

Biological invasions in soil: DNA barcoding as a monitoring tool in a multiple taxa survey targeting European earthworms and springtails in North America

David Porco · Thibaud Decaëns · Louis Deharveng · Samuel W. James · Dariusz Skarżyński · Christer Erséus · Kevin R. Butt · Benoit Richard · Paul D. N. Hebert

Received: 8 October 2011 / Accepted: 31 August 2012 / Published online: 12 September 2012
© Springer Science+Business Media B.V. 2012

Abstract Biological invasions are increasingly recognized as a potent force altering native ecosystems worldwide. Many of the best documented cases involve the massive invasions of North America by plant and animal taxa native to Europe. In this study, we use DNA barcoding to survey the occurrence and genetic structure of two major groups of soil invertebrates in both their native and introduced ranges:

Collembola and earthworms. Populations of ten species of earthworms and five species of Collembola were barcoded from both continents. Most of these species exhibited a similar genetic structure of large and stable populations in North America and Europe, a result supporting a scenario of multiple invasions. This was expected for earthworm species involved in human economic activities, but not foreseen for Collembola species de facto unintentionally introduced. This study also establishes that invasive species surveys employing DNA barcoding gain additional resolution over those based on morphology as they

Electronic supplementary material The online version of this article (doi:[10.1007/s10530-012-0338-2](https://doi.org/10.1007/s10530-012-0338-2)) contains supplementary material, which is available to authorized users.

D. Porco (✉) · T. Decaëns · S. W. James · B. Richard
Laboratoire d'Ecologie, EA 1293 ECODIV, FED
SCALE, Bâtiment IRESE A, UFR Sciences et
Techniques, Université de Rouen, Place Emile Blondel,
76821 Mont Saint Aignan, Cedex, France
e-mail: david.porco.gm@gmail.com

D. Porco · P. D. N. Hebert
Canadian Centre for DNA Barcoding, University
of Guelph, 50 Stone Road East, Guelph,
ON N1G 2W1, Canada

L. Deharveng
Museum National d'Histoire Naturelle,
UMR7205 «Origine, Structure et Evolution de la
Biodiversité», 45 rue Buffon, CP50, 75005 Paris, France

Present Address:

S. W. James
Department of Biology, University of Iowa,
Iowa City, IA 52242, USA

D. Skarżyński
Department of Evolutionary Biology and Ecology,
Wrocław University, Przybyszewskiego 63/77,
51-148 Wrocław, Poland

C. Erséus
Department of Biology and Environmental Sciences,
University of Gothenburg, Box 463, 405 30 Göteborg,
Sweden

K. R. Butt
School of Built and Natural Environment, University
of Central Lancashire, Preston PR1 2HE, UK

B. Richard
Laboratoire BioSol, Esitpa-Ecole d'Ingénieurs en
Agriculture, 3 rue du Tronquet, 76134 Mont-Saint-
Aignan, Cedex, France

allow evaluation of cryptic lineages exhibiting different invasion histories.

Keywords Invasion survey · Underground · Molecular taxonomy · Springtails · Lumbricidae · Anthropogenic dispersion

Introduction

Both the pace of non-indigenous species European colonization (Addison 2009), and the success of biological invasions have dramatically increased with the rise of international trade. These biotic exchanges are leading to an increasing homogenization of communities (Olden et al. 2004; Winter et al. 2010), and represent one of the main causes of biodiversity decline (Primack 2000; Wilson 2002). Accordingly, scientific interest in this topic has expanded rapidly over the last 30 years. While soil invasions have generally received little attention, some taxa have recently been studied, e.g. earthworms (Hendrix et al. 2008). This neglect is largely due to the ‘taxonomic impediment’ (Rodman and Cody 2003): in most soil groups, species identification is a difficult task and specialist expertise is often lacking (Decaëns et al. 2006). In the past, earthworms, among other soil invertebrates, received more attention because they represent the most conspicuous element of the soil biological invasions due to their large body size and their role as soil engineers. However, it is undoubtedly the case that smaller-bodied invaders, having no such perceivable activity, also occur in soil. In this study, the invasive patterns in two model taxa of different size are compared: earthworms and Collembola from their native (European) and invasive (North American) ranges.

Over past decades, there has been increasing documentation of the occurrence and consequences of the invasions of European earthworms in North America (Addison 2009). After Pleistocene glaciations, the extinction of native North American earthworms, except in southern refuges, left many regions of North America devoid of native species at the time of European colonization (Addison 2009), and allowed non-native species spreading (Reynolds 1973). Various human activities have led to recurrent invasions (Cameron et al. 2007; Cameron et al. 2008). The effects on soil ecosystem properties and functioning include major impacts on litter decomposition,

horizon mixing, the cycling of carbon, nitrogen (Burtelow et al. 1998; Groffman et al. 2004; Suárez et al. 2006; Wironen and Moore 2006), phosphorous (Suárez et al. 2004) and calcium (Holdsworth et al. 2008). These changes have major impacts on the plant (Hale et al. 2008), microbial, mesofaunal (Eisenhauer et al. 2007; McLean et al. 2006) and fungal communities (McLean and Parkinson 2000).

By contrast, biological invasions by Collembola and their impact on habitats have been much less documented in spite of their ecological significance (Hopkins 1997). To date, Collembola invasion has only been documented in a few sub-Antarctic Islands (Greenslade and Wise 1984; Gabriel et al. 2001; Greenslade 2002a; Myburgh et al. 2007; Greenslade 2008; Greenslade and Convey 2012), in Australia (Womersley 1939; King et al. 1985; Greenslade et al. 2002; Greenslade 2002b; Oliver et al. 2005; Greenslade 2008) and in New Zealand (Salmon 1941). Shown to competitively exclude certain native species (Convey et al. 1999), invasive Collembola display a greater ability to adapt to rising temperature and so could be favored by climate changes (Slabber et al. 2007; Janion et al. 2009). Disturbance, land use changes (Yeates 1991; Yeates et al. 2000), the use of fertilizers, and high grazing intensity (Oliver et al. 2005) also favors invasive species.

The lack of taxonomic expertise and data on existing invasive taxa in soil biota, emphasize the need for new tools to rapidly assess both local and global situations. In this context, molecular systematic, especially DNA barcoding (Hebert et al. 2003), is valuable for detecting or identifying invasive species of plants (Van de Wiel et al. 2009) and animals (Darling and Blum 2007). Successfully tested in earthworms and Collembola (Rougerie et al. 2009, Decaëns et al. 2012), this approach also allows the identification of juvenile specimens (Richard et al. 2010; Chown et al. 2008), preventing an underestimation of the propagule pressure for invasive species.

