

Contents lists available at [SciVerse ScienceDirect](http://www.sciencedirect.com)

## Molecular Phylogenetics and Evolution

journal homepage: [www.elsevier.com/locate/ympev](http://www.elsevier.com/locate/ympev)

## Reappraisal of the “Molecular phylogeny of Western Palaeartic Helicidae s.l. (Gastropoda: Stylommatophora)”: When poor science meets GenBank

Dick S.J. Groenenberg<sup>a,\*</sup>, Eike Neubert<sup>b</sup>, Edmund Gittenberger<sup>a,c</sup>

<sup>a</sup> Netherlands Centre for Biodiversity Naturalis, P.O. Box 9517, NL 2300 RA Leiden, The Netherlands

<sup>b</sup> Naturhistorisches Museum der Burgergemeinde Bern, Bernastrasse 15, CH-3005 Bern, Switzerland

<sup>c</sup> Institute of Biology, Leiden University, Sylvius Lab., P.O. Box 9505, NL 2300 RA Leiden, The Netherlands

## ARTICLE INFO

## Article history:

Received 20 April 2011

Revised 10 August 2011

Accepted 26 August 2011

Available online xxxx

## Keywords:

Helicidae

GenBank

'Pseudo-voucher'

Misidentification

COI

16S

## ABSTRACT

A paper on the 'Molecular phylogeny of Western Palaeartic Helicidae s.l.', published by Steinke et al. (2004) in this journal, is critically analysed. Several obvious errors are corrected and methodological weaknesses are revealed. BLAST searches on the sequences published in that paper and now in GenBank, showed high percentages of similarity of the alleged species with taxa that are considered only distantly related in the literature. Inspection of the so-called voucher specimens showed that some shells were misidentified, whereas others contained dirt or were bleached, indicating that these had been collected empty. Obviously the sequences published for those species could not have originated from those specimens, which cannot be considered vouchers therefore, even if they are from the same locality. In other instances, spurious sequences were published for correctly identified voucher specimens. For several species for which we collected specimens ourselves, the COI or the 16S sequence, or both, clearly differed from the results published by Steinke et al. The consequences of our results for the molecular data on helioid gastropods and their classification are listed.

© 2011 Elsevier Inc. All rights reserved.

### 1. Introduction

Some years ago a paper which was published in this journal, provided an interesting scope on what was referred to as the 'Molecular phylogeny of the Western Palaeartic Helicidae s.l.' (Steinke et al., 2004). The authors claimed to present the first comprehensive molecular phylogeny of this family. With a skewed taxon sampling (arguably overrepresenting the genera *Candidula* and *Trochoidea*) and only about a fifth of the genera of Helicidae (according to Zilch, 1959) included, this claim seemed somewhat far-fetched. Nevertheless, the paper did include representatives of the subfamilies Helicodontinae, Ariantinae, Helicinae, Hygromiinae and Helicellinae and as such could have provided a scaffold for the molecular phylogeny of this family.

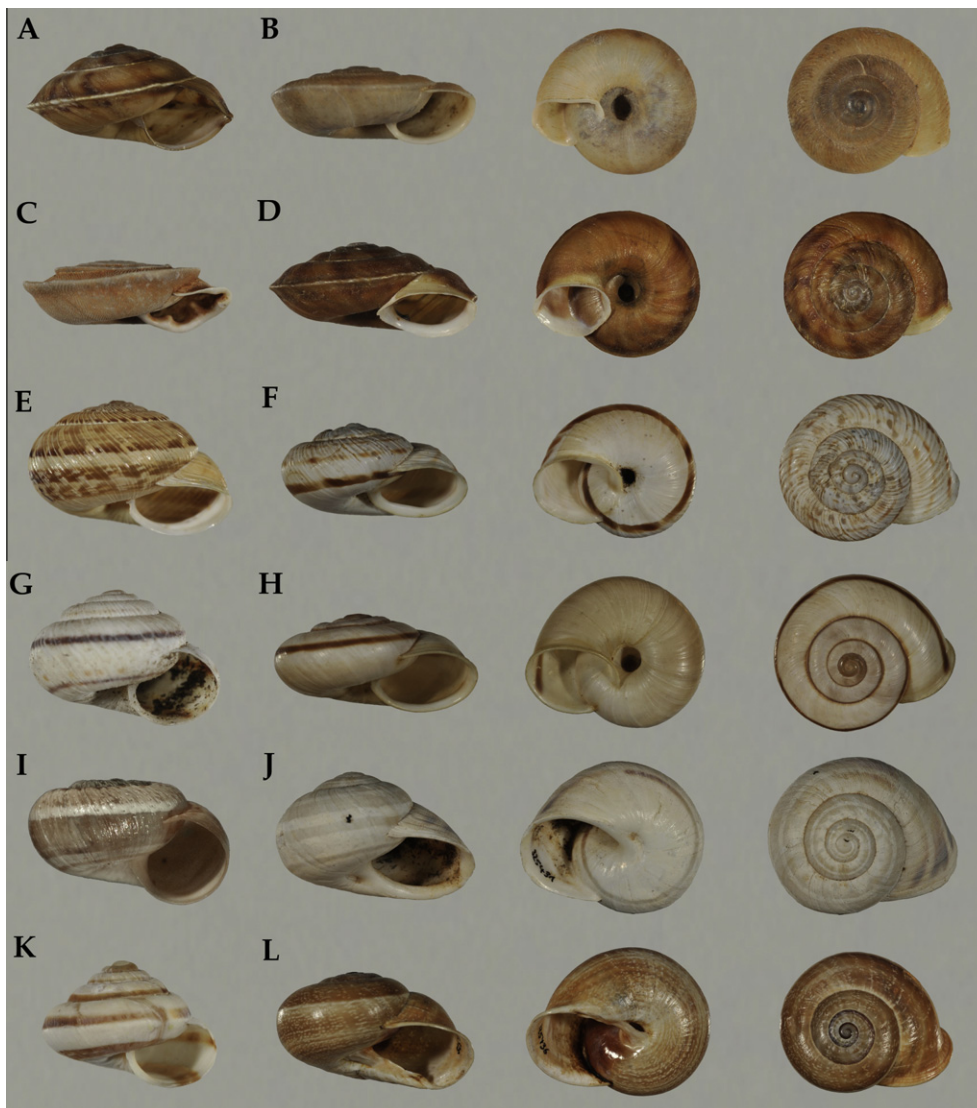
Public databases (like GenBank) provide a great resource nowadays for the comparison of sequence-data. NCBI-BLAST (Basic Local Alignment Search Tool) searches (Zhang et al., 2000), for instance, can confirm species identifications or indicate sources of contamination and incomplete taxon-sampling might be reduced or overcome by adding sequences obtained through GenBank. However, the added value of these open source databases strongly depends on the trustworthiness of the data that have been put into them. Usage of GenBank sequences which have not been safeguarded during each

step in the process from the collection of the specimens all the way to the submission of the resulting sequences is a potential source of error. GenBank contains sequences of which the vouchers were misidentified, or worse, sequences that do not correspond to their supposed voucher specimens. Occasionally, researchers may not be able to unequivocally identify a species, or it can be impossible to preserve a voucher (destructive sampling for DNA extractions). But even when errors in GenBank records are evident, GenBank only allows the primary authors to reannotate those data. Moreover, removal of corrupt sequences might be undesirable, once they have been used in the literature. Therefore it has been suggested to 'wikify' GenBank, in order to allow for the correction of inaccuracies in GenBank records (Pennisi, 2008).

Initial BLAST searches with the COI and 16S sequences of Steinke et al. (2004) in several cases revealed high percentages (>90%) of sequence similarity between taxa that are expected to be only distantly related (e.g. *Oestophora turriplana* AY546289 and *Candidula codia* AY238622, or *Trochoidea pyramidata* AY546297 and *Otala punctata* AY546290). In other instances the sequences of Steinke et al. (2004) differed substantially from the sequences on GenBank for which the vouchers were identified as conspecific (e.g. *Helicigona lapicida* AY546280 and AF297000; *Trochoidea elegans* AY546295 and FJ627176). Based on the remarkable results for some initial BLAST-searches, our own interest in the Ariantinae (Gittenberger et al., 2004) and the peculiar position of *Caracollina lenticula* (Fig. 1 in Steinke et al., 2004), we decided to obtain sequences for a number of species ourselves.

\* Corresponding author. Fax: +31 071 5274900.

E-mail address: [Dick.Groenenberg@ncbnaturalis.nl](mailto:Dick.Groenenberg@ncbnaturalis.nl) (D.S.J. Groenenberg).



