

Title: Global Diversification Dynamics Since the Jurassic: Low Dispersal and Habitat-Dependent Evolution Explain Hotspots of Diversity and Shell Disparity in River Snails (Viviparidae)

Running title: BIOGEOGRAPHY AND SHELL EVOLUTION IN RIVER SNAILS

Authors: Björn Stelbrink<sup>1,2\*</sup>, Romy Richter<sup>3</sup>, Frank Köhler<sup>4</sup>, Frank Riedel<sup>5,6</sup>, Ellen E. Strong<sup>7</sup>, Bert Van Bocxlaer<sup>8,9</sup>, Christian Albrecht<sup>1</sup>, Torsten Hauffe<sup>1</sup>, Timothy J. Page<sup>10</sup>, David C. Aldridge<sup>11</sup>, Arthur E. Bogan<sup>12</sup>, Li-Na Du<sup>13</sup>, Marivene R. Manuel-Santos<sup>14</sup>, Ristiyanti M. Marwoto<sup>15</sup>, Alena A. Shirokaya<sup>16</sup> & Thomas von Rintelen<sup>3</sup>

<sup>1</sup>Systematics & Biodiversity Lab, Department of Animal Ecology and Systematics, Justus Liebig University Giessen, Giessen, Germany

<sup>2</sup>Zoological Institute, Department of Environmental Sciences, University of Basel, Basel, Switzerland

<sup>3</sup>Museum für Naturkunde, Leibniz Institute for Evolution and Biodiversity Science, Berlin, Germany

<sup>4</sup>Australian Museum, Sydney, New South Wales, Australia

<sup>5</sup>Palaeontology, Institute of Geological Sciences, Freie Universität Berlin, Berlin, Germany

<sup>6</sup>Institute for Ecological Research and Pollution Control of Plateau Lakes, School of Ecology and Environmental Science, Yunnan University, Kunming, China

<sup>7</sup>Department of Invertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA

<sup>8</sup>CNRS, Univ. Lille, UMR 8198 Evo-Eco-Paleo, Lille France

<sup>9</sup>Limnology Unit, Department of Biology, Ghent University, Ghent, Belgium

<sup>10</sup>Australian Rivers Institute, Griffith University, Nathan, Queensland, Australia

<sup>11</sup>Aquatic Ecology Group, Department of Zoology, Cambridge University, Cambridge, UK

<sup>12</sup>Research Laboratory, North Carolina State Museum of Natural Sciences, Raleigh, USA

<sup>13</sup>State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China

<sup>14</sup>Zoology Division, National Museum of the Philippines, Manila, Philippines

<sup>15</sup>Zoology Division (Museum Zoologicum Bogoriense), Research Center for Biology, LIPI, Cibinong, Bogor, Indonesia

<sup>16</sup>Limnological Institute, Siberian Branch of Russian Academy of Sciences, Irkutsk, Russia

\*To whom correspondence should be addressed:

Present address: Björn Stelbrink, Zoological Institute, Department of Environmental Sciences, University of Basel, Basel, Switzerland. Tel: +41 61 2075723, Fax: +41 61 2070301. Email: [bjoern.stelbrink@allzool.bio.uni-giessen.de](mailto:bjoern.stelbrink@allzool.bio.uni-giessen.de), [bjoern.stelbrink@unibas.ch](mailto:bjoern.stelbrink@unibas.ch)

## ABSTRACT

The Viviparidae, commonly known as River Snails, is a dominant group of freshwater snails with a nearly worldwide distribution that reaches its highest taxonomic and morphological diversity in Southeast Asia. The rich fossil record is indicative of a probable Middle Jurassic origin on the Laurasian supercontinent where the group started to diversify during the Cretaceous. However, it remains uncertain when and how the biodiversity hotspot in Southeast Asia was formed. Here, we used a comprehensive genetic dataset containing both mitochondrial and nuclear markers and comprising species representing 24 out of 28 genera from throughout the range of the family. To reconstruct the spatiotemporal evolution of viviparids on a global scale, we reconstructed a fossil-calibrated phylogeny. We further assessed the roles of cladogenetic and anagenetic events in range evolution. Finally, we reconstructed the evolution of shell features by estimating ancestral character states to assess whether the appearance of sculptured shell morphologies was driven by major habitat shifts. The molecular phylogeny supports the monophyly of the three subfamilies, the Bellamyinae, Lioplacinae, and Viviparinae, but challenges the currently accepted genus-level classification in several cases. The almost global distribution of River Snails has been influenced both by comparatively ancient vicariance and more recent founder events. In Southeast Asia, Miocene dispersal was a main factor in shaping the modern species distributions. A recurrent theme across different viviparid taxa is that many species living in lentic waters exhibit sculptured shells, whereas only one strongly sculptured species is known from lotic environments. We show that such shell sculpture is habitat-dependent and indeed evolved several times independently in lentic River Snails. Considerably high transition rates between shell types in lentic habitats probably caused the co-occurrence of morphologically distinct shell types in several lakes. In contrast, directional evolution towards smooth shells in lotic habitats, as identified in the present analyses, explains why sculptured shells are rarely found in these

habitats. However, the specific factors that promoted changes in shell morphology require further work.

**KEYWORDS:** Southeast Asia, fossil-calibrated phylogeny, biogeographical analyses, fossil-constrained analyses, stochastic character mapping

## INTRODUCTION

Integrating molecular phylogenetic methods and biogeographical theory promises to improve our understanding of the distribution of life on Earth by providing new insights into the timing and spatial extent of distributional shifts on a global scale (e.g., Ree and Smith 2008; Ronquist and Sanmartín 2011; Matzke 2013a, 2014; Ho et al. 2015; Yu et al. 2015). Such insights are crucial to discern the factors and processes that have shaped current patterns of biodiversity. The importance of geological events as drivers of vicariance was recognized early (e.g., Wegener 1912a, 1912b, 1912c; Arldt 1917; du Toit 1937), which led to the rise of cladistic biogeography in the 1980s (see e.g., Hallam 1981; Wiley 1988; Upchurch 2008; Briggs 2009; Crisp et al. 2011; McIntyre et al. 2017). However, this focus on vicariance has increasingly been challenged with the recognition that additional factors, both intrinsic (e.g., dispersal ability) and extrinsic (e.g., wind and vectors), also play influential roles in determining species distributions (e.g., de Queiroz 2005; Cowie and Holland 2008; Gillespie et al. 2012; Shaw and Gillespie 2016). Many freshwater taxa are particularly suitable for the study of biogeographical patterns and processes and how they affect diversification dynamics as they inhabit an environment that is delimited by well-defined boundaries that act as effective dispersal barriers (i.e., both terrestrial and marine environments). Those taxa with a long evolutionary history, a rich fossil record, and a nearly global distribution, such as various freshwater gastropods (see e.g., Strong et al. 2008), are particularly suited to examine the biogeographical context of diversification across large spatial and temporal scales.

One such gastropod family is the Viviparidae, or River Snails (and also sometimes referred to as Mystery or Pond Snails), which currently occur on almost all continents from sea level to about 2,700 m a.s.l., where they frequently dominate the mollusk fauna (e.g., Strong et al. 2008; Van Bocxlaer and Strong 2019). Viviparidae is a well-supported family within the Caenogastropoda based on morpho-anatomical and molecular evidence (see e.g.,

Colgan et al. 2007; Ponder et al. 2008; Strong et al. 2011). It has a rich fossil record, with the oldest taxa known from the Mesozoic, including several Jurassic and Cretaceous representatives (see Van Bocxlaer and Strong 2019 and references therein). River Snails are commonly subdivided into three subfamilies, the Bellamyinae, Lioplacinae, and Viviparinae (= Campelomatinae, *vide* Bouchet et al. 2017). This classification is mainly based on anatomical features (see Rohrbach 1937; Vail 1977), which suggest a close relationship between Lioplacinae and Viviparinae (see Vail 1977). Whereas Lioplacinae are indigenous to eastern North America, the Viviparinae occur in North America, the western Palearctic and eastern Asia. The most diverse and widespread subfamily is Bellamyinae comprising numerous species from Africa, Asia, and Australia (Fig. 1). River Snails mainly inhabit lentic waters, such as (ancient) lakes, rice paddies, wetlands, and ponds, but are also common in large rivers and streams (i.e., lotic habitats), such as the Nile and Zambezi in Africa (genus *Bellamyia*), the Mekong in Asia (genus *Mekongia*), or the North American Alabama-Coosa river system (genus *Tulotoma*).

Viviparidae are represented by roughly 125-150 extant species worldwide and reach their highest taxonomic diversity (c. 40-60 species) in Southeast Asia (Strong et al. 2008). This region also harbours most currently accepted genera (Fig. 1). Encompassing several global biodiversity hotspots, Southeast Asia is of great interest to biogeographers and evolutionary biologists because it has experienced a dynamic and complex tectonic and climatic history, particularly since the Early Miocene (Woodruff 2010; Lohman et al. 2011; Metcalfe 2011; de Bruyn et al. 2014; Klaus et al. 2016). The interplay of tectonic and climatic processes has triggered extensive *in situ* diversification in several taxonomic groups, but it has also facilitated widespread biotic exchange via dispersal and vicariance across prominent biogeographical barriers in numerous taxonomic groups (see e.g., Stelbrink et al. 2012; de

Bruyn et al. 2014; Crayn et al. 2015 for meta-analyses), including freshwater mollusks (e.g., Köhler and Glaubrecht 2010; von Rintelen et al. 2014; Zielske et al. 2017).

Owing to their evolutionary age and widespread distribution, biogeographical patterns in River Snails across Southeast Asia should reflect the past environmental history well. Consequently, the study of diversification dynamics in this group would allow a better understanding of the mechanisms that have structured freshwater biodiversity throughout the region. Despite the ubiquity of viviparids, and their potential for biogeographical studies, surprisingly few modern studies have sought to unravel global diversification dynamics. Several recent studies have investigated the phylogeography and population genetic structure (e.g., Johnson and Bragg 1999; Johnson and Leefe 1999; Chiu et al. 2002; Carini and Hughes 2006; Carini et al. 2006; Schultheiß et al. 2014; Gu et al. 2015b, 2015a) or addressed the systematic relationships among species in a regional context (Sengupta et al. 2009; Schultheiß et al. 2011; Du et al. 2013; Hirano et al. 2015, 2019a, 2019b; Zhang et al. 2015; Gu et al. 2019; Rysiewska et al. 2019; Sil et al. 2019). These studies identified high levels of local and regional endemism and the importance of intralacustrine radiation in generating this diversity. They further indicated that past geological and climatic events and palaeohydrological connections have shaped the present-day distribution of viviparids. However, because of the lack of comprehensive taxonomic and genetic datasets, phylogenetic relationships, and global diversification dynamics of River Snails have remained poorly understood.

Irrespective of their considerable taxonomic diversity, Viviparidae are also renowned for their remarkable conservatism in shell morphology. Most taxa exhibit conical to subglobose shells that lack sculpture, and the observed variation in shell characters is comparatively small in comparison to the taxonomic diversity and evolutionary age of the family. However, there are notable exceptions of taxa, among both extant and fossil lineages, of taxa primarily from lacustrine (lentic) environments exhibiting highly divergent shell

morphotypes that include whorls covered in spiral lines, keels, or even noded spirals and spines. Such examples are found in the extant viviparid faunas of Lake Lanao, Philippines (*Torotaia* spp.; Bartsch 1907; Stelbrink et al. 2019) and Lake Inlé, Myanmar (*Taia* spp.; e.g., Annandale 1924), but also the fossil assemblages of *Neothauma* from the palaeo-Lake Obweruka in the East African Rift (see e.g., Van Damme and Pickford 1999; Salzburger et al. 2014) and *Margarya/Macromargarya* from the Chinese palaeo-Lake Nanning (Tian et al. 2013, 2018). Whereas shell sculpture has been emphasized in taxonomy, little is known about how often and when such sculptured shells have evolved, for how long they persisted, and whether their occurrence is dependent on the habitat type in which these forms are found.

Here, we present the first molecular phylogeny of viviparid snails based on a comprehensive genetic dataset with a global coverage that comprises most of the extant genera (24 out of 28). We calibrated the phylogeny using five fossil calibration points and considered distributional and shell shape information of both extant and extinct species to inform various analyses. The aims of this study are: (1) to identify the evolutionary centre of origin of River Snails and to shed light on the relative importance of cladogenetic range evolution (e.g., vicariance or founder events) and anagenetic dispersal in shaping modern distribution patterns, and (2) to assess whether shell sculpture evolved several times independently in River Snails, and if so, whether it was linked to major habitat shifts. On the basis of modern species distributions and life-history traits, we expect a considerable role of palaeogeographical processes and vicariance, whereas dispersal in River Snails may have been of less importance. Moreover, morphological disparity across the family suggests that shell evolution was closely linked to the differential ecology of major viviparid habitats.

## MATERIALS AND METHODS

### *Materials*



The bulk of material used in this study was collected throughout Southeast Asia and Australia between 2001-2010 and is deposited in the Malacological Collection of the Museum für Naturkunde Berlin (ZMB), the Bogor Zoological Museum (Cibinong, Indonesia), and the National Museum of the Philippines (Manila). Additional ethanol-fixed samples have been obtained from the Florida Museum of Natural History, Gainesville (FLMNH), the North Carolina State Museum of Natural Sciences, Raleigh (NCSM), the National Museum of Natural History, Smithsonian Institution, Washington (USNM), and the Systematics and Biodiversity Collection, University of Giessen (UGSB).

While most of the genetic data analysed herein was newly generated, complementary sequences have been obtained from GenBank (Supplementary Table S1 available on Dryad). For the African species (*Bellamya* spp. and *Neothauma tanganyicense*), we subsampled all major taxonomic groups and biogeographical clades that were published by Schultheiß et al. (2014). Our final dataset contained 61 nominal species from 24 out of 28 viviparid genera (sampling sites shown in Fig. 1 and Supplementary Fig. S1 available on Dryad). Two caenogastropod representatives belonging to the families Ampullariidae (*Pomacea canaliculata*) and Bithyniidae (*Bithynia* sp.) were used as outgroups to root the trees (see e.g., Ponder et al. 2008).

