UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS BIOLÓGICAS

TESIS DOCTORAL

Phylogeography and Biology of Eiseniella tetraedra (Savigny, 1826)

Filogeografía y Biología de Eiseniella tetraedra (Savigny, 1826)

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Irene de Sosa Carrasco

Directores

Darío Díaz Cosín
Ana Almodóvar Pérez
Marta Novo Rodríguez

Madrid

© Irene de Sosa Carrasco, 2022

Universidad Complutense de Madrid
Facultad de Ciencias Biológicas

Phylogeography and Biology of *Eiseniella tetraedra* (Savigny, 1826)

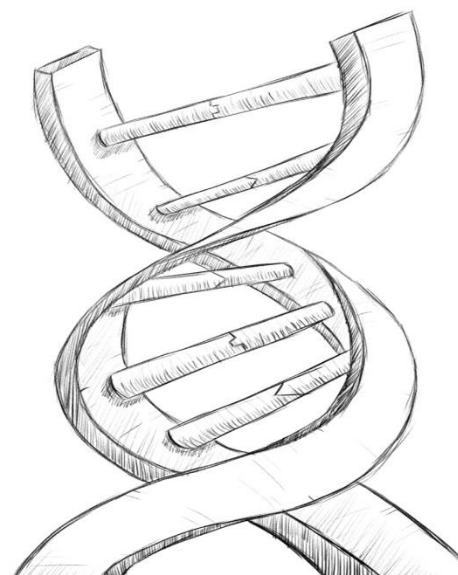
Filogeografía y Biología de *Eiseniella tetraedra* (Savigny, 1826)

Tesis Doctoral

Autora:
Irene de Sosa Carrasco

Directores:
Darío Díaz Cosín, Ana Almodóvar Pérez, Marta Novo Rodríguez

2022



UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS BIOLÓGICAS

Phylogeography and biology of *Eiseniella tetraedra* (Savigny, 1826)

Filogeografía y biología de *Eiseniella tetraedra* (Savigny, 1826)

MEMORIA PARA OPTAR AL TÍTULO DE DOCTOR

PRESENTADA POR

Irene de Sosa Carrasco

Los directores:

Darío Díaz Cosín   Ana Almodóvar Pérez   Marta Novo Rodríguez

Madrid, 2022

©Irene de Sosa Carrasco, 2022

La presente Tesis Doctoral ha sido financiada por:

- Ayuda para contratos predoctorales de personal investigador en formación que forma parte del Programa de Formación de Personal Investigador de la Universidad Complutense de Madrid (UCM) y del Programa de Financiación de la UCM y el Banco Santander (CT42/18-CT43/18).

Los estudios que se desarrollan en esta Tesis Doctoral han sido financiados por los siguientes proyectos de investigación:

- CGL2013-42908-P: Filogeografía y especiación críptica en dos grupos de invertebrados de vagilidad reducida: oligoquetos y kinorrincos- Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia; Subprograma Estatal de Generación de Conocimiento.
- FAUNA IBERICA XII. Lumbricidae, Hormogastridae (Oligochaeta, Lumbricoidea). Programa Estatal de Generación del Conocimiento y Fortalecimiento Científico y Tecnológico del Sistema de I+D+i.

Agradecimientos

Pues parece que la tesis si tiene fin, y se va acercando. Hoy no estoy muy motivada para seguir escribiendo, así que me he dicho “¡vamos a por los agradecimientos!”, pero no sé yo qué requiere más inspiración. A lo largo de estos años sois muchas las personas que me habéis ayudado y apoyado durante este proceso o que simplemente cuando me veíais me preguntabais un “¿qué tal va la tesis?”; sirva este párrafo inicial para un agradecimiento general a todas ellas.

Esta travesía comienza allá por 2013, cuando cursé la asignatura Análisis de la Biodiversidad Animal. Primero la salida a Galicia…espectacular (quien me diría que años después acabaría yendo como becaria). Posteriormente, el trabajo que teníamos que realizar en la asignatura. Lombrices de tierra, con Darío Díaz Cosín. He de admitir que al principio no estaba muy ilusionada, pero después de leer varios artículos de unas tal R. Fernández y M. Novo, aquello empezó a llamarme mucho la atención. Y así fue el principio de este fin. A la primera persona que quiero dar las gracias es a Darío. Jefe, eres de piedra. ¡Menos mal que siempre tus sustos quedan en nada! Espero verte en la defensa de esta tesis, aunque sea a través de la pantalla. Gracias por darme la oportunidad de trabajar en el grupo, y de enseñarme tanto a lo largo de todos estos años. Gracias también por proponerme trabajar con *Eiseniella*. Aunque a veces haya maldecido el no trabajar con una especie un poco más normalita, en los retos está lo interesante. Ana, gracias a ti también por tu ayuda en estos años. Tu visión de la biogeografía y la genética de poblaciones ha sido clave para todo lo que ha salido de esta tesis. También te agradezco tu dedicación, aún en momentos de estrés siempre estabas para una llamada telefónica que solventara las dudas. Por último, a esa tal M.

Novo de los artículos de ABA. Marta, gracias infinitas. Gracias por estar tan implicada y por aguantar mis malos ratos. Gracias por estar siempre presente. Queda pendiente un anexo en el que cuente los problemas a los que nos hemos tenido que enfrentar durante esta tesis, que no han sido pocos. Espero que nadie me quite el título de gafe nunca.

Después de agradecer a mis directores de tesis, me gustaría dedicar un párrafo a mi familia. A esa que nunca me dijo “¿pero en biología no hay mucha tasa de paro?” y siempre me apoyo en todo lo que quise hacer, tanto moral como económicamente durante los años que no tuve ni contrato, ni beca. Me acompañaron a muestrear por toda la Comunidad de Madrid, fin de semana sí y fin de semana también. Aunque he decir que el único que acabó sabiendo diferenciar las lombrices de esta especie en el campo fue El Niño. Gracias por vuestro apoyo incondicional.

He trabajado en esta tesis durante varios años, y siempre en la planta 10 de la facultad, en el Departamento de Biodiversidad, Ecología y Evolución. A lo largo de los años, han pasado muchos becarios por allí, y todos habéis aportado un granito de arena para que esto sea posible. Diego, has sido mi compañero de batalla todos estos años. Gracias por todos y cada uno de los días que hemos pasado mano a mano quejándonos de los días grises, tomando cafés y cerves en forestales y recorriéndonos Malasaña a golpe de gintonic y música de los ochenta. Dani, gracias por enseñarme tanto y por estar siempre para cualquier cosa, somos un buen tándem. Dani Romero, ¿qué habría sido de mi sin tu compañía y ayuda en todos los muestreos por España? Nuri, Italia quiere que volvamos a hacer un roadtrip juntas, gracias por ese viaje de muestreo tan productivo, divertido y lleno de sorpresas (muestreos en Pompeya, castillo de

Sant'Angelo y visitas a la embajada incluidas). Sheila, gracias por estar siempre presente en el labo, para ayudarme con cualquier problema que surgía en las PCRs y para reírnos contándonos nuestra vida. Fer, gracias por todos tus consejos y por todas esas cerves en forestales arreglando el mundo. Raúl, gracias por tus dibujos, ¡eres un artista! No me olvido de Sergio, Marisa y Joaquin. Ale, gracias por el trabajo que te ha supuesto revisar la tesis y por hacerlo con tanta dedicación. Quedan pendientes los hipervolúmenes. A las últimas incorporaciones, Sergio y Natasha (el RNA seq contigo se veía más fácil). Por supuesto agradecer a todos los profesores de la Cátedra de Invertebrados no Artrópodos su ayuda durante todos estos años: Dolores, Mónica, Juan, Fernando, Jesús, Carmen, Tito y Pablo (ojalá algún día ser como vosotros). Of course, some of this work would not be possible without the help of all the researchers who have provided me specimens from many parts of Europe: Dr. Erséus, Dr. Jelic, Dra. Jablonska, Dr. Csudzi, Dr. Lappied and Dra. Szederjesi, thank you all. In particular, I would like to thank Dr. Erseús for his cooperation and for hosting me in Gothenburg to study his collection of worms. His help was of fundamental importance.

Pero también mucha gente fuera de la facultad me ha dado su apoyo incondicional durante estos años de tesis y también quiero agradecérselo a ellos. Patio, que decir de vosotros. Siempre aquí, estemos donde estemos. Gracias Ainhoa, Euge, Fra, Justo, Marta. Siempre habéis tenido tiempo para ayudarme y escucharme, aunque no supierais ni de lo que estaba hablando. Miriam, Ángel, Eva y Edu, gracias por estar juntos desde la adolescencia, pasar juntos el instituto, la carrera y lo que nos echen. Nerea, gracias por acompañarme en este camino que está viendo su fin, que esta tesis esté escrita y yo no esté loca es gracias a ti.

Pero antes de que toda esta gran historia empezara, llegaste tú. Y estuviste, y estás, y estarás. Me rio al pensar que una de las primeras veces que quedábamos fue a coger lombrices en Toledo y acabamos hasta arriba de barro. Siempre has estado ahí. Acompañándome en durísimos viajes de muestreo (como aquel de Portugal, donde no teníamos tiempo ni para comer), machacando y tamizando suelo en el labo 21, aguantando historias interminables de cómo no me salían las PCRs, escuchando ilusionado cada vez que te contaba algún avance, ayudándome a maquetar la tesis, ofreciéndome un hombro donde llorar cuando no podía más y haciéndome reír cuando más lo necesitaba. Gracias por estos ocho años juntos y ya sabes… ¡Para nosotros es la vida! Te quiero Manué.

Caricature from the journal Punch, 1882

“*It may be doubted if there are any other animals which have played such an important part in the history of the world as these lowly organized creatures*." The Formation of Vegetable Mould, Through the Action of Worms. Charles Darwin, 1881.

Index

Abstract 7

Resumen 11

Introduction 15

Soil fauna: earthworms 15

How climate change affects earthworms? 18

The genus *Eiseniella* Michaelsen, 1900 20

Cosmopolitan and parthenogenetic earthworms: the case of *Eiseniella tetraedra* .. 22

Phylogeographic studies on earthworms 26

Objectives 29

Objectives of Chapter 1 29

Objectives of Chapter 2 30

Objectives of Chapter 3 30

Objectives of Chapter 4 30

Objectives of Chapter 5 30

1. Chapter 1: Bless this phylogeographic mess - Comparative study of *Eiseniella tetraedra* (Annelida, Oligochaeta) between an Atlantic area and a continental Mediterranean area in Spain. 33

Abstract 35

1.1 Introduction 37

1.2 Material and methods 40

1.2.1 Sampling and morphological studies.. 40

1.2.2 DNA extraction, gene amplification and sequencing 42

1.2.3 Data analysis 42

1.3 Results 43

1.3.1 Phylogenetic analysis.. 43

1.3.2 Haplotype and lineage distribution 45

1.3.3 Genetic diversity 46

1.3.4 Morphological studies 49

1.4 Discussion 49

1.5 Conclusions 53

2. Chapter 2: Phylogeography of *Eiseniella tetraedra* (Savigny, 1826) in the Iberian Peninsula shows a genetic distribution based on environmental factors. 55

Abstract 57

2.1 Introduction 59

2.2 Material and methods 61

2.2.1 Sampling and morphological studies.. 61

2.2.2 DNA extraction, gene amplification and sequencing 62

2.2.3 Genetic data analyses.. 63

2.2.4 Environmental factor analyses 64

2.2.5 Morphological analyses.. 65

2.3 Results 66

2.3.1 Phylogenetic analyses 66

2.3.2 Lineages distribution 67

2.3.3 Genetic diversity, genetic divergence and population structure 69

2.3.4 Historic demography 73

2.3.5 Environmental factors analyses 75

2.3.6 Morphological analyses 79

2.4 Discussion 80

2.5 Conclusions 84

3. Chapter 3: The nunatak and tabula rasa hypotheses may be compatible: the European phylogeography of a riparian earthworm 87

Abstract 89

3.1 Introduction 91

3.2 Material and methods 94

3.2.1 Earthworm sampling and morphological analyses 94

3.2.2 DNA extraction, gene amplification and sequencing 94

3.2.3 Phylogenetic analyses and genetic variability 95

3.2.4 Ecological niche modeling 96

3.3 Results 97

3.3.1 Phylogenetic analysis 97

3.3.2 Lineages distribution 99

3.3.3 Genetic diversity and genetic divergence 102

3.3.4 Ecological niche modeling 104

3.3.5 Morphological data 106

3.4 Discussion 106

3.5 Conclusions 111

4. Chapter 4: Back to the past: Revisiting *Eiseniella neapolitana* and *Norealidys andaluciana* (Annelida, Lumbricidae) corroborates Bouché's proposal of Norealidys. 113

Abstract 115

4.1 Introduction 117

4.2 Material and methods 118

4.2.1 Earthworm sampling and morphological studies 118

4.2.2 DNA extraction, amplification and sequencing 120

4.2.3 Data analyses 120

4.3 Results 122

4.3.1 Phylogenetic analyses 122

4.3.2 Genetic diversity and genetic divergence 124

4.3.3 Morphological studies 129

4.4 Discussion 132

4.5 Conclusions 136

5. Chapter 5: How to thrive in unstable environments: gene expression profile of a riparian earthworm under abiotic stress. 137

Abstract 139

5.1 Introduction 141

5.2 Material and methods 145

5.2.1 Experimental animals and laboratory conditions 145

5.2.2 Freezing experiment 145

5.2.3 Desiccation experiment 146

5.2.4 RNA extraction and species verification 148

5.2.5 Transcriptome sequencing 148

5.2.6 Establishing a reference transcriptome 149

5.2.7 Quantitative transcriptomic analysis 150

5.3 Results 151

5.3.1 Sequencing output 151

5.3.2 Differential gene expression under freezing conditions 151

5.3.3 Differential gene expression under desiccation conditions 155

5.4 Discussion 159

5.4.1 Cold tolerance 159

5.4.2 Desiccation tolerance 163

5.5 Conclusions 167

Discussion 169

Conclusions 173

Chapter 1: Bless this phylogeographic mess - Comparative study of *Eiseniella tetraedra* (Annelida, Oligochaeta) between an Atlantic area and a continental Mediterranean area in Spain 173

Chapter 2: Phylogeography of *Eiseniella tetraedra* (Savigny, 1826) in the Iberian Peninsula shows a genetic distribution based on environmental factors. 174

Chapter 3: The nunatak and tabula rasa hypotheses may be compatible: the European phylogeography of a riparian earthworm. 175

Chapter 4: Back to the past: Revisiting *Eiseniella neapolitana* and *Norealidys andaluciana* (Annelida, Lumbricidae) corroborates Bouché's proposal of Norealidys. 175

Chapter 5: How to thrive in unstable environments: gene expression profile of a riparian earthworm under abiotic stress. 176

References 177

Supplementary Material 219

Abstract

The aim of this Ph.D. dissertation is to increase the knowledge about cosmopolitan and parthenogenetic earthworms. For this purpose, the earthworm *Eiseniella tetraedra* (Savigny, 1826) was chosen as a model. Although there have been studies on other earthworm species with the same characteristics, no work has ever been done that includes so many individuals and populations. On the other hand, its relationship with other species of the same genus has been studied using molecular markers and traditional techniques such as the study of its morphology. Finally, new generation sequencing techniques were used to determine changes in the gene expression profile of *E. tetraedra* under different abiotic stressors. Thus, the work can be divided into three sections:

- **Phylogeography and study of genetic diversity of *E. tetraedra* (Chapters 1, 2 and 3)**

In the first three chapters, various aspects related to the genetic diversity of this earthworm are examined using three molecular markers (COI, 16S and 28S). Chapter 1 proposes a smaller scale study comparing two biogeographical areas of the Iberian Peninsula. A large genetic diversity was found, without a clear pattern of geographic structuring. In this first chapter, the six main lineages of this species are already presented. By increasing the number of populations studied and including the whole Iberian Peninsula in the study, two new lineages with very restricted distribution were presented in Chapter 2. In addition, a phylogeographic pattern for the lineages of *E. tetraedra* began to emerge as differences in certain environmental preferences such as

temperature, precipitation, pH, and habitat stability were noted. Finally, in Chapter 3, a total of 304 populations studied across Europe were included, yielding a clearer insight into the evolutionary history of *E. tetraedra* lineages. Thus, we find that new lineages, again with restricted distributions, are found in new areas of Europe such as the Scandinavian Peninsula and Eastern Europe. Therefore, we conclude that despite the fact that the nunatak theory (which defends the idea that there were ice-free glacial refugia in northern parts of Europe where certain species could survive) is not usually accepted for earthworms, in the case of *E. tetraedra* it played as important a role in post-glacial recolonization as the *tabula rasa* theory (which explains recolonization from refugia in southern Europe).

Moreover, thanks to all the molecular studies carried out in this work, it has been established that the subspecies described for *E. tetraedra*, all based on the different positions of the male pore, have no genetic basis, so that they are completely without systematic value.

- **Relationship between *E. tetraedra* and *Eiseniella neapolitana* (Örley, 1885) and position of the genus *Eiseniella* in the phylogeny of the family Lumbricidae (Chapter 4)**

Although there have been doubts about the validity of the species *E. neapolitana* over the years, Chapter 4 confirms that *E. neapolitana* is indeed a different species from *E. tetraedra* by examining molecular markers (COI, 16S, and 28S9). The validity of the species *Norealidys andaluciana* (Qui and Bouché, 1998) has also been questioned for years, as it is considered a synonym of *E. neapolitana* by the vast majority of scientists. In this chapter, this idea is rejected, as phylogenetic analyzes, genetic divergence

(based on the COI gene), and differences in the number of segments not only support the existence of the genus *Norealidys*, but also argue that *E. neapolitana* is found within this genus and becomes *Norealidys neapolitana*. Finally, the genera *Eiseniella* and *Norealidys* appear in the phylogeny of lumbricids in a polytomy with the genus *Iberoscolex.*

- **Genetic expression profile of *E. tetraedra* under different abiotic stressors (Chapter 5)**

In the face of current climate emergency, all animals are facing greatly altered environmental conditions. For this reason, two experiments were conducted in which *E. tetraedra* was exposed to different abiotic stressors and their changes in the gene expression profile were studied. In the first experiment, individuals were exposed to low temperatures up to freezing point. It was seen that this species was able to acclimatize to low temperatures, thanks to the uregulation of genes involved in increasing glucose reserves. However, acclimation to freezing conditions was less successful, as the expression of genes related to the respiratory chain was downregulated compared to control individuals. The second experiment consisted of reducing humidity to desiccation conditions. *E. tetraedra* used several mechanisms to acclimatize to these adverse conditions, such as upregulatoin of genes related to DNA repair and reduced apoptosis.

Resumen

La presente Tesis Doctoral pretende incrementar el conocimiento sobre las lombrices de tierra cosmopolitas y partenogenéticas. Para ello, se eligió como modelo la lombriz *Eiseniella tetraedra* (Savigny, 1826). A pesar de que existen estudios sobre otras especies de lombriz con las mismas características, nunca se había desarrollado un trabajo que incluyera tantos individuos y poblaciones. Por otro lado, se estudió su relación con otras especies del mismo género en base a marcadores moleculares y técnicas tradicionales como el estudio de su morfología. Finalmente, se emplearon técnicas de secuenciación de nueva generación para poder definir cambios en el perfil de expresión genética en *E: tetraedra* bajo distintos estresores abióticos. Así, la tesis se puede dividir en tres secciones:

- **Filogeografía y estudio de la diversidad genética de *E. tetraedra* (Capítulos 1,2 y 3)**

A lo largo de los tres primeros capítulos, se estudian distintos aspectos relacionados con la diversidad genética de esta lombriz basada en tres marcadores moleculares (COI, 16S y 28S). En el Capítulo 1, se plantea un estudio a menor escala, comparando dos zonas biogeográficas de la Península Ibérica. Se encontró una gran diversidad genética, sin un patrón claro de estructuración geográfica. En este primer capítulo ya se presentan los seis principales linajes de esta especie. Aumentando el número de poblaciones estudiadas, y esta vez ya incluyendo toda la Península Ibérica en el estudio, en el Capítulo 2, aparecieron dos nuevos linajes de distribución muy restringida. También, se comenzó a apreciar un patrón filogeográfico para los linajes

de *E. tetraedra*, puesto que se encontraron diferencias en cuanto a ciertas preferencias ambientales como temperatura, precipitación, pHy establidad del hábitat. Finalmente, en el Capítulo 3, se incluyeron un total de 304 poblaciones estudiadas en toda Europa, lo que nos dio una visión más clara sobre la historia evolutiva de los linajes de *E: tetraedra*. Así, observamos que nuevos linajes, nuevamente con distribución restringida se encuentran en nuevas zonas de Europa como la Península Escandinava y la Europa del Este. Por eso, concluimos que a pesar de que la teoría de nunatak (que defiende de la idea de que en zonas norte de Europa hubo refugios glaciares libres de hielo donde determinadas especies pudieron sobrevivir) no suele ser aceptada para lombrices de tierra, en el caso de *E. tetraedra* jugó un papel tan importante en la recolonización post glaciar como la teoría de la *tabula rasa* (que explica la recolonización desde refugios al sur de Europa).

Además, gracias a todos los estudios moleculares llevados a cabo en esta tesis, se pudo comprobar que las subespecies descritas para *E. tetraedra*, todas ellas basadas en las distintas posiciones del poro masculino, no tienen ninguna base genética, por lo que carecen completamente de valor sistemático.

- **Relación de *E. tetraedra* con *Eiseniella neapolitana* (Örley, 1885) y ubicación del género *Eiseniella* en la filogenia de la familia Lumbricidae (Capítulo 4)**

A pesar de que a lo largo de los años ha habido dudas sobre la autenticidad de la especie *E. neapolitana*, en el Capítulo 4 se confirma mediante en el estudio de marcadores moleculares (COI, 16S y 28S9, que efectivamente, *E. neapolitana* es una especie distinta a *E. tetraedra*. La validez de la especie *Norealidys andaluciana* (Qui and Bouché, 1998), también ha estado durante años en entredicho, siendo considerada por

la gran mayoría de científicos como un sinónimo de *E. neapolitana*. En este Capítulo, se rechaza esta idea, puesto que los análisis filogenéticos, la divergencia genética (basada en el gen COI) y las diferencias en número de segmentos no solo apoya la existencia del género *Norealidys*, si no que sostienen que *E. neapolitana* se encontraría dentro de este género, pasando a ser *Norealidys neapolitana*. Finalmente, los géneros *Eiseniella* y *Norealidys* aparecen en la filogenia de los lumbrícidos formando una politomía con el género *Iberoscolex.*

- **Perfil de expresión genética en *E: tetraedra* bajo distintos estresores abióticos (Capítulo 5**)

Dada la situación actual de Cambio Climático, todos los animales se han de enfrentar a unas condiciones ambientales muy cambiantes. Es por esto, que se hicieron dos experimentos sometiendo a *E. tetraedra* a distintos estresores abióticos y se estudiaron sus cambios en el perfil de expresión genética. En el primer experimento se expuso a los individuos a condiciones de baja temperatura, hasta el punto de la congelación. Se pudo ver como esta especie se pudo aclimatar a condiciones de baja temperatura, gracias a la sobre expresión de genes implicados en aumentar las reservas de glucosa. Sin embargo, la aclimatación a condiciones de congelación, fue menos exitosa, viéndose como genes relacionados con la cadena respiratoria veían su expresión reducida respecto a los individuos de control. El segundo experimento consitió en reducir la humedad hasta condiciones de desecación. *E. tetraedra* utilizó diversos mecanismos para la aclimatación a estas condiciones desfavorables, como la sobre expresión de genes relacionados con reparación de ADN y disminución de la apoptosis.

Introduction

Soil fauna: earthworms

Soil fauna includes those organisms that spend all or part of their lives in the soil, on the immediate surface of the soil, on surface litter, on rotting logs, and in other environmental appendages called floating soils (Brown *et al*., 2001). From a functional point of view, it is useful to distinguish three main size classes based on body width (Swift *et al*., 1979): the microfauna (< 100 µm), the mesofauna, mainly microarthropods (100 µm to 2 mm), and the macrofauna (> 2 mm). The soil biota, especially the larger soil animals, contributes significantly to rates of soil turnover, soil maintenance, formation of stable soil aggregates, total porosity, infiltration of water, and water retention (Lee and Foster, 1991). The abiotic physical and chemical soil processes caused by soil fauna regulate soil fertility and counteract the processes of soil degradation (Maldague 1970, Hole 1981, Lee 1983, Lee and Ladd 1984, Lal 1988). Soil biodiversity is thus vital to humans as it supports a variety of ecosystem processes, functions and services (Blouin *et al*., 2013, Skubala, 2013, Bardgett and van der Putten, 2014, Jouquet *et al*., 2014, Wall *et al*., 2015).

Earthworms may be the most important components of the soil biota in terms of maintenance of soil structure and fertility. Although they do not dominate in numbers, their size makes them one of the most important components of soil invertebrate biomass (Edwards, 2004). Aristotle was one of the first to draw attention to the role of earthworms in turning over the soil; he aptly called them "the intestines of the earth." The ancient Egyptians also held earthworms responsible for the fertility of their land,

and even Cleopatra elevated them to the category of minor gods. But it was not until the late 1800s that Charles Darwin really drew attention to the extreme importance of earthworms in the soil ecosystem in his work "The Formation of vegetable mould through the action of worms" (1881). Nowadays, the ecological importance of these ecosystem engineers (Lavelle *et al*., 1997) is widely known. Earthworms improve soil structure for the benefit of soil productivity (Barley, 1959; Lee and Foster, 1991; Edwards *et al*., 1996) and, as mentioned earlier, also improve soil fertility by accelerating the decomposition of plant litter and soil organic matter and consequently converting nutrients into forms available for uptake by plants (Curry and Boyle, 1987). In addition, their activities are vital for maintaining soil fertility in forests, grasslands and agroecosystems in a variety of ways (Edwards, 2004; Blouin *et al*., 2006; Bertrand *et al*., 2015; Johnston *et al*., 2015; Gavinelli *et al*., 2018). Apart from their immediate importance to agriculture as soil processors, earthworms also have considerable economic importance: compost-dwelling species are used for food waste processing, urban waste treatment, and sewage sludge treatment. In addition, some species are sold as bait for fish and accumulate heavy metals (Wang *et al*., 2018; Osioma and Hamilton-Amachree, 2019; Hattab *et al*., 2020). Earthworms are food for many other species such as planarians, leeches, mollusks, insects, amphibians, lizards, snakes, birds and mammals and are therefore an important link in many terrestrial food webs (Anderson *et al*., 2017).

Earthworms are a diverse group of mainly terrestrial burrowing annelids comprising more than 6,000 extant species classified in 18 families and found on all continents except Antarctica (Anderson *et al*., 2017). Most earthworm species live in soil, but some also live in rotting logs, leaf litter, stream mud, and river banks, as well as in

arboreal (e.g., epiphytic root masses) and even marine littoral habitats. Depending on the habitat, therefore, three different ecological categories are usually formed: epigean, anecic, and endogeic, with different effects on the soil environment (Sheehan *et al*., 2008; Ernst *et al*., 2009). Bouché (1972) developed this classification using two main criteria, vertical distribution of the earthworms and morpho-physiological characteristics. However, some earthworm species could be classified between the three main categories. Thus, four additional classes were defined and termed 'epi-endogeic', 'epi-anecic', 'endo-anecic' and 'intermediate' (Bottinelli *et al*., 2020). However, these ecological categories do not correspond to functional groups (Bottinelli and Capowiez, 2021).

The family Lumbricidae Rafinesque-Schmaltz, 1815 is the most diverse and widespread in the Palearctic region. It includes 44 genera and about 670 species (Blakemore, 2008). They are also the most abundant animals in temperate soils, where they account for 90% of the invertebrate biomass (Edwards, 2004). The Lumbricidae are considered a monophyletic group belonging to the monophyletic Crassiclitellata (which includes all earthworms except the Moniligastridae) (Jamieson, 1988, Jamieson, 2006, James and Davidson, 2012), united by the presence of a multilayered clitellum. Domínguez *et al*. (2015) highlighted the Palearctic origin of this family. They state that the monophyletic Lumbricidae genera diversified from the Lower Miocene 20.5 (14.6-27.1) Mya to the Paleocene 61.5 (53.5-68.6) Mya.

Because of their abundance and importance to the edaphic ecosystem, studies on earthworms, especially Lumbricidae, are essential for good management of the soil ecosystem.

How climate change affects earthworms?

As early as the late 19th century, scientists were arguing that human emissions of the greenhouse effect could change the climate, but the calculations were challenged. Nowadays, living its consequences, very few people dare to deny it. Indeed, global climate change is recognized as one of the greatest threats to biodiversity over the next century with significant consequences for the functioning and service provisioning of many ecosystems (Sala *et al*., 2000; Maxwell *et al*., 2016; IPBES 2018).

Soil processes influence climatic changes directly through the production and consumption of CO_2, CH_4 and N_2O and indirectly through the production and consumption of NH_3, NO_x and CO. Although CO_2 is primarily produced by the burning of fossil fuels, land use changes and conversion of forests and grasslands to agricultural land have contributed significantly to the increase of CO_2 in the atmosphere (Mosier, 1998). Changes in land use and management can also lead to a net uptake or sequestration of CO_2 in the atmosphere. CH_4 and N_2O are produced in the soil (30% and 70%, respectively), and soil processes are likely to regulate future changes in atmospheric concentrations of these gases (Cole *et al*. 1996). The exchange of CO_2, CH_4, and N_2O between soil and atmosphere is interconnected. Changes in the C cycle, for example, can lead to changes in the N cycle by altering rates of N mineralization and immobilization, which in turn can alter the exchange of N_2O between soil and atmosphere. Conversely, N supply increases C sequestration (Holland *et al*. 1997).

Climate change may also be reflected in changes in the mean and/or variability of its characteristics, such as temperature, precipitation and wind, that persist over a longer period of time (usually decades or longer) and are associated with increased likelihood

and/or intensity of extreme climate events such as droughts and floods (IPCC 2013). Climate change impacts on soil food webs may be caused by changes in the activity and mortality of soil organisms. Rising temperatures (which increase metabolic demands), frequency of extreme precipitation events, and droughts may additionally lead to mortality by altering the life cycle and diet of soil animals (Bates *et al*., 2008; Thakur *et al*., 2018). Phillips *et al*. (2020) found that precipitation, followed by habitat cover and temperature, are the most important factors affecting earthworm diversity and distribution patterns at a global scale. Soil carbon and pH also influence earthworm communities (Rutgers *et al*., 2009; Rutgers *et al*., 2016). It would be important to understand what environmental factors might explain earthworms species distribution in order to predict how changes in their communities may alter ecosystem functioning.

Temperature and soil moisture are the most important factors limiting earthworm survival, growth, and reproduction (Lee 1985). Earthworms employ a variety of strategies to cope with these conditions. Many earthworms often migrate to deeper soil layers (Jiménez *et al*., 2000; Kretzschmar and Bruchou, 1991; McDaniel *et al*., 2013) and other species, such as epigeic, produce drought-resistant cocoons (Holmstrup and Loeschcke, 2003; Petersen *et al*., 2008). Endogeic species enter a dormant state called aestivation (Díaz Cosín *et al*., *2006*).

The importance of winter conditions is often overlooked (Kreyling, 2010), although single extreme cold events can offs*et al*l distributional adjustments to the general warming trend (Jalili *et al*., 2010). Despite the lower frequency of their occurrence, both the intensity and duration of such extreme cold events may not decrease this

century. This is due to changes in atmospheric circulation and internal atmospheric variability that counteract the warming trend caused by greenhouse gases (Kodra *et al*., 2011). In fact, these extreme weather events have already increased in recent years due to climate change and will continue to increase. This creates new extreme habitats for species that need to develop adaptation mechanisms for their survival. It is known that earthworms cannot live in permafrost for long periods of time (Holmstrup *et al*., 1991), and epigeics are the most vulnerable in these conditions. However, they can tolerate low temperatures employing two main strategies: freeze avoidance either by migration or physiological adaptation; or allowing the formation of extracellular ice. The latter results in an osmotic flow of water out of the cells which prevents cell freezing (Mazur, 1963; Holmstrup and Zachariassen, 1996).

Given the importance of earthworms resulting from the ecosystem services they provide, it would be necessary to study the mechanisms and procedures by which they adapt to such events in order to preserve them while protecting an ecosystem as important as that of the soil.

The genus *Eiseniella* Michaelsen, 1900

Within the family Lumbricidae we find the genus *Eiseniella*. The last revision of the genus was proposed by Omodeo and Rota (1991). It includes two quite distinct groups of species, all of which are semiaquatic and have a quadrangular cross-section. The first group includes only the species *Eiseniella tetraedra* (Savigny, 1826) and its varieties (Omodeo and Rota, 1989). It is a parthenogenetic and cosmopolitan earthworm with morphological variability. The second group includes five species with biparental reproduction: *E. neapolitana* Örley, 1885, *E. ochridana* Cernosvitov, 1931, *E.*

eutypica Svetlov, 1924, *E. kuzencoi* Michaelsen, 1910 and *E. paradoxoides* Álvarez, 1971. The male pores are almost imperceptible, without porophores and are generally in XV. The female pores open dorsally to setae b or behind, as in the other lumbricids (with the exception of the Spermophorodrilinae), and the number of segments usually exceeds 95 and may increase to 160 (Perel, 1967; Zicsi, 1972; Omodeo and Rota, 1991). The distribution of each species of this second group is geographically restricted. *E. neapolitana* has been recorded from France, Italy, Spain, Turkey, Palestine and Algeria. *E. ochridana* is endemic to Lake Ochrida (North Macedonia); *E. eutypica* has been recorded from Permian (eastern Russia); *E. kucenkoi* near Lake Issik-Kul (Kyrgyzstan) (Omodeo and Rota, 1991) and *E. paradoxoides* is endemic to Spain (Álvarez, 1971).

In the large-scale molecular phylogenetic analysis of Lumbricidae proposed by Domínguez *et al*. (2015), they noted that the only member of the genus in the study, *E. tetraedra*, appears to be closely related to the Iberian members of the genus *Eiseniona* Omodeo, 1956 (included by other authors in *Iberoscolex* Qiu and Bouché, 1998). The provision of new specimens and species of the genus *Eiseniella* and related genera for this analysis could prove or reject this hypothesis.

E. neapolitana was long considered a subspecies of *E. tetraedra* (Stephenson, 1924; Bodenheimer, 1937; Cernosvitov, 1938, Cernosvitov, 1940; Pavlicek *et al*., 2003). Finally, Csuzdi and Pavlicek (2005) considered it as a valid species. Qiu and Bouché (1998) described *Reynoldsia andaluciana* on the basis of a few specimens from the locality Salobreña (Granada, Spain). This new genus is characterized by the absence of calciferous glands (glands of Morren), rudimentary or absent typhlosole and virtually

absent nephridial vesicles. Blakemore (2008) did not accept *Reynoldsia* due to homonymy with a fly and named it *Norealidys andaluciana*. In DriloBase (a database resulting from a collaborative work between worldwide members of the earthworm scientific community) it is considered a synonym of *E. neapolitana*. A solution to the systematics of this group of species could come from the study of molecular markers, as molecular tools help with these taxonomic problems (e.g. Huang *et al*., 2007; Marchán *et al*., 2018; de Sosa *et al*., 2019; Jiménez-Pinadero *et al*., 2021).

Cosmopolitan and parthenogenetic earthworms: the case of *Eiseniella tetraedra*

An earthworm community potentially includes four components: native species, often highly endemic; introduced exotics, which tend to be more widespread; translocated native species, which are endemic species that have been transported or translocated outside their natural range within a bioregion; and neoendemics, which are members of non-native genera that have persisted long enough after their introduction to reproduce in a new region (or that are currently unknown or extinct in their places of origin) (Blakemore, 1999). Cosmopolitan species are species that are both exotic to a region and widely distributed through transport. Only 120 species (plus synonyms) are widespread and common throughout the world-including 47 Megascolecidae (mainly pheretimoids), 33 Holarctic Lumbricidae, and about 40 species from some of the 18 other families (Blakemore, 2006). Widespread earthworms tend to be transported and usually exhibit some or all of the following characteristics: small size, parthenogenetic or asexual reproduction such that a single specimen can establish a new population (Gates, 1972; Sims and Gerard, 1999). However, the ability of an individual or species

to survive after its introduction to a site is subsequently influenced by local climate, soil ecology, and soil management practices. Parthenogenesis, often accompanied by polyploidy, allows for a more rapid accumulation of mutations than if reproduction had remained amphimictic (Gates, 1968; Gates, 1972). Those parthenogenetic morphs appear to resist a higher parasitic load, possibly because fewer physical resources are devoted to reproduction (Blakemore, 2009). High fecundity, hardy cocoons, great environmental or food tolerance, and rapid dispersal rates are also features of most cosmopolitan earthworms (Lee, 1985; Lee, 1987; Blakemore, 2002).

Parthenogenetic earthworms are automictic and thelytokous (Omodeo, 1951; Muldal, 1952; Omodeo, 1952, Casellato and Rodighiero, 1972). This implies a premeiotic doubling of chromosome number in the last oogonial division, followed by the formation of chiasmatic bivalent and regular meiosis with removal of the two polar bodies. By duplicating the chromosomes and mating with their identical copy, this mode of reproduction is perfectly compatible with any ploidy, even odd. Thus a single individual could leave offspring without having to find a mate. Nevertheless, cases have also been discovered in which a series of physico-chemical stimuli is necessary to trigger the parthenogenetic process. This is the case with *Octolasion tyrtaeum* (Savigny, 1826), where individuals copulate and exchange empty spermatophores, which, however, appear to act as a necessary stimulus to trigger parthenogenetic reproduction (Muldal, 1952; Jaenicke and Selander, 1979).

Parthenogenetic lineages can arise from species with biparental reproduction in several ways: bacterial infections such as *Wolbachia* or *Rickettsia* (Huigens *et al.*, 2000), spontaneous loss of sex due to mutations in genes related to mating and

fertilization of eggs (Carson *et al*., 1982), or in genes involved in sexual forms (Simon *et al*., 2003) and contagious origin, with incomplete reproductive isolation between sexual individuals and pre-existing parthenogenetic lineages (Simon *et al*., 2003) and hybridization origin between individuals of the same or closely related species (Lorenzo-Carballa and Cordero-Rivera, 2009).

The earthworm (Annelida, Oligochaeta) *Eiseniella tetraedra* reproduces by parthenogenesis (Casellato, 1987) and has a cosmopolitan distribution (Blakemore, 2006). It is a riparian species closely associated with the margins of water bodies of various types, both natural (lakes, rivers, streams and ponds) and artificial (sources, drains and the like), as well as stable (lakes or rivers) and unstable (drains or sources). Its strict reliance on aquatic environments may mean that some aspects of its biology resemble those of an aquatic animal rather than a typical earthworm. For example, its aquatic distribution in some of the rivers studied may result in greater clonal diversity downstream than upstream (Terhivuo and Saura, 2002).

Studies on genetic variability of parthenogenetic earthworms are scarce and have yielded surprising results, see e.g. Fernández *et al*. (2011, 2012) on *Aporrectodea trapezoides* (Dugés, 1828), where genetic variability was greater than expected. This genetic variability could be explained by the multiple and delocalized origin of parthenogenesis in this species (Fernández *et al*. 2012; de Sosa *et al*., 2017), so that the genetic variability of parthenogenetic forms could be largely inherited from their sexual ancestors. Other parthenogenetic species such as *Aporrectodea rosea* (Savigny, 1826) and *Dendrobaena octaedra* (Savigny, 1826) also show high genetic variability with a high number of clones in populations in Scandinavia, while species such as

Octolasion cyaneum (Savigny, 1826) and *O. tyrtaeum* have a low number of clones and thus low variability in the same area (Terhivuo and Saura, 2006). Thus, there is a possibility that *E. tetraedra* follows one of these patterns, with the added uncertainty of whether it has sexual forms (not yet found). During years, there was no precursor to the study of the genetic structure of *E. tetraedra* based on sequences, but in enzymatic polymorphisms, a method developed by Terhivuo and co-workers in Finland, Sweden and other Nordic countries. Terhivuo *et al*. (1994, 2011) pointed out the advantages and problems of studies based on enzymatic polymorphisms and called for the study of this species using other molecular methods. Because polyploidy can complicate study with nuclear genes, they recommended the use of mitochondrial markers. Thus, Javidkar *et al*. (2020) studied the genetic diversity of this species using the cytochrome oxidase subunit I gene and found six different lineages. The presence of identical haplotypes at European, North American, Australian and Iranian sites, the sympatric clustering of several different intraspecific lineages in the same rivers and the lack of a phylogeographic pattern strengthen the hypothesis of a possible unintentional anthropogenic introduction (Jadvikar *et al*., 2020).

Morphological and taxonomic aspects must not be forgotten. Within this species, several subspecies have been described, differing mainly in the number and position of the male pores. According to Blakemore (2006), the most common subspecies are:

- *E. t. cerni* (Blakemore, 2004), male pores at 14.
- *E. t. tetraedra* (Savigny, 1826), male pores at 13.
- *E. t. pupa* (Eisen, 1874) (syn *hercynius* Michaelsen, 1890; *tetraedra quadripora* Cernosvitov, 1942), male pores at 12 or 15.

Considering the high degree of polymorphism of this trait at the population level, their study in several distant populations where individuals with male pores could be found in different locations, this could support (or refute) their use to determine subspecies or species after comparing genetic variability and morphological results.

Phylogeographic studies on earthworms

Phylogeography is the field of research concerned with the principles and processes that determine the geographic distribution of gene lineages, especially between and within closely related species (Avise, 2000). The analysis and interpretation of lineage distributions requires extensive contributions from the fields of molecular genetics, population genetics, demography, phylogenetics, paleontology, geology, and historical geography. Thus, phylogeography is an integrative discipline located at an important crossroads of various micro- and macroevolutionary disciplines (Avise, 2000).

Earthworms are suitable candidates for phylogeographic studies among edaphic animals due to their low vagility in geological time: paleogeographic events seem to have been of great importance for their present distribution (Fernández *et al*., 2013; Novo *et al*., 2011). However, the biogeographic models and evolutionary history of this group are still largely unknown, which is particularly important at the systematic level. Species delimitation and the establishment of a target genus system are highly dependent on information about these evolutionary processes. However, the few existing studies have almost always been limited to endemic species (Chang and James, 2011; Pérez-Losada *et al*., 2011; Novo 2010), although contrary to popular belief, even the widespread edaphic species show a high degree of genetic differentiation among populations (Fernández *et al*., 2013; Fernández *et al*., 2015) and

their study may be even more interesting than that of endemics to reveal general evolutionary processes.

The first attempt at a phylogeographic study of a cosmopolitan and parthenogenetic earthworm, *Aporrectodea trapezoides*, on a global scale was made by Fernández *et al*. 2011. They found two geographically distinct lineages, one with a Eurosiberian distribution and the other with a Mediterranean distribution. Originally, *Aporrectodea rosea* also had these geographical lineages in Europe (Fernández *et al*., 2016). However, Shekhovtsov *et al*. (2020) conclude that the patterns of genetic diversity in *A. rosea* appear to be more complex than previously thought. The Eurosiberian lineage of this species has probably diversified in Eastern Europe and spread further east and west.

The phylogeography of earthworms may have been significantly influenced by Pleistocene glaciation cycles. It is generally believed that northern western Europe and North America were covered by ice sheets that wiped out most of the fauna (Hewitt, 2000), but some species were able to survive in nunataks (Provan and Bennett,2000). The latter variant seems unlikely for earthworms, but such hypotheses have been proposed for Fennoscandia (Fridolin, 1936; Stöp-Bowitz, 1969) and Greenland (Hansen *et al*., 2006). Northeastern Eurasia, on the other hand, had limited glaciation but a harsher climate (Shekhovtsov *et al*., 2018).

Thus, a comprehensive phylogeographic study of *E. tetraedra* could shed light on biogeographic patterns in earthworms, especially cosmopolitan, parthenogenetic, and riparian species.

Objectives

The main objectives of this study are:

- To investigate the genetic diversity of *Eiseniella tetraedra*.
- To understand the evolutionary history of this species in Europe and its relationships with environmental factors.
- To corroborate or refute its subspecies with molecular analyses.
- To test whether *Eiseniella neapolitana* is really a true species or a subspecies of *E. tetraedra*.
- To understand the genetic mechanisms of adaptation of *E. tetraedra* in the context of climate change.

Objectives of Chapter 1

- To study of the genetic structure of *E. tetraedra* in two different biogeographical zones in Spain (Eurosiberian and Mediterranean).
- To assess whether this species shares the general traits of most soil fauna, such as high genetic diversity, strong population structure and high intraspecific divergence.
- To test whether different lineages are found in European and Mediterranean areas as previously shown in other species.
- To study the distribution of genetic variability along Guadarrama river basin to test whether *E. tetraedra* follows described patterns for other riparian animals.
- To evaluate the taxonomic subspecies division based on male pore position, by correlating genetic and morphological variability.

Objectives of Chapter 2

- To study the genetic diversity of *E. tetraedra* in the Iberian Peninsula.
- To understand its demographic history in the Iberian Peninsula.
- To investigate putative relationships between genetic diversity and environmental factors.
- To test whether genetic diversity is related or not to adaptation to stable or unstable habitats.
- To investigate whether the morphological variations are related to genetic diversity and other variables such as altitude, longitude and latitude.

Objectives of Chapter 3

- To examine the genetic diversity of *E. tetraedra* in Europe.
- To infer its phylogeography in this continent.
- To test macroecological preferences of the most consistent clades.

Objectives of Chapter 4

- To test if *E. neapolitana* is a valid species, distinct from *E. tetraedra*.
- To study the genetic diversity of *E. neapolitana.*
- To prove whether *Reynoldsia andaluciana* is really a synonym of *E. neapolitana*.
- To place the genus *Eiseniella* in the phylogenetic tree of lumbricids

Objectives of Chapter 5

- To explore the transcriptional changes in *E. tetraedra* during freezing conditions.

- To explore the transcriptional changes in *E. tetraedra* during desiccation conditions.
- To investigate possible common responses between both treatments.

1. Chapter 1: Bless this phylogeographic mess - Comparative study of *Eiseniella tetraedra* (Annelida, Oligochaeta) between an Atlantic area and a continental Mediterranean area in Spain.

Abstract

Due to the influence of Atlantic Ocean and Mediterranean Sea, Spain has different climates, from desert to Atlantic. We sampled the parthenogenetic earthworm *Eiseniella tetraedra* in two different biogeographical zones in Spain, in order to study their genetic diversity and test their potential distinctiveness. Moreover, we evaluated the presence or absence of two different lineages (Eurosiberian and Mediterranean) found in other parthenogenetic earthworms such as *Aporrectodea trapezoides* and *Aporrectodea rosea*. We studied the molecular markers COI, 16S and 28S. *E. tetraedra* presents a high diversity in Spain (one COI haplotype every two individuals were found) and no clear geographical patterns except for diffuse patterns along the Guadarrama River basin. In contrast, worldwide localities were more homogeneous with low diversity, to be confirmed with further samples. After morphological study, no correlation was found between phylogenetic relationships and the diagnostic characters for the previously described subspecies in *E. tetraedra*.

1.1 Introduction

Spain is a highly suitable scenario for phylogeographic studies due to its complex geological history and the variety of its environments (Hewitt, 1996). Due to its geographical position, Spain is under the influence of the Atlantic Ocean and the Mediterranean Sea, resulting in a wide range of climates including desert, Mediterranean, Alpine and Atlantic (Gómez and Lunt, 2007) comprising two biogeographic regions: Eurosiberian and Mediterranean (Alcaraz *et al*., 2006). Some differences in presence of earthworm species between these regions have been observed, but many lumbricids (like *Eiseniella tetraedra*) are present in both (Rodríguez *et al*., 1997). Moreover, Spain was one of the most important Pleistocene glacial refugia in the European subcontinent (Hewitt, 1999; Hewitt 2001) and acted as a species repository of northern lands (Beebee and Rowe, 2000; Jaarola and Searle, 2002; Vernesi *et al*., 2002). The phylogeographic study of *E. tetraedra* within Spain and Europe could shed light on these processes and be useful for elaborating European settlement models.

The focus of this piece of work is *Eiseniella tetraedra* (Savigny, 1826) (Annelida, Oligochaeta), a parthenogenetic tetraploid (Casellato, 1987) and cosmopolitan earthworm (Blakemore, 2006) with a riparian lifestyle (closely linked to edges of water). It is possible that *E. tetraedra*'s biology is closer to aquatic rather than terrestrial animals, due to its strict dependence on water. Together with its wide distribution and its parthenogenetic reproduction, this makes *E. tetraedra* an

interesting candidate species for investigating biogeographic and genetic diversity patterns in soil, water and ecotone systems.

Knowledge on evolutionary processes in soil fauna is only starting to flourish. Some studies of different groups such as harvestmen (Boyer *et al*., 2007; Clouse *et al*., 2015), oribatid mites (Heethoff *et al*., 2007), myriapods (Vélez *et al*., 2012), earthworms (Fernández *et al*., 2013), or caecilians (Stoelting *et al*., 2014) have revealed shared characteristics: high genetic variability, low vagility in a particular geological time, strong population structure, and high intraspecific divergence (Costa *et al*., 2013). Because of these attributes, earthworms have been proposed as good candidates for phylogeographic studies, since paleogeographic events appear to have great relevance in their current distribution (Novo *et al*., 2011; Fernández *et al*., 2013).

However, this is not true for all earthworms, and examples have been found in which little or no population structure exists, like *Aporrectodea icterica* Savigny, 1826, *Allolobophora chlorotica* Savigny, 1826 (Torres-Leguizamon, 2014) or the case of invasive species like *Amynthas cortices* Kinberg, 1867 (Novo *et al*., 2015). Relative absence of population structure was also found in the semi aquatic earthworms *Glyphidrilus vangviengensis* Panha and Chanabun, 2011 and *Glyphidrilus mekongensis* Panha and Chanabun, 2012. They showed isolation by distance, but not vicariance, due to river flow connecting their populations (Jirapatrasilp *et al*., 2015). Studies of cosmopolitan, parthenogenetic earthworms have found different results: *Octolasion tyrteum* showed high homogeneity with a low number of clones across northern Europe while *Aporrectodea rosea* showed high genetic variability in the same region (Terhivuo and Saura, 2006). Moreover, studies of molecular markers found two deep,

differentiated lineages, one present in Eurosiberian region and the other one present in Mediterranean region, not only in *A. rosea* but also in *Aporrectodea trapezoides* (Fernández *et al*., 2012; Fernández *et al*., 2015) Therefore, there is the possibility that *E. tetraedra* follow either of these patterns, with the added uncertainty of the presence of sexual forms (which have not yet been found).

Only studies based on enzymatic polymorphisms have been conducted for *E. tetraedra* (Terhivuo *et al*., 1994; Terhivuo *et al*., 2011), which showed high clonal variability with no clones shared between Sweden, Finland and Russia (Terhivuo *et al*., 2002), yet they found shared clones between Northern Norway and Southern Finland with no clear distribution patterns (Terhivuo *et al*., 1994). These works had two additional interesting outcomes. Firstly, morphological variability showed no correlation to the enzyme patterns in the studied populations. Secondly, a higher clonal diversity was found in the lower course compared to that upstream in the studied rivers (Terhivuo *et al*., 2002), due to its dispersion through water. This pattern has also been observed in other clonal and flightless invertebrates (Stenberg *et al*., 1997; Stenberg *et al*., 2000)

The morphological and anatomical simplicity of soil dwelling animals, such as earthworms, has limited the establishment of a robust taxonomy. It remains anchored, to some degree, in subjective criteria of each author. Within this context, integrative approaches including molecular information are becoming more popular to solve the phylogenetic position of conflictive taxa (Díaz Cosín *et al*., 2014). Morphological variability in *Eiseniella tetraedra* and its taxonomy are an example of a field open to such research. At least three subspecies have been described (Savigny, 1826; Eisen, 1874; Blakemore, 2004) the position of the male pores being the diagnostic character.

Due to the high polymorphism of this character in different populations (pers. observ.), the correlation of morphological and genetic variability of specimens collected in different localities could support or contradict the use of this character to differentiate subspecies.

With this background, *E. tetraedra* constitutes a suitable model for testing not only different hypotheses about riparian and soil fauna but also for testing some animal distribution patterns in Europe. The aim of this work is the study of the genetic structure of *E. tetraedra* in two different biogeographical zones (Eurosiberian and Mediterranean) with the following objectives: i) to assess whether this species shares the general traits of most soil fauna (high genetic variability and strong population structure). ii) To test whether different lineages are found in European and Mediterranean areas as previously shown in other species. iii) To study the distribution of genetic variability along a river basin to test whether *E. tetraedra* follows described patterns for other riparian animals. iv) To evaluate the taxonomic subspecies division based on male pore position, by correlating genetic and morphological variability.

1.2 Material and methods

1.2.1 Sampling and morphological studies

We collected specimens from thirty localities from the Spain: 14 from the Northwestern area and 16 from the central region (Figure 1, see geographical coordinates in Supplementary Tables 1 and 2). For comparative purposes, all sequences of the mitochondrial cytochrome c oxidase subunit I gene (COI) belonging

to *Eiseniella tetraedra* were downloaded from the BOLD system web (Supplementary Table 3). Moreover, we collected one more population in Wales, United Kingdom.

Figure 1. Localities sampled. More detailed maps are shown in Figure 3.

All individuals were hand collected, washed in distilled water, fixed in 96% EtOH and preserved at -20°C. A fragment of the body wall from the posterior end was separated for genetic analysis. Later, morphological studies and dissections were performed, focusing on: length, dry weight, clitellum position, tubercula pubertatis, male pore position, seminal vesicles number and position. Presence of iridescence in spermathecae and male funnels indicates presence of sperm (Plisko, 2002). In order to test the existence of sexual specimens in this supposed strictly parthenogenetic species, presence or absence of iridescence of spermathecae and male funnels was studied.

1.2.2 DNA extraction, gene amplification and sequencing

Total genomic DNA was extracted using the DNeasy Blood and Tissue Kit (QIAGEN). Furthermore, two mitochondrial markers (COI and 16S + tRNAs Leu, Ala and Ser) and one nuclear marker (28S) were amplified. For COI (632 bp) we used the primers and PCR conditions used by Pop *et al.* (2003). For 16S-tRNAs (775 bp) and for 28S (806 bp) primers and PCR conditions from Fernández *et al*. (2015) were used. PCR products were sequenced by Macrogen Europe Inc. (Holland).

1.2.3 Data analysis

Sequences were aligned in MAFFT v.7 (Katoh and Standley, 2003) and concatenated in Bioedit (Hall, 1999). Haplotypes of each gene were retrieved by DNAsp v.5 (Librado and Rozas, 2009). Phylogenetic trees based on the concatenated sequence (2,270 bp) were built through Maximum Likehood using the software RaxML v7.0.3 (Stamatakis, 2006) and Bayesian Inference with MrBayes v3.1.2 (Ronquist and Huelsenbeck, 2003) both implemented in the CIPRES Science Gateway [43]. GTR+Γ+I was chosen by jModelTest2 (Darriba *et al*., 2012) as the best-fit substitution model for each partition. Bayesian analysis consisted of two parallels runs of 10 millions of generations and 20% of the trees were discarded as burn-in. For the ML analysis rapid bootstrapping was conducted including 1000 replicates. *Carpetania matritensis* Marchán *et al*., 2020, *Lumbricus rubellus* Hoffmeister, 1843, *Dendrobaena byblica* Rosa, 1893, *Iberoscolex oliveirae* (Rosa, 1894) and *Prosellodrilus biauriculatus* Bouché, 1972 were chosen as outgroups and their sequences were retrieved from GenBank (Supplementary Table 4). Furthermore, uncorrected pairwise distances for 16S-tRNAs and COI were computed in Arlequin v.3.5. (Excoffier and Lischer, 2010).

A COI based haplotype network was constructed using TCS version 1.21 (Clement *et al.*, 2000) with statistical parsimony and a connection limit of 95%.

We also obtained haplotype and nucleotide diversities of the localities from Guadarrama river basin (in order to compare the different parts of the river) and of each lineage retrieved from phylogenetic analysis by DNAsp v.5. (Librado and Rozas, 2009).

Finally, statistical analyses of morphological data (Kruskal-Wallis test) were conducted in Statistica v.7. in order to test whether genetic and morphological data were congruent and thus the taxonomic value of morphology in this case.

1.3 Results

A total of 113 haplotypes were identified among 271 sequences for COI gene, 40 haplotypes within 58 sequences for 16S-tRNAs and 3 haplotypes within 28 sequences for 28S.

1.3.1 Phylogenetic analysis

Both Maximum Likelihood and Bayesian approaches presented trees with congruent topology (Figure 2). Haplotypes nested in six well supported lineages (A to F). Lineages A, B, E and F showed higher genetic diversity (with 18.18%, 19.93%, 19.69% and 22.27% of total haplotypes respectively) than lineages C and D (with 9.09% and 11.36% of haplotypes). As shown by the short branch lengths of the tree and reticulated structure of haplotype networks (Figure 2) lineages presented a high internal homogeneity. In contrast, the different lineages showed deep divergences with long branches.

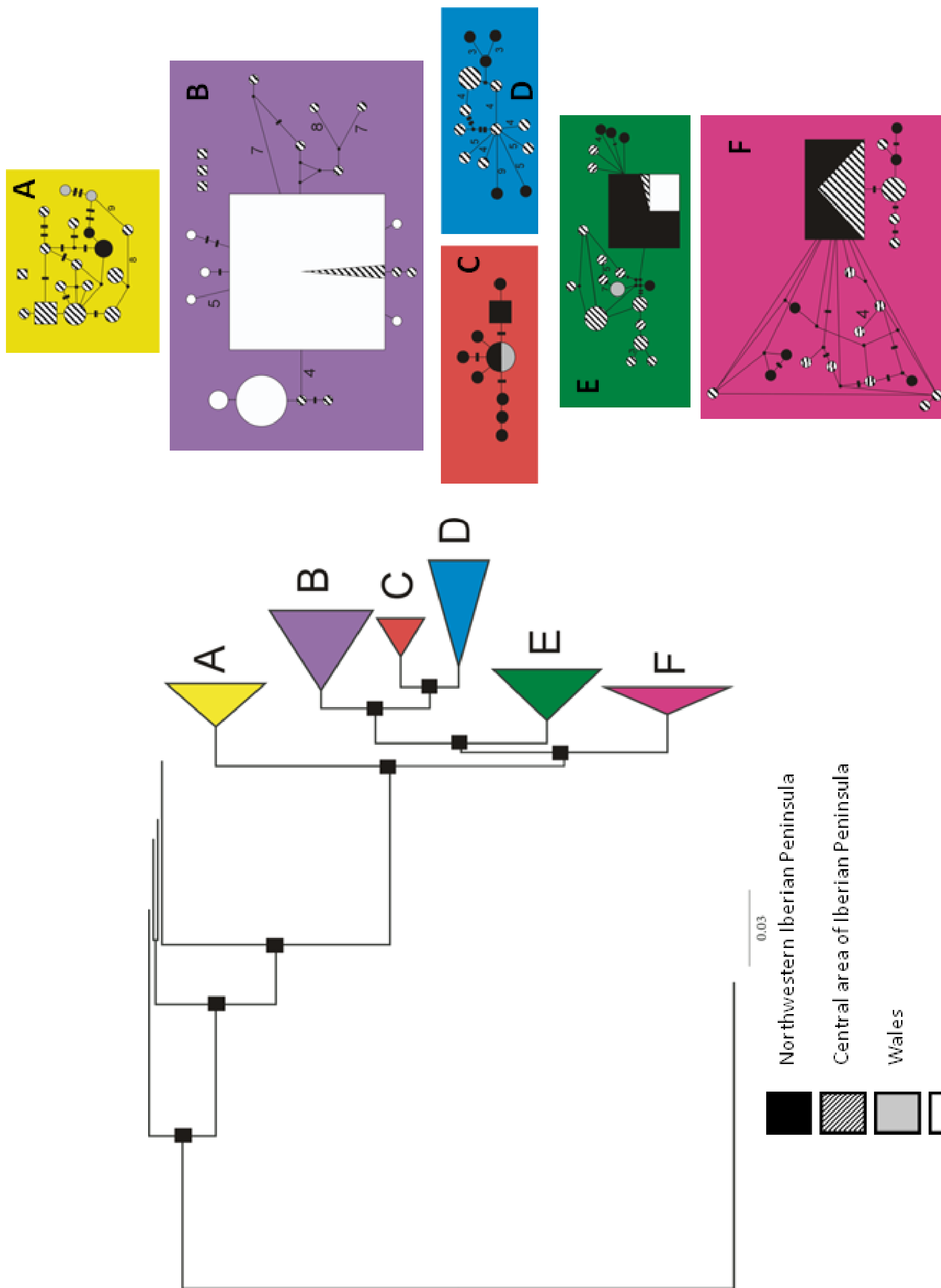


Figure 2. Bayesian inference (BI) of the phylogenetic tree based on the concatenated sequences of COI, 16S-tRNAs and 28S. Posterior probability/bootstrap support values (of Maximum Likehood Analysis, ML) are shown when they are higher than 0.99/0.7 (BI/ML) as black squares. The scale bar represents 0.03 substitutions per position.

Haplotype networks are based on COI sequence. Size of the ellipses represents the frequency of each haplotype. Squares represent hypothetical ancestral haplotypes according to TCS software. Intermediate circles are hypothetical intermediated haplotypes. Each branch indicates a mutational step and its length contains no information. More than two steps are represented by thick black lines or their value.

1.3.2 Haplotype and lineage distribution

Studied localities showed high haplotypic variability. Most lineages, except B and C, appeared in both regions (Eurosiberian and Mediterranean), lacking a pronounced geographical structure. In Spain, lineage B is exclusively Mediterranean and C Eurosiberian. The rest of the sequences of the world (RSW), were retrieved only in lineages B and E and showed rather low haplotypic variability. The number of sequences/specimens included in each lineage is: A, 33; B, 98; C, 13; D, 18; E, 68; F, 45. As haplotype networks show (Figure 2) regional haplotypes were found, either region-specific (Eurosiberian or continental Mediterranean) or even locality-specific. On the other hand, interregional haplotypes (shared between both regions) were found as well. Figure 3 shows the distribution of the lineages in the studied regions (maintaining the color code of Figure 2). Both localities with presence of a single lineage and localities with several lineages were found. In the central area of Spain (Figure 3C) the lineages are differentially distributed along the River Guadarrama basin: some lineages predominate in the upper course, while others have a stronger presence in the middle and lower course. Figure 3D shows higher lineage diversity in Spain and Wales, while the rest of localities around the world are very homogeneous. Lineage B is distributed in the western localities of Europe and North America (with as many as 90 sequences),

and lineage E was found in the eastern localities of Europe (Austria) western Asia (Turkey) and Oceania (New Zealand).

