Insect Systematics and Diversity, 6(6), 3; 2022, 1–21
https://doi.org/10.1093/isd/ixac028
Research

ENTOMOLOGICAL SOCIETY OF AMERICA
SHARING INSECT SCIENCE GLOBALLY

OXFORD

Evolution

Climate Relicts: Asian Scorpion Family Pseudochactidae Survived Miocene Aridification in Caves of the Annamite Mountains

Stephanie F. Loria,[1,2,3,7,] Valentin L. Ehrenthal,[3,4] Anh D. Nguyen,[5,6] and Lorenzo Prendini[1]

[1]Scorpion Systematics Research Group, Division of Invertebrate Zoology, American Museum of Natural History, New York, NY, USA, [2]Richard Gilder Graduate School, American Museum of Natural History, New York, NY, USA, [3]Section Arachnology and Myriapodology, Museum of Nature Hamburg – Zoology, Leibniz Institute for the Analysis of Biodiversity Change, Hamburg, Germany, [4]Department of Biology, University of Hamburg, Hamburg, Germany, [5]Institute of Ecology and Biological Resources, Vietnam Academy of Science and Technology, Caugiay, Hanoi, Vietnam, [6]Graduate University of Science and Technology, Vietnam Academy of Science and Technology, Caugiay, Hanoi, Vietnam, and [7]Corresponding author, e-mail: sloria@amnh.org

Subject Editor: Lauren Esposito

Received 23 February 2022; Editorial decision 12 October 2022

Abstract

Southeast Asia is a hotspot of karst systems in the tropics and many relictual taxa have been documented in caves across the region. The ancient, relictual scorpion family Pseudochactidae Gromov 1998 has a disjunct distribution and includes two hypogean subfamilies from caves in the Khammouan-Phong Nha-Kẻ Bàng Karst in the northern Annamite (Trường Sơn) Mountains of Laos and Vietnam, and one epigean subfamily from Central Asia. A recent revision identified six species in the family; however, how these taxa dispersed and diversified into Southeast Asian cave systems has not been tested. In the present contribution, the phylogeny of Pseudochactidae is reconstructed using three mitochondrial and three nuclear markers and 140 morphological characters, divergence time and ancestral range estimation analyses are conducted, and the evolution of troglomorphic characters is investigated. Results confirm a previous hypothesis that Pseudochactidae originated in Eurasia, most likely near the Tajik block in the Carboniferous, supporting the 'Out of Eurasia' hypothesis and contradicting the 'Eurogondwana' and 'Out of India' hypotheses for the origin of Southeast Asian scorpions. Pseudochactidae dispersed across Southeast Asia after the collision of the Cimmerian continent and Indochina with Eurasia in the Late Jurassic. Colonization of Southeast Asian caves began in the Late Cretaceous and was completed by the Miocene. The onset of aridification in Southeast Asia during the Late Miocene resulted in the extinction of epigean Pseudochactidae, whereas hypogean members of the family likely survived within caves in the limestone massifs of the Annamite Mountains, supporting the 'Climate Relict' hypothesis.

Key words: evolution, biogeography, systematics, morphology

Many caves harbor relictual species (Howarth 1973, 1993; Day and Urich 2000; Deharveng and Bedos 2000; Gibert and Deharveng 2002; Clements et al. 2006; Wessel et al. 2007; Niemiller et al. 2008; Assmann et al. 2010; Juan et al. 2010; Prendini et al. 2010; Deharveng and Bedos 2012). These are mostly ancient taxa with disjunct distributions, often lacking sympatric epigean relatives (Howarth 1973, 1993; Deharveng and Bedos 2000; Gibert and Deharveng 2002; Clements et al. 2006; Wessel et al. 2007; Assmann et al. 2010; Juan et al. 2010; Prendini et al. 2010; Deharveng and Bedos 2012). Many of these taxa exhibit troglomorphic characters (cave adaptations), including loss or reduction of pigmentation, loss of eyes, and elongation of appendages, the result of millions of years of isolation in darkness (Howarth 1993, Deharveng and Bedos 2000, Gibert and Deharveng 2002, Wessel et al. 2007, Niemiller et al. 2008, Assmann et al. 2010, Juan et al. 2010, Prendini et al. 2010, Deharveng and Bedos 2012). The presence of relicts in caves is often explained by the 'Climate Relict' model (Barr 1968, Romero and Green 2005, Niemiller et al. 2008, Juan et al. 2010). According to

this hypothesis, caves acted as 'museums' or refugia when environmental stresses resulted in extinction of their epigean relatives (Barr 1968, Romero and Green 2005, Niemiller et al. 2008, Assmann et al. 2010, Juan et al. 2010), e.g., during Pleistocene glaciation (Barr 1968, Howarth 1973, Romero and Green 2005, Wessel et al. 2007, Assmann et al. 2010, Juan et al. 2010, Bryson et al. 2014). The 'Climate Relict' model has mostly been tested on relictual taxa in temperate caves, which experienced drastic changes in climate during Pleistocene glaciation, but how this aligns with the growing body of evidence demonstrating that tropical caves also house relictual taxa has not been tested (Clements et al. 2006, Deharveng and Bedos 2012).

Southeast Asia is a hotspot for tropical caves (Deharveng and Bedos 2000, 2012; Clements et al. 2006). More than 400,000 km^2 of karst occurs throughout this region in southern China, Indonesia, Laos, Malaysia, Myanmar, the Philippines, Thailand, and Vietnam (Clements et al. 2006). Karst age varies with some cave systems dating to the Cambrian and others to the Quaternary (Clements et al. 2006). The vast differences in karst age can be explained by the region's complex geological history. Southeast Asia is a mosaic of landmasses that rifted from Gondwana and collided with Eurasia at different times (Metcalfe 2011). Given the complex geological history, cave biodiversity in the region is expected to be high. Preliminary surveys ranked Southeast Asian cave systems among the top 10 in terrestrial species richness, but this is probably an underestimate as many cave systems remain poorly studied (Deharveng and Bedos 2012). Surveys of Southeast Asian caves resulted in the discovery of numerous relictual species among various arthropod taxa (e.g., beetles, crabs, millipedes, and scorpions), contradicting hypotheses that tropical caves are unlikely to harbor relicts because surface and cave environments share similar climatic conditions (Deharveng and Bedos 2012). Data from geology and palynology suggest that parts of Southeast Asia experienced drastic changes in precipitation, temperature, and sea level in the early Cenozoic, prior to the Pleistocene, but how this coincides with the timing of evolution among Southeast Asian cave organisms is poorly understood (Hall 1998, Morley 2012). The relictual scorpion family Pseudochactidae Gromov 1998, comprising six species in four genera and three subfamilies (Table 1), has a disjunct distribution with one epigean genus inhabiting arid valleys near the Tajik Depression and north Pamir Mountains in Central Asia, and three hypogean genera restricted to the Khammouan-Phong Nha-Kẻ Bàng Karst in the Annamite (Trường Sơn) Mountains of Southeast Asia (Fig. 1). The two Central Asian species of *Pseudochactas* Gromov 1998 are lapidicolous (Prendini 2001) with an affinity for humid microhabitats, occurring under stones or in mud cracks along riverbanks in semi-arid mountainous areas (Prendini et al. 2006, Soleglad et al. 2012). The three Southeast Asian genera, *Aemngvantom* Prendini et al. 2021, *Troglokhammouanus* Lourenço 2007, and *Vietbocap* Lourenço and Pham 2010, occur below the surface in limestone caves, and were collected along walls, under stones or in rock cracks (Prendini et al. 2021). Although the three Southeast Asian genera are exclusively subterranean, records of their distance from the surface environment vary. *Troglokhammouanus steineri* Lourenço 2007 was collected from the twilight zone of the cave to ca. 3 km from the downstream entrance, whereas *Vietbocap canhi* Lourenço and Pham 2010 and the two species of *Aemngvantom* were recorded from 200 m to 5 km from the cave entrance (Lourenço 2007, 2012, 2017; Lourenço and Pham 2012; Lourenço et al. 2018; Prendini et al. 2021). This variation in ecology among pseudochactid genera is reflected in their morphology. *Pseudochactas* and *Troglokhammouanus* exhibit well-developed ocelli (eyes) and pigmentation, and no noticeable attenuation of the appendages, whereas *Aemngvantom* and *Vietbocap* are markedly troglomorphic, lacking ocelli and pigmentation, and exhibiting marked attenuation of the appendages, among other troglomorphic characters (Prendini et al. 2021).

Prendini et al. (2021) reviewed the list of cavernicolous, troglomorphic, and troglobitic scorpions in light of the revised Schiner-Racovitza system (Trajano 2012, Trajano and Carvalho 2017), which recognizes three ecological categories of cavernicolous organisms based on the location of source populations: 1) troglobites, with exclusively hypogean source populations, although sink populations may be found in epigean habitats; 2) troglophiles, with both hypogean and epigean source populations, allowing gene flow between these environments; and 3) trogloxenes, with source populations in epigean habitats, although sink populations may be present in hypogean habitats. Prendini et al. (2021) followed Volschenk et al. (2008) in recognizing a distinction between scorpions that are cavernicolous, i.e., found in subterranean environments, and those that are troglomorphic, and restricted the definition of troglobitic scorpions to species that are both restricted to cavernicolous (hypogean) habitats and exhibit pronounced troglomorphies. Prendini et al. (2021) presented a key to classify cavernicolous scorpions according to their ecology and morphology. The difference in ecology and morphology among pseudochactid genera led Prendini et al.

Table 1. Subfamilies, genera, and species of the relictual Asian scorpion family Pseudochactidae Gromov 1998 with known countries of occurrence, based on Prendini et al. (2021)

Subfamily	Species	Distribution
Pseudochactinae Gromov 1998	*Pseudochactas mischi* Soleglad et al. 2012	Afghanistan
	Pseudochactas ovchinnikovi Gromov 1998	Tajikistan, Uzbekistan
Troglokhammouaninae Prendini et al. 2021	*Troglokhammouanus steineri* Lourenço 2007	Laos
	= *Troglokhammouanus louisanneorum* Lourenço 2017	
Vietbocapinae Lourenço 2012	*Aemngvantom lao* (Lourenço 2012)	Laos
	Aemngvantom thamnongpaseuam Prendini et al. 2021	Laos
	Vietbocap canhi Lourenço and Pham 2010	Vietnam
	= *Vietbocap thienduongensis* Lourenço and Pham 2012	
	= *Vietbocap auriantiacus* Lourenço et al. 2018	
	= *Vietbocap quinquemilia* Lourenço et al. 2018	

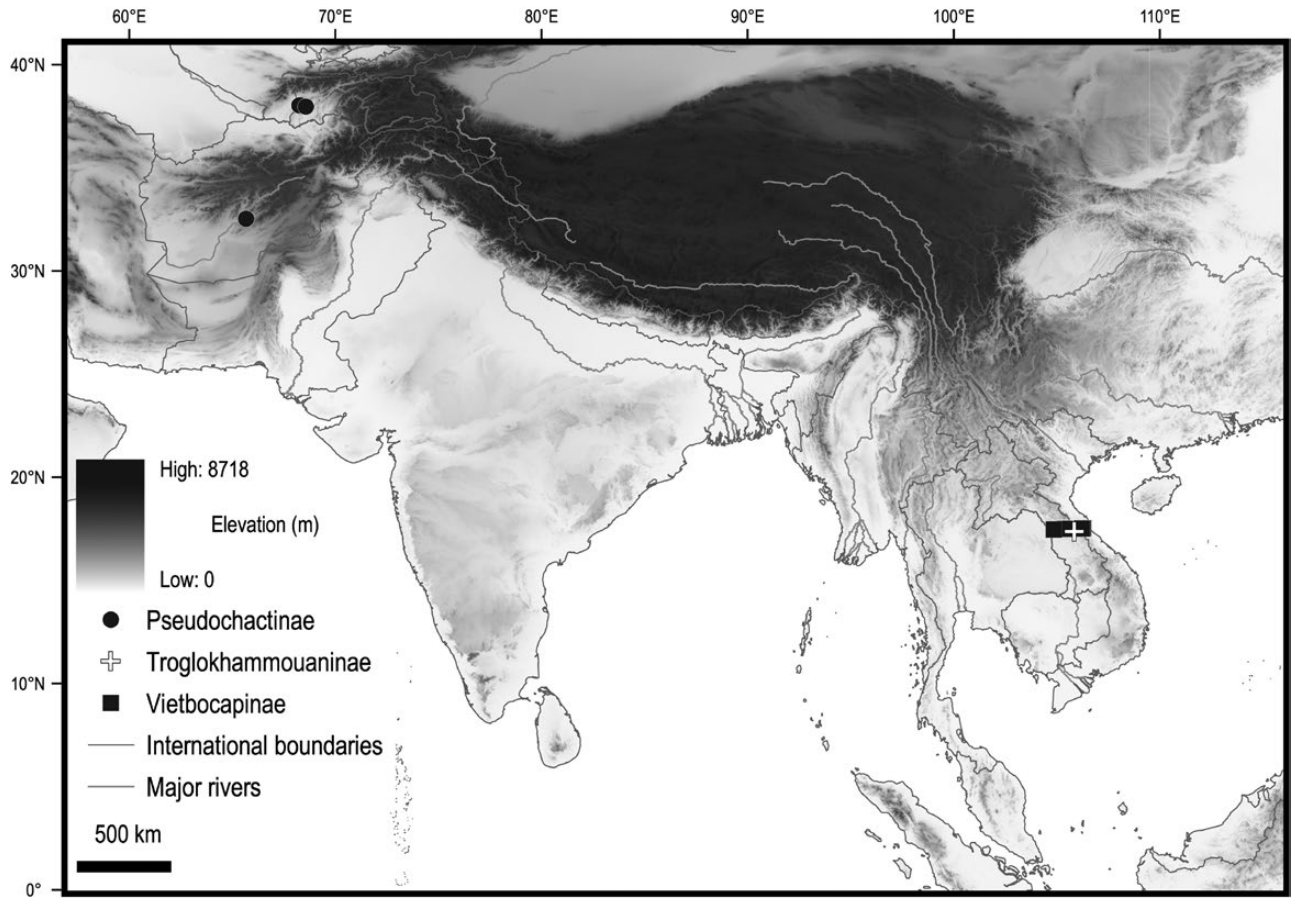

Fig. 1. Distribution of subfamilies of the relictual Asian scorpion family Pseudochactidae Gromov 1998: Pseudochactinae Gromov 1998, Troglokhammouaninae Prendini et al. 2021, Vietbocapinae Lourenço 2007.

(2021) to classify *Pseudochactas* as epigean; *Troglokhammouanus* as a hypogean troglophile; and *Aemngvantom* and *Vietbocap* as hypogean troglobites.

Pseudochactidae are thought to constitute an ancient, relictual lineage that was once more widespread (Soleglad and Fet 2003a, Fet et al. 2004, Prendini et al. 2006, Lourenço 2007). These scorpions have long been considered 'living fossils' due to their unique morphology (Prendini et al. 2006, 2021). Although its phylogenetic position has been extensively debated, Pseudochactidae has consistently been placed basal in the scorpion tree with three alternative hypotheses proposed (Coddington et al. 2004, Prendini et al. 2006, 2021): (1) sister to all Recent (extant) scorpions; (2) sister to Buthidae C.L. Koch 1837; or (3) sister to Chaerilidae Pocock 1893. Fet (2000) suggested Pseudochactidae was closely related to Buthidae or Chaerilidae based on its unique trichobothrial pattern. Lourenço (2000) placed Pseudochactidae in superfamily Chaeriloidea Pocock 1893, implying a close relationship to Chaerilidae. Soleglad and Fet (2001) and Prendini et al. (2006) suggested a close relationship between Pseudochactidae and Buthidae based on trichobothria and other characters, including the reproductive system. Soleglad and Fet (2003a, b) placed Pseudochactidae sister to all Recent scorpions, supporting the original suggestion of Gromov (1998). Recent phylogenomic analyses placed Pseudochactidae and Chaerilidae in a clade sister to Buthidae (Sharma et al. 2015, 2018).