#### *DNA Isolation, PCR Amplification and DNA Sequencing*

Genomic DNA was isolated from c. 1 mm<sup>3</sup> of muscle tissue from the foot using a mollusk-specific protocol by Winnepenninckx et al. (1993). A single mitochondrial (COI, 660 base pairs, bp) and two nuclear markers (28S rRNA, c. 1,078 bp; H3, 328 bp) were amplified by using primers and PCR cycling conditions described by Van Bocxlaer et al. (2018) to resolve relationships from within species to between subfamilies and genera.

In some individuals (N = 24; Supplementary Table S1-S2 available on Dryad), the COI sequences contained apparently heterozygous sites, possibly indicating the presence of nuclear copies of mitochondrial DNA (NUMTs; e.g., Bensasson et al. 2001). In these cases, the PCR amplifications were repeated and the PCR products were subsequently purified using magnetic beads (SPRI beads, Agencourt AMPure XP). Amplicon sequencing was then performed on a Roche 454 GS FLX Titanium platform at the Berlin Center for Genomics in Biodiversity Research (BeGenDiv) using the direct multiplex sequencing (DMPS) protocol of Stiller et al. (2009). Multiple fragments were aligned against a reference sequence obtained by Sanger sequencing, translated into amino acids to check for stop codons or non-coding positions and were subjected to a phylogenetic analysis by maximum likelihood (ML) to allow unambiguous identification of the mitochondrial fragment.

#### *Alignment and Identification of Best-Fit Partition Models*

The ribosomal sequences were aligned using the MAFFT WEB SERVICE (Kato and Toh 2008; Kato and Standley 2013) with default settings. Ambiguous alignment sites were removed using the GBLOCKS SERVER 0.91b (Castresana 2000) with all three options enabled for a less-stringent selection (i.e., smaller final blocks, gap positions within the final block and less strict flanking positions). The resulting 28S rRNA alignment included 1,078 bp, or 93% of the original 1,150 positions.

The final three-gene dataset was investigated for the best-fit partition models in PARTITIONFINDER 2 (Guindon et al. 2010; Lanfear et al. 2012, 2016) by using a greedy algorithm with codon-position data blocks (for COI and H3) and linked branch lengths. The best-fit scheme revealed by PARTITIONFINDER 2 using the corrected Akaike information criterion (AICc) suggested seven partitions for both criteria: 28S rRNA = GTR+I+ $\Gamma$ , COI<sup>1st</sup> = GTR+I+ $\Gamma$ , COI<sup>2nd</sup> = GTR+I+ $\Gamma$ , COI<sup>3rd</sup> = HKY+I+ $\Gamma$ , H3<sup>1st</sup> = GTR+ $\Gamma$ , H3<sup>2nd</sup> = GTR+I, and

H3<sup>3rd</sup> JC+I, whereas the Bayesian information criterion (BIC) favoured a similar but slightly less complex partitioning scheme: 28S rRNA+ H3<sup>2nd</sup> = GTR+I+ $\Gamma$ , COI<sup>1st</sup> = GTR+ $\Gamma$ , COI<sup>2nd</sup> = GTR+I+ $\Gamma$ , COI<sup>3rd</sup> = HKY+I+ $\Gamma$ , H3<sup>1st</sup> = K80+ $\Gamma$ , and H3<sup>3rd</sup> JC+I. Phylogenetic reconstructions were highly similar under both partitioning schemes and we applied the AICc partitioning scheme in the following analyses.

### *Phylogenetic Analyses*

Phylogenetic analyses were conducted on the CIPRES Science Gateway (Miller et al. 2010) using ML as implemented in RAXML-HPC BLACKBOX (8.2.10; Stamatakis 2014), and Bayesian inference (BI) using MRBAYES 3.2.6 (Ronquist et al. 2012) with the above-mentioned partitioning scheme. However, for ML we used the only available substitution model GTR+ $\Gamma$  for all partitions (Supplementary Fig. S2 available on Dryad). ML bootstrapping was performed by employing a stop rule (the analysis was automatically stopped after 456 bootstrap replicates). The settings for BI were: ngen = 10,000,000, samplefreq = 500, nchains = 4, burn-in = 10,001; Supplementary Fig. S3 available on Dryad). Convergence of the BI analysis was assessed using TRACER 1.5 (Rambaut and Drummond 2007).

### *Fossil-Calibrated Molecular-Clock Analyses*

For the molecular-clock analyses, we reduced the dataset and retained the most basal lineage of each species according to the ML and BI analyses, resulting in a total of 74 operational taxonomic units (OTUs; see Fig. 2). The discrepancy between the number of nominal species vs. OTUs (i.e., 61 vs. 74) is related to the non-monophyly of various *Bellamyia* species (see also Schultheiß et al. 2014) and the presence of undetermined species in the dataset. Because *Filopaludina javanica* was consistently recovered to be non-monophyletic, we predefined two

OTUs for this taxon: whereas '*F. javanica* 1' included individuals from Sumatra, Java, and Sulawesi, '*F. javanica* 2' comprises specimens from Bali, Borneo, and Sulawesi plus *F. decipiens* from New Guinea. We re-ran PARTITIONFINDER and used the best-fit substitution models and partitioning scheme as suggested by the AICc, which was very similar compared to the full dataset: 28S rRNA = GTR+I+ $\Gamma$ , COI<sup>1st</sup>+COI<sup>2nd</sup>+COI<sup>3rd</sup> = GTR+I+ $\Gamma$ , H3<sup>1st</sup> = K80+I+ $\Gamma$ , H3<sup>2nd</sup>+H3<sup>3rd</sup> = GTR+I.

The rich fossil record of Viviparidae provides excellent opportunities for time calibration, but also a risk, as several fossils have a questionable taxonomic status, are difficult to place within the phylogeny, or are derived from strata with poor age constraints. Here, we only used well-preserved taxa from non-marine facies, which could be confidently allocated based on synapomorphic shell features and are stratigraphically well constrained (following Parham et al. 2012; Supplementary Tables S3-S4 available on Dryad).

Divergence times were estimated using BEAST 1.8.4 (Drummond et al. 2012) by employing the ages of five fossils from the Jurassic, Cretaceous, Oligocene, and Miocene to calibrate the phylogeny (Fig. 2, Supplementary Fig. S4 and Supplementary Table S3 available on Dryad). The oldest known and reliably classified fossil viviparid is †*Viviparus langtonensis* from England, UK (Middle Jurassic; Hudleston 1896; Tracey et al. 1993). It was used to constrain the age of the most recent common ancestor (MRCA) of all extant viviparids (crown node; calibration point, CP1). The remaining fossils were used to calibrate internal nodes: †*Campeloma harlowtonense* from Montana, USA (Early Cretaceous; Stanton 1903; Yen 1950) was used to calibrate the split between *Campeloma* and *Lioplax* (CP2); †*Margarya nanningensis* from Guangxi, China (Early Oligocene; Tian et al. 2013, 2018; Quan et al. 2016) for the MRCA of *Margarya melanioides*, *M. oxytropoides*, and *Cipangopaludina* spp. from East Asia and New Guinea based on similar embryonic shell features (CP3); *Bellamyia* cf. *unicolor* from Napak, Uganda (Early Miocene; Pickford 2004) for the MRCA of *Bellamyia*

spp., *Neothauma tanganyicense*, and *Idiopoma* sp. (CP4; see also Schultheiß et al. 2014); and †*Neothauma hattinghi* from the Albertine Rift Valley in Uganda (Middle/Late Miocene; Van Damme and Pickford 1999) for the stem age of the two *Neothauma tanganyicense* lineages, which represents the MRCA of *Neothauma tanganyicense*, *Bellamyia* cf. *capillata* ('Northern', 'Southern', and 'Zambezi'), *B.* cf. *monardi* 'Northern', *B. crawshayi*, and *B. pagodiformis* based on the topology inferred from the ML and BI analyses (CP5; see also Schultheiß et al. 2014).

A gamma prior was chosen for all fossil calibrations with a hard minimum age and a soft maximum age corresponding to the fossil's stratigraphic age (Supplementary Table S3 available on Dryad). Accordingly, scale values were selected that ensure that the median of the prior represented the maximum age of the respective fossil, whereas shape values were always set to 1.0. By doing so, a broad confidence interval was selected that accounts for the incompleteness of the fossil record (see e.g., Marshall 1990): CP 1: offset = 166.0, scale = 11.5 → 95% quantile of the prior = 200.5; CP 2: offset = 112.0, scale = 13.0 → 95% = 150.9; CP 3: offset = 25.0, scale = 7.3 → 95% = 46.9; CP 4: offset = 18.5, scale = 2.2 → 95% = 25.1; and CP 5: offset = 10.0, scale = 1.5 → 95% = 14.5. Analyses were run on the CIPRES server in two replicates with an uncorrelated lognormal relaxed-clock (UCLN) model, with codon partitions, unlinked clock models (but linked among codon positions), and a birth–death tree prior (settings: ngen = 200,000,000, samplefreq = 10,000). Separate log and tree files were combined using LOGCOMBINER 1.8.4 and a 50% burn-in (resulting in a total of 20,000 trees). The maximum clade credibility (MCC) tree was identified using TREEANNOTATOR 1.8.4. Because convergence was not reached using the best-fit substitution models from PARTITIONFINDER for each of the seven data partitions, less complex substitution models (i.e., HKY instead of GTR) were used for the relevant partitions (see above), which resulted in considerably higher effective sample size (ESS) values (>200) for all parameters.

Branch-specific rates for 28S rRNA, COI, and H3 derived from the BEAST MCC tree are shown in Supplementary Fig. S5 available on Dryad.

To assess how the different fossil calibration priors interact with each other and how they influence divergence times, we performed additional BEAST analyses, in which: (1) we sampled from the prior only, (2) only CP1 was enforced, or (3) only CP3-CP5 were enforced (Supplementary Fig. S6 and Supplementary Table S5 available on Dryad).

### *Biogeographical Analyses*

Ancestral areas were estimated using BIOGEOBEARS 1.1 (Matzke 2013b, 2013a) for the R statistical environment 3.5.1 (R Core Team, 2018). We delimited thirteen geographical areas; Southeast Asia was subdivided into eight geographical areas following the study by de Bruyn et al. (2014), whereas some large areas with sparse representation such as North America and Africa were not further subdivided for the purpose of this study: (A) = North America, (B) = Africa, (C) = Europe, (D) = East Asia (China, Far East Russia, Japan, and South Korea), (E) = India, (F) = Indochina (incl. the Thai-Malay Peninsula and Singapore), (G) = Sumatra, (H) = Philippines, (J) Palawan, (K) = Borneo, (L) = Java (incl. Bali), (M) = Sulawesi, and (N) = ‘East of Wallacea’ (Australia and New Guinea).

In a second, fossil-constrained biogeographical analysis, we assessed the impact of the geographical occurrence of fossils on the estimation of ancestral areas. Thus, the BIOGEOBEARS arguments ‘fixnode’ and ‘fixlikes’ were used to apply hard geographical constraints to the five nodes used to time-calibrate the molecular phylogeny: CP1 = Europe (C), CP2 = North America (B), CP3 = East Asia (D), and CP4 and CP5 = Africa (B).

Three different biogeographical models were tested (i.e., BayAreaLIKE, DEC, and DIVALIKE) plus their +J version, which implements jump dispersal at cladogenetic events and represents founder-event speciation (see Matzke 2013a, 2014). Lineages were allowed to

occur in a maximum of four combined areas. The best-fit biogeographical model for each approach was identified using the AIC (Supplementary Table S6 available on Dryad).

Because estimated ancestral areas alone do not necessarily reflect the history of biogeographical events, an additional analysis using biogeographical stochastic mapping (BSM; see Dupin et al. 2017) was performed in BIOGEOBEARS (settings: nsim = 1,000). Parameters estimated for the best-fit model for both the unconstrained and fossil-constrained analyses were used as priors, and the frequency of cladogenetic and anagenetic events was counted.

#### *Analysis of Shell Characters and Habitat Types*

Representative specimens of all 74 OTUs included in the biogeographical analyses were photographed in a standardized view (Supplementary Fig. S7 available on Dryad). Standard shell parameters (height, width, width/height ratio; Supplementary Table S7 available on Dryad) were plotted against each other per genus and clade (Supplementary Fig. S8 available on Dryad). Phylogenetic ANOVAs (Garland et al. 1993) were performed using the R package PHYTOOLS 0.6-44 (Revell 2012) to test for clade-specific shell differences.

We also constructed a matrix, in which three discrete states representing gradations of spiral sculpture (i.e., ‘absent’, ‘fine’, ‘coarse’ incl. noded spirals) and two discrete habitat categories (i.e., lentic = rice paddies, wetlands, ponds, lakes; lotic = creeks, rivers) were assigned to the 74 OTUs (Supplementary Table S7 available on Dryad). Shell and habitat information were coded based on sequenced specimens and general information from the literature. Accordingly, a total of six possible categories were obtained: ‘LeA’ = lentic with shell type A (‘absent’), ‘LeB’ = lentic with shell type B (‘fine’), ‘LeC’ = lentic with shell type C (‘coarse’), ‘LoA’ = lotic with shell type A, ‘LoB’ = lotic with shell type B, and ‘LoC’ = lotic with shell type C.

