



**QUEEN'S  
UNIVERSITY  
BELFAST**

## **Oribatid mites show how climate and latitudinal gradients in organic matter can drive large-scale biodiversity patterns of soil communities**

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. <https://doi.org/10.1111/jbi.13501>

**Published in:**  
Journal of 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**

© 2018 John Wiley & Sons Ltd.

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](mailto: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 **Article type: Research Article**

2

3 **Oribatid mites show how climate and latitudinal gradients in organic matter can drive**  
4 **large-scale biodiversity patterns of soil communities**

5

6 **Short running title:** *Macroecological determinants of soil animals*

7

8 **Authors:** Tancredi Caruso<sup>a</sup>, Ina Schaefer<sup>b</sup>, Frank Monson<sup>c</sup>, Aidan M Keith<sup>d</sup>

9

10 <sup>a</sup> School of Biological Sciences and Institute for Global Food Security, Queen's University  
11 Belfast. Medical Biology Centre, 97 Lisburn Road, Belfast BT9 7BL, Northern Ireland, UK

12

13 <sup>b</sup>J.F.Blumenbach Institute of Zoology and Anthropology University of Göttingen, Untere  
14 Karspüle 2, 37073 Göttingen, Germany

15

16 <sup>c</sup> Associate, World Museum Liverpool, William Brown Street, Liverpool, L3 8EN

17

18 <sup>d</sup>Centre for Ecology & Hydrology, Lancaster Environment Centre, Library Avenue, Bailrigg,  
19 Lancaster LA1 4AP, UK

20 **Corresponding Author:** Tancredi Caruso, *t.caruso@qub.ac.uk*, + 44 (0) 28 90972271

21

## 22 **Acknowledgments**

23 TC and IS were supported by the project SENSE (Structure and Ecological Niche in the Soil  
24 Environment; EC FP7 - 631399 - SENSE). We thank AG Scheu (University Göttingen) for  
25 kindly providing sequences of Brachychthoniidae specimens. We acknowledge the UKCP09  
26 data made freely available by the Met Office and supported by the Department for  
27 Environment, Food and Rural Affairs (Defra). We also wish to thank all those responsible for  
28 coordinating and conducting the 1998 Countryside Survey and subsequent processing of  
29 samples. We thank three anonymous reviewers for their constructive comments.

30

31 **Abstract**

32 Aim: The factors determining spatial distributions and diversity of terrestrial invertebrates are  
33 typically investigated at small scales. Large-scale studies are ~~particularly missing~~ lacking for  
34 soil animals, which control microbial communities and represent one of the most diverse yet  
35 poorly known animal assemblages. Here, we analyzed a major group (Oribatida) to test  
36 whether belowground macroecological patterns can be predicted by climatic variables,  
37 vegetation, and large-scale variation in key soil properties.

38 Location: ~~we~~We modelled the multivariate distribution of more than 100 species using  
39 biodiversity data collected across Great Britain in the framework of the Countryside Survey  
40 (<http://www.countryside.gov.uk>).

41 Methods: We analysed species-level data from 582 samples collected across 162 hectads  
42 (10 × 10 km) covering the largest possible range of vegetation types, soil properties and  
43 climatic conditions within GB. We created the first large-scale maps of soil animal diversity  
44 metrics at the GB scale, including novel estimates of metrics of phylogenetic diversity. Using  
45 structural equation modelling, we quantified the direct and indirect effects of location  
46 (latitude, longitude), plant community structure, and abiotic factors such as precipitation on  
47 species composition, richness, and phylogenetic diversity.

48 Results: We found that variation in species composition follows a latitudinal gradient with  
49 diversity generally increasing northward. The latitudinal variation in species composition  
50 drives phylogenetic diversity via changes in both species richness and phylogenetic distance  
51 between species. This gradient is mostly determined by latitudinal variation in precipitation  
52 and organic matter, which were very good predictor of species composition. Precipitation  
53 and organic matter were, however, relatively weak while statistically significant predictors of  
54 diversity metrics.

55 Conclusions: Past studies have emphasized the unpredictability of species distributions and  
56 variation in species composition in hyper diverse soil animal communities. However, past  
57 studies were conducted at small scales, where stochastic factors may weaken the signal of  
58 deterministic factors. Oribatid mites in our study show for the first time that large scale  
59 latitudinal gradients in climate and organic matter predict not only variation in species  
60 composition but also taxonomic and phylogenetic diversity of soil animal communities.

61 **Keywords:** soil macroecology; animals; diversity; distribution; community phylogenetics;  
62 Oribatida

63 **1. Introduction**

64 In the last two decades, one of the major goals of ecology has been to understand the  
65 relative roles of the many factors that structure ecological communities in space and time but  
66 the majority of studies have focused on aboveground communities, particularly plant  
67 communities (Chesson, 2000; Clark & McLachlan, 2003; Hubbell, 2005; Levine &  
68 HilleRisLambers, 2009). More recently, ecologists have started to better investigate the  
69 interactions between aboveground and belowground communities and how these  
70 interactions drive the composition and diversity of both communities (Kardol et al., 2006; Van  
71 Der Heijden et al., 2008; de Vries et al., 2012; Prober et al., 2015). Traditionally, most  
72 studies investigating aboveground-belowground linkages have been conducted at relatively  
73 local spatial scales but some regional and global scale analyses of soil microbial  
74 communities have shed light on the large scale determinants of these communities (Fierer  
75 et al., 2009; de Vries et al., 2012; Ramirez et al., 2014). Local and fine scale variation seems  
76 mostly due to interactions determined by the patchy distribution of resources and plant  
77 species (Bezemer et al., 2010; Thomson et al., 2010) while spatial gradients in vegetation  
78 types and abiotic factors such as pH and climatic conditions are the major correlates of  
79 microbial distribution at regional scales (Fierer et al., 2009; Griffiths et al., 2011). Protists,  
80 too, follow similar macroecological patterns (Soininen, 2011). For some groups such as  
81 arbuscular mycorrhizal fungi (AMF), global studies have started to reveal the relative roles of  
82 local, regional and historical factors on community structure and diversity (Davison et al.,  
83 2015) but for soil animals large-scale studies are missing despite some synthesis data  
84 having provided insight in the macroecology of soil arthropods, nematodes and oligochaetes  
85 (Decaëns, 2010; Brusaard et al., 2012; Nielsen, 2014). Studies focusing on selected  
86 assemblages at relatively local scales (Lindo & Winchester, 2009; Nielsen et al., 2010;  
87 Caruso et al., 2011) have shown that, similarly to microbial communities, soil animal  
88 communities are structured at multiple spatial scales, with many species being dispersal-  
89 limited over certain scales (Ettema & Wardle, 2002; Wardle, 2006) and soil environmental

90 heterogeneity being high already at very small scale (e.g. <100cm; Ettema & Wardle 2002).  
91 This small scale heterogeneity promotes community diversity in both animals and microbes  
92 and is mostly due to small scale variation in soil properties such as pH, the concentration of  
93 organic matter and key nutrient such as P and N, and also structural variation in soil such as  
94 variation in the physical distribution of aggregates (Dumbrell et al., 2010; Nielsen et al.,  
95 2010). Nevertheless, much spatial variation in the structure and diversity of soil communities,  
96 especially animals, is often left unexplained by variation in environmental variables or other  
97 biotic factors. Previous studies have hypothesised that stochastic population dynamics,  
98 including dispersal limitation, may sometimes play a major role in this variation (Lindo &  
99 Winchester, 2009; Caruso et al., 2011). At large scales, the few studies available on soil  
100 animals have concentrated on classical macroecological patterns such as species-area  
101 relationships, altitudinal and latitudinal gradients in diversity and some insight on patterns of  
102 phylogenetic diversity (see review in Brusaard et al., 2012) but most datasets focused on the  
103 highest taxonomic ranks (e.g., family or class level; Nielsen, 2014) or, as noted by Decaëns,  
104 (2010) are biased in terms of sampling efforts towards temperate countries, and in general  
105 lack the resolution necessary to disentangle the contribution of multiple factors at multiple  
106 spatial scales. The only quantitative study on the regional variation of soil animal  
107 communities at the species level (Zaitsev *et al.* 2013) showed that studies conducted at  
108 relatively small scales cannot capture the long-term effects of the historical processes that  
109 contribute to large scale gradients in species richness and community composition. Overall,  
110 large-scale studies that help disentangle the roles of abiotic and biotic factors that structure  
111 soil communities at regional scales are in their infancy.

