Article type: Research Article

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

Short running title: *Macroecological determinants of soil animals*

Authors: Tancredi Caruso[a], Ina Schaefer[b] , Frank Monson[c], Aidan M Keith[d]

[a] School of Biological Sciences and Institute for Global Food Security, Queen's University Belfast. Medical Biology Centre, 97 Lisburn Road, Belfast BT9 7BL, Northern Ireland, UK

[b]J.F.Blumenbach Institute of Zoology and Anthropology University of Göttingen, Untere Karspüle 2, 37073 Göttingen, Germany

[c] Associate, World Museum Liverpool, William Brown Street, Liverpool, L3 8EN

[d]Centre for Ecology & Hydrology, Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster LA1 4AP, UK

Corresponding Author: Tancredi Caruso, *t.caruso@qub.ac.uk*

Acknowledgments

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

Abstract

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

Location: we modelled the multivariate distribution of more than 100 species using biodiversity data collected across Great Britain in the framework of the Countryside Survey (http://www.countrysidesurvey.org.uk).

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

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

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

Keywords: soil macroecology; animals; diversity; distribution; community phylogenetics; Oribatida

1. Introduction

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

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

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

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

Methods

Database Background

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

2.2 Oribatid extraction and identification

Cores were processed over five days using a dry Tullgren extraction method and all invertebrate specimens collected into 70% ethanol preservative (Emmett et al., 2010) . Once 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 ×400 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**.** Oribatid species
177 records and taxonomic details were then collated into a dataset that is available upon
178 request from the NERC Environmental Information Centre (see Keith et al. 2018 for details
179 to access the data).

180

181 *2.3 Associated environmental data*

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

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

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

196

197 *2.4 Statistical analyses*

198 *Community and environmental data*

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

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

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

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

Phylogenetic methods

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

Genomic DNA was extracted from single individuals using the DNeasy® Blood and Tissue Kit (Qiagen, Manchester, UK) following the manufacturer’s protocol for animal tissue. Amplification of the 18S region was performed in 25 µl volumes containing 12.5 µl HotStarTaq Mastermix (Qiagen), 5 µl of template DNA, 1 µl of each primer (100 pM) and 5.5 µl H_2O. Primers for PCR were 5’ -TAC- CTGGTTGATCCTGCCAG-3’ (forward) and 5’ -TAATGATCCTTCCGC AGGTTCAC-3’ (reverse) (Domes et al., 2007). The PCR protocol consisted of an initial activation step at 95 °C for 15 min, 35 amplification cycles (95°C for 45 s, 57° C for 60 s, 72°C for 60 s) and a final elongation step at 72 °C for 10 min. All PCR

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

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

We used this oribatid mite phylogenetic tree to calculate metrics of phylogenetic diversity. The resolution of the available phylogenetic information constrained us to calculate these metrics at the family level. Specifically, we calculated the Faith's Index PD (Faith, 1992) and

two distance based metrics: the mean phylogenetic distance (MPD) and the mean nearest taxon distance (MNTD). The index PD estimates the phylogenetic diversity of a community as the sum of the tree branch lengths connecting all species in the assemblage and as such it can be considered an estimate of point diversity with two components: species richness and amount of phylogenetic information across all the species in the assemblage. The indices MPD and MNTD measure the average phylogenetic distance between species in an assemblage. The MPD is based on mean distance of any taxon from every other taxon while MNTD is average of the distance between any taxon and its closest relative. Community phylogenetic metrics were calculated using the R package “picante” and other packages that support phylogenetically informed statistical analyses (“ape”, “phylobase”,”adephylo”,”phytools”; Swenson, 2014)

Structural equation modelling (SEM)

In order to quantify direct and indirect effects of climatic and soil variables (temperature, plant communities, pH, precipitation, organic matter) on community species composition (PCoA ordination axes) and diversity (richness and metric of phylogenetic diversity) we used Structural Equation Modelling (Grace, 2006). We started from an *a priori* model (Appendix S2, Fig. S1 in the Supporting Information) assuming that latitude, longitude and elevation drive the spatial variation of climate (i.e. precipitation and temperature), which correlates with spatial variation in organic matter. The variation in climate and organic matter then drives spatial variation in oribatid mite species composition. However, other factors that may vary with latitude and longitude, including biogeographical factors, may drive the spatial distribution of oribatid species. Biogeographical factors, which are implicitly accounted for by latitude and longitude, include the major spatially structured features of the geology of Britain (Toghil, 2005), with northern areas (e.g, Scottish highlands) being generally older but also more affected by last glacial maxima than southern areas (e.g., southeast England).

