

# Into-India or out-of-India? Historical biogeography of the freshwater gastropod genus *Pila* (Caenogastropoda: Ampullariidae)

MAITREYA SIL<sup>\*1,2,☉</sup>, N. A. ARAVIND<sup>2,3</sup> and K. PRAVEEN KARANTH<sup>1</sup>

<sup>1</sup>Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560012, India

<sup>2</sup>Ashoka Trust for Research in Ecology and the Environment, Bangalore 560064, India

<sup>3</sup>Yenepoya Research Centre, Yenepoya (Deemed to be University), University road, Derlakatte, Mangalore 575018, India

Received 22 May 2019; revised 11 October 2019; accepted for publication 15 October 2019

The biota of the Indian subcontinent was assembled through multiple associations with various landmasses during a period spanning the Late Cretaceous to the present. It consists of Gondwanan elements that subsequently dispersed ‘out-of-India’ and biota that dispersed ‘into-India’ after the subcontinent collided with Asia. However, the relative contribution of these connections to the current biotic assembly of the subcontinent has been under-explored. Our aim here was to understand the relative importance of these various routes of biotic assembly in India by studying the historical biogeography of the tropical Old World freshwater snail genus *Pila*. We reconstructed a near-complete phylogeny, based on nuclear and mitochondrial markers, of Ampullariidae including all the described *Pila* species from India and Ampullariids worldwide. Thereafter, molecular dating and ancestral range estimation analyses were carried out to ascertain the time frame and route of colonization of India by *Pila*. The results showed that *Pila* dispersed into India as well as other parts of tropical Asia from Africa after both India and Africa collided with Eurasia. Furthermore, multiple dispersals took place between Southeast Asia and India. These findings corroborate increasing evidence that much of the current Indian assemblage of biota actually dispersed ‘into-India’ after it collided with Asia.

ADDITIONAL KEYWORDS: biogeography – biotic assembly – Eurasian route – freshwater gastropod – into-India – out-of-India – molecular dating.

## INTRODUCTION

In any given landmass, the current assembly of flora and fauna represents an outcome of a combination of present and historical processes (Brown & Lomolino, 1998). Geological processes such as continental drift have been major drivers of floral and faunal assemblages. In this regard, the Indian subcontinent is of considerable interest to historical biogeographers because of its plate tectonic history. Around 300 Mya, India was part of the Gondwanaland supercontinent along with Africa, South America, Antarctica, Australia, the Seychelles and Madagascar. Indo-Madagascar–Seychelles separated from the supercontinent about 121 Mya (Sanmartín *et al.*, 2004). Consequently, Indo-Seychelles broke off from Madagascar around 80 Mya

(Chatterjee and Scotese, 1999), and India, in turn, broke off from the Seychelles ~65 Mya (Aitchison *et al.*, 2008). Thereafter, India drifted northwards and came into contact with Asia around 55–35 Mya (Aitchison *et al.*, 2008). The complex geological history of India thus provides an opportunity to understand how plate tectonic processes shape the assembly of biota.

Mani (1974) classified the Indian biota into several categories based on their purported origins: for example, Ethiopian (African) elements in the west, Palaeartic elements in the north, Indo-Chinese or Sundaic (Southeast Asian) elements in the east and endemic peninsular Indian elements in the south. However, the origins of the peninsular Indian forms are much debated. It was speculated and later established that some of these forms [e.g. the endemic frog family Nasikabatrachidae (Biju & Bossuyt, 2003)] were present in peninsular India (PI) (and Sri Lanka) before the Indian plate separated

\*Corresponding author. E-mail: [maitreya.sil@gmail.com](mailto:maitreya.sil@gmail.com)

from the Seychelles around 65 Mya. Some other PI groups are not found in the Seychelles but trace their ancestry back to before the break up of Indo(-Seychelles)–Madagascar; one example is the Hyrophylid cascade beetles (Toussaint *et al.*, 2016). Some Indian lineages of Gondwanan origin dispersed into Southeast Asia (SEA) ‘out-of-India’, such as Ichthyophiid caecilians (Gower *et al.*, 2002), Dipterocharaceae (Dayanandan *et al.*, 1999) and Crypteroniaceae (Conti *et al.*, 2002). Furthermore, some of the lineages found in the wet evergreen forests of the Western Ghats in PI appear to have sister/closely related lineages in north-east India (NEI) and SEA (Mani 1974). Indeed, molecular phylogenetic studies have shown that many PI lineages, not exclusively wet zone taxa, in fact have their origins in SEA (Honda *et al.*, 1999; Noonan & Chippindale, 2006; Köhler & Glaubrecht, 2007; Van Bocxlaer *et al.*, 2009; Adamson *et al.*, 2010; Jansen *et al.*, 2010; Rehan *et al.*, 2010; Datta-Roy *et al.*, 2012; Agarwal & Karanth, 2015; Barley *et al.*, 2015; Klaus *et al.*, 2016; Sil *et al.*, 2019). Several theories propose that during India’s northward course some groups dispersed into the Indian plate from Africa when the two landmasses were close together (Briggs, 2003), possibly via a land bridge, in the form of a chain of islands (the Oman–Kohistan–Dras Island Arc) around 75–65 Mya (Chatterjee & Scotese, 2010). Furthermore, once the African plate sutured with Eurasia around 40 Mya (Van Yperen, 2005), a host of diverse African elements dispersed into India through the ‘Eurasian route’ (Yoder & Yang, 2004). Thus, the Indian biota has diverse origins; some have an ancient Gondwanaland affinity while others have dispersed more recently from Asia or Africa ‘into-India’ (Karanth, 2006; Datta-Roy & Karanth, 2009; Jana & Karanth, 2019). However, the relative importance of these different routes of faunal assembly in India is still not well understood. Further investigation into understudied groups found in PI that have African as well as SEA affinities will provide better insight into the historical patterns of faunal assemblages in PI.

The signatures of continental drift on biogeographical events are particularly identifiable in freshwater biota. A physical link between the freshwater bodies situated in different landmasses requires a terrestrial connection, and thus most of the dispersals undertaken by freshwater organisms would be geodispersal events. Hence, the phylogenies of such organisms often mirror continental breakups and connections. However, freshwater organisms (especially gastropods) are often secondarily dispersed by birds (Woodruff & Mulvey, 1997; Kappes & Hasse, 2012; Von Oheimb *et al.*, 2013; Altaba, 2015). Under such circumstances, dated phylogenies can often be useful in refuting a continental drift origin of their distribution (Leys *et al.*, 2002; Friedman *et al.*, 2013; Opell *et al.*, 2016).

Snails of the genus *Pila* are distributed in tropical and sub-tropical Asia including India, Africa and Madagascar. The genus belongs to the family Ampullariidae, which is found in many Gondwanan landmasses such as India, Africa, Madagascar and South America as well as in non-Gondwanan landmasses such as Central America and SEA. The presence of *Pila* in SEA, Madagascar and Africa makes it a useful model system to ascertain the role of the drifting Indian plate on the assembly of tropical Asian biota. Previously it has been speculated that many tropical Asian species with African affinities are Gondwanan elements that were present in PI before fragmentation of the supercontinent and were carried to Asia by the drifting Indian plate. This biotic ferry hypothesis has been supported by several studies on various taxonomic groups (Mani, 1974; Hedges, 2003; Bossuyt *et al.*, 2006). However, it is not supported by a molecular phylogenetic study undertaken on Ampullariidae (Hayes *et al.*, 2009). Hayes *et al.* (2009) suggested that the split between African and Asian *Pila* was much more recent than the split between the Old and New World Ampullariids, which is contrary to the order of fragmentation of Gondwanaland. Hence, they concluded that *Pila* colonized Asia later, after commencement of the continental collision between Africa and Eurasia. Nevertheless, we believe this result is tentative as the sampling from Asia was sparse: no species from India were included and no molecular dating analysis was carried out. Incomplete sampling often leads to erroneous estimation of phylogenetic relationships. As we have mentioned above, molecular dating can be instrumental in distinguishing between vicariant patterns resulting from continental drift and secondary dispersals. Because much of India’s geological history has been in isolation from that of mainland Asia, it seems likely that Indian species will have a different biogeographical origin than mainland Asian species. Should the Gondwana origin hypothesis be rejected, there are additional scenarios that need to be tested: (1) that *Pila* colonized India from Africa when there was a brief connection between the two drifting landmasses (~75–65 Mya) and (2) that *Pila* took the ‘Eurasian route’ – dispersing from African into tropical Asia after the Africa–Eurasia collision (~40 Mya).