112 Here we focused on a unique dataset of the species of oribatid mites collected during the  
113 first assessment of soil biodiversity across Great Britain undertaken in 1998 (known as the  
114 GB Countryside Survey or CS: <http://www.countrysidesurvey.org.uk>). This survey produced  
115 a baseline dataset across all major soil types and habitats (Black et al., 2002) and showed  
116 that populations of microbes and microarthropods varied across major environmental zones,

117 vegetation classes and soil types (Black et al., 2003; Griffiths et al., 2011; Keith et al., 2015).  
118 Oribatid mites (Acari, Acariformes) are a cornerstone of soil food webs worldwide: over  
119 10,000 species have been described and they can reach densities of up to 400,000 ind./m<sup>2</sup>  
120 in forest soils, although they occur in all biomes including continental Antarctica (Coleman et  
121 al., 2004). Oribatid mites are one of the most ancient groups of terrestrial animals and have  
122 been part of the soil food webs ever since soil have appeared on the geological record about  
123 400 mya (Shear et al., 1984; Norton et al., 1988). They appear for the first time in the fossil  
124 record of the Devonian site of Rhynie Chert (407-397 mya, Aberdeenshire, Scotland)  
125 although a relatively recent molecular clock suggests a much earlier origin (Schaefer et al.,  
126 2010). For all these reasons, oribatid mites provide an excellent model to analyze the role of  
127 abiotic and biotic factors in structuring diversity and composition of belowground animals at  
128 regional scales. We used the dataset of oribatid mites to conduct a species-level analysis of  
129 the determinants of community structure and diversity of this major group of soil animals.  
130 The CS is a unique audit of vegetation, soils, habitats and landscape across GB that began  
131 in 1978 (Firbank et al., 2003; Keith *et al.* 2015). Using a spatially explicit approach, we  
132 created the first maps of diversity metrics at the GB scale for a major soil animal group,  
133 including the first estimates of phylogenetic diversity (Faith, 1992; Cadotte et al., 2010) and  
134 used structural equation modelling (Grace, 2006) to estimate the direct and indirect effects of  
135 location (latitude, longitude), abiotic factors such as precipitation, and plant community  
136 structure on oribatid mite species composition and diversity. We hypothesized that large-  
137 scale gradients in this belowground community are directly driven by large-scale gradients in  
138 abiotic factors (e.g., climatic variables) but also via the effects of these factors on plant  
139 community structure and edaphic properties such as organic matter.

140

141

142

143 **Methods**

144 ~~Database~~ Background background

145 The data analysed in this study were collected in the framework of the Country Survey audit  
146 (Firbank et al., 2003; Keith *et al.* 2015). This environmental audit is based on a stratification  
147 of GB into land classes, each land class being characterised by a combination of climate,  
148 altitude and location (Firbank et al., 2003; Sheail & Bunce, 2003). Each sampling location  
149 was assigned ~~a Broad Habitat (BH) and~~ an Aggregate Vegetation Class (AVC). ~~BH is a~~  
150 ~~classification consisting of 27 habitats that are used in the Land Cover Map accounting for~~  
151 ~~the entire land surface of GB, and;~~ AVC is a high-level grouping of vegetation types  
152 produced from a classification of plant communities from the original CS vegetation plots and  
153 includes eight categories that group common habitats (crops and weeds, tall grass and herb,  
154 fertile grassland, infertile grassland, lowland woodland, upland woodland, moorland-grass  
155 mosaic, and, heath and bog (Bunce et al., 1999). Specifically, we analysed 582 samples  
156 mostly collected between 29 May and 28 October in 1998 with some samples collected  
157 between June and August in 1999. The samples analysed in this study were collected  
158 across 162 10 x 10 km plots (hectads), with an average of 4 locations sampled within each  
159 hectad. Each sample location was associated with information on vegetation, soil properties  
160 and land-use, produced during CS. For the collection of each soil sample, surface vegetation  
161 was removed leaving the litter layer intact and a soil core (4 cm diameter, 8 cm depth) was  
162 taken. Cores were placed immediately in cool boxes and sent to the laboratory at the Centre  
163 for Ecology & Hydrology Lancaster for extraction of invertebrates.

164

165 *2.2 Oribatid extraction and identification*

166 Cores were processed over five days using a dry Tullgren extraction method and all  
167 invertebrate specimens collected into 70% ethanol preservative (Emmett et al., 2010) . Once  
168 collected, the soil invertebrates were identified to broad taxa, separated and counted under a

169 stereomicroscope. Specimens of Acari (mites) from each extract were removed into another  
170 vial and sent for separation and identification of oribatid mites species. Specimens were  
171 identified at x400 magnification and, where necessary, were cleared for 24h using lactic acid  
172 at room temperature before being mounted in glass cavity slides. The unpublished  
173 monograph of British oribatid mites by M. Luxton and other specialist primary literature were  
174 used to identify oribatid mites to the species level; identifications have since been checked  
175 against Weigmann, (2006). Weigmann (2006) plus several other specialist publications were  
176 also used to define geographic distributions and ecological traits of the taxa. Plot codes,  
177 ~~Oribatid-oribatid~~ species records and taxonomic details were then collated into a dataset ~~that~~  
178 ~~is available upon request from the NERC Environmental Information Centre (see Keith et al.~~  
179 ~~2018 for details to access the data).~~

Formatted: Font: Not Bold

### 181 2.3 Associated environmental data

182 Existing ~~data on soil (Black et al. 2016), and vegetation (Barr et al. 2014) and habitat data~~  
183 ~~from CS were collated for the 582 plots with oribatid mite records samples of soil fauna and~~  
184 ~~are available upon request from the NERC Environmental Information Centre (see Barr et al.~~  
185 ~~2014 and Black et al. 2016 for details to access the vegetation and habitat data).~~ Soil  
186 properties were collected from a separate core taken adjacent to the core used to extract soil  
187 animals; soil data included moisture content, pH, organic matter (loss-on-ignition), total C  
188 content and total N content. The sampling protocol and detailed methods used for these soil  
189 analyses can be found in Emmett et al. (2008) and the data are reported in more detail  
190 elsewhere (Reynolds et al., 2013). For vegetation composition, ordination scores were used  
191 from the first three axes of a Detrended Correspondence Analysis (DCA) using binary plant  
192 species data from the same plots.

193 Climate data associated with each sampling plot was derived from the UKCP09: Met Office  
194 gridded land surface climate observations at 5 × 5km resolution (Met Office, 2017). These

195 data were used to calculate average values of mean annual temperature and mean annual  
196 rainfall for the period 1992–1997, in order to incorporate recent climatic trends.

