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

– CENTRE OF BIODIVERSITY AND SUSTAINABLE LAND USE –
SECTION: BIODIVERSITY, ECOLOGY AND NATURE CONSERVATION

Genetic diversity of sexual and parthenogenetic soil living arthropods (Collembola) in Europe: colonization patterns, pre-glacial diversifications and founder effects

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

vorgelegt von

Diplom-Biologe

Helge von Saltzwedel

aus
Uelzen

Göttingen, Dezember 2015

Referent: Prof. Dr. Stefan Scheu

Korreferent: Prof. Dr. Mark Maraun

Tag der mündlichen Prüfung: 18.03.2016

No serious biologist doubts the fact that evolution has happened, nor that all living creatures are cousins of one another.
Richard Dawkins

The answer to life, universe and everything: 42.
Douglas Adam

Table of contents

Summary 7

CHAPTER 1

GENERAL INTRODUCTION **11**

Phylogeograhy and soil 12
Sex and parthenogenesis 15
Molecular markers 17
Collembola 18
Study objectives and hypotheses 20
References 22

CHAPTER 2

FOUNDER EVENTS AND PRE-GLACIAL DIVERGENCES SHAPE THE GENETIC STRUCTURE OF EUROPEAN COLLEMBOLA SPECIES **35**

Abstract 36
Introduction 37
Materials and Methods 39
Sampling of animals and DNA extraction *39*
Phylogenetic analyses, divergence time estimation and population structure *42*
Results 43
Population structure *43*
Haplotype networks and phylogenetic analyses *48*
Estimation of divergence times *53*
Discussion 56
Acknowledgements 59
References 60
Data Accessibility 66
Supporting information 66

CHAPTER 3

GENETIC STRUCTURE AND COLONIZATION PATTERNS OF SEXUAL AND PARTHENOGENETIC SPRINGTAILS (HEXAPODA, COLLEMBOLA) IN EUROPE **81**

Abstract 82
Introduction 83

Materials and Methods 85
Data collection *85*
Phylogenetic analyses, divergence time estimation and population structure *88*
Clusters of low genetic distances *89*
Results 89
Population structure *89*
Phylogenetic lineages, molecular age estimates and phylogeographic patterns *91*
Demographic changes and genetic variation *97*
Discussion 98
Refugia and dispersal *98*
Dispersal and vicariance *99*
Mitochondrial and nuclear markers *100*
Conclusion 102
Acknowledgements 102
References 103
Data Accessibility 107
Supporting information 108

CHAPTER 4

COLONIZATION OF EUROPE BY *PARISOTOMA NOTABILIS* (COLLEMBOLA): CRYPTIC DIVERSITY AND ANTHROPOGENIC IMPACT 117

Abstract 118
Introduction 119
Materials and Methods 120
Ethics statement *120*
Sampling of animals and DNA extraction *121*
Data analysis *123*
Results 125
Phylogeny and independent evolutionary units *125*
Genetic diversity *128*
Molecular divergence times *131*
Discussion 131
Acknowledgements 135
References 135
Data Accessibility 142
Supplementary material 142

GENERAL DISCUSSION 151

References 160

ACKNOWLEDGEMENTS 165

LIST OF PUBLICATIONS 167

Published in peer-reviewed journals 167

Submitted 167

THESIS DECLARATIONS 169

Declaration of the author's own contribution to manuscripts with multiple authors 169

Plagiarism declaration 170

Summary

Large parts of northern and central Europe were covered by ice sheets and permafrost due to climate changes in Europe during the last ice age (2.7 million to 11.7 kya). Plant and animal species had to adapt to lower temperatures, retreated to warmer areas in the south or went extinct. Once, after the Last Glacial Maximum (LGM, 26.5 kya to 19 kya) higher temperatures induced ice free habitats and these new habitats could be recolonized from different refugia. Collembola are one of the most abundant soil living decomposer animals and play a major role in aboveground - belowground interactions. Surprisingly, little is known about genetic and phylogeographic patterns, dispersal routes and anthropogenic influences of sexual and parthenogenetic reproducing European Collembola species, neglecting one important part of the global biodiversity, the belowground system. This thesis focuses on genetic patterns of four Collembola species with different reproductive modes and overlapping ecology across Europe. Collembola existed millions of years in stable habitats, as Eocene fossils show only little variation to extant taxa. In contrast, Cenozoic and Quaternary climatic changes reduced diversity and changed genetic structure of above living animals and plants.

In Chapter 2 I investigated the phylogeographic patterns of three common species of Collembola (*Ceratophysella denticulata*, *Folsomia quadrioculata* and *Isotomiella minor*) at a pan-European scale to identify glacial refuges and post-glacial colonization patterns with three genetic markers to cover different time scales. Results suggested density dependent processes for the establishment of new populations, as genetic diversity was high between but low within populations. This founder-takes-it-all principle is common in animal and plant species and suggests that only few early colonizing individuals founded the populations which grew and expanded rapidly. Arrival and invasion of other alleles into these populations was prevented by competition. Surprisingly and in contrast to the post-glacial recolonization patterns of aboveground organisms, the last ice age little affected the genetic composition of the studied Collembola species, indicating that soil provided habitat and resources for survival. The results show that divergence of populations took place during the Miocene (20-5 mya), when climatic conditions were favorable (warm and humid) for little sclerotized arthropods, susceptible to desiccation. Thus, the Miocene facilitated large scale expansion of European Collembola species. The

results suggest that evolutionary processes of soil-living species are slowed-down, compared to above the ground living species, resulting in stable populations for millions of years.

In chapter 3 I investigated differences in phylogeographic patterns due to different reproductive modes. As no partner is needed for reproduction, parthenogenesis provides a colonization advantage. To investigate the significance of reproductive modes for colonization, I compared the genetic structure of one sexual (*Folsomia quadrioculata*) and one parthenogenetic (*Isotomiella minor*) Collembola species with similar ecology across Europe, using one mitochondrial and two nuclear markers. Molecular variance was similar in both species and genetic differences were high between populations, indicating old diversifications. Northern and central Europe populations of *I. minor* were genetically homogenous suggesting that few lineages of this parthenogenetic species colonized these regions after LGM. Compared to *I. minor* the genetic structure of *F. quadrioculata* was more complex with more synonymous substitutions in protein coding genes. The results suggest that in addition to founder-effects and old diversifications, different forces affected sexual and parthenogenetic species, resulting in different phylogeographic patterns. In addition, mitonuclear compatibility among mating partners likely contributed to the more complex genetic structure in *F. quadrioculata*, whereas gene-environment interactions were of greater importance in *I. minor*. Overall, results indicate that the widespread view of central and northern European species being shaped by postglacial colonization patterns does not hold for both parthenogenetic and sexual soil-living species.

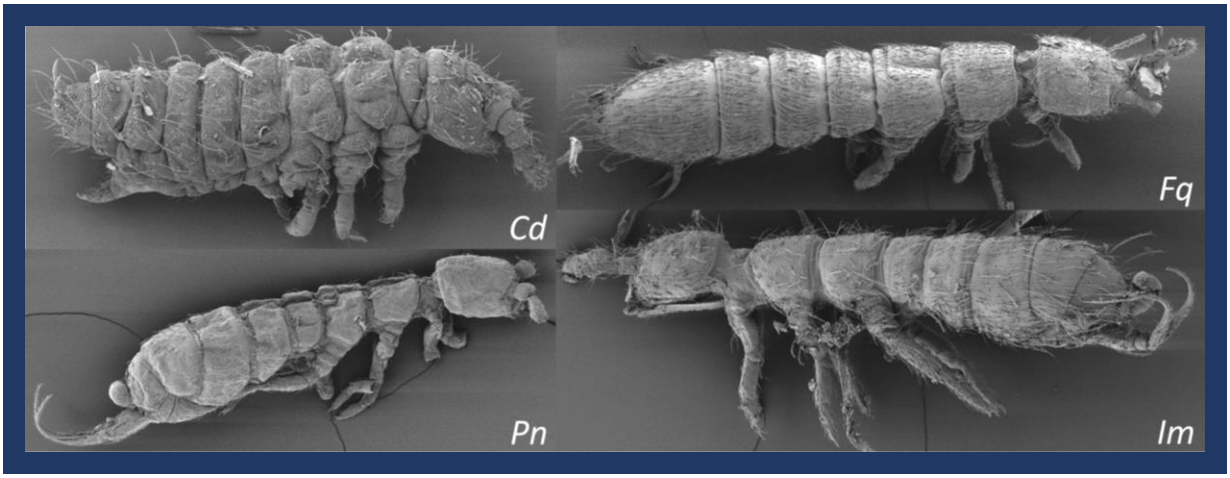
In chapter 4 I investigated colonization patterns including cryptic diversity and the anthropogenic influence of the ubiquitous Collembola species *Parisotoma notabilis* in Europe. *P. notabilis* is the most widespread and abundant Collembola species in Europe colonizing many anthropogenic and disturbed habitats. Three molecular markers were used to investigate how anthropogenic factors and climate affected the present-day genetic structure of *P. notabilis* in Europe. The results showed that *P. notabilis* forms one morphologically coherent species comprising of several discrete genetic lineages. Molecular divergence estimates suggest that these lineages diverged in the Miocene during wet and warm climate, and a biome change in central and Eastern Europe from

forest to grassland. The results further suggest that human activities favored the dispersal of *P. notabilis* as genetic lineages reflect human trading and migration routes such as the Channel and Mediterranean areas. These lineages are rather young and genetically uniform as compared to other soil-living arthropods.

Overall, the results suggest that phylogeographic patterns of soil-living European Collembola follow the 'southern-richness and northern-purity' scenario, but in contrast to this scenario, the pattern is not due to climate driven extinction of populations in central Europe during the Quaternary and recolonization thereafter. Rather, the pattern originated during the Miocene (20-5 mya) as indicated by divergence times of most clades. This indicates that the soil forms a stable habitat buffering climatic variability. Moreover, the results suggest that the dispersal of Collembola in Europe was affected by human activities. Overall, colonization in the Miocene and human activities in the Holocene resulted in a more complex genetic structure in soil-living species as compared to those living above the ground.

Chapter 1

General Introduction



Phylogeograhy and soil

Phylogeography investigates phylogenetic relatedness of populations in a geographic context (Hickerson *et al.* 2010) and has been an active research field for more than 20 years. Phylogeography solved a range of scientific questions, such as issues about speciation (Avise 2000; Moritz & Patton 2000; Hewitt 2001; Kohn 2005), human evolution (Beaumont 2004; Templeton 2005; Torroni *et al.* 2006) and taxonomy and biodiversity (Avise & Ball 1990; Taberlet 1998b; Beheregaray & Caccone 2007). The combination of molecular markers with intraspecific variance, such as the mitochondrial *COI* (cytochrome c oxidase subunit I) gene, and a geographic sampling of species revealed colonization routes, diversification events of populations and changes in distribution ranges that were triggered by climate changes or by geographical barriers for numerous species (Avise 1998; Beheregaray 2008).

Due to changes in European climate during the last ice age (2.7 mya to 11.7 kya) and the Last Glacial Maximum (LGM, 26.5 kya to 19 kya), ice sheets and permafrost covered large parts of northern and central Europe. Accordingly, populations of animals and plants adapted to cold temperatures or retreated to warmer areas, survived in southern refugia like the Balkan Peninsula, Italy, Iberian Peninsula or Greece, or went extinct (Hewitt & Ibrahim 2001; Hewitt 2004). Influences on the European fauna and flora were investigated by phylogeographic methods. Based on phylogeographic studies the following patterns were proposed:

1. The genetic pattern for aboveground species of ‘southern richness and northern purity’ (Hewitt 2000). This theory predicts a decrease of genetic diversity among populations from southern refugia to northern regions and is a consequence of postglacial recolonization of northern European countries (Hewitt 2000). With the beginning of the warmer Holocene (117 kya) period surviving species recolonized the formerly frozen and glaciated areas in central and northern Europe from refugia in the south of Europe. These dispersal routes can be inferred with molecular markers, and genetic differences of populations, due to changes in climatic conditions, can be evaluated. This pattern was found in many

aboveground species, e.g. *Chorthippus paralleus* (grasshopper), *Lissotriton vulgaris meridionalis* (newt; Maura *et al.* 2014), the brown bear *Ursus arctos* and *Alnus glutinosa* (alder) that show less haplotype diversity in northern populations than in southern ones (Hewitt 1999; Sommer & Benecke 2005).

2. The existence of southern refugia for most of the above mentioned species during the last ice age like Iberian Peninsula, Italy and the Balkans including Romania (Taberlet 1998a; Sommer & Benecke 2005; Magri 2008; Stewart & Cooper 2008; Fløjgaard *et al.* 2009; Homburg *et al.* 2013) and additionally for the insectivore mammals *Erinacues europeaus* (hedgehog; Seddon *et al.* 2001) and *Sorex araneus* (shrew; Hewitt 2001), the red deer *Cervus elatus* (Sommer & Zachos 2009), the tree species *Fagus sylvatica* (beech; Magri 2008) and *Quercus* spp. (oak; Hewitt 2001) and the insects *Arcynopteryx dichroa* (stonefly; Theissinger *et al.* 2013) and *Carabus irregularis* (beetle; Homburg *et al.* 2013).
3. Documentation of colonization routes and immigrations from Eurasia after the LGM with the help of genetic marker for most of the above mentioned species (Hewitt 1999, 2001).
4. The existence of hybrid zones, as suggested for species like grasshoppers, newts, frogs and hedgehogs due to geographical barriers of ice sheets in Europe (Hewitt 2000, 2001; Seddon *et al.* 2001; Babik *et al.* 2005).

In addition to large southern refugia, data of molecular and radiocarbon approaches indicated that species also survived in 'cryptic refugia' such as ice free nunataks (Stewart & Lister 2001; Provan & Bennett 2008). These potential refugia were small, existed in central and northern Europe and are supported by few fossilized tracks (Schmitt 2007, 2009; Provan & Bennett 2008) and molecular data (Sutkowska *et al.* 2014; McInerney *et al.* 2014; Maura *et al.* 2014).

Until now, the above mentioned phylogeographic patterns were only investigated for aboveground and freshwater species. This is surprising as it neglects one important part of the global biodiversity, the belowground system. Soil invertebrates are very diverse with up to 1,000 species per square meter in woodland soils (Giller 1996). Next to oribatid mites, Collembola are one of the most abundant groups in soils (Behan-Pelletier

2003), with more than 100,000 individuals per square meter in forest soils (Petersen & Luxton 1982). Below- and aboveground systems are linked via nutrient cycling (Hooper *et al.* 2000; Wardle *et al.* 2004; Kardol & Wardle 2010) and up to 90% of terrestrial primary production is decomposed in the soil (Giller 1996). Soil differs in many respects from freshwater and aboveground ecosystems. It buffers temperature changes and organisms that live in soil can avoid environmental stress by vertical movement into deeper soil layers (Healey 1967; Gass *et al.* 2006). Thus, the soil offers rather stable conditions for many organisms and soil-living animals are presumably less affected by climatic fluctuations compared to above-living animals. Therefore, it remains to be examined if the above phylogeographic patterns apply to soil-living species.

Collembola are suitable as model organisms for phylogeographic studies. They are living in different soil layers and are among the most abundant microarthropods (Petersen & Luxton 1982; Hopkin 1997; Rusek 1998; Paul 2011). As fungal feeders, bacterivores and detritivores, they do not depend on dead organic matter and may evade climatic changes. In addition Collembola are easy to sample and molecular markers are well established.

Although Collembola are suitable for phylogeographic studies, only few studies investigated the genetic structure of Collembola in Europe. The existing studies focus on large, epedaphic species in a geographically restricted area (Frati *et al.* 2000; van der Wurff *et al.* 2003, 2005; Cicconardi *et al.* 2010). Until now, there are only two studies showing genetic patterns of soil-living species at a large European scale, one investigating the parthenogenetic and ubiquitous Collembola species *Parisotoma notabilis* (Porco *et al.* 2012b), the other the soil-living oribatid mite *Steganacarus magnus* (Rosenberger *et al.* 2013). Results of the two studies agree in showing that intraspecific genetic diversity of these soil-living species is remarkably high in *COI* with 21% in *P. notabilis* and 32% in *S. magnus*. Porco *et al.* (2012b) found four genetically different, but morphologically similar lineages with the genetic markers *COI* and 28S rDNA. They concluded that these lineages built a ‘cryptic species complex’ because the genetic differences were as high within *P. notabilis* as in other species of the genus *Parisotoma*. While the genetic difference of *COI* also is very high in *S. magnus,* Rosenberger *et al.* (2013) concluded that individuals with high genetic differences in *COI* belong to the same species as ribosomal 18S rDNA differs

little. In addition to Porco *et al.* (2012b), Rosenberger *et al.* (2013) analyzed deep population splits using a molecular clock approach. They found that pre-glacial diversification events during the Miocene or earlier shaped the genetic patterns of *S. magnus*. However, post-glacial recolonization of central Europe also occurred. Due to high genetic variance across Europe and the coexistence of different lineages in the same population, this species contradicts the 'southern richness and northern purity' hypothesis.

The lack of phylogeographic studies on soil-living organisms is in part likely due to difficulties in handling of small organisms, often < 1 mm, determination of species requiring expert knowledge and challenging genetic analyzes due to low amounts of DNA after extraction. To investigate phylogeographic patterns in soil-living animals it is important to analyze genetic variance of the highly abundant soil-living groups such as Collembola and, in particular, more than one species has to be analyzed to find general patterns.

Sex and parthenogenesis

Many theories have been developed for explaining evolutionary advantages of sexual over parthenogenetic reproduction (Bell 1982; Kondrashov 1993; Barton & Charlesworth 1998). Parthenogenetic, or asexual, reproduction enables organisms to quickly colonize new or disturbed habitats due to faster population growth in the short-term (Williams 1975; Bell 1982; Scheu & Schulz 1996; Lindberg & Bengtsson 2005; Ingimarsdóttir *et al.* 2012). However, one advantage of sexual species is that they can eliminate harmful mutations by genetic recombination. Further, favorable mutations can be combined in following generations leading to a faster response to environmental changes in sexual populations (Williams 1975; Hamilton 1980). In contrast, parthenogenetic lineages do not have the possibility to purge accumulating deleterious mutations, therefore Muller's ratchet predicts that these lineages are doomed to extinction in the long-term (Muller 1964).

Nuclear- and mitochondrial genes are linked and mitonuclear processes are influenced by these linked genes, evolving differently in closely related hybrids (Ellison *et*

al. 2008; Montooth *et al.* 2010; Gagnaire *et al.* 2012). Mitochondria in sexually reproducing species experience different nuclear backgrounds every generation and maintaining synonymous mutations within populations may be an effective way for keeping multiple non-interfering allelic combinations. In contrast, mitonuclear complexes in parthenogenetic species are intimately linked for generations, reducing the strength of selection on neutral substitutions. Indeed, a number of studies demonstrated that mitonuclear interactions affect the fitness of species with some allelic combinations being more efficient in certain environments (Dowling *et al.* 2007; Arnqvist *et al.* 2010; Hoekstra *et al.* 2013; Wolff *et al.* 2014).

Studies showed that despite theoretical considerations and the over-whelming evidence for the prevalence of sexual reproduction in the field, some taxa appear to have survived evolutionary periods of time without sexual reproduction. These obligate long-term parthenogenetic species include bdelloid rotifers (Mark Welch & Meselson 2000; Fontaneto *et al.* 2008), Darwinulid ostracods (van Doninck *et al.* 2002; Martens *et al.* 2003), some groups of oribatid mites (Norton & Palmer 1991; Heethoff *et al.* 2007) and walking sticks of the genus *Timema* (Schwander *et al.* 2011). Additional candidate taxa for successful long-term parthenogenesis may include obligate parthenogenetic species of Collembola. Although most Collembola species are sexual, parthenogenesis is widespread (Goto 1960; Petersen 1978; Chahartaghi *et al.* 2006). Only few studies investigating parthenogenesis of Collembola are available, but it is known that parthenogenesis is most common in euedaphic taxa, like Onychiuridae (10%) and Isotomidae (7%) that live deeper in soil, where environmental conditions are stable but resources are harder to reach (Petersen 2002; Chernova *et al.* 2010). However, species living deeper in soil not only have to resolve resource problems. One of the main problems is the limited space of soil pores, impacting on morphology of Collembola species resulting in reduced number of ocelli and size of body appendages. Further, these challenges also affect the reproductive mode, favoring parthenogenetic reproduction in deeper soil layers due to difficulties in finding spermatophores or sexual partners (Chernova *et al.* 2010).

Oribatid mites resemble Collembola in ecological perspectives, but more information about parthenogenesis is available in the former group. Approximately 10% of all oribatid mites (~10,000 species) reproduce via automictic thelytoky (Norton &

Palmer 1991; Subías 2009). For Collembola more research is needed to investigate which taxa are parthenogenetic, which mode of parthenogenesis is common and why. Fossils from the Rhynie Chert suggest, that Collembola have lived approximately for 400 my in Europe (Hirst & Maulik 1926), which is similar to oribatid mites, and they may carry information on the history of the earth's climate in their genetic material. In addition, members of recent Collembola families were found as fossils in Baltic amber (Rapoport 1971; Zawischa 1993; Hädicke *et al.* 2013), indicating that Collembola have been living in Europe without large morphological modifications for more than 50 million years. Thus, they are interesting model organisms for genetic studies to investigate potential deep splits in species and information about changes in environmental conditions.

Molecular markers

With the help of molecular markers it is possible to trace diversification of populations and migration routes of organisms over time. The conservative 28S rDNA is a suitable marker to discriminate species of Collembola (D'Haese 2002; Greenslade *et al.* 2011b; Porco *et al.* 2012a) and oribatid mites (Cruickshank 2002; Maraun *et al.* 2004). The gene is evolving slowly and allows tracing of old diversifications (Avise 1994; Giribet *et al.* 1996).

Mitochondrial DNA of animals is passed on maternally (Avise *et al.* 1987) and generally has a faster mutation rate than most nuclear genes used for phylogenetic and phylogeographic studies. Due to its faster mutation rate, it allows to detect splits within species with high resolution for the past few million years. The mitochondrial protein coding gene *COI* allows inferring genetic changes between populations and has been widely used to investigate genetic changes between populations during the last ice age (Weisrock & Janzen 1999; Stevens & Hogg 2003, 2006; Hewitt 2004). Most phylogeographic studies on animals used mitochondrial DNA due to its lack of recombination, strict maternal inheritance and known substitution rate for arthropods of 1.5 to 2.3% per million years (Avise 1994; Brower 1994). While *COI* is used for barcoding of species (Hebert *et al.* 2003), genetic distances of *COI* are often high for non-insect arthropods, ranging from 11% to 23% in soil-living taxa (Edmands 2001; Heethoff *et al.*

2007; Boyer *et al.* 2007; Torricelli *et al.* 2010; Schäffer *et al.* 2010; Porco *et al.* 2012b; Rosenberger *et al.* 2013).

The nuclear protein coding gene *H3* is more variable than 28S rDNA but more conserved than *COI* (Avise 1994) and therefore suitable to resolve divergences at intermediate time-scales between those of 28S and *COI*. The combination of mitochondrial markers with the nuclear marker *Histone H3* was successfully used in different phylogeographic studies of invertebrates, e.g. for Crustacea (Villacorta *et al.* 2008; Bauzà-Ribot *et al.* 2011), Ephemeroptera (Sekiné *et al.* 2013) or Acari (Mortimer *et al.* 2011). Protein coding genes further have the advantage that non-synonymous mutations with possible direct effects on protein functions can be observed via translation of the nucleotide sequence into amino acids to find differences, e.g. between reproductive modes.

Collembola

Approximately 8,000 species of Collembola are described but possibly more than 50,000 species exist (Cicconardi *et al.* 2013). Collembola are small (0.1 mm to 17 mm), wingless and have been on earth for more than 400 million years as the Devonian springtail *Rhyniella praecursor* suggests (Hirst & Maulik 1926; Whalley & Jarzembowski 1981). Their term springtails is based on an abdominal, tail-like appendage, the furca, used for jumping. They occur in most soil habitats on earth, including extreme ones, like the Arctic and Antarctic. Springtails live in pore spaces of soil particles and in leaf litter, about 90% of all species inhabit the upper 10 cm of the soil (Hopkin 1997; Fountain & Hopkin 2004). They can reach very high densities in temperate forest soils with more than 100,000 ind./m^2 (Petersen & Luxton 1982). Using stable isotope ratios of $^{15}N/^{14}N$, different feeding guilds of Collembola were determined, including herbivores and primary and secondary decomposers that feed on dead and living plant material, or fungi and associated bacteria, respectively (Chahartaghi *et al.* 2005). Thus, most Collembola are generalist feeders that hold different trophic niches. Thereby, they are an important functional group of the soil animal community by affecting litter decomposition and nutrient cycling (Filser 2002; Kaneda & Kaneko 2006).

One characteristic of Collembola is that many species are quick colonizers of disturbed habitats (Lindberg & Bengtsson 2005; Ingimarsdóttir *et al.* 2012). Further, freezing avoidance has been reported for several species and probably is a common trait in many Collembola (Ohlsson & Verhoef 1988; Worland 2005; Bahrndorff *et al.* 2006; Konestabo *et al.* 2007). Many species are also cold-adapted and active in winter, like the Antarctic springtails *Cryptopygus antarcticus* (Pilipovic *et al.* 2008) and *Desoria saltans* (Sattler *et al.* 2012), living on glacial surfaces, or the winter-active European Collembola *Ceratophysella sigillata* (Block & Zettel 2003). Thus, Collembola potentially survived in central or northern European regions during the last ice age.

To acquire a general picture of large-scale genetic diversity within and among soil-living organisms and the effect of severe climatic changes like the LGM on soil organisms in Europe (Clark *et al.* 2009), I investigated the genetic structure of four species of Collembola that are widely distributed and common in the northern Hemisphere: (1) The hypogasturid *Ceratophysella denticulata* (Bagnall, 1941) is distributed in the Palearctic (Nitzu *et al.* 2010), feeds as secondary decomposer on bacteria, fungi and smaller animals like protozoa, nematodes and rotifers (Chahartaghi *et al.* 2005; Heidemann *et al.* 2014) and has limited dispersal potential, due to a weakly developed furca. (2) The isotomid *Folsomia quadrioculata* (Tullberg, 1871) has a Holarctic distribution (Potapov & Babenko 2000), feeds as primary decomposer especially on litter material and linked microorganisms (Chahartaghi *et al.* 2005) and has a well-developed furca for jumping several centimeters. Therefore it can escape predators or reach other habitats more easily as compared to *C. denticulata*. Both species are of intermediate body size (up to 2.5 mm) and live in the uppermost soil layer (hemiedaphic) and reproduce sexually. (3) *Isotomiella minor* (Schäffer, 1896) is a small euedaphic species (~1.3 mm) that is distributed worldwide, feeds probably on bacteria (Ponge 1991; Langeneckert 2013) and has a well-developed furca. (4) The small (~1.1 mm) hemiedaphic species *Parisotoma notabilis* (Schäffer, 1896) has a cosmopolitan distribution (Potapov 2001; Greenslade *et al.* 2011a; Potapov *et al.* 2011; Wang *et al.* 2014), feeds as secondary decomposer on bacteria, fungi and small soil animals and has a well-developed furca for dispersal and escaping predators. It is the most widespread Collembola species in Europe (Fiera & Ulrich 2012), occupying various habitats like grasslands, agricultural fields, forests and

urban soils. *I. minor* and *P. notabilis* belong to the family Isotomidae and are obligate parthenogens contrasting *C. denticulata* and *F. quadrioculata* which reproduce sexually (Chahartaghi *et al.* 2006). In general, hemiedaphic species likely are more mobile than euedaphic species and colonize new habitats faster.

Species of both Hypogasturidae and Isotomidae have been found in Baltic amber (Rapoport 1971; Lawrence 1985; Christiansen & Pike 2002) indicating their occurrence in Europe for more than 35 million years.

Study objectives and hypotheses

This thesis investigated the genetic structure of four species of Collembola and focused on ancient and recent changes in genetic variance and diversity among populations, shaped by changes of the European climate during the past three million years. In addition, this thesis aimed at investigating migration patterns related to the reproductive mode with the help of genetic markers. Two sexual and two parthenogenetic Collembola with overlapping ecological and environmental preferences were investigated. To cover the temporal resolution of genetic divergences among lineages from relatively recent to more distant time scales I analyzed one mitochondrial (*COI*) and two nuclear markers (*H3* and 28S rDNA D3-D5 region) with different substitution rates. The conserved nuclear markers allow inspecting old diversifications of lineages and are useful for resolving time scales for which the mitochondrial marker cannot be used due to its higher intraspecific variance. Further, molecular clock analyses were performed for all species, to ascribe major radiation events to geological time periods and events.

In Chapter 2 I analyzed three species in a comparative way to identify general patterns in population structure and genetic diversity that relate to major climate changes on continental scale. If Collembola were equally affected by changing environmental conditions, all three species will have similar genetic patterns of high endemism of distinct haplotypes in southern countries and closely related haplotypes will be rather homogeneously distributed in central and northern countries. In contrast, if Collembola survived during Quaternary glaciations in southern refugia, local molecular

variance and genetic distances within populations from southern Europe presumably will be higher than in northern and central Europe. If the soil buffered abiotic fluctuations in the past and Collembola survived in local patches north of the Alps, genetic diversity in local populations in central and northern Europe is expected to resemble that of southern refugia (Rosenberger *et al.* 2013). To test these hypotheses leaf litter of deciduous and coniferous forests were collected at 19 locations in 13 countries in Europe, including northwest and middle-west Russia, and the known Pleistocene refuge areas of the Balkan Peninsula, Iberian Peninsula and Italy.

In Chapter 3 I analyzed differences in recolonization patterns and genetic variance of one sexual and one parthenogenetic Collembola species. During Quaternary ice-ages glaciers covered wide areas of central and northern Europe and excluded animals from these areas. After warming, populations of Collembola likely recolonized the formerly glaciated areas. However, due to differences in population growth, colonization potential and adaptability and mutation accumulation which are related to the reproductive mode, the recolonization pattern is expected to differ between sexual and parthenogenetic species. The genetic structure will be lower in the parthenogenetic than the sexual species, because only few haplotypes invaded the new habitats, originating from areas south of the Alps. The number of haplotypes in locations will be higher in the sexual species as recombination generates genetic variance and enables faster adaption to changing environments. Further, recombination can facilitate the access to a wider range of resources and therefore coexistence of genetic lineages that differ slightly in their spectrum of resource utilization. Accordingly, local endemism of lineages will be more pronounced in the sexual species due to slower population growth and lower colonization potential.

To test these hypotheses I collected leaf litter of deciduous and coniferous forests in 37 locations in Europe. Compared to Chapter 2 the sampling was expanded to locations from northern Europe (Denmark, northern Germany, Great Britain and Sweden), central Europe (Czech Republic, middle and south-eastern France, Hungary and Slovakia), southern Europe (Macedonia, Slovenia and Ukraine) and Greenland.

In Chapter 4 I investigated the genetic patterns of the ubiquitous, generalist, hemiedaphic and parthenogenetic species *P. notabilis* in Europe. A previous study of

Porco *et al.* (2012b) found four lineages within this species that represent 'cryptic species', due to pronounced genetic differences but morphological uniformity. Patterns of deep divergences, possible survival in southern refugia during the last ice-age, ancient and recent divergences were not investigated in this study. I expected that the generalist life-style of *P. notabilis* and its ubiquitous presence in anthropogenic locations, such as urban fields, agricultural sites, forests and meadows, and the high activity on soil surfaces have consequences on the dispersal of *P. notabilis*. Human-mediated dispersal may be reflected in genetic patterns and colonization patterns. If human history will be reflected in genetic patterns of *P. notabilis,* the phylogenetic analysis will show a mixture of northern, central and southern European locations. If *P. notabilis* survived during Quaternary glaciations in southern refugia, local molecular variance and genetic distances within populations from southern Europe is expected to be higher than in northern and central Europe. If founder effects exist due to few, fast colonizing individuals that grow fast and expanded rapidly, the majority of sampling locations will be dominated by a single haplotype. In addition, I expected that *P. notabilis* is comprised of more than the four genetic lineages suggested by the previous study, because known locations with high genetic variance such as southern refuge regions of the Balkan Peninsula were not included in the study of Porco *et al.* (2012b). To test these hypotheses, about two square meters of deciduous and coniferous forests were sampled in 26 locations in Europe. Compared to the study of Porco *et al.* (2012b), the sampling was extended to more southern regions, like the Balkan Peninsula, Ukraine and Turkey and to more northern and eastern European locations like Norway, Great Britain, Greenland and Russia.

References

Arnqvist G, Dowling DK, Eady P *et al.* (2010) Genetic architecture of metabolic rate: environment specific epistasis between mitochondrial and nuclear genes in an insect. *Evolution; International Journal of Organic Evolution*, **64**, 3354–3363.

Avise JC (1994) *Molecular Markers, Natural History and Evolution* (JC Avise, Ed,). Chapman & Hall, New York.

Avise JC (1998) The history and purview of phylogeography: a personal reflection. *Molecular Ecology*, **7**, 371–379.

Avise J (2000) *Phylogeography: The History and Formation of Species.* Harvard University Press, Cambridge.

Avise JC, Arnold J, Ball RM *et al.* (1987) Intraspecific the Phlyogeography : Mitochondrial DNA Bridge Between Population Genetics and Systematics. *Annual Review of Ecology and Systematics*, **18**, 489–522.

Avise J, Ball R (1990) Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surveys In Evolutionary Biology*.

Babik W, Branicki W, Crnobrnja-Isailović J *et al.* (2005) Phylogeography of two European newt species--discordance between mtDNA and morphology. *Molecular Ecology*, **14**, 2475–2491.

Bahrndorff S, Holmstrup M, Petersen H, Loeschcke V (2006) Geographic variation for climatic stress resistance traits in the springtail *Orchesella cincta*. *Journal of Insect Physiology*, **52**, 951–959.

Barton NH, Charlesworth B (1998) Why sex and recombination? *Science (New York, N.Y.)*, **281**, 1986–1990.

Bauzà-Ribot MM, Jaume D, Fornós JJ, Juan C, Pons J (2011) Islands beneath islands: phylogeography of a groundwater amphipod crustacean in the Balearic archipelago. *BMC Evolutionary Biology*, **11**, 221.

Beaumont M (2004) Recent developments in genetic data analysis: what can they tell us about human demographic history? *Heredity*, **92**, 365–79.

Behan-Pelletier VM (2003) Acari and Collembola biodiversity in Canadian agricultural soils. *Canadian Journal of Soil Science*, **83**, 279–288.

Beheregaray LB (2008) Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology*, **17**, 3754–3774.

Beheregaray LB, Caccone A (2007) Cryptic biodiversity in a changing world. *Journal of Biology*, **6**, 9.

Bell G (1982) *The Masterpiece of Nature: The Evolution and Genetics of Sexuality.* (G Bell, Ed,). University of California Press, Berkeley.

Block W, Zettel J (2003) Activity and dormancy in relation to body water and cold tolerance in a winter-active springtail (Collembola). *European Journal of Entomology*, **100**, 305-312.

Boyer SL, Baker JM, Giribet G (2007) Deep genetic divergences in *Aoraki denticulata* (Arachnida, Opiliones, Cyphophthalmi): a widespread "mite harvestman" defies DNA taxonomy. *Molecular Ecology*, **16**, 4999–5016.

Brower A V (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 6491–6495.

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

Chahartaghi M, Scheu S, Ruess L (2006) Sex ratio and mode of reproduction in Collembola of an oak-beech forest. *Pedobiologia*, **50**, 331–340.

Chernova NM, Potapov MB, Savenkova YY *et al.* (2010) Ecological Significance of Parthenogenesis in Collembola. *Entomological Review*, **90**, 23–38.

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

Cicconardi F, Fanciulli PP, Emerson BC (2013) Collembola, the biological species concept and the underestimation of global species richness. *Molecular Ecology*, **22**, 5382–5396.

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

Clark PU, Dyke AS, Shakun JD *et al.* (2009) The Last Glacial Maximum. *Science (New York, N.Y.)*, **325**, 710–714.

Cruickshank RH (2002) Molecular markers for the phylogenetics of mites and ticks. *Society*, **7**, 3–14.

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 Biological Sciences*, **269**, 1143–1151.

Van Doninck K, Schön I, De Bruyn L *et al.* (2002) A general purpose genotype in an ancient asexual. *Oecologia*, **132**, 205–212.

Dowling DK, Abiega KC, Arnqvist G (2007) Temperature-specific outcomes of cytoplasmic-nuclear interactions on egg-to-adult development time in seed beetles. *Evolution; International Journal of Organic Evolution*, **61**, 194–201.

Edmands S (2001) Phylogeography of the intertidal copepod *Tigriopus californicus* reveals substantially reduced population differentiation at northern latitudes. *Molecular Ecology*, **10**, 1743–1750.

Ellison CK, Niehuis O, Gadau J (2008) Hybrid breakdown and mitochondrial dysfunction in hybrids of *Nasonia* parasitoid wasps. *Journal of Evolutionary Biology*, **21**, 1844–1851.

Fiera C, Ulrich W (2012) Spatial patterns in the distribution of European springtails (Hexapoda: Collembola). *Biological Journal of the Linnean Society*, **105**, 498–506.

Filser J (2002) The role of Collembola in carbon and nitrogen cycling in soil. *Pedobiologia*, **46**, 234–245.

Fløjgaard C, Normand S, Skov F *et al.* (2009) Ice age distributions of European small mammals: insights from species distribution modelling. *Journal of Biogeography*, **36**, 1152–1163.

Fontaneto D, Boschetti C, Ricci C (2008) Cryptic diversification in ancient asexuals: evidence from the bdelloid rotifer *Philodina flaviceps*. *Journal of Evolutionary Biology*, **21**, 580–587.

Fountain MT, Hopkin SP (2004) A comparative study of the effects of metal contamination on Collembola in the field and in the laboratory. *Ecotoxicology*, **13**, 573–587.

Frati F, Dell'Ampio E, Casasanta S *et al.* (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**, 456–461.

Gagnaire P-A, Normandeau E, Bernatchez L (2012) Comparative genomics reveals adaptive protein evolution and a possible cytonuclear incompatibility between European and American Eels. *Molecular Biology and Evolution*, **29**, 2909–2919.

Gass F, Gillet S, Ponge J (2006) The use of directional traps for the assessment of short-term phenanthrene effects upon soil springtail communities. *Environmental Pollution*, **2**, 364–370.

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

Giribet G, Carranza S, Baguna J *et al.* (1996) First molecular evidence for the existence of a Tardigrada + Arthropoda clade. *Molecular Biology and Evolution*, **13**, 76–84.

Goto H (1960) Facultatice Parthenogenesis in Collembola (Insecta). *Nature*, **180**,958-959.

Greenslade P, Bell L, Florentine S (2011a) Auditing revegetated catchments in southern Australia: Decomposition rates and collembolan species assemblages. *Soil Organisms*, **83**, 433–450.

Greenslade P, Stevens M, Torricelli G *et al.* (2011b) An ancient Antarctic endemic genus restored: morphological and molecular support for *Gomphiocephalus hodgsoni* (Collembola: Hypogastruridae). *Systematic Entomology*, **36**, 223–240.

Hädicke C, Haug C, Haug J (2013) Adding to the few: a tomocerid collembolan from Baltic amber. *Paleodiversity*, **6**, 149–156.

Hamilton W (1980) Sex versus Non-Sex versus Parasite. *Oikos*, **35**, 282–290.

Healey I (1967) An ecological study of temperatures in a Welsh moorland soil. *Journal of Animal Ecology*, **36**, 425–434.

Hebert PDN, Ratnasingham S, deWaard JR (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings. Biological sciences / The Royal Society*, **270**, 96–99.

Heethoff M, Domes K, Laumann M *et al.* (2007) High genetic divergences indicate ancient separation of parthenogenetic lineages of the oribatid mite Platynothrus peltifer (Acari, Oribatida). *Journal of Evolutionary Biology*, **20**, 392–402.

Heidemann K, Hennies A, Schakowske J *et al.* (2014) Free-living nematodes as prey for higher trophic levels of forest soil food webs. *Oikos*, **23**, 1-13.

Hewitt G (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hewitt GM (2001) Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Molecular Ecology*, **10**, 537–49.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **359**, 183–195.

Hewitt GM, Ibrahim KM (2001) Inferring glacial refugia and historical migrations with molecular phylogenies. In: *Integrating Ecology and Evolution in a Spatial Context* (eds Silvertown J, Antonovics J), pp. 271–294. Blackwell Science Ltd., Oxford.

Hickerson MJ, Carstens BC, Cavender-Bares J *et al.* (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution*, **54**, 291–301.

Hirst S, Maulik S (1926) On some Arthropod Remains from the Rhynie Chert (Old Red Sandstone). *Geological Magazine*, **63**, 69–71.

Hoekstra LA, Siddiq MA, Montooth KL (2013) Pleiotropic effects of a mitochondrial-nuclear incompatibility depend upon the accelerating effect of temperature in *Drosophila*. *Genetics*, **195**, 1129–39.

Homburg K, Drees C, Gossner MM *et al.* (2013) Multiple Glacial Refugia of the Low-Dispersal Ground Beetle *Carabus irregularis*: Molecular Data Support Predictions of Species Distribution Models (J Chave, Ed,). *PLoS ONE*, **8**, e61185.

Hooper DU, Bignell DE, Brown VK *et al.* (2000) Interactions between Aboveground and Belowground Biodiversity in Terrestrial Ecosystems: Patterns, Mechanisms, and Feedbacks. *BioScience*, **50**, 1049.

Hopkin SP (1997) *Biology of the Springtails*. Oxford University Press, Oxford.

Ingimarsdóttir M, Caruso T, Ripa J *et al.* (2012) Primary assembly of soil communities: disentangling the effect of dispersal and local environment. *Oecologia*, **170**, 745–54.

Kaneda S, Kaneko N (2006) Effects of collembola on nutrient cycling and microbial biomass on intact soil microcosm. *Soil Science and Plant Nutrition*, **52**, 577–577.

Kardol P, Wardle DA (2010) How understanding aboveground-belowground linkages can assist restoration ecology. *Trends in Ecology and Evolution*, **25**, 670–679.

Kohn LM (2005) Mechanisms of fungal speciation. *Annual review of phytopathology*, **43**, 279–308.

Kondrashov A (1993) Classification of hypotheses on the advantage of amphimixis. *Journal of Heredity*, **84**, 372–387.

Konestabo HS, Michelsen A, Holmstrup M (2007) Responses of springtail and mite populations to prolonged periods of soil freeze-thaw cycles in a sub-arctic ecosystem. *Applied Soil Ecology*, **36**, 136–146.

Langeneckert A (2013) Lipid composition of Collembola from different forest types (beech, spruce) as indicated by fatty acid analyses.

Lawrence P (1985) Ten species of Collembola from Baltic amber. *Prace Muzeum Ziemi*, **37**, 101–104.

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

Magri D (2008) Patterns of post-glacial spread and the extent of glacial refugia of European beech (*Fagus sylvatica*). *Journal of Biogeography*, **35**, 450–463.

Maraun M, Heethoff M, Schneider K *et al.* (2004) Molecular phylogeny of oribatid mites (Oribatida, Acari): evidence for multiple radiations of parthenogenetic lineages. *Experimental & Applied Acarology*, **33**, 183–201.

Mark Welch D, Meselson M (2000) Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. *Science*, **288**, 1211–1215.

Martens K, Rossetti G, Horne DJ (2003) How ancient are ancient asexuals? *Proceedings. Biological sciences / The Royal Society*, **270**, 723–729.

Maura M, Salvi D, Bologna MA, Nascetti G, Canestrelli D (2014) Northern richness and cryptic refugia: phylogeography of the Italian smooth newt *Lissotriton vulgaris meridionalis*. *Biological Journal of the Linnean Society*, **113**, 590–603.

McInerney CE, Maurice L, Robertson AL *et al.* (2014) The ancient Britons: groundwater fauna survived extreme climate change over tens of millions of years across NW Europe. *Molecular Ecology*, **23**, 1153–1166.

Montooth KL, Meiklejohn CD, Abt DN *et al.* (2010) Mitochondrial-nuclear epistasis affects fitness within species but does not contribute to fixed incompatibilities between species of Drosophila. *Evolution; International Journal of Organic Evolution*, **64**, 3364–3379.

Moritz C, Patton J (2000) Diversification of rainforest faunas: an integrated molecular approach. *Annual Review of Ecology and Systematics*, **31**, 533-563

Mortimer E, Jansen van Vuuren B, Lee JE *et al.* (2011) Mite dispersal among the Southern Ocean Islands and Antarctica before the last glacial maximum. *Proceedings. Biological sciences / The Royal Society*, **278**, 1247–1255.

Muller H (1964) The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **1**, 2–9.

Nitzu E, Nae A, Giurginca A *et al.* (2010) Invertebrate communities from the mesovoid shallow substratum of the carpatho-euxinic area: eco-faunistic and zoogeographic analysis. *Travaux de l'Institut de Spéologie "Émile Racovitza"*, **49**, 41–79.