This study represents the first multi-taxa application of DNA barcoding to the investigation of biological invasions in soil. The effectiveness of this tool for detection and confirmation of invasive status in these two target groups is tested. In addition, we aim to use the DNA barcoding data to compare genetic structures between native and invaded areas as a proxy of invasion history (number of introductions, propagule pressure).

Materials and methods

Sampling

In total, 941 specimens belonging to 10 species of Lumbricidae and 275 specimens from 5 species of Collembola were collected from several localities in Europe and North America (Supplementary Material Table 1). In addition, 94 sequences of Lumbricidae (AM774234–AM774293, EU035477–EU035492, FJ214209–FJ214213, FJ214215, FJ214218, FJ214221, FJ214222, FJ214227, FJ214230–FJ214232, FJ214234, FJ214235, FJ214237, FJ374776, FJ374777) and one sequence of Collembola (EU869805) were downloaded from GenBank. All the other specimens were processed in the context of the International Barcode of life project (iBOL) for the campaigns ‘Barcoding Collembola’ and ‘Barcoding Earthworms’ (<http://www.earthwormbol.org>). Specimen data are available in the project ‘DATASET-INVSOIL1’ on the Barcode of Life Data System (www.boldsystems.org). Each sequence in these libraries is linked to a specimen identified by taxonomic specialists and deposited as a voucher in a public institution.

Preliminary remarks concerning the species origin

For Lumbricidae the direction of circulation between North America and Europe is clear as it is well established that the invasive species present in North America originated from Europe (Gates 1966). The situation is more difficult to assess for Collembola. The five species that were chosen as models for this study have all been flagged as introduced European species in the literature (Maynard 1951; Stach 1966; Christiansen and Bellinger 1998), but without reference to a broader biogeographical context. Here we first evaluated this hypothesis with all the elements available for the five species of Collembola studied: *Neanura muscorum*, *Orchesella villosa*, *Orchesella cincta*, *Parisotoma notabilis* and *Sminthurinus elegans*.

For *Neanura muscorum*, the hypothesis of a European origin is strongly supported by the fact that all other species of the genus are restricted to Eastern Europe (Bellinger et al. 2011). The situation is more complex for the four other species, as they belong to genera widely distributed across the Holarctic, making the biogeographical arguments less clear.

Eastern Russia is the main center of diversification for *Parisotoma* with the highest number of species (Potapov 2001). Moreover *P. notabilis* was not found in the most oriental part of the genus’ distribution in Eastern Asia (Potapov 2001; Yosii 1977), while it is very abundant in Western Europe. From these chorological data, it can be assumed that *P. notabilis* was introduced from Europe to North America; a phylogenetic framework would nevertheless be useful to investigate this hypothesis further.

Orchesella villosa and *O. cincta* were already claimed to be introduced species in North America (Maynard 1951; Christiansen and Bellinger 1998). Like *Parisotoma*, *Orchesella* exhibits its highest diversity in the old world, with a majority of species found in Europe (Bellinger et al. 2011), by contrast, it is only represented in the western regions. So, these two species are also likely to originate from Europe.

Sminthurinus could be another old world genus with its highest species diversity in Europe (Bellinger et al. 2011). This pattern suggests that the species *S. elegans* used in this study is likely to be an introduced species in North America.

So the European origin is a reliable working hypothesis for the species of Lumbricidae and Collembola examined for this study.

Molecular analysis

DNA extraction was achieved in 50 µL of lysis buffer with proteinase K incubated at 56 °C overnight. It was performed following a standard automated protocol using 96-well glass fiber plates (Ivanova et al. 2006). For Collembola, the entire animal was lysed, and specimens were recovered after DNA extraction using a specially designed work flow (Porco et al. 2010a). The 5’ region of COI used as a standard DNA barcode was amplified using M13 tailed primers LCO1490 and HCO2198 (Folmer et al. 1994). Samples that failed to generate an amplicon were subsequently amplified with a pair of internal primers combined with full length ones LepF1-MLepR1 and MLepF1-LepR1 (Hajibabaei et al. 2006). A standard PCR reaction protocol was used for amplifications, and products were checked on a 2 % E-gel 96Agarose (Invitrogen). Unpurified PCR amplicons were sequenced in both directions using M13 tailed primers (Hajibabaei et al. 2005), with products subsequently purified using Agencourt CleanSEQ protocol and processed using

BigDye version 3.1 on an ABI 3730 DNA Analyzer (Applied Biosystems). Sequences were assembled with Sequencher 4.5 (GeneCode Corporation, Ann Arbor, MI, USA) and aligned by eye using BIOEDIT version 7.0.5.3 (Hall 1999). As we observed no indels in the COI sequences, sequence alignment was unambiguous. Sequences are publicly available on BOLD in the project DATASET-INVSOIL1 and on GenBank (accession numbers AM774234–AM774293, EU035477–EU035492, EU869805, FJ214209–FJ214213, FJ214215, FJ214218, FJ214221, FJ214222, FJ214227, FJ214230–FJ214232, FJ214234, FJ214235, FJ214237, FJ374776, FJ374777, FJ937284, FJ937285, FJ937292, FJ937295–FJ937297, FJ937299, FJ937300, FJ937304–FJ937325, GQ373501–GQ373508, GQ373667–GQ373670, GQ373867–GQ373872, GU013792–GU013814, GU013817, GU013819–GU013825, GU013830, GU013831, GU013834–GU013839, GU014221–GU014232, GU206152–GU206159, GU206161, GU206162, GU206169–GU206173, GU206175, GU206178–GU206182, GU206184, GU206187, GU206190, GU206191, GU206213–GU206239, GU656217, GU656242, GU656408–GU656411, GU656418–GU656423, GU656430–GU656433, GU656446, GU656448–GU656451, GU656502, GU656503, GU656507, GU656612, GU656614–GU656619, GU656621, GU656625, GU656627–GU656629, GU656705, GU656706, GU656709–GU656711, GU656778, GU656886, HM388349–HM388353, HM397729, HM397730, HM397758–HM397761, HM398181, HM398303–HM398307, HM398316–HM398319, HM398981–HM398989, HQ024539, HQ024542–HQ024671, HQ682449–HQ682497, HQ942779, JQ908634–JQ935043, JQ935046–JQ935066, JQ935069–JQ935072, JQ935074–JQ935076, JQ935090, JQ935092, JQ935101–JQ935110, JQ935112, JQ935113, JQ935115, JQ935116, JQ935118, JQ935121, JQ935123, JQ935125, JQ935127, JQ935130–JQ935132, JQ935134, JQ935135, JQ935137, JQ935138, JQ935140–JQ935147, JQ935151, JQ935152, JQ935155, JQ935156, JQ935160–JQ935163, JQ935165–JQ935172, JQ935181–JQ935203, JX008093–JX008098).