We sampled *Pila* throughout the Indian biogeographical subregion and NEI, which is part of the Indo-Chinese subregion, using over one-third of the described species. Molecular phylogenetic, molecular dating and ancestral range estimation analysis were undertaken to investigate the following questions: (1) Where did the genus *Pila* originate? (2) Did Indian lineages come ‘into India’ from SEA or Africa after India collided with Eurasia or is *Pila* a lineage that existed in India ever since it was a part

of Gondwanaland and then dispersed ‘out of India’? (3) When did *Pila* colonize India and the rest of tropical Asia?

## MATERIALS AND METHODS

### SAMPLE COLLECTION

Whole animal samples were collected from throughout the Indian subregion and NEI, covering all the major river basins. The samples were preserved in absolute ethanol. Five species are described from within the political boundary of India: *Pila globosa*, *Pila virens*, *Pila saxea*, *Pila cf. nevilleiana* and *Pila cf. olea* (Subba Rao, 1989). Of these, the first three are widespread and the rest are data deficient and known from only their type locality (Subba Rao, 1989). Because many widely distributed species may consist of species complexes with many cryptic species (Casu & Kurini-Galletti, 2006; Thabah *et al.*, 2006; Lajmi *et al.*, 2016; Karanth, 2017; Deepak & Karanth, 2017), the three widespread species were collected from multiple locations within their distribution range. We also collected individuals from the type locality of data-deficient species as a proxy for these species. Furthermore, a potentially new species of hill stream-dwelling species of *Pila* was collected from the NEI state of Mizoram, bordering Myanmar. This species is referred to as *Pila* sp. 1 throughout this paper. Supporting Information Table A1 gives a complete list of species with their sampling locations and Figure A1 provides a map of sampling locations. To understand the placement of the Indian *Pila* with respect to *Pila* species from SEA and Africa, published sequences of five *Pila* species from SEA and Africa were included in the phylogenetic reconstruction (Table A2). Thus, of the 29 described species of *Pila*, 11 were included in the dataset. Additionally, three African and two South American Ampullariid genera were included in the dataset; their sequences were obtained from GenBank (Table A2).

### MOLECULAR STUDY

Total genomic DNA was extracted using the CTAB extraction method from the foot muscle tissue of individuals (Williams *et al.*, 2003). Two mitochondrial (*COI* and 16S rRNA) and two nuclear markers (18S rRNA and Histone H3) were amplified. Sequencing was carried out at Medauxin Inc., Bangalore, India. Sequences were aligned thereafter using ClustalW in MEGA 7 (Kumar *et al.*, 2016). Table 1 provides more details on molecular markers used for this study.

### PHYLOGENETIC ANALYSIS

The mitochondrial and nuclear datasets were analysed to check for mito-nuclear discordance and then concatenated and analysed together. The models of sequence evolution and the partitioning scheme was derived from partition analysis using the Byaesian information criterion (BIC) employed in PartitionFinder (Lanfear *et al.*, 2017). Maximum-likelihood (ML) analysis was carried out in RAXML HPC 1.8.2 (Stamatakis, 2014) implemented in raxmlGUI 1.5 (Silvestro & Michalak, 2011). Ten independent runs were initiated, along with 1000 thorough bootstrap replicates. MrBayes 3.2 (Ronquist & Huelsenbeck, 2003) was used to carry out the Bayesian analysis. Two independent runs each consisting of four chains were undertaken for 5 million generations each and sampled every 500 generations. Lowering of the standard deviation of split frequency below 0.01 was taken as a proxy for convergence. The software Tracer 1.5 (Rambaut & Drummond, 2009) was used to check whether all the parameters had reached a stationary phase [expected sample size (ESS) >200]. The final consensus tree was summarized from both runs.

The overall mitochondrial and nuclear phylogenies of the family Ampullariidae were similar, except for some rearrangements within the genus *Pila*. The phylogeny reconstructed from the combined dataset more closely resembled the mitochondrial dataset than the nuclear dataset. A nuclear topology is

**Table 1.** Details of the molecular markers used in the study

Gene name	Abbreviation	Length (bp)	Percentage included	Primers	Reference(s)
Cytochrome <i>c</i> oxidase subunit 1	<i>COI</i>	561	37	LCO and HCO	Folmer <i>et al.</i> (1994), Williams <i>et al.</i> (2003)
16S ribosomal RNA	16S rRNA	500	37	16Sar-L and 16Sbr-H	Palumbi <i>et al.</i> (1991)
18S ribosomal RNA	18S rRNA	468	26	18SYLMFOR and 18SYLMREV	Stothard <i>et al.</i> (2000)
Histone H3	Histone H3	278	75	H3F and H3R	Colgan <i>et al.</i> (1998)

preferred especially when one is investigating deeper divergences because mitochondrial genes evolve faster than nuclear genes and thus reach saturation sooner than the latter (Brown *et al.*, 1979; Lajmi *et al.*, 2018). In particular, third codon positions of the mitochondrial coding genes reach saturation faster and are responsible for the discordance between the mitochondrial and nuclear topology, as observed in both vertebrates (Lajmi *et al.*, 2018) and invertebrates (Blouin *et al.*, 1998). Hence, we excluded the third position from all the codons of the *COI* gene. Furthermore, the *COI* gene was checked for substitution saturation implemented in the program DAMBE6 (Xia, 2017). Mito-nuclear discordance can arise from incomplete lineage sorting and hybridization (Toews & Brelsford, 2012; Franco *et al.*, 2015). However, it is impossible to distinguish between these phenomena using a small dataset.

We also carried out an extensive search to select a suitable outgroup. ML analyses were carried out using different combinations of the outgroup families used in Hayes *et al.* (2009). When the tree was rooted using only Ampullariids, the ingroup was retrieved with the highest support. Furthermore, it also supported the relationship shown in Hayes *et al.* (2009). Hence, the Ampullariid taxa alone were used as outgroups.

#### MOLECULAR DATING

Molecular dating was carried out using BEAST 1.8.3 (Drummond *et al.*, 2012) with the concatenated dataset. The models of sequence evolution and the partitioning scheme were followed as per the results from partition analysis. An uncorrelated lognormal relaxed clock was used for each partition with continuous time markov chain rate reference prior. A Yule tree prior was used as the tree prior. The tree was calibrated using two Old World Ampullariid fossils. First, the most recent common ancestor (MRCA) of the genus *Pila* was calibrated with the oldest fossil of *Pila* unearthed from Oman and dates back to the Priabonian stage (37.8–33.9 Mya) of the Eocene (Harzhauser *et al.*, 2016). Second, the MRCA of the genus *Lanistes* was calibrated with the oldest *Lanistes* fossil described from the Maastrichtian (Late Cretaceous, 72.1–66 Mya) deposits of Oman (Pickford, 2017). The fossil calibration can only inform us about the minimum bound of the calibration with certainty. Hence, we set gamma priors on both calibrations. Different combinations of the shape ( $\alpha$ ) parameter were used to incorporate the uncertainty about the mode of the distribution in four independent runs, whereas the scale ( $\beta$ ) parameters were fixed at 2 for both calibrations (see Supporting Information, Appendix B for more details). Marginal-likelihood estimates (MLEs) of these analyses were calculated following the path sampling/stepping stone sampling method. Thereafter, the

marginal-likelihood scores of each of these runs were compared using the Bayes Factor test. The saturation (ESS > 200) of the analyses was checked in Tracer v.1.5 (Rambaut, 2009). The trees were summarized in TreeAnnotator 1.8.0 (Rambaut & Drummond, 2013) and 25% of the trees were discarded as burn-in. The trees were summarized only from the analysis with the best MLE score based on the Bayes Factor test.