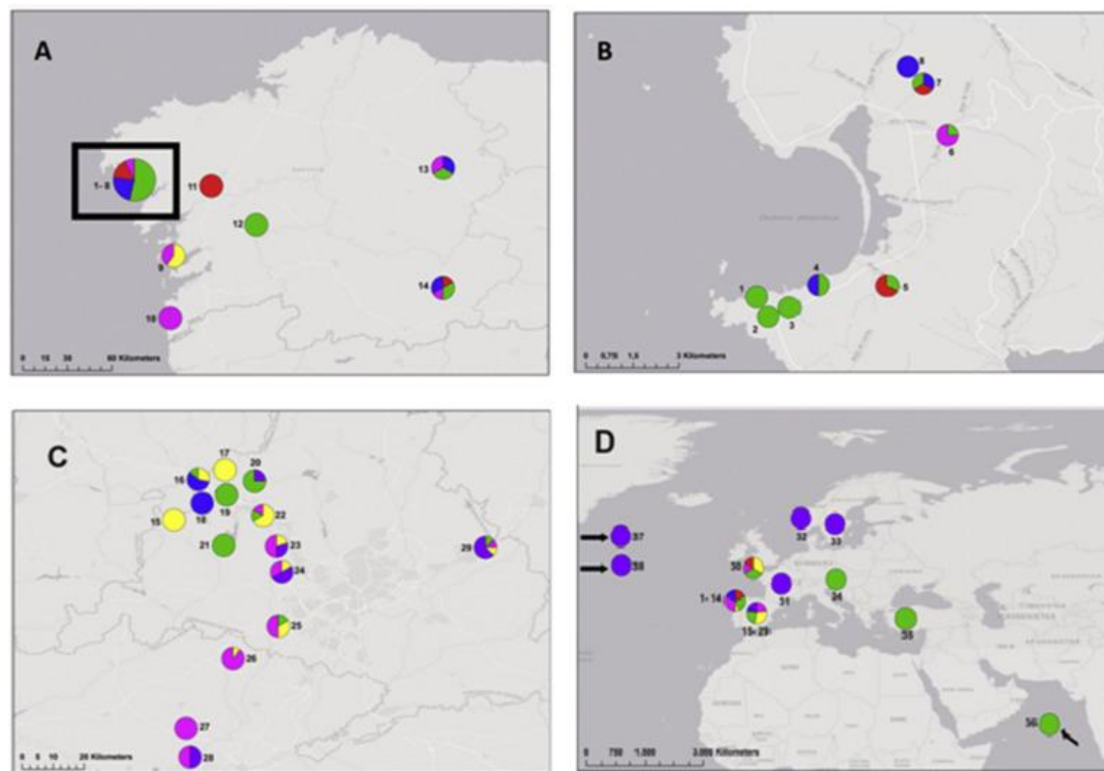


Figure 3. Lineage distribution in A: the localities from Northwestern Spain, B: A micro-scale study in Carnota, A Coruña, Galicia, Spain from localities 1 to 8 (area indicated by a square in A). C: Lineages distribution in the central area of Spain. D: Lineages distribution in all studied areas. Colors used are the same as in Figure 2.

1.3.3 Genetic diversity

As it shown in Table 1, genetic distances based on 16S-tRNAs were lower than those based on COI, due to 16S-tRNAs being a more preserved region of the mitochondrial genome. A certain degree of variability within lineages was observed, but divergence between them was remarkably higher. Values between lineages were considerably lower than those found between the outgroups, although some of them (specially

lineage A) are near (or within) the ambiguous gap between intraspecific and interspecific divergence in earthworms proposed by Chang and James (2011) −9 to 15%-.

COI/16S	**A**	**B**	**C**	**D**	**E**	**F**
A	0.97/0.33	2.22	3.21	3.45	2.52	2.77
B	9.79	0.69/0.25	2.67	2.3	2.06	2.79
C	9.71	4.93	0.58/0.57	2.64	3.04	4.06
D	10.25	5.53	5.2	1.65/0.33	2.84	3.66
E	7.94	8.13	8.41	9.42	0.47/0.51	2.9
F	7.75	8.19	7.88	8.86	5.86	0.84/0.19

Table 1. Percentage of uncorrected average pairwise genetic distances (COI below the diagonal and 16S-tRNAs above) between lineages (within lineages in diagonal) retrieved for *Eiseniella tetraedra*.

Haplotype and nucleotide diversities were high at all localities from the Guadarrama river basin (Table 2). The highest diversity was found in the middle course, while the lowest diversity was found in the lower course.

	Number of samples	Number of haplotypes	H diversity	Π diversity
Upper course	49	29	0,82	0,11458
Middle course	22	19	0,97	0,20719
Lower course	21	8	0,59	0,06364

Table 2. Genetic diversity parameters (H diversity: haplotype diversity; π diversity: nucleotide diversity) in Guadarrama river basin localities based on COI sequences of *Eiseniella tetraedra*.

High haplotypic but low nucleotide diversity values were found within lineages, meaning that haplotypes within each lineage were abundant but very similar to each other (Table 3).

Lineage	Number of samples	Number of haplotypes	Number of polymorphic sites	H diversity	Π diversity
A	31	11	22	0,828	0,00691
B	96	12	42	0,52	0,00286
C	13	5	5	0,705	0,00222
D	19	9	13	0,678	0,00929
E	62	17	28	0,744	0,00421
F	46	12	56	0,701	0,00636

Table 3. Genetic diversity parameters (H diversity: haplotype diversity; π diversity: nucleotide diversity) of lineages retrieved for *Eiseniella tetraedra* based on COI sequences.

1.3.4 Morphological studies

Collected individuals showed wide variability in the studied morphological characters. In terms of internal characters, spermathecae and male funnels were never iridescent, even being absent in many individuals.

Length, weight and number of segments also showed a great degree of variability, but no significant differences were found between lineages (Supplementary Figures 1–3). Morphological data is presented in Supplementary Table 5.

Regarding the position of the male pore, we found identical haplotypes with different state of this character.

1.4 Discussion

Phylogenetic analysis and pairwise distances showed six divergent, internally homogeneous lineages of *Eiseniella tetraedra*. This pattern suggests that not enough time would have elapsed for the differentiation of the haplotypes in each lineage due to regular bottlenecks caused by a constant founder effect or selective sweeps.

The most basal divergence was between lineage A and the rest of lineages. These results suggest that lineage A could represent an ancestral group with a stable demographic history due to its high nucleotide and haplotype diversity (Grant and Bowen, 1998).

The geographic distribution of lineages (all present in Eurosiberian and Mediterranean areas except B and C) and the dispersed geographical distribution of haplotypes within lineages support the absence of strong population structure and geographical structure, despite some lineages showing predominance in certain regions. However,

clearly defined Eurosiberian and Mediterranean lineages as in *Aporrectodea rosea* and *A. trapezoides* (Fernández *et al*., 2012; Fernández *et al*., 2015) were not found. The most extreme case of this is lineage C. It presents haplotypes from Northwestern Spain and from Wales, which could be explained by human-mediated dispersal. The historical connection between these areas (i.e. the Britonia Breton enclave in northern Lugo (Ferreiro, 2003)) supports this possibility.

Lineages B and E included specimens of the rest of the world (RSW), in addition to specimens from Spain. It is known that recent glaciations could have led to the extinction of most of the North European populations of earthworms (Mathieu and Davies, 2014). A possible explanation could be a recolonization from both Spanish regions (Eurosiberian and Mediterranean), which acted as glacial shelters, to Europe. A recolonization from other refugia (like Italic or Balcanic peninsulas) is also probable, as seen in other groups of animals and plants (Hewitt, 1999). Including individuals from those refugia in future studies is necessary in order to provide a wider view of the biogeography of this species in Europe. *E. tetraedra* has expanded from glacial shelters to Scandinavia in approximately 15.000 years, which indicates a not exceedingly low vagility. While such active colonization seems unlikely, it could be explained by passive dispersion along rivers or by human mediated transport (as suggested for *Sparganophilus tamesis* Benham, 1892 (Rota *et al*., 2016)). The distribution of the Lineages B and E in the westernmost and easternmost localities of the studied area suggest two independent colonization events from the highly diverse Spain or other refugia. These findings suggest a certain geographical ancestral structure modified by climate and dispersive history of the species.

The majority of haplotypes found were regional or even unique in each locality, as previously found by Terhivuo and Saura (1997) using allozymes. Interregional haplotypes were also found, shared between Northwestern and central area of Spain. Moreover, haplotypes (based in COI) shared between Northwestern Spain and Wales were found. Terhivuo *et al*. (2011) didn't find shared clones between Russia, Finland and Sweden but they found shared clones between north Norway and south Finland, separated by about 1,200 km (Terhivuo *et al*., 1994).

In the present study, two opposite patterns of genetic diversity were found. On one hand, localities with representatives from only one lineage or even only one haplotype were found. On the other hand we found heterogeneous localities, with representatives from two or more lineages and/or a great variety of haplotypes. This result suggests at least two possible interpretations: competitive exclusion between lineages (or haplotypes), and the sequential arrival of founder populations followed by the coexistence of lineages and/or haplotypes. Terhivuo and Saura (1996) thought that the absence of interregional clones in a country is due to dispersive history. Possibility of different colonization events in each locality suggests complex evolutionary histories in each population, following a random pattern (Terhivuo *et al*., 2002). The co-existence of both diversity patterns in geographically close populations, together with the long-range (more than 2,000 km) presence of shared haplotypes, give as a result an unexpectedly ‘messy’ phylogeography for *Eiseniella tetraedra.*

A very small subset of the *E. tetraedra* genetic variability was represented, with the available data, in the areas outside Spain and Wales (RWS): only 15 haplotypes in 90 sequences from 8 different countries, compared to the total 129 haplotypes in 271

sequences. This pattern is similar to the one described by Mathieu and Davies (2014) in France based on species diversity instead of genetic diversity: according to the authors this latitudinal gradient of diversity would be caused by the post-glaciation dispersive history of the different lineages.

In terms of haplotype and nucleotide diversity in the Guadarrama river basin in the central area of Spain, localities didn't seem to follow the dispersive pathway found by Terhivuo and Saura (1996) in Ume and Vindel rivers (Sweden) where the lower course presented the highest genetic diversity. Instead, in our study, the middle course of the Guadarrama river basin showed the highest variability being the lower course the less diverse. Environmental conditions in the lower course are expected to be less suitable for *E. tetraedra* (like intermittent water availability in the tributary streams, use of water for irrigated crops and other human consumption, and presence of contamination, which will be tested in further studies), which could negatively affect their genetic variability. Thus, the middle course would be more representative of the lower course in said study, accumulating haplotypes from upstream which would not thrive downstream due to the unfavorable conditions.

Uniparental and parthenogenetic reproduction of *Eiseniella tetraedra* were inferred by Gavrilov (1939) and the chromosomal study performed by Muldal (1952) confirmed said condition Morphological studies showed anomalies in structures related to this kind of reproduction, like variability in number and absence of iridescence in spermathecae, reduction or absence of male funnels, and male pores in different positions. All these varieties were found in other studies (Gavrilov, 1939; Bouché, 1972; Gates, 1977; Terhivuo *et al*., 1994; Blakemore, 2006). Some of these works

found the presence of spermatophores (structures associated with biparental reproduction), but they were absent from our samples.

Male pore position was used as a taxonomical character separating different subspecies (Bouché, 1972; Gates, 1977; Blakemore, 2006). According to our results, subspecies division based in male pore position does not appear to have a phylogenetic base, as no correlation between this character and the relationships recovered by the three molecular markers studied was found.

Further studies will shed light on the complete phylogeography of this species in whole Spain.

1.5 Conclusions

Some of the initial questions can be answered. The high genetic variability and population structure were variable between populations. *Eiseniella tetraedra* showed a great genetic variability in the two studied areas in Spain with no clear patterns of population structure as a result of two opposite trends (monohaplotypic and polihaplotypic populations in both areas). There were some genetic differences between populations of *E. tetraedra* from the Eurosiberian and Mediterranean studied regions. We found six deeply divergent lineages, all of them represented in both zones of Spain except lineage C, which was only found in Northwestern Spain and Wales, and lineage B, only present in central area of Spain and worldwide localities. The samples from the rest of the world showed limited variability. The studied samples from the rest of the world showed limited variability, which can be biased by the number of samples. We suggested a recolonization from Spain to Europe even though other

possible refugia such as Italian and Balcanic Peninsulas should be explored with further sampling.

An apparent dispersion pattern was detected along the Guadarrama river basin, with lineages upstream and middle/low stream being different.

Finally, actual subspecies subdivision based on position of male pores did not show correlation with genetic data.

2. Chapter 2: Phylogeography of *Eiseniella tetraedra* (Savigny, 1826) in the Iberian Peninsula shows a genetic distribution based on environmental factors.

Abstract

The Iberian Peninsula is a highly suitable scenario for phylogeographic studies due to its complex geological history and the diversity of its environments. It comprises two biogeographical zones: Eurosiberian and Mediterranean. We sampled the cosmopolitan and parthenogenetic earthworm *Eiseniella tetraedra* at 65 localities throughout Iberian Peninsula to study its genetic and morphological diversity, and possible relationships with environmental factors. We examined the molecular markers COI, 16S and 28S. *E. tetraedra* showed high diversity in the Iberian Peninsula, with eight different lineages nested in two clades. Environmental factors, such as precipitation, temperature and soil pH, influenced the distribution of genetic lineages, so that two lineages mainly present in the Eurosiberian region, one in the Mediterranean area, and three were widely throughout Iberian Peninsula. In contrast, clades were distributed according to a north-south pattern. Habitat stability by means of permanent or not availability of water also showed differences between clades and some lineages. After morphological studies of 739 specimens of *E. tetraedra*, a high diversity was found. Most of the variation showed no phylogenetic basis, although differences in length and weight were found among lineages.

2.1 Introduction

The Iberian Peninsula is considered a biodiversity hotspot, owing to a high level of endemism and a great diversity of habitats and landscapes (Myers *et al*., 2000). This could be explained by the complex geographical and geological history of the peninsula, such as Pleistocene glacial events that made it a glacial refuge (Gómez and Lunt, 2007) and as center of dispersal during interglacial periods. Despite this great diversity, the number of species could be expected to decline from the north of the Peninsula (proximal) to the south of the Peninsula (distal) due to peninsula effect (Simpson, 1964). Due to its geographical position, the Iberian Peninsula is under the influence of Atlantic Ocean and Mediterranean Sea, resulting in a wide range of climates (Gómez and Lunt, 2007) and comprising two biogeographical regions, the Eurosiberian and the Mediterranean (Alcaraz *et al*., 2006).

Eiseniella tetraedra (Savigny, 1826) is a cosmopolitan (Blakemore, 2006), parthenogenetic tetraploid (Casellato, 1987) earthworm with truly aquatic or semiaquatic habitat (Omodeo and Rota, 1991). Phylogeographic studies can help us determine the history of diversification and dispersal of a species. The high genetic diversity of some earthworms has shown that they are good models for phylogeographic studies (Fernández *et al*., 2013), and this has been the case for *E. tetraedra* (Chapter 1; Terhivuo *et al*., 1994; Terhivuo *et al*., 2011; Javidkar *et al*., 2020). In preliminary Iberian studies, high haplotypic diversity was found nested in six different lineages, without clear patterns of population structure (Chapter 1). The

same results were found in Alborz Mountains (Iran), with the presence of the same six lineages probably introduced by human activities (Javidkar *et al*., 2020).

Although several studies on earthworms showed genetic differences between Eurosiberian and Mediterranean regions in Iberian Peninsula (Fernández *et al*., 2012; Fernández *et al*., 2015; Rodríguez *et al*., 1997), *E. tetraedra* did not show a clear pattern (Chapter 1). Rodríguez *et al*. (1997) studied different environmental factors that could influence the distribution of earthworm species in the Iberian Peninsula. They suggested that transition zones between Mediterranean and Atlantic regions are characterized by mean annual precipitation in the range of 700-1000 mm. It is possible that the distribution ranges of *E. tetraedra* are not explained by large dimensions in time and space historical biogeography processes, but rather by more ecological factors that allow or limit their existence in current time.

As a riparian earthworm*, E. tetraedra* appears to be closely tied to water margins. It is found in rivers as well as streams or even unstable habitats that may be frozen or dried out depending on the season. This can be explained by the fact that it is parthenogenetic: a single propagule suffices to establish a new population (Terhivuo and Saura, 2006). Moreover, in the Aland archipelago (Baltic Sea), all individuals of *E. tetraedra* disappeared every year due to the freezing of their habitat, followed by recolonization by different clones of the species (Terhivuo and Saura, 1997). The same pattern could be present in unstable habitats of Iberian Peninsula. Thus, the difference in the distribution of lineages of *E. tetraedra* in stable or unstable habitats could provide information about this colonization pattern.

This species is characterized by a quadrangular transverse section, a small number of segments and a variable position of the male pores (more often XIII) (Omodeo and Rota, 1991). This variation has been evaluated in several taxonomic works: some (Michaelsen, 1900; Blakemore, 2006) considered the divergent forms as subspecies, others (Michaelsen, 1910; Michaelsen, 1932; Omodeo, 1952; Omodeo, 1956; Pop, 1952; Zicsi, 1960; Plisko, 1965) considered them vaguely as "formae" or "varietates", and finally others (Michaelsen, 1932; Cernosvitov, 1942; Omodeo and Rota, 1989) considered them as genetic mutations. However, it was shown in Chapter 1 that different positions in the male pore have no phylogenetic basis. Length, weight and number of segments also showed a high degree of variability, but no significant differences were found between lineages (Chapter 1).

The aim of this work is: i) to improve current knowledge of the genetic diversity of *Eiseniella tetraedra* in the Iberian Peninsula; ii) to understand its demographic history; iii) to investigate putative relationships between genetic diversity and environmental factors; iv) to test whether some lineages are better adapted to unstable habitats than others, and v)to further investigate whether the morphological variations are related to genetic diversity through the study of a high number of specimens.

2.2 Material and methods

2.2.1 Sampling and morphological studies

We collected 739 specimens of *Eiseniella tetraedra* from 65 localities around the Iberian Peninsula between 2012 and 2016. A subset of 29 localities from the central and northwestern areas of the Iberian Peninsula was previously sampled for a

microscale study (Chapter 1). For the selection of the remaining 36 localities, we used random sampling in 50x50 km UTM cells to represent its distribution along the entire Iberian Peninsula using ArcMap 9.3 software (Environmental Systems Resource Institute, ArcMap 9.3 ESRI, Redlands, California). (Supplementary Figure 1. See geographic coordinates in Supplementary Table 1).

All individuals were collected by manual sorting, washed in distilled water, fixed in 96% ethanol and stored at -20°C in the earthworm collection of the Department of Biodiversity, Ecology and Evolution, Complutense University of Madrid (UCM-LT). Morphological studies were performed in all 739 individuals focusing on: length, dry weight, number of segments, position of the clitellum and tubercula pubertatis, position of male pores, number and position of seminal vesicles. With the aim of testing the existence of individuals with biparental reproduction, the presence or absence of iridescence in spermathecae and male funnels were studied. The presence of iridescence would indicate the presence of sperm (Plisko, 2002).

Whenever possible, we selected ten individuals per locality, and a portion of the posterior body section was excised and carefully cleaned under a stereomicroscope to remove gut and soil particles. Samples were then stored in ethanol and preserved at -20°C for genetic analysis.

2.2.2 DNA extraction, gene amplification and sequencing

Total genomic DNA was extracted from the integument sample using the Speedtools Tissue DNA Kit (Biotools). Two mitochondrial markers, a fragment of cytochrome *c* oxidase subunit I (COI) and a fragment containing 16S rRNA + tRNAs Leu, Ala and Ser, and one nuclear marker (a fragment of 28S rRNA) were amplified.

For COI (632 bp) primer sequences and polymerase chain reactions (PCR) followed Pop *et al*. (2003). For 16S-tRNAs (775 bp) and 28S (806 bp) primer sequences and PCR conditions followed Fernández *et al*. (2015). All PCRs were specific and resolved via 1% agarose gel electrophoresis; they were visualised with GelRed stain (Biotium). All products were purified using ExoSAP-IT reagent (ThermoFisher Scientific).

PCR products were sequenced by Macrogen Spain Inc. Chromatograms were visualized and edited in BioEdit v7.0.9 (Hall, 1999).

2.2.3 Genetic data analyses

Sequences for each fragment were aligned in MAFFT v.7 (Katoh and Standley, 2003) using default settings and concatenated with BioEdit v7.0.9 (Hall, 1999). Haplotypes of single and concatenated genes were recovered in DNAsp v.6 (Rozas *et al*., 2017) and it was tested that the possible reading frames did not have a stop codon, to avoid the presence of pseudogenes in the dataset (Buhay, 2009). Phylogenetic trees based on the concatenated sequences of the three genes (1,906 bp) and each single gene were constructed by Bayesian Inference (BI) with MrBayes v3.1.2 (Ronquist and Huelsenbeck, 2003) and Maximum Likelihood (ML) using RaxML v7.03 software (Stamatakis, 2006) both implemented in Cipres Science Gateway v.3.3 (Miller *et al*., 2010). Phylogenetic trees obtained were visualised in FigTree v1.3.1 (Morariu *et al*., 2008). The best-fitting substitution model selected by jModelTest2 (Darriba *et al*., 2012) was GTR+Γ+I for each partition. ML analysis with rapid bootstrapping was performed with 1,000 replicates. Parameters in MrBayes were set to ten million generations and 10,000 trees were sampled every 1,000th generation, initiating the analysis with a random tree. Two independent analyses were performed and 20% of

the trees were discarded as burn-in. The remaining trees were combined to find the maximum *a posteriori* probability estimate of phylogeny. Sequences of *Carpetania matritensis* Marchán *et al*., 2020, *Lumbricus rubellus* Hoffmeister, 1843, *Dendrobaena byblica* Rosa, 1893, *Iberoscolex oliveirae* (Rosa, 1894) and *Prosellodrilus biauriculatus* Bouché, 1972 were retrieved from GenBank and used as outgroups (Supplementary Table 2).

Haplotype networks based on COI and 28S were constructed in PopART 1.7 (Leigh and Bryant, 2015) using statistical parsimony. Furthermore, Arlequin v.3.5 (Excoffier and Lischer, 2010) was used to perform an Analysis of the Molecular Variance (AMOVA) with 10,000 permutations for statistical confidence. In order to test whether population genetic structure existed at different levels, the analysis was performed following a hierarchical structure: first with main clades and then with lineages (see Results). Also, uncorrected pairwise distances for COI and 16S tRNAs were calculated within and between main clades and lineages.

We also examined haplotype and nucleotide diversity for lineages, clades and localities, and mismatch distributions, and neutrality tests such as Fu and Li's D, Fu and Li's F and Tajima's D for lineages, all calculated with DNAsp v.6 (Rozas *et al*., 2017).

2.2.4 Environmental factor analyses

Correlations between genetic diversity (haplotypic and nucleotidic) and parameters such as latitude, longitude and altitude of the localities were tested using Statgraphics Centurion 18 (StatPoint Technologies Inc., USA). Also, nineteen bioclimatic variables of WorldClim (Bio1, Bio2, Bio3…Bio19) (www.worldclim.org) and three soil properties (pH, sand and soil organic carbon) obtained from Soilgrids (https://soilgrids.org/) were

studied (Supplementary Table 3). Differences between lineages and clades (see Results) regarding environmental variables values for their localities were compared by one-way analyses of variance (ANOVA) and T-test followed by Fisher LSD post hoc test. Also, non-parametric analyses (Kruskall-Wallis and U Mann-Withney tests) with subsequent Fisher LSD post hoc test were run for those who did not fulfill assumptions of normality and homoscedasticity (verified through Kolmogorov-Smirnov and Levene's test respectively). The significance level for all statistical tests was set at α = 0.05.

We also explored possible differences in the proportion of presence between lineages and clades due to the stability of the water body in which they were found. Two categories were established: stable (rivers, permanent streams or lakes) and unstable (non-permanent streams fountains, wash tubs, etc). Thus, $\chi 2$ tests were performed in IBM SPSS Statistics v.24.

2.2.5 Morphological analyses

Statistical analyses of morphological data were conducted in Statgraphics Centurion 18 (StatPoint Technologies Inc., USA). We used length, dry weight (after letting it drip on filter paper for 30 seconds) and number of segments of mature specimens to investigate differences between lineages and clades and morphological diversity through non-parametric analyses (Kruskal-Wallis) followed by Fisher LSD post hoc test. We also examined the effects of altitudinal gradient on morphological variation using simple-linear-regression model.

2.3 Results

2.3.1 Phylogenetic analyses

The analysis of the most variable gene (COI) revealed eight distinct and well supported lineages (labeled A to H, Figure 1). These lineages were clustered into two clades (labeled I and II). Clade I included lineages B, C, D, and G and was strongly supported (0.94/0.75 for BI/ML), while clade II comprised lineages A, E, F, and H and support values were lower (Figure 1). Clade II showed higher genetic diversity with 67% of total haplotypes. The different lineages showed deep divergence as shown by the long branches of the tree (Figure 1). The tree based on concatenated sequences, COI, 16S-tRNAs and 28S (including only specimens with 28S sequences available) recovered the same lineages (except for lineage B) with high support values, and the same two distinct clades but without high support (Supplementary Figure 2). In order to improve the concatenated tree, we decided to remove sequences with high average of missing data in rake-like polytomies. This new tree recovered well supported clades and lineages (except lineage E, included now in lineage F) (Supplementary Figure 3).

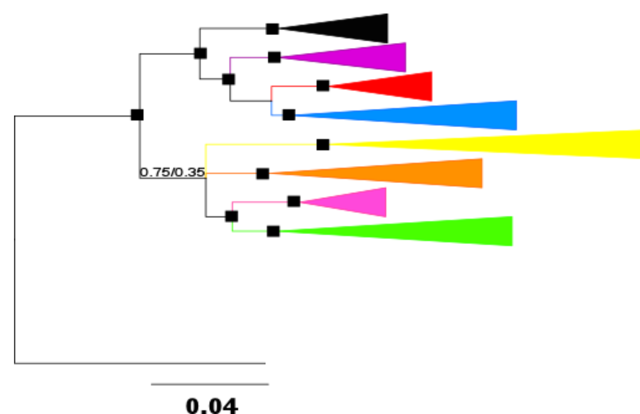


Figure 1. Bayesian inference (BI) of the phylogenetic tree based on sequences of the COI gene. Posterior probability/bootstrap support values (of Maximum Likelihood

Analysis, ML) are shown as black squares when higher than 0.9/0.7 (BI/ML). The scale bar represents 0.04 substitutions per position. Colors and names of lineages (A-H) are the same as in Chapter 1.

2.3.2 Lineages distribution

The distribution of lineages in the Iberian Peninsula is shown in Figure 2 and Supplementary Figure 4. The eight lineages were found at least in the northern half of the Iberian Peninsula. Most individuals of lineage B were distributed in the northeastern area and most individuals of lineages C and D in the northwestern area. In contrast, only lineages A, F and E were present in the southern half of the Iberian Peninsula (with the exception of lineage B in one locality). In Majorca (Balearic Islands), lineages E, F and G were found. Thus, clade I was restricted to the northern half of the Iberian Peninsula (with the exception of two localities in Cádiz and Balearic Islands), while Clade II was widely distributed throughout the Iberian Peninsula and the Balearic Islands. Localities with occurrences of a single lineage as well as localities with several lineages were found (Figure 2).

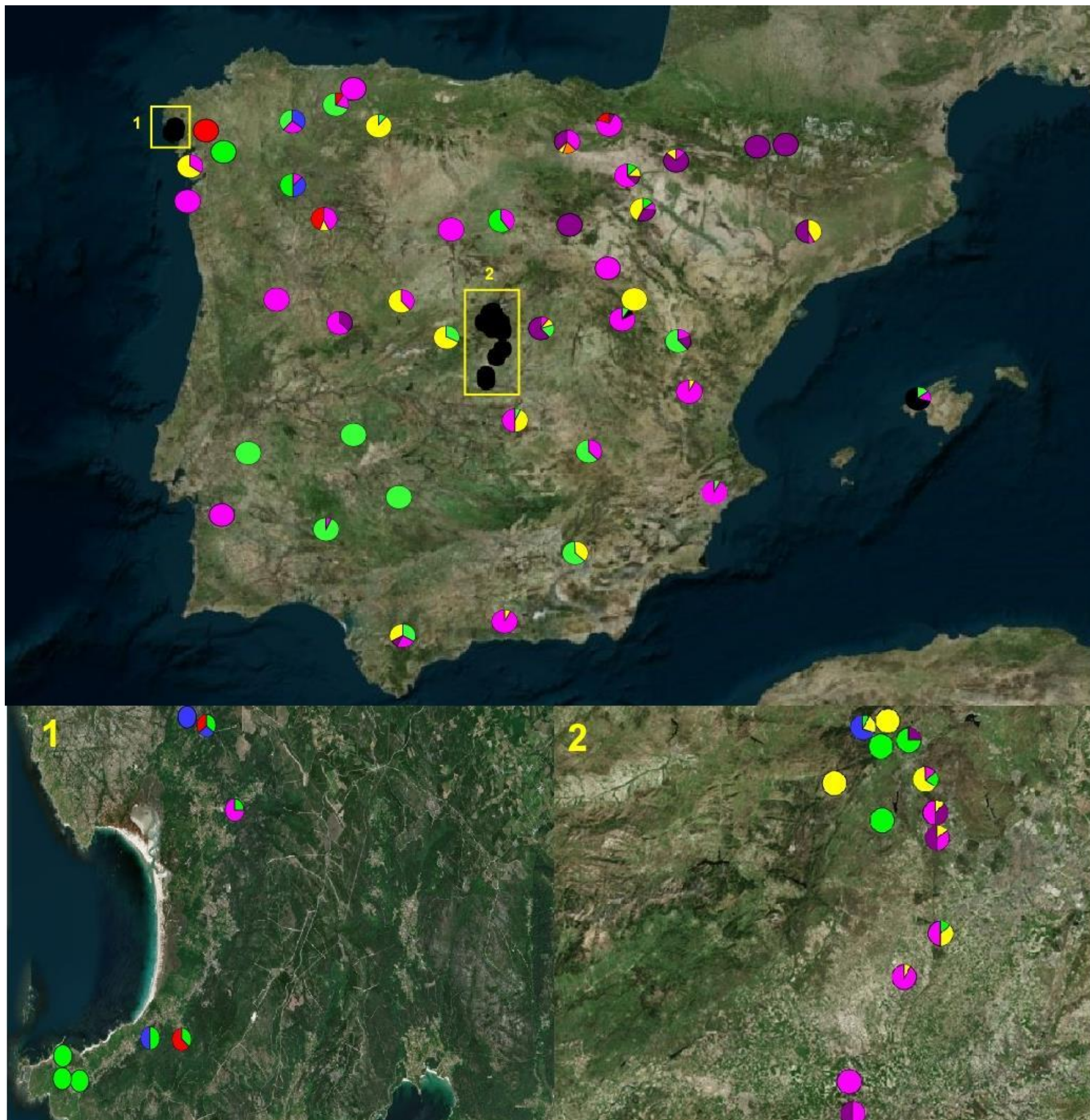


Figure 2. Lineage distribution in the Iberian Peninsula. Proportion of individuals from each genetic lineage in each locality is represented in pie charts. 1: Localities from a lower-scale study in Carnota, A Coruña, Spain (Chapter 1). 2: Localities from a lower-scale study in Guadarrama river basin, Madrid, Spain (Chapter 1). Colors used are the same as in Figure 1.

2.3.3 Genetic diversity, genetic divergence and population structure

AMOVA results indicated that most of the observed genetic variation (82.90%) was explained by differences among lineages (Table 1). Only 22.30% of the variance was explained by differences between clades and there was no genetic structure due to differences among localities. Individuals from the same locality showed haplotypes belonging to different lineages or even clades (Supplementary Table 4).

Source of Variation	**d.f.**	**Sum of squares**	**Variance components**	**Percentage of variation**
Among lineages	7	4255.067	19.91240 Va	82.90
Within lineages	261	1072.019	4.10735 Vb	17.10
Total	268	5327.086	24.01976	
Among clades	1	1609.422	6.77931 Va	22.30
Among lineages within clades	6	4326.671	17.53355 Vb	57.68
Within lineages	364	2214.289	6.08321 Vc	20.01
Total	371	8150.382	30.39607	

Table 2. AMOVA results based on the COI gene. d.f. , degrees of freedom

A total of 102 haplotypes were identified among 372 sequences for the COI gene, 34 haplotypes within 148 sequences for 16S-tRNAs, and 16 haplotypes within 128

sequences for 28S. Values of haplotype and nucleotide diversity for each lineage and clade are shown in Table 3. Haplotype diversity (H) and nucleotide diversity (π) based on COI including all the specimens within the study were 0.82 and 0.055 respectively. H and π based on 16S-tRNAs were 0.88 and 0.022. Finally, genetic diversity parameters based on 28S were 0.39/0.006.

Lineage	**Number of COI sequences**	**Number of COI haplotypes**	**H**	**π**
A	75	18	0.79	0.007
B	41	11	0.57	0.074
C	14	8	0.82	0.004
D	18	12	0.90	0.009
E	103	21	0.71	0.004
F	113	21	0.70	0.003
G	6	4	0.80	0.021
H	2	2	1.00	0.087
Clade	**Number of COI sequences**	**Number of COI haplotypes**	**H**	**π**
I	79	35	0.77	0.027
II	293	67	0.80	0.023

Table 3. Genetic diversity parameters of each lineage and each clade based on the COI gene. (H: haplotype diversity; π: nucleotide diversity).

Genetic distances within lineages based on COI was in the range of 0.25 to 3.64% showing moderate variability. Divergence between lineages was remarkably higher (2.91 to 7.79%) (Supplementary Table 5). However, values between lineages were lower than 9%-15%, the ambiguous gap between intraspecific and interspecific

divergence in earthworms proposed by Chang and James (2011). Genetic divergence between clade I and II based on COI were 10.26% for COI and 4.16% for 16S. Average divergence within clade I was 7.33%/3.24% (COI/16S), while within clade II average genetic distance was 5.21%/2.69% (COI/16S). Genetic distances based on 16S-tRNAs were lower than those based on COI, due to 16s-tRNAs being a more preserved region of the mitochondrial genome.

Most localities studied presented high haplotype diversity (average 0.70). As displayed by haplotype network based on COI from each lineage (Figure 3), most of the lineages showed a star-shaped network topology structured around highly frequent haplotypes.

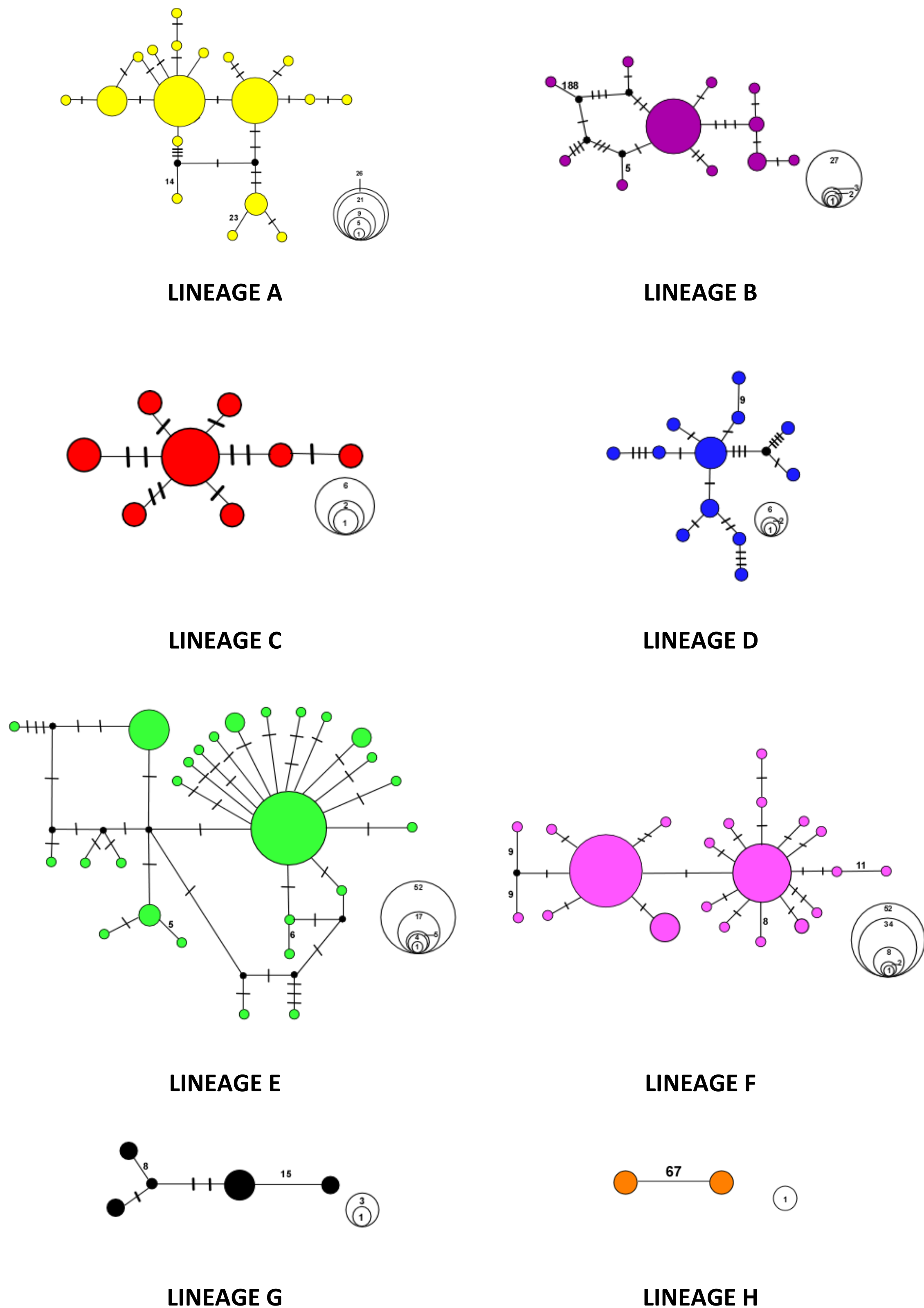


Figure 3. Haplotype network based on COI. Each colored circle represents a different haplotype; their size is proportional to the number of samples belonging to that haplotype. Each perpendicular line indicates a mutational step. Intermediate circles

are hypothetical intermediate haplotypes. Branch length does not contain any information. Colors are the same as those used in Figure 1.

The haplotype network based on the nuclear marker (28S) displayed the relationships among the sixteen haplotypes (Figure 4). A star-shaped network topology with a central main haplotype was observed. No clear differentiation between clades or lineages was observed.

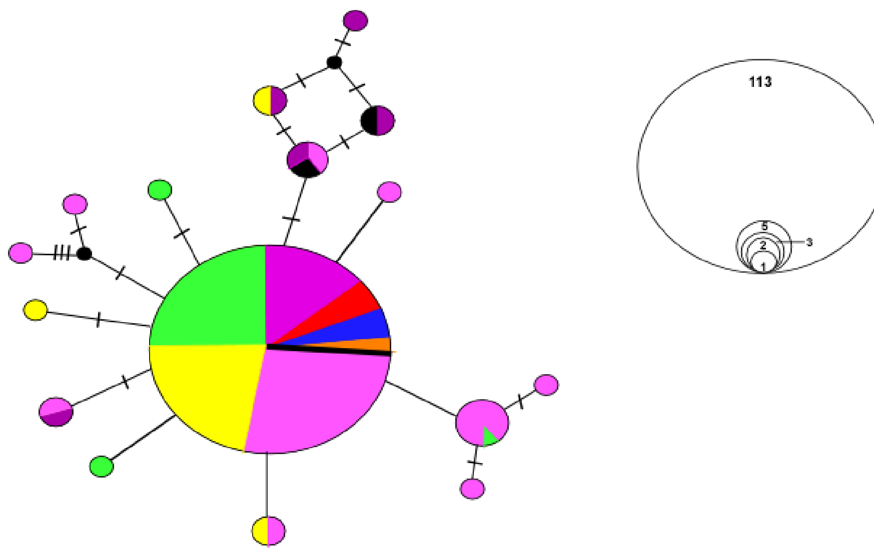


Figure 4. Haplotype network based on the 28S gene. Each circle represents a different haplotype; their size is proportional to the number of samples belonging to that haplotype. Each perpendicular line indicates a mutational step. Intermediate circles are hypothetical intermediate haplotypes. Branch length contains no information. Colors are the same as those used in Figure 1.

2.3.4 Historic demography

Mismatch distributions based on COI were tested for each of the eight lineages recovered in phylogenetic analyses (Figure 5). The distributions for lineages B, C, and D (Clade I) were not significantly different ($p>0.05$) from expectations under the sudden

expansion model. Nevertheless, the distribution for lineage E differed significantly ($p < 0.05$), and was consistent with a constant population size. Lineages A and F presented a transition pattern between unimodal and bimodal distributions. No results could be obtained for lineages G and H because of the small sample size.

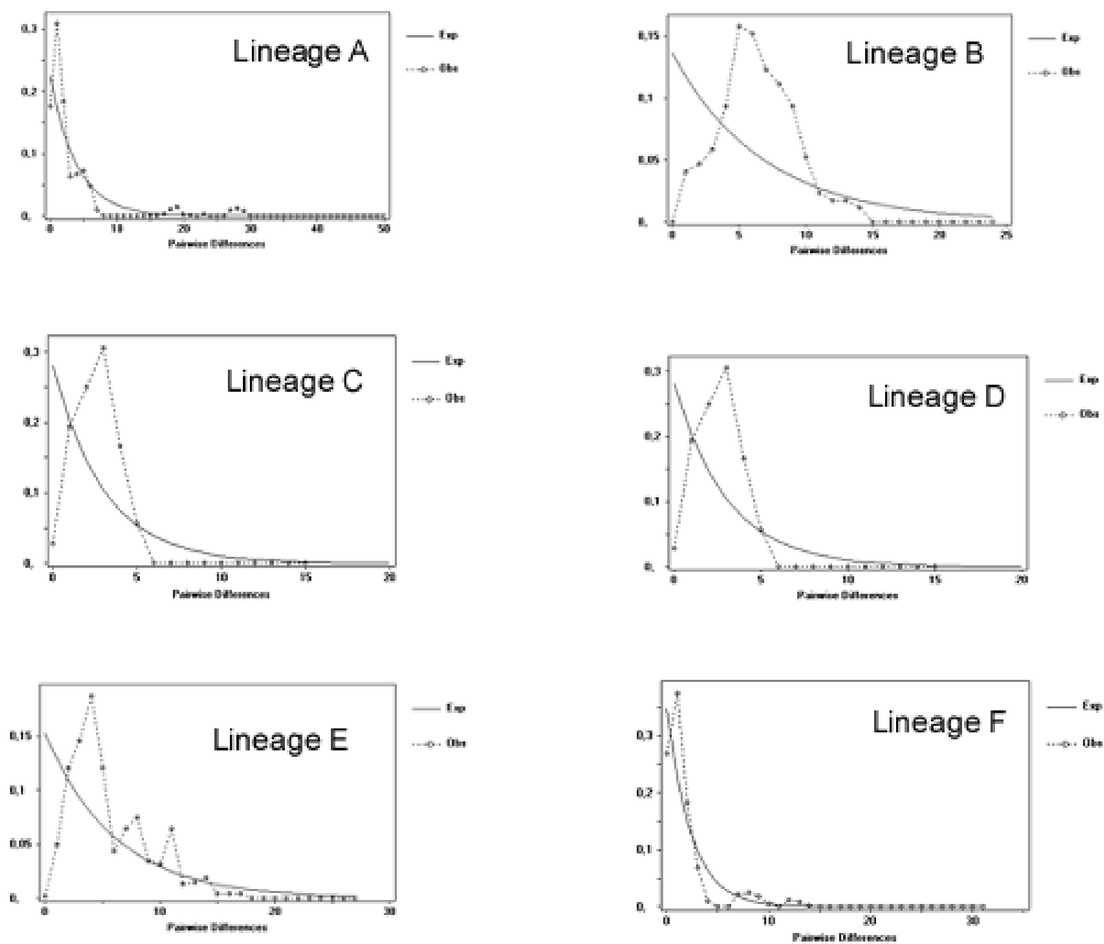


Figure 5. Mismatch distributions for lineages based on the COI gene. The abscissa represents the number of pairwise differences and the ordinate the number of observations. The smoothed line represents the observed distribution of mismatches and the discontinuous line represents the expected distribution under the sudden-expansion model of Rogers (1995) as modified by Schneider and Excoffier (1999).

The neutrality tests showed different results (Table 4). The parameters for lineages A, E, and F (Clade II) suggested a demographic expansion. For lineages B, C, and D (Clade I) only Fu's F indicated a demographic expansion. No results were obtained for lineages G and H due to the small sample size.

LINEAGE	Tajimas'D	Fu's F	Fu and Li's D*	Fu and Li's F
A	**-2,45**	**-5,30**	**-5,82**	**-5,44**
B	-1,72	**-15,37**	-2,31	-2,49
C	-1,62	**-4,26**	-1,62	-1,84
D	-1,73	**-5,97**	-1,71	-1,96
E	**-2,26**	**-19,03**	**-4,10**	**-4,08**
F	**-2,62**	**-13,66**	**-7,00**	**-6,32**

Table 4. Demographic parameters for lineages based on the COI gene. Significant values are indicated in bold.

2.3.5 Environmental factors analyses

No statistical correlations were found between altitude, longitude, latitude of the localities and genetic diversity (haplotypic and nucleotidic). Significant differences were found between lineages regarding eleven climatic variables from WorldClim (Table 5). For the temperature factors, the annual range and seasonality showed the same pattern, five different groups belonged to A, B-F, C-D, E, and G-H. The mean daily range and the mean temperature of the warmest quarter also showed five groups, but they were composed differently: A-B-F, C, D, G-E, and H; A-F, B-G-H, C, D, and E, respectively. Three groups were found for the maximum temperature of the warmest month, nesting A-B-E-F, C-D, and G-H.

In the precipitation factors, there were three groups, A-E-F, B-G-H, and C-D, for the driest and warmest quarters. In the same way, the annual values included three different homogeneous groups, A-B-E-F-G, C-D and H. For the coldest quarter we could distinguish five groups, A, B-F-G, C-D, E and H. The wettest month showed seven groups, A, B-F, C, D, E, G and H, and each lineage formed a group for the wettest quarter. Within soil factors, only pH showed significant differences between lineages with three different homogeneous groups (Table 6). In most of the cases, the statistically significant differences were due to lineages C and D. Finally, eight temperature and precipitation factors showed significant differences between clades (Table 6).

Factors	A (MEAN±SD)	B (MEAN±SD)	C (MEAN±SD)	D (MEAN±SD)	E (MEAN±SD)	F (MEAN±SD)	G (MEAN±SD)	H (MEAN)	F/K	P-VALUE
Temperature Factors (°C/10)										
Max. T Warmest Month	29.4±3.1^{B}	28.3±3.9^{B}	24.4±1.9^{A}	23.5±0.8^{A}	27.9±4.1^{B}	28.7±3.4^{B}	28.5±0.3^{AB}	27.1^{AB}	K=17.98	0.01
Mean Diurnal Range	11.0±1.8^{C}	10.9±1.4^{C}	8.3±2.1^{AB}	7.7±1.7^{A}	9.9±2.5^{BC}	10.7±2.0^{C}	11.1±2.7^{BC}	10.8^{ABC}	F=3.21	0.004
Mean T Warmest Quarter	21.2±2.1^{C}	20.1±3.6^{ABC}	18.5±0.9^{AB}	17.8±1.1^{A}	20.6±2.5^{BC}	20.8±2.2^{C}	20.7±2.3^{ABC}	19.35^{ABC}	K=17.18	0.01
Annual Range	28.5±3.6^{C}	28.2±2.8^{BC}	20.5±4.9^{A}	20.1±4.8^{A}	25.5±6.2^{B}	26.7±5.1^{BC}	27.3±5.3^{ABC}	25.6^{ABC}	K=19.32	0.007
Seasonality	626.7±76.4^{C}	622.8±60.2^{BC}	440.2±97.9^{A}	453.2±127.1^{A}	558.6±136.4^{B}	573.7±111.9^{BC}	589.9±83.3^{ABC}	544.2^{ABC}	K=18.95	0.008
Precipitation Factors (kg·m-2)										
Annual	549.0±248.8^{A}	634.2±320.5^{A}	1066.8±244.8^{B}	1101.3±226.7^{B}	716.9±356.1^{A}	686.1±354.7^{A}	540.5±38.9^{A}	675^{AB}	K=20.59	0.004
Coldest Quarter	169.2±107.5^{A}	182.1±108.5^{AB}	370.2±113.4^{C}	385.5±95.8^{C}	239.7±145.8^{B}	228.3±152.7^{AB}	151.5±3.5^{AB}	184^{ABC}	K=20.86	0.003
Driest Quarter	69.5±31.8^{A}	94.6±68.9^{AB}	124.5±28.5^{B}	117.3±18.6^{B}	78.4±40.8^{A}	77.3±38.6^{A}	65.0±24.1^{AB}	113^{AB}	K=15.60	0.02
Warmest Quarter	74.2±34.4^{A}	100.7±66.8^{AB}	141.0±31.9^{B}	135.0±26.3^{B}	90.5±46.5^{A}	87.4±43.0^{A}	85±0^{AB}	113^{AB}	K=16.23	0.02
Wettest Month	70.8±35.4^{A}	77.9±36.7^{AB}	145.3±38.9^{CD}	149.7±30.7^{D}	98.6±50.5^{B}	92.5±50.9^{AB}	74.5±9.2^{ABC}	74^{ABCD}	K=22.11	0.002
Wettest Quarter	194.9±101.9^{A}	212.7±102.9^{AB}	403.8±110.4^{CD}	423.0±89.2^{D}	272.5±144.8^{B}	2558±147.2^{AB}	197.0±7.1^{ABC}	202^{ABCD}	K=21.09	0.003
Soil factors (pH·10)										
pH	72.7±4.2^{C}	70.7±5.9^{BC}	57.5±5.8^{A}	57.3±8.5^{A}	66.0±8.9^{B}	68.1±9.1^{B}	76^{BC}	72^{ABC}	K= 21.7	0.002

Table 5. Statistical results for each lineage regarding temperature, precipitation and soil factors in ANOVA (F) or Kruskal-Wallis (K). SD: standard deviation. Different letters in the same row indicate different groups in multiple range test.

Factors	CLADE I (MEAN±SD)	CLADE II (MEAN±SD)	W	P-VALUE
Temperature Factors (°C/10)				
Max. T Warmest Month	26.6±3.7	28.6±3.6	1897.5	0.01
Mean T Driest Quarter	18.3±4.2	19.8±3.8	1886.0	0.01
Mean T Warmest Quarter	19.4±2.8	20.8±2.2	1913.5	0.009
Precipitation Factors (kg·m-2)				
Annual	807.9±354.4	663.1±333.8	1076.0	0.03
Driest Month	26.4±16.9	18.7±11.6	1041.0	0.01
Driest Quarter	103.2±54.4	72.2±36.7	1008.5	0.01
Warmest Quarter	114.6±54.8	85.5±42.4	983.5	0.007
Wettest Month	105.5±48.4	89.2±53.8	1118.0	0.05

Table 6. Statistical results for each clade regarding temperature, precipitation and soil factors in U Mann Whitney test. SD: standard deviation.

We distinguished 26 stable and 38 unstable habitats within the sampled localities (Supplementary Table 1). The results of $\chi 2$ test are presented in Figure 6. Lineage D was significantly more frequent in stable habitats, whereas lineage E was significantly more frequent in unstable habitats. Moreover, Clade I appeared more in stable habitats, while Clade II was found more frequently in unstable habitats.

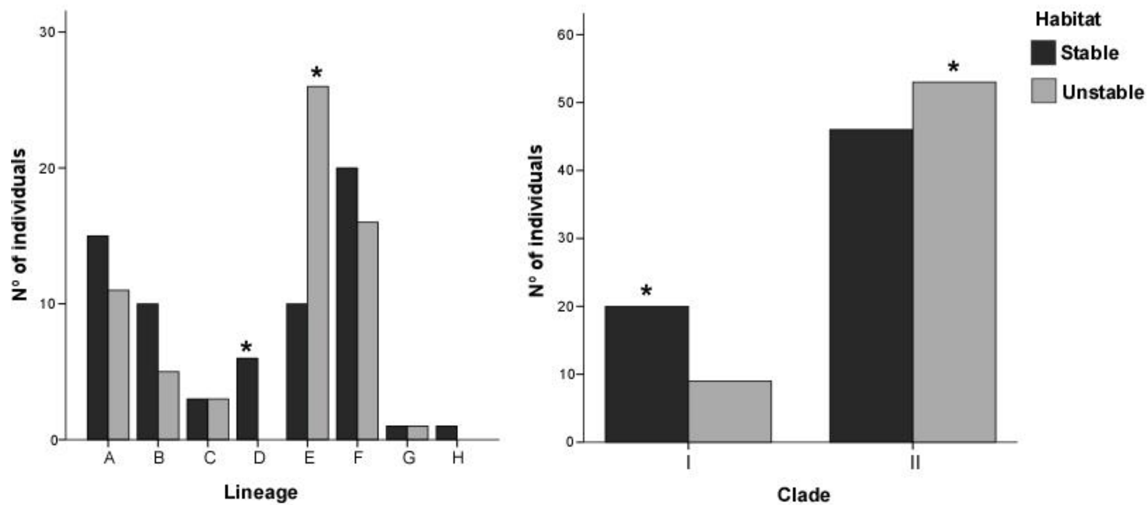


Figure 6. Results of χ2 test for lineages and clades and number of habitats stable or unstable. Asterisks indicate statistically significant differences according to χ2 test ($p<0.05$).

2.3.6 Morphological analyses

High morphological variability was observed in the examined specimens. Only mature individuals were used for the morphological analyses. Spermathecae and male funnels were never iridescent and were even absent in 66.25% and 69.16% of the individuals, respectively. We found variability in the position of the male pore. In most of the specimens (95.22%), the male pore was located in segment 13. However, in 2.57%, male pore was found in segment 12; in 0.73%, in segment 15; and in 0.37% of the specimens, in segments 8, 9, 11 or 14. Individuals sharing the same haplotype had different states of this trait, and no association was found between the position of the male pore, position of clitellum or tubercula pubertatis (also showing some variability, Supplementary Table 6) and genetic lineages.

Although the number of segments showed a high degree of variability between individuals, no significant differences were found between lineages or clades. However, statistically significant differences ($p<0.05$) were found in the length and weight of specimens between lineages (Supplementary Figure 5), but none between clades. Regarding length, specimens of lineage C were significantly longer than specimens from the others lineages. Within weight lineage E was significantly heavier than lineage A. Also, lineages D and E were significantly heavier than lineage B. Lineages G and H were excluded from these analyses because mature specimens were not available.

No significant results were obtained when studying the effects of altitudinal gradient on morphological variation.

2.4 Discussion

Several studies on earthworms showed that most species have high number of distinct genetic lineages (King *et al*., 2008; Porco *et al*., 2013). In a previous study of *Eiseniella tetraedra* in two different regions in the Iberian Peninsula (Chapter 1) six different lineages were found. Owing to the fact that in this work the localities sampled (covering the whole peninsula) have been increased, phylogenetic analysis based on COI of Iberian *E. tetraedra* revealed eight distinct lineages nested in two clades. Since COI is a more variable region of the mitochondrial genome, the lineages had only good support and were recovered in the phylogenetic tree based on this gene. Genetic divergence also showed that these lineages were internally homogeneous, suggesting that not enough time may have elapsed for differentiation of haplotypes within each lineage due to evolutionary forces such as regular bottlenecks caused by a constant

founder effect or selective sweeps. Additionally, high haplotype diversity was found for *E. tetraedra* across the Iberian Peninsula. The majority of haplotypes were unique (found in one individual at one site), as Knott and Haimi (2010) found for *Dendrobaena octaedra* Savigny, 1826 in Finland.

We found that lineages of *E. tetraedra* clustered within two clades (I and II) might be in a process of cryptic speciation, due to their high COI divergences, 10.26%, which are within the 9% and 15% interval between intraspecific and interspecific divergences proposed by Chang and James (2011), although the lack of support in phylogenetic trees suggest that this process may be incipient.

Other parthenogenetic species have also shown cryptic divergent lineages, such as *Octolasion tyrtaeum* Savigny, 1826 (Heethoff *et al*., 2004) or *Aporrectodea trapezoides* (Dugés, 1828)*,* which showed underlying speciation for lineages I and II (Fernández *et al*., 2011). This also seems to be the case for *E. tetraedra*.

The three sources of information for historic demography examined in this paper (haplotype and nucleotide diversity, mismatch distributions, and neutrality tests) showed contrasting results, except for lineage C, which clearly appears to be in demographic expansion. Regarding lineages B and D, Fu's F test and mismatch distributions indicated demographic expansion, although their high haplotype and nucleotide diversity suggested demographic stability. Neutrality tests of the most widespread lineages in the Iberian Peninsula (A, E, and F) suggested demographic expansion. However, this is not clear in their mismatch distributions. Harpending (1994) noted that an excessively recent population expansion, e.g. at the end of the Pleistocene, would not result in a smooth unimodal mismatch distribution (as would

be expected in such a scenario). It is known that the Iberian Peninsula served as a refuge during Pleistocene glaciation (Gómez and Lunt, 2007). Thus, these lineages may have suffered bottleneck events in the past and their population expansion in the Iberian Peninsula may have begun recently.

The geographic distribution of lineages and the scattered geographic distribution of haplotypes within lineages support the lack of a strong population structure, which was also supported by AMOVA results. However, three patterns can be defined, first, the lineages that were present throughout the peninsula: A, E and F (Clade II). In contrast, there were lineages that were present mainly or only in half of the northern region of the peninsula: B, C, and D (Clade I). Lineage C was found only in the northwestern region of the Iberian Peninsula (which corresponds to Eurosiberian region). Lineage D was mainly present in the Eurosiberian region with the exception of one locality in the central area which has similar environmental factors to this region. In contrast, lineage B was distributed in the northeastern region (appearing mainly in Mediterranean areas). The presence of lineage B at a site in the south could be due to human introduction activity, which has been previously reported for *E. tetraedra* (Blakemore *et al*., 2006; Brown *et al*., 2006; Javidkar *et al*., 2020). These patterns suggest the existence of two Eurosiberian lineages (C and D), one Mediterranean lineage (B) and three lineages distributed throughout the peninsula. Other phylogeographic studies of cosmopolitan and polyploid earthworms in the Iberian Peninsula, such as *A. trapezoides* and *Aporrectodea rosea* Dugés, 1828, showed similar phylogeographic patterns (Fernández *et al*., 2012; Fernández *et al*., 2015). They found two distinct clades: one present in Eurosiberian region and the other in the Mediterranean region. Finally, lineages G and H were present in only one or two

localities, so no robust conclusions can be drawn about them, although they could be explained by point mutations in these populations.

The geographical distribution of the clades was more distinct. Clade I was present only in the northern half of the peninsula (with the exception of one locality in the south, which can be explained by human activity, as mentioned before), while clade II was distributed throughout the peninsula. We found more genetic variability in the northern half of Iberian Peninsula, which may respond to adaptations to local environments of individuals and not to peninsula effect (Simpson, 1964), as a gradual loss of genetic variability from the continent was not found.

The distribution of clades and lineages within *E. tetraedra* could be partially explained by the detected ecological preferences as temperature, precipitation and pH. Phillips *et al*., (2020) found that precipitation, followed by habitat cover and temperature were the most important drivers shaping diversity and distribution patterns in earthworms. Also, pH influence earthworms' communities (Rutgers *et al*., 2009; Rutgers *et al*., 2016). Lineages C and D occurred in damp and cold localities with more acidic soil, due to their limited distribution in the northwestern quadrat, corresponding to the Eurosiberian region. In contrast, lineages A, B, E, and F occurred in dry and hot localities. Differences between clades were found for temperature and precipitation, probably due to their north-south distribution.

As is commonly known, *E. tetraedra* is closely tied to edges of water, regardless of their stability. Clade I was found to be more adapted to stable habitats, while Clade II appeared more unstable habitats. Since Clade II occurred throughout the Iberian Peninsula (including the south where environmental conditions are less idoneous and

more unstable habitats are found) it can be assumed to have greater resistance to poor conditions and greater colonization potential, which could also explain its distribution.

Variability in morphological characters related to sexual reproduction, such as seminal vesicles or spermathecae, was observed in Iberian specimens. Gavrilov (1939) found variability in the number of spermathecae and attributed it to the possibly gradual evolution of parthenogenesis in *E. tetraedra*. Different positions of the male pores, also found in other studies (Gavrilov, 1939; Bouché, 1972; Gates, 1977; Terhivuo *et al.*, 1994; Blakemore, 2006), could be explained by the same reason. Although several varieties or even subspecies have been described based on this trait (Bouché, 1972; Gates, 1977; Blakemore, 2006), no phylogenetic basis has been found (Chapter 1). No evidence of sexual reproduction was found in 739 specimens, so *E. tetraedra* appears to be strictly pathenogenetic in the Iberian Peninsula. This was not the case for *A. trapezoides*, with extremely rare sexual forms in the Iberia Peninsula and Algeria (Fernández *et al.*, 2011; de Sosa *et al.*, 2017).

Only individuals from lineage C can be clearly distinguished from the others by length (>50 mm). In addition, individuals from lineages D and E were heavier than individuals from lineage B. Specimens from lineage E were also heavier than specimens from lineage A.

2.5 Conclusions

Eight different lineages (two of them extremely rare) for *Eiseniella tetraedra* were found in the Iberian Peninsula, nested in two different clades. The distribution of

clades was neatly differentiated, with Clade I limited to the northern half and Clade II present through the whole peninsula. The distribution of lineages may respond to environmental and soil factors such as temperature, precipitation, and pH. Lineages C and D occurred mainly in the Eurosiberian region, while lineages A, B, E, and F were predominant in the Mediterranean region, although they were present in the whole Iberian Peninsula. Genetic divergence based on the COI gene for clades suggested an underlying cryptic speciation of *E. tetraedra*. Historical demography showed demographic expansion for most lineages, and recent dispersal of lineages A, E and F in the Iberian Peninsula, probably after Pleistocene glaciation. Most of the morphological variation showed no phylogenetic basis, although differences in length and weight were found among lineages.

3. Chapter 3: The nunatak and tabula rasa hypotheses may be compatible: the European phylogeography of a riparian earthworm.

Abstract

The *tabula rasa* hypothesis of postglacial immigration supports the notion that species now found in northern European areas must have been recently recolonized from historical refugia. Until the 1960s, however, there was almost complete consensus that disjunctions and endemism in the North Atlantic could not be explained without in situ survival during glacial periods (the nunatak hypothesis). Although some earthworms can survive in permafrost and tolerate cold conditions, it is generally believed that most earthworms were eradicated from northern latitudes during the Last Glacial Maximum. To test which hypothesis explains the phylogeography of the riparian and parthenogenetic earthworm *Eiseniella tetraedra*, we collected 1,640 specimens from 19 different countries in Europe. We examined three molecular markers (COI, 16S and 28S) and their morphology. Eleven lineages were found, nested in five clades. Clade I was more prevalent in cold biogeographical regions such as the Continental, the Atlantic or even the Arctic, while clade II was prevalent in Mediterranean regions. We investigated their potential niches through Species Distribution Models, which agreed with the distribution trends. The presence of restricted clades in the Iberian and Scandinavian peninsulas, as well as in Eastern Europe, suggests that these three regions served as refugia during the Last Glacial Maximum. Thus, both hypotheses were necessary to explain the actual distribution of this shore-dwelling earthworm.

