Article (refereed) - postprint

Keith, Aidan M.; Boots, Bas; Hazard, Christina; Niechoj, Robin; Arroyo, Julio; Bending, Gary D.; Bolger, Tom; Breen, John; Clipson, Nicholas; Doohan, Fiona M.; Griffin, Christine T.; Schmidt, Olaf. 2012 Cross-taxa congruence, indicators and environmental gradients in soils under agricultural and extensive land management. *European Journal of Soil Biology*, 49 (Mar-Apr). 55-62. 10.1016/j.ejsobi.2011.08.002

Cross-taxa congruence, indicators and environmental gradients in soils under agricultural and extensive land management

Aidan. M. Keith[a,b,c,*, †], Bas Boots[a,†], Christina Hazard[a], Robin Niechoj[d], Julio Arroyo[a], Gary D. Bending[e], Tom Bolger[a], John Breen[d], Nicholas Clipson[a], Fiona M. Doohan[a], Christine T. Griffin[b] and Olaf Schmidt[f]

[a]*UCD School of Biology and Environmental Science, University College Dublin, Belfield, Dublin 4, Ireland.*

[b]*Department of Biology, National University of Ireland, Maynooth, Kildare, Ireland.*

[c]*Centre for Ecology and Hydrology, Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster, LA1 4AP, UK.*

[d]*Department of Life Sciences, University of Limerick, Limerick, Ireland.*

[e]*School of Life Sciences, University of Warwick, Wellesbourne, Warwick, UK..*

[f]*UCD School of Agriculture, Food Science and Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland.*

*Corresponding author at: Centre for Ecology & Hydrology, Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster, LA1 4AP, United Kingdom. Tel: +44 (0)1524 595871; Fax: +44 (0)1524 61536. *E-mail address*: ake@ceh.ac.uk (A. M. Keith).

†AMK and BB contributed equally to this work.

Abstract

Important steps in developing reliable bioindicators for soil quality are characterising soil biodiversity and determining the response of its components to environmental factors across a range of land uses and soil types. Baseline data from a national survey in Ireland were used to explore relationships between diversity and composition of micro-organisms (bacteria, fungi, mycorrhiza), and micro-, meso- and macro-fauna (nematodes; mites; earthworms, ants) across a general gradient representing dominant land-uses (arable, pasture, rough-grazing, forest and bogland). These diversity data were also linked to soil physico-chemical properties. Differences in diversity and composition of meso- and macrofauna, but not microbes, were clear between agriculturally-managed (arable and pasture) and extensively-managed (rough-grazing and bogland) soils corresponding to a broad division between 'mineral' and 'organic' soils. The abundance, richness and composition of nematode and earthworm taxa were significantly congruent with a number of the other groups. Further analysis, using significant indicator species from each group, identified potential target taxa and linked them to soil environmental gradients. This study suggests that there is potential surrogacy between the diversity of key soil taxa groups and that different sets of bioindicators may be most effective under agricultural and extensive land-use.

Keywords: Soil monitoring, land use, biodiversity, physico-chemical gradients, bioindicators, soil community structure.

1. Introduction

Large-scale soil monitoring schemes that include biological measurements are already established in many European countries [e.g. 1,2,3]. These are important in detecting impacts of broader environmental changes but also in assessing more specific effects of land management practices on soil organisms and the ecosystem services they support. The EU thematic strategy on soil protection has identified major threats to soil quality and biodiversity [4]. However, no integrated EU-wide programme of biological monitoring exists and therefore recent impetus has been towards a reliable and harmonised programme across different countries [5,6,7,8].

While the advantages of a harmonised system are clear, it is challenging to reach consensus on which groups of taxa, or particular "keystone" taxa, act as good indicators of soil quality and should be monitored [5,9]. Indeed, there are different types of bioindicator, and the appropriate measures may depend on whether the need is for an indicator of soil biodiversity itself, the ecological soil status, or an environmental change imposed on the soil ecosystem [10]. A number of studies have examined cross-taxon congruency in aquatic systems e.g. [11,12] and above-ground terrestrial systems [13,14,15], but such assessments for below-ground biodiversity are scarce. This type of assessment can subsequently be used to identify potential surrogacy in soil bioindicators.

Understanding how the diversity of different groups of soil taxa may provide information on the quality and status of soils remains a challenge, because for many ecosystems we lack biological typologies and the opportunity for comparative analyses. Consequently, an important step in developing reliable bioindicators for soil health is the characterisation of soil biodiversity and then determining the response of its components to environmental factors across a range of land uses and soil types.

Systematic biodiversity surveys require co-located data including a representative range of soil taxa, covering dominant land use and soil types over an extensive geographical area in order to make inferences about potential soil bioindicators. Here, we use data from a national survey of soil biodiversity carried out in Ireland to a) characterise soil taxa assemblages across five major land uses (classified as arable, pasture, forest, rough-grazing and bogland), b) examine how abundance, richness

and composition of different major groups of soil taxa are related to each other across land uses, and **c)** determine potential indicator taxa for land use and management and their relationship with soil environmental factors.

2. Material and Methods

2.1 National soil biodiversity survey

A baseline soil biodiversity survey ('CréBeo' project) was undertaken to contribute to the development of a national soil monitoring network in Ireland. This was linked with an earlier initiative in soil chemical monitoring, the National Soil Database (NSD) project [16], which contains site information, a suite of chemical soil measurements and GIS-supported mapping for 1310 locations. A sub-set of the NSD sites was selected, based on a number of criteria including the inclusion of major land uses and soil types in proportion to their known frequency in Ireland and geographical spread. In total, 61 sites were sampled during the soil biodiversity survey including arable (n=14), pasture (n=21), forest (n=10; 5 each of coniferous and broadleaved forest), rough-grazing (n=8) and bogland (n=8) land uses (Table 1; Supplementary Fig. A1). The major soil types were classified following Gardiner and Radford [17] and included: Acid brown earths (n=10), shallow brown earths (n=3), brown podzolics (n=9), grey-brown podzolics (n=10), podzolics (n=3), gleys (n=10), lithosols (n=3) and peats (n=13). Soil data held in the NSD were utilised to examine relationships between physico-chemical properties and soil taxa. Much of these soil data was produced by the 'SoilC' project [18] which had 55 sites in common with the present soil biodiversity baseline survey.

2.2 Sampling and processing of soil organisms

A 20 m × 20 m plot was centered on the NSD [16] GPS coordinates of each site. The different groups of soil taxa were sampled within this plot using separate protocols as briefly outlined below (see Supplementary File A for detailed methods):

1. Soil bacteria and fungi were surveyed at all sites. Twenty soil cores (20 cm depth) were collected and bulked per site, sieved (4 mm) and stored at –20ºC for DNA extraction. Molecular fingerprinting techniques were used to assess general bacterial and fungal diversity.

2. Arbuscular mycorrhizal fungi (AMF) were surveyed within 45 NSD locations in 2006. Bulked soil samples (obtained from step 1.) were used for bioassays with *Trifolium repens* L. (White clover) and molecular fingerprinting techniques were used to characterise the AMF diversity in the plant roots.

3. Nematodes were surveyed at all sites by sugar centrifugation extraction from a 100 cm^3 sub-sample of bulked soil (obtained from step 1.). Nematodes were counted and approximately 100 nematodes from each site were identified to at least genus level (except for Rhabditidae and Neodiplogasteridae).

4. Micro-arthropods (Collembola and Acari) were extracted from 4 intact soil cores (5 cm diameter, 5 cm depth) per site using a Kempson apparatus. Oribatid (mainly detritivorous) and mesostigmatid (predatory) mites were identified to species level.

5. Earthworms were extracted in the field by hand-sorting four 25 cm × 25 cm × 25 cm soil blocks and, where feasible, by chemical expellant from four 50 cm × 50 cm quadrats. Mature individuals were identified to species level.

6. Soil-dwelling ant diversity was assessed using 20-metre-line of crumb baits to attract species that forage and by visual searches (30–60 min) within a 100 metre-radius of each GPS location. All ants were identified to species level.

2.3 Statistical analyses

Unless stated otherwise, all analyses were conducted in the R statistical environment [19]. The effect of land use on the richness of each soil taxa group was analysed using a Kruskal-Wallis non-parametric (χ^2) test since replication of land use was unbalanced. Patterns of site compositional similarity were investigated using Non-metric Multidimensional Scaling (NMDS). Similarity matrices were calculated using Bray-Curtis associations on square-root transformed data and clustering of sites according to soil type and land use was tested by PERMANOVA using the distance matrices in the *adonis* function of the *vegan* package [20]. Homogeneity of multivariate dispersion [21,22] was tested

using the *betadisper* function in *vegan* [20]. However, soil-dwelling ants were omitted for the *adonis* analysis due to their sparse coverage and low diversity. The same analyses were repeated using only the arable and pasture sites to examine whether the patterns were consistent within only agricultural systems. The effect of soil type was also examined only within arable and pasture sites since it tends to be confounded by land use in organic soils (e.g. boglands contain peats).

Congruence between different taxa groups was assessed using Spearman correlation of abundance, richness, Shannon diversity and Bray-Curtis similarity. Spearman coefficients and significance of correlations for abundance, richness and Shannon diversity were calculated using the *Rcorr* function of the *Hmisc* package [23]. In addition, Mantel tests were used to determine the significance of rank correlations between Bray-Curtis matrices of different taxa groups in the *vegan* package [20].

Indicator species analysis (IndVal) was conducted to examine the fidelity and specificity of individual taxa to the different land uses [24] within the *indicspecies* package [25]. Group-equalized options were used to account for differences in numbers of sites between each land use. The number of indicator taxa significant at $P < 0.05$ within each different group of soil taxa and land use were recorded. This analysis was repeated using only arable and pasture sites to assess potential indicators within agricultural land uses. We acknowledge that this represents a large number of individual analyses but consider this as a liberal method of identifying the potential pool of indicator taxa and of reducing the dataset to taxa likely to be important as indicators.

The correlation between abundances of all significant indicator taxa (as identified above) and soil physico-chemical gradients was assessed using Redundancy Analyses (RDA). RDA is a constrained ordination, aiming to find linear combinations of the predictor variables that explain the greatest variation in the data cloud [26], based on the smallest residual sum of squares. Small differences in values of abiotic data between samples can have large impacts on the outcome of multivariate analyses [27]. Therefore, in order to reduce variation between samples, all abiotic factors were square-root transformed and standardised. The abundance of all indicator taxa were also standardised (subtract minimum from value and divide by the range) to account for the different

scales of measurement between taxa groups. The model to explain variability encompassed a selection of properties including relatively easy to obtain information (moisture content, pH, bulk density, C, N and P concentrations), and those that did not show any co-linearity (i.e. where correlation between variables was <0.80. The RDA was repeated using those indicator taxa identified within IndVal analyses using arable and pasture sites. RDA analyses were visualised in two dimensional ordinations using CANOCO for Windows v.4.5 [28]

3. Results

3.1 The biota

A total of 1148 bacterial, 874 fungal, 446 AMF, 94 nematode, 108 mite, 19 earthworm and 8 ant taxa were recorded across all sites. The greatest number of taxa recorded at one site was 356 for bacteria 159 for fungi, 78 for AMF, 25 for nematodes, 27 for mites, 11 for earthworm, and 5 for ants. The greatest number of taxa recorded did not occur at an arable site for any of the taxa groups. The smallest number of bacteria, fungi and AMF taxa were all recorded at an arable site. The smallest richness of nematode taxa was recorded at a bogland site, while low richness of mites and earthworms occurred in several land uses, and all land uses had sites where no ant species were recorded (Table 1).

