



**HAL**  
open science

## Quaternary landscape dynamics boosted species dispersal across Southeast Asia

Tristan Salles, Claire Mallard, Laurent Husson, Sabin Zahirovic, Anta-Clarisse Sarr, Pierre Sepulchre

► **To cite this version:**

Tristan Salles, Claire Mallard, Laurent Husson, Sabin Zahirovic, Anta-Clarisse Sarr, et al.. Quaternary landscape dynamics boosted species dispersal across Southeast Asia. *Communications Earth & Environment*, 2021, 2 (1), 10.1038/s43247-021-00311-7 . hal-03443584

**HAL Id: hal-03443584**

**<https://hal.science/hal-03443584>**

Submitted on 23 Nov 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Quaternary landscape dynamics boosted species dispersal across Southeast Asia

Tristan Salles <sup>1</sup>✉, Claire Mallard<sup>1</sup>, Laurent Husson<sup>2</sup>, Sabin Zahirovic <sup>1</sup>, Anta-Clarisse Sarr <sup>3</sup> & Pierre Sepulchre <sup>4</sup>

Sundaland, the inundated shelf separating Java, Sumatra and Borneo from the Malay Peninsula, is of exceptional interest to biogeographers for its species richness and its position at the junction between the Australasian and Indomalay biogeographic provinces. Owing to its low elevation and relief, its physiography is contingent on relative sea-level change, which drove Quaternary species burst in response to flooding episodes. New findings show that the region was predominantly terrestrial during the Late Pleistocene requiring a reassessment of the drivers of its recent biodiversity history. Here we show that physiographic changes have modified the regional connectivity network and remodelled the pathways of species dispersal. From combined landscape evolution and connectivity models, we found four phases of drainage reorganisation and river captures. These changes have fragmented the environment into multiple habitats connected by migratory corridors that cover 8% of the exposed shelf and stretch across the biogeographic provinces. Our results support the theory that rapidly evolving physiography could foster Quaternary biodiversification across Southeast Asia.

<sup>1</sup>School of Geosciences, University of Sydney, Sydney, Australia. <sup>2</sup>CNRS, ISTERre, Université Grenoble-Alpes, Grenoble, France. <sup>3</sup>CL-Climat, CEREGE, Aix-en-Provence, France. <sup>4</sup>LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris Saclay, Paris, France. ✉email: [tristan.salles@sydney.edu.au](mailto:tristan.salles@sydney.edu.au)

Based on the purported Quaternary geodynamic stability of the Sunda Shelf<sup>1–5</sup>, eustatic sea level fluctuations have generally been regarded as an important contributor to the recent Southeast Asia extraordinary biological diversification<sup>6–9</sup>. Divergence and speciation would increase during eustatic highstands and geographic dispersal during glacial sea level lowstands; this alternation would overall remodel the taxonomic composition of the regional biotas<sup>3,10–13</sup>. Under these circumstances, sea level oscillations would have acted as a species pump to increase the regional terrestrial biodiversity<sup>1,10,14–16</sup>. New findings<sup>17,18</sup> have challenged the idea of the prevalence of eustatic controls and have demonstrated that the Sunda Shelf was subsiding throughout the Quaternary, being permanently subaerial before 400 ka. With Sundaland exposed for such a long period, the assumption of sea level dominance is called into question implying that the distribution of Sundaland biogeography either predates the Quaternary<sup>10,19,20</sup> or is modulated by additional factors including variations of edaphic properties<sup>12,14,21</sup>, fragmentation of forested habitats<sup>4,8,21–23</sup> or changes in paleoclimate conditions<sup>4,24,25</sup>.

In this study, we evaluate the contributions of overlooked surface processes on the last 500 kyr biodiversity dynamic of the region. Although investigations on the role that landscapes might have played in the biological evolution of Southeast Asia can be traced back to Darwin and Wallace<sup>1</sup>, limited work has been conducted on the potential roles of geomorphology in shaping the relationships observed in the present-day assembly of the region's biota<sup>5,13,26,27</sup>.

Previous work<sup>28–30</sup> has shown that changing landscape morphology is a major driver of species dispersal strategies as they track their optimum habitat<sup>31,32</sup>. Species running out of suitable habitats either become isolated, go extinct or are forced to coexist with others and adapt. Fragmentation of habitats could favour local endemism and higher speciation rates, both of which contribute positively to increase biodiversity<sup>28,33</sup>. Here we test the role of landscape dynamic in modifying connectivity and dispersal corridors across the exposed Sunda Shelf using a series of calibrated surface evolution simulations<sup>34</sup> forced with eustatic, climatic and tectonic conditions. From changing physiographic characteristics, we then estimate the spatio-temporal permeability of the shelf to lowland evergreen rainforests species movement using a connectivity model<sup>35</sup>. Finally, we use a combination of statistical approaches to quantify preferential and persistent paleo-migration pathways across the shelf since the Late Pleistocene.

## Results

**Sensitivity of Sundaland flooding history to tectonic and sea level conditions.** To evaluate paleo-stream paths and associated paleo-drainage basins evolution, we performed a series of landscape evolution simulations<sup>34</sup> over the last one million years (Fig. 1a). Each simulation accounts for fluvial incision, sediment transport and deposition (in endorheic basins and in the marine realm) as well as hillslope processes (Methods, Fig. 1b). All scenarios are constrained by three external forcing mechanisms: (1) climate, as rainfalls, (2) eustatic sea level fluctuations, and (3) vertical tectonics. In this study, we tested five combinations of these forcing conditions (Supplementary Table 2). Two sea level curves were assessed (Supplementary Fig. 1c), the first three scenarios used the reconstruction from<sup>36</sup> and the last two the sea level stack proposed by<sup>37</sup>. We explored three different tectonic regimes. In scenario 1, we assumed no vertical deformation; in 2, a uniform subsidence rate of  $-0.25$  mm/yr<sup>17</sup>, and for the last 3 we imposed a tectonic map (i.e., vertical land motion, Supplementary Fig. 1b) built from available regional uplift and subsidence rates (Methods, Supplementary Table 1). Finally, in all cases except

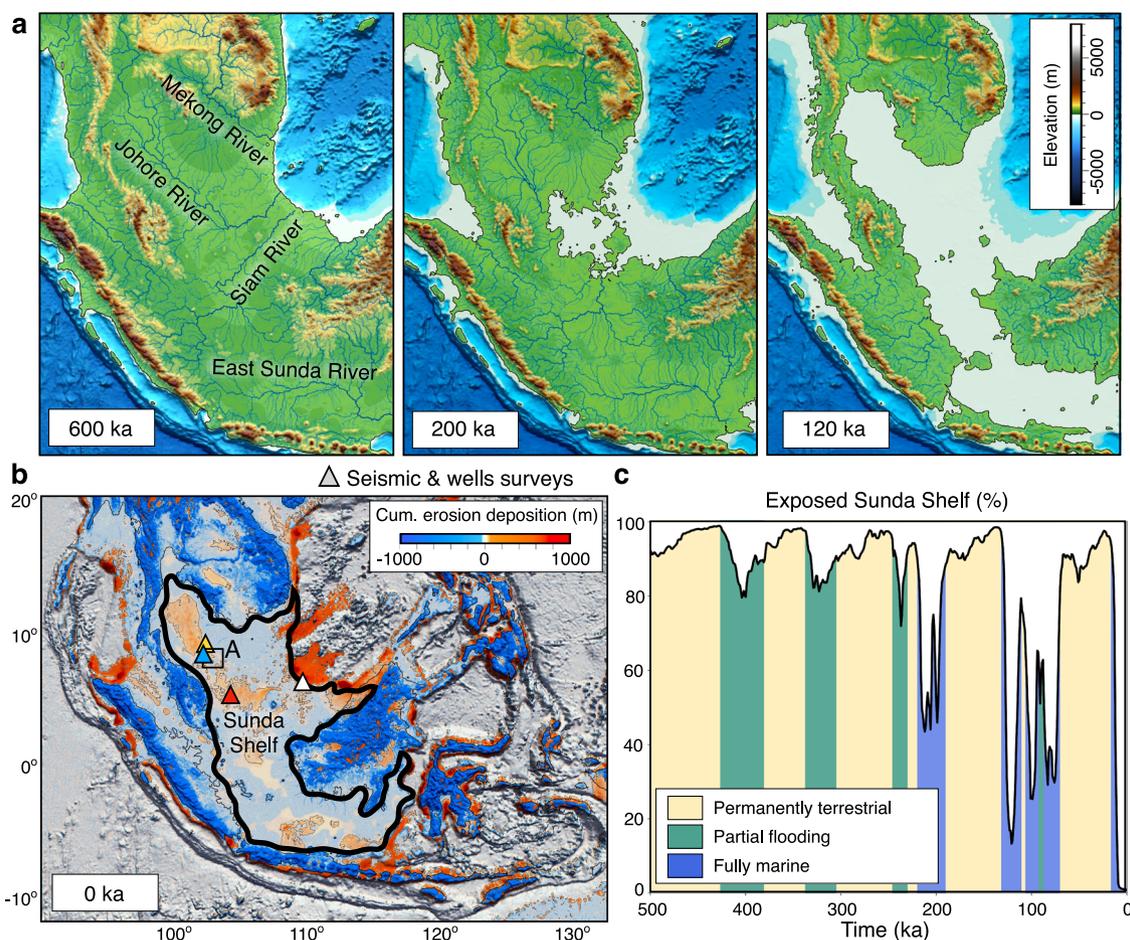
scenario 4 (where we assumed a uniform rainfall of 2 m/yr), we imposed derived precipitation estimates from the PaleoClim database<sup>38–40</sup>. Landscape evolution models were calibrated against paleo-rivers derived from phylogenetic studies<sup>9,13</sup> and seismic surveys as well as flooding events retrieved from boreholes data<sup>41</sup> in the Malay Basin (box A in Fig. 1b) and sediment accumulation across the Sunda Shelf<sup>41–44</sup> (Methods, Supplementary Table 3). In summary, scenarios 1 and 2 are used to isolate the effects of tectonics, while scenario 4 is used as a control for rainfall regime. Scenarios 3 and 5 are used to evaluate the role of sea level on the flooding history of the shelf.

