

ECOGRAPHY

Research

Testing the causes of richness patterns in the paleotropics: time and diversification in cycads (Cycadaceae)

Jian Liu, Anders J. Lindstrom, Nathalie S. Nagalingum, John J. Wiens and Xun Gong

J. Liu (<https://orcid.org/0000-0003-1828-4024>) and X. Gong (<https://orcid.org/0000-0003-1705-6935>) ✉ (gongxun@mail.kib.ac.cn), CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Inst. of Botany, Chinese Academy of Sciences, Kunming, Yunnan, China; and: Dept of Economic Plants and Biotechnology, Yunnan Key Laboratory for Wild Plant Resources, Kunming Inst. of Botany, Chinese Academy of Sciences, Kunming, China. – A. J. Lindstrom (<https://orcid.org/0000-0003-0419-9136>), Global Biodiversity Conservancy, Soi Bua Thong, Bangsalae, Sattabip, Chonburi, Thailand. – N. S. Nagalingum (<https://orcid.org/0000-0002-0221-9650>), California Academy of Sciences, San Francisco, CA, USA. – J. J. Wiens (<https://orcid.org/0000-0003-4243-1127>), Dept of Ecology and Evolutionary Biology, Univ. of Arizona, Tucson, AZ, USA. XG also at: Univ. of Chinese Academy of Sciences, Beijing, China.

Ecography

44: 1606–1618, 2021

doi: 10.1111/ecog.05910

Subject Editor: Wolf L. Eiserhardt

Editor-in-Chief: Miguel Araújo

Accepted 9 August 2021



The paleotropics harbor many biodiversity hotspots and show many different species richness patterns. However, it remains unclear which factors are the most important in directly shaping richness patterns among regions in the paleotropics (i.e. diversification rates, colonization times, and dispersal frequency). Here we used Cycadaceae as a model system to test the causes of regional richness patterns in the paleotropics. Specifically, we tested the roles of dispersal frequency, colonization time, diversification rates, and their combined role in explaining richness patterns among regions. We generated a well-sampled, time-calibrated phylogeny and then used this to estimate dispersal events, colonization times, and diversification rates. Richness patterns were significantly associated with the timing of the first colonization of each region and were best explained by the combined effects of colonization time and diversification rates. The number of dispersal events into each region and the diversification rates of species in each region were not significantly related to richness. Ancestral-area reconstructions showed frequent migrations across Wallace's line, with a higher diversification rate east of Wallace's line than west of it. Overall, our study shows that colonization time can be an important factor for explaining regional richness patterns in the paleotropics.

Keywords: Cycadaceae, diversification, paleotropics, species richness, time-for-speciation effect, Wallace's line

Introduction

Variation in species richness among regions has intrigued ecologists, biogeographers, and evolutionary biologists for centuries (Darwin 1859, Wallace 1869). Richness patterns are shaped directly by speciation, extinction, and dispersal, the three processes that directly change species numbers in a region (Ricklefs 1987). Given this idea, two non-exclusive hypotheses have been proposed to explain why more species occur in some regions than others (Ricklefs 2006, Wiens 2011). One hypothesis is that lineages in high-richness regions have higher diversification rates than those in low-richness



www.ecography.org

regions (Condamine et al. 2012), where the diversification rate is the speciation rate minus the extinction rate (Ricklefs 2007, Morlon 2014). The other hypothesis is that high-richness regions were colonized earlier, allowing more time for richness to build up through in-situ speciation (Stephens and Wiens 2003). Many previous studies have supported the role of time (Jansson et al. 2013, Li and Wiens 2019), whereas others supported the diversification-rate hypothesis (Pyron and Wiens 2013, Schluter 2016). Yan et al. (2018) recently found that the combined effects of diversification rates and colonization time can explain regional richness patterns (in Primulaceae), even when neither factor alone is significantly related to richness. More colonizations of a region can also increase its richness (Jablonski et al. 2006). Recently, Li and Wiens (2019) examined the roles of colonization times, colonization frequencies, and diversification rates in explaining regional richness patterns for 15 plant and animal clades. They found that richness patterns in most clades were best explained by colonization times.

The paleotropics include tropical Africa, Asia, and Oceania. They include biodiversity hotspots such as the rainforests of the Philippines, Wallacea, and Sundaland (Myers et al. 2000). In plants, many different richness patterns have been found in this region, such as high richness in eastern and southeastern Asia (Alangiaceae; Feng et al. 2009) and diversity hotspots in Pacific archipelagoes (Pittosporaceae; Gemmill et al. 2002). However, to our knowledge, no study has focused on which factors are most important in shaping regional richness patterns in the paleotropics, particularly the relative roles of colonization time, diversification rates, and dispersal frequency.

A major biogeographic boundary in the paleotropics is Wallace's line, which separates the Asian (Sunda plate) and Australian (Sahul plate) biotic regions (Voris 2000, Hanebuth et al. 2011). This line coincides with the ancient, deep-water channels of the Lombok Strait and the Makassar Strait. These acted as dispersal barriers even during the Pleistocene, when sea levels were much lower than present (Voris 2000, Hanebuth et al. 2011). For plants, some studies have assessed patterns of diversification and colonization across Wallace's line and have challenged the importance of this barrier in restricting dispersal (Thomas et al. 2012, Chantarasuwan et al. 2016). However, it remains unclear whether or not dispersal across Wallace's line is significantly reduced relative to dispersal among other regions. Furthermore, a study in palms (Bacon et al. 2013) suggested that lineages on the east side of Wallace's line (Sahul) might have increased diversification rates relative to those on the west side (Sunda). However, this hypothesis has not been tested in other groups.

The cycad family Cycadaceae is an excellent model system for examining patterns of species richness and biogeography in paleotropics. Cycadaceae consists of one genus *Cycas* with 117 accepted, extant species (Calonje et al. 2020). These species occur only in the paleotropics and show strong disparities in richness among regions (Fig. 1). The Indochina Peninsula and northern Australia have the highest richness with >70% of the species (Lindstrom et al. 2009). A recent

study proposed that Cycadaceae originated in the Indochina Peninsula (Mankga et al. 2020) and dispersed to other regions, but no studies have tested the causes of their richness patterns among regions.

Here we use Cycadaceae as a model to understand the processes shaping richness patterns in the paleotropics and across Wallace's line. To do this we 1) infer a densely sampled, time-calibrated phylogeny, 2) estimate regional diversification rates, 3) reconstruct biogeographic history, and 4) test relationships between regional species richness and colonization times, diversification rates, and the number of colonization events.

Material and methods

Taxon sampling and sequence data

We obtained molecular data from a recent phylogeny (Liu et al. 2018), using only accepted species (Calonje et al. 2020). For species with two or more subspecies, we randomly selected one subspecies to represent the species. In total, we sampled 103 *Cycas* species, representing 88% of the 117 extant species, and spanning the geographic distribution of the genus. The 14 unsampled species are listed in Supporting information. A ten-marker dataset (seven nuclear genes and three plastid intergenic spacers; Supporting information) for 113 species (103 ingroups, 10 outgroups) was used (Liu et al. 2018). Details on taxon sampling, accession numbers, and species distributions are given in Supporting information.

Phylogeny and divergence-time estimation

Likelihood trees were estimated in IQ-TREE v2.1.1 (Minh et al. 2020) using separate plastid and nuclear datasets and a concatenated dataset. We checked for nodes with strongly supported incongruence (bootstrap >90% in both) between the plastid and nuclear trees. We then used BEAST v2.6.1 (Bouckaert et al. 2019) for topology and age estimation (details in Supporting information), and compared the two concatenated estimates.