197

## 198 2.4 Statistical analyses

### 199 *Community and environmental data*

200 Records of oribatid mites across sampling locations were collated at 10km × 10km/hectad  
201 resolution for a total of 162 10 x 10 km squares, and the associated environmental data were  
202 averaged at this resolution (~~see Barr et al. 2014, Black et al. 2016 and Keith et al. 2018 to~~  
203 ~~access the data from NERC Environmental Information Centre).~~

204 We used the spatial interpolation method of kriging (Matheron, 1963; Wagner, 2003) to  
205 illustrate spatial variation in community structure and metric of diversity (see below for the  
206 metrics used). The spatial structure of the variables was quantified with the empirical  
207 semivariogram (Wagner, 2003; Bivand et al., 2008) and then fitted with a theoretical  
208 variogram model (i.e., exponential or Gaussian, or spherical models) to estimate values at  
209 unmeasured locations. We used the R library "~~geoR", "maps", "mapdata" and "gstat"~~  
210 for [calculating](#) variograms, kriging estimation and mapping of results (~~See also~~ Bivand et al.  
211 ~~(2008)~~ [for further details](#).

212 We used a multivariate regression approach based on Principal Coordinate Analysis  
213 (Legendre & Legendre, 1998; Borcard et al., 2004) to quantify the relative importance of  
214 location (space) and environment (temperature, plant community composition, pH,  
215 precipitation, organic matter) on oribatid mite community structure. PCoA was applied to the  
216 Jaccard distance matrix obtained by the presence-absence distribution of species, and a  
217 distance based RDA (db-RDA) was used to estimate the effect of space and environment on  
218 the multivariate distribution of species.

**Commented [KA1]:** Do these all have separate citations or do they come as one?

219 To quantify the effect of “space” (i.e., location), we used latitude, longitude and the method  
220 of principal coordinate analysis of neighbour matrices (PCNM; Borcard & Legendre, 2002),  
221 which defines a set of spatial factors that parsimoniously account for patterns in species  
222 distribution at multiple scales. The final set of PCNM vectors was defined using a  
223 multivariate extension of the Akaike information criterion (AIC; Dray et al., 2006). Variance  
224 partitioning was computed to estimate the amount of fraction uniquely attributable to space  
225 and environment, and the variation shared between space and environment (Borcard et al.,  
226 1992; Legendre & Legendre, 1998). Besides observed species number per hectad we also  
227 calculated species rarefaction curve (Gotelli & Colwell, 2001) for each hectad and estimated  
228 the hectad asymptotic richness using the Chao estimator (O’Hara 2005; Chiu *et al.* 2014). All  
229 multivariate analyses and estimates of species richness were performed using the R  
230 package ‘vegan’ ~~“vegan”~~ (Oksanen et al., 2007).

231

#### 232 Phylogenetic methods

233 The phylogenetic tree was reconstructed based on 18S rDNA. Sequences were downloaded  
234 from NCBI (<https://www.ncbi.nlm.nih.gov>) or, if not available, were newly generated  
235 sequenced at the J.F. Blumenbach Institute of Zoology and Anthropology, University of  
236 Göttingen.

237 Genomic DNA was extracted from single individuals using the DNeasy® Blood and Tissue  
238 Kit (Qiagen, Manchester, UK) following the manufacturer’s protocol for animal tissue.  
239 Amplification of the 18S region was performed in 25 µl volumes containing 12.5 µl  
240 HotStarTaq Mastermix (Qiagen), 5 µl of template DNA, 1 µl of each primer (100 pM) and 5.5  
241 µl H<sub>2</sub>O. Primers for PCR were 5’ -TAC- CTGGTTGATCCTGCCAG-3’ (forward) and 5’ -  
242 TAATGATCCTTCCGC AGGTTTAC-3’ (reverse) (Domes et al., 2007). The PCR protocol  
243 consisted of an initial activation step at 95 °C for 15 min, 35 amplification cycles (95°C for 45  
244 s, 57° C for 60 s, 72°C for 60 s) and a final elongation step at 72 °C for 10 min. All PCR

245 products were visualized on a 1% agarose gel, purified with the QIAquick PCR Purification  
246 Kit (Qiagen), and sequenced by Microsynth Seqlab (Göttingen, Germany), using the  
247 additional sequencing primers 18S554f 5'-AAGTCTGG TGCCAGCAGCCGC-3' , 18S1282r  
248 5'-TCACTCCACCAACTA AGAACGG C-3' , 18S1150f 5' -  
249 ATTGACGGAAGGGCACCACCAG-3' and 18S614r 5'- TCCAACACTACGAGCTTTTAAACC-3'  
250 (Domes et al., 2007). In total, we used 51 species for the phylogenetic tree, including four  
251 outgroup taxa. All taxa and accession numbers are available at GenBank (Supporting  
252 Information, Appendix S1, Table S1). We aimed to represent each family in the GB dataset  
253 with at least one species but very few rarer species could not be represented either because  
254 sequences are not available in public database or because we did not have sufficient  
255 material to sequence them. In total, the dataset represents 31 out of the 34 families found in  
256 the GB dataset and the three families we could not represent were very rare and present  
257 with very low abundances. Sequences were assembled and ambiguous positions were  
258 corrected in Sequencher 5.3 (Gene Codes Corporation, Ann Arbor, Michigan, USA) and  
259 aligned using ClustalW implemented in BioEdit v7.0.1 (Hall, 1999) with the multiple  
260 alignment parameters gap opening = 30 and gap extension = 0.3.

261 The alignment was truncated to the shortest sequence resulting in a length of 1,743 bp  
262 including gaps. Evolutionary model parameters were determined with jModelTest v2.0  
263 (Guindon & Gascuel, 2003; Darriba et al., 2012) using the AIC. The best-fit model for  
264 sequence evolution for 18S was GTR + I + G. The phylogenetic tree was constructed in R  
265 using packages '[ape](#)' (Paradis et al., 2004) and '[phangorn](#)' "~~ape~~" and "~~phangorn~~" (Paradis et  
266 al., 2004; (Schliep 2011, Schliep et al., 2017) using Maximum Likelihood and 1,000  
267 bootstrap replicates. The phylogenetic tree was reduced to 31 oribatid mite taxa  
268 representing one species per family using the drop.tip function, the R script is provided in the  
269 Appendix S1 (Supporting Information)

270 We used this oribatid mite phylogenetic tree to calculate metrics of phylogenetic diversity.

271 The resolution of the available phylogenetic information constrained us to calculate these

272 metrics at the family level. Specifically, we calculated the Faith's Index PD (Faith, 1992) and  
273 two distance based metrics: the mean phylogenetic distance (MPD) and the mean nearest  
274 taxon distance (MNTD). The index PD estimates the phylogenetic diversity of a community  
275 as the sum of the tree branch lengths connecting all species in the assemblage and as such  
276 it can be considered an estimate of point diversity with two components: species richness  
277 and amount of phylogenetic information across all the species in the assemblage. The  
278 indices MPD and MNTD measure the average phylogenetic distance between species in an  
279 assemblage. The MPD is based on mean distance of any taxon from every other taxon while  
280 MNTD is average of the distance between any taxon and its closest relative. Community  
281 phylogenetic metrics were calculated using the R package "picante" and other packages  
282 ('ape' and 'phytools') that support phylogenetically informed statistical analyses (~~'ape',~~  
283 ~~'phylobase', 'adephylo', 'phytools'~~ "ape", "phylobase", "adephylo", "phytools"; Kembel Paradis  
284 [et al., 2004](#), [Kembel 2010](#), [Revell 2012](#), Swenson, 2014)

