ZENTRUM
FÜR BIODIVERSITÄT UND NACHHALTIGE LANDNUTZUNG
SEKTION
BIODIVERSITÄT, ÖKOLOGIE UND NATURSCHUTZ

– CENTRE OF BIODIVERSITY AND SUSTAINABLE LAND USE –

SECTION: BIODIVERSITY, ECOLOGY AND NATURE CONSERVATION

Assembly processes in soil animal communities:

Integrating phylogeny and trait-based approaches

Dissertation zur Erlangung des Doktorgrades der
Mathematisch-Naturwissenschaftlichen Fakultäten der
Georg-August-Universität Göttingen

vorgelegt von

Master of Science

Ting-Wen Chen

aus
Kaohsiung City, Taiwan

Göttingen, Dezember 2017

Referent: Prof. Dr. Stefan Scheu
Korreferent: Prof. Dr. Holger Kreft
Tag der mündlichen Prüfung: 26.01.2018

Nothing in biology makes sense except in the light of evolution.

Theodosius Dobzhansky

Table of contents

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of publications

Thesis declarations

Summary

Soils are among the most biodiverse systems on earth. The coexistence in soils of a multitude of animal species has long puzzled soil ecologists. How can so many species co-occur, and what are the processes driving and maintaining species coexistence in soil? Using a deductive approach, I propose that (1) there are assembly processes, (2) that work on, or are related to, certain objects, i.e., functional traits, to (3) produce particular patterns. I use a conceptual model combining patterns of evolution of species traits, trait similarity and phylogenetic relatedness between coexisting species, from which to infer assembly processes in soil Collembola (springtail) communities collected from habitats characterized by different disturbance regimes.

In **Chapter 2**, I reconstruct a Collembola phylogeny and use phylogenetic comparative methods to explore phylogenetic signal, model of evolution and ancestral state for a variety of traits, including body shape, body length, pigmentation, number of ommatidia, vertical stratification and reproductive mode. The results demonstrate that body shape of Collembola evolved quickly early in their diversification but slowed down afterwards. In contrast, evolutionary transitions in pigmentation, number of ommatidia and reproductive mode depended on how deep in the soil that species live. Ancestral Collembola traits were likely slender body, hemiedaphic way of life, sexual reproduction, possession of many ommatidia and bright color, but these traits presumably changed several times during species diversification. The phylogenetic signal detected in these traits forms the basis of further community phylogenetic analyses.

In **Chapter 3**, I propose the neutral lipid fatty acid composition of Collembola as a functional trait related to both food resources and physiological functions and test phylogenetic signal in fatty acid profiles. Long-chain polyunsaturated fatty acids related to physiological functions demonstrated phylogenetic signal. In contrast, most food resource biomarker fatty acids and the ratios between bacterial, fungal and plant biomarker fatty acids exhibited no phylogenetic signal. These results suggest that Collembola with close phylogenetic affinity experienced similar environments during divergence, while niche partitioning in food resources among closely related species favored species coexistence.

In **Chapter 4**, I use both community phylogenetic and trait-based approaches to infer the assembly processes of Collembola communities inhabiting arable fields, grasslands and forests. The results indicate that Collembola communities in arable fields were mainly structured by environmental filtering, while niche partitioning dominated in forests. Epedaphic (surface-living) species showed phylogenetic clustering in grasslands and forests, while in forests they also possessed similar traits. Hemiedaphic (sub-surface-dwelling) species were phylogenetically clustered in arable fields and grasslands, but in forests they were phylogenetically overdispersed and had different traits. However, the assembly of euedaphic (soil-dwelling) communities did not differ from random patterns. Furthermore, different phylogenetic groups of Collembola showed different patterns in the three habitats. These results suggest that Collembola assemblages are driven by different mechanisms in different habitats, with the relative importance of these mechanisms different between soil strata and between phylogenetic lineages.

Furthermore, applying community phylogenetic approaches to a manipulative soil block experiment (Auclerc et al. 2009; Soil Biology and Biochemistry 41, 1596–1604) in **Chapter 5** shows that Collembola community composition during their succession in forest and meadow soil was determined by the interaction of dispersal and selection processes. Niche partitioning gradually strengthened at later successional stages, offsetting the effects of environmental filtering. As a consequence of dispersal, community composition changed gradually from that resembling the original habitats to that of the new habitats.

In the final chapter I ascribe the above-mentioned patterns to the scenarios presented in the conceptual model and discuss the likely mechanisms, with reference to the four high-level processes, selection, dispersal, drift and speciation, proposed in *The Theory of Ecological Communities* (Vellend 2016). I provide a roadmap for integrating phylogenetic comparative methods, community phylogenetic analyses and trait-based approaches in studies on the assembly processes of soil communities. Overall, this thesis is the first application of new methods developed in community ecology and evolutionary biology to the study on assembly processes in the soil communities. Future studies using the conceptual model and roadmap proposed in this thesis will advance our understanding of the mechanisms driving and maintaining soil biodiversity from both ecological and evolutionary perspectives.

Chapter 1

General Introduction

The nature of soil biodiversity

Soils are among the most biodiverse systems on earth and have been viewed as "the poor man's tropical rainforest" (Giller 1996). The coexistence in soils of a multitude of animals has long puzzled soil ecologists, as reflected in the phrase "the enigma of soil animal species diversity" (Anderson 1975). For example, the density of soil mesofauna, animals with a body width between 0.2 mm and 2 mm, typically ranges between 10,000 and 200,000 individuals m^{-2}, and local species richness in temperate deciduous forests is usually between 60 and 200 species (Petersen and Luxton 1982). How can so many species co-occur, and what are the processes driving and maintaining species coexistence in soil?

Inferring processes from patterns

For natural communities, assembly processes are hidden. What is observable in nature are the final patterns, the results of various mechanisms working on the members of communities. Using a deductive reasoning approach, I propose that (1) there are assembly **processes**, (2) that work on, or are related to, certain **objects**, to (3) produce particular **patterns**. By designing models of specific assembly processes and deriving the subsequent patterns, it is possible to relate the patterns observed in real communities with those derived from the models, thus inferring the assembly processes.

Community assembly processes

Theories posit that communities are affected by a number of **processes**, including niche-related (Chase and Leibold 2003), neutral (Hubbell 2001) and biogeographical processes (Ricklefs 1987). Recently, Vellend (2010, 2016) in *The Theory of Ecological Communities* linked community ecology and evolutionary biology and summarizes a variety of processes into four overarching high-level processes: selection, drift, dispersal and speciation. In this theory, niche-related processes, such as abiotic environmental filtering and interspecific competition, usually considered as major drivers of community composition, are ascribed to selection processes. Studies on soil biotic communities usually fall into this category and focus on e.g., community-environment relationships (Scheu and Schulz 1996, Scheu et al. 2003, Eissfeller et al. 2013). Ecological drift, in line with the neutral theory of biodiversity (Hubbell 2001), emphasizes stochasticity (i.e., rare or unpredictable fluctuations of populations) of local communities. Some studies suggest that the contribution of stochastic drift to soil community assembly is similar to that of deterministic processes (Minor 2011, Caruso et al. 2011, Caruso et al. 2012). Furthermore, dispersal as a high-level process (i.e., movement of individuals between local communities or from source populations) interacts with selection processes and so together influence local community composition. Metacommunity studies of soil invertebrates point to the generality of frequent dispersal of individuals between local communities, resulting in mass effects being a dominant metacommunity scenario for soil biota (Ingimarsdóttir et al. 2012, Heiniger et al. 2014). Finally, speciation, as a consequence of selection and dispersal but not usually discussed as part of community ecology, is the only process generating new ecologically relevant phenotypes or traits—the objects on which selection processes can work.

Functional traits—the objects

Functional traits are properties of species which influence their performance and fitness (Violle et al. 2007, Pey et al. 2014). Ecologically, functional traits regulate the occurrence of species in habitats and the coexistence with other species (McGill et al. 2006, Ackerly and Cornwell 2007, Adler et al. 2013). They are the **objects** (or targets, media) on which work selection processes, such as environmental filtering and interspecific competition, resulting in certain patterns of traits in local communities. There are two categories of ecological traits, α and β niche traits. β niche traits determine the environmental tolerance of species, while α niche traits relate to resource exploitation (Ackerly and Cornwell 2007). Similar β niche traits but different α niche traits allow species to live under similar environmental conditions but utilize different resources thereby promoting coexistence (Silvertown et al. 2006).

From an evolutionary perspective, traits are the products of the adaptation of species during their evolutionary history. They may exhibit phylogenetic signal (i.e., a statistic pattern where closely related species resemble each other in their trait values), since species inherit similar traits from their common ancestors. Where this occurs, variations in traits between species are predicted by phylogenetic distances. However, environmental constraints in the past may result in more conserved traits than predicted, while other diversifying mechanisms, such as adaptive radiation or competition, may result in trait divergence and therefore in traits being phylogenetically labile. As a consequence, β niche traits are usually phylogenetically conserved and exhibit phylogenetic signal, while α niche traits tend to be evolutionarily labile or divergent (Ackerly et al. 2006, Best and Stachowicz 2013). The different evolutionary consequences of α and β niche traits allow species coexistence (Silvertown et al. 2006, Ackerly and Cornwell 2007). The phylogenetic signal in functional traits therefore forms a mechanistic link between the evolutionary history of species and the contemporary ecological processes to which they are exposed (Cavender-Bares et al. 2009). Contemporary selection processes, such as environmental filtering and biotic interactions, work on existing traits, resulting in similar and/or different traits among the members of local communities.

Trait-based approaches aim at inferring community assembly processes from the patterns of traits within and between communities. A community with species possessing similar traits is usually inferred to result from environmental filtering, while coexisting species with different traits indicate interspecific competition or niche partitioning (Widenfalk et al. 2015, Widenfalk et al. 2016). Soil ecologists have recently adopted the concept of functional traits, in addition to simple species identity, to investigate belowground community structure and its association with the environment (Vandewalle et al. 2010, Pey et al. 2014, Moretti et al. 2017). Commonly used functional traits of soil invertebrates include morphological characters (e.g., body size, pigmentation and eye morphology), life history (or performance) traits (e.g., reproductive mode and fitness), physiological traits (e.g., metabolic rate and desiccation resistance), behavioral traits (e.g., dispersal mode), as well as ecological preferences that interrelate with other traits (Pey et al. 2014, Moretti et al. 2017). Using functional traits has been shown to be more powerful than simply using species identity for predicting the environmental associations of communities in soil (Makkonen et al. 2011, Bokhorst et al. 2012). Therefore, trait-based approaches are increasingly adopted in analyzing soil communities.

However, not all traits relevant to assembly processes can be measured in soil invertebrates (Moretti et al. 2017). Given that functional traits exhibit phylogenetic signal, phylogenetic information about species can be used as a surrogate for functional traits to infer assembly processes (Kembel 2009, Mouquet et al. 2012, Cadotte et al. 2013). Soil invertebrates likely exhibit phylogenetic signal (Pachl et al. 2012, Ponge and Salmon 2013, Potapov et al. 2016, Malcicka et al. 2017), and therefore, in this thesis I adopt the methods of community phylogeny which have been developed among plant ecologists (Webb et al. 2002, Cavender-Bares et al. 2009) and applied these methods to the studies on soil animal communities.

Community phylogenetic patterns

Community phylogenetic approaches explore the phylogenetic **patterns** of local communities to investigate the relative contribution of different processes to community assembly (Webb et al. 2002, Cavender-Bares et al. 2009). The essence of the community phylogenetic approach is to compare the observed pattern of phylogenetic distances between species in local communities with that derived from null model communities by randomly drawing species from a pre-defined species pool. If the assumption of phylogenetic conservatism of ecologically relevant traits is accepted, a community composed of phylogenetically closely related species can be inferred to be structured by environmental filtering. The environment may select for species possessing certain conserved traits that cope with specific abiotic conditions. In contrast, low relatedness among coexisting species reflected in each species having different traits points to the dominance of competitive interactions, as species compete for the same resources thereby limiting local coexistence or occupying different niches (Webb et al. 2002, Cavender-Bares et al. 2009, but see Gerhold et al. 2015).

Furthermore, community phylogenetic approaches are sensitive to spatial and taxonomic scales and the definition of species pool (Cavender-Bares et al. 2006, Swenson et al. 2006, Cavender-Bares et al. 2009). At larger spatial scales, environmental filtering influences local communities more strongly than species interactions, while the species interactions are more influential at finer taxonomic or spatial scales (Cavender-Bares et al. 2006, Swenson et al. 2006, Cavender-Bares et al. 2009). Therefore, using different specific pool definitions at different scales may help to gain deeper insight into the processes working at different spatial, temporal and taxonomic levels (Swenson et al. 2006, Emerson and Gillepsie 2008, Lessard et al. 2012).

Community phylogenetic approaches have shown their value for inferring assembly processes in aboveground communities inhabiting various environments characterized by disturbance regimes or harshness. For example, phylogenetic clustering in communities of plants (Webb 2000, Dinnage 2009, Ding et al. 2012), birds (Gianuca et al. 2014), amphibians (Brum et al. 2013) and bees (Pellissier et al. 2013, Sydenham et al. 2016) indicates that intensive disturbance and harshness in environments work predominantly as filtering processes leading to the coexistence of phylogenetically closely related species possessing similar disturbance-adapted traits (Ding et al. 2012, Gianuca et al. 2014). However, community phylogenetic approaches have not yet been commonly applied to soil biota [but see Bässler et al. (2014) and Thorn et al. (2016) for fungi; Li et al. (2014) for nematodes; Hausberger and Korb (2015) and Hausberger and Korb (2016) for termites; Andújar et al. (2015) and Thorn et al. (2016) for beetles], despite the exceptional diversity of soil communities and the varied traits possessed by different species.

A conceptual model for soil biodiversity studies

In this thesis, I use the conceptual model of Emerson and Gillespie (2008) that considers (1) **evolution of species traits**, (2) **trait similarity** and (3) **phylogenetic relatedness** between coexisting species (**Figure 1.1**). Patterns of these three measurable elements together produce scenarios of coexisting species that situate somewhere between the four extreme cases:

(a) Conserved traits + similar states + closely related species;

(b) Conserved traits + different states + distantly related species;

(c) Divergent traits + different states + closely related species;

(d) Convergent traits + similar states + distantly related species.

Here, conserved, divergent and convergent indicate how traits have evolved during species diversification, while similar or different indicate the differences in trait states between coexisting species.

Various processes can produce the above-mentioned patterns. In **case (a)** local species possess similar traits that have evolved in a conserved manner, and the local species are phylogenetically close relatives. This suggests that environmental filtering is the predominant process by selecting species that possess certain traits and therefore resulting in similar traits between coexisting species. The strength of species dispersal is weak, compared to local environmental filtering. This scenario may also suggest sympatric speciation, since species within local communities belong to the same phylogenetic clades and each community is assembled of species from a single clade all of which inherit their traits from the common ancestor. **Case (b)** provides an example in which the local community is assembled from species of various phylogenetic clades each possessing different traits inherited from their ancestors. This suggests that contemporary competition drives species toward niche partitioning; the local species possess different traits. Here, dispersal is intermediate, allowing species from different phylogenetic clades to meet together. In **case (c)** species possess different traits but assemble from the same phylogenetic clades. Divergent traits resulting from species diversification may suggest historical competition, while different traits among coexisting species indicate niche partitioning. The dispersal ability of the species is likely limited. Finally, in **case (d)**, environmental filtering is likely a predominant process that selects local species with similar traits that evolved convergently, presumably reflecting past competition or adaptation. This scenario also points to substantial dispersal allowing species to colonize new habitats where selection is at work.

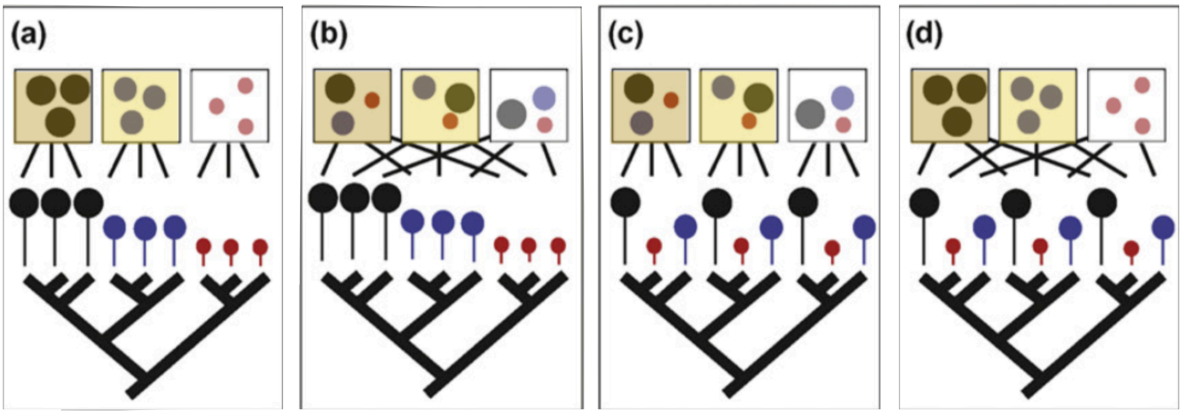

Figure 1.1
The conceptual model used in this thesis. Possible processes resulting in the four scenarios are (a) sympatric speciation + phylogenetic niche conservatism of trait + environmental filtering, (b) phylogenetic niche conservatism of trait + intermediate dispersal ability + contemporary competition, (c) historical competition + niche partitioning + low dispersal ability and (d) environmental filtering + convergent evolution of trait + substantial dispersal. Figures are modified from Emerson and Gillespie (2008). Squares represent local communities composed of species with different states of a trait (circles with different sizes and colors). Trait states are mapped on the phylogenetic tree, with connecting lines to the communities that indicate phylogenetic belonging of the species in the local communities.

Given that the target objects are known, the patterns are used to infer the processes. This thesis aims at integrating community phylogenetic approaches with phylogenetic comparative and trait-based methods, to study assembly processes in one of the most widespread and abundant soil arthropods, Collembola.

Collembola—the model organisms

Collembola (springtails) are early-derived Hexapoda characterized among others by a "jumping organ" (furca) at the ventral side of abdomen. The evolution of the furca likely contributed to species diversification (Beutel et al. 2017). The furca presumably evolved to escape predators, but it is reduced or lost in some soil-dwelling species. Another feature that characterizes Collembola is the ventral tube, from which Collembola got their name from Greek *colla* (glue) and *embolus* (piston). The ventral tube functions in fluid balance and is presumably important because the habitats of Collembola are usually moist, such as soil, tree canopies, aquatic surfaces, glaciers and caves (Hopkin 1997).

In soil, Collembola are among the most abundant and diverse mesofauna. Their density may reach 10,000–100,000 individuals m^{-2}, with a typical species richness of 60–80 species in temperate deciduous forests (Petersen and Luxton 1982). They are decomposers, feeding mainly on fungi and decaying organic matter but may also function as omnivores consuming a variety of food resources (Scheu and Falca 2000, Berg et al. 2004, Chahartaghi et al. 2005, Ferlian et al. 2015). They regulate microbial activity and therefore contribute to decomposition processes and nutrient cycling (Petersen and Luxton 1982, Rusek 1998, Schaefer et al. 2009).

Community compositions of Collembola typically vary with habitat characteristics. Different Collembola communities have been reported inhabiting arable fields, grasslands and forests, three types of habitats dominating mosaic landscapes in Central Europe (Ponge et al. 2003, Sousa et al. 2006, Martins da Silva et al. 2012, Heiniger et al. 2014). This thesis aims at investigating soil Collembola communities sampled from habitats characterized by different disturbance regimes by integrating phylogenetic comparative methods, community phylogenetic and trait-based approaches. The challenges in applying such approaches include the lack of knowledge on phylogenetic relationships of Collembola. Also, traits relevant to processes need to be identified and phylogenetic signal of traits needs to be tested.

Collembola phylogeny

The first fossil record of Collembola is from the Devonian, ca. 400 million years ago (Hirst and Maulik 1926, Whalley and Jarzembowski 1981). Today about 8,600 Collembola species have been described and assigned to four orders, i.e., Poduromorpha, Entomobryomorpha, Neelipleona and Symphypleona (Bellinger et al. 1996–2017; www.collembola.org), with their phylogenetic relationships still in debate. Molecular phylogenetic studies suggest that Symphypleona are sister to the other Collembola taxa, either paraphyletic (D'Haese 2002, Luan et al. 2005) or monophyletic (Xiong et al. 2008, Schneider et al. 2011, Yu et al. 2016) and that Entomobryomorpha are paraphyletic due to the position of Tomoceridae (Schneider et al. 2011). Furthermore, the phylogenetic position of Neelipleona is not resolved, although they are likely more closely related to Entomobryomorpha than to Symphypleona (Deharveng 2004).

However, previous phylogenetic studies of Collembola either were based on few genetic markers (D'Haese 2002) or few taxa (Xiong et al. 2008), or focused on single taxonomic groups (Schneider et al. 2011, Yu et al. 2016). Different marker selection, unbalanced taxon sampling and choice of outgroup taxa may lead to inconsistency in phylogenetic inference (Heath et al. 2008, Rosenfeld et al. 2012). Since phylogenetic relationships of Collembola species are still not resolved, in the thesis I first construct phylogenetic trees for locally occurring Collembola species sampled from arable fields, grasslands and forests, and then apply the trees in the subsequent studies on trait evolution and community phylogeny.

Collembola traits

Collembola are among the few soil taxa for which trait databases are available (Vandewalle et al. 2010, Pey et al. 2014, Matty P. Berg, *unpublished data*). Collembola traits, such as reproductive modes and morphological characters, are associated with the environment (Makkonen et al. 2011, Bokhorst et al. 2012, Salmon and Ponge 2012, Salmon et al. 2014, Widenfalk et al. 2015). Spherical body shape, large body size, dark pigmentation and sexual reproduction are characteristic of species occurring in open habitats and at the soil surface, whereas small body size, lack of eye spots, pale color and asexual reproduction are typical traits of species inhabiting forests and living in soil (Salmon et al. 2014). That particular traits are associated with habitats suggest that community assembly processes are driven by selection processes (Vellend 2016), e.g. environmental filtering or interspecific competition (Widenfalk et al. 2015, Widenfalk et al. 2016).

In this thesis, I use body shape, body length, pigmentation, number of ommatidia, vertical stratification and reproductive mode as traits of Collembola. Phylogenetic signal of these traits is analyzed using comparative methods. To link traits that presumably underlie contemporary assembly processes to evolutionary processes, I estimate ancestral character states of the traits and compare models of trait evolution. Furthermore, I propose neutral lipid fatty acid composition as a trait that can reflect both food resources and physiological attributes of different species.

Fatty acid composition as a trait

Neutral lipid fatty acids (NLFAs) in fat deposits of consumers are commonly used to identify diets of soil animals (Ruess and Chamberlain 2010, Ferlian et al. 2015). Specific NLFAs directly incorporated from food resources without modification in consumers carry signatures of food resources ('dietary routing'). Thus, these NLFAs are used as biomarkers to identify the origin of food. As omnivores in soil, Collembola consume a wide range of food resources including detritus, roots and root exudates, bacteria, fungi and algae (Hopkin 1997), which are difficult to measure directly in the field. By examining biomarker fractions of NLFAs in Collembola, the dietary routes can be identified (Chamberlain et al. 2005, Ferlian et al. 2015). Proportions of different biomarker NLFAs may be considered as a trophic trait , i.e., an α niche trait that reflects food resources.

Other NLFAs are synthesized or modified by consumers from precursors and then further integrated into other compounds. For example, long-chain polyunsaturated fatty acids are essential for biosynthesis of prostaglandins and eicosanoids, which are associated with reproduction, immune response and temperature regulation (Chamberlain et al. 2004, Chamberlain and Black 2005, Haubert et al. 2008, Ruess and Chamberlain 2010). Proportions of these fatty acids may thus represent physiological attributes, i.e., β niche traits that reflect environmental requirements of species.

In this thesis, I analyze NLFA profiles of field-sampled Collembola. A further compilation of literature-reported NLFA profiles in other Collembola species allows phylogenetic signal to be measured in a broader context.

Case studies using Collembola

To conduct community phylogenetic analyses for studying assembly processes of Collembola communities, three datasets are needed: a phylogenetic tree, a trait matrix and a species-site (community) matrix. Assembly processes are inferred from the patterns of phylogenetic relatedness and trait similarity of the coexisting species in local communities (an α-diversity approach; Webb et al. 2002). The observed patterns in communities are compared with those generated by null models, where species are randomly selected from a pre-defined species pool. The species pool can be defined in various ways, for example, at different taxonomic levels or for species with certain traits e.g., vertical stratification in soil profile. In this thesis I focus on the comparison of species coexistence patterns between Collembola communities inhabiting arable fields, grasslands and forests. Since these habitats are characterized by distinct disturbance regimes and vegetation, community assembly processes in soil are likely different between habitats. In a disturbed habitat, Collembola assemblages are likely to be driven by environmental filtering, resulting in similar traits and phylogenetically related species coexisting locally (Widenfalk et al. 2015). In contrast, different traits of distant relatives coexisting in a stable environment suggest that communities are predominantly influenced by niche partitioning or interspecific competition (Widenfalk et al. 2016).

I also use a metacommunity phylogenetic approach (phylogenetic β-diversity; Graham and Fine 2008). In distinction to traditional β-diversity studies where species are treated independently, phylogenetic β-diversity considers similarity in traits and phylogeny between species when exploring dissimilarities between local communities thus allowing regional or evolutionary processes to be connected to local processes, such as environmental filtering and interspecific competition (Graham and Fine 2008). Accounting for non-independence between coexisting species in the traits and phylogeny can help to investigate community-environmental associations from an evolutionary perspective (Pillar and Duarte 2010, Duarte 2011, Duarte et al. 2016).

Furthermore, inferring processes from patterns may benefit from experimental manipulations, compensating for the insufficiency of process inference in pure observational studies (Weber and Agrawal 2012). Controlling for factors that potentially influence community assembly is a necessity if processes are to be explicitly tested. In this respect, soil animal communities are more easy to manipulate than those of higher plants or vertebrates, for example, by defaunation (removing animals from soil) followed by the translocation of soil blocks (Ponge et al. 2008, Auclerc et al. 2009, Heiniger et al. 2015). Taking the data of Collembola communities from the experiment conducted by Auclerc et al. (2009) that aimed at ascribing species to different groups of dispersal ability and habitat preferences, I reanalyze species compositions using (meta)community phylogenetic approaches, to understand assembly processes that result from dispersal, are determined by soil properties, and differ at different successional stages.

Structure of the thesis

The aim of this thesis is to investigate assembly processes of soil Collembola communities in different types of habitats using phylogenetic comparative methods, community phylogenetic approaches and trait-based analyses. The general hypothesis tested here is that species in disturbed habitats are determined by environmental filtering that results in phylogenetic clustering and similar traits, while in relatively stable habitats interspecific competition or niche partitioning is predominant that leads to phylogenetic overdispersion with species possessing different traits, given that traits exhibit phylogenetic signal.

In **Chapter 2**, I construct Collembola phylogeny and use phylogenetic comparative methods to explore phylogenetic signal, model of evolution and ancestral state for a variety of traits, including body shape, body length, pigmentation, number of ommatidia, vertical stratification and reproductive mode. In **Chapter 3**, fatty acid composition is considered as a trait of Collembola and its phylogenetic signal is measured. **Chapter 4** aims at inferring community assembly processes of Collembola inhabiting arable fields, grasslands and forests using community phylogenetic approaches. Further in **Chapter 5**, I test for phylogenetic signal in habitat/soil preferences and dispersal abilities of Collembola species reported in Auclerc et al. (2009), and with a specific focus on assembly processes I explore the phylogenetic patterns in communities derived from the experimental designs. In **Chapter 6**, I ascribe the patterns of Collembola communities found in this thesis to the scenarios in the conceptual model (**Figure 1.1**), discuss the likely processes referring to the four high-level processes in *The Theory of Ecological Communities* (Vellend 2010, 2016) including selection, dispersal, drift and speciation, and propose a roadmap for soil ecologists to integrate phylogenetic comparative methods, community phylogenetic analyses and trait-based approaches in studies on the assembly processes of soil communities.

References

Ackerly, D.D. & Cornwell, W.K. (2007) A trait-based approach to community assembly: Partitioning of species trait values into within- and among-community components. *Ecology Letters*, **10**, 135–145.

Ackerly, D.D., Schwilk, D.W. & Webb, C.O. (2006) Niche evolution and adaptive radiation: testing the order of trait divergence. *Ecology*, **87**, S50-61.

Adler, P.B., Fajardo, A., Kleinhesselink, A.R. & Kraft, N.J.B. (2013) Trait-based tests of coexistence mechanisms. *Ecology Letters*, **16**, 1294–1306.

Anderson, J.M. (1975) The enigma of soil animal species diversity. Progress in Soil Zoology: Proceedings of the 5th International Colloquium on Soil Zoology (ed J. Vaněk), pp. 51–58. Springer Netherlands, Dordrecht.

Andújar, C., Arribas, P., Ruzicka, F., Platt, A.C., Timmermans, M.J.T.N. & Vogler, A.P. (2015) Phylogenetic community ecology of soil biodiversity using mitochondrial metagenomics. *Molecular Ecology*, **24**, 3603–3617.

Auclerc, A., Ponge, J.-F., Barot, S. & Dubs, F. (2009) Experimental assessment of habitat preference and dispersal ability of soil springtails. *Soil Biology and Biochemistry*, **41**, 1596–1604.

Bässler, C., Ernst, R., Cadotte, M., Heibl, C. & Müller, J. (2014) Near-to-nature logging influences fungal community assembly processes in a temperate forest (ed J Barlow). *Journal of Applied Ecology*, **51**, 939–948.

Bellinger, P.F., Christiansen, K.A. & Janssens, F. Checklist of the Collembola of the world (www.collembola.org).

Berg, M.P., Stoffer, M. & van den Heuvel, H.H. (2004) Feeding guilds in Collembola based on digestive enzymes. *Pedobiologia*, **48**, 589–601.

Best, R.J. & Stachowicz, J.J. (2013) Phylogeny as a proxy for ecology in seagrass amphipods: which traits are most conserved? *PLoS ONE*, **8**, e57550.

Beutel, R.G., Yavorskaya, M., Mashimo, Y., Fukui, M. & Meusemann, K. (2017) The phylogeny of Hexapoda (Arthropoda) and the evolution of megadiversity. *Proceedings of the Arthropodan Embryological Society of Japan*, **51**, 1–15.

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.

Brum, F.T., Gonçalves, L.O., Cappelatti, L., Carlucci, M.B., Debastiani, V.J., Salengue, E.V., dos Santos Seger, G.D., Both, C., Bernardo-Silva, J.S., Loyola, R.D. & da Silva Duarte, L. (2013) Land use explains the distribution of threatened New World amphibians better than climate. *PLoS ONE*, **8**, 4–11.

Cadotte, M., Albert, C.H. & Walker, S.C. (2013) The ecology of differences: assessing community assembly with trait and evolutionary distances. *Ecology Letters*, **16**, 1234–1244.

Caruso, T., Chan, Y., Lacap, D.C., Lau, M.C., McKay, C.P. & Pointing, S.B. (2011) Stochastic and deterministic processes interact in the assembly of desert microbial communities on a global scale. *The ISME Journal*, **5**, 1406–1413.

Caruso, T., Taormina, M. & Migliorini, M. (2012) Relative role of deterministic and stochastic determinants of soil animal community: a spatially explicit analysis of oribatid mites. *The Journal of Animal Ecology*, **81**, 214–221.

Cavender-Bares, J., Keen, A. & Miles, B. (2006) Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology*, **87**, S109-22.

Cavender-Bares, J., Kozak, K.H., Fine, P. V & Kembel, S.W. (2009) The merging of community ecology and phylogenetic biology. *Ecology Letters*, **12**, 693–715.

Chahartaghi, M., Langel, R., Scheu, S. & Ruess, L. (2005) Feeding guilds in Collembola based on nitrogen stable isotope ratios. *Soil Biology and Biochemistry*, **37**, 1718–1725.

Chamberlain, P.M. & Black, H.I.J. (2005) Fatty acid compositions of Collembola: unusually high proportions of C20 polyunsaturated fatty acids in a terrestrial invertebrate. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, **140**, 299–307.

Chamberlain, P.M., Bull, I.D., Black, H.I., Ineson, P. & Evershed, R.P. (2004) Lipid content and carbon assimilation in Collembola: implications for the use of compound-specific carbon isotope analysis in animal dietary studies. *Oecologia*, **139**, 325–335.

Chamberlain, P.M., Bull, I.D., Black, H.I.J., Ineson, P. & Evershed, R.P. (2005) Fatty acid composition and change in Collembola fed differing diets: identification of trophic biomarkers. *Soil Biology and Biochemistry*, **37**, 1608–1624.

Chase, J.M. & Leibold, M.A. (2003) Ecological Niches: Linking Classical and Contemporary Approaches. University of Chicago Press, Chicago, IL.

D'Haese, C.A. (2002) Were the first springtails semi-aquatic? A phylogenetic approach by means of 28S rDNA and optimization alignment. *Proceedings of the Royal Society B*, **269**, 1143–1151.

Deharveng, L. (2004) Recent advances in Collembola systematics. *Pedobiologia*, **48**, 415–433.

Ding, Y., Zang, R., Letcher, S.G., Liu, S. & He, F. (2012) Disturbance regime changes the trait distribution, phylogenetic structure and community assembly of tropical rain forests. *Oikos*, **121**, 1263–1270.

Dinnage, R. (2009) Disturbance alters the phylogenetic composition and structure of plant communities in an old field system. *PLoS ONE*, **4**, e7071.

Duarte, L.D.S. (2011) Phylogenetic habitat filtering influences forest nucleation in grasslands. *Oikos*, **120**, 208–215.

Duarte, L.D.S., Debastiani, V.J., Freitas, A.V.L. & Pillar, V.D. (2016) Dissecting phylogenetic fuzzy weighting: theory and application in metacommunity phylogenetics. *Methods in Ecology and Evolution*, **7**, 937–946.

Eissfeller, V., Langenbruch, C., Jacob, A., Maraun, M. & Scheu, S. (2013) Tree identity surpasses tree diversity in affecting the community structure of oribatid mites (Oribatida) of deciduous temperate forests. *Soil Biology and Biochemistry*, **63**, 154–162.

Emerson, B.C. & Gillespie, R.G. (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology & Evolution*, **23**, 619–630.

Ferlian, O., Klarner, B., Langeneckert, A.E. & Scheu, S. (2015) Trophic niche differentiation and utilisation of food resources in collembolans based on complementary analyses of fatty acids and stable isotopes. *Soil Biology and Biochemistry*, **82**, 28–35.

Gerhold, P., Cahill, J.F., Winter, M., Bartish, I. V & Prinzing, A. (2015) Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Functional Ecology*, **29**, 600–614.

Gianuca, A.T., Dias, R.A., Debastiani, V.J. & Duarte, L.D.S. (2014) Habitat filtering influences the phylogenetic structure of avian communities across a coastal gradient in southern Brazil. *Austral Ecology*, **39**, 29–38.

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

Graham, C.H. & Fine, P. V. (2008) Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecology Letters*, **11**, 1265–1277.

Haubert, D., Häggblom, M.M., Scheu, S. & Ruess, L. (2008) Effects of temperature and life stage on the fatty acid composition of Collembola. *European Journal of Soil Biology*, **44**, 213–219.

Hausberger, B. & Korb, J. (2015) A phylogenetic community approach for studying termite communities in a West African savannah. *Biology Letters*, **11**, 20150625.

Hausberger, B. & Korb, J. (2016) The impact of anthropogenic disturbance on assembly patterns of termite communities. *Biotropica*, **48**, 356–364.

Heath, T.A., Hedtke, S.M. & Hillis, D.M. (2008) Taxon sampling and the accuracy of phylogenetic analyses. *Journal of Systematics and Evolution*, **46**, 239–257.

Heiniger, C., Barot, S., Ponge, J.-F., Salmon, S., Botton-Divet, L., Carmignac, D. & Dubs, F. (2014) Effect of habitat spatiotemporal structure on collembolan diversity. *Pedobiologia*, **57**, 103–117.

Heiniger, C., Barot, S., Ponge, J.-F., Salmon, S., Meriguet, J., Carmignac, D., Suillerot, M. & Dubs, F. (2015) Collembolan preferences for soil and microclimate in forest and pasture communities. *Soil Biology and Biochemistry*, **86**, 181–192.

Hirst, S. & Maulik, S. (1926) On some arthropod remains from the Rhynie Chert (old red sandstone). *Geological Magazine*, **63**, 69–71.

Hopkin, S.P. (1997) Biology of the Springtails (Insecta: Collembola). Oxford University Press, Oxford.

Hubbell, S.P. (2001) The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press, Princeton and Oxford.

Ingimarsdóttir, M., Caruso, T., Ripa, J., Magnusdottir, O.B., Migliorini, M. & Hedlund, K. (2012) Primary assembly of soil communities: disentangling the effect of dispersal and local environment. *Oecologia*, **170**, 745–754.

Kembel, S.W. (2009) Disentangling niche and neutral influences on community assembly: assessing the performance of community phylogenetic structure tests. *Ecology letters*, **12**, 949–60.

Lessard, J.-P., Belmaker, J., Myers, J.A., Chase, J.M. & Rahbek, C. (2012) Inferring local ecological processes amid species pool influences. *Trends in Ecology & Evolution*, **27**, 600–607.

Li, J., Li, S., Chen, Y., Jia, P., Hua, Z., Wang, S., Song, Y., Liao, B. & Shu, W. (2014) Phylogenetic structures of soil nematode communities along a successional gradient in an unreclaimed copper mine tailings site. *Soil Biology and Biochemistry*, **77**, 179–186.

Luan, Y., Mallatt, J.M., Xie, R.D., Yang, Y.M. & Yin, W. (2005) The phylogenetic positions of three basal-hexapod groups (Protura, Diplura, and Collembola) based on ribosomal RNA gene sequences. *Molecular Biology and Evolution*, **22**, 1579–1592.

Makkonen, M., Berg, M.P., van Hal, J.R., Callaghan, T. V., Press, M.C. & Aerts, R. (2011) Traits explain the responses of a sub-arctic Collembola community to climate manipulation. *Soil Biology and Biochemistry*, **43**, 377–384.

Malcicka, M., Berg, M.P. & Ellers, J. (2017) Ecomorphological adaptations in Collembola in relation to feeding strategies and microhabitat. *European Journal of Soil Biology*, **78**, 82–91.

Martins da Silva, P., Berg, M.P., Serrano, A.R.M., Dubs, F. & Sousa, J.P. (2012) Environmental factors at different spatial scales governing soil fauna community patterns in fragmented forests. *Landscape Ecology*, **27**, 1337–1349.

McGill, B.J., Enquist, B.J., Weiher, E. & Westoby, M. (2006) Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, **21**, 178–185.

Minor, M.A. (2011) Spatial patterns and local diversity in soil oribatid mites (Acari: Oribatida) in three pine plantation forests. *European Journal of Soil Biology*, **47**, 122–128.

Moretti, M., Dias, A.T.C., de Bello, F., Altermatt, F., Chown, S.L., Azcarate, F.M., Bell, J.R., Fournier, B., Hedde, M., Hortal, J., Ibanez, S., Öckinger, E., Sousa, J.P., Ellers, J. & Berg, M.P. (2017) Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. *Functional Ecology*, **31**, 558–567.

Mouquet, N., Devictor, V., Meynard, C.N., Munoz, F., Bersier, L.F., Chave, J., Couteron, P., Dalecky, A., Fontaine, C., Gravel, D., Hardy, O.J., Jabot, F., Lavergne, S., Leibold, M.A., Mouillot, D., Munkemuller, T., Pavoine, S., Prinzing, A., Rodrigues, A.S., Rohr, R.P., Thebault, E. & Thuiller, W. (2012) Ecophylogenetics: advances and perspectives. *Biological Reviews*, **87**, 769–785.

Pachl, P., Domes, K., Schulz, G., Norton, R., Scheu, S., Schaefer, I. & Maraun, M. (2012) Convergent evolution of defense mechanisms in oribatid mites (Acari, Oribatida) shows no "ghosts of predation past." *Molecular Phylogenetics and Evolution*, **65**, 412–420.

Pellissier, L., Pradervand, J.-N., Williams, P.H., Litsios, G., Cherix, D. & Guisan, A. (2013) Phylogenetic relatedness and proboscis length contribute to structuring bumblebee communities in the extremes of abiotic and biotic gradients (ed A Baselga). *Global Ecology and Biogeography*, **22**, 577–585.

Petersen, H. & Luxton, M. (1982) A comparative analysis of soil fauna populations and their role in decomposition processes. *Oikos*, **39**, 288–388.

Pey, B., Nahmani, J., Auclerc, A., Capowiez, Y., Cluzeau, D., Cortet, J., Decaëns, T., Deharveng, L., Dubs, F., Joimel, S., Briard, C., Grumiaux, F., Laporte, M.-A.A., Pasquet, A., Pelosi, C., Pernin, C., Ponge, J.-F., Salmon, S., Santorufo, L. & Hedde, M. (2014) Current use of and future needs for soil invertebrate functional traits in community ecology. *Basic and Applied Ecology*, **15**, 194–206.

Pillar, V.D. & Duarte, L.D.S. (2010) A framework for metacommunity analysis of phylogenetic structure. *Ecology Letters*, **13**, 587–596.

Ponge, J.-F., Gillet, S., Dubs, F., Fedoroff, E., Haese, L., Sousa, J.P. & Lavelle, P. (2003) Collembolan communities as bioindicators of land use intensification. *Soil Biology and Biochemistry*, **35**, 813–826.

Ponge, J.-F. & Salmon, S. (2013) Spatial and taxonomic correlates of species and species trait assemblages in soil invertebrate communities. *Pedobiologia*, **56**, 129–136.

Ponge, J.-F., Tully, T. & Gins, A. (2008) Short-term responses of two collembolan communities after abrupt environmental perturbation: A field experimental approach. *Pedobiologia*, **52**, 19–28.

Potapov, A.A., Semenina, E.E., Korotkevich, A.Y., Kuznetsova, N.A. & Tiunov, A. V. (2016) Connecting taxonomy and ecology: Trophic niches of collembolans as related to taxonomic identity and life forms. *Soil Biology and Biochemistry*, **101**, 20–31.

Ricklefs, R.E. (1987) Community diversity: Relative roles of local and regional processes. *Science*, **235**, 167–171.

Rosenfeld, J.A., Payne, A. & DeSalle, R. (2012) Random roots and lineage sorting. *Molecular Phylogenetics and Evolution*, **64**, 12–20.

Ruess, L. & Chamberlain, P.M. (2010) The fat that matters: Soil food web analysis using fatty acids and their carbon stable isotope signature. *Soil Biology and Biochemistry*, **42**, 1898–1910.

Rusek, J. (1998) Biodiversity of Collembola and their functional role in the ecosystem. *Biodiversity and Conservation*, **7**, 1207–1219.

Salmon, S. & Ponge, J.-F. (2012) Species traits and habitats in springtail communities: A regional scale study. *Pedobiologia*, **55**, 295–301.

Salmon, S., Ponge, J.-F., Gachet, S., Deharveng, L., Lefebvre, N. & Delabrosse, F. (2014) Linking species, traits and habitat characteristics of Collembola at European scale. *Soil Biology and Biochemistry*, **75**, 73–85.

Schaefer, M., Migge-Kleian, S. & Scheu, S. (2009) The role of soil fauna for decomposition of plant residues. *Functioning and Management of European Beech Ecosystems* (eds R. Brumme & P.K. Khanna), pp. 207–230. Springer Berlin Heidelberg, Berlin, Heidelberg.

Scheu, S., Albers, D., Alphei, J., Buryn, R., Klages, U., Migge, S., Platner, C. & Salamon, J.-A. (2003) The soil fauna community in pure and mixed stands of beech and spruce of different age: trophic structure and structuring forces. *Oikos*, **101**, 225–238.

Scheu, S. & Falca, M. (2000) The soil food web of two beech forests (*Fagus sylvatica*) of contrasting humus type: stable isotope analysis of a macro- and a mesofauna-dominated community. *Oecologia*, **123**, 285–296.

Scheu, S. & Schulz, E. (1996) Secondary succession, soil formation and development of a diverse community of oribatids and saprophagous soil macro-invertebrates. *Biodiversity and Conservation*, **5**, 235–250.

Schneider, C., Cruaud, C. & D'Haese, C.A. (2011) Unexpected diversity in Neelipleona revealed by molecular phylogeny approach (Hexapoda, Collembola). *Soil Organisms*, **83**, 383–398.

Silvertown, J., McConway, K., Gowing, D., Dodd, M., Fay, M.F., Joseph, J.A. & Dolphin, K. (2006) Absence of phylogenetic signal in the niche structure of meadow plant communities. *Proceedings of the Royal Society B*, **273**, 39–44.

Sousa, J.P., Bolger, T., da Gama, M.M., Lukkari, T., Ponge, J.-F., Simón, C., Traser, G., Vanbergen, A.J., Brennan, A., Dubs, F., Ivitis, E., Keating, A., Stofer, S. & Watt, A.D. (2006) Changes in Collembola richness and diversity along a gradient of land-use intensity: A pan European study. *Pedobiologia*, **50**, 147–156.

Swenson, N.G., Enquist, B.J., Pither, J., Thompson, J. & Zimmerman, J.K. (2006) The problem and promise of scale dependency in community phylogenetics. *Ecology*, **87**, 2418–2424.

Sydenham, M.A.K., Moe, S.R., Stanescu-Yadav, D.N., Totland, Ø. & Eldegard, K. (2016) The effects of habitat management on the species, phylogenetic and functional diversity of bees are modified by the environmental context. *Ecology and Evolution*, **6**, 961–973.

Thorn, S., Bässler, C., Bernhardt-Römermann, M., Cadotte, M., Heibl, C., Schäfer, H., Seibold, S. & Müller, J. (2016) Changes in the dominant assembly mechanism drives species loss caused by declining resources. *Ecology Letters*, **19**, 163–170.

Vandewalle, M., de Bello, F., Berg, M.P., Bolger, T., Dolédec, S., Dubs, F., Feld, C.K., Harrington, R., Harrison, P.A., Lavorel, S., Silva, P.M., Moretti, M., Niemelä, J., Santos, P., Sattler, T., Sousa, J.P., Sykes, M.T., Vanbergen, A.J. & Woodcock, B.A. (2010) Functional traits as indicators of biodiversity response to land use changes across ecosystems and organisms. *Biodiversity and Conservation*, **19**, 2921–2947.

Vellend, M. (2010) Conceptual synthesis in community ecology. *The Quarterly review of biology*, **85**, 183–206.

Vellend, M. (2016) The Theory of Ecological Communities. Princeton University Press, Princeton.

Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. (2007) Let the concept of trait be functional! *Oikos*, **116**, 882–892.

Webb, C.O. (2000) Exploring the Phylogenetic Structure of Ecological Communities: An Example for Rain Forest Trees. *The American Naturalist*, **156**, 145–155.

Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, **33**, 475–505.

Weber, M.G. & Agrawal, A. a. (2012) Phylogeny, ecology, and the coupling of comparative and experimental approaches. *Trends in Ecology & Evolution*, **27**, 394–403.

Whalley, P. & Jarzembowski, E.A. (1981) A new assessment of *Rhyniella*, the earliest known insect, from the Devonian of Rhynie, Scotland. *Nature*, **291**, 317.

Widenfalk, L.A., Bengtsson, J., Berggren, Å., Zwiggelaar, K., Spijkman, E., Huyer-Brugman, F. & Berg, M.P. (2015) Spatially structured environmental filtering of collembolan traits in late successional salt marsh vegetation. *Oecologia*, **179**, 537–549.

Widenfalk, L.A., Malmström, A., Berg, M.P. & Bengtsson, J. (2016) Small-scale Collembola community composition in a pine forest soil - Overdispersion in functional traits indicates the importance of species interactions. *Soil Biology and Biochemistry*, **103**, 52–62.

Xiong, Y., Gao, Y., Yin, W. & Luan, Y. (2008) Molecular phylogeny of Collembola inferred from ribosomal RNA genes. *Molecular phylogenetics and evolution*, **49**, 728–735.

Yu, D., Zhang, F., Stevens, M.I., Yan, Q., Liu, M. & Hu, F. (2016) New insight into the systematics of Tomoceridae (Hexapoda, Collembola) by integrating molecular and morphological evidence. *Zoologica Scripta*, **45**, 286–299.

Chapter 2

Collembola Phylogeny and Trait Evolution

Ting-Wen Chen, Jo-Fan Chao, Matty P. Berg, Ina Schaefer, Stefan Scheu

Abstract

Collembola (springtails) are among the most diverse soil mesofauna. As basal Hexapoda derived from a crustacean ancestor they diversified on land and occupied a wide variety of ecological niches, with distinct morphological and ecological traits among different taxonomic groups. However, information on trait evolution during Collembola diversification is scarce. At least in part this is due to the lack of knowledge on phylogeny across taxonomic levels. In this study, we first reconstructed phylogenetic trees of Collembola from various taxonomic groups that locally co-occur in arable fields, grasslands and forests. We then used phylogenetic comparative methods to investigate the evolution of functional traits of Collembola. Results demonstrate that body shape of Collembola evolved quickly early in their diversification but slowed down afterwards. In contrast, evolutionary transitions of pigmentation, number of ommatidia and reproductive mode correlated with vertical stratification of species living in soils. Ancestral traits of Collembola were likely slender body, hemiedaphic way of life, sexual reproduction, possession of many ommatidia and bright body color, but these traits presumably changed several times during diversification of species. The traits with phylogenetic signal can help to investigate assembly processes in soil animal communities using community phylogenetic approaches and our study paves the way for integrating of evolutionary approaches into soil ecological studies.

Keywords

ancestral state reconstruction; Brownian motion model; Early Burst model; evolutionary constraint; functional trait; life form; phylogenetic comparative method; phylogenetic signal; reproductive mode; springtail; soil

Introduction

Diversification processes of soil organisms are less investigated than aboveground biota, albeit the tremendous diversity of soil animals has puzzled ecologists for decades (Anderson 1975, Petersen and Luxton 1982, Giller 1996). Extant soil invertebrates include Nematoda, Annelida, Chelicerata, Myriapoda and Hexapoda that terrestrialized several times independently (von Reumont et al. 2012, Rota-Stabelli et al. 2013, Minter et al. 2017). Among these, Hexapoda derived from a crustacean ancestor with the fossil record dating back to the Devonian or earlier (Hirst and Maulik 1926, Whalley and Jarzembowski 1981, Rota-Stabelli et al. 2013, Wolfe et al. 2016) and rapidly diversified into various ecological niches. Today, Collembola (springtails) are the most abundant and diverse Hexapoda living in soil (Hopkin 1997).

About 8,600 species of Collembola have been described (Bellinger et al. 1996–2017; www.collembola.org). Systematically, they are classified into the four orders: Poduromorpha, Entomobryomorpha, Symphypleona and Neelipleona. Species of different orders are characterized by distinct morphological characters and usually prefer different ecological environments. For example, Entomobryomorpha are elongate in body shape, while Symphypleona are globular, and both usually live on the soil surface. Similarly, elongated Poduromorpha and globular Neelipleona predominantly dwell in the soil (Salmon et al. 2014). Morphological characters possessed by different Collembola species presumably are associated with adaptation to the habitat they colonize, and thus represent functional traits on which ecological processes can work (Violle et al. 2007, Pey et al. 2014).

However, traits associated with environmental factors are also shaped by evolutionary processes. As being inherited from a common ancestor, functional traits of species show phylogenetic signal, i.e., closely related species possess similar traits. As a result of different evolutionary mechanisms, a trait can be phylogenetically conserved (i.e., shaped by evolutionary constraints), convergent (i.e., evolved repeatedly in distantly related species) or labile. These evolutionary mechanisms can be inferred using phylogenetic comparative methods (Blomberg and Garland 2002, Losos 2008, Revell et al. 2008, Cooper et al. 2010) such as phylogenetic signal measurements, ancestral character state reconstruction and likelihood comparisons of evolutionary models. In this study we examined the patterns of evolution in Collembola functional traits including body length, body shape, intensity of pigmentation, number of ommatidia (eyes), vertical stratification in soil profile and reproductive mode—all associated with environmental factors (Widerfalk et al. 2015).

Although it appeals intuitively that variations in functional traits of different species have evolutionary bases, few studies have tested evolutionary hypotheses of traits in soil animals. Previous studies indicated that desiccation resistance of Isopoda (Dias et al. 2013) and defense mechanisms of Oribatida (Pachl et al. 2012) exhibit phylogenetic signal. In Collembola, ecological preferences (Ponge and Salmon 2013) and stable isotopic signatures (Potapov et al. 2016) are likely to exhibit phylogenetic signal. Recently, Chen et al. (2017) demonstrated that fatty acid profiles of Collembola exhibit phylogenetic signal. Furthermore, Malcicka et al. (2017) suggested that trophic guilds of Collembola and their mouthpart structures evolved in parallel. However, the results of these studies are limited, as they only used a single comparative method (Dias et al. 2013, Malcicka et al. 2017, Chen et al. 2017) or used taxonomy as proxy for phylogenetic relationships without basing them on phylogenetic trees inferred by genetic markers (Ponge and Salmon 2013, Potapov et al. 2016).

One challenge in applying phylogenetic comparative approaches to soil animals is the lack of phylogenetic information for most taxa. In Collembola phylogenetic relationships between the four orders are still debated. Symphypleona, either paraphyletic (D'Haese 2002, Luan et al. 2005) or monophyletic (Xiong et al. 2008, Schneider et al. 2011, Yu et al. 2016), are recovered as the sister group to the other Collembola. Species of Entomobryomorpha are paraphyletic, presumably due to the position of Tomoceridae. Furthermore, the phylogenetic position of Neelipleona is not resolved, although they are likely more closely related to Entomobryomorpha than to Symphypleona (Deharveng 2004). However, previous studies on Collembola phylogeny were based on few genetic markers (D'Haese 2002), few taxa (Xiong et al. 2008) or focused on a narrow spectrum of taxonomic groups (Frati and Carapelli 1999, Frati et al. 2000, Soto-Adames 2002, Burkhardt and Filser 2005, Park 2009, Cicconardi et al. 2010, Greenslade et al. 2011, Schneider et al. 2011, Cicconardi et al. 2013, Zhang et al. 2014, Yu et al. 2016). A phylogeny of Collembola including all major evolutionary lineages and various taxonomic levels is still lacking.

In this study, we reconstructed phylogenetic trees of Collembola from various taxonomic groups that locally co-occur in arable fields, grasslands and forests in Central Europe. We further included sequences deposited in GenBank for additional Central European species, to construct a more comprehensive phylogeny. Then, we mapped Collembola traits on the phylogenetic tree, measured phylogenetic signal, tested models of trait evolution and reconstructed ancestral states. We hypothesized that (1) Collembola functional traits show phylogenetic signal, with closely related species having similar traits, and that (2) the evolution of major traits of Collembola is correlated with the vertical stratification of species in soil, i.e. their depth distribution.

Materials and Methods

Taxa sampling

Collembola were sampled between March and June 2014 from arable fields, grasslands and forests at six sites near Göttingen, Germany (**Figure 2.1**, **Table S2.1**). In each of the arable fields and grasslands, one suction sample equal to a surface area of 154 cm^2 was taken for 10 s to collect surface-living individuals. Then, to sample soil-dwelling individuals, a soil core (5 cm diameter, 5 cm depth) was taken at the center of the area from which the suction sample was collected. In each forest, litter material (L and F layer) from an area of 154 cm^2 was sampled by hand, followed by a 10 sec suction sample of the humus layer (H layer). Animals from this suction sample later were added to the animals from the litter sample forming the full sample of organic layers. Further, to sample soil dwelling Collembola, a soil core (5 cm diameter, 5 cm depth) was taken at the center of the area sampled for Collembola in organic layers. Collembola from the suction samples from arable fields and grasslands were directly transferred into 96% ethanol, while those in the organic layers and soil cores were extracted by heat (Kempson et al. 1963), collected in water and then transferred into 96% ethanol every two days over a period of ten days. Samples were kept at 4°C until identification and then stored at -80°C. Collembola identification was based on Hopkin (2007), Fjellberg (1998, 2007) and Gisin (1960); nomenclature followed Bellinger et al. (1996–2017; www.collembola.org).

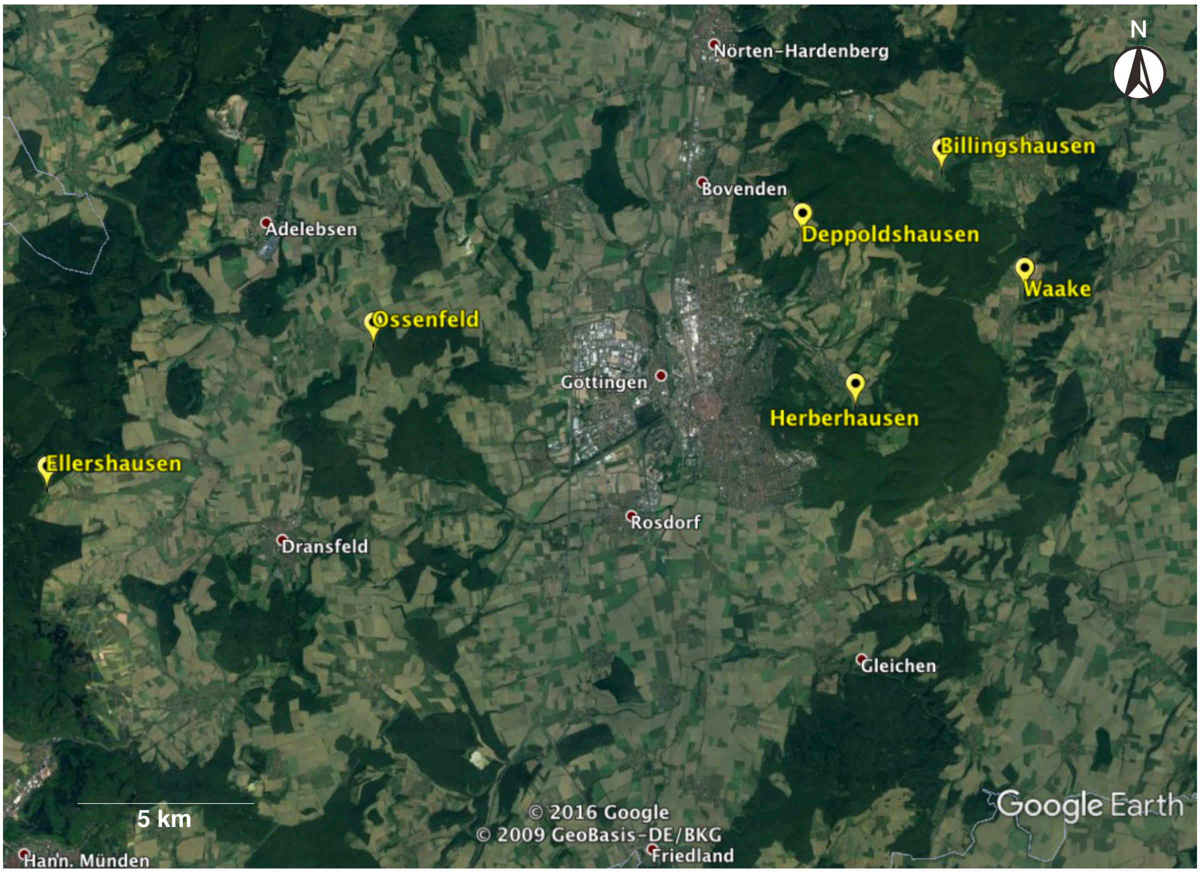

Figure 2.1
Sampling sites of this study. For more details see text and Table S2.1.

DNA extraction, PCR and sequencing

In total, 75 morphological species representing 51 genera and 18 families were collected from the study sites and used for phylogenetic reconstruction. To cover cryptic species, multiple individuals of each species were sequenced, if replicates were available. Genomic DNA from a single individual was extracted using the DNeasy® Blood and Tissue Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol for animal tissue. Two ribosomal genes, partial 18S rRNA (~680 bp; McGaughran et al. 2010) and partial 28S rRNA covering the D1-D2 region (~780 bp; D'Haese 2002) and D3-D5 region (~570 bp; Luan et al. 2005), and two protein coding genes, Histone H3 (H3; nuclear gene,

~372 bp; von Saltzwedel et al. 2016) and Cytochrome Oxidase I (COI; mitochondrial gene, ~681 bp; Schneider et al. 2011), were amplified separately in 25 µl volumes containing 12.5 µl SuperHot Taq Mastermix (Genaxxon Bioscience GmbH, Ulm, Germany), 1.5 µl of each primer (10 pM; **Table S2.2**), 2 µl $MgCl_2$ (25 mM) and 3–5 µl template DNA. PCR conditions for 18S rRNA, 28S rRNA and H3 included an initial activation step at 95°C for 15 min, followed by 35 amplification cycles (with a denaturation step at 94°C for 30 s, a primer-annealing step at the optimal temperature for 45 s and a elongation step at 72°C for 30 s; **Table S2.2**) and ended with a final elongation step at 72°C for 6 min. The PCR program for COI was identical to that used in Anslan and Tedersoo (2015). Positive PCR products were purified with the PCR DNA Purification Mini Prep Kit (Genaxxon Bioscience GmbH, Ulm, Germany) following the manufacturer's protocol and sent for sequencing to the Göttingen Genome Laboratory (Institute for Microbiology and Genetics, University of Göttingen, Germany). The obtained sequences were checked and ambiguous positions were corrected using Sequencher 5.0 (Gene Code Corporation, Ann Arbor, Michigan, USA) aided by chromatograms.

Phylogeny of field-collected Collembola

Individual-based unrooted phylogeny

First, individual-based phylogenetic trees were built to remove redundant sequences of the same species. Sequences of the five genetic markers were aligned separately in R 3.2.2 (R Core Team 2015) using the functions *AlignSeqs* and *AdjustAlignment* with the default parameter settings (package "DECIPHER"; Wright 2015). In each marker set, terminal gaps at the beginning and the end of sequences were replaced by "?". The alignments of the five markers were concatenated in a supermatrix (3,084 bp) using SequenceMatrix 1.8 (Vaidya et al. 2011). An unrooted phylogeny was inferred using Bayesian Inference (BI) in MrBayes 3.2.4 (Ronquist et al. 2012), setting the model of sequence evolution as GTR+I+G. Bayesian Inference was conducted by two independent runs of four chains, 5,000,000 generations, 0.02 temperature and 0.5 burn-in fraction; all other parameters remained set as default. The resulting consensus tree was checked to remove redundant sequences of the same species for the following species-based phylogenetic inference. Redundant sequences were excluded if genetic distances between individuals were less than 5% across all five markers. Different lineages of the same species were retained if they showed genetic distances more than 5% using the R function *otuPhylo* (Steven Kembel *pers. comm.*). Overall, a total of 102 operational taxonomic units (OTUs) of different species and within-species lineages were obtained for the species-based phylogenetic reconstruction (**Table S2.3**; Accession Number KY230697–KY231137).

Species-based phylogeny

Sequences of the 102 OTUs were aligned with the outgroup taxa Zygentoma (Insecta), *Machilis* (Insecta: Archaeognatha), *Callibaetis* (Insecta: Palaeoptera), *Baculentulus* (Protura), *Parajapyx* (Diplura) and

Speleonectes (Crustacea; **Table S2.3**) using the R functions mentioned above. For each genetic marker the aligned sequences were trimmed to the same length. The best model of sequence evolution for all tested markers were fitted with GTR+I+G according to Akaike information criterion (AIC) estimated in jModelTest 2.1.4 (Guindon and Gascuel 2003, Darriba et al. 2012). Terminal gaps of each alignment were replaced by "?", and a supermatrix including all genetic markers was generated using SequenceMatrix 1.8 (Vaidya et al. 2011), resulting in a total length of 3,650 bp of the alignment. Bayesian Inference (BI; MrBayes 3.2.4, Ronquist et al. 2012) was applied for phylogenetic reconstruction by setting the five markers (all GTR+I+G) unlinked, two independent runs, four chains, 10,000,000 generations, 0.05 temperature and 0.5 burn-in fraction. Other parameters were set as default. Furthermore, Collembola phylogeny was inferred using Maximum Likelihood (ML) algorithm in RAxML 7.0.3 (Stamatakis 2006) setting the GTR+I+G model and 1,000 bootstrap replicates. BI and ML trees were similar in topology and we continued the analyses of trait evolution with the BI tree.