Norton R, Palmer S (1991) The distribution, mechanisms, and evolutionary significance of parthenogenesis in oribatid mites. In: *The Acari: Reproduction, Development and Life-history Strategies* (eds Schuster R, Murphy P), pp. 107–136. Chapman and Hall, London.

Ohlsson L, Verhoef H (1988) Effects of diet composition on cold adaptation in temperate Collembola. *Comparative Biochemistry and Physiology*, **60,** 1152-1159.

Paul D (2011) Collembolan Density and Diversity in a Forest and an Agroecosystem. *Open Journal of Soil Science*, **01**, 55–61.

Petersen H (1978) Sex-ratios and the extant of parthenogenetic reproduction in some collembolan populations. In: *First International Seminary on Apterygota* (ed Dallai R), pp. 19–35. Accademia delle Scienze di Siena detta de' Fisiocritici, Siena.

Petersen H (2002) General aspects of collembolan ecology at the turn of the millennium. *Pedobiologia*, **46**, 246–260.

Petersen H, Luxton M (1982) A Comparative Analysis of Soil Fauna Populations and Their Role in Decomposition Processes. *Oikos*, **39**, 288.

Pilipovic J, Burns GP, Thorne MA *et al.* (2008) Cold hardening processes in the Antarctic springtail, *Cryptopygus antarcticus*: clues from a microarray. *Journal of Insect Physiology*, **54**, 1356-1362

Ponge JF (1991) Food resources and diets of soil animals in a small area of Scots pine litter. *Geoderma*, **49**, 33–62.

Porco D, Bedos A, Greenslade P (2012a) Challenging species delimitation in Collembola: cryptic diversity among common springtails unveiled by DNA barcoding. *Invertebrate Systematics*, **26**, 470–477.

Porco D, Potapov M, Bedos A *et al.* (2012b) Cryptic Diversity in the Ubiquist Species *Parisotoma notabilis* (Collembola, Isotomidae): A Long-Used Chimeric Species? *PLoS ONE*, **7**, e46056.

Potapov M (2001) Synopses on Palaearctic Collembola, Volume 3: Isotomidae. *Abhandlungen und Berichte des Naturkundemuseums Goerlitz*, **73**, 1–603.

Potapov M, Babenko A (2000) Species of the genus *Folsomia* (Collembola: Isotomidae) of northern Asia. *European Journal of Entomology*, **97**, 51–74.

Potapov M, Janion C, Deharveng L (2011) Two new species of *Parisotoma* (Collembola: Isotomidae) from the Western Cape, South Africa. *Zootaxa*, **24**, 5326.

Provan J, Bennett KD (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution*, **23**, 564–571.

Rapoport E (1971) The geographical distribution of Neotropical and Antarctic Collembola. *Pacific Insects Monograph*, **4**, 99–118.

Rosenberger M, Maraun M, Scheu S *et al.* (2013) Pre- and post-glacial diversifications shape genetic complexity of soil-living microarthropod species. *Pedobiologia*, **56**, 79–87.

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

Sattler B, Post B, D. R *et al.* (2012) *Life at Extremes: Environments, Organisms, and Strategies for Survival* (E Bell, Ed,). Cab International, Dunebeg, United Kingdom.

Schäffer S, Koblmüller S, Pfingstl T, Sturmbauer C, Krisper G (2010) Contrasting mitochondrial DNA diversity estimates in Austrian *Scutovertex minutus* and *S. sculptus* (Acari, Oribatida, Brachypylina, Scutoverticidae). *Pedobiologia*, **53**, 203–211.

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

Schmitt T (2007) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology*, **4**, 11.

Schmitt T (2009) Biogeographical and evolutionary importance of the European high mountain systems. *Frontiers in Zoology*, **6**, 9.

Schwander T, Henry L, Crespi BJ (2011) Molecular evidence for ancient asexuality in timema stick insects. *Current Biology : CB*, **21**, 1129–34.

Seddon JM, Santucci F, Reeve NJ *et al.* (2001) DNA footprints of European hedgehogs, *Erinaceus europaeus* and *E. concolor*: Pleistocene refugia, postglacial expansion and colonization routes. *Molecular Ecology*, **10**, 2187–2198.

Sekiné K, Hayashi F, Tojo K (2013) Phylogeography of the East Asian polymitarcyid mayfly genus *Ephoron* (Ephemeroptera: Polymitarcyidae): a comparative analysis of molecular and ecological. *Biological Journal of the Linnean Society*, **109**, 181–202.

Sommer RS, Benecke N (2005) The recolonization of Europe by brown bears *Ursus arctos* Linnaeus, 1758 after the Last Glacial Maximum. *Mammal Review*, **35**, 156–164.

Sommer RS, Zachos FE (2009) Fossil evidence and phylogeography of temperate species: "glacial refugia" and post-glacial recolonization. *Journal of Biogeography*, **36**, 2013–2020.

Stevens MI, Hogg ID (2003) Long-term isolation and recent range expansion from glacial refugia revealed for the endemic springtail *Gomphiocephalus hodgsoni* from Victoria Land, Antarctica. *Molecular Ecology*, **12**, 2357–2369.

Stevens MI, Hogg ID (2006) Contrasting levels of mitochondrial DNA variability between mites (Penthalodidae) and springtails (Hypogastruridae) from the Trans-Antarctic Mountains suggest long-term effects of glaciation and life history on substitution rates, and speciation processes. *Soil Biology and Biochemistry*, **38**, 3171–3180.

Stewart J, Cooper A (2008) Ice Age refugia and Quaternary extinctions: An issue of Quaternary evolutionary palaeoecology. *Quaternary Science Reviews*, **27**, 2443–2448.

Stewart JR, Lister AM (2001) Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology & Evolution*, **16**, 608–613.

Subías LS (2009) Listado sistemático, sinonímico y biogeográfico de los ácaros oribátidos (Acariformes: *Graellsia)*, **60**, 3–305.

Sutkowska A, Pasierbiński A, Warzecha T, Mitka J (2014) Multiple cryptic refugia of forest grass Bromus benekenii in Europe as revealed by ISSR fingerprinting and species distribution modelling. *Plant Systematics and Evolution*, **300**, 1437–1452.

Taberlet P (1998a) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Taberlet P (1998b) Biodiversity at the intraspecific level: The comparative phylogeographic approach. *Journal of Biotechnology*, **64**, 91–100.

Templeton AR (2005) Haplotype trees and modern human origins. *American Journal of Physical Anthropology*, **Suppl 41**, 33–59.

Theissinger K, Bálint M, Feldheim KA *et al.* (2013) Glacial survival and post-glacial recolonization of an arctic-alpine freshwater insect (*Arcynopteryx dichroa*, Plecoptera, Perlodidae) in Europe (H-P Comes, Ed,). *Journal of Biogeography*, **40**, 236–248.

Torricelli G, Carapelli A, Convey P *et al.* (2010) High divergence across the whole mitochondrial genome in the "pan-Antarctic" springtail *Friesea grisea*: Evidence for cryptic species? *Gene*, **449**, 30–40.

Torroni A, Achilli A, Macaulay V *et al.* (2006) Harvesting the fruit of the human mtDNA tree. *Trends in genetics : TIG*, **22**, 339–345.

Villacorta C, Jaume D, Oromí P *et al.* (2008) Under the volcano: phylogeography and evolution of the cave-dwelling *Palmorchestia hypogaea* (Amphipoda, Crustacea) at La Palma (Canary Islands). *BMC Biology*, **6**, 7.

Wang B, Chang L, Wu D *et al.* (2014) Effect of Transgenic Corn Cultivation and Sampling Location on Feeding Habits of Collembola. *Journal of Agricultural Science and Technology A 4*, **4**, 235–242.

Wardle DA, Bardgett RD, Klironomos JN *et al.* (2004) Ecological linkages between aboveground and belowground biota. *Science*, **304**, 1629–1633.

Weisrock DW, Janzen FJ (1999) Thermal and fitness-related consequences of nest location in Painted Turtles (*Chrysemys picta*). *Functional Ecology*, **13**, 94–101.

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

Williams G (1975) *Sex and Evolution* (G Williams, Ed,). Princeton University Press, Princeton.

Wolff JN, Ladoukakis ED, Enríquez JA *et al.* (2014) Mitonuclear interactions: evolutionary consequences over multiple biological scales. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*, **369**, 20130443.

Worland MR (2005) Factors that influence freezing in the sub-Antarctic springtail *Tullbergia antarctica*. *Journal of Insect Physiology*, **51**, 881–94.

Van der Wurff AWG, Gols R, Ernsting G *et al.* (2005) Population genetic structure of *Orchesella cincta* (Collembola; Hexapoda) in NW Europe, as revealed by microsatellite markers. *Pedobiologia*, **49**, 167–174.

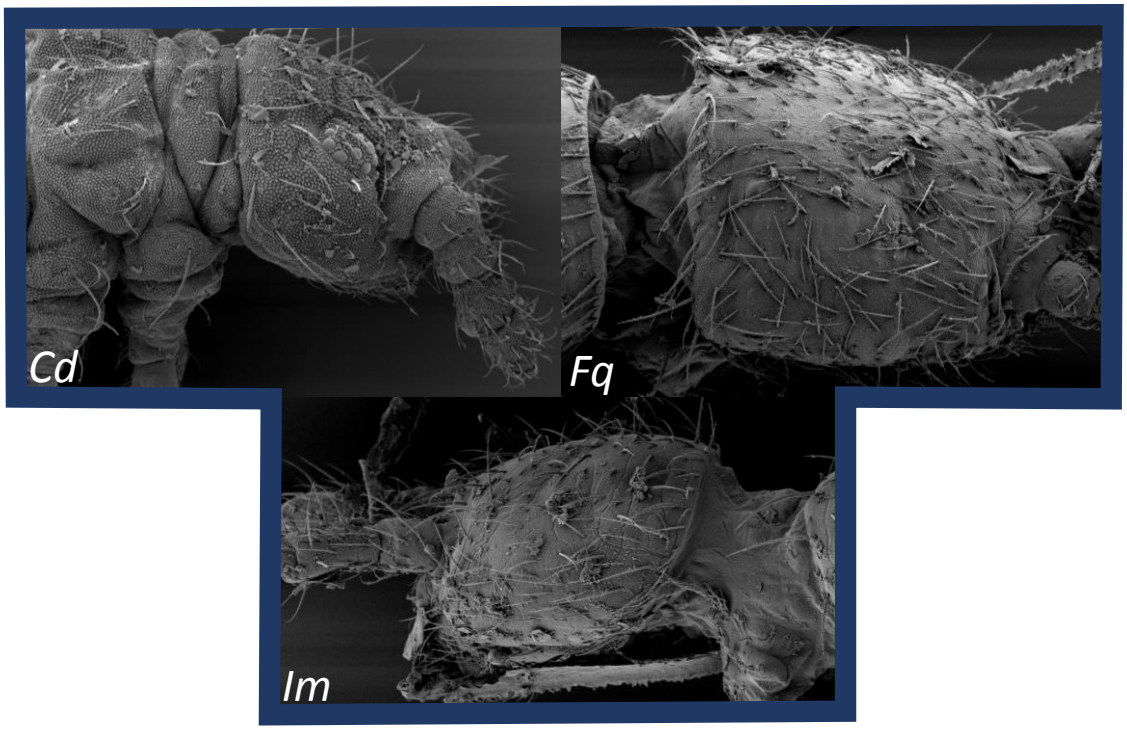
Van der Wurff AWG, Isaaks JA., Ernsting G *et al.* (2003) Population substructures in the soil invertebrate *Orchesella cincta*, as revealed by microsatellite and TE-AFLP markers. *Molecular Ecology*, **12**, 1349–1359.

Zawischa D (1993) Einschlüsse im Bernstein. *Arbeitskreis Paläontologie Hannover*, **21**, 1–32.

Chapter 2

FOUNDER EVENTS AND PRE-GLACIAL DIVERGENCES SHAPE THE GENETIC STRUCTURE OF EUROPEAN COLLEMBOLA SPECIES

Helge von Saltzwedel, Stefan Scheu and Ina Schaefer



submitted

Abstract

Collembola are ubiquitous and the most abundant arthropods in temperate soils. Eocene fossils show little morphological variation to extant taxa, suggesting persistence in stable habitats for millions of years. During the Cenozoic, extensive fluctuations in climate changed flora and fauna in central Europe, and Quaternary ice ages in particular reduced diversity and genetic structure of species living above the ground today. To evade adverse climatic conditions, Collembola move into deeper soil layers, and some species evolved frost and draught tolerances. However, if these adaptations sufficed for surviving glacial periods remains open. We investigated the phylogeographic patterns of three common species of Collembola at a pan-European scale to identify glacial refuges and post-glacial colonization patterns with three genetic markers to cover different time scales. All genes revealed remarkable genetic structure between but not within populations, suggesting density dependent processes for establishment of populations (founder-takes-all principle) which is common for European animals and plants. In contrast to the post-glacial recolonization patterns of many aboveground organisms, divergence times of most geographic lineages indicate preservation of genetic structure since the Miocene. Presumably, buffering of climatic conditions in soil and evading adverse climatic conditions enabled Collembola to survive climatic changes including those during Quaternary glaciations. This suggests that selection due to abiotic forces is weaker in the soil than above the ground, resulting in slowed down evolutionary changes in soil animal communities.

Keywords: colonization, springtail, Quaternary, founder takes it all, Miocene divergence, climate change, genetic diversity

Introduction

Collembola are small wingless hexapods that have been among the first arthropods on land, oldest fossils date to the Early Devonian ~400 mya (Hirst & Maulik 1926; Whalley & Jarzembowski 1981). Fossils from Baltic amber (55-35 mya) have been assigned to extant species (Rapoport 1971; Zawischa 1993; Weitschat & Wichard 2002; Hädicke *et al.* 2013), indicating persistence with little morphological modification over very long periods of time. They are worldwide distributed and are ubiquitous in leaf litter with more than 90% of the individuals inhabiting the upper 10 cm of the soil (Fountain & Hopkin 2004). They reach high local density in temperate forests (> 10^5 ind./m^2; Petersen & Luxton 1982) and significantly contribute to decomposition processes, soil respiration and nutrient cycling (Seastedt 1984; Wolters 1991).

The above- and belowground food web is intimately linked (Wardle *et al.* 2004) but also differs in many respects. Mobility of soil arthropods is limited due to the porous structure of soils. Further, abiotic constraints, such as temperature fluctuations and drought, are less severe in soil, and soil-living animals can evade adverse climatic conditions by moving deeper into the soil (Healey 1967; Gass *et al.* 2006). Indeed, freezing avoidance but also frost and draught tolerance is widespread in soil invertebrates including Collembola (Ohlsson & Verhoef 1988; Worland 2005; Bahrndorff *et al.* 2006; Konestabo *et al.* 2007). Consequently, Collembola are little affected by low temperature conditions and in temperate ecosystems they typically remain active during winter (Block 1982; Hopkin 1997; Zettel *et al.* 2000). This suggests that Collembola in soil may have suffered less from Quaternary climate changes than animals above the ground.

Collembola form part of the decomposer system and predominantly feed on dead organic matter and associated microorganisms (Maraun *et al.* 2003; Chahartaghi *et al.* 2005), resources which likely were at least temporarily available during Quaternary glaciations of central Europe. Therefore, present day populations of central European Collembola may well descend from relict populations that survived glacial periods, rather than from south European populations that recolonized empty habitats when glaciers retreated northwards. Effects of the recurrent and severe abiotic disturbances during Quaternary ice-ages on the genetic structure of soil-living animals has been hardly

investigated (Garrick *et al.* 2007; Rosenberger *et al.* 2013), disregarding one of the most species-rich terrestrial animal communities.

Few studies investigated the genetic structure of Collembola in Europe and the existing studies focused on large species living on the soil surface rather than soil-living species (Frati *et al.* 2000; van der Wurff *et al.* 2003, 2005; Cicconardi *et al.* 2010). Generally, there is very limited information on phylogeographic patterns of soil-living invertebrates across Europe. The only study available analyzing the genetic structure of soil-living arthropods on a wide geographic area investigated the oribatid mite species *Steganacarus magnus* (Rosenberger *et al.* 2013). Conform to the above considerations the results suggest that *S. magnus* survived Quaternary glaciations in central Europe in cryptic refugia and indicates that phylogeographic patterns of aboveground animals cannot easily be transferred to those of the belowground system. Collembola resemble oribatid mites in various respects but typically are faster reproducing, more mobile and less frost tolerant (Block 1982; Webb & Block 1993; Siepel 1994; Petersen 2002; Salomone *et al.* 2002; Lindberg & Bengtsson 2005). Therefore, the ability to survive Quaternary glaciations and the recolonization potential from glacial refuges likely differs between Collembola and oribatid mites.

We investigated the genetic structure of three species of Collembola that are widely distributed and common in the northern hemisphere across thirteen European countries, including potential refuge areas south of the Alps. *Ceratophysella denticulata* (Bagnall, 1941) (Hypogasturidae) and *Folsomia quadrioculata* (Tullberg, 1871) (Isotomidae) are large species (up to 2.5 mm) that live in the uppermost soil layer (hemiedaphic) and reproduce sexually. In contrast, the third species, *Isotomiella minor* (Schaeffer, 1896) (Isotomidae), is small (up to 1.3 mm), lives deeper in soil (euedaphic) and reproduces via parthenogenesis. The investigated species are generally described as palearctic or holarctic in their distribution ranges, only *I. minor* is considered cosmopolitan, but its taxonomic status needs further attention (Deharveng & Olieveira 1990; Bedos & Deharveng 1994; Gao & Potapov 2011). In fact, however, all three species are more widely distributed but it remains uncertain to what extent their occurrence in other regions of the world is based on introduced or indigenous populations (Greenslade & Convey 2012). *Ceratophysella denticulata* feeds on bacteria and fungi but presumably

also on other soil animals such as nematodes (Chahartaghi *et al.* 2005; Heidemann *et al.* 2014). While *C. denticulata* functions as secondary decomposer, *F. quadrioculata* is a typical primary decomposer feeding predominantly on litter (Chahartaghi *et al.* 2005). In contrast, *I. minor* presumably feeds predominantly on bacteria (Ponge 1991; Langeneckert 2013). Due to its parthenogenetic mode of reproduction, *I. minor* likely recovers more quickly from disturbances and colonizes new habitats faster than sexual species (Lindberg & Bengtsson 2005). Species of Hypogasturidae and Isotomidae have been found in Baltic amber (55-35 mya, Rapoport 1971; Lawrence 1985; Zawischa 1993; Christiansen & Pike 2002; Hädicke *et al.* 2013) suggesting that they have been present in Europe for millions of years.

To explore divergences of lineages at a wide temporal window, we analyzed one mitochondrial and two nuclear markers with different substitution rates. Assuming that soil buffered Quaternary temperature extremes, Collembola presumably survived in local patches in central Europe and expanded from these patches. Accordingly, endemic haplotypes are present in central and northern Europe. Further, we expected different genetic lineages to coexist locally due to expansions from isolated central European refuge populations. Due to more extensive survival during Quaternary glaciations of Collembola in southern Europe, the local molecular variance and genetic distances within populations from southern Europe should be higher than those within central and northern European populations.

Materials and Methods

Sampling *of animals and DNA extraction*

Leaf litter including humus layers from about two square meters of deciduous and coniferous forests was collected in 19 locations in 13 countries in Europe, including northwest and middle-west Russia, and the Pleistocene refuge areas of the Balkan (Montenegro, Serbia, Croatia, Greece), Italy and Spain, and transferred to the University of Göttingen. Animals were extracted by heat (Kempson et al. 1963), collected in 96% EtOH and stored at -20°C until further analyses. For species identification specimens were sorted under a dissecting microscope and determined by light microscopy following

Hopkin (2007). If possible, five individuals per species and sampling location were sequenced (**Table 1**). Genomic DNA was extracted from single individuals of *C. denticulata* (n=54)*, F. quadrioculata* (n=56) and *I. minor* (n=55) using the DNeasy® Blood and Tissue Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol for animal tissue. Purified DNA was eluted in 30 µl buffer AE and stored at -20°C until further preparation. Two nuclear genes (*Histone 3*, ~374 bp; D3-D5 region of 28S rDNA, ~570 bp) and the barcoding fragment of the mitochondrial COI gene (709 bp) were amplified in 25 µl volumes containing 12.5 µl SuperHot Taq Mastermix (Genaxxon Bioscience GmbH, Ulm, Germany) with 1.5 µl of each primer (10 pM), 4.5 µl H_2O, 2 µl $MgCl_2$ (25 mM) and 3 µl template DNA. The primers used and the PCR programs are given in **Table S1, Supporting information**. Positive PCR products were purified with the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol and sent for sequencing to the Göttingen Genome Laboratory (Institute for Microbiology and Genetics, Georg August University of Göttingen). All sequences are available at GenBank (KF684371-KF684865, **Table S2, Supporting information**). DNA was extracted from entire specimens but secondary vouchers (same morphological species from the same population) were deposited at our collections at J.F. Blumenbach Institute of Zoology and Anthropology, Georg August University Göttingen, Germany.

Table 1. Summary of sampling sites of three species of Collembola sampled form forests across Europe. Abbreviations of sampling locations and number of individuals sequenced for this study are listed. Sampling locations are grouped into north, central and south European regions.

	country	location	abbre-viation	*Ceratophysella denticulata*	*Folsomia quadrioculata*	*Isotomiella minor*	coordinates (N, E)
north	**Estonia**	Tallinn	EE	-	5	-	59.44° 24.69°
	Norway	Rod	NO	5	5	5	59.07° 10.23°
	Russia	Letnerechenskiy	RU1	-	4	5	64.27° 34.44°
		Nischni Novgorod	RU2	5	-	-	56.37° 43.98°
central	**Austria**	Tirol, Sonnenberg Alm	AT1	-	-	5	47.46° 12.24°
		Hittisau	AT2	5	-	-	47.46° 9.95°
		Holzgau	AT3	-	3	-	47.29° 10.33°
	France	Chartreuse	FR	5	5	4	45.42° 5.81°
	Germany	Solling, Neuhaus	DE	5	5	4	51.71° 9.64°
	Poland	Warsaw	PL	4	-	4	52.33° 20.76°
south	**Croatia**	Sljeme	HR	3	4	4	45.90° 15.95°
	Greece	Chrysovitsi	GR	5	5	5	37.56° 22.20°
	Italy	Felitto	IT1	-	5	5	40.37° 15.22°
		Berceto	IT2	3	-	-	44.50° 10.00°
	Montenegro	Bar	ME	5	5	5	42.13° 19.09°
	Serbia	Markovac	RS	5	5	-	44.22° 21.09°
	Spain	Oviedo	ES1	-	-	5	43.36° -6.00°
		Ponga	ES2	4	-	4	43.19° -5.16
		Martiartu	ES3	-	5	-	43.20° -2.90
		total no. of individuals.		**54**	**56**	**55**	

Phylogenetic analyses, divergence time estimation and population structure

Sequences were edited, ambiguous positions were corrected by hand and nucleotide sequences were translated into amino acid sequences using the invertebrate mitochondrial code implemented in Sequencher 4.10 (Gene Codes Corporation, USA). Nucleotide (28S) and protein sequences (*COI* and *H3*) were aligned separately and combined (concatenated nucleotide sequences of all three genes) for each species with Clustal W (Thompson *et al.* 1994) implemented in BioEdit 7.0.1 (Hall 1999). Each of the species was analyzed separately.

The best fit model of sequence evolution for each alignment (*COI*, 28S, *H3*, combined matrix) was inferred with to the hLRT in TOPALi v2.5 (Milne *et al.* 2009) using the PHYML algorithm. Phylogenetic trees were calculated with RAxML v7.0.3 (Stamatakis *et al.* 2005) and MrBayes v3.1.2 (Ronquist & Huelsenbeck 2003). For Maximum likelihood analyses the model of sequence evolution was GTR+I+G (all four alignments) and 10,000 bootstrap replicates were calculated. For Bayesian Inference lset parameters were nst=6, rates=invgamma (all four alignments), the mcmc chains were run for ten million generations that were sampled every 1,000th generation and a burnin of 2,500 (25%) was used. In the absence of fossil or biogeographic calibration points for the investigated species in Europe, a strict molecular clock for the *COI* nucleotide alignment was applied in BEAST v1.7.4 (Drummond *et al.* 2012). As tree prior we used the Yule Process (Gernhard *et al.* 2008) as preliminary analyses indicated quicker convergence and higher probabilities and likelihoods than coalescent tree priors. However, topologies with different tree priors did not vary. The Yule Process also is more appropriate for the genetically highly diverged lineages as the substitution rate among branches is more variable than with coalescent priors. We used the widely adopted substitution rate of 0.023 substitutions/site per million years that corresponds to a rate of 0.0115 (Avise 1994; Brower 1994).

Convergence of the mcmc chain after 600 million generations (sampled every 60,000th generation) with a burnin of 25% was confirmed using Tracer v1.4 (Drummond & Rambaut 2007). Divergence estimates were calculated with three datasets: (1) all *COI* and 28S sequences of this study and NCBI combined, with strict clock settings for *COI* (fixed rate of 0.0115) and estimated rates for 28S, (2) all *COI* sequences obtained in this

study with a strict clock (fixed rate of 0.0115), and (3) all *COI* sequences of this study extended with sequences from non-European countries obtained from NCBI and BOLD databanks with a strict clock (fixed rate of 0.0115). The extended dataset (2) and (3) included additional sequences from Antarctica, Australia, Canada, Chile, New Zealand, South Africa and northern France and a more detailed outgroup sampling for better estimation of the substitution rate, covering additional species within the respective genera and families (**Table S3, Supporting information**). For *F. quadrioculata* and *I. minor*, non-European 28S sequences of the D3-D5 region were not available, the combined datasets were extended with additional outgroup taxa only. The outgroup settings of the two isotomid species were identical, except for five additional sequences of the parthenogenetic species *Parisotoma notabilis*, to account for variance in the substitution rate due to the reproductive mode of *I. minor*.

Median-joining haplotype networks for the nucleotide datasets of *COI*, *H3*, 28S and the concatenated dataset were generated for all three species using the program Network 4.6 (Fluxus Technology, Suffolk, Great Britain). Molecular variance (AMOVA) within and between populations and isolation by distance (Mantel test) of all three genes (uncorrected p-distances) were analyzed separately in ARLEQUIN (Excoffier & Lischer 2010) with 20,000 permutations. To infer the molecular divergence times of *C. denticulata*, *F. quadrioculata* and *I. minor*, we generated a phylogenetic tree of 10 families with 23 genera and eight outgroup taxa including the taxa studied by D'Haese (2002), using 28 *COI* sequences available at NCBI and a strict molecular clock in BEAST as described above.

Results

Population structure

Isolation by distance was rejected for all datasets and all species as not significant; however, genetic variance among populations was high, explaining 96-97% (*C. denticulata*), 87-88% (*F. quadrioculata*) and 92-99% (*I. minor*) of the genetic variance of *COI*, *H3* and the concatenated dataset (*COI* and *H3*; **Table 2**). Genetic distances between populations were very high (**Table 3a**), ranging for *COI* between 14-20% in *C. denticulata*

and *I. minor* and 11-17% in *F. quadrioculata*. Within population distances were not-existing or very low (<2%). The only exceptions were one population of *C. denticulata* (Croatia: *H3*, 5.5%), two populations of *F. quadrioculata* (Croatia: *COI*, 8 % and Greece: *COI*, 9.5%; *H3*, 4.3%) and two populations of *I. minor* (Russia: *COI*, 5.8% and Germany: *COI*, 4.9%) in which distinct *COI* and *H3* lineages co-occurred. Genetic distances between and within regions (*COI* and *H3*) were highest in southern Europe and decreased towards the north, but were lowest within central Europe (**Table 3b**). Between several populations in central and northern Europe genetic distances were lower than average (**Table 3c**). Distances were low for *H3* but high for *COI* between populations from Norway, Germany and France (*F. quadrioculata;* 14%), from Montenegro and Italy (*C. denticulata*, 16%), and from Norway and Spain (*I. minor*; 14%).

The nuclear 28S rDNA (D3-D5 region) had the lowest genetic variance with little (6.2% and 1.8% in *C. denticulata* and *I. minor*, respectively) or almost no variation between populations (0.2% for *F. quadrioculata*). Accordingly, the network analysis separated only haplotypes of *C. denticulata* into distinct geographic clusters (**Fig. S1, Supporting information**).

Table 2. Variance partitioning among and within sampling locations (AMOVA) of three Collembola species (*Ceratophysella denticulata*, *Folsomia quadrioculata* and *Isotomiella minor*) sampled across Europe based on sequence variance of the mitochondrial *COI* gene, the nuclear *H3* gene and a concatenated dataset of three genes (*COI*, *H3* and 28S). Asterisks indicate significant differences at p<0.05, df, degrees of freedom.

		COI		***H3***		**conc (*COI* + *H3* + 28S)**	
	source of variation	among pop.	within pop.	among pop.	within pop.	among pop.	within pop.
C. denticulata	d.f.	11	42	11	42	11	42
	sum of squares	2,831.46	77.67	674.86	25.07	3,980.26	103.48
	variance components	56.94 Va***	1.85 Vb***	13.54 Va***	0.60 Vb***	80.07 Va***	2.46 Vb***
	% variation	96.85	3.15	95.78	4.22	97.01	2.99
	fixiation indices	Fst 0.969***		Fst 0.958***		Fst 0.970***	
F. quadrioculata	d.f.	11	44	11	44	11	44
	sum of squares	2,413.07	311.2	486.74	53.6	2,867.11	364.8
	variance components	45.57 Va***	7.07 Vb***	9.24 Va***	1.22 Vb***	54.16 Va***	8.29 Vb***
	% variation	86.57	13.43	88.35	11.65	86.72	13.28
	fixiation indices	Fst 0.866***		Fst 0.883***		Fst 0.867***	
I. minor	d.f.	11	43	11	43	11	43
	sum of squares	2,784.45	193.35	453.47	4.95	3,346.90	198.3
	variance components	54.30 Va***	4.50 Vb***	8.98 Va***	0.12 Vb***	65.45 Va***	4.61 Vb***
	% variation	92.35	7.65	98.73	1.27	93.42	6.58
	fixiation indices	Fst 0.923***		Fst 0.987***		Fst 0.934***	

Table 3. Summary of genetic distances of *COI* and *H3* in three European Collembola species (*Ceratophysella denticulata*, *Folsomia quadrioculata* and *Isotomiella minor*). Mean p-distances are in percent with standard deviation (A) between and within populations, (B) geographic regions, and of (C) geographic clusters with exceptionally low genetic distances.

(A)

			between populations						**within populations**					
			C. denticulata		***F. quadrioculata***		***I. minor***		***C. denticulata***		***F. quadrioculata***		***I. minor***	
	location		***COI***	***H3***	***COI***	***H3***	***COI***	***H3***	***COI***	***H3***	***COI***	***H3***	***COI***	***H3***
north	**EE**	**Estonia**			14 ± 1	5 ± 2					0.0	1.39		
	NO	**Norway**	16 ± 4	7 ± 2	15 ± 2	5 ± 2	15 ± 1	4 ± 2	0.06	0.0	0.06	0.0	0.11	0.0
	RU1	**Russia**	17 ± 3	7 ± 2	14 ± 1	5 ± 2	14 ± 6	4 ± 3	0.06	0.0	0.99	0.0	5.84	0.16
central	**AT**	**Austria**	15 ± 5	7 ± 3	13 ± 3	5 ± 2	15 ± 7	4 ± 3	0.0	0.0	0.0	0.0	0.0	0.0
	FR	**France**	15 ± 5	7 ± 3	14 ± 2	4 ± 2	17 ± 1	5 ± 2	0.73	0.0	1.02	0.0	0.21	0.0
	DE	**Germany**	15 ± 5	6 ± 3	13 ± 3	5 ± 2	15 ± 5	5 ± 3	1.16	0.0	0.06	0.11	4.87	0.13
	PL	**Poland**	15 ± 5	7 ± 3			15 ± 7	5 ± 3	1.03	0.0			0.0	0.0
south	**HR**	**Croatia**	14 ± 4	7 ± 2	15 ± 1	6 ± 1	18 ± 1	5 ± 2	0.09	5.53	8.04	0.0	1.6	0.53
	GR	**Greece**	19 ± 1	7 ± 1	15 ± 1	7 ± 1	17 ± 5	4 ± 2	1.27	0.48	9.51	4.33	0.0	0.0
	IT	**Italy**	17 ± 2	10 ± 1	14 ± 1	4 ± 2	17 ± 5	4 ± 2	0.0	0.0	0.0	0.0	0.06	0.0
	ME	**Montenegro**	17 ± 2	9 ± 2	15 ± 1	7 ± 1	18 ± 1	5 ± 2	0.39	0.11	1.35	0.0	0.71	0.0
	RS	**Serbia**	17 ± 3	8 ± 2	15 ± 1	6 ± 1			0.0	0.0	1.97	1.23	0.0	0.0
	ES1	**Spain**					18 ± 1	9 ± 0					1.86	0.0
	ES2	**Spain**	19 ± 1	10 ± 1			17 ± 1	5 ± 1	1.32	0.0			0.07	0.0
	ES3	**Spain**			16 ± 1	8 ± 2					1.21	0.11		

(B)

	COI									*H3*								
	C. denticulata			***F. quadrioculata***			***I. minor***			***C. denticulata***			***F. quadrioculata***			***I. minor***		
region	**N**	**C**	**S**	**N**	**C**	**S**	**N**	**C**	**S**	**N**	**C**	**S**	**N**	**C**	**S**	**N**	**C**	**S**
North (N)	10			9			11			4.2			2.7			2.8		
Central (C)	14	9		13	7		13	9		5.7	2.9		3.1	1.6		3.2	2.5	
South (S)	17	17	16	16	16	14	18	18	14	8.0	8.7	7.4	6.2	6.0	5.8	5.2	5.5	4.5

(C)

	C. denticulata		***F. quadrioculata***		***I. minor***	
geographic cluster	***COI***	***H3***	***COI***	***H3***	***COI***	***H3***
NO-ES2					14.6	0.5
ME-IT	16	5.7				
GR-IT					1.9	0.3
AT-FR	0.3	2.7				
AT-DE			5.2	2	7.3	0.5
RU2-RS	9.8	2.9				
AT-PL-RU1					0.9-4.3	0.6-1
NO-DE-FR			14.1-14.7	1.7-2.7		
HR-DE-PL-NO	6.3-10.5	2.4-7.9				

Haplotype networks and phylogenetic analyses

Haplotype diversity of the mitochondrial gene was generally high in the three investigated species (**Table 4**) but most haplotypes differed in only few positions. Haplotype networks of the *COI* and *H3* gene were complex in each of the species as the majority of populations were separated by >10 mutation steps (**Figs. S1-S3, Supporting information**). For simplifying the dataset, we grouped the *COI* haplotypes and *H3* and 28S alleles into lineages according to clades with very high [posterior probabilities (pp): 1, bootstrap (bs): 100] or high node support (pp: 0.97-99, bs: 89-100) in the phylogenetic trees based on single gene alignments (**Figs. S4-S5, Supporting information**).

Phylogenetic trees based on the combined matrix gave the best resolution and statistical support for internal and terminal nodes (**Figs. 1-3**). Topologies of ML and BI trees were similar with nearly all sampling locations clustering in separate clades with high support (pp: 0.97-1; bp: 95-100). Populations from southern European countries comprised a number of isolated *COI* and *H3* lineages which derived early in each of the trees. Central and northern European populations were more homogeneous with shorter and more derived branches. Common patterns in each of the three species were the presence of one or two closely related phylogeographic clades with a wide distribution range across central and north-east Europe (blue and green lineages). Sampling locations in the Mediterranean were not covered by phylogeographic clades but rather by genetically distinct and phylogenetically isolated lineages (orange lineages). Interestingly, both isotomid species had phylogenetically related lineages with disjunct distributions. In *F. quadrioculata*, individuals from Spain and Croatia (black lineage) were monophyletic in the single gene phylogenetic trees and carried identical 28S sequences. In *I. minor*, individuals from Norway and Spain had distinct mitochondrial haplotypes, but nearly identical *H3* and 28S alleles. With one exception (population HR in *F. quadrioculata*) coexistence of genetic lineages did not occur in each of the three species. Different to the other species, *C. denticulata* had a central eastern phylogeographic clade (black lineage) but the geographic range and spatial coherence of this lineage remains open due to the limited number of sampling locations.

Table 4. Number of haplotypes (*COI*) and different alleles (*H3*, 28S) of the three investigated species of Collembola.

species	no. individuals	no. of haplotypes / alleles		
		COI	***H3***	**28S**
C. denticulata	54	27	18	8
F. quadrioculata	56	31	23	3
I. minor	55	22	15	8

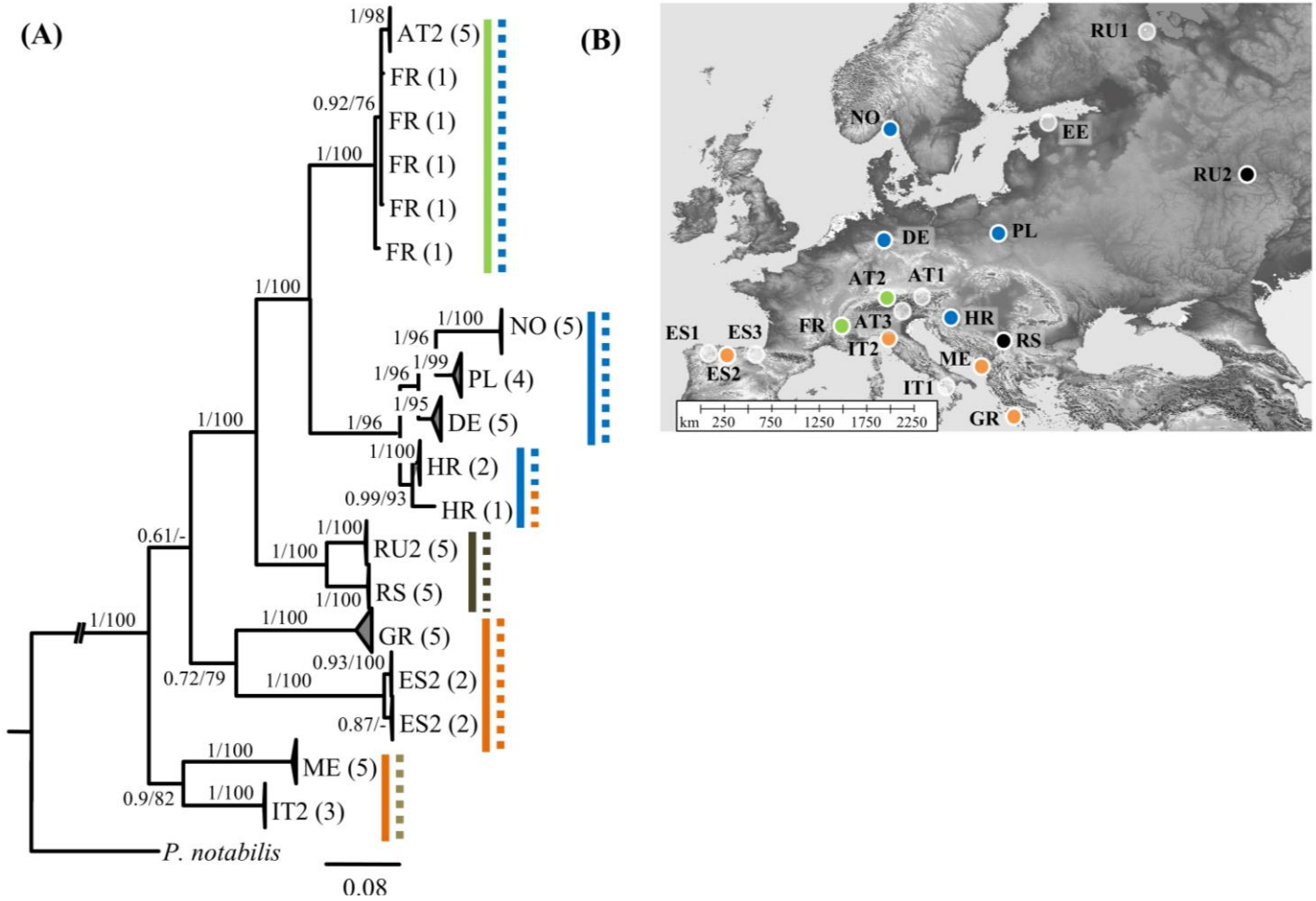


Figure 1. Relationships of lineages of *Ceratophysella denticulata* in Europe derived from (A) Bayesian analysis of sequence data of three genes; distinct genetic lineages are indicated by solid (*COI*) and dashed (*H3*) lines. (B) Sampling locations and geographic distribution of *COI* lineages; white sampling locations are not represented by this species. Terminal clades in the phylogenetic tree have been collapsed and numbers of individuals included in each clade are indicated by numbers in brackets. Numbers on nodes are posterior probabilities (Bayesian Inference) and bootstrap values (Maximum Likelihood).

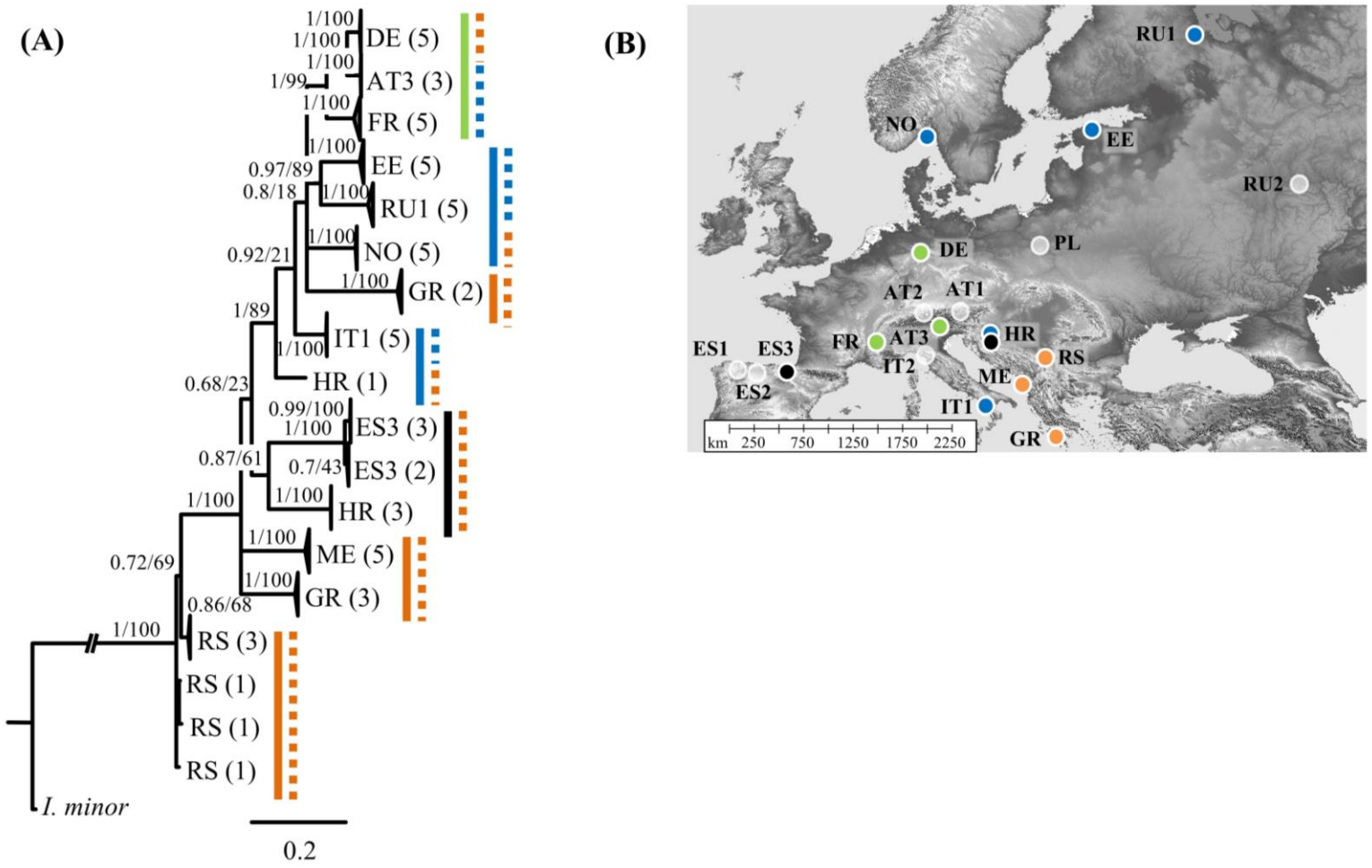


Figure 2. Relationships of lineages of *Folsomia quadrioculata* in Europe derived from (A) Bayesian analysis of sequence data of three genes; distinct genetic lineages are indicated by solid (*COI*) and dashed (*H3*) lines. (B) Sampling locations and geographic distribution of *COI* lineages; white sampling locations are not represented by this species.

