

Body size determines soil community assembly in a tropical forest

Lucie Zinger¹  | Pierre Taberlet²  | Heidy Schimann³  | Aurélie Bonin² |
 Frédéric Boyer²  | Marta De Barba² | Philippe Gaucher⁴ | Ludovic Gielly² |
 Charline Giguet-Covex²  | Amaia Iribar¹ | Maxime Réjou-Méchain^{1,5}  | Gilles Rayé² |
 Delphine Rioux² | Vincent Schilling¹ | Blaise Tymen¹  | Jérôme Viers⁶ | Cyril Zouiten⁶ |
 Wilfried Thuiller²  | Eric Coissac²  | Jérôme Chave¹ 

¹CNRS, IRD, UMR 5174 Evolution et Diversité Biologique (EDB), Université Toulouse 3 Paul Sabatier, Toulouse, France

²CNRS, Laboratoire d'Ecologie Alpine (LECA), Univ. Grenoble Alpes, Grenoble, France

³UMR Ecologie des Forêts de Guyane (AgroParisTech, CIRAD, CNRS, Université des Antilles, Université de la Guyane), INRA, Kourou, France

⁴USR 3456, CNRS, Centre de recherche de Montabo IRD, CNRS-Guyane, Cayenne, France

⁵UMR AMAP, IRD, Montpellier, France

⁶CNRS, IRD, UMR 5563 GET, Université Toulouse 3 Paul Sabatier, Toulouse, France

Correspondence

Lucie Zinger, CNRS, IRD, UMR 5174 Evolution et Diversité Biologique (EDB), Université Toulouse 3 Paul Sabatier, Toulouse, France.
 Email: lucie@zinger.fr

Present addresses

Lucie Zinger, Institut de biologie de l'Ecole normale supérieure (IBENS), Ecole normale supérieure, CNRS, INSERM, PSL Research University, Paris, France.

Charline Giguet-Covex, CNRS, Ministère de la Culture et de la Communication, EDYTEM, Univ. Savoie Mont Blanc, Chambéry, France.

Funding information

Agence Nationale de la Recherche, Grant/Award Number: ANAEE-France: ANR-11-INBS-0001, CEBA: ANR-10-LABX-25-01, METABAR: ANR-11-BSV7-0020, OSUG@2020: ANR-10-LABX-56 and TULIP: ANR-10-LABX-0041

Abstract

Tropical forests shelter an unparalleled biological diversity. The relative influence of environmental selection (i.e., abiotic conditions, biotic interactions) and stochastic-distance-dependent neutral processes (i.e., demography, dispersal) in shaping communities has been extensively studied for various organisms, but has rarely been explored across a large range of body sizes, in particular in soil environments. We built a detailed census of the whole soil biota in a 12-ha tropical forest plot using soil DNA metabarcoding. We show that the distribution of 19 taxonomic groups (ranging from microbes to mesofauna) is primarily stochastic, suggesting that neutral processes are prominent drivers of the assembly of these communities at this scale. We also identify aluminium, topography and plant species identity as weak, yet significant drivers of soil richness and community composition of bacteria, protists and to a lesser extent fungi. Finally, we show that body size, which determines the scale at which an organism perceives its environment, predicted the community assembly across taxonomic groups, with soil mesofauna assemblages being more stochastic than microbial ones. These results suggest that the relative contribution of neutral processes and environmental selection to community assembly directly depends on body size. Body size is hence an important determinant of community assembly rules at the scale of the ecological community in tropical soils and should be accounted for in spatial models of tropical soil food webs.

KEYWORDS

DNA metabarcoding, eDNA, French Guiana, multitaxa, neutral assembly, niche determinism, propagule size, soil diversity

1 | INTRODUCTION

During community assembly, ecological communities build up to utilize environmental resources, assemble into trophic webs and co-exist in time. Major determinants of community assembly include selection by abiotic and biotic environmental factors, ecological drift (i.e., demographic processes) and dispersal (Hubbell, 2001; Leibold et al., 2004; Ricklefs, 2004; Vellend, 2010). Their relative contribution has been the subject of a vast literature, and there is growing awareness that their influence depends on the spatial scale of observation (Leibold et al., 2004; Levin, 1992; Ricklefs, 2004). For example, the influence of environmental selection has been shown to increase as the observational scale increases, due to the inclusion of greater environmental heterogeneity (Barton et al., 2013; Chase, 2014; De Cáceres et al., 2012; Dini-Andreote, Stegen, van Elsas, & Salles, 2015). However, much less is known on how community assembly processes vary across organisms, and especially when it comes to the whole soil biota.

Soils are structurally complex environments (Hinsinger, Bengough, Vetterlein, & Young, 2009) and harbour a large and elusive diversity of organisms (Bardgett & van der Putten, 2014). Amplicon-based DNA analysis of environmental samples or metabarcoding (Taberlet, Coissac, Pompanon, Brochmann, & Willerslev, 2012) has recently enabled unravelling novel macroecological patterns for soil fauna, nematodes, bacteria, fungi, protists and archaea across biomes and habitats (Bardgett & van der Putten, 2014; Fierer, Strickland, Liptzin, Bradford, & Cleveland, 2009; Mahé et al., 2017; Tedersoo et al., 2014; Wu, Ayres, Bardgett, Wall, & Garey, 2011). These patterns often covary with environmental conditions such as pH and nutrient quality/availability. They also depend on plant cover due to trophic and mutualistic/pathogenic interactions (Dini-Andreote & van Elsas, 2013; Lavelle et al., 1997; Wardle, 2002), but this relationship appears context-dependent (Barberán et al., 2015; Peay, Baraloto, & Fine, 2013; Tedersoo et al., 2016). More generally, soil communities are often reported to be driven by environmental selection (Bardgett & van der Putten, 2014; Hanson, Fuhrman, Horner-Devine, & Martiny, 2012), although certain studies also unveiled dispersal limitation and/or ecological drift in fungi, mesofauna and bacteria to a lesser extent (Bahram et al., 2016; Dumbrell, Nelson, Helgason, Dytham, & Fitter, 2010; Peay, Garbelotto, & Bruns, 2010; Powell et al., 2015; Stegen, Lin, Konopka, & Fredrickson, 2012).

Much of the difficulty in drawing general conclusions about soil community assembly lies in that empirical studies often focus on single taxonomic groups (but see Ramirez et al., 2014; Powell et al., 2015; Schuldt et al., 2015; Tedersoo et al., 2016). Soil organisms exhibit large differences in life history traits, in particular their body size, which spans six orders of magnitude (0.1 μm to 10 cm; Bardgett & van der Putten, 2014). Organisms of different body size have contrasting metabolisms, ontogenies, population sizes and diversities (Brown, Gillooly, Allen, Savage, & West, 2004; Woodward et al., 2005). Body size is also indicative of organism trophic status in its broadest sense (Briones, 2014; Brose et al., 2012) and is a key trait in the modelling of food webs (Petchey,

Beckerman, Riede, & Warren, 2008; Williams & Martinez, 2000). Finally, body size also influences the way organisms perceive—or move across—space (Berg, 2012; Ettema & Wardle, 2002; Hubbell, 2001). This suggests that community assembly processes depend not only on the spatial scale of the study, but also on the size of the focal organisms.

For example, dispersal of microorganisms is mediated by external agents and is assumed to occur across large spatial distances in most cases. This, together with their large population size and short generation time, contributes to creating a homogeneous regional microbial species pool. Local microbial communities are then selected from the regional species pool by local environmental conditions (Barberán, Casamayor, & Fierer, 2014; Hanson et al., 2012), although this selection can be offset by high immigration rates, a process previously coined “mass effect” or “homogenizing dispersal” (Leibold et al., 2004; Stegen et al., 2013). In contrast, larger organisms, such as trees or mesofauna, are thought to be limited in their ability to disperse, have smaller population sizes and have longer generation times. They hence tend to lag behind their optimal habitat conditions and are more prone to local extinction. In addition, they may not perceive small-scale environmental variations as compared to microorganisms (Berg, 2012; Ettema & Wardle, 2002; Woodward et al., 2005). As a consequence, their distribution patterns depend less, or not at all, on local environmental conditions (Bie et al., 2012 and references within). Hence, body size differences may have enormous consequences for the spatial scaling of soil biodiversity and its associated major biogeochemical cycles and ecosystem services (Bardgett & van der Putten, 2014; Lavelle et al., 1997; Wardle, 2002). However, existing knowledge on the scaling relationship between body sizes and soil communities' assembly processes is often based on meta-analyses and is therefore indirect (Berg, 2012; Ettema & Wardle, 2002; Lavelle et al., 1997).

Here, we addressed the question of how body size determines community assembly in a tropical forest, by assessing soil community assembly across taxonomic groups spanning several body size classes. In these ecosystems, a large part of our knowledge on spatial patterns of biodiversity and their underlying processes comes from large permanent forest plots (Anderson-Teixeira et al., 2015; Condit et al., 2002; Hubbell, 2001). In that spirit, we assessed the processes governing soil community assemblages by combining an extensive characterization of abiotic and plant cover conditions with a comprehensive survey of soil biodiversity using DNA metabarcoding in a 12-ha plot, where samples were collected every ten metres. We sought to (a) determine which factors, both abiotic (e.g., soil chemistry) and biotic (plant diversity, identity), influence soil community composition, (b) evaluate the relative importance of environmental selection versus ecological drift in shaping soil communities and (c) determine how these effects relate to taxon body size. We hypothesized that in communities of small-bodied organisms, environmental selection should be more important relative to larger organisms, which should instead display spatial patterns resulting from ecological drift.

2 | MATERIALS AND METHODS

2.1 | Study site and sampling

The study site is located at the Nouragues Ecological Research Station, in the lowland rain forest of French Guiana (latitude: 4°4'28"N, longitude: 52°40'45"W, Figure 1). Rainfall is 2861 mm/year (average: 1992–2012), with a two-month dry season (<100 mm/month), from late August to early November, and a shorter dry season in March. Our sampling campaign was conducted on November 7–20, 2012, towards the end of the dry season, which lasted from early September to late November. Cumulative rainfall during the 60 days preceding the sampling session was 134 mm, with 44 days without rain, and over 90% of the rainfall concentrated in seven days.

We surveyed a 12-ha (300 × 400 m) plot established in 1992. This plot extends on a gentle slope (7.5°) between a ridge and a small

creek (Chave et al., 2008). The 5,640 trees occurring in the plot (diameter at breast height ≥10 cm) belong to over 600 species, with the two dominant species accounting each for only 2.3% of individuals (Baraloto et al., 2012). Sand and clay fractions are about 40% each in the soil top 10 cm. The parent material is Caribbean granite. Soil edaphic conditions are typical of tropical oxisols, with an acidic pH (pH = 5.0) and low exchangeable cation content (ECEC = 3.5 cmolc/kg). The C:N and N:P ratios are of 13.4 and 40.5, respectively (median values), typical of tropical forests (Turner & Engelbrecht, 2011).

