Comparative Phylogeography of select scolopendrids from Western Ghats

A Thesis

submitted to

Indian Institute of Science Education and Research Pune in partial fulfilment of the requirements for the BS-MS Dual Degree Programme

by

S S Dhanush

Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,

Pashan, Pune 411008, INDIA

Date: March, 2024

Under the guidance of

Supervisor: **Dr. Jahnavi Joshi**,

Senior Scientist, CCMB Hyderabad

From May 2023 to Mar 2024

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

Certificate

This is to certify that this dissertation entitled '**Comparative Phylogeography of select scolopendrids from Western Ghats**' towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by **S S Dhanush** at the Department of Wildlife Conservation and Ecology, CCMB, Hyderabad under the supervision of **Dr. Jahnavi Joshi**, Senior Scientist, CCMB Hyderabad, during the academic year **2023-2024**.

Dr. Jahnavi Joshi

Senior Scientist

CCMB, Hyderabad

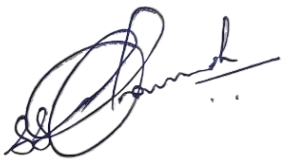
Committee:

Name of your Guide: **Dr. Jahnavi Joshi**, Senior Scientist, CCMB Hyderabad

Name of Your TAC: **Dr. Ramana Athreya**, Associate Professor, Department of Biology, IISER Pune

Declaration

I hereby declare that the matter embodied in the report entitled '**Comparative Phylogeography of select scolopendrids from Western Ghats'** are the results of the work carried out by me at the Department of Wildlife Conservation and Ecology, CCMB, Hyderabad under the supervision of **Dr. Jahnavi Joshi** and the same has not been submitted elsewhere for any other degree. Wherever others contribute, every effort is made to indicate this clearly, with due reference to the literature and acknowledgement of collaborative research and discussions.



S S Dhanush

Roll no. 20191048

Table of Contents

List of Tables

List of Figures

Abstract

Comparative phylogeography examines the patterns and drivers of genetic variation and population structure in co-distributed species from a geographical and historical perspective. It helps us elucidate how biotic and abiotic factors have historically shaped the community structure and biodiversity across an entire landscape. The Western Ghats is a mountain range that runs parallel to the west coast of India and has also been recognised as a biodiversity hotspot. Biogeographically, it can be divided into three subregions: Southern, Central and Northern Western Ghats. Further, this landscape has three geographical and climatic gaps: the Palghat Gap, Shencottah Gap and Goa Gap and multiple riverine barriers. In this study, I examined the comparative phylogeography of a few selected scolopendrids from the Western Ghats. Specifically, we examined the comparative phylogeography of the Western Ghats endemic species from the genera *Ethmostigmus* and *Rhysida*, which differ significantly in their body size, dispersal ability, and species ecology. We meticulously chose co-distributed species from across the Western Ghats with varying degrees of geographical overlap in their distribution. We hypothesised that the phylogeographical concordance would also increase as the degree of geographical overlap increases. We used mtDNA and nuDNA data to reconstruct molecular phylogeny and mtDNA to reconstruct haplotype networks and to assess the population structure. Our results support our expectation of phylogeographic concordance in co-distributed species and highlight that geography dictated the phylogeography of Scolopendrids rather than species-specific traits. Our study identifies geography, climate, and rivers as important dispersal barriers for scolopendrids, driving their genetic variation and population structure.

Acknowledgments

First and foremost, I would like to thank my Guide Dr. Jahnavi Joshi for her unwavering support, patience and guidance throughout this project. Next, I would like to thank my Thesis supervisor, Dr. Ramana Athreya for his valuable suggestions and feedbacks through this project. Next, I would like to thank my friend Sudhanshu Kumar, who is responsible for generating about half of the molecular dataset and carrying out the species delimitation analyses that I have used in my thesis. I would also like to thank our previous lab members Payal Dash and Maya Manivannan, for training me to do molecular laboratory work. I would also like to than Payal Dash for the molecular data that she had previously generated that I had used in my Thesis. Next, I would like to thank Pragyadep Roy and Dr. Aniruddha Marathe for their valuable feedback and helping me out with various analyses, which are part of my thesis. I would also like to thank Dr. Bharti D K and Pooja Pawar, whose previous work formed the basis for my current study. I would also like to thank Abhishek Gopal, Dr. Mihir Kulkarni, Pooja Pawar, Aditi Sinha, Karunakar Majhi and Satabdi Mandal for their valuable feedbacks and suggestions. I would also like to thank the HLS staff at LaCONES, CCMB for getting us necessary supplies from CCMB and for transporting our DNA samples to the Sanger Sequencing facility at CCMB. I would also like to thank Sanger Sequencing Facility at CCMB without whom this work would not have been possible. Last but most importantly I would like to thank my parents for being with me through thick and thin, without whose support I would never have been able to pursue science.

Contributions

Contributor name	Contributor role
JJ, SSD, SK, AM	Conceptualization Ideas
Contributed equally* - SSD* and SK*, PD, AM, JJ	Methodology
SSD, SK, PR, AM, JJ	Software
SSD, SK, PR, AM, JJ	Validation
SSD, SK, PR, AM, JJ	Formal analysis
JJ, AM, PR	Investigation
JJ	Resources
SSD, SK, PD, AM, JJ	Data Curation
SSD, JJ	Writing - original draft preparation
AM, JJ	Writing - review and editing
SSD, SK, AM, JJ	Visualization
JJ	Supervision
JJ	Project administration
JJ	Funding acquisition

SK: Sudhanshu Kumar, IISER Pune

SSD: SS Dhanush, IISER PUNE

JJ: Jahnavi Joshi, CSIR- CCMB, Hyderabad

AM: Aniruddha Marathe, CSIR- CCMB, Hyderabad

PR: Pragyadeep Roy, CSIR- CCMB, Hyderabad

PD: Payal Dash, Jagiellonian University, Kraków, Poland

This contributor syntax is based on the Journal of Cell Science CRediT Taxonomy[1].

[1] https://journals.biologists.com/jcs/pages/author-contributions

1. Introduction

The field of phylogeography studies the geographical distribution of genetic variation of populations from a historical perspective (Avise et al., 1987). Intraspecific phylogeographic studies help us understand how biotic and abiotic factors structure patterns of genetic variation in a single species, and comparative phylogeography helps us to understand how historical factors have shaped biodiversity and community structure across the entire landscapes by looking at co-distributed species (Moritz et al., 1998; Avise, 2000; Riddle et al., 2000). Phylogeographic pattern or structure refers to the spatial distribution of genetic variation in populations of a species. It could be inferred by looking at the genetic variation within a species in terms of phylogenetic pattern or in a haplotype network and correlating them with their geographical distribution. Phylogeographic congruence refers to the degree of similarity in phylogeographic patterns in co-distributed species. If we observe similar phylogeographic pattern in co-distributed species, we refer to it as a concordant pattern and if we observe a dissimilar pattern we refer to it as a discordant pattern. Generally, we expect concordance in the phylogeographic structure of co-distributed species, as shown in many studies involving multiple geo-climatic events or geographical barriers (Schneider et al., 1998; Avise & Walker, 1998), although several studies have also reported discordance in the phylogeographic structure of co-distributed species which is rather due to idiosyncrasies related to the species ecology or stochastic processes (Avise, 2000; Schneider et al., 1998). Moreover, there has been a paradigm shift from this concordance-discordance dichotomy towards more nuanced and biologically motivated hypothesis testing, accommodating species biology into the picture (Papadopoulou & Knowles, 2016).

Comparative phylogeography, since its days of inception, has investigated the role of biogeographic barriers, glacial refugia, and other landscape and environmental factors in structuring patterns of genetic variation in co-distributed species (Salter et al., 2021). The most common hypotheses in comparative phylogeography involve global temperature variations, past climatic fluctuations, and glacial cycles, particularly during the quaternary period (Gutiérrez-García & Vázquez-Domínguez,

2011). Few phylogeographic studies have been conducted on terrestrial invertebrates (like grasshoppers, beetles, and moths) (Salter et al., 2021). Short life cycles, the ability to maintain viable populations with relatively small habitats, limited dispersal and habitat specialisation resulting in somewhat patchy distributions make it an excellent study model for understanding how environment and landscape structure genetic variations (Salter et al., 2021). However, despite these apparent advantages, the proportion of comparative phylogeographic studies relative to the total number of phylogeographic studies in terrestrial invertebrates is relatively low (<8% of total terrestrial invertebrate phylogeography studies as of 2008). Additionally, most phylogeographic studies, including those for invertebrates, are from the Northern Hemisphere and the Southern Hemisphere, and the tropics, despite being a cradle of biodiversity, are generally neglected (Beheregaray, 2008).

Within the tropics, the diverse soil arthropods of the Western Ghats are an excellent study system for comparative phylogeography studies. The Western Ghats mountain range runs along the western coast of peninsular India for a distance of 1600 km from 8°N to 21°N (Subramanyam and Nayar, 1974; Ali & Ripley,1987). Western Ghats is recognised as a biodiversity hotspot due to its high diversity and endemicity (Myers et al., 2000). Biogeographically, Western Ghats hosts taxa with Gondwanan affinity as it was part of the Gondwana supercontinent and also taxa with Asian affinity, which dispersed into India after the collision (Mani,1974; Subramanyam and Nayar,1974). The Western Ghats comprises three distinct zones: Southern Western Ghats (SWG), Central Western Ghats (CWG), and Northern Western Ghats (NWG) (Biswas and Karanth, 2021). There are three geographical gaps in the Western Ghats mountain ranges: the Palghat Gap (PG), the Shencottah Gap (SG), and the Goa Gap (GG). Of these, the GG is more of a climatic than a geographical barrier and has a more recent origin than SG and PG. The duration of the dry period also shows a significant positive gradient from south to north within the Western Ghats, especially north of the Goa gap (Pascal*,1986*). Geographical gaps have been identified as potential dispersal barriers, given that they have considerably lower elevations than nearby areas in a mountain range. Several studies have examined these gaps' role in the Western Ghats as a dispersal barrier and found that PG is a more prominent dispersal barrier across taxa than GG and SG. However, the number of studies that have examined the role of SG and GG is limited (Biswas and

Karanth, 2021 & references therein). Previous studies on the centipede genus *Digitipes* have shown that geographical gaps such as PG have not significantly influenced the diversification of the genus but within species variation (e.g. *D. coonoorensis*) (Joshi and Karanth, 2012, 2013). However, phylogeographic studies on other vertebrate groups, both large and small, have shown that SG and PG could have acted as a significant dispersal barrier (Biswas and Karanth, 2021 & references therein). In addition to this, GG is important, given that there is a stark gradient in seasonality from south to north (Pascal,1986). Thus, we expect to observe concordance in the geography of phylogenetic breaks across co-distributed species, i.e., Type III concordance, according to Avise, 2000 in most of our focal species.

