



## Evolution of microgastropods (Ellobioidea, Carychiidae): integrating taxonomic, phylogenetic and evolutionary hypotheses

Weigand *et al.*

RESEARCH ARTICLE

Open Access

# Evolution of microgastropods (Ellobioidea, Carychiidae): integrating taxonomic, phylogenetic and evolutionary hypotheses

Alexander M Weigand<sup>1\*</sup>, Adrienne Jochum<sup>1</sup>, Rajko Slapnik<sup>2</sup>, Jan Schnitzler<sup>1,3</sup>, Eugenia Zarza<sup>1,3</sup> and Annette Klussmann-Kolb<sup>1,3</sup>

## Abstract

**Background:** Current biodiversity patterns are considered largely the result of past climatic and tectonic changes. In an integrative approach, we combine taxonomic and phylogenetic hypotheses to analyze temporal and geographic diversification of epigeal (*Carychium*) and subterranean (*Zospeum*) evolutionary lineages in Carychiidae (Eupulmonata, Ellobioidea). We explicitly test three hypotheses: 1) morphospecies encompass unrecognized evolutionary lineages, 2) limited dispersal results in a close genetic relationship of geographical proximally distributed taxa and 3) major climatic and tectonic events had an impact on lineage diversification within Carychiidae.

**Results:** Initial morphospecies assignments were investigated by different molecular delimitation approaches (threshold, ABGD, GMYC and SP). Despite a conservative delimitation strategy, carychiid morphospecies comprise a great number of unrecognized evolutionary lineages. We attribute this phenomenon to historic underestimation of morphological stasis and phenotypic variability amongst lineages. The first molecular phylogenetic hypothesis for the Carychiidae (based on COI, 16S and H3) reveals *Carychium* and *Zospeum* to be reciprocally monophyletic. Geographical proximally distributed lineages are often closely related. The temporal diversification of Carychiidae is best described by a constant rate model of diversification. The evolution of Carychiidae is characterized by relatively few (long distance) colonization events. We find support for an Asian origin of *Carychium*. *Zospeum* may have arrived in Europe before extant members of *Carychium*. Distantly related *Carychium* clades inhabit a wide spectrum of the available bioclimatic niche and demonstrate considerable niche overlap.

**Conclusions:** Carychiid taxonomy is in dire need of revision. An inferred wide distribution and variable phenotype suggest underestimated diversity in *Zospeum*. Several *Carychium* morphospecies are results of past taxonomic lumping. By collecting populations at their type locality, molecular investigations are able to link historic morphospecies assignments to their respective evolutionary lineage. We propose that rare founder populations initially colonized a continent or cave system. Subsequent passive dispersal into adjacent areas led to *in situ* pan-continental or mountain range diversifications. Major environmental changes did not influence carychiid diversification. However, certain molecular delimitation methods indicated a recent decrease in diversification rate. We attribute this decrease to protracted speciation.

**Keywords:** Integrative taxonomy, Protracted speciation, Subterranean environment, Cryptic diversity, DNA barcoding, Allopatric diversification, *Carychium*, *Zospeum*, Phylogeny, Gastropoda

\* Correspondence: A.Weigand@bio.uni-frankfurt.de

<sup>1</sup>Department of Phylogeny and Systematics, Institute for Ecology, Evolution and Diversity, Biosciences, Goethe-University Frankfurt, Max-von-Laue Straße 13, 60438, Frankfurt am Main, Germany

Full list of author information is available at the end of the article

## Background

Climatic and geological changes are considered to be major drivers of biological diversification. Many well-characterized radiations were initiated in aftermath of major geologic events [1-3]. Current biodiversity patterns reflect these consequential processes. While taxa with low dispersal ability may be particularly sensitive to changes in their environment, historically-formed patterns within these taxa are known to remain well preserved [4,5]. The taxon Ellobioidea (Gastropoda, Eupulmonata) comprises a group of morphologically and ecologically highly diverse snails, known to have successfully invaded the marine, brackish water and terrestrial habitats [6,7]. Species are traditionally classified into five taxonomic groups, the Pythiidae, the Laemodontidae, the Melampodidae, the Ellobiidae and the Carychiidae. These taxa have been recognized either as families within Ellobioidea or as sub-families within the family Ellobiidae [6-8]. To avoid confusion, we will here refer to the taxon Ellobioidea and to families. Taxonomic descriptions and systematic classifications were exclusively based on morphological (anatomical and conchological) characters while extant species show a mosaic pattern of plesiomorphic and apomorphic features resulting from convergent evolution [6,7,9,10]. Due to the tenuous nature of morphological characters, phylogenetic reconstructions are extremely difficult [6,11]. Moreover, the high degree of homoplasy in morphological characters and frequent low variability has led to the description of approximately 800 species names available in the literature, whereby 250 are likely to be valid [12]. The most comprehensive molecular study for the Ellobioidea suggests a monophyletic origin of the entire group. However, the relationships among the five traditional ellobioid taxa are still unclear [13].

One lineage of the Ellobioidea, the Carychiidae Jeffreys, 1830 has successfully accomplished a complete transition onto land. Extant carychiid snails inhabit aphotic and permanently wet epigeal (*Carychium*) or subterranean (*Zospeum*) environments throughout their Holarctic distribution. This dramatic shift from a marine to a terrestrial habitat has occurred independently of the stylommatophoran land-snails of the Eupulmonata [13,14]. As for all Ellobioidea, taxonomic and systematic descriptions of Carychiidae are based upon characters of the mature shell, which in the case of carychiid gastropods, are suspected to vary according to environmental conditions [15-17]. The first attempt to characterize carychiid taxa using DNA barcodes supported 90% of traditional morphospecies assignments [18]. Nevertheless, the same study only addressed a single population or a few populations per morphospecies. Phenotypic variability (in *Carychium*) and morphological stasis (in *Zospeum*) were identified as potential explanations for discrepancies in morphological and molecular taxonomy. In particular, *Zospeum* displayed

high intraspecific genetic diversity within single morphospecies, as reflected by several cave-endemic evolutionary lineages (ELs). In the case of European *Carychium*, DNA barcoding revealed a previously overlooked taxonomic entity and helped to reevaluate the taxonomic status of questionable morphospecies [18,19]. Additionally, the phenotypically variable *Carychium* shell was shown to span a wide range of shell dimension and proportions, encompassing three morphospecies. Hence, it is very likely that the Carychiidae harbor a considerable number of morphologically unrecognized ELs.

A gastropod's migratory ability is correlated with its shell size [20,21]. Due to their small size, the active dispersal abilities of carychiid snails are highly limited. Slapnik [22] conducted investigations on *Zospeum isselianum* activity within a cave showing moving distances measuring 1 to 15 cm per week (on average 0.7 cm per day). In a comparative phylogeography of two European *Carychium* species, Weigand et al. [23] revealed that the local population structure is often formed by only a few mtDNA haplotypes suggesting that adjacent populations may only exhibit infrequent gene flow. Nevertheless, minute gastropods are well adapted to passive dispersal (e.g. [24,25]), a mechanism that best explains the foundation of transatlantic European *Carychium* populations in North America [26,27] and the postglacial recolonization of Northern Europe [23]. Although passive dispersal events are relatively rare, these, along with limited mobility, could well contribute to a higher incidence of isolated populations and narrow endemism.

Here, we use an integrative approach combining taxonomic, phylogenetic and evolutionary hypotheses to assess the diversification of ELs in Carychiidae. This process first assigned all specimens to phenotype hypotheses. A conservative genetic delimitation method is used to evaluate the initial morphospecies assignments and to identify distinct ELs [28,29]. This approach allows us to include otherwise morphologically unrecognized ELs, whose absence during phylogenetic tree reconstructions could yield misleading results [30]. Based on the identified lineages, a molecular phylogenetic hypothesis for Carychiidae is then reconstructed to test three evolutionary hypotheses: 1) morphospecies encompass unrecognized ELs, 2) limited dispersal results in a close genetic relationship of geographical proximally distributed taxa and 3) major climatic and tectonic events had an impact on lineage diversification within Carychiidae.

## Results

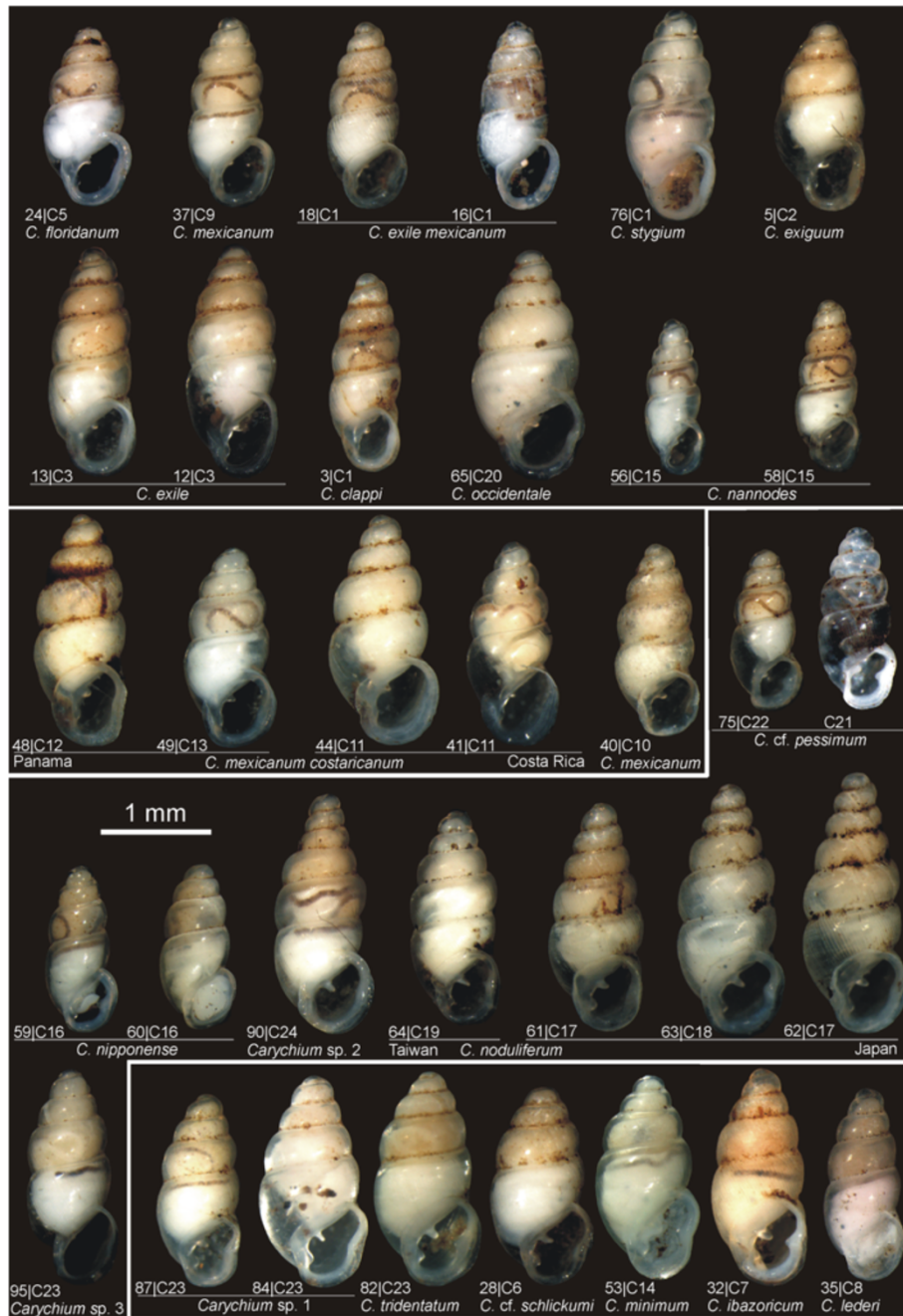
### Molecular identification of evolutionary lineages and species delimitation

To address morphologically unrecognized ELs, the initial morphospecies identifications (Figures 1 and 2, Table 1) were examined applying a combination of five molecular

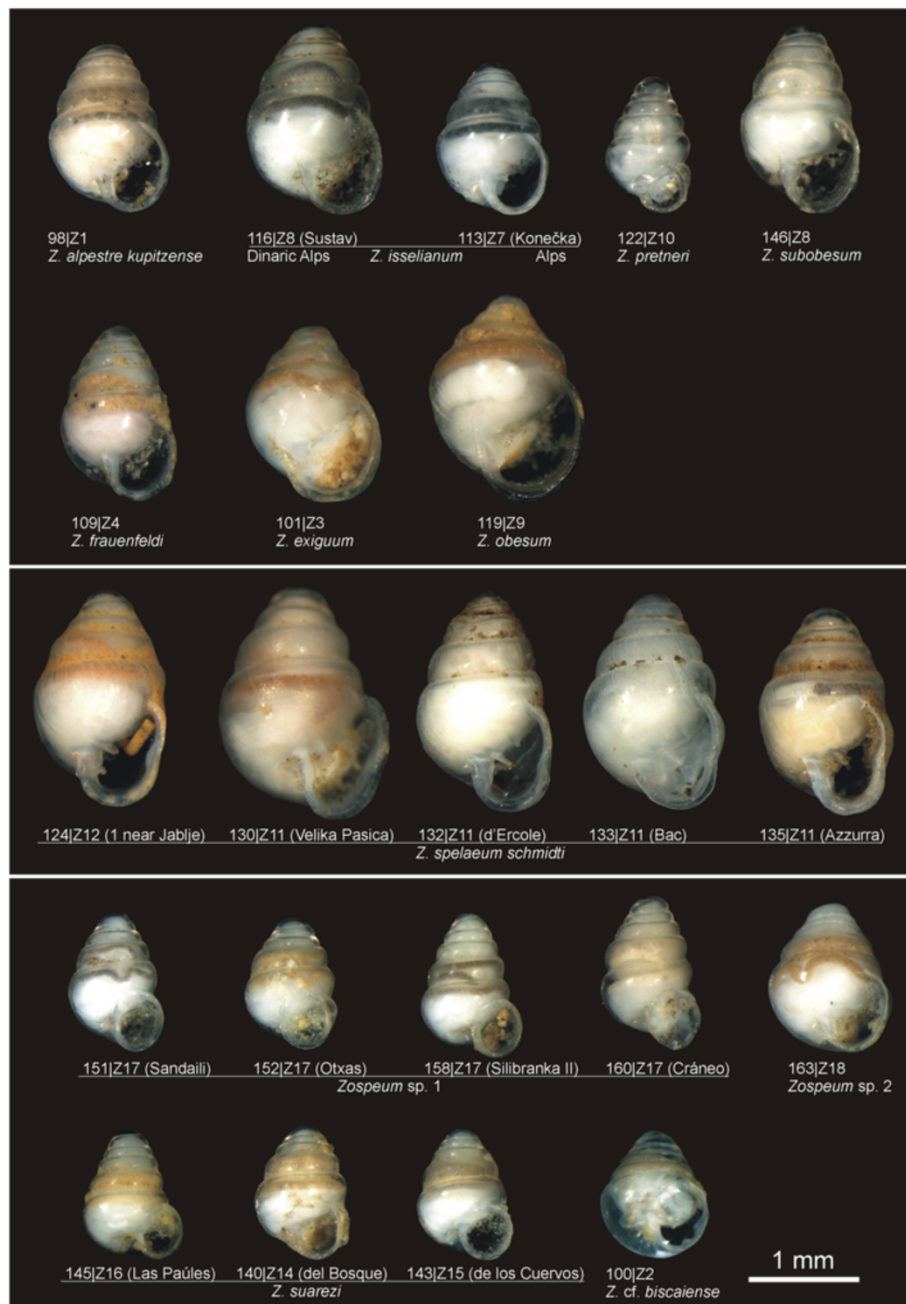


delimitation approaches on mitochondrial sequence data. The 3.2% K2P threshold value identified 45 partitions (Figures 3 and 4). The Automatic Barcoding Gap Detection (ABGD) method consistently revealed 43 ELs for all tested combinations. Two versions of the General Mixed Yule-Coalescent (GMYC) model were applied. Both variants performed significantly better than the null model of a single coalescent population ( $L_0 = 1275.858$ ,

both p-values < 0.0001), but led to a high partitioning of the dataset resulting in 64 ( $L_{GMYCs} = 1315.46$ ) and 78 clusters ( $L_{GMYCm} = 1321.515$ ), respectively. No significant improvement was found when applying GMYCm instead of GMYCs (Chi-square = 12.1081, df = 6 and p = 0.0596). Finally, the Statistical Parsimony (SP) approach delimited 51 ELs mostly congruent to the 3.2% threshold and ABGD results.