We sampled the plot following a regular grid scheme with a 10-m resolution grid, excluding bordering points, hence resulting in a total of 1,132 sampling points. At each point, 50–100 g soil cores were collected with an auger at ~10 cm depth, excluding the O-horizon. We did so because the organic and mineral horizons harbour different arthropods and microbial communities (Baldrian et al., 2012; Yang et al., 2014). Lumping together these

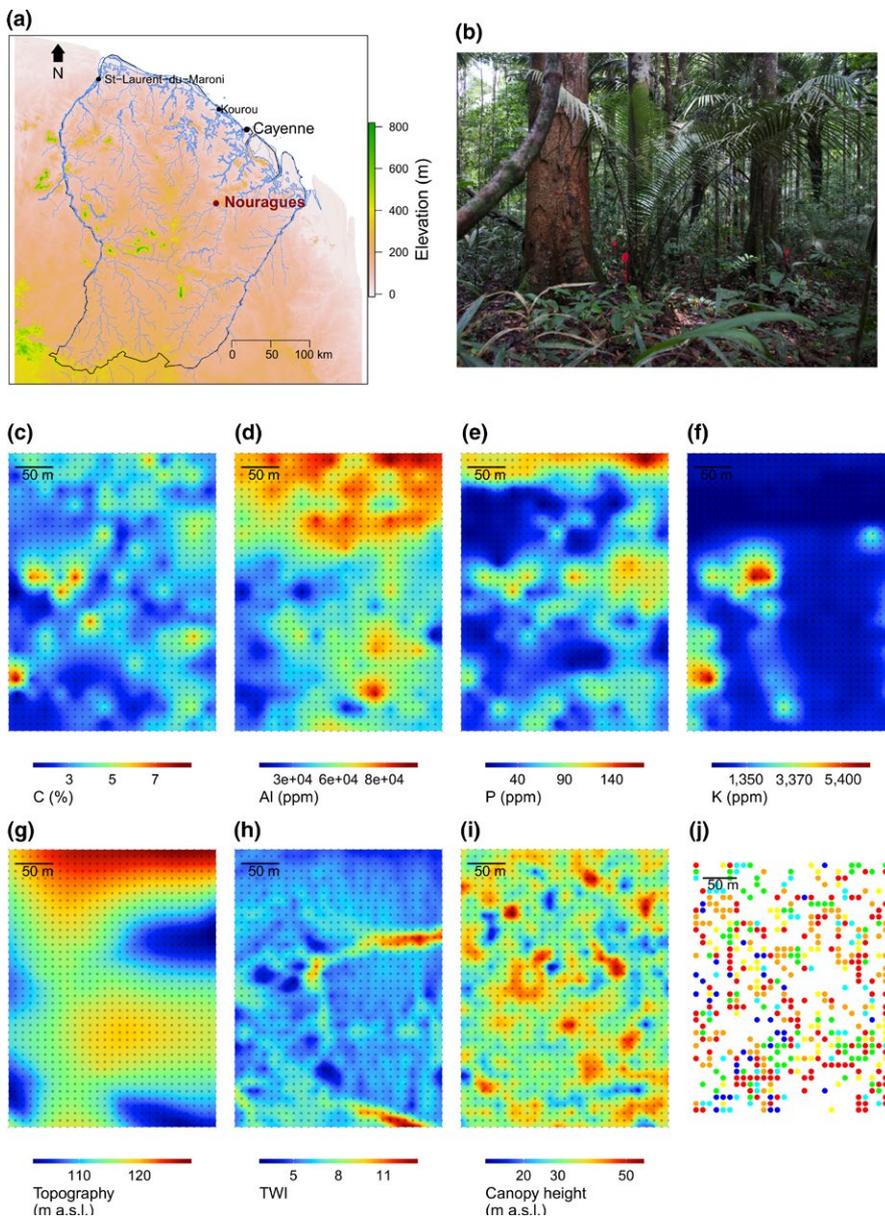


FIGURE 1 Study site description. (a) Geographic location (red) of the study site in South America (top right) and in French Guiana. Elevation and hydrologic networks were retrieved from the Worldclim data set (www.worldclim.org) and GIS Guiana layers (gisguyane.brgm.fr). (b) Photography of the studied plot (12 ha, 300 × 400 m). Sampling points (every 10 m) are marked with red sticks. Photo credit: P. Taberlet. (c–f) Soil content in carbon, aluminium, phosphorus and potassium. (g) Site topography, (h) soil wetness (TWI index, unitless) and (i) canopy height as inferred from Lidar data. (j) Distribution of the most dominant plant taxon in each soil sample as inferred from the plant molecular data set. Only the six most frequent dominant taxon are shown: red: Apocynoideae; orange: Andira; yellow: Ingeae; green: Brosimum; light blue: Sapotaceae; blue: Drypetes. x-axis and y-axis are not shown for representation purposes and range between 10 and 290 and 10 and 390 m, respectively

compartments may hence complicate the interpretation of the spatial distribution of certain soil clades. Consequently, we focused on the surface soil layer, which is the most biogeochemically active in the mineral horizon (Silver et al., 2000). The soil cores were stored and sealed in sterile plastic bags after collection and transported to the field station laboratory. Extracellular DNA was extracted twice from 15 g of soil per soil core as described previously (Taberlet, Prud'homme et al., 2012; Zinger et al., 2016) within 4 hours after sample collection to prevent microbial growth. This method was used due to its easiness of implementation for large-scale samplings. It also yields ecological signals similar to those obtained with total DNA, although it may under-sample rare taxa (Zinger et al., 2016). Its use for tropical soils is relevant because both microorganisms and macroorganisms continuously release DNA in the environment that is most likely recycled rapidly in tropical soils, but persistent enough to buffer microbial diel-scale dynamics (Taberlet, Bonin, Zinger, & Coissac, 2018). Soil DNA was extracted twice for each soil core, and the remaining soil material was dried and stored for analytical chemistry analyses.

2.2 | Molecular analyses

Soil biodiversity was surveyed through DNA metabarcoding using three DNA markers, with primers targeting three hypervariable regions of the *ssu* rRNA gene in Archaea, Bacteria and Eukaryota domains, respectively (see Table 1). The corresponding primers have not been reported to present amplification biases (see Supporting information Appendix S1: Material S1 for a detailed description of priming site conservatism across phyla). We conducted duplicated

PCRs for each marker and each DNA extract, hence representing a total of 18,112 independent PCRs. To discriminate PCR products after sequencing, forward and reverse primers were tagged with a combination of two different 8-nucleotide labels. PCR and sequencing protocols are provided in the Supporting information Appendix S1: Methods. To control for potential contaminants (Salter et al., 2014) and false positive caused by tag-switching events (Esling, Lejzerowicz, & Pawlowski, 2015), the sequenced multiplexes comprised extractions/PCR blank controls, unused tag combinations and positive controls for which specific analyses are provided in Supporting information Appendix S1: Material S2.

2.3 | Sequence analyses and curation

About one billion sequencing reads were produced and curated mainly using the *OBITOOLS* package (Boyer et al., 2016) and the *R* software v.3.3.2 (R Core Team, 2016) following the protocols provided in the Supporting information Appendix S1: Methods (see also Supporting information Appendix S1: Material S2 for a validation of the data curation pipeline). Briefly, paired-end reads were assembled, assigned to their respective samples/marker and dereplicated. Low-quality sequences were excluded; the remaining ones were clustered into operational taxonomic units (OTUs) and assigned to taxonomic clade. We paid particular attention to minimize PCR/sequencing errors, contaminant and false-positive sequences and potential nonfunctional PCRs by using conservative quality check criteria. Raw and curated data set statistics are reported in Supporting information Appendix S1: Table S1. Finally, the sequencing depth of each sample was standardized for each marker by randomly resampling

TABLE 1 Barcode characteristics and associated PCR conditions

	Archaea	Bacteria	Eukaryota	Viridiplantae
Primers set				
Targeted region	V8-V9 regions, 16S rRNA gene	V5-V6 regions, 16S rRNA gene	V7 region, 18S rRNA gene	trnL-P6 loop, chloroplastic DNA
Forward primer (5'-3')	CCTGCTCCTTGACACAC ^a	GGATTAGATACCTGGTAGT ^b	TTTGTCTGTTAATTSCG ^c	GGGCAATCCTGAGCCAA ^d
Reverse primer (5'-3')	CCTACGGCTACCTGTTC ^a	CACGACACGAGCTGACG ^b	TCACAGACCTGTTATTGC ^c	CCATTGAGTCTCTGCACCTATC ^d
PCR conditions				
Number of cycles	40	35	45	40
PCR cycle characteristics ^e (denaturation, annealing, elongation—final elongation)	95°C (30 s), 58°C (30 s), 72°C (90 s)—72°C (7 min)	95°C (30 s), 57°C (30 s), 72°C (90 s)—72°C (7 min)	95°C (30 s), 45°C (30 s), 72°C (60 s)—72°C (7 min)	95°C (30 s), 50°C (30 s), 72°C (60 s)—72°C (7 min)
Illumina platform	HiSeq 2×100	MiSeq 2×250	HiSeq 2×150	HiSeq 2×100

^aThis study. ^bFliegerova et al. (2014). ^cGuardiola et al. (2015). ^dTaberlet, Gielly, Pautou, & Bouvet (1991). ^eAll PCRs were preceded by an initial denaturation at 95°C (10 min).

a number of reads equal to the first quartile of read number across samples. Although such procedure has been recently questioned (McMurdie & Holmes, 2014), the data loss caused by standardization is here minimal and had no or weak effects on the data set characteristics and the retrieved patterns of diversity compared to raw data sets (Supporting information Appendix S1: Figure S1).

2.4 | Focus taxonomic groups and body size

Within the Eukaryota and Bacteria domains, we distinguished groups on the basis of their taxonomic affiliation at the phylum or class level (Table 2). We did so because broadly defined functional traits such as body size and trophic categories are relatively well conserved within phyla (Briones, 2014; Martiny, Jones, Lennon, & Martiny, 2015). This approach circumvents the difficulties in retrieving body size for each OTU, for which fine-level taxonomic assignment and/or morphological descriptions are seldom available. We restricted our analysis to the most abundant phyla (i.e., representing $\geq 1\%$ of the total bacterial or eukaryotic OTU diversity). The OTUs

unambiguously assigned to nonfungal unicellular eukaryote clades (e.g., Stramenopiles, Rhizaria, Alveolata, Cercozoa, Amoebozoa) constituted only a low fraction of the eukaryotic data set (1.95%). To include in our analysis this often-neglected, yet important part of the soil biota (Geisen et al., 2018), we therefore considered them as one single group, hereafter referred to as “protists.” Archaea were analysed as a single group due to imprecise taxonomic assignments.

Average body sizes per focus taxonomic group were inferred from Portillo, Leff, Lauber, and Fierer (2013) for bacterial cells, from Bécard and Pfeffer (1993) and Ingold (2001) for fungal propagules, and from Briones (2014) for the other groups (Table 2). Body size classes were defined as the size of the dispersal unit (propagule), rather than that of the mature individual. In fungi, spore size can be used as the operational definition of body size (Peay, Kennedy, & Bruns, 2008). Indeed, mature bodies may extend over large areas through mycelium growth (i.e., the vegetative part of fungi). The intraspecific variability of these mature forms span one to two orders of magnitude (Douhan, Vincenot, Gryta, & Selsosse, 2011), and it is thus difficult to decide on an effective definition

TABLE 2 Soil biota characteristics in the 12-ha plot

Clade	Total rel. abundance (%)	Total richness (per sample)	Total diversity (per sample)	OTUs distribution range (m)	Propagule size (μm)
Archaea	n.a.	2,502 (43 \pm 30)	30 (9 \pm 8)	170 \pm 163	0.5
Bacteria	n.a.	19,101 (1,391 \pm 232)	885 (453 \pm 102)	289 \pm 158	n.a.
<i>Acidobacteria</i>	19.81	1,601 (231 \pm 46)	132 (80 \pm 32)	352 \pm 136	0.5
<i>Actinobacteria</i>	20.58	1,696 (167 \pm 37)	86 (54 \pm 13)	313 \pm 151	5
<i>Bacteroidetes</i>	1.02	767 (31 \pm 12)	97 (22 \pm 8)	242 \pm 165	0.5
<i>Chloroflexi</i>	5.22	2,147 (136 \pm 48)	268 (73 \pm 29)	308 \pm 148	1
<i>Firmicutes</i>	2.66	873 (19 \pm 6)	4 (3 \pm 2)	186 \pm 164	1
<i>Alphaproteobacteria</i>	19.86	3,234 (303 \pm 57)	242 (129 \pm 25)	303 \pm 155	0.5
<i>Betaproteobacteria</i>	3.47	482 (45 \pm 13)	24 (14 \pm 7)	297 \pm 160	5
<i>Deltaproteobacteria</i>	4.36	1,357 (75 \pm 31)	73 (34 \pm 15)	279 \pm 151	0.5
<i>Gammaproteobacteria</i>	8.59	888 (72 \pm 11)	30 (21 \pm 5)	260 \pm 169	5
<i>Verrucomicrobia</i>	4.99	353 (51 \pm 12)	33 (22 \pm 5)	344 \pm 139	0.5
Eukaryota	n.a.	11,470 (475 \pm 137)	228 (51 \pm 23)	254 \pm 160	n.a.
<i>Ascomycota</i>	3.49	317 (25 \pm 9)	22 (8 \pm 3)	312 \pm 150	100
<i>Basidiomycota</i>	16.94	387 (31 \pm 8)	44 (8 \pm 4)	305 \pm 149	10
<i>Glomeromycota</i>	1.08	38 (7 \pm 2)	8 (5 \pm 1)	360 \pm 141	200
<i>Annelida</i>	6.62	49 (8 \pm 3)	6 (3 \pm 1)	336 \pm 155	20e ³
<i>Arthropoda</i>	17.65	1,777 (63 \pm 21)	91 (15 \pm 7)	235 \pm 168	10e ³
<i>Nematoda</i>	0.47	359 (9 \pm 5)	35 (6 \pm 2)	258 \pm 153	100
<i>Platyhelminthes</i>	0.54	117 (4 \pm 3)	15 (3 \pm 1)	226 \pm 169	20e ³
Protists	1.95	1,610 (53 \pm 40)	126 (27 \pm 18)	265 \pm 154	100

