MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 3776–3793 doi: 10.1111/j.1365-294X.2012.05648.x

Appearances can be deceptive: different diversification patterns within a group of Mediterranean earthworms (Oligochaeta, Hormogastridae)

MARTA NOVO,*† ANA ALMODÓVAR,† ROSA FERNÁNDEZ,† DOLORES TRIGO,† DARÍO J. DÍAZ-COSÍN† and GONZALO GIRIBET*

**Department of Organismic and Evolutionary Biology, Museum of Comparative Zoology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA, †Departamento de Zoología y Antropología Física, Facultad de Biología, Universidad Complutense de Madrid, C/José Antonio Nováis, 28040, Madrid, Spain*

Abstract

Many recent studies on invertebrates have shown how morphology not always captures the true diversity of taxa, with cryptic speciation often being discussed in this context. Here, we show how diversification patterns can be very different in two clades of closely related earthworms in the genus *Hormogaster* stressing the risk of using nonspecific substitution rate values across taxa. On the one hand, the *Hormogaster elisae* species complex, endemic to the central Iberian Peninsula, shows morphological stasis. On the other hand, a clade of *Hormogaster* from the NE Iberian Peninsula shows an enormous morphological variability, with 15 described morphospecies. The *H. elisae* complex, however, evolves faster genetically, and this could be explained by the harsher environmental conditions to which it is confined—as detected in this study, that is, sandier and slightly poorer soils with lower pH values than those of the other species in the family. These extreme conditions could be at the same time limiting morphological evolution and thus be responsible for the observed morphological stasis in this clade. Contrarily, *Hormogaster* species from the NE Iberian Peninsula, although still inhabiting harsher milieu than other earthworm groups, have had the opportunity to evolve into a greater morphological disparity. An attempt to delimit species within this group following the recently proposed general mixed Yule-coalescent method showed a higher number of entities than expected under the morphospecies concept, most probably due to the low vagility of these animals, which considerably limits gene flow between distant conspecific populations, but also because of the decoupling between morphological and genetic evolution in the *H. elisae* complex.

Keywords: Annelida, environmental analyses, evolutionary rates, general mixed Yule-coalescent species delimitation

Received 17 October 2011; revision received 1 April 2012; accepted 4 April 2012

Introduction

Recent use of molecular techniques in taxonomy has often led to the detection of unbalanced molecular and morphological evolution, with the study of cryptic speciation becoming an important topic in evolutionary biology (e.g. McGuigan & Sgrò 2009). The discovery of species difficult to distinguish by morphological traits is not uncommon, especially within invertebrate groups (e.g. Müller 2000; Pfenninger *et al.* 2003; Stoks *et al.* 2005; Hogg *et al.* 2006; Tully *et al.* 2006; Challis *et al.* 2007; Finston *et al.* 2007) and specifically in annelid worms such as leeches (Bely & Weisblat 2006), freshwater oligochaetes (Gustafsson *et al.* 2009) and earthworms (e.g. Chang *et al.* 2008; King *et al.* 2008; James *et al.* 2010). Mayr (1948) was the first to discuss the difficulties in distinguishing certain species based solely on morphological

Correspondence: Marta Novo, Fax: +1 617 496 5854;
E-mail: mnrodrig@fas.harvard.edu

characters and he used this to argue against the morphological species concept (Mayr 1963). This is particularly true in earthworms, whose taxonomy suffers from their structural simplicity (Pop *et al.* 2003).

Earthworms are key organisms for the correct functioning of soil systems. They captured Darwin's attention who stated that 'It may be doubted whether there are many other animals which have played so important a part in the history of the world, as have these lowly organized creatures' (Darwin 1881: 316). Earthworms have also been targeted in applied research for a long time (e.g. Lavelle & Spain 2001; Edwards 2004). More recently, molecular data on earthworms have flourished, including DNA barcoding for understanding taxonomy of these soil organisms (Rougerie *et al.* 2009), but their evolutionary biology is still poorly known. Several recent studies have shown large genetic diversity, suggesting some interesting cases of cryptic speciation, while others have questioned morphology-based taxonomy (e.g. Briones *et al.* 2009; Fernández *et al.* 2011). Discordant patterns of morphological and molecular evolution can be common in the soil environment, where chemical signalling may play a key role in sexual selection (Lee & Frost 2002), perhaps more important than morphology itself. In addition, it has been proposed that extreme subsurface conditions could constrain morphological evolution (e.g. Jones *et al.* 1992; Caccone & Sbordoni 2001; Wiens *et al.* 2003).

Hormogastrid earthworms, endogeic and endemic to the Mediterranean region (Cobolli-Sbordoni *et al.* 1992), have shown to be challenging from a genetic perspective. Specifically, a very high genetic diversity was found in the central Iberian Peninsula for *Hormogaster elisae* (Novo *et al.* 2009, 2010a), the only morphospecies described in this area, but now thought to constitute a species complex. This contrasts with the much larger species number of the genus in other areas of the Mediterranean. The phylogenetic placement of this group within the family indicates that *H. elisae* constitutes an independent lineage of the remaining species of *Hormogaster*, a genus shown to be paraphyletic (Novo *et al.* 2011). The family Hormogastridae currently comprises four genera and 22 species of large to middle-sized earthworms, most of which inhabit the NE Iberian Peninsula (15 *Hormogaster* species and some varieties or subspecies), whose phylogeny was studied in Novo *et al.* (2011).

In this work, we markedly increase the sampling for hormogastrids since Novo *et al.* (2011) with the aim to address an important ecological question—whether the diversity of soil habitats is related to genetic and morphological variability—and its evolutionary consequences. We therefore wanted to test the hypothesis that morphological change in earthworms is limited by their environment (i.e. pH values, poor and sandy soils). This constraint of morphological change is exemplified by the *H. elisae* lineage. In addition, we wanted to compare the substitution rates of the different clades defined in the family Hormogastridae because their morphological diversity and probably the characteristics of the soil they inhabit vary. Finally, we investigated the delimitation of hormogastrid species using a general mixed Yule-coalescent (GMYC) model approach to test whether the limited information provided by morphology could be complemented by this method.

Materials and methods

Sampling and morphological study

We collected 376 mature individuals representing 20 of the 22 described species of Hormogastridae in 46 localities (Table 1, see a similar map in Novo *et al.* 2011) from the Iberian Peninsula, France (including Corsica) and Italy (Sardinia). Additional data from the study of Novo *et al.* (2010a) in central Spain were included for cytochrome *ca.* oxidase subunit I gene (COI) and 16S-tRNA genes and soil analyses.

All individuals were collected by hand, washed in distilled water and preserved in ca. 96% EtOH at −20 °C for subsequent molecular work. A portion of the integument (*ca.* 25 mg) was cleansed under a stereomicroscope to remove soil particles. Subsequently, integument samples were hydrated and preserved at −80 °C until DNA extraction. All specimens were dissected and examined morphologically for their taxonomic identification following Qiu & Bouché (1998).

DNA extraction, gene amplification and sequencing

Total genomic DNA was extracted from the integument tissue sample using the DNeasy Tissue Kit (QIAGEN) eluting twice with 70 μL of buffer. Molecular markers included mitochondrial regions of the COI, 16S rRNA gene and tRNA Leu, Ala, and Ser (16S-tRNA) and two nuclear protein-encoding genes (histone H3 and histone H4). Primer sequences, polymerase chain reactions (PCR) and sequencing reactions are the same as in Novo *et al.* (2011).

Chromatograms were visualized in Sequencher v.4.7 (Gene Codes Corporation, Ann Arbor, MI, USA) to assemble sequences. All amplicons were compared against the GenBank database with the BLAST algorithm (Altschul *et al.* 1997) for potential contaminants.

Genetic diversity

Genetic variability estimates and mean genetic differentiation between and within studied populations were

Table 1 Hormogastrid species (alphabetically ordered within the clades recovered by Novo *et al.* 2011), localities of collection and GenBank Accession numbers

Phylogenetic Clade	Species	Locality	Code	Region, Country	COI: GenBank Numbers	16S: GenBank Numbers	H3: GenBank Numbers	H4: GenBank Numbers	GPS Coordinates	Altitude (m)
HCL	*Hormogaster najaformis*	Ordal	ORD	Barcelona, Spain	HQ621985, **JN209571–JN209575**	HQ621878, **JN209316–JN209321**	HQ622023, **JN209695–JN209697**	HQ622069, **JN209768–JN209770**	N41°23′27.3″ E001°49′39.3″	391
	*Hormogaster pretiosa c.f. hispanica**	Río Ter	HPA	Girona, Spain	—	HQ621892	HQ622044	—	—	—
	Hormogaster pretiosa pretiosa	Villamassargia	VIL	Carbonia-Iglesias, Sardinia, Italy	HQ621998, **JN209576–JN209578**	HQ621893, **JN209439–JN209445**	HQ622045, **JN209698–JN209700**	HQ622090, **JN209771–JN209777**	N39°15′29.7″ E8°40′17.3″	290
HE	*Hormogaster elisae* Sp1	Cabrera	CAB	Madrid, Spain	GQ409689.1	GQ409729.1, **JN209221–JN209226**	HQ622007, **JN209648–JN209649**, **JN209659**	HQ622053, **JN209747**, **JN209752**	N40°51′25.9″ W03°37′18.2″	1029
	H. elisae Sp1	Cubillo de Uceda	UCE	Guadalajara, Spain	GQ409692.1–GQ409697.1	GQ409720.1–GQ409722.1, **JN209407–JN209412**	HQ622039, **JN209660–JN209662**	HQ622085, **JN209739–JN209740**	N40°49′38.7″ W03°25′19.5″	883
	H. elisae Sp1	Fresno del Torote	FRE	Madrid, Spain	GQ409698.1–GQ409699.1	GQ409723.1–GQ409724.1	HQ622009, **JN209637**	HQ622055, **JN209723**	N40°35′51.8″ W03°24′42.0″	660
	H. elisae Sp1	Lozoyuela	LOZ	Madrid, Spain	EF653886.1–EF653890.1, GQ409690.1, GQ409691.1	GQ409725.1–GQ409728.1, **JN209271–JN209285**	HQ622016, **JN209650–JN209652**, **JN209656–JN209658**	HQ622062, **JN209727–JN209730**, **JN209748**	N40°56′51.9″ W03°37′16.2″	1036
	H. elisae Sp1	Molar	MOL	Madrid, Spain	EF653874.1–EF653880.1	GQ409732.1–GQ409736.1, **JN209294–JN209295**	HQ622019, **JN209626–JN209627**, **JN209629–JN209633**, **JN209636**	HQ622065, **JN209731–JN209732**, **JN209741–JN209744**, **JN209746**	N40°44′22.9″ W03°33′53.1″	753
	H. elisae Sp1	Navas de Buitrago	NAV	Madrid, Spain	GQ409683.1–GQ409688.1	GQ409730.1–GQ409731.1, **JN209303–JN209308**	HQ622021, **JN209653–JN209655**	HQ622067, **JN209733–JN209735**	N40°56′21.0″ W03°35′38.1″	994
	H. elisae Sp1	Paracuellos del Jarama	JAR	Madrid, Spain	GQ409665.1–GQ409670.1	GQ409745.1–GQ409749.1, **JN209252–JN209256**	HQ622013, **JN209625**, **JN209628**, **JN209640**	HQ622059, **JN209724–JN209726**	N40°30′36.9″ W03°31′59.1″	674

