Cenozoic imprints on the phylogenetic structure of palm species assemblages worldwide

W. Daniel Kissling[a,1,2], Wolf L. Eiserhardt[a,1], William J. Baker[b], Finn Borchsenius[a], Thomas L. P. Couvreur[c], Henrik Balslev[a], and Jens-Christian Svenning[a]

[a]Ecoinformatics and Biodiversity Group, Department of Bioscience, Aarhus University, DK-8000 Aarhus C, Denmark; [b]Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, United Kingdom; and [c]Institut de Recherche pour le Développement, UMR DIADE, F-34394 Montpellier cedex 5, France

Edited by Robert E. Ricklefs, University of Missouri, St. Louis, MO, and approved March 28, 2012 (received for review December 14, 2011)

Despite long-standing interest in the origin and maintenance of species diversity, little is known about historical drivers of species assemblage structure at large spatiotemporal scales. Here, we use global species distribution data, a dated genus-level phylogeny, and paleo-reconstructions of biomes and climate to examine Cenozoic imprints on the phylogenetic structure of regional species assemblages of palms (Arecaceae), a species-rich plant family characteristic of tropical ecosystems. We find a strong imprint on phylogenetic clustering due to geographic isolation and in situ diversification, especially in the Neotropics and on islands with spectacular palm radiations (e.g., Madagascar, Hawaii, and Cuba). Phylogenetic overdispersion on mainlands and islands corresponds to biotic interchange areas. Differences in the degree of phylogenetic clustering among biogeographic realms are related to differential losses of tropical rainforests during the Cenozoic, but not to the cumulative area of tropical rainforest over geological time. A largely random phylogenetic assemblage structure in Africa coincides with severe losses of rainforest area, especially after the Miocene. More recent events also appear to be influential: phylogenetic clustering increases with increasing intensity of Quaternary glacial-interglacial climatic oscillations in South America and, to a lesser extent, Africa, indicating that specific clades perform better in climatically unstable regions. Our results suggest that continental isolation (in combination with limited long-distance dispersal) and changing climate and habitat loss throughout the Cenozoic have had strong impacts on the phylogenetic structure of regional species assemblages in the tropics.

biodiversity | biogeography | climate change | evolution | extinction

Despite long-standing interest, the mechanisms behind the origin and maintenance of high tropical biodiversity remain elusive (1, 2). Much recent macroecological research has focused on explaining large-scale species richness gradients by contemporary climate (3–5), but the importance of evolutionary diversification and past environments has also been highlighted (2, 6–8). Over the past decades it has become increasingly clear that local community structure depends on both local processes and large-scale factors that influence regional species diversity (9, 10). Regional species assemblages are jointly shaped by within-region diversification and dispersal between regions, the former being constrained by time for speciation (7, 11) and by climatic or other factors influencing net diversification rates (12, 13), and the latter by the formation and disappearance of dispersal barriers (14, 15), time for dispersal (16), and phylogenetic niche conservatism (17). Understanding the historical assembly and present-day structure of regional species assemblages thus requires integration of ecological, paleogeographic, and phylogenetic information (10).

Increasing availability of phylogenies has ignited interest in the phylogenetic structure of species assemblages (18–20). Numerous studies have used this approach to examine the assembly of local communities (21–24). However, measures of assemblage phylogenetic structure may also reveal large-scale biogeographic processes (25, 26), although this approach has rarely been taken in biogeography (27). Several consequences of large-scale biogeographic processes can be predicted for the phylogenetic structure of species assemblages (Table 1). First, strong dispersal barriers should lead to in situ diversification within major biogeographic realms or continents. As a consequence, regional species assemblages are expected to consist predominantly of species that are relatively closely related (hypothesis H1). Second, it has been suggested that lineages inhabiting large areas over extended geological time experience higher speciation rates and lower extinction rates than lineages inhabiting small areas ("time-integrated area effect") (12). The resulting higher net diversification should lead to phylogenetically clustered assemblages at regional scales (35) (hypothesis H2). Third, the contraction of major habitat types ("biomes") during the Cenozoic should have resulted in severe extinction (31, 32). If net diversification decreases with the loss of biome area, most lineage divergences should be ancient and we could expect a tendency toward random phylogenetic structuring of species assemblages (hypothesis H3). Finally, the frequency and magnitude of Quaternary climatic oscillations has also been suggested as a major driver of assemblage structure, specifically for species richness and endemism (8, 36, 37). Although never assessed on a global scale, this could also have major consequences for the phylogenetic structure of species assemblages, for example, by favoring clades that perform well in such variable environments (hypothesis H4) (33).

Here, we combine community phylogenetic and macroecological methods to test these four core hypotheses in historical biogeography (Table 1). We use a unique dataset of global species distributions and a dated phylogeny of an important tropical plant lineage: the palm family (Arecaceae). Palms are diverse (>2,400 species in 183 genera), are a characteristic element of tropical ecosystems worldwide (38), and have served as a model system in geographical ecology (39) and rainforest evolution (40). Global distribution patterns in palms are well documented (37, 41), and phylogenetic relationships within the family are well understood (42). Most present-day palm diversity has evolved during the Cenozoic (40). For these reasons, palms are ideally suited to study the evolutionary imprints of Cenozoic history on tropical biotic assemblages at a global scale. Using the Net Relatedness Index (NRI) (18), we quantify the phylogenetic structure of palm assemblages and relate it to paleo-reconstructions of climate and biomes across most of the Cenozoic. By quantifying to what extent species that co-occur are more (NRI > 0) or less (NRI < 0) closely related than expected by random sampling from a species pool, NRI can reveal ecological and evolutionary mechanisms of assemblage structure. Scaling sampling pools to different spatial extents provides insights into the critical scales at which the assembly processes operate (26).

Author contributions: W.D.K., W.L.E., F.B., and J.-C.S. designed research; W.D.K., W.L.E., W.J.B., and F.B. performed research; W.D.K., W.J.B., and T.L.P.C. contributed new reagents/analytic tools; W.D.K. and W.L.E. analyzed data; and W.D.K., W.L.E., W.J.B., F.B., T.L.P.C., H.B., and J.-C.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]W.D.K. and W.L.E. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: danielkissling@web.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1120467109/-/DCSupplemental.

Table 1. Four core hypotheses in historical biogeography and their predictions for the phylogenetic structure of regional species assemblages

Hypotheses	Prediction	References
H1: The long-term geographic isolation of continents and biogeographic realms has caused dispersal limitation and in situ diversification.	Phylogenetic clustering within continents and biogeographic realms increases with an increasing spatial extent of the sampling pool (e.g., from continental to hemispheric and global extent).	20, 23, 28, 29
H2: Large habitat areas over deep geological time increase speciation rates and decrease extinction rates ["time-integrated species-area effect" *sensu* (12)].	Regions with large time-integrated areas of suitable habitat show higher phylogenetic clustering than regions where suitable habitat is limited over deep geological time.	12, 29, 30
H3: Strong loss of biome area over Cenozoic time scales leads to decreasing diversification rates (less speciation, more extinction).	Regions with high initial, but strongly decreasing, biome area—and thus diversification rates—will show less phylogenetic clustering than areas with constant or increasing diversification rates.	31, 32
H4: Areas with high Quaternary climate change harbor assemblages of survivors and/or species that were able to recolonize such areas after local extinction.	Areas with high paleoclimatic amplitudes (low stability) are characterized by phylogenetic clustering because assemblages predominantly consist of species with phylogenetically conserved traits that enable them to survive in or quickly recolonize such environments.	33, 34

Results

Phylogenetic Structure with a Global Sampling Pool. Across regional assemblages, NRI values calculated with a global sampling pool (Fig. 1*A*) ranged from −1.5 to 40.9 (median = 2.0). The global mean was significantly > 0 (*SI Appendix*, Table S1), indicating an overall predominance of phylogenetic clustering. A number of islands stood out with remarkably high NRI values (*SI Appendix*, Table S2; Fig. 1*A*), including Madagascar (NRI = 40.9), New Caledonia (NRI = 22.0), Hawaii (NRI = 20.7), and Cuba (NRI = 18.1). However, the majority of islands had small or intermediate NRI values (median = 2.6), and there was no general statistical difference in mean NRI between islands and continental geographic units at a global scale [$t = -1.4$, df = 150, $P = 0.160$, data ln $(x + 2)$ transformed]. NRI values calculated with a global sampling pool (blue boxes in Fig. 2) were on average significantly larger than zero in South America, Indomalaya, and Australasia, indicating phylogenetic clustering. In contrast, mean NRI did not statistically differ from zero in Africa, indicating an overall random phylogenetic structure. Moreover, Africa harbored the only three regional assemblages (Uganda, Burundi, and the Cape Province) with phylogenetic overdispersion (Fig. 1*A*) using a global sampling pool.

Geographic Isolation and Sampling Pool Scaling (H1). Restricting the sampling pool to a given hemisphere or biogeographic realm caused significant changes in NRI (*SI Appendix*, Fig. S1 and Table S3). NRI values in South America, Indomalaya, and Australasia decreased consistently with a decreasing sampling pool extent (Fig. 2), indicating a strong effect of geographic isolation (with limited dispersal among regions) and in situ diversification within three of four realms (supporting hypothesis H1). The strongest decrease in NRI was observed in South America (Fig. 2). Africa, in contrast, showed a predominantly random phylogenetic structure independent of the sampling pool extent (Fig. 2). Across realms, changes in the spatial distribution of NRI values were minor between global and hemispheric sampling pools (Fig. 1 *A* and *B*). However, geographically localized phylogenetic structuring became evident when reducing the sampling pool extent to realms (Fig. 1*C*). Notably, phylogenetic overdispersion emerged in biogeographic contact zones (Colombia, Wallacea).

Tropical Rainforest Distribution Through Time (H2 and H3). We quantified the historical extent of tropical rainforests throughout the Cenozoic (Fig. 3). All four biogeographic realms have harbored large rainforest areas since the Eocene, with Africa having by far the highest time-integrated area [area under the curve (AUC)] (Fig. 3). However, in disagreement with hypothesis H2, mean NRI values of regional palm assemblages did not consistently increase with AUC (Fig. 4*A*). Rainforest area decreased throughout the Cenozoic in all realms; this loss was much more dramatic in Africa ($>18 \times 10^6$ km^2 loss) than in Australasia (6×10^6 km^2), South America (4×10^6 km^2), and Indomalaya (3×10^6 km^2). In three of four regions, rainforest losses were most pronounced after the Miocene (11 Mya) (Fig. 3 *B*–*D*), especially in Africa (Fig. 3*B*). In agreement with hypothesis H3, the realm with the highest rate of biome loss (Africa) showed the lowest mean NRI (Fig. 4*B*). Also consistent with hypothesis H3, the realm with the largest minimum Cenozoic rainforest area (South America) showed the highest mean NRI (Fig. 4*C*).

Quaternary Climate Change (H4). Climatic oscillations (temperature anomalies) during the Quaternary were strongest at the northern range boundary of the palm family (*SI Appendix*, Fig. S2). Within realms, Quaternary climate change peaked in southern and eastern South America, eastern Africa, northern Indomalaya, and Australia (*SI Appendix*, Fig. S2). In contrast to continental sites, islands generally showed low temperature anomalies. Globally, the relationship between temperature anomaly and phylogenetic structure of palm assemblages was not significant (*SI Appendix*, Fig. S3). However, at the realm scale, South American and African palm assemblages showed increasing NRI with increasing temperature anomaly, with no significant relationship in Indomalaya and Australasia (Fig. 5; *SI Appendix*, Table S4). The Quaternary climate change effects in South America and Africa remained in multiple-predictor models when accounting for covariation with present-day environment (*SI Appendix*, Table S5). Furthermore, the commonly observed increase of species richness with contemporary precipitation was not reflected in NRI (*SI Appendix*, Fig. S4).