#### TOPOLOGY TEST

A Bayes Factor test was also carried out to establish whether the topologies retrieved from the ML and Bayesian analyses were significantly better than topologies which could potentially support other biogeographical scenarios. Thus, we carried out two additional molecular dating analyses in BEAST 1.8.3 following the best calibration scheme in the molecular dating section. However, in these analyses, the African and the SEA *Pila* species were constrained to be sisters to all other *Pila* taxa respectively (Supporting Information, Figure A2). Such constrained trees might reconstruct a different biogeographical history of the genus *Pila* and an altered time frame of dispersal into India. Furthermore, the constrained topology, in which the African *Pila* are sisters to all other species from that genus, also resembles the nuclear topology obtained from the ML analysis. Similar to the previous section, the analyses were run for 100 million generations with sampling every 1000 generations using the best calibration scheme picked (*Pila* MRCA  $\alpha$  parameter value = 2, *Lanistes* MRCA  $\alpha$  parameter value = 4; *Pila* MRCA  $\beta$  parameter value = 2, *Lanistes* MRCA  $\beta$  parameter value = 2). MLE scores were estimated from the path sampling/stepping stone sampling method using 100 path steps run for 5 00 000 generations. MLE scores were compared pairwise to calculate the Bayes Factor to select the best model.

#### ANALYSIS OF ANCESTRAL RANGE ESTIMATES

Analyses of ancestral range estimates were carried out in R 3.4.2 (available at <https://www.r-project.org/>) using the package BioGeoBEARS (Matzke, 2013). Both DEC (Ree & Smith, 2008) and DEC+*j* (Matzke, 2014) models were used to determine the range evolution and time of colonization of India. The DEC model is particularly effective for revealing the biogeographical patterns of groups that have very different ecologies as well as evolutionary origins. It is also conceptually a better alternative than other similar models such as DIVA (Kodandaramaiah, 2010). The DEC+*j* model is an extension of the principles of the DEC model, although recently it has received criticism (Ree & Sanmartin, 2018). However, this criticism has been targeted mainly at the choice of the best model out of the many implemented



in BioGeoBEARS by comparing their ML scores using Akaike's information criterion (AIC) (Ree & Sanmartin, 2018), and it has often been implemented to explore the range evolution of island taxa and of dispersal-limited taxa (Kitson *et al.*, 2018; Hendriks *et al.*, 2019).

Freshwater snails are known to be dispersal-limited. Moreover, many of the landmasses where the genus *Pila* is found did not have any land connection with each other for much of their geological history. Although geodispersal is more likely for freshwater species, we cannot rule out the possibility of jump-dispersal. Ree & Sanmartin (2018) further suggest that it is permissible to use DEC-based models as long as model selection is based on taxon life-history traits and distribution rather than AIC score. Hence, we decided to implement both the DEC and the DEC+*j* model to uncover the biogeographical history of the taxa of interest and compare the results. A Maximum Clade Credibility (MCC) tree, obtained from the BEAST analysis, was pruned using the package APE (Paradis *et al.*, 2004) in R 3.4.2 to include only Old World taxa. Analysis of ancestral range estimation was based on the MCC tree with the highest MLE score, out of the one unconstrained tree and two constrained trees. The distribution range of the family was broken up into three regions: Africa, SEA including NEI, and India covering all of India south of the Himalayas except NEI (which is part of the Indo-Chinese subregion) (Wallace, 1876). We carried out time-stratified analyses where the adjacency and dispersal probability between different landmasses changed over time based on the plate tectonic movement of the said landmasses. Dispersal from Africa to India and vice versa was assigned a high probability (1.0) at certain points of time and a low probability (0.01) at other times as their adjacency changed over time. Likewise, different values were assigned to the dispersal between SEA and India and their adjacency as the relative position of these landmasses evolved over time. Direct dispersal from SEA to Africa was unlikely for most of their geological history because these landmasses were almost never in close proximity to each other, especially considering the root-age of our focal taxon. Hence dispersal between these two landmasses was given a low probability (0.01) for most of the time periods, following previous studies in which a

similar dispersal probability was assigned between areas where direct dispersal was less likely (Barley *et al.*, 2015) (see Table A3 for details on the dispersal multiplier and adjacency between areas).

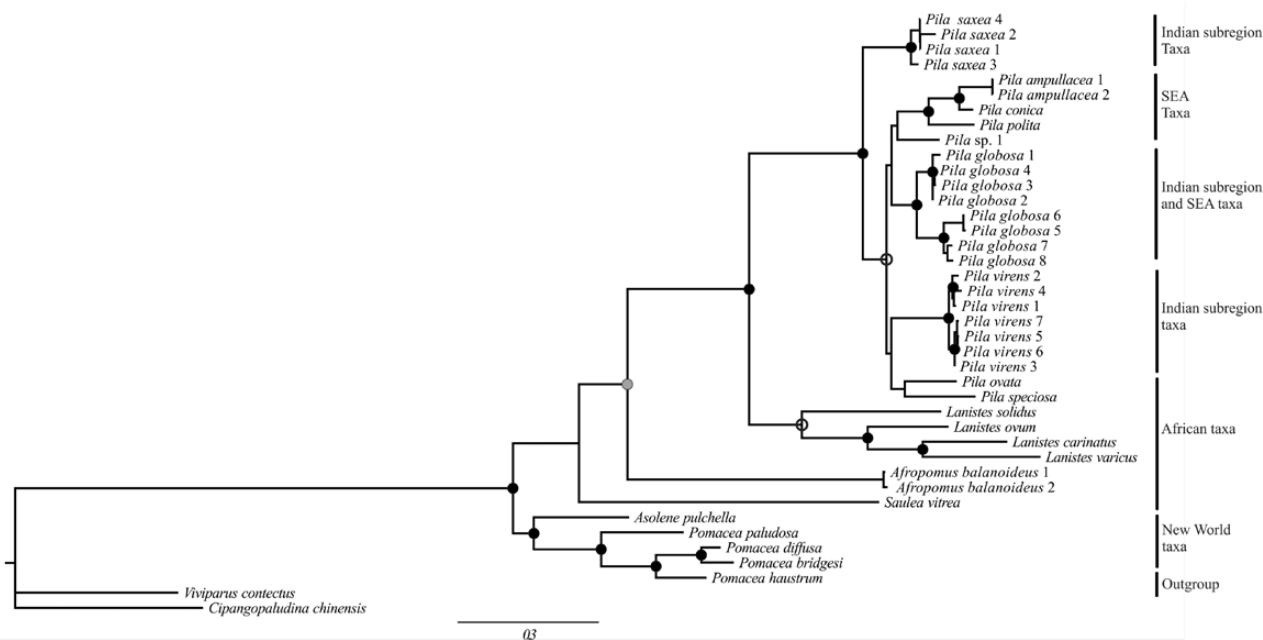
## RESULTS

### PHYLOGENETIC ANALYSIS

The partition scheme and models of sequence evolution estimated for ML, MrBayes and BEAST analyses all varied from each other (see Table 2 for a complete list of partitions). As outlined above, the concatenated tree was more similar to the mitochondrial tree than to the nuclear tree. However, the *COI* gene exhibited little substitution saturation. Moreover, the ML topologies derived from the complete dataset and the dataset after dropping the third codon position of the *COI* gene were identical. Hence, we carried out further analyses based on the topology derived from the complete dataset. The resulting Bayesian tree and ML tree (Fig. 1) exhibited overall congruence. The Old and New World taxa form two clades. The African genus *Saulea* is sister to all other Old World taxa. The genus *Pila* is monophyletic and sister to the African genus *Lanistes*; this clade was in turn sister to the African genus *Afropomus*. Overall, the reconstructed phylogeny is in agreement with that of Hayes *et al.* (2009). Within the *Pila* radiation, the PI species *Pila saxea* is sister to all the remaining *Pila* species. The clade consisting of the remaining *Pila* species constitutes four lineages: African species *P. speciosa* and *P. ovata*; PI species *P. virens* and *P. cf. nevilleiana*; *P. globosa* distributed in India and SEA and *P. cf. olea* from NEI; and *P. conica*, *P. polita* and *P. ampullacea* from SEA and *Pila* sp. 1 from NEI. Among these four the first two were retrieved to be sister lineages, as well as the last two lineages in the ML tree. In the Bayesian tree, however, the relationship between the lineages was unresolved. There were two well-supported clades in the *P. globosa* group: one consisting of species sampled from the Ganges basin and the other composed of individuals collected from Orissa and NEI. The data-deficient species described