To estimate transition rates among these categories, three different models for discrete character evolution (i.e., ‘ER’ = equal rates, ‘SYM’ = symmetric rates, and ‘ARD’ = all rates different) were fitted in PHYTOOLS and compared using the AIC; Supplementary Table S8 available on Dryad). To account for these different models and to differentiate between habitat-independent and habitat-dependent transition rates, a total of twelve Q (transition) matrices were created that constrain particular transition rates (Supplementary Fig. S9 available on Dryad). Whereas models with habitat-independent transitions assume equal rates for shell type transitions in both lentic and lotic habitats (e.g., ‘LeA’→‘LeB’ = ‘LoA’→‘LoB’), habitat-dependent models suppose different shell transition rates (‘LeA’→‘LeB’ ≠ ‘LoA’→‘LoB’). Simultaneous transitions in habitat and shell sculpture (e.g., ‘LeA’→‘LoB’) were not allowed *a priori*. This is a multistate extension of the model of correlated evolution of binary traits (Pagel 1994). Ancestral state estimation was performed using stochastic character mapping (Huelsenbeck et al. 2003) as implemented in PHYTOOLS, which samples character histories from the posterior probability distribution (settings: nsim = 1,000).

As for the biogeographical analysis, a fossil-constrained analysis of habitat-shell evolution was performed by constraining the ancestral states for the five nodes used as fossil calibration points: CP1 = ‘LeA’/‘LoA’, CP2 and CP4 = ‘LoA’, CP3 = ‘LeC’, and CP5 = ‘LeB’/‘LoB’ (see Supplementary Material available on Dryad). Accordingly, five additional tips with the mentioned character states were incorporated into the phylogeny at the nodes of interest using the PHYTOOLS function ‘bind.tip’. After re-running the above-mentioned analyses (settings: nsim = 1,000), the five additional tips were removed and the fossil-constrained ancestral states were mapped onto the original phylogeny. Because of the general difficulty to infer habitat types of fossil species, we evaluated the impact of our fossil states on the analysis by an additional analysis, in which the habitat type remained uncertain for all



five fossil species (i.e., ‘LeA’/LoA’, ‘LeB’/LoB’, and ‘LeC’/LoC’). In the fossil-constrained analyses, we estimated considerably high transition rates (e.g., ‘LeA’→‘LeB’) for few of the models of character evolution. These corresponded to an instantaneous state change and prevented the subsequent stochastic mapping of ancestral states. Because transition rates higher than 10 did not substantially improve model likelihoods, we fixed their upper bound during the maximum likelihood search to 10.

In contrast to the biogeographical models that only contained 3 parameters, up to 14 free parameters needed to be estimated by the models of habitat-shell evolution. Together with our incomplete taxon sampling, this increases the chance of selecting the wrong model. We therefore used a bootstrapping approach in which we performed the following steps: 1) we estimated speciation and extinction rates under consideration of the incomplete taxon sampling with the R package DIVERSITREE 0.9-13 (FitzJohn 2012); although c. 125-150 extant species are currently recognized (Strong et al. 2008), we applied a very conservative sampling fraction of c. 0.40 based on the number of extant taxa found at MolluscaBase (2019; N = 186); 2) speciation and extinction rates were used to simulate 250 taxonomically complete phylogenies (i.e., number of taxa = 186) with PHYTOOLS; 3) habitat and shell states were simulated for all taxa using the parameters from the best-fit habitat-shell evolution model; 4) phylogenies and states were pruned to the number of observed taxa (N = 74); and 5) all abovementioned models of habitat-shell evolution were fitted and ranked.

## RESULTS

### *Phylogenetic Relationships and Divergence Time Estimates*

Our dataset included 193 sequences from 74 OTUs representing 61 nominal species (see explanation above) in 24 genera and is the most comprehensive dataset for River Snails analysed thus far. The topologies of the unconstrained (ML and BI; Supplementary Figs. S2-

S3 available on Dryad) and fossil-calibrated trees (BEAST) were almost identical and were consistent with the anatomy-based taxonomic subdivision of the family into three subfamilies. Whereas prior and posterior distributions were very similar for CP1 and CP3, the comparison for the remaining calibration points indicated that divergence times were mainly informed by the sequence information and not only by the priors (Supplementary Fig. S6 available on Dryad). Mean ages for CP2-CP5 only differed by c. 2-17% (except for CP3, which was considerably younger) when only the oldest fossil calibration point (CP1) was used. Similarly, a reduced set of fossil calibration points (CP3-CP5) resulted in only c. 11-18% younger mean ages for the nodes corresponding to CP1 and CP2, respectively (Supplementary Table S5 available on Dryad). Both suggesting that the placement of the fossils in the tree was generally reasonable and that age estimates are robust.

The monophyly of Lioplacinae and Viviparinae, respectively, was highly supported (Fig. 2 and Supplementary Figs. S2-S3 available on Dryad). However, branch support for their sister group relationship was low (ML bootstrap value: <0.50; BI posterior probability: 0.76; BEAST posterior probability: 0.68; see also Fig. 2 and Supplementary Figs. S2-S3 available on Dryad). The estimated age of the Lioplacinae–Viviparinae split (mean = 155.2 million years ago, Ma; 95% highest posterior density, 95% HPD = 129.4-178.7 Ma) and the initial diversification within each of these clades dated back to the Mesozoic (Fig. 2). Within the Lioplacinae, *Campeloma* and *Lioplax* were recovered reciprocally monophyletic, whereas the three genera within the Viviparinae fall into two distinctive and highly supported sister groups: the European *Viviparus* species are sister to the Chinese *Rivularia*, whereas the North American *Viviparus* and *Tulotoma* are sister within a separate clade. The split between the European/East Asian and North American Viviparinae was estimated at c. 106.1 Ma (95% HPD = 67.8-148.3).

The sister clade of Lioplacinae + Viviparinae is the highly supported Bellamyinae (Fig. 2 and Supplementary Figs. S2-S3 available on Dryad). The initial diversification of Bellamyinae probably began during the Late Cretaceous, c. 87.2 Ma (95% HPD = 63.3-113.9), i.e., considerably later than the initial diversification of the other two subfamilies. Within the Bellamyinae, two distinct clades were recovered: Bellamyinae-clade A includes only Asian genera (*Angulyagra*, *Anularya*, *Celetaia*, *Cipangopaludina*, *Margarya*, *Sinotaia*, and *Torotaia*), whereas Bellamyinae-clade B is geographically more widespread comprising African (*Bellamyia*, *Neothauma*), Asian (*Angulyagra*, *Anulotaia*, *Filopaludina*, *Idiopoma*, *Mekongia*, *Taia*, *Tchangmargarya*, and *Trochotaia*), and Australian genera (*Larina* and *Notopala*). However, the support for Bellamyinae-clade B is comparatively low given the uncertain phylogenetic positions of *Mekongia* and *Tchangmargarya* (Fig. 2). Moreover, the monophyly of the Australian clade remains uncertain because of low node support. The MRCA of each of Bellamyinae clades A and B is estimated to have originated in the Eocene, c. 46.9 Ma (95% HPD = 31.4-63.7), and in the Late Cretaceous c. 78.7 Ma (95% HPD = 55.9-103.7), respectively.

Most of the genera studied here are reciprocally monophyletic and highly supported (Fig. 2 and Supplementary Figs. S2-S3 available on Dryad). However, several paraphyletic and polyphyletic groups were recovered, including e.g., *Viviparus* (Viviparinae), *Cipangopaludina* (Bellamyinae-clade A), *Angulyagra* (Bellamyinae-clades A and B), and perhaps *Filopaludina* (Bellamyinae-clade B), although low support values do not allow to draw final conclusions on this latter taxon.

### *Major Biogeographical Patterns*

In both analyses (unconstrained vs. fossil-constrained), DEC+J was identified as the best-fit biogeographical model (Supplementary Table S6 available on Dryad). The biogeographical

models BayAreaLIKE, DEC, and DIVALIKE consistently revealed a lower fit to the data than their +J versions. However, the jump-dispersal weight was generally low in all analyses (Supplementary Table S6 available on Dryad). According to the BSM analyses, the largest share of cladogenetic events (unconstrained: c. 74% vs. fossil-constrained: c. 73%) was attributed to sympatry ('y', *sensu* (Matzke 2014). In contrast, range-changing jump-dispersal ('j'), subset sympatry ('s'), and vicariance ('v') were only rarely identified in the unconstrained and fossil-constrained analyses ('j': c. 10% vs. 16%, 's': c. 7% vs. 5%, and 'v': 10% vs. 3%, corresponding 19-20 out of 73 cladogenetic events; see colour-coded nodes in Fig. 3). In contrast, only few anagenetic events were recovered (unconstrained vs. fossil-constrained: 5 vs. 7 dispersal events out of a total of 146 branches; no extinction events, i.e., range contraction; see colour-coded branches in Fig. 3).

The unconstrained analysis suggested that the MRCA of all viviparids inhabited North America + East Asia + Indochina (Fig. 3; see Supplementary Material available on Dryad). The first diversification event that gave rise to the MRCA of Lioplacinae and Viviparinae (North America) and the MRCA of the Bellamyinae (East Asia + Indochina) was likely caused by a vicariant event (BSM: c. 82%). A combination of cladogenetic events (sympatry and vicariance/jump-dispersal) resulted in the colonization of Europe via East Asia by Viviparinae. The split between the Lioplacinae and Viviparinae (sympatry, BSM: c. 68%) and within the Viviparinae occurred in North America, followed by either a vicariant (BSM: c. 57%) or jump-dispersal event (BSM: c. 43%) that gave rise to the North American and European/East Asian clades in the Viviparinae (Fig. 3). Similarly, the split between the European *Viviparus* and the Chinese *Rivularia* was either caused by vicariance or jump dispersal (BSM: c. 53 and c. 47%, respectively). Within the Bellamyinae, clade A probably originated in East Asia, whereas clade B probably emerged in East Asia + Indochina. In clade A, different SE Asian islands or archipelagos, such as the Philippines, Sulawesi, and New

Guinea, were colonized mainly by jump dispersal events during the Middle and Late Miocene. In clade B, ancestral lineages from Indochina colonized Africa via vicariance (BSM: c. 58%) or jump dispersal (BSM: c. 42%) and Australia + New Guinea (sympatry and subset sympatry) during the Late Oligocene to Early Miocene, the latter of which gave rise to the geographically widespread genus *Filopaludina*. This genus may have originated in Indochina during the Early Miocene where it still occurs today (*F. doliaris*, *F. filosa*, *F. martensi*, *F. polygramma*, and *F. sumatrensis*). *Filopaludina* also colonized several islands within the Indo-Australian Archipelago (IAA) from Indochina during the Late Miocene (*F. javanica* 1, *F. javanica* 2, and *F. luzonica*). These colonization events were mainly associated with anagenetic dispersal (Fig. 3).

The fossil-constrained analysis revealed similar results to the unconstrained analysis described above, however, the reconstruction of the centre of origin of River Snails changed. Accordingly, the MRCA of all viviparids originated in Europe, whereas both the Lioplacinae and Viviparinae originated in North America + Europe and the split within the North American and European/East Asian Viviparinae was caused by vicariance (BSM: c. 81%). For the Bellamyinae, a jump dispersal-event from Europe into Indochina (BSM: 100%), followed by dispersal into Indochina + East Asia is suggested. Other, more recent, nodes indicated similar cladogenetic events regardless of whether or not fossil constraints were used, except for nodes related to *Mekongia* and *Tchangmargarya*. In summary, adding fossil constraints reduced the number of possible ancestral areas estimated for the MRCA of Viviparidae and several descendent nodes in comparison to the unconstrained analyses (Fig. 3).

### *Shell Character Evolution across Habitat Types*

General shell measurements (height, width, and width/height ratio) did not allow separation of viviparid species into taxonomic groups at the genus or subfamily level (Supplementary Fig. S8 available on Dryad). Similarly, despite the fact that clade A may be characterized by larger and potentially more slender species on average in comparison to clade B (Supplementary Fig. S8 available on Dryad), no significant differences were identified between the two clades by the phylogenetic ANOVAs ( $F_{\text{height}} = 79.8$ ,  $p = 0.19$ ;  $F_{\text{width}} = 72.9$ ,  $p = 0.21$ ;  $F_{\text{width/height ratio}} = 7.9$ ,  $p = 0.68$ ).

Forty-eight of the 74 OTUs (c. 65%) represented in the phylogeny possess a smooth shell, 17 OTUs (c. 23%) exhibit a fine spiral sculpture, and 9 OTUs (c. 12%) a coarse spiral sculpture that may include noded spirals (Fig. 4 and Supplementary Fig. S7 available on Dryad). With the exception of *Tulotoma magnifica*, all extant members of the Lioplacinae and Viviparinae lack significant sculpture. This is also true for many species belonging to Bellamyinae—clade B. In contrast, the extant representatives of Bellamyinae—clade A display considerably greater disparity in shell sculpture compared to the three above-mentioned clades (Fig. 4). Species with a smooth shell or a fine spiral sculpture mainly inhabit lotic but also lentic habitats. In comparison, taxa with a coarse spiral shell sculpture are almost exclusively found in lacustrine environments (except *Tulotoma magnifica*; Fig. 4). Within the Bellamyinae, most species within clade A inhabit lentic habitats. In contrast, clade B mainly consists of lotic species except for *Neothauma* and some *Bellamyia* and *Filopaludina* species.

According to the AIC comparisons, the habitat-dependent HabDep-Hab<sub>ER</sub>-Shell<sub>ARD</sub> model featuring equal rates for habitat transitions and different rates for shell transitions (Supplementary Fig. S9 available on Dryad) was identified as the best-fit model for both the unconstrained and fossil-constrained analyses (Supplementary Table S8 available on Dryad). Both the unconstrained and fossil-constrained analyses identified smooth shells and a predominantly lotic habitat for the root of the tree (Fig. 4). The shell of the MRCA of all three

subfamilies was also estimated to be smooth, but the ancestral habitat remains largely inconclusive. A smooth shell in combination with a lotic habitat was estimated for the MRCA of Bellamyinae—clade B and most of its descendent nodes. In contrast, species belonging to clade A may have been derived from an ancestor with fine or coarse spiral shell sculpture. This pattern is even more pronounced in the fossil-constrained analysis (Fig. 4). Remarkably, the occurrence of sculptured shells in clade A coincided with a major shift from lotic to lentic habitats as indicated by the colour gradient across the branches. However, such habitat shifts did not coincide with a change in shell sculpture, for example, in *Bellamya* species (clade B) inhabiting ancient lakes such as Lake Malawi (Fig. 4; see also Schultheiß et al. 2011).

