

Methods and approaches to advance soil macroecology

White, H. J., León-Sánchez, L., Burton, V. J., Cameron, E. K., Caruso, T., Cunha, L., Dirilgen, T., Jurburg, S. D., Kelly, R., Kumaresan, D., Ochoa-Hueso, R., Ordonez, A., Phillips, H. R. P., Prieto, I., Schmidt, O., & Caplat, P. (2020). Methods and approaches to advance soil macroecology. *Global Ecology & Biogeography*. https://doi.org/10.1111/geb.13156

Published in: Global Ecology & Biogeography

Document Version: Peer reviewed version

Queen's University Belfast - Research Portal:

Link to publication record in Queen's University Belfast Research Portal

Publisher rights

Copyright 2020 Wiley. This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Open Access

This research has been made openly available by Queen's academics and its Open Research team. We would love to hear how access to this research benefits you. – Share your feedback with us: http://go.qub.ac.uk/oa-feedback

¹ Methods and approaches to advance

² soil macroecology

3 Abstract

4 Motivation and aim: Soil biodiversity is central to ecosystem function and services. It represents 5 most of terrestrial biodiversity and at least a quarter of all biodiversity on Earth. Yet research into 6 broad, generalisable spatial and temporal patterns of soil biota has been limited compared to 7 aboveground systems due to complexities of the soil system. We review the literature and identify key 8 considerations necessary to expand soil macroecology beyond the recent surge of global maps of soil 9 taxa, so that we can gain greater insight into the mechanisms and processes shaping soil biodiversity. 10 We focus primarily on three groups of soil taxa (earthworms, mycorrhizal fungi and soil bacteria) that 11 represent a range of body sizes and ecologies, and, therefore, interact with their environment at 12 different spatial scales.

Results: The complexities of soil, including fine-scale heterogeneity, 3D habitat structure, difficulties
with taxonomic delimitation, and the wide-ranging ecologies of its inhabitants, require the classical
macroecological toolbox to be expanded to consider novel sampling, molecular identification,
functional approaches, environmental variables, and modelling techniques.

Main conclusion: Soil provides a complex system within which to apply macroecological research, yet it is this property which itself makes soil macroecology a field ripe for innovative methodologies and approaches. To achieve this, soil-specific data, spatiotemporal, biotic, and abiotic considerations are necessary at all stages of research, from sampling design to statistical analyses. Insights into whole ecosystems and new approaches to link genes, functions and diversity across spatial and temporal scales, alongside methodologies already applied in aboveground macroecology, invasion ecology and aquatic ecology, will facilitate the investigation of macroecological processes in soil biota, which is
key to understanding the link between biodiversity and ecosystem functioning in terrestrial
ecosystems.

26 Introduction

27 Macroecology strives to understand the generality of emergent patterns of aggregate ecological entities (McGill, 2019). The area has moved beyond large scale biodiversity maps for many 28 29 aboveground taxa, addressing process-based patterns including, latitudinal gradients of diversity 30 (Stehli, Douglas & Newell, 1969), species-area relationships (Rosenzweig, 1995), species-energy relationships (Gaston, 2000) and broad-scale responses of diversity to anthropogenic pressures 31 32 (Newbold et al., 2015), across spatial and temporal scales. There are, however, still significant 33 knowledge gaps surrounding many questions of macroecology and biogeography, and new techniques 34 to investigate these, in soil ecology (Eisenhauer et al., 2017).

35 Soil biodiversity regulates nutrient cycling and makes other central contributions to net primary 36 production and carbon sequestration (Bardgett & Wardle, 2010; Wagg, Bender, Widmer & van der 37 Heijden, 2014) across spatial scales (Delgado-Baquerizo et al. 2020). Given the importance of soil 38 biodiversity, knowledge of spatial and temporal distributions of soil organisms is essential for both 39 management of terrestrial ecosystems and predicting future ecosystem functioning under global 40 environmental change (Wagg et al., 2014; Crowther et al., 2019). Recently, there has been an 41 increased interest in describing the generality of emergent patterns of soil biodiversity (e.g. 42 Shoemaker, Locey & Lennon, 2017; Shade et al., 2018; Thakur et al., 2019), although, difficulties in 43 observing many soil organisms directly has limited this. Further, it has been challenging to formulate 44 a common macroecological framework across soil taxa given the sheer range in body sizes and 45 ecologies, from bacteria to macrofauna, which vary in the scale at which they perceive the environment. Methodological and taxonomic challenges have hindered macroecological research into 46 47 belowground systems concerning species richness, spatial distributions, ecosystem functioning and

spatial scaling (Schröder, 2008; Decaëns, 2010; Shade et al., 2018). The challenges of soil
macroecology, therefore, arise not only across spatial and temporal scales, but across the

organisational hierarchy of species, communities, and ecosystems.

50

At a species level, species distribution models (SDMs) are frequently employed to relate the spatial 51 52 distribution of aboveground species to their environment (Elith & Leathwick, 2009), yet there are very few examples of SDMs applied to soil-dwelling species other than earthworms (e.g. Palm, van 53 Schaik & Schröder, 2013; Marchán et al., 2015). For many soil taxa, this is a result of not having 54 55 sufficient species-level occurrence data from various geographic locations spanning environmental 56 gradients at relevant spatial scales (Schröder, 2008). For other taxa (i.e. collembola, but see Caruso et 57 al., 2009) these data are already available, which begs the question of why the modelling of 58 belowground species distributions has been so limited. There is, therefore, a huge opportunity to 59 expand soil macroecological research through species distributions modelling, as long as we are able 60 to overcome the challenges of the soil system itself.

61 Community biodiversity patterns include well-established macroecological relationships, such as the species-area relationship and latitudinal gradients (e.g. Rosenzweig, 1995), distance-decay curves 62 63 (Nekola and White, 1999), as well as models linking biodiversity patterns to environmental variables 64 such as the species-energy relationship (e.g. Evans et al. 2008). These relationships have formed the 65 basis of macroecological theory and our knowledge of how diversity is arranged in space and time, at 66 a range of spatial, temporal, and hierarchical scales. Although community assembly in soil organisms 67 has been investigated (Caruso et al., 2012; Dirilgen et al. 2018), generalisable spatial and temporal 68 patterns of soil biodiversity have yet to be determined to the degree that they have above ground (but 69 see Caruso et al., 2019). For microbes in particular, studies have tended to focus on individual taxa 70 and metabolic genes, yet it is the emergent properties of communities which influence ecosystem functioning and services (Ladau and Eloe-Fadroush, 2019), and, therefore, need to be considered if 71 we are to expand the field of soil macroecology. 72

73 At the ecosystem level, there have been huge advancements in our ability to model whole ecosystem functioning, particularly within the context of global change (Allan et al., 2015; De Laender et al., 74 2016). The development of measures of functional diversity (the range and variation of ecological 75 traits present within an area) has facilitated this. Aboveground, these measures have been used to 76 77 investigate key conservation issues such as biotic homogenisation (White et al., 2018a). Given the 78 importance of soil and soil biodiversity in ecosystem functioning and global cycles (Delgado-79 Baquerizo et al., 2020), functional traits have much to offer soil macroecology. The measurement of 80 functional traits is becoming increasingly feasible from a technical perspective (e.g. proteomics and 81 isotope tracer-enabled analysis of proteins and metabolites, trait database availability), and they can 82 be directly linked to ecosystem level processes (Hooper et al., 2005). There is a need, therefore, to 83 continue to develop methods to measure functional diversity and broad scale functioning of soil 84 organisms, and understand, for example, how these respond to environmental gradients. Applied 85 properly to soil organisms, macroecological approaches will provide us with another tool to tackle global challenges related to soil functioning and soil protection. 86

⁸⁷ Aims and scope of this review

88 To move beyond mapping of soil taxa in space and identify broad, generalisable spatial and temporal 89 patterns of soil biodiversity and their underlying mechanisms, we need to consider: the biodiversity 90 data; the spatio-temporal context of the system and mechanisms under investigation; the biotic 91 components shaping these patterns; and the abiotic components shaping these patterns. We discuss 92 each of these sets of considerations and review current methods and future directions for each. Our 93 intention is not to provide a systematic review of macroecological patterns of multiple soil taxa, but 94 instead, to focus on the specific characteristics of the soil system that have limited soil macroecology, 95 and review and propose methods that would help address these limitations. To accomplish this, we 96 use earthworms, mycorrhizal fungi and soil bacteria as focal taxa. These groups have all had their 97 global biodiversity modelled (earthworms, Phillips et al., 2019; fungi, Kivlin et al., 2011; Tedersoo et al., 2014; Pärtel et al., 2017; soil bacteria, Delgado-Baquerizo et al., 2018), and they reflect the 98

99 diversity of taxa which inhabit soil through their variation in body size, life forms and ecology which
100 leads to differences in the spatial scale at which they perceive and interact with their environment.
101 They, therefore, provide a range of robust examples of the issues that have hindered soil
102 macroecological research to date, and the diversity in approaches required to overcome these
103 difficulties.

¹⁰⁴ Data considerations

105 Data availability

106 Initiatives exist to establish standardised sampling protocols to collect new data on soil biodiversity 107 (Philippot, Ritz, Pandard, Hallin & Martin-Laurent, 2012), e.g. the Earth Microbiome Project 108 (Thompson et al., 2017) and EcoFINDERS (Lemanceau, 2011), as well as bring together existing soil 109 biodiversity data (see Table 1) into accessible databases and frameworks, e.g. the Global Soil 110 Biodiversity Initiative (www.globalsoilbiodiversity.org) and Ramirez et al. (2015). Pärtel et al. (2017) use a standardised dataset of metabarcode information (MaarjAM database, Öpik et al., 2010) to 111 model spatial variation in species pool, local and dark diversity (members of the species pool which 112 are absent locally) alongside their respective environmental correlates. These standardised data 113 provide a suitable means for investigating large scale biodiversity patterns (e.g. species-area 114 relationships, latitudinal gradients etc.), and, in fact, better satisfy the criteria of comparability across 115 116 sites and studies than many datasets of macro-organism distributions which show large variation in sampling methodology and effort. 117

The Global Soil Biodiversity Initiative and Ramirez et al. (2015), on the other hand, aim to maximise the potential of existing data by establishing global platforms that combine databases of all types of soil biodiversity data, including molecular, taxonomic, and morphological measurements and traits. Assembling data from multiple sources is complicated due to variability in taxonomic resolutions, unresolved taxonomies (Cameron, Decaëns, Lapied, Porco & Eisenhauer, 2016) and lack of 123 standardisation of sampling techniques that cause technical factors (e.g. sampling protocol, primer and 124 sequencing platform) to be an important source of inter-study variability (Ramirez et al., 2018). Some 125 studies choose to only include data collected using specific methodologies to reduce inter-study 126 variability (e.g. European earthworm diversity maps, Rutgers et al., 2016; 2019), yet, non-127 standardised datasets can still provide important global insights into the ecological preferences and 128 geographical ranges of species (Ramirez et al., 2015). These data will complement standardised 129 sampling protocols when analysed appropriately, for example using meta-analytical or machine 130 learning approaches (Hendershot, Read, Henning, Sanders & Classen, 2017; Ramirez et al. 2018). 131 Substantial geographic gaps in sampling, however, need to be addressed (Cameron et al., 2018) if we are to determine generalisable patterns of biodiversity, and model predictions should be refined and 132 updated as empirical datasets become increasingly available. 133

134 Measuring and defining soil biodiversity

To study many macroecological patterns such as species-abundance distributions, species-area 135 136 relationships and latitudinal gradients, we first need to identify the species. However, challenges 137 remain with identification of many soil taxa (Soininen, 2012); the huge diversity of soil organisms 138 combined with their small size (Schaefer et al., 2010), has led to a distinct lack of taxonomic knowledge of soil biodiversity (Decaëns, Lavelle & Jiménez, 2008; Phillips et al., 2017). 139 Biogeographic realms at multiple spatial scales, therefore, exist, where every new place sampled, 140 141 especially remote locations, can give huge volumes of undescribed species or sequences that cannot 142 be matched to any known species or genus-level taxon (Decaëns, 2010; Fierer, 2017). Indeed, even urban systems such as Central Park in New York City, can harbour large volumes of undescribed soil 143 biodiversity (Ramirez et al. 2014), and earthworms, a comparatively well-studied group with 144 145 relatively low diversity, have had a large number of cryptic species revealed through DNA barcoding (King, Tibble & Symondson, 2008). 146

147 Gene-based diversity assessments of environmental DNA (eDNA) are a promising toolset for

148 facilitating large scale sampling of soil diversity (Bik et al., 2012; Deiner et al., 2017). Species can be

149 identified using short metabarcodes and small samples of soil (e.g. earthworms, Bienert et al., 2012). eDNA has already furthered macroecological research, for example, through the detection of 150 earthworm diversity at the landscape-scale, revealing impacts of abiotic factors not detected using 151 traditional survey methods (Pansu et al., 2015). Nevertheless, the degree to which eDNA-based 152 153 diversity estimates capture the living soil biota (rather 'relic' DNA) is unclear (Lennon et al., 2018; 154 Ruppert, Kline and Rahman, 2019). Furthermore, this technique can vary with soil organic matter 155 content and type, complicating the comparison of biodiversity patterns across environments and 156 highlighting the continued, pressing need for the standardization of methodologies (Philippot et al., 157 2012; Geisen et al., 2019).

