Mapping the imprint of biotic interactions on β-diversity

Marc Ohlmann, Florent Mazel, Loïc Chalmandrier, Stéphane Bec, Eric Coissac, Ludovic Gielly, Johan Pansu, Vincent Schilling, Pierre Taberlet, Lucie Zinger, et al.

Mapping the imprint of biotic interactions on β-diversity

Marc Ohlmann[1], Florent Mazel[2], Loïc Chalmandrier[3,4], Stéphane Bec[1], Eric Coissac[1], Ludovic Gielly[1], Johan Pansu[5], Vincent Schilling[6], Pierre Taberlet[1], Lucie Zinger[7], Jérome Chave[6], Wilfried Thuiller[1]

[1]Univ. Grenoble Alpes, CNRS, Univ. Savoie Mont Blanc, CNRS, LECA, Laboratoire d'Écologie Alpine, F-38000 Grenoble, France

[2]Department of Botany and Biodiversity Research Centre, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

[3]Landscape Ecology, Institute of Terrestrial Ecosystems, ETH Zürich, Zürich, Switzerland.

[4]Swiss Federal Research Institute WSL, 8903 Birmensdorf, Switzerland.

[5] Princeton University, 110 Morrison Hall, Princeton, NJ 08544, United States

[6]Université Toulouse 3 Paul Sabatier, CNRS, IRD, UMR 5174 Evolution et Diversité Biologique (EDB), F-31062 Toulouse, France.

[7]Ecole Normale Supérieure, PSL Research University, CNRS, Inserm, Institut de Biologie de l'Ecole Normale Supérieure (IBENS), F-75005 Paris, France.

Email addresses: Marc Ohlmann (marc.ohlmann@univ-grenoble-alpes.fr), Florent Mazel(flo.mazel@gmail.com), Loïc Chalmandrier(loic.chalmandrier@usys.ethz.ch), Stéphane Bec (stephane.bec@univ-grenoble-alpes.fr), Eric Coissac (eric.coissac@gmail.com), Ludovic Gielly (ludovic.gielly@univ-grenoble-alpes.fr), Johan Pansu (johan.pansu@gmail.com), Vincent Schilling (vincent.schilling@hotmail.fr), Pierre Taberlet (pierre.taberlet@univ-grenoble-alpes.fr), Lucie Zinger (zinger@biologie.ens.fr), Jérôme Chave (jerome.chave@univ-tlse3.fr), Wilfried Thuiller (wilfried.thuiller@univ-grenoble-alpes.fr)

Corresponding author: Marc Ohlmann, LECA, UMR UGA-USMB-CNRS 5553, Université Grenoble Alpes, CS 40700 38058 Grenoble cedex 9, France, tel : +33 4 76 51 42 78, fax :+33 4 76 51 42 79

Type of article: Letters

Running title: Imprint of biotic interactions on β-diversity

Keywords: *β*-diversity, meta-communities, graphical model, interaction network, partial correlation networks, graphical lasso

Number of words in the abstract: 149
Number of words in the main text: 4981
Number of references: 76
Number of figures, tables: 3 figures in the main text, 13 figures and 6 tables in Supplementary Information

Statement of authorship: WT, FM and MO conceived the ideas with the help of LC. LC, EC, LG, JP, FM, PT, LZ and WT collected the soil and environmental data in the field. LG, JP and PT ran the PCR and DNA extractions. LZ and VS ran the bioinformatic analyses. MO ran all statistical analyses and developed the method together with FM and WT. MO, FM and WT wrote the first version of the paper, and all authors contributed substantially to the revisions.

Data accessibility statement: We confirm that, should the manuscript be accepted, the data supporting the results will be archived on Dryad and the data DOI will be included at the end of the article.

Abstract

Investigating how trophic interactions influence the β-diversity of metacommunities is of paramount importance to understanding the processes shaping biodiversity distribution. Here, we propose a statistical method for inferring the strength of spatial dependencies between pairs of species groups. Using simulated data generated from a model of trophic community assemblage, we showed that this method can recover the global structure of the trophic network from the β-diversity patterns of multiple trophic groups. When applied to multi-trophic communities in the soil along an elevational gradient in the French Alps, we found that fungi make a major contribution to the structuring of β-diversity across trophic groups, that there were strong spatial dependencies between groups known to interact specifically (e.g. plant- symbiotic fungi, bacteria-nematodes), and that the influence of environmental gradient was less significant than expected. Our method should help interpret a wide range of multi-trophic biodiversity patterns and yield new testable hypotheses.

Introduction

Understanding the processes that determine the spatial structure of biodiversity is one of the overarching goals of ecology (Ricklefs 1987). In particular, the study of β-diversity, the change in species (or species groups) identities across sampled locations, sheds light on different ecological, evolutionary, and biogeographic processes (e.g. Graham & Fine 2008; Anderson *et al.* 2011). For a given regional species pool, the processes responsible for β-diversity are usually assumed to be environmental filtering, dispersal limitations, and biotic interactions (HilleRisLambers *et al.* 2012; Meynard *et al.* 2013).

Previous studies have sought to analyze β-diversity by teasing apart the effects of environmental filtering from biotic interactions along environmental gradients on a single group of species (e.g. Peay *et al.* 2016 for fungi; Hanson *et al.* 2012 for bacteria; Mazel *et al.* 2017 for mammals; Chalmandrier *et al.* 2015 for plants). However, biotic interactions *across* groups are also expected to drive the structure and distribution of biodiversity. For example, plant-pollinator, trophic (e.g. prey-predator or plant-decomposers) and host-symbiont (including pathogens, mutualistic and commensal organisms) interactions strongly impact diversity distribution and ecosystem functioning (Brose & Hilleband 2016). These interactions among groups of species, although they are not necessarily species-specific (Walker *et al.* 2011; Peay *et al.* 2015), harbor some degree of specificity due to trait constraints and shared habitat preferences (e.g. Allesina *et al.* 2008; Gonzales-Varo 2016). This interdependence among groups implies that the β-diversity of a single group is likely to be contingent to that of the other groups. Hence, studying how the β-diversity of multiple groups covaries along environmental gradients should help better understand their spatial distribution and uncover their interactions.

Soil ecosystems are ideal for documenting these multi-group β-diversity patterns. Indeed, they provide several examples of coupled biological systems (Wardle *et al.* 2006; Bardgett & Wardle 2010), for example plant-mycorrhiza associations (Smith & Read 2008), direct plant control on fungal communities (Broeckling *et al.* 2008), predator-prey relationships (Hedlund & Öhrn

2000) or plant-decomposer relationships (Hattenschwiler *et al.* 2005). Empirical evidence suggests that these biological couplings do indeed produce spatial dependencies between the β-diversity of the different groups (e.g. plants and fungi or bacterial turnover, Zinger *et al.* 2011; Prober *et al.* 2015; Geremia *et al.*, 2016). However, data on soil meta-communities remain scarce and there are no appropriate statistical tools for exploring the β-diversity patterns in multiple soil groups (i.e. more than two or three) while taking into account environmental variation.

Here we develop a generic method for analysing how the β-diversity of one group depends on the β-diversity of the other groups and on the spatial variation in environmental conditions. This is challenging due to potential confounding effects between different groups, which can generate spurious correlations. Variance partitioning methods aim to disentangle the influence of several explanatory variables on a response variable and they are, in principle, able to deal with confounding effects by providing partial correlations among partitions (Borcard *et al.* 1992; Dray *et al.* 2012). However, the number of partitions grows exponentially with the number of explanatory variables, making the method impractical for studies encompassing a large number of species groups. This is problematic since the amplicon-based DNA analysis of environmental samples (i.e. environmental DNA) holds the promise of more streamlined and consistent all-biodiversity environmental surveys (Taberlet *et al.* 2012; Kress *et al.* 2015) and, consequently, new statistical tools are needed. One possible solution to this problem is to use statistical tools called graphical models (Koller & Friedman 2009), in which each node is a random variable and edges represent conditional dependencies between these random variables. Partial correlation networks represent a particular type of graphical model where the edges of the network are weighted by the partial correlation coefficients (i.e. the correlation between two variables while controlling the others). Here, we propose studying multi-group β-diversity patterns in soil communities using partial correlation networks.

Specifically, we hypothesise that interrelated β-diversity patterns among groups are partly explained by biotic interactions along abiotic gradients, and that these spatial dependencies can be

detected using a method called graphical lasso to infer a partial correlation network (Friedman *et al.* 2007). Using simulated multi-trophic communities in a given, known trophic network, we first test and validate our hypothesis and statistical method. We show that biotic interactions may indeed produce interrelated β-diversity patterns (i.e. non-zero partial correlation coefficients) that can be uncovered using a partial correlation network. We then apply our method to an empirical dataset including bacteria, microeukaryotes, meso/macrofauna, plants and abiotic factors along an elevation gradient in the French Alps. We jointly explore the co-variation between the β-diversity of multiple functional and trophic groups to unravel known and unknown potential biotic interactions, while controlling for the relative role of the abiotic environment and the β-diversity of the other groups.

Material and Methods

The method: applying the Graphical lasso to multi-group β-diversity patterns

From species to relevant groups – Species grouping can be defined by known trophic position (e.g. symbiotic fungi), or by taxonomy (e.g. bacteria) or function. This step is essential and has to be implemented in light of the prior knowledge of the study system (see our case study for an example).

Measuring β-diversity and environmental distances – One possible measure of β-diversity between two local communities, A and B, is the Jaccard dissimilarity index, defined as one minus the ratio of the number of species present in both A and B over the number of species present in either A or B. The Jaccard dissimilarity index equals 0 when A and B share the same species, and 1 when they do not share any. We used the R package ‘vegan’ (Ricklefs 2008) to compute the Jaccard dissimilarity index for multiple species groups independently (see below). Moreover, we partitioned the Jaccard index into the true turnover component and the nested component (Baselga 2010), using the R package 'betapart' (Baselga & Orme 2012). Environmental distances between pairs of local communities were computed using Euclidean distances.

The graphical lasso method – The overarching goal of our approach is to use a network to parsimoniously represent the correlations between the β-diversity of each group as a function of the

143 others and the environmental distances. This is challenging due to potential confounding effects (i.e.
144 collinearity among multiple variables), which can generate spurious correlations. Working with
145 partial correlations (correlation between two variables when knowing the value of the others)
146 instead of marginal correlations (Pearson correlation) avoid these confounding effects.
147 Consequently, a suitable description of the system consists of using a class of models that (1)
148 represent the conditional dependencies between random variables (here the *β*-diversity of multiple
149 species groups) using partial correlations while (2) allowing for a parsimonious representation of
150 the dependencies using a network. The Lasso approach (Tibshirani 1994) has been developed to
151 produce precisely this type of parsimonious set of variables. More recently, its multivariate form,
152 the Graphical lasso (Glasso), has made it possible to represent the partial correlations between
153 multiple variables in a network (here the *β*-diversity of multiple groups and the environmental
154 distances, Friedman *et al.* 2007; Mazumder & Hastie 2011). As a consequence, Glasso should be an
155 appropriate tool for representing the structure of the partial correlations between multi-trophic *β*-
156 diversity patterns.

157 In short, the Glasso uses the empirical variance-covariance matrix *S* (computed from the *β*-
158 diversity values of multiple species groups and the environmental distances) to estimate a partial
159 correlation matrix that quantifies the degree of association between pairs of variables conditional to
160 the other variables. In order to estimate the partial correlation matrix in Glasso, the S matrix is
161 inverted to produce the precision matrix P with the particular constraint that *P* must be sparse (i.e.
162 have many zeros). This is equivalent to the variable selection step in conventional statistical
163 approaches (see Friedman *et al.* 2007 for mathematical details). The partial correlation matrix is
164 then computed from P as follows (Eq. 1):

165 $$cor(y_i, y_j | y_{I \setminus i,j}) = -\frac{p_{i,j}}{p_{i,i} p_{j,j}} \qquad \text{(Eq.1)}$$

166

167 where $cor(y_i, y_j | y_{I \setminus i,j})$ represents the partial correlation between the components i and j of a
168 random variable Y given all the other components, and $p_{i,j}$, $p_{i,i}$, $p_{j,j}$ are the elements of P.

As the precision matrices have been inverted with the constraint in order to ensure sparsity, it follows that the partial correlation matrix is also sparse. This sparse representation of the relationships between the β-diversity of multi-trophic groups and the environmental distances means it can be represented simply and intuitively using a network. In the Glasso, this parsimonious representation (i.e. the number of 0 coefficients in the partial correlation matrix) is not optimally determined but is arbitrary and depends on a coefficient λ. Here, we used the Extended Bayesian Information Criterion (Foygel & Drton 2010) to select optimal λ. We used the R package 'qgraph' to estimate the partial correlation matrix with graphical lasso (Epskamp & Fried 2016).

Representing conditional dependencies between β-diversity using a network representation – We plotted the inferred network using Gephi (Bastian *et al.* 2009) and analysed its properties. In particular, we quantified the connectivity of each group by analysing the degree and the weighted degree for each group in the partial correlation network. The degree represents the number of times the β-diversity of a given group is conditionally dependent on the β-diversity of other groups. If the β-diversity of two groups is conditionally independent (i.e. has a zero partial correlation coefficient), they cannot causally influence each other (Murphy 2012). Consequently, the more connected a group is, the more central it is to structuring the β-diversity of all groups. The weighted degree represents the total sum of partial correlations in β-diversity between a given group and all other connected groups. The higher the sum, the more important the group.

Validating the method using simulations

In order to test whether the Glasso approach was able to recover known interactions among species groups from local community β-diversity patterns, we first built a set of simulated data. We did so by constructing a regional trophic web (step 1) from which local multitrophic communities were sampled using a stochastic model (step 2). Then, we measured the partial correlation between the β-

diversity of each trophic level (step 3) and tested whether these patterns matched the simulated regional trophic web.

Step 1 – The regional web was assumed to have six trophic levels (one basal level and five consumer levels), containing twenty species each. We assumed some degree of specialization in the relationships between trophic levels: each consumer species had a number of prey equals to one plus a random number drawn in a Poisson law of parameter 1 (Fig. 1a). Thus, each consumer species had one prey species at least, and on average two prey species. Once the number of prey species had been drawn for a given consumer species, prey were drawn randomly from the lower trophic level.