**Fig. 1.** Compilation of shells including some of the inspected vouchers of Steinke et al. (2004) and a selection of shells from the conchological collection of NCB Naturalis. A. “*Caracollina lenticula* sensu Steinke et al., 2004” (published as SMF 325411) = *Helicigona lapicida andorrlica* (Bourguignat, 1876), juvenile (without apertural lip) B 15.0 mm [SMF 325448/1] Spain, “Seteases”; M. Pfenninger leg. B. *Caracollina lenticula* (Michaud, 1831) B 7.2 mm [RMNH 100372] Italy, Lampedusa island, Punta Sotille; A.W. Janssen leg., 18.v.2005. C. “*Oestophora turriplana*” = *Gittenbergeria turriplana* (Morelet, 1845) B 14.7 mm [SMF 325435/1] Portugal, “Boliouenne”; M. Pfenninger leg. D. *Helicigona lapicida* (L., 1758) B 17.6 mm [SMF325426/1] Germany, Hessen, Schlüchtern; M. Pfenninger leg. E. *Candidula codia* (Bourguignat, 1859) B 12.0 mm [RMNH G1980] Portugal, Algarve, 5 km NE of Boliouenne; E. Gittenberger leg., iv.1980. F. “*Chilostoma cingulatum* sensu Steinke et al., 2004” = *Marmorana (Ambigua) signata* Férussac, 1821. B 20.2 mm [SMF 325420/1]; Italy, Passo di Furlò; M. Pfenninger leg. G. “*Helicopsis striata*” (Müller, 1774)? B 7.0 mm [SMF 325428/1] T. *geyeri* (Soós, 1926)? Germany, “Thüringen, Kyffhäuser”; D. Steinke leg. H. *Chilostoma cingulatum* (Studer, 1820) B 20.5 mm [RMNH H1938] Switzerland, Tessin, Melide along Lago di Lugano; J.T. Henrard leg., 28.viii.1938. I. *Xerotracha apicina* (Lamarck, 1822) B 7.8 mm (after Soes and de Winter, 2005) Netherlands, IJmuiden; A.J. de Winter leg. J. “*Iberus gualtierianus*” = *Iberus gualtierianus sensu lato* (L., 1758) B 29.6 mm [SMF 325431/1] Spain, “Ronda”; M. Pfenninger leg. K. *Trochoidea pyramidata* (Draparnaud, 1805) B 7.4 mm [SMF325443/1] France, “St. Maximin”; A. Eppenstein leg. L. *Otala punctata* (Müller, 1774) B 31.3 mm [SMF 325436/1] Spain, “Zaragoza”; M. Pfenninger leg.

We compared the sequence data published by Steinke et al. (2004) with our data and those published by other research groups (in particular Manganelli et al., 2005). For the Hygromiidae we restrict ourselves here to indicating anomalies or clear errors. Going into more detail would take things too far in this speciose and partly still problematic family. For the Helicidae *sensu stricto* however, we checked all the *COI* and *16S* sequence data published by Steinke et al. (2004).

## 2. Material and methods

### 2.1. Taxon sampling

In order to verify some of the results of Steinke et al. (2004) the following taxa were used: *Arianta arbustorum* (Linnaeus, 1758), C.

*lenticula* (Michaud, 1831), *Chilostoma cingulatum* (Studer, 1820), *Helicigona lapicida andorrlica* (Bourguignat, 1876), *Helicigona lapicida lapicida* (Linnaeus, 1758), *Helicodonta obvolvata* (Müller, 1774), *Isognomostoma isognomostomos* (Schröter, 1874), *Marmorana signata* (A. Férussac, 1821), *O. punctata* (Müller, 1774) and *Zonites algirus* (Linnaeus, 1758). For sampling localities, see Table 1. For the various species, the nomenclature proposed by CLECOM (Bank et al., 2001) is adopted, but while referring to sequence data in GenBank, the generic classification of the species is not altered.

### 2.2. Voucher material

Next to sequencing specimens of the above mentioned (sub)species, the following voucher specimens of Steinke et al. (2004), which are kept in the molluscan collection of the

**Table 1**  
Taxa and sampling localities.

Species	RMNH	Country	Localitiy	Lat/Long	Date	Collectors	GenBank Accession number	
							COI	16S
<i>Arianta arbustorum</i>	139193	Netherlands	Zuid-Holland, banks of the Wijde Aa river	52°11'00" N, 4°38'2" E	3.viii.1992	E. Gittenberger	AF296940	JF17809
<i>Arianta arbustorum</i>	139194	France	Aude, Gorges du Rébenty	-	15.vi.1985	A.J. de Winter	AF296945	JF17810
<i>Caracollina lenticula</i>	100400	Italy	Lampedusa island, Punta Sottile residence	-	2005	A.W. Janssen	JF177793	JF17811
<i>Chilostoma cingulatum</i>	114170	Austria	Tirol, Halltal, 9 km NE of Innsbruck	47°19'29" N, 5°45'24" E	1.v.2009	D.S.J. Groenenberg & M. Kolet	JF177794	JF17812
<i>Chilostoma cingulatum</i>	62177	Italy	Belluno, Valle d'Ampezzo, near Rio Felizon	-	18.vii.1968	W.H. Neuteboom	JF177795	JF17813
<i>Helicigona lapicida andorrca</i>	116481/1	Spain	Catalunya, Barcelona, Castellar de N'Hug, Fonis de Liobregat	42°16'59" N, 2°0'47" E	11.vi.2009	B. Kokshoorn & N. Snijders	JF177796	JF17814
<i>Helicigona lapicida andorrca</i>	116481/2	Spain	Catalunya, Barcelona, Castellar de N'Hug, Fonis de Liobregat	42°16'59" N, 2°0'47" E	11.vi.2009	B. Kokshoorn & N. Snijders	JF177797	JF17815
<i>Helicigona lapicida andorrca</i>	116481/3	Spain	Catalunya, Barcelona, Castellar de N'Hug, Fonis de Liobregat	42°16'59" N, 2°0'47" E	11.vi.2009	B. Kokshoorn & N. Snijders	JF177798	JF17816
<i>Helicigona lapicida lapicida</i>	96401	Belgium	Luxembourg, La Roche-en-Ardenne, chateau	50°10'56" N, 5°34'35" E	ix.2003	D.S.J. Groenenberg	JF177799	JF17817
<i>Helicigona lapicida lapicida</i>	99439	Germany	Nordrhein-Westfalen, Blankenheim castle	50°26'16"N, 6°39'4"E	24.vii.1983	P. Subai	JF17800	JF17818
<i>Helicodonta obvoluta</i>	139195	Netherlands	Limburg, Magraten, Savelsbos	50°47'21" N, 5°45'24" E	30.vii.2009	D.S.J. Groenenberg & P.J.L. Cuijpers	JF17801	JF17819
<i>Isoptomostoma isopomostomos</i>	139196	Italy	Trento-Alto Adige, between Predazzo and Bellamonte	46°18'37" N, 11°37'30" E	9.ix.2009	D.S.J. Groenenberg & I.M. Voermans	JF17802	JF17820
<i>Isoptomostoma isopomostomos</i>	139197	Italy	Trento-Alto Adige, between Predazzo and Bellamonte	46°18'37" N, 11°37'30" E	9.ix.2009	D.S.J. Groenenberg & I.M. Voermans	JF17803	JF17821
<i>Marmorana signata</i>	139198	Italy	Umbria, Perugia, Bevagna town wall	42°55'51" N, 12°36'16"E	25.v.2010	A.S.H. Breure	JF17804	JF17822
<i>Otala punctata</i>	139199	Spain	Valencia, Carlet, orange grove	-	31.v.2010	A. Martínez-Ortí	JF17805	JF17823
<i>Otala punctata</i>	139200	Spain	Valencia, Carlet, orange grove	-	31.v.2010	A. Martínez-Ortí	JF17806	JF17824
<i>Otala punctata</i>	139201	Spain	Valencia, Carlet, orange grove	-	31.v.2010	A. Martínez-Ortí	JF17807	JF17825
<i>Zonites algirus</i>	Pers.coll. Kornilios & Giokas	Greece	Peloponnisos, Lakonia prefecture, Gerolimenas	36°28'55" N, 22°24'02" E	29.5.2009	Páll-Gergely & G. Kornilios	JF17808	JF17826

Forschungsinstitut Senckenberg, Frankfurt am Main, Germany (=SMF), were studied by all authors of this paper: *C. lenticula* (SMF325448), *Xerosecta cespitum* (Draparnaud, 1801) (SMF325417), *C. cingulatum* (SMF325420), *H. lapicida* (SMF325426), *Helicopsis striata* (Müller, 1774) (SMF325428), *Iberus gualtierianus* (Linnaeus, 1758) (SMF325431), *Gittenbergeria turriplana* (Morelet, 1845) (SMF325435), *O. punctata* (SMF325436), *T. elegans* (Gmelin, 1791) (SMF325441), *T. pyramidata* (Draparnaud, 1805) (SMF325447) and *Zonites algirus* (SMF325447). The remainder 29 vouchers were inspected by Neubert only.

### 2.3. DNA-isolation, PCR and sequencing

Tissues were always stored in ethanol (70% or 96%), except for *C. cingulatum* from Italy (Table 1) which was stored in methylated spirits. Genomic DNA was extracted with a DNeasy blood and tissue kit of Qiagen, following the manufacturer's protocol. PCRs were carried out in 25 µl volumes using 1.25 units of Taq DNA polymerase from Qiagen, 0.4 mM of each primer and 0.2 mM dNTPs. For the amplification reactions (COI and 16S) the same primers were used as described in Steinke et al. (2004). We concentrated on mtDNA, in particular COI and 16S, since these can be considered as a single locus and most data in GenBank refer to those markers.

For both markers we used the following PCR thermoprofile: initial denaturation 3 min. @ 94 °C, followed by 40 cycles of – denaturation 15 s. @ 94 °C, annealing 30 s. @ 50 °C, extension 40 s. @ 72 °C – and a final extension of 5 min. @ 72 °C. The obtained PCR products were sent to Macrogen Inc. Europe (Amsterdam), where they were purified with a Montage purification kit (Millipore) and subsequently sequenced in both directions on an ABI3730XL using the same primers as used for PCR. The resulting forward and re-

verse sequences were assembled with Sequencher 4.10.1 (Gene Codes Corporation), checked for irregularities (including amino acid translation for COI) and submitted to GenBank.

