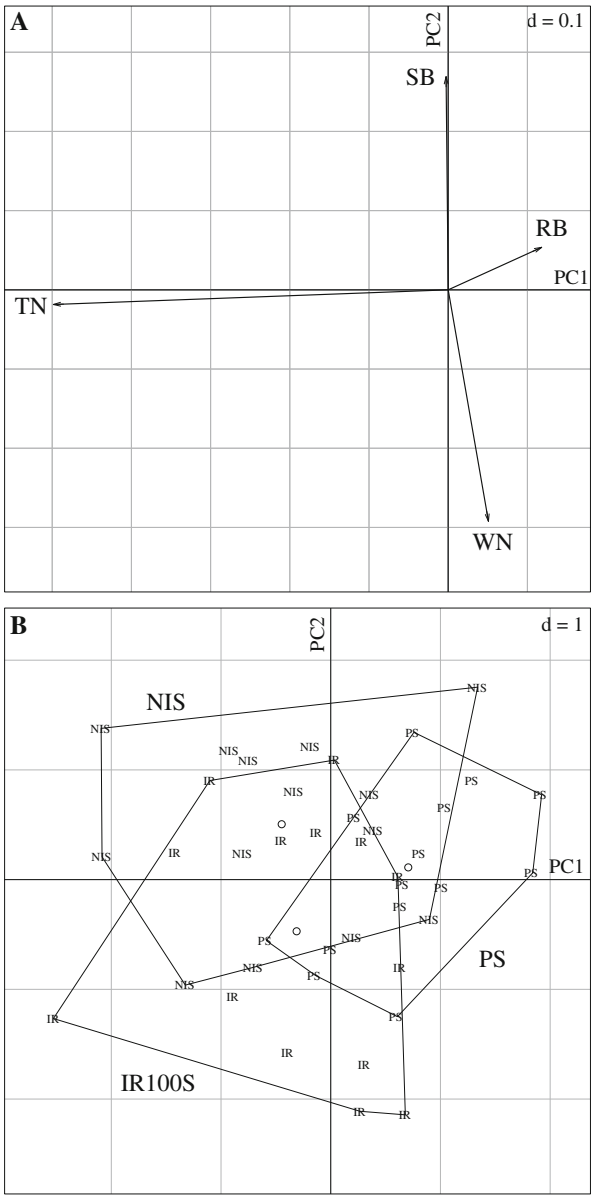


Fig. 1 Between-group analysis (BGA) of *Faidherbia albida* growth (shoot and root biomass: SB and RB, respectively) and nodule formation (total number and dry weight of nodules per plant: TN and WN, respectively). **a** Plot of variable loadings. **b** Plot of sample scores. The scale is given by the value in the *upper right corner*: this value represents the length of the side of background grid squares. The second principal component opposes the shoot biomass (*up*) to the nodule dry weight (*down*). The plot of sample scores (**b**) is split in three groups, according to the origin of the *Bradyrhizobia* isolates: PS, soil of *F. albida* parkland collected outside the *A. holosericea* plantation, NIS, soil of plantation with not inoculated trees, and IR100S, soil of plantation with IR100-inoculated trees. The circle inside each convex hull gives the position of the gravity center of each group. (Reprinted from Faye et al. (2009) with kind permission from Elsevier)