#### 285

286 *Structural equation modelling (SEM)*

287 In order to quantify direct and indirect effects of climatic and soil variables (temperature,  
288 plant communities, pH, precipitation, organic matter) on community species composition  
289 (PCoA ordination axes) and diversity (richness and metric of phylogenetic diversity) we used  
290 Structural Equation Modelling (Grace, 2006). We started from an *a priori* model (Appendix  
291 S2, Fig. S1 in the Supporting Information) assuming that latitude, longitude and elevation  
292 drive the spatial variation of climate (i.e. precipitation and temperature), which correlates with  
293 spatial variation in organic matter. The variation in climate and organic matter then drives  
294 spatial variation in oribatid mite species composition. However, other factors that may vary  
295 with latitude and longitude, including biogeographical factors, may drive the spatial  
296 distribution of oribatid species. Biogeographical factors, which are implicitly accounted for by  
297 latitude and longitude, include the major spatially structured features of the geology of Britain

298 (Toghil, 2005), with northern areas (e.g, Scottish highlands) being generally older but also  
299 more affected by last glacial maxima than southern areas (e.g., southeast England).  
300 Eventually, all these environmental and geological changes in space determined spatial  
301 variation both in species composition and metrics of diversity, including species richness and  
302 phylogenetic diversity metrics that combined both richness and compositional information.  
303 Starting from this conceptual model, we fitted various versions of the model to the data until  
304 we obtained a parsimonious model that could adequately fit the data. Model fit was  
305 evaluated using the Chi-square test, and the RMSEA and CFI index, while amount of  
306 explained variation in community metrics and diversity indices (R-square) was used to  
307 measure the predictive power of the model (Grace, 2006; Shipley, 2016). SEM was  
308 performed using the R package '~~lavaan~~ [lavaan](#)' (Rosseel, 2011).

309

## 310 **Results**

311 A total of 141 species were found in this study, which represented more than one third of  
312 known oribatid mites in the British Isles (Luxton, 1996) and the vast majority of belowground  
313 species (the CS survey specifically focused on soil species while oribatid mites also live on  
314 aboveground moss and tree canopy). Observed species richness ranged from 1 to 28 while  
315 the Chao's estimator ranged from 1 to 161 species. Hectads with very low species richness  
316 always included arable and very infertile grassland soil, where environmental conditions  
317 typically supported only very poor oribatid mite communities or no oribatids at all. These  
318 soils were colonised only by very few opportunistic species such as some of the species in  
319 the genera *Tectocepheus*, *Liochthonius* and *Pantelozetes*. On the contrary, hectads with  
320 high species richness tended to be characterised by woodland or organic soils, where  
321 oribatid mites are known to be abundant. Observed hectad species richness displayed both  
322 clear latitudinal and longitudinal gradients with hectads in north-west Scotland being richer  
323 than in south-east England (Fig. 1). Instead, Chao's estimator showed a very patchy

324 distribution suggesting the existence of hotspots of species richness, mostly located in  
325 central and northern England, and Scotland (Fig. 1).

326 The species-area relationship was best fitted by a sigmoidal model (Fig. 2) meaning the  
327 existence of an upper limit below which species richness is relatively, but not completely,  
328 independent of area. Also, as area increases species richness is predicted to reach an upper  
329 asymptotic level (Lomolino, 2000). The classical power model and semi-log model provided  
330 a much poorer fit to the data.

331 The variation in oribatid mite species composition was mostly driven by the covariation  
332 between organic matter (measured as Loss-on-ignition), pH, and precipitation and  
333 variation in plant species composition, which was almost collinear with amount of organic  
334 matter. Although both community structure and environmental variables follow clear  
335 latitudinal patterns (Fig. 3), the total amount of variance accounted for by measured  
336 environmental variables was only 8%. Yet, this fraction of variation was statistically  
337 significant at  $P < 0.05$ . There was also 6% of variance accounted for by the spatially  
338 structured effect of environmental and plant variables. The pure effect of latitude, longitude  
339 and PCNMs (i.e., after removing environment) accounted only for 1% of community  
340 variance, meaning that the observed spatial variation in the assemblage is mostly co-varying  
341 with the spatial structure observed in the environmental variables (6%).

342 Metrics of phylogenetic diversity showed different types of spatial patterns (Fig. 4). The  
343 Faith's index showed gradients that were highly correlated to the same ones observed for  
344 plot species richness (compare Fig. 4 with Fig. 1) while MPD and MNTD mostly reflected  
345 longitudinal gradients. MPD is higher in the North and the East while MNTD seems more  
346 variable and reaching the highest value in the South-East (Fig. 4).

347 Structural equation modelling indicated that models including just latitude as a descriptor of  
348 position generally outperformed models with both latitude and longitude in terms of global fit  
349 metrics. For example, all models with both latitude and longitude resulted in Chi-square with

350 p-values much lower than 0.05 (i.e., model rejected) and with very poor CFI (<0.9) and  
351 RMSEA (>0.2) values. We therefore retained latitude and removed longitude from the  
352 subsequent models. Although latitude could affect indices of diversity both directly and  
353 indirectly, models with a direct link between latitude and diversity indices returned very poor  
354 global fit metrics and were therefore not considered further. Details on the models  
355 considered during the SEM exercise and their performances are provided Appendix S2. The  
356 optimal model (Fig. 5) suggests that organic matter is the major driver of oribatid mite  
357 community composition and that variation in species composition determines metrics of  
358 phylogenetic diversity. Specifically, greater shifts in oribatid mite community assemblages  
359 towards that typical of heath, bog and highly organic soil, were associated with higher  
360 phylogenetic diversity (positive correlation between PCoA1 and PD) but also lower mean  
361 nearest taxon distance (negative correlation between PCoA1 and MNTD). However, there is  
362 also an indirect positive effect of PCoA1 on MNTD via PD (positive correlation between PD  
363 and MNTD). The model could account for 50, 16 and 5 % of variance in PCoA 1 (major  
364 changes in species composition), PD and MNTD, respectively. The full ~~lavaan~~ output of the  
365 SEM is in Appendix S2 (Supporting Information)

366

#### 367 4. Discussion

368 Soil animal assemblages tend to be very species rich even at small scales. This has been  
369 explained as an effect of the high environmental and microbiological heterogeneity that  
370 some soil can display already from the 10 m to the sub-metre scale (Anderson, 1975; Giller,  
371 1996; Ettema & Wardle, 2002; Nielsen et al., 2010). A surprisingly large fraction of the  
372 variation observed in the distribution of soil species is very often left unexplained by variation  
373 in key soil variables such as pH and organic C, or even pollutants (Maraun & Scheu, 2000;  
374 Caruso et al., 2011, 2017; Maaß et al., 2015). Also, high degrees of stochasticity seem to  
375 characterise assembly dynamics of soil animals such as oribatid mites and collembolans at  
376 least at small to medium scales (Maaß et al., 2014; Dirilgen et al., 2018). Still, species