### 2.4. Alignment

Our sequences were aligned with the sequences of Steinke et al. (2004) and other sequences from GenBank. For the analyses that focussed on the Helicidae s.s. (Ariantinae, Helicinae, Helicodontinae), COI sequences were aligned directly in MacClade 4.08 (Maddison and Maddison, 2005) and 16S sequences were aligned with MAFFT v.6.847b (Katoh et al., 2002). For an analysis that focussed on the Hygromiidae we used all of the sequences of Manganelli et al. (2005) and some noteworthy 16S sequences of Steinke et al. (2004). Because the Hygromiidae sequences were more diverged than the Helicidae sequences, this alignment clearly improved when structural information was included. Therefore MAFFT-Q-INS-i (Katoh and Toh, 2008) was used for a structural alignment of the 16S Hygromiidae dataset.

### 2.5. Phylogenetic analyses

For the Helicidae s.s. COI and 16S were analysed separately to test if both datasets would yield similar overall topologies and to check for inconsistencies within taxa. In order to make a comparison with the phylogeny presented by Steinke et al. (2004), the same 526 bases of COI were included; our 16S alignment had a length of 326 bases (compared to 323 = 380–57 of Steinke et al., 2004). For both datasets a NJ (Neighbor–Joining) analysis was performed in PAUP v. 4.0b10 for Unix (Swofford, 2002). We imposed a 3:3:1 weighting scheme on the COI-dataset to compensate for the

relaxed mutation rate at third base positions, rather than completely excluding these positions (as done by Steinke et al., 2004). For the analysis that focussed on the Hygromiidae 342 presumably homologous positions (on par with Manganeli et al., 2005) were included. Two Markov chain Monte Carlo (MCMC) runs were executed simultaneously in MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003) using the general time reversible model with a proportion of invariant sites and Gamma-distributed among site rate variation (GTR+I+G) as selected by MrModeltest 2.2 (Nylander, 2004). These analyses were set to run for 10,000,000 generations or until the MCMC runs converged (diagnostic stop value 0.009).

2.6. BLAST searches

MegaBLAST searches (that look for highly similar sequences) were performed on all of the COI (AY546262–AY546301) and 16S (AY546342–546381) sequences from Steinke et al. (2004) as currently included in GenBank. By default the output of a BLAST search is sorted by the bit score of high-scoring segment pairs (HSPs). Using any sequence from GenBank as a query sequence, obviously will result in that query sequence showing the largest bit score. When referring to ‘most similar’ sequences, we do not show the query sequences, but the second-best matches based on bit score. Since bit scores by themselves are not very informative between different BLAST searches, we report the identity scores (which are more intuitive for comparisons between sequences) for these second-best matches instead. The results for a selection of sequences (including all the Helicidae s.s. from Steinke et al., 2004) are summarized. Throughout this paper, whenever we refer to a BLAST search, we automatically imply MegaBLAST.

2.7. Genetic distances

Uncorrected *p*-distances (all positions included and weighted equally) were calculated with PAUP v. 4.0b10 for both the COI and 16S dataset (length 526 and 326 nucleotides, respectively, see above). These distances were converted into “similarity scores” (calculated as 100 minus the uncorrected *p*-distances times 100), so the shown values can be interpreted as percentages.

3. Results

Fig. 1 shows a compilation of some of the inspected vouchers of Steinke et al. (2004) and a selection of shells from the conchological collection of NCB Naturalis. Fig. 2 focusses mainly on the Helicidae s.s. and shows the COI and 16S NJ phylogenies based on sequences of Steinke et al. (2004; in blue), sequences from various studies (GenBank; in green) and from this study (in black). Fig. 3 depicts a majority rule consensus tree (showing all compatible partitions) of a Bayesian analysis that mainly focused on the Hygromiidae. After 3,340,000 generations both MCMC runs converged (the average standard deviation of split frequencies reached 0.008965), resulting in two tree files, each consisting of 33,401 trees of which the first 10,020 trees (30%) were discarded as burnin. Only posterior probabilities higher than 0.5 have been indicated. In this figure, sequences of Steinke et al. (2004) are in blue, sequences from various studies (Genbank) are in green and sequences from Manganeli et al. (2005) are in black. Table 2 gives an overview of some of the noteworthy MegaBLAST results based on the COI and 16S sequences (including all the Helicidae s.s.) of Steinke et al. (2004). Table 3 gives an overview of the genetic distances (in percentages, calculated as described above).

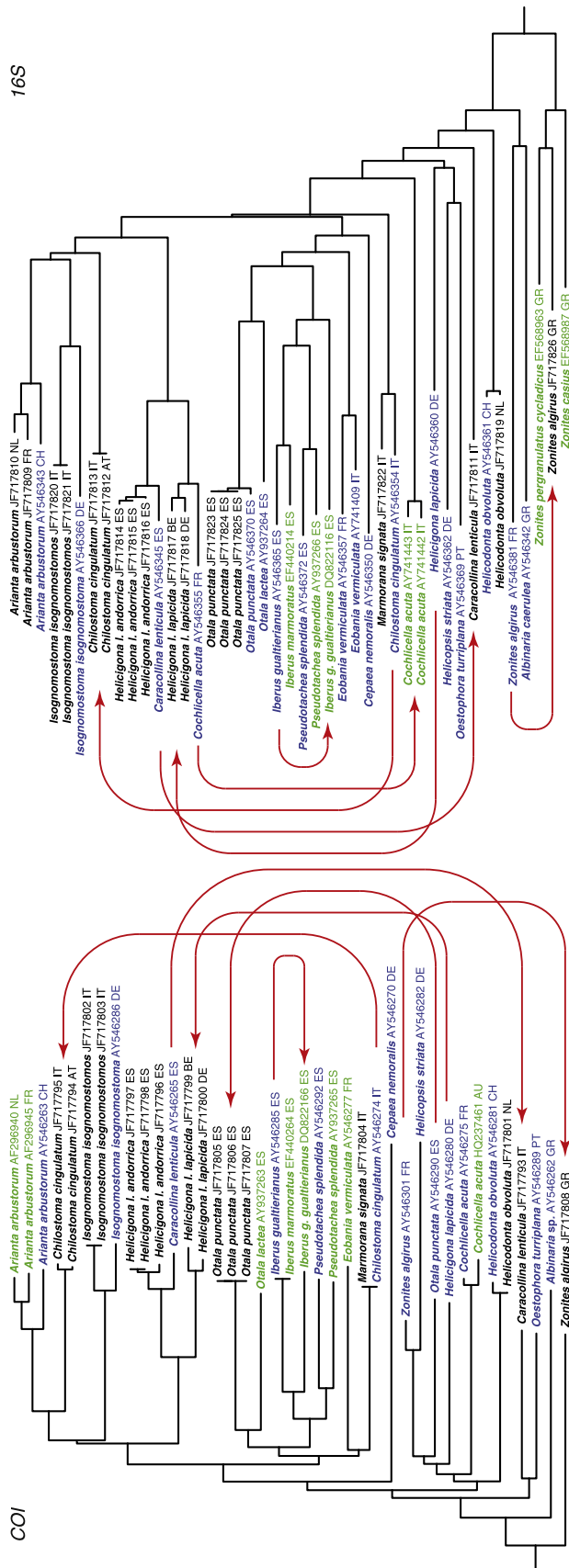
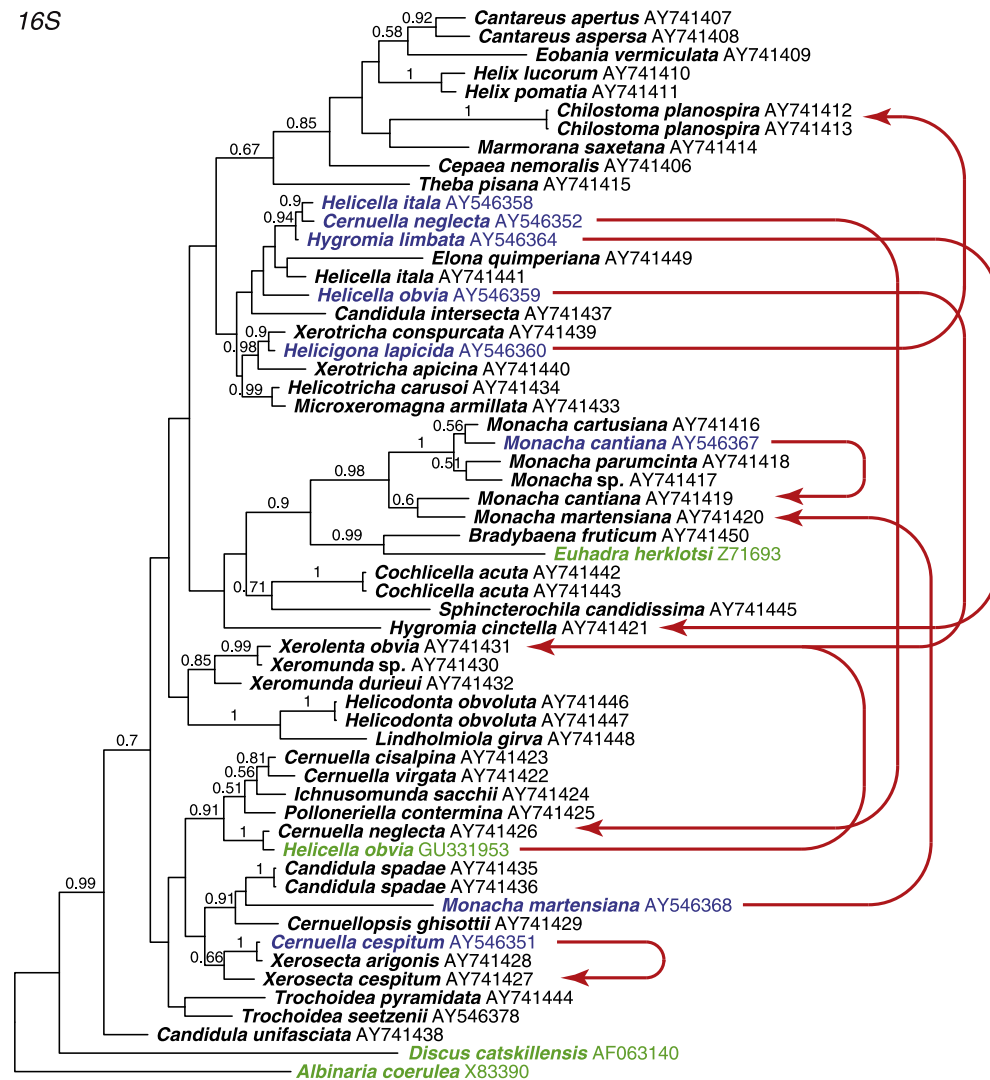