Calculations

Distance analyses were conducted with MEGA4 (Tamura et al. 2007) using a Neighbor-Joining algorithm (Saitou and Nei 1987). Distances were estimated using the Kimura-2 parameter model (Kimura 1980).

The robustness of nodes was evaluated through bootstrap re-analysis of 1000 pseudoreplicates. The trees were replotted using the online utility iTOL (Letunic and Bork 2007).

For the five species of Collembola and five species of earthworms for which the sampling effort was comparable in terms of population number on each continent (Supplementary Material Table 1), genetic diversity (haplotype diversity (h), nucleotide diversity (π)) and gene flow/population exchange between the two continents (Nm, Fst) were analysed along with descriptive statistics, mismatch distributions (population growth-decline model, Raggedness, R2 values were estimated), and neutrality tests (Tajima's D, Fu's Fs). All these values were calculated with DnaSP v5.10 (Librado and Rozas 2009).

Results

Some well characterized species among those included in the present study consist of several distinct genetic lineages already found in previous studies (Chahartaghi 2007; King et al. 2008; Pérez-Losada et al. 2009; James et al. 2010), but a few additional lineages were detected here (Table 1). For the remaining species, the COI barcode clusters obtained corresponded to the morphological identifications of the specimens. The neighbor joining trees for both groups, showed North American matches for the different European species and their lineages (Figs. 1 and 2). Although all 15 nominal species of Collembola and Lumbricidae were detected in North America, some of their COI lineages were not (*Parisotoma notabilis* L3; *Allolobophora chlorotica* L1, L3, L4, L5; *Aporrectodea rosea* L3; *Aporrectodea caliginosa* L1; *Lumbricus rubellus* L1; *Lumbricus herculeus*—throughout the text, we will consider *L. terrestris* and *L. herculeus* as lineages to allow one to keep track of the fact that we address entities that were until recently considered as cryptic lineages within *L. terrestris*). A more exhaustive sampling of the target species might reveal more haplotypes than found in this study and possibly more lineages. A clear gap between the interspecific and intraspecific variation was found in both datasets. For Collembola, the mean intraspecific variation was 2.32 % versus 26.15 % interspecific divergence and 21.43 % for interlineage distances in *P. notabilis* (Supplementary Material Table 2; Supplementary

Table 1 COI lineages within nominal species

Species	Number of COI lineages	Lineages names	Source
Earthworms			
<i>Allolobophora chlorotica</i>	5	L1, L2, L3, L4, L5	King et al. (2008)
<i>Aporrectodea caliginosa</i>	3	L1, L2, L3	Pérez-Losada et al. (2009)
<i>Aporrectodea longa</i>	1		
<i>Aporrectodea rosea</i>	4	L1, L2, L3, L4	King et al. (2008)
<i>Dendrobaena octaedra</i>	1		
<i>Dendrodrilus rubidus</i>	1		
<i>Lumbricus castaneus</i>	1		
<i>Lumbricus rubellus</i>	2	L1, L2	King et al. (2008)
<i>Lumbricus terrestris</i>	2	<i>L. terrestris</i> / <i>L. herculeus</i>	James et al. (2010)
<i>Octolasion cyaneum</i>	1		
Springtails			
<i>Neanura muscorum</i>	1		
<i>Orchesella cincta</i>	1		
<i>Orchesella villosa</i>	1		
<i>Parisotoma notabilis</i>	3	L1, L2, L3	Porco et al. (2012)
<i>Sminthurinus elegans</i>	1		

Material Fig. 1). The mean intraspecific variation was 1.47 % in earthworms versus 22.09 % for interspecific divergence and 15.88 % for interlineage distances (Supplementary Material Table 3; Supplementary Material Fig. 2).

Table 2 displays the results of population genetics analyses for the species and lineages having a comparable sampling effort in Europe and North America. For most species, both haplotype and nucleotide diversity are high and equivalent between European and North American populations. There were a few exceptions in genetic diversity evenness between the two continents. The Collembola species *S. elegans*

exhibited lower nucleotide diversity in North America. All Collembola species except *O. villosa* showed a high value of Nm and a low Fst. The earthworm lineages *Ap. rosea* L4 and *Ap. caliginosa* L2 had similar haplotype diversity on both continents but lower nucleotide diversity values in North America. Among those species only *Ap. caliginosa* L2 exhibited a low value of Nm and relatively high Fst. All species with an evenness of both haplotype and nucleotide diversity exhibited a high Nm value and low Fst. For the two groups, comparable mismatch distributions were found (multimodal or unimodal) in both areas (Supplementary Material Figs. 3, 4). Fu's Fs, Tajima's D, raggedness and R2 values were not significant (Table 2), allowing one to reject the hypothesis of recent demographic expansion, thus confirming that these populations are well settled and stable.

Discussion

Barcode efficiency and its relevance for monitoring soil invasions

Our results showed that COI sequence data provide an appropriate mean of detecting biological invasions in two major groups of soil invertebrates where such studies would normally be constrained by the taxonomic impediment. Such DNA barcode analyses will enable large-scale studies requiring substantial numbers of routine identifications. The ability of DNA barcodes to distinguish closely related species in earthworms and Collembola has been established in several prior studies (Hogg and Hebert 2004; Chang et al. 2009; Richard et al. 2010; Porco et al. 2010b). However, this approach also brings a finer resolution than morphology by revealing potential cases of cryptic diversity in well studied species (Table 1). Our study revealed a number of cases in which COI divergences between lineages of a single nominal 'species' were as great as typical interspecific distances (Supplementary Material Figs. 1 and 2), suggesting the likely presence of cryptic species. The specific status of many of these lineages has, in fact, already been confirmed through supplemental molecular or morphological analysis (King et al. 2008; Pérez-Losada et al. 2009; James et al. 2010; Porco et al. 2012). This high discrimination power has sometimes been considered problematic (Armstrong