**Figure 1** *Carychium* morphospecies and corresponding genetic lineages. Traditionally identified morphospecies and their respective delimited evolutionary lineage (C1-C25) and specimen identifier are visualized. The lineage C21 (*C. cf. pessimum*) is figured by an empty shell of the same population instead of an analyzed individual.



**Figure 2** *Zospeum* morphospecies and corresponding genetic lineages. Traditionally identified morphospecies and their respective delimited evolutionary lineage (Z1-Z18) and specimen identifier are visualized. In some cases the cave name is provided in brackets.

A conservative delimitation strategy was used to combine the partitions of all five approaches. Since all molecular methods tended to split morphospecies, we decided to risk taxonomic lumping, and classified ELs as the most comprehensive grouping of specimens predicted by any of the five delimitation methods. Thus, the 28 carychiid morphospecies (Figures 1 and 2, Table 1) comprised 43 distinct ELs (Figures 3 and 4, Table 2). In total, 17 morphospecies were each recovered as a single

EL (Table 2; 'matches'). The morphologically distinct, but so far, undescribed *Carychium* sp. 2, *Carychium* sp. 3 (both from China), *Zospeum* sp. 1 and *Zospeum* sp. 2 (both from Cantabrian Mts.) were also genetically revealed as separate clusters. Divergent ELs within a single morphospecies hypothesis were found for eight morphospecies (32%): *Carychium exile mexicanum* (2 lineages), *C. mexicanum* (2), *C. mexicanum costaricanum* (3), *C. noduliferum* (3), *C. cf. pessimum* (2), *Zospeum*

**Table 1 Individuals and sampling localities**

morphospecies	#	locality	geographical position		EL
			latitude	longitude	
<i>Carychium</i> O.F. Müller, 1774					
<i>C. clappi</i>	1	USA, Tennessee, La Follette	36.332	-83.998	C1
Hubricht, 1959	2	USA, Tennessee, La Follette	36.332	-83.998	C1
	3	USA, Tennessee, La Follette	36.332	-83.998	C1
	4	USA, Tennessee, La Follette	36.332	-83.998	C1
<i>C. exiguum</i>	5	USA, New York, Naples, Ontario County, Grimes Glen	42.61591	-77.41355	C2
Say, 1822	6	USA, New York, Naples, Ontario County, Grimes Glen	42.61591	-77.41355	C2
	7	USA, New York, Naples, Ontario County, Grimes Glen	42.61591	-77.41355	C2
<i>C. exile</i>	8	USA, Indiana, Lawrence County, Williams Cave Sinkhole	38.7555	-86.59105	C3
H. C. Lea, 1842	9	USA, Indiana, Lawrence County, Williams Cave Sinkhole	38.7555	-86.59105	C3
	10	USA, Indiana, Lawrence County, Williams Cave Sinkhole	38.7555	-86.59105	C3
	11	USA, New York, Portageville, Letchworth State Park	42.57909	-78.04945	C3
	12	USA, New York, Portageville, Letchworth State Park	42.57909	-78.04945	C3
	13	USA, New York, Watkins Glen	42.375828	-76.871115	C3
	14	USA, New York, Watkins Glen	42.375828	-76.871115	C3
	15	USA, New York, Watkins Glen	42.375828	-76.871115	C3
<i>C. exile mexicanum</i>	16	USA, Georgia, Adairsville, Bartow County, Barnsley Gardens	34.311233	-84.9866	C1
Pilsbry, 1891	17	USA, Georgia, Adairsville, Bartow County, Barnsley Gardens	34.311233	-84.9866	C1
	18	USA, Alabama, Little River Mouth Park, Cherokee County	34.312233	-85.685733	C1
	19	USA, Alabama, Little River Mouth Park, Cherokee County	34.312233	-85.685733	C1
	20	USA, Illinois, Jackson County, Gorham	37.6865	-89.490667	C1
	21	USA, Florida, Marianna	30.810556	-85.226667	C4
<i>C. floridanum</i>	22	USA, Florida, Wakulla Springs	30.23548	-84.303087	C5
Clapp, 1918	23	USA, Florida, Wakulla Springs	30.23548	-84.303087	C5
	24	USA, Florida, Wakulla Springs	30.23548	-84.303087	C5
	25	USA, Florida, Wakulla Springs	30.23548	-84.303087	C5
	26	USA, Florida, Wakulla Springs	30.23548	-84.303087	C5
<i>C. cf. schlickumi</i>	27	Greece, Epirus, Ioánnina Prov., Métsovo	39.8213	21.1294	C6
Strauch, 1977	28	Greece, Epirus, Ioánnina Prov., Métsovo	39.8213	21.1294	C6
	29	Greece, Epirus, Ioánnina Prov., Métsovo	39.8213	21.1294	C6
<i>C. ibazoricum</i>	30	Portugal, Azores, San Miguel, Sete Cidades	37.847033	-25.780217	C7
Bank & Gittenberger, 1985	31	Portugal, Azores, San Miguel, Furnas	37.770383	-25.306583	C7
	32	Portugal, Azores, San Miguel, Furnas	37.770383	-25.306583	C7
	33	Portugal, Estremadura, near Sao Pedro de Moel	39.774167	-9.016667	C7
	34	Portugal, Estremadura, near Sao Pedro de Moel	39.774167	-9.016667	C7
<i>C. lederi</i> O. Boettger, 1880	35	Iran, Mazandaran, Nowshahr, Kheiroudkanar forest	36.605833	51.568333	C8
<i>C. mexicanum</i>	36	USA, Georgia, Flovilla, Butts County, Indian Springs State Park	33.242367	-83.92035	C9
Pilsbry, 1891	37	USA, Georgia, Flovilla, Butts County, Indian Springs State Park	33.242367	-83.92035	C9
	38	USA, Georgia, Flovilla, Butts County, Indian Springs State Park	33.242367	-83.92035	C9
	39	Belize, Maya Mountains, Bladen Nature Reserve	16.557167	-88.707833	C10
	40	Belize, Maya Mountains, Bladen Nature Reserve	16.557167	-88.707833	C10
<i>C. mexicanum costaricanum</i>	41	Costa Rica, Puntarenas, Santa Elena	10.369833	-84.804167	C11
Von Martens, 1898	42	Costa Rica, Puntarenas, Monte Verde	10.301333	-84.790167	C11
	43	Costa Rica, San José, San Gerardo de Dota	9.5496	-83.8032	C11

**Table 1 Individuals and sampling localities (Continued)**

	44	Costa Rica, San José, San Gerardo de Dota	9.5496	-83.8032	C11
	45	Costa Rica, San José, San Gerardo de Dota	9.5496	-83.8032	C11
	46	Panama, Chiriquí, Sendero el Retoño, Parque internacional la Amistad	8.8913	-82.628611	C12
	47	Panama, Chiriquí, Sendero el Retoño, Parque internacional la Amistad	8.88815	-82.620667	C12
	48	Panama, Chiriquí, Sendero el Retoño, Parque internacional la Amistad	8.88815	-82.620667	C12
	49	Panama, Chiriquí, near Boquete	8.824767	-82.495833	C13
	50	Panama, Chiriquí, near Boquete	8.824767	-82.495833	C13
	51	Panama, Chiriquí, near Boquete	8.824767	-82.495833	C13
	52	Panama, Chiriquí, near Boquete	8.824767	-82.495833	C13
<i>C. minimum</i>	53	Poland, Stubice	52.3467	14.5806	C14
O.F. Müller, 1774	54	Italy, Upper Adige, near Auer	46.3385	11.3503	C14
	55	Spain, Asturias, Gijón	43.5204	-5.6164	C14
<i>C. nannodes</i>	56	USA, Tennessee, Overton County, Slit Cave	36.456833	-85.375067	C15
G.H. Clapp, 1905	57	USA, Tennessee, Unicoi County, near Unaka Springs	36.0986	-82.4466	C15
	58	Canada, Ontario, Crawford Lake	43.4712	-79.9464	C15
<i>C. nipponense</i>	59	Japan, Honshu Is., Iwate, Ichinoseki, Geibikei	38.9845	141.255667	C16
Pilsbry & Hirase, 1904	60	Japan, Honshu Is., Fukushima, Urabandai	37.687783	140.1205	C16
<i>C. noduliferum</i>	61	Japan, Honshu Is., Iwate, Ichinoseki, Geibikei	38.9845	141.255667	C17
Reinhardt, 1877	62	Japan, Honshu Is., Fukushima, Urabandai	37.687783	140.1205	C17
	63	Japan, Honshu Is., Kanagawa, Yamakita, Oomatazawa	35.423333	139.005	C18
	64	Taiwan, Taichung County, Heping, Shei-Pa N.P., Huanshan	24.37295	121.310567	C19
<i>C. occidentale</i>	65	USA, Washington, Mason County	47.3038	-123.0918	C20
Pilsbry, 1891	66	USA, Washington, Mason County	47.3038	-123.0918	C20
	67	USA, Washington, Mason County	47.3038	-123.0918	C20
	68	USA, Washington, Mason County	47.3038	-123.0918	C20
	69	USA, Washington, Mason County	47.3038	-123.0918	C20
<i>C. cf. pessimum</i>	70	Russia, Vladivostok, Promorsky Kray	43.193417	132.0511	C21
Pilsbry, 1902	71	Russia, Vladivostok, Promorsky Kray	43.193417	132.0511	C21
	72	Russia, Vladivostok, Promorsky Kray	43.193417	132.0511	C21
	73	Russia, Vladivostok, Promorsky Kray	43.193417	132.0511	C21
	74	Russia, Vladivostok, Promorsky Kray	43.193417	132.0511	C21
	75	Japan, Honshu Is., Iwate, Sarusawa	39.021167	141.292833	C22
<i>C. stygium</i>	76	USA, Kentucky, Hart County, Horse Cave (Hidden River Cave)	37.174167	-85.903833	C1
Call, 1897	77	USA, Kentucky, Hart County, Horse Cave (Hidden River Cave)	37.174167	-85.903833	C1
	78	USA, Tennessee, Slit Cave (cave entrance)	36.456833	-85.375067	C1
	79	USA, Tennessee, Slit Cave (cave entrance)	36.456833	-85.375067	C1
<i>C. tridentatum</i>	80	Switzerland, St. Gallen, St. Gallen	47.36	9.14	C23
(Risso, 1826)	81	Italy, Upper Adige, near Auer	46.3385	11.3503	C23
	82	France, Brittany, near Lopreden	48.596667	-3.97	C23
<i>Carychium</i> sp. 1	83	Bulgaria, Mostovo Village, entrance Gargina Dupka Cave (Rhodopi Mnt.)	41.850667	24.926183	C23
	84	Bulgaria, Mostovo Village, entrance Gargina Dupka Cave (Rhodopi Mnt.)	41.850667	24.926183	C23