We analysed the flooding history of Sundaland for the five proposed scenarios (Methods - Supplementary Fig. 1c) and found that except for the case with no subsidence (scenario 1), the shelf was fully exposed prior to 400 ka (MIS 11) in agreement with previous work<sup>17,18</sup>. At low sea level periods, sediments bypass the shelf and accumulate in more distal offshore regions<sup>43</sup> (Fig. 1b), while minimal shelf erosion is driven by river incision. Conversely, during marine transgressions, terrestrial sediments fluxes mostly rest on the shelf (Supplementary Fig. 1d). Over 1 Myr, our simulations predict up to 105 m of deposition (for scenario 1, Supplementary Table 3) in the Gulf of Thailand and around 300 m sediment accumulation on the continental slope facing the southeast South China Sea (scenarios 2, 3 and 5 - Fig. 1b).

After 400 ka, level of marine incursions across the shelf are variable between tested scenarios (Fig. 1c and Supplementary Fig. 1c), with incomplete (<80% of the shelf flooded) flooding events during sea level highstands. For variable tectonic subsidence and even when considering the sea level curve with the highest amplitudes<sup>37</sup> (scenario 5), pervasive (more than half the surface of the shelf) flooding prevails for <20% of the time (corresponding to ~80 kyr since MIS 11, Fig. 1c).

We found a clear sensitivity of the flooding history to both tectonic and sea level conditions. For all scenarios with imposed tectonic forcing (2–5), land bridges between Borneo and surrounding regions are disrupted on several occasions. Conversely, most of the shelf remains subaerial over the last million years, and the terrestrial connection from mainland Asia to Java persists up to the Holocene. Our simulations refine recent work on Sundaland subsidence<sup>17</sup> and provide, for the first time, quantitative estimates of the timing and extent of the flooding episodes across the shelf (shaded grey areas in Fig. 2a). From Fig. 2b–d, compiled timing of the divergences for different species groups and between populations varied widely over the last 500 kyr and cannot be attributed to a single event or mechanism such as eustatic sea level fluctuations<sup>18</sup>. Instead, this suggests that the divergences that did happen during the Late Pleistocene were more likely caused by a combination of different drivers acting on species diversification over tens of thousands of years<sup>6</sup>. In the following sections, we investigate the role that physiographic and climatic changes might have played on this regional biological diversification and how it affected Late Pleistocene diversification and post-divergence flow from older speciation events.

**Drainage basin reorganisation and river capture.** Although paleo-drainage systems have been invoked to explain the spatial distribution of Southeast Asia freshwater aquatic species<sup>5,9,13</sup>, quantifying how physiographic changes might have affected regional biota has never been done. Here our first step is to model the region's geomorphological evolution by characterising the main catchments dissecting the shelf and detailing in space and time the major phases of drainage basin reorganisations (Fig. 3).



**Fig. 1 Simulated landscape evolution, sedimentation and Sunda Shelf flooding history under tectonic, eustatic and rainfall forcing conditions.**

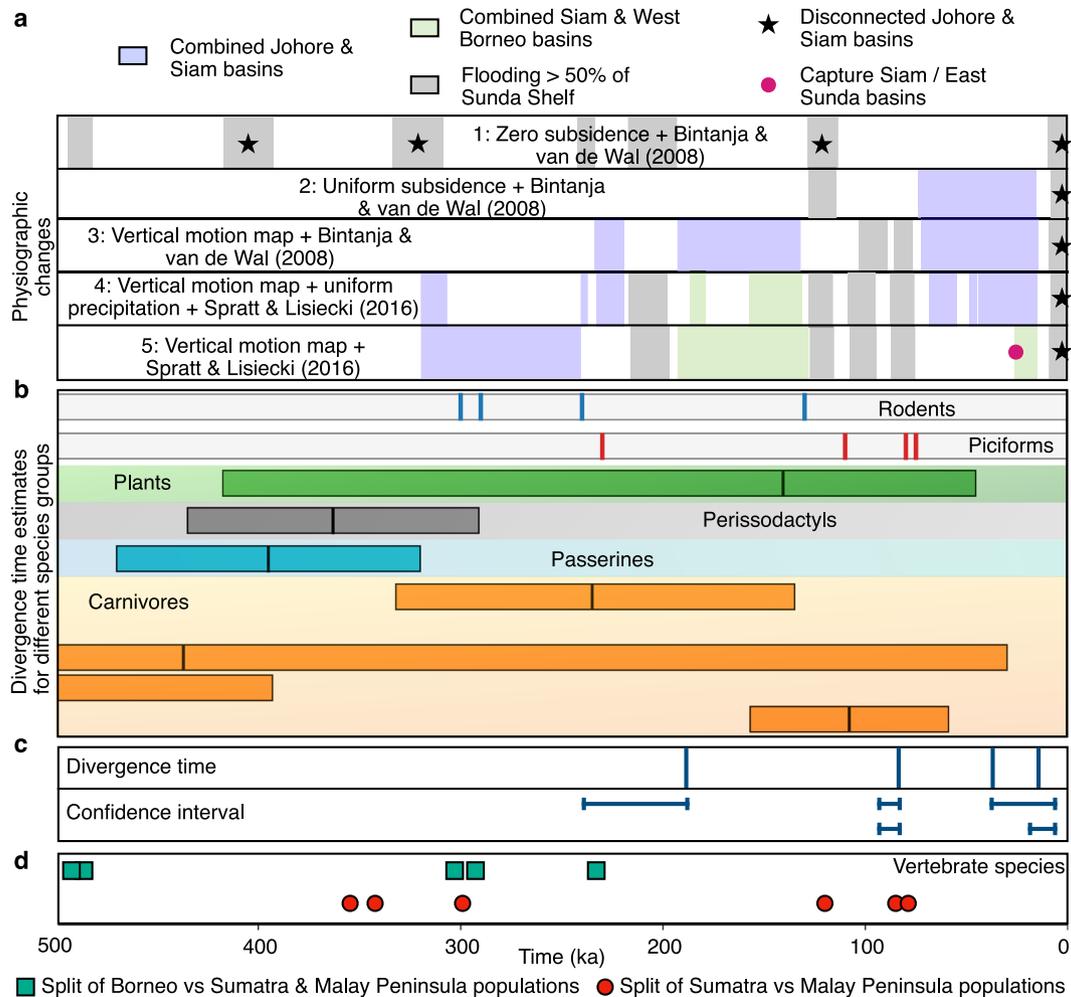
**a** Outputs showing simulated physiography changes for scenario 5 (Supplementary Table 2) induced by spatially variable tectonics (Supplementary Fig. 1b), eustatic (sea level<sup>37</sup>) and atmospheric (precipitation<sup>40</sup>) forcing. **b** Sunda shelf extent prior to 400 ka is delineated by the black contour. Cumulative erosion (blues) and deposition (reds) map obtained after calibration of landscape evolution parameters from subsurface data (seismic surveys and boreholes<sup>41–44</sup>, see Methods). **c** Subaerial exposure of the shelf is calculated at each time interval by computing the ratio between exposed area and Sunda shelf area prior to 400 ka. The shelf alternates three settings: exposed (>80%), partially flooded (>50%), and fully marine periods coloured in yellow, green and blue, respectively.

In all models accounting for subsidence (scenarios 2–5), we found that successive marine regressions and transgressions promote expansions and contractions of individual drainage basins. This is indicated by the robust statistical correlations between sea level fluctuations and catchments characteristics (Pearson's coefficients in Fig. 3b and Supplementary Fig. 2). We also found a 30–50% decrease in drainage area for both the Siam and Johore rivers during the largest intermittent flooding episodes (Fig. 3b and Supplementary Fig. 2).

In our simulations, the Mekong and Singapore main drainage divides remain stable over time (Fig. 3a). When tectonic forcing is considered, the shelf experiences at least one phase of drainage basin capture of the Johore and Siam catchments (shaded blue areas in Figs. 2a and 3b, Supplementary Figs. 2 and 3). For the model presented in Fig. 3 (scenario 5), the Johore, Siam and West Borneo rivers undergo several phases of drainage reorganisations. First, at around 330 ka, the Johore River captures neighbouring headwaters from the Siam Basin. Following a phase of flooding around 240 ka, the newly formed large drainage basin splits into two distinct catchments which have similar characteristics as the ones before the capture (Fig. 3b). Since 200 ka, the Siam and West Borneo basins also experienced two captures lasting for 60 and 15 kyr, respectively. Finally, we notice a large capture of parts of

the Siam headwaters by the East Sunda Basin over the last 30 kyr. This capture lasts until the complete flooding of the Sunda Shelf 10 kyr ago. The simulation from scenario 5, which includes tectonics, sea level and precipitation maps constrained with Earth data, reproduces the number of interpreted marine intrusions<sup>41</sup> (Supplementary Table 2). In addition, the simulated cumulative deposit thicknesses from this scenario is in the range of the ones observed across the shelf (Supplementary Table 2). Finally, it also shows the maximum number of reorganisations between the main catchments, making it the best candidate to estimate geomorphological impacts on species migration (Fig. 2a).

These robust simulated geomorphic changes could be critical to some species. For example, lowland freshwater biotas<sup>6,9,13,26</sup> (Supplementary Fig. 1a and Table 4) have the ability to migrate across the shelf from mainland Southeast Asia to both the Malay Peninsula, Sumatra, Borneo and Java based on simulated drainage basins evolution and river captures (Fig. 3a and Supplementary Fig. 2). By comparing the major physiographic changes from our simulations with divergence times from different groups of terrestrial organisms over the past 320 kyr (Fig. 2b–d), we found that many of the Johore, Siam and West Borneo reorganisations often predate measured divergence times and population splits for the considered time period.