Discussion

Diversification in Isolation (H1). The overall phylogenetic clustering of regional palm assemblages in South America, Indomalaya, and Australasia (Fig. 2) is consistent with hypothesis H1 and reflects that many higher-level palm taxa are endemic to continents, biogeographic realms, or islands (38, 39) (*SI Appendix*, Table S6). This provides evidence for a strong role of in situ diversification within biogeographic realms—and limited dispersal between them—in the formation of regional biota. Limited long-distance dispersal in palms is evident from few species having seeds suitable for oceanic drift, few palm genera being represented on both sides of the Atlantic, and a high degree of endemism (*SI Appendix*, *SI Text S1*). The decrease of NRI with sampling pool extent is further consistent with a strong dispersal limitation at higher taxonomic levels (39, 43) (*SI Appendix*, Table S6). Phylogenetic clustering might further be influenced by dispersal limitation at finer spatial scales.

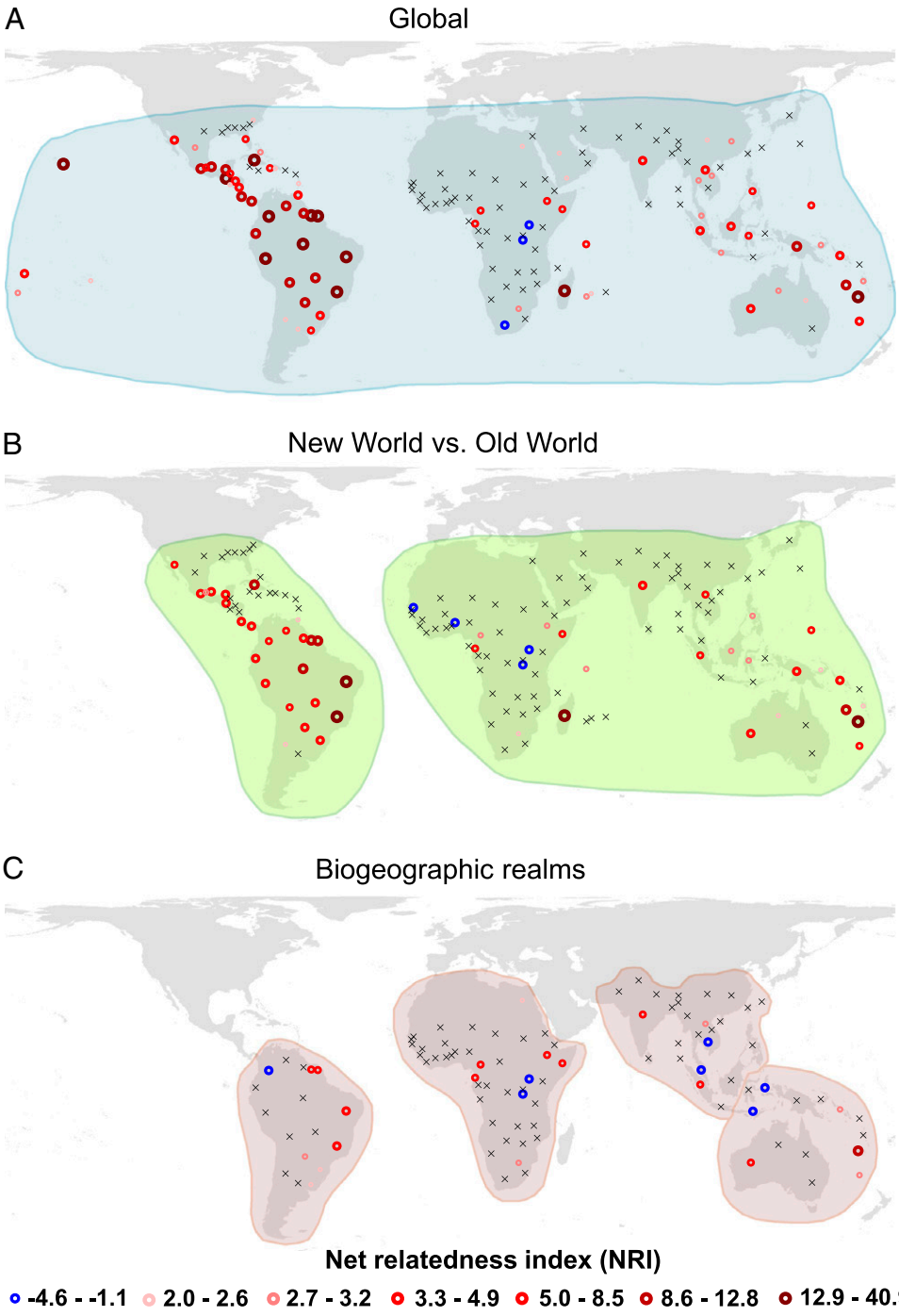

Fig. 1. The phylogenetic structure of palm assemblages given species pools restricted to (*A*) global (shaded blue), (*B*) hemispheric (New World vs. Old World) (shaded green), and (*C*) biogeographic realm (shaded brown) extent. The NRI is plotted for the mass centroid of each sample unit ("botanical country"). Blue circles indicate significantly ($P < 0.05$) negative NRI values (phylogenetic overdispersion). Red circles represent significantly positive NRI values (phylogenetic clustering) with darker red and larger circles indicating increases in NRI (quantile classification based on values in *A*). An "x" indicates nonsignificant NRI values. Maps are in Behrmann projection.

Specifically, frugivorous understory birds with low dispersal capacities might have promoted palm speciation in the undergrowth of tropical rainforests (44) (*SI Appendix*, *SI Text S1*). On islands, outstandingly high NRI values coincide with spectacular in situ radiations (*SI Appendix*, Table S2), for example, *Dypsis* on Madagascar (45), *Pritchardia* on Hawaii (46), *Coccothrinax* on Cuba (38), and three subtribes, Archontophoenicinae, Basseliniinae, and Clinospermatinae on New Caledonia (47). However, there are also many islands with low-to-medium NRI values, probably because these islands are rather small and too recently colonized for major radiations to happen.

Availability of Time-Integrated Area (H2). We used the distribution of tropical rainforests to characterize habitat availability for palms over geological time because palms are used as paleo-indicators for this biome (31, 32, 48, 49) and their present-day distribution, species richness, and net diversification rate are related to humid megathermal climates (37, 39, 40, 50). Time-integrated biome area

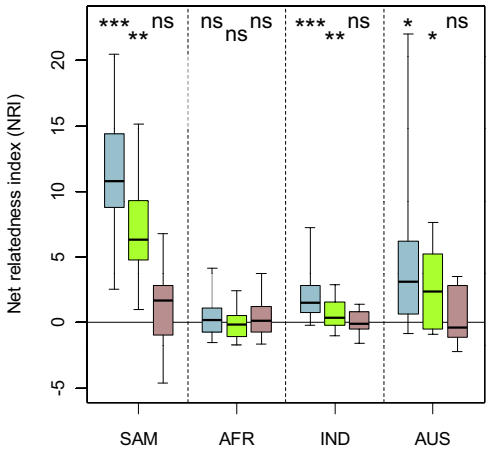

Fig. 2. Effect of sampling pool extent on phylogenetic structure of palm assemblages in South America (SAM), Africa (AFR), Indomalaya (IND), and Australasia (AUS). Box plots summarize values of the NRI within each biogeographic realm for a given sampling pool. Colors of sampling pool extent as in Fig. 1. Statistical significance of whether mean NRI differs from zero is given above each box plot (ns, not significant; $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; see *SI Appendix*, Table S1 for details of the tests). Boxes represent the interquartile range (IQR), horizontal lines within the boxes represent medians, and whiskers extend to 1.5 times the IQR.

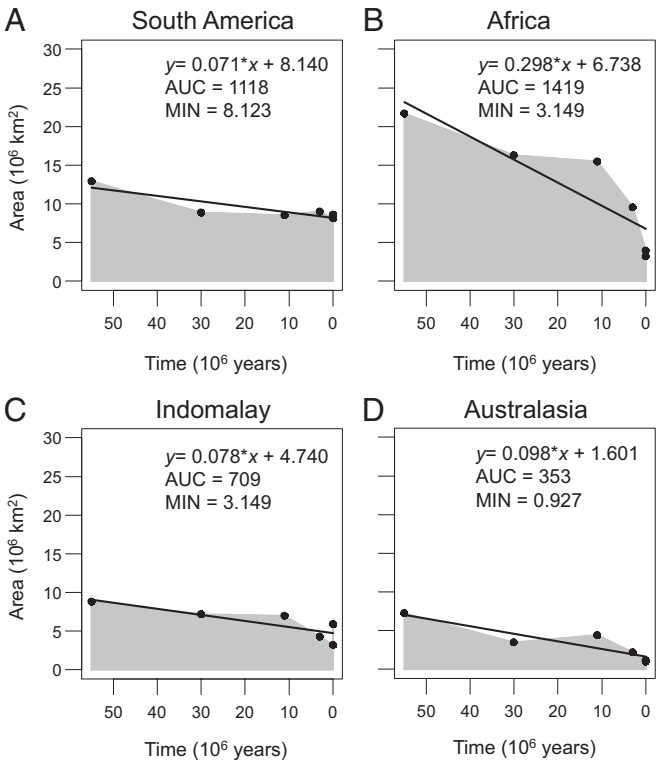

Fig. 3. Area-time plots representing the area of tropical rainforest over geological time: (*A*) South America, (*B*) Africa, (*C*) Indomalaya, and (*D*) Australasia. The area under the curve (AUC) is illustrated in gray and measures the time-integrated area of this biome from the Eocene to the present (10^6 km^2/55 Mya). Black lines indicate a simple linear regression. The steepness (i.e., negative slope β) of these regression lines (given in the regression formulas) approximates the rate of biome loss over geological time. The minimum area (MIN, in 10^6 km^2) during the Cenozoic is also provided.

has been suggested as an important driver of differences in species diversity between tropical and temperate regions (12). This effect should be reflected in assemblage phylogenetic structure (hypothesis H2). For South America, the pronounced phylogenetic clustering is consistent with a relatively large and stable rainforest area over geological time (Fig. 4*A*). However, our results for Africa suggest that phylogenetic assemblage structure at a realm scale can be more sensitive to temporal dynamics in biome area than to absolute time-integrated biome area (Fig. 4), indicating that the latter is insufficient to explain major differences between realms.

Biome Loss over Geological Time (H3). The major continental-scale differences in phylogenetic assemblage structure seem to be partly driven by Cenozoic change of biome area (Fig. 4 *B* and *C*). Africa stands out with the strongest biome loss (Fig. 4*B*), i.e., a very large extent of tropical rainforest in the Mid-Eocene, which subsequently declined and substantially so from the Mid-Miocene onward (Fig. 3*B*). The African fossil record of palms indicates severe Tertiary extinction (51). It is therefore likely that decreasing speciation and increasing extinction rates due to dramatic rainforest decline have caused an overall random phylogenetic structure of African palm assemblages (Fig. 2). Interestingly, there is a paucity of small-fruited palms in Africa (*SI Appendix*, Fig. S5), as can functionally be expected from its Cenozoic drying (*SI Appendix*, *SI Text S1*). In South America, the strong phylogenetic clustering might further be related to a relatively large minimum area of tropical rainforest (Fig. 4*C*), suggesting a smaller bottleneck effect in this realm. In addition, the upheaval of the Andes and other paleogeographical reorganizations in the Miocene and Pliocene (52, 53) might also play an important role in the Neotropics.

Quaternary Survival and Recolonization (H4). Our results show that Quaternary climatic oscillations can affect the phylogenetic structure of tropical species assemblages (hypothesis H4). For palm assemblages in South America, the strong observed relationship between NRI and Quaternary temperature anomaly (Fig. 5*A*) could be driven by tribe Cocoseae (*SI Appendix*, Fig. S6), which dominates assemblages in eastern South America. These areas are drier and more seasonal than the Amazon basin and characterized by high Quaternary temperature oscillations (*SI Appendix*, Fig. S2). We suggest that strong Quaternary climate oscillations have favored the survival and diversification of palm lineages adapted to dry and seasonal climates (e.g., Cocoseae in South America; cf. *SI Appendix*, Fig. S6) and have prevented the survival of, or colonization by, species from lineages adapted to warm, wet rainforest environments (cf. 33). A similar filtering might explain the relationship between NRI and Quaternary climate change in Africa. The absence of such an effect in Indomalaya and Australasia might be due to these regions' predominance of islands where Quaternary climate change was less severe (*SI Appendix*, Fig. S2) than on continents, possibly due to oceanic buffering (54) and the expansion of rainforests in these areas during glacials (55).

