SEABIRD RANGE CONTRACTION AND DISPERSAL UNDER CLIMATE CHANGE

Jorge Avaria-Llautureo[1,2*], Marcelo M. Rivadeneira[2,3,4], Chris Venditti[1], Guillermo Luna-Jorquera[2,3,4*]

1. School of Biological Sciences, University of Reading, UK.
2. Center for Ecology and Sustainable Management of Oceanic Islands (ESMOI).
3. Departamento de Biología Marina, Universidad Católica del Norte, Coquimbo, Chile.
4. Centro de Estudios Avanzados en Zonas Áridas (CEAZA), Universidad Católica del Norte, Coquimbo, Chile.

*Correspondence: Jorge Avaria-Llautureo (j.l.avaria@reading.ac.uk); Guillermo Luna-Jorquera (gluna@ucn.cl)

Abstract

Many marine ectotherms have historically adapted to local climate change by evolving smaller body sizes, reducing their energy demands in warmer waters but limiting their dispersal and speciation rate. Whether endothermic marine species respond similarly remains unclear, as temperature minimally affects their size diversity, and the drivers of their dispersal and speciation are poorly understood. Here we show that globally distributed seabirds (albatrosses, petrels, shearwaters, and storm petrels), facing rapid historical climate change, responded by shifts in geographic range size rather than body mass. Additionally, where the rate of warming is high, geographic ranges contracts most, intensifying the selective forces for higher dispersal capacity, and increasing speciation rate. Our findings reveal a triple threat to extant seabirds from human-induced global warming: shrinking ranges increase their extinction risk, push them to their maximum limit of dispersal capacity, and subject them to unprecedented warming rates. These insights reveal a distinct endothermic response to ocean warming, underscoring the vulnerability of seabirds and the urgent need to integrate range dynamics into conservation strategies for marine biodiversity under accelerating global change.

Introduction

Seabirds play a key role in supporting ecosystems services related to human economies[1] but they are one of the most threatened marine taxonomic groups, largely owing to indirect climate-mediated changes in the distribution and abundance of their prey[2]. Although numerous hypotheses predict direct effects of current climate change on seabirds' morphology[3], physiology[4], and geographic distribution[5], our ability to test these hypotheses is limited by lack of data[6,7] and suitable methodological approaches for modelling it. Seabirds are highly sensitive to small-scale environmental factors, such as local temperature (LT)[2] and environmental unpredictability[8]. Therefore, understanding how seabirds have adapted to long-term local warming is crucial for their conservation and the sustainable management of resources in the context of human-induced global warming.

Recent research indicates that marine ectothermic species, such as fish, have evolved smaller body sizes in response to historical local warming, reducing their dispersal and speciation capabilities[9]. However, it remains unclear whether this evolutionary trend applies universally across all marine biodiversity. Extrapolating findings from ectothermic species to endothermic ones, like seabirds, is challenging because they respond differently to temperature changes[4,10,11]. Ectothermic metabolic rate scales directly with LT[12,13]; for example, a tropical fish at 30°C requires six times more oxygen for resting metabolism than a polar fish at 0°C[12]. Warmer conditions thus increase energetic demands, favouring smaller body sizes that reduce total energy expenditure[14]. Conversely, endotherms maintain stable metabolic rates across a wide LT range (the thermoneutral zone) with increases only outside this range [4,10,13], which is a mass-independent adaptation to thermal environment across all birds [11,15,16]. Additionally, LT accounts for only ~6% of mass variation in seabirds (Extended Data Fig. 1), thus we should expect adaptive shifts in seabirds' physiology instead of body mass in face of rapid local warming.

However, seabirds (Procellariiformes) are already operating near their physiological limits, under a narrow range of hot temperatures for optimal performance, and limited water availability[4,10]. Thus, physiological shifts may not be a viable adaptation for coping with future warming. Additionally, given that physiological tolerance and capacities correlates positively with geographic range size[5], seabirds may face an imminent risk of extinction owing to range contractions mediated by rapid warming. Finally, the dispersal and speciation capacity of seabirds should be unaffected by warming if these variables are linked to their body mass[17–19].

To evaluate how seabirds have adapted to local climate change we developed a novel geographic-phylogenetic approach allowing us to reconstruct seabirds' ancestral species locations, their dispersal routes, and their LT though time and across the world's palaeo-oceans. For this, we reconstructed ancestral locations at phylogenetic nodes, using a posterior sample of the most complete phylogeny of extant procellariforms to date[20] to account for phylogenetic uncertainty (Methods; Extended Data Fig. 2). We used a new version of the Geographical (Geo) model in BayesTraits [21–23] which allows variation in species dispersal ability as well as continental drift by the inclusion of palaeo-map restrictions when inferring ancestral locations[23]. We explicitly identified the palaeo-coordinates of the places where seabirds spend most of their lives, i.e., the oceans[24] (Extended Data Fig. 3). Although seabirds return to land to breed, they spend long periods at sea and are therefore effectively pelagic organisms[24]. Thus, we assigned a zero-prior probability of ancestral seabirds being present across palaeo-continents, providing a more realistic recreation of ancestral locations though the palaeo-oceans (restricting ancestral location to continents had a worst fit to the data, see Supplementary Information). Using these reconstructions, we can infer the geographic path and distance that each species travelled over the oceans, from the root of the phylogenetic tree to its present location (Pathwise distance; Methods). Finally, we obtained the palaeo-temperature from the HadCM3 palaeo-climate model for the reconstructed locations at phylogenetic nodes (Methods; Extended Data Fig. 2). This allowed us to evaluate the expected seabirds' responses to historical local climate change in a far more nuanced fashion than has ever been possible previously.