Extended Collembola phylogeny

To construct a more comprehensive phylogeny, we further included sequences of the above-mentioned genes reported in the literature for other Collembola species with recorded Central European occurrence (Bellinger et al. 1996–2017; www.collembola.org). Here 242 OTUs of 167 species, 82 genera and 18 families were downloaded from NCBI, covering most of the common Collembola in Central Europe (**Table S2.4**). Since outgroups always resulted in polytomies in the backbone of the tree in preliminary trials, we constructed an unrooted phylogeny using the extended dataset. The downloaded sequences were aligned with the field-collected dataset following the steps mentioned above, except that D1 and D2 regions of 28S rRNA were aligned separately and the three codon positions of protein-coding genes (H3 and COI) were spilt into three alignments. The model of sequence evolution of each marker was estimated using jModelTest 2.1.4. Terminal gaps of each alignment were replaced by "?", and concatenated in a supermatrix. The total length of the alignment was 3,083 bp. Bayesian Inference was applied, setting the genetic markers unlinked and other parameters as default except for 10,000,000 generations, 0.02 temperature and 0.5 burn-in fraction.

Ultrametric tree transformation

The species-based BI tree was transformed to a chronogram (ultrametric tree) using a penalized likelihood approach by setting different models of substitution rate variation among branches, i.e., correlated, relaxed, discrete and strict clock models, using the function *chronos* implemented in the R package "ape" (Paradis et al. 2004). The most appropriate ultrametric tree was selected based on the smallest value among all PHIIC generated from different models (Paradis 2013). The ultrametric tree based on the strict clock model was selected. For the morphological species including different genetic lineages (cryptic species), only one OTU was retained. Traits of species were mapped onto the ultrametric phylogenetic tree using the function *plot.phylo* implemented in the R package “ape” (Paradis et al. 2004).

Collembola traits

Traits of the field-sampled Collembola species were extracted from a trait database (Matty P. Berg, *unpublished data*) compiled from literature, including maximum body length, overall body shape, intensity of pigmentation, number of ommatidia (eyes), vertical stratification in soil profile and reproductive mode. These traits are associated with environmental gradients or stress (Vandewalle et al. 2010, Makkonen et al. 2011, Salmon et al. 2014, Widenfalk et al. 2015). Collembola body length was used as continuous variable spanning from 0.4 to 6.5 mm, whereas body shape (slender, stocky, spheric), pigmentation (pale, bright, dark), number of ommatidia [many (6–8), few (1–5), none (0)], vertical stratification (epedaphic, hemiedaphic, euedaphic) and reproductive mode (parthenogenetic, bisexual) were used as categorical variables. Species occurrence in each of the three habitats was further included as a binary variable [0 (absence), 1 (presence)], and species logarithmic density summed across habitat types and sites was treated as further continuous variable (**Table S2.4**).

Phylogenetic signal

Phylogenetic signal in continuous, categorical and binary variables was analyzed using Blomberg's K (Blomberg et al. 2003), Pagel's lambda (Pagel 1999, Freckleton et al. 2002) and D statistic (Fritz and Purvis 2010), respectively. Blomberg's K was calculated using the function *phylosig* implemented in the R package “phytools” (Blomberg et al. 2003, Revell 2012). Significance tests were done by randomizing species on the phylogeny 1,000 times to test whether trait values showed phylogenetic signal or not (i.e., H0 = 0). In case of significant K-values of traits, the observed K-value was further compared with 10,000 simulated K-values to test whether phylogenetic signal significantly differed from the level expected under Brownian motion evolution model (i.e., H0 = 1; Revell et al. 2007). Simulations of trait values were conducted using the function *fastBM* in the R package “phytools” (Revell 2012). Lower and higher phylogenetic signal than predicted by a Brownian motion model was defined as a K-value in the 0.025 and 0.975 quantiles of the log-transformed simulated K-values, respectively. For categorical variables, Pagel's lambda was estimated using the function *fitDiscrete* implemented in R package “geiger” (Pagel 1999, Freckleton et al. 2002, Harmon et al. 2008). An appropriate trait evolution model was estimated from one of the equal-rates (ER), symmetric (SYM) and all-rates-different (ARD) models using likelihood comparison. The ER model was accepted for all categorical traits. Then, a star-like tree (lambda 0) was transformed from the original tree (lambda 1). Likelihoods of the distribution of trait states among species were compared, given the lambda 1 and lambda 0 trees. If the lambda 0 tree was accepted, the trait showed no phylogenetic signal (Pagel 1999, Freckleton et al. 2002). Phylogenetic signal in binary variables was measured using the D statistic with 10,000 permutations by the function *phylo.d* implemented in the R package "caper" (Fritz and Purvis 2010, Orme et al. 2013). A D-value smaller than zero indicated a trait with conserved phylogenetic signal, while a value greater than one suggested that a trait is divergent. The observed D-value was compared with simulated D-values generated by two models, phylogenetic randomness and Brownian threshold models (Fritz and Purvis 2010).

Ancestral state estimation

For continuous traits with phylogenetic signal, ancestral character states were estimated using Maximum Likelihood method by the function *fastAnc* and then followed by a traitgram plotted using the function *phenogram* implemented in the R package "phytools" (Revell 2012). For the categorical traits with phylogenetic signal, stochastic character mapping was used to estimate ancestral character states using the function *make.simmap* implemented in the R package "phytools" (Huelsenbeck et al. 2003, Bollback, 2006, Revell 2012). Trait state transition was assumed following the ER model. Prior distribution on root node was estimated from tip character states. Transition matrix Q was sampled 1,000 times from the posterior probability distribution using Bayesian MCMC. Then, 1,000 stochastic maps were simulated which were conditioned by the sampled value of Q. Numbers of trait state transitions in the tree were calculated based on the mean and median. Posterior probabilities of trait states were mapped to the tree nodes.

Evolution model of traits

For continuous traits, the fits of trait evolution under Brownian motion model (BM) and Ornstein-Uhlenbeck model (OU) were compared using the function *fitContinuous* implemented in the R package "geiger", while for discrete traits, the fits of trait evolution under the rate constancy model, white-noise model, Pagel's lambda model (lambda), time-dependent model (delta), Early-burst model (EB) and punctuational model (kappa) were compared using the function *fitDiscrete*. Akaike Weights (AIC_w) were used to estimate the support of each tested model to the trait states.

Correlated evolution of traits

To explore the evolution of traits Collembola body length was converted into a categorical variable and coded as small (0.4–1.5 mm), medium (1.6–3.1 mm) and large (> 3.2 mm). Then, all the categorical trait variables were coded as binary variables to estimate relationships of evolutionary changes between any two traits using Pagel's general method as implemented in the function *fitPagel* in the R package "phytools" (Pagel 1994, Revell 2012). The model of evolution for each trait was set as ER. Using likelihood ratio test, the independent model was compared with the dependent model. Three dependent models ("x", "y" and "xy") were used to explore the dependency of substitution rate of one variable on the other ("X" on "Y"), that of the opposite direction ("Y" on "X") and that in both directions ("X" on "Y" and "Y" on "X"; Pagel 1994). *P*-values were derived from 1,000 simulations and then adjusted using Benjamini and Hochberg corrections (BH; Benjamini and Hochberg 1995).

Results

Phylogeny of field-collected Collembola

Both BI tree and ML tree using the data on local Collembola species supported monophyletic origin of major taxonomic groups including Entomobryoidea, Isotomidae, Tomoceridae, Neelidae, Symphypleona and Poduromorpha. Entomobryoidea was sister to all other Collembola lineages. Isotomidae was sister to the remaining taxa, while Poduromorpha formed the sister lineage to Symphypleona. Neelidae and Tomoceridae formed one monophyletic clade positioned between Isotomidae and the clade comprising Poduromorpha and Symphypleona, although this positioning was based on weak support. Within Poduromorpha, Neanuridae plus Hypogastruridae and Brachystomellidae formed a monophyletic clade, sister to the clade composed of Tullbergiidae, Onychiuridae and Odontellidae. Notably, monophyly of Hypogastruridae was not supported, and Brachystomellidae was nested within Neanuridae. Symphypleona comprised two clades, one including Sminthurididae, Arrhopalitidae and Katiannidae, and the other including Bourletiellidae and Sminthuridae (with *Sminthurides parvulus*, however). The position of Dicyrtomidae in the BI tree (sister to Bourletiellidae + Sminthuridae) differed from that in the ML tree (sister to Sminthurididae + Arrhopalitidae + Katiannidae). In Isotomidae most genera formed monophyletic clades including *Parisotoma, Isotoma, Vertagopus* and *Folsomia*; however, species of the genus *Desoria* were separated, with *D. violacea* Fjellberg, 1979 forming the sister taxon of *Vertagopus* and *D. trispinata* forming the sister taxon of *Parisotoma*. Within Entomobryoidea, Orchesellidae was sister group to Lepidocyrtidae + Entomobryidae (**Figure 2.2**).

In addition to the position of Dicyrtomidae, the topology of the ML tree differed from that of the BI tree in the relationships within Tullbergiidae, relationship of Brachystomellidae with the other clades, monophyly of *Neelus murinus*, the position of *Isotomodes productus* within Isotomidae and that of *Pseudosinella* within Lepidocyrtidae (**Figure S2.1**).

Phylogeny of species occurring in Central Europe

Including the sequences of the other European Collembola species, the phylogenetic tree showed several monophyletic groups: Poduromorpha, Symphypleona, Tomoceridae, Neelipleona, Isotomidae, and Entomobryoidea. Within Poduromorpha, except *Triacanthella perfecta* (Hypogastruridae) of which the sequence was taken from GenBank, the monophyly of Onychiuroidea was supported and included three monophyletic families, Onychiuridae, Odontellidae and Tullbergiidae. Odontellidae was the sister group to Tullbergiidae, while Onychiuridae was the sister group to Odontellidae plus Tullbergiidae. Relationships of the remaining Poduromorpha were complex. Both Hypogastruridae and Neanuridae were not monophyletic, while Poduridae and Brachystomellidae were close to Neanuridae. Symphypleona was the sister group to Poduromorpha. In Symphypleona, Sminthurididae formed the sister clade to Katiannoidea which composed of Arrhopalitidae and Katiannidae. Dicyrtomidae was the sister group to

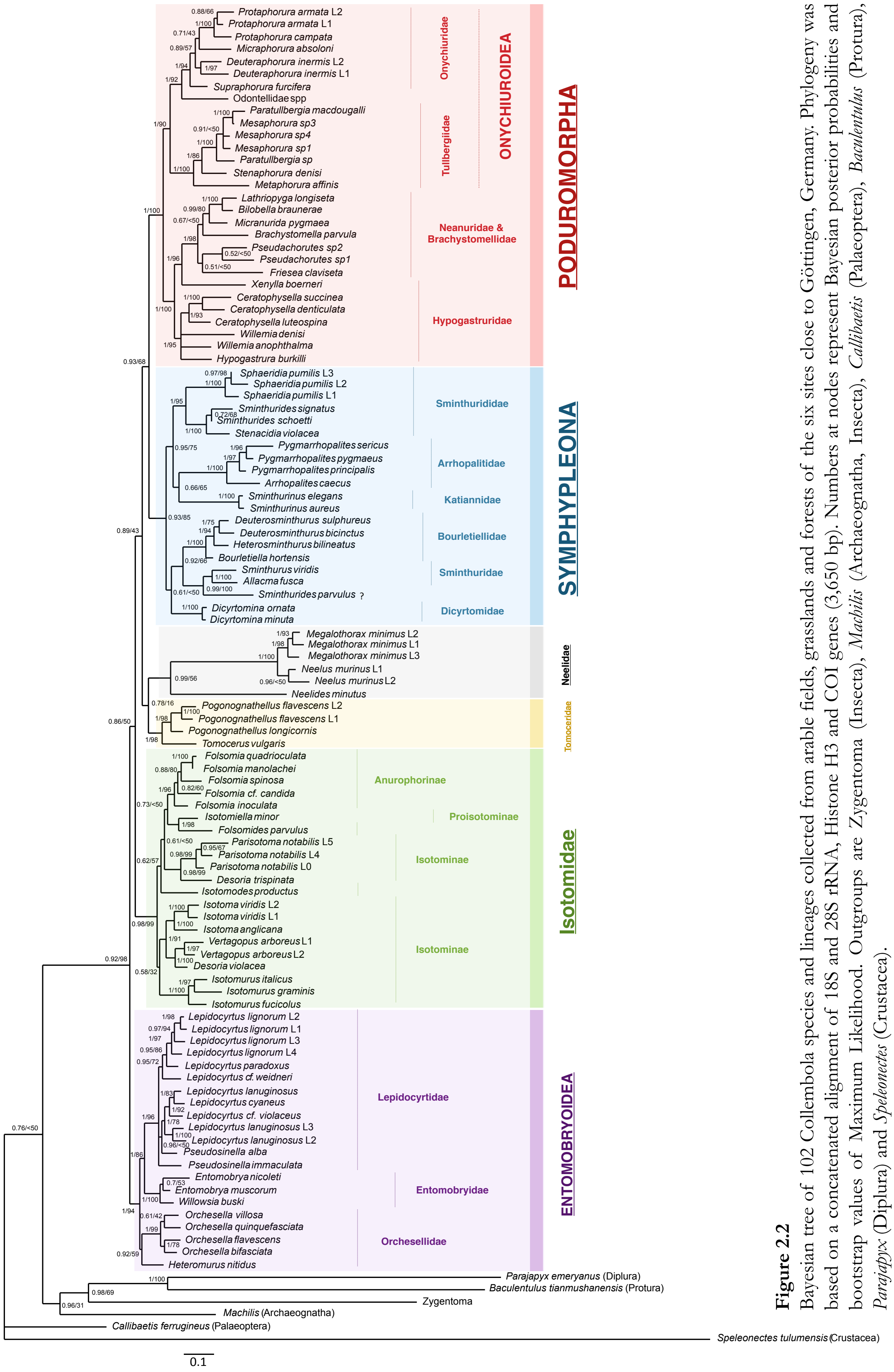

Figure 2.2
Bayesian tree of 102 Collembola species and lineages collected from arable fields, grasslands and forests of the six sites close to Göttingen, Germany. Phylogeny was based on a concatenated alignment of 18S and 28S rRNA, Histone H3 and COI genes (3,650 bp). Numbers at nodes represent Bayesian posterior probabilities and bootstrap values of Maximum Likelihood. Outgroups are Zygentoma (Insecta), *Machilis* (Archaeognatha, Insecta), *Callibaetis* (Palaeoptera), *Baculentulus* (Protura), *Parajapyx* (Diplura) and *Speleonectes* (Crustacea).

Sminthuroidea which included Sminthuridae and Bourletiellidae. Isotomidae seemed closely related to Entomobryoidea. In Isotomidae, *Archisotoma* was sister to the other species. Within Entomobryoidea, Entomobryidae and Lepidocyrtidae were sister to each other, forming a monophyletic group sister to Orchesellidae. However, relationships of Tomoceridae and Neelipleona to the other Collembola were not resolved (**Figure 2.3**).

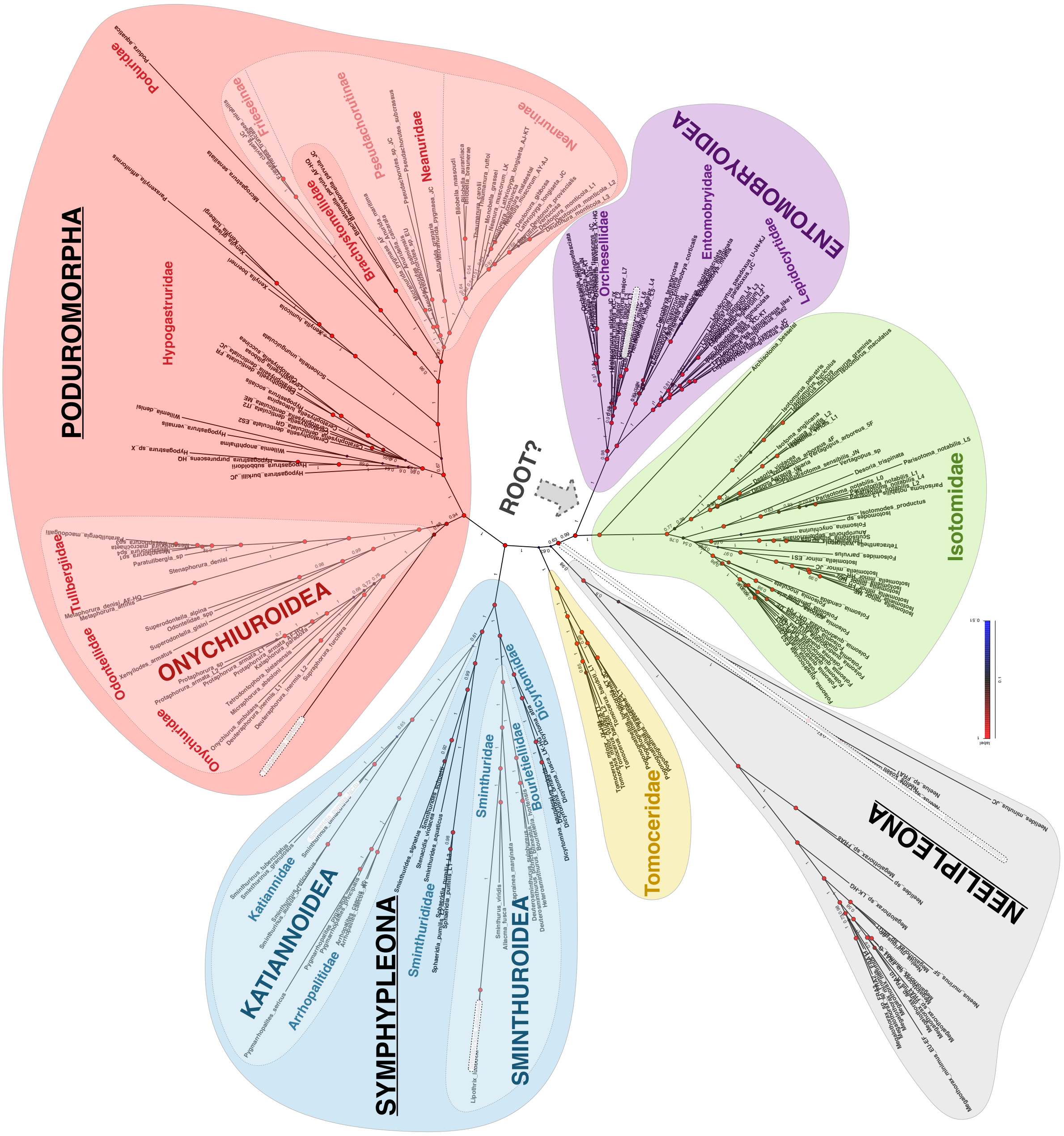

Figure 2.3
Unrooted Bayesian tree of 242 Collembola taxa (including 167 species, 82 genera and 18 families) based on a concatenated alignment of 18S and 28S rRNA, Histone H3 and COI genes (3,083 bp). Numbers next to branches and colors on nodes represent posterior probabilities of 20,002 sampled trees.

Phylogenetic signal of traits

Collembola body length exhibited phylogenetic signal following Brownian motion model (K = 0.786, $P_{(H0 = 0)}$ = 0.001; 2.5%–97.5% simulated K: 0.587–1.770, $P_{(H0 = 1)}$ = 0.520; **Figure 2.4**, **S2.2**). Body shape, vertical stratification, pigmentation, number of ommatidia and reproductive mode all showed phylogenetic signal as indicated by Pagel's lambda which approached 1.0 (**Table 2.1**). Species abundance, however, exhibited lower phylogenetic signal than that predicted by Brownian motion model (K = 0.412, $P_{(H0 = 0)}$ = 0.029; 2.5%–97.5% simulated K: 0.581–1.762, $P_{(H0 = 1)}$ < 0.001), indicating abundant species comprised species from different clades. Species occurrence in different types of habitats, however, exhibited no phylogenetic signal, as indicated by the D-values of 0.889, 0.790 and 0.661 for occurrence in arable fields, grasslands and forests, respectively. All D-values deviated from the Brownian threshold model but followed the phylogenetic randomness model.

Table 2.1

Phylogenetic signal reported as Pagel's lambda with maximum log likelihood test in Collembola categorical traits. Maximum log likelihood of a trait fit to the given phylogeny (logL) was tested against that fit to a lambda transformed phylogeny (logL0, lambda = 0). A significant *P*-value in Pagel's lambda test indicates phylogenetic signal in that trait as predicted by the Brownian motion model.

Trait	States	lambda	logL	logL0	*P*
Body shape	3	1.000	-24.48	-73.36	**0.000**
Vertical stratification	3	1.000	-61.56	-81.87	**0.000**
Pigmentation	3	0.995	-70.63	-81.49	**0.000**
Number of ommatidia	3	1.000	-53.40	-73.36	**0.000**
Reproductive mode	2	0.899	-42.21	-42.45	**0.044**

Trait evolution

Ancestral states of the examined characters of Collembola were intermediate size, slender body shape, hemiedaphic, bright pigmentation, many ommatidia and sexual reproduction (**Figure 2.5**, **S2.2**, **S2.3**). During Collembola diversification trait states changed several times (**Table S2.6**). Variation in Collembola body length as a continuous variable was not different from that predicted by the BM model as compared to the OU model, suggesting that difference in body length of Collembola species resulted from a gradual and continuous drift when species diverged. However, evolution of Collembola body shape was supported by the EB model, suggesting that evolution of body shape was faster early in Collembola diversification and decelerated afterwards. The evolution models of the other categorical traits were not different from the rate constancy model (**Table 2.2**). Evolutionary transitions of Collembola traits depended on each other, especially for vertical stratification, number of ommatidia, pigmentation and reproductive mode, with euedaphic stratification correlated with pale pigmentation, none ommatidia and parthenogenesis (**Figure 2.6**, **S2.4**, **Table S2.7**). Interestingly, body shape did not correlate with any other examined traits.

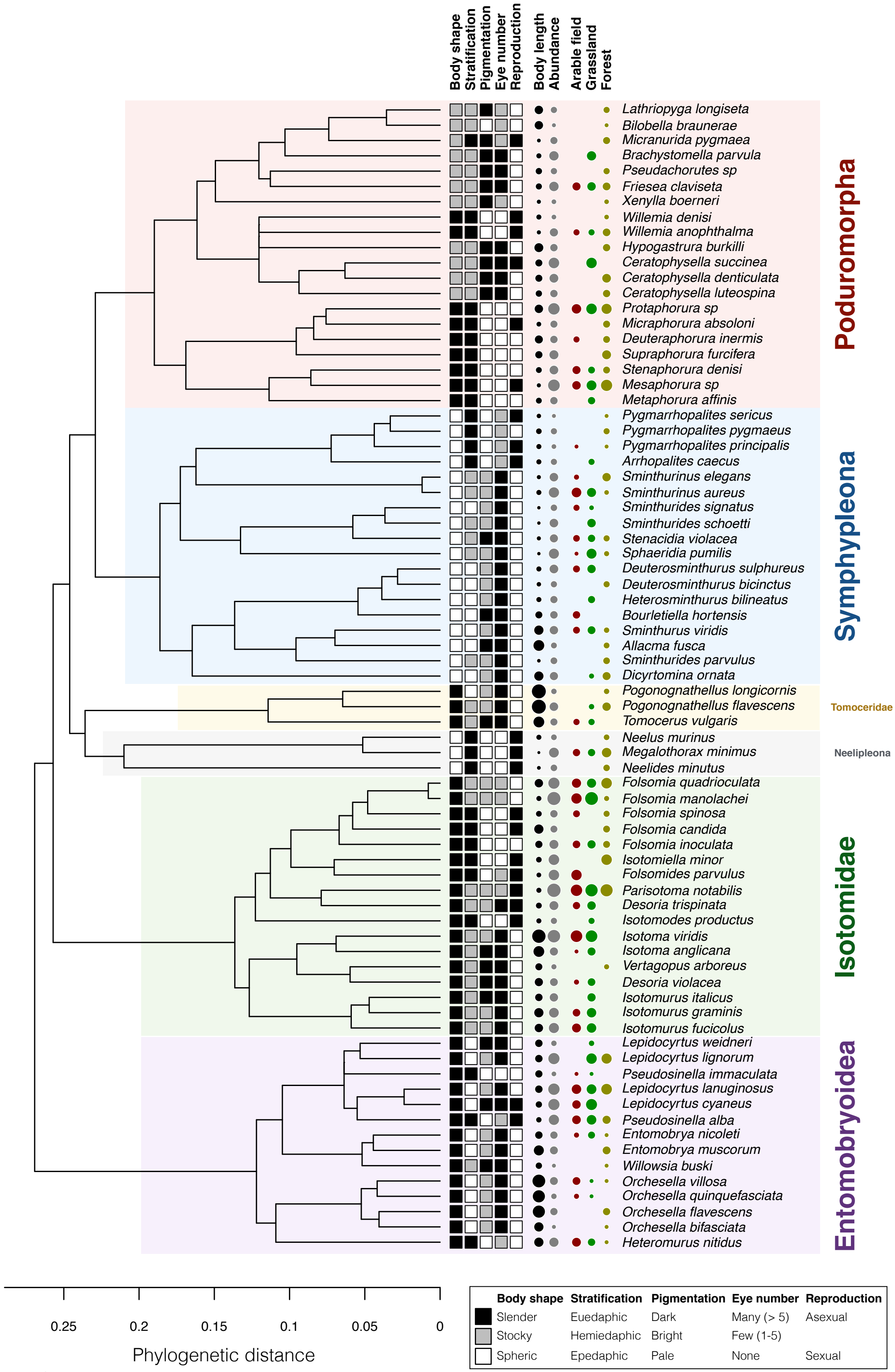

Figure 2.4

Ultrametric phylogenetic tree of the 75 Collembola species based on the Bayesian tree using penalized likelihood that assumes a strict clock model of substitution rate variation among branches. Trait states (squares) and trait values (circles) are plotted next to the tree. Size of a circle represents relative quantity of the trait.

Table2.2

Results of Akaike weights of trait evolution models of Collembola. *P*-value is given based on log-likelihood comparison between the corresponding model and the rate constancy model (none).

Model	**Body shape**				**Stratification**				**Pigmentation**				**Ommatidia**				**Reproduction**			
	AIC	**ΔAIC**	**AIC_w**	***P***	**AIC**	**ΔAIC**	**AIC_w**	***P***	**AIC**	**ΔAIC**	**AIC_w**	***P***	**AIC**	**ΔAIC**	**AIC_w**	***P***	**AIC**	**ΔAIC**	**AIC_w**	***P***
None	51.0	3.84	0.098		125.2	0.20	**0.261**		143.3	0.10	**0.323**		108.9	0.00	**0.396**		86.9	0.00	**0.330**	
White	147.7	100.53	0.000		166.4	41.43	0.000		165.4	22.18	0.000		150.4	41.52	0.000		91.0	4.05	0.044	
Lambda	53.1	5.95	0.034		127.3	2.31	0.091		145.4	2.20	0.113		111.0	2.11	0.138		88.6	1.64	0.145	
Delta	50.1	2.89	0.157	0.080	125.1	0.15	**0.268**	0.141	145.4	2.21	0.112	0.984	110.9	2.04	0.143	0.787	88.2	1.30	0.172	0.368
Early-burst	47.2	0.00	**0.668**	**0.015**	125.0	0.00	**0.289**	0.128	145.4	2.20	0.113	0.961	110.4	1.53	0.185	0.445	88.3	1.32	0.171	0.373
Kappa	52.7	5.52	0.042	0.510	127.3	2.31	0.091	1.000	143.2	0.00	**0.339**	0.137	111.0	2.11	0.138	1.000	88.7	1.75	0.137	0.550

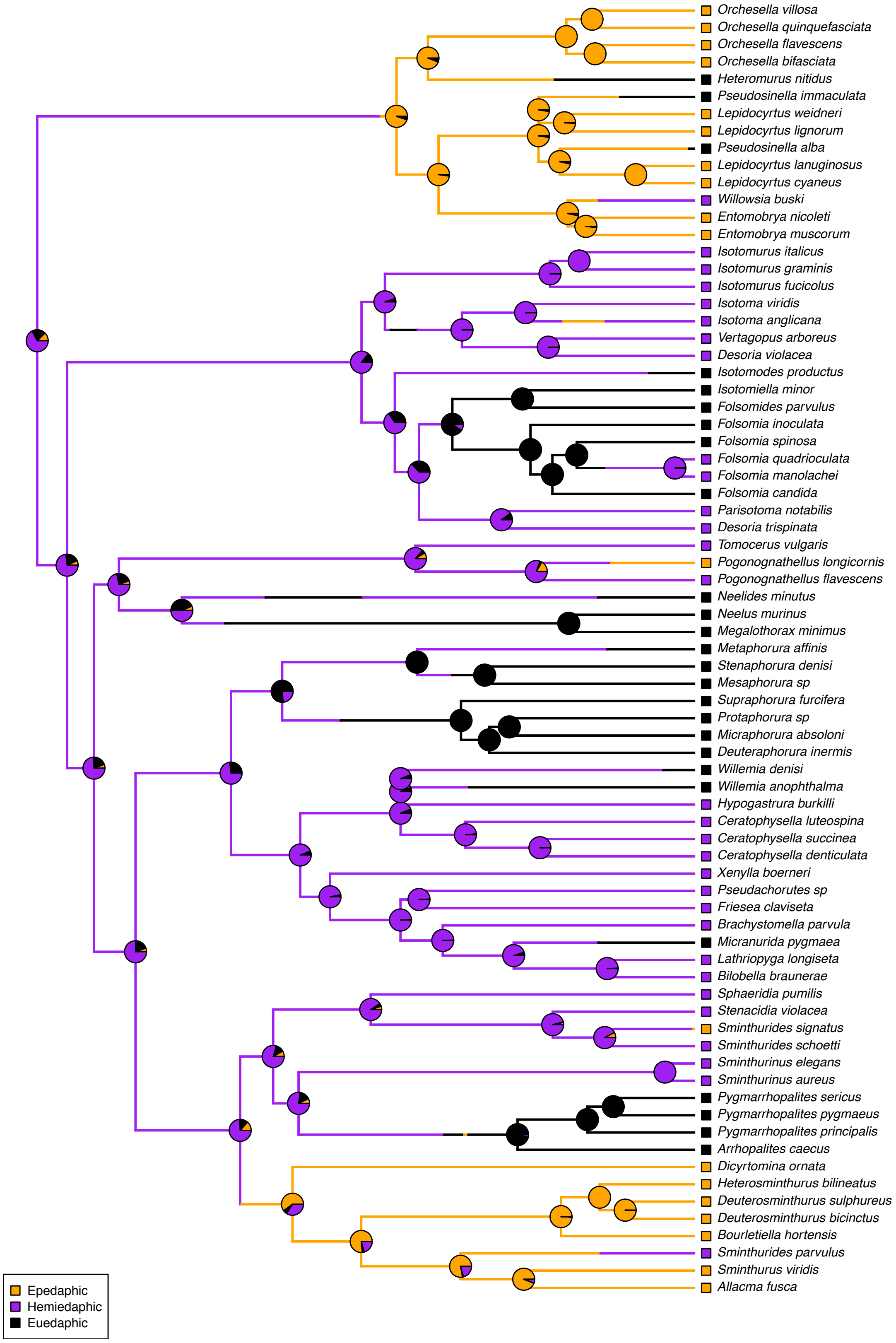

Figure 2.5
Ancestral state reconstruction of vertical stratification of Collembola. Colored squares at the tips of each branch represent the vertical stratification of each extant species. Pie charts on each node indicate the proportion of each character state summed across the posterior distribution of simulations. One of the 1,000 stochastic character maps is plotted.

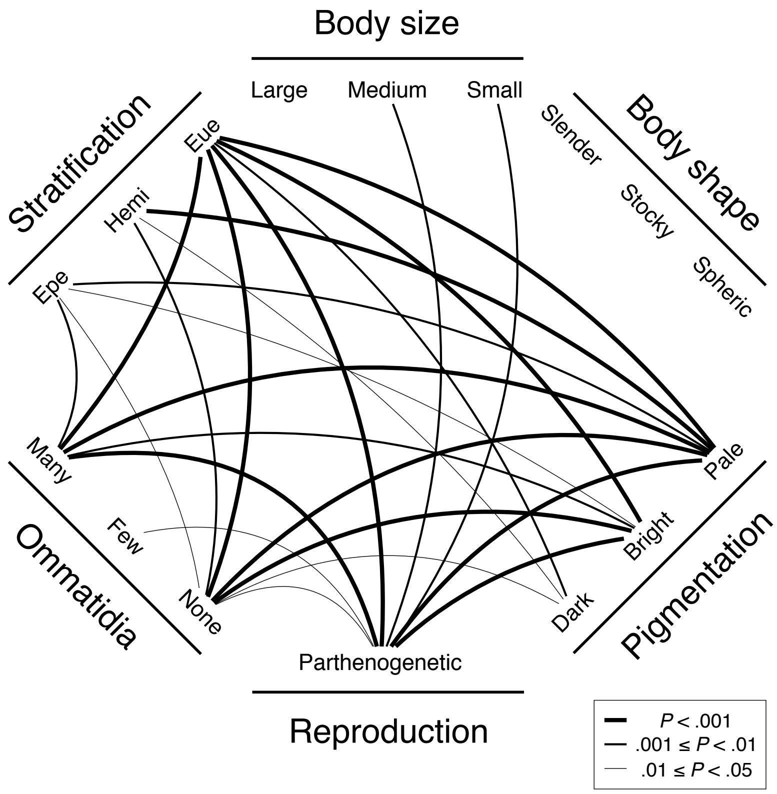

Figure 2.6
Results of Pagel's correlation test between the transition rate of two traits of Collembola; significant dependence between two traits (i.e., "xy" model) are indicated by lines with thickness reflecting the *P*-values. See Table S2.4 for the results of trait dependence by one direction.

Discussion

In most trait-based studies on soil animals, associations of traits of species with the habitats are usually descriptive but still have been interpreted as the result of adaptation to environmental factors, without testing it in an evolutionary context (Vandewalle et al. 2010, Ponge and Salmon 2012, Bokhorst et al. 2017, but see Malcicka et al. 2017). Here, we conducted phylogenetic comparative analyses to study the evolution of functional traits in soil Collembola. We first reconstructed phylogenetic trees for species occurring in various habitats, and erected a more comprehensive phylogeny as compared to existing Collembola phylogenies. We then focused on the evolution of functional traits of Collembola in light of their associations with different soil layers. Explicitly testing trait evolution is an important step toward integrating phylogeny into studies on assembly processes of soil communities using community phylogenetic approaches. See **Chapters 1 and 4** for more details on the community phylogenetic methods.

Collembola phylogeny

In the present study, both conserved (18S rRNA and H3) and variable (28S rRNA and COI) genetic markers were used to infer the phylogeny of 102 locally occurring Collembola species. Our tree provides the most comprehensive Collembola phylogeny to date that included 51 genera and 18 families and considered a variety of closely related outgroups (Misof et al. 2014). This tree indicates that Entomobryoidea and Isotomidae diverged earlier from the other Collembola lineages, while Poduromorpha and Symphypleona are sister groups, contrary to previous studies, which recovered Symphypleona as the sister group to the other Collembola (D'Haese 2002, Xiong et al. 2008, Schneider et al. 2011, Yu et al. 2016). Marker selection, unbalanced taxon sampling and choice of outgroup taxa can affect the topology of phylogenetic trees (Heath et al. 2008, Rosenfeld et al. 2012) and thus resulted in the inconsistency between our tree and those inferred in the previous studies. However, fossil records of Collembola support our findings of the early diverging positions of Entomobryoidea and Isotomidae. The first fossil record of Collembola, *Rhyniella praecursor* (396–407 million years ago; Hirst and Maulik 1926, Whalley and Jarzembowski 1981, Greenslade 1988), has been assigned to Isotomidae (Greenslade and Walley 1986), suggesting that Isotomidae are likely among the oldest Collembola. Another Paleozoic fossil, *Permobrya mirabilis* from the Upper Permian, shares characters with extant Entomobryoidea (Riek 1976), suggesting that this superfamily is also phylogenetically old. Fitting the more derived position in our tree, fossil Collembola found in mid and upper Cretaceous amber have been assigned to Collembola families such as Sminthuridae, Neanuridae and Tomoceridae (Christiansen and Pike 2002a, b, Christiansen and Nascimbene 2006). However, in particular early radiation of Collembola lineages needs further investigation using novel approaches, e.g. phylogenomics. Nevertheless, with the extensive taxon sampling and more genes included in this study and the support from the existing fossil record, our tree provides a solid basis for investigating the evolution of traits in Collembola.

Functional traits and their evolution

Collembola body length, body shape and vertical stratification all showed phylogenetic signal, supporting our first hypothesis. Intuitively plausible, body shape and body length of Collembola may relate to the three-dimensional structure of soil pores (Larsen et al. 2004), as pore size declines typically with soil depth and differs between mineral soil and organic layers. If the structure of soil pores is related to evolutionary adaptation of Collembola species to the vertical distribution in soil profile, we would expect to find evolutionary correlations between body shape, body size and vertical stratification. However, our study does not support the existence of such relationships, suggesting that other mechanisms resulted in the phylogenetic signal of these traits. The evolution of body shape followed the early burst model, suggesting a rapid morphological change in the early evolution followed by relative stasis, presumably under stable evolutionary or ecological constraints. The finding contrasts the conclusion of Harmon et al. (2010) that early evolutionary changes in body shape are rare. In contrast to body shape, body size followed the rate constancy model, suggesting that Collembola gradually changed their body size via drift, an evolutionary process in which trait variation accumulated during evolutionary time. Collembola body size has been shown to relate to desiccation resistance of species (Kærsgaard et al. 2004). Phylogenetic signal detected in the body size therefore supports the idea that phylogenetically related species share physiological attributes (Chen et al. 2017). The evolution of vertical stratification of Collembola, though not related to either body shape or body size, was likely related to pigmentation, number of ommatidia and reproductive mode, supporting the second hypothesis.

Evolutionary correlation between pigmentation and vertical stratification suggests that these two traits were functionally related during species diversification. Dark pigmentation is likely a product of adaptation of Collembola to open habitats or surface-living (Salmon et al. 2014), since it reduces damages caused by solar ultraviolet radiation (UV). Also, species with spots and stripes of different colors probably function to mislead predators. Therefore, surface living (epedaphic) species possess shiny or dark pigments, such as many species of Entomobryoidea and Symphypleona. Species dwelling in soil, on the contrary, are usually pale, such as Onychiuroidea. This is further supported by reconstruction of pale coloration as ancestral state in Onychiuroidea. Similar to the pigmentation, significant correlation between the number of ommatidia and vertical stratification indicates that these two traits were functionally related in Collembola evolution. This is also reflected in the eight evolutionary shifts from many ommatidia to no ommatidia, corresponding closely to the nine evolutionary shifts in vertical stratification from hemiedaphic to euedaphic. Surface living Collembola may use their ommatidia as sensory organ to orient themselves according to sunlight (Hågvar 1995) and to detect predators (Salmon et al. 2014). In contrast, euedaphic species rely more on other sensory organs, such as sensory structures on antennae or the postantennal organ (Hopkin 1997, Salmon et al. 2014). Furthermore, contrary to the findings of Malcicka et al. (2017), our study shows that reproductive mode correlated with pigmentation, number of ommatidia and vertical stratification. Indeed, most parthenogenetic Collembola species are euedaphic (Chernova et al. 2010), pale and lack ommatidia. Parthenogenesis in deep soil may reflect the increased difficulty in finding sexual partners due to restricted movement and location of olfactory cues of spermatophores of respectively male partners. In addition, less pronounced resource limitation and dominance of density independent factors in soil may also contribute to the dominance of parthenogenetic Collembola deeper in soil (Chahartaghi et al. 2009).

We did not find phylogenetic signal of species occurrence in different types of habitats, i.e., arable fields, grasslands and forests. Compared to the broadly defined habitats investigated in this study, soil microhabitats, especially the vertical heterogeneity in the soil profile, may have played a more important role for evolutionary adaptations in Collembola than habitat types such as forests and grasslands. Interestingly, in each phylogenetic clade, there was one or few species reaching high abundance, resulting in the significantly lower phylogenetic signal measured in total abundance than that predicted by the Brownian motion model. Since abundance of soil microarthropods positively correlates with the amount of food resources available (Domes et al. 2007, Chahartaghi et al. 2009), this might indicate niche partitioning in food resources between closely related species. Considering the many microhabitats in soil and associated niches, niche partitioning likely contributed to species diversification and this may offer an explanation for the enigma of soil animal biodiversity (Anderson 1975).

Conclusions and Outlooks

Overall, phylogenetic signal in body length, body shape, vertical stratification, pigmentation, number of ommatidia and reproductive mode support our first hypothesis that phylogenetically related species of Collembola share similar functional traits. Further, the results suggest that the body shape of Collembola evolved fast during early diversification of lineages but slowed down thereafter. Transitions of pigmentation, number of ommatidia and reproductive mode depended on vertical stratification of species during Collembola diversification, supporting our second hypothesis. The ancestral state of Collembola traits are likely slender body shape, hemiedaphic lifestyle, sexual reproduction, possession of many ommatidia and bright color, but these traits changed several times during Collembola evolution.

Phylogenetic signal in functional traits of soil species provides an evolutionary perspective to soil biodiversity and community assembly processes. The traits of species are products of ecological processes in the past resulting in evolutionary processes. Although trait patterns in communities may reflect environmental associations of the species resulting from contemporary ecological processes, e.g. environmental filtering (Widenfalk et al. 2015), evolution of traits also needs to be considered. If traits evolved in a constrained manner or following Brownian motion model, trait similarity between species in communities cannot be simply inferred as a result of contemporary ecological processes, since also evolutionary mechanisms may have resulted in species possessing similar traits. Our study on trait evolution of soil biota, therefore, paves the way for integrating evolutionary approaches and perspectives into soil ecological studies. Future studies on functional traits and assembly processes of soil communities will benefit from multiple approaches including phylogenetic comparative and community phylogenetic methods.

Author Contributions

TWC, IS and SS conceived and designed the study; TWC, JFC, and MPB performed the study and collected the data; TWC analyzed the data; TWC, MPB, IS and SS wrote the manuscript.

Data Accessibility

All DNA sequences generated from this study were deposited in GenBank (Accession Number KY230697–KY231137).

References

Anderson, J.M. (1975) The enigma of soil animal species diversity. Progress in Soil Zoology: Proceedings of the 5th International Colloquium on Soil Zoology (ed J. Vaněk), pp. 51–58. Springer Netherlands, Dordrecht.

Anslan, S. & Tedersoo, L. (2015) Performance of cytochrome c oxidase subunit I (COI), ribosomal DNA Large Subunit (LSU) and Internal Transcribed Spacer 2 (ITS2) in DNA barcoding of Collembola. *European Journal of Soil Biology*, **69**, 1–7.

Bellinger, P.F., Christiansen, K.A. & Janssens, F. Checklist of the Collembola of the world (www.collembola.org).

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*, **57**, 289–300.

Blomberg, S.P. & Garland Jr., T. (2002) Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *Journal of Evolutionary Biology*, **15**, 899–910.

Blomberg, S.P., Garland Jr., T. & Ives, A.R. (2003) Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution*, **57**, 717–745.

Bokhorst, S., Berg, M.P. & Wardle, D.A. (2017) Micro-arthropod community responses to ecosystem retrogression in boreal forest. *Soil Biology and Biochemistry*, **110**, 79–86.

Bollback, J.P. (2006) SIMMAP: stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics*, **7**, 88.

Burkhardt, U. & Filser, J. (2005) Molecular evidence for a fourth species within the *Isotoma viridis* group (Insecta, Collembola). *Zoologica Scripta*, **34**, 177–185.

Carapelli, A., Convey, P., Nardi, F. & Frati, F. (2014) The mitochondrial genome of the antarctic springtail *Folsomotoma octooculata* (Hexapoda; Collembola), and an update on the phylogeny of collembolan lineages based on mitogenomic data. *Entomologia*, **2**.

Chahartaghi, M., Maraun, M., Scheu, S. & Domes, K. (2009) Resource depletion and colonization: A comparison between parthenogenetic and sexual Collembola species. *Pedobiologia*, **52**, 181–189.

Chen, T.-W., Sandmann, P., Schaefer, I. & Scheu, S. (2017) Neutral lipid fatty acid composition as trait and constraint in Collembola evolution. *Ecology and Evolution*, **7**, 9624-9638.

Chernova, N.M., Potapov, M.B., Savenkova, Y.Y. & Bokova, A.I. (2010) Ecological significance of parthenogenesis in Collembola. *Entomological Review*, **90**, 23–38.

Christiansen, K. & Nascimbene, P. (2006) Collembola (Arthropoda, Hexapoda) from the mid Cretaceous of Myanmar (Burma). *Cretaceous Research*, **27**, 318–363.

Christiansen, K. & Pike, E. (2002a) A preliminary report on the Cretaceous Collembola. *Pedobiologia*, **46**, 267–273.

Christiansen, K. & Pike, E. (2002b) Cretaceous Collembola (Arthropoda, Hexapoda) from the Upper Cretaceous of Canada. *Cretaceous Research*, **23**, 165–188.

Cicconardi, F., Fanciulli, P.P. & Emerson, B.C. (2013) Collembola, the biological species concept and the underestimation of global species richness. *Molecular Ecology*, **22**, 5382–96.

Cicconardi, F., Nardi, F., Emerson, B.C., Frati, F. & Fanciulli, P.P. (2010) Deep phylogeographic divisions and long-term persistence of forest invertebrates (Hexapoda: Collembola) in the North-Western Mediterranean basin. *Molecular Ecology*, **19**, 386–400.

Cooper, N., Jetz, W. & Freckleton, R.P. (2010) Phylogenetic comparative approaches for studying niche conservatism. *Journal of Evolutionary Biology*, **23**, 2529–2539.

D'Haese, C.A. (2002) Were the first springtails semi-aquatic? A phylogenetic approach by means of 28S rDNA and optimization alignment. *Proceedings of the Royal Society of London B*, **269**, 1143–1151.

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

Deharveng, L. (2004) Recent advances in Collembola systematics. *Pedobiologia*, **48**, 415–433.

Dias, A.T.C., Krab, E.J., Mariën, J., Zimmer, M., Cornelissen, J.H.C., Ellers, J., Wardle, D.A. & Berg, M.P. (2013) Traits underpinning desiccation resistance explain distribution patterns of terrestrial isopods. *Oecologia*, **172**, 667–677.

Domes, K., Scheu, S. & Maraun, M. (2007) Resources and sex: Soil re-colonization by sexual and parthenogenetic oribatid mites. *Pedobiologia*, **51**, 1–11.

Fjellberg, A. (1998) *The Collembola of Fennoscandia and Denmark Part I: Poduromorpha* (eds NP Kristensen and V Michelsen). Koninklijke Brill, Leiden. The Netherlands.

Fjellberg, A. (2007) The Collembola of Fennoscandia and Denmark Part II: Entomobryomorpha and Symphypleona (eds NP Kristensen and V Michelsen). Koninklijke Brill, Leiden. The Netherlands.

Frati, F. & Carapelli, A. (1999) An assessment of the value of nuclear and mitochondrial genes in elucidating the origin and evolution of *Isotoma klovstadi* Carpenter (Insecta, Collembola). *Antarctic Science*, **11**, 160–174.

Frati, F., Dell'Ampio, E., Casasanta, S., Carapelli, A. & Paolo Fanciulli, P. (2000) Large amounts of genetic divergence among Italian species of the genus *Orchesella* (Insecta, collembola) and the relationships of two new species. *Molecular Phylogenetics and Evolution*, **17**, 456–61.

Freckleton, R.P., Harvey, P.H. & Pagel, M. (2002) Phylogenetic analysis and comparative data: a test and review of evidence. *The American Naturalist*, **160**, 712–726.

Fritz, S.A. & Purvis, A. (2010) Selectivity in mammalian extinction risk and threat types: A new measure of phylogenetic signal strength in binary traits. *Conservation Biology*, **24**, 1042–1051.

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

Gisin, H. (1960) *Collembolenfauna Europas*. Muséum d'Histoire Naturelle, Geneva.

Greenslade, P.J.M. (1988) Reply to R.A. Crowson's "Comments on Insecta of the Rhynie Chert." *Entomologia Generalis*, **13**, 115–117.

Greenslade, P., Stevens, M.I., Torricelli, G. & D'Haese, C. a. (2011) An ancient Antarctic endemic genus restored: Morphological and molecular support for *Gomphiocephalus hodgsoni* (Collembola: Hypogastruridae). *Systematic Entomology*, **36**, 223–240.

Greenslade, P. & Whalley, P. (1986) The systematic position of *Rhyniella praecursor* Hirst and Maulik (Collembola). The earliest known hexapod. *Second International Seminar on Apterygota*, pp. 319–323. University of Siena, Siena, Italy.

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

Hågvar, S. (1995) Long distance, directional migration on snow in a forest collembolan, *Hypogastrura socialis* (Uzel). *Acta Zoologica Fennica*, **196**, 200–205.

Harmon, L.J., Losos, J.B., Jonathan Davies, T., Gillespie, R.G., Gittleman, J.L., Bryan Jennings, W., Kozak, K.H., McPeek, M.A., Moreno-Roark, F., Near, T.J., Purvis, A., Ricklefs, R.E., Schluter, D., Schulte, J.A., Seehausen, O., Sidlauskas, B.L., Torres-Carvajal, O., Weir, J.T. & Mooers, A.T. (2010) Early bursts of body size and shape evolution are rare in comparative data. *Evolution*, **64**, 2385–2396.

Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E. & Challenger, W. (2008) GEIGER: Investigating evolutionary radiations. *Bioinformatics*, **24**, 129–131.

Heath, T.A., Hedtke, S.M. & Hillis, D.M. (2008) Taxon sampling and the accuracy of phylogenetic analyses. *Journal of Systematics and Evolution*, **46**, 239–257.

Hirst, S. & Maulik, S. (1926) On some arthropod remains from the Rhynie Chert (old red sandstone). *Geological Magazine*, **63**, 69–71.

Hopkin, S.P. (1997) Biology of the Springtails (Insecta: Collembola). Oxford University Press, Oxford.

Hopkin, S.P. (2007) A Key to the Collembola (Springtails) of Britain and Ireland (ed Field Studies Council (FSC)). Shrewsbury.

Huelsenbeck, J.P., Nielsen, R., Bollback, J.P. & Schultz, T. (2003) Stochastic mapping of morphological characters. *Systematic Biology*, **52**, 131–158.

Kærsgaard, C.W., Holmstrup, M., Malte, H. & Bayley, M. (2004) The importance of cuticular permeability, osmolyte production and body size for the desiccation resistance of nine species of Collembola. *Journal of Insect Physiology*, **50**, 5–15.

Kempson, D., Lloyd, M. & Ghelardi, R. (1963) A new extractor for woodland litter. *Pedobiologia*, **3**, 1–21.

Larsen, T., Schjønning, P. & Axelsen, J. (2004) The impact of soil compaction on euedaphic Collembola. *Applied Soil Ecology*, **26**, 273–281.

Losos, J.B. (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters*, **11**, 995–1003.

Luan, Y., Mallatt, J.M., Xie, R.D., Yang, Y.M. & Yin, W. (2005) The phylogenetic positions of three basal-hexapod groups (Protura, Diplura, and Collembola) based on ribosomal RNA gene sequences. *Molecular Biology and Evolution*, **22**, 1579–1592.

Macfadyen, A. (1961) Improved funnel-type extractors for soil arthropods. *The Journal of Animal Ecology*, **30**, 171–184.

Makkonen, M., Berg, M.P., van Hal, J.R., Callaghan, T. V., Press, M.C. & Aerts, R. (2011) Traits explain the responses of a sub-arctic Collembola community to climate manipulation. *Soil Biology and Biochemistry*, **43**, 377–384.

Malcicka, M., Berg, M.P. & Ellers, J. (2017) Ecomorphological adaptations in Collembola in relation to feeding strategies and microhabitat. *European Journal of Soil Biology*, **78**, 82–91.

McGaughran, A., Stevens, M.I. & Holland, B.R. (2010) Biogeography of circum-Antarctic springtails. *Molecular Phylogenetics and Evolution*, **57**, 48–58.

Minter, N.J., Buatois, L.A., Mángano, M.G., Davies, N.S., Gibling, M.R., MacNaughton, R.B. & Labandeira, C.C. (2017) Early bursts of diversification defined the faunal colonization of land. *Nature Ecology & Evolution*, **1**, 175.

Misof, B., Liu, S., Meusemann, K., Peters, R.S., Donath, A., Mayer, C., Frandsen, P.B., Ware, J., Flouri, T., Beutel, R.G., Niehuis, O., Petersen, M., Izquierdo-Carrasco, F., Wappler, T., Rust, J., Aberer, A.J., Aspöck, U., Aspöck, H., Bartel, D., Blanke, A., Berger, S., Böhm, A., Buckley, T.R., Calcott, B., Chen, J., Friedrich, F., Fukui, M., Fujita, M., Greve, C., Grobe, P., Gu, S., Huang, Y., Jermiin, L.S., Kawahara, A.Y., Krogmann, L., Kubiak, M., Lanfear, R., Letsch, H., Li, Y., Li, Z., Li, J., Lu, H., Machida, R., Mashimo, Y., Kapli, P., McKenna, D.D., Meng, G., Nakagaki, Y., Navarrete-Heredia, J.L., Ott, M., Ou, Y., Pass, G., Podsiadlowski, L., Pohl, H., von Reumont, B.M., Schütte, K., Sekiya, K., Shimizu, S., Slipinski, A., Stamatakis, A., Song, W., Su, X., Szucsich, N.U., Tan, M., Tan, X., Tang, M., Tang, J., Timelthaler, G., Tomizuka, S., Trautwein, M., Tong, X., Uchifune, T., Walzl, M.G., Wiegmann, B.M., Wilbrandt, J., Wipfler, B., Wong, T.K.F., Wu, Q., Wu, G., Xie, Y., Yang, S., Yang, Q., Yeates, D.K., Yoshizawa, K., Zhang, Q., Zhang, R., Zhang, W., Zhang, Y., Zhao, J., Zhou, C., Zhou, L., Ziesmann, T., Zou, S., Li, Y., Xu, X., Zhang, Y., Yang, H., Wang, J., Wang, J., Kjer, K.M. & Zhou, X. (2014) Phylogenomics resolves the timing and pattern of insect evolution. *Science*, **346**, 763–767.

Orme, D., Freckleton, R.P., Thomas, G., Petzoldt, T., Fritz, S.A., Isaac, N. & W. Pearse. (2013) Caper: comparative analyses of phylogenetics and evolution in R.

Pachl, P., Domes, K., Schulz, G., Norton, R., Scheu, S., Schaefer, I. & Maraun, M. (2012) Convergent evolution of defense mechanisms in oribatid mites (Acari, Oribatida) shows no "ghosts of predation past." *Molecular Phylogenetics and Evolution*, **65**, 412–420.

Pagel, M. (1994) Detecting Correlated Evolution on Phylogenies: a General Method for the Comparative Analysis of Discrete Characters. *Proceedings of the Royal Society of London*, **255**, 37–45.

Pagel, M. (1999) Inferring the historical patterns of biological evolution. *Nature*, **401**, 877–884.

Paradis, E. (2013) Molecular dating of phylogenies by likelihood methods: A comparison of models and a new information criterion. *Molecular Phylogenetics and Evolution*, **67**, 436–444.

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

Park, K.-H. (2009) Molecular biological study on speciation and phylogeny of the order Entomobryomorpha (Collembola: Hexapoda). *Entomological Research*, **39**, 334–340.

Pey, B., Nahmani, J., Auclerc, A., Capowiez, Y., Cluzeau, D., Cortet, J., Decaëns, T., Deharveng, L., Dubs, F., Joimel, S., Briard, C., Grumiaux, F., Laporte, M.-A.A., Pasquet, A., Pelosi, C., Pernin, C., Ponge, J.-F., Salmon, S., Santorufo, L. & Hedde, M. (2014) Current use of and future needs for soil invertebrate functional traits in community ecology. *Basic and Applied Ecology*, **15**, 194–206.

Ponge, J.-F. & Salmon, S. (2013) Spatial and taxonomic correlates of species and species trait assemblages in soil invertebrate communities. *Pedobiologia*, **56**, 129–136.

Potapov, A.A., Semenina, E.E., Korotkevich, A.Y., Kuznetsova, N.A. & Tiunov, A. V. (2016) Connecting taxonomy and ecology: Trophic niches of collembolans as related to taxonomic identity and life forms. *Soil Biology and Biochemistry*, **101**, 20–31.

R Core Team. (2015) R: A language and environment for statistical computing.

von Reumont, B.M., Jenner, R.A., Wills, M.A., Dell'Ampio, E., Pass, G., Ebersberger, I., Meyer, B., Koenemann, S., Iliffe, T.M., Stamatakis, A., Niehuis, O., Meusemann, K. & Misof, B. (2012) Pancrustacean phylogeny in the light of new phylogenomic data: Support for Remipedia as the possible sister group of Hexapoda. *Molecular Biology and Evolution*, **29**, 1031–1045.

Revell, L.J. (2012) phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, **3**, 217–223.

Revell, L., Harmon, L. & Collar, D. (2008) Phylogenetic signal, evolutionary process, and rate. *Systematic Biology*, **57**, 591–601.

Revell, L.J., Johnson, M.A., Schulte, J.A., Kolbe, J.J. & Losos, J.B. (2007) A phylogenetic test for adaptive convergence in rock-dwelling lizards. *Evolution*, **61**, 2898–2912.

Riek, E.F. (1976) An entomobryid collembolan (Hexapoda: Collembola) from the Lower Permian of southern Africa. *Palaeontologia Africana*, **19**, 141–143.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A. & Huelsenbeck, J.P. (2012) Mrbayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.

Rosenfeld, J.A., Payne, A. & DeSalle, R. (2012) Random roots and lineage sorting. *Molecular Phylogenetics and Evolution*, **64**, 12–20.

Rota-stabelli, O., Daley, A.C. & Pisani, D. (2013) Molecular timetrees reveal a Cambrian colonization of land and a new scenario for Ecdysozoan evolution. *Current Biology*, **23**, 392–398.

Salmon, S. & Ponge, J.-F. (2012) Species traits and habitats in springtail communities: A regional scale study. *Pedobiologia*, **55**, 295–301.

Salmon, S., Ponge, J.-F., Gachet, S., Deharveng, L., Lefebvre, N. & Delabrosse, F. (2014) Linking species, traits and habitat characteristics of Collembola at European scale. *Soil Biology and Biochemistry*, **75**, 73–85.

von Saltzwedel, H., Scheu, S. & Schaefer, I. (2016) Founder events and pre-glacial divergences shape the genetic structure of European Collembola species. *BMC Evolutionary Biology*, **16**, 148.

Schneider, C., Cruaud, C. & D'Haese, C.A. (2011) Unexpected diversity in Neelipleona revealed by molecular phylogeny approach (Hexapoda, Collembola). *Soil Organisms*, **83**, 383–398.

Soto-Adames, F.N. (2002) Molecular phylogeny of the Puerto Rican *Lepidocyrtus* and *Pseudosinella* (Hexapoda: Collembola), a validation of Yoshii's "color pattern species". *Molecular Phylogenetics and Evolution*, **25**, 27–42.

Stamatakis, A. (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.

van Straalen, N.M., Timmermans, M.J.T.N., Roelofs, D. & Berg, M.P. (2008) Apterygota in the spotlights of ecology, evolution and genomics. *European Journal of Soil Biology*, **44**, 452–457.

Vaidya, G., Lohman, D.J. & Meier, R. (2011) SequenceMatrix: Concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, **27**, 171–180.

Vandewalle, M., de Bello, F., Berg, M.P., Bolger, T., Dolédec, S., Dubs, F., Feld, C.K., Harrington, R., Harrison, P.A., Lavorel, S., Silva, P.M., Moretti, M., Niemelä, J., Santos, P., Sattler, T., Sousa, J.P., Sykes, M.T., Vanbergen, A.J. & Woodcock, B.A. (2010) Functional traits as indicators of biodiversity response to land use changes across ecosystems and organisms. *Biodiversity and Conservation*, **19**, 2921–2947.

Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. (2007) Let the concept of trait be functional! *Oikos*, **116**, 882–892.

Whalley, P. & Jarzembowski, E.A. (1981) A new assessment of *Rhyniella*, the earliest known insect, from the Devonian of Rhynie, Scotland. *Nature*, **291**, 317.

Widenfalk, L.A., Bengtsson, J., Berggren, Å., Zwiggelaar, K., Spijkman, E., Huyer-Brugman, F. & Berg, M.P. (2015) Spatially structured environmental filtering of collembolan traits in late successional salt marsh vegetation. *Oecologia*, **179**, 537–549.

Wolfe, J.M., Daley, A.C., Legg, D.A. & Edgecombe, G.D. (2016) Fossil calibrations for the arthropod Tree of Life. *Earth-Science Reviews*, **160**, 43–110.

Wright, E.S. (2015) DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics*, **16**, 322.

Xiong, Y., Gao, Y., Yin, W. & Luan, Y. (2008) Molecular phylogeny of Collembola inferred from ribosomal RNA genes. *Molecular Phylogenetics and Evolution*, **49**, 728–735.

Yu, D., Zhang, F., Stevens, M.I., Yan, Q., Liu, M. & Hu, F. (2016) New insight into the systematics of Tomoceridae (Hexapoda, Collembola) by integrating molecular and morphological evidence. *Zoologica Scripta*, **45**, 286–299.

Zhang, F., Chen, Z., Dong, R.-R., Deharveng, L., Stevens, M.I., Huang, Y.-H. & Zhu, C.-D. (2014) Molecular phylogeny reveals independent origins of body scales in Entomobryidae (Hexapoda: Collembola). *Molecular Phylogenetics and Evolution*, **70**, 231–9.

Supplementary Materials

Table S2.1

Sampling locations of this study.

Site	Habitat	Latitude (°N)	Longitude (°E)
1 Herberhausen	Arable	51.53408	10.00058
	Grassland	51.53294	9.99273
	Forest	51.53072	9.99066
2 Deppoldshausen	Arable	51.57565	9.97312
	Grassland	51.57612	9.97195
	Forest	51.57506	9.97443
4 Ossenfeld	Arable	51.54789	9.79804
	Grassland	51.54731	9.79733
	Forest	51.54898	9.80041
5 Waake	Arable	51.56308	10.05845
	Grassland	51.56293	10.06232
	Forest	51.55943	10.07080
6 Billingshausen	Arable	51.59006	10.02655
	Grassland	51.59372	10.03101
	Forest	51.59235	10.03254
8 Ellershausen	Arable	51.51264	9.66830
	Grassland	51.50857	9.66414
	Forest	51.51325	9.66628

Table S2.2

Primer pairs used in this study.

Region	Primer	Sequence 5' - 3'	Annealing temperature	Reference
18S rRNA	18SA2_F	ATGGTTGCAAAGCTGAAAC	50°C	Whiting 2002
	18S9_R	GATCCTTCCGCAGGTTCACCTAC		
28S rRNA				
D1-D2	C1'_F	ACCCGCTGAATTTAAGCAT	50°C	D'Haese 2002
	D2coll_R	ACCACGCATGCWTTAGATTG		
D3-D5	28SA_F	GACCCGTCTTGAAGCACG	52°C	Tully et al. 2006
	28Sbout_R	CCCACAGCGCCAGTTCTGCTTACC		
Histone H3	H3F2	ATGGCTCGTACCAAGCAGAC	56°C	Colgan et al. 1998
	H3R2	ATRTCCTTGGGCATGATTGTTAC		
COI	LCO1490_F	GGTCAACAAATCATAAAGATATTGG	45°C (5 cycles) followed by 51°C (35 cycles)	Folmer et al. 1994
	HCO2198_R	TAAACTTCAGGGTGACCAAAAAATCA		

References

Colgan, D.J., McLauchlan, a., Wilson, G.D.F., Livingston, S.P., Edgecombe, G.D., Macaranas, J., Cassis, G. & Gray, M.R. (1998) Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. *Australian Journal of Zoology*, **46**, 419.

