



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Phylogenomics and evolutionary history of Oreocnide (Urticaceae) shed light on recent geological and climatic events in SE Asia

Citation for published version:

Wu, Z-Y, Milne, RI, Liu, J, Slik, JWF, Yan, Y, Luo, Y-H, Monro, A, Wang, WT, Hong, W, Kessler, P, Cadotte, MW, Nathan, R & Li, D-Z 2022, 'Phylogenomics and evolutionary history of Oreocnide (Urticaceae) shed light on recent geological and climatic events in SE Asia', *Molecular phylogenetics and evolution*, vol. 175, 107555. <https://doi.org/10.1016/j.ympev.2022.107555>

Digital Object Identifier (DOI):

[10.1016/j.ympev.2022.107555](https://doi.org/10.1016/j.ympev.2022.107555)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Molecular phylogenetics and evolution

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



1 **Phylogenomics and evolutionary history of *Oreocnide* (Urticaceae) shed light**
2 **on recent geological and climatic events in SE Asia**

3
4
5 **Zeng-Yuan Wu¹, Richard I. Milne⁴, Jie Liu^{1,2}, Ferry Slik³, Yan Yu⁵, Ya-Huang Luo², Alexandre K.**
6 **Monro⁶, Wan-Ting Wang¹, Hong Wang², Paul J. A. Kessler⁹, Marc W. Cadotte⁷, Ran Nathan⁸, De-**
7 **Zhu Li^{1*}**

8
9 ¹Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences,
10 Kunming, Yunnan 650201, China

11 ²Key Laboratory for Plant and Biodiversity of East Asia, Kunming Institute of Botany, Chinese Academy
12 of Sciences, Kunming, Yunnan 650201, China

13 ³Environmental and Life Sciences, Faculty of Science, Universiti Brunei Darussalam, Jalan Tungku Link,
14 Gadong, BE1410, Brunei Darussalam

15 ⁴Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh,
16 Edinburgh EH9 3JH, UK

17 ⁵Key Laboratory of Bio-Resources and Eco-Environment of Ministry of Education, College of Life
18 Sciences, Sichuan University, Chengdu, Sichuan 610207, China

19 ⁶Identification & Naming Department, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AE, UK

20 ⁷Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, ON, M5S 3B2,
21 Canada

22 ⁸Movement Ecology Laboratory, Department of Ecology, Evolution and Behavior, The Hebrew
23 University of Jerusalem, Jerusalem 91904, Israel.

24 ⁹University of Leiden Hortus botanicus Leiden, PO Box 9500, 2300 RA Leiden, The Netherlands.

25
26 ***Corresponding author:**

27 De-Zhu Li
28 132 Lanhei Road, Kunming, Yunnan 650201, China
29 E-mail: dzl@mail.kib.ac.cn;
30 Tel: +86-871-65223503;
31 Fax: +86-871-6521779

32
33 **Declarations of interest:** none.

34
35 **Author contributions:** DZL, ZYW and FS conceived the study; ZYW, HW, JL and
36 AKM did field work; ZYW conducted lab work and wrote the first draft; ZYW, JL,
37 YHL, YY and WTW analyzed data; RIM, MWC, RN and PJAK helped improve
38 language; all authors contributed the ideas presented in the paper and to revisions.

43 **Abstract**

44 Climate change and geological events have long been known to shape biodiversity,
45 implying that these can likewise be viewed from a biological perspective. To study
46 whether plants can shed light on this, and how they responded to climate change there,
47 we examined *Oreocnide*, a genus widely distributed in SE Asia. Based on broad
48 geographic sampling with genomic data, we employed an integrative approach of
49 phylogenomics, molecular dating, historical biogeography, and ecological analyses. We
50 found that *Oreocnide* originated in mainland East Asia and began to diversify ~6.06 Ma,
51 probably in response to a distinct geographic and climatic transition in East Asia at
52 around that time, implying that the last important geological change in mainland SE
53 Asia might be 1 Ma older than previously suggested. Around four immigration events
54 to the islands of Malesia followed, indicating that immigration from the mainland could
55 be an underestimated factor in the assembly of biotic communities in the region. Two
56 detected increases of diversification rate occurred 3.13 and 1.19 Ma, which strongly
57 implicated climatic rather than geological changes as likely drivers of diversification,
58 with candidates being the Pliocene intensification of the East Asian monsoons, and
59 Pleistocene climate and sea level fluctuations. Distribution modelling indicated that
60 Pleistocene sea level and climate fluctuations were inferred to enable inter-island
61 dispersal followed by allopatric separation, underpinning radiation in the genus. Overall,
62 our study, based on multiple lines of evidence, linked plant diversification to the most
63 recent climatic and geological events in SE Asia. We highlight the importance of
64 immigration in the assembly and diversification of the SE Asian flora, and underscore
65 the utility of plant clades, as independent lines of evidence, for reconstructing recent
66 climatic and geological events in the SE Asian region.

67
68 **Key words** Biogeography, evolutionary history, *Oreocnide*, phylogeny, SE Asia,
69 Urticaceae

70 **Introduction**

71 Climatic change, especially changes in temperature and precipitation regimes, and
72 geological events such as tectonic collisions, are important and interconnected factors
73 that profoundly affect the evolution and biogeography of biotas (Antonelli et al., 2018).
74 Climate change alters the geographical location of suitable climatic niches, resulting in
75 shifts in species distributions and/or local adaptation, whereas geological events can
76 fundamentally change the regional landscape, creating new niches and erecting or
77 removing dispersal barriers. Therefore, climate and geological changes can jointly
78 promote speciation and large-scale biotic interchange (Cody, Richardson, Rull, Ellis, &
79 Pennington, 2010; Vermeij, 1991), and affect the distribution of biodiversity across
80 scales (Craven, Knight, Barton, Bialic-Murphy, & Chase, 2019; Kissling et al., 2012;
81 Ricklefs, 2004). Such biological responses might, in turn, cause significant changes to
82 regional climate, with profound feedback effects on local biodiversity (e.g. Hoorn et al.,
83 2010; Liu et al., 2013). Biodiversity hotspots are hence commonly associated with areas

84 that have undergone recent (i.e. a few to tens of millions of years) changes in geological
85 or regional climate regimes (Favre et al., 2015). Likewise, global climate shifts affect
86 biogeographic and phylogeographic patterns in plant communities via sea-level
87 changes that are important for our understanding of distribution patterns of both aquatic
88 and terrestrial organisms (Guo et al., 2018; Slik et al., 2011; Woodruff, 2010; He et al.,
89 2019). Therefore, understanding the evolutionary history of biotic groups can provide
90 an additional source of evidence for elucidating the nature and timing of geological and
91 climatic events in their region of occurrence (Diaz et al., 2019; Mao et al., 2012; Zhao,
92 Xia, Cannon, Kress, & Li, 2016).

93 Present anthropogenic climate change is already reshaping species distributions,
94 biodiversity patterns and ecosystem structure (Hooper et al., 2012), leading to
95 observable range shifts (Nathan et al., 2011; Trakhtenbrot, Nathan, Perry, &
96 Richardson, 2005) and rising sea levels (IPCC, 2013) with potentially major impacts
97 upon biogeography (Pfeifer-Meister et al., 2016). However, much about how plants
98 respond to climate change over longer timescales remains uncertain. Understanding
99 the interaction between biological evolution and climate change is particularly relevant
100 in islands and archipelagos, which have long been considered natural laboratories for
101 the study of evolution (Lohman et al., 2011; Santos, Field, & Ricklefs, 2016).

102 Stretching from southern China to New Guinea (Gower et al., 2012), Southeast
103 Asia harbors four of the Earth's 25 major biodiversity hotspots (Myers, Mittermeier,
104 Mittermeier, Da Fonseca, & Kent, 2000), and ranks as one of the highest in the world
105 in terms of species richness and endemism (Cámara-Leret et al., 2020). This region lies
106 on the intersection of several tectonic plates and thus has both a complex climatic
107 history and very high rates of geological change (Robert Hall, 1998, 2002). Here, unlike
108 some other biodiversity hotspots, geological and climatic changes caused landmasses
109 to vary in size and interconnectedness over time. Changing sea levels due to Pleistocene
110 climate oscillations drastically altered the land areas and connectivity of SE Asia
111 (Meltzner et al., 2017; Voris, 2000). Our understanding of past sea levels and their
112 impacts upon island connections in SE Asia is rapidly increasing (Lambeck, Esat, &
113 Potter, 2002; E. J. Rohling et al., 2019). In SE Asia, a few studies have revealed effects
114 of geography, climate and sea levels upon both zoogeography (Clouse & Giribet, 2010;
115 Li & Li, 2018; Zarowiecki et al., 2014) and phytogeography for certain taxonomic
116 groups (Janssens et al., 2016; Sirichamorn, Thomas, Adema, van Welzen, & Parmakelis,
117 2014; Stelbrink, Albrecht, Hall, & von Rintelen, 2012). Nonetheless, the provenance
118 and dispersal direction of the biota in SE Asia remain controversial (Kooyman et al.,
119 2019), and the causal effects underlying plant biogeography there have rarely been
120 examined using multiple lines of evidence. Therefore, inference of geological and
121 climatic events from plant evidence in this region has rarely been possible so far.
122 Addressing this challenge requires integrating interdisciplinary lines of evidence
123 (Gower et al., 2012; Robert Hall, 2017), and using diverse analytical approaches
124 (Favre et al., 2015). An ideal subject for such analysis would be a widespread clade
125 across the whole region, for which detailed knowledge of the systematics and ecology
126 of all taxa exist.

127 The Angiosperm genus *Oreocnide* (Urticaceae) contains ten recognized species

128 distributed throughout SE Asia (Fig. 1), with highest diversity in continental south
129 China, the island of Taiwan and the Philippine Sea plate; it is ideal for inferring the
130 geological and climatic history of SE Asia. First, preliminary examination of *Oreocnide*
131 species distributions indicates distribution patterns closely related to specific
132 tectonomorphological features, making it a suitable plant model to investigate past
133 geological and climatic events. Second, potential capabilities for dispersal by rats and
134 local birds (Nago et al., 2019; Osuri et al., 2017) and for at least some species across
135 long distances in seawater (Wu et al., 2018), means that the genus could be
136 representative of typical modes and directions of dispersal for biota in SE Asia. Finally,
137 the genus consistently inhabits tropical and subtropical forests (Chen, Lin, Friis,
138 Wilmot-Dear, & Monro, 2003), where both the mean annual temperature and
139 precipitation are very high (Fig. 2). The genus would hence be profoundly affected by
140 expansions and contractions of this habitat type, and therefore its evolution and
141 biogeography should exhibit potentially strong signals from past environmental shifts.

142 Here, we aimed to determine how plants respond to the climate and geological
143 change in SE Asia, so we examined the evolutionary history of *Oreocnide*, using an
144 integrative approach to provide insights into the geological and climatic history of SE
145 Asia. Phylogenetic relationships and divergence times were estimated using genomic
146 data, with multiple calibration points, from which ancestral area reconstructions (AARs)
147 were deployed, and shifts in diversification rate were tested for. Finally, ecological
148 niche modeling (ENM) and ecological analysis were carried out to examine past,
149 present and future range shifts for all species of *Oreocnide*.