Within soil arthropods, the class Chilopoda, comprising centipedes, is a fascinating group for their diversity and distribution and evolutionary history of over 420 million years coupled with limited dispersal ability (Edgecombe and Giribet, 2019). Of these, the family *Scolopendridae* (*Scolopendromorpha)* exhibit high levels of diversity and endemicity in the Western Ghats and is represented by three tribes: *Scolopendrini*, *Asanadini* and *Otostigmini* (Joshi and Karanth, 2011). We are interested in two genera, *Rhysida* and *Ethmostigmus,* within Otostigmini. The biogeography and phylogenetic relationships within these two genera have been well studied within the Indian subcontinent, especially within the Western Ghats. Given these, we have chosen multiple endemic species within these two genera with varying levels of overlap in their distribution for our study. If geographic and historical factors majorly drive phylogeographic patterns, we would observe a positive relationship between the degree of range overlap and phylogeographic congruence. And if species ecology mainly drives phylogeographic patterns, then the relationship between phylogeographic congruence and range overlap would be stochastic. In other words, species with even high levels of geographic range overlap would not necessarily show phylogeographic congruence. Thus, our null model, as in most comparative phylogeographic studies, is that co-distributed species or species with overlapping geographic ranges would show discordance in phylogeographic patterns, assuming that they have responded to historical events and geographical factors majorly in an idiosyncratic manner.

Geographical range overlap between any two species could be visualised as any of the four scenarios. In Scenario 1 (SC1), the geographical ranges are mutually exclusive. In Scenario 2 (SC2), the geographical ranges partially overlap. In Scenario 3 (SC3), the geographical range of one species is nested within the other. Moreover, finally, in Scenario 4 (SC4), the geographical ranges of the two species overlap completely (or nearly completely). Thus, if geographical factors dictate phylogeography, then as we go from SC1 to SC4, the phylogeographic concordance would increase. If we do not observe such a pattern, species ecology could explain the idiosyncratic phylogeographic patterns.

To incorporate species ecology into our study design, we have meticulously selected species with overlapping species distribution belonging to genera *Ethmostigmus* and *Rhsyida* within the family Scolopendridae in our dataset. Species belonging to the genus *Ethmostigmus* are giant centipedes, which could be up to 30 cm long and are much more prominent in body size and body mass than species of *Rhysida* (Joshi et al., 2020). Since species with larger body masses have higher energy requirements, they usually have a wider range size (McNab,1963). This means the larger species are more likely to disperse than the smaller ones. This positive relationship between body size and dispersal distance is generally true for most active dispersers, as shown in a meta-analysis by Jenkins et al. (2007). Thus, we expect higher levels of dispersal in species of *Ethmostigmus* than in *Rhysida*. Higher dispersal would imply high gene flow, and as a result, genetic structuring might be less pronounced within the species of *Ethmostigmus* than in the *Rhysida*. This could influence asynchronies in the phylogeographic patterns of overlapping species belonging to the genera *Rhysida* and *Ethmostigmus*. However, within a genus, body size variability is negligible and hence is not expected to be an important factor dictating the distribution of genetic variation and structuring patterns.

Furthermore, body size in centipedes was inversely related to genetic diversity in centipedes in a meta-analysis (Bharti et al., 2023). It might be driving asynchronous patterns while comparisons are made between species with overlapping geographic distributions belonging to the two genera. Species also show a differential preference for elevation and habitat based on our limited understanding. These differences in elevational range also indicate differences in species niche and could be responsible

for the asynchronous patterns if observed. It is also important to note that these factors driving the distribution and clustering of intraspecific genetic variation genetic variation might not be acting independently of each other and could be acting in a concerted manner.

Another plausible driver of phylogeographic structure is the climatic fluctuations associated with Quaternary Glaciations. C4 plant signals in the Western Ghats during the Last Glacial Maxima (LGM) (20,000 B.P) indicate that it might have been a period of aridity in the Western Ghats. This was followed by the interglacial period when aridity slowly declined (Sukumar et al.,1993). Thus, the Scolopendrid centipede species would have occupied a refugium and expanded following the LGM. Thus, population demographic statistics like Tajima's D and R2 would indicate signatures of recent population expansions. This expectation should likely be valid for all our species, given that their origins within the Western Ghats pre-date the LGM. Also, if the species population had expanded following LGM from a refugial population just after the LGM, we would expect the species haplotype network to follow a star topology. Furthermore, it is interesting to determine whether the known geographical gaps in the Western Ghats landscape represent regions of unsuitable condition using species distribution models.

Species could be classified into a spectrum, from habitat specialists to habitat generalists. It has been hypothesised that specialists usually have smaller effective population sizes and lesser genetic diversity than generalist species. This hypothesis has been termed a specialist-generalist variation hypothesis (SGVH) (Li et al., 2014). This suggests that specialists require specialised resources, as a result of which, they have patchy distributions. Due to this, specialists have lower effective population sizes and are thus more susceptible to meta-population dynamics and higher odds of extirpations, in turn reducing their overall genetic diversity. We could hypothesise that species occupying a wider geographical and elevational range have wider niches. Moreover, a meta-analysis examining the relationship between ecological niche breadth and geographical range recovered a significant positive relation, making our argument pertinent. Under this assumption, we expect a generalist and a specialist (w.r.t to the geographical and elevational range) with overlapping distributions to have discordant phylogeographic structures in co-

distributed regions. Based on the SGVH hypothesis, we expect specialists to have a larger number of smaller populations than generalists, where we expect a smaller number of larger populations. In addition, we expect the overall intraspecific-genetic diversity in specialists to be lower than that of generalists.

In this light, we have chosen eight species from the Western Ghats Landscape for our analysis. Of these, *Ethmostigmus agasthyamalaensis*, *Rhysida pazhuthara* and *Rhysida aspinosa* are endemic and co-distributed in the southern western ghats. *Ethmostigmus praveeni* and *Rhysida lewisi* are endemic to CWG, while *Ethmostigmus coonooranus* distribution is spread across the CWG and SWG. *Rhysida sada* and *Ethmostigmus sahyadrensis* are NWG endemics. Thus, these species represent co-distributed species across the Western Ghats landscape, allowing us to understand the factors shaping the diversity and community structure spanning the entire Western Ghats. We hypothesise that geography, mainly geographical barriers, would play an important role in structuring species' genetic variation within the Western Ghats for these low-dispersing organisms. We also hypothesise that body size might also be a driving factor of genetic variation and would play an important role in structuring patterns of genetic variation. Through this study, we hope to shed light on the role of geography and species ecology in driving patterns of genetic diversity and population structure within the Western Ghats Landscape.

2. Materials and Methods

2.1 Sampling

We sampled scolopendrid species from multiple sites within protected and non-protected areas covering the states of Kerala, Tamil Nadu, Karnataka, and Maharashtra across the Western Ghats. The fieldwork to collect the centipede samples was carried out over three years, from 2020 to 2023, by the Evolutionary Ecology Lab members. We carried out active sampling using forceps to turn rocks, wooden logs, and leaf litter on the forest floor with rakes and caught centipedes. We brought live to the lab in plastic containers with soil and then preserved in 70-75 % Ethanol.

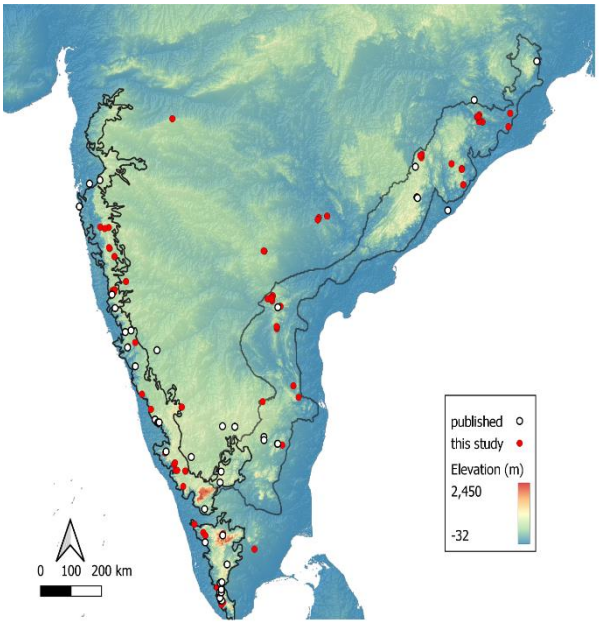


Fig. 1a

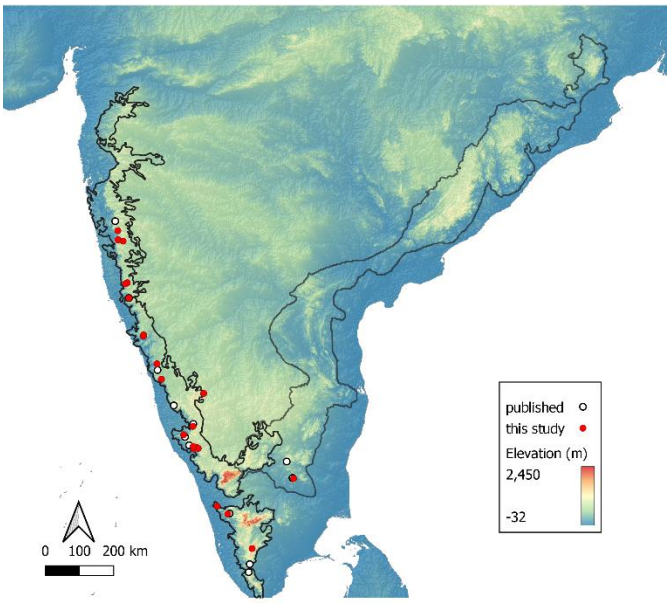


Fig. 1b

Fig 1. Map showing sampling locations of the genera *Rhysida* and *Ethmostigmus* within peninsular India. **Fig. 1a** on the left is the map for the genus *Rhysida*, while **Fig. 1b** is the map for the genus *Ethmostigmus*. The white and the red solid circle represent locations from published studies and current study respectively. A north arrow and scalebar is provided for reference.

2.2 Taxonomy

We carefully examined the preserved samples under an Olympus compound microscope and then classified them up to species level using the diagnostic characters mentioned in key (Joshi et al., 2020 and Joshi and Edgecombe, 2018). In cases where the diagnostic characters were unclear, samples were identified up to the genera level under a microscope, and then species were assigned based on genetic data.