3.1 Introduction

Nowadays, earthworms include about 6,500 species described, of which 3,000-3,500 are valid (Csuzdi, 2012). The family Lumbricidae Rafinesque- Schmaltz, 1815 includes about 300 species (Csuzdi, 2012) and originated in the Lower Cretaceous in the Holarctic region (Dominguez *et al*., 2015). Earthworms belonging to this family are the most abundant invertebrates in the soil of temperate regions and account for 90% of invertebrate biomass (Edwards, 2004). Earthworms are not able to live in permafrost for long periods of time (Holmstrup *et al*., 1991). Therefore, it is generally assumed that earthworms were eradicated from northern latitudes during the Last Glacial Maximum (LGM) (Tiunov *et al*., 2006), so that species now found in northern European areas must have been recently recolonized from historical refugia such as the southern European peninsulas. This basic expansion-contraction model is known as the "*tabula rasa*" hypothesis. In contrast, the nunatak hypothesis suggests that some ice-free refuge existed in northern Europe, such as mountains rising above inland ice or coastal ice-free refuges, where the biota could survive. When the ice melted, plants and animals recolonized previously ice-covered areas from these northern refugia. There is now strong geological evidence suggesting that some nunataks and ice-free coastal shelves existed within the maximum limits of the last glacial period 25,000–10,000 years ago (Hansen *et al*., 2006). Although nunatak hypothesis seems unlikely for earthworms, it was proposed for Fennoscandia (Fridolin 1936, Stöp-Bowitz 1969) and Greenland (Hansen *et al*. 2006). Moreover, some hardy Lumbricidae species could survive the last glacial period in ice-free refugia in association with some arctic plants

(Julin, 1949; Stöp-Bowitz, 1969). *Dendrobaena octaedra* (Savigny, 1826) overwinters in frozen ground either as adults or in cocoons. In Greenland, it exhibited high genetic diversity, which, combined with its high frost tolerance, suggests its survival in ice-refuge in Greenland (Hansen *et al*., 2006), following the nunatak hypothesis. Phylogeographic studies can help us determine the history of diversification and dispersal of a species and earthworms with high genetic diversity are good models for these studies (Chapter 1; Chapter 2; Fernández *et al*., 2013; Shekhovtsov *et al*., 2020). A phylogeographical study covering the whole Europe, might therefore throw light on these hypotheses.

Eiseniella tetraedra (Savigny, 1826) is a parthenogenetic and tetraploid (Casellato, 1897) earthworm with a riparian lifestyle that inhabits margins of freshwater. It has a worldwide distribution and is therefore referred as a cosmopolitan earthworm (Blakemore, 2006). *E. tetraedra* showed a high genetic diversity in the Iberian Peninsula with eight different lineages nested in two clades. This genetic diversity was distributed according to three environmental factors: temperature, precipitations and pH. Thus, three linages were found in the northern half of the Iberian Peninsula (namely B, C and D), three lineages were distributed throughout the Peninsula (A,E and F) and two of them were restricted geographically and appeared in only one (H) or two (G) populations (Chapter 2). Jadvikar *et al*. (2020) found only six lineages in Iran, probably introduced by human activity, while in Chapter 1, only one lineage in Scandinavia was found, which could be due to a limited number of samples. *E. tetraedra* was found to the north of the 65 parallel (Haraldsen and Engelstad 1998), and was detected as far as the northern coast of the Scandinavian Peninsula (Terhivuo 1988). It was also found in Iceland (Blakemore, 2007). According to these presences, *E.*

tetraedra seems to be a cold tolerant earthworm. However, Terhivuo and Saura (1997) showed in the Aland Archipelago (Baltic Sea) that all individuals of *E. tetraedra* disappeared each year due to habitat freezing, followed by recolonization by different clones of the species.

Ecological Niche Modeling (ENM) with MaxEnt (Phillips *et al*., 2006) has facilitated ecological inference in soil due to its high power when only presence data are included. It has been implemented in several groups such as termites (Maynard *et al*., 2015), beetles (Crawford and Hoagland, 2010), millipedes (Marek *et al*., 2012) and earthworms (Marchán *et al*., 2015; Marchán *et al*., 2016). Moreover, macroecological preferences in soil communities were studied by Decaëns (2010). The macroecological preferences of *E. tetraedra* have been studied (Si-Moussi, 2020) and soil texture was the most discriminating factor for this species, but nothing is known about its clades. Marchán *et al*. (2016) found ecological divergence among four different clades of the earthworm family Hormogastridae, which showed different environmental responses and preferred different habitats.

The present study investigated the genetic variation and phylogeographic relationships of a large number of populations of *E. tetraedra* collected from nineteen countries in Europe. The aim of this study was: i) to examine the genetic diversity of *E. tetraedra* in Europe; ii) to infer its phylogeography in this continent and iv) to test macroecological preferences for clades.

3.2 Material and methods

3.2.1 Earthworm sampling and morphological analyses

We collected 1,640 specimens from 19 different countries: Belgium, Bulgaria, Czech Republic, Finland, France, Germany, Greece, Israel, Italy, Netherlands, Norway, Poland, Portugal, Russia, Slovakia, Spain, Sweden, Turkey and United Kingdom (Supplementary Tables 1, 2, 3, 4, 5). Earthworms were collected by hand-sorting and fixed in 96% ethanol and stored at -20°C. Whenever possible, we selected ten individuals per locality, and a portion of the posterior body section was excised and carefully cleaned under a stereomicroscope to remove gut and soil particles. Tegument samples were stored in ethanol and preserved at -20°C until DNA extraction. Morphological characters were studied on 1,147 specimens, focusing on: length, dry weight, number of segments, position of clitellum and tubercula pubertatis, position of male pores, number and position of seminal vesicles, spermathecae and spermiducal funnels.

3.2.2 DNA extraction, gene amplification and sequencing

Total genomic DNA was extracted using the Speedtools Tissue DNA Kit (Biotools). We amplified a fragment of cytochrome *c* oxidase subunit I (COI) in all specimens, and we chose two specimens per site and lineage (see Results) for amplification of a fragment containing 16S rRNA + tRNAs Leu, Ala and Ser (16S), and a fragment of 28S rRNA (28S). For COI (632 bp) primer sequences and polymerase chain reactions (PCR) followed Pop *et al*. (2003). For 16S-tRNAs (775 bp) and 28S (804 bp) primer sequences and PCR conditions followed Fernández *et al*. (2015). All PCRs were specific, resolved via 1% agarose gel electrophoresis and were visualised with GelRed stain (Biotium). PCR

products were purified using ExoSAP-IT reagent (ThermoFisher Scientific) and sequenced by Macrogen Spain Inc.

3.2.3 Phylogenetic analyses and genetic variability

Sequences of each gene were aligned in MAFFT v.7 (Katoh and Standley, 2003) using default settings and concatenated with BioEdit v7.0.9 (Hall, 1999). Phylogenetic analyses with the concatenated sequence of the three genes (2,213 bp) included Bayesian inference (BI) using MrBayes v.3.1.2 (Ronquist and Huelsenbeck, 2003) and Maximum Likelihood (ML) with RaxML v.7.03 software (Stamatakis, 2006) both implemented in CIPRES Science Gateway v.3.3 (Miller *et al*., 2010). Phylogenetic trees obtained were visualised in FigTree v1.3.1 (Morariu *et al*., 2008). The best-fitting substitution model selected by jModelTest2 (Darriba *et al*., 2012) for each partition was GTR+Γ+I. The Bayesian analyses consisted of two parallel runs of ten million of generations, sampling 10,000 trees every 1,000th generation, starting the analysis with a random tree. 20% of the trees were discarded as burn-in. ML analysis with rapid bootstrapping was performed with 1,000 replicates. Sequences of *Lumbricus rubellus* Hoffmeister, 1843, *Dendrobaena byblica* Rosa, 1893, *Iberoscolex oliveirae* (Rosa, 1894), *Prosellodrilus biauriculatus* Bouché, 1972 and *Carpetania matritensis* Marchán *et al*., 2020, were retrieved from GenBank and used as outgroups (Supplementary Table 6).

Uncorrected pairwise distances for COI were calculated within and between main clades (see Results). We also examined haplotype and nucleotide diversity for clades and for the three peninsulas sampled: Iberian, Italian and Scandinavian.

3.2.4 Ecological niche modeling

Presence data were aggregated according to the two principal clades (I and II) recovered in the phylogenetic analyses (see Results). The following large-scale variables were chosen as predictor variables. They are the same (or equivalent) as the variables found as the most influential for the distribution of *Eiseniella tetraedra* by Si-moussi (2020).

- *Bioclimatic variables* (downloaded from Worldclim, http://www.worldclim.org/ accessed 01/12/2020)
    - Min Temperature of Coldest Month (BIO6)
    - Temperature Annual Range (BIO7)
    - Precipitation of Wettest Month (BIO13)
- *Soil variables*
    - Lithology (PARMA), represented by the PAR-MATDOM2 (Second level code for the dominant parent material of the STU) layer obtained from the European Soil Database Raster Library 1 km _ 1 km (http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDB_data_1k_raster_intro/ESDB_1k_raster_data_intro.html accessed 01/12/2020).
    - Soil crusting class (CRUST) layer obtained from the European Soil Database Raster Library 1 km _ 1 km (http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDB_data_1k_raster_intro/ESDB_1k_raster_data_intro.html accessed 01/12/2020).
    - Topsoil available water capacity (AWC) layer obtained from the European Soil Database Raster Library 1 km _ 1 km

(http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDB_data_1k_raster_intro/ESDB_1k_raster_data_intro.html accessed 01/12/2020).

- *Biotic variables*
    - Vegetation and dominant land use (CLC) were represented by the CORINE 2018 Land Cover layer (version v.2020_20u1: https://land.copernicus.eu/pan-european/corine-land-cover/clc2018?tab=download).

Ecological niche models were obtained using the R package 'SSDM' (Schmitt *et al*. 2017). Presence (localities were the target clade was found) and absence (localities were the other clade -but not the target clade- was found) data was used as input. Ensemble species distribution models (ESDMs) were built by combining the algorithms ('MAXENT', 'GLM', 'CTA' and 'MARS') producing kappa values greater than 0.5, with three repetitions for each algorithm.

3.3 Results

3.3.1 Phylogenetic analysis

Bayesian and Maximum Likelihood approaches yielded trees with congruent topology for *Eiseniella tetraedra* (Figure 1). All sequences, except one (LSWED303 from Sweden) were nested in five different and well supported clades (labeled I to V). Clades I and II included four different lineages (B-C-D-G, A-H-E-F, respectively), whereas clades III to V comprised only one each (I, J and K). Clade IV showed the lowest genetic diversity with only three different haplotypes, while clade II was the most diverse with 35.76% of the total haplotypes.

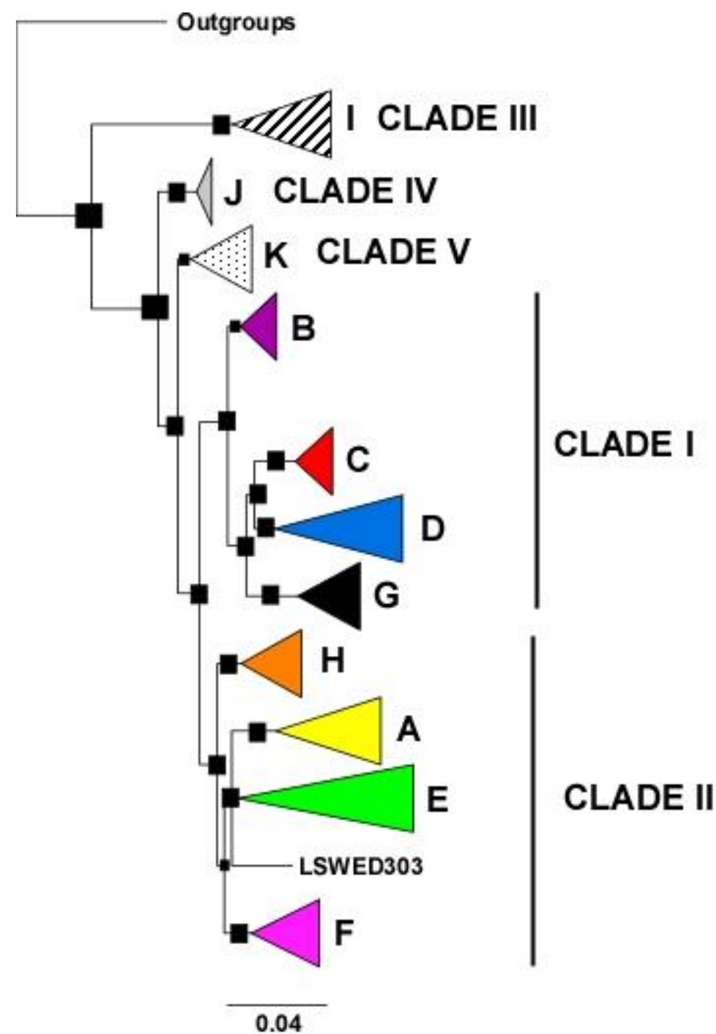


Figure 1. Bayesian inference (BI) of the phylogenetic tree based on concatenated sequences of COI, 16S and 28S. Posterior probability/bootstrap support values (of Maximum Likelihood Analysis, ML) are shown as black squares when higher than 0.9/0.7 (BI/ML). The scale bar represents 0.04 substitutions per position. Colors and names of lineages (A-H) and number of clades (I and II) are the same as in Chapter 2 and in Figures 2 and 3. LSWED303 correspond to a mature specimen from Frösslundabäcken Stream, Öland, Sweden.

3.3.2 Lineages distribution

The distribution of lineages in Europe is shown in Figures 2 and 3. In the Iberian Peninsula, the northern area was more diverse than the southern one, and lineages D and H were present only in this area of Europe (Figures 2 and 3). We found all lineages in the Scandinavian Peninsula (Figure 2.5), except those which were present in only a few populations (D, G, H and I). Lineage B was clearly predominant in all areas and lineage J and K occurred mostly in this countries. The Italian Peninsula is probably one of the least diverse areas (Figure 2.2). Only lineages from the clade II (A, E and F) were found. North-South differences were observed in the Italian Peninsula. Lineage A predominated in the northern area, while lineage F was more numerous in the central and southern areas. The two populations of United Kingdom showed high diversity (Figure 2.1). Only lineages A, E and F were represented at Central Europe with the exception of two individuals of lineages G in the Netherlands and J in Slovakia (Figure 2.3). The eastern regions of Europe were less represented and lineages A, F and J were found (Figure 2.1). The distribution of lineages in France is shown in Figure 2.4. In northern France, lineage B was predominant, while South France showed greater diversity, especially in the Pyrenees. Therefore, clades I and II were widespread in Europe, although clade I was most common in cold areas and clade II in temperate ones. In contrast, clades III, IV and V had a restricted distribution, with clade III occurring mainly in Eastern Europe and clades IV and V found mainly in Scandinavian Peninsula.

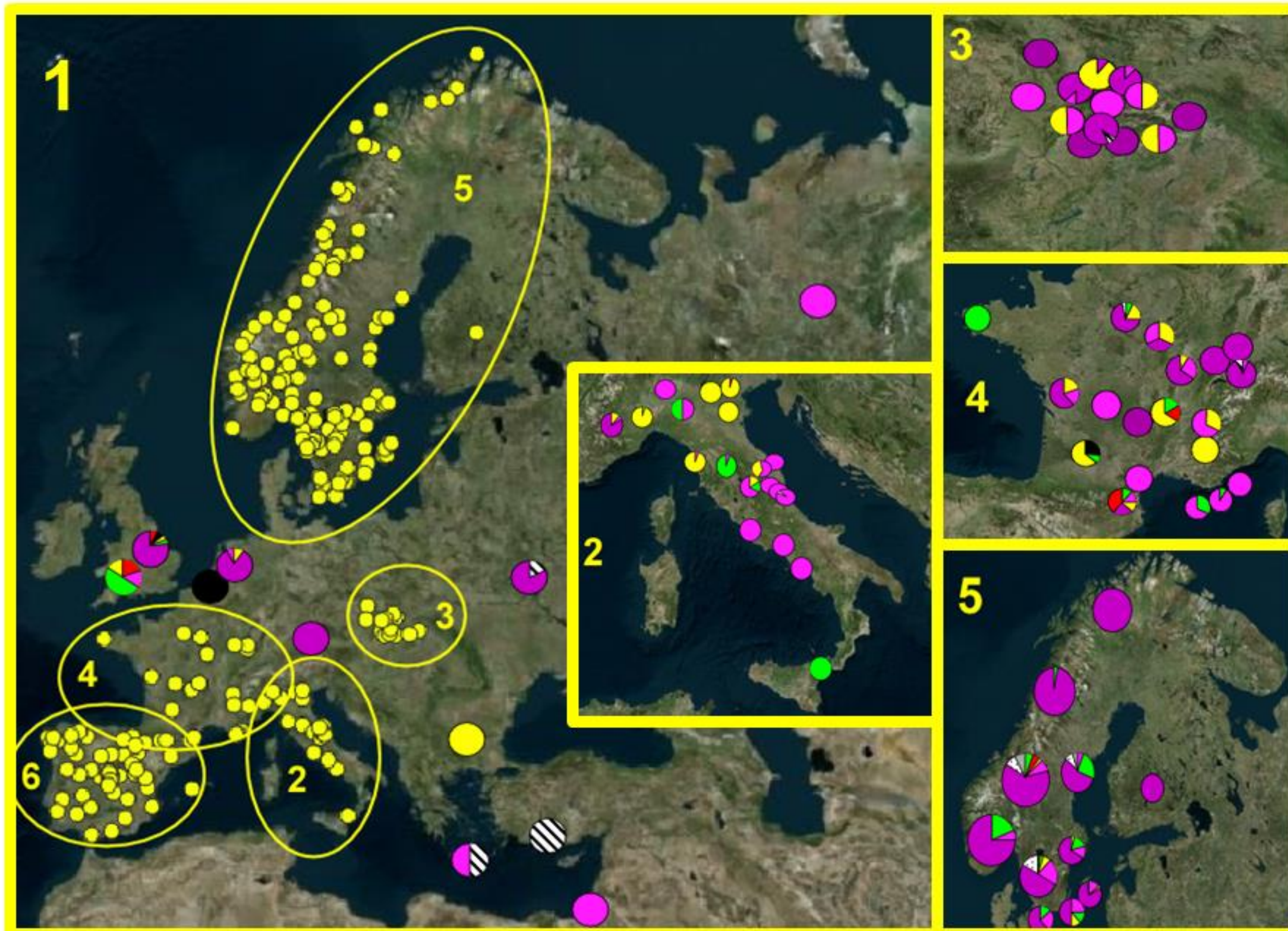

1
5
3
4
6
2
2
3
4
5

Figure 2. Lineages distribution in Europe. Proportion of individuals from each genetic lineage in each locality is represented in pie charts. 1: Lineages distribution in distinct parts of Europe. 2: Lineages distribution in the Italian Peninsula. 3: Lineages distribution in Slovakia, Poland and Czech Republic. 4: Lineages distribution in France. 5: Lineages distribution in the Scandinavian Peninsula. Lineages pie charts show information on several closed populations. The size of circles is proportional to the number of localities it includes. Colors used are the same as in Figure 1. 6: It is showed in Figure 3.

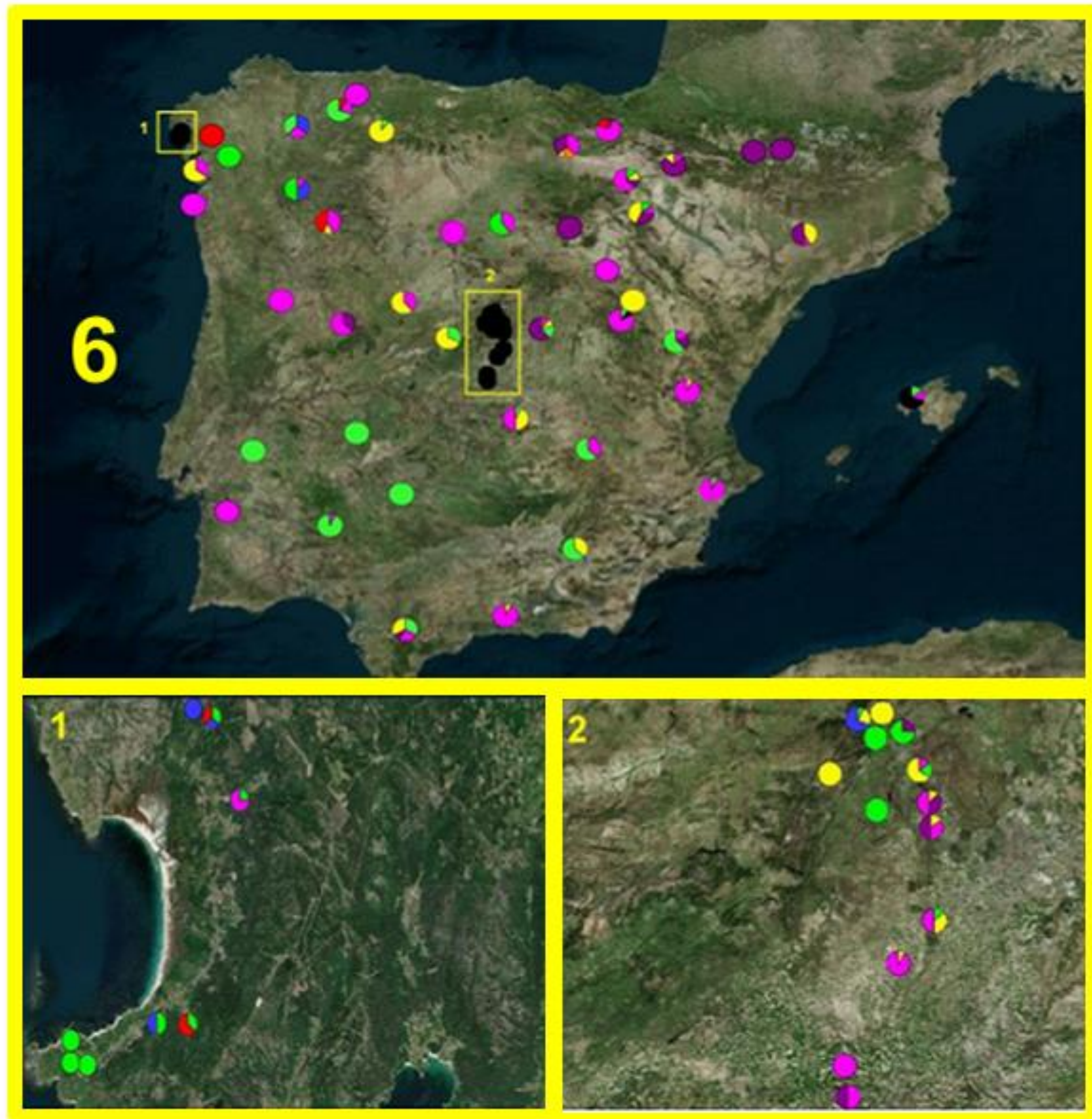


Figure 3. Lineages distribution in the Iberian Peninsula (modified from Chapter 2). 5.1: Localities from a lower-scale study in Carnota, A Coruña, Spain (Chapter 1). 5.2: Localities from a lower-scale study in Guadarrama river basin, Madrid, Spain (Chapter 1). Colors used are the same as in Figure 1.

3.3.3 Genetic diversity and genetic divergence

A total of 485 haplotypes were found among 1,038 sequences for the COI gene, 153 haplotypes within 262 sequences for 16S and 187 haplotypes within 378 sequences for

28S. Values of haplotype and nucleotide diversity for each clade are shown in Table 1. The clade II showed the greatest diversity, while all specimens of clade IV belonged to the same haplotype. Haplotype diversity (H) and nucleotide diversity (π) based on COI including all the specimens within the study were 0.97 and 0.059 respectively. H and π based on 16S-tRNas were 0.88 and 0.022. Finally, genetic diversity parameters based on 28S were 0.33 and 0.003. Moreover, the genetic diversity (H/π) parameters in the three peninsulas studied were: 0.82/0.055 for the Iberian Peninsula, 0.91/0.06 for the Italian Peninsula and 0.65/0.06 for the Scandinavian Peninsula.

	Number of COI sequences	**Number of COI haplotypes**	**H**	**π**
CLADE I	415	193	0.93	0.027
CLADE II	553	281	0.97	0.035
CLADE III	5	4	0.9	0.038
CLADE IV	8	1	0	0
CLADE V	18	6	0.67	0.006

Table 1. Genetic diversity parameters of each clade based on the COI gene. H: haplotype diversity. π: nucleotide diversity.

Genetic distances within clades based on COI genes (Table 2), ranged from 0% to 5% and showed little variability. In contrast, inter-clade distances were higher, 6.55-15.22%. The clade III showed the greatest variability with other clades, even into the ambiguous gap between intraspecific and interspecific divergence in earthworms proposed by Chang and James (2011), 9-15%.

	CLADE I	CLADE II	CLADE III	CLADE IV	CLADE V
CLADE I	2.39	8.57	13.8	7.47	8.22
CLADE II		5	15.11	8.8	9.33
CLADE III			3.79	14.77	15.22
CLADE IV				0	6.55
CLADE V					0.58

Table 2. Percentage of uncorrected pairwise genetic distances based on COI retrieved for *E. tetraedra.*

3.3.4 Ecological niche modeling

The two ecological niche models obtained displayed different predictive power, with higher AUC and kappa values (0.71 vs 0.79 and 0.41 vs 0.57 respectively), higher sensitivity and specificity (0.72 vs 0.79 and 0.71 vs 0.80 respectively) and lower omission rates (0.28-0.21) for the Clade II model (Supplementary Table 7).

The geographical representation of the predicted suitability values is shown in Figure 4. Highly suitable areas were more widespread for Clade II, covering most of the Mediterranean countries and Britain while being scarce in Scandinavia and other northern countries. For Clade I, highly suitable areas corresponded to countries in the same latitude as Britain and higher, being especially widespread in Scandinavia.

The relative contributions of the predictor variables to each model are shown in Supplementary Table 7. BIO7, BIO6, BIO13, AWC and CLC were the most influential variables for Clade I, while BIO6, PARMA, AWC and CRUST were the most influential variables for Clade II (Supplementary Table 7).

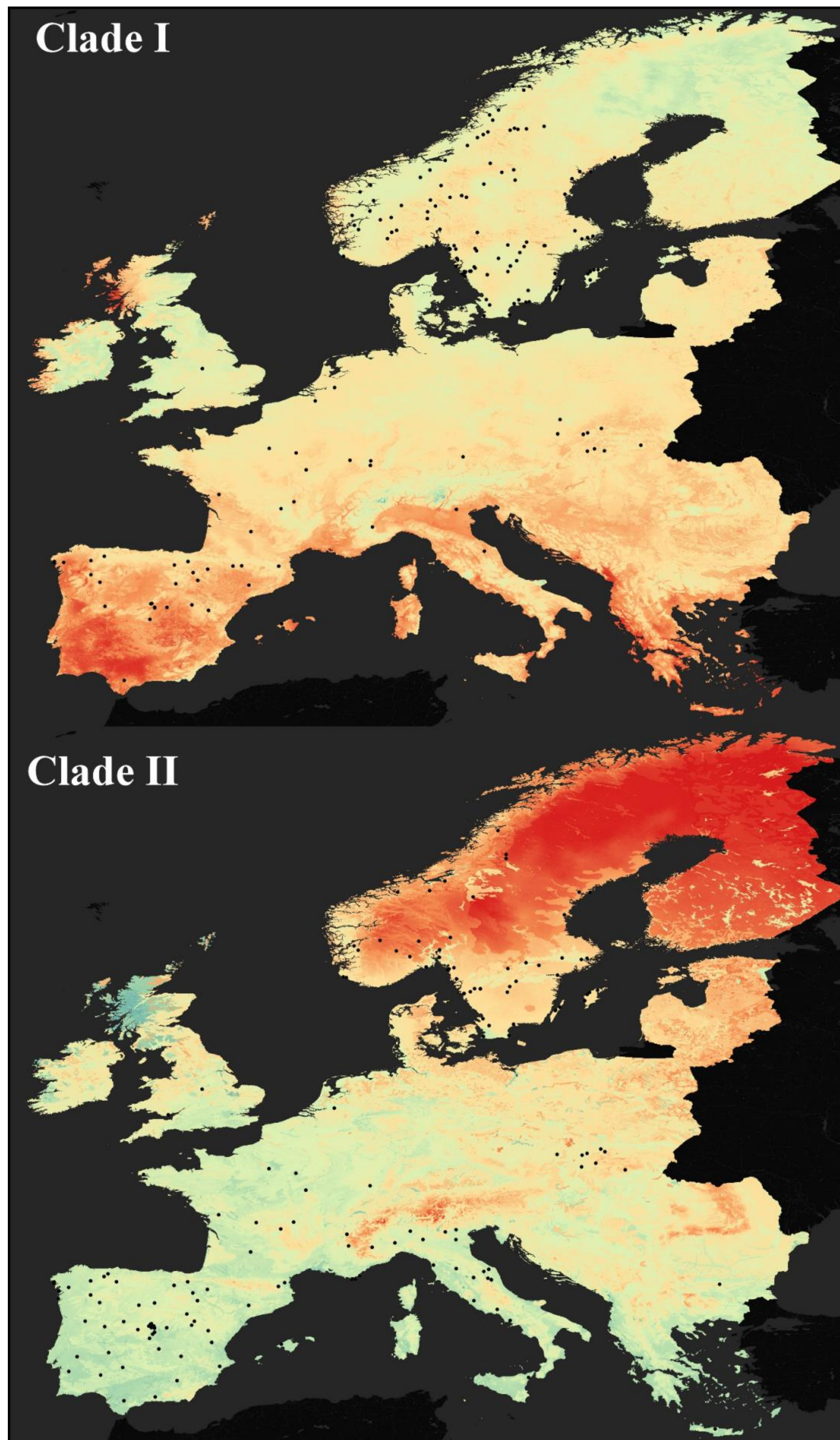


Figure 4. Ecological niche models obtained for the Clade I and Clade II localities (shown as black dots). Red color indicates the lowest estimated suitability and blue-green color indicate the maximum suitability.

3.3.5 Morphological data

Morphological data is shown in Supplementary Table 8. The sexual characters belonging to the male apparatus were the most diverse (e.g. male pore, seminal vesicles and spermathecae) and no presence of sperm was found in any specimen. However, 37 earthworms had empty spermatophores. These individuals are nested in the same lineages or even haplotypes with other specimens that do not have spermatophores. Thus, no genetic basis for the presence of this trait can be assumed.

3.4 Discussion

Several genetic studies on earthworms have shown that some morphologically distinct species turn out to be complexes of genetically well individualized lineages (King *et al*., 2008; Porco *et al*., 2013). It is still controversial whether there are morphological or ecological differences between these lineages and whether they can be considered as cryptic species (Marchán *et al*., 2018; Shekhovtsov *et al*., 2018). Parthenogenetic lineages can arise from sexual species in a variety of ways: bacterial infections such as *Wolbachia* or *Rickettsia* (Huigens *et al*., 2000), spontaneous loss of sex due to mutations in genes related to mating and fertilization of eggs (Carson *et al*., 1982), or in genes involved in sexual forms (Simon *et al*., 2003) and contagious origin, with incomplete reproductive isolation between sexual individuals and pre-existing parthenogenetic lineages (Simon *et al*., 2003) and hybridization origin between individuals of the same or closely related species (Lorenzo-Carballa and Cordero-Rivera, 2009).

The genus *Eiseniella* Michaelsen, 1900 includes two quite distinct groups (Omodeo and Rota, 1991). One includes only the parthenogenetic and cosmopolitan species

Eiseniella tetraedra and the other includes sexual species such as *Eiseniella lacustris* Cernosvitov, 1931 or *Eiseniella ochridana* Cernosvitov, 1931, most of which occur only in Eastern Europe. This biogeographic pattern may indicate the existence of geographic parthenogenesis (Butlin, 2002), in which sexual forms are restricted to areas around the Mediterranean Sea, while parthenogens spread to the rest of the world, presumably due to a greater capacity for colonization and a high potential for adaptation to new environmental conditions.

Although several subspecies had been defined for *E. tetraedra* based on the position of the male pores, no genetic basis for this has been found (Chapter 1). However, eight different lineages nested in two clades have been identified for this species in the Iberian Peninsula (Chapter 2). In the present study, we found eleven lineages in five clades in Europe. As sexual forms of *E. tetraedra* were not ever found, this high genetic diversity in a parthenogenetic species could be due to an ancestral hybridization origin between individuals of the related species of the genus.

According to European Environmental Agency, eleven biogeographical regions in Europe could be defined: Alpine, Anatolian, Atlantic, Artic, Black Sea, Boreal, Continental, Macaronesian, Mediterranean, Pannonian and Steppic. In this study, we sampled *E. tetraedra* in all regions except four (Black Sea, Anatolian, Macaronesian and Steppic).

Clade I was widespread in Europe. However, it was clearly dominant in cold biogeographical regions such as the Alpine, Arctic, Atlantic, Boreal, Continental and Pannonian regions. To a lesser extent, clade I also occurred in the Mediterranean region, only in Iberian Peninsula. This distribution explains the ecological niche model

derived for this clade, which is particularly prevalent in Scandinavian Peninsula. The most influential variables for this clade were also consistent with these results. BIO6 is related to the minimum temperature of the coldest month, which is likely to be low in clade I niches. In contrast, BIO13 corresponds to the precipitation of the wettest month, which is high, as expected. According to Chapter 2, temperature and precipitation were the most important factors affecting the distribution of *E. tetraedra* in Iberian Peninsula. It seems that these variables remain important in Europe as well. Nested in clade I, four linages were found. The most numerous and widespread was clade B. It clearly dominated in the Scandinavian Peninsula and Pannonian regions, confirming the best adaptation to these zones. In contrast, lineages C, D and G were geographically restricted. Lineage C occurred only in Atlantic regions, mainly in the Iberian Peninsula. The presence of this lineage in United Kingdom, Norway and France may be due to human transport, which has been previously reported in this species (Gates, 1977; Javidkar *et al*., 2020). Lineage D occurred only at Iberian Peninsula, and mainly in the Atlantic region. We found only a single occurrence in the Mediterranean region, which could also be attributed to human transport. Finally, lineage G occurred in the Atlantic and in one Mediterranean region, although it was scattered in the southern half of Europe (United Kingdom, Netherlands, France and Balearic Islands).

Like clade I, clade II was widespread in Europe and included four lineages. It was predominant in Mediterranean regions, but also occurred in Alpine, Atlantic, Boreal, Continental and Pannonian regions. This dominance explains its distribution in potential niches, which includes all Mediterranean regions and is scarce in the Scandinavian Peninsula and northern countries. Considering its potential distribution, it seems reasonable that the most influential variable for this clade was the minimum

temperature of the coldest month. In Chapter 2 pH was found one of the most important factors affecting the distribution of lineages in this species. In agreement with this, lithology also appeared to be an influential variable for the clade II. Lineages A, E, F and H were nested in the clade II. All lineages occurred predominantly in the Mediterranean region in Europe. However, lineage A also occurred in Atlantic, Pannonian and Continental regions; lineage F in Pannonian, Continental (only in France) and to a lesser extent in Boreal and Atlantic regions; lineage E in Atlantic, Continental and Boreal regions. Finally, lineage H, which occurs only in one place on Iberian Peninsula, belonged to the Mediterranean region.

The parthenogenetic and cosmopolitan earthworms *Aporrectodea trapezoides* Dugés, 1828 and *Aporrectodea rosea* Dugés, 1828 appeared to be divided into two distinct clades: one present in Eurosiberian region and the other in the Mediterranean region (Fernández *et al*., 2012; Fernández *et al*., 2015). Although clades I and II of *E. tetraedra* were not restricted to one region; the trend seems similar to those species. While paleographic events seem to be of great importance for the earthworm's present-day distribution due to its low vagility (Fernández *et al*., 2013; Novo *et al*., 2011), the ability of *E. tetraedra* to disperse by hydrochory, possibly also by zoochory (Terhivuo and Saura, 2006) and by anthropochory (Gates, 1977; Javidkar *et al*., 2020) may contribute to the underlying processes being hidden in a confusing phylogeography.

The predominant presence of the clade III (lineage I) in Eastern Europe, its position on the phylogenetic tree, and its high genetic distance from the other clades of *E. tetraedra* indicate the ancestry of the clade and the possible origin of *E. tetraedra* in Eastern Europe, where most of the sexually related species of the genus were found.

This hypothesis needs to be confirmed by future studies such as ancestral territory reconstruction (Fernández *et al*., 2016). Although all genetic distances between clades were high, only the distances between the clade III and the rest were in the ambiguous gap between intraspecific and interspecific divergence in earthworms proposed by Chang and James (2011), namely 9-15%. Blakemore (1999) described the species delimitation of parthenogenetic earthworms as a "systematic nightmare". The biological term "species" is not applicable to parthenogenetic earthworms because of the reproductive isolation of each individual. However, the authors tried some species delimitation rules (as GMYC) with no conclusive results.

Clades IV and V included a minor number of individuals of *E. tetraedra*, mostly restricted to the Scandinavian Peninsula, although they had punctual presences in France and Italy (only one specimen in each country) probably due to human transport.

The model of glacial refugia as core areas for the survival of thermophilic and/or temperate animal and plant species during unfavourable Pleistocene environmental conditions and as sources of postglacial recolonization processes is widely accepted in biogeography (Hewitt, 2000; Willis and Whittaker, 2000). It is generally believed that the main hypothesis of the recolonization of Europe after the LGM for earthworms is the "*tabula rasa*" hypothesis. It states that earthworms became extinct in northern latitudes and recolonized these areas after the ice melted. However, there are examples in earthworms, such as *Dendrobaena octaedra*, that did not follow this pattern (Hansen *et al*., 2006). The nunatak hypothesis suggests that there were some ice-free refuges in northern Europe, such as mountains rising above the ice sheet or

ice-free refuges on the coast, where the biota could survive. According to our results, *E. tetraedra* could follow both patterns. Although the high haplotype diversity of Italian Peninsula, 0.96, could be due to its role as a refuge during the LGM, the number of individuals from Italy was lower than on other sampled peninsulas. In Chapter 2, we found lower haplotype diversity than in Chapter 1 in the Iberian Peninsula, which expands the sampling locations. Therefore, we think that this high haplotype diversity could be explained by the smaller number of samples. The presence of restricted lineages at Iberian Peninsula and even clades at Scandinavian Peninsula and Eastern Europe suggest that the three areas served as refugia during the LGM, bringing *E. tetraedra* to the rest of Europe after the ice melted. The Iberian Peninsula was one of the most important glacial refugia of the Pleistocene in Europe (Hewitt, 1999; Hewitt, 2001) and served as a species repository for northern countries (Beebee and Rowe, 2000; Vernesi *et al.*, 2002). Eastern Europe was also an important refuge during the LGM (Sommer and Zachos, 2009) and biota, even earthworms, survived in some ice-free refugia in northern Europe (Hansen *et al.*, 2006).

3.5 Conclusions

Eiseniella tetraedra has been found to have a high genetic diversity in Europe. This diversity was classified into eleven lineages nested in five clades. Clades I and II were widely distributed in Europe, while the others had a limited distribution. Clade I was more represented in cold biogeographical regions such as the continent, the Atlantic or even the Arctic, while clade II was prevalent in Mediterranean regions. Potential niches were also consistent with distribution trends. This is consistent with the phylogeographic patterns of other cosmopolitan and parthenogenetic earthworms.

The clade III is largely restricted to Eastern Europe and appears to be the original clade. Clades IV and V were mostly present in Scandinavian Peninsula. The presence of restricted clades in the Iberian and Scandinavian Peninsula and Eastern Europe, suggests that the three acted as refugia during the LGM. Thus, both the "*tabula rasa*" and nunatak hypotheses could apply to *E. tetraedra* in Europe.

4. Chapter 4: Back to the past: Revisiting *Eiseniella neapolitana* and *Norealidys andaluciana* (Annelida, Lumbricidae) corroborates Bouché's proposal of Norealidys.

Abstract

Eiseniella neapolitana is a semiaquatic and diploid earthworm that for many years was related to the cosmopolitan species *Eiseniella tetraedra* and even considered a subspecies of it. *Norealidys andaluciana* was described in Spain and is usually synonymized with *E. neapolitana*. We collected 69 specimens from Italy, Spain and Cyprus and studied five molecular markers (COI, 16S, 28S, 12S and ND1) and their morphology to solve this taxonomic problem. Genetic studies and differences in the number of segments validate the genus *Norealidys* and reject the widely accepted synonymy. Moreover, this study shows that *N. andaluciana* and *E. neapolitana* were closely related and therefore we renamed the latter *Norealidys neapolitana*. Similar morphological appearance despite clear genetic differences of the three species should be explained by convergence to the aquatic habitat. Despite the expected low haplotype diversity based on the 28S gene, we found a surprisingly high variability in the *N. andaluciana* population in Spain. However, its stable predicted secondary structure and its high content of G+C reject the presence of a pseudogenes in our dataset.

4.1 Introduction

The genus *Eiseniella* Michaelsen, 1900 includes two distinct groups of species, all of which belong to the riparian earthworms. The first group includes *Eiseniella tetraedra* (Savigny, 1826), a parthenogenetic and cosmopolitan earthworm with morphological variability, but in general: quadrangular cross-section, less than 95 segments, male pore usually in XIII, tumid porophores and unique position of the female pores below the line of *setae a* (Omodeo and Rota, 1991). Perel (1967) and Zicsi (1972) revised the other group. It includes about ten species and they are also characterized by a quadrangular cross-section, the number of segments ranging from 95 to 160, male pores without porophores and female pores open dorsally to *setae b* or behind (Omodeo and Rota, 1991). The distribution of the individual species of this second group is geographically restricted.

Eiseniella neapolitana (Örley, 1885) belongs to the last group. It is a semiaquatic, circum Mediterranean and diploid earthworm with biparental reproduction (Omodeo, 1952). It was long considered a subspecies of *E. tetraedra* (Stephenson, 1924; Bodenheimer, 1937; Cernosvitov, 1938, Cernosvitov, 1940; Pavlicek *et al*., 2003). Finally, Csuzdi and Pavlicek (2005) considered it as a valid species.

Qiu and Bouché (1998) described *Reynoldsia andaluciana* on the basis of a few specimens from the locality Salobreña (Granada, Spain). This new genus, *Reynoldsia*, is characterized by the absence of calciferous glands (glands of Morren), rudimentary or absent typhlosole and practically nonexistent nephridial vesicles. Blakemore (2008) did

not accept *Reynoldsia* due to homonymy with a fly and named it *Norealidys andaluciana*. It is considered a synonym of *E. neapolitana* in DriloBase.

In recent years, molecular tools are helping with these taxonomic problems (e.g., Huang *et al*., 2007; Marchán *et al*., 2018; de Sosa *et al*., 2019; Jiménez-Pinadero *et al*., 2021). Thus, a solution for the systematics of this group of species could arise from the study of molecular markers.

In the large-scale molecular phylogenetic analysis of Lumbricidae proposed by Domínguez *et al*. (2015) they found that the sole member of the genus in the study, *E. tetraedra*, appeared closely related to Iberian representatives of the genus *Eiseniona* Omodeo, 1956 (included by other authors in *Iberoscolex* Qiu and Bouché, 1998). The provision of new specimens and species of the *Eiseniella* and related genus to this analysis could prove or reject this hypothesis.

The aim of the present study is: i) to verify whether *Eiseniella neapolitana* is a valid species, distinct from *Eiseniella tetraedra*; ii) to investigate the genetic diversity of *E. neapolitana* for the first time; iii) to prove whether *Reynoldsia andaluciana* is really a synonym of *E. neapolitana*; iv) to place the genus *Eiseniella* in the phylogenetic tree of lumbricids.

4.2 Material and methods

4.2.1 Earthworm sampling and morphological studies

We collected 56 specimens of *Eiseniella neapolitana* from nine localities in Italy (including the type locality), twelve specimens of *Norealidys andaluciana* from Salobreña (Granada, Spain) (type locality of the species), and one undetermined

specimen from Cyprus (Figure 1, Supplementary Table 1). All individuals were collected by manual sorting, washed in distilled water, fixed in 96% ethanol and stored at -20°C in the earthworm collection of the Department of Biodiversity, Ecology and Evolution, Complutense University of Madrid (UCM-LT). All specimens, except the Cypriot, were examined morphologically with emphasis on: length, weight after fixation, number of segments, position of the clitellum and tubercula pubertatis, position of male pores, number and position of seminal vesicles and spermathecae, glands of Morren, thyphlosole and nephridial vesicles. A portion of the posterior body section was collected and carefully cleaned under a stereomicroscope to remove gut and soil particles. Samples were then stored in ethanol and preserved at -20°C until DNA extraction.

Figure 1. Localities sampled. GPS coordinates and number of specimens collected can be found in Supplementary Table 1.

4.2.2 DNA extraction, amplification and sequencing

Total genomic DNA was extracted using the Speedtools Tissue DNA Kit (Biotools). Four mitochondrial markers, a fragment of cytochrome *c* oxidase subunit I (COI), 16S + tRNAs Leu, Ala and Ser (16S), 12S rRNA (12S) and NADH dehydrogenase (ND1), and one nuclear marker (a fragment of 28S rRNA) were amplified. For COI (632 bp), primer sequences and polymerase chain reactions (PCR) followed Pop *et al*. (2003). For 16S-tRNAs (774 bp) and 28S (805 bp) primer sequences and PCR conditions followed Fernández *et al*. (2015). For 12S (974 bp) and ND1 (947 bp), primer sequences and PCR conditions followed Pérez-Losada *et al*. (2015). Due to the unusual high variability found for the 28S gene (see Results), four primer pairs were designed to verify the sequences and exclude a possible pseudogene. However, none of the attempts worked.

All PCRs were specific and resolved via 1% agarose gel electrophoresis; they were visualised with GelRed stain (Biotium). All products were purified using ExoSAP-IT reagent (ThermoFisher Scientific). PCR products were sequenced by Macrogen Spain Inc.

4.2.3 Data analyses

Sequences were aligned in MAFFT v.7 (Katoh and Standley, 2003) using default settings and concatenated with BioEdit v7.0.9 (Hall, 1999). Two phylogenetic trees were constructed, one with sequences of the studied populations only (COI-16S+tRNAs+28S) and another with these populations and other lumbricid species (COI-16S+tRNAs+28S+12S+ND1) to place them in a wider phylogenetic context.

Phylogenetic trees based on the concatenated sequences of COI-16S+tRNAs+28S (2,211 bp) and COI-16S+tRNAs+28S+12S+ND1 (3,296 bp) were constructed by Bayesian Inference (BI) with MrBayes v3.1.2 (Ronquist and Huelsenbeck, 2003) and Maximum Likelihood (ML) using RaxML v7.03 software (Stamatakis, 2006) both implemented in CIPRES Science Gateway v.3.3 (Miller *et al*., 2010). The best-fitting substitution model selected by jModelTest2 (Darriba *et al*., 2012) for all markers in the short dataset was GTR+Γ+I, while for the longer dataset HKY+Γ+I was chosen for 16S and GTR+Γ was chosen for 12S. Maximum Likelihood analysis with rapid bootstrapping was performed with 1,000 replicates. Parameters in MrBayes were set to ten million generations and 10,000 trees were sampled every 1,000th generation, initiating the analysis with a random tree. Two independent analyses were performed and 20% of the trees were discarded as burn-in. A sequence of *E. tetraedra* was retrieved from GenBank and used as outgroup (Supplementary Table 2). Phylogenetic trees obtained were visualised in FigTree v1.3.1 (Morariu *et al*., 2008).

We calculated haplotype and nucleotide diversity for populations and genes using DNAsp v.6 (Rozas *et al*., 2017). We also estimated uncorrected pairwise distances for COI and 16S-tRNAs within and between populations, with *E. tetraedra* and two more lumbricids: *Dendrobaena byblica* Rosa, 1893 and *Iberoscolex oliveirae* (Rosa, 1894) (Supplementary Table 2). Haplotype networks based on each gene were constructed in PopART 1.7 (Leigh and Bryant, 2015) using the TCS inference method.

The RNAfold web server (Gruber *et al*., 2008) was used in order to estimate the secondary structure of the 28S rRNA for Spanish and Italian specimens and six other different species: *Aporrectodea trapezoides* (Dugès, 1828), *Allolobophora dubiosa*

(Örley, 1881), *Dendrobaena octaedra* (Savigny, 1826), *Dendrobaena byblica*, *I. oliveirae* and *E. tetraedra* (Supplementary Table 2). Both MFE (Minimum Free Energy) and Centroid secondary structures were estimated. We calculated the G+C content of Spanish specimens by DNAsp v.6 (Rozas *et al*., 2017).

Statistical analyses of morphological data were conducted in Statgraphics Centurion 19 (StatPoint Technologies Inc., USA). We used length, dry weight (after letting it drip on filter paper for 30 seconds) and number of segments of mature specimens to investigate differences between Spanish and Italian populations and *E. tetraedra* through non-parametric analyses (Kruskal-Wallis) followed by Fisher LSD post hoc test.

4.3 Results

4.3.1 Phylogenetic analyses

Bayesian and Maximum Likelihood approaches reveal trees with congruent topology for the studied populations (Figure 2). Sequences are nested in three well supported clades corresponding to each country studied.

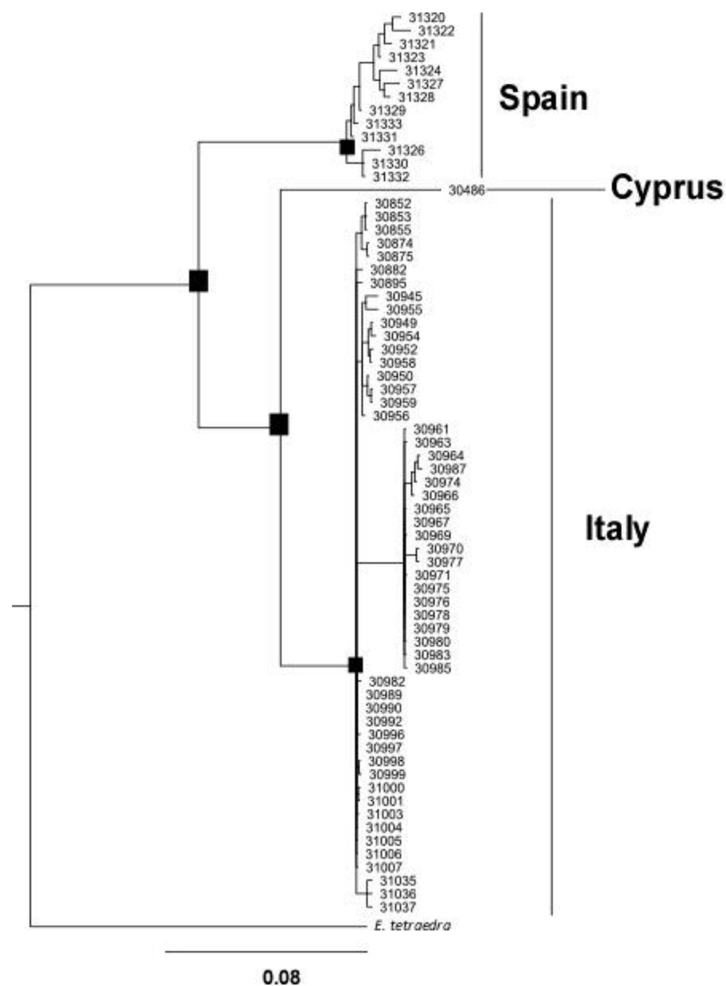


Figure 2. Bayesian inference (BI) of the phylogenetic tree based on concatenated sequences of COI, 16S and 28S of the studied populations. Posterior probability/bootstrap support values (of Maximum Likelihood Analysis, ML) are shown as black squares when higher than 0.9/0.7 (BI/ML). The scale bar represents 0.08 substitutions per position. Sample numbers correspond to references in the UCM-LT collection and can be found at Supplementary Table 1.

Bayesian inference for the studied lumbricid species shows a politomy for the genus *Iberoscolex*, *Eiseniella* and the studied populations (Figure 3). The Spanish population appears to be more related to the Italian and Cypriot populations than to *E. tetraedra*, supporting the idea of a different genus, *Norealidys*, other than *Eiseniella*, for the populations studied.

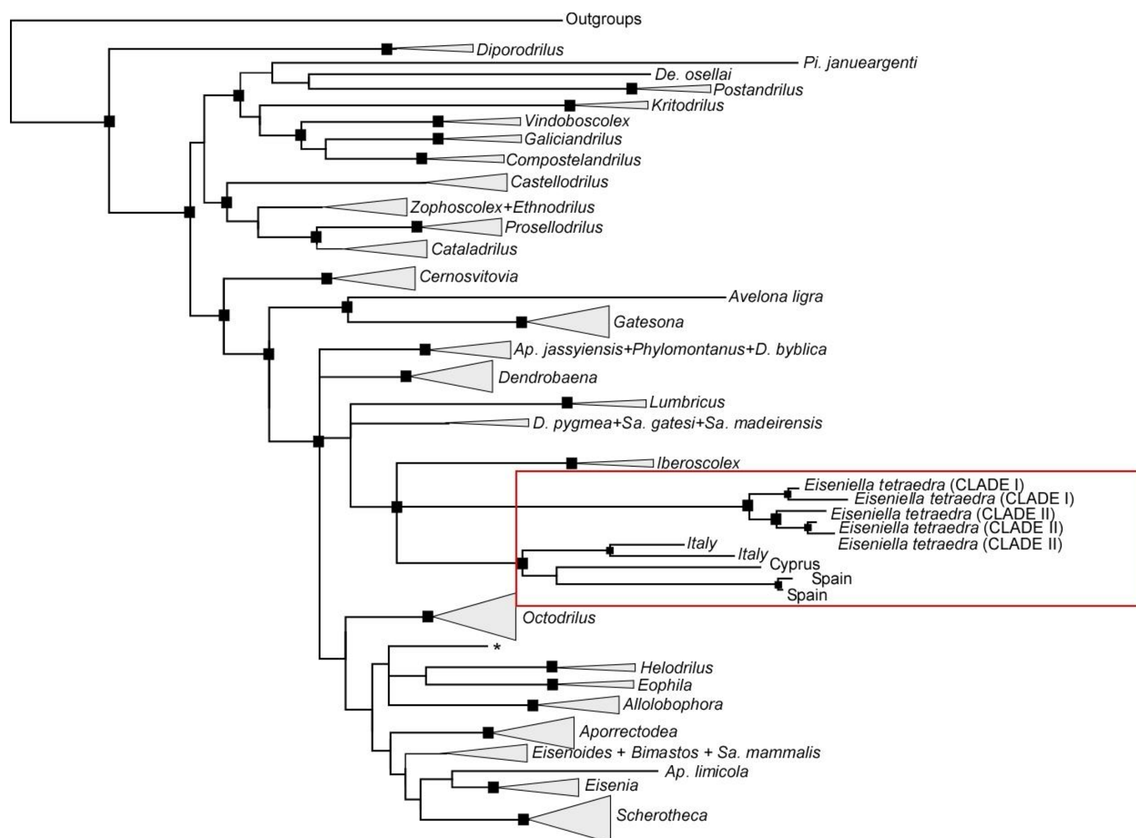


Figure 3. Bayesian inference (BI) of the phylogenetic tree of Lumbricids based on concatenated sequences of COI, 16S, 28S, 12S and ND1 (long dataset). Posterior probability/bootstrap support values (of Maximum Likelihood Analysis, ML) are shown as black squares when higher than 0.9/0.7 (BI/ML). Grey triangles represent two or more species of a genus-level clade collapsed to facilitate interpretation of the figure. *: several singleton species in an unresolved polytomy -*Aporrectodea rosea*, *Aporrectodea georgii*, *Panoniona leoni*, *Allolobophora bartolii*, *Helodrilus patriarchalis* and *Proctodrilus antipai*. Complete names of species are shown in Supplementary Table 2.

4.3.2 Genetic diversity and genetic divergence

A total of 17 haplotypes have been identified among 42 sequences for the COI gene, 22 haplotypes within 60 sequences for 16S-tRNAs, and 21 haplotypes within 68

sequences for 28S. Haplotype diversity (H) and nucleotide diversity (π) based on COI including all the specimens within the study are 0.91 and 0.11 respectively. H and π based on 16S-tRNAs are 0.90 and 0.04. Genetic diversity parameters based on 28S are 0.58 (H) and 0.01 (π). Values of haplotype and nucleotide diversity for each population are shown in Table 1. A remarkably high haplotype diversity based on 28S has been found for the Spanish population.

Populations	**COI (H/π)**	**16S (H/π)**	**28S (H/π)**
Spain	0.69/0.005	0.67/0.004	0.98/0.01
Cyprus	NA	NA	NA
Italy 1	-	0	0
Italy 2	-	0.66/0.004	0.66/0.0008
Italy 3	NA	NA	NA
Italy 4	0.90/0.01	-	0.2/0.0002
Italy 5	0.92/0.01	0.52/0.005	0
Italy 6	0.53/0.0008	0	0.27/0.0003
Italy 7	-	-	0

Table 1. Haplotype and nucleotide diversity values of each population studied. NA indicates populations with only one specimen and sequence. Hyphens indicate no sequences for a gene.

Genetic divergence among populations based on COI ranges from 1.57 to 17.81% (Table 2). Values of more than 9-15% are the ambiguous gap between intraspecific and interspecific divergence in earthworms suggested by Chang and James (2011). Thus, *E. neapolitana* from Italy seems to represent a cryptic species complex, but according to

genetic divergence, the samples from the Spanish population, Cyprus and Italy could be three different species. Values obtained for the populations based on 16S are in the range to 0.38 to 10.22% and also show a high variability (Table 2). Genetic distances between the studied populations and *E. tetraedra*, *D. byblica*, and *I. oliveirae* are shown in Table 2. COI gene based distances among *E. tetraedra* and the studied populations are higher than the threshold of 9-15% proposed by Chang and James (2011) for interspecific distances, confirming that they are different species. Moreover, distances based on 16S are even higher, reaching about 30%.

16S/COI	**Spain**	**Cyprus**	**Italy 1**	**Italy 2**	**Italy 4**	**Italy 5**	**Italy 6**	***E. tetraedra***	***I. oliveirae***	***D. byblica***
Spain	0.57/0.43	16.04	-	-	12.17	17.81	15.14	20.51	18.26	19.44
Cyprus	10.22	0/0	-	-	10.09	16.33	13.13	20.61	17.87	18.67
Italy 1	7.33	7.49	-/0	-	-	-	-	-	-	-
Italy 2	7.24	7.62	0.38	/0.43	-	-	-	-	-	-
Italy 4	7.62	7.5	1.01	0.89	0.49/0.9	9.11	1.57	16.27	13.38	14.14
Italy 5	7.89	8.62	5.09	5.17	5.28	1.52/0.55	11.85	21.38	18.17	19.24
Italy 6	7.46	7.49	0.51	0.47	0.70	5.22	0.83/0	20.34	17.29	16.98
E.tetraedra	29.75	29.47	29.51	29.52	29.36	29.33	29.46	7.10/0.93	19.20	19.50
I. oliveirae	31.07	31.26	30.74	29.58	30.90	30.70	31.00	5.98	0/0	12.65
D. byblica	29.78	29.97	29.45	29.58	29.64	29.28	29.71	5.22	3.22	0/0

Table 2. Percentage of uncorrected pairwise distances based on COI (above diagonal) and 16S (below diagonal) between the studied populations and three different species: *Eiseniella tetraedra*, *Dendrobaena byblica* and *Iberoscolex oliveirae*. Hyphens indicate no sequences for a gene in one or two of the compared populations.

The haplotype network for COI exhibits each population is highly distant from other neighboring populations, with number of mutational steps from 14 (Italy 4 – Italy 6) to

98 (Cyprus – Italy 5) (Figure 4). Samples from each country are separated from the others by a high number of mutational steps. Italy 5 is the most heterogeneous population. Otherwise, the haplotype network based on 28S, shows high similarity for Italian haplotypes, although Spanish haplotypes show remarkably high differences, as evidenced in their haplotype diversity (Figure 5). Also, populations of each country appear separated from the others.

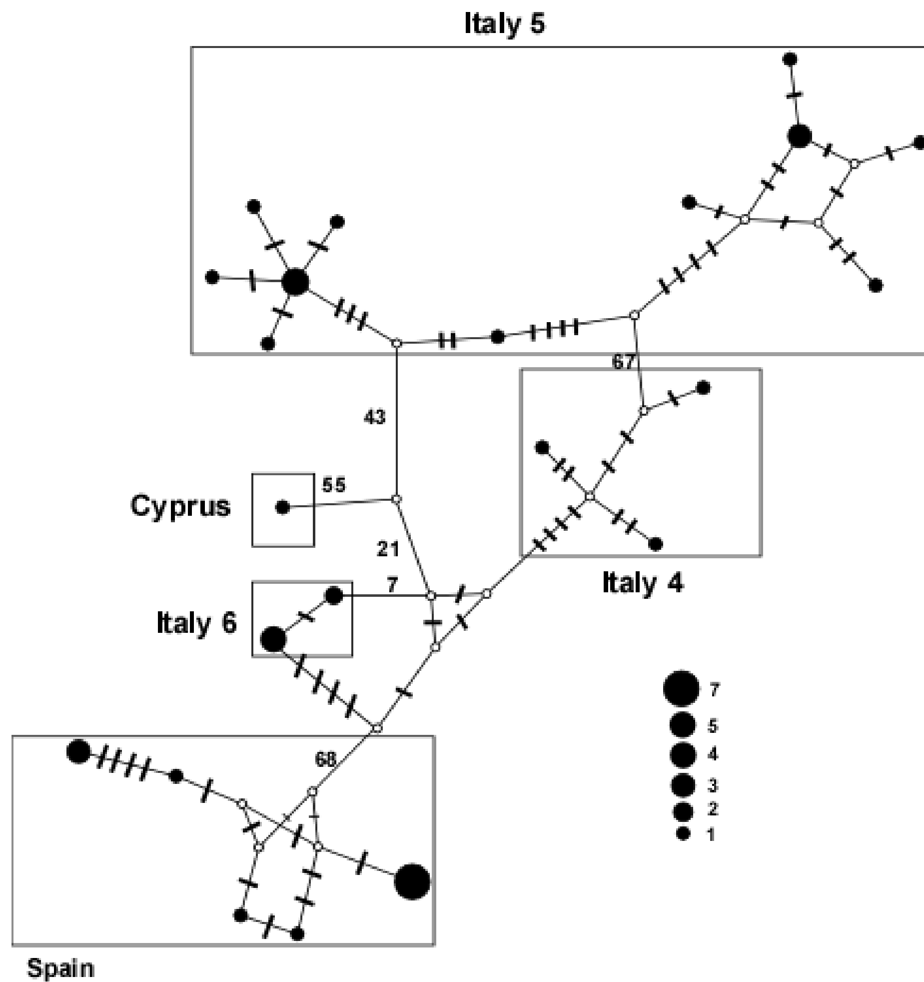


Figure 4. Haplotype network based on COI. Each black circle represents a different haplotype; their size is proportional to the number of samples belonging to that haplotype. Each perpendicular line indicates a mutational step, if there were more than four it is shown with the number of mutational steps. White circles are

hypothetical intermediate haplotypes. Branch length does not contain any information.

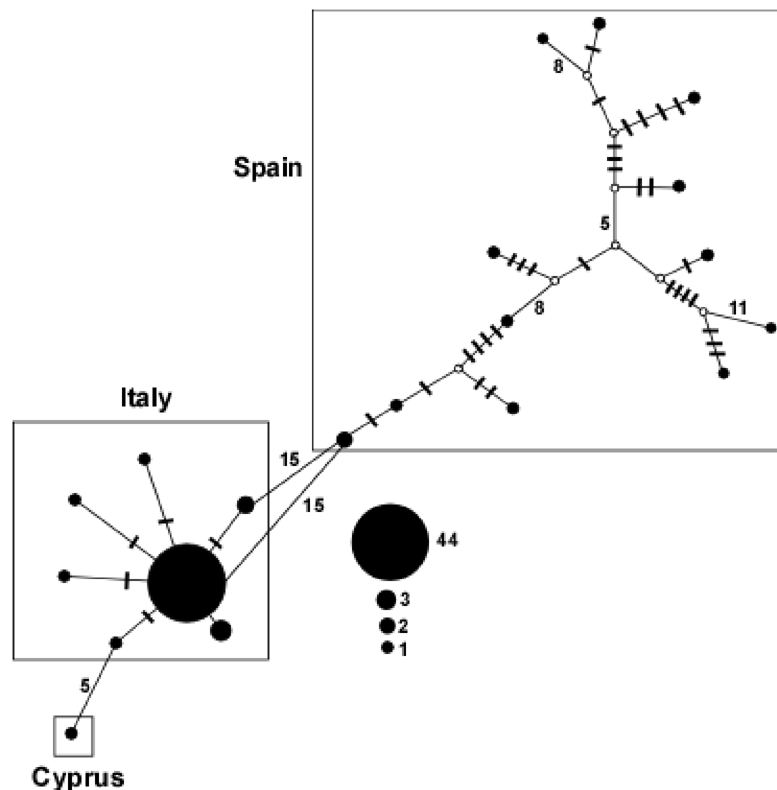


Figure 5. Haplotype network based on 28S. Each black circle represents a different haplotype; their size is proportional to the number of samples belonging to that haplotype. Each perpendicular line indicates a mutational step, if there were more than four it is shown with the number of mutational steps. White circles are hypothetical intermediate haplotypes. Branch length does not contain any information.

