Zoological Journal of the Linnean Society, 2018, **183**, 749–762. With 3 figures.

Cryptic diversity in supposedly species-poor genera of Enchytraeidae (Annelida: Clitellata)

SVANTE MARTINSSON and CHRISTER ERSÉUS*

Systematics and Biodiversity, Department of Biological and Environmental Sciences, University of Gothenburg, Box 463, SE-405 30 Göteborg, Sweden

Received 19 May 2017; revised 1 September 2017; accepted for publication 27 September 2017

Using a two-step workflow, we test the species boundaries in three genera of enchytraeid worms, *Globulidrilus*, *Hemifridericia* and *Stercutus* (Clitellata: Enchytraeidae), which contains one to three nominal species each. For the species discovery phase, DNA barcode-based clustering analyses in ABGD (Automatic Barcode Gap Discovery) are performed, using mitochondrial cytochrome *c* oxidase subunit I (*COI*) data. The clusters from these analyses are then used as input in the species validation phase, where multispecies coalescent-based multilocus species delimitation analyses are performed in BPP (Bayesian Phylogenetics and Phylogeography) using nuclear Histone 3 (*H3*) and Internal Transcribed Spacer (ITS) data. For all BPP analyses, several species delimitation arrangements are included in the 95% credibility interval, and no complete sets of species are well supported. However, we conclude that our data set comprises at least seven species of *Globulidrilus*, all attributed to the nominal morpho-species *G. riparius*, three species of *Hemifridericia*, whereof two are cryptic lineages within *H. parva* and one is *H. bivesiculata*, and at least six species of the previously monotypic *Stercutus*. The species are not formally described here due to a lack of mature specimens of many of them in combination with low support for some of them in the genetic analyses. However, this is the first step towards a better understanding of the diversity within these groups.

ADDITIONAL KEYWORDS: BPP – BP&P – *Globulidrilus* – *Hemifridericia* – multispecies coalescent – species delimitation – *Stercutus*.

INTRODUCTION

With the introduction of molecular data, together with the development of species delimitation methods, it has been shown that many nominal species actually are complexes of morphologically very similar, or identical, species, the so-called cryptic species (Bickford *et al.*, 2007). Such species have been found in most animal groups (Pfenninger & Schwenk, 2007), including clitellate annelids (e.g. Erséus & Gustafsson, 2009; Novo *et al.*, 2010; Matamoros, Rota & Erséus, 2012; Donnelly *et al.*, 2013; Liu *et al.*, 2017; Martinsson & Erséus, 2017).

In this study, we focus on three assumedly species-poor genera (one to three nominal species in each) of the family Enchytraeidae (Annelida: Clitellata): *Hemifridericia* Nielsen & Christensen, 1959, *Globulidrilus* Christensen & Dózsa-Farkas, 2012 and *Stercutus* Michaelsen, 1888. Today, three species are referred to *Hemifridericia*, one of which however, *H. varanensis* Lal, Singh & Prasad, 1981, does not seem to belong to this genus (Dózsa-Farkas & Felföldi, 2015). *Hemifridericia parva* Nielsen & Christensen, 1959 is widespread in Europe and is also reported from China and USA, whereas the third species, *Hemifridericia bivesiculata* Christensen & Dózsa-Farkas, 2006, is so far only known from Hungary and Arctic Canada. *Hemifridericia* is found in a clade together with *Buchholzia* and *Fridericia* (Martinsson *et al.*, 2017a). *Globulidrilus* consists of *G. helgei* Christensen & Dózsa-Farkas, 2012 (type species) that has so far only been found in South Korea and the Holarctic *G. riparius* (Bretscher, 1889) that was recently transferred from *Marionina* (Christensen & Dózsa-Farkas, 2012). *Globulidrilus* is found in a clade together with *Bryodrilus*, *Henlea*, *Oconnorella* as well as *Claparedrilus semifuscoides* (as *Lumbricillus*

*Corresponding author. E-mail: christer.erseus@bioenv.gu.se

semifuscus) and *Marionina communis*; the latter is not found together with the rest of their respective congenerics (Erséus *et al.*, 2010; Klinth, Martinsson & Erséus, 2017; Martinsson *et al.*, 2017a). The third genus in this study, *Stercutus*, is monotypic for *S. niveus* Michaelsen, 1888, which is found in both Europe and North America. This genus is found in a clade together with *Chamaedrilus* and *Euenchytreus*, two genera that were previously combined into *Cognettia* (Martinsson *et al.*, 2017a).

The species delimitation process can be divided into two steps, species discovery and species validation (Carstens *et al.*, 2013). In the first, specimens are grouped into groups/putative species, usually using a single data source, for example morphology or DNA barcoding, and these species hypotheses are then tested in the latter with additional data and analyses (Carstens *et al.*, 2013). It has been shown that the delimitation success increases with the number of markers and that a single marker is not enough for a solid well-supported delimitation (Dupuis, Roe & Sperling, 2012). The development of the methods used for species delimitation using molecular data has been rapid (see e.g. Sites & Marshall, 2003; Fujita *et al.*, 2012), and in recent years, methods to analyse several loci together in a single analysis for species delimitation has become available (Rannala, 2015). Some of these are based on the multispecies coalescent model (Rannala & Yang, 2003). In this model, genes evolve inside a species phylogeny where the branches are species and the properties of the branches restrict the gene trees. One of these restrictions is that the divergence times between species have to be more recent than the coalescent times for any genes shared between them (Rannala & Yang, 2003). This model can be used for statistical testing of species assignments (Fujita *et al.*, 2012; Rannala, 2015), and it is based on a clearly defined species concept in which a species constitutes a branch of a species tree, which is defined by abrupt speciation and no genetic exchange after the speciation event (Aydin *et al.*, 2014). Under this definition, a species is a separately evolving meta-population lineage, that is, the unified species concept suggested by De Queiroz (2007). Many multilocus species delimitation methods require the user to assign the specimens to putative species that are then tested; the software usually does this by collapsing the species tree and joining sister species, and thereby testing which of the species assignments fits the model the best (Fujita *et al.*, 2012; Rannala, 2015). In theory, it is possible to assign each specimen to its own putative species, but at least for some software products, this may increase the computational time so that the analyses are not practically possible to run (Yang & Rannala, 2014). Additionally, starting with single specimens as input species may lead to low posterior probabilities (PPs) for each of the delimited species, due to the high number of possible species assignments, together with the limited data for each of the putative species (Olave, Sola & Knowles, 2014).

In this study, we test the species diversity of *Globulidrilus*, *Hemifridericia* and *Stercutus*, using a combination of DNA-barcode-based clustering analyses for the species discovery phase, followed by multispecies coalescent-based multilocus species delimitation analyses of nuclear markers in the software BPP (also known as BP&P, Bayesian Phylogenetics and Phylogeography; Yang, 2015) for the species validation phase.

MATERIAL AND METHODS

Specimens, DNA sequencing and assembly

DNA data from a total of 101 specimens from Norway, Sweden, Hungary and USA were included in the study: 32 each of *S. niveus* and *G. riparius*, 35 of *H. parva* and 2 of *H. bivesiculata* (for details, see Table 1).