Ancestral-area reconstructions and dispersal

We used seven biogeographic regions (Supporting information). These were A: east Asia and Indochina; B: India including Sri Lanka; C: Sunda Shelf (west of Wallace's line, including the Andaman and Nicobar Islands); D: the Philippines; E: Wallacea (east of Wallace's line); F: Australasia (Sahul Shelf, including New Guinea and northern Australia); and G: Africa (including Madagascar). These regions were based on previous biogeographic studies of paleotropical *Begonia* and legume genera (Thomas et al. 2012, Sirichamorn et al. 2014). These regions have distinct species compositions, with numerous *Cycas* species endemic to each (Table 1 and Supporting information). We found a very strong relationship between the number of described species in each region

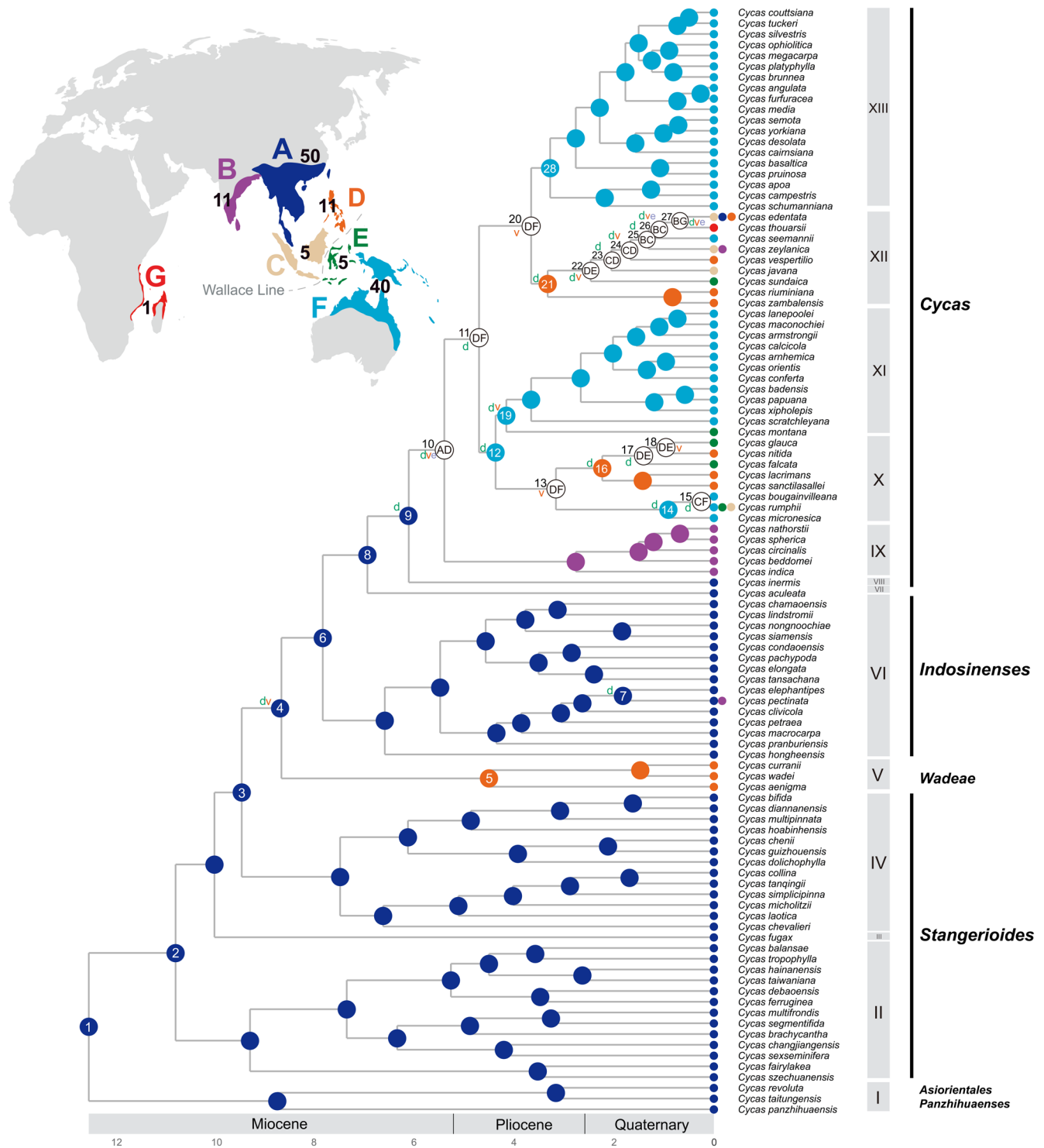


Figure 1. Phylogeny and biogeographic reconstructions for Cycadaceae. The tree is the maximum clade-credibility tree from BEAST, and colored dots at nodes represent most likely ancestral areas from the dispersal–extinction–cladogenesis (DEC) model in RASP allowing up to two regions per ancestral node. Clades of interest are represented with Latin numbers (Table 2). The historical events inferred by DEC are shown on the nodes of the tree. For node identity: d=dispersal, v=vicariance, e=extinction. Inset is the geographic distribution of extant Cycadaceae (*Cycas*), including the seven geographic regions used and the number of accepted, extant species in each region (Calonje et al. 2020). The map was produced using equal-area projection. See Supporting information for details on how the map was generated.

Table 1. Summary of species richness (including sampled size in this study), the timing of colonization, and weighted diversification rate and time by net diversification rate values of *Cycas* of each region. ϵ , the relative extinction fraction (ϵ = speciation rate/extinction rate); diversification rates are in species per millions of years (Mya).

Regions	Species richness (sampled species)	Sample fraction	Age of first colonization AFC (Mya)	Summed ages of colonization SAC (Mya)	Number of colonization events (NCE, two-area per region)	Net diversification rates				Time (AFC) by net diversification rates			
						MuSSE ($\epsilon=0.1$)	MS ($\epsilon=0.1$)	MS ($\epsilon=0.5$)	MS ($\epsilon=0.9$)	MuSSE ($\epsilon=0.1$)	MS ($\epsilon=0.1$)	MS ($\epsilon=0.5$)	MS ($\epsilon=0.9$)
East Asia and Indochina (A)	50 (48)	0.96	12.50	12.86	1	0.1952	0.2466	0.1955	0.0813	2.4397	3.0813	2.444	1.016
India and Sri Lanka (B)	11 (7)	0.64	4.07	5.63	4	0.5597	0.3199	0.2446	0.0892	2.2779	1.3020	0.996	0.363
Sunda Shelf (C)	5 (4)	0.8	1.02	2.15	6	0.5434	0.286	0.201	0.058	0.5543	0.3215	0.205	0.059
The Philippines (D)	11 (10)	0.91	6.57	12.53	4	0.3768	0.2495	0.1732	0.0486	2.4754	1.7253	1.138	0.319
Wallacea (E)	5 (5)	1	2.07	4.12	4	0.2608	0.1292	0.0815	0.0192	0.5398	0.2674	0.169	0.0397
Australasia (F)	40 (34)	0.85	4.54	5.05	4	0.6618	0.6792	0.5481	0.2451	3.0046	3.0836	2.4884	1.1126
Madagascar, Africa (G)	1 (1)	1	0.36	0.36	1	0	0	0	0	0	0	0	0

(Calonje et al. 2020) and those sampled in the tree from that region ($r^2=0.989$; $p < 0.0001$, Supporting information). Among regions, the proportion of species sampled ranged from 0.64 to 1.00 (mean = 0.88, Table 1).

Biogeographic analyses were conducted using RASP v4 (Yu et al. 2019) based on the dispersal–extinction–cladogenesis (DEC) model (Ree and Smith 2008). We pruned out-group taxa from the tree using the package *ape* (Paradis and Schliep 2019) in R v3.6.3 (<www.r-project.org>). For DEC analysis, baseline rates of dispersal and local extinction were set to ‘estimated’. We used an adjacency matrix to constrain dispersal between regions (Supporting information). Each ancestor was allowed to occur in two or fewer regions. This option was chosen because only two (< 2%) *Cycas* species (*C. rumphii*, *C. edentata*) occur in two or three regions, and ~96% are limited to one region (Supporting information, Fig. 1). Nevertheless, we also performed an alternative analysis that allowed each ancestor to occur in four regions.