3.2 Land use and soil biodiversity

There were significant differences in the richness of nematode, mite, earthworm and ant taxa between land uses, but not in the richness of bacteria, fungi or AMF (Table 1). Mean taxon richness was greatest in pasture for nematodes and earthworms, rough-grazing for mites, and both rough-grazing and bogland for ants (Table 1). This pattern across soil taxa was similar in the land uses where the greatest number of taxa were recorded (Table 1). There were no differences in the richness of any taxa between soil types within arable and pasture land uses (data not shown).

There was no significant effect of land use on bacteria composition ($F_{4,35}$ = 1.02, P = 0.357) or AMF composition ($F_{4,35}$ = 1.42, P = 0.065)(Supplementary Fig. B1). However, there was a highly

178 significant influence of land use on fungi ($F_{4,35}$ = 1.20, P = 0.001), nematode ($F_{4,35}$ = 6.36, P = 0.001),
179 mite ($F_{4,33}$ = 1.58, P = 0.001) and earthworm ($F_{4,33}$ = 3.05, P = 0.001) composition. Although
180 multivariate dispersion was significantly different between land uses for nematodes (F = 3.9, P =
181 0.006) and mites (F = 1.6, P = 0.008), visual inspection of the axes of the principal coordinate
182 indicates that there were clear differences between land uses for nematodes (Supplementary Fig. B2).
183 Land use explained 11.8%, 13.9% and 12.8% of the variation in bacteria, fungi and mycorrhiza
184 composition, respectively. In contrast, land use explained almost three times as much of the variation
185 (31.2%) in nematode composition (Fig. 1) in comparison to that of the microbial taxa. The same
186 pattern was present across the different taxa when only agricultural sites (arable and pasture) were
187 included in the analyses, except that the percentage sum of squares explained by land use was lower,
188 and there were no differences in the composition of any taxa between soil types (data not shown).
189
190 *3.3 Congruency between soil taxa groups*
191 Consistent correlations between particular taxa across the different measures were evident for bacteria
192 and earthworms, fungi and nematodes, fungi and earthworms, and nematodes and earthworms
193 (Supplementary Table B1). The only significant correlations in the abundance of soil taxa were
194 between bacteria and earthworms (Fig. 2A), and nematodes and earthworms (Fig. 2B), being
195 negatively and positively correlated, respectively. There were significant positive correlations in taxon
196 richness between fungi and earthworms (Fig. 2C), and between nematodes and earthworms (Fig. 2D).
197 Conversely, there were significant negative correlations between nematodes and earthworms, and ants
198 (Supplementary Table B1). Positive correlations in composition (Bray-Curtis similarity) were highly
199 significant for fungi and nematodes, and, as with taxon richness, for fungi and earthworms (Fig. 2E),
200 and nematodes and earthworms (Fig. 2F).
201
202 *3.4 Potential indicator taxa across land uses*
203 IndVal analyses identified 14, 10, 22, 34 and 61 significant indicators for arable, pasture, forest,
204 rough-grazing and bogland, respectively (Table 2). Bacteria, AMF and ants had no indicators of

arable and pasture and their greatest number of indicators in bogland, fungi and mites had indicator taxa in four land uses and their greatest number in rough-grazing; nematodes had indicators in all land uses except the forest land use, earthworms had indicators in pasture (Table 2). Interestingly, analysis using only arable and pasture sites resulted in far greater significant results for bacteria and fungi, being 15 and 11 respectively for bacteria, and 20 and 1 for fungi respectively (Table 2). However, it is noted that the percentage of significant taxa in bacteria and fungi was not greater than would be expected by chance at $P = 0.05$.

3.5 Indicator taxa across environmental gradients

Indicator taxa were correlated with several physico-chemical soil properties characteristic of the different land uses (Fig. 3 and 4). Including all land uses, 28% and 20% of variation in species-environment relation was explained by axes 1 and 2, respectively (Table 3). Microbial indicator taxa (bacteria, fungi, mycorrhiza) were more generally associated with boglands, whereas nematodes and earthworms indicator taxa were more strongly associated with arable and pasture (Fig. 3; colour version in Supplementary Fig. B1). Mean bulk density significantly correlated ($F = 4.31$, $P < 0.001$) with the indicator taxa data, being typically lower in the rough-grazing and bogland (extensive land uses) compared to arable (intensive land use). In addition, Fe and Al significantly correlated with the indicator data ($F = 2.24$, $P = 0.015$ and $F = 2.37$, $P = 0.007$, respectively). Al and pH showed a similar correlation, albeit pH was not significant.

When only arable and pasture (intensively managed land) were included, 37% and 22% of variation in species-environment relation was explained by axes 1 and 2, respectively (Table 3). Again, microbial indicator taxa (bacteria and fungi) were associated together, with arable land use in this case, and earthworm indicators associated with pasture (Fig. 4). Two mite indicator taxa were also associated with a small outlier group of pasture sites which appeared to have high concentrations of Ca and P (Fig. 4; Supplementary Fig. B2). With only arable and pasture sites, mean bulk density was also significantly correlated ($F = 1.96$, $P = 0.043$) with the species data, being lower in the arable than the pasture soils (Fig. 4). Al was significantly correlated with the indicator taxa data ($F = 2.13$, P

= 0.040) with the greatest concentration in the opposite direction to the pasture outlier group (Fig. 4), and N correlated significantly with the indicator taxa data (F = 3.06, P = 0.002) being higher in the pasture soils.

4. Discussion

McGeoch [10] discussed different types of biological indicators including those that are typical of a habitat or ecological status and those that are representative of the diversity of other taxa. Here, we have explored these categories of indicator in the soil using a national baseline survey of a range of different taxa groups (e.g. microbes, micro-, meso- and macrofauna).

The potential value of these different taxa as indicators of habitat or ecological status was first gauged by examining their richness and composition across sites, and assessing whether a significant amount of variation could be explained by land use. Land use appeared to have a stronger influence on the richness of soil fauna (nematodes, mites, earthworms and ants) compared to microbes (bacteria, fungi, mycorrhiza). It has been suggested that microbes do not respond to large-scale environmental gradients as do meso- and macrofauna [29]. Therefore, it is likely that specific management practices such as crop types within a land use had a stronger relationship with microbial diversity [30,31]. Although, within arable and pasture sites soil type did not influence richness of any soil taxa. Changes in richness of faunal groups were generally evident between agriculturally-managed (arable and pasture) and extensively-managed (rough-grazing and bogland) soils, and this corresponded to a division between ‘mineral’ and ‘organic’ soils. Greater nematode and earthworm richness was associated with arable and pasture, and greater mite and ant richness was associated with rough-grazing and bogland. This is similar to findings by Rutgers *et al*. [3] from a national soil monitoring scheme in different habitats in the Netherlands with generally greater abundance and richness of nematodes and earthworms in dairy systems. A similar pattern was also evident when examining taxon composition with land use accounting for a lower proportion of variation in microbial taxa groups and soil type having no effect within arable and pasture. Although broad differences in soil communities are greatly appreciated [1,3,8,9,29] it is less well understood how particular taxa, within

these broad groups, may respond to soil environmental gradients and contribute to patterns across these land uses.

A second approach to examining these different taxa as potential indicators of habitat or ecological status was based upon the fidelity and specificity of individual taxa to the different land uses [24,25,32]. A comparison of the taxa identified in this way showed that generally greater numbers of microbial taxa were indicators of the extensive land uses (forest, rough-grazing and bogland) and almost none were characteristic of intensive land uses (arable and pasture). However, when using only arable and pasture in the analysis, many microbial taxa appear as indicators of these land uses. This implies that the microbial indicator taxa found associated with intensive land uses are also found in extensive land uses. Nematodes had indicator taxa across intensive and extensive land uses, and this is in agreement with the greatest amount of variation in nematode composition being explained by land use, whereas ant taxa were not generally good indicators and only one indicator taxon for bogland was identified. Though the number of analyses differed between the taxa (because of different numbers of recorded taxa), the indicator values of individual taxa are derived independently of other taxa and therefore this type of analysis is valuable for exploring the pool of potential indicators in different land uses. A wide range of studies have used indicator value analysis to examine invertebrates characteristic of habitats or land management but fewer have attempted to make links to their traits [e.g. 33,34]. A more detailed examination of indicator traits of soil taxa was beyond the scope of this study but could generate more mechanistic insights. Furthermore, indicator taxa may reveal stronger affinities across several land uses [32].

The indicator taxa identified were utilised to reduce the datasets to taxa likely to be important indicators across land uses. O'Neill *et al*. [35] used this type of analysis with a soil micro-invertebrate dataset and found that classification efficiency for vegetation cover decreased only marginally using only the significant indicator morphotaxa. Moreover, the variability explained by the first two axes of a principal components analysis increased when using only the significant indicator taxa compared to the full complement of taxa. [35]. We combined the significant indicators from all taxa groups to explore the correlation of their abundances with soil physico-chemical gradients. The primary axis of

286 variation was generally associated with the change from intensive (arable) through to extensive
287 (bogland) land use; though mean bulk density was the only significant soil characteristic that showed
288 a strong correlation with this axis, it clearly masked the significance of similarly strong relationships
289 with moisture, carbon and nitrogen in the opposite direction. The ordinations also highlighted how
290 individual indicator taxa were related to the main axes of variation and this may be a useful
291 exploratory tool to identify taxa that are most responsive to particular gradients.

292 Studies of cross-taxon congruency from aboveground systems have found inconsistent
293 relationships [13,14,15]. For example, in grasslands, Oertli *et al*. [14] found no significant congruency
294 between taxonomic richness of three insect groups (bees, aculeate wasps and grasshoppers) but
295 significant congruency in community similarity of bees and grasshoppers. Lovell *et al*. [13] reported
296 mostly weak correlations in congruency of richness and compositional similarity of above-ground
297 invertebrates. We may expect that congruency is both more likely and stronger in the soil given the
298 importance of local environmental conditions and the physical nature of soil as a habitat. Indeed, we
299 found consistent correlations between several taxa groups, in particular, positive correlations between
300 fungi, nematodes and earthworms, thus demonstrating that there is a level of congruency across
301 different measures of soil biodiversity. However, congruency between other taxa was limited.
302 Different soil taxa may be more dominant at different times of the year, for example, microbes can
303 show high seasonal variation [36]. The activity of ecosystem engineering organisms such as
304 earthworms can also impact upon other smaller-bodied taxa and these effects should not be ignored in
305 assessing soil biodiversity.