Several biogeographical hypotheses have been proposed to explain the distribution of Pseudochactidae. Soleglad and Fet (2003a) suggested the family originated on Pangaea in the Permian/Triassic but did not speculate on its paleodistribution as no other extant or fossil taxa were known at that time. Fet et al. (2004) hypothesized that the ancestor of Pseudochactidae inhabited littoral forests on islands in the Tethys Sea near the modern-day Tajik Depression during the Cretaceous. These island forests were elevated during mountain uplift in the Tertiary due to the collision of the Indian subcontinent with Eurasia, shielding them from the onset of aridification in the Eocene caused by this collision. Prendini et al. (2006) noted that a more recent divergence was plausible if Pseudochactidae were the sister group of Buthidae. Following the discovery of *T. steineri* in Southeast Asia, Lourenço (2007) suggested that a Pangaean origin for Pseudochactidae in the Permian/Triassic should be reconsidered as the only land connection between Central Asia and Laos was the old Asia core in the Paleozoic. A dated phylogenomic analysis of scorpions, calibrated using fossil scorpions and other arachnids under an autocorrelated lognormal model, proposed an even earlier divergence between Pseudochactidae and Chaerilidae, ca. 329 Ma in the Carboniferous (Sharma et al. 2015), whereas a second analysis, applying an uncorrelated gamma rates model, placed this divergence much later, ca. 198 Ma in the Jurassic (Sharma et al. 2018). Sharma et al. (2018) preferred the results of the autocorrelated lognormal model due to its better performance; however, no biogeographical hypotheses were proposed to explain either of these divergence dates.

Given their unique morphology, disjunct distribution, and variable levels of troglomorphic adaptation, Pseudochactidae represent a model system for comparative analysis of the origins and evolution of cave organisms in Southeast Asia. In the present contribution, the phylogeny of Pseudochactidae is reconstructed using three mitochondrial markers (12S rDNA, 16S rDNA, Cytochrome

Oxidase Subunit I), three nuclear markers (18S rDNA, 28S rDNA, Internal Transcribed Spacer) and 140 morphological characters. Divergence time and ancestral range estimation analyses are conducted and the evolution of troglomorphic characters investigated to understand how this unique, relictual scorpion lineage dispersed and diversified into Southeast Asian cave systems.

Materials and Methods

Taxon Sampling

The ingroup dataset comprised 23 terminals representing seven populations and all six species of Pseudochactidae (Prendini et al. 2021). *Pseudochactas ovchinnikovi* Gromov 1998 was represented by three terminals from a single population in Dikhana Canyon in the Babatag Mountains, Uzbekistan, ca. 3 kilometers from the type locality. *Pseudochactas mischi* Soleglad et al. 2012 was represented by the subadult female holotype from Uruzugan Province, Afghanistan, scored from the original description. *Troglokhammouanus steineri* was represented by five terminals from the type locality, Tham Xe Bang Fai. *Aemngvantom lao* (Lourenço 2007) was represented by four terminals from the type locality, Tham Nam Lot and *Aemngvantom thamnongpaseuam* Prendini et al. 2021 represented by the holotype and only known specimen. *Vietbocap canhi* was represented by nine terminals from two caves: three from the type locality, Tiên Sơn, and six from Thiên Đường, the type locality of its three synonyms, *Vietbocap aurantiacus* Lourenço et al. 2018, *Vietbocap quinquemilia* Lourenço et al. 2018, and *Vietbocap thienduongensis* Lourenço and Pham 2012. All ingroup terminals were represented by both morphological and molecular data except for *P. mischi*, which lacked molecular data. *Chaerilus sejnai* Kovařík 2005, representing Chaerilidae, and *Buthus atlantis* Pocock 1889, representing Buthidae, were included as outgroups. The tree was rooted on *B. atlantis*.

Material Examined

Samples of all species of Pseudochactidae, except *P. mischi*, were collected for the present analyses. Scorpions were collected on the surface at night, or in caves during the day and night using ultraviolet (UV) light detection with portable light emitting diode UV lamps. Specimens were initially preserved in and injected with 95% ethanol. DNA voucher tissues are stored in 95% ethanol at –150°C in the Ambrose Monell Cryocollection for Molecular and Microbial Research (AMCC) at the American Museum of Natural History (AMNH). All other material is stored in 70% ethanol at the AMNH. Additional material examined for the present study is deposited in the following institutions (Supp Appendix 1 [online only]): Alexander V. Gromov Private Collection (AGPC), Almaty, Kazakhstan; Institute of Ecology and Biological Resources (IEBR), Hanoi, Vietnam; Muséum National d'Histoire Naturelle (MNHN), Paris, France; Phong Nha-Kẻ Bàng National Park (PNKB), Bố Trạch, Vietnam.

Mapping

Maps were produced using QGIS ver. 3.14 (http://www.qgis.org), by superimposing locality records and borders, coastlines and rivers from Natural Earth (https://www.naturalearthdata.com) on the global multi-resolution terrain elevation data (GTMTED2010; https://www.usgs.gov/core-science-systems/euros/coastal-changes-and-impacts/gmted2010). Palaeogeographical reconstructions were adapted and modified from Metcalfe (2011) and Morley (2018).

Morphological Dataset

A morphological data matrix comprising 140 characters (Supp Appendices 2 and 3 [online only]) scored for the 25 terminals and incorporating characters from previously published matrices by Lamoral (1980), Stockwell (1989), Prendini (2000), Soleglad and Sissom (2001), Soleglad and Fet (2001, 2003a), and Prendini (2004), augmented by observations in Prendini et al. (2006) was prepared using WinClada ver. 1.00.08 (Nixon 2002), and deposited in Morphobank (http://morphobank.org/permalink/P3916). Characters scored in the matrix included a variety of traits capturing the morphological diversity within Pseudochactidae, most pertaining to the adult stage: coloration and infuscation, 4 (2.9%); chelicerae, 5 (3.6%); carapace, 24 (17.1%); pedipalps, 60 (42.9%); coxosternum, 4 (2.9%); legs, 5 (3.6%); genital operculum, 1 (0.7%); male and female reproductive anatomy, 4 (2.9%); pectines, 12 (8.6%); tergites, 1 (0.7%); sternites, 3 (2.1%); metasoma, 11 (7.9%); telson, 6 (4.3%). The breakdown of morphological characters by character system is as follows: trichobothria, 36 (25.7%); surface macrosculpture, 35 (25%); shape and size, 14 (10%); pectines, 12 (8.6%); surface topography, 7 (5%); ocelli, 7 (5%); pedipalp finger dentition, 6 (4.3%); cheliceral dentition, 5 (3.6%); genitalia, 5 (3.6%); leg setation and spination, 5 (3.6%); color patterns, 4 (2.9%); sutures, 2 (1.4%); respiratory spiracles (stigmata), 1 (0.7%); venom glands, 1 (0.7%). Thirty-four (24%) multistate characters were treated as unordered/nonadditive (Fitch 1971) to avoid a priori character state transformations, and 58 (41%) characters were uninformative.

DNA Extraction, PCR, and Sequencing

DNA extraction was performed using the Qiagen (Germantown, MD) DNEasy Blood and Tissue Kit. Muscle tissue was dissected from the leg, digested in an incubator for 24–56 h and eluted in 50–200 µl of AE buffer, depending on the quality of specimen preservation and amount of tissue. PCR was performed using a combination of reagents, including Taq (Fisher Scientific, Hampton, NH), TopTaq (Qiagen), REDTaq (Sigma-Aldrich, St. Louis, MO) and Illustra PuReTaq Ready-To-Go PCR Beads (GE Healthcare, Chicago, IL), and primers (Table 2). Amplified loci were purified using Agencourt Ampure (Beckman Coulter, Brea, CA) or HighPrep (MagBio Genomics, Gaithersburg, MD). Fluorescent nucleotides were added to amplified loci using the Big Dye Terminator ver. 3.1 cycle sequencing kit (Life Technologies, Carlsbad, CA) and cleaned using Agencourt CleanSEQ (Beckman Coulter) or ethanol precipitation. Purified PCR products were sequenced using an ABI Prism 3730 XL DNA Sequencer (Perkin-Elmer, Melville, NY) at the AMNH Sackler Institute of Comparative Genomics or GeneWiz (South Plainfield, NJ). Loci were sequenced in forward and reverse directions to substantiate results. Forward and reverse strands were assembled and edited into consensus sequences using Sequencher ver. 5.4.6 (Gene Codes, Corporation, Ann Arbor, MI) and deposited in Genbank (https://www.ncbi.nlm.nih.gov/genbank; Table 3).

Molecular Dataset

The molecular dataset comprised three mitochondrial loci, small subunit ribosomal RNA (12S rDNA, hereafter '12S'), large subunit ribosomal RNA (16S rDNA, hereafter '16S') and Cytochrome *c* Oxidase Subunit I (hereafter 'COI'), and three nuclear loci, large subunit ribosomal RNA (28S rDNA, hereafter '28S'), small subunit ribosomal RNA (18S rDNA) and Internal Transcribed Spacer (ITS2, hereafter '18S-ITS2'). The 12S, 16S and COI loci were completely sequenced for all terminals

Table 2. Primers used to generate DNA sequences from the 18S rDNA (18S), Internal Transcribed Spacer (ITS2), 28S rDNA (28S), 12S rDNA (12S), 16S rDNA (16S), and Cytochrome *c* Oxidase Subunit I (COI) loci for phylogenetic analysis of the relictual Asian scorpion family Pseudochactidae Gromov 1998

Primer	Sequence (5′–3′)	Other Name	Citation
18S			
18S1F	TACCTGGTTGATCCTGCCAGTAG		Giribet et al. (1996)
18S5R	CTTGGCAAATGCTTTCGC		Giribet et al. (1996)
18S3F	GTTCGATTCCGGAGAGGGA		Giribet et al. (1996)
18Sbi	GAGTCTCGTTCGTTATCGGA		Wheeler et al. (1993)
18Sa2.0	ATGGTTGCAAAGCTGAAAC		Wheeler et al. (1993)
18S9R	GATCCTTCCGCAGGTTCACCTAC		Giribet et al. (1996)
ITS2			
CAS18sF1	TACACACCGCCCGTCGCTACTA		Ji et al. (2003)
CAS5p8sB1d	ATGTGCGTTCRAAATGTCGATGTTCA		Ji et al. (2003)
CAS5p8dFc	TGAACATCGACATTTYGAACGCACAT		Ji et al. (2003)
CAS28sB1d-spid	TTCTTTTCCTCCGCTTATTTATATGCTTAA		Ji et al. (2003)
28S			
28S-RD1A	CCCSCGTAAYTTAGGCATAT		Nunn et al. (1996)
28S-RD4B	CCTTGGTCCGTGTTTCAAGAC		Nunn et al. (1996)
28S-RD3.2A	AGTACGTGAAACCGTTCASGGGT		Nunn et al. (1996)
28SB	TCGGAAGGAACCAGCTAC	D3B	Nunn et al. (1996)
28SA	GACCCGTCTTGAAGCACG	D3A	Nunn et al. (1996)
28SBout	CCCACAGCGCCAGTTCTGCTTACC		Prendini et al. (2005)
28S-RD4.5A	AAGTTTCCCTCAGGATAGCTG		Nunn et al. (1996)
28S-RD4.8A	ACCTATTCTCAAACTTTAAATGG		Nunn et al. (1996)
28S-RD7B1	GACTTCCCTTACCTACAT		Nunn et al. (1996)
12S			
12Sai	AAACTAGGATTAGATACCCTATTAT	SR-N-14588	Kocher et al. (1989)
12Sbi	AAGAGCGACGGGCGATGTGT	SR-J-14233	Kocher et al. (1989)
16S			
16Sa	CTCCGGTTTGAACTCAGATCA	LR-N-13398	Simon et al. (1991)
16Sb	CGCCTGTTTATCAAAAACAT	LR-J-12887	Simon et al. (1991)
COI			
LCO	GGTCAACAAATCATAAAGATATTGG	LCO1490-J-1514	Folmer et al. (1994)
COImodF	ATCATAAGGATATTGGGACTATGT		Bryson et al. (2013)
Nancy	CCCGGTAAAATTAAAATATAAACTTC	C1-N-2191	Harrison et al. (1987)
HCO	TAAACTTCAGGGTGACCAAAAAATCA	HCO2198-N-2175	Folmer et al. (1994)
HCOoutout	GTAAATATATGRTGDGCTC		Prendini et al. (2003)
CruzR	CATACCCAAAGARCCAAAAGG		Valdez-Cruz et al. (2004)
CI-J01718-spider	GGNGGATTTGGAAATTGRTTRGTTCC	C1-J-1718	Harrison et al. (1987)
CI-N02776-spider	GGATAATCAGAATANCGNCGAGG		Simon et al. (1991)
EXTB	CCTATTGAWARAACATARTGAAAATG		Simon et al. (1991)
EXTA	GAAGTTTATATTTTAATTTTACCTGG		Simon et al. (1991)

(Table 3). The 18S-ITS2 locus was partially sequenced for one terminal of *P. ovchinnikovi*, and completely sequenced for one terminal each of *T. steineri*, *A. lao* and *A. thamnongpaseuam*, and two terminals of *V. canhi*, representing both cave populations, Tiên Sơn and Thiên Đường. The 28S locus was completely sequenced for all except three terminals of *P. ovchinnikovi*, which were partially sequenced.

Sequence length was calculated using the 'gc_calculator.py' script in Biopython (Cock et al. 2009). Sequences of the COI locus were length invariant (1078 bp). The number of haplotypes was calculated using DNACollapser (https://users-birc.au.dk/palle/php/fabox/dnacollapser.php) in FaBox (Villesen 2007). The 18S, 16S, and COI loci each comprised twelve haplotypes, followed by 12S with ten haplotypes and 28S with eight.

Table 3. Tissue samples and GenBank accession codes for DNA sequences used for phylogenetic analysis of the relictual Asian scorpion family Pseudochactidae Gromov 1998

Species	AMCC	Country	18S/*18S-ITS2*	28S	12S	16S	COI
Buthidae C.L. Koch 1837							
Buthus atlantis Pocock 1889	LP 3489	Morocco	MZ041717	MZ041765	MZ041686	MZ041735	MZ032877
Chaerilidae Pocock 1893							
Chaerilus sejnai Kovařík 2005	LP 3685	Malaysia	MZ041718	MZ041766	MZ041687	MZ041736	MZ032878
Pseudochactidae Gromov 1998							
Pseudochactas ovchinnikovi Gromov 1998	LP 2303A	Uzbekistan	*MZ041759**	MZ041767*	MZ041688	MZ041737	MZ032879
	LP 2303B		MZ041719	MZ041768*	MZ041689	MZ041738	MZ032880
	LP 2303C		MZ041720	MZ041769*	MZ041690	MZ041739	MZ032881
Troglokhammouanus steineri Lourenço 2007	LP 11271	Laos	*MZ041760*	MZ041770	MZ041691	MZ041740	MZ032882
	LP 11272		MZ041721	MZ041771	MZ041692	MZ041741	MZ032883
	LP 11273		MZ041722	MZ041772	MZ041693	MZ041742	MZ032884
	LP 11274		MZ041723	MZ041773	MZ041694	MZ041743	MZ032885
	LP 11275		MZ041724	MZ041774	MZ041695	MZ041744	MZ032886
Aemngvantom lao (Lourenço 2012)	LP 11349		*MZ041761*	MZ041775	MZ041696	MZ041745	MZ032887
	LP 11350		MZ041725	MZ041776	MZ041697	MZ041746	MZ032888
	LP 15570		MZ041726	MZ041777	MZ041698	MZ041747	MZ032889
	LP 15591		MZ041727	MZ041778	MZ041699	MZ041748	MZ032890
Aemngvantom thamnongpaseuam Prendini et al. 2021	LP 11351		*MZ041762*	MZ041779	MZ041700	MZ041749	MZ032891
Vietbocap canhi Lourenço and Pham 2010	LP 11270	Vietnam	*MZ041763*	MZ041780	MZ041701	MZ041750	MZ032892
	LP 16750		MZ041733	MZ041787	MZ041708	MZ041757	MZ032899
	LP 16751		MZ041734	MZ041788	MZ041709	MZ041758	MZ032900
	LP 11348		MZ041728	MZ041781	MZ041702	MZ041751	MZ032893
	LP 16533		*MZ041764*	MZ041782	MZ041703	MZ041752	MZ032894
	LP 16534		MZ041729	MZ041783	MZ041704	MZ041753	MZ032895
	LP 16535		MZ041730	MZ041784	MZ041705	MZ041754	MZ032896
	LP 16536		MZ041731	MZ041785	MZ041706	MZ041755	MZ032897
	LP 16537		MZ041732	MZ041786	MZ041707	MZ041756	MZ032898

Sequences of the 18S rDNA (18S), Internal Transcribed Spacer (ITS2), 28S rDNA (28S), 12S rDNA (12S), 16S rDNA (16S), and Cytochrome *c* Oxidase Subunit I (COI) loci. Tissue samples deposited in the Ambrose Monell Collection for Molecular and Microbial Research (AMCC) at the American Museum of Natural History, New York. Asterisks denote partial sequences. Italicized accession codes represent the 18S-ITS locus whereas all others represent the 18S locus only.