Regional species richness: the role of time

To test the time-for-speciation hypothesis, we carried out regression between richness and the estimated age of first colonization (AFC) of each region. To obtain the AFC, we integrated the dated phylogeny with the reconstructed ancestral areas to estimate when each region was first colonized. For example, for region B, we identified the oldest node on the tree on which that region was inferred to be present and then repeated this for each region. We used reconstructions allowing ancestral species to occur in no more than two regions, which is more realistic (above). We assumed that colonization events happened in the middle of the branch (mean of crown and stem ages) given that it is impossible to infer when exactly a region was colonized. This assumption should have little impact on the results since most branches were short (<13 million years).

A region was inferred to be ancestral for a given node when that region had the highest proportional likelihood relative to other regions. When the inferred region was ambiguous (i.e. different regions equally likely), we used the next (most recent) node on the tree. In some cases, the region’s oldest colonization was represented by only one extant species, and we used half the species’ age in these cases.

We used the estimated crown age of extant *Cycas* (A) as the AFC for the ancestral region for the genus. However, this estimated crown age is somewhat uncertain given the broad confidence interval (Supporting information). Therefore, we randomly sampled 10 trees from the BEAST posterior distribution of trees (using *ape*), and repeated our analyses of AFC on each tree.

We also examined the relationship between richness and summed ages of colonization (SAC) of each region (Hutter et al. 2013). This approach incorporates both the number of colonization and the age of each colonization. We used the DEC analysis to infer the time of each colonization of each region, as described above. We then summed these numbers for each region.

To assess the relationship between colonization time and richness we performed two analyses using linear regression in R. Both used the richness of regions as the dependent variable. The independent variables were either AFC or SAC. We did not correct these regression analyses for phylogeny since the units for these analyses are regions (not organisms).

Regional species richness: the role of diversification rates

To examine the relationship between the richness of each region and the net diversification rates of clades occurring there, we estimated net diversification rates for each region using two methods.

One approach used the method-of-moments estimator for stem-group ages (Magallon and Sanderson 2001). This approach (MS hereafter) was implemented in the R package GEIGER v2.0.6.4 (Harmon et al. 2008). Based on simulations, this method can give accurate estimates of diversification rates for clades when the true diversification rates vary strongly between subclades (Meyer and Wiens 2018), vary strongly within clades over time (Meyer et al. 2018), and when rates are faster in younger clades (Kozak and Wiens 2016). Thus, it does not require constant rates within or between clades to be accurate. Furthermore, rate estimates from this approach are strongly related to richness patterns among clades across all of life (Scholl and Wiens 2016), and such relationships are not inevitable or artifactual (Kozak and Wiens 2016). These crucial findings have been ignored in criticisms of the method (Henaó Diaz et al. 2019, Rabosky and Benson 2021).

For the MS estimator, we first divided all the species in the genus into previously recognized, nonoverlapping, endemic higher taxa (e.g. section, subsection and species groups/complex). We selected higher taxa to represent the largest monophyletic groups in each region (Supporting information). For species not included in our tree, we assumed that they belonged to the higher taxa to which they were assigned based on previous taxonomy (Hill and Osborne 2001, Hill 2004, Singh and Radha 2006, Jutta et al. 2010, Forster 2011, Srivastava 2014, Singh et al. 2015, Rao et al. 2016, Forster and Gray 2017, Calonje et al. 2020). We used the stem-group estimator since it is generally more accurate in simulations than the crown-group estimator (Meyer and Wiens 2018). We initially applied a relatively low extinction fraction (ϵ , epsilon) of 0.1, but alternative values ($\epsilon=0.5$, $\epsilon=0.90$) gave similar results. To estimate a single diversification rate for regions with multiple clades, we weighted the diversification rate for each clade based on its proportional richness in the region (the clade's richness in that region divided by the region's total richness), following Li and Wiens (2019). We then summed weighted rates across clades to yield the weighted diversification rate for each region. This is equivalent to a diversification rate averaged among the species in the region.

The other method used to estimate diversification rates was the multiple state speciation extinction (MuSSE) approach

implemented in the R package *diversitree* v0.9 (FitzJohn 2012). We treated occurrence in each region as a character state. MuSSE does not require assigning species to clades, nor weighting rates among clades within a region. Since MuSSE allows only one state for each species, we coded species that occur in more than one region based on their inferred ancestral region. However, only four species occurred in >1 region. Three rates were estimated for each region: speciation (λ), extinction (μ), and transitions between regions (q). We compared eight models (Supporting information), each with parameters that were either variable or constrained to be equal among regions. To account for incomplete species sampling in the tree, we incorporated the sampling fraction for each region (Table 1): the number of species sampled in the tree relative to the total number in the region. The size-corrected AIC (AICc; Burnham and Anderson 2002) was compared to find the optimal model.

Note that we focused on estimating an overall diversification rate for each region, not estimating the effects of 'hidden states' on diversification (e.g. HiSSE: Beaulieu and O'Meara 2016). We could not use MuHiSSE (Nakov et al. 2019) since it only allows up to four states (seven used here), nor GeoSSE (Goldberg et al. 2011) given the many regions used. Furthermore, few species were shared between regions (which GeoSSE corrects for).

After identifying the best-fitting MuSSE model, we estimated the posterior density distribution of parameters for that model and conducted Bayesian analyses (20% burn-in after 10^5 generations) to estimate speciation, extinction, and dispersal rates. Diversification rates were obtained by subtracting the extinction rates from the speciation rates.

Finally, for both methods (MS and MuSSE), we tested the relationship between the ln-transformed richness of each region and its estimated diversification rate, using linear regression in R. Again, we did not correct for phylogeny in these analyses because the units were regions. The diversification rate for Africa was treated as zero for both methods since there was no in situ diversification there (balanced speciation and extinction, or no speciation or extinction). Furthermore, the stem-group MS estimate of the diversification rate for a clade with one species is zero.

Regional species richness: time and rates

We also tested the hypothesis that time and diversification rates act together to explain richness patterns among regions (Yan et al. 2018). To test this combined effect, we simply multiplied the diversification rate inferred for each region by its oldest colonization time (using both MS and MuSSE methods), following Yan et al. (2018). The unit for this combined variable is the expected number of species in each region. Note that a strong relationship between this combined variable and richness is not inevitable, especially since diversification rates and richness of regions need not be positively related.

Regional species richness: the role of dispersal

Colonization frequency may also influence richness patterns (Jablonski et al. 2006, Li and Wiens 2019). Therefore, we

used the DEC analysis to infer the number of colonization events for each region. We then conducted a regression of regional richness against colonizations per region (above).

Since the biogeographic regions used to differ in size, we also tested for a relationship between the area and richness of each region (Supporting information for details of area estimation). We then regressed estimated areas of regions (Supporting information) against their richness. We also tested the relationship of region size to colonization age, diversification rate, and their combined effect (above).

Diversification and dispersal across Wallace's line

We also compared diversification rates of lineages occurring east and west of Wallace's line. We first redefined four biogeographic regions (Supporting information). We used MuHiSSE (Nakov et al. 2019) implemented in R package *hisse* v1.9.19 (Beaulieu and O'Meara 2016) to build 22 models with or without hidden states (Supporting information). We then estimated the averaged speciation, extinction, and net diversification rates for states based on the best-fit model. We also used MuSSE to estimate the posterior distribution of diversification rates of regions. We then computed differences in diversification rates between regions (details in Supporting information).

If Wallace's line is a biogeographic barrier, we expected significantly fewer dispersal events that crossed Wallace's line relative to the average number of dispersal events between other pairs of regions. We first assigned inferred dispersal events (above) to different pairs of regions. We then tallied the number of dispersal events that crossed Wallace's line and those that did not. For dispersal events from nodes whose most probable ancestral region was inferred as two that span Wallace's line (e.g. region DE), it was uncertain if these dispersal events to descendent areas had crossed Wallace's line or not because both were possible (e.g. either from region D or E to its descendent areas). In these ambiguous cases, we treated them as half a dispersal event for both categories (crossing versus not). Finally, we tested for a significant difference between the means of these categories using a t-test in the R package *ggpubr* v0.4.0.