Notes. Only the studied taxonomic groups (most abundant ones) are shown. Plant OTUs retrieved with the eukaryote marker were not included in our analysis. They represented only 2% of the eukaryotic OTUs (10% of reads). Relative abundances were calculated for each marker separately and correspond to the per cent of reads of each clade. Richness values correspond to the number of OTUs (mean \pm standard deviation [SD] for values per sample is reported in parentheses). Diversity values correspond to the effective number of OTUs (exponential Shannon diversity index; mean \pm SD for values per sample). Note that the latter provides better estimates of species diversity than the number of OTUs per se for metabarcoding data (Haegeman et al., 2013; Supporting information Appendix S1: Material S2). OTU distribution range corresponds to the maximum distance between samples where an OTU is detected (mean \pm SD). n.a.: not applicable. The study plot is located in the Nouragues Reserve, French Guiana.

of a “body size.” In contrast, spore size is more conserved and was shown to be proportional to the reproduction rate and to the fructification size in certain fungal groups (Douhan et al., 2011; Meerts, 1999). This definition also cuts across the domains of life as it corresponds to average cell size for unicellular organisms (i.e., archaea, bacteria and protists) and average body size of adults for the soil mesofauna.

2.5 | Environmental parameters

We measure a total of 60 contextual parameters (Supporting information Appendix S1: Figure S2) corresponding essentially to soil chemistry (including important nutrients such as Mg, S, B, Ca, Mn), but also to lidar and plant DNA-derived data.

2.5.1 | Soil chemistry and airborne lidar

Total content in soil chemical elements was assessed every 20 m, corresponding to a quarter of the collected soil samples. Soil samples were ground, and 100 mg of powder was digested in Teflon beakers in a clean room following this protocol: 1 ml H₂O₂ + 1 ml HNO₃ during 24 hr at room temperature and during 24 hr at 80°C; evaporation to dryness at 80°C; 1 ml HNO₃ + 1 ml HF during 24 hr at 80°C; evaporation to dryness; 20 drops of HCl + 10 drops of HNO₃ (aqua regia) during 24 hr at 115°C; evaporation to dryness. The residue was then removed in a matrix HNO₃ 2% for analysis by ICPMS (Agilent 7500ce). Indium and rhenium were used as internal standards to correct for instrumental drift and eventual matrix effects during ICPMS analyses. The quality of the analysis was checked by analysing international certified reference waters (CNRC SLRS-5, NIST SRM 1643e). The accuracy was better than 5% relative to the certified values, and the analytical error (relative standard deviation) was better than 5% for concentrations ten times higher than the detection limits. Carbon and nitrogen concentrations were measured by a CHN elemental analyser (NA 2100 Protein, CE Instruments). In total, our data set included 55 elemental concentrations per sample. These were krigged using an exponential variogram model so as to obtain values for all the points of the initial sampling design. This analysis was conducted with the *sp* (<http://rsatial.r-forge.r-project.org/>) and *gstat* (<http://gstat.r-forge.r-project.org/>) R packages. In our study, we did not assess soil pH due to strong methodological constraints, although it is a common predictor of soil microbial diversity. However, previous studies in French Guiana revealed limited spatial variation compared to what could be expected in oxisols in general, and negligible variation in our study plot (pH = 4.2 ± 0.1, Supporting information Appendix S1: Figure S3). In a previous study at a site encompassing greater soil pH variation than at our site (Barberán et al., 2015), soil pH was found to have a limited effect on tropical soil microbial communities. We hence expect pH to have a small explanatory power of community composition at our site.

A 1-m² digital elevation model (DEM) was obtained at our site using airborne lidar scanning (Réjou-Méchain et al., 2015; Tymen et al., 2016). From the DEM, we derived slope and topographic

wetness indices (see Supporting information Appendix S1: Methods). We also used the lidar data to define a canopy closure index, which quantifies how much light reaches the ground level at each point.

2.5.2 | Plant sequencing data

Plants can influence soil communities through direct biotic interactions or local modification of the soil chemistry. To quantify this effect, we characterized the plant composition at the scale of the sampling point (i.e., root or litter presence) by amplifying a plastid DNA marker (P6 loop of the *trnL* intron) from soil DNA extracts, following the same analytical protocol as explained above (see also Table 1). We preferred the DNA metabarcoding approach over traditional plant census because it is a direct local measure, and because plant censuses usually include only trees ≥10 cm dbh, and thus exclude understory plants and lianas. Moreover, inferring the influence of plants on a local soil sample based on the stem locations is highly challenging. Since PCR failed for 30% of the sampling points (330 of the 1,132), we decided to infer plant composition by a multinomial resampling of reads from neighbouring points. This approach is reasonable because tree root influence is detectable up to 20 m from their corresponding stems in tropical forests (Barberán et al., 2015). We then inferred several plant-related characteristics to explain soil community assembly. These corresponded to plant diversity (Shannon index) and the identity of the three most dominant plant genera in each soil sample. In the molecular data set, plant dominant genera corresponded to 66 different plant taxa, 42 of which were unambiguously identified. The three dominant plant genera represented on average 70 ± 15% of the plant reads in each sample, hence providing a good description of the local dominant plants.

2.6 | Statistical analyses

We assessed the spatial autocorrelation of each focus taxonomic group using Mantel correlograms (999 permutations with FDR correction of *p*-values). This was done on both OTU diversity and community compositional dissimilarities. OTU diversity was calculated as the exponential of the Shannon entropy, an effective number of species that is less sensitive to rare OTUs (Jost, 2006) and hence more robust to describe microbial diversity (Haegeman et al., 2013). Community dissimilarities were obtained with the Bray–Curtis index. Correlation of spatial patterns between focus taxonomic groups and with plants was assessed with the Pearson product-moment correlation or Mantel tests for OTU diversity and community dissimilarity, respectively.

To disentangle the relative importance of niche-based and neutral processes in the assembly of soil communities, we did not test neutral models (Hubbell, 2001) against our data because they assume an unbiased resampling of individuals and species. This assumption is not necessarily met with metabarcoding data (e.g., molecular biases, sampling of DNA fragments instead of individuals), which can bias estimates of Hubbell’s neutral theory parameters (Sommeria-Klein, Zinger, Taberlet, Coissac, & Chave, 2016). Instead,

we used two complementary methods (as reviewed in Vellend et al., 2014). First, we used a null model approach based on community pairwise dissimilarities, as proposed by Chase and Myers (2011) and extended by Stegen et al. (2013) to incorporate species abundances. This approach consists in determining whether community dissimilarity patterns deviate from a null scenario. Consequently, it does not require a priori knowledge of the local environmental conditions. For each sample, the null scenario was generated by random resampling of OTUs and reads in the total community matrix, while preserving the sample richness and number of reads in the sample. OTU total occurrences and abundances were used as probabilities of selecting an OTU and its associated number of reads, respectively. Dissimilarities across null communities were then computed with the Bray–Curtis index. This process was repeated 1,000 times, and the resulting dissimilarities were averaged out to obtain null expectations of community dissimilarities for each pair of samples. We then defined the null deviation as the difference between observed and averaged null dissimilarities (Tucker, Shoemaker, Davies, Nemergut, & Melbourne, 2016), which returns values between -1 and 1 and deviating from 0 when the observed dissimilarities differ from the null model. Values departing from 0 usually indicate environmental selection in well-sampled communities (Supporting information Appendix S1: Figure S1; Tucker et al., 2016), but can also correspond to high dispersal limitation (values > 0) or homogenizing dispersal (values < 0 ; Stegen et al., 2013). To better differentiate these processes, null model inferences can be complemented with analyses of the community phylogenetic structure (Stegen et al., 2013). We did not conduct such an analysis here because obtaining phylogenetic trees for the 19 taxonomic groups studied proved an insurmountable challenge, due to the unavailability of backbone phylogenies for most groups, and the uncertainty associated with phylogenetic placement methods for short sequences.

The null model approach cannot identify potential biotic or abiotic drivers of community assembly. To complement it, we applied in parallel another approach based on variation partitioning and redundancy analyses (RDA; Legendre, 2008). This analysis consists in decomposing the variance of community composition explained by environmental or spatial variables alone as well their combined effects. Initially, the contribution of pure environmental vs. spatial component was considered to be indicative of the relative importance of niche-based processes vs. dispersal limitation (Cottenie, 2005; Legendre, 2008). This interpretation has been questioned because it applies only when the environmental context is well characterized, spatially not structured (Smith & Lundholm, 2010; Vellend et al., 2014), and when neutral processes are not correlated with the environment. We will consider these limits in our interpretations.

Our preliminary analyses indicated a linear relationship between dependent variables (soil diversity or community distribution) and explanatory variables, a prerequisite for using RDA. We Hellinger-transformed the OTU tables of each taxonomic group to downweigh rare OTUs, as well as to preserve the Euclidean distance among samples (Legendre & Gallagher, 2001). We then constructed three parsimonious models corresponding to three

nonexclusive hypotheses: (a) community composition is driven by the measured abiotic parameters, (b) community composition is driven by direct or indirect interactions with plants, or (c) community composition is driven by spatial processes, linked or not to environmental factors. Model 1 included soil chemistry and lidar-derived data as predictive factors. Model 2 included plant explanatory variables (Figure 1; Supporting information Appendix S1: Table S2). For these two models, explanatory variables were preselected to reduce multicollinearity ($|\text{Pearson's } r| \geq 0.7$; Supporting information Appendix S1: Figure S2) and normalized with a Box–Cox transformation to meet normality assumptions. Finally, model 3 included spatial eigenvectors as explanatory variables, derived from a Principal Coordinate analysis of Neighbour Matrices (PCNM) (Dray, Legendre, & Peres-Neto, 2006). PCNMs allow modelling the spatial structure of community composition variation. To avoid inflation of R^2 statistics caused by the large number of PCNM eigenvectors (Gilbert & Bennett, 2010), we preselected them prior the forward selection procedure. We retained only PCNMs explaining a significant fraction of the variation in the biological response (at a significance level $p \leq 0.02$), as assessed through partial canonical redundancy analysis (pRDA). This preselection was independently performed for each study clade. Geographic coordinates were also included in the model to account for possible linear trends along the study area (Dray et al., 2006). Finally, we applied a classical forward selection procedure on models 1, 2 and 3 independently for each taxonomic group to limit inflation of the amount of explained variance and type I error (Blanchet, Legendre, & Borcard, 2008). The resulting abiotic, plant and spatial matrices were used to conduct a variation partitioning analysis. Significance of the total RDA models and of variation partitioning pure components was determined with 1,000 Monte Carlo permutations. We here report adjusted R^2 statistics (R^2_{adj}), which are less inflated when the numbers of explanatory variables are high (Peres-Neto, Legendre, Dray, & Borcard, 2006). To better interpret the combined effects of environmental/spatial parameters, we identified the plant and abiotic parameters that correlated with PCNM eigenvector using pRDA.