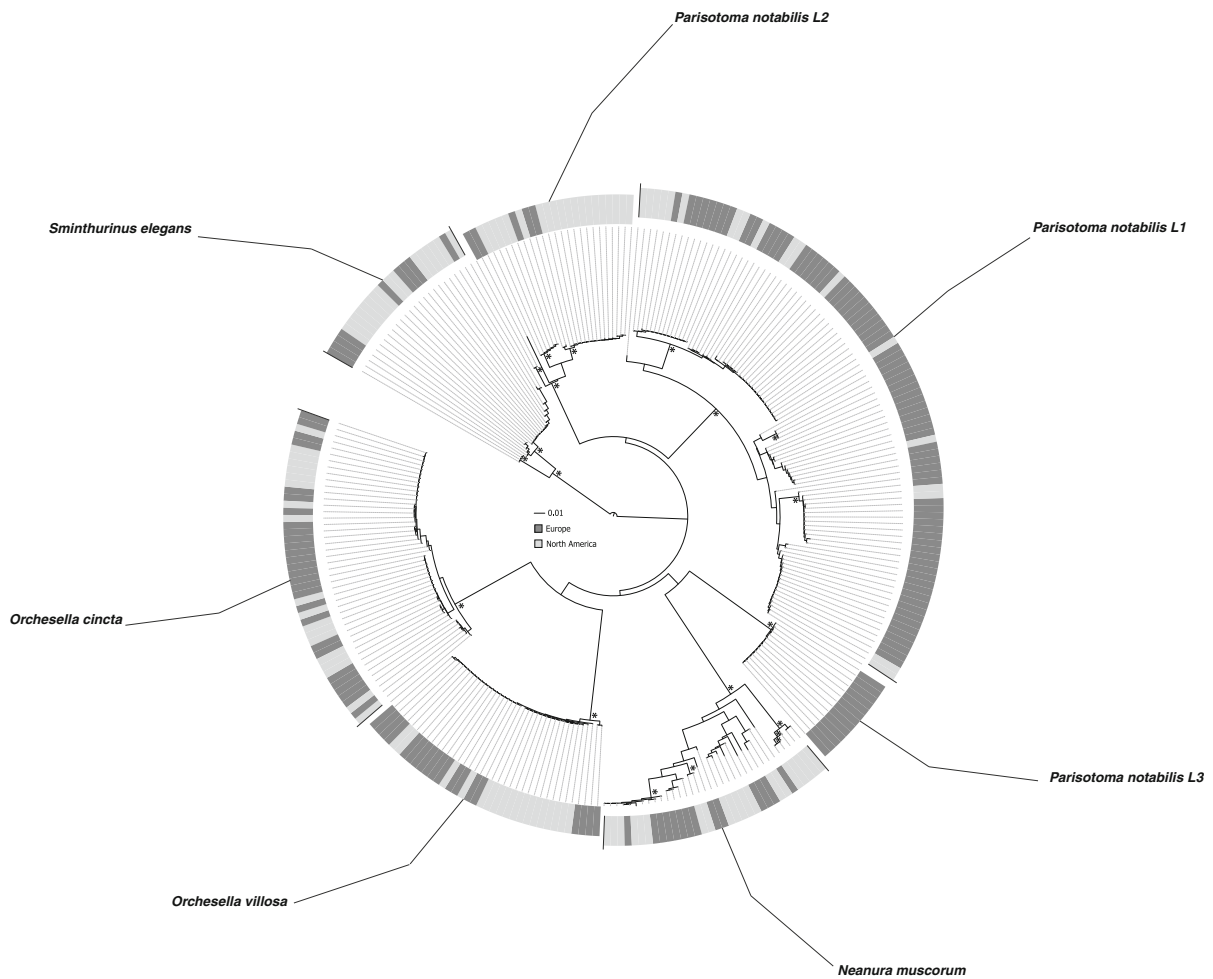


Fig. 1 Neighbor joining K2P tree obtained from 275 specimens among 5 species of Collembola. Bootstrap values $\geq 99\%$ are marked by an asterisk over the branch

and Ball 2005), but this conclusion ignores the growing evidence that most cases of high intraspecific divergence reflect overlooked species or incorrect identifications. By contrast, the importance of considering these cryptic lineages is crucial to gain a proper understanding of biological invasions. Indeed, the COI lineages detected within a single species can exhibit different invasive patterns, occurring sympatrically in some of the localities (Supplementary material Table S1). Consideration of these lineages makes it possible to discern biogeographic trends of invasions that would be overlooked through morphological analysis alone. For example, some lineages of our target species were not detected in North America (*L. herculeus*, *L. rubellus* L1, and *P. notabilis* L3; see Supplementary material Table S1), suggesting potential ecological or biological differences, further

supporting their status as distinct species. It is important to note that the use of provisional names enables the assembly and comparison of distributional data on each of these distinct genetic entities long before the description of any new taxa or the connection of described taxa to their valid name. It was for this reason that we adopted the provisional names for five lineages of *Al. chlorotica* (King et al. 2008) and introduced a similar number system for the lineages in other species.

A second criticism concerning the utility of DNA barcodes for the detection of invasive species involved the current limitations in taxonomic coverage (Darling and Blum 2007). However, barcode coverage is expanding rapidly for numerous taxonomic groups (International Barcode of Life project—<http://ibol.org/>) overcoming this disadvantage. This ongoing effort will