Results

Phylogeny and divergence-time estimation

BEAST generated a well-resolved phylogeny (Supporting information) that was largely consistent with the likelihood tree (Supporting information) and a recent estimate (Liu et al. 2018). Some nodes were weakly supported in both concatenated phylogenies (Supporting information). We found no strongly supported incongruence between chloroplast and nuclear datasets (Supporting information). The section *Stangerioides* remained non-monophyletic and appeared as four subclades (clades II–IV and VII in Supporting information but see Supporting information). The previous Bayesian

tree revealed three subclades, with several taxa clustered in the core *Stangerioides* clade (Liu et al. 2018). Except for sect. *Stangerioides*, all five previously described sections within *Cycas* were well supported as monophyletic (*Panzhihuaenses*, *Asiorientales Wadeae*, *Indosinenses* and *Cycas*).

Estimated ages of major nodes are summarized in Table 2. The crown age of Cycadaceae was estimated as ~12.50 Mya, with a long branch connecting it to Zamiaceae (Table 2, Supporting information).

Ancestral-area reconstruction

Ancestral-area reconstructions indicated that the region consisting of east Asia and Indochina (region A, node 1 in Fig. 1) was the most likely ancestral area for extant *Cycas* (proportional likelihood=1.00, Table 2). Results allowing four ancestral areas per node showed some differences from the two-area analysis (nodes 4, 10–16 and 19–27 in Supporting information versus Table 2, and Supporting information versus Fig. 1), and the inferred ancestral area of some internal nodes showed low relative probabilities in the four-area scheme (nodes 9, 11–13 and 20 in Supporting information). However, an assumption of four ancestral areas is less realistic based on the observed data (Supporting information).

Regional richness: colonization time, diversification rates, colonization frequency, and area

We found a significant relationship between the richness of regions and their ages of first colonization (AFC: $r^2=0.639$, $p=0.031$; Supporting information, Fig. 2a). The AFC from 10 randomly sampled trees were all strongly and positively related to richness (Supporting information), despite the broad confidence interval for the age of the *Cycas* crown node. The relationship between richness and summed colonization ages (SAC) was weaker and not significant ($r^2=0.493$, $p=0.079$; Supporting information, Fig. 2b).

The best-fitting MuSSE model supported different speciation rates but equal extinction and transition rates among regions (Supporting information). There was no significant relationship between richness and net diversification rates among regions (MuSSE: $r^2=0.254$, $p=0.249$, Supporting information, Fig. 2c; MS: $r^2=0.543$, $p=0.059$, Supporting information, Fig. 2d; Supporting information). The weak relationship may be explained by the ancestral region (east Asia and Indochina) having high richness but lower diversification rates, whereas the Sunda Shelf has low richness but higher diversification rates (Table 1).

We found a strong correlation between richness and the combined effect of time and diversification rate, compared to time alone (MuSSE: $r^2=0.791$, $p=0.007$, Supporting information, Fig. 2e; MS: $\epsilon=0.1$, $r^2=0.879$, $p=0.002$, Supporting information, Fig. 2f; Supporting information for $\epsilon=0.5, 0.9$). Regression models including this combined effect had a better fit than those including a single factor (Supporting information).

We also found a significant relationship between the area of regions and their richness ($r^2=0.666$, $p=0.025$;

Table 2. Biogeographic reconstructions and estimated ages for major nodes of *Cycas*. All events are extracted from ancestral-state reconstruction in RASP. Figure 1 for node identity. d=dispersal, v=vicariance, e=extinction, RP: relative probability. Descendant areas (DA): estimated regions of the two descendant lineages from each node inferred from the DEC analysis. HPD: highest probability density. Mya, millions of years ago.

Node	Ancestral area (RP)	Descendant areas (DA)	Estimated node			age (Mya)	95% HPD (Mya)
			d	v	e		
1 <i>Cycas</i> crown	A (1.0)	A A	0	0	0	12.50	22.18–0.79
2 <i>C. sect. Stangerioides</i> stem	A (1.0)	A A	0	0	0	10.77	18.77–0.75
3 <i>Stangerioides</i> Indochina stem	A (1.0),	A A	0	0	0	9.43	17.16–5.40
4 <i>C. sect. Wadeae</i> stem	A (0.60), AD (0.40)	D A	2	1	0	8.63	14.69–4.87
5 <i>Wadeae</i> crown	D (0.73), AD (0.27)	D D	0	0	0	4.50	8.41–0.07
6 <i>C. sect. Indosinenses</i> stem	A (1.0)	A A	0	0	0	7.81	13.23–4.42
7 Dispersal from A to B	A (1.0)	A AB	1	0	0	1.82	3.36–0.80
8 <i>C. aculeata</i>	A (0.85), AD (0.15)	A AD	0	0	0	6.92	11.98–0.50
9 <i>C. sect. Cycas</i> crown	A (0.74), AB (0.26)	A AD	1	0	0	6.10	10.59–0.47
10 Indian clade stem	AD (1.0)	B DF	2	1	1	5.38	9.21–0.35
11 Clades X–XIII crown	DF (0.79), D (0.21)	F DF	1	0	0	4.70	8.02–0.31
12 Clades X–XI crown	F (0.62), CF (0.38)	DF F	1	0	0	4.37	–
13 Clades X crown	DF (1.0)	F D	0	1	0	3.16	5.70–0.23
14	F (0.61), CF (0.39)	CF F	1	0	0	0.91	2.24–0.01
15	CF (1.0)	F CEF	2	0	0	0.22	0.83–0.0001
16	D (0.74), DE (0.26)	DE D	1	0	0	2.23	4.26–0.12
17	DE (1.0)	E DE	1	0	0	1.41	2.81–0.034
18	DE (1.0)	E D	0	1	0	0.96	2.07–0.014
19	F (0.79), EF (0.21)	E F	2	1	0	4.14	–
20 Clades XII–XIII crown	DF (1.0)	D F	0	1	0	3.65	6.19–0.25
21 Clades XII crown	D (1.0)	D DE	1	0	0	3.32	–
22	DE (0.56), D (0.23), CD (0.21)	E CD	1	1	0	2.47	6.17–0.20
23	CD (1.0)	C CD	1	0	0	2.04	4.67–0.15
24	CD (1.0)	D BC	1	1	0	1.66	3.54–0.07
25	BC (0.51), C (0.49)	BC BC	2	0	0	1.31	2.83–0.03
26	BC (1.0)	F BG	2	1	1	1.18	–
27 Asian–African disjunction	BG (0.51), G (0.49)	ACD G	3	1	1	0.72	2.19–0.01
28 Clades XIII crown	F (1.0)	F F	0	0	0	3.27	5.48–0.19

Supporting information, Fig. 2g). Area was not related to colonization time (Supporting information) or diversification rate (Supporting information), but showed significant relationships with the combined time-rate variable (Supporting information, MS; but see Supporting information for MuSSE).

There was no significant relationship between richness and colonization frequency (two ancestral regions: $r^2 < 0.001$, $p=0.953$; Supporting information, Fig. 2h; four regions: $r^2=0.159$, $p=0.375$, Supporting information).

Diversification and dispersal across Wallace's line

In the MuHiSSE analysis focusing on Wallace's line, models with hidden states were rejected. Instead, the best-fit model was a MuSSE model assuming diversification parameters are shared among lineages in the same region (Supporting information). This model estimated extremely low extinction rates in both regions across Wallace's line ($<10^{-8}$) with higher diversification rates east of Wallace's line (i.e. EW, 0.592, Fig. 3a) than west (WW, 0.188, Fig. 3a).

Density distributions for diversification rates across Wallace's line were plotted (Fig. 3b) based on the best-fitting

model in a separate MuSSE analysis (Supporting information), which favored different speciation and transition rates between regions but equal extinction rates. Similar to the seven-region results (i.e. C–F in Fig. 1, Table 1), regions east of Wallace Line showed higher diversification rates than west, using two regions (the 95% confidence interval of rate differences did not overlap with zero, Supporting information).