Results

Seabirds originated in temperate and tropical temperatures. We first carried out the Geo model analysis with map restrictions using the maximum clade credibility (MCC) tree of Procellariiformes (Methods). The Geo model with variable rates fit the data better than the constant rate model (Bayes Factor (BF) > 10, very strong support[25]). This result implies that the historical geographic expansion of seabirds' diversity was shaped by species dispersing at variable dispersal speeds (distance per unit on time). Our model indicates that the most recent common ancestor (MRCA) of procellariforms originated in the area encompassing the present-day Coral Sea and the ancient, submerged, Zealandia continent [26], in the Paleocene (Fig. 1a, white filled dots). This suggests that the MRCA likely breed and nested along the coast of East Australia and nearby islands (Fig. 1). The LT, extracted from the HadCM3 palaeo-climate model and using the MRCA posterior distribution of geographic coordinates (Methods), ranges from 7.9 to 29.5°C (Fig. 1b), which encompass temperate and tropical temperatures. Finally, we carried out the Geo model analyses on each of the 500 phylogenetic trees, and we found consistent results for the MRCA location as well as for the main families (Extended Data Fig. 4). For example, in over 80% of the 500 analysed trees, we recover the procellariform MRCA in what is now the Coral Sea, east of Australia (Extended Data. Fig. 4a).

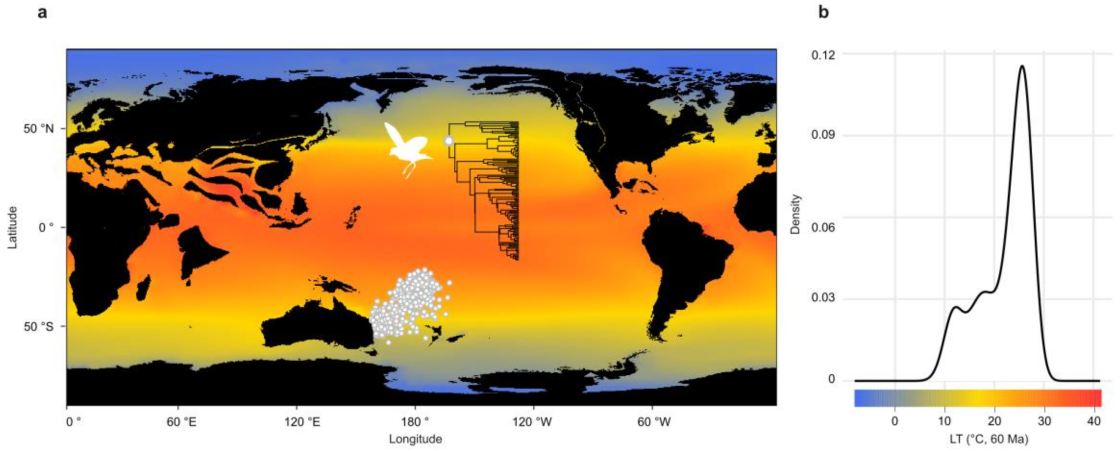


Figure 1. Procellariiformes palaeo-location and local palaeo-temperature. a, white filled dots indicate the posterior distribution of geographic coordinates for the MRCA (white dot in the tree), which were obtained from the Geo model analysis with palaeo-map restriction. The MRCA was distributed in what is now the Coral Sea, east of Australia. **b,** sea surface palaeo-temperature extracted from the posterior distribution of coordinates estimated with the Geo model. These temperature values represent the local sea surface temperature (LT) for the MRCA. The plot indicates that the local environments of the MRCA was included temperate and hot LT (n = 500 coordinates).

Body mass is decoupled from temperature but positively associated with range size. The following results are based on a Bayesian PGLS multiple regression with variable rate (best fitted model, Supplementary Table 1) and using the MCC tree (Methods). As main explanatory variables for body mass, we included LT (using samples of LT data obtained from the entire geographic distribution of each species, see Methods), range size, and an interaction term between range size and LT (Methods). We also included the rate of local climate change ($\mathrm{LT}_{\mathrm{RATE}}$)[23] as additional covariate in our phylogenetic regressions. The $\mathrm{LT}_{\mathrm{RATE}}$ is the cumulative change of LT across the phylogenetic branches linking the common ancestor of seabirds with every extant species (Methods), divided by time. Crucially, these changes are not directional, i.e., changes can be to either cooler or warmer local conditions. We used samples of 500 $\mathrm{LT}_{\mathrm{RATE}}$ data per species as obtained from the posterior distribution of coordinates inferred with the Geo model (see Methods).

Results show that only range size had a positive and significant effect on body mass (mean slope = 0.06; $pMCMC_{99.6} > 95$, Fig. 2), indicating that geographically widespread species tends to be larger in size. LT had a non-significant effect on body mass ($pMCMC_{85} < 95$), revealing that body mass is decoupled from LT. Finally, we found that range size had a positive and significant effect on body mass in over 90% of the regressions analyses carried out on each of the 500 phylogenetic trees.

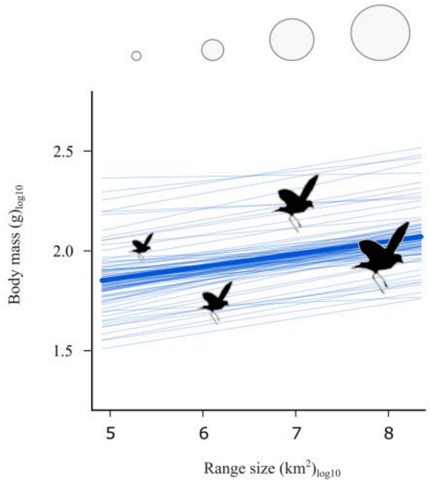


Figure 2. Bigger seabirds in the order Procellariiformes are geographically more widespread. The plot shows the predicted body mass by range size. Light-blue lines indicate the posterior distribution of the regression coefficient linking body mass with range size. The dark-blue line indicates the coefficient posterior mean.

Range size gets smaller under rapid local warming and correlates positively with body mass. The following results were obtained from a Bayesian PGLS multiple regression with variable rate (best fitted model, Supplementary Table 1) and using the MCC tree. Main explanatory variables for range size included LT, $\mathrm{LT}_{\mathrm{RATE}}$, body mass, the Pathwise distance, and absolute latitude. We also included an interaction term between Pathwise distance and $\mathrm{LT}_{\mathrm{RATE}}$.