377 distributions seem structured at small and medium scales even when spatial structure  
378 cannot be explained by spatial gradients in environmental variables (Caruso et al., 2011;  
379 Zaitsev et al., 2013). At the regional scale of the Netherlands, Zaitsev *et al.* (2013) found that  
380 oribatid mite communities significantly changed along the East-West direction in the absence  
381 of a significant variation in precipitation and mean annual temperature. However, geological  
382 age (bedrock) and amount and quality of organic matter did change from East to West  
383 supporting richer communities in the older forest sites (Zaitsev *et al.* 2013). Our dataset  
384 supports this idea at the much broader scale of Great Britain, which is characterised by a  
385 relationship between climatic gradients and organic matter: in Britain very organic rich soils  
386 (i.e. bogs and peatlands) are ~~mostly found~~predominant in the North and West, and are  
387 characterised by a colder winter climate with more precipitation. Thus, as mean annual  
388 precipitation increases with latitude so does organic matter. This is reflected in our data by  
389 statistically significant, latitudinal changes in the oribatid mite communities, which prefer  
390 organic soil and woodland over low fertile grassland and cropland. Land use could also  
391 contribute to these patterns because, in GB, land is generally much more exploited for  
392 intensive farming in the south (e.g., England) than the north (Highlands in Scotland).  
393 However, our analysis independently accounted for vegetation types and latitudinal  
394 gradients in other properties and our results suggest a prominent role of organic matter *per*  
395 *se*. That means that, given the same land use and vegetation type, sites with higher organic  
396 matter are associated to specific oribatid mite composition and higher diversity overall.  
397 Species richness and metrics of phylogenetic diversities, too, follow this latitudinal gradient  
398 in community structure although metrics of phylogenetic diversity that take into account  
399 phylogenetic distance between species (MPD and MNTD) show patterns more complex than  
400 just a latitudinal gradient. The SEM showed that variation in distance based metrics of  
401 phylogenetic diversity (e.g., MNTD) seemed mostly explained by latitudinal changes in  
402 species composition rather than accumulation of species richness and phylogenetic diversity  
403 (PD). In fact, the direct and negative effect of the latitudinal changes in species composition  
404 on MNTD was statistically significant while the direct and positive effect of PD on MNTD was

405 not. The negative correlation between the latitudinal gradient in oribatid mite composition  
406 and MNTD suggests that the more the oribatid community moves to the species composition  
407 typical of woodland and highly organic soils the less the phylogenetic distance is between a  
408 species and its closest relatives in the local assemblages. This result suggests a process of  
409 environmental filtering and convergence toward specific assemblages (Webb, 2000). The  
410 SEM, however, could explain only 5% of the variance observed in MNTD and 16% of the  
411 variance observed in PD suggesting that the measured environmental variables are  
412 generally weak predictors of these broad biodiversity metrics. On the contrary, the SEM  
413 could explain about 50% of the variance observed in the latitudinal gradient in species  
414 composition, which implies species composition is much more predictable than compound  
415 metrics of biodiversity such as phylogenetic diversity (PD). Specifically, the latitudinal  
416 changes in species composition seem best explained by latitudinal variation in organic  
417 matter and precipitation, regardless of variation in phylogenetic diversity.

418 Latitude directly correlates with precipitation and organic matter distribution merely because  
419 of the north-south climatic gradient. When taking into account the direct and indirect effects  
420 of latitude, precipitation and organic matter on oribatid mite species composition, the  
421 strongest effect was that of organic matter. Precipitation, too, had a statistically significant,  
422 direct effect on community structure but the effect was much smaller than that of organic  
423 matter, which is consistent with Zaitsev *et al.* (2013). Instead, the direct effect of latitude on  
424 species composition was small and not statistically significant, which implies that latitudinal  
425 changes in species composition are driven by latitudinal changes in other variables, namely  
426 precipitation and organic matter. Alternative SEMs that linked latitude, longitude, organic  
427 matter and precipitation directly to metrics of diversity had a very poor global fit supporting  
428 the notion that large-scale gradients in soil oribatid mite diversity are driven by the factors  
429 that drive changes in species composition. Still, changes in species composition explained  
430 only a relatively small fraction of changes in species richness and phylogenetic diversity,  
431 suggesting a potential role for smaller scale heterogeneity. This heterogeneity is not

432 captured by our predictors and suggests that microscale variation in edaphic properties  
433 remain a fundamental driver of species distribution and diversity in these communities. This  
434 is confirmed by the fact that some hectads resulted to be biodiversity hotspot in terms of  
435 estimated species richness. We could not resolve the variables driving this patchy pattern  
436 but we speculate that this is driven by soil environmental heterogeneity within hectads, which  
437 could be caused by unmeasured variation in habitat fragmentation and land-use intensity  
438 (see also supplementary results in the Supporting Information, Appendix S3, Table S2)

439 Despite the latitudinal patterns observed in oribatid mites and contrary to what has been  
440 observed in small- and medium-scale studies (Caruso et al., 2011; Maaß et al., 2015), the  
441 investigated community had limited spatial structure, even when considering spatial variation  
442 that is not explainable by spatial structure in environmental variables. In comparison, the  
443 microbial communities of GB seem to be much more spatially structured (Griffiths et al.,  
444 2011), which suggests the interesting hypothesis of a decoupling between large-scale  
445 patterns in soil microbes and animals.

446

## 447 **Conclusions**

448 Latitudinal gradients in organic matter are the most important predictor of latitudinal changes  
449 in species composition of oribatid mites across the spatial extent of Great Britain. These  
450 changes partially drive variation in species richness and phylogenetic diversity but a  
451 significant fraction of the variation observed in these metrics remained unexplained,  
452 suggesting a potential role for unmeasured environmental heterogeneity at medium and  
453 small scales. Despite small and medium scale heterogeneity, macroecological patterns in  
454 this major group of soil animals are predictable by the climatic factors that control variation in  
455 plant community structure and organic matter.

456

457 **Data accessibility**

Formatted: Font: Bold

458 [All Countryside Survey data used in this study are available from the NERC Environmental](#)  
459 [Information Data Centre. Data on oribatid mite records \(\[https://doi.org/10.5285/05cae9ae-\]\(https://doi.org/10.5285/05cae9ae-86c4-4a2f-a628-6c7fd0882459\)](#)  
460 [86c4-4a2f-a628-6c7fd0882459](#)) are linked to soil properties  
461 [\(<https://doi.org/10.5285/9d1eada2-3f8b-4a7b-a9b0-a7a04d05ff72>\) and vegetation](#)  
462 [\(<https://doi.org/10.5285/07896bb2-7078-468c-b56d-fb8b41d47065>\) by unique plots codes in](#)  
463 [each dataset. Access to spatial coordinates for plot locations in Countryside Survey is](#)  
464 [restricted to minimise risk of compromising future surveys and preserve anonymity of](#)  
465 [locations: plot locations at 10 × 10 km resolution, required to collate oribatid records at a](#)  
466 [hectad level and retrieve linked climatic data \(Met Office\), were therefore obtained under a](#)  
467 [bespoke data licence agreement between TC and the Centre for Ecology & Hydrology.](#)

## 468 References

469 Anderson, J. (1975) The enigma of soil animal species diversity. *Progress in soil zoology*, ~~pp~~ 51–58.  
470 Springer.

471 Barr, C.J., Bunce, R.G.H., Gillespie, M.K., Howard, D.C., Maskell, L.C., Norton, L.R., Scott, R.J., Shield,  
472 E.R., Smart, S.M., Stuart, R.C., Watkins, J.W., Wood, C.M., (2014) Countryside Survey 1998  
473 vegetation plot data. NERC Environmental Information Data  
474 Centre. <https://doi.org/10.5285/07896bb2-7078-468c-b56d-fb8b41d47065>

475  
476 Bezemer, T.M., Fountain, M.T., Barea, J.M., Christensen, S., Dekker, S.C., Duyts, H., van Hal, R.,  
477 Harvey, J.A., Hedlund, K., Maraun, M., Mikola, J., Mladenov, A.G., Robin, C., de Ruiter, P.C.,  
478 Scheu, S., Setälä, H., Šmilauer, P., & van der Putten, W.H. (2010) Divergent composition but  
479 similar function of soil food webs of individual plants: plant species and community effects.  
480 *Ecology*, **91**, 3027–3036.

