Received: 14 September 2022 | Revised: 15 September 2022 | Accepted: 24 September 2022

DOI: 10.1111/jbi.14518

Journal of Biogeography

WILEY

RESEARCH ARTICLE

Phylogeography of a riparian earthworm shows environmental factors influence genetic structure

Irene de Sosa[1] | Daniel F. Marchán[2] | Marta Novo[1] | Ana Almodóvar[1] | Darío J. Díaz Cosín[1]

[1]Biodiversity, Ecology and Evolution Department, Faculty of Biology, Complutense University of Madrid, Madrid, Spain

[2]Centre d'Ecologie Fonctionnelle et Evolutive UMR 5175 CNRS 1919 Route de Mende, Montpellier Cedex 5, France

Correspondence
Irene de Sosa, Biodiversity, Ecology and Evolution Department, Faculty of Biology, Complutense University of Madrid, Madrid, Spain.
Email: iscarrasco@ucm.es

Funding information
European Regional Development Fund; Universidad Complutense de Madrid; Spanish Government, Grant/Award Number: PGC2018-094112-A-100 and FJCI-2017-32895; Ministry for Europe and Foreign Affairs, in collaboration with the Ministry for Higher Education and Research, Grant/Award Number: mopga-postdoc-3--6111272103; Ramón y Cajal Fellowship, Grant/Award Number: RYC2018-024654-I

Handling Editor: Michael Dawson

Abstract

Aim: The study of cosmopolitan earthworms could be even more interesting than that of endemic species in revealing evolutionary processes. Previous research on the cosmopolitan worm *Eiseniella tetraedra* has indicated some phylogeographic structure among populations, but the factors responsible remain unresolved. We hypothesized that environmental factors and dispersal have shaped the distribution of the species' lineages.

Location: Spain and Portugal; Iberian Peninsula.

Taxon: *Eiseniella tetraedra* (Lumbricidae, Oligochaeta, Annelida).

Methods: We collected 739 specimens of *Eiseniella tetraedra* from 65 localities around the Iberian Peninsula between 2012 and 2016. We performed phylogenetic analysis (Bayesian Inference and maximum likelihood) using two mitochondrial (COI and 16S) and one nuclear marker (28S). Furthermore, we studied their genetic diversity and historic demography based on the COI gene. Correlations between genetic diversity and 22 environmental factors were tested.

Results: *Eiseniella tetraedra* showed high diversity in the Iberian Peninsula, with eight different lineages nested in two clades. We found lineages mostly restricted to the northern region, while others were distributed throughout the Iberian Peninsula. Habitat stability, that is, constant availability or lack of water, also correlated with genetic diversity. Thus, although no clear phylogeographic pattern was found, environmental factors (such as precipitation, temperature, and soil pH) and habitat stability influenced the distribution of genetic variability.

Main Conclusions: *Eiseniella tetraedra* is an earthworm with great genetic variability. We show that the ranges of species with high relative dispersal ability and ambiguous phylogeographic patterns may be better explained by influence of environmental conditions rather than specific geographic features. Adaptation to unstable conditions has been shown to confer more success on one of the two major genetic clades recovered, pointing to ecological plasticity as a key for evolutionary success.

KEYWORDS
Annelida, dispersion, ecological diversity, *Eiseniella tetraedra*, genetic variability, semi-aquatic habitat

1 | INTRODUCTION

Phylogeographic studies suggest that southern Europe has acted as a refuge for species survival during periods of unfavourable climatic conditions (Hewitt, 2004). The topographic complexity and geographic mosaic of habitats in European southern refugia favour the occurrence of multiple disjunct refugia, allowing isolated populations to survive in these areas during glacial periods (Cooper & Hewitt, 1993; Hewitt, 1993, 1996). Within the Iberian Peninsula, complex species histories have been uncovered for a variety of taxa, with some showing remarkable patterns of phylogeographic concordance (Gómez & Lunt, 2007) that include deep genetic subdivisions, high haplotype richness, and distinct hybrid zones. The Iberian Peninsula not only facilitated the northern redistribution of species after the Ice Age but also diversification through patterns of repeated fragmentation, contraction, expansion, and mixing of populations. Thus, the Iberian Peninsula is considered a biodiversity hotspot (as a part of the Mediterranean biodiversity hotspot), owing to a high level of endemism and a great diversity of habitats and landscapes (Myers et al., 2000). 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 the peninsula effect (Simpson, 1964). Due to its geographical position, the Iberian Peninsula is under the influence of the Atlantic Ocean and Mediterranean Sea, resulting in a wide range of climates (Gómez & Lunt, 2007) and comprises two biogeographical regions, the Eurosiberian and the Mediterranean (Alcaraz et al., 2006).

As a result of glaciation in the Last Ice Age, as well as human activities, high local earthworm species richness has typically been found in the mid-latitudes where the Iberian Peninsula is located (Phillips et al., 2020). Rodriguez 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 characterised by mean annual precipitation in the range of 700–1000 mm. Moreover, several studies on earthworms showed genetic differences between Eurosiberian and Mediterranean regions in the Iberian Peninsula (Fernández et al., 2012, 2015; Rodriguez et al., 1997). Phylogeographic studies can help determine the history of the 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 *Eiseniella tetraedra* (Savigny, 1826) (de Sosa, Marchán, Novo, Almodóvar, et al., 2017; Javidkar et al., 2020; Terhivuo et al., 1994, 2011).

Eiseniella tetraedra is a cosmopolitan (Blakemore, 2006), parthenogenetic tetraploid (Casellato, 1987) earthworm associated with aquatic or semiaquatic habitats (Omodeo & Rota, 1991). In preliminary Iberian studies, high haplotype diversity was nested in six different lineages without a clear pattern in the population structure (de Sosa, Marchán, Novo, Almodóvar, et al., 2017). The same results were found in the Alborz Mountains (Iran), with the presence of the same six lineages probably introduced by human activities (Javidkar et al., 2020). However, *E. tetraedra* did not show a clear distribution pattern (de Sosa, Marchán, Novo, Almodóvar, et al., 2017; Javidkar et al., 2020). The fact that this earthworm is distributed by hydrochory (Terhivuo & Saura, 2006), anthropochory (Gates, 1977; Javidkar et al., 2020), and even zoochory (Terhivuo & Saura, 2006) indicates that it may show ambiguous phylogeographic patterns. For this reason, our hypothesis is that current environmental factors may better explain the distribution of their lineages.