Our results show that only LT, $\mathrm{LT}_{\mathrm{RATE},}$ and body mass, had a significant effect on range size. Warmer LT and higher $\mathrm{LT}_{\mathrm{RATE}}$ values were associated with smaller range size (LT mean slope = -0.81, $pMCMC_{100} > 95$; $\mathrm{LT}_{\mathrm{RATE}}$ mean slope = -1.65, $pMCMC_{100} > 95$; Fig. 3a, b). Additionally, body mass was positively associated with range size (mean slope = 0.38, $pMCMC_{100} > 95$, Fig. 3c). These results suggest rapid local warming reduce seabirds' range size, likely owing to their limited adaptive responses to local climate change, as supported by previous studies[4,10]. Furthermore, since range size correlates positively with body mass, rapid warming may indirectly select for smaller seabirds by contracting their geographic range.

The variable rate PGLS regression model that considers LT, LTRATE, and body mass, explained 67% (mean R^2) of the variance in range size. When exploring the effect size of each independent variable, we found that LT, LTRATE, and body mass, explained 1, 35, and 2% of the variance, respectively (after accounting for shared ancestry). Finally, body mass had a significant and positive effect on range size in 79% of the regression analyses carried out on each of the 500 phylogenetic trees. LT had a negative and significant effect in 100% of the analyses while LTRATE had a negative and significant effect in 97% of the analyses.

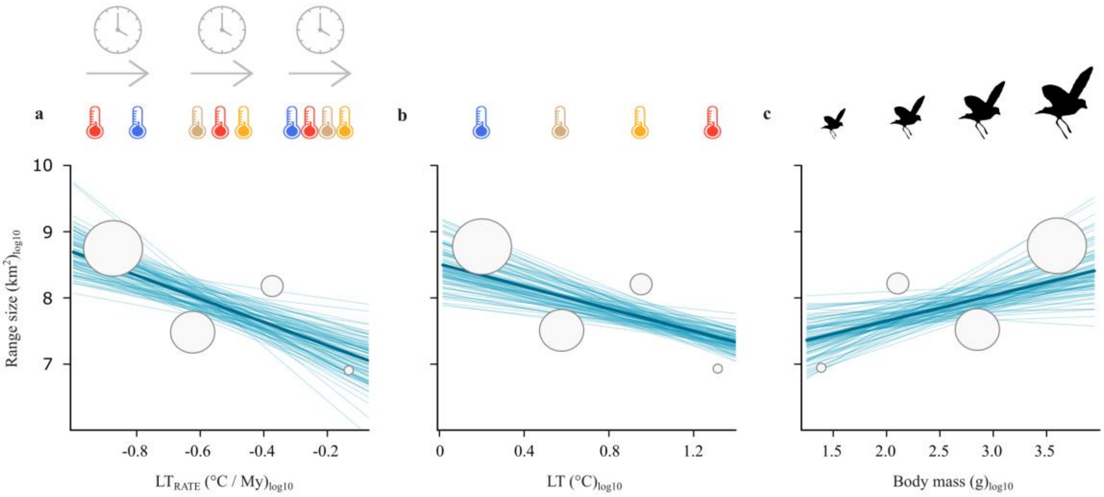


Figure 3. Procellariiformes' range size gets smaller under rapid warming and correlates positively with body mass. a, plot of the predicted range size by LTRATE, for a mean LT and mean body mass. **b,** plot of the predicted range size by LT, for a mean LTRATE and mean body mass. **c,** plot of the predicted range size by body mass, for a mean LT and mean LTRATE. Light-blue lines indicate the posterior distribution of the regression coefficients, and the dark-blue lines indicate the posterior mean.

To obtain a general view of range size evolution in face of local climate change, we predicted the ancestral range size by integrating the additive effect of LT, LTRATE, phylogeny, and the ancestral location inferred with the Geo model (Methods). This approach brings a simplified yet robust reconstruction of seabirds' geographic expansion in hace of historical climate change (Fig. 4). Originating from a common ancestor with a range spanning approximately 125 million km^2 (Fig. 4a, b), seabirds predominantly dispersed and diversified across the Pacific Ocean (Fig. 4b–e). The diversification of species' range sizes with a general tendency to smaller ranges was driven by the combined influence of temperature and its rate of change, with these temperature dynamics themselves shaped by species dispersal over evolutionary time.