A vast majority of the sequences were newly generated (Table 1). DNA was extracted from the posterior ends of ethanol-preserved worms, and the anterior parts were mounted in Canada balsam to be used as physical vouchers. DNA was extracted using either Qiagen's DNeasy Blood & Tissue Kit or Epicentre QuickExtract DNA Extraction Solution 1.0, following the manufacturer's instructions. Three markers, the mitochondrial cytochrome *c* oxidase subunit I (*COI*) gene, the complete nuclear ribosomal Internal Transcribed Spacer (ITS) region and the nuclear gene Histone 3 (*H3*), were amplified using primers and PCR programmes listed in Supporting Information, Table S1. Sequencing was carried out by Macrogen Inc. (Seoul, Korea) and Eurofins MWG Operon (Ebersberg, Germany). As specified in Table 1, a few sequences are from previously published works (Erséus *et al.*, 2010; Martinsson & Erséus, 2014; Dózsa-Farkas & Felföldi, 2015; Martinsson *et al.*, 2017a) and were all downloaded from GenBank. Sequences were assembled in Geneious Pro v. 7.1. (Biomatters Ltd.; http://www.geneious.com) and aligned separately for each genus using MAFFT v7.017 (Katoh *et al.*, 2002) as implemented in Geneious Pro v. 7.1., using the auto-algorithm and default settings. Information about the alignments (length and number of variable positions) can be found in Table 2. All sequences produced in this study are deposited in GenBank, and the vouchers are deposited in either the Swedish Museum of Natural History (SMNH), Stockholm, Sweden, or the University Museum of Bergen (ZMBN), Bergen, Norway (accession numbers in Table 1).

Table 1. Specimens included in the study, with individual specimen numbers, collection data and GenBank accession numbers

Specimen no	Museum voucher no	Species	*COI* cluster	Locality (country, state/province)	Coordinate		GenBank accession number		
					Latitude	Longitude	*COI*	*H3*	ITS
CE694	No voucher	*Globulidrilus riparius*	A	SE: Uppland	59.3609	18.0802	**MF801990**	**MF802085**	**MF802180**
CE2985	SMNH162823	*G. riparius*	A	SE: Öland	56.9855	16.8764	**MF801976**	**MF802071**	**MF802166**
CE2987	SMNH162824	*G. riparius*	A	SE: Öland	56.9855	16.8764	**MF801977**	**MF802072**	**MF802167**
CE3900	SMNH162825	*G. riparius*	A	SE: Västergötland	57.7761	12.2411	**MF801988**	**MF802083**	**MF802178**
CE11435	SMNH162826	*G. riparius*	A	SE: Öland	56.5444	16.6095	**MF801961**	**MF802056**	**MF802151**
CE27365	SMNH162827	*G. riparius*	A	SE: Skåne	55.5321	14.2701	**MF801971**	**MF802065**	**MF802160**
CE27369	SMNH162828	*G. riparius*	A	SE: Skåne	55.5321	14.2701	**MF801973**	**MF802068**	**MF802163**
CE27370	SMNH162829	*G. riparius*	A	SE: Skåne	55.5321	14.2701	**MF801974**	**MF802069**	**MF802164**
CE20290	ZMBN110441	*G. riparius*	A	NO: Østfold	59.0859	11.4359	**MF801963**	**MF802058**	**MF802153**
CE1127	SMNH108427	*G. riparius*	B	US: Alaska	60.4846	–150.0414	GU902096*	**MF802054**	**MF802149**
CE1128	SMNH162830	*G. riparius*	H	US: Alaska	60.4846	–150.0414	**MF801960**	**MF802055**	**MF802150**
CE23044	ZMBN110952	*G. riparius*	H	NO: Finnmark	23.3002	69.9497	**MF801964**	**MF802059**	**MF802154**
CE23046	ZMBN120602	*G. riparius*	C	NO: Finnmark	23.3002	69.9497	**MF801965**	**MF802060**	**MF802155**
CE30036	ZMBN120603	*G. riparius*	C	NO: Akershus	59.8624	11.2178	**MF801979**	**MF802074**	**MF802169**
CE2923	SMNH162831	*G. riparius*	C	SE: Öland	56.8621	16.8539	**MF801975**	**MF802070**	**MF802165**
CE11436	SMNH162832	*G. riparius*	C	SE: Öland	56.5444	16.6095	**MF801962**	**MF802057**	**MF802152**
CE27362	SMNH162833	*G. riparius*	C	SE: Skåne	55.5321	14.2701	**MF801968**	**MF802063**	**MF802158**
CE27367	SMNH162834	*G. riparius*	C	SE: Skåne	55.5321	14.2701	**MF801969**	**MF802067**	**MF802162**
CE3595	SMNH162835	*G. riparius*	E	US: Tennessee	36.1540	–86.2723	**MF801985**	**MF802080**	**MF802175**
CE3596	SMNH162836	*G. riparius*	E	US: Tennessee	36.1540	–86.2723	**MF801986**	**MF802081**	**MF802176**
CE3597	SMNH162837	*G. riparius*	E	US: Tennessee	36.1540	–86.2723	**MF801987**	**MF802082**	**MF802177**
CE3956	SMNH162838	*G. riparius*	D	SE: Västergötland	57.7761	12.2411	**MF801989**	**MF802084**	**MF802179**
CE30115	ZMBN120604	*G. riparius*	D	NO: Akershus	60.1252	11.4649	**MF801982**	**MF802077**	**MF802172**
CE27360	SMNH162839	*G. riparius*	F	SE: Skåne	55.5321	14.2701	**MF801966**	**MF802061**	**MF802156**
CE27361	SMNH162840	*G. riparius*	F	SE: Skåne	55.5321	14.2701	**MF801966**	**MF802062**	**MF802157**
CE27363	SMNH162841	*G. riparius*	F	SE: Skåne	55.5321	14.2701	**MF801970**	**MF802064**	**MF802159**
CE27366	No voucher	*G. riparius*	F	SE: Skåne	55.5321	14.2701	**MF801972**	**MF802066**	**MF802161**
CE29939	ZMBN120605	*G. riparius*	G	NO: Østfold	59.3347	11.6375	**MF801978**	**MF802073**	**MF802168**
CE30037	ZMBN120606	*G. riparius*	G	NO: Akershus	59.8624	11.2178	**MF801980**	**MF802075**	**MF802170**
CE30040	ZMBN120607	*G. riparius*	G	NO: Akershus	59.8624	11.2178	**MF801981**	**MF802076**	**MF802171**
CE30248	ZMBN120608	*G. riparius*	G	NO: Hedmark	60.1930	12.0288	**MF801983**	**MF802078**	**MF802173**
CE30249	ZMBN120609	*G. riparius*	G	NO: Hedmark	60.1930	12.0288	**MF801984**	**MF802079**	**MF802174**
CE794	No voucher	*Hemifridericia parva*	A	SE: Västergötland	58.6195	13.4258	GU902081*	KX644882‡	**MF802210**
CE9509	SMNH162842	*H. parva*	A	SE: Lappland	66.7996	21.8130	**MF801991**	**MF802115**	**MF802211**
CE9548	SMNH162843	*H. parva*	A	SE: Lappland	67.8546	20.2173	**MF801992**	**MF802116**	**MF802212**