Conclusions. Our study demonstrates that broad-scale patterns in phylogenetic assemblage structure are consistent with differences in long-term historical drivers: continental isolation in combination with limited long-distance dispersal, Cenozoic habitat loss, and Quaternary climate instability. The relative importance of Quaternary and deep-time climate change depends on spatial scale (global, within vs. between realms), geographic settings (barriers and degree of isolation, continents vs. islands), and the unique history of biogeographic realms (e.g., dramatic biome changes in Africa, Quaternary climate change effects in South America). In addition to the four major biogeographic hypotheses tested here, phylogenetic assemblage structure might provide additional insights into other large-scale biogeographic processes: for example, phylogenetic overdispersion might indicate areas of biotic interchange (e.g., Colombia, Wallacea), and high levels of phylogenetic clustering might be facilitated through biotic interactions (e.g., via dispersers, herbivores, pollinators, and pathogens). Measures of phylogenetic assemblage structure such as NRI can thus provide insights into biogeographic and evolutionary processes at the assemblage level. We see great potential for

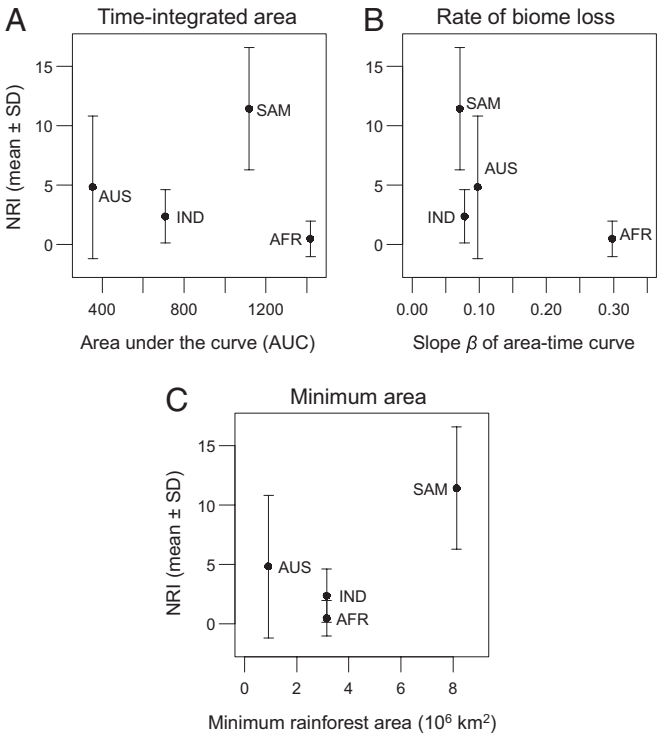

Fig. 4. The relationship between net relatedness (NRI) of palm assemblages and habitat availability over geological time (AFR, Africa; AUS, Australasia; IND, Indomalay; SAM, South America). Habitat availability was specified as (*A*) area under the curve (AUC) of area-time plots, (*B*) rate of biome loss (measured as slope β of a simple linear regression of area vs. time), and (*C*) minimum area of tropical rainforest during the Cenozoic. Compare with Fig. 3.

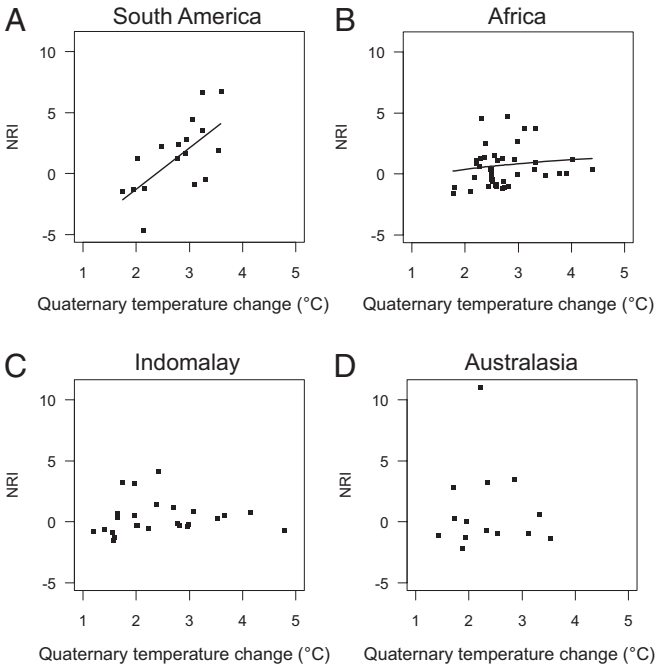

Fig. 5. The effect of Quaternary climatic oscillations on the net relatedness index (NRI, calculated with a realm sampling pool): (*A*) South America, (*B*) Africa, (*C*) Indomalaya, and (*D*) Australasia. Quaternary climatic oscillations were quantified as the change in mean annual temperature (anomaly) between the Last Glacial Maximum (~0.021 Mya) and the present (in °C; *SI Appendix*, Fig. S2). See *SI Appendix*, Tables S4 and S5, for statistical results.

better understanding the origins of tropical biodiversity by integrating geological, geographic, and paleoclimatic reconstructions with phylogenetic and species distribution data at large spatiotemporal scales.

Materials and Methods

Assemblage Data. Presence and absence of all palm species (n = 2,440) in all level 3 geographic units ("botanical countries") of the World Geographical Scheme for Recording Plant Distributions (56) was obtained from the World Checklist of Palms (41). The units are often countries, but very small countries are omitted whereas very large countries are subdivided according to states or provinces. We used an updated version of the World Checklist of Palms (downloaded on March 9, 2009 from http://apps.kew.org/wcsp) excluding introduced occurrences.

Phylogeny. We used a dated phylogeny of the 183 palm genera (40) (*SI Appendix*, Fig. S7), which is based on a recent supertree, the most extensive phylogenetic study of the family published to date (42). The phylogeny was dated using a Bayesian relaxed molecular clock approach with uncorrelated rates and calibrated using four palm fossil taxa (40). Below the genus level, species were appended as polytomies with a divergence age arbitrarily set at two-thirds the stem node age of the genus. A sensitivity analysis indicated that our results are not dependent on this arbitrarily set mean divergence age within genera (*SI Appendix*, Fig. S8).

Phylogenetic Assemblage Structure. We calculated the NRI (18) for each assemblage with more than one palm species (n = 151). NRI measures how mean phylogenetic distance (MPD) between all species pairs in the assemblage deviates from random. The random expectation is computed from "null" assemblages that are randomly sampled from a predefined species pool. NRI is calculated as:

$$\mathrm{NRI} = -1 \times (\mathrm{MPD}_{\mathrm{obs}} - \mathrm{meanMPD}_{\mathrm{rnd}})/\mathrm{sdMPD}_{\mathrm{rnd}},$$

where $\mathrm{MPD}_{\mathrm{obs}}$ is the observed MPD of a given assemblage measured in million years, $\mathrm{meanMPD}_{\mathrm{rnd}}$ the mean of the MPD values of the null assemblages, and $\mathrm{sdMPD}_{\mathrm{rnd}}$ the SD of the MPD values of the null assemblages. Values near zero indicate phylogenetically random assemblages, and deviations indicate overdispersion (<0) or clustering (>0) (18, 24). NRI was calculated with "picante" (57) in R (58) using the function ses.mpd(). We simulated null assemblages using the "taxa.labels" null model (n = 999 randomizations), which randomizes taxon labels on the phylogeny for the species included in the sampling pool. We tested three different spatial extents for the sampling pools: global (all species in the phylogeny), hemispheric (Old World vs. New World species), and continental/biogeographic realm (South America, continental Africa, Indomalaya, and Australasia). Note that we used NRI and not the nearest taxon index (NTI) (18) because NTI represents mainly recent clustering or overdispersion (near the tips of the phylogeny) and therefore is less relevant for assessing deep-time hypotheses.

Biome Reconstructions and Time-Integrated Area. We used paleogeographic vegetation and biome reconstructions to estimate the area and distribution of tropical rainforests during the Cenozoic (*SI Appendix, SI Text S2* and Table S7). We obtained biome estimates for the Eocene (~ 55 Mya), Oligocene (~30 Mya), Miocene (~ 11 Mya), Middle Pliocene (~ 3 Mya), Last Glacial Maximum (LGM) (0.021 Mya), and the present (*SI Appendix*, Table S8). We digitized the distribution of rainforest for all time steps and plotted rainforest area against geological time [area-time plots *sensu* (12)]. We then estimated the AUC (12), the rate of area loss (measured as the slope β of a simple linear regression of area vs. time), and the minimum area during the Cenozoic as summary statistics for each biogeographic realm (*SI Appendix*, Table S8).

Quaternary Climate Change. We compiled two paleoclimatic reconstructions for the LGM (0.021 Mya), namely from the Community Climate System Model version 3 (CCSM3) and the Model for Interdisciplinary Research on Climate version 3.2 (MIROC3.2) (available at http://pmip2.lsce.ipsl.fr/) (59). We used the mean of the anomaly between LGM and present-day annual temperature across the two paleoclimatic simulations (CCSM3, MIROC3.2) to represent Quaternary climatic oscillations (37). These temperature anomalies cover almost the full Quaternary (past 2.6 Mya) temperature range with a geographic pattern that is representative for at least a large portion of the period (36). We used nonspatial as well as spatial regression models (60) to relate NRI to Quaternary climate change (*SI Appendix*, Tables S4 and S5).

EVOLUTION

ACKNOWLEDGMENTS. We thank Ulrich Salzmann and Sandy Harrison for information on biome reconstructions; Paul Fine for discussion of the time-integrated species-area effect; Peder K. Bøcher for assistance with geographic information system analysis; and John Dransfield, Rafaël Govaerts, and the Royal Botanic Gardens for providing the World Checklist of Palms. We further acknowledge the international modeling groups for providing the LGM climate data and the Laboratoire des Sciences du Climat et de l'Environnement for collecting and archiving them. The data archive of the Paleoclimate Modeling Intercomparison Project (PMIP2) is supported by Commissariat à l'Energie Atomique, Centre National de la Recherche Scientifique, the European Union project MOTIF (EVK2-CT-2002-00153), and the Programme National d'Etude de la Dynamique du Climat. This work was supported by Villum Kahn Rasmussen Foundation Grant VKR09b-141 (to J.-C.S.), Danish Council for Independent Research–Natural Sciences Grant 10-083348 (to H.B.) and starting independent researcher Grant 11-106163 (to W.D.K.), European Community Seventh Framework Programme Grant 212631 (to H.B.), and the Aarhus University Research Foundation (postdoctoral stipend to W.L.E.).