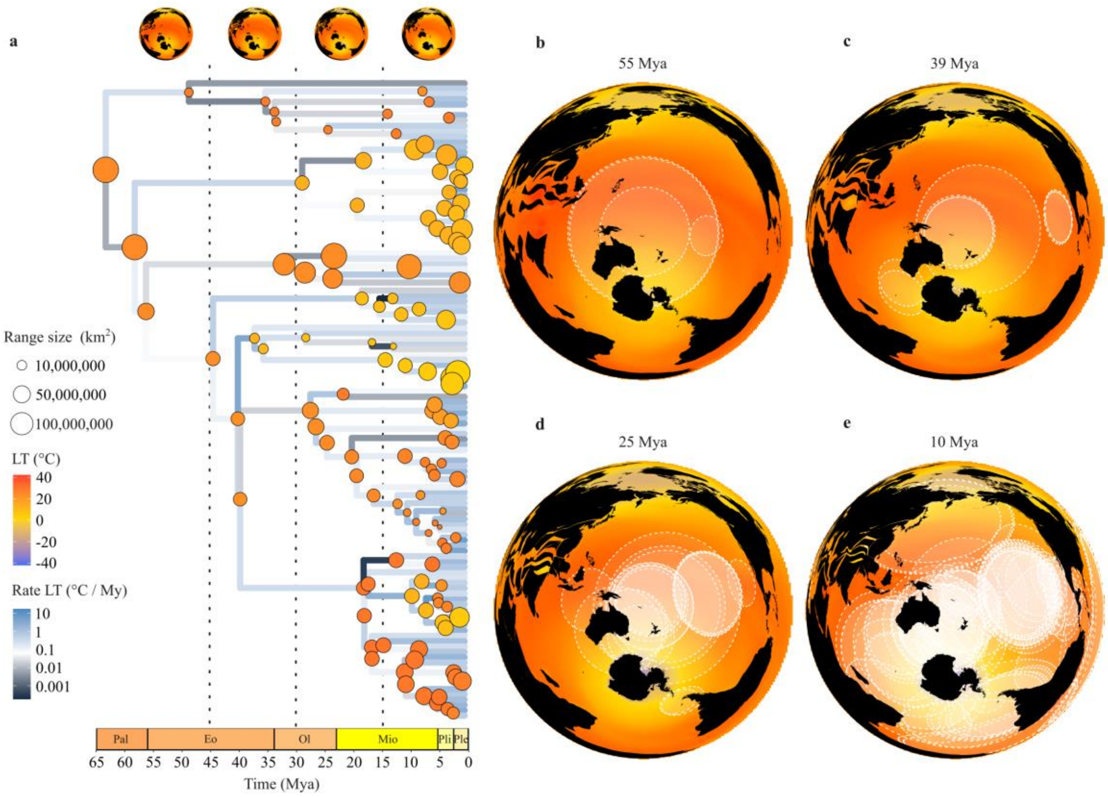


Figure 4. Phylogenetic range size prediction for Procellariiformes based on the additive effect of LT and LT_{RATE}. a, MCC tree with palaeo-LT and estimated range size at phylogenetic nodes. Coloured branches indicate the LT rate per million years. The vertical dashed lines split the predictions on four-time intervals. **b, c, d, e,** predicted geographic range size projected on world palaeo-maps using the Lambert azimuthal equal-area projection. For representation only, we show the temperature and continental configuration of mid-time intervals. Each geographic range was centred on the geographic centroid that was obtained from the Geo model posterior locations. For simplicity, we projected the distribution range as a circle on the sphere. Circles illustrate the range size only, not its shape. By including range shape, all the modelled ranges would border some land.

Seabirds dispersed longer distances when their geographic range got smaller under higher rates of local climate change. We carried out a PGLS multiple regression, based on the lambda model (best fitted model, Supplementary Table 1), to test the expected correlates of seabirds' Pathwise distance (Methods). As main explanatory variables, we included body mass, range size, and LT_{RATE}. Additionally, we included the interaction between LT_{RATE} and range size as explanatory variable.

We found that range size (mean slope = -0.12; $pMCMC_{98.5} > 95$), LT_{RATE} (mean slope = 3.24; $pMCMC_{100} > 95$), and the interaction between them (mean slope = -0.35; $pMCMC_{99.9} > 95$), had a significant effect on Pathwise distance (Fig. 5). The negative interaction between LT_{RATE} and range size in predicting dispersal distance suggests that as species are more geographically restricted, they are more sensitive to the rate of local climate. More rapid climate change increases the selective pressure on dispersal to track suitable habitat elsewhere, which agrees with observations from individual-based model simulations[27].

The best fitting PGLS model explained 59% of the variance in Pathwise distance. Regarding the effect size of each variable, LT_{RATE} explained 21% of the variance in Pathwise distance, while range size explained only 1%. However, the interaction between range size and LT_{RATE} explained 8% of the variance in Pathwise distance (after accounting for shared ancestry). Finally, we found that the interaction term between range size and LT_{RATE} was negative and significant in 85% of the regression analyses carried out on each of the 500 phylogenetic trees.

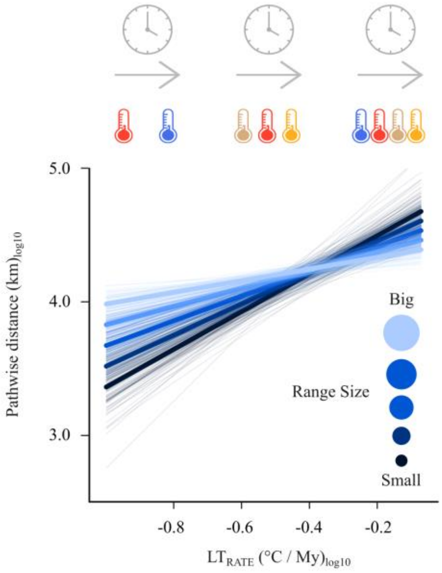


Figure 5. Procellariiformes dispersed longer distances when their geographic range contracted under rapid local climate change. Interaction plot of Pathwise distance dispersed under variable rate of local climate change at five different range size values. Ligh transparent lines indicate the posterior distribution of the regression coefficient linking Pathwise distance with LT_{RATE}. The dark lines indicate the coefficient posterior mean. Circles and lines colours represent the five different range size categories.

Speciation rate increased under rapid climate change and for larger species. We ran a Bayesian PGLS multiple regression based on the Brownian motion model (best fitted model, Supplementary Table 1), to test the correlates of speciation rate. We included range size, body mass, Pathwise distance, LT, and the LT_{RATE} as predictor variables, while using sampling of trait data (Methods).

After reducing our model to have only significant predictors, we found that LT_{RATE} affected positively speciation rate (mean slope = 0.04; $pMCMC_{97.6} > 95$; Fig. 6a) while body mass had similar effect (mean slope = 0.05; $pMCMC_{99.4} > 95$; Fig. 6b). The mean R^2 of this regression model was 8.9%. The effect size of LT_{RATE} was ~5.6% while the effect size of body mass was ~3.4%. Finally, we found that LT_{RATE} had a positive and significant effect in 89% of the regressions carried out on each of the 500 phylogenetic trees. However, body mass was positive and significant in 53% of the regressions only, leaving large degrees of uncertainty about the effects of body mass.

