European Journal of Soil Biology 64 (2014) 23–29

Contents lists available at ScienceDirect

European Journal of Soil Biology

journal homepage: http://www.elsevier.com/locate/ejsobi

Original article

Mitochondrial DNA and morphological variation in the sentinel earthworm species *Lumbricus rubellus*

Robert K. Donnelly [a, b, *], Georgina L. Harper [a], A. John Morgan [b], Gabriela A. Pinto-Juma [b], Michael W. Bruford [b]

[a] Faculty of Health, Sport and Science, University of Glamorgan, Llantwit Road, Trefforest, Mid Glamorgan CF37 1DL, UK
[b] Cardiff School of Biosciences, Cardiff University, PO Box 915, Cardiff CF10 3TL, UK

ARTICLE INFO

Article history:
Received 6 May 2014
Received in revised form
24 July 2014
Accepted 25 July 2014
Available online 28 July 2014

Keywords:
Lumbricus rubellus
Earthworm
Cryptic lineages
mtDNA
Cytochrome oxidase I
Morphology

ABSTRACT

The common epigeic earthworm *Lumbricus rubellus* has been found to comprise two genetically distinct lineages and this genetic heterogeneity could compromise its currently widespread use as a sentinel in soil ecotoxicology. An extensive analysis of lineage diversity was conducted on UK populations of *L. rubellus*, comprising 137 individuals collected from England and Wales. Sequencing of the mitochondrial COI region revealed the widespread occurrence of the two described lineages throughout Britain, which were often found to co-exist at the same site. Morphological characters were investigated to differentiate the two lineages. A rapid genetic test (mitochondrial PCR amplification and restriction digestion) was applied to determine the lineage of each specimen. A blind trial revealed a characteristic glandular tumescence to be effective in differentiating the lineages, particularly in identifying 'lineage B' worms. COI sequence analysis was also conducted upon three other *Lumbricus* species, which failed to uncover any genetic lineages comparable to those found within *L. rubellus*.

1. Introduction

The increasing evidence that components of soil fauna are sensitive to a variety of anthropogenic activities [30,41] promotes the need for a more accurate understanding of species diversity and its relationship to ecosystem services [37]. Overlooking cryptic species leads to underestimation of ecosystem complexity, and could potentially confound biomonitoring and conservation initiatives [5]. Therefore a reliance on the use of only morphological criteria to identify species, especially in taxa such as earthworms where juvenile life-stages present particular difficulties and others where good keys and taxonomic expertise is limiting [12,38], has resulted in a change towards the widespread application of molecular genetic approaches [16].

The application of molecular techniques to taxonomy has revealed that many common invertebrate morphotypes may actually comprise highly divergent cryptic lineages [13,17,21,22,35,42]. The phenomenon appears to be especially prevalent in sexually reproducing lumbricid earthworms [25,27,28,32,33].

Phylogenetic analyses of European lumbricid earthworms using mitochondrial barcoding primers have revealed evidence of deeply-divergent lineages in several species: *Allolobophora chlorotica*, *Aporrectodea longa*, *Aporrectodea rosea*, *Aporrectodea trapezoides*, *Lumbricus rubellus* and *Lumbricus terrestris* [1,15,18,25,27,28,32,33]. Two distinct genetic lineages have been uncovered within UK populations of *L. rubellus*, with as many as four other identifiable lineages in mainland Europe (P. Sechi, pers. comm.), each displaying sequence divergences comparable to distinct species. Moreover, the two UK lineages have been confirmed as being present in ecotoxicity studies employing combinations of mitochondrial and nuclear markers [2,14,24]. These high levels of genetic divergence within *L. rubellus* could have major implications for the role of this epigeic species as a biomonitoring sentinel, especially in highly discriminatory 'omics' studies [4,10]. For example [11], emphasised that the substantial genetic heterogeneity that often exists in field populations should be fully considered in the design and interpretation of field-based ecotoxicological assessments. Studies demonstrating genotype-related differential sensitivities and resilience to toxicant exposures highlight the need to account for this variation [3,34,39,44].

From the above examples, the practical and interpretive problems associated with the inclusion of undetected cryptic speciation in molecular-genetic monitoring regimes are clear. For this reason

* Corresponding author. Present address: Cardiff School of Biosciences, Cardiff University, PO Box 915, Cardiff CF10 3TL, UK.
E-mail address: robertkdonnelly@gmail.com (R.K. Donnelly).

http://dx.doi.org/10.1016/j.ejsobi.2014.07.002
1164-5563/

[2], advocated that every individual sampled from wild populations should be genotyped prior to eco-toxicogenomic assay. However this will often be impractical due to logistic and financial constraints in routine monitoring, therefore it could be highly advantageous to be able to identify cryptic species in the field The recognition of putative *L. rubellus* sub-species based on minor differences in external morphological traits [7] highlights that it might be similarly possible to discriminate between individuals of the two UK *L. rubellus* lineages.

The first aim of this study was to apply 'universal' COI primers [19] to a relatively large number of *L. rubellus* from a variety of habitat types across a wide geographic range to evaluate genetic diversity within the two lineages to see whether *L. rubellus* includes more cryptic diversity than previously recorded in the UK. Mitochondrial analysis was also applied to individuals of three congeners (*Lumbricus castaneus*, *Lumbricus festivus*, *L. terrestris*) in order to establish whether the lineage diversification so commonly encountered in *L. rubellus* is a feature shared by other members of the genus. This is especially relevant in a European context, since studies have recently indicated the presence of cryptic genetic variation within *L. terrestris* sampled from Canada, the US and mainland Europe [28].