1. Hill JL, Hill RA (2001) Why are tropical rain forests so species rich? Classifying, reviewing and evaluating theories. *Prog Phys Geogr* 25:326–354.
2. Mittelbach GG, et al. (2007) Evolution and the latitudinal diversity gradient: Speciation, extinction and biogeography. *Ecol Lett* 10:315–331.
3. Hawkins BA, et al. (2003) Energy, water, and broad-scale geographic patterns of species richness. *Ecology* 84:3105–3117.
4. Allen AP, Brown JH, Gillooly JF (2002) Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. *Science* 297:1545–1548.
5. Currie DJ, et al. (2004) Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecol Lett* 7:1121–1134.
6. Dynesius M, Jansson R (2000) Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proc Natl Acad Sci USA* 97:9115–9120.
7. Wiens JJ (2011) The causes of species richness patterns across space, time, and clades and the role of "ecological limits". *Q Rev Biol* 86:75–96.
8. Sandel B, et al. (2011) The influence of Late Quaternary climate-change velocity on species endemism. *Science* 334:660–664.
9. Zobel M (1997) The relative role of species pools in determining plant species richness: An alternative explanation of species coexistence? *Trends Ecol Evol* 12:266–269.
10. Ricklefs RE (2005) Phylogenetic perspectives on patterns of regional and local species richness. *Tropical Rainforests: Past, Present and Future*, eds Bermingham E, Dick CW, Moritz C (University of Chicago Press, Chicago, London), pp 16–40.
11. Stephens PR, Wiens JJ (2003) Explaining species richness from continents to communities: The time-for-speciation effect in emydid turtles. *Am Nat* 161:112–128.
12. Fine PVA, Ree RH (2006) Evidence for a time-integrated species-area effect on the latitudinal gradient in tree diversity. *Am Nat* 168:796–804.
13. Svenning JC, Borchsenius F, Bjorholm S, Balslev H (2008) High tropical net diversification drives the New World latitudinal gradient in palm (Arecaceae) species richness. *J Biogeogr* 35:394–406.
14. Renner S (2004) Plant dispersal across the tropical Atlantic by wind and sea currents. *Int J Plant Sci* 165:S23–S33.
15. Webb SD (2006) The Great American Biotic Interchange: Patterns and processes. *Ann Mo Bot Gard* 93:245–257.
16. Paul JR, Morton C, Taylor CM, Tonsor SJ (2009) Evolutionary time for dispersal limits the extent but not the occupancy of species' potential ranges in the tropical plant genus *Psychotria* (Rubiaceae). *Am Nat* 173:188–199.
17. Wiens JJ, Donoghue MJ (2004) Historical biogeography, ecology and species richness. *Trends Ecol Evol* 19:639–644.
18. Webb CO, Ackerly DD, McPeek MA, Donoghue MJ (2002) Phylogenies and community ecology. *Annu Rev Ecol Syst* 33:475–505.
19. Emerson BC, Gillespie RG (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends Ecol Evol* 23:619–630.
20. Cavender-Bares J, Kozak KH, Fine PVA, Kembel SW (2009) The merging of community ecology and phylogenetic biology. *Ecol Lett* 12:693–715.
21. Cooper N, Rodríguez J, Purvis A (2008) A common tendency for phylogenetic overdispersion in mammalian assemblages. *Proc Biol Sci* 275:2031–2037.
22. Kamilar JM, Guidi LM (2010) The phylogenetic structure of primate communities: Variation within and across continents. *J Biogeogr* 37:801–813.
23. Swenson NG, Enquist BJ, Pither J, Thompson J, Zimmerman JK (2006) The problem and promise of scale dependency in community phylogenetics. *Ecology* 87:2418–2424.
24. Webb CO (2000) Exploring the phylogenetic structure of ecological communities: An example for rain forest trees. *Am Nat* 156:145–155.
25. Pennington RT, Richardson JE, Lavin M (2006) Insights into the historical construction of species-rich biomes from dated plant phylogenies, neutral ecological theory and phylogenetic community structure. *New Phytol* 172:605–616.
26. Vamosi SM, Heard SB, Vamosi JC, Webb CO (2009) Emerging patterns in the comparative analysis of phylogenetic community structure. *Mol Ecol* 18:572–592.
27. Cardillo M (2011) Phylogenetic structure of mammal assemblages at large geographical scales: Linking phylogenetic community ecology with macroecology. *Philos Trans R Soc Lond B Biol Sci* 366:2545–2553.
28. Crisp MD, Trewick SA, Cook LG (2011) Hypothesis testing in biogeography. *Trends Ecol Evol* 26:66–72.
29. Lomolino MV, Riddle BR, Whittaker RJ, Brown JH (2010) *Biogeography* (Sinauer Associates, Sunderland, MA).
30. Rosenzweig ML (1995) *Species Diversity in Space and Time* (Cambridge University Press, Cambridge, UK).
31. Morley RJ (2000) *Origin and Evolution of Tropical Rain Forests* (John Wiley & Sons, Chichester, UK).
32. Morley RJ (2007) Cretaceous and Tertiary climate change and the past distribution of megathermal rainforests. *Tropical Rainforest Responses to Climate Change*, eds Bush MB, Flenley JR (Springer, Berlin), pp 1–31.
33. Hortal J, et al. (2011) Ice age climate, evolutionary constraints and diversity patterns of European dung beetles. *Ecol Lett* 14:741–748.
34. Jansson R, Dynesius M (2002) The fate of clades in a world of recurrent climatic change: Milankovitch oscillations and evolution. *Annu Rev Ecol Syst* 33:741–777.
35. Davies TJ, Buckley LB (2011) Phylogenetic diversity as a window into the evolutionary and biogeographic histories of present-day richness gradients for mammals. *Philos Trans R Soc Lond B Biol Sci* 366:2414–2425.
36. Jansson R (2003) Global patterns in endemism explained by past climatic change. *Proc Biol Sci* 270:583–590.
37. Kissling WD, et al. (2012) Quaternary and pre-Quaternary historical legacies in the global distribution of a major tropical plant lineage. *Glob Ecol Biogeogr*, 10.1111/j.1466-8238.2011.00728.x.
38. Dransfield J, et al. (2008) *Genera Palmarum: The Evolution and Classification of Palms* (Royal Botanical Gardens, Kew, UK).
39. Eiserhardt WL, Svenning J-C, Kissling WD, Balslev H (2011) Geographical ecology of the palms (Arecaceae): Determinants of diversity and distributions across spatial scales. *Ann Bot* 108:1391–1416.
40. Couvreur TLP, Forest F, Baker WJ (2011) Origin and global diversification patterns of tropical rain forests: Inferences from a complete genus-level phylogeny of palms. *BMC Biol* 9:44.
41. Govaerts R, Dransfield J (2005) *World Checklist of Palms* (Royal Botanic Gardens, Kew, UK).
42. Baker WJ, et al. (2009) Complete generic-level phylogenetic analyses of palms (Arecaceae) with comparisons of supertree and supermatrix approaches. *Syst Biol* 58:240–256.
43. Bjorholm S, Svenning J-C, Baker WJ, Skov F, Balslev H (2006) Historical legacies in the geographical diversity patterns of New World palm (Arecaceae) subfamilies. *Bot J Linn Soc* 151:113–125.
44. Givnish TJ (2010) Ecology of plant speciation. *Taxon* 59:1326–1366.
45. Dransfield J, Beentje H (1995) *The Palms of Madagascar* (Royal Botanic Gardens and International Palm Society, Kew, UK).
46. Hodel DR (2007) A review of the genus *Pritchardia*. *Palms* 51:S1–S53.
47. Pintaud J-C, Baker WJ (2008) A revision of the palm genera (Arecaceae) of New Caledonia. *Kew Bull* 63:61–73.
48. Willis KJ, McElwain JC (2002) *The Evolution of Plants* (Oxford University Press, Oxford).
49. Greenwood DR, Wing SL (1995) Eocene continental climates and latitudinal temperature gradients. *Geology* 23:1044–1048.
50. Svenning J-C, Borchsenius F, Bjorholm S, Balslev H (2008) High tropical net diversification drives the New World latitudinal gradient in palm (Arecaceae) species richness. *J Biogeogr* 35:394–406.
51. Pan AD, Jacobs BF, Dransfield J, Baker WJ (2006) The fossil history of palms (Arecaceae) in Africa and new records from the Late Oligocene (28-27 Mya) of northwestern Ethiopia. *Bot J Linn Soc* 151:69–81.
52. Rull V (2011) Neotropical biodiversity: Timing and potential drivers. *Trends Ecol Evol* 26:508–513.
53. Hoorn C, et al. (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science* 330:927–931.
54. Cronk QCB (1997) Islands: Stability, diversity, conservation. *Biodivers Conserv* 6:477–493.
55. Cannon CH, Morley RJ, Bush ABG (2009) The current refugial rainforests of Sundaland are unrepresentative of their biogeographic past and highly vulnerable to disturbance. *Proc Natl Acad Sci USA* 106:11188–11193.
56. Brummitt RK (2001) *World Geographical Scheme for Recording Plant Distributions* (Hunt Institute for Botanical Documentation Carnegie Mellon University, Pittsburgh), 2nd Ed.
57. Kembel SW, et al. (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26:1463–1464.
58. R Development Core Team (2010) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna).
59. Braconnot P, et al. (2007) Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum—Part 1: Experiments and large-scale features. *Climate of the Past* 3:261–277.
60. Kissling WD, Carl G (2008) Spatial autocorrelation and the selection of simultaneous autoregressive models. *Glob Ecol Biogeogr* 17:59–71.

Supporting Information

Kissling et al. (2012): Cenozoic imprints on the phylogenetic structure of palm species assemblages worldwide.

Table S1: Statistical results for testing whether the mean net relatedness index (NRI, response variable) for a given spatial extent (global, South America, Africa, Indomalaya, Australasia) and sampling pool (global, hemispheric, realm) differs significantly from zero. Significant results (i.e., non-random phylogenetic structure) are highlighted in **bold**. An intercept-only linear ordinary least squares (OLS) regression model was used to test whether the intercept significantly departs from zero (equivalent to a non-spatial one sample *t*-test). In the presence of spatial autocorrelation in OLS model residuals, a spatial model of the simultaneous autoregressive error type (SAR) was used instead (equivalent to a 'spatial' one sample *t*-test). SAR models were fitted with a row standardized coding scheme and a neighborhood based on a Gabriel connection (global spatial extent) or a minimum distance to connect each sample unit to at least one other site (South America, Africa, Indomalaya, Australasia). Moran's *I* values were used to assess residual spatial autocorrelation in OLS and SAR models based on the four nearest neighbors of each site. Significance of Moran's *I* statistics was determined from permutation tests (n = 1000 permutations). All Moran's *I* values of reported OLS and SAR models were not significant ($P > 0.05$).

Model				Results		
Spatial extent	Sample size	Sampling pool	Model type	Intercept (±SE)	*z* or *t*	*P*-value
Global	**152**	**global**	**SAR**	**3.842 (±0.752)**	**5.113**	**<0.001**
South America	**17**	**global**	**SAR**	**11.946 (±3.189)**	**3.746**	**<0.001**
South America	**17**	**hemispheric**	**SAR**	**6.806 (±2.152)**	**3.163**	**0.002**
South America	17	realm	SAR	1.164 (±1.415)	0.822	0.411
Africa	43	global	SAR	0.439 (±0.315)	1.393	0.164
Africa	43	hemispheric	SAR	0.038 (±0.283)	0.134	0.893
Africa	43	realm	SAR	0.475 (±0.320)	1.485	0.138
Indomalaya	**25**	**global**	**OLS**	**2.362 (±0.448)**	**5.266**	**<0.001**
Indomalaya	**25**	**hemispheric**	**OLS**	**1.066 (±0.349)**	**3.058**	**0.005**
Indomalaya	25	realm	OLS	0.375 (±0.280)	1.344	0.192
Australasia	**14**	**global**	**OLS**	**4.822 (±1.604)**	**3.007**	**0.010**
Australasia	**14**	**hemispheric**	**OLS**	**3.666 (±1.377)**	**2.663**	**0.020**
Australasia	14	realm	OLS	0.930 (±0.913)	1.019	0.327

Table S2: Insular geographic units (TDWG level 3 units) which show high phylogenetic clustering, i.e., the top 10 highest positive values of the net relatedness index (NRI) in palm assemblages on islands worldwide. The species pool of the null assemblages to calculate the NRI values is global (compare Fig. 1A in main text).