Dispersal events among these regions are summarized in Fig. 3c. There was no significant difference ($p=0.692$) between mean dispersal events among pairs of regions that crossed Wallace's line and pairs that did not (Fig. 3d). This extensive dispersal across Wallace's line occurred despite the relatively low richness of *Cycas* in regions adjacent to it (Fig. 1).

Discussion

The impact of time and diversification rates on richness patterns

What causes disparities in species richness among regions? We explored whether richness patterns in paleotropical

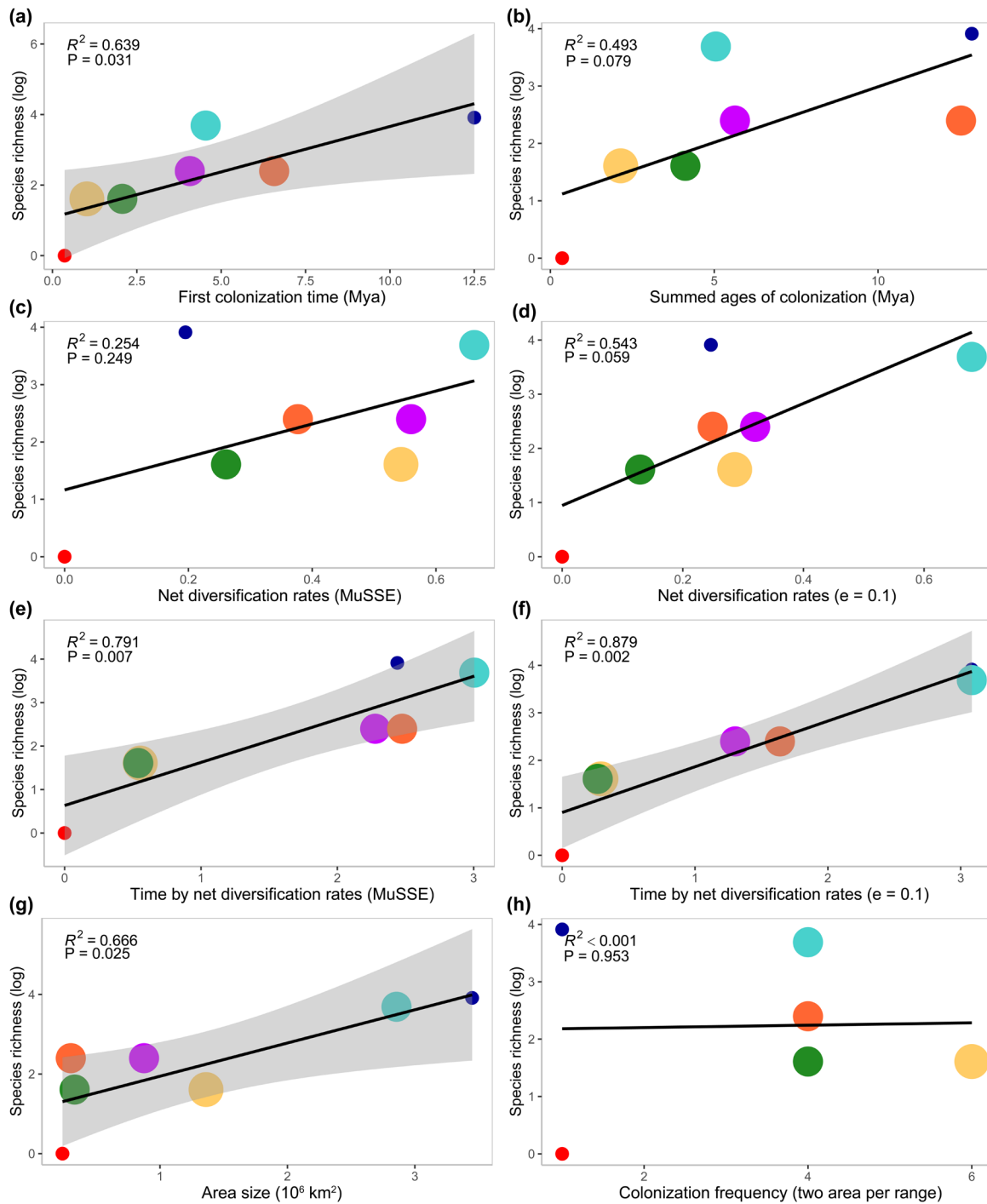


Figure 2. Relationships between regional species richness and colonization time, diversification rates, and their combined effect. (a) Strong relationship between the age of first colonization (AFC) and richness for each region ($r^2 = 0.639$, $p = 0.031$). (b) Relationship between the summed age and species richness for each region ($r^2 = 0.493$, $p = 0.079$). (c) Relationship between the net diversification rate (based on the best MuSSE model) and richness for each region ($r^2 = 0.254$, $p = 0.249$). (d) Relationship between the weighted net diversification rate (based on $\epsilon = 0.1$) and richness for each region ($r^2 = 0.543$, $p = 0.059$). (e) Strong relationship between the combined effect of time (first colonization of each region) and diversification rate (based on the best-fit MuSSE model) and richness of regions ($r^2 = 0.791$, $p = 0.007$). (f) Strong relationship between the combined effect of time (first colonization of each region) and weighted net diversification rate ($\epsilon = 0.1$) and richness of regions ($r^2 = 0.879$, $p = 0.002$). (g) Strong relationship between the area of occurrence of *Cycas* and richness of each region ($r^2 = 0.666$, $p = 0.025$). (h) Relationship between the number of colonizations of each region and regional richness ($r^2 < 0.001$, $p = 0.953$). Colored dots in each graph indicates the colored regions in Fig. 1, and the size of dots represents the relative frequency of colonizations of each region summarized by RASP. Grey areas show 95% credibility intervals. Mya, millions of years ago.