Step 2 – Based on this regional network, we generated 1,000 local multi-trophic communities. Communities were simulated using a stochastic model of multi-trophic community assembly inspired by the Trophic Theory of Island Biogeography (TTIB, Gravel *et al.* 2011; Massol *et al.* 2017). The TTIB assumes bottom-up sequential dependencies (Holt 1997, 2009; Dunne *et al.* 2002) with two phases. In phase 1, each species can colonise a local community if at least one of its prey species is present. In phase 2, a species which has lost its last prey species goes extinct. For the sake of clarity, we assumed a homogeneous environment. The probability of each basal species being present in the local community was assumed to be constant and set to $p_0 = 0.5$. The probability of each consumer species C being present in the local community is related to the fraction of its prey available through the relation $p_C=(k/g)^r$ where g is the diet breadth of C (i.e. the number of potential prey species), k is the number of its prey species present in the community, and r is a constant that controls the shape of the relation. In the TTIB, having more prey species present in the community does not increase the probability of consumer presence, and the probability of survival is either 0 (when k=0) or 1 (when k>0). This corresponds to the limit case r=0. For the simulation, we used r=1, assuming that p_C grows linearly with the number of prey species present in the community. We also studied the case r=1/3 presented in the appendices.

Step 3 – We then computed β-diversity patterns for each trophic level and inferred the partial correlations between these β-diversities using the Glasso method explained above. We thus obtained a distribution of partial correlations between the β-diversity at the different trophic levels. We expect these partial correlations to be high between trophic levels which interact directly and low between trophic levels which do not interact. Since here only two successive trophic levels interact (see Fig. 1a), we expect the partial correlation between successive trophic levels to be high compared to the partial correlation between non-successive trophic levels. We expect these partial correlations to be more informative than marginal correlations (Pearson correlations), because they avoid spurious correlations due to confounding effects.

Analysing multi-trophic patterns in multi-trophic soil ecosystems in the French Alps.

Study site and soil sampling – The study was conducted in the northern French Alps (Arves Massif, 45.12°N, 6.40°E) along a 977 m elevation gradient (1748m to 2725m a.s.l.) located in a single cow-grazed pasture, above the tree-line. The vegetation at the bottom of the gradient corresponds mainly to subalpine grasslands, while alpine meadows with sparse vegetation dominate at high elevation (Chalmandrier *et al.* 2017). Ten plots were established at 100m altitude intervals along the gradient, each of them composed of two 10 x 10 m sub-plots. All plots were placed on the same south-facing slope with a similar bedrock type and land-use to ensure a relatively homogeneous gradient. Mean annual temperature along the gradient ranges between 8°C at the bottom and 3°C at the top, while mean annual rainfall is 473mm over the period 2000-2012. The soil sampling field campaign was conducted in September 2012.

We collected twenty-one soil samples per sub-plot following the two diagonals with the distance increasing exponentially from the corners to the centre on one diagonal and from the centre to the corners on the other. Central points were sampled only once. This sampling scheme was implemented to cover as much local biodiversity as possible at the sub-plot scale (Taberlet *et al.* 2012; Chalmandrier *et al.* 2017) and provided an almost regular distribution of pairwise spatial distances between points. Each sample contained 50g of soil from the uppermost 10 cm of soil and

was placed in individual plastic bags to be processed individually. Corers were cleaned and sterilised using a blowtorch between each sample to prevent cross-contamination (Taberlet *et al.* 2012). Extracellular DNA was extracted from 15g of soil as previously described (Taberlet *et al.* 2012; Zinger et al. 2016) within four hours after sample collection to prevent microbial growth. DNA was extracted twice for each sample. Blank extraction controls were included in the extraction process (using solely phosphate buffer as the template).

Molecular analyses – Soil biodiversity was estimated using four DNA markers. Universal markers such as 18S (amplifying all Eukaryotes, 18S nuclear rDNA) and 16S (amplifying all Bacteria, 16S rRNA) were used to obtain a general overview of the multi-trophic composition of the sites. Another two markers focus on Eukaryota diversity by targeting fungi (ITS1) and vascular plants (Chloroplast trnL-P6 loop), respectively. For each DNA extract, PCRs were run in duplicate leading to four PCR replicates per core sample. For each marker, PCM products were purified (MinElute™ PCR purification kit, Qiagen) with extraction and PCR blank controls included in the mix. High-throughput sequencing of eukaryotes and plant amplicons was performed on an Illumina HiSeq 2500 platform (2x150 bp paired-end reads for Eukaryotes and 2*200pb for plants) while fungal, archae and bacterial amplicons (200-350 bp) were sequenced on an Illumina MiSeq (2 x 250 bp paired-end reads) platform. Data curation is presented in Appendix 1. We pooled the samples together per subplot in order to obtain a single community per sub-plot and converted the data into presence-absences. The raw and curated sequencing data as well as associated metadata are available on the Dryad Digital Repository under accession XXX (to be provided on manuscript acceptance) and the summary statistics are presented in table S1. As explained above, we computed β-diversity using the Jaccard dissimilarity index. Since this index is sensitive to sequencing depth difference between samples, we also used the true turnover component of the Jaccard index (Baselga 2010), to remove this potential bias. The results obtained using the true turnover component of the Jaccard index are similar to those using the Jaccard index so we only refer to the

Jaccard index in the main text, and present the results of the true turnover component of the Jaccard index in Appendix 3.

Defining trophic groups of species

We selected *a priori* groups of soil taxa based on their distinctive role in the functioning of the soil ecosystem, and in biogeochemical cycles (especially the carbon and nitrogen cycles). Our focal taxa were: plants, fungi, bacteria, oribatid mites, nematodes, and springtails(Bardgett 2005). We included plants since they are the primary producers and their diversity and identity drive the functioning and the stability of most terrestrial ecosystems (Hooper, 1997; van der Heijen 1998). We also included litter feeders that contribute to dead material fragmentation, in a particular oribatid mites and springtails. Oribatid mites form one of the most abundant groups of arthropods in soil, (Behan-Pelletier 1999) with up to several hundreds of thousands of individuals per square meter (Norton 1990). Springtails are the most numerous group of hexapods in most terrestrial ecosystems (Deharveng 2004). We also included a group of taxa that mineralise the fragmented litter, such as saprophytic fungi and bacteria (Bardgett 2005). Fungi are found in different compartments of the soil trophic web so we classified fungal OTUs into three main functional groups, namely, symbiotic fungi, saprophytic fungi and pathogenic fungi using the FUNguild database (Nguyen *et al.* 2016). Symbiotic fungi represent a pivotal group in the soil ecosystem because over 90% of terrestrial plant families benefit from symbiotic associations between their roots and fungi (Wang & Qiu 2006). Moreover, in arctic and alpine systems, 60% to 80% of the nitrogen available for plants is supplied by mycorrhizal fungi (Bjorbækmoet *et al.* 2010). Bacteria also contribute to this supply by fixing atmospheric nitrogen (Bonfante & Anca 2009; Haq *et al.* 2014). Finally, we included a group of predators, here nematodes, the most abundant belowground multicellular animals (Bardgett 2005). Nematode OTUs were divided into bacterivore nematodes and herbivore/fungivore nematodes using the NEMAguild database (Nguyen *et al.* 2016). In summary, the nine groups included in this study were: plants, symbiotic fungi, pathogenic fungi, saprophytic fungi,

bacterivore nematodes, herbi/fungivore nematodes, bacteria, springtails, and oribatid mites. Most of these groups interact via trophic interactions, so that this study case matches the simulations described above, but not all of them. For the sake of clarity, we will hereafter refer to them as trophic groups.

Environmental characteristics – Mean annual soil temperature was estimated from field meteorological stations placed at the centre of each plot. We also estimated growing season length (GSL) and number of frost days based on daily maps of snow cover at 15 m resolution for 5 years falling between 2000 and 2014 and air temperature values extracted from the SAFRAN meteorological model developed by Meteo France for the French Alps (Durand *et al.* 2009). More methodological details and validation results for the snow cover model are available in Carlson *et al.* (2015). For each sub-plot, topsoil (0-10 cm) characteristics were determined from the average values obtained from three soil samples collected in July/August 2012, air-dried and then sieved at 2mm. Fine-scale topography and associated parameters (topographic wetness index and slope) were inferred from airborne LIDAR data acquired the year of sampling. Mean soil pH over the gradient was 5.40 (sd 0.300) whereas mean soil temperature over the year was 6.06°C (sd 2.84). The environmental distances between sub plots were estimated with Euclidean distances from the first two axes of a principal component analysis run for all produced (and normalised) environmental variables (the variance captured by the first two axes was 34% and 25% respectively, the first axis roughly represented the climatic conditions whereas the second axis roughly represented the soil conditions, see Appendix 2 for more details).

Results

Validation of the statistical method – Our simulations showed that trophic interactions produce significant relationships between β-diversity on consecutive trophic levels. Moreover, the Glasso method detected the conditional relationships between the β-diversity of the different trophic groups corresponding to their trophic position in the network (Fig. 1b) contrary to the marginal correlations

(Fig. 1c). The median of both marginal and partial correlation coefficients between pairs of trophic levels decreases with the shortest path length between these trophic levels. Nevertheless, while the median of the marginal correlation coefficients (Fig. 1e) decreased slowly (the distribution of marginal correlation coefficients overlapped), the median of the partial correlation coefficients (Fig. 1d) was close to 0 once the trophic level distance was higher than 1. In other words, marginal correlation analyses detected spurious correlation between non-adjacent trophic levels, but this bias was strongly alleviated when considering partial correlations. Although these simulation results remain qualitative, they demonstrate that specialised biotic interactions between trophic group leave detectable imprints on compositional patterns and that these patterns can be detected using our method. Changing the shape of the relationship linking the probability of presence of a consumer species with its number of available prey species did not alter this conclusion (Fig. S4).

β-diversity modelling of empirical soil communities – In the French Alps, the partial correlations estimated between the *β*-diversity of each predefined trophic group and environmental distances were all positive (Fig. 2, Fig. S5, Fig. S6, table S2). The estimated partial correlation network had a connectance of 0.618 and was composed of 34 undirected edges (i.e. partial correlation coefficients > 0) out of 55 possible edges and 11 nodes (9 trophic groups and 2 environmental variables).

Saprophytic fungi were the most influential group in conditioning the *β*-diversity of the other groups (highest degree value, 8, and highest weighted degree value, 1.30, Figure 3). Plants and oribatid mites also had a strong influence on the *β*-diversity of other groups, as did pathogenic and symbiotic fungi. In contrast, environmental variables had a relatively small impact on the *β*-diversity of the trophic groups.

The probability of observing a non-null partial correlation between the *β*-diversity of a trophic group and the environmental distance was 0.44 (8 edges linking environmental nodes to the trophic group nodes and 18 potential edges) whereas the probability of observing a link between the *β*-diversity of any two trophic groups was 0.69 (25 edges and 36 potential edges). Since the variables

associated with disconnected nodes are conditionally independent and that conditionally independent variables cannot causally influence each other, this highlights the reduced influence of environmental variables on the β-diversity of the trophic groups.

Discussion

In this study, we proposed a method for dissecting the joint spatial structure of multiple trophic groups. This method builds on observed patterns of β-diversity in multiple trophic groups to infer the conditional dependencies between pairs of groups and with the environment, in order to pinpoint potential biotic interactions. Simulations confirmed that our method is able to recover the overall structure of a trophic network using partial correlations. This method does not directly infer the interaction network at species level, and the inferred conditional dependencies at the group level do not imply causality. A similar method for producing directed networks is path analysis (Wright 1921; Shipley 2000 application on soil diversity Schuldt *et al.* 2017). However, the ability of path analysis to draw causal inference is controversial (Shipley 2000 but see Dawid 2010; Bollen & Pearl 2013). The correlative method presented here should help interpret a wide range of multi-trophic biodiversity patterns. Although our approach is likely to be sensitive to the effect of a missing predictor (since the structure of the partial correlation network would be affected by the addition of a variable), environmental DNA sampling should facilitate comprehensive assessments. One key issue remains the mapping of OTUs into functional groups, which is only possible using integrative biology knowledge combined with large and comprehensive databases. An interesting perspective for this method is that it could also be used to study temporal changes in biodiversity (i.e., temporal β-diversity) using environmental DNA approaches (Bohan *et al.* 2017).

When applied to soil multi-trophic diversity along an elevation gradient of the French Alps, we were able to quantify the relative importance of biotic interactions and the environment in shaping the spatial structure of the meta-communities. Pairwise environmental distances displayed a low correlation with the β-diversity of each group (as measured with the few non-zero partial

correlations). In other words, the ecological community was found to be shaped more by biotic interactions than by abiotic environmental constraints. This result is rather surprising given the sharpness of the elevational gradient, and the expected importance of environmental filtering in shaping above and below-ground communities along elevation gradients (Meynard *et al.* 2013). However, many previous studies have focused on intra-guild biotic interactions, or on environmental effect only, and they have ignored the importance of inter-guild interactions (e.g. Chalmandrier *et al.*, 2015). As our climatic and soil properties were sampled at plot level, they might not have been measured at the appropriate scale to reflect the fine conditions experienced by below-ground organisms (Falconer *et al.* 2015; Baveye *et al.* 2016; Zinger *et al.* 2017). A single plot-scale measure of abiotic conditions may not be sufficient to explain the β-diversity of organisms with heterogeneous sizes and capture the fine-scale heterogeneity of soil conditions. Additionally, we cannot rule out the fact that some relatively important missing environmental factors might explain the low predictive power of environmental distances (e.g. phosphorous for symbiotic mushrooms, Liu *et al.* 2012; Camenzind *et al.* 2014). Nevertheless, while the greater β-diversity was due to true-turnover (a turnover measure independent of variation of species richness; on average, 80% of the β-diversity across trophic groups), climatic and soil properties had a stronger influence on the true-turnover component of the β-diversity of the different trophic groups.

The approach described here should help interpret a wide range of multi-trophic biodiversity patterns and yield new testable hypotheses. For example, in multiple pairs of trophic groups we found strong spatial associations which can be interpreted in light of functional associations. In the soil, plants are directly affected by symbiotic/pathogenic fungi and root herbivores, whereas the complex network of detritivorous organisms affects them indirectly (Bardgett 2005; Wardle *et al.* 2011). Our method showed that plant β-diversity was strongly linked to the β-diversities of symbiotic fungi, pathogenic fungi, oribatid mites and bacterivore nematodes. While the links with the two groups of fungi most likely reflect the strong direct associations between these groups and the vegetation structure, the link with oribatid mites may be explained by the fact that this group

feeds on plant litter fungi, especially saprophytic fungi (Schneider *et al.* 2004, Crowther *et al.* 2012; 2011). The graphical lasso method detected this link showing that changes in the composition of saprophytic fungi lead to a change in the oribatid mite assemblages. As expected, the β-diversity of bacteria correlated to that of bacterivore nematodes, which reflects this known trophic interaction (Ettema 1998; Wardle 2006). Moreover, the β-diversity of herbi/fungivore nematodes correlated to that of symbiotic fungi. The link between the β-diversity of oribatid mites and springtails might be explained by mite predation of springtails (Ferguson & Joly 2002). Interestingly, we revealed a link between the β-diversity of oribatid mites and symbiotic fungi. This result suggests that mite assemblages could be influenced by fungi spatial distribution through trophic interactions, which have been so far very poorly documented (Gange & Brown 2003). Interestingly, our analysis also showed a strong partial correlation (the strongest partial correlation: 0.365) between the β-diversities of saprophytic fungi and bacteria. Many clades of these two groups exploit the same resource (in particular cellulose, which represents 30-50 % of plant dry weight, Whipps 2001; De Boer et al. 2005) which thus creates an indirect relationship between bacteria and saprophytic fungi, through an unmeasured confounding effect. The relationship uncovered between bacteria and symbiotic fungi could be attributed to the fact that bacteria can assist mycorrhiza by colonising the extraradical hyphae or by living in the cytoplasm of mycorrhizal fungi (Bonfante *et al.* 2009; Haq *et al.* 2014).