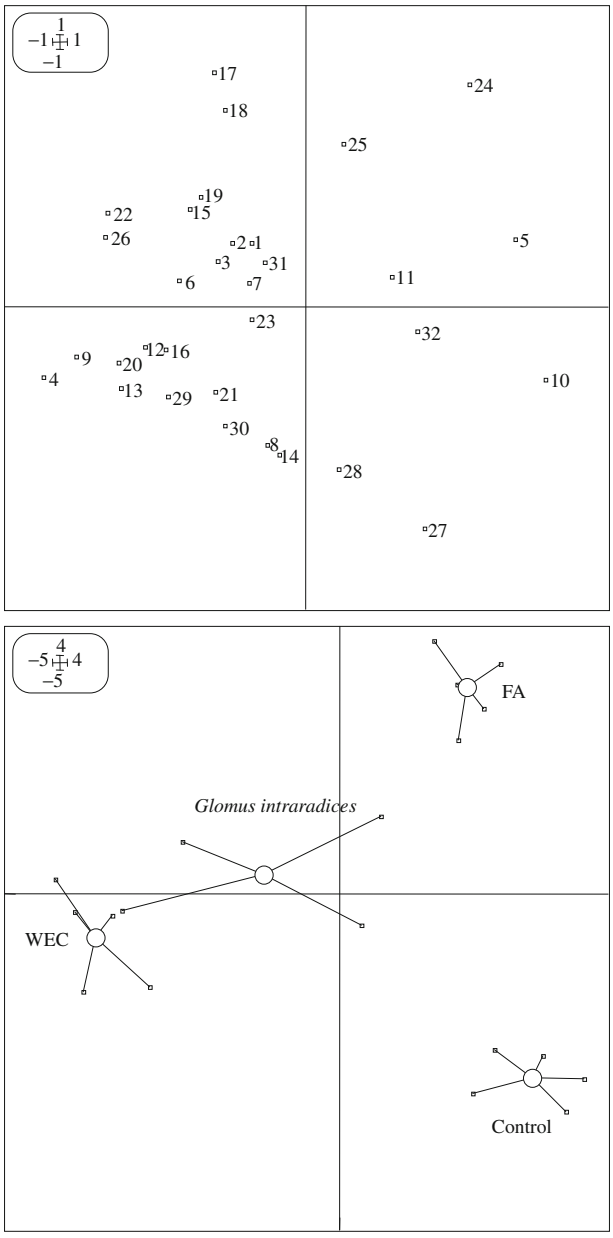


Fig. 2 Between-group analysis (BGA) of the SIR responses with respect to the soil treatments. WEC, without *Eucalyptus camaldulensis* seedlings. FA, preplanting fertilizer application. 1, L-phenylalanine; 2, L-glutamine; 3, L-serine; 4, L-arginine; 5, L-asparagine; 6, L-histidine; 7, L-lysine; 8, L-glutamic acid; 9, L-tyrosine; 10, L-cystein; 11, D-glucose; 12, D-mannose; 13, sucrose; 14, D-glucosamine; 15, *N*-methyl-D-glucamine; 16, succinamide; 17, ascorbic acid; 18, citric acid; 19, fumaric acid; 20, gluconic acid; 21, quinic acid; 22, malonic acid; 23, formic acid; 24, ketoglutaric acid; 25, ketobutyric acid; 26, succinic acid; 27, tartaric acid; 28, uric acid; 29, oxalic acid; 30, gallic acid; 31, malic acid; 32, DL-hydroxy-butyric acid. (Reprinted from Kisa et al. (2007) with kind permission from John Wiley and Sons)

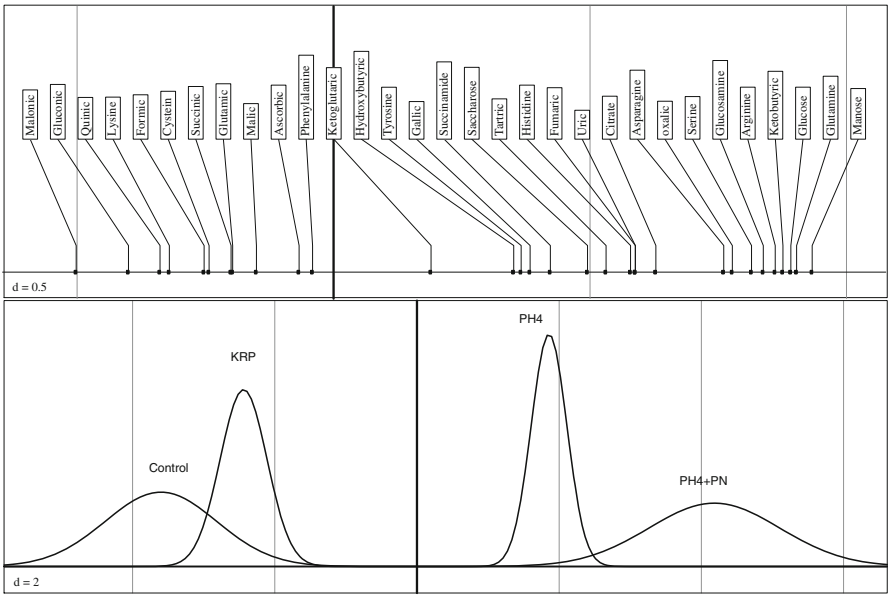


Fig. 3 Graphical display (biplot) of BGA axes showing the Substrate Induced Respirations with respect to the soil treatments. The *upper part* of the figure shows the scores of the 31 substrates on the first BGA axis. The four Gauss curves in the *lower part* of the figure represent the mean and the variance of the scores of the soil samples on the first BGA axis. Control: un-inoculated and un-amended soil, KRP: soil amended with Khouribga Rock Phosphate, PH4: soil inoculated with *Pisolithus* sp. strain PH4, PH4+PN: soil amended with Khouribga Rock Phosphate and inoculated with *Pisolithus* sp. strain PH4. (Reprinted from Ouahmane et al. (2009) with kind permission from Springer Science+Business Media)

3.1.3 BGA graphics

The aim of graphical displays after a BGA is to underline the differences between groups when these differences are statistically significant. Three examples of BGA graphics are presented here: convex hulls (Fig. 1) (Faye et al. 2009), stars (Fig. 2) (Kisa et al. 2007), and Gauss curves (Fig. 3) (Ouahmane et al. 2009).

In the first example (Fig. 1), Faye et al. (2009) use BGA to show that the biomass increase of *Faidherbia albida* seedlings is positively linked to the inoculation of *Bradyrhizobia spp*. Furthermore, this effect varies according to the origin of *Bradyrhizobia* isolates. *Bradyrhizobia* strains were isolated from a controlled mycorrhization experiment with an exotic Acacia species (*A. holosericea*) and an ectomycorrhizal fungus, *Pisolithus albus* IR100. This plantation was located in Senegal. Three origins of isolates were compared, and four variables were measured on *F. albida* seedlings: shoot and root biomass (SB and RB) and total number and dry weight of nodules (TN and DW). The three isolate origins were:

1. Bacterial strains isolated from the soil of a plantation of *A. holosericea* previously inoculated with the ectomycorrhizal fungus *P. albus* IR100 (IR100S in Fig. 1)
2. Bacterial strains isolated from the soil of a plantation of *A. holosericea* uninoculated with the ectomycorrhizal fungus (NIS in Fig. 1)

3. Bacterial strains isolated from the soil of the *F. albida* parkland surrounding the *A. holosericea* plantation (PS in Fig. 1).

On Fig. 1, the three origins were represented with convex hulls surrounding the corresponding samples. A multivariate permutation test showed that the differences were statistically significant ($p < 0.01$), and the use of convex hulls on Fig. 1 helped underline these differences. Faye et al. (2009) concluded that exotic plant species introduction (*A. holosericea* is an Australian Acacia) can drastically affect the structure and symbiotic effectiveness of native *Bradyrhizobia* populations and noted that this could limit the natural regeneration of native (Sahelian) plant species such as *F. albida*.