Table 1. Continued

Specimen no	Museum voucher no	Species	*COI* cluster	Locality (country, state/province)	Coordinate		GenBank accession number		
					Latitude	Longitude	*COI*	*H3*	ITS
CE22423	SMNH162844	*H. parva*	A	SE: Lappland	68.4417	22.4794	**MF801997**	**MF802099**	**CE22423**
CE22424	SMNH162845	*H. parva*	A	SE: Lappland	68.4417	22.4794	**MF801998**	**MF802100**	**CE22424**
CE22425	SMNH162846	*H. parva*	A	SE: Lappland	68.4417	22.4794	**MF801999**	**MF802101**	**CE22425**
SM66	SMNH162847	*H. parva*	A	SE: Uppland	59.7584	17.6461	**MF802006**	**MF802117**	**MF802213**
CE19219	ZMBN110086	*H. parva*	A	NO: Sogn og Fjordane	60.8308	7.1196	**MF801993**	**MF802086**	**MF802196**
CE19220	ZMBN110087	*H. parva*	A	NO: Sogn og Fjordane	60.8308	7.1196	**MF801994**	**MF802087**	**MF802197**
CE20554	No voucher	*H. parva*	A	NO: Akershus	59.5645	10.6511	**MF801995**	**MF802094**	**MF802199**
CE20563	No voucher	*H. parva*	A	NO: Akershus	59.5645	10.6511	**MF801996**	**MF802097**	**MF802200**
CE23116	ZMBN110966	*H. parva*	A	NO: Troms	69.7463	21.0604	**MF802000**	**CE23116**	**MF802204**
CE23232	ZMBN110996	*H. parva*	A	NO: Troms	69.2244	19.5047	**MF802001**	**CE23232**	**MF802205**
CE23235	ZMBN110997	*H. parva*	A	NO: Troms	69.2244	19.5047	**MF802002**	**CE23235**	**MF802206**
CE23632	ZMBN111071	*H. parva*	A	NO: Nordland	68.4676	15.8931	**MF802003**	**MF802108**	**MF802207**
CE23633	No voucher	*H. parva*	A	NO: Nordland	68.4676	15.8931	**MF802004**	**MF802109**	**MF802208**
CE24916	ZMBN111358	*H. parva*	A	NO: Nordland	67.2352	14.6142	**MF802005**	**MF802114**	**MF802209**
CE19722	ZMBN110253	*H. parva*	B	NO: Sør-Trøndelag	62.606	11.6599	**MF802007**	**MF802089**	**MF802182**
CE22545	ZMBN110893	*H. parva*	B	NO: Finnmark	69.1938	23.5756	**MF802016**	**MF802102**	**MF802190**
CE24459	ZMBN111280	*H. parva*	B	NO: Nordland	67.2784	14.4343	**MF802021**	**MF802112**	**MF802198**
CE24299	ZMBN111246	*H. parva*	B	NO: Nordland	66.3182	14.1655	**MF802020**	**MF802111**	**MF802194**
CE19496	ZMBN110169	*H. parva*	C	NO: Møre og Romsdal	62.7558	7.2659	**MF802008**	**MF802088**	**MF802181**
CE19723	ZMBN110256	*H. parva*	C	NO: Sør-Trøndelag	62.606	11.6599	**MF802009**	**MF802090**	**MF802183**
CE19754	ZMBN110266	*H. parva*	C	NO: Sør-Trøndelag	62.5763	11.3745	**MF802010**	**MF802091**	**MF802184**
CE19755	ZMBN110267	*H. parva*	C	NO: Sør-Trøndelag	62.5763	11.3745	**MF802011**	**MF802092**	**MF802185**
CE19756	ZMBN110268	*H. parva*	C	NO: Sør-Trøndelag	62.5763	11.3745	**MF802012**	**MF802093**	**MF802186**
CE20559	ZMBN110528	*H. parva*	C	NO: Akershus	59.5645	10.6511	**MF802013**	**MF802095**	**MF802187**
CE20562	ZMBN110529	*H. parva*	C	NO: Akershus	59.5645	10.6511	**MF802014**	**MF802096**	**MF802188**
CE22224	ZMBN110803	*H. parva*	C	NO: Hordaland	60.3633	6.743	**MF802015**	**MF802098**	**MF802189**
CE23068	No voucher	*H. parva*	C	NO: Finnmark	69.9496	23.3003	**MF802018**	**MF802103**	**MF802191**
CE23220	ZMBN110992	*H. parva*	C	NO: Troms	69.5421	20.5499	**MF802017**	**MF802105**	**MF802192**
CE23653	ZMBN111080	*H. parva*	C	NO: Nordland	68.5683	14.963	**MF802019**	**MF802110**	**MF802193**
CE24882	ZMBN111348	*H. parva*	C	NO: Nordland	67.7232	15.8982	**MF802022**	**MF802113**	**MF802195**
511a	–	*H. parva*	B	HUN			KM591923[§]	KM591931[§]	KM591940[§]
615	–	*H. parva*	D	HUN			KM591925[§]	KM591929[§]	KM591941[§]
507	–	*H. bivesiculata*	E	HUN			KM591922[§]	KM591927[§]	KM591933[§]
508b	–	*H. bivesiculata*	E	HUN			KM591921[§]	KM591928[§]	KM591935[§]

Table 1. Continued

Specimen no	Museum voucher no	Species	*COI* cluster	Locality (country, state/province)	Coordinate		GenBank accession number		
					Latitude	Longitude	*COI*	*H3*	ITS
CE841	No voucher	*Stercutus niveus*	A	SE: Västergötland	57.779	12.286	GU902112*	KF672507†	KF672547†
CE6486	SMNH162848	*S. niveus*	A	SE: Uppland	59.4967	18.2732	**MF802028**	**MF802121**	**MF802227**
SM54	SMNH162849	*S. niveus*	A	SE: Uppland	59.7584	17.6461	**MF802023**	**MF802139**	**MF802221**
SM55	SMNH162850	*S. niveus*	A	SE: Uppland	59.7584	17.6461	**MF802024**	**MF802140**	**MF802222**
SM56	SMNH162851	*S. niveus*	A	SE: Uppland	59.7584	17.6461	**MF802025**	**MF802141**	**MF802223**
SM60	SMNH162852	*S. niveus*	A	SE: Uppland	59.7584	17.6461	**MF802026**	**MF802148**	**MF802224**
SM62	No voucher	*S. niveus*	A	SE: Uppland	59.7584	17.6461	**MF802027**	**MF802142**	**MF802225**
CE11433	SMNH162853	*S. niveus*	A	SE: Öland	56.5444	16.6095	**MF802029**	**MF802123**	**MF802228**
CE20498	ZMBN110501	*S. niveus*	A	NO: Akershus	59.5521	10.6912	**MF802034**	**MF802131**	**MF802226**
CE22393	ZMBN110849	*S. niveus*	A	NO: Oppland	61.1488	8.7387	**MF802035**	**MF802132**	**MF802214**
CE22394	ZMBN110850	*S. niveus*	A	NO: Oppland	61.1488	8.7387	**MF802036**	**MF802133**	**MF802215**
CE23845	ZMBN111115	*S. niveus*	A	NO: Akershus	60.3139	11.1799	**MF802032**	**MF802144**	**MF802218**
CE23846	ZMBN111116	*S. niveus*	A	NO: Akershus	60.3139	11.1799	**MF802033**	**MF802145**	**MF802219**
CE23075	ZMBN110960	*S. niveus*	H	NO: Finnmark	69.9493	23.299	**MF802030**	**MF802135**	**MF802216**
CE23080	ZMBN110963	*S. niveus*	H	NO: Finnmark	69.9493	23.299	**MF802031**	**MF802136**	**MF802217**
CE848	SMNH162854	*S. niveus*	B	SE: Västergötland	57.768	12.255	**MF802037**	**MF802119**	**MF802230**
CE13854	SMNH162855	*S. niveus*	B	SE: Västergötland	57.768	12.255	**MF802038**	**MF802124**	**MF802229**
CE27637	SMNH162856	*S. niveus*	B	SE: Dalsland	58.9387	12.4975	**MF802039**	**MF802137**	**MF802220**
CE5803	SMNH162857	*S. niveus*	C	SE: Västergötland	57.6830	11.9542	**MF802045**	**MF802120**	**MF802234**
CE6647	SMNH162858	*S. niveus*	C	SE: Södermanland	58.6117	16.7598	**MF802041**	**MF802122**	**MF802235**
CE20421	ZMBN110489	*S. niveus*	C	NO: Akershus	59.5262	10.69	**MF802042**	**MF802130**	**MF802233**
CE24263	ZMBN111233	*S. niveus*	I	NO: Nordland	66.3181	14.1658	**MF802040**	**MF802146**	**MF802231**
CE24264	ZMBN111234	*S. niveus*	I	NO: Nordland	66.3181	14.1658	**MF802043**	**MF802147**	**MF802232**
CE17826	SMNH162859	*S. niveus*	E	SE: Västergötland	58.069	12.689	**MF802046**	**MF802125**	**MF802237**
CE17827	SMNH162860	*S. niveus*	E	SE: Västergötland	58.069	12.689	**MF802047**	**MF802126**	**MF802238**
CE17828	SMNH162861	*S. niveus*	E	SE: Västergötland	58.069	12.689	**MF802048**	**MF802127**	**MF802239**
CE17829	SMNH162862	*S. niveus*	E	SE: Västergötland	58.069	12.689	**MF802049**	**MF802128**	**MF802240**
CE17830	SMNH162863	*S. niveus*	E	SE: Västergötland	58.069	12.689	**MF802050**	**MF802118**	**MF802241**
CE17831	SMNH162864	*S. niveus*	E	SE: Västergötland	58.069	12.689	**MF802051**	**MF802143**	**MF802242**
CE18648	No voucher	*S. niveus*	D	SE: Skåne	56.0379	13.2524	**MF802044**	**MF802129**	**MF802236**
CE23074	ZMBN110959	*S. niveus*	F	NO: Finnmark	69.9493	23.299	**MF802052**	**MF802134**	**MF802243**
CE3735	SMNH162865	*S. niveus*	G	US: Illinois	36.1531	−86.2725	**MF802053**	**MF802138**	**MF802244**

Accession numbers in bold are newly generated in this study.
*From Erséus *et al.* (2010).
†From Martinsson & Erséus (2014).
‡From Martinsson *et al.* (2017a).
§From Dózsa-Farkas & Felföldi (2015).