Conclusions

The rise of environmental DNA metabarcoding and the ever-increasing availability of databases on species co-occurrence have opened up a new era in quantitative and predictive ecology. Comprehensive species lists of taxa are necessary, but not sufficient to provide information on how co-occurring species interact and whether multi-trophic interactions shape community assembly. We have developed a method that could reveal how trophic groups influence each other and respond to environmental variation. As such our method helps uncover potential causes of the compositional

turnover of species groups from a multi-trophic interaction perspective. Gathering informative expert knowledge from literature review or using automatic methods should then pave the way to understanding the fine-scale structure of multi-trophic assemblages and obtaining a more holistic picture of biodiversity.

Acknowledgements

We would like to thank all the students and colleagues who participated in the intensive field season that generated the data used in the manuscript, as well as the Genotoul bioinformatics platform Toulouse Midi-Pyrenees (Bioinfo Genotoul) and Pierre Solbes (EDB-Calc Cluster) for providing computing and storage resources. The research received funding from the French Agence Nationale de la Recherche (ANR) through the METABAR (ANR-11-BSV7-0020) and GlobNets (ANR-16-CE02-0009) projects, and from “Investissement d’Avenir” grants managed by the ANR (Trajectories: ANR-15-IDEX-02; Montane: OSUG@2020: ANR-10-LAB-56; CEBA: ANR-10-LABX-25-01, TULIP: ANR-10- 599 LABX-0041).

Competing interests

L.G. and P.T. are co-inventors of patents related to the gh primers and the use of the P6 loop of the chloroplast trnL (UAA) intron for plant identification using degraded template DNA. These patents only restrict commercial applications and have no impact on the use of this locus by academic researchers.

Figure caption

Figure 1: Design and results of the simulation assessing the impact of biotic interactions on the marginal and partial correlation of the *β*-diversity values of different trophic levels. *(a)* the trophic network used for the simulation, composed of six trophic levels. The indegree of each node is one plus an integer randomly drawn from a Poisson law of parameter 1.0 (except for the basal species). (b) Partial correlation network built using the median values of the partial correlations (inferred using the graphical lasso) between the *β*-diversity values of different trophic levels. (c) Marginal correlation networks built using the median values of the Pearson correlations between the *β*-diversity values of different trophic levels. *(d)* Partial correlation coefficient (inferred using the graphical lasso) between the *β*-diversity values of two trophic levels as a function of the shortest path length between these trophic levels. *(e)* Marginal Pearson correlation coefficient between the *β*-diversity values of two trophic levels as a function of the shortest path length between these trophic levels.

Figure 2: Undirected partial correlation network inferred using the graphical lasso method between the *β*-diversity of the major trophic groups constituting soil biodiversity and the environmental distances. Each node represents the *β*-diversity of a trophic group or an environmental distance. Here, only the partial correlations above the median value of the non-null partial correlation coefficients (0.106) are shown. The non-filtered network is presented in Figure S1. Edge thickness is proportional to the value of the partial correlation coefficient and the partial correlation coefficients are all positive.

Figure 3: Properties of the inferred network. The degree (left panel) is the number of neighbours of nodes in a graph (here, the undirected partial correlation network). It measures the number of variables that are conditionally dependent on the variable associated with this node. The weighted degree (right panel) is the sum of the partial correlation coefficients attached to the edges adjacent to this node. Dashed lines represent the mean values of degree and weighted degree.

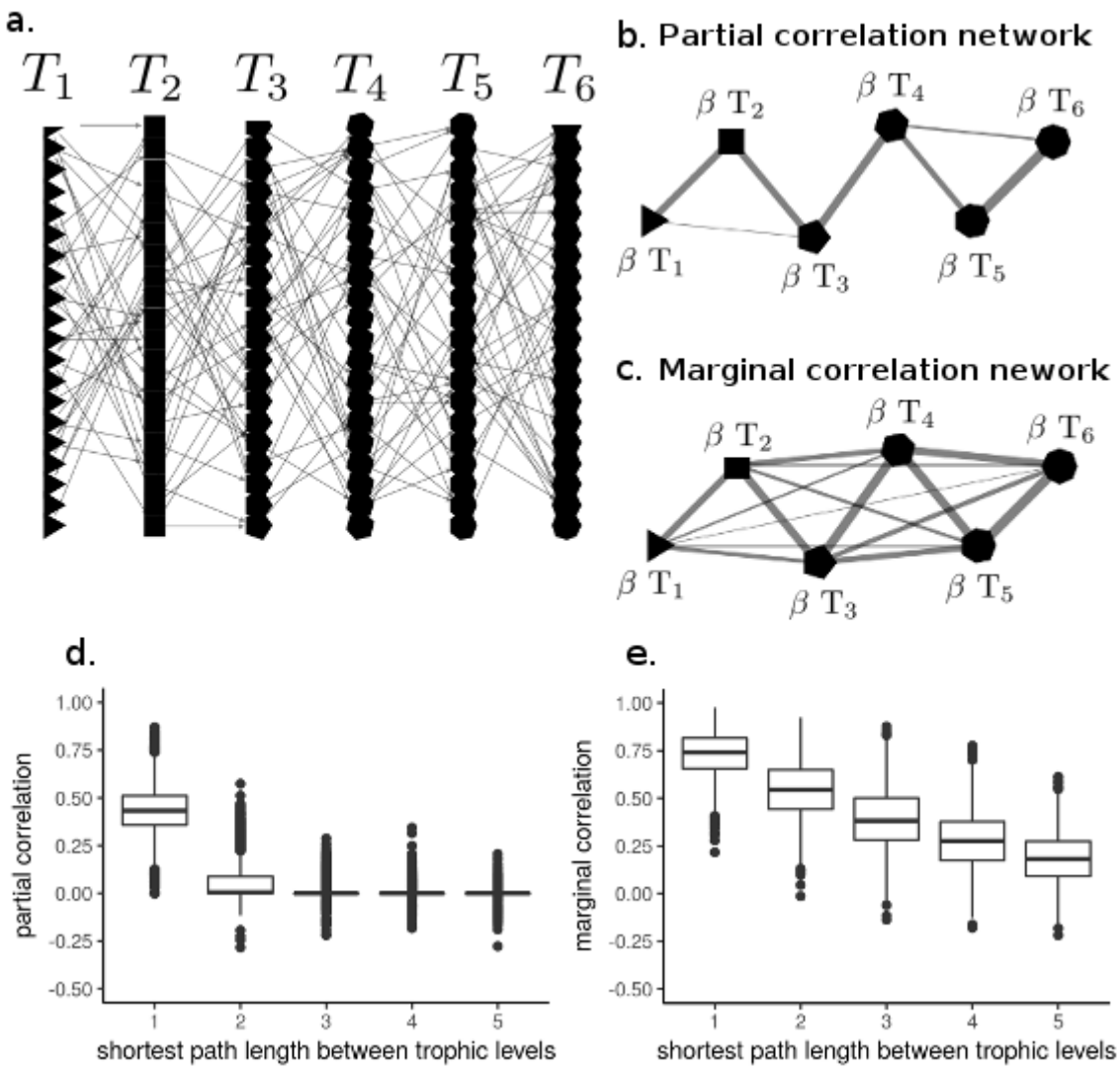

a.
T_1 T_2 T_3 T_4 T_5 T_6
b. Partial correlation network
β T$_2$
β T$_4$
β T$_6$
β T$_1$
β T$_3$
β T$_5$
c. Marginal correlation nework
β T$_2$
β T$_4$
β T$_6$
β T$_1$
β T$_3$
β T$_5$
d.
partial correlation
1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
1
2
3
4
5
shortest path length between trophic levels
e.
marginal correlation
1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
1
2
3
4
5
shortest path length between trophic levels

Figure 1

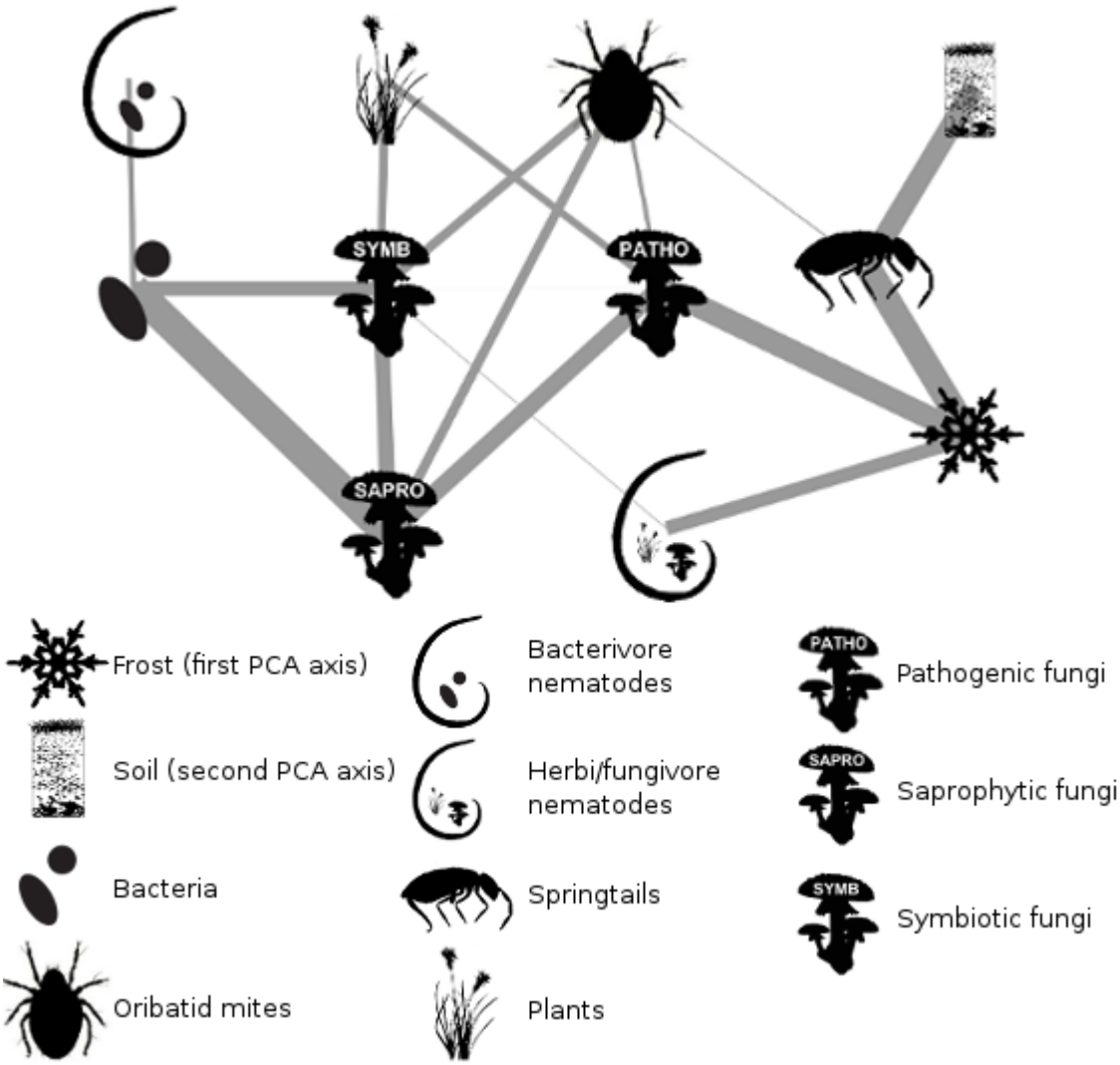
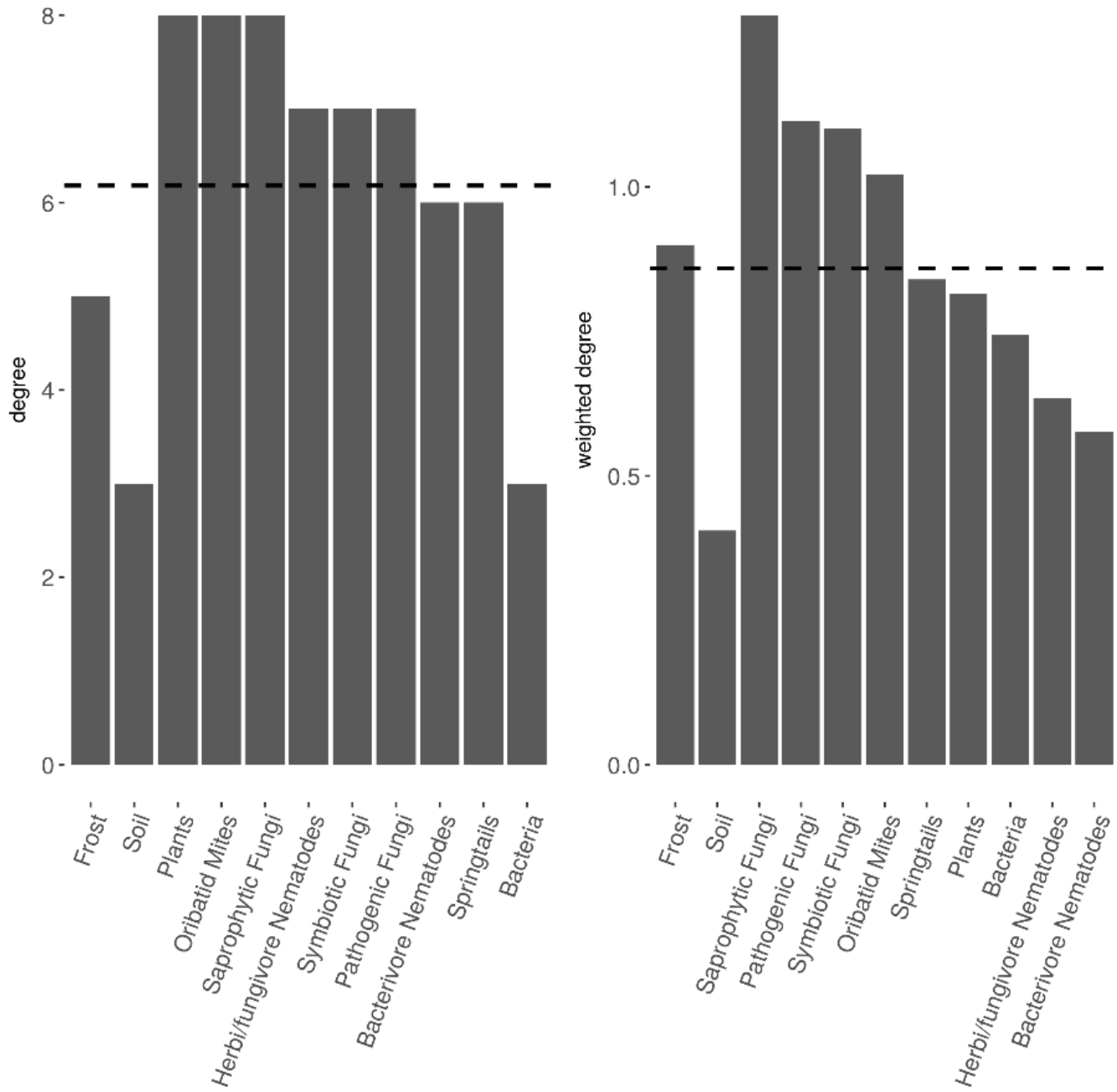

SYMB
PATHO
SAPRO
Frost (first PCA axis)
Soil (second PCA axis)
Bacteria
Oribatid mites
Bacterivore nematodes
Herbi/fungivore nematodes
Springtails
Plants
PATHO
Pathogenic fungi
SAPRO
Saprophytic fungi
SYMB
Symbiotic fungi

Figure 2

502

503

504



506 **Figure 3**

507

References

Allesina, S., Alonso, D., & Pascual, M. (2008). A General Model for Food Web Structure. *Science*, **320**, 658-661

Anderson, M. J., Crist, T. O., Chase, J. M., Vellend, M., Inouye, B. D., Freestone, A. L. *et al.* (2011). Navigating the multiple meanings of β diversity: A roadmap for the practicing ecologist. *Ecology Letters*, **14**, 19–28.

