Main Manuscript for

Idiosyncrasies unveiled: examining the pace, patterns and predictors of biotic diversification in peninsular India

Authors: Pragyadeep Roy[1,2,*] and Jahnavi Joshi[1,2,*]

1. Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, India
2. CSIR-Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad, India

*Corresponding authors: pragyadeeproy@gmail.com; jahnavi@csirccmb.org

Author Contributions: JJ and PR conceptualised the study; PR collected the data and conducted the analyses with inputs from JJ; PR and JJ wrote the manuscript.

Competing Interest Statement: We declare we have no competing interests.

Classification: Biological Sciences; Evolution

Keywords: diversification; biogeography; paleoclimate; animals and plants; species richness

This PDF file includes:

Main Text

Figures 1 to 5

Tables 1 to 2

Abstract

Understanding the tempo and mode of diversification and its consequence for tropical biodiversity remains a key challenge as different clades and regions exemplify distinct patterns and processes. The evolution of the Peninsular Indian biota, one of the oldest regions of differentiation in tropical Asia with a complex geological history and high species diversity with endemism, remains poorly explored. Using birth-death models and their variations, we examined diversification rates and their drivers across 34 well-studied endemic lineages in peninsular India. We show that 19 lineages supported gradual species accumulation, suggesting

that the historical stability of the landscape was an important driver. Additionally, paleotemperature, Miocene aridification, monsoon intensification and existing species diversity explained time-varying diversification patterns among the other 15 lineages. The species richness across clades was driven by net-diversification rates than clade ages, where speciation mainly drove diversification, as extinction rates remained low across taxa. Our results highlight the importance of regional biogeographic, phylogenetic, and geoclimatic history on diversification dynamics. This is critical to understanding the evolution of high tropical diversity and predicting how it may respond to global change.

Significance Statement

Peninsular India, an ancient Gondwanan landmass in tropical Asia, harbours highly diverse and endemic biota, although much of which is threatened by global change. Our understanding of how the peninsular Indian biota has evolved over a long evolutionary time using molecular phylogenies in a quantitative framework is limited. We perform multi-taxa quantitative analyses on endemic peninsular Indian biota, including 633 endemic species of plants and animals, to document the patterns and drivers of diversification. We show the prevalence of idiosyncratic diversification patterns explained by stability and multiple geoclimatic processes. Being the first comprehensive macroevolutionary synthesis for Oriental taxa, this study contributes to the understanding of diversification in the region and gives insights into the evolutionary assembly of tropical diversity.

Main Text

Introduction

Diversification, a balance between speciation and extinction, is known to shape biodiversity patterns. The tropical biome, the most species-rich terrestrial biome, has been central to diversification studies (1–4). The high tropical diversity has been explained through multiple hypotheses, including a long time for speciation - being the oldest biome (5), greater climatic stability (4), habitat, and climatic heterogeneity (6) have been invoked to explain the high diversity in the tropics. For example, a recent study from Neotropics showed that gradual expansion was the most supported scenario, irrespective of biogeographic subregions and that diversification rates differed substantially across clades (7). On the other hand, in Australia, aridification and biogeographic barriers were important drivers of diversification (8). In Madagascar, a continental island, wet-forest-adapted species showed accelerated diversification (9). These studies across the tropical regions suggest that patterns and drivers of diversification depend on regional biogeographic and geoclimatic processes and vary across the Tree of Life. Therefore, there is a need to identify and characterise the processes responsible for generating differential species diversity among distinct geographical regions and clades for a nuanced understanding of the processes contributing to species diversity.

In this regard, Peninsular India (PI) of tropical Asia remains relatively unexplored despite being one of the oldest regions for diversification in Tropical Asia (10, 11). The geological and climatic history of PI has played an important role in generating this complex landscape (12–14) and has influenced the diversification of flora and fauna in PI and also in Tropical Asia (10, 15–17) (Table 1). PI harbours the Western Ghats biodiversity hotspot (WG), which is an escarpment along the west coast, along with the 66 million-year-old Deccan trap or peninsula, and the Eastern Ghats (EG), which is a broken chain of mountains along the east coast (18, 19). Historical biogeographic studies indicate that the peninsular Indian biota has distinct biogeographic and evolutionary affinities with Gondwana, Asia or Eurasia (11, 20, 21). Interestingly, irrespective of biogeographic origins, many clades have undergone extensive in-situ speciation in peninsular India, making peninsular India a distinct region within Tropical Asia (11). However, our understanding of diversification tempo, mode and their driver remains limited for peninsular Indian flora and fauna. Like most tropical areas, PI diversity and habitats are threatened by

global change. Therefore, it is imperative to know the diversification dynamics to assess the evolutionary past and predict the future of peninsular India's highly diverse endemic biota.

Given this, we examine the relative roles of geoclimatic processes, biogeographic affinities, and species traits likely to have shaped the extant diversity in peninsular India in a comparative framework. We use state-of-the-art methods to examine the tempo and mode of *in-situ* diversification for PI endemic plants and animal lineages (22, 23). We compiled a dataset for 34 lineages (clades) belonging to diverse taxonomic groups - arthropods, amphibians, birds, plants, and reptiles where data is available on species distribution, molecular phylogenetic relationships, and divergence time estimates (11, 24). Many of these lineages are endemic to peninsular India and have undergone extensive in-situ diversification, making the landscape ideal for examining biotic diversification patterns and drivers in a comparative framework. These lineages also have distinct biogeographic origins, such as Gondwanan or Asian, occupying distinct habitats within peninsular India, wet or dry forests. We build on these systematic and phylogenetic studies to infer diversification rates, patterns, and potential drivers for peninsular Indian biota. Importantly, we synthesise the hypotheses and outline testable predictions for the diversification dynamics of the endemic peninsular Indian biota.

1. How do diversification rates vary through time and across lineages? - We examined four scenarios to determine the pattern of diversification within each lineage (7, 25) relevant to peninsular India. The gradual accumulation (SC1) scenario predicts species accumulate gradually over time and have constant speciation and extinction rates, leading to constant diversification rates over time (7, 26). Prolonged environmental stability can potentially explain the gradual accumulation of lineages (27–29) (SI Fig. 1 SC1). Within PI, southern parts of the WG and the mountaintops of the EG have remained stable through evolutionary time and could have served as refugia (12, 30–32). Alternatively, lineages can show time-varying diversification with an early increase in diversification followed by saturation in diversity (26, 33, 34). This mode of diversification is attributed to ecological limits/diversity-dependence (35–37) or drastic geoclimatic perturbations that decelerate species accumulation or due to damped increases (37) (SI Fig. 1. SC2). Additionally, species diversity can decline following bouts of increase, ultimately leading to a loss in diversity. Such waxing and waning

diversification patterns (SI Fig. 1. SC3) can be attributed to the inability of lineages to cope with a changing environment and associated decrease in available niches (26, 33). It is generally characterised by negative net-diversification rates (i.e. extinction rates exceeding speciation rates) in a certain period(s) across a lineage (7, 26, 33). PI has undergone drastic climate changes through the Cenozoic, which might be relevant for many lineages. Lastly, lineages can exhibit an exponential increase of diversity towards the present (SI Fig. 1. SC4) and are primarily attributed to geoclimatic perturbations and suitable environmental conditions promoting adaptive radiations. The most important geoclimatic events, like the Mid-Miocene aridification, monsoon intensification and expansion of C4-plants in peninsular India, could explain this pattern (15, 38).

Furthermore, the lineages may also experience episodic shifts in diversification rates, which these four scenarios may not be able to capture (39–41). Diversification rate shifts can also be attributed to geological episodes where certain geoclimatic events can cause rapid increase and/or decrease in diversification rates (39–41). Species diversity within a lineage is expected to show fluctuations from general trends, as seen for SC1-4 at episodes relevant to rate shifts. Hence, we also examined episodic rate shifts in diversification scenarios. We investigated the diversification scenarios using birth-death models and their variations to explicitly explore diversification tempo, which is currently lacking for PI biota.

2. What are the abiotic and biotic drivers of diversification rates and patterns for PI biota? - Among abiotic factors, we assessed the role of environmental variables, given the complex geological past of the peninsular Indian landmass over the last 100 million years (13, 14). Additionally, across different regions and taxa, environmental variables have been shown to influence the diversification rates (7, 23, 26). We specifically examined the temperature in the Cenozoic (67 Mya to present) (42), pedogenic C-content (43, 44), a proxy for vegetation cover/habitat, and the Himalayan orogeny (45) as a proxy for seasonality on the speciation and extinction rates (Fig. 3b), which are thought to be important events for peninsular Indian biota. For the biotic drivers, we assessed the diversity-dependent or ecological limit hypothesis, which suggests that the diversity within a lineage limits the species' diversity. These different drivers were

assessed by fitting birth-death models described in SI Fig. 1 and SI Table 2 in the likelihood framework. The biogeographic and habitat affinity could also influence a clade's diversification pattern and rates (7, 9, 46). Hence, we also assessed the role of biogeographic origin (Asian or Gondwanan) and habitat affinity (wet and dry) on diversification rates in each lineage, given the complex geological past and landscape in PI.

3. What are the relative roles of clade- age and diversification rates on species richness in PI? - Historical biogeographic studies in peninsular India highlighted that taxa with Gondwanan origins are generally of older clade age than the Asian origins (11, 47). This has implications for the clade age hypothesis, which predicts that there could be a disparity in species richness among lineages attributed to clade age, i.e., clades with older ages would be more species-rich (5, 48, 49). Alternatively, differences in species diversity may also result from differences in net-diversification rates between clades, irrespective of clade age, explaining high species richness(5, 50, 51). The rate disparity could also arise due to key innovations, habitat shifts, and diversity dependence. These hypotheses have received support from multiple global, regional and taxon-specific studies (5, 50, 52).

Results

2.1. Endemic Peninsular Indian biota: Diversity, Age and Biogeographic Affinities - The dataset consists of well-sampled dated phylogenetic trees of 34 lineages encompassing 669 species, of which 94% (633 species) are endemic to peninsular India (Fig. 1). These 34 lineages comprised five predatory soil arthropods (50 species - centipedes and scorpions), three freshwater molluscs (22 species - mussels), eight amphibians (153 species - frogs and caecilians), 12 reptiles (304 species - turtles, geckos, agamids, skinks, lacertids and snakes), one bird (4 species), and five plants (136 species - angiosperms). These lineages are distributed in both terrestrial (30 lineages) and freshwater habitats (four lineages). Among terrestrial lineages, 18 lineages are restricted to the wet zones of the Western Ghats, Eastern Ghats, and Sri-Lankan Forests, and 12 are distributed across both wet and dry zones in Peninsular India and/or Sri Lanka (Fig. 1). These

lineages also have varied biogeographic origins, where 21 of the lineages are of Asian origin and 13 are of Gondwanan origin (Fig. 1).