D'Haese, C.A. (2002) Were the first springtails semi-aquatic? A phylogenetic approach by means of 28S rDNA and optimization alignment. *Proceedings of the Royal Society of London B*, **269**, 1143–1151.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**, 294–299.

Tully, T., D'Haese, C.A., Richard, M. & Ferrière, R. (2006) Two major evolutionary lineages revealed by molecular phylogeny in the parthenogenetic collembola species Folsomia candida. *Pedobiologia*, **50**, 95–104.

Whiting, M.F. (2002) Mecoptera is paraphyletic: Multiple genes and phylogeny of Mecoptera and Siphonaptera. *Zoologica Scripta*, **31**, 93–104.

Table S2.3

NCBI Accession Numbers of sequences generated in this study to construct the molecular phylogeny of Collembola.

Taxonomic group	Family	Species	18S rRNA	28S rRNA		Histone H3	COI
				D1-D2	D3-D5		
Outgroups							
Crustacea		*Speleonectes*	EU370431		EU370446	KC989979	JF297647
Protura		*Baculentulus*	AY037169		EF192433		HQ882817
Diplura		*Parajapyx*	AY037168		EF192440		JQ796635
Archaeognatha (Insecta)		*Machilis*	AY338689		AY338646	AY338614	JF826083
Palaeoptera (Insecta)		*Callibaetis*	AF370791		AY859557	AY749703	GU711466
Zygentoma (Insecta)		Zygentoma	EU368615		EU376048	AY555568	JN970940
Collembola							
Poduromorpha	Brachystomellidae	*Brachystomella parvula*	KY230724	KY230822	KY230925	KY231017	KY231088
	Hypogastruridae	*Ceratophysella denticulata*	KY230747	KY230847	KY230948	KY231036	KY231107
		Ceratophysella luteospina	KY230762	KY230859	KY230962		KY231118
		Ceratophysella succinea		KY230885	KY230990	KY231065	KY231136
		Hypogastrura burkilli	KY230725	KY230823	KY230926	KY231018	KY231089
		Willemia anophthalma	KY230726	KY230869	KY230972		KY231090
		Willemia denisi	KY230757	KY230855	KY230957		KY231114
		Xenylla boerneri	KY230728	KY230824			
	Neanuridae	*Bilobella braunerae*	KY230721	KY230819	KY230922		
		Friesea claviseta	KY230730	KY230826	KY230929	KY231020	KY231092
		Lathriopyga longiseta	KY230722	KY230820	KY230923		KY231086
		Micranurida pygmaea	KY230723	KY230821	KY230924		KY231087
		Pseudachorutes sp1	KY230754		KY230954		KY231113
		Pseudachorutes sp2		KY230886		KY231067	
	Odontellidae	*Odontellidae spp*	KY230796		KY230998	KY231074	
	Onychiuridae	*Deuteraphorura inermis L1*	KY230761	KY230858	KY230961	KY231068	KY231117
		Deuteraphorura inermis L2	KY230791	KY230887	KY230991		
		Micraphorura absoloni		KY230868	KY230971	KY231049	KY231127
		Protaphorura armata L1	KY230759	KY230856	KY230959	KY231064	KY231116
		Protaphorura armata L2	KY230788	KY230884	KY230988	KY231063	
		Protaphorura campata	KY230770	KY230867	KY230970	KY231048	KY231126
		Supraphorura furcifera	KY230792	KY230888	KY230992	KY231069	KY231085
	Tullbergiidae	*Mesaphorura sp1*	KY230786	KY230828	KY230931	KY231022	
		Mesaphorura sp3	KY230785	KY230882	KY230987	KY231062	
		Mesaphorura sp4	KY230760	KY230857	KY230960		
		Metaphorura affinis	KY230731	KY230827	KY230930	KY231021	KY231093
		Paratullbergia macdougalli	KY230790	KY230817	KY230920		
		Paratullbergia sp	KY230787	KY230883			
		Stenaphorura denisi	KY230720	KY230818	KY230921		KY231135
Symphypleona	Arrhopalitidae	*Arrhopalites caecus*	KY230780	KY230879	KY230981	KY231060	
		Pygmarrhopalites principalis		KY230800	KY230901	KY231002	
		Pygmarrhopalites pygmaeus	KY230746	KY230846	KY230947		KY231106
		Pygmarrhopalites sericus		KY230799	KY230900		
	Bouletiellidae	*Bourletiella hortensis*	KY230700		KY230898		
		Deuterosminthurus bicinctus	KY230732	KY230829	KY230932	KY231023	KY231094
		Deuterosminthurus sulphureus	KY230733	KY230830	KY230933	KY231024	KY231095
		Heterosminthurus bilineatus	KY230734	KY230831	KY230934	KY231025	KY231096
	Dicyrtomidae	*Dicyrtomina minuta*	KY230769	KY230866	KY230969	KY231047	KY231125
		Dicyrtomina ornata	KY230768	KY230865	KY230968	KY231046	

Table S2.3 Continued

Taxonomic group	Family	Species	18S rRNA	28S rRNA		Histone H3	COI
				D1-D2	D3-D5		
	Katiannidae	*Sminthurinus aureus*	KY230741	KY230839	KY230942	KY231031	KY231102
		Sminthurinus elegans	KY230782	KY230881	KY230984	KY231061	
	Sminthuridae	*Allacma fusca*	KY230735	KY230832	KY230935		KY231097
		Sminthurus viridis	KY230701	KY230798	KY230899	KY231001	
	Sminthurididae	*Sminthurides parvulus*	KY230699	KY230797	KY230897	KY231000	
		Sminthurides schoetti	KY230781		KY230982		
		Sminthurides signatus	KY230697		KY230895	KY230999	
		Sphaeridia pumilis L1	KY230783		KY230985		KY231133
		Sphaeridia pumilis L2	KY230758		KY230958		KY231115
		Sphaeridia pumilis L3	KY230698		KY230896		
		Stenacidia violacea	KY230784		KY230986		KY231134
Neelipleona	Neelidae	*Megalothorax minimus L1*	KY230793	KY230889	KY230993		
		Megalothorax minimus L2		KY230890		KY231070	
		Megalothorax minimus L3	KY230794	KY230891	KY230994	KY231071	
		Neelides minutus	KY230716	KY230813	KY230916		
		Neelus murinus L1		KY230892	KY230995	KY231072	
		Neelus murinus L2	KY230745	KY230844			KY231104
Tomoceridae	Tomoceridae	*Pogonognathellus flavescens L1*	KY230717	KY230814	KY230917	KY231016	KY231083
		Pogonognathellus flavescens L2	KY230718	KY230816	KY230919	KY231075	KY231084
		Pogonognathellus longicornis		KY230845	KY230946	KY231035	KY231105
		Tomocerus vulgaris	KY230777	KY230877	KY230979	KY231058	KY231131
Isotomidae	Isotomidae	*Desoria trispinata*	KY230773	KY230874	KY230976	KY231054	
		Desoria violacea	KY230736	KY230833	KY230936	KY231026	KY231098
		Folsomia cf. candida	KY230756	KY230854	KY230956	KY231042	
		Folsomia inoculata	KY230743	KY230842	KY230944	KY231033	
		Folsomia manolachei	KY230706	KY230803	KY230906	KY231007	KY231077
		Folsomia quadrioculata	KY230755	KY230853	KY230955		
		Folsomia spinosa	KY230707	KY230804	KY230907	KY231008	
		Folsomides parvulus	KY230742	KY230841	KY230943	KY231032	
		Isotoma anglicana	KY230703	KY230801	KY230903	KY231004	KY231076
		Isotoma viridis L1	KY230752	KY230852	KY230953	KY231041	KY231112
		Isotoma viridis L2	KY230774	KY230875	KY230977	KY231055	KY231129
		Isotomiella minor	KY230744	KY230843	KY230945	KY231034	KY231103
		Isotomodes productus	KY230709	KY230805	KY230908	KY231009	
		Isotomurus fucicolus	KY230737	KY230834	KY230937	KY231027	KY231099
		Isotomurus graminis	KY230738	KY230836	KY230939		KY231100
		Isotomurus italicus	KY230739	KY230837	KY230940	KY231029	KY231101
		Parisotoma notabilis L0	KY230772	KY230872	KY230974	KY231052	KY231128
		Parisotoma notabilis L4	KY230771	KY230870	KY230973	KY231050	
		Parisotoma notabilis L5		KY230871		KY231051	
		Vertagopus arboreus L1	KY230705	KY230802	KY230905	KY231006	
		Vertagopus arboreus L2	KY230775	KY230876	KY230978	KY231056	KY231130
Entomobryoidea	Entomobryidae	*Entomobrya muscorum*	KY230710	KY230806	KY230909	KY231010	
		Entomobrya nicoleti	KY230740	KY230838	KY230941	KY231030	KY231120
		Willowsia buski	KY230767	KY230864	KY230967		KY231124
	Lepidocyrtidae	*Lepidocyrtus cyaneus*	KY230751	KY230851	KY230952	KY231040	KY231111
		Lepidocyrtus lanuginosus	KY230765	KY230862	KY230965	KY231044	KY231123

Table S2.3 Continued

Taxonomic group	Family	Species	18S rRNA	28S rRNA		Histone H3	COI
				D1-D2	D3-D5		
		Lepidocyrtus lanuginosus L2	KY230748	KY230848	KY230949	KY231037	KY231108
		Lepidocyrtus lanuginosus L3	KY230763	KY230860	KY230963		KY231121
		Lepidocyrtus lignorum L1	KY230764	KY230861	KY230964	KY231043	KY231122
		Lepidocyrtus lignorum L2	KY230766	KY230863	KY230966	KY231045	
		Lepidocyrtus lignorum L3	KY230750	KY230850	KY230951	KY231039	KY231110
		Lepidocyrtus lignorum L4	KY230749	KY230849	KY230950	KY231038	KY231109
		Lepidocyrtus paradoxus	KY230795	KY230894	KY230996	KY231073	
		Lepidocyrtus cf. violaceus	KY230778	KY230878	KY230980	KY231059	KY231132
		Lepidocyrtus cf. weidneri		KY230893			KY231137
		Pseudosinella alba	KY230711	KY230807	KY230910	KY231011	
		Pseudosinella immaculata	KY230712		KY230911	KY231012	KY231078
	Orchesellidae	*Heteromurus nitidus*	KY230776	KY230808	KY230912	KY231013	KY231079
		Orchesella bifasciata	KY230713	KY230810	KY230914		KY231081
		Orchesella flavescens	KY230714	KY230811	KY230915	KY231015	KY231082
		Orchesella quinquefasciata		KY230809	KY230913	KY231014	KY231080
		Orchesella villosa	KY230729	KY230825	KY230928	KY231019	KY231091

Table S2.4

NCBI Accession Numbers of sequences used and generated in this study to construct the molecular phylogeny of Collembola.

Taxonomic group	Family	Genus	Species	Code	Source	18S rRNA	28S rRNA			Histone H3	COI
							D1	D2	D3-5		
Poduromorpha	Brachystomellidae	*Brachystomella*	*parvula*	AF-HQ	NCBI		AF483360	AF483418	HQ731912		
				JC	This study	KY230724	KY230822	KY230822	KY230925	KY231017	KY231088
	Hypogastruridae	*Ceratophysella*	*denticulata*	ES2	NCBI				KF684422	KF684752	KF684587
				FR	NCBI				KF684382	KF684712	KF684547
				GR	NCBI				KF684390	KF684720	KF684555
				IT2	NCBI				KF684395	KF684725	KF684560
				JC	This study	KY230747	KY230847	KY230847	KY230948	KY231036	KY231107
				ME	NCBI				KF684399	KF684729	KF684564
			gibbosa		NCBI		AF483362	AF483420	HQ731924		HQ732039
			luteospina		This study	KY230762	KY230859	KY230859	KY230962		KY231118
			succinea		This study		KY230885	KY230885	KY230990	KY231065	KY231136
		Hypogastrura	*burkilli*	JC	This study	KY230725	KY230823	KY230823	KY230926	KY231018	KY231089
			purpurescens	HQ	NCBI		HQ732118	HQ732009	HQ731942		HQ732058
			socialis		NCBI		LK024330	LK024330			HG422630
			sp	X	NCBI	X90677	X90679		X90680		
			subboldorii		NCBI		HQ732124	HQ732015	HQ731948		HQ732064
			vernalis		NCBI		AF483371	AF483429	HQ731949		HQ732065
		Microgastrura	*sensiliata*		NCBI		AF483379	AF483437	HQ731958		
		Paraxenylla	*affiniformis*		NCBI		AF483387	AF483445	HQ731962		
		Schoettella	*ununguiculata*		NCBI		AF483395	AF483453	HQ731967		HQ732079
		Triacanthella	*perfecta*		NCBI		AF483407	AF483465	HQ731972		
		Willemia	*anophthalma*		This study	KY230726	KY230869	KY230869	KY230972		KY231090
			denisi		This study	KY230757	KY230855	KY230855	KY230957		KY231114
		Xenylla	*boerneri*		This study/NCBI	KY230728	KY230824	KY230824			KT808349
			grisea		NCBI	AY555517	AF483409	AF483467	HQ731974	AY555563	
			humicola		NCBI		LK024440	LK024440			HG422632
			tullbergi		NCBI		AF483410	AF483468	HQ731980		HQ732088
		Xenyllodes	*armatus*		NCBI		AF483411	AF483469			HQ732089
	Neanuridae	*Anurida*	*granaria*		NCBI		AF483356	AF483414	HQ731909		KT808325
			maritima		NCBI	EU368603	AF483357	AF483415	HQ731910		HQ732028
		Anuridella	*calcarata*		NCBI		AF483358	AF483416	HQ731911		
		Bilobella	*aurantiaca*		NCBI	EU368602	AJ251729	AJ251729	AJ251729		GU656564
			braunerae		This study/NCBI	KY230721	KY230819	KY230819	KY230922		KT808394
			massoudi		NCBI		AJ251732	AJ251732	AJ251732		JN283905
		Cansilianura	*malatestai*		NCBI		AJ251737	AJ251737	AJ251737		
		Deutonura	*conjuncta*		NCBI		AJ251735	AJ251735	AJ251735		
			gibbosa		NCBI			JX261725			HM208637
			monticola	L1	NCBI			JX261695			JX261842
				L2	NCBI			JX261677			JX261817
				L3	NCBI			JX261656			HM397763
			provincialis		NCBI			JX261655			JX261795
		Friesea	*claviseta*	JC	This study	KY230730	KY230826	KY230826	KY230929	KY231020	KY231092
			mirabilis		NCBI		AF483368	AF483426			
			truncata		NCBI		AF483369	AF483427			

Table S2.4 Continued

Taxonomic group	Family	Genus	Species	Code	Source	18S rRNA	28S rRNA			Histone H3	COI
							D1	D2	D3-5		
		Lathriopyga	*longiseta*	AJ-KT	NCBI		AJ251736	AJ251736	AJ251736		KT808385
				JC	This study	KY230722	KY230820	KY230820	KY230923		KY231086
		Micranurida	*pygmaea*	AF	NCBI		AF483377	AF483435			
				JC	This study	KY230723	KY230821	KY230821	KY230924		KY231087
		Monobella	*grassei*		NCBI		AF483380	AF483438			KT808343
		Morulina	*verrucosa*		NCBI	AY555521	AF483381	AF483439	HQ731959	AY555565	
		Neanura	*muscorum*	AY-AJ	NCBI	AY555520	AJ251733	AJ251733	AJ251733		AY555544
				LK	NCBI		LK024367	LK024367			LK024486
		Pseudachorutes	*palmiensis*		NCBI		AF483392	AF483450	HQ731966		
			sp	EU	NCBI		EU371045	EU371045	EU371045		
				JC	This study	KY230754	KY230886	KY230886	KY230954	KY231067	KY231113
			subcrassus		NCBI		LK024410	LK024410			LK024515
		Thaumanura	*carolii*		NCBI		AF483403	AF483461			
			ruffoi		NCBI		AJ251734	AJ251734	AJ251734		
	Odontellidae	*Odontelidae*	*spp*		This study	KY230796			KY230998	KY231074	
		Superodontella	*alpina*		NCBI		AF483400	AF483458	HQ731968		HQ732080
			gisini		NCBI		AF483401	AF483459	HQ731969		HQ732081
	Onychiuridae	*Deuteraphorura*	*inermis*	L1	This study	KY230761	KY230858	KY230858	KY230961	KY231068	KY231117
				L2	This study	KY230791	KY230887	KY230887	KY230991		
		Kalaphorura	*paradoxa*		NCBI		AF483373	AF483431	HQ731955		HQ732071
		Micraphorura	*absoloni*		This study		KY230868	KY230868	KY230971	KY231049	KY231127
		Onychiurus	*ambulans*		NCBI	AY555518	AF483384	AF483442	HQ731961	AY555564	HQ732075
		Protaphorura	*armata*	AF-HQ	NCBI		AF483391	AF483449	HQ731965		HQ732078
				L1	This study	KY230759	KY230856	KY230856	KY230959	KY231064	KY231116
				L2	This study/NCBI	KY230788	KY230884	KY230884	KY230988	KY231063	HG422585
			sp		This study	KY230770	KY230867	KY230867	KY230970	KY231048	KY231126
		Supraphorura	*furcifera*		This study	KY230792	KY230888	KY230888	KY230992	KY231069	KY231085
		Tetrodontophora	*bielanensis*		NCBI	AY555519	EU376051	EU376051	EU376051		AY555543
	Poduridae	*Podura*	*aquatica*		NCBI	EU368604	AF483390	AF483448	HQ731964		HQ732077
	Tullbergiidae	*Mesaphorura*	*macrochaeta*		NCBI		AF483375	AF483433			
			sp1		This study	KY230786	KY230828	KY230828	KY230931	KY231022	
			sp3		This study	KY230785	KY230882	KY230882	KY230987	KY231062	
			sp4		This study	KY230760	KY230857	KY230857	KY230960		
		Metaphorura	*affinis*		This study	KY230731	KY230827	KY230827	KY230930	KY231021	KY231093
			denisi	AF-HQ	NCBI		AF483376	AF483434	HQ731956		HQ732072
		Paratullbergia	*macdougalli*		This study	KY230790	KY230817	KY230817	KY230920		
			sp		This study	KY230787	KY230883	KY230883			
		Stenaphorura	*denisi*		This study	KY230720	KY230818	KY230818	KY230921		KY231135
Symphypleona	Arrhopalitidae	*Arrhopalites*	*caecus*	AY	NCBI			AY239037			
				JC	This study	KY230780	KY230879	KY230879	KY230981	KY231060	
		Pygmarrhopalites	*principalis*		This study		KY230800	KY230800	KY230901	KY231002	
			pygmaeus		This study	KY230746	KY230846	KY230846	KY230947		KY231106
			sericus		This study		KY230799	KY230799	KY230900		

Table S2.4 Continued

Taxonomic group	Family	Genus	Species	Code	Source	18S rRNA	28S rRNA			Histone H3	COI
							D1	D2	D3-5		
	Bourletiellidae	*Bourletiella*	*hortensis*		This study	KY230700			KY230898		
		Deuterosminthurus	*bicinctus*		This study	KY230732	KY230829	KY230829	KY230932	KY231023	KY231094
			sulphureus		This study	KY230733	KY230830	KY230830	KY230933	KY231024	KY231095
		Heterosminthurus	*bilineatus*		This study	KY230734	KY230831	KY230831	KY230934	KY231025	KY231096
	Dicyrtomidae	*Dicyrtoma*	*atra*		NCBI		LK024412	LK024412			LK024518
			fusca	LK-HG	NCBI		LK024301	LK024301			HG422654
		Dicyrtomina	*minuta*		This study	KY230769	KY230866	KY230866	KY230969	KY231047	KY231125
			ornata		This study/NCBI	KY230768	KY230865	KY230865	KY230968	KY231046	KT808331
			saundersi		NCBI	EU368611	EF199974	EF199974	EF199974		
	Katiannidae	*Sminthurinus*	*aureus*	JC	This study	KY230741	KY230839	KY230839	KY230942	KY231031	KY231102
			bimaculatus		NCBI	AY555522	AF483398	AF483456	AY555536	AY555566	AY555545
			granulosus		NCBI		HQ592774	HQ592774			HQ592714
			reticulatus		This study/NCBI	KY230782	KY230881	KY230881	KY230984	KY231061	KT808391
			tuberculatus		NCBI		HQ592770	HQ592770			HQ592709
	Sminthuridae	*Allacma*	*fusca*		This study	KY230735	KY230832	KY230832	KY230935		KY231097
		Caprainea	*marginata*		NCBI		AF483361	AF483419			
		Lipothrix	*lubbocki*		NCBI		LK024363	LK024363			HG422652
		Sminthurus	*viridis*		This study/NCBI	KY230701	KY230798	KY230798	KY230899	KY231001	KC150080
	Sminthurididae	*Sminthurides*	*aquaticus*		NCBI	AY596364			AY596391		AY665315
			parvulus		This study	KY230699	KY230797	KY230797	KY230897	KY231000	
			schoetti		This study	KY230781			KY230982		
			signatus		This study	KY230697			KY230895	KY230999	
		Sphaeridia	*pumilis*	L1	This study	KY230783			KY230985		KY231133
				L2	This study	KY230758			KY230958		KY231115
				L3	This study	KY230698			KY230896		
				NCBI	NCBI	AY145140	EF192443	EF192443	EF192443		
		Stenacidia	*violacea*		This study	KY230784			KY230986		KY231134
Neelidae	Neelidae	*Megalothorax*	*minimus*	A	This study	KY230793	KY230889	KY230889	KY230993		
				EU-EF	NCBI	EU368608	EF199975	EF199975	EF199975		
				F	This study	KY230794	KY230891	KY230891	KY230994	KY231071	
				FRA11	NCBI		JN970990	JN970990			JN970914
				G	This study		KY230890	KY230890		KY231070	
			sp	BEG1	NCBI		JN971001	JN971001			JN970925
				FRA1	NCBI		JN970988	JN970988			JN970912
				FRA10	NCBI		JN970999	JN970999			JN970923
				FRA2	NCBI		JN970993	JN970993			JN970917
				FRA3	NCBI		JN970987	JN970987			JN970911
				FRA4	NCBI		JN970996	JN970996			JN970920
				FRA8	NCBI		JN970989	JN970989			JN970913
				LK-HG	NCBI		LK024365	LK024365			HG422646
		Neelides	*minutus*	JC	This study	KY230716	KY230813	KY230813	KY230916		
			sp		NCBI		EF422366	EF422366	EF422366		

Table S2.4 Continued

Taxonomic group	Family	Genus	Species	Code	Source	18S rRNA	28S rRNA			Histone H3	COI
							D1	D2	D3-5		
		Neelus	*koseli*		NCBI		JN971008	JN971008			JN970932
			murinus	5F	This study	KY230745	KY230844	KY230844			KY231104
				6F	This study		KY230892	KY230892	KY230995	KY231072	
			sp	FRA1	NCBI		JN971006	JN971006			JN970930
				FRA2	NCBI		JN971007	JN971007			JN970931
Tomoceridae	Oncopoduridae	*Oncopodura*	*crassicornis*		NCBI	DQ016563	AF483383	AF483441			
	Tomoceridae	*Pogonognathellus*	*flavescens*	L1	This study	KY230717	KY230814	KY230814	KY230917	KY231016	KY231083
				L2	This study	KY230718	KY230816	KY230816	KY230919	KY231075	KY231084
			longicornis	JC	This study		KY230845	KY230845	KY230946	KY231035	KY231105
				KT	NCBI	EU368607	EU376053	EU376053	EU376053		KT808358
		Tomocerus	*baudoti*	L1	NCBI			JX261697			JX261845
				L2	NCBI			JX261684			GU656231
			minor	AT-AF-KT1	NCBI	AY555516	AF483406	AF483464	HQ731971	AY555562	KT808379
				JX-HM	NCBI			JX261700			HM398041
			vulgaris		This study	KY230777	KY230877	KY230877	KY230979	KY231058	KY231131
Isotomidae	Isotomidae	*Agrenia*	*riparia*		NCBI	JN981032	JN981055	JN981055	JN981055		JN981076
		Anurophorus	*septentrionalis*		NCBI		LK024296	LK024296			LK024446
		Archisotoma	*besselsi*		NCBI		JN981045	JN981045	JN981045		JN981071
		Desoria	*tigrina*	JN	NCBI	JN981033	JN981056	JN981056	JN981056		JN981078
			trispinata		This study/NCBI	KY230773	KY230874	KY230874	KY230976	KY231054	KT808356
			violacea		This study	KY230736	KY230833	KY230833	KY230936	KY231026	KY231098
		Folsomia	*candida*		This study/NCBI	KY230756	KY230854	KY230854	KY230956	KY231042	AY555541
			fimetaria		NCBI		LK024319	LK024319			HG422609
			inoculata		This study	KY230743	KY230842	KY230842	KY230944	KY231033	
			manolachei		This study	KY230706	KY230803	KY230803	KY230906	KY231007	KY231077
			penicula		NCBI	JN981026	JN981049	JN981049	JN981049		
			quadrioculata	DE	This study/NCBI	KY230755	KY230853	KY230853	KY230955	KF684772	KF684607
				ES3	NCBI				KF684479	KF684809	KF684644
				GR_Fq3	NCBI				KF684449	KF684779	KF684614
				GR_Fq4	NCBI				KF684450	KF684780	KF684615
				HR_Fq3	NCBI				KF684430	KF684760	KF684595
				ME	NCBI				KF684461	KF684791	KF684626
				NO	NCBI				KF684464	KF684794	KF684629
				RS	NCBI				KF684474	KF684804	KF684639
				RU1	NCBI				KF684468	KF684798	KF684633
			spinosa		This study	KY230707	KY230804	KY230804	KY230907	KY231008	
		Folsomides	*parvulus*		This study/NCBI	KY230742	KY230841	KY230841	KY230943	KY231032	JN981069
		Folsomina	*onychiurina*		NCBI	JN981028	JN981051	JN981051	JN981051		JN981073
		Isotoma	*anglicana*		This study	KY230703	KY230801	KY230801	KY230903	KY231004	KY231076
			riparia		NCBI		LK024337	LK024337			HG422621
			viridis	L1	This study	KY230752	KY230852	KY230852	KY230953	KY231041	KY231112
				L2	This study	KY230774	KY230875	KY230875	KY230977	KY231055	KY231129

Table S2.4 Continued

Taxonomic group	Family	Genus	Species	Code	Source	18S rRNA	28S rRNA			Histone H3	COI
							D1	D2	D3-5		
		Isotomiella	*minor*	ES1	NCBI				KF684534	KF684863	KF684698
				ES2	NCBI				KF684530	KF684865	KF684700
				FR	NCBI				KF684492	KF684822	KF684657
				GR	NCBI				KF684500	KF684830	KF684665
				HR	NCBI				KF684489	KF684819	KF684654
				JC	This study	KY230744	KY230843	KY230843	KY230945	KY231034	KY231103
				ME	NCBI				KF684511	KF684841	KF684676
				NO	NCBI				KF684516	KF684846	KF684681
		Isotomodes	*productus*		This study	KY230709	KY230805	KY230805	KY230908	KY231009	
			sp		NCBI	JN981023	JN981042	JN981042	JN981042		JN981070
		Isotomurus	*fucicolus*		This study	KY230737	KY230834	KY230834	KY230937	KY231027	KY231099
			graminis		This study	KY230738	KY230836	KY230836	KY230939		KY231100
			italicus		This study	KY230739	KY230837	KY230837	KY230940	KY231029	KY231101
			maculatus		NCBI				AJ251740		KT808362
			palustris		NCBI	DQ016560					KT808340
		Parisotoma	*notabilis*	L0	This study	KY230772	KY230872	KY230872	KY230974	KY231052	KY231128
				L1	NCBI				KJ792144	KJ792384	KJ792264
				L2	NCBI		LK024387	LK024387	KJ792148	KJ792388	KJ792268
				L3	NCBI				KJ792162	KJ792402	KJ792282
				L4	This study/NCBI	KY230771	KY230870	KY230870	KY230973	KY231050	KJ792230
				L5	This study/NCBI		KY230871	KY230871		KY231051	JN970935
		Pseudisotoma	*sensibilis*	JN	NCBI		JN981065	JN981065	JN981065		JN981086
		Scutisotoma	*subarctica*		NCBI	JN981030	JN981053	JN981053	JN981053		
		Tetracanthella	*sp*		NCBI	JN981025	JN981044	JN981044	JN981044		
			wahlgreni		NCBI	JN981024	JN981043	JN981043	JN981043		
		Vertagopus	*arboreus*	4F	This study	KY230775	KY230876	KY230876	KY230978	KY231056	KY231130
				5F	This study	KY230705	KY230802	KY230802	KY230905	KY231006	
			sp		NCBI	JN981039	JN981066	JN981066	JN981066		
Entomobryoidea	Entomobryidae	*Coecobrya*	*tenebricosa*		NCBI	KC236231	KC236272	KC236272	KC236272		
		Entomobrya	*corticalis*		NCBI		LK024307	LK024307			LK024454
			lanuginosa		NCBI		AF483365	AF483423			JN970907
			marginata		NCBI		LK024309	LK024309			HG422600
			multifasciata		NCBI	KC236237	KC236276	KC236276	KC236276		KM610123
			muscorum		This study	KY230710	KY230806	KY230806	KY230909	KY231010	
			nicoleti		This study	KY230740	KY230838	KY230838	KY230941	KY231030	KY231120
			nivalis		NCBI		LK024313	LK024313			HG422598
		Sinella	*curviseta*		NCBI	DQ016565	EF192441	EF192441	EF192441		
		Willowsia	*buski*		This study	KY230767	KY230864	KY230864	KY230967		KY231124
			nigromaculata		NCBI	KC236263	KC236304	KC236304	KC236304		

Table S2.4 Continued

Taxonomic group	Family	Genus	Species	Code	Source	18S rRNA	28S rRNA			Histone H3	COI
							D1	D2	D3-5		
	Lepidocyrtidae	*Lepidocyrtus*	*cyaneus*	JC	This study	KY230751	KY230851	KY230851	KY230952	KY231040	KY231111
			lanuginosus	like1	This study	KY230748	KY230848	KY230848	KY230949	KY231037	KY231108
				like2	This study	KY230763	KY230860	KY230860	KY230963		KY231121
				std	This study	KY230765	KY230862	KY230862	KY230965	KY231044	KY231123
			lignorum	L1	This study	KY230764	KY230861	KY230861	KY230964	KY231043	KY231122
				L2	This study	KY230766	KY230863	KY230863	KY230966	KY231045	
				L3	This study	KY230750	KY230850	KY230850	KY230951	KY231039	KY231110
				L4	This study	KY230749	KY230849	KY230849	KY230950	KY231038	KY231109
			paradoxus	JC	This study	KY230795	KY230894	KY230894	KY230996	KY231073	
				U-JN-KJ	NCBI	U61301	JN981068	JN981068	JN981068		KJ207897
			sp1		This study	KY230778	KY230878	KY230878	KY230980	KY231059	KY231132
			sp2		This study		KY230893	KY230893			KY231137
		Pseudosinella	*alba*	JC	This study	KY230711	KY230807	KY230807	KY230910	KY231011	
				KC-KT	NCBI	KC236253	KC236295	KC236295	KC236295		KT808338
			immaculata		This study	KY230712			KY230911	KY231012	KY231078
	Orchesellidae	*Heteromurus*	*major*	KC	NCBI	KC236241	KC236282	KC236282	KC236282		
				L1	NCBI			JX261711			HM397731
				L2	NCBI			JX261671			JX261806
				L3	NCBI			JX261679			JX261820
				L4	NCBI			JX261657			GU656499
				L5	NCBI			JX261654			GU656225
				L6	NCBI			JX261653			JX261794
				L7	NCBI			JX261732			JX261885
			nitidus	JC	This study	KY230776	KY230808	KY230808	KY230912	KY231013	KY231079
				KC-JX	NCBI	KC236242	KC236283	KC236283	KC236283		JX261868
		Orchesella	*bifasciata*		This study	KY230713	KY230810	KY230810	KY230914		KY231081
			cincta	AF-JN	NCBI		AF483385	AF483443			JN970933
				KC-KT	NCBI	KC236250	KC236290	KC236290	KC236290		KT808383
			flavescens	JC	This study	KY230714	KY230811	KY230811	KY230915	KY231015	KY231082
				LK-HG	NCBI		LK024378	LK024378			HG422627
			quinquefasciata		This study		KY230809	KY230809	KY230913	KY231014	KY231080
			villosa		This study	KY230729	KY230825	KY230825	KY230928	KY231019	KY231091

Table S2.5

Trait attributes of the 75 Collembola species used in this study.

Species	Body		Vertical stratification	Pigmentation		Ommatidia		Reproductive mode**	Density (individuals/m^2)				Occurrence		
	*Lengtl	Shape		Color	Level	Number	Level		Arable	Grassland	Forest	All	Arable	Grassland	Forest
Allacma fusca	4.0	spheric	epedaphic	dark brown	dark	8	many	bisexual	0	0	390	390	0	0	1
Arrhopalites caecus	1.0	spheric	euedaphic	white or pinkish	pale	1	few	parthenogenetic	0	509	0	509	0	1	0
Bilobella braunerae	2.5	stocky	hemiedaphic	salmon pink	bright	2	few	?	0	0	65	65	0	0	1
Bourletiella hortensis	1.8	spheric	epedaphic	bluish black	dark	8	many	bisexual	1689	0	0	1689	1	0	0
Brachystomella parvula	1.0	stocky	hemiedaphic	bluish violet	dark	8	many	bisexual	0	9373	0	9373	0	1	0
Ceratophysella denticulata	1.8	stocky	hemiedaphic	bluish grey	dark	8	many	bisexual	0	0	5665	5665	0	0	1
Ceratophysella luteospina	1.5	stocky	hemiedaphic	bluish grey	dark	8	many	?	0	0	1538	1538	0	0	1
Ceratophysella succinea	1.8	stocky	hemiedaphic	bluish grey	dark	8	many	parthenogenetic	0	27512	0	27512	0	1	0
Desoria trispinata	1.3	slender	hemiedaphic	light blue	bright	8	many	parthenogenetic	1528	4649	0	6177	1	1	0
Desoria violacea	2.4	slender	hemiedaphic	violet black	dark	8	many	bisexual	195	2102	0	2297	1	1	0
Deuteraphorura inermis	2.3	slender	euedaphic	white	pale	0	none	bisexual	509	0	1019	1528	1	0	1
Deuterosminthurus bicinctus	0.8	spheric	epedaphic	yellow with violet patches	bright	8	many	bisexual	0	0	520	520	0	0	1
Deuterosminthurus sulphureus	1.0	spheric	epedaphic	uniformly yellow	bright	8	many	bisexual	1094	4028	0	5122	1	1	0
Dicyrtomina ornata	3.0	spheric	epedaphic	irregular brown patches of pigment	bright	8	many	bisexual	0	195	2923	3118	0	1	1
Entomobrya muscorum	3.5	slender	epedaphic	characteristic pattern	bright	8	many	?	0	0	2533	2533	0	0	1
Entomobrya nicoleti	2.0	slender	epedaphic	characteristic pattern	bright	8	many	?	195	1169	65	1429	1	1	1
Folsomia candida	3.0	slender	euedaphic	white	pale	0	none	parthenogenetic	0	0	1019	1019	0	0	1
Folsomia inoculata	1.9	slender	euedaphic	white	pale	0	none	bisexual	2037	3565	1084	6686	1	1	1
Folsomia manolachei	1.0	slender	hemiedaphic	greyish brown	bright	2	few	bisexual	24955	246824	195	271974	1	1	1
Folsomia quadrioculata	2.5	slender	hemiedaphic	greyish brown	bright	2	few	bisexual	8788	8658	31179	48625	1	1	1
Folsomia spinosa	1.3	slender	euedaphic	white	pale	0	none	parthenogenetic	1019	0	509	1528	1	0	1
Folsomides parvulus	0.9	slender	euedaphic	white	pale	2	few	parthenogenetic	25974	0	0	25974	1	0	0
Friesea claviseta	1.4	stocky	hemiedaphic	bluish grey	dark	8	many	?	3121	4584	2362	10066	1	1	1
Heteromurus nitidus	3.0	slender	euedaphic	pale brownish to white	pale	1	few	?	6218	1658	65	7941	1	1	1
Heterosminthurus bilineatus	1.0	spheric	epedaphic	yellowish with dark bands	bright	8	many	bisexual	0	1299	0	1299	0	1	0
Hypogastrura burkilli	2.8	stocky	hemiedaphic	greyish blue	dark	8	many	?	0	0	1549	1549	0	0	1
Isotoma anglicana	4.0	slender	hemiedaphic	bluish or violet	dark	8	many	bisexual	65	2947	0	3012	1	1	0
Isotoma viridis	6.0	slender	hemiedaphic	greenish	bright	8	many	bisexual	79307	79959	0	159267	1	1	0
Isotomiella minor	1.2	slender	euedaphic	white	pale	0	none	parthenogenetic	0	0	29123	29123	0	0	1
Isotomodes productus	1.0	slender	euedaphic	white	pale	0	none	parthenogenetic	0	509	0	509	0	1	0
Isotomurus fucicolus	2.5	slender	hemiedaphic	greenish	bright	8	many	bisexual	6813	11701	0	18514	1	1	0
Isotomurus graminis	3.0	slender	hemiedaphic	greenish	bright	8	many	bisexual	2058	11160	0	13218	1	1	0
Isotomurus italicus	2.0	slender	hemiedaphic	violet	bright	8	many	bisexual	0	3703	0	3703	0	1	0
Lathriopyga longiseta	2.5	stocky	hemiedaphic	bluish grey	dark	2	few	?	0	0	520	520	0	0	1
Lepidocyrtus cyaneus	1.5	slender	epedaphic	bluish violet	dark	8	many	parthenogenetic	4495	46465	0	50961	1	1	0
Lepidocyrtus lanuginosus	2.0	slender	epedaphic	white brown	bright	8	many	?	11568	14723	33673	59964	1	1	1
Lepidocyrtus lignorum	2.0	slender	epedaphic	white brown	bright	8	many	?	0	27712	21222	48934	0	1	1

Table S2.5 Continued

Species	Body		Vertical stratification	Pigmentation		Ommatidia		Reproductive mode**	Density (individuals/m²)				Occurrence		
	*Lengtl	Shape		Color	Level	Number	Level		Arable	Grassland	Forest	All	Arable	Grassland	Forest
Lepidocyrtus weidneri	2.0	slender	epedaphic	bluish violet	dark	8	many	?	0	260	0	260	0	1	0
Megalothorax minimus	0.4	spheric	euedaphic	white	pale	0	none	parthenogenetic	1528	1528	11150	14206	1	1	1
Mesaphorura sp	0.7	slender	euedaphic	white	pale	0	none	parthenogenetic	4584	16297	55014	75895	1	1	1
Metaphorura affinis	1.3	slender	euedaphic	white	pale	0	none	bisexual	0	1528	0	1528	0	1	0
Micranurida pygmaea	0.6	stocky	euedaphic	white to bluish grey	pale	2	few	parthenogenetic	0	0	1213	1213	0	0	1
Micraphorura absoloni	0.9	slender	euedaphic	white	pale	0	none	parthenogenetic	0	0	1149	1149	0	0	1
Neelides minutus	0.6	spheric	euedaphic	pale brown	pale	0	none	parthenogenetic	0	0	715	715	0	0	1
Neelus murinus	1.0	spheric	euedaphic	pale brown	pale	0	none	parthenogenetic	0	0	325	325	0	0	1
Orchesella bifasciata	3.0	slender	epedaphic	characteristic pattern, brown	bright	8	many	bisexual	0	0	65	65	0	0	1
Orchesella flavescens	5.0	slender	epedaphic	characteristic pattern, brown	bright	8	many	bisexual	0	0	1429	1429	0	0	1
Orchesella quinquefasciata	5.0	slender	epedaphic	characteristic pattern, brown	bright	8	many	bisexual	195	65	0	260	1	1	0
Orchesella villosa	5.5	slender	epedaphic	characteristic pattern, brown	bright	8	many	bisexual	2393	65	65	2523	1	1	1
Parisotoma notabilis	1.0	slender	hemiedaphic	greyish brown	bright	4	few	parthenogenetic	65395	146285	129481	341161	1	1	1
Pogonognathellus flavescens	6.5	slender	hemiedaphic	greyish brown	bright	6	many	bisexual	0	260	4093	4352	0	1	1
Pogonognathellus longicornis	6.5	slender	epiedaphic	greyish brown	bright	6	many	bisexual	0	0	260	260	0	0	1
Protaphorura sp	2.5	slender	euedaphic	white	pale	0	none	bisexual	11205	30558	20382	62144	1	1	1
Pseudachorutes sp	1.5	stocky	hemiedaphic	dark blue	dark	8	many	?	0	0	574	574	0	0	1
Pseudosinella alba	1.1	slender	euedaphic	white	pale	2	few	parthenogenetic	5428	12969	3251	21648	1	1	1
Pseudosinella immaculata	1.9	slender	euedaphic	white	pale	0	none	?	65	65	0	130	1	1	0
Pygmarrhopalites principalis	1.0	spheric	euedaphic	pale reddish	pale	1	few	parthenogenetic	65	0	65	130	1	0	1
Pygmarrhopalites pygmaeus	1.2	spheric	euedaphic	white pinkish	pale	1	few	bisexual	0	0	509	509	0	0	1
Pygmarrhopalites sericus	0.8	spheric	euedaphic	pale grey	pale	1	few	parthenogenetic	0	0	65	65	0	0	1
Sminthurides parvulus	0.6	spheric	hemiedaphic	light blue	bright	8	many	bisexual	0	0	844	844	0	0	1
Sminthurides schoetti	0.5	spheric	hemiedaphic	yellow with blue band	bright	8	many	bisexual	0	4584	0	4584	0	1	0
Sminthurides signatus	0.5	spheric	epedaphic	yellowish with violet patches	bright	8	many	bisexual	520	130	0	650	1	1	0
Sminthurinus aureus	1.0	spheric	hemiedaphic	yellow, brown	bright	8	many	bisexual	13408	7842	130	21380	1	1	1
Sminthurinus elegans	0.7	spheric	hemiedaphic	yellow with bluish bands	bright	8	many	bisexual	195	0	4158	4352	1	0	1
Sminthurus viridis	3.0	spheric	epedaphic	greenish	bright	8	many	bisexual	1039	2523	260	3822	1	1	1
Sphaeridia pumilis	0.5	spheric	hemiedaphic	pinkish or greyish yellow	bright	8	many	bisexual	65	17737	195	17997	1	1	1
Stenacidia violacea	0.7	spheric	hemiedaphic	violet	dark	8	many	bisexual	834	1593	520	2947	1	1	1
Stenaphorura denisi	1.3	slender	euedaphic	white	pale	0	none	bisexual	3056	1019	1019	5093	1	1	1
Supraphorura furcifera	1.9	slender	euedaphic	white	pale	0	none	bisexual	0	0	6738	6738	0	0	1
Tomocerus vulgaris	4.0	slender	hemiedaphic	greyish black	bright	6	many	bisexual	574	585	0	1159	1	1	0
Vertagopus arboreus	1.8	slender	hemiedaphic	bluish violet	dark	8	many	bisexual	0	0	195	195	0	0	1
Willemia anophthalma	0.7	slender	euedaphic	white	pale	0	none	parthenogenetic	509	509	2167	3186	1	1	1
Willemia denisi	1.0	slender	euedaphic	white	pale	0	none	parthenogenetic	0	0	130	130	0	0	1
Willowsia buski	1.5	slender	hemiedaphic	brownish violet	dark	8	many	?	0	0	65	65	0	0	1
Xenylla boerneri	0.7	stocky	hemiedaphic	bluish grey	dark	5	few	?	0	0	130	130	0	0	1

* Body length in mm.

** Species with unknown reproductive mode "?" were pre-defined as bisexual in phylogenetic signal measurement and trait similarity matrix.

Table S2.6(a)

Estimated number of transitions across 1,000 stochastic character mapping simulations of body shape of Collembola (N = 6, median; 5.8 ± 1.0, mean ± SD).

To:	**Slender**				**Stocky**				**Spheric**			
From:	**Median**	**Mean**	**±**	**SD**	**Median**	**Mean**	**±**	**SD**	**Median**	**Mean**	**±**	**SD**
Slender					**1**	1.2	±	0.7	**2**	2.3	±	0.6
Stocky	**2**	1.9	±	0.7					**0**	0.1	±	0.4
Spheric	**0**	0.2	±	0.7	**0**	0.1	±	0.3				

Table S2.6(b)

Estimated number of transitions across 1,000 stochastic character mapping simulations of vertical stratification of Collembola (N = 25, median; 26.1 ± 4.8, mean ± SD).

To:	**Epedaphic**				**Hemiedaphic**				**Euedaphic**			
From:	**Median**	**Mean**	**±**	**SD**	**Median**	**Mean**	**±**	**SD**	**Median**	**Mean**	**±**	**SD**
Epedaphic					**3**	3.4	±	1.8	**4**	4.2	±	1.5
Hemiedaphic	**5**	5.0	±	1.9					**9**	8.4	±	2.5
Euedaphic	**1**	1.4	±	1.6	**3**	3.7	±	2.8				

Table S2.6(c)

Estimated number of transitions across 1,000 stochastic character mapping simulations of pigmentation of Collembola (N = 33, median; 34.3 ± 7.1, mean ± SD).

To:	**Dark**				**Bright**				**Pale**			
From:	**Median**	**Mean**	**±**	**SD**	**Median**	**Mean**	**±**	**SD**	**Median**	**Mean**	**±**	**SD**
Dark					**4**	4.4	±	2.6	**5**	5.1	±	2.3
Bright	**8**	8.6	±	2.4					**8**	7.9	±	2.9
Pale	**3**	3.3	±	2.5	**4**	5.0	±	3.1				

Table S2.6(d)

Estimated number of transitions across 1,000 stochastic character mapping simulations of number of ommatidia of Collembola (N = 19, median; 19.8 ± 3.3, mean ± SD).

To:	**Many**				**Few**				**None**			
From:	**Median**	**Mean**	**±**	**SD**	**Median**	**Mean**	**±**	**SD**	**Median**	**Mean**	**±**	**SD**
Many					**6**	6.5	±	1.4	**8**	7.7	±	1.6
Few	**1**	0.8	±	1.0					**1**	1.0	±	1.2
None	**1**	1.3	±	1.4	**3**	2.7	±	1.1				

Table S2.6(e)

Estimated number of transitions across 1,000 stochastic character mapping simulations of reproductive mode of Collembola (N = 19, median; 19.9 ± 4.26, mean ± SD).

To:	**Parthenogenetic**				**Sexual**			
From:	**Median**	**Mean**	**±**	**SD**	**Median**	**Mean**	**±**	**SD**
Parthenogenetic					**5**	6.0	±	2.8
Sexual	**13**	13.9	±	2.5				

Table S2.7

P-values of Pagel's correlation test between the substitution rate of the traits in the first column and that of the traits in the first row. Significance is determined by one direction, i.e., traits in the first column depends on the traits in the first row. NS, non-significant.

	Dependent on:	**Body size**			**Body shape**			**Stratification**			**Pigmentation**			**Ommatidia**			**Repr.**
		Large	**Medium**	**Small**	**Slender**	**Stocky**	**Spheric**	**Epe.**	**Hemi.**	**Eue.**	**Dark**	**Bright**	**Pale**	**Many**	**Few**	**None**	**Parth.**
Body size	**Large**		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
	Medium	NS		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
	Small	NS	NS		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Body shape	**Slender**	NS	NS	NS		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
	Stocky	NS	NS	NS	NS		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
	Spheric	NS	NS	NS	NS	NS		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Stratification	**Epedaphic**	NS	NS	NS	NS	NS	NS		NS	NS	NS	**0.008**	NS	**0.018**	NS	NS	NS
	Hemiedaphic	NS	NS	NS	NS	NS	NS	NS		NS	NS	NS	**0.018**	NS	NS	NS	NS
	Euedaphic	NS	NS	NS	NS	NS	NS	NS	NS		NS	**0.009**	**0.000**	**0.000**	**0.040**	**0.001**	**0.001**
Pigmentation	**Dark**	NS	NS	NS	NS	NS	NS	NS	**0.039**	**0.007**		NS	NS	NS	NS	**0.026**	NS
	Bright	NS	NS	NS	NS	NS	NS	**0.037**	NS	**0.002**	NS		NS	NS	NS	**0.009**	NS
	Pale	NS	NS	NS	NS	NS	NS	NS	**0.014**	**0.000**	NS	NS		**0.000**	**0.034**	**0.001**	**0.001**
Ommatidia	**Many**	NS	NS	NS	NS	NS	NS	NS	NS	**0.001**	NS	**0.027**	**0.001**		NS	NS	**0.035**
	Few	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS		NS	**0.035**
	None	NS	NS	NS	NS	NS	NS	NS	**0.040**	**0.000**	NS	**0.019**	**0.000**	NS	NS		**0.004**
Reproduction	**Parthenogenetic**	NS	NS	**0.008**	NS	NS	NS	NS	NS	**0.000**	NS	**0.001**	**0.000**	**0.000**	**0.029**	**0.012**	

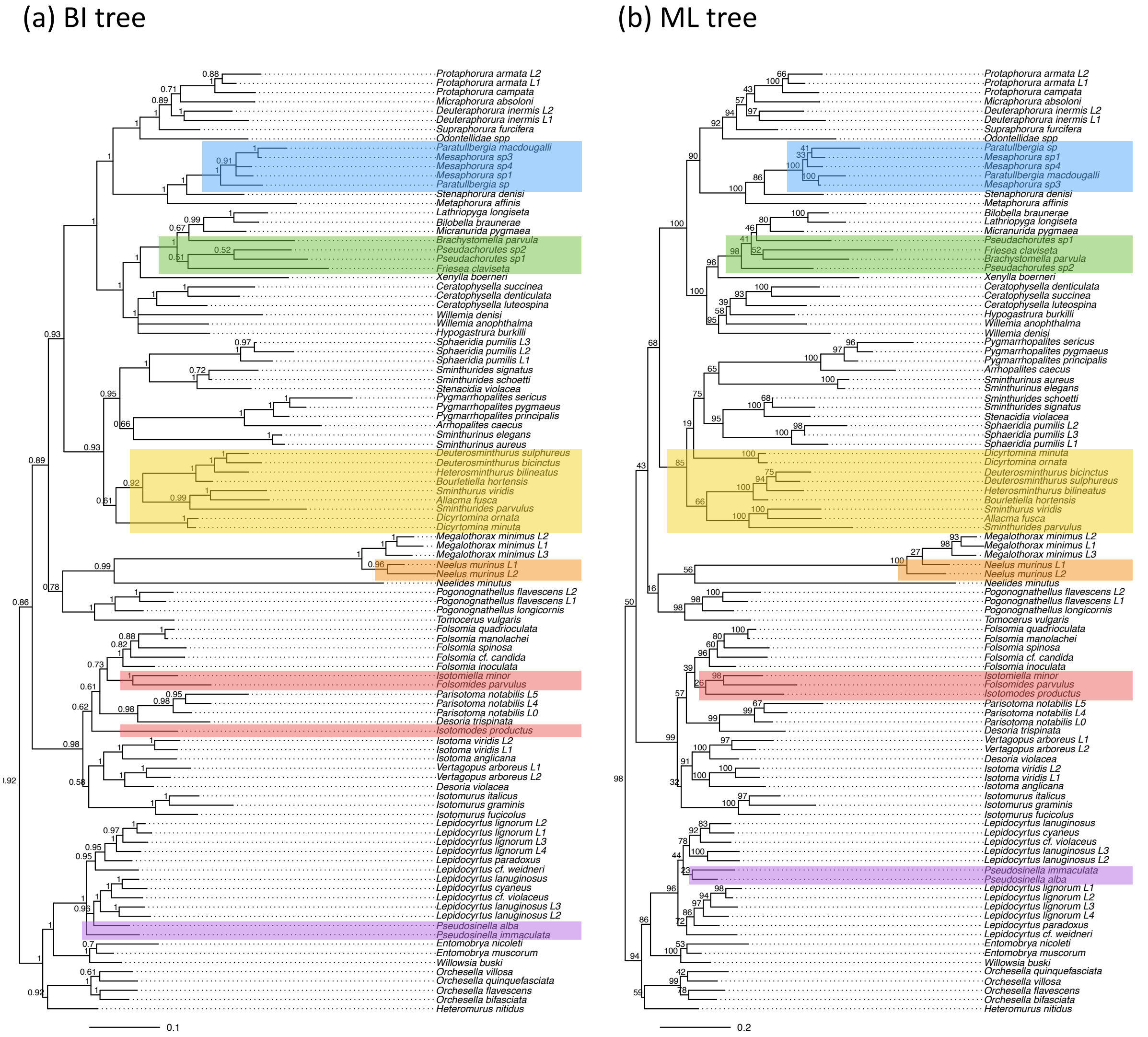

Figure S2.1
Comparison between Bayesian Inference (BI) tree and Maximum Likelihood (ML) tree based on 102 Collembola species and lineages. Coloring indicates inconsistency of phylogenetic relationships of taxa between BI and ML trees.

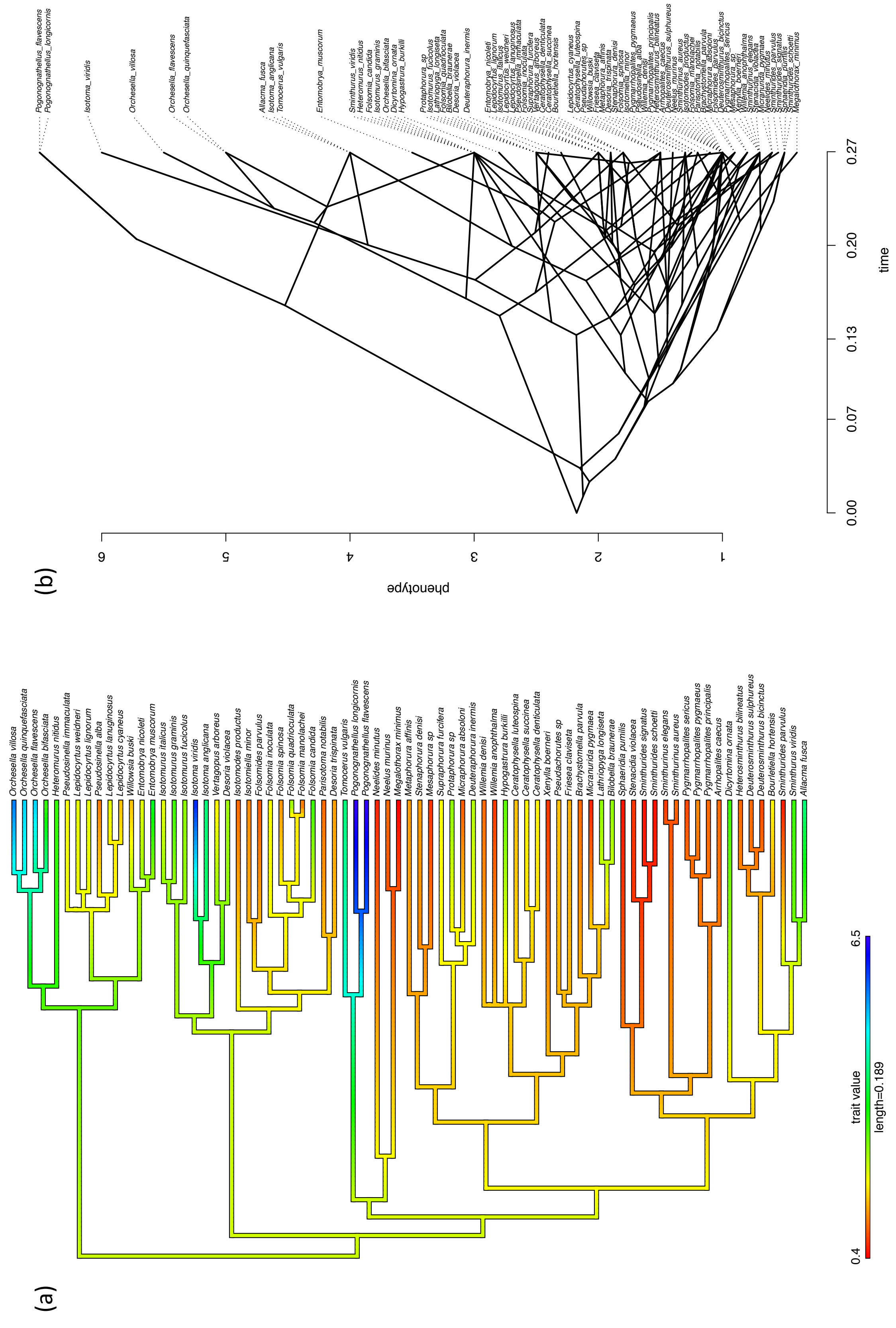

Figure S2.2
(a) Ancestral state of body length (in mm) of Collembola. (b) Traitgram shows variation of body length of Collembola during their evolution.

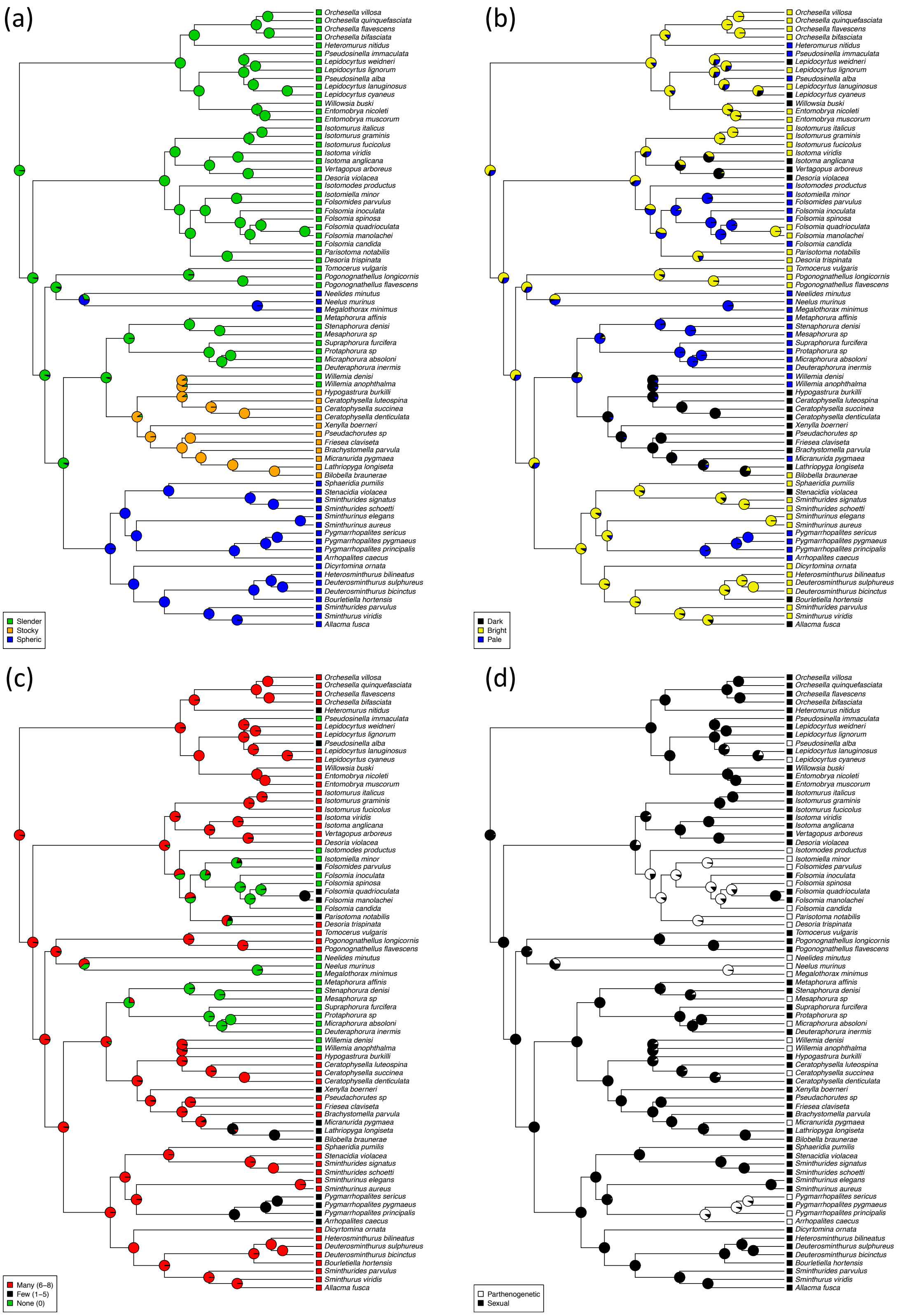

Figure S2.3

Ancestral state reconstruction of (a) body shape, (b) pigmentation, (c) number of ommatidia and (d) reproductive mode of Collembola. Colored squares at the tips of each branch represent the trait state of extant species. Pie charts on each node indicate the proportion of each character state summed across the posterior distribution of simulations.

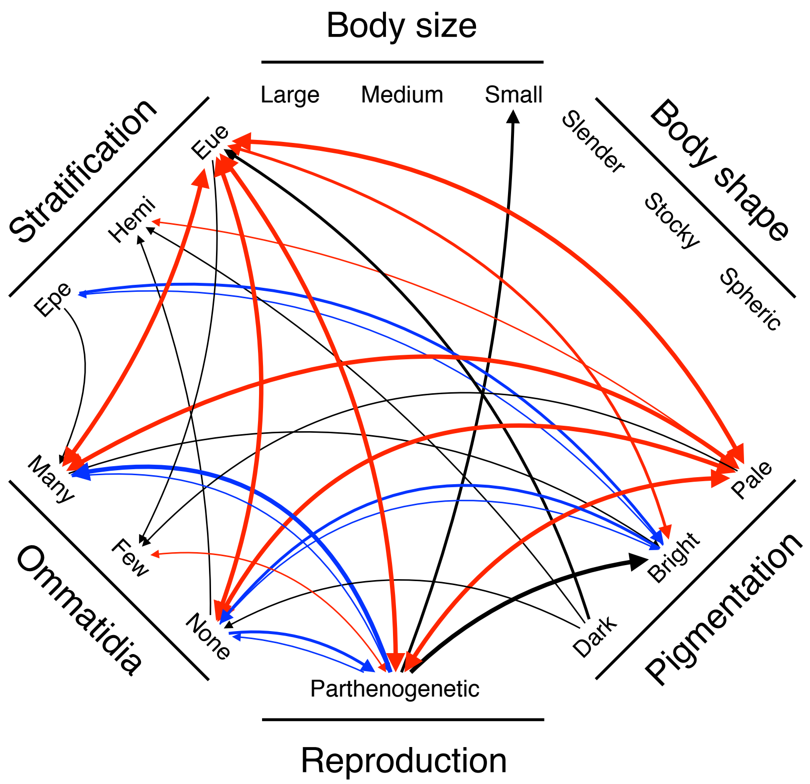

Figure S2.4
Results of Pagel's correlation test between the substitution rate of Collembola traits; significant dependence between two traits (i.e., "x" or "y" model) are indicated by arrows and lines with thickness reflecting the *P*-values. Red lines indicate that both directions ("x" to "y" and "y" to "x") are significant with the same level of *P*-values (< 0.001, 0.001–0.01, or 0.01–0.05). Blue lines indicate that "x" and "y" as dependent variables give different *P*-values. Black lines indicate that only one direction is significance.

Chapter 3

Published in *Ecology and Evolution* (2017) **7**, 9624–9638.

Neutral lipid fatty acid composition as trait and constraint in Collembola evolution

Ting-Wen Chen, Philipp Sandmann, Ina Schaefer, Stefan Scheu

Abstract

Functional traits determine the occurrence of species along environmental gradients and their coexistence with other species. Understanding how traits evolved among coexisting species helps to infer community assembly processes. We propose fatty acid composition in consumer tissue as a functional trait related to both food resources and physiological functions of species. We measured phylogenetic signal in fatty acid profiles of 13 field-sampled Collembola (springtail) species and then combined the data with published fatty acid profiles of another 24 species. Collembola fatty acid profiles generally showed phylogenetic signal, with related species resembling each other. Long-chain polyunsaturated fatty acids, related to physiological functions, demonstrated phylogenetic signal. In contrast, most food resource biomarker fatty acids and the ratios between bacterial, fungal and plant biomarker fatty acids exhibited no phylogenetic signal. Presumably, fatty acids related to physiological functions have been constrained during Collembola evolutionary history: species with close phylogenetic affinity experienced similar environments during divergence, while niche partitioning in food resources among closely related species favored species coexistence. Measuring phylogenetic signal in ecologically relevant traits of coexisting species provides an evolutionary perspective to contemporary assembly processes of ecological communities. Integrating phylogenetic comparative methods with community phylogenetic and trait-based approaches may compensate for the limitations of each method when used alone and improve understanding of processes driving and maintaining assembly patterns.