As a riparian earthworm, *E. tetraedra* appears to be closely tied to water margins. It is found in rivers, streams, or even unstable habitats that may be frozen or dried out, depending on the season. The presence of *E. tetraedra* in this type of habitats can be explained by the fact that it is parthenogenetic: a single propagule suffices to establish a new population (Terhivuo & Saura, 2006). This ability, as well as the ease of dispersal already mentioned, may be key to the cosmopolitan distribution of the species. Moreover, in the Aland archipelago (Baltic Sea), all individuals of *E. tetraedra* disappear every year due to freezing in their habitat, followed by recolonisation by different clones of the species (Terhivuo & Saura, 1997). The same pattern could be present in the unstable habitats of the Iberian Peninsula. Thus, examining the differential distribution of *E. tetraedra* lineages in stable and unstable habitats may shed light on this colonisation pattern.

This species is characterised by a quadrangular posterior transverse section, a small number of segments, and a variable position of the male pores (more often XIII) (Omodeo & Rota, 1991). This variation has been evaluated in several taxonomic works; some (Blakemore, 2006; Michaelsen, 1900) considered the divergent forms as subspecies, others (Michaelsen, 1910, 1932; Omodeo, 1952, 1956; Plisko, 1965; Pop, 1952; Zicsi, 1960) considered them vaguely as "formae" or "varietates," and finally others (Černosvitov, 1942; Michaelsen, 1932; Omodeo & Rota, 1989) considered them genetic mutations. However, (de Sosa, Marchán, Novo, Almodóvar, et al., 2017) showed that different positions of male pores have no phylogenetic basis. Length, weight and the number of segments also showed a high degree of variability, but no significant differences were found between lineages (de Sosa, Marchán, Novo, Almodóvar, et al., 2017). Bergmann's rule states that animals in cold climates tend to increase their body size (Bergmann, 1848). It is possible that *E. tetraedra* could follow this trend.

Using a combination of genetic markers and a complete sample in the Iberian Peninsula, we tested several hypotheses. Based on previous studies, we hypothesized that current environmental

factors better explain the distribution of their lineages. Second, considering that *E. tetraedra* inhabits both stable and unstable habitats, lineages that resist unfavourable areas will be the most successful in terms of variables such as diversity and dispersal. Finally, we hypothesize that the morphological differences in this species are due to environmental factors and do not have a phylogenetic basis.

2 | MATERIALS AND METHODS

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 (de Sosa, Marchán, Novo, Almodóvar, et al., 2017). For the selection of the remaining 36 localities, we used random sampling in 50 × 50 km UTM cells to represent their distribution along the entire Iberian Peninsula using ArcMap 9.3 software (Environmental Systems Resource Institute, ArcMap 9.3 ESRI, Redlands, California; see Appendix S1 in Supporting Information. See geographic coordinates in Appendix S2).

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 739 individuals, focusing on length, dry weight and number of segments. Due to the maturity of the specimens, whenever possible, the position of the clitellum (237 specimens) and *tubercula pubertatis* (254 specimens), the position of male pores (272 specimens), and the number and position of seminal vesicles (240 specimens) were also studied. To test the existence of individuals with biparental reproduction, the presence or absence of iridescence in spermathecae and male funnels (240 specimens) was studied. The presence of iridescence indicates the presence of sperm (Plisko, 2002).

Whenever possible, we selected 10 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. The samples were then stored in ethanol and preserved at −20°C for genetic analysis.

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.

Primer sequences and polymerase chain reaction (PCR) for COI (632 bp) followed Pop et al. (2003), and those for 16S-tRNAs (775 bp) and 28S (806 bp) followed Fernández et al. (2015). PCR was specific and resolved via 1% agarose gel electrophoresis; gels were visualised with GelRed stain (Biotium, Fremont, California, United States). All products were purified using ExoSAP-IT reagent (ThermoFisher Scientific).

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

2.3 | Genetic data analyses

Sequences for each fragment were aligned in MAFFT v.7 (Katoh & Standley, 2013) 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 the presence of a reading frame that did not have a stop codon was determined to avoid the presence of pseudogenes in the dataset (Buhay, 2009). Phylogenetic trees based on the concatenated sequences of the three genes (1906 bp) and each gene were constructed by Bayesian inference (BI) with MrBayes v3.1.2 (Ronquist & Huelsenbeck, 2003) and maximum likelihood (ML) using RaxML v7.03 software (Stamatakis, 2006). Both were implemented in Cipres Science Gateway v.3.3 (Miller et al., 2010). To test the effects of missing data on phylogenetic signals and tree support, we also built a phylogenetic tree with individuals having all three markers. Then, we built a phylogenetic tree with the whole dataset but used a multifurcating constraining tree for those nodes having strong support in the first analysis. We inferred the maximum likelihood tree using IQ-TREE2 (Minh et al., 2020). The obtained phylogenetic trees were visualised in FigTree v1.3.1 (Morariu et al., 2008). The best-fitting substitution model selected by jModelTest2 (Darriba et al., 2012) for codon partition of COI, 16S-tRNAs, and 28S was GTR + Γ + I.ML analysis with rapid bootstrapping and was performed with 1000 replicates. Parameters in MrBayes were set to 10 million generations, and 10,000 trees were sampled every 1000 generations, 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), *Eiseniona oliveirae* (Rosa, 1894), and *Prosellodrilus biauriculatus* (Bouché, 1972) were retrieved from GenBank and used as outgroups (see Appendix S2).

Haplotype networks based on COI and 28S were constructed in PopART 1.7 (Leigh & Bryant, 2015) using statistical parsimony. Furthermore, Arlequin v.3.5 (Excoffier & Lischer, 2010) was used to perform an analysis of the molecular variance (AMOVA) with 10,000 permutations for statistical confidence. To determine whether the population genetic structure existed at different levels, the analysis was performed following a hierarchical structure: first with main

clades and then with lineages (see Section 3). In addition, uncorrected pairwise distances for COI and 16S tRNAs were calculated within and between the 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, were calculated with DNAsp v.6 (Rozas et al., 2017).