To reject the amplification of a pseudogene in the sequences of 28S for the Spanish population, we examined the secondary structure of this gene in the studied populations and other Lumbricidae species (Figure 6). The Italian specimens show the same structure as other species such as *A. trapezoides*, *D. octaedra* and *I. albolineatus*.

However, the Spanish specimens, *A. dubiosa* and *D. byblica* show different forms. The minimum free energy prediction of all sequences was low, ranged from -397 kcal/mol to -420 kcal/mol. The G+C content for the Spanish population is 66.50%.

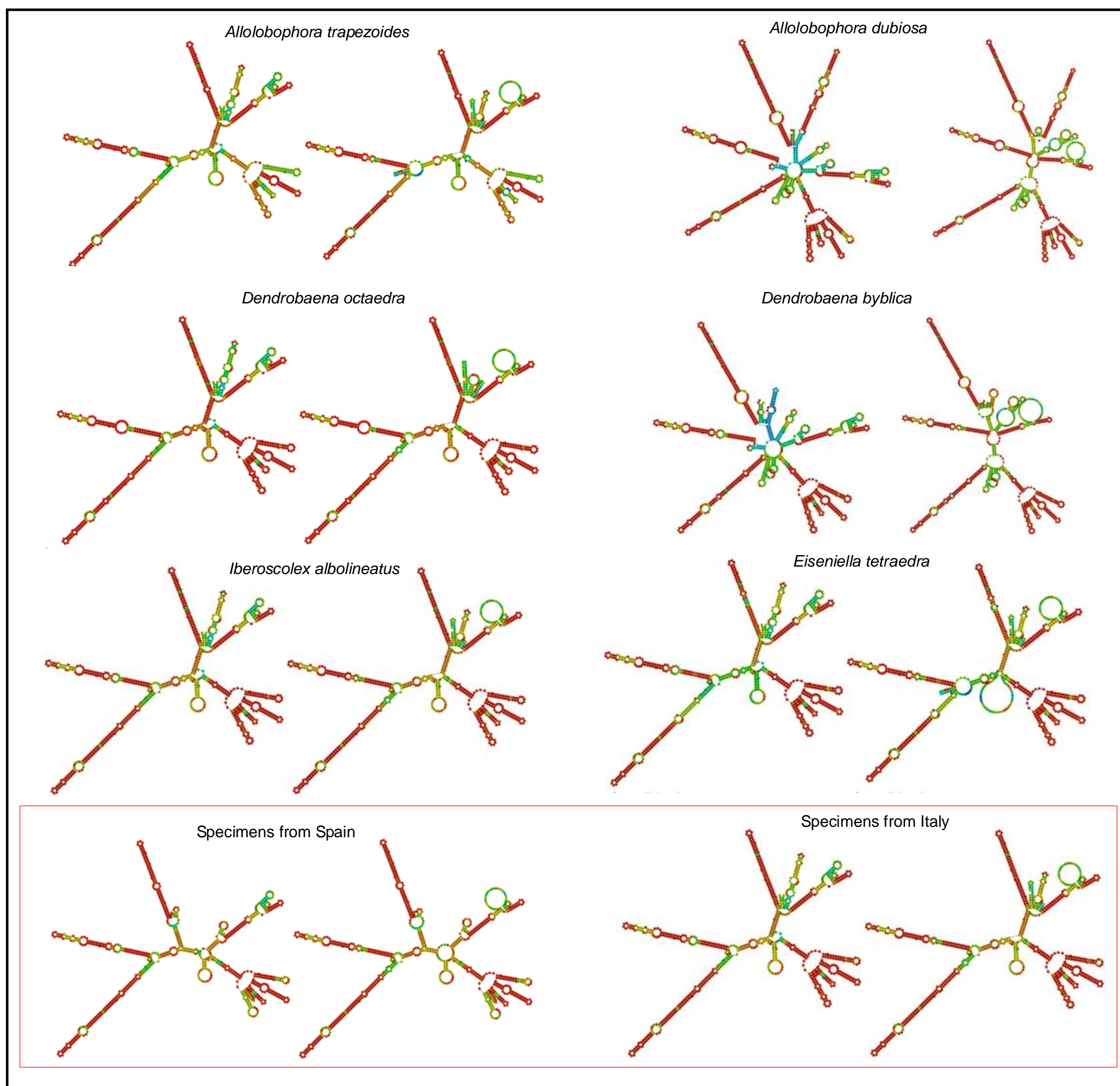


Figure 6. Predicted secondary structures of 28S gene for Spanish and Italian specimens from the studied populations and of six different species: *A. trapezoides, A. dubiosa, D. octaedra*, *D. byblica*, *I. albolineatus* and *E. tetraedra*.

4.3.3 Morphological studies

Morphological characters of the specimens are shown in Table 3. Only mature individuals are used for the morphological analyses. The Cypriot specimen is not

included in the morphological analysis because we only have the posterior section of the body. Spermathecae and male funnels are in 10 and 11 and are iridescent in all specimens. The clitellum position of all individuals from the Spanish population is 21-25, while the Italian populations show some variability in this trait. There is also some variability in the position of the tubercula pubertatis in the Italian populations, but in all Spanish specimens they are 22-23-24. No variability has been found for the male pore, which is always located at segment 15.

ID UCM-LT	Locality	Length (mm)	Dry weight (mg)	Nº of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Male funnels
31320	Spain	40	180	127	21-25	22-23-24	15	9-10-11-12	iridiscent	iridiscent
31321	Spain	37	150	118	21-25	22-23-24	15	9-10-11-12	iridiscent	iridiscent
31324	Spain	25	80	101	21-25	22-23-24	15	9-10-11-12	iridiscent	iridiscent
31329	Spain	30	110	102	21-25	22-23-24	15	9-10-11-12	iridiscent	iridiscent
31332	Spain	31	100	103	21-25	22-23-24	15	9-10-11-12	iridiscent	iridiscent
31333	Spain	29	80	100	21-25	22-23-24	15	9-10-11-12	iridiscent	iridiscent
30853	Italy 1	52	226.8	168	21-25	21-22-23-24	15	9-10-11-12	iridiscent	iridiscent
30964	Italy 5	21	45.2	97	21-26	22-23-24	15	9-10-11-12	iridiscent	iridiscent
30966	Italy 5	30	45.8	132	21-26	22-23-24-25	15	9-10-11-12	iridiscent	iridiscent
30976	Italy 5	35	75.7	135	21-27	22-23-24-1N25	15	9-10-11-12	iridiscent	iridiscent
30997	Italy 6	33	114	114	21-25	22-23-24-25	15	9-10-11-12	iridiscent	iridiscent

Table 3. Morphological traits for mature specimens of *N. neapolitana* and *N. andaluciana*.

The Italian specimens show glands of Morren in 10-14 with diverticula in 10, lamellae in 11-13 and little distinct in 14, while none of these structures are observed in the individuals from the Spanish population (except for a very doubtful structure in 11, which should be confirmed histologically) (Figure 7A). In the nephridia, the glandular part is broadened in both the Italian and Spanish specimens, and a well-differentiated nephridial vesicle do not appear, but the nephridial tube gradually tapers to flow through the nephridial pore (Figure 7C). The typhlosole is present in all specimens. It is

small and simple at first and then a slightly pronounced central furrow appears, giving it a certain bilobed appearance (Figure 7E).

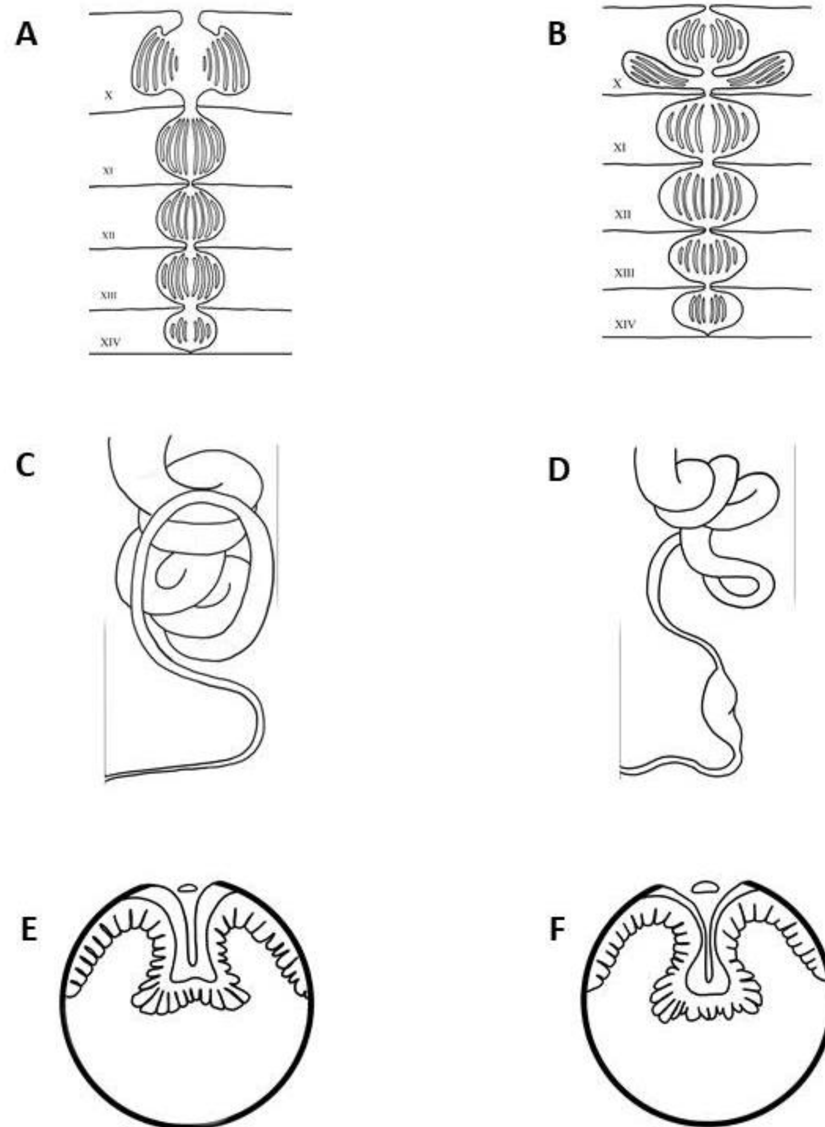


Figure 7. Illustrations of internal anatomy of specimens studied. A: Glands of Morren of Italian specimens. B: Glands of Morren of *E. tetraedra*. C: Nephridia of the Italian and Spanish specimens. D: Nephridia of *E. tetraedra*. E: Thyphlosole of Italian and Spanish specimens. E: Thyphlosole of *E. tetraedra.*

No statistically significant differences ($p>0.05$) are found between the Spanish and Italian populations and *E. tetraedra* for length and weight. In contrast, number of segments show statistically significant differences ($p= 0.002$) between two homogenous groups: Italian and Spanish populations *vs*. *E. tetraedra* (Figure 8).

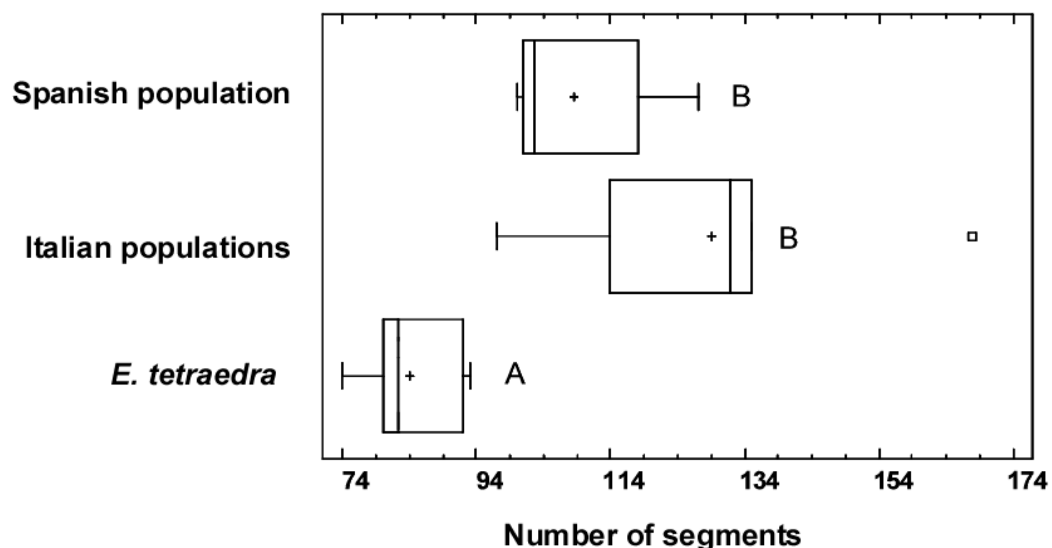


Figure 8. Box-plot diagram of number of segments in the studied populations and *E. tetraedra*. Different letters indicate different groups in multiple range tests.

4.4 Discussion

Eiseniella neapolitana was described by Örley (1885) as *Allurus neapolitanus* and considered by some authors (Stephenson, 1924; Bodenheimer, 1937; Cernosvitov, 1938, Cernosvitov, 1940; Pavlicek *et al*., 2003) as a subspecies of *Eiseniella tetraedra*, while Csuzdi and Pavlicek (2005) consider it a valid species. The holotype of *E. neapolitana* was found by Örley (1885) in Sebeto River (Napoli, Italy). Our specimens from Italy 1 correspond to this river. However, all specimens collected in this study

correspond to the morphological characters used by Örley (1885) to describe this species.

Specimens from Italy and Spain were nested in the same clade in our phylogenetic tree, but with deep and well-separated long branches, and their genetic distances, ranging from 16.27% to 21.38%, were as high among themselves as with the other species studied *(E. tetraedra*, *D. byblica* and *E. oliveirae*). Thus, it can be assumed that they are different species. The haplotype networks also support this hypothesis. Moreover, the depth of the branches in the phylogenetic tree corresponds to a genus separate from *Eiseniella*. Thus, both species could be included in the genus *Reynoldsia* established by Qiu and Bouché (1998). This name enters into homonymy with the Diptera *Reynoldsia* Malloch (1934), so Blakemore (2008) proposed the name *Norealidys*, with the only species *N. andaluciana*.

In this study, we propose that *N. andaluciana* is a valid species and that *E. neapolitana* is related to it. Furthermore, *E. neapolitana* should not be included in the genus *Eiseniella* and should be renamed *Norealidys neapolitana*. Moreover, we found a high genetic divergence among the studied populations of *N. neapolitana*. According to the ambiguous gap between intraspecific and interspecific divergence in earthworms (9%-15%) proposed by Chang and James (2011), the Italian populations of *N. neapolitana* seemed to form a cryptic complex, with one side belonging to Italy 4 and Italy 6 and the other to Italy 5.

The Cypriot specimen included in this work may represent another species within the genus *Norealidys*. However, since it is a single specimen and we only had the posterior

section of the body, we preferred not to describe or name this species until we have examined other complete specimens from this population.

The genera *Eiseniella* and *Norealidys* appear in the phylogenetic tree in a politomy with *Iberoscolex* (= *Eiseniona*), a genus whose relationships with the other lumbricids and internal phylogeny will be the subject of future studies. Although the general appearance of *Norealidys* and *Eiseniella* specimens is very similar, such as the quadrangular posterior section with dorsal groove, the anterior position of the clitellum or even individuals with male pore in segment 13, there are differences in the position of clitellum and *tubercula pubertatis*. We also found significant differences in the number of segments between *N. neapolitana* and *N. andaluciana* with *E. tetraedra*. Their similar external appearance (small size, anterior clitellum and posterior dorsal sulcus) could be due to convergence to aquatic habitat. Earthworms from the semiaquatic family Almidae present quadrangular body at the posterior end and round body at the anterior and an anterior clitellum (Chanabun *et al*., 2020). Other semiaquatic species, such as those of the semiaquatic family Criodrilidae, also share some of these morphological characteristics (Blakemore, 2008). However, the most closely related genus in the phylogenetic tree, *Iberoscolex*, does not present these characters. Therefore, these traits do not appear to have a phylogenetic signal, being shared by species from different families.

N. neapolitana showed well-developed glands of Morren, which would place it close to *E. tetraedra*. However, molecular markers placed it with *N. andaluciana*, which did not show these developed structures. Future studies could resolve these inconsistencies and show whether structures such as the glands of Morren could correspond to

analogous organs and what evolutionary pressures intervene in their development, similar to those indicated by Marchán *et al*. (2016) for the typhlosole of the Hormogastridae.

As a result of concerted evolution, all copies of rDNA families are generally rapidly homogenized within individuals and species, but interspecific divergence can be high (Hillis and Dixon 1991). Intraspecific divergence has been found to be high in animals such as corals (Márquez *et al*., 2003), grasshoppers (Keller *et al*., 2006), and nematodes (Hugall *et al*., 1999; Pereira and Baldwin, 2016). At protein-coding loci, pseudogenes can be detected by the presence of stop codons, frame shifts, the absence of substitution bias at third position, or changes in otherwise invariant amino acid residues (e.g., Arctander 1995). These indicators cannot be used for nonprotein-coding genes (such as 28S rRNA); however, selection acts to preserve the secondary structure of functional RNA molecules. Current methods for predicting RNA secondary structure are mainly based on the minimum free energy algorithm, which finds the optimal folding state of RNA *in vivo* using an iterative method to satisfy the minimum energy or other constraints (Zhang *et al*., 2019). In general, secondary structure stability (low free energies) and pattern of nucleotide substitution in other ribosomal genes such as 5.8S or ITS, appeared to be the most powerful approaches to distinguish putative pseudogenes from presumed functional sequences (Razafimandimbison *et al*., 2004). Thus, the presence of a stable predicted secondary structure (Pereira and Baldwin, 2016) and high G+C content (Zheng *et al*., 2008) do not support the presence of pseudogenes. Therefore, the high intraspecific divergence for the 28S rRNA gene in the Spanish populations studied may not be explained by the presence of pseudogenes. The mode of reproduction has also been associated with rRNA heterogeneity. Cross-

fertilization may increase intraspecific variation due to recombination in the cyclic parthenogenetic *Daphnia pulex* Leydig, 1860 (Crease and Lynch, 1991) and in polyploid obligate mitotic parthenogens of the nematode genus *Meloidogyne* (Hugall *et al*., 1999). Although the Spanish population studied appears to have biparental reproduction based on the presence of sperm in the spermathecae, suggesting copulation with another individual, laboratory experiments are required to investigate whether it may also be parthenogenetic. Therefore, its hybrid origin could explain the hypervariability in the 28S rRDNA.

Differences in predicted secondary structure of 28S rRNA between different earthworms' species may be due to hypervariability rather than pseudogenes or lack of function. They also do not appear to have a phylogenetic signal, as species of the same genus have different structures.

4.5 Conclusions

Phylogenetic analysis, genetic divergence based on COI, and significant differences in the number of segments support the earthworm genus *Norealidys*. According to the phylogenetic analysis for lumbricids, *E. neapolitana* is different, but closely related to *N. andaluciana* and renamed *Norealidys neapolitana,* which invalidates the idea of being a subspecies of *E. tetraedra* and its synonymy with *N. andaluciana*. Also, in the studied populations *N. neapolitana* appeared as a cryptic complex. The high variability of the 28S rRNA gene for *N. andaluciana* was not explained by the presence of pseudogenes, due to its stable secondary structure and high G+C content.

5. Chapter 5: How to thrive in unstable environments: gene expression profile of a riparian earthworm under abiotic stress.

Abstract

Nowadays, extreme weather events caused by climate change are becoming more frequent. This leads to the occurrence of extreme habitats to which species must adapt. This challenge becomes crucial for species living in unstable environments, such as the riparian earthworm *Eiseniella tetraedra*. Its cosmopolitan distribution exposes it to various environmental changes, such as freezing in subarctic regions or droughts in Mediterranean areas. Transcriptional changes under cold and desiccation conditions could therefore shed light on the adaptive mechanisms of this species. An experiment was performed for each condition. In the cold experiment, the temperature was lowered to -14 °C ± 2 °C (compared to 8 °C for control samples), and in the desiccation treatment, humidity was lowered from 60% to 15%. Comparisons of gene expression levels between earthworms under freezing conditions and control earthworms revealed a total of 84 differentially expressed genes and comparisons between the desiccation experiment and the control yielded 163 differentially expressed genes. However, no common responses were found between the two treatments. The results suggest that *E. tetraedra* can acclimate to low temperatures due to the upregulation of genes involved in glucose accumulation. However, downregulation of the respiratory chain suggests that this earthworm does not tolerate freezing conditions. Under desiccation conditions, genes involved in cell protection from apoptosis and DNA repair were upregulated. In contrast, lipid metabolism was downregulated, presumably to conserve resources by reducing the rate at which they are consumed.

5.1 Introduction

Recent climate change has impacted a wide range of organisms worldwide, responding at population and species levels with changes in phenology, physiology, and range shifts, as well as changes in community and ecosystem structure and dynamics (Walther *et al*., 2002; Brown *et al*., 2016). Projected climate changes are likely to significantly affect the spatial extent, distribution, and function of wetlands (IPCC 1992, 2001). Changes in precipitation will alter water availability, river flows and affect ecosystem productivity. Populations of aquatic organisms are sensitive to the effects of floods and droughts (Dawson *et al*., 2003). In assessing recent climate change and that projected for the coming decades, the 3rd IPCC report on climate change (Houghton *et al*., 2001) notes that increasing summer drought over most continental mid-latitude areas and associated drought risk was likely in the 20th century and is to continue in the 21st century (Dubrovsky *et al*., 2009). Moreover, the importance of winter conditions is often overlooked, especially in temperate ecology (Kreyling, 2010). A single extreme cold event can offs*et al*l distributional adjustments to the general warming trend (Jalili *et al*., 2010). Despite the average warming and the lower frequency of their occurrence, both the intensity and duration of such extreme cold events may not decrease this century due to changes in atmospheric circulation and internal atmospheric variability that counteract the warming trend due to greenhouse gases (Kodra *et al*., 2011). These extreme weather events have increased in recent years due to climate change. This is creating new extreme habitats for species, which need to develop adaptive mechanisms for their survival.

Earthworms are probably the most important animals of the soil biota in terms of soil formation and maintenance of soil structure and fertility. Although they do not dominate in numbers, their large size makes them one of the major components of soil invertebrate biomass (Edwards, 2004). Their activities are important for maintaining soil fertility in a variety of ways in forest, grassland and agroecosystems (Edwards, 2004). The local species richness and abundance of earthworm communities typically peak at mid-latitudes, and their global distribution is mostly determined by climatic variables and habitat cover rather than soil properties Phillips *et al*. (2020), suggesting that climate and habitat changes can have serious impacts on earthworm communities and the functions they provide.

Soil moisture and temperature are the main factors limiting earthworm survival, growth and reproduction (Lee 1985). Below a critical soil moisture level, there is an adverse osmotic effect (desiccation) on earthworms (Grant, 1955). This critical moisture level depends on the physical properties of the soil and it is species-specific, as desiccation tolerance varies among earthworm species (Grant, 1955; Buckman and Brady, 1969; Lee, 1985). Moderately low soil moisture decreases aerobic metabolism (Diehl and Williams, 1992), fertility (Reinecke and Venter, 1987) and whole organism growth rate (Viljoen and Reinecke, 1989), whereas high soil moisture lowers fertility and growth rate (Reinecke and Venter, 1987). Levels beyond the critical point for desiccation, interfere mainly with the respiratory system (Williams and Diehl, 1992) due to both the difference in oxygen diffusion rate in water compared to air (Cameron, 1986) and the effects of soil moisture on cutaneous oxygen uptake (Lee, 1985).

Temperature also affects earthworm physiology and ecology (Lee, 1985), with upper thermal limits for survival varying among species, ranging from 25 to 33°C (Wolf, 1938; Miles, 1963). When temperatures are low, earthworms employ two main strategies: (1) freeze avoidance either by migration or physiological adaptation, such as the synthesis of cryoprotectants to avoid the formation of ice within the animal; and (2) allowing the formation of extracellular ice. This leads to dehydration (and thus a lower freezing point) of the cells, as the concentration of unfrozen extracellular fluid is increased (since only pure water forms ice), resulting in an osmotic flow of water out of the cells. At normal cooling rates, the loss of water from the cells due to extracellular freezing is sufficient to bring the freezing point of the intracellular fluids into equilibrium with that of the extracellular fluids, thus preventing cell freezing (Mazur, 1963; Holmstrup and Zachariassen, 1996). Glucose loading appears to be essential for freezing tolerance in earthworms, but other factors may also be involved (Berman and Leirikh, 1985; Holmstrup and Petersen, 1997; Holmstrup *et al*., 1999). According to Holmstrup (2014) there seems to be a common response between cold tolerance and desiccation tolerance in soil invertebrates, such as initiating same modifications in membrane composition.

Eiseniella tetraedra (Savigny, 1826) is a parthenogenetic earthworm (Casellato, 1897) that presents high genetic and morphological diversity (Chapter 1; Chapter 2; Chapter 3) and a cosmopolitan distribution (Blakemore, 2006), being found in a variety of climates, such as mediterranean or subarctic, and thus coping with a wide range of climatic conditions. It is considered a semi-aquatic earthworm (Omodeo and Rota, 1991) and inhabits both stable habitats, such as rivers or streams, and unstable waters, which may be frozen or dried up depending on the season. Terhivuo and Saura (1997)

showed in the Aland Archipelago (Baltic Sea) that all individuals of *E. tetraedra* disappeared each year due to habitat freezing, followed by recolonization by different clones of the species. However, some cocoons of *E. tetraedra* are able to survive at -28°C and under desiccation conditions (one week completely dehydrated in Petri dishes, pers. observ), which suggests that some *E. tetraedra* individuals might be able to survive under extreme circumstances. The abundance of *E. tetraedra* populations can reach 1000 ind/m2 (Malevich, 1956) and it occurs mainly in soils with mull humus, under leaf litter or moss in the upper 2 cm of the soil profile (Terhivuo *et al*., 1994). In the Russian habitats studied by Barne and Striganova (2004), adults predominated in the populations, accounting for 70-75% of the total abundance. The proportion of juveniles and subadults ranged from 10 to 16 %. This earthworm can be adapted to laboratory cultures with relatively high survival and reproduction rates (Barne and Striganova, 2004).

RNA-Seq is a powerful high-throughput sequencing method that provides rapid and comprehensive gene expression data (Chen *et al*., 2017; Jin *et al*., 2017). It is an effective way to identify a range of novel protein-coding and non-coding genes/transcripts of organisms of interest (Roberts *et al*., 2011; Pauli *et al*., 2012; Chettoor *et al*., 2014). Additionally, RNA-Seq technology can detect gene transcription, characterise gene expression, and discover novel genes and biomarkers at the aggregate level (Diao *et al*., 2019). Advances in high-throughput sequencing technology along with relevant bioinformatics tools and pipelines have allowed us to study the transcriptome profile of non-model species (Ekblom and Galindo, 2011). Unfortunately, whole transcriptomic studies for earthworms are rare. However, Paul *et al*. (2018) examined differential gene expression under freezing and warm conditions

in *Dendrobaena octaedra* Savigny, 1826, identifying numerous genes involved in the response to each abiotic stressor, but very few shared between the two conditions. In other animals, thermal responses showed differential expression of stress-responsive genes, such as genes regulating metabolism, cell cycle, protein folding repair systems, or oxidation-reduction processes (e.g., Xu *et al*., 2018; Zheng *et al*., 2019).

The aim of this study was to explore i) the transcriptional changes in the riparian earthworm *E. tetraedra* during freezing and ii) during desiccation conditions. Also, iii) to investigate possible common responses to both abiotic stressors.

5.2 Material and methods

5.2.1 Experimental animals and laboratory conditions

In February 2018 we collected by manual sorting juveniles and adults of *Eiseniella tetraedra* in the surroundings of Fuente de las Hondillas, in Guadarrama, Madrid, Spain (40°41'50.79"N, 4° 8'22.75"O). In order to establish a culture of *E. tetraedra* in the laboratory we prepared boxes with dry soil from the collection site and distilled water up to 60% humidity. We stored the earthworms in this medium at 8°C for 1 year and 4 months, changing soil every month approximately. We tried other culture methods, such as the one followed by Barne and Striganova (2004) or those described in Supplementary Table 1 but earthworms died.

5.2.2 Freezing experiment

Eight mature earthworms from the laboratory culture were weighted and ranged between 0.13 g and 0.28 g. Eight petri dishes were prepared with 30 g of soil and distilled water at 60% of humidity. One earthworm was placed on each dish and they

were introduced in a freezer at -14 °C ± 2 °C. The reaction of earthworms was checked every ten minutes for the first two hours and every twenty minutes afterwards for less manipulation by touching them with tweezers for a further three hours. Then, earthworms were returned to the initial control conditions and two of them died the next day. Final conditions for the RNA-Seq experiment were based on the results from this preliminary trial (Supplementary Table 2) and earthworms were exposed to freezing conditions for one hour (the first time point at which all of the earthworms were less reactive). For the RNA-Seq experiment six petri dishes were prepared with 30g of soil and distilled water (60% moisture). Six mature earthworms from laboratory culture were weighted and ranged between 0.08 g and 0.22 g. One earthworm was introduced per petri dish; three of them were placed at 8°C (control) and three of them were introduced in the freezer at -14 °C ± 2 °C (cold). After one hour all individuals were flash frozen in dry ice and stored at -80°C (Figure 1).

5.2.3 Desiccation experiment

Similar to the freezing experiment, a preliminary experiment was conducted to understand the limiting desiccation conditions of earthworms. Eight earthworms were weighted and ranged from 0.13 g to 0.27 g. Eight petri dishes were prepared with 16 g of soil and distilled water of 30%. One earthworm was placed per plate and kept at 8°C for 24 hours. New petri dishes were prepared with 15% moisture and earthworms were placed in the new petri dishes. They also remained at 8 °C for 24 hours, after which time they were no longer reactive. The conditions for the RNA-Seq experiment were exactly the same as in this preliminary experiment. Six petri dishes were prepared with 16 g of soil. Six mature earthworms from the laboratory culture were

weighted and ranged in weight from 0.16 g to 0.39 g. One earthworm was introduced per petri dish; three of them at 30% moisture (desiccation) and three at 60% moisture (control) and kept at 8 °C for 24 hours. To ensure that all individuals had the same manipulation, the six earthworms were placed in new petri dishes; three of them (desiccation) at 15% humidity and three (control) at 60% humidity and kept at 8 °C. After 24 hours, all individuals were flash frozen in dry ice and stored at -80 °C (Figure 1)

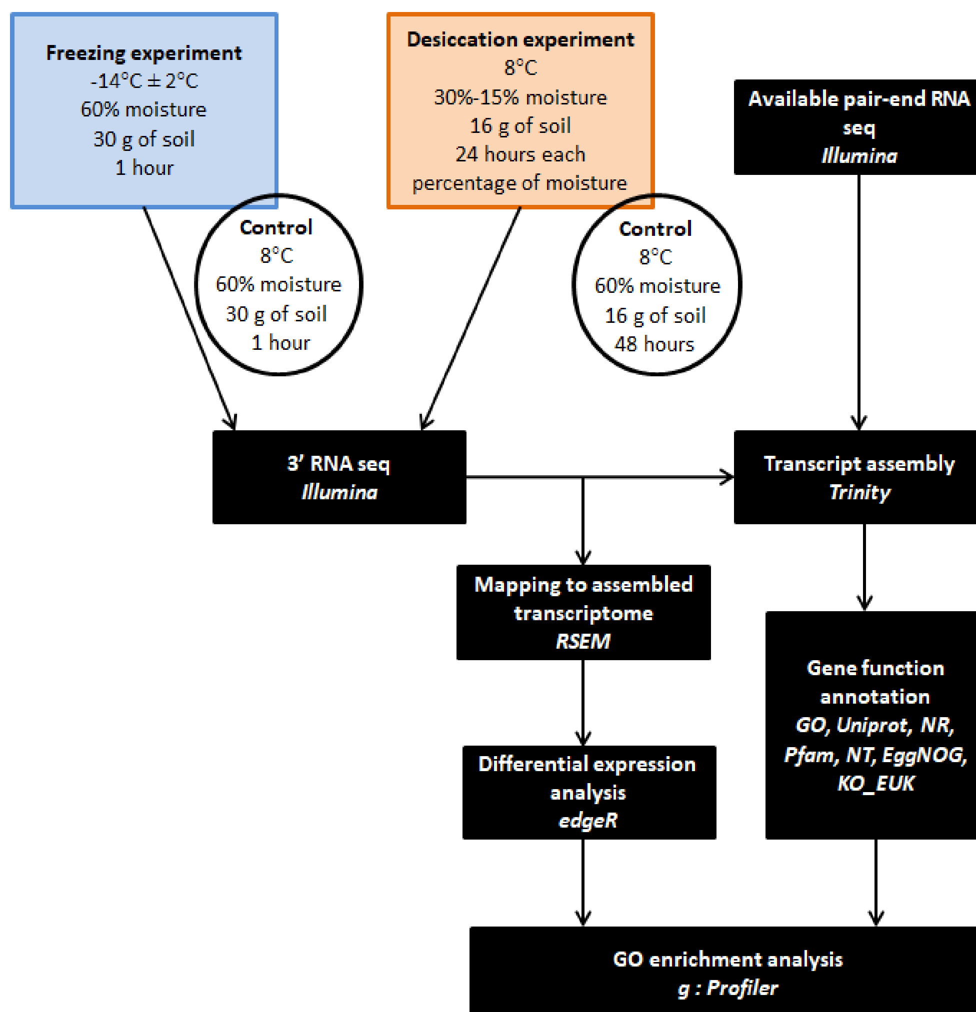


Figure 1. Schematic representation of the experimental design and data analysis workflow performed for desiccation and cold conditions in the earthworm *Eiseniella tetraedra*.

5.2.4 RNA extraction and species verification

The frozen tissue from six earthworms per experiment was powdered with a mortar and pestle and 50 mg were used for RNA and DNA extractions. Powdered tissue was homogenized in 1.5 ml of Trizol (Invitrogen) and RNA was extracted according to the manufacturer's protocol. Subsequently, samples were treated with RNAse-free DNAse (Roche) for 90 min and organic extraction was performed using phenol-chloroform-isoamyl alcohol and Phase Lock Gel Light tubes (5Prime). RNA integrity was verified using Agilent 2100 Bioanalyzer and quality and concentration of RNA was estimated with Nanodrop. For DNA extraction, the organic phases left from the RNA extractions were used and DNA was precipitated with ethanol.

A fragment of cytochrome c oxidase subunit I (COI) was amplified using primer sequences and polymerase chain reactions (PCR) following Pop *et al*. (2003). All PCRs were specific and resolved via 1% agarose gel electrophoresis; they were visualised with GelRed stain (Biotium). All products were purified using ExoSAP-IT reagent (ThermoFisher Scientific). PCR products were sequenced by Macrogen Spain Inc and species identity was verified through blasting with previous sequences from Chapter 1.

5.2.5 Transcriptome sequencing

Once verified the identity of the animals for which RNA was extracted, three samples per treatment were further processed for RNA-seq. 3’RNAseq libraries were prepared from ~500ng of total RNA per sample using the Lexogen QuantSeq 3’ mRNA-Seq Library Prep Kit FWD for Illumina (https://www.lexogen.com/quantseq-3mrna-sequencing/). The libraries were quantified on a Molecular Devices Spectra Max M2 plate reader (with the intercalating dye QuantiFluor) and pooled accordingly for

maximum evenness. The pool was quantified by digital PCR and sequenced on one lane of an Illumina NextSeq500 sequencer, single-end 1x86bp, and de-multiplexed based upon six base i7 indices using Illumina bcl2fastq2 software (version 2.18; Illumina, Inc., San Diego, CA). Illumina adapters, poly-A tails, and poly-G stretches of at least 10 bases in length were removed from the de-multiplexed fastq files using the BBDuk program in the package BBMap (https://sourceforge.net/projects/bbmap/), keeping reads at least 18 bases in length after trimming. Poly-G stretches result from sequencing past the ends of short fragments (G = no signal).

5.2.6 Establishing a reference transcriptome

Trinity v2.8.4 (Grabherr *et al*., 2011) was used for establishing two different reference transcriptomes. Firstly, a transcriptome was assembled with default options using available 101bp paired-end reads from an ongoing work (Planelló and Herrero unpublished, ENA reference number PRJEB49685 (ERP134202)). These reads were sequenced from an individual of *E. tetraedra* collected in North Western Spain. A second transcriptome was generated by combining the mentioned paired-end reads with 3´RNAseq reads from this experiment in order to check whether we could improve mapping. We normalised reads according to Trinity protocol (insilico_read_normalization.pl) and then combined 3´RNAseq reads with R1 from the paired-end set and proceeded with the assembly including the R2. We then assessed both assemblies with BUSCO using metazoa_odb10 database (Seppey *et al*., 2019) and performed some mapping trials with Bowtie 1 and 2 (Langmead *et al*., 2009; Langmead and Salzberg 2012). We found that the combined transcriptome included too much duplication, which provoked a dispersion of the reads in the mapping and therefore

lower mapping rates (398,361 contigs, 91% complete BUSCOS, 16% complete and single, 75% complete and duplicate). The transcriptome assembled only with the pair-end reads on the other hand showed lower duplication rates and higher mapping (143,339 contigs, 91% complete BUSCOS, 87% complete and single, 4% complete and duplicate). For this reason, we decided to continue the analyses with the transcriptome assembled only with the paired-end reads.

5.2.7 Quantitative transcriptomic analysis

A differential gene expression analysis was performed with the Trinity module (Grabherr *et al*., 2011) which integrates Bowtie 1.3.0 (Langmead *et al*., 2009) to align input reads from each individual library against the reference transcriptome, RSEM (Li and Dewey, 2011) to estimate transcript abundance using the alignments, and edgeR (Robinson *et al*., 2010) to perform pairwise comparisons of expression levels and extract Differentially Expressed Genes (DEGs). edgeR was run with a p-value cut off of 0.001, a false discovery rate <0.05 and min abs(log2(a/b)) change of 1 (therefore, minimally, 2-fold change).

Enrichment analyses were performed in g: Profiler (Reimand *et al*., 2019) using the detected differentially expressed transcripts (FDR<0.05) for each experiment, specifically for upregulated and downregulated genes subsets separately, using the complete transcriptome of *E. tetraedra* as reference. GO terms with adjusted P value < 0.05 were considered significantly enriched.

5.3 Results

5.3.1 Sequencing output

Reads were submitted to the European Nucleotide Archive (ENA) under study number PRJEB46360 (Samples ERS6655476-ERS6655487). For each sample, 2-40 million clean single-ended reads were used for analyses (Supplementary Table 3). Between 45 - 65% of the reads from each sample mapped against the reference transcriptome, and were therefore informative for subsequent quantitative analyses.

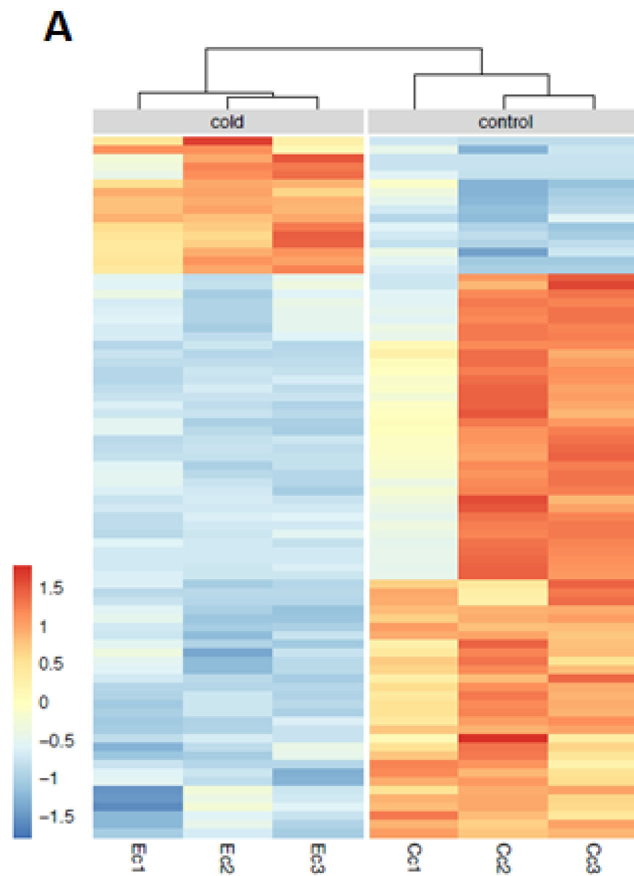
5.3.2 Differential gene expression under freezing conditions

Hierarchical clustering analysis of the differentially expressed transcripts revealed differences in gene expression between control earthworms and those under freezing conditions (Figure 2A). Comparisons of gene expression levels between earthworms under freezing conditions and control earthworms generated a total of 84 DEGs. Of those, 17 DEGs were upregulated, while 67 were downregulated and a total of 21 could be functionally annotated (Table 1). Among the downregulated DEGs, seven genes related to the respiratory chain were expressed: chains 4, 5, and 6 of NADH-ubiquinone oxidoreductase, cytochrome c oxidase (subunits 1 and 2), cytochrome b, and ATP synthase subunit a. Moreover, glutamyl aminopeptidase and C-type lectin domain family 4 member D were also found downregulated. In contrast, the genes encoding giant extracellular hemoglobin linker 2 chain, uridine-cytidine kinase-like 1, GPI-linked NAD (P)(+)-arginine ADP -ribosyltransferase 1-like (GLP1R) and lectin, were upregulated. To improve our knowledge of the biological functions represented by DEGs, a GO enrichment analysis was performed (Figure 2B). Among the downregulated

DEGs, nine GO terms belonging to the Biological Processes (BP) category were enriched. Transcriptional changes under freezing conditions were characterized in E. tetraedra by downregulation of aerobic respiration (GO:0009060), including several steps of this process, such as oxidative phosphorylation (GO:0006119) which is an ATP biosynthetic process (GO:0006754), proton transmembrane transport (GO:1902600) and mitochondrial electron transport, cytochrome c to oxygen (GO:0006123). In addition, purine biosynthesis processes were downregulated (GO:0009206 and GO:0009145). In contrast, pyrimidine ribonucleotide and nucleotide salvage (GO:0010138 and GO:0032262) and pyrimidine biosynthesis and metabolism processes such as the pyrimidine ribonucleoside triphosphate biosynthesis process (GO:0009209) and the pyrimidine ribonucleoside monophosphate metabolism process (GO:0009173) were upregulated. As for the molecular function category, uridine kinase activity (GO:0004849) and nucleoside kinase activity (GO:0019206) were upregulated. As expected, most of the GO terms belonging to the category of molecular functions were consistent with the downregulation of aerobic cellular respiration. Thus, detected that genes associated with seven aerobic cellular respiration-related GO terms were downregulated, such as proton transmembrane transporter activity (GO:0015078), cytochrome c oxidase activity (GO:0004129), and electron transfer activity (GO:0009055), among others (Figure 2B). Genes implied in copper ion binding (GO:0005507) were also downregulated.

Gene	Protein	logFC
ART1	GPI-linked NAD(P)(+)--arginine ADP-ribosyltransferase 1-like	3,66501667
LEC	Lectin	2,71902448
UCKL1	Uridine-cytidine kinase-like 1	7,82968242
N/A	Giant extracellular hemoglobin linker 2 chain	2,53081466
ENPEP	Glutamyl aminopeptidase	-5,18492199
Clec4d	C-type lectin domain family 4 member D	-3,47079202
COI	Cytochrome c oxidase subunit 1	-7,47972063
MT-CYB	Cytochrome b	-7,15222671
ND5	NADH-ubiquinone oxidoreductase chain 5	-7,93382089
ND6	NADH-ubiquinone oxidoreductase chain 6	-7,4433782
ATP6	ATP synthase subunit a	-6,40678916
N/A	Vasotocin-neurophysin VT 1	-4,4569175
N/A	Genome polyprotein	-4,80785628
N/A	Retrovirus-related Pol polyprotein from type-1 retrotransposable element R2	-5,99953981
Ttn	Titin	-7,90297656
COII	Cytochrome c oxidase subunit 2	-6,25792581
ND4	NADH-ubiquinone oxidoreductase chain 4	-7,08596604

Table 1. Annotated differentially expressed genes (DEGs) for the cold experiment in the earthworm *Eiseniella tetraedra*. Positive log fold change values (logFC) represent upregulated DEGs (in dark blue), while negatives represent downregulated DEGs (in light blue).



B

Down-DEGs

proton transmembrane transporter activity
cytochrome-c oxidase activity
oxidoreductase activity, acting on a heme group of donors
inorganic cation transmembrane transporter activity
electron transfer activity
proton-transporting ATP synthase activity, rotational mechanism
copper ion binding
oxidative phosphorylation
proton transmembrane transport
aerobic respiration
cellular respiration
mitochondrial electron transport, cytochrome c to oxygen
energy derivation by oxidation of organic compounds
ATP biosynthetic process
purine ribonucleoside triphosphate biosynthetic process
purine nucleoside triphosphate biosynthetic process
respiratory chain complex IV
cytochrome complex
mitochondrial envelope
inner mitochondrial membrane protein complex
mitochondrial respiratory chain complex IV
mitochondrial proton-transporting ATP synthase complex
proton-transporting ATP synthase complex, coupling factor F(o)
mitochondrial respiratory chain complex III
mitochondrial respirasome
oxidoreductase complex

0,00 0,50 1,00 1,50 2,00 2,50 3,00 3,50 4,00 4,50 5,00

-log10 (padj)

Up-DEGs

uridine kinase activity
nucleoside kinase activity
pyrimidine nucleotide salvage
pyrimidine ribonucleotide salvage
pyrimidine ribonucleoside triphosphate biosynthetic process
pyrimidine ribonucleoside monophosphate biosynthetic process
pyrimidine ribonucleoside monophosphate metabolic process
pyrimidine ribonucleoside triphosphate metabolic process
glycosyl compound metabolic process
ribose phosphate biosynthetic process

0,00 0,50 1,00 1,50 2,00 2,50

-log10 (padj)

Figure 2. A: Hierarchically-clustered heatmap of all differentially expressed genes under freezing conditions in the earthworm *Eiseniella tetraedra*. Red color indicates high expression values and blue color indicates low expression values. B: Results of GO enrichment analysis performed with differentially expressed genes (DEGs). Terms in the Molecular Function (MF) category are shown in green, Biological Processes (BP) in purple, and Cellular Components (CC) in blue. All shown GO terms were significantly enriched (adjusted p value <0.05). For a complete list of results of GO enrichment analysis see Supplementary Table 3 and 4.

5.3.3 Differential gene expression under desiccation conditions

Hierarchical clustering analysis of the differentially expressed transcripts revealed clearly differences in gene expression between control and earthworms under desiccation conditions (Figure 3A). Comparison of genes between earthworms under desiccation conditions and control earthworms yielded a total of 163 differentially expressed genes (DEGs), with 58 upregulated and 105 downregulated, of which we could functionally annotate 43 (Table 2). Several genes were found upregulated under desiccation conditions, including genes encoding enzymes involved in protein modification, such as UDP-N-acetylglucosamine-peptide-N-acetylglucosaminyltransferase (110 kDa subunit), which glycosylates a large and diverse number of proteins, and peptidyl-prolyl cis-trans isomerase B. In addition, maternal embryonic leucine zipper kinase, which is involved in various processes such as cell cycle regulation, stem cell self-renewal, apoptosis, and splicing regulation, and the protein CREBRF homolog, a transcriptional regulator that acts in the TORC1 pathway to regulate energy homeostasis and promote survival during nutrient

deprivation, were also upregulated. In contrast, several genes involved in protein transport and degradation were downregulated, including sorting nexin-27 and UBX domain-containing protein. Apoptosis regulators such as serine/threonine protein kinase 17A were also found downregulated. To improve our knowledge of the biological functions represented by DEGs, a GO enrichment analysis was performed (Figure 3B). We only found enriched GO terms belonging to the Molecular Function (MF) among the downregulated DEGs. Enriched GO terms were related to dehydrogenase activities such as testosterone dehydrogenase (NAD +) activity (GO:0047035) and alcohol dehydrogenase [NAD (P)+] activity (GO:0018455). Also fucosyltransferase activity (GO:0008417), such as galactoside 2-alpha-L-fucosyltransferase activity (GO:0008107). In addition, GO terms such as protein folding chaperone (GO:0044183) and long chain fatty acid transporter activity (GO:0005324) appeared enriched. Isomerase activity (GO:0016853) and racemase and epimerase activity (GO:0016854) were also enriched among the downregulated DEGs.

Gene	Protein	logFC
melk	Maternal embryonic leucine zipper kinase	2,42595792
Ogt	UDP-N-acetulglucosamine peptide N-acetylglusaminyltransferase 110 kDa subunit	2,32822656
DIO3	Thyrosine 5-deiodinase	2,57000343
RNF11	RING finger protein 11	2,17304219
Ppib	Peptidyl-prolyl cis-trans isomerase B	3,03414854
REPTOR	Protein CREBRF homolog	1,95570455
gag-pol	Gag-Pol polyprotein	1,81829307
SEC61A2	Protein transport protein Sec61 subunit alpha isoform 2	-2,99656266
vps25	Vacuolar protein-sorting-associated protein 25	-9,57780365
Mtorc2	Mitochondrial amodoxime reducing component 2	-5,35907783
UGT2A3	UDP-glucoronosyltransferase 2A3	-3,15670948
Igf2bp1	insuline-like growth factor 2 mRNA-binding protein 1	-5,79504615
BSG	Basigin	-2,79364374
adcy5	Adenylate cyclase type 5	-5,26211640
STK17A	Serine/threonine-proteine kinase 17A	-5,00263744
N/A	Soma ferritin	-2,15117738
N/A	Parvalbumin alpha	-4,78357177
SNX27	Sorting nexin-27	-4,11397031
KIDINS220	Kinase D-interacting substrate of 220 kDa	-8,94636595
clec4a	C-type lectin domain family 4 member A	-3,22700454
Rmdn1	Regulator of microtubule dynamics protein 1	-4,52941819
Plod3	Multifunctional procollagen lysine hydroxylase and glycosyltransferase LH3	-3,58851204
TM1	Tropomyosin	-2,28803424
CD109	CD109 antigen	-3,49888030
dld	Delta-like protein D	-4,69949756
N/A	Extracellular globin-1	-4,20395984
DHRS9	Dehydrogenase/reductase SDR family member 9	-2,91725090
TMEM26	Transmembrane protein 26	-4,25613240
FUT1	Galactoside alpha-(1,2)-fucosyltransferase 1	-3,93638172
AMY2	Pancreatic alpha-amylase	-5,54200867
chmp3	Charged multivesicular body protein 3	-2,51423075
ABCD4	Lysosomal cobalamin transporter ABCD4	-7,77174965
N/A	C-type lectin mannose-binding isoform	-2,80057574
SLC1A2	Excitatory amino acid transporter 2	-2,27316235
DAZAP2	DAZ-associated protein 2	-2,12466453
Mocs2	Molybdopterin synthase catalytic subunit	-2,48858001
Ubxn4	UBX domain-containing protein 4	-7,62518905
FASTKD3	FAST kinase domain-containing protein 3, mitochondrial	-3,35739905
RAB14	Ras-related protein Rab-14	-1,78479830
CREB3L2	Cyclic AMP-responsive element-binding protein 3-like protein 2	-2,73883959
SPRYD3	SPRY domain-containing protein 3	-2,37690372
PI16	Peptidase inhibitor 16	-2,75902813
FKBP8	Peptidyl-prolyl cis-trans isomerase FKBP8	-2,57298054

Table 2. Annotated differentially expressed genes (DEGs) for the desiccation experiment in the earthworm *Eiseniella tetraedra*. Positive log fold change values (logFC) represent upregulated DEGs (in dark orange), while negatives represent downregulated DEGs (in light orange).

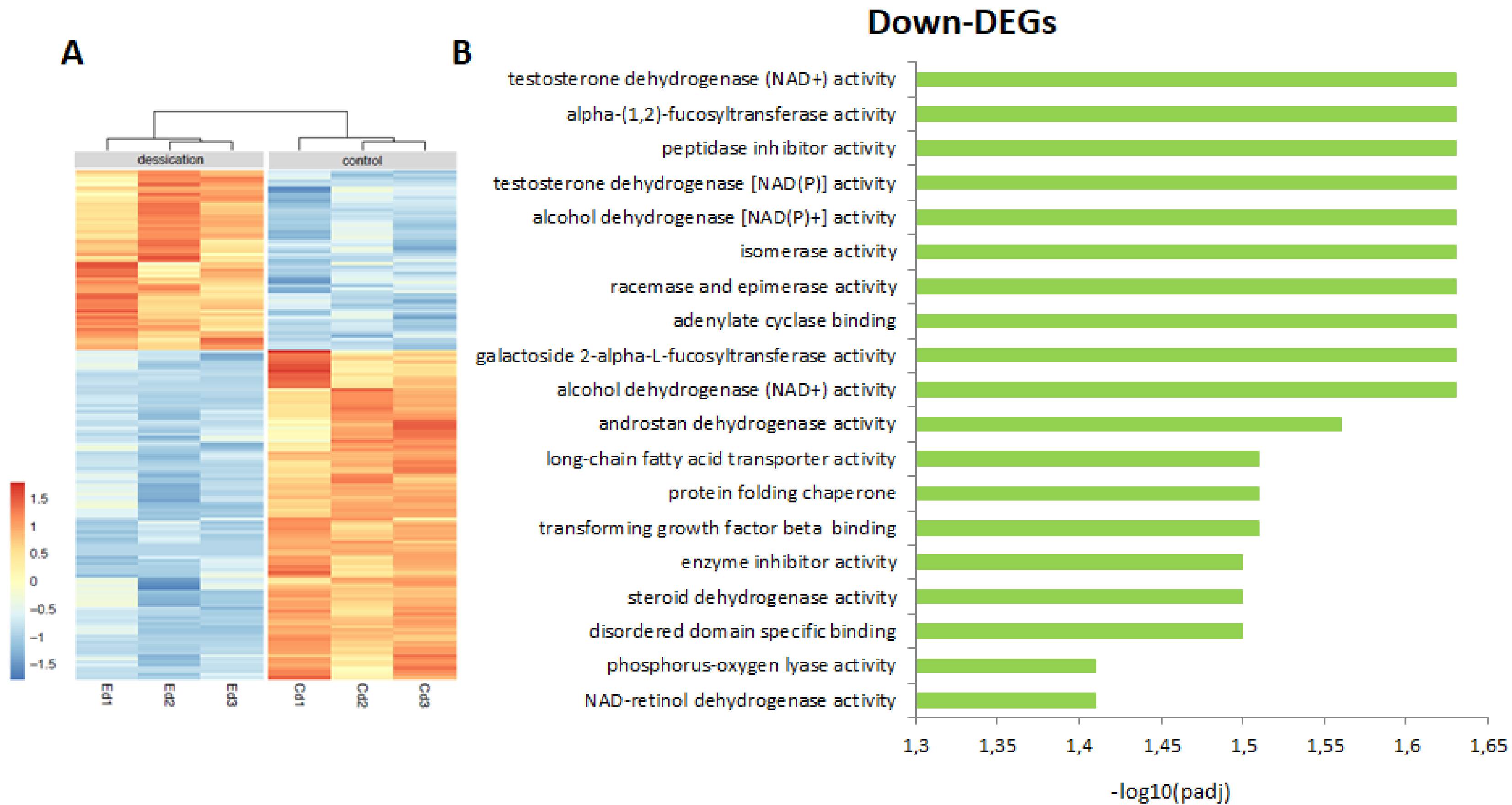

A
B
Down-DEGs
dessication
control
Ed1
Ed2
Ed3
Cd1
Cd2
Cd3
1.5
1
0.5
0
−0.5
−1
−1.5
testosterone dehydrogenase (NAD+) activity
alpha-(1,2)-fucosyltransferase activity
peptidase inhibitor activity
testosterone dehydrogenase [NAD(P)] activity
alcohol dehydrogenase [NAD(P)+] activity
isomerase activity
racemase and epimerase activity
adenylate cyclase binding
galactoside 2-alpha-L-fucosyltransferase activity
alcohol dehydrogenase (NAD+) activity
androstan dehydrogenase activity
long-chain fatty acid transporter activity
protein folding chaperone
transforming growth factor beta binding
enzyme inhibitor activity
steroid dehydrogenase activity
disordered domain specific binding
phosphorus-oxygen lyase activity
NAD-retinol dehydrogenase activity
1,3
1,35
1,4
1,45
1,5
1,55
1,6
1,65
-log10(padj)

Figure 3. A: Hierarchically-clustered heatmap of all differentially expressed genes under desiccation conditions in *Eiseniella tetraedra* earthworms. Red color indicates high expression values and blue color indicates low expression values. B: Enriched GO terms obtained from downregulated differentially expressed genes (DEGs) in dessication samples. Terms corresponding to the molecular function (MF) category are shown in green. All shown GO terms were significantly enriched (adjusted p value <0.05). No terms for Biological Process (BP) or Cellular Component (CC) were enriched. No enriched GO terms were obtained from upregulated DEGs. For a complete list of results of GO enrichment analysis see Supplementary Table 5.

5.4 Discussion

A complex machinery of gene expression and biochemical adaptive responses is induced by thermal stress in most organisms (Lindquist, 1986; Fujita, 1999). This response enables organisms to survive and adapt to thermal stress and it is a fundamental requirement for species exposed to temperature fluctuations, such as *Eiseniella tetraedra*. Its cosmopolitan distribution implies survival in a wide range of climatic conditions. Cold and desiccation have many similar effects at the cellular level in other invertebrates, such as dehydration and osmotic stress (Sinclair *et al*., 2013; Paul *et al*., 2018). However, we found no common responses of both treatments in *E. tetraedra*. Further experiments will unveil whether those common responses exist.

5.4.1 Cold tolerance

The ecophysiology of cold tolerance in many terrestrial invertebrates is based on water and its activity at low temperatures affecting the functions of cells, tissues and the

whole organism (Block, 2003). Even though cold shock response has been studied for several decades in a number of different organisms, we have only recently begun to understand the molecular mechanisms that control the adaptation to cold stress. Remarkably, all organisms, from prokaryotes to plants to higher eukaryotes, respond to cold shock in comparatively similar ways (Al-Fageeh and Smales, 2006).

Post-translational modifications of cellular proteins by ADP-ribosylation have been associated with a number of cellular processes (Fabrizio *et al*., 2015). The *ART1* gene, which encodes the protein GPI-linked NAD (P)(+)-arginine ADP-ribosyltransferase 1-like, was found to be upregulated in our freezing experiments. It has ADP - ribosyltransferase activity towards the glucagon-like peptide-1 receptor (GLP1R), which is involved in the control of blood glucose levels. In mammals, GLP1R has the ability to induce insulin secretion in response to high glucose levels (Gutniak *et al*., 1992). In general, ADP-ribosyltransferases inhibit the function of the target protein (Álvarez-Herrera, 2004). Since the accumulation of glucose can slow down the freezing process and even reduce the amount of ice formed (Holstrump, 2003), it seems reasonable that *ART1* is upregulated to inhibit the function of GLP1R and thus accumulate glucose to cope with freezing temperatures.

Earthworms have cutaneous respiration, they perform gas exchange through the skin. Therefore, air dissolves on the mucus of their skin, so they must remain moist to breathe. The skin of *E. tetraedra* was nearly frozen during the freezing experiment, causing it to lose much of its moisture and making gas exchange more difficult. The giant extracellular hemoglobin of annelids has been studied in detail because of its exceptional oxygen transport properties. Interestingly, expression of the gene for giant

extracellular hemoglobin of *Eisenia fetida* Eisen, 1874 is induced in the regenerating tissue (Bhambri *et al*., 2018). It is possible that the upregulation of the gene encoding the linker 2 chain of extracellular giant hemoglobin contributes to the enhancement of respiration in *E. tetraedra* at low temperatures due to its oxygen carrier activity. Also, it may be involved in tissue repair due to its destruction by freezing.

Uridine kinase activity (GO:0004849) is involved in the pyrimidine salvage pathway (GO:0032262), and both molecular functions have been upregulated. In the cell, nucleotides are produced either by the de novo pathway or by the salvage pathway. Uridine kinase phosphorylates uridine to uridine monophosphate using ATP as a phosphate donor (Kashuba *et al*., 2002). It has been suggested that de novo synthesized UTP is preferentially used for the production of UDP-sugars and phospholipids, whereas UTP produced via the salvage pathway is used for RNA synthesis (Anderson and Parkinson, 1997). There seems to be general agreement in the literature that increased RNA and protein synthesis may be part of the resistance mechanism of plants to low temperatures (Sarhan and D'aoust, 1975. Ribosomal RNA has been shown to be the RNA class most responsive to the physiological stimuli that causes acclimation (Khan *et al*., 1968; Brown, 1972; Gusta and Weiser, 1972). This is consistent with the upregulation of pyrimidine ribonucleotide salvage (GO:0010138) found in the freezing experiment.

The C-type lectin (CTL) protein family has been extensively studied in both vertebrates and invertebrates (Wei *et al*., 2010). CTL is Ca2+ dependent, and all members of the CTL family of invertebrates have one or more carbohydrate recognition domains (CRD) (Li *et al*., 2020). CTLs from *Bombyx mori* Linnaeus, 1758 play an important role in the

activation of immune signaling pathways (Zhu *et al*., 2016). They are also associated with tissue regeneration (Gao *et al*., 2017). In addition, it has been demonstrated that there is some relationship between herring antifreeze protein (AFP) type II and the carbohydrate binding site of C-type lectin CRDs. The ice binding site of AFP corresponds to the CRD binding site of CTL, showing that the carbohydrate binding site of C-type lectin evolved into an ice binding site (Ewart and Fletcher, 1993; Ewart *et al*., 1998). Thus, CTL has a potential function in low temperature tolerance and cold resistance. In samples of *E. tetraedra* subjected to cold treatment, the gene *CLEC4D* encoding C-type lectin domain family 4 member D was downregulated, similarly to the transcription of CTL mRNA in the liver of *Oreochromis niloticus* Linnaeus, 1758 under cold stress, which is significantly decreased (Yang *et al*., 2016). However, the expression level of CTLs in rainbow trout was hardly affected by cold stress (Rebl *et al*., 2012) and in *Venerupis* (*Ruditapes*) *philippinarum* (Adams and Reeve, 1850) low-temperature stress promotes the synthesis of CTLs in the hepatopancreas, which could explain the upregulation of the *LEC* gene in *E. tetraedra* that may be involved in protection of eggs and embryos against microorganisms.