Keywords

community phylogenetics; comparative method; functional traits; phylogenetic signal; springtails; trophic niche

Introduction

Functional traits are measurable properties of species which influence their performance and fitness (Violle et al. 2007, Pey et al. 2014). They in part regulate the occurrence of species along environmental gradients and coexistence with other species in local communities (McGill et al. 2006, Ackerly and Cornwell 2007, Adler et al. 2013), where coexisting species may possess similar or different traits. Ecological traits have been assigned to two categories, α and β niche traits. While β niche traits determine species' environmental tolerance, α niche traits relate to resource exploitation (Ackerly and Cornwell 2007). Similar β niche traits but different α niche traits thus allow species to live under similar environmental conditions but utilize different resources (Silvertown et al. 2006).

Understanding evolution of traits in coexisting species helps to infer community assembly processes (Webb et al. 2002, Silvertown et al. 2006, Best and Stachowicz 2013). Species' traits may exhibit phylogenetic signal, i.e., phylogenetically related species share similar traits derived from a common ancestor (Harvey and Pagel 1991). In contrast, traits may evolve convergently, resulting in closely related species with dissimilar traits or distantly related species with similar traits (Cavender-Bares et al. 2004). However, species' traits may also be labile, i.e., varying among species irrespective of phylogenetic relationships. Further, α and β niche traits may evolve in different ways and thus exhibit different phylogenetic signal: β niche traits are usually phylogenetically conserved, while α niche traits tend to be evolutionarily labile (Silvertown et al. 2006, Ackerly et al. 2006, Best and Stachowicz 2013). In this study, we measured phylogenetic signal in a ubiquitous trait of terrestrial microarthropods, i.e., fatty acid composition.

Fatty acids (FAs) are major components of lipids, serving as a source of energy (i.e., neutral lipids) and structural components of cell membranes (i.e., phospholipids; Ruess and Chamberlain 2010). Neutral lipid fatty acids (NLFAs) in animal fat deposits carry the signal of the diet. Some NLFAs are incorporated directly and unmodified from food resources and are useful as biomarkers to distinguish between major food resources in animals living in soil (Ruess and Chamberlain 2010, Buse et al. 2013, Ferlian et al. 2015). These biomarker FAs include absolute bacterial biomarkers which are only synthesized by prokaryotes, such as a15:0, i15:0, 16:1ω5, 16:1ω7, i16:0, i17:0, cy17:0, 18:1ω7 and cy19:0, as well as relative biomarkers, such as plant biomarker 18:1ω9 and fungal biomarker 18:2ω6,9, which are found in high proportions when the consumer mainly feeds on plant or fungi, respectively. Thus, proportions of biomarker FAs imply α niche traits related to food resources. Other NLFAs, such as C20 polyunsaturated FAs 20:4ω6 and 20:5ω3, can be synthesized or modified from precursors by consumers (Chamberlain and Black 2005, Ruess and Chamberlain 2010). These FAs are essential for biosynthesis of other compounds such as prostaglandins and eicosanoids, which are associated with reproduction, immune response and temperature regulation (Chamberlain et al. 2004, Chamberlain and Black 2005, Haubert et al. 2008). They thus represent β niche traits reflecting species environmental requirements.

Springtails (Hexapoda: Collembola) are among the most abundant soil invertebrates. They occur in virtually every terrestrial habitat reaching particularly high densities in soil and contribute to decomposition processes and nutrient cycling in terrestrial ecosystems (Rusek 1998). They are ideal for exploring phylogenetic signal of FAs as they consume a wide range of food resources including detritus, roots and root exudates, bacteria, fungi and algae (Hopkin 1997). Fatty acid profiles have been used to identify food resources of Collembola and their association with different decomposition channels based on bacteria, fungi or root exudates as basal resources (Ruess et al. 2005, Pollierer et al. 2012, Ferlian et al. 2015). Distinct FA profiles of different Collembola species suggest trophic niche differentiation among co-occurring species (Chamberlain and Black 2005, Ruess et al. 2007, Ferlian et al. 2015). This may be attributed to (1) taxonomic or evolutionary relationships between different phylogenetic groups (Chamberlain and Black 2005), reflecting fixation of the physiology of species and their way of feeding over evolutionary time, and/or (2) ecological characteristics, such as life-forms (eu-, hemi-, and epedaphic) or availability of food resources in a habitat (Ruess et al. 2007). Further, species assigned to different soil strata may have similar FA profiles, indicating the use of similar resources (Ferlian et al. 2015). Overall, FA composition of Collembola may be similar in closely related species (phylogenetic signal present) and/or determined by available resources and thus not related to phylogenetic affinity (phylogenetic signal absent).

In this study we consider FA composition as a functional trait and analyze its phylogenetic signal using a comparative method (Harvey and Pagel 1991, Freckleton et al. 2002). Based on the α and β niche trait concept, we tested the following hypotheses: (1) C20 polyunsaturated FAs exhibit phylogenetic signal in Collembola, suggesting that closely related species have similar physiological attributes. (2) Food resource FA biomarkers in Collembola are phylogenetically independent as different species utilize different resources. We used two FA datasets: FA profiles measured in this study from 13 field-sampled Collembola species and our data combined with published FA profiles of another 24 species (**Table 3.1**). We constructed a phylogenetic tree for all 37 Collembola species and measured phylogenetic signal in both FA datasets using two common comparative phylogenetic metrics, Blomberg's K (Blomberg et al. 2003) and Pagel's lambda (Pagel 1999, Freckleton et al. 2002).

Materials and Methods

Sampling

Collembola were sampled from two sites near Göttingen, Germany, Deppoldshausen (51.575°N, 9.973°E) and Ossenfeld (51.548°N, 9.798°E). Each sampling site was composed of three adjacent habitats: arable field, pasture and forest. In each habitat, five samples (1 m^2, at least 5 m apart) were taken in June and July 2015. Collembola in arable fields and pastures were sampled using an aspirator then immediately brought to the laboratory at the University of Göttingen and sorted. Collembola in forests were extracted from leaf litter by heat (Kempson et al. 1963) at constant 35°C for one week. Collembola were sampled

Table 3.1

Taxonomy and collection habitat of the Collembola species used in this study. The 13 species collected for this study are marked in bold.

Phylogenetic group	Family	Species*	Habitat**	Reference
Symphypleona	**Sminthuridae**	***Allacma fusca***	**Arable field (1)**	**This study**
			Forest (3)	**This study**
			Forest	Chamberlain and Black (2005)
		Sminthurus viridis	**Arable field (1)**	**This study**
			Grassland (5)	**This study**
	Bouletiellidae	***Deuterosminthurus sulphureus***	**Arable field (1)**	**This study**
			Grassland (2)	**This study**
	Dicyrtomidae	*Dicyrtomina ornata*	Forest	Chamberlain and Black (2005)
			Forest	Ruess et al. (2005)
			Forest	Ruess et al. (2007)
		Dicyrtomina sp. (D. saundersi)	Forest	Ruess et al. (2005)
Poduromorpha	**Hypogastruridae**	***Ceratophysella denticulata***	**Forest (3)**	**This study**
			Forest	Ruess et al. (2005)
			Forest	Ruess et al. (2007)
			Forest	Ferlian et al. (2015)
		Ceratophysella succinea[1]	Grassland	Sechi et al. (2014)
		Willemia anophthalma	Arable field	Ngosong et al. (2009)
			Arable field	Ngosong et al. (2011)
	Brachystomellidae	*Brachystomella parvula*[1]	Grassland	Sechi et al. (2014)
	Neanuridae	*Neanura muscorum*	Forest	Ruess et al. (2005)
			Forest	Ruess et al. (2007)
		Polyacanthella (Friesea claviseta)	Arable field	Ngosong et al. (2009)
	Onychiuridae	*Onychiurus spp. (O. ambulans)*	Forest	Ruess et al. (2005)
		Protaphorura armata	Forest	Ferlian et al. (2015)
		Protaphorura fimata (P. sp1)	Arable field	Haubert et al. (2009)
		Protaphorura spp. (P. sp2)	Forest	Ruess et al. (2007)
Tomoceridae	**Tomoceridae**	***Pogonognathellus flavescens***	**Grassland (1)**	**This study**
			Forest (6)	**This study**
		Pogonognathellus longicornis	Forest	Chamberlain and Black (2005)
			Forest	Ruess et al. (2005)
			Forest	Ruess et al. (2007)
		Tomocerus vulgaris	**Forest (4)**	**This study**
		Tomocerus baudoti	Forest	Pollierer et al. (2012)
		Tomocerus minor	Forest	Chamberlain and Black (2005)
Isotomidae	**Isotomidae**	***Isotoma viridis***	**Arable field (4)**	**This study**
			Grassland (4)	**This study**
			Arable field	Ngosong et al. (2009)
			Arable field	Ngosong et al. (2011)
			Forest	Chamberlain and Black (2005)
		Isotoma viridis[2]	Grassland	Sechi et al. (2014)
		Isotoma anglicana[2]	Grassland	Sechi et al. (2014)
		Desoria violacea	Forest	Ruess et al. (2005)
			Forest	Ruess et al. (2007)
		Folsomia quadrioculata	Forest	Ruess et al. (2005)
			Forest	Ruess et al. (2007)
			Forest	Ferlian et al. (2015)
		Isotomiella minor	Forest	Ferlian et al. (2015)
		Isotomurus palustris (I. fucicolus)	Forest	Chamberlain and Black (2005)
		Parisotoma notabilis	Forest	Ruess et al. (2005)
			Forest	Ferlian et al. (2015)

Table 3.1 Continued

Phylogenetic group	Family	Species*	Habitat**	Reference
Entomobryoidea	**Entomobryidae**	***Entomobrya muscorum***	**Grassland (2)**	**This study**
			Forest (5)	**This study**
			Forest	Ruess et al. (2005)
		Entomobrya nicoleti	**Grassland (2)**	**This study**
		Entomobrya nivalis	Forest	Ruess et al. (2005)
	Lepidocyrtidae	***Pseudosinella immaculata***	**Grassland (1)**	**This study**
		Lepidocyrtus cyaneus	**Arable field (4)**	**This study**
			Grassland (5)	**This study**
			Grassland	Sechi et al. (2014)
		Lepidocyrtus lanuginosus	**Arable field (2)**	**This study**
			Grassland (4)	**This study**
			Forest (1)	**This study**
			Forest	Pollierer et al. (2012)
			Forest	Ferlian et al. (2015)
		Lepidocyrtus lignorum	Forest	Ruess et al. (2005)
			Forest	Ruess et al. (2007)
		Lepidocyrtus curvicollis (L. sp)	Forest	Chamberlain and Black (2005)
	Orchesellidae	***Orchesella villosa***	**Arable field (8)**	**This study**
			Forest (1)	**This study**
			Forest	Chamberlain and Black (2005)
			Arable field	Haubert et al. (2009)
		Orchesella flavescens	Forest	Ruess et al. (2005)
			Forest	Ruess et al. (2007)

* Name in parenthesis indicates the congeneric species used in the phylogeny constructed by sequences listed in Table S3.1.

** Number in parenthesis indicates replicate number in fatty acid measurements of field derived Collembola in this study (pooled for sites)

[1] Fatty acid data compiled using Poduromorpha in Sechi et al. 2014

[2] Fatty acid data compiled using *Isotoma spp.* in Sechi et al. 2014

alive daily and immediately stored at -80°C until identification and lipid extraction. Species were identified according to Hopkin (2007). In total, sufficient biomass for FA extraction was obtained for 13 species.

Fatty acid analysis

Soil and organic matter was removed from the surface of each Collembola using a brush prior to FA extraction. Depending on body size of individuals and species, three to 36 individuals of the same species and sample were pooled for one FA extraction. In total, 70 FA measurements were obtained, ranging from one to four replicates for each species per habitat and site.

NLFAs were extracted as described in Haubert et al. (2004). Neutral lipid fractions were dried at 50°C using a rotation vacuum concentrator (RVC 2-25, Chris, Osterode am Harz, Germany). The lipid fractions were then saponified, methylated and washed. The obtained FA methyl esters were transferred into vials, capped and stored at -21°C until gas chromatography (GC) analysis. The gas chromatograph (Clarus 500, Perkin Elmer, Waltham, USA) was equipped with a flame ionization detector (PE-5 capillary column, 30 m × 0.32 mm i.d., 0.25 mm film thickness, Perkin Elmer, Waltham, USA) and helium as carrier gas. The analysis program followed Ferlian and Scheu (2014). FA methyl esters were identified by comparing retention times of samples and standard mixtures comprising unbranched and branched FA methyl esters.

Collembola phylogeny

In addition to the above 13 Collembola species, published NLFA data were available for 24 additional species (**Table 3.1**). A phylogeny of all 37 Collembola species, spanning 12 families, was constructed using six genetic markers: 18S rRNA, 28S rRNA D1, D2 and D3 regions, cytochrome oxidase subunit I (COI) and Histone H3 genes. *Callibaetis* (Insecta: Ephemeroptera), *Machilis* (Insecta: Archaeognatha) and Zygentoma (Insecta: Zygentoma) were used as outgroups. Sequences were downloaded from GenBank (www.ncbi.nih.gov; Accession Number see **Table S3.1**). Species without sequence data available in GenBank were replaced by the taxonomically closest species, usually a congener (**Table 3.1**). The six genetic markers were aligned separately in R 3.2.2 (R Core Team 2015) using functions *AlignSeqs* and *AdjustAlignment* for 18S and 28S rRNA and Histone H3 genes (package "DECIPHER"; Wright 2015) and function *msaClustalW* for COI by setting gap opening as 15 and gap extension as 6.6 (package "msa"; Bodenhofer et al. 2015). The aligned sequences were trimmed to the same length in BioEdit 7.2.5 (Hall 1999). Models of sequence evolution for each marker were obtained using jModelTest 2.1.4 and based on the Alkaike information criterion (Darriba et al. 2012). Terminal gaps in each marker set were replaced by "?" and the six markers were concatenated in a supermatrix (3,053 bp) using SequenceMatrix 1.8 (Vaidya et al. 2011). The phylogeny was inferred using Bayesian Inference (BI) in MrBayes 3.2.4 (Ronquist et al. 2012), setting models of sequence evolution for each marker separately as suggested by jModelTest. Bayesian Inference was conducted using two independent runs of four chains for 1,000,000 generations and the consensus tree generated using a burn-in of 0.25. A second phylogenetic tree was constructed using Maximum Likelihood (ML) in RAxML 7.0.3 (Stamatakis 2006) based on the GTR+I+G model and 1,000 bootstrap replicates. The topologies of the phylogenetic trees of BI and ML were similar, except for the sister taxon of Poduromorpha. In the BI tree, Tomoceridae was sister of Poduromorpha (**Figure S3.1**), while in the ML tree it was Symphypleona (**Figure S3.2**). The BI tree was selected and transformed to an ultrametric tree using a penalized likelihood approach assuming different models of substitution rate variation among branches, including correlated, relaxed, discrete or strict clock models, using the function *chronos* implemented in the R package "ape" (Paradis et al. 2004). The ultrametric tree for downstream phylogenetic signal analyses was selected based on the smallest PHIIC value, a criterion analogous to Alkaike information criterion reflecting the best model fit to the data (Paradis 2013). Concomitantly, a strict clock model was used in the phylogenetic analyses. This tree was then used in the phylogenetic signal measurement.

Statistical analysis

For our field data, rare FAs present in only single measurement and FAs contributing less than 1% of total FAs were eliminated from the analyses. The remaining FAs were summed to 100% and the proportions of single FAs were logit-transformed using the function *logit* in the R package "car" (Fox and Weisberg 2011). To test for differences in FA compositions between Collembola species and habitats, multivariate analysis of variance (MANOVA) and discriminant function analysis (DFA, function *lda* implemented in the R package "MASS"; Venables and Ripley 2002) were used, with sites and habitats set as error terms in the model, followed by ANOVA with Holm's adjusted *P*-values (Holm 1979). For the FAs showing significant differences between Collembola species, Tukey's honestly significant difference (HSD) test was conducted. Fatty acid profiles of species were also explored using eigen decomposition principle components analysis (PCA). Species mean logit-transformed FA proportions were calculated and then multiplied by the eigenvectors based on a covariance matrix using the species mean. Individual observational logit-transformed FA proportions were multiplied by the same eigenvectors to examine intraspecific variation. Principle components (PCs) were selected if the variance explained by each axis was more than predicted by a broken stick model. Pearson correlation coefficients of FAs and PCs were calculated using function *cor.test* in R with Holm's *P*-value adjustment.

Three types of FA data were used to measure phylogenetic signal: (1) Species mean scores on the PC axes, irrespective of site and habitat; (2) species mean proportion of individual FAs; (3) species mean values of FA indices, including sums of bacterial FAs, plant-to-fungal FA marker ratios (P:F ratio), bacterial-to-fungal FA marker ratio (B:F ratio), bacterial-to-plant FA marker ratio (B:P ratio), Unsaturation Index (UI; Haubert et al. 2004), sums of saturated FAs (SFAs), monounsaturated FAs (MUFAs), polyunsaturated FAs (PUFAs) and C20 PUFAs and ratio of unsaturated to saturated FAs (U:S ratio). Phylogenetic signal was detected and quantified using both Blomberg's K (Blomberg et al. 2003) and Pagel's lambda (Pagel 1999, Freckleton et al. 2002). These two metrics assume a Brownian motion model of trait evolution, i.e., variance in trait values is directly proportional to branch length of a given phylogeny (Pagel 1999, Blomberg et al. 2003). Both methods were used because they have different sensitivities in detecting phylogenetic signal for traits evolved with various strengths of Brownian motion and for trees with different size (Münkemüller et al. 2012). Phylogenetic signal analyses were conducted using the function *phylosig* implemented in the R package "phytools" (Revell 2012). Standard errors of FA measurements were considered in Blomberg's K statistics (Ives et al. 2007). Significance tests were done by randomizing species on the phylogeny 10,000 times, to test whether trait values show phylogenetic signal or not (i.e., H0 = 0). In case of significant K-values of traits, the observed K-value was further compared with 5,000 simulated K-values to test whether phylogenetic signal was significantly different from the level expected under Brownian motion evolution model (i.e., H0 = 1; Revell et al. 2007). Simulations of trait values were conducted using the function *fastBM* in the R package "phytools" (Revell 2012). Lower and higher phylogenetic signal than predicted by a Brownian motion model was defined as a K-value in the 0.025 and 0.975 quantiles of the log-transformed simulated K-values, respectively. All *P*-

values in phylogenetic signal measurement were adjusted using Benjamini and Hochberg's method (Benjamini and Hochberg 1995). Phylogenetic signal of FAs was accepted only when both Blomberg's K and Pagel's lambda were significant.

Since a small phylogenetic tree (13 species in our field-sampled dataset) may lack power to detect phylogenetic signal (Freckleton et al. 2002, Blomberg et al. 2003, Münkemüller et al. 2012), FA phylogenetic signal was also measured using a combined dataset comprising data of our field-sampled Collembola and published FA data (**Table 3.1**). Mean FA proportions were calculated for each species at each site and habitat for our FA data. Data from the literature were compiled at species level for each treatment or site by extracting the published mean values or recalculating original data provided by the authors. Due to inconsistency of FAs measured in different studies, only biomarker FAs, C20 unsaturated FAs and saturated FAs 16:0 and 18:0 were included. Unavailable values of these FAs in literature data were replaced by zero assuming that they were not reported due to being present in trace amounts only. Fatty acids contributing less than 1% of total FAs and those occurring in only one sample were eliminated. The remaining FAs were summed to 100% and logit-transformed, resulting in a final dataset of 37 species and 149 data points for phylogenetic signal measurements. Principle components and phylogenetic signal in species mean scores on PCA axes, mean proportion of individual FAs and FA indices were analyzed as above.

Results

Fatty acid composition of Collembola

Thirty-two FAs were identified from the 13 field-sampled Collembola species (**Table S3.2**). Frequent FAs (occurring in > 30 of the 70 measurements) were 18:1ω9, 18:2ω6,9, 16:0, 18:0, 20:5ω3, 20:4ω6, 16:1ω7, 14:0 and 18:1ω7. Overall, the lipid composition of Collembola predominantly differed between species, whereas the effect of habitat was not significant (MANOVA, $F_{384,288} = 1.65$, $P < 0.001$ for species and $F_{256,160} = 1.26$, $P = 0.058$ for habitat). The DFA plot clearly separated the FA profiles between different species (**Figure 3.1**). *Allacma fusca*, *Deuterosminthurus sulphureus*, *Sminthurus viridis*, *Ceratophysella denticulata* and *Isotoma viridis* were separated from the remaining species along the first two axes. The proportions of individual FAs differed among species (**Table S3.2**, **S3.3**). Fatty acid 18:1ω9, a predominant FA in Collembola, was lower in *I. viridis* (12.4%), while it contributed 26.5%–42.2% to total FAs in all other species. Another major FA, 18:2ω6,9, was highest in the three Symphypleona species, *S. viridis* (37.6%), *A. fusca* (34.9%) and *D. sulphureus* (32.3%). Fatty acid 16:0 was low in *A. fusca* (9.6%), but high in all Entomobryoidea (23.3%–26.4%), except for *Orchesella villosa* (17.6%). Fatty acid 18:0 was present in trace proportions in *D. sulphureus* (0.9%), but was one of the main FAs in *I. viridis* (15.8%). C20 PUFAs 20:4ω6 and 20:5ω3 were not detected in any of the three Symphypleona species, while 20:5ω3 was high in the Tomoceridae, *Pogonognathellus flavescens* (8.1%) and *Tomocerus vulgaris* (6.1%). Fatty acid 16:1ω7 was highest in *C. denticulata* (8.3%), while 18:1ω7 was highest in the two tomocerids, *T. vulgaris*

(6.7%) and *P. flavescens* (5.7%). *Pseudosinella immaculata* had a relatively high proportions of FA 14:0 (11.4%).

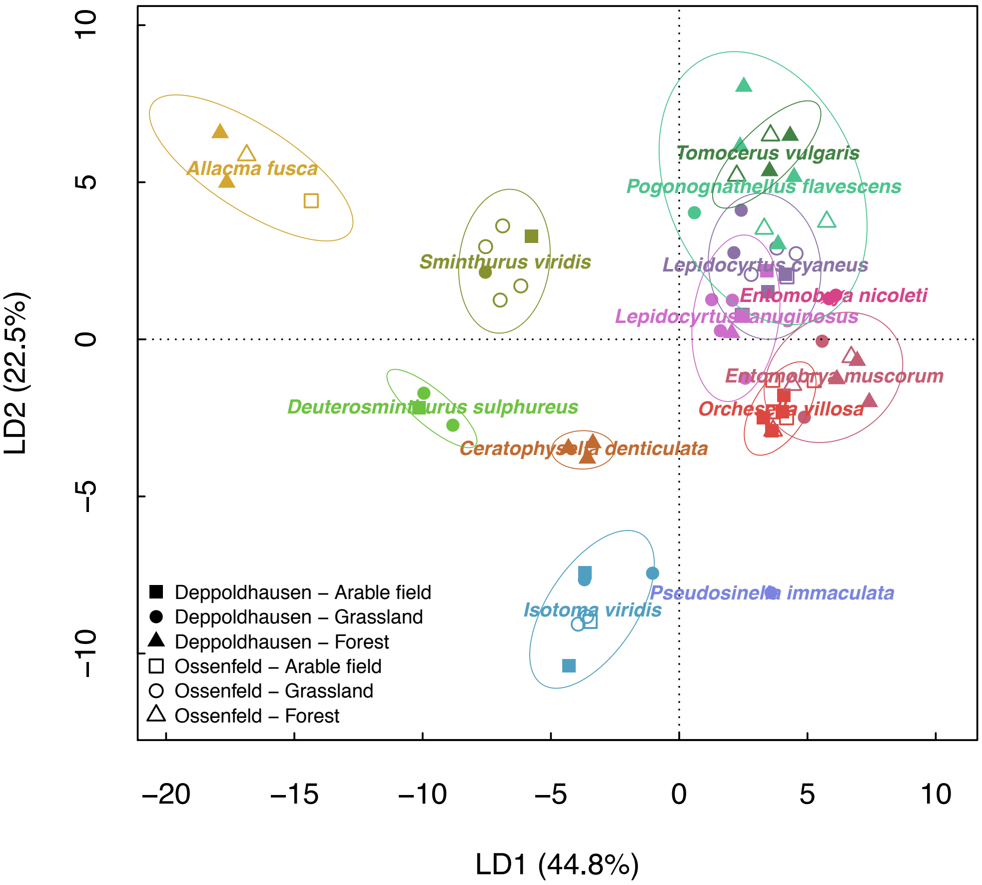

Figure 3.1
Discriminant function analysis of fatty acid profiles of 13 field-sampled Collembola species. Ellipses represent confidence ranges at $P = 0.05$.

Phylogenetic signal in FAs of sampled species

The first four PCs together explained 84.2% of the variation in the FA profiles of the Collembola. PC1, representing 44.9% of the variation, showed phylogenetic signal consistent with predictions from the Brownian motion model, as indicated by both Blomberg's K and Pagel's lambda (**Table 3.2**). The PCA biplots indicated that three Symphypleona, *A. fusca*, *D. sulphureus* and *S. viridis*, had higher scores along PC1, which was negatively correlated with FA 18:0, 20:5ω3, 20:4ω6 and 16:0, and positively correlated with FA 18:2ω6,9 (**Table S3.4**, **Figure 3.2**, **3.3a**). The remaining PCs, however, exhibited no phylogenetic signal, except PC3 using Blomberg's K without *P*-value adjustment.

Proportions of the FAs 16:0, 18:1ω7, 18:2ω6,9, 20:1ω9 and 20:5ω3 showed significant phylogenetic signal as indicated by Blomberg's K after *P*-value adjustment. Pagel's lambda further indicated that the FAs 2-OH 10:0, 12:0, 14:1, 15:0, 16:1ω7, i16:0, 18:0, 20:2ω6,9, 20:3ω6, 20:4ω6 and 22:2 also showed phylogenetic signal after P-value adjustment (**Table 3.2**). Phylogenetic signal in FA 16:0 resulted from higher proportions in the clade composed of Lepidocyrtidae and Entomobryidae and lower proportions in *C. denticulata* and Symphypleona. Phylogenetic signal in FA 20:1ω9 resulted from the lack in the clades of Lepidocyrtidae (*Lepidocyrtus* and *Pseudosinella*), Entomobryidae (two *Entomobrya* species) and Sminthuridae (*Allacma* and *Sminthurus*). Notably, the K-value of 20:1ω9 was larger than the 97.5% quantile of simulated K-values, suggesting stronger phylogenetic signal than predicted by the Brownian motion model. Fatty acid 20:5ω3 showed phylogenetic signal due to its consistently lower proportions in *C. denticulata* and Symphypleona, intermediate proportions in Entomobryoidea, higher proportions in Tomoceridae, and even higher proportions in *I. viridis*. Phylogenetic signal in the bacterial biomarker 18:1ω7 reflected higher relative proportion in Tomoceridae and lower in Symphypleona and Entomobryoidea (**Table 3.2**, **Figure 3.3b**). The fungal biomarker 18:2ω6,9 showed phylogenetic signal, reflecting higher proportions in Symphypleona as well as lower proportions in most Entomobryoidea (**Table 3.2**, **Figure 3.3c**).

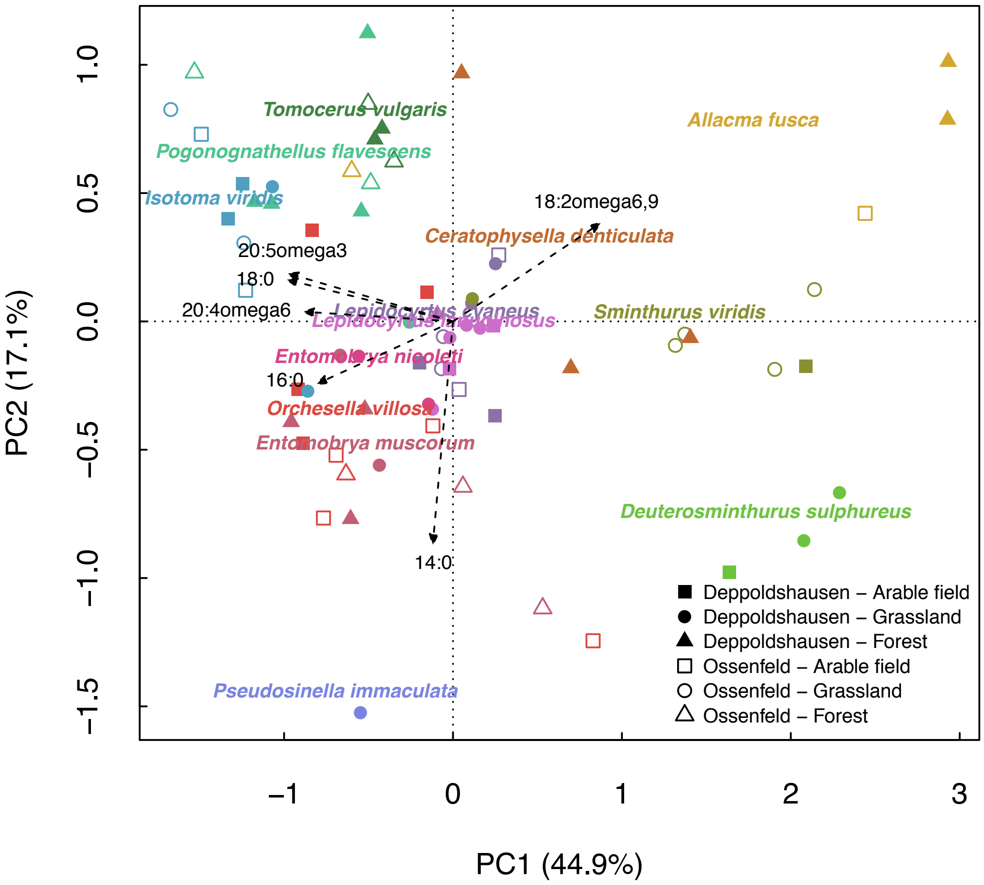

Figure 3.2
Biplots of principle components analysis using whole fatty acid profiles of 13 field-sampled Collembola species. Variation explained by each axis is given in parentheses. Position of species name represents its mean score on the axis irrespective of site and habitat. Only fatty acids significantly correlated to the PCs are plotted.

Table 3.2

Phylogenetic signal in 13 field-collected Collembola using whole fatty acid profiles. Phylogenetic signal was measured for PCA axes, individual fatty acid proportions and FA indices, and is reported as Blomberg's K combined with permutation significance test (H0 = 0) and simulation test (H0 = 1) and as Pagel's lambda with maximum log likelihood test. *P*-values based on permutation testing were corrected using Benjamini & Hochberg' (BH) method. A significant K-value ($p < 0.05$) within the 2.5% and 97.5% quantiles of simulated K indicates trait evolution as expected under a Brownian motion model. Maximum log likelihood of a trait fit to the given phylogeny (logL) was tested against the fit to a lambda transformed phylogeny (logL0, lambda = 0). A significant *P*-value in Pagel's lambda test indicates phylogenetic signal in that trait. Individual fatty acid differed in proportions between species as indicated by ANOVA (Table S3.3) is marked in italic. Significant phylogenetic signal of the trait detected by at least one method is marked in bold and by both methods with an asterisk.

	Blomberg's K					**Pagel's lambda**				
	Observed K	**Permutated *P* (H0 = 0)**		**Simulated *P* (H0 = 1)**	**Simulated K (2.5% - 97.5%)**	**lambda**	**logL**	**logL0**	***P***	
		Unadjusted	**BH**						**Unadjusted**	**BH**
PCA axis (explained variation)										
***PC1 (44.9%)**	**1.593**	**0.001**	**0.003**	0.121	0.550 - 1.829	**1.278**	**-10.77**	**-18.62**	**< 0.001**	**< 0.001**
PC2 (17.0%)	0.694	0.311	0.311			0.262	-12.10	-12.30	0.519	0.693
PC3 (13.1%)	**0.993**	**0.037**	0.073	0.842	0.543 - 1.817	1.075	-9.60	-10.61	0.155	0.309
PC4 (9.1%)	1.059	0.121	0.161			0.000	-8.25	-8.25	1.000	1.000
Individual fatty acid										
8:0	1.696	0.794	0.794			0.000	36.36	36.36	1.000	1.000
10:0	1.854	0.122	0.250			**1.264**	**49.65**	**47.31**	**0.031**	0.058
2-OH 10:0	7.778	0.075	0.199			**1.278**	**78.54**	**69.63**	**< 0.001**	**< 0.001**
12:0	4.754	0.118	0.250			**1.278**	**79.56**	**72.41**	**< 0.001**	**0.001**
14:0	0.500	0.679	0.721			0.000	27.15	27.15	1.000	1.000
14:1	7.580	0.074	0.199			**1.278**	**78.21**	**69.30**	**< 0.001**	**< 0.001**
15:0	6.775	0.160	0.301			**1.278**	**63.03**	**54.89**	**< 0.001**	**< 0.001**
a15:0 (biomarker)	1.476	0.184	0.326			0.000	53.48	53.48	1.000	1.000
i15:0 (biomarker)	1.431	0.546	0.624			0.000	40.02	40.02	1.000	1.000
****16:0***	**1.393**	**0.001**	**0.029**	0.241	0.537 - 1.820	**0.973**	**22.90**	**19.31**	**0.007**	**0.016**
i16:0 (biomarker)	2.950	0.125	0.250			**1.278**	**76.28**	**71.99**	**0.003**	**0.009**
16:1ω5	4.840	0.397	0.508			0.000	64.51	64.51	1.000	1.000
16:1ω7 (biomarker)	1.564	0.416	0.512			**1.276**	**36.78**	**31.14**	**0.001**	**0.003**
17:0	2.315	0.698	0.721			0.000	60.91	60.91	1.000	1.000
cy17:0 (biomarker)	3.338	0.277	0.403			0.000	55.60	55.60	1.000	1.000
i17:0 (biomarker)	2.039	0.260	0.396			0.000	33.87	33.87	1.000	1.000
17:1ω8	2.820	0.250	0.396			0.000	53.26	53.26	1.000	1.000
18:0	**0.801**	**0.019**	0.070	0.592	0.545 - 1.833	**1.273**	**28.55**	**24.78**	**0.006**	**0.014**
****18:1ω7*** (biomarker)	**1.482**	**0.005**	**0.044**	0.170	0.532 - 1.829	**1.244**	**36.51**	**31.79**	**0.002**	**0.007**

Table 3.2 Continued

	Blomberg's K					Pagel's lambda				
	Observed K	Permutated *P* (H0 = 0)		Simulated *P* (H0 = 1)	Simulated K (2.5% - 97.5%)	lambda	logL	logL0	*P*	
		Unadjusted	BH						Unadjusted	BH
18:1ω9 (relative biomarker)	**0.920**	**0.020**	0.070	0.968	0.525 - 1.826	1.136	15.54	14.65	0.182	0.307
****18:2ω6,9*** (relative biomarker)	**1.336**	**0.006**	**0.044**	0.279	0.538 - 1.837	**1.165**	**16.04**	**11.88**	**0.004**	**0.010**
19:0	2.126	0.632	0.698			0.056	55.40	55.39	0.909	1.000
cy19:0 (biomarker)	1.521	0.368	0.490			0.000	48.56	48.56	1.000	1.000
***20:1ω9**	**3.364**	**0.007**	**0.044**	**< 0.001**	**0.547 - 1.854**	**1.278**	**60.43**	**51.84**	**< 0.001**	**< 0.001**
20:2ω6,9	**2.124**	**0.017**	0.070	0.026	0.547 - 1.799	**1.278**	**69.80**	**60.18**	**< 0.001**	**< 0.001**
20:3ω6	3.530	0.082	0.202			**1.278**	**68.24**	**60.19**	**< 0.001**	**< 0.001**
20:4ω6	**1.225**	**0.014**	0.070	0.406	0.538 - 1.847	**1.116**	**36.73**	**34.01**	**0.020**	**0.039**
****20:5ω3***	**1.593**	**0.003**	**0.044**	0.109	0.541 - 1.798	**1.229**	**30.54**	**26.02**	**0.003**	**0.008**
22:1ω9	4.097	0.201	0.339			**1.171**	**57.28**	**55.06**	**0.035**	0.062
22:2	**3.348**	**0.032**	0.101	< 0.001	0.540 - 1.852	**1.278**	**65.05**	**56.89**	**< 0.001**	**< 0.001**
23:0	4.459	0.343	0.477			0.000	43.74	43.74	1.000	1.000
24:1	0.898	0.460	0.545			0.000	52.84	52.84	1.000	1.000
FA indices										
FA number	0.975	0.110	0.135			1.127	-26.71	-27.62	0.177	0.277
Sums of bacterial FAs	0.934	0.068	0.107			0.498	25.77	25.17	0.273	0.334
Plant-to-fungal FA ratio (P:F ratio)	0.855	0.299	0.299			0.368	-23.84	-24.47	0.264	0.334
Bacterial-to-fungal FA ratio (B:F ratio)	1.758	0.097	0.133			0.000	-7.22	-7.22	1.000	1.000
Bacterial-to-plant FA ratio (B:P ratio)	2.471	0.065	0.107			1.050	-5.99	-7.49	0.083	0.183
Unsaturation Index (UI)	0.659	0.068	0.107			**1.100**	**71.46**	**68.17**	**0.010**	**0.038**
Sums of saturated FAs (SFAs)	0.788	0.053	0.107			1.110	16.16	14.85	0.105	0.193
Sums of monounsaturated FAs (MUFAs)	0.933	0.201	0.221			1.253	17.18	16.73	0.341	0.375
***Sums of polyunsaturated FAs (PUFAs)**	**1.128**	**0.012**	**0.044**	0.563	0.535 - 1.763	**0.937**	**19.82**	**16.25**	**0.007**	**0.038**
***Sums of C20 PUFAs**	**1.687**	**0.001**	**0.009**	0.088	0.537 - 1.874	**1.278**	**29.43**	**20.61**	**< 0.001**	**< 0.001**
Unsaturated-to-saturated FA ratio (U:S ratio)	**1.373**	**0.012**	**0.044**	0.245	0.539 - 1.863	**0.898**	**-19.99**	**-22.24**	**0.034**	0.093

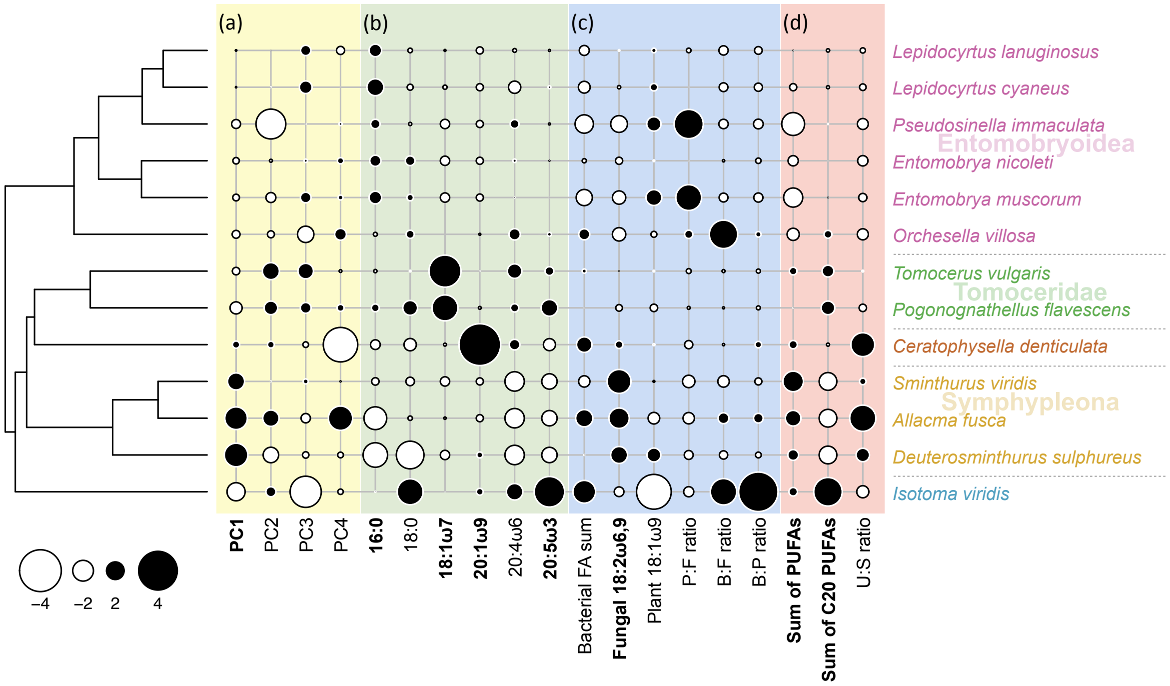

Figure 3.3
Relationship between phylogeny and selected trait values for field-sampled Collembola. Trait values were scaled and centralized before plotting. The size of the white and black circles indicates more negative or positive values, respectively. (a) Mean values of the scores of fatty acid profiles on the first four axes in principle components analysis (PCA), (b) proportions of individual fatty acids, (c) proportions of fatty acids derived from bacteria, fungi or plants, and the ratios between these three, and (d) summed proportions of polyunsaturated fatty acid and C20 polyunsaturated fatty acid, and ratio of unsaturated to saturated FAs; see Table 3.2 for abbreviations. Traits exhibiting phylogenetic signal as indicated by both Blomberg's K and Pagel's lambda are marked in bold.

The sum of C20 PUFAs and of all PUFAs exhibited phylogenetic signal according to both Blomberg's K and Pagel's lambda after *P*-value adjustment. The sum of C20 PUFAs was low in Symphypleona but high in *I. viridis* and the two Tomoceridae species (**Table 3.2**, **Figure 3.3d**). The sum of all PUFAs, however, was high in Symphypleona, *C. denticulata* and *I. viridis* but low in Entomobryoidea. The other FA indices, such as ratios between bacterial, fungal and plant biomarker fatty acids, showed no phylogenetic signal.

Phylogenetic signal in FAs of combined dataset

The first four PCs explained 76.4% of variation in the FA profiles of the 37 species of the combined dataset. PC1 explained 31.8% of the variation in the FA profiles which was positively correlated with FA 18:2ω6,9 and 18:1ω9 and negatively with 18:1ω7, 20:5ω3, 20:4ω6 and 18:0 (**Table S3.4**, **Figure 3.4**). Phylogenetic signal in species mean scores at PC1 were driven by low scores in Tomoceridae and high scores in the clade of *A. fusca*, *D. sulphureus* and *S. viridis* (**Table 3.3**, **Figure 3.5a**). No phylogenetic signal was detected in species mean scores at the other three PCs.

Analyses of phylogenetic signal in the mean proportions of individual FAs (**Table S3.5**) suggested that FA 18:0 and 20:5ω3 exhibited phylogenetic signal as indicated by both Blomberg's K and Pagel's lambda (**Table 3.3**). Fatty acid 18:0 was high in the clade composed of *Isotomurus palustris*, *Parisotoma notabilis*, *Isotomiella minor* and *Folsomia quadrioculata*, while FA 20:5ω3 was mainly present in Tomoceridae but absent in Poduromorpha (**Figure 3.5b**). The fungal biomarker 18:2ω6,9 exhibited phylogenetic signal only by Pagel's lambda; however, plant biomarker 18:1ω9, the sum of bacterial FAs and the ratios between fungal, plant and bacterial FAs showed no phylogenetic signal (**Table 3.3**, **Figure 3.5c**). The sum of C20 PUFAs showed phylogenetic signal as indicated by both Blomberg's K and Pagel's lambda (**Table 3.3**). It was high in *Tomocerus*, intermediate in Entomobryoidea, and low in the clade of *I. minor*, *F. quadrioculata* and *P. notabilis*, as well as the clade of *S. viridis*, *A. fusca* and *D. sulphureus* (**Figure 3.5d**).

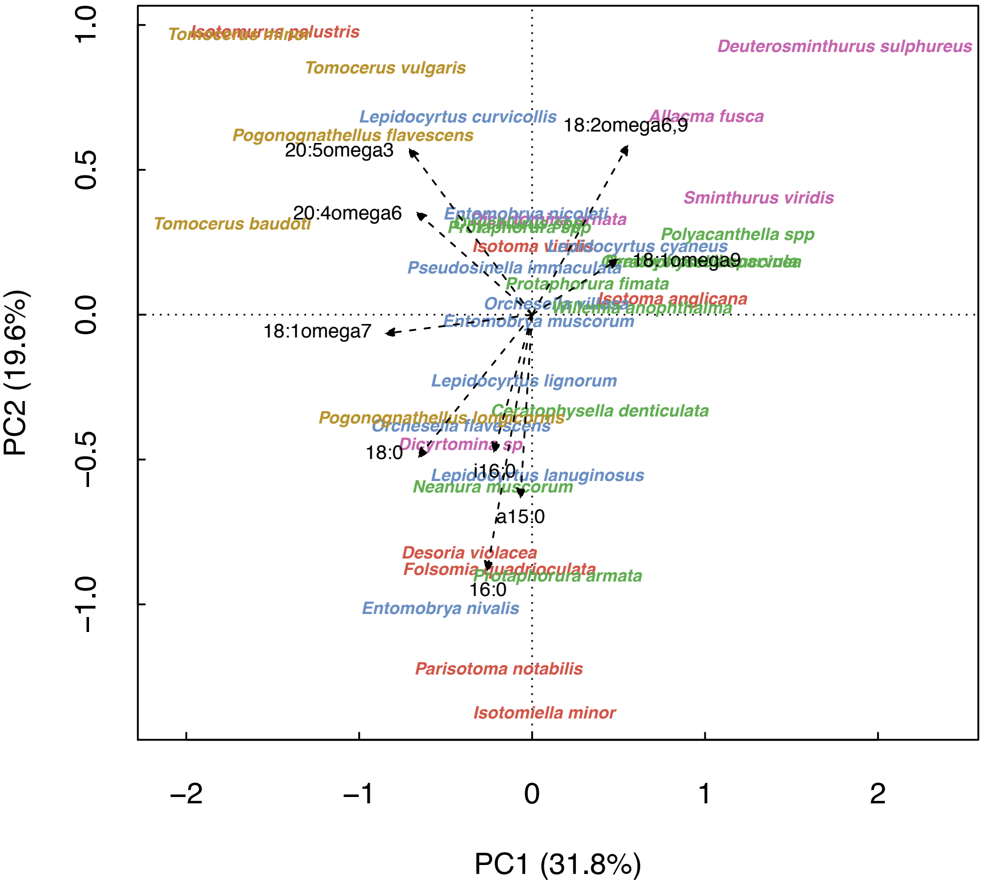

Figure 3.4
Principle components biplots of the variation in fatty acid profiles of the combined dataset. Variation explained by each axis is given in parentheses. Position of species name represents its mean score on the axis irrespective of reference, site, habitat and treatment; only fatty acids significantly correlated to the PCs are plotted.

Table 3.3

Phylogenetic signal in the expanded 37-species set of Collembola fatty acid profiles. Phylogenetic signal was measured for PCA axes, individual fatty acid proportions and FA indices, and is reported as Blomberg's K combined with permutation significance test (H0 = 0) and simulation test (H0 = 1) and as Pagel's lambda with maximum log likelihood test. *P*-values based on permutation testing were corrected using Benjamini & Hochberg' (BH) method. A significant K-value ($p < 0.05$) within the 2.5% and 97.5% quantiles of simulated K indicates trait evolution as expected under a Brownian motion model. Maximum log likelihood of a trait fit to the given phylogeny (logL) was tested against the fit to a lambda transformed phylogeny (logL0, lambda = 0). A significant *P*-value in Pagel's lambda test indicates phylogenetic signal in that trait. Individual fatty acid differed in proportions between species as indicated by ANOVA (Table S3.5) is marked in italic. Significant phylogenetic signal of the trait detected by at least one method is marked in bold and by both methods with an asterisk.

	Blomberg's K					**Pagel's lambda**				
	Observed K	**Permutated *P* (H0 = 0)**		**Simulated *P* (H0 = 1)**	**Simulated K (2.5% - 97.5%)**	**lambda**	**logL**	**logL0**	***P***	
		Unadjusted	**BH**						**Unadjusted**	**BH**
PCA axis (explained variation)										
***PC1 (31.4%)**	**1.036**	**< 0.001**	**< 0.001**	0.785	0.598 - 1.698	**0.983**	**-34.91**	**-43.14**	**0.000**	**< 0.001**
PC2 (19.4%)	0.513	0.330	0.440			0.249	-33.88	-34.15	0.461	0.615
PC3 (13.3%)	0.541	0.183	0.366			0.013	-27.11	-27.11	0.963	0.963
PC4 (11.6%)	0.477	0.444	0.444			0.166	-24.04	-24.51	0.332	0.615
Individual fatty acid										
a15:0 (biomarker)	0.367	0.903	0.903			0.000	142.10	142.10	1.000	1.000
i15:0 (biomarker)	0.973	0.379	0.587			0.023	119.77	119.77	0.964	1.000
16:0	0.533	0.293	0.587			0.189	45.67	45.45	0.515	0.735
i16:0 (biomarker)	0.433	0.719	0.885			0.000	147.66	147.66	1.000	1.000
16:1ω7 (biomarker)	0.551	0.329	0.587			0.356	90.66	89.16	0.084	0.209
i17:0 (biomarker)	0.859	0.854	0.903			0.000	134.92	134.92	1.000	1.000
18:0	**0.742**	**0.002**	**0.022**	0.338	0.591 - 1.664	**1.086**	**75.62**	**71.19**	**0.003**	**0.015**
18:1ω7 (biomarker)	**0.661**	**0.025**	0.125	0.209	0.605 - 1.686	**0.638**	**78.38**	**75.82**	**0.024**	0.095
18:1ω9 (relative biomarker)	0.480	0.432	0.587			0.000	46.33	46.33	1.000	1.000
18:2ω6,9 (relative biomarker)	0.683	0.130	0.417			**0.619**	**43.22**	**40.00**	**0.011**	**0.037**
cy19:0 (biomarker)	1.180	0.373	0.587			0.000	120.35	120.35	1.000	1.000
20:1ω9	0.608	0.440	0.587			0.219	104.49	104.26	0.498	0.735
20:2ω6,9	3.473	0.328	0.587			0.026	171.24	171.23	0.870	1.000
20:3ω6	2.916	0.840	0.903			0.044	107.74	107.73	0.912	1.000
20:4ω6	**0.644**	**0.031**	0.125	0.104	0.599 - 1.711	0.492	89.31	89.03	0.454	0.735
20:5ω3	**0.853**	**0.003**	**0.022**	0.641	0.601 - 1.678	**0.820**	**78.16**	**73.26**	**0.002**	**0.015**
FA indices										
Sum of bacterial FAs	0.524	0.296	0.453			0.213	70.39	69.60	0.208	0.498
Plant-to-fungal FA ratio (P:F ratio)	0.584	0.370	0.453			0.187	-60.64	-61.02	0.381	0.544
Bacterial-to-fungal FA ratio (B:F ratio)	0.620	0.228	0.453			0.370	-25.67	-26.61	0.169	0.498
Bacterial-to-plant FA ratio (B:P ratio)	0.518	0.452	0.453			0.102	11.25	11.02	0.498	0.553
Unsaturation Index (UI)	0.565	0.090	0.448			0.358	170.89	170.43	0.337	0.544
Sum of saturated FAs (SFAs)	0.569	0.183	0.453			0.270	33.43	33.16	0.468	0.553
Sum of monounsaturated FAs (MUFAs)	0.486	0.453	0.453			0.000	46.46	46.46	1.000	1.000
Sum of polyunsaturated FAs (PUFAs)	0.529	0.340	0.453			0.273	40.21	39.55	0.249	0.498
Sum of C20 PUFAs*	**0.811**	**0.003**	**0.034**	0.525	0.603 - 1.642	**0.803**	**76.98**	**72.96**	**0.005**	**0.046**
Unsaturated-to-saturated FA ratio (U:S ratio)	1.166	0.148	0.453			**0.698**	**-60.38**	**-63.37**	**0.014**	0.072

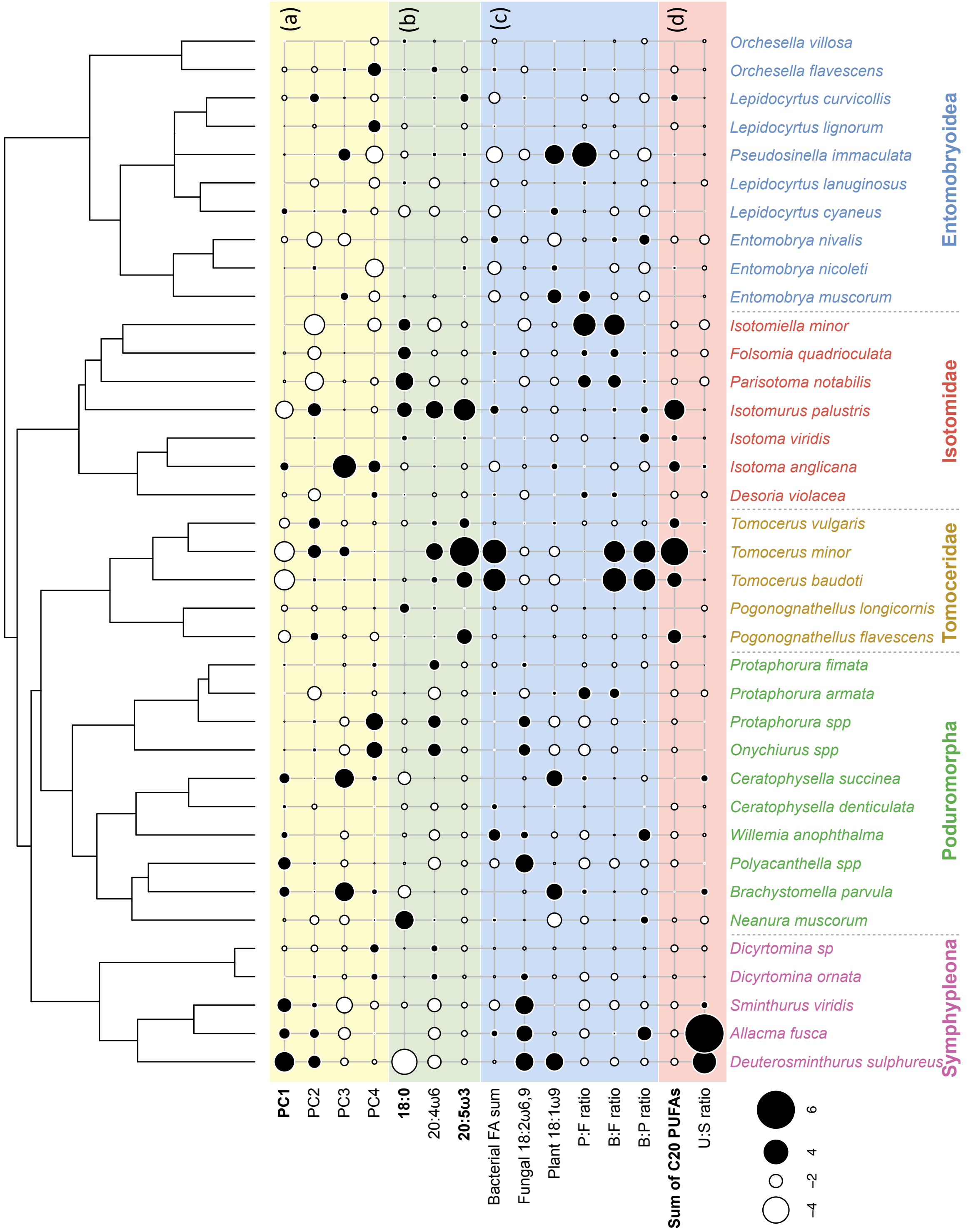

Figure 3.5
Relationship between phylogeny and selected trait values of Collembola using the combined dataset. Trait values were scaled and centralized before plotting. The size of the white and black circles indicate more negative or positive values, respectively. (a) Mean values of the scores of fatty acid profiles on the first four axes in principle components analysis (PCA), (b) proportions of fatty acids, (c) proportions of fatty acids derived from bacteria, fungi or plants, as well as the ratios between these three, and (d) summed proportions of C20 polyunsaturated fatty acid and ratio of unsaturated to saturated FAs; see Table 3.3 for abbreviations. Traits exhibiting phylogenetic signal as indicated by both Blomberg's K and Pagel's lambda are marked in bold.

Discussion

Studies of FAs in Collembola usually have used one or two species in laboratory cultures under different conditions (Chamberlain et al. 2005, Haubert et al. 2008, van Dooremalen and Ellers 2010) or analyzed FAs of field-sampled species but with limited numbers of species sampled from one habitat type, i.e., forest (Chamberlain and Black 2005, Ruess et al. 2007, Ferlian et al. 2015) or arable fields (Haubert et al. 2009, Ngosong et al. 2009, Sechi et al. 2014). This study is the first to measure phylogenetic signal in FA compositions of field-sampled Collembola from different habitats using a phylogenetic comparative method. Our results suggest that although habitat effects on FA profiles were minor, FA compositions differed significantly between species and generally displayed phylogenetic signal, as indicated by the first PC axis for both field-sampled and combined datasets.

Fatty acids, animal physiology and phylogenetics (β niche traits)

Phylogenetic signal was detected in C20 PUFAs and proportions of 20:5ω3 in both our field-sampled and combined datasets, supporting the first hypothesis that closely related Collembola species have similar proportions of C20 PUFAs. In field-sampled Collembola, Symphypleona contained lower proportions of C20 PUFAs than Entomobryomorpha, consistent with previous findings (Chamberlain and Black 2005). Collembola may have the ability to synthesize C20 PUFAs from precursors, as indicated by laboratory experiments in which a high proportion of C20 PUFAs was found in Isotomidae and Onychiuridae fed with food containing no PUFAs (Chamberlain and Black 2005). In insects, C20 PUFAs are essential for biosynthesis of prostaglandins and eicosanoids, which are important for reproduction and immune response, and related to temperature and humidity of the habitat (Stanley-Samuelson et al. 1992, Stanley-Samuelson 1994). Accordingly, the phylogenetic signal of C20 PUFA in different Collembola lineages presumably reflects an evolutionary constraint of physiological functions related to these FAs. Symphypleona predominantly live at the soil surface where humidity fluctuates with some dry periods, while the other taxa, such as Isotomidae, Tomoceridae and Poduromorpha, predominantly dwell in soil where humidity is high and relatively stable. Physiological constraints on the proportions of C20 PUFAs within phylogenetic lineages likely reflect the different soil horizons the species live in. However, the linkage between C20 PUFAs and the adaptation of species to different soil layers requires further examination of the functions of C20 PUFAs in Collembola.

Fatty acids, food resources and phylogenetics (α niche traits)

Among biomarker FAs, only three markers (18:1ω7, 18:2ω6,9 and 20:1ω9) exhibited phylogenetic signal in the field-sampled dataset, while the combined dataset showed phylogenetic signal in PC1 that correlated with 18:1ω7 and two other biomarker FAs (18:1ω9 and 18:2ω6,9). Fatty acid 18:1ω7 is an absolute bacterial biomarker synthesized exclusively by bacteria (Ruess and Chamberlain 2010, Ferlian et al. 2015). High proportions of 18:1ω7 in Tomoceridae of our field-sampled dataset indicate that they fed heavily on bacteria at the study sites, whereas Entomobryoidea and Symphypleona consumed less of food resources containing this FA. Presumably, feeding on bacteria has been restricted to certain Collembola phylogenetic groups during evolutionary history, but this hypothesis needs further testing.

Phylogenetic signal was detected in the proportion of 18:2ω6,9 in the field-sampled dataset. Fatty acid 18:2ω6,9 was higher in Symphypleona, consistent with findings of Chamberlain and Black (2005) where two Symphypleona species also had higher proportions of it than the other species sampled from a deciduous woodland. High proportions of 18:2ω6,9 are found in body tissue under a fungus-based diet and thus have been used as indicator of fungal food resources (Ruess and Chamberlain 2010, Ferlian et al. 2015). However, 18:2ω6,9 can be synthesized by higher insects (Cripps et al. 1986) and therefore may also be related to species' physiology. Several groups of Collembola are able to synthesize 18:2ω6,9, including Isotomidae, Poduromorpha and Entomobryoidea (Chamberlain et al. 2004, Chamberlain and Black 2005, Haubert et al. 2006), but this has not been tested for Symphypleona. Nevertheless, high proportions of 18:2ω6,9 still may reflect a fungal based diet in Symphypleona (Ruess et al. 2005, Ruess and Chamberlain 2010), but biosynthesis must be excluded by laboratory experiments before concluding that there is an evolutionary constraint in fungal feeding among different Collembola phylogenetic groups.

Strong phylogenetic signal was detected in the proportion of 20:1ω9 of the field-sampled dataset. Collembola unlikely are able to biosynthesize 20:1ω9 *de novo* but rather incorporate it from food, presumably from nematodes (Ruess et al. 2004, Ruess et al. 2005). The lack of 20:1ω9 in the clade of Lepidocyrtidae and Entomobryidae and the clade of Sminthuridae indicates that at our study sites these Collembola did not feed on nematodes, while the remaining species, especially *C. denticulate*, may have consumed nematodes. However, when more species and measurements were included from other studies (the combined dataset), no phylogenetic signal was found in proportion of 20:1ω9, nor in site scores on PC3 and PC4 which were correlated with 20:1ω9. Phylogenetic signal found in our field-sampled dataset may therefore be an exception. Indeed, Collembola from different forest sites have been shown with different proportions of 20:1ω9, presumably related to the amounts of resources in the environment (Ruess et al. 2005).

The ratios of bacterial, fungal and plant FAs, which have been used to assign species to feeding guilds, did not show phylogenetic signal. These results partially support our second hypothesis that food resource FAs are a phylogenetically independent trait, implying niche partitioning in food resources among closely related species, thereby favoring species coexistence. Phylogenetic signal may be reduced due to a mixture of convergent evolution and conservatism in traits, or a developed trait irrespective of species' evolution (i.e., a phylogenetically random trait). Our analyses used ratio as a continuous variable and the ability to detect phylogenetic signal may be reduced due to large intraspecific variation or measurement errors (Ives et al. 2007). Indeed, Collembola are described as generalists able to consume a broad spectrum of food resources, exhibiting a considerable intraspecific variation in biomarker FA proportions from laboratory experiments (Chamberlain et al. 2005, Ruess et al. 2005, Haubert et al. 2011). In field samples, the variation is expected to be even larger, and it is possible that consumption of food resources is influenced by other co-occurring species.

Fatty acid composition complements stable isotopes in analyzing the trophic niche of soil biota (Ferlian et al. 2015). Using taxonomy as a surrogate of phylogenetic relationships with stable isotope data suggests conservatism in Collembola trophic niches (Potapov et al. 2016), in contrast to the findings of the current study. Thus, Collembola feeding traits are, on one hand, likely to have been constrained along species' evolutionary history; on the other hand, they may retain variability to reduce competition. More data on trophic niches and food resources of Collembola species from different phylogenetic groups are necessary to test this hypothesis.

Traits and species coexistence in soil

Species can coexist when they have similar β niche traits and different α niche traits (Silvertown et al. 2006). Phylogenetic signal detected in C20 PUFAs (β niche) but general lability in biomarker FAs and bacterial, fungal and plant FA ratios (α niche) may explain how different Collembola species coexist. Moreover, explicitly testing phylogenetic conservatism in functional traits is crucial for community phylogenetic and trait-based approaches, because the traits are mechanistic links by which phylogenetic history can influence contemporary ecological processes in communities (Cavender-Bares et al. 2009). Phylogenetic signal measurement in this study, therefore, represents a starting point to further investigate evolutionary hypotheses on the adaptation of soil animals to environmental conditions (Revell et al. 2008, Cooper et al. 2010), thereby linking community phylogenetic and trait-based approaches with coexistence studies on soil biota.