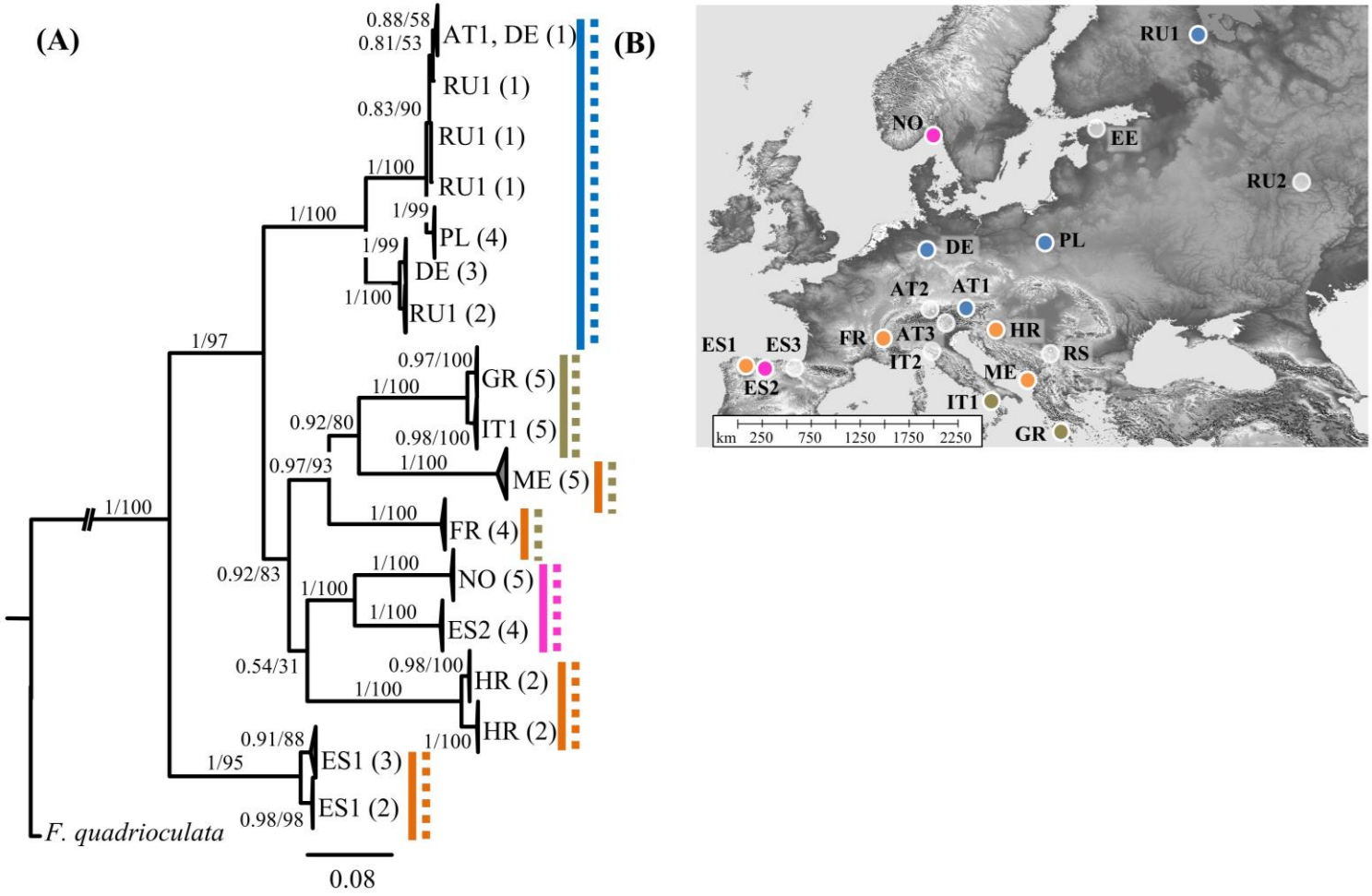


Figure 3. Relationships of lineages of *Isotomiella minor* in Europe derived from (A) Bayesian analysis of sequence data of three genes; distinct genetic lineages are indicated by solid (*COI*) and dashed (*H3*) lines. (B) Sampling locations and geographic distribution of *COI* lineages; white sampling locations are not represented by this species.

Estimation of divergence times

Trees for molecular divergence estimates calculated with *COI* in BEAST differed only slightly from those generated with the combined matrix (**Fig. S6, Supporting information**). All populations and clades that comprised larger geographic areas were recovered, only positions of populations from Italy and Montenegro (*C. denticulata*), Greece (*F. quadrioculata*), Croatia and Montenegro (*I. minor*) differed from the trees calculated with the combined matrices. Divergence estimates of major lineages fell into the Miocene in each of the three species, considerably predating Quaternary glaciations. Results were consistent among the three datasets, demonstrating robustness of the molecular clock analyses (**Fig. 4**).

In each of the three species and datasets three major radiation events occurred (**Fig. 4, Figs. S6, S8-S10, Supporting information**). First, large phylogeographic clades of lineages now present in central and north-east Europe separated from south European lineages during the Early and Middle Miocene (23.3-11.6 mya); second, clades in central and northern Europe separated during the Late Miocene (11.6-5.4 mya), and third, populations, i.e. sampling locations, were of Pleistocene origin (0.011-1.8 mya). Divergence estimates of the *COI* datasets were consistent, divergences in dataset (2) without additional outgroups were even more conserved, i.e. lineages in general were younger. All three datasets of *C. denticulata* suggest that the lineages diverged about 0.03-2.3 mya. In the two isotomid species (*F. quadrioculata* and *I. minor*) divergence estimates of larger phylogeographic clades were younger in dataset (1) (combined *COI* and 28S) compared to datasets (2)-(3) (*COI* only), with predominantly Middle Miocene origin (16-11.5 mya). However, this probably is an artefact of the 28S gene which is strongly conserved among genera of *Folsomia* (H. von Saltzwedel, unpublished data) and probably other Isotomidae. In general, divergences of clades and populations were more recent in *F. quadrioculata* than in the other two species but still concentrated in the Miocene, independent of the dataset analyzed.

The phylogeographic dataset can only estimate the molecular age of European lineages of the respective species, which ranged between 22.5 and 25.8 my for *C. denticulata*, between 12.8 and 20.6 my for *F. quadrioculata* and between 17.2 and 24.9 my for *I. minor*, depending on the dataset analyzed. Estimated divergence times of the

three Collembola species based on a phylogeny of 28 *COI* sequences can only infer the ages of the respective genera, which were about twice as old with 48 ± 13 my for *C. denticulata*, 56 ± 14 my for *F. quadrioculata* and 49 ± 15 my for *I. minor* (**Fig. S7, Supporting information**). The age of the species must be between these two age estimates. As both analyses do not conflict with the fossil record, we take this as support for our molecular clock analyses and the Miocene origin of large phylogeographic clades in European Collembola.

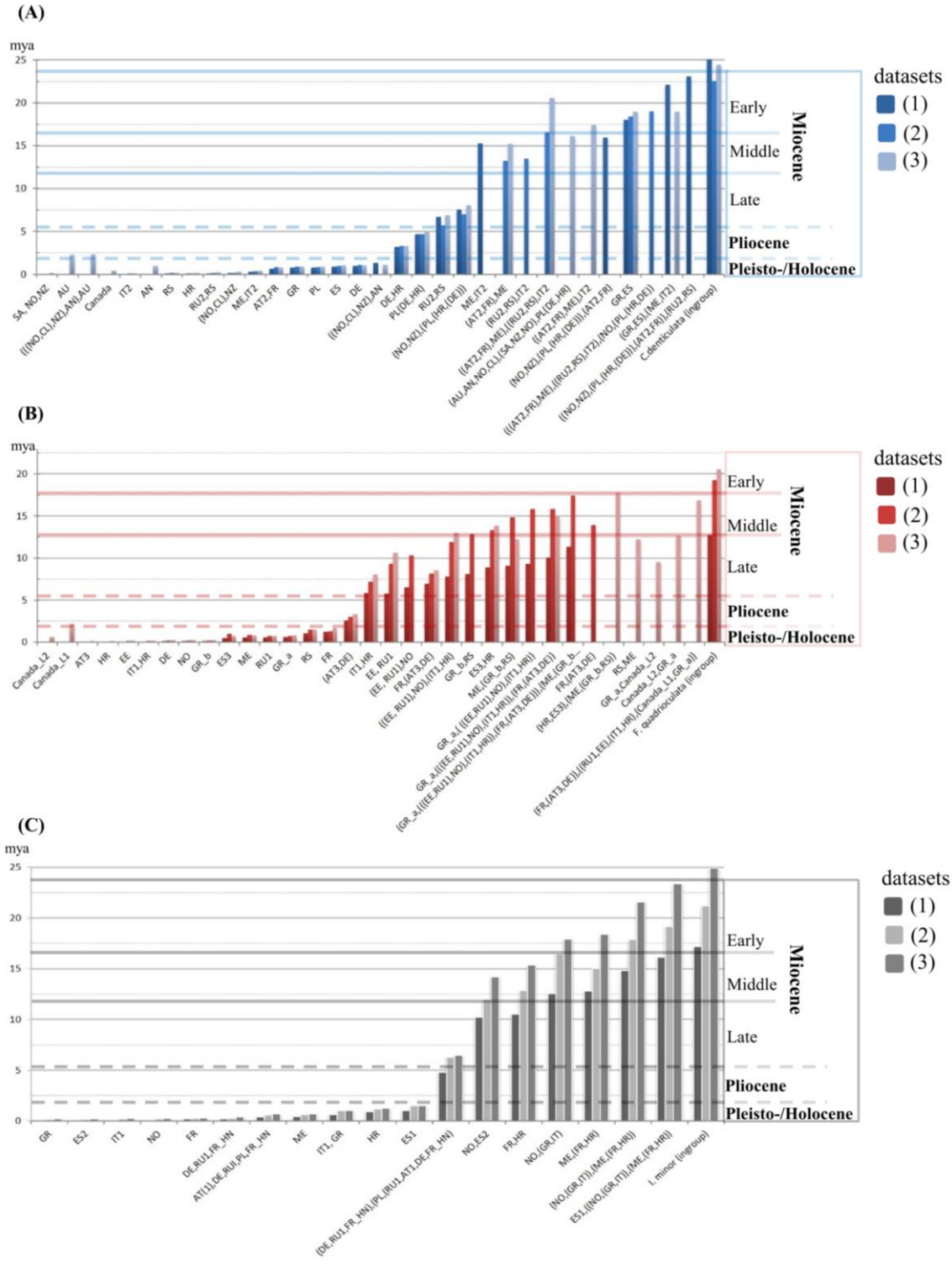


Figure 4. Comparison of molecular divergence time estimates of three datasets of (A) *Ceratophysella denticulata*, (B) *Folsomia quadrioculata* and (C) *Isotomiella minor* calculated with BEAST. Adjacent columns indicate divergence times per node, the left column (dark color) represents dataset (1; 28S+*COI*), the central column (light color) refers to dataset (2; *COI* only this study) and the right column (intermediate color) dataset (3; *COI* with additional non-European taxa). Sampling locations included in nodes are indicated on the x-axis, for topology of the phylogenetic trees see Figs. S6 and S8-S10.

Discussion

This study investigated the genetic structure of Collembola at a pan-European scale. The results suggest three largely consistent phylogeographic patterns of the three investigated species. First, populations of Collembola in Europe are highly structured with high genetic variance between, but low variance within populations. Second, molecular divergence estimates and genetic distances suggest that common ancestors of present day populations in central and southern Europe probably persisted in isolated populations for millions of years. This is in agreement with our hypothesis that Quaternary climate changes were less severe for soil animals than for aboveground animals. Third, Miocene divergences dominate among geographic lineages indicating that diversification events correlate with wet and warm climate, and a biome change in central and eastern Europe from forest to grassland. The presence of the three species of Collembola in Europe during the Miocene is supported by their estimated Eocene origin that also correlates well with the fossil record from Baltic amber (Rapoport 1971; Zawischa 1993; Weitschat & Wichard 2002; Hädicke *et al.* 2013).

In each of the three species, the majority of sampling locations were dominated by a single haplotype, suggesting that only few early colonizing individuals founded the populations which grew and expanded rapidly thereby preventing invasion of other lineages. This is conform to the founder-takes-it-all process (Waters *et al.* 2013) resulting in low genetic variance within but high variance between populations. Molecular divergence estimates of populations substantially predated Quaternary glaciations and originated in the Miocene. Long-term persistence of distinct genetic lineages without mixing of populations presumably is related to low mobility and high local abundances of springtails.

Strong molecular differentiation between populations indicates the existence of cryptic species. Indeed, recent molecular studies suggest that due to the widespread occurrence of cryptic species the number of Collembola species may be magnitudes higher than current estimates based on morphology (Emerson *et al.* 2011; Porco *et al.* 2012; Cicconardi *et al.* 2013). Despite the distinctness of genetic lineages in the investigated Collembola species, we consider each of them as morphologically coherent species complex with high intraspecific genetic variability in the mitochondrial gene as

evidence for the existence of different biological species is missing. High genetic variation in *COI* have been reported from other arthropods in particular those living in soil, including Collembola (Torricelli *et al.* 2010; Porco *et al.* 2012), oribatid mites (Heethoff *et al.* 2007; Schäffer *et al.* 2010; Rosenberger *et al.* 2013) and Opilionida (Boyer *et al.* 2007). In the sexual species *C. denticulata* genetic variance in nuclear genes was high but genetic distances within the slowly evolving 28S gene still were below the intra-lineage threshold previously estimated for this species (Porco *et al.* 2012).

Genetic distances between and within regions were high in the south compared to distances between and within the central and northern European regions. This suggests that Collembola follow the "southern richness and northern purity" scenario of recolonization of central and northern Europe (Hewitt & Ibrahim 2001). Lower distances within the central region suggest more recent expansion of populations or recent colonization of genetic lineages from east Europe or central Asia. However, expansions of geographic lineages in Collembola are considerably older than those for most aboveground animals. Pleistocene or Holocene extinctions of populations during the Quaternary and recolonization of central Europe thereafter (Hewitt 1999) little affected population structure. Rather, divergence times of most clades indicate pronounced radiations into distinct geographic lineages during the Miocene (20-5 mya).

The Miocene was warm and humid resembling subtropical climate with central Europe separating later into marine and continental climatic zones along west – east gradients with seasonality increasing to the east (Bruch *et al.* 2007). For little sclerotized arthropods, susceptible to desiccation such as Collembola, the warm climate and high precipitation during the Miocene potentially facilitated large scale expansion. Accordingly, our divergence estimates indicate that southern and central European lineages separated during the Mid Miocene Climate Optimum (MMCO, 17-14 mya) when precipitation and temperature were higher than in the Early Miocene (Zachos *et al.* 2008). Increasing seasonality and decreasing precipitation during the Miocene also changed the vegetation across Europe favoring the spread of deciduous forests, open woodlands and grassland vegetation in eastern and southern Europe, which expanded considerably during the Late Miocene (Messinian, 7.2-5.3 mya; Mosbrugger *et al.* 2005; Bruch *et al.* 2007). Interestingly, clades of *C. denticulata* and *F. quadrioculata* that cover the

geographic regions from south-east, central-east and central Europe (Croatia, Russia, Serbia Austria, Germany, Poland) were dated to be of Late Miocene origin. The lower than average genetic distance in these clades suggest that a single or few lineages expanded into these areas, possibly due to fragmentation of forest habitats into open woodlands which caused local extinctions and adaptations of sexually reproducing Collembola. In contrast, lineages of the parthenogenetic species *I. minor* were less separated into geographic regions, but rather split continuously from a ubiquitous and widespread ancestral lineage. This is also reflected by the equal distances of haplotypes in the network of the nuclear marker (28S), indicating that no region is more ancestral to others and suggesting separation of these lineages from a common ancestor. Accordingly, the two closely related (28S and *H3*) but geographically separated populations of Norway and eastern Spain likely represent relict populations that separated from an ancestral lineage about 12 mya and survived in isolated patches. Related lineages either went extinct or were not sampled. Further, divergences into geographic lineages mainly occurred during the Mid Eocene and Pleistocene in *I. minor*, with only one divergence within the central European clade during the Late Miocene. In contrast, divergences of central European lineages of *C. denticulata* and *F. quadrioculata* predominantly took place in the Late Miocene and Early Pliocene.

Collembola are fast colonizers of newly formed and young habitats such as glacier forelands (Hågvar 2010), but the geographic origin of these founder populations are unknown. The distinctness of populations, including the most closely related populations from the east of Europe, suggests a regional recruitment of colonizers from local source populations that survived in glacial refugia and consequently diverged from more distant populations for millions of years. A more detailed sampling likely would reveal even more divergent haplotypes providing a more coherent picture of distribution ranges of genetic lineages. Further, a more detailed sampling in north-east and central-east Europe may uncover if north European Collembola descended from geographically distant source populations in Eurasia or from genetically distinct source populations from within Europe.

Southern European populations were dominated by isolated lineages, suggesting that the Alps function as major dispersal barrier. Accordingly, regional extinctions in central Europe and rapid recolonization of empty habitats from the surrounding rather

than from southern refugia, likely resulted in the strong phylogeographic structure of populations north of the Alps. Strong genetic differentiation, little or no mixing between populations and ancient diversifications resemble the pattern reported for epigeic springtail species of the genus *Lepidocyrtus* in the Mediterranean basin (Cicconardi *et al.* 2010) and the soil-living oribatid mite species *S. magnus* (Rosenberger *et al.* 2013). This indicates that evolutionary changes in soil are much slower and follow different trajectories as in aboveground invertebrates. However, coexistence of genetically distinct lineages in populations of the oribatid mite species *S. magnus* was more common than in the studied Collembola. This suggests that different factors regulate the establishment of populations among belowground invertebrates and that competition and early habitat colonization is more important for springtails than for oribatid mites (Lindberg & Bengtsson 2005; Hågvar 2010; Ingimarsdóttir *et al.* 2012).

Overall, the results suggest that populations of Collembola across Europe are highly structured and that diversifications into regional lineages predominantly occurred in the Miocene. Divergence dates and geographic structure suggest that soil microarthropods are less affected by global and long-term climatic changes than aboveground animals and plants. Buffering of adverse climatic conditions, the presence of organic matter colonized by bacteria and fungi in soil, large population sizes and the ability to tolerate freezing conditions likely contributed to the lower responsiveness of soil microarthropods. This suggests that in soil evolutionary processes are slowed down compared to the aboveground system, rendering soil-living detritivorous microarthropods “living fossils” with communities in central Europe reflecting Miocene rather than Quaternary diversification events.

Acknowledgements

We thank Patrick Pachl for help with sampling of the Balkan region. We are grateful to Jens Bast, Julian Bock, Elena Corral Hernández, Guido Humpert, Sven Marhan, Jordi Moya-Laraño and Lucija Šerić Jelaska for sending litter and soil material

References

Avise JC (1994) *Molecular Markers, Natural History and Evolution* (JC Avise, Ed,). Chapman & Hall, New York.

Bahrndorff S, Holmstrup M, Petersen H *et al.* (2006) Geographic variation for climatic stress resistance traits in the springtail *Orchesella cincta*. *Journal of insect physiology*, **52**, 951–959.

Bedos A, Deharveng L (1994) The Isotomiella of Thailand (Collembola: Isotomidae), with description of five new species. *Entomologica Scandinavica*, **25**, 451–460.

Block W (1982) Cold hardiness in invertebrate poikilotherms. *Comparative Biochemistry and Physiology Part A: Physiology*, **73**, 581–593.

Boyer SL, Baker JM, Giribet G (2007) Deep genetic divergences in *Aoraki denticulata* (Arachnida, Opiliones, Cyphophthalmi): a widespread “mite harvestman” defies DNA taxonomy. *Molecular Ecology*, **16**, 4999–5016.

Brower AV (1994) Rapid morphological radiation and convergence among races of the butterfly Heliconius erato inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 6491–5.

Bruch AA, Uhl D, Mosbrugger V (2007) Miocene climate in Europe — Patterns and evolution. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **253**, 1–7.

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

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

Cicconardi F, Fanciulli PP, Emerson BC (2013) Collembola, the biological species concept and the underestimation of global species richness. *Molecular Ecology*, **22**, 5382–5396.

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

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

Deharveng L, Olieveira E (1990) Isotomiella (Collembola: Isotomidae) d´Amazonie: les espèces du groupe delamarei. *Annales de la Société Entomologique de France*, **26**, 185–201.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Drummond AJ, Suchard MA, Xie D *et al.* (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 3–6.

Emerson BC, Cicconardi F, Fanciulli PP *et al.* (2011) Phylogeny, phylogeography, phylobetadiversity and the molecular analysis of biological communities. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences*, **366**, 2391–2402.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources*, **10**, 564–567.

Fountain MT, Hopkin SP (2004) Biodiversity of Collembola in urban soils and the use of Folsomia candida to assess soil "quality". *Ecotoxicology*, **13**, 555–72.

Frati F, Dell'Ampio E, Casasanta S *et al.* (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**, 456–461.

Gao Y, Potapov M (2011) *Isotomiella* (Isotomidae: Collembola) of China. *Annales de la Societe Entomologique de France*, **47**, 350–356.

Garrick RC, Sands CJ, Rowell DM *et al.* (2007) Catchments catch all: long-term population history of a giant springtail from the southeast Australian highlands - a multigene approach. *Molecular Ecology*, **16**, 1865–1882.

Gass F, Gillet S, Ponge J (2006) The use of directional traps for the assessment of short-term phenanthrene effects upon soil springtail communities. *Environmental Pollution*, **2**, 364–370.

Gernhard T, Hartmann K, Steel M (2008) Stochastic properties of generalised Yule models, with biodiversity applications. *Journal of Mathematical Biology*, **57**, 713–735.

Greenslade P, Convey P (2012) Exotic Collembola on subantarctic islands: Pathways, origins and biology. *Biological Invasions*, **14**, 405–417.

Hädicke C, Haug C, Haug J (2013) Adding to the few: a tomocerid collembolan from Baltic amber. *Paleodiversity*, **6**, 149–156.

Hågvar S (2010) Primary Succession of Springtails (Collembola) in a Norwegian Glacier Foreland. *Arctic, Antarctic, and Alpine Research*, **42**, 422–429.

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

Healey I (1967) An ecological study of temperatures in a Welsh moorland soil. *Journal of Animal Ecology*, **36**, 425–434.

Heethoff M, Domes K, Laumann M *et al.* (2007) High genetic divergences indicate ancient separation of parthenogenetic lineages of the oribatid mite *Platynothrus peltifer* (Acari, Oribatida). *Journal of Evolutionary Biology*, **20**, 392–402.

Heidemann K, Hennies A, Schakowske J *et al.* (2014) Free-living nematodes as prey for higher trophic levels of forest soil food webs. *Oikos*, 1-13.

Hewitt G (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hewitt GM, Ibrahim KM (2001) Inferring glacial refugia and historical migrations with molecular phylogenies. In: *Integrating Ecology and Evolution in a Spatial Context* (eds Silvertown J, Antonovics J), pp. 271–294. Blackwell Science Ltd., Oxford.

Hirst S, Maulik S (1926) On some Arthropod Remains from the Rhynie Chert (Old Red Sandstone). *Geological Magazine*, **63**, 69–71.

Hopkin SP (1997) *Biology of the Springtails*. Oxford University Press, Oxford.

Hopkin SP (2007) *A Key to the Collembola (Springtails) of Britain and Ireland*. Field Studies Council.

Ingimarsdóttir M, Caruso T, Ripa J *et al.* (2012) Primary assembly of soil communities: disentangling the effect of dispersal and local environment. *Oecologia*, **170**, 745–54.

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

Konestabo HS, Michelsen A, Holmstrup M (2007) Responses of springtail and mite populations to prolonged periods of soil freeze-thaw cycles in a sub-arctic ecosystem. *Applied Soil Ecology*, **36**, 136–146.

Langeneckert A (2013) Lipid composition of Collembola from different forest types (beech, spruce) as indicated by fatty acid analyses.

Lawrence P (1985) Ten species of Collembola from Baltic amber. *Prace Muzeum Ziemi*, **37**, 101–104.

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

Maraun M, Martens H, Migge S *et al.* (2003) Adding to "the enigma of soil animal diversity": fungal feeders and saprophagous soil invertebrates prefer similar food substrates. *European Journal of Soil Biology*, **39**, 85–95.

Milne I, Lindner D, Bayer M *et al.* (2009) TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics*, **25**, 126–127.

Mosbrugger V, Utescher T, Dilcher DL (2005) Cenozoic continental climatic evolution of Central Europe. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 14964–9.

Ohlsson L, Verhoef H (1988) Effects of diet composition on cold adaptation in temperate Collembola. *Comparative Biochemistry and Physiology*, **91**, 475-479.

Petersen H (2002) General aspects of collembolan ecology at the turn of the millennium. *Pedobiologia*, **46**, 246–260.

Petersen H, Luxton M (1982) A Comparative Analysis of Soil Fauna Populations and Their Role in Decomposition Processes. *Oikos*, **39**, 288.

Ponge JF (1991) Food resources and diets of soil animals in a small area of Scots pine litter. *Geoderma*, **49**, 33–62.

Porco D, Bedos A, Greenslade P (2012) Challenging species delimitation in Collembola: cryptic diversity among common springtails unveiled by DNA barcoding. *Invertebrate Systematics*, **26**, 470–477.

Rapoport E (1971) The geographical distribution of Neotropical and Antarctic Collembola. *Pacific Insects Monograph*, **4**, 99–118.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Rosenberger M, Maraun M, Scheu S *et al.* (2013) Pre- and post-glacial diversifications shape genetic complexity of soil-living microarthropod species. *Pedobiologia*, **56**, 79–87.

Salomone N, Emerson BC, Hewitt GM *et al.* (2002) Phylogenetic relationships among the Canary Island Steganacaridae (Acari, Oribatida) inferred from mitochondrial DNA sequence data. *Molecular Ecology*, **11**, 79–89.

Schäffer S, Koblmüller S, Pfingstl T *et al.* (2010) Contrasting mitochondrial DNA diversity estimates in Austrian *Scutovertex minutus* and *S. sculptus* (Acari, Oribatida, Brachypylina, Scutoverticidae). *Pedobiologia*, **53**, 203–211.

Seastedt TR (1984) The Role of Microarthropods in Decomposition and Mineralization Processes. *Annual Review of Entomology*, **29**, 25–46.

Siepel H (1994) Life-history tactics of soil microarthropods. *Biol Fertile Soils*, 263–278.

Stamatakis A, Ott M, Ludwig T (2005) *Parallel Computing Technologies* (V Malyshkin, Ed,). Springer Berlin Heidelberg, Berlin, Heidelberg.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680.

Torricelli G, Carapelli A, Convey P *et al.* (2010) High divergence across the whole mitochondrial genome in the “pan-Antarctic” springtail *Friesea grisea*: Evidence for cryptic species? *Gene*, **449**, 30–40.

Wardle DA, Bardgett RD, Klironomos JN *et al.* (2004) Ecological linkages between aboveground and belowground biota. *Science (New York, N.Y.)*, **304**, 1629–1633.

Waters JM, Fraser CI, Hewitt GM (2013) Founder takes all: density-dependent processes structure biodiversity. *Trends in Ecology & Evolution*, **28**, 78–85.

Webb NR, Block W (1993) Aspects of cold hardiness in *Steganacarus magnus* (Acari: Cryptostigmata). *Experimental & Applied Acarology*, **17**, 741–748.

Weitschat W, Wichard W (2002) Diplurans and Springtails - Insecta: Diplura and Collembola. In: *Atlas of Plants and Animals in Baltic Amber* , pp. 86–87. Dr. Friedrich Pfeil, München.

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

Wolters V (1991) Soil Invertebrates - Effects on Nutrient Turnover and Soil Structure - A Review. *Zeitschrift für Pflanzenernährung und Bodenkunde*, **154**, 389–402.

Worland MR (2005) Factors that influence freezing in the sub-Antarctic springtail *Tullbergia antarctica*. *Journal of insect physiology*, **51**, 881–94.

Van der Wurff AWG, Gols R, Ernsting G *et al.* (2005) Population genetic structure of *Orchesella cincta* (Collembola; Hexapoda) in NW Europe, as revealed by microsatellite markers. *Pedobiologia*, **49**, 167–174.

Van der Wurff AWG, Isaaks JA, Ernsting G *et al.* (2003) Population substructures in the soil invertebrate *Orchesella cincta*, as revealed by microsatellite and TE-AFLP markers. *Molecular Ecology*, **12**, 1349–1359.

Zachos JC, Dickens GR, Zeebe RE (2008) An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature*, **451**, 279–83.

Zawischa D (1993) Einschlüsse im Bernstein. *Arbeitskreis Paläontologie Hannover*, **21**, 1–32.

Zettel J, Zettel U, Egger B (2000) Jumping technique and climbing behaviour of the collembolan *Ceratophysella sigillata* (Collembola: Hypogasturidae). *European Journal of Entomology*, **97**, 41–45.

Data Accessibility

DNA sequences: Genbank accessions KF684371-KF684865.

Supporting information

Table S1. (A) Primer sequence of and (B) PCR conditions for the ribosomal D3-D5 region of 28S rDNA, and the protein coding Histone 3 (*H3*) and *COI* genes.

(A)

gene	**primer name *(organism)***	**sequence (5' - 3')**	**reference**
28S (D3-D5 region)	28Sa	GAC CCG TCT TGA AGC ACG	Tully *et al.* 2006
	28Sbout	CCC ACA GCG CCA GTT CTG CTT ACC	
Histone 3	H3F1 (C*. denticulata, F. quadrioculata*)	ATG GCT CGT ACC AAG CAG ACV GC	Colgan & McLauchlan 1998
	H3R1 (C*. denticulata, F. quadrioculata*)	ATA TCC TTR GGC ATR ATR GTG AC	
	H3F2 (*I. minor*)	ATG GCT CGT ACC AAG CAG AC	
	H3F2 (*I. minor*)	ATR TCC TTG GGC ATG ATT GTT AC	
COI	LCO1490	GGT CAA CAA ATC ATA AAG ATA TTG G	Folmer *et al.* 1994
	HCO2198	TAA ACT TCA GGG TGA CCA AAA AAT CA	

(B)

PCR steps	**28S (D3-D5 region)**	***H3***	***COI***
	temp - time	temp - time	temp - time
1 - initial denaturation step	95°C - 15 min	95°C - 15 min	95°C - 15 min
2 - denaturation	94°C - 15 sec	94°C - 15 sec	94°C - 15 sec
3 - annealing	49°C - 15 sec	59°C - 15 sec (*F. quadrioculata, I. minor*)	45°C - 15 sec
		42°C - 15 sec (*C. denticulata*)	
4 - elongation	72°C - 15 sec	72°C - 15 sec	72°C - 15 sec
5 - final elongation	72°C - 6 min	72°C - 6 min	72°C - 6 min
number of cycles (2-4)	35	35	35

Table S2. DNA sequences available at GenBank**.**

abbre-viation	*C. denticulata*			*F. quadrioculata*			*I. minor*		
	COI	***H3***	**28S**	***COI***	***H3***	**28S**	***COI***	***H3***	**28S**
EE	-	-	-	KF684597-601	KF684762-6	KF684432-6	-	-	-
NO	KF684567-71	KF684732-6	KF684402-6	KF684627-31	KF684792-6	KF684462-66	KF684678-82	KF684843-7	KF684513-7
RU1	-	-	-	KF684632-5	KF684797-800	KF684467-70	KF684687-91	KF684852-6	KF684522-26
RU2	KF684576-80	KF684741-5	KF684411-5	-	-	-	-	-	-
AT1	-	-	-	-	-	-	KF684646-50	KF684811-5	KF684481-85
AT2	KF684536-40	KF684701-5	KF684371-5	-	-	-	-	-	-
AT3	-	-	-	KF684590-2	KF684755-7	KF684425-7	-	-	-
FR	KF684544-8	KF684709-13	KF684379-83	KF684602-6	KF684767-71	KF684437-41	KF684655-8	KF684820-3	KF684490-3
DE	KF684549-53	KF684714-8	KF684384-88	KF684607-11	KF684772-6	KF684442-6	KF684659-62	KF684824-7	KF684494-7
PL	KF684572-5	KF684737-40	KF684407-10	-	-	-	KF684683-6	KF684848-51	KF684518-21
HR	KF684541-3	KF684706-8	KF684376-8	KF684593-6	KF684758-61	KF684428-31	KF684651-4	KF684816-9	KF684486-9
GR	KF684554-8	KF684719-23	KF684389-93	KF684612-6	KF684777-81	KF684447-51	KF684663-7	KF684828-32	KF684498-502
IT1	-	-	-	KF684617-21	KF684782-6	KF684452-6	KF684668-72	KF684833-7	KF684503-7
IT2	KF684559-61	KF684724-6	KF684394-96	-	-	-	-	-	-
ME	KF684562-6	KF684727-31	KF684397-401	KF684622-6	KF684787-91	KF684457-61	KF684673-7	KF684838-42	KF684508-12
RS	KF684581-5	KF684746-50	KF684416-20	KF684636-40	KF684801-5	KF684471-75	-	-	-
ES1	-	-	-	-	-	-	KF684695-99	KF684860-64	KF684531-5
ES2	KF684586-9	KF684751-4	KF684421-4	-	-	-	KF684692-4, 700	KF684857-9, 65	KF684527-30
ES3	-	-	-	KF684641-5	KF684806-10	KF684476-80	-	-	-

Table S3. Summary of additional locations and outgroup taxa included in the extended datasets (1) and (3) of the molecular clock analyses. Accession numbers from NCBI and BOLD databanks are listed. For some isotomid species, the complete D3 fragment of 28S was not available, missing positions were filled with "?" in the alignments.

species	location	NCBI accession numbers	
		COI	**28S**
Hypogastruridae			
Ceratophysella denticulata	Antarctica	HQ732034.1	HQ731919.1
		HQ732033.1	HQ731918.1
		HQ732032.1	HQ731917.1
		HQ732031.1	HQ731915.1
	Australia	GU656655	-
		GU656653	-
		GU656652	-
	Canada	KF642457	-
		KF642321	-
		KF642319	-
		KF642278	-
		KF642149	-
	Chile	HQ732035.1	HQ731920.1
	New Zealand	HQ732038.1	HQ731923.1
		HQ732037.1	HQ731922.1
		HQ732036.1	HQ731921.1
	South Africa	JX261832	-
		JX261831	-
		JX261793	-
		JX261791	-
		JX261790	-
congeneric species			
Ceratophysella gibbosa	Ireland	HQ732029.1	HQ731913.1
	Argentina	HQ732040.1	HQ731925.1
outgroups			
Hypogastrura viatica	New Zealand	HQ732070.1	HQ731954.1
Hypogastrura vernalis	France	HQ732065.1	HQ731949.1
Hypogastrura pubescens	New Zealand	HQ732059.1	HQ731943.1
Hypogastrura subboldorii	France	HQ732064.1	HQ731948.1
Xenylla grisea	Antarctica	HQ732084.1	HQ731975.1
Gomphiocephalus hodgsoni	Antarctica	HQ732052.1	HQ731936.1

Table S3 continued

Isotomidae			
Folsomia quadrioculata	Canada (L1)	KF642009	-
		KF642087	-
		KF642519	-
		KF642580	-
	Canada (L2)	KF641961	-
		KF642017	-
		KF642370	-
		KF642443	-
Isotomiella minor	France (Haute Normandie)	(BOLD accession numbers)	
		GENHP910-12	-
		GENHP911-12	-
		GENHP941-12	-
		GENHP942-12	-
		GENHP944-12	-
congeneric species			
Folsomia candida		KF592092	DQ279734
		HQ732049	HQ731934
Folsomia fimetaria		LK024463	LK024320
Folsomia octoculata		HM366595	-
outgroups			
Parisotoma notabilis	Denmark (DK_Pn1 DK_Pn1)	KJ792235	KJ792115
	Turkey (TR_Pn1)	KJ792335	KJ792215
	Germany (DE3_Pn1, DE3_Pn1)	KJ792272	KJ792152
	Greece (GR_Pn1)	KJ792282	KJ792162
	Croatia (HR_Pn5)	KJ792234	KJ792114
Anurophorus septentrionalis		LK024447	LK024297
Cryptopygus antarcticus travei		HQ592681	HQ592739
Isotoma riparia		LK024475	LK024340
		LK024474	LK024339
		LK024473	LK024338
		HG422621	LK024337
Isotoma viridis		HG422615	LK024346
Isotomurus maculatus		HQ592702	HQ592760
Isotoma marionensis		HQ592673	HQ592731
		HQ592672	HQ592730
		HQ592670	HQ592729
		HQ592671	HQ592727
		HQ592669	HQ592728

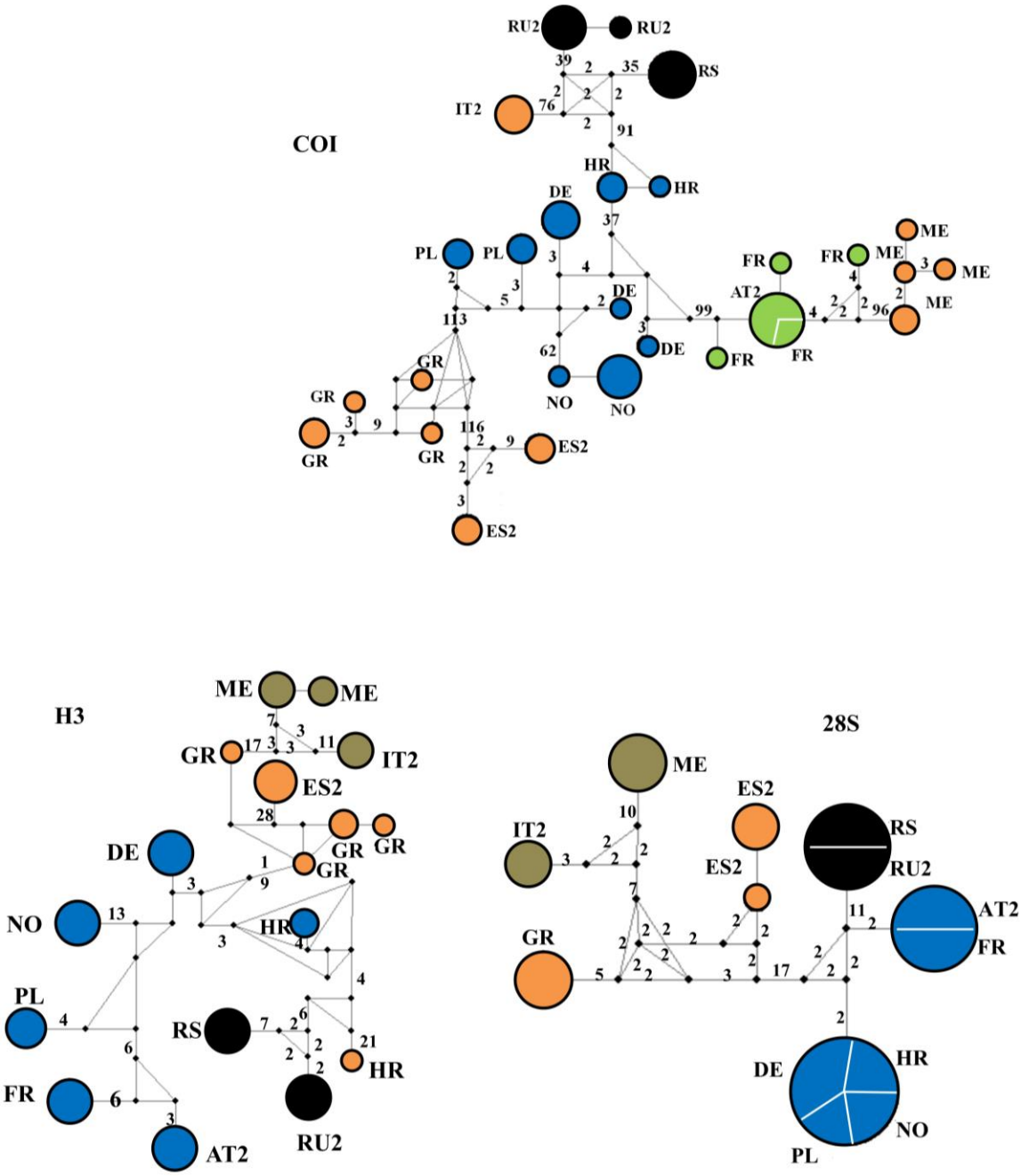


Figure S1. Median-joining networks of *Ceratophysella denticulata* of *COI* (709 bp), *H3* (374 bp) and the D3-D5 region of 28S rDNA (570 bp). Numbers on lines represent hypothetical mutational steps between haplotypes, lines without numbers indicate a single mutation step. For abbreviations of sampling locations see Table 1, colors of locations correspond to the genetic lineages from Fig. 1 and Figs. S4A and S5A.

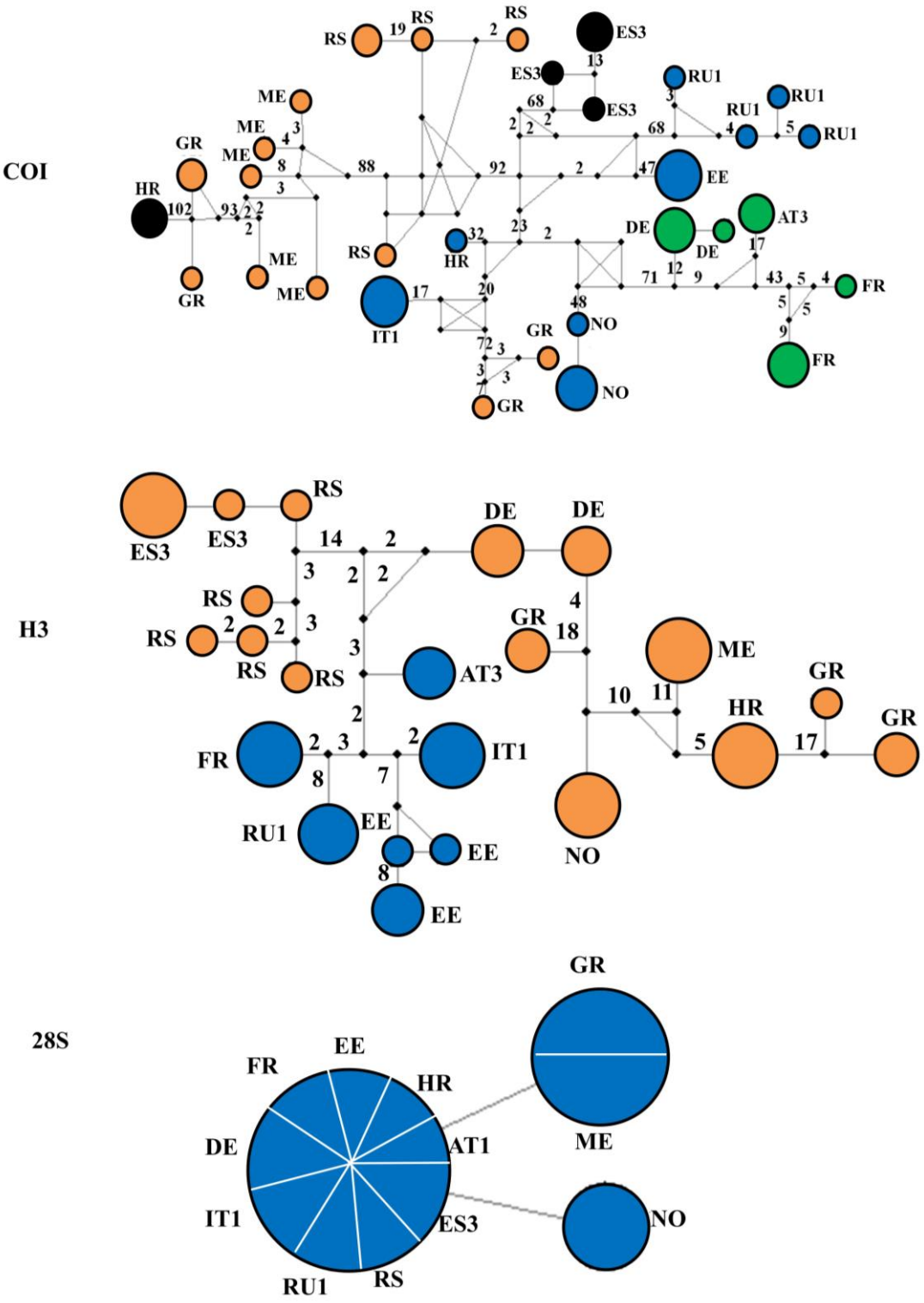


Figure S2. Median-joining networks of *Folsomia quadrioculata* of *COI* (709 bp), *H3* (374 bp) and the D3-D5 region of 28S rDNA (570 bp). Numbers on lines represent hypothetical mutational steps between haplotypes, lines without numbers indicate a single mutation step. For abbreviations of sampling locations see Table 1, colors of locations correspond to the genetic lineages from Fig. 2 and Figs. S4B and S5B.