297 Eventually, all these environmental and geological changes in space determined spatial
298 variation both in species composition and metrics of diversity, including species richness and
299 phylogenetic diversity metrics that combined both richness and compositional information.
300 Starting from this conceptual model, we fitted various versions of the model to the data until
301 we obtained a parsimonious model that could adequately fit the data. Model fit was
302 evaluated using the Chi-square test, and the RMSEA and CFI index, while amount of
303 explained variation in community metrics and diversity indices (R-square) was used to
304 measure the predictive power of the model (Grace, 2006; Shipley, 2016). SEM was
305 performed using the R package lavaan (Rosseel, 2011).

306

307 **Results**

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

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

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

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

344 Structural equation modelling indicated that models including just latitude as a descriptor of
345 position generally outperformed models with both latitude and longitude in terms of global fit
346 metrics. For example, all models with both latitude and longitude resulted in Chi-square with
347 p-values much lower than 0.05 (i.e., model rejected) and with very poor CFI (<0.9) and
348 RMSEA (>0.2) values. We therefore retained latitude and removed longitude from the
349 subsequent models. Although latitude could affect indices of diversity both directly and

indirectly, models with a direct link between latitude and diversity indices returned very poor global fit metrics and were therefore not considered further. Details on the models considered during the SEM exercise and their performances are provided Appendix S2. The optimal model (Fig. 5) suggests that organic matter is the major driver of oribatid mite community composition and that variation in species composition determines metrics of phylogenetic diversity. Specifically, greater shifts in oribatid mite community assemblages towards that typical of heath, bog and highly organic soil, were associated with higher phylogenetics diversity (positive correlation between PCoA1 and PD) but also lower mean nearest taxon distance (negative correlation between PCoA1 and MNTD). However, there is also an indirect positive effect of PCoA1 on MNTD via PD (positive correlation between PD and MNTD). The model could account for 50, 16 and 5 % of variance in PCoA 1 (major changes in species composition), PD and MNTD, respectively. The full lavaan ouput of the SEM is in Appendix S2 (Supporting Information)

4. **Discussion**

Soil animal assemblages tend to be very species rich even at small scales. This has been explained as an effect of the high environmental and microbiological heterogeneity that some soil can display already from the 10 m to the sub-metre scale (Anderson, 1975; Giller, 1996; Ettema & Wardle, 2002; Nielsen et al., 2010). A surprisingly large fraction of the variation observed in the distribution of soil species is very often left unexplained by variation in key soil variables such as pH and organic C, or even pollutants (Maraun & Scheu, 2000; Caruso et al., 2011, 2017; Maaß et al., 2015). Also, high degrees of stochasticity seem to characterise assembly dynamics of soil animals such as oribatid mites and collembolans at least at small to medium scales (Maaß et al., 2014; Dirilgen et al., 2018). Still, species distributions seem structured at small and medium scales even when spatial structure cannot be explained by spatial gradients in environmental variables (Caruso et al., 2011; Zaitsev et al., 2013). At the regional scale of the Netherlands, Zaitsev *et al.* (2013) found that

oribatid mite communities significantly changed along the East-West direction in the absence of a significant variation in precipitation and mean annual temperature. However, geological age (bedrock) and amount and quality of organic matter did change from East to West supporting richer communities in the older forest sites (Zaitsev *et al.* 2013). Our dataset supports this idea at the much broader scale of Great Britain, which is characterised by a relationship between climatic gradients and organic matter: in Britain very organic rich soils (i.e. bogs and peatlands) are mostly found in the North and West, and are characterised by a colder winter climate with more precipitation. Thus, as mean annual precipitation increases with latitude so does organic matter. This is reflected in our data by statistically significant, latitudinal changes in the oribatid mite communities, which prefer organic soil and woodland over low fertile grassland and cropland. Land use could also contribute to these patterns because, in GB, land is generally much more exploited for intensive farming in the south (e.g., England) than the north (Highlands in Scotland). However, our analysis independently accounted for vegetation types and latitudinal gradients in other properties and our results suggest a prominent role of organic matter *per se*. That means that, given the same land use and vegetation type, sites with higher organic matter are associated to specific oribatid mite composition and higher diversity overall. Species richness and metrics of phylogenetic diversities, too, follow this latitudinal gradient in community structure although metrics of phylogenetic diversity that take into account phylogenetic distance between species (MPD and MNTD) show patterns more complex than just a latitudinal gradient. The SEM showed that variation in distance based metrics of phylogenetic diversity (e.g., MNTD) seemed mostly explained by latitudinal changes in species composition rather than accumulation of species richness and phylogenetic diversity (PD). In fact, the direct and negative effect of the latitudinal changes in species composition on MNTD was statistically significant while the direct and positive effect of PD on MNTD was not. The negative correlation between the latitudinal gradient in oribatid mite composition and MNTD suggests that the more the oribatid community moves to the species composition typical of woodland and highly organic soils the less the phylogenetic distance is between a species and its closest relatives in the