2.3 DNA Extraction

We performed DNA extraction in 154 and 33 individuals representing the genera *Rhysida* and *Ethmostigmus*, respectively, preserved in 70-75% ethanol. We extracted DNA, mainly using leg tissue and segments, in case the samples were too small. The number of legs used for extraction was determined according to the size and the quality of the preserved sample. On average, we have used about 4-5 legs for most samples of the *Rhysida* genera and one leg for samples of the genus *Ethmostigmus*. We cut the tissue using a scalpel and transferred it to a 1.5 ml or 2 ml tube with 180 ul of lysis buffer (T1 buffer) present in MN (Macherey Nagel) blood and tissue DNA Extraction kit and crushed it with the help of a micro-pestle so that the interior tissue devoid of chitin layer comes in contact with the lysis buffer. Further, 25 ul of proteinase K was added to the crushed tissue so the proteins could degrade and not affect our further reactions. This reagent mixture was then incubated for overnight digestion in a thermal shaker at 56°C at 800 r.p.m. Further, the DNA extractions were carried out using the MN kit following the steps specified in the MN blood and tissue DNA extraction kit for animal tissue. Finally, the samples were eluted in 50ul of the kit elution buffer and stored at 4°C in a refrigerator.

2.4 DNA Quantification

We quantified the extracted DNA in a NanoDrop® ND-1000 spectrophotometer. Samples with satisfactory 260/280 (~1.8) and 260/230 (2-2.2) values and concentrations (>80 ng/ul) were diluted to 30ng/ul by adding the kit elution buffer. We then stored the sample dilutions at 4°C in a refrigerator. If the quality was not satisfactory, the DNA extractions were repeated.

2.5 Polymerase Chain Reaction (PCR)

We amplified the diluted samples for 16s, 28s and COI DNA markers using a 3-step DNA PCR reaction. The 16s and COI are mitochondrial markers, while the 28s is a nuclear marker. We used these markers as the PCR conditions have been standardised for these markers for centipedes, and the same markers have been used in previous studies as well, thus facilitating the collation of molecular data from multiple studies for this and future studies. In a PCR reaction, the template DNA is amplified for a specific region using deoxyribonucleotides(dNTPs) and DNA primers. The standardised reagent concentrations and volumes for the PCR reaction mixture we have used in this study (for 25 ul of the reaction mixture) are given in the supplementary. The temperature and durations of the PCR steps have already been standardised and were used in the PCR reactions. The information about markers used, reagent concentrations and reaction conditions are specified in Table 2,3 and 4 respectively in the Supplementary. The PCR products are stored at 4°C in a refrigerator following the PCR reaction.

2.6 DNA Gel Electrophoresis and DNA Sequencing

We then subjected the DNA products to DNA gel electrophoresis, and the gel image was visualised using EtBr under UV light in a gel camera. In a DNA gel electrophoresis, the DNA moves according to the charge/mass ratio from the negative to the positive node of the electrophoresis machine. Thus, the smaller bands get separated most. If clear bands were observed, they were compared with the DNA ladder to verify if the band is of the appropriate size. If the gel bands were of appropriate size, the respective PCR product was sent to the Sanger sequencing facility at CCMB, Hyderabad, for sequencing. The PCR products were further purified in the facility and subjected to Sanger sequencing.

2.7 DNA Sequence Cleaning and Pairwise Alignment

The Forward and Reverse DNA Chromatograms for each sample for each marker were carefully inspected in Chromas (version 2.6.7) and inspected for sequence quality. Subsequently, the forward and the reverse sequences were aligned using the ClustalW algorithm (Thompson et al., 1994) in Mega7 (Kumar et al., 2016) under default settings, trimmed, edited and stored as FASTA files.

2.8 Multiple Sequence Alignment and Maximum Likelihood Phylogeny

We constructed individual gene trees and concatenated gene trees for *Ethmostigmus* and *Rhysida* using the maximum likelihood framework in the IQTree web server (Trifinopoulos et al.,2016). We reconstructed phylogenetic trees for multiple datasets, firstly for the genus *Rhysida* and *Ethmostigmus* individually for 16s, COI, mitochondrial concatenated (16s and COI) and all markers concatenated (16s, COI and 28s). For this, we made multiple sequence alignments (MSA) with forward sequences in Mega11(Tamura et al., 2021) using default settings for the Muscle Algorithm (Edgar, 2004). Closely related species were chosen as outgroups. For the genus *Ethmostigmus*, we also added species from Australasia and Africa and species from the genus *Otostigmus* and *Alluropus carinulata* as outgroups. For the genus *Rhysida* we used species from the genus *Ethmostigmus* and *Otostigmus* in outgroups.

We made alignments separately from each marker. We checked for stop codons in the alignment for the COI MSA by translating the sequences in Mega11 using the invertebrate mitochondrial codon table to ensure that pseudogenes were not amplified. We assigned partitions to the COI alignment using codon position. For the concatenated mitochondrial (16s, COI) and nuclear-mitochondrial concatenated, we partitioned the alignment into four partitions (1 for 16s and 3 for COI based on codon position) and five partitions (1 for 16s,1 for 28s and 3 for COI) respectively. We ran the Phylogenetic tree in the IQTree web server with 1000 Ultra-fast Bootstraps (Hoang et al., 2018) rooted the trees using outgroup, and visualised them in Figtree(v 1.4.4). The model of sequence evolution was selected in IQTree.

2.9 Species Delimitation

We performed species delimitation for *Ethmostigmus* and *Rhysida* using the mitochondrial marker data separately. We used the maximum likelihood tree we had generated earlier using IQTree and the multiple sequence alignment used in phylogeny as the input data. We ran Multi-rate Poisson Tree Process (mPTP) (Kapli et al., 2017) on a Linux Kali device using bash script. However, since the outgroup, as well as species within the ingroup located exclusively outside peninsular India, were not sampled well enough in our phylogeny, carrying out delimitation with the entire dataset would result in erroneous species delimitation of the outgroup as well

as ingroup taxa outside peninsular India, which would, in turn, affect the delimitation of our Indian peninsular ingroup. Thus, we have extracted only our Indian peninsular clade and one outgroup using the ape package in R. In contrast, only the peninsular Indian clade was maintained in the alignment. The final results were mapped onto the concatenated IQ Tree (16s, COI and 28s), and we used these results along with their occurrence data to select species for our further analysis.

2.10 Species Haplotype Network

The forward sequences of all samples belonging to a single species from the cleaned pairwise alignment were aligned using the Muscle algorithm in Mega11 under default settings for each mitochondrial marker (16s, COI) separately. The multiple sequence alignments were saved as a NEXUS file, to which a GEOTAG block was added, containing information about each sample's latitude and longitude and the number of clusters the haplotype network is expected to have. This nexus file generated a TCS (Templeton, Crandall and Sing, 1992) haplotype network in PopART (version 1.7) (Leigh and Bryant, 2015). A TCS algorithm was used to construct the haplotype network as it allows the inference of ancestral haplotypes as existing haplotypes, unlike other popular methods. The haplotypes generated in PopART using the nexus file were then clustered by latitude and longitude in PopART for the selected number of clusters in the nexus file. The expected number of clusters was inferred based on the number of phylogenetic clusters observed for each species in the concatenated maximum likelihood tree.

2.11 Population Genetic Statistics

Intraspecific Nucleotide diversity (pi) (Miraldo et al., 2016), Tajima's D (Tajima, 1989) and R2 statistic (Ramos-Onsins and Rozas, 2002) were computed using the functions nuc.div, Tajima.test and R2.test respectively for each species for 16 sequence data using pegas package (version 1.3) (Paradis, 2010) in R (version 4.2.2). Further, we computed Tajima's D and R2 statistics for each population cluster and reported the p-values. Tajima's D statistic indicates species populations' deviation from neutrality assumptions. A negative value <-1 of Tajima's D suggests

that the population has recently expanded. The R2 statistic is also similar, with values close to zero indicating recent population expansion.

2.12 Inferring comparative phylogeographic pattern

We constructed haplotype networks as mentioned above and then compared them with their geographical distribution using maps generated in QGIS (version 3.14.15). We examined the haplotype networks for signatures of population expansion, the presence of biogeographic barriers and examined the distribution of the haplotypes with respect to the geography of the landscape. We then assessed these patterns across co-distributed species in our dataset to classify the phylogeographic patterns as concordant or discordant.

3. Results

3.1 Data

We generated sequence data for 154 individuals of the genus *Rhysida* and 33 individuals of *Ethmostigmus,* respectively. Additionally, we used secondary molecular data for *Rhysida* and *Ethmostigmus* from Joshi et al. (2020) and Joshi and Edgecombe 2018), respectively. Altogether, our compiled dataset consisted of molecular data for 224 individuals of the genus *Rhysida* and 55 individuals of *Ethmostigmus*, respectively. The species-wise details of our compiled molecular dataset for Peninsular India for both genera are summarised in Table 1 below. We have also used additional sequences from published literature outside India as outgroups in our phylogeny. The details of the entire dataset including outgroups used in phylogeny are reported in Supplementary Table 1.

Genus	*Species*	*Total samples*	*Molecular data generated*			*Molecular data from existing literature*		
			CO1	*16S*	*28S*	*CO1*	*16S*	*28S*
Rhysida	*longipes*	13	5	5	6	7	7	7
Rhysida	*crassispina*	3	0	1	1	2	2	1
Rhysida	*pazhutharа*	24	11	15	15	8	9	8
Rhysida	*konda*	32	21	26	26	3	4	3
Rhysida	*trispinosa*	72	43	49	52	17	18	18
Rhysida	*aspinosa*	5	0	1	1	2	4	3
Rhysida	*sada*	7	0	5	5	1	2	2
Rhysida	*immarginata*	5	0	1	1	4	4	4
Rhysida	*lewisi*	28	13	16	16	11	12	12
Rhysida	*ikhalama*	3	0	0	0	2	3	2
Rhysida	*sp.1*	4	1	1	0	3	3	3
Rhysida	*sp.2*	14	5	10	12	2	2	2
Rhysida	*sp.3*	2	1	2	2	0	0	0
Rhysida	*sp.4*	10	6	10	10	0	0	0
Rhysida	*sp.5*	2	2	2	2	0	0	0
Ethmostigmus	*agasthyamal-aiensis*	4	1	1	1	3	3	3
Ethmostigmus	*sahyadrensis*	20	15	16	13	3	3	3
Ethmostigmus	*coonoorанus*	13	6	9	8	3	4	1
Ethmostigmus	*praveeni*	4	7	6	4	4	4	4
Ethmostigmus	*tristis*	6	2	1	1	4	4	3

Table 1: Species-wise summary of compiled molecular data for the genera *Ethmostigmus* and *Rhysida* from Peninsular India

3.2 Species Delimitation and Likelihood Phylogeny

The concatenated maximum likelihood tree showed high support for the monophyly of the clade for Peninsular Indian species of both *Ethmostigmus* (BS =100) and *Rhysida* (BS=96). The phylogeny for both the genera had further split within the species level, which we have interpreted as distinct population clusters. We used this as the number of population clusters in the 16S haplotype network, provided the BS was greater than 75. Also, if the overall sample size of the species haplotype network was less than 8, we did not assign population clusters.