From our results, we can conclude that the respiratory chain of *E. tetraedra* is clearly affected under cold conditions. Seven genes were down-regulated: chains 4, 5 and 6 of NADH-ubiquinone oxidoreductase, cytochrome c oxidase (subunits 1 and 2), cytochrome b and ATP synthase subunit a. NADH-ubiquinone oxidoreductase and cytochrome c oxidase are involved in the transfer of electrons to oxygen. Therefore, they are considered as the key components of the respiratory chain (Hatefi, 1985). They were also downregulated under other stress conditions such as different light intensities (Li *et al*., 2020) or anoxia (Cai and Storey, 1996). Moreover, several

biological processes and molecular functions involved in the respiratory chain were enriched among the downregulated DEGs in our results, such as oxidative phosphorylation (GO:0006119), cytochrome c oxidase activity (GO:0004129) or cellular respiration (GO:0045333). Nevertheless, some studies have shown the importance of energy in maintaining metabolic homeostasis at low temperatures for poikilotherms. For example, animals with a high capacity for ATP production under cold conditions also show good tolerance to cold stress (Wang *et al*., 2014; Lu *et al*., 2017). Thus, the downregulation of a variety of genes involved in the respiratory chain suggests that *E. tetraedra* may not be a good candidate to sustain low temperature conditions. However, it is a cosmopolitan species present in even subarctic climates. It is possible that the upregulation of *ART1*, which presumably leads to glucose accumulation, counteracts the lack of energy from ATP and allows *E. tetraedra* to live in low temperature zones, but not in freezing conditions. This may explain why Terhivuo and Saura (1997) reported the yearly disappearance of *E. tetraedra* caused by habitat freezing in the Aland Archipelago (Baltic Sea). However, some cocoons of E. tetraedra are able to survive at -28°C (pers. observ.). Whether the cocoons may be the frost-resistant form of this species needs to be investigated in further studies.

5.4.2 Desiccation tolerance

Water accounts for 70% of the weight of a living cell (Alberts *et al*., 2008) and water is the medium in which most cellular reactions occur. Lack or scarcity of water is a big stressor for organisms. A loss of more than 20-50% of their water content is fatal to most higher plants (Kranner *et al*., 2002) and most animals die when they lose more than 15-20% of their body water (Barrett, 1991). Seasonal variation in earthworm

habitats is a natural phenomenon. As a semi-aquatic species (Omodeo and Rota, 1991), *E. tetraedra* is highly dependent on the availability of water. Therefore, it is believed to be extremely susceptible to drought stress.

Phosphorylation and dephosphorylation of proteins are regulatory mechanisms in many organisms in which protein kinases play an important role (Hunter, 1995). Depending on the substrate, protein kinases are divided into serine/threonine kinases and tyrosine kinases (Schenk and Snaarjagalska, 1999). Maternal embryonic leucine zipper kinase (MELK), encoded by the gene also known as *MELK*, is a serine/threonine kinase involved in various cellular processes such as cell division (Nakano *et al*., 2005; Cordes *et al*., 2006; Niesler *et al*., 2007; Le Page *et al*.,2011; Chien *et al*., 2013), transcription (Seong *et al*., 2002), pre-mRNA splicing (Vulsteke *et al*., 2004), DNA repair (Bensimon *et al*., 2011) and apoptosis (Lin *et al*., 2007; Jung *et al*., 2008). In addition, *MELK* has been implicated in DNA damage response pathways (Hurov *et al*., 2010; Bensimon *et al*., 2011). Thus, elevated MELK protein levels increase resistance to DNA-damaging events (Choi and Ku, 2011). In severe cases, abiotic stress can trigger apoptosis or programmed cell death (Wang and Kaufman, 2012) and DNA damage response pathways (Nisa *et al*., 2019). Serine threonine protein kinase 17A, encoded by the gene *STK17A*, appeared downregulated in our results. It belongs to a member of the death associated protein kinase family, and acts as a positive regulator of apoptosis (Sanjo *et al*., 1998; Mao *et al*., 2011). Therefore, upregulation of *MELK* and downregulation of *STK17A* in *E. tetraedra* during desiccation conditions could protect the animal from apoptosis and DNA damage.

Peptidyl-prolyl cis-trans isomerases (PPiase) are a group of chaperones found in all organisms. They are involved in many biochemical processes such as signal transduction, protein folding and development. Their occurrence in plants is comparable to that of heat shock proteins (Schiene-Fisher and Yu, 2001). As in the desiccation treatment of *E. tetraedra*, PPiases were up-regulated under drought stress in rice (Wang *et al*., 2012) and under heat conditions in the springtail *Folsomia candida* Willem, 1902 (Nota *et al*., 2010). Moreover, they are involved in the anhydrobiosis process of the nematode *Panagrolaimus superbus* Fuchs, 1930. In contrast, in tardigrades subjected to desiccation stress, the concentration of another chaperone family, namely Hsp70, was reduced, whereas elevated levels were detected after rehydration. This suggests that the role of Hsp70 may be related to repair processes after desiccation rather than biochemical stabilization in the dry state (Jönsson *et al*., 2007). Consequently, this could be the reason why the GO: 0044183, which belongs to the protein folding chaperones, appeared downregulated in *E. tetraedra*.

The gene *REPTOR*, which encodes the protein CREBRF homolog, plays a key role in regulating energy homeostasis and promotes survival during nutrient deprivation (Tiebe *et al*., 2015). It maintains the organism's metabolism by activating the expression of target genes of the stress response, including genes involved in glycogenesis and triglyceride biosynthesis. Therefore, the upregulation of this gene under the drought conditions in our experiment seems to be reasonable, as it maintains homeostasis and covers the lack of nutrients caused by the absence of moss or plant roots in the petri dishes, which *E. tetraedra* also eats (pers. observ.).

The UBX domain-containing protein family is evolutionarily conserved and is present in several model organisms (Hartmann-Petersen *et al*., 2003; Schuberth *et al*., 2004; Yamauchi *et al*., 2007; Lee *et al*., 2010) where it is involved in oxidative stress and osmotic stress response (Zhang *et al*., 2017). Two other domains are found in the UBX domain: ubiquitin-associated (UBA) and ubiquitin-associated (UAS) (Zhang *et al*., 2017). The UAS domain has been shown to interact with long-chain unsaturated fatty acids (Kim *et al*., 2013). Like UBX, the transporter activity for long-chain fatty acids (GO:0005324) and several GO terms related to dehydrogenase activities (such as GO:0047035, GO:0018455 and GO:0047044) were also downregulated in our desiccation experiment. The gene *DHRS9*, which is involved in all these dehydrogenase activities, is also related to lipid metabolism.

The *FUT1* gene is also involved in the fatty acid biosynthesis pathway (Maheswary *et al*., 2016). It encodes galactoside-2-alpha-L-fucosyltransferase, which mediates the attachment of L-fucose to the terminal β-D-galactose residues of the glycan via an α1,2 linkage (Kim *et al*., 2020). Fucosylation is a biological process that is widely distributed in vertebrates, invertebrates, plants, bacteria, and fungi (Li *et al*., 2018). It plays an important role in molecular functions such as cell adhesion and immune regulation (Kim *et al*., 2020). Fucosylated carbohydrates are involved in the regulation of several cellular functions such as cell trafficking, immune cell development, and interaction with gut microbes (Aplin *et al*., 2012; Goto *et al*., 2016). Galactoside 2-alpha-L-fucosyltransferase is also needed for the glycosylation of N-glycans, O-glycans, and glycosphingolipids (Taniguchi and Yoshida, 2017). *FUT1* appeared to be upregulated under water stress in plants (Maheswary *et al*., 2016), whereas it appeared downregulated in our desiccation experiments.

Survival under stress conditions can be maximized by two physiological mechanisms: increasing the storage of resources (energy or water) aimed for consumption during stress, or conserving resources by reducing the rate at which they are consumed (Marron *et al*., 2003). Experiments with *Drosophila melanogaster* Meigen, 1830 show that selection for stress resistance in the laboratory can lead to a reduction in metabolic rate (Hoffmann and Parsons, 1993). Lipids provide more than twice as much energy per gram as carbohydrates (Withers, 1992), so they would be a suitable fuel for starvation resistance and explain the downregulation of lipid metabolism and transport in *E. tetraedra*.

5.5 Conclusions

In this study, we identified 84 differentially expressed genes in *Eiseniella tetraedra* under cold conditions and 163 in the desiccation experiment. Some mechanisms such as dehydration and osmotic stress have been reported as common at low temperatures and low humidity, but in this case we did not detect common responses between the two experiments. Despite its cosmopolitan distribution, even in subarctic zones, *E. tetraedra* did not tolerate freezing conditions as demonstrated by the downregulation of the respiratory chain. However, upregulation of genes related to glucose accumulation suggests acclimation to low temperatures. We hypothesized that cocoons are the frost-resistant forms of this species, but this needs to be confirmed in future experiments. Earthworms perform gas exchange through their skin, and under cold conditions their skin remains nearly frozen, making it difficult for them to breathe. The upregulation of extracellular giant hemoglobin could be helpful in this process due to its exceptional oxygen transport properties. In contrast, under desiccation

conditions, genes that protect earthworms from apoptosis and DNA damage were upregulated. Lipid metabolism, however, appeared to be downregulated, presumably due to resource accumulation in case the adverse conditions persisted for a prolonged period of time.

Discussion

The studies on the species *Eiseniella tetraedra* compiled in this thesis are probably among the most extensive available on earthworms. The high number of populations studied (304) and, consequently, the number of individuals treated in each analysis, allowed for a robust test of our initial hypothesis and the conclusions obtained.

We can assert that *E. tetraedra* is an earthworm with great genetic variability (Chapter 1, Chapter 2, and Chapter 3). This genetic diversity appeared nested in different lineages which were clustered in different clades.This diversity increased as we increased the sample size of each study. However, the main lineages were consistent throughout our work and were already highlighted by the first microscale study (Chapter 1). Thus, this increase in diversity was reflected in the appearance of geographically restricted lineages (Chapter 2 and Chapter 3). Genetic distances between the lineages were high, yet remaing within the range of intraspecific variation reported for other earthworms. On the contrary, genetic distances between some clades were so large as to indicate a possible cryptic speciation (Chapter 2 and Chapter 3). However, the concept of species in parthenogenetic animals is a complicated subject (Blakemore, 1999), so no new species or subspecies have been described in this work since further evidence need to be gathered.

The phylogeography of *E. tetraedra* was not easy to describe. Probably the fact that this earthworm is spread by hydrochory (Terhivuo and Saura, 2002), anthropochory (Gates, 1977; Javidkar *et al*., 2020), and even zookory (Terhivuo and Saura, 2006) makes the evolutionary history of its lineages and clades in the areas studied not so

clear in studies with few populations. Although no geographically consistent pattern was found in the first chapter, by the second chapter it was possible to determine how the distribution of these lineages and clades appeared to respond to environmental conditions such as temperature, precipitation, pH, and habitat stability. In the third chapter, which referred to phylogeography at the European level, a clearer pattern was already evident. Clades I and II appeared to be prevalent in Eurosiberian and Mediterranean region respectively. In addition, the Ecological Niche Modelling study supported this idea, showing how cold areas were suitable for Clade I and warm areas for Clade II. Moreover, the presence of clades restricted to specific and widely known glacial refugia such as the Iberian Peninsula and Eastern Europe provided us with the basis for postglacial recolonization of certain lineages. Although it seems unlikely that earthworms inhabited the northern parts of Europe during the Last Glacial Maximum (Tiunov *et al*., 2006), the presence of other lineages of *E. tetraedra*, mainly found in the Scandinavian Peninsula, confirmed the usefulness of the famous nunataks in northern Europe as refugia for this earthworm. Despite the fact that these nunataks were ice-free zones, *E. tetraedra* must have been able to withstand the low temperatures that prevailed in these areas. Currently, we find this species even in areas further north, suggesting that this idea is not far-fetched. Furthermore, the results of Chapter 5 support this idea, where it is confirmed that *E. tetraedra* is capable of acclimating to low temperatures through changes in gene expression, although not to freezing conditions. This supports the idea that recolonization from refugia was possible in both the north and south.

As mentioned earlier in this dissertation, *E. tetraedra* occurs also in unstable environments. Depending on the season, the habitats of this earthworm can change,

so that individuals face desiccation in summer and low temperatures or even ground frost in winter. Because of the current situation on the planet as a result of man-made climate change, these temperature fluctuations are becoming more frequent. After exposing specimens of *E. tetraedra* to freezing and drying conditions, we found that they were able to acclimatize to low temperatures and low humidity in the environment, as mentioned above, by down and/or upregulating various genes involved in different metabolic pathways such as respiratory chain, lipid metabolism and different pathways involved in DNA repair and against apoptosis (Chapter 5).

In this work, the existence of subspecies described for the species *E. tetraedra* was mentioned (Chapter 1, Chapter 2 and Chapter 3). They were all based on the position of the male pore. However, our molecular work showed that this variability is unrelated to the studied markers, highlighting this subespecies as meaningless from a systematic perspective (Chapter 1). A tremendous variability has been found in traits related to the male reproductive system (Chapter 3). This suggests that individuals with this variability are not under selection pressure, which is to be expected in a parthenogenetic earthworm. Thus, the fact that these traits do or do not occur, or where they occur, has no a priori effect on the fitness of the individuals that exhibit these modifications, since they are traits that are of no use to them. The presence of spermatophores was not common (Chapter 3). Out of 1,147 individuals examined morphologically, only 37 exhibited them. In all of them the spermatophore was empty, containing no spermatozoa. Although the presence of spermatophores in animals that do not have gamete exchange may be contradictory, we found evidence in the literature for the presence of these structures in other parthenogenetic earthworms. Apparently, the exchange of empty spermatophores would be necessary to activate

parthenogenesis (Muldal, 1952; Jaenicke and Selander, 1979). It is possible that this action is an intermediate step prior to pure parthenogenesis, and based on the number of individuals found with these structures; this would be a process that is being lost in *E. tetraedra*.

As mentioned in Chapter 4, *Eiseniella neapolitana* has been considered a subspecies of *E. tetraedra* by some authors (Stephenson, 1924; Bodenheimer, 1937; Cernosvitov, 1938, Cernosvitov, 1940; Pavlicek *et al*., 2003) and a valid species by others (Csuzdi and Pavlicek, 2005). It was also considered as a synonym of *Norealidys andaluciana*. From the results of Chapter 4, we can conclude that *E. tetraedra* and *E. neapolitana* are distinct species. More so, the phylogenetic analysis, genetic divergence based on COI and significant differences in the number of segments support the earthworm genus *Norealidys* proposed by Qiu and Bouché (1998). Thus, according to the phylogenetic analysis for lumbricids, *E. neapolitana* is different but closely related to *N. andaluciana* and has been renamed *Norealidys neapolitana*. In the populations studied in Chapter 4, *N. neapolitana* occurred as a cryptic complex, and therefore its genetic diversity should be studied in more detail by including new populations. In the future, it would be interesting to investigate why *N. neapolitana* occurs in other countries with a Mediterranean distribution, but not in the Iberian Peninsula, while the species *N. andaluciana* occurs in this area. The Cypriot specimen included in Chapter 4 may represent another species within the genus *Norealidys*. All this suggests that what was known as *E. neapolitana* was probably a group of species distributed in the Mediterranean region. However, since it is a single specimen and we only had the posterior section of the body, we preferred not to describe or name this species until we have examined other complete specimens from this population.

Conclusions

The main conclusions we can draw from this work are that *Eiseniella tetraedra* possess a high genetic diversity in Europe, geographically located due to recolonization from both southern and northern refugia and that it can withstand unstable environments through changes in gene expression. Distribution of clades and lineages is influenced by environmental and soil factors such as temperature, precipitations, pH, and lithology. Furthermore, the subspecies previously described for this species have no genetic basis, so they are not considered valid from a systematic point of view in this thesis. The conclusions of each chapter are then listed:

Chapter 1: Bless this phylogeographic mess - Comparative study of *Eiseniella tetraedra* (Annelida, Oligochaeta) between an Atlantic area and a continental Mediterranean area in Spain.

- *E. tetraedra* showed a great genetic variability in the two studied areas in Spain with no clear patterns of population structure as a result of two opposite trends (monohaplotypic and polyhaplotypic populations in both areas). Thus, this species share only some general traits of moist soil fauna.
- The studied samples from the rest of the world showed limited variability, which can be biased by the number of samples.
- *E. tetraedra* showed no clear difference in the presence of lineages between Eurosiberian and Mediterranean areas.
- An apparent dispersion pattern was detected along the Guadarrama river basin, with lineages upstream and middle/low stream being different. The

middle course of the Guadarrama river basin showed the highest variability being the lower course the less diverse, apparently due to unfavorable conditions.

- No genetic basis for the morphological variation was found. The position of the male pore and thus the subspecies of *E. tetraedra* described had no molecular basis.

Chapter 2: Phylogeography of *Eiseniella tetraedra* (Savigny, 1826) in the Iberian Peninsula shows a genetic distribution based on environmental factors.

- Eight different lineages for *E. tetraedra* have been found in the Iberian Peninsula, nested in two different clades.
- Historical demography showed demographic expansion for most lineages, and recent dispersal of lineages A, E and F in the Iberian Peninsula, probably after Pleistocene glaciation.
- Genetic divergence based on the COI gene for clades suggested an underlying cryptic speciation of *E. tetraedra*.
- The distribution of lineages may respond to environmental and soil factors such as temperature, precipitation, and pH.
- Clade I appeared more in stable habitats, while Clade II was found more frequently in unstable habitats.
- Morphological variation did not correlate with our phylogenies, although differences in length and weight were found among lineages. No significant

results were obtained when studying the effects of altitudinal gradient, latitude and longitude on morphological variation.

Chapter 3: The nunatak and tabula rasa hypotheses may be compatible: the European phylogeography of a riparian earthworm.

- Eleven genetic lineages for *E. tetraedra* have been found in Europe, nested in five different clades.
- Clades I and II were widely distributed in Europe, while the others had a limited distribution. Thus, Clade I was more represented in cold biogeographical regions such as the continent, the Atlantic or even the Arctic and Clade II was prevalent in Mediterranean regions. Clade III is largely restricted to Eastern Europe and appears to be the ancestral clade.
- Macroecological preferences of clades I and II match their distribution pattern.
- Both "*tabula rasa*" and nunatak hypotheses are congruent to *E. tetraedra* phylogeography in Europe.

Chapter 4: Back to the past: Revisiting *Eiseniella neapolitana* and *Norealidys andaluciana* (Annelida, Lumbricidae) corroborates Bouché's proposal of Norealidys.

- Phylogenetic analysis, genetic divergence based on COI, and significant differences in the number of segments support the earthworm genus *Norealidys*.

- *E. neapolitana* is different, but closely related to *N. andaluciana* and was renamed *Norealidys neapolitana*. This point invalidates the idea of being a subspecies of *E. tetraedra* and its synonymy with *N. andaluciana*.
- *N. neapolitana* appeared as a cryptic complex.
- The genera *Eiseniella* and *Norealidys* appear in the phylogenetic tree in a polytomy with *Iberoscolex.*

Chapter 5: How to thrive in unstable environments: gene expression profile of a riparian earthworm under abiotic stress.

- *E. tetraedra* adults can acclimate to low temperatures due to the upregulation of genes involved in glucose accumulation but cannot acclimate to freezing conditions due to the downregulation of the respiratory chain.
- Under desiccation conditions, genes that protect earthworms from apoptosis and DNA damage were upregulated in *E. tetraedra* whereas lipid metabolism appeared to be downregulated, presumably due to resource accumulation in case the adverse conditions persisted for a prolonged period of time.
- No common responses were found between cold and desiccation treatments.

References

2001: Climate Change 2001 – The Scientific Basis. Contributions of working group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. – Cambridge University Press, Cambridge, UK and New York, NY, USA.

Alberts, B., Johnson, A., Walter, P., Lewis, J., Raff, M., Roberts, K. 2008. Molecular cell biology. New York: Garland Science.

Alcaraz, D., Paruelo, J., Cabello, J. 2006. Identification of current ecosystem functional types in the Iberian Peninsula. *Global Ecology and Biogeography*, 15(2): 200-212.

Al-Fageeh, M. B., Smales, C. M. 2006. Control and regulation of the cellular responses to cold shock: the responses in yeast and mammalian systems. *Biochemical Journal*, 397: 247-259.

Álvarez, J. 1971. Los Oligoquetos Terricolas de la Peninsula Iberica. Tesis Doctoral, Serie A. Publicaciones de la Facultad de Ciencias, Madrid.

Álvarez, A. H. 2004. Mono-adp-ribosilación: implicación en la fisiología de los organismos. *Revista de educación bioquímica*, 23: 149-156.

Anderson, C. M., Parkinson, F. E. 1997. Potential signalling roles for UTP and UDP: sources, regulation and release of uracil nucleotides. *Trends in Pharmacological Sciences*, 18: 387-392.

Anderson, F. E., Williams, B. W., Horn, K. M., Erséus, C., Halanych, K. M., Santos, S. R., James, S. W. 2017. Phylogenomic analyses of Crassiclitellata support major Northern

and Southern Hemisphere clades and a Pangaean origin for earthworms. *BMC evolutionary biology*,17(1): 1-18.

Aplin, J. D., Jones, C. J. 2012. Fucose, placental evolution and the glycocode. *Glycobiology*, 22: 470–478

Arctander, P. 1995. Comparison of a mitochondrial gene and a corresponding nuclear pseudogene. *Proceedings of the Royal Society of London B: Biological Science,* 262: 13–19.

Avise, J. C. 2000. Phylogeography: the history and formation of species. Harvard university press.

Bardgett, R. D., Van Der Putten, W. H. 2014. Belowground biodiversity and ecosystem functioning. *Nature*, 515(7528): 505-511.

Barley, K. P. 1959. Earthworms and soil fertility. IV. The influence of earthworms on the physical properties of a red-brown earth. *Australian journal of agricultural research*, 10(3): 371-376.

Barne, A. Z., Striganova, B. R. 2005. Evaluation of production parameters of earthworms *Eiseniella tetraedra* Sav. in a laboratory culture. *Biological Bulletin*, 32: 264-267.

Barrett, J. 1991. Water and fertilizer movement in greenhouse subirrigation systems. *Greenhouse Manager,* 10: 89–90.

Bates, B.C., Kundzewicz, Z.W., Wu, S., Palutikof, J.P. 2008. ‘Climate Change and Water’. Technical Paper of the Intergovernmental Panel on Climate Change. Geneva: IPCC Secretariat.

Beebee, T.J.C., Rowe, G. 2000. Microsatellite analysis of natterjack toad *Bufo calamita* Laurenti populations: consequences of dispersal from a Pleistocene refugium. *Biological Journal of the Linnean Society,* 69: 367-381.

Bensimon, A., Aebersold, R., Shiloh, Y. 2011. Beyond ATM: the protein kinase landscape of the DNA damage response. *FEBS Letters*, 585: 1625-1639.

Berman, D. I., Leirikh, A. N. 1985. The ability of earthworm *Eisenia-nordenskioldi* (Eisen)(Lumbricidae, Oligochaeta) to resist minus temperatures*. Doklady Akademii Nauk SSSR*, 285: 1258-1261.

Bertrand, M., Barot, S., Blouin, M., Whalen, J., De Oliveira, T., Roger-Estrade, J. 2015. Earthworm services for cropping systems. A review. *Agronomy for Sustainable Developmen*t, 35(2): 553-567.

Bhambri, A., Dhaunta, N., Patel, S. S., Hardikar, M., Bhatt, A., Srikakulam, N., ..., Pillai, B. 2018. Large scale changes in the transcriptome of *Eisenia fetida* during regeneration. *PLoS One*, 13: e0204234.

Blakemore, R. J. 1999. Diversity of exotic earthworms in Australia–a status report. *Proceedings of The other*, 99: 182-187.

Blakemore, R.J. 2002 . Cosmopolitan Earthworms — an Eco - Taxonomic Guide to the Peregrine Species of the World. CD - ROM. Canberra, Australia: VermEcology, p. 500.

Blakemore, R.J. 2006 . Cosmopolitan Earthworms — an Eco - Taxonomic Guide to the Peregrine Species of the World, 2nd edn. Japan: VermEcology , p. 600.

Blakemore, R. J., Chang, C. H., Chuang, S. C., Ito, M. T., James, S., Chen, J. H., 2006. Biodiversity of earthworms in Taiwan: a species checklist with the confirmation and

new records of the exotic lumbricids *Eisenia fetida* and *Eiseniella tetraedra*. *Taiwania*, 51: 226-236.

Blakemore, R. J. 2007. Checklist of megadrile earthworms from Greenland and Iceland. Yokohama: Yokohama National University, Japan.

Blakemore, R. J. 2008. An updated list of valid, invalid and synonymous names of Criodriloidea and Lumbricoidea (Annelida: Oligochaeta: Criodrilidae, Sparganophilidae, Ailoscolecidae, Hormogastridae, Lumbricidae, Lutodrilidae). A Series of Searchable Texts on Earthworm Biodiversity, Ecology and Systematics from Various Regions of the World-Supplemental. CD publication under ICZN (1999: Art. 8) also presented online http://www.annelida.net/earthworm/ Accessed October, 2012.

Blakemore, R. J. 2009. Cosmopolitan earthworms–a global and historical perspective. *Annelids in modern biology*, 257-283.

Block, W. 2003. Water or ice? —the challenge for invertebrate cold survival. *Science Progress*, 86: 77-101.

Blouin, M., Barot, S., Lavelle, P. 2006. Earthworms (Millsonia anomala,Megascolecidae) do not increase rice growth through enhanced nitrogen mineralization. *Soil Biology and Biochemistry*, 38: 2063-2068.

Blouin, M., Hodson, M.E., Delgado, E.A., Baker, G., Brussard, L., Butt, K.R., Dai, J., Dendooven, L., Peres, G., Tondoh, J.E., Cluzeau, D., Brun, J.J. 2013. A review of earthworm impact on soil function and ecosystem services. *European Journal of Soil Science,* 64: 161–182.

Bodenheimer, F.S. 1937. Prodromus faunae Palestinae. Essai sur les éléments zoogeographiques et historiques du sud-ouest du sous-règne Paléarctique. *Mémoires présentés a l'Institut d'Egypte,* 33: 1–286.

Bottinelli, N., Hedde, M., Jouquet, P., Capowiez, Y. 2020. An explicit definition of earthworm ecological categories–Marcel Bouché's triangle revisited. *Geoderma*, 372: 114361.

Bottinelli, N., Capowiez, Y. 2021. Earthworm ecological categories are not functional groups. *Biology and Fertility of Soils*, 57(2): 329-331.

Bouché, M. B. 1972. Lombriciens de France: écologie et systématique. *Annales de Zoologie Ecologie Animale*, 214-218.

Boyer, S. L., Baker, J. M., Giribet, G. 2007. Deep genetic divergences in *Aoraki denticulata* (Arachnida, Opiliones, Cyphophthalmi): a widespread 'mite harvestman'defies DNA taxonomy. *Molecular Ecology*, 16(23): 4999-5016.

Brown, C. J., O'Connor, M. I., Poloczanska, E. S., Schoeman, D. S., Buckley, L. B., Burrows, M. T., ..., Richardson, A. J. 2016. Ecological and methodological drivers of species' distribution and phenology responses to climate change. *Global change biology,* 22: 1548-1560.

Brown, G., Fragoso, C., Barois, I., Rojas, P., Patrón, J.C., Bueno, J., …, Rodríguez, C. 2001. Diversidad y rol funcional de la macrofauna edáfica en los ecosistemas tropicales mexicanos. *Acta Zoológica Mexicana*, 1: 79-110.

Brown, G. G., James, S. W., Pasini, A., Nunes, D. H., Benito, N. P., Martins, P. T., Sautter, K. D., 2006. Exotic, peregrine, and invasive earthworms in Brazil: diversity, distribution, and effects on soils and plants. *Caribbean Journal of Science*, 42: 339-359.

Brown, N. G. 1972. Changes in ribosomal patterns and a related membrane fraction during induction of cold hardiness in mimosa epicotyl tissues. *Plant Cell Physiology*, 13: 345-351.

Buckman, H.O., Brady, N.C. 1969. The nature and properties of soils. The MacMillan Company, London, UK.

Buhay, J. E., 2009. “COI-like” sequences are becoming problematic in molecular systematic and DNA barcoding studies. *Journal of Crustacean Biology*, 29: 96-110.

Butlin, R. 2002. The costs and benefits of sex: new insights from old asexual lineages. *Nature Reviews Genetics*, 3(4): 311-317.

Cai, Q., Storey, K. B. 1996. Anoxia-Induced Gene Expression in Turtle Heart: Upregulation of Mitochondrial Genes for NADH-Ubiquinone Oxidoreductase Subunit 5 and Cytochrome c Oxidase Subunit 1. *European Journal of Biochemistry*, 241: 83-92.

Cameron, J. N. 2012. Principles of Physiological Measurement. Academic Press, Orlando, USA.

Carson, H. L., Chang, L. S., Lyttle, T. W. 1982. Decay of female sexual behavior under parthenogenesis. *Science*, 218(4567): 68-70.

Casellato, S., Rodighiero, R., 1972. Karyology of Lumbricidae. III Contribution. *Caryologia*, 25: 513-524.

Casellato, S. 1987. On polyploidy in Oligochaetes with particular reference to Lumbricids. In *On Earthworms. Modena: Selected symposia and monographs UZI* (pp. 75-87).

Černosvitov L. 1938. The Oligochaeta. In Washbourn R, Jones RF (eds): Report of the Percy Sladen Expedition to Lake Huleh; a contribution to the study of fresh waters of Palestine, *Annals and Magazine of Natural History* 11: 535–550.

Černosvitov L. 1940. On some Oligochaeta from Palestine. *Annals and Magazine of Natural History,* 11: 438–447.

Černosvitov, L., 1942. Oligochaeta from various parts of the world. *Proceedings of Zoological Society of London*, 111: 197-236.

Chanabun, R., Lin, A., Jirapatrasilp, P., Bantaowong, U. 2020. The Semi-Aquatic Freshwater Earthworm Genus *Glyphidrilus* Horst, 1889 from Myanmar (Oligochaeta: Almidae) with Description of a New Species. *Tropical Natural History, 20*: 203-218.

Chang, C.H., James, S. 2011. A critique of earthworm molecular phylogenetics. *Pedobiologia,* 54: S3-S9.

Chen, L., Sun, F., Yang, X., Jin, Y., Shi, M., Wang, L., ..., Wang, Q. 2017. Correlation between RNA-Seq and microarrays results using TCGA data. *Gene,* 628: 200-204.

Chettoor, A. M., Givan, S. A., Cole, R. A., Coker, C. T., Unger-Wallace, E., Vejlupkova, Z., ..., Evans, M. M. 2014. Discovery of novel transcripts and gametophytic functions via RNA-seq analysis of maize gametophytic transcriptomes. *Genome Biology*, 15: 1-23.

Chien, S. C., Brinkmann, E. M., Teuliere, J., Garriga, G. 2013. Caenorhabditis elegans PIG-1/MELK acts in a conserved PAR-4/LKB1 polarity pathway to promote asymmetric neuroblast divisions. *Genetics,* 193: 897-909.

Choi, S., Ku, J. L. 2011. Resistance of colorectal cancer cells to radiation and 5-FU is associated with MELK expression. *Biochemical and biophysical research communications,* 412: 207-213.

Clement, M., Posada, D. C. K. A., Crandall, K. A. 2000. TCS: a computer program to estimate gene genealogies. *Molecular ecology*, 9(10): 1657-1659.

Clouse, R. M., Sharma, P. P., Stuart, J. C., Davis, L. R., Giribet, G., Boyer, S. L., Wheeler, W. C. 2015. Phylogeography of the harvestman genus *Metasiro* (Arthropoda, Arachnida, Opiliones) reveals a potential solution to the Pangean paradox. *Organisms Diversity & Evolution*, 16(1): 1-18.

Cole, C.V., Cerri, C., Minami, K., Mosier, A., Rosenberg, N., Sauerbeck, D., …, Zhao, Q. 1996. Agricultural options for mitigation of greenhouse gas emissions. In: Watson RT, Zinyowera MC, Moss RH (eds) Climate change 1995. Impacts, adaptations and mitigation of climate change: scientific technical analyses. Published for the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, pp 745–771.

Cordes, S., Frank, C. A., Garriga, G. 2006. The *C. elegans* MELK ortholog PIG-1 regulates cell size asymmetry and daughter cell fate in asymmetric neuroblast divisions. *Development*, 133: 2747-2756.

Costa, D., Timmermans, M. J., Sousa, J. P., Ribeiro, R., Roelofs, D., Van Straalen, N. M. 2013. Genetic structure of soil invertebrate populations: Collembolans, earthworms and isopods. *Applied soil ecology*, 68: 61-66.

Crawford, P. H., Hoagland, B. W. 2010. Using species distribution models to guide conservation at the state level: the endangered American burying beetle (*Nicrophorus americanus*) in Oklahoma. *Journal of Insect Conservation*, 14(5): 511-521.

Crease, T.J., Lynch, M. 1991. Ribosomal DNA variation in *Daphnia pulex*. *Molecular Biology and Evolution,* 8: 620-640.

Csuzdi, C., Pavlíček, T. 2005. Earthworms from Israel. II. Remarks on the genus *Perelia* Easton, 1983 with descriptions of a new genus and two new species. *Acta Zoologica Academiae Scientiarum Hungaricae*, 51: 75-96.

Csuzdi, C. 2012. Earthworm species, a searchable database. *Opuscula Zoologica Budapest*, 43(1): 97-99.

Curry, J. P., Boyle, K. E. 1987. Growth rates, establishment, and effects on herbage yield of introduced earthworms in grassland on reclaimed cutover peat. *Biology and Fertility of soils*, 3(1): 95-98.

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

Dawson, T. P., Berry, P. M., Kampa, E. 2003. Climate change impacts on freshwater wetland habitats. *Journal for Nature Conservation*, 11: 25-30.

De Sosa, I., Marchán, D. F., Novo, M., Cosín, D. J. D., Giribet, G., Fernández, R., 2017. Insights into the origin of parthenogenesis in oligochaetes: strong genetic structure in

a cosmopolitan earthworm is not related to reproductive mode. *European Journal of Soil Biology*, 81: 31-38.

de Sosa, I., Díaz Cosín, D.J., Csuzdi, C., Paoletti, M.G., Marchán, D.F. 2019. Placing Eophila tellinii (Oligochaeta, Lumbricidae) in a molecular phylogenetic context advances the century-old controversy around the problematic genus. *European Journal of Soil Biology,* 94: 103-114.

Decaëns, T. 2010. Macroecological patterns in soil communities. *Global Ecology and Biogeography*, 19(3): 287-302.

Diao, P., Song, Y., Ge, H., Wu, Y., Li, J., Zhang, W., ..., Cheng, J. 2019. Identification of 4-lncRNA prognostic signature in head and neck squamous cell carcinoma. *Journal of Cellular Biochemistry*, 120: 10010-10020.

Díaz Cosín, D.J., Ruiz, M.P., Ramajo, M., Gutierrez, M. 2006. Is the aestivation of the earthworm *Hormogaster elisae* a paradiapause? *Invertebrate Biology*, 125: 250-255.

Díaz Cosín, D. J., Novo, M., Fernández, R., Marchán, D. F., Gutiérrez, M. 2014. A new earthworm species within a controversial genus: *Eiseniona gerardoi* sp. n.(Annelida, Lumbricidae)-description based on morphological and molecular data. *ZooKeys*, 399: 71.

Diehl, W. J., Williams, D. L. 1992. Interactive effects of soil moisture and food on growth and aerobic metabolism in *Eisenia fetida* (Oligochaeta). *Comparative Biochemistry and Physiology - Part A: Molecular & Integrative Physiology*, 1021: 179-184.

Domínguez, J., Aira, M., Breinholt, J. W., Stojanovic, M., James, S. W., Pérez-Losada, M. 2015. Underground evolution: new roots for the old tree of lumbricid earthworms. *Molecular phylogenetics and evolution*, 83: 7-19.

Dubrovsky, M., Svoboda, M. D., Trnka, M., Hayes, M. J., Wilhite, D. A., Zalud, Z., Hlavinka, P. 2009. Application of relative drought indices in assessing climate-change impacts on drought conditions in Czechia. *Theoretical and Applied Climatology*, 96: 155-171.

Dupont, L., Grésille, Y., Richard, B., Decaëns, T., & Mathieu, J. 2015. Dispersal constraints and fine-scale spatial genetic structure in two earthworm species. *Biological Journal of the Linnean Society*, 114(2): 335-347.

Edwards, C. A., Bohlen, P. J. 1996. Biology and ecology of earthworms. 3rd edn. Springer Science & Business Media.

Edwards, C. A. 2004. The importance of earthworms as key representatives of the soil fauna. *Earthworm ecology*, 2: 3-11.

Ekblom, R., Galindo, J. 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, 107: 1-15.

Ernst, G., Felten, D., Vohland, M., Emmerling, C. 2009. Impact of ecologically different earthworm species on soil water characteristics. *European Journal of Soil Biology*, 45(3): 207-213.

Ewart, K. V., Fletcher, G. L. 1993. Herring antifreeze protein: primary structure and evidence for a C-type lectin evolutionary origin. *Molecular Marine Ecology and Biotechnology*, 2: 20-27.

Ewart, K. V., Li, Z., Yang, D. S., Fletcher, G. L., Hew, C. L. 1998. The ice-binding site of Atlantic herring antifreeze protein corresponds to the carbohydrate-binding site of C-type lectins. *Biochemistry*, 37: 4080-4085.

Excoffier, L., Lischer, H. E. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources*, 10(3): 564-567.

Fabrizio, G., Di Paola, S., Stilla, A., Giannotta, M., Ruggiero, C., Menzel, S., ...,Di Girolamo, M. 2015. ARTC1-mediated ADP-ribosylation of GRP78/BiP: a new player in endoplasmic-reticulum stress responses. *Cellular and Molecular Life Science*, 72, 1209-1225.

Fernández, R., Almodóvar, A., Novo, M., Gutiérrez, M., Díaz Cosín, D. J., 2011. A vagrant clone in a peregrine species: phylogeography, high clonal diversity and geographical distribution in the earthworm *Aporrectodea trapezoides* (Dugès, 1828). *Soil Biology and Biochemistry*, 43: 2085-2093.

Fernández, R., Bergman, P., Almodóvar, A., Díaz Cosín, D.J., Heethoff, M., 2011. Ultrastructural and molecular insights into three populations of *Aporrectodea trapezoides* (Dugés, 1828) (Oligochaeta, Lumbricidae) with different reproductive modes. *Pedobiologia*, 54: 281–290.

Fernández, R., Almodóvar, A., Novo, M., Simancas, B., Díaz Cosín, D. J. 2012. Adding complexity to the complex: new insights into the phylogeny, diversification and origin of parthenogenesis in the *Aporrectodea caliginosa* species complex (Oligochaeta, Lumbricidae). *Molecular phylogenetics and evolution*, 64(2): 368-379.

Fernández, R., Almodóvar, A., Novo, M., Gutiérrez, M., Díaz Cosín, D.J. 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.

Fernández R, Novo M, Marchán DF, Díaz Cosín DJ. 2015. Diversification patterns in cosmopolitan earthworms: similar mode but different tempo*. Molecular phylogenetics and evolution,* 94: 701-708.

Ferreiro, A. 2003. Britonia: Camiños Novos. [Reseña]

Fridolin, V. Y. 1936. Animal and plant communities of the Khibiny mountains I. *Proceedings of the Kola Station AS USSR,* 3: 19–295

Fujita, J. 1999. Cold shock response in mammalian cells. *Journal of Molecular Microbiology and Biotechnology*, 1: 243-255.

Gao, L., Han, Y., Deng, H., Hu, W., Zhen, H., Li, N., ..., Pang, Q. 2017. The role of a novel C-type lectin-like protein from planarian in innate immunity and regeneration. *Developmental & Comparative Immunology*, 67: 413-426.

Gates, G.E. 1968. On a new anthropochorous species of the earthworm genus *Pheretima* (Megascolecidae, Oligochaeta). *Journal of Natural History*, 2: 253 – 261.

Gates, G.E. 1972. Burmese earthworms, an introduction to the systematics and biology of Megadrile oligochaetes with special reference to South - East Asia. *Transactions of the American Philosophical Society*, 62: 1 – 326.

Gates, G. E. 1977. Contribution to a revision of the earthworm family Lumbricidae. XX. The genus *Eiseniella* in North America. *Megadrilogica*, 3 (5): 71-79.

Gavinelli, F., Barcaro, T., Csuzdi, C., Blakemore, R. J., Marchan, D. F., De Sosa, I., ..., Paoletti, M. G. 2018. Importance of large, deep-burrowing and anecic earthworms in forested and cultivated areas (vineyards) of northeastern Italy. *Applied Soil Ecology,* 123: 751-774.

Gavrilov, K. 1939. Sur la reproduction de Eiseniella tetraedra (Sav.) forma typica. *Acta Zoologica*, 20(2-3): 439-464.

Gómez, A., Lunt, D. H. 2007. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In *Phylogeography of southern European refugia* (pp. 155-188). Springer Netherlands.

Goto, Y., Uematsu, S., Kiyono, H. 2016. Epithelial glycosylation in gut homeostasis and inflammation. *Nature Immunology*, 17: 1244–1251.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., ..., Regev, A. 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature Biotechnology*, 29: 644-652.

Grant, W. C. 1955. Studies on moisture relationships in earthworms. *Ecology*, 36: 400-407.

Gruber, A.R., Lorenz, R., Bernhart, S.H., Neuböck, R., Hofacker, I.L. 2008. The Vienna RNA Websuite. *Nucleic Acids Research*, 36:70-74.

Gusta, L. V., Weiser, C. J. 1972. Nucleic acid and protein changes in relation to cold acclimation and freezing injury of Korean Boxwood leaves. *Plant Physiology*, 49: 91-96.

Gutniak, M., Orskov, C., Holst, J.J., Ahren, B.B., Efendic, S. 1992. Antidiabetogenic effect of glucagon-like peptide-1(7–36) amide in normal subjects and patients with mellitus. *The New England Journal of Medicine*, 326, 1316–1322.

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

Hansen, P. L., Holmstrup, M., Bayley, M., Simonsen, V. 2006. Low genetic variation for *Dendrobaena octaedra* from Greenland compared to populations from Europe and North America: Refuge or selection? *Pedobiologia*, 50(3): 225-234.

Haraldsen, T. K., Engelstad, F. 1998. Influence of earthworms on soil properties and crop production in Norway. – Centre for Soil and Environmental Research, Oslo, 12 pp

Harpending, H. C., 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human biology*, 66: 591-600.

Hartmann-Petersen, R., Semple, C. A., Ponting, C. P., Hendil, K. B., Gordon, C. 2003. UBA domain containing proteins in fission yeast. *The International Journal of Biochemistry & Cell Biology*, 35: 629-636.

Hatefi, Y. 1985. The mitochondrial electron transport and oxidative phosphorylation system. *Annual Review of Biochemistry*, 54: 1015-1069.

Hattab, S., Boughattas, I., Mkhinini, M., Banni, M. 2020. Impact of intensive farming on soil heavy metal accumulation and biomarkers responses of earthworms *Eisenia andrei. Bulletin of Environmental Contamination and Toxicology*, 105(4): 559-564.

Heethoff, M., Domes, K., Laumann, M., Maraun, M., Norton, R. A., Scheu, S. 2007. High genetic divergences indicate ancient separation of parthenogenetic lineages of the

oribatid mite *Platynothrus peltifer* (Acari, Oribatida). *Journal of evolutionary biology*, 20(1): 392-402.

Hewitt, G. M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological journal of the Linnean Society*, 58(3): 247-276.

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

Hewitt, G.M. 2001. Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Molecular Ecology,* 10: 537-549.

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

Hillis, D.M., Dixon, M.T. 1991. Ribosomal DNA: molecular evolution and phylogenetic inference, *The Quarterly review of biology,* 66: 411-453.

Hoffmann, A. A., Parsons, P. A. 1993. Direct and correlated responses to selection for desiccation resistance: a comparison of *Drosophila melanogaster* and *D. simulans*. *Journal of Evolutionary Biology*, 6: 643-657.

Hole, F.D. 1981. Effects of animals on soil. *Geoderma*, 25: 75-112.

Holland, E.A., Braswell, B.H., Lamarque, J.F., Townsend, A., Suylzman, J., Mueller, J.F., …, Roelofs, G. 1997. Variations in the predicted spatial distribution of atmospheric nitrogen deposition and their impact on carbon uptake by terrestrial ecosystems. *Journal of Geophysical Research: Atmospheres*, 102(D13): 15849-15866

Holmstrup, M., Ostergaard, I.K., Nielsen, A., Hansen, B.T. 1991. The relationship between temperature and cocoon incubation time for some lumbricid earthworm species. *Pedobiologia*, 35: 179-184.

Holmstrup, M., Zachariassen, K. E. 1996. Physiology of cold hardiness in earthworms. *Comparative Biochemistry and Physiology – Part A: Molecular & Integrative Physiology*, 115: 91-101.

Holmstrup, M., Petersen, B. F. 1997. Freeze-tolerance in the subarctic earthworm *Eisenia nordenskioeldi* (Eisen). *Cryo-letters,* 18: 153-156.

Holmstrup, M., Costanzo, J. P., Lee Jr, R. E. 1999. Cryoprotective and osmotic responses to cold acclimation and freezing in freeze-tolerant and freeze-intolerant earthworms. *Comparative Biochemistry and Physiology – Part B: Biochemistry & Molecular Biology*, 169: 207-214.

Holmstrup, M., Loeschcke, V. 2003 Genetic variation in desiccation tolerance of *Dendrobaena octaedra* cocoons originating from different climatic regions*. Soil Biology and Biochemistry*, 35: 119-124.

Holmstrup, M. 2003. Overwintering adaptations in earthworms: The 7th international symposium on earthworm ecology Cardiff, Wales, 2002. *Pedobiologia,* 47: 504-510.

Holmstrup, M. 2014. The ins and outs of water dynamics in cold tolerant soil invertebrates. *Journal of Thermal Biology*, 45: 117-123.

Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., Van der Linden, P. J., Dai, X., Maskell, K., and Johnson, C. A.: Climate Change 2001, The Scientific Basis, Cambridge University Press, Cambridge, UK.

Huang, J., Xu, Q., Sun, Z.J., Tang, G.L., Su, Z.Y. 2007. Identifying earthworms through DNA barcodes. *Pedobiologia, 51*: 301-309.

Hugall, A., Stanton J, Moritz C. 1999. Reticulate evolution and the origins of ribosomal internal transcribed spacer diversity in apomictic *Meloidogyne*. *Molecular Biology and Evolution*, 16: 157-164.

Huigens, M. E., Luck, R. F., Klaassen, R. H. G., Maas, M. F. P. M., Timmermans, M. J. T. N., Stouthamer, R. 2000. Infectious parthenogenesis. *Nature*, 405(6783), 178-179.

Hunter, T. 1995. Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell*, 80: 225-236.

Hurov, K. E., Cotta-Ramusino, C., Elledge, S. J. 2010. A genetic screen identifies the Triple T complex required for DNA damage signaling and ATM and ATR stability. *Genes & Development*, 24: 1939-1950.

Intergovernmental Panel on Climate Change (IPCC) *Confidence and Likelihood in the IPCC Fifth Assessment Report.* IPCC; Geneva, Switzerland: 2013.

Intergovernmental Panel on Climate Change (IPCC). 1992. Climate Change 1992. The Suplementary Report to the IPCC Scientific Assessment. Cambridge University Press. Cambridge, UK.

Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) *The Assessment Report on Land Degradation and Restoration.* 2018

Jiménez, J.J., Brown, G.G., Decaëns, T., Feijoo, A., Lavelle, P. 2000. Differences in the timing of diapause and patterns of aestivation in tropical earthworms. *Pedobiologia,* 44: 677-694.

Jaarola, M., Searle, J. B. 2002. Phylogeography of field voles (*Microtus agrestis*) in Eurasia inferred from mitochondrial DNA sequences. *Molecular ecology*, 11(12): 2613-2621.

Jaenicke, J., Selander, K., 1979. Evolution and ecology of parthenogenesis in earthworms. *American Zoologist*, 19: 729-737.

Jalili, A., Jamzad, Z., Thompson, K., Araghi, M. K., Ashrafi, S., Hasaninejad, M., ..., Parvaneh, K. 2010. Climate change, unpredictable cold waves and possible brakes on plant migration. *Global Ecology and Biogeography*, 19: 642-648.

James, S. W., Davidson, S. K. 2012. Molecular phylogeny of earthworms (Annelida: Crassiclitellata) based on 28S, 18S and 16S gene sequences. *Invertebrate Systematics*, 26(2): 213-229.

Jamieson, B. G. M. 1988. On the phylogeny and higher classification of the Oligochaeta. *Cladistics*, *4*(4): 367-401.

Jamieson, B. G., Ferraguti, M. 2006. Non-leech clitellata. *Reproductive biology and phylogeny of Annelida*, *4*: 235-392.

Javidkar, M., Abdoli, A., Ahmadzadeh, F., Nahavandi, Z., Yari, M., 2021. Molecular evidence reveals introduced populations of *Eiseniella tetraedra* (Savigny, 1826)(Annelida, Lumbricidae) with European origins from protected freshwater ecosystems of the southern Alborz Mountains. *Marine and Freshwater Research*, 72: 44-57.

Ji, K., Wang, Y., Sun, W., Lou, Q., Mei, H., Shen, S., Chen, H. 2012. Drought-responsive mechanisms in rice genotypes with contrasting drought tolerance during reproductive stage. *Journal of Plant Physiology*, 169: 336-344.

Jiménez Pinadero, S., Marchán, D.F.,Novo, M., Trigo, D., Domínguez, J., Díaz Cosín, D.J. 2021. Sorry atlanticus, you are not my type: molecular assessment splits *Zophoscolex* (Lumbricidae: Crassiclitellata) into French and Iberian genera, *Zoological Journal of the Linnean Society*, zlab011.

Jin, H., Wan, Y. W., Liu, Z. 2017. Comprehensive evaluation of RNA-seq quantification methods for linearity. *BMC bioinformatics*, 18: 51-59.

Jirapatrasilp, P., Prasankok, P., Chanabun, R., & Panha, S. 2015. Allozyme data reveal genetic diversity and isolation by distance in sympatric *Glyphidrilus* Horst, 1889 (Oligochaeta: Almidae) of the Lower Mekong River Basin. *Biochemical Systematics and Ecology*, 61: 35-43.

Johnston, A. S., Sibly, R. M., Hodson, M. E., Alvarez, T., Thorbek, P. 2015. Effects of agricultural management practices on earthworm populations and crop yield: validation and application of a mechanistic modelling approach. *Journal of Applied Ecology,* 52(5): 1334-1342.

Jönsson, K. I., Schill, R. O. 2007. Induction of Hsp70 by desiccation, ionising radiation and heat-shock in the eutardigrade Richtersius coronifer. *Comparative Biochemistry & Physiology – Part B: Biochemistry and Molecular Biology*, 146: 456-460.

Jouque,P., Blancart, E., Capowiez, Y. 2014. Utilization of earthworms and termites for the restoration ecosystem functioning. *Appiled Soil Ecology*, 73: 34–40.

Julin, E. (1950). De svenska daggmaskarterna (pp. 58).

Jung, H., Seong, H. A., Ha, H. 2008. Murine protein serine/threonine kinase 38 activates apoptosis signal-regulating kinase 1 via Thr838 phosphorylation. *Journal of Biological Chemistry*, 283: 34541-34553.

Kashuba, E., Kashuba, V., Sandalova, T., Klein, G., Szekely, L. 2000. Epstein-Barr virus encoded nuclear protein EBNA-3 binds a novel human uridine kinase/uracil phosphoribosyltransferase. *Oncogene*, 3: 1-12.

Katoh, K., Standley, D. M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4): 772-780.

Keller, I., Chintauan-Marquier, I.C., Veltsos, P., Nichols, R.A. 2006. Ribosomal DNA in the grasshopper *Podisma pedestris*: escape from concerted evolution. *Genetics,* 174: 863-874.

Khan, A. A., Heit, C. E., Lippolo, P. C. 1968. Increase in nucleic acid synthesizing capacity during cold treatment of dormant pear embryos. *Biochemical and Biophysical Research Communications*, 33: 391-396.

Kim, H., Zhang, H., Meng, D., Russell, G., Lee, J. N., Ye, J. 2013. UAS domain of Ubxd8 and FAF1 polymerizes upon interaction with long-chain unsaturated fatty acids. *Journal of Lipid Research*, 54: 2144-2152.

Kim, K. W., Ryu, J. S., Ko, J. H., Kim, J. Y., Kim, H. J., Lee, H. J., ..., Oh, J. Y. 2020. FUT1 deficiency elicits immune dysregulation and corneal opacity in steady state and under stress. *Cell Death & Disease*, 11: 1-11.

King, R. A., Tibble, A. L., Symondson, W. O., 2008. Opening a can of worms: unprecedented sympatric cryptic diversity within British lumbricid earthworms. *Molecular ecology*, 17: 4684-4698.

Knott, K. E., Haimi, J., 2010. High mitochondrial DNA sequence diversity in the parthenogenetic earthworm *Dendrobaena octaedra*. *Heredity*, 105: 341-347.

Kodra, E., Steinhaeuser, K., Ganguly, A. R. 2011. Persisting cold extremes under 21st-century warming scenarios. *Geophysical research letters*, 38(8).

Kranner, I., Beckett, R. P., Wornik, S., Zorn, M., Pfeifhofer, H. W. 2002. Revival of a resurrection plant correlates with its antioxidant status. *The Plant Journal*, 31: 13-24.

Kretzschmar, A., Bruchou, C. 1991. Weight response to the soil water potential of the earthworm *Aporrectodea longa*. *Biology and fertility of soils*, 12: 209-212.

Kreyling, J. 2010. Winter climate change: a critical factor for temperate vegetation performance. *Ecology,* 91: 1939-1948.

Lal, R. 1988. Effects of macrofauna on soil properties in tropical ecosystems. *Agriculture, Ecosystems and Environment,* 24: 101-116.

Langmead, B., Trapnell, C., Pop, M., Salzberg, S. L. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, 10, 1-10.

Langmead, B., Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods*, 9: 357-359.

Lavelle, P., Bignell, D., Lepage, M., Wolters, V., Roger, P. A., Ineson, P. O. W. H., ..., Dhillion, S. 1997. Soil function in a changing world: the role of invertebrate ecosystem engineers. *European Journal of soil biology*, 33(4): 159-193.

Le Page, Y., Chartrain, I., Badouel, C., Tassan, J. P. 2011. A functional analysis of MELK in cell division reveals a transition in the mode of cytokinesis during *Xenopus* development. *Journal of Cell Science*, 124: 958-968.

Lee, K.E. 1983. Soil animals and pedological processes. In 'Soils: an Australian Viewpoint'. (Division of Soils, CSIRO.) pp. 629-44. (CSIRO, Melbourne: Academic Press, London).

Lee, K.E., and Ladd, J.N. 1984. Some recent advances in soil biology and biochemistry. Proc. National Soils Conference, Australian Society of Soil Science, Brisbane 1984, pp. 83-103.

Lee, K.E. 1985 . Earthworms—Their Ecology and Relationships with Soils and Land Use. Sydney, Australia: Academic Press , p. 411 .

Lee , K.E. 1987 . Peregrine species of earthworms . In On Earthworms , edited by A.M. Bonvicini Pagliai and P. Omodeo . Modena, Italy : U.Z.I., 2, Mucchi , pp. 315 – 327.

Lee, K. E., Foster, R. C. 1991. Soil fauna and soil structure. *Soil Research*, 29(6): 745-775.

Lee, J. N., Kim, H., Yao, H., Chen, Y., Weng, K., Ye, J. 2010. Identification of Ubxd8 protein as a sensor for unsaturated fatty acids and regulator of triglyceride synthesis. *Proceedings of the National Academy of Science*, 107: 21424-21429.

Leigh, J. W., Bryant, D., 2015. Popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6: 1110–1116.

Li, B., Dewey, C. N. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics*, 12: 1-16.

Li, D., Nie, H., Jahan, K., Yan, X. 2020. Expression analyses of C-type lectins (CTLs) in Manila clam under cold stress provide insights for its potential function in cold resistance of *Ruditapes philippinarum*. *Comparative Biochemistry and Physiology – Part C: Toxicology & Pharmacology*, 230: 108708.

Li, J., Hsu, H. C., Mountz, J. D., Allen, J. G. 2018. Unmasking fucosylation: from cell adhesion to immune system regulation and diseases. *Cell Chemical Biology*, 25: 499-512.

Li, N., Zhou, J., Wang, H., Mu, C., Shi, C., Liu, L., Wang, C. 2020. Transcriptome analysis of genes and pathways associated with metabolism in *Scylla paramamosain* under different light intensities during indoor overwintering. *BMC genomics*, 21: 1-15.

Librado, P., Rozas, J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25(11): 1451-1452.

Lin, M. L., Park, J. H., Nishidate, T., Nakamura, Y., Katagiri, T. 2007. Involvement of maternal embryonic leucine zipper kinase (MELK) in mammary carcinogenesis through interaction with Bcl-G, a pro-apoptotic member of the Bcl-2 family. *Breast Cancer Research*, 9: 1-13.

Lindquist, S. 1986. The heat-shock response. *Annual Review of Biochemistry*, 55: 1151-1191.

Lorenzo-Carballa, M. O., Cordero-Rivera, A. 2009. Thelytokous parthenogenesis in the damselfly *Ischnura hastata* (Odonata, Coenagrionidae): genetic mechanisms and lack of bacterial infection. *Heredity*, 103(5): 377-384.

Lu, X., Zhou, X., Cao, Y., Zhou, M., McNeil, D., Liang, S., Yang, C. 2017. RNA-seq Analysis of Cold and Drought Responsive Transcriptomes of *Zea mays* ssp. *mexicana* L. *Frontiers in Plant Science*, 8: 136.

Maldague, M.E. 1970. RBle des animaux kdaphiques dans la fertilité des sols forestiers. PubL Inst. National Etude Agronomique Congo, SCr. Sci. 112, 1-245.

Malevich, I. I. 1956. Oligohetes of the Moscow Region. Uch. Zap. MGPI im. *VP Potemkina*, 61: 403-437.

Mao, P., Hever, M. P., Niemaszyk, L. M., Haghkerdar, J. M., Yanco, E. G., Desai, D., ..., Spinella, M. J. 2011. Serine/threonine kinase 17A is a novel p53 target gene and modulator of cisplatin toxicity and reactive oxygen species in testicular cancer cells. *Journal of Biological Chemistry*, 286: 19381-19391.

Marchán, D. F., Refoyo, P., Novo, M., Fernández, R., Trigo, D., Díaz Cosín, D. J. 2015. Predicting soil micro-variables and the distribution of an endogeic earthworm species through a model based on large-scale variables. *Soil Biology and Biochemistry*, 81: 124-127.

Marchán, D.F., Novo, M., Fernández, R., De Sosa, I., Trigo, D., Díaz Cosín, D.J. 2016. Evaluating evolutionary pressures and phylogenetic signal in earthworms: a case study-the number of typhlosole lamellae in Hormogastridae (Annelida, Oligochaeta). *Zoological Journal of the Linnean Society,* 178: 4-14.

Marchán, D. F., Refoyo, P., Fernández, R., Novo, M., de Sosa, I., Díaz Cosín, D. J. 2016. Macroecological inferences on soil fauna through comparative niche modeling: the case of Hormogastridae (Annelida, Oligochaeta). *European Journal of Soil Biology*, 75: 115-122.

Marchán, D.F., Fernández, R., de Sosa, I., Sánchez, N., Díaz Cosín, D.J., Novo, M. 2018. Integrative systematic revision of a Mediterranean earthworm family: Hormogastridae (Annelida, Oligochaeta) *Invertebrate Systematics*, 32: 652-671.

Marek, P. E., Shear, W. A., Bond, J. E. 2012. A redescription of the leggiest animal, the millipede *Illacme plenipes*, with notes on its natural history and biogeography (Diplopoda, Siphonophorida, Siphonorhinidae). *ZooKeys*, (241): 77.

Marquez, L.M., MacKenzie, J.B., Takabayashi, M., Smith, C.R., Chen, C.A. 2003. Difficulties in obtaining microsatellites from acroporid corals. *Proceedings of the 9th International Coral Reef Symposium,* 1: 139-143.

Marron, M. T., Markow, T. A., Kain, K. J., Gibbs, A. G. 2003. Effects of starvation and desiccation on energy metabolism in desert and mesic *Drosophila*. *Journal of Insect Physiology*, 49: 261-270.

Mathieu, J., Jonathan Davies, T. 2014. Glaciation as an historical filter of below-ground biodiversity. *Journal of biogeography*, 41(6): 1204-1214.

Maxwell, J.T., Harley, G.L., Robeson, S.M. 2016. On the declining relationship between tree growth and climate in the Midwest United States: the fading drought signal. *Climatic Change,* 138: 127–142.

Maynard, D. S., Crowther, T. W., King, J. R., Warren, R. J., Bradford, M. A. 2015. Temperate forest termites: ecology, biogeography, and ecosystem impacts. *Ecological Entomology*, 40(3): 199-210.

Mazur, P. 1963. Kinetics of water loss from cells at subzero temperatures and the likelihood of intracellular freezing. *Journal of General Physiology*, 47: 347-369.

McDaniel, J.P., Stromberger, M.E., Barbarick, K.A., Cranshaw, W. 2013 Survival of *Aporrectodea caliginosa* and its effects on nutrient availability in biosolids amended soil. *Applied soil ecology*, 71: 1-6

Michaelsen, W. 1900. Oligochaeta. In: Das Tierreich. Friedländer & Sohn, Berlin.

Michaelsen, W. 1910. Zur Kenntnis der Lumbriciden und ihrer Verbreitung. *Annuarie du Musée Zoologique de l'Academie des Sciences de St. Pétersbourg*, 15: 1-74.

Michaelsen, W. 1932. Variations und Mutationsverhältnisse bei den Arten der Lumbricidengattung *Eiseniella*. *Zeitschrift für Naturwissenschaften Jena*, 67: 141-157.

Miles, H. B. 1963. Heat-death temperature in *Allolobophora terrestris* (Sav.) forma longa (Ude) and *Eisenia foetida* (Sav.). *Nature*, 199: 826-826.

Miller, M. A., Pfeiffer, W., Schwartz, T. 2010- Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Gateway Computing Environments Workshop (GCE), 2010* (pp. 1-8). IEEE.

Morariu, V.I., Srinivasan, B.V., Raykar, V.C., Duraiswami, R., Davis, L.S. 2008. Automatic online tuning for fast Gaussian summation *Advances in Neural lnformation Processing Systems,* 21: 1113-1120.

Mosier, A. R. 1998. Soil processes and global change. *Biology and fertility of soils*, 27(3): 221-229.

Muldal, S. 1952. The chromosomes of the earthworms. I. The evolution of polyploidy. *Heredity*, 6(1): 55-76.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., Kent, J., 2000. Biodiversity hotspots for conservation priorities. *Nature*, 403: 853-858.

Nakano, I., Paucar, A. A., Bajpai, R., Dougherty, J. D., Zewail, A., Kelly, T. K., ..., Kornblum, H. I. 2005. Maternal embryonic leucine zipper kinase (MELK) regulates multipotent neural progenitor proliferation. *Journal of Cell Biology*, 170: 413-427.

Niesler, C. U., Myburgh, K. H., Moore, F. 2007. The changing AMPK expression profile in differentiating mouse skeletal muscle myoblast cells helps confer increasing resistance to apoptosis. *Experimental Physiology*, 92: 207-217.