Bardgett, R. D. (2005). *The Biology of Soil: A Community and Ecosystem Approach*. Oxford University Press.

Bardgett, R. D., & Wardle, D. A. (2010). *Aboveground-belowground linkages : biotic interactions, ecosystem processes, and global change*. Oxford University Press.

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, **19**, 134–143.

Baselga, A., & Orme, C. D. L. (2012). betapart : an R package for the study of beta diversity. *Methods in Ecology and Evolution*, **3**, 808–812.

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. *Third International AAAI Conference on Weblogs and Social Media*, 361–362.

Baveye, P. C., Berthelin, J., & Munch, J.-C. (2016). Too much or not enough: Reflection on two contrasting perspectives on soil biodiversity. *Soil Biol. Biochem.*, **103**, 320-326.

Behan-Pelletier, V. M. (1999). Oribatid mite biodiversity in agrosystems:role for bioindication. *Agric. Ecosyst. Environ.*, **74** , 411–423.

Bjorbækmo, M., Carlsen, T., Brysting, A., Vrålstad, T., Høiland, K., Ugland, K. *et al.* (2010). High diversity of root associated fungi in both alpine and arctic Dryas octopetala. *BMC Plant Biology*, **10**, 244.

Bohan, D. A., Vacher, C., Tamaddoni-Nezhad, A., Raybould, A., Dumbrell, A. J., & Woodward, G. (2017). Next-Generation Global Biomonitoring: Large-scale, Automated Reconstruction of Ecological Networks. *Trends in Ecology and Evolution*, **32**, 477–487.

Bollen K.A., Pearl J. (2013) Eight Myths About Causality and Structural Equation Models. In: Morgan S. *Handbook of Causal Analysis for Social Research. Handbooks of Sociology and Social Research.* Springer, Dordrecht (pp. 301–328)

Bonfante, P., & Anca, I.-A. (2009). Plants, Mycorrhizal Fungi, and Bacteria: A Network of Interactions Rhizosphere: the narrow zone of soil surrounding living roots. *Annu. Rev. Microbiol*, **63**, 363–83.

Borcard, D., Legendre, P., & Drapeau, P. (1992). Partialling out the Spatial Component of Ecological Variation. *Ecology*, **73**, 1045–1055.

Broeckling, C. D., Broz, A. K., Bergelson, J., Manter, D. K., & Vivanco, J. M. (2008). Root exudates regulate soil fungal community composition and diversity. *Applied and Environmental Microbiology*, **74**, 738–44.

Brose, U., & Hillebrand, H. (2016). Biodiversity and ecosystem functioning in dynamic landscapes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **371**, 20150267.

Camenzind, T., Hempel, S., Homeier, J., Horn, S., Velescu, A., Wilcke, W. *et al.* (2014). Nitrogen and phosphorus additions impact arbuscular mycorrhizal abundance and molecular diversity in a tropical montane forest. *Global Change Biology*, **20**, 3646–3659.

Carlson, B. Z., Choler, P., Renaud, J., Dedieu, J. P., & Thuiller, W. (2015). Modelling snow cover duration improves predictions of functional and taxonomic diversity for alpine plant communities. *Annals of Botany*, **116**, 1023–1034.

Chalmandrier, L., Münkemüller, T., Colace, M. P., Renaud, J., Aubert, S., Carlson, B. Z. *et al*.(2017). Spatial scale and intraspecific trait variability mediate assembly rules in alpine grasslands. *Journal of Ecology*, **105**, 277–287.

549 Chalmandrier, L., Münkemüller, T., Lavergne, S., & Thuiller, W. (2015). Effects of species' similarity and dominance
550 on the functional and phylogenetic structure of a plant meta-community. *Ecology*, **96**, 143–153.

551 Crowther, T. W., & A'Bear, A. D. (2012). Impacts of grazing soil fauna on decomposer fungi are species-specific and
552 density-dependent. *Fungal Ecology*, **5**, 277–281.

553 Crowther, T. W., Boddy, L., & Jones, T. H. (2011). Species-specific effects of soil fauna on fungal foraging and
554 decomposition. *Oecologia*, **167**, 535–545.

555 Dawid, A. P. (2008). Beware of the DAG! *NIPS 2008Workshop on Causality*, 59–86.

556 De Boer, W., Folman, L. B., Summerbell, R. C., & Boddy, L. (2005). Living in a fungal world: Impact of fungi on soil
557 bacterial niche development. *FEMS Microbiology Reviews*, **29**, 795–811.

558 Deharveng, L. (2004). Recent advances in Collembola systematics. *Pedobiologia*, **48**, 415–433.

559 Dray, S., Pélissier, R., Couteron, P., Fortin, M.-J., Legendre, P., Peres-Neto *et al.* (2012). Community ecology in the age
560 of multivariate multiscale spatial analysis. *Ecological Monographs*, **82**, 257–275.

561 Dunne, J. A., Williams, R. J., & Martinez, N. D. (2002). Network structure and biodiversity loss in food webs:
562 robustness increase with connectance. *Ecology Letters*, **5**, 558–567.

563 Durand, Y., Laternser, M., Giraud, G., Etchevers, P., Lesaffre, B., & Mérindol, L. (2009). Reanalysis of 44 yr of climate
564 in the French Alps (1958-2002): Methodology, model validation, climatology, and trends for air temperature and
565 precipitation. *Journal of Applied Meteorology and Climatology*, **48**, 429–449.

566 Epskamp, S., & Fried, E. I. (2016). A Tutorial on Regularized Partial Correlation Networks. arXiv:1607.01367

567 Ettema, C. H. (1998). Soil Nematode Diversity: Species Coexistence and Ecosystem Function. *Journal of Nematology*.
568 **30**, 159-169

569 Falconer, R. E., Battaia, G., Schmidt, S., Baveye, P., Chenu, C., & Otten, W. (2015). Microscale heterogeneity explains
570 experimental variability and non-linearity in soil organic matter mineralisation. *PLoS ONE*, **10**, 1–12.

571 Ferguson, S. H., & Joly, D. O. (2002). Dynamics of springtail and mite populations: The role of density dependence,
572 predation, and weather. *Ecological Entomology*, **27**, 565–573.

573 Foygel, R., & Drton, M. (2010). Extended Bayesian Information Criteria for Gaussian Graphical Models. *Adv Neural*
574 *Inf Process Syst.*, 604–612.

575 Friedman, J., Hastie, T., & Tibshirani, R. (2007). Sparse inverse covariance estimation with the lasso. *Biostatistics*, **9**,
576 432-441

577 Gange, A. C., & Brown, V. K. (2003). Actions and Interactions of Soil Invertebrates and Arbuscular Mycorrhizal Fungi
578 in Affecting the Structure of Plant Communities. In M. G. A. van der Heijden & I. R. Sanders (Eds.), *Mycorrhizal*
579 *Ecology* (pp. 321–344). Berlin, Heidelberg: Springer Berlin Heidelberg.

580 Geremia, R. A., Puscas, M., Zinger, L., Bonneville, J. M., & Choler, P. (2016). Contrasting microbial biogeographical
581 patterns between anthropogenic subalpine grasslands and natural alpine grasslands. *New Phytologist*, **209**, 1196–
582 1207

583 Gonzalez-Varo, J. P., & Traveset, A. (2016). The Labile Limits of Forbidden Interactions. *Trends in Ecology and*
584 *Evolution*, **31**, 700–710.

585 Graham, C. H., & Fine, P. V. A. (2008). Phylogenetic beta diversity: Linking ecological and evolutionary processes
586 across space in time. *Ecology Letters*, **11**, 1265–1277.

587 Gravel, D., Massol, F., Canard, E., Mouillot, D., & Mouquet, N. (2011). Trophic theory of island biogeography.
588 *Ecology Letters*, **14**, 1010–1016.

589 Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., & Martiny, J. B. H. (2012). Beyond biogeographic patterns:
590 Processes shaping the microbial landscape. *Nature Reviews Microbiology*, **10**, 497–506.

591 Haq, I. U., Zhang, M., Yang, P., & van Elsas, J. D. (2014). The interactions of bacteria with fungi in soil: emerging
592 concepts. *Advances in applied microbiology*, **89**, 432-441.

593 Hättenschwiler, S., Tiunov, A. V, & Scheu, S. (2005). Biodiversity and Litter Decomposition in Terrestrial Ecosystems.
594 *Annual Review of Ecology, Evolution, and Systematics*, **36**, 191–218.

595 Hedlund, K., & Öhrn, M. S. (2000). Tritrophic interactions in a soil community enhance decomposition rates. *Oikos*, **88**,
596 585–591.

597 HilleRisLambers, J., Adler, P. B., Harpole, W. S., Levine, J. M., & Mayfield, M. M. (2012). Rethinking Community
598 Assembly through the Lens of Coexistence Theory. *Annual Review of Ecology, Evolution, and Systematics*, **43**,
599 227–248.

600 Holt, R.D. (1997) From Metapopulation Dynamics to Community Structure: Some Consequences of Spatial
601 Heterogeneity. In : *Metapopulation Biology: Ecology, Genetics and Evolution.* Academic Press, New York, pp.
602 149-164.

603 Holt, R.D. (2009) Towards a trophic island biogeography: reflections on the interface of island biogeography and food
604 web ecology. in J.B. Losos and R.E. Ricklefs *The Theory of Island Biogeography Revisited* Princeton University
605 Press pp. 143-185

606 Hooper, D. U. (1997). The Effects of Plant Composition and Diversity on Ecosystem Processes. *Science*, **277**, 1302–
607 1305.

608 Koller, D., & Friedman, N. (2009). *Probabilistic Graphical Models: Principles and Techniques - Adaptive Computation*
609 *and Machine Learning*. The MIT Press.

610 Kress, W. J., García-Robledo, C., Uriarte, M., & Erickson, D. L. (2015). DNA barcodes for ecology, evolution, and
611 conservation. *Trends in Ecology and Evolution*, **30**, 25–35.

612 Liu, Y., Shi, G., Mao, L., Cheng, G., Jiang, S., Ma, X. *et al.* (2012). Direct and indirect influences of 8 yr of nitrogen
613 and phosphorus fertilization on Glomeromycota in an alpine meadow ecosystem. *New Phytologist*, **194**, 523–535.

614 Massol, F., Dubart, M., Calcagno, V., Cazelles, K., Jacquet, C., Kéfi, S., & Gravel, D. (2017). Chapter Four – Island
615 Biogeography of Food Webs. In *Advances in Ecological Research*, **56**, 183–262.

616 Mazel, F., Wüest, R. O., Gueguen, M., Renaud, J., Ficetola, G. F., Lavergne, S. *et al.* (2017). The Geography of
617 Ecological Niche Evolution in Mammals. *Current Biology*, **27**, 1369-1374.

618 Mazumder, R., & Hastie, T. (2011). The Graphical Lasso: New Insights and Alternatives. arXiv:1111.5479

619 Meynard, C. N., Lavergne, S., Boulangeat, I., Garraud, L., Van Es, J., Mouquet, N., & Thuiller, W. (2013).
620 Disentangling the drivers of metacommunity structure across spatial scales. *Journal of Biogeography*, **40**, 1560–
621 1571.

622 Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. The MIT Press. pp. 933-935.

623 Norton, R.A., 1990. Acarina: Oribatida. In: *Soil Biology Guide*. Wiley, New York, pp. 779–803.

624 Nguyen, N. H., Song, Z., Bates, S. T., Branco, S., Tedersoo, L., Menke, J. *et al.* (2016). FUNGuild: An open annotation
625 tool for parsing fungal community datasets by ecological guild. *Fungal Ecology*, **20**, 241–248.

626 Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D. *et al.* (2015). vegan: Community
627 Ecology Package. *R Package Version 2.3-2*, https://CRAN.R-project.org/package=vegan.

628 Peay, K. G., Kennedy, P. G., & Talbot, J. M. (2016). Dimensions of biodiversity in the Earth mycobiome. *Nature*
629 *Reviews Microbiology*, **14**, 434–447.

630 Peay, K. G., Russo, S. E., Mcguire, K. L., Lim, Z., Chan, J. P., Tan, S. *et al.* (2015). Lack of host specificity leads to
631 independent assortment of dipterocarps and ectomycorrhizal fungi across a soil fertility gradient. *Ecology Letters*,
632 **18**, 807–816.

633 Prober, S. M., Leff, J. W., Bates, S. T., Borer, E. T., Firn, J., Harpole, W. S *et al.* (2015). Plant diversity predicts beta but
634 not alpha diversity of soil microbes across grasslands worldwide. *Ecology Letters*, **18**, 85–95.