To further gain insights into the spatial scales at which the community composition may covary with environmental conditions, we decomposed the community variation with a second “full” RDA model that included four categories of explanatory variables. The first category comprised environmental features, that is abiotic and plant variables. The three remaining categories corresponded to PCNM eigenvectors with wavelengths ≥ 100 m, between 100 and 25 m, and < 25 m (Supporting information Appendix S1: Figure S6). The R^2_{adj} statistics obtained for pure or combined effects of each category of variable were used to calculate their relative contribution to the total explained variance. Variables used in these models corresponded to those preselected with the initial forward selection procedure, and tests for significance were conducted as described above. All variation partitioning analyses were repeated by considering OTU diversity as a response variable to identify predictors of soil biodiversity.

Finally, we tested the effect of body size on community assembly by comparing taxon's propagule sizes with the various statistics obtained above using the Pearson product-moment correlation tests. All analyses were conducted with the *vegan* R package (<http://vegan.r-forge.r-project.org/>).

3 | RESULTS

We found a total of 2,502 archaeal, 19,101 bacterial and 11,470 eukaryotic OTUs within the 12-ha plot (Table 2). Many of these OTUs were rare (relative abundance <0.1%): they accounted for $36 \pm 10\%$ (mean \pm standard deviation values hereafter) of archaeal OTUs, $86 \pm 2\%$ of bacterial OTUs and $80 \pm 6\%$ of eukaryotic OTUs. Bacterial OTUs identified at the phylum level (88% of OTUs, 94% of reads) corresponded primarily to Actinobacteria,

Alphaproteobacteria and Acidobacteria. Eukaryotic OTUs identified at least at the phylum level (51% of OTUs, 70% of reads) belonged to fungi (mainly Agaricomycetes, Glomeromycetes and Orbiliomycetes; Supporting information Appendix S1: Table S3), arthropods (mainly termites, mites and springtails) and annelids (Oligochaeta). For Archaea, the OTUs being assigned at least at the phylum level (3% of OTUs, 64% of reads) mainly corresponded to Nitrososphaeria and Methanomicrobia.

Figure 2 illustrates the spatial variation in local diversity for the focal taxonomic groups. Both positive and negative spatial autocorrelation were observed at various distances, depending on the taxon (Fig. 2; Supporting information Appendix S1: Figure S4). Dissimilarities of soil community increased with body size (Figure 3a, Pearson's $r = 0.52$, $p = 0.02$). Compositional turnover also displayed positive spatial autocorrelation (Supporting information Appendix S1: Figure S5) for distance classes that decreased with taxon body

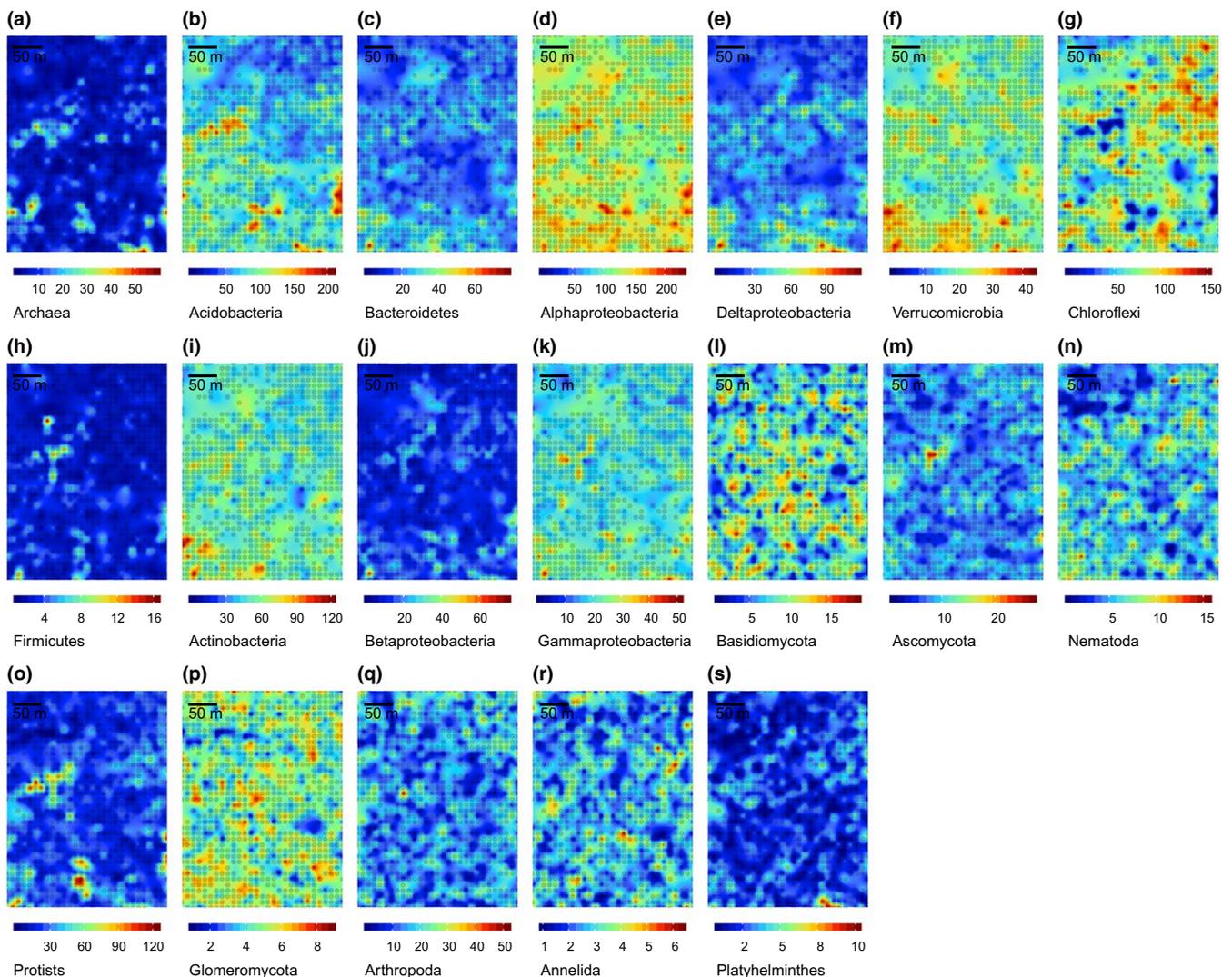


FIGURE 2 Spatial distribution of OTU diversity per focal clade. The colour scale is expressed as effective numbers of OTUs (exponential Shannon diversity index) for archaea (a), bacterial clades (b–k) and eukaryotic clades (l–s), ranked by increasing body size (Table 2). Maps were obtained by ordinary kriging. Grey dots represent the samples available for each clade. The study plot is located in the Nouragues Reserve, French Guiana

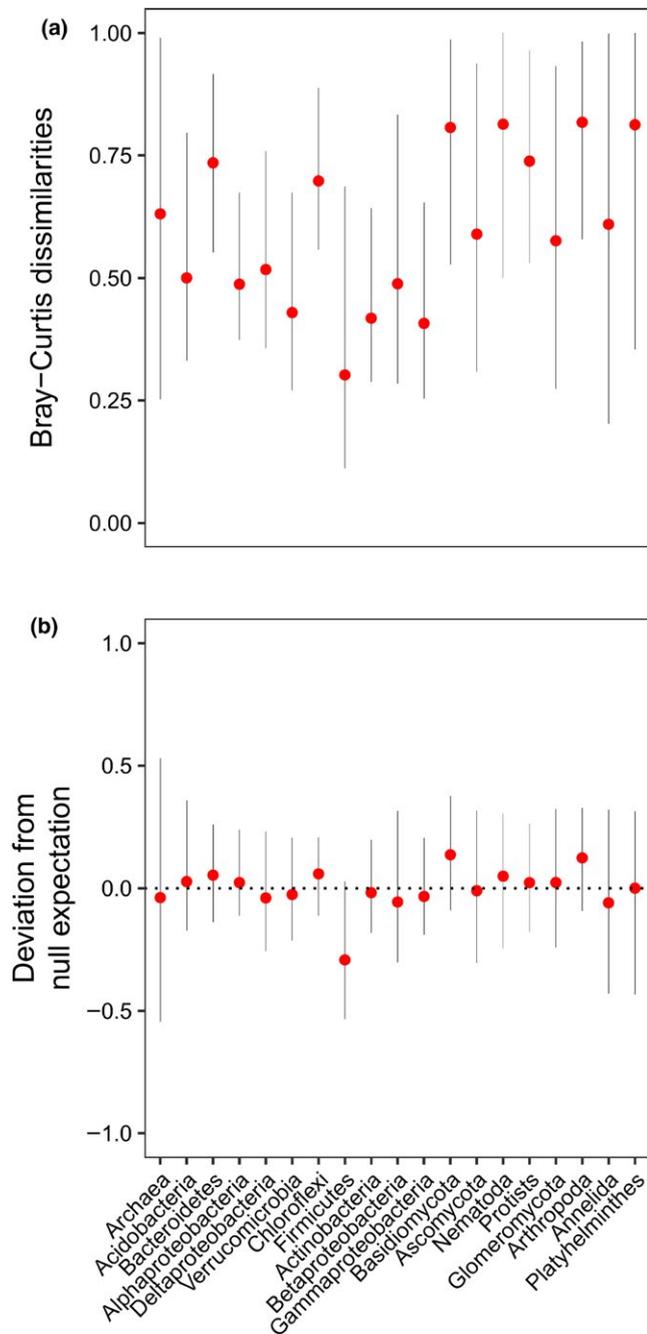


FIGURE 3 Community dissimilarities in the study plot (a) and their deviation from null model expectations (b) for each focus taxonomic group. Values of both observed dissimilarity and null deviation are summarized by their average (red dots) and 95% confidence intervals. In (b), the horizontal black dashed line indicates neutral assembly. Focal taxonomic groups on the x-axis are ranked by increasing body size (Table 2). The study plot is located in the Nouragues Reserve, French Guiana

size (Pearson's $r = -0.59$, $p = 0.007$). Negative spatial autocorrelation was observed at larger distances (>100 m) mostly in microorganisms (Supporting information Appendix S1: Figure S5). The average OTU distribution range of each taxonomic group was >200 m in general, and showed no trend with body size (Table 2). Pairwise correlation

of OTU diversity or compositional dissimilarity between focal taxonomic groups was moderate to low (|Pearson's r | < 0.7), except for several bacterial clades. They showed low or no correlation with plant diversity or community turnover (Supporting information Appendix S1: Tables S4, S5). Finally, the observed community dissimilarity values did not deviate from null expectations for all taxonomic groups, except Firmicutes (Figure 3b). Null deviation values did not show any particular trend with respect to body size (Pearson's $r = 0.18$, $p = 0.45$).

Abiotic and biotic conditions were noticeably heterogeneous, with abiotic gradients detectable at broad and intermediate spatial scales (Figure 1; Supporting information Appendix S1: Figure S6). However, the full RDA models, that is abiotic, plant and spatial descriptors, explained only 3%–13% of the variation in OTU composition across the 19 focal taxonomic groups (Figure 4a, Supporting information Appendix S1: Table S6). Only $3 \pm 1\%$ of the variation was due to total abiotic effects, mainly through abiotic parameters that were spatially structured at broad-to-intermediate spatial scales (Figure 1; Supporting information Appendix S1: Figure S5, Table S6), such as soil aluminium and topographic variables (i.e., elevation, convexity and wetness). These abiotic parameters correlated negatively with the OTU diversity of most unicellular taxonomic groups, and worms for topographic variables (Supporting information Appendix S1: Figures S7, S8). When exploring plants' influence on species composition across groups, we found that plant variables explained $2 \pm 1\%$ of the variation. The identity of the dominant plant genera in soil samples as detected with the plant DNA marker was the best plant-related predicting factor. Finally, spatial total effects explained $7 \pm 2\%$ of the variation, half of which corresponded to pure spatial effects.