Nisa, M. U., Huang, Y., Benhamed, M., Raynaud, C. 2019. The plant DNA damage response: signaling pathways leading to growth inhibition and putative role in response to stress conditions. *Frontiers in Plant Science*, 10: 653.

Nota, B., Van Straalen, N. M., Ylstra, B., Roelofs, D. 2010. Gene expression microarray analysis of heat stress in the soil invertebrate *Folsomia candida*. *Insect Molecular Biology*, 19: 315-322.

Novo, M. 2010. Algunos aspectos de la Biología Evolutiva de la Familia Hormogastridae Michaelsen, 1900: Filogenia, Filogeografía y Selección Sexual. Tesis de Doctorado, Universidad Complutense de Madrid, Madrid, España. 312 pp.

Novo, M., Almodovar, A., Fernandez, R., Giribet, G., Díaz Cosín, D.J. 2011. Understanding the biogeography of a group of earthworms in the Mediterranean basin: the phylogenetic puzzle of Hormogastridae (Clitellata: Oligochaeta). *Molecular Phylogenetics and Evolution*, 61: 125-135.

Novo, M., Cunha, L., Maceda-Veiga, A., Talavera, J. A., Hodson, M. E., Spurgeon, D., ..., Kille, P. (2015). Multiple introductions and environmental factors affecting the establishment of invasive species on a volcanic island. *Soil Biology and Biochemistry*, 85: 89-100.

Omodeo, P., 1951. Corredo cromosomico e spermatogenesi aberrante in *Allolobophora caliginosa trapezoides. Estratto dal Bolletino di Zoologia*, 18: 27-39.

Omodeo P. 1952. Cariologia dei Lumbricidae. *Caryologia* 4: 173-275.

Omodeo, P., 1956. Contributo alla revisione dei Lumbricidae. *Archivio zoologico italiano: pubblicato sotto gli auspicii della unione zoologica*, 41: 129-212.

Omodeo, P., Rota, E. 1989. Earthworms of Turkey. *Italian Journal of Zoology*, 56(2): 167-198.

Omodeo P, Rota E. 1991. Earthworms of Turkey. II, *Italian Journal of Zoology,* 58: 171-181.

Örley L. 1885. *A palaearktikus övben élő Terrikolák-nak revisiója és elterjedése*. *Magyar Tudományos Akadémia,* 15-18.

Osioma, E., Hamilton-Amachree, A. 2019. Heavy metal accumulation and biomarker responses in the earthworm *Lumbricus terrestris* collected from Kolo Creek, Bayelsa State, Nigeria. *FUW Trends in Science & Technology Journal,* 4: 319-323.

Paul, S., Heckmann, L. H., Sørensen, J. G., Holmstrup, M., Arumugaperumal, A., Sivasubramaniam, S. 2018. Transcriptome sequencing, de novo assembly and annotation of the freeze tolerant earthworm, *Dendrobaena octaedra*. *Gene Reports*, 13: 180-191.

Pauli, A., Valen, E., Lin, M. F., Garber, M., Vastenhouw, N. L., Levin, J. Z., ..., Schier, A. F. 2012. Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome Research*, 22: 577-591.

Pavlícek, T., Csuzdi, C., Nevo, E. 2003. Species richness and zoogeographic affinities of earthworms in the Levant: The 7th international symposium on earthworm ecology· Cardiff, Wales, 2002. *Pedobiologia*, 47: 452-457.

Pereira, T.J., Baldwin, J.G. 2016. Contrasting evolutionary patterns of 28S and ITS rRNA genes reveal high intragenomic variation in *Cephalenchus* (Nematoda): implications for species delimitation. *Molecular Phylogenetics and Evolution,* 98: 244-260.

Perel, T.S. 1967. Earthworms of the relict forests of Western Transcaucasia and Talysh region. (Russian), *Pedobiologia* 7: 93-120.

Pérez-Losada, M., Breinholt, J.W., Porto, P.G., Aira, M., Dominguez, J. 2011. An earthworm riddle: systematics and phylogeography of the Spanish lumbricid *Postandrilus*. *PLoS One*, 6: e28153.

Pérez-Losada, M., Breinholt, J.W., Aira, M., Domínguez, J. 2015. An updated multilocus phylogeny of the Lumbricidae (Annelida: Clitellata: Oligochaeta) earthworms. *Journal of Phylogenetics and Evolutionary Biology,* 3: 140.

Petersen, C.R., Holmstrup, M., Malmendal, A., Bayley, M., Overgaard, J. 2008. Slow desiccation improves dehydration tolerance and accumulation of compatible osmolytes in earthworm cocoons (*Dendrobaena octaedra* Savigny). *Journal of Experimental Biology*, 211: 1903-1910.

Phillips, H. R. P., Guerra, C. A., Bartz, M. L. C., Briones, J. I., Brown, G., Crowther, T. W., ..., Eisenhauer, N. 2020. Global distribution of earthworm diversity. *Science*, 366: 480.

Plisko, J. D., 1965. Materialy do rozmieszczenia geograficznego i ekologii dzdzownia w Polsce (Oligochaeta, Lumbricida). *Fragmenta Faunistica*, 12: 57-108.

Plisko, J. D. 2002. Three new earthworm species of *Microchaetus* Rapp, 1849, and new data on two earlier known species of this genus (Oligochaeta: Microchaetidae). *African Invertebrates*, 43: 205-214.

Pop, V., 1952. Revizuirea sistematica a genului de Lumbricidae Eiseniella. *Anelele Academiei Republicii Populare Române*, 3: 170-186.

Pop, A. A., Wink, M., Pop, V. V. 2003. Use of 18S, 16S rDNA and cytochrome c oxidase sequences in earthworm taxonomy (Oligochaeta, Lumbricidae): The 7th international symposium on earthworm ecology· Cardiff, Wales, 2002. *Pedobiologia*, 47: 428-433.

Porco, D., Decaëns, T., Deharveng, L., James, S. W., Skarżyński, D., Erséus, C., ... , Hebert, P. D., 2013. Biological invasions in soil: DNA barcoding as a monitoring tool in a multiple taxa survey targeting European earthworms and springtails in North America. *Biological Invasions*, 15: 899-910.

Provan, J., Bennett, K. D. 2008. Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution*, 23(10): 564-571.

Qiu, J.P., Bouché, M.B, 1998. Contribution à la taxonomie des Eiseniellini (Oligochaeta: Lumbricidae). *Reynoldsia* gen. nov. et deux nouveaux *Eiseniella*. Institut National de la Recherche Agronomique (ed.), *Documents pédozoologiques et intégrologiques*, 4: 102-108.

Razafimandimbison, S.G., Kellogg, E.A., Bremer, B. 2004. Recent origin and phylogenetic utility of divergent ITS putative pseudogenes: a case study from *Naucleeae* (Rubiaceae). *Systematic Biology*, 177-192.

Rebl, A., Verleih, M., Korytář, T., Kühn, C., Wimmers, K., Köllner, B., Goldammer, T. 2012. Identification of differentially expressed protective genes in liver of two rainbow trout strains. *Veterinay Immunology and Immunopathology*, 145: 305-315.

Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-Lopes, C., Rostamianfar, A., ..., Bader, G. D. 2019. Pathway enrichment analysis and visualization of omics data using g: Profiler, GSEA, Cytoscape and EnrichmentMap. *Nature Protocols*, 14: 482-517.

Reinecke, A. J., Venter, J. M. 1987. Moisture preferences, growth and reproduction of the compost worm *Eisenia fetida* (Oligochaeta). *Biology and Fertility Soils*, 3: 135-141.

Roberts, A., Pimentel, H., Trapnell, C., Pachter, L. 2011. Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics*, 27: 2325-2329.

Robinson, D.M., McCarthy, D.J. Smyth, G.K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1): 139–140.

Rodriguez, T., Trigo, D., Díaz Cosín, D.J. 1997. Biogeographical zonation of the western Iberian peninsula on the basis of the distribution of earthworm species. *Journal of Biogeography*, 24(6): 893-901.

Rogers, A. R., 1995. Genetic evidence for a Pleistocene population explosion. *Evolution*, 49: 608-615.

Ronquist, F., Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19(12): 1572-1574.

Rota, E., Martinsson, S., Bartoli, M., Beylich, A., Graefe, U., Laini, A., Wetzel, M.J., Erséus, C. 2016. Mitochondrial evidence supports a Nearctic origin for the spreading limicolous earthworm *Sparganophilus tamesis* Benham, 1892 (Clitellata, Sparganophilidae). *Contributions to Zoology*, 85 (1): 113-119.

Rozas, J., Ferrer-Mata, A., Sánchez-Del Barrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sánchez-Gracia, A. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution,* 34: 3299-3302.

Rutgers, M., Schouten, A. J., Bloem, J., Van Eekeren, N., De Goede, R. G. M., Jagersop Akkerhuis, G. A. J. M., ... , Breure, A. M., 2009. Biological measurements in a nationwide soil monitoring network. *European Journal of Soil Science*, 60: 820-832.

Rutgers, M., Orgiazzi, A., Gardi, C., Römbke, J., Jänsch, S., Keith, A. M., ... ,De Zwart, D., 2016. Mapping earthworm communities in Europe. *Applied Soil Ecology*, 97: 98-111.

Phillips, S. J., Anderson, R. P., Schapire, R. E. 2006. Maximum entropy modeling of species geographic distributions. *Ecological modelling*, 190(3-4): 231-259.

Sala, O. E., Chapin, F. S., Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., ..., Wall, D. H. 2000. Global biodiversity scenarios for the year 2100. *Science*, 287(5459): 1770-1774.

Sanjo, H., Kawai, T., Akira, S. 1998. DRAKs, novel serine/threonine kinases related to death-associated protein kinase that trigger apoptosis. *Journal of Biological Chemistry*, 273: 29066-29071.

Sarhan, F., D'aoust, M. J. 1975. RNA synthesis in spring and winter wheat during cold acclimation. *Physiologia Plantarum*, 35: 62-65.

Schenk, P. W., Snaar-Jagalska, B. E. 1999. Signal perception and transduction: the role of protein kinases. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1449(1): 1-24.

Schiene-Fischer, C., Yu, C. 2001. Receptor accessory folding helper enzymes: the functional role of peptidyl prolyl cis/trans isomerases. *FEBS Letters*, 495: 1-6.

Schmitt, S., Pouteau, R., Justeau, D., De Boissieu, F., Birnbaum, P. 2017. ssdm: An r package to predict distribution of species richness and composition based on stacked species distribution models. *Methods in Ecology and Evolution*, 8(12): 1795-1803.

Schneider, S., Excoffier, L., 1999. Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, 152: 1079-1089.

Schuberth, C., Richly, H., Rumpf, S., Buchberger, A. 2004. Shp1 and Ubx2 are adaptors of Cdc48 involved in ubiquitin-dependent protein degradation. *EMBO reports*, 5: 818-824.

Seong, H. A., Gil, M., Kim, K. T., Kim, S. J., Ha, H. 2002. Phosphorylation of a novel zinc-finger-like protein, ZPR9, by murine protein serine/threonine kinase 38 (MPK38). *Biochemical Journal*, 361: 597-604.

Seppey M., Manni M., Zdobnov E.M. 2019. BUSCO: Assessing Genome Assembly and Annotation Completeness. In: Kollmar M. (eds) Gene Prediction. *Methods in Molecular Biology*, vol 1962. Humana, New York, NY.

Sheehan, C., Kirwan, L., Connolly, J., Bolger, T. 2008. The effects of earthworm functional diversity on microbial biomass and the microbial community level physiological profile of soils. *European journal of soil biology*, 44(1), 65-70.

Shekhovtsov, S. V., Berman, D. I., Bulakhova, N. A., Vinokurov, N. N., Peltek, S. E. 2018. Phylogeography of *Eisenia nordenskioldi nordenskioldi* (Lumbricidae, Oligochaeta) from the north of Asia. *Polar Biology*, 41(2): 237-247.

Shekhovtsov, S. V., Rapoport, I. B., Poluboyarova, T. V., Geraskina, A. P., Golovanova, E. V., Peltek, S. E. 2020. Morphotypes and genetic diversity of *Dendrobaena schmidti* (Lumbricidae, Annelida). *Vavilov Journal of Genetics and Breeding*, 24(1): 48.

Simon, J. C., Delmotte, F., Rispe, C., Crease, T. 2003. Phylogenetic relationships between parthenogens and their sexual relatives: the possible routes to parthenogenesis in animals. *Biological Journal of the Linnean Society*, 79(1): 151-163.

SI-Moussi, S. (2010). Apports de la fouille de données à la modélisation des communautés écologiques.

Simpson, G. G., 1964. Species density of North American recent mammals. *Systematic Zoology*, 13: 57-73.

Sims , R.W. , Gerard , B.M. 1985 . Earthworms. Keys and Notes to the Identification and Study of the Species. Synopsis of the British Fauna (New Series), vol. 31. Leiden: EJ Brill , p. 171 .

Sinclair, B. J., Ferguson, L. V., Salehipour-Shirazi, G., MacMillan, H. A. 2013. Cross-tolerance and cross-talk in the cold: relating low temperatures to desiccation and immune stress in insects. *Integrative and Comparative Biology*, 53: 545-556.

Skubała, P. 2013. Biodiversity and ecosystem services in soil under threat. *Journal of Pollution Effects & Control,* 1: e101.

Sommer, R. S., Zachos, F. E. 2009. Fossil evidence and phylogeography of temperate species:'glacial refugia'and post-glacial recolonization. *Journal of Biogeography,* 36: 2013-2020.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22(21): 2688-2690.

Stenberg, P., Terhivuo, J., Lokki, J., Saura, A. 1997. Clone diversity of tetraploid Otiorhynchus scaber in northern Europe. *Hereditas*, 126(2): 169-172.

Stenberg, P., Terhivuo, J., Lokki, J., Saura, A. 2000. Clone diversity in the polyploid weevil *Otiorhynchus scaber*. *Hereditas*, 132(2), 137-142.

Stephenson J. 1924. On some Indian Oligochaeta, with a description of two new genera of Ocnerodrilinae. *Records of the Indian Museum,* 26: 317–365.

Stoelting, R. E., Measey, G. J., Drewes, R. C. 2014. Population genetics of the São Tomé caecilian (Gymnophiona: Dermophiidae: Schistometopum thomense) reveals strong geographic structuring. *PloS one*, 9(8): e104628.

Stöp-Bowitz, C. 1969. A contribution to our knowledge of the systematics and zoogeography of Norwegian earthworms (Annelida, Oligochaeta: Lumbricidae). *Nytt Magasin for Zoologi*, 17: 169–280.

Swift, M. J., Heal, O. W., Anderson, J. M., Anderson, J. M. 1979. Decomposition in terrestrial ecosystems (Vol. 5). Univ of California Press.

Taniguchi, M., Yoshida, H. 2017. TFE3, HSP47, and CREB3 pathways of the mammalian Golgi stress response. *Cell Structure and Function*, e16023.

Terhivuo, J. 1988, January. The Finnish Lumbricidae (Oligochaeta) fauna and its formation. In *Annales Zoologici Fennici* (pp. 229-247). Finnish Academy of Sciences, Societas Scientiarum Fennica, Societas pro Fauna et Flora Fennica and Societas Biologica Fennica Vanamo.

Terhivuo, J., Saura, A., Hongell, K. 1994. Genetic and morphological variation in the parthenogenetic earthworm *Eiseniella tetraedra* (Sav.)(Oligochaeta: Lumbricidae) from South Finland and North Norway. *Pedobiologia*, 38(1): 81-96.

Terhivuo, J., Saura, A. 1996. Clone pool structure and morphometric variation in endogeic and epigeic North-European parthenogenetic earthworms (Oligochaeta: Lumbricidae). *Pedobiologia*, 40(3): 226-239.

Terhivuo, J., Saura, A. 1997. Island biogeography of North European parthenogenetic Lumbricidae: I. Clone pool affinities and morphometric differentiation of Åland populations. *Ecography*, 20(2): 185-196.

Terhivuo, J., Lundqvist, E., Saura, A. 2002. Clone diversity of *Eiseniella tetraedra* (Oligochaeta: Lumbricidae) along regulated and free-flowing boreal rivers. *Ecography*, 25(6): 714-720.

Terhivuo, J., Saura, A. 2006. Dispersal and clonal diversity of North-European parthenogenetic earthworms. *Biological Invasions*, 8(6): 1205-1218.

Terhivuo, J., Halmepuro, A. M., Saura, A. 2011. Clonal diversity and morphometry in the parthenogenetic earthworm *Eiseniella tetraedra* (Sav.) as affected by habitat characteristics including radioactive pollution. *Pedobiologia*, 54: S11-S18.

Thakur, M. P., Reich, P. B., Hobbie, S. E., Stefanski, A., Rich, R., Rice, K. E., ..., Eisenhauer, N. 2018. Reduced feeding activity of soil detritivores under warmer and drier conditions. *Nature climate change*, 8(1): 75-78.

Tiebe, M., Lutz, M., De La Garza, A., Buechling, T., Boutros, M., Teleman, A. A. 2015. REPTOR and REPTOR-BP regulate organismal metabolism and transcription downstream of TORC1. *Developmental Cell*, 33: 272-284.

Tiunov, A. V., Hale, C. M., Holdsworth, A. R., Vsevolodova-Perel, T. S. 2006. Invasion patterns of Lumbricidae into the previously earthworm-free areas of northeastern Europe and the western Great Lakes region of North America. In *Biological Invasions Belowground: Earthworms as Invasive Species* (pp. 23-34). Springer, Dordrecht.

Torres-Leguizamon, M., Mathieu, J., Decaëns, T., Dupont, L. 2014. Genetic structure of earthworm populations at a regional scale: inferences from mitochondrial and microsatellite molecular markers in *Aporrectodea icterica* (Savigny 1826). *PloS one*, 9(7): e101597.

Vélez, S., Mesibov, R., Giribet, G. 2012. Biogeography in a continental island: population structure of the relict endemic centipede *Craterostigmus tasmanianus* (Chilopoda, Craterostigmomorpha) in Tasmania using 16S rRNA and COI. *Journal of Heredity*, 103(1): 80-91.

Vellupillai, M., Nasir, K. H., Sivarajah, V., Othman, N. Q. 2016. Digital gene expression of water-stressed *Phyllanthus urinaria* (Dukung anak). *Journal of Tropical Agriculture*, 44: 121-136.

Vernesi, C., Pecchioli, E., Caramelli, D., Tiedermann, R., Randi, E., Bertorelle, G. 2002 The genetic structure of natural and reintroduced roe deer (*Capreolus capreolus*) populations in the Alps and central Italy, with reference to the mitochondrial DNA phylogeography of Europe. *Molecular Ecology,* 11: 1285-1297.

Viljoen, S. A., Reinecke, A. J. 1989. Moisture and growth, maturation and cocoon production of Eudrilus eugeniae (Oligochaeta). *Revue d'écologie et de biologie du sol*, 26(3): 291-303.

Vulsteke, V., Beullens, M., Boudrez, A., Keppens, S., Van Eynde, A., Rider, M. H., ..., Bollen, M. 2004. Inhibition of spliceosome assembly by the cell cycle-regulated protein kinase MELK and involvement of splicing factor NIPP1. *Journal of Biological Chemistry*, 279: 8642-8647.

Wall, D. H., Nielsen, U. N., Six, J. 2015. Soil biodiversity and human health. *Nature*, 528(7580): 69-76.

Walther, G. R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J., ..., Bairlein, F. 2002. Ecological responses to recent climate change. *Nature,* 416: 389-395.

Wang, K., Qiao, Y., Zhang, H., Yue, S., Li, H., Ji, X., Liu, L. 2018. Bioaccumulation of heavy metals in earthworms from field contaminated soil in a subtropical area of China. *Ecotoxicology and Environmental Safety*, 148: 876-883.

Wang, Q., Tan, X., Jiao, S., You, F., Zhang, P. J. 2014. Analyzing cold tolerance mechanism in transgenic zebrafish (*Danio rerio*). *PloS one*, 9: e102492.

Wang, S., Kaufman, R. J. 2012. The impact of the unfolded protein response on human disease. *Journal of Cell Biology*, 197: 857-867.

Wei, J., Xu, D., Zhou, J., Cui, H., Yan, Y., Ouyang, Z., ..., Qin, Q. 2010. Molecular cloning, characterization and expression analysis of a C-type lectin (Ec-CTL) in orange-spotted grouper, *Epinephelus coioides*. *Fish and Shellfish Immunology*, 28: 178-186.

Williams, D. L., Diehl, W. J. 1992. Interactive effects of soil moisture and food on glycolytic metabolism in *Eisenia fetida* (Oligochaeta). *Comparative Biochemistry and Physiology – Part B: Biochemistry and Molecular Biology*, 102: 911-917.

Willis, K. J., Whittaker, R. J. 2000. The refugial debate. *Science*, 287(5457): 1406-1407.

Withers, P. C. 1992. Comparative animal physiology. Philadelphia: Saunders College Pub.

Wolf, A. V. 1938. Notes on the effect of heat in *Lumbricus terrestris* L. *Ecology*, 19: 346-348.

Xu, D., Zhou, S., Sun, L. 2018. RNA-seq based transcriptional analysis reveals dynamic genes expression profiles and immune-associated regulation under heat stress in *Apostichopus japonicus*. *Fish and Shellfish Immunology*, 78: 169-176.

Yamauchi, S., Sasagawa, Y., Ogura, T., Yamanaka, K. 2007. Differential expression pattern of UBX family genes in *Caenorhabditis elegans*. *Biochememical and Biophysical Research Communications*, 358: 545-552.

Yang, C., Jiang, M., Wu, F., Yu, L., Tian, J., Liu, W., ..., Wen, H. 2016. Identification of a C-type lectin from tilapia (*Oreochromis niloticus*) and its functional characterization under low-temperature stress. *Fish and Shellfish Immunology*, 58: 631-640.

Zhang H, Zhang C, Li Z, Li C, Wei X, Zhang B, Liu Y. 2019. A new method of RNA secondary structure prediction based on convolutional neural network and dynamic programming. *Frontiers in genetics*, 10: 467.

Zhang, M., Yu, Q., Liu, Z., Liang, C., Zhang, B., Li, M. 2017. UBX domain-containing proteins are involved in lipid homeostasis and stress responses in *Pichia pastoris*. *The International Journal of Biochemistry & Cell Biology*, 90: 136-144.

Zheng, J., Cao, J., Mao, Y., Su, Y., Wang, J. 2019. Comparative transcriptome analysis provides comprehensive insights into the heat stress response of *Marsupenaeus japonicus. Aquaculture*, 502, 338- 346.

Zheng X, Cai D, Yao L, Teng Y. 2008. Non-concerted ITS evolution, early origin and phylogenetic utility of ITS pseudogenes in *Pyrus*. *Molecular Phylogenetics and Evolution,* 48: 892-903.

Zhu, Y. T., Zhang, X., Wang, S. C., Li, W. W., Wang, Q. 2016. Antimicrobial functions of EsLecH, a C-type lectin, via JNK pathway in the Chinese mitten crab, *Eriocheir sinensis*. *Developmental & Comparative Immunology*, 61: 225-235.

Zicsi, A., 1960. Die Regenwurmfauna des oberen ungarischen Donau-Ufergebietes. *Annales Universitatis Scientiarium Budapestinensis*, 3: 427-440.

Zicsi A. 1972. Eiseniella koreana eine neue Regenwurm-Art (Oligochaeta: Lumbricidae) aus Korea, *Annales Musei historico-naturalis hungarici*, 64: 129-132.

Supplementary Material

Chapter 1

Localities	Number in map (Fig. 3)	GPS coordinates	Specimens studied morphologically	Sequences obtained
Portocubelo (A Coruña)	1	42°48'15.10"N 9°8'7.88"O	19	8/-/-
Lavadero de Lira (A Coruña)	2	42°48'0.63"N 9° 8'4.42"O	36	8/-/-
Miñarzos (A Coruña)	3	42°47'57.43"N 9°7'59.05"O	16	5/1/1
Arroyo Playa Cons (A Coruña)	4	42º48'26.2" N 9 º 6'29.5" O	8	2/1/-
Mallou (A Coruña)	5	42°48'25.99"N 9° 5'56.55"O	3	3/3/-
O Cruceiro (A Coruña)	6	42°51'5.61"N 9° 4'55.27"O	8	4/1/-
Fuente de Cornido (A Coruña)	7	42°52'5.58"N 9° 5'27.10"O	11	6/1/-
Cornido (A Coruña)	8	42°52'5.58"N 9° 5'27.10"O	1	1/1/1
Montalvo (Pontevedra)	9	42°23'57.98"N 8°50'46.82"O	13	8/4/-
Lavadero de Oia (Pontevedra)	10	42° 0'11.33"N 8°52'21.53"O	10	5/3/-
Lavadero de Casalonga (A Coruña)	11	42º49'17.15"N 8º36'49.69"O	7	7/2/1
Forcarei (Pontevedra)	12	42°34'59.39"N 8°20'39.17"O	40	4/-/-
Souto (Lugo)	13	42°55' 6.5" N 7° 12' 53.93"O	3	3/2/1
Chaguazoso (Orense)	14	42°11'31.29"N 7°12'59.30"O	18	7/3/3

Supplementary Table 1. Localities sampled in Northwestern Spain.

Localities	Number in map (Fig. 4)	GPS coordinates	Specimens studied morphologically	Sequences obtained (COI/16S/28S)
Santa María de la Alameda (Madrid)	15	40°35'44.76"N 4°13'46.40"O	2	2/1/1
Fuente de las Hondillas (Madrid)	16	40°41'50.79"N 4° 8'22.75"O	36	19/4/1
Los Molinos. Arroyo de los Irrios (Madrid)	17	40º42'26.19"N 4º0.5'42.14"O	1	1/1/-
La Jarosa (Madrid)	18	40°39'45.83"N 4° 7'44.12"O	6	1/1/-
Fuente de Guadarrama (Madrid)	19	40°40'23.66"N 4° 6'17.52"O	14	5/-/1
Los Molinos. Arroyo Majaltobar (Madrid)	20	40°43'4.00"N 4° 5'9.16"O	18	12/-/1
Valdemorillo (Madrid)	21	40°30'57.41"N 4° 5'41.11"O	2	2/1/-
Parquelagos (Madrid)	22	40°35'57.48"N 3°58'8.40"O	12	11/-/-
Las Rozas (Madrid)	23	40°31'3.64"N 3°56'24.59"O	10	10/8/3
Villanueva del Pardillo (Madrid)	24	40°29'8.29"N 3°56'11.47"O	7	6/3/-
Arroyomolinos (Madrid)	25	40°16'40.37"N 3°55'47.95"O	13	6/1/3
Casarrubios del Monte (Toledo)	26	40°10'59.82"N 4° 1'43.24"O	21	16/3/-
Rielves (Toledo)	27	39°57'38.10"N 4°11'29.16"O	1	1/-/-
Albarreal del Tajo (Toledo)	28	39°53'54.24"N 4°11'2.25"O	7	4/4/2
Alcalá de Henares (Madrid)	29	40º30'40.26"N 3º18'29.15"O	14	9/5/3
Arroyo el Soto-Móstoles (Madrid)	No data	No data	3	3/2/2

Supplementary Table 2. Localities sampled in central area of Spain.

Country	Number in map (Fig. 5)	COI Sequences
Wales	30	6 (16S=5;28S=3)
France	31	10
Norway	32	3
Sweden	33	36
Austria	34	1
Turkey	35	1
New Zealand	36	1
Canada	37	22
United States	38	12

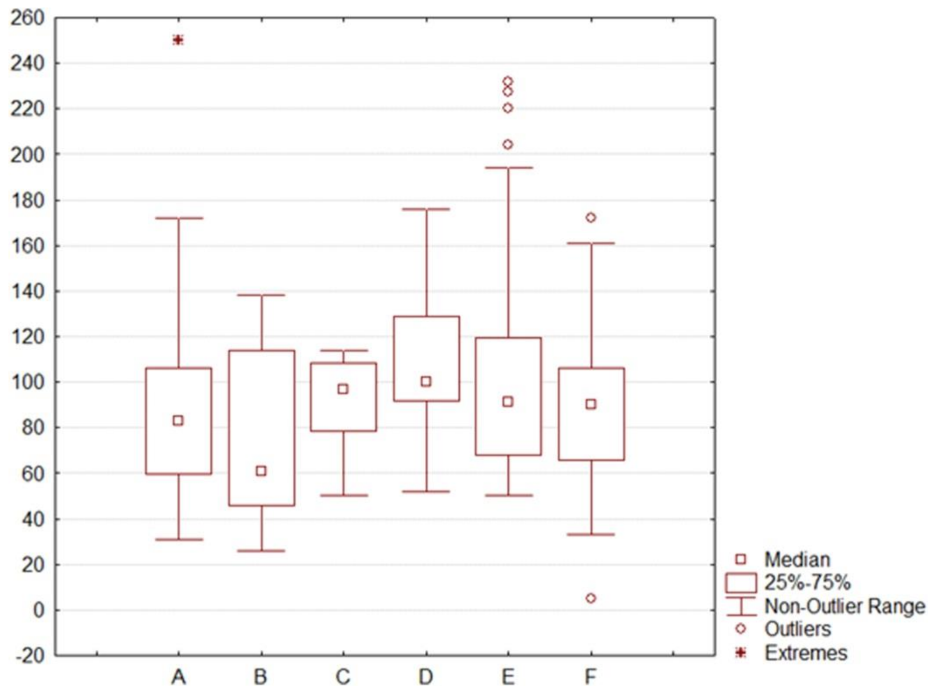
Supplementary Table 3. Worldwide sequences studied.

Species	Accesion number COI	Accesion number 16S	Accesion number 28S
Eiseniella tetraedra	KY284162-KY284343	KX212907-KX212965	KX212966-KX212990
Carpetania matritensis	GQ409661.1	JN209218.1	GQ409652.1
Lumbricus rubellus	KM611946.1	KJ912567.1	KJ912213.1
Dendrobaena byblica	Domínguez et al. 2015; Pérez-Losada et al. 2015	KJ912523.1	KJ912165.1
Iberoscolex oliveirae	Domínguez et al. 2015; Pérez-Losada et al. 2015	KJ912555.1	Domínguez et al. 2015; Pérez-Losada et al. 2015
Prosellodrilus biariculatus	Domínguez et al. 2015; Pérez-Losada et al. 2015	KJ912586.1	JN871950.1

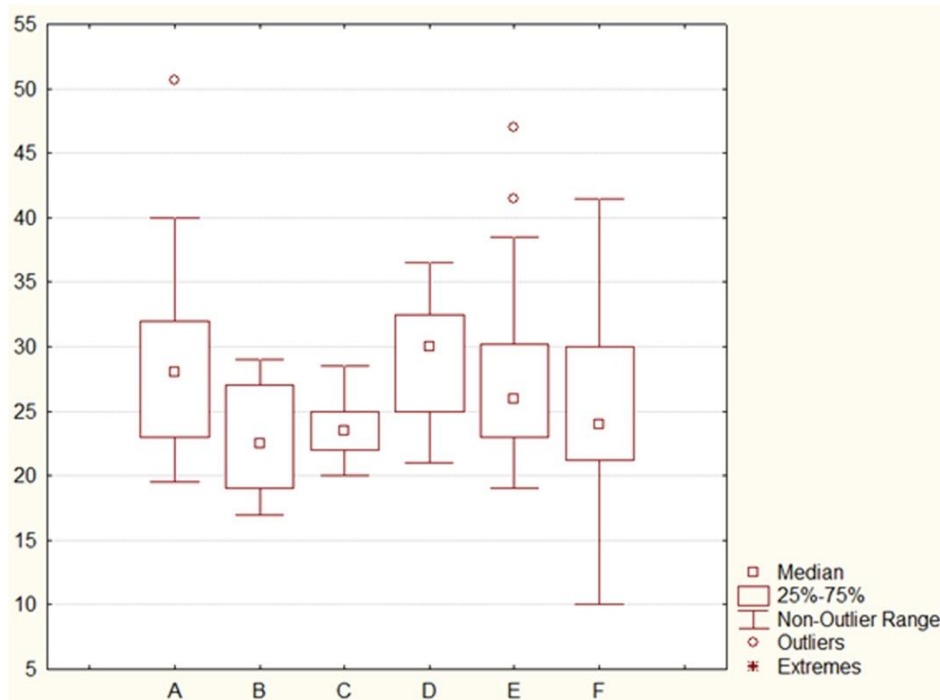
Supplementary Table 4. Accession numbers and references of all sequences used.

CLITELLUM	A(%)	A(n)	B(%)	B(n)	C(%)	C(n)	D(%)	D(n)	E(%)	E(n)	F(%)	F(n)
22-26	7.4	2	12.5	1	40	4	18.75	3	18.36	9	6.25	2
22-27	66.66	18	50	4	50	5	81.25	13	55.1	27	56.25	18
22-28	11.11	3	37.5	3					14.28	7	25	8
others	14.81	4			10	1			12.24	6	12.5	4
TUBERCULA PUBERTATIS												
23-25	31.03	9	44.4	4	40	4	40	6	62	31	39.47	15
23-1n26	31.03	9		0	40	4	53.33	8	10	5	15.79	6
23-26	24.13	7	11.2	1			6.66	1	16	8	21.05	8
others	13.79	4	44.4	4	20	2			12	6	23.68	9
MALE PORE												
9									1.92	1		
11												
12	3.44	1							3.85	2	5.26	2
13	96.55	28	100	9	75	9	93.75	15	92.3	48	94.74	36
14					8.3	1						
15					16.66	2	6.25	1	1.92	1		
SEMINAL VESICLES												
9-10-11-12	40.74	11	33.3	3	27.27	3	66.66	10	42.85	21	38.7	12
9--11--12	18.51	5	33.3	3	36.36	4	20	3	28.57	14	16.12	5
others	40.74	11	33.3	3	36.36	4	13.33	2	28.57	14	45.16	14

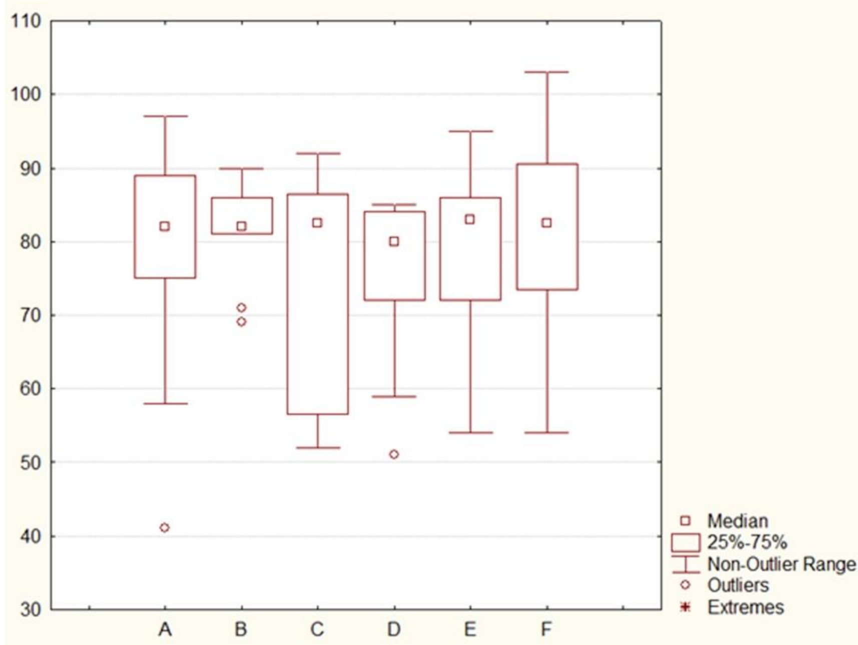
Supplementary Table 5. Morphological data of specimens of *E. tetraedra* studied. % indicates percentage of specimens and n absolute numbers.



Supplementary Figure 1. Statistical analysis of weight. P (Kruskal-Wallis) = 0.1344. Lineages are represented in the x axis.



Supplementary Figure 2. Statistical analysis of length. P (Kruskal-Wallis) = 0.0155. Lineages are represented in the x axis.



Supplementary Figure 3. Statistical analysis of number of segments. P (Kruskal-Wallis) = 0.6607. Lineages are represented in the x axis.

Chapter 2

Locality	Longitude	Latitude	Sampling	Number	Habitat
Burgohondo	-4.7877	40.4139	Chapter 2	1	Stable
Soria	-2.8566	41.7273	Chapter 2	2	Stable
Ordesa	0.0784	42.6381	Chapter 2	3	Unstable
Posets-Maladeta	0.5282	42.6522	Chapter 2	4	Unstable
Alba de Tormes	-5.5179	40.8268	Chapter 2	5	Stable
Sanjuanejo	-6.4928	40.5683	Chapter 2	6	Stable
Portalegre	-7.9096	39.0246	Chapter 2	7	Stable
Casas de Don Antonio	-6.2904	39.2339	Chapter 2	8	Unstable
Viseu	-7.4949	40.8478	Chapter 2	9	Unstable
Alcacer do Sal	-8.3372	38.2635	Chapter 2	10	Stable
Jabugo	-6.7001	38.0791	Chapter 2	11	Unstable
Ziordia	-2.2303	42.8700	Chapter 2	12	Stable
Miranda del Ebro	-2.9282	42.6885	Chapter 2	13	Stable
Olmedillo de Roa	-3.9333	41.7822	Chapter 2	14	Unstable
Valladolid	-4.7373	41.6674	Chapter 2	15	Stable
Río Cauxa	-6.3229	43.2797	Chapter 2	16	Stable
El Castril	-2.7769	37.7899	Chapter 2	17	Unstable
Urda	-3.7116	39.4165	Chapter 2	18	Unstable
Arenas del Rey	-3.8968	36.9572	Chapter 2	19	Stable
Peraleda del Zaucejo	-5.5664	38.4714	Chapter 2	20	Unstable
Chelva	-0.9975	39.7582	Chapter 2	21	Unstable
Tarazona	-1.7275	41.9023	Chapter 2	22	Stable
Poveda de la Sierra	-2.0283	40.6412	Chapter 2	23	Stable
El Palau d'Anglesola	0.8994	41.6566	Chapter 2	24	Unstable
Tibi	-0.5874	38.5283	Chapter 2	25	Unstable
Arcos de Jalón	-2.2737	41.2158	Chapter 2	26	Stable
Sos del Rey Católico	-1.2194	42.4768	Chapter 2	27	Unstable
San Blas	-1.1774	40.3582	Chapter 2	28	Stable
Calahorra	-1.9674	42.3002	Chapter 2	29	Stable
Molina de Aragón	-1.8894	40.8427	Chapter 2	30	Stable
Sotuélamos	-2.5723	39.0408	Chapter 2	31	Stable
Esporlas	2.5795	39.6657	Chapter 2	32	Unstable
Bragança	-6.7540	41.8045	Chapter 2	33	Stable
Mallo de Luna	-5.8873	42.8721	Chapter 2	34	Stable
Puenticella	-6.5265	43.1383	Chapter 2	35	Stable
El Bosque	-5.5034	36.7622	Chapter 2	36	Stable
Miñarzos	-9.1330	42.7992	Chapter 1	37	Unstable
Portocubelo	-9.1355	42.8041	Chapter 1	38	Unstable
Mallou	-9.0990	42.8072	Chapter 1	39	Unstable
Lavadero de Lira	-9.1345	42.8001	Chapter 1	40	Unstable
Fuente de Cornido	-9.0908	42.8682	Chapter 1	41	Unstable
Ocruceiro	-9.0820	42.8515	Chapter 1	42	Unstable

Locality	Longitude	Latitude	Sampling	Number	Habitat
Cornido	-9.0908	42.8682	Chapter 1	43	Unstable
Chaguazoso	-7.2164	42.1920	Chapter 1	44	Unstable
Arroyo Playa Cons	-9.1080	42.8072	Chapter 1	45	Unstable
Forcarei	-8.3442	42.5831	Chapter 1	46	Unstable
Montalvo	-8.8463	42.3994	Chapter 1	47	Unstable
Lavadero de Oia	-8.8726	42.0031	Chapter 1	48	Unstable
Lavadero de Casalonga	-8.6119	42.8217	Chapter 1	49	Unstable
Souto	-7.2151	42.9182	Chapter 1	50	Unstable
Fuente de las Hondillas	-4.1396	40.6974	Chapter 1	51	Unstable
Alcalá de Henares	-3.3080	40.5111	Chapter 1	52	Stable
Las Rozas	-3.9401	40.5176	Chapter 1	53	Stable
Villanueva del Pardillo	-3.9365	40.4856	Chapter 1	54	Stable
Valdemorillo	-4.0947	40.5159	Chapter 1	55	Unstable
Río Hornillo	-4.2295	40.5957	Chapter 1	56	Unstable
El Soto	no data	no data	Chapter 1	57	No data
Arroyo de los Irrios	-4.0858	40.7177	Chapter 1	58	Unstable
Los Molinos	-4.0858	40.7177	Chapter 1	59	Unstable
Parquelagos	-3.9690	40.5993	Chapter 1	60	Unstable
Arroyomolinos	-3.9299	40.2778	Chapter 1	61	Unstable
Fuente de Guadarrama	-4.1048	40.6732	Chapter 1	62	Unstable
Rielves	-4.1902	39.9634	Chapter 1	63	Unstable
Casarrubios del Monte	-4.0286	40.1832	Chapter 1	64	Unstable
Albarreal del Tajo	-4.1839	39.8984	Chapter 1	65	Stable

Supplementary Table 1. Localities sampled, geographic coordinates and type of habitat.

Species	Accesion number COI	Accesion number 16S	Accesion number 28S
Eiseniella tetraedra	KY284162-KY284343 MZ229064-MZ229298	KX212907-KX212965 MZ156460-MZ156566	KX212966-KX212990 MZ156353-MZ156459
Carpetania matritensis	GQ409661.1	JN209218.1	GQ409652.1
Lumbricus rubellus	KM611946.1	KJ912567.1	KJ912213.1
Dendrobaena byblica	Domínguez et al. 2015; Pérez-Losada et al. 2015	KJ912523.1	KJ912165.1
Iberoscolex oliveirae	Provided by the authors	KJ912555.1	Domínguez et al. 2015; Pérez-Losada et al. 2015
Prosellodrilus biariculatus	Domínguez et al. 2015; Pérez-Losada et al. 2015	KJ912586.1	JN871950.1

Supplementary Table 2. Accession numbers and references of all sequences used.

Variable	Source
BIO1 = Annual Mean Temperature	www.worldclim.org
BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))	www.worldclim.org
BIO3 = Isothermality (BIO2/BIO7) (×100)	www.worldclim.org
BIO4 = Temperature Seasonality (standard deviation ×100)	www.worldclim.org
BIO5 = Max Temperature of Warmest Month	www.worldclim.org
BIO6 = Min Temperature of Coldest Month	www.worldclim.org
BIO7 = Temperature Annual Range (BIO5-BIO6)	www.worldclim.org
BIO8 = Mean Temperature of Wettest Quarter	www.worldclim.org
BIO9 = Mean Temperature of Driest Quarter	www.worldclim.org
BIO10 = Mean Temperature of Warmest Quarter	www.worldclim.org
BIO11 = Mean Temperature of Coldest Quarter	www.worldclim.org
BIO12 = Annual Precipitation	www.worldclim.org
BIO13 = Precipitation of Wettest Month	www.worldclim.org
BIO14 = Precipitation of Driest Month	www.worldclim.org
BIO15 = Precipitation Seasonality (Coefficient of Variation)	www.worldclim.org
BIO16 = Precipitation of Wettest Quarter	www.worldclim.org
BIO17 = Precipitation of Driest Quarter	www.worldclim.org
BIO18 = Precipitation of Warmest Quarter	www.worldclim.org
BIO19 = Precipitation of Coldest Quarter	www.worldclim.org
pH water *10 (Depth= 0-5 cm)	https://soilgrids.org/

Variable	Source
Soil organic carbon in dg/kg (Depth= 0-5 cm)	https://soilgrids.org/
Sand in g/kg (Depth= 0-5 cm)	https://soilgrids.org/

Supplementary Table 3. Environmental and soil variables retrieved from Worldclim and Soilgrids used in the study.

	Nº OF HAPLOTYPES	Nº OF HAPLOTYPES PRESENT IN ONE POPULATION	Nº OF HAPLOTYPES RESENT IN TWO OR MORE POPULATIONS
CLADE I	69	67	2
LINEAGE B	40	39	1
LINEAGE C	11	10	1
LINEAGE D	12	12	0
LINEAGE G	6	6	0
CLADE II	200	188	12
LINEAGE A	58	55	3
LINEAGE E	58	56	2
LINEAGE F	82	75	7
LINEAGE H	2	2	0

Supplementary Table 4. Numbers of COI haplotypes included in each clade and lineage. Number of local haplotypes (present in one population) and present in two or more populations.

16S/COI	A	B	C	D	E	F	G	H
A	0.59/0.72	7.04	7.04	7.79	5.79	5.2	7.46	6.13
B	0.91	0.5/0.51	2.91	3.48	5.05	5.26	4.25	5.45
C	1.58	1.69	0.18/0.25	3.97	5.88	5.36	5.09	6.7
D	1.36	1.13	1.65	0.07/1.58	6.83	6.39	5.28	7.24
E	1.04	0.83	1.6	1.11	0.44/0.52	3.95	5.92	4.24
F	1.13	1.1	1.8	1.38	1.06	0.19/0.27	6.18	4.6
G	1.35	1.28	1.92	1.61	1.21	1.32	0/0.79	6.78
H	0.82	0.65	1.59	1.03	0.71	0.97	1.22	0.13/3.64

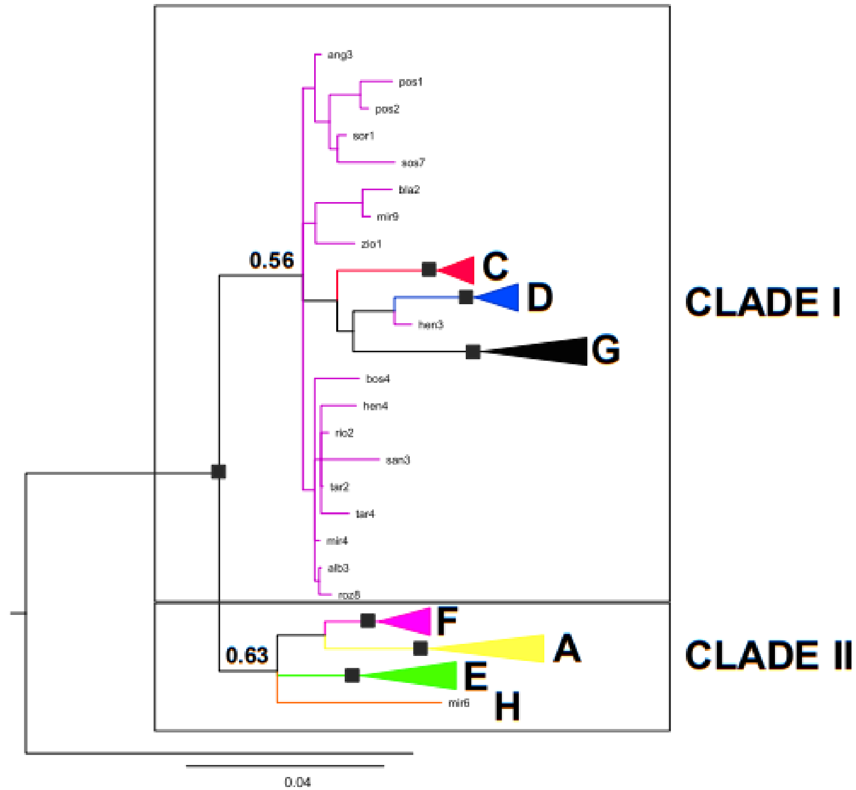
Supplementary Table 5. Percentage of uncorrected pairwise genetic distances based on COI (values above diagonal) and 16S (below diagonal) between lineages (within lineages in diagonal) retrieved for *E.tetraedra.*

CLITELLUM	A(%)	A(n)	B(%)	B(n)	C(%)	C(n)	D(%)	D(n)	E(%)	E(n)	F(%)	F(n)
22-26	12.25	6	9.09	2	28.57	2	18.75	3	11.59	8	6.49	5
22-27	59.18	29	68.18	15	57.14	4	81.25	13	65.22	45	67.53	52
22-28	20.4	10	13.64	3	0	0	0	0	14.49	10	18.18	14
others	8.17	4	9.09	2	14.29	1	0	0	8.7	6	7.18	6
TUBERCULA PUBERTATIS												
23-25	48.15	26	64	16	62.5	5	40	6	59.72	43	74.68	59
23-1n26	16.67	9	4	1	12.5	1	46.66	7	9.72	7	5.06	4
23-26	16.67	9	4	1	0	0	6.67	1	18.06	13	10.13	8
others	18.51	10	28	7	25	2	6.67	1	12.5	9	10.13	8
POSITION OF MALE PORES												
8	0	0	0	0	0	0	0	0	0	0	1.16	1
9	0	0	0	0	0	0	0	0	1.28	1	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	1.28	1	0	0
12	1.75	1	0	0	0	0	0	0	2.57	2	4.65	4
13	98.25	56	100	27	75	6	93.75	15	94.87	74	94.19	81
14	0	0	0	0	12.5	1	0	0	0	0	0	0
15	0	0	0	0	12.5	1	6.25	1	0	0	0	0
SEMINAL VESICLES												
9-10-11-12	51.92	27	40.9	9	42.86	3	62.5	10	44.12	30	33.33	25
10-11--12	3.85	2	4.55	1	14.28	1	6.25	1	10.29	7	16	12
11--12	25	13	9.09	2	28.57	2	6.25	1	5.88	4	20	15
no present	0	0	0	0	0	0	0	0	0	0	1.34	1
others	19.23	10	45.46	10	14.29	1	25	4	39.71	27	29.33	22

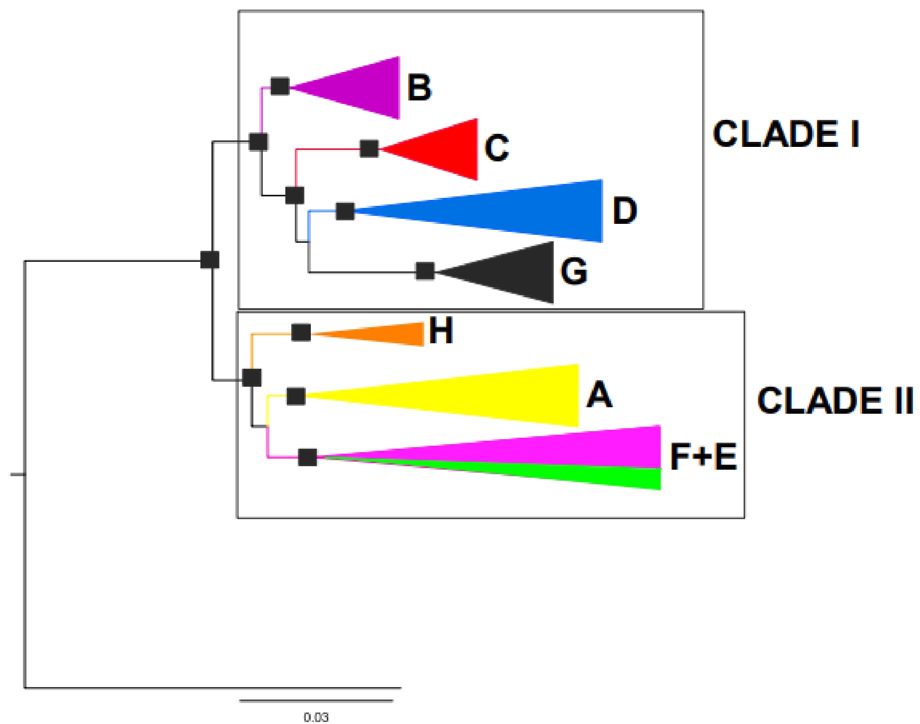
	A(%)	A(n)	B(%)	B(n)	C(%)	C(n)	D(%)	D(n)	E(%)	E(n)	F(%)	F(n)
SPERMATHECAE												
no present	82.7	43	45.45	10	57.14	4	75	12	54.41	37	70.66	53
no iridiscent	17.3	9	54.55	12	42.86	3	25	4	45.59	31	29.34	22
MALE FUNNELS												
no present	90.38	47	90.9	20	71.43	5	43.75	7	47.06	32	73.33	55
no iridiscent	9.62	5	9.1	2	28.57	2	56.25	9	52.94	36	26.67	20

Supplementary Table 6. Number and percentage of individuals presenting different states of morphological characters for each lineage. No results could be obtained for lineages G and H because of all individuals were immatures.

Supplementary Figure 1. Localities sampled. Numbers of each locality correspond to Supplementary Table 1. 1: Localities from a lower-scale study in Carnota, A Coruña, Spain (Chapter 1). 2: Localities from a lower-scale study in Guadarrama river basin, Madrid, Spain (Chapter 1).

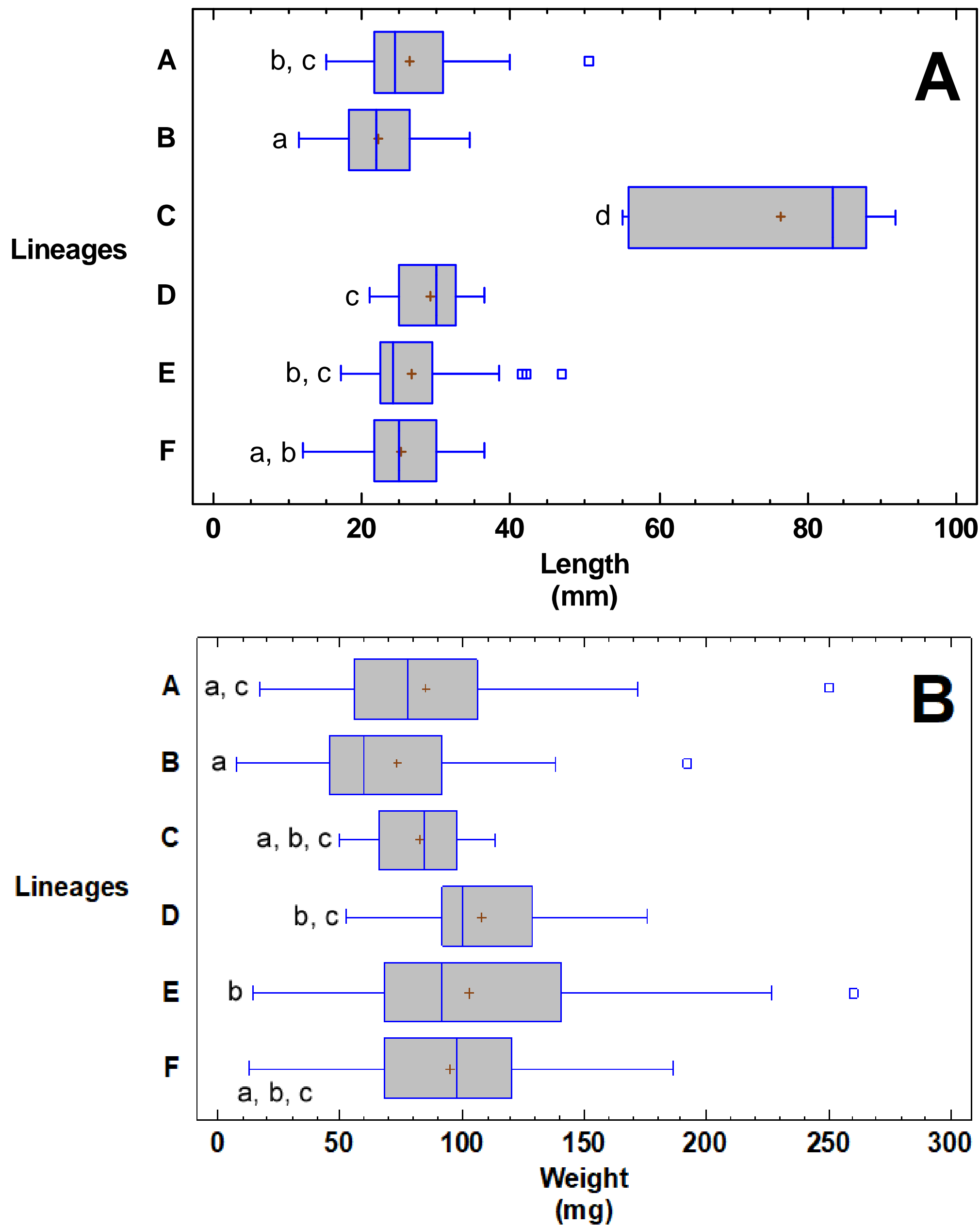


Supplementary Figure 2. Bayesian inference (BI) of the phylogenetic tree based on the concatenated sequences of COI, 16S-tRNAs and 28S. Posterior probability/bootstrap support values (of Maximum Likelihood Analysis, ML) are shown when they are higher than 0.9/0.7 (BI/ML) as black squares. The scale bar represents 0.04 substitutions per position. Colors used are the same as in Figure 1.



Supplementary Figure 3. Bayesian inference (BI) of the phylogenetic tree based on the concatenated sequences of COI, 16S-tRNAs and 28S (rake-like polytomies removed). Posterior probability/bootstrap support values (of Maximum Likelihood Analysis, ML) are shown when they are higher than 0.9/0.7 (BI/ML) as black squares. The scale bar represents 0.03 substitutions per position. Colors used are the same as in Figure 1.

Supplementary Figure 4. Lineages and clades distribution (based on presences) in the Iberian Peninsula. Colors are the same used as in Figure 1.



Supplementary Figure 5. Box-plot diagram of length (A) and weight (B). Different letters in each character indicate different groups in multiple range tests.

Chapter 3

Locality	Longitude	Latitude	Sampling
Burgohondo	-4,7877	40,4139	Chapter 2
Soria	-2,8566	41,7273	Chapter 2
Ordesa	0,0784	42,6381	Chapter 2
Posets-Maladeta	0,5282	42,6522	Chapter 2
Alba de Tormes	-5,5179	40,8268	Chapter 2
Sanjuanejo	-6,4928	40,5683	Chapter 2
Portalegre	-7,9096	39,0246	Chapter 2
Casas de Don Antonio	-6,2904	39,2339	Chapter 2
Viseu	-7,4949	40,8478	Chapter 2
Alcacer do Sal	-8,3372	38,2635	Chapter 2
Jabugo	-6,7001	38,0791	Chapter 2
Ziordia	-2,2303	42,87	Chapter 2
Miranda del Ebro	-2,9282	42,6885	Chapter 2
Olmedillo de Roa	-3,9333	41,7822	Chapter 2
Valladolid	-4,7373	41,6674	Chapter 2
Río Cauxa	-6,3229	43,2797	Chapter 2
El Castril	-2,7769	37,7899	Chapter 2
Urda	-3,7116	39,4165	Chapter 2
Arenas del Rey	-3,8968	36,9572	Chapter 2
Peraleda del Zaucejo	-5,5664	38,4714	Chapter 2
Chelva	-0,9975	39,7582	Chapter 2
Tarazona	-1,7275	41,9023	Chapter 2
Poveda de la Sierra	-2,0283	40,6412	Chapter 2
El Palau d'Anglesola	0,8994	41,6566	Chapter 2
Tibi	-0,5874	38,5283	Chapter 2
Arcos de Jalón	-2,2737	41,2158	Chapter 2
Sos del Rey Católico	-1,2194	42,4768	Chapter 2
San Blas	-1,1774	40,3582	Chapter 2
Calahorra	-1,9674	42,3002	Chapter 2
Molina de Aragón	-1,8894	40,8427	Chapter 2
Sotuélamos	-2,5723	39,0408	Chapter 2
Esporlas	2,5795	39,6657	Chapter 2
Bragança	-6,754	41,8045	Chapter 2
Mallo de Luna	-5,8873	42,8721	Chapter 2
Puenticella	-6,5265	43,1383	Chapter 2
El Bosque	-5,5034	36,7622	Chapter 2
Miñarzos	-9,133	42,7992	Chapter 1
Portocubelo	-9,1355	42,8041	Chapter 1
Mallou	-9,099	42,8072	Chapter 1
Lavadero de Lira	-9,1345	42,8001	Chapter 1

Locality	Longitude	Latitude	Sampling
Fuente de Cornido	-9,0908	42,8682	Chapter 1
Ocruceiro	-9,082	42,8515	Chapter 1
Cornido	-9,0908	42,8682	Chapter 1
Chaguazoso	-7,2164	42,192	Chapter 1
Arroyo Playa Cons	-9,108	42,8072	Chapter 1
Forcarei	-8,3442	42,5831	Chapter 1
Montalvo	-8,8463	42,3994	Chapter 1
Lavadero de Oia	-8,8726	42,0031	Chapter 1
Lavadero de Casalonga	-8,6119	42,8217	Chapter 1
Souto	-7,2151	42,9182	Chapter 1
Fuente de las Hondillas	-4,1396	40,6974	Chapter 1
Alcalá de Henares	-3,308	40,5111	Chapter 1
Las Rozas	-3,9401	40,5176	Chapter 1
Villanueva del Pardillo	-3,9365	40,4856	Chapter 1
Valdemorillo	-4,0947	40,5159	Chapter 1
Río Hornillo	-4,2295	40,5957	Chapter 1
El Soto	no data	no data	Chapter 1
Arroyo de los Irrios	-4,0858	40,7177	Chapter 1
Los Molinos	-4,0858	40,7177	Chapter 1
Parquelagos	-3,969	40,5993	Chapter 1
Arroyomolinos	-3,9299	40,2778	Chapter 1
Fuente de Guadarrama	-4,1048	40,6732	Chapter 1
Rielves	-4,1902	39,9634	Chapter 1
Casarrubios del Monte	-4,0286	40,1832	Chapter 1
Albarreal del Tajo	-4,1839	39,8984	Chapter 1

Supplementary Table 1. Localities sampled in the Iberian Peninsula.

Locality	Longitude	Latitude
Napoli	14,3306	40,8849
Caorso-delta roncaglia	9,8526	45,04356
Masio	8,41	44,8712
Pisa	10,3892	43,729
Verona	10,9992	45,4339
Roma	12,4657	41,9012
Via Boalto a Levante	11,545	45,0388
Lusserna	7,2433	44,6444
Pontecorvo	13,6652	41,4542
Milan	9,2006	45,4739
Vicenza	11,5511	45,5471
Valdarno	11,5606	43,5499
Ferriera	12,4508	43,0904
Jesi	13,2291	43,5134
San Lorenzo	13,1845	43,0538
Conca Cave	15,2733	37,9567
Amandola	13,3574	42,9842
Antillo	13,3674	42,966
Affluent of Fiume Corno, near Norcia	13,0797	42,7949
near Genga Railway Station, Genga (Ancona), Marche, Italy	12,9773	43,4029

Supplementary Table 2. Localities sampled in Italy by Irene de Sosa, Christer Érseus and co-workers (blue) and Misel Jelic (purple).