Table 1 *Continued*

Phylogenetic Clade	Species	Locality	Code	Region, Country	COI: GenBank Numbers	16S: GenBank Numbers	H3: GenBank Numbers	H4: GenBank Numbers	GPS Coordinates	Altitude (m)
	H. elisae Sp1	Redueña	RED	Madrid, Spain	EF653881.1–EF653885.1, GQ409673.1, GQ409675.1–GQ409677.1	GQ409741.1–GQ409744.1, GQ409752.1, **JN209338–JN209357**	HQ622029, **JN209624**, **JN209634**	HQ622075, **JN209736**	N40°48′46.7″ W03°36′06.2″	797
	H. elisae Sp1	Tres Cantos	TRE	Madrid, Spain	GQ409678.1–GQ409682.1	GQ409737.1–GQ409740.1, **JN209401–JN209406**	HQ622038, **JN209622**, **JN209639**, **JN209641**	HQ622084, **JN209737–JN209738**, **JN209745**	N40°36′46.9″ W03°40′41.1″	675
	H. elisae Sp1	Venturada	VEN	Madrid, Spain	GQ409671.1, GQ409672.1, GQ409674.1	GQ409750.1–GQ409751.1, GQ409753.1, **JN209420–JN209425**	HQ622041, **JN209623**, **JN209635**, **JN209638**	HQ622087, **JN209749–JN209751**	N40°48′07.7″ W03°37′19.6″	890
	H. elisae Sp2	Siguero	SIG	Segovia, Spain	EF653891.1–EF653897.1	GQ409710.1–GQ409715.1, **JN209376–JN209380**	HQ622033, **JN209642–JN209644**	HQ622079, **JN209716–JN209719**	N41°11′06.1″ W03°37′07.4″	1073
	H. elisae Sp2	Soto del Real	SOT	Madrid, Spain	GQ409700.1–GQ409702.1	GQ409716.1–GQ409719.1, **JN209381–JN209386**	HQ622034, **JN209645–JN209647**	HQ622080, **JN209720–JN209722**	N40°46′30.5″ W03°46′42.6″	985
	H. elisae Sp3	Anchuelo	ANC	Madrid, Spain	EF653868.1–EF653873.1	GQ409754.1, **JN209205–JN209207**	HQ622001, **JN209663–JN209665**	HQ622047, **JN209713–JN209715**	N40°28′50.2″ W03°14′33.5″	780
	H. elisae Sp4	Boadilla del Monte	BOA	Madrid, Spain	GQ409661.1–GQ409664.1	GQ409704.1–GQ409706.1, JN209215-JN209220	HQ622004, **JN209666–JN209668**	HQ622050, **JN20975–JN2097553**	N40°25′50.2″ W03°55′30.9″	667
	H. elisae Sp4	Sevilla la Nueva	SEV	Madrid, Spain	EF653903.1–EF653905.1	GQ409707.1–GQ409708.1, **JN209365–JN209370**	HQ622031, **JN209669–JN209671**	HQ622077, **JN209756–JN209758**	N40°20′41.9″ W04°00′48.9″	644
	H. elisae Sp5	Pardo	PAR	Madrid, Spain	EF653898.1–EF653902.1	GQ409709.1, HQ621879, **JN209322–JN209325**	HQ622024, **JN209672–JN209673**	HQ622070, **JN209759–JN209761**	N40°31′11.0″ W03°47′42.7″	662

Table 1 *Continued*

Phylogenetic Clade	Species	Locality	Code	Region, Country	COI: GenBank Numbers	16S: GenBank Numbers	H3: GenBank Numbers	H4: GenBank Numbers	GPS Coordinates	Altitude (m)
HIS	*Hormogaster redii insularis*	Volpajola	VPJ	Haute-Corse, Corsica, France	HQ621996, **JN209565–JN209570**	HQ621890, **JN209426–JN209431**	HQ622042, **JN209683–JN209685**	HQ622088, **JN209813–JN209815**	N42°31′18.6″ E009°21′05.7″	401
	Hormogaster redii redii	Alghero	ALG	Sassari, Sardinia, Italy	HQ621971	HQ621863	HQ622000	—	N40°35′59.9″ E008°17′56.7″	0
	Hormogaster redii redii	Ghilarza	GHI	Oristano, Sardinia, Italy	HQ621976, **JN209453–JN209457**	HQ621869, **JN209234–JN209237**	HQ622010, **JN209674–JN209676**	HQ622056, **JN209804–JN209806**	N40°06′58.9″ E008°51′53.3″	247
	Hormogaster redii redii	Iglesias	IGL	Carbonia-Iglesias, Sardinia, Italy	HQ621978, **JN209458–JN209464**	HQ621871, **JN209245–JN209251**	HQ622012, **JN209677–JN209679**	HQ622058, **JN209807–JN209809**	N39°19′23.8″ E008°31′17.7″	276
	Hormogaster samnitica lirapora	Col de la Testa	TES	Corse-du-Sud, Corsica, France	HQ621993, **JN209465–JN209469**	HQ621887, **JN209389–JN209393**	HQ622036, **JN209680–JN209682**	HQ622082, **JN209810–JN209812**	N41°28′24.4″ E009°06′08.6″	54
HM	*Hemigastrodrilus monicae*	Cardayre	LAC	Lot-et-Garonne, France	HQ621979, **JN209579–JN209585**	HQ621872, **JN209257–JN209263**	HQ622014, **JN209689–JN209691**	HQ622060, **JN209823–JN209825**	N44°18′46.8″ E000°52′45.0″	210
	Hemigastrodrilus monicae	Mende	MND	Lozére, France	HQ621982, **JN209446–JN209452**	HQ621875, **JN209287–JN209293**	HQ622018, **JN209692–JN209694**	HQ622064, **JN209826**, **JN209828**	N44°29′43.6″ E003°27′41.9″	710
HNE	*Hormogaster arenicola*	Biosca	BIO	Lleida, Spain	HQ621972, **JN209493–JN209499**	HQ621865, **JN209208–JN209214**	HQ622003, **JN209596–JN209598**	HQ622049, **JN209820–JN209822**	N41°51′ 04.6″ E001°19′40.4″	454
	Hormogaster catalaunensis	El Brull	BRU	Barcelona, Spain	HQ621973	HQ621866	HQ622005	HQ622051	N41°48′04.9″ E002°20′51.6″	1145
	Hormogaster castillana	Puerto Querol	QUE	Castellón, Spain	HQ621989, **JN209507–JN209513**	HQ621883, **JN209331–JN209337**	HQ622028, **JN209602–JN209604**	HQ622074, **JN209793**, **JN209795–JN209796**	N40°33′49.1″ W000°01′03.5″	957
	Hormogaster eserana	Graus	GRA	Huesca, Spain	HQ621977, **JN209530–JN209536**	HQ621870, **JN209238–JN209244**	HQ622011, **JN209608–JN209610**	HQ622057, **JN209785–JN209787**	N42°10′11.8″ E000°19′53.2″	454
	Hormogaster gallica	Banyuls Sur Mer	BSM	Pyrénées-Orientales, France	HQ621974	HQ621867	HQ622006	HQ622052	N42°28′08.0″ E003°09′08.2″	105
	Hormogaster huescana	Loporzano	LOP	Huesca, Spain	HQ621980, **JN209523–JN209529**	HQ621873, **JN209264–JN209270**	HQ622015, **JN209614**, **JN209621**	HQ622061, **JN209788–JN209789**, **JN209829**	N42°07′04.6″ W000°14′54.6″	558
	Hormogaster ireguana	Torrecilla en Cameros	TOR	La Rioja, Spain	HQ621994, **JN209486–JN209492**	HQ621888, **JN209394–JN209400**	HQ622037, **JN209593–JN209595**	HQ622083, **JN209765–JN209767**	N42°13′54.7″ W002°37′35.2″	789