The second aim of this study was to determine whether variation in the anterior segment number (an evolutionarily conserved trait widely used in earthworm taxonomy [43]), and anterior-located glandular tumescences are distinguishing morphological features in *L. rubellus* lineages. A blind trial was conducted to determine if these traits could be successfully applied to differentiate between the lineages in the field, with a lineage-specific mitochondrial Restriction Fragment-Length Polymorphism (RFLP) test subsequently used to ascertain whether the inferences based on morphology were reliable.

2. Material and methods

2.1. MtDNA analysis

Specimens of four *Lumbricus* species were collected from a field site located at Pontcanna fields, Cardiff (Tables 1 and 2). Further collections were conducted at two sites at Clydach, Swansea, two sites located in Mid-Wales (Wemyss and Ystwyth Source) and two sites in South-West England (Glovers Field and Hallen Hill). The sample sites represented a range of different habitats including forests, pastureland and parkland (Table 2).

Approximately 25 mg of earthworm tissue was used for each DNA extraction. DNA was extracted using a QIAGEN DNeasy tissue extraction kit (QIAGEN, UK). A 580 bp sub-unit of the mitochondrial cytochrome oxidase I gene (COI) was then PCR-amplified using the general invertebrate primers LCO1490 (5′-GGTCAACAAATCATAAAGATATTGG-3′) and HCO2198 (5′-TAAACTTCAGGGTGACCAAAAAATCA-3′), [19]. PCR reactions were performed using the reaction volumes and times suggested by Ref. [25]. All reactions were performed using a GeneAmp PCR System 9700 (Applied Biosystems, UK).

Table 1
Number of sampled individuals at each site.

Collection site	Site code	*L. castaneus*	*L. festivus*	*L. rubellus*	*L. terrestris*
Pontcanna fields, Cardiff	PT	17	25	25	22
Clydach (site 1), Swansea	CL1	0	0	22	0
Clydach (site 2), Swansea	CL2	0	11	22	9
Glovers Field, Somerset	GF	0	0	19	0
Hallen Hill, Bristol	HH	0	0	9	0
Ystwth Source, Powys	YS	0	0	10	0
East Cottage, Ceredigion	EC	0	0	16	0
Wemyss, Ceredigion	WM	0	0	14	0
Total		17	36	137	31

Table 2
Location and description of sample sites.

Site	Latitude	Longitude	Site description
Glovers field	51:18:42N	02:47:33W	Rough pastureland near abandoned Zn mine
Hallen Hill	51:30:57N	02:39:28W	Deciduous woodland near abandoned Pb/Zn smelter
Ystwyth Source	52:21:57N	03:41:44W	Stream bank
East Cottage	52:21:35N	03:44:59W	Derelict building at abandoned Pb/Zn mine
Wemyss	52:20:58N	03:53:14W	Stream bank near abandoned Pb/Zn mine
Clydach 1	51:41:47N	03:53:18W	Deciduous woodland and rough grassland close to Ni smelter
Clydach 2	51:41:48N	03:52:49W	Playing field
Dinas Powys 6/Pontcanna	51:29:37N	03:12:18W	Mound of topsoil leaves on mowed parkland

The general COI primers were found to successfully amplify only two of the *L. rubellus* lineage B individuals. A primer pair was therefore developed based upon the COI sequences from these two individuals, which successfully amplified a smaller 356 bp fragment of the COI region of the lineage B individuals (LRB COIF 5′-TCTTCTTTCTTGTCATGCCTGT-3′, LRB COIR 5′-TGAAGTATTTAGATTTCGGTCAGTT-3′). PCR reactions were again performed using the reaction volumes of [25]. All samples were initially denatured for 2.5 min at 94 °C with 35 cycles of 94 °C for 30s, 45 °C for 30s and 72 °C for 45s, followed by 72 °C for 10 min. Following PCR amplification, PCR products were purified and sequenced in accordance with the procedure detailed by Ref. [14].

Sequencher 3.1.2. (Gene Codes Corporation, USA) and CodonCode Aligner 3.0.1 (CodonCode Corporation, USA) were used to align and edit all sequences and to determine COI haplotypes for each species. The program MEGA 4.0 [45] was used to compute pairwise genetic distances between all haplotypes. Pairwise distances were calculated both as uncorrected p-distances and also using the Tamura-Nei model of substitution (gamma shape distribution parameter = 0.28), selected on the basis of hierarchical likelihood ratio tests performed using the program jModelTest [36]. A maximum-likelihood phylogeny was produced using the program PHYML [20]. This was based upon Tamura-Nei corrected distances between haplotypes with a bootstrap replication of 1000. A Bayesian phylogenetic analysis of haplotypes was also conducted using MrBayes 3.1.2 [40]. The general time reversible model was selected for this analysis with gamma and invariable sites. Bayesian analysis was run using two independent runs consisting of four chains. These analyses were run for 400,000 generations with trees saved every 100 generations. Following the analysis, 1000 trees were removed as burn-in. Consensus trees were viewed using MEGA 4.0 [45].

2.2. Blind trial

One hundred and five *L. rubellus* individuals were collected from four field sites located across South Wales. The sites were selected by the collectors to include populations of both mitochondrial lineages of *L. rubellus* based upon the findings of previous studies. All earthworms were placed individually into polythene bags labelled with an identifying code. The code was initially known only to the collector, allowing a blind trial to be conducted by the experimenter.