635 Ricklefs, R. E. (2008). Disintegration of the ecological community. *The American Naturalist*, *172*, 741–750.

636 Schneider, K., Renker, C., Scheu, S., & Maraun, M. (2004). Feeding biology of oribatid mites: a minireview.
637 *Phytophaga XIV: Acarine Biodiversity in the Natural and Human Sphere (Proceedings of the Symposium of the*
638 *European Association of Acarologists)*, 247–256.

639 Shipley, B. (2000). *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and*
640 *Causal Inference*. Ecology, Cambridge University Press, pp. 48-55

641 Smith, S. E.,Read, D. J., David J. . (2008). *Mycorrhizal symbiosis*. Academic Press.

642 Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, **21**,
643 1789–1793.

644 Tibshirani, R., & Tibshirani, R. (1994). Regression Shrinkage and Selection Via the Lasso. J*ournal of the royal*
645 *statistical society,* **58**, 267--288.

646 van der Heijden, M. G. A., Klironomos, J. N., Ursic, M., Moutoglis, P., Streitwolf-Engel, R., Boller, T. *et al.* (1998).
647 Mycorrhizal fungal diversity determines plant biodiversity , ecosystem variability and productivity. *Nature*, **396**,
648 69–72.

649 Walker, J. F., Aldrich-Wolfe, L., Riffel, A., Barbare, H., Simpson, N. B., Trowbridge, J. *et al.* (2011). Diverse helotiales
650 associated with the roots of three species of arctic ericaceae provide no evidence for host specificity. *New*
651 *Phytologist*, **191**, 515–527.

652 Wang, B., & Qiu, Y. L. (2006). Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza*, **16**,
653 299–363.

654 Wardle, D. A. (2006). The influence of biotic interactions on soil biodiversity. *Ecology Letters*, **9**, 870–886.

655 Wardle, D. A., Bardgett, R. D., Klironomos, J. N., Setälä, H., Wim, H., Putten, V. Der, & Wall, D. H. (2011). Ecological
656 Linkages between Aboveground and Belowground Biota. *American Association for the Advancement of Science*
657 *Stable*, **304**, 1629–1633.

658 Whipps, J. M. (2001). Microbial interactions and biocontrol in the rhizosphere. *Journal of Experimental Botany*, ***52***,
659 487–511.

660 Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, **20**, 557– 585.

661 Zinger, L., Lejon, D. P. H., Baptist, F., Bouasria, A., Aubert, S., Geremia, R. A., & Choler, P. (2011). Contrasting
662 diversity patterns of crenarchaeal, bacterial and fungal soil communities in an alpine landscape. PLOS ONE, **6**, 1-
663 7.

664 Zinger, L., Taberlet, P., Schimann, H., Bonin, A., Boyer, F., De Barba, M. *et al.* (2017). Soil community assembly varies
665 across body sizes in a tropical forest. *bioRxiv* 154278; doi: https://doi.org/10.1101/154278

666

667

668

Appendices

Appendix 1 : Molecular and bioinformatics analyses

Data were then curated using the OBITools software package (Boyer *et al.* 2016) together with customs R scripts (R Core Team 2015) following the procedure described in Zinger *et al.* (2017). Paired-end reads were assembled, assigned to their respective samples/marker and dereplicated. Low-quality sequences were excluded. For the remaining ones we computed pairwise dissimilarities between sequences (i.e. the number of mismatches, allowed to be 0-3) using the Sumatra algorithm (Mercier *et al.* 2013), and formed Operational Taxonomic Units OTUs using the Infomap community detection algorithm (Rosvall & Bergstrom 2008). Abundance of OTUs was defined as the sum of read abundances of sequences belonging to them. In subsequent analyses, each OTU was represented by its most abundant sequence. Each OTU was assigned a taxonomic clade using the *ecotag* program (Boyer *et al.* 2016) using a set of reference databases to refine taxonomic annotations. (e.g. GENBANK, UNITE database, Abarenkov *et al.* 2010) We paid particular attention to minimize PCR/sequencing errors, contaminant and false positive sequences as well as potential non-functional PCRs by using the procedure described in Zinger *et al.* (2017). PCR replicates were finally summed for each samples.

	Bacteria	**Eukaryota**	**Fungi**	**Viridiplantae**
Primers set				
Targeted region	V5-V6 regions, 16S rRNA gene	V7 region, 18S rRNA gene	ITS1	P6 loop of the chloroplast trn*L* intron
Forward primer (5'–3')	GGATTAGATACCC TGGTAGT‡	TTTGTCTGSTTAATTSCG§	CAAGAGATCCGTTGTT GAAAGTK†	GGGCAATCCTGAG CCAA¶
Reverse primer (5'–3')	CACGACACGAGCT GACG‡	TCACAGACCTGTTATTGC§	GGAAGTAAAAGTCGTA ACAAGG§§	CCATTGAGTCTCT GCACCTATC¶
PCR conditions				
Number of cycles	35	45	40	40
PCR cycles characteristics†† (Denaturation, Annealing, Elongation – Final elongation)	95°C(30s), 57°C(30s), 72°C(90s) – 72°C(7min)	95°C(30s), 45°C(30s) 72°C(60s) – 72°C(7min)	95°C (30 sec), 55°C (30 sec), 72°C (60s) – 72°C (7 min)	95°C(30s), 50°C(30s), 72°C(60s) – 72°C(7min)
Illumina platform	MiSeq 2x250	HiSeq 2x150	MiSeq 2x250	HiSeq 2x100
Reads and Otus statistics				
Initial number of OTUs	25383	7242	4193	2153

Final number of OTUs	24058	5833	2987	555
Initial number of reads	5042640	68893592	4159074	37509301
Final number of reads	4357053	61818803	3626819	35950935

† this study

§§ White, T. J., T. Bruns, S. Lee, and J. W. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pp. 315-322 In: PCR Protocols: A Guide to Methods and Applications, eds. Innis, M. A., D. H. Gelfand, J. J. Sninsky, and T. J. White. Academic Press, Inc., New York.

‡ Fliegerova, K., Tapio, I., Bonin, A., Mrazek, J., Callegari, M. L., Bani, P., *et al.* (2014). Effect of DNA extraction and sample preservation method on rumen bacterial population. *Anaerobe*, **29**, 80-84.

§ Guardiola, M., Uriz, M. J., Taberlet, P., Coissac, E., Wangensteen, O. S., & Turon, X. (2015). Deep-Sea, Deep-Sequencing: Metabarcoding Extracellular DNA from Sediments of Marine Canyons. PLoS One, 10, e0139633.

¶ Taberlet, P., Gielly, L., Pautou, G., & Bouvet, J. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, 17, 1105-1109.

†† All PCRs were preceded by an initial denaturation at 95°C (10 min)

Table S1 Markers characteristics, PCR conditions and sequencing technology

Appendix 2 : Climatic and soil variables

Climatic variables: We described the bioclimatic conditions of the gradient using four variables describing both thermal and nival regimes: soil temperature, solar radiations, Growing Season Length (hereafter “GSL”) and number of frost days. Mean annual soil temperature was recorded using HOBO stations located at the center of each sampling site. Solar radiation was averaged from simulations of solar radiation using a 2-m digital elevation model (DEM) (Carlson et al. 2015) derived from a LIDAR acquisition. The number of frost days, i.e. the number of snow-free days with air temperature below 0°C, was computed as described in Carlson et al. (2015). GSL, the number of snow-free days with a daily mean air temperature over 0°C between snow melt-out and August 15, was computed as described in Chalmandrier et al. 2016.

Topographical variables: Slope, Topographic Wetness Index (TWI) and Topographic Position Index (TPI) were inferred with the DEM derived from the airborne LiDAR data acquired the year of sampling.

Soil analyses were conducted in summer 2012 from soil cores taken at the center of each subplot (1m25x1m25). Each of them were weighed and 5mm-sieved (Legay et al. 2014) and processed for quantifying soil moisture, soil organic matter content, pH, soil texture (gravel mass and apparent density after 2mm sieving) and soil nutrients (ammonium (soil NH4+), nitrates (soil NO3-), total dissolved nitrogen (TDN), and dissolved organic nitrogen (DON)) from 0.5m K2SO4 soil extracts (Jones et al. 2004; Legay et al. 2014). Soil subsamples were added 100 mL distilled water to saturate each soil core and calculate the waterfilled pore space (WFPS) and then air dried and ground to measure total

soil C and N (Legay et al. 2014). Microbial biomass nitrogen (N) was measured using the chloroform-fumigation extraction technique of Vance et al. 1987). Potential N mineralization rates were estimated using anaerobic incubations of fresh soil subsamples (Legay et al. 2014). Potential ammonium and nitrates leaching was measured from the percolate of a site central supplementary core leached in a given volume of distilled water (Legay et al. 2016).

	PCA axis 1	PCA axis 2
pH	-0.417	-0.368
Gravel mass	0.791	-0.277
Apparent density	0.235	-0.781
Total porrosity rate	0.264	0.789
Water Filled pore space	-0.76	-0.23
Leaching NO3	-0.398	-0.029
Leaching NH4	0.0901	0.791
N rate	0.0542	-0.933
C rate	-0.587	0.713
C/N ratio	-0.332	0.896
Rate organic matter	0.332	0.863
N NO3	-0.341	-0.368
N NH4	-0.512	0.264
N TDN	0.257	-0.106
N NH4 product	-0.158	-0.0632
Year soil temperature	-0.823	0.00538
FrostDays	-0.858	-0.196
Growing Season Length	-0.983	-0.0919
Topographic wetness index	-0.652	-0.0833
slope	0.28	-0.343
Topographic position indexPI10m_mean_topo	0.35	0.101
Altitude Digital Elevation Model	0.98	0.0867
Mean solar radiation	-0.964	-0.058
Variance Growing Season Length	0.834	0.347

Table S2 Correlations between the environmental variables and the first two axes of the Principal Component Analysis

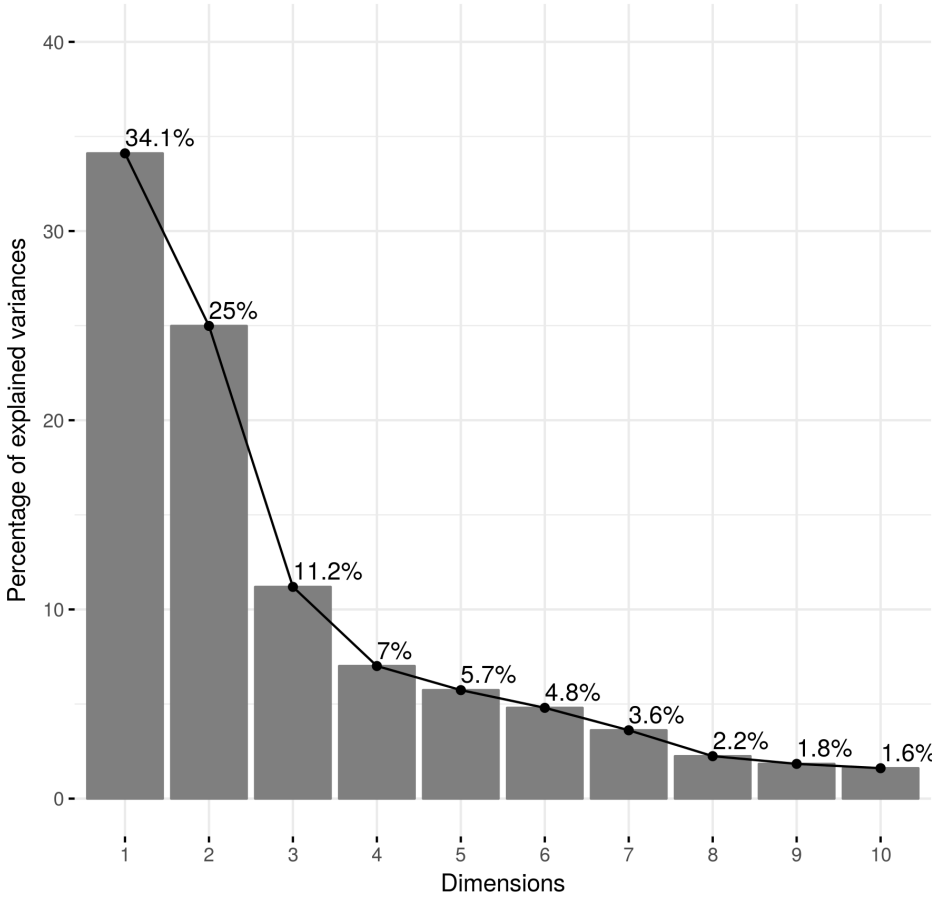


Figure S1 Scree plot of eigenvalues on the principal component on environmental variables

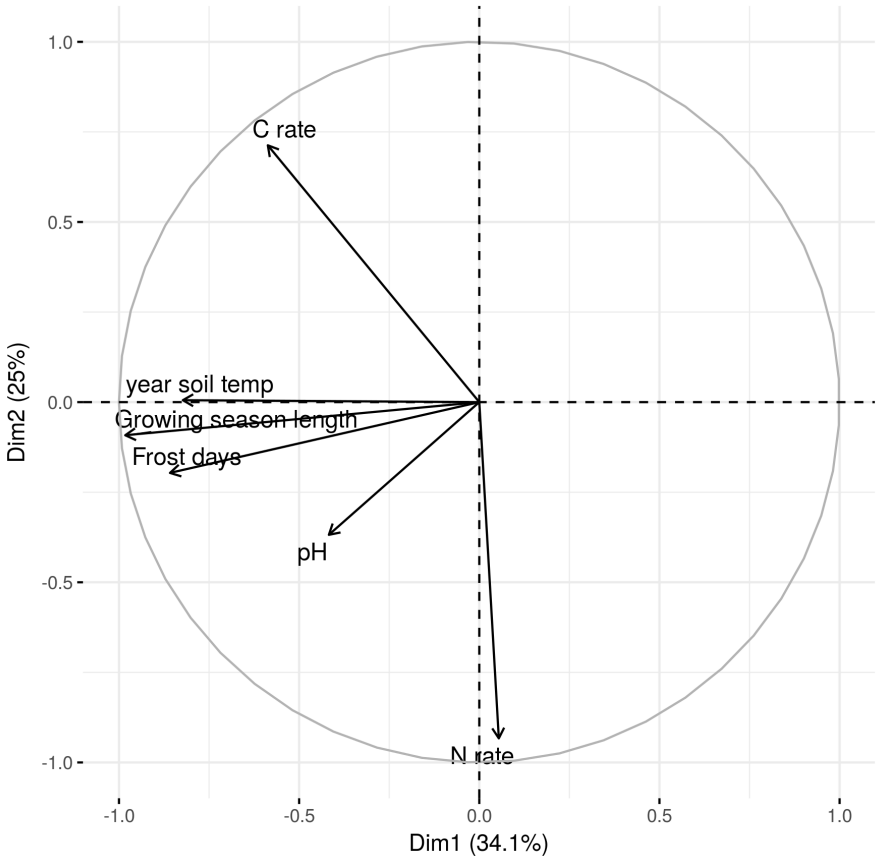


Figure S2 Correlation circles of the principal component analysis. Only 6 variable over the 24 variables are represented here.

