Applied Soil Ecology 65 (2013) 35–42

Contents lists available at SciVerse ScienceDirect

Applied Soil Ecology

journal homepage: www.elsevier.com/locate/apsoil

Review

Potential of DNA barcoding for earthworm research in taxonomy and ecology

Thibaud Decaëns[a,*], David Porco[a,b], Rodolphe Rougerie[a,b], George G. Brown[c], Samuel W. James[d]

[a] Laboratoire d'Ecologie, EA 1293 ECODIV, FED SCALE, UFR Sciences et Techniques, Université de Rouen, F-76821 Mont Saint Aignan cedex, France
[b] Canadian Centre for DNA Barcoding, Biodiversity Institute of Ontario, University of Guelph, 579 Gordon Street, Guelph, Ontario, N1G 2W1, Canada
[c] EMBRAPA Floresta, Estrada da Ribeira, km. 111, Caixa-Postal: 319, CEP 83411-000, Curitiba, PR, Brazil
[d] Department of Biology, University of Iowa, Iowa City, IA 52242, USA

ARTICLE INFO

Article history:
Received 4 January 2011
Received in revised form
29 December 2012
Accepted 3 January 2013

Keywords:
Earthworm
DNA barcoding
Cryptic diversity
Species identification
Taxonomic crisis
Taxonomic deficit

ABSTRACT

The biodiversity of soil animal communities is still poorly known. Most taxa, from the smaller body-sized to the large invertebrates of the macrofauna, suffer a strong taxonomic deficit. Earthworms comprise about 3700 described species, but this number probably only represents half of the actual worldwide diversity of the group. In many cases, earthworm species identification is impeded by the lack of stable and easily observable morphological characters, a high level of phenotypic variability, and the lack of diagnostic characters in juvenile stages. Another problem is the high level of expertise required for these identifications, in addition to the lack of expert identification services. These limitations are a serious issue in studies that focus on this group and which require reliable identifications and/or species lists (e.g. taxonomy, biogeography, community ecology, etc.). DNA barcoding, the use of a short DNA fragment as a genetic tag for species identification, offers both a better circumscription of species and a solution to streamline identifications. Preliminary studies have demonstrated the value of this approach for species discrimination, identification of new taxa, identification of juveniles, detection of cryptic diversity, and rapid surveys of biodiversity at different spatial scales. In this review, we illustrate these aspects with examples taken from published studies as well as from unpublished preliminary results of the "Earthworm Barcode of Life" (EarthwormBOL) campaign of the "International Barcode of Life" initiative (iBOL).

Contents

1. Introduction

Soils probably represent one of the most diversified habitats of terrestrial ecosystems, with their biota comprising representatives of almost all major taxa and trophic groups that form terrestrial biodiversity (Decaëns et al., 2006; Decaëns, 2010; Swift et al., 1979; Wolters, 2001). At a local scale, a single gram of soil may contain several thousands of different bacterial and fungal genotypes, and a square meter of soil may host several hundreds of arthropods species (Hawksworth, 2001; Schaefer and Schauermann, 1990; Torsvik et al., 1994). Considered in a global perspective, soil organisms represent about 25% of the described forms of life on earth (Decaëns et al., 2006; Wardle, 2002). On the other hand, soil organisms have weakly engaged the attention of taxonomists

* Corresponding author. Tel.: +33 235146771.
E-mail address: thibaud.decaens@univ-rouen.fr (T. Decaëns).

0929-1393/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.apsoil.2013.01.001

compared with other groups of aboveground organisms (Wolters, 2001), and only a very small part of their taxonomic diversity has been formally documented (Decaëns et al., 2006). The paucity of our taxonomic knowledge for most soil taxonomic groups has been repeatedly stressed in the scientific literature (André et al., 2001, 2002; Behan-Pelletier, 1999; Brussaard et al., 1997; Dance, 2008; Decaëns, 2010; Decaëns et al., 2006; Giller, 1996; Wall et al., 2001; Wolters, 2001) and soils are considered as the last biotic frontier after oceanic abysses and tropical forest canopies (André et al., 1994; Hagvar, 1998).

Earthworms are generally considered as a relatively well known group of soil invertebrates from both ecological and taxonomical perspectives. These relatively large animals have been extensively studied in most soil types worldwide, where they often represent a significant if not the dominant part of the soil biomass (Lavelle and Spain, 2001). They drew scientific attention due to their ecological importance as major regulators of important soil processes, or because of their economical importance in agroecosystems and/or in commercial markets where some species are sold for different purposes (Decaëns et al., 2006; Lavelle et al., 2006; Tomlin, 1983). Many species are reported as invasive in different regions of the world and are currently monitored in order to predict their potential impact on soil functioning (Addison, 2009; Bohlen et al., 2004; Fragoso et al., 1997; Gonzalez, 2006; Hale et al., 2005; Hendrix and Bohlen, 2002). Species such as *Lumbricus terrestris* L. have also been used as model organism in education and in many domains of biological science such as molecular biology and physiology (James et al., 2010). This strong interest is supposed to have benefited earthworm taxonomy, and as a consequence these organisms are generally less affected – or considered so – by the taxonomic deficit deplored in poorly studied groups of smaller soil organisms. About 3700 species have been described so far, a number supposed to represent a significant proportion of the estimated total biodiversity of the group (ca. 6000 species; Fragoso et al., 1997; Reynolds and Cook, 1976). The actual taxonomic deficit, however, may be higher than this optimistic expectation. This is suggested by the high levels of local endemism and undescribed species found in inter-tropical areas such as the Amazonian basin (Lavelle and Lapied, 2003). The recent use of molecular approaches to explore earthworm biodiversity also revealed that many described species comprise several distinct genetic lineages which may represent cryptic species (Chang et al., 2009; Huang et al., 2007; James et al., 2010; King et al., 2008; Rougerie et al., 2009). These recent findings suggest that our taxonomic knowledge of the group may not be as accurate as initially thought.