Island name	TDWG unit	NRI	Rank	Phylogenetic structure and composition	References
Borneo	BOR	6.26	9	Several large palm genera are represented by many species (including *Calamus*, *Daemonorops*, *Iguanura*, *Areca*, *Pinanga*, *Licuala*) and some genera are closely related within higher groups (e.g., tribes Calameae, Areceae, Trachycarpeae)	(1, 2)
Cuba	CUB	18.13	4	Palm flora dominated by major radiations in *Coccothrinax* and *Copernicia*, as well as the smaller genus *Roystonea*. Some higher lineages generally well represented (e.g., tribe Cryosophileae)	(1)
Fiji	FIJ	11.61	5	The majority of genera fall within one tribe (Areceae), with the two most diverse genera (*Balaka* and *Veitchia*) being closely related	(1)
Hawaii	HAW	20.65	3	One genus that has radiated extensively (*Pritchardia*; 22 of the 27 recognized species are Hawaiian endemics)	(1)
Madagascar	MDG	40.9	1	Spectacular radiation of the genus *Dypsis* (>140 *Dypsis* species accounting for 80% of all species). Remaining species also include the genus *Ravenea*, 16 of which are endemic to Madagascar.	(1, 3, 4)
New Caledonia	NWC	22.03	2	Apart from one species, all taxa in New Caledonia are restricted to three subtribes (Archontophoenicinae, Basseliniinae, Clinospermatinae) within the Pacific clade of tribe Areceae.	(1, 5)
New Guinea	NWG	10.78	6	Majority of genera of New Guinea palms fall within subtribe Areceae, including some significant radiations in *Heterospathe*, *Hydriastele*, *Calyptrocalyx*, and subtribe Ptychospermatinae. Important radiations also in *Calamus*, *Licuala* and *Saribus*.	(1, 2)
Lord Howe-Norfolk Is.	NFK	5.52	10	All species within Areceae, including both genera of Rhopalostylidinae (*Hedyscepe* and *Rhopalostylis*) and both species of *Howea*.	(1, 6)
Solomon Islands	SOL	8.46	7	Palm flora dominated by genera from tribe Areceae (76%)	(1, 2)
Sumatra	SUM	7.22	8	Similar to Borneo, though species diversity is lower in all groups (much lower in some cases, e.g., *Iguanura* and *Licuala*)	(1, 2)

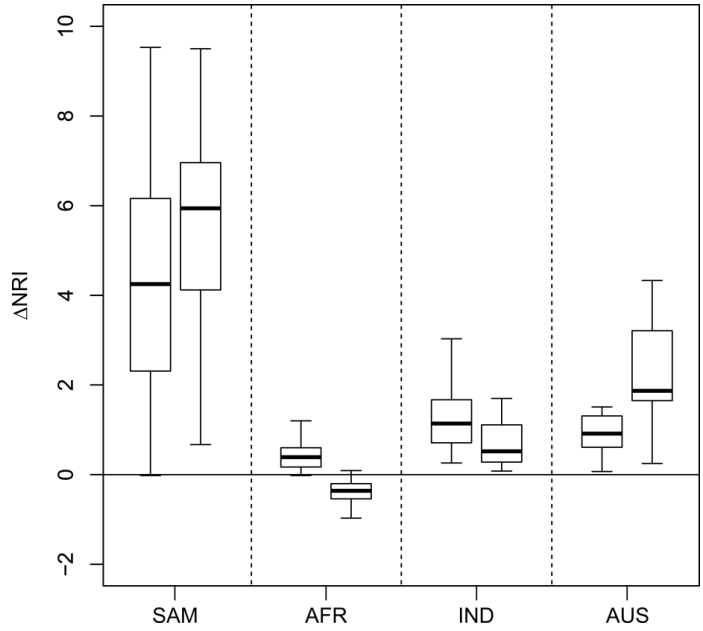

Fig. S1: Effects of sampling pool scaling on NRI within biogeographic realms (SAM = South America; AFR = Africa; IND = Indomalaya; AUS = Australasia). Within each panel, the left box plot shows the change in NRI between global and hemispheric sampling pools (i.e., $NRI_{global} - NRI_{hemispheric}$). The right box plot shows the change in NRI between hemispheric and realm sampling pools ($NRI_{hemispheric} - NRI_{realm}$). Values > 0 indicate stronger phylogenetic clustering with the geographically more extensive sampling pool. All distributions are significantly $\neq 0$ ($P \leq 0.001$) (see Table S3).

Table S3: Statistical results for testing whether changes in the mean net relatedness index (NRI) in relation to sampling pools are significantly different from zero. For each assemblage within a realm, we computed $NRI_{global\ pool} - NRI_{hemispheric\ pool}$ and $NRI_{hemispheric\ pool} - NRI_{realm\ pool}$ and used an intercept-only linear ordinary least squares (OLS) regression model to test whether the intercept significantly departs from zero (equivalent to a paired two sample *t*-test). In the presence of spatial autocorrelation in OLS model residuals, a spatial model of the simultaneous autoregressive error type (SAR) was used instead (equivalent to a 'spatial' paired two sample *t*-test). SAR model fitting and Moran's *I* calculations as in Table S1. Note that all changes in NRI values in relation to sampling pool scaling are significant.

Model				Results		
Spatial extent	Sample size	Sampling pool change	Model type	Intercept (±SE)	*z* or *t*	*P*-value
South America	17	global -> hemispheric	SAR	4.931 (±1.290)	3.824	<0.001
South America	17	hemispheric -> realm	SAR	5.797 (±1.440)	4.027	<0.001
Africa	43	global -> hemispheric	SAR	0.392 (±0.075)	5.200	<0.001
Africa	43	hemispheric -> realm	SAR	-0.419 (±0.111)	-3.786	<0.001
Indomalaya	25	global -> hemispheric	SAR	1.299 (±0.262)	4.952	<0.001
Indomalaya	25	hemispheric -> realm	SAR	0.691 (±0.121)	5.728	<0.001
Australasia	14	global -> hemispheric	OLS	1.156 (±0.266)	4.349	<0.001
Australasia	14	hemispheric -> realm	OLS	2.735 (±0.675)	4.053	0.001

Quaternary climate change

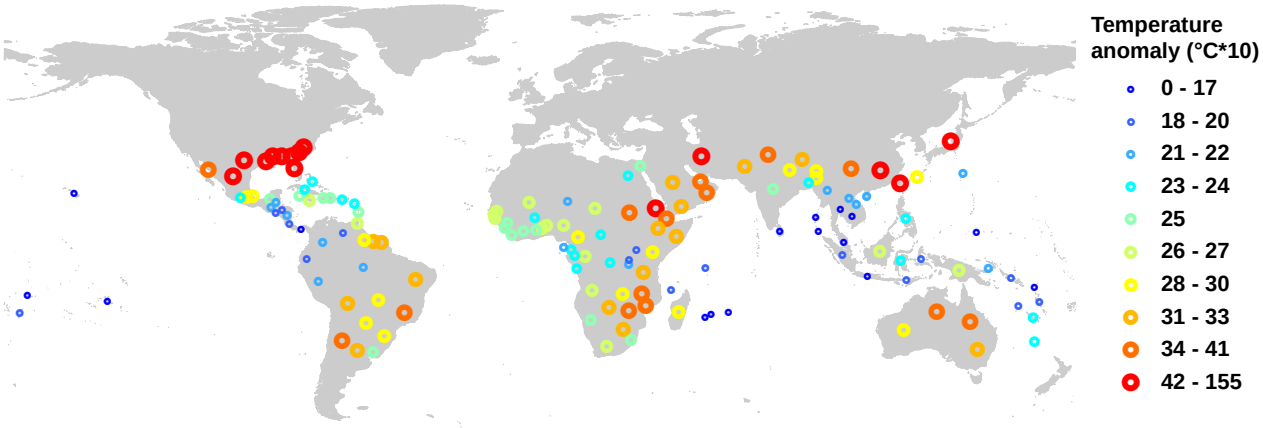

Fig. S2: Mean temperature changes (anomalies) between the Last Glacial Maximum (LGM, ca. 0.021 mya) and the present plotted for the centroid of each 'botanical country' that has more than one palm species present. Data are ensemble (means) of the temperature anomalies of two paleoclimatic simulations (CCSM3 and MIROC3.2; see methods). Classification uses quantiles and maps are in Behrmann projection.

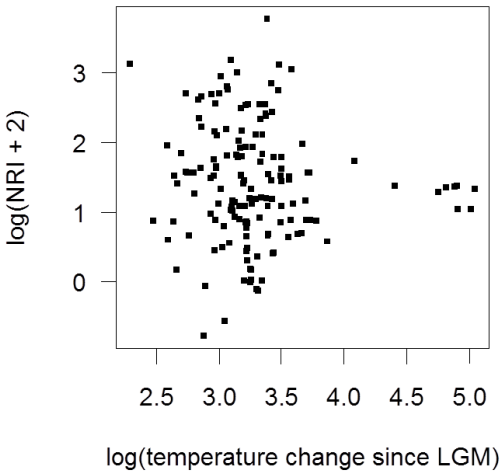

Fig. S3: Global relationship between Quaternary climatic oscillations and the net relatedness index (NRI) in palms (calculated with a global sampling pool). Quaternary climatic oscillations were quantified as the change in mean annual temperature (anomaly, in °C*10) between the Last Glacial Maximum (~0.021 mya) and the present (compare Fig. S2 and methods). The relationship between NRI and Quaternary temperature change is not statistically significant at a global scale (see Table S4 for statistical results).

Table S4: Single-predictor models of the effect of Quaternary temperature change (predictor) on the net relatedness index (NRI; response) in palms globally and within four biogeographic realms (South America, Africa, Indomalaya, and Australasia). The relationship is only significant within South America and Africa (highlighted in bold). Quaternary temperature change was quantified as the anomaly in mean annual temperature between the Last Glacial Maximum (LGM, ~0.021 mya) and the present. For the calculation of NRI values, the spatial extent of sampling pools was global for the global analysis and restricted to a continental/biogeographic scale for the realm analyses. Results from ordinary least square regression models (OLS) and spatial autoregressive error models (SAR) are given. SAR models were fitted to account for spatial autocorrelation in model residuals using a spatial weights matrix with a row standardized coding scheme (7) and a Gabriel connection (global) or a minimum distance to connect each sample unit to at least one other site (realm). The explained variance (R^2) of the SAR model includes only the environmental component to make it comparable to OLS (i.e., it excludes the explanatory power of the spatial weights matrix). Moran's *I* values quantify residual spatial autocorrelation based on the four nearest neighbors of each site. Significance of Moran's *I* statistics was determined from permutation tests (n = 1000 permutations). Sample sizes (n geographic units) are provided after the region names.

Model statistics	Global (n = 151)		South America (n = 17)		Africa (n = 43)		Indomalaya (n = 25)		Australasia (n = 14)	
	OLS	SAR	OLS	SAR	OLS	SAR	OLS	SAR	OLS	SAR
Intercept	1.891	1.524	-7.873	-4.224	-2.900	-3.040	-0.614	-0.618	0.636	-0.912
Coefficient	-0.142	-0.045	0.338	0.201	1.102	1.142	0.422	0.423	0.150	0.642
Std. Error	0.145	0.176	0.100	0.101	0.498	0.519	0.333	0.318	0.778	0.870
t or *z* value	-0.976	-0.237	3.383	1.991	2.213	2.199	1.270	1.330	0.193	0.738
P-value	0.330	0.813	**0.004**	**0.046**	**0.033**	**0.028**	0.217	0.183	0.850	0.460
R^2	0.006	0.006	0.433	0.433	0.107	0.107	0.065	0.065	0.003	0.003
Moran's *I*	0.503	0.034	0.187	0.135	0.181	0.049	-0.049	-0.049	0.093	0.041
P of Moran's *I*	<0.001	0.205	0.051	0.077	0.017	0.214	0.490	0.494	0.133	0.175

For South America, the frequency distribution of model residuals approximated a normal distribution, so data could be analyzed with untransformed variables. For the other realms and the global analysis, the response variable (NRI+2) as well as the predictor variable (Quaternary temperature change) was log-transformed.

Table S5: Coefficients from the most parsimonious multiple-predictor model to explain the net relatedness index (NRI) in palms within South America and Africa in relation to Quaternary temperature anomaly (ANOM TEMP) and present-day environment (ELEV, PREC, PREC SEAS, TEMP, TEMP SEAS). A full model with all predictor variables was subject to a step-wise model selection based on the Akaike information criterion (AIC). '–', not selected by step-wise model selection. Moran's *I* statistics indicate no spatial autocorrelation in model residuals.

	South America (*n* = 17)		Africa (*n* = 43)	
	Coefficient	*P*	Coefficient	*P*
Intercept	20.258	*	-4.571	*
ANOM TEMP	0.234	*	1.379	*
ELEV	-0.001	***	–	
PREC	-0.003	*	–	
PREC SEAS	–		-0.007	n.s.
TEMP	-0.057	*	0.006	n.s.
TEMP SEAS	-0.002	**	–	
R^2	0.840		0.191	
Moran's *I*	-0.178		0.080	
P of Moran's *I*	n.s.		n.s.	

ANOM TEMP, anomaly in temperature between Last Glacial Maximum and present (in °C * 10); ELEV, elevation range (in m); PREC, annual precipitation (in mm $year^{-1}$); PREC SEAS, precipitation seasonality measured as coefficient of variation of monthly values (in mm); TEMP, annual mean temperature (in °C * 10); TEMP SEAS, temperature seasonality measured as standard deviation of monthly means (in °C * 10).
Significance levels: ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$; n.s., not significant.
NRI calculation: NRI values were calculated with a realm sampling pool.
Data transformations: Data for South America were analyzed with untransformed variables as model residuals approximated a normal distribution. For Africa, the response variable (NRI + 2) as well as the predictor variable (ANOM TEMP) was log-transformed.
Spatial autocorrelation: Moran's *I* values quantify residual spatial autocorrelation based on the four nearest neighbors of each site. Significance of Moran's *I* statistics was determined from permutation tests (*n* = 1000 permutations).

