Linking species, traits and habitat characteristics of Collembola at European scale
Sandrine Salmon, Jean-François Ponge, Sophie Gachet, Louis Deharveng, Noella Lefebvre, Florian Delabrosse

To cite this version:
Sandrine Salmon, Jean-François Ponge, Sophie Gachet, Louis Deharveng, Noella Lefebvre, et al.. Linking species, traits and habitat characteristics of Collembola at European scale. Soil Biology and Biochemistry, Elsevier, 2014, 75, pp.73-85. 10.1016/j.soilbio.2014.04.002. hal-00983926

HAL Id: hal-00983926
https://hal.archives-ouvertes.fr/hal-00983926
Submitted on 26 Apr 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Linking species, traits and habitat characteristics of Collembola at European scale

Salmon S.¹*, Ponge J.F.¹, Gachet S.², Deharveng, L.³, Lefebvre N.¹, Delabrosse F.¹

¹Muséum National d'Histoire Naturelle, CNRS UMR 7179, 4 avenue du Petit-Château, 91800 Brunoy, France
²Muséum National d'Histoire Naturelle, CNRS UMR 7205, 45 rue Buffon, 75005 Paris, France
³Aix-Marseille Université, Institut Méditerranéen de Biodiversité et d’Écologie Marine et Continentale, CNRS UMR 7263, Campus Saint-Jérôme, Case 421, 13397 Marseille Cedex 20, France

* Corresponding author: Muséum National d'Histoire Naturelle, UMR CNRS 7179, 4 Avenue du Petit-Château, 91800 Brunoy, France
Tel: +33 (0)1 60 47 92 21. E-mail address: ssalmon@mnhn.fr
Abstract

Although much work has been done on factors which influence the patterning of species and species trait assemblages in a variety of groups such as plants, vertebrates and invertebrates, few studies have been realized at a broad geographic scale. We analyzed patterns of relationships between species, species trait distribution/assembly, and environmental variables from the west of Europe to Slovakia, Poland and Sweden. We created a database by compiling traits and occurrence data of European collembolan species, using literature and personal field studies embracing a large range of environmental gradients (vertical stratification, habitat closure, humus form, soil acidity and moisture, temperature, rainfall, altitude) over which Collembola are supposed to be distributed. Occurrences of the 58 best-documented species, environmental variables and species traits allowed us to (i) show which environmental variables impact the distribution of the 58 species at broad scale and (2) document to what extent environmental variables and species trait assemblages are related and which trends could be found in trait/environment relationships. The impact of vertical stratification, habitat closure, humus form, soil acidity, soil moisture, temperature, and to a lesser extent rainfall and altitude on species distribution, firstly revealed by indirect gradient analysis (correspondence analysis, CA), was further shown to be significant by direct gradient analysis (canonical correspondence analysis, CCA). RLQ analyses were performed to find linear combination of variables of table R (environmental variables) and linear combinations of the variables of table Q (species traits) of maximum covariance weighted by species occurrence data contained in table L. RLQ followed by permutation tests showed that all tested environmental variables apparently contributed significantly to the assemblages of the twelve species traits studied. A convergence was observed between traits related to vertical stratification and those related to habitat closure/aperture. Well-developed locomotory organs
(furcula, legs), presence of sensorial organs sensitive to air movements and light (e.g. trichobotria and eye spots), spherical body, large body size, pigmentation (UV protection and signalling) and sexual reproduction largely occur in epigeic and open habitats, while most of woodland and edaphic habitats are characterized by short locomotory appendages, small body size, high number of defense organs (pseudocelli), presence of post-antennal organs and parthenogenesis. Climate and especially temperature exert an effect on the assemblage of traits that are mostly present above-ground and in open habitats. The contribution of combinations of some environmental variables to the occurrence of each species trait was tested by linear, logistic or multinomial regression (Generalized Linear Models). Vertical stratification, followed by temperature, played a dominant role in the variation of the twelve studied traits. Relationships between traits and environment tested here shows that it is possible to use some traits as proxies to identify potential ecological preferences or tolerances of invertebrate species. However, a significant part of species distribution remained unexplained, probably partly because some traits, like ecophysiological ones, or traits involved in biotic interactions (e.g. competition) were unavailable. The present work is thus a first step towards the creation of models predicting changes in collembolan communities. Further studies are required to inform ecophysiological traits, in order to complete such models. Moreover the niche width of species will have to be determined.

**Keywords:** Collembola; environmental filtering; habitats; broad scale distribution; species traits; species assemblages; sensory organs
1. Introduction

Identifying the main factors that drive the composition of communities and the
distribution of species is a fundamental goal in community ecology and is of particular
importance for predicting biodiversity responses to environmental changes (Belyea and
Lancaster, 1999). Selection of species by habitat constraints (deterministic process) is one of
the four classes of processes that influence patterns in the composition and diversity of
species (Vellend, 2010). Functional traits, (named “traits” hereafter), are well-defined,
measurable properties of organisms, used comparatively across species, and that strongly
influence organismal performance (McGill et al., 2006). Focusing on the selection of species
functional traits rather than only on species identity, allows to (1) identify mechanisms
implied in the distribution of species and in the dynamics of biodiversity, (2) understand the
mechanisms that shape communities comprised of many species (3) identify general patterns
and hence, help to predict potential changes in the composition of communities, and
consecutive ecosystem functioning, following disturbance (McGill et al., 2006, Vellend,
2010). The use of functional traits of species allowed to understand species responses to
varied disturbances such as fragmentation, land use change or agricultural practices (Cole et
al., 2002; Barbaro and van Halder, 2009; Ozinga et al. 2009, Vandewalle et al., 2010). For
example, Ozinga et al. (2009) showed that differences between plant species in characteristics
(traits) involved in dispersal processes contribute significantly to explaining losses in plant
diversity in response to habitat degradation.

Species traits of diverse communities (plants, carabids, butterflies, birds, spiders) also
have been shown to vary with environmental factors such as habitat fragmentation (Barbaro
and van Halder, 2009), presence of planted hedgerows in highway verges (Le Viol et al.,
2008), post-fire age (Langlands et al., 2011), salinity (Pavoine et al., 2011), agricultural land
use and urbanisation (Vandewalle et al., 2010). Nevertheless, the role of habitat constraints and dispersal abilities as filters, allowing only species with similar traits to assemble, has never been demonstrated at broad spatial scales, due to lack of suitable data, especially in soil invertebrates (Barbaro and van Halder, 2009; Decaëns et al., 2011; Makkonen et al., 2011; Pavoine et al., 2011). This may bias to a great extent the relationships between habitat preferences and species traits. Because the overall species response to habitat constraints involves trade-offs (Uriarte et al., 2012) between responses to different environmental factors (e.g. bedrock and climate, habitat openness and humidity, or temperature, or soil pH), it is easy to correlate erroneously a trait to an environmental factor. For example, the collembolan species *Heteromurus nitidus*, thought to strictly depend on soil pH since it was never found in soils at pH <5 in North and West of France (Ponge, 1980, 1993), was later found in soils at pH <4 in south-western mountains of France (Cassagne et al., 2003, 2004). One way of avoiding this error risk is to determine habitat preferences of species over a wide range of habitats, encompassing a variety of temperature and altitude levels, at a scale close to the geographic distribution range of the species.

Moreover, despite the abundance, high diversity and essential functional role of soil invertebrates (Hopkin, 1997; Coleman et al., 2004), trait-based approaches were not explicitly used to study species/environment patterns and processes in these animal groups (Vandewalle et al 2010). Only studies focusing either on a restricted number of traits (especially dispersal), or of habitats have been made to assess the effects of land-use disturbance or climate change on soil communities (Ponge et al., 2006; Vandewalle et al., 2010; Decaëns et al., 2011; Makkonen et al., 2011; Bokhorst et al., 2012).

The taxonomic Class of Collembola is a good model to address such questions, because it comprises a high number of species, occupying highly diverse habitats over a broad biogeographic area (Hopkin, 1997). Moreover, some authors have hypothesized, from field
observations, the existence of five or more “eco-morphological groups” based on conspicuous morphological differences among Collembola living in diverse habitats (Gisin, 1943; Delamare-Deboutteville, 1951; Rusek, 2007). They classified collembolan species according to the relationships between some morphological characteristics and different gradients of vertical stratification (edaphic, hemiedaphic, epigeic) and soil moisture (hydrophilic, xerophilic), but no attempt was made to rely statistically morphological characteristics (traits) to environmental variables. Europe, as a wide area including a high diversity of landscape and vegetation types, is a favourable terrain for exploring multivariate relationship between species trait values, assembly processes, and environmental factors.