**Table 1 Individuals and sampling localities (Continued)**

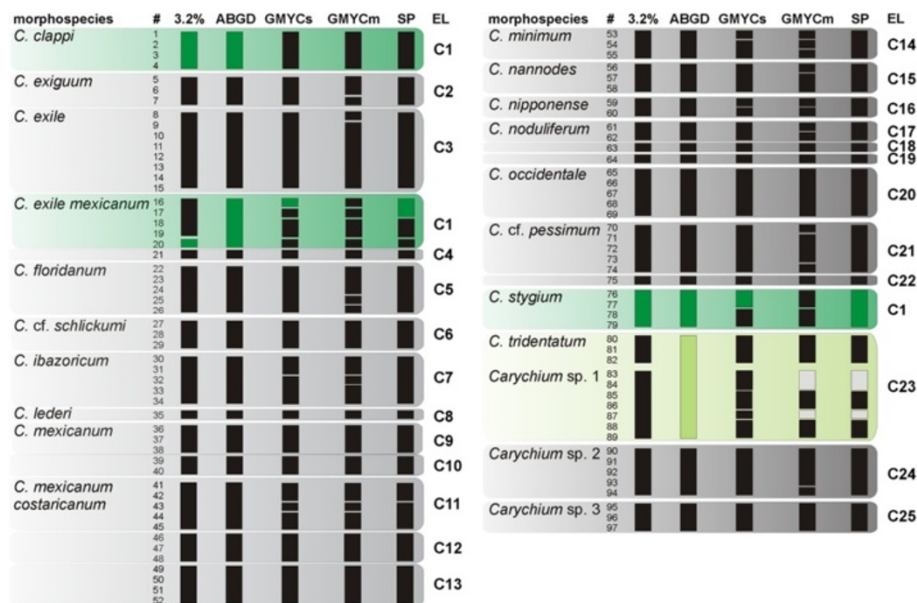
	85	Italy, Lombardy, Lago Maggiore, Laveno-Mombello	45.90718	8.6625	C23
	86	Italy, Lombardy, Lago d'Iseo, Vigolo	45.7367	10.0025	C23
	87	Romania, Bacăia	46.0249	23.1734	C23
	88	Georgia, Kakheti Province, Lagodekhi Nature Reserve	41.88	46.31	C23
	89	Georgia, Kakheti Province, Lagodekhi Nature Reserve	41.88	46.31	C23
<i>Carychium</i> sp. 2	90	China, Yunnan, Zhongdian County, Shangri-La	27.501017	100.033833	C24
	91	China, Yunnan, Zhongdian County, Shangri-La	27.501017	100.033833	C24
	92	China, Yunnan, Zhongdian County, Shangri-La	27.501017	100.033833	C24
	93	China, Yunnan, Zhongdian County, Shangri-La	27.501017	100.033833	C24
	94	China, Yunnan, Zhongdian County, Shangri-La	27.501017	100.033833	C24
<i>Carychium</i> sp. 3	95	China, North Sichuan, Songpan County, Huanglong	32.7373	103.824383	C25
	96	China, North Sichuan, Songpan County, Huanglong	32.7373	103.824383	C25
	97	China, North Sichuan, Songpan County, Huanglong	32.7373	103.824383	C25
<b><i>Zospeum</i> Bourguignat, 1856</b>					<b>cave name</b>
<i>Z. alpestre kupitzense</i>	98	Slovenia, Kamnik-Savinja Alps, Solčava	Ložekarjeva zijalka	Z1	
A. Stummer, 1984	99	Slovenia, Kamnik-Savinja Alps, Solčava	Ložekarjeva zijalka	Z1	
<i>Z. cf. biscaiense</i> Gómez & Prieto, 1983	100	Spain, Garaimendi, Yurre	Cueva de Otxas*	Z2	
<i>Z. exiguum</i>	101	Slovenia, Cerknica, Lož	Križna jama*	Z3	
Kusčar, 1932	102	Slovenia, Cerknica, Lož	Križna jama*	Z3	
	103	Slovenia, Cerknica, Lož	Križna jama*	Z3	
	104	Slovenia, Cerknica, Lož	Križna jama*	Z3	
	105	Slovenia, Cerknica, Lož	Križna jama*	Z3	
<i>Z. frauenfeldi</i>	106	Slovenia, Dobropolje, Podpeč	Podpeška jama*	Z4	
(Freyer, 1855)	107	Slovenia, Dobropolje, Podpeč	Podpeška jama*	Z4	
	108	Slovenia, Dobropolje, Podpeč	Podpeška jama*	Z4	
	109	Slovenia, Dobropolje, Podpeč	Podpeška jama*	Z4	
	110	Slovenia, Dobropolje, Podpeč	Podpeška jama*	Z4	
<i>Z. isselianum</i>	111	Slovenia, Kamnik, Kamniška Bistrica	Jama pod Farjevim plazom	Z5	
Pollonera, 1887	112	Slovenia, Kobarid*, Robič	Turjeva jama	Z6	
	113	Slovenia, Kamnik-Savinja Alps, Šmihel nad Mozirjem	Konečka zijalka	Z7	
	114	Slovenia, Kamnik-Savinja Alps, Šmihel nad Mozirjem	Konečka zijalka	Z7	
	115	Croatia, Kordun, Karlovac, Krnjak, Brebornica Mts.	Jopićeva špilja	Z8	
	116	Croatia, Kordun, Karlovac, Krnjak, Brebornica Mts.	Jopićeva špilja	Z8	
	117	Croatia, Kordun, Karlovac, Krnjak, Brebornica Mts.	Jopićeva špilja	Z8	
	118	Croatia, Kordun, Karlovac, Krnjak, Brebornica Mts.	Jopićeva špilja	Z8	
<i>Z. obesum</i>	119	Slovenia, Gradiček, Krška vas	Krška jama*	Z9	
(Frauenfeld, 1854)	120	Slovenia, Gradiček, Krška vas	Krška jama*	Z9	
	121	Slovenia, Gradiček, Krška vas	Krška jama*	Z9	
<i>Z. pretneri</i> Bole, 1960	122	Croatia, Gračac, Kesići	Donja Cerovačka špilja*	Z10	
<i>Z. spelaum schmidti</i>	123	Slovenia, Veliki Otok, Postojna	Betalov Spodmol	Z11	
(Frauenfeld, 1854)	124	Slovenia, Loka pri Mengšu	Jama 1 pri Jabljah	Z12	
	125	Slovenia, Loka pri Mengšu	Jama 1 pri Jabljah	Z12	
	126	Slovenia, Loka pri Mengšu	Jama 1 pri Jabljah	Z12	
	127	Slovenia, Loka pri Mengšu	Jama 1 pri Jabljah	Z12	



**Table 1 Individuals and sampling localities (Continued)**

	128	Slovenia, Loka pri Mengšu	Jama 1 pri Jabljah	Z12
	129	Slovenia, Krim Region, Gornji Ig	Velika Pasica*	Z11
	130	Slovenia, Krim Region, Gornji Ig	Velika Pasica*	Z11
	131	Italy, Trieste, near Gabrovizza San Primo	Grotte d'Ercole	Z11
	132	Italy, Trieste, near Gabrovizza San Primo	Grotte d'Ercole	Z11
	133	Italy, Trieste, near Basovizza	Grotte Bac	Z11
	134	Italy, Trieste, near Basovizza	Grotte Bac	Z11
	135	Italy, Trieste, near Samatorza	Grotte Azzurra	Z11
<i>Z. suarezi</i>	136	Spain, Cantabria, Novales, La Busta	Cueva del Linar	Z13
Gittenberger, 1980	137	Spain, Cantabria, Novales, La Busta	Cueva del Linar	Z13
	138	Spain, Cantabria, Novales, La Busta	Cueva del Linar	Z13
	139	Spain, Cantabria, Novales, La Busta	Cueva del Linar	Z13
	140	Spain, Asturias, Inguanzo	Cueva del Bosque/Cueva Inguanzo*	Z14
	141	Spain, Asturias, Inguanzo	Cueva del Bosque/Cueva Inguanzo*	Z14
	142	Spain, Bizkaia, San Pedro de Galdames, Barranco de Aranaga	Cueva de los Cuervos	Z15
	143	Spain, Bizkaia, San Pedro de Galdames, Barranco de Aranaga	Cueva de los Cuervos	Z15
	144	Spain, Bizkaia, San Pedro de Galdames, Barranco de Aranaga	Cueva de los Cuervos	Z15
	145	Spain, Castilla y León, Monte Santiago	Cueva de Las Paúles	Z16
<i>Z. subobesum</i>	146	Croatia, Ogulin, Tounj	Tounjčica*	Z8
Bole, 1974	147	Croatia, Ogulin, Tounj	Tounjčica*	Z8
	148	Croatia, Ogulin, Tounj	Tounjčica*	Z8
<i>Zospeum</i> sp. 1	149	Spain, Bizkaia, Guipúzcoa Mts., Valle de Araotz	Cueva de Ermita de Sandaili	Z17
	150	Spain, Bizkaia, Guipúzcoa Mts., Valle de Araotz	Cueva de Ermita de Sandaili	Z17
	151	Spain, Bizkaia, Guipúzcoa Mts., Valle de Araotz	Cueva de Ermita de Sandaili	Z17
	152	Spain, Garaimendi, Yurre	Cueva de Otxas	Z17
	153	Spain, Garaimendi, Yurre	Cueva de Otxas	Z17
	154	Spain, Garaimendi, Yurre	Cueva de Otxas	Z17
	155	Spain, Garaimendi, Yurre	Cueva de Otxas	Z17
	156	Spain, Garaimendi, Yurre	Cueva de Otxas	Z17
	157	Spain, Bizkaia, Manaria, Urkuleta Valley	Cueva Silibranka II	Z17
	158	Spain, Bizkaia, Manaria, Urkuleta Valley	Cueva Silibranka II	Z17
	159	Spain, Bizkaia, Manaria, Urkuleta Valley	Cueva Silibranka II	Z17
	160	Spain, Bizkaia, Dima, Indusi	Cueva del Cráneo	Z17
<i>Zospeum</i> sp. 2	161	Spain, Castilla y León, Monte Santiago	Cueva de Las Paúles	Z18
	162	Spain, Castilla y León, Monte Santiago	Cueva de Las Paúles	Z18
	163	Spain, Castilla y León, Monte Santiago	Cueva de Las Paúles	Z18
<b>Outgroup taxa</b>				
Pythiidae, <i>Laemodonta cubensis</i>	164	Bermuda, Hamilton Parish, Walsingham Pond, outside Cliff Cave	32.34773	-64.70965
Melampodidae, <i>Microtralia occidentalis</i>	165	Bermuda, St. George Parish, St. George Island, Lover's Lake	32.36750	-64.70990
Veronicellidae, <i>Veronicella cubensis</i>	166	Bermuda, Devonshire Parish, Winfried Gibbons Nature Reserve, South Road	32.30230	-64.74450

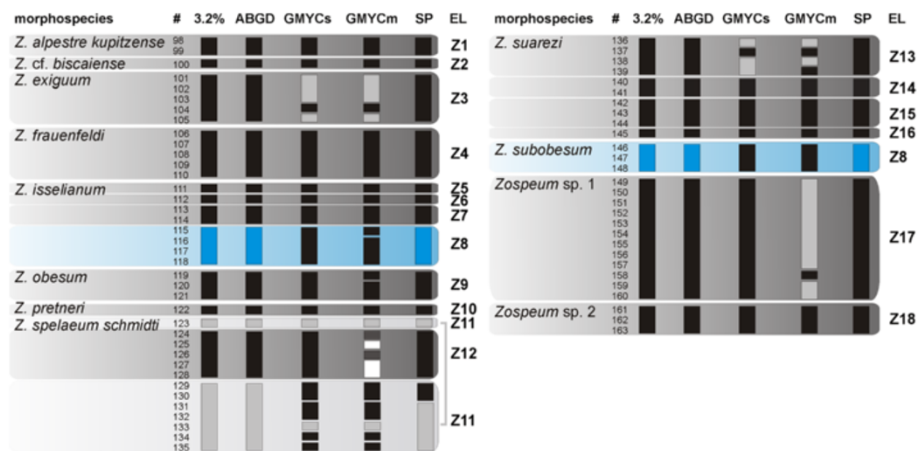
Data on the morphological identification, taxonomic first description, specimen number (#), sampling locality and evolutionary lineage (EL) are provided. Type locality populations or localities referred to in the phenotype descriptions of *Zospeum* morphospecies are marked by an asterisk.



**Figure 3 Molecular delimitation of *Carychium* evolutionary lineages.** Results of the five genetic delimitation approaches are indicated. For *Carychium*, 97 specimens (#) comprising 18 morphospecies and three undescribed taxa are analysed. A conservative strategy identified 25 evolutionary lineages (EL; C1–C25). Specimens clusters identified under each single delimitation method (3.2%, ABGD, GMYCs, GMYCm and SP) are indicated by black boxes. Grey boxes refer to clustered specimens within a single morphospecies (e.g. SP, # '83' + '84' + '87'), colored boxes to specimens between different morphospecies (e.g. SP, # '16' + '17' + '76-79').

*isselianum* (3), *Z. spelaenum schmidti* (2) and *Z. suarezi* (4). Hence, we uncovered at least 20 ELs that could not clearly be distinguished with the initial morphospecies hypotheses (i.e. 47% of all ELs). For *Zospeum isselianum*, *Z. spelaenum schmidti*, *Z. suarezi*, *Z. subobesum* and *Carychium mexicanum costaricanum*, our analysis of topotypic populations enabled us to link a single EL to the initial morphospecies hypothesis (Tables 1 and 2). The three

morphospecies, *Zospeum isselianum*, *Z. spelaenum schmidti* and *Z. suarezi* were found over a large geographical area and possessed moderately variable shell phenotypes, two characteristics that make them prime candidates for taxonomic lumping. Specimens of *C. exile mexicanum*, *C. clappi* and *C. stygium* shared closely related barcodes and consequently, were treated as a single taxon (C1). The European *Carychium* sp. 1 (C23)



**Figure 4 Molecular delimitation of *Zospeum* evolutionary lineages.** Results of the five genetic delimitation approaches are indicated. For *Zospeum*, 66 specimens (#) comprising 10 morphospecies and two undescribed taxa are analysed. A conservative strategy identified 18 evolutionary lineages (EL; Z1–Z18). Specimens clusters identified under each single delimitation method (3.2%, ABGD, GMYCs, GMYCm and SP) are indicated by black boxes. Grey boxes refer to clustered specimens within a single morphospecies (e.g. 3.2%, # '123' + '129-135'), colored boxes to specimens between different morphospecies (e.g. 3.2%, # '115-118' + '146-148').

**Table 2 Integrative identification of evolutionary lineages in Carychiidae**

EL	cross-validation with MA
C1	includes <i>C. clappi</i> , some specimens of <i>C. exile mexicanum</i> and <i>C. stygium</i>
C2	matches <i>C. exiguum</i>
C3	matches <i>C. exile</i>
C4	lumped with specimens from C1 as <i>C. exile mexicanum</i>
C5	matches <i>C. floridanum</i>
C6	matches <i>C. cf. schlickumi</i>
C7	matches <i>C. ibazoricum</i>
C8	matches <i>C. lederi</i>
C9	lumped with C10 as <i>C. mexicanum</i>
C10	lumped with C9 as <i>C. mexicanum</i>
C11	regarded as <i>C. costaricanum</i> ; lumped with C12 + C13 as <i>C. mexicanum costaricanum</i>
C12	lumped with C11 + C13 as <i>C. mexicanum costaricanum</i>
C13	lumped with C11 + C12 as <i>C. mexicanum costaricanum</i>
C14	matches <i>C. minimum</i>
C15	matches <i>C. nannodes</i>
C16	matches <i>C. nipponense</i>
C17	lumped with C18 + C19 as <i>C. noduliferum</i>
C18	lumped with C17 + C19 as <i>C. noduliferum</i>
C19	lumped with C17 + C18 as <i>C. noduliferum</i>
C20	matches <i>C. occidentale</i>
C21	lumped with C22 as <i>C. cf. pessimum</i>
C22	lumped with C21 as <i>C. cf. pessimum</i>
C23	includes <i>C. tridentatum</i> and <i>Carychium</i> sp. 1
C24	matches <i>Carychium</i> sp. 2
C25	matches <i>Carychium</i> sp. 3
Z1	matches <i>Z. alpestre kupitzense</i>
Z2	matches <i>Z. cf. biscaiene</i>
Z3	matches <i>Z. exiguum</i>
Z4	matches <i>Z. frauenfeldi</i>
Z5	lumped with Z6 + Z7 as <i>Z. isselianum</i>
Z6	regarded as <i>Z. isselianum</i> ; lumped with Z5 + Z7 as <i>Z. isselianum</i>
Z7	lumped with Z5 + Z6 as <i>Z. isselianum</i>
Z8	matches <i>Z. subobesum</i> ; specimens 146–148 identified as <i>Z. isselianum</i>
Z9	matches <i>Z. obesum</i>
Z10	matches <i>Z. pretneri</i>
Z11	regarded as <i>Z. spelaenum schmidti</i> ; lumped with Z12 as <i>Z. spelaenum schmidti</i>
Z12	lumped with Z11 as <i>Z. spelaenum schmidti</i>
Z13	lumped with Z14, Z15 + Z16 as <i>Z. suarezi</i>
Z14	regarded as <i>Z. suarezi</i> ; lumped with Z13, Z15 + Z16 as <i>Z. suarezi</i>
Z15	lumped with Z13, Z14 + Z17 as <i>Z. suarezi</i>
Z16	lumped with Z13, Z14 + Z15 as <i>Z. suarezi</i>

**Table 2 Integrative identification of evolutionary lineages in Carychiidae (Continued)**

Z17 matches *Zospeum* sp. 1

Z18 matches *Zospeum* sp. 2

Overview of cross-validated evolutionary lineages (EL) and initial morphospecies assignments (MA). 'matches' = direct match between EL and initial MA; 'lumped' = more than one EL in initial MA; 'includes' = EL includes more than one morphologically delineated taxon; 'regarded as' = used for a 'lumped' EL in case type locality populations have been analyzed.