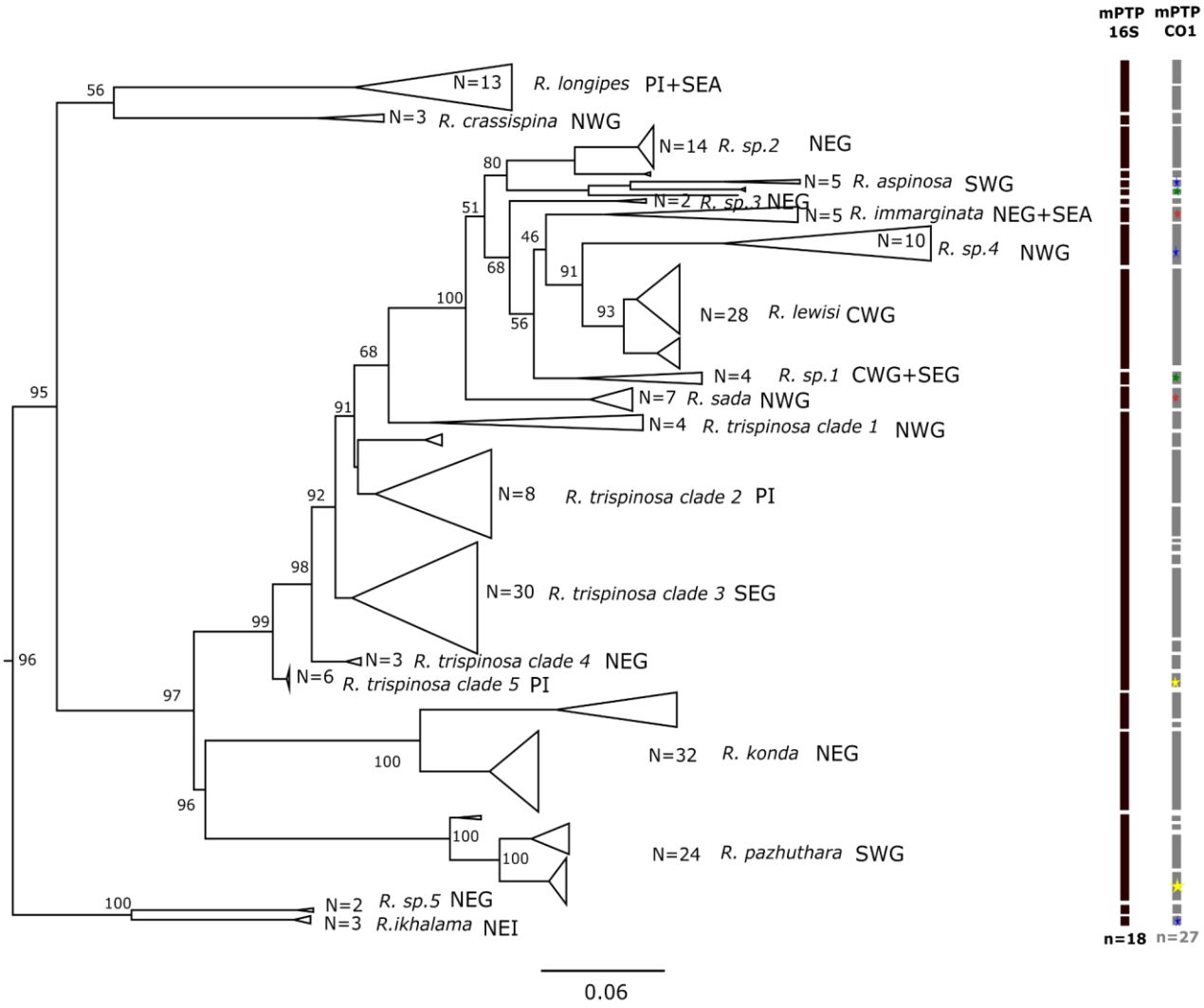


Fig 2a. Species delimitation results mapped on to the concatenated maximum likelihood Phylogeny for *Rhysida* species within peninsular India. The solid black and grey lines represent the species identified in 16S and COI mPTP analyses. The node values in the phylogeny represent Bootstrap Support (BS), and the tip labels represent species identity and sample size. The abbreviations used in the figure are **PI: Peninsular India, SEA: South East Asia, SWG: Southern Western Ghats, NWG; Northern Western Ghats, CWG; Central Western Ghats, NEG; Northern Eastern Ghats, SEG; Southern Eastern Ghats and NEI: North East India** respectively.

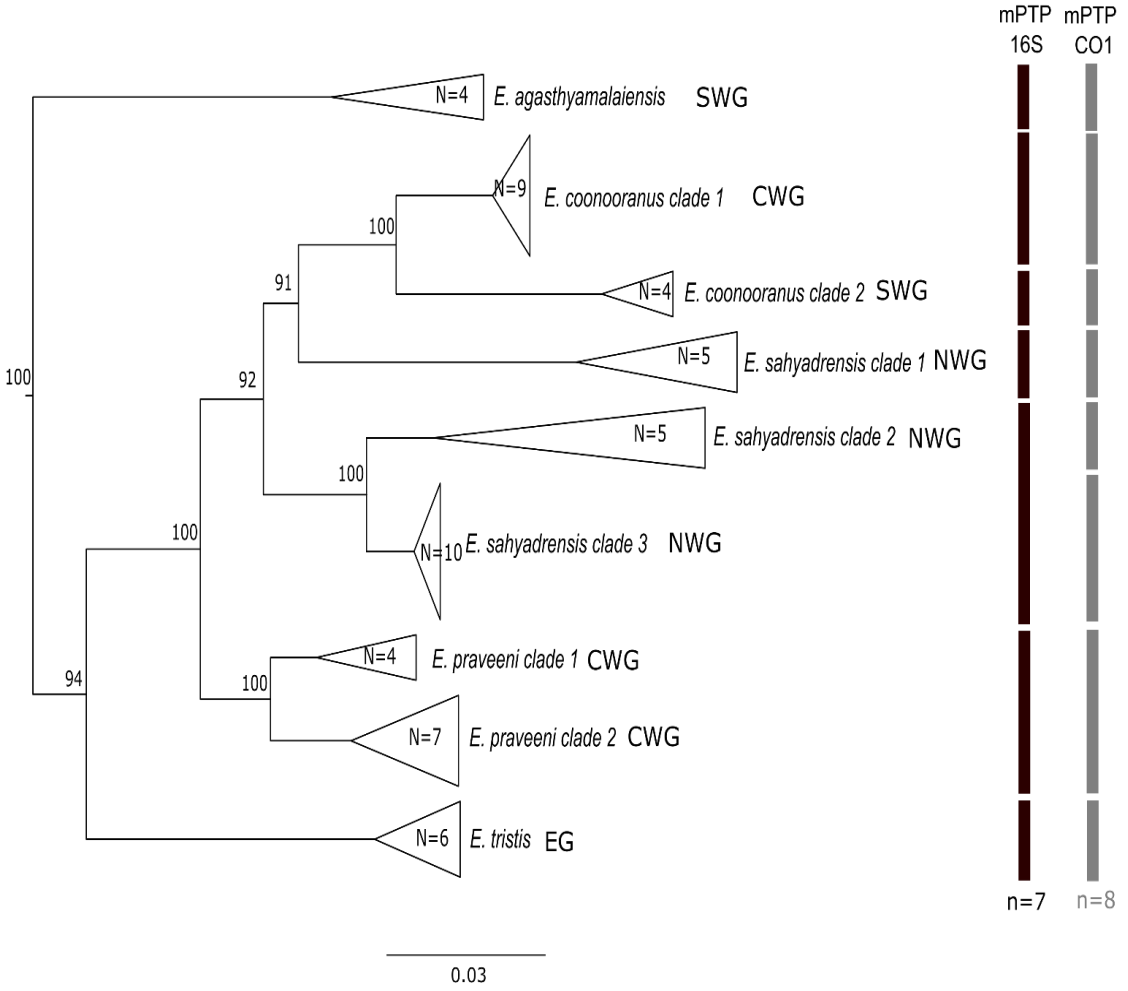
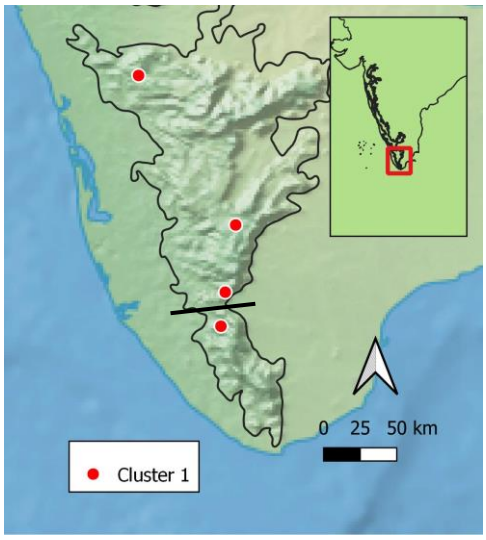


Fig 2b. Species delimitation results mapped on to the concatenated maximum likelihood Phylogeny for *Ethmostigmus* species within peninsular India. The solid black and grey lines represent the species identified in 16S and COI mPTP analyses respectively. The node values in the phylogeny represent Bootstrap Support (BS) and the tip labels represent species identity and sample size. The abbreviations used in the figure are **SWG: Southern Western Ghats, NWG: Northern Western Ghats, CWG; Central Western Ghats and EG: Eastern Ghats** respectively.

3.3 Phylogeography

3.3.1 *Ethmostigmus agasthyamalaiensis*

The 16S TCS Network recovered showed a star haplotype topology, which indicates that the species populations have undergone a population expansion from a single ancestral population at a somewhat recent period. The star topology also indicates no apparent directionality in dispersal from the ancestral population and follows a uniform trajectory. This was also supplemented by Tajima's D statistic of -2.2611 (p-value < 0.05) for 16s data. However, R2 statistic of 0.171, although close to zero, indicating population expansion, was non-significant. The overall intraspecific genetic diversity calculated using nucleotide diversity (Pi) of *E. agasthyamalaensis* based on the 16S marker was 0.029. We did not further assign population clusters to the samples to examine patterns of gene flow between population clusters due to the small sample size for *E. Agasthyamalaensis* (n=4).