405 local assemblages. This result suggests a process of environmental filtering and
406 convergence toward specific assemblages (Webb, 2000).The SEM, however could explain
407 only 5% of the variance observed in MNTD and 16% of the variance observed in PD
408 suggesting that the measured environmental variables are generally weak predictors of
409 these broad biodiversity metrics. On the contrary, the SEM could explain about 50% of the
410 variance observed in the latitudinal gradient in species composition, which implies species
411 composition is much more predictable than compound metrics of biodiversity such as
412 phylogenetic diversity (PD). Specifically, the latitudinal changes in species composition seem
413 best explained by latitudinal variation in organic matter and precipitations, regardless of
414 variation in phylogenetic diversity.

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

estimated species richness. We could not resolve the variables driving this patchy pattern but we speculate that this is driven by soil environmental heterogeneity within hectads, which could be caused by unmeasured variation in habitat fragmentation and land-use intensity (see also supplementary results in the Supporting Information, Appendix S3, Table S2)

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

Conclusions

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

References

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

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

461

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

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

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

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

476 Black, H.I.J., Chaplow, J.S., Ainsworth, G., Coward, P.A., Creamer, R., Ellwood, S., Horne, J.,
477 Hornung, M., Kennedy, V.H., Raine, L., Osborn, D., Parekh, N.R., Parrington, J., Poskitt,
478 J.M., Potter, E., Reeves, N., Rowland, A.P., Self, P., Turner, S., Watkins, J.W., Wood,
479 C.M., Woods, C., Wright, J. (2016). Soil physico-chemical properties 1998 [Countryside
480 Survey]. NERC Environmental Information Data
481 Centre. https://doi.org/10.5285/9d1eada2-3f8b-4a7b-a9b0-a7a04d05ff72

482 Borcard, D. & Legendre, P. (2002) All-scale spatial analysis of ecological data by means of principal
483 coordinates of neighbour matrices. *Ecological Modelling*, **153**, 51–68.

484 Borcard, D., Legendre, P., Avois-Jacquet, C., & Tuomisto, H. (2004) Dissecting the spatial structure of
485 ecological data at multiple scales. *Ecology*, **85**, 1826–1832.

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

488 Brusaard, L., Aanen, D.K., Briones, M.J.I., Decaëns, T., De Deyn, G.B., Fayle, T.M., James, S.., & Nobre,
489 T. (2012) Biogeography and phylogenetic community structure of soil invertebrate
490 ecosystem engineers: global to local patterns, implications for ecosystem functioning and
491 services, and global environmental change impacts. In: Wall, D.H. et al. (eds.), Soil Ecology
492 and Ecosystem Services. Oxford University Press, UK, 201–232. *Wall, D.H. et al. (eds.), Soil*
493 *Ecology and Ecosystem Services* pp. 201–232. Oxford University Press.

494 Bunce, R.G.H., Barr, C., Gillespie, M., Howard, D., Scott, W., Smart, S., Van de Poll, H., & Watkins, J.
495 (1999) *Vegetation of the British countryside-the Countryside Vegetation System. ECOFACT*
496 *Volume 1.* DETR.

497 Cadotte, M.W., Jonathan Davies, T., Regetz, J., Kembel, S.W., Cleland, E., & Oakley, T.H. (2010)
498 Phylogenetic diversity metrics for ecological communities: integrating species richness,
499 abundance and evolutionary history. *Ecology Letters*, **13**, 96–105.

500 Caruso, T., Migliorini, M., Rota, E., & Bargagli, R. (2017) Highly diverse urban soil communities: Does
501 stochasticity play a major role? *Applied Soil Ecology*, **110**, 73–78.

502 Caruso, T., Taormina, M., & Migliorini, M. (2011) Relative role of deterministic and stochastic
503 determinants of soil animal community: a spatially explicit analysis of oribatid mites. *Journal*
504 *of Animal Ecology*, **81**, 214–221.

505 Chesson, P. (2000) Mechanisms of maintenance of species diversity. *Annual Reviews in Ecology and*
506 *Systematics*, **31**, 343–366.

507 Chiu, C., Wang, Y., Walther, B.A., & Chao, A. (2014) An improved nonparametric lower bound of
508 species richness via a modified Good–Turing frequency formula. *Biometrics*, **70**, 671–682.

509 Clark, J.S. & McLachlan, J.S. (2003) Stability of forest biodiversity. *Nature*, **423**, 635–638.

510 Coleman, D.C., Crossley, D., & Hendrix, P.F. (2004) *Fundamentals of soil ecology.* Academic press.

511 Darriba, D., Taboada, G.L., Doallo, R., & Posada, D. (2012) jModelTest 2: more models, new heuristics
512 and parallel computing. *Nature methods*, **9**, 772.