2.2. Mode and tempo of species accumulation within each lineage

Speciation, extinction, and net diversification rates were calculated fitting the constant birth-death (CBD) model across all 34 lineages. Among the different taxonomic groups examined in our dataset, arthropods showed significantly low diversification rates, and plants had significantly higher diversification rates than other taxonomic groups (p<0.01, Fig. 2.a.i, c.i). Within different habitat types, lineages distributed in the freshwater habitats (n=4) showed a slower pace of diversification than lineages in terrestrial habitats (n=30) (p<0.05 - Fig. 2.a.ii, c.ii). We also found no difference in diversification rates with respect to habitat type (i.e. terrestrial: wet and wet+dry habitats). The diversification rates were lower among the Gondwanan than in Asian lineages ($p<0.05$, Fig. 2.a.iii, c.iii). Nevertheless, the credible intervals for diversification rates overlapped for habitat types and biogeographic origins. Differences in speciation rates mainly contributed to disparities in net-diversification rates because extinction rates across groups are very low (~0), except for four lineages (*Ophisops* and Fan-throated lizards, *Montecincla* birds and *Parreysia* mussels) out of 34 lineages (Fig. 2.d, e, f).

<u>Diversification patterns based on γ-statistic and diversification models</u> - Of the 34 lineages, the null hypothesis that "$\gamma = 0$" could not be rejected for 29 taxa, suggesting support for the gradual accumulation of lineages depicted in scenario 1 (SC1). The remaining five lineages showed a slowdown in accordance with diversification scenarios 2 and 3 ($\gamma < 0$, *gammatest* and *MCCR test* - $p < 0.05$, SI Table 3). Further, the constant-rate (CB and CBD) and time-varying models were fitted to all lineages. We found that 19 lineages supported the constant-rate birth (CB) model with gradual expansion (SC1), and 15 supported time-varying birth-death models with both slowdowns (SC2) and Waxing and Waning (SC3) (Table 2; SI Table 1, SI Fig. 3). These 19 lineages were spread across taxonomic groups, biogeographic origins and habitat types (SI Fig. 3). Among 15 lineages with time-varying birth-death models, 13 also showed episodic shifts in the diversification rates and two lineages (*Hemidactylus* lizards and *Pseudophilautus* frogs) showed saturated accumulation (SC2) without any episodic rate shifts. Among the 13 lineages with episodic rate shifts, one lineage showed slowdown (SC2), four lineages with waxed and

waned (SC3) and eight lineages showed episodic rate shifts. Episodic rate shifts were inferred by CoMET and RPANDA (symmetrical shift model) in 9 and 8 lineages, respectively, out of which both methods concordantly inferred four. Parameter estimates of Γ-statistic, birth-death model-fitting (RPANDA) and CoMET analysis are provided in SI Table 3.

2.3. Drivers of diversification and species diversity

There was no association between the diversification scenarios and taxonomic groups, habitat types or biogeographic origins (Fig. 4; SI Fig. 3), except in molluscs, where all three lineages supported the gradual accumulation scenario (SC1). The time-varying diversification rates were primarily associated with paleo-temperature and existing diversity through time (Table 2; SI Table 1 and 4; Fig. 4). Of 15 lineages that supported time-varying models and/or shifts over constant rates, 10 showed temperature dependence (TD). Among these 10 TD lineages, three lineages equally supported (ΔAICc ≤ 2) diversity-dependence (DD). Furthermore, two of these three lineages supported Himalayan-orogeny-dependence, and none supported C4-plant-expansion-dependence, respectively.

Diversification rate shifts were inferred for 13 lineages, and 12 of them were found to occur in Neogene, as shown in Fig. 3. Also, nine lineages among the 12 lineages above showed diversification rate shifts within the geological period that experienced a mid-Miocene intensification of aridification, and monsoons and expansion of C4-plants in PI. Four lineages showed rate shifts younger than 3 Mya, and two lineages showed rate shifts in the late Neogene. Sharp declines in global temperatures from around 2 Mya to the present are concurrent with these recent rate shifts in these lineages, all of which are also declines in diversification rates (Fig. 3, SI Fig. 4.b).

In the FAMD analysis, 36% (Fig. 4 - Axis 1: 19.1%, Axis 2: 16.9%) of the total variation was explained by the first two dimensions. Majorly, diversification scenarios, species richness, temperature-dependence and diversity-dependence contributed to dimension 1 and biogeographic origins, taxonomic groups and clade ages contributed to dimension 2 (SI Fig. 5). Additionally, species richness and net diversification rates were associated. However, clade ages did not show any relationship with species richness (SI Fig. 5 and Fig. 4), similar to the PGLS analyses (Fig. 5).

Boxplots and the Wilcoxon test reconfirmed the association between TD, DD, species richness and net diversification as the lineages showing Temperature-dependent and Diversity-dependent diversification had significantly higher species richness and net-diversification rates (SI Fig. 5.c). Additionally, among diversification scenarios, gradual accumulation favoured lower species richness and net-diversification rates (SI Fig. 5.c), reconfirming the association between scenarios, species richness and net-diversification rates, as indicated by the FAMD.

Convex hulls around clusters of diversification scenarios illustrating time-varying diversification were predominantly concordant with temperature- and diversity-dependence. Also, clusters of lineages pertaining to different biogeographic origins showed minimal overlap in the multidimensional space. Additionally, all arthropod (n=5) and mollusc (n=3) lineages were found to form distinct clusters, with no overlap with clusters composed of amphibians, reptiles and plants (Fig. 5.b). *P-values* of PERMANOVA tests, indicating the significance of overlaps between clusters pertaining to each categorical variable, have been provided in SI Table 5.

The relationship between species richness and clade age was not significant across lineages in PGLS regression analyses (Fig. 5.a, $R^2 = 0.1$, $p = 0.141$) as well as for different biogeographic and habitat affinities (SI Fig. 6.b), except for lineages distributed across both wet and dry habitats (SI Fig. 5.b, Wet+Dry habitat - $R^2 = 0.47$, $p = 0.02$). On the other hand, species richness and net-diversification rates showed a significant association (Fig. 5.b, $R^2 = 0.31$, $p = 0.002$) across lineages and biogeographic and habitat affinities, except for lineages from wet and dry habitats (SI Fig. 6.b, Wet+Dry habitat - $R^2 = 0.28$, $p = 0.887$). Also, a statistically significant negative relationship was observed between net diversification rates and clade ages across lineages with a high phylogenetic signal (Fig. 5.c, SI Fig. 6.b).

Discussion

Our multi-clade analyses examining the tempo and mode of diversification using birth-death models suggested that more than 50% of lineages show gradual species accumulations through time with high speciation rates accompanied by low extinction rates, indicating that the historical stability of the landscape played an important role. However, we also showed that time-varying diversification patterns were evident among the other lineages, and paleotemperature, Miocene aridification, ecological limits, and species traits influenced them. In addition, the biogeographic and taxonomic affinities impacted diversification and diversity dynamics. For the first time, we demonstrate that the diversification of peninsular biota follows idiosyncratic patterns and is influenced by multiple factors, including regional biogeographic and phylogenetic history, historical stability, and habitat heterogeneity. We also show that the disparity in diversification rates is the primary driver of differences in species diversity among peninsular-Indian lineages, driven by speciation rates as extinction rates remained low across lineages.

3.1. Disparities in the tempo of diversification in PI biota: The taxonomic identity, a proxy for clade-specific factors, such as range size, dispersal abilities, and ecological niches, is thought to influence net diversification rates (50). Among the 34 lineages studied, angiosperms (flowering plants) had the highest diversification rates. In contrast, soil arthropods had the lowest net diversification rates (Fig. 4). Angiosperms are known to have higher diversification rates, primarily because they typically have larger clade-level range sizes (53), polyploidy (54) and higher dispersal ability (55). This means that angiosperms can occupy a wide range of ecological niches and are more likely to colonise new areas, promoting their diversification. On the other hand, soil arthropods, specifically centipedes and scorpions, show specialisation for ecological niches, being one of the top predators in soil ecosystems and fossorial (burrowing), which could limit their opportunities for speciation. They also have relatively low dispersal abilities (24, 56), which could further restrict their ability to colonise new areas and diversify (57).

The biogeography affinities or evolutionary origins could also influence the diversification rates. Peninsular Indian biota has an ancient Gondwanan or young Asian affinity followed by in-situ speciation (11). This offers us a unique opportunity to explore the role of distinct biogeographic affinities in diversification rates. Ancient Gondwanan lineages (n = 13) diversified at lower rates

than younger Asian lineages. Interestingly, most Gondwanan lineages are wet-zone specialists and show less dispersal and more vicariance-mediated diversification than Asian lineages. Gondwanan lineages also show a high degree of ecological specialisation and retention of their ancestral niches, such as fossoriality by arthropods, blind snakes and caecilians and preference for freshwater habitat by molluscs. Ecological specialisations or niche conservatism may limit diversification rates, leading to lower diversification rates among Gondwanan lineages (57–59). On the other hand, many Asian lineages occupy diverse ecological and climatic niches, contributing to higher diversification rates. This indicates that biogeographic affinities and ecological specialisation influenced the diversification rates in peninsular India.

Habitat preference is another important factor influencing the diversification rates (60, 61). Although wet and dry forests are well characterised by their distinct climatic conditions, wet and dry forest-dwelling taxa remained relatively similar in their diversification rates (10, 18). On the other hand, freshwater lineages (n=4) showed significantly lower diversification rates than terrestrial lineages, and three out of these four lineages are molluscs. Hence, species traits could have influenced the diversification rate more than habitat preference. However, a larger sample size is required to strengthen this claim.

To summarise, unique ecological traits associated with these varying affinities (taxonomic, biogeographic and habitat) have driven the disparities in diversification regimes of biota.