Table 2. Details of alignments

	Alignment length	Number of variable positions
Globulidrilus		
COI	658	221
H3	328	15
ITS	949	56
Hemifridericia		
COI	658	151
H3	328	8
ITS	877	39
Stercutus		
COI	625	179
H3	320	11
ITS	949	36

The variation in length in *COI* and *H3* between genera is due to trimming of the ends of the alignments to minimize the amount of missing data; this is also true for ITS, but this marker also varies in length between taxa.

Species discovery

The three *COI* alignments were used to divide the specimens into clusters representing hypothetical species. The *COI* alignments were analysed in the web version of ABGD (Automatic Barcode Gap Discovery) (Puillandre *et al.*, 2012) (available at http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html), using simple distances, P_{min} = 0.001, P_{max} = 0.10 and X (relative gap width) = 0.5. The highest numbers of clusters from the initial partition were selected as input for the species validation step, as BBP (see Species validation) only can lump groups into more inclusive species, not split them into smaller groups; therefore, we rather risk starting with too many species than with few input species. However, we compare the result with the clusters given using a maximum intraspecific *p*-distance of 7%, this value is chosen, based on the genetic distances in the data sets, to represent a more reasonable threshold between intraspecific and interspecific genetic distances. However, as for most threshold values, it is somewhat arbitrarily chosen. Uncorrected intra-cluster and inter-cluster *p*-distances were calculated in MEGA v6.06 (Tamura *et al.*, 2013), missing data and gap was excluded using pairwise deletions, for between-cluster comparisons the minimum distances were used and for within-cluster comparison the maximum distances were used. The clusters were visualized on *COI* gene trees estimated with maximum likelihood in PhyML (Anisimova & Gascuel, 2006; Guindon *et al.*, 2010) as implemented at the South of France Bioinformatics platform (http://www.atgc-montpellier.fr/). The automatic model selection using Smart Model Selection with Bayesian information criterion as the selection criterion was chosen; Subtree Pruning and Regrafting was used for tree improvement. Branch support was calculated with the chi-square-based approximate likelihood ratio test (Anisimova & Gascuel, 2006) in PhyML. The same settings were used for all three *COI* analyses. The trees (Fig. 1) were drawn in FigTree 1.4.2 (Rambaut, 2014) and further edited in Adobe Illustrator. The variation in *H3* and ITS was visualized by haplotype networks created in PopART v1 (Leigh & Bryant, 2015) using statistical parsimony (Templeton, Crandall & Sing, 1992; Clement *et al.*, 2002); sites with missing data or gaps were masked and not included in the networks (Fig. 2).

Species validation

The two nuclear markers (*H3* and ITS) were used for the validation step and were analysed using the BPP v.3.3 program (Yang, 2015). Including the *COI* data in the validation phase would make the result more robust. However, doing so would allow the *COI* data to take over, as the *COI* data set matches the groups found in the discovery phase perfectly. There is a risk that the reasoning would be circular when the same data are used to find and validate species; therefore, we perform the BPP analyses with the nuclear data only.

Joint Bayesian species delimitations and species tree estimations were conducted, a method using the multispecies coalescent model to compare different arrangements of species delimitation and species phylogeny in a Bayesian framework, accounting for incomplete lineage sorting due to ancestral polymorphism and gene tree-species tree conflicts (Yang & Rannala, 2010; Rannala & Yang, 2013; Yang & Rannala, 2014). Three analyses were run for each genus, varying the population size (θs) and divergence time (τ0) priors. The priors in the different analyses, respectively, were the same for all genera:

In analysis A, the population size parameters (θs) were assigned the gamma prior G(2, 400), with mean 2/400 = 0.005. The divergence time at the root of the species tree (τ0) was assigned the gamma prior G(2, 200), while the other divergence time parameters were assigned the Dirichlet prior (Yang & Rannala, 2010: equation 2).

In analysis B, the population size parameters (θs) were assigned the gamma prior G(2, 1000), with mean 2/1000 = 0.002. The divergence time at the root of the species tree (τ0) was assigned the gamma prior G(2, 200), while the other divergence time parameters were

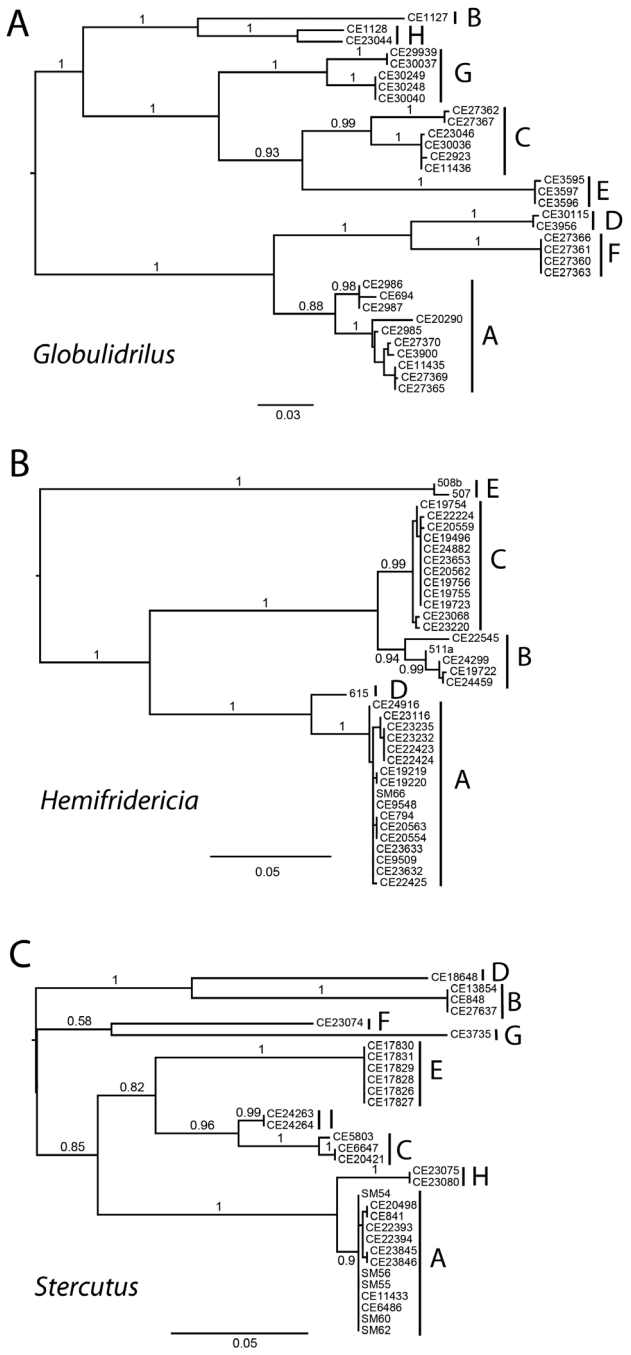

Figure 1. *COI* gene trees, showing clusters from ABGD analyses. All trees are rooted using midpoint rooting. A, *Globulidrilus*, all clusters represent *G. riparius s.l.* B, *Hemifridericia*, cluster E corresponds with *H. bivesiculata*, the other clusters represent *H. parva s.l.* C, *Stercutus*, all clusters represent *S. niveus s.l.* Numbers at branches denote branch support by the approximate likelihood ratio test. Scale bars show estimated numbers of nucleotide substitutions per site.

assigned the Dirichlet prior (Yang & Rannala, 2010: equation 2).