Estimated transition rates were generally very similar between the unconstrained and fossil-constrained analyses, but were highly asymmetric (Supplementary Table S9 available on Dryad). Habitat transition and shell transition rates were estimated to be comparatively low. However, elevated rates were identified for bidirectional transitions between smooth and finely sculptured shells in lentic ('LeA'↔'LeB') habitats. In lotic habitats, transition rates towards more weakly sculptured shells ('LoC'→'LoB' and 'LoB'→'LoA') were substantially higher than transitions towards sculptured shells. Finally, the magnitude of the above-mentioned transition rates in lentic habitats were considerably higher than in lotic habitats. Interestingly, direct transitions between 'LeA'↔'LeC' and 'LoA'↔'LoC' were not observed (Fig. 4, inset).

Our bootstrapping approach showed that, independent of our incomplete taxon sampling and the high number of transition rates, a scenario of habitat-dependent asymmetric shell transitions remained the best-fit model and we obtained a reasonable precision of rate estimation (Supplementary Fig. S10 available on Dryad).

## DISCUSSION

*Taxon Sampling and Robustness of Phylogenetic Inference and Downstream Analyses*

Whereas only ~50% of recognized viviparid species are represented in our analysis (61 out of c. 125-150 recognized species or 74 OTUs), our dataset at the genus level is more comprehensive and covers 24 out of the 28 currently valid genera. The missing genera are *Amuropaludina* Moskvicheva, 1979 (Far East Russia), *Boganmargarya* Thach, 2018 (Myanmar) as well as *Eyriesia* P. Fischer, 1885 and *Glaucostracia* Ancey, 1898 from Australia. However, all these geographical regions are represented in our dataset and thus the inferred regional and global biogeographical patterns should remain robust. At the species level, the inclusion of additional taxa is generally considered to improve phylogenetic accuracy and ancestral state reconstructions (e.g., Salisbury and Kim 2001; Hug and Roger 2007; Heath et al. 2008a, 2008b; but see Nabhan and Sarkar 2012 for a recent review on this debate). However, evolutionary rates and thus time estimates seem to be robust to differences in taxon sampling if rate variation among branches is low (Soares and Schrago 2015), and we did not observe systematic differences among sampled clades and genera for any of the genetic markers used (Supplementary Fig. S5 available on Dryad). A potentially more critical issue is how even taxa were sampled across the phylogeny and thus how tree shape may be affected by missing data (e.g., Li et al. 2008). For our dataset, most of the missing (extant) taxa belong to terminal nodes and thus we assume that the general tree shape should remain robust. Moreover, shell morphologies and geographical areas are well covered in our dataset, suggesting that missing taxa would not bias biogeographical and morphological reconstructions. For example, all members of the Lioplacinae and Viviparinae occur in the same geographical region (see Fig. 1) and are also morphologically very similar, at least in terms of the defined shell sculpture types. Within the Bellamyinae, the majority of missing species belong to the genera *Bellamyia*, *Cipangopaludina*, *Idiopoma*, *Mekongia*, *Notopala*, *Sinotaia*, and *Taia*. However, these genera are all represented in the phylogeny and the



missing species mainly occur in the same geographical area and possess shell features similar to their sampled congeners. The case of *Mekongia* is somewhat different. Because it represents a more widespread genus (we included two species from Indochina and Borneo, respectively), additional *Mekongia* species may improve biogeographical reconstructions.

Very importantly, our simulations on the models of habitat-shell evolution, which are generally more sensitive to incomplete taxon sampling compared to the biogeographical models due to higher number of free parameters, supported a robust model selection. Accordingly, both best-fit models and levels of transition rates simulated for a taxon-complete phylogeny were also identified by our reduced dataset. Besides the general robustness of this approach, the random pruning of taxa from the simulated taxon-complete phylogenies also recovered almost all genera examined. We are therefore convinced that our findings very likely reflect the evolutionary history of River Snails.

#### *Phylogenetic Relationships and Systematic Implications*

The present study provides strong molecular support for the classical, mainly anatomy-based, subdivision of the Viviparidae into the three subfamilies Bellamyinae, Lioplacinae, and Viviparinae (Rohrbach 1937; Vail 1977; Bouchet et al. 2017) (Fig. 2). Of these subfamilies, a sister group relationship between Lioplacinae and Viviparinae was returned in most but not all phylogenetic analyses, although it is not highly supported. The phylogeny further confirms the placement of *Tulotoma* and *Rivularia* within the Viviparinae, as already proposed for *Tulotoma* based on comparative anatomy (Vail 1977) and for *Rivularia* by considering both anatomical and restricted molecular data (Van Bocxlaer et al. 2018). Contrary to Bouchet et al. (2017), our analysis also suggests that Campelomatinae is a synonym of Lioplacinae rather than Viviparinae. The genus *Viviparus* as traditionally conceived, with both European and North American representatives, was recovered as polyphyletic. The North American species

are more closely related to the American *Tulotoma*, and the European species more closely related to the Asian *Rivularia*. The genus-group name *Callinina* Thiele, 1931 (a replacement name for *Callina* Hannibal, 1912 *non* Lowe, 1855), with *Viviparus intertextus* (Say, 1829) as type species by typification of a replaced name, is available for the American species of *Viviparus*.

Traditionally, the systematics of viviparids at both the genus and species level has been based largely on shell characters. However, the analysis of character evolution indicates that the current shell morphology-based taxonomy does not accurately reflect phylogenetic relationships in this group. At the genus level, several cases of paraphyly and polyphyly were identified. However, this finding is understandable in the light of multiple independent origins of such morphological features. In some reconstructions, the genus *Filopaludina* was found to be monophyletic with low support, whereas in others (see Fig. 2), it was paraphyletic, with *F. tricostata* as sister to a clade that includes representatives of *Angulyagra*, *Larina*, and *Notopala*, but also with low support. Similarly, the monophyly of *Torotaia* remains uncertain owing to the ambiguous placement of *Angulyagra costata* between the different phylogenetic analyses. Assessing the monophyly of these genera will require the inclusion of additional taxa and more informative molecular markers. Moreover, in several cases resolving the systematic status of the species will require the inclusion of types. For example, both *Cipangopaludina* and *Angulyagra* were supported as polyphyletic, but resolving the question of which clade carries the genus name awaits inclusion of the respective type species. These systematic implications highlight the need for an integrative taxonomic revision of the Viviparidae, as recently initiated for the polyphyletic genus *Margarya* from China (see Du et al. 2013; Zhang et al. 2015). The present molecular phylogeny provides a robust family-level framework for such an endeavour on a global scale.

### *The Fossil Record and Evolutionary Centre of River Snails*

Both the unconstrained and fossil-constrained analyses reconstructed very similar biogeographical histories for the Viviparidae, except for the origin of the family and subsequent diversification events within the Lioplacinae and Viviparinae (Fig. 3). In our opinion, the incorporation of fossil constraints considerably improved the overall plausibility of ancestral area estimations (see also Matzke, 2013a) and thereby allowed a more robust biogeographical reconstruction for River Snails. Note, however, that the majority of cladogenetic events discussed below were not constrained by fossils (see Fig. 3).

Based on the biogeographical analyses, ancestral lineages that gave rise to the Lioplacinae and Viviparinae originated on the Laurasian supercontinent (i.e., Europe and North America; Fig. 3). Whereas a Laurasian origin contradicts the hypothesis of a Pangaeon origin for Viviparidae (see e.g., Strong et al. 2008), the former interpretation is supported by the fossil record: River Snails first appeared in Middle Jurassic deposits of western Europe (Hudleston 1896; Tracey et al. 1993). Initial diversification of the Viviparidae during the Jurassic-Cretaceous transition (Huckriede 1967; Bandel 1991; Radley and Allen 2012; see Scotese 2014a, 2014b and Supplementary Fig. S11a,b available on Dryad) was followed by the colonization of large areas of Laurasia (Valanginian; see Scotese 2014b and Supplementary Fig. S11c available on Dryad), North America (Stanton 1903; Yen 1950, 1952; Tozer 1956), and far east Eurasia (Matsukawa et al. 2006) during the Early Cretaceous. The oldest *Campeloma* was described from the Early Cretaceous (Aptian; see Scotese 2014b and Supplementary Fig. S11d available on Dryad) of Montana, USA (Yen 1952) providing evidence of the rise of the Lioplacinae during that period.

### *Age Estimates and Biogeographical Implications*

The split between North American genera and the remaining members of the Viviparinae (95% HPD = 67.8-148.3) was inferred to have been caused by vicariance. However, the large credibility interval (error bar) does not allow to make a final conclusion of whether the onset of spreading in the North Atlantic c. 95-110 Ma (e.g., Sclater et al. 1977; Jones et al. 1995) caused the initial separation within the Viviparinae (i.e., *Viviparus* + *Rivularia* vs. *Viviparus* [= *Callinina*] + *Tulotoma*; Fig. 3) or whether other geological events can be associated with this split.

Within the Viviparinae, a close relationship between the European *Viviparus* species and the Chinese *Rivularia auriculata* is strongly supported (see also Van Bocxlaer et al. 2018). Our analyses indicate that this split was either caused by vicariance or a jump-dispersal event (Fig. 3). The most parsimonious explanation would be a single colonization by a founder population out of Europe. Alternatively, *Rivularia auriculata*, which is currently restricted to the Hunan Province, could represent a relic of an ancestral lineage that expanded its range into East Asia during the Eocene (Fig. 3). Following the India-Asia and the Asia-Australia collision during the Eocene and Oligocene, respectively (see Scotese 2014c and Supplementary Fig. S11e,f available on Dryad), this region changed considerably in terms of its geology, geography, climate, and fauna (see e.g., Lohman et al. 2011; de Bruyn et al. 2014). European and Asian populations may have become geographically separated from each other (vicariance) as a consequence of these changes. However, given the absence of both extinct lineages and additional extant populations of *Rivularia*, we consider both scenarios equally likely.

Based on the present analyses, Bellamyinae-clade A probably originated in Indochina + East Asia, whereas clade B emerged in Indochina (Fig. 3). Similar patterns have been found in spiny frogs within the Dicroglossidae in the East Asian-Indochinese transition zone (Che et al. 2010). The interplay of geological and climatic dynamics during the Oligocene and

Miocene related to the uplift of the Tibetan Plateau has been hypothesized to have facilitated the diversification of these amphibians. However, the split between the two clades of Bellamyinae was estimated to have occurred in the Late Cretaceous c. 87.2 Ma (95% HPD = 63.3-113.9). Because of the complex palaeogeographical and climatic history of Southeast Asia, it remains difficult to identify the exact cause of this divergence.

The biogeographical analyses further revealed some remarkable jump-dispersal events, including the colonization of Africa and India from Indochina in the Early and Late Miocene, respectively (Fig. 3); each of these events was followed by subsequent diversification in these regions. Whereas our study supports the proposed Asian Miocene origin of the African clade (Sengupta et al. 2009; Schultheiß et al. 2014; Van Bocxlaer et al. 2018), it rejects the hypothesis that the Bellamyinae are of Gondwanan origin (*sensu* Davis 1982; see also Sengupta et al. 2009) as has been hypothesized for other widespread freshwater gastropod families such as the Ampullariidae, Pachychilidae, Pomatiopsidae, and Thiaridae (e.g., Davis 1982; Strong et al. 2008; Hayes et al. 2009). Moreover, the probable Miocene origin of the Australian and Indian species implies that the Cretaceous fossils described from these areas likely do not represent members of the Viviparidae (see Supplementary Material available on Dryad for details).

Whereas the recent study of Sil et al. (2019) estimated comparatively similar divergence times for the Bellamyinae, their calibration strategy based on two considerably young fossil calibration points (i.e., Miocene for *Bellamyia* and Pleistocene for *F. bengalensis*) suggested a much younger (Paleocene) origin for the Viviparidae and considerably larger credibility intervals towards the base of the phylogeny. This pattern is very similar to the divergence time estimates obtained from our modified BEAST analysis, in which only CP3-CP5 were enforced (Supplementary Fig. S6 and Supplementary Table S5 available on Dryad). Significantly younger ages for both the Bellamyinae and the Viviparidae have been suggested

by Gu et al. (2019) and Hirano et al. (2019a). These authors focussed on African *Bellamyia*, mainly Chinese *Sinotaia* species (Gu et al. 2019), and other East Asian taxa (Hirano et al. 2019a). According to both studies, the Bellamyinae and even the Viviparidae originated as recently as the Early Miocene (c. 20 Ma). However, the dataset of Gu et al. (2019) is taxonomically very incomplete, *Filopaludina* and *Sinotaia* species have been misplaced in *Bellamyia*, taxa previously recognized as belonging to the ingroup (i.e., *Neothauma tanganyicense*; see Sengupta et al. 2009; Schultheiß et al. 2014) have been forced into the outgroup, and their three geological and fossil calibration points are questionable and do not follow the guidelines of Parham et al. (2012). Moreover, the authors used a COI substitution rate for calibrating the tree as an alternative calibration strategy, but this rate was reported explicitly for relatively small, annual species and is likely saturated for diversification events older than 10 million years (My; see Wilke et al. 2009). Likewise, Hirano et al. (2019a) applied the fast COI rate inferred from the fossil-calibrated *Bellamyia* phylogeny of Schultheiß et al. (2014) to East Asian taxa. Whereas Gu et al. (2019) refrained from discussing their age estimates in the light of the fossil record, Hirano et al. (2019) suggested that all pre-Miocene viviparid lineages either went extinct or were simply misidentified. Substitution saturation may also have affected the divergence time estimation by Hirano et al. (2019a) and in any case led to the considerably younger divergence estimates compared to our study. This comparison highlights the challenge of incorporating fossil information in molecular phylogenies, particularly for taxa with a long evolutionary history.