In this study, we asked the following questions: (1) What is the pattern of relationships between species assemblages and environmental variables at broad geographic scale? (2) Which environmental variables are associated with trait variation in Europe and which environmental variables contribute to the assemblage of local communities?

To this end, we compiled a large volume of data about species traits and environmental characteristics of sites where species have been collected throughout Europe. To enable this, we created ‘Coltrait’, a database collating traits and occurrence data of European collembolan species across a wide range of habitats, mostly from Northwest Europe. Occurrence data and associated descriptions of samples and sampling sites were either provided by our own studies, or collected in the literature. For traits, we selected collembolan characteristics expected to explain the distribution of species and the subsequent composition of species communities through three processes that drive patterns of community composition, namely (1) Abiotic components of habitat, (i.e. environmental variables) adaptation/selection (e.g. sensorial organs, cuticle protection, reproduction type); (2) Dispersal ability (e.g. locomotory appendages); (3) Biotic components of habitat selection (predator defence, e.g. detection by sensory organs, excretion of repulsive substances).
We firstly analyzed the impact of environmental variables on the distribution of species in Europe, and then we analyzed patterns of trait/environment relationships.

2. Materials and methods

2.1. Data collection

2.1.1. Habitat characteristics and species occurrences

The Coltrait database comprises four tables that were used for the present study: a species traits table, a sample description table (environmental variables observed in samples or in sample sites to determine habitat characteristics), an occurrence/sample table and a bibliography table.

Habitat characteristics and occurrences of collembolan species in these habitats were provided either by our own studies (Arpin et al., 1984, 1985, 1986; Ponge, 1980, 1993; da Gama et al., 1994, 1997; Ponge, 2000; Loranger et al., 2001; Ponge et al., 2003; Gillet and Ponge, 2005) or were extracted from articles dealing with field studies on collembolan communities (e.g. Hågvar, 1982; Rusek, 1989, 1990; da Gama et al., 1994, 1997; Cassagne et al., 2003, 2004; Dunger et al., 2004; Chauvat et al., 2007; see Appendix 1 for the complete list of references). We selected habitat characteristics (environmental variables) that were described for a large amount of samples, proved to be linked to the composition of communities at local scale in previous studies and/or were susceptible to “filter” species traits.

We collected qualitative and quantitative data regarding site and sample descriptions (i.e. habitat characteristics) of 926 samples. Quantitative data (temperature and rainfall, altitude, soil pH and C/N) were directly incorporated in the data base (Table 1). We aggregated qualitative data into binary classes, assigning each sample to one of two classes.
for each modality of a given habitat: “1” if sampling occurred in the modality (e.g. close habitat, mull humus...), “0” if sampling did not occur in the modality and this for each field study. Modalities were (1) for habitat closure “close” (forest, wood), “open” (pasture, grassland, meadow, cultivated field) and “intermediate” (hedgerow, forest clearing, forest edge, heathland), (2) for soil characteristics “mull”, “moder” and “mor” humus, “peat” and “cultivated” soil, “organic” and “organo-mineral” horizons and “hydromorphic” soils, (3) for vertical stratification “edaphic” (soil), “hemiedaphic” (litter), “epigeic-1” (ground surface and mosses), “epigeic-2” (herb layer, boulder), “epigeic-3” (tree trunk and canopy) (Table 1).

These data allowed us determining the occurrence of species and their traits along several environmental gradients: vertical stratification, habitat closure, soil acidity, decomposition rate (humus form and C/N), moisture, minimum annual air temperature (hereafter named “temperature”), minimum annual rainfall (hereafter named “rainfall”) and altitude.

As sampling strategies varied between different studies and was not always precisely described regarding every sampling area, volume or depth of soil, we only compiled the presence/absence of species and not their abundance. We selected species that were recorded in at least 10 studies and 20 samples, providing a list of 58 most frequent collembolan species in Europe (Table 2). Habitat and species occurrence data covered the following countries: Germany, United Kingdom, Austria, Belgium, Denmark, Spain, Finland, France, Italy, Norway, the Netherlands, Poland, Portugal, Czech Republic, Slovakia, Sweden, two thirds of samples being from France. Available data from Greece were discarded because the 58 selected species were generally present in the 17 above-cited countries while 33 species out of 58 were not recorded in Greece according to Fauna Europaea. Consequently, in our analysis, biogeographic segregation does not bias species distribution.
2.1.2 Species traits

We first listed traits that were most likely to influence community assembly either through habitat characteristics, dispersal abilities, or biotic interactions. We then eliminated traits for which detailed and complete information was not available for a high number of species. This selection step provided a set of 11 morphological traits and one life-history trait (reproduction mode). Physiological traits would have been highly relevant, however, as far as we are aware this information is only available for a few species reared in laboratory conditions. In the same way, life-history traits were poorly informed except for reproduction mode.

Traits were collected from a number of specified synopses and identification keys (Gisin, 1960; Jordana et al., 1997; Fjellberg, 1998, 2007; Bretfeld, 1999; Potapow, 2001; Thibaud et al., 2004; Chahartaghi et al., 2006; Hopkin, 2007; Chernova et al., 2009; Dunger and Schlitt, 2011). Visual and jumping apparatus and leg lengths are supposed to be related to dispersal abilities of species (Ponge et al., 2006). Antennal length, eye (ocelli) number, presence of trichobothria and presence and complexity of post-antennal organs play a role in sensory functions (Hopkin, 1997) and are expected to vary between close versus open, and epigeic versus endogeic habitats. Body pigmentation and presence of scales are involved in UV protection, thermodynamic buffering and signalling (Hopkin, 1997) and are also suspected to vary with habitat characteristics, as well as body shape and length and length of jumping apparatus. Pseudocelli are circular structures allowing Collembola to extrude repulsive fluids from specialized glands (Hopkin, 1997; Rusek and Weyda, 1981) and thus play a role in protection against predation (Rusek and Weyda, 1981). At last sexual/parthenogenetic reproduction is associated to survival or colonizing strategy (Chernova et al., 2009; Lavelle et al., 1987). Most of these traits are species- or clade- (e.g. pseudocelli) specific and we did not consider here intra-specific variability since data on
morphometric variation over broad geographical areas are unavailable in Collembola. Data may be quantitative (e.g. body length), binary (e.g. presence/absence of scales), or semi-quantitative (e.g. furcula length, see Table 3). We computed traits of the 58 selected collembolan species. All trait attributes (e.g. furcula length categories) were issued from precise definitions (Table 3) and not from expert appreciations. Some of these traits, currently in the Coltrait database, will be integrated in the BETSI database (Hedde et al., 2012).

2.2. Statistical analyses

Correspondence analysis (CA) was used to analyze species-environment relationships without constraining species distribution (samples as observations, species as active variables, environmental variables as passive variables). As an explanatory step, CA allowed us to visualize patterns of species distribution with environmental variables superposed on the revealed gradients.

Canonical correspondence analysis (CCA) and Monte Carlo permutation tests were used to verify whether species distribution was significantly explained by environmental factors and to know the importance of each factor. The type of horizon (organo-mineral or organic) and the C/N ratio were discarded because they were not available for a high number of samples. As Spearman correlation tests revealed that some environmental variables or levels of environmental variables were correlated we discarded some of them and we tested only the effect of temperature and rainfall (for climate), soil pH, hydromorphy and only some modalities of habitat closure (Open), vertical stratification (Hemi, Epi-1, Epi-2, Epi-3), and humus forms (Mull, Moder). Missing values were estimated by the nearest neighbor method. Seven partial CCAs were then performed to test the effect of above-mentioned variables while discarding the effect of one of them on species distribution.