**Table 2.** Details of the partition scheme used for each analysis

	RAxML	RAxML (third codon position dropped)	MrBayes	BEAST
<i>COI</i> codon position 3	GTR+I	GTR+I	GTR+G	TrN+G
<i>COI</i> codon position 3	GTR+I	GTR+I	F81+I	HKY+I
<i>COI</i> codon position 3	GTR+I	GTR+I	HKY+G	HKY+G
16S rRNA	GTR+I	GTR+I	GTR+I+G	HKY+I+G
18S rRNA	GTR+I	GTR+I	K80+G	K80+G
Histone H3 codon position 3	GTR+I	GTR+I	GTR	K80+I
Histone H3 codon position 1 and 2	GTR+I	GTR+I	K80+I	K80+I



**Figure 1.** Maximum-likelihood and Bayesian inference phylogeny of Ampullariidae (combined dataset: *COI*, 16S, 18S, Histone H3). Black circles represent nodes with high Bayesian posterior probability (>0.9) and high bootstrap support (>75), grey circles represent only high posterior probability and open circles represent only high bootstrap support. Multiple individuals from the same species are tagged with serial numbers (Supporting Information, Appendix A). Scale bar represents number of substitutions per site. Here the scale bar is 0.3.

from Cachar, Assam, namely *P. cf. olea*, branched with the other individual collected from NEI. This individual is therefore referred to as *P. globosa* 8. The other data-deficient species, *P. cf. nevilliana*, reported from Tranquibar, Tamil Nadu, was likewise nested within *P. virens*. Hereafter, this individual is referred to as *P. virens* 7. Although bootstrap values strongly suggested a sister relationship between *Pila* and *Lanistes* and the monophyly of *Pila*, support for relationships within the genus *Pila* was low.

#### MOLECULAR DATING AND TOPOLOGY TEST

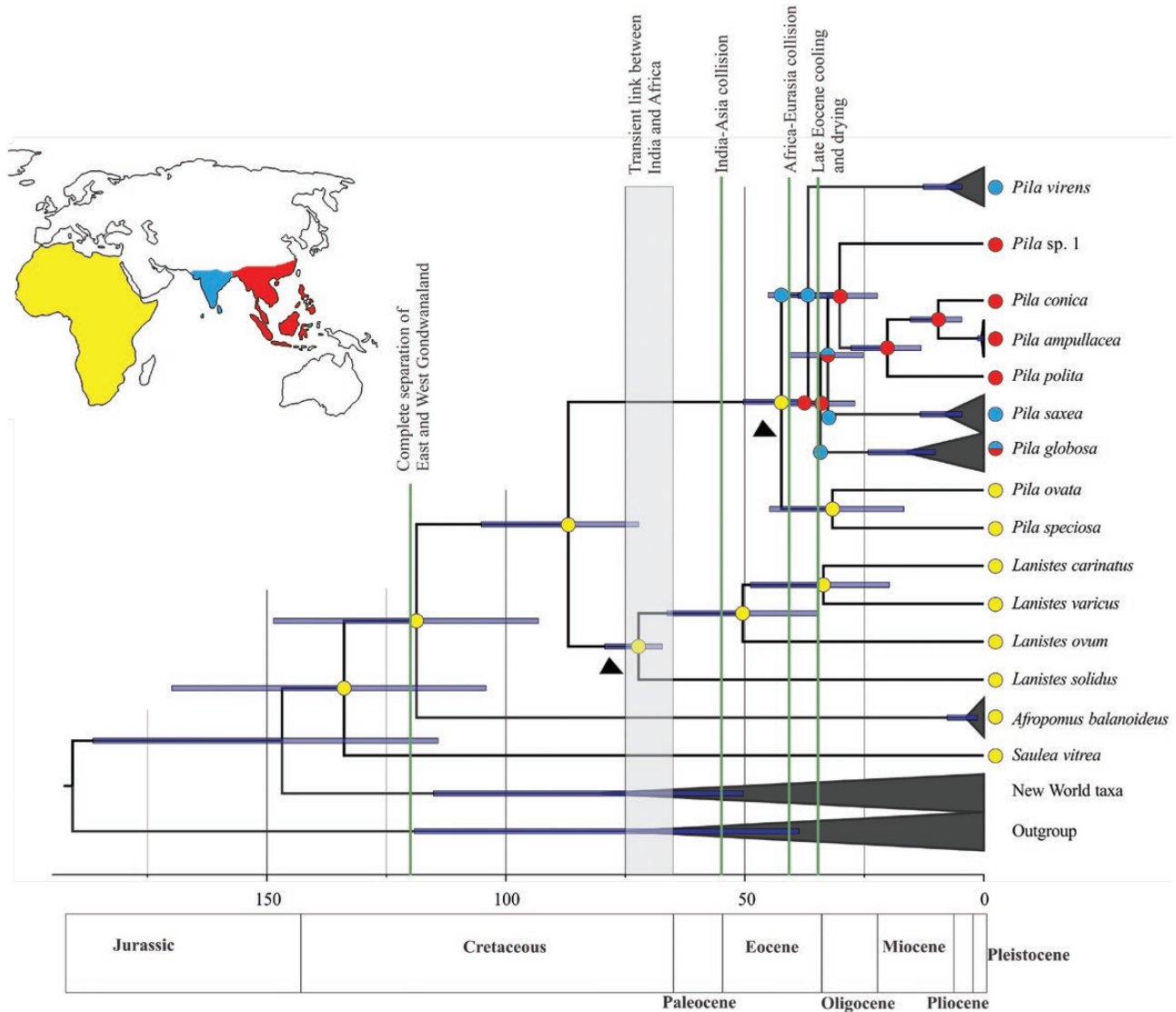
The model in which the alpha parameter of the gamma distribution priors of both the MRCA of *Pila* as well as *Lanistes* was fixed at 4, was retrieved as the best fit model in the Bayes Factor test (Supporting Information, Table A4). Hence, we discuss results from only this particular analysis. The time tree obtained was identical to the ML tree (Figure A3). The split between the Old World and the New World groups dates back to 182.8–112.9 Mya. The split between *Pila* and *Lanistes* was found to have taken place 105.5–73.0 Mya. Radiation started in the genus *Pila* at around 51.3–38.3 Mya. *Pila virens* separated from the African lineage 38.8–20.9 Mya. The SEA species, which were retrieved to be sisters to *P. globosa*, split off from the latter at around the same time (39.7–24.1 Mya).

In both the African and the SEA species, radiation commenced immediately afterward (33.1–10.9 and 37.7–21.0 Mya, respectively). The two clades within the *P. globosa* group diverged around 23.1–9.9 Mya.

The constrained tree in which the African *Pila* species *P. speciosa* and *P. ovata* were constrained to be sister to all other *Pila* species (the tree that resembles the nuclear topology) was found to be a better fit than both the unconstrained topology and the other constrained topology as shown in the second Bayes Factor test (Supporting Information, Table A5). The relationship within the *Pila* clade in the MCC tree, where the African *Pila* were constrained to be sisters to all other *Pila* (Fig. 2), varied from the unconstrained tree topology. Here the radiation started between 50.3 and 36.1 Mya. *Pila virens* branched off from a clade containing *P. globosa*, *P. saxea* and the SEA species around 45.1–29.3 Mya. *Pila globosa*, in turn, separated from the rest 42.4–27.0 Mya. Lastly, *P. saxea* diverged from the SEA species 40.7–25.2 Mya.