3.2. Climatic stability, paleotemperature, ecological limits and biogeography: More than half of the studied lineages (approximately 56%) supported the constant-rate pure-birth model, indicating a gradual accumulation of diversity over time. This gradual increase in diversity is attributed to the stable climate in regions of diversity expansion (7, 28, 29), especially in tropical biomes like the tropical forests of peninsular India. This is an intriguing pattern given that peninsular India (PI) has experienced significant climatic changes throughout the Cenozoic era, during which diversity has accumulated (13, 14). Therefore, we suspect some regions in peninsular India may have served as refugia, and it would be worth identifying them in the future through spatially explicit diversification analyses. Also, how diversification rates vary across space in peninsular India would be worth exploring. Among the tropical regions, Neotropics, 50–67% of endemic plant and tetrapod lineages were supported by gradual diversity

expansions (7). This suggests that stable climate refugia supporting diversity accumulation is not unique to peninsular India but may have global relevance in tropical regions, which need further exploration.

A large proportion of the current species diversity in peninsular India is of Asian origin, except for a few small vertebrates and soil arthropods (11). The plant fossil record suggests diverse Gondwanan flora in the region until at least the late Paleocene, which now has extant relatives in Southeast Asia (12, 62, 63). Relictual Gondwanan floral lineages are yet to be identified in peninsular India, or all have gone extinct. Despite the potential role of peninsular India as an "evolutionary graveyard" for floral lineages (63), our results indicate a predominant historic stability in the landscape. One potential way to provide this stability is through rapid turnover of the forests where ecologically similar species of Asian origins replace Gondwanan lineages, as has also been suggested in earlier palynological studies specifically for the Indian landmass (64). Globally, rapid turnover following high species-level extinctions in angiosperms during the Cretaceous-Paleogene (K-Pg) boundary has been inferred by both paleobotanical records and diversification analyses using molecular phylogenies (65). This prediction can be tested using extensive South and Southeast Asian plant clades. Also, similar assessments across other tropical biomes and more taxa from peninsular India would be necessary to assess the generality of stability in tropical forests.

On the other hand, lineages with time-varying diversification rates show a complex interplay between climatic events, episodic rate shifts, and species traits. The Neogene period, especially during the mid-Miocene, witnessed intensified aridification and seasonality in the Indian subcontinent, which is believed to have significantly impacted diversification, particularly among lizards (14, 15, 38). While previous studies suggested a connection between climatic events and diversification, a robust statistical framework to test these hypotheses was lacking. Of the 14 lineages with episodic rate shifts, 12 occurred during the Neogene (23–2.6 Mya), and three occurred more recently, younger than 3 Mya. These shifts involve either an overall decline in diversification rates due to ecological limits or symmetrical episodic shifts (SES), which result from alternating periods of increased and decreased diversification. Nine of these lineages showed rate shifts that consistently overlapped with the temporal range of relevant climatic events, suggesting a potential link between climatic changes and diversification in these

lineages. Additionally, we also reported strict declines in diversification rates in four lineages, which appear to be concurrent with drastic declines in global temperatures during the late Neogene and throughout the Quaternary.

Another important diversification pattern observed among multiple taxa globally and in tropical regions is the slowdowns in diversification rates (26, 33). A global study assessing the drivers of slowdowns inferred the role of diversity-dependence (or ecological limits) and declining global paleotemperature (26). As previously described, diversification slowdown and diversity dependence were observed in *Hemidactylus* geckos. Similar results were also found in two more lineages, namely scorpions and bush frogs (Heterometrinae and *Pseudophilautus*). These patterns were also linked to diversity dependence and temperature dependence, suggesting that changes in species diversity and temperature fluctuations play a role in diversification rates. In four lineages, waxing and waning (diversity declines) were detected and influenced by temperature fluctuations, and only two of these lineages showed equal support for diversity dependence. Again, variation in speciation rates contributed to diversity declines in these lineages, except in one where the extinction rate was relatively high (*Ophisops*).

Mid-Miocene intensification of aridification and monsoons have also been hypothesised to drive exponential species accumulation (SI Fig. 1) (15, 38). However, none of the lineages in our dataset supported a recent rise or exponential accumulation. Also, the role of Himalayan orogeny and the expansion of C4 plants in explaining diversification among PI biota remained limited. Only a few lineages depended on these factors, and their influence may have been episodic and inconsistent throughout lineage history. Hence, derivative curves of these factors, highlighting periods of instability, can be utilised to check the dependence on diversification. Other factors, such as paleo-temperature, could have played a more significant role in diversification. Analytical methods to determine the dependence of diversification in each time interval (piecewise dependence) on factors like paleo-climate and existing diversity can be useful in these scenarios.

3.3. Drivers of diversification rates, clade ages and species diversity: Species diversity across taxa was strongly associated with their intrinsic diversification rates (net-diversification rate) rather than the evolutionary time for diversification (clade-age), supporting the expectation of

the diversification-rate hypothesis. The strong relationship between the diversification rates and species richness has been suggested in clade-specific studies (e.g., plant family Annonaceae) (51) as well as across the Tree of Life (50). Interestingly, we also found a significant negative relationship between diversification rates and clade age, where older lineages have lower diversification rates. Older lineages, persisting in a region for long, are expected to have limited ecological niches to diversify further. Further, regional-level processes could also impose ecological limits on speciation with increasing diversity (54). Therefore, in peninsular India, old lineages, mainly of Gondwanan origins, are expected to show diversity-dependent diversification. However, only six out of 34 lineages showed strong evidence for diversity-dependence, of which only three had Gondwanan (total n=13) affinity, indicating that other factors may also be at play in shaping diversification patterns.

Phylogenetic niche conservatism (PNC) could also explain low diversification rates among older lineages. Given most old lineages are terrestrial Gondwanan lineages (except Heterometrinae and *Cnemaspis*) and are distributed solely in the wet zones (i.e. they exhibit PNC), ecological speciation would be limited by high niche saturation (57, 59). In contrast, high diversification rates for young PI lineages, which are of Asian origins, can be explained by higher dispersal rates allowing lineages to occupy diverse ecological and climatic niches (many have Pan-PI and SL distribution) (11, 13, 66).

Phylogenetic relatedness also contributed significantly to the negative relationship between diversification rates and clade ages, as indicated by Pagel's lambda values in the PGLS analyses. In our dataset, young lineages (geckos, frogs and plants) share biogeographic and ecological (climatic niche, range, dispersal ability, trophic levels, etc.) similarities, and so do old lineages (centipedes and scorpions) among themselves. Ecology and evolutionary history of biota are strongly tied in peninsular India, which may pose ecological constraints on diversification for young and old lineages. Hence, disparities in diversification rates are also very pronounced.

It is also likely that the biomes and ranges that comprise the old lineages constitute museums of biodiversity, whereas the ones composed of young lineages constitute cradles of biodiversity. A recent study on endemic woody plants suggests that the Western Ghats is both a museum and a cradle of diversity (67). Given many endemic lineages (e.g. - centipedes, scorpions, geckos)

found in the WG are not restricted to the biodiversity hotspot but have pan-peninsular distributions, the entirety of PI may potentially be a museum and a cradle of biodiversity. The old lineages studied here follow the museum model, with low diversification rates, longer persistence times and constant diversification. In contrast, younger lineages show characteristics of the cradle model of more dynamic and higher diversification rates. This could potentially explain the negative relationship between diversification rates and clade ages (Fig. 5). It would be worth exploring this further to validate this relationship with more lineages with distinct ages, richness patterns and ecologies in peninsular India and tropical regions.

3.4 Understanding diversification in tropical areas: limitations and the way forward

There has been tremendous growth in methods that infer the diversification tempo and mode recently. However, one of the limiting factors to applying these across taxa and regions has been the requirement of large datasets, where the number of tips analysed in each lineage should be high. Many clades studied in our study are species-poor; therefore, we used multiple methods, including summary statistics (Γ-statistics) and birth-deaths models implemented in different frameworks. Also, a default sampling fraction of 1 was kept in the AICc-based birth-death model fitting and the CoMET analyses across all lineages for simplicity. We acknowledge that this might contribute to erroneous assessments in diversification models, especially in slowdowns when taxa are incompletely sampled. Hence, future studies should incorporate more precise sampling fractions for each lineage for robust results. Additionally, comparing these canonical rate estimates with pulled diversification rates (PDR) would help avoid the issue of non-identifiability among many congruent models and provide robust estimates for extinction rates (68). Nevertheless, simulations through our parameter estimates were able to recover the branching patterns and the number of tips of the empirical phylogenies (SI Fig. 7).

Our dataset comprising 633 endemic peninsular Indian species is strongly biased towards small vertebrates (~70 % - 437 species), and birds, freshwater biota and soil arthropods are among the most underrepresented groups. However, it is worth noting that the endemicity among birds (~5%) and mammals (~15%) itself is low in peninsular India (IUCN maps) as compared to the soil arthropods (~ 90%) (56, 69), small vertebrates (~ 52% - amphibians and reptiles) and flowering plants (~ 26%) (70). This could be because bird and mammal assemblages in PI are mainly

dispersal-driven and haven't undergone in-situ speciation like soil arthropods, small vertebrates and angiosperms.

Furthermore, speciation rates drove diversification rates among peninsular Indian taxa, as extinction rates remained low across peninsular Indian lineages. Given the current peninsular India biota largely comprises groups which have evolved post the K/T mass extinction event (ca. 66 Mya), we do not find the signature of extinctions in our dataset (12, 71). The current diversity estimates could be underestimated for soil arthropods, herpetofauna and freshwater biota, as many new species are still being described (24). Also, many studies have explicitly focused on the Western Ghats, the known biodiversity hotspots, and other regions of peninsular India remain unexplored and understudied. Even in global datasets, soil and freshwater habitats are poorly represented, urging us to focus on unexplored taxa and habitats for a holistic understanding of the tempo and mode of diversification across tropical forests. Additionally, phylogenetic reconstructions using more sophisticated tree priors like the fossilised birth-death (FBD) or the cladogenetic diversification rate shifts (ClaDS) model ought to influence the estimates for the diversification dynamics and need to be integrated into future studies.

The role of historical stability in the biotic diversification in tropical forests has been recognised for a long time (7, 72). In addition, species traits and geoclimatic events influencing diversification have also received support across taxa and regions (8, 73). We show that diversification in Peninsular India has a signature of historical ecosystem stability, paleotemperature, Miocene aridification and intensification, and existing diversity, the known global drivers of time-varying diversification. The complex tempo and mode of diversification shown in our results suggest that regional biogeographic, geoclimatic and phylogenetic history needs to be examined across tropical regions, including peninsular India. Furthermore, exploring among-clade rate-heterogeneity within endemic lineages using methods like - BAMM and ClaDS can provide insights into the intricacies of their cladogenesis and possibly shed light on the bioregionalisation of diversification within a biome. Studies like this, where multi-taxa comparisons are carried out from different tropical areas, are critical for gaining insights into the evolution of biodiversity in a tropical biome.