158 In particular, many molecular studies use targeted sequencing of specific marker genes or gene regions, which serve as 'barcodes'. Sequence differences within the "variable regions" of the marker 159 160 genes are used to identify operational taxonomic units (OTUs), which may serve to delineate taxa 161 below the genus level (e.g. earthworms, Pansu et al., 2015; fungi, Pellissier et al., 2014; Tedersoo et 162 al., 2014). However, there is no universal threshold for an eukaryotic OTU, and this may well be clade-specific, and dependent on the barcode length and region of choice (Mysara et al., 163 2017). Despite recent technological advances in sequencing and bioinformatic pipelines to process 164 high-throughput amplicon sequences, considerable challenges remain when using gene-based 165 methodologies (Nesme et al., 2016). Errors or biases can arise from different DNA extraction methods 166 (İnceoğlu, Hoogwout, Hill & van Elsas, 2010), primer specificity, sequencing chemistries (i.e. short 167 read length, Tedersoo et al., 2015; Tremblay et al., 2015), and bioinformatic processing, as well as 168 169 difficulties with OTU delimitation and chimeras (Orgiazzi, Dunbar, Panagos, de Groot & Lemanceau, 170 2015), and lack of matching current database entries (e.g. fungi; Tedersoo et al., 2014, bacteria; Ramirez et al., 2014; Thompson et al., 2017). These can confound diversity estimates in cross-study 171 comparisons (Fierer, 2017; Thompson et al., 2017). Recent developments in clustering algorithms, i.e. 172 173 a reference-free approach to resolve sequence variants at a single nucleotide resolution (Callahan, 174 McMurdie & Holmes, 2017; e.g. Deblur, Amir et al., 2017), could offer stable identifiers across different studies (Thompson et al., 2017), facilitating data reusability and integration of multiple data 175

sets (Amir et al., 2017), thus maximising the potential of gene-based surveys in soil macroecology
(Geisen et al., 2019). Alternatively, DNA-targeted enrichment (bait capture) allows more efficient
recovery of sequence information, not relying on sequence affinity as strongly as PCR (Dowle et al.
2016). DNA enrichment can be applied to a range of different sample sources, in individual or pooled
samples and can be used for recovering exogenous DNA present in environmental samples (Shokralla
et al. 2016).

182 Many studies now focus on alternative dimensions of biodiversity to taxonomic diversity. Functional 183 diversity, which measures the diversity and range of traits within a community, can be closely linked 184 to ecosystem functioning (Heemsbergen et al., 2004; Hooper et al., 2005). Shifting from taxonomic to 185 functional information can provide a way to unify the study of soil macroecology: the units are no 186 longer species, or taxa, but functions and the diversity of functions expressed, which are relevant 187 across taxonomic groups. Macroecological patterns that emerge from traits include trait-area and -188 time relationships (e.g. White et al., 2018b), and functional diversity gradients (e.g. Meynard et al., 189 2011), which further our knowledge on the spatial, temporal and environmental structuring of 190 ecological traits and ecosystem functioning (Violle et al., 2014), and provide new insights into community assembly theory (Smith et al., 2013). Due to the key ecosystem functions that soil 191 biodiversity provides (Heemsbergen et al., 2004), we advocate a concerted shift towards functional 192 approaches within soil macroecology through the following methodological and data options. 193

194 Trait databases facilitate the investigation of trait composition of soil communities. Morphological, 195 physiological and phenological traits are present in the literature for many soil invertebrates due to 196 their relative ease of measurement (e.g. carabid beetles, Barbaro & van Halder, 2009: earthworms, Hedde et al., 2012; collembola, Martins da Silva et al., 2016; Bonfanti et al., 2018: ants, Bishop et al., 197 198 2016). For soil bacteria and fungi, on the other hand, characterisation of morphological, physiological and phenological traits can be more challenging since isolation of individual species for trait 199 measurements is not feasible in most cases (Krause et al., 2014, Zanne et al., in press). Two recently 200 published, large global databases, FungalRoot (Soudzilovskaia et al. 2020) and the Fungal Functional 201 Database (FUNfun, Zanne et al. 2020) however, will facilitate better understanding of trait 202

composition of fungal communities, their interactions with plants, and their effects on worldwide on
ecosystem functioning (Powell & Rillig 2018). However, as with many taxa, fungal trait databases are
often incomplete (Zanne et al. 2020). There has also been a concerted effort to quantify microbial
functions in soil due to their importance for key ecosystem functions (Aguilar-Trigueros et al. 2015;
Zanne et al. 2020) through standardized 'omics' and enzymatic approaches (Dawson et al., 2019), as
well as the measure of climate tolerances to investigate life history trait trade offs at large spatial
scales (Maynard et al., 2019).

210 'Omic'-based functional analyses (e.g. proteomics, metabolomics) have been used to determine 211 broad-scale patterns in fungi (Tedersoo et al., 2014) and bacteria (Fierer & Jackson, 2006; Delgado-212 Baquerizo et al., 2018). These methods enable the categorisation of soil microbes by their ecological 213 strategies (e.g. Fierer, 2017), therefore bypassing the taxonomic deficit problem and focusing on 214 functional diversity. Functional 'omics' approaches can be used in experimental community ecology 215 studies to provide information on potential ecosystem functioning (Maron, Ranjard, Mougel & 216 Lemanceau, 2007), such as the diversity of protein-coding gene categories e.g. antibiotic resistance 217 genes and osmoregulatory genes (Fierer et al., 2012; Bahram et al., 2018), and soil enzyme substrate specificity (Caldwell, 2005) in nutrient cycling, as well as community structure and dynamics within 218 the soil (Arsène-Ploetze, Bertin & Carapito, 2015). Unlike gene-based approaches, protein-based 219 'omic' approaches (metaproteomics) capture the active component of soil biomass (Blagodatskaya & 220 Kuzyakov, 2013), and thus avoid overestimating diversity from dormant or dead biomass (Carini et 221 222 al., 2016; but see Papp et al., 2018).

Despite limitations (i.e., transcriptomics may reveal potential, rather than fulfilled functions, Prosser,
2015), elucidating the underlying molecular mechanisms supporting functions could help to unlock
the functional networks that interact to sustain soil properties (e.g. Bonfante and Genre, 2010). Semicontrolled experiments can build novel bridges with complex natural systems. For example, we
recommend mesocosm experiments coupled with "omics" approaches designed to identify
longitudinal biological responses of soil biota to microbiomes (e.g. soil and rhizosphere) and plants,
although destructive sampling is a limitation. Further still, defining a global transcriptome-based

230 delineation of functional entities, and obtaining a landscape of their similarities and differences based on differential expression of genes across different combinations of hierarchical levels and abiotic 231 factors, for example, could help to derive the sets of genes involved in specific ecological processes 232 that could be targeted in the field using metagenomics as a ground-truthing approach. Integrating 233 234 multiple "omic" approaches, therefore, into community studies will lead to a rapid transformation in our empirical understanding of soil functioning and interactions (Swenson and Jones, 2017), and 235 236 complementary approaches to DNA sequencing remain a priority for soil biodiversity research 237 (Eisenhauer et al., 2017) as they will reveal new information on the mechanisms underlying 238 ecosystem processes.

239

240 Spatio-temporal considerations

241 Scale and structure

Intra- and interspecific aggregations (Ettema & Wardle, 2002), which result from the microscale 242 243 heterogeneity of the soil system and limited dispersal of organisms, structure soil communities across spatial scales (Noguez et al. 2005; Decaëns, 2010; Bach et al., 2018; Thakur et al., 2019). In a 244 Brazilian agricultural system, for example, geographical distance explained nearly 18-times more 245 variance in soil fungal community composition than environmental factors such as soil and climate 246 247 characteristics (Gumiere, Durrer, Bohannan & Andreote, 2016). Biogeography of soil bacteria has been investigated at a range of scales from continental and global scales (e.g. Fierer & Jackson, 2006; 248 249 Barberán et al., 2012; Bahram et al., 2018) to the landscape scale (Bru et al., 2011; Pasternak et al., 2013), and even at the centimeter scale (O'Brien et al., 2016). By taking multiple soil cores separated 250 251 by only a few centimetres at sampling plots spaced 6 m apart, O'Brien et al. (2016) showed that there was extreme patchiness in community structure at the centimetre scale, but more general patterns in 252 253 diversity, linked to fertilisation, were observed at the plot-level. This scale-dependent heterogeneity 254 arises from the operating scale of mechanisms underlying community assembly. This needs to be

255 accounted for in sampling design for investigations of spatial patterns of soil diversity with multiscale sampling protocols (e.g. Rasmussen et al., 2018). Additionally, macroecological analyses of soil 256 diversity need to use techniques that model space, such as autoregressive models (Beale, Lennon, 257 258 Yearsley, Brewer & Elston, 2010), to address the strong spatial structuring in soil ecosystems, as 259 demonstrated by spatial analyses of earthworm abundance, biomass and diversity (Joschko et al. 260 2006), and ectomycorrhizal fungi (Pickles, 2010). Similarly, techniques such as Principal Coordinates 261 of Neighbourhood Matrices (PCNM) are useful in differentiating between spatial and environmental 262 structuring of communities at different spatial scales (e.g. Columbian earthworm communities, 263 Jiménez, Decaëns, Lavelle and Rossi, 2014).