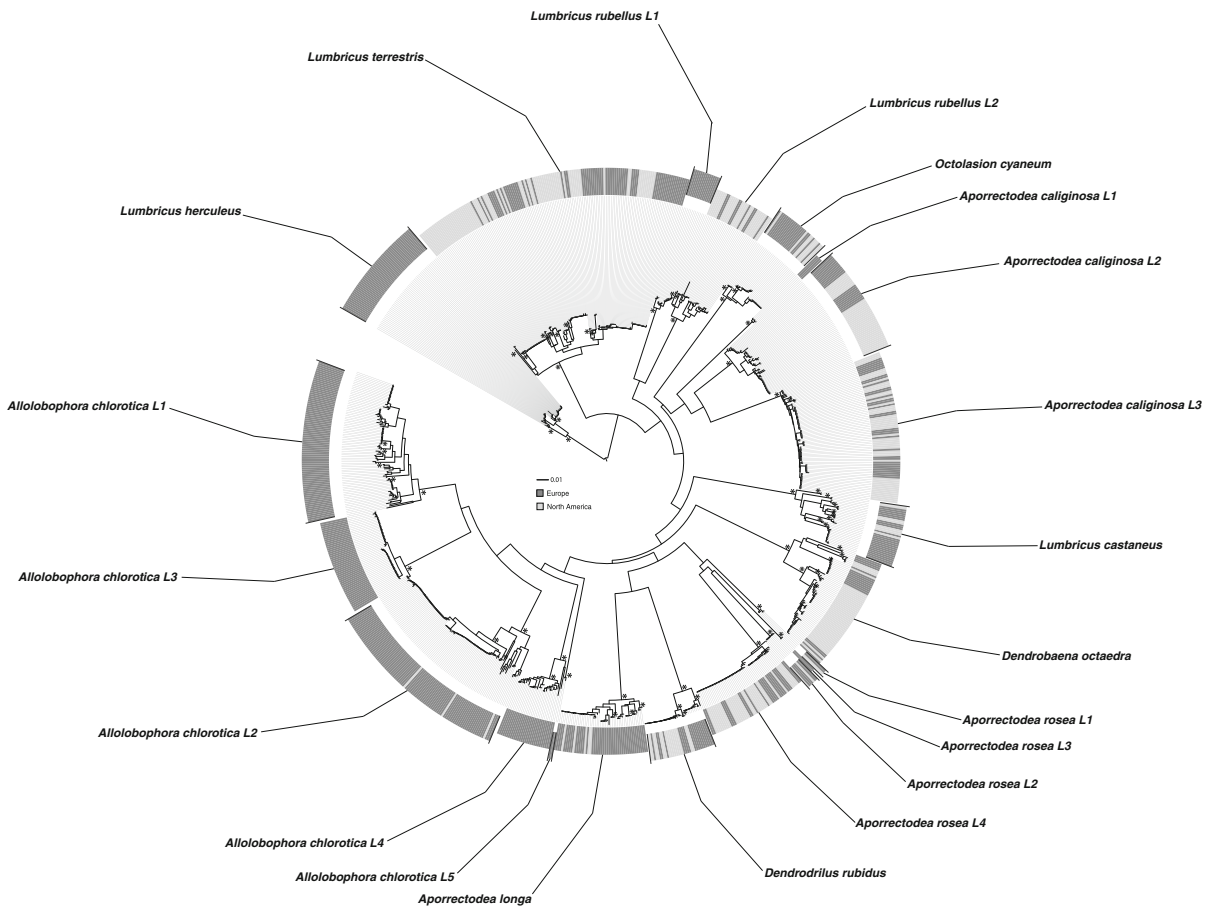


Fig. 2 Neighbor joining K2P tree obtained from 941 specimens among 10 species of Lumbricidae. Bootstrap values $\geq 99\%$ are marked by an *asterisk* over the branch

undoubtedly increase the range of this approach both geographically and taxonomically, leading to the discovery of many cases of overlooked biological introductions or invasions in all biota. The utility of DNA barcoding in this regard has already been revealed in Lepidoptera (Chown et al. 2008; deWaard et al. 2009), the insect order with the best parameterized barcode library.

Comparing the genetic structure of native and invasive populations

COI data generated for the detection of invasive species can also allow the comparison of their genetic structure with that in their native distribution. Such analyses reveal information on the genetic diversity and demography in the different species for both areas,

providing an indirect assessment of the propagule pressure in the invaded one.

In this study, all species showed similar patterns of mismatch distributions in native and invaded areas. If the mismatch distributions in the European native populations reflect historical events, the similar profiles for introduced populations suggest recurrent introduction events involving large numbers of individuals from across European source populations. Certainly, mutation and drift are unlikely to be responsible for the presence of so much variation given the brief interval since the introduction in North America. We detected erratic mismatch distributions in six species (*D. octaedra*, *L. terrestris*, *Ap. caliginosa* L2, *N. muscorum*, *O. cincta*, *P. notabilis* L1) indicating stable rather than expanding populations over a substantial period of time, a conclusion supported by the high average number of pairwise

Table 2 Genetic diversity indices and neutrality test statistics (N = number of individuals; H = number of haplotypes; N pop. = number of populations; π = nucleotide diversity; h = haplotype diversity; SE = Standard error; Nm = gene flow estimate; Fst = genetic differentiation; D = Tajima's D; Fs = Fu's Fs; r = Ragedness; R2 = Ramos-Onsins and Rozas R2)

	N	H	N	N	h	(SE)	π	(SE)	Nm	Fst	D*	Fs*	r*	R2*	Mean Pairwise Differences
<i>Lumbricus terrestris</i> Europe	69	23	21	21	0.893	(0.022)	0.02316	(0.00227)	2.07	0.19421	-1.0598	8.8145	0.0216	0.0758	10.723
<i>Lumbricus terrestris</i> North America	82	14	9	9	0.836	(0.026)	0.0337	(0.00142)			-0.1632	-0.276	0.1048	0.1021	13.177
<i>Dendrobaena octaedra</i> Europe	34	20	4	4	0.900	(0.041)	0.04294	(0.00322)	2.56	0.16346	-1.5007	-1.4168	0.0277	0.1371	26.111
<i>Dendrobaena octaedra</i> North America	44	22	5	5	0.932	(0.023)	0.01985	(0.00443)			0.175	1.6269	0.0292	0.0772	12.071
<i>Aporrectodea caliginosa</i> L3 Europe	33	11	9	9	0.877	(0.034)	0.00699	(0.00057)	6.8	0.06852	-0.9934	-0.793	0.055	0.0832	4.045
<i>Aporrectodea caliginosa</i> L3 North America	49	8	7	7	0.832	(0.024)	0.00529	(0.00048)			-0.46247	1.5977	0.0741	0.0949	3.446
<i>Aporrectodea caliginosa</i> L2 Europe	21	10	5	5	0.862	(0.057)	0.01376	(0.00125)	0.53	0.48655	1.05499	0.327	0.0388	0.1702	6.352
<i>Aporrectodea caliginosa</i> L2 North America	35	8	5	5	0.709	(0.073)	0.00672	(0.00120)			-0.0802	1.0022	0.0805	0.1221	3.19
<i>Aporrectodea rosea</i> L4 Europe	22	3	7	7	0.567	(0.051)	0.00379	(0.00025)	1.48	0.2524	2.5208	3.5366	0.5667	0.2603	2.082
<i>Aporrectodea rosea</i> L4 North America	33	3	4	4	0.322	(0.097)	0.00195	(0.00057)			0.1329	1.8106	0.5024	0.1297	1.038
<i>Parisotoma notabilis</i> L1 Europe	76	14	5	5	0.874	(0.019)	0.03196	(0.00099)	21.52	0.0227	2.42353	4.215	0.0735	0.1817	9.108
<i>Parisotoma notabilis</i> L1 North America	17	8	7	7	0.853	(0.066)	0.03121	(0.00572)			1.084	3.052	0.0735	1817	11.235
<i>Neanura muscorum</i> Europe	14	8	3	3	0.868	(0.068)	0.03301	(0.00640)	3.33	0.13056	-0.5549	3.3187	0.1005	0.1221	16.769
<i>Neanura muscorum</i> North America	20	9	3	3	0.895	(0.041)	0.03969	(0.00390)			1.2518	3.444	0.0462	0.1943	10.533
<i>Orchesella cincta</i> Europe	28	10	5	5	0.791	(0.059)	0.00842	(0.00079)	16.35	0.03154	0.82505	0.5175	0.1309	0.152	5.407
<i>Orchesella cincta</i> North America	19	4	5	5	0.702	(0.065)	0.01105	(0.00094)			1.65115	6.644	0.3378	0.2062	6.187
<i>Orchesella villosa</i> Europe	19	3	4	4	0.374	(0.130)	0.00114	(0.00046)	0.39	0.55992	-0.8295	-0.085	0.2264	0.0994	5.96
<i>Orchesella villosa</i> North America	18	3	2	2	0.451	(0.117)	0.00073	(0.00021)			-0.43794	-0.377	0.1362	0.0998	0.477
<i>Sminthurinus elegans</i> Europe	9	3	3	3	0.6667	(0.105)	0.03519	(0.00534)	0.96	0.34278	2.5682	9.0513	0.4568	0.2674	14.111
<i>Sminthurinus elegans</i> North America	16	2	3	3	0.500	(0.074)	0.00284	(0.00042)			1.6871	2.7081	0.75	0.25	1