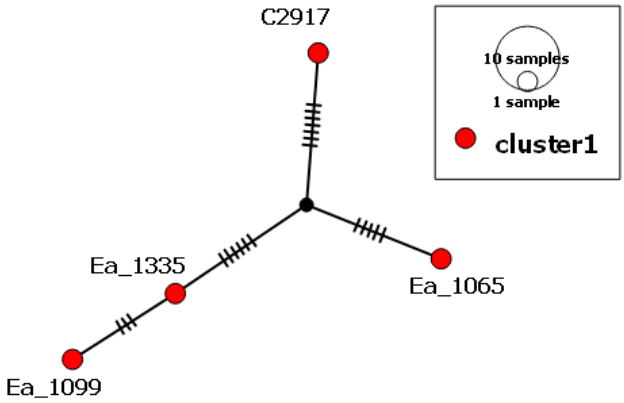
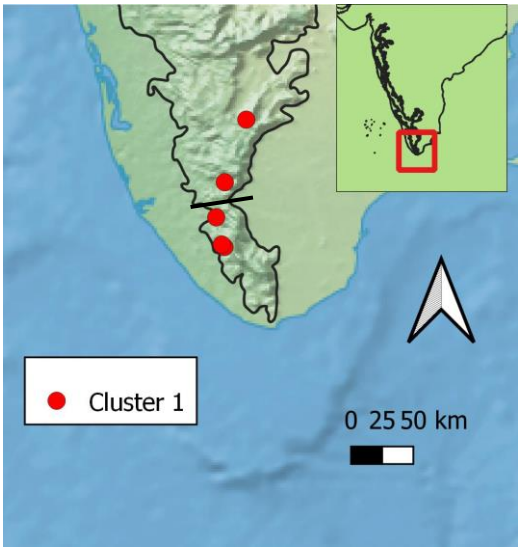


Fig 3a **Fig 3b**

Fig 3. a) map showing geographical distribution of haplotype clusters in the Western Ghats for ***E. agasthyamalaiensis***. The dark solid black line in the map represents the Shencottah gap. The inset map shows the sampling locations with respect to the Peninsular Indian Landscape b)16s TCS haplotype network for ***E. agasthyamalaiensis*** (n=4).

3.3.2 *Rhysida aspinosa*

The 16s TCS Network recovered seven distinct haplotypes. Of the five samples, the samples from Periyar Tiger Reserve, Achankovil Reserve Forest and Shendurney Wildlife Sanctuary were quite distinct from other samples, and the node connecting the haplotypes of all other samples from the node connecting these two samples differed by 13 mutations. Moreover, the haplotype from Periyar Tiger Reserve differs by 45 and 46 mutations, respectively, from the other two haplotypes. Thus, there must be at least three distinct populations with minimal gene flow within *R. aspinosa*. The distinct haplotypes from Periyar and Achankovil are samples above the Shencottah gap, indicating the role of the Shencottah gap acting as a dispersal barrier. The sample sizes are, however, quite low; hence, no further inferences were made. However, the value of Tajima's D statistic for the species was 0.1, and the R2 statistic was 0.2056, which was statistically nonsignificant. The value of overall intraspecific genetic diversity nucleotide diversity (Pi) was 0.072. Due to the low sample size, we did not assign populations or clusters to the samples.



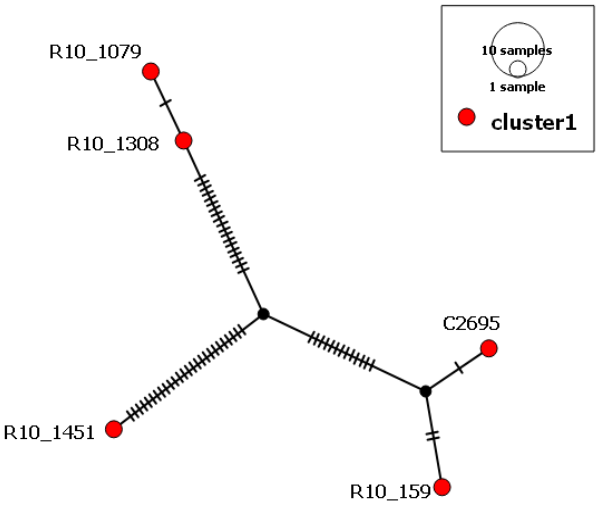


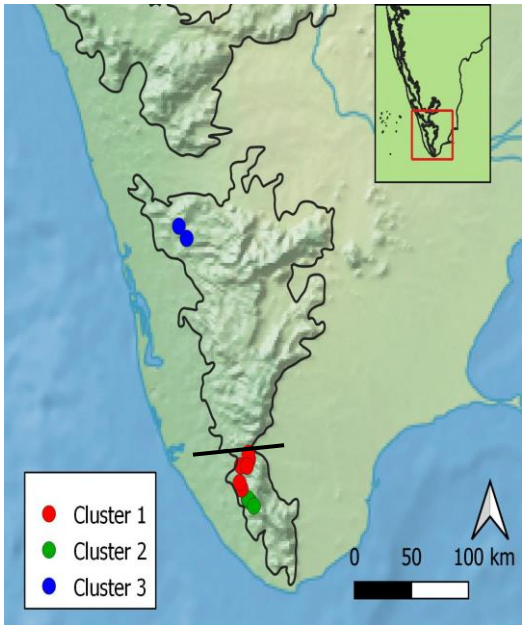
Fig 4a **Fig 4b**

Fig 4. a) map showing geographical distribution of haplotype clusters in the Western Ghats for ***R. aspinosa***. The dark solid line in the map represents the Shencottah

gap. The inset map shows the sampling locations with respect to the Peninsular Indian Landscape. b) 16s TCS haplotype network for ***R. aspinosa*** (n = 5).

3.3.3 *Rhysida pazhuthara*

The 16s TCS Network for *Rhysida pazhuthara* recovered 17 haplotypes. Moreover, the network follows a star topology, indicating recent range expansion from an ancestral population and uniform dispersal without directionality. The haplotypes were clustered into three clusters based on the concatenated maximum likelihood phylogeny, which recovered three distinct clades for *Rhysida pazhuthara*, indicative of three distinct population clusters. However, population clusters 1 and 2 overlap geographically and show very low levels of admixture between them. This indicates that there are unique populations within a connected geographical landscape comprising regions within or in close proximity of the Agasthyamalai Biosphere Reserve. The other cluster is representative of samples from Parambikulam Tiger Reserve, which is geographically quite distant from the Agasthyamalai Biosphere Reserve and is above the Shencottah gap, indicating that the Shencottah gap might be acting as an important dispersal barrier for populations of *Rhysida pazhuthara*. Tajima's D and R2 statistics were non-significant for all clusters except for Cluster 2, for which Tajima's D value was -0.7196 ($p < 0.05$), indicative of a recent population expansion. We did not compute population statistics for Cluster 3 as sample size was small ($n < 4$).



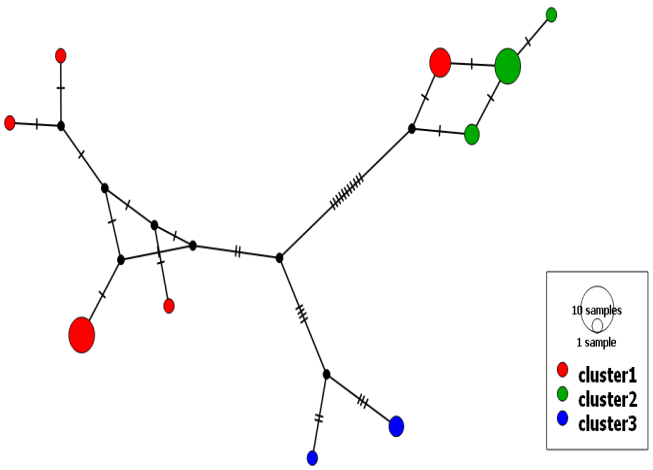
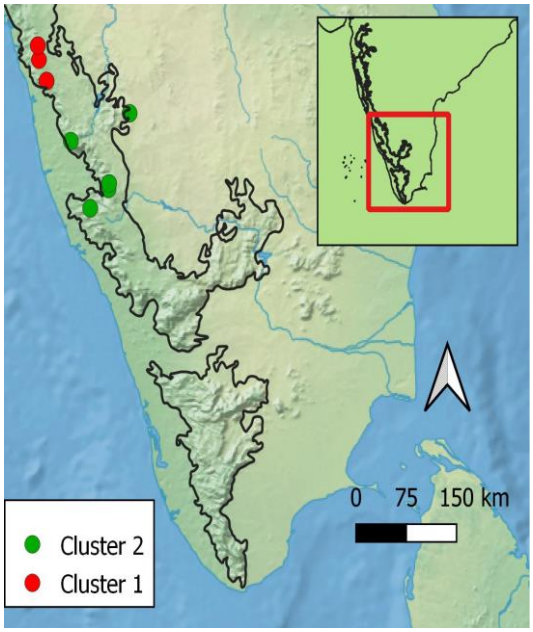


Fig 5a **Fig 5b**

Fig 5. a) map showing geographical distribution of haplotype clusters in the Western Ghats for ***R. pazhuthara***. The dark solid line in the map represents the Shencottah gap. The inset map shows the sampling locations with respect to the Peninsular Indian Landscape. b)16s TCS haplotype network for ***R. pazhuthara*** (n = 25).

3.3.4 *Ethmostigmus praveeni*

The 16S *Ethmostigmus Praveeni* network recovered 22 haplotypes, further assigned into two clusters based on the concatenated maximum likelihood phylogeny. The two distinct clusters corresponded to the north and south of the Sharavathi River, indicating that the Sharavathi River could act as a dispersal barrier between the two distinct population clusters. However, the haplotype network also indicates some level of gene flow between the two distinct population clusters. The overall nucleotide diversity was reported to be 0.0453. Tajima's D and R2 statistics were statistically insignificant for all population clusters except the R2 statistic for Cluster 2, which was reported to be 0.1177.