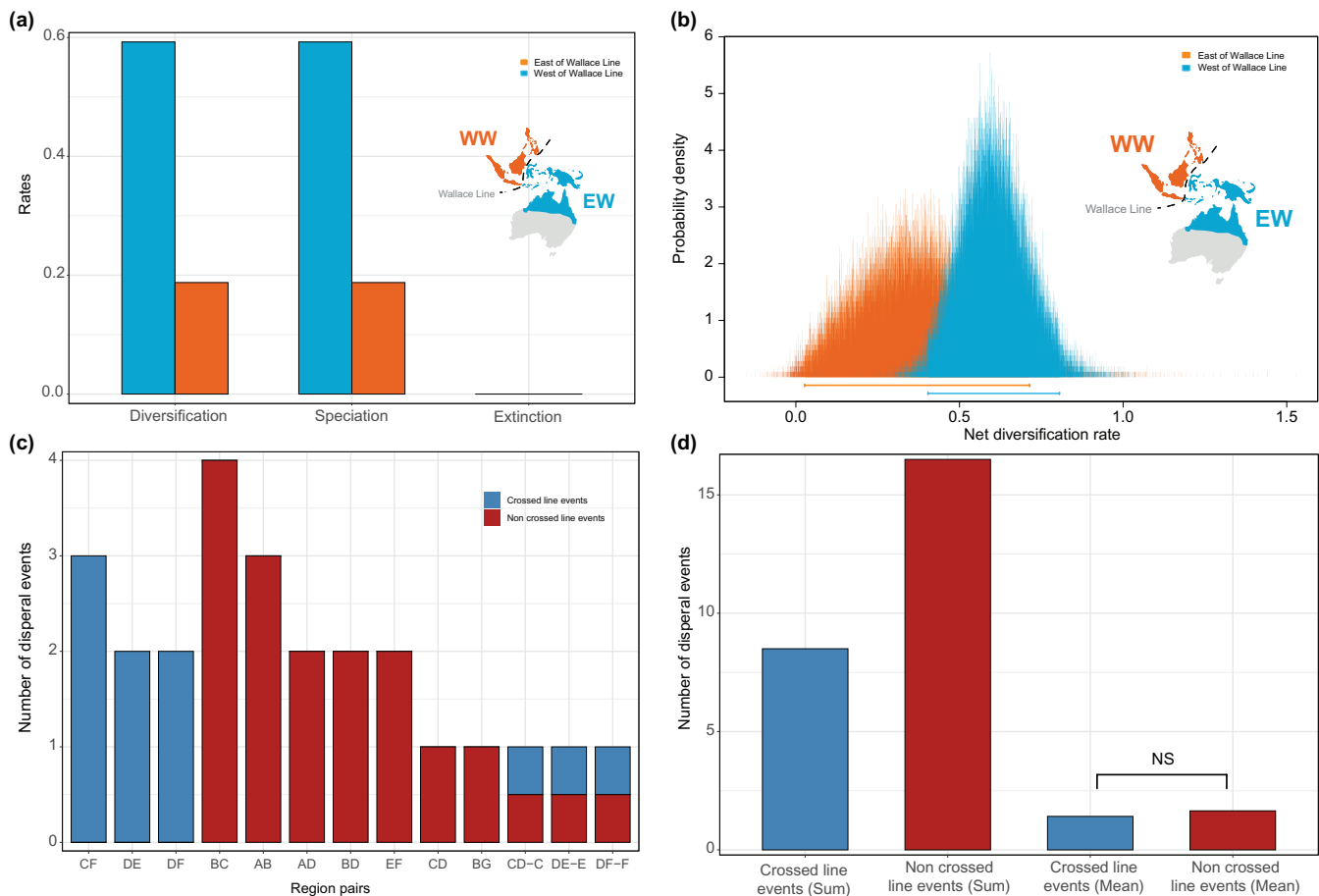


Figure 3. Overall probability density of net diversification rates and dispersal events across Wallace's Line as calculated with MuHiSSE, MuSSE and DEC respectively. (a) Net diversification, speciation and extinction rates estimated by MuHiSSE. (b) Probability density of net diversification rates based on MuSSE. Insets are defined regions. Credibility intervals of net diversification rates with each region are depicted as bars below the plot. (c) The number of dispersal events between different region pairs. (d) Summed and mean number of dispersal events that crossed Wallace's Line or not. See Supporting information for the distribution of overall net diversification-rate differences among regions across Wallace's Line.

cycads are explained by colonization time, colonization frequency, or diversification rates. Our study is one of only a handful to consider all three factors (Li and Wiens 2019). We also tested whether a variable combining colonization time and diversification rates could explain richness patterns, which only one previous study has done (Yan et al. 2018). Our results support the importance of colonization time over diversification rates or colonization frequency, but also show that the combination of time and rates best explained richness patterns.

Time (colonization age) appears to be more important than other factors because it can take substantial time to build up richness within a region through in situ speciation, even if diversification rates are high (Pontarp and Wiens 2017, Li and Wiens 2019). We recognize that a much older crown age of Cycadaceae could be inferred based on alternative calibration schemes. However, our results on richness were largely insensitive to different age estimates (Supporting information). The importance of colonization time in driving

richness patterns has been observed in many clades and regions (Jansson et al. 2013, Li and Wiens 2019).

Colonization time may also explain the low richness of *Cycas* in biodiversity hotspots. For example, studies have suggested that Borneo has high biodiversity and endemism because it is a 'cradle' for diversification (Klaus et al. 2013, Grismer et al. 2016, Williams et al. 2017). However, Borneo has only a single *Cycas* species, the coastally distributed *C. edentata* (Lindstrom et al. 2009). This species arose from dispersal in the last one million years (Fig. 1, node 27). Another similar pattern occurs in Africa (Madagascar), where colonization appears to have been too recent for in situ speciation to occur. We recognize that more ancient colonization may have occurred in some regions, followed by extinction and then more recent recolonization (Miller and Wiens 2017). However, colonizations that failed to survive to the present day have no impact on present-day richness, and our goal is to explain present-day richness patterns.

Time may dominate richness patterns at shallower temporal scales, but diversification rates may explain patterns at deeper scales, both across regions (Pyron and Wiens 2013, Wiens 2015, Schluter 2016) and among clades (Scholl and Wiens 2016). This idea is strongly supported by simulations (Pontarp and Wiens 2017). At a more recent timescale, we found that diversification rates alone were unable to explain richness patterns in *Cycas*. Nonetheless, high diversification rates were coupled with high richness for some recent groups, such as the endemic Australasian group. This group had the highest diversification rates and the second-highest richness. Similar patterns were found in the *Cycas* group endemic to India and Sri Lanka (Table 1).

We also found that the combined effects of colonization time and diversification rates best explained richness patterns. Although time alone can explain most variation in richness among regions ($r^2 = 0.639$), a model also including diversification rates had the best fit (Supporting information). This combined variable can explain some richness patterns that colonization time alone cannot. For example, the colonization of Australasian (region F) is the third oldest (AFC) compared to other regions. Yet, this region has the second-highest richness. This may be attributable to the higher diversification rate in Australasia, and the combined effects of colonization time and diversification (Table 1).

Overall, our results further support the importance of testing the combined effect of colonization time and diversification rates on spatial richness patterns. A recent review showed that studies testing the importance of colonization time remain relatively rare (Li and Wiens 2019). Studies that test this combined effect are rarer still.

We note that cycads are not angiosperms, whereas most plant species are. Importantly, previous plant studies showing the effect of colonization time (and time + diversification) focused on angiosperms (Yan et al. 2018, Li and Wiens 2019). Our results show that these patterns are not restricted to angiosperms. Indeed, impacts of colonization time have mostly been studied in animals (Li and Wiens 2019). Future studies on other paleotropical clades (both angiosperms and animals) would be invaluable for testing the patterns found here.

Other factors

Many other factors may influence spatial richness patterns in cycads (e.g. area, climate). However, other factors can only influence richness through their impacts on speciation, extinction, and dispersal, which we estimated here. Area of regions was related to richness in *Cycas* (Fig. 2g). Interestingly, area was not significantly related to colonization time or diversification rates separately, but was to their combined effect (Supporting information).

Yet, our results do not support a straightforward relationship between area and richness. In some regions (e.g. Africa), very recent colonization may explain both their limited richness and limited area occupied within the

region. Other regions were colonized long ago and have only limited area and richness (Philippines), suggesting that area might constrain diversification. Unsuitable climate (e.g. cold, aridity) might limit the area occupied to a small portion of some regions (e.g. Australia). Further untangling these relationships could be an intriguing topic for future research. However, we emphasize that area is not a competing explanation for richness relative to colonization time or diversification rate.

Climate can also be important for richness patterns, but (like area) can only impact richness by affecting speciation, extinction, and dispersal. Furthermore, climate seems unlikely to be the main cause of richness patterns in *Cycas*. For example, the two highest-latitude regions have the most species (A and F; Fig. 1), whereas regions supporting (originally) wet equatorial rainforest have limited richness (C, D, and E) as does one with considerable climatic heterogeneity (G). Furthermore, a region with extensive arid habitat (F) has the second highest richness. Climate may also impact the habitable area of some regions (above). Climatic heterogeneity in some regions might contribute to speciation (e.g. China, New Guinea). These should all be interesting topics for future research.