In analysis C, the population size parameters (θs) were assigned the gamma prior G(2, 2000), with mean 2/2000 = 0.001. The divergence time at the root of the species tree (τ0) was assigned the gamma prior G(2, 200), while the other divergence time parameters were assigned the Dirichlet prior (Yang & Rannala, 2010: equation 2).

Each analysis was run three times to confirm consistency between runs. For the species arrangements, the 95% credibility intervals, that is the number of arrangements with the highest PP that together makes up 95% of the PP, were calculated. With regard to the individual species, we considered species delimited with a PP > 0.90 in all analyses to be well supported. For clusters with a PP < 0.90, we used a conservative approach and instead accepted the best-supported more inclusive species. The PP values for the species were mapped on the species tree with the highest PP in the majority of analyses.

Genetic uncorrected *p*-distances were calculated for *H3* and ITS the same way as for *COI* (see Species discovery), but sorted on the delimited species instead of the initial clusters.

RESULTS

SPECIES DISCOVERY

The ABGD analyses of *COI* divided the *Globulidrilus* data set into eight clusters, the *Hemifridericia* data set into five clusters and the *Stercutus* data set into nine clusters (Fig. 1A–C) using the maximum number of clusters. Using the 7% threshold, the *Globulidrilus* data set is divided into two groups, combining clusters A, D and F into one group and clusters B, C, E, G and H into another group; the *Hemifridericia* data set was divided into three groups, combining groups A and D into one group and groups B and C into another group, and group E still being a separate group. The *Stercutus* data set was divided into seven groups, clusters A and H were combined into one group and C and I into another group, remaining clusters were unchanged.

The uncorrected *p*-distances presented here are the maximum intra-cluster distances and the minimum intra-cluster distances for each comparison; the values are summarized in Supporting Information, Tables S2–S4. For the *Globulidrilus* clusters (Supporting Information, Table S2), the largest intra-cluster *p*-distance was 6.1% (within cluster C), the smallest inter-cluster *p*-distance was 8.5% (between clusters D and F)

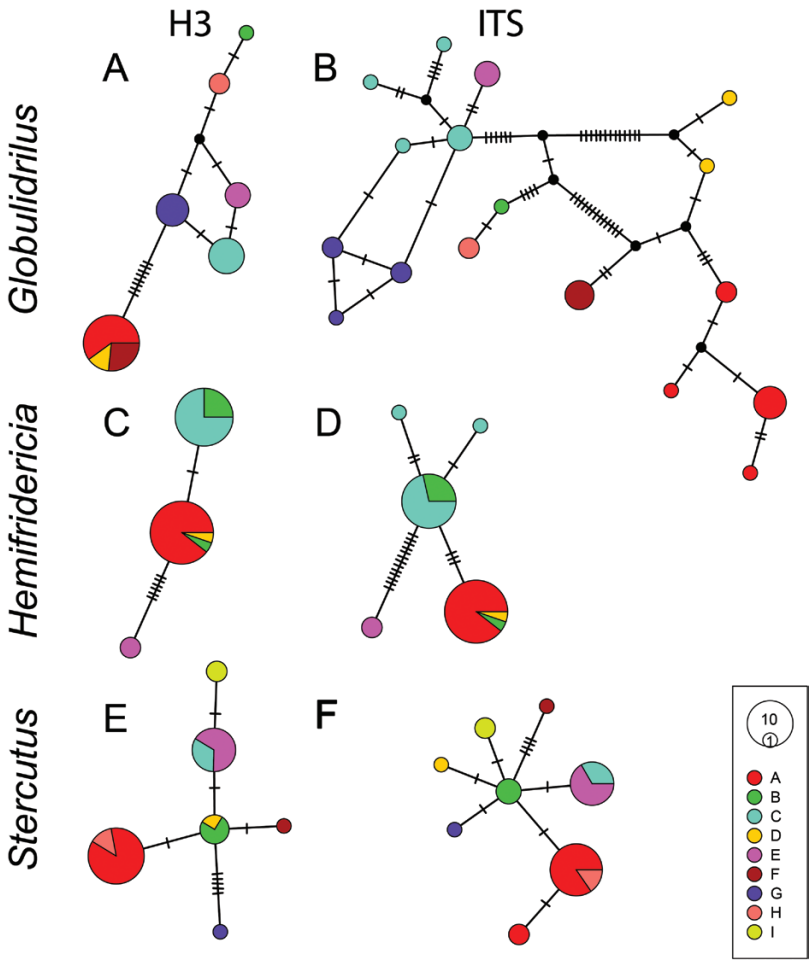

Figure 2. Statistical parsimony haplotype networks. A, B, *Globulidrilus*. C, D, *Hemifridericia*. E, F, *Stercutus*. A, C, D, *H3* networks. B, D, F, ITS networks. The size of the circles is relative to the number of sequences sharing that haplotype, the colours correspond to *COI* clusters and the hatch marks correspond to substitutions.

and the largest inter-cluster *p*-distance was 18.3% (between clusters D and E). For the *Hemifridericia* data set (Supporting Information, Table S3), the largest intra-cluster *p*-distance was 3.0% (within cluster B), the smallest inter-cluster *p*-distance was 3.0% (between clusters B and C) and the largest inter-cluster *p*-distance was 15.9% (between clusters B and E). For the *Stercutus* data set (Supporting Information, Table S4), the largest intra-cluster *p*-distance was 1.0% (within clusters C), the smallest inter-cluster *p*-distance was 3.2% (between clusters A and H) and the largest inter-cluster *p*-distance was 15.7% (between clusters B and G). The haplotype networks (Fig. 2) showed that the specimens of most *COI* clusters have unique haplotypes in both their nuclear genes (*H3* and ITS), not shared with other clusters, and that the *H3*/ITS haplotypes from the same *COI*-based cluster are close to each other. In a few cases, however, the *H3*/ITS haplotypes are shared by two or three clusters. Moreover, one specimen (CE24459) of *H. parva* cluster B is not even found together with the other cluster B specimens, neither in *H3* nor in ITS; instead, it shares its nuclear haplotypes with specimens from clusters A and D (Fig. 2C, D).