150 **2 Materials and Methods**

151 **2.1 Taxon sampling and DNA sequencing**

152 Because there are few taxonomic studies on *Oreocnide* to date, some species are still to
153 be verified as natural taxa, so we recognized ten species following The Plant List (2010)
154 and Flora of China (Chen, Lin, Friis, Wilmot-Dear, & Monro, 2003). A total of 63
155 individuals including all ten species and covering the entire distribution range
156 of *Oreocnide* were sampled, and assigned to species based on morphology. Each
157 species was sampled from as many localities as possible, across as much of its range as
158 possible. For seven species, we were able to collect multiple individuals from different
159 locations across their whole range, especially for the most widespread species *O.*
160 *rubescens*. From *O. integrifolia* it was only possible to obtain material from one region
161 of its range, i.e. Indochina (area B, see below), but we did gather seven samples of it
162 spread across that region. However, for *O. boniana* and *O. obovata*, due to their narrow
163 distribution and sequencing failures, only one individual was successfully sequenced
164 from each species. Outgroup taxa were selected based on our previous phylogenetic
165 studies of Urticaceae (Wu et al., 2018; Wu et al., 2013), and we included 23 species
166 from 17 genera which represented four main clades (Clade I-IV) of Urticaceae. A
167 complete list of accessions with sample number, collection localities, voucher specimen
168 numbers and GenBank accession numbers is given in Table S1.

169 Total genomic DNA was extracted from silica-gel dried leaves or herbarium
170 materials using the CTAB method (Doyle & Doyle, 1987). Library construction was
171 performed with NEBNext® Ultra™ II DNA Library Prep Kit for Illumina (New
172 England BioLabs) following the manufacturer's instructions. Sequencing was
173 implemented using the Illumina HiSeq X Ten platform, which produced 150 bp paired-
174 end reads. The expected sequencing quantity of each sample was ca. 4 Gigabyte.

175 **2.2 Sequence assembly, annotation and alignment**

176 Raw data of sequences were assembled firstly by *de novo* assembly which was
177 conducted in SPAdes v3.9.1 (Bankevich et al., 2012), setting the kmer length to 85-115
178 bp; followed by connection conducted with Bandage v0.8.1 (Wick, Schultz, Zobel, &
179 Holt, 2015). *Boehmeria umbrosa* (GenBank accession number MF990291) was used as
180 reference for assembling and annotation. Inverted repeat boundaries were determined
181 by blast, and verified by reads mapping in GENEIOUS v9.1.4 (Kearse et al., 2012). All
182 sequences were annotated by PGA (Qu, Moore, Li, & Yi, 2019). The quality of the final
183 plastome scaffolds was assessed and they were adjusted manually to correct for errors
184 or ambiguities. Using GENEIOUS, the coding sequences (CDS) of 78 protein-coding
185 genes were extracted from each sampled plastome, and 18S ribosomal RNA - internal
186 transcribed spacer 1 - 5.8S ribosomal RNA-internal transcribed spacer 2 - 28S ribosomal
187 RNA genes (18S-ITS1-5.8S-ITS2-26S) were also extracted. We aligned sequences in
188 GENEIOUS, and non-alignable regions were excluded from the alignments using
189 Gblocks version 0.91b (Castresana, 2000); the best alignment was selected by IQ-
190 TREE (Nguyen, Schmidt, von Haeseler, & Minh, 2015) with default settings.

191 **2.3 Phylogenetic analyses**

192 Phylogenetic analyses were conducted using four datasets, (a) complete plastomes, (b)
193 coding DNA sequences (CDS), (c) 18S-ITS1-5.8S-ITS2-26S (nrDNA) sequences and
194 (d) complete plastomes + nrDNA sequences. Both maximum likelihood (ML) and
195 Bayesian inference (BI) analyses were employed. The models of best fit for the single
196 or combined datasets were determined by jModelTest2 (Darriba, Taboada, Doallo, &
197 Posada, 2012). ML analyses were implemented in RAxML v8.2.12 (Stamatakis, 2014)
198 on the CIPRES Science Gateway v3.3 (Miler, Pfeiffer, & Schwartz, 2010). Tree support
199 was assessed through 1000 bootstrap replicates. The GTR + G substitution model was
200 used in the ML analyses. BI analyses were performed in MrBayes v3.2.2 (Ronquist et
201 al., 2012) on the CIPRES Science Gateway (Miler et al., 2010). Four Markov Chain
202 Monte Carlo (MCMC) chains were run for one million generations until the average
203 standard deviation of split frequencies fell well below 0.01. Each run started with a
204 random tree and sampled every 1000 generations. After discarding the first 25% of trees
205 as burn-in, a 50% majority-rule consensus tree was constructed from the remaining
206 trees to estimate posterior probabilities (PP). No well-supported conflicts were
207 observed among the chloroplast genomic and nuclear ribosomal phylogenies, therefore,
208 we combined the complete chloroplast genomes with the nrDNA datasets for further
209 analyses.

210

211 **2.4 Divergence time estimation**

212 **2.4.1 Calibration points selection**

213 Although no fossil is unequivocally assignable to *Oreocnide*, Urticaceae has an
214 extensive fossil record. Therefore, age estimates within *Oreocnide* were inferred using
215 a large-scale, time-calibrated dating method that covered the four main clades of the
216 Urticaceae (Wu et al., 2013). Such an approach allowed us to incorporate multiple fossil
217 calibrations and consequently reduced bias in the results, relative to using a single
218 calibration point. Three well-identified fossils that fell into different clades of
219 Urticaceae were used to calibrate the topology: a staminate flower of Forsskaleae, and
220 fossil achenes of *Pilea* and *Urtica*. Their minimum age constraints followed our
221 previous work on the whole family Urticaceae (Wu et al., 2018). Our sampling across
222 Urticaceae was unbalanced, which can affect topology and lead to bias in the dates
223 calculated (Milne, 2009). Therefore, we additionally used three secondary calibration
224 points taken from Wu et al. (2018). For the first, the most recent common ancestor
225 (MRCA) of Clade I + Clade IV was constrained using a normal distribution with a mean
226 of 54.9 Ma and a standard deviation of 1.0. For the second, the MRCA of Clade II +
227 Clade III was constrained with a normal distribution, a mean of 59.2 Ma and a standard
228 deviation of 1.0. For the third, the divergence of *Oreocnide* from its sister groups was
229 constrained with a normal distribution, a mean of 42.87 Ma and a standard deviation of
230 1.0.

231

232 **2.4.2 Molecular dating analysis**

233 A temporal framework for the evolution of *Oreocnide* was provided using two widely
234 adopted methods, the Bayesian method (BEAST) (Drummond & Rambaut, 2007) and
235 Penalized Likelihood (treePL) (Smith & O'Meara, 2012). First, to estimate divergence
236 time using BEAST, the best-fit evolution model (GTRmodel + gamma + invariable sites)
237 was selected using the Akaike information criterion (AIC) as implemented in
238 jModelTest (Darriba et al., 2012). An uncorrelated lognormal relaxed clock was
239 selected, and given the phylogenetic shape of *Oreocnide*, which had a very long stem
240 branch, we assumed that the evolution of *Oreocnide* had not only been affected by
241 speciation, but also by extinction. Therefore, we set the tree prior under the Speciation:
242 Birth-Death process (Bromham et al., 2018). Four independent Markov chain Monte
243 Carlo runs of 1×10^9 generations were performed, with sampling at every 1000th
244 generation. The adequacy of parameters was checked using Tracer v1.7 (Rambaut,
245 Drummond, Xie, Baele, & Suchard, 2018), noting effective sample size values > 200.
246 The first 20% of trees were discarded as burn-in. Then, the time-calibrated species tree
247 was summarized as a Maximum Clade Credibility (MCC) tree using TreeAnnotator
248 v1.7.5 (BEASTS package). Divergence times with 95% highest posterior density (HPD)
249 intervals were visualized in Figtree v1.4 (Rambaut, 2009).

250 For treePL, confidence intervals were generated using 1000 RAxML bootstrap
251 trees. The same calibration points as for BEAST were used, and the random subsample
252 and replicate cross-validation (RSRCV) analyses were conducted from 10^{20} to 10^{-20} to

253 determine the best smoothing value, which was found to be 0.00001. Confidence
254 intervals for the dating estimates were calculated from the 1,000 bootstrap replicates,
255 with branch lengths generated by RAxML by using TreeAnnotator as implemented in
256 BEAST (Drummond & Rambaut, 2007). Additionally, to assess the robustness of age
257 estimates by removing one or a series of calibration points, a reduced set of three
258 calibration points was used for this purpose in both BEAST and treePL analyses, i.e.
259 only the Forsskaoleae fossil and two secondary constraints (root of Urticaceae and the
260 crown node of *Oreocnide* + sister groups).

261

262 **2.5 Diversification and rate shifts**

263 To examine diversification through time and to identify shifts in the diversification rates
264 along branches in *Oreocnide*, we performed two different kinds of diversification rate
265 analyses, both employing the BEAST results with seven calibrations. Prior to
266 diversification analyses, each of the seven species resolved as monophyletic in our
267 phylogenetic analysis was pruned to one individual each, whereas the three species
268 determined to be paraphyletic (*O. trinervis*, *O. kwangsiensis*, and *O. tokinensis*) (Fig.
269 3), we pruned each distinct clade of each species to one individual each. Therefore, this
270 analysis considered 15 individual samples, each representing a particular species or part
271 thereof.

272 First, we conducted Bayesian analysis of macro-evolutionary mixtures (BAMM)
273 using BAMM v. 2.5.0 (Rabosky et al., 2014) to estimate speciation and extinction rates
274 through time and to identify shifts in diversification rate. In total, 5000000 generations
275 were run, and we sampled every 1000 generations using the chronogram generated from
276 BEAST as the input tree, with the outgroup pruned using APE (Paradis, Claude, &
277 Strimmer, 2004). Post-run analysis of BAMM output was performed using the
278 “BAMMTOOLS” package in R (Rabosky et al., 2014). The convergence was checked
279 by plotting the log likelihood trace of the MCMC output file and also by calculating the
280 effective sample size values (ESS) (> 200), and the first 10% of the sampled data were
281 discarded as burn-in. We tested the shifts in the diversification rate across the tree in
282 two ways: (1) by mapping speciation rates across all clades of the phylogeny; and (2)
283 by calculating the specific speciation, extinction, and net diversification rates for all
284 clades from BAMM output, then using these data to generate diversification through
285 time plots.

286 Second, we further examined lineage diversification rates of each clade using
287 MEDUSA (Brown, FitzJohn, Alfaro, & Harmon, 2018), which allowed clade-specific
288 birth–death models, and could detect regions in a time-calibrated phylogeny where
289 diversification rate shifts likely occurred (Hernández-Hernández, Brown,
290 Schlumpberger, Eguiarte, & Magallón, 2014).