#### ANALYSIS OF ANCESTRAL RANGE ESTIMATION

Analysis of ancestral range estimation following both the DEC and the Dec+*j* models revealed overlapping dispersal time frames. The analyses showed that the Old World Ampullariids as well as *Pila* originated in Africa as per expectation, given



**Figure 2.** The constrained Maximum Clade Credibility tree obtained from the BEAST analysis where the African *Pila* were constrained to be sisters to all other species of *Pila*. Coloured circles represent ancestral areas reconstructed at respective nodes and their inheritance as inferred from BioGeoBEARS analysis following the DEC model. Yellow, blue and red circles represent Africa, Indian subregion and SEA respectively, whereas the bicoloured circles indicate presence in any two of all the areas, depending on the constituent colours. Blue bars are the 95% highest posterior density estimates for the age of each node. Black triangles show the placement of the fossil calibrations in the phylogeny. Ages are in Mya.

the higher genus-level diversity in Africa. The Indian subcontinent was colonized from Africa after the break up of Gondwanaland, which was speculated in previous studies (Hayes *et al.*, 2009). Furthermore, dispersals have taken place from SEA into India and back.

Analysis using the constrained MCC tree following the DEC model (Fig. 2) revealed that the lineage consisting of all Asian *Pila* colonized India from Africa during the Eocene (50.3–36.1 Mya). This was followed by dispersal into SEA from India by the lineage composed of all extant Asian *Pila* species except

*P. virens* (45.0–29.3 Mya). Subsequently, there was reverse dispersal back to India by a lineage that gave rise to *P. globosa* (42.6–27.0 Mya), following its range expansion into SEA. Lastly, the MRCA of the clade containing the SEA species and *P. saxea* dispersed into India around 42.4–25.2 Mya.

Results obtained from the DEC+*j* model were almost identical to those of the DEC model analysis, and the range inheritance pattern showed overall congruence except for only one case: only *P. saxea* dispersed back to India from SEA, not the whole clade containing

*P. saxea* and SEA species. The dispersal took place 40.7–25.2 Mya. As these differences did not have any major implications for the questions addressed in the study, we focus our discussion on the results based on the DEC model.

## DISCUSSION

This study aimed to understand the range evolution of the tropical Old World freshwater gastropod genus *Pila*, in particular the origin and biogeography of the Indian species. The results demonstrated that the genus has its origin in Africa, colonizing tropical Asia during the Eocene. Thereafter, multiple dispersals have taken place between India and SEA. Below we discuss how the palaeogeology and palaeoclimate of the three colliding continents, Africa, Eurasia and India, could have governed these dispersal events.

### FROM AFRICA INTO INDIA?

Africa has shared a long biogeographical link with India. Even after the separation of Indo-Madagascar from Africa, there is evidence for a dispersal route between these continents. Geological and palaeontological evidence suggests that India might have been connected to Africa by means of a land bridge during the Late Cretaceous (Chatterjee & Scotese, 2010). However, much of the biota present on the drifting Indian plate at the end of the Cretaceous is thought to have been wiped out, probably by a volcanic explosion in the Deccan region that took place ~65 Mya (Officer *et al.*, 1987; Joshi & Karanth, 2011). Hence much of the biota that has African affinities probably colonized the subcontinent much later, after Africa collided with Eurasia.

Analysis based on the constrained tree following both the DEC and the DEC+*j* models (Fig. 2) showed that only one dispersal event occurred from Africa to India to give rise to all the extant Asian species during the Eocene (50.3–36.1 Mya). Incidentally, this time frame coincides with the collision of the Afro-Arabian plate with the Eurasian plate in the late Eocene (~40 Mya) (Van Yperen, 2005). This suggests that the dispersal might have been through the Eurasian route, the Africa–Eurasia collision (~40 Mya).

The fossil record suggests that this scenario is plausible. The earliest African fossil that resembles *Pila* was dated to the Late Cretaceous (Neubert & van Damme, 2012). However, the assignment of this fossil to *Pila* is dubious (Neubert & van Damme, 2012). The earliest confirmed *Pila* fossil was excavated in Oman and dates to late Eocene times when the Afro-Arabian plate was still conjoined. Thus, it is likely that the *Pila* lineages have followed a Eurasian route as

outlined by Yoder & Yang (2004) and not an earlier Cretaceous dispersal. This is also concurrent with the closure of several major seaways connecting the Tethys with the Atlantic and Arctic oceans which would have made dispersal across Eurasia possible (Meulenkamp & Sissingh, 2003). More complete species and gene sampling needs to be carried out to distinguish between the different competing scenarios. Phylogenetic reconstruction and biogeographical inference will greatly benefit from inclusion of more members of the genus *Pila*.

The results support the Eurasian route suggested for the dispersal of *Lorinae* from Africa (Yoder & Yang 2004), although the possible existence of such a route for dispersal of terrestrial biota at that period of time is contested by Conti *et al.* (2002). The time frame of the dispersal is also concurrent with the completion of the suturing of the Asian and Indian plates (Aitchison *et al.*, 2008).

### ROLE OF PALAEOCLIMATE

*Pila* is primarily a lentic habitat species (Harzhauser *et al.*, 2016). Hence, humid climatic conditions that foster such habitats are likely to have aided in its dispersal. Oman in the late Eocene was characterized by the presence of vast swamplands which harboured Ampullariid snails such as *Pila* (Neubert & van Damme, 2012). The present-day Arabian Peninsula receives much less precipitation and hosts mainly arid habitats (Kotwicki & Al Sulaimani, 2009). The subsequent arid conditions in this region might have been responsible for the absence of extant Ampullariid taxa in the Middle East. A major shift in global climate from warm and humid to cold and arid took place around the Eocene–Oligocene boundary (~34 Mya) (Zachos *et al.*, 2001). The impact of these events on Central Asia and the Arabian peninsula is known from various sedimentological, biogeochemical, magnetostratigraphic and palaeontological studies (Kraatz & Geisler, 2010; Sun & Windley, 2015). Our results from molecular dating and ancestral range estimation analyses suggest that dispersal from Africa into India and SEA took place during the Eocene (50.3–36.1 Mya). The warm and humid climatic conditions of that time were likely to have aided in the dispersal into India from Africa, before the climatic shifts across the Eocene–Oligocene boundary.

### DISPERSALS ACROSS TROPICAL ASIA

Our ancestral range analysis estimates that several bouts of dispersal and back-dispersal have taken place within tropical Asia, i.e. between SEA and India. The first dispersal to SEA from India took place in Eocene–Oligocene times (45.0–29.3 Mya). This was closely



followed by three more dispersal and back-dispersal events. Interestingly, all these events occurred after the completion of the suturing of the Indian and Asian plates or at least overlapped with it. It has been suggested previously that there was an increase in dispersal rate from continental Asia to India after the completion of suturing (Klaus *et al.*, 2016). Hence, our study corroborates the finding that most of the dispersals from continental Asia into India and vice versa occurred after the aforementioned event.

#### RAPID RADIATION

The genus *Pila* was identified (with high support) to be monophyletic, but the relationships within the clade were not well supported. The reason behind the low bootstrap support could be attributed to the mito-nuclear discordance in the *Pila* lineage, caused by hybridization (Francesca *et al.*, 2006; Morii *et al.*, 2015) and/or random sorting of the ancestral alleles. Random sorting of ancestral alleles take place when more than one speciation event takes place within a very short window of time and alleles are distributed randomly between these lineages such that the gene trees do not reflect the true species history (Whitfield & Lockhart, 2007; Rothfels *et al.*, 2012). Rapid radiation has been observed across the breadth of the taxonomic spectrum (Whitfield & Lockhart, 2007; Wiegmann, 2011; Rothfels *et al.*, 2012). In our study the very short internode distance hints at rapid radiation early in the *Pila* lineage (Tsigenopoulos *et al.*, 2003; Han *et al.*, 2004), which might have facilitated random sorting of ancestral alleles, leading to the present case. Rapid radiation can take place in response to increased ecological opportunity such as colonization of a new habitat with available niches (Seehausen, 2006; Sundue *et al.*, 2015, Lajmi & Karanth, 2019; but see Alhajeri *et al.*, 2016).