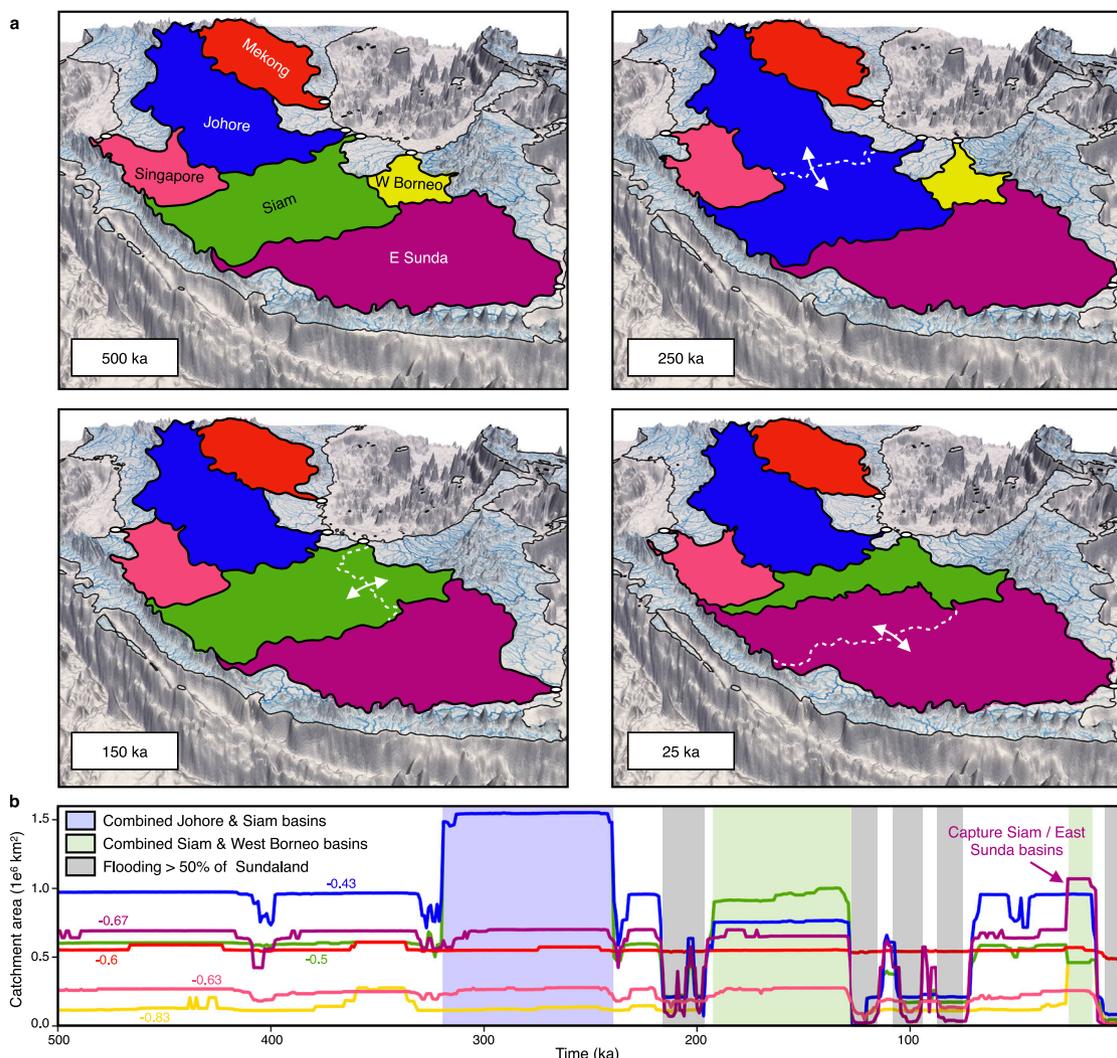
**Fig. 2 Physiographic changes and regional species dispersal.** **a** Physiographic changes for each landscape evolution simulation from scenarios 1–5 (basin reorganisations, drainage captures, flooding and disconnections between main landmasses). **b** Estimated divergence through time (coloured rectangle: left and right sides are the min and max divergence time, and the middle line is the mean) for different terrestrial organisms (carnivores<sup>62–65</sup>, passerines<sup>3,66,67</sup>, barbets<sup>68</sup>, perissodactyls<sup>69</sup> and plants<sup>70</sup>). Estimated divergence time for piciforms<sup>6</sup> and rodents<sup>16,71</sup> without confidence intervals are in red and blue, respectively. **c** Estimated time of divergence and confidence interval for songbird species<sup>21</sup>. **d** Age estimates of population split within vertebrate species<sup>6</sup> between Borneo, Sumatra and the Malay Peninsula. Main divergence times presented in **b** are extracted from Table 1 in<sup>18</sup>.

**Geomorphically derived cost surfaces for lowland evergreen rainforests species.** From the landscape evolution models, we assessed the roles of geomorphology on past movements across the exposed Sunda Shelf over the last million years. Models based on structural connectivity can be used to evaluate species dispersal, gene flow, and other ecological functions of a particular landscape<sup>35,45,46</sup>. Often, a focal-species approach is implemented and considers only one or few species and their specific dispersal thresholds to evaluate the connectivity of a larger suite of species based on different resistance values derived from chosen landscape features. For our large-scale, long-term analysis, this approach may require choosing a large number of focal species to represent diverse habitat types<sup>35,46</sup>. An alternative technique, which does not depend on specific biological or ecological species characteristics is employed in this study. It consists in generating corridors for structural connectivity using a species-agnostic approach<sup>45,47</sup> focused on particular landscape geomorphological attributes (Methods).

Even though species-agnostic methods do not rely on any particular species, they are often used to predict functional connectivity across large regions for multispecies favouring a similar type of habitat<sup>46</sup> (e.g., closed forests, edge habitats, open

prairies). Some paleo-environmental reconstructions suggest lowland rainforests across Sundaland were extensive during sea level lowstands<sup>4,24</sup> while others hypothesize that the Pleistocene land bridges were much drier and substantially covered by grasslands or savannah<sup>22,23</sup>. Conversely, even in the presence of a transequatorial savannah, studies have proposed that gallery forests along paleo-rivers running through the xeric exposed land could have acted as corridors for forest-dependent taxa<sup>5,8,27</sup>. As a result, our geomorphically derived cost surfaces have been designed to evaluate connectivity pathways for lowland evergreen rainforests species since these organisms would be impacted by barriers to migration created by flooding episodes during sea level highstands and would also be affected by changes associated to drainage basin reorganisation and river captures under the assumption of limited rainforest habitats. In addition and depending on species, rivers network could be viewed either as corridor<sup>48,49</sup> or as barrier<sup>50,51</sup> to migration. To account for these two cases, we defined different cost surfaces related to closeness to rivers when evaluating landscape connectivity across Sundaland.

At specific time steps deemed representative of the different stages of Sundaland basins physiography and for the terrestrial domain, we compute cost surfaces from specific geomorphic



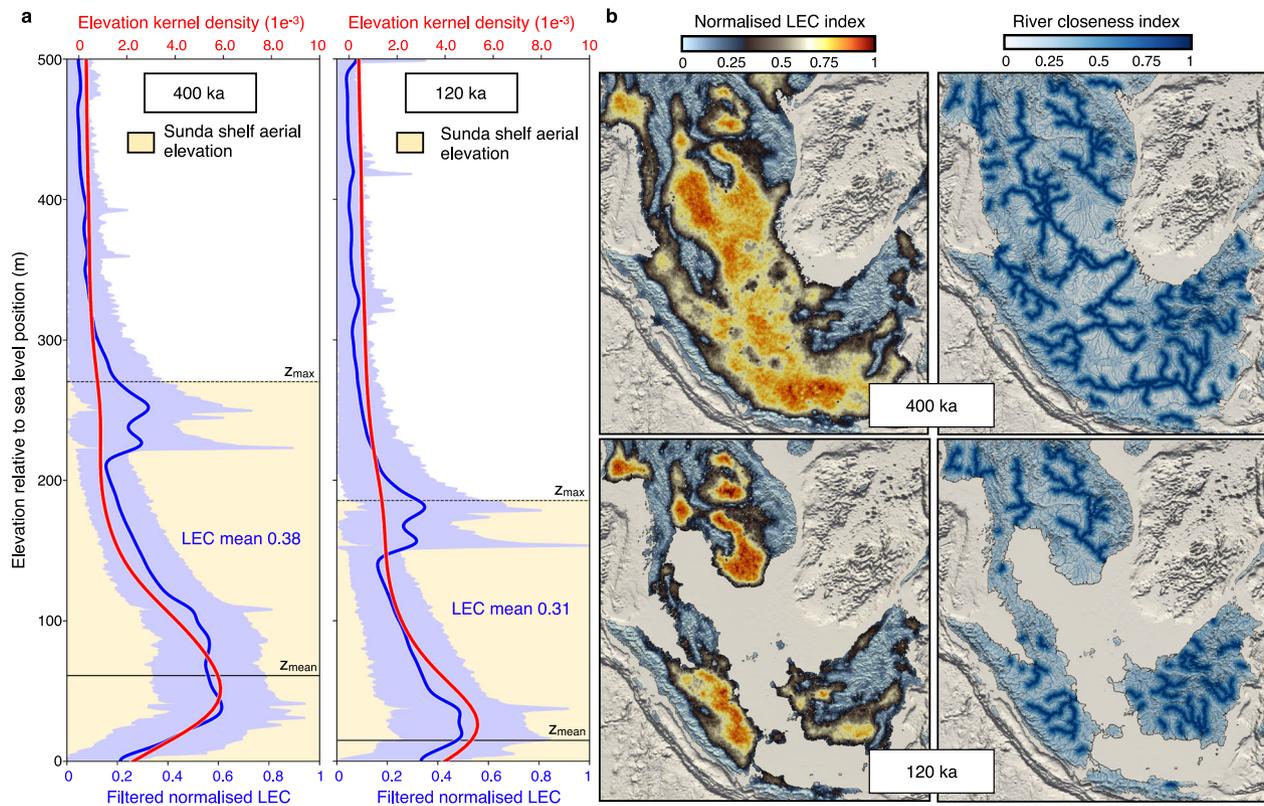
**Fig. 3 Temporal evolution of Sundaland drainage basins distribution and changes in main catchments characteristic.** **a** Evolution of drainage basins tessellating the Sunda Shelf for scenario 5. Basins are generated from the simulation results based on flow accumulation and direction. The drainage divide positions are changing over time but these changes, expect during reorganisations and captures, are limited due to small tectonic rates and shelf physiography. **b** Catchment areas over the last 500 kyr for the Johore, Siam, Mekong, East Sunda, Singapore and West Borneo basins (coloured lines refer to coloured basins in (a)). Pearson’s coefficients of correlation between area and sea level fluctuations<sup>37</sup> are also provided.

features derived from our landscape evolution models. As previously mentioned, these cost surfaces are designed to represent long-term suitability of the landscape to the movement of lowland evergreen rainforests species and combine the following three natural landscape features: (1) the normalized landscape elevational connectivity (LEC) metric<sup>33</sup> explained in the Methods, (2) the distance to main river systems where rivers are used either as corridors (Fig. 4b) or barriers to species movement, and (3) the average local slope.