291

292 2.6 Biogeographical inference

293 The biogeography of *Oreocnide* was inferred using a parsimony-based statistical
294 dispersal-variance (S-DIVA; Yu et al., 2010) and a likelihood analysis implemented in
295 the R package BioGeoBEARS (Nicholas Joseph Matzke, 2013; 2014), both analyses
296 were implemented in RASP v4.2 (Yu, Blair, & He, 2020). The analysis was based on
297 the MCC tree with the birth-death model and seven calibration points. To understand
298 the historical spatiotemporal evolution of *Oreocnide*, eight geological areas were
299 defined based on geological history of SE Asia (Hall, 2002) and the extant distribution
300 and endemism of *Oreocnide*: A, Eurasia; B, Indo-China and Malay Peninsula; C,
301 Taiwan; D, Sumatra and Java; E, Borneo; F, Philippines; G, Wallacea; H, New Guinea
302 and surrounding islands. Taiwan was recognized as a distinct geological entity because
303 it has an endemic *Oreocnide* species (*O. pedunculata*) and has been an island for most
304 of the ~6 million years since *Oreocnide* began to diversify. Conversely, we followed
305 Thomas et al (2012) in treating most of mainland Eurasia (other than Indochina) as a
306 single area. The remaining islands and groups thereof are separated by various
307 biogeographical dividing lines (e.g. see Van Welzen et al., 2011). To eliminate the
308 influence of the outgroup on the root ancestral area, all outgroups were pruned prior to
309 the analysis. *Oreocnide* species were pruned to 15 individual samples that represented
310 species or clades, in the same way as the diversification analysis (see above). Those
311 that represented species resolved as monophyletic were coded as present in all areas
312 where that species occurred. However, coding polyphyletic species in this way would
313 produce false results, assigning clade members to areas that clade had never occupied.
314 Hence for paraphyletic species (*O. trinervis*, *O. tonkinensis*, and *O. kwangsiensis*), each
315 representative individual was coded only for the distribution areas occupied by the
316 clade that it represented. An addition issue was that *O. rubescens*, although
317 monophyletic, has a distribution covering all areas. Hence coding this species as present
318 throughout the range potentially meant the loss of infraspecific data concerning its
319 likely point of origin, which might in turn affect analysis of the whole genus. To test
320 this possibility, we ran an alternate analysis in which every accession was coded
321 separately with its collection locality. However, it yielded very similar results (data
322 not shown).

323 S-DIVA analyses were performed based on a sample of 1000 random BEAST trees,
324 with ancestral areas limited to no more than two areas. Biogeographical results were
325 summarized on the BEAST consensus tree (fully bifurcating) using majority rule. For
326 the BioGeoBEARS analysis, a comparison among models was firstly executed using
327 AIC to determine the best fit model among the six available BioGeoBEARS models
328 (DEC, DEC+J, DIVALIKE, DIVALIKE+J, BAYAREALIKE, BAYEARELIKE+J). We
329 performed estimations by delimiting two time slices in concordance with major plate
330 tectonic events and environmental transformations: (1) before 5 Ma (Taiwan has not
331 yet formed, Sunda islands were joined together); (2) 5–0 Ma (Taiwan Island formed,
332 Sunda islands separated). Three possible dispersal probabilities (0.01, 0.5, and 1.0)
333 were assigned between all adjacent areas based on the connectivity and distance of

334 different land masses during each time slice (Table S2). The maximum number of areas
335 in the distribution range of each species was set to three.
336

337 **2.7 Ecological analyses**

338 (1) Phylogenetic signal test

339 Incorporating phylogenetic information into ecology is enlightening because it allows
340 ecological questions to be addressed in an evolutionary context, and vice versa, leading
341 to a deeper understanding of the processes that give rise to patterns of biological
342 diversity (Webb, Ackerly, McPeck, & Donoghue, 2002). To test whether bioclimatic
343 niches were more similar for related species, we quantified the phylogenetic signals of
344 species bioclimatic variables using Blomberg's K values (Blomberg, Garland Jr, & Ives,
345 2003) with 9999 random replicates using “phytools” package in R software (Revell,
346 2012). Higher values for Blomberg's K indicated bioclimatic variables that were
347 strongly linked to the evolutionary history of the species, and $K = 1$ indicated a
348 phylogenetic pattern compatible with a Brownian motion model of evolution, whereas
349 $K = 0$ is indicative of truly random patterns (Blomberg et al., 2003).

350

351 (2) Ecological niche modelling

352 Ecological niche modelling (ENM), paleoclimatic inference, or geological
353 reconstructions can be used to generate biogeographic hypotheses for the distribution
354 of suitable habitats or island terranes at a given point in the past (Lohman et al., 2011).
355 To predict the potential distribution of *Oreocnide* during the Last Interglacial (LIG),
356 ecological niche modelling was performed for the Last Glacial Maximum (LGM), the
357 present and the future. For the future, we took the year 2070 under the pessimistic
358 RCP8.5 scenario (IPCC 2013) which seems applicable given current geopolitics,
359 whereas partial emissions cuts would see this scenario being achieved later, e.g. 2100
360 (Meinshausen et al., 2011; Schwalm et al., 2020). Our sampling was limited for the
361 three paraphyletic species *O. trinervis*, *O. kwangiensis* and *O. tokinensis*, especially
362 with only a few individuals from the last two species, and hence treating the clades as
363 distinct species in this analysis would be premature, and moreover their ranges would
364 be severely under-sampled. Therefore, for this analysis, we used the existing taxonomy,
365 while acknowledging that the results would be improved upon by wider sampling of
366 these taxa. Localities were also added from the literature, and from our own fieldwork.
367 All occurrences were carefully evaluated, and we pruned those where either the
368 identification or the locality (e.g. occurrence in oceans) was likely to be incorrect. Once
369 this pruning was complete, additional records were removed so that we only kept
370 records of the same species that were separated from each other by ≥ 5 km. The final
371 dataset consisted of 1329 unique sample locations. For each sample location, we
372 obtained climatic data from the WorldClim database v1.4 (Hijmans, Cameron, Parra,
373 Jones, & Jarvis, 2005), comprising 19 bioclim variables with 2.5 minutes resolution (ca.
374 5 km at the equator). To reduce the influence of collinearity among bioclimatic variables
375 on model prediction, we kept only those bioclimatic variables with Pearson's
376 correlation coefficient $r < 0.8$. Hence each species had different variables retained in

377 the downstream ENM analysis (Table 1). We then limited our model extent to the
378 distributional range of each *Oreocnide* species with a buffer zone of 200 km. ENM was
379 conducted using the MAXENT v3.4.1 software package (Phillips, Anderson, &
380 Schapire, 2006), with 10 bootstrap replicates, employing 75% of the localities to train
381 the model, and applying the ‘Equal training sensitivity and specificity threshold’ rule
382 (Liu, Berry, Dawson, & Pearson, 2005) to define the minimum threshold of suitable
383 habitat.

384

385 (3) Spatial distribution dynamic analysis

386 We conducted spatial distribution dynamic analysis (SDDA), to post analyze the
387 distribution shifts of each species in response to climate change, based on the results
388 produced by MAXENT, employing ArcGIS v10.3 (ESRI, Redlands, CA, USA). First,
389 we calculated the potential distribution area of each species across the four epochs
390 defined above (LIG, LGM, present and future). We used the ‘Equal training sensitivity
391 and specificity threshold’ rule to decide the threshold of occurrence in the prediction of
392 species distributions, and this was then used to reclassify the raster into suitable and
393 unsuitable habitats, with the help of Spatial Analyst Tools’s Reclassify tool. Then, by
394 deleting regions determined to be unsuitable, we defined suitable regions using
395 Conversion Tools’s Raster to polygon. Subsequently, we projected the suitable regions
396 from the original WGS 1984 geographic coordinate system into Asia North Albers
397 Equal Area Conic, so that we could calculate the area of potential distribution for each
398 species.

399 Second, the mean elevation was calculated for each species across each time period
400 (LIG, LGM, present and future), using Spatial Analyst Tools’s Extract by mask. Third,
401 to explore potential shifts for each species across each time period (LIG, LGM, present
402 and future) in latitude and longitude, we determined the mean latitudinal and
403 longitudinal center of the distribution range of each species in each time period using
404 the Spatial Statistics Tools’s Mean center.

405

406 **3 Results**

407 **3.1 Characteristics of datasets**

408 A total of 76 plastomes were newly sequenced in this study, of which 62 were of
409 *Oreocnide*, to which were added one, *Oreocnide frutescens*_GLGE12243, that had been
410 published before (Wang, Milne, Du, Liu, & Wu, 2020). The plastome size ranged from
411 156,663 bp (*O. rubescens*_W322) to 157,464 bp (*O. rubescens*_W310) (Table S4). The
412 structure of the plastomes of *Oreocnide* was conservative in overall size and in the order
413 and size of each gene and intergene region, i.e. all plastomes possessed the typical
414 quadripartite structure with two identical copies of a large inverted repeat (IR) separated
415 by a large single copy (LSC) and a small single copy (SSC). The LSC in *Oreocnide* was
416 ca. 86,560 bp, the IR ca. 25,697 bp, and the SSC ca. 19,038 bp. The GC content of all
417 individuals of *Oreocnide* was ca. 36.3%. All reconstructed plastomes of *Oreocnide*

418 contained 112 different genes, including four distinct rRNA, 30 tRNA, and 78 protein-
419 coding genes (PCG). The aligned consensus length of the complete plastome was
420 160,544 bp, with one of the inverted repeat (IR) regions excluded from all of the
421 phylogenetic analyses. After cleaning by Gblocks, the length became 138,324 bp, with
422 29% variable and 17.5% PI sites. The aligned length of the nrDNA dataset was 7,462
423 bp, with the length of 18S, nrITS and 26S rDNA being 2,396 bp, 992 bp, and 3,955 bp,
424 respectively.

425 For the nr DNA dataset, sequences of 83 individuals were generated, including 59
426 ingroup and 24 outgroups. The numbers of variable sites and PI sites of the nrDNA
427 dataset were 20% and 12.6%, respectively.

428 **3.2 Phylogenetic relationships within *Oreocnide***

429 Sequence characteristics and best-fit model determined by jModelTest for all four
430 datasets are given in Table S5. ML and BI analyses gave almost identical topologies for
431 each dataset. For chloroplast data, support values for the vast majority of the clades
432 were high, and the monophyly of *Oreocnide* received maximum support, whereas the
433 branch lengths within the genus were relatively short (Figs. 3, S1 & S2). There were,
434 however, three notable conflicts between relationships shown by nrDNA data (18S-
435 ITS1-5.8S-ITS2-26S; Fig. S3) and those based on plastid data. First, *O. trinervis*_W350
436 and all individuals of *O. rubescens* formed a monophyletic clade (Clade II) with
437 maximum support in the CDS and plastome topologies (Figs. S1 & S2), but *O.*
438 *trinervis*_W350 was not in Clade II according to nrDNA sequences (Fig. S3). Second,
439 plastid results showed that individuals from *O. kwangsiensis*, *O. tokinensis*, *O. boniana*,
440 *O. serrulata*, *O. trinervis*, *O. frutescens*, *O. obovata*, and *O. integrifolia* formed a
441 monophyletic clade (Clade III), sister to Clade IV (Figs. 3 & S1), but this clade seemed
442 to be polyphyletic according to nrDNA data (Fig. S3). Finally, all sampled individuals
443 of *O. integrifolia* formed two well supported clades in plastid trees (Figs. 3 & S1), one
444 comprising W297 and W299 while the other formed by the remaining five individuals;
445 but these clades were well supported as sister to each other in the nrDNA tree (Fig. S3)
446 and the combined tree (Fig. 2). Visual checks of phylogenetic trees from the complete
447 plastome dataset (Fig. S1), and the combined dataset (Fig. 3), generated by ML and BI
448 analyses, revealed no obvious differences between them that had statistical support. The
449 support values for the combined dataset were relatively higher than those of the
450 complete plastome dataset. Therefore, the relationships described here are from the
451 combined tree.