In the second example (Fig. 2), Kisa et al. (2007) use BGA to show that the functional diversity of soil microbial communities (measured by ISCP) is altered by the exotic tree species *Eucalyptus camaldulensis*, and that the inoculation of an arbuscular mycorrhizal fungus (*Glomus intraradices*) can counterbalance this negative influence. Figure 2 shows substrate induced respiration (SIR) substrates (top) and the position of soil samples on which SIR profiles were measured (bottom). The five pointed irregular stars on this figure show the five experimental repeats and their mean position (circle at the center of the star). The permutations test of BGA confirmed that the difference between the four groups was highly significant ($p < 0.001$). The effect of *Eucalyptus camaldulensis* on bacterial functional diversity (difference between WEC and FA), and the influence of *Glomus intraradices* inoculation, are indeed very clear. Kisa et al. (2007) conclude that arbuscular mycorrhizal symbiosis with *Glomus intraradices* can counterbalance the negative influence of exotic tree species on soil microbial communities.

In Fig. 3, Ouahmane et al. (2009) shows a third example of BGA graphical display, with only one principal axis. The aim of this analysis was to show that the inoculation of *Pinus halepensis* with the ectomycorrhizal fungus *Pisolithus* sp. strain PH4 had a strong effect on soil microbial functional diversity and on rock phosphate (Khodjari Rock Phosphate, KRP) solubilisation. The first axis of BGA very clearly shows the effect of PH4 inoculation on functional diversity (ISCP profiles), so using the second axis to draw a factor map was not appropriate. In the upper part of the graph, substrate labels are ordered according to the substrate score on the first BGA axis. In the lower part, Gauss curves are adjusted to the parameters (mean and standard deviation) of sample scores in each treatment. The mean and standard deviation of the fives samples belonging to each of the four treatments (Control, KRP, PH4, PH4+PN) are computed and the corresponding Gauss curves are drawn. This presentation shows, for each treatment, the optimal substrates (position of Gauss curves) and the functional diversity (Gauss curve width). The permutation test showed that the difference between treatments was highly significant ($p < 0.001$).

3.2 Co-inertia analysis

There are many methods to analyze the relationships between two data tables. In ecology, these methods play a major role because they can be used to analyse the relationships between species distribution and environmental variables. These methods are applied to a species data table, containing the number of individuals of various species

(columns) found in a series of places (rows), and an environmental data table, containing the values of environmental variables (columns) measured at the same places (rows). The statistical significance of the relationships between the two tables can be tested by a permutation test, using a criterion that depends on the particular method. In coinertia analysis, this permutation test is based on the total coinertia criterion (i.e., the sum of the squared cross-covariances between the variables of the two tables, see Sect. 3.2.1).

3.2.1 CoIA and other methods

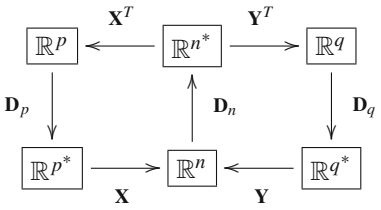
Dray et al. (2003) gives a detailed mathematical description of CoIA and compares it with several methods, particularly canonical correlation analysis (CANCOR), CCA, and RDA. They note that CoIA is the only method that has no constraint on the number of samples compared to the number of species or environmental variables. They also underline the problems occurring when the number of samples is low, or when "explanatory variables" are correlated. For example, if the number of samples is lower than the number of environmental variables, then CCA is equivalent to a simple CA, and the potential relationship with environmental variables disappears. In the same way, RDA is reduced to a simple PCA, and the relationship with the environment is lost.

The main difference between CoIA and constrained methods like CCA and RDA is the difference between the "descriptive strategy" and the "predictive strategy" (Thioulouse 2011). The aim of the first strategy is an objective description of the data set and of the relationships between its components. The second approach is orientated toward the prediction of "explained" (or "dependent") variables by "explanatory" (or "independent") ones. This distinction implies an asymmetry of predictive methods and a symmetry of descriptive ones. This difference also implies computational constraints: predictive methods have a matrix inversion step that is not present in descriptive methods. This matrix inversion step has negative consequences on the data sets that can be analyzed: it means that "explanatory" variables must be independent (in the statistical sense), because the rank of their correlation matrix must not be less than its dimension. This also implies that the number of cases (samples) must not be less than the number of explanatory variables.

In the same way as BGA can be viewed as the analysis of the table of group means in the case of quantitative variables, CoIA can be seen as the analysis of the table of cross-covariances between the variables of the two tables. The number of rows and columns of this crossed table is equal to the number of variables of the two tables. The sample scores are computed by projection of the rows of the two tables as additional elements into the vector space defined by this analysis. This means that CoIA provides two sets of sample scores (one for each of the two initial tables). A description of CoIA in terms of duality diagram is given by Thioulouse (2011), and we give here a short summary of this presentation.

Let $\mathbf{X}$ be the first table (environment variables table), with n rows (sampling sites) and p columns (variables), and let $\mathbf{Y}$ be the second table (species data), with the same n rows, and q columns (species). $\mathbf{X}^T$ and $\mathbf{Y}^T$ are the transpose of $\mathbf{X}$ and $\mathbf{Y}$. Let $\mathbf{D}_n$ be the diagonal matrix (n x n) of site weights: $\mathbf{D}_n = diag(w_1, \ldots, w_n)$, and let $\mathbf{D}_p$

and $\mathbf{D}_q$ be two diagonal metrics on $\mathbb{R}^p$ and $\mathbb{R}^q$ respectively. Co-Inertia Analysis is defined by its "coupled diagram", that shows that CoIA is the eigenanalysis of matrix $\mathbf{X}^T\mathbf{D}_n\mathbf{Y}\mathbf{D}_q\mathbf{Y}^T\mathbf{D}_n\mathbf{X}\mathbf{D}_p$:



The triplet notation of this diagram is $(\mathbf{Y}^T\mathbf{D}_n\mathbf{X}, \mathbf{D}_p, \mathbf{D}_q)$. If the columns of both tables are centered, then the total inertia of each table is simply a sum of variances: $I_{\mathbf{X}} = trace(\mathbf{X}\mathbf{D}_p\mathbf{X}^T\mathbf{D}_n)$ and $I_{\mathbf{Y}} = trace(\mathbf{Y}\mathbf{D}_q\mathbf{Y}^T\mathbf{D}_n)$. And the co-inertia between $\mathbf{X}$ and $\mathbf{Y}$ is in this case a sum of squared covariances:

$$\mathrm{CoI}_{\mathbf{XY}} = trace(\mathbf{X}\mathbf{D}_p\mathbf{X}^T\mathbf{D}_n\mathbf{Y}\mathbf{D}_q\mathbf{Y}^T\mathbf{D}_n)$$

It is this criterion that is used in the permutation test to check for the statistical significance of the relationship between the two tables.

Co-inertia analysis is also related to partial least squares (PLS) regression: the first step of PLS regression is equivalent to the first CoIA axis. It is similar to WA-PLS (ter Braak and Juggins 1993) and has the same advantages, allowing the use of any number of variables without having to select some of them through questionable methods like stepwise regression with forward and/or backward selection.

3.2.2 CoIA examples

Seven of the 16 articles used CoIA (see Table 2). In six analyses, it was used to study the relationship between microbial functional diversity (ISCP data) and another type of data, and in one analysis (Diallo et al. 2006), it was used to study the effects of native and exotic plant residues (leaf powders) on plant growth and microbial communities. CoIA was used to analyze the relationship between the microbial functional diversity, the plant growth and the mycorrhizal symbiosis establishment under different soil treatments (mycorrhizal inoculation and phosphorus amendment) (Ouahmane et al. 2007). It was also used in a study on the relationship between microbial functional diversity, Sorghum growth and Striga development with or without termite mount powder amendment (Andrianjaka et al. 2007). Another example is given by Ouahmane et al. (2006a) in a study focused on the effect of nurse plants (*Lavandula species*) on *Cupressus arizonica* growth, and soil microbial functional diversity. Duponnois et al. (2005a,b, 2006b) use CoIA to analyze the relationships between soil microbial functional diversity, plant growth and mycorrhizal variables in various conditions of termite mound powder amendment, rock phosphate amendment, and inoculation with an ectomycorrhizal fungus, *Scleroderma dictyosporum*. The last

example is the use of CoIA to analyze the effect of rock phosphate amendment and *Glomus intraradices* inoculation on the relationships between microbial functional diversity and plant growth and some soil microbial characteristics (Duponnois et al. 2005b).

In most of these examples, the number of samples is low: these are field data and each sample represents a lot of time and work. While many data analysis methods could not be applied on such data sets, CoIA allows researchers to analyze them and to test the significance of observed structures.

3.2.3 CoIA graphics

The aim of graphical display in CoIA is to reveal the relationships between the two data tables. CoIA provides four sets of coordinates: one set for the rows and one set for the columns of the two tables. Figure 4 is taken from Ouahmane et al. (2007) and shows an example of the four graphics that can be drawn with these four sets of coordinates. In this example, the objective of the authors was to show the difference between native (AM) and allochtonous (GI) arbuscular mycorrhizal fungi inoculation on soil bacterial functional diversity and on rock phosphate alteration. They used CoIA to analyze the relationship between a table of SIR profiles and a table of plant variables. In the SIR profiles table, the columns correspond to the 28 substrates, and the rows to the 18 soil samples. In the plant variables table, the columns are *Cupressus atlantica* seedling height (H), shoot and root biomass (SB, RB), leaf P content (P), and mycorrhizal colonization (MC). The rows correspond to the same 18 soils as in the first table. The CoIA permutation test showed that the relationship between these two tables was highly significant.

The first graph (Fig. 4a) is the factor map of plant variables. The five variables are all oriented toward the left of the graph. This is a "size effect", meaning that the left side of the graph corresponds to samples where plant growth is better, while conversely the right side corresponds to a lesser growth of *Cupressus atlantica* seedlings. On Fig. 4c, we can see that this better growth is correlated with the inoculation of native arbuscular mycorrhizal fungi (CAM), and that this effect is even stronger when rock phosphate amendment is done (CAMP). Inoculation with the allochtonous fungi *Glomus intraradices* (GI) is also correlated with a better plant growth, but there is no additional rock phosphate amendment effect (GIP).

On Fig. 4b, d, we can see that plant growth is also linked to the functional diversity of the soil microbial community. The SIR substrates located on the left side of Fig. 4b (particularly organic acids) correspond to a better plant growth, and are correlated with the inoculation of native arbuscular mycorrhizal fungi (CAM), alone or combined with rock phosphate amendment (CAMP). The effect of *Glomus intraradices* inoculation (GI) alone or combined with rock phosphate amendment (GIP) is also positive on plant growth, but clearly less than the effect of native arbuscular mycorrhizal fungi. Rock phosphate amendment alone (CP) is also positive, but less than in combination with fungi inoculation.