RLQ analyses (Dolédec et al., 1996) were performed to assess whether species traits distribution was significantly correlated with environmental variables, and to determine
patterns observed when variables were constrained. RLQ analysis allows to perform a double inertia analysis of two arrays (R: environmental variables, and Q: species traits) with a link expressed by a contingency table (L: species occurrences). We discarded C/N, organo-mineral and organic horizons and soil pH, that were not sufficiently informed. As missing observations were deleted when variables (e.g. mull, moder) were not fully informed (Table 1), and as informed variables varied among studies (either humus form, or climate) we performed two RLQs to minimize the loss of data in each analysis. We transformed quantitative data (temperature, altitude and rainfall, leg and furcula lengths, ocelli and PAO vesicle numbers) into semi-quantitative data with three or two classes each. This transformation was performed by discretizing data over two classes when the trait might be absent and consequently it comprises a class “0” (e.g. ocelli, PAO vesicle numbers) in addition to the two classes created by the discretization. Discretization was performed over three classes for the other traits (e.g. lengths). To limit deleting observations we pulled together samples taken either in edaphic or in hemiedaphic levels (S.Soil), creating a vertical stratification category gathering samples taken either from the soil or from the litter. We also created another level of vertical stratification (S.Soil-Epi) that includes moss cushions or grass tufts with adhering humus or soil. Other variables were the same as those used in CA (Table 3). The first RLQ tested the interaction of traits with climate, vertical stratification, habitat closure, and soil moisture, while the second RLQ tested the interaction of traits with humus form, habitat closure, and soil moisture.

At last, traits were analyzed separately using statistical models to test the effect of environmental variables highlighted by RLQ. We first calculated the percent occurrence of each species for each level of binary environmental variables (e.g. edaphic level, hemidaphic level, mull humus, moder humus, etc…), and for continuous environmental variables (soil pH, altitude, temperature and rainfall) we selected minimum or maximum values for each
species. We then implemented linear models on continuous trait values, binary logit models on binary traits, and a multinomial logit model on pigmentation which included three categories. We excluded variables with a variance inflation factor (VIF) higher than four to avoid multi-collinearity. We selected models with contributive variables that significantly explained the greatest part of trait variability (i.e. with highest adjusted R² and log-likelihood for linear and logit models, respectively). All variables bringing significant information to the model were included, even if their impact was not significant when evaluated separately (Type III analysis). In order to fulfill linear model assumptions, some variables had to be log-transformed. As pseudocell number did not follow a normal distribution even after transformation, we analyzed their presence/absence with a binary logit model.

RLQ analysis was performed with the package “ade4” in R software (R Development Core Team, 2010). CA, CCA, general linear models (GLM) and other calculations (e.g. data normalization, discretization) were done using XLSTAT® (Addinsoft®, Paris, France).

3. Results

3.1. Species and environmental variables

The relationship between species, environmental variables and traits was first analyzed by CA using samples as observations, species as active variables and environmental variables as passive variables (Figs. 1a, 1b). The first three components of CA extracted 17.4% of the total variance (6.3%, 5.9% and 5.2% for F1, F2 and F3, respectively). The projection of species and environmental variables along axes 1 and 2 (Fig. 1a) shows that species are spread along axis 1 mainly according to humus form, soil hydromorphy and habitat opening/closure and to a lesser extent soil acidity, temperature and altitude. Along axis 2, species are distributed according to vertical stratification, and to a lesser extent habitat closure/opening, soil hydromorphy, organic matter, and temperature. In the upper left side of the biplot, we find
edaphic and hemiedaphic species living in acid soils with a low rate of organic matter decomposition (peat, mor, moder): *Micaphorura absoloni, Mesaphorura tenuisensillata, Willemia denisi, Willemia anophthalma, Mesaphorura yosii, Micranurida pygmaea, Folsomia quadrioculata, Protaphorura armata* and *Pogonognathellus flavescens*. Opposed to this group along axis 1, we find edaphic and hemiedaphic species living in mull on calcareous or neutroacidocline soils: *Stenaphorura denisi, Mesaphorura hylophila, Mesaphorura krausbaueri, Heteromurus nitidus, Mesaphorura italic a, Paratullbergia callipygos*, and *Pseudosinella alba*.

Opposed along axis 2 to these two groups of edaphic and hemi-edaphic species, we find epigeic species living on the soil surface (epigeic-1), herb layers and boulders (epigeic-2) and trees (epigeic-3) (Figs 1a, 1b): *Orchesella cincta, Heteromurus major, Xenylla tullbergi Lepidocyrtus curvicollis, Xenylla grisea, Brachystomella parvula, Isotomurus palustris, Sminthurinus elegans, Sminthurides schoetti, Entomobrya multifasciata* and *Tomocerus minor*. Species distribution according to soil characteristics (humus form, acidity) thus occurs mainly for endogeic and hemiedaphic species.

Epigeic species are separated along axis 1 by low temperature levels and hydromorphy gradients, *O. cincta, H. major, X. tullbergi*, *X. grisea* and *L. curvicollis* occurring more often in slightly drier and/or warmer habitats (Temp), while *B. parvula, S. elegans, E. multifasciata, I. palustris* and *L. cyaneus* prefer hydromorphic soils (Hydro). Forests (close) and open habitats are better separated by axis 3 (Fig. 1b) that opposes epigeic species living preferentially in herb layers or exceptionally on boulders (epigeic-2) in open habitats (*S. elegans, B. parvula, L. cyaneus, E. multifasciata, S. schoetti, I. palustris*), to epigeic species living either at the soil surface (epigeic-1) or on trees (epigeic-3) in forests (*X. tullbergi, O. cincta, L. curvicollis, Entomobrya nivalis, X. grisea, Allacma fusca*). Species segregation according to habitat closure thus mainly occurs for epigeic species. *Parisotoma notabilis*
shows no preferences for the studied factors, like the very common species *Isotomiella minor*, which tends nevertheless to be observed more frequently in non-acidic soils.

CCA indicated that the bulk of tested environmental variables, i.e. humus form, habitat closure, vertical stratification, hydromorphy, temperature, rainfall and soil pH significantly impacted species distribution (Table 4). The seven partial CCAs computed without either vertical stratification (Hemie, Eda, Epi1, Epi2, Epi3), or humus form (Mull, Moder), or habitat closure (Open), or hydromorphy, or rainfall, or temperature, or soil pH showed that remaining factors impacted significantly the distribution of species (all p < 0.0001) but with a decrease in the rate of constraint inertia. This indicates that all factors impacted the distribution of species directly and not only through interactions with other factors. However, discarding the effect of one or the other variable did not affect to the same extent the constrained inertia (12.5% for total CCA): 5.2%, 8.0%, 10.5%, 10.6%, 10.6%, 11.5%, and 11.6% of the total inertia were explained when discarding vertical stratification, humus form, rainfall, temperature, habitat closure, hydromorphy and pH effect, respectively.

### 3.2. Traits and environmental variables

A first RLQ analysis was applied to climate (temperature and rainfall), altitude, vertical stratification, habitat closure, soil moisture, and traits (Figs. 2a, 2b). Axis 1 separates levels of rainfall, temperature, vertical stratification, soil hydromorphy, and to a lesser extent, altitude. Axis 2 separates modalities of the same factors than for axis 1 in addition to habitat opening/closure. Traits mostly represented in deep soil and closed habitats (on the positive side of axes 1 and 2) are the absence of pigmentation, furcula, ocelli, trichobothria and scales, the presence of a post-antennal organ (PAO) comprising a great number of vesicles, of pseudocelli (defense organs), short legs and antennae and small body length. These habitats also display a parthenogenetic dominant reproduction. On the soil surface, the
number of ocelli and furcula length increased slightly. When habitats are less edaphic, more open and in less favorable climatic conditions (higher altitude, associated to low temperature levels positive side of axis 1 and negative side of axis 2) the following trait modalities are more frequent: stocky body, variable pigmentation, PAO present but less developed, and furcula, present but usually regressed and not fully functional. In epigeic habitats and more favorable climatic conditions (i.e. high and intermediate temperature levels and lower altitude, negative side of axis 1) most common traits are larger body size, elongated legs and antennas, spherical and pigmented body, well-developed eyes, presence of trichobothria and scales, and absence of PAO and pseudocella, with sexual reproduction dominant.