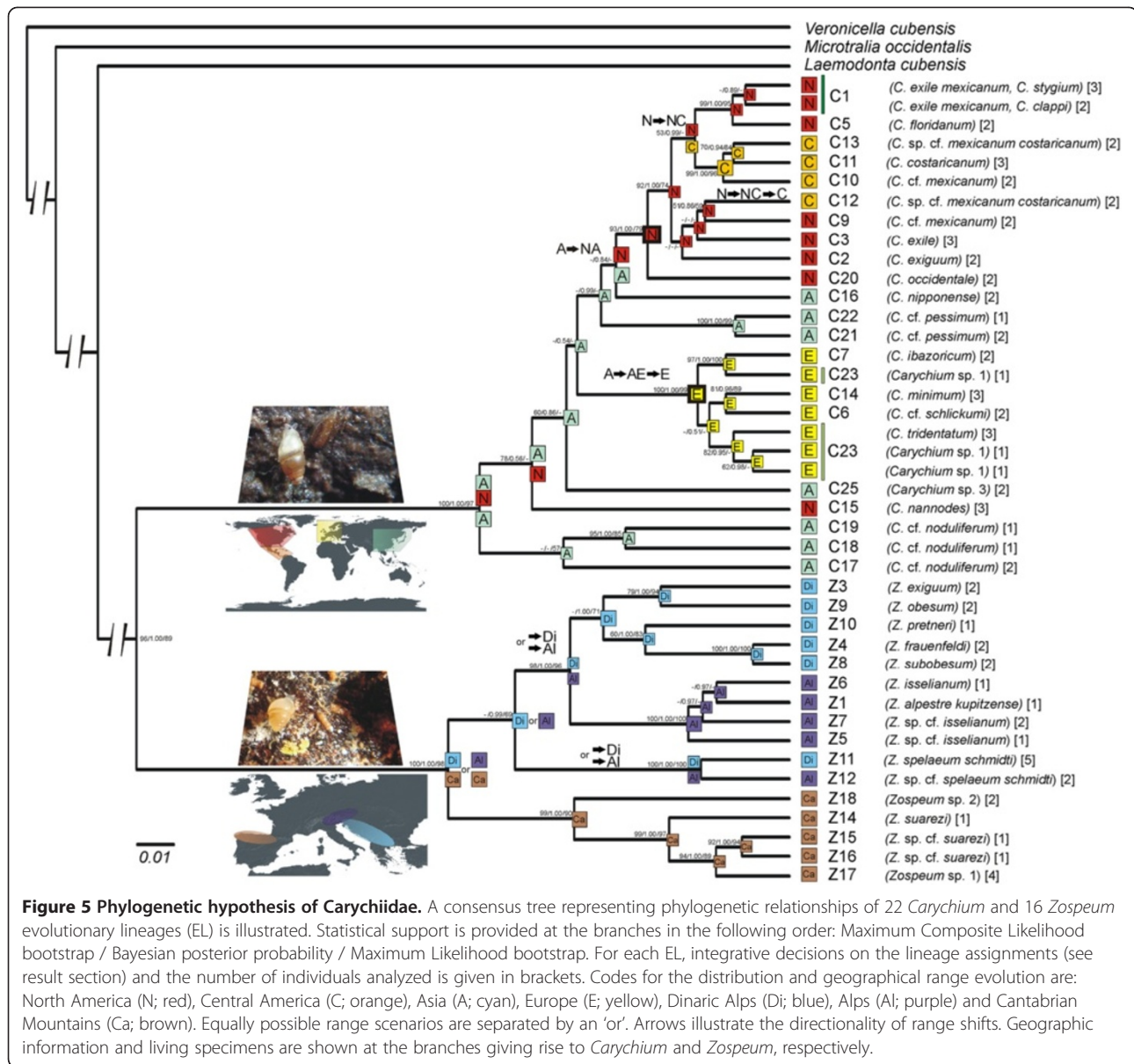
needs further consideration since results of the barcoding approach and phylogenetic reconstruction were not congruent. Thus, two ELs (C1 and C23) included more than one morphospecies each.

### Phylogenetic tree hypothesis

Phylogenetic relationships of 38 ELs based on 1210 bp of nuclear (H3) and mitochondrial sequence data (COI and 16S) were estimated with three different phylogenetic inference methods. Maximum Likelihood (ML), Maximum Composite Likelihood (MCL) and Bayesian inference (BI) yielded mainly congruent results (Figure 5). Lineage assignments correspond to the identification scheme for ELs described in the previous chapter however, with the following exceptions: morphologically unrecognized ELs, for which topotypic populations of the same morphospecies were analysed, were marked with 'sp. cf.' and the morphospecies name (e.g. Z16, *Z. sp. cf. suarezi*). The EL of the same morphospecies, which included specimens from the type locality or from localities used for the phenotype description, was named after the morphospecies hypothesis (e.g. Z14, *Z. suarezi* from Cueva Inguanzo referred to in the first description [31]). Due to the molecular distinctiveness of *C. costaricanum* (C11) from its type locality [32], we address this taxon at the species level.

The morphologically and ecologically defined designation into *Carychium* and *Zospeum* was supported by molecular data (MCL: 100; PP: 1.00; ML  $\geq$  97). All analysed mitochondrial ELs (C1-25 and Z1-18, Figure 5) were monophyletic after incorporation of ncDNA with the exception of *Carychium* sp. 1 (C23). However, sister relationship between lineages were not resolved entirely. Only the relationship of one individual of *Carychium* sp. 1 to *C. ibazoricum* (C7) was highly supported (97; 1.00; 100). Another cluster formed by two lineages of *Carychium* sp. 1 and *C. tridentatum* was only weakly supported (82; 0.95; -).

Except for *C. nannodes*, the North + Central (N+C) American taxa can be traced back to one most recent common ancestral lineage (93; 1.00; 79). The only *Carychium* taxon with a distribution along the West Coast of North America, *C. occidentale* forms the sister group to all other taxa in North and Central America



(92; 1.00; 79). Extant European *Carychium* have descended from a single lineage (100; 1.00; 99).

Well resolved divergence events for ELs of *Carychium* comprised: i) a clustering of C1 + C5, including the morphospecies *C. exile mexicanum*, *C. stygium*, *C. clappi* and *C. floridanum* (99; 1.00; 95), ii) a clade formed by three C-American lineages C10 + C11 + C13 and the morphospecies *C. cf. mexicanum*, *C. costaricanum* and *C. cf. mexicanum costaricanum* (99; 1.00; 96) and iii) the sister group relationship of *C. minimum* (C14) and *C. cf. schlickumi* (C6) (81; 0.96; 89).

Support values for diversification events within *Zospeum* were high. *Zospeum* from the Cantabrian Mountains were monophyletic (99; 1.00; 90). A sister group relationship between *Zospeum* sp. 2 (Z18) and a clade comprising the

morphospecies *Z. suarezi* and *Zospeum* sp. 1 (Z17) was revealed (99; 1.00; 97). Within this clade, three distinct ELs form the morphospecies *Z. suarezi* (Z14-Z16), which was paraphyletic with respect to *Zospeum* sp. 1 (94; 1.00; 89). The widely distributed morphospecies *Z. spelaenum schmidti* was monophyletic but contained two deeply separated ELs (100; 1.00; 100). Monophyly between the Alpine *Z. isselianum* (Z5-Z7) and *Z. alpestre kupitzense* (Z1) was strongly supported (100; 1.00; 100). The relationship between the Dinaric morphospecies *Z. exiguum*, *Z. obesum*, *Z. pretneri*, *Z. frauenfeldi* and *Z. subobesum* received only partial support (-; 1.00; 71). Sister group relationships between the Dinaric *Z. frauenfeldi* (Z4) and *Z. subobesum* (Z8) (100; 1.00; 100) and between *Z. exiguum* (Z3) and *Z. obesum* (Z9) (79; 1.00; 94) were well



resolved. Remarkably, the Dinaric *Z. isselianum* (specimens 115–118, Z8, see Figure 4) did not cluster with other Alpine *Z. isselianum* lineages (Z5–Z7) but fell within the Dinaric clade. Dinaric *Z. isselianum* were probably misidentified. They clustered with *Z. subobesum* (Z8) for which topotypic specimens have been analyzed (Figures 4 and 5, Tables 1 and 2).

### Biogeographical reconstruction

Geographical range evolution of Carychiidae was reconstructed to evaluate geographical transitions and the presence or absence of geographically monophyletic clades. Results are shown only for the model where taxa inhabit a maximum of two regions (continents or mountain ranges). Differences to alternative models, where the maximum range was allowed to encompass all ranges (3 or 4), were largely restricted to basal nodes. However, in some cases these scenarios resulted in extremely wide ancestral ranges which seems biologically unreasonable. Generally, geographical proximally distributed taxa were closely related. Only few colonization events were discovered. The range reconstruction of the most likely (conservative) scenario for the root node of *Carychium* revealed an ‘Asian’ (A) or ‘Asian + North American’ (A+N) origin (Figure 5).

The oldest diversification events within *Carychium* are characterized by relatively short branches with low statistical support, giving rise mostly to deep Asian lineages. Central America could have been colonized by more than one lineage. The European branch originated out of a long separately evolving lineage. We further tested two alternative tree hypotheses for *Carychium* in a model selection framework: i) constrained monophyletic Asian *Carychium* and ii) constrained monophyletic American *Carychium* (see Additional file 1). Both constraints received higher support than the unconstrained phylogenetic hypothesis. Monophyletic Asian and/or American lineages should not be ruled out (Table 3).

The geographic origin of *Zospeum* could not be unambiguously resolved (Figure 5). Our reconstructions equally supported an ancestral distribution in the ‘Cantabrian Mountains + Alps’ or ‘Cantabrian Mountains + Dinaric Alps’. In comparison to *Carychium*, *Zospeum* arrived much earlier in Europe. Only two colonization events were discovered: Both lineages of *Z. spelaicum schmidti* were geographically restricted to the Alps (Z12) and the Dinaric Alps (Z11), respectively. Moreover, the sister clade (comprised of all other Dinaric and Alpine taxa) demonstrated a similar pattern with separate colonization of the Alps or Dinaric Alps; depending upon the ancestral state. We tested our phylogenetic hypothesis against two constrained hypotheses for *Zospeum*: i) monophyletic Dinaric *Zospeum* and ii) monophyletic Alpine *Zospeum*. Both scenarios

**Table 3 Model selection results of alternative phylogenetic hypothesis testing**

hypothesis	$\Delta$ AICM	ln BF <sub>PS</sub>	ln BF <sub>SS</sub>
monophyletic American <i>Carychium</i>	0.00	0.00	0.00
monophyletic Asian <i>Carychium</i>	−14.07	17.24	17.66
unconstrained phylogeny	−18.96	44.86	45.14
monophyletic Alpine <i>Zospeum</i>	−157.41	85.25	85.81
monophyletic Dinaric <i>Zospeum</i>	−345.81	191.58	190.53

Results of the model selection approach to test the monophyly of geographically closely distributed taxa. Hypotheses are ranked according to model fit.

AICM = modified Akaike Information Criterion; BF<sub>PS</sub> = Bayes Factor based on path simulation sampling; BF<sub>SS</sub> = Bayes Factor based on stepping stone sampling. A  $\Delta$ AICM > 7 and a ln BF > 2.3 are seen as strong model support.

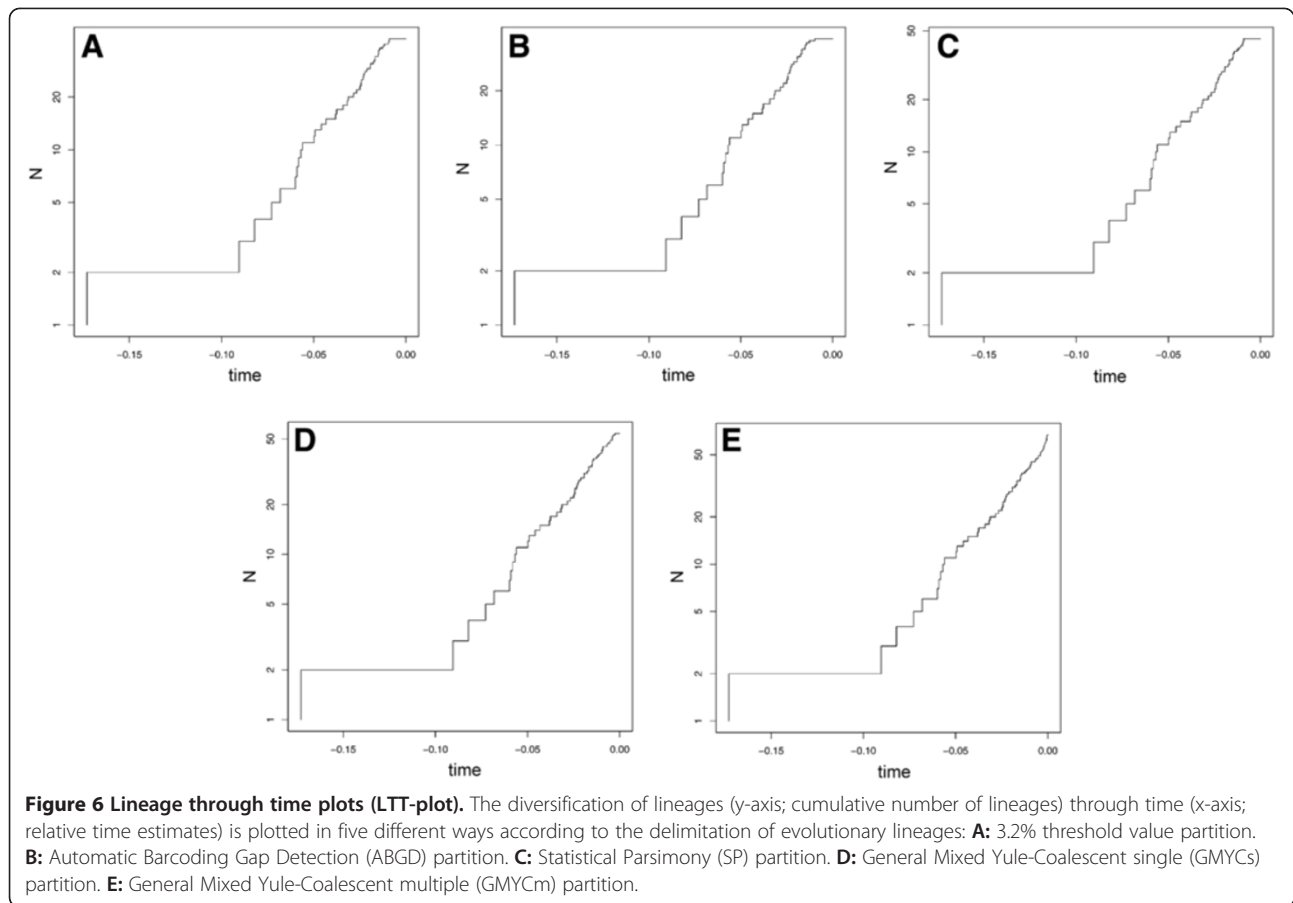
received considerably lower support than the unconstrained hypothesis (Table 3).