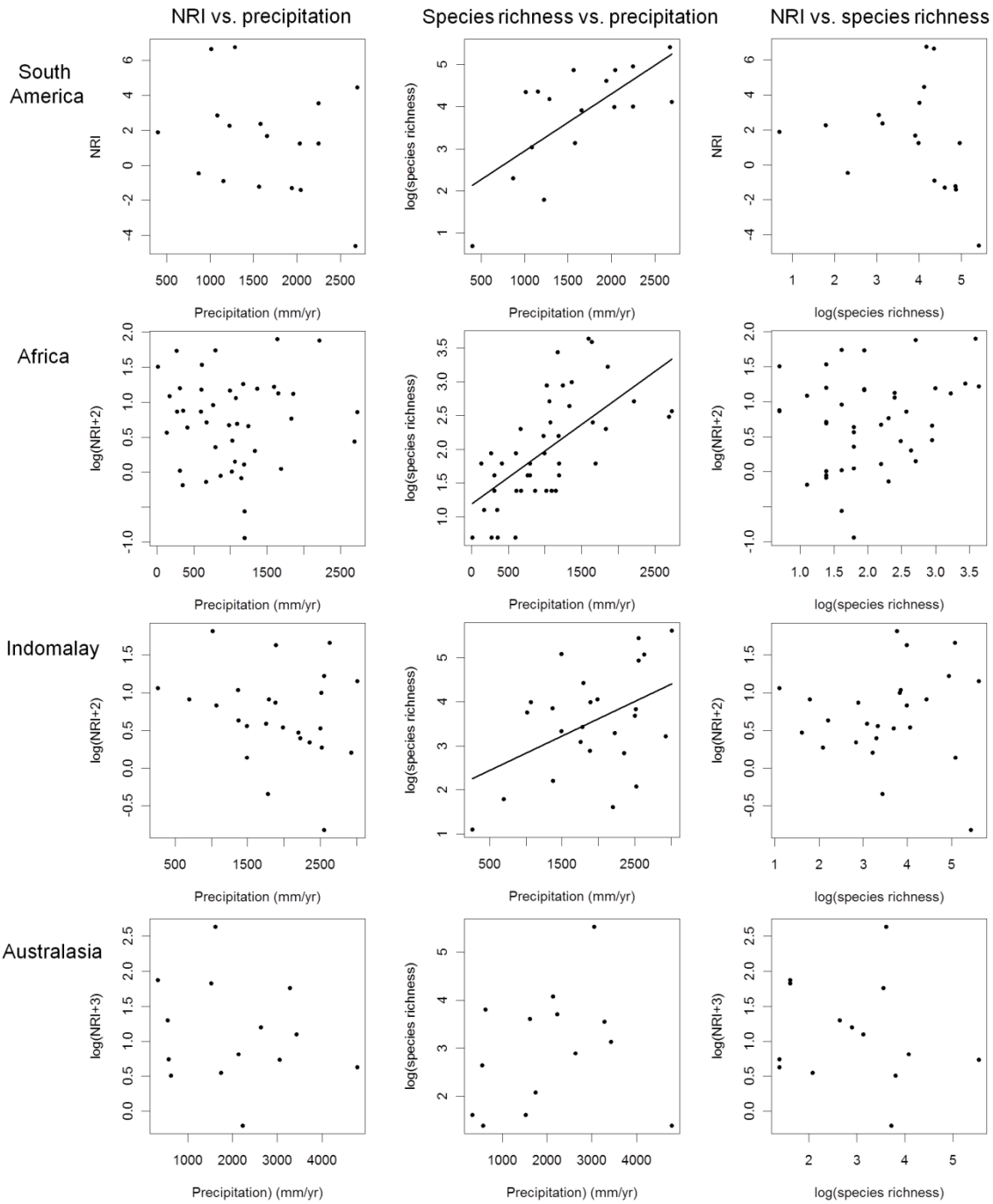

Fig. S4: The relationships between the net relatedness index (NRI), precipitation, and species richness within four biogeographic realms (South America, Africa, Indomalaya, and Australasia). For the calculation of NRI values, the spatial extent of sampling pools was restricted to a continental/biogeographic scale. Note that the commonly observed increase of species richness with precipitation is not reflected in the NRI.

Table S6: Endemism in palms at different taxonomic levels and various biogeographic scales. Data are based on the World Checklist of Palms (downloaded on 9th March 2009). Biogeographic scales (global, New World vs. Old World, and within four realms) correspond to Fig. 1 in the main text. Taxonomic levels distinguish species, genera and tribes. Note the extremely high level of endemism at the species level, but also the high degree of endemism at the tribe level in the New World and the Old World.

	Species			Genus			Tribe		
Biogeographic scale	Endemic	Total	Percent	Endemic	Total	Percent	Endemic	Total	Percent
Global	2440	2440	100.0	183	183	100.0	29	29	100
New World	782	784	99.7	65	69	94.2	11	17	64.7
Old World	1656	1658	99.8	114	118	96.6	12	18	66.7
South America	371	451	82.2	17	51	33.3	1	15	6.7
Africa	50	65	76.9	5	16	31.3	1	9	11.1
Indomalay	872	897	97.2	23	46	50.0	1	12	8.3
Australasia	451	472	95.6	28	56	50.0	0	11	0.0

Text S1: Dispersal and phylogenetic assemblage structure in palms

Most palm species have limited capacities for long-distance dispersal via oceanic drift because floating seeds, though well known, are the exception (notably *Cocos nucifera* and *Nypa fruticans*) (1). Instead, palm seeds are generally heavy, dense, and most commonly sink. Palm fruits are mostly dispersed by small- to large-bodied frugivorous birds and mammals (8, 9) and fruit sizes typically range between 1–3 cm (Fig. S5). Nevertheless, a few palm genera (e.g., *Lodoicea*, *Cocos* and *Borassus*) have large fruits (e.g., >10 cm). As a consequence of seed dispersal by birds and mammals, palm fruits are mostly dispersed at short to medium distances (<100 km; (10)) while seed dispersal among continents and realms is apparently very rare. The limited long-distance dispersal capacity of most palms is reflected in a high degree of endemism (Table S6) and in the fact that only few palm genera (*Elaeis* and *Raphia*) are present at both sides of the Atlantic (in Africa and South America). Of the 2,440 palm species analyzed in this study a total of 1,648 (68%) occur in only one of the level 3 geographic units ("botanical countries") of the World Geographical Scheme for Recording Plant Distributions. Moreover, there is not only an extremely high degree of endemism at the species level in all realms, but also a high degree of endemism at the tribe level in the New World and the Old World (Table S6). Overall, this limited transoceanic dispersal capacity (together with *in situ* diversification) results in pronounced phylogenetic clustering of regional palm assemblages in South America, Indomalaya, and Australasia when using a global (or New World vs. Old World) species pool for NRI calculations (Fig. 1 and Fig. 2 in main text).

Beyond cross-continental dispersal limitation, phylogenetic clustering in palms might also be influenced by dispersal processes within realms. Notably, strong dispersal limitation (for both seeds and pollen) is likely to increase speciation because intraspecific gene flow at the population level is low (11). Two basic traits in particular may influence dispersal limitation in animal-dispersed plants, namely stem height and fruit (seed) size (12-14). If stem height of fleshy-fruited plants is small, such as in understory palms, the sedentary nature of most frugivorous birds in the understory of tropical rainforests has been proposed to lead to strong dispersal limitation at the local or landscape scale (11). Empirical evidence supports this idea for palms, as understory palms show a stronger decay with environmental and geographical distance than canopy palms (15), suggesting that the former might be more dispersal limited than the latter. The high species diversity of some palm genera in the understory of tropical rainforests (e.g., *Bactris* (16), *Chamaedorea* (17), and *Geonoma* (18) in South America) and its spatial congruence with high phylogenetic clustering is consistent with this hypothesis. Besides stem height, fruit size may also influence dispersal limitation in palms, e.g., by shaping dispersal mode. Fleshy fruits of ≥4 cm are often predominantly dispersed by non-flying mammals, including scatter-hoarding rodents, primates, elephants, and tapirs (8, 19, 20). However, non-flying mammals are less effective than flying animal dispersers (birds and bats) for seed dispersal across major barriers (rivers, mountain ranges, or vast stretches of sea). This has been demonstrated empirically by studying the colonization of volcanic islands (e.g., Krakatau or Long Island) through palms (21) and other fleshy-fruited plants (22, 23). Hence, the presence, abundance and predominance of flying vs. non-flying dispersers could have important consequences for allopatric speciation and phylogenetic clustering within realms.

Comparing fruit sizes and stem heights of palms across realms (Fig. S5) suggests that Africa generally shows a paucity of small-fruited palms, both in short-statured species (stem heights <15 m) as well as tall-statured species. Instead, the African palm flora is dominated by large-fruited palms (Fig. S5) which are often dispersed by non-flying mammals (8). The paucity of small-fruited palms in Africa (including both short-stemmed and tall-growing species) may be related to the

drying of this region during the Cenozoic (Fig. 3*A* in main text) because a larger seed mass enables seedlings to better survive hazards such as drought (24) and because tropical rainforest understory plants are especially sensitive to drought (25). Hence, the dramatic drying of Africa during the Cenozoic could have had a strong effect on dispersal trait composition of palm species assemblages in this region. However, it remains unclear to what extent seed dispersal of non-flying mammals affects speciation and phylogenetic structuring of fleshy-fruited plants, e.g., relative to seed dispersal by birds and bats or when considering seed dispersal services of extinct megafauna (26).

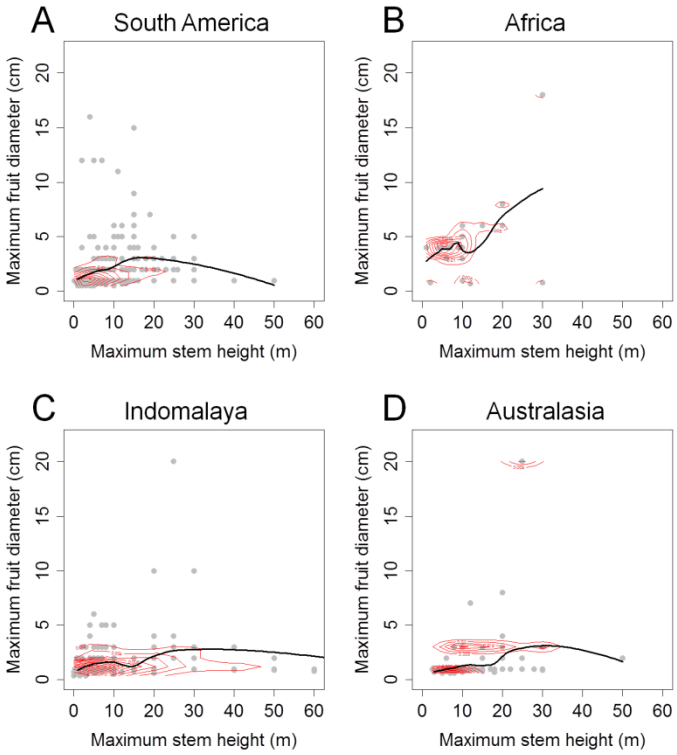

Fig. S5: Plots of fruit size vs. stem height for palm species in (*A*) South America (*n* = 295), (*B*) Africa (*n* = 25), (*C*) Indomalaya (*n* = 204), and (*D*) Australasia (*n* = 90). Note that Africa generally misses small-fruited palms (< 4cm fruit diameter), both short-stemmed as well as tall-stemmed species. Red contour lines indicate two-dimensional kernel densities as derived from the R function kde2d(). The black lines represent local polynomial regression fits as obtained from the R function loess(). Data were obtained from (9). This data source provides information on fruit size and stem height for >1,100 palm species (45%). Only those species were included which occur within the four realms and for which data on both fruit size and stem height were available.