Locality	Longitude	Latitude
EW426	11,2258	58,8999
EW336	11,9602	57,6873
Jättadalen Nature Reserve	13,6986	58,4325
ew455	11,9551	57,6822
EW338	12,2706	57,7614
NOEW26	5,7543	58,5383
NOEW27	5,9183	58,5021
EW352	18,4223	57,7752
EW354	18,16	57,314
NOEW48	12,0539	61,5706
NOEW55	11,2264	61,5678
NOEW56	12,2002	61,3025
EW343	12,6142	59,7612
NOEW70	8,2084	60,5307
NOEW74	7,6466	60,4119
NOEW81	6,1986	60,6388
NOEW86	7,29949	61,40259

Locality	Longitude	Latitude
NOEW97	7,143	62,1949
NOEW110	7,1079	60,8604
EW363	12,012	57,7439
EW365	12,79829	57,92529
EW364	12,2251	57,7702
EW369	12799	57931
EW458	11,9604	57,6805
EW373	14,7193	61,3272
EW376	17,1118	61,72969
EW379	17,0481	61,3001
NOEW131	9,1798	59,5788
NOEW135	8,486	59,8724
NOEW136	8,2297	59,7712
NOEW138	8,0122	59,4455
NOEW141	6,5263	59,9079
EW388	15,80799	59,132
NOEW161	10,2189	61,4274
NOEW165	7,7058	62,5667
NOEW167	8,5508	62,8917
NOEW169	10,9794	63,4639
NOEW170	12,1039	64,2318
NOEW171	12,91807	64,79248
NOEW172	13,4041	65,53308
NOEW174	13,703	66,05419
NOEW182	17,2609	68,53631
NOEW183	16,5625	68,6303
NOEW186	15,9844	69,1223
NOEW192	15,2939	67,2656
NOEW194	14,9989	67,0747
NOEW198	13,245	65,8025
NOEW200	13,1569	64,9264
NOEW214	10,0739	60,8231
EW431	11,9548	57,6796
NOEW221	10,7059	59,9281
EW412	17,9874	59,4306
Kvillebäcken Stream at Hökälla-Ekehöjd	11,943	57,7563
EW461	12,5007	57,9332
EW463	13,2459	56,0359
EW464	15,282	56,199
EW465	15,105	56,179
EW466	12,4257	62,6453
NOEW247	11,8164	61,6568
NOEW253	11,521	58,9099

Locality	Longitude	Latitude
NOEW254	11,4359	59,0859
NOEW258	10,6956	59,542
EW475	16,6074	57,6593
EW433	11752	57732
NOEW268	10,016	59,043
NOEW278	6,3153	60,1295
NOEW288	8,7287	61,1488
EW 1	11,8452	57,6297
NOEW298	25,7633	71,1447
NOEW301	24,0789	70,2332
NOEW302	23,3	69,95
NOEW304	22,0121	69,8409
NOEW312	17,98608	69,6395
NOEW183	16,5625	68,6303
NOEW158	11,2569	60,5628
NOEW160	10,4783	61,12
NOEW328	9,704	62,5918
NOEW329	12,6195	64,6366
NOEW332	14,4343	67,2784
NOEW343	10,1945	62,9701
EW487	11,9591	57,6807
EW494	14,27	55,5321
EW496	12,49704	58,93791
EW501	12,4473	59,0002
EW502	12,2338	59,1386
EW4	16,0425	59,0853
EW503	15,0186	59,0398
NOEW351	10,06376	60,47608
NOEW358	6,6094	61,8464
NOEW366	9,6369	60,3704
EW499	12,5737	58,8211
EW437	16,8764	56,9855
EW38	15,1431	59,2664
EW439	16,8539	56,8621
EW440	16,66111	56,8195
EW441	16,6468	56,8056
EW442	16,8764	56,9855
EW5	16,0872	59,0875
EW4	16,0425	59,0853
EW445	12,241	57,776
EW 15	14,7456	56,2997
EW 14	14,5028	56,1058
EW 29	14,3127	58,0372

Locality	Longitude	Latitude
EW 34	14,6339	58,31
EW 30	14,3084	58,0414
EW 37	14,7958	58,7306
EW 38	15,1431	59,2664
EW 47	14,5564	62,4525
EW51	14,4806	62,9831
EW 52	13,87	62,8506
EW 55	12,9486	62,2472
EW 59	14,1792	59,1169
ew64	11,6154	58,3695
EW 100	11,5744	58,4327
EW 95	11,1275	58,8903
EW 105	13,1247	57,59
EW 106	13,1233	57,5928
EW109	11,983	57,5
EW 115	11,9542	57,6831
EW 116	11,9567	57,6817
EW 124	12,9353	56,3887
EW 126	12,96	55,4797
EW 130	13,0752	56,3136
EW 133	12,7978	57,9253
EW 136	18,1603	57,3139
EW 140	12,453	57,993
EW 144	18,2743	59,496
EW 151	17,6282	59,8515
EW419	18,4046	57,7402
EW 159	18,2742	59,4973
EW 162	18,5502	59,571
EW 164	16,7598	58,6298
EW420	18,5988	57,8409
EW213	11,969	57,683
EW 223	17,4782	62,5235
Locality	**Longitude**	**Latitude**
EW 225	17,8746	62,8919
EW 227	18,4314	62,8797
EW 229	19,6624	63,6038
EW 265	15,1348	65,0873
EW 270	14,4652	65,0947
EW 271	14,2979	65,0176
EW424	12,2845	57,7736
EW 303	18,4103	57,32
EW 307	18,285	57,635
EW 310	18,8163	57,8507

Locality	Longitude	Latitude
EW 317	11,4408	59,0078
EW 318	11,1931	58,9453
EW 319	11,2261	58,9003
EW 324	12,285	57,774
EW 248	19,0383	68,3484
EW 257	16,1578	65,9594
EW 263	16,0426	65,2151
EW 267	14,6968	65,0635
EW 269	14,5219	65,1194
EW 273	14,1278	64,8268
EW 274	14,1218	64,619
EW 278	13,2783	63,3153
EW 290	14,2116	58,9656
EW 284	16,6082	56,5443
EW 293	11,5423	58,0225
EW 297	14,1823	57,7529
EW332	16,9983	59,5103
Lake Alvajärvi, Jyväskylä	25,7109	62,3186
EW 187	15,2247	56,7827
EW 191	11,9562	57,6813
NOEW54	10,7676	62,1722
NOEW111	7,5585	60,8073

Supplementary Table 3. Localities sampled in the Scandinavian Peninsula by Christer Érseus and co-workers.

Locality	Longitude	Latitude
Brest	No data	No data
FR15829.2.1, Francia	2,5489	45,5854
FR15829.2.1, Francia	2,5489	45,5854
FR14807.1.1, Francia	1,2914	45,9135
FR15808 1.1, Francia	5,9462	45,3158
FR14812.1.1, Francia	2,42923	42,6314
FR15L31.1.1, Francia	2,4292	42,6315
FR15507.1a.1, Francia	3,8379	47,5797
FR14824.1B.1, Francia	3,2109	45,9402
FR15507.1c.1, Francia	3,8379	47,5798
FR15L28.1.1, Francia	7,1557	48,0433
FR14824.1b.1, Francia	3,2109	45,9402
FR15825.1.1, Francia	1,0018	44,4063
FR15502.1.1, Francia	-0,6405	46,313
FR14421.1a.1, Francia	3,3255	48,4495
FR16222.3.1, Francia	6,0319	44,7123
FR15L27.1b.1, Francia	6,0876	48,0718
FR15517.1b.1, Francia	1,9426	48,6821
FR14816.2.1, Francia	2,7487	42,8265
Porquerolles island	6,2023	42,9888
Port-Cros island	6,3839	43,0068
Cap Lardier	6,6045	43,1809

Supplementary Table 4. Localities sampled in France by Emmanuelle Lapied, Daniel F. Marchán (orange) and Nuria Sánchez (pink).

Locality	Longitude	Latitude
Germany, Bavaria	11,8823	48,249
Germany	No data	No data
Wadi Kelt (Quilt), Israel	No data	No data
Stara Planina, Bulgaria	No data	No data
Crete 1	24,85584	35,30808
Crete 2	24,83732	35,31051
River near Velky Folkmar, Slovakia	21,01145	48,84588
Stream near Zvolen, Slovakia	19,1397	48,5592
River Muran, Jelsava, Slovakia	20,23907	48,62386
Stream in Selec, Slovakia	17,99599	48,77677
Stream in Krnca, Slovakia	18,26541	48,53594
Stream in Pdhradie, Slovakia	18,63131	48,66155
Vrica Stream in Predvrecko, Slovakia	18,72669	48,96505
Moravia, near Brno, Blanska, River Punkva banks, Czech Republic	16,7403	49,4173
Jasénka (Kotrlé), Czech Republic	18,0233	49,3781

Locality	Longitude	Latitude
Pardubický, Czech Republic	16,8624	50,1499
Mala Bystricka Stream, Czech Republic	18,05408	49,3948
Stream Knehyne in Prosteredni Becva, Czech Republic	18,27853	49,46681
Zabniczanka stream, Poland	19,19397	49,55338
Zylica stream in Szczyrk, Poland	18,98406	49,69355
UA14J02.1a.1, Russia	29,9294	51,3885
UA14J05.2.1,Russia	30,12828	51,41435
UA14J06.2b.1, Russia	29,7038	51,48437
UA14J04.2a.1, Russia	30,1362	51,35207
RU12728.1.1, Russia	53,5102	63,5476
Derbyshire, United Kingdom	-1,4900	52,766
Wales, United Kingdom	-3,2278	51,4499
Yukari Dueden Selalesi waterfall, N of Antalya	No data	No data
Scheldt River, near Wintam, Jan Soors, Belgium	4,3047	51,1073
Rossum, The Netherlands	5,2980	51,793

Supplementary Table 5. Localities sampled in several countries by Christer Érseus and co-workers (blue), Emmanuelle Lapied (purple), Aleksandra Jablonska and Misel Jelic (green), Csaba Csuzdi and co-workers (black) and Marta Novo (grey).

Species	Accesion number COI	Accesion number 16S	Accesion number 28S
Eiseniella tetraedra	OL457985-OL458610		OL471766-OL471904
Carpetania matritensis	GQ409661.1	JN209218.1	GQ409652.1
Lumbricus rubellus	KM611946.1	KJ912567.1	KJ912213.1
Dendrobaena byblica	Domínguez et al. 2015; Pérez-Losada et al. 2015	KJ912523.1	Domínguez et al. 2015; Pérez-Losada et al. 2015
Iberoscolex oliveirae	Domínguez et al. 2015; Pérez-Losada et al. 2015	KJ912555.1	Domínguez et al. 2015; Pérez-Losada et al. 2015
Prosellodrilus biariculatus	Provided by the authors	KJ912586.1	JN871950.1

Supplementary Table 6. Accesion numbers an references of all sequences used.

	AUC	omission.rate	sensitivity	specificity	prop.correct	Kappa	
Clade I	0,71	0,28	0,72	0,71	0,72	0,41	
Clade II	0,79	0,21	0,79	0,80	0,79	0,57	
	awc	**bio13**	**bio6**	**bio7**	**clc**	**crust**	**parma**
Clade I	15,09	14,65	15,47	16,97	15,06	11,95	10,80
Clade II	15,78	7,94	27,41	9,69	7,38	13,27	18,54

Supplementary Table 7. Above: comparison of the performance of the models for clades I and II. Below: relative contribution of the predictor variables to each model.

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidrucal funnels	Observations
30000	35	143	83	22-27	23-24-25	13	9-10-11-12	not present	not present	
30001	34	140	56	22-27	23-24-25	13	9-10-11-12	not present	not present	
30002	47	220	90	22-27	23-24-25	13	9-10-11-12	not present	very small	
30003	34.5	141	82	22-27	23-24-25	13	9-10-11-12	not present	very small	
30004	31	99	54	22-26	23-24-25	13	9-10-11-12	not present	not present	
30005	28	100	81	22-27	23-24-25	13	9-10-11-12	not present	very small	
30006	33	115	83	22-1n27	23-24-25-1n26	13	10-11--12	not present	maybe on 10 without sperm	
30007	16.5	40	46	22-26	23-24-25	13	9-10-11-12	not present	very small	
30008	15	56	52	23-26	23-24-25	13	8-10-11-12	not present	maybe on 10 without sperm	
30009	27	78	89			13				semi-mature
30010	25	70	92		23-24-25					semi-mature
30011	16	27	88							inmature
30012	18	15	80							inmature
30013	14	9	74							inmature
30014	11	10	87							inmature
30015	12.5	14	82							inmature
30016	17.5	80	89	22-26	23-24-25	13	10-11--12	not present	not present	the earthworm was broken
30017	15.5	44	92	22-26	23-24-25	13	9-10-11-12	not present	very small	the earthworm was broken
30018	20	45	79							inmature
30019	21.5	49	85							inmature
30020	24	92	76	22-26	23-24-25	13	9-10-11-12	very small	very small	
30021	22	85	85	22-27	23-24-25-1n26	13	9-10-11-12	not present	maybe on 10 without sperm	
30022	21	63	72	1n22-1n27	23-1n26	13	9-10-11-12	without sperm	not present	
30023	24	78	88	22-27	23-24-25	13	9-10-11-12	not present	not present	

UCM-LT number	**Length (mm)**	**Dry weight (mg)**	**Number of segments**	**Clitellum**	**Tubercula pubertatis**	**Male pore**	**Seminal vesicles**	**Spermatechae**	**Spermidical funnels**	**Observations**
30024	23.5	92	62	22-27	23-24-25	13	8-10-11-12	very small	very small	
30025	24	68	83	22-26	23-24-25-26	13	10-11--12	not present	maybe on 10 without sperm	spermatophore without sperm between 11 and 12
30026	26.5	99	85	22-27	23-24-25-26	13	10-11--12	not present	without sperm on left side, not present on right side	
30027	19	53	84	23-26	23-24-25	13	10-11--12	not present	not present	spermatophore without sperm on 11
30028	21	49	65		23-24-25	13				semi-mature
30029	21.5	50	84	22-27	23-24-25	13	9-10-11-12	not present	very small	
30030	19	44					11 -- 12	not present	very small	the earthworm was broken
30031	21	73	65	22-27	23-24-25-1n26	13	left side: 11-12 right side:10-11--12	not present	without sperm	
30032	22	68	65	22-27	1n22-23-24-25	13	left side: 11-12 right side: 10-11--12	not present	without sperm	
30033	19	42	69	22-26	23-24-25	13	9-10-11-12	not present	without sperm	
30034										
30035	15	26	65	22-27						inmature
30036	20	56	71	20-25	21-22-23-24	11	8-9--10	not present	without sperm	
30037	21.5	54	82	18-24	20-21-22-1n23	9	left side: 8-7--6 right side: 8-6	not present	without sperm	
30038	23.5	73	74	22-28	23-24-25-26	13	9-10-11-12	not present	without sperm	
30039	27	86	78	22-27	23-24-25-1n26	13	9-10-11-12	not present	very small	
30040	22	86	80	22-27	23-24-25	13	9-10-11-12	not present	very small on left side	
30041	26	85	86	22-27	23-24-25	13	9-10-11-12	not present	very small	
30042	19	53	88	22-27	23-24-25-1n26	13	11--12	not present	without sperm	
30043	24	81	80	22-27	23-24-25	13	9-10-11-12	not present	not present	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidlucal funnels	Observations
30044	27	91	88	22-27	23-24-25	13	10--11--12	not present	maybe on 10 without sperm	
30045	23.5	80	85	23-27	23-24-25-1n26	13	10-11--12	without sperm	without sperm	
30046	27	86	87	22-27	23-24-25	13	9--12	without sperm	without sperm	
30047	23	62	80	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30048	26	90	85	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30049	24.5	71	91	22-27	23-24-25	13	10--11-12	without sperm	without sperm	
30050	25.5	85	83	22-27	23-24-25	13	10--11--12	without sperm	without sperm	
30051	24	80	84	22-27	23-24-25	13	10--11--12	not present	without sperm	
30052	25	79	82	22-27	23-24-25-1n26	13	9-10-11-12	without sperm	without sperm	
30053	22.5	65	81	20-26	23-24-25 (right), 20 to 26 left	13	9-10-11-12	without sperm	not present	
30054	25.5	85	88	22-27	23-24-25-1n26	13	9-10--12	without sperm	very small	
30055	25	70	84							inmature
30056	21	38	88							inmature
30057	22.5	43	84							inmature
30058	20.5	60	83							inmature
30059	22	43	86							inmature
30060	25	71	83							inmature
30061	22.5	59	79							inmature
30062	20.5	49	81							inmature
30063	19.5	57	80							inmature
30064	16	28	80							inmature
60065	17	25	76							inmature
30066	22	49	82							inmature
30067	20.5	45	89							inmature

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidiucal funnels	Observations
30068	14.5	24	68							inmature
30069	19.5	29	87							inmature
30070	16	30	77							inmature
30071	17	27	88							inmature
30072	24	108	82	22-28	23-24-25	13	9--10-12	not present	maybe on 10 without sperm	
30073	23	67	83	22-27	23-24-25	13	9-10-11-12	one on the left side without sperm	very small	
30074	28.5	50	55	22-26	23-24-25	15	9--11-12	one on the left side without sperm	very small	
30075	17	21	83							inmature
30076	17	24	82							inmature
30077	20	33	87							inmature
30078										the earthworm was broken, inmature
30079	12	15	52							inmature
30080	14.5	17	83							inmature
30081	13	12	62							inmature
30082	17.5	33	63							inmature
30083	25	63	85	22-26	23-24-25-1n26	13	11 -- 12	not present	very small	
30084	23	51	91	22-27	23-24-25-1n26	15	9 -- 12	without sperm	very small	
30085	18	42	87							inmature
30086	14	15	90							inmature
30087	11.5	11	89							inmature
30088										inmature
30089	19	31	87		23-24-25					semi-mature
30090	15	12	88							inmature
30091				22-27	23-24-25	13	9-10-11-12	not present	not present	the earthworm was broken

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidecal funnels	Observations
30092	29.5	106	86	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30093				22-27	23-24-25	13	9--11--12	not present	not present	the earthworm was broken
30994	33	94	79	22-27	23-24-25	13	9-10-11-12	without sperm	very small	
30095	20	49	68	22-27	23-24-25-1n26	13	8-10-11-12	not present	not present	
30096	34.5	113	88	22-26	23-24-25-1n26	13	11 -- 12	not present	not present	
30097	27.5	66	83	22-27	23-24-25-26	13	11 -- 12	without sperm	not present	
30098	31.5	115	82	22-1n28	23-24-25	13	9--11--12	not present	not present	
30099	29.5	82	87	22-27	23-24-25-1n26	13	9-10-11-12	not present	very small	
30100	30	100	73	22-28	23-24-25-26	13	10--11--12	without sperm	not present	
30101	17	26	86							
30102	20	46	78							
30103	18	28	81							
30104	19	57	69	22-27	23-24-25	13	9--11--12	without sperm	not present	
30105	19.5	36	87			13				semi-mature
30106										inmature the earthworm was broken
30107	10	8	90							inmature
30108	19	33	103							inmature
30109	16	19	69							inmature
30110	10	5	90							inmature
30111	35	176	85	22-27	23-24-25	15	9--11-12	not present	not present	
30112	35	232		22-27	23-24-25	13	9--11--12	without sperm	very small	the earthworm was broken
30113	35	194	89	22-27	23-24-25-26	13	9-10-11-12	not present	very small	
30117	22	100	52	22-26	23-24-25	15	9--11--12	without sperm	very small	
30118	19.5	61	60	22-27	23-24-25	15	9--11--12	not present	very small	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidiucal funnels	Observations
30119	33.5	143	77	22-27	23-24-25	15	9--11--12	without sperm	very small	
30120	17.5	30	84							inmature
30121	26	129	77	22-27	23-24-25-1n26	13	9--11--12	not present	very small	
30122	26.5	115	86	22-27	23-24-25-1n26	13	9--11--12	not present	not present	
30123	25	108	79	22-27	23-24-25	13	9--11--12	not present	very small	
30124	28	128	93	22-27	23-24-25	13	9--11--12	not present	not present	
30125	33	168	72	22-27	23-24-25	13	9--11--12	without sperm	very small	
30126	29	150	84	22-28	23-24-25	13	9--11--12	without sperm	very small	
30127	16.5	74	46	22-26	23-24-25	13	9--11--12	not present	maybe on 10 without sperm	
30128	32	150	82	22-27	23-24-25	13	9--12	without sperm	maybe on 10 without sperm	
30129	28.5	114	85	22-27	23-24-25	13	9-10-11-12	without sperm	maybe on 10 without sperm	
30130	23	105	65	22-26	23-24-25-1n26	13	9--11--12	not present	not present	
30131	21	63	79			13				semi-mature
30132	20	47	82			13				semi-mature
30133	18	45	84							inmature
30134	13.5	20	77							inmature
30135	24.5	86	89	22-27	23-24-25-1n26	13	10--11--12	not present	not present	
30136	23	44	83	22-27	23-24-25-26	13	10--11--12	not present	not present	
30137	20	57	69		23-24-25	13				semi-mature
30138	21	98	59	22-27	23-24-25-26	13	9--11--12	without sperm	very small	
30139	32	85	81	22-27	23-24-25	13	9-10-11-12	not present	not present	
30140	16	26	84							inmature
30141	27	81	82							inmature
30142	15	16	84							inmature

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidducal funnels	Observations
30143	22.5	82	59	22-27	23-24-25-1n26	13	9--11--12	not present	maybe on 10 without sperm	
30144	28	76	75	22-27	23-24-25-1n26	13	8-10-11-12	not present	not present	
30145	36	145	83	22-27	23-24-25	13	9-10-11-12	not present	not present	
30146	31.5	90	84	22-27	23-24-25-1n26	13	9--11--12	not present	not present	
30147	15	17	41	22-27	No presentes	13	9--12	without sperm	very small	
30148	32	92	81	22-27	1n22-23-24-25-1n26	13	9-10-11-12	without sperm	not present	
30149	30	96	78	22-27	23-24-25-1n26	13	9-10-11-12	without sperm	very small	
30150	38	134	83	22-27	23-24-25	13	9-10-11-12	not present	very small	
30151	29.5	78	78	22-27	23-24-25	13	9-10-11-12	one on the right side without sperm	not present	
30152	30	83	65	22-27	23-24-25-26	13	9-10-11-12	not present	not present	
30153	28.5	105	63	22-26	23-24-25	13	9--11--12	without sperm	very small	
30154	20	56	51	22-26	23-24-25	13	9-10-11-12	not present	very small	
30155	32.5	121	85	22-27		13	10--11--12	not present	not present	
30156	28	93	72	22-27	23-24-25	13	9-10-11-12	not present	very small	
30157	26	53	78	22-27	23-24-25-1n26	13	9-10-11-12	without sperm	not present	
30158	26.5	67	56	22-26	23-24-25-1n26	13	9-10-11-12	without sperm	not present	
30159	36.5	133	80	22-26	23-24-25	13	9-10-11-12	not present	not present	
30160	24.5	66	69	22-27	26-24-25-1n26	13	9-10-11-12	not present	not present	
30161	23	69	55	22-27	23-24-25	13	9-10-11-12	not present	not present	
30162	22	39	84							semi-mature
30163	27	56	84	22-27	23-24-25	13	10--11--12	not present	not present	
30164	28	64	78	22-27	23-24-25-26	13	9-10-11-12	without sperm	not present	
30165	31	100	84	22-27	23-24-25-1n26	13	9-10-11-12	not present	not present	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidiucal funnels	Observations
30166	25	85	62	22-26	23-24-25-1n26	13	9-10-11-12	not present	very small	
30167	29.5	79	70	22-27	23-24-25-1n26	13	9-10-11-12	without sperm	not present	
30168	21.5	53	68	22-27	23-24-25	13	9--11--12	without sperm	very small	
30169	23	52	51	22-26	23-24-25-1n26	13	9-10-11-12	not present	very small	
30170				22-27	23-24-25-1n26	13	11--12	not present	very small	the earthworm was broken
30171	19	25	83							inmature
30172	18.5	20	82							inmature
30173	15	19	64							inmature
30174	13	14	74							inmature
30175	22	64	87	22-27	23-24-25	13	10--11--12	without sperm	not present	
30176	15.5	56	33	22-27	23-24-25-26	13	9-10-11-12	without sperm	very small	
30177	41.5	159	81	22-28	23-24-25-26	13	9-10-11-12	not present	not present	
30178	22.5	61	82	22-28	24-25-26	13	9-10-11-12	without sperm	not present	
30179	29	114	81	22-27	24-25-26	13	9-10-11-12	without sperm	not present	
30180	28	90	82	22-28	24-25-26	13	9--11--12	without sperm	maybe on 10 without sperm	
30181	28.5	67	75	22-27	23-24-25-26	13	9-10-11-12	not present	not present	
30182	21.5	46	86	22-27	23-24-25	13	10--11--12	without sperm	not present	
30183				22-27	23-24-25-1n26	13	9-10-11-12	without sperm	maybe on 10 without sperm	the earthworm was broken
30184	11	7	80							inmature
30185										inmature
30186	11.5	8	86							inmature
30187	12	9	79							inmature
30188	18.5	30	84							inmature
30189										inmature

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidical funnels	Observations
30190	17	19	80							inmature
30191	12	9	73							inmature
30192	24	106	82		23-24-25	13				semi-mature
30193	20	64	80	22-27	23-24-25-26	13	9-11--12	without sperm	not present	
30194	16.5	54	61			13				semi-mature
30195	17	31	86							inmature
30196	17.5	48	85							inmature
30197	19	61	71							inmature
30198	28.5	106	76	22-27	23-24-25-26	13	9-10-11-12	without sperm	not present	
30199	26.5	134	73	22-27	23-24-25	13	9-10-11-12	without sperm	not present	
30200	27	101	84	22-27	23-24-25	13	8-9-10-11-12	without sperm	not present	
30201	28	134	77	22-27	23-24-25	13	9--11--12	without sperm	not present	
30202	28.5	131	83	22-27	23-24-25-26	13	9-10-11-12	without sperm	very small	
30203	26.5	142	76	22-27	23-24-25	13	9-10-11-12	without sperm	very small	
30204	27	122	81	22-27	23-24-25-1n26	13	9--11--12	without sperm	not present	
30205	26.5	115	81	22-27	23-24-25-1n26	13	9--11--12	without sperm	very small	
30206	24	91	84	22-27	23-24-25	13	9--11--12	without sperm	not present	
30207	21	74	65	22-27	23-24-25-1n26	13	10--11--12	without sperm	not present	
30208	25	106	80	22-27	23-24-25-1n26	13	10--11--12	without sperm	not present	
30209	23	117	59	22-27	23-24-25	13	9-10-11-12	without sperm	not present	
30210				22-27	23-24-25	13	9--11--12	without sperm	very small	the earthworm was broken
30211				22-26	23-24-25	13	8-10-11-12	without sperm	not present	the earthworm was broken
30212	25.5	116	82		23-24-25	13				semi-mature
30213	25	96	80		23-24-25	13				semi-mature

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
30214	24	99	81		23-24-25-1n26	13				semi-mature
30215	21.5	90	80		23-24-25-1n26	13				semi-mature
30216	24	95	83		23-24-25-1n26	13				semi-mature
30217	21	86	66		23-24-25-1n26	13				semi-mature
30218	20.5	85	63		23-24-25-1n26	13				semi-mature
30219	17	59	60		23-24-25-1n26	13				semi-mature
30220	20	62	77		23-24-25-1n26	13				semi-mature
30221	23.5	93	81		23-24-25	13				semi-mature
30222	17	66	75		23-24-25-1n26	13				semi-mature
30223					23-24-25	13				semi-mature
30224	21.5	68				13				semi-mature and the earthworm was broken
30225	22	72	81			13				semi-mature
30226	12.5	44	65							inmature
30227	19.5	74	69							inmature
30228	16.5	38	84							inmature
30229	21.5	60	96							inmature
30230	18.5	36	76							inmature
30231	17.5	41	83							inmature
30232	24.5	87	81	22-26	23-24-25	13	9--11--12	not present	not present	
30233	21	47	83		24-25	13				semi-mature
30234	21	41	84	22-26	23-24-25	13	11 -- 12	not present	not present	
30235	19.5	35	75	22-25	1n22-23-24-1n25	13	11 -- 12	not present	not present	
30236				22-27	24-25-26	13	9-10-11-12	not present	very small	broken

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
30237	22	82	64	22-27	23-24-25	13	10--11--12	without sperm	very small	
30238	27	75	86	22-27	23-24-25-1n26	13	11 -- 12	without sperm	not present	
30239	21.5	84	58	22-27	23-24-25-1n26	13	9--11--12	not present	not present	
30240	20	69	63	22-27	23-24-25-1n26	13	9-10-11-12	without sperm	not present	
30241	22	104	65	22-28	23-24-25-26	13	9-10-11-12	without sperm	very small.	
30242	24	70	85	22-27	23-24-25	13	9-10-11-12	without sperm	not present	
30243	25	75	87	22-27	1n23-24-25-26	13	9--11--12	without sperm	not present	
30244	22	65	80	22-27	23-24-25	13	9-10-11-12	without sperm	very small	
30245	22.5	62		22-27	23-24-25	13	9--11--12	without sperm	very small	the earthworm was broken
30246	22	67	88	22-28	24-25-26	13	9-10-11-12	without sperm	not present	
30247	29	100	90	22-28	23-24-25-1n26	13	9-10-11-12	not present	not present	
30248	19	71	66	22-28	23-24-25-1n26	13	9--11--12	without sperm	not present	
30249	29.5	79	89	22-28	23-24-25-1n26	13	9--11--12	not present	not present	
30250	28.5	102	70	22-28	23-24-25-1n26	13	9--11--12	not present	maybe on 10 without sperm	
30251	26	87	88	22-28	23-24-25	13	11 -- 12	not present	not present	
30252				22-27	23-24-25-26	13	9-11--12	not present	not present	the earthworm was broken
30253	27	107	75	22-28	24-25-26	13	9--11--12	not present	not present	
30254	41.5	161	86	22-28	24-25-26	13	9--11--12	not present	not present	
30255	50.7	250	94	22-27	23-24-25	13	9--11--12	one on the right side without sperm	not present	
30258	28	109	66	22-1n27	23-24-25-1n26	13	9-10-11-12	not present	not present	
30259	40	172	92	22-26	23-24-25	13	9-10-11-12	one on the left side without sperm	not present	
30260	33.5	117	88	22-27	23-24-25-26	13	9-10-11-12	not present	not present	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidiucal funnels	Observations
30262	20.5	39.3	84		23-24-25-26	13				semi-mature
30263	16.5	22	80							inmature
30264	12	10	87							inmature
30265				22-28	24-25-26	13	9-10-11-12	not present	not present	the earthworm was broken
30266	23	71	88	22-27	23-24-25	13	9-10-11-12	one on the left side without sperm	maybe on 11 without sperm	
30267	22	67	83	22-27	23-24-25	13	9--10--12	not present	without sperm	
30268	21.5	46	89			15				semi-mature
30269	25	92	85	22-27	24-25-26	13	9--11--12	one on the right side without sperm	not present	
30270				22-27	23-24-25-1n26	15	9--11--12	without sperm	not present	the earthworm was broken
30271	25	76	87	22-27	24-25-26	13	10--11--12	not present	not present	
30272	14	40	54	22-27	23-24-25	13	9--10--11--12	not present	not present	
30273	26	112	86	22-27	23-24-25-1n26	13	9-10-11-12	not present	very small	
30274	16	34	62		23-24-25-1n26	13				semi-mature
30275	25.5	96	81	23-28	24-25-26	14	9-10-11-12	without sperm	not present	
30276	26.5	110	84	22-27	23-24-25-26	13	9--11--12	without sperm	maybe on 10 without sperm	
30277	20	59	84			13				semi-mature
30278	18.5	36	78							inmature
30279	17	38	84		24-25-26	13				semi-mature
30280	34	114	89	22-28	23-24-25-26	13	12 -- 11	not present	not present	
30281	23.5	116	73	22-28	23-24-25-1n26	13	9--11--12	not present	not present	
30282	29.5	141	86	22-27	23-24-25-26	13	9--11--12	not present	not present	
30283	23	80	96	22-28	23-24-25-1n26	13	9-10-11-12	not present	not present	
30284	30	150	90	22-28	23-24-25-1n26	13	9-10-11-12	without sperm	without sperm	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
30285				22-27-	23-24-25-1n26	13	9-10-11-12	without sperm	without sperm	the earthworm was broken
30286	29	70	81	22-28	23-24-25-26	13	8-9-11-12	without sperm	without sperm	
30287	24.5	60	87	22-26	23-24-25	13	9-10-11-12	not present	not present	
30288	23	100	66	22-28	23-24-25-26	13	9-10-11-12	not present	not present	
30289	12.5	10								inmature and broken
30290	22.5	76	89	22-27	23-24-25	13	9-10-11-12	not present	not present	
30291	23	104	89	22-27	23-24-25-1n26	13	9-10-11-12	not present	not present	
30292	31	122		22-28	24-25-26	13	9-10-11-12	not present	not present	
30293	22.5	98	58	22-27	23-24-25	13	9-10-11-12	without sperm	very small	
30299	25	113.5	86	22-27	24-25-26	13	10--11--12	not present	not present	
30300	24	113.7	84	22-27	23-24-25-1n26	13	9-10-11-12	not present	not present	
30301	23	72.1	88	22-26	23-24-25	13	11---12	not present	not present	
30302	24.5	108.2	87	22-27	23-24-25-1n26	13	8-10-11-12	not present	not present	
30303	23	97.7	92	22-27	23-24-25-1n26	13	11---12	not present	without sperm	
30304	20	78.6	84	22-26	23-24-25	13	9-10-11-12	not present	not present	
30305						13	9---11-12	not present	not present	the earthworm was broken
30306	16	43.3	74			13				semi-mature
30307	37	93.5	81	21-26	23-24-25	12	9--11-12	not present	not present	
30308	22	45.5	73	20-24	22-23-24	11	9-11--12	without sperm	without sperm	
30309	27	24.4	45			13				semi-mature
30310	14	29.2	71			13				semi-mature
30311	16	6.9	38							inmature
30312	14	12.4	81							inmature
30313	17.5	54.6	72	22-26	23-24-25	13	9--11--12	not present	not present	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidučal funnels	Observations
30314	12.5	16.2	74							inmature
30315	14	19.2	91							inmature
30316	9	7.1	85							inmature
30317	20.5	58.5	79	22-27	23-24-25	13	9---11-12	not present	not present	
30318	19	41.1	85	22-27	23-24-25-1n26	13	9--11--12	not present	not present	
30319	16.5	35.7	61	22-27	23-24-25	13	11---12	not present	not present	
30320	16	50.5	56	22-27	23-24-25-26-27	13	9--11--12	not present	not present	
30321										
30322										
30323										
30324										
30325	23.5	97.5	73	22-27	23-24-25	13	9--10	not present	not present	
30326	30	172	76	22-27	26-24-25-26	13	9-10-11-12	not present	not present	
30327	31	83	91	22-28	23-25-24-25-1n27	13	9-10-11-12	not present	not present	
30328	24	117.2		22-27	23-24-25	13	11--12	not present	not present	
30329	19	39.7	93		23-24-25-1n26	13 (and 14?)				semi-mature
30330	23.5	66.6	83	22-27	23-24-25-26	13	9--11-12	without sperm	not present	
30331	19	58.7	76			13				semi-mature
30332	28	109.1	78	22-27	23-24-25	13	11--12	not present	not present	
30333	22	101.8	60	22-26	23-24-25	12	9--11--12	not present	not present	
30334	27	138.1	82	22-26	23-24-25	13	11--12	without sperm	not present	
30335				22-28	23-24-25	13	9--11--12	not present	not present	the earthworm was broken
30336	26	109.7	80	22-27	24-25-26	13	9-10-11-12	not present	not present	
30337				22-28	23-24-25	13	9-10-11-12	not present	not present	the earthworm was broken

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidual funnels	Observations
30338				22-27	23-24-25-26	13	8-10-11-12	not present	not present	the earthworm was broken
30339	24	101.7	92	22-27	24-25-26	13	9-10-11-12	not present	not present	
30340										the earthworm was broken, semi-mature
30341	11	14.8	85							inmature
30342				22-27	23-24-25-26	13	10--11--12	not present	not present	the earthworm was broken
30343	19	61.4	81		22-23-24	13				semi-mature
30344						13	9-10-11-12	without sperm	without sperm	the earthworm was broken
30345	14.5	58.2	40	22-27	1n23-24-25-1n26	13	9-10-11-12	without sperm	without sperm	the earthworm was broken
30346						13				the earthworm was broken
30348	24	59.5	78	22-27	23-24-25-1n26	13	9-10-11-12	not present	not present	
30349	24.5	72.5	80	22-27	23-24-25-1n26	13	11 -- 12	not present	not present	
30350	21	47.6	101		23-24-25	13				semi-mature
30351	32	106	78	22-27	23-24-25-1n26	13	9--11--12	not present	not present	
30352	20	73.8	95	22-27	23-24-25-26	13	9-10-11-12	not present	not present	
30353	20	54.3	97	22-27	23-24-25-26	13	9-10-11-12	not present	not present	
30354	18	31.2	89							inmature
30355	26.5	94.3	97	22-27	23-24-25	13	9-10-11-12	not present	not present	
30356	22	55.3	85	22-27	23-24-25-26	13	9-10-11-12	not present	not present	
30357	20	58.4	82							inmature
30358	21.5	34.2								inmature
30360	24	102.5	67	22-28	24-25-26	13	9-10-11-12	not present	not present	
30361	19.5	55.5	78	21-28	22-23-24-25-26	13	9-10-11-12	not present	not present	
30362	21.5	61.5	96	22-27	23-24-25-26	13	11 -- 12	not present	not present	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
30363	23.5	73.4	77	21-27	23-24-25	13	9-10-11-12	not present	not present	
30364	24	92.3	97	22-27	23-24-25-26	13	9--11--12	one on the right side without sperm	not present	
30365	28	94	87	22-27	23-24-25-1n26	13	9--11--12	not present	not present	
30366	24	109.7	62	22-28	24-25-26	13	12 -- 11	not present	not present	
30367	28	103	97	22-27	23-24-25	13	12 -- 11	not present	not present	
30368	21	68.2	60	22-26	23-24-25	13	13--11--12	not present	not present	
30369	19.5	87.8	58	22-28	23-24-25-26	13	9-10-11-12	one on the right side without sperm	not present	
30370	16.5	24.2	92							inmature
30371	19	54.9	76	22-27	23-24-25	13	9 --12	without sperm	not present	
30372	16	36.7	71			9				semi-mature
30373	18	27.1	74							inmature
30374	22.5	71.9	98	22-28	23-24-25-26	13	9-10-11-12	without sperm	without sperm	
30375	13.5	13.9	73							inmature
30376	12	11.8	77							inmature
30377	29.5	99	59	22-26	23-24-25	13	9--11--12	not present	not present	
30378	20.5	43.1	92	22-27	24-25-26	13	10--11-12	not present	without sperm	
30379	32.5	118.8	84	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30380	22	102.2	58	22-27	23-24-25	13	12	without sperm	not present	
30381	25.5	98.7	93	22-27	23-24-25	13	12 -- 11	without sperm	without sperm	
30382	32	128.4	82	23-28	24-25-26	13	12 -- 11	not present	not present	
30383	23	97.8	68	22-28	23-24-25	13	9-10-11-12	one on each side without sperm	without sperm	
30384	35.5	158.8	86	21-26	1n22-23-24-1n25	12	12	not present	not present	
30385				22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	the earthworm was broken
30386	29	123.7	80	22-27	23-24-25	13	9--11--12	not present	not present	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
30387	24.5	116.1	83	22-28	24-25-26	13	9-10-11-12	one on each side without sperm	without sperm	
30388	32.5	148.8	86	22-26	23-24-25	13	9--11--12	one on each side without sperm	not present	
30389	22.5	83.7	79	22-27	23-24-25-26	13	9--11--12	without sperm	without sperm	
30390	21.5	82.5	60	22-27	23-24-25-26	13	9 -- 12	without sperm	without sperm	
30391	10	8.8	83							inmature
30392	6	4.5	67							inmature
30399	34.5	227.2	71	22-27	23-24-25	13	9--11--12	not present	not present	
30401	40	289.7	93	22-28	23-24-25-1n26	13	9-10-11-12	one on each side without sperm	not present	
30402	39	260.6	86	22-28	24-25-26	13	10--11--12	one on right side without sperm	without sperm	
30403	30	150.7	82	22-27	1n23-24-25-1n26	13	9--11--12	not present	not present	
30404				22-27	23-24-25	13	10--11-12	not present	not present	the earhtworm was broken
30405	30.5	158.4	83	22-27	24-25-26	13	9-10-11-12	one on right side without sperm	without sperm	
30406	26.5	101.2	84			13				semi-mature
30407	23	81.1	91	22-26	23-24-25	13	9-10-11-12	one on left side without sperm	without sperm	
30408	38.5	204.3	82	22-28	24-25-26	13	12	not present	without sperm	
30409	32	123.1		21-25	22-23-24	12	9-11--12	not present	without sperm	the earthworm was broken
30410	22	99		22-28	23-24-25	13	9-10-11-12	not present	without sperm	the earthworm was broken
30411				22-27	24-25-26	13	9-10-11-12	not present	not present	the earthworm was broken
30412	22	61	91							inmature
30413	30	93.2	74	22-28	23-24-25	12	9-10-11-12	without sperm	without sperm	
30414	14	30.8	69							inmature
30425	26.5	119.3	90	22-28	23-24-25-26-27	13	9-10-11-12	without sperm	without sperm	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidrucal funnels	Observations
30426	19	87.1	79		23-24-25	13				semi-mature
30427	18.5	85.6	75	22-28	24-25-26-27	13	9--11--12	without sperm	not present	
30428				22-27	24-25-26	13	9--11--12	without sperm	not present	
30429	13	74.1	81			13				semi-mature
30434	25	77.6	88	22-27	23-24-25	13	9-10-11-12	one on right side without sperm	not present	
30435	19	49	61	22-27	23-24-25	13	9-10-11-12	without sperm	not present	
30436	21	46.2	70		24-25-26	13				semi-mature
30437	27	79.4	76	22-27	23-24-25	13	9-10-11-12	not present	not present	
30438	17	69.7	59	19-23	20-21-22	9	9-10-11-12	one on left side without sperm	not present	
30439	19	58.3	80	22-28	23-24-25	13	9-10-11-12	without sperm	without sperm	
30440	26	87.6	70	22-27	23-24-25	13	9--11--12	not present	not present	
30441	22	102.8	88	22-27	23-24-25	13	9--11--12	without sperm	without sperm	
30442	25	93.2	91	21-25	21-22-23-24	13	9-10-11-12	one on right side without sperm	not present	
30443	24.5	87.5	84	23-28	24-25-26	13	10--11-12	not present	not present	
30444	25	90.2	89	22-27	23-24-25	13	9--11--12	without sperm	not present	
30445	26	97.2	91	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30446	27	94.1	78	22-28	23-24-25	13	10--11--12	not present	not present	
30447	23.5	58.8	74			13				semi-mature
30448	24	91.6	90	22-27	23-24-25	13	9-10-11-12			
30449	22.5	73.8	79	22-27	23-24-25	13	9-10-11-12	not present	not present	
30450	26	88.6	66	22-27	23-24-25	13	9-10-11-12	one on right side without sperm	not present	
30451	22.5	64.4	77	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30452	25.5	101.8	72	22-28	23-24-25	12	9-10-11-12	one on right side without sperm	not present	
30453	22.5	51.3	103	22-28	23-24-25	13	9-10-11-12	without sperm	without sperm	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
30454	24.5	88.7	76	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30455	25	91.3	55	22-27	23-24-25	13	9--11--12	without sperm	not present	
30456	22	79.9	81	22-27	23-24-25	13	9-10-11-12	without sperm	not present	
30457	21	68.7	65	22-27	23-24-25	13	9-10-11-12	not present	not present	
30458	23	91.3	83	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30459	26	86.7	77	22-28	24-25-26	13	9-10-11-12	without sperm	without sperm	
30460	23.5	91.4	80	22-27	23-24-25	13	9-10-11-12	without sperm	not present	
30461	24	69.9	89	22-27	23-24-25	13	9-10-11-12	not present	not present	
30462	23.5	73	73	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30463				22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	the earthworm was broken
30464	21.5	65.3	82	23-28	24-25-26	13	9--11--12	without sperm	not present	
30487	24	70.2	68	22-27	23-24-25	13	11 -- 12	not present	not present	
30488	26	61.9	62	22-27	23-24-25	13	11 -- 12	not present	not present	
30489	21.5	54.8	52	22-26	23-24-25	13	11 -- 12	not present	not present	
30490				22-28	23-24-25	13	10--11--12	without sperm	not present	the earthworm was broken
30491				22-28	24-25-26	13	10--11--12	no presents	not present	the earthworm was broken
30492	26	73.4	79	22-28	24-25-26	13	11 -- 12	not present	not present	
30493	23.5	55.1	85	22-28	24-25-26	13	11 -- 12	not present	not present	
30494	16	43.4	56	22-26	23-24-25	13	9-10-11-12	without sperm	without sperm	
30496	17.5	49.7	70	22-27	23-24-25	13	9--11-12	not present	not present	
30497						13	11 -- 12	not present	not present	the earthworm was broken
30498	29	93.5	82	22-28	23-24-25	13	10--11--12	not present	not present	
30499				22-28	23-24-25	13	11--12	not present	not present	the earthworm was broken
30500	25	79.1	80	22-27	23-24-25	13	10--11--12	not present	not present	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
30501	22	81.3	85		23-24-25	13				semi-mature
30502	16	41.4	77							inmature
30503										inmature, the earthworm was broken
30504	18	54.1	68							inmature
30505	17	41.4	72	22-27	23-24-25	13	10--11--12	not present	not present	
30506	18	59.5	71	22-27	24-25-26	13	11 -- 12	not present	not present	
30507	21	62.4	80	22-27	23-24-25	13	11 -- 12	not present	not present	
30508	12	20.1	69							inmature
30509	14	19.1	75							inmature
30510	11	6.5	86							inmature
30511	12.5	12.7	68							inmature
30512	14	19.5	70							inmature
30513										posterior part of the body
30514	32	128.1	99	22-27	23-24-25	13	9--11--12	not present	not present	
30515	35	132.4	82	22-27	23-24-25	13	9--11--12	not present	not present	
30516	30	84.5	96	22-28	24-25-26	13	9,11,12	not present	not present	
30517	36.5	144.7	100	22-27	23-24-25-1n26	13	9-10-11-12	not present	not present	
30518	18.5	36	77							inmature
30519	12	52.1	56	22-26	23-24-25	13	9--11--12	without sperm	without sperm	
30520	26	129.7	81	22-27	23-24-25-26	13	9--11--12	without sperm	not present	
30521	14	28.4	82							inmature
30522				22-27	23-24-25	13	9--11--12	not present	not present	the earthworm was broken
30523	16	26.5	92		23-24-25	13				semi-mature
30524				22-27	23-24-25	12	9--11--12			the earthworm was broken

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidical funnels	Observations
30525	16	40.8	95		23-24-25	13				semi-mature
30526						13				the earthworm was broken semi-mature
30527	10	13.2	74							inmature
30528	10.5	11.8	80							inmature
30529										the earthworm was broken e inmature
30530	16	20.6	77							inmature
30531	19.5	40.4	70			13				semi-mature
30532	21	69.4	75	22-27	23-24-25	14	9-10-11-12	without sperm	not present	
30533										inmature the earthworm was broken
30534	12	16.3	72							inmature
30535	15	22.8	96							inmature
30536	10	12.4	82							inmature
30537	18	38.3	88							inmature
CE2518				22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
ce2771				22-29	24-25-26	13	9-10-11-12	without sperm	not present	
ce2772				22-28	23-24-25	13	9-10-11-12	not present	not present	
ce2779										
ce2780						13				
ce2864				22-26	23-24-25	13	9-10-11-12	not present	not present	
ce2865						13				
ce2876				22-27	24-25-26	13	9-10-11-12	not present	not present	
ce2877				22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
ce2878				22-27	23-24-25	13	9-10-11-12	not present	not present	
ce2879				22-27	23-24-25	11	9-10-11-12	one on left side without sperm	not present	
ce3813				23-28	25-26-27	13 , 14	9-10-11-12	without sperm	without sperm	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
ce3814				22-28	24-25-26	13	9-10-11-12	without sperm	not present	
ce4070				22-27	23-24-25	13	9-10-11-12	not present	not present	
ce4207				22-27	23-24-25	13	9-10-11-12	not present	not present	
ce4208				22-27	23-24-25	13	11 -- 12	not present	not present	
ce4209				22-28	23-24-25-26-27-28	13	11 --12	not present	not present	
ce4211				22-27	23-24-25	13	11 -- 12	not present	not present	
ce4213					23-24-25	13				
ce4361				22-27	23-24-25	13	9--11--12	not present	not present	
ce4362				22-27	23-24-25	13	11 -- 12	not present	not present	
ce4395				22-27	23-24-25	13	11 -- 12	without sperm	very small	
ce4396				22-27	23-24-25	13	11 -- 12	not present	not present	
ce4398				22-27	23-24-25	13	11 -- 12	without sperm	without sperm	
ce4399				22-27	13-24-25	13	11 -- 12	not present	not present	
ce4400				22-27	23-24-25	13	11 -- 12	not present	not present	
ce4401				22-27	23-24-25	13	11 -- 12	without sperm	without sperm	
ce4448				22-27	23-24-25	13	11 -- 12	without sperm	without sperm	
ce4454				22-27	23-24-25	13	11 -- 12	without sperm	not present	
ce4455				21-28	24-25-26-27	13	10--11--12	without sperm	without sperm	
ce4456				22-27	23-24-25	13	left side: 11 right side: 10--11--12	not present	not present	
ce4457				23-28	24-25-26	13	11	without sperm	not present	
ce4458				22-27	23-24-25	13	11 --12	without sperm	without sperm	
ce4460					23-24-25	13				
ce4573				22-27	23-24-25	13	9--11--12	not present	not present	
ce4574				22-27	23-24-25	13	11--12	not present	not present	
ce4575				22-26	23-24-25	13	11--12	not present	not present	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidbeing funnels	Observations
ce4587				23-29	24-25-26-27	13	11--12	without sperm	very small	
ce4597				22-27	23-24-25	13	9--11--12	not present	not present	
ce4725						13				
ce4726				22-27	23-24-25	13	11--12	without sperm	without sperm	
ce4727				21-25	22-23-24	13	11--12	without sperm	not present	
ce4728				22-27	23-24-25	13	9-10-11-12	not present	not present	
ce4729						13				
ce4993				22-27	23-24-25	13	11 -- 12	not present	not present	
ce5110				22-27	23-24-25	13	10-11--12	not present	very small	
ce5111				22-27	23-24-25	13	11--12	not present	not present	
ce5112				22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
ce5248						13				
ce5685						13				
ce5711					23-24-25	13				
ce5712				22-27	23-24-25	13	9-10-11-12	not present	very small	
ce5713						13				
ce5714				22-27	23-24-25	13	9-10-11-12	not present	not present	
ce5715				22-27	23-24-25	13	10--11--12	not present	without sperm	
ce5716					23-24-25	13				
ce5717				22-27	23-24-25	13	11--12	not present	not present	
30747	19.5	96.8	54	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	spermatophore between 19 and 20
30748	26.5	148	79	22-26	22-23-24	13	9-10-11-12	without sperm	without sperm	
30749	27.5	136.6	86	23-28	24-25-26	13	9-10-11-12	one on left side without sperm	without sperm	
30750	30	133.7	85	23-28	24-25-26	13	9-10-11-12	without sperm	not present	
30751	34.5	192	90	22-27	23-24-25	13	9-10-11-12	without sperm	not present	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
30752				22-27	23-24-25	13	11	one on each side without sperm	not present	the earthworm was broken
30753	32	150	84	22-27	23-24-25	13	9-10-11-12	not present	not present	
30754	23	87.5	107	22-28	23-24-25-26	13	9-10-11-12	not present	not present	
30755	30.5	167.1	90	22-26	23-24-25-26	13	11 -- 12	not present	not present	
30756	30	144.7	65	22-27	23-24-25	13	9-10-11-12	not present	not present	
30757	27	96.4	79	22-27	23-24-25	13	9-10-11-12	not present	not present	
30758	18	74.5	90							inmature
30650	12	50.3	88			13				semi-mature
30651	19	81.7	80		24-25-26	13				semi-mature
30652						13				semi-mature, the earthworm was broken
30653	15	77	91							inmature
30654	17	71	84							inmature
30655	11	30.8	86							inmature
30656	16	40.5	94							inmature
30657	12	20.8	77							inmature
30658	12.5	19.7	88							inmature
30663	15	24.3	60	22-27	23-24-25	13	9-10-11-12	not present	without sperm	
30667	16	39.1	84	22-27	23-24-25	13	9--11--12	not present	not present	
30668	18	44.2	77	22-27	23-24-25	13	11 -- 12	not present	not present	
30646	23	99.1	75	22-27	23-24-25	13	9--11--12	not present	not present	
30647	15	30.8	90							inmature
30648				22-28	24-25-26	13	9--11--12	without sperm	not present	the earthworm was broken
30649	24	40.4	84			13				semi-mature
30698	14.5	31.1	60	22-28	23-24-25-26	13	11 -- 12	not present	not present	
30699	25	59.7	82	22-28	23-24-25	13	9-10-11-12	not present	not present	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
30700				22-27	23-24-25	13	9-10-11-12	without sperm	not present	the earthworm was broken
30701	15	31.3	67	22-27	23-24-25	13	9-10-11-12	without sperm	not present	
30702	10	22.7	94							inmature
30703				23-28	24-25-26	13	9-10-11-12	without sperm	not present	the earthworm was broken
30538	26	93.6	82	22-27	23-24-25	13	9-10-11-12	without sperm	not present	
30539	25	65.9	56	22-27	23-24-25	13	9-10-11-12	without sperm	not present	
30540	29.5	142.9	79	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30541	27	117.5	95	22-27	23-24-25	13	9-10-11-12	not present	not present	
30542	26.5	110.8	89	22-27	23-24-25	13	9-10-11-12	not present	not present	
30543	32	161.7	94	22-27	23-24-25	13	9-10-11-12	not present	not present	
30544	31	158.7	93	22-27	23-24-25	13	9-10-11-12	without sperm	not present	
30545				22-27	23-24-25	13	9-10-11-12	not present	not present	the earthworm was broken
30546				22-27	23-24-25	13	9-10-11-12	not present	not present	the earthworm was broken
30547	35	147.7	87	22-28	23-24-25-26	13	10--11-12	not present	not present	
30548				22-27	23-24-25	15	9-10-11-12	not present	not present	the earthworm was broken
30599	19.5	44.3	60	22-27	23-24-25	13	10--11-12	not present	not present	
30719				22-28	23-24-25-26	13	9-10-11-12	not present	not present	the earthworm was broken
30720	20	44.2	92	22-29	24-25-26	13	10--11--12	not present	not present	
30721	11	9.4	87							
30722	13	12.1	77							
30723				22-27	23-24-25	13	9-10-11-12	not present	not present	the earthworm was broken
30724	18.5	34.9	81		23-24-25	13				
30725	6	6.9	74							
30726				22-27	23-24-25	13				broken

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
30727	11.5	7.8	83				9-10--11-12	not present	not present	
30728										the earthworm was broken
30729						14				the earthworm was broken
30788	13	8.8	77							inmature
30789	12	8.3	65							inmature
30790	9	5.1	60							inmature
30791	13	9.5	84							inmature
30792										inmature, the earthworm was broken
30793										inmature, the earthworm was broken
30794										inmature, the earthworm was broken
30613	34	186	85	22-27	24-25-26	12	9-10-11-12	not present	not present	
30614	29.5	131.7	69	22-27	23-24-25	13	9-10-11-12	not present	not present	
30615	12	31.1	87							
30616										the earthworm was broken
30617	23	68.4	83							
30618	20	33.1	96							
30619										the earthworm was broken
30620	22.5	83.1	77							
30621	19	75.3	82							
30622	9	10.3	71							
30623										the earthworm was broken
30624	32	165.9	94	23-30	24-25-26-27	13	9--11--12	not present	not present	
30625	25	94.8	61	18-23	20-21-22	8	not present	not present	not present	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidical funnels	Observations
30626	28	127.1	91	22-28	23-24-25	13	11 -- 12	not present	not present	
30627	21.5	77.4	69	22-27	23-24-25	13	9-10-11-12	not present	not present	
30628	36	110.7	81	22-28	23-24-25	13	9-10-11-12	not present	not present	
30629	27.5	102.7	77	22-27	23-24-25	13	10--11--12	not present	not present	
30630	22	88.9	66	22-27	23-24-25	13	9--11--12	not present	not present	
30631	27	107.9	72	22-27	23-24-25	13	9-10-11-12	not present	not present	
30632	29	110.2	100	22-28	23-24-25-26	13	11 -- 12	not present	not present	
30633	25	104.7	98	22-27	23-24-25	13	9--11--12	not present	not present	
30562	26	85.7	74	22-27	23-24-25	13	11 -- 12	not present	not present	
30563	17	43.3	73		23-24-25	13		not present	not present	the earthworm was broken
30564						13				semi-mature, the earthworm was broken
30565				22-27	23-24-25	13	9-10-11-12			semi-mature, the earthworm was broken
30566								without sperm	without sperm	
30567	24.5	79.8	64	22-27	23-24-25	13	9-10-11-12			
30568	21	71.3	52	22-27	23-24-25	13	9-10-11-12	without sperm	not present	
30569	20	38.6	82							
30570	16	22.4	85							
30571	22	65	75	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30572				22-27	23-24-25-26	13	9-10-11-12	not present	not present	the earthworm was broken
30573	20	27.9	79			13				
30574	22	112.4	62	22-28	24-25-26	13	9-10-11-12	not present	not present	
30575					23-24-25	13				the earthworm was broken
30576	21	67.3	68	22-27	24-25-26	13	11 -- 12	not present	not present	
30577	21.5	65.8	77							

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidical funnels	Observations
30578				22-27	23-24-25	13	9 -- 10	not present	not present	the earthworm was broken
30603										the earthworm was broken
30604	33	187.7	94		22-27	13	9-10-11-12	not present	not present	
30605	39	157.1	90		22-27	13	9-10-11-12	not present	not present	
30606	42	260.3	76		22-27	13	9-10-11-12	without sperm	without sperm	the earthworm was broken
30609					22-27	13	9-10-11-12	without sperm	without sperm	the earthworm was broken
30610						13				
30611	33.5	134.5	74		22-27	13	11 -- 12	not present	not present	
30612	32	121.4	81		22-27	13	10--11--12	not present	not present	
30591	25	118	75	22-27	23-24-25	13	9-10-11-12	not present	not present	
30592	31	122.6	71	22-27	23-24-25	13	10--11-12	not present	not present	
30593						13				the earthworm was broken
30594	25	66.7	87		23-24-25	13				
30595				22-27	23-24-25	13	9-10-11--12	not present	not present	the earthworm was broken
30596	31	120.2	88	22-27	23-24-25	13	11 -- 12	not present	not present	
30597				22-27	23-24-25	13	9-10-11--12	not present	not present	the earthworm was broken
30598	13	13	82							
30580						13				the earthworm was broken
30581				22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	the earthworm was broken
30582	18	88.7	39	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30583	22	82.2	73	22-27	23-24-25-26	13	9-10-11-12	without sperm	without sperm	
30584				22-27	23-24-25	13	11 -- 12	not present	not present	the earthworm was broken
30585	15	21.9	93							

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
30587						13				the earthworm was broken
30588	14.5	20.3	88							
30589	11	11.1	80							
30590	9	5.3 g	77							
30761	22	63.1	93	22-27	24-25-26	13	9-10-11-12	not present	not present	
30762	22	64.8	87	22-26	23-24-25	13	9-10-11-12	not present	not present	
30763										the earthworm was broken
30765	18	49.6	59	22-27	23-24-25	13	11 -- 12	not present	not present	
30766	20	38.2	88			13				
30767	14	25.4	56							
30768	21.5	53.9	94		23-24-25	13				
30769				22-27	23-24-25	13	9-10-11-12	not present	not present	the earthworm was broken
30770	18	25.4	74	22-27	23-24-25	13	9-10-11-12	not present	not present	
30771	20	77.8	90	22-28	23-24-25-26	13	10--11--12	not present	not present	
30772	20	75.4	87		23-24-25	13				
30773				22-26	23-24-25	13	9-10-11-12	not present	not present	the earthworm was broken
30774	20	76.8	88	22-27	23-24-25	13	9-10-11-12	not present	not present	
30775	15	30.4	90							
30776										inmature, the earthworm was broken
30777	19	65.3	84		23-24-25	13				
30778					23-24-25	13				the earthworm was broken
30843										inmature, the earthworm was broken
30844	22	76.2	90		23-24-25	13				semi-mature
30845	19	32.1	83							

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
30846	37	160.2	89	23-28	24-25-26	13	11 -- 12	one on each side without sperm	without sperm	
30847	20	63.7	77		23-24-25	13				
30848					23-24-25	13				the earthworm was broken
30849										inmature, the earthworm was broken
30850	18	30.3	91			13				
30851										
30854	20	59.3	66	22-27	23-24-25	13	9-10-11-12	one on each side without sperm	without sperm	
30856	4	30	57							inmature
30857	5	32	61							inmature
30858	17									
30859	20	150	62	24-27	25-26-27	13	11 -- 12	not present	without sperm	
30860										
30861	21	45.6	65	22-26	22-23-24-25-26	13	9-10-11-12	not present	not present	
30862	24	51.5	83	22-27	23-24-25	13	10--11--12	not present	not present	
30863				22-26	23-24-25	13	11 -- 12	not present	without sperm	the earthworm was broken
30864				22-26	23-24-25	13	11	not present	not present	the earthworm was broken
30865				21-26	22-23-24	10	11 -- 12	not present	not present	the earthworm was broken
30866					21-22-23	10	8--9--10	not present	without sperm	the earthworm was broken
30867				22-26	23-24-25	13	10--11--12	without sperm	without sperm	the earthworm was broken
12	14	12.3	72							inmature
30869	9	9.8	81							inmature
30870	10	11.5								inmature
30871										

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidical funnels	Observations
30872										inmature, the earthworm was broken
30873	6	3.2	65							
30876										inmature, the earthworm was broken
30877	21	43.5	67		23-24-25	13				semi-mature
30878	8	6.7	73							
30879										inmature, the earthworm was broken
30880						13				semi-mature, the earthworm was broken
30883	20	35.2	77							inmature
30884				21-26	22-23-24	13	9-10-11-12	without sperm	without sperm	the earthworm was broken
30885	20	47.4	69	22-27	23-24-25-26	13	10 -- 11	without sperm	without sperm	
30886	10	11.3	62							inmature
30887	18	21.1	65							inmature
30888	20.5	60.6	71	21-26	22-25	13	9-10-11-12	not present	not present	
30889					23-24-25	13				semi-mature
30890	19.5	33.1	86	22-28	24-25-26	13	9-10-11-12	not present	not present	the earthworm was broken
30891										the earthworm was broken
30892	14	31.2	74							inmature
30894						13				semi-mature, the earthworm was broken
30896	29	67..7	87	22-28	23-24-25	12	10--11--12	not present	without sperm	
30897										inmature and the earthworm was broken
30898										inmature and the earthworm was broken
30899										inmature and the earthworm was broken

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidical funnels	Observations
30900	24	55.9	82	22-28	23-24-25	13	9-10-11-12	without sperm	without sperm	
30901				21-26	22-23-24-25	12	9-10-11-12	not present	not present	the earthworm was broken
30902	15	34.2	69							inmature
30903	29	79.2	89			13				semi-mature
30904	25	56.7	84	21-27	23-24-25	13	10--11--12	without sperm	without sperm	
30905	19.5	34.2	91			13				semi-mature
30906	22.5	46.3	97	21-28	22-23-24-25-26-1n27	11	9-10-11-12	without sperm	without sperm	
30907	17	27.1	87							inmature
30908				22-28	23-24-25	13	11 -- 12	one on each side without sperm	without sperm	the earthworm was broken
30909				20-26	22-23-24-25	13	9-10-11-12	not present	not present	the earthworm was broken
30910										the earthworm was broken
30911	23	58.5	91		23-24-25	13				semi-mature
30912										
30913	18	32.3	82							inmature
30914										
30915	29	117.9	85	22-27	23-24-25-1n26	13	11 -- 12	not present	not present	
30916	26.5	101.6	90	22-28	23-24-25-1n26	13	10--11--12	one on each side without sperm	without sperm	
30917	20	36.4	87			13				semi-mature
30918	19	30.5	77							inmature
30919										
30920	29	123.5	88	22-27	23-24-25-1n26	13	11 -- 12	without sperm	not present	
30921	14.5	19.8	83							inmature
30922	18	38.1	79							inmature