### Temporal dynamics of lineage diversification

By fitting different models of diversification using a maximum likelihood approach, temporal dynamics of the diversification of lineages, such as rate shifts due to historic climatic or tectonic changes, were investigated. To account for uncertainties in species delimitation, the analyses were performed on each of the five trees produced by the genetic delimitation approaches (threshold, ABGD, SP, GMYCs and GMYCm). A constant rate model of diversification was preferred for the GMYCm tree (Figure 6, Table 4), whereas a rate-variable model (yule2rate) provided a better fit for all other trees (but not significantly in case of the GMYCs tree). A two-rate pure-birth model best explained the data for the threshold, ABGD and SP trees (Table 4) with a recent decrease (relative shift time approx. −0.01) of the speciation rate.

### Bioclimatic niche differentiation

Bioclimatic niche modeling was performed to study niche differences between two well supported *Carychium* clades and to identify potential climatic factors responsible for lineage persistence and diversification. Maxent models based on occurrence data of European and North + Central American *Carychium* behaved realistically and were statistically well supported (AUC<sub>Europe</sub>: 0.769 ± 0.033; AUC<sub>NorthCentralAmerica</sub>: 0.881 ± 0.016) (see Additional file 2). Bioclimatic niche models for native *Carychium* taxa in America and in Europe, respectively, generally predicted wide areas of suitable habitat (see Additional file 2A, B). The American model was characterized by a sharp longitudinal transition from suitable habitats in the East to unsuitable conditions in the West. Visual inspection of the geographical distribution of all 19 bioclimatic variables throughout this region identified diurnal range (bio2) to best explain the observed pattern (see Additional file 3). Modeling results for Europe highlighted a continuous region of high habitat suitability, stretching from the



Mediterranean Sea to Northern Scandinavia and from the Atlantic Coast to the Black Sea.

Compared to the niche model of the native clade in Europe, the projected model for America suspected European *Carychium* to inhabit regions at lower and higher latitude (see Additional file 2C). Variable degrees of predicted habitat suitability were modeled for populations of non-native European *Carychium minimum* (CM) and *C. tridentatum* (CT) populations in North America (see Additional files 2C and Additional file 4). The modeled latitudinal distribution in Eastern North America was interrupted by a large territory of unsuitable bioclimatic conditions. Similar to the native American clade, this phenomenon could be attributed to changes in the variable states of diurnal range (see Additional file 3). Projection of the North and Central American bioclimatic envelope onto Europe largely corresponded with the distribution range of native *Carychium* taxa (see Additional file 2D). However, a slight shift towards more Mediterranean (to the South) and continental conditions (to the East) was visible.

Bioclimatic niche envelopes of both *Carychium* clades indicated moderate niche overlap in America (D: 0.3878, I: 0.6885, RR: 0.5998; Table 5). Niche

overlap was considerably higher on the European continent (D: 0.6162, I: 0.8570, RR: 0.6589). Niche breadth, i.e. the flatness of the distribution of suitability scores (D. Warren, pers. comm.) differed between both clades with the American clade demonstrating lower values of niche breadth throughout both regions ( $B1_{\text{America}} < B1_{\text{EuropeOntoAmerica}}$ ;  $B1_{\text{AmericaOntoEurope}} < B1_{\text{Europe}}$ ) (Table 6).

## Discussion

### Unrecognized evolutionary lineages

While most species concepts view a species as ‘a separately evolving metapopulation lineage’ [33], they disagree on which characters should be applied to organize this speciation continuum [33]. In the case of the microgastropod taxon Carychiidae, variable environmental conditions can lead to phenotypic variability (e.g. lineages inhabit large geographical regions), whereas relatively stable environments can lead to morphological stasis (e.g. cave endemic lineages) [34-38]. Given that traditional taxonomic classifications completely relied upon characters of the mature shell, unrecognized ELs had to be expected.

Integration of molecular data in a conservative delimitation approach revealed that the 28 analyzed morphospecies actually comprise 43 distinct ELs. From our results,

**Table 4 Model results for the temporal lineage diversification of *Carychium* and *Zospeum***

model	P	df	mtype	LH	r1	r2	a	xp	k	st	AIC	ΔAIC
<b>threshold</b>												
yule2rate	r1, r2, st	3	RV	212.45	30.83	6.04	n.a.	n.a.	n.a.	0.0117	-418.90	0
DDL	r1, k	2	RV	208.09	37.57	n.a.	n.a.	n.a.	68.32	n.a.	-412.18	6.73
pureBirth	r1	1	RC	206.58	23.71	n.a.	n.a.	n.a.	n.a.	n.a.	-411.16	7.74
DDX	r1, xp	2	RV	206.61	26.44	n.a.	n.a.	0.04	n.a.	n.a.	-409.21	9.69
bd	r1, a	2	RC	206.58	23.71	n.a.	0.00	n.a.	n.a.	n.a.	-409.16	9.74
<b>ABGD</b>												
yule2rate	r1, r2, st	3	RV	188.50	30.78	3.65	n.a.	n.a.	n.a.	0.0142	-371.00	0
DDL	r1, k	2	RV	182.55	39.38	n.a.	n.a.	n.a.	53.63	n.a.	-361.10	9.90
pureBirth	r1	1	RC	180.28	21.97	n.a.	n.a.	n.a.	n.a.	n.a.	-358.56	12.44
DDX	r1, xp	2	RV	180.40	27.75	n.a.	n.a.	0.08	n.a.	n.a.	-356.81	14.20
bd	r1, a	2	RC	180.28	21.97	n.a.	0.00	n.a.	n.a.	n.a.	-356.56	14.44
<b>SP</b>												
yule2rate	r1, r2, st	3	RV	226.28	31.07	4.65	n.a.	n.a.	n.a.	0.0096	-446.56	0
DDL	r1, k	2	RV	221.20	36.80	n.a.	n.a.	n.a.	78.07	n.a.	-438.40	8.16
pureBirth	r1	1	RC	220.00	24.58	n.a.	n.a.	n.a.	n.a.	n.a.	-438.01	8.56
bd	r1, a	2	RC	220.00	24.58	n.a.	0.00	n.a.	n.a.	n.a.	-436.01	10.56
DDX	r1, xp	2	RV	220.00	24.58	n.a.	n.a.	0.00	n.a.	n.a.	-436.01	10.56
<b>GMYCs</b>												
yule2rate	r1, r2, st	3	RV	285.59	31.68	9.37	n.a.	n.a.	n.a.	0.0040	-565.17	0
pureBirth	r1	1	RC	283.47	29.02	n.a.	n.a.	n.a.	n.a.	n.a.	-564.93	0.24
DDX	r1, xp	2	RV	283.78	20.50	n.a.	n.a.	-0.11	n.a.	n.a.	-563.56	1.61
bd	r1, a	2	RC	283.52	26.48	n.a.	0.15	n.a.	n.a.	n.a.	-563.04	2.14
DDL	r1, k	2	RV	283.47	29.44	n.a.	n.a.	n.a.	1927.73	n.a.	-562.94	2.23
<b>GMYCm</b>												
bd	r1, a	2	RC	385.03	16.10	n.a.	0.74	n.a.	n.a.	n.a.	-766.06	0
DDX	r1, xp	2	RV	384.37	12.05	n.a.	n.a.	-0.35	n.a.	n.a.	-764.75	1.31
yule2rate	r1, r2, st	3	RV	385.23	30.56	71.35	n.a.	n.a.	n.a.	0.0042	-764.47	1.60
pureBirth	r1	1	RC	381.35	35.94	n.a.	n.a.	n.a.	n.a.	n.a.	-760.69	5.37
DDL	r1, k	2	RV	381.35	35.94	n.a.	n.a.	n.a.	11.20x10 <sup>-3</sup>	n.a.	-758.69	7.37

Results are shown for the five different molecular delimitation approaches. Model = model name; P = parameters included, df = degrees of freedom, mtype = model type; rate-constant (RC) or rate-variable (RV), LH = model log likelihood, r1 = initial diversification rate, r2 = second diversification rate, a = extinction fraction, xp = the x-parameter from the DDX model, k = the k-parameter from the DDL model, st = shift-time, AIC = Akaike Information Criterion, ΔAIC = delta AIC, the difference in AIC scores between given model and the overall best-fit model, n.a. = not available.

three different taxonomic scenarios can be distinguished: i) ELs directly matching the morphospecies, including morphologically delimited, but so far, undescribed taxa (21 lineages), ii) distinct ELs being lumped into a single morphospecies (8 morphospecies encompassing 20 lineages), and iii) single ELs including more than one morphospecies hypothesis (4 morphospecies, 1 undescribed taxon, 2 lineages). This suggests that basing carychiid taxonomic delineations only on conchological characters may be inadequate. We emphasize that vague taxonomic (under-) descriptions and semantic tradition, with morphospecies concepts often only referring to a few

specimens, have widely neglected aspects such as intraspecific shell variability. In congruence with overlapping, interspecific shell dimensions, this has led to a vague situation in carychiid taxonomy and systematics.

Carychiidae appear to still harbor a considerable amount of undiscovered diversity, especially in biogeographic regions that are underexplored due to political strife, non-access and challenging geography, such as Asia or Central America. For example, East Asian *C. noduliferum*, *C. cf. pessimum* and the Central American *C. mexicanum* and *C. mexicanum costaricanum* morphospecies for which, at least one unrecognized EL was discovered. In *Zospeum*,

**Table 5 Niche overlap within *Carychium* in North America and Europe**

	Niche overlap statistic		
	D	I	relative rank
Overlap in North America	0.3878	0.6885	0.5998
Overlap in Europe	0.6162	0.8570	0.6589

Habitat suitability maps were calculated based on species occurrences in each region (Europe and North America) independently and projected into the other region (North America and Europe, respectively).

similar as in other cave animals [39-41], ELs are likely to be morphologically cryptic, possibly due to morphological stasis.

On the other hand *Carychium exile mexicanum* most likely served as a 'taxonomic lumping bin' built up by the frequent nomenclatural intermixture of ELs of the morphospecies *C. stygium*, *C. clappi*, *C. floridanum*, *C. mexicanum* and *C. mexicanum costaricanum* (e.g. [42-46]). The relevant taxonomic literature is puzzling and contains contradictory statements [42]. In general, individual morphospecies within the *C. exile mexicanum* + *C. stygium* + *C. clappi* EL showed highly similar barcode sequences and are thus, difficult to distinguish by molecular delimitation methods, suggesting that these are probably in the process of speciation [18]. A similar complex situation is evident for *C. tridentatum* and *Carychium* sp. 1, which still need further taxonomic investigation. An alternative explanation for the clustering of morphologically distinguishable taxa may be an artifact created by our conservative genetic delimitation strategy. However, if we expect these 'true' species to be artificially lumped into a single EL, we may equally expect other ELs to include even more 'true' species, e.g. as indicated by the two GMYC models.

Based on our taxonomic investigations, we suggest that apparently widespread and assumedly variable *Zospeum* morphospecies must be revised. Moreover, since several *Zospeum* spp. are already listed as vulnerable or endangered [47], we anticipate high conservation value amidst cave-endemics. Furthermore, we recommend that future molecular analyses should focus on populations collected at the type localities, to link an EL with the historic morphospecies hypothesis. A more comprehensive geographic

**Table 6 Niche breadth statistic for native and projected models**

	Niche breadth
America	0.326
Europe onto America	0.374
Europe	0.549
America onto Europe	0.472

Niche breadth for *Carychium* species in Europe and North America as well as according to the projected niches into the other region is shown.

sampling would most likely uncover yet even more unrecognized ELs.

#### Diversification of evolutionary lineages

Appropriate taxon sampling is crucial for the reconstruction of phylogenetic relationships [30]. Our taxon sampling not only covered large parts of the known Holarctic distribution of the Carychiidae but we integrated molecular data to uncover morphologically unrecognized ELs. As expected, the geographic evolution of Carychiidae suggests that the majority of ELs in a geographic area are phylogenetically more closely related than taxa between distant regions (i.e. continents / mountain ranges). This general pattern is most obvious for the European *Carychium* and Cantabrian *Zospeum*. After the initial colonization of new areas due to rare (long-distance) passive dispersal, the ancestral lineages diversified *in situ*.

Based on five different genetic delimitation approaches, our analyses of diversification modes provided mixed results, and in some cases, indicate that diversification rates may have changed over time. While the two GMYC models favored a constant-rate diversification model, the three more conservative genetic delimitation strategies (threshold, ABGD and SP) point to a relatively recent rate slowdown. Such a rate shift can be interpreted in different ways. First, a decrease in the rate of speciation may be due to diversity dependence, e.g. a niche-filling process, where new species reduce the probability of future speciation events [48]. Second, incomplete taxon sampling can result in a spurious rate slowdown due to an overrepresentation of deeper nodes [49]. Third, since speciation is a continuous process, relatively young divergence events are likely to remain unobserved. This effect has recently been described as 'protracted speciation' [50]. Given that the more conservative delimitation approaches are likely to underestimate the number of ELs, we suggest that the observed pattern is more likely an artifact than an actual slowdown. The other two causes may be difficult to distinguish as they can lead to similar patterns, but we suggest protracted speciation as the most likely cause here. First, while taxon sampling in *Carychium* and *Zospeum* is indeed incomplete, several of the missing morphospecies in fact seem to belong to ELs included in this study (for example, *C. mariae* and *C. riparium*; [19] and Jochum & Weigand, unpublished data). Second, the very recent shift in diversification rates is expected under protracted speciation, but could result from incomplete taxon sampling only if the missing species were those that have originated most recently. After omitting the youngest 5th percentile of total branch lengths (i.e. excluding the most recent 5% of the evolutionary history), a constant-rate model is also favored for the threshold, ABGD and SP partitions (data not shown). This not only highlights the very recent timing of the rate



shift (consistent with protracted speciation), but it also shows that no additional rate shifts throughout the evolutionary history of Carychiidae have occurred, suggesting that major environmental changes did not affect the rate of diversification of this clade.

The geographic evolution of *Carychium* implies an Asian or Asian + North American origin. Asian *Carychium* are the genetically most distinct. Since high regional genetic diversity of a taxon can be taken as evidence for lineage persistence in ancestral areas [51], this provides support for an out-of-Asia hypothesis. Niche models suggest significant overlap and occupancy of a wide range of environmental conditions among two geographically and phylogenetically distant *Carychium* clades (Europe and America). There is, however, no evidence for an intermixture of the East and West North American *Carychium* lineages. Thus, larger-scale bioclimatic factors are likely to have been of lesser importance than microhabitat conditions in the diversification of *Carychium*. Nonetheless, differences in the larger-scale bioclimatic niche can result from independent evolutionary histories in isolated biomes and, affect the marginal distribution of lineages. For example diurnal range patterns, an important factor for North American *Carychium*, might affect the colonization of habitats featuring a high variation between day and night. Temperature extremes, moisture and relative humidity levels in typical microhabitats occupied by *Carychium* (e.g. moist leaf-litter, crevices or superficial subterranean habitats) are much less pronounced than those on the surface [52-54]. In cases where the temperature diurnal range is high, microhabitats will not be able to stabilize a given condition.

Since the available bioclimatic layers do not cover subterranean habitats, the influence of bioclimatic parameters on the distribution and diversification of *Zospeum* could not be addressed. The ancestral area reconstruction indicated the 'Cantabrian Mountains + Alps' or 'Cantabrian Mountains + Dinaric Alps' as the ancestral area. We identified two independent colonizations of the Alps and/or the Dinaric Alps but the geographical directionality of these colonizations could not yet be deduced. The incorporation of a recently (re)discovered Asian cf. *Zospeum* from Chinese and South Korean caves (RS and [55]) may provide further insight into the evolutionary history of *Zospeum*.