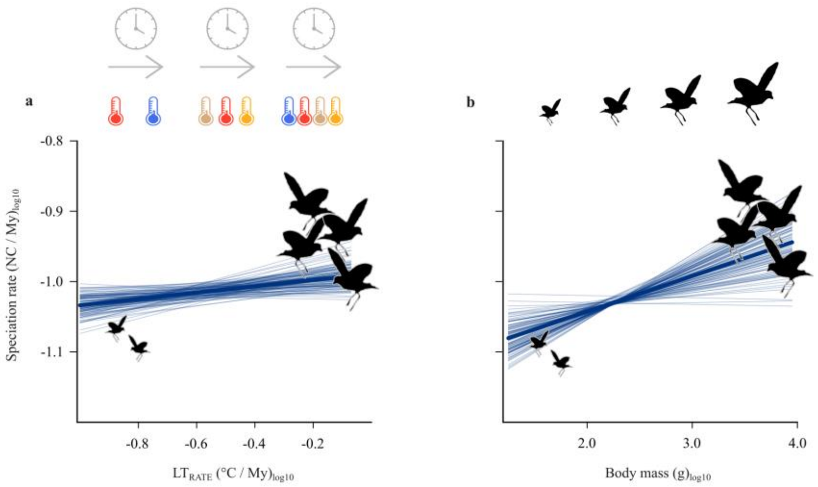


Figure 6. Rapid climate change and bigger body mass increased seabirds' speciation rate. a, plot of the predicted speciation rate under variable rate of local climate change, for a mean body mass. Light transparent lines indicate the posterior distribution of the regression coefficient linking speciation rate with LT_{RATE}. Dark line represents the coefficient posterior mean. **b,** plot of the predicted speciation by body mass, for a mean LT_{RATE}. Light transparent lines indicate the posterior distribution of the regression coefficient linking speciation rate with body mass. Dark line represents the coefficient posterior mean.

Discussion

This study demonstrates that seabirds in the order Procellariiformes adapted to rapid local climate change by evolving smaller geographic ranges, dispersing longer distances, and diverging into new species. This is distinct from what has been observed in marine ectotherms like fish. However, as in fish[9] and other studies investigating terrestrial species[28], our study show that seabirds are facing unprecedented rates of climate change in their evolutionarily history. The estimated LT_{RATE}, across the phylogenetic branches of the sample of 500 phylogenetic trees, shows that seabirds have historical adapted to a median rate of 2.2°C/Myr (ranging from 0 to 53°C/Myr). This equals to 2.2E-5 (°C/decade) which is ~4 orders of magnitude slower than the current 0.13 (°C/decade) warming rates[29].

Our results also highlight a process where rapid local climate change is pushing seabirds closer to their maximum dispersal capacity – small ranged species that are facing rapid climate change moved the longest distances (Fig. 5). Rapid climate change increased seabird speciation rates (Fig. 6a), suggesting that future global warming might create new opportunities for species origination. However, this might not help if extinction rates are higher, outpacing speciation, owing to range size contraction.

Our study emphasizes the idea that the rate of local climate change is far more relevant than its direction in driving species evolution and, therefore, it is critical to predict future responses under human-induced climate change. We show that LT_{RATE} had the greatest effect size in the Bayesian PGLS regressions predicting range size, dispersal distance, and speciation rate. This implies that how fast the local temperature change is more important than whether the local climate becomes warmer or cooler over time and space, lending support to studies of climate change velocity[30], and to predictions from the evolutionary theory[31].

We demonstrate that the historical rapid local climate change consistently contracted seabirds' geographic ranges, with surviving species likely dispersing to track suitable habitats[27,32].

However, current rates of climate change may outpace their ability to disperse and adapt. Conservation efforts should prioritise highly restricted species, such as the Mascarene, Beck's, and Fiji petrels, and areas like the Southeast Pacific, which experienced high warming and cooling rates in 2023[33].

Methods

There are several methodological limitations hindering explicit testing of the effect of climate change on seabird's diversity. First, empirical data on local temperature change across the million years of seabird evolutionary history and their global distribution are lacking. Second, dispersal capacity data is scarce: natal and adult dispersal records are available for just three seabirds in the order procellariiforms[6], while latitudinal shift data cover only 19 species of this order[7]. Indirect proxies like the Hand-Wing Index[34,35] is available for most extant seabirds, and it offers some insight into dispersal ability of extant species, but its relationship with species dispersal in deep evolutionary scales remain uncertain. The following methodology addresses all these concerns.

The order Procellariiformes as a study model. The seabird species of the order Procellariiformes have a cosmopolitan distribution across all the world's oceans. This is advantageous for testing our hypothesis because climate change, in terms of direction and rates, varies widely on the global scale[36]. Almost all Procellariiformes are conspicuous and wide-ranging birds, and as top marine predators, they integrate information from the bottom of the food web and thus respond to changes in sea surface temperature[37,38], providing insight into marine ecosystem processes. Procellariiformes is a clade that emerged from a common ancestor in the mid-Eocene (i.e., 40-45 Ma, sensu[39]) or early-Eocene (i.e., ca. 63 Ma, sensu[20]). The phylogeny of this order is in permanent revision, and the taxonomic status of most of the species is well known (e.g.[40]), which is a requirement for understanding evolutionary processes and biogeographical patterns of diversification. The extant Procellariiformes are estimated to be 148 species[41], which is roughly 44% of all extant seabirds totalled ca. 335 species. Thus, Procellariiformes represents a high proportion of the total number of seabirds, and the number of species included is high enough to ensure that the output of our models is robust. The order is grouped into four families: the albatrosses (Diomedeidae), the petrels and shearwaters (Procellariidae), the northern storm-petrels (Hydrobatidae), and the southern storm-petrels (Oceanitidae)[41]. They share life-history traits, such as being central place foragers, colonial breeders, highly philopatric, monogamist, and are long-lived species. They are highly diverse in morphological traits; for example, adult body mass ranges between 20 g in the Least storm-petrel to 6 kg in the Wandering Albatross[42].

Procellariiformes have diversified and expanded geographically across all the world oceans. They switch from their breeding habitat on land to their foraging habitat at sea, relying on marine resources like planktonic, invertebrates, and fish preys [42]. Although they return to land to breed, the Procellariiformes spend long periods at sea and are therefore effectively pelagic organisms[24]. Thus, seabirds are globally essential drivers of nutrient cycling, transferring nutrients from their pelagic feeding grounds to islands on which they roost and breed [43]. As top marine predators, they react to changes in ocean temperature, ocean productivity, and the ecology of their prey (e.g., squid and fish)[44]. The main threat to seabirds' diversity is climate change (along with invasive alien species and bycatch), being the most threatened species globally [45,46].