Materials and Methods

4.1. Data compilation - endemic lineages from peninsular India

We sifted through the published molecular phylogenetic studies on peninsular Indian taxa and identified well-sampled dated phylogenetic trees for the PI lineages where at least 65% of species were endemic. We found 42 time-calibrated phylogenetic trees, of which 34 trees had at least four species. We collected data on the biogeographic origin, distribution, crown age and species richness of these 34 peninsular Indian endemic lineages of animals (including arthropods, amphibians, reptiles, and birds) and plants representing 669 species, of which 633 species are endemic to PI. We requested and obtained time-calibrated tree files (maximum clade credibility trees) for 26 lineages from respective authors (see acknowledgements and SI Table 1), and for the remaining lineages (n = 8) *newick* species trees were manually written based on median node ages, from published time trees.

These lineages do not necessarily pertain to one specific taxonomic rank (ranges from species groups to family level). This was done to compare lineages based on their ages and not of specific taxonomic ranks (like genus), as time (clade age) is thought to be one of the major drivers of diversification (5), taxa of different taxonomic ranks can have similar ages and taxonomic ranks are arbitrary (50).

4.2 Assessing the tempo and mode of species accumulation within each of the lineages

<u>The tempo of diversification</u> - The net diversification rates were calculated to assess the relative pace of diversification for each lineage. The rates were calculated by fitting the constant-rate birth-death model (CBD) using the package *RPANDA* (42). The Wilcoxon test was performed using the *stat_compare_means()* function from the *ggplot2* package (74) in R to check if the tempo of diversification for each lineage type is significantly different from the other and the overall mean.

<u>Assessing mode of diversification</u> - The constant-rate and time-varying models were fitted to all 34 lineages. Constant-rate models included constant-rate pure birth (CB - only a constant speciation rate, no extinction) and constant-rate birth and death models (CBD - constant and finite speciation and extinction rates). Time-varying models included combinations of linearly ($\lambda(t) = \lambda_0 + \lambda \cdot t$, $\mu(t) = \mu_o + \mu \cdot t$) or exponentially ($\lambda = \lambda_0 \cdot e^{\lambda \cdot t}$, $\mu = \mu_o \cdot e^{\mu \cdot t}$) varying (with time - t) birth (λ) and death (μ) rates (λ_0 and μ_o are the expected speciation and extinction rates at t=0, λ and μ

are the coefficients of gain or decay in the functions and time, here, is interpreted as time from the present) (26). These combinations also included constant-rate models for scenarios where one rate (birth or death) varies with time, and the other remains constant. In the models above, the absolute values of speciation and extinction rates are considered even when the models infer negative values. The negative values for speciation and extinction rates are not biologically informative but are feasible mathematically. Therefore, we followed the approach described by Morlon et al, (2020) (75) to include a set of models with the same mathematical formulations but instead take a value of 0 in cases where rates become negative. We refer to these models as the "non-negative models" hereafter. These functions capture the broad trends in diversification mode within a lineage.

In addition, to model the possibility of symmetrical local peaks and dips (sharp increase followed by a decrease in diversification rates or *vice versa*) in diversification rate in time, we developed a continuous time-varying mathematical function "witch of agnesi" curve, to model symmetrical local peaks in diversification, diverging from a baseline constant rate: $\lambda(t) = \lambda_0 + \frac{\alpha\omega}{p} \cdot \left(\frac{t}{(t-p)^2 + \omega}\right)$, where λ_0 represents a baseline rate of speciation, α represents the intensity of the peak (or dip), ω represents the spread of the peak or dip, and p represents the position of the peak (or dip) in Ma (76). Similar rate shifts in extinction were also modelled using the same formulation where μ_0, α and ω represented the baseline rate, intensity and spread of the peak or dip in extinction rate, respectively. This was done to check the prevalence of symmetrical local peaks (or dips) in diversification, which might be best fitted by the constant-rate models otherwise, among other models. Additionally, using this function, we intended to qualitatively assess the role of specific geoclimatic events (Table 1) on the diversification of lineages corresponding to the time of occurrence of a peak (or dip). We refer to this model as the "symmetrical local shift"/"symmetrical episodic shift" model henceforth, and it was implemented in the *RPANDA* model-fitting framework. Model fitting using constant-rate and time-varying birth-death models was executed in *RPANDA* using the *fit_bd()* function to determine the diversification pattern of each lineage.

Best-fitting models were selected based on AICc scores to account for over-fitting. ΔAICc scores were calculated as absolute differences in the AICc scores between each model and the model

with the lowest AICc score. The models with ⍰AICc scores lesser than two were chosen for each lineage. If multiple models showed ⍰AICc below two, i.e., all models were equally likely, then the model with the least parameter complexity was chosen as the best-fitting model. If multiple models of similar AICc and the same complexity existed, the model with the least AICc was chosen as the best-fitting model. However, if CB or CBD were found along with other models to be equally likely, the corresponding constant-rate model was considered the best-fitting model for a lineage, as the simplest model could not be rejected. We used the parameter values (⍰$_0$, ⍰, ⍰$_o$ and ⍰) of the selected diversification models to evaluate the scenarios mentioned in SI Table 2 and SI Fig. 1.

In addition, we performed the CoMET (CPP on Mass-Extinction Times) analysis using the R package *TESS* (22, 39) to detect significant rate shifts in a lineage's history and generate rates-through-time (RTT) plots. The CoMET analysis involves binning the history of a clade into 100 time points at equal intervals. For each lineage, we extracted the Bayes Factor (BF) supports for rate shifts at these 100 time points and chose the ones where 2 * ln(BF) scores were substantially higher than constant rates (≥4.6 (40, 77)).

<u>⍰-statistic</u> - In addition to birth-death models, we used the ⍰-statistic to assess the branching patterns within a lineage, as some clades had fewer tips (78). For a phylogenetic tree following a pure-Birth model or a Yule process (no extinction), ⍰ = 0, i.e. each internal node is equidistant from the tips and the root. If ⍰ < 0, the internal nodes are closer to the root (i.e. early rise in diversification). If ⍰ > 0, then the internal nodes are closer to the tips than in comparison to a tree under a pure-birth model (i.e. recent rise in diversification). However, very few lineages would show an exact value of ⍰ = 0 in empirical datasets. Therefore, we used the *gammatest()* function from the *phytools* package in R to calculate the metric for every lineage and to compute the *p*-value to check the probability of a null hypothesis of constant rate birth-death, which can explain the branching pattern within a clade (79).

To summarise, diversification scenarios were inferred from the results of these three analyses - ⍰-statistic, *RPANDA* models, and CoMET.

4.3. Determining drivers of diversification rates and diversity

Assessing drivers of diversification patterns: We assessed the role of important environmental variables on the diversification patterns of biota from PI, mainly temperature, changes in seasonality patterns and Miocene aridification on the diversification patterns. We obtained paleoclimate data for temperature (42), reconstructed Himalayan elevations (45) (as a proxy for the onset of seasonality patterns) and pedogenic carbon content (43, 44) (as a proxy for aridification and expansion of C4 plants - majorly grasses) from the Potwar plateau. The curves for all the paleo-climatic data were smoothened through spline interpolation using the *ss()* function from the *npreg* package (80) in R. Subsequently, we used the same combinations for rate-dependencies (birth and death) as on time (except the local shift model), on temperature, paleo-elevation of the Himalayas and the pedogenic Carbon content and fitted these models using the *fit_env()* function from *RPANDA* package in R. Additionally, we included the functions ($\lambda(T) = \lambda_0 \cdot e^{-\lambda/T}$, $\mu(T) = \mu_0 \cdot e^{-\mu/T}$) inspired by the metabolic theory of biodiversity as provided in Condamine et al., (2019) for temperature-dependence (TD) (26, 81). Among biotic drivers of diversification, we checked for diversity-dependence (DD) on speciation and extinction rates, incorporating both linear and exponential dependencies (26, 82). The code for calculating AICc for all DDD models referred to relevant sections of source codes for *fit_bd()* and *fit_env()* functions from *RPANDA* (42).

Factor analysis of mixed data (FAMD) was performed to assess the correlations among different ecological and evolutionary drivers explaining the patterns of species diversity (or richness) and diversification across lineages (83). Both quantitative and qualitative (categorical) variables were used in this analysis. Quantitative variables included species richness, clade ages and net diversification rates. Categorical variables included taxonomic groups, biogeographic origins, habitat types, presence of habitat structuring within a lineage (i.e. presence of geographically structured sub-clades, indicating potential adaptations within a lineage), diversification scenarios, and whether the lineages show temperature-dependence (TD), diversity-dependence (DD), Himalayan-orogeny-dependence and C4-plant-expansion-dependence on diversification rates.

First, correlations among variables were assessed, and then clustering patterns of lineages based on categories relevant to each qualitative variable were assessed by estimating convex hulls around the centroids in the multidimensional space for each category. We also performed

the permutational multivariate analysis of variance (PERMANOVA), utilising the dissimilarities (Gower's distance) between lineages calculated from FAMD coordinates of each lineage to assess the statistical significance of clustering (84). Lastly, concordance between clusters within each qualitative variable and that of diversification scenarios was checked visually. All the analyses were performed using functions from the R packages - *FactoMineR* (85, 86) and *factoextra* (86).

Clade age vs diversification rate: The relationship between species richness with clade age and diversification rates was examined using Phylogenetic Generalised Least Squares regression (PGLS) (79, 87). It also accounts for the influence of phylogenetic relationships on the association among studied taxa. We performed pairwise PGLS for clade age and species richness, then net-diversification rate and species richness, and lastly, between diversification rate and clade age for all the lineages. The phylogenetic tree used in PGLS was time-calibrated and constructed in www.timetree.org (88) and using respective taxa-specific phylogenies (SI Fig. 6.a). The PGLS was performed in R using the packages - *nlme* (89), *ape* (90) and *rr2* (91), and relevant plots were made using *ggplot2* (74). The strength and statistical significance of regression was assessed using the R^2 and p-value (significance level - 0.05), and the phylogenetic signal for the relationships was assessed by estimating Pagel's λ. Log-transformed (natural log) values were used for species-richness and clade-ages to improve linearity in regression. Ordinary least squares linear regression was performed for cases where PGLS could not reach convergence.

The credibility of parameter estimates: 1000 phylogenetic trees were simulated using each combination of parameter estimates for the best-fitting constant-rate or time-varying birth-death models per lineage. Furthermore, the distributions of the differences in the [illegible]-statistic and number of tips between the simulated and respective empirical trees were plotted as violin plots. Simulations using accurate estimates are expected to show the distribution mean and median within two standard deviations from zero.