The diversification of *Zospeum* is characterized by rare, long-distance colonization events with *in situ* (mountain range) radiations into several (sometimes morphologically cryptic) lineages, occupying isolated cave systems. Troglone cave animals like bats and cave crickets, returning to the surface periodically, represent potential vectors for the passive dispersal of cave-dwelling microgastropods [56,57]. Such an allopatric diversification without phenotypic (and perhaps ecological) divergence is referred to as

morphostatic radiation [58] or non-adaptive radiation [59]. Once established at a new locality, ancestral populations may reach adjacent habitats by floating via underground drainage systems or by active subsurface migration [40]. As an example, seasonal flooding events are known triggers for the wash-out of the cave salamander *Proteus anguinus* [60]. As has been shown for human-dispersed transatlantic populations of *Carychium* [26,27], the transportation of only a few hermaphroditic individuals is needed for a successful population foundation. Finally, the remarkably wide distribution of the Cantabrian *Zospeum* sp. 1 with the presence of identical DNA barcodes in four distant caves (up to 30 km apart) suggests recent long-distance dispersal and merits further investigation.

In contrast to the aforementioned, dispersal colonization, cave lineages can arise from multiple diversification events of surface populations via vicariance colonization [56,61]. The geographic distribution of extant *Zospeum* allows room for speculation about the maximum age and evolutionary history for this group. All caves inhabited by European *Zospeum* are embedded in sediments of the Alpine belt initially formed during the Early Cenozoic Alpine orogeny [62,63]. Our most parsimonious assumption suggests that European *Zospeum* originated no earlier than the beginning of the Early Cenozoic (approximately 65 mya), coinciding with the beginning of the Alpine orogeny. Their last common ancestor (LCA) could have descended from the non-cave-dwelling *Carychiopsis*, for which fossils are known since the European Paleocene (66-56 mya) [64-66].

## Conclusions

Carychiidae harbor a substantial number of morphologically unrecognized ELs. In particular, several of the assumedly widespread cave-dwelling *Zospeum* as well as Asian and Central American *Carychium* species resulted from past taxonomic lumping. Future studies should focus on specimens from type localities to link the ELs with the initial phenotype descriptions. Rare long-distance colonization and *in situ* radiations within the newly inhabited areas, i.e. continents (*Carychium* spp.) and mountain ranges (*Zospeum* spp.) represent likely diversification processes in the Carychiidae. However, East Asia provides a notable exception, exhibiting high regional genetic diversity formed by several distinct *Carychium* lineages: indicating a potential origin for *Carychium* and a potentiality for all the Carychiidae.

Although global climatic conditions could influence distribution, microhabitat structure most likely determined local presence and promoted allopatric diversification. Land invasion and desiccation avoidance by the LCA of Carychiidae could have been achieved by colonizing aphotic, permanently humid microenvironments. The occurrence of a true, cave-dwelling lineage (*Zospeum*) and

sporadically-observed cave populations in *Carychium* only demonstrate ecological extremes for survival during the adaptive shift onto land. Shallow subterranean habitats, providing a connection between surface and subsurface realms, could well have promoted ecological transitions within the Carychiidae. Future species-specific ecological studies will allow the identification of micro-environmental parameters shaping the distribution and promoting lineage diversification.

## Methods

### Sampling and identification

In total, 166 individuals were collected during the years 2007–2012 (Table 1). Specimens were immediately stored in 70–99 % ethanol after collection. Our dataset comprises 28 morphologically-described (sub-) species (referred to as morphospecies) of the Carychiidae, including 18 *Carychium* (Figure 1) and 10 *Zospeum* taxa (Figure 2). Additional data were retrieved from a previous DNA barcoding study [18]. Initial taxonomic assignments are based upon conchological characters using taxonomic first descriptions, expert opinions (AJ and AMW [*Carychium*+*Zospeum*]; E. Gittenberger and R. Bank [European *Carychium*]; Y. Kano [Japanese *Carychium*]; RS and C.E. Prieto [*Zospeum*]) and relevant taxonomic keys (e.g. [42,46,67–69]). Morphospecies assignments are marked with 'cf.' in case of juvenile specimens or tenuous morphological characteristics. In particular, the specimens 27–29 from Epirus (Greece) do not match any description of extant taxa but very much resemble *C. schlickumi* described from the Pliocene [70].

### DNA extraction, PCR and sequencing

DNA extraction was carried out on ethanol-preserved individuals using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) protocol. Shell and visceral material were removed to lower cross-contamination risk. Polymerase chain reactions (PCR) were performed to amplify nuclear Histone 3 (H3), a partial fragment of the mitochondrial 16S rRNA (16S) and the DNA barcoding fragment of the mitochondrial-encoded Cytochrome C Oxidase Subunit I (COI). For COI, we used the standard invertebrate primer pair LCO1490 – 5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3' and HCO2198 – 5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3' [71]. Each 25  $\mu$ L PCR mixture included 1  $\mu$ L (10 pmol) of each primer, 2.5  $\mu$ L 10x PCR buffer, 2  $\mu$ L (100 mM) MgCl<sub>2</sub>, 0.3  $\mu$ L (20 mM) dNTPs, 0.3  $\mu$ L *Taq*-polymerase, 0.25  $\mu$ L (0.5 M) tetramethylammonium chloride, 1.5  $\mu$ L (10 mg / mL) bovine serum albumin, 11.15  $\mu$ L ddH<sub>2</sub>O and 5  $\mu$ L template DNA. PCR cycles were run at the following conditions: 1 min at 95°C, followed by 30 cycles of 30 s at 95°C, 30 s at 52°C and 30 s at 72°C, and finally, 3 min at 72°C. For 16S, we used the

same PCR conditions and the primer pair 16S-H – 5'-CGC CTG TTT ATC AAA AAC AT-3' and 16S-R – 5'-CCG GTC TGA ACT CAG ATC ACG T-3' [72]. For H3, we used the degenerated primer pair H3-F – 5'-ATG GCT CGT ACC AAG CAG AC(ACG) GC-3' and H3-R – 5'-ATA TCC TT(AG) GGC AT(AG) AT(AG) GTG-3' [73]. In principle, the same PCR conditions have been used for the amplification of the H3 marker. Modifications contain: the use of 0.1  $\mu$ L (20 mM) dNTPs, 0.14  $\mu$ L *Taq*-polymerase and no tetramethylammonium chloride and bovine serum albumin. PCR cycles were run at the following conditions: 5 min at 95°C, followed by 34 cycles of 30 s at 95°C, 25 s at 52°C and 45 s at 72°C, and finally, 5 min at 72°C. Visualization of single PCR products was performed on a 1.4% agarose gel. They were cleaned using the GeneJET PCR Purification Kit (Fermentas, St. Leon-Rot, Germany). In cases multiple PCR products were detected, the QIAquick Gel Extraction protocol (Qiagen) was used. PCR products were bidirectionally sequenced using the PCR primer pair (5 pmol) and the BigDye<sup>®</sup> Terminator v.3.1 Cycle Sequencing Kit (Applied Biosystems, Inc.) on an ABI 3730 xl capillary sequencer following the manufacturer's instructions.

### Molecular delimitation strategies

COI sequences of all 166 individuals were aligned using MAFFT 6.814 [74] implemented in the Geneious software under the G-INS -i algorithm proposed for less than 200 sequences with global homology. Ambiguous characters (Ns) were treated as missing data. The alignment was further modified by manual primer deletion and a 3' and 5' trimming conducted with GBLOCKS [75]. The length of the final alignment was 607 bp. DNA barcodes are deposited in the BOLD project 'Barcoding Carychiidae microsnails [BARCA]'.

Initial morphospecies assignments were tested using five genetic delimitation approaches: DNA barcoding via a threshold value [76], the Automatic Barcoding Gap Detection (ABGD) method [77], the General Mixed Yule-Coalescent single (GMYCs) and multiple (GMYCm) models [78,79] and Statistical parsimony network analysis (SP) [80]. A threshold value of 3.2% K2P genetic distance was used to separate intra- and interspecific variability in Carychiidae, as recently established by Weigand et al. [18]. The ABGD method separates DNA sequences based on an automatic procedure of barcode gap discovery. Three user defined input variables are requested: The minimum (Pmin) and maximum intraspecific variability (Pmax), which refer to the area where the barcode gap should be detected; and the minimum gap width (X) which relates to the sensitivity of the method to gap width. We tested model combinations of X ranging from 0.01 to 0.9 with Pmax of 0.001 and 0.9, respectively. All runs were

performed using Kimura (K89) genetic distances and 50 screening steps. The GMYC delimitation method combines phylogenetic and phylogeographical approaches to estimate the number of well-separated entities in a sample. It uses an ultrametric input tree to define partitions according to transitional points between speciation and coalescence within species rates. A model based on a single (GMYC) or multiple (GMYCm) threshold values can be tested. Analyses and model comparisons were performed in the R package 'Splits' using the 'gmyc' and 'compare' functions. The ultrametric input tree was obtained with BEAST v1.7.4. [81]. Statistical parsimony network analysis (SP) is commonly used to cluster haplotypes in a phylogeographical framework [82]. However, an inverse consideration of this method is proposed to allow the delimitation of coalescent populations [80]. The program TCS 1.21 [83] was used to delimitate entities on the basis of 95% statistical confidence (i.e. connection probability).

#### Alignment optimization and phylogenetic tree reconstruction

Phylogenetic hypotheses were reconstructed using a concatenated dataset of three phylogenetic markers (mitochondrial 16S, COI and nuclear H3) resulting in 1210 bp. In total, 86 individuals comprising 26 carychiid morphospecies (17 *Carychium*, 9 *Zospeum*) and 38 ELs as well as three outgroup taxa (Ellobioidea: Pythiidae: *Laemodonta cubensis*; Ellobioidea: Melampodidae: *Microtralia occidentalis* and Veronicelloidea: Veronicellidae: *Veronicella cubensis*) were analyzed for the concatenated dataset (see Additional file 5). Alignment optimization was performed separately for each phylogenetic marker: 16S sequences were aligned with MAFFT 6.814 under the FFT-NS -i x 1000 algorithm implemented in the Geneious software. The initial 16S alignment had a length of 548bp. Primer sequences were deleted and the initial alignment was further modified with GBLOCKS [75] to remove ambiguously aligned internal positions and to trim the alignment at the 5' and 3' ends. The final 16S alignment had a length of 368 bp (67% of the initial alignment). For the nuclear H3 marker, the G-INS -i algorithm was used. Primer deletion and GBLOCKS 5' and 3' trimming resulted in 235 bp of the initial 330 bp (71%). The already trimmed barcoding alignment was used as the COI alignment (607 bp, 93% of initial alignment).

Topologies were estimated under three different phylogenetic reconstruction methods: Maximum Composite Likelihood (MCL), Maximum Likelihood (ML) and Bayesian inference (BI). Runs for MCL were performed in MEGA5 under the pairwise deletion option, a gamma distribution (G) with rate parameter 1 and 1,000 bootstrap replicates. RaxML 7.0.3 [84] was used to estimate the ML topology. To account for varying substitution rates between different loci and nucleotide positions,

three marker-specific partitions under the GTR+G+I substitution model were set. MrModeltest 2.3 was used to distinguish between competing substitution models [85]. A thorough ML bootstrapping with 1,000 replicates was conducted. A Bayesian phylogenetic analysis was performed with MrBayes 3.2.1 [86,87] using *Veronicella cubensis* as outgroup. Three gene partitions were defined keeping the estimation of all parameters of the GTR+G+I model of evolution unlinked during the analysis. Two runs of 2,000,000 generations of the MCMC (Markov Chain Monte Carlo) were executed, sampling every 500 generations. The first 25% of the samples were discarded as burn-in to ensure sampling from the stationary phase of the model runs. The chain temperature parameter was set at 0.1. At the 2,000,000<sup>th</sup> generation, the average standard deviation of split frequencies had already fell below 0.01, thus the analysis was stopped.

The concatenated alignment and the phylogenetic consensus hypothesis are deposited in TreeBASE (<http://purl.org/phylo/treebase/phyloids/study/TB2:S13629>).

#### Alternative hypotheses testing

A model selection approach using BEAST v1.7.4 [81,88] was followed to test the monophyly of geographical closely-distributed taxa. Four taxon sets of species inhabiting a certain geographic area were created and constrained to be monophyletic, i.e. i) American *Carychium*, ii) Asian *Carychium*, iii) Dinaric *Zospeum* and iv) Alpine *Zospeum*. The monophyletic constraints were analyzed independently and compared with the results of the unconstrained topology. The Markov Chain Monte Carlo was run for 30 million generations, sampling trees and parameters every 1,000 generations. After verifying that appropriate effective sample sizes were achieved, three model selection methods were applied: a posterior simulation-based analogue of the Akaike Information Criterion (AICM) [89,90], and Bayes Factors (BF) between marginal likelihoods estimated through Path Sampling (PS) [91] and Stepping Stone Sampling (SS) [92]. We used the settings suggested on the BEAST website ([http://beast.bio.ed.ac.uk/Model\\_selection](http://beast.bio.ed.ac.uk/Model_selection)). Competing topological hypotheses were ranked according to the results of the AICM and the marginal likelihood values obtained with PS and SS. Differences between AIC were calculated ( $\Delta$  AICM) as were Bayes Factors between competing hypotheses. A  $\Delta > 7$  between AICM values of the best ranked hypothesis and the other hypotheses suggests that the latter are very unlikely [93]. A Bayes Factor<sub>in</sub> > 2.3 was considered as strong support for the hypothesis (modified guidelines of [94]).

#### Estimation of relative times of divergence

The scarcity of reliable fossils of *Carychium* hinders the estimation of absolute times of divergence. *Carychium*



*brotianum* De Loriol, 1865 from the Upper Jurassic in France is assumed to be by far the oldest carychiid fossil [95]. Our own investigations of the type material of *Carychium brotianum* deposited in the Musée Cantonal de Géologie Lausanne reject a close affiliation to Carychiidae. We rather regard the extinct lineage *Carychiopsis* Sandberger, 1872 known from the Paleocene until the Neogene or *Carychium munieri* Briart & Cornet, 1889 reported from the Early Paleocene as the oldest representatives of Carychiidae [64-66,96]. In respect to this problem, we chose to estimate only relative ages. The analysis was performed with the program BEAST v1.6.1 [81] using an uncorrelated, relaxed lognormal molecular clock model. The three genetic markers were concatenated but the parameters of the substitution and molecular clock models were independently estimated for each gene partition. The taxa *Veronicella cubensis*, *Microtralia occidentalis* and *Laemodonta cubensis* formed the outgroup taxa set. Trees were sampled from a Birth-Death tree-prior and following the GTR+G+I substitution model. The MCMC was run for 30,000,000 generations sampling trees and parameters every 1,000 generations. Effective sample size and convergence were evaluated in Tracer 1.5 [97]. The first 10% of samples were discarded as burn-in before building the maximum clade credibility (MCC) tree.