We analyse LEC<sup>33,52</sup> distribution for the simulated terrestrial part of the landscapes at specific time intervals corresponding to different sea level positions (Fig. 4a). Higher LECs correspond to places where a pool of species adapted to a particular habitat will have the ability to move up and down their initial and preferred habitat elevation and to migrate across the landscape more easily than lower LEC regions. The distribution of normalised LECs is maximum for regions within the Sunda Shelf elevation range. Graphs, in Fig. 4a, indicate that high connectivities are distributed around two regions. During full exposure and partial flooding, we found that LEC peaks around mid-elevation (between 10 and

80 m above sea level—blue line close to  $z_{mean}$  in Fig. 4a). Mid-elevations provide maximum surface area (peak in elevation kernel density—red line in Fig. 4a), and species can easily move up or down to accommodate terrestrial habitats waxing and waning. Not surprisingly, higher elevations are characterized by low landscape elevational connectivities (Fig. 4a) which promote isolation and increase endemism<sup>29</sup>. A second peak is located on the higher elevational range of the Sunda Shelf ( $z_{max}$ ). This second peak is within the shelf elevation range but corresponds to regions outside the shelf and relates to flood plains and plateaus found in the northern part of the simulated domain (i.e., Chao Phraya flood plain as well as the Khorat Basin in Thailand and to a lesser extent the Tonle Sap Lake in Cambodia).

For landscape connectivity computation, we assumed that costs related to LEC but also distance to river and local slope are equally weighted with categorical values ranging from 0–20 (Fig. 4b). Our choice of resistance and weight values implies assumptions about species movement ability, and many examples can be found where a natural topographic feature may act as a barrier for one species but not for another. The resistance maps are then combined to



**Fig. 4** Geomorphic parameters used for cost surface calculation. **a** Plots of normalized landscape elevational connectivity (LEC) for the entire aerial domain and within elevational bands (blue lines and standard deviation in light blue), and elevation kernel density (red lines) for the entire exposed domain at chosen time intervals for the simulation forced with variable tectonic map and the sea level curve from<sup>37</sup>. Yellow areas represent the Sunda Shelf aerial elevation extent, with lines showing the exposed shelf mean and maximum elevations ( $Z_{mean}$  and  $Z_{max}$ ). **b** Examples of LEC and river closeness indexes maps for fully exposed (400 ka) and flooded (120 ka) Sunda Shelf conditions. These indexes combined with local slope distribution are then used to produce cost surfaces at specific time intervals over which preferential pathways are computed.

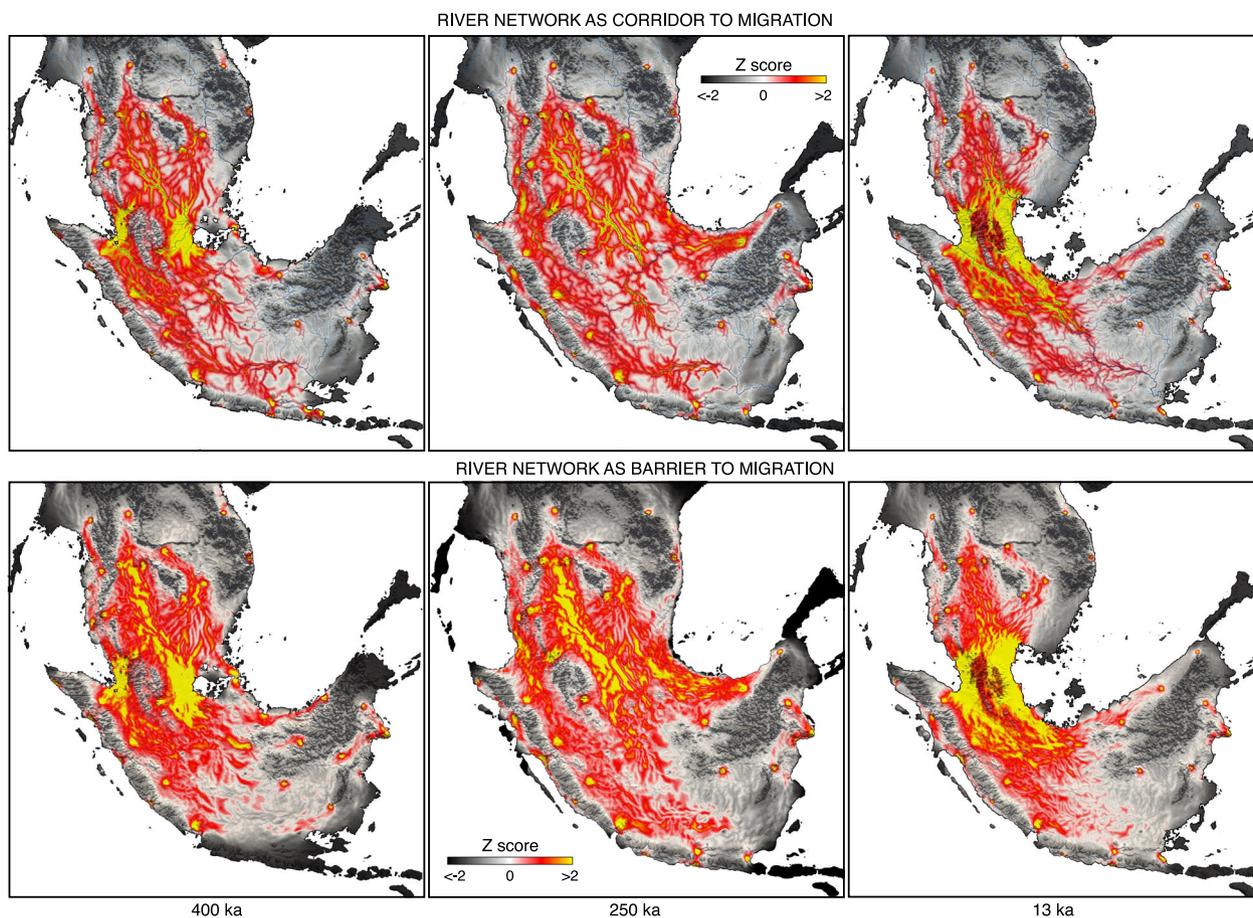
produce regular cost surfaces grids (5 km side squared cells), ranked from high cost (>30) that are impermeable to movement (e.g., ocean, remote/close from river drainage network, steep local slope and low elevational connectivity) to low cost (<10) when permeable to movement (e.g. terrestrial, close/remote to rivers, low local slope, and high elevational connectivity).

**Landscape connectivity in the exposed Sunda Shelf.** We used circuit theory<sup>35</sup> to extract preferential pathways over time. The approach relies on the temporal evolution of the cost surfaces from the species-agnostic approach<sup>46</sup> presented in the previous section. For our study, we connected the eight neighbouring cells as an average cost using the pairwise mode<sup>35</sup>. In each pair of points, one point is given a current source of 1 ampere and the other is connected to the ground<sup>35</sup>. Effective resistance, voltage and current are then calculated over the landscape between these points. The operation is repeated between all pairs, and the final map is produced by summing current values calculated between all pairs.

The resultant current density maps are a prediction of functional connectivity, whereby high values (high current flow) represent a high probability of use by random walkers (Supplementary Fig. 4). In the case where rivers network represents corridors to migration, high current flows are mostly associated with main river systems dissecting the Sunda Shelf around mean shelf elevation range when the shelf is exposed (Supplementary Fig. 4a, b). During partial or full marine transgression events, the remaining exposed shelf exhibits

extensive high flow regions showing that the limited shelf areas in such cases would become preferential corridors for species dispersal (Supplementary Fig. 4c). Highly constrained regions (i.e., high current flow - Supplementary Figs. 4 and 5a) represent corridors where movement would be funnelled. In these corridors, even a small loss of area would disproportionately compromise connectivity and thus may jeopardize species migration<sup>35</sup>.

To normalise connectivity over time, we compute current flow z-scores (Methods, Fig. 5 and Supplementary Fig. 5b). During full shelf exposure and when considering rivers as corridors to migration, the region is characterised by narrow high z-scores pathways which follow the Johore and Siam main paleo-drainage systems (Fig. 5 at 250 ka - top panel). At 250 ka, these pathways form a continuous migratory corridor across Sundaland and result from the capture of the Siam Basin by the Johore River (Fig. 3a). When rivers are considered as barriers to species movement, Sundaland exhibits patches of high connectivity areas delimited by the major streams flowing across the Johore and Siam basins (Fig. 5 at 250 ka - bottom panel). During partially flooded periods, the shelf is characterised by extended regions of high connectivity limited by the Titiwangsa Mountains on the Malay Peninsula to the west and the shallow marine incursion to the east (Fig. 5 at 400 ka). High connectivity regions become even more restricted when the shelf is nearly fully submerged (Fig. 5 at 13 ka) and represent essential areas for species migration due to habitat reduction and lack of alternative pathways<sup>5,53</sup>. It is worth mentioning that grey regions in Fig. 5 are not areas that did not support gene flow (e.g. between Java and Borneo) but exhibit low



**Fig. 5 Regional distribution of current flows highlighting connected regions at different time intervals for rivers network considered as corridor or barrier to migration.** Cost surfaces combining landscape elevational connectivity, distance to the main river and local slope are used to compute current flow fields<sup>35</sup>. Current flow values are then standardised at each time step by calculating z-scores to evaluate temporal changes in landscape connectivity across the different maps. Positive z-scores (red to yellow) represent high connectivity areas acting as preferential pathways across the Sunda Shelf.

resistance to movement for geomorphology-derived cost surfaces and therefore are not showing up as high current density areas. Our results show that physiographic changes exert a strong control on the location of high connectivity pathways, especially when the shelf is fully exposed. Rivers either trigger new dispersal routes (Fig. 5 top panel) or fragment the environment in multiple habitats (Fig. 5 bottom panel), both of which are critical to biodiversity evolution.