Table 1 *Continued*

Phylogenetic Clade	Species	Locality	Code	Region, Country	COI: GenBank Numbers	16S: GenBank Numbers	H3: GenBank Numbers	H4: GenBank Numbers	GPS Coordinates	Altitude (m)
	Hormogaster oroeli	Peña Oroel	OEL	Huesca, Spain	HQ621984, **JN209545–JN209551**	HQ621877, **JN209309–JN209315**	HQ622022, **JN209618–JN209620**	HQ622068, **JN209798–JN209800**	N42°31′20.1″ W000º2́09.1́	1090
	*Hormogaster pretiosa arrufati**	Vall d'Uixó	UIX	Castelló, Spain	HQ621995, **JN209514–JN209520**	HQ621889, **JN209413–JN209419**	HQ622040, **JN209605–JN209607**	HQ622086, **JN209778**, **JN209802**, **JN209803**	N39°49′39.4″ W000°15′40.2″	168
	*Hormogaster pretiosa nigra**	Quillan	QLL	Aude, France	HQ621988	HQ621882	HQ622027	HQ622073	N42°52′48.8″ E002°10′12.0″	396
	*Hormogaster pretiosa var.**	Peralba	PRB	Lleida, Spain	HQ621987, **JN209537**	HQ621881, **JN209330**	HQ622026, **JN209613**	HQ622072, **JN209792**	N41°59′22.7″ E000°54′50.1″	740
	Hormogaster pretiosiformis	Monrepós	MON	Huesca, Spain	HQ621983, **JN209538–JN209544**	HQ621876, **JN209296–JN209302**	HQ622020, **JN209615–JN209617**	HQ622066, **JN209791**, **JN209797**, **JN209801**	N42°23′26.0″ W000°22′20.9″	860
	Hormogaster riojana	Alesanco	ALE	La Rioja, Spain	HQ621970, **JN209477–JN209485**	HQ621862, **JN209196–JN209204**	HQ621999, **JN209590–JN209592**	HQ622046, **JN209762–JN209764**	N42°26′21.7″ W002°50′18.4″	596
	Hormogaster sp.	Talarn	TAL	Lleida, Spain	HQ621992, **JN209521–JN209522**	HQ621886, **JN209387–JN209388**	HQ622035, **JN209611–JN209612**	HQ622081, **JN209779–JN209780**	N42°11′05.5″ E000°54′11.7″	464
	Hormogaster sp.	Cervera del Maestre	CER	Castelló, Spain	HQ621975, **JN209500–JN209506**	HQ621868, **JN209227–JN209233**	HQ622008, **JN209599–JN209601**	HQ622054, **JN209783–JN209784**, **JN209827**	N40°27′23.1″ E000°16′59.0″	214
	Hormogaster sylvestris	Montmajor	MAJ	Barcelona, Spain	HQ621981, **JN209552**	HQ621874, **JN209286**	HQ622017, **JN209706**	HQ622063, **JN209790**	N42°01′43.3″ E001°42′43.7″	785
	New species (author's work in progress)	Sant Joan de les Abadesses	SAN	Girona, Spain	HQ621990, **JN209553–JN209559**	HQ621884, **JN209358–JN209364**	HQ622030, **JN209707–JN209709**	HQ622076, **JN209781–JN209782**, **JN209794**	N42°13′30.0″ E002°14′57.5″	735
VG	*Vignysa popi*	Saint- Gely-du-Fesc	SGF	Hérault, France	HQ621991, **JN209560–JN209564**	HQ621885, **JN209371–JN209375**	HQ622032, **JN209704–JN209705**	HQ622078, **JN209816–JN209817**	N43°42′19.0″ E003°48′03.7″	152
	Vignysa vedovinii	Pignans	PIG	Var, France	HQ621986, **JN209586–JN209589**	HQ621880, **JN209326–JN209329**	HQ622025, **JN209701–JN209703**	HQ622071, **JN209710–JN209712**	N43°18′04.6″ E006°12′35.9″	204

Table 1 *Continued*

Phylogenetic Clade	Species	Locality	Code	Region, Country	COI: GenBank Numbers	16S: GenBank Numbers	H3: GenBank Numbers	H4: GenBank Numbers	GPS Coordinates	Altitude (m)
XN	*Xana omodeoi*	San Román	XAN	Asturias, Spain	HQ621997, **JN209470–JN209476**	HQ621891, **JN209432–JN209438**	HQ622043, **JN209686–JN209688**	HQ622089, **JN209818–JN209819**	N43°15′20.9″ W005°05′10.3″	644

HCL, *Hormogaster* clade used for calibration; HE, *Hormogaster elisae* species complex; HIS, *Hormogaster* from the islands (Corsica and Sardinia); HM, *Hemigastrodrilus*; HNE, *Hormogaster* from the NE Iberian Peninsula; VG, *Vignysa*; XN, *Xana*; COI, cytochrome *c* oxidase subunit I gene.
Sequences in bold are additions for this study. The same species names as in Novo *et al.* (2011) are used although as indicated in that study, a systematic revision of the family is overdue and some forms named *Hormogaster pretiosa* (with asterisk) may require a new name (M.N. work in progress). For a specimen donated by Pietro Omodeo, no exact coordinates are available.

calculated using Arlequin v.3.5 (Excoffier & Lischer 2010) and DnaSP v. 5 (Librado & Rozas 2009) using a Kimura 2-parameter correction, as in previous studies.

Haplotype networks were constructed for each gene region. The statistical parsimony procedure (Templeton *et al.* 1992; Crandall *et al.* 1994) with a 95% connection threshold was used in TCS version 1.2.1 (Clement *et al.* 2000) for histones. Mitochondrial genes showed large divergence, and thus, these sequences were analysed using SplitsTree4 v.4.10 (Huson & Bryant 2006). Default settings were used, thus constructing a Neighbour-net with uncorrected *p*-distances.

Species boundaries based on GMYC approach and molecular rates

Species boundaries were evaluated using a general mixed Yule-coalescent (GMYC) model approach (Pons *et al.* 2006; Fontaneto *et al.* 2007) based on a separate analysis of the two mitochondrial genes. A Bayesian analysis was performed in BEAST 1.6.1 for COI and 16S-tRNA (Drummond & Rambaut 2007) to reconstruct a fully resolved topology with branch length estimates. A relaxed lognormal clock was used with the Coalescence prior. This prior has proven more conservative, thus suggesting more accurate species boundaries, because the GMYC uses a coalescent as the null model for explaining branching patterns (Monaghan *et al.* 2009). GTR+I+G, the best-fit model as indicated by jModeltest (Posada 2008), was used.

The phylogenetic trees were calibrated using the separation between *Hormogaster pretiosa pretiosa* from Sardinia (VIL) and the continental species clustering in the same clade (named HCL)—*Hormogaster pretiosa* cf. *hispanica* collected near the shore of the river Ter in Girona (HPA) and *Hormogaster najaformis* (ORD)—assuming a divergence time for these lineages at least dating back to the separation of the Corso-Sardinian microplate from continental Europe. As a calibration point, we use the most current estimates for the separation of the occidental Mediterranean microplates (33 Ma; Schettino & Turco 2006).

Bayesian analyses were run from for 50 million generations, sampling from every 5000th generation, always resulting in 10 000 trees. After checking for stationarity with Tracer v. 1.5 (Rambaut & Drummond 2007), 20% of the trees were discarded. A maximum clade credibility tree was then built with the remaining trees and analysed in the R package SPLITS (http://r-forge.r-project.org/projects/splits), following the GMYC approach for species delimitation with single and multiple threshold optimizations.

Molecular rates (in terms of substitution rates) were calculated for the four genes combined to compare the

evolution of the different clades. A similar approach as for the GMYC was followed by constructing a fully resolved topology with branch length estimates in BEAST 1.6.1. In this case, two priors were compared (Coalescence with constant size and Speciation-Yule) with the aim to evaluate the credibility of the rates yielded by the different analysis. We also calculated relative timings by calibrating the tree with a 1 in the root for Hormogastridae. Differences between the substitution rates (recovered by BEAST 1.6.1 for each node) of distinct clades within Hormogastridae (namely HCL, HE, HIS, HM, HNE, VG, XN, see Table 1) were analysed by means of a one-way ANOVA in STATISTICA v. 6.1 (StatSoft, Inc. (2001), http://www.statsoft.com).

Geographical and environmental assessment

A pattern of isolation-by-distance was tested for the big clades HE and HNE by means of a Mantel (1967) test, correlating the matrix of genetic distance between localities (Φ_{ST} with Kimura 2-parameter correction) based on the most variable gene, COI, and geographical distance (here, straight line between study sites). The test was performed in Arlequin v. 3.5 (Excoffier & Lischer 2010), and the significance of matrices correlation was evaluated comparing the Mantel test statistic *Z*, for which random distributions were obtained with 10 000 permutations.

Some soil factors were measured in the sampled locations to explore the role of the environment in the genetic diversification of this group of Mediterranean earthworms. These were selected because they are among the most important factors affecting earthworm distribution (Edwards & Bohlen 1996; Hernández *et al.* 2003). Soil texture and pH were determined as described by Guitián & Carballas (1976). Total nitrogen content was determined by the Kjeldahl method as indicated in Page *et al.* (1982) and expressed as a percentage. Organic oxidable carbon analysis was based on Anne's (1945) method, adapted for a microplate reader (Microplate Bio-Rad, 590 nm), employed using glucose as a standard and expressed as a percentage.

A Mantel test with 10 000 permutations was implemented in Arlequin v.3.5 (Excoffier & Lischer 2010) to correlate the genetic distances (Φ_{ST} with a Kimura 2-parameter correction), based on all the genes separately, with the differences between localities for soil factors (in absolute values). This was performed to indentify the soil factors that could be related to the earthworm genetic diversity (i.e. do the most genetically distinct populations live in the most unique soils? what soil characteristics could be involved in the evolution of the group?)

As the Mantel test including the genetic sequence information was not conclusive, a principal component analysis (PCA) was performed only with the soil variables and sampling localities to reduce these variables to Factors. This analysis was executed via a correlation matrix after standardization (43 localities × 8 soil variables), which permits a visualization of the variations among localities. Localities belonging to different phylogenetic clades (previously recovered by Novo *et al.* 2011; i.e. HCL, HE, HIS, HM, HNE, VG, XN, see Table 1) were represented with different colours in the PCA graph. Afterwards, the mean values of factor scores and separate soil variables for these phylogenetic clades were compared by an ANOVA (STATISTICA v. 6.1, StatSoft, Inc. 2001, http://www.statsoft.com). This is aimed to help ascertain whether the type of soil has an influence on the composition of the earthworms inhabiting each locality.

Results

Characteristics of the used genes and genetic variability

Genetic variability values for each gene region are shown in Table 2 for the whole data set of Hormogastridae. Detailed diversity descriptions by locality are shown in Table S1 (Supporting information). Also the mean genetic divergence between localities (Kimura 2-parameter corrected) can be found in Table S2 (mitochondrial data) and Table S3 (nuclear data) (Supporting information). Mitochondrial regions showed to be more variable than the nuclear genes, with a mean divergence in the range of 13.6–16.4%. Histones were the least variable markers, showing a mean divergence of *ca.* 3%. Indel events were only detected in the 16S-tRNA fragment.