A second RLQ tested the relationships between humus forms and soil types, habitat closure, soil moisture, and traits (Figs. 3a, 3b). Axis 1 separates closed and open habitats while axis 2 separates humus forms. Open habitats (positive side of axis 1) hosted species displaying traits previously mostly observed in epigeic and favorable climate conditions: increased body, leg and antennae length, spherical and pigmented body, with scales and trichobothria, well-developed eye spots, absence of pseudocelli, absence or reduction of PAO vesicles, and dominance of sexual reproduction. Furcula may be either very long or regressed. Near the positive end of axis 1, higher body length is found in well moistened soils. Closed (forest) and intermediate (heathland and scrub) habitats (negative side of axis 1) are separated along axis 2 according to humus form and thickness of organic horizon. Small body, leg and antennae length, with pseudocella and well developed PAO, but no pigmentation, scales and trichobothria, are observed in both mull and moder humus (negative side of axis 1 and axis 2). Some traits varied along axis 2 in relation to the humus form, especially the increase in organic horizon thickness. From mull to moder the following changes are observed: mixed reproduction shifts towards parthenogenesis, absence of eye towards simple eye, absence of furcula towards mean-length furcula. In humus forms with
thickest organic layers (Mor and Peat) particular traits are observed, e.g. stocky body, short furcula and variable pigmentation. This change of traits from mull with thin organic layers to humus forms with thicker organic layers (moder, then mor and peat) coincides with the passage from deep organo-mineral horizons with small interconnected pores (earthworm galleries, allowing Collembola to go deeper in the soil as in mull) to upper thick organic horizons that offer more space, allowing the movement of large bodies and the use of a medium-size or short furcula, while eyes are larger in relation to a little more light.

Monte-Carlo tests based on 1000 replicates were highly significant (p < 0.001) for the two RLQs, indicating a significant relationship between traits and environmental variables.

Linear and logistic models showed that environmental variables explained antenna, leg, body and furcula lengths, number of eyes and PAO vesicles, reproduction mode, spherical and cylindrical body shape, and presence of trichobothria, pseudocelli, scales and pigment to significant levels (Tables 5 and 6). Only the stocky body shape (McFadden R² = 0.099; Chi² = 3.362; p-value = 0.067) was not significantly explained by environmental variables. Logistic and linear regressions corroborated that temperature plays a significant part in the explanation of most trait variation (Tables 5 and 6). Antenna, furcula, leg, and body lengths, number of ocelli, sexual reproduction, spherical body shape, and presence of trichobothria were significantly and positively explained by minimum temperatures. Some of these traits tended to be (body length) or were (trichobothria) negatively explained by altitude. Vertical stratification (mostly through epigeic and hemiedaphic levels) contributed to the explanation of all traits (except stocky body). Both the number of PAO vesicles and the presence of pseudocelli were significantly and negatively explained by minimum temperatures and epigeic-1 level, and were positively explained by soil pH. Humus form and
habitat openness contributed to explain variations of only five and two traits, respectively, and
this contribution was significant in two cases only (peat and mor explained trichobothria
presence and antenna length, respectively).

4. Discussion

4.1. Species and environmental variables

We showed that the distribution of species is conditioned, first by vertical stratification, then
by humus form, air temperature, rainfall, and habitat closure. Soil moisture and acidity were
also shown to affect species distribution but more indirectly, probably through the effect of
humus form (Ponge et al., 2002; Cassagne et al., 2004), while the effect of altitude was
correlated to air temperature and rainfall (and more probably to insolation). These results arise
from an analysis performed at a broad scale, that covers most part of the biogeographic
distribution area of species. The multiplicity of sites over such a broad geographical scale
allows to overcome eventual interactions between spatial autocorrelation and environmental
variables that could arise at lower scales, e.g. at the scale of habitat (e.g. a forest) or
microhabitat (e.g. soil characteristics) (Caruso et al., 2012).

Our results show that these habitat constraints contributed in structuring collembolan
communities to an extent of about 12%. Consequently, at least 12% collembolan species
assemblies were explained in a deterministic way (Streit et al., 1985; Weiher and Keddy,
1995) while about 88% species distribution remained unexplained. A part of this unexplained
variation is probably due to the lack of data about biotic interactions like competition (McGill
et al., 2006; Comita et al., 2009).
We characterized habitat preferences for many species (see section 3.1.). The three levels of epigeic species were well separated between soil surface (epige-1), herb layers and boulders (epige-2) and trees (epige-3) while hemiedaphic species were closely related to edaphic species. This could in part be due to the fact that some samples contained a mixture of litter and organo-mineral or even mineral soil (soil), leading to a connection of species living in one or the other level, but also to the vertical movement of animals through horizons (Hassall et al., 1986). The distribution of species in the three epigeic levels was related to habitat closure, above-ground species that live in forests being preferentially located either at the soil surface (epige-1) or on trees (epige-3), contrary to above-ground species living in open areas which are located preferentially in herb layers or on boulders (epige-2). Species segregation according to habitat closure thus especially occurs for epigeic species because they are living far from the protection ensured by soil/litter layers, and occupy different substrata, some of them being more frequent in forests (trees) or in open environments (herb layers). Edaphic species appear in both closed and open habitats while hemiedaphic species seem more linked to closed habitats (due to thicker litter layers). However, their central location on the F2-F3 biplot of CA makes hemiedaphic habitats more difficult to interpret in terms of relationships with habitat closure. Edaphic species were actually segregated according to humus form and acidity.

Our results support studies done at a more local scale on the vertical distribution of species (Delamare-Deboutteville, 1951; Hale, 1966; Hågvar, 1983; Ponge, 2000), preferences for humus forms and soil acidity levels (Ponge, 1983; Hågvar and Abrahamsen, 1984; Loranger et al., 2001), and comparisons between forests and open habitats (Ponge, 1993). They reveal an impact of the minimum annual air temperature at which species have been found as well as the segregation of epigeic species among the three levels of epigeic habitats (see above). However, as our analysis encompasses a broad spatial area, and consequently
more diverse levels of environmental factors, we found some contrasted results compared to more local studies. For example *Mesaphorura macrochaeta*, previously observed in more relative abundance in acidic soils in a more local study (Ponge, 1993), is here found not really responding to acidity level, and even tending to occur more often in subacidic or non-acidic soils. The collembo lan *H. nitidus*, proved to prefer edaphic and hemiedaphic habitats in mull humus and calcareous or neutroacidocline soils (see 2.2 Results) on a broad scale, was able to live in acidic soil in some places (Cassagne et al., 2003, 2004). Such contrasted responses of species to varied environmental conditions show that the response (reaction norm) of species results from trade-offs between preferences for a variety of environmental factors and varies according to the value of each factor (Fry, 2003; Uriarte et al., 2012). They also highlight that in spite of marked habitat preferences, the studied collembo lan species were able to live in extremely diverse conditions. Consequently, the niche width will remain to be defined for the present 58 species in order to determine precisely the habitat limits of each species and thus provide a tool that could be used for eventual predictions of changes in collembo lan communities. Nevertheless, we have to keep in mind that some species appearing in our study as ubiquitous, such as *Parisotoma notabilis* or *Isotomiella minor*, might in fact be species complexes composed of two or more species with narrower niches as suggested by recent sequencing of a barcode DNA fragment (5’ end of COI gene) (Porco et al, 2012).

4.2. Traits and environmental variables

RLQ and GLM analyses showed that all tested environmental variables contributed to the assemblage of species traits. Vertical stratification followed by annual air temperature play a dominant role in the explanation of variation of the 12 studied traits.

The relationship between traits and vertical distribution of species, which was previously hypothesized by Collembola specialists (Gisin, 1943; Delamare-Deboutteville,
on the basis of their observations, is now statistically tested over a wide range of ecosystems. The contribution of habitat closure/openness and humus form to traits assemblages is showed at broad scale for the first time. We also showed that traits of species living in open habitats (e.g. grassland) or above-ground are both adapted to surface life, i.e. to light, a potentially dry environment and a wide space: high mobility (long legs and furcula), big size often associated with spherical body, scales, pigments (protection against UV), sensorial organs sensitive to air and light (trichobothria, ocelli and antennal organs), and sexual reproduction. On the opposite side, traits that are dominant in woodland are quite similar to those observed in hemiedaphic and edaphic levels and are mostly associated with subterranean concealed life or at least with a stable microclimate (litter in forests): small size, small locomotory appendices, toxic excreta (pseudocelli), post-antennal organs as main sensorial organs, and parthenogenesis. Such variation of trait modalities between open habitats and forests probably explain the higher soil fauna diversity (γ-diversity) observed in landscapes composed of a mixture of forests and open habitats in regard to homogeneous landscapes (Vanbergen et al, 2007).