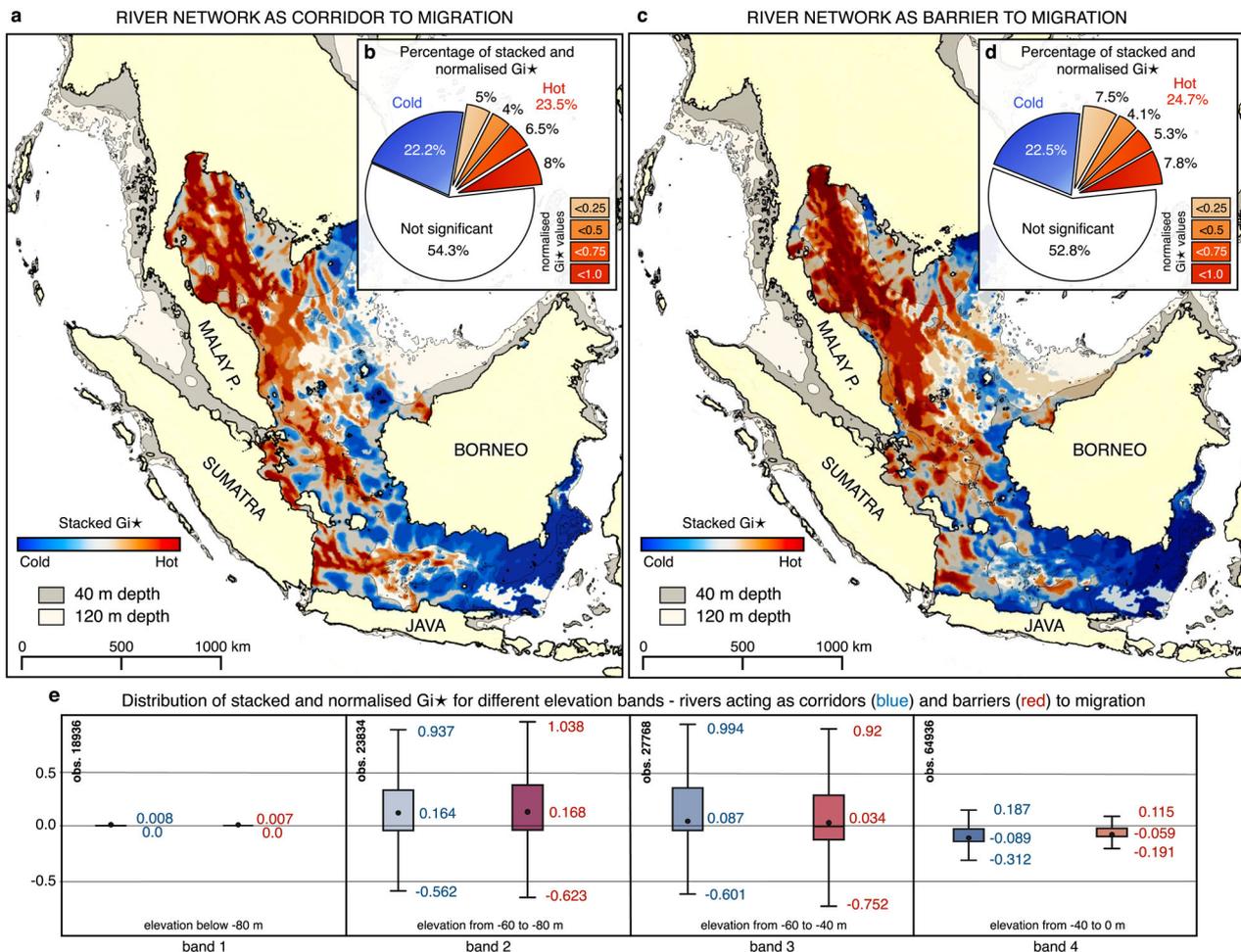
**Hotspot analysis of preferential migration pathways.** To explore spatio-temporal migration paths, we measure the degree of spatial clustering in successive current flow fields during periods of marine regression (corresponding to at least 50% of shelf exposure—scenario 5 in Supplementary Fig. 1c). For each map, we compute the Getis-Ord  $G_i^*$  index showing statistically significant connectivity hotspots (Methods, Supplementary Fig. 5c). We then stack all time slices to evaluate persistent corridors over geological time scales (Fig. 6a and c). Hotspots cover about 24% of Sundaland in both cases, with normalized  $G_i^*$  above 0.75 making ~8% of the region (Fig. 6b and d). Each combined hotspots map predicts that hotspots are preferentially located across the paleo Johore (Gulf of Thailand) and Siam basins and are currently between 40 and 80 m below sea level (Fig. 6e). Cold spots (blue regions in Fig. 6a and c) correspond to areas where gene flow is not constrained; hence current density in these regions appears to be diffused. Geomorphology-controlled hotspot regions (Fig. 6a and c) highlight a network

of preferential, well-connected biodiversity corridors that favour species migration between mainland Southeast Asia and Sumatra, Borneo and Java.

Our approach provides an innovative way of reframing the dynamics of species evolution on our planet, by comprehensively integrating several components of the Earth system (tectonics, climate, and surface processes) over geological time scales. We found that the fast pace of physiographic changes in the region has the ability to drive diversification processes by facilitating habitat fragmentation and promoting preferential pathways across the shelf. From a methodological standpoint, we show that extracting high connectivity areas from landscape evolution simulations provides an opportunity to evaluate how surface processes influence species movements and drive diversification. Conceptually, we demonstrate that accounting for these dynamics to delineate past biodiversity hotspots is critical to advance our comprehension of species migration both across Southeast Asia and globally and to further our understanding of the dynamics of life on Earth.

## Methods

**Surface processes model and forcing mechanisms.** Landscape evolution over the last one million years interval is performed with the open-source modelling code Badlands<sup>34</sup>. It simulates the evolution of topography induced by sediment erosion, transport, and deposition (Fig. 1a). Amongst the different capabilities available in Badlands, we applied the fluvial incision and hillslope processes, which are described by geomorphic equations and explicitly solved using a finite volume discretisation. In this study, soil properties are assumed to be spatially and



**Fig. 6 Hotspot analysis based on spatio-temporal distribution of Sunda Shelf connectivity.** **a** and **c** Getis-Ord ( $G_i^*$ ) index obtained by stacking temporal hot and cold spot analysis of current flow fields over multiple time steps. Red regions correspond to hotspot areas where computed connectivity from geomorphological features is high and favours species migration. **b** and **d** Pie chart presenting percentages of the shelf hot/cold spots. The hotspot category is sliced based on normalised stacked  $G_i^*$  from warm (normalised from 0. to 0.25) to hot (from 0.75 to 1). **e** Boxplots displaying the distribution of normalised stacked  $G_i^*$  for four different elevation bands. Each boxplot shows the minimum, first quartile (Q1), median, third quartile (Q3), and maximum for each band.

temporally uniform over the region, and we do not differentiate between regolith and bedrock. It is worth noting that the role of flexural responses induced by erosion and deposition is also not accounted for. Under these assumptions, the continuity of mass is governed by vertical land motion ( $U$ , uplift or subsidence in m/yr), long-term diffusive processes and detachment-limited fluvial runoff-based stream power law:

$$\frac{\partial z}{\partial t} = U + \kappa \nabla^2 z + \epsilon (PA)^m \nabla z^n \quad (1)$$

where  $z$  is the surface elevation (m),  $t$  is the time (yr),  $\kappa$  is the diffusion coefficient for soil creep<sup>34</sup> with different values for terrestrial and marine environments,  $\epsilon$  is a dimensional coefficient of erodibility of the channel bed,  $m$  and  $n$  are dimensionless empirical constants, that are set to 0.5 and 1, respectively, and  $PA$  is a proxy for water discharge that numerically integrates the total area ( $A$ ) and precipitation ( $P$ ) from upstream regions<sup>34</sup>.

Both  $\kappa$  and  $\epsilon$  depend on lithology, precipitation, and channel hydraulics and are scale dependent<sup>34</sup>. All our landscape evolution simulations are running over a triangular irregular network of  $\sim 18. e^6 \text{ km}^2$  with a resolution of  $\sim 5 \text{ km}$ , and outputs are saved every 1000 yr.

The detachment-limited fluvial runoff-based stream power law is computed with a  $\mathcal{O}(n)$ -efficient ordering approach<sup>54</sup> based on a single-flow-direction approximation where water is routed down the path of the steepest descent. The flow routing algorithm and associated sediment transport from source to sink depend on surface morphology, and sediment deposition occurs under three circumstances: (1) existence of depressions or endorheic basins, (2) if local slope is less than the aggregational slope in land areas and (3) when sediments enter the marine realm<sup>34</sup>. Submerged sediments are then transported by diffusion processes defined with a constant marine diffusion coefficient<sup>34</sup>.

All landscape simulations are constrained with different forcing mechanisms, and five scenarios were tested (Supplementary Table 2).

First, we impose precipitation estimates from the PaleoClim database<sup>38–40</sup>. These estimates are products from paleoclimate simulations (coupled atmosphere-ocean general circulation model) downsampled at approximately the same resolution as our landscape model ( $\sim 5 \text{ km}$  at the equator). Annual averages of precipitation rates are then used to provide rainfall trends in our simulations based on the ten specific snapshots available (from the mid-Pliocene warm period to late Holocene and present day). Between two consecutive snapshots, we assume that precipitation remains constant for the considered time interval. For exposed regions that are considered flooded in the PaleoClim database, we define offshore precipitations using a nearest neighbour algorithm where closest precipitation estimates are averaged from PaleoClim inland regions. To evaluate the role of precipitation variability on landscape dynamics, we also run a uniform rainfall scenario (2 m/yr obtained by averaging the annual precipitation rates from the PaleoClim database).

Secondly, the models are forced with sea level fluctuations known to play a major role in the flooding history of the Sunda Shelf<sup>11,13,53</sup>. Two sea level curves are tested (Supplementary Fig. 1d). To account for the inherent uncertainty in reconstructed sea level variations, we chose a first curve<sup>37</sup> obtained from a sea level stack constructed from five to seven individual reconstructions that agrees with isotopically adjusted coral-based sea level estimates at both 125 and 400 ka. The second one is taken from the global sea level curve reconstruction<sup>36</sup> based on benthic oxygen isotope data and has been recently used to reconstruct the subsidence history of Sundaland<sup>17,18</sup>.

The last forcing considered in our study is the tectonic regime. First, we chose to explore a non-tectonic model based on the default assumption of stability for the shelf<sup>17</sup>. Secondly, we assumed a uniform subsidence rate of  $-0.25 \text{ mm/yr}$  recently derived from a combination of geomorphological observations, coral reef growth numerical simulations and shallow seismic stratigraphy interpretations<sup>17</sup>. Then, to

represent the regional variations in the tectonic regime, we have compiled and digitised a number of calibration points (Supplementary Fig. 1b and Table 1) that were used to produce a subsidence and uplift map by geo-referencing calibration points and available tectonic polygons, and by Gaussian-smoothing and normalising the uplift and subsidence rates between the calibrated range to avoid sharp transitions in regions without observations. The resulting map does not account for fine spatial scale tectonic features such as fault systems<sup>43,55</sup> or orogenic and sedimentary related isostatic responses. It rather represents a regional vertical tectonic trend with an overall uplift of Wallacea and NW Borneo regions and long-wavelength subsidence of Sunda Shelf and Singapore Strait<sup>17</sup>.