When comparing these results across the focus taxonomic groups, we found that the unexplained component of RDA models correlated positively with dispersal unit size (Figure 4b) and was strongly related to the averaged dissimilarities expected under the null model (Pearson's $r = 0.82$, $p < 0.001$). We also found that the explained components corresponding to both pure environmental (i.e., abiotic and plant variables) and broad-/intermediate-scale effects (combined with environmental effects or not) all negatively correlated with dispersal unit size (Supporting information Appendix S1: Figure S9, Table S7). The fraction of community variation explained by fine-scale spatial variable was unrelated to environmental ones and tended to increase with organism propagule size. Using OTU diversity instead of OTU composition as the dependent variable resulted in greater amount of variation explained by the tested factors (4%–46%), and their relative importance for the two "full" RDA models, as well as their variation across body size, was similar to what is reported above (Supporting information Appendix S1: Tables S8, S9, Figures S10, S11).

4 | DISCUSSION

4.1 | Environmental drivers of soil communities

In this study, we assessed the environmental predictors that significantly explained soil community composition. Among them, we

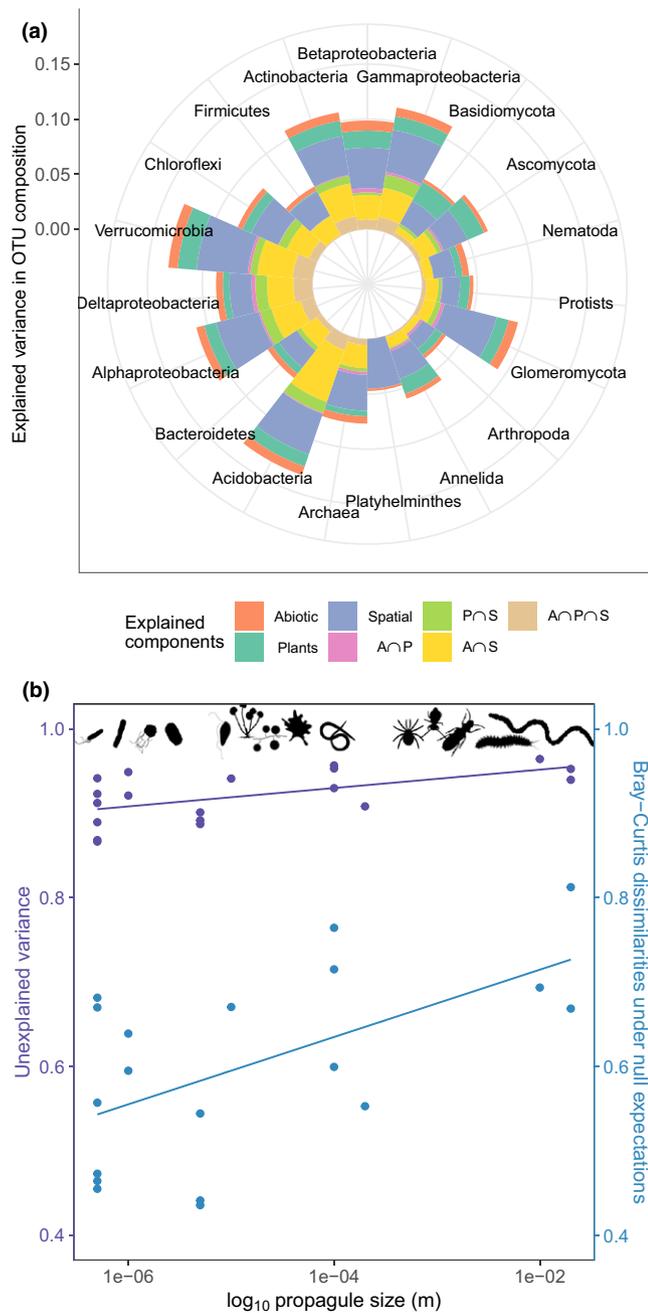


FIGURE 4 Variation partitioning of OTU composition in each focus taxonomic group and across propagule size. (a) Variations in OTU composition are partitioned into pure (i.e., abiotic, A; plant, P or spatial, S) and shared components ($A \cap P$, $P \cap S$, $A \cap S$ and $A \cap P \cap S$). See Supporting information Appendix S1: Table S6 for corresponding R^2_{adj} values and their significance (for pure components and full models only, shared components not testable). (b) Relationship between propagule size and the amount of unexplained variance (i.e., residual R^2_{adj} , left y-axis) by the RDA model (Pearson's $r = 0.58$, $p = 0.009$) and Bray-Curtis dissimilarities obtained with the null model (right y-axis; Pearson's $r = 0.57$, $p = 0.01$). The study plot is located in the Nouragues Reserve, French Guiana

identified the locally dominant plant genera, an effect previously termed the "tree identity effect" (Bardgett & van der Putten, 2014; Tedersoo et al., 2016). Neotropical tree species can be indeed

involved in specific mutualistic/trophic interactions, as they greatly differ in the quantity, quality and chemistry of their litter (Hättenschwiler, Aeschlimann, Coûteaux, Roy, & Bonal, 2008), and most likely of their root exudates and physical structure. We found that soil total aluminium content was also a predictor of microbial community composition and correlated negatively with microbial diversity. Aluminium concentration strongly correlated with that of many other metals (e.g., Ti, Cu, Zn), all well known for their toxic properties (Lemire, Harrison, & Turner, 2013; Piña & Cervantes, 1996) and their capacity to immobilize soil nutrients and organic matter (Fujii, 2014). Soils with high aluminium content also tend to be more acidic (Abreu, Muraoka, & Lavorante, 2003; Driscoll & Schecher, 1990), conditions under which microbial diversity is usually lower (Barberán et al., 2015; Ramirez et al., 2014). Finally, topography, slope and the topographic wetness index, all closely related to soil moisture, also predicted the community composition for many taxonomic groups. They were the most important predictor of OTU diversity and composition of flat/earthworms, which are well known to have particular soil moisture optima (Fragoso & Lavelle, 1992; Jiménez, Decaëns, Lavelle, & Rossi, 2014). These variables also correlated with the spatial diversity and composition patterns of Deltaproteobacteria, an anaerobic clade favoured in wet conditions (Martiny et al., 2015). We hence interpret the decline of diversity in metal-rich and/or drier sites as a response to stressful conditions.

However, environmental parameters explained only a minor fraction of the variation in soil community composition. This result contrasts with previous studies, where soil organisms are often reported to covary relatively strongly with, for example, soil pH, moisture, nutrient or organic matter availability/quality (Dumbrell et al., 2010; Ramirez et al., 2014; Tedersoo et al., 2016; Wu et al., 2011) or plant diversity (Basset et al., 2012; Peay et al., 2013). Although we have characterized the environmental conditions to the best of our ability, we inevitably did not quantify directly all environmental predictors or biotic interactions. For example, lidar-derived measures, including soil wetness index, may lack precision for our sampling grain. Also, we did not measure soil pH, but expect it to be homogeneous spatially and to have a negligible effect (Section 2). More generally, many edaphic parameters are often spatially structured and their effects should be indirectly detected with PCNM variables (Baldeck et al., 2013; De Cáceres et al., 2012; Smith & Lundholm, 2010). Here, these spatial variables also explained a minor part of soil community variation, hence suggesting that environmental gradients have only a weak influence on our study site. Yet, other parameters inaccessible with our sampling grain such as soil microstructure or the geometry of roots may be also important in explaining the small-scale distribution of soil organisms (Baveye, Berthelin, & Munch, 2016; Berg, 2012).

The unexplained component of community variation, which was prominent in our study, has often been interpreted as resulting from local stochastic demographic processes (De Cáceres et al., 2012; Legendre, 2008). We suggest that in our study, this interpretation is valid. Indeed, we also found that the observed community

dissimilarity patterns did not depart markedly from null expectations in any of the focal taxonomic groups. Also, the amount of unexplained variation across the focus taxonomic groups correlated remarkably well with corresponding averaged dissimilarities expected under the null model, a feature already observed (De Cáceres et al., 2012). Consequently, although the environmental context was not fully characterized in our study, it is unlikely that we underestimated the importance of environmental selection. Rather, all our analyses point towards the prominence of stochastic processes in the assembly of soil communities at the scale of our study plot, hence differing from most soil diversity studies (reviewed in Berg, 2012; Hanson et al., 2012; Barberán et al., 2014). On the other hand, environmental predictors (in particular soil aluminium content and soil moisture) better explained soil diversity patterns. This suggests that low moisture and high aluminium sites do not only recruit particular taxa, as suggested by their weak, yet significant effects on community composition. These conditions also impose energetic constraints on the system carrying capacity that would limit local community size, causing a random species loss. Further analyses of the community phylogenetic structure could help confirming this hypothesis (Hurlbert & Stegen, 2014).

4.2 | Underlying causes of community stochasticity

A major problem in attempts to explain patterns of biodiversity is the difficulty of observing ecological processes at the precise scale at which they manifest themselves (Levin, 1992). Soils are complex systems where ecological processes operate at a hierarchy of scales, and any attempt to interpret them should carefully examine the spatial scale of study and the grain of the sampling unit (Baveye et al., 2016; Berg, 2012; Ettema & Wardle, 2002; Hinsinger et al., 2009). We suggest that at least part of the discrepancies between our findings and previous studies are related to a problem of size and sampling grain. Previous analyses were based on sampling units of typically, for example, 20 × 30 m distributed across landscapes or regions with potentially steep environmental gradients. For example, previously observed relationships between the species diversity/composition of soil organisms and local plant community composition (Basset et al., 2012; Peay et al., 2013) are observable at spatial scales at which local plant community composition correlates with environmental conditions. However, the strength of this relationship decreases with decreasing sizes of sampling units and vanishes at the scale of the sampling point (Barberán et al., 2015). In our study, we did not observe such a relationship because we sampled local soil cores of ca. 15 g from which we inferred both local soil richness and plant community composition. Thus, we were able to relate the cooccurrence of microbial cells and plant cells at the centimetric scale (Dini-Andreote & van Elsas, 2013), and concluded there was a weak, yet significant “tree identity effect” as reported above.

Although our study plot was large and displayed gradients of abiotic conditions and trees species distribution, these gradients are most likely not steep enough to drive soil community assembly as

in other forests. Accordingly, previous studies have reported weak effects of environmental selection at smaller spatial scales (0.04–0.2 ha) for soil microeukaryotes (Bahram et al., 2016), earthworms (Jiménez et al., 2014) and bacteria (Barberán et al., 2015). This implies that soil communities appear to be primarily driven by ecological drift and/or dispersal processes. In our study plot, dispersal is unlikely to be limited. At this scale, dispersal of microbes occurs through the movement of soil particles caused by, for example, raindrops and water flows, aboveground and belowground animal activity, and filamentous fungi or plant root growth (Joung, Ge, & Buie, 2017; Rillig et al., 2015). Similarly, passive dispersal of soil mesofauna can spread through the entire plot and beyond, through transport of cocoons by water flow, or through phoresy on the macrofauna (Decaëns et al., 2016). Accordingly, we found that OTUs had relatively large distribution ranges (>200 m on average) and that communities were not more dissimilar than null ones, contrary to the dispersal limitation scenario (Stegen et al., 2013). In contrast, high levels of dispersal (i.e., homogenizing dispersal) should offset both environmental effects and ecological drift by increasing the chance for a soil-inhabiting species to occupy any location through a source-sink dynamics (Stegen et al., 2013; Vellend, 2010). This should result in communities that are more similar than expected by chance, which was not the case here. Priority effects, that is the fact that early colonizing species determine the local assemblage, may also cause stochastic patterns. However, this process is unlikely in our study because it is linked to environmental selection and would lead to communities more dissimilar than under null expectations. Hence, the stochasticity in the community assembly in our study plot is most likely caused by ecological drift without dispersal limitation.