#### *Biogeographical Patterns in the IAA*

With the possible exception of *Mekongia* sp. from Borneo, the colonization of insular Southeast Asia occurred comparatively late, from the Middle Miocene to the Plio-Pleistocene and thus closely reflects the dynamic geological history of this area (e.g., de Bruyn et al.

2014; see also Scotese 2014d and Supplementary Fig. S11g,h available on Dryad). East Asia and Indochina, which were identified as the main source for emigration events, have largely been established since the Middle Jurassic (e.g., Metcalfe 2011), allowing the accumulation of species over a long geological period. In contrast, several islands in the IAA were largely submerged during the Miocene and later (e.g., Java and Sumatra) or had not yet formed (e.g., the Philippines and Sulawesi; see e.g., de Bruyn et al. 2014; Nugraha and Hall 2018). The intensified colonization of the archipelago since the Late Miocene, facilitated by the extensive formation of land bridges across the IAA (see e.g., Hall 2009; de Bruyn et al. 2014), is thus consistent with palaeogeographical reconstructions and with inter-island colonization patterns observed in other invertebrates and vertebrates (Stelbrink et al. 2012; de Bruyn et al. 2014).

The analyses also identified multiple independent long-distance dispersal events from East Asia, Indochina and the Sunda Shelf into the Philippines, Sulawesi, and the Sahul Shelf (New Guinea and Australia; Fig. 3) that again highlight the permeability of renowned ‘faunal boundaries’ (i.e., Wallace’s Line and Lydekker’s Line; see e.g., Lohman et al. 2011). Moreover, the OTU-based tree used to infer biogeographical processes was constructed very conservatively, resulting in the grouping of genetically divergent island populations (Fig. 3). Consequently, some of the species, particularly within the genus *Filopaludina*, were treated as widespread and thus covered three to four of the predefined geographical areas. It is thus likely that the number of dispersal events is underestimated and that colonization within the IAA (particularly between islands within Sundaland) can be attributed to dispersal at the species level.

Whereas major, particularly older, speciation events are compatible with a vicariant hypothesis, dispersal seems to be more frequent in River Snails as previously thought (see e.g., Prasad 1928), despite their size and mobility. The biogeographical analyses showed that jump-dispersal was the most frequent range-changing cladogenetic process, whereas

anagenetic dispersal was rare (Supplementary Table S6 available on Dryad). Current global biodiversity patterns in River Snails can thus be only partly attributed to dispersal, as may perhaps be expected for such an ancient taxon. The most important process in shaping biodiversity hotspots in the Viviparidae, however, was *in situ* diversification (i.e., sympatry; see Fig. 3), particularly in isolated (and comparatively ancient) systems such as the African Great Lakes and the islands within the IAA.

### *Patterns and Drivers of Shell Evolution*

Diversification within the Viviparidae was accompanied by major habitat shifts. The present study suggests that River Snails evolved from a lotic and smooth-shelled ancestor (Fig. 4). This scenario is supported in both unconstrained and fossil-constrained analyses, and it is also confirmed by the fossil record: the oldest fossil viviparids from the Jurassic and Cretaceous all lack spiral shell sculpture (Hudleston 1896; Stanton 1903; Yen 1950; Huckriede 1967; Bandel 1991; Tracey et al. 1993; Radley and Allen 2012), with the possible exception of the North American *Tulotomops* and *Lioplacodes*, both of which cannot be unambiguously assigned to Viviparidae (see also Supplementary Material available on Dryad). However, as the habitat type for fossils is often questionable (see Supplementary Material available on Dryad), the environmental settings in which these earliest representatives originated remain uncertain.

Species with unsculptured shells predominate in extant River Snails, particularly in the oldest subfamilies Lioplacinae and Viviparinae, and smooth shells were also estimated to be the most likely ancestral state for all three major clades (Fig. 4). In contrast, the highest variability in shell sculpture is found in the Bellamyinae, particularly within clade A. Interestingly, various extinct taxa also evolved heavily sculptured shells. Renowned examples include the morphologically disparate Plio-Pleistocene fauna of the genus *Viviparus* from



Greece (Willmann 1985), the African fossil assemblage of *Neothauma* since the Miocene (see e.g., Van Damme and Pickford 1999; Salzburger et al. 2014), and heavily sculptured *Margarya* forms identified from Oligocene deposits in China (e.g., Tian et al. 2013, 2018). The present results thus strongly indicate that such fine and heavily sculptured shells evolved multiple times independently across various viviparid lineages (Fig. 4), providing the first phylogenetic evidence for iterative shell evolution in River Snails. However, the repeated evolution of shell sculpture in both the extant and extinct faunas raises the question of whether underlying factors can be identified.

Shell sculpture in freshwater gastropods may not necessarily represent an adaptive response (see e.g., Gorthner 1992), but it has often been attributed to limnological parameters including habitat differences (e.g., Annandale 1919, 1924) and predatory pressure (e.g., Vermeij and Covich 1978; Geary et al. 2002; Rasser and Covich 2014). Whereas the latter factor cannot be addressed here, habitat information available for the extant and some of the fossil species used to calibrate the phylogeny allows a discussion on environmental differences. In extant River Snails, species with smooth or finely sculptured shells mainly inhabit lotic habitats (Fig. 4). In contrast, those with coarse spiral sculpture are exclusively found in lentic environments (except for *Tulotoma magnifica*), and predominantly in (putative) ancient lakes (i.e., Yunnan Plateau lakes, China: *Anularya*, *Margarya*, and *Tchangmargarya*; Lake Inlé, Myanmar: *Taia*; Lake Poso, Indonesia: *Celetaia*; and Lake Lanao, Philippines: *Torotaia*). These species mainly belong to the Bellamyinae-clade A, which may also be characterized by larger and more slender shells compared to clade B (Supplementary Fig. S8 and Supplementary Table S7 available on Dryad). However, these apparent differences are not supported by the phylogenetic ANOVAs, suggesting that neutral morphological evolution explains the observed size differences.

The present analysis of habitat-shell evolution suggests that the evolution of sculptured shells was associated with major habitat shifts in some cases (Fig. 4). This is most noticeable in Bellamyinae–clade A, in which the highest shell disparity is found today (Fig. 4). In contrast, a habitat transition from lotic to lentic identified in the extant Lake Malawi species flock of *Bellamyia* (clade B; see also Schultheiß et al. 2011), for example, did not give rise to a similar extant variability in shell sculpture. This pattern can be best explained by the estimated habitat-shell transition rates (Fig. 4, inset). According to the best-fit model (and other well supported models; Supplementary Table S8 available on Dryad), shell transitions were habitat-dependent, i.e., they differed between lentic and lotic environments. Moreover, the best-fit model revealed very low habitat transition rates. This finding suggests that changes in shell form mainly occurred *in situ*. However, in lotic habitats, coarsely sculptured shells (‘LoC’) were solely gained via habitat shifts and never derived *in situ*. Because transition rates between habitats were considerably lower than those between shell forms, coarse sculpture is very rare in lotic species and in extant viviparids only represented by *Tulotoma magnifica* (Viviparinae; see Fig. 4). This low prevalence is further affected by high transition rates from ‘coarse’ to ‘smooth’ via ‘finely’ sculptured shells (‘LoC’→‘LoB’→‘LoA’), resulting in a net increase of smooth shells in lotic habitats. In lentic habitats, however, the highest transition rates were reciprocal between smooth and finely sculptured shells (‘LeA’↔‘LeB’), and the magnitude of these shell transition rates was much higher than those in lotic habitats. This finding implies an evolutionary versatility between these two shell types as a response to different environmental conditions in lentic habitats.

Similar to our results, data from the viviparid fossil record, for example, on European *Viviparus* (e.g. Willmann 1985; Posilović and Bajraktarević 2010), indicates that changes from smooth to sculptured shells relate to a change between lotic and lentic habitats. Whereas

our findings suggest a link between habitat and sculpture, more research is required to understand the mechanisms driving shell sculpture and how these processes are affected by abiotic and biotic factors. High rates of gain and loss in shell sculpture explain why coarsely sculptured shells are predominantly found in Bellamyinae–clade A in several independent lineages, and not more extensively in clade B, which comprises various extinct species from the African Great Lakes that are renowned for their high shell disparity (see e.g., Van Damme and Pickford 1999; Salzburger et al. 2014). Such high rates of shell evolution also explain the sympatric occurrence of closely related but morphologically disparate species in a single lake system such as Lake Inlé in Myanmar (Annandale 1919) and Lake Lanao in the Philippines (Bartsch 1907; Stelbrink et al. 2019), and why both weakly and heavily sculptured genera (*Sinotaia* vs. *Margarya*) co-occur in several Yunnan Plateau lakes (see e.g., Zhang et al. 2015). Whereas the generally low dispersal ability together with high transition rates in lentic habitats triggered both *in situ* diversification and a high versatility in shell evolution, a selective pressure towards smooth shells seems to predominate in lotic habitats.

## REFERENCES

- Annandale N. 1919. The gastropod fauna of old lake-beds in Upper Burma. *Rec. Geol. Surv. India* 50:209–240.
- Annandale N. 1924. The evolution of the shell-sculpture in fresh-water snails of the family Viviparidae. *Proc. R. Soc. London B* 96:60–76.
- Arlt T. 1917-1922. *Handbuch der Palaeogeographie*. Volume I (Palaeaktologie) and Volume II (Palaeogeographie). Leipzig: Gebrüder Borntraeger.
- Bandel K. 1991. Gastropods from brackish and fresh water of the Jurassic-Cretaceous transition (a systematic reevaluation). *Berliner Geowissenschaftliche Abhandlungen, Reihe A* 134:9–55.
- Bartsch P. 1907. The Philippine pond snails of the genus *Vivipara*. *Proc. United States Natl. Museum* 32:135–150.
- Bensasson D., Zhang D.-X., Hartl D.L., Hewitt G.M. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends Ecol. Evol.* 16:314–321.
- Bouchet P., Rocroi J.-P., Hausdorf B., Kaim A., Kano Y., Nützel A., Parkhaev P., Schrödl M., Strong E.E. 2017. Revised classification, nomenclator and typification of gastropod and monoplacophoran families. *Malacologia* 61:1–526.
- Briggs J.C. 2009. Darwin's biogeography. *J. Biogeogr.* 36:1011–1017.
- Carini G., Hughes J.M. 2006. Subdivided population structure and phylogeography of an endangered freshwater snail, *Notopala sublineata* (Conrad, 1850) (Gastropoda: Viviparidae), in Western Queensland, Australia. *Biol. J. Linn. Soc.* 88:1–16.
- Carini G., Hughes J.M., Bunn S.E. 2006. The role of waterholes as “refugia” in sustaining genetic diversity and variation of two freshwater species in dryland river systems (Western Queensland, Australia). *Freshw. Biol.* 51:1434–1446.
- Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in

- phylogenetic analysis. *Mol. Biol. Evol.* 17:540–552.
- Che J., Zhou W.-W., Hu J.-S., Yan F., Papenfuss T.J., Wake D.B., Zhang Y.-P. 2010. Spiny frogs (Paini) illuminate the history of the Himalayan region and Southeast Asia. *Proc. Natl. Acad. Sci. USA* 107:13765–13770.
- Chiu Y.W., Chen C.A., Chen H.C. 2002. Genetic variation of the viviparid snail, *Sinotaia quadrata* (Gastropod: Viviparidae), in Taiwan. *Acta Zool. Taiwanica* 13:1–10.
- Colgan D.J., Ponder W.F., Beacham E., Macaranas J. 2007. Molecular phylogenetics of Caenogastropoda (Gastropoda: Mollusca). *Mol. Phylogenet. Evol.* 42:717–737.
- Cowie R.H., Holland B.S. 2008. Molecular biogeography and diversification of the endemic terrestrial fauna of the Hawaiian Islands. *Philos. Trans. R. Soc. London B* 363:3363–3376.
- Crayn D.M., Costion C., Harrington M.G. 2015. The Sahul-Sunda floristic exchange: dated molecular phylogenies document Cenozoic intercontinental dispersal dynamics. *J. Biogeogr.* 42:11–24.
- Crisp M.D., Trewick S.A., Cook L.G. 2011. Hypothesis testing in biogeography. *Trends Ecol. Evol.* 26:66–72.
- Davis G.M. 1982. Historical and ecological factors in the evolution, adaptive radiation, and biogeography of freshwater mollusks. *Am. Zool.* 22:375–395.
- de Bruyn M., Stelbrink B., Morley R.J., Hall R., Carvalho G.R., Cannon C.H., van den Bergh G., Meijaard E., Metcalfe I., Boitani L., Maiorano L., Shoup R., von Rintelen T. 2014. Borneo and Indochina are major evolutionary hotspots for Southeast Asian biodiversity. *Syst. Biol.* 63:879–901.
- de Queiroz A. 2005. The resurrection of oceanic dispersal in historical biogeography. *Trends Ecol. Evol.* 20:68–73.
- Drummond A.J., Suchard M.A., Xie D., Rambaut A. 2012. Bayesian phylogenetics with

- BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29:1969–1973.
- Du L., Yang J., von Rintelen T., Chen X., Aldridge D. 2013. Molecular phylogenetic evidence that the Chinese viviparid genus *Margarya* (Gastropoda: Viviparidae) is polyphyletic. *Chinese Sci. Bull.* 58:2154–2162.
- Dupin J., Matzke N.J., Särkinen T., Knapp S., Olmstead R.G., Bohs L., Smith S.D. 2017. Bayesian estimation of the global biogeographical history of the Solanaceae. *J. Biogeogr.* 44:887–899.
- du Toit A.L. 1937. *Our wandering continents: an hypothesis of continental drifting*. London: Oliver & Boyd.
- FitzJohn R.G. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods Ecol. Evol.* 3:1084–1092.
- Garland T., Dickerman A.W., Janis C.M., Jones J.A. 1993. Phylogenetic analysis of covariance by computer simulation. *Syst. Biol.* 42:265–292.
- Geary D.H., Staley A.W., Müller P., Magyar I., Spring N. 2002. Iterative changes in Lake Pannon *Melanopsis* reflect a recurrent theme in gastropod morphological evolution. *Paleobiology* 28:208–221.
- Gillespie R.G., Baldwin B.G., Waters J.M., Fraser C.I., Nikula R., Roderick G.K. 2012. Long-distance dispersal: a framework for hypothesis testing. *Trends Ecol. Evol.* 27:47–56.
- Gorthner A. 1992. Bau, Funktion und Evolution komplexer Gastropodenschalen in Langzeit-Seen. Mit einem Beitrag zur Paläobiologie von *Gyraulus* “*multiformis*” im Steinheimer Becken. *Stuttgarter Beiträge zur Naturkunde, Serie B* 190:1–173.
- Gu Q.-H., Husemann M., Ding B., Luo Z., Xiong B.-X. 2015a. Population genetic structure of *Bellamyia aeruginosa* (Mollusca: Gastropoda: Viviparidae) in China: weak divergence across large geographic distances. *Ecol. Evol.* 5:4906–4919.