513 Davison, J., Moora, M., Öpik, M., et al. (2015) Global assessment of arbuscular mycorrhizal fungus
514 diversity reveals very low endemism. *Science*, **349**, 970.

515 de Vries, F.T., Manning, P., Tallowin, J.R.B., Mortimer, S.R., Pilgrim, E.S., Harrison, K.A., Hobbs, P.J.,
516 Quirk, H., Shipley, B., Cornelissen, J.H.C., Kattge, J., & Bardgett, R.D. (2012) Abiotic drivers
517 and plant traits explain landscape-scale patterns in soil microbial communities. *Ecology*
518 *Letters*, **15**, 1230–1239.

519 Decaëns, T. (2010) Macroecological patterns in soil communities. *Global Ecology and Biogeography*,
520 **19**, 287–302.

521 Dirilgen, T., Jucevič̆a, E., Melecis, V., Querner, P., & Bolger, T. (2018) Analysis of spatial patterns
522 informs community assembly and sampling requirements for Collembola in forest soils. *Acta*
523 *Oecologica*, **86**, 23–30.

524 Domes, K., Norton, R.A., Maraun, M., & Scheu, S. (2007) Reevolution of sexuality breaks Dollo's law.
525 *Proceedings of the National Academy of Sciences*, **104**, 7139–7144.

526 Dray, S., Legendre, P., & Peres-Neto, P.R. (2006) Spatial modelling: a comprehensive framework for
527 principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, **196**, 483–
528 493.

529 Dumbrell, A.J., Nelson, M., Helgason, T., Dytham, C., & Fitter, A.H. (2010) Relative roles of niche and
530 neutral processes in structuring a soil microbial community. The *ISME J*, **4**, 337–345.

531 Emmett, B., Reynolds, B., Chamberlain, P., Rowe, E., Spurgeon, D., Brittain, S., Frogbrook, Z., Hughes,
532 S., Lawlor, A., & Poskitt, J. (2010) Countryside survey: soils report from 2007.

533 Ettema, C.H. & Wardle, D.A. (2002) Spatial soil ecology. *Trends in Ecology & Evolution*, **17**, 177–183.

534 Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological conservation*, **61**, 1–
535 10.

536 Fierer, N., Strickland, M.S., Liptzin, D., Bradford, M.A., & Cleveland, C.C. (2009) Global patterns in
537 belowground communities. *Ecology Letters*, **12**, 1238–1249.

538 Firbank, L., Barr, C., Bunce, R., Furse, M., Haines-Young, R., Hornung, M., Howard, D., Sheail, J., Sier,
539 A., & Smart, S. (2003) Assessing stock and change in land cover and biodiversity in GB: an
540 introduction to Countryside Survey 2000. *Journal of environmental management*, **67**, 207–
541 218.

542 Giller, P.S. (1996) The diversity of soil communities, the 'poor man's tropical rainforest.' *Biodiversity*
543 *and Conservation*, **5**, 135–168.

544 Gotelli, N.J. & Colwell, R.K. (2001) Quantifying biodiversity: procedures and pitfalls in the
545 measurement and comparison of species richness. *Ecology letters*, **4**, 379–391.

546 Grace, J.B. (2006) *Structural equation modeling and natural systems.* Cambridge University Press,

547 Griffiths, R.I., Thomson, B.C., James, P., Bell, T., Bailey, M., & Whiteley, A.S. (2011) The bacterial
548 biogeography of British soils. *Environmental microbiology*, **13**, 1642–1654.

549 Guindon, S. & Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large
550 phylogenies by maximum likelihood. *Systematic biology*, **52**, 696–704.

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

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

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

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

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

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

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

566 Lindo, Z. & Winchester, N. (2009) Spatial and environmental factors contributing to patterns in
567 arboreal and terrestrial oribatid mite diversity across spatial scales. *Oecologia*, **160**, 817–
568 825.

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

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

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

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

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

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

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

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

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

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

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

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

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

Ramirez, K.S., Leff, J.W., Barberán, A., Bates, S.T., Betley, J., Crowther, T.W., Kelly, E.F., Oldfield, E.E., Shaw, E.A., Steenbock, C., Bradford, M.A., Wall, D.H., & 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 of London B: Biological Sciences*, **281**, .

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

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

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**, 113–121.

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

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

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

617 Shipley, B. (2016) *Cause and correlation in biology: a user's guide to path analysis, structural*
618 *equations and causal inference with R.* Cambridge University Press.

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

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

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

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

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

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

631 Wardle, D.A. (2006) The influence of biotic interactions on soil biodiversity. *Ecology Letters*, **9**, 870–
632 886.

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

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

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

638

639 ***Data accessibility statement****:* The data supporting the results are available from the
640 Environmental and Information Data Centre of NERC (UK). The spatial coordinates require a
641 license agreement to be accessed.