4.3 | Organism body size determines assembly processes

Despite this apparent stochasticity, we still found some coherence in the spatial patterns across taxonomic groups. An important result from our analysis was the significant and linear increase in both the amount of unexplained variation in diversity or composition, and null community dissimilarities with increasing propagule size. We did not retrieve this trend for null deviation values because the body size dependence of both observed and null community dissimilarities cancels out using this metric. One criticism of null models is that they preserve species abundances and local diversity distributions of the data, which can result from assembly mechanisms. Besides, null deviation values are difficult to compare across ecosystems or taxonomic groups due to their sensitivity to regional diversity and sampling effort (Bennett & Gilbert, 2015; Tucker et al., 2016). Finally, we also observed a negative correlation of body size with the spatial autocorrelation extent in community turnover, as well as with observed community dissimilarity. Our result implies that community stochasticity increases with body size.

The smallest body size category includes bacteria, whose distribution showed greater spatial autocorrelation at broad-to-intermediate scales. The variation partitioning approach including

explicitly spatial scales as explanatory variables suggests that this distribution pattern can be explained by two nonexclusive processes: (a) a more pronounced environmental selection by the broad- to intermediate-scale environmental gradients at our site, and (b) a higher level of homogenizing dispersal at intermediate scales, as suggested by the negative relationship between body size and the pure intermediate-scale spatial component. Microeukaryote (i.e., fungi and protists) distributions were less well explained by environmental gradients than bacteria. Similar differences have been observed between bacterial and fungal community assembly (Powell et al., 2015), which were explained by a greater dispersal limitation and stronger priority effects in fungi. That this interpretation holds true for our study site remains to be determined. Also, in fungi, the mycelium can spread across heterogeneous environments and the apparent distribution of fungi may be less related to environmental conditions at the scale of observation. Protists, archaea and certain bacterial groups (i.e., Firmicutes and Bacteroidetes) displayed similar features. These groups harbour low abundances in soils relative to other clades, they do not necessarily exhibit dormant stages, and some of them can be associated with soil invertebrates (Barberán et al., 2015; Engel & Moran, 2013; Fierer et al., 2009; Geisen et al., 2018; Ramirez et al., 2014). This suggests that these microbial groups have smaller population sizes in soils and/or lower dispersal rates. They could hence experience more extinction and be more prone to ecological drift compared to most bacterial groups.

The distribution of OTUs of large-bodied organisms, that is arthropods, annelids and flatworms, was even less related to environmental parameters and did not exhibit broad-to-intermediate spatial structures. Soil mesofauna communities were hence the most stochastic of the focus taxonomic groups at the scale of our 12-ha plot and grain studied. We also found that the contribution of the pure fine-scale spatial effects to the community variation tended to increase with increasing propagule size. This result is in agreement with the clumped spatial distribution of soil mesofauna, which usually displays spatial aggregation below 10 m (Bahram et al., 2016; Berg, 2012; Jiménez et al., 2014). This roughly corresponds to the horizontal distance earthworms can actively disperse per year, and suggests that although dispersal is not limited in the study plot, its rates may be lower for mesofauna than for microbes. In addition, studies using a finer sampling grain found that these aggregates covary poorly with environmental factors (Bahram et al., 2016; Jiménez et al., 2014). Indeed, at such small scales relative to mesofauna body size, population densities are low and hence more driven by drift (Chase, 2014). On the opposite end of the range, environmental selection of large-bodied organisms emerges at scales larger than 12 ha when soil properties are highly contrasted, as reported for tropical trees (Baldeck et al., 2013) or mesofauna (Basset et al., 2015; Decaëns et al., 2016).

The body size dependence of drift, dispersal and environmental selection is in line with Levin's argument on the problem of scales (Levin, 1992), and confirms our initial hypothesis. For the studied

area and sampling grain, we considered patterns of diversity at a range of spatial scales owing to differences in body size of the studied taxa relative to the studied area. We find that microbial groups, for which 12 ha represents a huge area, are better predicted by the environment than mesofauna ones, for which 12 ha represents a smaller spatial extent. This mirrors the increases in environmental selection with the observational spatial scale reported for single taxonomic groups (Barton et al., 2013; Chase, 2014; De Cáceres et al., 2012; Dini-Andreote et al., 2015). This result points to an interesting juncture between two fields of ecological theory that have heretofore been loosely connected, namely the role of body size in explaining the scaling rules of metabolism (Brown et al., 2004) and that of environment and dispersal in explaining the assembly of ecological communities (Ricklefs, 2004). In that light, the fact that large-bodied organisms are more stochastic than small-bodied organism assembly is consistent with the main predictions of macroecology (Ricklefs, 2004) and soil science (Berg, 2012; Ettema & Wardle, 2002). We are aware that the taxonomic groups studied here harbour a large variation in propagule size, but this variation does not exceed the between-group body size variation (Briones, 2014; Portillo et al., 2013). Consequently, our finding of a log-linear relationship between the contribution of environmental and/or pure spatial effects in community assembly and propagule size is robust, and generalizes previous empirical observations in freshwater ecosystems (Bie et al., 2012 and references within). It provides an empirical evidence of spatial scaling rules across the soil food web. Although other important biological features also likely explain the differences in spatial distribution among soil taxonomic groups (e.g., clonal vs. sexual reproduction, dormancy, mutualistic/pathogenic interactions, active vs. passive transport), body size constitutes an operational parameter that cuts across the tree of life. Therefore, and perhaps unsurprisingly, body size has been considered as a key trait in the modelling of species interaction networks, as it constrains several essential species interaction variables such as handling times, densities or attack rates (Brose, 2010; Petchey et al., 2008).

The recent explosion of DNA-based studies has considerably increased our knowledge on the taxonomic, genetic and functional diversity of soil organisms (Bardgett & van der Putten, 2014), but has yet to provide a comprehensive understanding of the mechanisms shaping soil biodiversity (Baveye et al., 2016). Here, we have evidence that this tropical soil community at the scale of a 12-ha plot is prominently neutral, mirroring tree communities in similar environments (Hubbell, 2001). However, it is likely that soil community assembly in other forests with more heterogeneous soil abiotic condition may be more prone to environmental selection than here. Still, one direct and important consequence of our result confirms that soil diversity, distribution and extinction rates cannot be inferred from the patterns of plant or any other surrogate group at the field scale (Schuldt et al., 2015; Veresoglou, Halley, & Rillig, 2015). Finally, our results also suggest that accounting for the scaling relationship between body size and community assembly processes will help unravel spatial patterns in such complex, multitrophic communities and contribute to the development of

currently emerging predictive spatial models of soil food webs (Galiana et al., 2018).

ACKNOWLEDGEMENTS

We are indebted to the staff of the Nouragues Research Field Station (CNRS, French Guiana) and the Nouragues Natural Reserve, for commodities and technical help during the field session. We are grateful to Benjamin L. Turner for providing us soil pH data and advice on the manuscript. We also thank Johan Pansu, Guilhem Sommeria-Klein, Mélanie Roy and Renato Lima for helpful discussions, as well as the Genotoul bioinformatics platform of Toulouse Midi-Pyrenees (Bioinfo Genotoul) and Pierre Solbes (EDB-Calc Cluster) for providing computing and storage resources. We are also grateful to anonymous referees for their constructive comments on an earlier version of the manuscript. The work was funded by the METABAR project (ANR-11-BSV7-0020) and has benefitted from “Investissement d’Avenir” grants managed by Agence Nationale de la Recherche (CEBA: ANR-10-LABX-25-01; TULIP: ANR-10-LABX-0041; OSUG@2020: ANR-10-LABX-56; ANAEE-France: ANR-11-INBS-0001). This work is dedicated to the memory of our colleague Serge Aubert.

AUTHOR CONTRIBUTIONS

L.Z., P.T. and J.C. conceived the study. L.Z., P.T., H.S., A.B., M.D.B., P.G., L.G., C.G.C., A.I., M.R.M., G.R., E.C. and J.C. contributed to the fieldwork and/or to DNA extractions. P.T., A.B., A.I. and D.R. conducted the laboratory work to produce the metabarcoding data. J.V. and C.Z. performed the chemical analyses. B.T. provided the LiDAR data. L.Z. performed the bioinformatics and statistical analyses with the help of V.S., F.B., E.C., W.T. and J.C. L.Z. and J.C. wrote the manuscript with input from all co-authors.

COMPETING INTERESTS

L.G. and P.T. are coinventors 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.

DATA ACCESSIBILITY

Raw and curated sequencing data as well as associated metadata and bioinformatic codes are available on the Dryad Digital Repository (<https://doi.org/10.5061/dryad.dc28m>).

ORCID

Lucie Zinger  <https://orcid.org/0000-0002-3400-5825>

Pierre Taberlet  <https://orcid.org/0000-0002-3554-5954>

Heidy Schimann  <https://orcid.org/0000-0002-9139-920X>

Frédéric Boyer  <https://orcid.org/0000-0003-0021-9590>

Charline Giguet-Covex  <https://orcid.org/0000-0002-5141-4815>

Maxime Réjou-Méchain  <https://orcid.org/0000-0003-2824-267X>

Blaise Tymen  <https://orcid.org/0000-0001-7341-4085>

Wilfried Thuiller  <https://orcid.org/0000-0002-5388-5274>

Eric Coissac  <https://orcid.org/0000-0001-7507-6729>

Jérôme Chave  <https://orcid.org/0000-0002-7766-1347>

REFERENCES

- Abreu, C. H. Jr, Muraoka, T., & Lavorante, A. F. (2003). Relationship between acidity and chemical properties of Brazilian soils. *Scientia Agricola*, 60(2), 337–343. <https://doi.org/10.1590/S0103-90162003000200019>
- Anderson-Teixeira, K. J., Davies, S. J., Bennett, A. C., Gonzalez-Akre, E. B., Muller-Landau, H. C., Joseph Wright, S., ... Basset, Y. (2015). CTFs-ForestGEO: A worldwide network monitoring forests in an era of global change. *Global Change Biology*, 21(2), 528–549. <https://doi.org/10.1111/gcb.12712>
- Bahram, M., Kohout, P., Anslan, S., Harend, H., Abarenkov, K., & Tedersoo, L. (2016). Stochastic distribution of small soil eukaryotes resulting from high dispersal and drift in a local environment. *The ISME Journal*, 10(4), 885. <https://doi.org/10.1038/ismej.2015.164>
- Baldeck, C. A., Harms, K. E., Yavitt, J. B., John, R., Turner, B. L., Valencia, R., ... Thomas, D. W. (2013). Soil resources and topography shape local tree community structure in tropical forests. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 280(1753), 20122532.
- Baldrian, P., Kolařík, M., Štursová, M., Kopecký, J., Valášková, V., Větrovský, T., ... Voříšková, J. (2012). Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. *The ISME Journal*, 6(2), 248. <https://doi.org/10.1038/ismej.2011.95>
- Baraloto, C., Hardy, O. J., Paine, C. E., Dexter, K. G., Cruaud, C., Dunning, L. T., ... Chave, J. (2012). Using functional traits and phylogenetic trees to examine the assembly of tropical tree communities. *Journal of Ecology*, 100(3), 690–701. <https://doi.org/10.1111/j.1365-2745.2012.01966.x>
- Barberán, A., Casamayor, E. O., & Fierer, N. (2014). The microbial contribution to macroecology. *Frontiers in Microbiology*, 5, 203.
- Barberán, A., McGuire, K. L., Wolf, J. A., Jones, F. A., Wright, S. J., Turner, B. L., ... Fierer, N. (2015). Relating belowground microbial composition to the taxonomic, phylogenetic, and functional trait distributions of trees in a tropical forest. *Ecology Letters*, 18(12), 1397–1405. <https://doi.org/10.1111/ele.12536>
- Bardgett, R. D., & van der Putten, W. H. (2014). Belowground biodiversity and ecosystem functioning. *Nature*, 515(7528), 505. <https://doi.org/10.1038/nature13855>
- Barton, P. S., Cunningham, S. A., Manning, A. D., Gibb, H., Lindenmayer, D. B., & Didham, R. K. (2013). The spatial scaling of beta diversity. *Global Ecology and Biogeography*, 22(6), 639–647. <https://doi.org/10.1111/geb.12031>
- Basset, Y., Cizek, L., Cuénoud, P., Didham, R. K., Guilhaumon, F., Missa, O., ... Tishechkin, A. K. (2012). Arthropod diversity in a tropical forest. *Science*, 338(6113), 1481–1484. <https://doi.org/10.1126/science.1226727>
- Basset, Y., Cizek, L., Cuénoud, P., Didham, R. K., Novotny, V., Ødegaard, F., ... Roubik, D. W. (2015). Arthropod distribution in a tropical rainforest: Tackling a four dimensional puzzle. *PLoS ONE*, 10(12), e0144110. <https://doi.org/10.1371/journal.pone.0144110>
- Baveye, P. C., Berthelin, J., & Munch, J. C. (2016). Too much or not enough: Reflection on two contrasting perspectives on soil biodiversity. *Soil*