Phylogenetic trees. We obtained a Bayesian sample of 500 phylogenetic trees for 121 species of Procellariiformes (Hackett source) from BirdTree[20]. We excluded the Jamaican Petrel

(*Pterodroma caribbaea*) because its IUCN status is "*Possibly Extinct*". We also obtained the maximum clade credibility (MCC) tree using the *maxCladeCred* function of the *phangorn* R-package version 2.12.1[47]. The phylogenetic uncertainty associated to divergence time (branch lengths) and topology can be found graphically as a DensiTree (Extended Data Fig. 2).

Species body mass and geographic range size. We obtained species body mass (in grams) from the AVONET database[35]. We also obtained the species range size from the species geographic distribution polygons, available in the Bird species distribution maps of the world[41]. We obtained the polygons database on February 8th, 2024. To get the polygons range size, we obtained their areas (in square kilometres), using the *area* function of the *raster* R-package version 3.6-30[48]. Most of the estimated range area was identical to the range area data available in the AVONET data base. However, there were ~10 species for which the AVONET data showed over and under estimation of the range area, probably due to the use of an older version of the polygon data base.

Species geographic distribution data. We generated geographic coordinate data (longitude and latitude) within the geographic polygon of each species in the phylogenetic tree. This dataset was used to infer the posterior distribution of ancestral coordinates at phylogenetic nodes and the species dispersal ability (see the *Species dispersal ability* section below). We obtained those coordinates from the species geographic distribution polygons available in the Bird species distribution maps of the world[41]. For this, we generated random samples of coordinates within each geographic polygon. This approach allowed us to get a more exhaustive representation of the extent of the geographic distribution for each species in comparison to approaches that relies on the observed geographic occurrences or the distribution centroids. We considered the polygons geographic area (in square kilometres) to define the number of random coordinates to be generated within them. Specifically, we generated random coordinates (50, 100, 200, 300, 400, and 500) for different ranges of area in km^2 (20 to 100,000; 100,000 to 200,000; 200,000 to 300,000; 300,000 to 400,000; 400,000 to 500,000; and > 500,000). All the coordinates generated at random within each species polygon are plotted as the Extended Data Fig. 3.

Species dispersal ability (Pathwise distance). To get a measure of species dispersal ability, we employed a novel, advance phylogenetic geographical method, the Geo model with map restrictions in BayesTraits[21,23]. The original version of the Geo model[21] reconstructs the posterior distribution of longitude and latitude onto a three-dimensional Cartesian coordinates system (x, y, and z), so that the model assumes that species change their locations as a geographic point moving in a spherical space. It estimates the posterior distribution of ancestral coordinates at phylogenetic nodes while sampling across all the geographic coordinates for phylogenetic tips (the coordinates generated at random within the distribution polygons in this case) – thus considering the variance in the species' geographic distribution. The sampling of coordinates data is simultaneously integrated with the estimation of model parameters. Coordinate changes along the branches of the phylogenetic tree – i.e., species dispersal - are modelled using Brownian motion, which assumes that species (geographic points) disperse across the globe at a constant speed (distance per time unit). However, the Geo model can also estimate ancestral locations while considering the continuous variation in species dispersal speed across phylogenetic branches. The speed of movement ranges from species quiescence or no movement per time unit, through constant movement in direct proportion to the passage of time (Brownian motion), to long-distance dispersal per time unit. Estimations of species dispersal speed across phylogenetic branches are based on the variable rate model [49]. For the Geo model, the variable rate model detects shifts away from a (background) constant speed, expected under Brownian motion.

We compared the constant and variable speed models by means of Bayes Factors (BF). The BF takes the model marginal likelihood for comparison which is estimated by the steppingstone sampling in BayesTraits. The steppingstone sampling estimated the model marginal likelihood while considering the number of parameters of the model (i.e., model complexity) [50]. The BF is calculated as the double of the difference between the log marginal likelihood of two models. Higher values of the log marginal likelihood represent better-fitted models. By convention, BF > 2 indicates positive support, BF = 5–10 indicates strong support, and BF > 10 is considered very strong support for a model over the other [25].

For this study, we used an updated version of the Geo model[23], which restricts reconstructed locations to points found only on lands and that recovers simulated data with high accuracy[23]. This is an extension to the model developed by O'Donovan et al. [21]. In this study, we restrict reconstructed locations to points found only on the world’s oceans. Initially, reconstructed locations are placed on oceans, when proposing a new location, the closest point to the proposed location is identified on the map. If the closest point is found to be on the land, the new location is assigned as zero probability (rejected), otherwise it is accepted or rejected based on its likelihood. Geography is not static through time, palaeo-maps were created for different time periods. As the phylogeny is time-calibrated each internal node is assigned a palaeo-map based on its age. To restrict the space for ancestral location inferences across phylogenetic nodes, we used the global maps of the PALEOMAP project [51]. The PALEOMAP project contains global maps for every million years, during the past 1,100 million years. We matched every phylogenetic node in the MCC tree with the closet palaeo-map given their ages.

For seabirds of the order Procellariiformes, our approach ensured that the reconstructed longitudes and latitudes for the phylogenetic internal nodes (i.e., ancestral species) fell within the ancestral configuration of the oceans. The ability to restrict reconstructions to valid palaeo-coordinates means that, for the first time, we could consider continental drift in the reconstruction of the seabirds’ ancestral geographic locations. We ran 10 MCMC chains of 400,000,000 million iterations each, discarding the first 300,000,000 million iterations as burn-in, and sampling every 200,000 iterations. We obtained a final sample of 500 coordinates per phylogenetic node.

Having the reconstructed locations across all phylogenetic nodes, we estimated the branchwise distance, which is a measure of the geographic distance that each ancestral species has moved across phylogenetic branches[9,21]. We calculated the branchwise distance using the Great Circle distance, which is the shortest geographic distance between two geographic points on a spherical surface. As geographic points linked to each branch, we used each coordinate from the posterior distribution of 500 coordinates per phylogenetic node.

Finally, we added up the branchwise distances along the paths that link the common ancestor of Procellariiformes with each extant species. This variable, the Pathwise Distance[9,21], allows us to have a measure of the geographic distance that each extant species has moved over the globe since and from the origin of the common ancestor of Procellariiformes. As we have 500 data points of branchwise distance, then we obtained a final dataset of 500 Pathwise Distances for each species. In this way, we can include the uncertainty in the ancestral location estimation when estimating the Pathwise Distance.