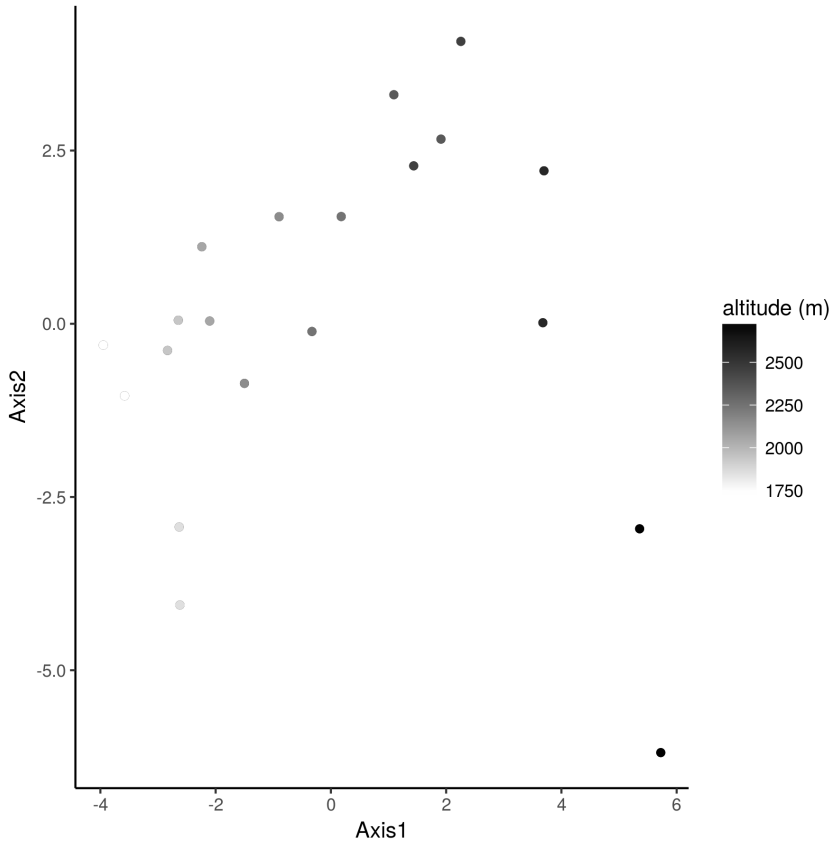


Figure S3 Plots represented on the first factorial plan on the principal component analysis. The color of the points correspond to their altitude.

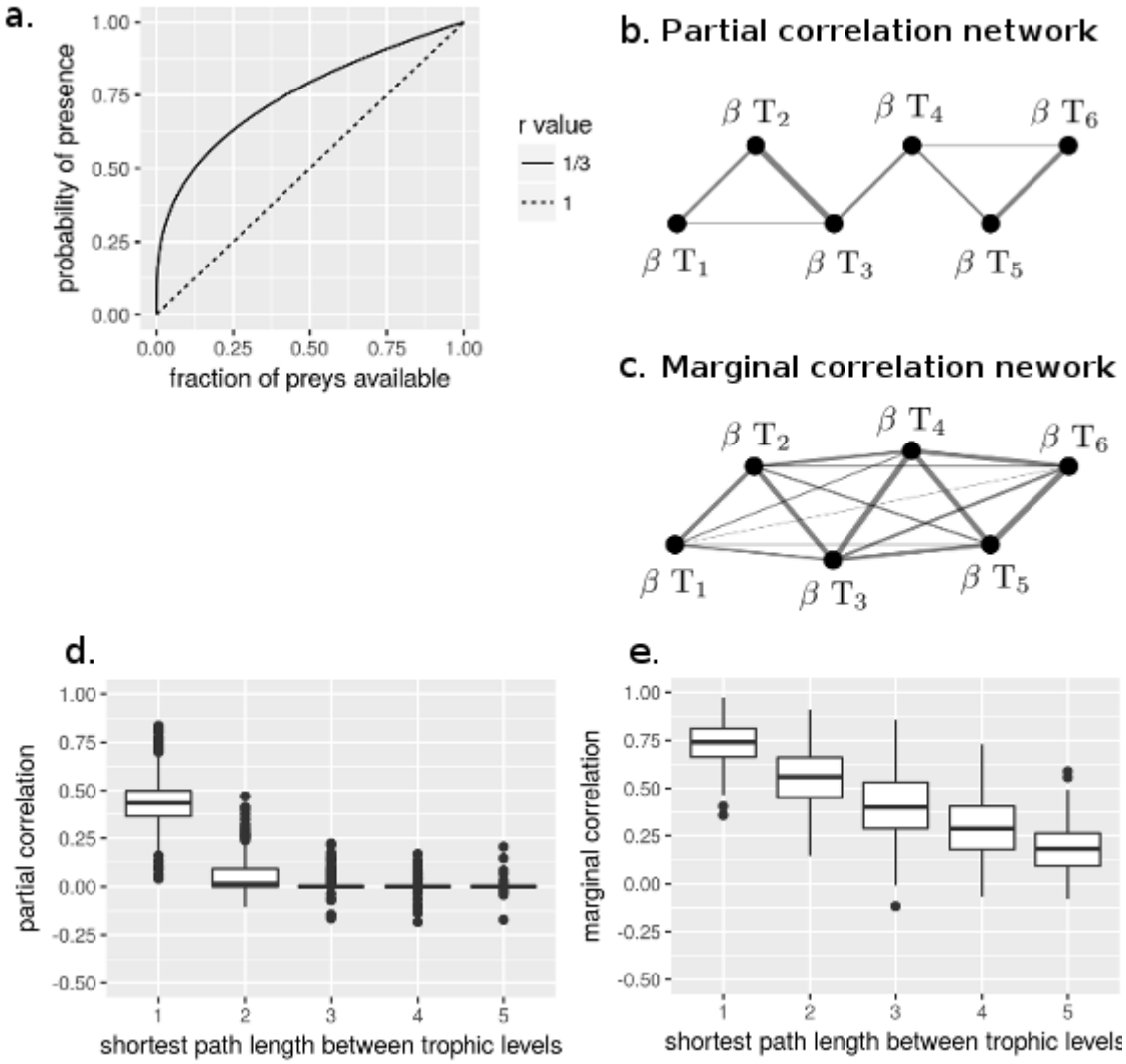


Figure S4 Design and results of the simulation assessing the impact of biotic interactions on the marginal and partial correlation of the *β*-diversity values of different trophic levels for another shape of the relationship linking the probability of presence of a consumer species with its number of available preys. *(a)* Two different shapes of the relationship, the value used in the following plots is r=1/3, while the main text r=1 (b) Partial correlation network built using the median values of partial correlations (inferred using Glasso) between the *β*-diversity values of different trophic levels. (c) Marginal correlation networks built using the median values of the Pearson correlations between the *β*-diversity values of different trophic levels. *(d)* Partial correlation coefficient (inferred using the graphical lasso) between the *β*-diversity values of two trophic levels in function of the shortest path length between these trophic levels. *(e)* Marginal Pearson correlation coefficient between the *β*-diversity values of two trophic levels in function of the shortest path length between these trophic levels.

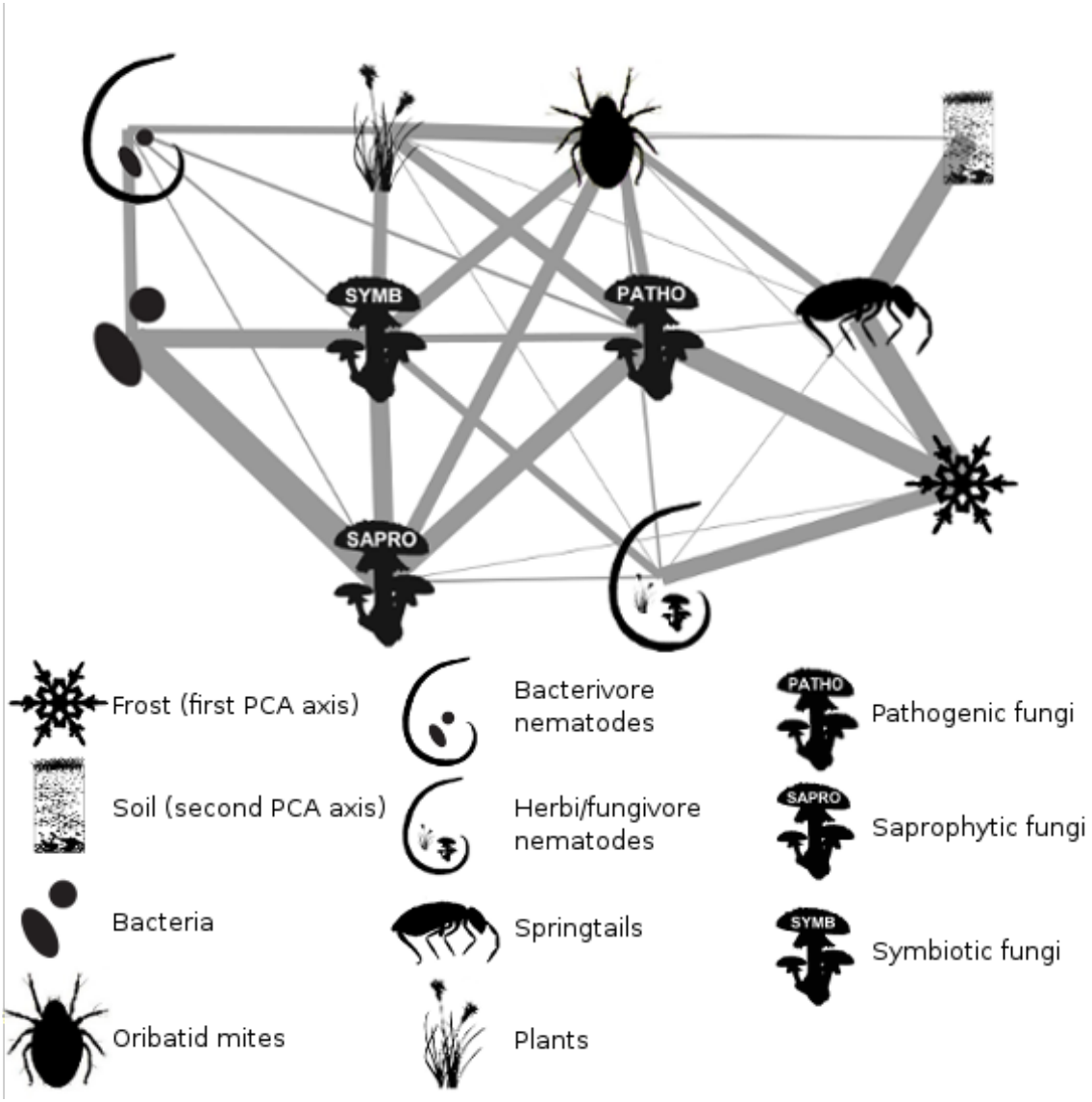


Figure S5 Undirected partial correlation network between β-diversity of the major trophic groups constituting of soil biodiversity and environmental distances. Each node represent the β-diversity of a trophic group or an abiotic environmental distance. Each edge represent a non-nul partial correlation between the variables associated to each node and was estimated using a graphical lasso. Consequently, if two nodes are disconnected, then the associated variables (β-diversity or environmental distances) are conditionally independent. The connectance of the network is 0.618.

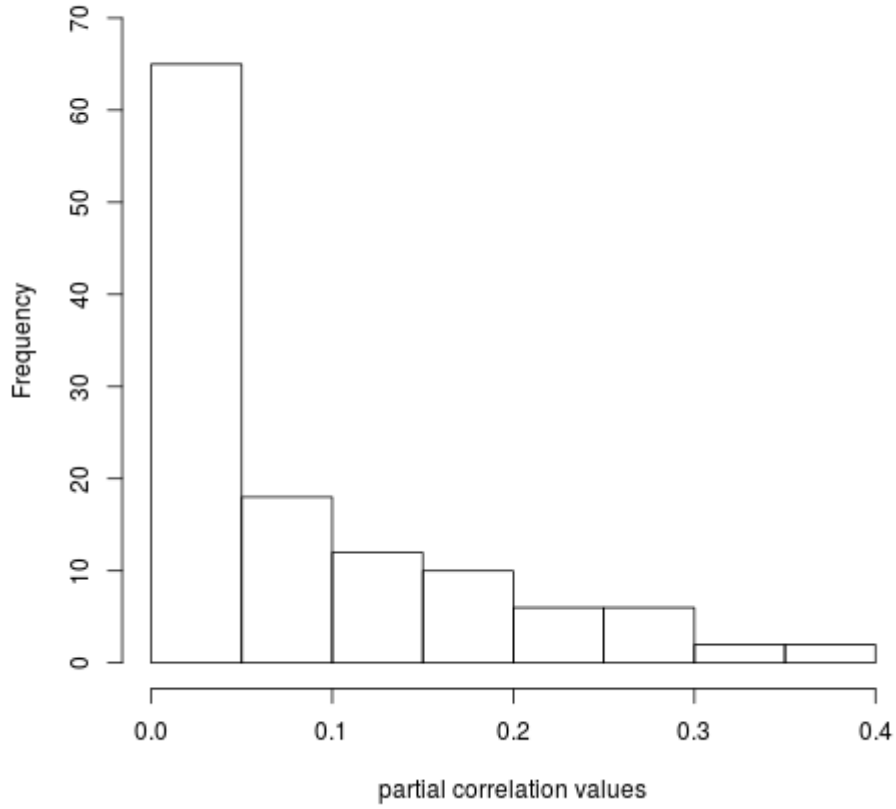


Figure S6 Histogram of the partial correlation coefficients estimated between β-diversity of major trophic groups constituting of soil biodiversity and environmental distances. Partial correlation coefficients are all positive.