2.4 | Environmental factor analyses

Correlations between genetic diversity (haplotype and nucleotide) and parameters such as latitude, longitude and altitude of the localities were tested using Statgraphics Centurion 18 (StatPoint Technologies Inc.). Additionally, 19 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 (see Appendix S2). Differences between lineages and clades (see Section 3) regarding environmental variable values for their localities were compared by one-way analyses of variance (ANOVA) and *T*-test, followed by Fisher's LSD post hoc test. Non-parametric analyses (Kruskal–Wallis and Mann–Whitney *U* tests) with subsequent Fisher's LSD post hoc tests were run for those who did not fulfil the assumptions of normality and homoscedasticity (verified through Kolmogorov–Smirnov and Levene's test, respectively). The significance level for all statistical tests was set at $\alpha = 0.05$.

We also explored possible differences in the proportion of presence between lineages and clades due to the stability of the water bodies 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, χ^2 tests were performed in IBM SPSS Statistics v.24.

2.5 | Morphological analyses

Statistical analyses of morphological data were conducted in Statgraphics Centurion 18 (StatPoint Technologies Inc.). We used length, dry weight (after letting it drip on filter paper for 30 s), 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's LSD post hoc test. We also examined the effects of the altitudinal gradient on morphological variation using a simple linear regression model.

3 | RESULTS

3.1 | Phylogenetic analyses

The analysis of the most variable gene (COI) revealed eight distinct and well-supported lineages (labelled A to H; Figure 1). These lineages were clustered into two clades (labelled 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 the support values were lower (Figure 1a). Clade II showed higher genetic diversity, with 67% of the total haplotypes. The different lineages showed deep divergence, as shown by the long branches of the tree (Figure 1a). 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 (see Appendix S1). To improve the concatenated tree, we decided to remove specimens whose sequences had a 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; Figure 1b). Similarly, the constraint tree recovered the two main clades and most of the lineages (see Appendix S1). The differences between these trees (Figure 1) are due to the fact that the COI gene is highly variable even at the species level, so it can detect lineages within *E. tetraedra* but not deeper phylogenetic relationships. The multigene tree with the 28S gene, a much more conserved gene, is therefore able to support subdivision between the two clades, but not between all lineages.

3.2 | Lineage distribution

The distribution of lineages in the Iberian Peninsula is shown in Figure 2 and in Appendix S1. Eight lineages were found, at least in the northern half of the Iberian Peninsula. Most individuals of lineage B were distributed in the north-eastern area, and most individuals of lineages C and D were distributed in the north-western 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).

3.3 | Genetic diversity, genetic divergence and population structure

The AMOVA results indicated that most of the observed genetic variation (82.90%) was explained by differences among lineages (see Appendix S2). 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 (see Appendix S2).

A total of 102 haplotypes were identified among 372 sequences for the COI gene, 34 haplotypes within 148 sequences for

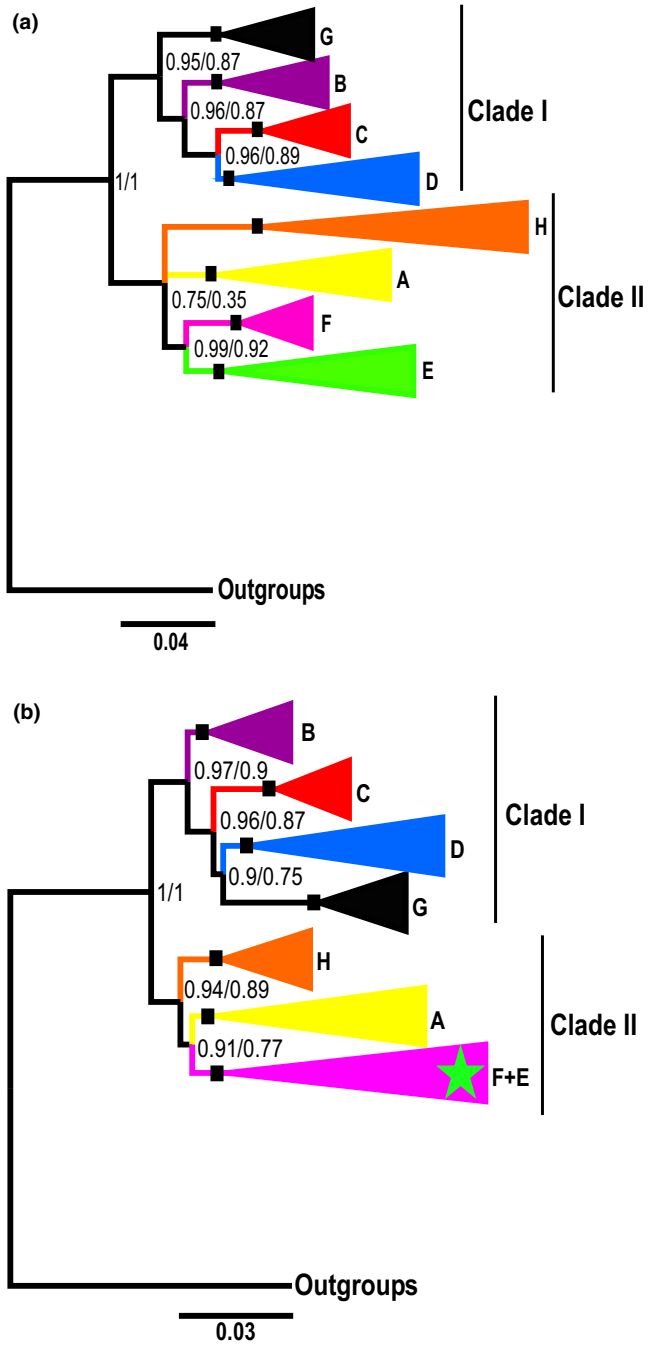

FIGURE 1 (a) Bayesian inference (BI) of the phylogenetic tree of *Eiseniella tetraedra* based on sequences of the COI gene. Posterior probability/bootstrap support values (of maximum likelihood analysis, ML) are shown. When this was not possible, they are shown as black squares. The scale bar represents 0.04 substitutions per position. Colours and names of lineages (A–H) are the same as in (de Sosa, Marchán, Novo, Almodóvar, & Díaz Cosín, 2017). (b) Bayesian inference (BI) of the phylogenetic tree of *E. tetraedra* based on sequences of COI, 16S and 28S. Specimens whose sequences had high average of missing data in rake-like polytomies were removed. Posterior probability/bootstrap support values (of maximum likelihood analysis, ML) are shown. The scale bar represents 0.03 substitutions per position. Posterior probability/bootstrap support values (of maximum likelihood analysis, ML) are shown. When this was not possible, they are shown as black squares. Colours and names of lineages (A–H) are the same as in (de Sosa, Marchán, Novo, Almodóvar, et al., 2017)