Acknowledgments: We thank Dr. Anieli Pereira, Dr. Aparna Lajmi, Dr. Ashok Malik, Dr. Chinta Sidharthan, Dr. Ishan Agarwal, Dr. Ivan Bolotov, Dr. Prabha Amarasinghe, Mr. R. Chaitanya, Dr.

Sandeep Sen, Mr. Saunak Pal, Dr. Siddharth Surveswaran, Dr. Stephanie Loria, Dr. Varun Torsekar, Dr. Vijay Ramesh, Dr. Vivek Cyriac, Mr. Vladislav Gorin, who shared the maximum clade credibility trees, enabling us to compile the dataset. We also thank Mr. Abhishek Gopal and Dr. Mihir Kulkarni for their suggestions on analyses. We are grateful to Dr. Jun Ying Lim, Dr. Rohit Naniwadekar, Mr. Abhishek Gopal, Dr. Bharti D. K. and Mr. R. Chaitanya for their comments on the manuscript.

Funding: This work was supported by the DBT/Wellcome Trust India Alliance Grant [grant number IA/I/20/1/504919] awarded to JJ. PR is supported by the Council for Scientific and Industrial Research (CSIR) Junior Research Fellowship.

Code availability: Custom codes used in the analyses and for plotting graphs would be shared upon acceptance of the manuscript.

References

1. R. E. Ricklefs, Estimating diversification rates from phylogenetic information. *Trends Ecol. Evol.* **22**, 601–610 (2007).
2. M. Cardillo, C. D. L. Orme, I. P. F. Owens, Testing for latitudinal bias in diversification rates: an example using new world birds. *Ecology* **86**, 2278–2287 (2005).
3. W. L. Eiserhardt, T. L. P. Couvreur, W. J. Baker, Plant phylogeny as a window on the evolution of hyperdiversity in the tropical rainforest biome. *New Phytol.* **214**, 1408–1422 (2017).
4. L. M. J. Dagallier, *et al.*, Cradles and museums of generic plant diversity across tropical Africa. *New Phytol.* **225**, 2196–2213 (2020).
5. M. A. McPeek, J. M. Brown, Clade Age and Not Diversification Rate Explains Species Richness among Animal Taxa. *Am. Nat.* **169**, E97–E106 (2007).
6. M. R. Willig, D. M. Kaufman, R. D. Stevens, Latitudinal Gradients of Biodiversity: Pattern, Process, Scale, and Synthesis. *Annu. Rev. Ecol. Evol. Syst.* **34**, 273–309 (2003).
7. A. S. Meseguer, *et al.*, Diversification dynamics in the Neotropics through time, clades, and biogeographic regions. *eLife* **11**, e74503 (2022).
8. M. Pepper, J. S. Keogh, Life in the "dead heart" of Australia: The geohistory of the Australian deserts and its impact on genetic diversity of arid zone lizards. *J. Biogeogr.* **48**, 716–746 (2021).
9. A. Crottini, *et al.*, Vertebrate time-tree elucidates the biogeographic pattern of a major biotic change around the K–T boundary in Madagascar. *Proc. Natl. Acad. Sci.* **109**, 5358–5363 (2012).
10. M. S. Mani, "Biogeographical Evolution in India" in *Ecology and Biogeography in India*, Monographiae Biologicae., M. S. Mani, Ed. (Springer Netherlands, 1974), pp. 698–724.x
11. K. P. Karanth, Dispersal vs. vicariance: the origin of India's extant tetrapod fauna. *Front. Biogeogr.* **13** (2021).
12. V. Prasad, A. Farooqui, S. K. M. Tripathi, R. Garg, B. Thakur, Evidence of Late Palaeocene-Early Eocene equatorial rain forest refugia in southern Western Ghats, India. *J. Biosci.* **34**, 777–797 (2009).
13. S. Chatterjee, A. Goswami, C. R. Scotese, The longest voyage: Tectonic, magmatic, and paleoclimatic evolution of the Indian plate during its northward flight from Gondwana to Asia. *Gondwana Res.* **23**, 238–267 (2013).
14. R. J. Morley, Assembly and division of the South and South-East Asian flora in relation to tectonics and climate change. *J. Trop. Ecol.* **34**, 209–234 (2018).
15. V. Deepak, P. Karanth, Aridification driven diversification of fan-throated lizards from the Indian subcontinent. *Mol. Phylogenet. Evol.* **120**, 53–62 (2018).
16. A. Lajmi, P. K. Karanth, Eocene–Oligocene cooling and the diversification of Hemidactylus geckos in Peninsular India. *Mol. Phylogenet. Evol.* **142**, 106637 (2020).
17. J. Joshi, G. D. Edgecombe, Evolutionary biogeography of the centipede genus Ethmostigmus from Peninsular India: testing an ancient vicariance hypothesis for Old World tropical diversity. *BMC Evol. Biol.* **19**, 1–1 (2019).
18. K. P. Karanth, Evolution of disjunct distributions among wet-zone species of the Indian subcontinent: Testing various hypotheses using a phylogenetic approach. *Curr. Sci.* **85**, 1276–1283 (2003).
19. K. P. Karanth, An island called India: phylogenetic patterns across multiple taxonomic groups reveal endemic radiations. *Curr. Sci.* **108**, 1847–1851 (2015).

20. S. Surveswaran, *et al.*, Origin and diversification of Indian Ceropegieae (Apocynaceae) and its possible relation to the Indian monsoon. *J. Syst. Evol.* **59**, 93–112 (2021).
21. J. Joshi, P. K. Karanth, G. D. Edgecombe, The out-of-India hypothesis: evidence from an ancient centipede genus, Rhysida (Chilopoda: Scolopendromorpha) from the Oriental Region, and systematics of Indian species. *Zool. J. Linn. Soc.* **189**, 828–861 (2020).
22. M. R. May, S. Höhna, B. R. Moore, A Bayesian approach for detecting the impact of mass-extinction events on molecular phylogenies when rates of lineage diversification may vary. *Methods Ecol. Evol.* **7**, 947–959 (2016).
23. F. L. Condamine, J. Rolland, H. Morlon, Macroevolutionary perspectives to environmental change. *Ecol. Lett.* **16**, 72–85 (2013).
24. J. Joshi, I. Agarwal, Integrative Taxonomy in the Indian Subcontinent: Current Progress and Prospects. *J. Indian Inst. Sci.* **101**, 125–149 (2021).
25. H. Morlon, Phylogenetic approaches for studying diversification. *Ecol. Lett.* **17**, 508–525 (2014).
26. F. L. Condamine, J. Rolland, H. Morlon, Assessing the causes of diversification slowdowns: temperature-dependent and diversity-dependent models receive equivalent support. *Ecol. Lett.* **22**, 1900–1912 (2019).
27. R. B. J. Benson, *et al.*, Near-Stasis in the Long-Term Diversification of Mesozoic Tetrapods. *PLOS Biol.* **14**, e1002359 (2016).
28. E. P. Derryberry, *et al.*, Lineage diversification and morphological evolution in a large-scale continental radiation: the neotropical ovenbirds and woodcreepers (Aves: Furnariidae). *Evolution* **65**, 2973–2986 (2011).
29. T. L. P. Couvreur, *et al.*, Early evolutionary history of the flowering plant family Annonaceae: steady diversification and boreotropical geodispersal. *J. Biogeogr.* **38**, 664–680 (2011).
30. D. J. Gower, *et al.*, The role of wet-zone fragmentation in shaping biodiversity patterns in peninsular India: insights from the caecilian amphibian *Gegeneophis*. *J. Biogeogr.* **43**, 1091–1102 (2016).
31. J. Joshi, P. Karanth, Did southern Western Ghats of peninsular India serve as refugia for its endemic biota during the Cretaceous volcanism? *Ecol. Evol.*, 3275–3282 (2013).
32. I. Das, S. K. Chanda, A new species of Philautus (Anura: Rhacophoridae) from the Eastern Ghats, south-eastern India. *J. South Asian Nat. Hist.* **3**, 103–112 (1998).
33. D. Moen, H. Morlon, Why does diversification slow down? *Trends Ecol. Evol.* **29**, 190–197 (2014).
34. R. S. Etienne, J. Rosindell, Prolonging the Past Counteracts the Pull of the Present: Protracted Speciation Can Explain Observed Slowdowns in Diversification. *Syst. Biol.* **61**, 204 (2012).
35. D. L. Rabosky, Ecological Limits on Clade Diversification in Higher Taxa. *Am. Nat.* **173**, 662–674 (2009).
36. J. J. Wiens, The Causes Of Species Richness Patterns Across Space, Time, And Clades And The Role Of "Ecological Limits." *Q. Rev. Biol.* **86**, 75–96 (2011).
37. D. L. Rabosky, Diversity-Dependence, Ecological Speciation, and the Role of Competition in Macroevolution. *Annu. Rev. Ecol. Evol. Syst.* **44**, 481–502 (2013).
38. I. Agarwal, U. Ramakrishnan, A phylogeny of open-habitat lizards (Squamata: Lacertidae: *Ophisops*) supports the antiquity of Indian grassy biomes. *J. Biogeogr.* **44**, 2021–2032 (2017).