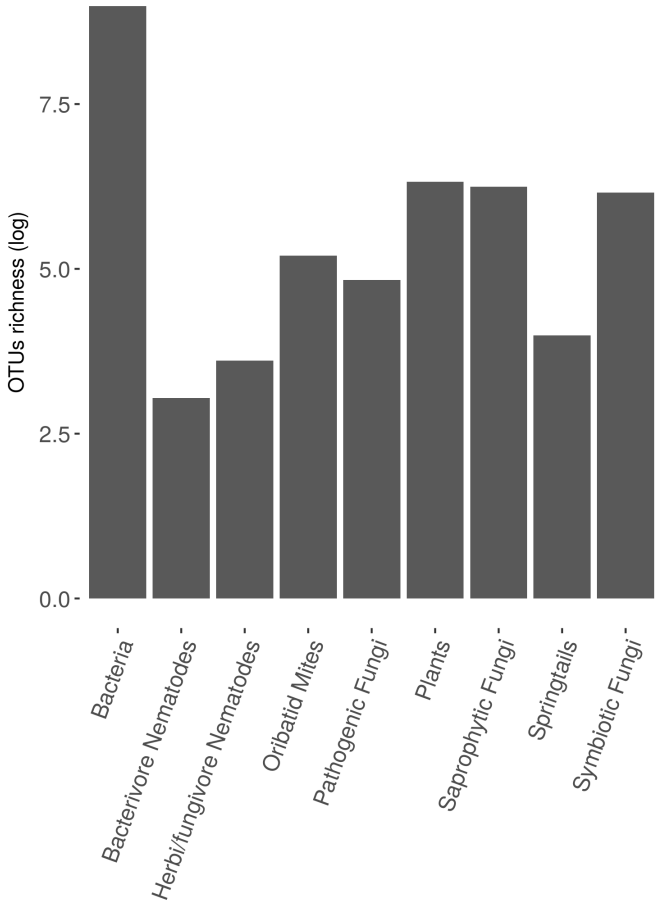


Figure S7 Histogram of the natural logarithm of the numbers of Operational Taxonomic Units (OTUs) assigned to each trophic group.

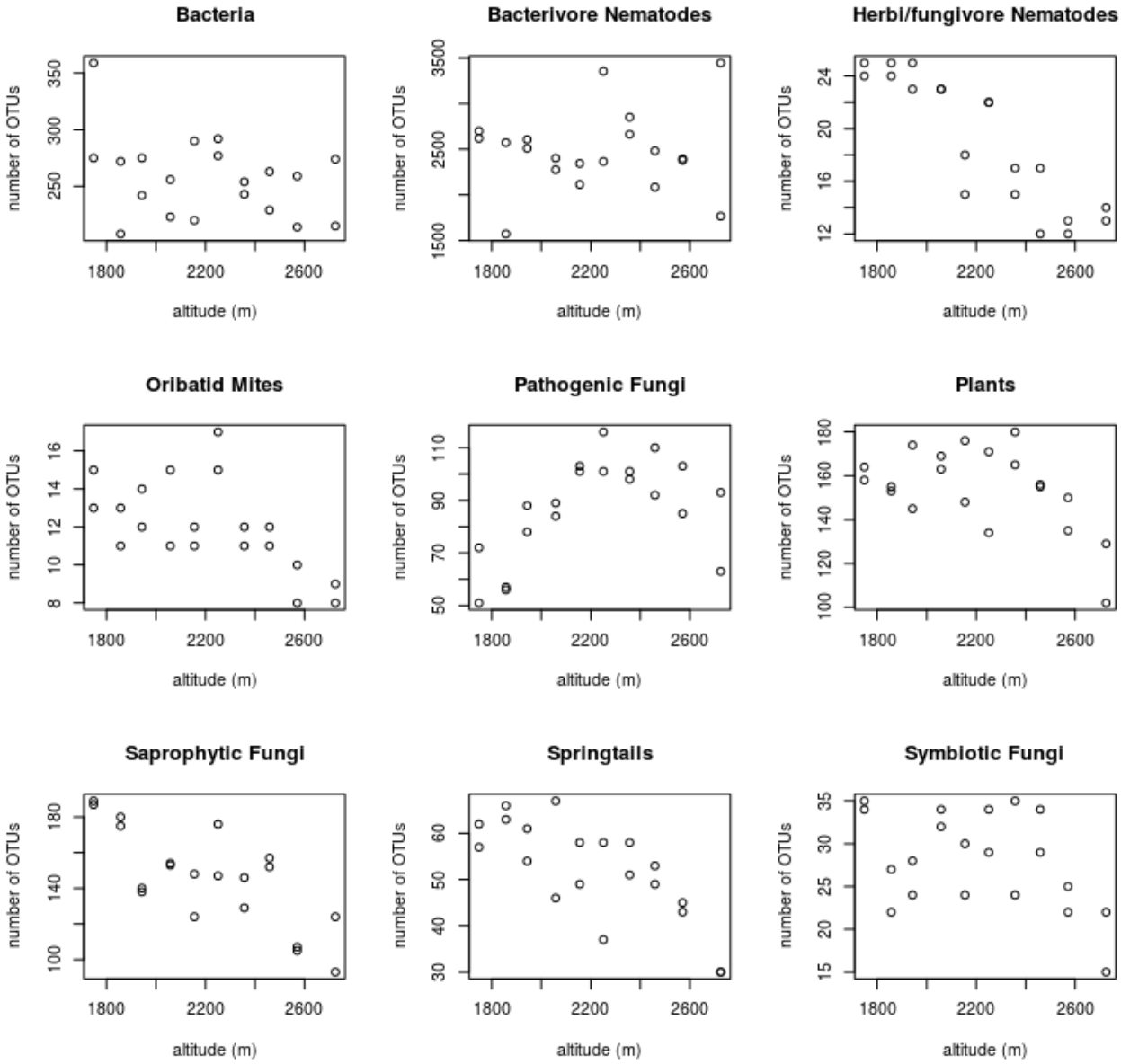


Figure S8 Number of OTUs per trophic group in function of the altitude of the plots

	Frost	Soil	Plants	Bacteria	Herbi/Fungivore nematodes	Bacterivore nematodes	Oribatida	Saprophytic fungi	Symbiotic fungi	Pathogenic fungi	Springtails
Frost	0	0	0	0	0.222	0	0.0358	0.0149	0	0.3	0.327
Soil	0	0	0.0347	0	0	0	0.0807	0	0	0	0.29
Plants	0	0.0347	0	0	0.0274	0.101	0.197	0.0914	0.169	0.166	0.0284
Bacteria	0	0	0	0	0	0.149	0	0.365	0.23	0	0
Herbi/Fungivore nematodes	0.222	0	0.0274	0	0	0.0833	0.084	0.084	0.11	0	0.0238
Bacterivore nematodes	0	0	0.101	0.149	0.0833	0	0	0.0708	0.0739	0.0994	0
Oribatida	0.0358	0.0807	0.197	0	0.084	0	0	0.185	0.182	0.14	0.116
Saprophytic fungi	0.0149	0	0.0914	0.365	0.084	0.0708	0.185	0	0.234	0.252	0
Symbiotic fungi	0	0	0.169	0.23	0.11	0.0739	0.182	0.234	0	0.102	0
Pathogenic fungi	0.3	0	0.166	0	0	0.0994	0.14	0.252	0.102	0	0.0541
Springtails	0.327	0.29	0.0284	0	0.0238	0	0.116	0	0	0.0541	0

Table S3 Partial correlation between β-diversity values of trophic groups constituting of soil diversity and abiotic environmental distances inferred using Glasso.

	Frost	Soil	Plants	Bacteria	Herbi/Fungivore Nematodes	Bacterivore nematodes	Oribatida	Saprophytic fungi	Symbiotic fungi	Pathogenic fungi	Springtails
Frost	1	0.305	0.581	0.405	0.648	0.497	0.672	0.678	0.575	0.772	0.703
Soil	0.305	1	0.392	0.235	0.325	0.0994	0.443	0.33	0.27	0.347	0.545
Plants	0.581	0.392	1	0.644	0.585	0.611	0.775	0.778	0.78	0.774	0.546
Bacteria	0.405	0.235	0.644	1	0.47	0.642	0.688	0.857	0.808	0.687	0.34
Herbi/Fungivore nematodes	0.648	0.325	0.585	0.47	1	0.531	0.64	0.664	0.657	0.626	0.513
Bacterivore nematodes	0.497	0.0994	0.611	0.642	0.531	1	0.573	0.675	0.656	0.645	0.346
Oribatida	0.672	0.443	0.775	0.688	0.64	0.573	1	0.836	0.818	0.808	0.622
Saprophytic fungi	0.678	0.33	0.778	0.857	0.664	0.675	0.836	1	0.888	0.864	0.553
Symbiotic fungi	0.575	0.27	0.78	0.808	0.657	0.656	0.818	0.888	1	0.81	0.476
Pathogenic fungi	0.772	0.347	0.774	0.687	0.626	0.645	0.808	0.864	0.81	1	0.628
Springtails	0.703	0.545	0.546	0.34	0.513	0.346	0.622	0.553	0.476	0.628	1

Table S4 Marginal (Pearson correlation) between β-diversity values of trophic groups constituting of soil diversity and abiotic environmental distances.

Appendix 3 : Statistical analysis using the true-turnover component of the β-diversity

We ran the statistical analysis on the true turnover component of the β-diversity (Baselga 2010) of each predefined trophic group. while the larger amount of β-diversity was due to true-turnover (Fig. 9) and on average, 80% of the β-diversity across trophic groups.The partial correlations estimated between the true turnover component of the β-diversity of each predefined trophic group and environmental distances were all positive except one coefficient (Fig. S10,Fig. S11, Fig. S12, table S4). The estimated partial correlation network had a connectance of 0.60 and was composed of 33 undirected edges (i.e. partial correlation coefficients>0) out of 55 possible edges and 11 nodes (9 trophic groups and 2 environmental variables). The mean value of non null partial correlation coefficients was 0.133, the median was 0.102 and the maximum 0.459.

Saprophytic fungi seemed to be the most influential group in conditioning the β-diversity of the others group (highest degree value, 9, and highest weighted degree value, 1.37, Fig. S12). Environmental variables had a moderate impact on the β-diversity of the trophic groups (degree of 6 for the first PCA axis, 4 for the second PCA axis and whereas the mean degree of the partial correlation network was 6; weighted degree of 1.13 for the first PCA-axis, 0.381 for the second pca axis node whereas the mean value is 0.759).

The probability of observing a non-null partial correlation between the β-diversity of a trophic group and the environmental distance was 0.44 (8 edges linking environmental nodes to the trophic groups nodes and 18 potential edges) whereas the probability of observing a link between the β-diversity of any two trophic groups was 0.67 (24 edges and 36 potential edges), highlighting the weaker influence of environmental variables on the true turnover component of the β-diversity of the trophic groups.

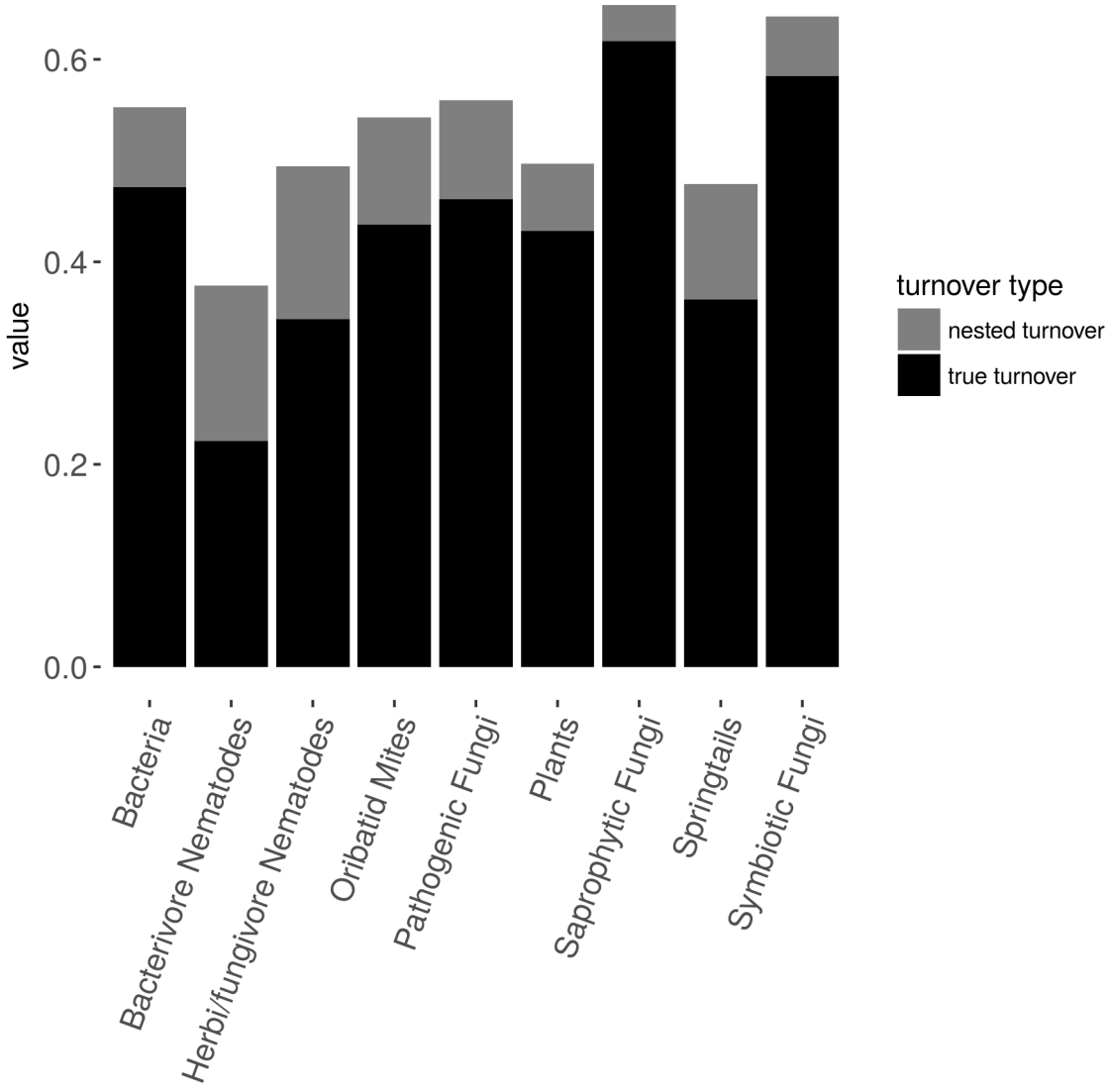


Figure S9 Mean true turnover and nested turnvoer components of the Jaccard index for each trophic group.