Conclusions

Our results show that Collembola FA profiles generally exhibit phylogenetic signal. We found phylogenetic signal in C20 PUFA proportions of Collembola, while biomarker FAs differed among species but were generally labile. These patterns suggest that (1) physiological properties of species may be constrained during evolutionary history, resulting in phylogenetically related species having similar physiologically related FAs, and (2) Collembola food resources are phylogenetically labile, favoring species coexistence. Our study is the first to report phylogenetic signal in the fatty acid compositions of animals in the context of species coexistence. The results form a starting point to further investigate evolutionary hypotheses on the adaptation of soil animals to environmental conditions. Integrating phylogenetic comparative methods and community phylogenetic and trait-based approaches may help identify evolutionary and ecological forces driving and maintaining communities in soil.

Acknowledgements

We thank Olga Ferlian, Dominique Haubert, Christopher Ngosong, Melanie M. Pollierer and Liliane Ruess for providing their data on Collembola fatty acids. We acknowledge the associate editor and three anonymous reviewers for their constructive comments. Special thanks to Guido Humpert and Jo-Fan Chao for support during lab work and Tamara Hartke for improving the English. The authors have no conflict of interest to declare.

Author Contributions

TWC and SS conceived and designed the study; TWC and PS performed the study; TWC, IS and SS wrote the manuscript. All authors revised and approved the manuscript.

Data Accessibility

Concatenated alignment and phylogenetic trees generated from this study were deposited in TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S20409).

References

Ackerly, D.D. & Cornwell, W.K. (2007) A trait-based approach to community assembly: Partitioning of species trait values into within- and among-community components. *Ecology Letters*, **10**, 135–145.

Ackerly, D.D., Schwilk, D.W. & Webb, C.O. (2006) Niche evolution and adaptive radiation: testing the order of trait divergence. *Ecology*, **87**, S50-61.

Adler, P.B., Fajardo, A., Kleinhesselink, A.R. & Kraft, N.J.B. (2013) Trait-based tests of coexistence mechanisms. *Ecology Letters*, **16**, 1294–1306.

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*, **57**, 289-300.

Best, R.J. & Stachowicz, J.J. (2013) Phylogeny as a proxy for ecology in seagrass amphipods: which traits are most conserved? *PLoS ONE*, **8**, e57550.

Blomberg, S.P., Garland Jr., T. & Ives, A.R. (2003) Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution*, **57**, 717–745.

Bodenhofer, U., Bonatesta, E., Horejs-Kainrath, C. & Hochreiter, S. (2015) msa: An R package for multiple sequence alignment. *Bioinformatics*, **31**, 3997–3999.

Buse, T., Ruess, L. & Filser, J. (2013) New trophic biomarkers for Collembola reared on algal diets. *Pedobiologia*, **56**, 153–159.

Cavender-Bares, J., Ackerly, D.D., Baum, D.A. & Bazzaz, F.A. (2004) Phylogenetic overdispersion in Floridian oak communities. *The American Naturalist*, **163**, 823–843.

Cavender-Bares, J., Kozak, K.H., Fine, P. V & Kembel, S.W. (2009) The merging of community ecology and phylogenetic biology. *Ecology Letters*, **12**, 693–715.

Chamberlain, P.M. & Black, H.I.J. (2005) Fatty acid compositions of Collembola: unusually high proportions of C20 polyunsaturated fatty acids in a terrestrial invertebrate. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, **140**, 299–307.

Chamberlain, P.M., Bull, I.D., Black, H.I., Ineson, P. & Evershed, R.P. (2004) Lipid content and carbon assimilation in Collembola: implications for the use of compound-specific carbon isotope analysis in animal dietary studies. *Oecologia*, **139**, 325–335.

Chamberlain, P.M., Bull, I.D., Black, H.I.J., Ineson, P. & Evershed, R.P. (2005) Fatty acid composition and change in Collembola fed differing diets: identification of trophic biomarkers. *Soil Biology and Biochemistry*, **37**, 1608–1624.

Cooper, N., Jetz, W. & Freckleton, R.P. (2010) Phylogenetic comparative approaches for studying niche conservatism. *Journal of Evolutionary Biology*, **23**, 2529–2539.

Cripps, C., Blomquist, G.J. & de Renobales, M. (1986) *De novo* biosynthesis of linoleic acid in insects. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism*, **876**, 572–580.

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

van Dooremalen, C. & Ellers, J. (2010) A moderate change in temperature induces changes in fatty acid composition of storage and membrane lipids in a soil arthropod. *Journal of Insect Physiology*, **56**, 178–184.

Ferlian, O., Klarner, B., Langeneckert, A.E. & Scheu, S. (2015) Trophic niche differentiation and utilisation of food resources in collembolans based on complementary analyses of fatty acids and stable isotopes. *Soil Biology and Biochemistry*, **82**, 28–35.

Ferlian, O. & Scheu, S. (2014) Shifts in trophic interactions with forest type in soil generalist predators as indicated by complementary analyses of fatty acids and stable isotopes. *Oikos*, **123**, 1182–1191.

Fox, J. & Weisberg, S. (2011) *An R Companion to Applied Regression*, 2nd ed. Sage, Thousand Oaks, CA.

Freckleton, R.P., Harvey, P.H. & Pagel, M. (2002) Phylogenetic analysis and comparative data. *The American Naturalist*, **160**, 712–726.

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.

Harvey, P.H. & Pagel, M.D. (1991) The Comparative Method in Evolutionary Biology. Oxford University Press, Oxford.

Haubert, D., Birkhofer, K., Fließbach, A., Gehre, M., Scheu, S. & Ruess, L. (2009) Trophic structure and major trophic links in conventional versus organic farming systems as indicated by carbon stable isotope ratios of fatty acids. *Oikos*, **118**, 1579–1589.

Haubert, D., Häggblom, M.M., Langel, R., Scheu, S. & Ruess, L. (2006) Trophic shift of stable isotopes and fatty acids in Collembola on bacterial diets. *Soil Biology and Biochemistry*, **38**, 2004–2007.

Haubert, D., Häggblom, M.M., Scheu, S. & Ruess, L. (2004) Effects of fungal food quality and starvation on the fatty acid composition of *Protaphorura fimata* (Collembola). *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, **138**, 41–52.

Haubert, D., Häggblom, M.M., Scheu, S. & Ruess, L. (2008) Effects of temperature and life stage on the fatty acid composition of Collembola. *European Journal of Soil Biology*, **44**, 213–219.

Haubert, D., Pollierer, M.M. & Scheu, S. (2011) Fatty acid patterns as biomarker for trophic interactions: Changes after dietary switch and starvation. *Soil Biology and Biochemistry*, **43**, 490–494.

Holm, S. (1979) A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6**, 65-70.

Hopkin, S.P. (1997) Biology of the Springtails (Insecta: Collembola). Oxford University Press, Oxford.

Hopkin, S.P. (2007) A Key to the Collembola (Springtails) of Britain and Ireland (ed Field Studies Council (FSC)). Shrewsbury.

Ives, A.R., Midford, P.E. & Garland, T. (2007) Within-species variation and measurement error in phylogenetic comparative methods. *Systematic Biology*, **56**, 252–270.

Kempson, D., Lloyd, M. & Ghelardi, R. (1963) A new extractor for woodland litter. *Pedobiologia*, **3**, 1–21.

McGill, B.J., Enquist, B.J., Weiher, E. & Westoby, M. (2006) Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, **21**, 178–185.

Münkemüller, T., Lavergne, S., Bzeznik, B., Dray, S., Jombart, T., Schiffers, K. & Thuiller, W. (2012) How to measure and test phylogenetic signal. *Methods in Ecology and Evolution*, **3**, 743–756.

Ngosong, C., Raupp, J., Scheu, S. & Ruess, L. (2009) Low importance for a fungal based food web in arable soils under mineral and organic fertilization indicated by Collembola grazers. *Soil Biology and Biochemistry*, **41**, 2308–2317.

Ngosong, C., Raupp, J., Richnow, H., & Ruess, L. (2011). Tracking Collembola feeding strategies by the natural ^{13}C signal of fatty acids in an arable soil with different fertilizer regimes. *Pedobiologia*, **54**, 225-233.

Pagel, M. (1999) Inferring the historical patterns of biological evolution. *Nature*, **401**, 877–884.

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

Paradis, E. (2013) Molecular dating of phylogenies by likelihood methods: A comparison of models and a new information criterion. *Molecular Phylogenetics and Evolution*, **67**, 436-444.

Pey, B., Nahmani, J., Auclerc, A., Capowiez, Y., Cluzeau, D., Cortet, J., Decaëns, T., Deharveng, L., Dubs, F., Joimel, S., Briard, C., Grumiaux, F., Laporte, M.-A.A., Pasquet, A., Pelosi, C., Pernin, C., Ponge, J.-F., Salmon, S., Santorufo, L. & Hedde, M. (2014) Current use of and future needs for soil invertebrate functional traits in community ecology. *Basic and Applied Ecology*, **15**, 194–206.

Pollierer, M.M., Dyckmans, J., Scheu, S. & Haubert, D. (2012) Carbon flux through fungi and bacteria into the forest soil animal food web as indicated by compound-specific ^{13}C fatty acid analysis. *Functional Ecology*, **26**, 978–990.

Potapov, A.A., Semenina, E.E., Korotkevich, A.Y., Kuznetsova, N.A. & Tiunov, A. V. (2016) Connecting taxonomy and ecology: Trophic niches of collembolans as related to taxonomic identity and life forms. *Soil Biology and Biochemistry*, **101**, 20–31.

R Core Team. (2015) R: A language and environment for statistical computing.

Revell, L.J. (2012) phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, **3**, 217–223.

Revell, L., Harmon, L. & Collar, D. (2008) Phylogenetic signal, evolutionary process, and rate. *Systematic Biology*, **57**, 591–601.

Revell, L.J., Johnson, M.A., Schulte, J.A., Kolbe, J.J. & Losos, J.B. (2007) A phylogenetic test for adaptive convergence in rock-dwelling lizards. *Evolution*, **61**, 2898–2912.

Ronquist, F., Teslenko, M., van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A. & Huelsenbeck, J.P. (2012) Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.

Ruess, L. & Chamberlain, P.M. (2010) The fat that matters: Soil food web analysis using fatty acids and their carbon stable isotope signature. *Soil Biology and Biochemistry*, **42**, 1898–1910.

Ruess, L., Häggblom, M.M., Langel, R. & Scheu, S. (2004) Nitrogen isotope ratios and fatty acid composition as indicators of animal diets in belowground systems. *Oecologia*, **139**, 336–346.

Ruess, L., Schütz, K., Haubert, D., Häggblom, M.M., Kandeler, E. & Scheu, S. (2005) Application of lipid analysis to understand trophic interactions in soil. *Ecology*, **86**, 2075–2082.

Ruess, L., Schütz, K., Migge-Kleian, S., Häggblom, M.M., Kandeler, E. & Scheu, S. (2007) Lipid composition of Collembola and their food resources in deciduous forest stands—Implications for feeding strategies. *Soil Biology and Biochemistry*, **39**, 1990–2000.

Rusek, J. (1998) Biodiversity of Collembola and their functional role in the ecosystem. *Biodiversity and Conservation*, **7**, 1207–1219.

Sechi, V., D'Annibale, A., Ambus, P., Sárossy, Z., Krogh, P.H., Eriksen, J. & Holmstrup, M. (2014) Collembola feeding habits and niche specialization in agricultural grasslands of different composition. *Soil Biology and Biochemistry*, **74**, 31–38.

Silvertown, J., McConway, K., Gowing, D., Dodd, M., Fay, M.F., Joseph, J.A. & Dolphin, K. (2006) Absence of phylogenetic signal in the niche structure of meadow plant communities. *Proceedings of the Royal Society B*, **273**, 39–44.

Stamatakis, A. (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.Stanley-Samuelson, D.W. (1994) Assessing the significance of prostaglandins and other eicosanoids in insect physiology. *Journal of Insect Physiology*, **40**, 3–11.

Stanley-Samuelson, D.W., Dell, T. & Ogg, C.L. (1992) Polyunsaturated fatty acid metabolism inferred from fatty acid compositions of the diets and tissues of the gypsy moth *Lymantria dispar*. *Comparative Biochemistry and Physiology*, **102A**, 173–178.

Vaidya, G., Lohman, D.J. & Meier, R. (2011) SequenceMatrix: Concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, **27**, 171–180.

Venables, W.N. & Ripley, B.D. (2002) *Modern Applied Statistics with S*, 4th ed. Springer, New York.

Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. (2007) Let the concept of trait be functional! *Oikos*, **116**, 882–892.

Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, **33**, 475–505.

Wright, E.S. (2015) DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics*, **16**, 322.

Supplementary Materials

Table S3.1

NCBI accession numbers of sequences used to construct the molecular phylogeny of Collembola.

Species	**Molecular markers**					
	18S rRNA	**28S rRNA**			**Histone H3**	**COI**
		D1	**D2**	**D3-D5**		
Zygentoma (outgroup)	AF370791	AY859557	AY859557	AY859557	AY749703	KM535783
Callibaetis (outgroup)	AY521826	AY521735	AY521735	AY521735	AY521695	JF735107
Machilis (outgroup)	AY210811	AY210810	AY210810	AY210810	AY338644	JN970940
Allacma fusca	KY230702	KY230832	KY230832	KY230935		KY231097
Brachystomella parvula	KY230724	KY230822	KY230822	KY230925	KY231066	KY231088
Ceratophysella denticulata	KY230747	KY230847	KY230847	KY230948	KY231036	KY231107
Ceratophysella succinea		KY230885	KY230885	KY230990	KY231065	KY231136
Desoria violacea	KY230736	KY230833	KY230833	KY230902	KY231003	KY231098
Deuterosminthurus sulphureus	KY230753	KY230830	KY230830	KY230933	KY231024	KY231095
Dicyrtomina ornata	KY230768	KY230840	KY230840	KY230968	KY231046	KY231125
Dicyrtomina saundersi	EU368611	EF199974	EF199974	EF199974		
Entomobrya muscorum	KY230710	KY230806	KY230806	KY230909	KY231010	
Entomobrya nicoleti	KY230740	KY230838	KY230838	KY230941	KY231030	KY231119
Entomobrya nivalis		LK024313	LK024313			HG422598
Folsomia quadrioculata	KY230755	KY230853	KY230853	KY230955	KF684772	KF684607
Friesea claviseta	KY230727	KY230826	KY230826	KY230929	KY231020	KY231092
Isotoma anglicana	KY230779	KY230873	KY230873	KY230975	KY231053	KY231076
Isotoma viridis	KY230708	KY230835	KY230835	KY230938	KY231028	KY231129
Isotomiella minor	KY230744	KY230843	KY230843	KY230945	KY231034	KY231103
Isotomurus fucicolus	KY230704	KY230834	KY230834	KY230937	KY231057	KY231099
Lepidocyrtus cyaneus	KY230751	KY230851	KY230851	KY230952	KY231040	KY231111
Lepidocyrtus lanuginosus	KY230748	KY230848	KY230848	KY230949	KY231037	KY231108
Lepidocyrtus lignorum	KY230749	KY230849	KY230849	KY230950	KY231038	KY231109
Lepidocyrtus sp	KY230750	KY230850	KY230850	KY230951	KY231039	KY231110
Neanura muscorum	AY555520	AJ251733	AJ251733	AJ251733		AY555544
Onychiurus ambulans	AY555518	AF483384	AF483442	HQ731961	AY555564	HQ732075
Orchesella flavescens	KY230714	KY230811	KY230811	KY230997	KY231015	KY231082
Orchesella villosa	KY230715	KY230812	KY230812	KY230904	KY231005	KY231091
Parisotoma notabilis	KY230772	KY230872	KY230872	KY230974	KY231052	KY231128
Pogonognathellus flavescens	KY230717	KY230814	KY230814	KY230917	KY231016	KY231083
Pogonognathellus longicornis		KY230845	KY230845	KY230946	KY231035	KY231105
Protaphorura armata		AF483391	AF483449	HQ731965		HQ732078
Protaphorura sp1	KY230789	KY230884	KY230884	KY230989	KY231063	HG422585
Protaphorura sp2	KY230719	KY230856	KY230856	KY230959	KY231064	KY231116
Pseudosinella immaculata	KY230712			KY230911	KY231012	KY231078
Sminthurus viridis	KY230701	KY230880	KY230880	KY230983	KY231001	JN970939
Tomocerus baudoti			JX261697			JX261845
Tomocerus minor	AY555516	AF483406	JX261700	HQ731971	AY555562	HM398041
Tomocerus vulgaris	KY230777	KY230815	KY230815	KY230918	KY231058	KY231131
Willemia anophthalma	KY230726	KY230869	KY230869	KY230927		KY231090

Table S3.2

NLFAs proportions (mean ± standard error, %) for the 13 Collembola species collected in this study, pooled for sites and habitats. Fatty acid names in bold indicate significant difference between different species on fatty acid proportions in ANOVA (Table S3.3). Different letters next to fatty acid proportions indicate significant differences ($P < 0.05$) between species in Tukey's HSD test.

Fatty acid	*Lepidocyrtus lanuginosus*		*Lepidocyrtus cyaneus*		*Pseudosinella immaculata*		*Entomobrya nicoleti*		*Entomobrya muscorum*		*Orchesella villosa*	
8:0	0.0 ± 0.0		1.8 ± 0.7		0.0		0.0 ± 0.0		1.2 ± 0.9		3.5 ± 1.2	
10:0	0.0 ± 0.0		0.1 ± 0.1		0.0		1.1 ± 1.1		0.3 ± 0.3		1.9 ± 0.8	
2-OH 10:0	0.0 ± 0.0		0.0 ± 0.0		0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0	
12:0	0.0 ± 0.0		0.0 ± 0.0		0.0		0.0 ± 0.0		0.0 ± 0.0		0.3 ± 0.3	
14:0	**1.6 ± 0.3**	**abc**	**0.6 ± 0.2**	**b**	**11.4**	**d**	**1.4 ± 0.1**	**abcd**	**1.7 ± 0.6**	**abc**	**5.0 ± 1.2**	**cd**
14:1	0.0 ± 0.0		0.0 ± 0.0		0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0	
15:0	**0.0 ± 0.0**	**b**	**0.0 ± 0.0**	**b**	**0.0**	**b**	**0.0 ± 0.0**	**b**	**0.0 ± 0.0**	**b**	**0.0 ± 0.0**	**b**
a15:0	0.0 ± 0.0		0.0 ± 0.0		0.0		0.9 ± 0.9		0.0 ± 0.0		0.9 ± 0.5	
i15:0	0.0 ± 0.0		0.0 ± 0.0		0.0		0.7 ± 0.7		0.0 ± 0.0		2.7 ± 1.0	
16:0	**24.6 ± 0.6**	**c**	**26.4 ± 0.7**	**c**	**23.3**	**abc**	**24.0 ± 0.4**	**bc**	**24.6 ± 1.1**	**c**	**17.6 ± 2.5**	**bc**
16:1ω5	0.0 ± 0.0		0.0 ± 0.0		0.0		0.6 ± 0.6		0.0 ± 0.0		0.2 ± 0.2	
16:1ω7	**1.9 ± 0.1**	**bc**	**2.0 ± 0.6**	**b**	**1.2**	**abc**	**0.8 ± 0.8**	**bc**	**1.4 ± 0.4**	**b**	**2.3 ± 0.6**	**bc**
i16:0	0.0 ± 0.0		0.0 ± 0.0		0.0		0.0 ± 0.0		0.2 ± 0.2		0.2 ± 0.2	
17:0	0.2 ± 0.2		0.0 ± 0.0		0.0		0.0 ± 0.0		0.0 ± 0.0		0.1 ± 0.1	
17:1ω8	0.0 ± 0.0		0.0 ± 0.0		0.0		1.0 ± 1.0		0.0 ± 0.0		0.2 ± 0.2	
i17:0	0.0 ± 0.0		0.0 ± 0.0		0.0		0.6 ± 0.6		0.3 ± 0.3		1.0 ± 0.6	
cy17:0	0.0 ± 0.0		0.0 ± 0.0		0.0		0.0 ± 0.0		0.0 ± 0.0		0.3 ± 0.2	
18:0	**7.3 ± 0.8**	**abc**	**6.9 ± 0.9**	**abc**	**7.7**	**abc**	**11.2 ± 1.8**	**abc**	**10.4 ± 1.7**	**bc**	**10.8 ± 1.9**	**bc**
18:1ω7	**1.2 ± 0.2**	**a**	**0.8 ± 0.4**	**a**	**0.0**	**abc**	**0.0 ± 0.0**	**a**	**0.0 ± 0.0**	**a**	**1.5 ± 0.8**	**a**
18:1ω9	**36.3 ± 1.8**	**a**	**38.0 ± 1.6**	**a**	**42.2**	**ab**	**34.6 ± 6.1**	**ab**	**43.1 ± 4.4**	**a**	**29.9 ± 3.5**	**a**
18:2ω6,9	**20.9 ± 1.5**	**abc**	**17.1 ± 1.7**	**acd**	**7.0**	**acd**	**14.2 ± 2.1**	**abcd**	**9.6 ± 1.4**	**d**	**9.6 ± 1.4**	**d**
19:0	0.0 ± 0.0		0.3 ± 0.2		0.0		0.0 ± 0.0		0.0 ± 0.0		0.4 ± 0.4	
cy19:0	0.4 ± 0.3		0.1 ± 0.1		0.0		1.3 ± 1.3		0.0 ± 0.0		0.0 ± 0.0	
20:1ω9	0.0 ± 0.0		0.0 ± 0.0		0.0		0.0 ± 0.0		0.0 ± 0.0		0.2 ± 0.2	
20:2ω6,9	0.0 ± 0.0		0.0 ± 0.0		0.0		0.0 ± 0.0		0.0 ± 0.0		0.1 ± 0.1	
20:3ω6	0.0 ± 0.0		0.0 ± 0.0		0.0		0.0 ± 0.0		0.0 ± 0.0		0.5 ± 0.4	
20:4ω6	**2.1 ± 0.2**	**abd**	**1.0 ± 0.3**	**bcd**	**3.9**	**abcd**	**3.1 ± 0.1**	**abcd**	**2.7 ± 0.6**	**ad**	**4.2 ± 0.7**	**a**
20:5ω3	**3.4 ± 0.2**	**cd**	**4.6 ± 0.4**	**de**	**3.4**	**abcdef**	**3.7 ± 0.4**	**bcdef**	**3.9 ± 0.9**	**cde**	**4.9 ± 1.1**	**de**
22:1ω9	**0.2 ± 0.2**	**ab**	**0.0 ± 0.0**	**a**	**0.0**	**ab**	**0.0 ± 0.0**	**ab**	**0.0 ± 0.0**	**a**	**0.0 ± 0.0**	**a**
22:2	0.0 ± 0.0		0.0 ± 0.0		0.0		0.0 ± 0.0		0.0 ± 0.0		0.7 ± 0.7	
23:0	**0.0 ± 0.0**	**a**	**0.0 ± 0.0**	**a**	**0.0**	**ab**	**0.5 ± 0.5**	**ab**	**0.2 ± 0.2**	**a**	**0.4 ± 0.2**	**a**
24:1	0.0 ± 0.0		0.1 ± 0.1		0.0		0.0 ± 0.0		0.3 ± 0.3		0.5 ± 0.3	

Table S3.2 Continued

Fatty acid	*Tomocerus vulgaris*		*Pogonognathellus flavescens*		*Sminthurus viridis*		*Allacma fusca*		*Deuterosminthurus sulphureus*		*Ceratophysella denticulata*		*Isotoma viridis*	
8:0	0.0 ± 0.0		0.9 ± 0.7		1.2 ± 1.2		4.9 ± 1.7		1.7 ± 1.7		0.0 ± 0.0		0.0 ± 0.0	
10:0	0.0 ± 0.0		0.5 ± 0.3		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		1.6 ± 0.7	
2-OH 10:0	0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.4 ± 0.3	
12:0	0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.2 ± 0.2	
14:0	**0.3 ± 0.3**	**ab**	**1.0 ± 0.8**	**ab**	**1.5 ± 0.3**	**abc**	**0.4 ± 0.4**	**ab**	**5.4 ± 0.6**	**cd**	**0.5 ± 0.5**	**ab**	**3.1 ± 0.6**	**acd**
14:1	0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.4 ± 0.3	
15:0	**0.0 ± 0.0**	**b**	**0.0 ± 0.0**	**b**	**0.0 ± 0.0**	**b**	**0.0 ± 0.0**	**b**	**0.0 ± 0.0**	**b**	**1.3 ± 0.8**	**a**	**0.0 ± 0.0**	**b**
a15:0	0.0 ± 0.0		0.0 ± 0.0		0.4 ± 0.4		0.3 ± 0.3		0.0 ± 0.0		0.0 ± 0.0		1.0 ± 0.4	
i15:0	0.0 ± 0.0		0.0 ± 0.0		0.9 ± 0.4		0.0 ± 0.0		3.4 ± 1.9		0.7 ± 0.7		2.0 ± 0.5	
16:0	**17.9 ± 0.5**	**bc**	**22.6 ± 1.4**	**bc**	**16.1 ± 2.3**	**bc**	**9.6 ± 7.9**	**a**	**8.9 ± 1.4**	**ab**	**15.2 ± 4.6**	**abc**	**19.8 ± 1.4**	**bc**
16:1ω5	0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0	
16:1ω7	**0.6 ± 0.3**	**b**	**0.5 ± 0.2**	**b**	**0.6 ± 0.3**	**b**	**0.8 ± 0.8**	**b**	**1.8 ± 0.4**	**bc**	**8.3 ± 1.5**	**a**	**5.7 ± 1.3**	**ac**
i16:0	0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.3 ± 0.2	
17:0	0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.7 ± 0.7		0.0 ± 0.0		0.0 ± 0.0		0.6 ± 0.4	
17:1ω8	0.0 ± 0.0		0.0 ± 0.0		0.2 ± 0.2		1.3 ± 1.3		0.0 ± 0.0		0.0 ± 0.0		0.2 ± 0.2	
i17:0	0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		6.8 ± 6.8		0.0 ± 0.0		0.4 ± 0.4		0.6 ± 0.4	
cy17:0	0.0 ± 0.0		0.0 ± 0.0		0.2 ± 0.2		0.0 ± 0.0		1.3 ± 0.0		0.0 ± 0.0		0.2 ± 0.2	
18:0	**9.0 ± 0.6**	**abc**	**12.7 ± 1.5**	**bc**	**6.2 ± 3.3**	**ac**	**7.3 ± 7.3**	**ac**	**0.9 ± 0.9**	**a**	**5.1 ± 2.7**	**abc**	**15.8 ± 0.9**	**b**
18:1ω7	**6.7 ± 0.9**	**c**	**5.7 ± 1.7**	**bc**	**0.0 ± 0.0**	**a**	**1.1 ± 0.8**	**a**	**0.0 ± 0.0**	**a**	**1.0 ± 0.5**	**ab**	**1.5 ± 0.8**	**a**
18:1ω9	**34.6 ± 0.8**	**a**	**29.0 ± 1.8**	**a**	**32.4 ± 3.1**	**a**	**26.5 ± 7.6**	**ab**	**41.9 ± 2.1**	**a**	**34.8 ± 11.7**	**a**	**12.4 ± 3.1**	**b**
18:2ω6,9	**19.3 ± 0.2**	**abcd**	**14.6 ± 1.2**	**acd**	**37.6 ± 3.8**	**b**	**34.9 ± 11.1**	**ab**	**32.3 ± 2.2**	**ab**	**25.4 ± 1.5**	**abc**	**12.4 ± 1.7**	**cd**
19:0	0.0 ± 0.0		0.0 ± 0.0		1.2 ± 0.8		0.4 ± 0.4		0.0 ± 0.0		0.0 ± 0.0		0.3 ± 0.3	
cy19:0	0.0 ± 0.0		0.0 ± 0.0		1.0 ± 0.6		1.7 ± 1.7		0.0 ± 0.0		0.0 ± 0.0		1.0 ± 0.5	
20:1ω9	0.3 ± 0.3		0.1 ± 0.1		0.0 ± 0.0		0.0 ± 0.0		0.5 ± 0.5		1.7 ± 0.9		0.5 ± 0.3	
20:2ω6,9	0.6 ± 0.3		0.2 ± 0.2		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.7 ± 0.7		0.5 ± 0.3	
20:3ω6	0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.7 ± 0.4	
20:4ω6	**4.6 ± 0.6**	**a**	**3.8 ± 1.0**	**a**	**0.0 ± 0.0**	**c**	**0.0 ± 0.0**	**bc**	**0.0 ± 0.0**	**bc**	**4.1 ± 0.4**	**a**	**4.9 ± 0.9**	**a**
20:5ω3	**6.1 ± 0.6**	**def**	**8.1 ± 1.2**	**ef**	**0.0 ± 0.0**	**a**	**0.0 ± 0.0**	**ab**	**0.0 ± 0.0**	**ab**	**0.8 ± 0.4**	**abc**	**11.6 ± 2.5**	**f**
22:1ω9	**0.0 ± 0.0**	**ab**	**0.0 ± 0.0**	**a**	**0.3 ± 0.3**	**ab**	**0.0 ± 0.0**	**ab**	**0.0 ± 0.0**	**ab**	**0.0 ± 0.0**	**ab**	**1.3 ± 0.7**	**b**
22:2	0.0 ± 0.0		0.2 ± 0.2		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		0.0 ± 0.0		1.0 ± 0.5	
23:0	**0.0 ± 0.0**	**a**	**0.2 ± 0.2**	**a**	**0.0 ± 0.0**	**a**	**3.2 ± 1.9**	**b**	**0.4 ± 0.4**	**ab**	**0.0 ± 0.0**	**a**	**0.0 ± 0.0**	**a**
24:1	0.0 ± 0.0		0.0 ± 0.0		0.2 ± 0.2		0.0 ± 0.0		1.5 ± 0.2		0.0 ± 0.0		0.0 ± 0.0	

Table S3.3

ANOVA table of logit-transformed proportions for each fatty acid from 13 field-collected Collembola species. Species and habitats are used as explanatory factors and sites and habitats as error terms in the model. *P*-values were obtained after Holm's correction. Fatty acids significantly different between species are marked in bold.

Fatty acid	**Species**			**Habitat:Species**		
	df	**F**	**Adjusted *P***	**df**	**F**	**Adjusted *P***
8:0	12	2.111	0.539	8	1.077	0.975
10:0	12	1.622	1.000	8	1.083	0.975
2-OH 10:0	12	2.050	0.539	8	2.288	0.211
12:0	12	0.182	1.000	8	0.235	1.000
14:0	**12**	**5.324**	**0.001**	8	2.063	0.277
14:1	12	2.258	0.423	8	2.520	0.193
15:0	**12**	**4.630**	**0.002**	8	0.000	1.000
a15:0	12	1.364	1.000	8	1.895	0.340
i15:0	12	2.894	0.100	8	1.331	0.738
16:0	**12**	**4.759**	**0.002**	8	0.214	1.000
16:1ω5	12	1.336	1.000	8	0.091	1.000
16:1ω7	**12**	**6.120**	**< 0.001**	8	1.470	0.626
i16:0	12	0.558	1.000	8	0.629	1.000
17:0	12	0.954	1.000	8	0.934	1.000
17:1ω8	12	1.135	1.000	8	0.623	1.000
i17:0	12	0.894	1.000	8	1.003	1.000
cy17:0	12	2.150	0.521	8	1.660	0.482
18:0	**12**	**3.657**	**0.018**	8	0.901	1.000
18:1ω7	**12**	**4.975**	**0.001**	8	0.855	1.000
18:1ω9	**12**	**4.761**	**0.002**	8	0.420	1.000
18:2ω6,9	**12**	**6.162**	**< 0.001**	8	0.705	1.000
19:0	12	2.571	0.214	8	2.514	0.193
cy19:0	12	2.508	0.237	8	2.272	0.211
20:1ω9	12	2.062	0.539	8	0.278	1.000
20:2ω6,9	12	0.962	1.000	8	0.087	1.000
20:3ω6	12	0.734	1.000	8	0.272	1.000
20:4ω6	**12**	**8.411**	**< 0.001**	8	0.354	1.000
20:5ω3	**12**	**16.549**	**< 0.001**	8	0.354	1.000
22:1ω9	**12**	**3.519**	**0.023**	8	2.538	0.193
22:2	12	0.688	1.000	8	0.043	1.000
23:0	**12**	**3.663**	**0.018**	8	2.642	0.193
24:1	12	2.109	0.539	8	0.493	1.000

Table S3.4

Correlations of different fatty acids with the first four axes (PCs) from principle component analyses. *P*-values were adjusted using Holm's method. Fatty acids correlated with the first PC are marked in bold.

Fatty acid	**PC1**		**PC2**		**PC3**		**PC4**	
	***cor**	***P***	**cor**	***P***	**cor**	***P***	**cor**	***P***
Field-derived species								
8:0	0.483	1.000	0.121	1.000	-0.225	1.000	0.649	0.476
10:0	-0.566	1.000	0.036	1.000	-0.575	0.832	0.248	1.000
2-OH 10:0	-0.404	1.000	0.224	1.000	-0.697	0.220	-0.121	1.000
12:0	-0.402	1.000	0.013	1.000	-0.744	0.107	0.141	1.000
14:0	-0.117	1.000	**-0.850**	**0.008**	-0.361	1.000	0.074	1.000
14:1	-0.404	1.000	0.224	1.000	-0.697	0.220	-0.121	1.000
15:0	0.161	1.000	0.151	1.000	-0.125	1.000	-0.770	0.065
a15:0	-0.259	1.000	0.103	1.000	-0.656	0.360	0.277	1.000
i15:0	0.175	1.000	-0.218	1.000	-0.704	0.203	-0.130	1.000
16:0	**-0.778**	**0.048**	-0.235	1.000	0.422	1.000	-0.178	1.000
16:1ω5	-0.185	1.000	-0.103	1.000	-0.013	1.000	0.202	1.000
16:1ω7	-0.095	1.000	0.033	1.000	-0.547	1.000	-0.751	0.092
i16:0	-0.495	1.000	-0.000	1.000	-0.627	0.478	0.077	1.000
17:0	0.073	1.000	0.413	1.000	-0.669	0.311	0.352	1.000
17:1ω8	0.237	1.000	0.238	1.000	-0.263	1.000	0.538	1.000
i17:0	0.298	1.000	0.339	1.000	-0.486	1.000	0.518	1.000
cy17:0	0.484	1.000	-0.354	1.000	-0.321	1.000	-0.073	1.000
18:0	**-0.953**	**< 0.001**	0.158	1.000	-0.022	1.000	0.135	1.000
18:1ω7	-0.357	1.000	0.674	0.358	0.252	1.000	-0.003	1.000
18:1ω9	0.235	1.000	-0.573	1.000	0.731	0.130	-0.081	1.000
18:2ω6,9	**0.849**	**0.007**	0.372	1.000	0.087	1.000	-0.231	1.000
19:0	0.382	1.000	0.178	1.000	-0.182	1.000	0.225	1.000
cy19:0	0.208	1.000	0.314	1.000	-0.294	1.000	0.339	1.000
20:1ω9	0.122	1.000	0.195	1.000	-0.331	1.000	**-0.805**	**0.029**
20:2ω6,9	-0.294	1.000	0.516	1.000	-0.243	1.000	-0.584	0.973
20:3ω6	-0.452	1.000	0.104	1.000	**-0.814**	**0.022**	0.047	1.000
20:4ω6	**-0.854**	**0.006**	0.037	1.000	-0.115	1.000	-0.276	1.000
20:5ω3	**-0.948**	**< 0.001**	0.181	1.000	-0.019	1.000	0.069	1.000
22:1ω9	-0.320	1.000	0.227	1.000	-0.631	0.478	-0.154	1.000
22:2	-0.508	1.000	0.189	1.000	-0.774	0.059	0.042	1.000
23:0	0.515	1.000	0.259	1.000	-0.237	1.000	0.627	0.613
24:1	0.493	1.000	-0.429	1.000	-0.171	1.000	-0.029	1.000
Literature 37 species								
a15:0	-0.066	1.000	**-0.618**	**0.001**	-0.355	0.336	-0.027	1.000
i15:0	0.148	1.000	-0.240	1.000	-0.300	0.569	-0.246	0.878
16:0	-0.256	1.000	**-0.870**	**< 0.001**	-0.040	1.000	-0.181	1.000
i16:0	-0.215	1.000	**-0.461**	**0.045**	-0.219	1.000	0.290	0.820
16:1ω7	0.207	1.000	0.029	1.000	**0.531**	**0.010**	0.351	0.400
i17:0	0.150	1.000	0.028	1.000	-0.214	1.000	-0.189	1.000
18:0	**-0.643**	**< 0.001**	**-0.481**	**0.031**	-0.046	1.000	-0.112	1.000
18:1ω7	**-0.825**	**< 0.001**	-0.063	1.000	-0.252	0.932	0.394	0.224
18:1ω9	**0.477**	**0.031**	0.183	1.000	**0.586**	**0.002**	-0.283	0.820
18:2ω6,9	**0.542**	**0.006**	**0.572**	**0.003**	**-0.513**	**0.014**	0.256	0.878
cy19:0	0.110	1.000	0.326	0.441	-0.311	0.551	-0.180	1.000
20:1ω9	0.052	1.000	-0.163	1.000	**0.523**	**0.012**	**0.526**	**0.013**
20:2ω6,9	-0.136	1.000	-0.002	1.000	-0.356	0.336	0.327	0.532
20:3ω6	0.343	0.379	0.071	1.000	**0.712**	**< 0.001**	0.277	0.820
20:4ω6	**-0.649**	**< 0.001**	0.344	0.370	0.182	1.000	0.427	0.125
20:5ω3	**-0.696**	**< 0.001**	**0.562**	**0.004**	0.168	1.000	-0.393	0.224

* Pearson correlation coefficient

Table S3.5

ANOVA table of logit-transformed proportions for each fatty acid in the expanded 37-species dataset. Species and habitats are used as explanatory factors in the model. *P*-values were obtained after Holm's correction. Fatty acids significantly different between species are marked in bold.

Fatty acid	**Habitat**			**Species**			**Habitat:Species**		
	df	**F**	**Adjusted P**	**df**	**F**	**Adjusted P**	**df**	**F**	**Adjusted P**
a15:0	2	3.090	0.327	36	0.877	1.000	9	0.415	1.000
i15:0	2	2.899	0.327	**36**	**2.146**	**0.011**	9	0.811	1.000
16:0	**2**	**23.017**	**< 0.001**	**36**	**3.557**	**< 0.001**	9	1.065	1.000
16:1ω7	2	1.545	0.825	**36**	**2.370**	**0.004**	9	0.347	1.000
i16:0	2	3.160	0.327	36	1.422	0.440	9	0.192	1.000
i17:0	2	0.535	0.825	36	0.716	1.000	9	0.845	1.000
18:0	**2**	**15.104**	**< 0.001**	**36**	**2.619**	**0.001**	**9**	**3.424**	**0.016**
18:1ω7	**2**	**35.324**	**< 0.001**	36	1.522	0.319	9	0.693	1.000
18:1ω9	**2**	**11.782**	**< 0.001**	**36**	**2.278**	**0.006**	9	1.174	1.000
18:2ω6,9	**2**	**31.610**	**< 0.001**	**36**	**3.176**	**< 0.001**	9	0.918	1.000
cy19:0	**2**	**12.049**	**< 0.001**	36	1.379	0.440	9	1.191	1.000
20:1ω9	**2**	**40.593**	**< 0.001**	**36**	**3.760**	**< 0.001**	9	2.943	0.054
20:2ω6,9	2	1.603	0.825	36	0.841	1.000	9	0.103	1.000
20:3ω6	**2**	**87.090**	**< 0.001**	**36**	**3.637**	**< 0.001**	**9**	**3.016**	**0.047**
20:4ω6	**2**	**5.363**	**0.049**	**36**	**4.431**	**< 0.001**	9	1.072	1.000
20:5ω3	2	1.244	0.825	**36**	**3.418**	**< 0.001**	9	1.961	0.671

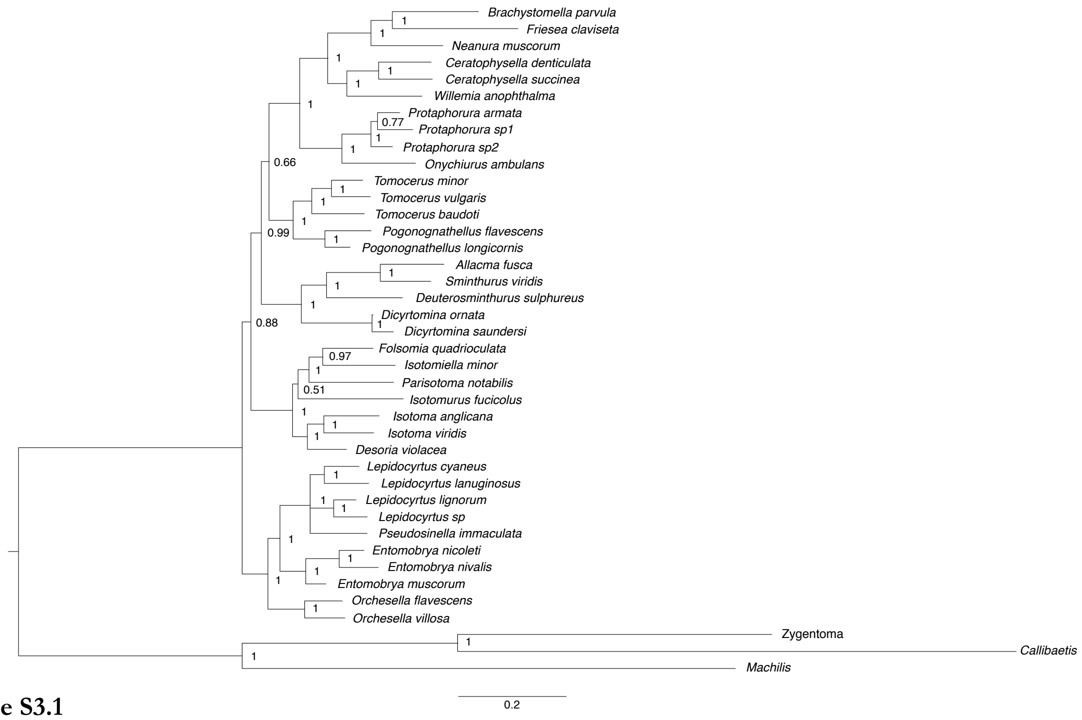

Figure S3.1

Bayesian Inference phylogeny of Collembola based on the concatenated DNA sequences of ribosomal 18S and 28S rRNA and cytochrome oxidase I and Histone H3 genes. *Callibaetis*, *Machilis* and Zygentoma served as outgroups. Numbers at nodes represent Bayesian posterior probabilities.

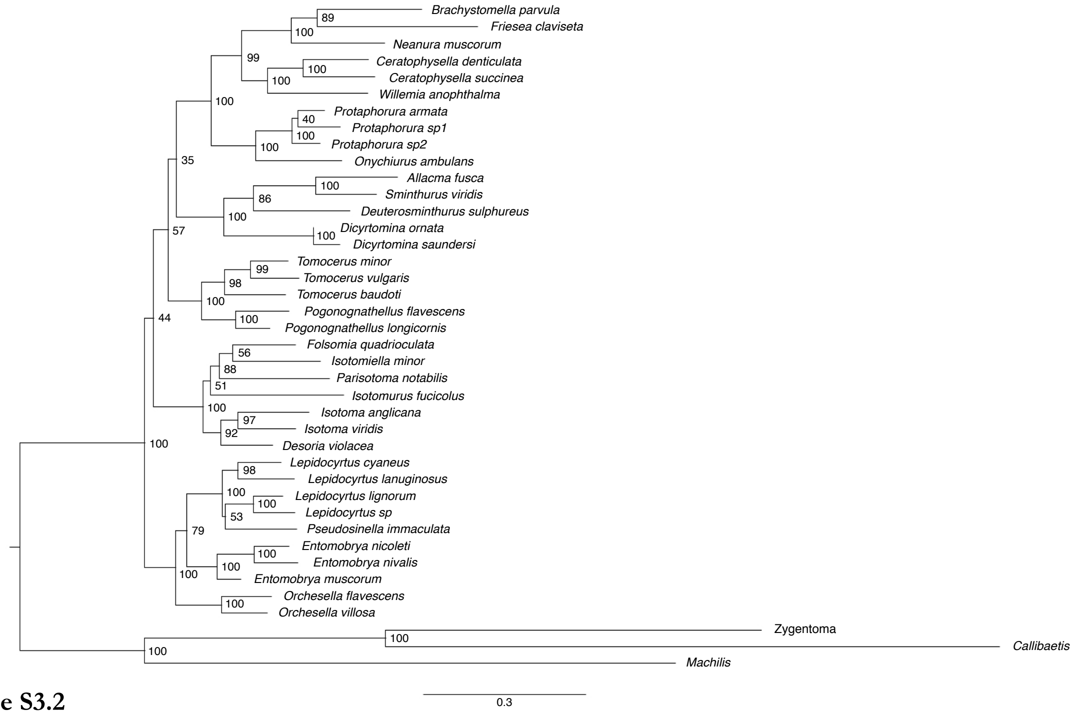

Figure S3.2

Maximum likelihood phylogeny of Collembola based on the concatenated DNA sequences of ribosomal 18S and 28S rRNA and cytochrome oxidase I and Histone H3 genes. *Callibaetis*, *Machilis* and Zygentoma served as outgroups. Numbers at nodes represent bootstrap values from Maximum Likelihood analyses.

Chapter 4

Mechanisms of Collembola species coexistence as indicated by phylogeny and functional traits

Ting-Wen Chen, Jo-Fan Chao, Victoria Kreipe, Ina Schaefer, Matty P. Berg, Stefan Scheu

Abstract

The mechanisms driving and maintaining species coexistence in soil have long puzzled ecologists. To investigate the relative contributions to community assembly of different processes, such as environmental filtering and interspecific competition, community phylogenetic and trait-based approaches have recently been developed. If process-related traits exhibit phylogenetic signal, the presence of phylogenetically closely related species within local communities point to environmental filtering as the major structuring force, while low relatedness among coexisting species point to the dominance of biotic interactions or niche partitioning. In this study we inferred assembly processes of Collembola communities sampled from arable fields, grasslands and forests by exploring phylogenetic relatedness and functional trait similarities of local communities. We found associations between habitat types and Collembola phylogenetic group at the metacommunity scale. The results indicate that Collembola in arable fields were mainly structured by environmental filtering, while niche partitioning dominated in forests. Epedaphic (surface-living) species showed phylogenetic clustering in grasslands and forests, while in forests they also possessed similar traits. Hemiedaphic (sub-surface-dwelling) species in arable fields and grasslands were phylogenetically clustered, but in forests they were phylogenetically overdispersed with different traits. However, the assembly patterns of euedaphic (soil-dwelling) species in each of the three habitat types did not differ from random patterns. Furthermore, different phylogenetic (taxonomic) groups of Collembola showed different patterns in the three habitats. Overall, the results suggest that Collembola assemblages were driven by different mechanisms in the studied habitats, with the relative importance of these mechanisms varying between soil strata and between phylogenetic lineages. Future studies integrating phylogenetic comparative methods, trait-based approaches and community phylogeny will allow novel insight into assembly processes of soil communities.

Keywords

community assembly; co-occurrence; disturbance; environmental filtering; habitat; niche partitioning; soil; springtail; stochasticity; trait

Introduction

Soils are among the most biodiverse ecosystems on earth and have been viewed as "the poor man's tropical rainforest" (Giller 1996). The coexistence of a multitude of animal species has puzzled soil ecologists for long, as reflected in the phrase "the enigma of soil animal species diversity" (Anderson 1975). How can so many species co-occur, and what are the mechanisms driving and maintaining species coexistence in local communities?

Community composition is influenced by a number of processes, including niche-related (Chase and Leibold 2003), neutral (Hubbell 2001) and biogeographical processes (Ricklefs 1987). Community phylogenetic framework (Webb et al. 2002, Cavender-Bares et al. 2009) and trait-based approaches (McGill et al. 2006; Adler et al. 2013) have been developed to investigate the relative importance of these processes for community assembly. In both frameworks, functional traits are crucial, since they form the mechanistic link between evolutionary processes and contemporary ecological processes (Cavender-Bares et al. 2009). On the one hand, traits are characters of species derived from their ancestors and thus, exhibit phylogenetic signal, i.e., phylogenetically related species possess similar traits. On the other hand, traits influence the performance and fitness of species in certain environments and are thus functional (Violle et al. 2007). If process-related traits exhibit phylogenetic signal, a community composed of phylogenetically closely related species is inferred to be structured by environmental filtering. In contrast, low relatedness among coexisting species points to the dominance of competitive interactions or niche partitioning (Webb et al. 2002, Cavender-Bares et al. 2009; but see Gerhold et al. 2015). Furthermore, using a phylogeny-weighted community matrix allows the measurement of phylogenetic dissimilarities between local communities and a trait-weighted matrix the measurement of trait dissimilarities (Pillar and Duarte, 2010). The phylogeny- or trait-weighted community matrix, combined with principle coordinate analysis, generates the principal coordinates of phylogenetic or trait structure of a metacommunity, allowing the identification of specific phylogenetic clades or species with certain traits that associate with habitats (Duarte, 2011, Duarte et al. 2016). In this study we adopted community phylogenetic approaches for the first time to investigate community assembly processes in one of the most abundant and diverse soil invertebrates, Collembola (springtails).

Collembola are basal Hexapoda that occur in high densities in soil with 10,000–100,000 individuals m^{-2} and local diversities of 60–80 species (Petersen and Luxton 1982). They occupy a wide variety of ecological niches, regulate soil microbial activity and contribute to decomposition processes and nutrient cycling (Petersen and Luxton, 1982, Rusek, 1998, Schaefer et al. 2009). Among the few soil taxa for which trait databases are available (Vandewalle et al. 2010, Pey et al. 2014, Moretti et al. 2017, Matty P. Berg, *unpublished data*), Collembola are ideal subjects for the application of trait-based and community phylogenetic approaches to investigate community assembly. The environmental associations of community structures and species functional traits suggest that both abiotic and biotic factors select for certain traits of Collembola (Ponge et al. 2003, Sousa et al. 2006, Makkonen et al. 2011, Bokhorst et al. 2012, Martins da Silva et al. 2012, Salmon and Ponge 2012, Heiniger et al. 2014, Widenfalk et al. 2015, Widenfalk et al. 2016, Pollierer and Scheu 2017). Spherical body shape, large body size, dark pigmentation and sexual reproduction are characteristics of species preferentially occurring in open habitats and at the soil surface, whereas small body size, lack of eye spots, pale color and asexual reproduction are typical traits of species in stable environment and in soil (Salmon et al. 2014). Furthermore, co-occurring Collembola species possessing similar traits suggest that disturbed habitats structure community assembly via abiotic factors (Widenfalk et al. 2015), while niche partitioning or interspecific competition are likely predominant processes in stable environments that result in co-occurring species with different traits (Widenfalk et al. 2016). Since the functional traits of Collembola derive from ancestors and thus exhibit phylogenetic signal (as being tested in **Chapter 2**), similar traits in local community also likely reflect close phylogenetic relationships between coexisting species.

In this study we explored habitat associations of Collembola phylogenetic groups and phylogenetic and trait patterns in local communities collected from arable fields, grasslands and forests, three types of habitats characterized by distinct disturbance regimes and dominating mosaic landscapes in Central Europe. We hypothesized that species co-occurring in heavily disturbed habitats, such as arable fields, show phylogenetic clustering and similar traits (Ding et al. 2012, Gianuca et al. 2014, Widenfalk et al. 2015), as disturbance functions as environmental filter selecting for specific traits that exhibit phylogenetic signal. In contrast, forests, representing stable habitats little disturbed by agricultural activities, allow complex soil food webs to be established (Scheu and Falca 2000, Digel et al. 2014), where different phylogenetic clades of Collembola occupy various niches and thus exhibit phylogenetic overdispersion with different traits in local communities. Since community phylogenetic analyses are sensitive to the spatial and taxonomical scales of the study (Cavender-Bares et al. 2006, Swenson et al. 2006), we further confined our analyses by defining species pools using habitat types (arable fields, grasslands and forests), vertical stratification of species (surface-living, sub-surface-dwelling and soil-dwelling species) and different phylogenetic (taxonomic) groups.

Materials and Methods

Sampling

Collembola were sampled between March and June 2014 from arable fields, grasslands and forests at six sites near Göttingen, Germany (**Figure 2.1**, **Table S2.1**, **Chapter 2**). In each of the arable fields and grasslands, one suction sample equal to a surface area of 154 cm^2 was taken for 10 sec to collect surface-living individuals. Then, to sample soil-dwelling individuals, a soil core sample (5 cm diameter, 5 cm depth) was taken at the center of the area from which the suction sample was collected. In each forest, both litter and fragmented litter were collected by hand in an area of 154 cm^2, followed by a 10 sec suction sample of the humus layer. This suction sample was added to the litter collection producing a full litter-sample. Thereafter, a soil core sample was taken at the center of the same area previously sampled by hand-collection and suction. Collembola from the suction samples from arable fields and grasslands were directly transferred into 96% ethanol, while those in the litter and soil cores were extracted by heat using a Kempson extractor (Kempson et al. 1963), collected in water and then transferred into 96% ethanol every two days over a period of ten days. Samples were kept at 4°C until identification and then stored at -80°C. Collembola identification was based on Hopkin (2007), Fjellberg (1998, 2007) and Gisin (1960). and the nomenclature followed Bellinger et al. (1996–2017; www.collembola.org). The density (individuals m^{-2}) of each species in the suction (litter) sample was added to that in the soil samples for the following analyses.

Phylogenetic and trait distances between species

Phylogenetic relatedness and trait similarities of between species were calculated based on species pairwise distance matrices referred to their phylogeny and traits, respectively. The phylogenetic distance matrix was obtained from the ultrametric phylogenetic tree (**Figure 2.3**, **Chapter 2**) using the function *cophenetic.phylo* implemented in the R package “ape” (Paradis et al. 2004). The trait distance matrix was calculated based on species characters in body shape and length, pigmentation, number of ommatidia, vertical stratification and reproductive mode, multiple traits that exhibited phylogenetic signal (**Chapter 2**). While body length was used as a continuous variable, body shape, pigmentation, number of ommatidia, vertical stratification and reproductive mode were coded by binary variables for each state of traits. Gower's distances were calculated to generate a trait distance matrix using the function *dist.ktab* implemented in the R package “ade4” (Dray and Dufour 2007).

Phylogenetic and trait structures of communities

To investigate habitat associations with species phylogeny and traits, phylogeny- and trait-weighted community matrices were generated using a fuzzy-weighting method (Pillar and Duarte, 2010). Community data were first standardized using the “Hellinger” transformation (Legendre and Gallagher 2001) by the function *decostand* implemented in the R package “vegan” (Oksanen et al. 2015) and then weighted by species phylogenetic or trait distance metrics using the function *matrix.p* implemented in the R package “SYNCSA” (Debastiani and Pillar 2012). Differences in community phylogeny- or trait-weighted compositions were calculated based on Bray-Curtis dissimilarity between communities, followed by distance-based multivariate analysis of variance (ADONIS) with 9,999 permutations using the function *adonis* implemented in the “vegan” package, to test habitat effects on community phylogeny- or trait-weighted compositions. If habitats had significant effects on community phylogeny- or trait-weighted compositions, the weighted community compositions were applied to principle coordinate analysis (PCoA) using the function *pcps* implemented in the R package “PCPS” (Debastiani and Duarte 2014). This resulted in principal coordinates of phylogenetic structures (PCPS) and trait structures (PCTS) of the metacommunity. Pearson correlation coefficients of site scores and species densities at each PCPS axis were calculated using the function *cor.test*. *P*-values of correlation tests were adjusted using

Benjamini and Hochberg corrections (BH; Benjamini and Hochberg 1995). Habitat effects on site scores of PCPS axes were further tested using Kruskal-Wallis rank sum test followed by Dunn's test with BH corrections for multiple comparisons.

Phylogenetic relatedness and trait similarities of coexisting species

To represent overall species relatedness of a local community, the Net Relatedness Index (NRI) was calculated based on mean pairwise distance (MPD) of a local community (Webb 2000). The NRI is a negative value of standardized effect size of MPD, for which the observed MPD was standardized by 999 simulated MPDs generated by drawing species with equal probability from the phylogenetic distance matrix ("phylogeny.pool" null model) representing the species pool, in which only species occurring in all the local communities were retained. A positive NRI indicates that the species in the observed community are more closely related than expected based on the null communities. A negative NRI, on the contrary, indicates that species within a local community were more distantly related to each other than expected. NRI was calculated using the function *ses.mpd* implemented in the R package "picante" (Kembel et al. 2010) for phylogenetic distance matrix or trait distance matrix using species presence/absence or density data.

Since considering different definitions of species pools may help to inspect assembly processes at different spatial or taxonomic levels (Emerson and Gillepsie 2008, Lessard et al. 2012), NRI was calculated based on the following definitions of species pools: (1) a total species pool considering all species recorded in the study (i.e., entire phylogenetic tree including all species); (2) habitat-specific species pools based on species occurrence in a certain habitat type: arable fields (37 species), grasslands (43 species) and forests (52 species; **Figure S4.1**). A significant pattern found in local communities after constraining the randomization process within a certain habitat-specific pool indicates that other factors beside the broad-defined habitat type influence community assembly at the local scale. These factors may relate to the microhabitat conditions, e.g. soil moisture, food resources, etc.; (3) soil horizontal species pools according to the data on vertical stratification of different species: epedaphic (surface-living, 19 species), hemiedaphic (sub-surface-dwelling, 30 species) and euedaphic (soil dwelling, 26 species; **Figure S4.2**). Species vertical stratification is relevant to the scale at which individuals likely interact; (4) different taxonomic (phylogenetic) scales at higher taxonomic levels including Entomobryomorpha (34 species), Poduromorpha (20 species) and Symphypleona (18 species), and at lower taxonomic levels for Entomobryoidea (14 species) and Isotomidae (17 species; **Figure S4.3**). Analyses across different taxonomic scales can help to inspect how evolution may have influenced contemporary species coexistence (Silver et al. 2012, Tanaka and Sato 2015), and assembly processes may differ between phylogenetic lineages (Ndiribe et al. 2013, Elliott et al. 2016). For each of the defined species pools, only the species belonging to the pool were retained in the community dataset as well as in the phylogenetic and trait distance matrices.

Since the aim of the study was to infer assembly processes of Collembola in different habitats, the six sites were treated as replicates, and phylogenetic and trait NRI in each habitat was tested using Wilcoxon signed rank test against “0”, which indicated a randomly assembled community. Kruskal-Wallis rank sum test was applied to compare NRI values among habitats.

Results

Phylogenetic and trait structures between communities

In total, 6,323 Collembola individuals were collected, with 75 morphologically defined species. Phylogeny-weighted community structures significantly differed among habitats ($R^2 = 0.321$, $P = 0.006$, ADONIS). However, trait-weighted community structures were not different among habitats ($R^2 = 0.292$, $P = 0.051$, ADONIS). Communities from forests were associated with Neelipleona and the clade comprising Poduromorpha and Symphypleona, with higher scores at PCPS 1 than that in grasslands and arable fields ($P = 0.019$, Kruskal-Wallis rank sum test), while at PCPS 2 communities in arable fields were associated with species of Isotomidae (**Figure 4.1**).

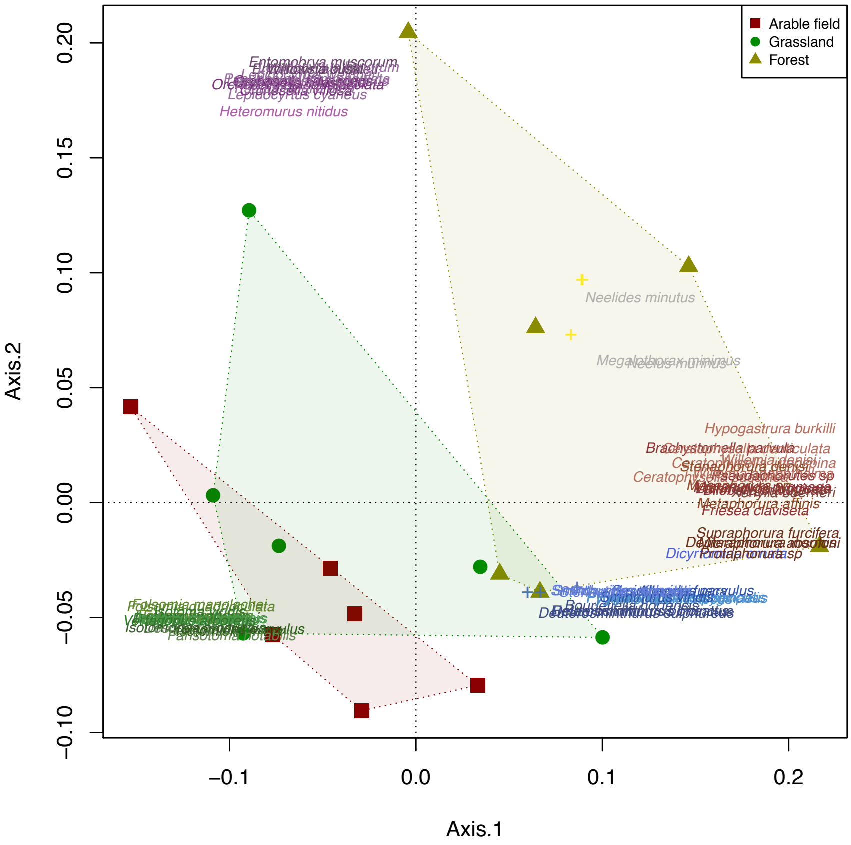

Figure 4.1
Scatter diagram of phylogenetic-weighted structure of Collembola communities (PCPS). Polygons encompass replicates of the same habitats. Only Collembola species significantly correlated to the axes are plotted with names and the others with cross symbols.

Phylogenetic relatedness and trait similarities of coexisting species

When all the species were used as species pool, co-occurring Collembola showed phylogenetic overdispersion in forests (phylogenetic NRI = -0.82 ± 0.21, mean ± SE; $P = 0.031$, Wilcoxon test). Mean phylogenetic relatedness of species inhabiting forests was significantly lower than those inhabiting arable fields and grasslands ($P = 0.019$, Kruskal-Wallis test). Accounting for species density, communities in arable fields showed phylogenetic clustering (phylogenetic NRI = 1.68 ± 0.37; $P = 0.031$, Wilcoxon test), indicating that abundant species in arable fields were closely related phylogenetically. Trait similarities in local communities, however, were not significantly different from random patterns in either of the three habitats, irrespective of using species presence/absence or density data (**Table 4.1**).

When the species pool was defined according to habitat types, similar patterns were detected to those using the entire species pool. In contrast, when the species pool was defined by vertical stratification of species, local communities exhibited different patterns of phylogenetic relatedness and trait similarity among different habitats. For epedaphic species, abundant species in grasslands and forests showed phylogenetic clustering, while forest species exhibited similar traits irrespective of whether species presence/absence or density data was considered. For hemiedaphic species, abundant species in arable fields and grasslands showed phylogenetic clustering, while species present in forests exhibited phylogenetic overdispersion and similar traits. Phylogenetic relatedness of the occurrence of hemiedaphic species was low in forests, intermediate in arable fields and high in grasslands. For euedaphic species, community phylogenetic relatedness and trait similarity in each of the three habitats were not different from random communities (**Table 4.1**).

At different taxonomic scales (phylogenetic clades), Collembola communities differed in their phylogenetic relatedness and trait similarities among the three habitats. At the order level, Entomobryomorpha, including Entomobryoidea, Tomoceridae and Isotomidae, showed similar patterns as compared to those using the entire species pool. When species densities were considered, weak but significantly similar traits were found in Symphypleona inhabiting arable fields, while phylogenetic overdispersion was detected in Poduromorpha living in forests. At the family level, Entomobryoidea in grasslands were from close relatives when species densities were considered, but displayed various traits when presence/absence data was used. In forests Entomobryoidea were phylogenetically clustered and carried similar traits when densities were considered. In contrast, Isotomidae in arable fields and grasslands showed phylogenetic overdispersion, while in forests they were phylogenetically closely related when densities were considered (**Table 4.1**, **Figure S4.3**).