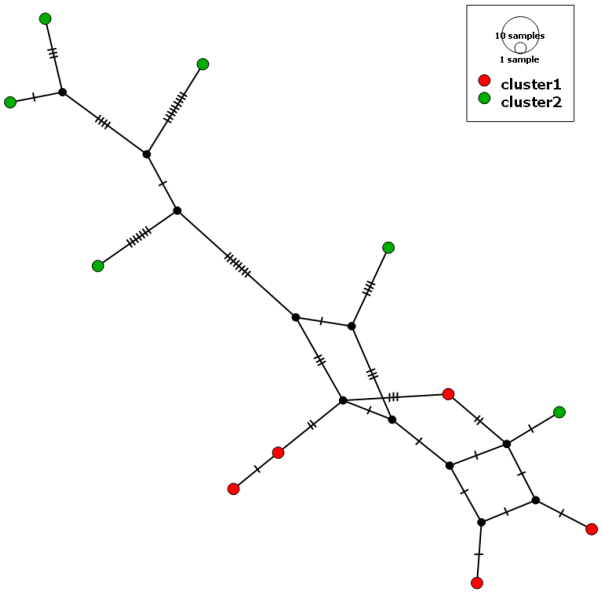
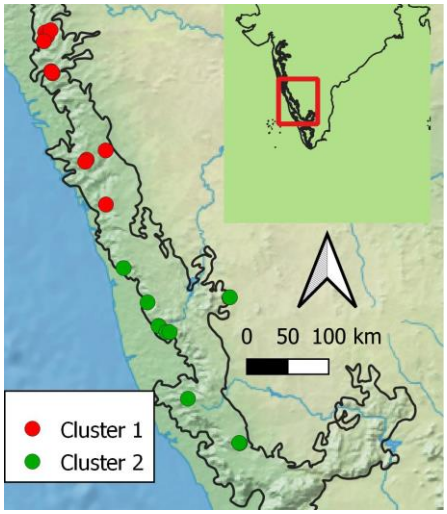


Fig 6a **Fig 6b**

Fig 6. a) map showing geographical distribution of haplotype clusters in the Western Ghats for ***E. praveeni***. The inset map shows the sampling locations with respect to the Peninsular Indian Landscape. b) 16s TCS haplotype network for ***E. praveeni*** (n = 11).

3.3.5 *Rhysida lewisi*

The 16s TCS Network for *Rhysida lewisi* recovered 15 haplotypes. Based on the concatenated maximum likelihood phylogeny, the haplotypes were clustered into two distinct population clusters. Population clusters 1 and 2 corresponded to samples from above and below the Sharavathi River, indicating that the Sharavathi River could be a barrier for gene flow. Moreover, in the haplotype network, there was no indication of admixture between the two population clusters. The TCS network indicates that the two clusters are not genetically very distant, indicating that the Sharavathi River might be a reasonably strong barrier to the gene flow of populations of *Rhysida lewisi*. The value of the nucleotide diversity statistic (pi) was reported to be 0.0128 based on the 16s marker Tajima's D and R2 statistics, which were significant for all the population clusters, indicating strong evidence for recent population expansion.



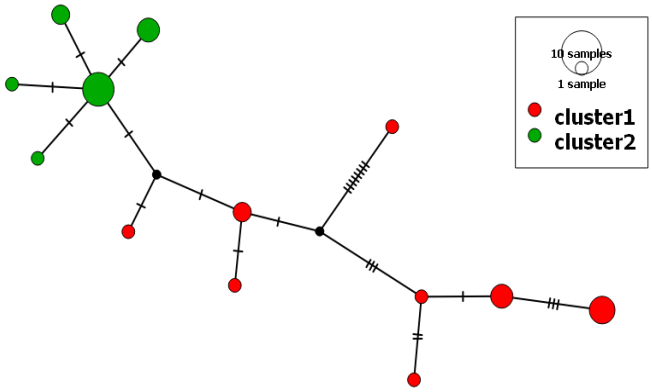
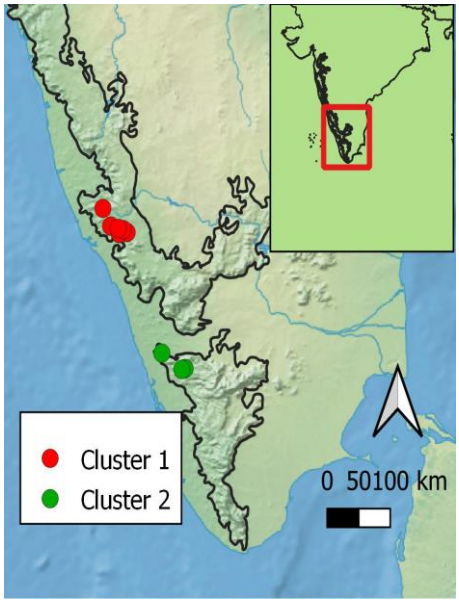


Fig 7a **Fig 7b**

Fig 7. a) map showing geographical distribution of haplotype clusters in the Western Ghats for ***R. lewisi***. The inset map shows the sampling locations with respect to the Peninsular Indian Landscape. b)16s TCS haplotype network for ***R. lewisi*** (n = 27).

3.3.6 *Ethmostigmus coonooranus*

The 16s TCS Networks for *Ethmostigmus coonooranus* recovered 16 haplotypes. The haplotypes were clustered into two distinct population clusters based on the maximum likelihood phylogeny. Cluster 1 and 2 were geographically separated by the Palghat Gap, indicating that the Palghat Gap might act as a barrier for dispersal between the two distinct population clusters. In the 16S haplotype network, there was no evidence of admixture between the two populations. Moreover, the two population clusters were genetically distinct by 24 mutations. This implies that the Palghat Gap is a strong geographical barrier to dispersal for *Ethmostigmus conooranus*. The overall intraspecific nucleotide genetic diversity was reported to be 0.0392. Tajima's D and R2 statistics were statistically insignificant except for population cluster 2, for which Tajima's D was reported to be -4.1611, indicating recent population expansion.



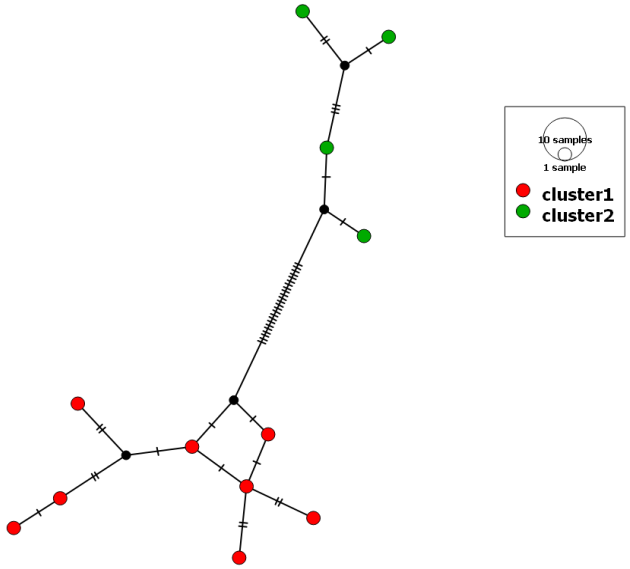


Fig 8a **Fig 8b**

Fig 8. a) map showing geographical distribution of haplotype clusters in the Western Ghats for ***E. coonooranus***. The inset map shows the sampling locations with respect to the Peninsular Indian Landscape. b)16s TCS haplotype network for ***E. coonooranus*** (n = 12).

3.3.7 *Ethmostigmus sahyadrensis*

The haplotype network recovered 16 haplotypes assigned into three distinct population clusters based on the Maximum Likelihood tree. Of these, population cluster 3 comprises samples below the Goa gap, while population clusters 1 and 2 are both above the Goa gap. The haplotype network indicates very little gene flow and admixture between the three distinct population clusters. Interestingly, population cluster 1, much above the Goa Gap, is genetically more distinct from the population clusters below the Goa Gap than population cluster 2, which comprises samples just above the Goa Gap. These results indicate that the Goa Gap as a climatic barrier is a reasonably strong dispersal barrier for populations of *Ethmostigmus sahyadrensis* and is indicative of the effect of the stark precipitation gradient across the Goa Gap. The overall intraspecific nucleotide diversity was reported to be 0.0456. Tajima's D statistic was statistically significant for population clusters 1 and 2, indicating recent population expansion.

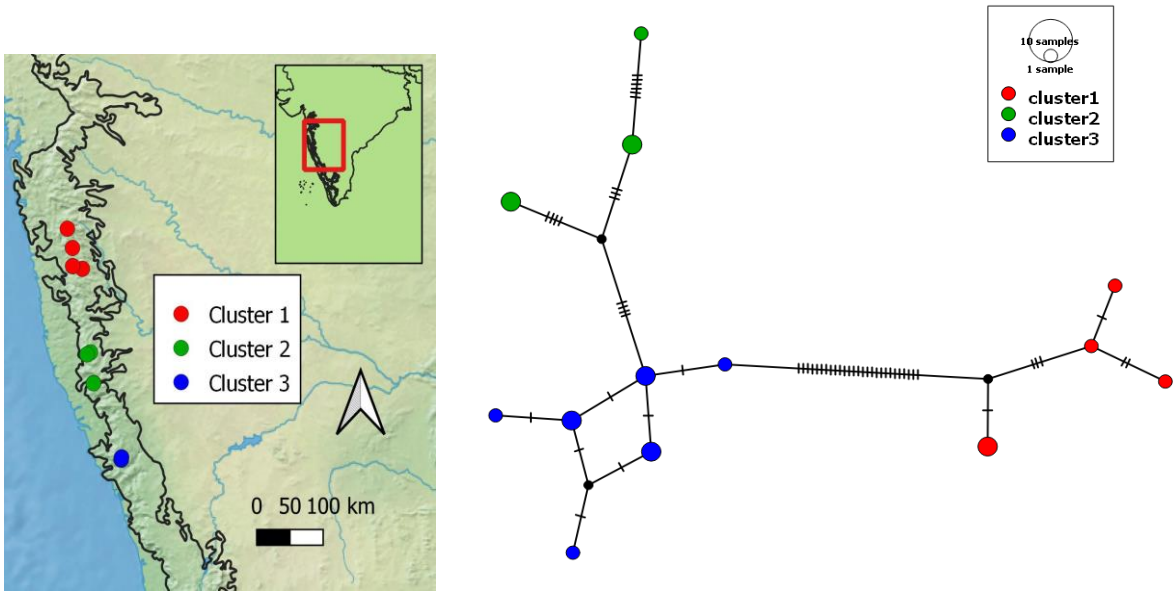
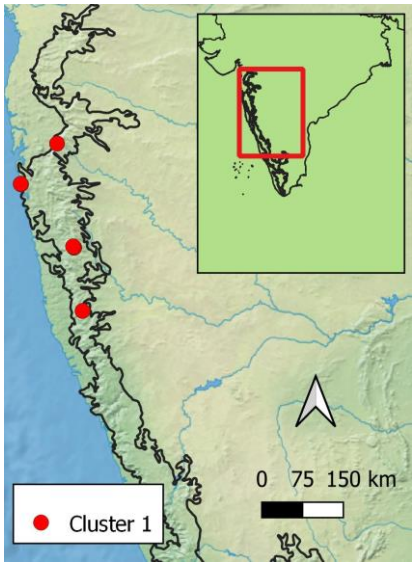


Fig 9a **Fig 9b**

Fig 9. a) map showing geographical distribution of haplotype clusters in the Western Ghats for ***E. sahyadrensis***. The inset map shows the sampling locations with respect to the Peninsular Indian Landscape. b)16s TCS haplotype network for ***E. sahyadrensis*** (n = 19).