452 The combined dataset produced strong support for monophyly of *Oreocnide* (Fig.
453 3). The ingroup fell into four major clades, here termed Clade I, II, III, and IV, with
454 clade composition as follows. Clade I included all sampled individuals of *O.*
455 *pedunculata*, and was strongly supported to be the earliest diverging clade of the genus.
456 Clade II comprised a large subclade which contained all samples of *O. rubescens*, and
457 one individual (W350) of *O. trinervis* (Fig. 3). Within Clade II, *O. rubescens* was split
458 across two subclades (subclade 2A, 2B), each of which had maximum support. Within
459 Clade II, all sampled individuals belonging to subclade 2A were from the west side of
460 Wallace's line, whereas all those in subclade 2B were either strictly from the east of

461 Wallace's line, or from the Philippines (Table S1). Within Clade III, subclade 3A had
462 strong support and consisted of all sampled individuals of *O. integrifolia*. Subclade 3B
463 was a poorly resolved subclade, in which *O. frutescens* was inferred to be a well-
464 supported monophyletic species, however, individuals from *O. kwansiensis*, *O. boniana*,
465 *O. serrulata*, *O. tonkinensis*, *O. obovata*, *O. trinervis* were nested within subclade 3B.
466 Sister to Clade III, Clade IV comprised two subclades (4A, 4B), each fully supported
467 as monophyletic. Of these, 4A comprised two individuals of *O. trinervis*, whereas 4B
468 comprised eight individuals of *O. trinervis*.

469 **3.3 Divergence time estimation and biogeographical analyses**

470 Similar estimates were obtained using different dating methods and with alternative
471 calibration scenarios and settings (Fig. 4); BEAST with birth/death priors estimated
472 older ages for the deep nodes within *Oreocnide* while treePL estimated older ages for
473 upper nodes, but all ages estimated from treePL fell within the 95% HPD of the ages
474 from BEAST (Table S6, Figs. 4, S4, S5, S6 and S7). Furthermore, between programs
475 and between calibration methods (3 vs 7 calibration points), the greatest difference in
476 mean age for any major node was 0.69 Million-years ago (Ma) (Table S6). However,
477 results from BEAST using yule priors were substantially different, with crown
478 divergence in *Oreocnide* estimated at 32.2 Ma (Fig. S8), well outside the confidence
479 ranges of dates for the birth-death model (1.27-13.72) or treePL (4.11-7.17). Hence the
480 birth-death model appears more suitable and accurate for *Oreocnide* based on both
481 corroborations of its dates by treePL, and the very long phylogenetic branch of the
482 genus that made the yule model (which excludes extinction) biologically unlikely.
483 Therefore, dating results were described throughout the rest of this section based only
484 on BEAST with the birth-death model and seven calibration points (Strategy A in Fig.
485 4, Fig. S4).

486 In the biogeographical reconstructions, the analysis of BioGeoBEARS (Fig. 5)
487 showed clearer origin area and dispersal routes compared to S-DIVA (Fig. S9).
488 Therefore, also considering the distinguished advantages of BioGeoBEARS, which
489 cannot be achieved by any other current software based on a parsimony framework or
490 Bayesian approach (Ding et al., 2019), we described the results here entirely based on
491 BioGeoBEARS.

492 In all three RASP model pairs, +J variants provided significantly different results
493 from their basic alternatives. BAYAREALIKE+J showed the lowest AICc and highest
494 AICc_wt and likelihood (LnL) values, all indicating it to be the best supported model
495 for our data (Table S7). Applying this model, crown diversification in *Oreocnide* began
496 in the late Miocene, c. 6.06 Ma (95% highest posterior density (HPD) 1.27-13.72), and
497 ancestral area reconstruction indicated a clear origin for *Oreocnide* on the Eurasian
498 mainland (Fig. 5). Clade I of *Oreocnide* began to diversify in the late Miocene (6.06
499 Ma), and in this clade, one dispersal from the mainland East Asia to Taiwan occurred,
500 making it the earliest detected dispersal event in the genus.

501 The second earliest detected dispersal within *Oreocnide* occurred in the MRCA of
502 Clade II at 2.19 (0.75-3.88) Ma; this was from mainland East Asia (areas A and B) to
503 somewhere in New Guinea (Fig. 5). Following this, numerous dispersal events within

504 Clade II were indicated; one directional arrival was from mainland into Wallacea (G),
505 followed by later dispersals from Wallacea to the New Guinea (H), and then from there
506 to the Philippines (F) and Taiwan (C) (Fig. 5). The other directional dispersal events
507 with Clade II were from the mainland East Asia to Borneo (E), and later to Sumatra +
508 Java (D) and then from there back to the Eurasia mainland (A+B).

509 The MRCA of Clade III likely to have inhabited the mainland East Asia, and three
510 dispersal events occurred in this clade. In the earliest event, the MRCA of *O. integrifolia*
511 was dispersed from the mainland to Sumatra + Java (D) around 1.19 (0.42-2.12) Ma;
512 the second dispersal from mainland East Asia to Borneo occurred in the MRCA of *O.*
513 *frutescens* around 0.22 (0.17-0.94) Ma; and most recently the MRCA of *O.*
514 *trinervis*_O20 dispersed from mainland East Asia to Taiwan (area C) around 0.62 (0.14-
515 1.23) Ma (Fig. 5).

516 The MRCA of Clade IV dispersed from the mainland to Borneo around (4.15-)
517 2.33-1.59 (-0.60), and then one daughter lineage dispersed from Borneo to Philippines
518 (Fig. 5).

519 **3.4 Diversification of *Oreocnide***

520 BAMM identified two rate shifts, the earlier of which occurred at 3.13 Ma, whereas the
521 later one was detected at 1.19 Ma (Fig. 6). Diversification rate through time analyses
522 obtained with MEDUSA also detected one distinguishable shift in diversification
523 relative to background level (Fig. 6).

524 **3.5 Ecological analyses**

525 (1) Phylogenetic signal test

526 Most bioclimatic variables had no phylogenetical signal, indicating that closely related
527 species did not exhibit more similar bioclimatic niches to one another than did less
528 closely related species. However, two bioclimatic variables (Precipitation of driest
529 month bio14 and Precipitation of driest quarter bio 17) had significant phylogenetic
530 signals ($K=0.535$ and 0.468 respectively, $P<0.01$) (Table S3).

531 (2) Ecological niche modelling

532 Each *Oreocnide* species had an area under the receiver operating characteristic curve
533 (AUC) value of ≥ 0.838 (Table 1), indicating a better than random prediction. Current
534 distribution predictions were generally good representations of the observed
535 distributions of all species, with *O. rubescens* showing the largest distribution range
536 size, and *O. pedunculata* the smallest. Palaeodistribution modeling showed that the
537 changes in the distribution of all species were highly influenced by the glacial and
538 interglacial periods (Fig. 5). Furthermore, all species except *O. kwangsiensis*, had their
539 broadest distribution ranges during the LGM, relative to LIG, Present and Future (Fig.
540 5). Moreover, ecological niche modelling predicts that future distributions would
541 become narrower than present for all species, except for *O. trinervis* whose range would
542 become broader.

543 (3) Spatial distribution dynamic analysis

544 The SDDA results concurred with the results above, supplying further details. The

545 potential distribution areas of all species were greatly influenced by glacial and
546 interglacial periods, with *O. integrifolia*, *O. trinervis* and especially the widespread *O.*
547 *rubescens* having an increased area during the LGM, relative to both present and LIG.
548 However, *O. serrulata*, *O. kwangsiensis*, *O. boniana* and *O. obovata* were all reduced
549 to below 25 km² during the LGM according to this analysis (Fig. 6). Most species were
550 predicted to decline by 2070, though the decline will be minimal in *O. kwangsiensis*,
551 whereas a slight increase was forecasted for *O. rubescens*; *O. trinervis* was predicted to
552 expand dramatically (Fig. 6), but this might not be accurate as this taxon appeared to
553 be comprised of three unrelated lineages (Fig. 3).

554 Unsurprisingly, regarding mean elevation, most species moved to lower altitudes
555 during the LGM, relative to the present, with the striking exception of *O. kwangsiensis*,
556 which moved higher. Furthermore, three species occupied lower mean elevations in the
557 LIG than the LGM, these were *O. kwangsiensis*, *O. obovata* and *O. serrulata*; all others
558 were higher, and often higher than the present elevational distribution as well. Most
559 species were predicted to increase in altitude by at least 200m by 2070, with smaller
560 increases predicted for *O. serrulata*, *O. rubescens* and *O. boniana*, while *O.*
561 *kwangsiensis* and the polyphyletic *O. trinervis* were predicted to move downwards (Fig.
562 6).

563 In the horizontal direction, based on mid-points of species distributions in each
564 period, the distributions of species were strongly affected by glaciation cycles and
565 associated climatic and land configuraional changes. Curiously, predicted migration
566 directions between LIG and LGM were more often east to west than towards the equator.
567 However, most were predicted to have moved northwards between the LIG and the
568 present, and were predicted to shift even further northwards in the future (Fig. 8).

569 **4 Discussion**

570 **4.1 The timing of diversification of *Oreocnide* reflects the most recent important** 571 **geological event in SE Asia**

572 Our data strongly supports the monophyly of *Oreocnide*, and suggests an origin in
573 mainland Eurasia. *Oreocnide* diverged from its sister group (a large clade of genera;
574 Fig. S3, Wu *et al.* 2018) around 42.59 Ma, but extant lineages did not begin to diversify
575 until the late Miocene (6.06 Ma). As we known, there were three important periods in
576 SE Asian biogeography: at about 45 Ma, 25 Ma and 5 Ma, during each of these periods,
577 plate boundaries and motions changed, probably as a result of major tectonic collision
578 events (Robert Hall, 1998, 2002). The plate reorganization of 45 Ma and 25 Ma may be
579 related to the India-Asia collision, and New Guinea's collision with the East Phillipine-
580 Halmahera-South Caroline Arc system, respectively (Robert Hall, 2002). The most
581 recent event was suggested by Michaux (2010) to have been around or perhaps slightly
582 earlier than 5 Ma, but the dates are not precisely known. Geological studies certainly
583 revealed that during the early Pliocene, plate motions and boundaries changed
584 dramatically in SE Asia, and most islands in the central Malesian region started to
585 emerge, possibly as a consequence of Pacific Plate motion changes, arc-continent

586 collision in Taiwan, or other boundary changes at Pacific margin, e.g. in the Philippines
587 (Hall, 1998, 2002; Van Welzen, Parnell, & Slik, 2011). Our results were consistent with
588 this drastic tectonic possibly having occurred 1 Ma earlier compared to Hall's (1998,
589 2002) estimate. This important tectonic event probably underpinned the diversification
590 of *Oreocnide*, by changing the regional landscape, creating new niches and erecting or
591 removing dispersal barriers, which might have allowed the diversification rate to
592 outstrip the extinction rate (which had not been the case before 6.06 Ma). Hence the
593 initiation of diversification within *Oreocnide* is one of the few detected case of plant
594 diversification linked to the most recent major geological event in SE Asia (Thomas et
595 al., 2012; Janssens et al., 2016).