39. S. Höhna, M. R. May, B. R. Moore, TESS: an R package for efficiently simulating phylogenetic trees and performing Bayesian inference of lineage diversification rates. *Bioinformatics* **32**, 789–791 (2016).
40. S. Hohna, M. R. May, B. R. Moore, Phylogeny Simulation and Diversification Rate Analysis with TESS (2015).
41. B. T. Kopperud, A. F. Magee, S. Höhna, Rapidly changing speciation and extinction rates can be inferred in spite of nonidentifiability. *Proc. Natl. Acad. Sci.* **120**, e2208851120 (2023).
42. H. Morlon, *et al.*, RPANDA: an R package for macroevolutionary analyses on phylogenetic trees. *Methods Ecol. Evol.* **7**, 589–597 (2016).
43. J. Quade, T. E. Cerling, J. R. Bowman, Development of Asian monsoon revealed by marked ecological shift during the latest Miocene in northern Pakistan. *Nature* **342**, 163–166 (1989).
44. P. D. Clift, A. A. G. Webb, A history of the Asian monsoon and its interactions with solid Earth tectonics in Cenozoic South Asia. *Geol. Soc. Lond. Spec. Publ.* **483**, 631–652 (2019).
45. L. Ding, *et al.*, Quantifying the rise of the Himalaya orogen and implications for the South Asian monsoon. *Geology* **45**, 215–218 (2017).
46. T. N. C. Vasconcelos, *et al.*, Fast diversification through a mosaic of evolutionary histories characterizes the endemic flora of ancient Neotropical mountains. *Proc. R. Soc. B Biol. Sci.* **287**, 20192933 (2020).
47. C. Sidharthan, K. P. Karanth, India's biogeographic history through the eyes of blindsnakes-filling the gaps in the global typhlopoid phylogeny. *Mol. Phylogenet. Evol.* **157**, 107064 (2021).
48. A. Gamisch, H. P. Comes, Clade-age-dependent diversification under high species turnover shapes species richness disparities among tropical rainforest lineages of Bulbophyllum (Orchidaceae). *BMC Evol. Biol.* **19**, 93 (2019).
49. J. Marin, *et al.*, Evolutionary time drives global tetrapod diversity. *Proc. R. Soc. B Biol. Sci.* **285**, 20172378 (2018).
50. J. P. Scholl, J. J. Wiens, Diversification rates and species richness across the Tree of Life. *Proc. R. Soc. B Biol. Sci.* **283**, 20161334 (2016).
51. B. Xue, *et al.*, Accelerated diversification correlated with functional traits shapes extant diversity of the early divergent angiosperm family Annonaceae. *Mol. Phylogenet. Evol.* **142**, 106659 (2020).
52. M. Tietje, *et al.*, Global variation in diversification rate and species richness are unlinked in plants. *Proc. Natl. Acad. Sci.* **119**, e2120662119 (2022).
53. T. Hernández-Hernández, J. J. Wiens, Why Are There So Many Flowering Plants? A Multiscale Analysis of Plant Diversification. *Am. Nat.* **195**, 948–963 (2020).
54. J. S. (Pat) Heslop-Harrison, T. Schwarzacher, Q. Liu, Polyploidy: its consequences and enabling role in plant diversification and evolution. *Ann. Bot.* **131**, 1–10 (2023).
55. D. Suárez, P. Arribas, E. Jiménez-García, B. C. Emerson, Dispersal ability and its consequences for population genetic differentiation and diversification. *Proc. R. Soc. B Biol. Sci.* **289**, 20220489 (2022).
56. S. F. Loria, L. Prendini, OPEN Out of India, thrice: diversification of Asian forest scorpions reveals three colonizations of Southeast Asia. *Sci. Rep.* **10**, 22301 (2020).
57. J. J. Wiens, Speciation and ecology revisited: Phylogenetic niche conservatism and the origin of species. *Evolution* **58**, 193–197 (2004).

58. V. P. Cyriac, U. Kodandaramaiah, Digging their own macroevolutionary grave: fossoriality as an evolutionary dead end in snakes. *J. Evol. Biol.* **31**, 587–598 (2018).
59. D. Da Silva, A. E. Aires, J. P. Zurano, M. A. Olalla-Tárraga, P. A. Martinez, Changing Only Slowly: The Role of Phylogenetic Niche Conservatism in Caviidae (Rodentia) Speciation. *J. Mamm. Evol.* **27**, 713–721 (2020).
60. A. Barlow, *et al.*, Ancient habitat shifts and organismal diversification are decoupled in the African viper genus Bitis (Serpentes: Viperidae). *J. Biogeogr.* **46**, 1234–1248 (2019).
61. A. Villastrigo, P. Abellán, I. Ribera, Habitat preference and diversification rates in a speciose lineage of diving beetles. *Mol. Phylogenet. Evol.* **159**, 107087 (2021).
62. E. A. Wheeler, R. Srivastava, S. R. Manchester, P. Baas, M. Wiemann, Surprisingly modern Latest Cretaceous– earliest Paleocene woods of India. *IAWA J.* **38**, 456–542 (2017).
63. C. D. Bacon, *et al.*, The origin of modern patterns of continental diversity in Mauritiinae palms: the Neotropical museum and the Afrotropical graveyard. *Biol. Lett.* **18**, 20220214 (2022).
64. R. J. Morley, Origin and evolution of tropical rain forests. *Orig. Evol. Trop. Rain For.* (2000) (January 12, 2024).
65. J. B. Thompson, S. Ramírez-Barahona, No phylogenetic evidence for angiosperm mass extinction at the Cretaceous–Palaeogene (K-Pg) boundary. *Biol. Lett.* **19**, 20230314 (2023).
66. J. J. Wiens, Patterns of Local Community Composition Are Linked to Large-Scale Diversification and Dispersal of Clades. *Am. Nat.* **191**, 184–196 (2018).
67. A. Gopal, *et al.*, Range restricted old and young lineages show the southern Western Ghats to be both a museum and a cradle of diversity for woody plants. *Proc. R. Soc. B Biol. Sci.* **290**, 20222513 (2023).
68. S. Louca, M. W. Pennell, Why extinction estimates from extant phylogenies are so often zero. *Curr. Biol.* **31**, 3168-3173.e4 (2021).
69. J. Joshi, K. P. Karanth, Cretaceous–Tertiary diversification among select Scolopendrid centipedes of South India. *Mol. Phylogenet. Evol.* **60**, 287–294 (2011).
70. International Union for Conservation of Nature (IUCN), The IUCN Red List of Threatened Species. **Version 2022-2** (2022).
71. G. Keller, A. Sahni, S. Bajpai, Deccan volcanism, the KT mass extinction and dinosaurs. *J. Biosci.* **34**, 709–728 (2009).
72. G. L. Stebbins, *Flowering plants: evolution above the species level* (Harvard University Press, 1974).
73. Y. Bouchenak-Khelladi, R. E. Onstein, Y. Xing, O. Schwery, H. P. Linder, On the complexity of triggering evolutionary radiations. *New Phytol.* **207**, 313–326 (2015).
74. H. Wickham, ggplot2. *WIREs Comput. Stat.* **3**, 180–185 (2011).
75. H. Morlon, J. Rolland, F. L. Condamine, Response to technical comment 'A cautionary note for users of linear diversification dependencies.' *Ecol. Lett.* **23**, 1172–1174 (2020).
76. R. C. Spencer, Properties of the Witch of Agnesi—Application to Fitting the Shapes of Spectral Lines. *J. Opt. Soc. Am.* **30**, 415–419 (1940).
77. H. Jeffreys, *The Theory of Probability* (OUP Oxford, 1998).
78. O. G. Pybus, P. H. Harvey, Testing Macro-Evolutionary Models Using Incomplete Molecular Phylogenies. *Proc. Biol. Sci.* **267**, 2267–2272 (2000).
79. L. J. Revell, phytools: an R package for phylogenetic comparative biology (and other things): *phytools: R package. Methods Ecol. Evol.* **3**, 217–223 (2012).
80. N. E. Helwig, *Multiple and Generalized Nonparametric Regression* (SAGE Publications Ltd,

2020) https:/doi.org/10.4135/9781526421036885885 (September 5, 2023).
81. J. H. Brown, J. F. Gillooly, A. P. Allen, V. M. Savage, G. B. West, Toward a metabolic theory of ecology. *Ecology* **85**, 1771–1789 (2004).
82. , DDD package for R: Diversity-Dependent Diversification (2023) (June 26, 2023).
83. J. Pagès, Analyse factorielle de donnees mixtes: principe et exemple d'application. *Rev. Stat. Appliquée* **52**, 93–111 (2004).
84. M. J. Anderson, "Permutational Multivariate Analysis of Variance (PERMANOVA)" in *Wiley StatsRef: Statistics Reference Online*, (John Wiley & Sons, Ltd, 2017), pp. 1–15.
85. S. Lê, J. Josse, F. Husson, FactoMineR: An R Package for Multivariate Analysis. *J. Stat. Softw.* **25**, 1–18 (2008).
86. I. Irnawati, F. D. O. Riswanto, S. Riyanto, S. Martono, A. Rohman, The use of software packages of R factoextra and FactoMineR and their application in principal component analysis for authentication of oils. *Indones. J. Chemom. Pharm. Anal.*, 1–10 (2021).
87. A. Grafen, The phylogenetic regression. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **326** (1989).
88. S. Kumar, G. Stecher, M. Suleski, S. B. Hedges, TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol. Biol. Evol.* **34**, 1812–1819 (2017).
89. J. Pinheiro, *et al.*, Package 'nlme.' *Linear Nonlinear Mix. Eff. Models* **3**, 274 (2017).
90. E. Paradis, J. Claude, K. Strimmer, APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* **20**, 289–290 (2004).
91. A. R. Ives, R^2 for Correlated Data: Phylogenetic Models, LMMs, and GLMMs. *Syst. Biol.* **68**, 234–251 (2019).