264 The scale at which similar processes act varies between taxonomic groups due to variation in body 265 size and life history, and is an important consideration when studying spatial patterns of soil 266 communities (Montagna et al., 2018). When studying landscape scale processes, Jackson and Fahrig 267 (2012) highlight the concept of the 'scale of effect' (the scale at which an ecological response is best 268 predicted by the habitat structure) and recommend sampling areas far enough apart to ensure sample 269 points are independent (i.e. taxa in one sample point do not directly interact with those in others). 270 Thakur et al. (2019) present a conceptual framework splitting soil into spatial compartments at which 271 different nested groups of soil organisms can be studied to test biodiversity theories: soil, where all size groups of organisms can be sampled; hotspots (rhizosphere and drilosphere), where meso- and 272 273 micro-fauna can be sampled; and microsites (root tips and aggregates), where micro-fauna can be sampled. This compartmentalisation allows integration of body size into the investigation of classical 274 275 ecological theory. Beyond simply considering 'scales of effect' and spatial compartmentalisation, we 276 suggest researchers further draw upon work on investigating scaling issues in complex landscapes to build on the conceptual framework of Thakur et al. (2019) and inform sampling designs. For example, 277 278 the problem of "coarse-graining" occurs when fine-scale information is aggregated to larger scales to 279 reduce model complexity (Newman et al. 2019). By tracking how this leads to loss of information 280 explicitly and investigating scaling relationships, corrections for statistical biases may be possible (Wu, 2004; Newman et al. 2019). 281

282 The species-area relationship is a classical macroecological relationship which is well-established in aboveground biota. Judas (1988) applied the species-area relationship to published data on European 283 lumbricid earthworm diversity, revealing a lower scaling exponent than those established in 284 285 aboveground systems. One approach to investigate species-area relationships for ectomycorrhizal 286 fungi is to use tree 'islands' of host plants, which, due to their obligate symbiosis, create an island 287 species-area relationship (Peay et al., 2007). This approach, however, may not be transferable to other 288 microbial organisms. Taxa-area relationships have also been applied to soil bacteria and fungi using 289 samples taken from the corners of four nested quadrats (Sayer et al., 2013). This method, however, is 290 likely to underestimate diversity at each quadrat size as it is only partially sampled and is more 291 representative of a taxa-areal extent relationship. It does, however, still provide a useful study design 292 to investigate key macroecological patterns of spatial scaling and turnover of community composition 293 of multiple soil taxonomic groups.

294 As well as accumulation in space, temporal accumulation of species is an important part of 295 macroecology, and community turnover and can be investigated using the species-time relationship. 296 In above ground literature, the species-time relationship has been studied far less frequently than the 297 species-area relationship (White et al., 2006), and belowground, this lack of studies is even more 298 pronounced (Ladau and Eloe-Fadrosh, 2019). The temporal component of soil biodiversity, however, is critical to consider. As well as hot spots of diversity, Kuzyakov and Blagodatskaya (2015) highlight 299 300 'hot moments' of soil microbial diversity which can occur either occasionally or regularly as a 301 component of periodic processes within a system, often dependent on the temporally dynamic input of 302 C into soil. Additionally, species interactions operate at multiple temporal scales and vary within and 303 between taxa. For example, the connectivity of soil networks on abandoned arable land varies 304 temporally during restoration (Morriën et al. 2017). Hence, for temporal questions, sampling is 305 required at temporal scales relevant for the organisms of interest and the ecosystem properties with 306 which they interact (Ettema & Wardle, 2002; Bardgett et al., 2005; De Deyn & Van der Putten, 2005). 307 However, the appropriate scale remains an outstanding question in temporal biodiversity change of most soil organism groups (Eisenhauer et al., 2017; Shade et al., 2018). We, therefore, recommend 308

- 309 sampling at multiple time scales depending on the question at hand to determine the temporal scales at
- 310 which soil biodiversity varies, from macroevolutionary investigations (e.g. Schaefer and Caruso,
- 311 2019) to short-term dynamics (Kuzyakov and Blagodatskaya, 2015).

312 Dimensionality

To date macroecological patterns have been primarily conceptualised in a 2D context (e.g. species-313 314 area relationships, distance-decay relationships in geographic space, latitudinal gradients). However, 315 where depth and/or height have a strong influence on the biodiversity estimates and community dissimilarity (e.g. marine systems, tropical forests and soils), there is much to be gained by explicitly 316 317 extending these macroecological concepts to a third dimension. Given that soil community 318 composition varies strongly with soil depth (e.g. fungi and bacteria, Fierer et al., 2003; Eilers et al., 319 2012; ants, Wong and Guénard, 2017; microfauna, Pausch et al., 2018), and key ecosystem processes 320 are reliant on sub-soil fauna at different depths (Rumpel & Kogel-Knabner, 2011; Ward, 2016), the 321 integration of depth as a third dimension into macroecological theory is a key area to which soil 322 ecologists can contribute. Expanding soil macroecology into this third dimension will be particularly 323 important for obtaining comparable estimates of true soil diversity, for example through extension of 324 species accumulation curves to depth profiles. Furthermore, it will enhance our understanding of species-environment relationships, and provide insights into interaction dynamics in 3D space. In 325 326 order to do so, state-of-the-art sampling designs and modelling approaches will be required.

327 To address these questions it is essential that both soil biota and abiotic properties be sampled across 328 depth profiles. For macrofauna, sampling must involve multiple methods (e.g. pitfall trapping, direct 329 sampling, Berlese extraction, subterranean baiting) as each is known to capture different species (Fig. 330 1). For example, in a review of sampling methods Wong and Guénard (2017) identified seven studies 331 in which more than 10% (range 12.3 to 44.4%) of all ant species recorded were unique to subterranean samples; and would have been missed with conventional sampling methods. For smaller 332 333 macro-organisms, such as collembolans and arthropods, core samples provide samples from across the depth profile, but these must be carefully sorted to maintain information on sample depths as 334

335 combining multiple soil depths into a single sample can homogenize microscale variation (Grundmann et al., 2001). For microbial organisms, microsampling of smaller soil quantities across 336 the depth profile is advised (e.g. Dechesne et al., 2003), or 2D images of soil thin sections at multiple 337 338 depths can be used to count bacterial cell distributions and construct 3D distributions (Raynaud and 339 Nunan, 2014). Sampling of root-associated species such as mycorrhizal fungi, may be facilitated by 340 incorporating information on the depth distribution of roots (e.g. Sosa-Hernandez et al., 2018). Depth 341 distributions of many taxa can vary seasonally, for example, some earthworm species aestivate in 342 lower soil layers during the summer (Gerard, 1967). Therefore, where feasible, sampling should be 343 conducted seasonally to capture the full breadth of species depth distributions (e.g. Martay & Pearce-344 Higgins, 2018).

345 A final opportunity for macro-ecological approaches related to dimensionality arises from the 346 existence of soil communities *ex situ*, i.e. in soil micro-habitats above the soil proper. For example, 347 true soil-dwelling invertebrates including springtails and mites occur in suspended soil at various 348 heights (up to 35 m) in canopies of subtropical and temperate rainforests (Rodgers and Kitching, 349 1998; Lindo and Winchester, 2006). In Canadian Western redcedar trees, 18 of the 53 springtail (Collembola) species recorded from suspended soils did not occur in the forest floor (Lindo and 350 Winchester, 2006). In Indonesian oil palm plantations, suspended soil in frond axils of palm oil trunks 351 352 supported much higher densities and biomasses of soil microfauna and mesofauna than belowground soil, with suspended soils contributing an estimated 28% of the overall soil fauna metabolism in 353 354 plantations (Potapov et al., 2020). These contributions of suspended soils to overall diversity and 355 community turnover, suggest that there is much to be gained by extending soil sampling above-356 ground in forest ecosystems (Fig. 1).

357 Information from vertical sampling may be integrated into the investigation of

358 important macroecological patterns through the explicit incorporation of depth/height as a parameter

359 standard macroecological approaches e.g. Does community similarity decay more rapidly with depth

- than horizontal distance?. However, more complex models will be required to investigate species-
- 361 environment relationships in 3D domains. For such applications, it may useful to turn to the marine

362 literature where development of SDM techniques for 3D systems is already underway (e.g. Duffy &

363 Chown, 2017; Pérez-Costas 2019). Similarly, investigations into the scaling of interaction strengths

364 within species networks can differ between two- and three-dimensional systems in terrestrial and

365 marine realms, where advanced non-linear statistical approaches, 3D statistical point pattern models,

and 3D agent based modelling have been successfully applied (see Raynaud and Nunan 2014;

367 Barrios-O'Neill et al., 2019; Pawar et al., 2019).

368

369 Biotic considerations

370 Dispersal

371 The dispersal capacity of many soil biota, and thus the environment in which soil communities develop and interact, is often at a much finer resolution than the scale at which environmental 372 variables are commonly measured (Grundmann & Debouzie, 2000; Hendershot et al., 2017), and 373 varies tremendously between and within different soil taxa (e.g. microbes, dispersing as spores, can 374 375 generally disperse at larger scales than most soil animals). However, dispersal syndromes vary at 376 multiple phylogenetic scales (e.g., fungi vs. bacteria and phylotypes within bacteria in Archer et al. 377 2019). Therefore, sudden environmental changes that favour specific dispersal modes, may then favour certain taxa and so drive community structure through an intensified propagule pressure of 378 379 those taxa, thus dictating macroecological patterns. This process, as well as others, may operate at 380 multiple scales. For example, some biogeographical 'rules' that apply to the distribution of vertebrates 381 at the continental scale, e.g. range size-frequency distributions, can be detected in prokaryotes within an 8 m x 8 m quadrat (Noguez et al. 2005), conflicting with the traditional idea that soil microbes are 382 383 not limited by dispersal (Finlay, 2002). Overall, it is now well established that dispersal limitation is central to explain the distribution of all soil organisms, and that dispersal may differ at different 384 phylogenetic levels (Davison et al., 2015; Archer et al., 2019). It is, therefore, crucial to include 385 386 dispersal capacity into models of belowground species and to to take into account the fact that

dispersal capacity varies greatly with size, ecology and life stage of the organism (Ettema & Wardle,
2002; Soininen, McDonald & Hillebrand, 2007).

Dispersal kernels are a frequently used tool in invasion ecology that can be used to incorporate 389 dispersal into species distribution models (Meentemeyer, Anacker, Mark & Rizzo, 2008). These 390 391 describe the probability distribution of the distance travelled by an individual from a parent source and can be used to estimate the probability of colonisation (Franklin, 2010). Novel tools for 392 implementing cellular automata models into SDMs which map accessibility from source cells are 393 394 another promising recent development (Nobis & Normand, 2014). Although implementing such 395 models can be challenging as the dispersal characteristics of many soil organisms remain largely 396 unknown (Schröder, 2008), molecular techniques provide tools to measure whole community 397 dispersal of microbial systems (Peay, Garbelotto & Bruns, 2010; Peay, Schubert, Nguyen & Bruns, 398 2012). For example, within a biogeographical framework of plant host 'islands', Peay et al. (2012) use 399 next-generation sequencing of propagules to demonstrate a dispersal limitation of one kilometre 400 across a whole ectomycorrhizal fungal community. For earthworms on the other hand, visual tagging 401 methods (Butt and Lowe, 2007) and X-ray scanning have been used to measure dispersal in addition to molecular methods (Mathieu, Caro & Dupont, 2018). Using these newer technologies to determine 402 403 active and passive dispersal of soil organisms will allow movement-based theories of ecology to be tested (Thakur et al., 2019) that contribute to spatial and temporal biodiversity patterns across scales 404 405 (Gumiere et al., 2016; Dirilgen et al., 2018).

406 Interactions

Microbiota, including soil fungi and bacteria, are almost always sampled at scales larger than that at
which individuals interact, therefore lumping together sets of taxa that partition different resources, or
occupy different microhabitats (Fierer & Lennon, 2011; Bach, Williams, Hargreaves, Yang &
Hofmockel, 2018). This can greatly influence observed macroecological patterns, such as the speciesabundance distribution (Shoemaker et al., 2017). Multiscale sampling is recommended to provide
novel insights into the processes and mechanisms underlying spatial organisation of communities of

soil taxa of various body sizes. Using soil aggregate isolation techniques when sampling, for example,
will help characterise biodiversity of soil microbes relevant to the spatial scale at which individuals
interact and use resources (Bach et al., 2018).