#### Temporal patterns of diversification

To assess the temporal dynamics of lineage diversification, we tested several constant- and variable-rate models of diversification using a maximum likelihood approach as implemented in the R package laser [98,99]. Model-fit was evaluated using the Akaike Information Criterion (AIC) [90]. Temporal changes in the net diversification rate were evaluated by calculating the difference in the AIC score between the best fit constant-rate (pure-birth, birth-death) and variable-rate models (DDL, DDX, yule2rate). Since these comparisons are susceptible to a high Type I error rate [98], a null distribution of the test statistic was generated by calculating the  $\Delta$ AIC score for 500 phylogenetic trees simulated under a pure-birth process.

#### Geographic range evolution

The dispersal-extinction-cladogenesis (DEC) model implemented in Lagrange [100] was used to infer geographic range evolution. The model assumes that geographic splits for a given lineage occur along the branches of a topology rather than at bifurcation points. Changes in the geographic range of a given taxon can be accounted for by dispersal events (range expansion) or local extinction (range contraction). Alternative Lagrange analyses were run with the maximum range size set to either two or four (the maximum number of areas in the model). The patterns were largely congruent and only the results of the

conservative approach are presented in which taxa are allowed a maximum range size of two, i.e. simultaneously inhabiting up to two geographic areas (here continents or mountain ranges). Given their low dispersal potential and the high proportion of continent- or cave-endemic taxa, this scenario seems more likely. The character state with the highest relative probability (fraction of the global likelihood) was plotted on the topology. In cases of similar relative probability values, all alternative scenarios were plotted to account for model uncertainty.

#### Bioclimatic niche modeling

Bioclimatic niche models were constructed with the program Maxent 3.3.3 [101]. This approach is based on the principle of least assumptions. In the absence of any further information, it prefers the model with the maximum entropy. Niche parameters are extracted from occurrence points and global environmental layers and combined into taxon-specific bioclimatic envelopes. We used georeferenced data for monophyletic lineages instead of species-specific data because ambiguous species-level identifications of *Carychium* taxa are likely to produce questionable entries in public databases or museum collections. The investigation of the bioclimatic niche and macro-evolutionary changes above the species level has proven suitable in earlier studies [102]. This is understandable as traits that allow taxa to persist tend to be conserved over time [102]. The bioclimatic niches for i) the monophyletic European *Carychium* clade and ii) the monophyletic North + Central American (NC) *Carychium* clade were estimated. In total, 136 sampling points for the European clade were collected either from our own collections or the GBIF database (www.gbif.org). Data from American museum collections (Carnegie Museum of Natural History, Pittsburgh; Field Museum of Natural History, Chicago; Florida Museum of Natural History, Gainesville) and entries within the GBIF database are used as distribution data for the NC-clade resulting in 241 sampling points for this model. All 19 bioclimatic variables of the WorldClim project [103] were used in highest resolution (~30 arc-seconds) in order to discover potential, fine-scale patterns affecting the distribution of carychiid microgastropods. Almost all included georeference points (>98%) were accurate to about less than 1 km, which justifies the implementation of bioclimatic variables in their highest resolution. We performed five cross-validated Maxent runs and considered grid cells with a cumulative probability of more than 10 (from a range of 0–100) as suitable [104,105]. The area under the ROC curve (AUC) gave an evaluation of the projections' overall quality. An AUC score above 0.7 is considered good model performance [106]. Outputs are generalized clade-based models. Such a model design can slightly overestimate taxon-specific



bioclimatic envelopes and can lead to false-positive model results (i.e. predicted habitat suitability but biological absence of taxa).

### Niche similarity

ENMtools [107] was used for the comparative analysis of environmental niche models of the European and NC-clade. Niche overlap between the two clades was assessed using three different similarity statistics: Schoener's D [108], the I statistic [109], and the relative rank test [110]. Schoener's D and I calculate the difference in standardized suitability scores for each grid cell. The relative rank test provides an estimate of the congruence of relative ranks of suitability for each grid cell. All three measurements can have a range between 0 (non-overlapping niches) and 1 (identical niches). To calculate niche similarity, climatic niche models were cross-projected onto the geographic region of the other clade.

### Additional files

**Additional file 1: Figure \*.tif.** Schematic visualization for the constrained model selection approaches. The unconstrained phylogenetic hypothesis was tested against four competing evolutionary scenarios in a model selection approach (refer to Table 3). Both scenarios concerning *Carychium* microgastropods are depicted, since the constraints likewise affected the splitting order of ancient nodes. E = Europe; A = Asia; N+C = North+Central America excl. C. *nannodes*; N = C. *nannodes*. Black dots indicate posterior probability  $\geq 0.98$ . A: phylogenetic unconstrained hypothesis.  $A_{1a} = C.$  cf. *pessimum*;  $A_{1b} = C.$  *nipponense*;  $A_2 = Carychium sp.3;  $A_3 = C.$  cf. *noduliferum*. B: monophyletic Asian *Carychium*. C: monophyletic American *Carychium*.  $A_1 = C.$  cf. *pessimum* + C. *nipponense*;  $A_2 = C.$  cf. *noduliferum* + *Carychium* sp.3.$

**Additional file 2: Figure \*.tif.** Bioclimatic niche models. The suitable bioclimatic conditions of the native monophyletic North + Central American (A; red) and European (B; yellow) clades are illustrated. Niche projections for these clades and between the areas are depicted in C and D. Black triangles indicated occurrence data of native (in A and B) and schematic snails of introduced taxa (in C). A: Potential distribution of the native North + Central American clade. B: Potential distribution of the native European clade. C: Projected distribution of the European clade in North + Central America. Non-native European *Carychium minimum* (CM) and *C. tridentatum* (CT) populations are indicated. D: Projected distribution of the North + Central American clade in Europe.

**Additional file 3: Figure \*.tif.** Distribution of bio2 (diurnal range) in North and Central America (TIFF 807 kb). North and Central American diurnal range values have a minimum of 1.8 (black) and a maximum of 21.4 (white) degrees Celsius (°C).

**Additional file 4: Table \*.xlsx.** Non-native *Carychium* populations in North America. Locality information and literature sources for non-native *Carychium* populations (based on morphospecies IDs) in North America are listed. Specimens from localities indicated by an asterisk (\*) are validated by DNA barcoding as the respective *Carychium* species [111-116].

**Additional file 5: Table \*.xlsx.** BOLD barcode identifier (COI) and NCBI accession numbers (16S and H3). EL = evolutionary lineage; # = specimen number. <sup>1</sup> type locality population, regarded as *C. costaricanum*. <sup>2</sup> cave population from region mentioned in phenotype description of *Z. isselianum*. <sup>3</sup> type locality population of *Z. spelaicum schmidti*. <sup>4</sup> cave population of *Z. suarezi* from cave locality mentioned in phenotype description, <sup>5</sup> H3 sequence of #152 upon request (too short for GenBank deposition).

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

AMW, JS, AJ and AKK designed the study. AJ, AMW and RS collected the specimens. AWM produced the sequences and together with JS acquired the biostatistical data. AMW, JS and EZ conducted the biostatistical analyses. AMW was responsible for the first draft of the manuscript. All authors gave conceptual comments, read and approved the final manuscript.

### Acknowledgements

AKK, EZ and JS were supported by the funding programme 'LOEWE-Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz' from the Ministry of Higher Education, Research, and the Arts, Hesse, Germany. We wish to thank all the following carychiid-savvy biologists, malacologists and museum curators who kindly collected material for us, provided valuable insights, lent us museum material or endured scanning marathons in order to relay us rare literature. Without their generous help, the comprehensive scope of this work could never have been realized: J. Ablett (NHM, UK), R. Bank (Netherlands), M. Bodon (Italy), S. Castillo (Panama), D. Dourson (Belize), R. Egorov (Russia), S. & S. Geiger (Germany), S. Cianfanelli (NHM "La Specola", Italy), J. Gerber (FMNH, USA), B. Gomez (Spain), B. Gregory (USA), A. Favre (Germany), F. Hardie (USA), D. Holyoak (Portugal), C. Holiday (USA), E. Gittenberger (Naturalis, Netherlands), K. Groh (Germany), R. Janssen (SMF, Germany), A. Komerički (HBSD, Croatia), F. Naggs (NHM, UK), J. Nekola (USA), E. Neubert (NMBE, Switzerland), R. Ozimec (Croatia), B. Páll-Gergely (Hungary), S. Pezet (MNHM, France), D. Geogiev (Bulgaria), D. Teixeira (Madeira), I. Richling (SMNS, Germany), J. J. Lewis (USA), J. Slapcinsky (FLNHM, USA), K. Kimura (Japan), C. E. Prieto (Spain), L. Prozorova (Russia), M. Prodan (Italy), M. Schilthuizen (Naturalis, Netherlands), M. Neiber (Germany), M. Horsák (Czech Rep.), S. R. Smith (BNHM, Bermuda), F. Stoch (Italy), F. Giusti (Italy), G. Rosenberg (ANSP, USA), T. Pearce (CMNH, USA), U. Boeßneck (Germany), Y. Kano (AORI, Japan), Z. Fehér (HNHM, Hungary), J. Valentinčič (Slovenia), T. Worsfold (UK), M. Wu (China), E. Zarza (Germany), and C. Zinßmeister (Germany). We also want to thank S. Telle, M. Piepenbring, T. Hofman, I. Richling and O. Caceres for helping us to obtain collection permits in Panama and Costa Rica. Specimens from the latter region refer to museum vouchers MZUCR250, MZUCR251 and MZUCR252. We acknowledge the support of the Bermuda Aquarium, Museum and Zoo as part of the Department of Conservation Services. This is contribution #196, Bermuda Biodiversity Project (BBP), Bermuda Aquarium, Natural History Museum and Zoo. Finally, we highly appreciate the constructive input of two anonymous reviewers and the editor on earlier versions of this manuscript.

### Author details

<sup>1</sup>Department of Phylogeny and Systematics, Institute for Ecology, Evolution and Diversity, Biosciences, Goethe-University Frankfurt, Max-von-Laue Straße 13, 60438, Frankfurt am Main, Germany. <sup>2</sup>Institute of Biology, Center for Scientific Research of the Slovenian Academy of Sciences and Arts, Novi trg 2, p.p. 306, 1000, Ljubljana, Slovenia. <sup>3</sup>Biodiversity and Climate Research Centre, Senckenberganlage 25, 60325, Frankfurt am Main, Germany.

Received: 20 September 2012 Accepted: 14 January 2013

Published: 23 January 2013

### References

1. Albertson RC, Markert JA, Danley PD, Kocher TD: **Phylogeny of a rapidly evolving clade: the cichlid fishes of Lake Malawi, East Africa.** *Proc Natl Acad Sci USA* 1999, **96**:5107-5110.
2. Robe LJ, Loreto ELS, Valente VLS: **Radiation of the, *Drosophilidae*, Diptera) in the Neotropics.** *J Zool Syst Evol Res* 2010, **48**:310-321.
3. Smith BT, Klicka J: **The profound influence of the Late Pliocene Panamanian uplift on the exchange, diversification, and distribution of New World birds.** *Ecography* 2010, **33**:333-342.
4. Avise JC: *Phylogeography: The history of formation of species.* Cambridge: Harvard University Press; 2000.
5. Zeisset I, Beebee TJC: **Amphibian phylogeography: a model for understanding historical aspects of species distributions.** *Heredity* 2008, **101**:109-119.
6. Morton JE: **The evolution of the Elobiidae with a discussion on the origin of the Pulmonata.** *P Zool Soc Lond* 1955, **125**:127-168.