Figures and Tables

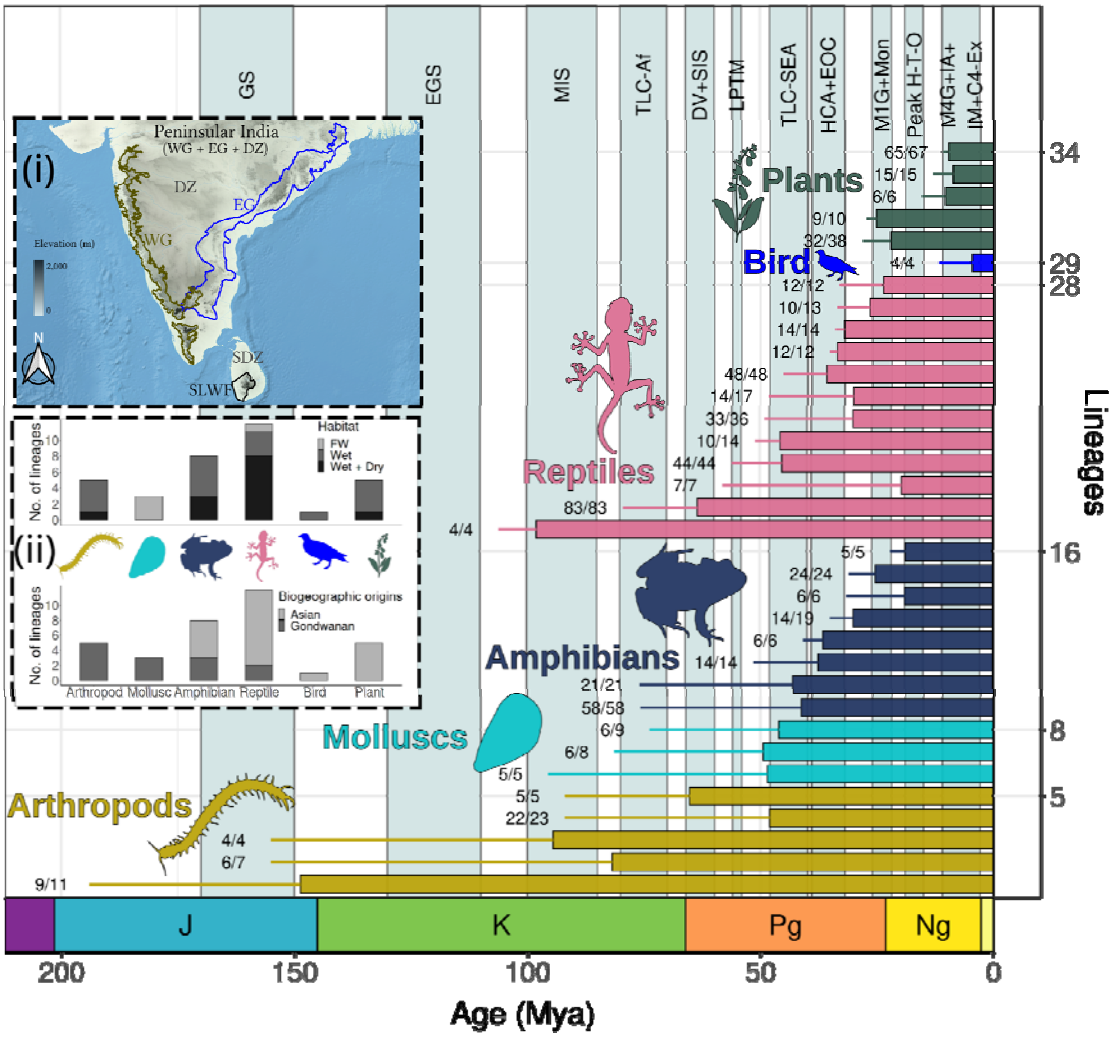

Fig. 1. Clade age, number of species, endemicity and biogeographic and habitat affinities for 34 endemic lineages from peninsular India (PI). The illustration of 34 endemic lineages of plants and animals from peninsular India with their divergence times. The numbers indicate the number of endemic species from PI by number of species present in the respective group. Key geological events (Table 1) have been highlighted using grey vertical bars. **(inset)(i)** A map of peninsular India showing a global biodiversity hotspot, the Western Ghats (denoted as WG), the Eastern Ghats (denoted as EG), and dry plains (DZ), **(ii)** Barplots indicating habitat type and biogeographic affinities for each lineage.

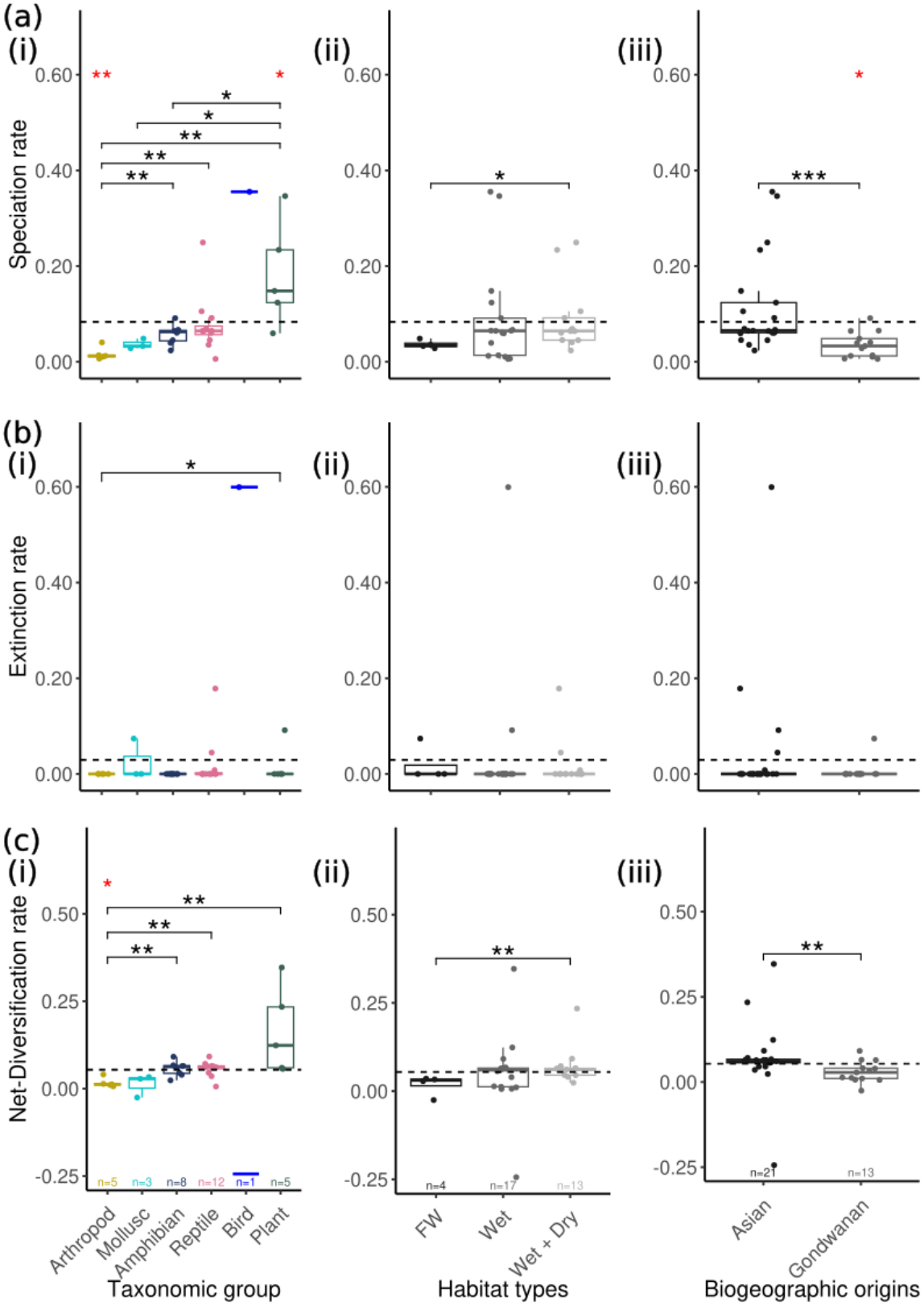
(a)
(i)
(ii)
(iii)
Speciation rate
0.60
0.40
0.20
0.00
(b)
Extinction rate
(c)
Net-Diversification rate
0.50
0.25
0.00
-0.25
n=5
n=3
n=8
n=12
n=1
n=5
n=4
n=17
n=13
n=21
n=13
Arthropod
Mollusc
Amphibian
Reptile
Bird
Plant
Taxonomic group
FW
Wet
Wet + Dry
Habitat types
Asian
Gondwanan
Biogeographic origins

Fig. 2. The pace of diversification. Comparisons of speciation **(a)**, extinction **(b)** and net-diversification rates **(c)** across taxonomic groups **(i)** (arthropods, molluscs, amphibians, reptiles, birds, plants), habitat types **(ii)** (FW - Freshwater habitat, Wet - Wet forests, Wet+Dry - Wet and dry habitats) and biogeographic origins **(iii)** (Asian and Gondwanan) in 34 lineages. Black asterisks on top of the boxplots indicate statistical significance in comparisons (* - $p < 0.05$, ** - $p < 0.01$, non-significant comparisons haven't been highlighted) and numbers indicate sample size within each category. The dotted line indicates the overall mean rate of diversification and the red asterisks indicate that the rates of the corresponding categories are significantly different from the overall mean, statistically.

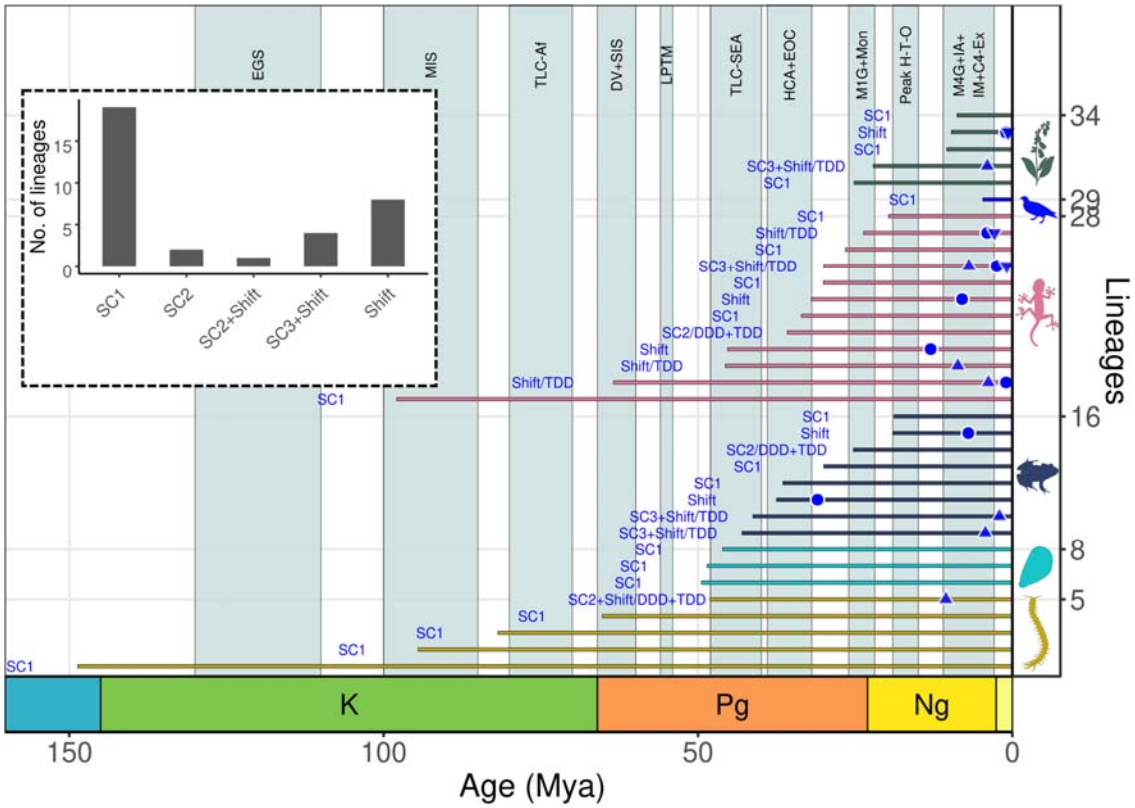

Fig. 3. Diversification patterns and drivers. Horizontal bar plots indicate clade-ages (crown ages) of 34 lineages, circles indicate symmetrical rate shifts and triangles represent one-way episodic rate shifts and their positions indicate their time of occurrence. Multiple opposing one-way rate shifts (sudden increase followed by sudden decrease or sudden decrease followed by sudden increase) within a lineage, inferred by the CoMET analysis are indicated by multiple oppositely faced triangles. Text in blue indicates the diversification scenario favoured by a lineage, followed by the prevalence of temperature (TDD) and/or diversity-dependence (DDD) in the lineage. Key geological events (Table 1) have been highlighted using grey vertical bars. **(inset)** Vertical Bar plot indicates the prevalence of different diversification scenarios in PI lineages, SC1 - gradual accumulation, SC2 - slowdowns, SC3 - Waxing and Waning and Episodic Rate Shifts.