306 It is also acknowledged that the outcomes of these analyses may in part depend on the
307 methods used to measure the richness and composition of the different soil taxa, and these outcomes
308 may change using different methods. For example, the AMF diversity investigated here was assessed
309 using a bait-plant method and this may have limited the richness and composition of taxa being
310 recorded [37]. Furthermore, the difference in 'taxonomic' resolution between molecular and
311 morphological approaches may influence differences between microbial and micro-, meso-, macro-
312 fauna. Nevertheless, these are standard and widespread methods to extract and measure soil

biodiversity and if we are looking for relative measures or fingerprints of soil assemblages, as opposed to an exhaustive cataloguing, then their comparison is informative. Developments in molecular techniques for the analysis of soil biodiversity [e.g. 38,39,40] will undoubtedly become particularly important as the choice of indicators is streamlined, but there is still the need to compare these with 'classic' approaches.

5. Conclusions

There are few soil biodiversity surveys that include the major land uses and a relatively large geographical spread with this range of belowground taxa [e.g. 3]. Characterising the richness and composition of different soil taxa groups and identifying potential indicators across land uses indicates that separate sets of taxa groups may be more useful as bioindicators in agriculturally and extensively managed land. The facts that land use accounted for the greatest amount of variation in nematode composition and that nematodes were indicator taxa in most land uses supports their potential as robust indicators across all land uses. Analysis of significant indicators can also help identify potential target taxa that are responsive to soil physico-chemical gradients and upon which future sampling could be focused. Further development of these types of analyses can inform soil monitoring programmes and increase their efficacy in being able to detect the effects of land management changes on soil status and the many ecosystem services supported by soil organisms.

Acknowledgements

We thank two anonymous reviewers whose comments greatly improved the manuscript. This study was funded by the Environmental ERDTI Programme 2000–2006, financed by the Irish Government under the National Development Plan and administered on behalf of the Department of Environment and Local Government by the Environmental Protection Agency ("CréBeo: Baseline data, response to pressures, functions and conservation of keystone micro- and macro-organisms in Irish soils", 2005-S-LS-8). We acknowledge the guidance and advice from the Steering Committee members, Dr. Alice Wemaere, Prof. Colin Campbell, Prof. Peter Loveland and Dr. John Scullion.

340

341 **References**

342 [1] H. I. J. Black, N. R. Parekh, J. S. Chaplow, F. Monson, J. Watkins, R. Creamer, E. D. Potter, J. M.
343 Potter, P. Rowland, G. Ainsworth, M. Hornung, Assessing soil biodiversity across Great Britain:
344 national trends in the occurrence of heterotrophic bacteria and invertebrates in soil, J. Environ.
345 Manage. 67 (2003) 255-266.
346 [2] L. Ranjard, S. Dequiedt, C. Jolivet, N. P. A. Saby, J. Thioulouse, J. Harmand, P. Loisel, A.
347 Rapaport, S. Fall, P. Simonet, R. Joffre, N. C.-P. Bouré, N., P.-A. Maron, C. Mougel, M. P. Martin, B.
348 Toutain, D. Arrouays, P. Lemanceau, Biogeography of soil microbial communities: A review and a
349 description of the ongoing French national initiative, Agron. Sustain. Dev. 30 (2010) 359–365.
350 [3] M. Rutgers, A. J. Schouten, J. Bloem, N. van Eekeren, R. G. M. de Goede, G. A. J. M. Jagers op
351 Akkerhuis, A. van der Wal, C. Mulder, L. Brussard, A. M. Breure, Biological measurements in a
352 nationwide soil monitoring network, Eur. J. Soil Sci. 60 (2009) 820-832.
353 [4] European Commission, Thematic Strategy for Soil Protection, Communication (COM(2006) 231)
354 final, Brussels, 2006.
355 [5] O. Andrén, R. Baritz, C. Brandao, T. Breure, I. Feix, U, Franko, A. Gronlund, J. Leifeld, S. Maly,
356 Volume III – Organic Matter and Biodiversity, Taskgroup 3 Soil biodiversity, in: L. Van-Camp, B.
357 Bujarrabal, A.-R. Gentile, R. J. A. Jones, L. Montanarella, C. Olazabal, S.-K. Selvaradjou (Eds.),
358 Reports of the Technical Working Groups Established under the Thematic Strategy for Soil
359 Protection, EUR 21319 EN/3, Office for Official Publications of the European Communities,
360 Luxembourg, 2004, pp. 357-397.
361 [6] X. Morvan, N. P. A. Saby, D. Arrouays, C. Le Bas, R. J. A. Jones, F. G. A. Verheijen, P. H.
362 Bellamy, M. Stephens, M. G. Kibblethwaite, Soil monitoring in Europe: A review of existing systems
363 and requirements for harmonisation, Sci. Total Environ. 391 (2008) 1-12.
364 [7] J. Römbke, J. P. Sousa, T. Schouten, F. Riepert, Monitoring of soil organisms: A set of
365 standardized field methods proposed by ISO, Eur. J. Soil Biol. 42 (2006) S61-S64.

366 [8] A. Turbé, A. De Toni, P. Benito, P. Lavelle, P. Lavelle, N. Ruiz, W. H. van der Putten, E.
367 Labouze, S. Mudgal, Soil biodiversity: functions, threats and tool for policy makers, Bio Intelligence
368 Service, IRD, and NIOO, Report for European Commission (DG Environment), 2010.
369 [9] K. Ritz, H. I. J. Black, C. D. Campbell, J. A. Harris, C. Wood, Selecting biological indicators for
370 monitoring soils: A framework for balancing scientific and technical opinion to assist policy
371 development, Ecol. Indicators 9 (2009) 1212-1221.
372 [10] M. A. McGeoch, The selection, testing and application of terrestrial insects as bioindicators,
373 Biol. Rev. 73 (1998) 181-201.
374 [11] D. T. Bilton, L. McAbendroth, A. Bedford, P. M. Ramsay, How wide to cast the net? Cross-
375 taxon congruence of species richness, community similarity and indicator taxa in ponds, Freshw. Biol.
376 51 (2006) 578-590.
377 [12] J. Heino, Are indicator groups and cross-taxon congruence useful for predicting biodiversity in
378 aquatic ecosystems? Ecol. Indicators 10 (2010) 112-117.
379 [13] S. Lovell, M. Hamer, R. Slotow, D. Herbert, Assessment of congruency across invertebrate taxa
380 and taxonomic levels to identify potential surrogates, Biol. Conserv. 139 (2007) 113-125.
381 [14] S. Oertli, A. Müller, D. Steiner, A. Breitenstein, S. Dorn, Cross-taxon congruence of species
382 diversity and community similarity among three insect taxa in a mosaic landscape, Biol. Conserv. 126
383 (2005) 195-205.
384 [15] N. Sauberer, K. P. Zulka, M. Abensperg-Traun, H-M. Berg, G. Bieringer, N. Milasowszky, D.
385 Moser, C. Plutzar, M. Pollheimer, C. Storch, R. Tröstl, H. Zechmeister, G. Grabherr, Surrogate taxa
386 for biodiversity in agricultural landscapes of eastern Austria, Biol. Conserv. 117 (2004) 181-190.
387 [16] D. Fay, D. McGrath, C. Zhang, C. Carrigg, V. O'Flaherty, G. Kramers, O. T. Carton, E.
388 Grennan, Towards A National Soil Database, Synthesis Report 2001-CD/S2-M2, Environmental
389 Protection Agency, Johnstown Castle, Wexford, Ireland, 2007.
390 [17] M. J. Gardiner, T. Radford, Soil associations of Ireland and their land use potential – Explanatory
391 bulletin to soil map of Ireland, The Agricultural Institute, Dublin, 1980.

392 [18] G. Kiely, N. M. McGoff, J. M. Eaton, X. Xu, P. Leahy, O. Carton, SoilC: Measurement and
393 modelling of soil carbon stocks and stock changes in Irish soils, STRIVE Report 2005-S-MS-26-M1,
394 Environmental Protection Agency, Johnstown Castle, Wexford, Ireland, 2009.
395 [19] R Development Core Team, R: A language and environment for statistical computing, R
396 Foundation for Statistical Computing, 2010.
397 [20] J. Oksanen, G. Blanchet, R. Kindt, P. Legendre, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H.
398 H. Stevens, H, Wagner, vegan: Community Ecology Package, R package version 1.17-2, 2010.
399 [21] Anderson, M. J. A new method for non-parametric multivariate analysis of variance. Austral
400 Ecol. 26 (2001) 32-46.
401 [22] Anderson, M. J. Distance-based tests for homogeneity of multivariate dispersions. Biometrics 62
402 (2006) 245-253.
403 [23] F. E. Harrell Jr, contributions from many other users, Hmisc: Harrell miscellaneous, R package
404 version 3.7-0, 2009.
405 [24] M. Dufrene, P. Legendre, Species assemblages and indicator species: The need for a flexible
406 asymmetrical approach, Ecol. Monogr. 67 (1997) 345-366.
407 [25] M. De Cáceres, P. Legendre, Associations between species and groups of sites: indices and
408 statistical inferences, Ecology 90 (2009) 3566-3574.
409 [26] B. H. McArdle, M. J. Anderson, Fitting multivariate models to community data: A comment on
410 distance-based redundancy analysis, Ecology 82 (2001) 290-297.
411 [27] M. W. Palmer, Putting things in even better order: The advantages of canonical correspondence
412 analysis, Ecology 74 (1993) 2215-2230.
413 [28] C. J. F. ter Braak, P. Šmilauer, CANOCO for Windows version 4.5, Biometris Plant Research
414 International, Wageningen, the Netherlands, 2002.
415 [29] T. Decaëns, Macroecological patterns in soil communities, Global Ecol. Biogeogr. 19 (2010)
416 287-302.

417 [30] M. Hartmann, A. Fliessbach, H.-R. Oberholzer, F. Widmer, Ranking the magnitude of crop and
418 farming system effects on soil microbial biomass and genetic structure of bacterial communities,
419 FEMS Microbiol. Ecol. 57 (2006) 378-388.
420 [31] D. Jordan, R. J. Miles, V. C. Hubbard, T. Lorenz, Effect of management practices and cropping
421 systems on earthworm abundance and microbial activity in Sanborn Field: A 115-year-old agricultural
422 field, Pedobiologia 48 (2004) 99-110.
423 [32] M. De Cáceres, P. Legendre, M. Moretti, Improving indicator species analysis by combining
424 groups of sites, Oikos 119 (2010) 1674-1684.
425 [33] A. Auclerc, J. F. Ponge, S. Barot, F. Dubs, Experimental assessment of habitat preference and
426 dispersal ability of soil springtails, Soil Biol. Biochem. 41 (2009) 1596-1604.
427 [34] M. Moretti, M. De Cáceres, C. Pradella, M. K. Obrist, B. Wermelinger, P. Legendre, P. Duelli,
428 Fire-induced taxonomic and functional changes in saproxylic beetle communities in fire sensitive
429 regions, Ecography 33 (2010) 760-771.
430 [35] K. O'Neill, H. W. Godwin, A. E. Jiménez-Esquilín, J. P. Battigelli, Reducing the dimensionality
431 of soil microinvertebrate community datasets using Indicator Species Analysis: Implications for
432 ecosystem monitoring and soil management, Soil Biol. Biochem. (2010) 145-154.
433 [36] M. Wolsing, A. Priemé, Observation of high seasonal variation in community structure of
434 denitrifying bacteria in arable soil receiving artificial fertilizer and cattle manure by determining T-
435 RFLP of *nir* gene fragments, FEMS Microbiol. Ecol. 48 (2004) 261-271.
436 [37] Z. Sykorova, K. Ineichen, A. Wiemken, D. Redecker, The cultivation bias: Different
437 communities of arbuscular mycorrhizal fungi detected in roots from the field, from bait plants
438 transplanted to the field, and from a greenhouse trap experiment, Mycorrhiza 18 (2007) 1-14.
439 [38] X. Y. Chen, T. J. Daniell, R. Neilson, V. O'Flaherty, B. S. Griffiths, A comparison of molecular
440 methods for monitoring soil nematodes and their use as biological indicators, Eur. J. Soil Biol. 42
441 (2010) 319-324.