642

643 **Biosketch**

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

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

647 **Aidan M Keith** is an ecologist investigating land use and climate impacts on soil biology

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

649

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

653

Figure Legends

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

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

Figure 3 PCoA ordination of oribatid mites (a). The first axis is a gradient that follow changes in vegetation, with more organic and woodland soil scoring on the positive site of PCoA1 and grasslands and arable soil on the negative side of PCoA1. This gradient is also correlated to organic matter, latitude, and precipitation, which are all positively correlated with PCoA1. In fact, a kriging interpolation of PCoA1 show a clear latitudinal gradient (b)

Figure 4. Kriging interpolation of three metric of Phylogenetic Diversity. The Faith's index (a) showed gradients that were very correlated to the same ones observed for plot species richness (Fig. 1a) while MPD and MNTD mostly reflected longitudinal gradients although MPD reaches the highest values in the North and the East while MNTD seems more variable and reaching the highest value in the South-East.

678 **Figure 5** Structural equation model linking latitude and abiotic parameters to oribatid mite
679 species composition (PCoA1 of Fig. 3) and diversity (Faith Index PD and Mean Nearest
680 Taxon Distance or MNTD of Fig. 4). Species richness was highly correlated to PD and was
681 thus excluded, while MNTD and MPD returned similar results in this SEM and we selected
682 MNTD, which provided the best fit. The model is supported by all metrics of global fit (Chi-
683 square = 8.809 with 9 df and p-value of 0.185, CFI = 0.989 and RMSEA = 0.059). Figures
684 besides the arrows are the path standardised coefficients. Black arrow stands for positive
685 coefficient and gray arrows for negative coefficients. Paths statistically significant at p-value
686 < 0.05 are in bold. All paths were statistically significant except for the direct effect of PD on
687 MNTD and the direct effect of Latitude on PCoA 1. See also Supporting Information b for the
688 full model output, including exact values of path coefficients, R-square values, standard
689 deviations and statistical significance of parameter estimates. The model could account for
690 50, 16 and 5 % of variance in PCoA 1, PD and MNTD respectively. The model could also
691 account for 55 and 17 % of variance in Organic Matter and Precipitation, respectively.