Fig. 2. A comparison of COI (left) and 16S (right) NJ phylogenies with a focus on the Helicidae s.s. In blue are sequences from Steinke et al. (2004), in green sequences from various studies (GenBank) and in black sequences from this study. The red arrow indicates erroneous or misidentified sequences and the correct sequences for the corresponding most closely related taxa in this figure.

16S



**Fig. 3.** Bayesian consensus tree based on 16S with a focus on the Hygrommidae. This cladogram shows all compatible partitions and posterior probabilities of 0.5 and higher. In blue are sequences from Steinke et al. (2004), in green sequences from various studies (Genbank) are in green and in black sequences from Manganelli et al. (2005). The red arrow indicates erroneous or misidentified sequences and the correct sequences for the corresponding most closely related taxa in this figure.

#### 4. Discussion

Helicidae, Ariantinae

##### 4.1. *Caracollina lenticula* & *Helicigona lapicida* andorrica

A BLAST search with the so-called *C. lenticula* COI sequence (AY546265) showed a relative high similarity (BLAST identity-score of 90%, Table 1) with sequence AF297000 for *H. lapicida* (Gittenberger et al., 2004). The voucher for *C. lenticula* (SMF325448, not SMF325411 as stated in Appendix A by Steinke et al. (2004): Fig. 1A) is a juvenile shell of *H. lapicida andorrica*, which differs from real *C. lenticula* (Fig. 1B) by its much larger size, sharper and more centrally situated peripheral keel and a granular surface. Despite the fact that Steinke et al. (2004) did include a correctly identified fully grown specimen of *H. lapicida lapicida* (Fig. 1D) in their study (see below), they did not recognize a juvenile specimen of the subspecies from Andorra and parts of the surrounding Pyrenees. The COI and 16S sequences (JF717796–JF717798 and JF717814–JF717816) that we independently obtained for three specimens of *H. lapicida andorrica* were on average

97.2% and 98.0% identical (Table 2) to the sequences AY546265 and AY546345 of *C. lenticula* sensu Steinke et al. (2004). The COI and 16S sequences that we ourselves obtained for the real *C. lenticula* (JF717793 and JF717811) differed 19.4% and 29.5%, respectively, from the alleged *C. lenticula* sequences of Steinke et al. (2004). Moreover, Fig. 2 shows that according to our COI data *C. lenticula* belongs to the Helicodontinae or Trissexodontinae, in accordance with the usual classification (e.g. Bank et al., 2001). The 16S dataset is somewhat indeterminate regarding the position of *C. lenticula*, but that taxon clearly does not belong to the monophyletic subfamily Ariantinae.

We consider it a methodological weakness that Steinke et al. (2004), instead of reconsidering their identification on the basis of their voucher specimen (or checking their surprising results by sequencing additional specimens), asserted that *C. lenticula* belongs to the Ariantinae instead of the Helicodontinae.

##### 4.2. *Chilostoma cingulatum* & *Marmorana signata*

BLAST searches with the COI (AY546274) and 16S (AY546354) sequences of Steinke et al.'s (2004) so-called *C. cingulatum* show

**Table 2**  
MegaBLAST search results based on a selection of COI and 16S sequences.

Query taxon	Accession Nr.	Most significant alignment	Accession Nr.	Score	Coverage (%)	Identity (%)	Reference
MegaBlast searches based on COI sequences of Steinke et al. (2004)				Date on which the search was performed: 28–03–2011			
<i>Arianta arbustorum</i>	AY546263	<i>Arianta arbustorum</i>	AF296945	800	100	94	Gittenberger et al. (2004)
<i>Caracollina lenticula</i>	AY546265	<i>Helicigona lapicida</i>	AF297000	660	95	90	Gittenberger et al. (2004)
<i>Cepaea nemoralis</i>	AY546270	<i>Cepaea nemoralis</i>	U23045	902	100	97	Yamazaki et al. (1997)
<i>Chilostoma cingulatum</i>	AY546274	<i>Marmorana signata strigata</i>	GU391375	654	67	100	Fiorentino et al. (2010)
<i>Cochlicella acuta</i>	AY546275	<i>Cochlicella acuta</i>	HQ237463	950	100	99	Unpublished
<i>Eobania vermiculata</i>	AY546277	<i>Eobania vermiculata</i>	JF277391	881	100	98	Unpublished
<i>Helicella itala</i>	AY546278	<i>Helicella obvia</i>	AY546279	822	95	96	Steinke et al. (2004)
<i>Helicigona lapicida</i>	AY546280	<i>Lozekia deubeli</i>	EU182491	523	100	84	Fehér et al. (2009)
<i>Helicodonta obvoluta</i>	AY546281	<i>Lozekia deubeli</i>	EU182493	440	99	81	Fehér et al. (2009)
<i>Helix aspersa</i>	AY546283	<i>Helix aspersa</i>	AY345052	913	100	98	Grande et al. (2004)
<i>Iberus gualtierianus</i>	AY546285	<i>Iberus marmoratus</i>	EF440264	972	100	100	Elejalde et al. (2008a)
<i>Isognomostoma isognomostoma</i>	AY546286	<i>Eobania vermiculata</i>	JF277394	444	98	82	Unpublished
<i>Monacha cantiana</i>	AY546287	<i>Helix aspersa</i>	HM179144	375	90	81	Unpublished
<i>Monacha martensiana</i>	AY546288	<i>Helix aspersa</i>	HM179144	497	99	84	Unpublished
<i>Oestophora turriplana</i>	AY546289	<i>Candidula codia</i>	AY238622	891	93	100	Pfenninger et al. (2003)
<i>Otala punctata</i>	AY546290	<i>Trochoidea pyramidata</i>	AY546297	894	99	97	Steinke et al. (2004)
<i>Pseudotachea splendida</i>	AY546292	<i>Pseudotachea splendida</i>	AY937265	617	100	87	Elejalde et al. (2008b)
<i>Trochoidea elegans</i>	AY546295	<i>Ceruellea cespitum</i>	AY546271	824	98	99	Steinke et al. (2004)
<i>Trochoidea pyramidata</i>	AY546297	<i>Otala punctata</i>	AY546290	894	98	97	Steinke et al. (2004)
<i>Zonites algirus</i>	AY546301	<i>Helicopsis striata</i>	AY546282	675	81	95	Steinke et al. (2004)
MegaBlast searches based on 16S sequences of Steinke et al. (2004)				Date on which the search was performed: 28–03–2011			
<i>Arianta arbustorum</i>	AY546343	<i>Helix aspersa</i>	AF126131	241	87	83	Guiller et al. (2001)
<i>Caracollina lenticula</i>	AY546345	<i>Cochlicella acuta</i>	AY546355	385	100	97	Steinke et al. (2004)
<i>Cepaea nemoralis</i>	AY546350	<i>Cepaea nemoralis</i>	AF249259	612	100	99	Wollschel d-Lengeling et al. (2001)
<i>Ceruellea neglecta</i>	AY546352	<i>Hygromia limbata</i>	AY546364	569	100	96	Steinke et al. (2004)
<i>Chilostoma cingulatum</i>	AY546354	<i>Marmorana signata</i>	GU391404	370	91	89	Fiorentino et al. (2010)
<i>Cochlicella acuta</i>	AY546355	<i>Caracollina lenticula</i>	AY546345	385	100	87	Steinke et al. (2004)
<i>Eobania vermiculata</i>	AY546357	<i>Eobania vermiculata</i>	JF277389	459	100	96	Unpublished
<i>Helicella itala</i>	AY546358	<i>Hygromia limbata</i>	AY546364	599	100	95	Steinke et al. (2004)
		<i>Ceruellea neglecta</i>	AY546352	542	92	95	Steinke et al. (2004)
		<i>Helicella itala</i>	AY741441	442	94	89	Manganelli et al. (2005)
<i>Helicella obvia</i>	AY546359	<i>Hygromia limbata</i>	AY546364	387	100	87	Steinke et al. (2004)
<i>Helicigona lapicida</i>	AY546360	<i>Xerotricha conspurcata</i>	AY741439	499	100	94	Manganelli et al. (2005)
<i>Helicodonta obvoluta</i>	AY546361	<i>Helicodonta obvoluta</i>	AY741446	457	100	94	Manganelli et al. (2005)
<i>Helix aspersa</i>	AY546363	<i>Helix aspersa</i>	AF126128	580	100	99	Guiller et al. (2001)
<i>Iberus gualtierianus</i>	AY546365	<i>Iberus marmoratus</i>	EF440214	416	100	91	Elejalde et al. (2008a)
<i>Isognomostoma isognomostoma</i>	AY546366	<i>Chilostoma cingulatum</i>	AY546354	255	96	82	Steinke et al. (2004)
<i>Monacha cantiana</i>	AY546367	<i>Monacha cartusiana</i>	AY741416	276	88	96	Manganelli et al. (2005)
<i>Monacha martensiana</i>	AY546368	<i>Candidula spadae</i>	AY741435	377	89	88	Manganelli et al. (2005)
<i>Oestophora turriplana</i>	AY546369	<i>Monacha martensiana</i>	AY741420	145	92	78	Manganelli et al. (2005)
<i>Otala punctata</i>	AY546370	<i>Otala lactea</i>	AY937264	425	100	89	Elejalde et al. (2008b)
<i>Pseudotachea splendida</i>	AY546372	<i>Pseudotachea splendida</i>	AY937266	499	100	93	Elejalde et al. (2008b)
<i>Trochoidea elegans</i>	AY546375	<i>Xerosecta arigonis</i>	AY741428	254	100	81	Manganelli et al. (2005)
<i>Trochoidea pyramidata</i>	AY546377	<i>Trochoidea pyramidata</i>	AY741444	459	94	90	Manganelli et al. (2005)
<i>Zonites algirus</i>	AY546381	<i>Zebrina detrita</i>	AY485907	191	99	78	Parmakelis et al. (2005)