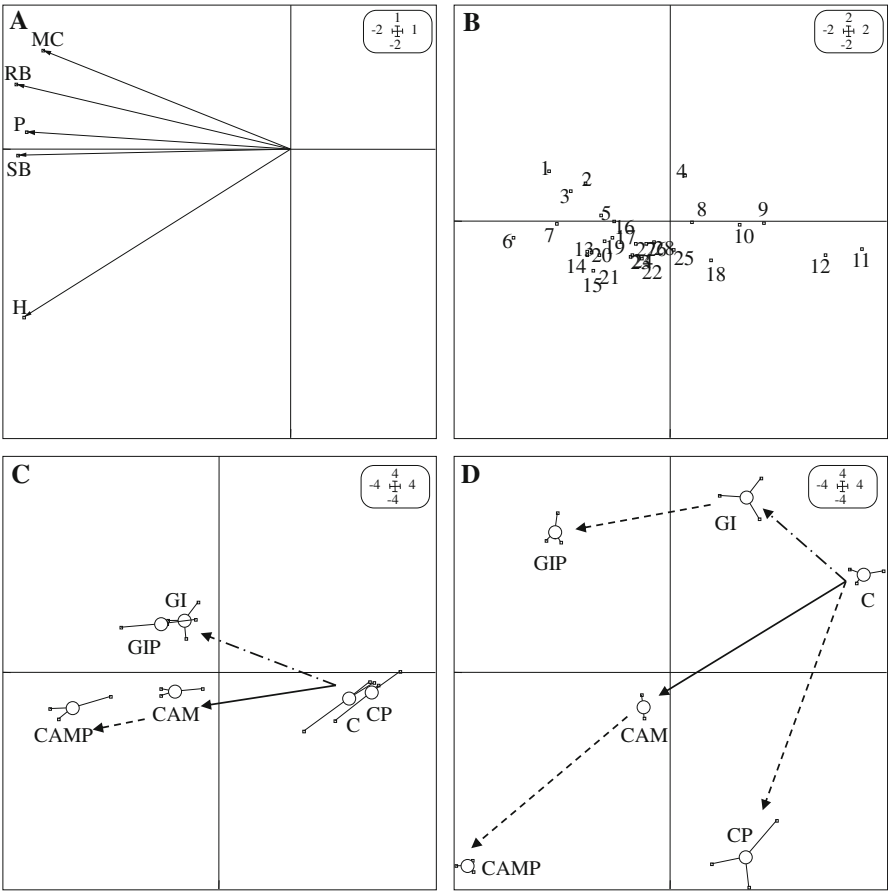


Fig. 4 Co-inertia analysis of the SIR responses of the soils inoculated with *Glomus intraradices* or the mixture of native arbuscular mycorrhizal fungi and/or Khouribga Rock Phosphate, plant growth, phosphorus leaf content and mycorrhizal colonization. **a** Factor map of plant growth and mycorrhizal colonization variables (H, height; SB, shoot biomass; RB, root biomass; P, leaf P content; MC, mycorrhizal colonization). **b** Factor map of SIR responses (D-mannose, 1; L-serine, 2; L- histidine, 3; L-tyrosine, 4; gluconic acid, 5; uric acid, 6; L-lysine, 7; L-glutamic acid, 8; sucrose, 9; succinamide, 10; cyclohexane, 11; L-glutamine, 12; citric acid, 13; ketobutyric acid, 14; tartaric acid, 15; DL-hydroxybutyric acid, 16; *N*-methyl-D-glucosamine, 17; D-glucose, 18; quinic acid, 19; L-asparagine, 20; succinic acid, 21; malic acid, 22; oxalic acid, 23; fumaric acid, 24; ascorbic acid, 25; malonic acid, 26; ketoglutaric acid, 27; L-arginine, 28). **c** Factor map of plant growth and mycorrhizal colonization (C, control (not inoculated); CP, Khouribga Rock Phosphate amendment; GI, *Glomus intraradices* inoculation; GIP, *Glomus intraradices* inoculation and Khouribga Rock Phosphate amendment; CAM, mixture of native arbuscular mycorrhizal fungi inoculation; CAMP, CAM inoculation and Khouribga Rock Phosphate amendment). **d** Factor map of SIR responses soil samples (for the legend, see **c**). (Reprinted from Ouahmane et al. (2007) with kind permission from Elsevier)

Authors concluded that the use of native arbuscular mycorrhizal fungi and their selective effect on soil microflora have to be considered in order to optimize the sustainable re-establishment of plant species in a degraded soil.

4 Discussion and conclusion

The results reported in these articles show that: (i) soil biofunctioning is driven by a multitude of microbiological components and biochemical pathways, (ii) the mycorrhizal symbiosis plays a key role in the complexity of microbial life in soil and (iii) it is necessary to use appropriate statistical tools to assess the patterns of soil microbiological and biochemical indicator aggregation for evaluating soil quality. Patterns of indicator aggregation are used to describe the state of a given ecosystem by integrating and summarizing the information contained in a larger set of indicators. It has been hypothesized that decreases in the diversity of soil organisms will lead to a lower resistance of soils to stress or disturbance (Brussaard et al. 1997). The quantification of the impacts of various treatments on soil quality and more particularly on soil microbial functional diversity, is therefore of great relevance in the cultural strategies required in conservation programs.

This paper has pointed out some original multivariate analysis techniques that offer opportunities to analyze soil data and to determine interactions between plants, microbial communities, mycorrhizal fungi and physico-chemical environmental variables. Data reviewed from the 16 articles have outlined the fundamental role of mycorrhizal symbiosis in soil biofunctioning and the importance of interactions with other variables such as soil chemical characteristics, amendments (for example rock phosphate amendment) or the composition of plant cover. All these studies have been realized in different environmental conditions, with different biological models and with replicates for each treatment.