#### TAXONOMIC IMPLICATIONS

Of the two data-deficient species (*P. cf. nevilliana* and *P. cf. olea*), *P. cf. nevilliana* (*P. virens* 7) is nested within the *P. virens* clade. There are two possibilities for this. First, it is possible that the original collectors wrongly assigned *P. cf. nevilliana* (*P. virens* 7) as a separate species. Some taxonomists indeed have expressed their reservations about whether this is a new species at all or a misidentified individual of *P. saxea* (Annandale, 1921). However, we acknowledge that our own sampling might have been inadequate. *Pila cf. olea*, similarly, is nested within *P. globosa* and it is sister to the other *P. globosa* individual collected from NEI. However, the original description alludes to morphological similarities between *P. cf. olea* and *P. virens* (Prashad, 1924; Subba Rao, 1989). Similar

to the previous case, this also could be an artefact of incomplete sampling or incorrect description. Given that the *P. globosa* individuals from NEI (including *P. cf. olea*) split off from the sister clade from Orissa 12.8–4.2 Mya, it is likely that these individuals warrant a separate species status. A thorough morphological study is needed before a firm conclusion can be reached.

#### CONCLUSION

Our study points to an African origin for the genus *Pila*. According to the molecular dating and ancestral range estimation analyses based on the constrained tree, the ancestors of tropical Asian *Pila* dispersed into India from Africa sometime in the Eocene (50.3–36.1 Mya). The fossil data and the process behind and direction of the dispersal events favour the Eurasian route theory which postulates dispersal of tropical Asian species from Africa after the latter docked with Eurasia during the Eocene. This dispersal event was followed by several dispersals across tropical Asia, from India into SEA and back. The current biota of the Indian subcontinent assembled via several routes at different time points: before India separated from the other Gondwanan fragments, when India was drifting northwards and only had transient contact with other landmasses, and after it collided with Asia. Our study further confirms that very few of the taxa were present in the subcontinent when India was still part of Gondwanaland and we provide further evidence that the biotic assemblages in India were shaped significantly after the India–Asia collision.

#### ACKNOWLEDGMENTS

Ishan Agarwal and Aparna Lajmi commented on the manuscript and assisted during analysis. Kunal Arekar, Chinta Sidharthan, Harshal Bhosale, Pradip Barua, Tarun Singh, Shantanu Kundu, Suhel Sheikh, Animesh Mondal, Harshil Patel and Lansothung Lotha helped during sample collection. The project was funded by the Department of Biotechnology (DBT), Govt. of India-(IISc) Indian Institute of Science, Bangalore partnership programme (22-0303-0007-05-469) and Department of Science and Technology (EMR/2017/001213) to P.K. The fieldwork was partially supported by a Rufford Small Grant for Nature Conservation (19805-1) to M.S. Parts of the fieldwork and molecular work were also funded by grants from the Department of Biotechnology, Government of India (BT/01/17/NE/TAX) and Institute of Bioresource and Sustainable

Development, Imphal, Manipur to N.A.A. We also acknowledge the three anonymous reviewers and the Editor for investing their time in providing useful comments.

## REFERENCES

- Adamson EAS, Hurwood DA, Mather PB. 2010.** A reappraisal of the evolution of Asian snakehead fishes (Pisces, Channidae) using molecular data from multiple genes and fossil calibration. *Molecular Phylogenetics and Evolution* **56**: 707–717.
- Agarwal I, Karanth KP. 2015.** A phylogeny of the only ground-dwelling radiation of *Cyrtodactylus* (Squamata, Gekkonidae): diversification of *Geckoella* across peninsular India and Sri Lanka. *Molecular Phylogenetics and Evolution* **82**: 193–199.
- Aitchison JC, Ali JR, Davis AM. 2008.** When and where did India and Asia collide? *Geological Bulletin of China* **27**: 1351–1370.
- Alhajeri BH, Schenk JJ, Stepan SJ. 2016.** Ecomorphological diversification following continental colonization in murid rodents (Rodentia: Muridae). *Biological Journal of the Linnean Society* **117**: 463–481.
- Altaba CR. 2015.** Once a land of big wild rivers: specialism is context-dependent for riparian snails (Pulmonata: Valloniidae) in central Europe. *Biological Journal of the Linnean Society* **115**: 826–841.
- Annandale N. 1921.** Materials for a generic revision of the freshwater gastropod molluscs of the Indian empire. No. 4. The Indian Ampullariidae. *Records of the Indian Museum* **22**: 7–22.
- Barley AJ, Datta-Roy A, Karanth KP, Brown RM. 2015.** Sun skink diversification across the Indian-Southeast Asian biogeographical interface. *Journal of Biogeography* **42**: 292–304.
- Biju SD, Bossuyt F. 2003.** New frog family from India reveals an ancient biogeographical link with the Seychelles. *Nature* **425**: 711–714.
- Blouin MS, Yowell CA, Courtney CH, Dame JB. 1998.** Substitution bias, rapid saturation, and the use of mtDNA for nematode systematics. *Molecular Biology and Evolution* **15**: 1719–1727.
- Bossuyt F, Brown RM, Hillis DM, Cannatella DC, Mi MC. 2006.** Phylogeny and biogeography of a cosmopolitan frog radiation: Late Cretaceous diversification resulted in continent-scale endemism in the family Ranidae. *Systems Biology* **55**: 579–594.
- Briggs JC. 2003.** Special paper : The biogeographic and tectonic history of India. *Journal of Biogeography* **30**: 381–388.
- Brown WM, George M, Wilson AC. 1979.** Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America* **76**: 1967–71.
- Brown JH, Lomolino MV. 1998.** *Biogeography*. Sunderland: Sinauer Associates.
- Casu M, Curini-Galletti M. 2006.** Genetic evidence for the existence of cryptic species in the mesopsammic flatworm *Pseudomonocelis ophiocephala* (Rhabditophora: Proseriata). *Biological Journal of the Linnean Society* **87**: 553–576.
- Chatterjee S, Scotese CR. 1999.** The breakup of Gondwana and the evolution and biogeography of the Indian plate. *Proceedings-Indian national science academy part A* **65**: 397–426.
- Colgan DJ, McLauchlan A, Wilson GDF, Livingston SP, Edgecombe GD, Macaranas J, Cassis G, Gray MR. 1998.** Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. *Australian Journal of Zoology* **46**: 419.
- Conti E, Eriksson T, Schönenberger J, Sytsma KJ, Baum DA. 2002.** Early Tertiary out-of-india dispersal of Crypteroniaceae: evidence from phylogeny and molecular dating. *Evolution* **56**: 1931–1942.
- Datta-Roy A, Praveen Karanth K. 2009.** The Out-of-India hypothesis: what do molecules suggest? *Journal of Biosciences* **34**: 687–697.
- Datta-Roy A, Singh M, Srinivasulu C, Karanth KP. 2012.** Phylogeny of the Asian *Eutropis* (Squamata: Scincidae) reveals an ‘into India’ endemic Indian radiation. *Molecular Phylogenetics and Evolution* **63**: 817–824.
- Dayanandan S, Ashton PS, Williams SM, Primack RB. 1999.** Phylogeny of the tropical tree family Dipterocarpaceae based on nucleotide sequences of the chloroplast *rbcL* gene. *American Journal of Botany* **86**: 1182–1190.
- Deepak V, Karanth P. 2017.** Aridification driven diversification of fan-throated lizards from the Indian subcontinent. *Molecular Phylogenetics and Evolution* **120**: 53–62.
- Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012.** Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29**: 1969–1973.
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994.** DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**: 294–299.
- Francesca Z, Roberta C, Giuseppe N. 2006.** Genetic relationships of the western Mediterranean painted frogs based on allozymes and mitochondrial markers: evolutionary and taxonomic inferences (Amphibia, Anura, Discoglossidae). *Biological Journal of the Linnean Society* **87**: 515–536.
- Franco FF, Lavagnini TC, Sene FM, Manfrin MH. 2015.** Mito-nuclear discordance with evidence of shared ancestral polymorphism and selection in cactophilic species of *Drosophila*. *Biological Journal of the Linnean Society* **116**: 197–210.
- Friedman M, Keck BP, Dornburg A, Eytan RI, Martin CH, Hulsey CD, Wainwright PC, Near TJ. 2013.** Molecular and fossil evidence place the origin of cichlid fishes long after Gondwanan rifting. *Proceedings of the Royal Society B: Biological Sciences* **280**: 20131733.
- Gower DJ, Kupfer A, Oommen OV, Himstedt W, Nussbaum RA, Loader SP, Presswell B, Müller H, Krishna SB, Boistel R, Wilkinson M. 2002.** A molecular phylogeny of ichthyophiid caecilians (Amphibia: Gymnophiona: Ichthyophiidae): out of India or out of South East Asia? *Proceedings of the Royal Society B: Biological Sciences* **269**: 1563–1569.
- Han D, Zhou K, Bauer AM. 2004.** Phylogenetic relationships among gekkotan lizards inferred from C-mos nuclear DNA