16S-tRNAs, and 16 haplotypes within 128 sequences for 28S. The values of haplotype and nucleotide diversity for each lineage and clade are shown in Appendix S2. Haplotype diversity (H) and nucleotide diversity (π) based on COI, including all specimens within the study, were 0.82 and 0.055, respectively. H and π based on 16S-tRNAs were 0.88 and 0.022, respectively. Finally, the genetic diversity parameters based on 28S were 0.39/0.006.

Genetic distances within lineages based on COI were in the range of 0.25 to 3.64%, showing moderate variability. The divergence between lineages was remarkably higher (2.91 to 7.79%) (see Appendix S2). However, values between lineages were lower than 9%–15%, which is the ambiguous gap between intraspecific and interspecific divergence in earthworms proposed by some authors (Chang & James, 2011; Decaëns et al., 2013; Rougerie et al., 2009). The genetic divergence between clades I and II was 10.26% for COI and 4.16% for 16S. The average divergence within clade I was 7.33%/3.24% (COI/16S), while within clade II, the average genetic distance was 5.21%/2.69% (COI/16S). Genetic distances based on 16S-tRNAs were lower than those based on COI, as 16 s-tRNAs is a more preserved region of the mitochondrial genome.

Most of the localities studied presented high haplotype diversity (average 0.70). As displayed by the haplotype network based on COI from each lineage (see Appendix S1), most of the lineages showed a star-shaped network topology structured around highly frequent haplotypes. The haplotype network based on the nuclear marker (28S) displayed the relationships among the 16 haplotypes (see Appendix S1). A star-shaped network topology with a central main haplotype was observed. No clear differentiation between clades or lineages was observed.

3.4 | Historic demography

Mismatch distributions based on COI were tested for each of the eight lineages recovered in phylogenetic analyses (see Appendix S1). 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.

The neutrality tests showed different results (see Appendix S2). 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.