Figure 1

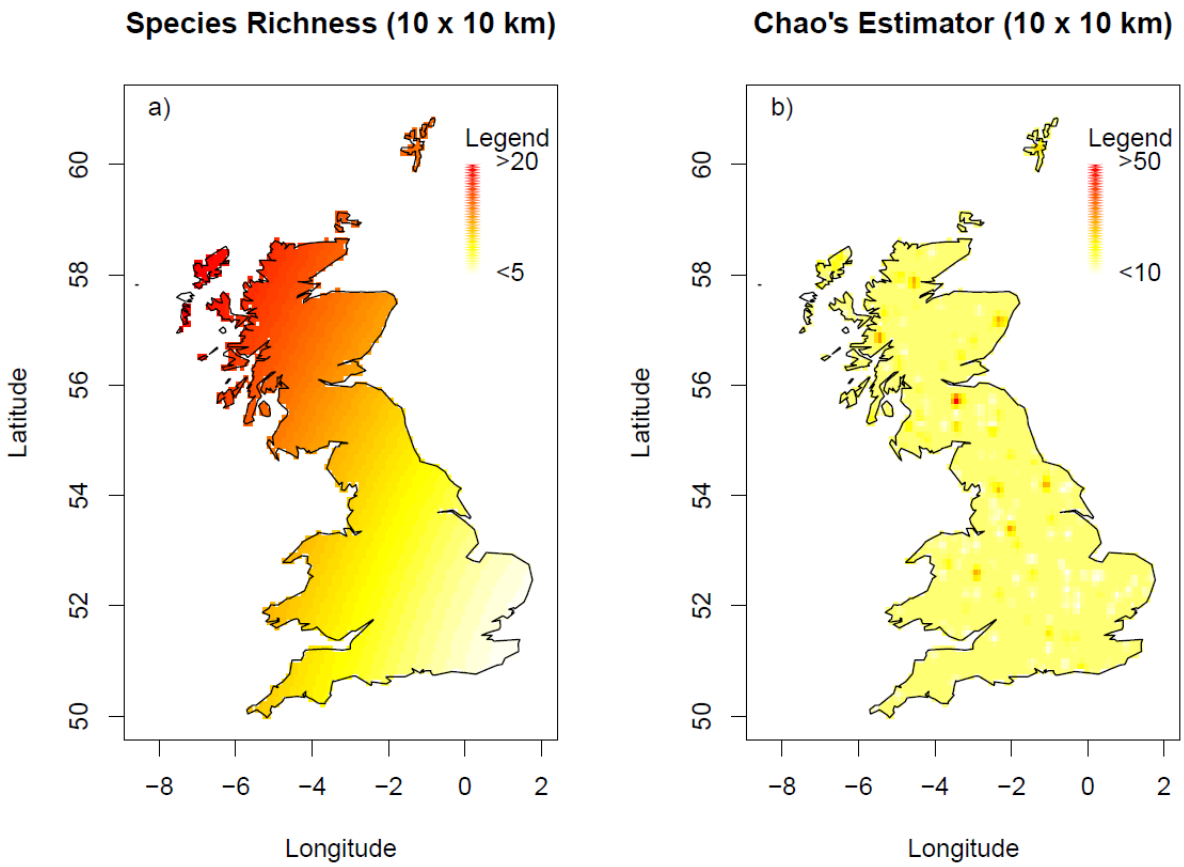

Species Richness (10 x 10 km)
a)
Legend
>20
<5
Latitude
Longitude
50
52
54
56
58
60
−8
−6
−4
−2
0
2
Chao's Estimator (10 x 10 km)
b)
Legend
>50
<10
Latitude
Longitude

Figure 2

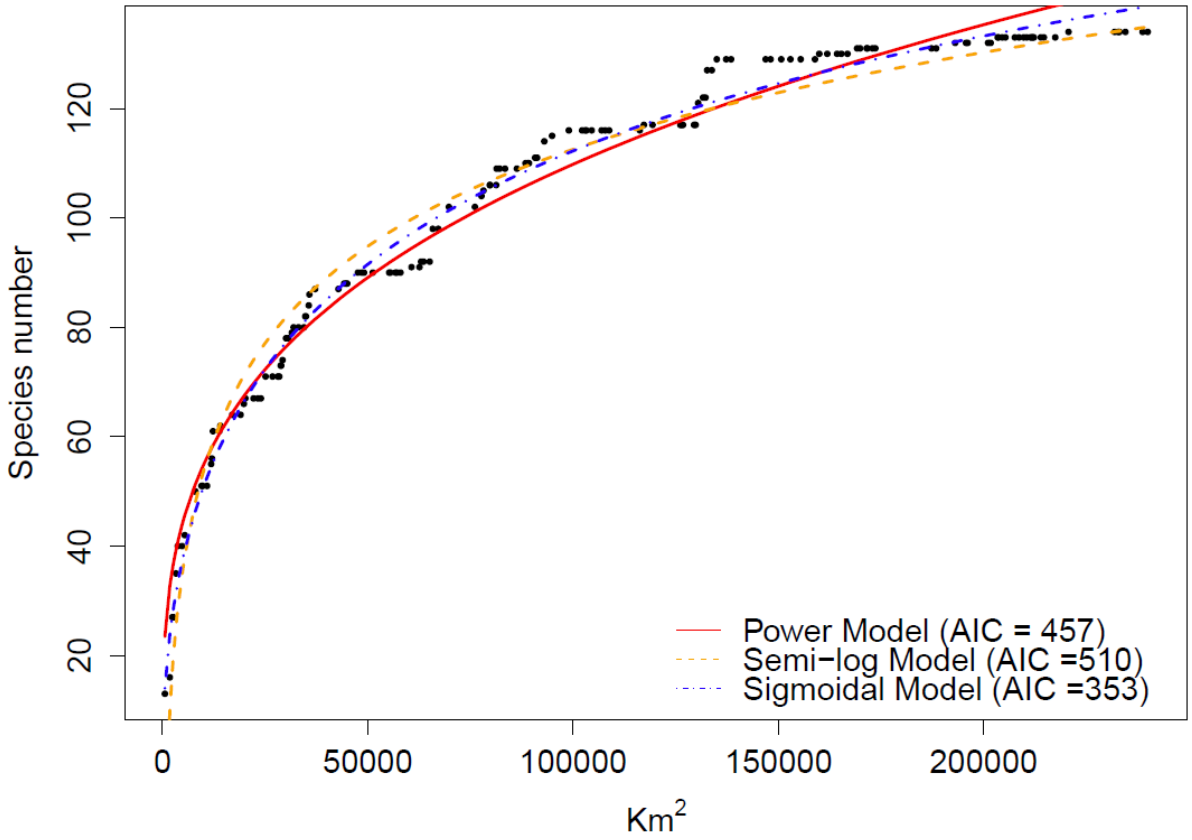

Species number
20
40
60
80
100
120
0
50000
100000
150000
200000
Km2
Power Model (AIC = 457)
Semi−log Model (AIC =510)
Sigmoidal Model (AIC =353)