* None of the values of D, Fs, r and R2 were found significant

differences and the results of the neutrality tests. By contrast, we found unimodal mismatch distributions in four species (*Ap. caliginosa* L3, *Ap. rosea* L4, *O. villosa*, *S. elegans*) with low numbers of pairwise differences, suggesting a bottleneck in their source origin (Europe) followed by recent expansion (Supplementary Material Fig. 3 and 4).

A similar level of genetic diversity was found in European and North American populations of most species (Table 2) indicating no bottleneck effects.

Accordingly, the combined native and invasive range genetic structure (F_{st}) and gene flow estimates (Nm) suggests massive and recurrent introductions of each species, supporting the loss of the “genetic paradox” in soil invasions as already emphasized for aquatic organisms (Roman and Darling 2007).

However, a few species in both groups had less genetic diversity in North America than in Europe (Table 2). For example, *D. octaedra* and *S. elegans* exhibited lower nucleotide diversity in North America suggesting their rapid expansion from a small number of individuals and the loss of variation following colonization due to low effective population numbers. *Ap. rosea* L4 had less nucleotide diversity in North America, but comparable haplotype diversity, suggesting rapid expansion after the introduction of individuals from a few genetically distinct populations. However, even these three species showed low F_{st} and high population exchange (Nm) between the two continents.

The low nucleotide diversity in North American populations of *Ap. caliginosa* L2 was associated with high F_{st} values suggesting genetic drift or that the native source of invading populations was not sampled in this study. Native populations of the collembolan *O. villosa* seem to have experienced a bottleneck event as both nucleotide and haplotype diversity values were low. In fact, these values were similar in populations from both Europe and North America, but F_{st} was high. This divergence likely also reflects the origin of North American populations from parts of Europe that were not sampled in this study.

Our genetic results, indicating either multiple and massive introduction events or expansion patterns can be diversely interpreted for the two target taxa of this study. Human activity likely explains high propagule pressure for some earthworm species that we analyzed. For example, *L. terrestris* and *D. octaedra*,

which are widely used as fishing bait, showed high genetic similarity between European and North American populations, a result consistent with multiple, large-scale introductions implied by such a trade. On the other hand, a species like *Ap. rosea*, not directly useful for any human activity, exhibited a genetic pattern suggestive of the introduction of a small number of individuals. Our data also confirmed the rarity of *Al. chlorotica* in North America (Whalen 2004), a fact that may reflect its low resistance to freezing (Tiunov et al. 2006) or low tolerance to soil acidity (Pearce 1972). However, this rarity may also reflect the absence of economic interest in this species, meaning that introduction events have been accidental, keeping the propagule pressure low. The fact that some individuals have survived suggests that this species could, in time, gain a broader invasive distribution.

The effect of human activity on collembolans is undoubtedly more homogenous as all introductions have been accidental. These unintentional introductions have generally resulted in similar genetic structure on both continents (*P. notabilis*, *N. muscorum*, *O. cincta* and *O. villosa*), suggesting massive and multiple introductions. However, this dispersal almost certainly reflects massive unidirectional exchanges of specimens from Europe to North America in soil used as ship ballast (Vazquez and Simberloff 2001).

Massive parallel sequencing technologies: future application?

The rapid expansion in species coverage for DNA barcode libraries will soon make it possible to employ rapidly-developing environmental sequencing approaches. This transition will enable inexpensive, rapid and accurate surveys permitting extensive screening of invasive species in various target groups. Massively parallelized sequencing will certainly be a useful tool to evaluate the impact of invasive species on the biological diversity of native fauna (Darling and Blum 2007). Because this technology cannot yet produce accurate quantitative data (Amend et al. 2010), propagule pressure cannot be assessed directly. However, the data generated by these techniques for population genetic analysis will allow a rapid indirect appraisal of propagule pressure on broader geographical and taxonomical scales.

Conclusion

With the taxonomic impediment affecting most soil taxa, molecular taxonomy tools such as DNA barcoding, appear to be an appropriate solution for the early, effective and accurate detection of biological invasions, especially for small and understudied taxa such as Collembola. This method also provides an opportunity to consider cryptic diversity in common and well-studied species, a benefit which was here proven critical as the various genetic entities exhibited diverse patterns of invasiveness possibly related to different ecological traits.

In most cases, the genetic structure of North American and European populations in both earthworms and Collembola were similar, fitting consistently with multiple and massive introduction events and thus a high propagule pressure. If this was anticipated in earthworm species involved in human trade, it was less expected in Collembola. These results indicates that passive introductions over a period of time can be as successful as intentional ones, suggesting this could be a general trend in other groups of the soil fauna, especially small body size ones which are easily overlooked.

Facilitation phenomena among invasive taxa, through the alterations of ecosystems, underlie a wider impact on a large scale and should be further investigated. Accordingly, this first study investigating biological invasions conjointly in two soil groups, advocates the necessity of further broad surveys accounting for several taxonomic groups of soil communities at the same time. DNA barcoding has the potential to make a significant contribution with extensive reference libraries available and a rapidly building capacity for those still lacking. Its use can be particularly critical to uncover and quantify 'cryptic invasions' such as that of the Collembola on a large scale. Moreover, this could also be the opportunity to monitor the many taxa rarely sampled in standard surveys.

Acknowledgments This work was supported by grants to PDNH from NSERC and from the government of Canada through Genome Canada and the Ontario Genomics Institute, (SCALE) research federation through the "Functions and Determinants of Biodiversity" (BIODIV) program. D. Porco was supported by post-doctoral fellowships grant from the Conseil Régional de Haute Normandie and from NSERC. C. Erséus was supported by the Swedish Taxonomy Initiative

(ArtDatabanken), and the Adlerbert Research Foundation. S. James was supported by a Marie Curie France Regions fellowship award to the Laboratoire EA 1293 ECODIV at the University of Rouen.