481 Bivand, R.S., Pebesma, E.J., Gomez-Rubio, V., & Pebesma, E.J. (2008) *Applied spatial data analysis*  
482 *with R*. Springer.

483 Black, H., Garnett, J., Ainsworth, G., Coward, P., Creamer, R., Ellwood, S., Horne, J., Hornung, M.,  
484 Kennedy, V., & Monson, F. (2002) *MASQ: Monitoring and Assessing Soil Quality in Great*  
485 *Britain. Survey Model 6: Soils and Pollution*. Environment Agency.

486 Black, H., Parekh, N., Chaplow, J., Monson, F., Watkins, J., Creamer, R., Potter, E., Poskitt, J.,  
487 Rowland, P., & Ainsworth, G. (2003) Assessing soil biodiversity across Great Britain: national  
488 trends in the occurrence of heterotrophic bacteria and invertebrates in soil. *Journal of*  
489 *Environmental Management*, **67**, 255–266.

490 Black, H.I.J., Chaplow, J.S., Ainsworth, G., Coward, P.A., Creamer, R., Ellwood, S., Horne, J., Hornung,  
491 M., Kennedy, V.H., Raine, L., Osborn, D., Parekh, N.R., Parrington, J., Poskitt, J.M., Potter, E.,

Formatted: Font: Calibri, 11 pt

Formatted: Font: Calibri, 11 pt, Font color: Auto

Formatted: Bibliography

Formatted: Font: Calibri, 11 pt, Font color: Auto

- 492 Reeves, N., Rowland, A.P., Self, P., Turner, S., Watkins, J.W., Wood, C.M., Woods, C., Wright,  
493 J. (2016). Soil physico-chemical properties 1998 [Countryside Survey]. NERC Environmental  
494 Information Data Centre. <https://doi.org/10.5285/9d1eada2-3f8b-4a7b-a9b0-a7a04d05ff72>
- 495 Borcard, D. & Legendre, P. (2002) All-scale spatial analysis of ecological data by means of principal  
496 coordinates of neighbour matrices. *Ecological Modelling*, **153**, 51–68.
- 497 Borcard, D., Legendre, P., Avois-Jacquet, C., & Tuomisto, H. (2004) Dissecting the spatial structure of  
498 ecological data at multiple scales. *Ecology*, **85**, 1826–1832.
- 499 Borcard, D., Legendre, P., & Drapeau, P. (1992) Partialling out the Spatial Component of Ecological  
500 Variation. *Ecology*, **73**, 1045–1055.
- 501 Brusaard, L., Aanen, D.K., Briones, M.J.I., Decaëns, T., De Deyn, G.B., Fayle, T.M., James, S., & Nobre,  
502 T. (2012) Biogeography and phylogenetic community structure of soil invertebrate  
503 ecosystem engineers: global to local patterns, implications for ecosystem functioning and  
504 services, and global environmental change impacts. In: Wall, D.H. et al. (eds.), *Soil Ecology  
505 and Ecosystem Services*. Oxford University Press, UK, 201–232. Wall, D.H. et al. (eds.), *Soil  
506 Ecology and Ecosystem Services*, pp. 201–232. Oxford University Press.
- 507 Bunce, R.G.H., Barr, C., Gillespie, M., Howard, D., Scott, W., Smart, S., Van de Poll, H., & Watkins, J.  
508 (1999) *Vegetation of the British countryside-the Countryside Vegetation System. ECOFACT  
509 Volume 1*. DETR.
- 510 Cadotte, M.W., Jonathan Davies, T., Regetz, J., Kembel, S.W., Cleland, E., & Oakley, T.H. (2010)  
511 Phylogenetic diversity metrics for ecological communities: integrating species richness,  
512 abundance and evolutionary history. *Ecology Letters*, **13**, 96–105.
- 513 Caruso, T., Migliorini, M., Rota, E., & Bargagli, R. (2017) Highly diverse urban soil communities: Does  
514 stochasticity play a major role? *Applied Soil Ecology*, **110**, 73–78.
- 515 Caruso, T., Taormina, M., & Migliorini, M. (2011) Relative role of deterministic and stochastic  
516 determinants of soil animal community: a spatially explicit analysis of oribatid mites. *Journal  
517 of Animal Ecology*, **81**, 214–221.
- 518 Chesson, P. (2000) Mechanisms of maintenance of species diversity. *Annual Reviews in Ecology and  
519 Systematics*, **31**, 343–366.
- 520 Chiu, C., Wang, Y., Walther, B.A., & Chao, A. (2014) An improved nonparametric lower bound of  
521 species richness via a modified Good–Turing frequency formula. *Biometrics*, **70**, 671–682.
- 522 Clark, J.S. & McLachlan, J.S. (2003) Stability of forest biodiversity. *Nature*, **423**, 635–638.
- 523 Coleman, D.C., Crossley, D., & Hendrix, P.F. (2004) *Fundamentals of soil ecology*. Academic Press,  
524 London, UK.
- 525 Darriba, D., Taboada, G.L., Doallo, R., & Posada, D. (2012) jModelTest 2: more models, new heuristics  
526 and parallel computing. *Nature methods*, **9**, 772.
- 527 Davison, J., Moora, M., Öpik, M., et al. (2015) Global assessment of arbuscular mycorrhizal fungus  
528 diversity reveals very low endemism. *Science*, **349**, 970–973.

Formatted: Font: Calibri, 11 pt, Font color: Auto

Formatted: Font: Calibri

- 529 de Vries, F.T., Manning, P., Tallowin, J.R.B., Mortimer, S.R., Pilgrim, E.S., Harrison, K.A., Hobbs, P.J.,  
530 Quirk, H., Shipley, B., Cornelissen, J.H.C., Kattge, J., & Bardgett, R.D. (2012) Abiotic drivers  
531 and plant traits explain landscape-scale patterns in soil microbial communities. *Ecology*  
532 *Letters*, **15**, 1230–1239.
- 533 Decaëns, T. (2010) Macroecological patterns in soil communities. *Global Ecology and Biogeography*,  
534 **19**, 287–302.
- 535 Dirilgen, T., Juceviča, E., Melecis, V., Querner, P., & Bolger, T. (2018) Analysis of spatial patterns  
536 informs community assembly and sampling requirements for Collembola in forest soils. *Acta*  
537 *Oecologica*, **86**, 23–30.
- 538 Domes, K., Norton, R.A., Maraun, M., & Scheu, S. (2007) Reevolution of sexuality breaks Dollo's law.  
539 *Proceedings of the National Academy of Sciences*, **104**, 7139–7144.
- 540 Dray, S., Legendre, P., & Peres-Neto, P.R. (2006) Spatial modelling: a comprehensive framework for  
541 principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, **196**, 483–  
542 493.
- 543 Dumbrell, A.J., Nelson, M., Helgason, T., Dytham, C., & Fitter, A.H. (2010) Relative roles of niche and  
544 neutral processes in structuring a soil microbial community. *The ISME J*, **4**, 337–345.
- 545 Emmett, B., Reynolds, B., Chamberlain, P., Rowe, E., Spurgeon, D., Brittain, S., Frogbrook, Z., Hughes,  
546 S., Lawlor, A., & Poskitt, J. (2010) Countryside survey: soils report from 2007.
- 547 Ettema, C.H. & Wardle, D.A. (2002) Spatial soil ecology. *Trends in Ecology & Evolution*, **17**, 177–183.
- 548 Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological conservation*, **61**, 1–  
549 10.
- 550 Fierer, N., Strickland, M.S., Liptzin, D., Bradford, M.A., & Cleveland, C.C. (2009) Global patterns in  
551 belowground communities. *Ecology Letters*, **12**, 1238–1249.
- 552 Firbank, L., Barr, C., Bunce, R., Furse, M., Haines-Young, R., Hornung, M., Howard, D., Sheail, J., Sier,  
553 A., & Smart, S. (2003) Assessing stock and change in land cover and biodiversity in GB: an  
554 introduction to Countryside Survey 2000. *Journal of environmental management*, **67**, 207–  
555 218.
- 556 Giller, P.S. (1996) The diversity of soil communities, the 'poor man's tropical rainforest.' *Biodiversity*  
557 *and Conservation*, **5**, 135–168.
- 558 Gotelli, N.J. & Colwell, R.K. (2001) Quantifying biodiversity: procedures and pitfalls in the  
559 measurement and comparison of species richness. *Ecology letters*, **4**, 379–391.
- 560 Grace, J.B. (2006) *Structural equation modeling and natural systems*. Cambridge University Press,  
561 [Cambridge, UK](http://www.cambridge.org/9780521876223)
- 562 Griffiths, R.I., Thomson, B.C., James, P., Bell, T., Bailey, M., & Whiteley, A.S. (2011) The bacterial  
563 biogeography of British soils. *Environmental microbiology*, **13**, 1642–1654.
- 564 Guindon, S. & Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large  
565 phylogenies by maximum likelihood. *Systematic biology*, **52**, 696–704.