- sequences and a new classification of the Gekkota. *Biological Journal of the Linnean Society* **83**: 353–368.
- Harzhauser M, Neubauer TA, Kadolsky D, Pickford M, Nordsieck H. 2016.** Terrestrial and lacustrine gastropods from the Priabonian (upper Eocene) of the Sultanate of Oman. *Palaontologische Zeitschrift* **90**: 63–99.
- Hayes KA, Cowie RH, Thiengo SC. 2009.** A global phylogeny of apple snails: Gondwanan origin, generic relationships, and the influence of outgroup choice (Caenogastropoda: Ampullariidae). *Biological Journal of the Linnean Society* **98**: 61–76.
- Hedges SB. 2003.** Biogeography: the coelacanth of frogs. *Nature* **425**: 669–670.
- Hendriks KP, Alciantore G, Schilthuizen M, Etienne RS. 2019.** Phylogeography of Bornean land snails suggests long-distance dispersal as a cause of endemism. *Journal of Biogeography* **46**: 932–944.
- Honda M, Ota H, Kobayashi M, Nabhitabhata J, Yong HS, Hikida T. 1999.** Phylogenetic relationships of the flying lizards, genus *Draco* (Reptilia, Agamidae). *Zoological Science* **16**: 535–549.
- Jana A, Karanth P. 2019.** Multilocus nuclear markers provide new insights into the origin and evolution of the blackbuck (*Antelope cervicapra*, Bovidae). *Molecular Phylogenetics and Evolution* **139**: 106560.
- Jansen G, Savolainen R, Vepsäläinen K. 2010.** Phylogeny, divergence-time estimation, biogeography and social parasite–host relationships of the Holarctic ant genus *Myrmica* (Hymenoptera: Formicidae). *Molecular Phylogenetics and Evolution* **56**: 294–304.
- Joshi J, Karanth KP. 2011.** Cretaceous–Tertiary diversification among select Scolopendrid centipedes of South India. *Molecular Phylogenetics and Evolution* **60**: 287–294.
- Kappes H, Haase P. 2012.** Slow, but steady: dispersal of freshwater molluscs. *Aquatic Sciences* **74**: 1–14.
- Karanth KP. 2017.** Species complex, species concepts and characterization of cryptic diversity: vignettes from Indian systems. *Current Science* **112**: 1320–1324.
- Kitson JJN, Warren BH, Thébaud C, Strasberg D, Emerson BC. 2018.** Community assembly and diversification in a species-rich radiation of island weevils (Coleoptera: Cratopini). *Journal of Biogeography* **45**: 2016–2026.
- Klaus S, Morley RJ, Plath M, Zhang YP, Li JT. 2016.** Biotic interchange between the Indian subcontinent and mainland Asia through time. *Nature Communications* **7**: 1–6.
- Kodandaramaiah U. 2010.** Use of dispersal–vicariance analysis in biogeography – a critique. *Journal of Biogeography* **37**: 3–11.
- Köhler F, Glaubrecht M. 2007.** Out of Asia and into India: on the molecular phylogeny and biogeography of the endemic freshwater gastropod *Paracrostoma* Cossmann, 1900 (Caenogastropoda: Pachychilidae). *Biological Journal of the Linnean Society* **91**: 627–651.
- Kotwicki V, Al Sulaimani Z. 2009.** Climates of the Arabian Peninsula – past, present, future. *International Journal of Climate Change Strategies and Management* **1**: 297–310.
- Kraatz BP, Geisler JH. 2010.** Eocene – Oligocene transition in Central Asia and its effects on mammalian evolution. *Geology* **38**: 111–114.
- Kumar S, Stecher G, Tamura K. 2016.** MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**: 1870–1874.
- Lajmi A, Bansal R, Giri V, Karanth P. 2018.** Phylogeny and biogeography of the endemic *Hemidactylus* geckos of the Indian subregion suggest multiple dispersals from Peninsular India to Sri Lanka. *Zoological Journal of the Linnean Society* **186**: 286–301.
- Lajmi A, Giri VB, Karanth KP. 2016.** Molecular data in conjunction with morphology help resolve the *Hemidactylus brookii* complex (Squamata: Gekkonidae). *Organisms Diversity and Evolution* **16**: 659–677.
- Lajmi A, Karanth PK. 2019.** Eocene–Oligocene cooling and the diversification of *Hemidactylus geckos* in Peninsular India. *Molecular Phylogenetics and Evolution* **142**: 106637.
- Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2017.** Partitionfinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution* **34**: 772–773.
- Leys R, Cooper SJ, Schwarz MP. 2002.** Molecular phylogeny and historical biogeography of the large carpenter bees, genus *Xylocopa* (Hymenoptera: Apidae). *Biological Journal of the Linnean Society* **77**: 249–266.
- Mani MS. 1974.** Biogeographical evolution in India. In: Mani MS, ed. *Ecology and biogeography of India*. The Hague: Dr W. Junk B. V. Publishers, 698–724.
- Matzke NJ. 2013.** Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography* **5**: 242–248.
- Matzke NJ. 2014.** Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Systematic Biology* **63**: 951–970.
- Meulenkamp JE, Sissingh W. 2003.** Tertiary palaeogeography and tectonostratigraphic evolution of the Northern and Southern Peri-Tethys platforms and the intermediate domains of the African–Eurasian convergent plate boundary zone. *Palaeogeography, Palaeoclimatology, Palaeoecology* **196**: 209–228.
- Morii Y, Yokoyama J, Kawata M, Davison A, Chiba S. 2015.** Evidence of introgressive hybridization between the morphologically divergent land snails *Ainohelix* and *Ezohelix*. *Biological Journal of the Linnean Society* **115**: 77–95.
- Neubert E, van Damme D. 2012.** Palaeogene continental molluscs of Oman. *Contributions to Natural History* **20**: 1–28.
- Noonan BP, Chippindale PT. 2006.** Dispersal and vicariance: the complex evolutionary history of boid snakes. *Molecular Phylogenetics and Evolution* **40**: 347–358.
- Officer CB, Hallam A, Drake CL, Devine JD. 1987.** Late Cretaceous and paroxysmal Cretaceous–Tertiary extinctions. *Nature* **326**: 143–149.
- Opell BD, Helweg SG, Kiser KM. 2016.** Phylogeography of Australian and New Zealand spray zone spiders