Haplotypic networks are shown in Fig. 1. A neighbour-net constructed with SplitsTree and based on mitochondrial data shows how the HE clade presents much deeper subdivisions than the HNE clade, despite the former including a single morphospecies and the latter many. Regarding the histone networks, both histones H3 and H4 presented unconnected networks for each of the main clades. The histone H4 for those networks with more than two terminals is presented as an example (Fig. 1), showing that the patterns of diversification for HE and HNE are similar for histones.

Species boundaries based on GMYC approach and molecular rates

General mixed Yule-coalescent analyses exhibited a higher subdivision for Hormogastridae than expected under a morphospecies concept. Results from both mitochondrial genes are similar, with 56 entities identified. The number of entities was artificially high in the HE clade, separating individuals collected in the same locale as different entities (24 entities in 16 populations).

Table 2 Genetic diversity values of the gene fragments used for the study

	COI	16S-tRNA	H3	H4
N	375	374	166 (332)	164 (328)
NHAP	170	162	65	37
H	0.99 (0.0001)	0.99 (0.001)	0.98 (0.002)	0.92 (0.01)
π	0.164 (0.078)	0.136 (0.065)	0.035 (0.018)	0.028 (0.015)
Positions	648	808	328	183
Substitutions	514	572	95	47
Indels	0	78	0	0
S	266	380	68	32
Ts	366	358	59	31
Tv	148	214	36	16
Dp (N)	106.35 (45.77)	110.20 (47.41)	11.47 (5.21)	5.12 (2.49)
Dp (%)	16.41 (7.06)	13.64 (5.87)	3.50 (1.59)	2.80 (1.36)

N, Number of individuals analysed (in the case of the histones, two alleles per individual were analysed, because of the presence of ambiguities); *NHAP*, number of identified haplotypes; *H*, haplotypic diversity; π, nucleotidic diversity; *S*, number of polymorphic sites; *Ts*, number of transitions; *Tv*, number of transversions, *Dp*, mean number of pairwise differences (N, total number, %, percentage); COI, cytochrome *c* oxidase subunit I gene.
Standard deviation of the estimates is shown in parenteses.

The HM clade also exhibited unexpected subdivisions (four entities in two populations). Regarding the remaining clades, nearly all localities were identified as a separate entity. The multiple threshold approach resulted in an even higher number of entities (126 with 16S-tRNA and 118 with COI).

Molecular rates were suitable for ANOVA analyses after the *Log*10 transformation. Only differences between the rates of clades HE and HNE were statistically significant, as the remaining clades had few values and great variability. Absolute molecular rates (after constructing the trees with the calibration in HCL at 33 Ma) were significantly higher (ANOVA: $F_{1,36} = 6.6988$, $P = 0.01$) in the HE clade (Mean rate = 0.002941, SD = 0.000735) than in the HNE clade (Mean rate = 0.002363, SD = 0.000779) for the Coalescent prior. No significant differences were found for the Speciation-Yule prior (ANOVA: $F_{1,35} = 1.3407$, $P = 0.254$), although mean values for HE (Mean rate = 0.00803, SD = 0.007611) were higher than in HNE (Mean rate = 0.005506, SD = 0.005808). Similar results were obtained when comparing the relative molecular rate values (obtained without calibration, setting a value of 1 in the root of Hormogastridae) showing significant differences with the Coalescent prior (ANOVA Coalescence $F_{1,35} = 13.499$, $P = 0.0008$) but not with the Speciation-Yule prior ($F_{1,33} = 2.9994$, $P = 0.092$). Note that the absolute molecular rate values are difficult to compare between different priors as they are very different.

Geographical and environmental assessment

A pattern of isolation-by-distance was found for the HE clade, as shown by the positive correlation of the Mantel test ($r = 0.372$, $P = 0.003$). No such pattern was detected for the HNE clade ($r = 0.077$, $P = 0.334$).

Soil characteristics for each locality are shown in Table 3. The results of the Mantel test between matrices of genetic distance and absolute differences in soil properties are shown in Table 4. All factors, except for the percentage of fine sand, are significantly correlated with each genetic fragment analysed.

Principal component analysis revealed three factors (Table 5) that explained the 77.52% of the total variance in the soil characteristics between localities. The first factor (36.35% of the total variance) was highly and positively correlated to the level of nutrients (carbon and nitrogen) and clay, and negatively correlated to the percentage of sand. pH and fine silt were positively related to the second factor (25.02% of the total variance). The graphical representation of the localities in the first two axes obtained is presented in Fig. 2, where localities are coloured according to the phylogenetic clade recovered in previous analyses by Novo *et al.* (2011). Localities included in the clade HE exhibit a tendency towards the more acidic soils, with higher content of coarse sand, and slightly poorer, whereas clade HNE is found in soils with finer texture and higher pH. *Xana* inhabits soils with higher content of clay than the ones of other hormogastrids and soil characteristics in Corsica and Sardinia (clade HIS) resulted very heterogeneous. These differences between the main clades (HE and HNE) are confirmed by the ANOVA. Scores of Factor 1 showed significant differences ($F_{6,36} = 4.55$, $P = 0.002$) only among XN and HE/HNE according to post hoc comparisons (Tukey's test). But when excluding XN and HCL (Phylogenetic clades represented only by one locality) the differences among

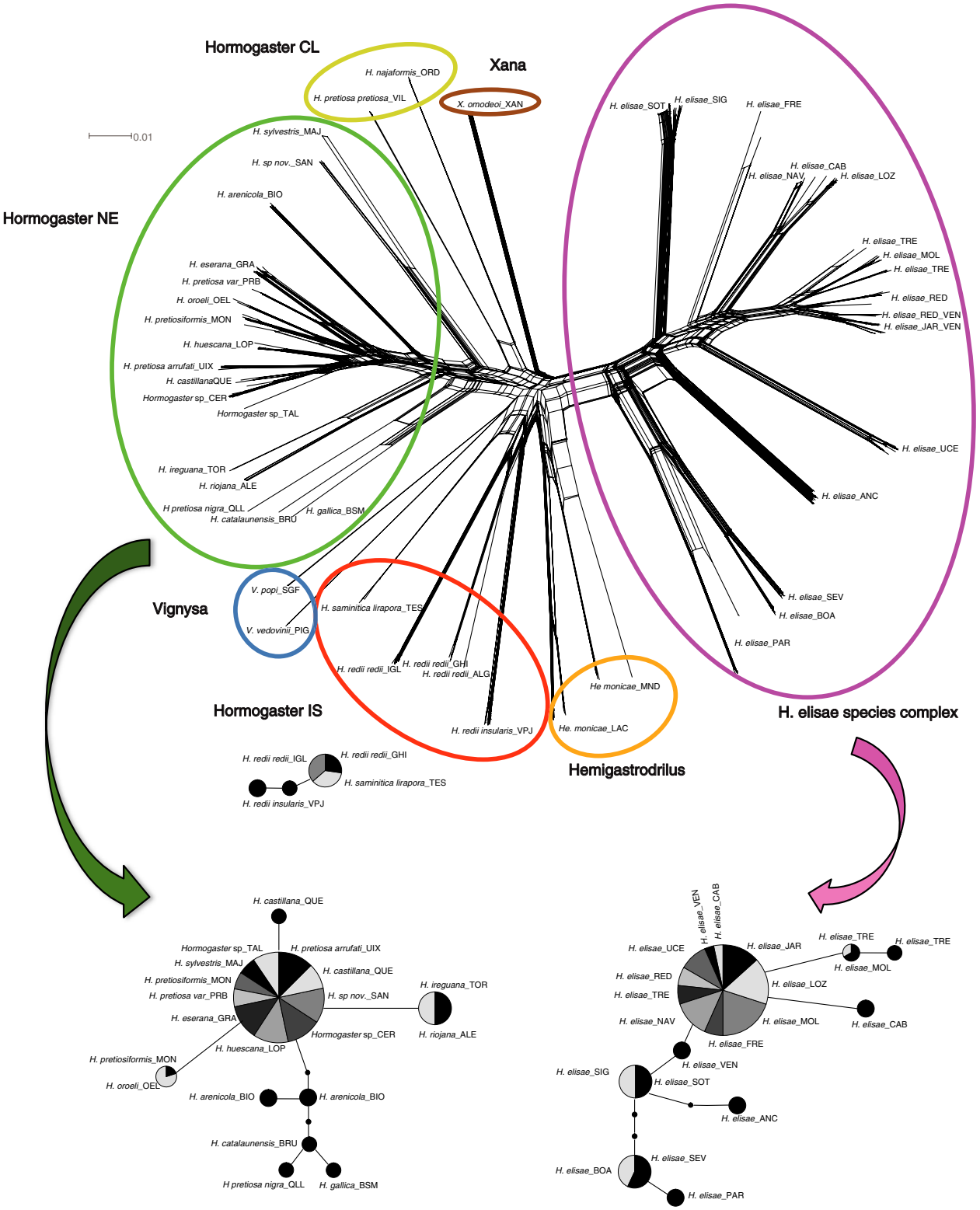

Fig. 1 Illustration of the genetic diversity within the different clades of Hormogastridae. In the upper part of the Figure, a network recovered by SplitsTree4 and based on the mitochondrial data is shown. Branch length is proportional to the genetic distance. Different clades recovered by Novo *et al.* (2011) are indicated. TCS network for some of the clades is shown, based on histone H4 data. The size of the circles reflects haplotype frequency. Circles with no names are intermediate inferred haplotypes. Each branch represents one mutational step, branch lenght being meaningless.