566 Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program  
567 for Windows 95/98/NT. *Nucleic acids symposium series*, **41**, 95–98.

568 Hubbell, S.P. (2005) Neutral theory in community ecology and the hypothesis of functional  
569 equivalence. *Functional Ecology*, **19**, 166–172.

570 Kardol, P., Martijn Bezemer, T., & van der Putten, W.H. (2006) Temporal variation in plant–soil  
571 feedback controls succession. *Ecology Letters*, **9**, 1080–1088.

572 Keith, A.M., Griffiths, R.I., Henrys, P.A., Hughes, S., Lebron, I., Maskell, L.C., Ogle, S.M., Robinson,  
573 D.A., Rowe, E.C., & Smart, S.M. (2015) Monitoring soil natural capital and ecosystem services  
574 by using large-scale survey data. **Soil Ecosystems Services**.

575 Keith, A.M., Monson, F.D., Chaplow, J.S., Black, H.I.J., Self, P., Rollett, A., Wood, C.M. (2018). Oribatid  
576 mites 1998 [Countryside Survey] . NERC Environmental Information Data  
577 Centre. <https://doi.org/10.5285/05cae9ae-86c4-4a2f-a628-6c7fd0882459>

578 [Kembel, S., Ackerly, D., Blomberg, S., Cowan, P., Helmus, M. & Webb, C. \(2008\). picante: Tools for](#)  
579 [Integrating Phylogenies and Ecology, Version 0.4-0. Software package available from](#)  
580 <http://picante.r-forge.r-project.org>.

581

582 Legendre, P. & Legendre, L. (1998) *Numerical Ecology*. Elsevier, Amsterdam, The Netherlands.

583 Levine, J.M. & HilleRisLambers, J. (2009) The importance of niches for the maintenance of species  
584 diversity. *Nature*, **461**, 254–257.

585 Lindo, Z. & Winchester, N. (2009) Spatial and environmental factors contributing to patterns in  
586 arboreal and terrestrial oribatid mite diversity across spatial scales. *Oecologia*, **160**, 817–  
587 825.

588 Lomolino, M.V. (2000) Ecology’s most general, yet protean 1 pattern: the species-area relationship.  
589 *Journal of Biogeography*, **27**, 17–26.

590 Luxton, M. (1996) Oribatid mites of the British Isles: a check-list and notes on biogeography (Acari:  
591 Oribatida). *Journal of Natural History*, **30**, 803–822.

592 Maaß, S., Maraun, M., Scheu, S., Rillig, M.C., & Caruso, T. (2015) Environmental filtering vs. resource-  
593 based niche partitioning in diverse soil animal assemblages. *Soil Biology and Biochemistry*,  
594 **85**, 145–152.

595 Maaß, S., Migliorini, M., Rillig, M.C., & Caruso, T. (2014) Disturbance, neutral theory, and patterns of  
596 beta diversity in soil communities. *Ecology and Evolution*, **4**, 4766–4774.

597 Maraun, M. & Scheu, S. (2000) The structure of oribatid mite communities (Acari, Oribatida):  
598 patterns, mechanisms and implications for future research. *Ecography*, **23**, 374–382.

599 Matheron, G. (1963) Principles of geostatistics. *Economic geology*, **58**, 1246–1266.

600 Nielsen, U.N. (2014) Global-scale patterns of soil nematode assemblage structure in relation to  
601 climate and ecosystem properties. *Glob. Ecol. Biogeogr.*, **23**, 968–978.

Formatted: Font: Bold

Formatted: Highlight

Formatted: Font: +Body (Calibri), 11 pt, Font color: Auto

Formatted: Font: +Body (Calibri), 11 pt, Font color: Auto

Formatted: None

Formatted: Font: +Body (Calibri), 11 pt

Formatted: Font: (Default) +Body (Calibri), 11 pt, Underline, English (United Kingdom), Pattern: Clear (White)

Formatted: English (United States)

Formatted: Normal, None

602 Nielsen, U.N., Osler, G.H.R., Campbell, C.D., Neilson, R., Burslem, D.F.R.P., & van der Wal, R. (2010)  
603 The Enigma of Soil Animal Species Diversity Revisited: The Role of Small-Scale Heterogeneity.  
604 *PLoS ONE*, **5**, e11567.

605 Norton, R.A., Bonamo, P.M., Grierson, J.D., & Shear, W.A. (1988) Oribatid mite fossils from a  
606 terrestrial Devonian deposit near Gilboa, New York. *Journal of Paleontology*, **62**, 259–269.

607 O'Hara, R.B. (2005) Species richness estimators: how many species can dance on the head of a  
608 pin? *Journal of Animal Ecology*, **74**, 375–386.

609 Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M.H.H., Oksanen, M.J., & Suggests, M.  
610 (2007) The vegan package. *Community ecology package*, **10**, 631–637.

611 Paradis, E., Claude, J., & Strimmer, K. (2004) APE: analyses of phylogenetics and evolution in R  
612 language. *Bioinformatics*, **20**, 289–290.

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

615 Ramirez, K.S., Leff, J.W., Barberán, A., Bates, S.T., Betley, J., Crowther, T.W., Kelly, E.F., Oldfield, E.E.,  
616 Shaw, E.A., Steenbock, C., Bradford, M.A., Wall, D.H., & Fierer, N. (2014) Biogeographic  
617 patterns in below-ground diversity in New York City's Central Park are similar to those  
618 observed globally. *Proceedings of the Royal Society of London B: Biological Sciences*, **281**, .

619 [Revell, L. J. \(2012\) phytools: An R package for phylogenetic comparative biology \(and other things\).](#)  
620 [Methods in Ecology and Evolution](#), **3**, 217-223.

621 Reynolds, B., Chamberlain, P., Poskitt, J., Woods, C., Scott, W., Rowe, E., Robinson, D., Frogbrook, Z.,  
622 Keith, A., & Henrys, P. (2013) Countryside Survey: National "Soil Change" 1978–2007 for  
623 Topsoils in Great Britain—acidity, carbon, and total nitrogen status. *Vadose Zone Journal*, **12**,  
624 .