7. De Frias Martins AM: Morphological and anatomical diversity within the Ellobiidae (Gastropoda, Pulmonata, Archaeopulmonata). *Vita Malacologica* 2007, **4**:1–28.
8. Barker GM: *The Biology of Terrestrial Molluscs*. UK: CABI Publishing, Oxon; 2001.
9. De Frias Martins AM: Relationships within the Ellobiidae. In *Origin and evolutionary radiation of the Mollusca*. Edited by Taylor J. Oxford: Oxford University Press, Oxford; 1996:285–294.
10. De Frias Martins AM: Anatomy and systematics of the western Atlantic Ellobiidae (Gastropoda: Pulmonata). *Vita Malacologica* 1996, **37**:163–332.
11. Odhner NH: *Marinula juanensis* n. sp., nebst Bemerkungen über die Systematik der Ellobiiden. *Ark Zool* 1925, **17A**:1–15.
12. White TR, Conrad MM, Tseng R, Balayan S, Golding R, de Frias Martins AM, Dayrat BA: Ten new complete mitochondrial genomes of pulmonates (Mollusca: Gastropoda) and their impact on phylogenetic relationships. *BMC Evol Biol* 2011, **11**:295.
13. Dayrat B, Conrad M, Balayan S, White TR, Albrecht C, Golding R, Gomes S, Harasewych MG, de Frias Martins AM: Phylogenetic relationships and evolution of pulmonate gastropods (Mollusca): new insights from increased taxon sampling. *Mol Phylogenet Evol* 2011, **59**:425–437.
14. Klussmann-Kolb A, Dinapoli A, Kuhn K, Streit B, Albrecht C: From sea to land and beyond: new insights into the evolution of euthyneuran Gastropoda (Mollusca). *BMC Evol Biol* 2008, **8**:57.
15. Bulman K: Shell variability in *Carychium tridentatum* (Risso, 1826) and its importance for infraspecific taxonomy (Gastropoda, Pulmonata: Ellobiidae). *Malak Abh Mus Tierkd Dresden* 1990, **15**:37–50.
16. Nekola JC, Barthel M: Morphometric analysis of the genus *Carychium* in the Great Lakes region. *J Conchol* 2002, **37**:515–531.
17. Jochum A: Evolution and diversity of the troglobitic Carychiidae - A morphological and phylogenetic investigation of the terrestrial ellobioid genera, *Carychium* and *Zospeum*. *Malacologist* 2011, **57**:16–18.
18. Weigand AM, Jochum A, Pfenninger M, Steinke D, Klussmann-Kolb A: A new approach to an old conundrum - DNA barcoding sheds new light on phenotypic plasticity and morphological stasis in microsnails (Gastropoda, Pulmonata, Carychiidae). *Mol Ecol Resour* 2011, **11**:255–265.
19. Weigand AM, Goetze M-C, Jochum A: Outdated but established?! Conchologically driven species delineations in microgastropods (Carychiidae, *Carychium*). *Org Divers Evol* 2012, **12**:377–386.
20. Nekola JC, Barker GM, Cameron RAD, Pokryszko BM: Latitudinal and longitudinal variation of body size in land snail populations and communities. In *Global patterns of body size*. Edited by Smith F, Lyons K. Chicago: University of Chicago Press; In press.
21. Aubry S, Labaune C, Magnin F, Roche P, Kiss L: Active and passive dispersal of an invading land snail in Mediterranean France. *J Anim Ecol* 2006, **75**:802–813.
22. Slapnik R: Activity and movements of *Zospeum isselianum* Pölloner 1886 (Gastropoda, Pulmonata, Carychiidae) in a cave in the Kamniške-Savinjske Alps (Slovenia). *Nat Croat* 2001, **10**:153–162.
23. Weigand AM, Pfenninger M, Jochum A, Klussmann-Kolb A: Alpine crossroads or origin of genetic diversity? Comparative phylogeography of two sympatric microgastropod species. *PLoS One* 2012, **7**(5):e37089.
24. Vagvolgyi J: Body size, aerial dispersal, and origin of the Pacific Land Snail Fauna. *Syst Biol* 1975, **24**:465–488.
25. Gittenberger E, Groenenberg DSJ, Kokshoorn B, Preece RC: Molecular trails from hitch-hiking snails. *Nature* 2006, **439**:409.
26. Pearce TA, Payne SL: First record of the European land snail *Carychium minimum* in Pennsylvania, USA. *Tentacle* 2011, **19**:13–14.
27. Weigand AM, Jochum A: Mollusca, Gastropoda, Ellobioidea, *Carychium minimum* O.F. Müller, 1774: filling gaps. New population record for the State of New York, northeastern United States. *Check List* 2010, **6**:517–518.
28. Puillandre N, Modica MV, Zhang Y, Sirovich L, Boisselier MC, Cruaud C, Holford M, Samadi S: Large-scale species delimitation methods for hyperdiverse groups. *Mol Ecol* 2012, **21**:2671–2691.
29. Haszprunar G: Species delimitations – not ‘only descriptive’. *Org Divers Evol* 2011, **11**:249–252.
30. Heath TA, Hedtke SM, Hillis DM: Taxon sampling and the accuracy of phylogenetic analyses. *J Syst Evol* 2008, **46**:239–257.
31. Gittenberger E: Three notes on Iberian terrestrial gastropods. *Zool Meded* 1980, **55**:201–213.
32. Von Martens E: Land and Freshwater Mollusca. In *Biologia Centrali-Americana*. Edited by Godman FD, Salvin O. London; 1890:1901–353.
33. De Queiroz K: Species concepts and species delimitation. *Syst Biol* 2007, **56**:879–886.
34. Trussell GC: Phenotypic plasticity in an intertidal snail: the role of a common crab predator. *Evolution* 1996, **50**:448–454.
35. Hollander J, Collyer ML, Adams DC, Johannesson K: Phenotypic plasticity in two marine snails: constraints superseding life history. *J Evol Biol* 2006, **19**:1861–1872.
36. Hollander J, Butlin RK: The adaptive value of phenotypic plasticity in two ecotypes of a marine gastropod. *BMC Evol Biol* 2010, **10**:333.
37. Wake DB, Roth G, Wake MH: On the problem of stasis in organismal evolution. *J Theor Biol* 1983, **101**:211–224.
38. West-Eberhard MJ: Phenotypic plasticity and the origins of diversity. *Annu Rev Ecol Syst* 1989, **20**:249–278.
39. Lefebvre T, Douady CJ, Gouy M, Trontelj P, Briolay J, Gibert J: Phylogeography of a subterranean amphipod reveals cryptic diversity and dynamic evolution in extreme environments. *Mol Ecol* 2006, **15**:1797–1806.
40. Culver DC, Pipan T: *Biology of Caves and Other Subterranean Habitats*. Oxford: U.K., Oxford University Press; 2009.
41. Wright S: *The roles of mutation, inbreeding, crossbreeding, and selection in evolution*. Genetics: Proceedings of the Sixth International Congress on; 1932:355–366.
42. Pilsbry HA: *Land Mollusca of North America (north of Mexico)*, 2(2). Philadelphia: Academy of Natural Sciences of Philadelphia, Monographs 3; 1948.
43. Pilsbry HA: Forms of American *Carychium*. *Nautilus* 1891, **4**:109–110.
44. Pilsbry HA: The American species of *Carychium*. *Nautilus* 1894, **8**:61–63.
45. Clapp GH: Notes on *Carychium* and description of a New variety. *Nautilus* 1906, **14**:138–140.
46. Burch JB, Van Devender AS: Identification of eastern North American land snails. The Prosobranchia, Opisthobranchia and Pulmonata (Actophila). *Walkerana* 1980, **1**:60–80.
47. Ozimec R, Bedek J, Gottstein S, Jalžić B, Slapnik R, Štamol V, Bilandžija H, Dražina T, Kletečki E, Komerički A, Lukić M, Pavlek M: *Red book of Croatian cave dwelling fauna*. Zagreb: Ministarstvo kulture, Državni Zavod za zaštitu prirode; 2009.
48. Phillimore AB, Price TD: Density-dependent cladogenesis in birds. *PLoS Biol* 2008, **6**:483–489.
49. Nee S, May RM, Harvey PH: The reconstructed evolutionary process. *Philos Trans R Soc Lond B* 1994, **344**:305–311.
50. Etienne RS, Rosindell J: Prolonging the past counteracts the pull of the present: protracted speciation can explain observed slowdowns in diversification. *Syst Biol* 2012, **61**:204–213.
51. Slatkin M: Isolation by distance in equilibrium and nonequilibrium populations. *Evolution* 1993, **47**:264–279.
52. Pipan T, López H, Oromí P, Polak S, Culver DC: Temperature variation and the presence of trogloliths in terrestrial shallow subterranean habitats. *J Nat Hist* 2011, **45**:253–273.
53. McGinnies WG, Goldman BJ, Paylore P: *Deserts of the World*. Tucson: Univ. Arizona Press; 1968.
54. Edney EB, Franco P, Wool R: The responses of *Arenivaga* investigate (Dictyoptera) to gradients of temperature and humidity in sand studied by tagging with Technetium 99m. *Physiol Zool* 1978, **51**:241–255.
55. Prozorova LA, Lee JS, Zasyapkina MO, Korean Hypogean Malacofauna: First record in Asia of Troglobitic *Zospeum*-like snails (Pulmonata, Ellobioidea, Carychiidae). *Korean J Soil Zool* 2011, **15**:1–4.
56. Barr TC Jr, Holsinger JR: Speciation in Cave Faunas. *Annu Rev Ecol Syst* 1985, **16**:313–337.
57. Allegrucci G, Todisco V, Sbordoni V: Molecular phylogeography of *Dolichopoda* cave crickets (Orthoptera, Rhaphidophoridae): a scenario suggested by mitochondrial DNA. *Mol Phylogenet Evol* 2005, **37**:153–164.
58. Davis GM: Evolution of prosobranch snails transmitting Asian *Schistosoma*; coevolution with *Schistosoma*: a review. *Prog Clin Parasitol* 1993, **3**:145–204.
59. Gittenberger E: What about non-adaptive radiation? *Biol J Linn Soc* 1991, **43**:263–272.
60. Aljančić G, Năpăruș M: Stygobionts washed out to surface: A case of *Proteus anguinus*. In *Proceedings of 21st International Conference on Subterranean Biology: 2–7 September 2012*. Edited by Pavol J. Slovakia: Kosice: Safarik University; 2012.
61. Juan C, Guzik MT, Jaume D, Cooper SJB: Evolution in caves: Darwin's ‘wrecks of ancient life’ in the molecular era. *Mol Ecol* 2010, **19**:3865–3880.
62. Rona PA, Richardson ES: Early Cenozoic global plate reorganization. *Earth Planet Sci Lett* 1978, **40**:1–11.

63. Hsü KJ: Time and place in Alpine orogenesis - the Fermor lecture. *Geological Society, London, Special Publications* 1989, **45**:421–443.
64. Sandberger F: *Die Land- und Süßwasser-Conchylien der Vorwelt*. Wiesbaden: C.W. Kreidel; 1870:1870–1875.
65. Villatte J: Nouvelles données sur les mollusques continentaux du Thanétien inférieur sous-pyrénéen. Interprétation stratigraphique et paléocéologique. *Geobios-Lyon* 1979, **12**:513–533.
66. Stworzewicz E: Miocene land snails from Belchatów (Central Poland), III: Carychiinae (Gastropoda; Pulmonata: Ellobiidae). *Palaeont Z* 1999, **73**:261–276.
67. Bole J: Rod *Zospeum* Bourguignat 1856 (Gastropoda, Ellobiidae) Jugoslaviji. Die Gattung *Zospeum* Bourguignat 1856 (Gastropoda, Ellobiidae) in Jugoslawien. *Razprave Slov Akad Znan Umetn* 1974, **17**:249–282.
68. Azuma M: *Colored Illustrations of the Land Snails of Japan*. Hoikusha, Osaka: Japan; 1982.
69. Bank RA, Gittenberger E: Notes on Azorean and European *Carychium* species (Gastropoda Basommatophora: Ellobiidae). *Basteria* 1985, **49**:85–100.
70. Strauch E: Die Entwicklung der europäischen Vertreter der Gattung *Carychium* O.F. MÜLLER seit dem Miozän (Mollusca: Basommatophora). *Arch Moll* 1977, **107**:149–193.
71. Folmer O, Black M, Heah W, Lutz R, Vrijenhoek R: DNA primers for amplification of mitochondrial cytochrome C oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* 1994, **3**:294–299.
72. Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P: Evolution, weighting, and phylogenetic utility of mitochondrial gene-sequences and a compilation of conserved polymerase chain-reaction primers. *Ann Entomol Soc Am* 1994, **87**:651–701.
73. Colgan DJ, McLauchlan A, Wilson GDF, Livingston SP, Edgecombe GD, Macaranas J, Cassis G, Gray MR: Histone H3 and U2 snRNA sequences and arthropod molecular evolution. *Aust J Zool* 1998, **46**:419–437.
74. Katoh K, Misasa K, Kuma K, Miyata T: MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 2002, **30**:3059–3066.
75. Castresana J: Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 2000, **17**:540–552.
76. Hebert PDN, Ratnasingham S, DeWaard JR: Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proc R Soc Lond B Biol Sci* 2003, **270**:S596–S599.
77. Puillandre N, Lambert A, Brouillet S, Achaz G: ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol Ecol* 2011, **21**:1864–1877.
78. Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP: Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst Biol* 2006, **55**:595–609.
79. Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, Inward DJG, Lees DC, Ranaivosolo R, Eggleton P, Barraclough TG, Vogler AP: Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Syst Biol* 2009, **58**:298–311.
80. Hart MW, Sunday J: Things fall apart: biological species form unconnected parsimony networks. *Biol Lett* 2007, **3**:509–512.
81. Drummond AJ, Rambaut A: BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 2007, **7**:214.
82. Templeton AR, Crandall KA, Sing CF: A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 1992, **132**:619–633.
83. Clement M, Posada D, Crandall KA: TCS: a computer program to estimate gene genealogies. *Mol Ecol* 2000, **9**:1657–1659.
84. Stamatakis A: RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 2006, **22**:2688–2690.
85. Nylander JAA: *MrModeltest 2.3*. Program distributed by the author. Uppsala: Evolutionary Biology Centre, Uppsala University; 2004.
86. Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP: Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 2001, **294**:2310–2314.
87. Ronquist F, Huelsenbeck JP: MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 2003, **19**:1572–1574.
88. Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV: Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol Biol Evol* 2012, **29**:2157–2167.
89. Raftery AE, Newton MA, Satagopan JM, Krivitsky PN: Estimating the integrated likelihood via posterior simulation using the harmonic mean identity. *Memorial Sloan-Kettering Cancer Center, Dept. of Epidemiology & Biostatistics Working Paper Series* 2007, **6**:1–45.
90. Turkheimer FE, Hinz R, Cunningham VJ: On the undecidability among kinetic models: from model selection to model averaging. *J Cereb Blood Flow Metab* 2003, **23**:490–498.
91. Lartillot N, Philippe H: Computing bayes factors using thermodynamic integration. *Syst Biol* 2006, **55**:195–207.
92. Xie W, Lewis PO, Fan Y, Kuo L, Chen M-H: Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Syst Biol* 2011, **60**:150–160.
93. Burnham KP, Anderson DR: *Model Selection and Multi-Model Inference*. New York: Springer-Verlag; 2002.
94. Kass RE, Raftery AE: Bayes factors. *J Am Stat Assoc* 1995, **90**:773–795.
95. De Loriol P, Jaccard A: Etude géologique et paléontologique de la formation d'eau douce infracrétacée du Jura et en particulier de Villers-le-Lac. *Mém Soc Phys et Hist Nat Genève* 1865, **18**:63–128.
96. Briart A, Cornet FL: Description des fossiles du calcaire grossier de Mons. Quatrième partie. *Mem Acad r Sci Lett Belg* 1889, **47**:1–128.
97. Rambaut A, Drummond AJ: *Tracer v1.4*. ; <http://beast.bio.ed.ac.uk/Tracer>.
98. Rabosky DL: Likelihood methods for detecting temporal shifts in diversification rates. *Evolution* 2006, **60**:1152–1164.
99. Rabosky DL: LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. *Evol Bioinform Online* 2006, **2**:257–260.
100. Ree RH, Smith SA: Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst Biol* 2008, **57**:4–14.
101. Phillips SJ, Anderson RP, Schapire RE: Maximum entropy modeling of species geographic distributions. *Ecol Modell* 2006, **190**:231–259.
102. Hadly EA, Spaeth PA, Li C: Niche conservatism above the species level. *Proc Natl Acad Sci USA* 2009, **106**:19707–19714.
103. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A: Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 2005, **25**:1965–1978.
104. Waltari E, Hijmans RJ, Peterson AT, Nyári AS, Perkins SL, Guralnick RP: Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS One* 2007, **2**:e563.
105. Pearson RG, Raxworthy CJ, Nakamura M, Peterson AT: Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *J Biogeogr* 2007, **34**:102–117.
106. Fielding AH, Bell JF: A review of methods for the assessment of prediction errors in conservation presence: absence models. *Environ Conserv* 1997, **24**:38–49.
107. Warren DL, Glor RE, Turelli M: ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* 2010, **33**:607–611.
108. Schoener TW: *Anolis* lizards of Bimini: resource partitioning in a complex fauna. *Ecology* 1968, **49**:704–726.
109. Warren DL, Glor RE, Turelli M: Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. *Evolution* 2008, **62**:2868–2883.
110. Warren DL, Seifert SN: Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecol Appl* 2011, **21**:335–342.
111. Holm G: Exotic species of land snails from Mount Vernon, WA and Queen Elizabeth Park, Vancouver, B.C. *Dredgings* 2010, **50**:3–4.
112. Roth B: European land mollusks in the San Francisco Bay area, California: *Carychium minimum* Müller and the *Arion hortensis* complex. *Veliger* 1982, **24**:342–344.
113. Forsyth RG: *Land Snails of British Columbia*. Royal BC Museum Handbook. Victoria: Royal British Columbia Museum; 2004.

114. Forsyth RG, Oldham MJ, Schueler FW: **Mollusca, Gastropoda, Ellobiidae, *Carychium minimum*, and Ferussaciidae, *Cecilioides acicula*: Distribution extension and first provincial records of two introduced land snails in Ontario, Canada.** *Check List* 2008, **4**:449–452.
115. Grimm FW, Forsyth RG, Schueler FW, Karstad A: *Identifying Land Snails and Slugs in Canada: Introduced Species and Native Genera*. Canadian Food Inspection Agency: Ottawa; 2009.
116. Clapp WF: *Carychium minimum* Mull. *Nautilus* 1912, **26**:24.

doi:10.1186/1471-2148-13-18

**Cite this article as:** Weigand et al.: Evolution of microgastropods (Ellobioidea, Carychiidae): integrating taxonomic, phylogenetic and evolutionary hypotheses. *BMC Evolutionary Biology* 2013 **13**:18.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