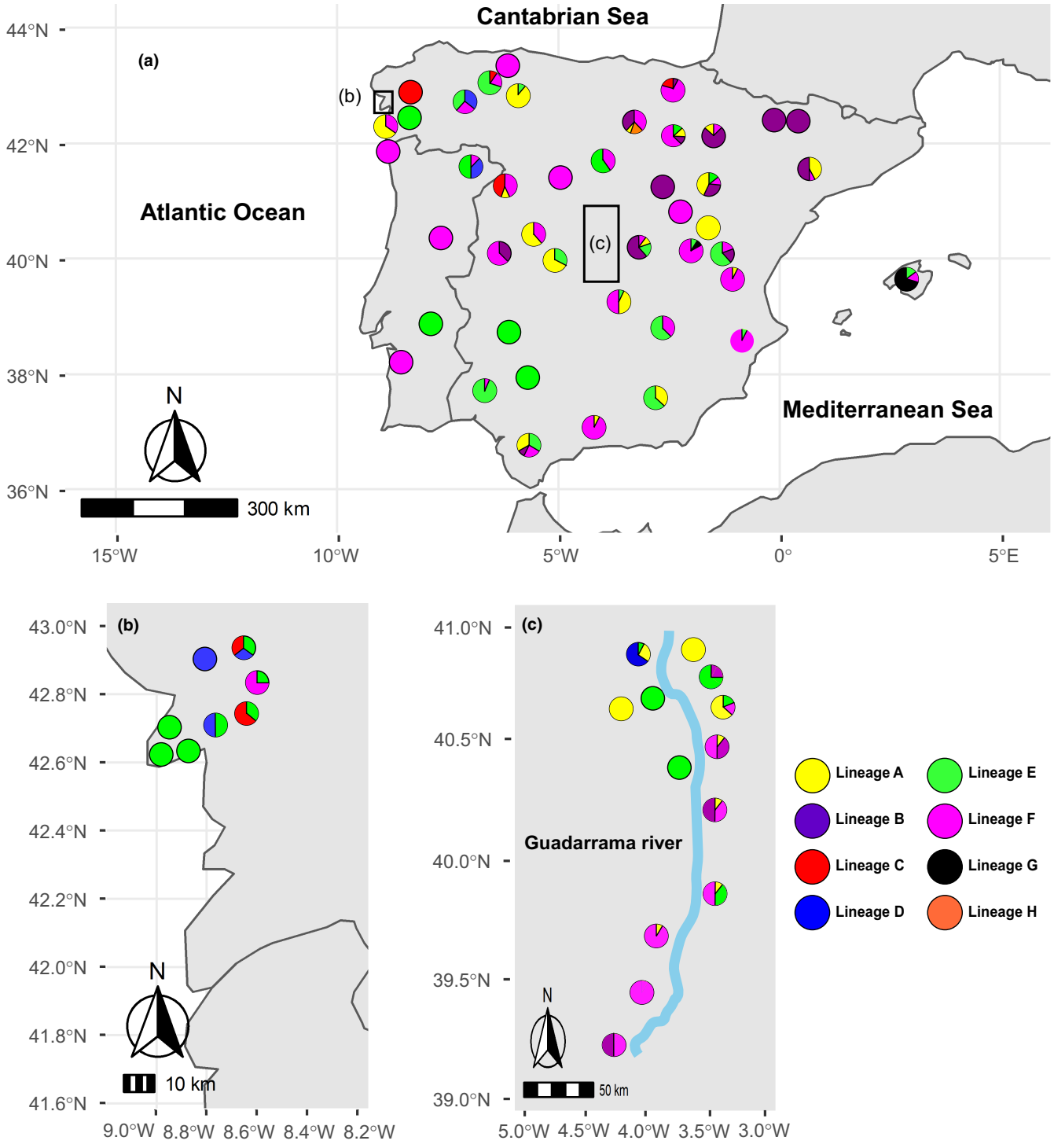

FIGURE 2 (a) Lineage distribution of *Eiseniella tetraedra* in the Iberian Peninsula. Proportion of individuals from each genetic lineage in each locality is represented in pie charts. Colours used are the same as in Figure 1. (b) Lineages distribution of *E. tetraedra* in localities from a lower-scale study in Carnota, A Coruña, Spain (de Sosa, Marchán, Novo, Almodóvar, et al., 2017). Proportion of individuals from each genetic lineage in each locality is represented in pie charts. Colours used are the same as in Figure 1. (c)Lineages distribution of *E. tetraedra* in localities from a lower-scale study in Guadarrama river basin, Madrid, Spain (de Sosa, Marchán, Novo, Almodóvar, et al., 2017). Proportion of individuals from each genetic lineage in each locality is represented in pie charts. Colours used are the same as in Figure 1

3.5 | Environmental factor analyses

No statistical correlations were found between altitude, longitude, latitude of the localities and genetic diversity (haplotype and nucleotide). Significant differences were found between lineages regarding 11 climatic variables from WorldClim (Table 1). Thus, lineages C and D occurred in damp and cold localities with more acidic soil, whereas lineages A, B, E and F occurred more in dry and hot

TABLE 1 Statistical results for each lineage regarding temperature, precipitation and soil factors in ANOVA (F) or Kruskal-Wallis (K)

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 ($kg \cdot m^{-2}$)										
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

Note: Different letters in the same row indicate different groups in multiple range test.

Abbreviation: SD, standard deviation.

TABLE 2 Statistical results for each clade regarding temperature, precipitation and soil factors in U Mann–Whitney 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 ($kg \cdot m^{-2}$)				
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

Abbreviation: SD, standard deviation.

localities. Eight temperature and precipitation factors showed significant differences between clades (Table 2).

We distinguished 26 stable and 38 unstable habitats within the sampled localities (see Appendix S2). The results of the χ^2 test are presented in Figure 3. Lineage D was significantly more frequent in stable habitats, whereas lineage E was significantly more frequent in unstable habitats. Moreover, Clade I appeared more frequently in stable habitats, while Clade II was found more frequently in unstable habitats.

3.6 | Morphological analyses

High morphological variability was observed in the examined specimens. Only mature individuals were used for morphological analyses. Spermathecae and male funnels were never iridescent and were absent in 66.25% and 69.16% of the individuals, respectively. We found variability in the position of the male pores. In most of the specimens (95.22%), the male pore was located in segment 13. However, in 2.57%, male pores were found in segment 12; 0.73% in segment 15; and 0.37% in segments 8, 9, 11 and 14. Individuals sharing the same haplotype had different states of this trait, and no association was found between the position of the male pore, the position of the clitellum or *tubercula pubertatis* (also showing some variability, see Appendix S2), 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 (see Appendix S1), but none were found between clades. Regarding length, specimens from lineage C were significantly longer than specimens from the other lineages. Within weight, lineage E was significantly heavier than lineage A. In addition, lineages D and E were significantly heavier than lineage B. Lineages G and H were excluded from these analyses because mature specimens were unavailable.

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

4 | DISCUSSION

Some morphologically distinct species are complexes of genetically well-individualised lineages, some of which may be considered cryptic species. This is particularly true of cosmopolitan species (King et al., 2008; Porco et al., 2013, 2018; Taheri et al., 2018). In a previous study of *Eiseniella tetraedra* in two different regions in the Iberian Peninsula (de Sosa, Marchán, Novo, Almodóvar, et al., 2017), six different lineages were found. Owing to the fact that, in this paper, the number of sample localities (covering the whole peninsula) increased, phylogenetic analysis based on the COI of Iberian *E. tetraedra* revealed eight distinct lineages nested in two clades. Because COI is a more variable region of the mitochondrial genome, the lineages had 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 the 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 clades of *E. tetraedra* (I and II) might be in a process of cryptic speciation due to their high COI divergences, 10.26% of which are within the 9% and 15% intervals between intraspecific and interspecific divergences proposed by Chang and James (2011), although the lack of support in phylogenetic trees suggests that this process may be incipient. Other parthenogenetic and cosmopolitan species have also shown cryptic divergence lineages, such as *Octolasion tyrtaeum* Savigny, 1826 (Heethoff et al., 2004), *Pontoscolex corethrurus* (Taheri et al., 2018), and *Aporrectodea trapezoides* (Dugès, 1828), which showed underlying speciation for lineages I and II (Fernández et al., 2011).

Neutrality tests of the most widespread lineages in the Iberian Peninsula (A, E and F) suggested demographic expansion. However, this pattern is not clear in the mismatch distributions. Harpending (1994) noted that an excessively recent population