Discussion

In this study we explored and compared phylogenetic and trait patterns of Collembola assemblages among arable fields, grasslands and forests, and inferred potential mechanisms driving Collembola

Table 4.1

Phylogenetic and trait structures represented by Net Relatedness Index (NRI; mean ± standard error), based on species occurrence (presence/absence, p/a) or density (abundance) of Collembola assemblages collected in different types of habitats. Different species pool definitions and taxonomic scales were applied. Numbers in parentheses indicate number of species in the species pool. A mean NRI value significantly different from zero (*P* < 0.05, Wilcoxon signed rank test) is marked in bold, with color in red indicating clustering and in blue overdispersion. Asterisks indicate NRI values significant different among habitats (*P* < 0.05, Kruskal-Wallis rank sum test).

Species pool definition	**Habitat**	**Phylogenetic NRI**		**Trait NRI**	
		p/a	**Abundance**	**p/a**	**Abundance**
All species pool (75)	Arable	-0.04 ± 0.27 *	**1.67 ± 0.37**	0.74 ± 0.51	0.80 ± 0.34
	Grassland	0.76 ± 0.40 *	1.13 ± 0.51	1.20 ± 1.04	0.94 ± 0.84
	Forest	**-0.82 ± 0.21** *	0.15 ± 0.38	-0.18 ± 0.29	0.22 ± 0.22
Habitat pool	Arable (37)	-0.26 ± 0.27 *	**1.42 ± 0.34**	0.17 ± 0.57	0.61 ± 0.34
	Grassland (43)	1.01 ± 0.47 *	1.14 ± 0.47	0.12 ± 0.87	0.44 ± 0.69
	Forest (52)	**-0.84 ± 0.22** *	0.17 ± 0.38	0.42 ± 0.37	0.41 ± 0.24
Vertical stratification pool					
Epedaphic (19)	Arable	0.03 ± 0.39	0.51 ± 0.54	-0.19 ± 0.52	-0.41 ± 0.40 *
	Grassland	0.72 ± 0.61	**1.09 ± 0.25**	-0.07 ± 0.39	-0.75 ± 0.48 *
	Forest	0.85 ± 0.59	**1.40 ± 0.30**	**1.67 ± 0.52**	**1.88 ± 0.17** *
Hemiedaphic (30)	Arable	1.17 ± 0.40 *	**1.60 ± 0.36**	-0.05 ± 0.40	-0.43 ± 0.32
	Grassland	**1.47 ± 0.47** *	**1.64 ± 0.38**	-0.36 ± 0.62	-0.36 ± 0.55
	Forest	**-0.88 ± 0.18** *	0.57 ± 0.50	**-1.32 ± 0.13**	-0.82 ± 0.34
Euedaphic (26)	Arable	-0.51 ± 0.19	-0.52 ± 0.27	-0.18 ± 0.31	-0.07 ± 0.28
	Grassland	-0.62 ± 0.19	-0.55 ± 0.26	0.02 ± 0.24	0.00 ± 0.12
	Forest	0.28 ± 0.39	-0.18 ± 0.26	0.30 ± 0.34	0.71 ± 0.35
Taxonomic scale					
Order level					
Entomobryomorpha (34)	Arable	0.07 ± 0.30	**0.80 ± 0.13**	-0.40 ± 0.35	0.19 ± 0.31
	Grassland	0.45 ± 0.36	0.67 ± 0.29	-0.27 ± 0.43	0.38 ± 0.41
	Forest	**-0.48 ± 0.17**	0.12 ± 0.24	0.32 ± 0.53	-0.15 ± 0.27
Symphypleona (18)	Arable	-0.08 ± 0.71	-0.06 ± 0.76	0.80 ± 0.35	**0.63 ± 0.26**
	Grassland	-0.35 ± 0.71	0.03 ± 0.68	1.23 ± 0.39	1.24 ± 0.19
	Forest	-0.38 ± 0.22	-0.33 ± 0.26	0.28 ± 0.53	0.48 ± 0.48
Poduromorpha (20)	Arable	-0.24 ± 0.50	-0.16 ± 0.39	0.56 ± 0.72	0.64 ± 0.57
	Grassland	-0.19 ± 0.27	-0.32 ± 0.34	-0.66 ± 0.13	-0.60 ± 0.25
	Forest	-0.50 ± 0.27	**-0.57 ± 0.12**	-0.06 ± 0.32	0.17 ± 0.22
Lower levels					
Entomobryoidea (14)	Arable	1.03 ± 0.97	1.05 ± 0.83	-0.66 ± 0.35 *	-0.61 ± 0.30 *
	Grassland	1.66 ± 0.75	**1.58 ± 0.31**	**-0.68 ± 0.13** *	-0.48 ± 0.32 *
	Forest	0.25 ± 0.37	**1.40 ± 0.11**	1.68 ± 0.61 *	**1.67 ± 0.17** *
Isotomidae (17)	Arable	**-0.52 ± 0.07** *	-0.46 ± 0.21 *	0.47 ± 0.62	0.40 ± 0.33
	Grassland	**-0.46 ± 0.15** *	-0.21 ± 0.21 *	0.81 ± 0.69	0.96 ± 0.50
	Forest	0.72 ± 0.31 *	**0.35 ± 0.06** *	-0.33 ± 0.24	0.03 ± 0.22

assemblages using community phylogenetic approaches. We found strong phylogeny-habitat associations, with Isotomidae generally associated with arable fields and Neelipleona, Poduromorpha and Symphypleona with forests. Our results generally support the hypothesis that Collembola communities in disturbed habitats such as arable fields are mainly driven by environmental filtering and those in stable habitats such as forests are based on competitive interactions or niche partitioning.

Phylogeny-habitat associations

Phylogeny-weighted community structures of Collembola separated well based on the habitats that associated strongly with phylogenetic lineages. In the phylogenetic tree, Isotomidae and Entomobryoidea derived earlier than the other Collembola groups. Their associations with disturbed environments, here represented by arable fields and grasslands with fluctuating temperature and humidity, likely reflected their early terrestrialization history and exposure to harsh environments. Development of forest habitats later on was likely accompanied by the diversification of other Collembola groups such as Symphypleona and Poduromorpha. This indicates that different phylogenetic groups remained associated with certain types of habitats. However, as pointed out in **Chapter 2**, the phylogenetic relationships between Collembola orders remain to be resolved.

Since disturbance regimes differed between arable fields and forests, phylogeny-habitat associations were likely mediated by disturbance-tolerant traits and related to moisture and habitat preferences of species (Makkonen et al. 2011, Widenfalk et al. 2015). Presumably, species with close phylogenetic affinities share similar physiological traits reflecting adaptation to disturbance regimes. However, to prove this idea, physiological traits need to be measured directly and the phylogenetic signal of the traits needs to be tested.

Community assembly in different habitats

The low phylogenetic relatedness in disturbed habitats, represented by the arable fields in this study, is consistent with the low relatedness in other communities at high altitudes or in disturbed habitats (Ding et al. 2012, Pellissier et al. 2013, Gianuca et al. 2014). Harsh environments likely select for species with similar traits e.g. those for coping with stress (Widenfalk et al. 2015), and these traits are likely to exhibit phylogenetic signal (Silvertown et al. 2006). Although all the traits analyzed in this study exhibited phylogenetic signal (**Chapter 2**), phylogenetic clustering in communities inhabiting arable fields did not translate into trait similarity, i.e., not different from randomness, suggesting that while some of the examined traits were filtered by disturbance in arable fields, other traits included in this study likely differed among coexisting species and were presumably driven by niche partitioning. Traits exhibiting phylogenetic signal but not included in this study, e.g. those related to drought and heat tolerance (Dias et al. 2013, Chen et al. 2017, **Chapter 3**), might be more relevant to the environmental filtering process.

In contrast, Collembola communities in forests were phylogenetically overdispersed, suggesting that assemblages of Collembola in stable habitats are based on competitive interactions, as argued in an earlier trait-based study on Collembola (Widenfalk et al. 2016). While little is known on the role of interspecific competition in community assembly of Collembola (Caruso et al. 2013), there is evidence that predator-prey interactions may regulate Collembola communities (Schneider and Maraun 2009, Birkhofer et al. 2010, Caruso et al. 2013). We therefore suggest that phylogenetic overdispersion and trait dissimilarity reflect resource-based niche partitioning. Stable habitats such as forests allow the establishment of complex soil food webs that include several trophic levels from primary decomposers to predators (Scheu and Falca 2000, Digel et al. 2014). Collembola species in forests span several trophic levels (Chahartaghi et al. 2005, Hishi et al. 2007, Pollierer et al. 2009, Hyodo et al. 2010) with taxonomically related species occupying similar trophic niches as represented by stable isotope signatures (Potapov et al. 2016). Phylogenetic overdispersion in forests therefore likely resulted from the presence of species affiliated to a variety of phylogenetic clades. This variety allows the community to cover a wide range of trophic niches and thereby completely exploit the available resources. Notably, when the species pool was defined based on the forest habitat, Collembola communities were still phylogenetically overdispersed in forests, suggesting that factors other than those associated with the broadly defined habitat, presumably local micro-habitat characteristics or food resources, influenced community assembly. Spatial heterogeneity across vertical soil layers in forests likely resulted in phylogenetic overdispersion (Berg et al. 1998, Berg and Bengtsson 2007). Sampling at a finer vertical spatial scale is needed to prove this idea.

Other scenarios also likely resulted in phylogenetic overdispersion in Collembola communities inhabiting forests, such as environmental filtering for phylogenetically convergent traits, a pattern in which distantly related species share similar traits (Emerson and Gillespie 2008). The morphological traits considered in this study, however, all exhibited phylogenetic signal and presumably more related to abiotic factors, pointing to the need to consider other types of traits such as those related to resource exploitation or interspecific competition. For soil animals, integrating stable isotope signatures, neutral lipid fatty acid composition and molecular gut content and microbiome analyses will shed new light on their trophic niches (Ferlian and Scheu 2014, Heidemann et al. 2014, Ferlian et al. 2015, Potapov et al. 2016, Chen et al. 2017, Gong et al. *submitted*), and these need to be considered as functional traits and included in future phylogenetic and trait-based analyses.

Species pool definitions

When the species pool was defined based on vertical stratification of species, a spatial scale where species are more likely to interact, the abundant epedaphic species in forests exhibited phylogenetic clustering with similar traits, suggesting that environmental filtering predominantly worked on the traits. For example, two congeneric species, *Lepidocyrtus lanuginosus* and *L. lignorum*, shared a number of traits and always coexisted in forests. In contrast, hemiedaphic Collembola in forests presumably were driven by niche partitioning that likely resulted in phylogenetic overdispersion. Species assigned to this category typically migrate between different soil layers and possess various traits. In deep soil (euedaphic communities), in contrast, community assembly was likely driven by stochastic processes or a balance between environmental filtering and niche partitioning, as indicated by randomness in both phylogenetic relatedness and trait similarity. However, in this study we used data on vertical stratification compiled from literature without proving if this also applied to the study sites. Further studies at a finer spatial scale of the vertical community composition along soil profiles are needed to explore coexistence patterns at the scale relevant to species interactions (Maaß et al. 2015).

Using different species pools defined by phylogenetic (taxonomic) groups may help to infer mechanisms driving communities across different hierarchical taxonomic groups and between different phylogenetic groups, although species may interact with each other irrespective of their phylogenetic affinities. Species of the same phylogenetic clade usually possess similar traits. Controlling the similarity at a higher level and then exploring patterns at a lower level may therefore reveal the patterns and processes working at the lower level, likely improving the detection of overdispersion (Swenson et al. 2006, Cavender-Bares et al. 2006, Elliott et al. 2016). Analyses confined to single phylogenetic clade in this study suggest that community assembly processes vary with taxonomic levels and between phylogenetic lineages. At the order level, phylogenetic and trait patterns in Entomobryomorpha resembled those if total Collembola was used as species pool. In contrast, community assembly processes in Symphypleona and Poduromorpha differed from those of total Collembola, with trait-based environmental filtering being important for Symphypleona in arable fields and niche partitioning for Poduromorpha in forests. At lower taxonomic level, phylogenetic clustering and similar traits of Entomobryoidea inhabiting forests point to the importance of environmental filtering working on the abundant species. In grasslands co-occurring Entomobryoidea species possessed divergent traits, including pigmentation and numbers of ommatidia, suggesting that niche partitioning played a major role. However, Entomobryoidea in grasslands exhibited phylogenetic clustering when species densities were accounted, presumably resulting from environmental filtering that selected other unmeasured traits. Interestingly, community assembly processes in Isotomidae differed from those of other Collembola taxa. Isotomidae inhabiting arable fields and grasslands were likely structured via niche partitioning, while environmental filtering dominated in Isotomidae inhabiting forests, the latter being most evident in *Folsomia* species.

Integration of phylogenetic, trait and comparative approaches

The community phylogenetic framework was proposed by plant ecologists 15 years ago and thereafter proven by its ability to infer assembly processes (Webb et al. 2002, Cavender-Bares et al. 2009, Vamosi et al. 2009). Application of this method to soil biota, however, just started recently and so far is restricted to fungi (Bässler et al. 2014, Thorn et al. 2016), nematodes (Li et al. 2014), termites (Hausberger and Korb 2015, Hausberger and Korb 2016) and beetles (Andújar et al. 2015, Thorn et al. 2016). Our study provides an example of integrating community phylogenetic and trait-based approaches in studies on the assembly processes of one of the most diverse soil microarthropod groups, Collembola. The phylogenetic signal first needs to be tested using comparative methods for process-related traits, such as Collembola body shape, body length, pigmentation, number of ommatidia, reproductive mode and vertical stratification that reflect species associations with the habitats (Salmon et al. 2014, Malcicka et al. 2017; **Chapter 2**). Significant differences in community phylogenetic relatedness then provide evidence on assembly processes, as shown in Collembola inhabiting arable fields and forests exhibiting different phylogenetic relatedness. In this approach phylogenetic information is used as a surrogate for functional similarity (Kembel 2009, Mouquet et al. 2012, Cadotte et al. 2013). However, other traits such as those related to dietary resources may be phylogenetically labile (Chen et al. 2017, **Chapter 3**) and therefore phylogeny may be of limited value as a proxy for traits. If these labile traits are at work, trait values of individual species need to be measured and trait-based approaches instead of community phylogeny need to be adopted. Nevertheless, since species phylogeny and traits are not mutually exclusive but complemented each other (Cadotte et al. 2013), integrating information on both functional traits and phylogenetic relationships promises major progress in understanding of assembly processes not only for aboveground biota but also for belowground organisms.

Acknowledgements

We thank Charlotte Seifert for her help in site selection. Special thanks to Guido Humpert for support during lab work. This study was supported by the German Ministry of Education, Science and Technology (BMBF) and by the Uyttenbogaard Eliasen Society via a personal grant to MPB. The authors have no conflict of interest to declare.

Author Contributions

TWC and SS conceived and designed the study; TWC, JFC, VK and MPB performed the study and collected the data; TWC analyzed the data; TWC, MPB, IS and SS wrote the manuscript.

References

Adler, P.B., Fajardo, A., Kleinhesselink, A.R. & Kraft, N.J.B. (2013) Trait-based tests of coexistence mechanisms. *Ecology Letters*, **16**, 1294–1306.

Anderson, J.M. (1975) The enigma of soil animal species diversity. Progress in Soil Zoology: Proceedings of the 5th International Colloquium on Soil Zoology (ed J. Vaněk), pp. 51–58. Springer Netherlands, Dordrecht.

Andújar, C., Arribas, P., Ruzicka, F., Platt, A.C., Timmermans, M.J.T.N. & Vogler, A.P. (2015) Phylogenetic community ecology of soil biodiversity using mitochondrial metagenomics. *Molecular Ecology*, **24**, 3603–3617.

Bässler, C., Ernst, R., Cadotte, M., Heibl, C. & Müller, J. (2014) Near-to-nature logging influences fungal community assembly processes in a temperate forest (ed J Barlow). *Journal of Applied Ecology*, **51**, 939–948.

Bellinger, P.F., Christiansen, K.A. & Janssens, F. Checklist of the Collembola of the world (www.collembola.org).

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*, **57**, 289–300.

Berg, M.P. & Bengtsson, J. (2007) Temporal and spatial variability in soil food web structure. *Oikos*, **116**, 1789–1804.

Berg, M.P., Kniese, J.P., Bedaux, J.J.M. & Verhoef, H.A. (1998) Dynamics and stratification of functional groups of micro- and mesoarthropods in the organic layer of a Scots pine forest. *Biology and Fertility of Soils*, **26**, 268–284.

Birkhofer, K., Scheu, S. & Wiegand, T. (2010) Assessing spatiotemporal predator–prey patterns in heterogeneous habitats. *Basic and Applied Ecology*, **11**, 486–494.

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.

Cadotte, M., Albert, C.H. & Walker, S.C. (2013) The ecology of differences: assessing community assembly with trait and evolutionary distances. *Ecology Letters*, **16**, 1234–1244.

Caruso, T., Trokhymets, V., Bargagli, R. & Convey, P. (2013) Biotic interactions as a structuring force in soil communities: evidence from the micro-arthropods of an Antarctic moss model system. *Oecologia*, **172**, 495–503.

Cavender-Bares, J., Keen, A. & Miles, B. (2006) Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology*, **87**, S109-22.

Cavender-Bares, J., Kozak, K.H., Fine, P. V & Kembel, S.W. (2009) The merging of community ecology and phylogenetic biology. *Ecology Letters*, **12**, 693–715.

Chahartaghi, M., Langel, R., Scheu, S. & Ruess, L. (2005) Feeding guilds in Collembola based on nitrogen stable isotope ratios. *Soil Biology and Biochemistry*, **37**, 1718–1725.

Chase, J.M. & Leibold, M.A. (2003) Ecological Niches: Linking Classical and Contemporary Approaches. University of Chicago Press, Chicago, IL.

Chen, T.-W., Sandmann, P., Schaefer, I. & Scheu, S. (2017) Neutral lipid fatty acid composition as trait and constraint in Collembola evolution. *Ecology and Evolution*, **7**, 9624-9638.

Debastiani, V.J. & Duarte, L.D.S. (2014) PCPS - an R-package for exploring phylogenetic eigenvectors across metacommunities. *Frontiers of Biogeography*, **6**, 144–148.

Debastiani, V.J. & Pillar, V.D. (2012) SYNCSA--R tool for analysis of metacommunities based on functional traits and phylogeny of the community components. *Bioinformatics*, **28**, 2067–8.

Dias, A.T.C., Krab, E.J., Mariën, J., Zimmer, M., Cornelissen, J.H.C., Ellers, J., Wardle, D.A. & Berg, M.P. (2013) Traits underpinning desiccation resistance explain distribution patterns of terrestrial isopods. *Oecologia*, **172**, 667–677.

Digel, C., Curtsdotter, A., Riede, J., Klarner, B. & Brose, U. (2014) Unravelling the complex structure of forest soil food webs: higher omnivory and more trophic levels. *Oikos*, **123**, 1157–1172.

Ding, Y., Zang, R., Letcher, S.G., Liu, S. & He, F. (2012) Disturbance regime changes the trait distribution, phylogenetic structure and community assembly of tropical rain forests. *Oikos*, **121**, 1263–1270.

Dray, S. & Dufour, A.-B. (2007) The ade4 Package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, **22**.

Duarte, L.D.S. (2011) Phylogenetic habitat filtering influences forest nucleation in grasslands. *Oikos*, **120**, 208–215.

Duarte, L.D.S., Debastiani, V.J., Freitas, A.V.L. & Pillar, V.D. (2016) Dissecting phylogenetic fuzzy weighting: theory and application in metacommunity phylogenetics. *Methods in Ecology and Evolution*, **7**, 937–946.

Elliott, T.L., Waterway, M.J. & Davies, T.J. (2016) Contrasting lineage-specific patterns conceal community phylogenetic structure in larger clades. *Journal of Vegetation Science*, **27**, 69–79.

Emerson, B.C. & Gillespie, R.G. (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology & Evolution*, **23**, 619–630.

Ferlian, O., Klarner, B., Langeneckert, A.E. & Scheu, S. (2015) Trophic niche differentiation and utilisation of food resources in collembolans based on complementary analyses of fatty acids and stable isotopes. *Soil Biology and Biochemistry*, **82**, 28–35.

Ferlian, O. & Scheu, S. (2014) Shifts in trophic interactions with forest type in soil generalist predators as indicated by complementary analyses of fatty acids and stable isotopes. *Oikos*, **123**, 1182–1191.

Fjellberg, A. (1998) The Collembola of Fennoscandia and Denmark Part I: Poduromorpha (eds NP Kristensen and V Michelsen). Koninklijke Brill, Leiden. The Netherlands.

Fjellberg, A. (2007) The Collembola of Fennoscandia and Denmark Part II: Entomobryomorpha and Symphypleona (eds NP Kristensen and V Michelsen). Koninklijke Brill, Leiden. The Netherlands.

Gerhold, P., Cahill, J.F., Winter, M., Bartish, I. V & Prinzing, A. (2015) Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Functional Ecology*, **29**, 600–614.

Gianuca, A.T., Dias, R.A., Debastiani, V.J. & Duarte, L.D.S. (2014) Habitat filtering influences the phylogenetic structure of avian communities across a coastal gradient in southern Brazil. *Austral Ecology*, **39**, 29–38.

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

Gisin, H. (1960) Collembolenfauna Europas. Muséum d'Histoire Naturelle, Geneva.

Gong, X., Chen, T.-W., Zieger, S., Bluhm, C., Heidemann, K., Schaefer, I., Maraun, M., Liu, M. & Scheu, S. Phylogenetic and trophic determinants of gut microbiota in soil detritivores. *Submitted.*

Hausberger, B. & Korb, J. (2015) A phylogenetic community approach for studying termite communities in a West African savannah. *Biology Letters*, **11**, 20150625.

Hausberger, B. & Korb, J. (2016) The impact of anthropogenic disturbance on assembly patterns of termite communities. *Biotropica*, **48**, 356–364.

Heidemann, K., Hennies, A., Schakowske, J., Blumenberg, L., Ruess, L., Scheu, S. & Maraun, M. (2014) Free-living nematodes as prey for higher trophic levels of forest soil food webs. *Oikos*, **123,** 1199–1211.

Heiniger, C., Barot, S., Ponge, J.-F., Salmon, S., Botton-Divet, L., Carmignac, D. & Dubs, F. (2014) Effect of habitat spatiotemporal structure on collembolan diversity. *Pedobiologia*, **57**, 103–117.

Hishi, T., Hyodo, F., Saitoh, S. & Takeda, H. (2007) The feeding habits of collembola along decomposition gradients using stable carbon and nitrogen isotope analyses. *Soil Biology and Biochemistry*, **39**, 1820–1823.

Hopkin, S.P. (2007) A Key to the Collembola (Springtails) of Britain and Ireland (ed Field Studies Council (FSC)). Shrewsbury.

Hubbell, S.P. (2001) The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press, Princeton and Oxford.

Hyodo, F., Matsumoto, T., Takematsu, Y., Kamoi, T., Fukuda, D., Nakagawa, M. & Itioka, T. (2010) The structure of a food web in a tropical rain forest in Malaysia based on carbon and nitrogen stable isotope ratios. *Journal of Tropical Ecology*, **26**, 205–214.

Kembel, S.W. (2009) Disentangling niche and neutral influences on community assembly: assessing the performance of community phylogenetic structure tests. *Ecology letters*, **12**, 949–60.

Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P. & Webb, C.O. (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, **26**, 1463–4.

Kempson, D., Lloyd, M. & Ghelardi, R. (1963) A new extractor for woodland litter. *Pedobiologia*, **3**, 1–21.

Legendre, P. & Gallagher, E.D. (2001) Ecologically meaningful transformations for ordination of species data. *Oecologia*, **129**, 271–280.

Lessard, J.-P., Belmaker, J., Myers, J.A., Chase, J.M. & Rahbek, C. (2012) Inferring local ecological processes amid species pool influences. *Trends in Ecology & Evolution*, **27**, 600–607.

Li, J., Li, S., Chen, Y., Jia, P., Hua, Z., Wang, S., Song, Y., Liao, B. & Shu, W. (2014) Phylogenetic structures of soil nematode communities along a successional gradient in an unreclaimed copper mine tailings site. *Soil Biology and Biochemistry*, **77**, 179–186.

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

Makkonen, M., Berg, M.P., van Hal, J.R., Callaghan, T. V., Press, M.C. & Aerts, R. (2011) Traits explain the responses of a sub-arctic Collembola community to climate manipulation. *Soil Biology and Biochemistry*, **43**, 377–384.

Malcicka, M., Berg, M.P. & Ellers, J. (2017) Ecomorphological adaptations in Collembola in relation to feeding strategies and microhabitat. *European Journal of Soil Biology*, **78**, 82–91.

Martins da Silva, P., Berg, M.P., Serrano, A.R.M., Dubs, F. & Sousa, J.P. (2012) Environmental factors at different spatial scales governing soil fauna community patterns in fragmented forests. *Landscape Ecology*, **27**, 1337–1349.

McGill, B.J., Enquist, B.J., Weiher, E. & Westoby, M. (2006) Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, **21**, 178–185.

Moretti, M., Dias, A.T.C., de Bello, F., Altermatt, F., Chown, S.L., Azcarate, F.M., Bell, J.R., Fournier, B., Hedde, M., Hortal, J., Ibanez, S., Öckinger, E., Sousa, J.P., Ellers, J. & Berg, M.P. (2017) Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. *Functional Ecology*, **31**, 558–567.

Mouquet, N., Devictor, V., Meynard, C.N., Munoz, F., Bersier, L.F., Chave, J., Couteron, P., Dalecky, A., Fontaine, C., Gravel, D., Hardy, O.J., Jabot, F., Lavergne, S., Leibold, M.A., Mouillot, D., Munkemuller, T., Pavoine, S., Prinzing, A., Rodrigues, A.S., Rohr, R.P., Thebault, E. & Thuiller, W. (2012) Ecophylogenetics: advances and perspectives. *Biological Reviews*, **87**, 769–785.

Ndiribe, C., Pellissier, L., Antonelli, S., Dubuis, A., Pottier, J., Vittoz, P., Guisan, A. & Salamin, N. (2013) Phylogenetic plant community structure along elevation is lineage specific. *Ecology and Evolution*, **3**, 4925–4939.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H. & Wagner, H. (2015) vegan: community ecology package.

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

Pellissier, L., Pradervand, J.-N., Williams, P.H., Litsios, G., Cherix, D. & Guisan, A. (2013) Phylogenetic relatedness and proboscis length contribute to structuring bumblebee communities in the extremes of abiotic and biotic gradients (ed A Baselga). *Global Ecology and Biogeography*, **22**, 577–585.

Petersen, H. & Luxton, M. (1982) A comparative analysis of soil fauna populations and their role in decomposition processes. *Oikos*, **39**, 288–388.

Pey, B., Nahmani, J., Auclerc, A., Capowiez, Y., Cluzeau, D., Cortet, J., Decaëns, T., Deharveng, L., Dubs, F., Joimel, S., Briard, C., Grumiaux, F., Laporte, M.-A.A., Pasquet, A., Pelosi, C., Pernin, C., Ponge, J.-F., Salmon, S., Santorufo, L. & Hedde, M. (2014) Current use of and future needs for soil invertebrate functional traits in community ecology. *Basic and Applied Ecology*, **15**, 194–206.

Pillar, V.D. & Duarte, L.D.S. (2010) A framework for metacommunity analysis of phylogenetic structure. *Ecology Letters*, **13**, 587–596.

Pollierer, M.M., Langel, R., Scheu, S. & Maraun, M. (2009) Compartmentalization of the soil animal food web as indicated by dual analysis of stable isotope ratios ($^{15}N/^{14}N$ and $^{13}C/^{12}C$). *Soil Biology and Biochemistry*, **41**, 1221–1226.

Pollierer, M.M. & Scheu, S. (2017) Driving factors and temporal fluctuation of Collembola communities and reproductive mode across forest types and regions. *Ecology and Evolution*, ece3.3035.

Ponge, J.-F., Gillet, S., Dubs, F., Fedoroff, E., Haese, L., Sousa, J.P. & Lavelle, P. (2003) Collembolan communities as bioindicators of land use intensification. *Soil Biology and Biochemistry*, **35**, 813–826.

Potapov, A.A., Semenina, E.E., Korotkevich, A.Y., Kuznetsova, N.A. & Tiunov, A. V. (2016) Connecting taxonomy and ecology: Trophic niches of collembolans as related to taxonomic identity and life forms. *Soil Biology and Biochemistry*, **101**, 20–31.

Ricklefs, R.E. (1987) Community diversity: Relative roles of local and regional processes. *Science*, **235**, 167–171.

Rusek, J. (1998) Biodiversity of Collembola and their functional role in the ecosystem. *Biodiversity and Conservation*, **7**, 1207–1219.

Salmon, S. & Ponge, J.-F. (2012) Species traits and habitats in springtail communities: A regional scale study. *Pedobiologia*, **55**, 295–301.

Salmon, S., Ponge, J.-F., Gachet, S., Deharveng, L., Lefebvre, N. & Delabrosse, F. (2014) Linking species, traits and habitat characteristics of Collembola at European scale. *Soil Biology and Biochemistry*, **75**, 73–85.

Schaefer, M., Migge-Kleian, S. & Scheu, S. (2009) The role of soil fauna for decomposition of plant residues. *Functioning and Management of European Beech Ecosystems* (eds R. Brumme & P.K. Khanna), pp. 207–230. Springer Berlin Heidelberg, Berlin, Heidelberg.

Scheu, S. & Falca, M. (2000) The soil food web of two beech forests (*Fagus sylvatica*) of contrasting humus type: stable isotope analysis of a macro- and a mesofauna-dominated community. *Oecologia*, **123**, 285–296.

Schneider, K. & Maraun, M. (2009) Top-down control of soil microarthropods - Evidence from a laboratory experiment. *Soil Biology and Biochemistry*, **41**, 170–175.

Silver, C.A., Vamosi, S.M. & Bayley, S.E. (2012) Temporary and permanent wetland macroinvertebrate communities: Phylogenetic structure through time. *Acta Oecologica*, **39**, 1–10.

Silvertown, J., McConway, K., Gowing, D., Dodd, M., Fay, M.F., Joseph, J.A. & Dolphin, K. (2006) Absence of phylogenetic signal in the niche structure of meadow plant communities. *Proceedings of the Royal Society B*, **273**, 39–44.

Sousa, J.P., Bolger, T., da Gama, M.M., Lukkari, T., Ponge, J.-F., Simón, C., Traser, G., Vanbergen, A.J., Brennan, A., Dubs, F., Ivitis, E., Keating, A., Stofer, S. & Watt, A.D. (2006) Changes in Collembola richness and diversity along a gradient of land-use intensity: A pan European study. *Pedobiologia*, **50**, 147–156.

Swenson, N.G., Enquist, B.J., Pither, J., Thompson, J. & Zimmerman, J.K. (2006) The problem and promise of scale dependency in community phylogenetics. *Ecology*, **87**, 2418–2424.

Tanaka, T. & Sato, T. (2015) Taxonomic, phylogenetic and functional diversities of ferns and lycophytes along an elevational gradient depend on taxonomic scales. *Plant Ecology*, **216**, 1597–1609.

Thorn, S., Bässler, C., Bernhardt-Römermann, M., Cadotte, M., Heibl, C., Schäfer, H., Seibold, S. & Müller, J. (2016) Changes in the dominant assembly mechanism drives species loss caused by declining resources. *Ecology Letters*, **19**, 163–170.

Vamosi, S.M., Heard, S.B., Vamosi, J.C. & Webb, C.O. (2009) Emerging patterns in the comparative analysis of phylogenetic community structure. *Molecular Ecology*, **18**, 572–592.

Vandewalle, M., de Bello, F., Berg, M.P., Bolger, T., Dolédec, S., Dubs, F., Feld, C.K., Harrington, R., Harrison, P. A., Lavorel, S., Silva, P.M., Moretti, M., Niemelä, J., Santos, P., Sattler, T., Sousa, J.P., Sykes, M.T., Vanbergen, A.J. & Woodcock, B.A. (2010) Functional traits as indicators of biodiversity response to land use changes across ecosystems and organisms. *Biodiversity and Conservation*, **19**, 2921–2947.

Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. (2007) Let the concept of trait be functional! *Oikos*, **116**, 882–892.

Webb, C.O. (2000) Exploring the Phylogenetic Structure of Ecological Communities: An Example for Rain Forest Trees. *The American Naturalist*, **156**, 145–155.

Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, **33**, 475–505.

Widenfalk, L.A., Bengtsson, J., Berggren, Å., Zwiggelaar, K., Spijkman, E., Huyer-Brugman, F. & Berg, M.P. (2015) Spatially structured environmental filtering of collembolan traits in late successional salt marsh vegetation. *Oecologia*, **179**, 537–549.

Widenfalk, L.A., Malmström, A., Berg, M.P. & Bengtsson, J. (2016) Small-scale Collembola community composition in a pine forest soil - Overdispersion in functional traits indicates the importance of species interactions. *Soil Biology and Biochemistry*, **103**, 52–62.

Supplementary Materials

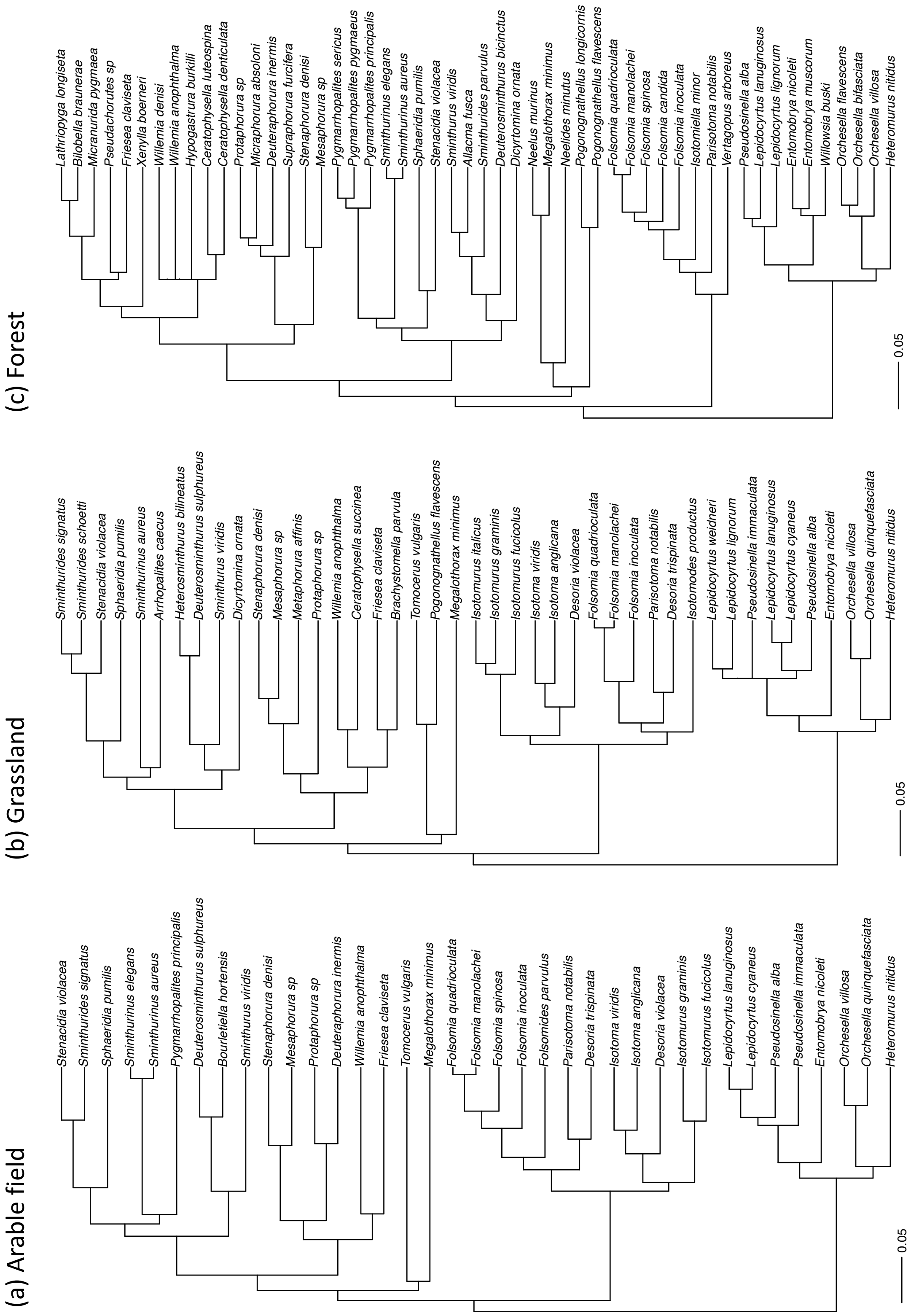

Figure S4.1
Subsets of phylogenetic trees used as species pools based on the three habitats: arable fields, grasslands and forests.

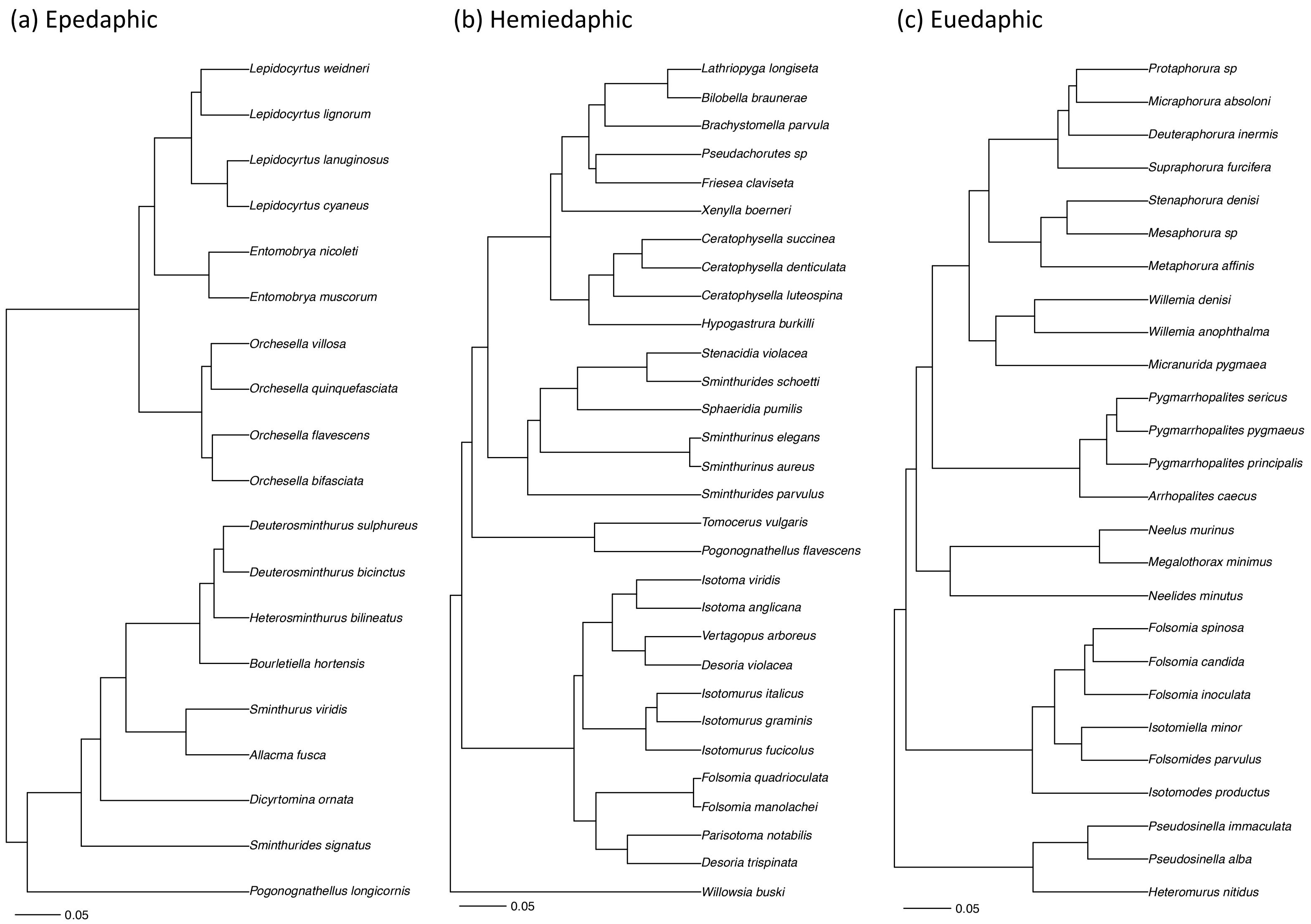

Figure S4.2
Subsets of phylogenetic trees used as species pools based on species vertical stratification: epedaphic, hemiedaphic and euedaphic.

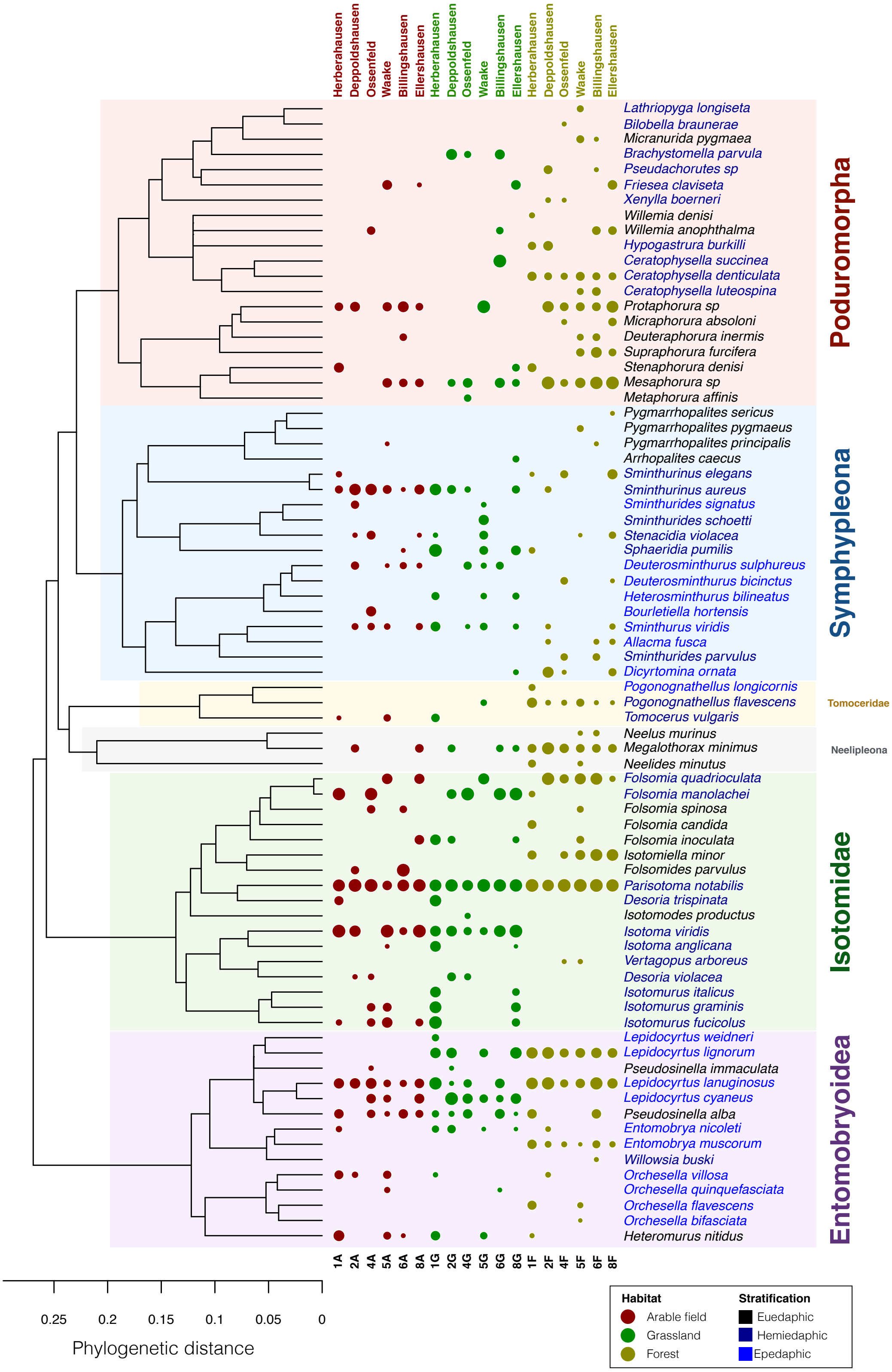

Figure S4.3
Densities of Collembola collected from the study sites plotted against the ultrametric phylogenetic tree. Size of circles represents standardized relative density of a species in a given community. Color of circles represent habitat type a species was sampled. Colors in the species names indicate vertical stratification of species. See Figure 2.4 and Chapter 2 for details on phylogenetic reconstruction of Collembola.

Chapter 5

Inferring assembly processes of Collembola communities along successional trajectories using phylogenetic approaches

Ting-Wen Chen, Gerrit Moser, Apolline Auclerc, Jean-François Ponge, Sébastien Barot, Florence Dubs, Stefan Scheu

Abstract

Understanding ecological and evolutionary processes in community assembly can help to explain species coexistence of soil biota. Along successional trajectories, dispersal of species interacting with selection processes determines species composition in local communities. Intermediate dispersal with local environmental selection results in "species sorting" according to metacommunity theory, while frequent dispersal leads to "mass effect". As a consequence, the community may show a random pattern in early stages of succession but may follow deterministic pathways in later stages. In this study, we applied community phylogenetic approaches to a manipulative soil block experiment (Auclerc et al. 2009; Soil Biology and Biochemistry 41, 1596–1604), to infer factors influencing community assembly of soil springtails (Hexapoda: Collembola) during succession in forest and meadow soil. Soil blocks were defaunated and/or exchanged, and Collembola colonization from surrounding habitat into soil blocks was monitored one week, one month and six months after soil block reinstallation to the fields. We found phylogenetic signal of species preference to the soil types and of their dispersal ability in meadow but not in forest. Community assembly processes in meadow were predominated by niche partitioning, while in forest environmental filtering was the main driver. While environmental filtering continuously influenced Collembola community assembly in forest during the whole succession period, niche partitioning worked during Collembola succession in meadow and was significant at the later successional stage. Our results also indicate that soil properties of the defaunated and transferred blocks influenced community assembly of Collembola at initial stages of succession, in line with the species sorting scenario of metacommunity theory, while at later stages community assembly was dominated by mass effect. Furthermore, indigenous Collembola species survived in forest soil blocks that were transferred to the meadow habitat. In contrast, while most indigenous meadow species transferred to forest did not survive, Isotomidae species survived in meadow soil blocks and likely also immigrated from the surrounding forest habitat.

Keywords

ancestral state estimation; community phylogeny; dispersal ability; experimental manipulation; habitat preference; phylogenetic signal; phylogeny; springtail; soil preference; succession; transfer experiment

Introduction

Composition of biotic community is influenced by both ecological and evolutionary processes (Vellend, 2010, 2016). In a metacommunity, local communities are connected via individual dispersal, interacting with habitat selection processes such as environmental filtering and species interactions that together drive local community composition (Leibold et al. 2004). In local communities, species possess a number of traits reflecting both habitat selection and dispersal processes. Traits associated with habitats, e.g. physiological attributes, may reflect environmental requirements of a species. Other traits such as morphological characters may directly link to species dispersal abilities in a habitat (Ponge et al. 2006, Auclerc et al. 2009, Pey et al. 2014). However, species traits not only are the characters which contemporary assembly processes are based on, but also reflect adaptation of species to the environment during evolutionary history, thus showing phylogenetic signal (Liu et al. 2015). Therefore, contemporary ecological processes, such as habitat selection and dispersal, work on (or relate to) species traits that have been shaped by past processes, and result in current patterns of community assembly.

Relative strengths of assembly processes may vary during colonization of species in new habitats, resulting in dynamic changes in community assembly patterns with time (Emerson and Gillespie 2008, Purschke et al. 2013, Li et al. 2014). At early successional stages, species arrive via dispersal from source habitats; species with high dispersal ability recolonize faster than poor dispersers (Emerson and Gillespie 2008). At these stages dispersal likely overwhelms the influence of local environments, resulting in the “mass effect” scenario, i.e., similar compositions of local communities. In later stages of succession, communities likely follow deterministic trajectories influenced by habitat selection, i.e., environmental filtering and biotic interactions (Purschke et al. 2013, Letten et al. 2014). Selection by the habitats results in the “species sorting” scenario, in which community compositions correlate with local abiotic and/or biotic factors.

To uncover the processes predominant during succession, community phylogeny offers an analytical framework, in which phylogenetic distances of species are used as a surrogate for the differences in traits, assuming that process-related traits exhibit phylogenetic signal (Cavender-Bares et al. 2009, Cadotte et al. 2013, de Bello et al. 2015). If traits are phylogenetically conserved (i.e., related species share similar traits), a community composed of phylogenetically closely related species suggests environmental filtering as major structuring force. In contrast, low relatedness between coexisting species suggests the dominance of biotic interactions, such as competition or niche partitioning (Webb et al. 2002, Cavender-Bares et al. 2009). Furthermore, a phylogeny-weighted community matrix (Pillar and Duarte, 2010) allows to measure phylogenetic dissimilarities between local communities. This matrix, combined with principle coordinate analysis, generates principal coordinates of phylogenetic structure of a metacommunity, allowing to identify phylogenetic clades associated with different habitats (Duarte 2011, Duarte et al. 2016).

Compared to studies on aboveground community succession, assembly processes of belowground community are little known, despite the tremendous biodiversity in soil ecosystems (Anderson 1975, Bardgett and van der Putten 2014). Collembola (springtails) are among the most abundant and diverse soil invertebrates which contribute to organic matter decomposition processes and nutrient cycling (Petersen and Luxton 1982, Rusek 1998, Schaefer et al. 2009). Distinct community composition of Collembola in different habitats suggests that assembly processes are related to local abiotic and biotic factors (Caruso et al. 2013, Salmon et al. 2014). Collembola communities also change along successional gradients (Dunger et al. 2004, Huebner et al. 2012, Perez et al. 2013), likely following the "mass effect" scenario via frequent dispersal from the source habitats (Ingimarsdóttir et al. 2012). After arrival in new habitats, however, local environmental conditions likely drive communities in certain directions. As indicated by a field observational study on Collembola assembly processes (**Chapter 4**) using community phylogenetic approaches, Collembola assemblages are likely driven by environmental filtering in disturbed environments, while niche partitioning is more prominent in a stable habitat. However, to further understand assembly processes of Collembola during succession, a combined approach of community phylogeny and manipulative experiments is needed.

The experiment of Auclerc et al. (2009) offers a model system to examine assembly processes during succession of soil animal community in contrasting habitats. In this study, the authors eradicated animals from meadow and forest soil blocks, and then transferred the soil blocks back to the original habitat or to new habitats (**Table 5.1**). Individual Collembola species were ascribed to different groups in respect of dispersal ability and habitat preferences based on monitoring species recolonization in the soil blocks. In the present study, we took a community perspective with a focus on assembly processes of Collembola recolonizing the soil blocks. We reanalyzed the data of Auclerc et al. (2009) using (meta)community phylogenetic approaches. We analyzed phylogenetic signal and ancestral states of dispersal abilities and environmental preferences of the species examined in Auclerc et al. (2009), to infer how the ecological preferences of Collembola and their dispersal abilities in habitats have evolved. Then, we inferred assembly processes of Collembola communities inhabiting meadow and forest and at different successional stages. We examined effects of soil origin on community succession patterns and assembly processes in defaunated blocks, as well as the fate of indigenous species after soil blocks were transferred to a new habitat.

We hypothesized that (1) community structures and phylogeny-weighted structures both differ between meadow and forest but not between sampling time; communities in meadow exhibit phylogenetic clustering due to environmental filtering resulting from disturbances, e.g. mowing and grazing, while forest communities show phylogenetic overdispersion resulting from niche partitioning in a stable environment. (2) At early successional stages drift predominates resulting in random patterns of phylogenetic relatedness in defaunated blocks, while selection drives communities to deterministic patterns at later successional stages. (3) Successional patterns in defaunated soil blocks differ between transferred soil blocks and those remained in the original habitat. (4) Community structures in untreated, transferred soil blocks change gradually from those installed in the original habitat to those in the transferred habitat (**Table S5.1**).

Materials and Methods

Collembola communities

We reanalyzed the data of Collembola community used in Auclerc et al. (2009). The original experiment was conducted between December 2005 and June 2006 in Morvan Regional Natural Park in Burgundy, France. Sixty soil blocks (15 cm in diameter and 10 cm in depth) were sampled from a meadow and another 60 soil blocks from a mixed forest nearby. For each habitat, 30 soil blocks were frozen at -20°C for one week to eradicate soil animals, and the other 30 blocks were left undisturbed. Fifteen defaunated and 15 undisturbed blocks were installed back to their original habitat, while the other 15 defaunated and 15 untreated blocks were transferred to the respective other habitat. The manipulations resulted in eight treatments with 15 soil blocks each (**Table 5.1**). The soil blocks were sampled again in three time intervals: one week (T1), one month (T2) and six months (T3) after installation. For each sampling interval, five blocks per treatment were randomly selected and used as replicates in the following statistical analyses. Soil animals were extracted from the blocks using heat and Collembola were identified to species level. For more details on the experiment see Auclerc et al. (2009).

Table 5.1

Eight treatments used in Auclerc et al. (2009). Soil blocks were sampled from meadow or forest, defaunated or untreated and installed to the original or the respective other habitat. For example, WMF indicates untreated soil blocks with fauna (W) originated from meadow (M) and installed in forest (F). Each of the treatments was replicated in 15 soil blocks. Five replicates per treatment were sampled at each of the three intervals, T1 (one week), T2 (one month) and T3 (six months) after soil block installation.

Habitat	**Defaunation**	**Soil block installation to**	
		Original habitat	**Respective other habitat**
Meadow	Untreated	WMM	WMF
	Defaunated	OMM	OMF
Forest	Untreated	WFF	WFM
	Defaunated	OFF	OFM

Collembola phylogeny

Forty-nine species of Collembola were found in the untreated blocks installed in the original habitats (WFF and WMM; **Table 5.1**) and were used as the regional species pool in the community phylogenetic analyses. Collembola phylogeny of the regional species pool was constructed using six genetic markers, including 18S rRNA, 28S rRNA (D1, D2 and D3-D5 regions), Histone H3 and Cytochrome Oxidase I (COI). Sequences were downloaded from Genbank with the Accession Number listed in **Table S5.2**. For species for which none of the above molecular sequences were available, sequences of congeneric species were used. In case of lack of congeneric sequences, taxa were grafted to the most recent common ancestors according to the taxonomy (**Table S5.2**).

Sequences of each marker were aligned with outgroup taxa Zygentoma (Insecta), *Machilis* (Insecta: Archaeognatha) and *Callibaetis* (Insecta: Palaeoptera) using R functions *AlignSeqs* and *AdjustAlignment* for 18S and 28S rRNA (package "DECIPHER"; Wright 2015) and *msa* for Histone H3 and COI, setting gap opening penalty to 15 and gap extension penalty to 6.6 (package "msa"; Bodenhofer et al. 2015). For each genetic marker the aligned sequences were trimmed to the same length. The best model of sequence evolution of each genetic marker was estimated using jModelTest 2.1.4 on the basis of the Akaike information criterion (AIC; Guindon and Gascuel, 2003, Darriba et al. 2012). Ribosomal 18S and 28S markers and COI were fitted with GTR+I+G model, while Histone H3 was fitted with SYM+I+G. In each marker set, terminal gaps at the beginning and the end of sequences varied in length and were replaced by "?", and all the six markers was concatenated in a supermatrix (3,073 bp) using SequenceMatrix 1.8 (Vaidya et al. 2011). Collembola phylogeny was inferred using Bayesian Inference (BI; MrBayes 3.2.4; Ronquist et al. 2012), setting the model of sequence evolution separately for the six markers, two independent runs, four chains, 2,000,000 generations, 0.1 temperature and 0.5 burn-in fraction; other parameters were set as default.

The resulting tree (**Figure S5.1a**) was transformed to an ultrametric tree using a penalized likelihood approach by assuming different models of substitution rate variation among branches, i.e., correlated, relaxed, discrete and strict clock models, using the function *chronos* implemented in the R package "ape" (Paradis et al. 2004). The best ultrametric tree was selected based on the smallest PHIIC value, a criterion analogous to AIC reflecting the best model fit to the data (Paradis 2013). The ultrametric tree based on the strict clock model was selected, and the species without available sequences (*Gisinianus flammeolus* and *Stenognathellus denisi*) were grafted to the most recent common ancestors according to the taxonomy using the function *add.species.to.genus* implemented in the R package "phytools" (Revell 2012). This tree was then used in comparative analyses of species preferences for habitats and phylogenetic analyses of community (**Figure S5.1b**).

Phylogenetic comparative methods

Collembola preferences for habitat and soil and dispersal abilities in meadow and forest reported in Auclerc et al. (2009) were tested for phylogenetic signal (**Table S5.2**). Species for which data on soil preference were lacking were removed from the corresponding analysis. Characters of species were mapped onto the phylogenetic tree and phylogenetic signal was measured using Pagel's lambda (Pagel 1999, Freckleton et al. 2002) by the function *fitDiscrete* implemented in R package "geiger" (Harmon et al. 2008). An appropriate model of character evolution was estimated from one of the equal-rates (ER), symmetric (SYM) and all-rates-different (ARD) models using likelihood comparison. The ER model was selected for habitat and soil preferences, while the SYM model was selected for dispersal ability in meadow and forest. A star-like tree (lambda 0) was then transformed from the original tree (lambda 1). Likelihoods of the distribution of character states among species were compared, given the lambda 1 and lambda 0 trees. If the lambda 0 tree was rejected, the character showed phylogenetic signal. An optimal lambda value was then estimated which represented the strength of phylogenetic signal (Pagel 1999, Freckleton et al. 2002).

For the characters exhibiting phylogenetic signal, ancestral states of characters were estimated using stochastic character mapping (Huelsenbeck et al. 2003, Bollback 2006) by the function *make.simmap* implemented in the R package "phytools" (Revell 2012). Transition of character state was assumed following the ER model. Prior distribution on root node was estimated from tip character states. Transition matrix Q was sampled 1,000 times from the posterior probability distribution using Bayesian MCMC. Then, 1,000 stochastic maps were simulated which were conditioned by the sampled value of Q. Numbers of character state transitions in the tree were reported as mean and median, and posterior probabilities of character states were mapped to the tree nodes.

Community composition and phylogenetic structure

For each community assembly hypothesis, a respective metacommunity was defined as a combination of several treatments (**Table S5.1**). Differences in community compositions were calculated based on Bray-Curtis dissimilarity between communities, resulting in a phylogeny-unweighted distance matrix. Effects of treatment and sampling interval on community compositions were tested by distance-based multivariate analysis of variance (ADONIS) with 9,999 permutations using the function *adonis* implemented in the R package "vegan" (Oksanen et al. 2015). If treatments had significant effects on community compositions, principle coordinate analysis (PCoA) was performed to visualize community compositions using the function *pcoa* implemented in the R package "ape" (Paradis et al. 2004). Pearson correlation coefficients of site scores and species abundance at each PCoA axis were calculated using function *cor.test*. *P*-values of correlation tests were adjusted using Benjamini and Hochberg corrections (BH; Benjamini and Hochberg 1995). Treatment effects on site scores of the PCoA axes were tested using Kruskal-Wallis rank sum test followed by the Dunn's test with BH corrections for multiple comparisons.

Furthermore, community compositions were weighted by species phylogenetic relationships using the phylogenetic fuzzy-weighting method (Pillar and Duarte 2010) implemented in the R package "SYNCSA"; Debastiani and Pillar 2012). ADONIS was then used to test treatment effects on the phylogeny-weighted community compositions. If treatments had significant effects, phylogeny-weighted community compositions were applied to PCoA using the function *pcps* implemented in the R package "PCPS" (Debastiani and Duarte 2014), resulting in principal coordinates of phylogenetic structures (PCPS) of the respective metacommunity. Pearson correlation coefficients of site scores and species abundance at each PCPS axis were tested as described above. Treatment effects on PCPS using Kruskal-Wallis rank sum test followed by the Dunn's test were tested as described above.

Community phylogenetic relatedness

Phylogenetic relatedness of coexisting species in local communities was calculated based on species pairwise phylogenetic distances obtained from the ultrametric tree using the function *cophenetic.phylo* implemented in R package "ape" (Paradis et al. 2004). Mean pairwise phylogenetic distance (MPD) and mean nearest taxon distance (MNTD) of a local community were calculated and Net Relatedness Index (NRI) and Nearest Taxon Index (NTI) were calculated based on MPD and MNTD to represent overall species relatedness and terminal species relatedness of a local community, respectively (Webb 2000). The NRI is a negative value of the standardized effect size of the MPD, for which the observed MPD was standardized by 999 simulated MPDs generated by drawing species with equal probability from the phylogenetic distance matrix ("phylogeny.pool" null model). Similarly, the NTI is a negative value of the standardized effect size of the MNTD, for which the observed MNTD was standardized by 999 simulated MNTDs. A positive NRI or NTI indicates more closely related species as compared to the "null" communities. A negative NRI or NTI, on the contrary, indicates more distantly related species in a local community. NRI and NTI were calculated using the function *ses.mpd* and *ses.mntd* implemented in R package "picante", respectively (Kembel et al. 2010), based on species presence/absence or abundance data.

To infer community assembly processes in each treatment, Student's t-test was used to test mean phylogenetic relatedness of the five replicates of a treatment against "0", which indicates a randomly assembled community. *P*-values were adjusted by the number of treatments within a metacommunity using BH corrections. Kruskal-Wallis rank sum test was applied to test effects of treatments on NRI or NTI, followed by the Dunn's test with BH corrections.

Results

Evolution of species preferences for habitats and dispersal abilities

Among the four characters published in Auclerc et al. (2009), only soil preference and dispersal ability in meadow exhibited phylogenetic signal (soil preference: lambda = 0.881, $P = 0.003$; dispersal ability in meadow: lambda = 0.498, $P = 0.020$; **Figure 5.1a**, **b**).

The most likely ancestral state of soil preference of Collembola was for the meadow soil. The median number of soil preference changes was 21. The median number of transitions from meadow specialist to soil generalist was eight and for the opposite direction it was three, while from meadow specialist to forest specialist it was five and in the opposite direction it was two. There was a single transition between soil generalist and forest specialist for both directions (**Table S5.3**). The ancestor of Neanuridae and the ancestor of Lepidocyrtidae were likely soil generalists, while the ancestor of Isotomidae was likely a meadow specialist (**Figure 5.1a**).