- Biology and Biochemistry*, 103, 320–326. <https://doi.org/10.1016/j.soilbio.2016.09.008>
- Bécard, G., & Pfeffer, P. E. (1993). Status of nuclear division in arbuscular mycorrhizal fungi during in vitro development. *Protoplasma*, 174(1–2), 62–68. <https://doi.org/10.1007/BF01404043>
- Bennett, J. R., & Gilbert, B. (2015). Contrasting beta diversity among regions: How do classical and multivariate approaches compare? *Global Ecology and Biogeography*, 25(3), 368–377.
- Berg, M. P. (2012). Patterns of biodiversity at fine and small spatial scales. In D. H. Wall, R. D. Bardgett, V. Behan-Pelletier, et al. (Eds.), *Soil ecology and ecosystem services* (pp. 136–152). Oxford, UK: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199575923.001.0001>
- Bie, T., Meester, L., Brendonck, L., Martens, K., Goddeeris, B., Ercken, D., ... Wichelen, J. (2012). Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. *Ecology Letters*, 15(7), 740–747. <https://doi.org/10.1111/j.1461-0248.2012.01794.x>
- Blanchet, F. G., Legendre, P., & Borcard, D. (2008). Forward selection of explanatory variables. *Ecology*, 89(9), 2623–2632. <https://doi.org/10.1890/07-0986.1>
- 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(1), 176–182. <https://doi.org/10.1111/1755-0998.12428>
- Briones, M. J. I. (2014). Soil fauna and soil functions: A jigsaw puzzle. *Frontiers in Environmental Science*, 2, 7.
- Brose, U. (2010). Body-mass constraints on foraging behaviour determine population and food-web dynamics. *Functional Ecology*, 24(1), 28–34. <https://doi.org/10.1111/j.1365-2435.2009.01618.x>
- Brose, U., Dunne, J. A., Montoya, J. M., Petchey, O. L., Schneider, F. D., & Jacob, U. (2012). Climate change in size-structured ecosystems. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 367(1605), 2903–2912. <https://doi.org/10.1098/rstb.2012.0232>
- Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., & West, G. B. (2004). Toward a metabolic theory of ecology. *Ecology*, 85(7), 1771–1789. <https://doi.org/10.1890/03-9000>
- Chase, J. M. (2014). Spatial scale resolves the niche versus neutral theory debate. *Journal of Vegetation Science*, 25(2), 319–322. <https://doi.org/10.1111/jvs.12159>
- Chase, J. M., & Myers, J. A. (2011). Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1576), 2351–2363. <https://doi.org/10.1098/rstb.2011.0063>
- Chave, J., Olivier, J., Bongers, F., Châtelet, P., Forget, P. M., Van der Meer, P., ... Charles-Dominique, P. (2008). Above-ground biomass and productivity in a rain forest of eastern South America. *Journal of Tropical Ecology*, 24(4), 355–366. <https://doi.org/10.1017/S0266467408005075>
- Condit, R., Pitman, N., Leigh, E. G., Chave, J., Terborgh, J., Foster, R. B., ... Muller-Landau, H. C. (2002). Beta-diversity in tropical forest trees. *Science*, 295(5555), 666–669. <https://doi.org/10.1126/science.1066854>
- Cottenie, K. (2005). Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters*, 8(11), 1175–1182. <https://doi.org/10.1111/j.1461-0248.2005.00820.x>
- De Cáceres, M., Legendre, P., Valencia, R., Cao, M., Chang, L. W., Chuyong, G., ... Kenfack, D. (2012). The variation of tree beta diversity across a global network of forest plots. *Global Ecology and Biogeography*, 21(12), 1191–1202. <https://doi.org/10.1111/j.1466-8238.2012.00770.x>
- Decaëns, T., Porco, D., James, S. W., Brown, G. G., Chassany, V., Dubs, F., ... Roy, V. (2016). DNA barcoding reveals diversity patterns of earthworm communities in remote tropical forests of French Guiana. *Soil Biology and Biochemistry*, 92, 171–183. <https://doi.org/10.1016/j.soilbio.2015.10.009>
- Dini-Andreote, F., Stegen, J. C., van Elsas, J. D., & Salles, J. F. (2015). Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proceedings of the National Academy of Sciences of the United States of America*, 112(11), E1326–E1332. <https://doi.org/10.1073/pnas.1414261112>
- Dini-Andreote, F., & van Elsas, J. D. (2013). Back to the basics: The need for ecophysiological insights to enhance our understanding of microbial behaviour in the rhizosphere. *Plant and Soil*, 373(1–2), 1–15. <https://doi.org/10.1007/s11104-013-1687-z>
- Douhan, G. W., Vincenot, L., Gryta, H., & Selosse, M. A. (2011). Population genetics of ectomycorrhizal fungi: From current knowledge to emerging directions. *Fungal Biology*, 115(7), 569–597. <https://doi.org/10.1016/j.funbio.2011.03.005>
- Dray, S., Legendre, P., & Peres-Neto, P. R. (2006). Spatial modelling: A comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, 196(3–4), 483–493. <https://doi.org/10.1016/j.ecolmodel.2006.02.015>
- Driscoll, C. T., & Schecher, W. D. (1990). The chemistry of aluminum in the environment. *Environmental Geochemistry and Health*, 12(1–2), 28–49. <https://doi.org/10.1007/BF01734046>
- Dumbrell, A. J., Nelson, M., Helgason, T., Dytham, C., & Fitter, A. H. (2010). Relative roles of niche and neutral processes in structuring a soil microbial community. *The ISME Journal*, 4(3), 337. <https://doi.org/10.1038/ismej.2009.122>
- Engel, P., & Moran, N. A. (2013). The gut microbiota of insects—diversity in structure and function. *FEMS Microbiology Reviews*, 37(5), 699–735. <https://doi.org/10.1111/1574-6976.12025>
- Esling, P., Lejzerowicz, F., & Pawlowski, J. (2015). Accurate multiplexing and filtering for high-throughput amplicon-sequencing. *Nucleic Acids Research*, 43(5), 2513–2524. <https://doi.org/10.1093/nar/gkv107>
- Ettema, C. H., & Wardle, D. A. (2002). Spatial soil ecology. *Trends in Ecology & Evolution*, 17(4), 177–183. [https://doi.org/10.1016/S0169-5347\(02\)02496-5](https://doi.org/10.1016/S0169-5347(02)02496-5)
- Fierer, N., Strickland, M. S., Liptzin, D., Bradford, M. A., & Cleveland, C. C. (2009). Global patterns in belowground communities. *Ecology Letters*, 12(11), 1238–1249. <https://doi.org/10.1111/j.1461-0248.2009.01360.x>
- Fliiegerova, K., Tapio, I., Bonin, A., Mrazek, J., Callegari, M. L., Bani, P., ... Boyer, F. (2014). Effect of DNA extraction and sample preservation method on rumen bacterial population. *Anaerobe*, 29, 80–84. <https://doi.org/10.1016/j.anaerobe.2013.09.015>
- Fragoso, C., & Lavelle, P. (1992). Earthworm communities of tropical rain forests. *Soil Biology and Biochemistry*, 24(12), 1397–1408. [https://doi.org/10.1016/0038-0717\(92\)90124-G](https://doi.org/10.1016/0038-0717(92)90124-G)
- Fujii, K. (2014). Soil acidification and adaptations of plants and microorganisms in Bornean tropical forests. *Ecological Research*, 29(3), 371–381. <https://doi.org/10.1007/s11284-014-1144-3>
- Galiana, N., Lurgi, M., Claramunt-López, B., Fortin, M. J., Leroux, S., Cazelles, K., ... Montoya, J. M. (2018). The spatial scaling of species interaction networks. *Nature Ecology & Evolution*, 2, 782–790. <https://doi.org/10.1038/s41559-018-0517-3>
- Geisen, S., Mitchell, E. A. D., Adl, S., Bonkowski, M., Dunthorn, M., Ekelund, F., ... Lara, E. (2018). Soil protists: A fertile frontier in soil biology research. *FEMS Microbiology Reviews*, 42(3), 293–323. <https://doi.org/10.1093/femsre/fuy006>
- Gilbert, B., & Bennett, J. R. (2010). Partitioning variation in ecological communities: Do the numbers add up? *Journal of Applied Ecology*, 47(5), 1071–1082. <https://doi.org/10.1111/j.1365-2664.2010.01861.x>
- Guardiola, M., Uriz, M. J., Taberlet, P., Coissac, E., Wangenstein, O. S., & Turon, X. (2015). Deep-sea, deep-sequencing: Metabarcoding extracellular DNA from sediments of marine canyons. *PLoS ONE*, 10(10), e0139633. <https://doi.org/10.1371/journal.pone.0139633>
- Haegeman, B., Hamelin, J., Moriarty, J., Neal, P., Dushoff, J., & Weitz, J. S. (2013). Robust estimation of microbial diversity in theory and