596 **4.2 The divergence and range dynamics of *Oreocnide* reflected the Pleistocene** 597 **climate fluctuations in SE Asia**

598 Most diversification in *Oreocnide* occurred from 3.13 Ma onwards, so, the geological
599 movements underlying Wallace's line appear to not be critical for *Oreocnide*. Instead,
600 dispersal patterns in *Oreocnide* are much more likely have been affected by existing
601 topography in SE Asia, especially land-sea configurations, plus Pleistocene climate and
602 sea level fluctuations.

603 BMM analyses detected an increase in diversification rate around 3.13 Ma (Fig.
604 6), which likely represented divergence events within mainland SE Asia and the arrival
605 of lineages into the Malesian archipelago (Fig. 6 and see below). This coincided with
606 the intensification of the East Asian summer monsoon (An, Kutzbach, Prell, & Porter,
607 2001), which profoundly affects summer rainfall across a wide area, including parts of
608 Malesia (Cheng et al., 2016). Given that *Oreocnide* occupied hot and humid habitats
609 (Fig. 2) (Chen et al., 2003), these events might have driven diversification in the genus
610 by opening up new habitats. Indeed, our phylogenetic signal analysis revealed that
611 precipitation and temperature were relatively conservative (Table S3), hence *Oreocnide*
612 lineages would appear to have tracked suitable conditions where they occurred, rather
613 than exploiting new ones, suggesting that speciation might have been more allopatric
614 than habitat-driven.

615 Furthermore, the onset of Pleistocene sea level fluctuations might have facilitated
616 *Oreocnide* dispersal between the islands as discussed above, and then allopatric
617 speciation might have been triggered when the land connections were submerged (Shih,
618 Hung, Schubart, Chen, & Chang, 2006; Thomas et al., 2012; Zhang et al., 2016).
619 Indeed, a second diversification rate shift in *Oreocnide* 1.19 Ma (Fig. 6), corresponded
620 with a global temperature peak at around 1.19 Ma, during an unusually warm
621 interglacial (see Fig. 6; Rohling *et al.* 2014); this might have increased isolation by
622 pushing species to higher altitudes (Fig. 6). However, other than in subclade 3B on the
623 Eurasian mainland, there have been no speciation events within *Oreocnide* since 2 Ma
624 (Fig. 5); detected diversification after this time is infraspecific differentiation.
625 *Oreocnide* achenes have fleshy receptacles, that might promote inter-island by birds or
626 rats (Balakrishnan, 2011; Nago et al., 2019; Osuri et al., 2017), whereas some
627 *Oreocnide* propagules appear capable of long-distance dispersal in seawater (Wu *et al.*
628 2018). These factors might have acted to prevent allopatric speciation during most of

629 the Pleistocene, and to maintain gene flow within the widespread species *O. rubescens*
630 and *O. integrifolia*.

631 Glacial maxima would also have influenced, via cooler climates, the altitudinal
632 distribution of species and ecosystems (Fernández-Palacios, 2016). This led, according
633 to our ENM analysis, to expanded distributions for all *Oreocnide* species during the
634 LGM relative to the present (Fig. 7). Range shifts continued through the LIG and LGM
635 up to the present (Figs. 7, 8 & 9), but the genetic signatures of these events would only
636 be detectable through more intensive sampling at the infraspecific level. Conversely,
637 most species are predicted to move upwards, northwards and/or westwards by 2070,
638 according to SDDA results (Figs. 8 & 9).

639 **4.3 Multiple immigration events into Malesia, and their drivers**

640 Ancestral area reconstruction analysis clearly indicated that all four major clades
641 diverged from one another within mainland Eurasia. At least six dispersals into Malesia
642 then followed, consistent with the general direction of biotic movement in this region
643 (Matuszak, Muellner-Riehl, Sun, & Favre, 2016; Thomas et al., 2012; Vermeij, 1991).
644 The first involved the MRCA of Clade II, (5.59-) 3.13-2.19 (-0.75) Ma, into New
645 Guinea (Fig. S1). Considering the relative geological stability of SE Asia during this
646 period, this dispersal might have been promoted by the intensification of the East Asian
647 summer monsoon at around this time (An et al., 2001), when the Eurasian climate
648 environment was particularly humid with abundant rainfall, increasing the availability
649 of the hot wet habitats favoured by *Oreocnide* (Fig. 2) (Chen et al., 2003), which
650 promoted rapid radiation and intra-species competition, which can drive emigration. At
651 the same time, the whole Malesia region has a climate that is suitable for the growth of
652 *Oreocnide*. The existence of a large area of suitable habitat into which propagules could
653 disperse makes long distance dispersal more likely to occur. In fact, the large area and
654 relative geological stability of Borneo, coupled with the persistence there of large tracts
655 of rainforest during Pleistocene glacial maxima might have made it a cradle of
656 diversification (de Bruyn et al., 2014; Robert Hall, 2012; Lohman et al., 2011), e.g. in
657 the exceptionally diverse genus *Cyrtandra* (Gesneriaceae) (Atkins et al., 2020). Either
658 way, after the divergence of *O. trinervis* C (*O. trinervis*_W350), the remainder of this
659 clade became *O. rubescens*, which spread via at least seven dispersal events across the
660 entire region, including a back-dispersal to mainland SE Asia via Sumatra (Fig. 5).
661 Considering this first dispersal into Malesia predated the Pleistocene, and crossed wider
662 ocean barriers to Borneo, seeds of this family appear to be capable of traversing these
663 sorts of distances (~1000 km or more) in seawater (Wu et al., 2018).

664 Clade IV (*O. trinervis*_A+B) was the second lineage to reach Malesia 1.59 (0.60-
665 2.83) Ma, arriving either into Borneo from which it then reached the Philippines.
666 Bornean and Philippine material diverged 1.59 (0.60-2.83) Ma, possibly due to
667 dispersal between the archipelagos being aided by a phase of low sea levels at that time.
668 Moreover, in Clade III, at least two further dispersals were indicated, each involving
669 species that originated on the mainland then dispersed into the Malesian Archipelago.
670 These were *O. integrifolia* (4.15-) 1.19-0.62 (-0.14) Ma, and *O. frutescens* (0.72-) 0.39-
671 0.22 (-0.06) Ma (Fig. 5).

672 According to ‘colonization window hypothesis’ (Carine, 2005), opportunities for
673 island colonization are temporally constrained to discrete waves of colonization, for
674 example due to the appearance of land bridges (Kim et al., 2008), or the opening of new
675 niches, e.g. due to the onset of the monsoon, as noted above. Therefore, repeated cycles
676 of sea level fluctuation from 2.7 Ma onwards (Woodruff, 2010), would have
677 periodically exposed vast areas of land, creating land bridges among the islands in SE
678 Asia (Voris, 2000), which might have promoted the more recent immigration events
679 into Malesia. Indeed, all *Oreocnide* species occupy relatively low elevation ranges
680 (≤ 2000 m; Fig. 2), so land bridges created by low sea levels may facilitate their
681 immigration, especially the widespread species *O. frutescens* and *O. integrifolia*, both
682 of which might have crossed the narrow ocean barrier between the Malay peninsula and
683 Sumatra.

684 **5 Conclusions**

685 Dense within-taxon sampling with genomic molecular markers yielded a solidly
686 monophyletic *Oreocnide*, but revealed at least two additional species to those currently
687 described. Furthermore, from timescales of tens of thousands to millions of years, our
688 study demonstrated that the evolutionary history of plants may possess important and
689 independent evidence to understand the geological history and climatic events of the
690 environments where they evolved. Moreover, this study also exemplifies how knowing
691 the timing of diversification and dispersal directions in this SE Asian genus is vital to
692 determining whether tectonic shifts or climate and sea level changes are likely to be
693 more important in influencing its biogeography. More parallel studies with extensive
694 taxon sampling in SE Asia will be crucial for improving our understanding of the
695 region’s biogeography, especially at the intraspecific level.

696

697

698

699 **Acknowledgements**

700 We thank Profs. Wen-Tsai Wang and Chia-Jui Chen for their great help in the
701 identification of the specimens, Dr. Jian-Jun Jin, Dr. Hong-Tao Li, Dr. Xiang-Qin Yu,
702 Dr. Chao-Nan Fu, Dr. Nir Horvitz, Mr. Guang-Fu Zhu and Miss Zhi-Qiong Mo for kind
703 help for software. The Royal Botanic Gardens Kew is thanked for providing some DNA
704 material. This work was facilitated by the Germplasm Bank of Wild Species, Kunming
705 Institute of Botany, Chinese Academy of Sciences (CAS). The study was supported by
706 the National Natural Science Foundation of China (31970356, 42171071, 41971071);
707 Youth Innovation Promotion Association, CAS (2019385), Yunnan Young & Elite
708 Talents Project (YNWR-QNBJ-2020-293), CAS Strategic Priority Research Program
709 (XDB31000000), the Biological Resources Programme, Chinese Academy of Sciences
710 (KFJ-BRP-017-34), and the open research project of “Cross-Cooperative Team” of the
711 Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of
712 Sciences (Y9283451Q1). Ran Nathan thanks Adelina and Massimo Della Pergola Chair
713 of Life Sciences and Minerva Center for Movement Ecology for the supports. Ran

714 Nathan and Richard I. Milne also thank the CAS President's International Fellowship
715 Initiative for its financial support (2022DB0001, 2022VBA0004).

716

717

718

719 **References**

720 An, Z., Kutzbach, J. E., Prell, W. L., & Porter, S. C. (2001). Evolution of Asian monsoons and phased
721 uplift of the Himalaya–Tibetan plateau since Late Miocene times. *Nature*, **411**, 62-66.