Species validation

The 95% credibility intervals from the BPP analyses of *Globulidrilus* contain three species delimitation arrangements in analysis A and two arrangements in analyses B and C (Table 3). Either a seven-species or an eight-species arrangement was preferred, the best seven-species arrangement had a mean PP for the three runs of each analysis of 0.58, 0.50 or 0.39 in analyses A, B and C, respectively, whereas the eight-species arrangement had a mean PP of 0.35, 0.47 or 0.60 in analyses A, B and C, respectively (Table 3). The difference between the seven-species and eight-species

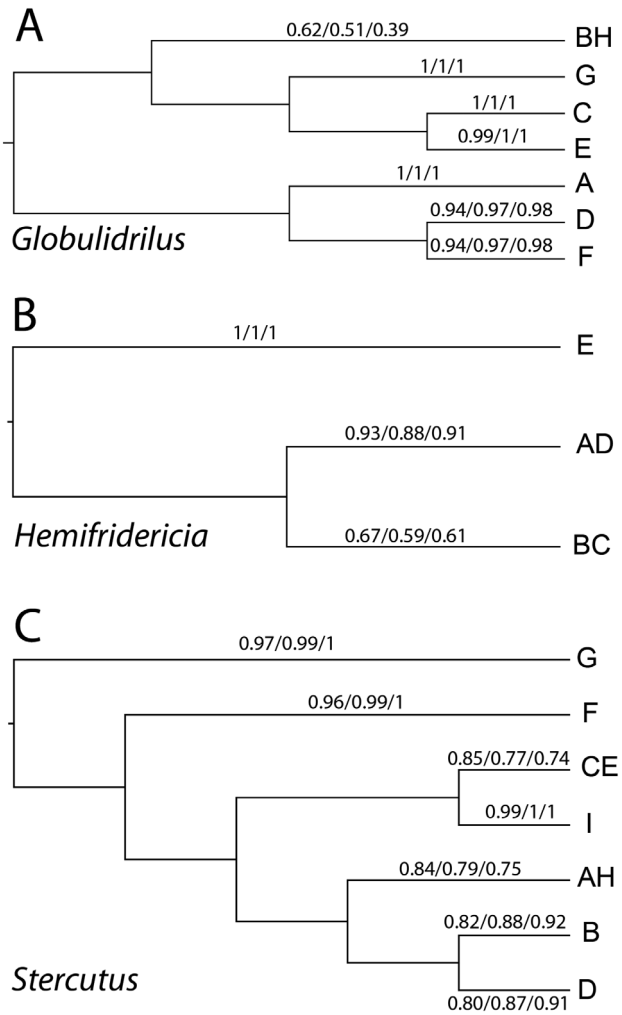

Figure 3. Species trees from BPP analyses, the trees shown are the species trees and species delimitations with the highest PP in a majority of analyses. A, *Globulidrilus*. B, *Hemifridericia*. C, *Stercutus*. Numbers at branches are mean PP for species in analyses A, B and C, respectively.

arrangements pertains to whether clusters B and H are united into one species or treated as two separate species. All clusters (putative species) except B and H were delimited with a PP > 0.90 in all analyses (Table 4).

The 95% credibility intervals from the BPP analyses of *Hemifridericia* contain three different species delimitation arrangements in analysis A and four arrangements in analyses B and C (Table 3). In all analyses, a three-species arrangement, uniting clusters A and D as well as B and C, was preferred with a mean PP of 0.64, 0.56 and 0.57 in analyses A, B and C, respectively (Table 3). Cluster E, that is *H. bivesiculata*, was delimited with a PP of 1 in all analyses, species A + D was delimited with a PP > 0.90 in analyses A and C and with a PP of 0.88 in analyses B and species B + C had a PP of 0.67, 0.59 and 0.61 in analyses A, B and C, respectively (Fig. 3B; Table 4).

The 95% credibility intervals from the BPP analyses of *Stercutus* contain nine different species delimitation arrangements in analysis A, six arrangements in analyses B and five arrangements in analysis C (Table 3). In all analyses, a seven-species arrangement, uniting clusters A and H as well as clusters C and E, was preferred with a mean PP of 0.55, 0.52 or 0.54 in analyses A, B and C, respectively (Table 3). Clusters F, G and I were delimited with a PP > 0.95 in all analyses, and clusters B and D were delimited with a PP > 0.90 in analysis C; species C + E had a PP of 0.85, 0.77 and 0.74 in analyses A, B and C, respectively, and species A + H had a PP of 0.84, 0.79 and 0.75 in analyses A, B and C, respectively (Fig. 3C; Table 4).

For *Globulidrilus*, the genetic distances within the delimited species in *H3* are maximally 0.3% (in BH) and in ITS are 0.6% (in C); the minimum distances between species vary in *H3* between 0.0 (between A, D and F) and 3.5% (between A and BH, C and F) and in ITS between 0.1 (between C and G) and 2.8% (between C and F) (Supporting Information, Table S5). For *Hemifridericia*, the genetic distances within the delimited species in *H3* are maximally 0.3% (in BC) and in ITS are 0.8% (in BC); the minimum distances between species vary in *H3* between 0.0 (between AD and BC) and 2.1% (between E and the other species) and in ITS between 0.0 (between AD and BC) and 2.4% (between AD and E) (Supporting Information, Table S6). If specimen CE24459, which is found together with different groups in the *COI* and the nuclear data sets, would be excluded, there would be a difference between species AD and BC in both data sets (0.3% in *H3* and 0.8% in ITS). For *Stercutus*, the maximum genetic distances within the delimited species in *H3* are 0.3% (in BD) and in ITS are 0.7% (in BD); the minimum distances between species vary in *H3* between 0.3 (between BD and AH, CE and F, and between CD and I) and 2.6% (between G and I) and in ITS between 0.3 (between AH and CE) and 1.5% (between BD and G) (Supporting Information, Table S7).

To summarize, the results of the BPP analyses suggest that our data set includes at least seven species of *Globulidrilus*, for the time being all attributed to the nominal species *G. riparius*; three species of *Hemifridericia*, whereof two are referred to as *H. parva* and the third as *H. bivesiculata*; and at least six species of *Stercutus*.

DISCUSSION

We have found that the four nominal taxa that we included in the analyses actually represent at least 16

Table 3. List of species delimitations and their mean posterior probability

	Species delimitations	PP analysis A	PP analysis B	PP analysis C
Globulidrilus	7 (A, BH, C, D, E, F, G)	**0.58**	**0.50**	**0.38**
	8 (A, B, C, D, E, F, G, H)	**0.35**	**0.47**	**0.60**
	6 (A, BH, C, DF, E, G)	**0.04**	0.02	0.01
	7 (A, B, C, DF, E, G, H)	0.02	0.01	0.01
Hemifridericia	3 (AD, BC, E)	**0.64**	**0.56**	**0.57**
	4 (AD, B, C, E)	**0.30**	**0.32**	**0.34**
	4 (A, BC, D, E)	**0.03**	**0.03**	**0.03**
	5 (A, B, C, D, E)	0.02	0.03	**0.02**
	4 (A, BD, C, E)	0.01	0.02	0.02
	3 (A, BCD, E)	0.01	**0.04**	0.01
Stercutus	7 (AH B CE D F G I)	**0.55**	**0.52**	**0.51**
	6 (AH BD CE F G I)	**0.12**	**0.07**	**0.05**
	8 (A B CE D F G H I)	**0.11**	**0.15**	**0.17**
	8 (AH B C D E F G I)	**0.10**	**0.16**	**0.18**
	7 (AH BD C E F G I)	**0.02**	**0.02**	0.02
	7 (A BD CE F G H I)	**0.02**	0.02	0.02
	9 (A B C D E F G H I)	**0.02**	**0.04**	**0.06**
	6 (AH B CE DF G I)	**0.01**	0.00	0.00
	6 (AH B CE DG F I)	**0.01**	0.00	0.00
	6 (AH B CE D FG I)	0.01	0.00	0.00

For each analysis and genus, the PP values in bold make up the 95% credibility interval.

species, with varying support. It is reasonable to question whether all these recognized units are species or if they represent something else. The multispecies coalescent approach has been criticized for delimiting population structure and not real species (Sukumaran & Knowles, 2017). However, the latter authors do not specify what they mean by 'species', and as so many species concepts have been put forward in the past (see e.g. De Queiroz, 2007), it is hard to understand what they mean with their claims. One of the strengths of the multispecies coalescent model is that it is based on a well-defined concept of species being lineages that no longer exchange genes following divergence from a common ancestor (Toprak *et al.*, 2016). It is possible that some of these lineages, in the future, will merge and hybridize to such extent that they will despeciate (as described by Turner, 2002). It is also possible that addition of more data, either in the form of more specimens or additional markers, will lead to another result than the one we have found in this study. Species assignments are hypotheses and as such may change over time, but, for the time being, the results of this study are the best hypotheses that we have for the species boundaries in our three genera.