In the present paper, most of the reviewed studies were based on examining patterns of ISCP. Numerous studies have outlighted the importance of the functional diversity of soil microbial communities for the sustained functioning of terrestrial ecosystems (Degens et al. 2000). The functional diversity of microbial communities includes a wide range of activities including decomposition, nutrient transformations, plant growth promotion or suppression and various soil physical processes. The ISCP assessment provides a more realistic measure compared with other methods since it reveals a direct measurement of substrate catabolism by soil microflora without prior culturing of microbes usually necessary in the culture-based methods. The main objective of this technical approach is to give an indicator of the soil microflora capacity to mobilize some nutrients (i.e. P and N) from the organic matter or minerals and to keep a level of soil fertility required for the productivity and stability of an ecosystem. Among soils subjected to different cultural practices or different plant covers (more or less degraded), ISCP patterns will differ according to the functional diversity of microbial communities. Hence ISCP measurement will give informations on the soil quality and the resistance of a soil to stress and disturbance (Remigi et al. 2008) and consequently on the resilience capacity of an ecosystem.

It is well known that ecological stability (resistance and resilience) of a soil system is a key factor influencing ecosystem properties and processes (Orwin and Wardle 2004). Resistance is usually considered as the amount of change caused by a disturbance and resilience as the speed with which an ecosystem returns to its pre-disturbance level following a disturbance (Pimm 1984). In this context, mycorrhizal development has a great role in the stability of soil ecosystems. Numerous indices on both resistance

and resilience have been proposed in the literature to quantify soil ecosystem stability (Lavorel 1999; Orwin and Wardle 2004), but they are not easy to use and interpret. It is necessary to have indices that provide a relative quantitative measure of resilience and resistance of a response variable in all possible scenarios to compare the stability of different systems. Most indices currently in use are not able to do this or are difficult to interpret particularly because of lack of standardization in experimental conditions.

All these approaches have to be compared and the variables hierarchised in order to provide robust indices measuring the level of an ecosystem stability that can be used in a wide range of environmental situations. Using the synthetic capacity of factor scores computed by multivariate analyses is probably a good choice in this area. With these statistical tools, different environmental conditions and different studies can be quantitatively compared more easily than with other univariate statistical methods. However, real operational indicators based on this approach are still to be developped. The idea here is to synthetize the information from these experimental studies in order to identify the cultural practices that are able to enhance some biological indicators potentially involved in the resistance and stability of the ecosystems. The main objective is to highlight or underline one or several indicators which contribute the most to the synthetic measures of stability.

Since a lot of biological and chemical variables can be analysed together and as links can be identified between variables, this approach can provide useful information from different environmental conditions and contribute to a generalization of the effect of a cultural practice (i.e. controlled mycorrhization, soil mycorrhizal potential management) on the stability of the ecosystem.

Lastly, the availability of appropriate software is also a key component of this approach. In this area, the R software offers a wide range of statistical methods, and the ade4 package includes many multivariate data analysis techniques in addition to PCA, BGA and CoIA. Other R packages are useful for ecological data analysis, particularly the vegan package (Oksanen et al. 2010), but the ade4 package has the particularity of being based on the duality diagram. This means that it proposes a synthetic and coherent theoretical framework for all the multivariate analysis methods. This is described in the forthcoming articles presented in de la Cruz and Holmes (2011). For example, Thioulouse (2011) present several k-table analysis methods like STATICO and COSTATIS, and a simple generalization of BGA and CoIA, named between-group coinertia analysis (BGCoIA) that could be particularly interesting for the study of the resilience capacity of an ecosystem after a stress or disturbance of soil variables.

Acknowledgments Authors are extremely grateful to the reviewer for the countless comments and very relevant suggestions that greatly helped improve the successive versions of this paper.

References

Andrianjaka Z, Bally R, Lepage M, Thioulouse J, Comte G, Kisa M, Duponnois R (2007) Biological control of *Striga hermonthica* by *Cubitermes* termite mound powder amendment in sorghum culture. Appl Soil Ecol 37:175–183

Azcon-Aguilar C, Palenzuela J, Roldan A, Bautista S, Vallejo R, Barea JM (2003) Analysis of the mycorrhizal potential in the rhizosphere of representative plant species from desertification-threatened Mediterra- nean shrublands. Appl Soil Ecol 14:165–175

Barrios E, Delve R, Bejunda M, Mowo J, Agunda J, Ramisch J, Trejo M, Thomas R (2006) Indicators of soil quality: a south-south development of a methodological guide for linking local and technical knowledge. Geoderma 135:248–259

Brundrett M (2002) Coevolution of roots and mycorrhizas of land plants. New Phytol 154:275–304

Brussaard L, Behanpelletier VM, Bignell DE, Brown VK, Didden W, Folgarait P, Fragoso C, Freckman DW, Gupta VVSR, Hattori T, Hawksworth DL, Klopatek C, Lavelle P, Malloch DW, Rusek J, Soderstrom B, Tiedje JM, Virginia RA (1997) Biodiversity and ecosystem functioning in soil. Ambio 26:563–570

Chessel D, Dufour AB, Thioulouse J (2004) The ade4 package-I- One-table methods. R News 4:5–10

Culhane A, Perriere G, Considine E, Cotter T, Higgins D (2002) Between-group analysis of microarray data. Bioinformatics 18(12):1600–1608

Dabire AP, Hien V, Kisa M, Bilgo A, Sangare KS, Plenchette C, Galiana A, Prin Y, Duponnois R (2000) Responses of soil microbial catabolic diversity to arbuscular mycorrhizal inoculation and soil disinfection. Mycorrhiza 17:537–545

Degens BP, Harris JA (1997) Development of a physiological approach to measuring the catabolic diversity of soil microbial communities. Soil Biol Biochem 29(9–10):1309–1320

Degens BP, Schipper LA, Sparling GP, Vojvodic-Vukovic M (2000) Decreases in organic C reserves in soils can reduce the catabolic diversity of soil microbial communities. Soil Biol Biochem 32:189–196