442 [39] C. L. Lauber, M. Hamady, R. Knight, N. Fierer, Pyrosequencing-based assessment of soil pH as
443 a predictor of soil bacterial community structure at the continental scale. Appl. Environ. Microbiol. 75
444 (2009) 5111-5120.
445 [40] V. Torsvik, L. Øvreås, Microbial diversity and function in soil: From genes to ecosystems. Curr.
446 Opin. Microbiol. 5 (2002) 240-245.
447
448

449 **Figure captions**

450 Fig. 1 - NMDS ordination of nematode composition across different land uses. Stress value = 0.18.
451 Each datapoint represents an individual site.

452

453 Fig. 2 - Examples of the strongest cross-taxon correlations between abundance (A and B), richness (C
454 and D) and composition (E and F) of soil taxa groups. For abundance and richness each point
455 represents an individual site; for composition each point is a pairwise similarity between two sites.
456 Spearman Rho coefficient inset; all correlations are significance at $P < 0.05$ after correction for
457 multiple comparisons.

458

459 Fig. 3 - Redundancy analyses (RDA) of taxa identified as indicators using IndVal and soil physic-
460 chemical variables across all land uses. Ellipses represent 95% confidence intervals of land uses using
461 site scores from axes 1 and 2. Arrows indicate gradients of soil physico-chemical variables; asterisks
462 denote variables significantly correlated with RDA axes.

463

464 Fig. 4 - Redundancy analyses (RDA) of taxa identified as indicators using IndVal and soil physic-
465 chemical variables across agricultural land uses (Arable and pasture only). Legend as in Fig. 3.
466 Ellipses represent 95% confidence intervals of land uses using site scores from axes 1 and 2. Arrows
467 indicate gradients of soil physico-chemical variables; asterisks denote variables significantly
468 correlated with RDA axes.

469

Table 1. Summary of taxa richness in the CréBeo baseline survey; minimum and maximum taxa richness recorded at a site, and the associated land use where these were recorded, mean taxa richness recorded within each land use and results of non-parametric Kruskal-Wallis ($\chi 2$) tests of the effect of land use on taxa richness. 'All sites' includes every site where the specific group of soil taxa were sampled; analyses of 'Shared sites' include only those sites where all soil taxa were sampled. Values are rounded to nearest integer for clarity. Significance: * = P<0.05; ** =P<0.01; *** =P<0.001. AMF = Arbuscular mycorrhizal fungi.

Soil organisms					**Land-use type**					**Kruskal-Wallis ($\chi 2$)**	
	Min. richness and associated land use		Max. richness and associated land use		Arable (A)	Pasture (P)	Forest (F)	Rough-grazing (RG)	Bogland (B)	All sites	Shared sites
Bacteria	24	A	356	B	160	200	184	187	216	2.55	2.76
Fungi	6	A	159	F	89	78	64	62	31	8.13	9.30
AMF	2	A	78	P	25	41	34	33	42	4.87	4.36
Nematodes	5	B	25	P, RG	18	19	17	17	12	19.23***	9.53*
Mites	0	A,B	27	RG	3	9	14	15	3	20.21***	11.28*
Earthworms	0	F, RG, B	11	P	6	7	4	3	0	30.31***	14.24**
Ants	0	all	5	RG	0	1	1	2	2	18.98***	13.49**

475 **Table 2.** Numbers of taxa identified by the 'IndVal' analyses as indicators of different land uses in the
476 different soil taxa groups. Indicators are significant at $P < 0.05$; % of significant taxa is calculated
477 within each group. Values in parentheses are numbers of indicator taxa identified in analysis of only
478 arable and pasture land uses. AMF = Arbuscular mycorrhizal fungi.

Soil organisms	**Land-use type**					
	Arable	Pasture	Forest	Rough-grazing	Bogland	% of significant taxa
Bacteria	0 (15)	0 (11)	13	11	41	5.7
Fungi	3 (20)	0 (1)	4	9	4	2.3
AMF	0 (0)	0 (2)	0	3	13	8.9
Nematodes	6 (1)	5 (4)	0	4	2	17.7
Mites	5 (1)	0 (2)	5	7	0	11.2
Earthworms	0 (1)	5 (3)	0	0	0	26.3
Ants	0 (0)	0 (0)	0	0	1	12.5

479

480

Table 3. Summary statistics from Redundancy Analyses (RDA) of taxa identified as indicators by indicator species analysis and soil physico-chemical variables.

RDA statistics	**All land uses**			**Arable + Pasture**		
	axis 1	axis 2	All axes	axis 1	axis 2	All axes
Eigenvalue	0.173	0.125		0.258	0.148	
Species-environment correlation	0.913	0.891		0.973	0.909	
Species-environment variation (Cumulative %)	27.8	47.9	62.0	37.4	58.9	69.5

Figure(s)

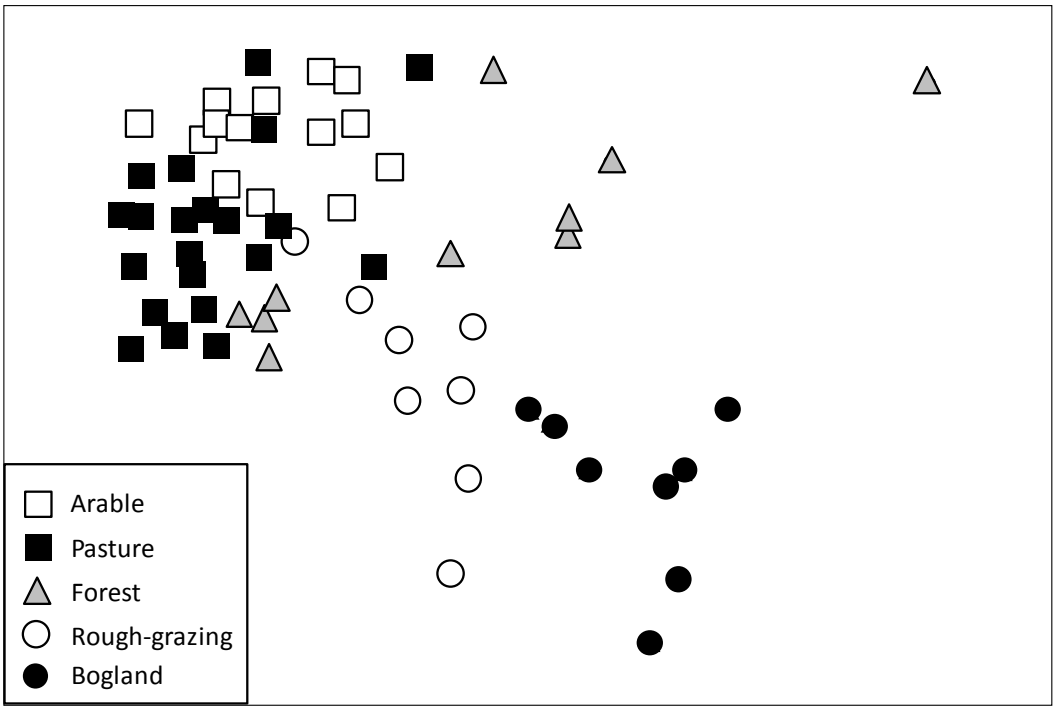

Figure 1

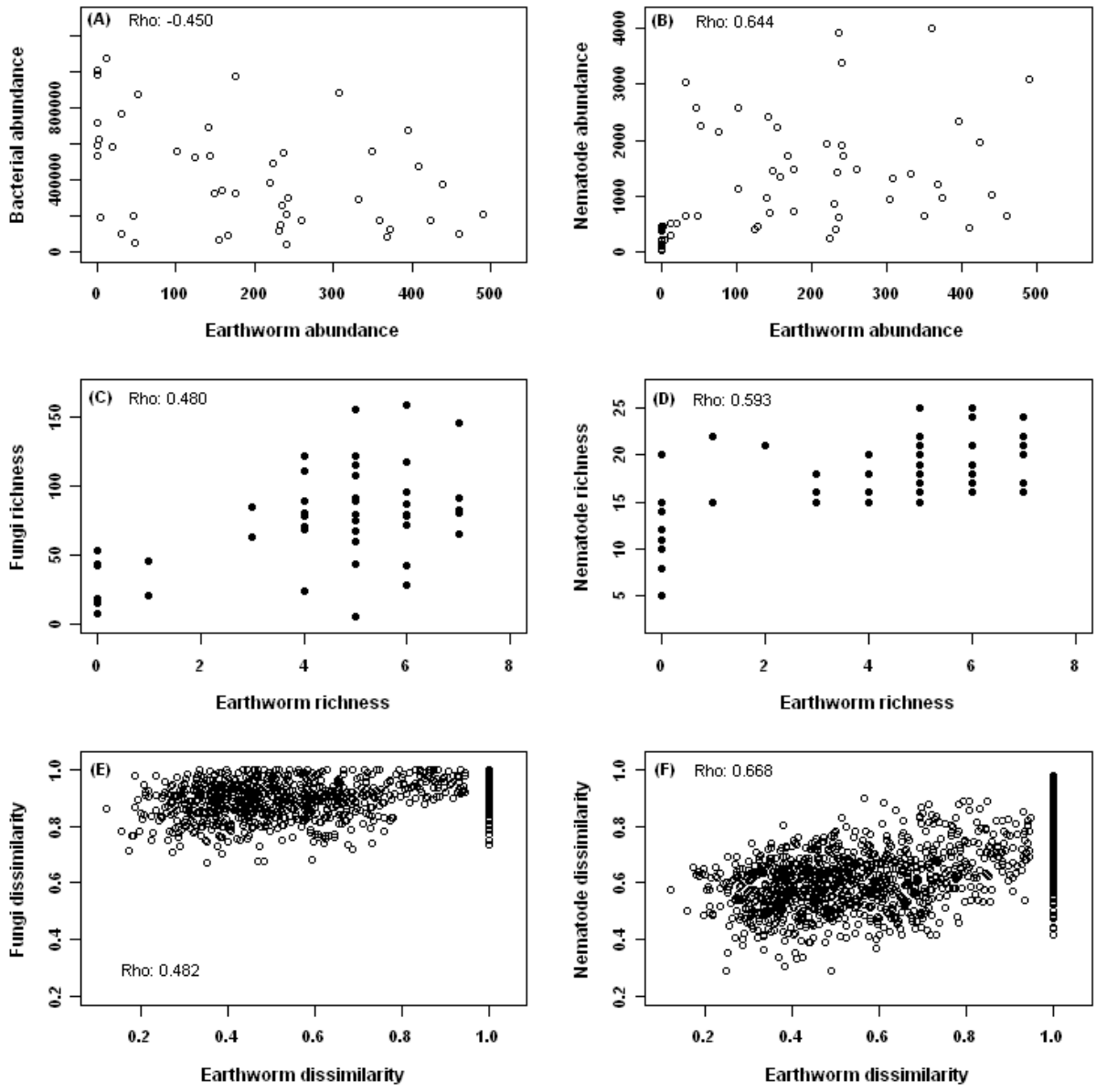
(A) Rho: -0.450
Bacterial abundance
Earthworm abundance
(B) Rho: 0.644
Nematode abundance
Earthworm abundance
(C) Rho: 0.480
Fungi richness
Earthworm richness
(D) Rho: 0.593
Nematode richness
Earthworm richness
(E)
Rho: 0.482
Fungi dissimilarity
Earthworm dissimilarity
(F) Rho: 0.668
Nematode dissimilarity
Earthworm dissimilarity