Table 3 Soil properties of the sampled locations of hormogastrids

Clade	Code	%Coarse sand	%Fine Sand	%Total Sand	%Coarse Silt	%Fine Silt	%Total Silt	%Clay	Texture classification	%C	%N	vC/N	pH
HCL	ORD	19.24	9.73	28.97	9.35	32.72	42.07	28.97	Clay loam	3.72	0.76	4.91	7.38
HE	CAB	65.46	9.28	74.75	4.3	11.62	15.92	9.34	Sandy loam	2.15	0.514	4.18	5.68
	UCE	22.94	16.53	39.47	26.2	20.14	46.34	14.19	Loam	2.3	0.476	4.83	5.57
	FRE	45.34	18.24	63.58	12.41	12.28	24.69	11.74	Sandy loam	0.86	0.179	4.79	7.45
	LOZ	50.37	20.21	70.58	10.23	10.76	21	8.43	Sandy loam	2	0.424	4.72	5.49
	MOL	52.47	19.91	72.38	2.92	9.61	12.53	15.1	Sandy loam	0.96	0.16	5.92	6.4
	NAV	51.76	18.22	69.97	10.1	10.5	20.6	9.42	Sandy loam	1.88	0.452	4.16	5.35
	JAR	36.01	11.48	47.49	12.36	13.69	26.06	26.45	Sandy clay loam	1.82	0.343	5.31	7
	RED	38.15	21.36	59.51	8.63	19.83	28.46	12.03	Sandy loam	1.77	0.399	4.45	7.62
	TRE	56.29	13.16	69.45	6.08	10.97	17.05	13.5	Sandy loam	1.86	0.41	4.55	6.24
	VEN	40.69	16.85	57.54	7.95	18.07	26.03	16.44	Sandy loam	2.08	0.543	3.83	7.41
	SIG	44.99	10.9	55.89	10.74	18.13	28.88	15.23	Sandy loam	2.14	0.606	3.54	5.05
	SOT	45.11	25.74	70.84	6.5	13.97	20.46	8.69	Sandy loam	1.87	0.326	5.74	5.43
	ANC	11.69	17.48	29.16	14.58	36.49	51.08	19.76	Clay	1.78	0.391	4.55	7.9
	BOA	59.67	11.93	71.6	2.32	7.61	9.93	18.47	Sandy loam	1.18	0.237	4.98	6.42
	SEV	66.19	8.22	74.41	2.93	6.58	9.52	16.08	Sandy loam	1.28	0.199	6.46	5.93
	PAR	65.51	8.65	74.16	3.16	7.49	10.65	15.19	Sandy loam	1.5	0.263	5.69	6.19
HIS	VPJ	30.61	17.82	48.43	7.57	34.68	42.25	9.31	Loam	2.59	0.29	8.93	5.73
	GHI	11.32	10.15	21.47	12.49	39.85	52.34	26.19	Silt loam	5.81	1.37	4.24	5.6
	IGL	30.54	14.40	44.94	19.20	5.34	24.54	30.52	Clay loam	1.83	0.51	3.56	7.26
	TES	69.76	12.69	82.45	2.23	7.92	10.16	7.40	Loamy Sand	1.21	0.37	3.31	5.78
HM	LAC	23.67	8.26	31.92	5.45	43.64	49.09	18.99	Loam	4.44	0.93	4.79	7.44
	MND	38.47	23.95	62.42	5.95	27.61	33.55	4.03	Sandy loam	2.15	0.52	4.14	7.6
HNE	BIO	7.53	14.10	21.62	4.36	69.01	73.37	5.01	Silt loam	3.86	1.20	3.21	7.19
	BRU	47.50	12.61	60.11	5.25	17.72	22.97	16.92	Sandy loam	4.76	1.32	3.61	4.76
	QUE	34.18	26.25	60.43	4.65	32.83	37.47	2.09	Sandy loam	3.79	0.84	4.49	7.47
	GRA	25.03	20.86	45.89	13.08	32.31	45.39	8.72	Loam	2.20	0.68	3.23	7.49
	BSM	25.21	8.33	33.55	7.11	36.49	43.60	22.85	Loam	3.88	0.94	4.13	5.51
	LOP	33.99	12.37	46.36	4.50	30.10	34.60	19.05	Loam	1.59	0.37	4.36	7.79
	TOR	19.78	25.04	44.82	15.87	17.87	33.74	21.44	Loam	2.32	0.53	4.37	7.45
	OEL	12.90	28.49	41.38	12.32	21.80	34.12	24.50	Loam	2.76	0.53	5.16	7.3
	UIX	33.24	16.93	50.17	4.57	21.41	25.98	23.85	Sandy clay loam	3.19	0.39	8.18	7.56
	QLL	11.04	11.09	22.13	3.46	72.08	75.54	2.33	Silt loam	1.81	0.52	3.52	7.62
	PRB	15.37	15.83	31.20	6.85	59.30	66.15	2.66	Silt loam	3.44	0.71	4.87	7.42
	MON	10.21	15.17	25.38	5.61	66.64	72.25	2.37	Silt loam	1.78	0.35	5.15	7.59
	ALE	9.24	25.12	34.36	55.38	1.86	57.24	8.40	Silt loam	1.63	0.30	5.33	7.33
	TAL	11.28	16.33	27.61	8.81	61.27	70.08	2.31	Silt loam	2.23	1.41	1.58	7.57
	CER	20.41	14.22	34.64	9.24	40.34	49.58	15.78	Loam	2.70	0.54	5.01	7.5
	MAJ	11.71	6.50	18.22	6.88	69.02	75.90	5.88	Silt loam	2.98	0.83	3.60	7.39
	SAN	13.57	9.62	23.18	6.27	32.37	38.64	38.18	Clay loam	4.48	1.32	3.39	7.09
VG	SGF	16.68	23.62	40.30	13.19	28.82	42.00	17.70	Loam	2.34	0.67	3.48	7.55
	PIG	20.21	11.09	31.31	7.78	33.55	41.33	27.36	Loam	4.26	0.85	5.02	7.4
XN	XAN	9.01	5.67	14.68	1.84	22.00	23.84	61.48	Clay	3.34	1.25	2.68	5.26

C, Carbon; N, Nitrogen.

ALG, HPA and VIL values are not included because specimens from those localities were not collected by the authors and soil samples were not available for study. For complete names see Table 1.

HE and HNE were significant ($F_{4,36} = 3.03$, $P = 0.03$). Scores of Factor 2 showed significant differences ($F_{6,36} = 6.11$, $P = 0.0002$) among HE and HNE according to post hoc comparisons. All the individually analysed soil variables showed significant differences among clades ($P < 0.05$) excepting fine sand and coarse silt content. Post hoc comparisons showed that coarse sand content was significantly higher in localities included in the HE clade than in the HNE clade, whereas fine silt content, Carbon, Nitrogen and pH were lower in the HE than in the HNE clade ($P < 0.05$ in all the cases). Locality XN showed a significantly higher content of clay when compared to the rest ($P < 0.01$, except for HCL).

Table 4 Results from the Mantel test of the correlation between Φ_{ST} based on different genes among localities, and the differences (in absolute values) of their soil factors

	COI		16S-tRNA		H3		H4	
	P	*r*	*P*	*r*	*P*	*r*	*P*	*r*
Coarse sand	***0.023***	***0.125***	0.120	0.070	***0.004***	***0.146***	***0.001***	***0.173***
Fine sand	0.630	−0.023	0.350	0.030	0.199	0.050	0.060	0.080
Total sand	0.076	0.086	***0.038***	***0.087***	***0.001***	***0.157***	***0.003***	***0.135***
Coarse silt	***0.003***	***0.158***	***0.009***	***0.130***	0.313	0.056	0.119	0.085
Fine silt	***0.014***	***0.155***	***0.031***	***0.121***	***0.000***	***0.194***	***0.010***	***0.135***
Total silt	***0.011***	***0.158***	0.080	0.094	***0.001***	***0.173***	***0.046***	***0.090***
Clay	***0.052***	***0.127***	***0.011***	***0.133***	***0.004***	***0.150***	0.311	0.039
C	0.118	0.108	***0.023***	***0.130***	***0.000***	***0.227***	***0.009***	***0.136***
N	***0.040***	***0.143***	***0.034***	***0.123***	***0.004***	***0.175***	0.097	0.074
pH	***0.034***	***0.119***	***0.022***	***0.112***	0.117	0.066	***0.003***	***0.147***

C, carbon; N, nitrogen; COI, cytochrome *c* oxidase subunit I gene.
Significant *P* values (<0.05) and the associated correlation coefficients (*r*) are shown in bold and italics.

Table 5 Correlation coefficients between soil variables and principal component analysis factors

Variables	Factor 1	Factor 2	Factor 3
%Coarse sand	***−0.580***	***−0.666***	***0.373***
%Fine sand	***−0.423***	0.196	***−0.643***
%Coarse silt	0.032	−0.031	***−0.886***
%Fine silt	0.277	***0.881***	***0.303***
%Clay	***0.688***	***−0.477***	−0.108
%C	***0.843***	0.200	0.155
%N	***0.837***	0.256	0.201
pH	−0.068	***0.721***	***−0.321***
Explained variance (%)	36.348	25.024	16.147

C, carbon; N, nitrogen.
Values in bold and italics are significant ($P < 0.05$).

Discussion

Characteristics of the used genes and genetic variability

The genetic variability found in these endogeic Mediterranean earthworms has proven to be high, when compared with other studied annelids at similar taxonomic levels. However, the evolutionary pattern does not seem constant along the different lineages in the family, as we have detected important differences between clades regarding genetic and morphological evolution. On one hand, the *Hormogaster elisae* complex (HE), whose distribution is limited to the central Iberian Peninsula, includes only one morphospecies showing a constancy in all morphological diagnostic characters, but exhibiting great genetic variability and a deep genetic subdivision. Other examples of large genetic divergence across short geographical distances (in the order of a few kilometres or even in the same site) have been reported for earthworms (James *et al.* 2010; Dupont *et al.* 2011; Férnandez 2011), in which cryptic speciation could be common.

On the other hand, the clade including *Hormogaster* from the NE Iberian Peninsula (HNE), with a similar sampling effort, includes 15 described species (in this work, we include 13 described species and some varieties, as well as a possible new species, but *H. multilamella* and *H. lleidana* were not found). Qiu & Bouché (1998) described the morphological diversity of this clade that occupies a slightly larger area in the northeastern part of the Iberian Peninsula. The genetic variability of HNE is not as great as in the HE clade and it does not present such deep subdivisions in mitochondrial data (see Splitstree in Fig. 1), despite showing much larger morphological diversity and including highly autapomorphic taxa, currently considered monotypic genera. Regarding the nuclear data, both clades show a core of sequence diversity that seems to be the origin of the remaining sequence types (see TCS networks for H4 in Fig. 1). All this evidence suggests that these two clades are undergoing different evolutionary processes and these could be attributed to the particular characteristics of the soil that they inhabit (see below), or to some rarely documented evolutionary processes.