- Gu Q.-H., Zhou C.-J., Cheng Q.-Q., Li X.-J., Zhu G.-R., Zhang M., Gao Y.-N., Dong J. 2015b. The perplexing population genetic structure of *Bellamya purificata* (Gastropoda: Viviparidae): low genetic differentiation despite low dispersal ability. *J. Molluscan Stud.* 81:466–475.
- Gu Q.H., Husemann M., Wu H.H., Dong J., Zhou C.J., Wang X.F., Gao Y.N., Zhang M., Zhu G.R., Nie G.X. 2019. Phylogeography of *Bellamya* (Mollusca: Gastropoda: Viviparidae) snails on different continents: contrasting patterns of diversification in China and East Africa. *BMC Evol. Biol.* 19:82.
- Guindon S., Dufayard J.F., Lefort V., Anisimova M., Hordijk W., Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59:307–321.
- Hall R. 2009. Southeast Asia's changing palaeogeography. *Blumea* 54:148–161.
- Hallam A. 1981. Plate tectonics, biogeography and evolution. *Nature* 293:31–32.
- Hayes K.A., Cowie R.H., Thiengo S.C. 2009. A global phylogeny of apple snails: Gondwanan origin, generic relationships, and the influence of outgroup choice (Caenogastropoda: Ampullariidae). *Biol. J. Linn. Soc.* 98:61–76.
- Heath T.A., Hedtke S.M., Hillis D.M. 2008a. Taxon sampling and the accuracy of phylogenetic analyses. *J. Syst. Evol.* 46:239–257.
- Heath T.A., Zwickl D.J., Kim J., Hillis D.M. 2008b. Taxon sampling affects inferences of macroevolutionary processes from phylogenetic trees. *Syst. Biol.* 57:160–166.
- Hirano T., Saito T., Chiba S. 2015. Phylogeny of freshwater viviparid snails in Japan. *J. Molluscan Stud.* 81:435–441.
- Hirano T., Saito T., Tsunamoto Y., Koseki J., Prozorova L., Do V.T., Matsuoka K., Nakai K., Suyama Y., Chiba S. 2019a. Role of ancient lakes in genetic and phenotypic diversification of freshwater snails. *Mol. Ecol.* 28:5032–5051.

- Hirano T., Saito T., Tsunamoto Y., Koseki J., Ye B., Do V.T., Miura O., Suyama Y., Chiba S. 2019b. Enigmatic incongruence between mtDNA and nDNA revealed by multi-locus phylogenomic analyses in freshwater snails. *Sci. Rep.* 9:6223.
- Ho S.Y.W., Tong K.J., Foster C.S.P., Ritchie A.M., Lo N., Crisp M.D. 2015. Biogeographic calibrations for the molecular clock. *Biol. Lett.* 11:20150194.
- Huckriede R. 1967. Molluskenfaunen mit limnischen und brackischen Elementen aus Jura, Serpulit und Wealden NW-Deutschlands und ihre paläogeographische Bedeutung. *Beihefte zum Geol. Jahrb.* 67:1–244.
- Hudleston W.H. 1887-1896. A monograph of the inferior Oolite Gasteropoda. Being part I of the British Jurassic Gasteropoda. London: Palaeontological Society London.
- Huelsensbeck J.P., Nielsen R., Bollback J.P. 2003. Stochastic mapping of morphological characters. *Syst. Biol.* 52:131–158.
- Hug L.A., Roger A.J. 2007. The impact of fossils and taxon sampling on ancient molecular dating analyses. *Mol. Biol. Evol.* 24:1889–1897.
- Johnson S.G., Bragg E. 1999. Age and polyphyletic origins of hybrid and spontaneous parthenogenetic *Campeloma* (Gastropoda: Viviparidae) from the southeastern United States. *Evolution* 53:1769–1781.
- Johnson S.G., Leefe W.R. 1999. Clonal diversity and polyphyletic origins of hybrid and spontaneous parthenogenetic *Campeloma* (Gastropoda: Viviparidae) from the southeastern United States. *J. Evol. Biol.* 12:1056–1068.
- Jones E.J.W., Cande S.C., Spathopoulos F. 1995. Evolution of a major oceanographic pathway: the equatorial Atlantic. *Geol. Soc. London, Spec. Publ.* 90:199–213.
- Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30:772–780.
- Katoh K., Toh H. 2008. Recent developments in the MAFFT multiple sequence alignment



- program. *Brief. Bioinform.* 9:286–298.
- Klaus S., Morley R.J., Plath M., Zhang Y.P., Li J.T. 2016. Biotic interchange between the Indian subcontinent and mainland Asia through time. *Nat. Commun.* 7:12132.
- Köhler F., Glaubrecht M. 2010. Uncovering an overlooked radiation: molecular phylogeny and biogeography of Madagascar's endemic river snails (Caenogastropoda: Pachychilidae: *Madagasikara* gen. nov.). *Biol. J. Linn. Soc.* 99:867–894.
- Lanfear R., Calcott B., Ho S.Y.W., Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* 29:1695–1701.
- Lanfear R., Frandsen P.B., Wright A.M., Senfeld T., Calcott B. 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* 34:772–773.
- Li G., Steel M., Zhang L. 2008. More taxa are not necessarily better for the reconstruction of ancestral character states. *Syst. Biol.* 57:647–653.
- Lohman D.J., de Bruyn M., Page T., von Rintelen K., Hall R., Ng P.K.L., Shih H.T., Carvalho G.R., von Rintelen T. 2011. Biogeography of the Indo-Australian archipelago. *Annu. Rev. Ecol. Evol. Syst.* 42:205–226.
- Marshall C.R. 1990. Confidence intervals on stratigraphic ranges. *Paleobiology* 16:1–10.
- Matsukawa M., Ito M., Nishida N., Koarai K., Lockley M.G., Nichols D.J. 2006. The Cretaceous Tetori biota in Japan and its evolutionary significance for terrestrial ecosystems in Asia. *Cretac. Res.* 27:199–225.
- Matzke N.J. 2013a. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Front. Biogeogr.* 5:242–248.
- Matzke N.J. 2013b. BioGeoBEARS: BioGeography with Bayesian (and Likelihood)

- evolutionary analysis in R scripts. University of California, Berkeley, CA.
- Matzke N.J. 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Syst. Biol.* 63:951–970.
- McIntyre S.R.N., Lineweaver C.H., Groves C.P., Chopra A. 2017. Global biogeography since Pangaea. *Proc. R. Soc. London B* 284:20170716.
- Metcalf I. 2011. Palaeozoic-Mesozoic history of SE Asia. In: Hall R., Cottam M.A., Wilson M.E.J., editors. *The Southeast Asian Gateway: history and tectonics of Australia-Asia collision*. London: The Geological Society of London. p. 7–35.
- Miller M.A., Pfeiffer W., Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop (GCE)*:1–8.
- MolluscaBase. 2019. *Viviparidae* Gray, 1847. Accessed at: <http://www.molluscabase.org/aphia.php?p=taxdetails&id=154000> on 2020-01-23.
- Nabhan A.R., Sarkar I.N. 2012. The impact of taxon sampling on phylogenetic inference: a review of two decades of controversy. *Brief. Bioinform.* 13:122–134.
- Nugraha A.M.S., Hall R. 2018. Late Cenozoic paleogeography of Sulawesi, Indonesia. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 490:191–209.
- Pagel M. 1994. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc. R. Soc. London B* 255:37–45.
- Parham J.F., Donoghue P.C.J., Bell C.J., Calway T.D., Head J.J., Holroyd P.A., Inoue J.G., Irmis R.B., Joyce W.G., Ksepka D.T., Patané J.S.L., Smith N.D., Tarver J.E., van Tuinen M., Yang Z., Angielczyk K.D., Greenwood J.M., Hipsley C.A., Jacobs L., Makovicky P.J., Müller J., Smith K.T., Theodor J.M., Warnock R.C.M., Benton M.J. 2012. Best practices for justifying fossil calibrations. *Syst. Biol.* 61:346–59.
- Pickford M. 2004. Palaeoenvironments of Early Miocene hominoid-bearing deposits at

- Napak, Uganda, based on terrestrial molluscs. *Ann. Paléontologie* 90:1–12.
- Ponder W.F., Colgan D.J., Healy J.M., Nützel A., Simone L.R.L., Strong E.E. 2008. Caenogastropoda. In: Ponder W.F., Lindberg D.R., editors. *Phylogeny and evolution of the Mollusca*. Berkeley: University of California Press. p. 331–383.
- Posilović H., Bajraktarević Z. 2010. Functional morphological analysis of evolution of ribbing in Pliocene viviparid shells from Croatia. *Lethaia* 43:457–464.
- Prashad B. 1928. Recent and fossil Viviparidae. A study in distribution, evolution and palaeogeography. *Mem. Indian Museum* 8:153–251.
- Quan C., Fu Q., Shi G., Liu Y., Li L., Liu X., Jin J. 2016. First Oligocene mummified plant Lagerstätte at the low latitudes of East Asia. *Sci. China Earth Sci.* 59:445–448.
- Radley J.D., Allen P. 2012. The Wealden (non-marine Lower Cretaceous) of the Weald Sub-basin, southern England. *Proc. Geol. Assoc.* 123:245–318.
- Rambaut A., Drummond A.J. 2007. Tracer v. 1.5. Accessed at: <http://tree.bio.ed.ac.uk/software/tracer>.
- Rasser M.W., Covich A.P. 2014. Predation on freshwater snails in Miocene Lake Steinheim: a trigger for intralacustrine evolution? *Lethaia* 47:524–532.
- R Core Team. 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Ree R.H., Smith S.A. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst. Biol.* 57:4–14.
- Revell L.J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3:217–223.
- Rohrbach F. 1937. Oekologische und morphologische Untersuchungen an *Viviparus* (*Bellamya*) *capillatus* Frauenfeld und *Viviparus* (*Bellamya*) *unicolor* Olivier, unter Berücksichtigung anderer tropischer Formen und im Hinblick auf phyletische

- Beziehungen. Arch. für Molluskenkd. 69:177–218.
- Ronquist F., Sanmartín I. 2011. Phylogenetic methods in biogeography. Annu. Rev. Ecol. Evol. Syst. 42:441–464.
- Ronquist F., Teslenko M., van der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61:539–542.
- Rysiewska A., Hofman S., Osikowski A., Beran L., Pešić V., Falniowski A. 2019. *Viviparus mammillatus* (Küster, 1852), and partial congruence between the morphology-, allozyme- and DNA-based phylogeny in European Viviparidae (Caenogastropoda: Architaenioglossa). Folia Malacol. 27:43–51.
- Salisbury B.A., Kim J. 2001. Ancestral state estimation and taxon sampling density. Syst. Biol. 50:557–564.
- Salzburger W., Van Bocxlaer B., Cohen A.S. 2014. Ecology and evolution of the African Great Lakes and their faunas. Annu. Rev. Ecol. Evol. Syst. 45:519–545.
- Schultheiß R., Van Bocxlaer B., Riedel F., von Rintelen T., Albrecht C. 2014. Disjunct distributions of freshwater snails testify to a central role of the Congo system in shaping biogeographical patterns in Africa. BMC Evol. Biol. 14:42.
- Schultheiß R., Wilke T., Jørgensen A., Albrecht C. 2011. The birth of an endemic species flock: demographic history of the *Bellamya* group (Gastropoda, Viviparidae) in Lake Malawi. Biol. J. Linn. Soc. 102:130–143.
- Sclater J.G., Hellinger S., Tapscott C. 1977. The paleobathymetry of the Atlantic Ocean from the Jurassic to the Present. J. Geol. 85:509–552.
- Scotese C.R. 2014a. Atlas of Jurassic Paleogeographic Maps (Mollweide Projection), Maps 32-42, Volume 4, The Jurassic and Triassic, PALEOMAP Atlas for ArcGIS, PALEOMAP Project, Evanston, IL.

- Scotese C.R. 2014b. Atlas of Early Cretaceous Maps (Mollweide Projection), Maps 23-31, Volume 2, The Cretaceous, PALEOMAP Atlas for ArcGIS, PALEOMAP Project, Evanston, IL.
- Scotese C.R. 2014c. Atlas of Paleogene Paleogeographic Maps (Mollweide Projection), Maps 8-15, Volume 1, The Cenozoic, PALEOMAP Atlas for ArcGIS, PALEOMAP Project, Evanston, IL.
- Scotese C.R. 2014d. Atlas of Neogene Paleogeographic Maps (Mollweide Projection), Maps 1-7, Volume 1, The Cenozoic, PALEOMAP Atlas for ArcGIS, PALEOMAP Project, Evanston, IL.
- Sengupta M.E., Kristensen T.K., Madsen H., Jorgensen A. 2009. Molecular phylogenetic investigations of the Viviparidae (Gastropoda: Caenogastropoda) in the lakes of the Rift Valley area of Africa. *Mol. Phylogenet. Evol.* 52:797–805.
- Shaw K.L., Gillespie R.G. 2016. Comparative phylogeography of oceanic archipelagos: hotspots for inferences of evolutionary process. *Proc. Natl. Acad. Sci. USA* 113:7986–7993.
- Sil M., Aravind N.A., Karanth K.P. 2019. Role of geography and climatic oscillations in governing into-India dispersal of freshwater snails of the family: Viviparidae. *Mol. Phylogenet. Evol.* 138:174–181.
- Soares A.E.R., Schrago C.G. 2015. The influence of taxon sampling on Bayesian divergence time inference under scenarios of rate heterogeneity among lineages. *J. Theor. Biol.* 364:31–39.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313.
- Stanton T.W. 1903. A new fresh-water molluscan faunule from the cretaceous of Montana. *Proc. Am. Philos. Soc.* 42:188–199.