Degens BP, Schipper LA, Sparling GP, Duncan LC (2001) Is the microbial community in a soil with reduced catabolic diversity less resistant to stress or disturbance? Soil Biol Biochem 33(9):1143–1153

de la Cruz O, Holmes S (2011) The duality diagram in data analysis: examples of modern applications. Ann Appl Stat 5:2266–2277

Diallo MD, Duponnois R, Guisse A, Sall S, Chotte JL, Thioulouse J (2006) Biological effects of native and exotic plant residues on plant growth, microbial biomass and n availability under controlled conditions. Eur J Soil Biol 42(4):238–246

Doledec S, Chessel D (1987) Seasonal successions and spatial variables in fresh-water environments .1. description of a complete 2-way layout by projection of variables. Acta Oecologica-Oecologia Generalis 8(3):403–426

Doledec S, Chessel D (1994) Co-inertia analysis—an alternative method for studying species environment relationships. Freshw Biol 31(3):277–294

Doran J, Parkin T (1994) Defining and assessing soil quality. In: Doran J, Coleman D, Bezdicek D, Stewart B (eds) Defining soil quality for a sustainable environment, vol 35. Soil Science Society of America, Madison pp 3–21

Dray S, Dufour AB (2007) The ade4 package: implementing the duality diagram for ecologists. J Stat Softw 22(4):1–20

Dray S, Chessel D, Thioulouse J (2003) Co-inertia analysis and the linking of ecological data tables. Ecology 84(11):3078–3089

Dumanski J, Pieri C (2000) Land quality indicators: research plan. Agric Ecosyst Environ 81:93–102

Duponnois R, Plenchette C, Ba AM (2001) Growth stimulation of seventeen fallow leguminous plants inoculated with *Glomus aggregatum* in Senegal. Eur J Soil Biol 124:181–186

Duponnois R, Paugy M, Thioulouse J, Masse D, Lepage M (2005a) Functional diversity of soil microbial community, rock phosphate dissolution and growth of *Acacia seyal* as influenced by grass-, litter- and soil-feeding termite nest structure amendments. Geoderma 124(3–4):349–361

Duponnois R, Colombet A, Hien V, Thioulouse J (2005b) The mycorrhizal fungus *Glomus intraradices* and rock phosphate amendment influence plant growth and microbial activity in the rhizosphere of *Acacia holosericea*. Soil Biol Biochem 37:1460–1468

Duponnois R, Kisa M, Assigbetse K, Prin Y, Thioulouse J, Issartel M, Moulin P, Lepage M (2006a) Fluorescent pseudomonads occuring in *Macrotermes subhyalinus* mound structures decrease cd toxicity and improve its accumulation in sorghum plants. Sci Total Environ 370(2–3):391–400

Duponnois R, Assigbetse K, Ramanankierana H, Kisa M, Thioulouse J, Lepage M (2006b) Litter-forager termite mounds enhance the ectomycorrhizal symbiosis between *Acacia holosericea* A. Cunn. Ex G. Don and *Scleroderma dictyosporum* isolates. FEMS Microbiol Ecol 56:292–303

Duponnois R, Hafidi M, Thioulouse J, Galiana A, Ouahmane L, Dreyfus B, Prin Y (2009) Monitoring the development of nurse plant species to improve the performances of reforestation programs in mediterranean areas. In: Khan MS, Zaidi A, Musarrat J (eds) Microbial strategies for crop improvement. Springer, Berlin

Escoufier Y (1987) The duality diagramm: a means of better practical applications. In: Legendre P, Legendre L (eds) Development in numerical ecology, NATO advanced Institute, Serie G. Springer, Berlin pp 139–156

Faye A, Krasova-Wade T, Thiao M, Thioulouse J, Neyra M, Prin Y, Galiana A, Ndoye I, Dreyfus B, Duponnois R (2009) Controlled ectomycorrhization of an exotic legume tree species *Acacia holosericea* affects the structure of root nodule bacteria community and their symbiotic effectiveness on *Faidherbia albida*, a native sahelian acacia. Soil Biol Biochem 41(6):1245–1252

Frey-Klett P, Chavatte M, Clausse ML, Courrier S, Roux CL, Raaijmakers J, Martinotti MG, Pierrat JC, Garbaye J (2005) Ectomycorrhizal symbiosis affects functional diversity of rhizosphere fluorescent pseudomonads. New Phytol 165(1):317–328

Giller KE, Beare MH, Lavelle P, Izac A-MN, Swift MJ (1997) Agricultural intensification, soil biodiversity and agrosystme function. Appl Soil Ecol 6:3–16

Heijden MGAvan der , Klironomos JN, Ursic M, Moutoglis P, Streitwolf-Engel R, Boller T, Wiemken A, Sanders IA (1998) Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. Nature 396:69–72

Hill MO, Smith AJE (1976) Principal component analysis of taxonomic data with multi-state discrete characters. Taxon 25:249–255

Holmes S (2006) Multivariate analysis: the French way. In: Nolan D, Speed T (eds) Festschrift for David Freedman. IMS, Beachwood pp 1–14

Huberty CJ (1994) Applied discriminant analysis. Wiley, New York

Johansson JF, Paul LR, Finlay RD (2004) Microbial interactions in the mycorrhizosphere and their significance for sustainable agriculture. FEMS Microbiol Ecol 48(1):1–13

Kiers HAL (1991) Simple structure in component analysis techniques for mixtures of qualitative and quantitative variables. Psychometrika 56(2):197–212

Kisa M, Sanon A, Thioulouse J, Assigbetse K, Sylla S, Spichiger R, Dieng L, Berthelin J, Prin Y, Galiana A, Lepage M, Duponnois R (2007) Arbuscular mycorrhizal symbiosis can counterbalance the negative influence of the exotic tree species *Eucalyptus camaldulensis* on the structure and functioning of soil microbial communities in a sahelian soil. FEMS Microbiol Ecol 62(1):32–44