For the remaining clades, no sound conclusions can be drawn due to the lack of sufficient information available. It would be particularly interesting to explore the hormogastrid fauna from Corsica and Sardinia to shed light on the potential cryptic speciation processes occurring there, as indicated by the great genetic divergences between sampled localities of *Hormogaster redii* and also

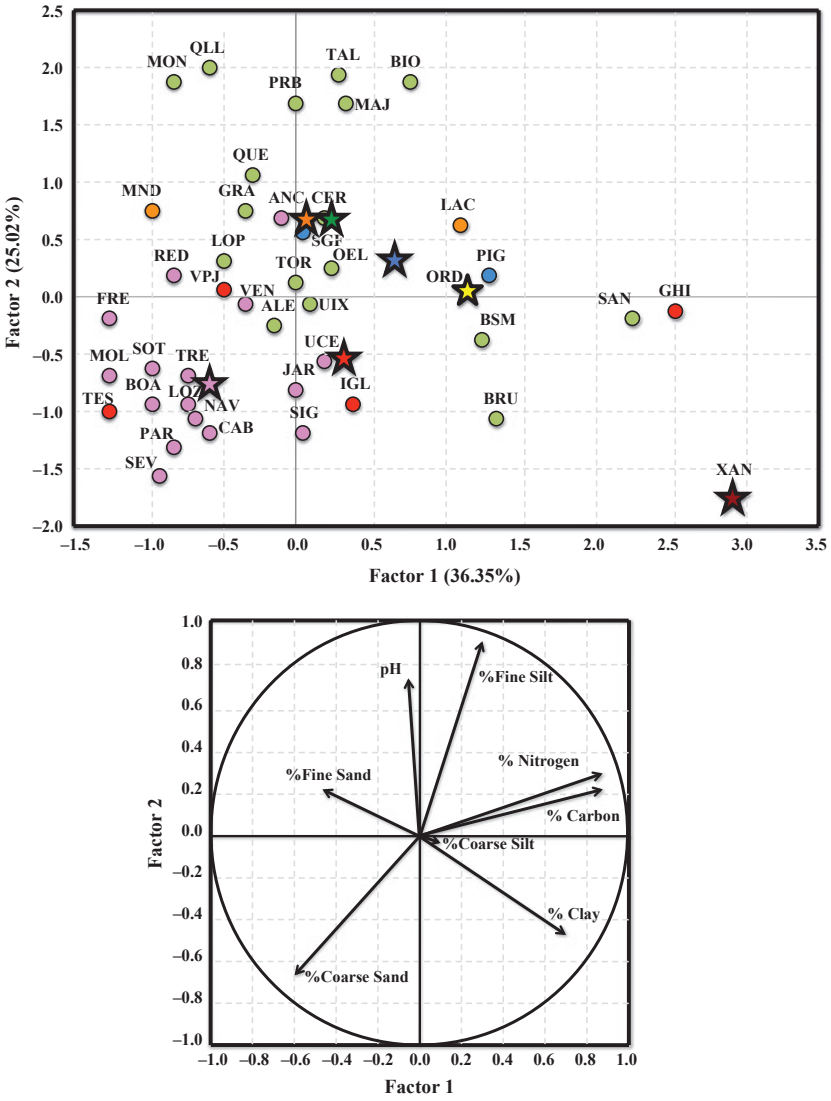

Fig. 2 Plot of the factor scores for soil analyses on the first two principal components for hormogastrid localities. Graphical illustration of the localities, indicating the variance explained by each axis (above) and correlation circle representation of the soil variables (below). Localities were coloured according to the clade in which they were included in previous phylogenetic analyses. Localities included in the clade HE are represented in pink, HCL in yellow, HIS in red, HM in orange, HNE in green, VG in blue and XN in brown. See codes for the localities and complete name of species and clades in Table 1. Barycentre coordinates of each clade are represented by a star of the corresponding colour. Note that the clades that include only one locality (HCL and XN) are only represented by this star.

by the GMYC results that identify each locality as a putative entity.

Species boundaries

Recently, the use of coalescent-based models of species delineation has flourished for studying difficult groups based on morphology alone (Nekola *et al.* 2009), for assigning immature specimens to their adult counterparts (Gattolliat & Monaghan 2010), to match spider genders where males and females are often described based on independent sets of character (M.A. Arnedo & G. Hormiga, unpublished data), or to perform rapid biodiversity estimates for taxonomic groups in areas where little taxonomic information is available (Pons *et al.* 2006; Kaya *et al.* 2009; Monaghan *et al.* 2009). The GMYC method has proven to work well for many of these examples, including arthropods and molluscs (e.g. Monaghan *et al.* 2009; Nekola *et al.* 2009; Gattolliat & Monaghan 2010), but almost no published account exists for annelids or oligochaetes. Only a very recent publication has explored the method in anecic earthworms

showing surprising results when compared to traditional taxonomy (Fernández *et al.* in press), and we found interesting to further test this method with these animals.

Papadopoulou *et al.* (2008) compared two groups of beetles occupying different habitats and with different dispersal habilities showing that in stable habitats (expected to have lower dispersal rates) there tends to be a higher subdivision of clades when applying the GMYC model, and lineage branching occurred more deeply in the tree. Also, the lineages from the more stable habitats had higher levels of nucleotide and haplotype diversity and greater geographical structure, perhaps indicating that the populations had had more time to diversify. This is indeed the case of hormogastrid earthworms. The soil is a very stable milieu, the dispersal rates of earthworms are thought to be very low (Ligthart & Peek 1997; Hale *et al.* 2005), and these animals tend to stay put in the same place for long periods of time, following the fate of their habitat and geographical area (Omodeo & Rota 2008). Specifically, in hormogastrid earthworms, this pattern is magnified because their active dispersal capabilities and passive dispersal opportunities are even lower (Novo *et al.* 2010b). Still, this is not an explanation for why the individuals from the same populations in clade HE are considered different entities by the GMYC model, but just an indication that they indeed show deep genetic divergences. Following Wiens & Penkrot's (2002) phylogenetic concept, these subdivisions make little sense. The genetic variability in this complex of species exceeds the predictable values and seems to be undergoing different evolutionary processes, as shown in some soil arthropods (Boyer *et al.* 2007). The HE clade already shows extreme subdivision when constructing statistical parsimony networks (with TCS, Novo *et al.* 2010a), also separating individuals from the same populations. Other authors stated that independent haplotype networks generally agree with named species or species subgroups (Pons *et al.* 2006; Hart & Sunday 2007), but again the HE clade may not follow this rule.

Lohse (2009) argued that incomplete sampling of demes involved in the coalescent process could artificially overestimate species numbers by the GMYC method. Papadopoulou *et al.* (2009) responded that indeed, if populations (demes) with intermediate haplotype composition are left unsampled, these results are an overestimate of species (oversplitting). Nevertheless, they argue that the problems raised by Lohse (2009) result from the fact that the offspring is produced in the vicinity of the parents, which in turn produces greater similarity of genotypes at a site compared with other sites. This is surely the case of hormogastrid earthworms (Novo *et al.* 2010b) but it only could explain the subdivision pattern found as one population corresponding to one entity.

Species delimitation within earthworms is still an open question because their morphology not always captures the true diversity owing to their structural simplicity and specific adaptations to the soil. Here, we show that coalescent methods overestimate the number of species, probably due to the marked genetic structure and scarce dispersion capacity. Further research in this important topic is therefore needed.

Molecular rates

The first conclusion regarding molecular rates is that a unique value cannot be generalized for this particular family. As shown in Fig. 2, these values not only vary along the tree, but they also differ depending on the prior used for the analyses. Thus, general statements such as the one proposed by Omodeo (2000), who indicated (based on continental drift) the need of 180 MY for an earthworm genus to differenciate, cannot be accepted and the rates proposed by other authors (e.g. Chang & Chen 2005; Chang *et al.* 2008; Buckley *et al.* 2011) working on megascolecids, should be taken with a grain of salt and limited to the species studied and to the concrete scenario of each particular case. They can definitely not be extrapolated to hormogastrids in general or to any particular clade. Our data suggest that the HE lineage has undergone quicker diversification than HNE as it presents higher rates, as shown by the ANOVA results. Conclusions about the remaining clades should await further collecting effort.

These results confirm cryptic speciation within the HE lineage in the central Iberian Peninsula, where a mismatch between morphological and molecular evolution has been reported (Novo *et al.* 2009, 2010a). Here, we propose that the characteristics of the soils where the members of clade HE live limits their morphological change, and thus, even molecular evolution is fast morphological changes do not have time to fixate. The harsher soil conditions (i.e. dryer and with less organic matter than other soil typically inhabited by earthworms) they inhabit could be implied also in this rate acceleration.

Geographical and environmental assessment

Hormogastrid evolutionary history is largely shaped by the geographical history of the landmasses they inhabit, probably because of their low vagility. Although confined to their particular areas, soil conditions could also be shaping the genetic processes of these earthworms, as shown by the correlation between the genetic divergences and the differences in soil characteristics

among collection sites. However, no single soil property is responsible for these differences, and instead, it is the combined effect of these properties what shows a correlation. Novo *et al.* (2010a) found that some soil texture characteristics have a larger influence in the differenciation between the *H. elisae* lineages. The higher diversity of soils included in this work shows relevance of other factors, such as carbon or nitrogen content, when taking into account the whole family Hormogastridae, as opposed to a single lineage.

The PCA permits a better understanding of the ecological scenario by combining the soil variables in factors. In general, hormogastrids are confined to poor soils, with low nitrogen and organic matter content, being relegated to the soils where other earthworms cannot survive, sometimes by means of exclusive competition (Ramajo 2010). The percentages in organic matter found in the studied localities were low in comparison with those observed by other authors studying earthworms in other areas of the Iberian Peninsula (Mariño *et al.* 1985). The localities included in the HE clade exhibit a tendency towards more acidic soils, with higher content of coarse sand, and slightly poorer, when compared to the localities included in the HNE clade, whose members inhabit soils with finer texture and higher pH.