The high number of species arrangements within the 95% credibility intervals, especially for *Stercutus* (see Table 2), as well as the low support for some of the delimited species (Table 3) are probably due to a lack of phylogenetic signal in our nuclear data, and it is possible that including more variable loci would give better support for some species. However, in other studies, BPP has successfully been used with an amount of data similar to what we have herein (e.g. Hambäck *et al.*, 2013; Parmakelis *et al.*, 2013; Fossen *et al.*, 2016; Martinsson, Rhodén & Erséus, 2017b). When using a higher, fixed threshold (7%) in the species discovery part the clusters changes, for *Hemifridericia*, the resulting clusters were the same as the delimited species, whereas for the other genera, it gave a different result, for *Globulidrilus*, a 7% threshold only gave two clusters, and interestingly, these two groups can be seen in the nuclear data as well, where there is a larger distance between groups ADF and BCEH than within the groups, nevertheless the BPP analyses gave strong support for at least seven species. For *Stercutus*, a 7% threshold gave mainly the same result as using the highest number of cluster, with the exception that it united clusters A and H as well as clusters C and I, and AD was found to be one species in the BPP analyses, but there is no support for combining C and I, instead C and E are found in the same species. This highlights the problem of using strict thresholds for single loci in species delimitation, even if it still can be a useful rule of thumb.

Of the nominal taxa in focus in this study, *G. riparius* has already been suggested to be a species complex based on morphological observations (Christensen & Dózsa-Farkas, 2012), and for *S. niveus*, morphological variation has been observed as well (Rota, 1995). In

Table 4. List of delimited species and their mean posterior probabilities

	Species	PP analysis A	PP analysis B	PP analysis C
Globulidrilus	**G**	**1.00**	**1.00**	**1.00**
	C	**1.00**	**1.00**	**1.00**
	E	**1.00**	**1.00**	**1.00**
	A	**0.99**	**1.00**	**1.00**
	F	**0.94**	**0.97**	**0.98**
	D	**0.94**	**0.97**	**0.98**
	BH	0.62	0.51	0.39
	H	0.38	0.49	0.61
	B	0.38	0.49	0.61
	DF	0.06	0.03	0.02
Hemifridericia	**E**	**1.00**	**1.00**	**1.00**
	AD	**0.93**	0.88	**0.91**
	BC	0.67	0.59	0.61
	C	0.32	0.37	0.38
	B	0.32	0.35	0.36
	A	0.07	0.12	0.09
	D	0.05	0.06	0.06
	BD	0.01	0.02	0.02
	BCD	0.01	0.04	0.01
	CD	0.00	0.00	0.00
Stercutus	**I**	**0.99**	**1.00**	**1.00**
	G	**0.97**	**0.99**	**1.00**
	F	**0.96**	**0.99**	**1.00**
	CE	0.85	0.77	0.74
	AH	0.84	0.79	0.75
	B	0.82	0.88	**0.92**
	D	0.80	0.87	**0.91**
	A	0.16	0.21	0.25
	H	0.16	0.21	0.25
	BD	0.17	0.12	0.08
	E	0.15	0.23	0.26
	C	0.15	0.23	0.26
	DF	0.02	0.01	0.00
	FG	0.01	0.00	0.00
	DG	0.01	0.00	0.00

PP > 0.90 are marked in bold. Accepted species are also marked in bold; note, for example, that although the two *Stercutus* clusters B and D have PP values between 0.80 and 0.92 in all analyses, the more conservative alternative of merging them into a single species (BD) is preferred.

H. parva, there is no notion of morphological variation, at least not in Europe. However, a description of this alleged species from China (Wang, Xie & Liang, 1999) differs from the description of material from Europe (e.g. Nielsen & Christensen, 1959) in detail about the septal glands; it could very well represent a different species. The odd specimen of *H. parva* (CE24459), which was found together with specimens of different clusters in the mitochondrial (cluster B) and nuclear data (in clusters A and D) sets, could be a case of hybridization and mitochondrial introgression between the two *H. parva* species. Hybridization has been reported between cryptic earthworm species (Dupont *et al.*, 2016; Martinsson & Erséus, 2017). At the locality where the specimen corresponding to CE24459 was found, there was also a specimen of cluster A (C.E., unpubl. data). In the material used in our present work, we had co-occurrence of more than one cluster at the same location in all species (see Table 1). In *Globulidrilus*, clusters A, C and F, clusters A and D, clusters B and H and clusters C and G are found together. In *Hemifridericia*, A and C are found together, and in *Stercutus*, F and H are found together. Despite this co-occurrence, only one case of mismatch between mitochondrial and nuclear markers is observed, indicating that reproductive barriers between the delimited species exist. It is worth noting that our two lineages of *H. parva* are found in Hungary

as well as in Scandinavia and they are probably well spread in Europe.

To summarize, we found high genetic diversity in the studied enchytraeid groups, compared to the previously recognized four species. We suggest that these three genera comprise at least 16 species. The delimited species are not formally revised, described and named here, mainly due to the combination of low support for some of them and lack of mature specimens of many of them, which makes morphological comparisons and good-quality descriptions difficult. However, it is likely that some of these species are truly cryptic and that it will be impossible to identify them based on morphology. Nevertheless, this study is a first step towards understanding the species diversity of these groups, and hopefully, more material suitable for morphological studies will be collected in the future.

ACKNOWLEDGEMENTS

We are grateful to Leyla Arsan, Stephen Atkinson, Mårten Klinth, Tryggve Persson, Mark Wetzel and Magdalena Zarowiecki for providing specimens and to Anna Ansebo, Marcus Carlberg, Daniel Gustafsson, Per Hjelmstedt, Jeff Hunt, Mårten Klinth, Emelie Lindquist, Maria Lindström and Urban Olsson for laboratory assistance. We thank Kerryn Elliott and two anonymous reviewers for valuable comments, which helped to improve the manuscript. The study was supported by the Royal Society of Arts and Sciences in Gothenburg, the Adlerbert Foundation to S.M. and C.E. and the Swedish and Norwegian Taxonomy Initiatives (ArtDatabanken, Uppsala, and Artsdatabanken, Trondheim) to C.E.

REFERENCES

Anisimova M, Gascuel O. 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Systematic Biology* **55:** 539–552.

Aydin Z, Marcussen T, Ertekin AS, Oxelman B. 2014. Marginal likelihood estimate comparisons to obtain optimal species delimitations in *Silene* sect. Cryptoneurae (Caryophyllaceae). *PLoS ONE* **9:** e106990.

Bickford D, Lohman DJ, Sodhi NS, Ng PK, Meier R, Winker K, Ingram KK, Das I. 2007. Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution* **22:** 148–155.

Carstens BC, Pelletier TA, Reid NM, Satler JD. 2013. How to fail at species delimitation. *Molecular Ecology* **22:** 4369–4383.

Christensen B, Dózsa-Farkas K. 2012. A new genus *Globulidrilus* and three new enchytraeid species (Oligochaeta: Enchytraeidae) from Seoraksan National Park (Korea). *Journal of Natural History* **46:** 2769–2785.

Clement M, Snell Q, Walke P, Posada D, Crandall K. 2002. TCS: estimating gene genealogies. *Proceedings of the 16th International Parallel & Distributed Processing Symposium* **2:** 184.

De Queiroz K. 2007. Species concepts and species delimitation. *Systematic Biology* **56:** 879–886.

Donnelly RK, Harper GL, Morgan AJ, Orozco-Terwengel P, Pinto-Juma GA, Bruford MW. 2013. Nuclear DNA recapitulates the cryptic mitochondrial lineages of *Lumbricus rubellus* and suggests the existence of cryptic species in an ecotoxicological soil sentinel. *Biological Journal of the Linnean Society* **110:** 780–795.

Dózsa-Farkas K, Felföldi T. 2015. Unexpected occurrence of *Hemifridericia bivesiculata* Christensen & Dózsa-Farkas, 2006 in Hungary, a species presumed to be endemic to Devon Island, Canada, and its comparative analysis with *H. parva* Nielsen & Christensen, 1959 (Enchytraeidae, Oligochaeta). *Zootaxa* **3914:** 185–194.