Lavorel S (1999) Ecological diversity and resilience of Mediterranean vegetation to disturbance. Divers Distrib 5:3–13

LeRoux B, Rouanet H (2004) Geometric data analysis. Kluwer, Dordrecht

Linderman RG (1988) Mycorrhizal interactions with the rhizosphere microflora—the mycorrhizosphere effect. Phytopathology 78(3):366–371

Nakatsu CH, Torsvik V, Ovreas L (2000) Soil community analysis using dgge of 16s rdna polymerase chain reaction products. Soil Sci Soc Am J 64(4):1382–1388

Oksanen J, Blanchet FG, Kindt R, Legendre P, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H (2010) vegan: community ecology package. URL http://CRAN.R-project.org/package=vegan, r package version 1.17-4

Orwin KH, Wardle DA (2004) New indices for quantifying the resistance and resilience of soil biota to exogenous disturbances. Soil Biol Biochem 36(11):1907–1912

Ouahmane L, Hafidi M, Plenchette C, Kisa M, Bournezzough A, Thioulouse J, Duponnois R (2006a) *Lavandula* species as accompanying plants in *Cupressus* replanting strategies: Effect on plant growth, mycorrhizal soil infectivity and soil microbial catabolic diversity. Appl Soil Ecol 34(2–3):190–199

Ouahmane L, Duponnois R, Hafidi M, Kisa M, Boumezouch A, Thioulouse J, Plenchette C (2006b) Some Mediterranean plant species (*Lavandula* spp. and *Thymus satureioides*) act as potential "plant nurses" for the early growth of *Cupressus atlantica*. Plant Eco 185:123–134

Ouahmane L, Thioulouse J, Hafidi M, Prin Y, Ducousso M, Galiana A, Plenchette C, Kisa M, Duponnois R (2007) Soil functional diversity and P solubilization from rock phosphate after inoculation with native or allochtonous arbuscular mycorrhizal fungi. For Ecol Manag 241(1–3):200–208

Ouahmane L, Revel JC, Hafidi M, Thioulouse J, Prin Y, Galiana A, Dreyfus B, Duponnois R (2009) Responses of *Pinus halepensis* growth, soil microbial catabolic functions and phosphate-solubilizing bacteria after rock phosphate amendment and ectomycorrhizal inoculation. Plant Soil 320(1–2):169–179

Pimm SL (1984) The complexity and stability of ecosystems. Nature 307(5949):321–326

R Development Core Team (2010) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, URL http://www.R-project.org, ISBN 3-900051-07-0

Ramanankierana N, Rakotoarimanga N, Thioulouse J, Kisa M, Randrianjohany E, Ramaroson L, Duponnois R (2006) The ectomycorrhizosphere effect influences functional diversity of soil microflora. Int J Soil Sci 1:8–19

Ramanankierana N, Ducousso M, Rakotoarimanga N, Prin Y, Thioulouse J, Randrianjohany E, Ramaroson L, Kisa M, Galiana A, Duponnois R (2007) Arbuscular mycorrhizas and ectomycorrhizas of *Uapaca bojeri* l. (euphorbiaceae): sporophore diversity, patterns of root colonization, and effects on seedling growth and soil microbial catabolic diversity. Mycorrhiza 17(3):195–208

Ramette A (2007) Multivariate analyses in microbial ecology. FEMS Microbiol Ecol 62:142–160

Remigi P, Faye A, Kane A, Deruaz M, Thioulouse J, Cissoko M, Prin Y, Galiana A, Dreyfus B, Duponnois R (2008) The exotic legume tree species *Acacia holosericea* alters microbial soil functionalities and the structure of the arbuscular mycorrhizal community. Appl Environ Microbiol 74(5):1485–1493

Schreiner RP, Mihara KL, McDaniel H, Bethlenfalvay GJ (1997) Mycorrhizal fungi influence plant and soil functions and interactions. Plant Soil 188(2):199–209

Smith SE, Read DJ (2008) Mycorrhizal symbiosis. Academic Press, London

ter Braak CJF, Juggins S (1993) Weighted averaging partial least squares regression (wa-pls): an improved method for reconstructing environmental variables from species assemblages. Hydrobiologia 269/270:485–502

Thioulouse J (2011) Simultaneous analysis of a sequence of paired ecological tables: a comparison of several methods. Ann Appl Stat 5:2300–2325

Venables WN, Ripley BD (2002) Modern applied statistics with S. Springer, Berlin

Wirsel SGR (2004) Homogenous stands of wetland grass harbour diverse consortia of arbuscular mycorrhizal fungi. FEMS Microbiol Ecol 48:129–138

Zelles L (1999) Fatty acid patterns of phospholipids and lipopolysaccharides in the characterisation of microbial communities in soil: a review. Biol Fertil Soils 29(2):111–129

Author Biographies

Jean Thioulouse is Senior Researcher in Biostatistics and Bioinformatics at the French CNRS. He is working in the Biometry and Evolutionary Biology department of the University of Lyon, France. His research activity is focused on the development of statistical methods and software applied to the analysis of environmental data.

Yves Prin is Senior Researcher in the French CIRAD in Montpellier. In the Laboratoire des Symbioses Tropicales et Méditerranéennes, he is conducting researches on the evolution and ecological functioning of both nitrogen-fixing and mycorrhizal symbioses within tropical and mediterranean ecosystems.

Robin Duponnois is Senior Researcher in Microbial Ecology at the French IRD. His research area is focussed on the ecology of mycorrhizal symbiosis, with emphasis on microbial interactions in the mycorrhizosphere of forest tree species in tropical and Mediterranean environment.