The convergence of traits observed along gradients of habitat closure and of vertical distribution is explained by a higher incidence of environmental stress (drought, UV and wind) in less protected open as well as aboveground habitats. Among species previously recorded as drought-sensitive (Lindberg and Bengtsson, 2005), many appears here as being edaphic and woodland species (Micraphorura absoloni, Willemia anophthalma, Micranurida pygmaea, Mesaphorura macrochaeta), while drought-tolerant species are found to be epigeic (Lepidocyrtus spp., Entomobrya nivalis).

The relatively large body size of epigeic species is a protection strategy against frost and desiccation (Kaersgaard et al., 2004, Bokhorst et al., 2012). Moreover, open unstable and stressful environments like agricultural fields require higher motility (long legs and furcula)
than shrubby fallows where litter begins to accumulate or woodland (Mebes and Filser, 1997).

Escaping from predators aboveground also relies on jumping, ensured by long furcula acting as a spring (Bauer, 1987), as well as visual (ocelli) or tactile (trichobothria) sensorial organs to detect their presence (Baatrup et al., 2006). In epigeic and open habitats longer antennae, bearing specialized tactile and chemosensory setae and vesicles, also allow detecting chemicals and air vibration at a higher distance (Hopkin, 1997). Increased ocelli number and furcula, antenna and body lengths were also used by Rusek (2007) to propose two epigeic collembolan “life form groups”, namely “Macrophytobionts” and “Microphytobionts”, corresponding to epigeic levels 2 and 3, respectively, in our study. Conversely, reduced pigmentation, furcula length and number of ocelli fit generally well with the “Euedaphobiont life form group” (Rusek, 2007). Ocelli number, antenna and furcula lengths had also previously been shown to be correlated with change in the vertical distribution of species in response to tree plantation affecting litter quality and quantity (Vandewalle et al., 2010).

Pseudocelli, mostly present in edaphic species, probably compensate for the difficulty to escape predators through active movement in deep soil by excreting chemicals that repel predators (Negri, 2004). In the same way the presence of post-antennal organs (PAO) in edaphic and hemiedaphic species only, support the hypothesis raised by Salmon and Ponge (2012) that these organs could compensate for the absence of other sensorial organs that are adapted to air (trichobothria), light (ocelli) and spacious environments (sensory organs on long antennae), PAO being supposed to be composed of thermo-, hygro-, or chemosensitive receptor cells (Altner, 1976). At last, the reproductive insurance hypothesis (Jain, 1976) predicts that parthenogenesis is favored in populations where mates are difficult to find. This hypothesis can explain the higher rate of parthenogenesis observed among euedaphic species, living in deep soil horizons and faced to deficit of space and need to explore narrow soil pore networks for mating (Chahartaghi et al., 2006; Chernova et al., 2009).
The effect of climate, and especially of temperature, on the distribution of traits is mostly observed in aboveground and open habitats, i.e. in least protected and more unstable environments. Moreover, as climate drives vegetation patterns and humus forms (Ponge et al, 2011), the effect of climate seems tightly linked to that of humus forms.

In this way, less favorable climatic conditions (high altitude associated with low temperature levels), when they allow vegetation development, often generate mor humus and peat, i.e. humus with thick organic horizons made of undecayed plant fragments where collembolan species with stocky body can move more easily than in small soil pores, and are impeded to go deep in the soil by waterlogging. Such conditions favor traits that are intermediate between strictly edaphic and epigeic life (pigmentation variable, PAO present but less developed, and short or regressed furcula). Species traits observed under more favorable climate conditions (moderately high temperature and moisture and low altitude), are the same as those present in epigeic levels because such conditions are favorable to epigeic life and because species adapted to drought can live above-ground in open habitats (see above; Siepel, 1996). This is supported by the positive correlation of the activity of epigeic Collembola with temperature in agricultural soils (Frampton et al., 2001). At last, the post-antennal organ seems more complex (larger numbers of vesicles, in relation to the number of dendritic branches of neural systems; Altner and Thies, 1976) in most species of mull humus and edaphic levels than in moder humus and hemiedaphic levels. This suggests that complex post-antennal organs, supposed to be more sensitive to chemical features of the immediate environment (Altner and Thies, 1976), could compensate for the complete absence of eyes and reduction of other sensory organs observed in euedaphic species, providing better adaptation to deeper soil layers. Nevertheless, although the presence of PAO was essentially observed in edaphic and hemiedaphic species, only a study of the variation of PAO
complexity in relation to habitat depth of each species along one or several lineages will be able to state on the importance of PAO complexity for adaptation to subterranean life.

Our results support some of the relationships between traits and environmental factors observed in a previous study realized at a more local scale (Salmon and Ponge, 2012). Nevertheless, some differences have to be noted since the cited study did not allow discerning a clear correlation between traits and humus types associated to soil acidity. The larger dataset used here allowed us to demonstrate such relationships.

The present study, by showing the change of traits along environmental gradients supports the importance of niche-based environmental filtering in the assembly of Collembolan species (McGill et al., 2006, Weiher et al, 2011). However, models showed that environmental variables explained 23% (cylindrical body) to 71% (pseudocelli; Table 6) of trait variation, meaning that a high part of trait and species assembly variation is explained by other factors.

Community phylogenetics will probably add an important contribution to the explanation of species distribution and assembly as this has been shown for plant communities (Prinzing et al., 2008; Pavoine et al., 2011). In addition, ecophysiological traits such as resistance to desiccation, enzymatic equipment for food digestion, colonization rate, etc., were missing in our analysis because they were only available for a few species (Hopkin, 1997) and lengthy researches are needed to assess them in a meaningful number of species. At last, although pseudocelli and furcula development allow taking into account repulsion to and escape from predators, respectively, the effects of other biotic interactions such as competition, or commensalism, known to be important drivers of community structuration (McGill et al., 2006; Comita et al., 2009; Vellend, 2010), were not taken into account because of the little amount of available data. Limits of models relating species range distributions to environmental variables have been underlined, in particular when using species- or trait-environment relationships to track large-scale effects of global change (Fridley and Wright,
2012; Urban et al., 2013; Zhu et al., 2014). We showed that within the studied group climate effects concern mainly traits associated to aboveground habitats in open environments, and that soil properties are the main driver of changes in belowground communities. However, predictions about the response of collembolan trait/species distributions at local scale to climate warming will have to take into account other factors like biotic interactions (Urban et al., 2013) and ecophysiological traits.

5. Conclusions

Our study shows that collembolan species assembly at a broad scale is conditioned by a set of interacting parameters including vertical stratification, humus form, habitat closure, air temperature, soil moisture and acidity, and to a lesser extent rainfall and altitude. Our results show that all these environmental factors significantly contributed to the assemblages of twelve studied species traits. Combinations of certain environmental factors, especially stratification level and air temperature, explain a large variety of species traits and are thereby thought responsible for a great part of collembolan “species sorting”. A convergence was observed between traits related to stratification level and habitat closure/openness, that we interpreted as the effect of similar physiological and behavioral constraints. The action of humus form would impact indirectly trait distribution/assemblies through the interaction between vertical stratification and habitat openness. Ecological preferences were identified for the 58 studied collembolan species relatively to this set of traits, their niche width and their degree of specialization remaining to be more precisely defined. Nevertheless, relationships between traits and environments provide interesting proxies to identify ecological preferences
of Collembola, even when not identified to species level. Beyond Europe (where the
invertebrate fauna is relatively well-characterized), this approach may offer new insight, for
example with respect to tropical soil communities where most species are still unknown to
science, and will likely remain so in the near future.

References

Altner, H., Thies, G., 1976. The postantennal organ: a specialized unicellular sensory input to
the protocerebrum in apterygotan insects (Collembola). Cell and Tissue Research 167, 97-110.

Influence du peuplement forestier sur la faune et la microflore du sol et des humus. Revue

populations animales et microbienues du sol à la privation des apports annuels de litière :

des collemboles pour caractériser des phénomènes pédobiologiques. Revue d'Écologie et de
Biologie du Sol 21, 243-268.

the springtail Folsomia fimetaria and the influence of sex, size, starvation, and poisoning.
Entomologia Experimentalis et Applicata 118, 61-70.

to habitat fragmentation in mosaic landscapes. Ecography 32, 321-333.

Bauer, T., Christian, E., 1987. Habitat dependent differences in the flight behavior of
Collembola. Pedobiologia 30, 233-239.