**Landscape evolution model calibration.** The landscape models start during the Calabrian in the Pleistocene Epoch, one million years before the present. At each time interval, the landscape evolves following Eq. (1) and the surface adjusts under the action of rivers and soil creep (Fig. 1a). In addition to surface changes, we extract morphometric characteristics such as drainage basins extents, river profiles lengths (Fig. 3 and Supplementary Fig. 2), distance between main rivers outlet (Supplementary Fig. 3) and tracks the cumulative erosion and deposition over time (Fig. 1b and Supplementary Fig. 1d).

For model calibration, we perform a series of steps consisting in adjusting the initial elevation and the erosion–deposition parameters (i.e.,  $\kappa$  and  $\epsilon$  in Eq. (1)) to match with different observations.

The initial paleo-surface is obtained by applying the uplift and subsidence rates backwards to calculate the total change in topography for the 1 Myr interval. Then, we test the simulated paleo-river drainages against results from a combination of phylogenetic studies<sup>9,13</sup> and paleo-river channels and valleys found from seismic and well surveys<sup>41,42,44</sup>. Iteratively, we modify our paleo-elevation to ensure those main river basins (e.g., Johore, Siam, Mekong, East Sunda) encapsulate the paleo-drainage maps reconstructed using lowland freshwater taxa described in<sup>13</sup> (Supplementary Fig. 1a and Table 4) and that the major rivers follow paleo-rivers systems derived from both 2D and 3D seismic interpretations (Fig. 1b).

For surface processes parametrisation, we tested different ranges of diffusion and erodibility coefficients and compared the final sediment accumulation across the Sunda Shelf (Fig. 1b) using estimated deposit thicknesses<sup>41–44</sup>. The Sunda Shelf is predominantly experiencing deposition over the past 500 kyr and increases in deposition are positively correlated with periods of sea level rise (i.e., Pearson's coefficients for correlation with sea level above 80%, Supplementary Fig. 1d). After exploring a range of values, we set  $\kappa$  values to  $1. e^{-2}$  and  $8. e^{-2} m^2/yr$  for terrestrial and marine environments and  $\epsilon$  between 2.5 and  $7.5 e^{-8} yr^{-1}$  for the different scenarios to fit with chosen surveys dataset (Supplementary Table 2 and 3).

Upon uniform subsidence case ( $-0.25 mm/yr$ ), flooding is limited, and the shelf only undergoes two full marine transgressions (>60% of the shelf flooded) around 125 ka and during the last 10–20 kyr (Supplementary Fig. 1c). Upon spatially variable tectonics (non-uniform subsidence), partial flooding events are more pervasive, with higher magnitudes and greater temporal durations. Due to the shallow and flat physiography of Sundaland, we also note that even small increases in sea level amplitudes (<10 m, as bore between our two sea level curves<sup>36,37</sup>) could trigger up to a 30% increase in shelf area inundation (for instance the red and blue curves in Supplementary Fig. 1c during MIS7, ~200 ka). In addition, we further tested our combination of forcing mechanisms, initial paleo-surface and model parameters by counting the number of flooding events simulated within the Malay Basin on the Sunda Shelf (box A in Fig. 1b). For both sea level curves and variable tectonic map, we recorded at least five marine incursion episodes, similar to the number of events found for the same area in<sup>41</sup> based on interpreted facies characteristics site survey borehole data and 3D seismic sections analysis.

**Landscape elevational connectivity.** The landscape elevational connectivity (LEC) is a measure of the energy required by a pool of adapted species to move outside its usual niche width, up and down their initial elevation range, to spread and colonise any other habitats<sup>33</sup>. This metric can predict local species richness ( $\alpha$  diversity) obtained from full metacommunity models when applied to real landscapes<sup>33</sup>. In a recent study<sup>52</sup>, comparisons between generated Badlands model landscapes using both uniform and orographic precipitation conditions have shown similar results when the metric is applied over geological time scales.

Landscape elevational connectivity at node  $i$  ( $LEC_i$ ) relies on the following set of equations<sup>33</sup>:

$$LEC_i = \sum_{j=1}^N C_{ji} \quad (2)$$

$$-\ln C_{ji} = \frac{1}{2\sigma^2} \min_{p \in (j \rightarrow i)} \sum_{r=2}^L (z_{k_r} - z_j)^2$$

where  $C_{ji}$  quantifies the closeness between sites  $j$  and  $i$  with respect to elevational connectivity and measures the cost for a given species adapted to cell  $j$  to move to cell  $i$ . This cost is a function of elevation and evaluates how often species adapted to the elevation of cell  $j$  have to travel outside their optimal species niche width ( $\sigma$ ) to reach cell  $i$ <sup>52</sup>.  $p = [k_1, k_2, \dots, k_L]$  (with  $k_1 = j$  and  $k_L = i$ ) are the cells in the path  $p$  from  $j$  to  $i$ .

Solving Eq. (2) relies on Dijkstra's algorithm accounting for diagonal connectivity between cells and implemented using the scikit-image library<sup>56</sup>. Even

with a balanced parallel implementation, the  $LEC_i$  calculation is slow as a Dijkstra tree for all nodes must be created, and least-cost distances between each node and all others have to be calculated.

In this study, to decrease computation cost, we have modified the initial algorithm<sup>52</sup> to limit the number of nodes that needs to be visited when calculating least-cost distances. We gave each cell an initial amount of energy that is consumed as the pool of species adapted to a specific cell is moving across the topography. The energy expenditure depends on (1) the differences in elevation and (2) the distances between neighbouring cells. Here, the initial amount of energy is set to 2000 and we weight the energy dissipation based on travelled distances by 0.4% assuming that species have the ability to move easily over long distances if they stay in their optimal niche range. This assumption is valid in our case as we are simulating the connectivity associated with a large area ( $\sim 18. e^6 km^2$ ) and over geological time scale.

**Connectivity mapping from current density modelling.** By considering all possible pathways, circuit theory offers a better modelling alternative to least-cost path approaches as it captures all the dynamics at play in travel decisions based on provided resistance maps. We chose Circuitscape<sup>35</sup> to model multiple pathways. The software uses random walk and electric current running through a circuit. Electric current runs across our cost surfaces between predefined source points. We position these points across Sundaland (approximately 250 km apart) chosen along the outer margin ( $\approx 100$  m above sea-level) of the maximum fully submerged shelf coastline. Circuitscape functions as a graph, where each cell centre is a node connected to neighbouring nodes with links<sup>35</sup>. The graph is interpreted as a circuit, and links have cost values combining the laws of electricity to estimate species flow. Effective resistance, voltage and current is then calculated over the landscape between the prescribed points with Ohm's and Kirchhoff's laws<sup>35</sup>.

In Ohm's law, voltage  $V$  applied to a resistor  $R$  gives the current  $I$  through  $I = V/R$ ; as a result, a lower resistance (e.g., low geomorphic cost) in the landscape will correspond to higher flow of species. Kirchhoff's law deals with effective resistance; when nodes are connected to several resistors the effective resistance will be the sum of the resistances:  $R = R_1 + R_2$ .

**Connectivity maps statistical analyses.** To evaluate the relative contribution of each of the three geomorphological features, we compute current density maps for different levels of Sunda Shelf exposure using different combinations of resistance surfaces. We then tested each feature independently (e.g., only landscape elevational connectivity, only distance to rivers (with rivers as barriers and as corridors) and only local slope), then pairwise costs (e.g., landscape elevational connectivity and distance to rivers, landscape elevational connectivity and local slope, local slope and distance to rivers) and finally all costs combined (Supplementary Fig. 5a).

Qualitative comparisons of current density maps rely on visual interpretation and are often altered by the choice of colour scale used to distinguish regions of high connectivity<sup>47</sup>. To perform a better evaluation, we first express current flow values ( $c$ ) as z-scores ( $z_{sc} = (c - \bar{c})/sd$ ) by subtracting the mean current value  $\bar{c}$  and dividing it by the standard deviation  $sd$ . We then used three different thresholds to estimate regions that have positive mean values (i.e.,  $z_{sc} > 0$ ) and values higher than one and two standard deviations ( $z_{sc} > 1$  and 2, respectively, Supplementary Fig. 1b). The approach provides a quantitative assessment of flow maps sensitivity to the chosen resistance maps.

To gain additional insights into the distribution of connectivity regions across the shelf, we also employed a local spatial autocorrelation indicator, namely the Getis-Ord  $G_i^*$  index<sup>57</sup>. This hotspot analysis method assesses spatial clustering of the obtained current density maps, and the resultant z-scores provide spatially and statistically significant high or low clustered areas. The approach consists in looking at each local current value relative to its neighbouring one. From this spatial analysis, we extract both statistically significant hot and cold spots for each combination of resistance surfaces (Supplementary Fig. 5c). To extract statistically significant and persistent biogeographic connectivity areas across the exposed Sunda Shelf, we then combine all hotspots together and define preferential migration pathways as regions having a positive  $G_i^*$  z-scores for all resistance surfaces combination.

We used the function  $zscore$  in the SciPy *stats* package to obtain the z-scores and the *ESDA* library for the  $G_i^*$  indicator computation.

**Modelling assumptions and limitations.** There are a number of important caveats for interpreting our modelling results.

First, we made several assumptions related to our transient landscape evolution simulations. A single-flow direction algorithm<sup>54</sup> was used to simulate temporal changes in river pathways. Recent work<sup>58</sup> has shown that this algorithm might lead to numerical diffusion, fast degradation of knickpoints and underestimation of river captures particularly in flat regions. One way to address this would be to use a multiple flow direction method<sup>59</sup> which allow for a better representation of flow distribution across the landscape. In this study, we also assumed a uniform and invariant soil erodibility coefficient for the entire domain and a detachment-limited erosion law. Even though the erodibility coefficient was calibrated independently for each simulation (Supplementary Table 3), soil cover and properties vary notably between Borneo, Sumatra, Java and the Malay Peninsula and soil

conditions for the exposed sea floor would have changed significantly over successive flooding events<sup>12</sup>. Badlands software<sup>34</sup> allows for multiple erodibility coefficients representing different soil compositions to be defined, and this functionality could be used to evaluate the impact of differential erosion on physiographic changes. Similarly, several transport-limited laws are also available and could be compared against our detachment-limited simulations.