Ancestral state of Collembola dispersal ability in meadow was likely either M0 or M1. Transitions of dispersal ability in meadow occurred more often between M0 and M1 and between M0 and M4 than the

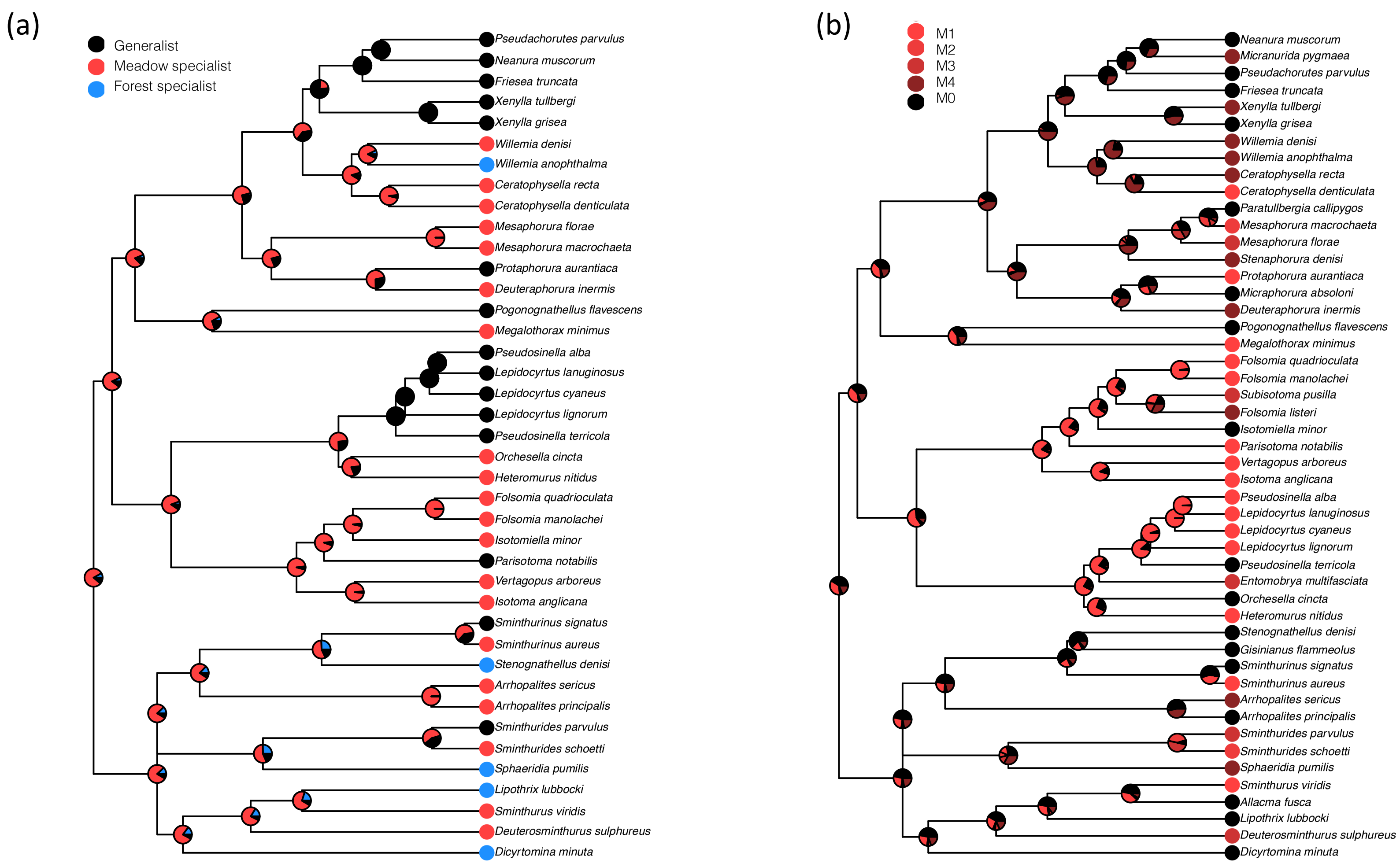

Figure 5.1
Ancestral state reconstruction of (a) soil preference and (b) dispersal ability in meadow of Collembola species reported in Auclerc et al. (2009). Colored circles at the tips of each branch represent trait state of each extant species. Pie charts on each node indicate the proportion of each character state summed across the posterior distribution of simulations. M1, species in defaunated meadow blocks after one week; M2, species in defaunated meadow blocks after one month; M3, species in defaunated meadow blocks after six months; M4, species which did not colonize defaunated meadow blocks within six months; M0, species absent in the meadow.

other state transitions (**Table S5.4**). The ancestor of Isotomidae and Entomobryoidea was likely a fast disperser in meadow (M1), while the ancestor of Poduromorpha was likely a slow or poor disperser in meadow (M4 and M0). The ancestor of Symphypleona, however, was likely a poor disperser in meadow (M0; **Figure 5.1b**).

Habitat effects on Collembola communities

Both community structures and phylogeny-weighted structures differed significantly between meadow (WMM) and forest (WFF) but not between sampling intervals (**Table S5.5**). Site scores on PCoA 1 and those on PCPS 1, PCPS 3 and PCPS 5 significantly differed between meadow and forest (**Figure 5.2a**, **b**, **Table S5.6**). Phylogenetic gradient along PCPS 1 was mostly correlated with Entomobryoidea and Isotomidae which were associated with forest, and Poduromorpha and Symphypleona which were associated with meadow. PCPS 3 was positively correlated with Poduromorpha and negatively with Symphypleona (**Figure 5.2b**). Communities in meadow showed phylogenetic overdispersion, while in forest they were phylogenetic clustered as indicated by NRI using species presence/absence data, with a significant difference between meadow and forest. When species abundances were considered, NRI increased in both habitats, indicating that abundant species in both habitats were close relatives. Mean NTI of communities did not differ significantly between meadow and forest when species presence/ absence data were used. However, abundant species of forest communities exhibited phylogenetic clustering, resulting in the NTI being higher in forest than in meadow (**Table 5.2**).

Successional patterns of community structures

Community structures differed during succession in both meadow and forest, mainly due to increased abundance of a few species (**Figure 5.3a**, **c**, **Table S5.5**). In contrast, phylogeny-weighted structures in meadow did not differ significantly with successional stages (**Figure 5.3b**, **Table S5.5**). However, phylogeny-weighted structures in forest differed between OFF at T1 and WFF (**Table S5.6**). The OFF communities at T1 were associated with Symphypleona and Entomobryoidea, while WFF communities were associated with Isotomidae and Poduromorpha (**Figure 5.3d**).

In meadow, NRI using species presence/absence data decreased from T1 to T3 and to the WMM where local communities showed phylogenetic overdispersion. However, when abundances were considered, although NRI did not differ from 0 in all successional stages, NRI at T3 was significantly higher than that at T2 and WMM. In contrast, NTI of OMM at T2 and of WMM show phylogenetic overdispersion but were not different from each other. In forest OFF communities, although phylogenetic relatedness based on species presence/absence data did not differ from that of a random pattern, abundance-weighted NRI and NTI showed significant phylogenetic clustering in each sampling interval, indicating that abundant species were phylogenetically closely related. The abundance-weighted NRI in OFF at T2 was the lowest among all the successional stages, while the abundance-weighted NTI in OFF at T1 was the highest. Interestingly, NTI using species presence/absence data decreased from T1 to T3 and to WFF the lowest (**Table 5.2**).

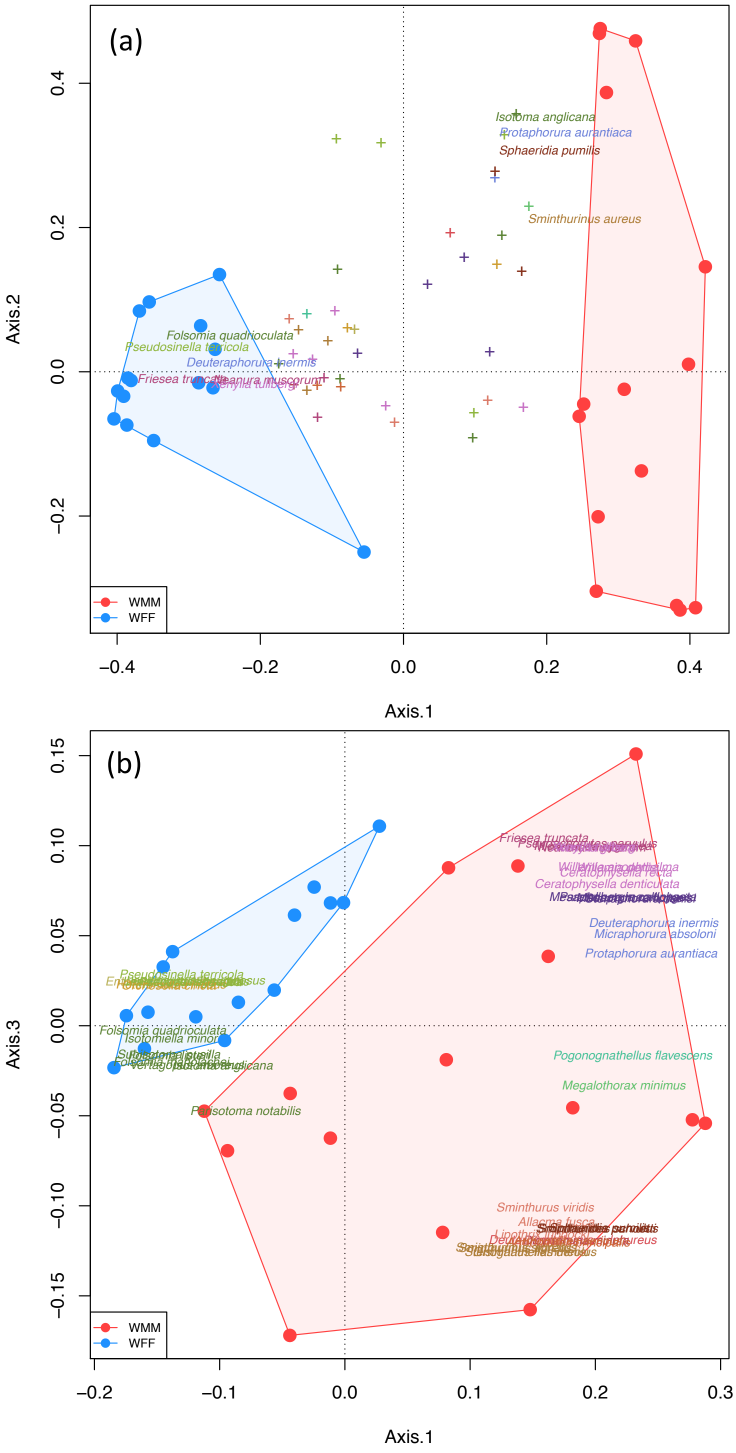

Figure 5.2
Scatter diagram of principle coordinates of (a) community structure (PCoA) and (b) phylogenetic-weighted structure (PCPS) of Collembola communities occurring in the untreated soil blocks installed in meadow (WMM) and forest (WFF) pooled for sampling times. Polygons encompass replicates of the same treatment. Only Collembola species significantly correlated with the axes are plotted with names and the others with cross symbols.

Table 5.2

Phylogenetic relatedness (mean ± SD) represented by Net Relatedness Index (NRI) and Nearest Taxon Index (NTI), based on species occurrence (presence/absence, p/a) or abundance data of Collembola assemblages collected in different treatments. Phylogenetic clustering is marked in red, while phylogenetic overdispersion is marked in blue. A mean value significantly different from zero (*P*-adj < 0.05, Student's t-test with Benjamini and Hochberg's (BH) adjustment) is marked in bold. Letters in lowercase indicate significant differences between treatments (*P* < 0.05, Dunn's test with BH adjustment). Asterisks in Dunn's test indicate significant differences without BH adjustment. Asterisks in treatments indicate the focused treatments compared with the other treatments. For abbreviation of treatments and definitions of pool see Table 5.1 and S5.1, respectively.

Hypotheses	Pool	Treatments	Time	NRI						NTI					
				p/a			Abundance			p/a			Abundance		
				Mean ± SD	*P*-adj	Dunn	Mean ± SD	*P*-adj	Dunn	Mean ± SD	*P*-adj	Dunn	Mean ± SD	*P*-adj	Dunn
(1) Habitat effects	Reference	WMM	-	**-0.44 ± 0.52**	**0.006**	**a**	0.23 ± 1.00	0.381	**a**	-0.42 ± 0.77	0.100		-0.13 ± 0.68	0.474	**a**
		WFF	-	**0.97 ± 1.18**	**0.006**	**b**	**1.74 ± 0.83**	**0.000**	**b**	0.01 ± 0.75	0.949		**1.03 ± 0.23**	**0.000**	**b**
(2) Succession patterns	M	T1 - OMM	-	0.40 ± 1.43	0.813	a*	0.26 ± 1.05	0.833	ab*	-0.29 ± 0.86	0.625		-0.10 ± 1.10	0.625	
		T2 - OMM	-	-0.66 ± 0.71	0.143	ab*	-0.13 ± 0.33	0.833	a*	**-0.98 ± 0.54**	**0.032**		-0.20 ± 0.47	0.523	
		T3 - OMM	-	**-0.78 ± 0.40**	**0.024**	ab*	1.13 ± 0.92	0.207	b*	-0.65 ± 0.80	0.191		-0.08 ± 0.18	0.523	
		WMM	-	**-0.89 ± 0.53**	**0.000**	b*	0.04 ± 1.01	0.873	a*	**-0.72 ± 0.83**	**0.019**		-0.28 ± 0.71	0.523	
	F	T1 - OFF	-	0.89 ± 0.87	0.171		**2.25 ± 0.90**	**0.005**	a*	1.97 ± 1.32	0.066	a*	**1.74 ± 0.67**	**0.004**	a*
		T2 - OFF	-	0.26 ± 0.76	0.481		**0.91 ± 0.28**	**0.003**	b*	0.87 ± 0.61	0.066	ab*	**1.07 ± 0.35**	**0.003**	b*
		T3 - OFF	-	0.33 ± 0.94	0.481		**1.68 ± 0.25**	**0.000**	a*	0.77 ± 0.84	0.149	ab*	**0.97 ± 0.28**	**0.003**	b*
		WFF	-	0.84 ± 1.13	0.051		**1.64 ± 0.77**	**0.000**	a*	0.25 ± 0.74	0.213	b*	**1.15 ± 0.25**	**0.000**	b*
(3) Soil origin effects	OM	OFM*	T1	**2.23 ± 0.71**	**0.010**	**a**	**1.30 ± 0.31**	**0.006**		**1.45 ± 0.41**	**0.013**	**a**	**1.12 ± 0.35**	**0.011**	**a**
			T2	1.41 ± 1.74	0.185	**ab**	0.90 ± 0.76	0.172		0.95 ± 0.82	0.110	**ab**	**1.21 ± 0.39**	**0.011**	**a**
			T3	-0.50 ± 0.67	0.189	**bcd**	0.42 ± 0.51	0.309		-0.05 ± 0.64	0.880	**abc**	0.44 ± 0.30	0.097	**ab**
		OMM	T1	0.38 ± 1.42	0.579	**abc**	0.23 ± 1.05	0.884		-0.28 ± 0.84	0.650	**abc**	-0.10 ± 1.14	0.855	**b**
			T2	-0.67 ± 0.72	0.156	**bcd**	-0.07 ± 0.35	0.884		-0.97 ± 0.56	0.053	**c**	-0.19 ± 0.48	0.631	**b**
			T3	**-0.77 ± 0.38**	**0.033**	**cd**	1.18 ± 1.00	0.172		-0.67 ± 0.78	0.188	**bc**	-0.08 ± 0.20	0.631	**b**
		WMM	T1	**-1.17 ± 0.37**	**0.010**	**d**	-0.12 ± 1.26	0.901		-1.05 ± 0.89	0.110	**c**	-0.49 ± 0.90	0.631	**b**
			T2	-0.86 ± 0.80	0.133	**cd**	0.06 ± 1.09	0.901		-1.05 ± 0.58	0.053	**c**	-0.25 ± 0.80	0.670	**b**
			T3	**-0.70 ± 0.41**	**0.041**	**bcd**	0.15 ± 0.80	0.884		-0.06 ± 0.70	0.880	**abc**	-0.04 ± 0.39	0.855	**b**
	OF	OMF*	T1	0.97 ± 0.65	0.176		**0.69 ± 0.41**	**0.021**	**c**	1.23 ± 0.78	0.078	**a**	0.86 ± 0.99	0.140	
			T2	0.50 ± 1.57	0.519		0.33 ± 0.82	0.419	**c**	0.84 ± 1.45	0.267	**abc**	0.54 ± 1.10	0.336	
			T3	-0.34 ± 0.44	0.270		**1.53 ± 0.54**	**0.007**	**abc**	-0.39 ± 0.57	0.230	**bc**	**0.71 ± 0.37**	**0.016**	
		OFF	T1	1.13 ± 0.93	0.176		**2.51 ± 0.93**	**0.007**	**a**	2.02 ± 1.30	0.078	**a**	**1.80 ± 0.71**	**0.007**	
			T2	0.48 ± 0.89	0.334		**0.95 ± 0.25**	**0.005**	**bc**	0.95 ± 0.62	0.078	**ab**	**1.11 ± 0.36**	**0.004**	
			T3	0.70 ± 1.05	0.270		**1.98 ± 0.28**	**0.001**	**ab**	0.77 ± 0.87	0.204	**abc**	**0.98 ± 0.28**	**0.003**	
		WFF	T1	1.44 ± 1.22	0.176		**1.56 ± 0.79**	**0.017**	**abc**	0.54 ± 0.65	0.204	**abc**	**1.20 ± 0.20**	**0.002**	
			T2	1.12 ± 1.58	0.270		**1.51 ± 0.80**	**0.017**	**abc**	0.59 ± 0.81	0.229	**abc**	**1.18 ± 0.31**	**0.003**	
			T3	0.89 ± 0.95	0.231		**2.36 ± 0.76**	**0.007**	**a**	-0.42 ± 0.37	0.140	**c**	**1.01 ± 0.24**	**0.003**	

Table 5.2 Continued

Hypotheses	Pool	Treatments	Time	NRI						NTI					
				p/a			Abundance			p/a			Abundance		
				Mean ± SD	*P*-adj	Dunn	Mean ± SD	*P*-adj	Dunn	Mean ± SD	*P*-adj	Dunn	Mean ± SD	*P*-adj	Dunn
(4) Indigenous species	FM	WFF	T1	1.26 ± 1.22	0.199	ab	**1.52 ± 0.80**	**0.034**	ab	0.31 ± 0.65	0.567	ab	**1.11 ± 0.18**	**0.002**	ab
			T2	0.93 ± 1.52	0.349	abc	**1.43 ± 0.77**	**0.034**	ab	0.33 ± 0.83	0.567	ab	**1.08 ± 0.25**	**0.004**	ab
			T3	0.66 ± 0.94	0.332	abc	**2.18 ± 0.66**	**0.011**	b	-0.63 ± 0.39	0.114	b	**0.92 ± 0.26**	**0.004**	abc
		WFM*	T1	1.24 ± 1.29	0.199	abc	**1.28 ± 0.75**	**0.038**	ab	0.26 ± 1.04	0.604	ab	0.60 ± 0.51	0.087	bcde
			T2	1.59 ± 1.27	0.196	ab	**1.43 ± 0.73**	**0.034**	ab	0.17 ± 0.57	0.604	ab	**0.79 ± 0.51**	**0.045**	abcd
			T3	0.59 ± 1.40	0.464	abc	1.02 ± 0.92	0.102	ab	-0.44 ± 0.97	0.567	b	0.62 ± 0.60	0.112	bcde
		OFM	T1	**2.54 ± 0.74**	**0.018**	a	**1.55 ± 0.40**	**0.011**	ab	**1.63 ± 0.40**	**0.010**	a	**1.29 ± 0.37**	**0.004**	ab
			T2	1.75 ± 1.79	0.199	ab	1.05 ± 0.80	0.072	ab	1.15 ± 0.77	0.114	a	**1.40 ± 0.39**	**0.004**	a
			T3	-0.11 ± 0.65	0.714	bc	0.59 ± 0.56	0.102	a	0.27 ± 0.63	0.567	ab	**0.59 ± 0.29**	**0.022**	bcde
		WMM	T1	**-0.70 ± 0.29**	**0.033**	c	0.04 ± 1.23	0.951	a	-0.70 ± 0.82	0.311	b	-0.33 ± 0.81	0.492	e
			T2	-0.42 ± 0.73	0.349	bc	0.28 ± 1.13	0.670	a	-0.71 ± 0.55	0.134	b	-0.09 ± 0.76	0.808	cde
			T3	-0.16 ± 0.40	0.464	bc	0.33 ± 0.81	0.499	a	0.16 ± 0.61	0.604	ab	0.05 ± 0.39	0.808	de
	MF	WMM	T1	-0.71 ± 0.31	0.084		0.04 ± 1.21	0.942	a	-0.69 ± 0.84	0.280	b	-0.32 ± 0.83	0.660	c
			T2	-0.44 ± 0.72	0.339		0.26 ± 1.09	0.682	a	-0.73 ± 0.57	0.188	b	-0.10 ± 0.78	0.889	bc
			T3	-0.15 ± 0.38	0.481		0.40 ± 0.86	0.482	a	0.17 ± 0.61	0.627	ab	0.05 ± 0.40	0.889	c
		WMF*	T1	1.05 ± 1.12	0.250		0.55 ± 1.15	0.482	a	-0.24 ± 0.37	0.376	ab	-0.09 ± 0.71	0.889	c
			T2	0.57 ± 1.15	0.396		0.60 ± 0.67	0.202	a	-0.72 ± 0.66	0.212	b	-0.00 ± 0.66	0.993	bc
			T3	0.33 ± 0.52	0.339		**2.41 ± 0.39**	**0.002**	b	-0.03 ± 0.64	0.925	ab	0.41 ± 0.24	0.053	abc
		OMF	T1	0.87 ± 0.65	0.158		0.64 ± 0.43	0.056	a	1.06 ± 0.76	0.188	a	0.71 ± 0.98	0.367	abc
			T2	0.39 ± 1.59	0.616		0.24 ± 0.84	0.669	a	0.68 ± 1.52	0.467	ab	0.43 ± 1.11	0.660	abc
			T3	-0.55 ± 0.40	0.158		**1.45 ± 0.55**	**0.017**	ab	-0.57 ± 0.59	0.229	b	0.61 ± 0.38	0.054	abc
		WFF	T1	1.22 ± 1.12	0.214		**1.44 ± 0.71**	**0.031**	ab	0.29 ± 0.62	0.467	ab	**1.09 ± 0.19**	**0.003**	a
			T2	0.89 ± 1.50	0.339		**1.51 ± 0.83**	**0.036**	ab	0.35 ± 0.81	0.467	ab	**1.09 ± 0.29**	**0.004**	a
			T3	0.67 ± 0.93	0.339		**2.27 ± 0.65**	**0.009**	b	-0.64 ± 0.41	0.188	b	**0.90 ± 0.24**	**0.004**	ab

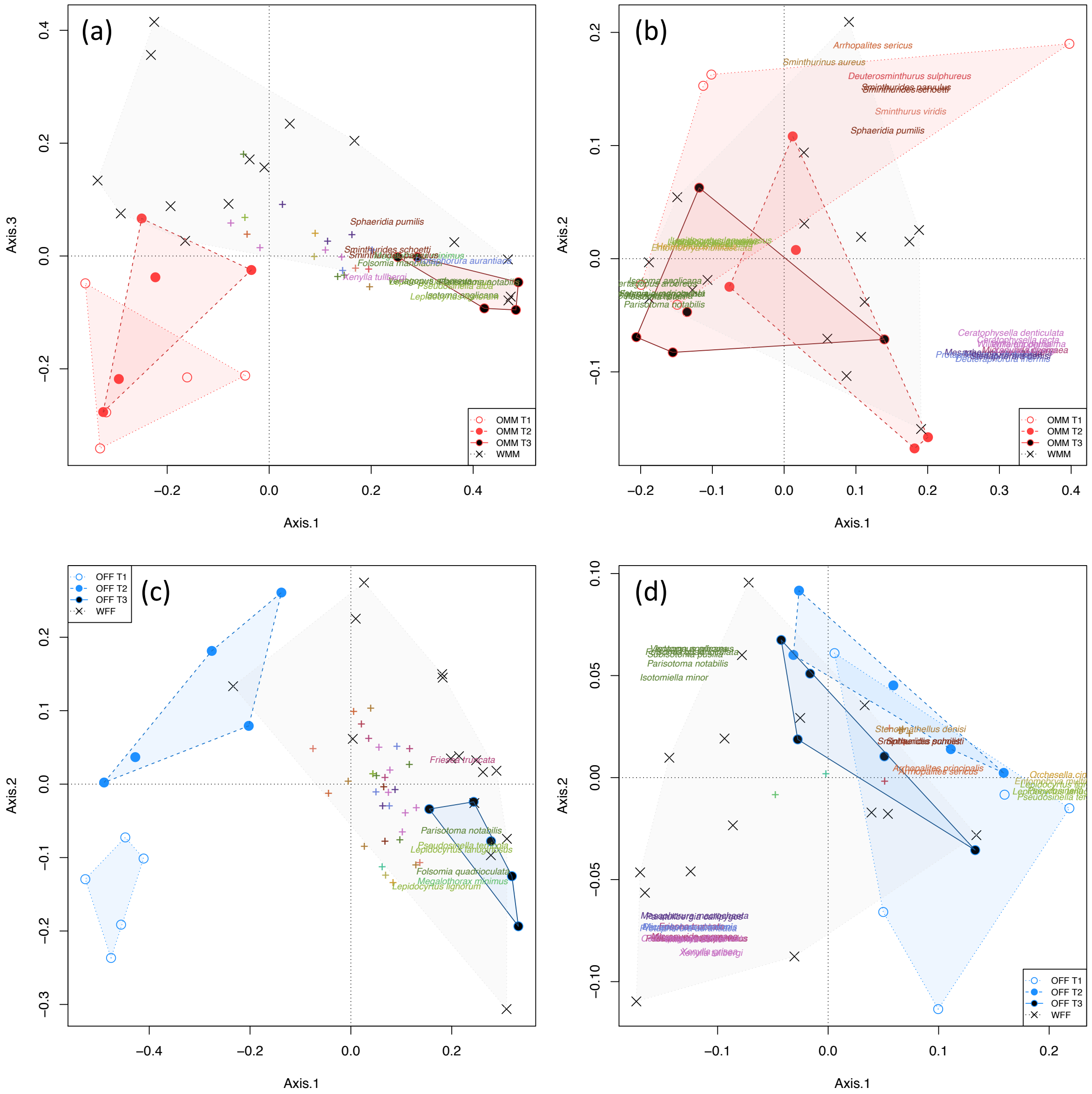

Figure 5.3
Scatter diagram of principle coordinates of (a) community structure (PCoA) and (b) phylogenetic-weighted structure (PCPS) of Collembola communities occurring in defaunated meadow soil blocks (OMM) sampled one week (T1), one month (T2) and six months (T3) after installation and in untreated meadow soil blocks (WMM). Scatter diagram of principle coordinates of (c) community structure (PCoA) and (d) phylogenetic-weighted structure (PCPS) of Collembola communities occurring in defaunated forest soil blocks (OFF) sampled one week (T1), one month (T2) and six months (T3) after installation and in untreated forest soil blocks (WFF). Polygons encompass replicates of the same sampling time. Only Collembola species significantly correlated to the axes are plotted with names and the others with cross symbols.

Effects of soil origin on community compositions

Communities of OFM, OMM and WMM, all surrounded by meadow soil, differed in their assembly patterns but this varied with sampling intervals (**Table S5.5**). Site scores of OFM communities at T1 differed from those at T3 on PCoA 1. OFM communities differed from OMM communities at T2 for site scores on PCoA 4 which were positively correlated with the abundance of *Ceratophysella denticulata* (**Figure 5.4a**, **Table S5.6**). In contrast, phylogeny-weighted structures were influenced by treatments but not by sampling intervals (**Table S5.5**, **S5.6**). Soil originating from forest was associated with Entomobryoidea that immigrated from the meadow surrounding (**Figure 5.4b**). OFM communities exhibited phylogenetic clustering that decreased with successional stages, as indicated by NRI using presence/absence data. However, NRI of OFM communities did not differ significantly from those of OMM in all three sampling intervals. Furthermore, NTI of OFM at T2 was significantly higher than that

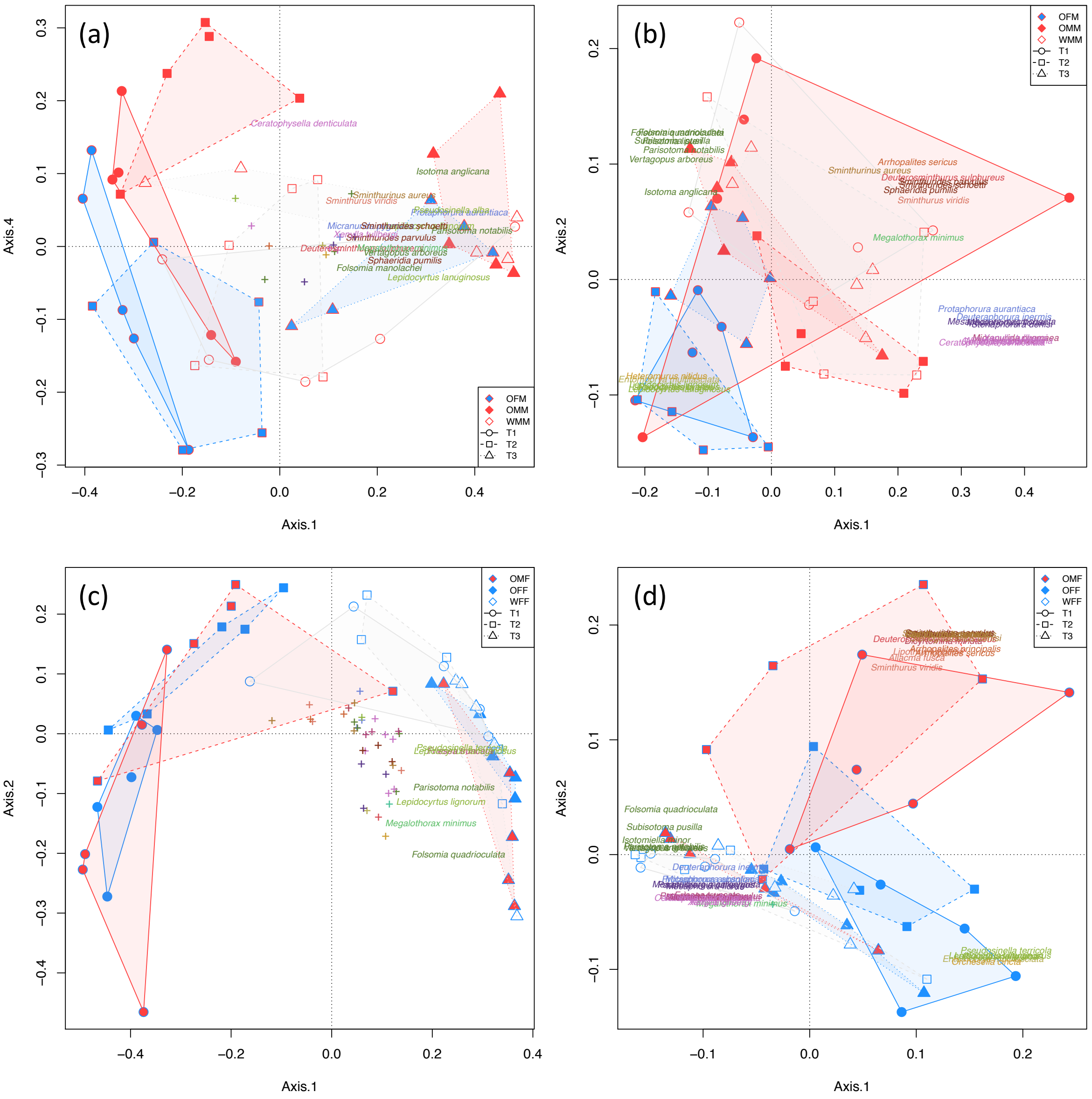

Figure 5.4
Scatter diagram of principle coordinates of (a) community structure (PCoA) and (b) phylogenetic-weighted structure (PCPS) of Collembola communities surrounded by meadow, and of (c) community structure (PCoA) and (d) phylogenetic-weighted structure (PCPS) of Collembola communities surrounded by forest. Polygons encompass replicates of the same treatment per sampling time. Only Collembola species significantly correlated to the axes are plotted with names and the others with cross symbols.

of OMM using species presence/absence data, while abundance-weighted NTI at both T1 and T2 were higher than that of OMM at the corresponding sampling intervals. This indicates that while OFM communities were phylogenetically clustered at T1, phylogenetic relatedness decreased with succession to a more random pattern at T3; compared with OMM, OFM exhibited more phylogenetic clustering (**Table 5.2**).

OMF, OFF and WFF communities, all surrounded by forest soil, differed in their assembly patterns but this again varied with sampling intervals (**Table S5.5**). OMF communities did not differ from OFF communities in any sampling interval (**Table S5.6**), although OMF communities changed with successional stages (**Figure 5.4c**). This indicates that meadow soil in the transferred blocks to forest habitat did not influence community structures significantly. Similarly, phylogeny-weighted structures differed between treatments and between successional stages, with these two factors depending on each other (**Table S5.5**). Differences between OMF and OFF communities were significant at T1 on PCP2 (**Table S5.6**). On that axis OMF communities were associated with Symphypleona while those of OFF with Entomobryoidea (**Figure 5.4d**). At T1 mean abundance-weighted NRI of OMF communities was significantly lower than that of OFF communities but both were higher than 0. The species presence/absence NTI of OMF communities decreased significantly with successional stage but did not differ from random. Abundant species in the OMF communities exhibited phylogenetic clustering at T3 as indicated by both NRI and NTI (**Table 5.2**).

Indigenous species

Indigenous species originating from one habitat but transferred to another habitat contributed to differences in community composition as compared with defaunated soil blocks, but this varied with time (**Table S5.5**). Collembola communities in forest blocks transferred to meadow (WFM) resembled those of their original forest habitat (WFF), but differed from those of the meadow (WMM) at each sampling interval, as indicated by site scores on PCoA 1 (**Figure 5.5a**, **Table S5.6**). Phylogeny-weighted structures differed between treatments but not between sampling intervals (**Table S5.5**). Furthermore, phylogeny-weighted structures of WFM communities at T3 resembled that of WMM communities but differed from that of WFF communities on PCPS 1, which was positively correlated with Poduromorpha, Tomoceridae and Neelidae, and negatively with Isotomidae and Entomobryoidea (**Figure 5.5b**). Phylogenetic relatedness in WFM communities did not differ from those of the other treatments at all sampling intervals (**Table 5.2**).

Collembola communities in meadow blocks transferred to forest (WMF) resembled those of their original meadow habitat (WMM) at T1 and T2, but were more similar to WFF communities at T3. Significant differences between WMF and OMF communities occurred at T2, when WMF communities resembled WMM communities and OMF communities resembled WFF communities. At T3 both WMF and OMF communities resembled WFF communities (**Figure 5.5c**, **Table S5.5**, **S5.6**). Phylogeny-weighted structures of WMF could not be differentiated from either WMM or WFF at each sampling interval. However, WMF communities differed from OMF communities at T1, as indicated by PCPS 1 and PCPS 2 (**Table S5.6**). WMF communities at T1 were associated with the clade composed of Poduromorpha plus Tomoceridae and Neelidae, while OMF communities at T1 were associated with the other basal phylogenetic clades, e.g. Entomobryoidea and Isotomidae (**Figure 5.5d**). Mean abundance-weighted phylogenetic relatedness of WMF communities at T3 was significantly higher than 0 as well as that of WMM communities, but not different from that of WFF and OMF communities as indicated by abundance-weighted NRI (**Table 5.2**).

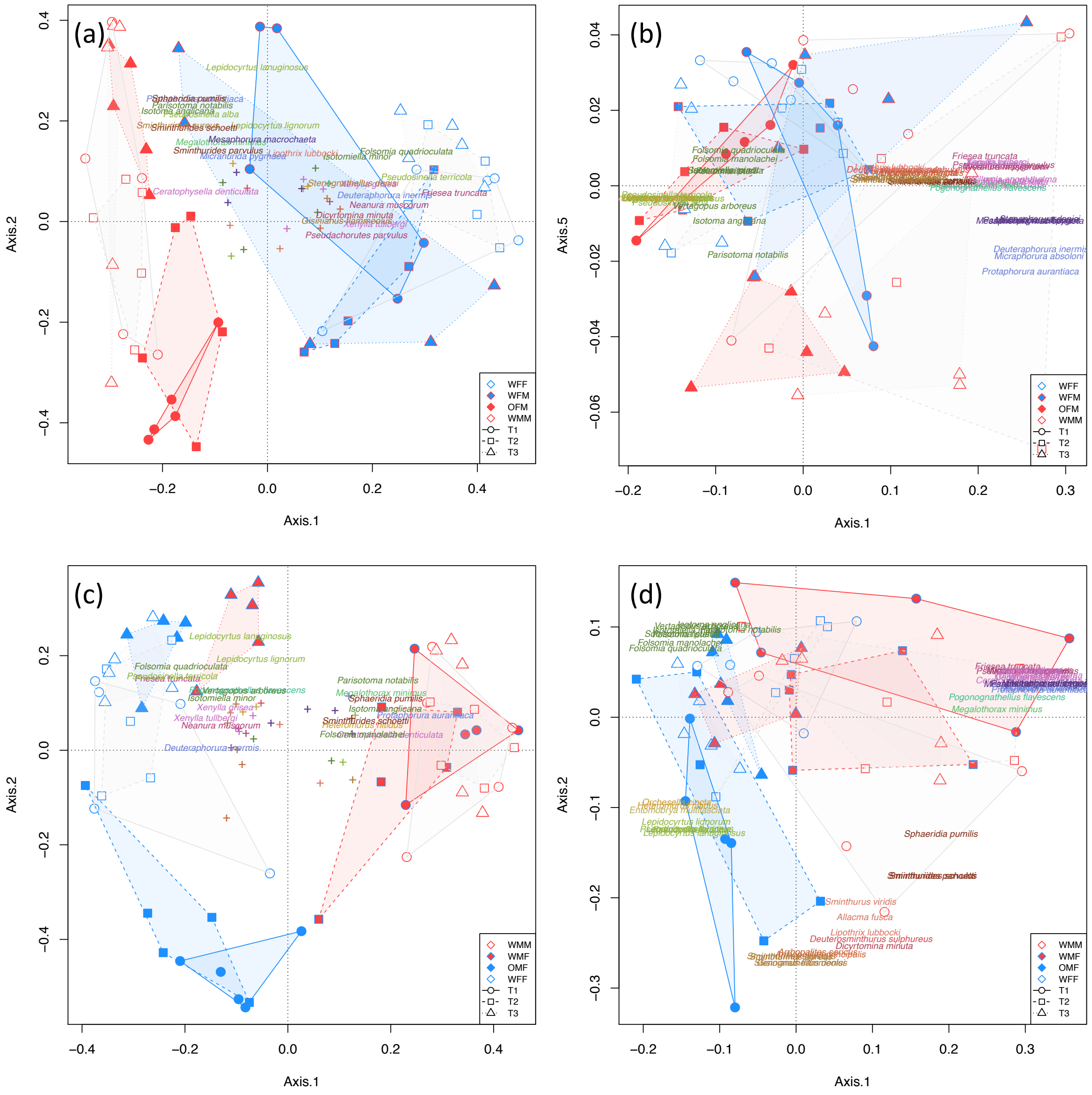

Figure 5.5
Scatter diagram of principle coordinates of community structure (a) and (c), and phylogenetic-weighted structure (b) and (d) of Collembola communities. Polygons encompass replicates of the same treatment per sampling time. Only Collembola species significantly correlated to the axes are plotted with names and the others with cross symbols.

Discussion

Evolution of Collembola dispersal abilities and environmental preferences

Phylogenetic signal found in Collembola dispersal ability in meadow suggests that related species displayed similar ability to move from the surrounding to the defaunated soil blocks in meadow habitat. Generally, most of Isotomidae and Entomobryoidea recolonized the defaunated soil blocks, while most species from Neanuridae and Hypogastruridae were poor dispersers. Dispersal ability in meadow likely reflected morphological characters of different taxa. Isotomidae and Entomobryoidea usually are elongate and possess long furca, while Neanuridae and Hypogastruridae are stout with short furca. Since these morphological characters were fixed within taxonomic groups (**Chapter 2**), ancestor of Isotomidae and Entomobryoidea in a meadow-like habitat was likely a fast disperser, while that of Poduromorpha was likely a poor disperser. Similarly, preference of Collembola for soil types exhibited phylogenetic signal. Ancestral character estimation showed that ancestor of Collembola was likely a meadow soil specialist but this changed to soil generalists several times during species diversification, while changes from generalist to meadow soil specialist occurred less frequently. Notably, changes from meadow soil specialists to forest soil specialists occurred frequently, while changes from soil generalists to forest soil specialists and vice versa rarely occurred. Compared to other soil animal taxa such as oribatid mites (Oribatida), Collembola recover faster after drought (Lindberg and Bengtsson 2005) and more tolerate fluctuations in temperature and soil moisture (Tsiafouli et al. 2005). Accordingly, Collembola might have adapted to meadow-like habitats characterized by fluctuations in abiotic environmental factors.

In the paper of Auclerc et al. (2009) Collembola dispersal ability in meadow and forest was defined by the time at which the species reappeared in the defaunated blocks (OFF, OMM), while soil preference was estimated by difference in abundance of species between forest and meadow soil cores installed in the same habitat. Species abilities to immigrate from the surrounding habitat and to survive in a new habitat are determined not only by their morphological characters, such as legs, antenna, furca and visual apparatus (Ponge et al. 2006), but also by the physiological attributes e.g. those related to moisture tolerance (Kuznetsova 2003). Soil preference identified in the study may be correlated with species dispersal from the surrounding. Therefore, dispersal ability and soil preference in the present study should be considered as a summarized response that is influenced by various functional traits of the species. Determination of these functional traits needs further measurements at individual or population level using both laboratory experiments and field observations (Pey et al. 2014, Moretti et al. 2017). Overall, our study indicates that closely related species shared similar soil preference and dispersal ability in meadow, and phylogenetic signal in these characters likely resulted from niche conservatism in morphological and/or physiological functional traits.

Hypothesis 1 - Habitat

Contrary to our Hypothesis 1, communities in meadow showed phylogenetic overdispersion, while those in forest exhibited phylogenetic clustering. These findings also contrast results presented in the previous study (**Chapter 4**) where Collembola from forests exhibited phylogenetic overdispersion, while those from grasslands showed random patterns. Presumably, the difference between the two studies was due to the different sampling designs. In the previous study, Collembola communities were sampled from replicated forest sites and grasslands, while in the present study communities of each type of habitats were replicated by soil blocks from a single site. This suggests that assembly processes of Collembola differed not only between habitats but may also be site-specific. In the present study, Collembola in forest comprised species of Isotomidae and Entomobryoidea, resulting in coexistence of phylogenetically closely related species, while meadow communities comprised Poduromorpha, Symphypleona, Neelipleona and Tomoceridae - species distantly related to each other. If traits responsible for these processes exhibited phylogenetic signal, Collembola communities of the studied forest in Morvan Regional Natural Park were likely structured by environmental filtering. Future studies need to identify the traits responsible for the observed patterns.

Hypothesis 2 - Succession

As tested in Hypothesis 1, community structures and phylogenetic relatedness of WMM and WFF did not significantly differed between the three sampling intervals, T1, T2 and T3. Temporal variations in the other treatments, therefore, reflected successional changes of Collembola communities. Consistent with our expectation, community phylogenetic relatedness in meadow as indicated by NRI decreased along successional stages when presence/absence data were used, although phylogeny-weighted structures did not differ significantly. At later stages, communities comprised the species assembled from various clades, though with low abundance. Presumably, Collembola communities in meadow needed a longer period of time, i.e., 6 months (T3), to fill the empty niches of defaunated soil blocks with distantly related species immigrated from the surrounding soil. Niche partitioning, instead of environmental filtering, was likely the predominant process at later successional stages of Collembola community in the meadow habitat.

In contrast, phylogeny-weighted community structures in forest differed between OFF at T1 and WFF, suggesting that habitat sorting was the main driver of Collembola community assembly early during succession (OFF, T1). Environmental filtering, interacting with frequent dispersal of a few species of Entomobryoidea throughout the study period, likely resulted in phylogenetic clustering in Collembola communities in forest. Interestingly, a decrease in abundance-weighted NTI at later successional stages suggests that niche partitioning gradually became stronger.

Hypothesis 3 - Soil origin

Community assembly patterns of defaunated soil blocks transferred to the respective other habitat differed from those in the original habitat, suggesting that soil properties influenced community assembly of Collembola. This is in line with the species sorting scenario of metacommunity theory (Leibold et al. 2004). However, the patterns varied with successional stages and also with the habitats where the soil blocks were inserted into. Soil originating from forest transferred to meadow was associated with Entomobryoidea. Presumably, forest soil attracted species of Entomobryoidea with some traits facilitating immigration from the surrounding meadow, resulting in different structures between OFM and OMM. Future studies need to investigate the traits corresponding to the species sorting by forest soil. Furthermore, decrease in phylogenetic relatedness in OFM suggests that environmental filtering ("species sorting") by forest soil in the meadow dominated at the beginning of succession but was less pronounced later on. Continuous immigration of species from the surrounding meadow habitat likely resulted in the "mass effect", and as a consequence, phylogenetic relatedness of OFM communities was similar to that of OMM communities at later successional stages.

The influence of species sorting in meadow soil blocks inserted into forest was weaker than that in forest soil blocks inserted into meadow, as indicated by the similarity between OMF and OFF communities at all three sampling intervals. This is further supported by the only significant difference in PCPS between OMF and OFF at T1, with no differences at later successional stages. At the beginning of succession, meadow soil blocks (OMF) were colonized predominantly by Symphypleona, while forest soil blocks (OFF) were colonized predominantly by Entomobryoidea. The lower abundance-weighted NRI of OMF at T1 than that of OFF indicates that meadow soil might hamper immigration of phylogenetically related species from forest habitat early during succession. Later, frequent immigration of species from the surrounding forest habitat likely resulted in the patterns in which OMF communities resembled OFF communities, and lower NTI (presence/absence data) in OMF communities at later successional stages. Overall, the results suggest that species sorting caused by transferring local soil likely influenced community assembly early in succession, while at later successional stages, mass effect predominantly structured Collembola communities in defaunated and transferred soil blocks.

Hypothesis 4 - Indigenous species

Community composition of Collembola in untreated forest soil blocks inserted into meadow soil (WFM) resembled those in the original forest habitat (WFF) but differed from those in the meadow (WMM). Although meadow habitat was characterized by more fluctuation in soil temperature and moistures, indigenous Collembola species of forest soil survived the translocation to meadow and might hamper colonization by meadow Collembola species. However, immigration of species from the surrounding meadow soil at T3 (e.g., Poduromorpha, Symphypleona, Tomoceridae and Neelipleona) likely resulted in resemblance between WFM and WMM communities at PCPS 1. In contrast, Collembola communities in meadow soil blocks inserted into forest soil (WMF) at T3 changed from meadow communities (WMM) to those of forest soil (WFF), indicating that indigenous meadow species could not survive, except for Isotomidae that not only remained in meadow blocks but also immigrated from the surrounding forest soil. Presumably, the change from a fluctuating to a more stable environment resulted in environmental filtering predominantly structuring Collembola communities.

Outlook

For the first time we applied community phylogenetic approaches to a field manipulative experiment on Collembola communities where the original animals were removed and recolonization of species was monitored in both native and translocated habitats characterized by different environmental variability. Assembly processes of Collembola along successional trajectories were inferred using phylogenetic relatedness as a surrogate for trait similarity between coexisting species. Since traits are objects that different processes relate to or work on, the next step is to explicitly measure functional traits including morphological, physiological as well as trophic attributes, and to test phylogenetic niche conservatism of the traits. Integrating phylogenetic comparative methods and community phylogenetic and trait-based approaches in both manipulative experiments and field observations allows a deeper understanding of the mechanisms driving and maintaining species coexistence in soil.

Author Contributions

TWC and SS conceived the ideas; AA, JFP, SB and FD collected the data; TWC and GM analyzed the data; TWC, GM and SS wrote the manuscript.

References

Anderson, J.M. (1975) The enigma of soil animal species diversity. Progress in Soil Zoology: Proceedings of the 5th International Colloquium on Soil Zoology (ed J. Vaněk), pp. 51–58. Springer Netherlands, Dordrecht.

Auclerc, A., Ponge, J.-F., Barot, S. & Dubs, F. (2009) Experimental assessment of habitat preference and dispersal ability of soil springtails. *Soil Biology and Biochemistry*, **41**, 1596–1604.

Bardgett, R.D. & van der Putten, W.H. (2014) Belowground biodiversity and ecosystem functioning. *Nature*, **515**, 505–511.

de Bello, F., Berg, M.P., Dias, A.T.C., Diniz-Filho, J.A.F., Götzenberger, L., Hortal, J., Ladle, R.J. & Lepš, J. (2015) On the need for phylogenetic "corrections" in functional trait-based approaches. *Folia Geobotanica*, **50**, 349–357.

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*, **57**, 289–300.

Bodenhofer, U., Bonatesta, E., Horejs-Kainrath, C. & Hochreiter, S. (2015) msa: An R package for multiple sequence alignment. *Bioinformatics*, **31**, 3997–3999.

Bollback, J.P. (2006) SIMMAP: stochastic character mapping of discrete traits on phylogenies. *BMC bioinformatics*, **7**, 88.

Cadotte, M., Albert, C.H. & Walker, S.C. (2013) The ecology of differences: assessing community assembly with trait and evolutionary distances. *Ecology Letters*, **16**, 1234–1244.

Caruso, T., Trokhymets, V., Bargagli, R. & Convey, P. (2013) Biotic interactions as a structuring force in soil communities: evidence from the micro-arthropods of an Antarctic moss model system. *Oecologia*, **172**, 495–503.

Cavender-Bares, J., Kozak, K.H., Fine, P. V & Kembel, S.W. (2009) The merging of community ecology and phylogenetic biology. *Ecology Letters*, **12**, 693–715.

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

Debastiani, V.J. & Duarte, L.D.S. (2014) PCPS - an R-package for exploring phylogenetic eigenvectors across metacommunities. *Frontiers of Biogeography*, **6**, 144–148.

Debastiani, V.J. & Pillar, V.D. (2012) SYNCSA--R tool for analysis of metacommunities based on functional traits and phylogeny of the community components. *Bioinformatics*, **28**, 2067–8.

Duarte, L.D.S. (2011) Phylogenetic habitat filtering influences forest nucleation in grasslands. *Oikos*, **120**, 208–215.

Duarte, L.D.S., Debastiani, V.J., Freitas, A.V.L. & Pillar, V.D. (2016) Dissecting phylogenetic fuzzy weighting: theory and application in metacommunity phylogenetics. *Methods in Ecology and Evolution*, **7**, 937–946.

Dunger, W., Schulz, H.-J., Zimdars, B. & Hohberg, K. (2004) Changes in collembolan species composition in Eastern German mine sites over fifty years of primary succession. *Pedobiologia*, **48**, 503–517.

Emerson, B.C. & Gillespie, R.G. (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology & Evolution*, **23**, 619–630.

Freckleton, R.P., Harvey, P.H. & Pagel, M. (2002) Phylogenetic analysis and comparative data: a test and review of evidence. *The American Naturalist*, **160**, 712–726.

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

Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E. & Challenger, W. (2008) GEIGER: Investigating evolutionary radiations. *Bioinformatics*, **24**, 129–131.

Huebner, K., Lindo, Z. & Lechowicz, M.J. (2012) Post-fire succession of collembolan communities in a northern hardwood forest. *European Journal of Soil Biology*, **48**, 59–65.

Huelsenbeck, J.P., Nielsen, R., Bollback, J.P. & Schultz, T. (2003) Stochastic Mapping of Morphological Characters. *Systematic Biology*, **52**, 131–158.

Ingimarsdóttir, M., Caruso, T., Ripa, J., Magnusdottir, O.B., Migliorini, M. & Hedlund, K. (2012) Primary assembly of soil communities: disentangling the effect of dispersal and local environment. *Oecologia*, **170**, 745–754.

Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P. & Webb, C.O. (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, **26**, 1463–4.

Kuznetsova, N. (2003) Humidity and distribution of springtails. *Entomological Review*, **83**, 230–238.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M. & Gonzalez, A. (2004) The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, **7**, 601–613.

Letten, A.D., Keith, D.A. & Tozer, M.G. (2014) Phylogenetic and functional dissimilarity does not increase during temporal heathland succession. *Proceedings of the Royal Society B*, **281**, 2014102.

Li, J., Li, S., Chen, Y., Jia, P., Hua, Z., Wang, S., Song, Y., Liao, B. & Shu, W. (2014) Phylogenetic structures of soil nematode communities along a successional gradient in an unreclaimed copper mine tailings site. *Soil Biology and Biochemistry*, **77**, 179–186.

Lindberg, N. & Bengtsson, J. (2005) Population responses of oribatid mites and collembolans after drought. *Applied Soil Ecology*, **28**, 163–174.

Liu, H., Xu, Q., He, P., Santiago, L.S., Yang, K. & Ye, Q. (2015) Strong phylogenetic signals and phylogenetic niche conservatism in ecophysiological traits across divergent lineages of Magnoliaceae. *Scientific Reports*, **5**, 12246.

Moretti, M., Dias, A.T.C., de Bello, F., Altermatt, F., Chown, S.L., Azcarate, F.M., Bell, J.R., Fournier, B., Hedde, M., Hortal, J., Ibanez, S., Öckinger, E., Sousa, J.P., Ellers, J. & Berg, M.P. (2017) Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. *Functional Ecology*, **31**, 558–567.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H. & Wagner, H. (2015) vegan: community ecology package.

Pagel, M. (1999) Inferring the historical patterns of biological evolution. *Nature*, **401**, 877–884.

Paradis, E. (2013) Molecular dating of phylogenies by likelihood methods: A comparison of models and a new information criterion. *Molecular Phylogenetics and Evolution*, **67**, 436–444.

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

Perez, G., Decaëns, T., Dujardin, G., Akpa-Vinceslas, M., Langlois, E. & Chauvat, M. (2013) Response of collembolan assemblages to plant species successional gradient. *Pedobiologia*, **56**, 169–177.

Petersen, H. & Luxton, M. (1982) A comparative analysis of soil fauna populations and their role in decomposition processes. *Oikos*, **39**, 288–388.

Pey, B., Nahmani, J., Auclerc, A., Capowiez, Y., Cluzeau, D., Cortet, J., Decaëns, T., Deharveng, L., Dubs, F., Joimel, S., Briard, C., Grumiaux, F., Laporte, M.-A.A., Pasquet, A., Pelosi, C., Pernin, C., Ponge, J.-F., Salmon, S., Santorufo, L. & Hedde, M. (2014) Current use of and future needs for soil invertebrate functional traits in community ecology. *Basic and Applied Ecology*, **15**, 194–206.

Pillar, V.D. & Duarte, L.D.S. (2010) A framework for metacommunity analysis of phylogenetic structure. *Ecology Letters*, **13**, 587–596.

Ponge, J.-F., Dubs, F., Gillet, S., Sousa, J. & Lavelle, P. (2006) Decreased biodiversity in soil springtail communities: the importance of dispersal and landuse history in heterogeneous landscapes. *Soil Biology and Biochemistry*, **38**, 1158–1161.

Purschke, O., Schmid, B.C., Sykes, M.T., Poschlod, P., Michalski, S.G., Durka, W., Kühn, I., Winter, M. & Prentice, H.C. (2013) Contrasting changes in taxonomic, phylogenetic and functional diversity during a long-term succession: Insights into assembly processes. *Journal of Ecology*, **101**, 857–866.

Revell, L.J. (2012) phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, **3**, 217–223.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A. & Huelsenbeck, J.P. (2012) Mrbayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.

Rusek, J. (1998) Biodiversity of Collembola and their functional role in the ecosystem. *Biodiversity and Conservation*, **7**, 1207–1219.

Salmon, S., Ponge, J.-F., Gachet, S., Deharveng, L., Lefebvre, N. & Delabrosse, F. (2014) Linking species, traits and habitat characteristics of Collembola at European scale. *Soil Biology and Biochemistry*, **75**, 73–85.

Schaefer, M., Migge-Kleian, S. & Scheu, S. (2009) The role of soil fauna for decomposition of plant residues. *Functioning and Management of European Beech Ecosystems* (eds R. Brumme & P.K. Khanna), pp. 207–230. Springer Berlin Heidelberg, Berlin, Heidelberg.

Tsiafouli, M.A., Kallimanis, A.S., Katana, E., Stamou, G.P. & Sgardelis, S.P. (2005) Responses of soil microarthropods to experimental short-term manipulations of soil moisture. *Applied Soil Ecology*, **29**, 17–26.

Vaidya, G., Lohman, D.J. & Meier, R. (2011) SequenceMatrix: Concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, **27**, 171–180.

Vellend, M. (2010) Conceptual synthesis in community ecology. *The Quarterly Review of Biology*, **85**, 183–206.

Vellend, M. (2016) *The Theory of Ecological Communities*. Princeton University Press, Princeton.

Webb, C.O. (2000) Exploring the Phylogenetic Structure of Ecological Communities: An Example for Rain Forest Trees. *The American Naturalist*, **156**, 145–155.

Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, **33**, 475–505.

Wright, E.S. (2015) DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics*, **16**, 322.

Supplementary Materials

Table S5.1

Hypotheses tested in this study and definition of metacommunity (species pool) for each hypothesis. Treatments in comparison are marked in bold. Asterisks in treatments indicate the focused treatments. For abbreviation of treatments see Table 5.1.

Hypotheses	Metacommunity (species pool)	
	Definition (number of species)	**Treatments**
(1) **Habitat and temporal effects:** Community structures differ between meadow and forest but not between sampling time. Communities in meadow exhibit phylogenetic clustering due to environmental filtering resulting from disturbances, while forest communities show phylogenetic overdispersion resulting from niche partitioning in a relatively stable environment.	**Reference:** Untreated soil blocks installed in original habitats (49)	**WMM** **WFF**
(2) **Successional patterns:** At early successional stages drift predominates, resulting in a random pattern of phylogenetic relatedness in defaunated blocks, while selection drives communities to a deterministic pattern at later successional stages.	**M:** Defaunated meadow soil blocks + untreated meadow soil blocks (32)	**T1** - OMM **T2** - OMM **T3** - OMM **WMM** (all time)
	F: Defaunated forest soil blocks + untreated forest soil blocks (40)	**T1** - OFF **T2** - OFF **T3** - OFF **WFF** (all time)
(3) **Soil origin effects:** Successional patterns in defaunated soil blocks differ between the transferred soil blocks and the soil blocks installed in the original habitat.	**OM:** Defaunated soil blocks from different origins surrounded by meadow + untreated meadow soil blocks (32)	**OFM*** **OMM** WMM
	OF: Defaunated soil blocks from different origins surrounded by forest + untreated forest soil blocks (44)	**OMF*** **OFF** WFF
(4) **Indigenous species:** Community structures in untreated but transferred soil blocks change gradually from those in soil blocks of the original habitat to those in the transferred habitat.	**FM:** Transferred forest soil blocks installed in meadow + untreated forest and meadow soil blocks (49)	WFF **WFM*** **OFM** WMM
	MF: Transferred meadow soil blocks installed in forest + untreated meadow and forest soil blocks (49)	WMM **WMF*** **OMF** WFF

Table S5.2

Collembola species reported in Auclerc et al. (2009) and species used for Collembola phylogenetic reconstruction in this study with Accession number in Genbank or local database. For species without sequences available, sequences of congeneric species were used. When sequences of congeneric species were also lacking, taxa were grafted to the most recent common ancestors according to the taxonomy. Habitat (land-use) preference, dispersal ability in meadow and in forest and soil preference of Collembola species reported in Auclerc et al (2009) are listed. Habitat preference: F, forest-specialist; FP, forest-preferring species; M, meadow-specialist; MP, meadow-preferring species; G, land-use generalist. Dispersal ability in meadow: M1, species in defaunated meadow blocks after one week; M2, species in defaunated meadow blocks after one month; M3, species in defaunated meadow blocks after six months; M4, species which did not colonize defaunated meadow blocks within six months; M0, species absent in the meadow. Dispersal ability in forest: F1, species in defaunated forest blocks after one week; F2, species in defaunated forest blocks after one month; F3, species in defaunated forest blocks after six months; F4, species which did not colonize defaunated forest blocks within six months; F0, species absent in the forest. Soil preference: FS, forest-soil-preferring species; MS, meadow-soil-preferring species; SG, soil-generalist species. NR, not sufficient specimens to define soil preference.

Species arrangement			**Accession number in Genbank/local database**						**Traits**			
Species reported in Auclect et al. (2009)	**Species used in phylogenetic reconstruction**	**Replaced by**	**18S**	**28S D1**	**28S D2**	**28S D3-5**	**H3**	**COI**	**Habitat (land-use) preference**	**Dispersal ability in meadow**	**Dispersal ability in forest**	**Soil preference**
	Archaeognatha (Outgroup)		AY338689	AY338646	AY338646	AY338646	AY338614	JF826083				
	Palaeoptera (Outgroup)		AF370791	AY859557	AY859557	AY859557	AY749703	GU711466				
	Zygentoma (Outgroup)		EU368615	EU376048	EU376048	EU376048	AY555568	JN970940				
Allacma fusca	*Allacma fusca*		JC372	JC372	JC372	JC372		JC372	F	M0	F4	NR
Arrhopalites principalis	*Pygmarrhopalites principalis*	Synonym		JC300	JC300	JC300	JC300		F	M0	F4	MS
Arrhopalites sericus	*Pygmarrhopalites sericus*	Synonym		JC299	JC299	JC299			FP	M4	F1	MS
Ceratophysella denticulata	*Ceratophysella denticulata*		JC390	JC390	JC390	JC390	JC390	JC390	MP	M1	F4	MS
Ceratophysella recta	*Ceratophysella luteospina*	Congener	JC416	JC416	JC416	JC416		JC416	FP	M4	F3	MS
Deuteraphorura inermis	*Deuteraphorura inermis*		JC415	JC415	JC415	JC415	JC496	JC415	FP	M4	F1	MS
Deuterosminthurus sulphureus	*Deuterosminthurus sulphureus*		JC369	JC369	JC369	JC369	JC369	JC369	M	M3	F0	MS
Dicyrtomina minuta	*Dicyrtomina minuta*		JC431	JC431	JC431	JC431	JC431	JC431	F	M0	F1	FS
Entomobrya multifasciata	*Entomobrya multifasciata*		KC236237	KC236276	KC236276	KC236276		KM610123	F	M3	F4	NR
Folsomia listeri	*Folsomia candida*	Synonym	JC404	JC404	JC404	JC404	JC404		M	M4	F0	NR
Folsomia manolachei	*Folsomia manolachei*		JC317	JC317	JC317	JC317	JC317	JC317	M	M1	F0	MS
Folsomia quadrioculata	*Folsomia quadrioculata*		JC403	JC403	JC403	JC403	KJ186565	KT808344	FP	M1	F1	MS
Friesea truncata	*Friesea claviseta*	Congener	JC363	JC363	JC363	JC363	JC363	JC363	F	M0	F1	SG
Gisinianus flammeolus		Ancestors at family							F	M0	F4	NR
Heteromurus nitidus	*Heteromurus nitidus*		JC454	JC329	JC329	JC329	JC329	JC329	M	M1	F0	MS
Isotoma anglicana	*Isotoma anglicana*		JC308	JC308	JC308	JC308	JC308	JC308	MP	M1	F4	MS
Isotomiella minor	*Isotomiella minor*		JC386	JC386	JC386	JC386	JC386	JC386	F	M0	F1	MS
Lepidocyrtus cyaneus	*Lepidocyrtus cyaneus*		JC396	JC396	JC396	JC396	JC396	JC396	M	M1	F0	SG
Lepidocyrtus lanuginosus	*Lepidocyrtus lanuginosus*		JC391	JC391	JC391	JC391	JC391	JC392	G	M1	F1	SG
Lepidocyrtus lignorum	*Lepidocyrtus lignorum*		JC393	JC393	JC393	JC393	JC393	JC393	G	M1	F1	SG
Lipothrix lubbocki	*Lipothrix lubbocki*			LK024363	LK024363			LK024485	F	M0	F2	FS
Megalothorax minimus	*Megalothorax minimus*		JC502	JC502	JC502	JC502	JC502	KC900195	MP	M1	F3	MS

Table S5.2 Continued

Species arrangement			Accession number in Genbank/local database						Traits			
Species reported in Auclect et al. (2009)	**Species used in phylogenetic**	**Replaced by**	**18S**	**28S D1**	**28S D2**	**28S D3-5**	**H3**	**COI**	**at (land-use) pr**	**al ability in**	**rsal ability in**	**oil preferenc**
Mesaphorura florae	*Mesaphorura sp4*	Congener	JC412	JC412	JC412	JC412			G	M3	F0	MS
Mesaphorura macrochaeta	*Mesaphorura sp1*	Congener	JC478	JC366	JC366	JC366	JC366		M	M2	F1	MS
Micranurida pygmaea	*Micranurida pygmaea*		JC352	JC352	JC352	JC352		JC352	MP	M4	F3	NR
Micraphorura absoloni	*Micraphorura absoloni*			JC434	JC434	JC434	JC434	JC434	F	M0	F4	NR
Neanura muscorum	*Neanura muscorum*		AY555520	AJ251733	AJ251733	AJ251733		AY555544	F	M0	F3	SG
Orchesella cincta	*Orchesella cincta*		KC236250	KC236290	KC236290	KC236290		EU869805	F	M0	F3	MS
Paratullbergia callipygos	*Paratullbergia sp*	Congener	JC479	JC479	JC479				F	M0	F4	NR
Parisotoma notabilis	*Parisotoma notabilis*		JC438	JC438	JC438	JC438	JC438	JC438	MP	M1	F1	SG
Pogonognathellus flavescens	*Pogonognathellus flavescens*		JC338	JC338	JC338	JC338	JC338	JC338	F	M0	F1	SG
Protaphorura aurantiaca	*Protaphorura campata*	Congener	JC433	JC433	JC433	JC433	JC433	JC433	MP	M1	F4	SG
Pseudachorutes parvulus	*Pseudachorutes sp*	Congener	JC401	JC491	JC491	JC401	JC491	JC401	F	M0	F1	SG
Pseudosinella alba	*Pseudosinella alba*		JC327	JC327	JC327	JC327	JC327	KM978368	MP	M1	F4	SG
Pseudosinella terricola	*Pseudosinella immaculata*	Congener	JC328			JC328	JC328	JC328	F	M0	F1	SG
Sminthurides parvulus	*Sminthurides signatus*	Congener	JC281			JC281	JC281		M	M3	F0	SG
Sminthurides schoetti	*Sminthurides schoetti*		JC463			JC463			MP	M2	F4	MS
Sminthurinus aureus	*Sminthurinus aureus*		JC382	JC382	JC382	JC382	JC382	JC382	MP	M1	F1	MS
Sminthurinus signatus	*Sminthurinus elegans*	Synonym	JC469	JC469	JC469	JC469	JC469		F	M0	F1	SG
Sminthurus viridis	*Sminthurus viridis*		JC293	JC465	JC465	JC466	JC293	JN970939	M	M1	F0	MS
Sphaeridia pumilis	*Sphaeridia pumilis*		JC470			JC470		JC470	M	M4	F3	FS
Stenaphorura denisi	*Stenaphorura denisi*		JC347	JC347	JC347	JC347		JC476	M	M4	F0	NR
Stenognathellus denisi		Ancestors at family							F	M0	F1	FS
Subisotoma pusilla	*Subisotoma sp*	Congener						DQ365795	F	M3	F4	NR
Vertagopus arboreus	*Vertagopus arboreus*		JC450	JC450	JC450	JC450	JC450	JC450	FP	M1	F4	MS
Willemia anophthalma	*Willemia anophthalma*		JC358	JC435	JC435	JC435		JC358	FP	M4	F2	FS
Willemia denisi	*Willemia denisi*		JC408	JC408	JC408	JC408		JC408	FP	M4	F3	MS
Xenylla grisea	*Xenylla grisea*		AY555517	HQ732134	HQ732026	HQ731974	AY555563	HQ732087	F	M0	F1	SG
Xenylla tullbergi	*Xenylla tullbergi*			AF483410	AF483468	HQ731980		HQ732088	FP	M4	F1	SG

Table S5.3

Estimated number of transitions across 1,000 stochastic character mapping simulations of soil preferences of Collembola reported in Auclerc et al. (2009).