Bokhorst, S., Phoenix, G.K., Bjerke, J.W., Callaghan, T.V., Huyer-Brugman, F., Berg, M.P.,
2012. Extreme winter warming events more negatively impact small rather than large soil
fauna: shift in community composition explained by traits not taxa. Global Change Biology
18, 1152-1162.

Abhandlungen und Berichte des Naturkundemuseums Goerlitz 71, 1-318.


complete framework for studying soil invertebrate functional traits. In: XVI ICSZ –
International Colloquium on Soil Zoology, Coimbra, Portugal.

Press, Oxford, UK.

Hopkin, S.P., 2007. A key to the Collembola (springtails) of Britain and Ireland. Field Studies
Council, Shrewsbury, UK.

Systematics 7, 469-495.


permeability, osmolyte production and body size for the desiccation resistance of nine species

Langlands, P.R., Brennan, K.E.C., Framenau, V.W., Main, B.Y., 2011. Predicting the post-
fire responses of animal assemblages: testing a trait-based approach using spiders. Journal of
Animal Ecology 80, 558-568.

Adaptive strategies of Pontoscolex corethrurus (Glossoscolecidae, Oligochaeta), a peregrine

Le Viol, I., Julliard, R., Kerbiriou, C., de Redon, L., Carnino, N., Machon, N., Porcher, E.,
2008. Plant and spider communities benefit differently from the presence of planted


altitudinal sequences in collembolan communities? Soil Biology and Biochemistry 33, 381-
393.

from functional traits. Trends in Ecology and Evolution 21, 178-185.Makkonen, M., Berg,
responses of a sub-arctic Collembola community to climate manipulation. Soil Biology and
Biochemistry 43, 377-384.


Pedobiologia 48, 585-588.

Ozinga, W.A., Romermann, C., Bekker, R.M., Prinzing, A., Tamis, W.L.M., Schaminée,
J.H.J., Henneksens, S.M., Thompson, K., Poschlod, P., Kleyer, M., Bakker, J.P., van


**Figure captions**

**Fig. 1.** Correspondence analysis showing the distribution and assembly of species (active variables) and environmental variables; a) Species and environmental variables along axes 1 and 2; b) Species and environmental variables along axes 2 and 3. See tables 1 and 2 for abbreviations.

**Fig. 2.** RLQ testing the direct relationships between climate (temperature and precipitation), altitude, vertical stratification, habitat closure and soil moisture, and traits. a) Environmental variables; b) Traits. Environmental variable classes were altitude: Alt.low: <= 50m, Alt.int: 50m<x<=100m, Alt.high >100m; rainfall: Pre.Low <=650mm, Pre.Int 650mm<x<=800mm, Pre.High >800mm; temperature: Tem.Low 2.8°C<x <8°C, Tem.Int 8°C<x<=11°C, Tem.High >11°C<x<15°C; habitat closure: close (H.Close), open (H.Open), intermediate (H.Int); stratification: epigeic levels of soil surface (S.Epi1), herb layers and boulders (S.Epi2), trees (S.Epi3), hemiedaphic and edaphic level (S.Soil) and soil and soil surface level (S.Epi.soil); hydromorphic (Hydro.1) and non-hydromorphic soils (Hydro.0). See Table 3 for abbreviations of species traits.

**Fig. 3.** RLQ testing the relationships between humus forms and soil types, habitat closure, soil moisture and traits. a) Environmental variables; b) Traits. Environmental variable classes were: habitat closure: close (H.Close), open (H.Open), intermediate (H.Int); stratification: epigeic levels of soil surface (S.Epi1), herb layers and boulders (S.Epi2), trees (S.Epi3), hemiedaphic and edaphic level (S.Soil) and soil and soil surface level (S.Epi.soil); hydromorphic (Hydro.1) and non-hydromorphic soils (Hydro.0); humus forms: mull (S.Mul), moder (S.Mod), mor (S.Mor), peat (S.Peat), mull or moder (S.MuMo), cultivated soil (S.Cult). See Table 3 for abbreviations of species traits.
Table 1
Modalities, abbreviations, information level, and data type of environmental variables.

<table>
<thead>
<tr>
<th>Environmental variables</th>
<th>Modalities</th>
<th>Abbreviations for CA graphics</th>
<th>Information level (%)</th>
<th>Data type (units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Habitat closure</td>
<td>Close habitat</td>
<td>Close</td>
<td>100</td>
<td>0/1</td>
</tr>
<tr>
<td></td>
<td>Open habitat</td>
<td>Open</td>
<td>100</td>
<td>0/1</td>
</tr>
<tr>
<td></td>
<td>Intermediate habitat</td>
<td>Int</td>
<td>100</td>
<td>0/1</td>
</tr>
<tr>
<td>Stratification</td>
<td>Epigeic 1 (soil surface and mosses)</td>
<td>Epi-1</td>
<td>100</td>
<td>0/1</td>
</tr>
<tr>
<td></td>
<td>Epigeic 2 (herb layers and boulders)</td>
<td>Epi-2</td>
<td>100</td>
<td>0/1</td>
</tr>
<tr>
<td>Humus form</td>
<td>Peat</td>
<td>Peat</td>
<td>100</td>
<td>0/1</td>
</tr>
<tr>
<td></td>
<td>Mull</td>
<td>Mull</td>
<td>65</td>
<td>0/1</td>
</tr>
<tr>
<td></td>
<td>Moder</td>
<td>Mod</td>
<td>63</td>
<td>0/1</td>
</tr>
<tr>
<td></td>
<td>Mor</td>
<td>Mor</td>
<td>63</td>
<td>0/1</td>
</tr>
<tr>
<td></td>
<td>Cultivated soil</td>
<td>Cult</td>
<td>95</td>
<td>0/1</td>
</tr>
<tr>
<td>Geoclimatic features</td>
<td>Mean annual temperature</td>
<td>Temp</td>
<td>58</td>
<td>value (°C)</td>
</tr>
<tr>
<td></td>
<td>Annual rainfall</td>
<td>Pre</td>
<td>78</td>
<td>value (mm/year)</td>
</tr>
<tr>
<td></td>
<td>Altitude</td>
<td>Alt</td>
<td>80</td>
<td>value (m)</td>
</tr>
<tr>
<td>Soil chemistry</td>
<td>Hydromorphy</td>
<td>Hydro</td>
<td>100</td>
<td>0/1</td>
</tr>
<tr>
<td></td>
<td>Soil pH</td>
<td>pH</td>
<td>61</td>
<td>0/1</td>
</tr>
<tr>
<td></td>
<td>Organic layer</td>
<td>Org</td>
<td>49</td>
<td>0/1</td>
</tr>
<tr>
<td></td>
<td>Organo-mineral layer</td>
<td>OM</td>
<td>47</td>
<td>0/1</td>
</tr>
<tr>
<td></td>
<td>C/N ratio</td>
<td>C/N</td>
<td>42</td>
<td>value</td>
</tr>
</tbody>
</table>
Table 2
Name, abbreviations for CA analysis, and number of observations of the 58 most commonly recorded collembolan species.