Sequence Alignment and Statistics

Sequences were aligned using MAFFT ver. 7.471 (Katoh et al. 2002, Katoh and Standley 2013) on the CIPRES Science Gateway (https://www.phylo.org; Miller et al. 2010). The COI locus was aligned using the G-INS-i alignment strategy, whereas the 12S, 16S, 18S-ITS2, and 28S loci were aligned using the Q-INS-i alignment strategy, which considers the secondary structure of RNA. Aligned sequences were checked in Mesquite (Maddison and Maddison 2019) and concatenated in a text editor to produce a molecular matrix of 8608 nucleotide base-pairs (bp). The aligned sequence length for each locus was as follows: 371 bp for 12S, 498 bp for 16S, 1078 bp for COI, 2245 bp for 28S, 4416 bp for 18S-ITS2. Mean nucleotide composition and number of variable positions (VP) and parsimony informative (PI) sites were calculated for each locus and the concatenated molecular dataset using MEGA ver. 10.1.7 (Tamura et al. 2004, 2013; Kumar et al. 2018; Stetcher et al. 2020). Mean GC content was higher (Table 4) for the nuclear 28S (59%) and 18S-ITS2 (52%) loci than for the mitochondrial COI (37%), 12S (28%) and 16S (28%) loci, whereas the proportion of variable positions (VP) and parsimony informative sites (PI) were higher for the mitochondrial 12S (58% and 41% for VP and PI, respectively), 16S (49% and 34%) and COI (36% and 30%) loci than for the nuclear, 18S-ITS2 (22% and 10%) and 28S (19% and 8%) loci. The third codon position of the COI locus was found to have the highest proportion of variable (27%) and parsimony informative sites (24%), followed by the first (7% and 5%) and second (2% and 1%) codon positions.

Phylogenetic Analyses

The morphological and molecular datasets were analyzed separately and simultaneously. Parsimony analyses were performed in TNT (Goloboff 1993; Goloboff et al. 2003, 2008; Goloboff and Catalano 2016). Analyses were performed with equal weighting and implied weighting with five values for the concavity constant (k = 1, 3, 10, 60, 100). Gaps were treated as missing data and uninformative characters deactivated in all analyses. A script from Dimitrov et al. (2012, 2013), modified by Santibáñez-López et al. (2014), which includes mixed sectorial search, tree drifting, and tree fusing, was used to

search for the most parsimonious tree. Bootstrap sampling with replacement for 1,000 iterations was used to evaluate nodal support in the analysis with equal weighting. For comparison, a jackknife analysis was used to evaluate nodal support in the analysis with equal weighting using a script from Dimitrov et al. (2012). Tree length, fit, retention indices, adjusted homoplasy, and consistency indices were also calculated (Table 5).

Maximum likelihood (ML) analyses were performed in RAxML-HPC ver. 8 on CIPRES (Stamatakis 2006, 2014). Each molecular locus and the morphological dataset were treated as a single partition. A GTRGAMMA model was applied to each partition in the separate molecular analyses whereas, in the simultaneous analyses, a MULTIGAMMA model was applied to the molecular partitions, and an MK model (Lewis 2001) to the morphological partition. An ascertainment bias using Lewis correction was applied to the morphological partition in one simultaneous analysis. The best scoring tree was searched for using a rapid bootstrap analysis (-f a) with 1,000 bootstrap iterations.

Analyses with Bayesian inference (BI) were performed using MrBayes ver. 3.2.7a (Ronquist and Huelsenbeck 2003) on CIPRES. As in the ML analyses, the molecular dataset was partitioned by loci and a separate partition was applied to the morphological dataset. Twenty-four nucleotide substitution models were tested for each locus using jModelTest2 ver. 2.1.6 (Guidon and Gascuel 2003, Darriba et al. 2012) on CIPRES and the best model chosen using the Akaike Information Criterion (AIC) or the corrected

Table 4. Statistics for nuclear loci, 18S rDNA (18S), Internal Transcribed Spacer (ITS2), and 28S rDNA, and mitochondrial loci, 12S rDNA, 16S rDNA and Cytochrome *c* Oxidase subunit I (COI) in phylogenetic analysis of the relictual Asian scorpion family Pseudochactidae Gromov 1998: number of haplotypes; length (nucleotide base-pairs) of unaligned and aligned sequences; average percent GC content; number and percent of variable positions (VP); number and percent of parsimony-informative sites (PI); and model selected for Bayesian Inference using the Akaike Information Criterion (AIC)

Locus	Haplotypes	Unaligned	Aligned	GC	VP	PI	AIC
18S-ITS2	12	1760–3926	4416	52	978 (22)	446 (10)	GTR + G
28S	8	1329–2193	2245	59	418 (19)	172 (8)	GTR + I + G
12S	10	337–344	371	28	217 (58)	151 (41)	GTR + G
16S	12	475–483	498	28	244 (49)	171 (34)	GTR + I + G
COI	12	1078	1078	37	388 (36)	319 (30)	GTR + I + G
COI 1st Codon	-	-	359	45	79 (7)	49 (5)	-
COI 2nd Codon	-	-	359	42	22 (2)	8 (1)	-
COI 3rd Codon	-	-	360	25	287 (27)	262 (24)	-

Table 5. Tree statistics for phylogenetic analysis of the relictual Asian scorpion family Pseudochactidae Gromov 1998. Length, consistency index (CI), retention index (RI), Fit and adjusted homoplasy (AH) of most parsimonious trees (MPTs) obtained by separate and simultaneous analyses of the morphological and molecular (concatenated aligned DNA sequences of three nuclear and three mitochondrial gene markers) datasets, with equal weighting (EW) and implied weighting (IW), applying five values of the concavity constant (*k*)

	Weight	MPTs	Length	CI	RI	Fit	AH
Morphology	IW: $k = 1$	1	162	0.784	0.930	66.5	15.5
	IW: $k = 3$	1	162	0.784	0.930	73.85	8.15
	IW: $k = 10$	1	162	0.784	0.930	78.91	3.09
	IW: $k = 60$	1	162	0.784	0.930	81.43	0.57
	IW: $k = 100$	1	162	0.784	0.930	81.65	0.35
	EW	1	162	0.784	0.930	73.85	8.15
Molecular	IW: $k = 1$	1	2258	0.769	0.899	1029.5	229.5
	IW: $k = 3$	1	2258	0.769	0.899	1138.1	120.9
	IW: $k = 10$	1	2258	0.769	0.899	1213.02	45.98
	IW: $k = 60$	1	2258	0.769	0.899	1250.49	8.51
	IW: $k = 100$	1	2258	0.769	0.899	1253.85	5.15
	EW	1	2258	0.769	0.899	1138.1	120.9
Simultaneous	IW: $k = 1$	1	2420	0.770	0.901	1096	245
	IW: $k = 3$	1	2420	0.770	0.901	1211.95	129.05
	IW: $k = 10$	1	2420	0.770	0.901	1291.93	49.07
	IW: $k = 60$	1	2420	0.770	0.901	1331.92	9.08
	IW: $k = 100$	1	2420	0.770	0.901	1335.51	5.49
	EW	1	2420	0.770	0.901	1211.95	129.05

AIC (AICc). A Lewis MK model was applied to the morphological partition. The following parameters were implemented: mcmc; nchains = 4; print frequency = 1,000; sample frequency = 1,000; diagnosing frequency = 1,000; burnin = 0.25; savebrlens = yes. Analyses ran for 25 million generations to ensure that the standard deviation of split frequencies of the two chains was below 0.01. All input data files are publicly available in DRYAD (doi: 10.5061/dryad.4xgxd25d8).

Divergence Time Estimation

A divergence time estimation analysis was performed on the molecular and morphological datasets using BEAST 2 ver. 2.6.4 (Bouckaert et al. 2019). The molecular dataset was trimmed to one terminal per species and sequences for all loci aligned in MAFFT, as described above. The morphological dataset was also trimmed to one terminal per species and, to capture morphological variation, characters fused among individuals of each species. xml files were prepared in Beauti (Drummond et al. 2012). The morphological dataset was imported without recording variable characters only and subsequently divided into four partitions, and the molecular dataset was partitioned by the five loci (18S-ITS2, 28S, 12S, 16S, COI), resulting in nine partitions. The tree model was linked and site models were unlinked across all partitions, whereas clock models were unlinked across molecular partitions but linked across morphological partitions. Vietbocapinae Lourenço 2007 was constrained to be monophyletic and Chaerilidae was constrained to be the sister group of Pseudochactidae, with *B. atlantis* as the outgroup. A simple HKY model was applied to all molecular partitions with the gamma category count set to four and base frequencies set to empirical. A Lewis MK model was applied to all morphological partitions, with the gamma rate category count set to four and the gamma shape parameter linked across all morphological partitions by manually editing the xml files. A Yule model was implemented for the tree prior and a gamma distribution was applied to the birthrate prior with $\alpha = 0.001$ and $\beta = 1{,}000$. A relaxed lognormal molecular clock was set for the clock model of each partition. Clock priors were set to the default uniform distribution for the morphological partition and a uniform distribution applied to all loci with the following parameters: ucld.min = 0.002, ucld.max = 0.5 for 12S, 16S, COI, and ucld.min = 0.0001, ucld.max = 0.01 for 28S and 18S-ITS2 (Bryson et al. 2013, Loria and Prendini 2020). The tree was calibrated using the estimated divergence date (ca. 329 Ma) between Chaerilidae and Pseudochactidae, following Sharma et al. (2018) who estimated divergence dates among scorpion families using genomic data from 30 taxa and five fossil calibrations under an autocorrelated lognormal model. Although Sharma et al. (2018) recovered a younger divergence date between Chaerilidae and Pseudochactidae (198 Ma) using an uncorrelated gamma rates model, the results of the autocorrelated lognormal model were preferred as simulations and empirical studies (Lepage et al. 2007, Paradis 2013) have shown this model to be superior compared to the uncorrelated gamma rates model (Sharma et al. 2018). A normal distribution was applied to the divergence between Chaerilidae and Pseudochactidae with $\mu = 329$ Ma and sigma $\sigma = 10$ Ma. Two MCMC runs were performed using different starting seed values, to test for convergence. Each analysis ran for 500 million generations on CIPRES with tracelog = 10,000 and treelog = 10,000. Convergence was assessed using Tracer ver. 1.7.2 (Rambaut et al. 2014). After multiple initial runs with poor convergence, operators were adjusted for several priors following recommendations in BEAST 2 output files. In the final analysis, trees were combined in LogCombiner by removing 25% of burnin from each run, and the maximum clade credibility tree selected using TreeAnnotator. Input data files are publicly available in DRYAD (doi: 10.5061/dryad.4xgxd25d8).

Ancestral Range Estimation

Ancestral range reconstruction for the time-calibrated phylogenetic tree was performed using the *BioGeoBEARS* ver. 1.1.1 package (Matzke 2013a, b, 2014) in R ver. 4.0.0 (R Core Team 2020). Outgroup taxa were removed as the ancestral areas of Buthidae and Chaerilidae are unknown. Four biogeographical areas were defined: 1) Tajik and north Pamir blocks (northern Afghanistan and southern China, Uzbekistan and Tajikistan), 2) western Cimmeria, including only the Farah, Helmand, south-central Pamirs, Karakorum and Band-e Bayan blocks (present-day Afghanistan and northern Pakistan), 3) northwest Annamite Mountains (Khammouan Karst, Laos), 4) northeast Annamite Mountains (Phong Nha-Kẻ Bàng Karst, Vietnam). Each species was scored present or absent in each area and the maximum number of areas a species was allowed to occupy, set to four. Three models were compared in the analysis: Dispersal-Extinction-Cladogenesis, DIVALIKE and BAYAREALIKE, and an additional parameter (*j*), which allows a founder speciation event, was tested for each. Statistical analyses were performed to determine the likelihood of the dataset for each model, and the model with lowest likelihood was selected as the best fit. Input data files and R scripts are publicly available in DRYAD (doi: 10.5061/dryad.4xgxd25d8).

Troglomorphism and Ancestral State Reconstruction

Ancestral state reconstruction was performed on fifteen morphological characters pertaining to cave adaptation to understand the evolution of troglomorphism in Pseudochactidae. Species isolated in cave environments are known to exhibit several troglomorphic characters, including reduced (pale) pigmentation, reduced ocelli (eyes), and elongated appendages or enhanced non-visual sensory organs to compensate for the absence of visual systems and/or assist with locomotion and predation in subterranean environments (Prendini et al. 2021). A morphological matrix was prepared including the following characters (Table 6): presence/absence of pigmentation on the carapace, tergites, and legs, 4 (26.7%); presence/absence of the median and lateral ocelli, 6 (40%); attenuation of the pedipalps, 1 (6.7%); presence/absence of spurs, spinules and macrosetae on the legs, 3 (20%); and shape of the telson vesicle, 1 (6.7%). Following Ojanguren-Affilastro et al. (2016) and Loria and Prendini (2021), character states were scored and summed for each species and scaled to create an ensemble index of cave adaptation with values ranging from 1 to 10. Scaled scores were treated as discrete characters and ancestral state reconstruction performed on the divergence time tree applying three models, all rates equal (ER), symmetric (SYM), and all rates different (ARD), using the *phytools* package (Revell 2012) in R. Absolute values of log-likelihoods were mapped onto nodes and chi-squared tests performed to select the best-fitting reconstruction. Scaled scores were then classified on a troglomorphic continuum from epigean species (score = 1) to hypogean species progressively more adapted to caves (score = 10): 1–5 (epigean), 6–8 (troglophilic), 9–10 (troglobitic). Input data files and R scripts are publicly available in DRYAD (doi: 10.5061/dryad.4xgxd25d8).

Table 6. Distribution of 15 morphological characters used to reconstruct ancestral ecomorphotypes for the relictual Asian scorpion family Pseudochactidae Gromov 1998

Species	Character Matrix	Score
Buthus atlantis	11000 00000 20000	[2]
Chaerilus sejnai	00000 10011 01100	[2.5]
Pseudochactas mischi	11110 11011 10110	[5.5]
Pseudochactas ovchinnikovi	10000 11011 10110	[4]
Troglokhammouanus steineri	00011 11111 10211	[6]
Aemngvantom lao	11112 11211 21211	[9.5]
Aemngvantom thamnongpaseuam	11112 11211 21211	[9.5]
Vietbocap canhi	11112 11211 31211	[10]

Terminology follows Supp Appendix 3 (online only).

1. Tegument base coloration: brown [0]; yellowish [+1].
2. Carapace infuscation: present [0]; absent (immaculate) [+1].
3. Mesosomal tergites infuscation: present (entirely or partially infuscate) [0]; absent (immaculate) [+1].
4. Legs, femoral and patellar prolateral surfaces infuscation: present [0]; absent (immaculate) [+1]
5. Carapace median ocelli: present, pigmented [0]; present, not pigmented in some individuals [+1]; absent [+2].
6. Carapace lateral ocelli, anterolateral major (ALMa) ocelli: present, pigmented [0]; absent [+1].
7. Carapace lateral ocelli, mediolateral major (MLMa) ocelli: present, pigmented [0]; absent [+1].
8. Carapace lateral ocelli, posterolateral major (PLMa) ocelli: present, pigmented [0]; present, not pigmented in some individuals [+1]; absent [+2].
9. Carapace lateral ocelli, anterodorsal minor (ADMi) ocelli: present, pigmented [0]; absent [+1].
10. Carapace lateral ocelli, posterodorsal minor (PDMi) ocelli: present, pigmented [0]; absent [+1].
11. Pedipalps shape: segments very robust, chela manus incrassate [0]; segments moderately robust, chela manus globose [+1]; segments gracile, chela manus globose [+2]; segments gracile, chela manus elongate [+3].
12. Legs III and IV, tibial spurs: present [0]; absent [+1].
13. Legs I–IV basitarsi and telotarsi, spinules: absent [0]; short [+1]; long [+2].
14. Legs I–IV telotarsi, ventrosubmedian rows of (socketed) macrosetae: present [0]; absent [+1].
15. Telson vesicle, shape: globose [0]; elongate [+1].

Results

Phylogenetic Analyses

Separate analyses of the molecular dataset with BI, ML, and parsimony with equal and implied weighting (k = 1, 3, 10, 60, 100) recovered all subfamilies, genera, species, and both cave populations of *V. canhi* (Tiên Sơn and Thiên Đường) as monophyletic (Table 7). Posterior probability and bootstrap support values were high for all taxa and clades in the analyses with BI (≥0.99), ML (≥96%) and parsimony with equal weighting (≥ 100%). Jackknife support values were generally high (≥85%) in the parsimony analysis with equal weighting except for intraspecific relationships among *A. lao*, *V. canhi*, *P. ovchinnikovi* and *T. steineri* (>10%).