Figure 2

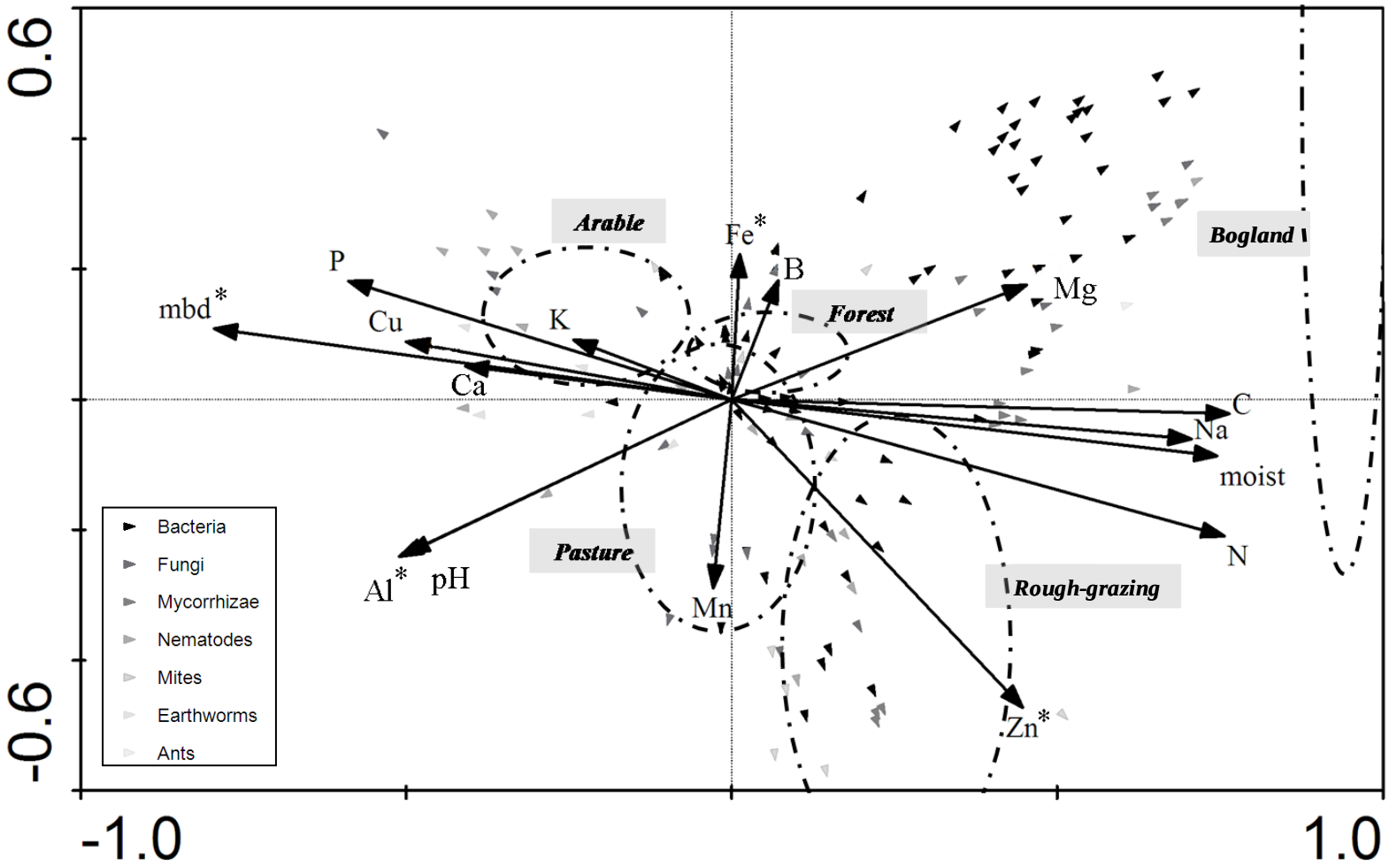
0.6
-0.6
-1.0
1.0
Arable
Forest
Bogland
Pasture
Rough-grazing
mbd*
P
Cu
K
Ca
Fe*
B
Mg
C
Na
moist
N
Al*
pH
Mn
Zn*
Bacteria
Fungi
Mycorrhizae
Nematodes
Mites
Earthworms
Ants

Figure 3

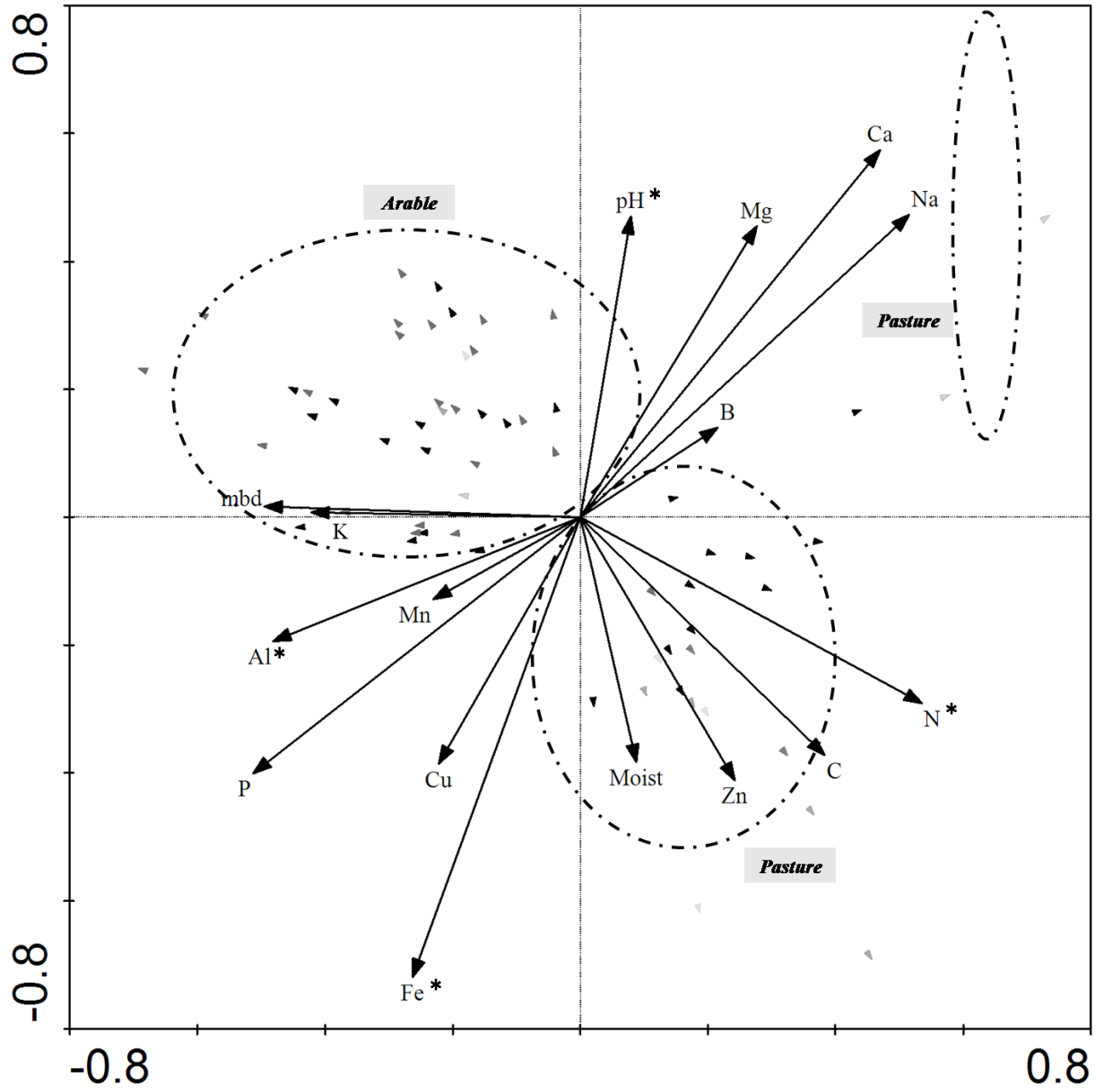
0.8
-0.8
-0.8
0.8
Arable
Pasture
Pasture
Ca
pH *
Mg
Na
B
mbd
K
Mn
Al*
P
Cu
Moist
Zn
C
N *
Fe *

Figure 4

Supplementary A. Detailed materials and methods for the sampling and processing of the different soil organism groups and a map of sampling locations classified by land use.

Soil bacteria and fungi

Soil samples were taken randomly from each GPS-located plot with a sterilised corer to a depth of 20 cm. From each plot, 20 cores were collected and bulked. Upon arrival in the laboratory, soil samples were immediately passed though a 4 mm aperture sieve and stored at –20ºC for DNA extraction and a sub-sample was preserved to determine soil moisture content at the time of sampling.