Figure 3

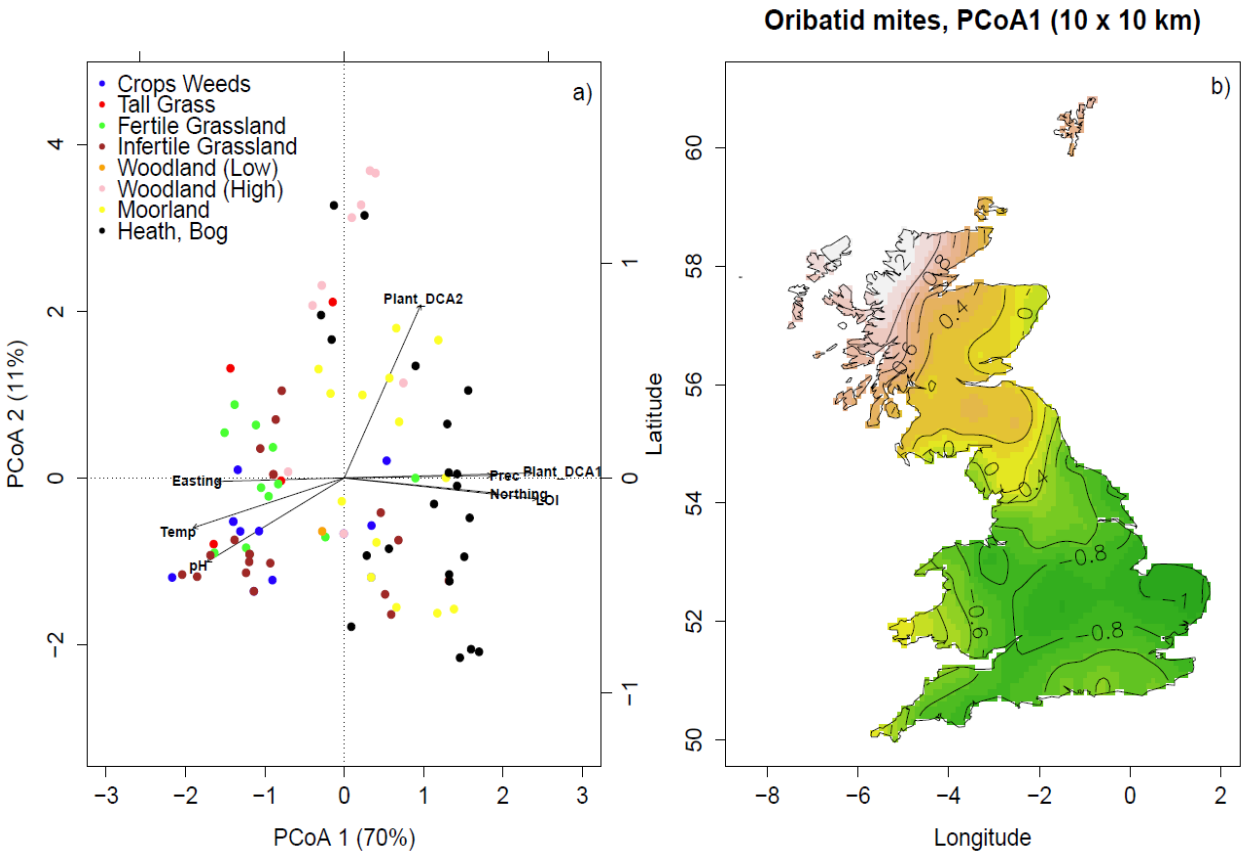

a)
Crops Weeds
Tall Grass
Fertile Grassland
Infertile Grassland
Woodland (Low)
Woodland (High)
Moorland
Heath, Bog
Plant_DCA2
Plant_DCA1
Prec
Northing
LOI
Easting
Temp
pH
PCoA 2 (11%)
PCoA 1 (70%)
Oribatid mites, PCoA1 (10 x 10 km)
b)
Latitude
Longitude