Separate analyses of the morphological dataset with parsimony and equal or implied weighting (k = 1, 3, 10, 60, 100) were congruent with separate analyses of the molecular dataset. Bootstrap support values in separate analyses of the morphological dataset with parsimony and equal weighting were mostly high (≥99%) except for the following taxa and clades, which received low to moderate support: *P. ovchinnikovi* (68%), *Pseudochactas* (58%), the Thiên Đường population of *V. canhi* (28%), and the Southeast Asian clade comprising Troglokhammouaninae Prendini et al. 2021 and Vietbocapinae. Jackknife support values in the separate analysis of the morphological dataset with parsimony and equal weighting were also high (≥ 79%) except for *P. ovchinnikovi* (70%), *A. lao* (66%), the Thiên Đường population of *V. canhi* (54%) and intraspecific relationships among *A. lao*, *P. ovchinnikovi*, *T. steineri* and *V. canhi* (> 10%).

Simultaneous analyses of the morphological and molecular datasets with BI, ML and parsimony with equal or implied weighting (k = 1, 3, 10, 60, 100) were congruent with separate analyses of the molecular and morphological datasets and recovered all subfamilies, genera, species and populations as monophyletic. Posterior probability and bootstrap support values obtained by simultaneous analysis of the morphological and molecular datasets were generally high in the analyses with BI (≥0.99), ML (≥99%), and parsimony with equal weighting (≥ 100%). Jackknife support values were also high (≥83%) in the parsimony analysis with equal weighting except for the following, which received low to moderate support: *P. ovchinnikovi* (63%) and intraspecific relationships among *V. canhi* and *T. steineri* (>10%). Unambiguous morphological synapomorphies were optimized on the ML tree obtained by simultaneous analysis of the morphological and molecular datasets, the preferred phylogeny (Fig. 2). Output files from phylogenetic analyses are publicly available in figshare: 10.6084/m9.figshare.21067528.

Divergence Time Estimation

The two independent MCMC chains from the divergence time estimation analysis in BEAST 2 produced 50,000 trees each. Examination of the logfiles in Tracer indicated convergence of the two chains. ESS values were above 200 for all parameters in both MCMC runs except with the gamma shape parameter, which was below 100 in one run, and the standard deviation of the morphological matrix, which was below 100 in both runs. The maximum clade credibility tree (Fig. 3), produced after removing 25% burnin and combining trees from both chains, was a fully dichotomous phylogeny. In accordance with the calibrated node, the divergence between Chaerilidae and Pseudochactidae occurred in the Late Mississippian of the Carboniferous (327 Ma, 95% HPD: 318–337 Ma). The divergence between Pseudochactinae Gromov 1998 and the Southeast Asian subfamilies, Troglokhammouaninae and Vietbocapinae, occurred in the Jurassic (157 Ma, 95% HPD: 115–184 Ma), whereas the divergence between Troglokhammouaninae and Vietbocapinae occurred in the Early Cretaceous (116 Ma, 95% HPD: 90–142 Ma). Within Vietbocapinae, the divergence between *Aemngvantom* and *Vietbocap* occurred in the Late Cretaceous (97 Ma, 95% HPD: 74–121 Ma), and within Pseudochactinae, the divergence between *P. mischi* and *P. ovchinnikovi* also occurred in the Late Cretaceous (79 Ma, 95% HPD: 4–109 Ma). The divergence between the two troglobitic species of *Aemngvantom*, *A. lao* and *A. thamnongpaseuam*, occurred in the Miocene (18 Ma, 95% HPD: 11–25 Ma). Output files from the divergence time estimation analysis are publicly available in figshare: 10.6084/m9.figshare.21067528.

Ancestral Range Estimation

Based on comparison of the AIC, AICc, and log-likelihood values from the ancestral range estimation analysis in *BioGeoBEARS*, the DIVALIKE + j model had the lowest log-likelihood and was the most appropriate model for the dataset (Table 8; Fig. 4). According to this model, by the Late Jurassic, the ancestral range of Pseudochactidae comprised western Cimmeria and

Table 7. Nodal stability for selected clades and taxa in phylogenetic analysis of the relictual Asian scorpion family Pseudochactidae Gromov 1998, based on separate and simultaneous (Simul) analyses of the morphological (Mor) and molecular (concatenated aligned DNA sequences of three nuclear and three mitochondrial gene markers, Mol) datasets using maximum likelihood (ML) including ascertainment bias with Lewis correction (ML-ASC), Bayesian inference (BI) and parsimony with equal weighting (EW) and implied weighting (IW) with five values of the concavity constant (*k*)

		Pseudochactidae	(P (V + T))	(V + T)	V	*Pseudochactas*	*Aemngvantom*	*T. steineri*	*P. ovchinnikovi*	*A. lao*	*V. canhi*	*V. canhi* (Thiên Đường)	*V. canhi* (Tiên Sơn)
Simul	ML	■	■	■	■	■	■	■	■	■	■	■	■
	ML-ASC	■	■	■	■	■	■	■	■	■	■	■	■
	BI	■	■	■	■	■	■	■	■	■	■	■	■
	EW	■	■	■	■	■	■	■	■	■	■	■	■
	IW (*k* = 1)	■	■	■	■	■	■	■	■	■	■	■	■
	IW (*k* = 3)	■	■	■	■	■	■	■	■	■	■	■	■
	IW (*k* = 10)	■	■	■	■	■	■	■	■	■	■	■	■
	IW (*k* = 60)	■	■	■	■	■	■	■	■	■	■	■	■
	IW (*k* = 100)	■	■	■	■	■	■	■	■	■	■	■	■
Mol	ML	■	■	■	■		■	■	■	■	■	■	■
	BI	■	■	■	■		■	■	■	■	■	■	■
	EW	■	■	■	■		■	■	■	■	■	■	■
	IW (*k* = 1)	■	■	■	■		■	■	■	■	■	■	■
	IW (*k* = 3)	■	■	■	■		■	■	■	■	■	■	■
	IW (*k* = 10)	■	■	■	■		■	■	■	■	■	■	■
	IW (*k* = 60)	■	■	■	■		■	■	■	■	■	■	■
	IW (*k* = 100)	■	■	■	■		■	■	■	■	■	■	■
Mor	EW	■	■	■	■	■	■	■	■	■	■	■	■
	IW (*k* = 1)	■	■	■	■	■	■	■	■	■	■	■	■
	IW (*k* = 3)	■	■	■	■	■	■	■	■	■	■	■	■
	IW (*k* = 10)	■	■	■	■	■	■	■	■	■	■	■	■
	IW (*k* = 60)	■	■	■	■	■	■	■	■	■	■	■	■
	IW (*k* = 100)	■	■	■	■	■	■	■	■	■	■	■	■

Highlighted black cells indicate clades and taxa recovered under the various parameters applied (classification follows Prendini et al. 2021). Highlighted grey cells inapplicable due to missing data. Abbreviations as follows: P: Pseudochactinae; T: Troglokhammouaninae; V: Vietbocapinae.

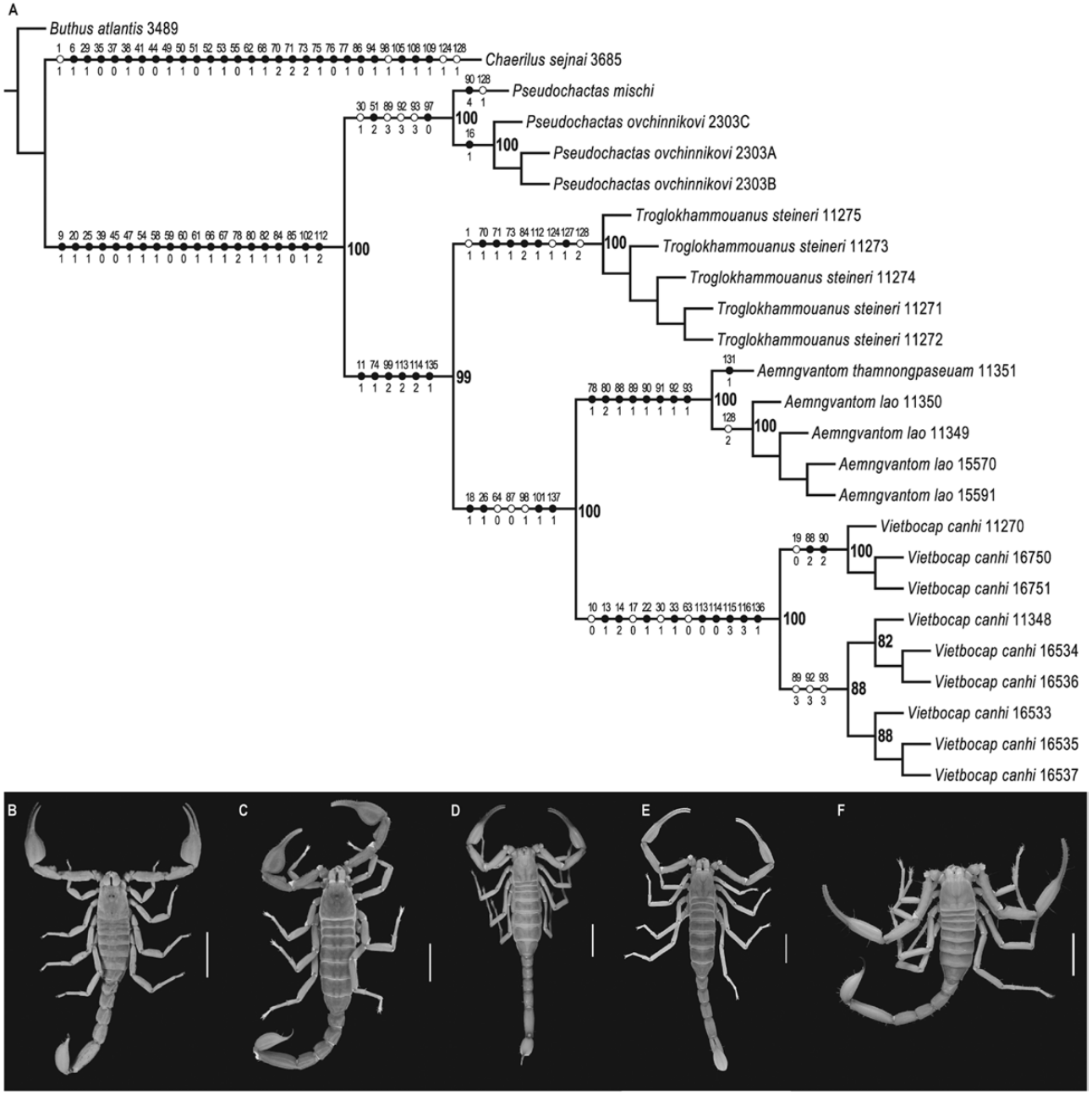

Fig. 2. Preferred phylogeny (**A**) of the relictual Asian scorpion family Pseudochactidae Gromov 1998 based on simultaneous analysis of the morphological and molecular datasets using Maximum Likelihood with unambiguous morphological synapomorphies optimized. Black circles indicate uniquely derived apomorphic states, white circles, parallel deviations of apomorphic states. Numbers above circles indicate characters, numbers below indicate states (Supp Appendices 2 and 3 [online only]). Dorsal habitus (**B–F**) of five species of Pseudochactidae: *Pseudochactas ovchinnikovi* Gromov 1998, Babatag, Uzbekistan, ♀ (AMNH) (**B**), *Troglokhammouanus steineri* Lourenço 2007, Tham Xe Bang Fai, Laos, ♀ (AMNH) (**C**), *Aemngvantom thamnongpaseuam* Prendini et al. 2021, Tham Nong Pa Seuam, Laos, holotype ♀ (AMNH) (**D**), *Aemngvantom lao* (Lourenço 2012), Tham Nam Lot (Lod), Laos, ♀ (AMNH) (**E**) and *Vietbocap canhi* Lourenço and Pham 2010, Tiên Sơn, Vietnam, ♀ (AMNH) (**F**). Scale bars = 10 mm.

the Tajik and north Pamir blocks, and the northwest Annamite Mountains (Khammouan Karst). Western Cimmeria and the Tajik and north Pamir blocks were recovered as ancestral for the Central Asian subfamily, Pseudochactinae, and the northwest Annamite Mountains (Khammouan Karst) as ancestral for the Southeast Asian subfamilies, Troglokhammouaninae and Vietbocapinae. Given criticism of the *j* parameter in DEC analyses (Ree and Sanmartín 2018), ancestral range estimations of the second best-fit model, BAYAREALIKE, were compared (Table 9). According to this model, although the ancestral range of Pseudochactinae was the same as the DIVALIKE + *j* model, the ancestral ranges of Pseudochactidae, Troglokhammouaninae and Vietbocapinae were broader (Table 9). However, these differences do not alter the biogeographical hypothesis proposed below. Output files from BioGeoBEARS are publicly available in figshare: 10.6084/m9.figshare.21067528.

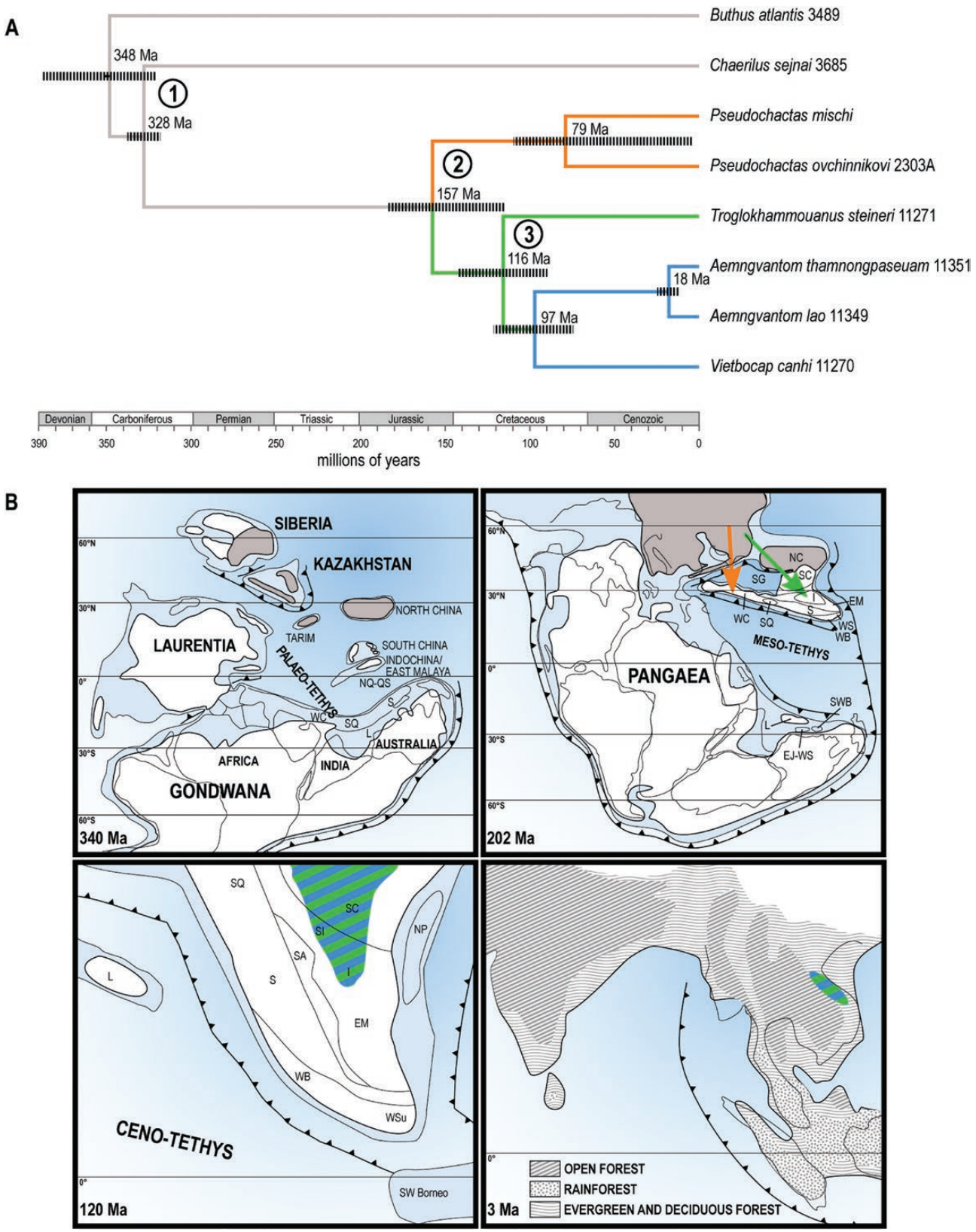

Fig. 3. Estimation of divergence times (**A**) in evolution of the relictual Asian scorpion family Pseudochactidae Gromov 1998 with palaeogeographical reconstructions (**B**) indicating tectonic positions and past and present plate boundaries. Geological time periods on phylogeny include Devonian, Carboniferous, Permian, Triassic, Jurassic, Cretaceous, and Cenozoic. Divergence ages (in millions of years) plotted on phylogeny. Circled numbers denote three divergence events leading to Pseudochactidae and its three subfamilies. Timeline for origins, dispersal, and diversification of Pseudochactidae as follows: Carboniferous (340 Ma), Pseudochactidae originates near Tajik block (colored grey area), situated among chain of continental blocks and island arcs between Turkestan and Palaeo-Tethys Oceans. Jurassic–Triassic Boundary (202 Ma): Cimmerian continent collides with Eurasia and Indochina in the Jurassic, allowing ancestor of