DNA was extracted with a modified method as described by Griffiths *et al*. (2000). Briefly, this involved a 0.5 g soil sub-sample in hexadecyltrimethylammonium bromide (CTAB) extraction buffer subjected to a heat treatment of 10 minutes at 70°C, subsequent physical cell lysis with a Ribolyser bead beater, while DNA was separated in a 25:24:1 phenol:chloroform:isoamylalcohol solution, followed with a clean-up with 24:1 chloroform:isoamylalcohol to remove impurities. The aqueous layer was removed and DNA was precipitated in 1 ml 95% ethanol after addition of 60 µl 3 M sodiumacetate and 1 µl glycogen and overnight incubation at –20°C before clean up with a high pure PCR product purification kit (Roche, Germany). Purified DNA, eluted to a final volume of 50 µl, was quantified on a spectrophotometer (Nanodrop) and diluted to 3–50 ng μl^{-1} suitable for PCR amplification without further treatment. Each extraction was replicated three times. Bacterial DNA was amplified using primers targeted on the intergenic spacer region (IGS) using the bacterial rRNA operon and amplified with the universal bacterial forward primer S-D-Bact-1522-b-S-20 (eubacterial rRNA small subunit, 5'-TGC GGC TGG ATC CCC TCC TT-3') and reverse primer L-D-Bact-132-a-A-18 (eubacterial rRNA large subunit 5'-CCG GGT TTC CCC ATT CGG-3') (Normand *et al.*, 1996). Fungal DNA was amplified using primers targeted on the fungal intergenic spacer region containing two internal transcribed spacers (ITS) and the 5.8S rRNA gene (ITS1-5.8S-ITS2) using universal fungal forward primer (ITS1-F) 5'-CTT GGT CAT TTA GAG GAA GTA A-3' (Gardes and Bruns, 1993) and reverse (ITS4) 5'-TCC TCC GCT TAT TGA TAT GC-3' (White *et al.*, 1990)

Each PCR reaction was done in 50 µl volumes, containing 10 µl 10X PCR buffer, 5 µl of 0.3

µM forward and reverse primer, 1.25 µl 10 mg ml^{-1} BSA, 1 µl dNTPs (10 mM each), 2.5 µl ultra clean H_2O and 0.25 µl 2.5 U *Taq* DNA polymerase. One µl template DNA was added to 25 µl ultra clean H_2O prior to adding the PCR mix. For bacterial ARISA, PCR conditions included a hot start at 94ºC for 3 min (1 cycle); 94ºC for 45 sec, 61.5ºC for 45 sec, 72ºC for 1 min (34 cycles) with a final annealing temperature at 72ºC for 7 min. DNA extractions of pure culture *E. coli* served as a positive control, while DNA free PCR mix was used as a negative control. For fungal ARISA, PCR conditions included a hot start at 95ºC for 4 min (1 cycle); 95ºC for 1 min, 56ºC for 30 sec, 72ºC for 1 min (35 cycles) with a final annealing temperature at 72ºC for 7 min. DNA extractions of a pure culture of a *Trichoderma* sp. served as a positive control, while DNA free PCR mix was used as a negative control. PCR products were confirmed on a 1% agarose gel and subsequently purified using a high pure PCR product cleanup kit (Roche) as per user manual instructions. Both forward primers were fluorescently labelled on the 5' side with Beckman Coulter dye D4. Products were purified with a high pure PCR product purification kit, and amplified nucleic acid was eluted in 50 µl sterile ultra clean H_2O at 55°C.

Intergenic spacer lengths were analysed using electrophoresis on a Beckman Coulter (CEQ 8000) automated sequencer, running 120 minutes at 60°C and 4 kV. A 20–1200 bp fragment sizing standard with a D2 dye was used to calculate reference curves. Beckman Coulter CEQ 8000 fragment analysis software was used to assess spacer profiles, and to identify peaks which correspond to ribotypes. Individual ribotypes were considered to represent taxa for the calculation of richness and similarity.

Mycorrhizal fungi

Arbuscular mycorrhizal fungi (AMF) were surveyed within forty-five NSD locations in 2006. Field moist soil, obtained as described before, was used for bioassays, with *Trifolium repens* L. (Fabaceae; White clover) as bait plants for AMF. For this, surface-sterilised seeds were sown in pots (8 cm × 8 cm × 8 cm) containing a 1:1 mix of soil and autoclaved sand replicated three times. All pots were then placed randomly into growth chambers and were grown for four months under environmentally

controlled conditions (8 h dark/16 h light cycle, and a constant temperature of 20°C). Negative control pots were grown in autoclaved field soil and sand (1:1 mix). At harvest, all soil was carefully and thoroughly removed from plant roots. Root samples were triple rinsed with sterile, de-ionised water, blotted dry and stored at –80°C for DNA extraction.

Molecular techniques were employed to characterise AMF diversity. Specifically, terminal restriction fragment length polymorphism (TRFLP) analysis was used. DNA was extracted from 100 mg of each sample using the DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) for AMF. A 550 bp region of the 18S rDNA was amplified using the universal eukaryotic primer NS31 (Simon *et al*., 1992) and the AMF specific primer AM1 (Helgason *et al*., 1998). For TRFLP analysis, purified polymerase chain reaction (PCR) products were digested with the restriction enzymes *Hin*fI and *Hsp*92II. Resulting TRFLP profiles were analysed using the program GeneMarker (SoftGenetics, State College, PA, USA). Only terminal restriction fragments with peak heights above 50 fluorescent units and between 75–450 bp in size were considered and used for further analyses.

Nematodes

Field moist soil, obtained as previously described, was mixed thoroughly and 500 cm^3 of soil was stored at 4ºC until extraction. Nematodes were then extracted from a 100 cm^3 sub-sample of soil from each site. This was suspended in water, sieved (through 600, 250, and 38 µm mesh sizes), and retained nematodes were extracted via sugar centrifugation (Southey, 1986). Nematodes were immediately counted under a stereomicroscope to estimate abundance, then killed by application of gentle heat, fixed in hot (65°C) buffered formalin:glycerine (FG 4:1) and stored in 4 ml glass vials. Nematodes were then processed to pure glycerine by slow evaporation and mounted in permanent mass slides for community analysis. Approximately 100 nematodes were identified for each site using Andrássy (1985, 1992, 1993), Bongers (1988) and Siddiqi (2000) to at least genus level (with the exception of Rhabditidae and Neodiplogasteridae).

Earthworms

Earthworms were sampled in the field using hand-sorting and chemical expellant approaches. For hand-sorting, earthworms were sampled from 25 cm × 25 x 25 cm square soil blocks at each of the four cardinal points in the plots (10 m from the GPS point). These soil blocks were placed on a plastic sheet and were sorted thoroughly by hand. Hand-sorting was standardised by limiting sorting time to 15 minutes. Specimens were placed in plastic bottles, kept cool (4°C) until they could be processed. The four sub-samples were kept separate throughout the sorting and identification process. For the chemical expellant four sub-samples were also taken using dilute mustard oil (2 mL allyl isothiocyanate) where feasible. This method stimulates earthworms to leave the soil so they can be collected on the surface. First, vegetation was clipped to ground level with hand shears and a 50 cm × 50 cm frame placed on the soil and pressed in to a depth of 1–2 cm. Then, 2 ml allyl isothiocyanate was dispersed in 40 ml isopropanol [2-propanol], then added to 20 L water and mixed thoroughly and was evenly applied 50 x 50 cm plots, and expelled earthworms were collected with forceps as they emerged. Application of the mustard oil solution was repeated after 10-15 minutes for each of the four sub-samples, adding approximately 5 L solution in total to each frame. Collected worms were placed in plastic jars containing a small amount of water to rinse off the irritant. In the laboratory, each sub-sample of worms was rinsed with tap water, blotted on paper towels and weighed live *en masse* for total biomass. After weighing, worms were fixed in 4% formalin until identification to species level.

Microarthropods

Four cores were taken at each site, one at each of four cardinal points (10 m from the GPS point). Cores were taken to a depth of 5 cm with a serrated coring device (approx. 5 cm diameter). These were placed in sample cups with a mesh screen bottom, and into plastic screw-cap jars for transport to the laboratory. Upon arrival in the laboratory, microarthropods were directly extracted from these for 7 days into 70% ethanol using a Kempson extractor. Mesostigmatid and oribatid mites were separated and identified to species level where possible.

Ants

The sampling sites for soil-dwelling ants represent a subset of the Irish National Soil Database and included 59 sites (Figure A1). At each site a 20 m line of crumb baits was set up at 1 m distances to attract ant species that forage (Agosti *et al.*, 2000). Furthermore, hand sampling within a 100 m radius of the site was conducted to include an active search for ants focussing on possible nesting sites. The time spent on each site was 30–60 min to standardise the method. The ants were collected with an aspirator and were immediately transferred into a vial with 70% alcohol for later identification following Seifert (2007) and Czechowski *et al.* (2002).

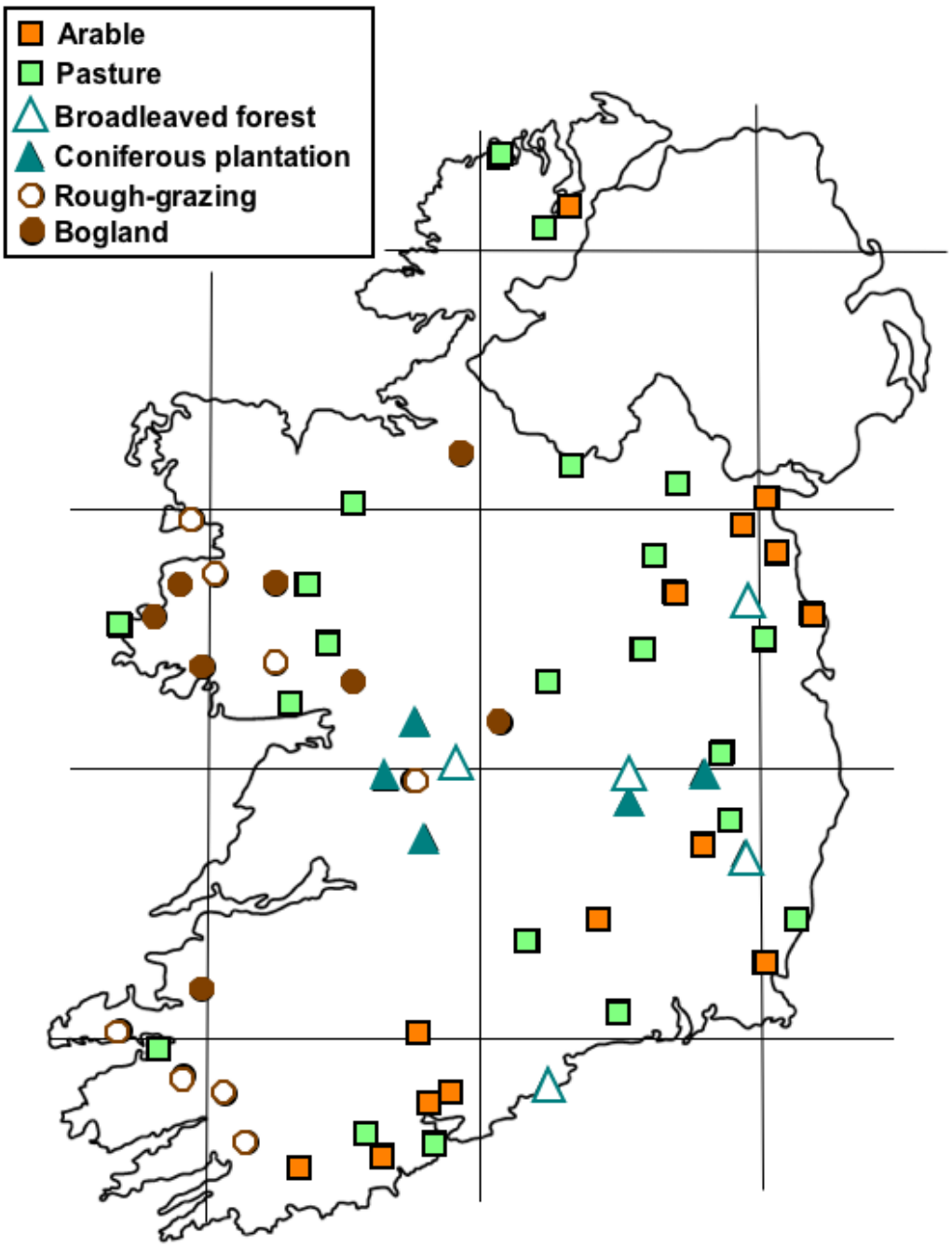

Figure A1. Map of sampling locations from the CréBeo soil biodiversity survey; sites are classified by land use.