**Table 3**  
Sequence similarities based on uncorrected p-distances.

From taxon (COI sequence/16S sequence):	To taxon (COI sequence/16S sequence):	Similarity score (%)	
		COI	16S
<i>Arianta arbustorum</i> AY546263/AY546343	<i>Arianta arbustorum</i> (AF296940 + AF296945)/(JF717809–JF717810)	93.2	95.1
<i>Caracollina lenticula</i> AY546365/AY546345	<i>Helicigona lapicida andorrica</i> (JF717796–JF717798)/(JF717814–JF717816)	97.2	98.0
<i>Caracollina lenticula</i> AY546365/AY546345	<i>Caracollina lenticula</i> JF717793/JF717811	80.6	70.5
<i>Chilostoma cingulatum</i> AY546354/AY54635	<i>Chilostoma cingulatum</i> (JF717794–JF717795)/(JF717812–JF717813)	81.2	76.9
<i>Chilostoma cingulatum</i> AY546354/AY54635	<i>Marmorana signata</i> JF717804/JF717822	100.0	94.9
<i>Cochlicella acuta</i> AY546275/AY546355	<i>Cochlicella acuta</i> HQ237461/(AY741442–AY741443)	98.9	69.1
<i>Cochlicella acuta</i> AY546275/AY546355	<i>Helicigona lapicida lapicida</i> (JF717796–JF717800)/(JF717817–JF717818)	75.9	95.2
<i>Helicigona lapicida</i> AY546280/AY546360	<i>Helicigona lapicida lapicida</i> (JF717799–JF717800)/(JF717817–JF717818)	76.7	74.0
<i>Helicodonta obvoluta</i> AY546281/AY546361	<i>Helicodonta obvoluta</i> JF717801/JF717819	99.6	99.6
<i>Isognomostoma isognomostoma</i> AY546286/AY546366	<i>Isognomostoma isognomostomos</i> (JF717802–JF717803)/(JF717820–JF717821)	94.1	88.2
<i>Otala punctata</i> AY546290/AY546370	<i>Otala punctata</i> (JF717805–JF717807)/(JF717823–JF717825)	80.8	99.6
<i>Zonites algirus</i> AY546301/AY546381	<i>Zonites algirus</i> JF717808/JF717826	78.4	65.7
<i>Zonites algirus</i> AY546301/AY546381	<i>Helicopsis striata</i> AY546282/AY546362	93.2	60.0

\*Similarity score = 100 – uncorrected p-distance \* 100.

a striking similarity (Table 2) with sequences GU391375 (100%) and GU391404 (89%), respectively, of *Marmorana signata strigata* (see Fiorentino et al., 2010). The voucher of *C. cingulatum* sensu Steinke et al., 2004 (SMF325420: Fig. 1F), is a species of *Marmorana*

Please cite this article in press as: Groenberg, D.S.J., et al. Reappraisal of the “Molecular phylogeny of Western Palaearctic Helicidae s.l. (Gastropoda: Stylommatophora)”: When poor science meets GenBank. Mol. Phylogenet. Evol. (2011), doi:10.1016/j.ympev.2011.08.024

indeed and not *C. cingulatum* (Fig. 1H). The *COI* and *16S* sequences (JF717804 and JF717822) that we obtained for a specimen which we identified as *M. signata* (Table 1) ourselves, were 100% and 94.9% identical (Table 3) to sequences AY546274 and AY546354, respectively. The average sequence divergences (Table 3) between our own *C. cingulatum* sequences (JF717794–JF717795 and JF717811–JF717812) and *C. cingulatum* sensu Steinke (AY546274 and AY546354) are 18.8% (*COI*) and 23.1% (*16S*, Table 3). The phylogeny reconstructions for both *COI* and *16S* datasets show that our specimens of *C. cingulatum* belong to the Ariantinae clade (Fig. 2), in conformity with the usual classification (Bank et al., 2001), whereas *M. signata*, or *C. cingulatum* sensu Steinke et al. (2004), has to be classified with another subfamily, possibly the Helicinae.

#### 4.3. *Helicigona lapicida* & *Xerotracha* sp.

Even though the shell voucher for *H. lapicida* (SMF325426; Fig. 1D) was correctly identified, it remains unclear what species or even genus Steinke et al. (2004) have actually sequenced here. A BLAST search with sequences AY546280 (*COI*) and AY546360 (*16S*), published and deposited in GenBank as connected with voucher SMF325426, showed a similarity of 84% with a sequence of *Lozekia deubeli* (M. von Kimakowicz, 1890) (EU182491, Fehér et al., 2009) and 94% with *Xerotracha conspurcata* (Draparnaud, 1801) (AY741439, Manganeli et al., 2005), respectively (Table 2). A phylogeny reconstruction based on a larger *COI* dataset (results not shown) indicates that it might belong to a *Leptaxis* species (Van Riel et al., 2005), but the sequences from that study are considerably shorter. No *COI* sequences have yet been deposited in Genbank for *Xerotracha*. However, the *16S* data show (Fig. 3) that sequence AY546360 is situated in between *X. conspurcata* and *X. apicina* (Lamarck, 1822) (Manganeli et al., 2005). On that basis we conclude that contrary to its supposed shell voucher, AY546360 belongs to a *Xerotracha* species (Fig. 1I), most certainly not to *H. lapicida* (Fig. 1D). The *COI* and *16S* sequences (JF717799–JF7177800 and JF717817–JF717818, Table 1) that we independently obtained for *H. l. lapicida* on average differed 23.3% and 26.0% from sequence AY546280 and AY546360, respectively (Table 3). Moreover, our *COI* sequences are on average 98.7% identical to the *H. lapicida* sequences (EF398129–EF398131) of Haase and Misof (2009). The NJ phylogeny reconstructions on the basis of both *COI* and *16S* (Fig. 2) show that *H. lapicida lapicida* and *H. lapicida andorrnica* form a monophyletic group, which is not closely related to *H. lapicida* sensu Steinke et al. (2004).

#### 4.4. *Arianta arbustorum* and *Isognomostoma isognomostomos*

The only *COI* and *16S* sequences in GenBank from representatives of the subfamily Ariantinae that were deposited by research groups other than Steinke et al. (2004) are *COI* sequences for *A. arbustorum* (Gittenberger et al., 2004; Haase and Misof, 2009). Hence only the BLAST search for sequence AY546263 (*A. arbustorum*, *COI*) yielded a high similarity with other sequences in GenBank (94% with sequence AF296945 of *A. arbustorum*, Table 2). The other three sequences (*A. arbustorum 16S* and *I. isognomostomos COI* and *16S*) were less similar (BLAST identity scores of less than 83%, Table 2) to best-matching sequences in GenBank, but were very similar to our sequences for those species. Our *16S* sequences for *A. arbustorum* (JF717809–JF717810) were on average 95.1% identical to sequence AY546343, whereas our *COI* and *16S* sequences for *I. isognomostomos* (JF717802–JF717803 and JF717820–JF717821) were on average 94.1% and 88.2% identical to sequences AY546286 and AY546366 (Steinke et al., 2004), respectively (Table 3). Given the fact that the vouchers for *A. arbustorum* (SMF325410) and *I. isognomostomos* (SMF325432) were identified correctly and that the sequences deposited by Steinke et al.

(2004) are consistent with our data, we assert that sequences AY546263, AY546286, AY546343 and AY546366 indeed belong to the indicated species.