3.3.8 *Rhysida sada*

The 16s TCS network recovered seven haplotypes for *Rhysida sada*. The network shows a star-like topology but deviates slightly from this topology. Haplotypes recovered were not very distinct, except for the sample from Radhanagari Wildlife Sanctuary, which is the northmost of all the samples in this network. Most individuals fall above the Goa Gap, and their structuring could be more evident across the Goa Gap. The overall intraspecific nucleotide genetic diversity(pi) was 0.0085. Tajima's D and R2 statistics were statistically insignificant, thus showing no evidence of recent population expansions.



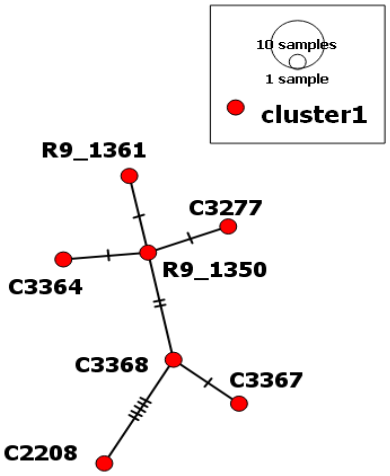


Fig 10a **Fig 10b**

Fig 10. a) map showing geographical distribution of haplotype clusters in the Western Ghats for ***R. sada***. The inset map shows the sampling locations with respect to the Peninsular Indian Landscape. b)16s TCS haplotype network for ***R. sada*** (n = 7).

Species	Population Cluster	Marker	Sample size	No. of Haplotypes	Nucleotide Diversity	Tajima's D	R2 statistic
E. agasthyamalaiensis	Combined	16s	4	5	0.029	-2.2611 (p <0.05)	0.171 (ns)
R. aspinosa	Combined	16s	5	7	0.072	0.2078 (ns)	0.1682 (ns)
R. pazhuthara	Combined	16s	25	17	0.0293	-1.807 (ns)	0.1104 (ns)
R. pazhuthara	Cluster 1	16s	13	-	-	-0.7196 (ns)	0.1104 (ns)
R. pazhuthara	Cluster 2	16s	9	-	-	-3.708 (p <0.05)	0.1527 (ns)
R. pazhuthara	Cluster 3	16s	3	-	-	-	-
E. praveeni	Combined	16s	11	22	0.0453	0.1058 (ns)	0.1318 (ns)
E. praveeni	Cluster 1	16s	5	-	-	0.9814 (ns)	0.2227 (ns)
E. praveeni	Cluster 2	16s	6			0.2687 (ns)	0.1177 (p-value < 0.05)
R. lewisi	Combined	16s	27	15	0.0128	-1.981 (p-value <0.05)	0.0507 (p-value < 0.05)
R. lewisi	Cluster 1	16s	14	-	-	-2.2214 (p-value <0.05)	0.0868 (p-value <0.05)
R. lewisi	Cluster 2	16s	13	-	-	-1.9661 (p-value <0.05)	0.092 (p-value < 0.05)
E. coonooranus	Combined	16s	12	16	0.0392	0.5372 (ns)	0.1508 (ns)
E. coonooranus	Cluster 1	16s	8	-	-	-1.3534	0.1065

						(ns)	(p-value < 0.05)
E. coonooranus	Cluster 2	16s	4	-	-	-4.1611 (p-value <0.05)	0.237 (ns)
E. sahyadrensis	Combined	16s	19	16	0.0456	-0.2430 (ns)	0.1102 (ns)
E. sahyadrensis	Cluster 1	16s	5	-	-	-1.9587 (p-value =0.05)	0.1652 (ns)
E. sahyadrensis	Cluster 2	16s	5	-	-	-4.1814 (p-value < 0.05)	0.2396 (ns)
E. sahyadrensis	Cluster 3	16s	9	-	-	-0.5157 (ns)	0.1489 (ns)
R. sada	Combined	16s	7	7	0.0085	-0.9342 (ns)	0.1526 (ns)

Table 2: Population genetic statistics of all species

3.4 Comparative Phylogeography

Species Name	**Distribution Scenario**	**Body size**	**Split across biogeographic barrier**	**Haplotype Network**	**Population Genetic parameter**
E. agasthyamalaiensis - SWG	SC3 w.r.t *R. aspinosa* & *R. pazhuthara*; SC2 w.r.t *E. coonooranus*	large	not inferred	uniform dispersal from an ancestral population**,** no structuring in haplotype	Signature of recent population expansion, genetic diversity = 0.029
R. aspinosa - SWG	SC3 w.r.t to *E. agasthyamalaiensi*	small	yes- SG	some level of structuring in	No signatur

	s; SC4 w.r.t *R. pazhuthara;* SC2 w.r.t to *E. coonooranus*			haplotype	e of recent populati on expansi on, genetic diversit y = 0.072
R. pazhuthara - SWG	SC3 w.r.t to *E. agasthyamalaiensi s;* SC4 w.r.t *R. aspinosa;* SC2 w.r.t *E. coonooranus*	small	yes-SG	Structuring of haplotypes into 3 clusters, almost uniform dispersal from an ancestral population	Signatu re of recent populati on expansi on for only one populati on cluster, genetic diversit y = .0293
E. praveeni - CWG	SC4 w.r.t R. *lewisi*; SC3 w.r.t *E. coonooranus*	large	yes- Sharavathi River	two haplotype clusters north and south of Sharavathi	Signatu re of recent populati on expansi on for only one populati on cluster, genetic diversit y = 0.0453
R. lewisi - CWG	SC4 w.r.t E. *praveeni* and SC3 w.r.t *E. coonooranus*	small	yes- Sharavathi River	two haplotype clusters north and south of Sharavathi river	Signatu re of recent recent populati

					on expansion for all population clusters, genetic diversity = 0.0128
E. coonooranus - CWG and SWG	SC3 w.r.t *R. lewisi* and *E. praveeni* and SC2 w.r.t to R. pazhuthara, *R. aspinosa* and *E. agasthyamalaiensis*	large	yes-Palghat Gap	two haplotype clusters north and south of Palghat Gap	Signature of recent population expansion for one population cluster, genetic diversity = 0.0392
E. sahyadrensis **-** NWG	SC3 w.r.t R. *sada*	large	yes- Goa Gap	three haplotype clusters, one above Goa Gap	Signature of recent population expansion in two population clusters, genetic diversity = 0.0456
R. sada - NWG	SC3 w.r.t *E. sahyadrensis*	small	not inferred	Star like topology indicating uniform	No signature of recent

				dispersal from ancestral population, no haplotype clustering	population expansion, genetic diversity = 0.0085

Table 3: Comparative Phylogeography Summary

4. Discussion

4.1 Geography dictates Phylogeography

Comparative phylogeographic studies help us understand the role of biotic and abiotic factors in driving community structure and biodiversity (Moritz et al., 1998; Avise, 2000; Riddle et al., 2000). In most studies involving a biogeographical barrier or geo-climatic events, the null expectation is concordance in co-distributed phylogeography. Our initial hypothesis was also on these lines, and we expected that the geoclimatic barriers within the Western Ghats would play a major role in explaining the observed patterns. However, if we were to observe discordant patterns in the phylogeography of co-distributed species, this could be attributed to idiosyncrasies in species ecology. In our study, given the lack of information on other species traits, the variation in body size would be able to explain the observed idiosyncratic patterns. Body size for actively dispersing organisms has been shown to scale with dispersal ability positively (Jenkins et al., 2007), and Bharti et al. (2021) have shown that body size is a major driver of genetic diversity in centipedes and scales inversely with genetic diversity. Thus, we expected that even among co-distributed centipede species, we might observe idiosyncrasies due to their difference in body size at the genus level and possibly different dispersal abilities.

In the Southern Western Ghats, our focal species are *R. pazhuthara*, *R. aspinosa*, and *E. agasthyamalaensis*. We expected the Shencottah gap to play a major role in the phylogeography in the region, and we expected to recover distinct population

clusters on either side of the Shencottah gap. We recovered distinct population clusters on either side of the Shencottah gap only in the case of *R. pazhuthara* from our haplotype network. Due to the small sample size, we couldn't assign clusters to the population of other species in the southern Western Ghats. However, our haplotype network recovered distinct haplotypes north of the Shencottah Gap, indicating the existence of distinct population clusters across the Shencottah Gap. However, the statistical confidence for this observation is quite low.

Further taxon sampling and sequence data are required to assign population clusters. We plan to add more sequence data to our analyses to make more conclusive inferences. Intraspecific nucleotide genetic diversity was the lowest for *E. agasthyamalaiensis* among the three SWG endemics, which could be attributed to the larger body size among the three. However, the differences in nucleotide diversity among them were similar. Our focal species in the Central Western Ghats were *R. lewisi* and *E. praveeni*. In both these species, we recovered distinct population clusters on either side of the Sharavathi River, indicating the role of the Sharavathi River as a dispersal barrier. The genetic diversity was lower for *R. lewisi* and was higher than that for *E. praveeni,* which deviates from our expectation of body size genetic diversity expectation. However, the difference was again negligible. For *E. coonooranus*, which has a distribution spread across the SWG and CWG, we recovered distinct population clusters on either side of the Palghat Gap, indicating the role of the Palghat Gap as a geographic barrier. Among NWG endemics, we recovered distinct population clusters on either side of the Goa Gap for *E. sahyadrensis*, indicating the role of the Goa Gap as a climatic barrier in the NWG. For *R. sada*, we could not assign haplotypes owing to the small sample sizes, which need to be supplemented by more taxon and sequence data.