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidducal funnels	Observations
30924										
30925										
30926	27	116.9	82	21-26	22-23-24	13	11 -- 12	one on each side without sperm	not present	
30927				22-27	23-24-25-26	13	11 -- 12	not present	not present	the earthworm was broken
30928										
30929	22	73.5	85	22-27	23-24-25	13	9-10-11-12	one on each side without sperm	without sperm	
30930										
30931										
30932	18	41.3	72	22-28	23-24-25-26	13	11 -- 12	not present	not present	
30933	16	39.5	57							inmature
30934	15	33.7				13				semi-mature
30935										
30936				21-26	23-24-25	12	9-10-11-12	without sperm	not present	the earthworm was broken
30937				22-28	23-24-25-26	13	9-10-11-12	not present	not present	the earthworm was broken
30938	15	43.1	53	22-27	23-24-25	13	10-- 11	not present	not present	
30939					23-24-25	13				semi-mature, the earthworm was broken
30940										
30941	18	39.8	82	21-27	23-24-25	13	11 -- 12	not present	not present	
30946	19	43.8	77							
30947										inmature, the eathworm was broken
30948										inmature, the eathworm was broken
30951						13				semi-mature, the earthworm was broken
30953	24	60	86	22-27	23-24-25	13	9-10-11-12	one on right side without sperm	without sperm	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
30960										inmature, the eathworm was broken
30991	22.5	63.1	77		23-24-25	13				
30993	21.5	62.7	82		23-24-25	13				
30995										inmature, the earthworm was broken
31002										
31008	20	64.3	77		23-24-25	13				
30818	21.5	55.3	93	22-27	23-24-25	13	11--12	not present	not present	
30811	19.5	43.6	85		23-24-25	13				semi-mature
30823	17	33.9	74		23-24-25	13				semi-mature
30814	29	97	87	22-27	23-24-25	13	11---12	not present	not present	
30820	14	16.5	77							inmature
30815	20	65.6	61	22-27	23-24-25	13	9--11--12	not present	without sperm	
30817	23	72.3	85		23-24-25	13				semi-mature
30816	24.5	84.1	101	22-28	24-25-26	13	11 -- 12	not present	without sperm	
30812	35	138.7	87	22-27	23-24-25	13	11 -- 12	not present	not present	
30813				22-27	23-24-25	13	11 -- 12	not present	not present	the earthworm was broken
30838	23.5	63.7	91	22-28	24-25-26	13	9-10-11-12	without sperm	without sperm	
30837	17	55.6	58		23-24-25	13				semi-mature
30841										inmature, the earthworm was broken
30842						13				semi-mature, the earthworm was broken
30836										inmature, the earthworm was broken
30839	11	33	78			13				semi-mature
30840										the earthworm was broken
30796					23-24-25	13				semi-mature, broken

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
30797	18	28.8	82							inmature
30795	24	68.4	93	22-27	23-24-25	13	11--12	without sperm	without sperm	
30800	13	17.8	77							inmature
30798				22-27	23-24-25	13	9--11--12	without sperm	without sperm	the earthworm was broken
30802					23-24-25	13				semi-mature, the earthworm was broken
30804	19	25	81							inmature
30801	16	21.3	78							inmature
30803						13				semi-mature, the earthworm was broken
30808	15	19.9	66							inmature
30670	24	111.5	81	22-27	23-24-25	13	9--11--12	not present	not present	spermatophore between 19 and 20
30692	25	113.8	93	22-27	23-24-25	13	10--11--12	without sperm	without sperm	
30687	18	100.7	88	22-27	23-24-25	13	9-10-11-12	not present	not present	
30696	30	114	89	22-27	23-24-25	13	9--11--12	without sperm	without sperm	
30694	12	76.8	57	22-27	23-24-25	13	9-10-11-12	not present	not present	
30688	21	74.4	84	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30686	20	86.4	64	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30693	26	109.7	87	22-28	24-25-26	13	9-10-11-12	without sperm	without sperm	
30685				22-27	23-24-25	13	9-10-11-12	not present	not present	the earthworm was broken
30684	18	61.2	65	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30736	32	160.4		22-27	23-24-25	13				the earthworm was broken
30737	28	157.4	94	22-27	23-24-25	13	11 -- 12	without sperm	not present	
30730	31	162.3	87	22-27	23-24-25	13	11 -- 12	without sperm	without sperm	
30738				22-26	23-24-25	13	10--11--12	without sperm	without sperm	the earthworm was broken
30739	23	75.1	87	22-27	23-24-25-26	13	9-10-11-12	without sperm	without sperm	

| UCM-LT number | Length (mm) | Dry weight (mg) | Number of segments | Clitellum | Tubercula pubertatis | Male pore | Seminal vesicles | Spermatechae | Spermid
ucal funnels | Observations |
|---|---|---|---|---|---|---|---|---|---|---|
| **30741** | 22.5 | 68.8 | 77 | | 23-24-25 | 13 | | | | semi-mature |
| **30735** | 19.5 | 44.7 | 82 | | | | | | | inmature |
| **30734** | 23 | 82.1 g | 81 | 22-27 | 23-24-25 | 13 | 9-10-11-12 | not present | not present | |
| **30733** | 19 | 60.1 | 89 | | 23-24-25 | 13 | | | | semi-mature |
| **30732** | 13 | 15.6 | 65 | | | | | | | inmature |
| **30716** | 24 | 102.3 | 88 | 22-26 | 23-24-25 | 13 | 9-10--11-12 | without sperm | without sperm | |
| **30713** | | | | 22-27 | 23-24-25 | 13 | 9-10-11-12 | not present | not present | the earthworm was broken |
| **30714** | 22.5 | 111.1 | 79 | 22-26 | 23-24-25 | 13 | 9-10-11-12 | not present | not present | |
| **30715** | 22 | 109 | 84 | 22-27 | 23/24/25 | 13 | | | | |
| **30717** | | | | | | | | | | inmature, the earthworm was broken |
| **31023** | 16 | 30 | 74 | | | | | | | inmature |
| **31024** | 14.5 | 11.2 | 77 | | | | | | | inmature |
| **31025** | 13 | 20 | | | | | | | | |
| **31026** | | | | | | | | | | inmature, the earthworm was broken |
| **31027** | | | | | | | | | | inmature, the earthworm was broken |
| **31028** | 8 | 11.2 | 71 | | | | | | | inmature |
| **31029** | | | | | | | | | | inmature, the earthworm was broken |
| **31030** | 13 | 20.1 | 60 | | | | | | | inmature |
| **31031** | | | | 22-28 | 23-27 | 13 | 9-10-11-12 | not present | not present | the earthworm was broken |
| **31032** | | | | 21-27 | 23-24-25 | 13 | 9-10-11-12 | not present | not present | the earthworm was broken |
| **31033** | | | | | | | | | | the earthworm was broken |
| **31034** | 20.5 | 57.6 | 63 | 22-28 | 23-24-25-1n26 | 13 | 9-10-11-12 | not present | not present | |
| **31038** | 30 | 118.8 | | | | 13 | 9-10-11-12 | not present | not present | the earthworm was broken |

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
31039	14	21.1	60							inmature
31040	13	18.1	72							inmature
31041	27	110.5	79							inmature
31043	15	20.8	64							inmature
31044	12	10.5	70							inmature
31045	25	111.5	69	22-27	23-24-25-26	13	10--11--12	without sperm	without sperm	
31046	17.5	50.8	66	22-27	23-24-25	13	10--11--12	without sperm	without sperm	
31047	19.5	30.3								inmature
31048										
31049	10	108	73							inmature
31050	9	10.9	71							inmature
31051	24	90.8	59	22-27	23-24-25-1n26	13	9-10-11-12	not present	without sperm	
31052	16	10.5	71							inmature
31053										
31054	17.5	54.8	82							inmature
31055										inmature, the earthworm was broken
31056	9	10.1	67							inmature
31057										inmature, the earthworm was broken
31058	6	7.4	65							inmature
31059	4	6.1	69							inmature
31060	12	20.4	65							inmature
31061	9	10.2	63							inmature
31062	9.5	11.4	77							inmature
31063	11	12.5	75							inmature
31064	10	16.5	67							inmature

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
31065	0.9	5.2	87							inmature
31066	0.8	11.9								inmature
31067				22-28	24-25-26-27	13	10--11--12	not present	not present	the earthworm was broken
31068				23-28	24-25-26	13	11 -- 12	not present	not present	the earthworm was broken
31069	0.8	11.3	87							inmature
31070	11	12.5	88							inmature
31071	23	50.9	83							inmature
31072										the earthworm was broken and nmature
31073										the earthworm was broken and inmature
31074	16	25.6	78							inmature
31075										the earthworm was broken and inmature
31076										
31077										the earthworm was broken and inmature
31078	0.7	3.5	82							inmature
31082	26	90.6	86	22-28	23-24-25	13	8-9-10-11	without sperm	not present	
31083	19	45.5	85							inmature
31084	21	50.5	77		23-24-25					semi-mature
31085	8	10.3	65							inmature
31086	17	30.5	71							inmature
31087	15	33.6	74							inmature
31088				23-28	24-25-26	13	9-10-11-12	one on each side without sperm	without sperm	the earthworm was broken
31089	20	43.2	82							inmature
31090										the earthworm was broken and inmature

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
31091				22-27	24-25-26	13	9-10-11-12	not present	not present	the earthworm was broken
31092	21	65.2	91	23-27	24-25-26	13	9-10-11-12	without sperm	without sperm	
31093	22	74.2	73	23-28	24-25-26	15	9-10-11-12	without sperm	not present	
31094	40	170.2	102	22-26	23-24-25	13	10--11--12	not present	not present	
31095	34	141.7	89	22-28	1n22-1n28	13	11 -- 12	not present	without sperm	
31096	30	111.4	93	22-28	24-25-26	13	10--11--12	not present	not present	
31097	26	95.6	80	23-27	24-25-26	13	9-10-11-12	not present	not present	
31098	21.5	78.4	51	22-26	23-24	13	10--11--12	not present	not present	
31099	33	120.3	88	22-27	23-24-25-26	13	11 -- 12	not present	not present	
31100	10	22.2	77							inmature
31101	0.8	24.5	74							inmature
31102	17	66.9	78		24-25-26	13				semi-mature
31103	29	71.8				13	9-10-11-12	not present	not present	
31104	27	74.5	80	22-27	23-24-25	13	11 -- 12	without sperm	not present	
31105				23-28	24-25-26	13	9-10-11-12	not present	not present	the earthworm was broken
31106	33	125.7	74	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
31107	19.5	66.5	60	23-28	24-25-26	13	11 -- 12	not present	not present	
31108	20	58.4	87							inmature
31109	14.5	10.9	83							inmature
31110	21	51								inmature
31111	32	141.2	93	23-28	24-25-26	13	10--11--12	not present	not present	
31112				23-28	24-25-26	13	10--11--12	not present	not present	the earthworm was broken
31113	23	55.3	94	22-27	13-24-25	13	11 -- 12	without sperm	without sperm	the earthworm was broken and inmature
31114										inmature
31115	10	11.6	79							inmature

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidlucal funnels	Observations
31116	22	80.5	86	22-27	23-24-25	13	9-10-11-12	not present	not present	
31117	29.5	171.2	90	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
31118	24.5	105.3	82	22-27	23-24-25	13	11 -- 12	not present	not present	
31119										
31120	22	57.6	79		23-24-25	13				semi-mature
31121	30	120.6	90	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
31122	23.5	50.3	83	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
31123	13	20.1	78							inmature
31124										inmature, the earthworm was broken
31125										the earthworm was broken
31126										inmature, the earthworm was broken
31127	29	71.2	78	22-27	23-24-25	13	9-10-11-12	not present	not present	
31128	26	45.6	78	17-22	19-20-21	8 y 9	3-4-5-6	without sperm	not present	maybe regenerated by the head
31129	24	88.9	82							inmature
31130	30	135.2	92		23-24-25	13				semi-mature
31131	16	33.8	77							inmature
31132										
31133				22-27	23-24-25-1n26	13	9-10-11-12	without sperm	without sperm	the earthworm was broken
31134	30	91.7	88	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
31135	24.5	63.8	77		23-24-25	13				semi-mature
31136	24	57.6	84							inmature
31137	15	35.2	91							inmature
31139	19	44.2	88							inmature
31140	28	109.6	93	22-27	23-24-25	13	9-10-11-12	not present	not present	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
31141	15	15.9	93							inmature
31142	19.5	36.5	90							inmature
31143										the earthworm was broken and inmature
31144	23	57.2	73			13				semi-mature
31145					24-25-26	13				the earthworm was broken and semi-mature
31146	30	115.2	75	23-28	24-25-26	13	9-10-11-12	without sperm	without sperm	
31147	24	63.8	70	23-28	24-25-26	13	10--11--12	not present	without sperm	
31148	20	48.3	81							inmature
31149	20	42.3	80		24-25-26	13				semi-mature
31169	27.5	57.6	82		13					semi-mature
31170	27	66.8	87	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
31171										inmature, the earthworm was broken
31172										inmature, the earthworm was broken
31173	18	25.2	82							inmature
31174				21-25	22-23-24	12	9-10-11-12	one on right side without sperm	not present	spermatophore between 19 and 20, the earthworm was broken
31175	22	47.9	87	22-27	23-24-25	15	10--11--12	not present	without sperm	
31176	21.5	39.5	77		23-24-25	13				semi-mature
31177					23-24-25	13				semi-mature, the earthworm was broken
31178	19	44.7	79		23-24-25	13				semi-mature
31179	20	42.9	79		23-24-25	13				semi-mature
31180				22-27	23-24-25	13	11 -- 12	without sperm	without sperm	the earthworm was broken
31181	25	47.6	84	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
31182				22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	the earthworm was broken

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidiucal funnels	Observations
31183	25	78.9	85	22-27	23-24-25	13	9-10-11-12	without sperm	no presente	
31184	12	45.8	79							inmature
31185	19	57.2	61			13				semi-mature
31186	24	54.6	81			13				semi-mature
31187	20	48.5	73		23-24-25	13				semi-mature
31188	20	75.6	94	21-27	23-24-25-1n26	13	11 -- 12	without sperm	without sperm	
31189	17.5	48.2	87	22-27	23-24-25	13	11 -- 12	without sperm	without sperm	
31190	17	55.2	64	22-29	23-24-25-26	13	10--11--12	without sperm	without sperm	
31191	19	58.3	101	22-28	24-25-26	13	11 -- 12	not present	without sperm	
31192	19	64.1	88	22-27	23-24-25	13	11 -- 12	without sperm	without sperm	
31193	19.5	61.9	71	22-27	23-24-25	13	10--11--12	without sperm	without sperm	
31194	11	22.1	52							inmature
31195	11	24.3	78							inmature
31196	19.5	59.1	83							inmature
31197	9	23.1	75							inmature
31198	18	35.6	79	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
31199	22	78	90	23-28	24-25-26	13	9-10-11-12	no presente	no presente	
31200	15	35.2	96	22-28	23-24-25	13	11 -- 12	without sperm	without sperm	
31201	8	16.4	36	22-27	23-24-25	13	not present	without sperm	not present	
31202	16	39.1	97	21-26	22-23-24	13	11 -- 12	not present	not present	
31203	11	33.5	83							inmature
31204	14.5	36.4	68							inmature
31205	13	37.5	57		23-24-25	13				semi-mature
31206	28	87.3	80		23-24-25	13				semi-mature, the earthworm was broken
31207										inmature

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidugal funnels	Observations
31208	16	39.5	71	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
31209	26.5	78.6	86	22-27	23-24-25	13	9-10-11-12	one on each side without sperm	without sperm	
31210	22.5	79.5	74	22-27	23-24-25	13	9-10-11-12	not present	not present	
31211	21.5	77.4	57	22-27	23-2425-1n26	13	11 -- 12	not present	not present	
31212	24	85.3	62	22-27		13	10--11--12	not present	not present	
31213	20	57.2	83	20-25	22-23-24	11	11 -- 12	not present	not present	
31214	23	65.2	83	22-27	23-24-25	13	9-10-1-12	without sperm	not present	
31215	18	59.2	85	22-27	23-24-25	13	9-10-11-12	not present	not present	
31216	25	84.1	92	22-27	23-24-25	13	11 -- 12	not present	not present	
31217	28.5	95.2	90	22-27	23-24-25	14	9-10-11-12	one on each side without sperm	not present	
31218										inmature, the earthworm was broken
31219	28	75.2	84		22-23-24	13				semi-mature
31220	23	49.3	91							inmature
31221	28	88.9	88			13				semi-mature
31222	20	36.7	77							inmature
31223	25	59.8	91							inmature
31224	21	36.8	82							inmature
31225	31	95.9	86		23-24-25	13				semi-mature
31226	23	96.7	91	22-28	23-24-25	13	9-10-11-12	without sperm	without sperm	
31227	21.5	75.8	87	22-28	23-24-25	14	10--11--12	not present	without sperm	
31228	25.5	93.2	87	22-27	23-24-25-26	13	9-10-11-12	not present	not present	
31229	42	193.6	84	22-27	23-24-25-26	13	9-10-11-12	not present	not present	
31230										the earthworm was broken
31231	26	112.8	81	22-28	23-24-25-26-1n27	13	9-10-11-12	without sperm	without sperm	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
31232	21	75.1	83	22-28	23-24-25	13	9-10-11-12	not present	not present	
31233	19	43.4	79			15				semi-mature
31234	21.5	87.9	65	22-28	23-24-25	13	9-10-11-12	not present	not present	
31235	22	57.3	89			15				semi-mature
31236	18	47.2	77			13				semi-mature
31237	27	118.5	86	22-28	23-24-25-1n26	13	11 -- 12	not present	not present	
31238	24.5	96.2	87	22-28	23-24-25	13	9-10-11-12	not present	not present	
31239	24.5	83.4	89	22-28	23-24-25	13	10 -- 12	not present	not present	
31240	25	87.2	87	22-28	23-24-25-26	13	11 -- 12	no presentes	not present	
31241	30	128.5	92	22-28	23-24-25-26	15	9-10-11-12	not present	not present	
31242	24.5	87.5	70	22-27	23-24-25	13	11 -- 12	not present	not present	
31243	20.5	74.9	82		23-24-25	15				semi-mature
31244	16	47.6	51	22-27	23-24-25	13	not present	not present	not present	
31245	15	19.8	80	22-27	24-25-1n26	13	9-10-11-12	not present	without sperm	
31246	19	54.9	80	23-28	24-25-26	13	11 -- 12	without sperm	without sperm	
31247	20	36.7	68							inmature
31248	21.5	78.5	78		23-24-25	13				semi-mature
31249	21	47.2	83			13				semi-mature
31250	24	75.2	77							inmature
31251	17	39.4	71							inmature
31253	14.5	47.6	77							inmature
31254	19	33.4	78							inmature
31255	17	37.2	63							inmature
31256	27	116.4	73	22-27	23-25	13	9-10-11-12	not present	not present	
31257	30	135.1	81	22-27	23-25	13	left side: 11 right side: 11--12	not present	not present	spermatophore between 15 and 16

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidurcal funnels	Observations
31258	28	97.4	85	22-27	22-25	13	10--11--12	not present	not present	
31259	28	114.2	82	23-29	24-27	13	11 -- 12	not present	not present	
31260	15	35.8	58		23-25	13				semi-mature
31262	17.5	36.7	89	22-27	23-25	13	9-10-11-12	not present	not present	
31263				22-27	23-25	13	10--11--12	not present	not present	the earthworm was broken
31264	13	37.5	74							inmature
31265	18	59.4	79		23-25	13				semi-mature
31266	18	36.5	77	22-27	23-25	13	11 -- 12	not present	not present	
31267	21	78.5	92	22-27	23-25	13	11 -- 12	not present	not present	
31268	18	39.6	82		23-25	13				semi-mature
31269	15	24.7	89							inmature
31270	18	47.2	63	22-27	23-24-25-1n26	13	10--11--12	not present	not present	
31271	22	89.6	76	22-27	23-24-25-1n26	13	9-10-11-12	without sperm	not present	
31272	21.5	64.7	83	22-27	23-24-25	13	not present	not present	not present	
31273					23-24-25	13				semi-mature, the earthworm was broken
31274	17	58.2	79		23-24-25	13				semi-mature
31275	17	74.3	81	22-27	23-24-25	13	not present	not present	not present	
31276	15	47.2	71							inmature
31277	15	26.1	66							inmature
31278	28	35.4	79			13				semi-mature
31279	25	89.4	88	22-27	23-24-25-1n26	15	11 -- 12	not present	not present	
31280				22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	the earthworm was broken
31281	25	94.3			23-24-25	13				semi-mature
31282	30	117.6	92	22-27	23-24-25	13	11 -- 12	without sperm	without sperm	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
31283	26	86.8	84	22-27	23-24-25	15	11 -- 12	not present	not present	
31284	18	54.9	51	22-27	24-25-26	13	11 -- 12	not present	not present	
31285	27	78.6	89	21-28	23-24-25-26-27	12	9-10-11-12	not present	not present	
31286	20	47.8	76		23-24-25	13				semi-mature
31287										inmature, the earthworm was broken
31288	19	48.2	73							inmature
31289	22	57.5	92			13				semi-mature
31290	17.5	35.9	69							inmature
31291	20	47.2	84		23-24-25	13				semi-mature
31292				22-27	23-24-25	13	11 -- 12	without sperm	without sperm	the earthworm was broken
31293	24	75.21	89	22-27	23-24-25	13	11 -- 12	not present	not present	
31294	17	45.6	61	22-28	24-25-26	13	11 -- 12	not present	not present	
31295	16	33.2	77	21-17	23-24-25	13	11 -- 12	not present	not present	
31296	23	69.4	82	22-27	23-24-25	13	11 -- 12	not present	not present	
31297	22	47.8	85	22-27	23-24-25	13	11 -- 12	not present	not present	
31298	14	45.1	71							inmature
31299	27.5	98.4	84	22-27	23-24-25	13	9-10-11-12	not present	without sperm	
31300	24.5	101.6	82	22-28	24-25-26	13	9-10-11-12	without sperm	without sperm	
31301	22	125.8	75	22-27	23-24-25	13	11 -- 12	not present	not present	
31302	22	74.3	79	22-27	23-24-25	13	10--11--12	not present	without sperm	
31303	17	42.1	77			13				semi-mature
31304	17.5	52.9	74	22-27	23-24-25	13	11 -- 12	not present	without sperm	the earthworm was broken
31305						13				semi-mature and the earthworm was broken
31306				22-27	23-24-25-1n26	13	10--11--12	without sperm	without sperm	

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermidducal funnels	Observations
31307	15	12.3	81							inmature
31308	14	17.4	79							inmature
31309	13	21.9	78							inmature
31310	16	54.6	85		13					semi-mature
31311	15	32.9	81							inmature
31312	16.5	25.8	81							inmature
31313	19.5	74.2	57	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
31314				22-27	23-24-25	13	10--11--12	not present	not present	the earthworm was broken
31315	23.5	71.3	77	22-27	23-24-25	13	10--11--12	not present	not present	
31316	25	79.5	92	22-28	23-24-25-26	13	11 -- 12	not present	not present	
30780	20	12.1	87	22-27	23-24-25	13	10--11-12	not present	not present	
30781	20	13.8	93	22-28	23-24-25-1n26	13	8-9-10-11	not present	not present	
30782	25	20.5	83	22-27	23-24-25	13	9-10-11-12	without sperm	without sperm	
30783	22.5	29.5	85	22-27	23-24-25	13	8-9-10-11	without sperm	without sperm	
30784	21	13.4	92	22-27	23-24-25	13	10--11-12	without sperm	not present	
30785	23.5	15.8	82	22-28	23-24-25-26	13	10--11-12	without sperm	without sperm	
30786	0.7	3.1	64							inmature
30673	17	19.2	74			13				semi-mature
30674	20	21	83		23-24-25	13				semi-mature
30675	12	10.2	90			13				semi-mature
30676	15	6.1	82							inmature
30677						13				semi-mature and the earthworm was broken
30678	11	5.8	74							inmature
30679										the earthworm was broken and inmature
30680	14	17.5	83			13				semi-mature, broken

UCM-LT number	Length (mm)	Dry weight (mg)	Number of segments	Clitellum	Tubercula pubertatis	Male pore	Seminal vesicles	Spermatechae	Spermiducal funnels	Observations
30682										inmature, the earthworm was broken
30683	14	8.7	95							inmature
31010	20	25.7	74	23-27	23-24-25	13	9-10-11-12	not present	not present	
31011	19	20.8	84	22-27	23-24-25	13	9 -- 10	not present	not present	
31012	22	19.8	85	22-27	22-23-24	13	9-10--11	without sperm	not present	
31013										the earthworm was broken and inmature
31014										inmature and the earthworm was broken
31015	9	3.7 mm	83							inmature
31016	9	3.5	80							inmature
31017										
31018	10	4.1	74							inmature
31019										the earthworm was broken
30704	22.5	131.8	94	22-27	23-24-25	13	11 -- 12	not present	not present	
30705	23	132.7	92	22-27	23-24-25	13	10--11-12	not present	not present	
30706				22-27	23-24-25	13	10--11--12	not present	not present	the earthworm was broken
30707	29	151.3	85	22-27	23-24-25	13	11 -- 12	not present	not present	
30708	24	120.7	77	22-27	23-24-25	13	11 -- 12	not present	not present	
30709				22-27	23-24-25	13	12	not present	not present	the earthworm was broken
30710	15	37.6	82							inmature
30711	26	100.4	91		23-24-25	13				semi-mature
30712	27	133.4	92	22-27	23-24-25	13	10--11--12	not present	not present	

Supplementary Table 7. Morphological traits of specimens included in the study.

Chapter 4

Locality	Longitude	Latitude	Number in map	UCM-LT references	Number of specimens used
Spain	-3,5726	36,7337	1	31320-31333	13
Cyprus	32,8617	34,8861	2	30486	1
Italy 1	14,3306	40,8848	3	30852-30855	3
Italy 2	10,9992	45,4340	4	30874-30882	7
Italy 3	13,2292	43,5134	5	30895	1
Italy 4	11,5512	45,5472	6	30945-30960	10
Italy 5	11,5606	43,5499	7	30961-30982	17
Italy 6	12,4508	43,0905	8	30989-31007	14
Italy 7	16,2934	39,7068	9	31035-31037	3

Supplementary Table 1. Localities sampled by Irene de Sosa and Nuria Sánchez (black), Csaba Csuzdi and Timea Szederjesi (blue), Aleksandra Jablonska and Misel Jelic (orange).

Name	**Reference or accession numbers (COI/16S/28S/12S/ND1)**
Allolobophora (Gatesona) chaetophora	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Allolobophora (Gatesona) sp1	Marchán *et al*. 2021
Allolobophora (Gatesona) sp2	Marchán *et al*. 2021
Allolobophora (Gatesona) sp3	Marchán *et al*. 2021
Allolobophora (Gatesona) sp4	Marchán *et al*. 2021
Allolobophora (Gatesona) sp5	Marchán *et al*. 2021
Allolobophora bartoli	Marchán *et al*. 2021

Name	Reference or accession numbers (COI/16S/28S/12S/ND1)
Allolobophora chlorotica	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Allolobophora dacica	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Allolobophora dubiosa	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Allolobophora mehadiensis	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Allolobophora moebii	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Allolobophora mollei	Pérez-Losada *et al*. 2009, 2015; Domínguez *et al*. 2015
Allolobophora robusta	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Aporrectodea caliginosa	Pérez-Losada *et al*. 2009, 2015; Domínguez *et al*. 2015
Aporrectodea georgii	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Aporrectodea jassyiensis	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Aporrectodea limicola	Pérez-Losada *et al*. 2009, 2015; Domínguez *et al*. 2015
Aporrectodea longa	Pérez-Losada *et al*. 2009, 2015; Domínguez *et al*. 2015
Aporrectodea nocturna	Pérez-Losada *et al*. 2009, 2015; Domínguez *et al*. 2015
Aporrectodea rosea	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Aporrectodea trapezoides	Pérez-Losada *et al*. 2009, 2015; Domínguez *et al*. 2015
Aporrectodea tuberculata	Pérez-Losada *et al*. 2009, 2015; Domínguez *et al*. 2015
Avelona ligra	Marchán *et al*. 2021
Bimastos rubidus	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Carpetania elisae	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Castellodrilus alavanensis	Jiménez *et al*. 2021
Castellodrilus chitae	Jiménez *et al*. 2021
Castellodrilus eurythrichos	Jiménez *et al*. 2021
Castellodrilus hongae	Jiménez *et al*. 2021
Castellodrilus ibericus	Jiménez *et al*. 2021
Castellodrilus opisthoporus	Jiménez *et al*. 2021
Castellodrilus pulvinus	Jiménez *et al*. 2021
Cataladrilus edwarsi	Jiménez *et al*. 2021
Cataladrilus monticola	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Cataladrilus porquerollensis	Marchán *et al*. 2020
Cataladrilus zhongi	Jiménez *et al*. 2021

Name	Reference or accession numbers (COI/16S/28S/12S/ND1)
Cernosvitovia dudichi	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Cernosvitovia rebeli	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Compostelandrilus bercianus	Domínguez *et al*. 2018
Compostelandrilus cyaneus	Domínguez *et al*. 2018
Compostelandrilus menciae	Domínguez *et al*. 2018
Criodrilus lacuum	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Dendrobaena attemsi	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Dendrobaena byblica	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Dendrobaena illyrica	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Dendrobaena jastrebensis	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Dendrobaena octaedra	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Dendrobaena pentheri	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Dendrobaena pygmaea	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Dendrobaena veneta	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Diporodrilus pilosus	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Diporodrilus sp.	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Eisenia andrei	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Eisenia balatonica	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Eisenia fetida	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Eisenia lucens	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Eiseniella tetraedra	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015; de Sosa *et al*., 2017
Eisenoides carolinensis	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Eisenoides lonnbergi	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Eophila gestroi	De Sosa *et al*. 2019
Eophila tellini	De Sosa *et al*. 2019
Ethnodrilus zajonci	Jiménez *et al*. 2021
Galiciandrilus bertae	Pérez-Losada *et al*. 2011, 2015; Domínguez *et al*. 2015
Galiciandrilus morenoe	Pérez-Losada *et al*. 2011, 2015; Domínguez *et al*. 2015
Helodrilus (Acystodrilus) cortezi	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Helodrilus (Acystodrilus) sp.	Marchán *et al*. 2021

Name	Reference or accession numbers (COI/16S/28S/12S/ND1)
Helodrilus cernosvitovianus	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Helodrilus patriarchalis	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Lumbricus castaneus	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Lumbricus rubellus	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Iberoscolex albolineatus	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Iberoscolex oliveirae	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Norealidys andaluciana	MZ578538-MZ578550/MZ661527-MZ661539/MZ661468-MZ661480/OK030852, OK030853/OK030857,OK030858
Norealidys neapolitana	MZ578502-MZ578537/MZ661482-MZ661526/MZ661415-MZ661467/OK030850,OK030851/OK030855,OK030856
Norealidys sp.	MZ578501/MZ661481/MZ661414/OK030849/OK030854
Octodriloides boninoi	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Octodrilus complanatus	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Octodrilus exacystis	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Octodrilus gradinescui	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Octodrilus pseudotranspadanus	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Octodrilus transpadanus	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Octolasion cyaneum	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Octolasion lacteum	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Octolasion montanum	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Panoniona leoni	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Philomontanus baloutchi	Bozorgui *et al*. 2019
Philomontanus sarii	Bozorgui *et al*. 2019
Pietromodeona januaeargenti	De Sosa *et al*. 2019
Postandrilus lavellei	Pérez-Losada *et al*. 2011, 2015; Domínguez *et al*. 2015
Postandrilus majorcanus	Pérez-Losada *et al*. 2011, 2015; Domínguez *et al*. 2015
Postandrilus medoakus	Pérez-Losada *et al*. 2011, 2015; Domínguez *et al*. 2015
Postandrilus palmensis	Pérez-Losada *et al*. 2011, 2015; Domínguez *et al*. 2015
Postandrilus sapkarevi	Pérez-Losada *et al*. 2011, 2015; Domínguez *et al*. 2015
Proctodrilus antipai	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Prosellodrilus biauriculatus	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Prosellodrilus biserialis	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Prosellodrilus pyrenaicus	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015

Satchelius gatesi	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Satchellius madeirensis	Provided by the authors
Scherotheca corsicana	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015
Scherotheca gigas	Domínguez *et al*. 2015; Pérez-Losada *et al*. 2015

Supplementary Table 2. References and accession numbers (GenBank) of all sequences used.

Chapter 5

Failed procedures for the laboratory culture of *Eiseniella tetraedra*
Dry soil + leaf litter infusion + distilled water
Dry soil + manure infusion + distilled water
Dry soil + leaf litter infusion + manure infusion + distilled water
Dry soil + leaf litter infusion (diluted ½) + distilled water
Dry soil + manure infusion (diluted ½) + distilled water
Dry soil + leaf litter infusion (diluted ½) + manure infusion (diluted ½) + distilled water
Dry soil + manure + compost + distilled water

Supplementary Table 1. Previous failed procedures for the laboratory culture of *E. tetraedra*. All have 60% moisture.

	E1	E2	E3	E4	E5	E6	E7	E8
10'	Reactive	Reactive	Reactive	Reactive	Reactive	Reactive	Reactive	Reactive
20'	Reactive	Reactive	Reactive	Reactive	Less reactive	Reactive	Less reactive	Less reactive
30'	Reactive	Reactive	Reactive	Reactive	Less reactive	Less reactive	Less reactive	Less reactive
40'	Less reactive	Reactive	Reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive
50'	Less reactive	Less reactive	Reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive
60'	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive
70'	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive
80'	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive
90'	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive
110'	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive
120'	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	No reaction. Back to 8°C.	Less reactive
140'	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	Less reactive	-	No reaction. Back to 8°C.
160'	Less reactive	No reaction. Back to 8°C.	Less reactive	Less reactive	Less reactive	Less reactive	-	-
180'	Less reactive	-	Less reactive	Less reactive	Less reactive	Less reactive	-	-
200'	Less reactive	-	Less reactive	Less reactive	Less reactive	Less reactive	-	-
220'	Less reactive	-	Less reactive	Less reactive	Less reactive	Less reactive	-	-
240'	Less reactive	-	No reaction. Back to 8°C	No reaction. Back to 8°C	Less reactive	Less reactive	-	-
260'	Less reactive	-	-	-	Less reactive	No reaction. Back to 8°C	-	-
280'	Frozen	-	-	-	Frozen	-	-	-
Next day	Dead	Alive	Alive	Alive	Dead	Alive	Alive	Alive

Supplementary Table 2. Results of preliminary cold experiment.

			RSEM Bowtie 2			
Treatment	**Sample Name**	**N reads (trimmed files)**	**aligned 0 times**	**alignedexactly1time**	**aligned>1times**	**overall alignment rate**
Experiment Cold	Ec1	29453955	11368458(36.60%)	17343559(58.88%)	741938(2.52%)	61,40%
	Ec3	16870983	6129964 (36.33%)	10327731 (61.22%)	413288 (2.45%)	66,66%
	Ec5	13048276	5831579 (44.69%)	6938896 (53.18%)	277801 (2.13%)	55,30%
Control Cold	Cc1	14605551	8054313(55.15%)	6245625(42.76%)	305613 (2.09%)	44,85%
	Cc2	31278380	11238661 (35.93%)	19296078 (61.69%)	743641 (2.38%)	64,07%
	Cc4	22167351	11853333 (53.47%)	9817925(44.29%)	496093(2.24%)	46,53%
Experiment Desiccation	Ed2	2397082	1015818 (42.38%)	1335682 (55.72%)	45582 (1.90%)	57,62%
	Ed5	11942468	4042759 (33.85%)	7551669 (63.23%)	348040 (2.91%)	66,15%
	Ed6	13464616	5458477 (40.54%)	7725400 (57.37%)	280739 (2.08%)	59,46%
Control desiccation	Cd2	31252535	9481695 (30.34%)	20917948 (66.93%)	852892 (2.73%)	69,66%
	Cd5	15699374	6303852 (40%)	8971868 (57%)	423654 (3%)	59,85%
	Cd6	40544664	16769996 (41.36%)	22772977 (56.17%)	1001691 (2.47%)	58,64%

Supplementary Table 3. Sequencing output.

source	term_name	term_id	adjusted_p_value	intersections
GO:MF	proton transmembrane transporter activity	GO:0015078	0.00002054141913122647	COX1,ATP6,COX2
GO:MF	cytochrome-c oxidase activity	GO:0004129	0.0000722472261620331	COX1,COX2
GO:MF	oxidoreductase activity, acting on a heme group of donors	GO:0016675	0.0000722472261620331	COX1,COX2
GO:MF	inorganic cation transmembrane transporter activity	GO:0022890	0.0006830234051501817	COX1,ATP6,COX2
GO:MF	inorganic molecular entity transmembrane transporter activity	GO:0015318	0.0006830234051501817	COX1,ATP6,COX2
GO:MF	cation transmembrane transporter activity	GO:0008324	0.0006830234051501817	COX1,ATP6,COX2
GO:MF	electron transfer activity	GO:0009055	0.0006830234051501817	COX1,COX2
GO:MF	ion transmembrane transporter activity	GO:0015075	0.0008711746489771326	COX1,ATP6,COX2
GO:MF	transmembrane transporter activity	GO:0022857	0.0013563656730277694	COX1,ATP6,COX2
GO:MF	transporter activity	GO:0005215	0.0015257141494689937	COX1,ATP6,COX2
GO:MF	proton-transporting ATP synthase activity,rotational mechanism	GO:0046933	0.0019921172136063803	ATP6
GO:MF	proton channel activity	GO:0015252	0.007299266020440657	ATP6
GO:MF	oxidoreductase activity	GO:0016491	0.012716818572527993	COX1,COX2
GO:MF	copper ion binding	GO:0005507	0.015619169570975835	COX2
GO:BP	oxidative phosphorylation	GO:0006119	0.00004734715577477944	COX1,ATP6,COX2
GO:BP	proton transmembrane transport	GO:1902600	0.00006836143065161999	COX1,ATP6,COX2
GO:BP	aerobic respiration	GO:0009060	0.00006836143065161999	COX1,ATP6,COX2
GO:BP	cellular respiration	GO:0045333	0.00007524454510147919	COX1,ATP6,COX2
GO:BP	ATP metabolic process	GO:0046034	0.00012625908206607866	COX1,ATP6,COX2
GO:BP	mitochondrial electron transport, cytochrome c to oxygen	GO:0006123	0.00013129276276461794	COX1,COX2
GO:BP	energy derivation by oxidation of organic compounds	GO:0015980	0.00013129276276461794	COX1,ATP6,COX2
GO:BP	ATP biosynthetic process	GO:0006754	0.0001608083421025898	ATP6,COX2
GO:BP	purine ribonucleoside triphosphate biosynthetic process	GO:0009206	0.00022046778176153746	ATP6,COX2
GO:BP	purine nucleoside triphosphate biosynthetic process	GO:0009145	0.00022046778176153746	ATP6,COX2
GO:BP	purine ribonucleoside triphosphate metabolic process	GO:0009205	0.00022961780750055307	ATP6,COX2
GO:BP	purine nucleoside triphosphate metabolic process	GO:0009144	0.00022961780750055307	ATP6,COX2
GO:BP	ribonucleoside triphosphate biosynthetic process	GO:0009201	0.0002590152238284992	ATP6,COX2
GO:BP	ribonucleoside triphosphate metabolic process	GO:0009199	0.0002885715408171042	ATP6,COX2
GO:BP	generation of precursor metabolites and energy	GO:0006091	0.0002988093068273851	COX1,ATP6,COX2
GO:BP	aerobic electron transport chain	GO:0019646	0.00034803447491391765	COX1,COX2
GO:BP	nucleoside triphosphate biosynthetic process	GO:0009142	0.0003778964796001131	ATP6,COX2
GO:BP	ATP synthesis coupled electron transport	GO:0042773	0.0003863596067271702	COX1,COX2
GO:BP	mitochondrial ATP synthesis coupled electron transport	GO:0042775	0.0003863596067271702	COX1,COX2
GO:BP	nucleoside triphosphate metabolic process	GO:0009141	0.0004678306557640051	ATP6,COX2
GO:BP	respiratory electron transport chain	GO:0022904	0.0005531267674892481	COX1,COX2
GO:BP	inorganic cation transmembrane transport	GO:0098662	0.0011500040247250328	COX1,ATP6,COX2

source	term_name	term_id	adjusted_p_value	intersections
GO:BP	inorganic ion transmembrane transport	GO:0098660	0.0014659345007680012	COX1,ATP6,COX2
GO:BP	cation transmembrane transport	GO:0098655	0.0015870908252496876	COX1,ATP6,COX2
GO:BP	purine ribonucleotide biosynthetic process	GO:0009152	0.0016242745295946919	ATP6,COX2
GO:BP	ribonucleotide biosynthetic process	GO:0009260	0.001722925341797247	ATP6,COX2
GO:BP	purine nucleotide biosynthetic process	GO:0006164	0.001722925341797247	ATP6,COX2
GO:BP	ribose phosphate biosynthetic process	GO:0046390	0.001722925341797247	ATP6,COX2
GO:BP	electron transport chain	GO:0022900	0.001722925341797247	COX1,COX2
GO:BP	purine-containing compound biosynthetic process	GO:0072522	0.0019205004584395093	ATP6,COX2
GO:BP	cation transport	GO:0006812	0.0024055899316878003	COX1,ATP6,COX2
GO:BP	response to abiotic stimulus	GO:0009628	0.0024597344073485116	COX1,ATP6,COX2
GO:BP	ion transmembrane transport	GO:0034220	0.0024597344073485116	COX1,ATP6,COX2
GO:BP	nucleotide biosynthetic process	GO:0009165	0.003599812109042362	ATP6,COX2
GO:BP	nucleoside phosphate biosynthetic process	GO:1901293	0.0036074023447451036	ATP6,COX2
GO:BP	ion transport	GO:0006811	0.004282264718044221	COX1,ATP6,COX2
GO:BP	transmembrane transport	GO:0055085	0.005075100277466673	COX1,ATP6,COX2
GO:BP	aging	GO:0007568	0.005075100277466673	COX1,ATP6
GO:BP	positive regulation of ATP biosynthetic process	GO:2001171	0.005533417925383506	COX2
GO:BP	positive regulation of purine nucleotide biosynthetic process	GO:1900373	0.005533417925383506	COX2
GO:BP	positive regulation of nucleotide biosynthetic process	GO:0030810	0.005533417925383506	COX2
GO:BP	mitochondrial ATP synthesis coupled proton transport	GO:0042776	0.005533417925383506	ATP6
GO:BP	energy coupled proton transport, down electrochemical gradient	GO:0015985	0.005533417925383506	ATP6
GO:BP	electron transport coupled proton transport	GO:0015990	0.005533417925383506	COX1
GO:BP	positive regulation of hydrogen peroxide biosynthetic process	GO:0010729	0.005533417925383506	COX2
GO:BP	ATP synthesis coupled proton transport	GO:0015986	0.005533417925383506	ATP6
GO:BP	energy coupled proton transmembrane transport, against electrochemical gradient	GO:0015988	0.005533417925383506	COX1
GO:BP	purine ribonucleotide metabolic process	GO:0009150	0.006225063416168587	ATP6,COX2
GO:BP	purine nucleotide metabolic process	GO:0006163	0.006468195340850997	ATP6,COX2
GO:BP	ribonucleotide metabolic process	GO:0009259	0.006575830430541989	ATP6,COX2
GO:BP	purine-containing compound metabolic process	GO:0072521	0.006575830430541989	ATP6,COX2
GO:BP	ribose phosphate metabolic process	GO:0019693	0.006575830430541989	ATP6,COX2
GO:BP	positive regulation of necrotic cell death	GO:0010940	0.006964535990132612	COX2
GO:BP	regulation of ATP biosynthetic process	GO:2001169	0.006964535990132612	COX2
GO:BP	regulation of hydrogen peroxide biosynthetic process	GO:0010728	0.006964535990132612	COX2
GO:BP	positive regulation of hydrogen peroxide metabolic process	GO:0010726	0.006964535990132612	COX2
GO:BP	nucleotide metabolic process	GO:0009117	0.010061968320321822	ATP6,COX2
GO:BP	nucleoside phosphate metabolic process	GO:0006753	0.010061968320321822	ATP6,COX2

source	term_name	term_id	adjusted_p_value	intersections
GO:BP	regulation of purine nucleotide biosynthetic process	GO:1900371	0.010312967254534093	COX2
GO:BP	organophosphate biosynthetic process	GO:0090407	0.010312967254534093	ATP6,COX2
GO:BP	response to hyperoxia	GO:0055093	0.010312967254534093	ATP6
GO:BP	regulation of hydrogen peroxide metabolic process	GO:0010310	0.010312967254534093	COX2
GO:BP	cristae formation	GO:0042407	0.010312967254534093	ATP6
GO:BP	nucleobase-containing small molecule metabolic process	GO:0055086	0.011158165346802749	ATP6,COX2
GO:BP	hydrogen peroxide biosynthetic process	GO:0050665	0.011462889920218633	COX2
GO:BP	regulation of nucleotide biosynthetic process	GO:0030808	0.011462889920218633	COX2
GO:BP	positive regulation of reactive oxygen species biosynthetic process	GO:1903428	0.011462889920218633	COX2
GO:BP	response to increased oxygen levels	GO:0036296	0.011462889920218633	ATP6
GO:BP	carbohydrate derivative biosynthetic process	GO:1901137	0.011713835976839768	ATP6,COX2
GO:BP	positive regulation of purine nucleotide metabolic process	GO:1900544	0.012454450080564382	COX2
GO:BP	response to electrical stimulus	GO:0051602	0.012454450080564382	COX1
GO:BP	positive regulation of nucleotide metabolic process	GO:0045981	0.012454450080564382	COX2
GO:BP	response to copper ion	GO:0046688	0.012454450080564382	COX1
GO:BP	positive regulation of ATP metabolic process	GO:1903580	0.014037999667666837	COX2
GO:BP	lactation	GO:0007595	0.015578505168081154	COX2
GO:BP	inner mitochondrial membrane organization	GO:0007007	0.017077665880659485	ATP6
GO:BP	regulation of reactive oxygen species biosynthetic process	GO:1903426	0.018299437792719857	COX2
GO:BP	regulation of necrotic cell death	GO:0010939	0.018299437792719857	COX2
GO:BP	hydrogen peroxide metabolic process	GO:0042743	0.020815799409644536	COX2
GO:BP	necrotic cell death	GO:0070265	0.020815799409644536	COX2
GO:BP	positive regulation of reactive oxygen species metabolic process	GO:2000379	0.020815799409644536	COX2
GO:BP	response to cold	GO:0009409	0.022138412738216335	COX2
GO:BP	organophosphate metabolic process	GO:0019637	0.022618044940423056	ATP6,COX2
GO:BP	response to stress	GO:0006950	0.022695730009552346	COX1,ATP6,COX2
GO:BP	reactive oxygen species biosynthetic process	GO:1903409	0.024112868393701463	COX2
GO:BP	body fluid secretion	GO:0007589	0.024112868393701463	COX2
GO:BP	carbohydrate derivative metabolic process	GO:1901135	0.0281936703686359	ATP6,COX2
GO:BP	positive regulation of small molecule metabolic process	GO:0062013	0.030892385491423015	COX2
GO:BP	regulation of purine nucleotide metabolic process	GO:1900542	0.03632910201892074	COX2
GO:BP	mammary gland development	GO:0030879	0.0369431613031256	COX2
GO:BP	mitochondrial membrane organization	GO:0007006	0.0369431613031256	ATP6
GO:BP	regulation of nucleotide metabolic process	GO:0006140	0.03753015408931815	COX2
GO:BP	regulation of reactive oxygen species metabolic process	GO:2000377	0.03753015408931815	COX2
GO:BP	mitochondrial transmembrane transport	GO:1990542	0.03849702180109943	ATP6

source	term_name	term_id	adjusted_p_value	intersections
GO:BP	regulation of ATP metabolic process	GO:1903578	0.03944289315818192	COX2
GO:BP	cerebellum development	GO:0021549	0.04170418494237256	COX1
GO:BP	metencephalon development	GO:0022037	0.045236449334031835	COX1
GO:CC	respiratory chain complex IV	GO:0045277	0.00012655567729787917	COX1,COX2
GO:CC	cytochrome complex	GO:0070069	0.00013284167391083507	COX1,COX2
GO:CC	organelle inner membrane	GO:0019866	0.00018424073999547653	COX1,ATP6,COX2
GO:CC	mitochondrial inner membrane	GO:0005743	0.00018424073999547653	COX1,ATP6,COX2
GO:CC	respiratory chain complex	GO:0098803	0.00019717852091603129	COX1,COX2
GO:CC	mitochondrial membrane	GO:0031966	0.00024243769652931423	COX1,ATP6,COX2
GO:CC	mitochondrial envelope	GO:0005740	0.0002595436065904971	COX1,ATP6,COX2
GO:CC	respirasome	GO:0070469	0.00026991729188410076	COX1,COX2
GO:CC	inner mitochondrial membrane protein complex	GO:0098800	0.00032395799216317816	COX1,ATP6
GO:CC	organelle envelope	GO:0031967	0.0009508092217745141	COX1,ATP6,COX2
GO:CC	envelope	GO:0031975	0.0009508092217745141	COX1,ATP6,COX2
GO:CC	mitochondrial protein-containing complex	GO:0098798	0.001085715461451425	COX1,ATP6
GO:CC	membrane protein complex	GO:0098796	0.001085715461451425	COX1,ATP6,COX2
GO:CC	mitochondrion	GO:0005739	0.002027756248327523	COX1,ATP6,COX2
GO:CC	mitochondrial respiratory chain complex IV	GO:0005751	0.0031594259637899432	COX1
GO:CC	mitochondrial proton-transporting ATP synthase complex	GO:0005753	0.0031594259637899432	ATP6
GO:CC	proton-transporting ATP synthase complex, coupling factor F(o)	GO:0045263	0.0031594259637899432	ATP6
GO:CC	mitochondrial respiratory chain complex III	GO:0005750	0.004027318637772337	COX1
GO:CC	respiratory chain complex III	GO:0045275	0.004027318637772337	COX1
GO:CC	proton-transporting ATP synthase complex	GO:0045259	0.004027318637772337	ATP6
GO:CC	proton-transporting two-sector ATPase complex, proton-transporting domain	GO:0033177	0.007665659304825296	ATP6
GO:CC	proton-transporting two-sector ATPase complex	GO:0016469	0.01096806811487927	ATP6
GO:CC	mitochondrial respirasome	GO:0005746	0.015139657510133716	COX1
GO:CC	organelle membrane	GO:0031090	0.017051852281193464	COX1,ATP6,COX2
GO:CC	oxidoreductase complex	GO:1990204	0.02670978761478042	COX1
GO:CC	integral component of membrane	GO:0016021	0.02948063301790454	COX1,ATP6,COX2
GO:CC	intrinsic component of membrane	GO:0031224	0.02988918317043817	COX1,ATP6,COX2

Supplementary Table 4. Results of the GO enrichment analysis for dowregulated DEGs in freezing conditions.

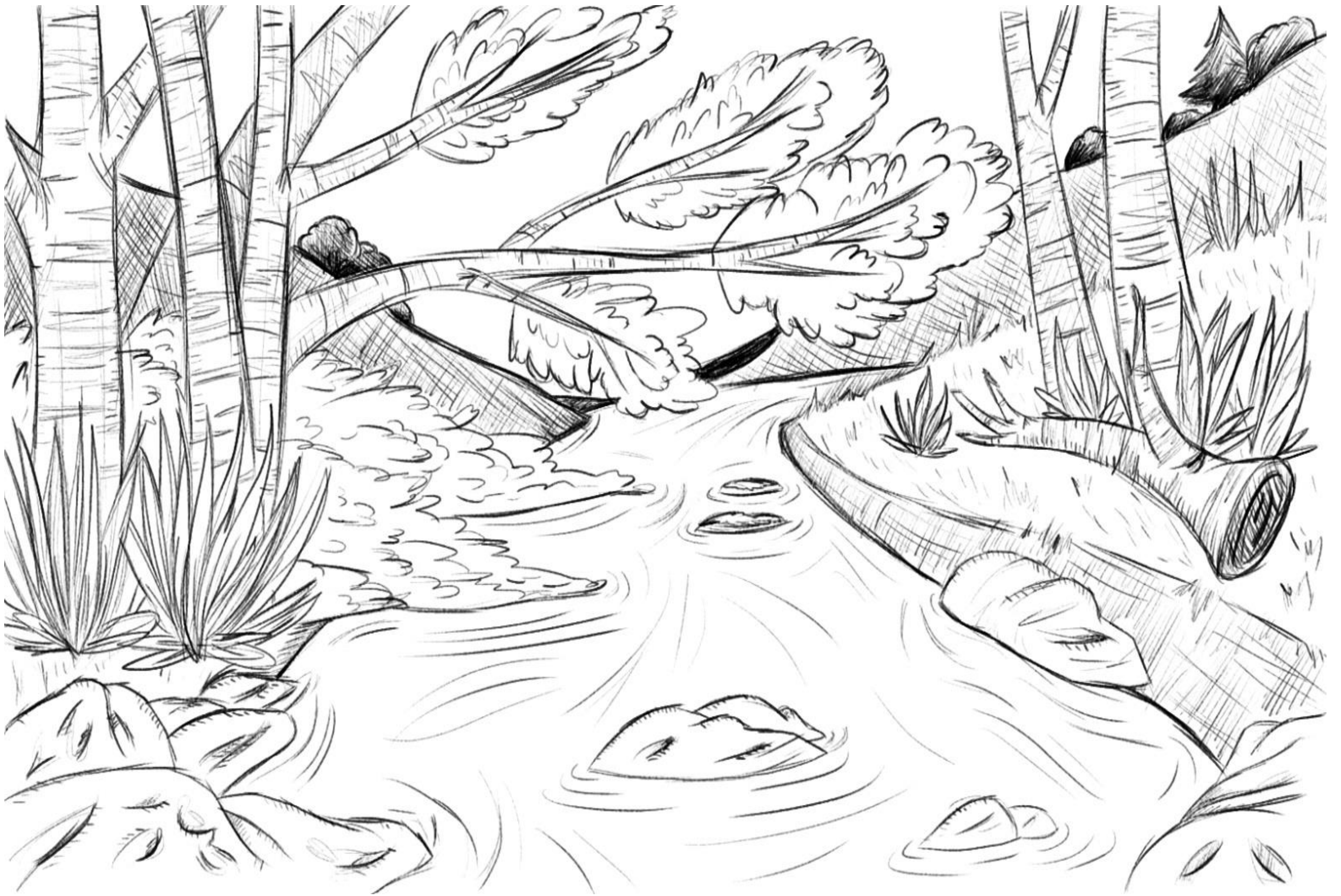
source	term_name	term_id	adjusted_p_value	intersections
GO:MF	uridine kinase activity	GO:0004849	0.005890669180018404	UCKL1
GO:MF	nucleoside kinase activity	GO:0019206	0.005890669180018404	UCKL1
GO:MF	nucleobase-containing compound kinase activity	GO:0019205	0.03141690229343612	UCKL1
GO:BP	UMP metabolic process	GO:0046049	0.006519007225887802	UCKL1
GO:BP	CTP metabolic process	GO:0046036	0.006519007225887802	UCKL1
GO:BP	CTP salvage	GO:0044211	0.006519007225887802	UCKL1
GO:BP	UMP salvage	GO:0044206	0.006519007225887802	UCKL1
GO:BP	pyrimidine nucleotide salvage	GO:0032262	0.006519007225887802	UCKL1
GO:BP	pyrimidine ribonucleotide salvage	GO:0010138	0.006519007225887802	UCKL1
GO:BP	pyrimidine ribonucleoside triphosphate biosynthetic process	GO:0009209	0.006519007225887802	UCKL1
GO:BP	pyrimidine ribonucleoside monophosphate biosynthetic process	GO:0009174	0.006519007225887802	UCKL1
GO:BP	pyrimidine ribonucleoside monophosphate metabolic process	GO:0009173	0.006519007225887802	UCKL1
GO:BP	pyrimidine ribonucleoside triphosphate metabolic process	GO:0009208	0.006519007225887802	UCKL1
GO:BP	CTP biosynthetic process	GO:0006241	0.006519007225887802	UCKL1
GO:BP	UMP biosynthetic process	GO:0006222	0.006519007225887802	UCKL1
GO:BP	pyrimidine nucleoside monophosphate biosynthetic process	GO:0009130	0.006519007225888058	UCKL1
GO:BP	pyrimidine nucleoside triphosphate biosynthetic process	GO:0009148	0.006519007225888058	UCKL1
GO:BP	nucleotide salvage	GO:0043173	0.006519007225888058	UCKL1
GO:BP	pyrimidine nucleoside salvage	GO:0043097	0.006519007225888058	UCKL1
GO:BP	pyrimidine ribonucleotide biosynthetic process	GO:0009220	0.006519007225888058	UCKL1
GO:BP	pyrimidine-containing compound salvage	GO:0008655	0.006519007225888058	UCKL1
GO:BP	nucleoside salvage	GO:0043174	0.0068449575871825374	UCKL1
GO:BP	pyrimidine ribonucleotide metabolic process	GO:0009218	0.0068449575871825374	UCKL1
GO:BP	pyrimidine nucleoside biosynthetic process	GO:0046134	0.007111644246423056	UCKL1
GO:BP	pyrimidine nucleoside triphosphate metabolic process	GO:0009147	0.007111644246423056	UCKL1
GO:BP	glycosyl compound biosynthetic process	GO:1901659	0.0071709079484765055	UCKL1
GO:BP	pyrimidine nucleoside metabolic process	GO:0006213	0.0071709079484765055	UCKL1

source	term_name	term_id	adjusted_p_value	intersections
GO:BP	pyrimidine nucleotide biosynthetic process	GO:0006221	0.0071709079484765055	UCKL1
GO:BP	cellular metabolic compound salvage	GO:0043094	0.0071709079484765055	UCKL1
GO:BP	pyrimidine nucleoside monophosphate metabolic process	GO:0009129	0.0071709079484765055	UCKL1
GO:BP	ribonucleoside triphosphate biosynthetic process	GO:0009201	0.0071709079484765055	UCKL1
GO:BP	ribonucleoside monophosphate biosynthetic process	GO:0009156	0.0071709079484765055	UCKL1
GO:BP	nucleoside biosynthetic process	GO:0009163	0.0071709079484765055	UCKL1
GO:BP	ribonucleoside triphosphate metabolic process	GO:0009199	0.007570460004256803	UCKL1
GO:BP	nucleoside monophosphate biosynthetic process	GO:0009124	0.008556196983977596	UCKL1
GO:BP	nucleoside triphosphate biosynthetic process	GO:0009142	0.008889555308028471	UCKL1
GO:BP	pyrimidine nucleotide metabolic process	GO:0006220	0.009514226762106398	UCKL1
GO:BP	pyrimidine-containing compound biosynthetic process	GO:0072528	0.009514226762106398	UCKL1
GO:BP	nucleoside triphosphate metabolic process	GO:0009141	0.009514226762106398	UCKL1
GO:BP	nucleoside metabolic process	GO:0009116	0.009514226762106398	UCKL1
GO:BP	glycosyl compound metabolic process	GO:1901657	0.011032166074578766	UCKL1
GO:BP	ribonucleoside monophosphate metabolic process	GO:0009161	0.011032166074578766	UCKL1
GO:BP	pyrimidine-containing compound metabolic process	GO:0072527	0.012712064090480868	UCKL1
GO:BP	nucleoside monophosphate metabolic process	GO:0009123	0.015264016919151703	UCKL1
GO:BP	nucleobase-containing small molecule biosynthetic process	GO:0034404	0.015831874691441476	UCKL1
GO:BP	ribonucleotide biosynthetic process	GO:0009260	0.018223588381458714	UCKL1
GO:BP	ribose phosphate biosynthetic process	GO:0046390	0.018223588381458714	UCKL1
GO:BP	nucleotide biosynthetic process	GO:0009165	0.027634921935827773	UCKL1
GO:BP	nucleoside phosphate biosynthetic process	GO:1901293	0.027634921935827773	UCKL1
GO:BP	ribose phosphate metabolic process	GO:0019693	0.04359586082312309	UCKL1
GO:BP	ribonucleotide metabolic process	GO:0009259	0.04359586082312309	UCKL1

Supplementary Table 5. Results of the GO enrichment analysis for upegulated DEGs in freezing conditions.

source	term_name	term_id	adjusted_p_value	intersections
GO:MF	adenylate cyclase activity	GO:0004016	0.02322933028094041	ADCY5
GO:MF	testosterone dehydrogenase (NAD+) activity	GO:0047035	0.02322933028094041	DHRS9
GO:MF	alpha-(1,2)-fucosyltransferase activity	GO:0031127	0.02322933028094041	FUT1
GO:MF	peptidase inhibitor activity	GO:0030414	0.02322933028094041	CD109,PI16
GO:MF	testosterone dehydrogenase [NAD(P)] activity	GO:0030283	0.02322933028094041	DHRS9
GO:MF	alcohol dehydrogenase [NAD(P)+] activity	GO:0018455	0.02322933028094041	DHRS9
GO:MF	peptidase regulator activity	GO:0061134	0.02322933028094041	CD109,PI16
GO:MF	isomerase activity	GO:0016853	0.02322933028094041	DHRS9,FKBP8
GO:MF	racemase and epimerase activity	GO:0016854	0.02322933028094041	DHRS9
GO:MF	adenylate cyclase binding	GO:0008179	0.02322933028094041	ADCY5
GO:MF	galactoside 2-alpha-L-fucosyltransferase activity	GO:0008107	0.02322933028094041	FUT1
GO:MF	alcohol dehydrogenase (NAD+) activity	GO:0004022	0.02322933028094041	DHRS9
GO:MF	fucosyltransferase activity	GO:0008417	0.027852205137735377	FUT1
GO:MF	androsterone dehydrogenase activity	GO:0047023	0.027852205137735377	DHRS9
GO:MF	androstan-3-alpha,17-beta-diol dehydrogenase activity	GO:0047044	0.027852205137735377	DHRS9
GO:MF	long-chain fatty acid transporter activity	GO:0005324	0.030921372910220124	ABCD4
GO:MF	protein folding chaperone	GO:0044183	0.030921372910220124	FKBP8
GO:MF	transforming growth factor beta binding	GO:0050431	0.030921372910220124	CD109
GO:MF	steroid dehydrogenase activity	GO:0016229	0.031598054275396595	DHRS9
GO:MF	enzyme inhibitor activity	GO:0004857	0.031598054275396595	CD109,PI16
GO:MF	steroid dehydrogenase activity	GO:0033764	0.031598054275396595	DHRS9
GO:MF	disordered domain specific binding	GO:0097718	0.031598054275396595	FKBP8
GO:MF	phosphorus-oxygen lyase activity	GO:0016849	0.038864639739918855	ADCY5
GO:MF	NAD-retinol dehydrogenase activity	GO:0004745	0.038864639739918855	DHRS9
GO:MF	cyclase activity	GO:0009975	0.038864639739918855	ADCY5

Supplementary Table 5. Results of the GO enrichment analysis for downregulatedDEGs in desiccation conditions.



Ilustraciones de portada y contraportada: Raúl Navarro©