A second set of simplifications lies in the climatic conditions (i.e., rainfall regimes) used to force our simulations. We relied on the PaleoClim database<sup>40</sup> which contains nine high-resolution paleoclimate datasets<sup>38–40</sup> corresponding to specific time periods (4.2–0.3 ka, 8.326–4.2 ka, 11.7–8.326 ka, 12.9–11.7 ka, 14.7–12.9 ka, 17.0–14.7 ka, ca. 130 ka, ca. 787 ka and 3.205 Ma). The climate simulations from which these time periods are extracted do not consider emerged Sunda Shelf for the oldest inter-glacial events which can result in incorrect climatic pattern<sup>60</sup>. From 0.3–17 ka, precipitation fields in PaleoClim are obtained from the TRaCE21ka transient simulations of the last 21 kyr run with the CCSM3 model<sup>40</sup>. Although Fordham et al.<sup>39</sup> show that precipitation errors range from 10–200% in their modern experiment, the paleoclim dataset provides a statistical downscaling method that includes a bias correction (namely the Change-Factor method, in which the anomaly between the modern simulation and observations is removed from the paleoclimate experiment) allowing the use of the model for paleoclimate studies<sup>40</sup>. The very same technique is applied for 130 ka and 787 ka fields that were obtained with different GCMs (namely HadCM3 and CCSM2). Given the absence of a million-year long transient climate simulation, we oversimplified the climatic conditions by considering that precipitation distribution and intensity remain constant between two consecutive intervals, generating an artificial stepwise evolution of rainfall through time. To evaluate the sensitivity of physiographic responses on the Sunda Shelf to precipitation, we ran a model with uniform rainfall over 1 Myr (scenario 4). Despite changes in the timing and extent of basins reorganisation (Supplementary Fig. 2 and Fig. 3b), we found limited differences in terms of flooding history and erosion/deposition patterns when compared with scenario 5 (Supplementary Fig. 1c, d and Supplementary Table 2). Recent work<sup>60</sup> suggests clear regional responses induced by the emerged Sunda Shelf with seasonal enhancement of moisture convergence and continental precipitation induced by thermal properties of the land surface. This could significantly impact our simulation results. However, and at the time of writing, more continuous high-resolution paleoclimatic simulations considering the shelf as an emerged continental platform were still unavailable. Using high-resolution multi-model outputs would allow to target the uncertainty on climatic maps<sup>4</sup> and will surely represent a significant advance for future studies. One approach would have used the orographic rainfall capability<sup>61</sup> available in Badlands. The method is better suited to run generic simulations but falls short when applied to real cases as it relies on imposing paleo-environmental boundary conditions (e.g., temporal changes in wind direction and speed, moisture stability frequency or depth of moist layer) difficult to obtain for Earth-like model applied over geological time scales.

Finally, our species-agnostic approach assumes an equally weighted cost between the three considered geomorphic features and does not account for additional factors (temperature, vegetation cover, solar radiation to cite a few), which are all important when assessing landscape connectivity for wildlife. Most importantly, we model connectivity at very large scales (5 km resolution). Often, species are highly influenced by microclimates and small-scale topography<sup>47</sup>. From our regional-scale simulations and hotspot analysis (Fig. 6), higher resolution models focusing on highly connected regions (across the Gulf of Thailand and Siam basin) could be applied to produce more detailed representations of species migration in the region. In addition, current flow field calculations from Circuitscape<sup>35</sup> rely on randomly selecting nodes around the region of interest. For connectivity analysis, we used 33 terrestrial nodes located around the perimeter of the buffered Sundaland area (white contour line in Fig. 1b). Using a selection of nodes in a buffered region allows to reduce the bias in current density estimates<sup>46</sup>. However, bias might depend on the buffer size compared to the study area as well as the number of nodes selected<sup>46,47</sup>. Because of memory limitations and the great number of computed grids used to cover the past 500 kyr, we made a trade-off between buffer size and the number of selected points for pairwise calculations. Additional experiments could possibly be tested to evaluate bias in the proposed connectivity maps potentially using a tiling approach to reduce cell number<sup>45</sup>.

### Data availability

For this study, we relied on paleo-precipitation maps obtained from PaleoClim (<http://www.paleoclim.org>) at 2.5 arc minutes grids from the Pleistocene to Pliocene<sup>40</sup>. The initial topography has been modified based on the ETOPO1 Bedrock dataset (<https://www.ngdc.noaa.gov/mgg/global/>). The sea level curves are from refs. <sup>36,37</sup> and are accessible from NOAA (<https://www.ncdc.noaa.gov/paleosearch/study/11933> and <https://www.ncdc.noaa.gov/paleosearch/study/19982> studies, respectively).

### Code availability

Two main scientific software are used in this study; Badlands<sup>34</sup> and Circuitscape<sup>35</sup>. A GitHub repository containing the required data (e.g., precipitation, paleo-topography and tectonic maps as well as eustatic sea level curves) has been created and is available from <https://github.com/badlands-model/badlands-sundaland>. In addition, the repository includes seven Jupyter Notebooks which have been designed to (1) run the

landscape evolution simulations; (2) extract and visualise the information relative to catchment dynamics, erosion–deposition patterns across the Sunda shelf as well as shelf exposure evolution; (3) compute cost surfaces used in the connectivity model (from the landscape elevational connectivity index to cost associated to elevation, distance to rivers or slope values); (4) evaluate and plot statistical significance of computed connectivity (based on current flow) using z-score and Getis-Ord indices. We have also built a Docker container (geodols/gospl:bio) that includes all the libraries required to run and analyse the simulations discussed in the manuscript as well as plot most of the presented graphs. The 3D figures in the paper are plotted with ParaView (<https://www.paraview.org>) using the outputs obtained with the above notebooks. Using the Docker container allows with ease the full reproducibility of the tested scenarios.

Received: 29 March 2021; Accepted: 2 November 2021;

Published online: 22 November 2021

### References

- Molengraaff, G. A. F. & Weber, M. On the relation between the Pleistocene glacial period and the origin of the Sunda sea (Java and South China-sea), and its influence on the distribution of coral reefs and on the land- and freshwater fauna. In *van Wetenschappen Proceedings*, vol. 23, 395–439 (Koninklijke Nederlandsche Akademie, 1921).
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B. & Kent, J. Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858 (2000).
- Lohman, D. J. et al. Biogeography of the Indo-Australian Archipelago. *Annu. Rev. Ecol. Evol. Syst.* **42**, 205–226 (2011).
- Raes, N. et al. Historical distribution of Sundaland's Dipteroecarp rainforests at Quaternary glacial maxima. *Proc. Nat. Acad. Sci.* **111**, 16790–16795 (2014).
- Voris, H. K. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *J. Biogeogr.* **27**, 1153–1167 (2000).
- Leonard, J. A. et al. Phylogeography of vertebrates on the Sunda Shelf: a multi-species comparison. *J. Biogeogr.* **42**, 871–879 (2015).
- Sheldon, F. H., Lim, H. C. & Moyle, R. G. Return to the Malay Archipelago: the biogeography of Sundaic rainforest birds. *J. Ornithol.* **156**, 91–113 (2015).
- Mason, V. C., Helgen, K. M. & Murphy, W. J. Comparative phylogeography of forest-dependent mammals reveals Paleo-forest corridors throughout Sundaland. *J. Hered.* **110**, 158–172 (2018).
- Sholihah, A. et al. Impact of Pleistocene eustatic fluctuations on evolutionary dynamics in Southeast Asian biodiversity hotspots. *Syst. Biol.* <https://doi.org/10.1093/sysbio/syab006> (2021).
- Gorog, A. J., Sinaga, M. H. & Engstrom, M. D. Vicariance or dispersal? Historical biogeography of three Sunda shelf murine rodents (Maxomys surifer, Leopoldamys sabanus and Maxomys whiteheadi). *Biol. J. Linn. Soci.* **81**, 91–109 (2004).
- Sathiamurthy, E. & Voris, H. K. Maps of Holocene sea level transgression and submerged lakes on the Sunda Shelf. *Trop. Nat. Hist.* **2**, 1–44 (2006).
- Slik, J. W. F. et al. Soils on exposed Sunda Shelf shaped biogeographic patterns in the equatorial forests of Southeast Asia. *Proc. Nat. Acad. Sci.* **108**, 12343–12347 (2011).
- de Bruyn, M. et al. Paleo-drainage basin connectivity predicts evolutionary relationships across three Southeast Asian biodiversity hotspots. *Syst. Biol.* **62**, 398–410 (2013).
- van den Bergh, G. D., de Vos, J. & Sondaar, P. Y. The late Quaternary palaeogeography of mammal evolution in the Indonesian Archipelago. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **171**, 385–408 (2001).
- Lucchini, V., Meijaard, E., Diong, C. H., Groves, C. P. & Randi, E. New phylogenetic perspectives among species of South-east Asian wild pig (*Sus* sp.) based on mtDNA sequences and morphometric data. *J. Zool.* **266**, 25–35 (2005).
- den Tex, R., Thorington, R., Maldonado, J. E. & Leonard, J. A. Speciation dynamics in the SE Asian tropics: putting a time perspective on the phylogeny and biogeography of Sundaland tree squirrels, *Sundasciurus*. *Mol. Phylogenet. Evol.* **55**, 711–720 (2010).
- Sarr, A.-C. et al. Subsiding Sundaland. *Geology* **47**, 119–122 (2019).
- Husson, L., Boucher, F. C., Sarr, A.-C., Sepulchre, P. & Cahyarini, S. Y. Evidence of Sundaland's subsidence requires revisiting its biogeography. *J. Biogeogr.* **47**, 843–853 (2020).
- Quek, S.-P., Davies, S. J., Ashton, P. S., Itino, T. & Pierce, N. E. The geography of diversification in mutualistic ants: a gene's-eye view into the Neogene history of Sundaland rain forests. *Mo. Ecol.* **16**, 2045–2062 (2007).
- Hinckley, A., Hawkins, M. T. R., Achmadi, A. S., Maldonado, J. E. & Leonard, J. A. Ancient divergence driven by Geographic Isolation and ecological adaptation in forest dependent Sundaland tree squirrels. *Front. Ecol. Evol.* **8**, 208 (2020).