To:	Meadow-soil-preferring		Soil-generalist		Forest-soil-preferring	
From:	**Median**	**Mean ± SD**	**Median**	**Mean ± SD**	**Median**	**Mean ± SD**
Meadow-soil-preferring			**8**	7.6 ± 2.0	**5**	5.1 ± 1.9
Soil-generalist	**3**	4.2 ± 3.4			**1**	2.0 ± 2.0
Forest-soil-preferring	**2**	2.7 ± 2.7	**1**	1.9 ± 2.0		

Table S5.4

Estimated number of transitions across 1,000 stochastic character mapping simulations of dispersal ability in meadow of Collembola reported in Auclerc et al. (2009). Dispersal ability in meadow: M1, species in defaunated meadow blocks after one week; M2, species in defaunated meadow blocks after one month; M3, species in defaunated meadow blocks after six months; M4, species which did not colonize defaunated meadow blocks within six months; M0, species absent in the meadow.

To:	M1		M2		M3		M4		M0	
From:	**Median**	**Mean ± SD**	**Median**	**Mean ± SD**	**Median**	**Mean ± SD**	**Median**	**Mean ± SD**	**Median**	**Mean ± SD**
M1			**0**	0.5 ± 0.9	**2**	2.0 ± 1.9	**2**	2.9 ± 3.0	**9**	10.1 ± 5.6
M2	**0**	0.4 ± 0.9			**1**	1.5 ± 1.5	**0**	0.6 ± 1.2	**1**	1.1 ± 1.7
M3	**0**	1.0 ± 1.7	**1**	1.4 ± 1.2			**1**	1.3 ± 1.9	**1**	1.4 ± 2.2
M4	**2**	2.7 ± 2.9	**0**	0.8 ± 1.2	**2**	2.1 ± 2.0			**8**	9.2 ± 6.7
M0	**8**	8.5 ± 4.8	**1**	1.3 ± 1.5	**2**	2.1 ± 2.1	**8**	9.1 ± 5.3		

Table S5.5

Effects of treatment and sampling interval on Collembola community compositions and phylogeny-weighted community compositions based on Bray-Curtis dissimilarity tested by distance-based multivariate analysis of variance (ADONIS) with 9,999 permutations. For definitions of pool see Table S5.1.

Hypotheses	Pool	Community composition												Phylogeny-weighted community composition											
		Treatment				**Time** (sampling interval)				**Treatment:Time**				**Treatment**				**Time** (sampling interval)				**Treatment:Time**			
		df	**F**	**R²**	***P***	**df**	**F**	**R²**	***P***	**df**	**F**	**R²**	***P***	**df**	**F**	**R²**	***P***	**df**	**F**	**R²**	***P***	**df**	**F**	**R²**	***P***
(1) Habitat effects	Reference	1	18.5	**0.39**	**0.000**	2	1.4	0.06	0.182	2	0.9	0.04	0.516	1	12.6	**0.31**	**0.000**	2	1.1	0.05	0.339	2	1.2	0.06	0.314
(2) Succession patterns	M	3	4.0	**0.32**	**0.000**	-	-	-	-	-	-	-	-	3	1.6	**0.16**	0.120	-	-	-	-	-	-	-	-
	F	3	9.4	**0.52**	**0.000**	-	-	-	-	-	-	-	-	3	4.6	**0.35**	**0.002**	-	-	-	-	-	-	-	-
(3) Soil origin effects	OM	2	3.2	**0.10**	**0.001**	2	6.8	**0.22**	**0.000**	4	1.7	**0.11**	**0.020**	2	5.8	**0.21**	**0.000**	2	2.2	0.08	0.056	4	1.1	0.08	0.379
	OF	2	7.1	**0.16**	**0.000**	2	13.8	**0.31**	**0.000**	4	2.7	**0.12**	**0.002**	2	8.4	**0.24**	**0.000**	2	2.4	**0.07**	**0.038**	4	2.8	**0.16**	**0.004**
(4) Indigenous species	FM	3	9.4	**0.31**	**0.000**	2	2.8	**0.06**	**0.001**	6	1.7	**0.11**	**0.011**	3	7.1	**0.27**	**0.000**	2	1.0	0.02	0.434	6	1.3	0.10	0.231
	MF	3	12.4	**0.34**	**0.000**	2	5.6	**0.10**	**0.000**	6	2.2	**0.12**	**0.000**	3	8.6	**0.29**	**0.000**	2	1.9	0.04	0.086	6	2.0	**0.13**	**0.017**

Table S5.6

Treatment effects on axes of principle coordinates of Collembola community and phylogeny-weighted community compositions based on Bray-Curtis dissimilarity. Krustak-Wallis rank sum test was performed to compare community scores at each axis between treatments. Only significant axes are reported with their relative eigenvalues. Asterisks in treatments indicate the focused treatments compared with the other treatments. Letters in lowercase indicate significant differences between treatments ($P < 0.05$, Dunn's test with BH adjustment) with the focused comparisons marked in bold that answer the question. For hypotheses (H) and definitions of pool see Table S5.1; for abbreviation of treatments see Table 5.1.

H	Pool	Statistics	Treatments	Time	Principle coordinates of community compositions (PCoA)								Principal coordinates of phylogenetic structures (PCPS)								
					Axis.1	**Axis.2**	**Axis.3**	**Axis.4**	**Axis.5**	**Axis.6**	**Axis.9**	**Axis.15**	**Axis.1**	**Axis.2**	**Axis.3**	**Axis.4**	**Axis.5**	**Axis.6**	**Axis.7**	**Axis.9**	**Axis.15**
(1)	Reference	*P*-value			0.000								0.000		0.014		0.017				
		Eigenvalue (%)			41.4%								57.9%		17.2%		3.3%				
		Dunn's test	WMM		**a**								**a**		**a**		**a**				
			WFF		**b**								**b**		**b**		**b**				
(2)	M	*P*-value			0.004		0.001					0.016				0.002				0.039	
		Eigenvalue (%)			35.4%		12.4%					0.8%				8.6%				1.1%	
		Dunn's test	T1 - OMM		**a**		**a**					**a**				**a**				ab	
			T2 - OMM		**ab**		**a**					**a**				**ab**				**a**	
			T3 - OMM		**c**		**a**					**b**				**bc**				ab	
			WMM		**b**		**b**					**a**				**c**				**b**	
	F	*P*-value			0.000	0.003					0.026		0.011								
		Eigenvalue (%)			51.6%	11.2%					2.3%		71.7%								
		Dunn's test	T1 - OFF		**a**	**a**					**a**		**a**								
			T2 - OFF		**a**	**b**					**b**		ab								
			T3 - OFF		**b**	**a**					**ab**		ab								
			WFF		**b**	**b**					**a**		**b**								

Table S5.6 Continued

H	Pool	Statistics	Treatments	Time	Principle coordinates of community compositions (PCoA)								Principal coordinates of phylogenetic structures (PCPS)								
					Axis.1	**Axis.2**	**Axis.3**	**Axis.4**	**Axis.5**	**Axis.6**	**Axis.9**	**Axis.15**	**Axis.1**	**Axis.2**	**Axis.3**	**Axis.4**	**Axis.5**	**Axis.6**	**Axis.7**	**Axis.9**	**Axis.15**
(3)	OM	*P*-value			0.001		0.006	0.022					0.023	0.014		0.004					
		Eigenvalue (%)			33.0%		9.5%	8.1%					54.5%	20.7%		6.8%					
		Dunn's test	OFM*	T1	**a**		ab	ab					a	ab		ab					
				T2	abc		a	**a**					a	a		abc					
				T3	**bd**		abc	ab					a	ab		abc					
			OMM	T1	ac		abc	ab					a	ab		a					
				T2	abc		abc	**b**					a	ab		ab					
				T3	d		abc	ab					a	b		abc					
			WMM	T1	bcd		c	a					a	b		c					
				T2	bcd		c	ab					a	ab		abc					
				T3	bd		bc	ab					a	b		bc					
	OF	*P*-value			0.000	0.032	0.023			0.001			0.011	0.007				0.035			
		Eigenvalue (%)			49.7%	12.9%	6.7%			4.6%			51.7%	30.7%				1.9%			
		Dunn's test	OMF*	T1	**a**	a	ab			abc			ab	**b**				ab			
				T2	**ab**	a	b			b			abc	b				ab			
				T3	**c**	a	ab			bc			bc	ab				b			
			OFF	T1	a	a	a			a			a	**a**				a			
				T2	ab	a	ab			bc			ab	ab				ab			
				T3	c	a	a			bc			abc	a				b			
			WFF	T1	bc	a	ab			a			c	ab				ab			
				T2	bc	a	ab			ac			bc	ab				ab			
				T3	c	a	a			abc			abc	a				b			

Table S5.6 Continued

H	Pool	Statistics	Treatments	Time	Principle coordinates of community compositions (PCoA)								Principal coordinates of phylogenetic structures (PCPS)								
					Axis.1	Axis.2	Axis.3	Axis.4	Axis.5	Axis.6	Axis.9	Axis.15	Axis.1	Axis.2	Axis.3	Axis.4	Axis.5	Axis.6	Axis.7	Axis.9	Axis.15
(4)	FM	*P*-value			0.000	0.006		0.000			0.012		0.002		0.011		0.009	0.003	0.047		
		Eigenvalue (%)			28.3%	20.6%		7.1%			3.9%		51.2%		18.4%		3.3%	2.5%	2.1%		
		Dunn's test	WFF	T1	d	abc		bcd			ab		abcd		ab		b	ab	a		
				T2	d	abc		abcd			a		abcd		a		ab	b	a		
				T3	cd	c		**abc**			c		**a**		ab		ab	b	a		
			WFM*	T1	**bcd**	**bc**		**bcd**			ab		bcd		ab		ab	ab	a		
				T2	**bcd**	abc		**d**			abc		abcd		ab		ab	b	a		
				T3	**bcd**	abc		**d**			abc		**bcd**		ab		**b**	ab	a		
			OFM	T1	ab	**a**		**a**			abc		abc		a		ab	a	a		
				T2	abc	ab		**ab**			abc		ab		a		ab	ab	a		
				T3	**a**	c		bcd			ab		abcd		ab		**a**	a	a		
			WMM	T1	**a**	abc		d			bc		cd		b		b	ab	a		
				T2	**a**	abc		d			ab		d		ab		ab	ab	a		
				T3	**a**	bc		cd			abc		d		ab		**a**	ab	a		
	MF	*P*-value			0.000	0.000	0.037		0.020	0.033			0.001	0.019	0.033		0.027		0.040		0.046
		Eigenvalue (%)			29.8%	21.4%	10.8%		6.2%	4.4%			53.3%	26.9%	14.3%		3.5%		2.2%		0.4%
		Dunn's test	WMM	T1	bc	abcd	ab		ab	**b**			bcd	ab	b		b		b		ab
				T2	c	abcd	b		ab	ab			d	ab	ab		ab		ab		ab
				T3	bc	bcd	ab		b	ab			cd	ab	ab		ab		ab		ab
			WMF*	T1	**bc**	abcd	ab		ab	**a**			**cd**	**b**	ab		ab		ab		ab
				T2	**bc**	**abc**	ab		a	ab			bcd	ab	ab		ab		a		**b**
				T3	abc	**d**	ab		ab	ab			abc	ab	ab		ab		ab		ab
			OMF	T1	ab	a	a		ab	a			**a**	**a**	ab		a		ab		ab
				T2	**a**	ab	ab		ab	ab			ab	ab	ab		a		ab		ab
				T3	a	cd	ab		ab	ab			ab	ab	ab		ab		ab		ab
			WFF	T1	**a**	abcd	b		ab	a			abcd	b	ab		ab		ab		ab
				T2	**a**	bcd	b		ab	a			abcd	ab	a		ab		ab		**a**
				T3	a	cd	ab		ab	ab			a	ab	a		ab		ab		ab

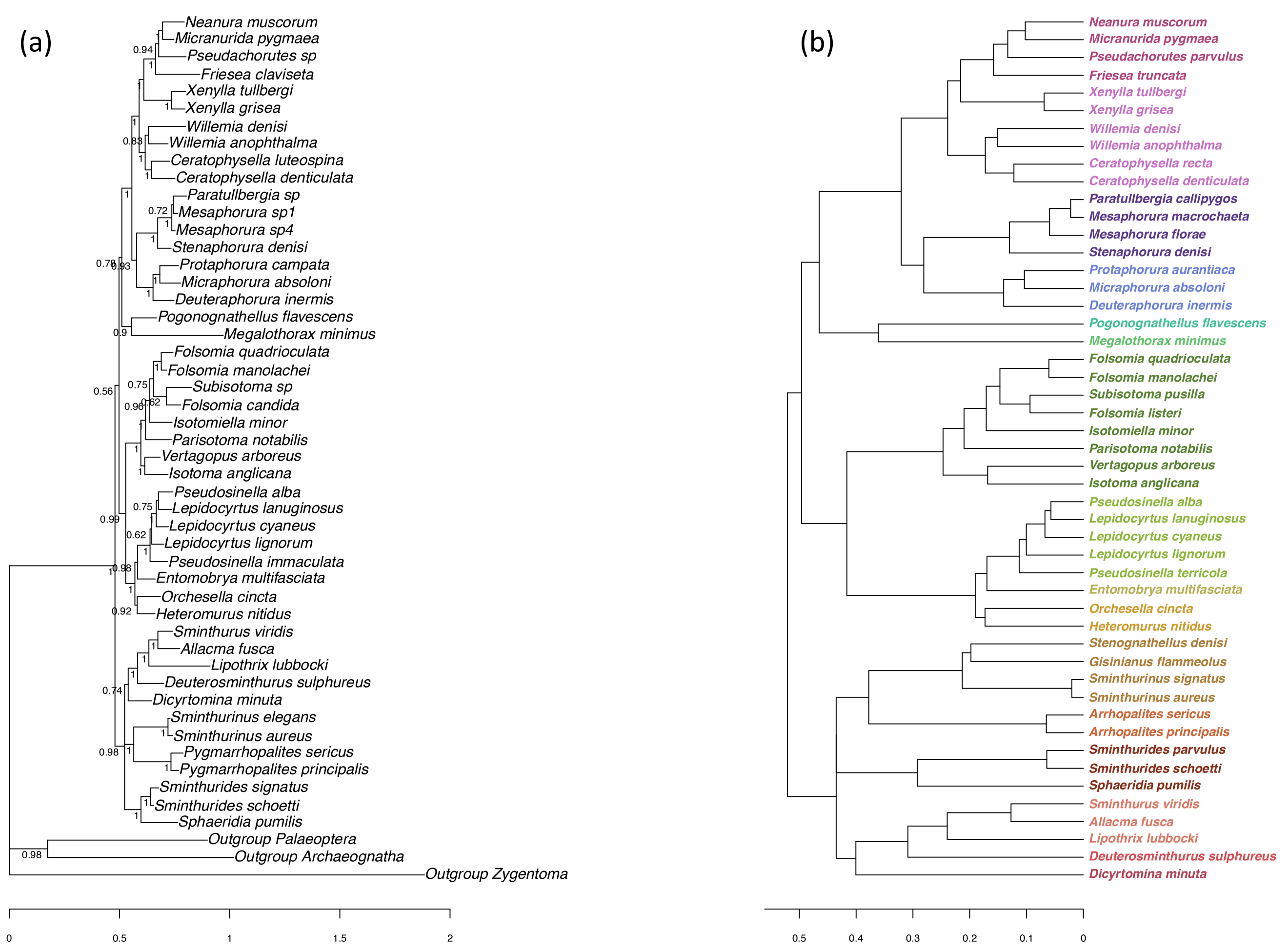

Figure S5.1
(a) Bayesian tree of Collembola species studied. The phylogeny was based on a concatenated alignment of 18S and 28S rRNA, Histone H3 and COI genes (3,073 bp). Numbers at nodes represent Bayesian posterior probabilities. Zygentoma (Insecta), *Machilis* (Archaeognatha, Insecta) and *Callibaetis* (Palaeoptera, Insecta) were used as outgroups. (b) Ultrametric tree of Collembola species used in the trait and community analyses. The ultrametric tree was generated based on the Bayesian tree using penalized likelihood that assumes a strict clock model with substitution rates varying among branches. *Gisinianus flammeolus* and *Stenognathellus denisi* were grafted to the most recent common ancestors according to the family. Different coloring indicates different Collembola families.

Chapter 6

General Discussion

This thesis is the first attempt to apply phylogenetic approaches to elucidate assembly processes of soil Collembola communities. By integrating community phylogenetic approaches with phylogenetic comparative and trait-based methods, this thesis presents major advances in understanding biodiversity of soil animals from both ecological and evolutionary perspectives. The conceptual model proposed in **Chapter 1** (**Figure 1.1**) can be applied to any other soil taxon and thus provides a framework for future studies.

In this thesis, I reconstructed the evolutionary history of several traits in Collembola, including morphological characters (**Chapter 2**), ecological preferences (**Chapters 2 and 5**), physiological attributes and food resources (**Chapter 3**). Phylogenetic relatedness and trait similarity of species within local communities (α-diversity) were used to infer assembly processes, given the assumption of phylogenetic signal in ecologically relevant traits. A phylogenetic β-diversity approach was used to elucidate community-environmental associations from an evolutionary perspective (**Chapters 4 and 5**). Furthermore, the community phylogenetic approaches were applied to both field observational (**Chapter 4**) and experimental (**Chapter 5**) studies.

In the following sections, I first ascribe the patterns of Collembola communities found in previous chapters to the scenarios in the conceptual model (**Figure 1.1**) by specifically exploring trait evolution, trait similarity and phylogenetic relatedness between coexisting species, then discuss the likely processes referring to the four high-level processes, selection, dispersal, drift and speciation, proposed in *The Theory of Ecological Communities* (Vellend 2010, 2016). Finally, I supply a roadmap for soil ecologists to integrate phylogenetic comparative methods, community phylogenetic analyses and trait-based approaches in studies on the assembly processes of soil communities.

Assembly processes of soil Collembola communities

Disturbance as a selecting factor

Phylogenetic clustering is revealed in the abundant Collembola species inhabiting arable fields near Göttingen (**Chapter 4)**, following scenario (a) in which environmental filtering is a predominant process, if the traits underlying community assembly processes are conserved relative to the ancestor and thus exhibit phylogenetic signal, or scenario (c) in which niche partitioning structures communities, if the niche traits diverged from the ancestral state (**Figure 1.1**). Traits such as body length, vertical stratification, pigmentation, number of ommatidia and reproductive mode are similar between phylogenetically related species and evolution of body shape was constrained (**Chapter 2**), suggesting that scenario (a) is more likely. As a logical consequence, traits between coexisting species should be more similar than that predicted by the null model. However, randomness in trait similarity suggests that while some of the examined traits are filtered by the environment, other traits are likely driven by other factors from the opposite direction. For example, niche partitioning may structure coexisting species which differed in traits, such as vertical stratification, pigmentation, number of ommatidia and reproductive mode, that co-evolved during Collembola diversification, and thus likely diminish the effect of environmental filtering, shifting the traits from similar to random patterns. Nevertheless, phylogenetic clustering in Collembola communities collected from arable fields is consistent with patterns in other taxa inhabiting disturbed environments (Ding et al. 2012, Pellissier et al. 2013, Gianuca et al. 2014) and with findings from trait-based analysis of Collembola communities inhabiting salt marshes (Widenfalk et al. 2015). Overall, environmental filtering is likely the predominant process in soil animal communities in habitats where disturbance may result in homogeneity of soil properties (Maaß et al. 2014).

Phylogenetic overdispersion is found in the forest Collembola communities collected in Göttingen (**Chapter 4**). Again, traits between coexisting species show a random pattern. Collembola communities in forests thus follow scenario (b) or (d) (**Figure 1.1**), depending on how the process-relevant traits have evolved and how they are structured by different processes. If the traits show phylogenetic signal, coexistence of distantly related species may result in different traits in local communities. Randomness in trait patterns, thus suggests that environmental filtering may still work but with limited influence, while niche partitioning is the predominant process working on Collembola living in forest soils. High spatial heterogeneity in forest soils likely results in different traits of communities (Maaß et al. 2014, Widenfalk et al. 2016). Also, soil food webs in stable habitats such as forests are complex (Scheu and Falca 2000, Digel et al. 2014). Given that Collembola species span several trophic levels (Chahartaghi et al. 2005, Pollierer et al. 2009) with taxonomically related species occupying similar trophic levels (Potapov et al. 2016), communities comprised of species from a variety of phylogenetic clades may be able to exploit a wide range of food resources and thus facilitate coexistence in forests. Overall, niche partitioning in micro-habitats and food resources is likely a predominant process in Collembola communities in stable habitats, although environmental filtering working on phylogenetically convergent traits may produce similar patterns.

The general hypothesis that Collembola communities in disturbed habitats are determined by environmental filtering, while in relatively stable habitats interspecific competition/niche partitioning is predominant, is supported by phylogenetic relatedness but not by trait similarity (**Chapter 4**), even though the tested traits exhibit phylogenetic signal (**Chapter 2**). In contrast, Collembola communities in Morvan Regional Natural Park (**Chapter 5, Hypothesis 1**) show phylogenetic clustering in forest but overdispersion in meadow habitats. Discrepancies between the two studies may be due to the sampling design. In **Chapter 4**, Collembola communities were sampled from replicated forest sites and grasslands, while in **Chapter 5** communities of each type of habitat were replicated by soil blocks from a single site. Assembly processes of Collembola likely differ not only between habitats but may also be site-specific. Furthermore, successional stages of communities may also influence relative strengths of the forces driving community assembly, as demonstrated in **Chapter 5** (**Hypothesis 2**) in which niche partitioning becomes stronger at later successional stages and may balance the effects of environmental filtering. Overall, the results of this thesis indicate that Collembola communities are structured by selection processes such as environmental filtering and niche partitioning, which vary in different habitats and with different successional stages.

Dispersal as a high-level process

Data on successional trajectories in the manipulative experiment (Auclerc et al. 2009; **Chapter 5**) show that species dispersal, interacting with selection processes, determine community composition of Collembola. Results indicate that soil properties of the defaunated blocks influence community assembly of Collembola at initial stages of succession, while at later stages community assembly is dominated by mass effects due to continuous immigration of species from the surrounding habitat (**Hypothesis 3**). As a consequence of dispersal (dispersal as a high-level process), community compositions change gradually from those resembling the original habitats to those of the new habitats (**Hypothesis 4**). These results are consistent with previous findings at different spatial scales, from plot (Åström and Bengtsson 2011) to landscape (Ingimarsdóttir et al. 2012), reemphasizing the importance of mass effects (i.e., consequences of dispersal) on Collembola community compositions. Collembola, considered a fast disperser among the soil animals (but see Ojala and Huhta 2001), can disperse actively at small spatial scales (Bengtsson et al. 1994) but may also be transmitted via other vectors over long geographical distances (Costa et al. 2013). Furthermore, dispersal ability of Collembola is likely related to their life forms, such as surface-living (epedaphic) or soil-dwelling (euedaphic) (Bengtsson et al. 1994, Hågvar 2000, Ojala and Huhta 2001, Zhang et al. 2017), and to food availability and quality in the habitats (Bengtsson et al. 1994, Stötefeld et al. 2012).

Stochasticity in communities—ecological drift

One prediction of *The Theory of Ecological Communities* (Vellend 2010, 2016) is that the signature of ecological drift as a high-level process on community assembly is random patterns of traits. In **Chapter 4**, Collembola collected from the grasslands and those dwelling in soil (euedaphic) exhibit random patterns in both trait similarity and phylogenetic relatedness. Considering that disturbance and spatial heterogeneity in grasslands are between that in the arable fields and forests, it may be that communities in moderately disturbed habitats are driven by environmental filtering and niche partitioning with similar strengths, therefore resulting in the random patterns. Interestingly, euedaphic Collembola are also likely influenced by the stochastic processes, irrespective of the habitat types. Whether ecological drift plays a major role in the coexistence of deep soil species needs further investigations. Although some have attempted to evaluate the importance of stochasticity in community assembly of Collembola (Ims et al. 2004, Ingimarsdóttir et al. 2012, Chen et al. 2015, Sha et al. 2015), to the best of my knowledge, no study so far explicitly tests or quantifies the contribution of drift (i.e., demographic stochasticity) to community assembly in soil, presumably due to the difficulty to discriminate between dispersal and demographic stochasticity in the spatial processes.

Speciation—with emphasis on trait evolution

Phylogenetic reconstruction of Collembola

Results of phylogenetic inference of Collembola species indicate that Entomobryoidea and Isotomidae diverged earlier from the other Collembola lineages, while Poduromorpha and Symphypleona are sister groups (**Chapter 2**), contrary to the previous studies which recover Symphypleona diverging earlier than the other Collembola (D'Haese 2002, Xiong et al. 2008, Schneider et al. 2011, Yu et al. 2016). Although marker selection, unbalanced taxon sampling and choice of outgroup taxa may affect the topology of phylogenetic trees (Heath et al. 2008, Rosenfeld et al. 2012), genetic markers ranging from conserved (18S rRNA, H3) to variable (28S rRNA, COI) are used in this study to generate a phylogenetic tree for 102 locally occurring Collembola species including 51 genera from 18 families. Thus, this tree provides the most comprehensive Collembola phylogeny to date. Future studies applying phylogenomic methods may help to resolve phylogenetic relationships between basal Collembola lineages (van Straalen et al. 2008, Misof et al. 2014, Carapelli et al. 2014).

Trait evolution in light of species coexistence

According to the α and β niche traits concept, species within a community possess both similar and different traits. Similar traits related to β niches allow community members to cope with certain environmental conditions, while different α niche traits avoid competition for resources (Silvertown et al. 2006). Since traits are evolutionary signatures of species diversification, these two types of traits may evolve in different ways, with β niche traits usually phylogenetically conserved, and α niche traits evolutionarily labile (Silvertown et al. 2006; Ackerly et al. 2006; Best and Stachowicz 2013).

Results of **Chapter 2** demonstrate that body shape of Collembola evolved quickly early in diversification but followed by relative stasis, presumably under stable evolutionary or ecological constraints. Pigmentation, number of ommatidia and reproductive mode of Collembola were all associated with vertical stratification during species diversification. Although ancestral traits of Collembola were likely slender body, hemiedaphic way of life, sexual reproduction, possession of many ommatidia and bright color, these traits may have changed several times during diversification of species. Overall, the traits considered in this chapter may be categorized as β niche traits, since they are more or less related to species adaptation to the abiotic conditions. Interestingly, in each phylogenetic clade, there is one or a few Collembola species with high abundance in the Göttingen region, resulting in lower phylogenetic signal measured in species total abundance compared to that predicted by Brownian motion model. Since abundance of soil microarthropods positively correlates with the amount of food resources available in the habitats (Domes et al. 2007, Chahartaghi et al. 2009), food resource exploitation of Collembola might have evolved convergently during species diversification.

In **Chapter 3**, neutral lipid fatty acid composition is proposed as a functional trait that relates to both physiological functions and food resources of Collembola. Long-chain polyunsaturated fatty acids, related to physiological functions, demonstrate phylogenetic signal, while most food resource biomarker fatty acids and the ratios between bacterial, fungal and plant biomarker fatty acids exhibit no phylogenetic signal. Presumably, species with close phylogenetic affinity experienced similar environments during diversification, while divergence in exploitation of food resources among closely related species may favor species coexistence. Since stable isotope data indicates phylogenetic (taxonomic) conservatism in Collembola trophic niches (Potapov et al. 2016) and fatty acid composition complements stable isotopes in analyzing trophic niche of soil animals (Ferlian et al. 2015), Collembola feeding traits are, on one hand, likely to relate to the evolutionary history of species; on the other hand, they may retain variability to reduce competition. More data on trophic niches and food resources identified indirectly by stable isotope signature and neutral lipid fatty acids and directly by gut content and microbiome analyses in various species from different phylogenetic groups are needed to test this hypothesis.

Dispersal ability is a composite trait. **Chapter 5** finds close Collembola relatives colonizing defaunated soil blocks with similar speed in a meadow habitat. However, species abilities to immigrate from the surrounding habitat and survive in a new habitat are determined not only by morphological characters, such as legs, antennae, furca and visual apparatus (Ponge et al. 2006, Auclerc et al. 2009), but also by physiological attributes such as moisture tolerance (Kuznetsova 2003) and presumably by food resource quality (Bengtsson et al. 1994, Stötefeld et al. 2012). Therefore, dispersal ability ascribed by Auclerc et al. (2009) and subsequently analyzed in this chapter should be considered as a summarized response, and phylogenetic signal measured in dispersal ability likely reflects phylogenetic niche conservatism in other functional traits. Future studies may estimate dispersal ability of different species using observations (Zhang et al. 2017) and genetic tools (van der Wurff et al. 2003) and further test phylogenetic signal in dispersal ability.

Except for the fatty acid composition, this thesis mainly relies on the trait data compiled from literature at species level. The situation is probably applicable to most soil animals where information on field-measured traits at individual or population levels is still lacking (Pey et al. 2014, Moretti et al. 2017). To understand intraspecific variations, there remains a need to directly measure field-derived individuals including different cohorts and populations. Furthermore, laboratory measurements of functional traits of soil animals may help to determine the range of species fundamental niches, which may be subsequently compared with field observation data to reveal the differences in realized niches under the influences of other coexisting species. Explicitly measuring individual properties using both laboratory experiments and field observations for multiple traits, including morphological, physiological and trophic attributes, may help uncover the mechanisms driving and maintaining species coexistence in soil animals and thus explain the enigma of soil animal biodiversity (Anderson 1975).

A roadmap for implementation

In this thesis I use phylogeny as a surrogate for traits, a technique that has been much debated recently among plant ecologists (Gerhold et al. 2015, de Bello et al. 2015, Šímová 2016, Prinzing 2016, Rosado et al. 2016, McPeek 2017). One main critique is that traits are the objects that relate to or underlie different processes, not phylogeny. Phylogeny should not be used to infer assembly processes without considering the many assumptions behind it (Gerhold et al. 2015, McPeek 2017). For example, traits underlying processes must be identified and included in the null model tests; the assumption of phylogenetic conservatism of traits must be validated; phylogenetic dispersion should reflect trait dispersion; similarity in traits should translate to interspecific competition; competition should lead to species exclusion and so on. These many assumptions have received very limited empirical support (Gerhold et al. 2015, Prinzing 2016). However, phylogeny may show its advantage as a representation of multiple unmeasurable traits (Cadotte et al. 2013, de Bello et al. 2015). Phylogeny may therefore be treated as an independent variable and combined with existing trait data in the case in which the process-related traits are phylogenetically convergent or labile (Cadotte et al. 2013). Furthermore, beyond its use as a surrogate for traits, phylogenetic information can be used for studies on adaptation of species in trait-based approaches (de Bello et al. 2015) and for asking evolutionary questions of community assembly (Gerhold et al. 2015, Prinzing 2016).

For most soil animal species, however, except for morphological traits, we lack information on other aspects of traits, such as physiological attributes or those related to food resources. The situation may be worse for rare species and those difficult to culture in the laboratory. Also, species possess multiple traits that may correlate statistically or evolutionarily. Therefore, as compared to measuring all possible relevant traits, phylogenetic relationships may be a more pragmatic approach to the problem, considering the current developing stage of trait databases of soil animals, although the same concerns raised by plant ecologists may likely also apply to soil animals. Nevertheless, when trait information is lacking, we may use phylogeny to develop a first impression of the likely distribution of species traits in the communities, from which we are able to further study ecological and evolutionary hypotheses of species coexistence in soil (Cadotte et al. 2013, Gerhold et al. 2015, de Bello et al. 2015; but see Rosado et al. 2016). To that end, I propose the following roadmap for analytical strategies to study community assembly of soil animals using phylogenetic information and functional traits (**Figure 6.1**):

1. Propose traits which are likely relevant for community assembly processes. Multiple traits of soil animals from various aspects need to be considered, including morphological characters, physiological attributes and those related to food resources. Dispersal ability and ecological preference of species may also be considered as traits. Here we may ask: *Are these traits measurable for all species (Q1)?* If traits underlying community assembly processes are known and measurable for all species, we then go on with step 2.

2. Measure traits for individuals/populations in the laboratory/field. Traits should be measured at individual, population or species levels. Both laboratory experiments and field observations are needed to achieve species fundamental and realized niches. We may ask: *Do trait variations reflect processes (Q2)?* For example, whether environmental gradients correlate with community-weighted mean traits (Widenfalk et al. 2015), or whether coexisting species possess similar traits. Here, ecological processes, usually selection processes, can be investigated using trait-based approaches (the **red** route; **I**). However, in a situation where no functional traits are known, or traits of rare species are difficult to measure (answer *no* to *Q1*), we need to go on with step 2a.

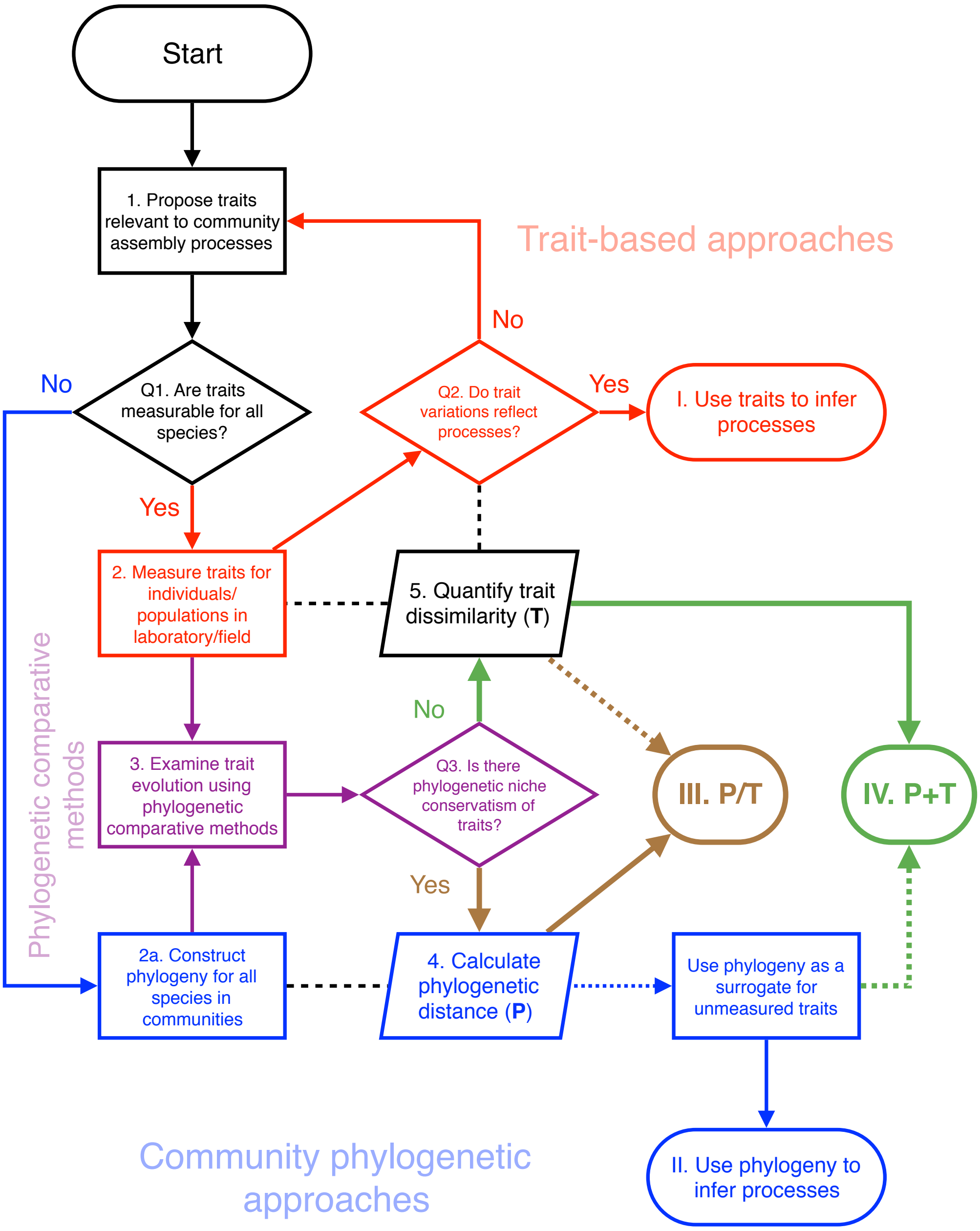

Figure 6.1
A guideline for analytical strategies to study community assembly of soil animals using phylogenetic information and functional traits. For details see the text.

2a. Construct phylogeny for all species in communities. We may take phylogeny as a surrogate for the unknown or unmeasured traits, assuming phylogenetic signal (de Bello et al. 2015), and get first insight into the phylogenetic (thus likely trait) structures in local communities (the **blue** route). In this situation, phylogeny provides an alternative tool to quantify differences between species, without *a priori* knowledge of traits (**II**). This approach may also apply to communities consisting of cryptic species that are unable to be differentiated by any morphological characters, a situation frequently occurring in soil animals (Zhang et al. *in revision*), in which only genetic distances between cryptic species are known. In such a case phylogeny shows its advantage to quantifying differences between (cryptic) species. Furthermore, from these patterns, other hypotheses may be proposed; for example, if phylogenetic clustering is found, we may hypothesize that some traits do differ between coexisting species which allow coexistence, and predict what these traits may be. This method is adopted, for example, in **Chapter 5** of this thesis.

3. Examine trait evolution. Then, having species trait data and their phylogenetic tree, we may investigate evolutionary processes that shape trait variations in different species (the **purple** route). Trait evolution can be analyzed using phylogenetic comparative methods, such as phylogenetic signal measurements, model tests and ancestral state estimation as demonstrated in **Chapters 2 and 3**. Taxonomic or phylogenetic scales should be considered in the analyses. Here, the question we may ask includes: *Is there phylogenetic niche conservatism of the traits (Q3)?* Different answers lead to different ways to the subsequent analyses: If *yes*, first go to step 4 and then consider step 5; if *no*, go to step 5 and combine this with step 4.

4. Calculate phylogenetic distances, and *5. quantify trait dissimilarity between species.* If traits are phylogenetically conserved or follow Brownian motion model of evolution (the **brown** route), phylogenetic distance matrix (**P**) is calculated and used to infer processes. At the same time, the trait dissimilarity matrix (**T**) is quantified to evaluate whether phylogenetic relatedness and trait similarity give congruent results (**III; P/T**), as demonstrated in **Chapter 4**. On the contrary, if traits evolved randomly or convergently (the **green** route), the trait distance matrix (**T**) should be treated as an independent dataset which can be combined with the phylogenetic distance matrix (**P**) to evaluate the relative contributions of each dataset (**P** and **T**) to the differences between community members (**IV**; **P+T**; Cadotte et al. 2013, Gong et al. *submitted*).

Conclusions

This thesis provides the first example of integrating phylogenetic comparative methods, community phylogenetic analyses and trait-based approaches in studies on the assembly processes of soil invertebrates using both field observations and experimental manipulations with Collembola as the model soil animal group. Phylogenetic signal in process-related functional traits is demonstrated by using comparative methods for body shape, body length, pigmentation, number of ommatidia, reproductive mode and vertical stratification that reflect adaptations to the environment. For the first time, neutral lipid fatty acid composition is established as a functional trait related to both food resources and physiological attributes of species. Phylogenetic signal in fatty acid composition suggests that species with close phylogenetic affinity experienced similar environments during divergence, while niche partitioning in food resources favored species coexistence. Furthermore, differences in phylogenetic relatedness and trait similarity of local communities provide hints on assembly processes, i.e., Collembola in arable fields are mainly structured by environmental filtering, while niche partitioning dominates in forests. In addition, the relative importance of these mechanisms varies between soil strata and between phylogenetic lineages. Furthermore, combining of field manipulative experiments with community phylogenetic approaches allows deeper understanding of assembly processes in Collembola communities along successional trajectories in different habitats. Both the conceptual model and analytical roadmap proposed in this thesis can be applied to other soil taxa. Future studies integrating different approaches may shed new light on the mechanisms driving and maintaining species coexistence and biodiversity in soil.

References

Ackerly, D.D., Schwilk, D.W. & Webb, C.O. (2006) Niche evolution and adaptive radiation: testing the order of trait divergence. *Ecology*, **87**, S50-61.

Anderson, J.M. (1975) The enigma of soil animal species diversity. Progress in Soil Zoology: Proceedings of the 5th International Colloquium on Soil Zoology (ed J. Vaněk), pp. 51–58. Springer Netherlands, Dordrecht.

Åström, J. & Bengtsson, J. (2011) Patch size matters more than dispersal distance in a mainland-island metacommunity. *Oecologia*, **167**, 747–757.

Auclerc, A., Ponge, J.-F., Barot, S. & Dubs, F. (2009) Experimental assessment of habitat preference and dispersal ability of soil springtails. *Soil Biology and Biochemistry*, **41**, 1596–1604.

de Bello, F., Berg, M.P., Dias, A.T.C., Diniz-Filho, J.A.F., Götzenberger, L., Hortal, J., Ladle, R.J. & Lepš, J. (2015) On the need for phylogenetic "corrections" in functional trait-based approaches. *Folia Geobotanica*, **50**, 349–357.

Bengtsson, G., Hedlund, K. & Rundgren, S. (1994) Food- and density-dependent dispersal: evidence from a soil collembolan. *Journal of Animal Ecology*, **63**, 513–520.

Best, R.J. & Stachowicz, J.J. (2013) Phylogeny as a proxy for ecology in seagrass amphipods: which traits are most conserved? *PLoS ONE*, **8**, e57550.

Cadotte, M., Albert, C.H. & Walker, S.C. (2013) The ecology of differences: assessing community assembly with trait and evolutionary distances. *Ecology Letters*, **16**, 1234–1244.

Carapelli, A., Convey, P., Nardi, F. & Frati, F. (2014) The mitochondrial genome of the antarctic springtail *Folsomotoma octooculata* (Hexapoda; Collembola), and an update on the phylogeny of collembolan lineages based on mitogenomic data. *Entomologia*, **2**.

Chahartaghi, M., Langel, R., Scheu, S. & Ruess, L. (2005) Feeding guilds in Collembola based on nitrogen stable isotope ratios. *Soil Biology and Biochemistry*, **37**, 1718–1725.

Chahartaghi, M., Maraun, M., Scheu, S. & Domes, K. (2009) Resource depletion and colonization: A comparison between parthenogenetic and sexual Collembola species. *Pedobiologia*, **52**, 181–189.

Chen, Y., Amundrud, S.L. & Srivastava, D.S. (2015) Spatial variance in soil microarthropod communities: Niche, neutrality, or stochasticity? *Ecoscience*, **21**, 1–14.

Costa, D., Timmermans, M.J.T.N., Sousa, J.P., Ribeiro, R., Roelofs, D. & van Straalen, N.M. (2013) Genetic structure of soil invertebrate populations: Collembolans, earthworms and isopods. *Applied Soil Ecology*, **68**, 61–66.

D'Haese, C.A. (2002) Were the first springtails semi-aquatic? A phylogenetic approach by means of 28S rDNA and optimization alignment. *Proceedings of the Royal Society of London B*, **269**, 1143–1151.

Digel, C., Curtsdotter, A., Riede, J., Klarner, B. & Brose, U. (2014) Unravelling the complex structure of forest soil food webs: higher omnivory and more trophic levels. *Oikos*, **123**, 1157–1172.

Ding, Y., Zang, R., Letcher, S.G., Liu, S. & He, F. (2012) Disturbance regime changes the trait distribution, phylogenetic structure and community assembly of tropical rain forests. *Oikos*, **121**, 1263–1270.

Domes, K., Scheu, S. & Maraun, M. (2007) Resources and sex: Soil re-colonization by sexual and parthenogenetic oribatid mites. *Pedobiologia*, **51**, 1–11.

Ferlian, O., Klarner, B., Langeneckert, A.E. & Scheu, S. (2015) Trophic niche differentiation and utilisation of food resources in collembolans based on complementary analyses of fatty acids and stable isotopes. *Soil Biology and Biochemistry*, **82**, 28–35.

Gerhold, P., Cahill, J.F., Winter, M., Bartish, I. V & Prinzing, A. (2015) Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Functional Ecology*, **29**, 600–614.

Gianuca, A.T., Dias, R.A., Debastiani, V.J. & Duarte, L.D.S. (2014) Habitat filtering influences the phylogenetic structure of avian communities across a coastal gradient in southern Brazil. *Austral Ecology*, **39**, 29–38.

Gong, X., Chen, T.-W., Zieger, S., Bluhm, C., Heidemann, K., Schaefer, I., Maraun, M., Liu, M. & Scheu, S. Phylogenetic and trophic determinants of gut microbiota in soil detritivores. *Submitted.*

Hågvar, S. (2000) Navigation and behaviour of four Collembola species migrating on the snow surface. *Pedobiologia*, **44**, 221–233.

Heath, T.A., Hedtke, S.M. & Hillis, D.M. (2008) Taxon sampling and the accuracy of phylogenetic analyses. *Journal of Systematics and Evolution*, **46**, 239–257.

Ims, R.A., Leinaas, H.P. & Coulson, S. (2004) Spatial and temporal variation in patch occupancy and population density in a model system of an arctic Collembola species assemblage. *Oikos*, **105**, 89–100.

Ingimarsdóttir, M., Caruso, T., Ripa, J., Magnúsdóttir, Ó.B., Migliorini, M. & Hedlund, K. (2012) Primary assembly of soil communities: disentangling the effect of dispersal and local environment. *Oecologia*, **170**, 745–754.

Kuznetsova, N. (2003) Humidity and distribution of springtails. *Entomological Review*, **83**, 230–238.

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

McPeek, M.A. (2017) Evolutionary Community Ecology. Princeton University Press.

Misof, B., Liu, S., Meusemann, K., Peters, R.S., Donath, A., Mayer, C., Frandsen, P.B., Ware, J., Flouri, T., Beutel, R.G., Niehuis, O., Petersen, M., Izquierdo-Carrasco, F., Wappler, T., Rust, J., Aberer, A.J., Aspöck, U., Aspöck, H., Bartel, D., Blanke, A., Berger, S., Böhm, A., Buckley, T.R., Calcott, B., Chen, J., Friedrich, F., Fukui, M., Fujita, M., Greve, C., Grobe, P., Gu, S., Huang, Y., Jermiin, L.S., Kawahara, A.Y., Krogmann, L., Kubiak, M., Lanfear, R., Letsch, H., Li, Y., Li, Z., Li, J., Lu, H., Machida, R., Mashimo, Y., Kapli, P., McKenna, D.D., Meng, G., Nakagaki, Y., Navarrete-Heredia, J.L., Ott, M., Ou, Y., Pass, G., Podsiadlowski, L., Pohl, H., von Reumont, B.M., Schütte, K., Sekiya, K., Shimizu, S., Slipinski, A., Stamatakis, A., Song, W., Su, X., Szucsich, N.U., Tan, M., Tan, X., Tang, M., Tang, J., Timelthaler, G., Tomizuka, S., Trautwein, M., Tong, X., Uchifune, T., Walzl, M.G., Wiegmann, B.M., Wilbrandt, J., Wipfler, B., Wong, T.K.F., Wu, Q., Wu, G., Xie, Y., Yang, S., Yang, Q., Yeates, D.K., Yoshizawa, K., Zhang, Q., Zhang, R., Zhang, W., Zhang, Y., Zhao, J., Zhou, C., Zhou, L., Ziesmann, T., Zou, S., Li, Y., Xu, X., Zhang, Y., Yang, H., Wang, J., Wang, J., Kjer, K.M. & Zhou, X. (2014) Phylogenomics resolves the timing and pattern of insect evolution. *Science*, **346**, 763–767.

Moretti, M., Dias, A.T.C., de Bello, F., Altermatt, F., Chown, S.L., Azcarate, F.M., Bell, J.R., Fournier, B., Hedde, M., Hortal, J., Ibanez, S., Öckinger, E., Sousa, J.P., Ellers, J. & Berg, M.P. (2017) Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. *Functional Ecology*, **31**, 558–567.

Ojala, R. & Huhta, V. (2001) Dispersal of microarthropods in forest soil. *Pedobiologia*, **45**, 443–450.

Pellissier, L., Pradervand, J.-N., Williams, P.H., Litsios, G., Cherix, D. & Guisan, A. (2013) Phylogenetic relatedness and proboscis length contribute to structuring bumblebee communities in the extremes of abiotic and biotic gradients (ed A Baselga). *Global Ecology and Biogeography*, **22**, 577–585.

Pey, B., Nahmani, J., Auclerc, A., Capowiez, Y., Cluzeau, D., Cortet, J., Decaëns, T., Deharveng, L., Dubs, F., Joimel, S., Briard, C., Grumiaux, F., Laporte, M.-A.A., Pasquet, A., Pelosi, C., Pernin, C., Ponge, J.-F., Salmon, S., Santorufo, L. & Hedde, M. (2014) Current use of and future needs for soil invertebrate functional traits in community ecology. *Basic and Applied Ecology*, **15**, 194–206.

Pollierer, M.M., Langel, R., Scheu, S. & Maraun, M. (2009) Compartmentalization of the soil animal food web as indicated by dual analysis of stable isotope ratios ($^{15}N/^{14}N$ and $^{13}C/^{12}C$). *Soil Biology and Biochemistry*, **41**, 1221–1226.

Ponge, J.-F., Dubs, F., Gillet, S., Sousa, J. & Lavelle, P. (2006) Decreased biodiversity in soil springtail communities: the importance of dispersal and landuse history in heterogeneous landscapes. *Soil Biology and Biochemistry*, **38**, 1158–1161.

Potapov, A.A., Semenina, E.E., Korotkevich, A.Y., Kuznetsova, N.A. & Tiunov, A. V. (2016) Connecting taxonomy and ecology: Trophic niches of collembolans as related to taxonomic identity and life forms. *Soil Biology and Biochemistry*, **101**, 20–31.

Prinzing, A. (2016) On the opportunity of using phylogenetic information to ask evolutionary questions in functional community ecology. *Folia Geobotanica*, **51**, 69–74.

Rosado, B.H.P., Matos, I.S. & Amorim, T. de A. (2016) A matter of scale and traits: a comment on "On the need for phylogenetic 'corrections' in functional trait-based approaches" by de Bello et al. (2015). *Folia Geobotanica*, **51**, 383–387.

Rosenfeld, J.A., Payne, A. & DeSalle, R. (2012) Random roots and lineage sorting. *Molecular Phylogenetics and Evolution*, **64**, 12–20.

Scheu, S. & Falca, M. (2000) The soil food web of two beech forests (*Fagus sylvatica*) of contrasting humus type: stable isotope analysis of a macro- and a mesofauna-dominated community. *Oecologia*, **123**, 285–296.

Schneider, C., Cruaud, C. & D'Haese, C.A. (2011) Unexpected diversity in Neelipleona revealed by molecular phylogeny approach (Hexapoda, Collembola). *Soil Organisms*, **83**, 383–398.

Sha, D., Gao, M., Sun, X., Wu, D. & Zhang, X. (2015) Relative Contributions of Spatial and Environmental Processes and Biotic Interactions in a Soil Collembolan Community. *Chinese Geographical Science*, **25**, 582–590.

Silvertown, J., McConway, K., Gowing, D., Dodd, M., Fay, M.F., Joseph, J.A. & Dolphin, K. (2006) Absence of phylogenetic signal in the niche structure of meadow plant communities. *Proceedings of the Royal Society B*, **273**, 39–44.

Šímová, I. (2016) Phylogenies are relevant when assessing environmental filtering. *Folia Geobotanica*, **51**, 65–68.

Stötefeld, L., Scheu, S. & Rohlfs, M. (2012) Fungal chemical defence alters density-dependent foraging behaviour and success in a fungivorous soil arthropod. *Ecological Entomology*, **37**, 323–329.

van Straalen, N.M., Timmermans, M.J.T.N., Roelofs, D. & Berg, M.P. (2008) Apterygota in the spotlights of ecology, evolution and genomics. *European Journal of Soil Biology*, **44**, 452–457.

Vellend, M. (2010) Conceptual synthesis in community ecology. *The Quarterly Review of Biology*, **85**, 183–206.

Vellend, M. (2016) *The Theory of Ecological Communities*. Princeton University Press, Princeton.

Widenfalk, L.A., Bengtsson, J., Berggren, Å., Zwiggelaar, K., Spijkman, E., Huyer-Brugman, F. & Berg, M.P. (2015) Spatially structured environmental filtering of collembolan traits in late successional salt marsh vegetation. *Oecologia*, **179**, 537–549.

Widenfalk, L.A., Malmström, A., Berg, M.P. & Bengtsson, J. (2016) Small-scale Collembola community composition in a pine forest soil - Overdispersion in functional traits indicates the importance of species interactions. *Soil Biology and Biochemistry*, **103**, 52–62.

van der Wurff, A.W.G., Isaaks, J.A., Ernsting, G. & van Straalen, N.M. (2003) Population substructures in the soil invertebrate *Orchesella cincta*, as revealed by microsatellite and TE-AFLP markers. *Molecular Ecology*, **12**, 1349–1359.

Xiong, Y., Gao, Y., Yin, W. & Luan, Y. (2008) Molecular phylogeny of Collembola inferred from ribosomal RNA genes. *Molecular Phylogenetics and Evolution*, **49**, 728–735.

Yu, D., Zhang, F., Stevens, M.I., Yan, Q., Liu, M. & Hu, F. (2016) New insight into the systematics of Tomoceridae (Hexapoda, Collembola) by integrating molecular and morphological evidence. *Zoologica Scripta*, **45**, 286–299.

Zhang, B., Chang, L., Ni, Z., Sun, X. & Wu, D. (2017) Directional migration of three *Desoria* species (Collembola: Isotomidae) on the snow surface in late winter. *European Journal of Soil Biology*, **81**, 64–68.

Zhang, B., Chen, T.-W., Mateos, E., Scheu, S. & Schaefer, I. Cryptic species in *Lepidocyrtus lanuginosus* (Collembola: Lepidocyrtidae) are sorted by habitat types. *Pedobiologia*. *In revision.*

Acknowledgements

I owe many thanks to **Stefan Scheu** for the opportunity to be part of his working group. He was and is the best supervisor with great support in many aspects during my whole PhD period. His knowledge and enthusiasm inspire me with new ideas and the motivation for going on. I still remember the warmest smile he gave to me at the entrance of the Zoological Institute when the first time I came to Göttingen for the interview.

I thank **Holger Kreft** for his efforts as referee for this dissertation. I appreciate **Mark Maraun**, **Marko Rohlfs**, **Elvira Hörandl** and **Kerstin Wiegand** for their willingness and the time they have invested in being part of the committee.

Special thanks to **Andrea Lambertz** and the technicians in the working group, **Christel Fischer**, **Susanne Böning-Klein**, **Guido Humpert**, **Theodora Volovei**, **Ingrid Kleinhans, Dieter Nünchert, Andrea Graebe and Elisabeth Opielka**. Without their efforts and organization, most things could not run smoothly in the labs or the working group.

I would like to thank the colleagues who helped me with field work, **Bernhard Eitzinger**, **Jens Bast**, **Patrick Pachl**, **Kathleen Lemanski**, **Odette Gonzalez**, **Paul Götsch**, **Andreas Klein** and **Charlotte Seifert**, and with Collembola identification, **Nicole Scheuermann** and **Xin Sun**. I also thank **Ina Schaefer** for her introductions to molecular phylogenetic analyses, as well as **Tamara Hartke** and **Andrew Davis** for their corrections to my writing in English. Special thanks to the people reviewing this dissertation; in addition to Stefan, Tamara and Andrew, this should also include **Mark**, **Ina** and **Gerrit Moser**; their comments on the manuscripts improved the quality of the thesis very much.

Thanks to the core members in the **paper writing group**, **Tamara**, **Odette** and **Nico Radermacher**; gathering together really speeded up the writing. Also thanks to the participants in the **paper/book reading group**, **Tamara**, **Odette**, **Ina**, **Bing Zhang**, **Zhipeng Li**, **Xin Gong**, **Bastian Heinburger** and those I forgot to mention; they have inspired me with lots of ideas and helped me to reformulate my thoughts. I was also glad to talk with **Huining Zhang**; the casual talk surprisingly led to a project within this thesis (Chapter 5), even though she is an expert in cicada taxonomy.

Thanks to all the **Bachelor and Master students** and **HiWis**, **Mona Laura Kahlfeld**, **Victoria Kreipe**, **Lars Blumenberg**, **Yu-De Lin**, **Einar Michael Unger**, **Denise Castle**, **Jo-Fan Chao**, **Philipp Sandmann**, **Sarah Wehrmaker**, **Birk Urmersbach** and **Gerrit M.**, for their time working with me and all they have accomplished.

I also appreciate the inspiring discussions with **Mark**, **Marko**, **Andrew**, **Tamara**, **Ina**, **Bing**, **Xin G.**, **Bernhard E.**, **Anton Potapov**, **Matty P. Berg**, **Tancredi Caruso**, **Stefanie Maaß, Meixiang Gao** and **Donghui Wu**, and still remember what we may go on with in the future.

I would like to thank the whole **AG Scheu** for such a nice working atmosphere. I have really enjoyed the time here.

What also impresses me a lot is the “German learning lunch” with **Jens**, **Bernhard E.** and **Alexandre Jousset** along the walk to the Mensa as well as our friendships. I received great support from my friends in Göttingen and would like to thank **Elissa H. Lee-Drude** and her family, **Che-Wei Chang**, **Chao-Chin Chan**, **Po-Yin Chiang, “Right Fist” Chang, Sheng-An Yang, Sheng-Yen Feng** and **Wei-An Sheng**; they have made life in Germany not lonely and boring but colorful and wonderful. I also thank **Liang Chang**, **Bing**, **Xin S.** and **Xin G.** for the relaxing talks in our fluent mother tongue. In addition, thanks to the Lord Jesus Christ for the support from the Brothers and Sisters in the Church life in Göttingen.

I am grateful for the scholarship given by the Ministry of Education, Taiwan and the BEST-project funded by the German Ministry of Education, Science and Technology (BMBF).

I would like to thank **Ping-Chun Lucy Hou**, who was a pioneer in the field, for bringing me to the world of soil invertebrates.

Last, I give my special thanks to my beloved **Jo-Fan Chao** for her encouragement, patience, company and love in all situations, and to my family in Taiwan for all kinds of their support.

謝謝大家！

List of publications

Published in peer-reviewed journals

Chen, T.-W., P. Sandmann, I. Schaefer and S. Scheu (2017) Neutral lipid fatty acid composition as trait and constraint in Collembola evolution. *Ecology and Evolution*, **7**, 9624–9638.

Talbot, B, **T.-W. Chen**, S. Zimmerman, S. Joost, A. J. Eckert, T. M. Crow, D. Semizer-Cuming, C. Seshadri, and S. Manel (2017) Combining genotype, phenotype, and environment to infer potential candidate genes. *Journal of Heredity*, **108**, 207–216.

Chen, T.-W., C.-P. Wang, and P.-C. L. Hou (2008) A review on the impacts of the invasive earthworms, *Pontoscolex corethrurus*, on forest ecosystems. *Quarterly Journal of Chinese Forestry*, **41**, 135–148 (in Chinese with English abstract).

Chang, W.-L., **T.-W. Chen,** C.-Y Yang, Y.-H Huang, and D. Chao (2004) Prevalence and observation of intestine-dwelling gregarines in the millipede *Trigoniulus corallinus* (Spirobolida: Pachybolidae) collected from Shoushan, Kaohsiung, Taiwan. *Formosan Entomologist*, **24**, 137–145.

Submitted

Zhang, B.†, **T.-W. Chen**†, E. Mateos, S. Scheu and I. Schaefer. Cryptic species in *Lepidocyrtus lanuginosus* (Collembola: Lepidocyrtidae) are sorted by habitat types. *Pedobiologia. In revision.*

(†These authors contributed equally to this paper.)

Gong, X.†, **T.-W. Chen**†, S. Zieger, C. Bluhm, K. Heidemann, I. Schaefer, M. Maraun, M. Liu and S. Scheu. Phylogenetic and trophic determinants of gut microbiota in soil detritivores. *Submitted.*

(†These authors contributed equally to this paper.)

Thesis declarations

Declaration of the author's own contribution to manuscripts with multiple authors

Chapter 3 has been published in a peer-reviewed journal; Philipp Sandmann collected the fatty acid data and performed his Bachelor thesis under my supervision. Trait data used in **Chapters 2 and 4** was provided by Matty P. Berg. Data on molecular sequences in **Chapter 2** was obtained with the help of Jo-Fan Chao. Abundance data on Collembola in **Chapter 4** was collected with the help of Jo-Fan Chao; Victoria Kreipe performed her Master thesis under my supervision which was related to **Chapter 4**. Community data used in **Chapter 5** was kindly provided by Apolline Auclerc, Jean-François Ponge, Sébastien Barot and Florence Dubs; Gerrit Moser conducted the preliminary analyses for the data and performed his Master thesis under my supervision.

I am the first author of all chapters; I have developed the main ideas, designed the studies, analyzed the data, written the manuscripts and created tables, figures and supplementary materials.

Plagiarism declaration

I, Ting-Wen Chen, declare that I have written this doctoral thesis independently. All persons contributing to the manuscripts have been named so. All sentences or passages quoted from other people's work have been specifically acknowledged by clear cross-referencing.

I have not submitted this thesis in any form for another degree at any university or institution.

Ting-Wen Chen
Göttingen, December 2017