Figure 4

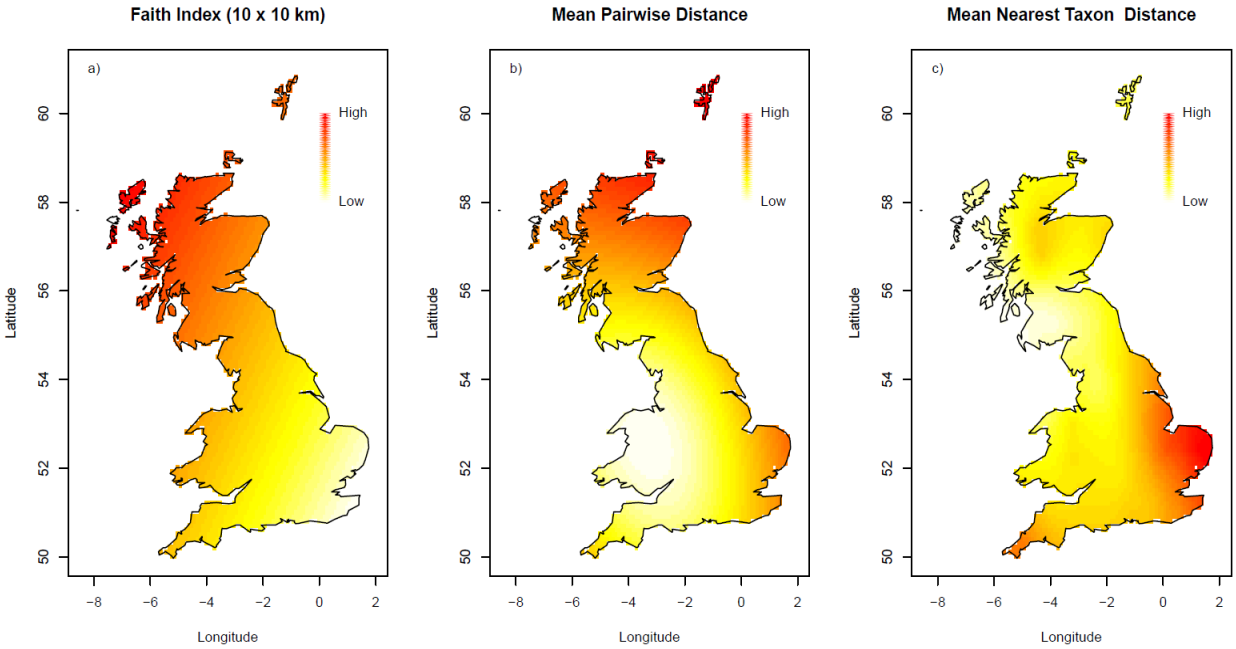
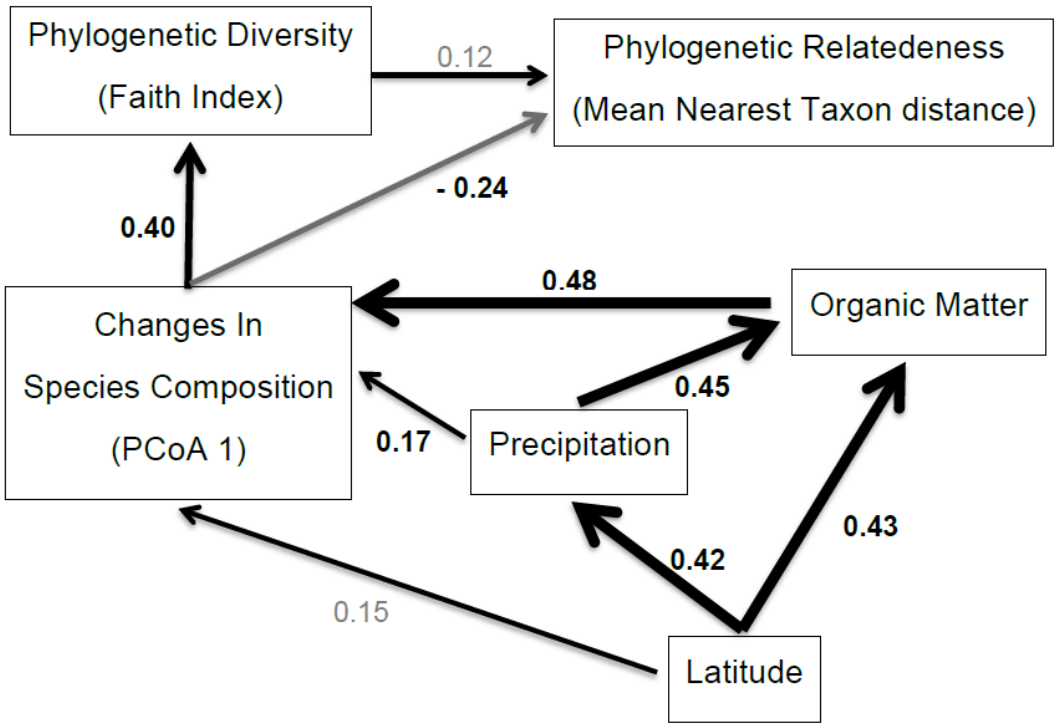

Faith Index (10 x 10 km)
Mean Pairwise Distance
Mean Nearest Taxon Distance
a)
b)
c)
High
Low
Latitude
Longitude
50
52
54
56
58
60
−8
−6
−4
−2
0
2

Figure 5


Phylogenetic Diversity
(Faith Index)
0.12
Phylogenetic Relatedeness
(Mean Nearest Taxon distance)
- 0.24
0.40
0.48
Organic Matter
Changes In
Species Composition
(PCoA 1)
0.45
0.17
Precipitation
0.43
0.42
0.15
Latitude