- Stelbrink B., Albrecht C., Hall R., von Rintelen T. 2012. The biogeography of Sulawesi revisited: is there evidence for a vicariant origin of taxa on Wallace's "anomalous island"? *Evolution* 66:2252–2271.
- Stelbrink B., von Rintelen T., Albrecht C., Clewing C., Naga P.O. 2019. Forgotten for decades: Lake Lanao and the genetic assessment of its mollusc diversity. *Hydrobiologia* 843:31–49.
- Stiller M., Knapp M., Stenzel U., Hofreiter M., Meyer M. 2009. Direct multiplex sequencing (DMPS) – a novel method for targeted high-throughput sequencing of ancient and highly degraded DNA. *Genome Res.* 19:1843–1848.
- Strong E.E., Colgan D.J., Healy J.M., Lydeard C., Ponder W.F., Glaubrecht M. 2011. Phylogeny of the gastropod superfamily Cerithioidea using morphology and molecules. *Zool. J. Linn. Soc.* 162:43–89.
- Strong E.E., Gargominy O., Ponder W.F., Bouchet P. 2008. Global diversity of gastropods (Gastropoda; Mollusca) in freshwater. *Hydrobiologia* 595:149–166.
- Tian Y., Fürsich F.T., Schneider S. 2013. Giant Viviparidae (Gastropoda: Architaenioglossa) from the Early Oligocene of the Nanning Basin (Guangxi, SE China). *Neues Jahrb. für Geol. und Paläontologie, Abhandlungen* 267:75–87.
- Tian Y., Shaw D., Schneider S. 2018. Oligocene fossil assemblages from Lake Nanning (Yongning Formation; Nanning Basin, Guangxi Province, SE China): biodiversity and evolutionary implications. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 505:100–119.
- Tozer E.T. 1956. Uppermost Cretaceous and Paleocene non-marine molluscan faunas of western Alberta. *Geol. Surv. Canada Mem.* 280:1–125.
- Tracey S., Todd J.A., Erwin D.H. 1993. Mollusca: Gastropoda. In: Benton M.J., editor. *The fossil record* 2. London: Chapman and Hall. p. 131–167.
- Upchurch P. 2008. Gondwanan break-up: legacies of a lost world? *Trends Ecol. Evol.*

23:229–236.

- Vail V.A. 1977. Comparative anatomy of 3 viviparid gastropods. *Malacologia* 16:519–540.
- Van Bocxlaer B., Strong E.E. 2019. Viviparidae, Gray 1847. In: Lydeard C., Cummings K.S., editors. *Freshwater mollusks of the world: a distribution atlas*. Baltimore: Johns Hopkins University Press. p. 43–50.
- Van Bocxlaer B., Strong E.E., Richter R., Stelbrink B., von Rintelen T. 2018. Anatomical and genetic data reveal that *Rivularia* Heude, 1890 belongs to Viviparinae (Gastropoda: Viviparidae). *Zool. J. Linn. Soc.* 182:1–23.
- Van Damme D., Pickford M. 1999. The late Caenozoic Viviparidae (Mollusca, Gastropoda) of the Albertine Rift Valley (Uganda-Congo). *Hydrobiologia* 390:171–217.
- Vermeij G.J., Covich A.P. 1978. Coevolution of freshwater gastropods and their predators. *Am. Nat.* 112:833–843.
- von Rintelen T., Stelbrink B., Marwoto R.M., Glaubrecht M. 2014. A snail perspective on the biogeography of Sulawesi, Indonesia: origin and intra-island dispersal of the viviparous freshwater gastropod *Tylomelania*. *PLoS ONE* 9:e98917.
- Wegener A. 1912a. Die Entstehung der Kontinente. *Petermanns Geogr. Mitt.*:185–195.
- Wegener A. 1912b. Die Entstehung der Kontinente. *Petermanns Geogr. Mitt.*:253–256.
- Wegener A. 1912c. Die Entstehung der Kontinente. *Petermanns Geogr. Mitt.*:305–309.
- Wiley E.O. 1988. Vicariance biogeography. *Annu. Rev. Ecol. Syst.* 19:513–542.
- Wilke T., Schultheiß R., Albrecht C. 2009. As time goes by: a simple fool's guide to molecular clock approaches in invertebrates. *Am. Malacol. Bull.* 27:25–45.
- Willmann R. 1985. Responses of the Plio-Pleistocene freshwater gastropods of Kos (Greece, Aegean sea) to environmental changes. *Lect. Notes Earth Sci.* 1:295–321.
- Winnepenninckx B., Backeljau T., De Wachter R. 1993. Extraction of high molecular weight DNA from molluscs. *Trends Genet.* 9:407.

- Woodruff D.S. 2010. Biogeography and conservation in Southeast Asia: how 2.7 million years of repeated environmental fluctuations affect today's patterns and the future of the remaining refugial-phase biodiversity. *Biodivers. Conserv.* 19:919–941.
- Yen T.-C. 1950. Fresh-water mollusks of Cretaceous age from Montana and Wyoming. *Geol. Surv. Prof. Pap.* 233-A:1–20.
- Yen T.-C. 1952. Molluscan fauna of the Morrison Formation. *U.S. Geol. Surv. Prof. Pap.* 223:21–51.
- Yu Y., Harris A.J., Blair C., He X. 2015. RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Mol. Phylogenet. Evol.* 87:46–49.
- Zhang L.-J., Chen S.-C., Yang L.-T., Jin L., Köhler F. 2015. Systematic revision of the freshwater snail *Margarya* Nevill, 1877 (Mollusca: Viviparidae) endemic to the ancient lakes of Yunnan, China, with description of new taxa. *Zool. J. Linn. Soc.* 174:760–800.
- Zielske S., Ponder W.F., Haase M. 2017. The enigmatic pattern of long-distance dispersal of minute freshwater gastropods (Caenogastropoda, Truncatelloidea, Tateidae) across the South Pacific. *J. Biogeogr.* 44:195–206.



## SUPPORTING INFORMATION

Data available from the Dryad Digital Repository: [http://dx.doi.org/10.5061/dryad.\[NNNN\]](http://dx.doi.org/10.5061/dryad.[NNNN]).

## FUNDING

This work was supported by DFG (German Science Foundation) grants to B.S. (DFG STE 2460/1-1 and STE 2460/2-1) and T.v.R. (DFG RI 1738/3-1). B.V.B. was supported by ANR-JCJC-EVOLINK of the French Agence Nationale de la Recherche. Sampling in Russia was partly funded by the Russian federal project no. 0345-2019-0009 to A.A.S.

## ACKNOWLEDGEMENTS

We thank Giovannella Carini and Ben Cook (Griffith University, Brisbane, Australia), Reuben Clements (Sunway University, Selangor, Malaysia), David Dudgeon (University of Hong Kong, China), Jeffrey Garner (Alabama Department of Conservation and Natural Resources, USA), Matthias Glaubrecht (Universität Hamburg, Centrum für Naturkunde, Germany), Paul Johnson (Alabama Aquatic Biodiversity Center, USA), Larisa Prozorova (Far Eastern Branch Russian Academy of Sciences, Vladivostok, Russia), Suriani Surbakti (Universitas Cenderawasih, Jayapura, Indonesia) and Daisy Wowor (Research Center for Biology, LIPI, Cibinong, Indonesia) for field assistance, providing specimens, and species determination. Nicholas J. Matzke (Australian National University, Canberra, Australia) provided useful advice on the biogeographical analyses and commented on a previous version of the manuscript. Catharina Clewing (Justus Liebig University Giessen, Germany) and Christine Zorn (Museum für Naturkunde Berlin, Germany) took pictures of additional voucher specimens. André Billion (Fraunhofer Institute for Molecular Biology and Applied Ecology, Giessen, Germany) kindly provided computer resources to run the biogeographical analyses. Further computational resources were provided by the BMBF-funded (Federal Ministry of Education and Research, Germany) de.NBI Cloud within the German Network for Bioinformatics Infrastructure (de.NBI). Finally, we thank Bryan Carstens and two anonymous reviewers for their helpful comments on improving this manuscript.

## FIGURE CAPTIONS

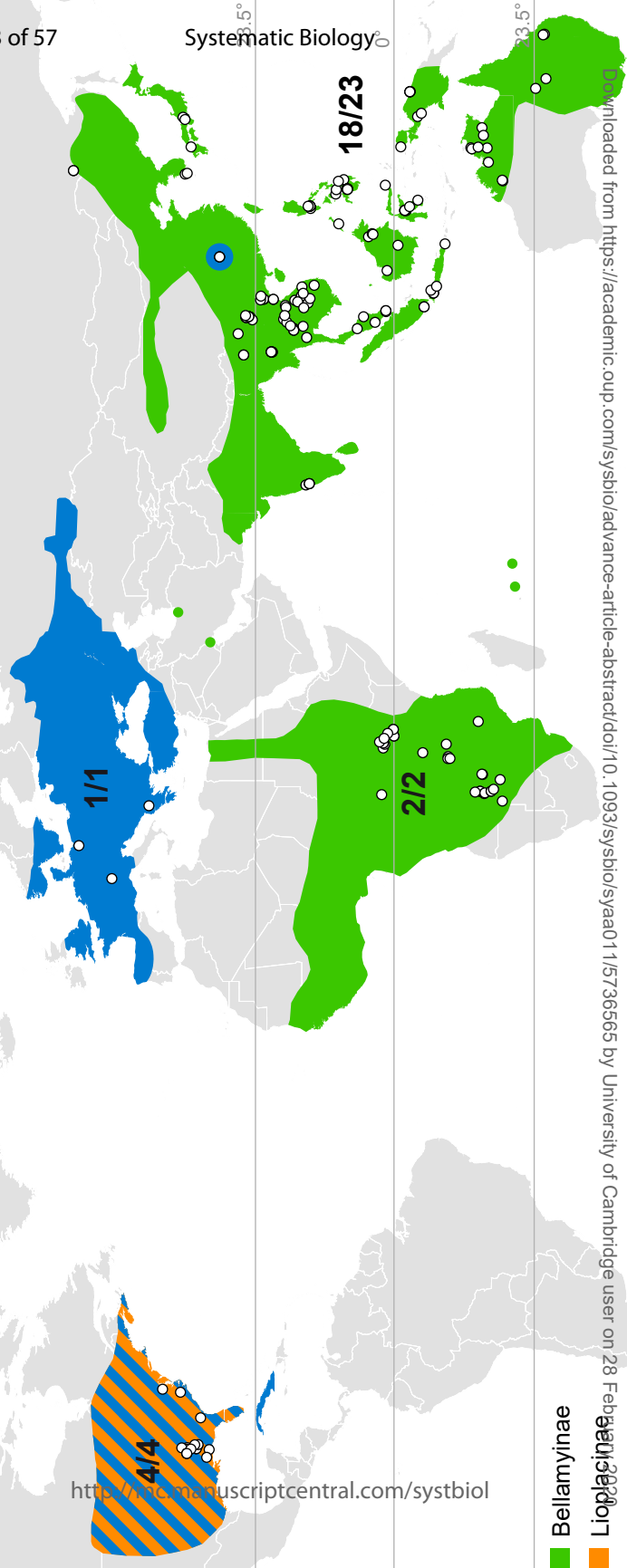
FIGURE 1. Distribution map of River Snails (Viviparidae) and their subfamilies (modified from Van Bocxlaer and Strong 2019; see Figure S1 for a detailed map of samples collected in Asia/Australia). White circles represent sample sites of specimens used in the present study. Note that *Viviparus* occurs in both North America (= *Callinina*; see Discussion) and the western Palearctic and that the monotypic genus *Rivularia* reported from the Hunan Province of China belongs to the Viviparinae (see Van Bocxlaer et al. 2018). Bold numbers represent number of extant genera described and sampled in our study.

FIGURE 2. OTU-based tree of Viviparidae created from the original BEAST MCC tree based on the genes 28S rRNA, COI, and H3 including node ages, error bars (95% HPD), and support values (posterior probabilities).

FIGURE 3. Unconstrained vs. fossil-constrained biogeographical analysis of Viviparidae based on 13 predefined geographical areas, the best-fit biogeographical model (both DEC+J), and error bars (95% HPD) obtained from the BEAST analysis. For simplicity, ancestral states are only shown when they differ from the ancestral node. Colour-coded branches and nodes refer to different cladogenetic and anagenetic events identified by the different analyses.

FIGURE 4. Ancestral state estimation of Viviparidae for habitat type and shell sculpture using stochastic character mapping and the best-fit model (HabDep-Hab<sub>ER</sub>-Shell<sub>ARD</sub>) for the fossil-constrained analysis (see text for details). For simplicity, ancestral states for the unconstrained analysis are only shown if they differ considerably from the fossil-constrained analysis. Colour gradients across the branches represent habitat transitions identified by the fossil-constrained analyses. Representative shell images for each of the three spiral sculpture

categories (from left to right): *Viviparus* cf. *contectus* ('absent'), *Torotaia* cf. *lanaonis* ('fine'), and *Celetaia persculpta* ('coarse'). The upper inset shows the transition rates calculated with the best-fit model for the six predefined habitat-shell states obtained from the fossil-constrained analysis: 'LeA', 'LeB', 'LeC', 'LoA', 'LoB', and 'LoC'. Horizontal arrows denote equally constrained transition rates for habitat transitions; transition rates between shell types were all allowed to be different as indicated by the colours (see text for details).