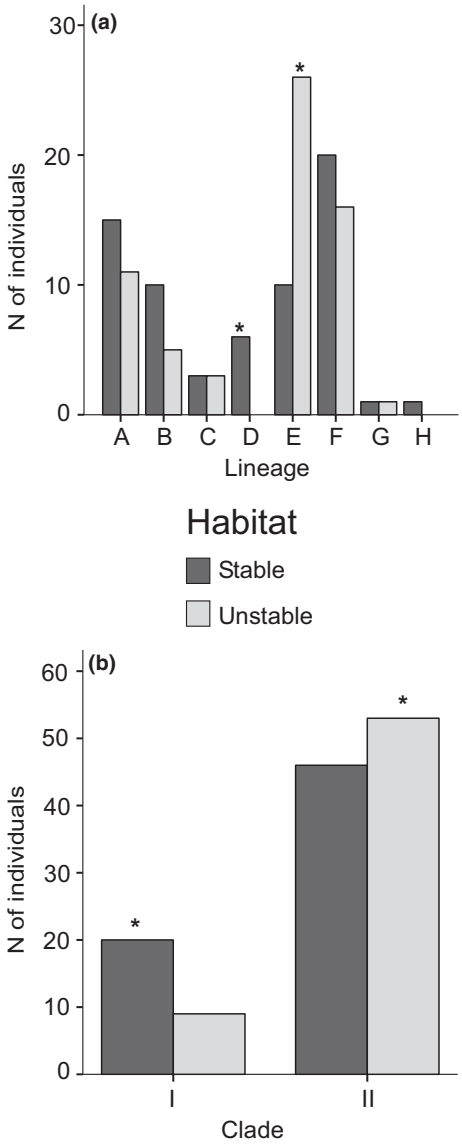

FIGURE 3 Results of χ^2 test for lineages (a) and clades (b) of *Eiseniella tetraedra* and number of habitats stable or unstable. Asterisks indicate statistically significant differences according to χ^2 test ($p < 0.05$)

expansion, for example at the end of the Pleistocene, would not result in a smooth unimodal mismatch distribution (as would be expected in such a scenario). The Iberian Peninsula served as a refuge during Pleistocene glaciation (Gómez & 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 the AMOVA results. However, two patterns can be discerned: first, the lineages that occurred throughout the peninsula, and second, the lineages that occurred mainly or only in the northern region of the peninsula. Within the northern lineages, we found a pattern suggesting the existence of two Eurosiberian lineages (C and D) and one Mediterranean lineage (B). Nevertheless, the presence of these lineages in specific localities outside their area of predominance 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). Other phylogeographic studies of cosmopolitan and polyploid earthworms in the Iberian Peninsula, such as *A. trapezoides* and *Aporrectodea rosea* (Savigny, 1826), showed similar phylogeographic patterns (Fernández et al., 2012, 2015). Finally, lineages G and H were present in only one or two localities, so no robust conclusions could be drawn about them, although they could be explained by point mutations in these populations. The geographic 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. Confirming our first hypothesis, the distribution of clades and lineages within *E. tetraedra* could be partially explained by the detected ecological preferences, such as temperature, precipitation and pH. Phillips et al. (2020) found that precipitation, followed by habitat cover and temperature, was the most important driver shaping diversity and distribution patterns in earthworms. Additionally, pH influences communities of earthworms (Rutgers et al., 2009, 2016). This genus is thought to date back to 60 million years ago (Domínguez et al., 2015). With the exception of *E. tetraedra* and *Eiseniella neapolitana* (Örley, 1885), which have cosmopolitan and circum-Mediterranean distributions, respectively, the remaining species of the genus occur in the Balkan Peninsula. It is therefore possible that this region is their original range. Thus, clades could have originated in the Balkanic Peninsula, probably due to a multiple and decentralised origin of parthenogenesis in this species (as in other earthworms; Fernández et al., 2011), and later spread to the Iberian Peninsula, so that their current distribution only reflects their ability to adapt to the local environment in the places where they were introduced. This pattern is probably due to the fact that it is a species with great dispersal ability. So much so that it was able to blur the biogeographic patterns typical of historical biogeography to follow those typical of ecological biogeography. In the case of microorganisms, it is generally assumed that geography has no influence on distribution, which is usually expressed as "everything is everywhere; the environment selects" (Becking, 1934). This affirmation does not mean that biogeographic patterns do not exist, but rather that since microbes disperse without boundaries, environmental conditions determine dispersal rather than anything specifically geographic. Thus, microbial evolution is ecologically controlled and geography plays no role (O'Malley, 2008). Thus, because of the great dispersal ability of this species, its pattern is more like that of microorganisms. This finding could open a door to all those species for which historical biogeography provides confusing answers.

Recent glaciations may have led to the extinction of most northern European earthworm populations (Mathieu & Davies, 2014). The high genetic diversity, mismatch distribution, and the presence of specific lineages (de Sosa, Marchán, Novo,

Almodóvar, et al., 2017) suggest that the Iberian Peninsula acted as a glacial refuge for *E. tetraedra* in the LGM. Furthermore, (de Sosa, Marchán, Novo, Almodóvar, et al., 2017) showed that lineage diversity is much lower in northern Europe than in the Iberian Peninsula. This pattern is similar to that described by Mathieu and Davies (2014) for France, based on species diversity rather than genetic diversity. According to the authors, this latitudinal gradient in diversity is due to the dispersal history of the different lineages after the Ice Age. However, to confirm or refute this hypothesis, the study needs to be extended by introducing new populations from Europe. We found more genetic variability in the northern half of the Iberian Peninsula, which may also respond to adaptations to the local environments of individuals and not to the peninsula effect (Simpson, 1964), as a gradual loss of genetic variability from the continent was not found.

Eiseniella tetraedra is closely tied to the edges of water, regardless of its stability. Clade I was found to be more adapted to stable habitats, while Clade II appeared to be adapted to more unstable habitats. Thus, it can be assumed that Clade II has greater resistance to poor conditions and greater colonisation potential. The second hypothesis is confirmed by the fact that it is the most successful clade, as shown by the greater diversity of haplotypes and their greater occurrence and distribution area in the Iberian Peninsula.

Variability in morphological characteristics related to sexual reproduction, such as seminal vesicles or spermathecae, was observed in the 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 (Blakemore, 2006; Bouché, 1972; Gates, 1977; Gavrilov, 1939; Terhivuo et al., 1994), could be explained by the same reason. Although several varieties or even subspecies have been described based on this trait (Blakemore, 2006; Bouché, 1972; Gates, 1977), no phylogenetic basis has been found (de Sosa, Marchán, Novo, Almodóvar, et al., 2017). No evidence of sexual reproduction was found in 739 specimens, so *E. tetraedra* appears to be strictly pathenogenetic in the Iberian Peninsula. Thus, these phenotypic traits are unlikely to be subject to selection pressure or to affect their fitness, so that different ranges occur in the morphological traits that are maintained over time. This was not the case for *A. trapezoides*, which has extremely rare sexual forms in the Iberian Peninsula and Algeria (de Sosa, Marchán, Novo, Cosín, et al., 2017; Fernández et al., 2011). No significant results were obtained when studying the effects of the altitudinal gradient on morphological characteristics. The high altitude areas are colder, so Bergmann's rule would not explain the morphological differences in *E. tetraedra*.