The Pathwise Distance effectively brings a measure of the variability in species dispersal capacity inferred for Procellariiformes, during the last ~63 million years of evolution. Interestingly, the Pathwise Distance does not correlate with neither the commonly used indirect metric of birds' dispersal ability – the Hand Wind Index[14,15] – nor the wing length (Extended Data Fig. 5). This lack of association suggests that the Hand Wind Index and wing length do not represent the distances that procellariiforms species have dispersed over evolutionary time.

Extant species temperature data (LT). We extracted present-day sea surface temperature from the Bio-ORACLE data base[52] by means of the *geodata* R-package version 0.6-2[53]. We extracted the sea surface temperature from the sample of coordinates generated at random within each species' geographic polygons so that we can consider the intraspecific variation in temperature within species geographic ranges. Given that the sea surface temperature is related to the species' geographic distribution, we called this variable the local sea surface temperature (LT).

Ancestral species palaeotemperature data. For phylogenetic nodes, we extracted the palaeo-climate data from the posterior distribution of coordinates inferred with the Geo model with palaeo-map restrictions. We obtained the annual mean two-meters palaeotemperature from world palaeo-climate simulations based on the Hadley Centre general circulation Coupled Model (the HadCM3BL-M2.1aD model) [54–56]. The performance of the HadCM3BL-M2.1aD in simulating modern climate is comparable to the Coupled Model Intercomparison Project 5 and 6, state-of-the-art models [56,57]. Most importantly for our objectives, the HadCM3BL-M2.1aD also recovers the pattern of global temperature change during the last 65 million years as expressed from fossil benthic foraminifera [58]. General circulation models like the HadCM3BL-M2.1aD have been widely used in current palaeo-climate research that have brough meaningful inferences about paleoclimate and diversity. Some examples include the FOAM and CESM models [59–63].

We used 17 dated-palaeoclimate layers, covering the evolutionary history of Procellariiformes as obtained from the sample of 500 phylogenetic trees. The ages for the climatic layers are as follows (million years ago): 0, 4, 10, 14, 19, 25, 31, 35, 39, 44, 52, 55, 60, 66, 69, 75, and 80 (Extended Data Fig. 2). To extract the nodes palaeo-temperature, we used the closest climatic layer to each node in the MCC tree given their ages. There were 94 nodes (79%) whose age was close to the climatic layer age (lower than two million years), and 25 (21%) nodes with a difference over two million years. To account for the potential bias introduced by these age differences in the 21% of nodes, we extracted the palaeo-temperature from the sample of 500 phylogenetic trees which contains variation in node ages. In this approach, we can include the node age uncertainty when estimating ancestral species' environmental temperature (Extended Data Fig. 2; see the testing for phylogenetic uncertainty section).

Species rate of local climate change ($\mathrm{LT_{RATE}}$). To obtain a measure of the rate at which the species' environmental temperature has changed during historical time, we used the *Pathwise Rate of Local Temperature change* metric ($\mathrm{LT_{RATE}}$). This metric was originally implemented in the context of continental temperature change[23]. The $\mathrm{LT_{RATE}}$ was obtained in a three-step approach[23]. First, we extracted the LT for each phylogenetic node in the MCC tree, using the posterior distribution of geographic coordinates for each node. The posterior distribution of geographic coordinates corresponds to the data obtained with the Geo model. Second, we calculated the palaeo-LT absolute difference between the ancestral and descendant nodes per phylogenetic branch. Third, we added up all the per branch absolute differences of LT, along the paths that link the common ancestor with each extant species, and we divided this variable by the total time of each path given the MCC tree.

The LT_{RATE} is the cumulative change of local temperature across the phylogenetic branches that link the common ancestor with every extant species, divided by time. Crucially, these changes are not directional, i.e., changes can be to either cooler or warmer waters. This variable gives us an estimation of the rate of change in local temperature. We also estimated the LT_{RATE} using the posterior distribution of 500 coordinates inferred with the Geo model. Therefore, we obtained a data set of 500 LT_{RATE} for each species given the MCC tree. Finally, we also estimated the LT_{RATE} using the sample of 500 phylogenetic tree to include the node age uncertainty in our downstream analyses.

Speciation rate. To study the correlates of speciation rates, we used the node count (NC) along phylogenetic paths. There are alternative non-model-based tip-rate metrics used to study the speciation rate correlates, such as the inverse of equal splits (ES) or the inverse of terminal branch length (TB)[64]. NC captures the average speciation rate over the entire phylogenetic path and weight equally all branch lengths along the paths. We did not use tip-rate speciation metric estimated from time-varying birth–death diversification models like BAMM[64], as the tip-rates metric is more suitable to study speciation in recent times (not in deep time as in our study). Additionally, it has also been shown in the context of phylogenetic regressions that NC is the response variable that exhibits the highest statistical power when compared to regressions using ES or TB as speciation metrics[65].

Phylogenetic regressions. We conducted Bayesian phylogenetic regression analyses in BayesTraits. We tested for multiple evolutionary models that can fit the regression residual variance, i.e., the Brownian Motion model (BM), the Lambda model (LA), the Variable Rate model (VR); and a combination of the LA and VR model[66]. We ran Phylogenetic Generalised Least Square regressions (PGLS) using body mass, range size, Pathwise Distance, and speciation rate, as response variables. For each response variable, we evaluated the influence of several predictors which has been proposed to correlate with each response variable (see below).

We used samples of trait data in our Bayesian PGLS, which allows us to consider within species variation in the traits of interest. The estimation of both the regression parameters and the evolutionary model parameters are integrated over the sample of trait data. We used the samples of trait data for the LT ($n \sim 500$ per species), Pathwise Distance ($n = 500$ per species), and LT_{RATE} ($n = 500$ per species).

We estimated the effect size of each predictor variable by taking the difference of the R^2 values between the regression with all the significant predictors and without the predictor of interest. We ran PGLS multiple regressions under Brownian motion and using the branch-scaled phylogenetic tree obtained from the PGLS with variable rate. We conducted PGLSs under Brownian motion and using the scaled tree to stabilize the inferred background variance when estimating the effect size of each predictor.