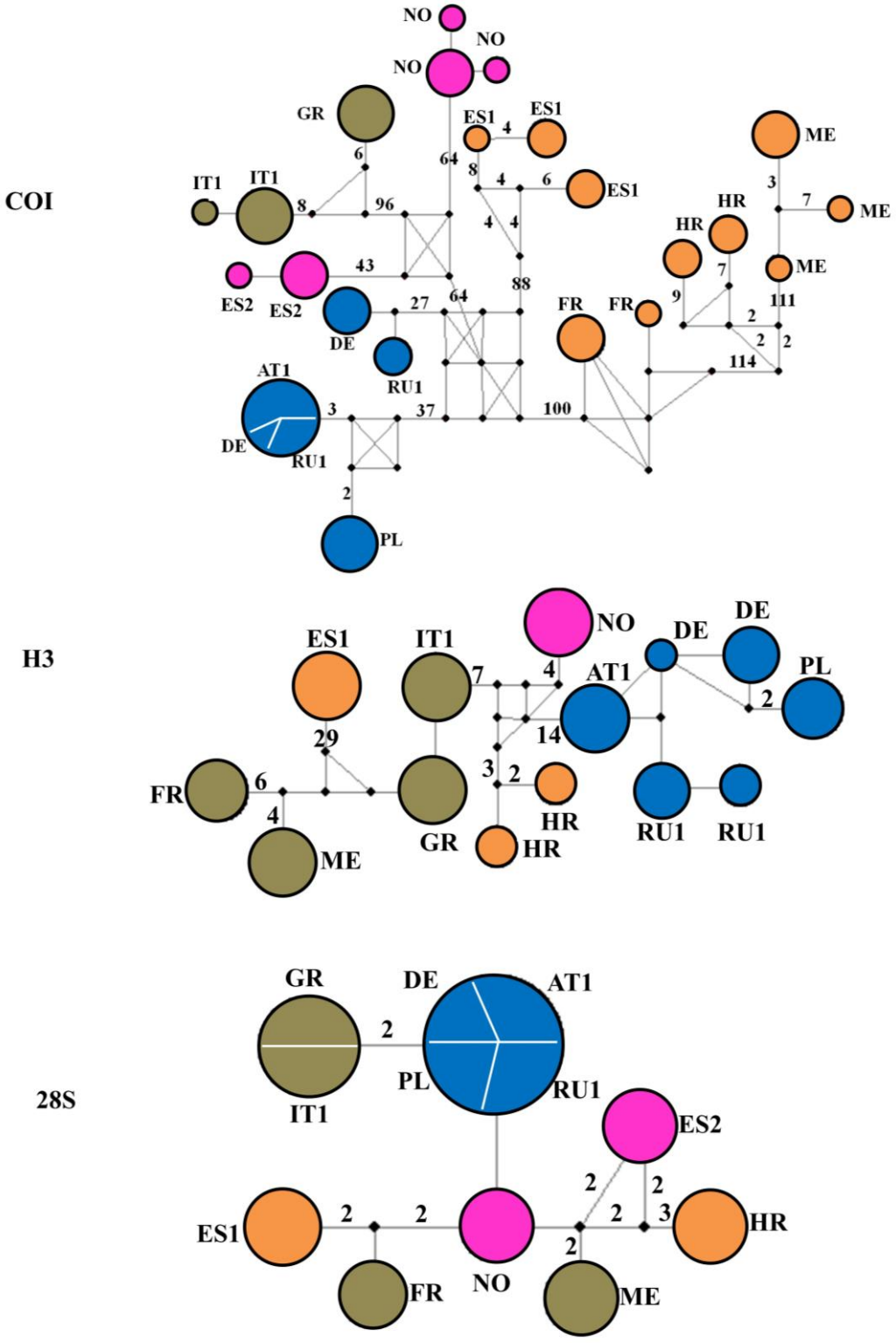


Figure S3. Median-joining networks of *Isotomiella minor* of *COI* (709 bp), *H3* (374 bp) and the D3-D5 region of 28S rDNA (570 bp). Numbers on lines represent hypothetical mutational steps between haplotypes, lines without numbers indicate a single mutation step. For abbreviations of sampling locations see Table 1, colors of locations correspond to the genetic lineages from Fig. 3 and Figs. S4C and S5C.

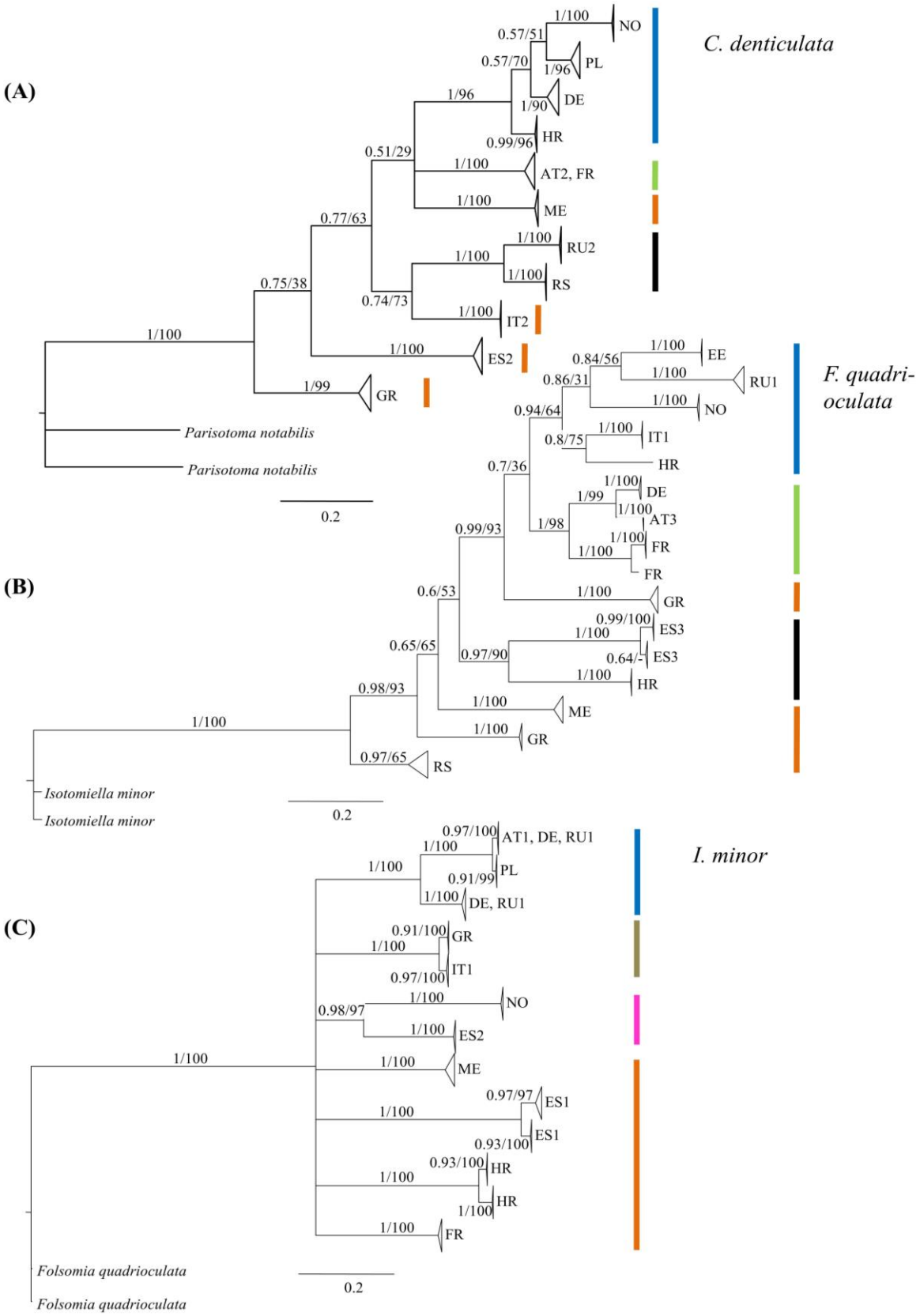


Figure S4. Bayesian analysis of nucleotide sequence data of *COI* of (A) *Ceratophysella denticulata*, (B) *Folsomia quadrioculata* and (C) *Isotomiella minor* in Europe and assignment of geographic lineages as indicated by colored lines. Numbers on nodes are posterior probabilities (Bayesian Inference) and bootstrap values (Maximum Likelihood).

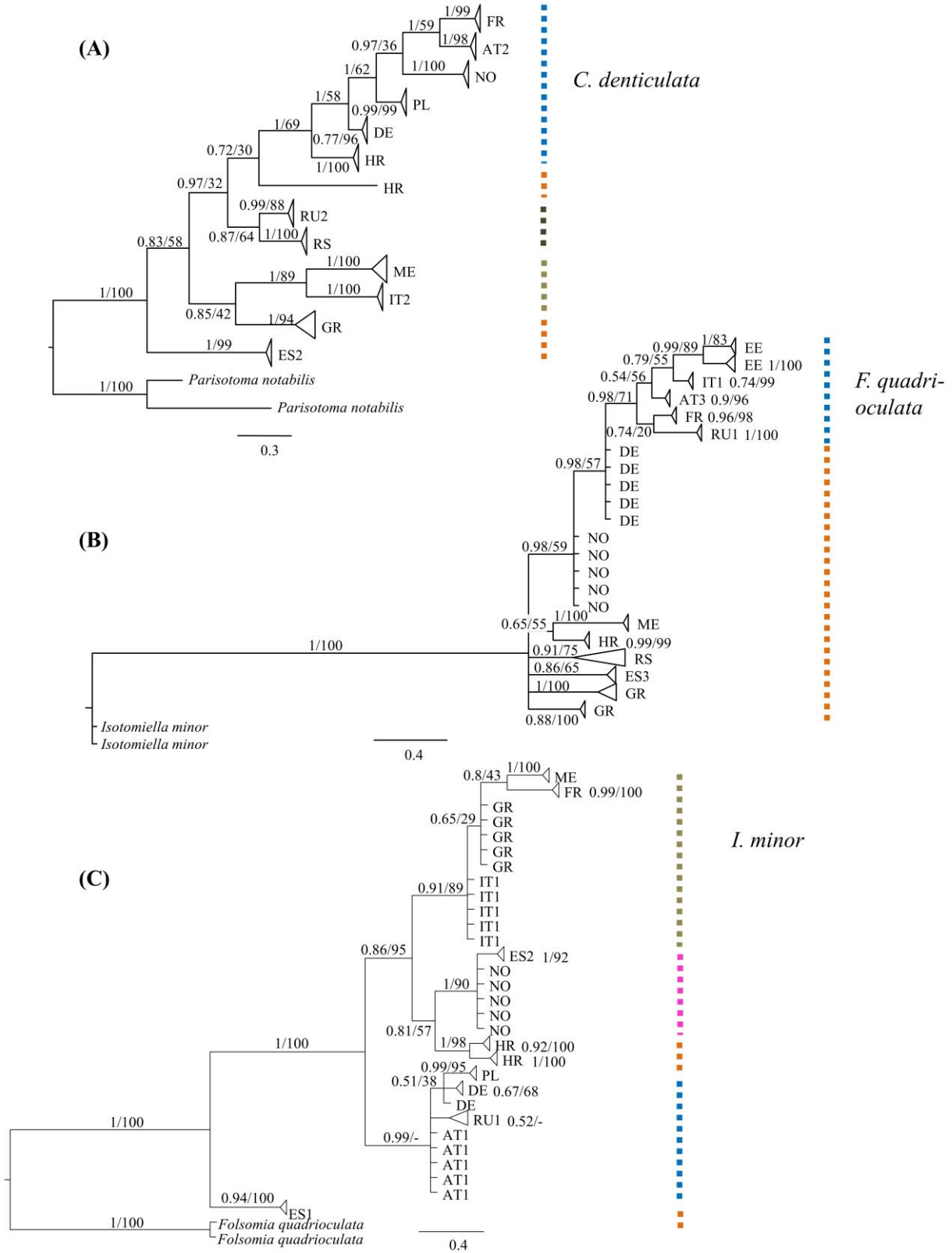


Figure S5. Bayesian analysis of nucleotide sequence data of *H3* of (A) *Ceratophysella denticulata*, (B) *Folsomia quadrioculata* and (C) *Isotomiella minor* in Europe and assignment of geographic lineages as indicated by colored lines. Numbers on nodes are posterior probabilities (Bayesian Inference) and bootstrap values (Maximum Likelihood).

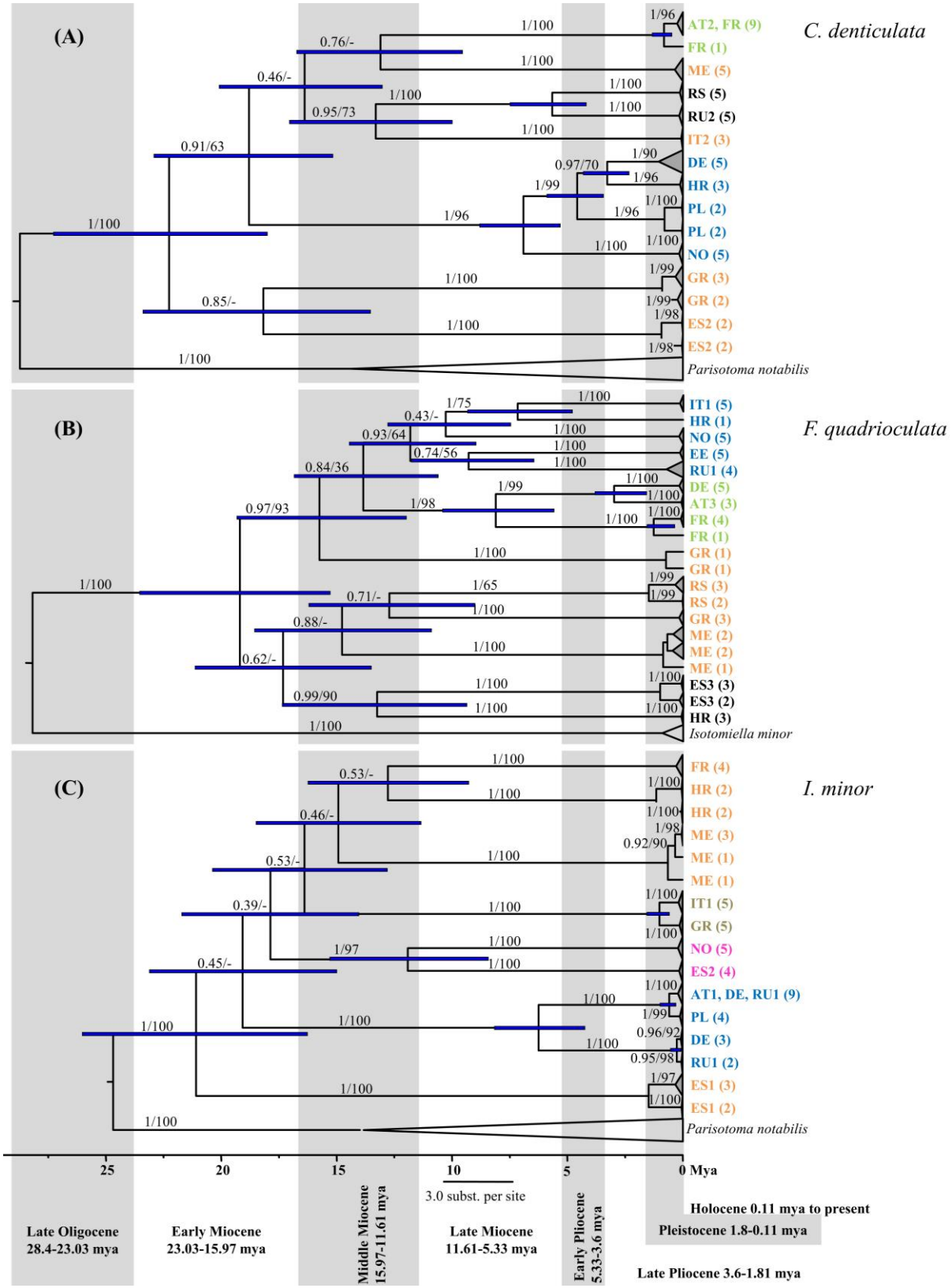


Figure S6. Molecular divergence estimates of dataset (2) for (A) *Ceratophysella denticulata*, (B) *Folsomia quadrioculata* and (C) *Isotomiella minor* in Europe, based on a strict clock and the *COI* sequences generated in this study with a single outgroup taxon.

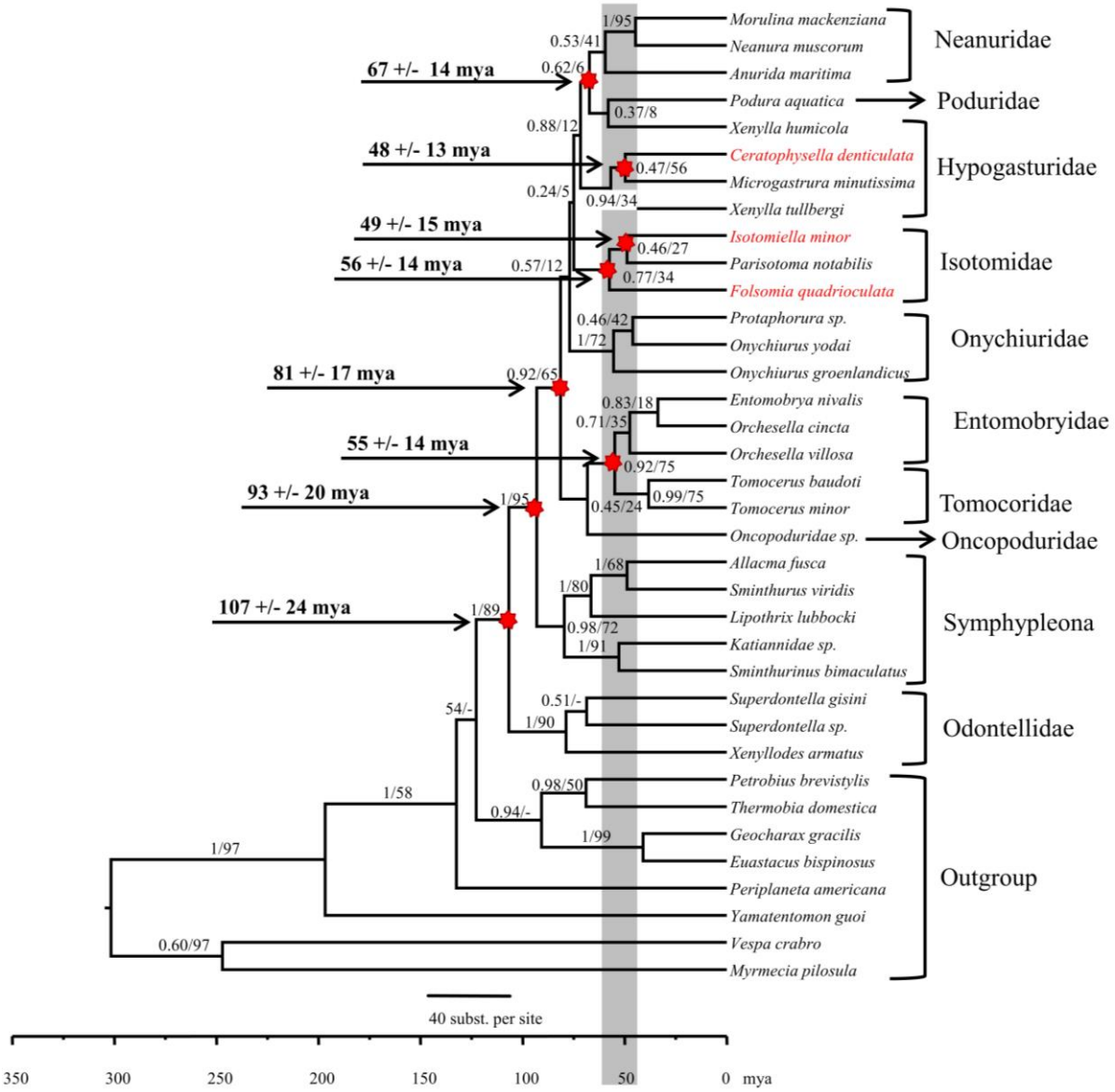


Figure S7. The *COI* tree for age estimation included all major taxa of Collembola and recovered all families according to the phylogeny of D'Haese (2002). Species of this study are marked in red. Numbers on arrows indicate divergence times of nodes highlighted with a red star. The timeline below the tree indicates divergence times using a strict molecular clock with a rate of 0.0115, i.e. the standard arthropod substitution rate for *COI* of 2.3 percent divergence per million years.

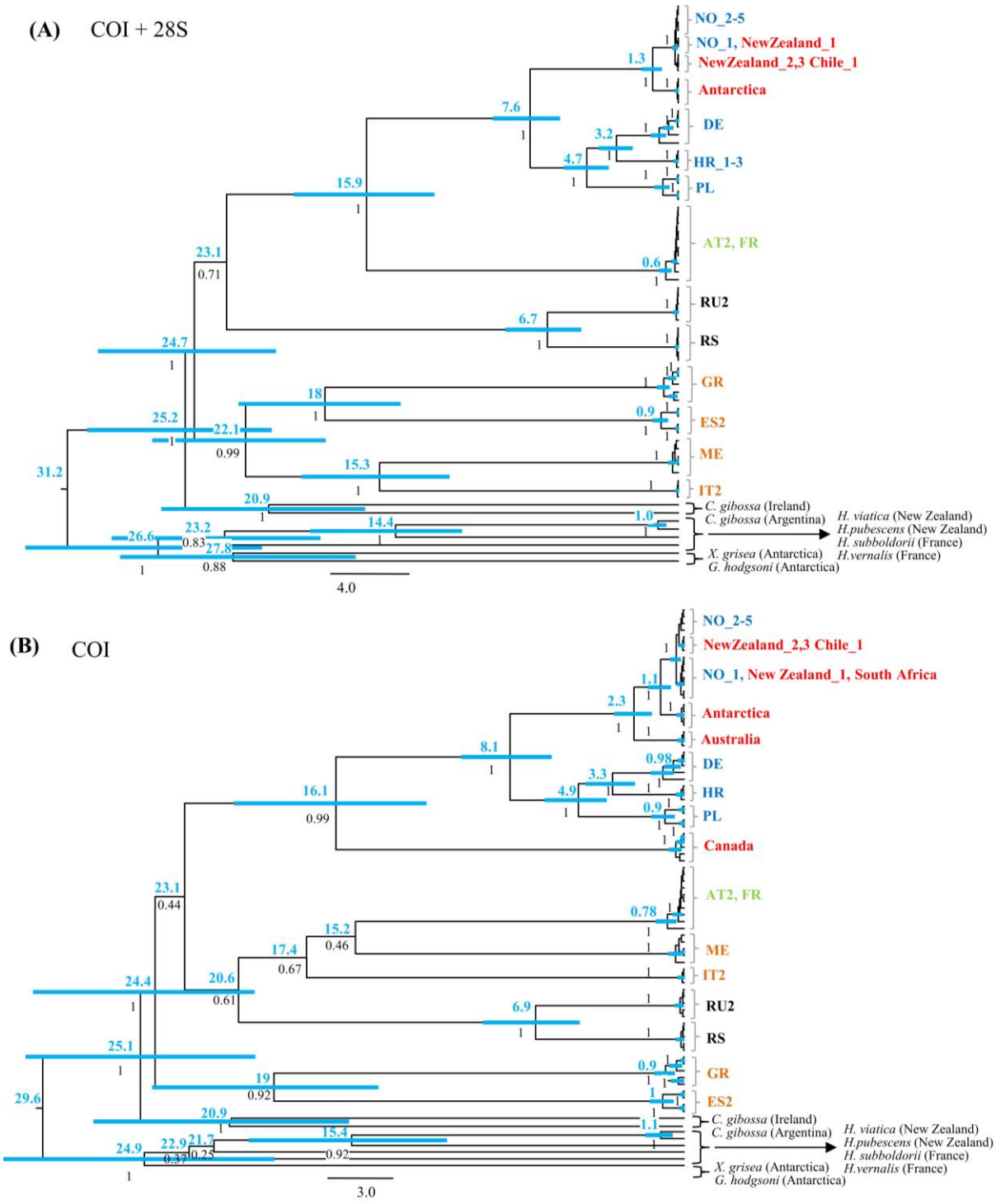


Figure S8. Molecular divergence estimates for *Ceratophysella denticulata* for (A) dataset (1), a combined analysis with two genes, applying a fixed rate for *COI* and an estimated rate for 28S and (B) dataset (3), using only *COI* with a fixed rate. Both datasets include additional sequences of *C. denticulata* from non-European locations and extended the number of outgroup taxa.

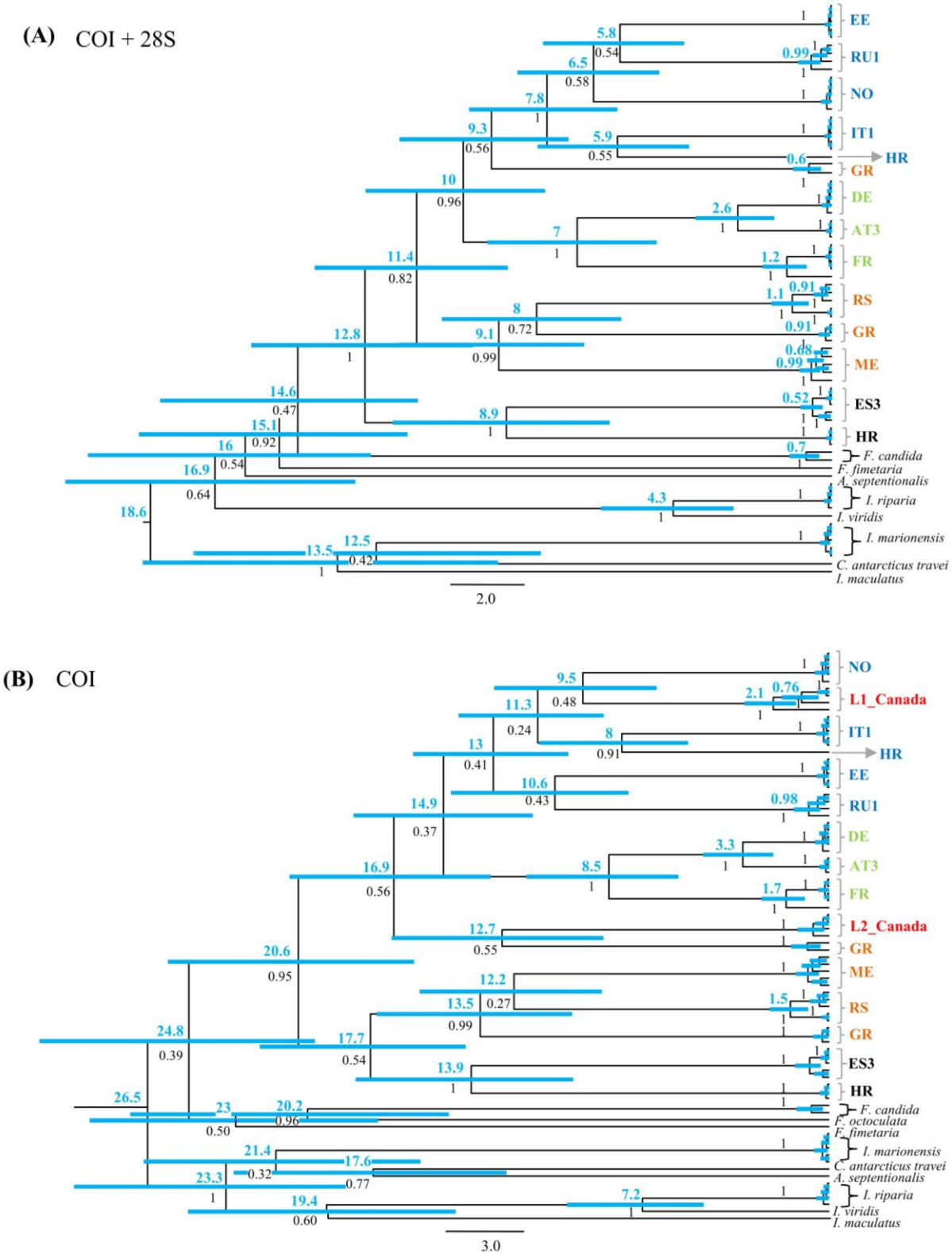


Figure S9. Molecular divergence estimates for *Folsomia quadrioculata* for (A) dataset (1), a combined analysis with two genes, applying a fixed rate for *COI* and an estimated rate for 28S and (B) dataset (3), using only *COI* with a fixed rate. Dataset (1) includes only species of *F. quadrioculata* from this study and an extended outgroup sampling, as no sequences of the D3 fragment were available at the NCBI databank. Dataset (3) includes additional sequences of *F. quadrioculata* from Canada and extended the number of outgroup taxa.

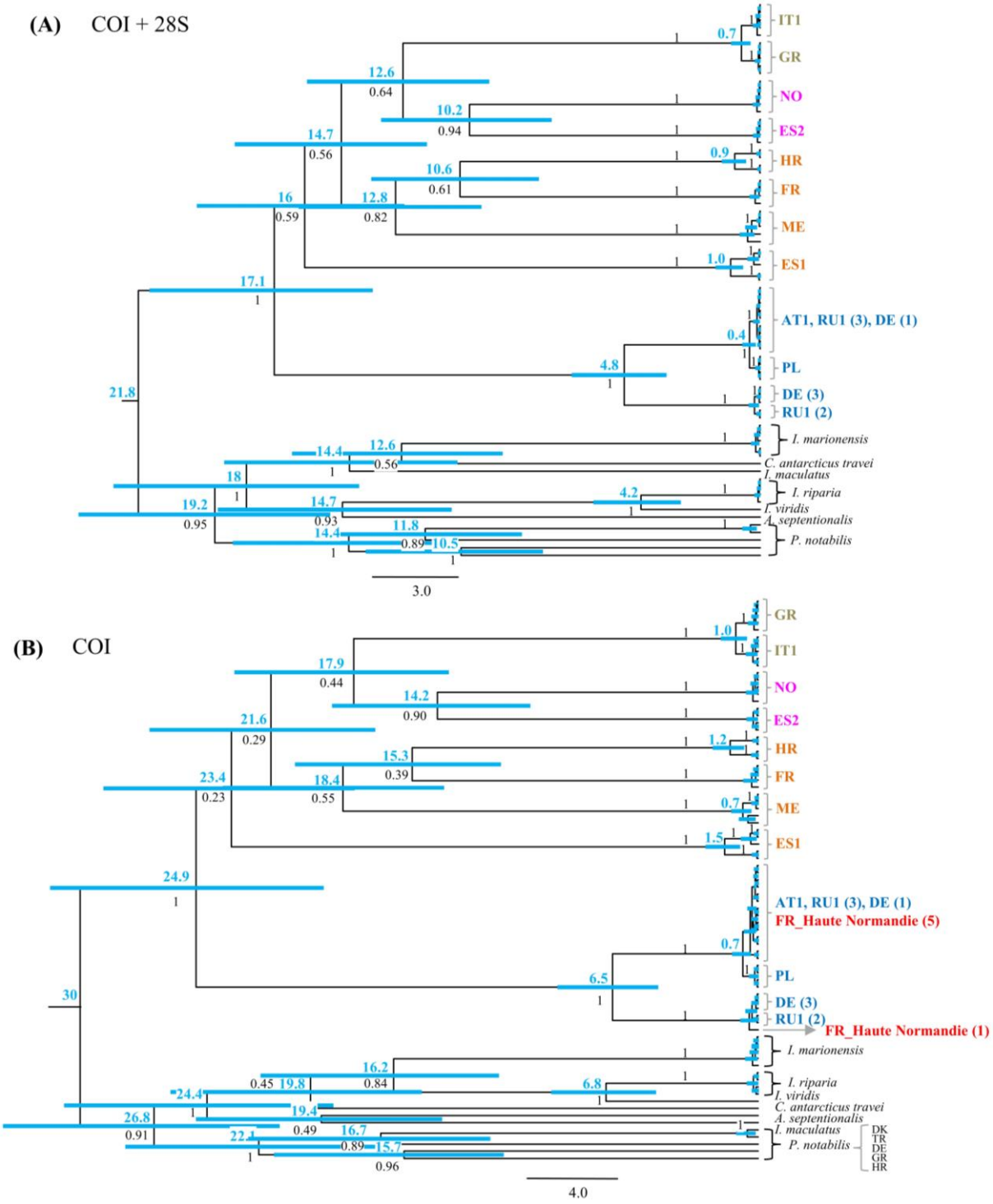
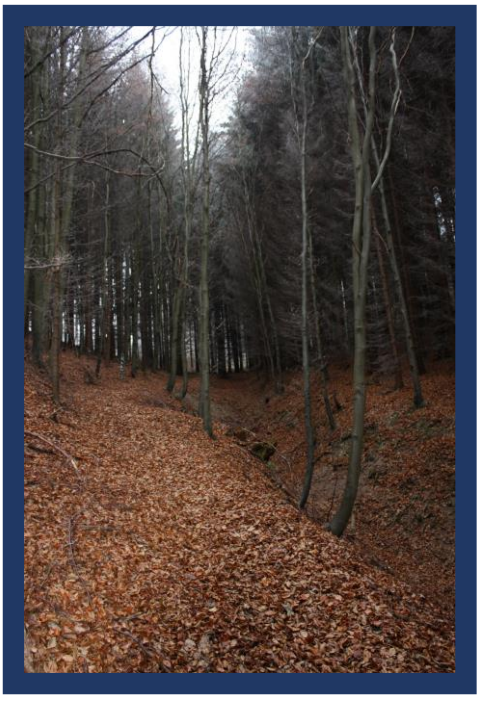


Figure S10. Molecular divergence estimates for *Isotomiella minor* for (A) dataset (1), a combined analysis with two genes, applying a fixed rate for *COI* and an estimated rate for 28S and (B) dataset (3), using only *COI* with a fixed rate. Dataset (1) includes only species of *I. minor* from this study and an extended outgroup sampling, as no sequences of the D3 fragment were available at the NCBI databank. Dataset (3) includes additional sequences of *I. minor* from northern France (Haute Normandie). The outgroup taxa were the same as for *F. quadrioculata* but adding 5 sequences of the parthenogenetic isotomid species *Parisotoma notabilis*.

Chapter 3

GENETIC STRUCTURE AND COLONIZATION PATTERNS OF SEXUAL AND PARTHENOGENETIC SPRINGTAILS (HEXAPODA, COLLEMBOLA) IN EUROPE

Helge von Saltzwedel, Stefan Scheu and Ina Schaefer



submitted

Abstract

Wide areas of central and northern Europe were virtually devoid of animals and plants during Quaternary glaciations, but were colonized during interglacial periods. Parthenogenetic reproduction provides a substantial colonization advantage compared to sexual reproduction. To investigate the significance of reproductive modes for colonization, we compared the genetic structure of one sexual and one obligate parthenogenetic springtail species with similar ecology across Europe, using mitochondrial and nuclear markers. Molecular variance was similar in both species with deep genetic divergences among populations, suggesting persistence of old lineages. Molecular divergence estimates and range expansions of genetic lineages indicated both survival in higher latitudes and recent anthropogenic distributions. However, compared to the parthenogenetic species genetic structure was considerably more complex in the sexual species and genetic variance was dominated by synonymous substitutions. In addition to founder effects and competition, maintenance of mitonuclear compatibility among mating partners and post-zygotic selection on interpopulation hybrids likely promoted genetic structure in the sexual species. In contrast, gene-environment interactions presumably were of greater importance in the parthenogenetic species, in which mitonuclear interactions are linked for generations. The results suggest that including mitonuclear and gene-environment interactions will add a new dimension to phylogeography and cryptic species concepts.

Key words: reproductive mode, phylogeography, Miocene, mitonuclear, gene-environment

Introduction

Parthenogenetic reproduction is rare in nature (Williams, 1975; Maynard Smith, 1978; Barton & Charlesworth, 1998), but comparatively common in the soil-system. Many taxa living in soil constitute of all-female populations, including oribatid mites, tardigrades, nematodes and springtails (Bell, 1982). Parthenogenesis potentially is advantageous in structured habitats such as soil. Females do not require searching for mating partners and populations can grow rapidly when resources are available. However, in the long-term populations likely become genetically homogeneous, which reduces their adaptive potential to changing environments or in using new resources (Rice, 2002; Song *et al.*, 2011). In contrast, genetic recombination in sexual species ensures genetic variance that allows purging deleterious mutations, combining beneficial mutations and facilitates access to a wider range of resources (Williams, 1975; Hamilton, 1980; Omilian *et al.*, 2006; Latta *et al.*, 2013). Collembola are abundant and omnipresent microarthropods in soil all over the world. Parthenogenetic reproduction is common in springtails and occurs in all phylogenetic lineages. It is common in soil living species such as Onychiuridae and Isotomidae with ten and seven percent of all species being parthenogenetic, respectively (Chahartaghi *et al.*, 2006; Chernova *et al.*, 2010). Collembola are early colonizers of disturbed habitats and disperse over long distances by wind (Dunger *et al.*, 2002; Hawes *et al.*, 2007). Both, sexual and parthenogenetic species are fast reproducing (Gregoire-Wibo, 1977; Hopkin, 1997) and coexist locally (Chahartaghi *et al.*, 2006; Chernova *et al.*, 2010). Similarities in dispersal ability and habitat preferences of sexual and parthenogenetic species make Collembola a good model system for comparative analyses on advantages and disadvantages of obligate parthenogenesis. In a stable habitat such as soil, the reproductive assurance of parthenogenetic species facilitates colonization of new habitats, while sexual reproduction enables access to a variety of habitats and resources and populations evade mutational meltdown (Williams, 1975; Bell, 1982; Gabriel *et al.*, 1993; Scheu & Drossel, 2007). Colonization success depends on several factors, such as resource quality and availability, mobility, fertility and habitat characteristics (Debouzie *et al.*, 2002). In soil detrital resources are uniformly distributed and omnipresent. Further, the buffering capacity of soil against unfavourable abiotic conditions, such as low temperatures or drought, provides more stable habitat

conditions than above the ground. Consequently, Collembola may be viewed as living in a stable habitat with omnipresent resources. Therefore, colonization success and distribution patterns likely are driven predominantly by species specific traits, such as reproductive mode and mobility, rather than by environmental filtering according to local habitat conditions.

We investigated the genetic structure of two species of Collembola, the sexual *Folsomia quadrioculata* (Tullberg, 1871) and the parthenogenetic *Isotomiella minor* (Schaeffer, 1896) across Europe. The two species coexist locally, being widely distributed and common in the northern hemisphere. Both are primary decomposers feeding on similar resources (Fiera, 2013) and predominantly live in the litter layer and below the soil surface but differ in reproductive modes (Ponge, 1991; Langeneckert, 2013).

During the Quaternary ice-ages northern and central Europe was to a great extent devoid of animals as glaciers and permafrost covered wide areas. Consequently, present day Collembola species recolonized these areas in similar time spans, but the speed of recolonization likely differed between species of different reproductive modes. Reproductive assurance in the parthenogenetic species likely facilitated faster recolonization of central and northern Europe from glacial refugia compared to the sexual species. Accordingly, we expected only few genetic lineages that expanded rapidly in open habitats of central and northern Europe after glaciers retreated. This pattern of lower genetic diversity will be more pronounced in the parthenogenetic species *I. minor* than in the sexual species *F. quadrioculata* due to the colonization advantage of parthenogenetic individuals. Accordingly, genetic structure across Europe will be low for Collembola and lower in the parthenogenetic compared to the sexual species. In contrast, genetic variance within sampling locations will be higher in the sexual than in the parthenogenetic species, because recombination likely enables sexual species to exploit a wider range of resources, permitting the coexistence of different genetic lineages by avoiding niche-overlap. Further, reproductive assurance enabled faster and therefore earlier recolonization of central and northern European habitats, resulting in older populations of the parthenogenetic species than those of the sexual species. To investigate these effects of long-term colonization of Europe we sampled at large geographic scale and sequenced one mitochondrial (*COI*, cytochrome c oxidase subunit I) and two nuclear genes (28S rDNA, D3-D5 region, and *Histone H3*) as genetic markers from

37 populations. The combination of three markers with different substitution rates allowed inferences on recolonization patterns and genetic structure at different time scales and increased the resolution of genetic structure across a wide geographic region.

Materials and Methods

Data collection

About one square meter of leaf litter in forests was collected at 37 locations in Europe, including northwest Russia (Karelia), Greenland, the Pleistocene refuge areas Italy, Spain, the Balkan (Montenegro, Serbia, Croatia, Greece) and Ukraine, and transferred to the University of Göttingen (**Table 1**). Animals were extracted by heat according to Kempson *et al.* (Kempson *et al.*, 1963), collected in 96% EtOH and stored at -20°C until further analyses. For species identification specimens were separated using a dissecting microscope and determined by light microscopy after Hopkin (Hopkin, 2007). Genomic DNA was extracted from single individuals of *F. quadrioculata* (n=143) and *I. minor* (n=104) using the DNeasy Blood and Tissue Kit (Qiagen; Hilden, Germany) following the manufacturer's protocol for animal tissue. Purified DNA was eluted in 30 µl buffer AE and stored at -20°C until further preparation. Two nuclear genes (*H3*, ~374 bp; D3-D5 region of 28S rDNA, ~570 bp) and the barcoding fragment of the mitochondrial *COI* gene (709 bp) were amplified in 25 µl volumes containing 12.5 µl SuperHot Taq Mastermix (Genaxxon Bioscience GmbH; Ulm, Germany) with 1.5 µl of each primer (10 pM), 4.5 µl H_2O (HPLC gradient grade, Roth; Karlsruhe), 2 µl $MgCl_2$ (25 mM) and 3 µl template DNA using the primers H3F1 5'-ATG GCT CGT ACC AAG CAG ACV GC-3' and H3R1 5'-ATA TCC TTR GGC ATR ATR GTG AC-3' for *F. quadrioculata* and the primers H3F2 5'-ATG GCT CGT ACC AAG CAG AC-3' and H3R2 5'-ATR TCC TTG GGC ATG ATT GTT AC-3' for *I. minor* (Colgan *et al.*, 1998); 28Sa 5'-GAC CCG TCT TGA AGC ACG-3' and 28Sbout 5'-CCC ACA GCG CCA GTT CTG CTT ACC-3' (Tully *et al.*, 2006) and LCO1490 5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3' and HCO2198 5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3' (Folmer *et al.*, 1994). The PCR conditions included one initial activation step at 95°C for 15 min, followed by 35 amplification cycles of denaturation at 94°C for 15 s, annealing at 45°C for *COI* (49°C for 28S, 59°C for *H3*) for 15 s, elongation at 72°C for 15 s and a final elongation step at 72°C for 6 min. Positive PCR products were purified with the QIAquick PCR Purification Kit

(Qiagen; Hilden, Germany) following the manufacturer's protocol and sent for direct sequencing to the Göttingen Genome Laboratory (Institute for Microbiology and Genetics, Georg August University Göttingen). All sequences are available at GenBank (**Table S1**, **Supporting information**).

For DNA extraction animals were destroyed but secondary vouchers (same morphological species, but different individuals sampled from the same population) were kept in our collections at J.F. Blumenbach Institute of Zoology and Anthropology, Georg August University Göttingen).

Table 1. Summary of sampling sites and their abbreviations of two species of Collembola (*Folsomia quadrioculata* and *Isotomiella minor*) across Europe. Number of individuals sequenced (n), haplotype diversity per sampling location of *COI* (H_d) and geographic coordinates are listed.