Network analyses based on Spearman's rank correlations have been used to determine co-occurrences 416 417 of soil bacteria (Barberán, Bates, Casamayor & Fierer, 2012), but due to the complex threedimensional structure of soil, these networks may not always be indicative of real species interactions 418 (Morriën, 2016). The simulation approach used by Raynaud and Nunan (2014), for example, shows 419 420 that interspecific interactions between bacterial species are substantially lower than what you may 421 expect given the bacterial diversity frequently measured in soil samples (see Dimensionality section 422 above). For earthworms, stable isotopes have been used to determine interactions between invasive species of earthworms within a 2 km² area in the USA (Zhang et al., 2010), providing a sophisticated 423 424 tool to address the challenge of non-transparency of the soil medium in identifying biotic interactions 425 where they cannot easily be observed.

426 There is a momentum to improve aboveground models of species distributions by incorporating biotic interactions (e.g. Staniczenko, Sivasubramaniam, Suttle & Pearson, 2017). While challenging, this 427 research track is promising for soil macroecology. Indeed, the particular importance of addressing this 428 429 challenge for soils is highlighted by the substantial contribution of biotic interactions in shaping soil 430 organisms' distributions (Raymond, Wharton & Marshall, 2013; Bahram et al., 2018), and the 431 complex aboveground-belowground interactions that have been observed globally (De Deyn & Van 432 der Putten, 2005). Joint SDMs account for species interactions via correlation matrices of co-433 occurrence (Pollock et al., 2014) and can disentangle species interactions from environmental correlates in large multispecies datasets (Warton et al., 2015), holding promise for soil 434 435 macroecological studies.

Aboveground and belowground systems are linked by food web interactions and nutrient fluxes
through plant litter decomposition (Wardle et al., 2004; de Vries et al., 2013), modulation of soil
physical properties by plant diversity (Gould, Quinton, Weigelt, De Deyn & Bardgett, 2016), and

439 direct species species interactions (Bardgett & Wardle, 2010). Soil biodiversity data, therefore, can be extremely powerful when used in conjunction with aboveground community data (Ramirez et al., 440 2015). This is particularly true for mycorrhizal fungi, which form mutualisms with plants, and for 441 442 which host specificity is an important geographical constraint on species distributions (Sato et al., 443 2012). Hence, studies of community structure and biogeographic patterns of soil fungi are 444 significantly improved when data on plant diversity and distributions are included within their models 445 (Pellissier et al., 2014; Vályi et al., 2016; Koyama, Maherali and Antunes, 2019). To this end, the 446 FungalRoot database (Soudzilovskaia et al., 2020) provides key data on mycorrhizal associations with 447 which to study the macroecology and biogeography of these interactions.

448 Feedbacks between the aboveground and belowground systems, however, are difficult to predict at 449 large scales, as species interactions are complex (Wardle et al., 2004; De Deyn & Van der Putten, 450 2005), vary along latitudinal gradients of biodiversity (De Deyn & Van der Putten, 2005) and operate 451 over a hierarchy of temporal scales (Bardgett, Bowman, Kaufmann & Schmidt, 2005). Small spatial 452 and temporal scale experimental manipulations (e.g. Gould et al., 2016; León-Sánchez et al., 2018), 453 therefore, may not truly capture the complexities or multifunctionality of aboveground-belowground processes (De Deyn & Van der Putten, 2005), which also pose a challenge to traditional modelling 454 approaches. 455

456 To overcome the challenge of integrating network ecology research into macroecology, we encourage 457 the use of modelling approaches that facilitate complex interactions and pathways between multiple 458 variable types (Kissling & Schleuning, 2015). For soil systems in particular, it is important to develop 459 methods, such as correlation networks and structural equation models (SEMs), that emphasise biotic interactions but also include environmental effects, e.g. through latent and composite variables, to 460 461 study causal mechanisms involving variables that are difficult to measure at a spatial scale relevant for soil organisms, or themselves exhibit complex interactions (e.g. as seen for soil fertility in Siciliano et 462 al., 2014). SEMs are a useful correlative approach that can be easily implemented and characterise 463 464 complex pathways at the ecosystem level, including the complex interaction networks and feedback loops observed in soil-to-aboveground systems (Eisenhauer, Bowker, Grace & Powell., 2015). Grace 465

466 et al. (2010) present SEMs as an approach to address the challenge of eliciting generalisable patterns, such as those sought in macroecology, from heterogeneous system components. For example, SEMs 467 468 have been used to partition causal influences and determine the direct and indirect relationships 469 between geographic variables, soil characteristics, plant productivity/diversity and soil diversity at 470 both the continental and global scale for soil bacteria and fungi (Tedersoo et al., 2014; Delgado-471 Baquerizo, Powell et al., 2017), as well as to determine the causal mechanisms underlying ecosystem 472 functioning (Eisenhauer, Reich & Isbell, 2012) and food web stability (De Vries et al., 2012). They, 473 therefore, provide a useful tool for soil scientists and macroecologists to combine biotic and abiotic 474 factors into studies of causal patterns of soil diversity and functioning.

475 Abiotic considerations

476 Climate and microclimate

Fine-scale spatial structuring of soil species may occur where the macroscopic environment appears 477 uniform (Nielsen et al., 2010; Caruso, Taormina & Migliorini, 2012). For example, the spatial 478 structure of grass tussocks in pasture alter the microclimate of the soil which impacts earthworm 479 480 diversity (Mathieu et al., 2009). To model soil biodiversity, therefore, measurement of environmental heterogeneity is required at fine scales (e.g. measuring and modelling microclimates). Microclimate 481 482 modelling can be expanded to the macro-scale by using gridded continental-scale soil and weather 483 data to accurately predict hourly local microclimates at multiple soil depths using a mechanistic modelling framework (Kearney et al. 2014). Microclimate modelling has become more accessible 484 485 through advances in remote sensing, e.g. LiDAR, to quantify environmental covariates at high 486 resolutions (Lembrechts, Nijs & Lenoir, 2019), and development of freely available software and 487 code, e.g. the R package microclima (Maclean et al. 2019). Furthermore, combining microclimate 488 modelling frameworks with soil moisture simulation algorithms can provide accurate, high resolution soil moisture estimates for entire continents (Kearney & Maino, 2018). Moving beyond simple 489 490 correlative models towards mechanistic modelling is being encouraged within the SDM literature

(Buckley et al., 2010) and offers a promising alternative for predicting distributions of soil organisms
in particular, where small-scale spatial and temporal heterogeneity of the environment is often more
important than large-scale climatic variables (Dauber et al., 2005; Kearney et al., 2014).

494 Geodiversity and pedodiversity

Environmental factors other than climate are likely to be particularly ecologically relevant 495 496 determinants of species distributions belowground, and geological as well as biological resources 497 should be considered (Ibáñez, Krasilnikov and Saldaña, 2012). The incorporation of soil types, texture and geochemistry (i.e. pedodiversity) into analyses of soil biodiversity has been encouraged (Parker, 498 2010) and carried out at the local scale (e.g. earthworms, Decaëns & Rossi, 2008; Solomou et al., 499 2013; bacteria, Ranjard et al., 2010), as well as larger scales for earthworms (Rutgers et al., 2016). 500 501 Integration of point level soil characteristics data, such as the LUCAS dataset (Orgiazzi, Ballabio, Panagos, Jones & Fernández-Ugalde, 2018), can provide environmental information at coarse 502 503 resolutions (2 km x 2 km) but large geographical extents, i.e. continental, whilst machine learning 504 techniques combining soil and earth observation data can provide global gridded soil information at a 505 resolution of 250 m x 250 m (Hengl et al. 2017). The latter has recently been incorporated into models 506 of global earthworm diversity, but did not appear important in shaping community diversity, likely 507 due to the scale of the study (Phillips et al. 2019). Diversity of geophysical properties (i.e. 508 geodiversity) and pedodiversity provide opportunities to scale up soil biodiversity analyses. Spatial 509 soil information science has become particularly advanced and sophisticated statistical tools to predict 510 spatial patterns of soil properties (including salinity, soil moisture content and soil bulk density) offer novel opportunities to obtain predictor variables of soil organism distributions and diversity (Padarian 511 et al., 2020). We may expect an obvious link between geodiversity or pedodiversity variables and soil 512 513 biodiversity, however, incorporation of these factors into models has often been restricted to plants (e.g. Tukiainen et al., 2017; Bailey et al., 2018), although geochemical variables including salinity and 514 515 soil nitrate variables have been linked to nematode abundances in Antarctica (Poage et al., 2008), and

global soil bacteria richness peaks in neutral soils (Fierer & Jackson, 2006). Pedodiversity analyses,
therefore, offer an exciting opportunity to overcome a severe gap in soil macroecological research.

518 Recommendations and Future Perspectives

Although there has been a recent surge in broad-scale papers mapping soil biodiversity, the field of 519 soil macroecology needs to be advanced by emphasising process over pattern (Hanson et al., 2012). 520 Soils differ from aboveground systems in ways that have been challenging for the generalisations dear 521 to macroecology: they are characterised by high heterogeneity making data collection and analysis 522 523 difficult, and are inhabited by organisms that are incredibly variable in size and trophic roles, while being poorly known taxonomically. We have discussed how recent methods and data management 524 525 initiatives might help soil ecologists and macroecologists to collaborate more often. We demonstrate 526 that methodological considerations need to be made at all stages of investigation spanning 527 delimitation and quantification of diversity, spatial and temporal context, biotic considerations and 528 abiotic properties, and propose multiple approaches to deal with challenges within each of these sets 529 of considerations (Fig. 2). Most of these methods are applicable (and some already applied) 530 aboveground, but are particularly suited to address the incredible spatial and temporal variability of 531 biotic and abiotic conditions, combined with the scarcity of data, in soils. When these are overcome, 532 we expect that new rules may emerge from macroecological analysis of soils.

533 References

- 534
- Aguilar-Trigueros, C. A., Hempel, S., Powell, J. R., Anderson, I. C., Antonovics, J., Bergmann, J., ...
- 536 Rillig, M. C. (2015). Branching out: Towards a trait-based understanding of fungal ecology.
- 537 *Fungal Biology Reviews*, 29(1), 34-41.
- Allan, E., Manning, P., Alt, F., Binkenstein, J., Blaser, S., Blüthgen, N., ... Fischer, M. (2015). Land
- use intensification alters ecosystem multifunctionality via loss of biodiversity and changes to
- 540 functional composition. *Ecology Letters*, *18*(8), 834–843.

541	Amir, A., McDonald, D., Navas-Molina, J. A., Kopylova, E., Morton, J. T., Zech Xu, Z., Knight,
542	R. (2017). Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns.
543	MSystems, 2(2), e00191-16.

- Archer, S. D. J., Lee, K. C., Caruso, T., Maki, T., Lee, C. K., Cary, S. C., ... Pointing, S. B. (2019).
 Airborne microbial transport limitation to isolated Antarctic soil habitats. *Nature Microbiology*, 4(6), 925–932.
- Arsène-Ploetze, F., Bertin, P. N., & Carapito, C. (2015). Proteomic tools to decipher microbial
 community structure and functioning. *Environmental Science and Pollution Research International*, 22(18), 13599–13612.
- Bach, E. M., Williams, R. J., Hargreaves, S. K., Yang, F., & Hofmockel, K. S. (2018). Greatest soil
 microbial diversity found in micro-habitats. *Soil Biology and Biochemistry*, *118*(January), 217–
 226.
- Bahram, M., Hildebrand, F., Forslund, S. K., Anderson, J. L., Soudzilovskaia, N. A., Bodegom, P. M.,
 ... Bork, P. (2018). Structure and function of the global topsoil microbiome. *Nature*, *560*(7717),
 233–237.
- Bailey, J. J., Boyd, D. S., & Field, R. (2018). Models of upland species' distributions are improved by
 accounting for geodiversity. *Landscape Ecology*, *33*(12), 2071–2087.
- Barbaro, L., & Van Halder, I. (2009). Linking bird, carabid beetle and butterfly life-history traits to
 habitat fragmentation in mosaic landscapes. *Ecography*, *32*(2), 321–333.
- 560 Barberán, A., Bates, S. T., Casamayor, E. O., & Fierer, N. (2012). Using network analysis to explore
- 561 co-occurrence patterns in soil microbial communities. *ISME Journal*, 6(2), 343–351.
- Bardgett, R. D., Bowman, W. D., Kaufmann, R., & Schmidt, S. K. (2005). A temporal approach to
 linking aboveground and belowground ecology. *Trends in Ecology and Evolution*, 20(11), 634–
- 564

641.