- in practice. *The ISME Journal*, 7(6), 1092. <https://doi.org/10.1038/ismej.2013.10>
- Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., & Martiny, J. B. (2012). Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nature Reviews Microbiology*, 10(7), 497. <https://doi.org/10.1038/nrmicro2795>
- Hättenschwiler, S., Aeschlimann, B., Couëteaux, M. M., Roy, J., & Bonal, D. (2008). High variation in foliage and leaf litter chemistry among 45 tree species of a neotropical rainforest community. *New Phytologist*, 179(1), 165–175. <https://doi.org/10.1111/j.1469-8137.2008.02438.x>
- Hinsinger, P., Bengough, A. G., Vetterlein, D., & Young, I. M. (2009). Rhizosphere: Biophysics, biogeochemistry and ecological relevance. *Plant and Soil*, 321(1–2), 117–152. <https://doi.org/10.1007/s11104-008-9885-9>
- Hubbell, S. P. (2001). *The unified neutral theory of biodiversity and biogeography*. Princeton, NJ: Princeton University Press.
- Hurlbert, A. H., & Stegen, J. C. (2014). When should species richness be energy limited, and how would we know? *Ecology Letters*, 17(4), 401–413. <https://doi.org/10.1111/ele.12240>
- Ingold, C. T. (2001). Range in size and form of basidiospores and ascospores. *Mycologist*, 15(4), 165–166. [https://doi.org/10.1016/S0269-915X\(01\)80010-0](https://doi.org/10.1016/S0269-915X(01)80010-0)
- Jiménez, J. J., Decaëns, T., Lavelle, P., & Rossi, J. P. (2014). Dissecting the multi-scale spatial relationship of earthworm assemblages with soil environmental variability. *BMC Ecology*, 14(1), 26. <https://doi.org/10.1186/s12898-014-0026-4>
- Jost, L. (2006). Entropy and diversity. *Oikos*, 113(2), 363–375. <https://doi.org/10.1111/j.2006.0030-1299.14714.x>
- Joung, Y. S., Ge, Z., & Buie, C. R. (2017). Bioaerosol generation by raindrops on soil. *Nature Communications*, 8, 14668. <https://doi.org/10.1038/ncomms14668>
- Lavelle, P., Bignell, D., Lepage, M., Wolters, V., Roger, P., Ineson, P. O. W. H., ... Dhillon, S. (1997). Soil function in a changing world: The role of invertebrate ecosystem engineers. *European Journal of Soil Biology*, 33(4), 159–193.
- Legendre, P. (2008). Studying beta diversity: Ecological variation partitioning by multiple regression and canonical analysis. *Journal of Plant Ecology*, 1(1), 3–8. <https://doi.org/10.1093/jpe/rtm001>
- Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, 129(2), 271–280. <https://doi.org/10.1007/s004420100716>
- Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., ... Loreau, M. (2004). The metacommunity concept: A framework for multi-scale community ecology. *Ecology Letters*, 7(7), 601–613. <https://doi.org/10.1111/j.1461-0248.2004.00608.x>
- Lemire, J. A., Harrison, J. J., & Turner, R. J. (2013). Antimicrobial activity of metals: Mechanisms, molecular targets and applications. *Nature Reviews Microbiology*, 11(6), 371. <https://doi.org/10.1038/nrmicro3028>
- Levin, S. A. (1992). The problem of pattern and scale in ecology: The Robert H. MacArthur award lecture. *Ecology*, 73, 1943–1967. <https://doi.org/10.2307/1941447>
- Mahé, F., de Vargas, C., Bass, D., Czech, L., Stamatakis, A., Lara, E., ... Siemensmeyer, T. (2017). Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. *Nature Ecology & Evolution*, 1(4), 0091. <https://doi.org/10.1038/s41559-017-0091>
- Martiny, J. B., Jones, S. E., Lennon, J. T., & Martiny, A. C. (2015). Microbiomes in light of traits: A phylogenetic perspective. *Science*, 350(6261), 9323. <https://doi.org/10.1126/science.aac9323>
- McMurdie, P. J., & Holmes, S. (2014). Waste not, want not: Why rarefying microbiome data is inadmissible. *PLoS Computational Biology*, 10(4), e1003531. <https://doi.org/10.1371/journal.pcbi.1003531>
- Meerts, P. (1999). The evolution of spore size in Agarics: Do big mushrooms have big pores? *Journal of Evolutionary Biology*, 12(1), 161–165. <https://doi.org/10.1046/j.1420-9101.1999.00020.x>
- Peay, K. G., Baraloto, C., & Fine, P. V. (2013). Strong coupling of plant and fungal community structure across western Amazonian rainforests. *The ISME Journal*, 7(9), 1852. <https://doi.org/10.1038/ismej.2013.66>
- Peay, K. G., Garbelotto, M., & Bruns, T. D. (2010). Evidence of dispersal limitation in soil microorganisms: Isolation reduces species richness on mycorrhizal tree islands. *Ecology*, 91(12), 3631–3640. <https://doi.org/10.1890/09-2237.1>
- Peay, K. G., Kennedy, P. G., & Bruns, T. D. (2008). Fungal community ecology: A hybrid beast with a molecular master. *AIBS Bulletin*, 58(9), 799–810.
- Peres-Neto, P. R., Legendre, P., Dray, S., & Borcard, D. (2006). Variation partitioning of species data matrices: Estimation and comparison of fractions. *Ecology*, 87(10), 2614–2625. [https://doi.org/10.1890/0012-9658\(2006\)87\[2614:VPOSDM\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[2614:VPOSDM]2.0.CO;2)
- Petchey, O. L., Beckerman, A. P., Riede, J. O., & Warren, P. H. (2008). Size, foraging, and food web structure. *Proceedings of the National Academy of Sciences of the United States of America*, 105(11), 4191–4196. <https://doi.org/10.1073/pnas.0710672105>
- Piña, R. G., & Cervantes, C. (1996). Microbial interactions with aluminum. *BioMetals*, 9(3), 311–316.
- Portillo, M. C., Leff, J. W., Lauber, C. L., & Fierer, N. (2013). Cell size distributions of soil bacterial and archaeal taxa. *Applied and Environmental Microbiology*, 79(24), 7610–7617. <https://doi.org/10.1128/AEM.02710-13>
- Powell, J. R., Karunaratne, S., Campbell, C. D., Yao, H., Robinson, L., & Singh, B. K. (2015). Deterministic processes vary during community assembly for ecologically dissimilar taxa. *Nature Communications*, 6, 8444. <https://doi.org/10.1038/ncomms9444>
- R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ramirez, K. S., Leff, J. W., Barberán, A., Bates, S. T., Betley, J., Crowther, T. W., ... Bradford, M. A. (2014). Biogeographic patterns in below-ground diversity in New York City's Central Park are similar to those observed globally. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 281(1795), 20141988. <https://doi.org/10.1098/rspb.2014.1988>
- Réjou-Méchain, M., Tymen, B., Blanc, L., Fauset, S., Feldpausch, T. R., Monteagudo, A., ... Chave, J. (2015). Using repeated small-footprint LiDAR acquisitions to infer spatial and temporal variations of a high-biomass Neotropical forest. *Remote Sensing of Environment*, 169, 93–101. <https://doi.org/10.1016/j.rse.2015.08.001>
- Ricklefs, R. E. (2004). A comprehensive framework for global patterns in biodiversity. *Ecology Letters*, 7(1), 1–15. <https://doi.org/10.1046/j.1461-0248.2003.00554.x>
- Rillig, M. C., Antonovics, J., Caruso, T., Lehmann, A., Powell, J. R., Veresoglou, S. D., & Verbruggen, E. (2015). Interchange of entire communities: Microbial community coalescence. *Trends in Ecology & Evolution*, 30(8), 470–476. <https://doi.org/10.1016/j.tree.2015.06.004>
- Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., ... Walker, A. W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology*, 12(1), 87. <https://doi.org/10.1186/s12915-014-0087-z>
- Schuldt, A., Wubet, T., Buscot, F., Staab, M., Assmann, T., Böhnke-Kammerlander, M., ... Pietsch, K. (2015). Multitrophic diversity in a biodiverse forest is highly nonlinear across spatial scales. *Nature Communications*, 6, 10169. <https://doi.org/10.1038/ncomms10169>

- Silver, W. L., Neff, J., McGroddy, M., Veldkamp, E., Keller, M., & Cosme, R. (2000). Effects of soil texture on belowground carbon and nutrient storage in a lowland Amazonian forest ecosystem. *Ecosystems*, 3(2), 193–209. <https://doi.org/10.1007/s100210000019>
- Smith, T. W., & Lundholm, J. T. (2010). Variation partitioning as a tool to distinguish between niche and neutral processes. *Ecography*, 33(4), 648–655. <https://doi.org/10.1111/j.1600-0587.2009.06105.x>
- Sommeria-Klein, G., Zinger, L., Taberlet, P., Coissac, E., & Chave, J. (2016). Inferring neutral biodiversity parameters using environmental DNA data sets. *Scientific Reports*, 6, 35644. <https://doi.org/10.1038/srep35644>
- Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., ... Konopka, A. (2013). Quantifying community assembly processes and identifying features that impose them. *The ISME Journal*, 7(11), 2069. <https://doi.org/10.1038/ismej.2013.93>
- Stegen, J. C., Lin, X., Konopka, A. E., & Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *The ISME Journal*, 6(9), 1653. <https://doi.org/10.1038/ismej.2012.22>
- Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). *Environmental DNA: For biodiversity research and monitoring*. Oxford, UK: Oxford University Press. <https://doi.org/10.1093/oso/9780198767220.001.0001>
- Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, 21(8), 2045–2050. <https://doi.org/10.1111/j.1365-294X.2012.05470.x>
- 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(5), 1105–1109. <https://doi.org/10.1007/BF00037152>
- Taberlet, P., Prud'homme, S. M., Campione, E., Roy, J., Miquel, C., Shehzad, W., ... Melodelima, C. (2012). Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. *Molecular Ecology*, 21(8), 1816–1820. <https://doi.org/10.1111/j.1365-294X.2011.05317.x>
- Tedersoo, L., Bahram, M., Cajthaml, T., Pölme, S., Hiiesalu, I., Anslan, S., ... Abarenkov, K. (2016). Tree diversity and species identity effects on soil fungi, protists and animals are context dependent. *The ISME Journal*, 10(2), 346. <https://doi.org/10.1038/ismej.2015.116>
- Tedersoo, L., Bahram, M., Pölme, S., Kõljalg, U., Yorou, N. S., Wijesundera, R., ... Smith, M. E. (2014). Global diversity and geography of soil fungi. *Science*, 346(6213), 1256688. <https://doi.org/10.1126/science.1256688>
- Tucker, C. M., Shoemaker, L. G., Davies, K. F., Nemergut, D. R., & Melbourne, B. A. (2016). Differentiating between niche and neutral assembly in metacommunities using null models of β -diversity. *Oikos*, 125(6), 778–789. <https://doi.org/10.1111/oik.02803>
- Turner, B. L., & Engelbrecht, B. M. (2011). Soil organic phosphorus in lowland tropical rain forests. *Biogeochemistry*, 103(1–3), 297–315. <https://doi.org/10.1007/s10533-010-9466-x>
- Tymen, B., Réjou-Méchain, M., Dalling, J. W., Fauset, S., Feldpausch, T. R., Norden, N., ... Chave, J. (2016). Evidence for arrested succession in a liana-infested Amazonian forest. *Journal of Ecology*, 104(1), 149–159. <https://doi.org/10.1111/1365-2745.12504>
- Vellend, M. (2010). Conceptual synthesis in community ecology. *The Quarterly Review of Biology*, 85(2), 183–206. <https://doi.org/10.1086/652373>
- Vellend, M., Srivastava, D. S., Anderson, K. M., Brown, C. D., Jankowski, J. E., Kleynhans, E. J., ... Myers-Smith, I. H. (2014). Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos*, 123(12), 1420–1430. <https://doi.org/10.1111/oik.01493>
- Veresoglou, S. D., Halley, J. M., & Rillig, M. C. (2015). Extinction risk of soil biota. *Nature communications*, 6, 8862. <https://doi.org/10.1038/ncomms9862>
- Wardle, D. A. (2002). *Communities and ecosystems: Linking the aboveground and belowground components*. New York, NY: Princeton University Press.
- Williams, R. J., & Martinez, N. D. (2000). Simple rules yield complex food webs. *Nature*, 404(6774), 180. <https://doi.org/10.1038/35004572>
- Woodward, G., Ebenman, B., Emmerson, M., Montoya, J. M., Olesen, J. M., Valido, A., & Warren, P. H. (2005). Body size in ecological networks. *Trends in Ecology & Evolution*, 20(7), 402–409. <https://doi.org/10.1016/j.tree.2005.04.005>
- Wu, T., Ayres, E., Bardgett, R. D., Wall, D. H., & Garey, J. R. (2011). Molecular study of worldwide distribution and diversity of soil animals. *Proceedings of the National Academy of Sciences of the United States of America*, 108(43), 17720–17725. <https://doi.org/10.1073/pnas.1103824108>
- Yang, C., Wang, X., Miller, J. A., de Blécourt, M., Ji, Y., Yang, C., ... Douglas, W. Y. (2014). Using metabarcoding to ask if easily collected soil and leaf-litter samples can be used as a general biodiversity indicator. *Ecological Indicators*, 46, 379–389. <https://doi.org/10.1016/j.ecolind.2014.06.028>
- Zinger, L., Chave, J., Coissac, E., Iribar, A., Louisanna, E., Manzi, S., ... Taberlet, P. (2016). Extracellular DNA extraction is a fast, cheap and reliable alternative for multi-taxa surveys based on soil DNA. *Soil Biology and Biochemistry*, 96, 16–19. <https://doi.org/10.1016/j.soilbio.2016.01.008>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Zinger L, Taberlet P, Schimann H, et al. Body size determines soil community assembly in a tropical forest. *Mol Ecol*. 2019;28:528–543. <https://doi.org/10.1111/mec.14919>