Many studies have looked at the role of geographical gaps in structuring intraspecific and interspecific genetic variation. Biswas and Karanth (2021) have reviewed multiple such studies within the Western Ghats landscape. Their results highlighted that the Palghat Gap is a significant intraspecific biogeographic barrier. This study also highlighted that studies across other barriers in the Western Ghats landscape, like the Shencottah Gap, Goa Gap, and other riverine barriers, have been limited. Moreover, many of these studies have been on birds and mammals that show relatively higher dispersal with respect to taxa like centipedes, for which population

structuring is expected to be much more prominent. In this light our study highlights the importance of other barriers as well within the landscape of the Western Ghats. Our study shows that geography, mainly biogeographic barriers, mainly drives phylogeographic patterns, and we have been able to recover distinct populations on either side of the barriers like Shencottah Gap, Palghat Gap, Goa Gap and even in the case of the Sharavathi River. We also could not observe any idiosyncratic patterns, thus indicating that climate and geography dictate phylogeographic patterns rather than species-specific traits. Globally as well, many other studies have shown similar patterns, where geography, especially geographical gaps, have been the major driver of phylogeographic structure like Schneider et al. (1998), which showed similar phylogeographic structure across the Black Mountain Corridor (BMC) for several herpetofauna in Australian tropical wet forest. The same study also recovered idiosyncratic patterns for some species under consideration. Idiosyncratic patterns are quite common in cases where the taxa under consideration differ in their species ecology, especially in their dispersal ability. Hence, we have built our hypotheses and study design to incorporate species ecology into the picture, especially with respect to body size, which is a proxy for dispersal ability. Despite this, we could not recover such patterns in our study, hinting that the difference in dispersal ability between the two genera under consideration might not be significant as per our assumption. Our results highlight the need to test for our recovered biogeographic barriers in other taxa as well as whose species distribution passes through these barriers, especially for taxa showing lower dispersal ability like centipedes.

4.2 Future directions and limitations

Although our study indicates that geography, especially geographical gaps, drives phylogeographic structure, our lines of evidence in this study are limited. Our inferences are based on haplotype networks, and population genetics statistics are based on a single marker. We need to include samples from their respective geographic range extremes for several of our species, which still need to be sampled. We used only a single marker, 16S, as our concatenated maximum likelihood phylogeny was heavily biased towards this marker, and it was

geographically the best sampled mitochondrial marker among the two. For statistical confidence, one needs to look at more mitochondrial markers, which we plan to do in the near future. Our study also recovered signatures of recent population expansion, which might be indicative of signatures from the Last Glacial Maxima (LGM), which was a period of aridity in the Western Ghats (Sukumar et al.,1993) when centipedes would have occupied refugia and would have dispersed out in our population genetic statistics and haplotype networks. Species Distribution Models (SDMs) are an excellent tool to address this and would show an expansion in suitable habitat if we compare LGM distribution with current SDM. SDMs are also an explicit way of testing for historical factors that have played a role in species phylogeography, which is missing from our current study. In addition, they can also be used to ascertain our recovered biogeographical barriers, as they would be regions with low habitat suitability scores in the SDMs. Although we have devised multiple geographical overlap scenarios, our current study needs to distinguish fine differences between these scenarios; for this, one needs a quantitative framework. Phylogeographic Concordance Factor (PCF) (Salter and Carstens, 2016), which gives a value between 0 and 1 based on the concordance between species based on their sampling location and species phylogeny, would be quite useful. Hence, we plan to calculate the same for our species. Above all, as we have indicated in previous sections, our study needs to be supplemented by more sequence data and possibly more marker data.

Supplementary

Link: https://drive.google.com/file/d/1E0bHPYs3VfJOslfuQM-ynuhHqDH8W3HN/view?usp=sharing

References

1. Avise, J. C., & Walker, D. (1998). Pleistocene phylogeographic effects on avian populations and the speciation process. Proceedings of the Royal Society of London. Series B: Biological Sciences, 265(1395), 457–463.
2. Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A., & Saunders, N. C. (1987, November). INTRASPECIFIC PHYLOGEOGRAPHY: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. *Annual Review of Ecology and Systematics*, *18*(1), 489–522.
3. AVISE, J. C. (2000). *Phylogeography: The History and Formation of Species*. Harvard University Press.
4. BEHEREGARAY, L. B. (2008). Twenty Years of phylogeography: The state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology*, *17*(17), 3754–3774.
5. Bharti, D. K., Edgecombe, G. D., Karanth, K. P., & Joshi, J. (2021). Spatial patterns of phylogenetic diversity and endemism in the Western Ghats, India: A case study using ancient predatory arthropods. *Ecology and Evolution*, *11*(23), 16499–16513.
6. Bharti, D. K., Pawar, P. Y., Edgecombe, G. D., & Joshi, J. (2023). Genetic diversity varies with species traits and latitude in predatory soil arthropods (myriapoda: Chilopoda). *Global Ecology and Biogeography*, *32*(9), 1508–1521.
7. Biswas, A., & Praveen Karanth, K. (2021). Role of geographical gaps in the Western Ghats in shaping intra- and interspecific genetic diversity. *Journal of the Indian Institute of Science*, *101*(2), 151–164.
8. Clement, M., Snell, Q., Walker, P., Posada, D., & Crandall, K. (2002). TCS: Estimating gene genealogies. Parallel and Distributed Processing Symposium, International Proceedings, 2, 184.
9. D.T. Hoang, O. Chernomor, A. von Haeseler, B.Q. Minh, and L.S. Vinh (2018) UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol., 35:518–522

10. Edgar R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, *32*(5), 1792–1797.
11. Giribet, G., & Edgecombe, G. D. (2019). The phylogeny and evolutionary history of arthropods. *Current Biology*, *29*(12), R592-R602.
12. Gutiérrez-García, T. A., & Vázquez-Domínguez, E. (2011). Comparative phylogeography: Designing studies while surviving the process. *BioScience*, *61*(11), 857–868.
13. J. Trifinopoulos, L.-T. Nguyen, A. von Haeseler, and B.Q. Minh (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res., 44 (W1):W232-W235.
14. Jenkins, D. G., Brescacin, C. R., Duxbury, C. V., Elliott, J. A., Evans, J. A., Grablow, K. R., ... & Williams, S. E. (2007). Does size matter for dispersal distance?. *Global Ecology and Biogeography*, *16*(4), 415-425.
15. Joshi, J., & Edgecombe, G. D. (2018). Molecular phylogeny and systematics of the centipede genus Ethmostigmus Pocock (Chilopoda: Scolopendromorpha) from peninsular India. *Invertebrate Systematics*, *32*(6), 1316-1335.
16. Joshi, J., & Karanth, K. P. (2011). Cretaceous–Tertiary diversification among select scolopendrid centipedes of South India. *Molecular Phylogenetics and Evolution*, *60*(3), 287-294.
17. Joshi, J., & Karanth, K. P. (2012). Coalescent method in conjunction with niche modeling reveals cryptic diversity among centipedes in the Western Ghats of South India.
18. Joshi, J., & Karanth, P. (2013). Did southern Western Ghats of peninsular India serve as refugia for its endemic biota during the Cretaceous volcanism?. *Ecology and Evolution*, *3*(10), 3275-3282.
19. Joshi, J., Karanth, P. K., & Edgecombe, G. D. (2020). The out-of-India hypothesis: evidence from an ancient centipede genus, Rhysida (Chilopoda: Scolopendromorpha) from the Oriental Region, and systematics of Indian species. *Zoological Journal of the Linnean Society*, *189*(3), 828-861.
20. Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., & Flouri, T. (2017). Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics*, *33*(11), 1630-1638.

21. Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular biology and evolution*, *33*(7), 1870-1874.
22. Leigh, J. W., & Bryant, D. (2015). POPART: full-feature software for haplotype network construction. *Methods in ecology and evolution*, *6*(9), 1110-1116.
23. Li, S., Jovelin, R., Yoshiga, T., Tanaka, R., & Cutter, A. D. (2014). Specialist versus generalist life histories and nucleotide diversity in Caenorhabditis nematodes. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1777), 20132858.
24. Mani, M. S. (Ed.). (2012). *Ecology and biogeography in India* (Vol. 23). Springer Science & Business Media.
25. McNab, B. K. (1963). Bioenergetics and the determination of home range size. *The American Naturalist*, *97*(894), 133-140.
26. Miraldo, A., Li, S., Borregaard, M. K., Flórez-Rodríguez, A., Gopalakrishnan, S., Rizvanovic, M., ... & Nogués-Bravo, D. (2016). An Anthropocene map of genetic diversity. *Science*, *353*(6307), 1532-1535.
27. Moritz, C., & Faith, D. P. (1998). Comparative phylogeography and the identification of genetically divergent areas for conservation. *Molecular ecology*, *7*(4), 419-429.
28. Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), 853-858.
29. Papadopoulou, A., & Knowles, L. L. (2016). Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. *Proceedings of the National Academy of Sciences*, *113*(29), 8018-8024.
30. Paradis, E. (2010). pegas: an R package for population genetics with an integrated–modular approach. *Bioinformatics*, *26*(3), 419-420.
31. Pascal, J. P. (1988). Wet evergreen forests of the Western Ghats of India.
32. Ramos-Onsins, S. E., & Rozas, J. (2002). Statistical properties of new neutrality tests against population growth. *Molecular biology and evolution*, *19*(12), 2092-2100.
33. Riddle, B. R., Hafner, D. J., Alexander, L. F., & Jaeger, J. R. (2000). Cryptic vicariance in the historical assembly of a Baja California Peninsular Desert

biota. *Proceedings of the National Academy of Sciences*, *97*(26), 14438-14443.

34. Satler, J. D., & Carstens, B. C. (2016). Phylogeographic concordance factors quantify phylogeographic congruence among co-distributed species in the Sarracenia alata pitcher plant system. *Evolution*, *70*(5), 1105-1119.
35. Satler, J. D., Carstens, B. C., Garrick, R. C., & Espíndola, A. (2021). The phylogeographic shortfall in hexapods: a lot of leg work remaining. *Insect Systematics and Diversity*, *5*(5), 1.
36. Schneider, C. J., Cunningham, M., & Moritz, C. (1998). Comparative phylogeography and the history of endemic vertebrates in the Wet Tropics rainforests of Australia. *Molecular Ecology*, *7*(4), 487-498.
37. Slatyer, R. A., Hirst, M., & Sexton, J. P. (2013). Niche breadth predicts geographical range size: a general ecological pattern. *Ecology letters*, *16*(8), 1104-1114.
38. Smith, M. L., Lang, C., Sneddon, D., Wallace, J., Espíndola, A., Sullivan, J., & Carstens, B. C. (2023). Comparative Phylogeography of Microsnails from the Pacific Northwest. *Northwest Science*, *96*(1-2), 117-132.
39. Subramanyam, K., & Nayar, M. P. (2001). Vegetation and phytogeography of the Western Ghats. *MEMOIRS-GEOLOGICAL SOCIETY OF INDIA*, (2), 945-960.
40. Sukumar, R., Ramesh, R., Pant, R. K., & Rajagopalan, G. (1993). A δ13C record of late Quaternary climate change from tropical peats in southern India. *Nature*, *364*(6439), 703-706.
41. Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, 123(3), 585-595.
42. Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular biology and evolution*, *38*(7), 3022-3027.
43. Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research*, *22*(22), 4673-4680.