Dispersal and diversification across Wallace's line

Our results suggest the Philippines (northern Sunda Shelf) might be the ancestral area for *Cycas* in Malesia and that there were subsequent dispersal events across Wallace's line during the Plio-Pleistocene (Fig. 1, nodes 16 and 21). These results imply that the Philippines acted as a 'stepping stone' for taxa to cross Wallace's line from the Sunda Shelf. This result was also found in other recent studies (Atkins et al. 2020, Yu and Van Welzen 2020). Moreover, we detected several back-dispersal events across Wallace's line, from Sulawesi (Wallacea) and the Philippines to the Sunda Shelf (Fig. 1, nodes 21–25). These back-dispersal events may be linked to the repeated exposure of the Sunda Shelf when sea level was low (40–120 m lower than present) during the Pleistocene (Voris 2000, Hanebuth et al. 2011). This exposure resulted in the Malay Peninsula being contiguous with Sumatra and Borneo, allowing terrestrial dispersal (Tougaard 2001, Lohman et al. 2011, De Bruyn et al. 2014). In addition, the buoyant seeds of some *Cycas* (Dehgan and Yuen 1983) may aid oceanic LDD, which may help explain frequent cross dispersal (Fig. 3c–d) and occurrence on both sides of Wallace's line. An asymmetrical and eastwards-dominated floristic exchange between Sahul and Sunda (Crayn et al. 2015) may have also facilitated the dispersal of *Cycas* to Australasia in this period.

After crossing Wallace's line from west to east, *Cycas* appears to have had accelerated diversification. This is supported by higher rates in the eastern Sahul region compared to those in the western Sunda Shelf (Table 1, Fig. 3a–b and Supporting information). A similar acceleration was also detected in two palm genera (subfamily Livistoninae; Bacon et al. 2013). This acceleration was attributed to both

intrinsic factors (higher dispersal abilities) and extrinsic factors (novel habitats created by tectonic movements). Likewise, a previous study suggested that some traits in Sunda species allow more efficient dispersal and faster growth than in Sahul species (Yap et al. 2018). The uplift of the New Guinean highlands in the last five million years and the stable and extensive landmass in northern Australia (Hall 2009), likely provided novel niches and opportunities for speciation and ultimately shaped diversity across Wallace's line.

Main conclusion

The causes of richness patterns among regions is a major unresolved question in ecology, biogeography, and evolutionary biology. We support colonization time as the main driver of richness patterns in paleotropical cycads, a hypothesis that remains rarely tested, and the combined effects of time and diversification (which is tested even more rarely). Our results also suggest that a major biogeographic barrier (Wallace's line) does little to impede dispersal.

Funding – This work is supported by the National Natural Science Foundation of China (no. 31900184), Natural Science Foundation of Yunnan Province (202001AT070072), West Light Foundation of the Chinese Academy of Sciences (Y8246811W1), and the China Scholarship Council project (201804910257).

Author contributions

Jian Liu: Conceptualization (equal); Data curation (lead); Formal analysis (lead); Investigation (equal); Methodology (lead); Validation (equal); Visualization (lead); Writing – original draft (lead). **Anders J. Lindstrom:** Investigation (equal); Resources (equal); Validation (equal); Writing – review and editing (equal). **Nathalie S. Nagalingum:** Conceptualization (equal); Supervision (equal); Validation (equal); Writing – review and editing (equal). **John J. Wiens:** Conceptualization (equal); Methodology (equal); Supervision (equal); Validation (equal); Writing – review and editing (lead). **Xun Gong:** Conceptualization (equal); Funding acquisition (lead); Resources (lead); Supervision (equal); Validation (equal); Writing – review and editing (equal).

Transparent Peer Review

The peer review history for this article is available at <<https://publons.com/publon/10.1111/ecog.05910>>.

Data availability statement

DNA alignments, distribution data, tree files and code are available from the Dryad Digital Repository: <<https://doi.org/10.5061/dryad.8gtht76pg>> (Liu et al. 2021).

References

- Atkins, H. J. et al. 2020. A molecular phylogeny of southeast Asian *Cyrtandra* (Gesneriaceae) supports an emerging paradigm for Malaysian plant biogeography. – *Front. Biogeogr.* 12: e44814.
- Bacon, C. D. et al. 2013. Geographic and taxonomic disparities in species diversity: dispersal and diversification rates across Wallace's line. – *Evolution* 67: 2058–2071.
- Beaulieu, J. M. and O'Meara, B. C. 2016. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. – *Syst. Biol.* 65: 583–601.
- Bouckaert, R. et al. 2019. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. – *PLoS Comput. Biol.* 15: e1006650.
- Burnham, K. P. and Anderson, D. R. 2002. Model selection and multimodel inference: a practical information–theoretical approach. – Springer.
- Calonje, M. A. et al. 2020. The World List of Cycads. – <<http://www.cycadlist.org>>.
- Chantarasuwan, B. et al. 2016. Palaeotropical intercontinental disjunctions revisited using a dated phylogenetic hypothesis with nearly complete species level sampling of *Ficus* subsect. *Urostigma* (Moraceae). – *J. Biogeogr.* 43: 384–397.
- Condamine, F. L. et al. 2012. What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. – *Ecol. Lett.* 15: 267–277.
- Crayn, D. M. et al. 2015. The Sahul-Sunda floristic exchange: dated molecular phylogenies document Cenozoic intercontinental dispersal dynamics. – *J. Biogeogr.* 42: 11–24.
- Darwin, C. 1859. On the origin of species by means of natural selection. – John Murray, London.
- De Bruyn, M. et al. 2014. Borneo and Indochina are major evolutionary hotspots for southeast Asian biodiversity. – *Syst. Biol.* 63: 879–901.
- Dehgan, B. and Yuen, C. 1983. Seed morphology in relation to dispersal, evolution and propagation of *Cycas* L. – *Bot. Gaz.* 144: 412–418.
- Feng, C. M. et al. 2009. Phylogeny and biogeography of Alangiaceae (Cornales) inferred from DNA sequences, morphology and fossils. – *Mol. Phylogenet. Evol.* 51: 201–214.
- FitzJohn, R. G. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. – *Methods Ecol. Evol.* 3: 1084–1092.
- Forster, P. I. 2011. *Cycas terryana* P. I. Forst. (Cycadaceae), a new species from central Queensland. – *Austrobaileya* 8: 356–363.
- Forster, P. I. and Gray, B. 2017. *Cycas distans* P. I. Forst. & B. Gray (Cycadaceae), a new species from southern Cape York Peninsula, Queensland. – *Austrobaileya* 10: 74–84.
- Gemmill, C. et al. 2002. Evolution of insular Pacific *Pittosporum* (Pittosporaceae): origin of the Hawaiian radiation. – *Mol. Phylogenet. Evol.* 22: 31–42.
- Goldberg, E. E. et al. 2011. Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. – *Syst. Biol.* 60: 451–465.
- Grismer, L. L. et al. 2016. Out of Borneo, again and again: biogeography of the Stream Toad genus *Ansonia* Stoliczka (Anura: Bufonidae) and the discovery of the first limestone cave-dwelling species. – *Biol. J. Linn. Soc.* 120: 371–395.
- Hall, R. 2009. Southeast Asia's changing palaeogeography. – *Blumea* 54: 148–161.
- Hanebuth, T. J. J. et al. 2011. Formation and fate of sedimentary depocentres on southeast Asia's Sunda Shelf over the past sea-