<table>
<thead>
<tr>
<th>Species name</th>
<th>Abbreviation</th>
<th>Number of observations</th>
<th>Species name</th>
<th>Abbreviation</th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allacma fusca (Linnaeus, 1858)</td>
<td>Afu</td>
<td>43</td>
<td>Mesaphorura tenuisensillata</td>
<td>Mte</td>
<td>97</td>
</tr>
<tr>
<td>Brachystomella parvula (Schäffer, 1896)</td>
<td>Bpa</td>
<td>66</td>
<td>Mesaphorura yosii (Rusek, 1967)</td>
<td>Myo</td>
<td>106</td>
</tr>
<tr>
<td>Ceratophysella armata (Nicolet, 1841)</td>
<td>Car</td>
<td>57</td>
<td>Metaphorura affinis (Börner, 1902)</td>
<td>Maf</td>
<td>20</td>
</tr>
<tr>
<td>Ceratophysella denticulata (Bagnall, 1941)</td>
<td>Cde</td>
<td>88</td>
<td>Micranurida pygmaea Börner, 1901</td>
<td>Mca</td>
<td>87</td>
</tr>
<tr>
<td>Dicyrtoma fusca (Lubbock, 1873)</td>
<td>Dfu</td>
<td>69</td>
<td>Oncopodura crassicornis Shoebotham, 1911</td>
<td>Ocr</td>
<td>35</td>
</tr>
<tr>
<td>Dicyrtomina minuta (O. Fabricius, 1783)</td>
<td>Dmi</td>
<td>53</td>
<td>Neanura muscorum (Templeton, 1835)</td>
<td>Nmu</td>
<td>147</td>
</tr>
<tr>
<td>Entomobrya multifasciata (Tullberg, 1871)</td>
<td>Emu</td>
<td>54</td>
<td>Paratullbergia callipygos (Börner, 1902)</td>
<td>Oci</td>
<td>111</td>
</tr>
<tr>
<td>Entomobrya nivalis (Linnaeus, 1758)</td>
<td>Eni</td>
<td>75</td>
<td>Parisotoma notabilis (Schäffer, 1896)</td>
<td>Pno</td>
<td>598</td>
</tr>
<tr>
<td>Folsomia candida Willem, 1902</td>
<td>Fca</td>
<td>36</td>
<td>Parisotoma minima (Absolon 1901)</td>
<td>Pmi</td>
<td>66</td>
</tr>
<tr>
<td>Folsomia manolachei Bagnall, 1939</td>
<td>Fma</td>
<td>268</td>
<td>Paratullbergia callipygos (Börner, 1902)</td>
<td>Pfl</td>
<td>77</td>
</tr>
<tr>
<td>Folsomia quadrioculata (Tullberg, 1871)</td>
<td>Fqu</td>
<td>310</td>
<td>Pogonognathellus flavescens (Tullberg, 1871)</td>
<td>Par</td>
<td>161</td>
</tr>
<tr>
<td>Friesea mirabilis (Tullberg, 1871)</td>
<td>Fmi</td>
<td>103</td>
<td>Protaphorura armata (Tullberg 1869)</td>
<td>Pminu</td>
<td>38</td>
</tr>
<tr>
<td>Friesea truncata Cassagnau, 1958</td>
<td>Ftr</td>
<td>239</td>
<td>Protaphorura armata (Tullberg 1869)</td>
<td>Pminu</td>
<td>38</td>
</tr>
<tr>
<td>Heteromurus major (Moniez, 1889)</td>
<td>Hma</td>
<td>124</td>
<td>Protaphorura armata (Tullberg 1869)</td>
<td>Pminu</td>
<td>38</td>
</tr>
<tr>
<td>Species</td>
<td>Code</td>
<td>ID</td>
<td>Species</td>
<td>Code</td>
<td>ID</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------</td>
<td>----</td>
<td>----------------------------------------------</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td>Heteromurus nitidus (Templeton, 1835)</td>
<td>Hni</td>
<td>68</td>
<td>Pseudosinella alba (Packard, 1873)</td>
<td>Pal</td>
<td>222</td>
</tr>
<tr>
<td>Isotomiella minor (Schäffer, 1896)</td>
<td>Imi</td>
<td>642</td>
<td>Pseudachorutes parvulus Börner, 1901</td>
<td>Ppa</td>
<td>96</td>
</tr>
<tr>
<td>Isotomurus palustris (Müller, 1776)</td>
<td>Ipa</td>
<td>73</td>
<td>Pseudisotoma sensibilis (Tullberg, 1876)</td>
<td>Pse</td>
<td>61</td>
</tr>
<tr>
<td>Isotoma viridis Bourlet, 1839</td>
<td>Ivir</td>
<td>92</td>
<td>Sminthurinus aureus (Lubbock, 1862)</td>
<td>Sau</td>
<td>208</td>
</tr>
<tr>
<td>Lepidocyrtus curvicollis Bourlet, 1839</td>
<td>Lcu</td>
<td>51</td>
<td>Sminthurinus elegans (Fitch, 1863)</td>
<td>Sel</td>
<td>44</td>
</tr>
<tr>
<td>Lepidocyrtus cyanus Tullberg, 1871</td>
<td>Lcy</td>
<td>119</td>
<td>Sminthurides schoetti Axelson, 1903</td>
<td>Ssc</td>
<td>52</td>
</tr>
<tr>
<td>Lepidocyrtus lanuginosus (Gmelin, 1788)</td>
<td>Lla</td>
<td>459</td>
<td>Sminthurinus signatus (Krausbauer, 1898)</td>
<td>Ssi</td>
<td>152</td>
</tr>
<tr>
<td>Lepidocyrtus lignorum (Fabricius, 1793)</td>
<td>Lli</td>
<td>207</td>
<td>Sphaeridia pumilis (Krausbauer, 1898)</td>
<td>Spu</td>
<td>221</td>
</tr>
<tr>
<td>Lepidocyrtus violaceus (Geoffroy, 1762)</td>
<td>Lvi</td>
<td>30</td>
<td>Stenaphorurella denisi (Bagnall, 1935)</td>
<td>Sde</td>
<td>93</td>
</tr>
<tr>
<td>Lipothrix lubbocki (Tullberg, 1872)</td>
<td>Llu</td>
<td>69</td>
<td>Tomocerus minor (Lubbock, 1862)</td>
<td>Tmi</td>
<td>79</td>
</tr>
<tr>
<td>Megalothorax minimus Willem, 1900</td>
<td>Mmi</td>
<td>453</td>
<td>Willemia anopthalma Börner, 1901</td>
<td>Wan</td>
<td>197</td>
</tr>
<tr>
<td>Mesaphorura hylophila Rusek, 1982</td>
<td>Mhy</td>
<td>161</td>
<td>Willemia denisi Mills, 1932</td>
<td>Wde</td>
<td>87</td>
</tr>
<tr>
<td>Mesaphorura italica (Rusek, 1971)</td>
<td>Mit</td>
<td>76</td>
<td>Willemia intermedia Mills, 1934</td>
<td>Win</td>
<td>38</td>
</tr>
<tr>
<td>Mesaphorura krausbaueri Börner, 1901</td>
<td>Mkr</td>
<td>174</td>
<td>Xenylla grisea Axelson, 1900</td>
<td>Xgr</td>
<td>45</td>
</tr>
<tr>
<td>Mesaphorura macrochaeta Rusek, 1976</td>
<td>Mma</td>
<td>513</td>
<td>Xenylla tullbergi Börner, 1903</td>
<td>Xtu</td>
<td>107</td>
</tr>
</tbody>
</table>
Table 3
Modalities of traits, abbreviations for RLQ and models, type of data, number of species for each category of categorical and binary traits and mean* for discrete and continuous data.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Trait modalities</th>
<th>Types of data for RLQ and models</th>
<th>Number of species per category or mean* for all species</th>
<th>Abbrevation of classes for RLQ</th>
<th>Abbrevation for regression models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reproduction mode</td>
<td>Parthenogenesis</td>
<td>0/1</td>
<td>10</td>
<td>Repro.part</td>
<td>Repro.part</td>
</tr>
<tr>
<td></td>
<td>Sexual reproduction</td>
<td>0/1</td>
<td>41</td>
<td>Repro.sex</td>
<td>Repro.sex</td>
</tr>
<tr>
<td></td>
<td>Mixed reproduction</td>
<td>0/1</td>
<td>7</td>
<td>Repro.mix</td>
<td>Repro.mix</td>
</tr>
<tr>
<td>Body shape</td>
<td>Spherical body</td>
<td>0/1</td>
<td>10</td>
<td>Body.sphe</td>
<td>Body.sphe</td>
</tr>
<tr>
<td></td>
<td>Cylindrical body</td>
<td>0/1</td>
<td>42</td>
<td>Body.cyl</td>
<td>Body.cyl</td>
</tr>
<tr>
<td></td>
<td>Stocky body</td>
<td>0/1</td>
<td>6</td>
<td>Body.stock</td>
<td>Body.stock</td>
</tr>
<tr>
<td>Body length (RLQ)</td>
<td>Categories</td>
<td>-1: less than 2mm</td>
<td>9</td>
<td>Body.L.1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2: between 2 and 3mm.</td>
<td>7</td>
<td>Body.L.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3: more than 3mm.</td>
<td>42</td>
<td>Body.L.3</td>
<td></td>
</tr>
<tr>
<td>Body length (models)</td>
<td>Length in mm</td>
<td>1.82*</td>
<td>-</td>
<td>Body.L</td>
<td></td>
</tr>
<tr>
<td>Locomotory organs</td>
<td>Furcula length</td>
<td>Categories:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-4: reaches the anterior limit of abd 2 when folded along body</td>
<td>29</td>
<td>Furca.L.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3: reaches the anterior limit of abd 3</td>
<td>5</td>
<td>Furca.L.3</td>
<td>Furca.L.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2: does not reach the anterior limit of abd 3</td>
<td>6</td>
<td>Furca.L.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1: rudimentary, not functional</td>
<td>3</td>
<td>Furca.L.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0: absent</td>
<td>5</td>
<td>Furca.L.0</td>
<td></td>
</tr>
<tr>
<td>Leg length</td>
<td>Ratio leg length / body length</td>
<td>0.36*</td>
<td></td>
<td>Leg.L.1</td>
<td>Legs-L</td>
</tr>
<tr>
<td>Sensory or sense-related organs</td>
<td>Antenna length</td>
<td>Ratio antenna length /head diagonal length</td>
<td>1.17*</td>
<td>Ant.L.1</td>
<td>Antenna-L</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ant.L.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ant.L.3</td>
<td></td>
</tr>
<tr>
<td>Trait</td>
<td>Trait modalities</td>
<td>Types of data for RLQ and models</td>
<td>Number of species per category or mean* for all species</td>
<td>Abbrevation of classes for RLQ</td>
<td>Abbrevation for regression models</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------------</td>
<td>----------------------------------</td>
<td>---------------------------------------------------------</td>
<td>--------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Sensory or sense-related organs</td>
<td>Number of ocelli</td>
<td>Number from 0 to 8</td>
<td>4-5*</td>
<td>Ocel.0 Ocel.1 Ocel.2</td>
<td>Ocelli</td>
</tr>
<tr>
<td></td>
<td>Max number of PAO vesicles</td>
<td>Number from 0 to 40</td>
<td>8*</td>
<td>PAO.ves.0 PAO.ves.1 PAO.ves.2</td>
<td>PAO vesicles-max</td>
</tr>
<tr>
<td>Trichobothria</td>
<td>Absent:0 Present : 1</td>
<td>24</td>
<td>Tricho.0 Tricho.1</td>
<td>Tricho</td>
<td></td>
</tr>
<tr>
<td>Protective features</td>
<td>Scales</td>
<td>Absent:0 Present : 1</td>
<td>11</td>
<td>Scale.0 Scale.1</td>
<td>Scales</td>
</tr>
<tr>
<td></td>
<td>Pigmentation</td>
<td>Absent:0 Present : 1 Presence variable (RLQ only)</td>
<td>36</td>
<td>Pigment.0 Pigment.1 Pigment.var</td>
<td>Pigment</td>
</tr>
<tr>
<td>Defense against predators</td>
<td>Pseudocelli</td>
<td>Absent:0 Present : 1</td>
<td>11</td>
<td>Pseudo.0 Pseudo.1</td>
<td>Pseudo</td>
</tr>
</tbody>
</table>
Table 4
Results of CCA followed with permutation tests with species occurrence as active variables and environmental factors as constraining variables. Humus form: Mull, Moder; Habitat closure: Open; Vertical stratification: Hemiedaphic, Epigeic-1, Epigeic-2, Epigeic-3; Hydromorphy, pH: soil pH, Climate: minimum Temperature and Precipitation.