	Sampling sites		*F. quadrioculata*		*I. minor*		Coordinates
			n	H_d (*COI*)	n	H_d (*COI*)	(N E)
north	**Denmark**	DK1	5	0.4 ± 0.2			55.98° 12.53°
		DK2			4	0.0 ± 0.0	55.69° 12.58°
	Estonia	EE	5	0.0 ± 0.0			59.44° 24.69°
	Greenland	GL	5	0.4 ± 0.2	5	0.4 ± 0.2	64.16° 51.52°
	Great Britain	GB1	4	1.0 ± 0.2			50.80° 1.39°
		GB2	4	0.0 ± 0.0	4	0.7 ± 0.2	55.62° 5.26°
	Norway	NO1	5	0.4 ± 0.2	5	0.7 ± 0.2	59.07° 10.23°
		NO2	3	1.0 ± 0.3			58.99° 9.94°
		NO3			5	0.0 ± 0.0	59.49° 9.98°
	Russia	RU1			5	0.4 ± 0.2	61.77° 34.22°
		RU2	4	1.0 ± 0.2	5	0.6 ± 0.2	64.27° 34.44°
	Sweden	SE	5	0.0 ± 0.0			57.68° 11.96°
central	**Austria**	AT1			5	0.0 ± 0.0	47.46° 12.24°
		AT2	3	0.0 ± 0.0			47.29° 10.33°
	Czech Republic	CZ1	3	0.7 ± 0.3			50.54° 13.99°
		CZ2	5	0.6 ± 0.2	4	0.8 ± 0.2	49.41° 15.64°
	France	FR1	5	0.4 ± 0.2			48.02° 7.27°
		FR2	5	0.4 ± 0.2	4	0.5 ± 0.3	45.42° 5.81°
		FR3	5	0.7 ± 0.2	5	0.4 ± 0.2	48.62° 1.86°
	Germany	DE1	5	0.6 ± 0.2			53.03° 14.27°
		DE2	5	0.4 ± 0.2	4	0.5 ± 0.3	53.58° 7.24°
		DE3	5	0.4 ± 0.2	4	0.5 ± 0.3	51.71° 9.64°
	Hungary	HU	5	0.4 ± 0.2			47.6° 18.41°
	Poland	PL			4	0.0 ± 0.0	52.33° 20.76°
	Slovakia	SK	4	0.8 ± 0.2			48.26° 17.09°
south	**Croatia**	HR	4	0.5 ± 0.3	4	0.7 ± 0.2	45.90° 15.95°
	Greece	GR	5	0.9 ± 0.2	5	0.0 ± 0.0	37.56° 22.20°
	Italy	IT1	5	0.0 ± 0.0	5	0.4 ± 0.2	40.37° 15.22°
		IT2	4	0.0 ± 0.0			46.67° 11.08°
	Macedonia	MK	5	0.4 ± 0.2			41.15° 20.62°
	Montenegro	ME	5	1.0 ± 0.1	5	0.7 ± 0.2	42.13° 19.09°
	Serbia	RS1	5	0.9 ± 0.2			44.22° 21.09°
		RS2	5	0.0 ± 0.0			43.03° 22.69°
	Slovenia	SI1	5	0.4 ± 0.2	5	0.0 ± 0.0	45.49° 14.28°
		SI2			3	0.0 ± 0.0	45.76° 14.21°
	Spain	ES1			5	0.8 ± 0.2	43.36° 6.00°
		ES2			4	0.5 ± 0.3	43.19° 5.16
		ES3	5	0.7 ± 0.2	5	0.0 ± 0.0	43.20° 2.90
	Ukraine	UA	5	0.9 ± 0.2			45.02° 35.07°
	total number of individuals		143		104		

Phylogenetic analyses, divergence time estimation and population structure

Sequences were edited in Sequencher 4.10 (Gene Codes Corporation; USA), ambiguous positions were corrected by hand aided by the chromatograms and nucleotide sequences were translated into amino acid sequences using the invertebrate mitochondrial code implemented in Sequencher. Consensus sequences were assembled in BioEdit 7.0.1 (Hall, 1999), nucleotide (28S) and protein sequences (*COI* and *H3*) were aligned with ClustalW v1.81 (Thompson & Gibson, 1997) separately and in a combined matrix concatenating nucleotide sequences of the three genes.

The best fit model of sequence evolution for each alignment was inferred according to the hLRT in TOPALi v2.5 (Milne *et al.*, 2009) using the PHYML algorithm. Phylogenetic trees were calculated with RAxML v7.0.3 (Stamatakis *et al.*, 2005) and MrBayes v3.1.2 (Ronquist & Huelsenbeck, 2003). The model of sequence evolution for maximum likelihood analysis was set to GTR+I+G for all alignments (28S, *COI*, *H3* and the combined matrix) with 10,000 bootstrap replicates. For Bayesian inference lset parameters for all alignments were nst=6 rates=invgamma and the MCMC chain was run for ten million generations sampling every 1,000th generation and applying a burnin of 2,500 generations. A strict molecular clock for the *COI* nucleotide alignments was applied in BEAST v1.7.4 (Drummond *et al.*, 2012) using the Yule Process as tree prior (Gernhard *et al.*, 2008) and a substitution rate of 0.0115, consistent with the general arthropod substitution rate for *COI* of $2.3*10^{-2}$ substitutions per site per million years (Avise, 1994; Brower, 1994). Convergence of the MCMC after 600 million generations (sampled every $60{,}000^{th}$ generation) and a burnin of 2,500 generations was confirmed using Tracer v1.4 (Drummond & Rambaut, 2007). Median-joining haplotype networks for the nucleotide datasets were generated for both species using the program Network 4.6 (Fluxus Technology, Suffolk, Great Britain). Molecular variance (AMOVA) within and between populations and isolation by distance (Mantel test) were analysed separately with all three genes (uncorrected p-distances) in ARLEQUIN v3.5 (Excoffier & Lischer, 2010) with 20,000 permutations. To check for positive selection within populations and their distribution range, *COI* nucleotide sequences of all sampling locations were compared using the McDonald Kreitman test (McDonald & Kreitman, 1991) in DNASP. The number of haplotypes (*n*) and haplotype diversity (*Hd*) was estimated for all sampling locations and all clusters of low genetic distance in ARLEQUIN. Further, pairwise genetic distances

between populations were rounded to the full integer and frequencies were plotted for all genes. The number of non-synonymous substitutions in the protein-coding genes (*H3* and *COI*) relative to the most common haplotype was counted in both species.

Clusters of low genetic distances

The number of haplotypes of *COI* and alleles of *H3* was very high. For better overview and to infer geographic patterns, populations that were separated by genetic distances below the mean genetic distance among all populations minus the standard deviation minus 25% were grouped as genetic lineages. Resulting lineages were labelled numerically (*COI*) and alphabetically (*H3*) and in *F. quadrioculata* applied to all populations with <10% (*COI*) and <3% (*H3*) sequence divergence. In *I. minor* the threshold corresponded to <7% (*COI*) and < 2% (*H3*) sequence divergence. In order to test for past demographic changes, Tajima's D (Tajima, 1989) and Fu's F_S (Fu, 1997) were calculated in ARLEQUIN v3.5 using 10,000 permutations. Further, R^2 statistics (Ramos-Onsins & Rozas, 2002), which has higher power for small sampling sizes, was calculated in DnaSP v5.10.01 (Librado & Rozas, 2009). Demographic changes were tested on seven (*COI*) and six (*H3*) clusters of low genetic distance of *F. quadrioculata* and four clusters (both *COI* and *H3*) of *I. minor*.

Results

Population structure

Genetic variance among populations (AMOVA) was very high explaining 92-93% (*F. quadrioculata*) and 89-95% (*I. minor*) of the genetic variance of *COI* and *H3* (**Table 2**). Genetic distances of *COI* were very high between populations with 15 ± 2% and 15 ± 6% for *F. quadrioculata* and *I. minor*, respectively (**Fig. 1, Table S2, Supporting information**). Within populations distances were zero or very low (≤ 2%) except for three populations of *F. quadrioculata* and five populations of *I. minor* in which distances ranged between 7-10% and 4-13%, respectively. For *H3* genetic distances between populations were much lower with 6 ± 2% in *F. quadrioculata* and 5 ± 3% in *I. minor*. Genetic distances of *H3* within sampling locations were very low (≤ 0.5%) except for five populations of *F.*

quadrioculata and one population of *I. minor*. Despite the pronounced structure between populations, isolation by distance was rejected for all datasets as being not significant.

The frequency distribution of genetic distances between sampling locations in *F. quadrioculata* was unimodal with a maximum at 15-16% in *COI* and 6-7% in *H3* (**Fig. 1a**). In contrast, genetic distances between populations in *I. minor* were multimodally distributed with peaks at 17-19, 10 and 5% in *COI*, and 9, 5-6 and 1% in *H3* (**Fig. 1b**). Distribution of pairwise distances of the protein sequences of *COI* and *H3* and the nucleotide sequence of 28S were unimodal in *F. quadrioculata* (**Fig. 1c**) and multimodal in *I. minor* (**Fig. 1d**).

Table 2. Analysis of molecular variance partitioning among and within sampling locations of two springtail species (*Folsomia quadrioculata* and *Isotomiella minor*) sampled across Europe, based on sequence variance of a mitochondrial (*COI*) and nuclear (*H3*) gene. Asterisks indicate significant differences at p<0.05; df, degrees of freedom; Fst, F-statistics**.**

		COI		***H3***	
		among	**within**	**among**	**within**
F. quadrioculata	d.f.	30	112	30	112
	sum of squares	6921	501	1432	82
	variance components	47 Va***	4 Vb***	10 Va***	1 Vb***
	% variation	92	8	93	7
	fixiation indices	Fst 0.916***		Fst 0.933***	
I. minor	d.f.	22	81	22	81
	sum of squares	4746	470	804	34
	variance components	46 Va***	6 Vb***	8 Va***	0 Vb***
	% variation	89	11	95	5
	fixiation indices	Fst 0.89***		Fst 0.95***	

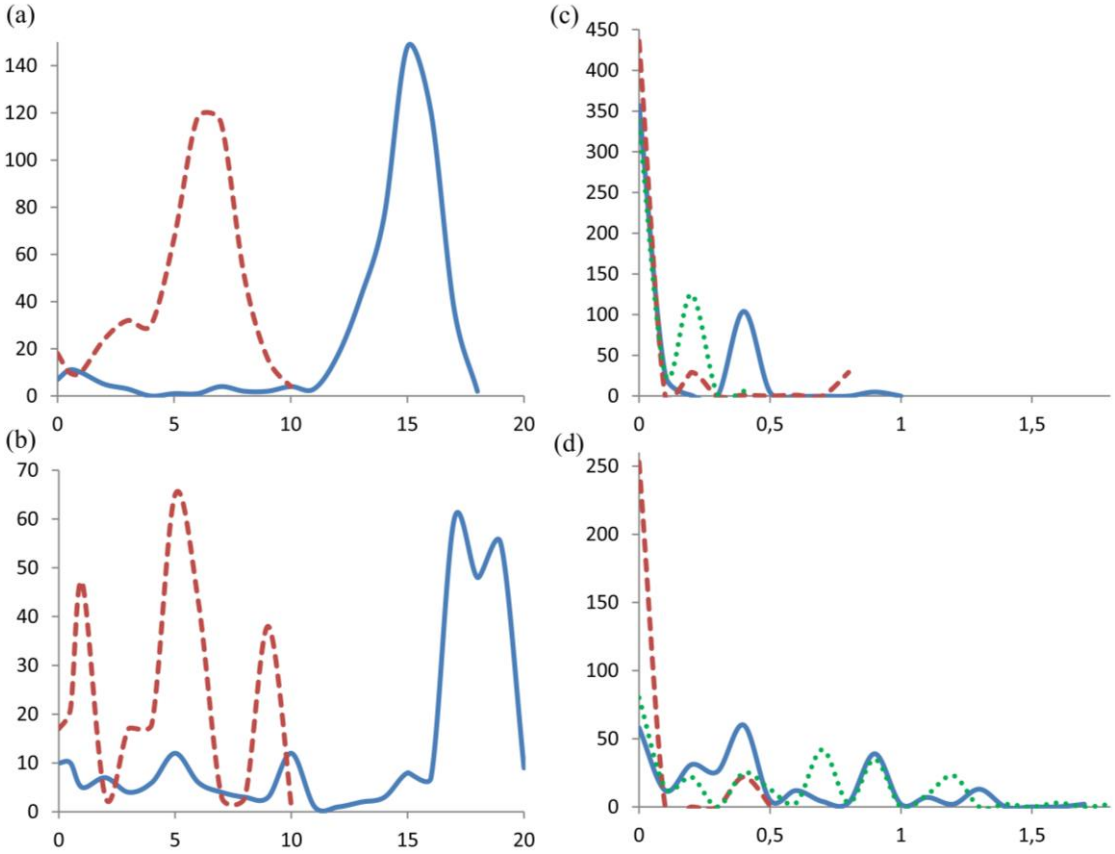


Figure 1. Frequency distributions of genetic distances between springtail populations in Europe. The sexual species *Folsomia quadrioculata* shows unimodal distributions (a, c), the parthenogenetic species *Isotomiella minor* shows multimodal distributions (b, d) of nucleotide (solid lines) and protein sequences (dashed lines) of *COI* (blue), *H3* (red) and nucleotide sequences of the partial region of 28S (green line).

Phylogenetic lineages, molecular age estimates and phylogeographic patterns

Phylogenetic trees based on the combined nucleotide sequences of *COI*, *H3* and 28S had the best resolution. Overall, sampling locations from southern Europe were early derived, and central and northern European sampling locations were later derived and more closely related to each other. Topologies differed markedly between species. In the sexual species *F. quadrioculata*, the majority of sampling locations formed monophyletic clades with very good posterior probabilities (pp: 0.96-1) and bootstrap supports (bt: 90-100) except for populations sampled from IT2 (pp: 0.81; bt: 100), FR2 (pp: 0.63; bt: 34) and DE1 (pp: 0.92; bt: 99; **Fig. 2**). The tree had 41 terminal branches, 13 sampling locations were isolated with long branches. In contrast, the tree of the parthenogenetic species *I. minor* had only 14 terminal branches, eight sampling locations were isolated

with long branches, but twelve sampling locations formed two large and closely related clades of which four sampling locations co-occurred in both clades (**Fig. 3**).

Genetic distances between several sampling locations were lower than average (**Table 3, Table S2, Supporting information**), which were pooled in low distance clusters. In *F. quadrioculata* seven *COI* lineages (1-7) and six *H3* alleles (A-F) had low genetic distances (**Fig. 2**). The most common *H3* lineage B was present in individuals with *COI* lineages 2 and 4, covering the largest geographic expansion ranging from south (IT1, IT2) to central Europe (FR2, FR3, CZ2). In contrast, *I. minor* had only four *COI* lineages (1-4) and four *H3* lineages (A-D) with low genetic distances (**Fig. 3)**. The *H3* lineage A occurred in all individuals of the two common and widespread *COI* lineages 1 and 2 that covered a large area from west to east to northern central Europe. The common *H3* lineage B was present in individuals from Norway, Germany and Spain. In Norway and Germany the *COI* lineage 3 was present but individuals from Spain carried two isolated *COI* lineages. Notably, one low distance cluster in *F. quadrioculata* also included individuals from Norway and Germany (cluster 6 E). Further, sampling locations DK2 and ES1 were distinct in *COI* (16% genetic distance) but had very similar *H3* alleles (1.1% genetic distance, cluster D) and identical 28S sequences.

Molecular age estimates based on a substitution rate of 2.3% per million years (**Fig. S1, Supporting information**) indicate that the majority of lineages separated in the Middle to Late Miocene (between 16-12 mya). During this time divergences were more pronounced in the sexual species *F. quadrioculata.* Clusters of lower than average genetic distance separated more recently during the Pliocene (5.3-1.8 mya) and Pleistocene (1.8-0.4 mya), and were older in *F. quadrioculata* (4.8-1.2 my) than in *I. minor* (1.1-0.7 my).

Haplotype Networks of *COI* were complex. Most haplotypes were identical or closely related within locations but separated from the next sampling location by many mutation steps (**Fig. S2, Supporting information**). Isolated haplotypes were most common in sampling locations from southern Europe, a pattern that was more pronounced in *F. quadrioculata* than in *I. minor*. In *I. minor* the two most common haplotypes were present in ~50% of all individuals. These common haplotypes corresponded to the *COI* low distance clusters 1 and 2 that had very similar *H3* alleles. Additionally, only cluster 2 had a star-like pattern of one common haplotype surrounded by four closely related haplotypes, a genetic pattern of recent expansion that was not

found for any other clade. Networks of *H3* were less complex, but similar to *COI* the nearest neighbours were separated by many mutation steps, a pattern that again was more common in the sexual species (**Fig. S3, Supporting information**). Networks of 28S were much simpler and differed markedly between species, being more complex in *I. minor* than in *F. quadrioculata* (**Fig. S4, Supporting information**). In *F. quadrioculata* sequences were similar (maximum genetic distance 0.4%) and constituted of only four very closely related alleles. In contrast, nine alleles were present in *I. minor* (maximum genetic distance 1.8%), being particularly different in southern Europe (HR, ME, SP3, SP2 and GR, IT) and separated from alleles from central and northern Europe. Further, all networks of *I. minor* demonstrated that two distinct genetic lineages coexist in CZ2 (28S, *H3* and *COI*).

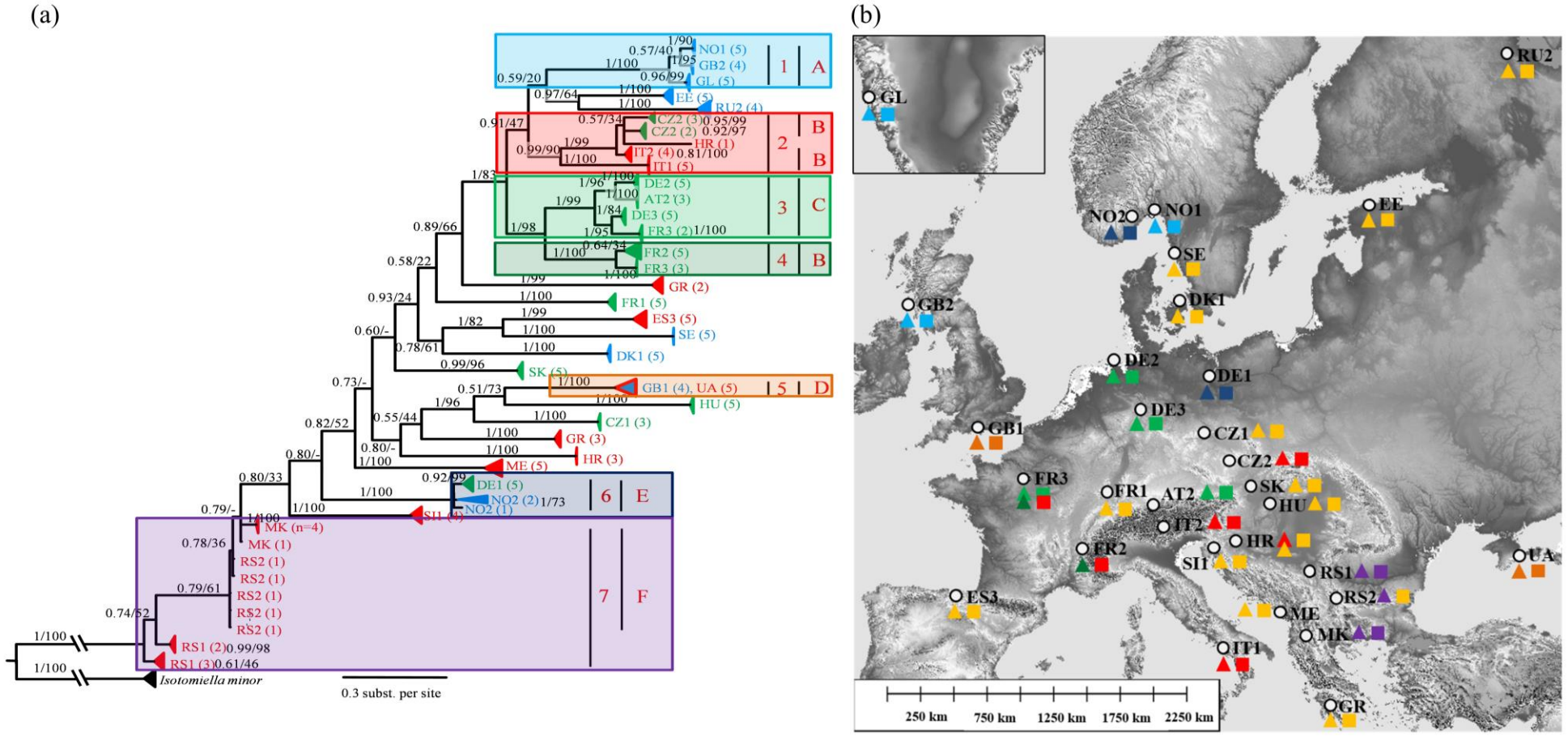


Figure 2. Bayesian phylogenetic tree (left) and sampling locations (right) of phylogenetic lineages of the sexual springtail *Folsomia quadrioculata* in Europe. Sampling locations are coloured corresponding to lineages in the phylogenetic tree. Mitochondrial (*COI*, triangle) and nuclear (*H3*, squares) lineages are distinguished into isolated lineages (orange) and clusters with low genetic distances (*COI*, lineages 1-7; *H3*, lineages A-F). The phylogenetic tree is based on the combined dataset (28S rDNA, *H3*, *COI*). Terminal clades have been collapsed, numbers of individuals of each clade are given as numbers in brackets. Numbers on nodes are posterior probabilities and bootstrap values.

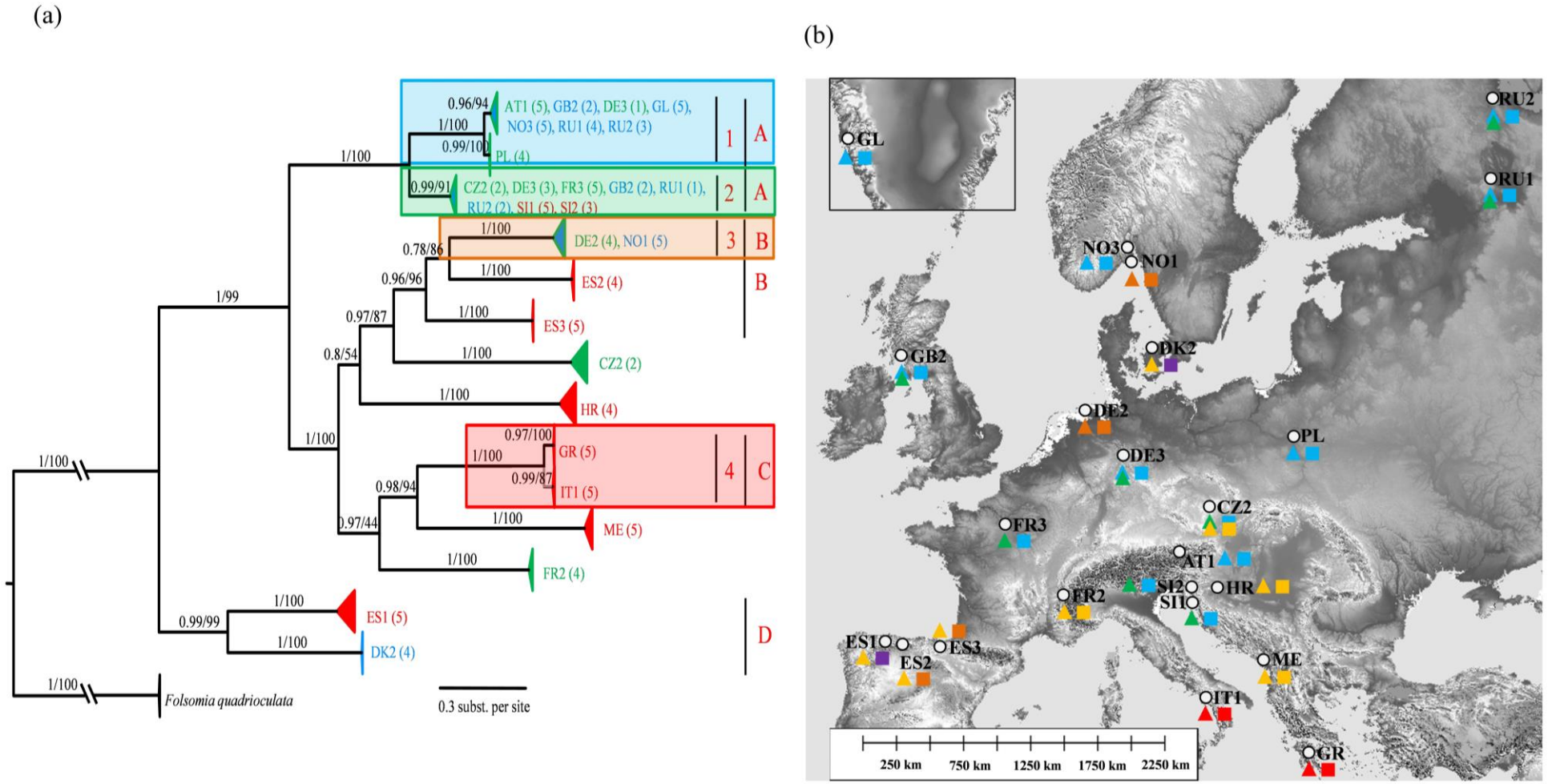


Figure 3. Bayesian phylogenetic tree (left) and sampling locations (right) of phylogenetic lineages of the parthenogenetic springtail *Isotomiella minor* in Europe. Sampling locations are coloured corresponding to lineages in the phylogenetic tree. Mitochondrial (*COI*, triangle) and nuclear (*H3*, squares) lineages are distinguished into isolated lineages (orange) and clusters with low genetic distances (*COI*, lineages 1-4; *H3*, lineages A-D). The phylogenetic tree is based on the combined dataset (28S rDNA, *H3*, *COI*). Terminal clades have been collapsed, numbers of individuals of each clade are given as numbers in brackets. Numbers on nodes are posterior probabilities and bootstrap values.

Table 3. Summary of tests for demographic changes and genetic variance within low distant cluster (*COI* and *H3*) of the sexual springtail *Folsomia quadrioculata* (a) and the parthenogenetic species *Isotomiella minor* (b). Genetic distances are uncorrected p-distances in percent. Number of individuals (n) and haplotype diversity (H_d) are given for each cluster. Names of clusters correspond to Figs. 2 and 3. Tajimas's D, Fu's F_S, R^2 and the corresponding p-values are listed, clades significant for Fu's Fs are highlighted in grey.

(a)

F. quadrioculata																			
cluster (*COI*)	**n**	***Hd***	**genetic distance %**	**Tajima's D**	**p**	**Fu's F_S**	**p**	**R^2**	**p**	**cluster (*H3*)**	**n**	***Hd***	**genetic distance %**	**Tajima's D**	**p**	**Fu's F_S**	**p**	**R^2**	**p**
1	14	0.8 ± 0.1	1.7-3.1	2.16	1	7.3	0.99	0.23	1	A	14	0.7 ± 0.1	0.5-1.4	1.70	0.96	2.11	0.87	0.23	0.97
2	10	0.8 ± 0.1	0.7-2.4	-0.62	0.28	3.49	0.94	0.21	0.86	B	22	0.9 ± 0.0	0.9-2.1	1.44	0.94	1.67	0.80	0.19	0.95
3	15	0.9 ± 0.1	3.2-6.8	1.93	0.99	8.36	1	0.22	1	C	15	0.8 ± 0.1	0.1-2.1	0.83	0.82	1.64	0.81	0.19	0.88
4	8	0.7 ± 0.1	3.4	1.49	0.96	8.46	1	0.23	0.92	B	22	0.9 ± 0.0	0.9-2.1	1.44	0.94	1.67	0.80	0.19	0.95
5	9	0.9 ± 0.1	1.4	1.73	0.98	0.27	0.5	0.23	0.95	D	9	0.6 ± 0.1	0.3	1.40	0.97	1.02	0.63	0.28	0.78
6	8	0.9 ± 0.1	1.9	-0.41	0.36	2.2	0.84	0.19	0.67	E	8	0.0 ± 0.0	0.0	0.00	1	-	-	-	-
7	15	0.8 ± 0.1	1.4-7.1	1.53	0.97	10.5	1	0.2	0.97	F	10	0.8 ± 0.1	1.9	2.16	1	2.49	0.89	0.27	0.98

(b)

I. minor																			
cluster (*COI*)	**n**	***Hd***	**genetic distance %**	**Tajima's D**	**p**	**Fu's F_S**	**p**	**R^2**	**p**	**cluster (*H3*)**	**n**	***Hd***	**genetic distance %**	**Tajima's D**	**p**	**Fu's F_S**	**p**	**R^2**	**p**
1	29	0.9 ± 0.0	0.0-0.8	-0.001	0.54	-3.4	0.02**	0.13	0.55										
2	23	0.9 ± 0.0	0.0-0.3	-1.30	0.08	-9.5	0***	0.19	0.98	A	52	1 ± 0.0	0.3-1.3	1.67	0.95	-11.66	0***	0.18	0.97
3	9	0.8 ± 0.1	1.5	0.93	0.84	2.21	0.84	0.2	0.74	B	18	0.8 ± 0.0	0.0-1.1	1.84	0.97	1.43	0.8	0.23	0.98
4	10	0.6 ± 0.1	1.9	2.29	1	6.58	0.99	0.26	1	C	10	0.6 ± 0.1	0.3	1.46	0.97	1.1	0.65	0.28	0.81
										D	9	0.6 ± 0.1	1.1	2.07	1	4.1	0.97	0.28	0.94

Demographic changes and genetic variation

For *I. minor* Fu's F_S values of *COI* clusters 1 and 2 and *H3* cluster B were negative and highly significant ($p < 0.02$ and $p < 0.001$, respectively; **Table 3**). For all other clades Fu's F_S was not significant. Tajima's D and R2 statistics were not significant for any clade.

Haplotype diversity (*Hd*) was higher in *F. quadrioculata* as compared to *I. minor* (**Table 1**). Seven of 31 sampling locations of *F. quadrioculata* had high haplotype diversity (*Hd*=0.9-1) and seven had no haplotype diversity (*Hd*=0). Maximum haplotype diversity in *I. minor* was lower with 0.8 and eight of 23 sampling locations containing only a single haplotype (*Hd*=0).

Both species constituted of high numbers of different haplotypes (**Fig. 4**). The number of *COI* nucleotide haplotypes was higher in the sexual *F. quadrioculata* with 70 haplotypes in 143 individuals as compared to the parthenogenetic *I. minor* with 32 haplotypes in 104 individuals. The more conserved nuclear genes *H3* and 28S had 54 and 4 alleles in *F. quadrioculata,* and 26 (*H3*) and 9 (28S) alleles in *I. minor*. However, the number of haplotypes of protein sequences was higher in the parthenogenetic *I. minor* with 16 (*COI*) and 5 (*H3*) haplotypes as compared to the sexual *F. quadrioculata* with 11 (*COI*) and 4 (*H3*) haplotypes.

The number of non-synonymous substitutions in the *COI* protein sequence was higher in *I. minor*, 17 of 235 amino acids differed from the most common haplotype, but only nine differed from the most common haplotype in *F. quadrioculata* (**Table S3, Supporting information**). In *I. minor* 15 amino acids remained in the same hydrophobicity class but five changed polarity; in *F. quadrioculata* five amino acids remained in the same hydrophobicity class and four changed polarity. The number of non-synonymous substitutions in the *H3* protein sequence were the same in both species (**Table S4, Supporting information**), only 4 of 124 amino acids differed from the most common haplotypes, but hydrophobicity changes were more pronounced in *I. minor* with two positions changing polarity, while in *F. quadrioculata* all four non-synonymous positions remained in the same hydrophobicity class.

Comparing dN/dS ratios between sampling locations using the McDonald-Kreitman test revealed significant deviations from neutrality for nine comparisons of *F. quadrioculata* and seven comparisons of *I. minor,* indicating predominantly negative selection on non-neutral substitution (NI > 1) in the Danish population of the sexual

species, but predominantly positive selection for non-neutral substitutions in the parthenogenetic species (NI < 1; **Table S5, Supporting information**).

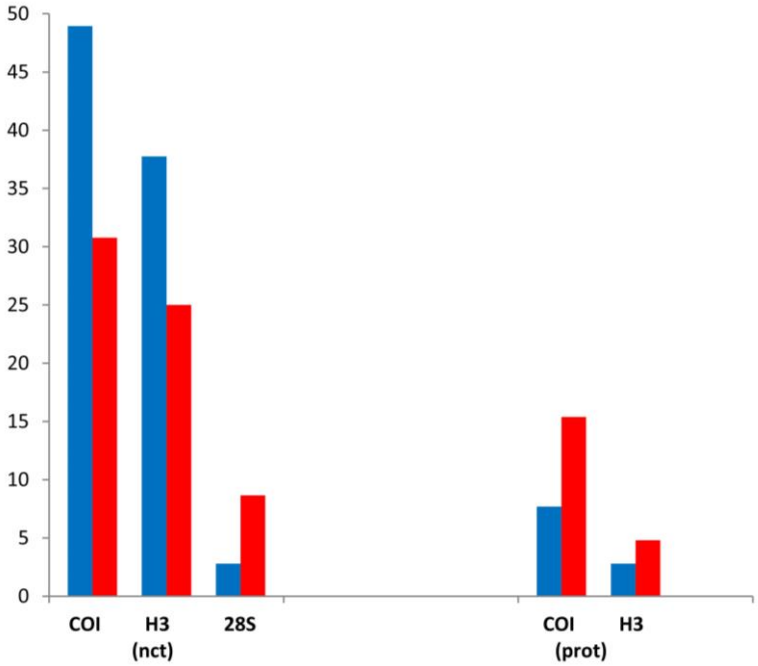


Figure 4. Percentage of nucleotide and protein haplotypes and alleles in the sexual species *Folsomia quadrioculata* (blue) and the parthenogenetic species *Isotomiella minor* (red).

Discussion

Refugia and dispersal

The genetic structure of the two springtail species was very high, irrespective of their reproductive mode. Across Europe many genotypes were isolated and most genetic lineages were separated by many mutation steps, indicating that a substantial part of present day genetic structure of European Collembola dates back to ancient divergences prior to Quaternary ice ages. Both species had isolated genotypes in areas south of the Alps and in central and northern Europe, indicating refuge sites in each of these areas. However, southern lineages contributed little to the genetic variance in central European populations as only few central European individuals were related with individuals from southern Europe. Rather, genetic diversification of lineages occurred deep in time suggesting that present day lineages in central Europe represent relict populations evolving since the Middle to Late Miocene. Tests for demographic changes were not

significant for these lineages, indicating persistence of stable populations at mutation-drift-equilibrium. Phylogeographic studies on subterranean amphipods (McInerney *et al.*, 2014) and soil-living oribatid mites (Rosenberger *et al.*, 2013) also revealed strong genetic and spatial structure and endemic lineages of Miocene and Late Pliocene origin. The high phylogenetic structure in European populations of groundwater isopods (Eme *et al.*, 2013) and a widespread parthenogenetic earthworm species (Fernández *et al.*, 2013) were also explained by local populations surviving Pleistocene glaciations. This suggests that environmental harshness is buffered in belowground systems allowing long-term persistence of populations. Notably, transplantation experiments demonstrated that Collembola communities are resistant against invasions (Ponge *et al.*, 2008) indicating stability of populations of soil-living detritivorous species. Accordingly, the high genetic structure between but not within springtail populations likely resulted from density-dependent processes, such as high-density blocking of local lineages and competition, as suggested by the founder takes all principle of colonization processes (Waters *et al.*, 2013).

Next to these rather local and old lineages, younger genetic lineages covered wider geographic areas and diverged from common ancestors in the Pliocene and Pleistocene. Notably, in both species the northern European regions of western Scotland, Norway and Greenland were colonized by a single genetic lineage of recent Pliocene origin, suggesting recent long-distance dispersal by similar, probably anthropogenic, means of transport. However, in contrast to our hypothesis, lineages were generally older in *F. quadrioculata* than in *I. minor*.

Dispersal and vicariance

Lineages of the sexual species *F. quadrioculata* were more regionally distributed, indicating limited long-distance dispersal ability. However, the presence of identical lineages in the Ukraine and southern Great Britain indicates that *F. quadrioculata* is able to disperse across wide ranges but local environmental factors or competition by autochthonous populations probably inhibited the establishment of novel lineages in most areas.

In contrast, two mitochondrial lineages of *I. minor* were common and widespread, covering a wide geographic area in central and northern Europe from west to east and

experienced population expansion according to Fu's F_S, indicating high dispersal ability of *I. minor*. These lineages coexisted in several sampling locations, suggesting that competition is less important in this species as compared to *F. quadrioculata*. The two lineages were associated with one nuclear lineage suggesting recent divergence. Negative values for Fu's F_S also indicate recent expansion in population size, supporting high colonization potential of this parthenogenetic species.

In both species, two distant sampling locations had identical or related genetic lineages, i.e. a vicariant distribution in Ukraine and Great Britain in *F. quadrioculata*, and in Denmark and Spain in *I. minor*. In *F. quadrioculata* individuals of both locations had identical genotypes, suggesting recent, human-mediated dispersal. In contrast, lineages of *I. minor* had closely related nuclear genotypes indicating descent from a common ancestral population but highly divergent mitochondrial haplotypes. Presumably, individuals from Denmark and Spain belong to relict populations in northern and southern refuges of a formerly widespread and potentially more continuously distributed lineage. According to the divergence of *COI*, lineages from Denmark and Spain have been evolving independently since the Middle or Late Miocene, supporting that ancient demographic changes structured populations of European Collembola and are still detectable today.

Mitochondrial and nuclear markers

The combination of three genetic markers allowed solid inferences on phylogeographic patterns of Collembola in Europe. The genetic variance of the mitochondrial marker alone was too high and of the 28S gene too conserved to infer geographic patterns, in particular in *F. quadrioculata*. The nuclear protein-coding gene *H3* had intermediate resolution with a substitution rate about three times slower than *COI* and corroborated geographic separations of mitochondrial lineages deeper in time. Interestingly, molecular variance and genetic distances were similar when comparing the complete datasets of both species. However, pairwise comparisons of genetic distances between sampling locations had a unimodal distribution in the sexual species, indicating a predominantly random distribution of substitutions in the mitochondrial and the nuclear genes. In contrast, genetic distances in *I. minor* were multimodally distributed, suggesting non-random increase in frequency of different *H3* alleles and *COI* haplotypes. Further, in

F. quadrioculata neutral molecular variance exceeded that in *I. minor*, with the latter having more haplotypes at the protein level, more non-synonymous substitutions in *COI* and more amino acid changes into different hydrophobicity classes, e.g. between electrically charged and uncharged amino acids. Presumably, different selective forces act on the sexual and parthenogenetic genomes which are supported by the neutrality index of the McDonald Kreitman test that indicated predominantly negative selection on non-neutral substitution in *F. quadrioculata* but predominantly positive selection for non-neutral substitutions in *I. minor*.

Several genes involved in mitochondrial functions are nuclear encoded and efficiency of mitonuclear interactions are impeded by mitochondrial and nuclear genes that evolved independently in hybrids of closely related species and even in hybrids from different populations within species (Ellison *et al.*, 2008; Montooth *et al.*, 2010; Gagnaire *et al.*, 2012). Mitochondria in sexually reproducing species experience different nuclear backgrounds every generation and maintaining synonymous variance within populations may be an effective way for keeping multiple non-interfering allelic combinations. In contrast, mitonuclear complexes in parthenogenetic species are intimately linked for generations, reducing the strength of selection on neutral substitutions. Indeed, a number of studies demonstrated that mitonuclear interactions affect the fitness of species with some allelic combinations being more efficient under certain environments (Dowling *et al.*, 2007; Arnqvist *et al.*, 2010; Hoekstra *et al.*, 2013; Wolff *et al.*, 2014).

The investigated nuclear genes are not part of the functional mitonuclear complex but differences in accumulation of synonymous and non-synonymous substitutions suggest, that additionally to competition and founder effects, mitonuclear and gene-by environment interactions between polymorphic populations may explain the high population structure in the two investigated species. Fitness differences between lineages may result in the dominance of one particular lineage in a particular environment but fitness may be reduced if local genotypes mix with immigrants. Notably, coexistence of mitochondrial haplotypes was considerably more pronounced in the parthenogenetic species *I. minor*, with the sexual species *F. quadrioculata* having only a narrow suture zone with closely related mitochondrial lineages in central France, possibly due to postzygotic selection or reduced fitness of hybrid offspring.

Conclusion

Results of the present study show that genetic variance in European Collembola species is remarkably high, with the sexual *F. quadrioculata* being more variable than the parthenogenetic *I. minor*. Genetic lineages of the two species separated deep in time contrasting aboveground species. Many populations are geographically isolated and have limited or vicariant distribution ranges, presumably due to Miocene to Pliocene separations. Lineages of *F. quadrioculata* tended to be older than those of *I. minor* suggesting that recombination and associated higher genetic variance facilitates persistence of populations, but the widespread occurrence of lineages of recent origin in *I. minor* indicates higher colonization potential of the parthenogenetic species. However, selective pressures on parthenogenetic and sexual genomes differ in order to maintain intragenomic or gene-by-environment interactions, and in addition to postglacial colonization patterns this likely also structured populations. Overall, the results indicate that the widely held view of postglacial colonization patterns predominate in central and northern European species does not hold for soil-living organisms including both parthenogenetic and sexual species. This suggests that in contrast to the view that parthenogenetic lineages are dead-ends in evolution, in soil they may persist over geological times calling for novel theories on the advantage of sexual reproduction also applying to soil-living taxa.

Acknowledgements

We thank Patrick Pachl for help with sampling of the Balkan region. We are grateful to Jens Bast, Julian Bock, Elena Corral Hernández, Guido Humpert, Sven Marhan, Jordi Moya-Laraño and Lucija Šerić Jelaska for sending litter and soil material.

This study was funded by the Deutsche Forschungsgemeinschaft (SCHE 376/13-1 u. 13-2)

References

Arnqvist G, Dowling DK, Eady P *et al.* (2010) Genetic architecture of metabolic rate: environment specific epistasis between mitochondrial and nuclear genes in an insect. *Evolution; International Journal of Organic Evolution*, **64**, 3354–333.

Avise JC (1994) *Molecular Markers, Natural History and Evolution* (JC Avise, Ed,). Chapman & Hall, New York.

Barton NH, Charlesworth B (1998) Why sex and recombination? *Science (New York, N.Y.)*, **281**, 1986–1690.

Bell G (1982) *The Masterpiece of Nature: The Evolution and Genetics of Sexuality.* (G Bell, Ed,). University of California Press, Berkeley.

Brower AV (1994) Rapid morphological radiation and convergence among races of the butterfly Heliconius erato inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 6491–6495.

Chahartaghi M, Scheu S, Ruess L (2006) Sex ratio and mode of reproduction in Collembola of an oak-beech forest. *Pedobiologia*, **50**, 331–340.

Chernova NM, Potapov MB, Savenkova YY *et al.* (2010) Ecological Significance of Parthenogenesis in Collembola. *Entomological Review*, **90**, 23–38.

Colgan DJ, McLauchlan A, Wilson GDF *et al.* (1998) Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. *Australian Journal of Zoology*, **46**, 419.

Debouzie D, Desouhant E, Oberli F *et al.* (2002) Resource limitation in natural populations of phytophagous insects. A long-term study case with the chestnut weevil. *Acta Oecologica*, **23**, 31–39.

Dowling DK, Abiega KC, Arnqvist G (2007) Temperature-specific outcomes of cytoplasmic-nuclear interactions on egg-to-adult development time in seed beetles. *Evolution; International Journal of Organic Evolution*, **61**, 194–201.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Drummond AJ, Suchard MA, Xie D *et al.* (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 3–6.

Dunger W, Schulz H, Zimdars B (2002) Colonization behaviour of Collembola under different conditions of dispersal. *Pedobiologia*, **327**, 316–327.

Ellison CK, Niehuis O, Gadau J (2008) Hybrid breakdown and mitochondrial dysfunction in hybrids of *Nasonia* parasitoid wasps. *Journal of Evolutionary Biology*, **21**, 1844–1851.

Eme D, Malard F, Konecny-Dupré L *et al.* (2013) Bayesian phylogeographic inferences reveal contrasting colonization dynamics among European groundwater isopods. *Molecular Ecology*, **22**, 5685–5699.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources*, **10**, 564–567.

Fernández R, Almodóvar A, Novo M *et al.* (2013) Earthworms, good indicators for palaeogeographical studies? Testing the genetic structure and demographic history in the peregrine earthworm *Aporrectodea trapezoides* (Dugès, 1828) in southern Europe. *Soil Biology and Biochemistry*, **58**, 127–135.

Fiera C (2013) Application of stable isotopes and lipid analysis to understand trophic interactions in springtails. *North-Western Journal of Zoology*, **10**.

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

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Gabriel W, Lynch M, Burger R (1993) Muller's ratchet and mutational meltdowns. *Evolution*, **47**, 1744–1757.

Gagnaire P-A, Normandeau E, Bernatchez L (2012) Comparative genomics reveals adaptive protein evolution and a possible cytonuclear incompatibility between European and American Eels. *Molecular Biology and Evolution*, **29**, 2909–2919.

Gernhard T, Hartmann K, Steel M (2008) Stochastic properties of generalised Yule models, with biodiversity applications. *Journal of Mathematical Miology*, **57**, 713–735.

Gregoire-Wibo C (1977) Biological aspects of the life history of *Folsomia quadrioculata* (Insecta, Collembola). *Annales de la Societe Royale Zoologique de Belgique*, **107**, 11–24.

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

Hamilton W (1980) Sex versus Non-Sex versus Parasite. *Oikos*, **35**, 282–290.

Hawes TC, Worland MR, Convey P *et al.* (2007) Aerial dispersal of springtails on the Antarctic Peninsula: implications for local distribution and demography. *Antarctic Science*, **19**, 3–10.

Hoekstra LA, Siddiq MA, Montooth KL (2013) Pleiotropic effects of a mitochondrial-nuclear incompatibility depend upon the accelerating effect of temperature in *Drosophila*. *Genetics*, **195**, 1129–39.

Hopkin SP (1997) *Biology of the Springtails*. Oxford University Press, Oxford.