#### Helicinae, Helicinae

#### 4.5. *Iberus gualtierianus* & *Iberus marmoratus* (A. Férussac, 1821)

A BLAST search with sequences AY546285 (*COI*) and AY546365 (*16S*) for *Iberus gualtierianus* sensu Steinke et al. (2004) showed a very high similarity (Table 2) with sequences EF440264 (100% identical) and EF440214 (BLAST identity score of 91%) for *I. marmoratus* (Elejalde et al., 2008a). Elejalde et al. (2008b) showed that the *I. gualtierianus* species complex is monophyletic and consists of at least six major clades, whereas *I. marmoratus* belongs to a different *Iberus* lineage and constitutes a separate clade with at least three other species (clade 5 in Elejalde et al., 2008a). An analysis with all of the sequences currently available for *I. gualtierianus* (78 for each *COI* and *16S*) and *I. marmoratus* (11 for each marker), resulted in two major clades, i.e. one for *I. gualtierianus* and one for *I. marmoratus*. The only exception are the alleged *I. gualtierianus* sequences of Steinke et al. (2004) (AY546285 and AY546365) which group within the *I. marmoratus*-clade (for both *COI* and *16S*) sensu Elejalde et al. (2008a). Inspection of the voucher for *I. gualtierianus* (SMF325431; Fig. 1J), according to the label collected near Ronda, Spain, brings another ‘problem’ to light. The shell in question is bleached and full of mud. Clearly, the snail that once sheltered inside cannot have been used for a molecular analysis. Shells that were severely damaged during the research process were replaced by better ones later on (pers. com. by Prof. Dr. M. Pfenninger to Dr. R. Janssen, Dr. E. Neubert and H. Nordsieck at a joint meeting on April 20th, 2005, which took place to discuss the observed problems with the voucher specimens). This procedure is fundamentally incorrect and potentially misleading however. A voucher should represent a part of the specimen that was actually sequenced. Occasionally labwork indeed is destructive and we empathise efforts that are being made to at least have a reference specimen, but if not the real voucher these should be marked as such (e.g. para-voucher) and originate from the same sample or population. In this case, the ‘pseudo-voucher’ is most probably from even another locality. From near Ronda, *I. gualtierianus* is not known (T.E.J. Ripken, pers. com.), but *I. marmorana* occurs in that area indeed and, obviously, the sequences AY546285 and AY546365 belong to that species. Apparently, Steinke et al. (2004) sequenced *I. marmorana* from near Ronda, recorded that locality for their species, and deposited the shell of a conchologically similar, somewhat larger species, viz. *I. gualtierianus*, from another locality in southern Spain as its voucher.

#### 4.6. *Otala punctata* & *Trochoidea pyramidata*

A BLAST search with sequence AY546290 (*COI*) for *O. punctata* yielded an identity score of 97% (Table 2) with sequence AY546297 for *T. pyramidata* (both sequences originate from Steinke et al., 2004). This is remarkable in view of a quite different *COI* sequence in GenBank for *Otala lactea* (Müller, 1774) (AY937263; Elejalde et al., 2008b). Inspection of vouchers SMF325443 and SMF325436 for *T. pyramidata* and *O. punctata*, respectively (Fig. 1K and L), made clear that these specimens had been identified correctly. An additional BLAST search with the *16S* sequence of *O. punctata* (AY546370) does show sequence AY937264 of *O. lactea* (Elejalde et al., 2008b) as most similar (89%, Table 2) for this marker. This result potentially indicated an inconsistency between the *COI* and *16S* datasets of Steinke et al. (2004). As a most likely explanation for the conflicting results we assume that the so-called *O. punctata COI* sequence resulted from

contamination by *T. pyramidata*. If “only” some sequences would have been swapped, each sequence should still be unique and at least one of them should be most closely related to *O. lactea* (AY937263). Since no other sequences of *O. punctata* are available to support this explanation, we sequenced some specimens of *O. punctata* (kindly put at our disposal by A. Martínez-Orti) ourselves. It turned out that our *COI* sequences for *O. punctata* (JF717805–JF717807) differed on average 19.2 % from the *O. punctata* sequence of Steinke et al., whereas our 16S sequences (JF717823–JF717825) were nearly identical (on average 99.6%, Table 3) to AY546370. In this case Steinke et al.’s (2004) 16S sequence for *O. punctata* is correct, but the *COI* sequence is from another species. Moreover, inclusion of our *O. punctata* sequences showed (Fig. 2) that *O. punctata* and *O. lactea* form a monophyletic group and that the *COI* and 16S datasets are congruent.

#### 4.7. Other Helicinae in Steinke et al. (2004)

Sequences of the four remaining Helicinae used by Steinke et al. (2004), showed most significant alignments in a BLAST search with sequences for the same species as indicated in that paper (Table 2). In fact, except for *Pseudotachea splendida* (Draparnaud, 1801) these sequences all have an identity score of 96% or higher (Table 2). Neither these BLAST results, nor the inspected vouchers indicate other errors within the Helicinae, than those here reported for *Iberus* and *Otala*.

#### Helicidae, Helicodontinae

#### 4.8. *Gittenbergeria turriplana* & *Candidula codia* (Bourguignat, 1859)

The only sequences in GenBank for *Oestophora*, or *Gittenbergeria* according to Schileyko (1991), are those deposited by Steinke et al. (2004). Consequently, we expected to yield only moderate similarities with our BLAST searches. Nevertheless, sequence AY546289 (*COI*) for *G. turriplana* was (100%, Table 2) identical to sequence AY238622 of *C. codia* (published by Pfenniger et al., 2003, according to the GenBank record, but not mentioned in that paper). Inspection of the voucher for *G. turriplana* (SMF325435: Fig. 1C) showed that it was identified correctly. To illustrate that the two species are strikingly different, a specimen of *C. codia* is also depicted (Fig. 1E). Clearly, the *COI* sequence cannot belong to both species. Given the position of sequence AY546289 in the *COI* based phylogeny reconstruction (the left tree in Fig. 2), i.e. next to our sequence for *C. lenticula* (both belonging to the Trissexodontidae; Nordsieck, 1987), together with the fact that voucher SMF325435 for *G. turriplana* was identified correctly (we did not inspect the voucher for *C. codia*), we presume that sequence AY546289 belongs to *G. turriplana* indeed and hence that sequence AY238622 does not represent *C. codia*. Maybe the fact that both *G. turriplana* and *C. codia* occur syntopically near Boliqueime, Portugal (‘Boliqenne’ after Steinke et al., 2004: Appendix A), partly explains this confusion.

Given the extreme similarity of the *H. obvoluta* sequences (99.6% for both datasets, Table 3) with our sequences for that species and the fact that voucher SMF325427 was identified correctly, the data deposited for *H. obvoluta* are shown to be correct.

#### Anomalies within the remainder data.

#### 4.9. *Zonites algirus* & *Helicella itala* (Linnaeus, 1758)

Although voucher SMF325447 was correctly identified as *Z. algirus*, a BLAST search showed that sequence AY546301 (*COI*) is strikingly similar (95%, Table 2) to sequence AY546282 for *Helicopsis striata* from the same publication. For neither species are additional

*COI* sequences available at this moment, but for *Zonites* several 16S sequences have been deposited in GenBank (among which *Z. algirus*; Kornilios et al., 2009). A BLAST search with sequence AY546381 (16S) supposedly also belonging to voucher SMF325447 did not show any of the just mentioned *Zonites* sequences as most similar. Instead sequence AY485907 for *Zebrina detrita* (Müller, 1774) (Parmakelis et al., 2005) was retrieved as being most similar (78%, Table 2). Particularly given the surplus of 16S data it was a pity not to have any additional *COI* sequences of *Zonites*. Upon request P. Kornilios and S. Giokas kindly put at our disposal a *COI* and 16S sequence for a specimen which they identified as *Z. algirus*. These sequences were used in the NJ phylogenies depicted in Fig. 2. The 16S phylogeny (the tree on the right side in Fig. 2) shows that sequence AY546381 does not belong to the monophyletic (after Kornilios et al., 2009) taxon *Zonites*. Moreover, these *Z. algirus* sequences (JF717808 and JF717826, Table 1) differed 21.6% and 34.3% (*COI* and 16S, respectively, Table 3) from the sequences for the same species (AY546301 and AY546381) according to Steinke et al. (2004). These results indicate that sequence AY546301 (*COI*) must be a contaminant, possibly originating from *H. striata*. This definitely needs confirmation however, since the ‘voucher’ for *H. striata* (SMF325428) is another example of a ‘pseudo-voucher’ (Fig. 1G). Given the fact that sequence AY546381 is unique within the 16S dataset of Steinke et al. and very different (40.0%, Table 3) from sequence AY546362 for *H. striata*, we cannot suggest to which species it actually belongs. It is certainly not a *Zonites* species.

#### 4.10. *Trochoidea elegans* & *Xerosecta cespitum*

A BLAST search with sequence AY546295 (*COI*) for *T. elegans* sensu Steinke et al. (2004) yielded an identity score of 99% (Table 2) with sequence AY546271 for “*Cernuella*” *cespitum*, published in the same paper. The only other *T. elegans* sequence on GenBank (FJ627176; Sauer and Hausdorf, 2009) differs 20.9% from sequence AY546295.

Moreover, a NJ analysis with all of the *COI* sequences of Steinke et al. (2004), thus including the five *Trochoidea* sequences and additional sequences for *T. elegans* (FJ627176), *T. spratti* (L. Pfeiffer, 1846) (FJ627177) and *T. geyeri* (Soós, 1926) (AY240948) never showed *Trochoidea* as a monophyletic group (phylogeny not depicted). The group of *Trochoidea* sequences that did form a clade always excluded *T. elegans* (AY546295), *T. trochoides* (Poiret, 1789) (AY546299) and *T. seetzeni* (L. Pfeiffer, 1847) (AY546298). This contrasts conspicuously with the monophyly of the genus as depicted in Fig. 1 in Steinke et al. (2004). As for *T. seetzeni* it might be true that this species indeed does not belong to *Trochoidea*; nowadays it is classified under the genus *Xerocrassa*.