The characteristics of the soil may be the reason why morphological adaptations are generally maintained in earthworms, causing thus morphological stasis. This is certainly the case of hormogastrids, as they are adapted to harsh soil conditions, and magnified in the case of *H. elisae* (HE), enduring very harsh soil conditions among oligochaetes. This, together with the extreme climatic characteristics (cold winters and hot and dry summers) in the central Iberian Peninsula makes the habitat of HE unsuited for most species of earthworms (Hernández *et al.* 2007). Other authors already proposed that extreme subsurface conditions could constrain morphological evolution and be responsible for convergence over large periods of time (e.g. for amphipods, beetles and salamanders: Jones *et al.* 1992; Caccone & Sbordoni 2001; Wiens *et al.* 2003). The morphological adaptations already evolved to inhabit these harsh habitats are probably optimal and any change could be nonadaptive. The specialized environment can impose stabilizing selection, thus minimizing or eliminating the morphological change that can occur during speciation (Bickford *et al.* 2007). Therefore, the high genetic divergences found within HE could be related to biochemical changes not detectable morphologically. We propose here, as observed for clade HE that the harsh soil conditions these earthworms have to endure contribute to accelerate genetic evolution (i.e. higher molecular rates).

Principal component analysis detected that *Xana* inhabits soils that differ from those of other hormogastrids, with a much higher clay content. Thus, the hypothesis of Novo *et al.* (2011) suggesting the reduction from three to two gizzards as an adaptation to the environment remains plausible. In that study, *Xana* and *Vignysa* (the only hormogastrids presenting two gizzards, while the others have three) were not related phylogenetically in most analyses (but they clustered in the analyses based on nuclear genes).

The soil collected in Corsica and Sardinia (where the clade HIS inhabits) is heterogeneous, as shown by the PCA, where the samples from both islands appear mixed without a clear tendency. This heterogeneity could be explained by the geological origin of the islands (Bacchetta *et al.* 2007; Omodeo & Rota 2008) and could also be the cause of the unexpected genetic diversity found within this clade. More populations of *H. redii* should be collected to study the real scope of this diversity.

A pattern of isolation-by-distance has been found in the HE clade but not for the HNE clade. Nevertheless, the relationships recovered within HNE show clusters related to geographical location. As mentioned earlier, hormogastrids show strong biogeographical patterns of vicariance and a diffusive-like colonization is expected via active dispersal (very slow) or via nonanthropogenic passive dispersal (e.g. by birds, mammals, wind or waterways) (Cameron *et al.* 2008). The fact that the isolation-by-distance pattern is significant in the case of HE but not for HNE demonstrates that in the former the great genetic diversity concentrated in a small geographical area has a deep genetic basis and that the animals show low vagility, as postulated for the members of this family.

The geographical distribution of HE could also explain its high diversity, as it is located in the westernmost limit of the Hormogastridae distribution (Cobolli-Sbordoni *et al.* 1992), exploiting and colonizing marginal habitats. Among sampled populations of HE, the ones in the mountainous area (Guadarrama Mountains, see Novo *et al.* 2009, 2010a for a detailed map) exhibit higher diversity and were easier to locate, which could be a sign of greater population densities and probably a more favourable area for hormogastrid development. Meanwhile, the conditions in the southern area may be less optimal, as reflected in the difficulties locating populations and their lower diversity values. These populations seem to represent the vanguard of this evolutionary lineage, which is the reason why the speciation and diversification is increasing in this southern area (including the localities PAR, SEV, BOA in the Tajo fluvial valley, see Novo *et al.* 2009, 2010a for a detailed map and more information). These hypotheses are consistent with the pattern found

by Novo *et al*. (2010a) who showed expansion processes in the southern populations and stability in mountainous populations.

Conclusions

Morphological changes in such a restrictive milieu as the soil are unlikely once the animals are adapted, and thus, morphology-based studies could confound the evolutionary patterns in certain groups of invertebrates, such as earthworms, showing sometimes high levels of homoplasy because of the convergence in morphological solutions to specific conditions. In this study, we found how within hormogastrid earthworms, morphological and molecular evolution can be decoupled in some clades. The high genetic diversification and morphological constancy in the HE clade from the central Iberian Peninsula contrasts with the higher morphological diversity in the HNE clade of NE Iberia, and these differences could be attributed to several factors. Higher molecular rates in HE could be the cause of this decoupling between the rapid molecular evolution and the morphological stasis found in this clade, and this could be accentuated by the extreme conditions it is subjected to, which at the same time could be forcing the higher evolutionary rates leading to changes only at the biochemical level. The GMYC analyses yielded a higher number of entities than expected under a morphospecies concept within Hormogastridae, generally showing each population as one different entity (i.e. species) probably due to the low vagility of these earthworms. The method yielded particularly high numbers of entities for HE, probably because of its high genetic diversity. Further studies for different invertebrates inhabiting similarly uniform and isolated environments would be necessary to shed light on these interesting evolutionary processes.

Acknowledgements

We are indebted to Carlos Bayón, Pablo Cabo, Inés García, Noa Novo and Juan José Molina, for assisting with fieldwork, Maigualida Ricoy and Pietro Omodeo for sharing additional samples, Irene Parra, Carlos Granado, Miquel Arnedo and Joan Pons for offering their suggestions in data analyses and Daniel Ayllón and Sheila Leal for technical support. The members from Giribet laboratory offered help for laboratory work and data analyses, especially Sebastián Vélez, Sónia Andrade and Prashant Sharma. The Editors and anonymous reviewers provided insightful comments that contributed to improve an earlier version of this work. M.N. was supported by a Grant from *Fundación Caja Madrid*. This research was funded by projects CGL2007-60715/BOS and CGL2010/16032 from the Spanish Government. Additional funding came from the Museum of Comparative Zoology and from the US National Science Foundation (#EF-0531757: AToL: Collaborative Proposal: Assembling the Protostome Tree of Life).

References

Anne P (1945) Sur le dosage rapide du carbone organique des sols. *Annales Agronomiques*, **2**, 162–172.

Altschul SF, Madden TL, Schaffer AA *et al*. (1997) Gapped BLAST and PSIBLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **25**, 3389–3402.

Bacchetta G, Mandis G, Pontecorvo C (2007) Contribution to the knowledge of the endemic vascular flora of Sulcis (SW Sardinia – Italy). *Bocconea*, **21**, 155–166.

Bely AE, Weisblat DA (2006) Lessons from leeches: a call for DNA barcoding in the lab. *Evolution and Development*, **8**, 491–501.

Bickford D, Lohman DJ, Sodhi NS *et al*. (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution*, **22**, 148–155.

Boyer SL, Baker JM, Giribet G (2007) Deep genetic divergences in *Aoraki denticulata* (Arachnida, Opiliones, Cyphophthalmi): a widespread 'mite harvestman' defies DNA taxonomy. *Molecular Ecology*, **16**, 4999–5016.

Briones MJI, Morán P, Posada D (2009) Are the sexual, somatic and genetic characters enough to solve nomenclatural problems in lumbricid taxonomy? *Soil Biology and Biochemistry*, **41**, 2257–2271.

Buckley TR, James S, Allwood J, Bartlam S, Howitt R, Prada D (2011) Phylogenetic analysis of New Zealand earthworms (Oligochaeta: Megascolecidae) reveals ancient clades and cryptic taxonomic Diversity. *Molecular Phylogenetics and Evolution*, **58**, 85–96.

Caccone A, Sbordoni V (2001) Molecular biogeography of cave life: a study using mitochondrial DNA from bathysciine beetles. *Evolution*, **55**, 122–130.

Cameron EK, Bayne EM, Coltman DW (2008) Genetic structure of invasive earthworms *Dendrobaena octaedra* in the boreal forests of Alberta: insights into introduction mechanisms. *Molecular Ecology*, **17**, 1189–1197.

Challis RJ, Mutun S, Nieves-Aldrey JL *et al*. (2007) Longitudinal range expansion and cryptic eastern species in the western Palaearctic oak gallwasp, *Andricus coriarius*. *Molecular Ecology*, **16**, 2103–2114.

Chang CH, Chen JH (2005) Taxonomic status and intraspecific phylogeography of two sibling species of Metaphire (Oligochaeta: Megascolecidae) in Taiwan. *Pedobiologia*, **49**, 591–600.

Chang CH, Lin SM, Chen JH (2008) Molecular systematics and phylogeography of the gigantic earthworms of the Metaphire formosae species group (Clitellata, Megascolecidae). *Molecular Phylogenetics and Evolution*, **49**, 958–968.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Cobolli-Sbordoni M, de Matthaeis E, Alonzi A, Matoccia M, Omodeo P, Rota E (1992) Speciation, genetic divergence and paleogeography in the Hormogastridae. *Soil Biology and Biochemistry*, **24**, 1213–1221.

Crandall KA, Templeton AR, Sing CF (1994) Intraspecific cladogram estimation: problems and solutions. In: *Models in Phylogeny Reconstruction* (eds Scotland RW, Siebert DJ and Williams DM), pp. 273–297. Clarendon Press, Oxford.

Darwin C (1881) *The Formation of Vegetable Mould through the Actions of Worms*. John Murray, London.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Dupont L, Lazrek F, Porco D *et al.* (2011) New insight into the genetic structure of the *Allolobophora chlorotica* aggregate in Europe using microsatellite and mitochondrial data. *Pedobiologia*, **54**, 217–224.

Edwards CA, Bohlen PJ (1996) *Biology and Ecology of Earthworms*, 3rd edn. Chapman and Hall, London.

Edwards CA (ed.) (2004) *Earthworm Ecology*, 2nd edn. CRC Press.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Férnandez R (2011) *Aporrectodea trapezoides (Dugés, 1828) (Oligochaeta, Lumbicidae): Filogeografía, Filogenia y Biología reproductiva*. PhD dissertation. Universidad Complutense de Madrid, Spain. 251 pp.

Fernández R, Almodóvar A, Novo M, Gutiérrez M, Díaz Cosín DJ (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, Almodóvar A, Novo M, Simancas B, Díaz Cosín DJ (in press) 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*, http://dx.doi.org/10.1016/j.ympev.2012.04.011.

Finston TL, Johnson MS, Humphreys WF, Eberhard SM, Halse SA (2007) Cryptic speciation in two widespread subterranean amphipod genera reflects historical drainage patterns in an ancient landscape. *Molecular Ecology*, **16**, 355–365.