Table 8. Statistics for models tested (DEC, DIVALIKE, and BAYAREALIKE) including an additional founder effect (*j*) parameter in ancestral range estimation of the relictual Asian scorpion family Pseudochactidae Gromov 1998: likelihoods (lnL), number of parameters in each model, dispersal (*d*), extinction (*e*), founder effect (*j*), Akaike Information Criterion (AIC), and corrected AIC (AICc)

Model	lnL	Parameters	*d*	*e*	*j*	AIC	AICc
DEC	−16.29	2	0.01000	0.001	0	36.6	40.6
DEC + *j*	−16.29	3	0.01000	0.001	0.00010	38.6	50.6
DIVALIKE	−16.34	2	0.01000	0.001	0	36.7	40.7
DIVALIKE + *j*	−5.15	3	0.00000	0.00000	0.14938	16.3	28.3
BAYAREALIKE	−11.31	2	0.00114	0.00878	0	26.6	30.6
BAYAREALIKE + *j*	−6.89	3	0.00000	0.00000	0.26504	19.8	31.8

Troglomorphism and Ancestral State Reconstruction

The ensemble index of troglomorphic characters identified a continuum from epigean species (score = 1) to hypogean species progressively more adapted to caves (score = 10): 1–5 (epigean), 6–8 (troglophilic), 9–10 (troglobitic). The two *Pseudochactas* were recovered as epigean (scores for *P. mischi* and *P. ovchinnikovi* = 5.5 and 4, respectively), *T. steineri* as troglophilic (score = 6), and the three species of Vietbocapinae, as troglobitic (scores for *A. lao*, *A. thamnongpaseuam* and *V. canhi* = 9.5, 9.5 and 10, respectively). The outgroup taxa, *B. atlantis* and *C. sejnai*, were recovered as epigean (scores = 2 and 2.5, respectively).

Ancestral state reconstruction of the cave adaptation index produced log-likelihood scores of −13.45, −11.83, −10.96 for the ER, SYM, and ARD models, respectively. Chi-squared tests of log-likelihood scores selected the ARD as the best fit, although reconstructions across all models were mostly congruent (Fig. 5). According to reconstructions from all three models, the ancestors of Pseudochactidae and Pseudochactinae were epigean, whereas the ancestors of Troglokhammouaninae and Vietbocapinae were troglobitic. Output files from the ancestral range estimation analysis are publicly available in figshare: 10.6084/m9.figshare.21067528.

Discussion

Pseudochactid Systematics

The type species of the family, *P. ovchinnikovi*, was discovered in the Babatag Mountains of southeastern Uzbekistan and southwestern Tajikistan. The distribution of Pseudochactidae was extended when the cavernicolous *T. steineri* was discovered in a cave in the Hin Nam No National Protected Area of central Laos. Three years later, another cavernicolous species, *V. canhi*, was described from a second cave on the Vietnamese side of the Annamite Mountains. *Pseudochactas mischi* from Afghanistan, another three species of *Vietbocap* and a second species of *Troglokhammouanus* from the Khammouan-Phong Nha-Kẻ Bàng Karst were subsequently described (Lourenço 2012, 2017; Lourenço and Pham 2012; Lourenço et al. 2018). However, a recent revision of Pseudochactidae by Prendini et al. (2021) synonymized *V. aurantiacus*, *V. quinquemilia* and *V. thienduongensis* with *V. canhi*; *Troglokhammouanus louiseanneorum* Lourenço 2017 with *T. steineri*; and created a new cavernicolous genus, *Aemngvantom*, to accommodate two species endemic to the Khammouan Karst of Laos: *A. lao* and *A. thamnongpaseuam*. Three subfamilies were recognized: Pseudochactinae accommodating *Pseudochactas*; Vietbocapinae accommodating *Aemngvantom* and *Vietbocap*; and Troglokhammouaninae accommodating *Troglokhammouanus*.

The phylogenetic analyses presented here confirm recent systematic changes in Pseudochactidae implemented by Prendini et al. (2021). The three subfamilies, Pseudochactinae, Troglokhammouaninae and Vietbocapinae, were monophyletic with high support across analyses. Within Pseudochactinae, *Pseudochactas* and *P. ovchinnikovi* were monophyletic. *Troglokhammouanus steineri*, the sole species of Troglokhammouaninae, was also recovered as monophyletic. The genera and species of Vietbocapinae were also monophyletic across the phylogenetic analyses.

Eurasian Origin in Carboniferous

Pseudochactidae is an ancient, relictual lineage (Soleglad and Fet 2003a, Fet et al. 2004, Prendini et al. 2006). Soleglad and Fet (2003a) proposed a Pangaean origin, during the Permian/Triassic, for the family. Fet et al. (2004) suggested that the ancestor of Pseudochactidae inhabited littoral forests on islands in the Tethys Sea in the Cretaceous, which were elevated during Tertiary uplift due to collision of the Indian subcontinent with Eurasia, shielding them from the onset of aridification in the Eocene. Prendini et al. (2006) stated that a more recent divergence was likely if Pseudochactidae was the sister group of Buthidae. A dated phylogenomic analysis of scorpions, calibrated using fossil scorpions and other arachnids under an autocorrelated lognormal model, proposed an even earlier divergence between Pseudochactidae and Chaerilidae, ca. 329 Ma in the Carboniferous, whereas a second analysis applying an uncorrelated gamma rates model, placed this divergence much later, ca. 198 Ma in the Jurassic (Sharma et al. 2018). However, Sharma et al. (2018) preferred the results of the autocorrelated lognormal model

Pseudochactidae to disperse into western Cimmeria (orange arrow), resulting in subfamily Pseudochactinae Gromov 1998, and Indochina (green arrow), leading to common ancestor of Southeast Asian subfamilies, Troglokhammouaninae Prendini et al. 2021 and Vietbocapinae Lourenço 2007. Early Cretaceous (120 Ma): Southeast Asian pseudochactids diverge into two subfamilies (blue-green stripes): Troglokhammouaninae and Vietbocapinae. Pliocene (3 Ma): Miocene aridification causes pseudochactids to retreat into subterranean refugia in Annamite Mountains (blue-green stripes) and widespread extinction of surface-dwelling pseudochactids across ancestral range in Southeast Asia. 340 Ma, 202 Ma, and 120 Ma palaeogeographical reconstructions from Metcalfe (2011, 2013) with Africa, Australia, East Java-West Sulawesi (EJ-WS), East Malaya (EM), India, Indochina (I), Kazakhstan, Laurentia, Lhasa (L), north China (NC), north Qiangtang-Qamdo-Simao (NQ-QS), Siberia, Simao (SI), south China (SC), south Qiangtang (SQ), southwest Borneo (SWB), Sibumasu (S), Sukhothai Arc (SA), Tarim, west Burma (WB), west Sumatra (WS), and western Cimmeria (WC) land blocks; 3 Ma palaeogeographical reconstruction from Morley (2018): open forest with grasses, scrub and savanna; seasonal evergreen, semi-evergreen and closed deciduous forest; and perhumid rainforest.

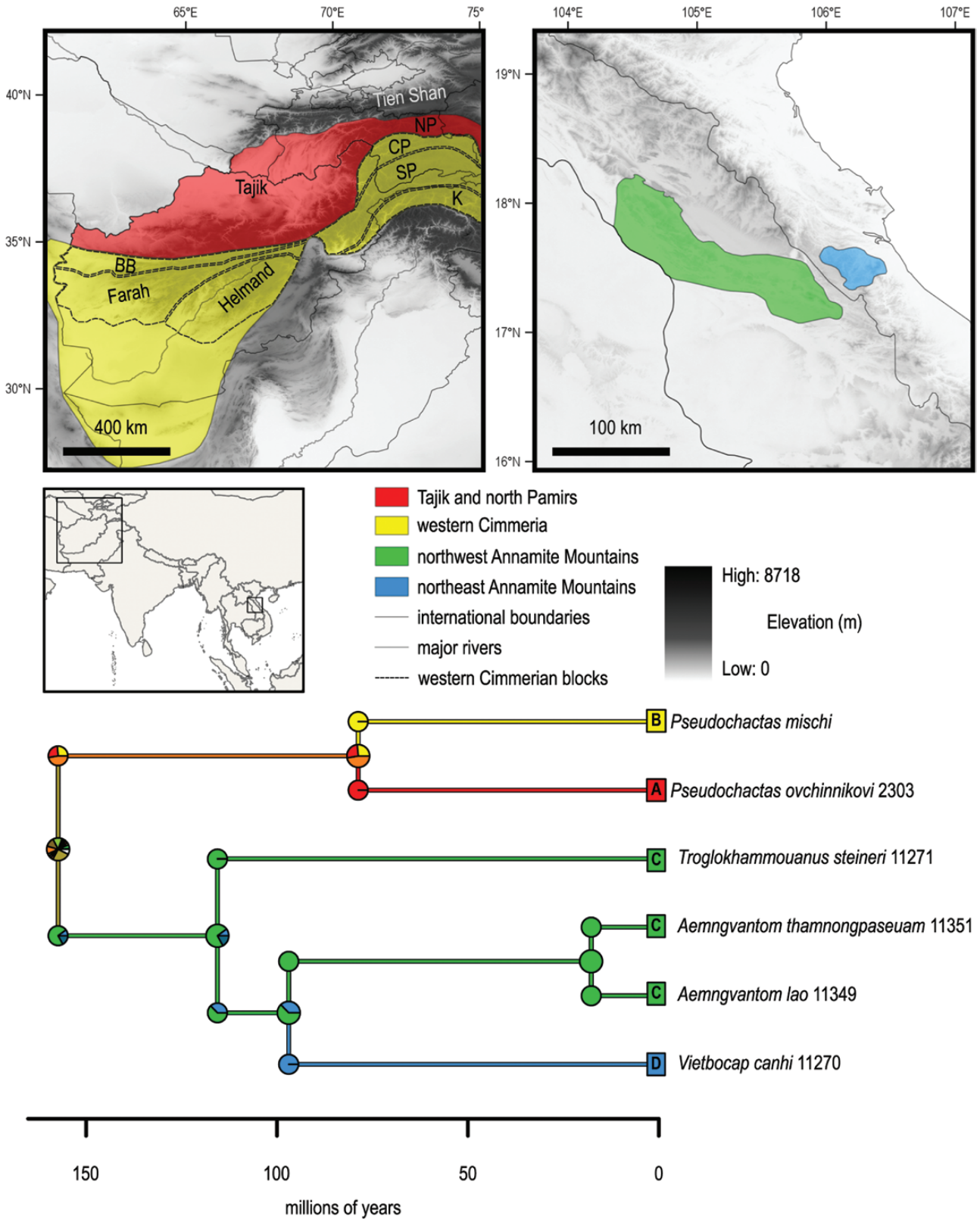

Fig. 4. Estimation of ancestral ranges for the relictual Asian scorpion family Pseudochactidae Gromov 1998 in four biogeographical regions: Tajik and north Pamir blocks (northern Afghanistan and southern China, Uzbekistan and Tajikistan) (**A**); western Cimmeria, including only the Farah, Helmand, south and central Pamirs (SP and CP, respectively), Karakorum (K) and Band-e Bayan (BB) blocks (present-day Afghanistan and northern Pakistan) (**B**); northwest Annamite Mountains (Khammouan Karst, Laos) (**C**); northeast Annamite Mountains (Phong Nha-Kẻ Bàng Karst, Vietnam) (**D**). Geological reconstructions following Montenat (2009), Collett et al. (2015), Robinson (2015), Siehl (2017), and Li et al. (2020).

Table 9. Most probable ancestral areas of the relictual Asian scorpion family Pseudochactidae Gromov 1998 and its subfamilies based on the best-fit DIVALIKE + *j* model and the second best-fit BAYAREALIKE model

Taxon	DIVALIKE + *j*	BAYAREALIKE
Pseudochactidae Gromov 1998	A, B, C	A, B, C, D
Pseudochactinae Gromov 1998	A, B	A, B
Troglokhammouaninae Prendini et al. 2021	C	C, D
Vietbocapinae Lourenço 2007	C	C, D

Biogeographical regions as follows: Tajik and north Pamir blocks (northern Afghanistan and southern China, Uzbekistan and Tajikistan) (**A**); western Cimmeria, including only the Farah, Helmand, south and central Pamirs (SP and CP, respectively), Karakorum (K) and Band-e Bayan (BB) blocks (present-day Afghanistan and northern Pakistan) (**B**); northwest Annamite Mountains (Khammouan Karst, Laos) (**C**); northeast Annamite Mountains (Phong Nha-Kẻ Bàng Karst, Vietnam) (**D**).

as this model is known to outperform the uncorrelated gamma rates model in simulations and empirical studies (Lepage et al. 2007, Paradis 2013).

The analyses presented here, which were calibrated using the (Chaerilidae + Pseudochactidae) node from the autocorrelated lognormal model of Sharma et al. (2018), support a Carboniferous origin for Pseudochactidae. Based on this divergence date and the geological literature, it is probable that the ancestor of Pseudochactidae was present on the Tajik and north Pamir blocks in the Paleozoic. During the Carboniferous and Permian, the Tajik and north Pamir blocks were located between the Turkestan and Palaeo-Tethys Oceans in the Northern Hemisphere and situated among a chain of continental blocks and island arcs, including the Tien Shan, Tarim, and Karakum blocks, among others (Fig. 3). Some geologists consider the Tajik block to be part of the Tajik-Karakum microcontinent, which was bounded to the south by north Pamir (Heubeck 2001, Biske and Seltman 2010, Alexeive et al. 2020, Nurtaev et al. 2021). It is likely that the ancestor of Pseudochactidae was located somewhere in this area because the Indochina and western Cimmerian blocks were still attached to Gondwana in the Southern Hemisphere at this time (Li et al. 2020; Fig. 3). Western Cimmeria was also largely submerged during the Permian (Metcalfe 2011) and limestone formations from the Khammouan-Phong Nha-Kẻ Bàng Karst date to the Carboniferous–Permian, suggesting this region was also submerged in the Paleozoic (Averyanov et al. 2003, Metcalfe 2011, Debevec et al. 2012, Nguyen et al. 2012, Bolger 2019, Kambesis 2019), and providing further support for a Tajik/north Pamir origin for Pseudochactidae.

Ancestral state reconstruction of the troglomorphic ecomorphotype also suggests that the ancestor of Pseudochactidae was epigean. Palynological data from the Tien Shan block, now adjacent to the Tajik block (Fig. 4), reveal a conifer-dominated forest that evolved in a xeric, seasonal environment during the Carboniferous (Zhou 1994). This corroborates additional palynological data from the Tarim block during the Permian which suggests a landscape dominated by conifers and pteridosperms adapted to drought conditions (Zhu et al. 2005). During the Permian–Triassic, the Tajik block became part of Asian plate and was later an inundated basin in the Mesozoic and Cenozoic (Faryad et al. 2013, Dedow et al. 2020, Kaya et al. 2020).

Late Jurassic Dispersal Across Western Cimmeria and Indochina

The phylogenetic analyses recovered two major clades in Pseudochactidae with a disjunct distribution, Pseudochactinae, endemic to Central Asia, and a clade comprising the two Southeast Asian subfamilies. Whereas Pseudochactinae are epigean, the Southeast Asian pseudochactids are strictly hypogean cave-dwellers (Prendini et al. 2021; Fig. 5). As Pseudochactidae comprises only six known species, including several cave relicts, disjunctly distributed across a vast area, it can only be assumed that the family was once more widespread and experienced high levels of extinction during its evolutionary history. Lourenço (2007) proposed a widespread distribution of Pseudochactidae in the Paleozoic and suggested that the only possible land connecting Uzbekistan and Tajikistan (Central Asia) to Laos (Southeast Asia) was the old Asia core.