References

- Addison J (2009) Distribution and impacts of invasive earthworms in Canadian forest ecosystems. *Biol Invasions* 11: 59–79
- Amend AS, Seifert KA, Bruns TD (2010) Quantifying microbial communities with 454 pyrosequencing: does read abundance count? *Mol Ecol* 19:5555–5565
- Armstrong KF, Ball SL (2005) DNA barcodes for biosecurity: invasive species identification. *Phil Trans R Soc B-Biol Sci* 360:1813–1823
- Bellinger PF, Christiansen KA, Janssens F (2011) Checklist of the Collembola of the World. In: <http://www.collembola.org>
- Burtelow AE, Bohlen PJ, Groffman PM (1998) Influence of exotic earthworm invasion on soil organic matter, microbial biomass and denitrification potential in forest soils of the northeastern United States. *Appl Soil Ecol* 9:197–202
- Cameron EK, Bayne EM, Clapperton MJ (2007) Human-facilitated invasion of exotic earthworms into northern boreal forests. *Ecoscience* 14:482–490
- Cameron EK, Bayne EM, Coltman DW (2008) Genetic structure of invasive earthworms *Dendrobaena octaedra* in the boreal forest of Alberta: insights into introduction mechanisms. *Mol Ecol* 17:1189–1197
- Chahartaghi M (2007) Trophic niche differentiation, sex ratio and phylogeography of European Collembola. Ph.D. Thesis
- Chang CH, Rougerie R, Chen JH (2009) Identifying earthworms through DNA barcodes: pitfalls and promise. *Pedobiologia* 52:171–180
- Chown SL, Sinclair BJ, van Vuuren BJ (2008) DNA barcoding and the documentation of alien species establishment on sub-Antarctic Marion Island. *Polar Biol* 31:651–655
- Christiansen K, Bellinger P (1998) The collembola of North America North of the Rio Grande. A taxonomic analysis. Grinnell College, Grinnell, pp 1–1520
- Convey P, Greenslade P, Arnold RJ et al (1999) Collembola of sub-Antarctic South Georgia. *Polar Biol* 22:1–6
- Darling JA, Blum MJ (2007) DNA-based methods for monitoring invasive species: a review and prospectus. *Biol Invasions* 9:751–765
- Decaëns T, Jiménez JJ, Gioia C et al (2006) The values of soil animals for conservation biology. *Eur J Soil Biol* 42:S23–S38
- Decaëns T, Porco D, Rougerie R et al. (2012) Potential of DNA barcoding for earthworm research in taxonomy and ecology. *Appl Soil Ecol* (in press)
- deWard JR, Landry JF, Schmidt BC et al (2009) In the dark in a large urban park: DNA barcodes illuminate cryptic and introduced moth species. *Biodivers Conserv* 18:3825–3839
- Eisenhauer N, Partsch S, Parkinson D et al (2007) Invasion of a deciduous forest by earthworms: changes in soil chemistry,

- microflora, microarthropods and vegetation. *Soil Biol Biochem* 39:1099–1110
- Folmer O, Black M, Hoeh W et al (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotech* 3:294–299
- Gabriel AGA, Chown SL, Barendse J et al (2001) Biological invasions of Southern Ocean islands: the Collembola of Marion Island as a test of generalities. *Ecography* 24: 421–430
- Gates GE (1966) Requiem for megadrile utopias. A contribution toward the understanding of the earthworm fauna of North America. *Proc Biol Soc Wash* 79:239–254
- Greenslade P (2002a) Assessing the risk of exotic Collembola invading subantarctic islands: prioritising quarantine management. *Pedobiologia* 46:338–344
- Greenslade P (2002b) Systematic composition and distribution of Australian cave collembolan faunas with notes on exotic taxa. *Helictite* 38:11–16
- Greenslade P (2008) Has survey effort of Australia's islands reflected conservation and biogeographical significance? An assessment using Collembola. *Eur J Soil Biol* 44:458–462
- Greenslade P, Convey P (2012) Exotic Collembola on subantarctic islands: pathways, origins and biology. *Biol Invasions* 14:405–417
- Greenslade P, Wise KAJ (1984) Additions to the Collembolan fauna of the Antarctic. *Trans R Soc S Aust* 108:203–206
- Greenslade P, Simpson JA, Grgurinovic CA (2002) Collembola associated with fungal fruit-bodies in Australia. *Pedobiologia* 46:345–352
- Groffman PM, Bohlen PJ, Fisk MC et al (2004) Exotic earthworm invasion and microbial biomass in temperate forest soils. *Ecosystems* 7:45–54
- Hajibabaei M, DeWaard JR, Ivanova NV et al (2005) Critical factors for assembling a high volume of DNA barcodes. *Phil Trans R Soc B-Biol Sci* 360:1959–1967
- Hajibabaei M, Janzen DH, Burns JM et al (2006) DNA barcodes distinguish species of tropical Lepidoptera. *Proc Natl Acad Sci USA* 103:968–971
- Hale CM, Frelich LE, Reich PB et al (2008) Exotic earthworm effects on hardwood forest floor, nutrient availability and native plants: a mesocosm study. *Oecologia* 155:509–518
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98
- Hebert PDN, Cywinska A, Ball SL et al (2003) Biological identifications through DNA barcodes. *Proc R Soc Lond Ser B-Biol Sci* 270:313–321
- Hendrix PF, Callahan MA, Drake JM et al (2008) Pandora's Box contained Bait: the global problem of introduced earthworms. *Annu Rev Ecol Evol Syst* 39:593–613
- Hogg ID, Hebert PDN (2004) Biological identification of springtails (Hexapoda : Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. *Can. J. Zool.-Rev. Can Zool* 82:749–754
- Holdsworth AR, Frelich LE, Reich PB (2008) Litter decomposition in earthworm-invaded northern hardwood forests: role of invasion degree and litter chemistry. *Ecoscience* 15:536–544
- Hopkin SP (1997) *Biology of the springtails (Insecta: Collembola)*. Oxford University Press, New York, Tokyo
- Ivanova NV, Dewaard JR, Hebert PDN (2006) An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Mol Ecol Notes* 6:998–1002
- James SW, Porco D, Decaëns T et al (2010) DNA barcoding reveals cryptic diversity in *Lumbricus terrestris* L., 1758 (Clitellata): resurrection of *L. herculeus* (Savigny, 1826). *PLoS ONE* 5:e15629
- Janion C, Worland MR, Chown SL (2009) Assemblage level variation in springtail lower lethal temperature: the role of invasive species on sub-Antarctic Marion Island. *Physiol Entomol* 34:284–291
- Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide-sequences. *J Mol Evol* 16:111–120
- King KL, Greenslade P, Hutchinson KJ (1985) Collembolan associations in natural versus improved pastures of the New-England tableland, NSW: distribution of native and introduced species. *Aust J Ecol* 10:421–427
- King RA, Tibble AL, Symondson WOC (2008) Opening a can of worms: unprecedented sympatric cryptic diversity within British lumbricid earthworms. *Mol Ecol* 17:4684–4698
- Letunic I, Bork P (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23:127–128
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452
- Maynard EA (1951) *A monograph of the Collembola or springtail insects of New York state*. Comstock Publishing Company Inc, Ithaca
- McLean MA, Parkinson D (2000) Field evidence of the effects of the epigeic earthworm *Dendrobaena octaedra* the microfungus community in pine forest floor. *Soil Biol Biochem* 32:351–360
- McLean MA, Migge-Kleian S, Parkinson D (2006) Earthworm invasions of ecosystems devoid of earthworms: effects on soil microbes. *Biol Invasions* 8:1257–1273
- Myburgh M, Chown SL, Daniels SR et al (2007) Population structure, propagule pressure, and conservation biogeography in the sub-Antarctic: lessons from indigenous and invasive springtails. *Divers Distrib* 13:143–154
- Olden JD, Poff NL, Douglas MR et al (2004) Ecological and evolutionary consequences of biotic homogenization. *Trends Ecol Evol* 19:18–24
- Oliver I, Garden D, Greenslade PJ et al (2005) Effects of fertiliser and grazing on the arthropod communities of a native grassland in South-Eastern Australia. *Agric Ecosyst Environ* 109:323–334
- Pérez-Losada M, Ricoy M, Marshall JC et al (2009) Phylogenetic assessment of the earthworm *Aporrectodea caliginosa* species complex (Oligochaeta: Lumbricidae) based on mitochondrial and nuclear DNA sequences. *Mol Phylogenet Evol* 52:293–302
- Pearce TG (1972) Acid intolerant and ubiquitous Lumbricidae in selected habitats in North Wales. *J Anim Ecol* 41:397
- Porco D, Bedos A, Deharveng L (2010a) Description and DNA barcoding assessment of the new species *Deutonura gibbosa* (Collembola: Neanuridae: Neanurinae), a common springtail of Alps and Jura. *Zootaxa* 2639:59–68
- Porco D, Rougerie R, Deharveng L et al (2010b) Coupling non-destructive DNA extraction and voucher retrieval for small