References

Agosti, D., Majer, J.U.D., Alonso, L.E. and Schultz, T.R., 2000, *Ants: Standard Methods for Measuring and Monitoring Biological Diversity*. Smithsonian Institute, Washington, USA.

Andrássy, I. (1985) On the genera *Mononchus* Bastian, 1865 and *Prionchulus* (Cobb, 1916) Wu & Hoeppli, 1929 (Nematoda: Mononchidae). Opuscula Zoologica Budapest **21**: 9-22.

Andrássy, I. (1992) A taxonomic survey of the family Mylonchulidae (Nematoda). Opuscula Zoologica Budapest **25**: 11-35.

Andrássy, I. (1993) A taxonomic survey of the family Anatonchidae (Nematoda). Opuscula Zoologica Budapest **26**: 9-25.

Bongers, T., 1988, *De Nematoden van Nederland*. Koninklijke Nederlandse Natuurhistorische Vereniging, Utrecht, the Netherlands.

Czechowski, W., Radchenko, A. and Czechowska, W., 2002, *The Ants (Hymenoptera, Formicidae) of Poland*. Warszawa Museum and Institute of Zoology PAS, Warszawa, Poland.

Gardes, M. and Bruns, T.D., 1993, ITS primers with enhanced specificity for Basidiomycetes: application to the identification of mycorrhiza and rusts. *Molecular Ecology* **2:** 113–118.

Griffiths, R.I., Whiteley, A. S., O'Donnell, A.G. and Bailey, M.J., 2000, Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. *Applied and Environmental Microbiology* **66:** 5488–5491.

Helgason, T., Daniell, T.J., Husband, R., Fitter, A.H. and Young, J.P.W., 1998, Ploughing up the wood-wide web? *Nature* 394**:** 431.

Normand, P., Ponsonnet, C., Nesme, X., Neyra, M. and Simonet, P., 1996, ITS analysis of prokaryotes. *In*: Akkermans, A. D. L., van Elsas, J. D. and De Bruijn, E. I. (Eds.) *Molecular Microbial Ecology Manual*. Kluwer Academic Press, Amsterdam, the Netherlands, pp.1–12.

Seifert, B., 2007, *Die Ameisen Mittel– und Nordeuropas*. lutra Verlags- und Vertriebsgesellschaft, Görlitz/Tauer, Germany.

Siddiqi, M.R. (2000) Tylenchida: Parasites of plants and insects. 2nd Edition. CABI Publishing, UK.

Simon, L., Lalonde, M. and Bruns, T.D., 1992, Specific amplification of 18S fungal ribosomal genes from vesicular-arbuscular endomycorrhizal fungi colonising roots. *Applied and Environmental Microbiology* **58:** 291–295.

Southey, J.F., 1986, *Laboratory Methods for Work with Plant and Soil Nematodes.* MAFF/ADAS, Her Majesty's Stationery Office, London, UK.

White, T.J., Bruns, T.D., Lee, S. and Taylor, J., 1990, Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In*: Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J., (Eds.) *PCR Protocols: a guide to methods and applications*. Academic Press, San Diego, CA, USA, pp. 315–322.

SUPPLEMENTARY B. ADDITIONAL DATA AND COLOUR VERSIONS OF FIGURES.

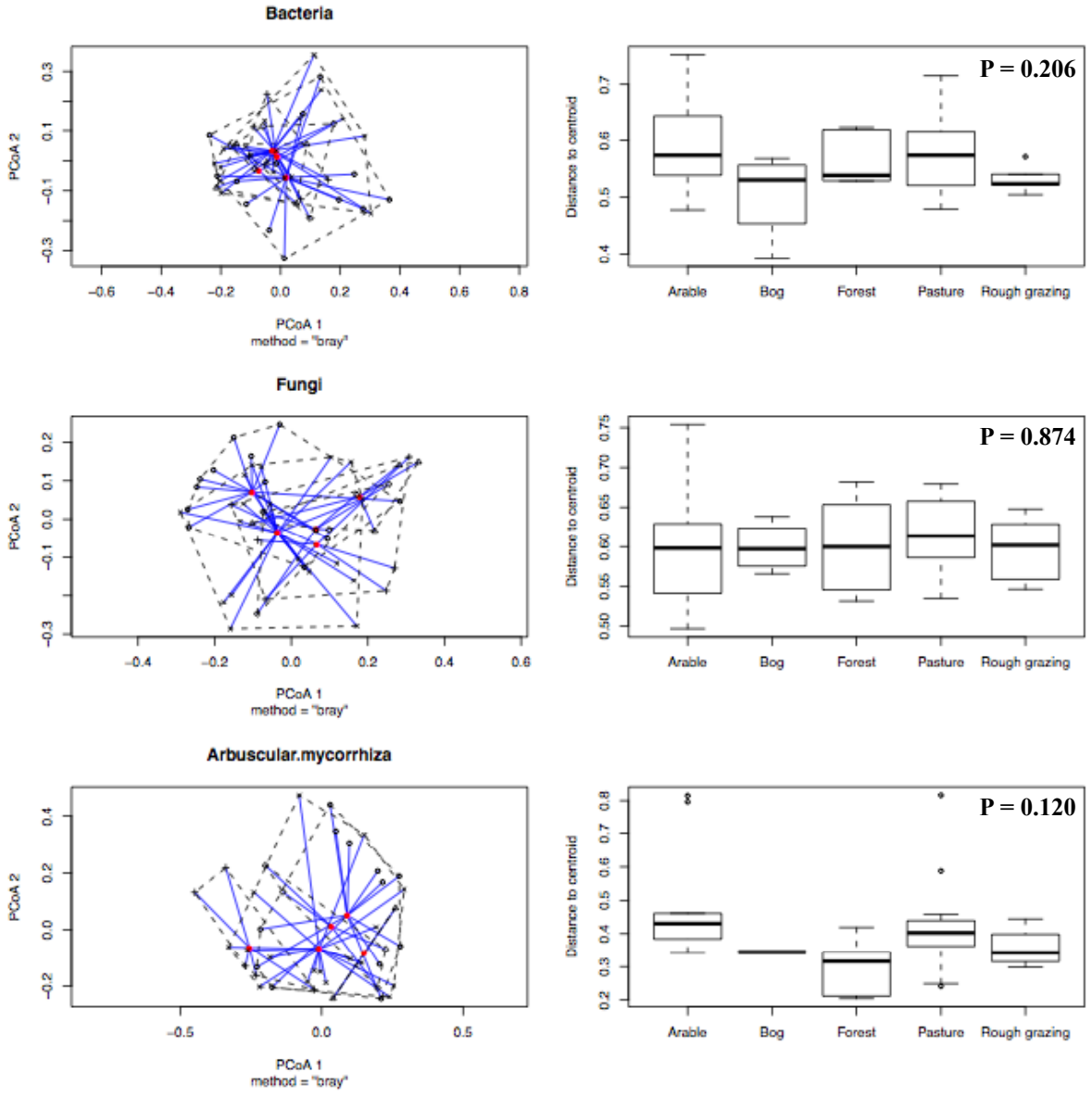

Figure B1. Plots of multivariate dispersion (distance to centroid) of bacteria, fungi and mycorrhiza composition across land uses derived following Anderson (2006). Plots derived using betadisp function in vegan (Oksanen *et al*., 2010).

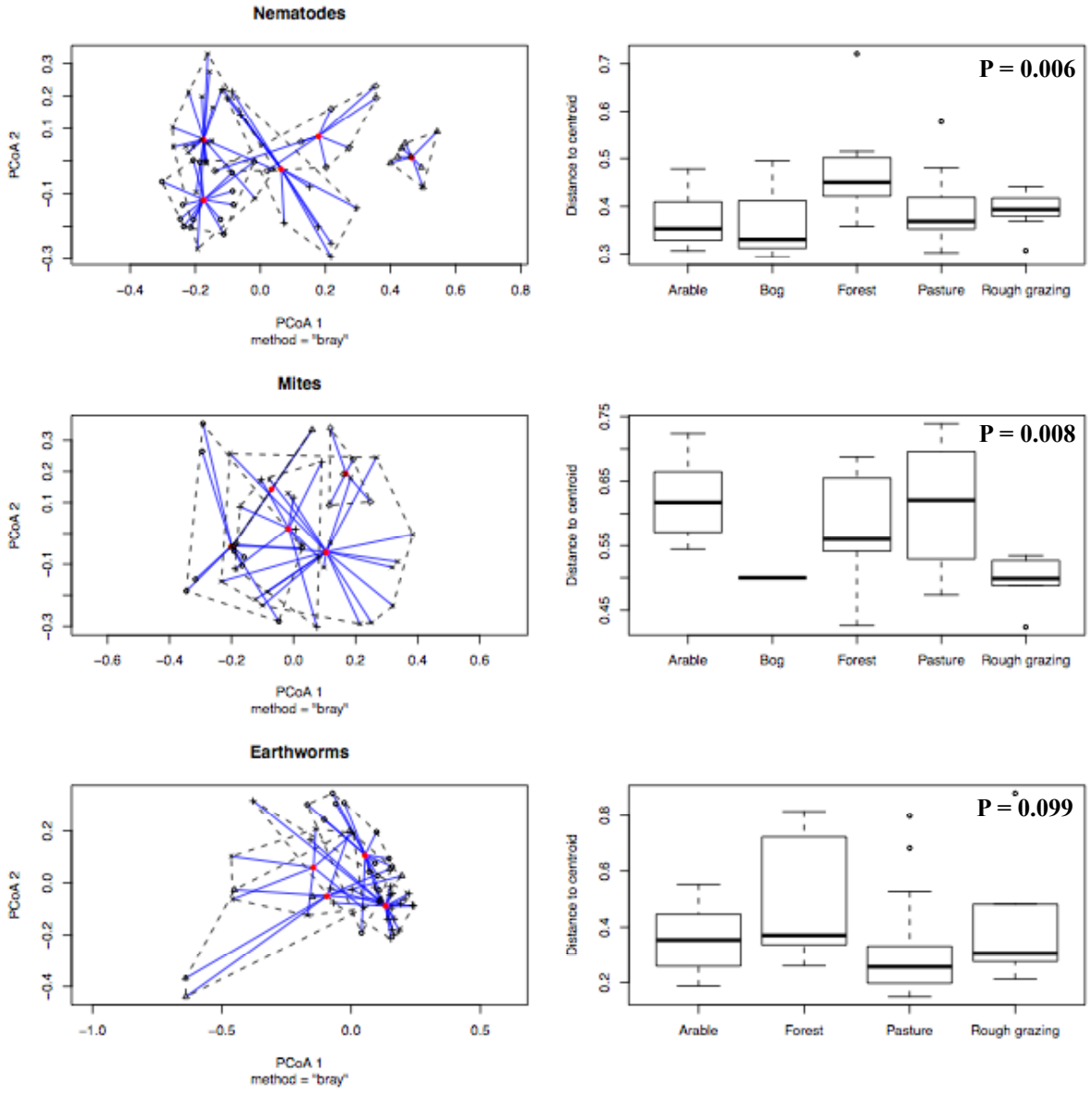

Figure B2. Plots of multivariate dispersion (distance to centroid) of nematode, mite and earthworm composition across land uses derived following Anderson (2006). Plots derived using betadisp function in vegan (Oksanen *et al.*, 2010).