According to the divergence time estimation and ancestral range estimation analyses, the divergence between Pseudochactinae and the Southeast Asian subfamilies occurred roughly 157 Ma in the Late Jurassic with the ancestor of Pseudochactinae occurring on the Tajik and western Cimmerian blocks and the ancestor of the Southeast Asian subfamilies occurring in Indochina. These results can best be explained by dispersal across the Cimmerian continent, a conglomeration of land blocks that rifted from Gondwana in the Early Permian and traveled across the Palaeo-Tethys during the Middle to Late Permian (Golonka 2004, Wang et al. 2021; Fig. 3). The Cimmerian continent comprised two parts, western and eastern Cimmeria. Although it is widely agreed that eastern Cimmeria comprised the Baoshan, Sibumasu, South Qiangtang, Tenchong, and possibly Lhasa land blocks (Angiolini et al. 2013, Metcalfe 2013, Wang et al. 2021), there is debate concerning the composition of western Cimmeria. Whereas most authors include south-central Iran, south-central Afghanistan, and south-central Pamir as part of western Cimmeria, some restrict western Cimmeria to southern Afghanistan, southeastern Pamir and southern Iran because central Afghanistan, central Iran, and central Pamir rifted from Gondwana in the Early Carboniferous and therefore do not have the same geological history as the rest of Cimmeria (Angiolini et al. 2013, Robinson 2015, Wang et al. 2021). Regardless, by the Late Triassic (ca. 200 Ma), western and eastern Cimmeria began to collide with each other, with the north China, south China and Indochina blocks, and with Eurasia, thereby closing the Palaeo-Tethys (Fig 3). This collision situated eastern Cimmeria adjacent to the north and south China blocks and Indochina and western Cimmeria just south of the Baltic continent (Miyahigashi et al. 2010; Metcalfe 2011, 2013), enabling faunal exchange between Cimmeria and proto-Eurasia. Based on the data presented here, it appears that the ancestor of the Southeast Asian pseudochactid subfamilies dispersed from the Tajik-north Pamir region of proto-Eurasia into Indochina via the north and south China blocks after this collision in the Jurassic (Metcalfe 2013; Fig. 3). This collision also allowed Pseudochactinae to disperse to the Farah and Helmand blocks (south-central Afghanistan) via the south-central Pamir blocks in western Cimmeria. The absence of Early Cretaceous marine deposits in south Pamir suggests it was above sea level during this time (Robinson 2015). Oceanic barriers among these western Cimmerian continental blocks may have caused divergence between the two Pseudochactinae species in the Late Cretaceous (Kaya et al. 2020; Fig. 4).

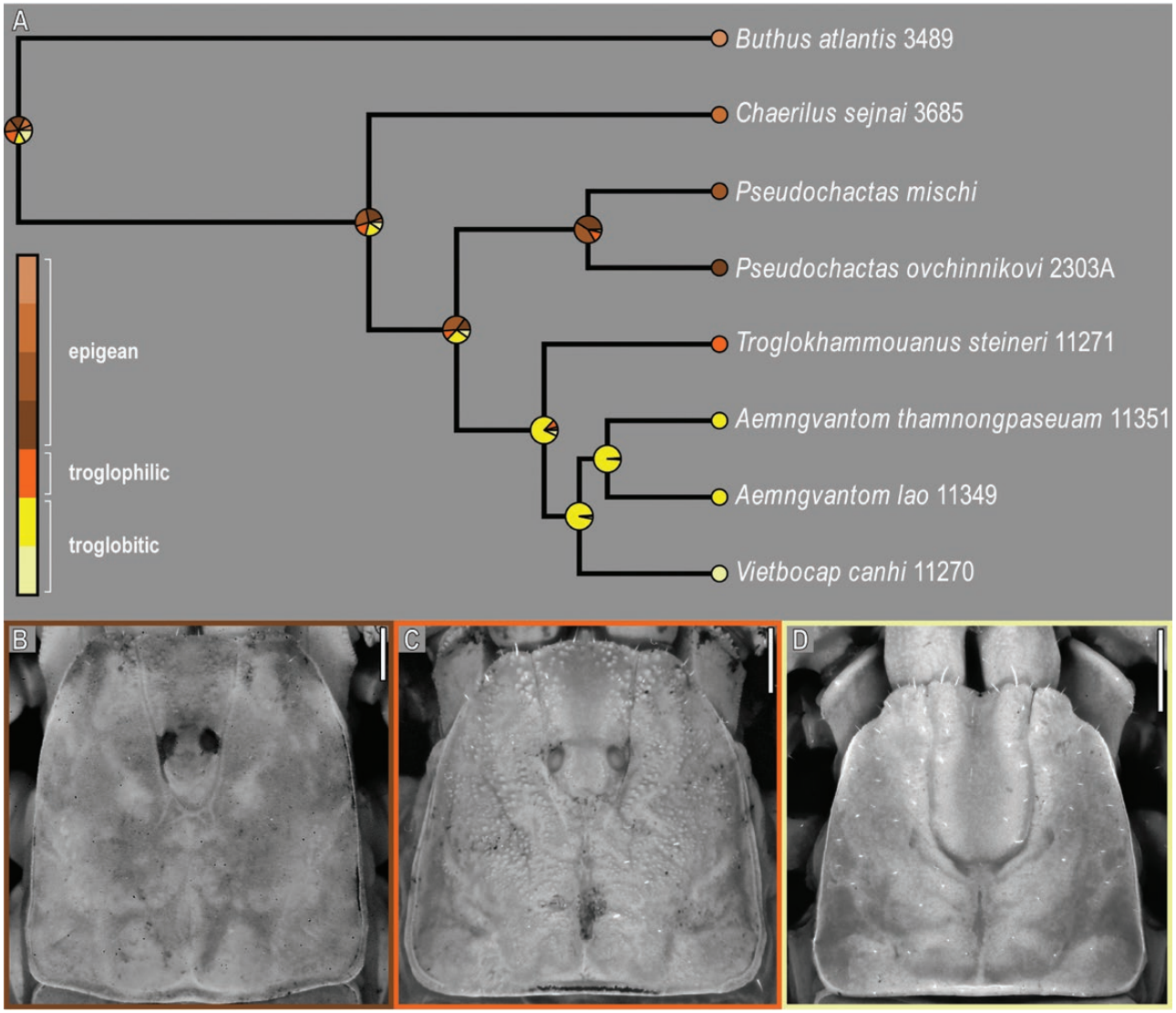

Fig. 5. Preferred ancestral state reconstruction (**A**) for the relictual Asian scorpion family Pseudochactidae Gromov 1998 using ARD model of troglomorphism index (scale = 1–10) ranging from epigean (score = 1) to hypogean, troglobitic (score = 10). Ecomorphotypes classified as epigean (1–5); hypogean, troglophilic (6–8); and hypogean, troglobitic (9–10). Carapace (**B–D**) of three pseudochactid species, representing each ecomorphotype: epigean: *Pseudochactas ovchinnikovi* Gromov 1998 (**B**); troglophilic: *Troglokhammouanus steineri* Lourenço 2007 (**C**); and troglobitic: *Vietbocap canhi* Lourenço 2010 (**D**).

Late Cretaceous Diversification in Southeast Asia and Late Miocene Extinction

The phylogeny recovered two subfamilies, Troglokhammouaninae and Vietbocapinae, within the Southeast Asian clade. Whereas *T. steineri*, the sole species of Troglokhammouaninae, is troglophilic, the three species of Vietbocapinae are troglobitic. All Southeast Asian pseudochactids are endemic to the Khammouan-Phong Nha-Kẻ Bàng Karst in the northern Annamite Mountains of Laos and Vietnam. This karstic region, 200 km long by 30 km wide, is similar in age and composition to the karst of southern China, and dates from the Middle Carboniferous to Lower Permian (Waltham and Middelton 2000, Laumanns and Price 2011). Few studies have examined the evolution of the fauna in this karstic region and the only research on the evolution of cave arachnids in East Asia concern species in southern China. Nicolas et al. (2012) suggested that the Laotian rock rat, *Laonastes aenigmanus* Jennkins et al. 2005, was widespread across limestone formations in Laos during the Late Miocene to Early Pliocene. However, karst fragmentation during the Pliocene/Pleistocene isolated populations resulting in a species complex. Zhang and Li (2013) found that spiders of the genus *Nesticella* Lehtinen and Saaristo 1980 from caves in southern China originated in the Early Miocene (ca. 20 Ma) and underwent substantial diversification between the Middle Miocene to Late Pliocene due to tectonic uplift in southern China. However, radiation into caves was only initiated in the Pleistocene as species took refuge from glaciation in subterranean habitats and colonized caves multiple times.

Based on the divergence time and ancestral range estimation analyses, the diversification of Southeast Asian pseudochactids began in the Early Cretaceous (ca. 116 Ma) and the ancestor of this clade was present in the northwestern Annamite Mountains (Khammouan Karst; Fig. 3). Ancestral state reconstruction of troglomorphism recovered the ancestors of both Troglokhammouaninae and Vietbocapinae as troglobitic, suggesting that *T. steineri*, a

troglophile that lacks cave-adapted features (Prendini et al. 2021), evolved from an ancestor resembling the markedly troglomorphic Vietbocapinae (Fig. 5); however, this hypothesis will need to be tested further. Divergence between *Aemngvantom* and *Vietbocap* occurred in the Late Cretaceous (ca. 96 Ma), whereas divergence between *A. lao* and *A. thamnongpaseuam* occurred in the Early Miocene (18 Ma).

Beginning in the Late Miocene, uplift of the Himalayas and Tibetan Plateau, closure of the Paratethys, and global cooling resulted in aridification across Asia (Gathorne-Hardy et al. 2002, Miao et al. 2013, Le et al. 2015, Tu et al. 2015, Loria and Prendini 2020). In Southeast Asia, forest was replaced with xeric-adapted plants as evidenced by palynology and the remains of many savannah herbivores (Hui et al. 2011, Le et al. 2015, Morley 2018, Jha et al. 2021). Morley (2018) indicated that between 20 Ma and 3 Ma, most of Indochina shifted from seasonal forest to open forests with grasses, scrub, and savanna (Fig. 3). Despite climatic and vegetational changes across the region, the Annamite (Trường Sơn) Mountains have been posited as a potential refugium in Southeast Asia during periods of climate change (Surridge et al. 1999, Morley 2018). This mountain chain, uplifted in the early Triassic (Sterling and Hurley 2008), is home to many ancient, endemic mammalian and plant lineages, including the saola antelope (*Pseudoryx nghetinhensis* Dung et al. 1993), the giant muntjac (*Muntiacus vuquangensis* Tuoc et al. 1994), the Annamite striped rabbit (*Nesolagus timminsi* Averianov et al. 2000) and Krempf's pine (*Pinus krempfii* Lecomte 1921), implying that the fauna and flora of these mountains remained relatively stable for a long period of time. A recent divergence time estimation analysis of the Laotian rock rat (*L. aenigmanus*) also suggested this species survived in the Annamite forests (Le et al. 2015). The results of the present study support previous hypotheses (Bain and Hurley 2011) in suggesting that epigean pseudochactids largely went extinct across their range in Central and Southeast Asia during aridification in the Late Miocene whereas hypogean pseudochactids survived in refugial limestone caves of the northern Annamites (Figs. 3–5).

Alternative Hypotheses for Origins of Southeast Asian Scorpions

Few biogeographical studies have investigated the origins and diversification of Southeast Asian scorpions. Loria and Prendini (2020) suggested an 'Out of India' origin for the Asian forest scorpions (Heterometrinae Simon 1879). According to their hypothesis, Heterometrinae diverged from other Scorpionidae Latreille 1802 on the African continent after the Indian subcontinent became separated in the Cretaceous. Environmental stresses during the Cretaceous–Tertiary mass extinction caused range contraction, restricting one clade of Heterometrinae to refugia in southern India (the Western Ghats) and Sri Lanka (the Central Highlands). Heterometrinae then dispersed to Southeast Asia three times during India's collision with Eurasia. The first dispersal event occurred in the Paleocene as the Indian subcontinent brushed up against the western side of Sumatra (ca. 55 Ma), the second in the Early Eocene as India moved closer to Sumatra and Southeast Asia (45 Ma), and the third in the Late Eocene as India touched the western side of Myanmar during its final collision with Eurasia (ca. 35 Ma). After the Indian-Eurasian collision, India's climate shifted from zonal to monsoon during the Miocene, resulting in aridification across the subcontinent. Previously confined to southern India and Sri Lanka during the KT mass extinction, Heterometrinae recolonized the Deccan Plateau and northern India, diversifying into new, more arid habitats after environmental conditions stabilized, until all niches were occupied by the Middle Miocene (ca. 10 Ma).

Although Monod and Prendini (2015) used the African-Indian rift to explain the origins of an Indian clade of Hormuridae Laurie 1896, the 'Eurogondwana' hypothesis was invoked to explain the origin and distribution of an Indo-Pacific hormurid clade. According to this hypothesis, Hormuridae originated on Gondwana in the Paleozoic and the ancestors of the Indo-Pacific hormurid clade diverged in isolation from the ancestors of the African and South American hormurid genera, when the Apulian microplate rifted from Africa in the Early Cretaceous. The Indo-Pacific hormurid clade subsequently arrived in Eurasia when the Apulian microplate collided with Europe and dispersed across the Palearctic, eventually reaching Southeast Asia and Australasia. No hormurids presently occur in the Palearctic; hence Monod and Prendini (2015) suggested that the 'icehouse' climate in the Neogene and the 'hothouse' climate in the late Paleocene–early Eocene may have caused the extinction of these mostly tropical or subtropical, humidity-dependent scorpions.

The present contribution provides the first divergence time and biogeographical analysis of the scorpion family Pseudochactidae. Unlike previous hypotheses (i.e., the 'Out of India' and 'Eurogondwana' hypotheses) concerning the origins of Southeast Asian scorpions, the analyses presented support an 'Out of Eurasia' origin for this family. Pseudochactids likely originated along a chain of continental blocks and island arcs adjacent to proto-Eurasia in the Carboniferous and dispersed into Indochina and Afghanistan after the Cimmerian continent and Indochina collided with Eurasia at the Triassic-Jurassic boundary. However, unlike Heterometrinae, which diversified into new environments with the onset of Late Miocene aridification across Asia, epigean pseudochactids experienced widespread extinction across their range, persisting mostly as hypogean relicts in subterranean refugia provided by the karstic landscape of the Annamite Mountains, consistent with the 'Climate Relict' hypothesis.

Note Added in Proof

While this paper was in press, a new monotypic genus of Pseudochactidae, *Qianxie solegladi* Tang, 2022, was described from a riparian habitat in Yunnan, China in: **Tang, V. 2022.** A new scorpion genus and species from China, *Qianxie solegladi* gen. et sp. n. (Scorpiones: Pseudochactidae). Euscorpius 351: 1–19. Tang (2022) suggested *Q. solegladi* is closely related to *T. steineri* and placed it in Troglokhammouaninae, a hypothesis that appears plausible but awaits confirmation with phylogenetic analysis. Although the new taxon was described based on a single female holotype, an adult male and several other observational records were noted by the author in both the Yunnan and Sichuan provinces of China. The discovery of this new species supports the biogeographical hypotheses presented herein as the area in which it occurs was formerly part of the South China block, suggesting that the ancestor of the Southeast Asian pseudochactid subfamilies (Troglokhammouaninae + Vietbocapinae) dispersed from the Tajik-north Pamir region of proto-Eurasia into Indochina via the north and south China blocks. *Quanxie solegladi* is lapidicolous and represents the first epigean member of the Southeast Asian pseudochactid subfamilies.

Supplementary Data

Supplementary data are available at *Insect Systematics and Diversity* online.