- (Anyphaenidae: Amaurobioides): Moa's Ark loses a few more passengers. *Biological Journal of the Linnean Society* **118**: 959–969.
- Palumbi SR, Martin A, Romano S, Mcmillan WO, Stice L, Grabowski G. 1991.** *Simple fool's guide to PCR*. Honolulu: University of Hawaii Press.
- Paradis E, Claude J, Strimmer K. 2004.** APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289–290.
- Pickford M. 2017.** Late Cretaceous *Lanistes* (Mollusca, Gastropoda) from Al Khodh, Oman. Al Hajar, 23rd edn. *Geological Society of Oman* 15–27.
- Prashad B. 1924.** Revision of the India Ampullariidae. *Memoirs of the Indian Museum* **8**: 69–89.
- Praveen Karanth K. 2006.** Out-of-India Gondwanan origin of some tropical Asian biota. *Current Science* **90**: 789–792.
- Rambaut A, Drummond AJ. 2009.** *Tracer 1.5.0*. MCMC trace analysis tool.
- Rambaut A, Drummond AJ. 2013.** *TreeAnnotator v1.7.0*. Available at: <http://beast.bio.ed.ac.uk>
- Ree RH, Smith SA. 2008.** Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* **57**: 4–14.
- Ree RH, Sanmartín I. 2018.** Conceptual and statistical problems with the DEC+J model of founder-event speciation and its comparison with DEC via model selection. *Journal of Biogeography* **45**: 741–749.
- Rehan SM, Chapman TW, Craigie AI, Richards MH, Cooper SJB, Schwarz MP. 2010.** Molecular phylogeny of the small carpenter bees (Hymenoptera: Apidae: Ceratinini) indicates early and rapid global dispersal. *Molecular Phylogenetics and Evolution* **55**: 1042–1054.
- Ronquist F, Huelsenbeck JP. 2003.** MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572–1574.
- Rothfels CJ, Larsson A, Kuo LY, Korall P, Chiou WL, Pryer KM. 2012.** Overcoming deep roots, fast rates, and short internodes to resolve the ancient rapid radiation of eupolypod II ferns. *Systematic Biology* **61**: 490–509.
- Sanmartín I, Ronquist F, Cunningham C. 2004.** Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Systematic Biology* **53**: 216–243.
- Seehausen O. 2006.** African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society B: Biological Sciences* **273**: 1987–1998.
- Sil M, Aravind NA, Karanth KP. 2019.** Role of geography and climatic oscillations in governing into-India dispersal of freshwater snails of the family: Viviparidae. *Molecular Phylogenetics and Evolution* **138**: 174–181.
- Silvestro D, Michalak I. 2011.** raxmlGUI: a graphic front-end for RAxML. *Organisms Diversity & Evolution* **12**: 335.
- Stamatakis A. 2014.** RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.
- Stothard JR, Brémond P, Andriamaro L, Loxton NJ, Sellin B, Sellin E, Rollins D. 2000.** Molecular characterization of the freshwater snail *Lymnaea natalensis* (Gastropoda: Lymnaeidae) on Madagascar with an observation of an unusual polymorphism in ribosomal small subunit genes. *Journal of Zoology* **252**: 303–315.
- Subba Rao NV. 1989.** *Handbook freshwater molluscs of India*. Kolkata: Zoological Survey of India.
- Sun J, Windley BF. 2015.** Onset of aridification by 34 Ma across the Eocene–Oligocene transition in Central Asia. *Geology* **43**: 1015–1018.
- Sundue MA, Testo WL, Ranker TA. 2015.** Morphological innovation, ecological opportunity, and the radiation of a major vascular epiphyte lineage. *Evolution* **69**: 2482–2495.
- Thabah A, Rossiter SJ, Kingston T, Zhang S, Parsons S, Mya KM, Akbar Z, Jones G. 2006.** Genetic divergence and echolocation call frequency in cryptic species of *Hipposideros larvatus* sl. (Chiroptera: Hipposideridae) from the Indo-Malayan region. *Biological Journal of the Linnean Society* **88**: 119–130.
- Toussaint EF, Fikáček M, Short AE. 2016.** India–Madagascar vicariance explains cascade beetle biogeography. *Biological Journal of the Linnean Society* **118**: 982–991.
- Toews DP, Brelsford A. 2012.** The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology* **21**: 3907–3930.
- Tsigenopoulos CS, Durand JD, Ünlü E, Berrebi P. 2003.** Rapid radiation of the Mediterranean *Luciobarbus* species (Cyprinidae) after the Messinian salinity crisis of the Mediterranean Sea, inferred from mitochondrial phylogenetic analysis. *Biological Journal of the Linnean Society* **80**: 207–222.
- Van Yperen GCN, Meijer PTh, Wortel MJR. 2005.** Rifting in a continental collision setting: the separation of Arabia from Africa. *Geophysical Research Abstracts* **7**: 07396.
- Wallace AR. 1876.** *The geographical distribution of animals, with a study of the relations of living and extinct faunas as elucidating the past chances of the Earth's surface, with maps and illustrations, 2 vols.* New York: Harper & Brothers.
- Whitfield JB, Lockhart PJ. 2007.** Deciphering ancient rapid radiations. *Trends in Ecology and Evolution* **22**: 258–265.
- Wiegmann BMB. 2011.** Episodic radiations in the fly tree of life. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 5690–5695.
- Williams ST, Reid DG, Littlewood DTJ. 2003.** A molecular phylogeny of the Littorininae (Gastropoda: Littorinidae): unequal evolutionary rates, morphological parallelism, and biogeography of the Southern Ocean. *Molecular Phylogenetics and Evolution* **28**: 60–86.
- Woodruff DS, Mulvey M. 1997.** Neotropical schistosomiasis: African affinities of the host snail *Biomphalaria glabrata* (Gastropoda: Planorbidae). *Biological Journal of the Linnean Society* **60**: 505–516.
- Von Oheimb PV, Albrecht C, Riedel F, Bössneck U, Zhang H, Wilke T. 2013.** Testing the role of the Himalaya Mountains as a dispersal barrier in freshwater gastropods (*Gyraulus* spp.). *Biological Journal of the Linnean Society* **109**: 526–534.



- Van Bocxlaer I, Biju S, Loader SP, Bossuyt F. 2009.** Toad radiation reveals into-India dispersal as a source of endemism in the Western Ghats–Sri Lanka biodiversity hotspot. *BMC Evolutionary Biology* **9**: 1–10.
- Xia X. 2017.** DAMBE6: new tools for microbial genomics, phylogenetics, and molecular evolution. *Journal of Heredity* **108**: 431–437.
- Yoder AD, Yang Z. 2004.** Divergence dates for Malagasy lemurs estimated from multiple gene loci: geological and evolutionary context. *Molecular Ecology* **13**: 757–773.
- Zachos J, Pagani M, Sloan L, Thomas E, Billups K. 2001.** Trends, global rhythms, aberrations in global climate 65Ma to present. *Science* **292**: 686–693.

### SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's website.

**Table A1.** List of samples collected during the course of the study with their sampling location.

**Table A2.** List of sequences downloaded from GenBank with accession numbers.

**Table A3.** Adjacency and dispersal probability across different areas used in analyses of ancestral range estimation.

**Table A4.** Result of the Bayes Factor test:  $\ln(\text{BF})$  is derived by subtracting marginal likelihood values of different Bayesian analyses based on different  $\alpha$  parameter values of the gamma distribution. Values above 1.1 represent positive support.

**Table A5.** Result of the Bayes Factor test:  $\ln(\text{BF})$  is derived by subtracting marginal likelihood values of an unconstrained analysis and two constrained Bayesian analyses. Values above 1.1 represent positive support.

**Figure A1.** Map showing the location of individuals collected and amplified during the course of the study.

**Figure A2.** The different scenarios that were tested using topology test (Bayes Factor test). The top left tree is the ML topology obtained from RAxML analysis and the tree summed over the posterior distribution of trees from the Bayesian (BEAST) analysis. In the bottom left scenario, African taxa *Pila ovata* and *P. speciosa* were constrained to be sister to the remaining members of the *Pila* clade. In the scenario on the right, Southeast Asian taxa *Pila ampullacea*, *P. conica*, *P. polita* and *P. a aperta* were constrained to be sisters to the remaining members of the *Pila* clade.

**Figure A3.** MCC tree obtained from the unconstrained BEAST analysis. Blue bars are the 95% highest posterior density estimates for the age of each node.