- Bellamyinae
- Viviparinae

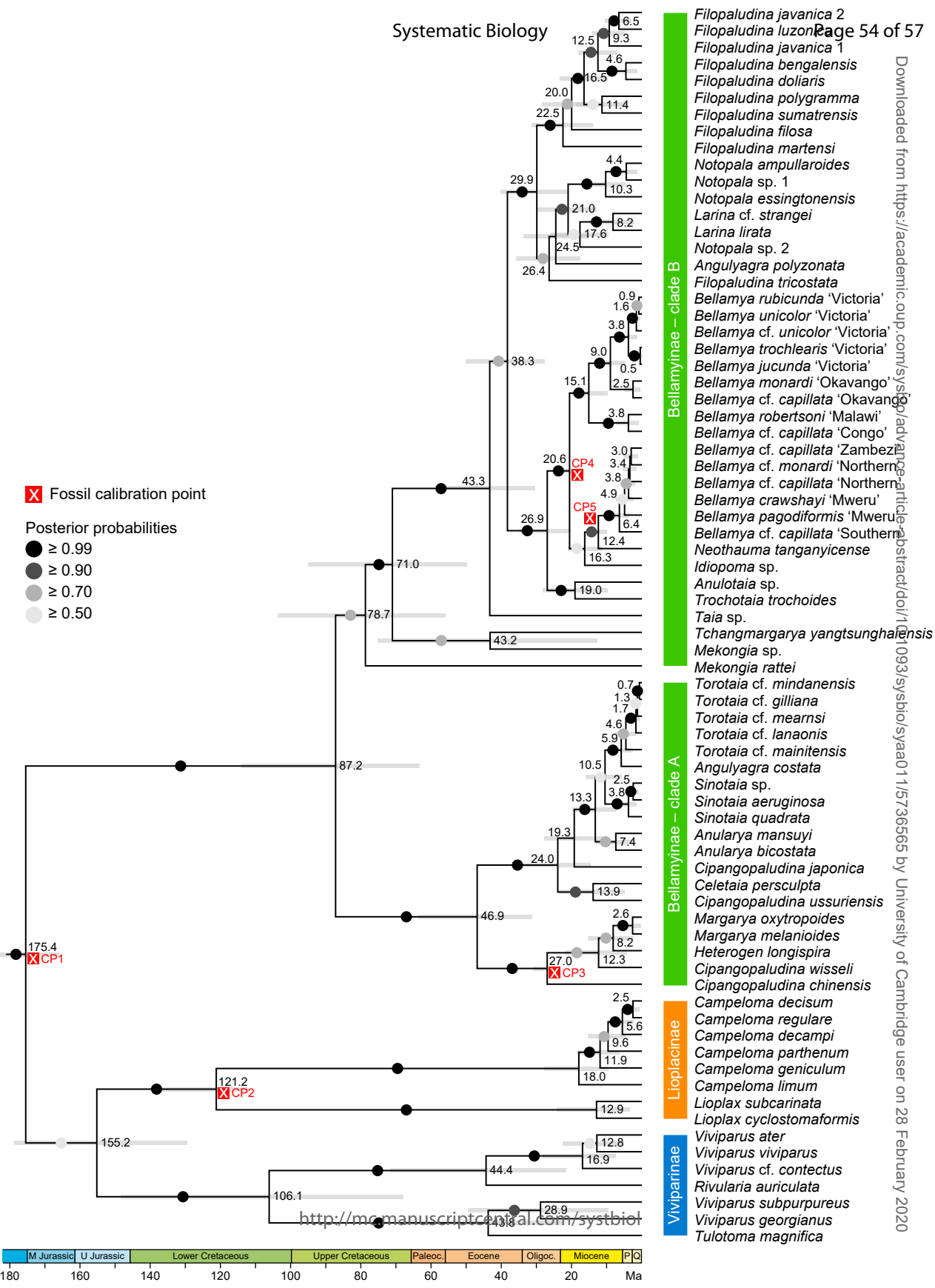
<http://mc.manuscriptcentral.com/systbiol>

Downloaded from <https://academic.oup.com/systbio/advance-article-abstract/doi/10.1093/systbio/syaa011/5736655> by University of Cambridge user on 28 February 2020

**X** Fossil calibration point

Posterior probabilities

- $\geq 0.99$
- $\geq 0.90$
- $\geq 0.70$
- $\geq 0.50$



http://mc.manuscriptcentral.com/sysbio

- North America
- Africa
- Europe
- East Asia (China, Far East Russia, Japan, and South Korea)
- India
- Indochina (incl. the Thai-Malay Peninsula and Singapore)
- Sumatra
- Philippines
- Palawan
- Borneo
- Java (incl. Bali)
- Sulawesi
- East of Wallacea (Australia and New Guinea)

X Fossil-constrained node

Cladogenetic events

- sympatry (y)
- subset (s)
- vicariance (v)
- jump dispersal (j)

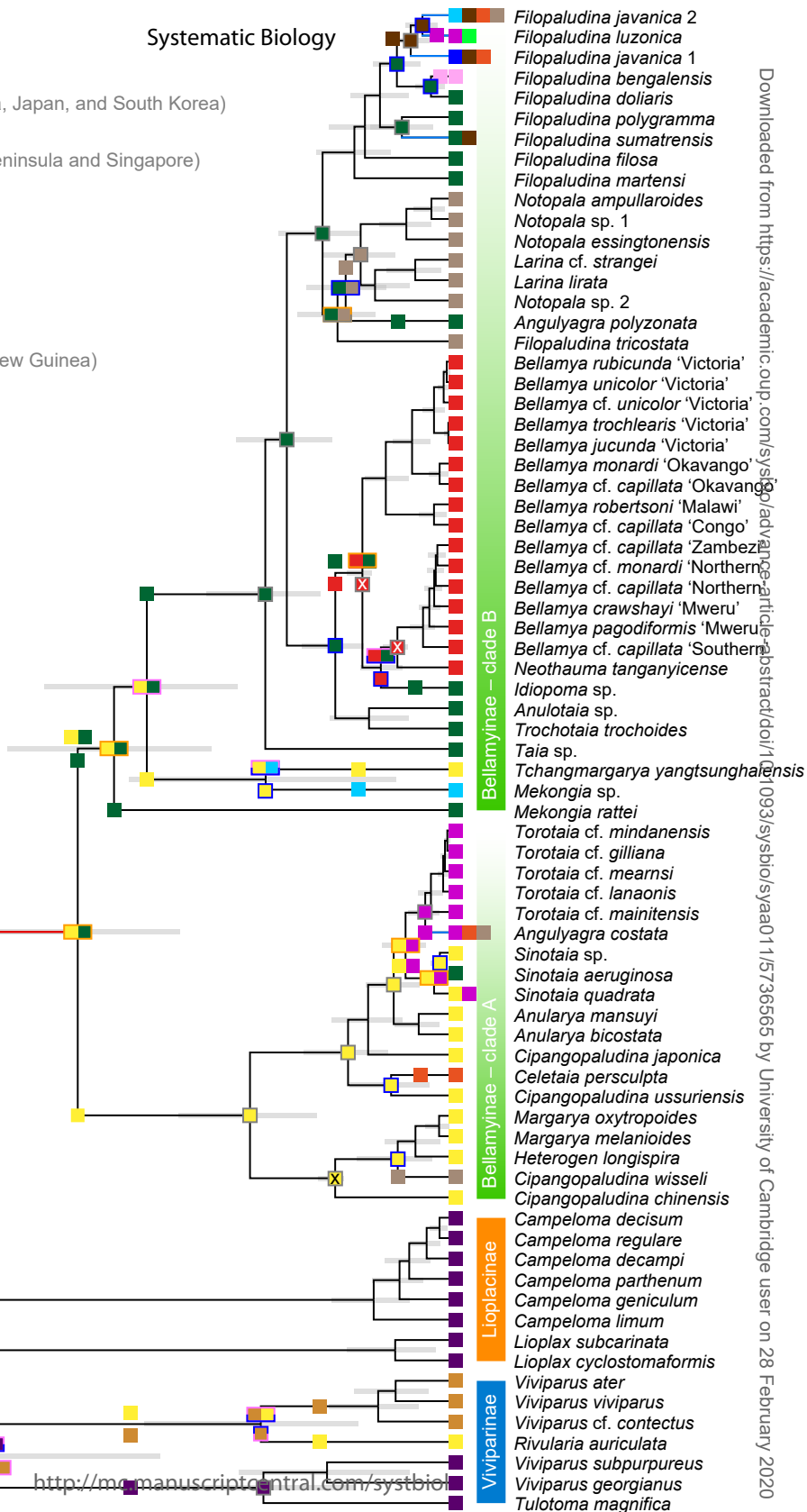
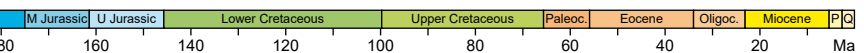
Anagenetic events

- no event
- dispersal (both)
- dispersal (fossil-constrained)

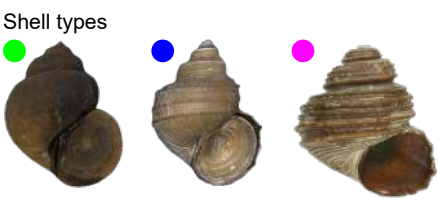
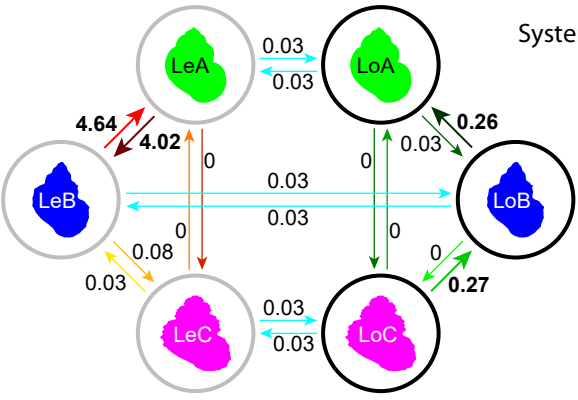
- unconstrained
- fossil-constrained

- unconstrained
- fossil-constrained

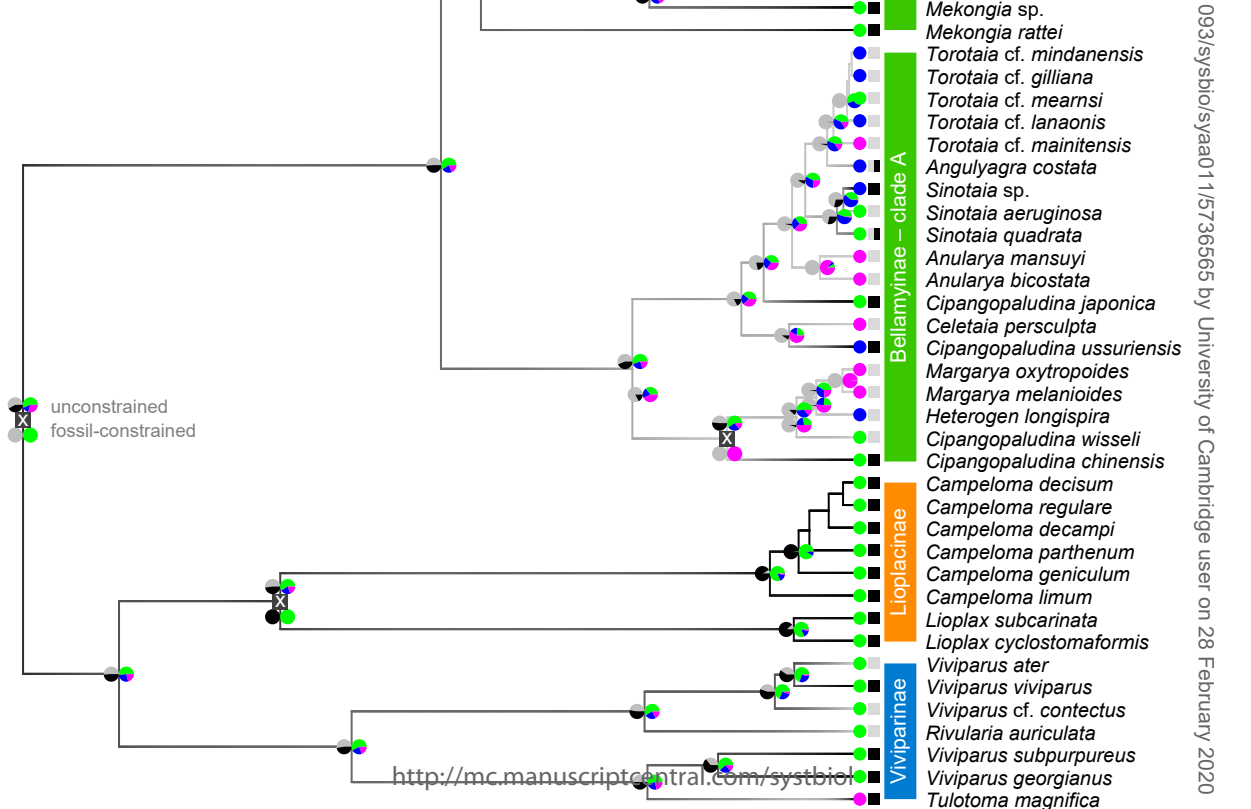
- unconstrained
- fossil-constrained



Downloaded from https://academic.oup.com/sysbio/advance-article-abstract/doi/10.1093/sysbio/syaa011/5736565 by University of Cambridge user on 29 February 2020



- Habitat types
- lentic (wetland, lakes, etc.)
  - lotic (creeks, rivers, etc.)
  - ⊗ Fossil-constrained node
  - Habitat transitions/shifts



<http://mc.manuscriptcentral.com/systbio>