- Bardgett, R. D., & Wardle, D. A. (2010). *Aboveground-belowground linkages: biotic interactions, ecosystem processes, and global change*. Oxford University Press.
- 567 Barrios-O'Neill, D., Kelly, R., Dick, J. T. A., Ricciardi, A., Macisaac, H. J., & Emmerson, M. C.
- 568 (2016). On the context-dependent scaling of consumer feeding rates. *Ecology Letters*, 19(6),
 569 668–678.
- Beale, C. M., Lennon, J. J., Yearsley, J. M., Brewer, M. J., & Elston, D. a. (2010). Regression
 analysis of spatial data. *Ecology Letters*, *13*(2), 246–64.
- 572 Bienert, F., De Danieli, S., Miquel, C., Coissac, E., Poillot, C., Brun, J. J., & Taberlet, P. (2012).
- 573 Tracking earthworm communities from soil DNA. *Molecular Ecology*, 21(8), 2017–2030.
- Bik, H. M., Porazinska, D. L., Creer, S., Caporaso, J. G., Knight, R., & Thomas, W. K. (2012, April).
 Sequencing our way towards understanding global eukaryotic biodiversity. *Trends in Ecology and Evolution*.
- 577 Bishop, T. R., Robertson, M. P., Gibb, H., van Rensburg, B. J., Braschler, B., Chown, S. L., ... Parr,
- 578 C. L. (2016). Ant assemblages have darker and larger members in cold environments. *Global*579 *Ecology and Biogeography*, 25(12), 1489–1499.
- Blagodatskaya, E., & Kuzyakov, Y. (2013). Active microorganisms in soil: Critical review of
 estimation criteria and approaches. *Soil Biology and Biochemistry*, 67, 192–211.
- Bonfante, P., & Genre, A. (2010). Mechanisms underlying beneficial plant–fungus interactions in
 mycorrhizal symbiosis. *Nature Communications*, 1(48).
- Buckley, L. B., Urban, M. C., Angilletta, M. J., Crozier, L. G., Rissler, L. J., & Sears, M. W. (2010).
 Can mechanism inform species' distribution models? *Ecology Letters*, *13*(8), 1041–1054.
- Butt, K. R., & Lowe, C. N. (2007). Presence of earthworm species within and beneath Lumbricus
 terrestris (L.) middens. *European Journal of Soil Biology*, *43*(SUPPL. 1), 57–60.

- 588Caldwell, B. A. (2005). Enzyme activities as a component of soil biodiversity: A review.
- 589 *Pedobiologia*, 49(6), 637–644.
- Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace
 operational taxonomic units in marker-gene data analysis. *ISME Journal*, *11*(12), 2639–2643.
- 592 Cameron, E. K., Decaëns, T., Lapied, E., Porco, D., & Eisenhauer, N. (2016). Earthworm databases
 593 and ecological theory: Synthesis of current initiatives and main research directions. *Applied Soil*
- 594 *Ecology*, *104*, 85–90.
- Cameron, E. K., Martins, I. S., Lavelle, P., Mathieu, J., Tedersoo, L. 2, Gottschall, F., ... Eisenhauer,
 N. (2018). Global gaps in soil biodiversity data. *Nature Ecology & Evolution*, *2*, 1042-1043.
- 597 Carini, P., Marsden, P. J., Leff, J. W., Morgan, E., Strickland, M. S., & Fierer, N. (2016). Relic DNA
 598 is abundant in soil and obscures estimates of soil microbial diversity. *Nature Microbiology*, *2*,
 599 16242.
- Caruso, T., Hogg, I. D., Carapelli, A., Frati, F., & Bargagli, R. (2009). Large-scale spatial patterns in
 the distribution of Collembola (Hexapoda) species in Antarctic terrestrial ecosystems. *Journal of Biogeography*, *36*(5), 879–886.
- Caruso, T., Schaefer, I., Monson, F., & Keith, A. M. (2019). Oribatid mites show how climate and
 latitudinal gradients in organic matter can drive large-scale biodiversity patterns of soil
 communities. *Journal of Biogeography*, *46*(3), 611–620.
- 606 Caruso, T., Taormina, M., & Migliorini, M. (2012). Relative role of deterministic and stochastic
 607 determinants of soil animal community : a spatially explicit analysis of oribatid mites. *Journal of*608 *Animal Ecology*, 81(1), 214–221.
- 609 Crowther, T. W., van den Hoogen, J., Wan, J., Mayes, M. A., Keiser, A. D., Mo, L., ... Maynard, D.
- 610 S. (2019). The global soil community and its influence on biogeochemistry. *Science*, *365*(6455),
 611 eaav0550.

- 612 Dauber, J., Purtauf, T., Allspach, A., Frisch, J., Voigtländer, K., & Wolters, V. (2005). Local vs.
- 613 landscape controls on diversity: A test using surface-dwelling soil macroinvertebrates of
 614 differing mobility. *Global Ecology and Biogeography*, *14*(3), 213–221.
- 615 Davison, J., Moora, M., Opik, M., Adholeya, A., Ainsaar, L., Ba, A., ... Zobel, M. (2015). Global
- assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism. *Science*, *349*(6251), 970–973.
- 618Dawson, S. K., Boddy, L., Halbwachs, H., Bässler, C., Andrew, C., Crowther, T. W., ... Jönsson, M.

619 (2019). Handbook for the measurement of macrofungal functional traits: A start with

basidiomycete wood fungi. *Functional Ecology*, *33*(3), 372-387.

- De Deyn, G. B., & Van Der Putten, W. H. (2005). Linking aboveground and belowground diversity.
 Trends in Ecology and Evolution, 20(11), 625–633.
- 623 De Laender, F., Rohr, J. R., Ashauer, R., Baird, D. J., Berger, U., Eisenhauer, N., ... Van den Brink,

624 P. J. (2016). Reintroducing Environmental Change Drivers in Biodiversity–Ecosystem

625 Functioning Research. *Trends in Ecology and Evolution*, *31*(12), 905–915.

- 626 De Vries, F. T., Thebault, E., Liiri, M., Birkhofer, K., Tsiafouli, M. A., Bjornlund, L., ... Bardgett, R.
- 627 D. (2013). Soil food web properties explain ecosystem services across European land use

628 systems. *Proceedings of the National Academy of Sciences*, *110*(35), 14296–14301.

629 De Vries, F. T., Liiri, M. E., Bjørnlund, L., Bowker, M. A., Christensen, S., Setälä, H. M., & Bardgett,

630 R. D. (2012). Land use alters the resistance and resilience of soil food webs to drought. *Nature*

- 631 *Climate Change*, 2(4), 276–280.
- 632 Decaëns, T., Lavelle, P., & Jiménez, J. J. (2008). Priorities for conservation of soil animals. CAB

633 *Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources,*

634 *3*(014), 1–18.

- 635 Decaëns, T. (2010). Macroecological patterns in soil communities. *Global Ecology and*636 *Biogeography*, *19*(3), 287–302.
- 637 Decaëns, T., & Rossi, J. P. (2008). Spatio-temporal structure of earthworm community and soil
 638 heterogeneity in a tropical pasture. *Ecography*, 24(6), 671–682.
- 639 Dechesne, A., Pallud, C., Debouzie, D., Flandrois, J. ., Vogel, T. ., Gaudet, J. ., & Grundmann, G. .
- 640 (2003). A novel method for characterizing the microscale 3D spatial distribution of bacteria in
 641 soil. *Soil Biology and Biochemistry*, *35*(12), 1537–1546.
- 642 Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., ...
- 643 Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey
- animal and plant communities. *Molecular Ecology*, 26(21), 5872-5895.
- 645 Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-González, A., Eldridge, D. J.,
- Bardgett, R. D., ... Fierer, N. (2018). A global atlas of the dominant bacteria found in soil. *Science*, *359*(6373), 320–325.
- 648 Delgado-Baquerizo, M., Powell, J. R., Hamonts, K., Reith, F., Mele, P., Brown, M. V, ... Bissett, A.
- 649 (2017). Circular linkages between soil biodiversity, fertility and plant productivity are limited to
 650 topsoil at the continental scale. *New Phytologist*, *215*(3), 1186–1196.
- 651 Delgado-Baquerizo, M., Reich, P. B., Trivedi, C., Eldridge, D. J., Abades, S., Alfaro, F. D., ... Singh,
- B. K. (2020). Multiple elements of soil biodiversity drive ecosystem functions across biomes.
- 653 *Nature Ecology & Evolution*, *4*(2), 210–220.
- Dirilgen, T., Juceviča, E., Melecis, V., Querner, P., & Bolger, T. (2018). Analysis of spatial patterns
- 655 informs community assembly and sampling requirements for Collembola in forest soils. *Acta*656 *Oecologica*, 86, 23–30.

- Dowle, E.J., Pochon, X., C. Banks, J., Shearer, K. and Wood, S.A., 2016. Targeted gene enrichment
 and high-throughput sequencing for environmental biomonitoring: A case study using
 freshwater macroinvertebrates. Molecular Ecology Resources, *16*(5), 1240-1254.
- Duffy, G., & Chown, S. (2017). Explicitly integrating a third dimension in marine species distribution
 modelling. *Marine Ecology Progress Series*, 564, 1–8.
- 662 Eilers, K. G., Debenport, S., Anderson, S., & Fierer, N. (2012). Digging deeper to find unique
- microbial communities: The strong effect of depth on the structure of bacterial and archaeal
 communities in soil. *Soil Biology and Biochemistry*, *50*, 58–65.
- 665 Eisenhauer, N., Antunes, P. M., Bennett, A. E., Birkhofer, K., Bissett, A., Bowker, M. A., ... Powell,
- J. R. (2017). Priorities for research in soil ecology. *Pedobiologia*, 63, 1–7.
- Eisenhauer, N., Bowker, M. A., Grace, J. B., & Powell, J. R. (2015). From patterns to causal
 understanding: Structural equation modeling (SEM) in soil ecology. *Pedobiologia*, 58(2–3), 65–
 72.
- Eisenhauer, N., Reich, P. B., & Isbell, F. (2012). Decomposer diversity and identity influence plant
 diversity effects on ecosystem functioning. *Ecology*, *93*(10), 2227–2240.
- Elith, J., & Leathwick, J. R. (2009). Species Distribution Models: Ecological explanation and
 prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40,
 674 677–697.
- Ettema, C. H., & Wardle, D. A. (2002). Spatial soil ecology. *Trends in Ecology & Evolution*, 17(4),
 177–183.
- Evans, K. L., Newson, S. E., Storch, D., Greenwood, J. J. D., & Gaston, K. J. (2008). Spatial scale,
 abundance and the species-energy relationship in British birds. *Journal of Animal Ecology*,
 77(2), 395–405.