- level cycle and biogeographic implications. – *Earth-Sci. Rev.* 104: 92–110.
- Harmon, L. J. et al. 2008. GEIGER: investigating evolutionary radiations. – *Bioinformatics* 24: 129–131.
- Heno Diaz, L. F. et al. 2019. Macroevolutionary diversification rates show time dependency. – *Proc. Natl Acad. Sci. USA* 116: 7403–7408.
- Hill, K. D. 2004. *Cycas candida* (Cycadaceae), a new species from Queensland together with an extension of range and amended description of *Cycas yorkiana*. – *Telopea* 10: 607–611.
- Hill, K. D. and Osborne, R. 2001. *Cycads of Australia*. – Kangaroo Press.
- Hutter, C. R. et al. 2013. Explaining Andean megadiversity: the evolutionary and ecological causes of glassfrog elevational richness patterns. – *Ecol. Lett.* 16: 1135–1144.
- Hutter, C. R. et al. 2017. Rapid diversification and time explain amphibian richness at different scales in the tropical Andes, earth's most biodiverse hotspot. – *Am. Nat.* 190: 828–843.
- Jablonski, D. et al. 2006. Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. – *Science* 314: 102–106.
- Jansson, R. et al. 2013. What can multiple phylogenies say about the latitudinal diversity gradient? A new look at the tropical conservatism, out of the tropics and diversification rate hypotheses. – *Evolution* 67: 1741–1755.
- Jutta, M. et al. 2010. A new species of *Cycas* (Cycadaceae) from peninsular Malaysia. – *Blumea* 55: 249–252.
- Klaus, S. et al. 2013. Out of Borneo: Neogene diversification of Sundaic freshwater crabs (Crustacea: Brachyura: Gecarcinucidae: *Parathelphusa*). – *J. Biogeogr.* 40: 63–74.
- Kozak, K. H. and Wiens, J. J. 2016. Testing the relationships between diversification, species richness and trait evolution. – *Syst. Biol.* 65: 975–988.
- Li, H. and Wiens, J. J. 2019. Time explains regional richness patterns within clades more often than diversification rates or area. – *Am. Nat.* 193: 514–529.
- Lindstrom, A. et al. 2009. The genus *Cycas* (Cycadaceae) in Indonesia. – *Telopea* 12: 385–418.
- Liu, J. et al. 2018. Phylogeny of the gymnosperm genus *Cycas* L. (Cycadaceae) as inferred from plastid and nuclear loci based on a large-scale sampling: evolutionary relationships and taxonomic implications. – *Mol. Phylogenet. Evol.* 127: 87–97.
- Liu, J. et al. 2021. Data from: Testing the causes of richness patterns in the paleotropics: Time and diversification in cycads (Cycadaceae). – Dryad Digital Repository, <<http://doi.org/10.5061/dryad.8gtht76pg>>.
- Lohman, D. J. et al. 2011. Biogeography of the Indo-Australian Archipelago. – *Annu. Rev. Ecol. Evol. Syst.* 42: 205–226.
- Magallon, S. and Sanderson, M. J. 2001. Absolute diversification rates in angiosperm clades. – *Evolution* 55: 1762–1780.
- Mankga, L. T. et al. 2020. The cycad genus *Cycas* may have diversified from Indochina and occupied its current ranges through vicariance and dispersal events. – *Front. Ecol. Evol.* 8: 1–13.
- Meyer, A. L. and Wiens, J. J. 2018. Estimating diversification rates for higher taxa: BAMM can give problematic estimates of rates and rate shifts. – *Evolution* 72: 39–53.
- Meyer, A. L. S. et al. 2018. BAMM gives misleading rate estimates in simulated and empirical datasets. – *Evolution* 72: 2257–2266.
- Miller, E. C. and Wiens, J. J. 2017. Extinction and time help drive the marine-terrestrial biodiversity gradient: is the ocean a death-trap? – *Ecol. Lett.* 20: 911–921.
- Minh, B. Q. et al. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. – *Mol. Biol. Evol.* 37: 1530–1534.
- Morlon, H. 2014. Phylogenetic approaches for studying diversification. – *Ecol. Lett.* 17: 508–525.
- Myers, N. et al. 2000. Biodiversity hotspots for conservation priorities. – *Nature* 403: 853–858.
- Nakov, T. et al. 2019. Diatoms diversify and turn over faster in freshwater than marine environments. – *Evolution* 73: 2497–2511.
- Paradis, E. and Schliep, K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. – *Bioinformatics* 35: 526–528.
- Pontarp, M. and Wiens, J. J. 2017. The origin of species richness patterns along environmental gradients: uniting explanations based on time, diversification rate and carrying capacity. – *J. Biogeogr.* 44: 722–735.
- Pyron, R. A. and Wiens, J. J. 2013. Large-scale phylogenetic analyses reveal the causes of high tropical amphibian diversity. – *Proc. R. Soc. B.* 280: 20131622.
- Rabosky, D. L. and Benson, R. B. 2021. Ecological and biogeographic drivers of biodiversity cannot be resolved using clade age-richness data. – *Nat. Commun.* 12: 1–10.
- Rao, P. V. C. et al. 2016. *Cycas seshachalamensis* (Cycadaceae), a new cycad from Andhra Pradesh, India. – *Asian J. Conserv. Biol.* 5: 55–58.
- Ree, R. H. and Smith, S. A. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction and cladogenesis. – *Syst. Biol.* 57: 4–14.
- Ricklefs, R. E. 1987. Community diversity: relative roles of local and regional processes. – *Science* 235: 167–171.
- Ricklefs, R. E. 2006. Evolutionary diversification and the origin of the diversity–environment relationship. – *Ecology* 87: S3–S13.
- Ricklefs, R. E. 2007. Estimating diversification rates from phylogenetic information. – *Trends Ecol. Evol.* 22: 601–610.
- Schluter, D. 2016. Speciation, ecological opportunity and latitude. – *Am. Nat.* 187: 1–18.
- Scholl, J. P. and Wiens, J. J. 2016. Diversification rates and species richness across the Tree of Life. – *Proc. R. Soc. B* 283: 20161334.
- Singh, R. and Radha, P. 2006. A new species of *Cycas* from the Malabar coast, western Ghats, India. – *Brittonia* 58: 119–123.
- Singh, R. et al. 2015. A new species, a new combination and a new subsection of *Cycas* from Odisha, northern Eastern Ghats of India. – *Asian J. Conserv. Biol.* 4: 3–14.
- Sirichamorn, Y. et al. 2014. Historical biogeography of *Aganope*, *Brachypterum* and *Derris* (Fabaceae, tribe Millettieae): insights into the origins of Palaeotropical intercontinental disjunctions and general biogeographical patterns in Southeast Asia. – *J. Biogeogr.* 41: 882–893.
- Srivastava, R. 2014. A new species of *Cycas* from India. – *Ind. J. Plant Sci.* 3: 109–110.
- Stephens, P. R. and Wiens, J. J. 2003. Explaining species richness from continents to communities: the time-for-speciation effect in emydid turtles. – *Am. Nat.* 161: 112–128.
- Thomas, D. et al. 2012. West to east dispersal and subsequent rapid diversification of the mega-diverse genus *Begonia* (Begoniaceae) in the Malesian archipelago. – *J. Biogeogr.* 39: 98–113.
- Tougaard, C. 2001. Biogeography and migration routes of large mammal faunas in south-east Asia during the late middle Pleistocene: focus on the fossil and extant faunas from Thailand. – *Paleogeogr. Paleoclimatol. Paleoecol.* 168: 337–358.
- Voris, H. K. 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. – *J. Biogeogr.* 27: 1153–1167.

- Wallace, A. R. 1869. The Malay Archipelago: the land of the orangutan and the bird of paradise; a narrative of travel, with studies of man and nature. – Macmillan and Co.
- Wiens, J. J. 2011. The causes of species richness patterns across space, time and clades and the role of ‘ecological limits’. – Q. Rev. Biol. 86: 75–96.
- Wiens, J. J. 2015. Faster diversification on land than sea helps explain global biodiversity patterns among habitats and animal phyla. – Ecol. Lett. 18: 1234–1241.
- Williams, E. W. et al. 2017. Out of Borneo: biogeography, phylogeny and divergence date estimates of *Artocarpus* (Moraceae). – Ann. Bot. 119: 611–627.
- Yan, H. F. et al. 2018. What explains high plant richness in east Asia? Time and diversification in the tribe Lysimachieae (Primulaceae). – New Phytol. 219: 436–448.
- Yap, J. Y. S. et al. 2018. Filters of floristic exchange: how traits and climate shape the rain forest invasion of Sahul from Sunda. – J. Biogeogr. 45: 838–847.
- Yu, R. Y. and Van Welzen, P. C. 2020. Historical biogeography of *Trigonostemon* and *Dimorphocalyx* (Euphorbiaceae). – Bot. J. Linn. Soc. 192: 333–349.
- Yu, Y. et al. 2019. RASP 4: Ancestral state reconstruction tool for multiple genes and characters. – Mol. Biol. Evol. 37: 604–606.