5 | CONCLUSIONS

Our results show that species with relatively high dispersal ability such as *E. tetraedra* are influenced by environmental conditions rather than specific geographic features making the presence of biogeographic patterns difficult to detect. However, when dispersal capacity is still limited, as in any soil organism, some patterns could be discerned. The split of *E. tetraedra* in two large genetic clades could be due to differentiation prior to their arrival in the Iberian Peninsula, probably in the area of origin of the genus, as a result of the multiple and decentralised origin of parthenogenesis in this species (as in other earthworms). We have found evidence that the Iberian Peninsula served as a glacial refuge for this species during LGM. To confirm or refute this hypothesis, it is necessary to complete the phylogeography of this species with new populations from Europe. As we hypothesized, the most successful clade was the one that could adapt to unstable conditions. Finally, most of the morphological variations did not show a phylogenetic basis or association with different altitudinal ranges, probably due to the parthenogenesis of this species, as they are not subjected to selection pressure.

ACKNOWLEDGEMENTS

We thank Daniel Romero for field support and Pablo Refoyo for his help in the selection of localities. IS was supported by a Predoctoral Fellowship grant from Universidad Complutense de Madrid, Spain. DF was supported by a Juan de la Cierva Formación grant from the Spanish Government (FJCI-2017-32895) and a MOPGA grant from the Ministry for Europe and Foreign Affairs, in collaboration with the Ministry for Higher Education and Research (mopga-postdoc-3--6111272103). MN was supported by a Ramón y Cajal Fellowship (RYC2018-024654-I) from MCIN/AEI/10.13039/501100011033 and "ESF: Investing in your future." This research was funded by project CGL2013-42908-P from the Spanish Government, by Grant PGC2018-094112-A-100 from MCIN/AEI/10.13039/501100011033, and by "ERDF: A way of making Europe." No permits were needed for this work.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in GenBank at https://www.ncbi.nlm.nih.gov/genbank/. Reference numbers are available in Appendix S2.

ORCID

Irene de Sosa https://orcid.org/0000-0002-4020-6375
Daniel F. Marchán https://orcid.org/0000-0001-7971-4329
Marta Novo https://orcid.org/0000-0001-7902-3819
Ana Almodóvar https://orcid.org/0000-0003-1465-3857
Darío J. Díaz Cosín https://orcid.org/0000-0001-8415-3001

REFERENCES

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.

Becking, L. G. M. B. (1934). *Geobiologie, of inleiding tot de milieukunde: Met literatuurlijst en ind*. Van Stockum.

Bergmann, C. (1848). Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse.

Blakemore, R. J. (2006). *Cosmopolitan earthworms-an eco-taxonomic guide to the peregrine species of the world* (2nd ed.). VermEcology.

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*(3), 226–236.

Bouché, M. B. (1972). Lombriciens de France: Écologie et systématique. *Annales de Zoologie Ecologie Animale*, *75*(2), 214–218.

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*(3), 339–359.

Buhay, J. E. (2009). "COI-like" sequences are becoming problematic in molecular systematic and DNA barcoding studies. *Journal of Crustacean Biology*, *29*(1), 96–110.

Casellato, S. (1987). On polyploidy in oligochaetes with particular references to Lumbricidae. In A. M. B. Pagliai, & P. Omodeo (Eds.), *Proceedings on the international symposium on earthworms* (pp. 75–84). Mucchi Editore.

Černosvitov, L. (1942). Oligochaeta from various parts of the world. *Proceedings of Zoological Society of London*, *111*(3–4), 197–236.

Chang, C. H., & James, S. (2011). A critique of earthworm molecular phylogenetics. *Pedobiologia*, *54*, S3–S9.

Cooper, S. J. B., & Hewitt, G. M. (1993). Nuclear DNA sequence divergence between parapatric subspecies of the grasshopper *Chorthippus parallelus*. *Insect Molecular Biology*, *2*, 185–194.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: More models, new heuristics and parallel computing. *Nature Methods*, *9*(8), 772.

de Sosa, I., Marchán, D. F., Novo, M., Almodóvar, A., & Díaz Cosín, D. J. (2017). Bless this phylogeographic mess–comparative study of *Eiseniella tetraedra* (Annelida, Oligochaeta) between an Atlantic area and a continental Mediterranean area in Spain. *European Journal of Soil Biology*, *78*, 50–56.

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.

Decaëns, T., Porco, D., Rougerie, R., Brown, G. G., & James, S. W. (2013). Potential of DNA barcoding for earthworm research in taxonomy and ecology. *Applied Soil Ecology*, *65*, 35–42.

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.

Excoffier, L., & Lischer, H. E. L. (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.

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*(10), 2085–2093.

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., 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., Novo, M., Marchán, D. F., & Díaz Cosín, D. J. (2015). Diversification patterns in cosmopolitan earthworms: Similar mode but different tempo. *Molecular Phylogenetics and Evolution*, *94*, 701–708.

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

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 S. Weiss, & N. Ferrand (Eds.), *Phylogeography of southern European refugia* (pp. 155–188). Springer.

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

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

Heethoff, M., Etzold, K., & Scheu, S. (2004). Mitochondrial COII sequences indicate that the parthenogenetic earthworm *Octolasion tyrtaeum* (Savigny 1826) constitutes of two lineages differing in body size and genotype. *Pedobiologia*, *48*(1), 9–13.

Hewitt, G. M. (1993). After the ice: Parallelus meets Erythropus in the Pyrenees. Hybrid zones and the evolutionary process. In R. G. Harrison (Ed.). Oxford University Press.

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

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *359*, 183–195.

Javidkar, M., Abdoli, A., Ahmadzadeh, F., Nahavandi, Z., & Yari, M. (2020). 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*(1), 44–57.

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.

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*(21), 4684–4698.

Knott, K. E., & Haimi, J. (2010). High mitochondrial DNA sequence diversity in the parthenogenetic earthworm *Dendrobaena octaedra*. *Heredity*, *105*(4), 341–347.