- 680 Fierer, N., Leff, J. W., Adams, B. J., Nielsen, U. N., Bates, S. T., Lauber, C. L., ... Caporaso, J. G.
- 681 (2012). Cross-biome metagenomic analyses of soil microbial communities and their functional
 682 attributes. *Proceedings of the National Academy of Sciences*, *109*(52), 21390–21395.
- Fierer, N. (2017). Embracing the unknown: disentangling the complexities of the soil microbiome.
 Nature Reviews Microbiology, *15*(10), 579–590.
- Fierer, N., & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences*, *103*(3), 626–631.
- Fierer, N., & Lennon, J. T. (2011). The generation and maintenance of diversity in microbial
 communities. *American Journal of Botany*, *98*(3), 439–448.
- Fierer, N., Schimel, J. P., & Holden, P. A. (2003). Variations in microbial community composition
 through two soil depth profiles. *Soil Biology and Biochemistry*, *35*(1), 167–176.
- Finlay BJ. (2002). Global dispersal of free-living microbial eukaryote species. *Science*, 296(5570),
 1061–1063.
- Franklin, J. (2010). Moving beyond static species distribution models in support of conservation
 biogeography. *Diversity and Distributions*, *16*(3), 321–330.
- 695 Gaston, K. J. (2000). Global patterns in biodiversity. *Nature*, 405(6783), 220–227.
- Geisen, S., Briones, M. J. I., Gan, H., Behan-Pelletier, V. M., Friman, V.-P., de Groot, G. A., ... Wall,
- 697 D. H. (2019). A methodological framework to embrace soil biodiversity. *Soil Biology and*
- 698 *Biochemistry*, *136*, 107536.
- Gerard, B. M. (1967). Factors affecting earthworms in pastures. *Journal of Animal Ecology*, *36*(1),
 235–252.

- Gould, I. J., Quinton, J. N., Weigelt, A., De Deyn, G. B., & Bardgett, R. D. (2016). Plant diversity and
 root traits benefit physical properties key to soil function in grasslands. *Ecology Letters*, *19*(9),
 1140–1149.
- Grace, J. B., Michael Anderson, T., Han, O., & Scheiner, S. M. (2010). On the specification of
 structural equation models for ecological systems. *Ecological Monographs*, 80(1), 67–87.
- Grundmann, G. L., & Debouzie, D. (2000). Geostatistical analysis of the distribution of NH ‡ 4 and
 NO 2 -oxidizing bacteria and serotypes at the millimeter scale along a soil transect. *FEMS Microbiology Ecology*, *34*(3), 57–62.
- 709 Grundmann, G. L., Dechesne, A., Bartoli, F., Flandrois, J. P., Chassé, J. L., & Kizungu, R. (2001).
- 710 Spatial Modeling of Nitrifier Microhabitats in Soil. *Soil Science Society of America Journal*,
 711 65(6), 1709–1716.
- Gumiere, T., Durrer, A., Bohannan, B. J. M., & Andreote, F. D. (2016). Biogeographical patterns in
 fungal communities from soils cultivated with sugarcane. *Journal of Biogeography*, 43(10),
 2016–2026.
- Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., & Martiny, J. B. H. (2012). Beyond
 biogeographic patterns: Processes shaping the microbial landscape. *Nature Reviews Microbiology*, *10*(7), 497–506.
- Hedde, M., Pey, B., Auclerc, A., Capowiez, Y., Cluzeau, D., et al.. BETSI, a complete framework for
 studying soil invertebrate functional traits. *XVI ICSZ International Colloquium on Soil Zoology*, Aug 2012, Coimbra, Portugal. 2012.
- Heemsbergen, D. A., Berg, M. P., Loreau, J. R., van Hal, J. R., Faber, J. H., & Verhoef, H. A. (2004).
 Biodiversity effects on soil processes explained by interspecific functional dissimilarity. *Science*,
- *306*(5698), 1019–1020.

- Hendershot, J. N., Read, Q. D., Henning, J. A., Sanders, N. J., & Classen, A. T. (2017). Consistently
 inconsistent drivers of microbial diversity and abundance at macroecological scales. *Ecology*,
 98(7), 1757–1763.
- Hengl, T., De Jesus, J. M., Heuvelink, G. B. M., Gonzalez, M. R., Kilibarda, M., Blagotić, A., ...
 Kempen, B. (2017). *SoilGrids250m: Global gridded soil information based on machine learning. PLoS ONE, 12*(2), e0169748.
- Hooper, D. U., Chapin, F. S., Ewel, J. J., Hector, A., Inchausti, P., Lavorel, S., ... Wardle, D. A.
- 731 (2005). Effects of biodiversity on ecosystem functioning: A consensus of current knowledge.

Ecological Monographs, 75(1), 3–35.

- 733 Ibáñez, J. J., Krasilnikov, P. V., & Saldaña, A. (2012). Archive and refugia of soil organisms:
- Applying a pedodiversity framework for the conservation of biological and non-biological
 heritages. *Journal of Applied Ecology*, 49(6), 1267–1277.
- ⁷³⁶ İnceoğlu, Ö., Hoogwout, E. F., Hill, P., & Van Elsas, J. D. (2010). Effect of DNA extraction method
 ⁷³⁷ on the apparent microbial diversity of soil. *Applied and Environmental Microbiology*, *76*(10),
 ⁷³⁸ 3378–3382.
- Jackson, H. B., & Fahrig, L. (2012). What size is a biologically relevant landscape? *Landscape Ecology*, 27(7), 929–941.
- 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.
- Joschko, M., Fox, C. A., Lentzsch, P., Kiesel, J., Hierold, W., Krück, S., & Timmer, J. (2006). Spatial
 analysis of earthworm biodiversity at the regional scale. *Agriculture, Ecosystems and Environment*, 112(4), 367–380. https://doi.org/10.1016/j.agee.2005.08.026

- Judas, M. (1988). The species-area relationship of European lumbricidade (annelida, oligochaeta). *Oecologia*, *76*(4), 579–587.
- Kearney, M. R., & Maino, J. L. (2018). Can next-generation soil data products improve soil moisture
 modelling at the continental scale? An assessment using a new microclimate package for the R
 programming environment. *Journal of Hydrology*, *561*, 662–673.
- Kearney, M. R., Shamakhy, A., Tingley, R., Karoly, D. J., Hoffmann, A. A., Briggs, P. R., & Porter,
 W. P. (2014). Microclimate modelling at macro scales: a test of a general microclimate model
 integrated with gridded continental-scale soil and weather data. *Methods in Ecology and*
- *Evolution*, *5*(3), 273–286.
- King, R. A., Tibble, A. L., & Symondson, W. O. C. (2008). Opening a can of worms: Unprecedented
 sympatric cryptic diversity within British lumbricid earthworms. *Molecular Ecology*, *17*(21),
 4684–4698.
- Kissling, W. D., & Schleuning, M. (2015). Multispecies interactions across trophic levels at
 macroscales: Retrospective and future directions. *Ecography*, *38*(4), 346–357.
- Kivlin, S. N., Hawkes, C. V., & Treseder, K. K. (2011). Global diversity and distribution of
 arbuscular mycorrhizal fungi. *Soil Biology and Biochemistry*, *43*(11), 2294–2303.
- Koyama, A., Maherali, H., & Antunes, P. M. (2019). Plant geographic origin and phylogeny as
 potential drivers of community structure in root-inhabiting fungi. *Journal of Ecology*, *107*(4),
 1720–1736.
- 766 Krause, S., Le Roux, X., Niklaus, P. A., Van Bodegom, P. M., Lennon, J. T., Bertilsson, S., ...
- Bodelier, P. L. E. (2014). Trait-based approaches for understanding microbial biodiversity and
 ecosystem functioning. *Frontiers in Microbiology*, 5(MAY), 1–10.
- Kuzyakov, Y., & Blagodatskaya, E. (2015). Microbial hotspots and hot moments in soil: Concept &
 review. *Soil Biology and Biochemistry*, 83, 184–199.

- Ladau, J., & Eloe-Fadrosh, E. A. (2019). Spatial, Temporal, and Phylogenetic Scales of Microbial
 Ecology. *Trends in Microbiology*, 27(8), 662–669.
- Lemanceau, P. (2011). EcoFINDERS: characterizing biodiversity and soil functioning in Europe. 23
 partners from 10 European countries and China. *Biofutur*, 326, 56-58.
- Lembrechts, J., Nijs, I., & Lenoir, J. (2018). Incorporating microclimate into species distribution
 models. *Ecography*, 42(7), 1267-1279.
- Lennon, J. T., Muscarella, M. E., Placella, S. A., & Lehmkuhl, B. K. (2018). How, When, and Where
 Relic DNA Affects Microbial Diversity. *MBio*, *9*(3), e00637-18
- 779 León-Sánchez, L., Nicolás, E., Goberna, M., Prieto, I., Maestre, F. T., & Querejeta, J. I. (2018). Poor
- plant performance under simulated climate change is linked to mycorrhizal responses in a semiarid shrubland. *Journal of Ecology*, *106*(3), 960–976.
- Lindo, Z., & Winchester, N. N. (2006). A comparison of microarthropod assemblages with emphasis
 on oribatid mites in canopy suspended soils and forest floors associated with ancient western
 redcedar trees. *Pedobiologia*, 50(1), 31–41.
- 785 Maclean, I. M. D., Mosedale, J. R., & Bennie, J. J. (2019). Microclima: An R package for modelling
- meso- and microclimate. *Methods in Ecology and Evolution*, *10*(2), 280–290.
- 787 Marchán, D. F., Refoyo, P., Novo, M., Fernández, R., Trigo, D., & Díaz Cosín, D. J. (2015).

Predicting soil micro-variables and the distribution of an endogeic earthworm species through a
model based on large-scale variables. *Soil Biology and Biochemistry*, *81*, 124–127.

Maron, P. A., Ranjard, L., Mougel, C., & Lemanceau, P. (2007). Metaproteomics: A new approach
for studying functional microbial ecology. *Microbial Ecology*, *53*(3), 486–493.

- Martay, B., & Pearce-Higgins, J. W. (2018). Using data from schools to model variation in soil
 invertebrates across the UK: The importance of weather, climate, season and habitat. *Pedobiologia*, 67, 1–9.
- Mathieu, J., Caro, G., & Dupont, L. (2018). Methods for studying earthworm dispersal. *Applied Soil Ecology*, *123*, 339–344.
- Mathieu, J., Grimaldi, M., Jouquet, P., Rouland, C., Lavelle, P., Desjardins, T., & Rossi, J. P. (2009).
 Spatial patterns of grasses influence soil macrofauna biodiversity in Amazonian pastures. *Soil Biology and Biochemistry*, *41*(3), 586–593.
- 800 Maynard, D. S., Bradford, M. A., Covey, K. R., Lindner, D., Glaeser, J., Talbert, D. A., ... Crowther,
- T. W. (2019). Consistent trade-offs in fungal trait expression across broad spatial scales. *Nature Microbiology*, 4(5), 846–853.
- McGill, B. J. (2019). The what, how and why of doing macroecology. *Global Ecology and Biogeography*, 28(1), 6–17.
- Meentemeyer, R. K., Anacker, B. L., Mark, W., & Rizzo, D. M. (2008). Early detection of emerging
 forest disease using dispersal estimation and ecological niche modeling. *Ecological Applications*, 18(2), 377–390.
- Meynard, C. N., Devictor, V., Mouillot, D., Thuiller, W., Jiguet, F., & Mouquet, N. (2011). Beyond
 taxonomic diversity patterns: how do α, β and γ components of bird functional and phylogenetic
 diversity respond to environmental gradients across France? *Global Ecology and Biogeography*,
 20(6), 893–903.
- Montagna, M., Berruti, A., Bianciotto, V., Cremonesi, P., Giannico, R., Gusmeroli, F., ... Gandini, G.
 (2018). Differential biodiversity responses between kingdoms (plants, fungi, bacteria and
 metazoa) along an Alpine succession gradient. *Molecular Ecology*, 27(18), 3671–3685.

- Morriën, E. (2016). Understanding soil food web dynamics, how close do we get? *Soil Biology and Biochemistry*, *102*, 10–13.
- 817 Mysara, M., Vandamme, P., Props, R., Kerckhof, F. M., Leys, N., Boon, N., ... Monsieurs, P. (2017).