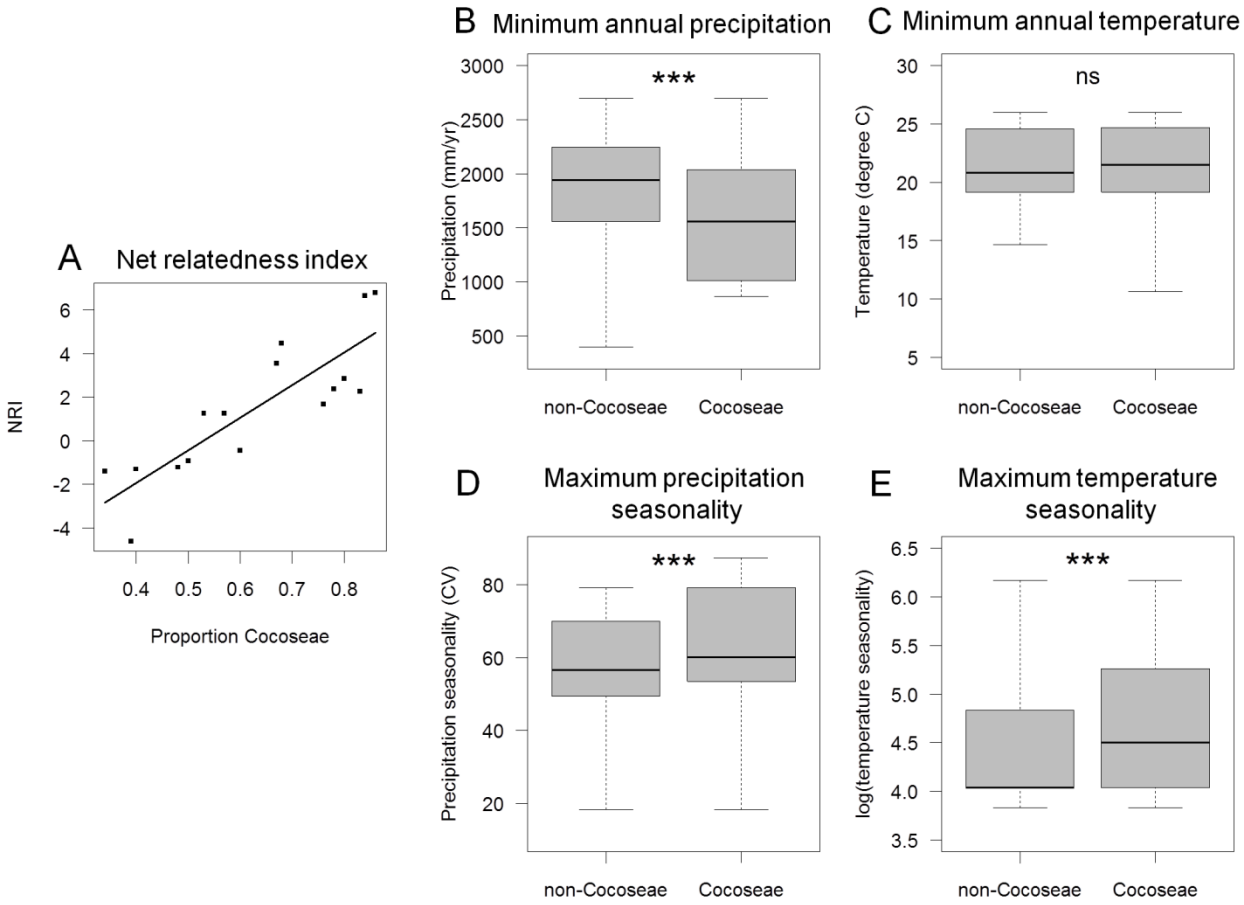

Figure S6: The Cocoseae effect in South America. (*A*) Relationship between the net relatedness index (NRI) and the proportion of species in the tribe Cocoseae within TDWG level 3 units (R^2 = 0.72, P<0.001), and (*B-E*) boxplots of species-climate sensitivities compared for species within the tribe Cocoseae vs. the rest. In (*B-E*), the minimum (*B,C*) or maximum (*D,E*) country-level climate mean value for each species within its geographic range (based on occurrences within level 3 geographic units of the World Geographical Scheme for Recording Plant Distributions) was taken to indicate species-level climatic sensitivity. Seasonality values (*D,E*) were quantified as coefficient of variation (CV) of monthly precipitation values and standard deviation (SD) of monthly temperature values. T-tests were used to compare differences in mean climate sensitivities between Cocoseae and non-Cocoseae species. NRI is largely driven by the proportion of species within the tribe Cocoseae. Cocoseae species tend to occur in areas that are drier and have higher precipitation and temperature seasonalities than areas were non-Cocoseae species occur. The strong correlation between NRI and Quaternary temperature anomaly (Spearman rank: r = 0.67, see also Fig. 5*A* main text) becomes weak once the proportion of Cocoseae species has been accounted for (partial Spearman rank: r = 0.34).

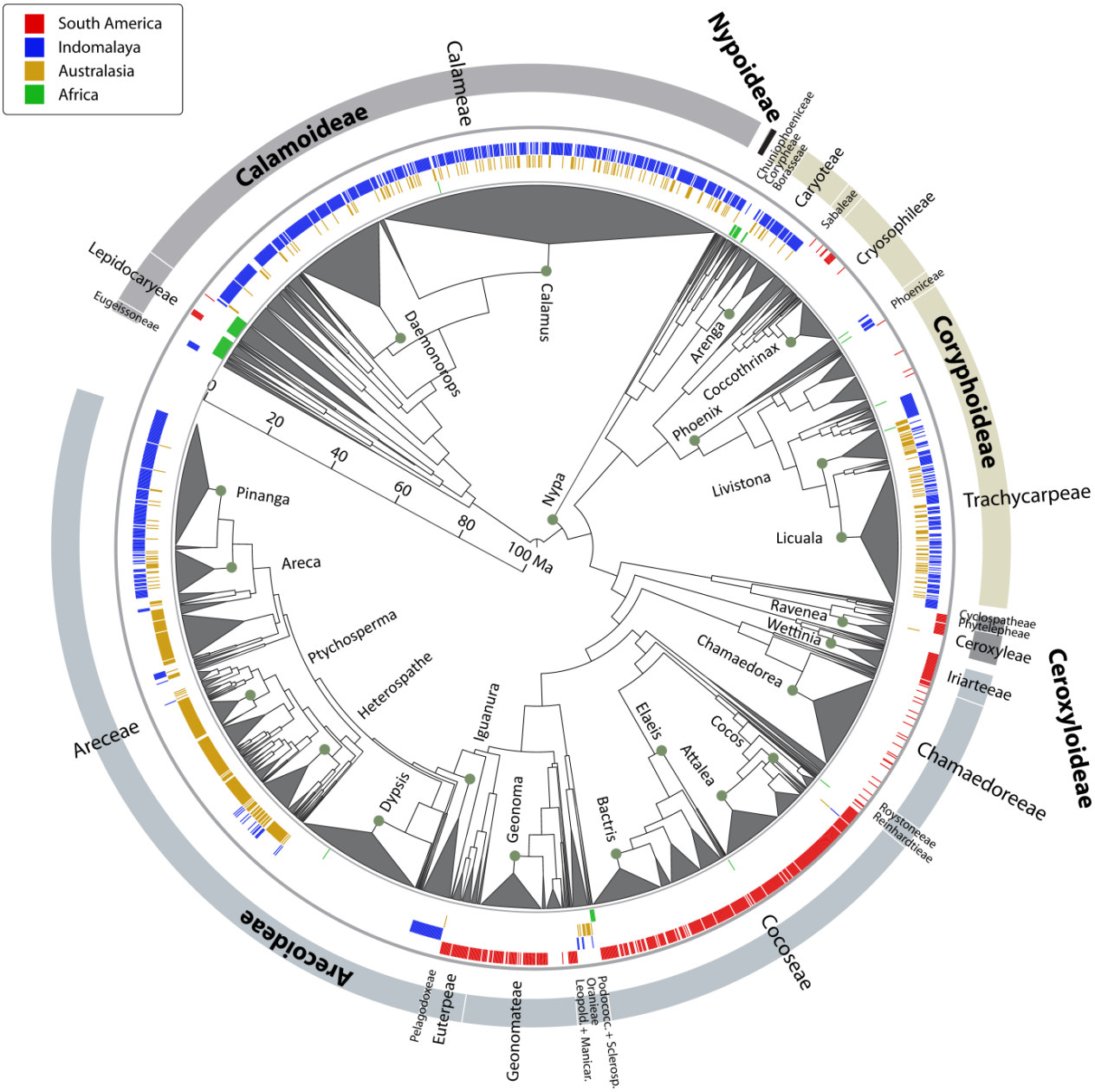

Fig. S7: The phylogenetic tree used in this study. The tree is based on a dated version (27) of a recent supertree of all palm genera, the most extensive phylogenetic study of the palm family published to date (28). Below the genus level, species were appended as polytomies with a divergence age arbitrarily set at two-thirds the stem node age of the genus. The width of the triangles corresponds to the number of species per genus. The time scale is in millions of years (Ma). The stem nodes of selected species-rich genera are indicated in the figure by a green dot adjacent to the generic name. Subfamilies (names in bold ending on -oideae) are separated by thick white lines in the outer gray ring. Tribes with 25 or more species (names ending in -eae) are delimited by thin white lines in the outer gray ring. Names of tribes are printed in a font size proportional to their size (smallest font 0–24 species; medium font 25–99 species; largest font > 100 species). Colors in the inner rings correspond to species occurring in the four biogeographic realms (South America, Indomalaya, Australasia, Africa) as used in this study (compare Fig. 1*C* in main text). Note that species within genera are ordered randomly. Species with no colour are found outside the four biogeographic realms in Fig. 1*C*, e.g., Madagascar (e.g., *Dypsis*), Central or North America, the Carribean (e.g., *Coccothrinax*), or the Pacific Islands.

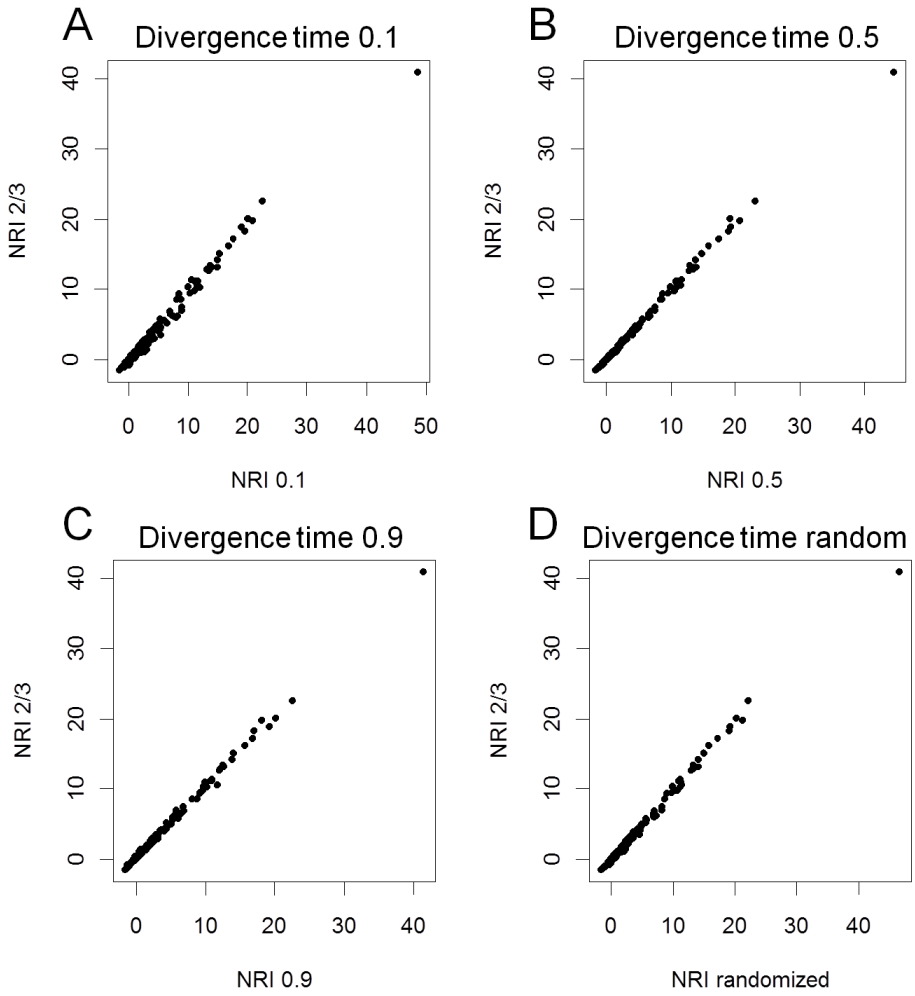

Fig. S8: Influence of the assumed divergence time among species within genera on the net relatedness index (NRI). In all cases, NRI was calculated on the same set of palm assemblages (n = 151 level 3 geographic units of the World Geographical Scheme for Recording Plant Distributions). The y-axis "NRI 2/3" represents NRI values calculated with a divergence time arbitrarily set at two-thirds of the stem node age of the genus (used throughout the paper). These values are compared to NRI values where divergence times within genera were assigned at (*A*) 0.1, (*B*) 0.5, and (*C*) 0.9 times the stem node of the genus, and (*D*) a random fraction between 0 and 2/3 of the stem node age of the genus. Note that intra-generic relationships are unresolved and thus a single divergence time per genus is used. For the calculation of (*D*), NRI calculation was repeated and averaged for 100 phylogenetic trees with randomly assigned divergence times. Note that NRI values from all these different approaches are highly correlated (all Pearson correlations $r > 0.99$). This sensitivity analysis shows that NRI calculations are robust to changing assumed divergence times within genera. This is expected because NRI values reflect deeper divergences in a phylogeny, not the shallow divergences of species within genera.