Hopkin SP (2007) *A Key to the Collembola (Springtails) of Britain and Ireland*. Field Studies Council.

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

Langeneckert A (2013) Lipid composition of Collembola from different forest types (beech, spruce) as indicated by fatty acid analyses.

Latta LC, Morgan KK, Weaver CS *et al.* (2013) Genomic background and generation time influence deleterious mutation rates in *Daphnia*. *Genetics*, **193**, 539–544.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics (Oxford, England)*, **25**, 1451–1452.

Maynard Smith J (1978) *The Evolution of Sex* (J Maynard Smith, Ed,). Cambridge Univ Press, Cambridge.

McDonald J, Kreitman M (1991) Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature*, **351**, 652–654.

McInerney CE, Maurice L, Robertson AL *et al.* (2014) The ancient Britons: groundwater fauna survived extreme climate change over tens of millions of years across NW Europe. *Molecular ecology*, **23**, 1153–1166.

Milne I, Lindner D, Bayer M *et al.* (2009) TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics*, **25**, 126–127.

Montooth KL, Meiklejohn CD, Abt DN *et al.* (2010) Mitochondrial-nuclear epistasis affects fitness within species but does not contribute to fixed incompatibilities between

species of *Drosophila*. *Evolution; International Journal of Organic Evolution*, **64**, 3364–3379.

Omilian AR, Cristescu MEA, Dudycha JL *et al.* (2006) Ameiotic recombination in asexual lineages of *Daphnia*. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 18638–18643.

Ponge JF (1991) Succession of fungi and fauna during decomposition of needles in a small area of Scots pine litter. *Plant and Soil*, **138**, 99–113.

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.

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092–2100.

Rice WR (2002) Experimental tests of the adaptive significance of sexual recombination. *Nature reviews. Genetics*, **3**, 241–51.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Rosenberger M, Maraun M, Scheu S *et al.* (2013) Pre- and post-glacial diversifications shape genetic complexity of soil-living microarthropod species. *Pedobiologia*, **56**, 79–87.

Scheu S, Drossel B (2007) Sexual reproduction prevails in a world of structured resources in short supply. *Proceedings. Biological sciences / The Royal Society*, **274**, 1225–1231.

Song Y, Drossel B, Scheu S (2011) Tangled Bank dismissed too early. *Oikos*, **120**, 1601–1607.

Stamatakis a, Ott M, Ludwig T (2005) *Parallel Computing Technologies* (V Malyshkin, Ed,). Springer Berlin Heidelberg, Berlin, Heidelberg.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Thompson J, Gibson T (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research,* **25**, 4876–4882.

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

Waters JM, Fraser CI, Hewitt GM (2013) Founder takes all: density-dependent processes structure biodiversity. *Trends in Ecology & Evolution*, **28**, 78–85.

Williams G (1975) *Sex and Evolution* (G Williams, Ed,). Princeton University Press, Princeton.

Wolff JN, Ladoukakis ED, Enríquez JA *et al.* (2014) Mitonuclear interactions: evolutionary consequences over multiple biological scales. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*, **369**, 20130443.

Data Accessibility

DNA sequences: GenBank accessions see Table S1.

Supporting information

Table S1. GenBank accession numbers, GenBank abbreviation (GB ab.) and location of the Collembola *Folsomia quadrioculata* and *Isotomiella minor*

		GB	*F. quadrioculata*			*I. minor*		
	location	**ab.**	***COI***	***H3***	**28S**	***COI***	***H3***	**28S**
DK1	Humlebaek	-	KJ186302-6	KJ186487-91	KJ186166-70	-	-	-
DK2	Kopenhagen	-	-	-	-	KJ186385-8	KJ186434-7	KJ186249-52
EE	Tallinn	EE	KF684597-601	KF684762-6	KF684432-6	-	-	-
GL	Kobbefjord	-	KJ186331-5	KJ186516-20	KJ186195-9	KJ186398-402	KJ186447-51	KJ186262-6
GB1	Southhampton	-	KJ186307-10	KJ186492-5	KJ186171-4	-	-	-
GB2	Island of Arran	-	KJ186353-6	KJ186538-41	KJ186217-20	KJ186413-6	KJ186462-5	KJ186277-80
NO1	Rod	NO	KF684627-31	KF684792-6	KF684462-6	KF684678-82	KF684843-7	KF684513-7
NO2	Aske / Vestfold	-	KJ186350-2	KJ186535-7	KJ186214-6	-	-	-
NO3	Reinemoen	-	-	-	-	KJ186403-7	KJ186452-6	KJ186267-71
RU1	Petrozavodsk	-	-	-	-	KJ186408-12	KJ186457-61	KJ186272-6
RU2	Letnerechenskiy	RU1	KF684632-5	KF684797-800	KF684467-70	KF684687-91	KF684852-6	KF684522-26
SE	Gothenburg	-	KJ186371-5	KJ186556-60	KJ186235-9	-	-	-
AT1	Tirol	AT1	-	-	-	KF684646-50	KF684811-5	KF684481-85
AT2	Holzgau	AT3	KF684590-2	KF684755-7	KF684425-7	-	-	-
CZ1	Velemin	-	KJ186294-6	KJ186479-81	KJ186158-60	-	-	-
CZ2	Nove Domky	-	KJ186297-301	KJ186482-6	KJ186161-5	KJ186381-4	KJ186430-3	KJ186245-8
FR1	Voegtlinshoffen	-	KJ186311-5	KJ186496-500	KJ186175-9	-	-	-
FR2	Chartreuse	FR	KF684602-6	KF684767-71	KF684437-41	KF684655-8	KF684820-3	KF684490-3
FR3	Rambouillet	-	KJ186316-20	KJ186501-5	KJ186180-4	KJ186389-93	KJ186438-42	KJ186253-7
DE1	Unteres Odertal	-	KJ186321-5	KJ186506-10	KJ186185-9	-	-	-
DE2	Norden	-	KJ186326-9	KJ186511-5	KJ186190-4	KJ186394-7	KJ186443-6	KJ186258-61
DE3	Neuhaus	DE	KF684607-11	KF684772-6	KF684442-6	KF684659-62	KF684824-7	KF684494-7
HU	Tatabanya	-	KJ186336-40	KJ186521-5	KJ186200-4	-	-	-
PL	Warsaw	PL	-	-	-	KF684683-6	KF684848-51	KF684518-21
SK	Borinka	-	KJ186362-5	KJ186547-50	KJ186226-9	-	-	-
HR	Sljeme	HR	KF684593-6	KF684758-61	KF684428-31	KF684651-4	KF684816-9	KF684486-9
GR	Chrysovitsi	GR	KF684612-6	KF684777-81	KF684447-51	KF684663-7	KF684828-32	KF684498-502
IT1	Felitto	IT1	KF684617-21	KF684782-6	KF684452-6	KF684668-72	KF684833-7	KF684503-7
IT2	Parcines	-	KJ186341-4	KJ186526-9	KJ186205-8	-	-	-
MK	Struga	-	KJ186345-9	KJ186530-4	KJ186209-13	-	-	-
ME	Bar	ME	KF684622-6	KF684787-91	KF684457-61	KF684673-7	KF684838-42	KF684508-12
RS1	Markovac	RS	KF684636-40	KF684801-5	KF684471-75	-	-	-
RS2	Sreckovac	-	KJ186357-61	KJ186542-6	KJ186221-5	-	-	-
SI1	Jelsane	-	KJ186366-70	KJ186551-5	KJ186230-4	KJ186422-6	KJ186471-5	KJ186286-90
SI2	Postojna	-	-	-	-	KJ186427-9	KJ186476-8	KJ186291-3
ES1	Oviedo	ES1	-	-	-	KF684695-99	KF684860-64	KF684531-5
ES2	Ponga	ES2	-	-	-	KF684692-4, 700	KF684857-9, 65	KF684527-30
ES3	Martiartu	ES3	KF684641-5	KF684806-10	KF684476-80	KJ186417-21	KJ186466-70	KJ186281-5
UA	Staryi Krym	-	KJ186376-80	KJ186561-5	KJ186240-4	-	-	-

Table S2. Genetic distances between and within populations of the springtails *Folsomia quadrioculata and Isotomiella minor*. Blue and green colored distances were higher than the average.

			between populations				within populations			
			F. quadrioculata		*I. minor*		*F. quadrioculata*		*I. minor*	
		location	*COI*	*H3*	*COI*	*H3*	*COI*	*H3*	*COI*	*H3*
north	DK1	Denmark	15 ± 1	8 ± 1			0.1	0.2		
	DK2				18 ± 1	3 ± 3			0.0	0.0
	EE	Estonia	14 ± 1	6 ± 2			0.0	1.4		
	GL	Greenland	14 ± 3	5 ± 2	12 ± 7	3 ± 3	0.1	0.0	0.1	0.0
	GB1	Great Britain	15 ± 3	6 ± 2			1.3	0.0		
	GB2		14 ± 4	5 ± 2	12 ± 7	4 ± 2	0.0	0.1	6.4	0.4
	NO1	Norway	15 ± 4	6 ± 2	17 ± 5	4 ± 2	0.1	0.0	0.1	0.0
	NO2		15 ± 3	6 ± 2			1.4	0.0		
	NO3				12 ± 7	4 ± 3			0.0	0.2
	RU1	Russia			12 ± 7	3 ± 3			3.9	0.0
	RU2		15 ± 1	6 ± 2	12 ± 7	3 ± 2	1.0	0.0	5.8	0.2
	SE	Sweden	15 ± 1	7 ± 1			0.0	0.0		
central	AT1	Austria			12 ± 7	5 ± 2			0.0	0.0
	AT2		14 ± 3	5 ± 2			0.0	0.0		
	CZ1	Czech Republik	16 ± 1	7 ± 1			0.1	0.0		
	CZ2		13 ± 3	5 ± 2	15 ± 3	4 ± 2	0.3	0.1	13.3	4.1
	FR1	France	16 ± 1	7 ± 1			0.1	0.9		
	FR2		14 ± 2	5 ± 2	17 ± 1	3 ± 2	1.0	0.0	0.2	0.0
	FR3		14 ± 3	5 ± 2	12 ± 6	4 ± 2	6.9	1.6	0.1	0.4
	DE1	Germany	15 ± 3	6 ± 2			0.3	0.0		
	DE2		14 ± 3	5 ± 2	17 ± 4	5 ± 3	0.1	0.0	1.0	0.0
	DE3		14 ± 3	5 ± 2	12 ± 6	4 ± 3	0.1	0.1	4.9	0.1
	HU	Hungary	16 ± 1	7 ± 2			0.1	0.0		
	PL	Poland			12 ± 7	4 ± 3			0.0	0.0
	SK	Slovakia	15 ± 1	7 ± 1			0.6	0.5		
south	HR	Croatia	15 ± 1	6 ± 1	19 ± 1	3 ± 3	8.0	0.0	1.6	0.5
	GR	Greece	15 ± 1	7 ± 1	17 ± 4	3 ± 3	9.5	4.3	0.0	0.0
	IT1	Italy	14 ± 2	5 ± 2	17 ± 4	3 ± 3	0.0	0.0	0.1	0.0
	IT2		13 ± 3	5 ± 2			0.0	0.1		
	MK	Macedonia	14 ± 3	6 ± 1			0.7	0.1		
	ME	Montenegro	15 ± 1	6 ± 1	18 ± 1	4 ± 3	1.4	0.0	0.7	0.0
	RS1	Serbia	14 ± 2	6 ± 1			2.0	1.2		
	RS2		14 ± 3	6 ± 2			0.0	0.3		
	SI1	Slovenia	15 ± 1	6 ± 1	12 ± 6	3 ± 3	0.3	0.0	0.0	0.0
	SI2				12 ± 6	3 ± 3			0.0	0.0
	ES1	Spain			18 ± 1	8 ± 2			1.9	0.0
	ES2				17 ± 1	4 ± 2			0.1	0.0
	ES3		16 ± 1	7 ± 1	17 ± 1	4 ± 2	1.2	0.1	0.0	0.0
	UA	Ukraine	15 ± 3	6 ± 1			1.6	0.0		
		Mean	15 ± 2	6 ± 2	15 ± 6	5 ± 3	1 ± 2.1	0.4 ± 0.9	1.2 ± 2.1	0.1 ± 0.2

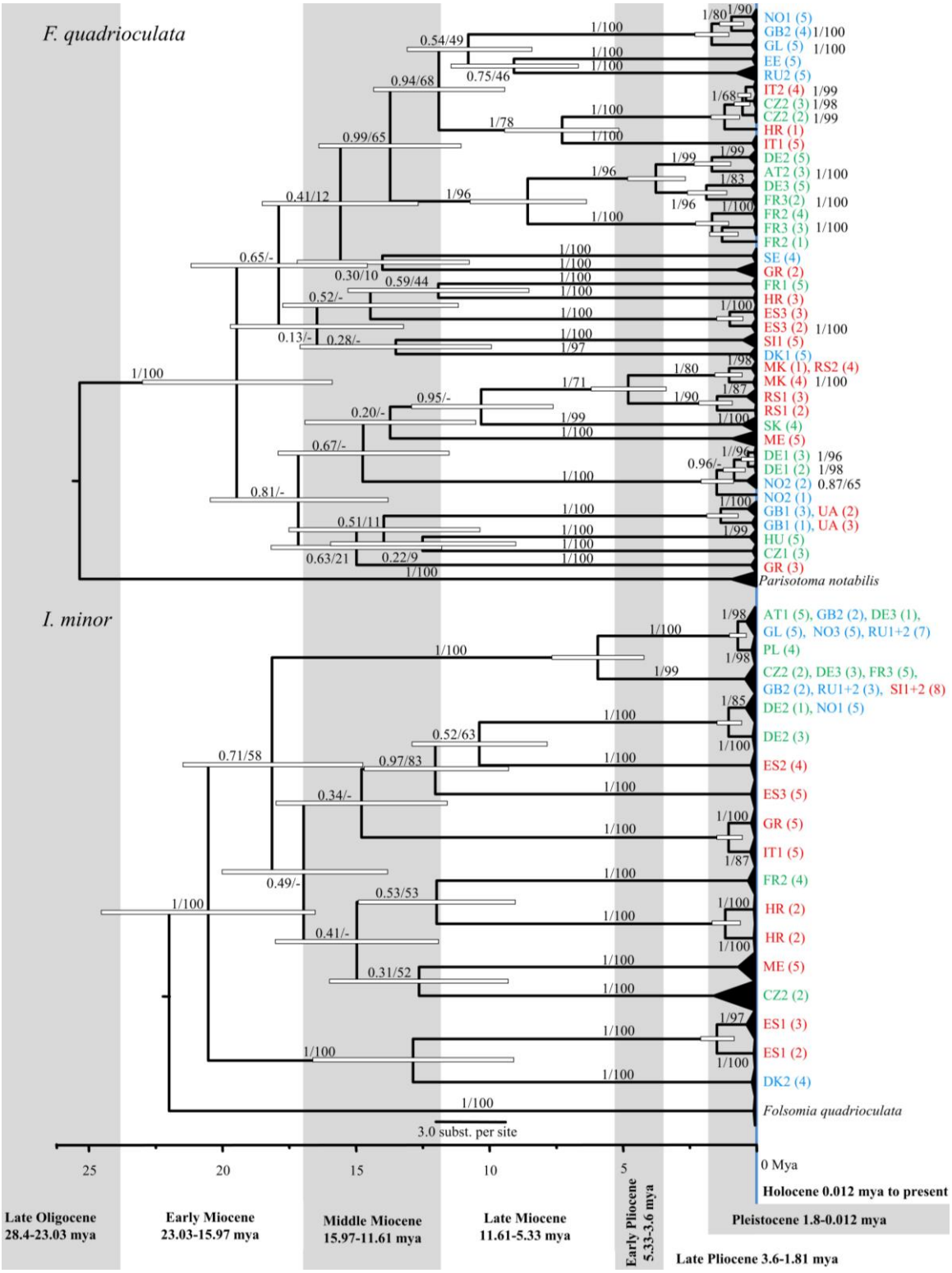


Figure S1. Molecular divergence times of *COI* of (a) *Folsomia quadrioculata* (sexual) and (b) *Isotomiella minor* (parthenogenetic) calculated with BEAST. Sampling locations from southern (red), central (green) and northern (blue) are abbreviated as in Table 1. Numbers on branches are posterior probabilities and bootstraps. Bars on nodes represent 95% confidence intervals; geological periods are indicated by grey and white areas.

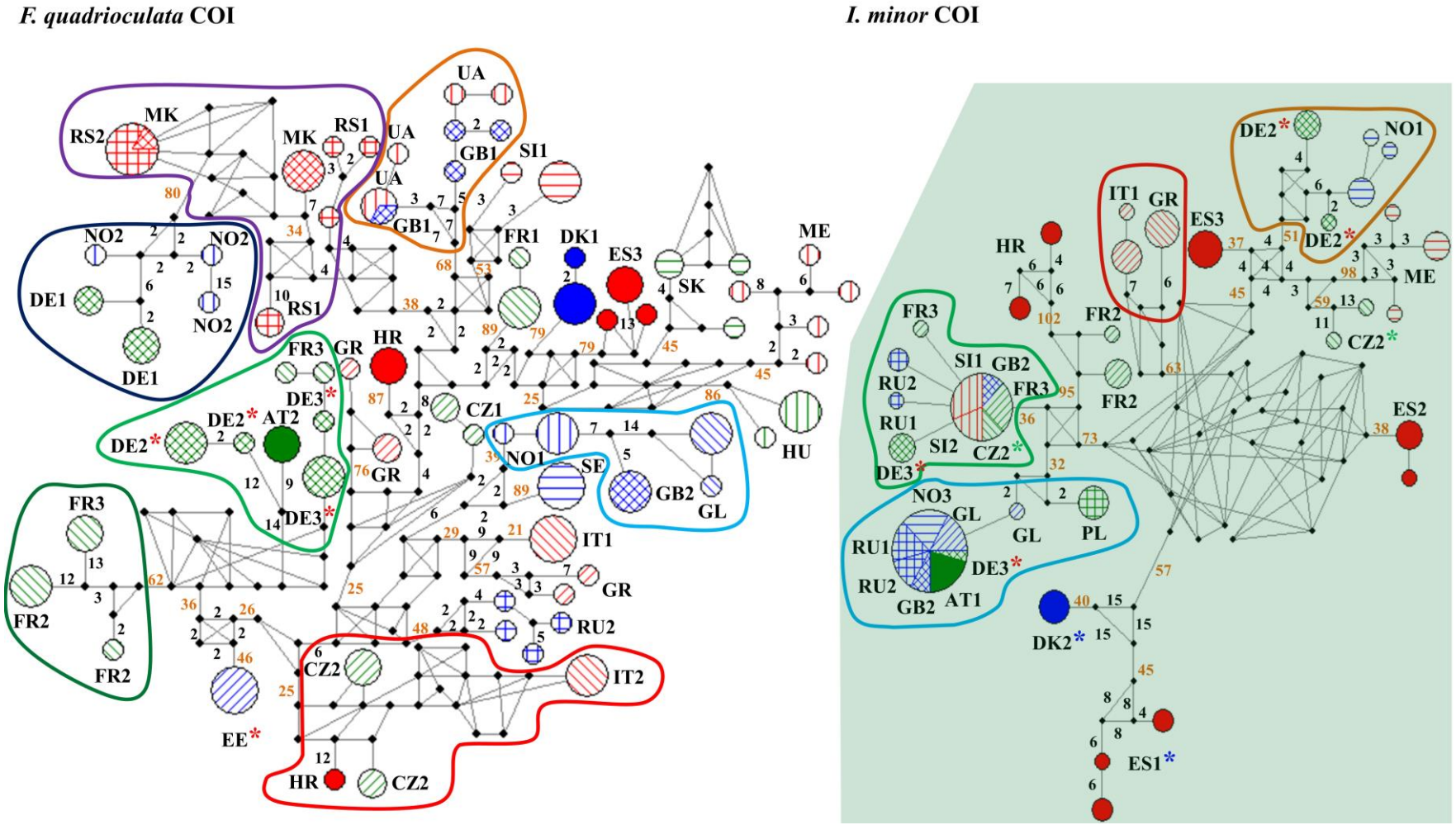


Figure S2. Median-joining networks of (a) the sexual species *Folsomia quadrioculata* and (b) the parthenogenetic species *Isotomiella minor* in Europe based on 709 bp of the mitochondrial protein coding gene *COI.* Numbers on lines represent hypothetical mutational steps between haplotypes, no number indicates a single mutation step, red numbers indicate high numbers of mutation steps (>20). Populations marked with red asterisk indicate that *COI* and *H3* networks did not correspond, i.e. within locations *COI* haplotypes were similar but *H3* alleles differed. Blue asterisk indicate identical 28S sequences and green asterisk indicate divergent lineages within sampling locations. Clusters of low genetic distances are outlined in colours corresponding to Figs. 2 and 3, for abbreviations of populations see Table 1.

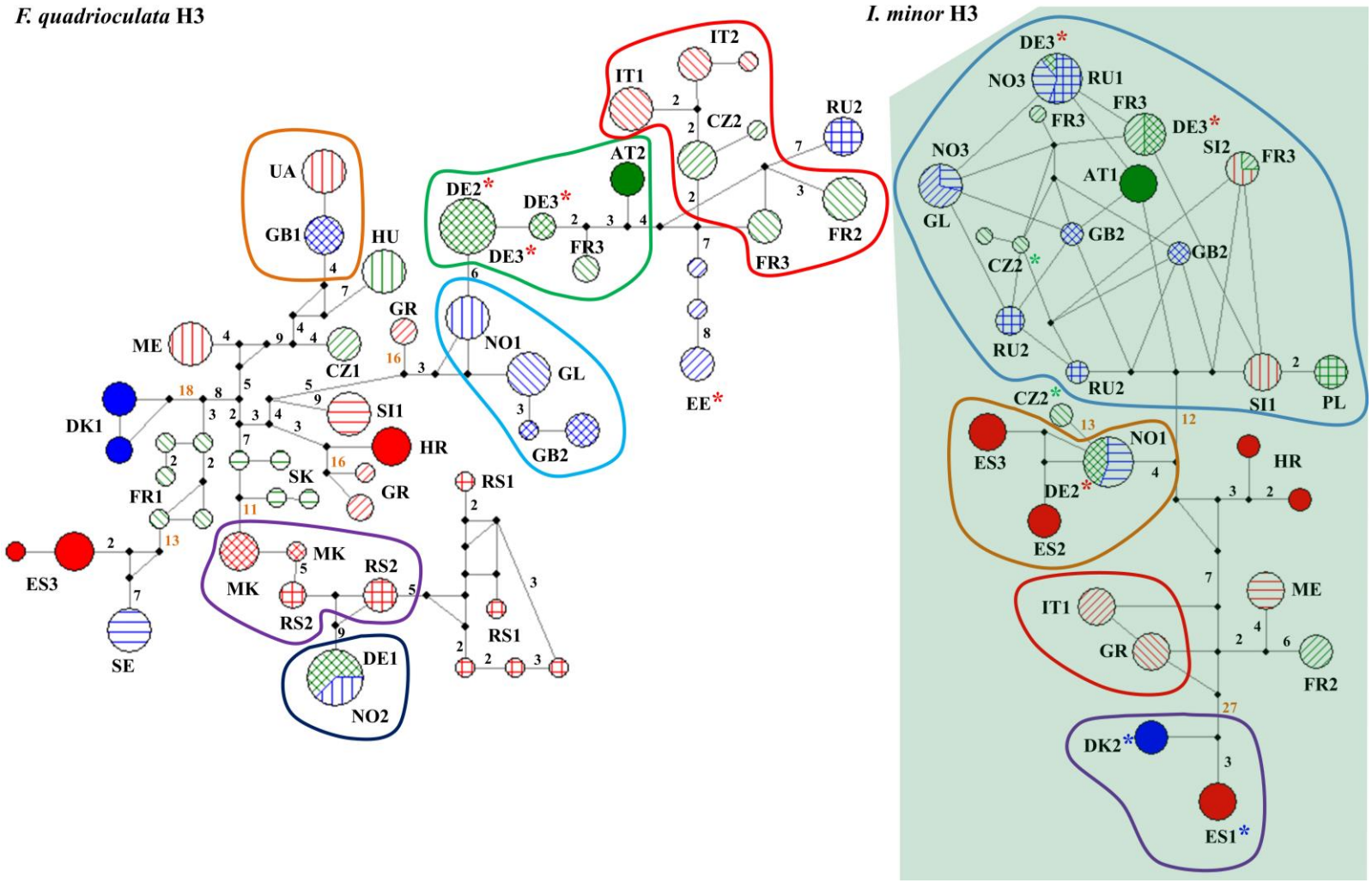


Figure S3. Median-joining networks of (a) the sexual species *Folsomia quadrioculata* and (b) the parthenogenetic species *Isotomiella minor* in Europe based on 374 bp of the nuclear protein coding gene *Histone H3.* Numbers on lines represent hypothetical mutational steps between haplotypes, no number indicates a single mutation step, red numbers indicate high numbers of mutation steps (>10). Populations marked with red asterisk indicate that *COI* and *H3* networks did not correspond, i.e. within locations *COI* haplotypes were similar but *H3* alleles differed. Blue asterisk indicate identical 28S sequences and green asterisk indicate divergent lineages within sampling locations. Clusters of low genetic distances are outlined in colours corresponding to Figs. 2 and 3, for abbreviations of populations see Tab. 1.

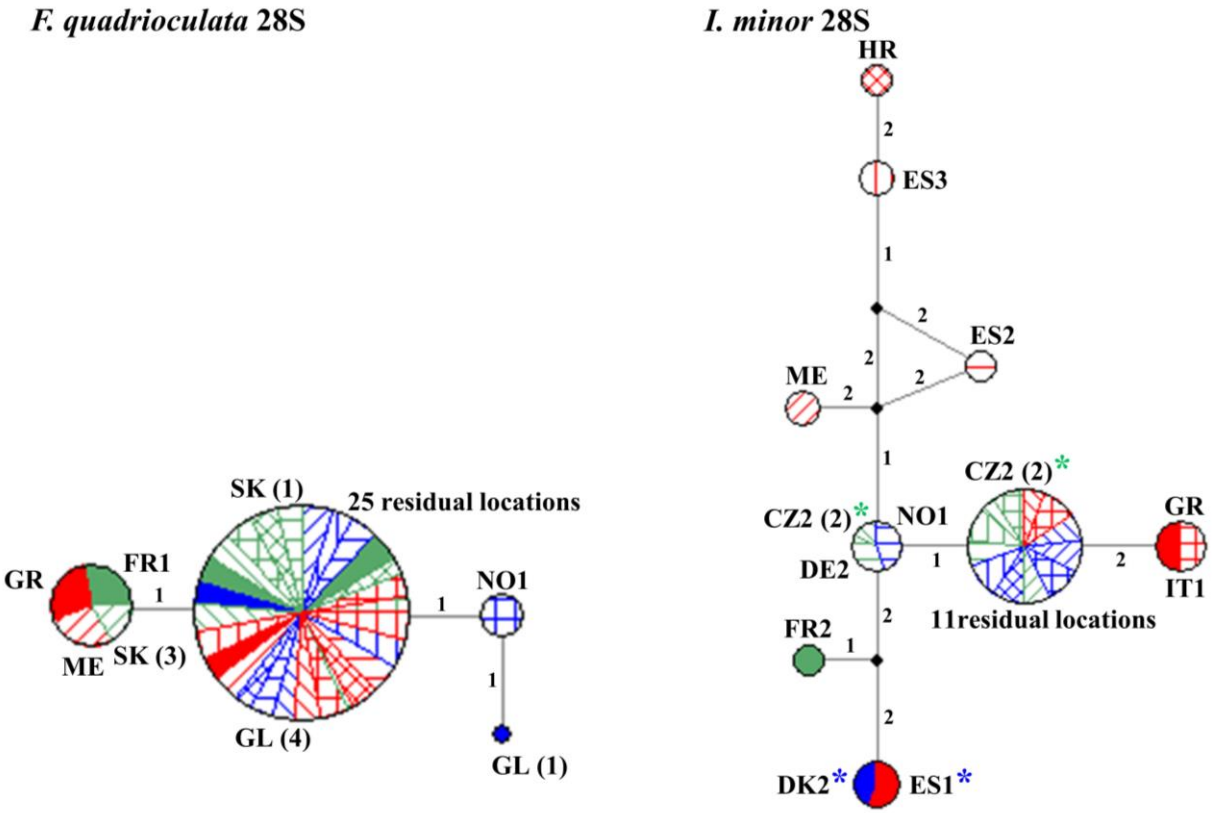


Figure S4. Median-joining networks of (a) the sexual species *Folsomia quadrioculata* and (b) the parthenogenetic species *Isotomiella minor* in Europe based on ~570 bp of the D3-D5 region of the nuclear 28S rDNA. Numbers on lines represent hypothetical mutational steps between haplotypes. Populations marked with blue asterisk indicate identical 28S sequences and green asterisk indicate divergent lineages within sampling locations. Clusters of low genetic distances are outlined in colours, corresponding to Figs. 2 and 3, for abbreviations of populations see Tab. 1.

Table S3. Non-synonymous amino acid substitutions among *COI* haplotypes (HT) of (a) *Folsomia quadrioculata* and (b) *Isotomiella minor*. Non-synonymous substitutions in the same hydrophobicity class are highlighted in blue and with changes in polarity are higlighted in red. Positions in the *COI* fragment are indicated.

(a)

		position of amino acid sequence						
HT	**n ind**	**09-11**	**17-19**	**55-57**	**74-76**	**101-103**	**164-173**	**n dif**
3	70	TMY	WSA	FIM	LVP	LIL	RTVGMTWDRT	0
1	23	TMY	WSA	FIM	LVP	LIL	RTEGMTWDRT	1
2	1	TMY	WSA	FIM	LVP	LVL	RTVGMTWDRT	1
4	1	TMY	WSA	FIV	LVP	LIL	RTVGMTWDRT	1
5	5	TMY	WSA	FIM	LVP	LIL	RTMGMTWDRT	1
6	5	TMY	WAA	FIM	LIP	LIL	RTVGMTWDRT	2
7	1	TMY	WSA	FIM	LVP	LIL	RAVGMTWDRT	1
8	19	TLY	WSA	FIM	LVP	LIL	RTVGMTWDRT	1
9	5	TMY	WAA	FIM	LVP	LIL	RTAGMTWDRT	2
10	9	TMY	WSA	FIM	LVP	LIL	RTVGMSWDRT	1
11	4	TMY	WSA	FIM	LVP	LIL	RTVGMTWDQT	1

(b)

		position of amino acid sequence									
HT	**n ind**	**15-28**	**55-57**	**78-80**	**99-103**	**108-110**	**126-129**	**164-174**	**188-190**	**206-208**	**214-216**	**n dif**
1	25	GVWSSMVGTAFSML	FVM	MIG	PSLIL	GMV	SMIA	RSVGMTWDRTP	LFS	NIN	PAG	0
2	4	GVWSAMVGTAFSVL	FIM	MIG	PSLML	GMV	SMMA	RSVGMTWDRTP	LLS	NIN	PAG	6
3	22	GVWSAMVGTAFSVL	FIM	MIG	PSLIL	GMV	SMIA	RSVGMTWDRTP	LLS	NIN	PAG	4
4	2	GVWSAMVGTAFSVL	FIM	MIG	PSLML	GLV	SMIA	RSVGMTWDRAP	LLS	NIN	PAG	7
5	4	GIWSAMVGTAFSVL	FIM	MIG	PSLAL	GMV	SAIA	RTTGMTWDRTP	LLS	NLN	PAG	10
6	8	GVWSAMVGTAFSVL	FVM	MIG	PSLML	GMV	SMIA	RSVGMTWDRTP	LLS	NIN	PAG	4
7	8	GVWSAMVGTAFSVL	FIM	MIG	PSLML	GMV	SLMA	RSAGMSWDRTP	LLS	NIN	PAG	9
8	5	GVWSAMVGTAFSVL	FVM	MIG	PSLML	GMV	SVIA	RSAGMTWDRTP	LLS	NIN	PAG	6
9	5	GVWSAMLGTAFSVL	FVM	MIG	PSLML	GMV	SVIA	RSAGMTWDRTP	LLS	NIN	PAG	7
10	5	GVWAAMVGTAFSVL	FIM	MIG	PSLML	GMV	SMIA	RTAGMTWDRTP	LLS	NIN	PAG	8
11	1	GVWSAMVGTAFSVL	FIM	MIG	PSLML	GMV	SLMA	RSAGMSWDRTP	LLS	NIN	PPG	10
12	4	GVWSSMVGTAFSVL	FVM	MIG	PALIL	GMV	SMIA	RSVGMTWDRTP	LLS	NIN	PAG	3
13	1	GVWSAMVGTAFSVL	FIM	MVG	PSLIL	GMV	SMIA	RSVGMTWDRTP	LLS	NIN	PAG	5
14	3	GVWSAMVGTAFSVL	FIM	MIG	PSLAL	GLV	SAIA	RTVGMTWDRTP	LLS	NLN	PAG	9
15	2	GIWSAMVGTAFSVL	FIM	MIG	PSLAL	GLV	SAIA	RTVGMTWDRTP	LLS	NLN	PAG	10
16	5	GVWSAMVGTAFSVL	FVM	MIG	PSLIL	GMV	SMIA	RSVGMTWDRTP	LLS	NIN	PAG	3

Table S4. Non-synonymous amino acid substitutions among *H3* haplotypes (HT) of (a) *Folsomia quadrioculata* and (b) *Isotomiella minor*. Non-synonymous substitutions in the same hydrophobicity class are highlighted in blue and with changes in polarity are highlighted in red. Positions in the *H3* fragment are indicated.

(a)

		position of amino acid sequence			
HT	**n ind**	**24-30**	**89-91**	**104-106**	**n dif**
1	135	KAARKSA	AVM	LFE	0
2	3	KAARKSA	AIM	LFE	1
3	1	KAARKSA	AVM	LLE	1
4	4	KSARKRA	AVM	LFE	2

(b)

		position of amino acid sequence				
HT	**n ind**	**10-12**	**17-19**	**42-44**	**72-74**	**n dif**
1	52	KST	PSK	YSP	VRE	0
2	22	KST	PRK	YRP	VRE	2
3	2	KWT	PSK	YRP	VRE	3
4	9	KST	PRK	YSP	VSE	2
5	19	KST	PRK	YSP	VRE	1

Table S5. Significant comparisions of McDonald-Kreitman test of (a) the sexual springtail *Folsomia quadrioculata* and (b) the parthenogenetic springtail *Isotomiella minor*.

(a)

Folsomia quadrioculata							
Population	**AT2**	**GB2**	**EE**	**ME**	**NO2**	**RS2**	**SI1**
DK1	111*	108*	101*			94*	97*
FR3				0.00***	0.00***		
GR				0.07**	0.04***		

(b)

Isotomiella minor							
Population	**AT1**	**DK2**	**DE2**	**GL**	**NO1**	**NO3**	**ES1**
HR	0.13*						
CZ2		0.10**	0.18*	0.13*		0.13*	015**
GR					30.25*		

Chapter 4

COLONIZATION OF EUROPE BY *PARISOTOMA NOTABILIS* (COLLEMBOLA): CRYPTIC DIVERSITY AND ANTHROPOGENIC IMPACT

Helge von Saltzwedel, Stefan Scheu and Ina Schaefer

submitted

Abstract

Climatic and biome changes of the past million years influenced the population structure and colonization routes of soil-living arthropods in Europe, but little is known on the colonization history and genetic structure of one of the most widespread and abundant Collembola species, *Parisotoma notabilis*. This generalistic and parthenogenetic species is a fast colonizer and often occurs in man-made environments. To investigate ancient climatic and recent anthropogenic influences on the genetic structure of *P. notabilis* we analyzed populations on a pan-European scale using three genetic markers differing in substitution rates. The results showed that *P. notabilis* comprises several genetic lineages with distinct distribution ranges that diverged in the Miocene, with the 4x rule indicating that the lineages form a single species. Further, the results indicate that anthropogenic activities likely influenced the present day genetic structure of *P. notabilis*. Compared to other soil-living arthropods, European lineages of *P. notabilis* are rather young and genetically rather uniform. The close association with anthropogenic habitats presumably contributed to rapid colonization of Europe by following human trading routes and migration. Human facilitated dispersal resulted in founder effects and the establishment of locally genetically homogenous lineages.

Key words: Colonization, Collembola, anthropogenic influence, cryptic species, phylogeography, Miocene divergence, soil arthropod

Introduction

The ubiquitous soil arthropod species *Parisotoma notabilis* (Schäffer, 1896) is one of the most successful species among Collembola being locally abundant in virtually any habitat in the temperate and boreal zone. Populations can reach densities of up to 10,000 and 6,000 individuals per square in forest soils and meadows, respectively, but also typically are present in arable fields, pastures, urban soils and caves (Soto-Adames 2002; Wanner & Dunger 2002; Fountain & Hopkin 2004; Kováč *et al.* 2005; Kuznetsova 2006; Ponge *et al.* 2008; García-Gómez *et al.* 2009; Salamon & Alphei 2009), and even in extreme habitats such as open glacier forelands at high elevation (Hågvar 2010). *P. notabilis* is the most abundant Collembola species in Europe (Fiera & Ulrich 2012) and together with *Isotomiella minor* (Schäffer, 1896) it often represents more than 50% of the total individuals in Collembola communities (Kuznetsova 2006; García-Gómez *et al.* 2009). It is morphologically well defined (Fjellberg 1977; Deharveng 1981; Rusek 1984; Potapov 1991), but exhibits inter-population differences in tolerance against low pH, mechanical disturbances and metal pollution (Fountain & Hopkin 2004; Salamon & Alphei 2009). According to stable isotope ratios of $^{14}N/^{15}N$ it feeds as generalist on bacteria, fungi and smaller soil animals including protozoa, nematodes and rotifers (Chahartaghi *et al.* 2005). Notably, *P. notabilis* reproduces asexually, no males have been found in natural populations (Chahartaghi *et al.* 2006) except for a Swedish population where males rarely occur (Fjellberg 1980). Wind dispersal (Wanner & Dunger 2002), the potential to start populations from a single female individual and generalist feeding predestine this species as fast and successful colonizer of new and disturbed habitats (Williams 1975; Bell 1982; Scheu & Schulz 1996; Lindberg & Bengtsson 2005; Ingimarsdóttir *et al.* 2012).

The genetic structure of *P. notabilis* populations is little known except for one study investigating genetic variation within and between European populations (Porco *et al.* 2012). Based on two genetic markers (*COI* and D2 region of 28S rDNA) they demonstrated that *P. notabilis* comprises four different lineages in Europe, with low genetic variance within (<3% for *COI* and zero for D2) but high variance between lineages (21% for *COI* and <3% for D2). The authors concluded that these four lineages represent 'cryptic species' which evolved independently but without morphological differentiation.

Deep genetic divergences in soil-living arthropods have been described previously and may be due to strong founder effects and genetic bottlenecks after long-distance dispersal combined with limited local dispersal within the soil-matrix (Rosenberger *et al.* 2013). Ancient divergences and survival in small patches during the Quaternary ice-ages may also generate patterns of deep divergence in *P. notabilis*, however, these questions have not been addressed yet.

In order to investigate the relevance of founder effects and historical dispersal patterns, we extended the geographic sampling of the previous study (Porco *et al.* 2012) and included west Russia, the Ukraine, Turkey, the Balkan Peninsula, Norway, Great Britain and Greenland. Thereby, the sampling included areas of northern Europe that were covered by glaciers during the last Ice Age, i.e. regions that must have been colonized in the Holocene by *P. notabilis*, resulting in populations of low genetic variation. Colonization likely occurred from southern Europe and south-eastern Russia, similar to the grasshopper *Chorthippus parallelus*, the hedgehog *Erinacues europeaus,* the bear *Ursus arctos*, the alder *Alnus glutinosa* and oaks *Quercus spp*. (Hewitt 1999; Hewitt & Ibrahim 2001; Sommer & Benecke 2005). Therefore, we expected northern lineages to be closely related to lineages of western and central Europe, but distantly to those south of the Alps.

We used three genetic markers, the mitochondrial *COI* gene, the D3-D5 region of 28S rDNA and the nuclear gene *Histone H3* that provided resolution intermediate to 28S and *COI*. This is the first study comparing two protein-coding genes and one ribosomal gene with different mutation rates to detect recent and old diversifications, independent evolutionary units (IEUs) and migration routes of Collembola in Europe.

Materials and Methods

Ethics statement

Sampling sites were outside Nature Reserve Areas and no permission for soil samples was required. The field study did not involve any endangered or protected species.

Sampling of animals and DNA extraction

Leaf litter and humus layers from about two square meters of deciduous and coniferous forests was collected in 26 locations in Europe, including Greenland, northwest Russia (Karelia), Ukraine, Turkey, the Balkan region (Bulgaria, Serbia, Croatia, Greece), Italy, Spain, and transferred to the University of Göttingen (**Table 1, Fig. 1**). Animals were extracted by heat, collected in water (Kempson *et al.* 1963), transferred into 96% EtOH and stored at -20°C until further analyses. For species identification specimens were sorted under a dissecting microscope and determined by light microscopy following (Hopkin 2007). Genomic DNA was extracted from single individuals of *P. notabilis* (n=120) using the DNeasy® Blood and Tissue Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol for animal tissue. Purified DNA was eluted in 30 µl buffer AE and stored at -20°C until further preparation. Two nuclear genes, *Histone H3* and the D3-D5 region of 28S rDNA, and the barcoding fragment of the mitochondrial *COI* gene were amplified in 25 µl volumes containing 12.5 µl SuperHot Taq Mastermix (Genaxxon Bioscience GmbH, Ulm, Germany) with 1.5 µl of each primer (10 pM), 4.5 µl H_2O, 2 µl $MgCl_2$ (25 mM) and 3 µl template DNA. A 374 bp fragment of the protein coding gene *H3* was amplified, using the primers H3F1 5'-ATG GCT CGT ACC AAG CAG ACV GC-3' and H3R1 5'-ATA TCC TTR GGC ATR ATR GTG AC-3' (Colgan *et al.* 1998). A ~573 bp fragment of the nuclear 28S rDNA was amplified using the primers 28Sa 5'-GAC CCG TCT TGA AGC ACG-3' and 28Sbout 5'-CCC ACA GCG CCA GTT CTG CTT ACC-3' (Tully *et al.* 2006). For the 709 bp fragment of the *COI* gene the primers LCO1490 5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3' and HCO2198 5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3' (Folmer *et al.* 1994) were used. PCR conditions included one initial activation step at 95°C for 15 min, followed by 35 amplification cycles of denaturation at 94°C for 15 s, annealing at 45°C (*COI*) or 49°C (28S) or 59°C (*H3*) for 15 s, elongation at 72°C for 15 s and a final elongation step at 72°C for 6 min. Positive PCR products were purified with the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol, eluted in 30 µl HPLC water and sent for direct sequencing to the Göttingen Genome Laboratory (Institute for Microbiology and Genetics, Georg August University of Göttingen). All sequences are available at GenBank (**Table S1, Supporting information**).

Genomic DNA was extracted from entire specimens but secondary vouchers (same morphological species from the same population) were deposited at our collections at J.F. Blumenbach Institute of Zoology and Anthropology, Georg August University Göttingen, Germany.