818 Reconciliation between operational taxonomic units and species boundaries. *FEMS*

- 819 *Microbiology Ecology*, *93*(4), 1–12.
- Nekola, J. C., & White, P. S. (1999). The distance decay of similarity in biogeography and ecology. *Journal of Biogeography 26*, 867-878.
- 822 Nesme, J., Achouak, W., Agathos, S. N., Bailey, M., Baldrian, P., Brunel, D., ... Simonet, P. (2016).

Back to the Future of Soil Metagenomics. *Frontiers in Microbiology*, 7, 1–5.

- Newbold, T., Hudson, L. N., Hill, S. L. L., Contu, S., Lysenko, I., Senior, R. A., ... Purvis, A. (2015).
 Global effects of land use on local terrestrial biodiversity. *Nature*, *520*(7545), 45–50.
- Newman, E. A., Kennedy, M. C., Falk, D. A., & McKenzie, D. (2019). Scaling and Complexity in
 Landscape Ecology. *Frontiers in Ecology and Evolution*, *7*, 293.
- 828 Nielsen, U. N., Osler, G. H. R., Campbell, C. D., Neilson, R., Burslem, D. F. R. P., & van der Wal, R.
- 829 (2010). The enigma of soil animal species diversity revisited: The role of small-scale
- 830 heterogeneity. *PLoS ONE*, *5*(7), 26–28.
- 831 Nobis, M. P., & Normand, S. (2014). KISSMig a simple model for R to account for limited
- migration in analyses of species distributions. *Ecography*, *37*(12), 1282–1287.
- 833 Noguez, A. M., Arita, H. T., Escalante, A. E., Forney, L. J., García-Oliva, F., & Souza, V. (2005).
- 834 Microbial macroecology: Highly structured prokaryotic soil assemblages in a tropical deciduous
- forest. *Global Ecology and Biogeography*, *14*(3), 241–248.

- 836 O'Brien, S. L., Gibbons, S. M., Owens, S. M., Hampton-Marcell, J., Johnston, E. R., Jastrow, J. D.,
- 837 ... Antonopoulos, D. A. (2016). Spatial scale drives patterns in soil bacterial diversity.
 838 *Environmental Microbiology*, 18(6), 2039–2051.
- 839 Öpik, M., Vanatoa, A., Vanatoa, E., Moora, M., Davison, J., Kalwij, J. M., ... Zobel, M. (2010). The
- 840 online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular
 841 mycorrhizal fungi (Glomeromycota). *New Phytologist*, *188*(1), 223–241.
- Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., & Fernández-Ugalde, O. (2018). LUCAS Soil, the
 largest expandable soil dataset for Europe: a review. *European Journal of Soil Science*, 69(1),
 140–153.
- Orgiazzi, A., Dunbar, M. B., Panagos, P., de Groot, G. A., & Lemanceau, P. (2015). Soil biodiversity
 and DNA barcodes: opportunities and challenges. *Soil Biology and Biochemistry*, 80, 244–250.
- Padarian, J., Minasny, B., & McBratney, A. B. (2020). Machine learning and soil sciences: a review
 aided by machine learning tools. *SOIL*, 6(1), 35–52.
- Palm, J., van Schaik, N. L. M. B., & Schröder, B. (2013). Modelling distribution patterns of anecic,
 epigeic and endogeic earthworms at catchment-scale in agro-ecosystems. *Pedobiologia*, 56(1),
 23–31.
- 852 Pansu, J., De Danieli, S., Puissant, J., Gonzalez, J.-M., Gielly, L., Cordonnier, T., ... Cécillon, L.
- 853 (2015). Landscape-scale distribution patterns of earthworms inferred from soil DNA. *Soil*854 *Biology and Biochemistry*, *83*, 100–105.
- 855 Papp, K., Mau, R. L., Hayer, M., Koch, B. J., Hungate, B. A., & Schwartz, E. (2018). Quantitative
- stable isotope probing with H218O reveals that most bacterial taxa in soil synthesize new
- ribosomal RNA. *The ISME Journal*, 18–20.

- Pärtel, M., Öpik, M., Moora, M., Tedersoo, L., Szava-Kovats, R., Rosendahl, S., ... Zobel, M. (2017). 858 Historical biome distribution and recent human disturbance shape the diversity of arbuscular 859 mycorrhizal fungi. New Phytologist, 227-238. 860
- Pasternak, Z., Al-Ashhab, A., Gatica, J., Gafny, R., Avraham, S., Minz, D., ... Jurkevitch, E. (2013). 861
- 862 Spatial and Temporal Biogeography of Soil Microbial Communities in Arid and Semiarid Regions. PLoS ONE, 8(7), e69705. 863
- Pausch, J., Hünninghaus, M., Kramer, S., Scharroba, A., Scheunemann, N., Butenschoen, O., Marhan, 864
- S., Bonkowski, M., Kandeler, E., Scheu, S., Kuzyakov, Y., & Ruess, L. (2018). Carbon budgets 865 of top- and subsoil food webs in an arable system. Pedobiologia, 69, 29-33. 866
- Pawar, S., Dell, A. I., Lin, T., Wieczynski, D. J., & Savage, V. M. (2019). Interaction dimensionality 867 868 scales up to generate bimodal consumer-resource size-ratio distributions in ecological communities. Frontiers in Ecology and Evolution, 7, 1–11. 869
- 870 Peay, K. G., Bruns, T. D., Kennedy, P. G., Bergemann, S. E., & Garbelotto, M. (2007). A strong
- species-area relationship for eukaryotic soil microbes: Island size matters for ectomycorrhizal 871 fungi. Ecology Letters, 10(6), 470-480. 872
- 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), 874 875 3631-3640.
- 876 Peay, K. G., Schubert, M. G., Nguyen, N. H., & Bruns, T. D. (2012). Measuring ectomycorrhizal fungal dispersal: macroecological patterns driven by microscopic propagules. Molecular 877 Ecology, 21(16), 4122–4136. 878
- 879 Pellissier, L., Niculita-Hirzel, H., Dubuis, A., Pagni, M., Guex, N., Ndiribe, C., ... Guisan, A. (2014). 880 Soil fungal communities of grasslands are environmentally structured at a regional scale in the Alps. *Molecular Ecology*, 23(17), 4274–4290. 881

- Philippot, L., Ritz, K., Pandard, P., Hallin, S., & Martin-Laurent, F. (2012). Standardisation of
 methods in soil microbiology: Progress and challenges. *FEMS Microbiology Ecology*, 82(1), 1–
 10.
- Phillips, H. R. P., Guerra, C. A., Bartz, M. L. C., Briones, M. J. I., Brown, G., Crowther, T. W., ...
 Eisenhauer, N. (2019). Global distribution of earthworm diversity. *Science*, *366*(6464), 480–
 485.
- Phillips, H. R. P., Cameron, E. K., Ferlian, O., Türke, M., Winter, M., & Eisenhauer, N. (2017). Red
 list of a black box. *Nature Ecology & Evolution*, 1(4), 0103.
- 890 Pickles, B. J., Genney, D. R., Potts, J. M., Lennon, J. J., Anderson, I. C., & Alexander, I. J. (2010).
- 891 Spatial and temporal ecology of Scots pine ectomycorrhizas. *New Phytologist*, *186*(3), 755–768.
- 892 Poage, M. A., Barrett, J. E., Virginia, R. A., & Wall, D. H. (2014). The Influence of Soil
- 893 Geochemistry on Nematode Distribution, McMurdo Dry Valleys, on Nematode Distri. *Arctic*,
 894 *Antarctic and Alpine Research*, 7(4), 331–340.
- Pollock, L. J., Tingley, R., Morris, W. K., Golding, N., O'Hara, R. B., Parris, K. M., ... Mccarthy, M.
- A. (2014). Understanding co-occurrence by modelling species simultaneously with a Joint
 Species Distribution Model (JSDM). *Methods in Ecology and Evolution*, 5(5), 397–406.
- Powell, J. R., & Rillig, M. C. (2018). Biodiversity of arbuscular mycorrhizal fungi and ecosystem
 function. *New Phytologist*, 220(4), 1059-1075.
- Prosser, J. I. (2015). Dispersing misconceptions and identifying opportunities for the use of "omics"
 in soil microbial ecology. *Nature Reviews Microbiology*, *13*, 439-446.
- 902 Ramirez, K. S., Knight, C. G., de Hollander, M., Brearley, F. Q., Constantinides, B., Cotton, A., ... de
- 903 Vries, F. T. (2018). Detecting macroecological patterns in bacterial communities across
- 904 independent studies of global soils. *Nature Microbiology*, *3*(2), 189–196.

905	Ramirez, K. S., Döring, M., Eisenhauer, N., Gardi, C., Ladau, J., Leff, J. W., Morgan, K. (2015).
906	Toward a global platform for linking soil biodiversity data. Frontiers in Ecology and Evolution,
907	3(91), 1–7.

908 Ramirez, K. S., Leff, J. W., Barberán, A., Bates, S. T., Betley, J., Thomas, W., ... Fierer, N. (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 B*, 281(1795), 20141988..

911 Ranjard, L., Dequiedt, S., Jolivet, C., Saby, N. P. A., Thioulouse, J., Harmand, J., ... Lemanceau, P.

912 (2009). Biogeography of soil microbial communities: A review and a description of the ongoing

913 French national initiative. *Sustainable Agriculture*, *2*, 857–865.

914 Rasmussen, P. U., Hugerth, L. W., Blanchet, F. G., Andersson, A. F., Lindahl, B. D., & Tack, A. J. M.

915 (2018). Multiscale patterns and drivers of arbuscular mycorrhizal fungal communities in the

916 roots and root-associated soil of a wild perennial herb. *New Phytologist*, 220(4), 1248–1261.

917 Raymond, M. R., Wharton, D. A., & Marshall, C. J. (2013). Factors determining nematode

- 918 distributions at Cape Hallett and Gondwana station, Antarctica. *Antarctic Science*, 25(3), 347–
 919 357.
- Raynaud, X., & Nunan, N. (2014). Spatial Ecology of Bacteria at the Microscale in Soil. *PLoS ONE*,
 921 9(1), e87217.
- 922 Rodgers, D. J., & Kitching, R. L. (1998). Vertical stratification of rainforest collembolan

923 (Collembola: Insecta) assemblages: Description of ecological patterns and hypotheses

- 924 concerning their generation. *Ecography*, 21(4), 392–400.
- 925 Rosenzweig, M. L. (1995). Species diversity in space and time. Cambridge University Press.
- 926 Rumpel, C., & Kögel-Knabner, I. (2011). Deep soil organic matter-a key but poorly understood
- 927 component of terrestrial C cycle. *Plant and Soil*, *338*(1), 143–158.