The 16S data show a slightly different pattern (also indicating however, that *Trochoidea* is not monophyletic); here *T. elegans* does group with *T. pyramidata* (the identity of the latter sequence is confirmed by sequence AY741444 of Manganelli et al., 2005). Inspection of voucher SMF325441 for *T. elegans* showed that this specimen was identified correctly. For now we can only conclude that the *COI* sequence AY546295 does not belong to *T. elegans* and possibly represents *Xerosecta cespitum* (SMF325417; Steinke et al., 2004).

#### 4.11. *Cochlicella acuta* (Müller, 1774) & *Helicigona lapicida*

A BLAST search with sequence AY546275 (*COI*) for *C. acuta* shows that it is nearly identical (average identity score of 99%, Table 2) to sequences HQ237461–HQ237463 for the same species (a study by Bon et al.; data unpublished). However, the 16S sequence (AY546355) is very different (on average 30.9%, Table 3) from other *C. acuta* 16S sequences (AY741442 and AY741443; Manganelli et al., 2005). The similarity of sequence AY546355 to our sequences



(JF717817–JF717818) for *H. lapicida lapicida* (average similarity score of 95.2%, Table 3) and the topological position of that sequence within the 16S NJ phylogeny (Fig. 2) are striking. Based on these results we conclude that the 16S sequence for *C. acuta* sensu Steinke et al. (2004) exemplifies contamination and actually originated from *H. lapicida* (voucher SMF325426; Fig. 1D). Here-with showing yet another inconsistency between the COI and the 16S datasets of Steinke et al. (2004).

#### 4.12. *Monacha* species

A BLAST search with Steinke et al.'s (2004) COI (AY546287) and 16S (AY546367) sequences for *Monacha cantiana* (Montagu, 1803) (voucher SMF325433) did not show sequences HQ204502 and HQ204543 (Duda et al., 2011) for the same marker and species, respectively, as most similar. By sequencing a number of *Monacha* species, including *Monacha cartusiana* (Müller, 1774) and *M. cantiana*, Manganelli et al. (2005) have shown that the genus is monophyletic (also Fig. 3; posterior probability 0.98). The BLAST search with sequence AY546367 (16S), however, did not show *M. cantiana* (AY741419), but *M. cartusiana* (AY741416) as most similar (identity score of 96%, Table 2). Since two research groups (Manganelli et al., 2005; Duda et al., 2011) independently obtained similar 16S sequences for *M. cantiana* (AY741419 and HQ204543) that clearly differ (Fig. 3) from the *M. cantiana* sequence (AY546367) sensu Steinke et al. (2004), we conclude that the latter authors did not sequence *M. cantiana*, but most likely *M. cartusiana*. The BLAST search with sequence AY546287 (COI) was most similar (81%, Table 2) to sequence HM179144 for *Cornu aspersum* (Müller, 1774) At this moment it is unclear if COI sequence AY546287 indeed belongs to a species of *Monacha*. In agreement with the BLAST result for 16S, the shell voucher (SMF325433) belongs to *M. cartusiana*.

A BLAST search with sequences AY546288 (COI) and AY546368 (16S), allegedly belonging to *M. martensiana* (Tiberi, 1869) (Steinke et al., 2004), showed the highest similarity with sequences HM179144 (84%) and AY741435 (88%) of *Cornu aspersum* (unpublished) and *Candidula spadae* (Calcara, 1845) (Manganelli et al., 2005), respectively. Currently the only other *M. martensiana* sequence available on GenBank is sequence AY741420 (16S) of Manganelli et al. (2005), which differs 22.5% from sequence AY546368. Hence the *M. martensiana* sequences of Steinke et al. (2004) are unlikely to belong to *Monacha* (see Fig. 3). However, voucher SMF 325434 indeed belongs to *Monacha*, possibly *M. cartusiana* as well.

#### 4.13. *Helicella itala* and *Xerolenta obvia* (Menke, 1828)

At this moment, three research groups deposited clearly different sequences in Genbank for *Xerolenta obvia* (Steinke et al., 2004; Manganelli et al., 2005 and Dinapoli et al., 2011), suggesting that the identification of this species requires specialist knowledge. Given the close relation of *Xerolenta* with *Xeromunda*, we assume that only the identification of Manganelli will be correct (see Fig. 3). In agreement with Steinke et al. (2004, Fig. 1), separate NJ analyses (including all taxa from that paper) of both the COI and 16S datasets (results not depicted) showed that *H. itala* (AY546358) and *X. obvia* (AY546359) form a clade (which is also shown in Fig. 3, albeit with a low posterior probability). In comparison with the Helicellinae sensu Steinke et al. (2004), the helicelline hygromiids and monachine hygromiids sensu Manganelli et al. (2005) form a broader sampling of the taxonomic diversity within this group. Manganelli et al. (2005) showed that different reconstruction methods yield different topologies (based on 16S), but each of their phylogeny reconstructions shows that *H. itala* and *X. obvia* do not constitute a monophyletic group. We used Fig. 3 to assess how the sequences for *H. itala* and *X. obvia* from the different studies relate to each other and conclude that *X. obvia* (GU331953) sensu Dinapoli et al. is most probably

the same species that Manganelli et al. identified as *Cernuella neglecta* (Draparnaud, 1805) (AY741426); the BLAST identity score between both sequences is 97%. As far as only the above mentioned sequences are concerned, Steinke et al. and Manganelli et al. seem to recognize the same taxon as *H. itala* (AY546359 and AY741441). Nevertheless BLAST identity scores (Table 2) for the 16S sequence of *H. itala* (AY546358) sensu Steinke et al. (2004) AY546359 are higher with *Hygromia limbata* (Draparnaud, 1805) (AY546364; 95%) and *C. neglecta* (AY546352; 95%) from the same study than with *H. itala* (AY741441; 87%) from Manganelli et al. (2005). This is also shown in Fig. 3 where *H. itala*, *H. limbata* and *C. neglecta* sensu Steinke et al. (2004) constitute a clade with a high posterior probability (0.94). Given the monophyly of *Cernuella* (as supported by *C. cisalpina* (Rossmässler, 1837) and *C. virgata* (Da Costa, 1778); Fig. 3), we may conclude that *C. neglecta* (AY546352) sensu Steinke et al. (2004) does not represent a *Cernuella* species. Given the limited data on *Hygromia* (Fig. 3 presumably includes only two species) we can not draw firm conclusions about sequence AY546364 (*H. limbata* sensu Steinke et al., 2004). However we can conclude that the 16S sequences of *H. limbata*, *C. neglecta* and *H. obvia* sensu Steinke et al. (2004), are all more similar to *H. itala* (based on sequences of both Steinke et al., 2004 and Manganelli et al., 2005) than each of them is to the most closely related taxon in Fig. 3 (red arrows). In general sequences and phylogenies published by Manganelli et al. (2005) strongly contrast with the data of Steinke et al. (2004) and much better reflect current anatomy-based classifications.

#### 4.14. Accuracy of deposited sequences

In order to align the COI sequences, gaps had to be postulated in sequences of *T. pyramidata* (AY546297), *Trochoidea geyeri* (AY546296) and *O. punctata* (AY546290) and one base had to be deleted from the sequence of *H. limbata* (AY546284). Since all of these indels will cause frame-shifts, we conclude that the sequences were not properly checked before submission to GenBank.

## 5. Conclusion

This study shows a variety of errors and methodological weaknesses in the 2004 paper of Steinke et al. published in Molecular Phylogenetics and Evolution. In that paper *H. lapicida andorrnica* and *M. signata* were incorrectly identified as *C. lenticula* and *C. cingulatum*, respectively. Here we add data (COI and 16S) for the real *C. lenticula* and *C. cingulatum* to GenBank. Furthermore we correct the wrong conclusions drawn by Steinke et al. (2004) on the bases of their misidentifications. *Caracollina* does not belong to the Ariantinae but most likely to the Trissexodontinae, in agreement with the usual classification. The so-called *C. lenticula* sequences of Steinke et al. (2004) show that *H. lapicida andorrnica* belongs to the Ariantinae, again in agreement with the usual classification.

In other cases we have shown that the sequences deposited in GenBank by Steinke et al. (2004) may not belong to the specified taxon, even in cases where the vouchers (shells) were identified correctly. The COI and 16S sequences that are allegedly of *H. lapicida* do certainly not refer to that species, despite the correctly identified shell voucher. In other instances (e.g. *O. lactea*, *C. acuta*) only one of the two sequences (of the COI and 16S datasets examined here) turned out to be correct, so errors had to be made either with the molecular analyses or with submission of the sequences. For a number of taxa (e.g. *T. elegans*, *M. cartusiana*, *X. obvia* and *C. neglecta*) the sequences deposited in GenBank by Steinke et al. (2004) differ strongly from those published by other authors, which are in conformity then with the usual classification, whereas the results published by Steinke et al. (2004) are not. A fundamentally incorrect and potentially misleading procedure employed by Steinke et al.

(2004) is the use of what is here introduced as a 'pseudo-voucher', i.e. a part (like a shell) of an individual that differs from the specimen that was actually used for DNA analysis. Steinke et al. (2004) apparently sequenced *Iberus marmorana* from Ronda, but while referring to that locality, deposited a shell of a conchologically somewhat similar species from elsewhere, i.e. *I. gualtierianus*, as its 'voucher'. As an alternative for the lack of an actual voucher (which indeed might be destroyed during DNA extractions) we here introduce the term 'para-voucher', which is (a part of) a different specimen than the one used for DNA analysis, but from the same sample or population. Although for the Ariantinae sensu Steinke et al. (2004) the sequences for *A. arbustorum* and *I. isognomostomos* are after all correct, still three out of the five taxa of that subfamily were either identified incorrectly or were published with incorrect sequences.