- soft-bodied Arthropods in a high-throughput context: the example of Collembola. *Mol Ecol Resour* 10:942–945
- Porco D, Potapov M, Bedos A et al (2012) Cryptic diversity in the ubiquitous species *Parisotoma notabilis* (Collembola, Isotomidae): a long used chimeric species? *PLoS One* (in press)
- Potapov M (2001) Isotomidae. Staatliches Museum für Naturkunde Görlitz
- Primack RB (2000) A primer of conservation biology, 2nd edn. Sinauer Associates, Sunderland
- Reynolds JW (1973) The earthworms of Connecticut (Oligochaeta: Lumbricidae, Megascolecidae and Sparganophilidae). *Megadrilogica* 1:1–4
- Richard B, Decaëns T, Rougerie R et al (2010) Re-integrating earthworm juveniles into soil biodiversity studies: species identification through DNA barcoding. *Mol Ecol Resour* 10:606–614
- Rodman JE, Cody JH (2003) The taxonomic impediment overcome: NSF's partnerships for enhancing expertise in taxonomy (PEET) as a model. *Syst Biol* 52:428–435
- Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions. *Trends Ecol Evol* 22:454–464
- Rougerie R, Decaëns T, Deharveng L et al (2009) DNA barcodes for soil animal taxonomy. *Pesqui Agropecu Bras* 44:789–802
- Saitou N, Nei M (1987) The neighbor-joining method—a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Salmon JT (1941) The Collembolan Fauna of New Zealand, including a discussion of its distribution and affinities. *Trans R Soc N Z* 70:282–431
- Slabber S, Worland MR, Leinaas HP et al (2007) Acclimation effects on thermal tolerances of springtails from sub-Antarctic Marion Island: indigenous and invasive species. *J Insect Physiol* 53:113–125
- Stach J (1966) On the Collembola of Newfoundland and Nova Scotia. *Acta Zool Crac* 11:211–222
- Suárez ER, Pelletier DM, Fahey TJ et al (2004) Effects of exotic earthworms on soil phosphorus cycling in two broadleaf temperate forests. *Ecosystems* 7:28–44
- Suárez ER, Fahey TJ, Yavitt JB et al (2006) Patterns of litter disappearance in a northern hardwood forest invaded by exotic earthworms. *Ecol Appl* 16:154–165
- Tamura K, Dudley J, Nei M et al (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
- Tiunov AV, Hale CM, Holdsworth AR et al (2006) Invasion patterns of Lumbricidae into the previously earthworm-free areas of northeastern Europe and the western Great Lakes region of North America. *Biol Invasions* 8:1223–1234
- Van De Wiel CCM, Van Der Schoot J, Van Valkenburg J et al (2009) DNA barcoding discriminates the noxious invasive plant species, floating pennywort (*Hydrocotyle ranunculoides* L.f.), from non-invasive relatives. *Mol Ecol Resour* 9:1086–1091
- Vazquez DP, Simberloff D (2001) Taxonomic selectivity in surviving introduced insects in the United States. Kluwer Academic/Plenum Publ, New York
- Whalen JK (2004) Spatial and temporal distribution of earthworm patches in corn field, hayfield and forest systems of southwestern Quebec, Canada. *Appl Soil Ecol* 27:143–151
- Wilson EO (2002) The future of life. Vintage Books, New York
- Winter M, Kuhn I, La Sorte FA et al (2010) The role of non-native plants and vertebrates in defining patterns of compositional dissimilarity within and across continents. *Glob Ecol Biogeogr* 19:332–342
- Wironen M, Moore TR (2006) Exotic earthworm invasion increases soil carbon and nitrogen in an old-growth forest in southern Quebec. *Can J For Res—Rev Can Rech For* 36:845–854
- Womersley H (1939) Primitive insects of South Australia. Silverfish Springtails and their allies, Adelaide
- Yeates GW (1991) Impact of historical changes in land-use on the soil fauna. *N Z J Ecol* 15:99–106
- Yeates GW, Hawke MF, Rijkse WC (2000) Changes in soil fauna and soil conditions under *Pinus radiata* agroforestry regimes during a 25-year tree rotation. *Biol Fertil Soils* 31:391–406
- Yosii R (1977) Critical checklist of the Japanese species of Collembola. *Contrib Biol Lab Kyoto Univ* 25:141–170