21. Cros, E. et al. Quaternary land bridges have not been universal conduits of gene flow. *Mol. Ecol.* **29**, 2692–2706 (2020).
22. Bird, M. I., Taylor, D. & Hunt, C. Palaeoenvironments of insular Southeast Asia during the Last Glacial Period: a savanna corridor in Sundaland? *Quat. Sci. Rev.* **24**, 2228–2242 (2005).
23. Wurster, C., Rifai, H., Zhou, B., Haig, J. & Bird, M. I. Savanna in equatorial Borneo during the late Pleistocene. *Sci. Rep.* **9**, 6392 (2019).
24. Cannon, C. H., Morley, R. J. & Bush, A. B. G. The current refugial rainforests of Sundaland are unrepresentative of their biogeographic past and highly vulnerable to disturbance. *Proc. Nat. Acad. Sci.* **106**, 11188–11193 (2009).
25. Morley, R. J. Assembly and division of the South and South-East Asian flora in relation to tectonics and climate change. *J. Trop. Ecol.* **34**, 209–234 (2018).
26. Rainboth, W. J. *The taxonomy, systematics, and zoogeography of Hypsibarbus, a new genus of large barbs (Pisces, Cyprinidae) from the rivers of southeastern Asia.* (University of California Press, 1996).
27. Heaney, L. R. A synopsis of climatic and vegetational change in Southeast Asia. *Climatic Change* **19**, <https://doi.org/10.1007/BF00142213> (1991).
28. Rahbek, C. The role of spatial scale and the perception of large-scale species-richness patterns. *Ecol. Lett.* **8**, 224–239 (2005).
29. Steinbauer, M. J. et al. Topography-driven isolation, speciation and a global increase of endemism with elevation. *Glob. Ecol. and Biogeogr.* **25**, 1097–1107 (2016).
30. Lyons, N. J., Val, P., Albert, J. S., Willenbring, J. K. & Gasparini, N. M. Topographic controls on divide migration, stream capture, and diversification in riverine life. *Earth Surf. Dyn.* **8**, 893–912 (2020).
31. Lomolino, M. V. Elevation gradients of species-density: historical and prospective views. *Glob. Eco. Biogeogr.* **10**, 3–13 (2001).
32. Gillespie, R. G. & Roderick, G. K. Geology and climate drive diversification. *Nature* **509**, 297 (2014).
33. Bertuzzo, E. et al. Geomorphic controls on elevational gradients of species richness. *Proc. Nat. Acad. Sci.* **113**, 1737–1742 (2016).
34. Salles, T. et al. A unified framework for modelling sediment fate from source to sink and its interactions with reef systems over geological times. *Sci. Rep.* **8**, 5252 (2018).
35. McRae, B. H., Dickson, B. G., Keitt, T. H. & Shah, V. B. Using circuit theory to model connectivity in ecology, evolution, and conservation. *J. Ecol.* **89**, 2712–2724 (2008).
36. Bintanja, R. & van de Wal, R. North American ice-sheet dynamics and the onset of 100,000-year glacial cycles. *Nature* **454**, 869–872 (2008).
37. Spratt, R. M. & Lisiecki, L. E. A late Pleistocene sea level stack. *Clim. Past* **12**, 1079–1092 (2016).
38. Otto-Bliesner, B. L. et al. Simulating Arctic climate warmth and icefield retreat in the last interglaciation. *Science* **311**, 1751–1753 (2006).
39. Fordham, D. A. et al. PaleoView: a tool for generating continuous climate projections spanning the last 21,000 years at regional and global scales. *Ecography* **40**, 1348–1358 (2017).
40. Brown, J., Hill, D., Dolan, A., Carnaval, A. C. & Haywood, A. M. PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. *Sci. Data* **5**, 180254 (2018).
41. Alqahtani, F. A., Johnson, H. D., Jackson, C. A.-L. & Som, M. R. B. Nature, origin and evolution of a late Pleistocene incised valley-fill, Sunda Shelf, Southeast Asia. *Sedimentology* **62**, 1198–1232 (2015).
42. Miall, A. D. Architecture and sequence stratigraphy of Pleistocene fluvial systems in the Malay Basin, based on seismic time-slice analysis. *AAPG Bulletin* **86**, 1201–1216 (2002).
43. Wong, H. K. et al. In *Tropical Deltas of Southeast Asia-Sedimentology, Stratigraphy, and Petroleum Geology* Vol. 76 <https://doi.org/10.2110/pec.03.76.0201> (SEPM Society for Sedimentary Geology, 2003).
44. Darmadi, Y., Willis, B. & Dorobek, S. Three-dimensional seismic architecture of fluvial sequences on the low-gradient Sunda Shelf, offshore Indonesia. *J. Sediment. Res.* **77**, 225–238 (2007).
45. Pelletier, D. et al. Applying circuit theory for corridor expansion and management at regional scales: tiling, pinch points, and omnidirectional connectivity. *PLoS One* **9**, 1–11 (2014).
46. Koen, E. L., Bowman, J., Sadowski, C. & Walpole, A. A. Landscape connectivity for wildlife: development and validation of multispecies linkage maps. *Methods Ecol. Evol.* **5**, 626–633 (2014).
47. Marrec, R. et al. Conceptual framework and uncertainty analysis for large-scale, species-agnostic modelling of landscape connectivity across Alberta, Canada. *Sci. Rep.* **10**, <https://doi.org/10.1038/s41598-020-63545-z> (2020).
48. Inger, R. F. & Voris, H. K. The biogeographical relations of the frogs and snakes of Sundaland. *J. Biogeogr.* **28**, 863–891 (2001).
49. Rinaldo, A., Gatto, M. & Rodriguez-Iturbe, I. River networks as ecological corridors: a coherent ecohydrological perspective. *Adv. Water Resour.* **112**, 27–58 (2018).
50. Gascon, C. et al. Riverine barriers and the geographic distribution of Amazonian species. *Proc. Nat. Acad. Sci.* **97**, 13672–13677 (2000).
51. Naka, L. N. & Brumfield, R. T. The dual role of Amazonian rivers in the generation and maintenance of avian diversity. *Sci. Adv.* **4**, <https://doi.org/10.1126/sciadv.aar8575> (2018).
52. Salles, T., Rey, P. & Bertuzzo, E. Mapping landscape connectivity as a driver of species richness under tectonic and climatic forcing. *Earth Surf. Dyn.* **7**, 895–910 (2019).
53. Woodruff, D. S. Biogeography and conservation in Southeast Asia: how 2.7 million years of repeated environmental fluctuations affect today's patterns and the future of the remaining refugial-phase biodiversity. *Biodivers. Conserv.* **19**, 919–941 (2010).
54. Braun, J. & Willett, S. D. A very efficient O(n), implicit and parallel method to solve the stream power equation governing fluvial incision and landscape evolution. *Geomorphology* **180–181**, 170–179 (2013).
55. Menier, D. et al. Landscape response to progressive tectonic and climatic forcing in NW Borneo: implications for geological and geomorphic controls on flood hazard. *Sc. Rep.* **7**, 2045–2322 (2017).
56. Walt, S. et al. scikit-image: image processing in python. *PeerJ* **2**, e453 (2014).
57. Getis, A. & Ord, J. Local spatial statistics: an overview. *Health* **3**, 261–277 (1996).
58. Campforts, B., Schwanghart, W. & Govers, G. Accurate simulation of transient landscape evolution by eliminating numerical diffusion: the TTLEM 1.0 model. *Earth Surf. Dyn.* **5**, 47–66 (2017).
59. Barnes, R. Accelerating a fluvial incision and landscape evolution model with parallelism. *Geomorphology* **330**, 28–39 (2019).
60. Sarr, A.-C., Sepulchre, P. & Husson, L. Impact of the Sunda Shelf on the climate of the maritime continent. *J. Geophys. Res. Atmos.* **124**, 2574–2588 (2019).
61. Smith, R. B. & Barstad, I. A linear theory of orographic precipitation. *J. Atmos. Sci.* **61**, 1377–1391 (2004).
62. Luo, S.-J. et al. Phylogeography and genetic ancestry of tigers (*Panthera tigris*). *PLoS Biol.* **2**, <https://doi.org/10.1371/journal.pbio.0020442> (2004).
63. Wilting, A. et al. Clouded leopard phylogeny revisited: support for species recognition and population division between Borneo and Sumatra. *Front. Zool.* **4**, <https://doi.org/10.1186/1742-9994-4-15> (2007).
64. Wilting, A. et al. Evolutionary history and conservation significance of the Javan leopard *Panthera pardus melas*. *J. Zool.* **299**, 239–250 (2016).
65. Patel, R. P. et al. Genetic structure and phylogeography of the leopard cat (*Prionailurus bengalensis*) Inferred from mitochondrial genomes. *J. Hered.* **108**, 349–360 (2017).
66. Moyle, R. G., Schilthuizen, M., Rahman, M. A. & Sheldon, F. H. Molecular phylogenetic analysis of the white-crowned forktail *Enicurus leschenaulti* in Borneo. *J. Avian Biol.* **36**, 96–101 (2005).
67. Sheldon, F. H. et al. Phylogeography of the magpie-robin species complex (Aves: Turdidae: Copsychus) reveals a Philippine species, an interesting isolating barrier and unusual dispersal patterns in the Indian Ocean and Southeast Asia. *J. Biogeogr.* **36**, 1070–1083 (2009).
68. den Tex, R.-J. & Leonard, J. A. A molecular phylogeny of Asian barbets: speciation and extinction in the tropics. *Mol. Phylogenet. Evol.* **68**, 1–13 (2013).
69. Steiner, C. C., Houck, M. & Ryder, O. A. G. Genetic variation of complete mitochondrial genome sequences of the Sumatran rhinoceros (*Dicerorhinus sumatrensis*). *Conserv. Genet.* **19**, 397–408 (2018).
70. Ohtani, M. et al. Nuclear and chloroplast DNA phylogeography reveals Pleistocene divergence and subsequent secondary contact of two genetic lineages of the tropical rainforest tree species *Shorea leprosula* (Dipterocarpaceae) in South-East Asia. *Mol. Ecol.* **22**, 2264–2279 (2013).
71. Achmadi, A. S., Esselstyn, J. A., Rowe, K. C., Maryanto, I. & Abdullah, M. T. Phylogeny, diversity, and biogeography of Southeast Asian spiny rats (Maxomys). *J. Mammal.* **94**, 1412–1423 (2013).

### Acknowledgements

We would like to thank the three anonymous reviewers for their valuable comments and suggestions on the drafts of this article. The presented study was supported by the ARC DARE Centre grant IC190100031 and the National Geographic Society grant NGS-50778R-18. This research was undertaken with the assistance of resources from the National Computational Infrastructure (NCI), which is supported by the Australian Government and from the Artemis HPC Grand Challenge supported by the University of Sydney.

### Author contributions

All authors conceived the experiments, T.S. and C.M. conducted the numerical experiments. T.S., C.M., L.H., S.Z., A.-C.S. and P.S. analysed the results, wrote and reviewed the manuscript.

### Competing interests

The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s43247-021-00311-7>.

**Correspondence** and requests for materials should be addressed to Tristan Salles.

**Peer review information** *Communications Earth & Environment* thanks Haw Chuan Lim, Richard Barnes and the other anonymous reviewer(s) for their contribution to the peer review of this work. Primary handling editor: Joe Aslin. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021