<table>
<thead>
<tr>
<th>Habitat closure-Vertical stratification-Humus form-Hydromorphy-Climate-pH</th>
<th>Totale</th>
<th>Constrained (%)</th>
<th>No-constrained (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inertia</td>
<td>4.987</td>
<td>0.625</td>
<td>4.362</td>
</tr>
<tr>
<td>Permutation test</td>
<td>100</td>
<td>12.528</td>
<td>87.472</td>
</tr>
<tr>
<td>Permutation number</td>
<td>900</td>
<td>Pseudo F</td>
<td>0.680</td>
</tr>
<tr>
<td>P-value</td>
<td>&lt; 0.0001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5
Results of linear models (R², F, p-value) and type III analysis of sum of squares (p-value) between discrete and continuous trait values, respectively and species occurrence in binary environmental variables or minimum values of continuous environmental variables (values in bold type indicate significant effect: p < 0.05; see Tables 1 and 3 for abbreviations).

<table>
<thead>
<tr>
<th></th>
<th>Antenna-L</th>
<th>Furcula-L</th>
<th>Ocelli</th>
<th>Legs-L</th>
<th>Log(Body-L)</th>
<th>Log(PAO vesicles-max)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R²</td>
<td>0.352</td>
<td>0.328</td>
<td>0.437</td>
<td>0.405</td>
<td>0.454</td>
<td>0.430</td>
</tr>
<tr>
<td>F</td>
<td>5.543</td>
<td>6.337</td>
<td>10.085</td>
<td>6.937</td>
<td>4.983</td>
<td>7.710</td>
</tr>
<tr>
<td>Pr &gt; F</td>
<td><strong>0.000</strong></td>
<td><strong>0.000</strong></td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td><strong>0.000</strong></td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Alt-min</td>
<td></td>
<td>0.359</td>
<td></td>
<td>0.285</td>
<td>0.057</td>
<td>0.378</td>
</tr>
<tr>
<td>Pre-min</td>
<td>0.062</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.041</td>
</tr>
<tr>
<td>Temp-min</td>
<td><strong>0.016</strong></td>
<td><strong>0.034</strong></td>
<td><strong>0.023</strong></td>
<td>0.099</td>
<td>0.107</td>
<td><strong>0.001</strong></td>
</tr>
<tr>
<td>Moder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.054</td>
</tr>
<tr>
<td>Mor</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eda</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.043</td>
</tr>
<tr>
<td>Hemi</td>
<td></td>
<td></td>
<td>0.809</td>
<td><strong>0.004</strong></td>
<td></td>
<td>0.692</td>
</tr>
<tr>
<td>Epi-1</td>
<td><strong>0.001</strong></td>
<td>&lt;0.0001</td>
<td></td>
<td><strong>0.001</strong></td>
<td><strong>0.001</strong></td>
<td></td>
</tr>
<tr>
<td>Epi-2</td>
<td></td>
<td></td>
<td></td>
<td><strong>0.006</strong></td>
<td>0.040</td>
<td></td>
</tr>
<tr>
<td>Epi-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>0.034</strong></td>
<td></td>
</tr>
<tr>
<td>pH-min</td>
<td>0.067</td>
<td>0.054</td>
<td>0.152</td>
<td>0.081</td>
<td><strong>0.001</strong></td>
<td></td>
</tr>
</tbody>
</table>
Table 6
Results of logistic models (Mc Fadden R², Khi², p-value) and type III analysis (p-value) between binary and qualitative trait values, respectively and species occurrence in binary environmental variables or minimum values of continuous environmental variables (values in bold type indicate significant effect: p <0.05; see Tables 1 and 3 for abbreviations).

<table>
<thead>
<tr>
<th>Scale</th>
<th>Pseudo</th>
<th>Repro.sex</th>
<th>Repro.part</th>
<th>Body.sphe</th>
<th>Body.cyl</th>
<th>Tricho</th>
<th>Pigment</th>
</tr>
</thead>
<tbody>
<tr>
<td>R²</td>
<td>0.084</td>
<td>0.707</td>
<td>0.316</td>
<td>0.348</td>
<td>0.310</td>
<td>0.229</td>
<td>0.447</td>
</tr>
<tr>
<td>McFadden</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Khi²</td>
<td>4.756</td>
<td>39.842</td>
<td>22.207</td>
<td>18.572</td>
<td>16.556</td>
<td>15.211</td>
<td>35.152</td>
</tr>
<tr>
<td>p-value</td>
<td>0.029</td>
<td>&lt;0.0001</td>
<td>0.000</td>
<td>&lt;0.001</td>
<td>&lt;0.005</td>
<td>0.019</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Alt-min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temp-min</td>
<td>0.000</td>
<td>0.030</td>
<td>0.030</td>
<td>0.077</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mull</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moder</td>
<td>0.623</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open</td>
<td>0.217</td>
<td>0.698</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eda</td>
<td>0.029</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemi</td>
<td>0.095</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epi-1</td>
<td>&lt;0.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epi-2</td>
<td>0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epi-3</td>
<td>0.053</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pH-min</td>
<td>0.001</td>
<td>0.281</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. 1

a)

b)
Fig. 2

a) 

b)
Fig. 3

(a)
Appendix 1

References dealing with field studies on collembolan communities and used to compile occurrences of collembolan species in diverse habitats and the description of these habitats in Coltrait database.