625 Rosseel, Y. (2011) lavaan: an R package for structural equation modeling and more Version 0.4-9  
626 (BETA).

627 Schaefer, I., Norton, R.A., Scheu, S., & Maraun, M. (2010) Arthropod colonization of land – Linking  
628 molecules and fossils in oribatid mites (Acari, Oribatida). *Molecular Phylogenetics and*  
629 *Evolution*, **57**, 113–121.

630 [Schliep, K.P. \(2011\) phangorn: Phylogenetic analysis in R. Bioinformatics](#), **27(4)**, 592-593

631 Schliep, K., Potts, A.J., Morrison, D.A., & Grimm, G.W. (2017) Intertwining phylogenetic trees and  
632 networks. *Methods in Ecology and Evolution*, **8**, 1212–1220.

633 Sheail, J. & Bunce, R. (2003) The development and scientific principles of an environmental  
634 classification for strategic ecological survey in the United Kingdom. *Environmental*  
635 *conservation*, **30**, 147–159.

636 Shear, W.A., Bonamo, P.M., Grierson, J.D., Rolfe, W.I., Smith, E.L., & Norton, R.A. (1984) Early land  
637 animals in North America: evidence from Devonian age arthropods from Gilboa, New York.  
638 *Science*, **224**, 492–494.

Formatted: Highlight

Formatted: Font: +Body (Calibri), 11 pt

Formatted: Font: +Body (Calibri), 11 pt

Formatted: Highlight

Formatted: Font: +Body (Calibri), 11 pt

Formatted: Font: +Body (Calibri), 11 pt

Formatted: Font: +Body (Calibri), 11 pt, Bold

Formatted: Font: +Body (Calibri), 11 pt

Formatted: Font: +Body (Calibri), 11 pt

639 Shipley, B. (2016) *Cause and correlation in biology: a user's guide to path analysis, structural*  
640 *equations and causal inference with R*. Cambridge University Press, [Cambridge, UK](#).

641 Soininen, J. (2011) Macroecology of unicellular organisms – patterns and processes. *Environmental*  
642 *Microbiology Reports*, **4**, 10–22.

643 Swenson, N.G. (2014) *Functional and phylogenetic ecology in R*. Springer.

Formatted: Highlight

644 Thomson, B.C., Ostle, N., McNamara, N., Bailey, M.J., Whiteley, A.S., & Griffiths, R.I. (2010)  
645 Vegetation Affects the Relative Abundances of Dominant Soil Bacterial Taxa and Soil  
646 Respiration Rates in an Upland Grassland Soil. *Microbial Ecology*, **59**, 335–343.

647 Toghil, P. (2005) *The Geology of Britain: an introduction*. The Crowood Press.

Formatted: Highlight

648 Van Der Heijden, M.G.A., Bardgett, R.D., & Van Straalen, N.M. (2008) The unseen majority: soil  
649 microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology*  
650 *Letters*, **11**, 296–310.

651 Wagner, H.H. (2003) Spatial covariance in plant communities: integrating ordination, geostatistics,  
652 and variance testing. *Ecology*, **84**, 1045–1057.

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

655 Webb, C.O. (2000) Exploring the phylogenetic structure of ecological communities: an example for  
656 rain forest trees. *The American Naturalist*, **156**, 145–155.

657 Weigmann, G. (2006) *Acari, Actinochaetida. Hornmilben (Oribatida)*. Goecke & Evers., Keltern.

Formatted: Highlight

658 Zaitsev, A.S., Straalen, N.M. van, & Berg, M.P. (2013) Landscape geological age explains large scale  
659 spatial trends in oribatid mite diversity. *Landscape Ecology*, **28**, 285–296.

660

661 ~~Data accessibility statement: The data supporting the results are available from the~~  
662 ~~Environmental and Information Data Centre of NERC (UK). The spatial coordinates require a~~  
663 ~~license agreement to be accessed.~~

664

## 665 Biosketch

666 **Tancredi Caruso** is a quantitative ecologist investigating the processes that structure  
667 terrestrial biodiversity in space and time

668 **Ina Schaefer** is an evolutionary biologist with expertise on soil fauna

669 **Aidan M Keith** is an ecologist investigating land use and climate impacts on soil biology and  
670 functional processes

671 **Frank Monson** is a soil animal taxonomist

672

673 All authors developed the concept of the paper; AMK and FM compiled and collected the  
674 data, TC analysed the data, IS compiled and collected molecular data and constructed the  
675 phylogenetic trees. All authors contributed substantially to the writing of the ms

676

677 **Figure Legends**

678 **Figure 1** a) species richness and b) Chao's estimator of species richness for each 10 × 10  
679 km plots. The maps were obtained via kriging interpolation at the hectad scale. Red  
680 represents high values, yellow low values. Species richness (a) displays a clear latitudinal  
681 gradient with richness increasing northward along the southwest-northeast direction whereas  
682 Chao's estimator (b) displays a patchy distribution suggesting the existence of hotspots if  
683 species richness.

684

685 **Figure 2** Species-area relationships for the oribatid mites of Great Britain. The three fitted  
686 models (power law, semi-log, sigmoidal) all fit the data reasonably well but the AIC criterion  
687 clearly shows that the sigmoidal model provides the best fit (blue dotted line).

688

689 **Figure 3** Principal Coordinates Analysis (PCoA) ordination of oribatid mites (a). The first axis  
690 (PCoA1) is a gradient that follow changes in mean Aggregate vegetation Vegetation Class,  
691 with more organic vegetation types and woodland soil scoring on the positive side of  
692 PCoA1 and agricultural vegetation types (crops and grasslands) and arable soil scoring on  
693 the negative side of PCoA1. This gradient is also correlated to organic matter, latitude, and  
694 precipitation, which are all positively correlated with PCoA1. ~~In fact,~~ a kriging interpolation  
695 of PCoA1 show a clear latitudinal gradient (b)

696

697 **Figure 4.** Kriging interpolation of three metrics of Phylogenetic Diversity. The Faith's index  
698 (a) showed gradients that were very correlated to the same ones observed for plot species  
699 richness (Fig. 1a) while Mean Pairwise Distance (MPD) and Mean Nearest Taxon Distance  
700 (MNTD) mostly reflected longitudinal gradients; although MPD reaches the highest values in

Formatted: Font: Not Bold

701 the North and the East while MNTD seems more variable and ~~reaching-reaches~~ the highest  
702 value in the South-East.

703

704 **Figure 5** Structural equation model linking latitude and abiotic parameters to oribatid mite  
705 species composition (PCoA1 of Fig. 3) and diversity (Faith Index PD and Mean Nearest  
706 Taxon Distance or MNTD of Fig. 4). Species richness was highly correlated to PD and was  
707 thus excluded, while MNTD and MPD returned similar results in this SEM and we selected  
708 MNTD, which provided the best fit. The model is supported by all metrics of global fit (Chi-  
709 square = 8.809 with 9 df and p-value of 0.185, CFI = 0.989 and RMSEA = 0.059). **Figures**  
710 **Values** beside ~~the~~ arrows are the path standardised coefficients. Black arrow stands for  
711 positive coefficient and gray arrows for negative coefficients. Paths statistically significant at  
712 p-value < 0.05 are in bold. All paths were statistically significant except for the direct effect of  
713 PD on MNTD and the direct effect of Latitude on PCoA 1. See also Supporting Information b  
714 for the full model output, including exact values of path coefficients, R-square values,  
715 standard deviations and statistical significance of parameter estimates. The model could  
716 account for 50, 16 and 5 % of variance in PCoA 1, PD and MNTD respectively. The model  
717 could also account for 55 and 17 % of variance in Organic Matter and Precipitation,  
718 respectively.