The existence of this taxonomic impediment is likely to be responsible for significant prejudices in all the domains of earthworm research relying on accurate species identification. One possible solution to address this constraint is the use of molecular approaches for species identification. DNA barcoding, the use of a standard genetic marker for species identification, has been increasingly used in studies of the biodiversity of a range of biota (Hebert et al., 2003, 2004), and the potential of this approach to address the taxonomic impediment of soil fauna has been recently stressed by several authors (Chang et al., 2009; Decaëns et al., 2008; Rougerie et al., 2009). Rougerie et al. (2009) provided a first overview of the value of this approach for earthworm species discrimination, identification of new taxa, species identification of juveniles and detection of cryptic diversity. Since this first general publication, a number of studies have been published that illustrate this usefulness and highlight how DNA barcoding can help opening new avenues in the domain of earthworm ecology, phylogeography and population genetics. In this review, we illustrate these aspects with examples taken from published studies as well as from unpublished preliminary results of the "Barcoding Earthworm" project (EarthwormBOL) campaign of the "International Barcode of Life" initiative (iBOL).

2. DNA barcoding

DNA barcoding is the use of a standardized region of 658 bp of the mitochondrial gene cytochrome c oxidase I (COI) for species discrimination (Hebert et al., 2003). The advantages of the method are multiple (Rougerie et al., 2009): (1) it is a testable and reproducible system as a link is maintained between any barcode and a voucher specimen; (2) for massive routine identifications it is in most cases faster and cheaper than traditional morphological identifications; (3) it is accessible for everybody and in any place where sequencing facilities exist; (4) it works for any life-stage and any kind of organic tissue types. The usefulness of DNA barcoding for the study of biodiversity (from species inventories to alpha taxonomy) at different levels of taxonomic resolution has now been revealed in a broad range of taxonomic groups of vertebrates and invertebrates (see Rougerie et al. (2009) for an extensive list of examples). On the other hand, some limitations and pitfalls of using a single genetic marker for species discrimination were pointed out (Rubinoff and Holland, 2005; Trewick, 2007; Wiemers and Fiedler, 2007; Will et al., 2005). The main caveats regard (1) potential false negatives, i.e. identical DNA barcodes in two actually different species due to short divergence time preventing the fixation of substitutions or to gene introgression; (2) potential false positives, i.e. different DNA barcodes between individuals belonging to the same species because of ancestral polymorphism or again genetic introgression. One of the most serious issues is the potential amplification of nonfunctional nuclear copies (Berthier et al., 2011) which can be overcomed by a posteriori quality control (Song et al., 2008) and a priori laboratory technics (Calvignac et al., 2011). The best solution to overcome these pitfalls is to use DNA barcodes in combination with other sets of data such as morphology, additional nuclear genetic markers, or ecological, ethological and biogeographical features (Rougerie et al., 2009). When these complementary data sets are not available, DNA barcoding should be used with the necessary caution relative to the use of a single marker.

DNA barcoding is supported by different national and international initiatives, of which the International Barcode of Life project (iBOL, http://ibol.org/) is the most ambitious in term of investment toward building large and comprehensive DNA barcode libraries, and also in setting a broad international framework for this endeavor, with more than 25 countries involved. The sequences generated are collectively compiled in a central integrative bioinformatics platform, the Barcode of Life Data System (BOLD, http://www.barcodinglife.org; Ratnasingham and Hebert, 2007). This database is also a scientific workbench supporting all phases of the analytical pathway from specimen collection to tightly validated barcode library (Ratnasingham and Hebert, 2007). Taxonomists play a key role in the assembly and curation of these reference libraries, thanks to their unique ability to scrutinize the results of DNA barcoding analyses and to use them when relevant within the classical workflow of new species descriptions and taxonomic revisions (Decaëns and Rougerie, 2008; Fisher and Smith, 2008; James et al., 2010; Porco et al., 2010; Stoev et al., 2010; Vaglia et al., 2008). This constitutes the cornerstone of a system enabling reliable identification through DNA barcodes, simply by comparing the barcode obtained from an unidentified specimen with those of formally named reference specimens in the library. The reference DNA sequences imbed taxonomic expertise and ensure the legacy of the incorporated expertise independently of further input from the taxonomists (Rougerie et al., 2009).

Earthworms were identified as the sentinel lineage for underground life forms within the "terrestrial biosurveillance" working