Dupont L, Porco D, Symondson WO, Roy V. 2016. Hybridization relics complicate barcode-based identification of species in earthworms. *Molecular Ecology Resources* **16:** 883–894.

Dupuis JR, Roe AD, Sperling FA. 2012. Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. *Molecular Ecology* **21:** 4422–4436.

Erséus C, Gustafsson D. 2009. Cryptic speciation in clitellate model organisms. In: Shain DH, ed. *Annelids in modern biology*. Hoboken: John Wiley & Sons, Inc, 31–46.

Erséus C, Rota E, Matamoros L, De Wit P. 2010. Molecular phylogeny of Enchytraeidae (Annelida, Clitellata). *Molecular Phylogenetics and Evolution* **57:** 849–858.

Fossen EI, Ekrem T, Nilsson AN, Bergsten J. 2016. Species delimitation in northern European water scavenger beetles of the genus *Hydrobius* (Coleoptera, Hydrophilidae). *Zookeys* **564**: 71–120.

Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C. 2012. Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution* **27:** 480–488.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59:** 307–321.

Hambäck PA, Weingartner E, Ericson L, Fors L, Cassel-Lundhagen A, Stenberg JA, Bergsten J. 2013. Bayesian species delimitation reveals generalist and specialist parasitic wasps on *Galerucella* beetles (Chrysomelidae): sorting by herbivore or plant host. *BMC Evolutionary Biology* **13:** 92.

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30:** 3059–3066.

Klinth MJ, Martinsson S, Erséus C. 2017. Phylogeny and species delimitation of North European *Lumbricillus* (Clitellata, Enchytraeidae). *Zoologica Scripta* **46:** 96–110.

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

Liu Y, Fend SV, Martinsson S, Erséus C. 2017. Extensive cryptic diversity in the cosmopolitan sludge worm *Limnodrilus hoffmeisteri* (Clitellata, Naididae). *Organisms Diversity & Evolution* **17:** 477–495.

Martinsson S, Dózsa-Farkas K, Rota E, Erséus C. 2017a. Placing the forgotten: on the positions of *Euenchytraeus* and *Chamaedrilus* in an updated enchytraeid phylogeny (Clitellata: Enchytraeidae). *Invertebrate Systematics* **31:** 85–90.

Martinsson S, Erséus C. 2014. Cryptic diversity in the well-studied terrestrial worm *Cognettia sphagnetorum* (Clitellata: Enchytraeidae). *Pedobiologia* **57:** 27–35.

Martinsson S, Erséus C. 2017. Cryptic speciation and limited hybridization within *Lumbricus* earthworms (Clitellata: Lumbricidae). *Molecular Phylogenetics and Evolution* **106:** 18–27.

Martinsson S, Rhodén C, Erséus C. 2017b. Barcoding gap, but no support for cryptic speciation in the earthworm *Aporrectodea longa* (Clitellata: Lumbricidae). *Mitochondrial DNA, Part A: DNA Mapping, Sequencing, and Analysis* **28:** 147–155.

Matamoros L, Rota E, Erséus C. 2012. Cryptic diversity among the achaetous *Marionina* (Annelida, Clitellata, Enchytraeidae). *Systematics and Biodiversity* **10:** 509–525.

Nielsen CO, Christensen B. 1959. The Enchytraeidae: critical revision and taxonomy of European species. *Natura Jutlandica* **8–9:** 1–160.

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

Olave M, Sola E, Knowles LL. 2014. Upstream analyses create problems with DNA-based species delimitation. *Systematic Biology* **63:** 263–271.

Parmakelis A, Kotsakiozi P, Stathi I, Poulikarakou S, Fet V. 2013. Hidden diversity of *Euscorpius* (Scorpiones: Euscorpiidae) in Greece revealed by multilocus species-delimitation approaches. *Biological Journal of the Linnean Society* **110:** 728–748.

Pfenninger M, Schwenk K. 2007. Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evolutionary Biology* **7:** 121.

Puillandre N, Lambert A, Brouillet S, Achaz G. 2012. ABGD, automatic barcode gap discovery for primary species delimitation. *Molecular Ecology* **21:** 1864–1877.

Rambaut A. 2014. FigTree v1.4.2. Available at: http://tree.bio.ed.ac.uk/software/figtree/ (last accessed 10 May 2017).

Rannala B. 2015. The art and science of species delimitation. *Current Zoology* **61:** 846–853.

Rannala B, Yang Z. 2003. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics* **164:** 1645–1656.

Rannala B, Yang Z. 2013. Improved reversible jump algorithms for Bayesian species delimitation. *Genetics* **194:** 245–253.

Rota E. 1995. Italian Enchytraeidae (Oligochaeta). I. *Bolletino di Zoologia* **62:** 183–231.

Sites JW, Marshall JC. 2003. Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology & Evolution* **18:** 462–470.

Sukumaran J, Knowles LL. 2017. Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences of the United States of America* **114:** 1607–1612.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30:** 2725–2729.

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.

Toprak Z, Pfeil BE, Jones G, Marcussen T, Ertekin AS, Oxelman B. 2016. Species delimitation without prior knowledge: DISSECT reveals extensive cryptic speciation in the *Silene aegyptiaca* complex (Caryophyllaceae). *Molecular Phylogenetics and Evolution* **102:** 1–8.

Turner GF. 2002. Parallel speciation, despeciation and respeciation: implications for species definition. *Fish and Fisheries* **3:** 225–229.

Wang H, Xie Z, Liang Y. 1999. Records of enchytraeidae (Clitellata) from the People's Republic of China. *Hydrobiologia* **406:** 57–66.

Yang Z. 2015. The BPP program for species tree estimation and species delimitation. *Current Zoology* **61:** 854–865.

Yang Z, Rannala B. 2010. Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences of the United States of America* **107:** 9264–9269.

Yang Z, Rannala B. 2014. Unguided species delimitation using DNA sequence data from multiple loci. *Molecular Biology and Evolution* **31:** 3125–3135.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Table S1. Primers, sequences and PCR programs used for amplification of the mitochondrial *COI* and nuclear ITS and *H3*.

Table S2. Uncorrected pairwise *COI* distances for *Globulidrilus riparius*, the values shown are the minimum values for the inter-cluster comparisons and the maximum for the intra-cluster comparisons. All distances are expressed as percents.

Table S3. Uncorrected pairwise *COI* distances for *Hemifridericia*, the values shown are the minimum values for the inter-cluster comparisons and the maximum for the intra-cluster comparisons. Clusters A–D represent *H. parva* and cluster E, *H. bivesiculata.* All distances are expressed as percents.
Table S4. Uncorrected pairwise *COI* distances for *Stercutus niveus*, the values shown are the minimum values for the inter-cluster comparisons and the maximum for the intra-cluster comparisons. All distances are expressed as percents.
Table S5. Uncorrected pairwise *H3* and ITS distances for *Globulidrilus riparius*, the values shown are the minimum values for the inter-cluster comparisons and the maximum for the intra-cluster comparisons. All distances are expressed as percents. The *H3* distances are at the lower left side, and the ITS distances are at the upper right side.
Table S6. Uncorrected pairwise *H3* and ITS distances for *Hemifridericia*, the values shown are the minimum values for the inter-cluster comparisons and the maximum for the intra-cluster comparisons. All distances are expressed as percents. The *H3* distances are at the lower left side, and the ITS distances are at the upper right side. Clusters A–D represent *H. parva* and cluster E, *H. bivesiculata.* All distances are expressed as percents.
Table S7. Uncorrected pairwise *H3* and ITS distances for *Stercutus niveus*, the values shown are the minimum values for the inter-cluster comparisons and the maximum for the intra-cluster comparisons. All distances are expressed as percents. The *H3* distances are at the lower left side, and the ITS distances are at the upper right side.