722 Antonelli, A., Kissling, W. D., Flantua, S. G., Bermúdez, M. A., Mulch, A., Muellner-Riehl, A. N., . . .
723 Fjeldså, J. (2018). Geological and climatic influences on mountain biodiversity. *Nature*
724 *Geoscience*, **11**, 718-725. <https://doi.org/10.1038/s41561-018-0236-z>

725 Atkins, H. J., Bramley, G. L., Johnson, M. A., Kartonegoro, A., Nishii, K., Kokubugata, G., . . . Hughes,
726 M. (2020). A molecular phylogeny of Southeast Asian *Cyrtandra* (Gesneriaceae) supports an
727 emerging paradigm for Malesian plant biogeography. *Frontiers of Biogeography*, **12**, e448.
728 <https://doi.org/10.21425/F5FBG44184>

729 Balakrishnan, P. (2011). Breeding biology of the Grey-headed Bulbul *Pycnonotus priocephalus* (Aves:
730 Pycnonotidae) in the Western Ghats, India. *Journal of Threatened Taxa*, **3**, 1415-1424.
731 <https://doi.org/10.11609/JoTT.o2381.1415-24>

732 Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., . . . Prjibelski, A. D.
733 (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell
734 sequencing. *Journal of Computational Biology*, **19**, 455-477.
735 <https://doi.org/10.1089/cmb.2012.0021>

736 Blomberg, S. P., Garland Jr, T., & Ives, A. R. (2003). Testing for phylogenetic signal in comparative data:
737 Behavioral traits are more labile. *Evolution*, **57**, 717-745. <https://doi.org/10.1111/j.0014-3820.2003.tb00285.x>

739 Bromham, L., Duchêne, S., Hua, X., Ritchie, A. M., Duchêne, D. A., & Ho, S. Y. (2018). Bayesian
740 molecular dating: Opening up the black box. *Biological Reviews*, **93**, 1165-1191.
741 <https://doi.org/10.1111/brv.12390>

742 Brown, J. W., FitzJohn, R. G., Alfaro, M. E., & Harmon, L. J. (2018). turbo-MEDUSA: Modelling
743 evolutionary diversification using stepwise AIC. R package version 0.953.

744 Cámara-Leret, R., Frodin, D. G., Adema, F., Anderson, C., Appelhans, M. S., Argent, G., . . . van Welzen,
745 P. C. (2020). New Guinea has the world's richest island flora. *Nature*, **584**, 579-583.
746 <https://doi.org/10.1038/s41586-020-2549-5>

747 Carine, M. A. (2005). Spatio-temporal relationships of the Macaronesian endemic flora: A relictual series
748 or window of opportunity? *Taxon*, **54**, 895-903. <https://doi.org/10.2307/25065476>

749 Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in
750 phylogenetic analysis. *Molecular Biology and Evolution*, **17**, 540-552.
751 <https://doi.org/10.1093/oxfordjournals.molbev.a026334>

752 Chen, C. J., Lin, Q., Friis, I., Wilmot-Dear, C. M., & Monro, A. K. (2003). Urticaceae. In Z. Y. Wu & P.
753 H. Raven (Eds.), *Flora of China* (Vol. 5, pp. 76-189). Beijing: Science Press & Missouri
754 Botanical Garden Press.

755 Cheng, H., Edwards, R. L., Sinha, A., Spötl, C., Yi, L., Chen, S., . . . Zhang, H. (2016). The Asian

756 monsoon over the past 640,000 years and ice age terminations. *Nature*, **534**, 640-646.
757 <https://doi.org/10.1038/nature18591>

758 Clouse, R. M., & Giribet, G. (2010). When Thailand was an island—the phylogeny and biogeography of
759 mite harvestmen (Opiliones, Cyphophthalmi, Stylocellidae) in Southeast Asia. *Journal of*
760 *Biogeography*, **37**, 1114-1130. <https://doi.org/10.1111/j.1365-2699.2010.02274.x>

761 Cody, S., Richardson, J. E., Rull, V., Ellis, C., & Pennington, R. T. (2010). The great American biotic
762 interchange revisited. *Ecography*, **33**, 326-332. <https://doi.org/10.1111/j.1600-0587.2010.06327.x>

764 Craven, D., Knight, T., Barton, K., Bialic-Murphy, L., & Chase, J. (2019). Dissecting macroecological
765 and macroevolutionary patterns of forest biodiversity across the Hawaiian archipelago.
766 *Proceedings of the National Academy of Sciences of the United States of America*, **116**, 16436-
767 16441. <https://doi.org/10.1073/pnas.1901954116>

768 Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: More models, new heuristics
769 and parallel computing. *Nature Methods*, **9**, 772-772. <https://doi.org/10.1038/nmeth.2109>

770 de Bruyn, M., Stelbrink, B., Morley, R. J., Hall, R., Carvalho, G. R., Cannon, C. H., . . . von Rintelen, T.
771 (2014). Borneo and Indochina are major evolutionary hotspots for Southeast Asian biodiversity.
772 *Systematic Biology*, **63**, 879-901. <https://10.1093/sysbio/syu047>

773 Diaz, F. P., Latorre, C., Carrasco-Puga, G., Wood, J. R., Wilmshurst, J. M., Soto, D. C., . . . Gutierrez, R.
774 A. (2019). Multiscale climate change impacts on plant diversity in the Atacama Desert. *Global*
775 *Change Biology*, **25**, 1733-1745. <https://doi.org/10.1111/gcb.14583>

776 Ding, A., Pittman, M., Upchurch, P., O'Connor, J., Field, D. J., & Xu, X. (2019). The Biogeography of
777 Coelurosaurian Theropods and its Impact on their Evolutionary History. *bioRxiv*, 634170.
778 <https://www.biorxiv.org/content/10.1101/634170v1>

779 Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf
780 material. *Phytochemical Bulletin*, **19**, 11-15.

781 Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees.
782 *BMC Evolutionary Biology*, **7**, 214. <https://doi.org/10.1186/1471-2148-7-214>

783 Favre, A., Packert, M., Pauls, S. U., Jahnig, S. C., Uhl, D., Michalak, I., & Muellner-Riehl, A. N. (2015).
784 The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas.
785 *Biological Reviews*, **90**, 236-253. <https://doi.org/10.1111/brv.12107>

786 Fernández-Palacios, J. M. (2016). Island biogeography: Shaped by sea-level shifts. *Nature*, **532**, 42-43.
787 <https://doi.org/10.1038/nature17880>

788 Gower, J. D., Johnson, G. K., Richardson, E. J., Rosen, R. B., Rüber, L., & Williams, T. S. (2012). *Biotic*
789 *evolution and environmental change in Southeast Asia*. Cambridge, England, UK: Cambridge
790 University Press.

791 Guo, Z. X., Li, X. N., He, Z. W., Yang, Y. C., Wang, W. Q., Zhong, C. R., . . . Shi, S. H. (2018). Extremely
792 low genetic diversity across mangrove taxa reflects past sea level changes and hints at poor
793 future responses. *Global Change Biology*, **24**, 1741-1748. <https://doi.org/10.1111/gcb.13968>

794 Hall, R. (1998). The plate tectonics of Cenozoic SE Asia and the distribution of land and sea. In R. Hall
795 & D. J. Holloway (Eds.), *Biogeography and geological evolution of SE Asia* (pp. 99-131).
796 Leiden, Netherlands: Backhuys.

797 Hall, R. (2002). Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific:
798 Computer-based reconstructions, model and animations. *Journal of Asian Earth Sciences*, **20**,
799 353-431. [https://doi.org/10.1016/S1367-9120\(01\)00069-4](https://doi.org/10.1016/S1367-9120(01)00069-4)

800 Hall, R. (2012). Sundaland and Wallacea: geology, plate tectonics and palaeogeography. In D. Gower, K.
801 Johnson, J. Richardson, B. Rosen, L. Rüber, & S. Williams (Eds.), *Biotic evolution and*
802 *environmental change in Southeast Asia* (pp. 32-78). Cambridge: UK: Cambridge Univ. Press.

803 Hall, R. (2017). Southeast Asia: New views of the geology of the Malay Archipelago. *Annual Review of*
804 *Earth Planetary Sciences*, **45**, 331-358. <https://doi.org/10.1146/annurev-earth-063016-020633>

805 He, Z. W., Li, X. N., Yang, M., Wang, X. F., Zhong, C. R., Duke, N. C., . . . Shi, S. H. (2019).
806 Speciation with gene flow via cycles of isolation and migration: insights from multiple
807 mangrove taxa. *National Science Review*, **6**, 275-288. <https://doi.org/10.1093/nsr/nwy078>

808 Hernández-Hernández, T., Brown, J. W., Schlumpberger, B. O., Eguiarte, L. E., & Magallón, S. (2014).
809 Beyond aridification: Multiple explanations for the elevated diversification of cacti in the New
810 World Succulent Biome. *New Phytologist*, **202**, 1382-1397. <https://doi.org/10.1111/nph.12752>

811 Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution
812 interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**,
813 1965-1978. <https://doi.org/10.1002/joc.1276>

814 Hooper, D. U., Adair, E. C., Cardinale, B. J., Byrnes, J. E., Hungate, B. A., Matulich, K. L., . . . O'Connor,
815 M. I. (2012). A global synthesis reveals biodiversity loss as a major driver of ecosystem change.
816 *Nature*, **486**, 105-108. <https://doi.org/10.1038/nature11118>

817 Hoorn, C., Wesselingh, F., Ter Steege, H., Bermudez, M., Mora, A., Sevink, J., . . . Figueiredo, J. (2010).
818 Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity.
819 *Science*, **330**, 927-931. <https://doi.org/10.1126/science.1194585>

820 IPCC. (2013). *Climate change 2013: The physical science basis*. Cambridge, England, UK: Cambridge
821 University Press.

822 Janssens, S. B., Vandeloek, F., De Langhe, E., Verstraete, B., Smets, E., Vandenhoutte, I., & Swennen,
823 R. (2016). Evolutionary dynamics and biogeography of Musaceae reveal a correlation between
824 the diversification of the banana family and the geological and climatic history of Southeast
825 Asia. *New Phytologist*, **210**, 1453-1465. <https://doi.org/10.1111/nph.13856>

826 Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., . . . Duran, C. (2012).
827 Geneious Basic: An integrated and extendable desktop software platform for the organization
828 and analysis of sequence data. *Bioinformatics*, **28**, 1647-1649.
829 <https://doi.org/10.1093/bioinformatics/bts199>

830 Kim, S.-C., McGowen, M. R., Lubinsky, P., Barber, J. C., Mort, M. E., & Santos-Guerra, A. (2008).
831 Timing and tempo of early and successive adaptive radiations in Macaronesia. *PloS ONE*, **3**,
832 e2139. <https://doi.org/10.1371/journal.pone.0002139>

833 Kissling, D. W., Baker, W. J., Balslev, H., Barfod, A. S., Borchsenius, F., Dransfield, J., . . . Svenning, J.
834 C. (2012). Quaternary and pre-Quaternary historical legacies in the global distribution of a
835 major tropical plant lineage. *Global Ecology and Biogeography*, **21**, 909-921.
836 <https://doi.org/10.1111/j.1466-8238.2011.00728.x>

837 Kooyman, R. M., Morley, R. J., Crayn, D. M., Joyce, E. M., Rossetto, M., Slik, J. W. F., . . . Wilf, P.
838 (2019). Origins and assembly of Malesian rainforests. *Annual Review of Ecology, Evolution,*
839 *and Systematics*, **50**, 119-143. <https://doi.org/10.1146/annurev-ecolsys-110218-024737>

840 Lambeck, K., Esat, T. M., & Potter, E.-K. (2002). Links between climate and sea levels for the past three
841 million years. *Nature*, **419**, 199-206. <https://doi.org/10.1038/nature01089>

842 Li, F. Y., & Li, S. Q. (2018). Paleocene–Eocene and Plio–Pleistocene sea-level changes as “species
843 pumps” in Southeast Asia: Evidence from *Altheopus* spiders. *Molecular Phylogenetics and*