Table B1. Congruence in soil assemblage measures (Pairwise correlations of abundance, richness, Shannon diversity and Bray-Curtis similarity) between groups of taxa across all sites. Bac = Bacteria, Fung = Fungi, Myco = Arbuscular mycorrhizae, Nem = Nematodes, Mite = Acarids, Worm = Earthworms; nd = no data;*= P<0.05 after correction for multiple comparisons following Benjamini and Hochberg (1995).

[a]Spearman rank correlations of raw data, see methods for details.

Taxa comparison	Soil assemblage measure			
	Abundance[a]	Richness[a]	Shannon[a]	Composition[b]
Bac v Fung	-0.198	-0.218	-0.105	-0.056
Bac v Myco	-0.263	-0.082	0.056	-0.079
Bac v Nem	-0.260	-0.029	0.081	-0.012
Bac v Mite	0.110	0.068	0.104	0.023
Bac v Worm	-0.450*	-0.160	-0.335*	-0.079
Bac v Ant	nd	0.197	nd	0.057
Fung v Myco	0.016	-0.040	0.099	-0.109
Fung v Nem	0.067	0.337	0.343*	0.430*
Fung v Mite	-0.232	-0.101	-0.079	0.007
Fung v Worm	0.088	0.480*	0.277	0.482*
Fung v Ant	nd	-0.372	nd	-0.119
Myco v Nem	-0.096	0.037	0.144	0.009
Myco v Mite	0.199	0.246	0.161	0.221
Myco v Worm	-0.025	0.186	0.298	0.006
Myco v Ant	nd	0.301	nd	0.013
Nem v Mite	-0.223	0.017	-0.074	0.145
Nem v Worm	0.644*	0.593*	-0.021	0.668*
Nem v Ant	nd	-0.342*	nd	-0.052
Mite v Worm	-0.049	-0.001	-0.150	0.097
Mite v Ant	nd	0.160	nd	0.012
Worm v Ant	nd	-0.415*	nd	-0.062

[b]Mantel correlation of Bray-Curtis matrices using square-root transformed abundance data.

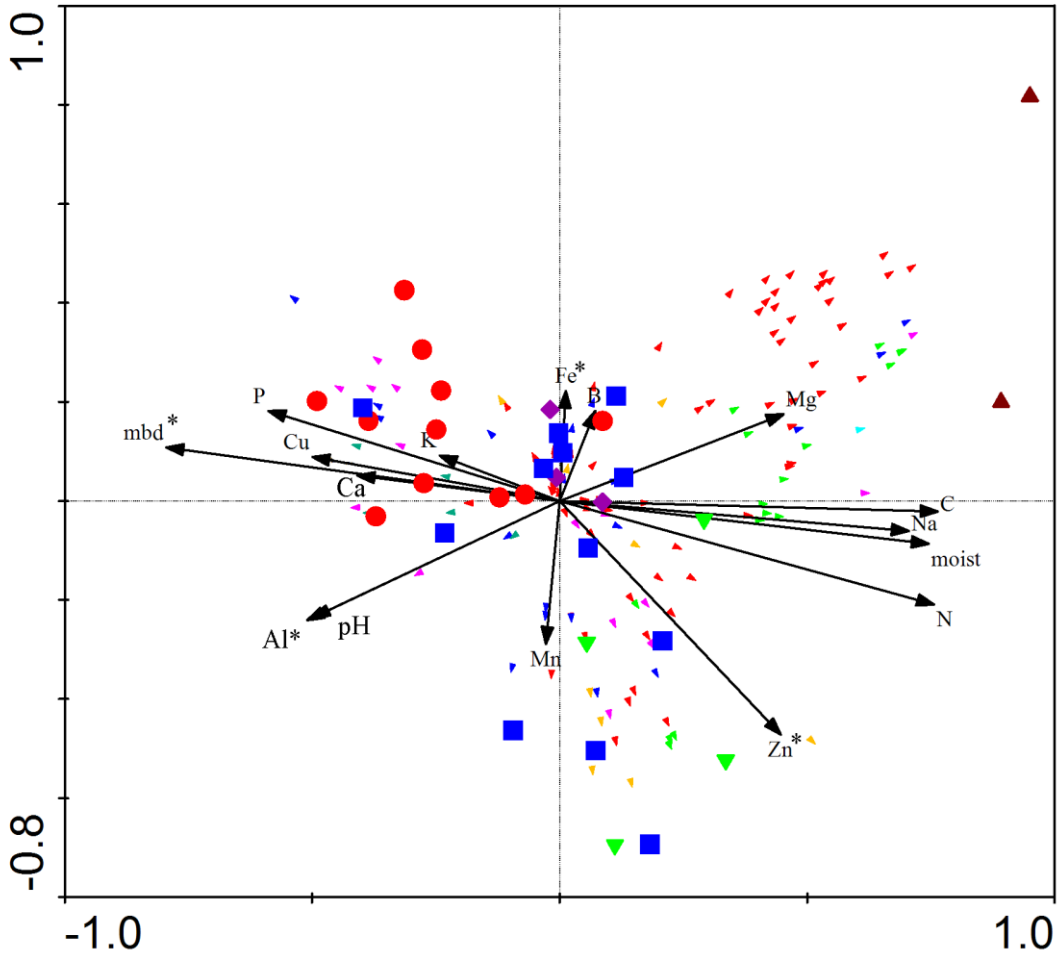

Figure B3. [*Colour version of analysis in Fig. 3*] Redundancy analyses (RDA) of taxa identified as indicators using IndVal and soil physico-chemical variables across all land uses. Arrows indicate gradients of soil physico-chemical variables; asterisks denote variables significantly correlated with RDA axes. Land use: ●= arable; ■= pasture; ♦= forest; ▼= rough-grazing; ▲= bog. Species: ►= bacteria; ►= fungi; ►= mycorrhizae; ►= nematodes; ►= mites; ►= earthworms; ►=ants.

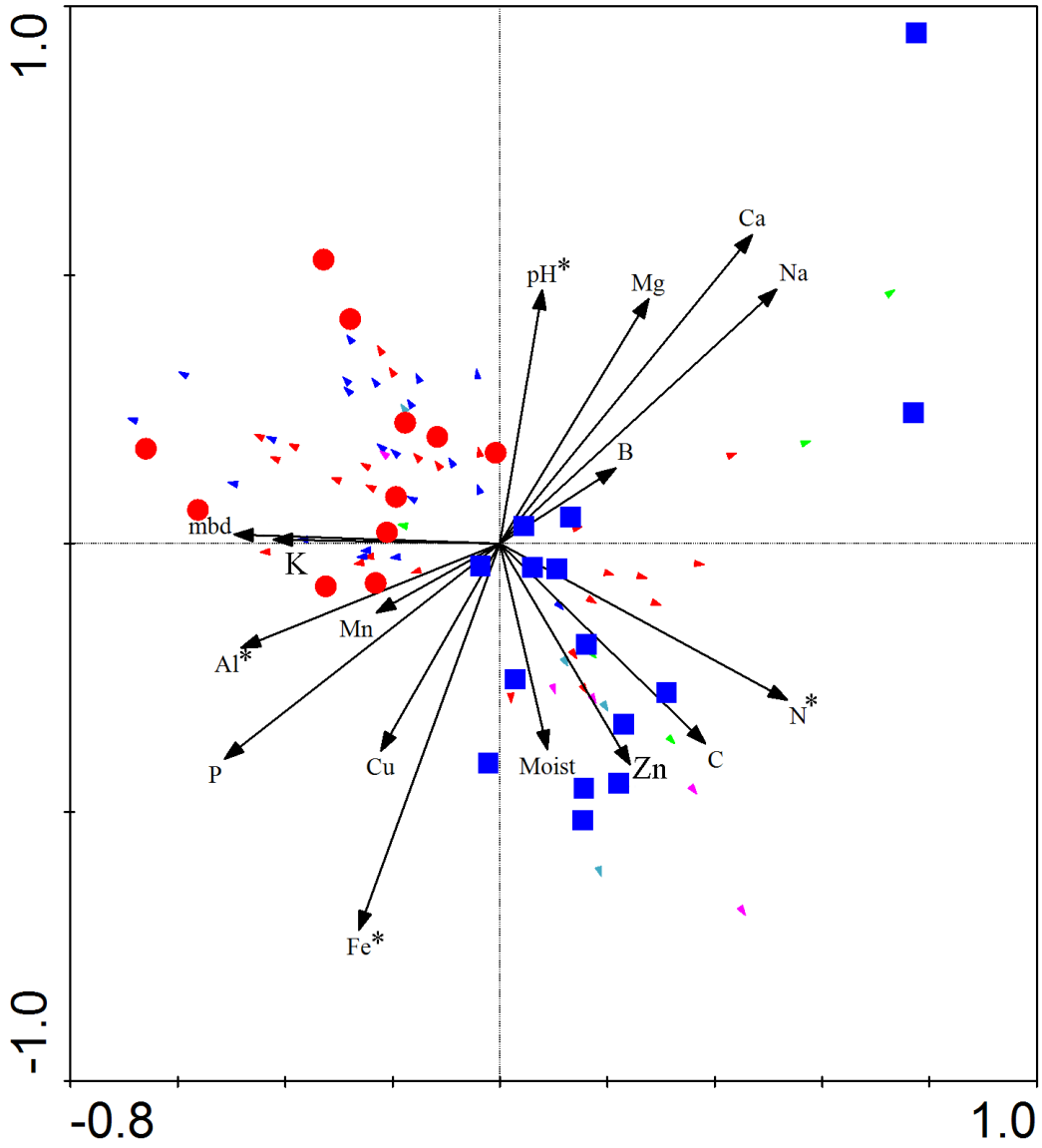

Figure B4. [*Colour version of analysis in Fig. 4*] Redundancy analyses (RDA) of taxa identified as indicators using IndVal and soil physic-chemical variables across agricultural land uses (Arable and pasture only). Arrows indicate gradients of soil physico-chemical variables; asterisks denote variables significantly correlated with RDA axes. P, N, pH and mean bulk density (mbd) explained significant amounts of the variation. Land use: ●= arable; ■= pasture. Species: ▶= bacteria; ▶= fungi; ▶= mycorrhizae; ▶= nematodes; ▶= mites; ▶= earthworms; ▶=ants.

References

Anderson, M. J. (2006) Distance-based tests for homogeneity of multivariate dispersions. Biometrics **62**: 245–253.

Benjamini, Y., Hochberg, Y. (1995) Controlling the False Discovery Rate: A practical and powerful approach to multiple testing, Journal of the Royal Statistical Society. Series B (Methodological) **57:** 289–300.

Oksanen, J., Blanchet, G., Kindt, R., Legendre, P., O'Hara, R. B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H. (2010) *vegan: Community Ecology Package, R package version 1.17-2*, 2010.