- 928 Ruppert, K. M., Kline, R. J., & Rahman, M. S. (2019). Past, present, and future perspectives of
- environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and
 applications of global eDNA. *Global Ecology and Conservation*, 17, e00547.
- 931 Rutgers, M., Orgiazzi, A., Gardi, C., Römbke, J., Jänsch, S., Keith, A. M., ... Zwart, D. de. (2016).
- 932 Mapping earthworm communities in Europe. *Applied Soil Ecology*, 97, 98–111.
- 933 Rutgers, M., van Leeuwen, J. P., Vrebos, D., van Wijnen, H. J., Schouten, T., & de Goede, R. G. M.
- 934 (2019). Mapping Soil Biodiversity in Europe and the Netherlands. *Soil Systems*, *3*(2), 39.
- 935 Sato, H., Tsujino, R., Kurita, K., Yokoyama, K., & Agata, K. (2012). Modelling the global
- 936 distribution of fungal species: New insights into microbial cosmopolitanism. *Molecular Ecology*,
 937 21(22), 5599–5612.
- 938 Sayer, E. J., Wagner, M., Oliver, A. E., Pywell, R. F., James, P., Whiteley, A. S., & Heard, M. S.
- 939 (2013). Grassland management influences spatial patterns of soil microbial communities. *Soil*940 *Biology and Biochemistry*, *61*, 61–68.
- 941 Schaefer, I., & Caruso, T. (2019). Oribatid mites show that soil food web complexity and close
- aboveground-belowground linkages emerged in the early Paleozoic. *Communications Biology*,
 2(1), 387.
- Schaefer, I., Norton, R. A., Scheu, S., & Maraun, M. (2010). Arthropod colonization of land Linking
 molecules and fossils in oribatid mites (Acari, Oribatida). *Molecular Phylogenetics and Evolution*, 57(1), 113–121.
- 947 Schröder, B. (2008) Challenges of species distribution modeling belowground. *Journal of Plant*948 *Nutrition and Soil Science*, 171, 325–337.
- Shade, A., Dunn, R. R., Blowes, S. A., Keil, P., Bohannan, B. J. M., Herrmann, M., ... Chase, J.
 (2018). Macroecology to Unite All Life, Large and Small. *Trends in Ecology & Evolution*, *33*(10), 731–744.

- Shoemaker, W. R., Locey, K. J., & Lennon, J. T. (2017). A macroecological theory of microbial
 biodiversity. *Nature Ecology & Evolution*, 1(5), 0107.
- 954 Shokralla, S., Gibson, J. F., King, I., Baird, D. J., Janzen, D. H., Hallwachs, W., & Hajibabaei, M.
- 955 (2016). Environmental DNA barcode sequence capture: targeted, PCR-free sequence capture for
 956 biodiversity analysis from bulk environmental samples. *BioRxiv*, (087437 [Preprint]), 28.
- 957 Siciliano, S. D., Palmer, A. S., Winsley, T., Lamb, E., Bissett, A., Brown, M. V., ... Snape, I. (2014).
- 958 Soil fertility is associated with fungal and bacterial richness, whereas pH is associated with
- 959 community composition in polar soil microbial communities. *Soil Biology and Biochemistry*, 78,
 960 10–20.
- 961 Smith, A. B., Sandel, B., Kraft, N. J. B., & Carey, S. (2013). Characterizing scale-dependent
- 962 community assembly using the functional-diversity-area-relationship. *Ecology*, 94(11), 2392–
 963 2402.
- Soininen, J. (2012). Macroecology of unicellular organisms patterns and processes. *Environmental Microbiology Reports*, 4(1), 10–22.
- Soininen, J., McDonald, R., & Hillebrand, H. (2007). The distance decay of similarity in ecological
 communities. *Ecography*, 30(1), 3–12.
- 968 Solomou, A. D., Sfougaris, A. I., Vavoulidou, E. M., & Csuzdi, C. (2013). Species richness and
- 969 density of earthworms in relation to soil factors in olive orchard production systems in Central
 970 Greece. *Communications in Soil Science and Plant Analysis*, 44(1–4), 301–311.
- 971 Sosa-Hernández, M. A., Roy, J., Hempel, S., & Rillig, M. C. (2018). Evidence for subsoil
- 972 specialization in arbuscular mycorrhizal fungi. *Frontiers in Ecology and Evolution*, 6, 1–7.
- 973 Soudzilovskaia, N. A., Vaessen, S., Barcelo, M., He, J., Rahimlou, S., Abarenkov, K., ... Tedersoo, L.
- 974 (*in press*). FungalRoot: Global online database of plant mycorrhizal associations. New
- 975 *Phytologist*.

- 976 Staniczenko, P. P. A., Sivasubramaniam, P., Suttle, K. B., & Pearson, R. G. (2017). Linking
- 977 macroecology and community ecology: refining predictions of species distributions using biotic
 978 interaction networks. *Ecology Letters*, 20(6), 693–707.
- Stehli, F. G., Douglas, R. G., & Newell, N. D. (1969). Generation and Maintenance of Gradients in
 Taxonomic Diversity. *Science*, *164*(3882), 947–949.
- Swenson, N. G., & Jones, F. A. (2017). Community transcriptomics, genomics and the problem of
 species co-occurrence. *Journal of Ecology*, *105*(3), 563–568.
- 983 Tedersoo, L., Anslan, S., Bahram, M., Põlme, S., Riit, T., Liiv, I., ... Abarenkov, K. (2015). Shotgun
- 984 metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in
 985 metabarcoding analyses of fungi. *MycoKeys*, *10*, 1–43.
- Tedersoo, L., Bahram, M., Polme, S., Koljalg, U., Yorou, N. S., Wijesundera, R., ... Abarenkov, K.
 (2014). Global diversity and geography of soil fungi. *Science*, *346*(6213), 1256688–1256688.
- 988 Thakur, M. P., Phillips, H. R. P., Brose, U., De Vries, F. T., Lavelle, P., Loreau, M., ... Cameron, E.
- 989 K. (2019). Towards an integrative understanding of soil biodiversity. *Biological Reviews*, 95(2),
 990 350-364.
- 991 Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., ... Zhao, H.
- 992 (2017). A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*, 551, 457–
 993 463.
- 994 Tremblay, J., Singh, K., Fern, A., Kirton, E. S., He, S., Woyke, T., ... Tringe, S. G. (2015). Primer
 995 and platform effects on 16S rRNA tag sequencing. *Frontiers in Microbiology*, *6*, 1–15.
- 996 Tukiainen, H., Alahuhta, J., Field, R., Ala-Hulkko, T., Lampinen, R., & Hjort, J. (2017). Spatial
 997 relationship between biodiversity and geodiversity across a gradient of land-use intensity in
 998 high-latitude landscapes. *Landscape Ecology*, *32*(5), 1049–1063.

- Vályi, K., Mardhiah, U., Rillig, M. C., & Hempel, S. (2016). Community assembly and coexistence in
 communities of arbuscular mycorrhizal fungi. *ISME Journal*, *10*(10), 2341–2351.
- Violle, C., Reich, P. B., Pacala, S. W., Enquist, B. J., & Kattge, J. (2014). The emergence and promise
 of functional biogeography. *Proceedings of the National Academy of Sciences*, *111*(38), 13690–
 13696.
- Wagg, C., Bender, S. F., Widmer, F., & van der Heijden, M. G. A. (2014). Soil biodiversity and soil
 community composition determine ecosystem multifunctionality. *Proceedings of the National Academy of Sciences*, *111*(14), 5266–5270.
- 1007 Ward, S. E., Smart, S. M., Quirk, H., Tallowin, J. R. B., Mortimer, S. R., Shiel, R. S., Wilby, A., &
- Bardgett, R. D. (2016). Legacy effects of grassland management on soil carbon to depth. *Global Change Biology*, 22(8), 2929–2938.
- Wardle, D. A., Bardgett, R. D., Klironomos, J. N., Setala, H., van der Putten, W. H., & Wall, D. H.
 (2004). Ecological Linkages Between Aboveground and Belowground Biota. *Science*, *304*(5677), 1629–1633.
- 1013 Warton, D. I., Blanchet, F. G., O'Hara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C., & Hui, F.
- 1014 K. C. (2015). So Many Variables: Joint Modeling in Community Ecology. *Trends in Ecology*1015 *and Evolution*, 30(12), 766–779.
- White, E., Adler, P., & Lauenroth, W. (2006). A comparison of the species-time relationship across
 ecosystems and taxonomic groups. *Oikos*, *112*(1), 185-195.
- 1018 White, H. J., Montgomery, W. I., Storchová, L., Hořák, D., & Lennon, J. J. (2018a). Does functional
- 1019 homogenization accompany taxonomic homogenization of British birds and how do biotic
- factors and climate affect these processes? *Ecology and Evolution*, 8(15), 7365–7377.

1021	White, H. J., Montgomery, W. I., Pakeman, R. J., & Lennon, J. J. (2018b). Spatiotemporal sca					
1022	plant species richness and functional diversity in a temperate semi-natural grassland. Ecography,					
1023	41(5), 845–856.					

Wong, M. K. L., & Guénard, B. (2017). Subterranean ants: Summary and perspectives on field
sampling methods, with notes on diversity and ecology (Hymenoptera: Formicidae).

1026 *Myrmecological News*, 25, 1–16.

- 1027 Wu, J. (2004). Effects of changing scale on landscape pattern analysis: Scaling relations. *Landscape* 1028 *Ecology*, 19(2), 125–138.
- 1029 Zanne, A. E., Abarenkov, K., Afkhami, M. E., Aguilar-Trigueros, C. A., Bates, S., Bhatnagar, J. M.,

1030 ... Treseder, K. K. (2020). Fungal functional ecology: bringing a trait-based approach to plant1031 associated fungi. *Biological Reviews*, 95(2), 409–433.

Zanne, A. E., Powell, J. R., Flores-Moreno, H., Kiers, E. T., van 't Padje, A., & Cornwell, W. K. (*in press*). Finding fungal ecological strategies: Is recycling an option? *Fungal Ecology*.

1034 Zhang, W., Hendrix, P. F., Snyder, B. A., Molina, M., Li, J., Rao, X., ... Fu, S. (2010). Dietary
1035 flexibility aids Asian earthworm invasion in North American forests. *Ecology*, *91*(7), 2070–

1036

2079.

1037

1038

1039

1040

1041

1042 1043

1044

Table 1 Examples of existing databases that include soil biodiversity or distribution data.

Database Website	Database Name	Таха	Geographic coverage	Taxonomic resolution	Year compiled
https://data.bioplatf orms.com/organiza tion/about/bpa-base	Biome of Australian Soil Environments	Microbes	Australia and Antarctica	Molecular (OTU)	2016
http://drilobase.org/	Drilobase - the World Earthworm Database	Earthworms	Global	Species	2014- present
http://www.earthmi crobiome.org/	Earth Microbiome Project	Microbes	Global	Molecular (OTU)	2017
https://edaphobase. org/	Edaphobase	Soil fauna	Germany (predominantly)	Species	2014- present
<u>http://maarjam.bota</u> <u>ny.ut.ee/</u>	MaarjAM	Arbuscular mycorrhizal fungi	Global	Molecular (small sub-unit rRNA sequences)	2010
https://catalogue.ce h.ac.uk/documents/ 53210c27-87fc- 46e4-a3d6- e731003dc541	Model estimates of topsoil microbes [Countryside Survey]	Bacteria	Great Britain	Molecular (terminal restriction fragments)	2007
https://nbnatlas.org	National Biodiversity Network Atlas	Various, including soil biota	Great Britain and Northern Ireland	Variable	2000- present
http://data.nhm.ac. uk/dataset/the- 2016-release-of- the-predicts- database	PREDICTS	Various, including soil biota	Global	Predominantly species	2016
https://catalogue.ce h.ac.uk/documents/ fccd86b0-f5b6- 4716-b4f7- f43ad82daeee	Soil Invertebrates Countryside Survey	Soil invertebrates	Great Britain	Class/Order	2007

1052 Figure Legends

1053 Figure 1. Considering different soil habitats and corresponding sampling methods adds1054 dimensionality to macroecological approaches.

1055

1056	Figure 2.	The challenges associate	d with the below	wground system	which contrib	uted to a	lack of so	oil
				<u> </u>				

1057 macroecological research, and potential approaches to address them. These challenges can be

separated into four sets of necessary considerations: data; spatio-temporal; biotic; and abiotic.

1059 Challenges for each solution are highlighted in central darker boxes whilst example approaches to

address these challenges are shown in the surrounding, lighter boxes. J-SDMs = joint species

- 1061 distribution modeles; AMF = arbuscular mycorrhizal fungi; PCNM = Principal Coordinates of
- 1062 Neighbourhood Matrices.