Acknowledgments

We acknowledge the following friends and colleagues for assistance with fieldwork and/or donations of material used in the study: Kazakhstan: A.V. Gromov and S. Mozorov; Laos: P. Kanyavong, S.T. Kangnavong, W. Phimmathong, and P.T. Thon Thondet; Uzbekistan: S. Chibutarov, A. and E. Kreuzberg, and S. Shanazarov; Vietnam: V.A. Dang, H.V. Dương, L. Ngo, D.-S. Pham, and B.Đ. Trân; V. Vignoli for material from Morocco; and the following colleagues, institutions and organizations for facilitating permissions to collect and export material: P. Kanyavong and S.T. Kangnavong (Green Discovery, Laos), Lao Ministry of Natural Resources and Environment Permits 91 and 297; A. Kreuzberg (Institute of Zoology, Uzbek Academy of Sciences, Tashkent) and P. Usmanov (Institute of Physiology and Biophysics, Uzbek Academy of Sciences, Tashkent); A.T.T. Nguyễn, N.V. Sinh, T.Đ. Bĩnh, S.T. Nguyễn, and D.-S. Pham (Institute of Ecology and Biological Resources, Vietnam Academy of Science and Technology, Hanoi), IEBR Permits 72/STTNSV, 73/STTNSV and 74/STTNSV; B.Đ. Trân and H.V. Dương (Phong Nha-Kẻ Bàng National Park), Vietnam Ministry of Agriculture and Rural Development, General Department of Forestry Collecting Permits 110/TCLN-BTTN and 1643/ TCLN-ĐDPH; Vietnam Ministry of Natural Resources and Environment Export Permit 3493/ QĐ-BTNMT; M. Judson, C. Rollard, and E.-A. Leguin (MNHN) for access to the scorpion collection during visits by S.F.L. and L.P.; P.A. Colmenares, O. Delgado, P. Horsley, M.M. Locke, and L. Sorkin for logistical support with collections at the AMNH; D. Harms for covering travel expenses for V.L.E. and assistance with the DAAD for S.F.L.; P. Rubi and T. Sharma for generating some of the DNA sequence data used in the study; S. Thurston for preparing the plates for this contribution; E.F. Gaughan for assistance with travel reimbursements and administering funds; and two reviewers for constructive comments on a previous draft of the manuscript.

Author Contributions

SFL: Conceptualization; Formal analysis; Funding acquisition; Investigation; Project Administration; Supervision; Writing – original draft; Writing – review & editing. VLE: Investigation; Visualization; Writing – review & editing. ADN: Data curation; Project administration; Writing – review & editing. LP: Conceptualization; Data curation; Funding acquisition; Investigation; Project administration; Resources; Supervision; Writing – review & editing.

Funding

S.F.L. was financially supported by a fellowship from the AMNH Richard Gilder Graduate School, a National Science Foundation (NSF) Graduate Research Fellowship, and a Postdoctoral Research Fellowship from NSF grant DEB 1655050 to L.P. Aspects of this research were funded by the following awards: National Science Foundation Doctoral Dissertation Improvement grant DEC 1310855 to L.P. and S.F.L.; a Deutscher Akademischer Austauschdienst (DAAD)/ German Academic Exchange Service grant to S.F.L.; NSF grants EAR 0228699, DEB 0413453 and DEB 1655050 to L.P.; a grant from the Richard Lounsbery Foundation to L.P.; two Constantine S. Niarchos Expedition grants from the Stavros Niarchos Foundation to L.P.; two grants from the Exploration Fund of the Explorers Club, one each to S.F.L. and V.L.E.; Young Explorers Grant 9816-15 from the National Geographic Society to S.F.L.; and the Vietnam Ministry of Science and Technology under the grant ĐTĐL.CN-113/21 (leader Vu Van Lien) to A.D.N.

References Cited

Alexeiev, D. V., G. S. Biske, A. Kröner, A. A. Tretyakov, V. P. Kovach, and **Y. Rojas-Agramonte. 2020**. Ediacaran, Early Ordovician and early Silurian arcs in the South Tianshan orogen of Kyrgyzstan. J. Asian Earth Sci. 190: 104194.

Angiolini, L., G. Crippa, G. Muttoni, and **J. Pignatti. 2013**. Guadalupian (middle Permian) paleobiogeography of the Neotethys Ocean. Gondwana Res. 24: 173–184.

Assmann, T., A. Casale, C. Drees, J. C. Habel, A. Matern, and **A. Schuldt. 2010**. The dark side of relict species biology: cave animals as ancient lineages, pp. 91–103. *In* J. C. Habel, T. Assmann (eds.), Relict species. Springer, Berlin, Heidelberg, Germany.

Averyanov, L. V., P. K. Loc, N. T. Hiep, and **D. K. Harder. 2003**. Phytogeographic review of Vietnam and adjacent areas of eastern Indochina. Komarovia 3: 1–83.

Bain, R. H., and **M. M. Hurley. 2011**. A biogeographic synthesis of the amphibians and reptiles of Indochina. Bull. Amer. Mus. Nat. Hist. 360: 1–138.

Barr, T. C. 1968. Cave ecology and the evolution of troglobites, pp. 35–102. *In* T. Dobzhansky, M. K. Hecht, W. C. Steere (eds.), Evolutionary biology. Springer, Boston, MA.

Biske, Y. S., and **R. Seltmann. 2010**. Paleozoic Tian-Shan as a transitional region between the Rheic and Urals-Turkestan oceans. Gondwana Res. 17: 602–613.

Bolger, T. 2019. Xe Bang Fai Cave, Laos, pp. 1185–1192. *In* W. B. White, D. C. Culver (eds.), Encyclopedia of caves, 3rd ed. Academic Press, Waltham, MA.

Bouckaert, R., T. G. Vaughan, J. Barido-Sottani, S. Duchêne, M. Fourment, A. Gavryushkina, J. Heled, G. Jones, D. Kühnert, and **N. De Maio, et al. 2019**. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 15: e1006650.

Bryson, R. W. Jr., B. R. Riddle, M. R. Graham, B. T. Smith, and **L. Prendini. 2013**. As old as the hills: montane scorpions in southwestern North America reveal ancient associations between biotic diversification and landscape history. PLoS One 8: e52822.

Bryson, R. W. Jr., L. Prendini, W. E. Savary, and **P. B. Pearman. 2014**. Caves as microrefugia: Pleistocene phylogeography of the troglophilic North American scorpion *Pseudouroctonus reddelli*. BMC Evol. Biol. 14: 1–16.

Clements, R., N. S. Sodhi, M. Schilthuizen, and **P. K. Ng. 2006**. Limestone karsts of Southeast Asia: imperiled arks of biodiversity. BioScience 56: 733–742.

Cock, P. J. A., T. Antao, J. T. Chang, B. A. Chapman, C. J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, and **B. Wilczynski, et al. 2009**. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics 25: 1422–1423.

Coddington, J. A., G. Giribet, M. S. Harvey, L. Prendini, and **D. E. Walter. 2004**. Arachnida, pp. 296–318. *In* J. Cracraft, M. Donoghue (eds.), Assembling the tree of life. Oxford University Press, Oxford, UK.

Collett, S., S. W. Faryad, and **A. M. Mosazai. 2015**. Polymetamorphic evolution of the granulite-facies Paleoproterozoic basement of the Kabul Block, Afghanistan. Mineral. Petrol. 109: 463–484.

Darriba, D., G. L. Taboada, R. Doallo, and **D. Posada. 2012**. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9: 772.

Day, M., and **P. Urich. 2000**. An assessment of protected karst landscapes in Southeast Asia. Cave Karst Sci. 27: 61–70.

Debevec, B., M. Knez, A. Kranjc, M. Pahor, M. Prelovšek, and **T. Slabe. 2012**. Preliminary study for the adaptation of the >>Heaven's Cave<< for tourist purposes (Phong Nha-Ke Bang National Park, Vietnam). Acta Carsologica 41: 115–127.

Dedow, R., M. Franz, A. Szulc, J. W. Schneider, J. Brückner, L. Ratschbacher, Ł. Gągała, J. C. Ringenbach, N. Rajabov, and **M. Gadoev, et al. 2020**. Tajik Basin and southwestern Tian Shan, northwestern India-Asia collision zone: 3. Preorogenic to synorogenic retro-foreland basin evolution in the eastern Tajik depression and linkage to the Pamir hinterland. Tectonics 39: e2019TC005874.

Deharveng, L., and **A. Bedos. 2000**. The cave fauna of Southeast Asia: origin, evolution and ecology, pp. 603–632. *In* H. Wilkens, D. C. Culver, W. F. Humphreys (eds.), Subterranean ecosystems, ecosystems of the World, vol. 30. Elsevier, Amsterdam, the Netherlands.

Deharveng, L., and **A. Bedos. 2012**. Diversity patterns in the tropics, pp. 238–250. *In* W. B. White, D. C. Culver (eds.), Encyclopedia of caves, 2nd ed. Academic Press, Waltham, MA.

Dimitrov, D., L. Lopardo, G. Giribet, M. A. Arnedo, F. Álvarez-Padilla, and **G. Hormiga. 2012**. Tangled in a sparse spider web: single origin of orb weavers and their spinning work unravelled by denser taxonomic sampling. Proc. R. Soc. B. 279: 1341–1350.

Dimitrov, D., J. J. Astrin, and **B. A. Huber**. **2013**. Pholcid spider molecular systematics revisited, with new insights into the biogeography and the evolution of the group. Cladistics 29: 132–146.

Drummond, A. J., M. A. Suchard, D. Xie, and **A. Rambaut**. **2012**. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29: 1969–1973.

Faryad, S. W., S. Collett, M. Petterson, and **S. A. Sergeev**. **2013**. Magmatism and metamorphism linked to the accretion of continental blocks south of the Hindu Kush, Afghanistan. Lithos 175: 302–314.

Fet, V. 2000. Family Pseudochactidae Gromov, 1998, p. 426. *In* V. Fet, W. D. Sissom, G. Lowe, M. E. Braunwalder (eds.), Catalog of the scorpions of the World (1758–1998). New York Entomological Society, New York, NY.

Fet, V., M. E. Soleglad, and **A. V. Gromov**. **2004**. The platypus of a scorpion: the genus *Pseudochactas* Gromov, 1998 (Scorpiones: Pseudochactidae). Euscorpius 17: 61–68.

Fitch, W. M. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. Syst. Zool. 20: 406–416.

Folmer, O., M. B. Black, W. Hoch, R. A. Lutz, and **R. C. Vrijehock**. **1994**. DNA primers for amplification of mitochondrial Cytochrome *c* Oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3: 294–299.

Gathorne-Hardy, F. J., R. G. Davies, P. Eggleton, and **D. T. Jones**. **2002**. Quaternary rainforest refugia in south-east Asia: using termites (Isoptera) as indicators. Biol. J. Linn. Soc. 75: 453–466.

Gibert, J., and **L. Deharveng**. **2002**. Subterranean ecosystems: a truncated functional biodiversity. BioScience 52: 473–481.

Giribet, G., S. Carranza, J. Baguña, M. Riutort, and **C. Ribera**. **1996**. First molecular evidence for the existence of a Tardigrada + Arthropoda clade. Mol. Biol. Evol. 13: 76–84.

Goloboff, P. A. 1993. Estimating character weights during tree search. Cladistics 9: 83–91.

Goloboff, P. A., and **S. A. Catalano**. **2016**. TNT version 1.5, including a full implementation of phylogenetic morphometrics. Cladistics 32: 221–238.

Goloboff, P. A., J. S. Farris, and **K. C. Nixon**. **2003**. TNT: tree analysis using new technology. Computer software and documentation. Available from https://cladistics.org/tnt

Goloboff, P. A., J. S. Farris, and **K. C. Nixon**. **2008**. TNT, a free program for phylogenetic analysis. Cladistics 24: 774–786.

Golonka, J. 2004. Plate tectonic evolution of the southern margin of Eurasia in the Mesozoic and Cenozoic. Tectonophysics 381: 235–273.

Gromov, A. V. 1998. A new family, genus and species of scorpions (Arachnida, Scorpiones) from southern Central Asia. Zool. Zh. 77: 1003–1008. [in Russian, English summary; English translation: Russian Journal of Zoology 2: 409–413].

Guindon, S., and **O. Gascuel**. **2003**. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52: 696–704.

Hall, R. 1998. The plate tectonics of Cenozoic SE Asia and the distribution of land and sea, pp. 99–131. *In* R. Hall, J. D. Holloway (eds.), Biogeography and geological evolution of SE Asia. Backhuys Publishers, Leiden, the Netherlands.

Harrison, R. G., D. M. Rand, and **W. C. Wheeler**. **1987**. Mitochondrial DNA variation in field crickets across a narrow hybrid zone. Mol. Biol. Evol. 24: 363–371.

Heubeck, C. 2001. Assembly of central Asia during the middle and late Paleozoic, pp. 1–22. *In* M. S. Hendrix, G. A. Davis (eds.), Paleozoic and Mesozoic tectonic evolution of central Asia: from continental assembly to intracontinental deformation. Memoirs of the Geological Society of America, Boulder, CO.

Howarth, F. G. 1973. The cavernicolous fauna of Hawaiian lava tubes, 1. Introduction. Pacific Insects 15: 139–151.

Howarth, F. G. 1993. High-stress subterranean habitats and evolutionary change in cave-inhabiting arthropods. Am. Nat. 142: S65–S77.

Hui, Z., J. Li, Q. Xu, C. Song, J. Zhang, F. Wu, and **Z. Zhao**. **2011**. Miocene vegetation and climatic changes reconstructed from a sporopollen record of the Tianshui Basin, NE Tibetan Plateau. Palaeogeogr. Palaeoclimatol. Palaeoecol. 308: 373–382.

Jha, A., S. Seneviratne, H. S. Prayag, and **K. Vasudevan**. **2021**. Phylogeny identifies multiple colonisation events and Miocene aridification as drivers of South Asian bulbul (Passeriformes: Pycnonotidae) diversification. Org. Divers. Evol. 21: 783–794.

Ji, Y.-J., D.-X. Zhang, and **L.-J. He**. **2003**. Evolutionary conservation and versatility of a new set of primers for amplifying the ribosomal internal transcribed spacer regions in insects and other invertebrates. Mol. Ecol. Notes 3: 581–585.

Juan, C., M. T. Guzik, D. Jaume, and **S. J. Cooper**. **2010**. Evolution in caves: Darwin's 'wrecks of ancient life' in the molecular era. Mol. Ecol. 19: 3865–3880.

Kambesis, P. 2019. Hang Son Doong and other caves of the Phong Nha-Kẻ Bàng karst, Quảng Bình Province, Vietnam, pp. 503–513. *In* W. B. White, D. C. Culver (eds.), Encyclopedia of caves, 3rd ed. Academic Press, Waltham, MA.

Katoh, K., and **D. M. Standley**. **2013**. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30: 772–780.

Katoh, K., K. Misawa, K. I. Kuma, and **T. Miyata**. **2002**. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30: 3059–3066.

Kaya, M. Y., G. Dupont-Nivet, J. N. Proust, P. Roperch, N. Meijer, J. Frieling, C. Fioroni, S. Özkan Altiner, M. Stoica, and **J. Aminov, et al. 2020**. Cretaceous evolution of the Central Asian proto-Paratethys Sea: tectonic, eustatic, and climatic controls. Tectonics 39: e2019–TC005983.

Kocher, T. D., W. K. Thomas, A. Meyer, S. V. Edwards, S. Pääbo, F. X. Villablanca, and **A. C. Wilson**. **1989**. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc. Natl. Acad. Sci. U.S.A. 86: 6196–6200.

Kumar, S., G. Stecher, M. Li, C. Knyaz, and **K. Tamura**. **2018**. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35: 1547–1549.

Lamoral, B. H. 1980. A reappraisal of the supragenerics classification of Recent scorpions and their zoogeography, pp. 439–444. *In* J. Grube (ed.), Verhandlungen 8. Internationaler Arachnologen-Kongress abgehalten an der Universität für Bodenkultur Wien, 7–12 Juli, 1980. H. Egermann, Vienna.

Laumanns, M., and **L. Price**. **2011**. A general assessment of the great caves and the karst of Southeast Asia, p. 424. *In* E. Haryono, T. N. Adjie, Suratman (eds.), Asian Trans-Disciplinary Karst Conference. 7–11 Jan. 2011, Yogyakarta, Indonesia.

Le, M., H. M. Nguyen, H. T. Duong, T. V. Nguyen, L. D. Dinh, N. X. Nguyen, L. D. Nguyen, T. H. Dinh, and **D. X. Nguyen**. **2015**. Phylogeography of the Laotian rock rat (Diatomyidae: *Laonastes*): implications for Lazarus taxa. Mamm. Study 40: 109–114.

Lepage, T., D. Bryant, H. Philippe, and **N. Lartillot**. **2007**. A general comparison of relaxed molecular clock models. Mol. Biol. Evol. 24: 2669–2680.

Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50: 913–925.

Li, Y. P., A. C. Robinson, M. Gadoev, and **I. Oimuhammadzoda**. **2020**. Was the Pamir salient built along a Late Paleozoic embayment on the southern Asian margin? Earth Planet. Sci. Lett. 550: 116554.