Leigh, J. W., & Bryant, D. (2015). Popart: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110–1116.

Mathieu, J., & Davies, T. (2014). Glaciation as an historical filter of belowground biodiversity. *Journal of Biogeography*, *41*(6), 1204–1214.

Michaelsen, W. (1900). Oligochaeta. In *Das Tierreich*. Friedländer & Sohn.

Michaelsen, W. (1910). Zur Kenntnis der Lumbriciden und ihrer Verbreitung. *Annuarie du Musée Zoologique de l'Academie des Sciences de St. Pétersbourg*, *15*(3), 1–74.

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

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES science gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop (GCE)*, *1*, 1–8.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, *37*, 1530–1534.

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 Information Processing Systems*, *21*, 1113–1120.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), 853–858.

O'Malley, M. A. (2008). 'Everything is everywhere: But the environment selects': Ubiquitous distribution and ecological determinism in microbial biogeography. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *39*(3), 314–325.

Omodeo, P. (1952). Cariologia dei Lumbricidae. *Caryologia*, *4*(2), 173–275.

Omodeo, P. (1956). Contributo alla revisione dei Lumbricidae. *Archivio zoologico italiano*, *41*(24), 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*(2), 171–181.

Phillips, H. R. P., Guerra, C. A., Bartz, M. L. C., Briones, M. J. I., Brown, G., Crowther, T. W., Ferlian, O., Gongalsky, K. B., van den Hoogen, J., Krebs, J., Orgiazzi, A., Routh, D., Schwarz, B., Bach, E. M., Bennett, J., Brose, U., Decaens, T., König-Ries, B., Loreau, M., & Mathieu, J. (2020). Erratum for the report "Global distribution of earthworm diversity". *Science*, *369*, 6503.

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*(1), 205–214.

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*(5–6), 428–433.

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

Porco, D., Chang, C.-H., Dupont, L., James, S. W., Richard, B., & Decaëns, T. (2018). A reference library of DNA barcodes for the earthworms from upper Normandy: Biodiversity assessment, new records, potential cases of cryptic diversity and ongoing speciation. *Applied Soil Ecology*, *124*, 362–371.

Porco, D., Decaëns, T., Deharveng, L., James, S. W., Skarżyński, D., Erséus, C., Burt, K. R., Richard, B., & 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*(4), 899–910.

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*, 893–901.

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

Rougerie, R., Decaëns, T., Deharveng, L., Chang, C.-H., James, S. W., Porco, D., Richard, B., & Hebert, P. D. N. (2009). DNA barcodes for soil animal taxonomy. *Pesquisa Agropecuaria Brasileira*, *44*(8), 789–801.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, 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*(12), 3299–3302.

Rutgers, M., Orgiazzi, A., Gardi, C., Römbke, J., Jänsch, S., Keith, A. M., Neilson, R., Boag, B., Schmidt, O., Murchie, A. K., Blackshaw, R. P., Pérès, G., Cluzeau, D., Guernion, M., Briones, M. J. I., Rodeiro, J., Piñeiro, R., Díaz Cosín, D. J., Sousa, J. P., & De Zwart, D. (2016). Mapping earthworm communities in Europe. *Applied Soil Ecology*, *97*, 98–111.

Rutgers, M., Schouten, A. J., Bloem, J., Van Eekeren, N., De Goede, R. G. M., Jagersop Akkerhuis, G. A. J. M., Van der Wal, A., Mulder, C., Brussaard, L., & Breure, A. M. (2009). Biological measurements in a nationwide soil monitoring network. *European Journal of Soil Science*, *60*(5), 820–832.

Simpson, G. G. (1964). Species density of north American recent mammals. *Systematic Zoology*, *13*(2), 57–73.

Stamatakis, A. (2006). RAxML-V1-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, *22*(21), 2688–2690.

Taheri, S., James, S., Roy, V., Decaëns, T., Williams, B. W., Anderson, F., Rougerie, R., Chang, C. H., Brown, G., Cunha, L., Stanton, D. W. G., Da Silva, E., Chen, J. H., Lemmon, A. R., Moriarty Lemmon, E., Bartz, M., Baretta, D., Barois, I., Lapied, E., … Dupont, L. (2018). Complex taxonomy of the 'brush tail' peregrine earthworm *Pontoscolex corethrurus*. *Molecular Phylogenetics and Evolution*, *124*, 60–70.

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.

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., & Saura, A. (2006). Dispersal and clonal diversity of north-European parthenogenetic earthworms. In *Biological invasions belowground: Earthworms as invasive species* (pp. 5–18). Springer.

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 (Jena)*, *38*(1), 81–96.

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

BIOSKETCHES

Irene de Sosa is a researcher at Complutense University of Madrid (Spain). She is interested in phylogeography, demographic history, biology, and the differential expression of genes in semiaquatic earthworms.

Daniel F. Marchán is a researcher at Centre d'Ecologie Fonctionelle et Evolutive (France). He is interested in the evolution, ecology, and biogeography of earthworms, focusing on the use of integrative systematics.

Marta Novo is a researcher at Complutense University of Madrid (Spain). She is interested in the evolutionary biology of earthworms, particularly their phylogeny, phylogeography, and adaptation to environmental stressors.

Ana Almodóvar is a full professor at the Complutense University of Madrid (Spain). Her interests include evolutionary biology and phylogeography of freshwater and terrestrial animals in the context of global change.

Darío J. Díaz Cosín is an Emeritus professor at the Complutense University of Madrid (Spain). His interests include soil biology,

specifically the biology, systematics, phylogeny, and phylogeography of earthworms.

Author contributions: IS, DFM, MN, AA and DJDC, conceived the project. IS carried out the field and laboratory work; analysed the data; and wrote the manuscript with assistance from DFM, MN, AA and DJDC.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: de Sosa, I., Marchán, D. F., Novo, M., Almodóvar, A., & Díaz Cosín, D. J. (2023). Phylogeography of a riparian earthworm shows environmental factors influence genetic structure. *Journal of Biogeography*, *50*, 156–168. https://doi.org/10.1111/jbi.14518