Text S2: Biome reconstructions and area-time plots

We used biome reconstructions of the tropical rainforest biome (Table S7) to estimate the available area of suitable habitat for palms since the Eocene (ca. 55 mya). We chose three reconstructions (29-31) for the deeper time periods (Eocene, Oligocene, Miocene) and used an ensemble (i.e., the average) to account for uncertainty of reconstructions during these time periods. The Eocene-to-Miocene biome reconstructions are primarily based on fossil records and palynological data and have partly been used to assess the time-integrated area effect for the latitudinal gradient in tree diversity (32). Other interpretations of the history of tropical regions (e.g., (33, 34) during these time periods rely primarily on geological data (evaporites and coal deposits) and do not explicitly reconstruct the tropical rainforest biome (32). For the more recent time periods we used reconstructions for the Middle Pliocene (35) and the Last Glacial Maximum (36) which were derived from new global biome reconstructions based on paleobotanical data and dynamic global vegetation models. The present-day distribution of tropical rainforests was based on global maps of broad vegetation types (37). The details of the sources for biome reconstructions are listed in Table S7.

Table S7: Details of sources for biome reconstructions of the tropical rainforest biome during different geological time periods.

Source	Time periods	Definition of tropical rainforest biome
(29)	Early Eocene (54–49 Mya), Oligocene (35–25 Mya), Miocene (16–10 Mya)	Closed-canopy tropical rain forests
(30)	Early Eocene (54–49 Mya), Oligocene (35–25 Mya), Miocene (16–10 Mya)	Closed canopy megathermal rainforests
(31)	Eocene (60–50 Mya), Oligocene (~30 Mya), Miocene (11.2–5.3 Mya)	Tropical everwet and subtropical summerwet forests
(35)	Middle Pliocene (3.6–2.6 Mya)	Tropical forests
(36)	Last Glacial Maximum (0.021 Mya)	Tropical forests
(37)	Present-day (0 Mya)	Tropical & subtropical moist broadleaf forests

Note: Time ranges are given for each map from each source (Mya = million years ago).

We scanned and georeferenced all maps from the biome reconstructions (see above) and digitized the rainforest areas using ArcGIS 9.3. Maps were projected to Behrmann projection to calculate the area of the rainforest polygons for each time period. From these maps we derived area-time plots for each of four biogeographic regions (South America, Africa, Indomalaya, Australasia) by plotting the area of tropical rainforests against geological time (Fig. 3 in main text). For the Early Eocene, Oligocene, and Miocene we used an ensemble (i.e., average) across the three available sources (29-31) to account for uncertainty in reconstructions. The time estimates for these periods were set to 55 Mya, 30 Mya, and 11 Mya, respectively. From the area-time plots we estimated (1) the area under the curve (AUC) as a summary statistic for the available biome area over geological time (*sensu* (32)), (2) the total loss of tropical rainforest area (Eocene minus Present), (3) the rate of habitat loss measured as the slope β of a simple linear regression of area vs. time, and (4) the minimum rainforest area available during the Cenozoic. All composite measures plus the available area of tropical rainforests during these time periods are provided in Table S8.

Table S8: Estimates of tropical rainforest area over geological time (from the Eocene to the present) for four major biogeographic regions. Four composite measures of tropical forest change over geological time are provided with (1) the time-integrated area measured as area under the curve (AUC), (2) the total loss of tropical rainforest area (Eocene minus Present), (3) the rate of habitat loss measured as the slope β of a simple linear regression of area vs. time, and (4) the minimum rainforest area available during the Cenozoic.

	Area of tropical rainforest biome (in 1,000 km^2)						Time-integrated area (AUC)	Total area loss (in 1,000 km^2)	Rate of area loss (slope β)	Minimum area (in 1,000 km^2)
Region	Eocene	Oligocene	Miocene	Pliocene	LGM	Present				
South America	12,902	8,804	8,498	8,984	8,123	8,569	1,118	4,333	0.071	8,123
Africa	21,640	16,263	15,457	9,530	3,925	3,149	1,419	18,491	0.298	3,149
Indomalay	8,762	7,155	6,945	4,267	3,149	5,889	709	2,873	0.078	3,149
Australasia	7,232	3,422	4,390	2,167	927	1,154	353	6,078	0.098	927

Note: Composite measures of tropical rainforest change over geological time were derived from area-time plots (*sensu* (32)) which plot biome area versus geological time (compare Fig. 3 in main text).

References

1. Dransfield J, *et al.* (2008) *Genera palmarum - the evolution and classification of palms* (Royal Botanical Gardens, Kew).
2. Baker WJ & Couvreur TLP (in press) Biogeography and distribution patterns of Southeast Asian palms. *Biotic evolution and environmental change in Southeast Asia*, eds Gower D, Johnson K, Richardson JE, Rosen B, Rüber L, & Williams S (Cambridge University Press, Cambridge).
3. Dransfield J & Beentje H (1995) *The palms of Madagascar* (Royal Botanic Gardens Kew and The International Palm Society, Kew).
4. Rakotoarinivo M (2010) New species of *Dypsis* and *Ravenea* (Arecaceae) from Madagascar. *Kew Bulletin* 65(2):279-303.
5. Pintaud J-C & Baker WJ (2008) A revision of the palm genera (Arecaceae) of New Caledonia. *Kew Bulletin* 63(1):61-73.
6. Savolainen V, *et al.* (2006) Sympatric speciation in palms on an oceanic island. *Nature* 441(7090):210-213.
7. Kissling WD & Carl G (2008) Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography* 17(1):59-71.
8. Zona S & Henderson A (1989) A review of animal mediated seed dispersal of palms. *Selbyana* 11:6-21.
9. Henderson A (2002) *Evolution and ecology of palms* (The New York Botanical Garden Press, Bronx).
10. Holland RA, Wikelski M, Kummeth F, & Bosque C (2009) The secret life of oilbirds: new insights into the movement ecology of a unique avian frugivore. *Plos One* 4(12).
11. Givnish TJ (2010) Ecology of plant speciation. *Taxon* 59(5):1326-1366.
12. Lord JM (2004) Frugivore gape size and the evolution of fruit size and shape in southern hemisphere floras. *Austral Ecology* 29(4):430-436.
13. Muller-Landau HC, Wright SJ, Calderón O, Condit R, & Hubbell SP (2008) Interspecific variation in primary seed dispersal in a tropical forest. *Journal of Ecology* 96(4):653-667.
14. Jordano P (1995) Angiosperm fleshy fruits and seed dispersers - a comparative analysis of adaptation and constraints in plant-animal interactions. *American Naturalist* 145(2):163-191.
15. Kristiansen T, *et al.* (in press) Environment versus dispersal in the assembly of western Amazonian palm communities. *Journal of Biogeography* doi:10.1111/j.1365-2699.2012.02689.x.
16. Henderson A (2000) *Bactris* (Palmae). *Flora Neotropica Monographs* 79:1-181.
17. Hodel DR (1992) *Chamaedorea palms: the species and their cultivation* (The International Palm Society, Allen Press, Lawrence, Kansas).
18. Henderson A (2011) A revision of *Geonoma* (Arecaceae). *Phytotaxa* 17:1–271.
19. Guimarães Jr PR, Galetti M, & Jordano P (2008) Seed dispersal anachronisms: rethinking the fruits extinct megafauna ate. *Plos One* 3(3):e1745.
20. Galetti M, Keuroghlian A, Hanada L, & Morato MI (2001) Frugivory and seed dispersal by the Lowland Tapir (*Tapirus terrestris*) in Southeast Brazil. *Biotropica* 33(4):723-726.
21. Morici C (2004) Palmeras e islas: la insularidad en una de las familias mas diversas del reino vegetal. *Ecologia Insular*, eds Fernandez-Palacios JM & Morici C (Asociacion Española de Ecología Terrestre, Badajoz), pp 81-122.
22. Thornton IWB, Compton SG, & Wilson CN (1996) The role of animals in the colonization of the Krakatau Islands by fig trees (*Ficus* species). *Journal of Biogeography* 23:577-592.
23. Shanahan M, Harrison RD, Yamuna R, Boen W, & Thornton IWB (2001) Colonization of an island volcano, Long Island, Papua New Guinea, and an emergent island, Motmot, in its caldera lake. V. Colonization by figs (*Ficus* spp.), their dispersers and pollinators. *Journal of Biogeography* 28(11-12):1365-1377.
24. Westoby M, Leishman M, & Lord J (1996) Comparative ecology of seed size and dispersal. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 351(1345):1309-1317.
25. Wright SJ (1992) Seasonal drought, soil fertility and the species density of tropical forest plant communities. *Trends in Ecology and Evolution* 7(8):260-263.

26. Hansen DM & Galetti M (2009) The forgotten megafauna. *Science* 324(5923):42-43.
27. Couvreur TLP, Forest F, & Baker WJ (2011) Origin and global diversification patterns of tropical rain forests: inferences from a complete genus-level phylogeny of palms. *BMC Biology* 9(1):44.
28. Baker WJ, *et al.* (2009) Complete generic-level phylogenetic analyses of palms (Arecaceae) with comparisons of supertree and supermatrix approaches. *Systematic Biology* 58(2):240-256.
29. Morley RJ (2000) *Origin and evolution of tropical rain forests* (John Wiley & Sons, Chichester).
30. Morley RJ (2007) Cretaceous and Tertiary climate change and the past distribution of megathermal rainforests. *Tropical rainforest responses to climate change*, eds Bush MB & Flenley JR (Springer, Berlin), pp 1-31.
31. Willis KJ & McElwain JC (2002) *The evolution of plants* (Oxford University Press, Oxford).
32. Fine PVA & Ree RH (2006) Evidence for a time-integrated species-area effect on the latitudinal gradient in tree diversity. *American Naturalist* 168(6):796-804.
33. Scotese CR (2003) *PALEOMAP project. http://www.scotese.com.*
34. Ziegler A, *et al.* (2003) Tracing the tropics across land and sea: Permian to present. *Lethaia* 36(3):227-254.
35. Salzmann U, Haywood AM, Lunt DJ, Valdes PJ, & Hill DJ (2008) A new global biome reconstruction and data-model comparison for the Middle Pliocene. *Global Ecology and Biogeography* 17(3):432-447.
36. Prentice IC, Harrison SP, & Bartlein PJ (2011) Global vegetation and terrestrial carbon cycle changes after the last ice age. *New Phytologist* 189(4):988-998.
37. Olson DM, *et al.* (2001) Terrestrial ecoregions of the world: a new map of life on earth. *Bioscience* 51(11):933-938.