Steinke et al. (2004) published a major paper on the phylogeny of the Gastropoda, Pulmonata. We intend to prevent the further spreading of an inexcusably high number of errors of various kinds in that article, because these incorrect data may result in wrong conclusions, such as already happened in Steinke et al., 2005. In the latter paper *COI* and *16S* sequences of Steinke et al. (2004) were used as a reference dataset (required by the program TaxI) for the identification of juvenile snails. Four of these snails being *C. acuta*, according to the authors. Remarkably the *16S* "divergence-value" (Steinke et al., 2005) between sequences from these snails and the reference dataset was only 0.4%. Given the fact that the 2004 study of Steinke et al. did not include any *16S* sequence of *C. acuta* (this paper), for this to be true would either mean that the authors identified a juvenile *H. lapicida andorrlica* as *C. acuta* four times, or their reference dataset (Steinke et al., 2005) did not consist of sequences AY546342–546381 and AY546262–AY546301 as currently included in GenBank. Either way, we strongly advise to be cautious when using sequence data from Steinke et al. (2004), if any.

## Acknowledgments

We are grateful to the following people for providing material or sequences: B. Kokshoorn for collecting specimens of *Helicigona lapicida andorrlica*, A.S.H. Breure for collecting *Marmorana signata*, A. Martínez-Ortí & M.A. Elejalde for providing *Otala punctata* and P. Kornilios & S. Giokas for providing sequences of *Zonites algirus*. We also like to thank R. Janssen and H. Nordsieck for their testimony regarding the discussion about the voucher material as deposited in the Senckenberg Museum in Frankfurt.

## References

Bank, R.A., Bouchet, P., Falkner, G., Gittenberger, E., Hausdorf, B., von Proschwitz, T., Ripken, T.E.J., 2001. Supraspecific classification of European non-marine Mollusca (CLECOM Sections I+II). *Heldia* 4, 77–128.

Dinapoli, A., Zinssmeister, C., Klusmann-Kolb, A., 2011. New insights into the phylogeny of the Pyramidellidae (Gastropoda). *J. Mollus. Stud.* 77, 1–7. doi:10.1093/mollus/eyq027.

Duda, M., Sattmann, H., Haring, E., Bartel, D., Winkler, H., Harl, J., Kruckenhauser, L., 2011. Genetic differentiation and shell morphology of *Trochulus oreinos* (Wagner, 1915) and *T. hispidus* (Linnaeus, 1758) (Pulmonata: Hygromiidae) in the northeastern Alps. *J. Mollus. Stud.* 77, 30–40.

Elejalde, M.A., Madeira, M.J., Arrebola, J.R., Muñoz, B., Gomez-Moliner, B.J., 2008a. Molecular phylogeny, taxonomy and evolution of the land snail genus *Iberus* (Pulmonata: Helicidae). *J. Zool. Syst. Evol. Res.* 46, 193–202.

Elejalde, M.A., Madeira, M.J., Muñoz, B., Arrébola, J.R., Gómez-Moliner, B.J., 2008b. Mitochondrial DNA diversity and taxa delineation in the land snails of the *Iberus gualtierianus* (Pulmonata, Helicidae) complex. *Zool. J. Linn. Soc.* 154, 722–737.

Fehér, Z., Szabó, K., Bozsó, M., Péntes, Z., 2009. Phylogeny and phylogeography of the *Lozekia-Kovacsia* species group (Gastropoda: Hygromiidae). *J. Zool. Syst. Evol. Res.* 47, 306–314.

Fiorentino, V., Salomone, N., Manganelli, G., Giusti, F., 2010. Historical biogeography of Tyrrhenian land snails: the *Marmorana-Tyrrheniberus* radiation (Pulmonata, Helicidae). *Mol. Phyl. Evol.* 55, 26–37.

Gittenberger, E., Piel, W.H., Groenenberg, D.S.J., 2004. The Pleistocene glaciations and the evolutionary history of the polytypic snail species *Arianta arbustorum* (Gastropoda, Pulmonata, Helicidae). *Mol. Phyl. Evol.* 30, 64–73.

Grande, C., Templado, J., Cervera, J.L., Zardoya, R., 2004. Phylogenetic relationships among *Opisthobranchia* (Mollusca: Gastropoda) based on mitochondrial *cox 1*, *trnV*, and *rnl* genes. *Mol. Phyl. Evol.* 33, 378–388.

Guiller, A., Coutellec-Vreto, M.A., Madec, L., Deunff, J., 2001. Evolutionary history of the land snail *Helix aspersa* in the Western Mediterranean: preliminary results inferred from mitochondrial DNA sequences. *Mol. Ecol.* 10, 81–87.

Haase, M., Misof, B., 2009. Dynamic gastropods: stable shell polymorphism despite gene flow in the land snail *Arianta arbustorum*. *J. Zool. Syst. Evol. Res.* 47, 105–114. doi:10.1111/j.1439-0469.2008.00488.x.

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

Katoh, K., Toh, H., 2008. Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. *BMC Bioinformatics* 9, 13.

Kornilios, P., Poulakakis, N., Mylonas, M., Vardinoyannis, K., 2009. The phylogeny and biogeography of the genus *Zonites* Montfort, 1810 (Gastropoda: Pulmonata): preliminary evidence from mitochondrial data. *J. Mollus. Stud.* 75, 109–117. doi:10.1093/mollus/eyp003.

Maddison, W.P., Maddison, D.R., 2005. MacClade 4.08: Analysis of Phylogeny and Character Evolution. Sinauer Associates, Sunderland, Massachusetts.

Manganelli, G., Salomone, N., Giusti, F., 2005. A molecular approach to the phylogenetic relationships of the western Palearctic *Helicoidea* (Gastropoda: Stylommatophora). *Biol. J. Lin. Soc.* 85, 501–512.

Nordsieck, H., 1987. Revision des Systemen der Helicoidea (Gastropoda: Stylommatophora). *Arch. Moll.* 118, 9–50.

Nylander, J.A.A., 2004. MrModeltest v. 2.2. Evolutionary Biology Centre, Uppsala University.

Parmakelis, A., Pfenninger, M., Spanos, L., Papagiannakis, G., Louis, C., Mylonas, M., 2005. Inference of a radiation in *Mastus* (Gastropoda, Pulmonata, Enidae) on the island of Crete. *Evolution* 59, 991–1005.

Pfenninger, M., Eppenstein, A., Magnin, F., 2003. Evidence for ecological speciation in the sister species *Candidula unifasciata* (Poiret, 1801), *C. rugosiuscula* (Michaud, 1931) (Helicellinae, Gastropoda). *Biol. J. Lin. Soc.* 79, 611–628.

Pennisi, E., 2008. DNA data. Proposal to 'Wikify' GenBank meets stiff resistance. *Science* 319, 1598–1599.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.

Sauer, J., Hausdorf, B., 2009. Sexual selection is involved in speciation in a land snail radiation on Crete. *Evolution* 63, 2535–2546.

Schileyko, A.A., 1991. Taxonomic status, phylogenetic relations and system of the *Helicoidea sensu lato* (Pulmonata). *Arch. Moll.* 120, 187–236.

Soes, D.M., de Winter, A.J., 2005. *Xerotracha apicina* (Lamarck, 1822), een nieuwe immigrant in Nederland (Gastropoda, Pulmonata, Hygromiidae). *Basteria* 69, 81–86.

Steinke, D., Albrecht, C., Pfenninger, M., 2004. Molecular phylogeny and character evolution in the western Palearctic *Helicidae s.l.* (Gastropoda: Stylommatophora). *Mol. Phyl. Evol.* 32, 724–734.

Steinke, D., Vences, M., Salzburger, W., Meyer, A., 2005. TaxI: a software tool for DNA barcoding using distance methods. *Philos. Trans. Roy. Soc. B* 360, 1975–1980. doi:10.1098/rstb.2005.1729.

Swofford, D.L., 2002. PAUP\*. Phylogenetic Analysis Using Parsimony (\* and other methods). Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts.

Van Riel, P., Jordaens, K., Houtte, N., van, Frias Martins, A.M., Verhagen, R., Backeljau, T., 2005. Molecular systematics of the endemic Leptaxini (Gastropoda: Pulmonata) from the Azores islands. *Mol. Phyl. Evol.* 37, 132–143.

Wollscheid-Lengeling, E., Boore, J., Brown, W., Wägele, H., 2001. The phylogeny of *Nudibranchia* (Opisthobranchia, Gastropoda, Mollusca) reconstructed by three molecular markers. *Org. Divers. Evol.* 1, 241–256.

Yamazaki, N., Ueshima, R., Terrett, J., Yokobori, S., Kaifu, M., Segawa, R., Kobayashi, T., Numachi, K.I., Ueda, T., Nishikawa, K., Watanabe, K., Thomas, R.H., 1997. Evolution of pulmonate gastropod mitochondrial genomes: comparisons of gene organizations of Euhadra, Cepaea and Albinaria and implications of unusual tRNA secondary structures. *Genetics* 145, 749–758.

Zhang, Z., Schwartz, S., Wagner, L., Miller, W., 2000. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 7, 203–214.