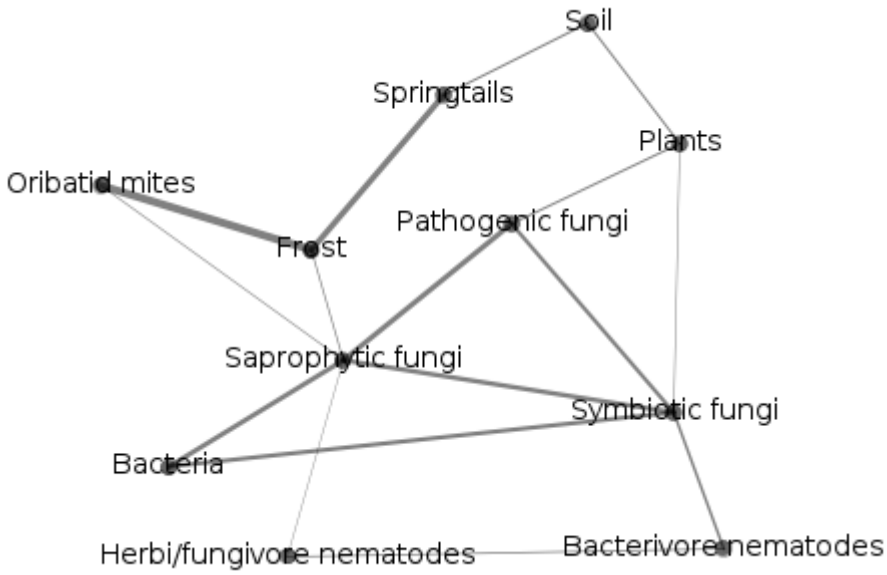


Figure S10 Undirected partial correlation network inferred by the graphical lasso method between the true turnover component of the β-diversity of the major trophic groups constituting of soil biodiversity and the environmental distances. Each node represent the β-diversity of a trophic group or an environmental distance. Here, are only shown the partial correlations above the median value of the non-null partial correlation coefficients (0.102). The non-filtered network is presented in Figure S11. Edge thickness is proportional to the value of the partial correlation coefficient and the partial correlation coefficients are all positive.

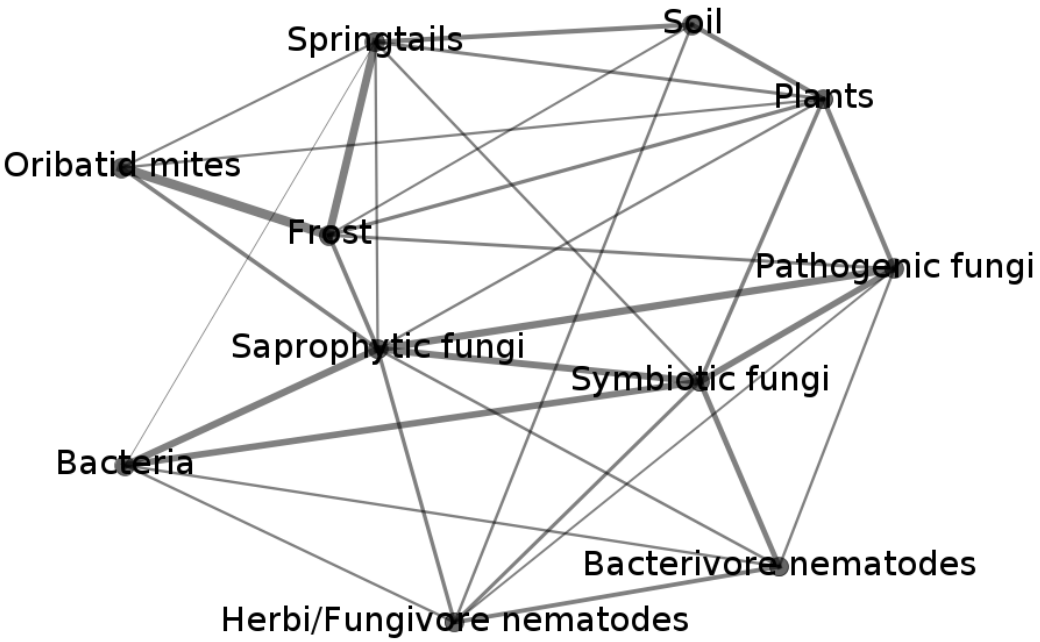


Figure S11 Undirected partial correlation network inferred by the graphical lasso method between the true turnover component of the β-diversity of the major trophic groups constituting of soil biodiversity and the environmental distances. Each node represent the β-diversity of a trophic group or an environmental distance. The connectance of the graph is 0.600 . Edge thickness is proportional to the value of the partial correlation coefficient and the partial correlation coefficients are all positive.

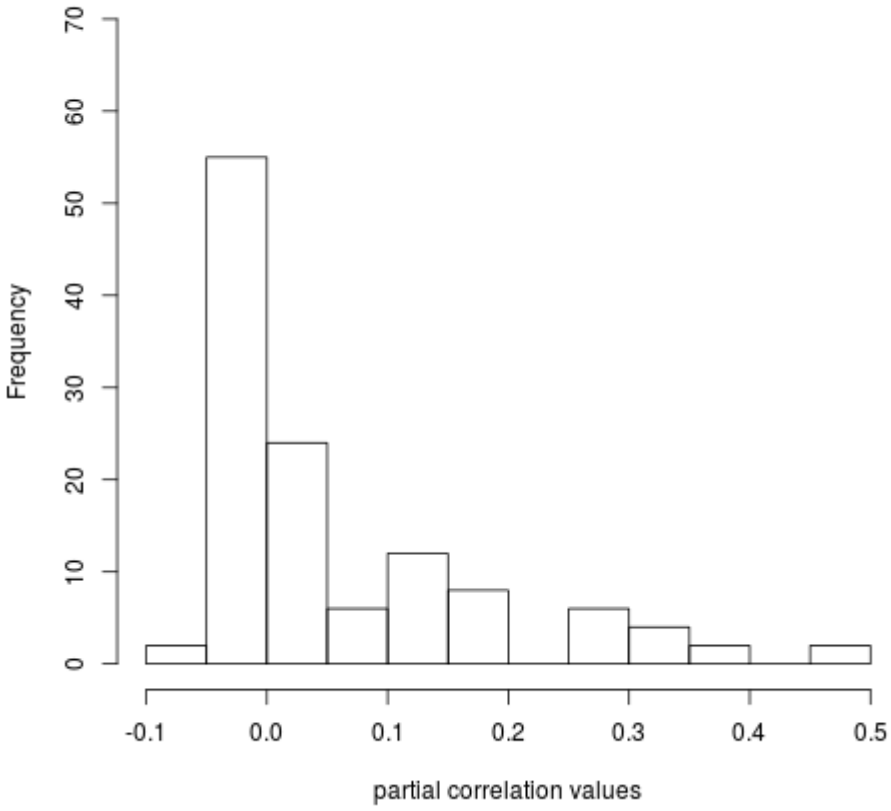


Figure S12 Histogram of the partial correlation coefficients estimated between the true turnover component of the β-diversity of major trophic groups constituting of soil biodiversity and environmental distances. Partial correlation coefficients are all positive.

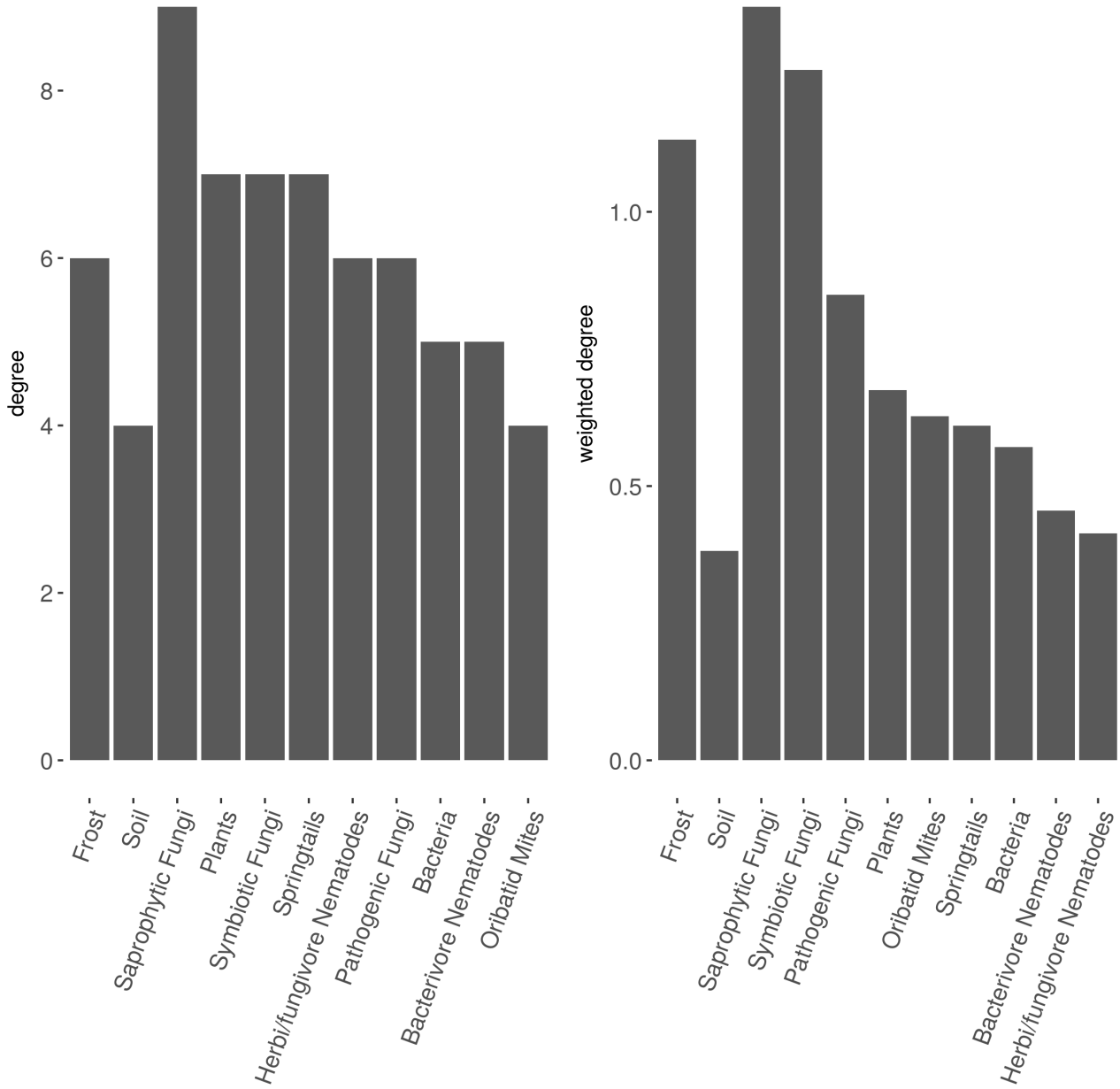


Figure S13 Properties of the inferred network. The mean degree (left panel) is the number of neighbors of nodes in a graph (here, the undirected partial correlation network). It measures the number of variables that are conditionally dependent to the variable associated to this node. The weighted degree (right panel) is the sum of the partial correlation coefficients attached to the edges adjacent this node.

	Frost	Soil	Plants	Bacteria	Herbi/Fungivore nematodes	Bacterivore nematodes	Oribatida	Saprophytic fungi	Symbiotic fungi	Pathogenic fungi	Springtails
Frost	0	0.0122	0.1	0	0	0	0.459	0.129	0	0.0725	0.359
Soil	0.0122	0	0.0161	0	0.0442	0	0	0	0	0	0.164
Plants	0.1	0.161	0	0	0	0	0.0289	0.0314	0.112	0.162	0.0796
Bacteria	0	0	0	0	0.0327	0.0354	0	0.0302	0.275	0	-0.0744
Herbi/Fungivore nematodes	0	0.0422	0	0.0327	0	0.139	0	0.104	0.0896	0.00466	0
Bacterivore nematodes	0	0	0	0.0354	0.139	0	0	0.0457	0.2	0.0353	0
Oribatida	0.459	0	0.0289	0	0	0	0	0.122	0	0	0.0181
Saprophytic fungi	0.129	0	0.0314	0.302	0.104	0.0457	0.122	0	0.292	0.321	0.0268
Symbiotic fungi	0	0	0.112	0.275	0.0896	0.2	0	0.292	0	0.254	0.0366
Pathogenic fungi	0.0725	0	0.162	0	0.00466	0.0353	0	0.321	0.254	0	0
Springtails	0.359	0.164	0.0796	-0.0744	0	0	0.0181	0.0268	0.0366	0	0

Table S5 Partial correlation between the true turnover component of β-diversity values of trophic groups constituting of soil diversity and abiotic environmental distances inferred using Glasso.

	Frost	Soil	Plants	Bacteria	Herbi/Fungivore Nematodes	Bacterivore nematodes	Oribatida	Saprophytic fungi	Symbiotic fungi	Pathogenic fungi	Springtails
Frost	1	0.305	0.503	0.313	0.102	0.117	0.734	0.619	0.471	0.551	0.637
Soil	0.305	1	0.363	0.0432	0.216	-0.0226	0.249	0.255	0.207	0.158	0.367
Plants	0.503	0.363	1	0.276	0.231	0.235	0.416	0.535	0.54	0.557	0.417
Bacteria	0.313	0.0432	0.276	1	0.395	0.439	0.346	0.743	0.732	0.593	0.0491
Herbi/Fungivore nematodes	0.102	0.216	0.231	0.395	1	0.398	0.0963	0.468	0.467	0.396	0.158
Bacterivore nematodes	0.117	-0.0226	0.235	0.439	0.398	1	0.196	0.499	0.557	0.456	0.0921
Oribatida	0.734	0.249	0.416	0.346	0.0963	0.196	1	0.556	0.448	0.463	0.461
Saprophytic fungi	0.619	0.255.	0.535	0.743	0.468	0.499	0.556	1	0.835	0.799	0.422
Symbiotic fungi	0.471	0.207	0.54	0.732	0.467	0.557	0.448	0.835	1	0.768	0.382
Pathogenic fungi	0.551	0.158	0.557	0.593	0.396	0.456	0.463	0.799	0.768	1	0.331
Springtails	0.637	0.367	0.417	0.0491	0.158	0.0921	0.461	0.422	0.382	0.331	1

Table S6 Marginal (Pearson correlation) between the true turnover component of β-diversity values of trophic groups constituting of soil diversity and abiotic environmental distances.

References

Abarenkov, K., Henrik Nilsson, R., Larsson, K.-H., Alexander, I. J., Eberhardt, U., Erland, S. et al. (2010). The UNITE database for molecular identification of fungi - recent updates and future perspectives. *New Phytologist*, *186*, 281–285.

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, *16*, 176–182.

Mercier, C., Boyer, F., Bonin, A., & Coissac, E. (2013). SUMATRA and SUMACLUST: fast and exact comparison and clustering of sequences. *Programs and Abstracts of the SeqBio 2013 Workshop, 27-29*.

Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 1118–23.

Zinger, L., Taberlet, P., Schimann, H., Bonin, A., Boyer, F., De Barba, M. *et al.* (2017). Soil community assembly varies across body sizes in a tropical forest. *bioRxiv* 154278; doi: https://doi.org/10.1101/154278