844 *Evolution* **127**, 545-555. <https://doi.org/10.1016/j.ympev.2018.05.014>

845 Liu, C. R., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in
846 the prediction of species distributions. *Ecography*, **28**, 385-393. [https://doi.org/10.1111/j.0906-](https://doi.org/10.1111/j.0906-7590.2005.03957.x)
847 [7590.2005.03957.x](https://doi.org/10.1111/j.0906-7590.2005.03957.x)

848 Liu, J., Möller, M., Provan, J., Gao, L. M., Poudel, R. C., & Li, D. Z. (2013). Geological and ecological
849 factors drive cryptic speciation of yews in a biodiversity hotspot. *New Phytologist*, **199**, 1093-
850 1108. <https://doi.org/10.1111/nph.12336>

851 Lohman, D. J., de Bruyn, M., Page, T., von Rintelen, K., Hall, R., Ng, P. K. L., . . . von Rintelen, T.
852 (2011). Biogeography of the Indo-Australian archipelago. *Annual Review of Ecology, Evolution,*
853 *and Systematics*, **42**, 205-226. <https://doi.org/10.1146/annurev-ecolsys-102710-145001>

854 Lohse, M., Drechsel, O., Kahlau, S., & Bock, R. (2013). OrganellarGenomeDRAW—a suite of tools for
855 generating physical maps of plastid and mitochondrial genomes and visualizing expression data
856 sets. *Nucleic Acids Research*, **41**, W575-W581. <https://doi.org/10.1093/nar/gkt289>

857 Mao, K. S., Milne, R. I., Zhang, L. B., Peng, Y. L., Liu, J. Q., Thomas, P., . . . Renner, S. S. (2012).
858 Distribution of living Cupressaceae reflects the breakup of Pangea. *Proceedings of the National*
859 *Academy of Sciences of the United States of America*, **109**, 7793-7798.
860 <https://doi.org/10.1073/pnas.1114319109>

861 Matuszak, S., Muellner-Riehl, A. N., Sun, H., & Favre, A. (2016). Dispersal routes between biodiversity
862 hotspots in Asia: the case of the mountain genus *Tripterospermum* (Gentianinae, Gentianaceae)
863 and its close relatives. *Journal of Biogeography*, **43**, 580-590. <https://doi.org/10.1111/jbi.12617>

864 Matzke, N. J. (2013). Probabilistic historical biogeography: New models for founder-event speciation,
865 imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of*
866 *Biogeography*, **5**. <https://doi.org/10.21425/F5FBG19694>

867 Matzke, N. J. (2014). Model selection in historical biogeography reveals that founder-event speciation is
868 a crucial process in island clades. *Systematic Biology*, **63**, 951-970.
869 <https://doi.org/10.1093/sysbio/syu056>

870 Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L., Lamarque, J. F., Matsumoto,
871 K., Montzka, S. A., Raper, S. C., Riahi, K., 2011. The RCP greenhouse gas concentrations and
872 their extensions from 1765 to 2300. *Climatic Change* **109** (1): 213-241.

873 Meltzner, A. J., Switzer, A. D., Horton, B. P., Ashe, E., Qiu, Q., Hill, D. F., . . . Majewski, J. M. (2017).
874 Half-metre sea-level fluctuations on centennial timescales from mid-Holocene corals of
875 Southeast Asia. *Nature Communications*, **8**, 1-16. <https://doi.org/10.1038/ncomms14387>

876 Michaux, B. (2010). Biogeology of Wallacea: Geotectonic models, areas of endemism, and natural
877 biogeographical units. *Biological Journal of the Linnean Society*, **101**, 193-212.
878 <https://doi.org/10.1111/j.1095-8312.2010.01473.x>

879 Miler, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference
880 of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop*
881 *(GCE)*, 14 Nov. 2010, New Orleans, LA pp 1-8.

882 Milne, R. I. (2009). Effects of taxon sampling on molecular dating for within-genus divergence events,
883 when deep fossils are used for calibration. *Journal of Systematics and Evolution*, **47**, 383-401.
884 <https://doi.org/10.1111/j.1759-6831.2009.00038.x>

885 Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity
886 hotspots for conservation priorities. *Nature*, **403**, 853-858. <https://doi.org/10.1038/35002501>

887 Nago, R., Kobayashi, S., Kinjo, T., Ogimi, K., Sinjo, A., Namisato, S., . . . Izawa, M. (2019). Seed feeding

888 behavior of *Diplothrix legata* (Muridae: Rodentia): Effects on germination of five plants in
889 Okinawajima Island, Japan. *Mammal Study*, **44**, 129-134. <https://doi.org/10.3106/ms2018-0049>

890 Nathan, R., Horvitz, N., He, Y., Kuparinen, A., Schurr, F. M., & Katul, G. G. (2011). Spread of North
891 American wind-dispersed trees in future environments. *Ecology Letters*, **14**, 211-219.
892 <https://doi.org/10.1111/j.1461-0248.2010.01573.x>

893 Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective
894 stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and*
895 *Evolution*, **32**, 268-274. <https://doi.org/10.1093/molbev/msu300>

896 Osuri, A. M., Chakravarthy, D., Mudappa, D., Raman, T. S., Ayyappan, N., Muthuramkumar, S., &
897 Parthasarathy, N. (2017). Successional status, seed dispersal mode and overstorey species
898 influence tree regeneration in tropical rain-forest fragments in Western Ghats, India. *Journal of*
899 *Tropical Ecology*, **33**, 270-284. <https://doi.org/10.1017/S0266467417000219>

900 Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R
901 language. *Bioinformatics*, **20**, 289-290. <https://doi.org/10.1093/bioinformatics/btg412>

902 Pfeifer-Meister, L., Bridgham, S. D., Reynolds, L. L., Goklany, M. E., Wilson, H. E., Little, C. J., . . .
903 Johnson, B. R. (2016). Climate change alters plant biogeography in Mediterranean prairies
904 along the West Coast, USA. *Global Change Biology*, **22**, 845-855.
905 <https://doi.org/10.1111/gcb.13052>

906 Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species
907 geographic distributions. *Ecological Modelling*, **190**, 231-259.
908 <https://doi.org/10.1016/j.ecolmodel.2005.03.026>

909 Qu, X. J., Moore, M. J., Li, D. Z., & Yi, T. S. (2019). PGA: A software package for rapid, accurate, and
910 flexible batch annotation of plastomes. *Plant Methods*, **15**. <https://doi.org/10.1186/s13007-019-0435-7>

911

912 Rabosky, D. L., Grudler, M., Anderson, C., Title, P., Shi, J. J., Brown, J. W., . . . Larson, J. G. (2014).
913 BAMM tools: An R package for the analysis of evolutionary dynamics on phylogenetic trees.
914 *Methods in Ecology and Evolution*, **5**, 701-707. <https://doi.org/10.1111/2041-210X.12199>

915 Rambaut, A. (2009). FigTree. *Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK*.
916 Available at: <http://tree.bio.ed.ac.uk/software/figtree/>.

917 Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization
918 in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, **67**, 901-904.
919 <https://doi.org/10.1093/sysbio/syy032>

920 Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things).
921 *Methods in Ecology and Evolution*, **3**, 217-223. <https://doi.org/10.1111/j.2041-210X.2011.00169.x>

922

923 Ricklefs, R. E. (2004). A comprehensive framework for global patterns in biodiversity. *Ecology Letters*,
924 **7**, 1-15. <https://doi.org/10.1046/j.1461-0248.2003.00554.x>

925 Rohling, E., Foster, G. L., Grant, K., Marino, G., Roberts, A., Tamisiea, M. E., & Williams, F. (2014).
926 Sea-level and deep-sea-temperature variability over the past 5.3 million years. *Nature*, **508**, 477-
927 482. <https://doi.org/10.1038/nature13230>

928 Rohling, E. J., Hibbert, F. D., Grant, K. M., Galaasen, E. V., Irvall, N., Kleiven, H. F., . . . Yu, J. (2019).
929 Asynchronous Antarctic and Greenland ice-volume contributions to the last interglacial sea-
930 level highstand. *Nature Communications*, **10**, 5040. <https://doi.org/10.1038/s41467-019-12874-3>

931

932 Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., . . . Huelsenbeck, J. P.
933 (2012). MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large
934 model space. *Systematic Biology*, **61**, 539-542. <https://doi.org/10.1093/sysbio/sys029>
935 Santos, A. M. C., Field, R., & Ricklefs, R. E. (2016). New directions in island biogeography. *Global*
936 *Ecology and Biogeography*, **25**, 751-768. <https://doi.org/10.1111/geb.12477>
937 Schwalm, C. R., Glendon, S., Duffy, P. B.. 2020. RCP8. 5 tracks cumulative CO2 emissions. *Proceedings*
938 *of the National Academy of Sciences* **117**(33): 19656-19657.
939 Shih, H. T., Hung, H. C., Schubart, C. D., Chen, C. A., & Chang, H. W. (2006). Intraspecific genetic
940 diversity of the endemic freshwater crab *Candidiopotamon rathbunae* (Decapoda, Brachyura,
941 Potamidae) reflects five million years of the geological history of Taiwan. *Journal of*
942 *Biogeography*, **33**, 980-989. <https://doi.org/10.1111/j.1365-2699.2006.01472.x>
943 Sirichamorn, Y., Thomas, D. C., Adema, F. A. C. B., van Welzen, P. C., & Parmakelis, A. (2014).
944 Historical biogeography of *Aganope*, *Brachypterum* and *Derris* (Fabaceae, tribe Millettieae):
945 Insights into the origins of Palaeotropical intercontinental disjunctions and general
946 biogeographical patterns in Southeast Asia. *Journal of Biogeography*, **41**, 882-893.
947 <https://doi.org/10.1111/jbi.12262>
948 Slik, J. W. F., Aiba, S.-I., Bastian, M., Brearley, F. Q., Cannon, C. H., Eichhorn, K. A. O., . . . Zweifel, N.
949 (2011). Soils on exposed Sunda Shelf shaped biogeographic patterns in the equatorial forests of
950 Southeast Asia. *Proceedings of the National Academy of Sciences of the United States of*
951 *America*, **108**, 12343. <https://doi.org/10.1073/pnas.1103353108>
952 Smith, S. A., & O'Meara, B. C. (2012). treePL: Divergence time estimation using penalized likelihood
953 for large phylogenies. *Bioinformatics*, **28**, 2689-2690.
954 <https://doi.org/10.1093/bioinformatics/bts492>
955 Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large
956 phylogenies. *Bioinformatics*, **30**, 1312-1313. <https://doi.org/10.1093/bioinformatics/btu033>
957 Stelbrink, B., Albrecht, C., Hall, R., & von Rintelen, T. (2012). The biogeography of Sulawesi revisited:
958 Is there evidence for a vicariant origin of taxa on Wallace's "anomalous island"? *Evolution*, **66**,
959 2252-2271. <https://doi.org/10.1111/j.1558-5646.2012.01588.x>
960 Thomas, D. C., Hughes, M., Phutthai, T., Ardi, W. H., Rajbhandary, S., Rubite, R., . . . Richardson, J. E.
961 (2012). West to east dispersal and subsequent rapid diversification of the mega-diverse genus
962 *Begonia* (Begoniaceae) in the Malesian archipelago. *Journal of Biogeography*, **39**, 98-113.
963 <https://doi.org/10.1111/j.1365-2699.2011.02596.x>
964 Trakhtenbrot, A., Nathan, R., Perry, G., & Richardson, D. M. (2005). The importance of long-distance
965 dispersal in biodiversity conservation. *Diversity and Distributions*, **11**, 173-181.
966 <https://doi.org/10.1111/j.1366-9516.2005.00156.x>
967 Van Welzen, P. C., Parnell, J. A., & Slik, J. F. (2011). Wallace's Line and plant distributions: Two or three
968 phytogeographical areas and where to group Java? *Biological Journal of the Linnean Society*,
969 **103**, 531-545. <https://doi.org/10.1111/j.1095-8312.2011.01647.x>
970 Vermeij, G. J. (1991). When biotas meet: Understanding biotic interchange. *Science*, **253**, 1099.
971 <https://doi.org/10.1126/science.253.5024.1099>
972 Voris, H. K. (2000). Maps of Pleistocene sea levels in Southeast Asia: Shorelines, river systems and time
973 durations. *Journal of Biogeography*, **27**, 1153-1167. <https://doi.org/10.1046/j.1365-2699.2000.00489.x>
974 Wang, R. N., Milne, R. I., Du, X. Y., Liu, J., & Wu, Z. Y. (2020). Characteristics and mutational hotspots
975