Table 1. Countries and sampling locations of *Parisotoma notabilis* collected in south, central and northern Europe. Abbreviations of sampling locations, the number of individuals analyzed, names and symbols for genetic lineages are listed.

	country	location	abbre-viation	n inds	IEUs (lineages)
north	**Denmark**	Humlebaek	DK	5	0, 1, 2
	Great Britain	Melrose	GB1	5	0
		Island of Arran	GB2	5	2
	Greenland	Kobbefjord / Nuuk	GL	5	2
	Norway	Rod	NO1	4	2
		Skjervenmoen, Fössa Öst	NO2	4	2
	Russia	Petrozavodsk, Karelia	RU1	5	2
		Znamenskoe	RU2	5	1
central	**France**	Voegtlinshoffen	FR1	4	2
		Salavas	FR2	5	1
		Chartreuse	FR3	5	2
		Rambouillet	FR4	4	2
	Germany	Uelzen - Elbeseitenkanal	DE1	5	1, 2
		Norden	DE2	4	2
		Solling, Neuhaus	DE3	5	0
south	**Bulgaria**	Bosnek	BG	5	1
	Croatia	Sljeme	HR	5	4
	France	Korsika, Olmi-Capella	FR5	5	1
	Greece	Chrysovitsi	GR	5	3
	Italy	Felitto	IT	4	1
	Serbia	Sreckovac	RS	5	1
	Slovenia	Postojna	SI	3	2
	Spain	Oviedo	ES1	3	2
		Ponga	ES2	5	2
	Turkey	Kayseri	TR	5	1
	Ukraine	Kubalach, Crimea	UA	5	1
		total no. of individuals.		**120**	

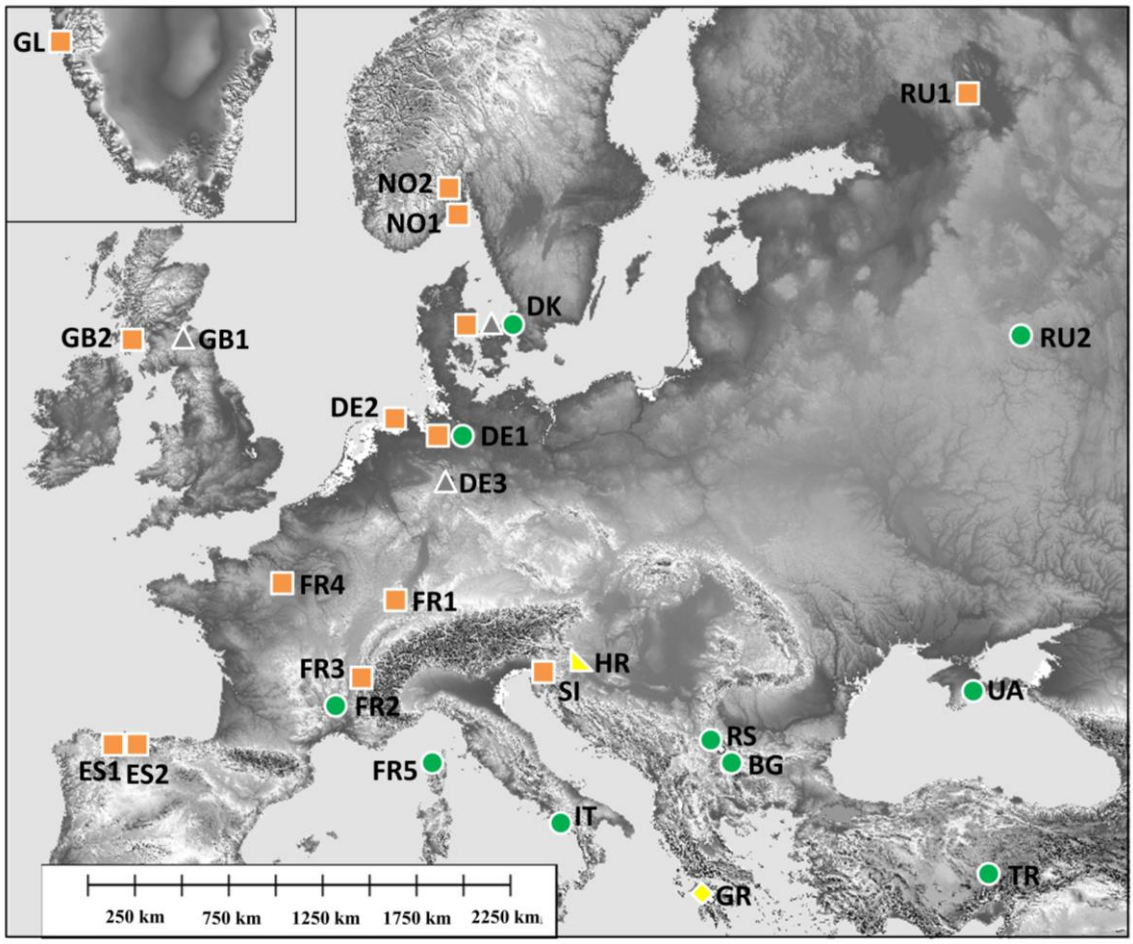


Figure 1. Sampling locations and distribution of lineages of *Parisotoma notabilis*. Genetic lineages were named following Porco et al. (2012), lineage L0 (grey triangles) occurs in three sampling locations (DE1, DK, GB1), while lineages L1 (green circles) and L2 are widespread in the southwest and the east of Europe, respectively. Lineages L3 (yellow diamond) and L4 (yellow turned triangle) are geographically isolated in the south of Europe.

Data analysis

Sequences were edited, ambiguous positions were corrected by hand and nucleotide sequences were translated into aminoacid sequences using the invertebrate mitochondrial code implemented in Sequencher v4.10 (Gene Codes Corporation, USA). Nucleotide (28S) and protein sequences (*COI* and *H3*) were aligned separately and as combined matrix (concatenated sequences of all three genes) with Clustal W (Thompson *et al.* 1994) implemented in BioEdit v7.0.1 (Hall 1999); protein alignments were retranslated to nucleotide sequences.

The best fit model of sequence evolution for each alignment (*COI*, 28S, *H3*, combined matrix) was inferred according to the hLRT in TOPALi v2.5 (Milne *et al.* 2009) using the PHYML algorithm. Phylogenetic trees were calculated with Maximum Likelihood in RAxML v8.0.0 (Stamatakis *et al.* 2005) and Bayesian Inference (BI) in MrBayes v3.1.2 (Ronquist & Huelsenbeck 2003). Phylogenetic analyses were performed for single genes (28S, *COI*, *H3*) and the combined matrix. The model of sequence evolution was GTR+I+Γ for the *COI* and the combined matrix, GTR+ Γ for 28S and JC for the *H3* matrix. For ML analyses, parameters were GTRGAMMAI and 10,000 bootstrap replicates. For Bayesian inference lset parameters were nst=6, rates=invgamma for the combined and the *COI* matrix, nst=6, rates=gamma for 28S and nst=1 for *H3*. Two independent MCMC chains were run for ten million generations that were sampled every 1,000th generation, a burnin of 2,500 was applied.

For all *COI* sequences obtained in this study (120 individuals, 709 bp), the number of independent evolutionary units (IEUs) was inferred with a GMYC (general mixed Yule-coalescent) analysis (Pons *et al.* 2006; Fontaneto *et al.* 2007; Fujisawa & Barraclough 2013); IEUs were tested for parthenogenetic speciation with the 4x rule (Birky *et al.* 2005; Birky 2013). For GMYC analysis, an ultrametric tree was generated in BEAST v1.8.0 (Drummond *et al.* 2012) with GTR+I+Γ as model of sequence evolution. The MCMC chain was run for 500 million generations and sampled every 5,000th generation and a burnin of 2,500 was applied. The GMYC analysis was performed with the splits package 1.0-19 (Ezard *et al.* 2009) in R v3.1.0 (R Development Core Team, 2008).

Molecular divergence times of major lineages were estimated with a molecular clock analysis in BEAST v1.8.0 (Drummond *et al.* 2012) based on a 709 bp alignment of 120 individuals from this study. In addition, a second molecular clock analysis was performed for a *COI* alignment that included 123 individuals of lineages L0-L3 of *P. notabilis* from(Porco *et al.* 2012) and all individuals from this study. Strict molecular clock analyses were conducted for the two *COI* nucleotide alignments using the Yule process as tree prior (Gernhard *et al.* 2008) with a fixed substitution rate of 0.0115 which corresponds to the common invertebrate rate of *COI* of 0.023 substitutions per site per million years (Avise 1994; Brower 1994). The Yule process allows higher rate variance among branches, which appeared to be more appropriate for this parthenogenetic

species, as preliminary analyses showed faster convergence and better likelihoods compared to analyses using coalescent tree priors. Convergence of the MCMC chain after 600 million generations (sampled every 60,000th generation) with a burnin of 2,500 was confirmed using Tracer v1.4 (Drummond & Rambaut 2007).

The number of nucleotide and protein haplotypes was determined using the online tool FaBox v1.41 (Villesen 2007). Haplotype alignments of the amino acid sequences for *COI* and *H3* and the nucleotide sequences of 28S were generated with FaBox v1.41, each gene was checked by eye for lineage specific substitutions. Lineage assignments corresponded to the IEU estimated by the GMYC analysis. Analysis of molecular variance (AMOVA) and genetic distance analyses (uncorrected p-distances) were performed separately for sampling locations and lineages and each gene in ARLEQUIN v3.5 (Excoffier & Lischer 2010) with 20,000 permutations.

Results

Phylogeny and independent evolutionary units

The phylogenetic tree based on the combined matrix of the three genes gave the best resolution and statistical support for internal and terminal nodes (**Fig. 2A**); topologies of ML and BI trees were very similar. The phylogenetic trees based on the combined matrix and the individual genes were always congruent i.e., *COI* haplotypes were linked with specific *H3* and 28S alleles.

The GMYC analysis estimated five IEU for the *COI* tree (**Fig. 2B**) with a threshold of <1% genetic distances as cut off. The number of ML entities was 7 with entities number 6 and 7 being outgroups. The likelihood ratio between the null model and the GMYC model was 21.05 ($p<0.001$). Branching patterns in the phylogenetic analyses with the single genes *H3* and 28S (**Fig. S1, Supporting information**) and the combined dataset (**Fig. 2A**) were consistent with these IEUs and corresponded to highly supported clades (posterior probabilities=1, bootstrap=97-100) in all four datasets. Results of the 4x rule, however, indicated that all IEU belong to a single species (**Table S2, Supporting information**).

All populations clustered with L0-3 from (Porco *et al.* 2012) and were named accordingly, except for the population from Greece which was not identical, but formed a

sister taxon to L3 from Paris, and the population from Croatia (L4) which differed considerably from all other lineages. Notably, the *COI* sequences of two individuals of L1 (Porco *et al.* 2012) clustered within L2 **(Fig. S2, Supporting information)**.

L0 comprised only three sampling locations in central and northern Europe, whereas L1 and L2 comprised large phylogeographic clades, covering wide regions in south and central Europe to the east (L1) and in western Europe from south to north (L2). Different lineages coexisted only in two sampling locations, in Denmark and Germany (Uelzen); however, each of the three German sampling locations was occupied by different lineages. The four French sampling locations comprised only two different lineages, with L1 being present only in southern France (FR2 and FR5) and L2 being present in all other French sampling locations. Further, two different lineages were sampled in Scotland; L0 was present in the east of Scotland (GB1) while the widespread L2 was present in the west of Scotland (GB2) (**Figs. 1, 2**).

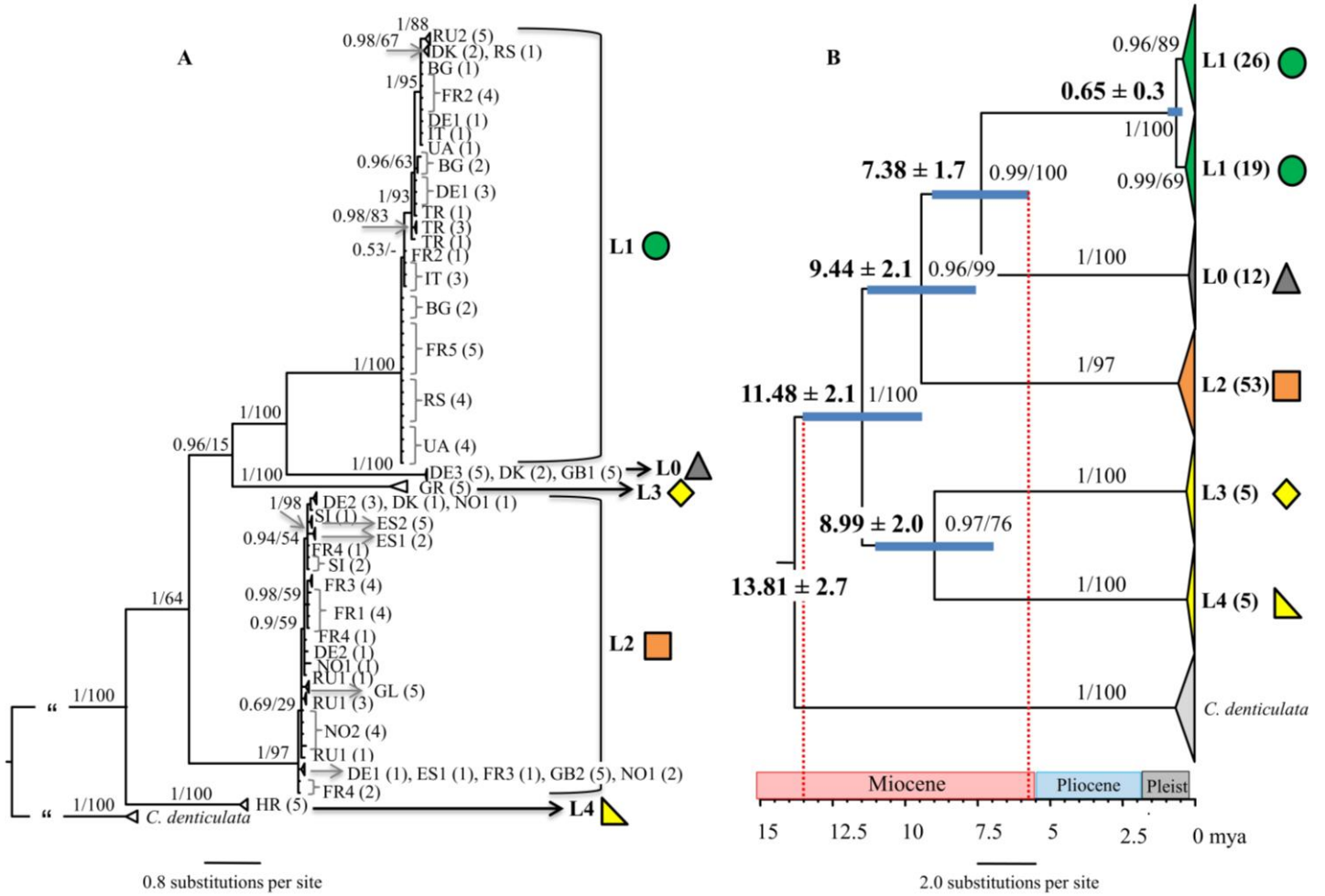


Figure 2. Phylogenetic relationships and molecular divergence estimates of five European lineages of *Parisotoma notabilis*. **A** Bayesian phylogeny based on nucleotide sequence data of three genes (28Sr DNA, *H3* and *COI*). **B** Molecular clock tree based on *COI* sequences (709 bp) calculated with BEAST. Blue bars on nodes represent 95% confidence intervals. Bold numbers on nodes are divergence times with confidence intervals in mya. A geological timescale is provided below the tree, indicating the Miocene (23-5.3 mya), the Pliocene (5.3-1.8 mya) and the Pleistocene (Pleist. 1.8-0.0114 mya). Dashed red lines indicate radiations of major lineages during the Miocene. Terminal clades are collapsed; numbers in brackets indicate the numbers of individuals of each sampling locations included in the respective clades. Numbers on nodes are posterior probabilities (Bayesian Inference) and bootstrap values (Maximum Likelihood).

Genetic diversity

Genetic distances between the lineages 0 to 4 (**Table 2A**) were very high, ranging between 15% and 18% for the *COI* gene, between 5% and 11% for *H3* and 0.5% to 1.9% for 28S. Genetic distances between sampling locations were also high, ranging for *COI* between 9 and 17%, for *H3* between 4 and 7% and for 28S between 0.7 and 1% (**Table 2B**). Within sampling locations genetic distances were generally low or non-existing, except for the population from Denmark with 13% (4% in *H3* and 1% in 28S) and Germany (DE1) with 7% distance in *COI* (3% in *H3* and 0.6% in 28S), the only locations with coexisting lineages. Accordingly, molecular variance was very high between locations with 90% variance for *COI*, 92% for *H3* and 92 % for 28S and low within locations with 10% variance for *COI*, 8% for *H3* and 8 % for 28S (**Table 3**).

Haplotype diversity of the mitochondrial gene was moderately high. The 120 sequenced individuals of *P. notabilis* separated for *COI* into 39 nucleotide and 18 amino acid haplotypes, the nuclear genes had twelve nucleotide and two amino acid haplotypes of *H3* and five nucleotide haplotypes of 28S. The amino acid haplotypes of *COI* had at least one non-synonymous and lineage specific substitution in each lineage **(Table S3, Supporting information)**. L1 and L2 were more variable, separating into sublineages, i.e. into regional lineages with different amino acids only being present in Bulgaria, Germany (Uelzen), south and central France (FR1-2, FR4), near Moscow (RU2) and Turkey. The *H3* gene had two non-synonymous and lineage specific substitutions between L4 and all other lineages and several lineage specific nucleotide substitutions. Several nucleotide substitutions in 28S were lineage specific and affected all individuals of a lineage **(Table S3, Supporting information)**.

Table 2. Mean genetic distances (observed p-distances in %) of *Parisotoma notabilis* in Europe analyzed for three genes analyzed in this study. (A) between genetic lineages (L0-L4) and (B) between and within locations of *Parisotoma notabilis* in Europe for three genes analyzed in this study. Genetic lineages refer to independent evolutionary units (IEUs) identified with GMYC.

A	***COI***					***H3***					**28S**				
Lineage	**L0**	**L1**	**L2**	**L3**	**L4**	**L0**	**L1**	**L2**	**L3**	**L4**	**L0**	**L1**	**L2**	**L3**	**L4**
Lineage 0	0					0					0				
Lineage 1	15.8	0.66				4.55	0.17				1.22	0			
Lineage 2	16.64	15.94	0.48			6.42	6.68	0.05			1.22	1.39	0		
Lineage 3	18.34	17.49	17.35	0.34		6.95	6.15	6.95	0.96		0.7	0.87	0.52	0	
Lineage 4	17.21	17.91	15.23	16.08	0.28	7.22	6.95	8.02	10.7	0	1.74	1.92	1.22	1.05	0

B			**between locations**			**within locations**		
	location		***COI***	***H3***	**28S**	***COI***	***H3***	**28S**
north	DK	**Denmark**	12.21 ± 2	4.38 ± 1	0.94 ± 0	13.17	4.28	1.01
	GB1	**Great Britain**	15.52 ± 4	5.15 ± 2	1.15 ± 0	0	0	0
	GB2		8.95 ± 8	3.94 ± 3	0.70 ± 1	0	0	0
	GL	**Greenland**	8.96 ± 8	4.19 ± 3	0.70 ± 1	0.16	0	0
	NO1	**Norway**	9.01 ± 8	3.94 ± 3	0.70 ± 1	0.73	0	0
	NO2		8.82 ± 8	3.94 ± 3	0.70 ± 1	0.14	0	0
	RU1	**Russia**	8.91 ± 8	5.15 ± 2	0.70 ± 1	0.2	0	0
	RU2		11.43 ± 7	4.42 ± 3	0.92 ± 1	0	0.16	0
central	FR1	**France**	9.01 ± 8	3.94 ± 3	0.70 ± 1	0.07	0	0
	FR2		11.45 ± 7	4.30 ± 3	0.92 ± 1	0.45	0	0
	FR3		8.85 ± 8	3.94 ± 3	0.70 ± 1	0.34	0	0
	FR4		9.02 ± 8	3.94 ± 3	0.70 ± 1	0	0	0
	DE1	**Germany**	10.53 ± 5	4.09 ± 2	0.86 ± 0	6.54	2.67	0.56
	DE2		9.14 ± 8	3.94 ± 3	0.70 ± 1	0.35	0	0
	DE3		15.52 ± 4	5.15 ± 2	1.15 ± 0	0	0	0
south	BG	**Bulgaria**	11.00 ± 7	4.23 ± 3	0.92 ± 1	0.73	0	0
	HR	**Croatia**	12.31 ± 2	4.42 ± 1	0.94 ± 0	0.28	0	0
	FR5	**France**	10.74 ± 7	4.23 ± 3	0.92 ± 1	0	0	0
	GR	**Greece**	17.41 ± 0	6.74 ± 1	0.69 ± 0	0.34	0.96	0
	IT	**Italy**	10.93 ± 7	4.23 ± 3	0.92 ± 1	0.56	0	0
	RS	**Serbia**	10.86 ± 7	4.22 ± 3	0.92 ± 1	0.51	0.11	0
	SI	**Slovenia**	8.99 ± 8	3.94 ± 3	0.70 ± 1	0.09	0	0
	ES1	**Spain**	9.19 ± 8	3.94 ± 3	0.70 ± 1	0.66	0	0
	ES2		9.08 ± 8	3.94 ± 3	0.70 ± 1	0	0	0
	TR	**Turkey**	11.09 ± 7	4.23 ± 3	0.92 ± 1	0.2	0	0
	UA	**Ukraine**	10.86 ± 7	4.23 ± 3	0.92 ± 1	0.51	0	0

Table 3. Analysis of molecular variance (AMOVA) among and within sampling locations of the *Parisotoma notabilis* sampled across Europe, based on sequence variance of three genes. Asterisks indicate significant differences at p<0.05; df, degrees of freedom; Fst, F-statistics.

	COI		***H3***		**28S**	
source of variation	**between locations**	**within locations**	**between locations**	**within locations**	**between locations**	**within locations**
df	25	94	25	94	25	94
sum of squares	4227.64	355.88	917.64	61.2	269.93	18
variance components	35.85 Va***	3.79 Vb***	7.82 Va***	0.65Vb***	2.30 Va***	0.19 Vb***
% variation	90.45	9.55	92.31	7.69	92.31	7.69
fixiation indices	Fst 0.904***		Fst 0.923***		Fst 0.923***	

Molecular divergence times

Estimates of divergence times refer to the *COI* dataset including all sequences of this study (**Fig. 2B**) because this alignment contained more informative sites compared to the combined alignment with sequences of (Porco *et al.* 2012). According to a constant substitution rate of 2.3% per million years (Avise 1994; Brower 1994), the five lineages diverged in the Miocene, 11.5-7.4 mya. The widespread south European L1 and the locally occurring L0 from central Europe were the youngest, with a Late Miocene origin about 7.4 ± 2.0 mya, and the western European L2 diverged from L0 about 9.4 ± 2.0 mya. At about the same time, 9.0 ± 2.0 mya, the two isolated south European lineages L3 (Greece) and L4 (Croatia) separated from a common ancestor. Despite the ancient diversifications of lineages, populations at sampling locations were rather young, between 0.1 and 0.023 my old, and of Pleistocene origin.

Molecular clock analyses based on the combined dataset of the present study and (Porco *et al.* 2012) differed from the molecular clock estimates above in having lower node supports of the phylogenetic groups, a lower resolution of the phylogenetic backbone and lineages separated much earlier **(Fig. S2, Supporting Information)**. However, radiations of lineages also occurred in the Miocene (17.8-13.0 mya), but were considerably older than in the analyses with fewer taxa and longer sequences (11.5-7.4 mya).

Discussion

This study analyzed the phylogeographic structure of *P. notabilis* in Europe, one of the most widespread and abundant species of Collembola. The sampling region covered southern, central and northern Europe from east (Ukraine) to west (Pyrenees). The GMYC analysis estimated five IEU (L0-L4) for the *COI* tree. Branching patterns in the phylogenetic analyses with the single genes *H3* and 28S and the combined dataset were consistent with these IEUs. L0 comprised only three sampling locations in central and northern Europe, whereas L1 and L2 comprised large phylogeographic clades covering wide regions from south to east Europe (L1) and western Europe from south to north (L2). The results suggest that local populations predominantly comprise of genetically uniform

populations; only in two sampling locations, in Denmark and in Uelzen (Germany) different lineages coexisted. The other two German populations comprised only a single lineage. Similarly, the four French sampling locations comprised only two different lineages, with L1 being present only in southern France (FR2 and FR5) and L2 being present in all other French sampling locations. Further, only two different lineages were sampled in Scotland; L0 was present in the east of Scotland (GB1) while the widespread L2 was present in the west of Scotland (GB2). In addition, in southern Europe (Greece and Croatia) we found two isolated lineages, i.e. L3 and L4, which were separated from each other. However, the analysis with the 4x rule did not support the distinction of genetic species among the five IEUs of *P. notabilis*. The method identifies species based on genetic distances and uses a more flexible species threshold than the GMYC method. This adds to other studies indicating that intraspecific genetic variance of *COI* is high in Collembola species (Shaw *et al.* 2013) and other species of soil invertebrates with intra-specific sequence divergences ranging between 11% and 32% (Heethoff *et al.* 2007; Boyer *et al.* 2007; Torricelli *et al.* 2010; Porco *et al.* 2012; Rosenberger *et al.* 2013).

Lineage L0, likely representing the type-species lineage, only occurred in three of our sampling locations and was restricted to the central European regions of southern UK, The Netherlands (Porco *et al.* 2012), central Germany and Denmark. Its distribution range strongly overlaps with the two other common lineages L1 and L2, but notably, most of the sampling locations of L0 are in the vicinity of The Channel and the east-coast of Great Britain suggesting that this lineage benefitted from dispersal along trading routes.

Results of this study indicate that the southern European lineage L1 of (Porco *et al.* 2012) is also common in the east of Europe (Russia2, Ukraine and Turkey), and lineage L2, assumed to be restricted to the Alpine-Carpathian mountain ranges by (Porco *et al.* 2012), in fact is widely distributed in western and northern Europe (France, Denmark, Norway, Greenland, west of Scotland) and the Pyrenees. Interestingly, these two widespread lineages are parapatric, suggesting the presence of local ecomorphs that outcompete other lineages within their respective distribution ranges. Further, additional to the previously described lineage L3 that only occurred in Paris (France) (Porco *et al.* 2012), we identified a related population to this lineage in southern Europe (Greece) and a new

lineage from a single location in Croatia. These lineages add to the cryptic genetic diversity of *P. notabilis* in Europe.

Each of the five lineages of *P. notabilis* identified in this study had specific, non-synonymous substitutions in the *COI* gene and coexistence of lineages was rare, indicating selection and subsequent spread of the most competitive genotype. However, all lineage specific substitutions in the *H3* gene were synonymous, suggesting that fixation of alleles due to bottlenecks and founder events also contributed to the high genetic structure between populations. Interestingly, lineages L0-L2 co-occurred in a small region in Canada (Porco *et al.* 2012), either due to multiple, independent anthropogenic introductions and/or due to different ecological conditions, allowing coexistence of lineages. This suggests that in Canada, compared to Europe, populations are more dynamic either because competitive exclusion among haplotypes is in progress due to recent establishment of *P. notabilis* or because strong (abiotic) disturbances structure populations facilitating maintenance of genetic variance within sampling locations.

The molecular divergence estimates indicated three radiation events. In the Late Miocene, the separation of the endemic southern lineages L3 and L4, the widespread lineages L0-L2 (11.5 mya) and the divergences into the five IEUs (L0-L4) (9-7.4 mya) occurred. Much later separation of lineage L1 into two sublineages occurred in the Pleistocene (0.65 mya). These radiation events coincide with climatic and biotic changes in Europe, i.e. changes from warm and wet climate during the Miocene and the extension of grassland together with the establishment of deciduous forests in Europe in the Late Miocene and Pliocene (Davis 1983; Retallack 2001; Böhme 2003; Osborne & Beerling 2006; Bruch *et al.* 2007; Böhme *et al.* 2008; Jiménez-Moreno *et al.* 2010; Pound *et al.* 2012). This suggests that *P. notabilis* benefitted from colder climatic conditions in the Late Miocene and associated changes in vegetation from grassland to woodland which allowed to expand its range size considerably.

Despite old diversifications of lineages L0-L3, their distribution ranges differ, with L0 being restricted to northern Europe at the eastern arm of the English Channel in areas with mild climate influenced by the Golf Stream and human trading. In contrast, the distribution range of L1 concentrates in the Mediterranean extending from Spain (Pyrenees) to Algeria to Turkey (Porco *et al.* 2012), another area influenced by trading.

Except for two locations in central and northern Europe, the most northern representatives of L1 occurred in the south of France and Russia (RU2), but the main distribution range remains south of the Alps. Lineage L2 is widely distributed in central and northern Europe and in the mountain ranges of the Pyrenees, Alps and Carpathians (Porco *et al.* 2012), but does not extend to the Mediterranean. Interestingly, both lineages L1 and L2 are present in the Pyrenees, indicating either the importance of northern Spain as refuge area during the Last Ice Age or the existence of a contact zone of two lineages (Hewitt 1999, 2001; Kutnik *et al.* 2004; Tastard *et al.* 2012; Milá *et al.* 2013; Salicini *et al.* 2013; Collins & Rawlins 2013; Patel *et al.* 2015) with otherwise distinct distribution ranges.

Despite the wide ranges of lineages, coexistence of lineages was very restricted; only five locations in Europe were colonized by two or more lineages of *P. notabilis*; two sites in France (Paris and Le Port, Ariege), one in Spain (Gerona, Catalonia) (Porco *et al.* 2012), one in Germany (Uelzen) and one in Denmark (Humblebaek). Notably, the sampling sites were close to urban areas suggesting that anthropogenic transport and disturbance favor coexistence of lineages. Generally, *P. notabilis* occurs in anthropogenic and disturbed habitats, suggesting a synanthropic distribution, i.e. passive dispersal by humans and establishment in human associated agricultural or managed systems. Therefore, the distribution ranges of lineages of *P. notabilis* may reflect early human trading and migration routes similar to neolithic distributions of the Pygmy shrew between Britain and Ireland (McDevitt *et al.* 2011) and the snail *Capaea nemoralis* with a strong Franco-Iberian distribution of haplotypes (Grindon & Davison 2013). This holds in particular for lineage L2, which is most variable in *COI*, indicating that selection and lineage sorting is still in progress. However, if human transport or local adaptations shaped the distribution ranges of the three widespread lineages of *P. notabilis* can only be resolved by using more variable markers such as microsatellites or SNPs combined with experiments testing the fitness of the different lineages under varying environmental conditions.

Overall, results of the present study show that the ubiquitous Collembola species *P. notabilis* comprises several genetic lineages with distinct distribution ranges, with all lineages likely forming part of a single species as indicated by the 4x rule. Anthropogenic

activities rather than climatic changes during the Last Ice Age likely shaped the present day genetic structure of this species in Europe. Compared to other species of Collembola (H. von Saltzwedel*, unpubl. data*), European lineages of *P. notabilis* are rather young, genetically uniform and depauperate. The strong human association of this species likely enabled rapid colonization of areas connected by human trading and migration. Human dispersal of few or single individuals resulted in founder effects and the establishment of genetically homogenous lineages. Thereby, *P. notabilis* is an interesting model organism to investigate population dynamics, adaptations and fitness differences among IEUs.

Further, its cosmopolitan distribution enables to compare these processes in independent geographic regions including Europe and North America, as the same lineages occur on both continents but gene flow is impeded between the continents. In contrast to many soil-living organisms, *P. notabilis* is easy to culture and has short generation times, making this species an ideal model organism for studying evolutionary processes and population genetics of soil invertebrates in both the field and laboratory.

Acknowledgements

We thank Patrick Pachl for help with sampling of the Balkan region. We are grateful to Professor Nusret Ayyildiz, Jens Bast, Elena Corral Hernández, Guido Humpert, Sven Marhan, Jordi Moya-Laraño, Lucija Šerić Jelaska and Annegret and Winfried Treptow for sending litter and soil material.

This study was funded by the Deutsche Forschungsgemeinschaft (SCHE 376/13-1 u. 13-2).

References

Avise JC (1994) *Molecular Markers, Natural History and Evolution* (JC Avise, Ed,). Chapman & Hall, New York.

Bell G (1982) *The Masterpiece of Nature: The Evolution and Genetics of Sexuality*. (G Bell, Ed,). University of California Press, Berkeley.

Birky CW (2013) Species Detection and Identification in Sexual Organisms Using Population Genetic Theory and DNA Sequences. *PLoS One*, **8**.

Birky CW, Wolf C, Maughan H *et al.* (2005) Speciation and Selection without Sex. *Hydrobiologia*, **546**, 29–45.

Böhme M (2003) The Miocene Climatic Optimum: evidence from ectothermic vertebrates of Central Europe. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **195**, 389–401.

Böhme M, Ilg A, Winklhofer M (2008) Late Miocene "washhouse" climate in Europe. *Earth and Planetary Science Letters*, **275**, 393–401.

Boyer SL, Baker JM, Giribet G (2007) Deep genetic divergences in *Aoraki denticulata* (Arachnida, Opiliones, Cyphophthalmi): a widespread "mite harvestman" defies DNA taxonomy. *Molecular Ecology*, **16**, 4999–5016.

Brower AV (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 6491–6495.

Bruch AA, Uhl D, Mosbrugger V (2007) Miocene climate in Europe — Patterns and evolution. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **253**, 1–7.

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

Chahartaghi M, Scheu S, Ruess L (2006) Sex ratio and mode of reproduction in Collembola of an oak-beech forest. *Pedobiologia*, **50**, 331–340.

Colgan DJ, McLauchlan A, Wilson GDF *et al.* (1998) Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. *Australian Journal of Zoology*, **46**, 419.

Collins MM, Rawlins JE (2013) A Transect for Reproductive Compatibility and Evidence for a "Hybrid Sink" in a Hybrid Zone of *Hyalophora* (Insecta: Lepidoptera: Saturniidae). *Annals of Carnegie Museum*, **82**, 193–223.

Davis M (1983) Quaternary History of Deciduous Forests of Eastern North America and Europe. *Annals of the Missouri Botanical Garden*, **70**, 550–563.

Deharveng L (1981) Collemboles des îles subantarctiques de l'Océan Indien: mission J. Travé 1972-1973. *CNFRA Biologie des Sols*, **48**, 33–108.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Drummond AJ, Suchard M a, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 3–6.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Ezard T, Fujisawa T, Barraclough TG (2009) SPecies LImits by Threshold Statistics (http://r-forge.r-project.org/projects/splits/).

Fiera C, Ulrich W (2012) Spatial patterns in the distribution of European springtails (Hexapoda: Collembola). *Biological Journal of the Linnean Society*, **105**, 498–506.

Fjellberg A (1977) On the identity of *Isotoma ekmani* nom. nov. pro *I. pallida* Agrell, 1939 (nec Nicolet, 1842, Moniez, 1894) (Collembola: Isotomidae). *Insect Systematics & Evolution*, **8**, 9–11.

Fjellberg A (1980) *Identification Keys to Norwegian Collembola* (A Fjellberg, Ed,). Norsk Entomologisk Forening, Cornell University.

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

Fontaneto D, Herniou E a, Boschetti C *et al.* (2007) Independently evolving species in asexual bdelloid rotifers. *PLoS Biology*, **5**, e87.

Fountain MT, Hopkin SP (2004) Biodiversity of Collembola in urban soils and the use of *Folsomia candida* to assess soil "quality". *Ecotoxicology*, **13**, 555–572.

Fujisawa T, Barraclough TG (2013) Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation on simulated data sets. *Systematic biology*, **62**, 707–724.

García-Gómez A, Castaño-Meneses G, Palacios-Vargas JG (2009) Diversity of springtails (Hexapoda) according to a altitudinal gradient. *Pesq Agropec Bras*, **44**, 911–916.

Gernhard T, Hartmann K, Steel M (2008) Stochastic properties of generalised Yule models, with biodiversity applications. *Journal of Mathematical Biology*, **57**, 713–735.

Grindon AJ, Davison A (2013) Irish *Cepaea nemoralis* Land Snails Have a Cryptic Franco-Iberian Origin That Is Most Easily Explained by the Movements of Mesolithic Humans. *PloS One*, **8**, e65792.

Hågvar S (2010) Primary Succession of Springtails (Collembola) in a Norwegian Glacier Foreland. *Arctic, Antarctic, and Alpine Research*, **42**, 422–429.

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

Heethoff M, Domes K, Laumann M *et al.* (2007) High genetic divergences indicate ancient separation of parthenogenetic lineages of the oribatid mite *Platynothrus peltifer* (Acari, Oribatida). *Journal of Evolutionary Biology*, **20**, 392–402.

Hewitt G (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hewitt GM (2001) Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Molecular ecology*, **10**, 537–549.

Hewitt GM, Ibrahim KM (2001) Inferring glacial refugia and historical migrations with molecular phylogenies. In: *Integrating Ecology and Evolution in a Spatial Context* (eds Silvertown J, Antonovics J), pp. 271–294. Blackwell Science Ltd., Oxford.

Hopkin SP (2007) *A Key to the Collembola (Springtails) of Britain and Ireland*. Field Studies Council.

Ingimarsdóttir M, Caruso T, Ripa J *et al.* (2012) Primary assembly of soil communities: disentangling the effect of dispersal and local environment. *Oecologia*, **170**, 745–754.

Jiménez-Moreno G, Fauquette S, Suc JP (2010) Miocene to Pliocene vegetation reconstruction and climate estimates in the Iberian Peninsula from pollen data. *Review of Palaeobotany and Palynology*, **162**, 403–415.

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

Kováč L, Kostúrová N, Miklisová D (2005) Comparison of collembolan assemblages (Hexapoda, Collembola) of thermophilous oak woods and *Pinus nigra* plantations in the Slovak Karst (Slovakia). *Pedobiologia*, **49**, 29–40.

Kutnik M, Uva P, Brinkworth L *et al.* (2004) Phylogeography of two European Reticulitermes (Isoptera) species: the Iberian refugium. *Molecular ecology*, **13**, 3099–3113.

Kuznetsova NA (2006) Long-term dynamics of Collembola in two contrasting ecosystems. *Pedobiologia*, **50**, 157–164.

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

McDevitt a D, Vega R, Rambau R V *et al.* (2011) Colonization of Ireland: revisiting "the pygmy shrew syndrome" using mitochondrial, Y chromosomal and microsatellite markers. *Heredity*, **107**, 548–557.

Milá B, Surget-Groba Y, Heulin B, Gosá A, Fitze PS (2013) Multilocus phylogeography of the common lizard *Zootoca vivipara* at the Ibero-Pyrenean suture zone reveals lowland barriers and high-elevation introgression. *BMC Evolutionary Biology*, **13**, 192.

Milne I, Lindner D, Bayer M *et al.* (2009) TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics*, **25**, 126–127.

Osborne CP, Beerling DJ (2006) Nature's green revolution: the remarkable evolutionary rise of C4 plants. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*, **361**, 173–194.

Patel S, Schell T, Eifert C *et al.* (2015) Characterizing a hybrid zone between a cryptic species pair of freshwater snails. *Molecular Ecology*, **24**, 643–655.

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.

Pons J, Barraclough T, Gomez-Zurita J *et al.* (2006) Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects. *Systematic Biology*, **55**, 595–609.

Porco D, Potapov M, Bedos A *et al.* (2012) Cryptic Diversity in the Ubiquist Species *Parisotoma notabilis* (Collembola, Isotomidae): A Long-Used Chimeric Species? *PLoS One*, **7**, e46056.

Potapov M (1991) Species of the genus Isotoma subgenus Parisotoma Bagnall, 1940 and *Sericeotoma* subgen. nov. (Collembola, Isotomidae) of USSR fauna. *Acta zoologica cracoviensia*, **34**, 267–301.

Pound MJ, Haywood AM, Salzmann U *et al.* (2012) Global vegetation dynamics and latitudinal temperature gradients during the Mid to Late Miocene (15.97–5.33Ma). *Earth-Science Reviews*, **112**, 1–22.

Retallack GJ (2001) Cenozoic expansion of grasslands and climatic cooling. *The Journal of Geology*, **109**, 407 – 426.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Rosenberger M, Maraun M, Scheu S *et al.* (2013) Pre- and post-glacial diversifications shape genetic complexity of soil-living microarthropod species. *Pedobiologia*, **56**, 79–87.

Rusek J (1984) New species and review of the *Isotoma notabilis* species-group (Collembola, Isotomidae). *Acta Entomologica Bohemoslovaca*, **81**, 343–369.

Salamon J-A, Alphei J (2009) The Collembola community of a Central European forest: Influence of tree species composition. *European Journal of Soil Biology*, **45**, 199–206.

Salicini I, Ibáñez C, Juste J (2013) Deep differentiation between and within Mediterranean glacial refugia in a flying mammal, the *Myotis nattereri* bat complex. *Journal of Biogeography*, **40**, 1182–1193.

Von Saltzwedel H, Scheu S, Schaefer I Founder events and pre-glacial divergences shape the genetic structure of European Collembola species. *in Preparation*.

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

Shaw P, Faria C, Emerson B (2013) Updating taxonomic biogeography in the light of new methods – examples from Collembola. *Soil Organisms*, **85**, 161–170.

Sommer RS, Benecke N (2005) The recolonization of Europe by brown bears *Ursus arctos* Linnaeus, 1758 after the Last Glacial Maximum. *Mammal Review*, **35**, 156–164.

Soto-Adames FN (2002) Molecular phylogeny of the Puerto Rican *Lepidocyrtus* and *Pseudosinella*. *Molecular Phylogenetics and Evolution*, **25**, 27–42.

Stamatakis a, Ott M, Ludwig T (2005) *Parallel Computing Technologies* (V Malyshkin, Ed,). Springer Berlin Heidelberg, Berlin, Heidelberg.

Tastard E, Ferdy JB, Burrus M *et al.* (2012) Patterns of floral colour neighbourhood and their effects on female reproductive success in an Antirrhinum hybrid zone. *Journal of Evolutionary Biology*, **25**, 388–399.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680.

Torricelli G, Carapelli A, Convey P *et al.* (2010) High divergence across the whole mitochondrial genome in the "pan-Antarctic" springtail *Friesea grisea*: Evidence for cryptic species? *Gene*, **449**, 30–40.

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

Villesen P (2007) FaBox: an online toolbox for fasta sequences. *Molecular Ecology Notes*, **7**, 965–968.

Wanner M, Dunger W (2002) Primary immigration and succession of soil organisms on reclaimed opencast coal mining areas in eastern Germany. *European Journal of Soil Biology*, **38**, 137–143.

Williams G (1975) *Sex and Evolution* (G Williams, Ed,). Princeton University Press, Princeton.

Data Accessibility

DNA sequences: GenBank accessions see S1 Table.