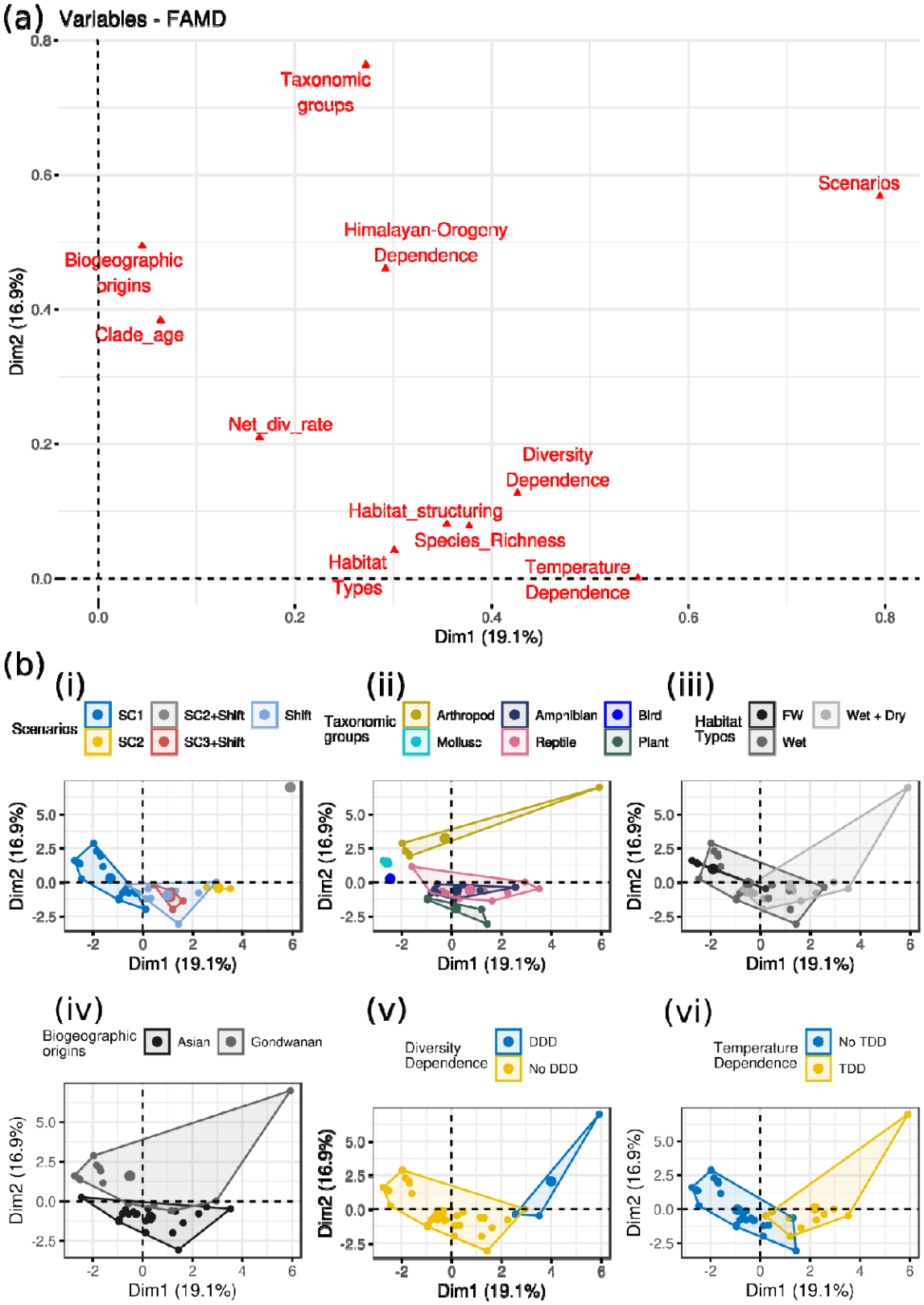
(a) Variables - FAMD
Taxonomic groups
Scenarios
Himalayan-Orogeny Dependence
Biogeographic origins
Clade_age
Net_div_rate
Diversity Dependence
Habitat_structuring
Species_Richness
Habitat Types
Temperature Dependence
Dim1 (19.1%)
Dim2 (16.9%)
(b)
(i)
Scenarios
SC1
SC2
SC2+Shift
SC3+Shift
Shift
(ii)
Taxonomic groups
Arthropod
Mollusc
Amphibian
Reptile
Bird
Plant
(iii)
Habitat Types
FW
Wet
Wet + Dry
(iv)
Biogeographic origins
Asian
Gondwanan
(v)
Diversity Dependence
DDD
No DDD
(vi)
Temperature Dependence
No TDD
TDD

Fig. 4. Ordination (Factor Analysis of Mixed Data - FAMD) plots made using 13 relevant ecological and evolutionary parameters for 34 lineages. a) The first two dimensions of the ordination analysis explain that the variables positioned in the space represent their relative contribution to the relevant axes. The positions of the variables relative to each other represent the strength of their associations. Dimension 1 is majorly explained by variables relevant to lineage identity (Taxonomic group, Biogeographic Origin and Clade age) and Dimension 2 is majorly explained by diversification and diversity dynamics (Species richness, Diversification Scenarios, etc.), **b)** Each panel represents a categorical variable and the clusters within them represent the categories. Clusters relevant to diversification scenarios are largely congruent with that of diversity-dependence and temperature-dependence. Check SI Table 5 for information relevant to PERMANOVA tests for inferring significantly distinct clusters within each categorical variable.

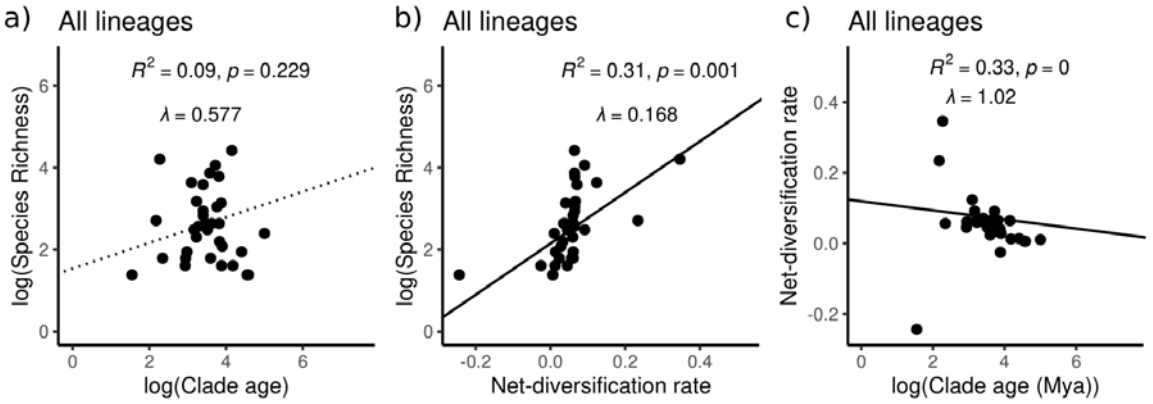

Fig. 5. Drivers of species richness. Phylogenetically corrected linear regression analysis of species richness (log-transformed) versus clade-age (log-transformed) and diversification-rates (n=34 for all the three plots). Each point is relevant to one lineage, R2 values indicate the goodness of fit of the regression lines, p-value indicates the significance of the regression, and λ indicates the degree of phylogenetic non-independence of the correlation (i.e. the phylogenetic signal in the correlation). Solid and dotted lines indicate statistically significant and non-significant linear models (significance level - $p < 0.05$), respectively.

Table 1: Major geoclimatic events relevant to peninsular India and their estimated periods of occurrence (11, 13, 14)

Geological event	Code used in further analysis	Period of occurrence (Mya)
Gondwanan split	“GS”	Between 170 and 150
East-Gondwana split	"EGS"	Between 130 and 110
Madagascar-India separation	"MIS"	Between 100 and 85
Transient land connection with Africa	"TLC-Af"	Between 80 and 70
Deccan volcanism and Seychelles-India separation	"DV+SIS"	Between 67 and 60
Last Paleocene Thermal Maxima	"LPTM"	Between 56 and 54
Transient land connection with south-east Asia	"TLC-SEA"	Between 48 and 40
Hard collision with Asia and Eocene-Oligocene cooling	"HCA+EOC"	Between 39 and 32
Mi-1 glaciation and Monsoon kicks in	"M1G+Mon"	Between 26 and 22
Peak Himalayan-Tibetan Orogeny	"Peak H-T-O"	Between 19 and 15
Mi-2 glaciation, Intensified aridification, Intensification of monsoons, Expansion of C4 plants	"M4G+IA+IM+C4-Ex"	Between 11 and 3 (overlapping ranges of events, hence clubbed together)

Table 2. Patterns and drivers of diversification. Numbers indicate the number of lineages favouring each diversification scenario and diversification-dependence on paleo-temperature, existing diversity, Himalayan orogeny and expansion of C4-plants.

Diversification scenarios	**Total no. of lineages**	**Temperature dependence**	**Diversity dependence**	**Himalayan orogeny dependence**	**Dependence on expansion of C4-plant**
Gradual accumulation (SC1)	19	0	0	0	0
Saturated accumulation (SC2)	2	2	2	0	0
Saturated accumulation + Episodic Shift (SC2+Shift)	1	1	1	1	0
Waxing-waning + Episodic Shift (SC3+Shift)	4	4	2	0	0
Symmetrical Shift (Shift)	8	3	0	0	0