Loria, S. F., and **L. Prendini**. **2020**. Out of India, thrice: diversification of Asian forest scorpions reveals three colonizations of Southeast Asia. Sci. Rep. 10: 1–19.

Loria, S. F., and **L. Prendini**. **2021**. Burrowing into the forest: phylogeny of the Asian forest scorpions (Scorpionidae: Heterometrinae) and the evolution of ecomorphotypes. Cladistics 37: 109–161.

Lourenço, W. R. 2000. Panbiogéographie, les familles des scorpions et leur répartition géographique. Biogeographica 76: 21–39.

Lourenço, W. R. 2007. First record of the family Pseudochactidae Gromov (Chelicerata, Scorpiones) from Laos and new biogeographic evidence of a Pangaean palaeodistribution. C.R. Biol. 330: 770–777.

Lourenço, W. R. 2012. The genus *Vietbocap* Lourenço and Pham, 2010 (Scorpiones: Pseudochactidae); proposition of a new subfamily and description of a new species from Laos. C. R. Biol. 335: 232–237.

Lourenço, W. R. 2017. Second record of the genus *Troglokhammouanus* Lourenço, 2007 from Laos, with the description of a new species (Scorpiones: Pseudochactidae). Acta Arachnol. 66: 19–24.

Lourenço, W. R., and **D.-S. Pham**. **2012**. A second species of *Vietbocap* Lourenço & Pham, 2010 (Scorpiones: Pseudochactidae) from Vietnam. C.R. Biol. 335: 80–85.

Lourenço, W. R., D.-S. Pham, T. H. Tran, and **T.-H. Tran. 2018**. The genus *Vietbocap* Lourenço & Pham, 2010 in the Thien Duong cave, Vietnam: a possible case of subterranean speciation in scorpions (Scorpiones: Pseudochactidae). C.R. Biol. 341: 264–273.

Maddison, W. P., and **D. R. Maddison. 2019**. Mesquite: a modular system for evolutionary analysis. Version 3.70. Computer software and documentation. Available from http://www.mesquiteproject.org

Matzke, N. J. 2013a. BioGeoBEARS: BioGeography with Bayesian (and Likelihood) evolutionary analysis in R Scripts. R Package. Version 0.2.1. Available from http://CRAN.R-project.org/package=BioGeoBEARS

Matzke, N. J. 2013b. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. Front. Biogeogr. 5: 242–248.

Matzke, N. J. 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. Syst. Biol. 63: 951–970.

Metcalfe, I. 2011. Tectonic framework and Phanerozoic evolution of Sundaland. Gondwana Res. 19: 3–21.

Metcalfe, I. 2013. Gondwana dispersion and Asian accretion: tectonic and palaeogeographic evolution of eastern Tethys. J. Asian Earth Sci. 66: 1–33.

Miao, Y. F., X. M. Fang, F. L. Wu, M. T. Cai, C. H. Song, Q. Q. Meng, and **L. Xu. 2013**. Late Cenozoic continuous aridification in the western Qaidam Basin: evidence from sporopollen records. Clim. Past 9: 1863–1877.

Miller, M. A., W. Pfeiffer, and **T. Schwartz. 2010**. Creating the CIPRES Science Gateway for inference of large phylogenetic trees, pp. 1–8. *In* Proceedings of the Gateway Computing Environments Workshop (GCE). 14 Nov. 2010, New Orleans, LA.

Miyahigashi, A., K. Ueno, T. Charoentitirat, Y. Sera, Y. Kamata, and **A. Sardsud. 2010**. Late Carboniferous-early Permian foraminiferal assemblages from the Doi Chiang Dao limestone in the Inthanon Zone, northern Thailand, pp. 94–98. *In* 6th Symposium of the International Geological Correlation Programme Project Vol. 516. 9–14 Nov. 2010, Kuala Lumpur, Malaysia.

Monod, L., and **L. Prendini. 2015**. Evidence for Eurogondwana: the roles of dispersal, extinction and vicariance in the evolution and biogeography of Indo-Pacific Hormuridae (Scorpiones: Scorpionoidea). Cladistics 31: 71–111.

Montenat, C. 2009. The mesozoic of Afghanistan. GeoArabia 14: 147–210.

Morley, R. J. 2012. A review of the Cenozoic palaeoclimate history of Southeast Asia, pp. 79–114. *In* D. Gower, K. Johnson, J. Richardson, B. Rosen, L. Rüber, S. Williams (eds.), Biotic evolution and environmental change in Southeast Asia. Cambridge University Press, Cambridge, U.K.

Morley, R. J. 2018. Assembly and division of the South and South-East Asian flora in relation to tectonics and climate change. J. Trop. Ecol. 34: 209–234.

Nguyen, H., V. P. Vu, and **H. Limbert. 2012**. Cave systems in Phong Nha-Ke Bang area. VNU J. Sci.: Earth Environ. Sci. 28: 217–229.

Nicolas, V., V. Herbreteau, A. Couloux, K. Keovichit, B. Douangboupha, and **J. -P. Hugot. 2012**. A remarkable case of micro-endemism in *Laonastes aenigmamus* (Diatomyidae, Rodentia) revealed by nuclear and mitochondrial DNA sequence data. PLoS One 7: e48145.

Niemiller, M. L., B. M. Fitzpatrick, and **B. T. Miller. 2008**. Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: *Gyrinophilus*) inferred from gene genealogies. Mol. Ecol. 17: 2258–2275.

Nixon, K. C. 2002. WinClada, ver. 1.00.08. Computer software and documentation. Available from http://www.diversityoflife.org/winclada

Nunn, G. B., B. F. Theisen, B. Christensen, and **P. Arctander. 1996**. Simplicity-correlated size growth of the nuclear 28S ribosomal RNA D3 expansion segment in the crustacean order Isopoda. J. Mol. Evol. 42: 211–223.

Nurtaev, B. S., O. G. Tsai, and **D. U. Kurbanova. 2021**. The structural positions of the ophiolite complex in the Earth crust of the western Tien-Shan. IOP Conf. Ser.: Earth Environ. Sci. 929: 012004.

Ojanguren-Affilastro, A. A., C. I. Mattoni, J. A. Ochoa, M. J. Ramírez, F. S. Ceccarelli, and **L. Prendini. 2016**. Phylogeny, species delimitation and convergence in the South American bothriurid scorpion genus *Brachistosternus* Pocock 1893: integrating morphology, nuclear and mitochondrial DNA. Mol. Phylogenet. Evol. 94: 159–170.

Paradis, E. 2013. Molecular dating of phylogenies by likelihood methods: a comparison of models and a new information criterion. Mol. Phylogenet. Evol. 67: 436–444.

Prendini, L. 2000. Phylogeny and classification of the superfamily Scorpionoidea Latreille 1802 (Chelicerata, Scorpiones): an exemplar approach. Cladistics 16: 1–78.

Prendini, L. 2001. Substratum specialization and speciation in southern African scorpions: the Effect Hypothesis revisited, pp. 113–138. *In* V. Fet, P. A. Selden (eds.), Scorpions 2001: *In memoriam* Gary A. Polis. British Arachnological Society, Burnham Beeches, Bucks, UK.

Prendini, L. 2004. Systematics of the genus *Pseudolychas* Kraepelin (Scorpiones: Buthidae). Ann. Entomol. Soc. Am. 97: 37–63.

Prendini, L., E. S. Volschenk, S. Maaliki, and **A. V. Gromov. 2006**. A 'living fossil' from Central Asia: the morphology of *Pseudochactas ovchinnikovi* Gromov, 1998 (Scorpiones: Pseudochactidae), with comments on its phylogenetic position. Zool. Anz. 245: 211–248.

Prendini, L., T. M. Crowe, and **W. C. Wheeler. 2003**. Systematics and biogeography of the family Scorpionidae (Chelicerata: Scorpiones), with a discussion on phylogenetic methods. Invertebr. Syst. 17: 185–259.

Prendini, L., P. Weygoldt, and **W. C. Wheeler. 2005**. Systematics of the *Damon variegatus* group of African whip spiders (Chelicerata: Amblypygi): evidence from behaviour, morphology and DNA. Org. Divers. Evol. 5: 203–236.

Prendini, L., O. F. Francke, and **V. Vignoli. 2010**. Troglomorphism, trichobothriotaxy and typhlochactid phylogeny (Scorpiones, Chactoidea): more evidence that troglobitism is not an evolutionary dead-end. Cladistics 26: 117–142.

Prendini, L., V. L. Ehrenthal, and **S. F. Loria. 2021**. Systematics of the relictual Asian scorpion family Pseudochactidae Gromov, 1998, with a review of cavernicolous, troglobitic, and troglomorphic scorpions. Bull. Amer. Mus. Nat. Hist. 453: 1–149.

R Core Team. 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from https://www.R-project.org

Rambaut, A., M. A. Suchard, D. Xie, and **A. J. Drummond. 2014**. Tracer 1.6. Available from http://beast.bio.ed.ac.uk/tracer

Ree, R. H., and **I. Sanmartín. 2018**. Conceptual and statistical problems with the DEC + J model of founder-event speciation and its comparison with DEC via model selection. J. Biogeogr. 45: 741–749.

Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3: 217–223.

Robinson, A. C. 2015. Mesozoic tectonics of the Gondwanan terranes of the Pamir plateau. J. Asian Earth Sci. 102: 170–179.

Romero, A., and **S. M. Green. 2005**. The end of regressive evolution: examining and interpreting the evidence from cave fishes. J. Fish Biol. 67: 3–32.

Ronquist, F., and **J. P. Huelsenbeck. 2003**. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.

Santibáñez-López, C. E., O. F. Francke, and **L. Prendini. 2014**. Phylogeny of the North American scorpion genus *Diplocentrus* Peters, 1861 (Scorpiones: Diplocentridae) based on morphology, nuclear and mitochondrial DNA. Arthropod Syst. Phylogeny 72: 257–279.

Sharma, P. P., R. Fernández, L. A. Esposito, E. González-Santillán, and **L. Monod. 2015**. Phylogenomic resolution of scorpions reveals multilevel discordance with morphological phylogenetic signal. Proc. R. Soc. B. 282: 20142953.

Sharma, P. P., C. M. Baker, J. G. Cosgrove, J. E. Johnson, J. T. Oberski, R. J. Raven, M. S. Harvey, S. L. Boyer, and **G. Giribet. 2018**. A revised dated phylogeny of scorpions: phylogenomic support for ancient divergence of the temperate Gondwanan family Bothriuridae. Mol. Phylogenet. Evol. 122: 37–45.

Siehl, A. 2017. Structural setting and evolution of the Afghan orogenic segment – a review, pp. 57–88. *In* M.-F. Brunet, T. McCann, E. R. Sobel (eds.), Geological evolution of Central Asian Basins and the Western Tien Shan Range. Geological Society, Special Publication 427, London, UK.

Simon, C., A. Francke, and **A. Martin. 1991**. The polymerase chain reaction: DNA extraction and amplification, pp. 329–355. *In* G. Witt, A. Johnson,

J. Young (eds.), Molecular techniques in taxonomy. Springer Verlag, New York, NY.

Soleglad, M. E., and **V. Fet**. **2001**. Evolution of scorpion orthobothriotaxy – a cladistic approach. Euscorpius 1: 1–38.

Soleglad, M. E., and **V. Fet**. **2003a**. High-level systematics and phylogeny of the extant scorpions (Scorpiones: Orthosterni). Euscorpius 11: ii + 1–175.

Soleglad, M. E., and **V. Fet**. **2003b**. The scorpion sternum: structure and phylogeny (Scorpiones: Orthosterni). Euscorpius 5: 1–34.

Soleglad, M. E., and **W. D. Sissom**. **2001**. Phylogeny of the family Euscorpiidae Laurie, 1896: a major revision, pp. 25–111. *In* V. Fet, P. A. Selden (eds.), Scorpions 2001: *In memoriam* Gary A. Polis. British Arachnological Society, Burnham Beeches, Bucks, UK.

Soleglad, M. E., **F. Kovařík**, and **V. Fet**. **2012**. A new species of *Pseudochactas* from Afghanistan (Scorpiones: Pseudochactidae). Bol. Soc. Entomol. Aragon. 50: 89–98.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688–2690.

Stamatakis, A. 2014. RaxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30: 1312–1313.

Stecher, G., **K. Tamura**, and **S. Kumar**. **2020**. Molecular evolutionary genetics analysis (MEGA) for macOS. Mol. Biol. Evol. 37: 1237–1239.

Sterling, E. J., and **M. M. Hurley**. **2008**. Biogeography of Vietnam, pp. 45–70. *In* E. J. Sterling, M. M. Hurley, D. M. Le (eds.), Vietnam: a natural history. Yale University Press, New Haven, CT.

Stockwell, S. A. 1989. Revision of the phylogeny and higher classification of scorpions (Chelicerata). Ph.D. dissertation, University of California, Berkeley, CA.

Surridge, A. K., **R. J. Timmins**, **G. M. Hewitt**, and **D. J. Bell**. **1999**. Striped rabbits in southeast Asia. Nature 400: 726–726.

Tamura, K., **M. Nei**, and **S. Kumar**. **2004**. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc. Natl. Acad. Sci. U.S.A. 101: 11030–11035.

Tamura, K., **G. Stecher**, **D. Peterson**, **A. Filipski**, and **S. Kumar**. **2013**. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30: 2725–2729.

Trajano, E. 2012. Ecological classification of subterranean organisms, pp. 275–277. *In* W. B. White, D. C. Culver (eds.), Encyclopedia of caves, 2nd ed. Academic Press, Waltham, MA.

Trajano, E., and **M. R. de Carvalho**. **2017**. Towards a biologically meaningful classification of subterranean organisms: a critical analysis of the Schiner-Racovitza system from a historical perspective, difficulties of its application and implications for conservation. Subterr. Biol. 22: 1–26.

Tu, V. T., **G. Csorba**, **T. Görföl**, **S. Arai**, **N. T. Son**, **H. T. Thanh**, and **A. Hasanin**. **2015**. Description of a new species of the genus *Aselliscus* (Chiroptera, Hipposideridae) from Vietnam. Acta Chiropt. 17: 233–254.

Valdez-Cruz, N. A., **S. Dávila**, **A. Licea**, **M. Corona**, **F. Z. Zamudio**, **J. García-Valdes**, **L. Boyer**, and **L. D. Possani**. **2004**. Biochemical, genetic and physiological characterization of venom components from two species of scorpions: *Centruroides exilicauda* Wood and *Centruroides sculpturatus* Ewing. Biochimie 86: 387–396.

Villesen, P. 2007. FaBox: an online toolbox for fasta sequences. Mol. Ecol. Notes 7: 965–968.

Volschenk, E. S., **C. I. Mattoni**, and **L. Prendini**. **2008**. Comparative anatomy of the mesosomal organs of scorpions (Chelicerata, Scorpiones), with implications for the phylogeny of the order. Zool. J. Linn. Soc. 154: 651–675.

Waltham, T., and **J. Middleton**. **2000**. The Khammouan karst of Laos. Cave Karst Sci. 27: 113–120.

Wang, X. D., **K. Y. Hu**, **Y. K. Shi**, **J. T. Chen**, **S. R. Yang**, **X. Y. Ye**, **X. M. Li**, **Y. F. Song**, **B. Chen**, and **X. L. Chang**, **et al**. **2021**. The missing upper Carboniferous in the Cimmerian continent: a critical review. Earth-Sci. Rev. 217: 103627.

Wessel, A., **P. Erbe**, and **H. Hoch**. **2007**. Pattern and process: evolution of troglomorphy in the cave-planthoppers of Australia and Hawai'i – preliminary observations (Insecta: Hemiptera: Fulgoromorpha: Cixiidae). Acta Carsologica 36: 199–206.

Wheeler, W. C., **P. Cartwright**, and **C. Hayashi**. **1993**. Arthropod phylogeny: a combined approach. Cladistics 9: 1–39.

Zhang, Y., and **S. Li**. **2013**. Ancient lineage, young troglobites: recent colonization of caves by *Nesticella* spiders. BMC Evol. Biol. 13: 1831–1810.

Zhou, Y. X. 1994. Earliest pollen-dominated microfloras from the early Late Carboniferous of the Tian Shan Mountains, NW China: their significance for the origin of conifers and palaeophytogeography. Rev. Palaeobot. Palynol. 81: 193–211.

Zhu, H. C., **S. Ouyang**, **J. Z. Zhan**, and **Z. Wang**. **2005**. Comparison of Permian palynological assemblages from the Junggar and Tarim Basins and their phytoprovincial significance. Rev. Palaeobot. Palynol. 136: 181–207.