- 976 of plastomes in *Debregeasia* (Urticaceae). *Frontiers in Genetics*, **11**, 729-729.
 977 <https://doi.org/10.3389/fgene.2020.00729>
- 978 Webb, C. O., Ackerly, D. D., McPeck, M. A., & Donoghue, M. J. (2002). Phylogenies and community
 979 ecology. *Annual Review of Ecology and Systematics*, **33**, 475-505.
 980 <https://doi.org/10.1146/annurev.ecolsys.33.010802.150448>
- 981 Wick, R. R., Schultz, M. B., Zobel, J., & Holt, K. E. (2015). Bandage: Interactive visualization of *de*
 982 *novo* genome assemblies. *Bioinformatics*, **31**, 3350-3352.
 983 <https://doi.org/10.1093/bioinformatics/btv383>
- 984 Woodruff, D. S. (2010). Biogeography and conservation in Southeast Asia: how 2.7 million years of
 985 repeated environmental fluctuations affect today's patterns and the future of the remaining
 986 refugial-phase biodiversity. *Biodiversity and Conservation*, **19**, 919-941.
 987 <https://doi.org/10.1007/s10531-010-9783-3>
- 988 Wu, Z. Y., Liu, J., Provan, J., Wang, H., Chen, C. J., Cadotte, M. W., . . . Milne, R. I. (2018). Testing
 989 Darwin's transoceanic dispersal hypothesis for the inland nettle family (Urticaceae). *Ecology*
 990 *Letters*, **21**, 1515-1529. <https://doi.org/10.1111/ele.13132>
- 991 Wu, Z. Y., Monro, A. K., Milne, R. I., Wang, H., Yi, T. S., Liu, J., & Li, D. Z. (2013). Molecular phylogeny
 992 of the nettle family (Urticaceae) inferred from multiple loci of three genomes and extensive
 993 generic sampling. *Molecular Phylogenetics and Evolution*, **69**, 814-827.
 994 <http://dx.doi.org/10.1016/j.ympev.2013.06.022>
- 995 Yu, Y., Blair, C., & He, X. J. (2020). RASP 4: Ancestral state reconstruction tool for multiple genes and
 996 characters. *Molecular Biology and Evolution*, **37**, 604-606.
 997 <https://doi.org/10.1093/molbev/msz257>
- 998 Yu, Y., Harris, A. J., & He, X. J. (2010). S-DIVA (Statistical Dispersal-Vicariance Analysis): A tool for
 999 inferring biogeographic histories. *Molecular Phylogenetics and Evolution*, **56**, 848-850.
 1000 <http://dx.doi.org/10.1016/j.ympev.2010.04.011>
- 1001 Zarowiecki, M., Linton, Y. M., Post, R. J., Bangs, M. J., Htun, P. T., Hlaing, T., . . . Sochantha, T. (2014).
 1002 Repeated landmass reformation limits diversification in the widespread littoral zone mosquito
 1003 *Anopheles sundaicus* sensu lato in the Indo-Oriental Region. *Molecular Ecology*, **23**, 2573-2589.
 1004 <https://doi.org/10.1111/mec.12761>
- 1005 Zhang, D., Ye, Z., Yamada, K., Zhen, Y., Zheng, C., & Bu, W. (2016). Pleistocene sea level fluctuation
 1006 and host plant habitat requirement influenced the historical phylogeography of the invasive
 1007 species *Amphiareus obscuriceps* (Hemiptera: Anthocoridae) in its native range. *BMC*
 1008 *Evolutionary Biology*, **16**, 174. <https://doi.org/10.1186/s12862-016-0748-3>
- 1009 Zhao, J. L., Xia, Y. M., Cannon, C. H., Kress, W. J., & Li, Q. J. (2016). Evolutionary diversification of
 1010 alpine ginger reflects the early uplift of the Himalayan–Tibetan Plateau and rapid extrusion of
 1011 Indochina. *Gondwana Research*, **32**, 232-241. <https://doi.org/10.1016/j.gr.2015.02.004>

1012
 1013
 1014
 1015

1016 Captions of figures

1017

1018 **Fig. 1** Present distribution of the ten currently recognized species of *Oreocnide*, based
1019 on all known herbarium records, plus sampling sites for the current study, shown as
1020 circles and triangles of the same color, respectively.

1021

1022 **Fig. 2** (a) Number of accessions of *Oreocnide* binned by 1°C according to average
1023 annual temperature in observed locations. (b) Number of accessions of *Oreocnide*
1024 binned by 50 mm according to average annual rainfall in observed locations. (c) The
1025 elevation of each species of *Oreocnide*. Data for (a), (b) and (c) is based on GBIF
1026 observation and Worldclim data.

1027

1028 **Fig. 3** ML tree inferred from combined sequences comprising complete plastome and
1029 18S-ITS1-5.8S-ITS2-26S sequences. Numbers above branches are bootstrap values
1030 (BS) from ML, and below branches are Bayesian posterior probability (PP). Maximum
1031 values (BS=100% and PP =1.0) are marked with star (*).

1032

1033 **Fig. 4** The effect of different calibrations on posterior divergence time estimates of
1034 *Oreocnide*. Horizontal bars show 95% high posterior density (HPD) credibility
1035 intervals. Node **a** represents the split time of *Oreocnide* from its sister lineage (stem
1036 divergence); Node **b** is crown divergence of *Oreocnide*. Circles with numbers indicate
1037 calibration points. Strategies A and B both employed seven calibrations (four secondary
1038 calibrations (1-3, 7) and three fossils (4-6)), but differ in that dates were calculated
1039 using BEAST and treePL, respectively. Likewise strategies C and D used BEAST and
1040 treePL, respectively, but these strategies employed only three calibrations, i.e. two
1041 secondary calibrations (1, 7) and one fossil (4).

1042

1043 **Fig. 5** Global biogeographical patterns of *Oreocnide* inferred by BioGeoBEARS. The
1044 map shows the eight biogeographical regions defined in this study. Clades are marked
1045 by numbers. Ancestral area reconstructions (AARs) with the highest likelihood values
1046 are shown via pie charts above nodes, with colors referring to the different regions as
1047 indicated in the map. Colored boxes behind tips indicate the sampled location of each
1048 individual accession, as indicated in the map. Sketches on the right show the most likely
1049 dispersal routes of the four main Clades (I-IV), based on the divergence times and
1050 biogeographic analyses.

1051

1052 **Fig. 6** (a) Rate through time plots for speciation, extinction and net diversification with
1053 95% confidence interval indicated by shaded areas, obtained from the BAMM analysis.
1054 The red, green and blue lines represent the extinction rate, speciation rate and net
1055 diversification rate respectively. (b) Phylorate plot of BAMM analysis. The branches
1056 are colored by speciation rate (lineages million yr⁻¹), where blue is lowest and red
1057 highest. (c) Results of the MEDUSA diversification rates analyses over the 1000
1058 randomly selected trees. Number of rate shift was indicated in gray circle with number,
1059 and back ground was colored by gray circle.

1060

1061 **Fig. 7** Potential distribution range of each species of *Oreocnide* simulated by Ecological
1062 Niche Models using bioclimatic variables and species occurrence points. For each
1063 species, red, purple, blue and green colors indicate the potential ranges (based on
1064 specific threshold (Table 1) during LIG (Last Interglacial), LGM (Last Glacial
1065 Maximum), Present and future (2070), respectively. Where an area was occupied during
1066 more than one period, the latest layer is shown on top.

1067

1068 **Fig. 8** (a) Mean elevation, and potential distribution area for each species across four
1069 epochs (LIG, LGM, present and future (2070 under RCP 8.5)). (b) Potential migration
1070 direction shift along the latitude and longitude for each species in four periods (LIG,
1071 LGM, present and future), the mean centers for each period were connected in
1072 chronological order, and the arrow represents the migration direction, and the points on
1073 the line represent the mean center of the potential distribution area in each period.

1074

1075

1076 **Fig. 9** A schematic diagram show the potential range shift scenarios of *Oreocnide* in
1077 response to climate change from LIG to 2070. (a) latitudinal expansion and contraction
1078 between continent and islands from past to future, and (b) elevation range shift in
1079 response to climate change and climate driven sea level fluctuations.

1080

1081

1082 **Fig. S1** ML tree inferred from complete plastome sequences. Numbers above and below
1083 branches are bootstrap values (BS) from ML and Bayesian posterior probability (PP),
1084 respectively. BS=100% and PP =1.0 are marked with star (*).

1085

1086 **Fig. S2** ML tree inferred from CDS sequences. Numbers above and below branches are
1087 bootstrap values (BS) from ML and Bayesian posterior probability (PP), respectively.
1088 BS=100% and PP =1.0 are marked with star (*).

1089

1090 **Fig. S3** ML tree inferred from 18S-ITS1-5.8S-ITS2-26S sequences. Numbers above
1091 and below branches are bootstrap values (BS) from ML and Bayesian posterior
1092 probability (PP), respectively. BS=100% and PP =1.0 are marked with star (*).

1093

1094 **Fig. S4** Result of BEAST analyses based on seven constraints using Birth-Death priors.

1095

1096 **Fig. S5** Result of treePL analyses based on seven constraints using Birth-Death priors.

1097

1098 **Fig. S6** Result of BEAST analyses based on three constraints using Birth-Death priors.

1099

1100 **Fig. S7** Result of treePL analyses based on three constraints using Birth-Death priors.

1101

1102 **Fig. S8** Result of BEAST analyses based on seven constraints using yule priors.

1103

1104 **Fig. S9** Global biogeographical patterns of *Oreocnide* inferred by S-DIVA. Map shows
1105 the eight biogeographical regions in colors as defined in this study. Clades are marked
1106 by numbers. Ancestral area reconstructions (AARs) with the highest likelihood values
1107 are shown via pie charts above nodes, with colors referring to the different regions as
1108 indicated in the map. Colored boxes behind tips indicate which region each accession
1109 was sampled from, as indicated in the map.