Fontaneto D, Herniou EA, Boschetti C *et al.* (2007) Independently evolving species in asexual bdelloid rotifers. *PLoS Biology*, **5**, 914–921.

Gattolliat JL, Monaghan MT (2010) DNA-based association of adults and larvae in Baetidae (Ephemeroptera) with the description of a new genus *Adnoptilum* in Madagascar. *Journal of the North American Benthological Society*, **29**, 1042–1057.

Guitián F, Carballas T (1976) *Técnicas de análisis de suelos*, 2nd edn. Editorial Pico Sacro, Santiago de Compostela.

Gustafsson DR, Price DA, Erseus C (2009) Genetic variation in the popular lab worm *Lumbriculus variegatus* (Annelida: Clitellata: Lumbriculidae) reveals cryptic speciation. *Molecular Phylogenetics and Evolution*, **51**, 182–189.

Hale CM, Frelich LE, Reich PB (2005) Exotic European earthworm invasion dynamics in northern hardwood forest of Minnesota, USA. *Ecological Applications*, **15**, 848–860.

Hart MW, Sunday J (2007) Things fall apart: biological species form unconnected parsimony networks. *Biology Letters*, **3**, 509–512.

Hernández P, Gutiérrez M, Ramajo M, Trigo D, Díaz Cosín DJ (2003) Horizontal distribution of an earthworm community at El Molar, Madrid (Spain). *Pedobiologia*, **47**, 568–573.

Hernández P, Fernández R, Novo M, Trigo D, Díaz Cosín DJ (2007) Geostatistical and multivariate analysis of the horizontal distribution of an earthworm community in El Molar (Madrid, Spain). *Pedobiologia*, **51**, 13–21.

Hogg ID, Stevens MI, Schnabel KE, Chapman MA (2006) Deeply divergent lineages of the widespread New Zealand amphipod *Paracalliope fluviatilis* revealed using allozyme and mitochondrial DNA analyses. *Freshwater Biology*, **51**, 236–248.

Huson DH, Bryant D (2006) Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology and Evolution*, **23**, 254–267.

James S, Porco D, Decäens T, Richard B, Rougerie R, Erséus C (2010) Barcoding reveals cryptic diversity in *Lumbricus terrestris* L., 1758 (Clitellata): resurrection of *L. herculeus* (Savigny, 1826). *PLoS One*, **5**, e15629.

Jones R, Culver DC, Kane TC (1992) Are parallel morphologies of cave organisms the result of similar selection pressure? *Evolution*, **46**, 353–365.

Kaya M, Herniou EA, Barraclough TG, Fontaneto D (2009) Inconsistent estimates of diversity between traditional and DNA taxonomy in bdelloid rotifers. *Organisms Diversity and Evolution*, **9**, 3–12.

King RA, Tibble AL, Symondson WOC (2008) Opening a can of worms: unprecedented sympatric speciation within British lumbricid earthworms. *Molecular Ecology*, **17**, 4684–4698.

Lavelle P, Spain AV (2001) *Soil Ecology*. Kluwer Academic Publishers, London, UK.

Lee CE, Frost BW (2002) Morphological stasis in the *Eurytemora affinis* species complex (Copepoda: Temoridae). *Hydrobiologia*, **480**, 111–128.

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

Ligthart TN, Peek GJCW (1997) Evolution of earthworm burrow systems alter inoculation of lumbricid earthworms in a pasture in the Netherlands. *Soil Biology and Biochemistry*, **29**, 453–462.

Lohse K (2009) Can mtDNA barcodes be used for delimit species? A response to Pons *et al.* (2006). *Systematic Biology*, **58**, 439–442.

Mantel N (1967) The detection of disease clustering and generalized regression approach. *Cancer Research*, **27**, 209–220.

Mariño F, Trigo MD, Díaz Cosín DJ, Calvin EB (1985) Lombrices de tierra de Galicia. Relación con los factores del suelo. *Anales de Edafología y Agrobiología*, **44**, 1641–1649.

Mayr E (1948) The bearing of the new systematics on genetical problems. The nature of species. *Advances in Genetics*, **2**, 205–237.

Mayr E (1963) *Animal Species and Evolution*. Harvard University Press, Cambridge.

McGuigan K, Sgrò CM (2009) Evolutionary consequences of cryptic genetic variation. *Trends in Ecology and Evolution*, **24**, 305–311.

Monaghan MT, Wild R, Elliot M *et al.* (2009) Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Systematic Biology*, **58**, 298–311.

Müller J (2000) Mitochondrial DNA variation and the evolutionary history of cryptic *Gammarus fossarum* types. *Molecular Phylogenetics and Evolution*, **15**, 260–268.

Nekola JC, Coles BF, Bergthorssson U (2009) Evolutionary pattern and process within the *Vertigo gouldii* (Mollusca: Pulmonata, Pupillidae) group of minute North American land snails. *Molecular Phylogenetics and Evolution*, **53**, 1010–1024.

Novo M, Almodóvar A, Díaz-Cosín DJ (2009) High genetic divergence of hormogastrid earthworms (Annelida, Oligochaeta) in the central Iberian Peninsula: evolutionary and demographic implications. *Zoologica Scripta*, **38**, 537–552.

Novo M, Almodóvar A, Fernández R, Trigo D, Díaz Cosín DJ (2010a) Cryptic speciation of hormogastrid earthworms revealed by mitochondrial and nuclear data. *Molecular Phylogenetics and Evolution*, **56**, 507–512.

Novo M, Almodóvar A, Fernández R, Gutiérrez M, Díaz Cosín DJ (2010b) Mate choice of an endogeic earthworm revealed by microsatellite markers. *Pedobiologia*, **53**, 375–379.

Novo M, Almodóvar A, Fernández R, Giribet G, Díaz Cosín DJ (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.

Omodeo P (2000) Evolution and biogeography of megadriles (Annelida, Clitellata). *The Italian Journal of Zoology*, **67**, 179–201.

Omodeo P, Rota E (2008) Earthworm diversity and land evolution in three mediterranean districts. *Proceedings of the California Academy of Sciences, Fourth Series*, **59**, 65–83.

Page AL, Miller RH, Keeney DR (1982) *Methods in Soil Analysis. Part 2: Chemical and Microbiological Properties*. Agronomy 9, ASA-SSSA, Madison, Wisconsin, USA.

Papadopoulou A, Bergsten J, Fujisawa T, Monaghan MT, Barraclough TG, Vogler AP (2008) Speciation and DNA barcotes: testing the effects of dispersal on the formation of discrete sequence clusters. *Philosophical transactions of the Royal Society B. Biological Sciences*, **363**, 2987–2996.

Papadopoulou A, Monaghan MT, Barraclough TG, Vogler AP (2009) Sampling error does not invalidate the Yule-Coalescent model for species delimitation. A response to Lohse (2009). *Systematic Biology*, **58**, 442–444.

Pfenninger M, Staubach S, Albrecht C, Schwenk K (2003) Ecological and morphological differentiation among cryptic evolutionary lineages in freshwater limpets of the nominal form-group *Ancylus fluviatilis* (O.F. Müller, 1774). *Molecular Ecology*, **12**, 2731–2745.

Pons J, Barraclough TG, Gómez-Zurita J *et al*. (2006) Sequence based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, **55**, 595–609.

Pop AA, Wink M, Pop VV (2003) Using of 18S, 16S rDNA and cyctochrome c oxidase sequences in earthworm taxonomy (Oligochaeta, Lumbricidae). *Pedobiologia*, **47**, 428–433.

Posada D (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Qiu JP, Bouché MB (1998) Contribution to the taxonomy of Hormogastridae (Annelida: Oligochaeta) with description of new species from Spain. *Documents Pedozoologiques et Integrologiques*, **4**, 164–177.

Ramajo M (2010) *Relaciones interespecíficas de lombrices de tierra en una parcela de El Molar (Madrid)*. PhD dissertation. Facultad de Biología. Universidad Complutense de Madrid, Spain, 301 pp.

Rambaut A, Drummond AJ (2007) Tracer v1.4. [Online]. Available from http://beast.bio.ed.ac.uk/Tracer.

Rougerie R, Decaëns T, Deharveng L *et al*. (2009) DNA barcodes for soil animal taxonomy. *Pesquisa Agrpecuaria Brasileira*, **44**, 789–801.

Schettino A, Turco E (2006) Plate kinematics of the western Mediterranean region during the Oligocene and early Miocene. *Geophysical Journal International*, **166**, 1398–1423.

Stoks R, Nystrom JL, May ML, McPeek MA (2005) Parallel evolution in ecological and reproductive traits to produce cryptic damselfly species across the Holarctic. *Evolution*, **59**, 1976–1988.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data III. Cladogram estimation. *Genetics*, **132**, 619–633.

Tully T, D'Haese CA, Richard M, Ferriere R (2006) Two major evolutionary lineages revealed by molecular phylogeny in the parthenogenetic collembola species *Folsomia candida*. *Pedobiologia*, **50**, 95–104.

Wiens JJ, Penkrot TA (2002) Delimitig species using DNA and morphological variation and discordant species limits in spiny lizards (Sceloporus). *Systematic Biology*, **51**, 69–91.

Wiens JJ, Chippindale PT, Hillis DM (2003) When are phylogenetic analyses misled by convergence? A case study in Texas cave salamanders. *Systematic Biology*, **52**, 501–514.

M.N., R.F. and D.J.D.C. are interested in systematics, phylogeny, phylogeography and reproduction of earthworms. A.A. is interested in molecular and evolutionary ecology of freshwater fishes and phylogeography of invertebrate animals. D.T. is interested in soil ecology and the role of earthworms in restoration. G.G. is interested in the origins and maintenance of animal diversity, using invertebrates as model organisms.

Data accessibility

DNA sequences: GenBank Accession nos: JN209196–JN209829.

See Table 1 for specific Accession nos on different loci and samples.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Intrapopulation diversity values for hormogastrids.

Table S2 Mean values of genetic divergence (K2P in percentage of changes) between pairs of populations within hormogastrids, based on mitochondrial data (COI below the diagonal; 16S-tRNAs above the diagonal).

Table S3 Mean values of genetic divergence (K2P in percentage of changes) between pairs of populations within hormogastrids, based on nuclear data (H4 below the diagonal; H3 above the diagonal).

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.