Supplementary material

Table S1A. Accessionnumbers of DNA sequences of *Parisotoma notabilis* from Europe obtained in this study. All sequences are available at NCBI GenBank.Countries, sampling locations and sampling coordinates are listed.

country	location	coordinates (N, E)	abbre-viation	*COI*	*H3*	28S
Bulgaria	Bosnek	42.50° 23.17°	BG	KJ792225-9	KJ792345-9	KJ792105-9
Gemany	Uelzen	52.93° 10.61°	DE1	KJ792263-7	KJ792383-7	KJ792143-7
	Norden	53.58° 7.24°	DE2	KJ792268-71	KJ792388-91	KJ792148-51
	Solling, Neuhaus	51.71° 9.64°	DE3	KJ792272-6	KJ792392-6	KJ792152-6
Denmark	Humlebaek	52.93° 10.61°	DK	KJ792235-9	KJ792355-9	KJ792115-9
Spain	Oviedo	43.36° -6.00°	ES1	KJ792324-6	KJ792444-6	KJ792204-6
	Ponga	43.19° -5.16	ES2	KJ792327-31	KJ79247-51	KJ792207-11
France	Voegtlinshoffen	48.02° 7.27°	FR1	KJ792240-3	KJ792360-3	KJ792120-3
	Salavas	44.39° 4.37	FR2	KJ792249-53	KJ792369-73	KJ792129-33
	Chartreuse	45.42° 5.81°	FR3	KJ792254-8	KJ792374-8	KJ792134-8
	Rambouillet	48.62° 1.86°	FR4	KJ792259-62	KJ792379-82	KJ792139-42
	Korsika, Olmi-Capella	42.52° 9.02°	FR5	KJ792244-8	KJ792364-8	KJ792124-8
Great Britain	Melrose	53.58° 7.24°	GB1	KJ792309-13	KJ792429-33	KJ792189-93
	Island of Arran	51.71° 9.64°	GB2	KJ792314-8	KJ792434-8	KJ792194-8
Greenland	Kobbefjord / Nuuk	64.16° -51.52°	GL	KJ792277-81	KJ792397-401	KJ792157-61
Greece	Chrysovitsi	37.56° 22.20°	GR	KJ792282-6	KJ792402-6	KJ792162-6
Croatia	Sljeme	45.90° 15.95°	HR	KJ792230-4	KJ792350-4	KJ792110-4
Italy	Felitto	40.37° 15.22°	IT	KJ792287-90	KJ792407-10	KJ792167-70
Norway	Rod	59.07° 10.23°	NO1	KJ792291-4	KJ792411-4	KJ792171-4
	Skjervenmoen, Fössa Öst	59.45° 10.07°	NO2	KJ792295-8	KJ792415-8	KJ792175-8
Serbia	Sreckovac	43.03° 22.69°	RS	KJ792319-23	KJ792439-43	KJ792199-203
Russia	Petrozavodsk, Karelia	61.77° 34.22°	RU1	KJ792299-303	KJ792419-23	KJ792179-83
	Znamenskoe	55.73° 37.18°	RU2	KJ792304-8	KJ792424-8	KJ792184-8
Slovenia	Postojna	45.76° 14.21°	SI	KJ792332-4	KJ792452-4	KJ792212-4
Turkey	Kayseri	38.67° 35.53°	TR	KJ792335-9	KJ792455-9	KJ792215-9
Ukraine	Kubalach, Crimea	45.00° 34.82°	UA	KJ792340-4	KJ792460-4	KJ792220-4

Table S1B. Assignment of individuals of *Parisotoma notabilis* sampled from seventeen sampling locations in Europe to genetic lineages (L0-L4) and haplotypes for three genes (*COI*, *Histone3*, 28S D3-D5 region).

gene	lineage	haplo-type	no. of sequences	sampling locations	indi-viduals
COI	**L0**	1	12	DE3	1-2
				DK	2,5
				GB1	1-5
	L1	2	14	BG	1
				DE1	2
				FR2	1, 3-5
				IT	2
				RS	1
				RU2	1-5
				UA	2
		3	1	BG	2
		4	5	BG	3
				DE1	1, 4-5
				TR	5
		5	15	BG	4-5
				FR5	1-5
				RS	3-5
				UA	1, 3-5
		6	2	DK	1,4
		7	4	FR2	2
				IT	1, 3-4
		8	3	TR	1, 3-4
		9	1	TR	2
	L2	10	8	DE1	3
				ES1	3
				GB2	1-5
				NO1	2
		11	4	DE2	1,4
				DK	3
				NO1	1
		12	1	DE2	2
		13	1	DE2	3
		14	2	ES1	1-2
		15	5	ES2	1-5
		16	1	DFR1	1
		17	3	FR1	2-4
		18	3	FR3	1
				SI	1,4
		19	2	FR3	2-4
				RU1	2
		20	2	FR3	3-4
		21	1	FR3	5
		22	4	FR4	1-4
		23	4	GL	1, 3-5
		24	1	GL	2
		25	1	NO1	3
		26	1	NO1	4
		27	3	NO2	1-3
		28	1	NO2	4
		29	1	RU1	1
		30	3	RU1	3-5
		31	1	SI	3
	L3	32	1	GR	1
		33	1	GR	2
		34	2	GR	3-4
		35	1	GR	5
	L4	36	2	HR	1,5
		37	1	HR	2
		38	1	HR	3
		39	1	HR	4

Table S1B. *continued*

gene	lineage	haplo-type	no. of sequences	sampling locations	indi-viduals
Histone3	**L0**	1	12	DE3	1-5
				DK	2.5
				GB1	1-5
	L1	2	37	BG	1-5
				DE1	1-2, 4-5
				FR2	1-5
				FR5	1-5
				IT	1-4
				RS	3-5
				TR	1-5
				UA	1-5
		3	3	DK	1.4
				RS	1
		4	2	RU2	1-2
		5	3	RU2	3-5
	L2	6	48	DE1	3-4
				DE2	1-4
				DK	3
				ES1	1-3
				ES2	1-5
				FR1	1-4
				FR3	1-5
				FR4	2-4
				GB2	2-5
				NO1	1-4
				NO2	1-4
				RU1	1-5
				SI	3-4
		7	5	GL	1-5
	L3	8	1	GR	1
		9	1	GR	2
		10	2	GR	3-4
		11	1	GR	5
	L4	12	5	HR	1-5

Table S1B. *continued*

gene	lineage	haplo-type	no. of sequences	sampling locations	indi-viduals
28S D3-5	**L0**	1	12	DE3	1-5
				DK	2.5
				GB1	1-5
	L1	2	45	BG	1-5
				DE1	2, 4-5
				DK	1.4
				FR2	1-5
				FR5	1-5
				IT	1-4
				RS	1-5
				RU2	1-5
				TR	1-5
				UA	1-5
	L2	3	53	DE1	3
				DE2	1-4
				DK3	3
				ES1	1-3
				ES2	1-5
				FR1	1-4
				FR3	1-5
				FR4	1-4
				GB2	1-5
				GL	1-5
				NO1	1-4
				NO2	1-4
				RU1	1-5
				SI	1, 3-4
	L3	4	5	GR	1-5
	L4	5	5	HR	1-5

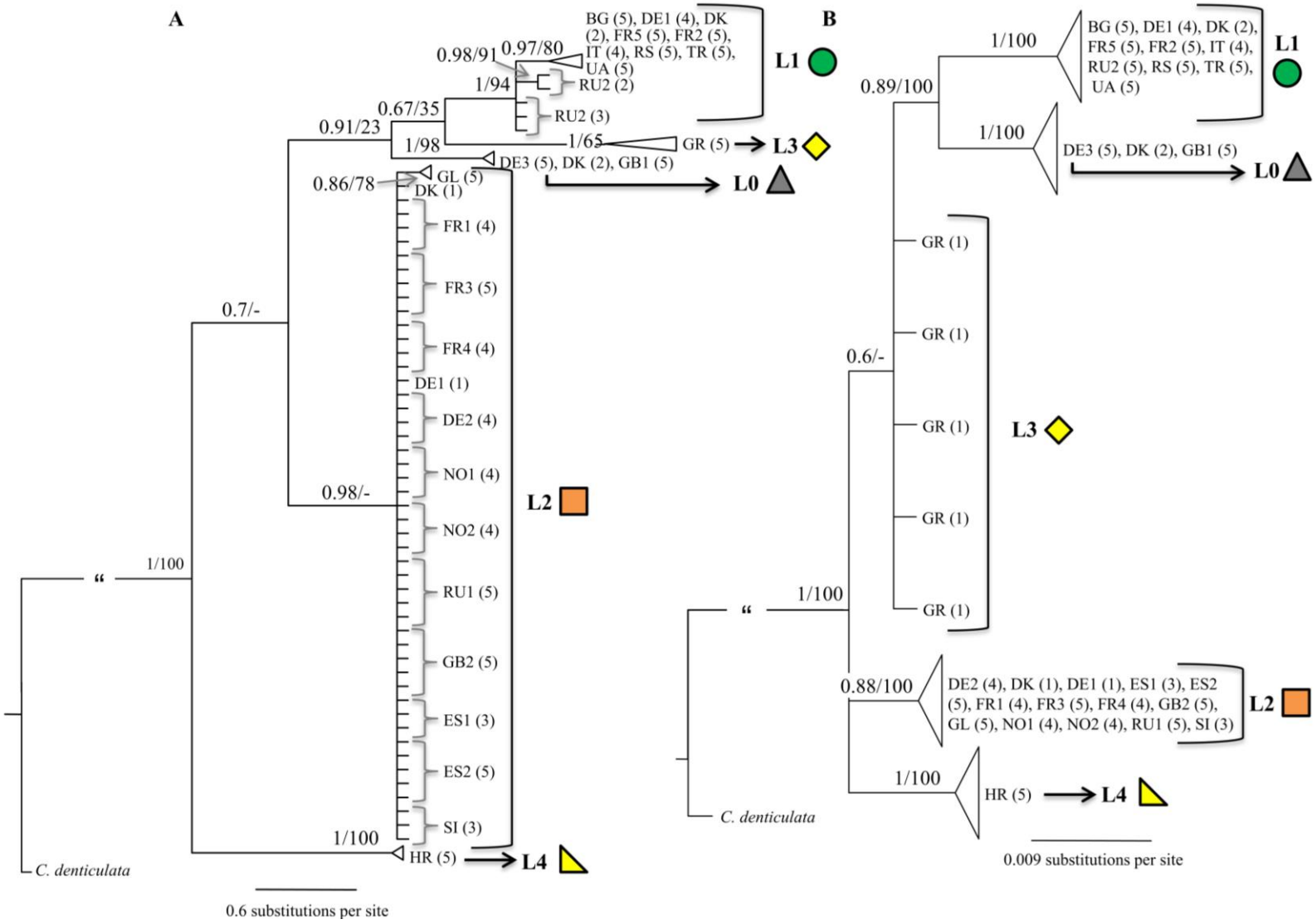


Figure S1. Bayesian phylogeny of European lineages of *Parisotoma notabilis* based on nucleotide sequences of the **A** *H3* gene and **B** 28S rDNA (D3-D5 region). Numbers on nodes are posterior probabilities and bootstrap values.

Table S2A. Values to estimate speciation among parthenogenetic lineages (Birky's 4x rule) for the five genetic lineages of *Parisotoma notabilis* sampled in Europe. Nucleotide diversity (π), pairwise difference between sequences (d), number of individuals (n), sequence length (L) and θ. Calculations are based genetic distances and the phylogeny generated with 120 individuals and a 709 bp fragment of *COI*.

	π	θ	d	n	L
Lineage 0	0.00	0.00	0.00	12	709
Lineage 1.1	2.35	-0.84	2.26	26	709
Lineage 1.2	0.37	-2.32	0.35	19	709
Lineage 2	3.51	-0.81	3.44	53	709
Lineage 3	3.00	-0.82	2.40	5	709
Lineage 4	2.5	-0.83	2.00	5	709
Lineage1	4.78	-0.75	4.67	45	709
Lineage 0-1	41.52	-0.75	40.79	57	709
Lineage 0-2	69.75	-0.75	69.12	110	709
Lineage 3-4	71.58	-0.75	64.42	10	709

Table S2B. K/θ between highly supported clades of *Parisotoma notabilis* from Europe to estimate K for Birky's 4x rule. K/θ ≥ 4 indicitate that samples are from different species and K/θ ≤ 4 indicate that samples are from the same species. K and θ values correspond to Table S2A.

	K/θ	K
L1.1/L1.2	-3.34	7.75
L1.1/L0	1772093.27	113.62
L1.2/L0	-47.33	109.79
L1/L0	1746826.67	112
L0-L1/L2	2.8	114.05
L3/L4	-139.58	114.2
L0-2/L3-4	-151.73	120.66

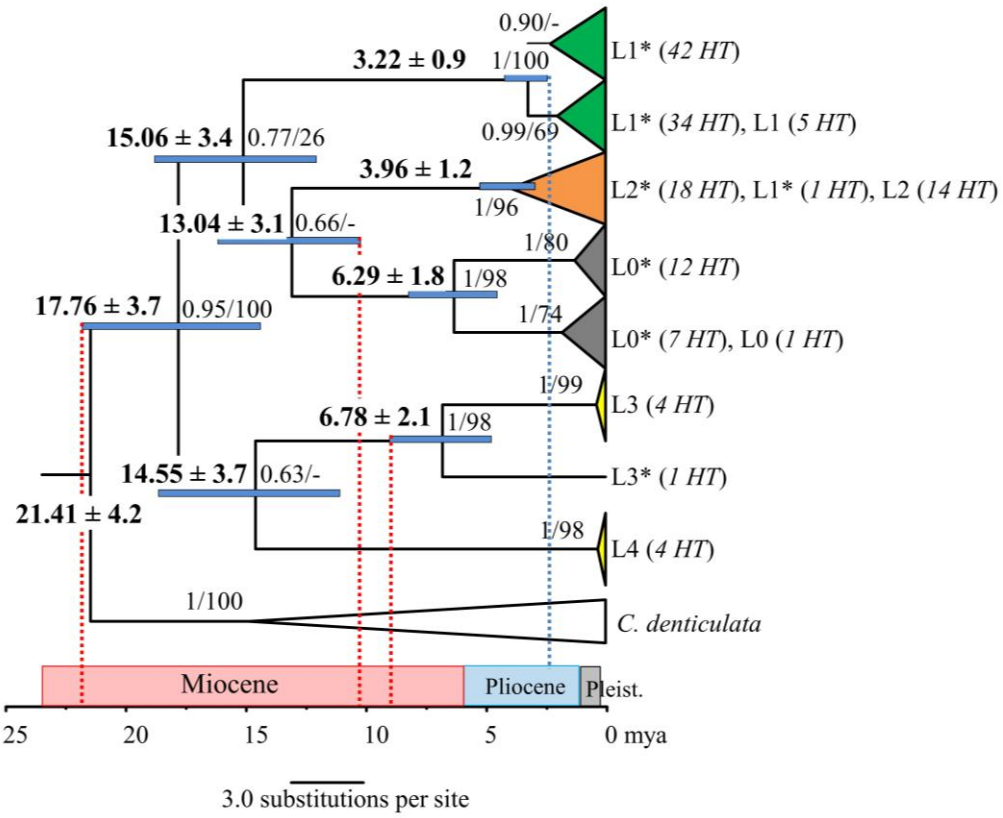


Figure S2. Bayesian trees based on *COI* sequences of *Parisotoma notablilis* from Europe. Molecular divergence estimates of European of *P. notabilis* calculated with BEAST based on a 500 bp alignment including all *COI* sequences from this study (n=120) and from Porco *et al*. (2012) (n=123). Numbers of different haplotypes (HT) are indicated in brackets next to the genetic lineage, * indicate lineages and haplotypes from Porco *et al.* (2012); dashed lines indicate radiations of major lineages in the Miocene (red) and Miocene-Pliocene (blue). Note that divergence estimates are several million years older than in Fig 2B.

Table S3. Positions of lineage specific substitutions in the alignments of three genes (*COI*, *Histone 3*, 28S rDNA) of 120 individuals of *Parisotoma notabilis* sampled across Europe. For *COI*, the amino acid alignment (aa) was analyzed, for *Histone 3* the nucleotide (nct) alignment was investigated as all amino acid sequences were identical. The common characters (amino acid for *COI*, nucleotide for *Histone 3* and 28S rDNA) are listed next to the specific substitution.

COI	**position**	**lineage**	**specific aa**	**common aa**	**specific haplotypes***
	360-362	L0	I	V	
	147-149	L1	A	M	2-4, 6
		L1	T	M	8-9
	498-500	L1	A	V	
	81-83	L2	A	V	
		L2	T	V	16-17, 22
	507-509	L3	K	T	
	48-50	L4	I	V	
Histone H3	**position**	**lineage**	**specific aa**	**common aa**	
	52-54	L4	R	S	
	127-129	L4	R	S	
Histone H3	**position**	**lineage**	**specific nct**	**common nct**	
	117	L0	G	C	
	228	L0	G	C	
	234	L0	T	C	
	351	L0	G	A	
	138	L1	C	T	
	159	L1	G	T	
	117	L2	G	C	
	243	L2	T	G	
	270	L2	C	G	
	279	L2	C	T	
	309	L2	A	T	

Table S5. *continued*

28S rDNA	position	lineage	specific nct	common nct
D3-D5	122	L0	C	T
	124	L0	C	T
	133	L0	A	T
	144	L0	T	A
	122	L1	C	T
	132	L1	T	C
	135	L1	C	T
	142	L1	C	A
	147	L1	C	T
	158	L2	A	G
	116	L2	A	T
	119	L2	G	A
	119	L4	G	A
	124	L4	T	C
	125	L4	T	gap
	130	L4	A	G
	148	L4	T	C
	154	L4	A	G

*sampling locations of specific haplotypes (no. of individuals), see also Table S1

L1_2-4, 6	BG (3), DE1 (4), DK (2), FR2 (4), IT (1), RS (1), RU2 (5), TR (1), UA (1)
L1_8-9	TR (4)
L2_16-17, 22	FR1 (4), FR4 (4)

Chapter 5

GENERAL DISCUSSION

The results of this thesis represent major advances in understanding large-scale genetic diversity within and among soil living animals and the effect of climatic changes on the genetic structure of Collembola populations during the past million years. By compiling and evaluating different phylogeographic and phylogenetic methods on four soil living Collembola species at European scale, I investigated genetic differences of populations and differences of reproductive modes due to climate changes and anthropogenic influences. Two sexual and two parthenogenetic Collembola species with overlapping ecological and environmental preferences were investigated. To get comparable data the four species were sampled from deciduous and coniferous forests at a European scale, including southern refugia. To cover the temporal resolution of genetic divergences among lineages from recent to more distant time scales I analyzed one mitochondrial (*COI*) and two nuclear (*H3* and 28S rDNA D3-D5 region) markers of different substitution rates. The conserved nuclear markers allow inspecting old diversifications of lineages and resolving time scales for which *COI* with its high intraspecific variation cannot be used. With this study design new insights into the formation of independent populations of soil-living species affected by climate changes, such as those during the last glaciation, were gained. Further, evidence for dispersal routes from south and east to central and northern Europe was obtained. The results indicate that in contrast to aboveground living species, Collembola were affected by climate changes during the Miocene (23 to 5 mya) and evaded the last Ice Age by escaping into deeper soil layers.

The first part of this thesis (Chapter 2) of three ecological different Collembola species (*Ceratophysella denticulata*, *Folsomia quadrioculata* and *Isotomiella minor*) investigated ancient diversifications, genetic distances among and between populations and distribution patterns at a pan-European scale. The results showed that Collembola have similar genetic and dispersal patterns, despite species specific differences. *C. denticulata* has limited dispersal potential due to its weakly developed furca. In addition, the three Collembola species are living in different soil layers and having different reproduction modes, as *C. denticulata* and *F. quadrioculata* are hemiedaphic and reproduce sexually compared to the euedaphic and parthenogenetic species *I. minor*.

Furthermore, Collembola follow the well-known 'southern richness and northern purity' scenario (Hewitt 2001; Sommer & Zachos 2009; Temunović *et al.* 2013), but extinction and re-colonization events took place during the Miocene (23 to 5 mya) and not during and after the last Ice Age, contrasting the pattern for species living above the ground. Furthermore, few closely related haplotypes dominated the majority of sampling locations, indicating that few individuals founded new populations and competition avoided the expansion of new alleles in these populations, resulting in low genetic variance within and high genetic variance between populations ('founder-takes-it-all' effect; Waters *et al.* 2013).

In the second part of this thesis (Chapter 3) population structure and genetic variance of one sexual (*F. quadrioculata*) and one parthenogenetic (*I. minor*) Collembola species were compared to investigate the significance of reproductive modes for colonization across Europe. In contrast to aboveground species, persistence of old lineages was supported by deep genetic divergences of both species, indicating that colonization of Europe by Collembola dates back to preglacial times supporting the findings of Chapter 2. Isolated genotypes in areas south of the Alps and in central and northern Europe suggest that refuge sites existed in these areas. However, in contrast to the sexual species, two widespread and common lineages of *I. minor* covered wide geographic areas in central and northern Europe. This supports theories that parthenogenetic reproduction facilitates colonization of new habitats. Overall, the genetic structure was more complex in the sexual species, likely due to founder effects, competition and mitonuclear processes.

As we know, nuclear- and mitochondrial genes are linked and mitonuclear processes are influenced by these linked genes, evolving differently in closely related hybrids (Ellison *et al.* 2008; Montooth *et al.* 2010; Gagnaire *et al.* 2012).

The investigated nuclear genes are not part of the functional mitonuclear complex but differences in accumulation of synonymous and non-synonymous substitutions suggest that additionally to competition and founder effects, mitonuclear and gene-by-environment interactions between polymorphic populations may explain the pronounced population structure in the two investigated species. Fitness differences between lineages

may result in the dominance of one particular lineage in a particular environment, but fitness may be reduced if local genotypes mix with immigrants. Notably, coexistence of mitochondrial haplotypes was considerably more pronounced in the parthenogenetic species *I. minor*, with the sexual species *F. quadrioculata* having only a narrow overlapping zone with closely related mitochondrial lineages in central France, possibly due to postzygotic selection or reduced fitness of hybrid offspring.

Furthermore, this study, together with the study on the oribatid mite species *Steganacarus magnus* (Rosenberger *et al.* 2013), suggests that changes in climate after the Last Glacial Maximum (LGM, 26.5 kya to 19 kya) likely were of little importance for soil-living species and that the last Ice Age less affected belowground as compared to aboveground living species, presumably as the former evaded adverse climatic conditions by moving into deeper soil layers. In addition, the parthenogenetic species *I. minor* persisted over geological times contrary to the hypothesis that parthenogenetic species are short-lived.

In the third part of this thesis (Chapter 4) the ubiquitous parthenogenetic species *Parisotoma notabilis* was investigated in detail. This Collembola species colonizes virtually any anthropogenically modified habitat suggesting human mediated colonization patterns. Results of this study showed that populations of *P. notabilis* in Europe can be divided into five independent evolving units (IEUs), but with all lineages belonging to a single species, as the 4x rule did not support the distinction of genetic species among the five IEUs. Compared to the other Collembola species studied in this thesis, anthropogenic activities appear to have had a stronger effect on the population structure of *P. notabilis* because genetic patterns suggest that this species more efficiently colonized areas connected by human trading and migration. European lineages of *P. notabilis* diverged during the Miocene, but were rather young (Late Miocene) compared to the other Collembola species studied in this thesis (Early and Middle Miocene).

There are important differences between species living above and below the ground, as the results of molecular clock analyses of this thesis clearly indicate that divergences of European Collembola populations predominated during the Miocene (23

to 5 mya) and not after LGM as described for aboveground living species in other studies (Hewitt 2000, 2001; Coyer *et al.* 2003; Sommer & Benecke 2005; Dépraz *et al.* 2008; Sommer & Zachos 2009). In fact, during the last Ice Age (2.7 mya to 11.7 kya) and the LGM populations of aboveground living species in northern and central Europe went extinct or escaped to warmer regions in the south. These unsettled areas were recolonized from southern habitats when the glaciers melted and went northwards with the beginning of the warmer Holocene (11.7 kya). As a consequence of recolonization of northern and central European habitats from southern refugia, the genetic diversity decreases from the south to the north. This famous pattern was found for many aboveground living species and is called "southern richness and northern purity"(Hewitt 2000). With molecular markers, these recolonization routes can be traced back and, in addition, local molecular variance together with genetic distances between and within populations give information about possible refugia.

The results of this thesis show higher genetic distances between and within populations in the south demonstrating that Collembola follow the above described scenario in general. However, lineages of Collembola from southern and central Europe diverged during the Middle Miocene, when the climate was favorable for little sclerotized organisms because more humidity and warmer temperature existed compared to the Early Miocene (Mosbrugger *et al.* 2005; Strömberg *et al.* 2007; Bruch *et al.* 2007; Zachos *et al.* 2008), and these favorable conditions may have resulted in the shown colonization patterns of Collembola, with stable populations over long periods of time. This scenario is supported by records of the investigated families (Hypogasturidae and Isotomidae) from Baltic amber (55 – 35 mya) (Rapoport 1971; Zawischa 1993; Hädicke *et al.* 2013) indicating that they have persisted over long periods of time with little morphological variation and have been present in Europe for millions of years. In addition, the suggested scenario is compatible with the fact that soil buffers temperature changes and organisms can escape from temperature fluctuations by moving into deeper soil layers (Healey 1967; Gass *et al.* 2006). This leads me to conclude that Collembola have survived the last Ice Age in the soil and that extinction events due to changes of the climate affected them far less than populations of above living species.

Furthermore, all investigated genes in Collembola revealed remarkable high genetic differences between but not within populations, suggesting density dependent processes for establishment of populations (founder-takes-all principle) which is common for European animals and plants (Waters *et al.* 2013). This is supported by the fact that most of the sampling locations were dominated by few and closely related haplotypes, indicating that few individuals founded the populations and prevented by competition that new alleles settled into the populations, resulting in low genetic variance within but high genetic variance between populations.

Due to strong molecular differentiation between populations a number of studies discussed the existence of cryptic species in Collembola (Cicconardi *et al.* 2010; Emerson *et al.* 2011; Porco *et al.* 2012). Genetic distances were high between populations for *COI* (up to 19%) but low within populations (up to 1%). This high genetic variation was also detected in other soil-living arthropods, e.g. in other Collembola species (Torricelli *et al.* 2010; Porco *et al.* 2012), oribatid mites (Heethoff *et al.* 2007; Schäffer *et al.* 2010; Rosenberger *et al.* 2013) and Opiliones (Boyer *et al.* 2007). In some populations the mixture of nuclear and mitochondrial markers indicates recombination in the sexual species. Furthermore, Birky's 4x rule did not support the distinction of genetic species among lineages of *Parisotoma notabilis*.

This is not to suggest that cryptic species within Collembola do not exist, but we should also consider that the pronounced structure of populations could be explained by mitonuclear and gene-by-environment interactions between populations, competition and founder effects. Despite high molecular differentiations between populations, this study indicates that each of the investigated individuals of the four species belong not only morphologically but also genetically to the same species as related patterns between different genes of different populations show.

However, the four Collembola species have not only shared, but also different genetic patterns. For example, lineages of the parthenogenetic species *Isotomiella minor* and *Parisotoma notabilis* were less separated into geographic regions, indicating that no region is more ancestral to others and suggesting separation of these lineages from a common ancestor. In addition, lineages of sexual Collembola tended to be older than lineages of parthenogenetic Collembola, indicating that recombination and thus, higher

genetic variance supports persistence of populations. Nevertheless, the widespread occurrence of lineages of recent origin indicated a higher spread rate due to being parthenogenetic. Comparing synonymous and non-synonymous substitutions to detect the effects of natural selection on molecular polymorphisms gave negative values for Fu's Fs and also indicated recent expansion in population size, supporting high colonization potential of *Isotomiella minor*. This higher colonization rate and success of populations of both parthenogenetic Collembola species is also supported by their high abundance, often representing more than 50% of the total individuals in Collembola communities (Kuznetsova 2006; García-Gómez *et al.* 2009). Accordingly, lineages of parthenogenetic species coexisted at more sampling locations compared to sexual species, suggesting that competition is less important in parthenogenetic as compared to sexual species. Overall, more non-synonymous mutations were found in the protein coding genes of the parthenogenetic species, indicating that, in addition to colonization patterns, populations were also structured by selective pressures on parthenogenetic and sexual genomes that differ in order to maintain intragenomic or gene-by-environment interactions. However, more isolated lineages were found for all of the four Collembola species studied in southern Europe, indicating that the Alps and the Pyrenees operate as dispersal barriers. Therefore, the strong phylogeographic pattern north of these mountains is due to extinctions of populations in central Europe and recolonization of empty habitats from nearby rather than from southern refugia.

In contrast to *C. denticulata*, *F. quadrioculata* and *Isotomiella minor*, lineages of *P. notabilis* are up to 10 million years younger at the same sampling location indicating that lineages of this ubiquitous Collembola species are fast colonizers, potentially due to tolerances against abiotic and biotic forces restricting the dispersal of Collembola (e.g., population size, reproductive mode, tolerances against temperature). Another difference between the four Collembola species is that populations of *P. notabilis* separated into clearly distinguishable genetic lineages, independent of which gene was used, and non-synonymous mutations in the *COI* gene represent these lineages. While lineages have a wide distribution range, coexistence of lineages in Europe was restricted to five locations. These lineages were probably bound to habitats nearby human trading routes or to habitats affected by humans. Coexistence of lineages of *P. notabilis* was more

pronounced in a small region in Canada (Porco *et al.* 2012), potentially due to multiple settlement events by humans or favorable environmental conditions allowing co-occurrence of lineages at small spacial scale. In total, different forces affect the population structure of Collembola, with, compared to the other three studied species, *P. notabilis* having less restrictions hampering dispersal (e.g., faster population growth, tolerances against metal pollution and temperature, reproductive mode). These tolerances may explain the success of this species.

In fact, *P. notabilis* is one of the most successful species of Collembola, being ubiquitous in virtually any habitat. It can reach high densities of up to 10,000 individuals per square meter and is frequently present in human mediated soils such as arable fields, pastures, urban soils and meadows (Soto-Adames 2002; Wanner & Dunger 2002; Fountain & Hopkin 2004; Kováč *et al.* 2005; Kuznetsova 2006; Ponge *et al.* 2008; García-Gómez *et al.* 2009; Salamon & Alphei 2009). The occurrence in these anthropogenic habitats suggests that its distribution is facilitated by human activities as well as passive dispersal. Thus, phylogeographic patterns of *P. notabilis* may reflect trading and migration routes of humans, similar to neolithic distributions of the Pygmy shrew (McDevitt *et al.* 2011) and the snail *Capaea nemoralis* (Grindon & Davison 2013). On the basis of tolerance against disturbances and fast colonization potential due to parthenogenesis, the results suggest that anthropogenic activities were more important for the dispersal of *P. notabilis* than climatic changes and passive dispersal. In addition to the reproductive mode, founder effects may also have contributed to the formation of genetically uniform lineages, as founder populations may prevent invasion of lineages arriving later. Human mediated dispersal could also have played a role in *F. quadrioculata* as in this species all three studied genes of the sampling locations Ukraine and Great Britain were identical.

In conclusion, the results suggest that populations of Collembola across Europe are highly structured and that diversifications into lineages predominantly occurred in the Miocene. Divergence dates and geographic structure suggest that soil microarthropods are less affected by long-term climatic changes than aboveground animals and plants. Soil has a buffering capacity against changing abiotic conditions, such as high temperature fluctuations or drought, providing more stable conditions compared to above the ground. Phylogeographic studies on amphipods (McInerney *et al.* 2014) and soil-living oribatid

mites (Rosenberger *et al.* 2013) also identified strong genetic structure and isolated lineages of Miocene and Late Pliocene origin, suggesting that belowground systems facilitated survival of species. Along with large population sizes and the ability to tolerate freezing, stable environmental conditions and resource availability likely contributed to the lower sensitivity to abiotic changes of belowground arthropods. These results suggest that evolutionary processes are slowed down in soil and populations of soil-living Collembola persisted as "living fossils" over long geological periods reflecting Miocene rather than Quaternary diversification events for both parthenogenetic and sexual species. Additionally, the results indicate that parthenogenetic lineages are not dead-ends in evolution when living in stable habitats. Old diversifications and strong genetic differentiation of populations of the investigated Collembola species resemble the phylogeographic pattern of the Collembola *Lepidocyrtus* (Cicconardi *et al.* 2010) and the oribatid mite *Steganacarus magnus* (Rosenberger *et al.* 2013), indicating strong genetic structures of populations and older diversifications of linages in soil and its associated species compared to aboveground living species. Coexistence of genetic lineages in populations of *S. magnus* was more pronounced than in the Collembola species studied, suggesting that different forces are at work for the establishment of different soil-living arthropods such as expansion potential or resource availability due to different trophic niches of the species, tolerances against temperature and other abiotic factors. However, the results suggest that competition and early colonization of new habitats are more important in Collembola than in oribatid mites (Lindberg & Bengtsson 2005; Hågvar 2010; Ingimarsdóttir *et al.* 2012).

More research is needed to investigate exact colonization routes of Collembola in Europe, to find out how the founder-takes-it-all principle and competition influence genetic differentiated populations of the same species and the establishment of new populations in empty habitats. For example, transplant experiments would allow to test if there is a genetic component to differences in populations (Molles & Molles 2002). In addition, laboratory experiments of different genetic lineages of one species will be informative concerning competition and founder effects. In general, a denser sampling in regions with high genetic difference of populations will help solving the colonization routes of Collembola and to understand the variable genetic patterns. Further, human

activities could have had influences on the dispersal and should be investigated, to support this hypothesis. Therefore, a denser sampling of populations close to human trading routes is needed, and whether humans shaped the distribution ranges can only be resolved by using more variable markers such as microsatellites or SNPs combined with experiments testing the fitness of the different lineages under varying environmental conditions.

References

Arnqvist G, Dowling DK, Eady P *et al.* (2010) Genetic architecture of metabolic rate: environment specific epistasis between mitochondrial and nuclear genes in an insect. *Evolution; International Journal of Organic Evolution*, **64**, 3354–3363.

Boyer SL, Baker JM, Giribet G (2007) Deep genetic divergences in *Aoraki denticulata* (Arachnida, Opiliones, Cyphophthalmi): a widespread "mite harvestman" defies DNA taxonomy. *Molecular Ecology*, **16**, 4999–5016.

Bruch AA, Uhl D, Mosbrugger V (2007) Miocene climate in Europe — Patterns and evolution. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **253**, 1–7.

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

Coyer JA, Peters AF, Stam WT, Olsen JL (2003) Post-ice age recolonization and differentiation of *Fucus serratus L.* (Phaeophyceae; Fucaceae) populations in Northern Europe. *Molecular Ecology*, **12**, 1817–1829.

Dépraz A, Cordellier M, Hausser J *et al.* (2008) Postglacial recolonization at a snail's pace (*Trochulus villosus*): confronting competing refugia hypotheses using model selection. *Molecular Ecology*, **17**, 2449–62.

Dowling DK, Abiega KC, Arnqvist G (2007) Temperature-specific outcomes of cytoplasmic-nuclear interactions on egg-to-adult development time in seed beetles. *Evolution; International Journal of Organic Evolution*, **61**, 194–201.

Ellison CK, Niehuis O, Gadau J (2008) Hybrid breakdown and mitochondrial dysfunction in hybrids of *Nasonia* parasitoid wasps. *Journal of Evolutionary Biology*, **21**, 1844–1851.

Emerson BC, Cicconardi F, Fanciulli PP *et al*. (2011) Phylogeny, phylogeography, phylobetadiversity and the molecular analysis of biological communities. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences*, **366**, 2391–2402.

Fountain MT, Hopkin SP (2004) Biodiversity of Collembola in urban soils and the use of *Folsomia candida* to assess soil "quality". *Ecotoxicology*, **13**, 555–72.

Gagnaire P-A, Normandeau E, Bernatchez L (2012) Comparative genomics reveals adaptive protein evolution and a possible cytonuclear incompatibility between European and American Eels. *Molecular Biology and Evolution*, **29**, 2909–2919.

García-Gómez A, Castaño-Meneses G, Palacios-Vargas JG (2009) Diversity of springtails (Hexapoda) according to a altitudinal gradient. *Pesq Agropec Bras*, **44**, 911–916.

Gass F, Gillet S, Ponge J (2006) The use of directional traps for the assessment of short-term phenanthrene effects upon soil springtail communities. *Environmental Pollution*, **2**, 364–370.

Grindon AJ, Davison A (2013) Irish *Cepaea nemoralis* Land Snails Have a Cryptic Franco-Iberian Origin That Is Most Easily Explained by the Movements of Mesolithic Humans. *PloS One*, **8**, e65792.

Hädicke C, Haug C, Haug J (2013) Adding to the few: a tomocerid collembolan from Baltic amber. *Paleodiversity*, **6**, 149–156.

Hågvar S (2010) Primary Succession of Springtails (Collembola) in a Norwegian Glacier Foreland. *Arctic, Antarctic, and Alpine Research*, **42**, 422–429.

Healey I (1967) An ecological study of temperatures in a Welsh moorland soil. *Journal of Animal Ecology*, **36**, 425–434.

Heethoff M, Domes K, Laumann M *et al.* (2007) High genetic divergences indicate ancient separation of parthenogenetic lineages of the oribatid mite *Platynothrus peltifer* (Acari, Oribatida). *Journal of Evolutionary Biology*, **20**, 392–402.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hewitt GM (2001) Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Molecular Ecology*, **10**, 537–549.

Hoekstra LA, Siddiq MA, Montooth KL (2013) Pleiotropic effects of a mitochondrial-nuclear incompatibility depend upon the accelerating effect of temperature in Drosophila. *Genetics*, **195**, 1129–39.

Ingimarsdóttir M, Caruso T, Ripa J *et al.* (2012) Primary assembly of soil communities: disentangling the effect of dispersal and local environment. *Oecologia*, **170**, 745–754.

Kováč L, Kostúrová N, Miklisová D (2005) Comparison of collembolan assemblages (Hexapoda, Collembola) of thermophilous oak woods and *Pinus nigra* plantations in the Slovak Karst (Slovakia). *Pedobiologia*, **49**, 29–40.

Kuznetsova NA. (2006) Long-term dynamics of Collembola in two contrasting ecosystems. *Pedobiologia*, **50**, 157–164.

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

McDevitt a D, Vega R, Rambau R V *et al.* (2011) Colonization of Ireland: revisiting "the pygmy shrew syndrome" using mitochondrial, Y chromosomal and microsatellite markers. *Heredity*, **107**, 548–557.

McInerney CE, Maurice L, Robertson AL *et al.* (2014) The ancient Britons: groundwater fauna survived extreme climate change over tens of millions of years across NW Europe. *Molecular ecology*, **23**, 1153–1166.

Molles MC. J, Molles MC (2002) *Ecology: Concepts and Applications*. The McGraw-Hill Companies, Inc., New York.

Montooth KL, Meiklejohn CD, Abt DN *et al*. (2010) Mitochondrial-nuclear epistasis affects fitness within species but does not contribute to fixed incompatibilities between species of *Drosophila*. *Evolution; international journal of organic evolution*, **64**, 3364–3379.

Mosbrugger V, Utescher T, Dilcher DL (2005) Cenozoic continental climatic evolution of Central Europe. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 14964–14969.

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.

Porco D, Potapov M, Bedos A *et al.* (2012) Cryptic Diversity in the Ubiquist Species *Parisotoma notabilis* (Collembola, Isotomidae): A Long-Used Chimeric Species? *PLoS One*, **7**, e46056.

Rapoport E (1971) The geographical distribution of Neotropical and Antarctic Collembola. *Pacific Insects Monograph*, 99–118.

Rosenberger M, Maraun M, Scheu S *et al.* (2013) Pre- and post-glacial diversifications shape genetic complexity of soil-living microarthropod species. *Pedobiologia*, **56**, 79–87.

Salamon J-A, Alphei J (2009) The Collembola community of a Central European forest: Influence of tree species composition. *European Journal of Soil Biology*, **45**, 199–206.

Schäffer S, Koblmüller S, Pfingstl T *et al.* (2010) Contrasting mitochondrial DNA diversity estimates in Austrian *Scutovertex minutus* and *S. sculptus* (Acari, Oribatida, Brachypylina, Scutoverticidae). *Pedobiologia*, **53**, 203–211.

Sommer RS, Benecke N (2005) The recolonization of Europe by brown bears *Ursus arctos* Linnaeus, 1758 after the Last Glacial Maximum. *Mammal Review*, **35**, 156–164.

Sommer RS, Zachos FE (2009) Fossil evidence and phylogeography of temperate species: "glacial refugia" and post-glacial recolonization. *Journal of Biogeography*, **36**, 2013–2020.

Soto-Adames FN (2002) Molecular phylogeny of the Puerto Rican *Lepidocyrtus* and *Pseudosinella*. *Molecular Phylogenetics and Evolution*, **25**, 27–42.

Strömberg CAE, Werdelin L, Friis EM *et al.* (2007) The spread of grass-dominated habitats in Turkey and surrounding areas during the Cenozoic: Phytolith evidence. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **250**, 18–49.

Temunović M, Frascaria-Lacoste N, Franjić J *et al.* (2013) Identifying refugia from climate change using coupled ecological and genetic data in a transitional Mediterranean-temperate tree species. *Molecular Ecology*, **22**, 2128–2242.

Torricelli G, Carapelli A, Convey P *et al.* (2010) High divergence across the whole mitochondrial genome in the "pan-Antarctic" springtail Friesea grisea: Evidence for cryptic species? *Gene*, **449**, 30–40.

Wanner M, Dunger W (2002) Primary immigration and succession of soil organisms on reclaimed opencast coal mining areas in eastern Germany. *European Journal of Soil Biology*, **38**, 137–143.

Waters JM, Fraser CI, Hewitt GM (2013) Founder takes all: density-dependent processes structure biodiversity. *Trends in Ecology & Evolution*, **28**, 78–85.

Wolff JN, Ladoukakis ED, Enríquez JA *et al.* (2014) Mitonuclear interactions: evolutionary consequences over multiple biological scales. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **369**, 20130443.

Zachos JC, Dickens GR, Zeebe RE (2008) An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature*, **451**, 279–83.

Zawischa D (1993) Einschlüsse im Bernstein. *Arbeitskreis Paläontologie Hannover*, **21**, 1–32.

Acknowledgements

Ganz besonders möchte ich mich bei Stefan Scheu für die Möglichkeit bedanken, Teil der Arbeitsgruppe gewesen zu sein und dafür, dass ich mein wissenschaftliches Projekt durchführen durfte. Herzlichen Dank für Deine Unterstützung, Dein offenes Ohr bei jeglichen Fragen und Dein unglaubliches Fachwissen, welches mir in vielen Situationen weitergeholfen hat.

Ein ebenso großer Dank geht an Ina Schäfer, die bereits seit meiner Diplomarbeit an mich und meine Arbeit geglaubt und mich immer unterstützt hat. Mit Deinem Fachwissen und Deiner Unterstützung hast Du maßgeblich zu meiner Dissertation beigetragen. Auch bei dem einen oder anderen Kaffee (und ‚Einreichungsbieren') hast Du mir immer den richtigen Weg aufgezeigt und mich betreut. Danke, dass Du meine Manuskripte von den ersten holprigen Versuchen an mit großer Geduld gelesen und verbessert hast, auch nachdem ich der Universität eigentlich schon den Rücken zugekehrt hatte. Danke dafür!

Bedanken möchte ich mich bei der gesamten Molligruppe für die vielen Stunden im Labor, dem einen oder anderen Tip während der Arbeit und die Kaffee- sowie Mittagspizzapausen. Vielen Dank an Jens Bast, Bernhard Eitzinger, Kerstin Heidemann, Guido Humpert, Andreas Klein und Garvin Schulz für die wirklich gute Zeit mit euch!

Ein ganz besonderer Dank geht an Patrick Pachl, der mit mir nicht nur bei Metalveranstaltungen, Grillevents, der Renovierung unseres Hauses und ähnlichem die Zeit verbracht hat, sondern mir auch bei einer besonderen Probennahme geholfen hat. Vielen Dank, dass Du mit mir die über 3.000 km lange Fahrt durch den Ostblock Europas und den Balkan gemacht hast. Diese Probennahme ist ein wichtiger Bestandteil meiner Arbeit.

Vielen Dank auch an die gesamte AG Scheu für viele nette Stunden während der Arbeit, Events und Pausen. Ein großer Dank geht auch an Marcel Graf und Christian Bluhm für die guten Diskussionen bei einigen ‚Kaffeepausen' sowie an Nicole Scheunemann.

Weiterhin möchte ich meinem Prüfungskommitee, bestehend aus Stefan Scheu, Mark Maraun und Ina Schäfer für die eine oder andere Verlängerung meiner Studienzeit danken.

Acknowledgements

Vielen Dank an Mark Maraun, Christoph Digel und Bernhard Klarner für ihre Hilfe bei der statistischen Auswertung meiner Datensätze.

Danke an Christel Fischer, Andrea Lambertz, Dieter Nünchert, Guido Humpert und Dora Volovei, die als gute Geister der AG tätig sind und unmögliche Dinge möglich machen.

Ein ganz besonderer Dank geht an Guido für die Vorbereitung meiner Proben für die Sequenzierung. Wir werden uns sicherlich noch das eine oder andere Mal beim NABU oder beim Angeln sehen!

Ein liebes Dankeschön geht an alle, die mir Proben aus allen Teilen Europas zugeschickt oder mitgebracht haben. Ohne euch wäre diese Arbeit in der Zeit nicht möglich gewesen!

Ganz besonders danken möchte ich einigen meiner Freunde, die viel Zeit mit mir während meiner Promotion verbracht haben. Ich bedanke mich besonders bei Familie Schulze (Fabian, Jasmin, meinem Patenkind Clara und ihrer Schwester Nora), Hauke und Steffen Maschmeyer, Jeldrik und Verena Moritz, Anett Rhone, Maurizio Amato, Sebastian und Yvonne Riesch und natürlich bei Frank und Hanna Keul.

Bedanken möchte ich mich bei der Firma Labotect Labor-Technik Göttingen GmbH, die es mir ermöglicht hat meine Dissertation zu beenden. Ohne die Arbeit bei euch wäre dies nicht möglich gewesen. Danke an Delia Schinkel-Fleitmann, Friederike und Peter Fleitmann!

Ich möchte mich herzlich bei meinen Brüdern Arne und Ulf Treptow bedanken. Ihr habt nicht gedacht, dass ich noch fertig werde, oder?

Ein ganz besonderer Dank geht an meine Eltern Annegret und Winfried Treptow. Vielen Dank dafür, dass ihr mich immer unterstützt und an mich geglaubt habt! Schade, dass Du es nicht mehr miterleben kannst, Papa. Diese Arbeit wäre ohne Dich und Mama nie möglich gewesen. Danke für alles!

Zuguterletzt möchte ich den größten Dank an meine Frau Fee. Du warst immer für mich da, hast meine Launen ertragen und mich dazu gebracht, meine Disseration zu beenden. Ohne Dein Mitgefühl und Deine Liebe wäre alles, was wir uns aufgebaut haben, nicht möglich gewesen. Danke dafür, dass es Dich gibt!

List of publications

Published in peer-reviewed journals

von Saltzwedel H, Maraun M, Scheu S, Schaefer I (2014) Evidence for Frozen-Niche Variation in a Cosmopolitan Parthenogenetic Soil Mite Species (Acari, Oribatida). PLoS ONE 9(11): e113268. doi:10.1371/journal.pone.0113268

Submitted

von Saltzwedel H, Scheu S, Schaefer I (under review) Founder events and pre-glacial divergences shape the genetic structure of European Collembola species.

von Saltzwedel H, Scheu S, Schaefer I (under revision) Genetic structure and colonization patterns of sexual and parthenogenetic springtails (Hexapoda, Collembola) in Europe.

von Saltzwedel H, Scheu S, Schaefer I (under revision) Colonization of Europe by *Paristoma notabili*s (Collembola): cryptic diversity and anthropogenic impact.

Thesis declarations

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

Chapter 2 is currently submitted to a peer-reviewed journal; I have collected all data. Chapter 3 is currently submitted to a peer-reviewed journal; I have collected all data. Chapter 4 is currently submitted to a peer-reviewed journal; I have collected all data.

I am the first author of all manuscripts; I have analyzed the data, written the manuscripts, developed the main ideas, created tables, figures and appendices and contributed significantly to the study design. All persons contributing to the manuscripts have been named so. All co-authors contributed to finalizing the manuscripts.

Plagiarism declaration

I, Helge von Saltzwedel, 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.

_____

Helge von Saltzwedel

Göttingen, December 2015