We did not use other regression approaches that examine and quantify the direct and indirect relationships between variables, like phylogenetic path analyses, because those approaches neither cannot fit Variable Rate models nor consider the intraspecific variation in the data when estimating regression parameters as BayesTraits does. We rather evaluated the effect of the hypothesised covariates for each response variables and we also tested for interactions between covariates when they were expected to be associated given previous hypotheses.

Body mass predictors. We included the range size and the LT as additive continuous variables into our Bayesian PGLS. Range size can directly affect body mass as widespread species often experience a wide variety of environmental conditions and larger body size can provide greater physiological tolerance to temperature fluctuations. Second, the effect of LT on body mass is expected given how endothermic physiology varies with ambient temperature. Bigger species have a lower surface-to-volume ratios, which allow them to lose less heat in comparison to smaller species in cold environments. This mechanism can generate a negative relationship between body mass and LT at the interspecific level.

We also included an interaction term for range size and LT in our regression (i.e., range size * LT). This interaction between range size and LT is expected under two hypotheses establishing that climatic variables account for much of the variation in range size[5]. First, the climate variability hypothesis postulates that as the range of climatic variability experienced by terrestrial animals increases with latitude; to survive, individuals need a broader range of physiological tolerance, which consequently allows these species to become more extensively distributed [5]. Second, the climate extreme hypothesis postulates that extreme climatic variables within the species range may relate to range size [5]. Given that higher temperature variability is found at higher latitudes then colder extreme temperatures should favour species with bigger geographic ranges

Additionally, given that there are statistical complications associated with estimating interactions without including quadratic terms[67], that is, nonlinear functions between the response and predictor variables, we added the quadratic terms for both the range size and the LT. Finally, we used the sample of LT data (n ~ 500 per species) to consider the intra specific variation of LT within the geographic range of each species.

Range size predictors. For range size, we evaluated the effect of body mass, LT, $\mathrm{LT_{RATE}}$, Pathwise Distance, and absolute latitude. The expected effect of body mass on range size comes from the common macroecological patterns of body mass and range size explained by the higher demands for food and space of big species[19]. The association of range size with LT is expected from the climate variability and climate extreme hypotheses describe above. The effect of $\mathrm{LT_{RATE}}$ on range size is an indirect effect expected from the hypothesis linking the evolution of longer dispersal distance under increased rates of climate change[27]. If higher dispersal capacity allow species to expand their geographic range[68], then, we should observe an indirect positive effect of $\mathrm{LT_{RATE}}$ on range size mediated by dispersal. The effect of Pathwise Distance on range size, therefore, is expected for the theorical link between dispersal ability and range size[27,68]. Finally, the effect of absolute latitude on range size is expected from the climate variability and extreme climate hypotheses (an also from the mechanism behind Rapoport's rule). We used samples of trait data for LT (n ~ 500 per species), $\mathrm{LT_{RATE}}$ (n = 500 per species), and Pathwise Distance (n = 500 per species).

Pathwise distance predictors. For species dispersal distance, we evaluated the effect of body mass, range size, $\mathrm{LT_{RATE}}$, and the interaction between range size and $\mathrm{LT_{RATE}}$ (including their respective quadratic terms). We expect an association between species dispersal distance and body mass given a previous study reporting a positive effect of body size on natal dispersal distance in terrestrial carnivorous birds[17]. In addition, Ferrer-Obiol et al.[18] proposed that heavier pelagic seabirds should have higher dispersal capacity as body mass is mainly predicted by the latitudinal range and migratory capacity. Nevertheless, evidence from the family Oceanitidae do not support a positive relationship between body mass and dispersal capacity across all the species of the order Procellariiformes. The Wilson's storm petrel is among the smaller species in the group (~ 40 g), but it is widely distributed across almost all south oceans, and it reaches a southern part of the North Atlantic Ocean. Range size should affect positively the dispersal distance given the positive evolutionary feedback between range expansion and dispersal distance evolution[69]. The effect of LT_{RATE} on dispersal distance is expected from the hypothesis linking the evolution of longer dispersal distance under increased rates of climate change[27]. Finally, the interaction term between range size and LT_{RATE} is expected given the indirect effect of LT_{RATE} on range size from the hypothesis linking the evolution of longer dispersal distance under increased rates of climate change[27].

Speciation rate predictors. We evaluated widely suggested factors that influence the rate of speciation in vertebrates. We included range size, body mass, Pathwise distance, LT, and the LT_{RATE} as predictor variables for seabird's speciation rates.

Theoretically, range size may correlate positively with speciation rates because larger ranges have higher probability of being dissected by barriers which promote allopatric speciation[70]. Additionally. some studies[18,24] have suggested that range size should correlate positively with seabird's speciation because the peak of seabird's species richness is found in large, continuous areas of ocean basins, and it is well-known that areas and opportunities for allopatric speciation positively impact the rate of diversification [24]. Body size is postulated to be an additive factor affecting positively speciation rate under the assumption that bigger species have higher range size[19]. Pathwise distance is expected to correlate positively with speciation rate because founder events by long distance dispersal are supported as the main mode of allopatric speciation in the highly mobile species of shearwaters[18]. LT is expected to have a positive effect on bird's speciation rate under the evolutionary speed hypothesis[71]. The evolutionary speed hypothesis postulates that higher kinetic energy increases the rate of speciation through temperature. The addition of LT_{RATE} is justified by its potential indirect effect on speciation rates. This is because higher rates of climate change can select for species with higher dispersal capacity[27] and bigger range size which, in turn, can increase the rate of speciation by range fragmentation. We also evaluated the effect of the following interaction terms: (range size * body mass), (Pathwise distance * LT_{RATE}), and (range size * LT_{RATE}). Finally, we included the quadratic terms for each of the interacting predictors.

All regression coefficients were considered to have significant effect on the response variable based on a calculated value of *pMCMC* for each of their posterior distribution. When > 95% of the estimated coefficients in the posterior distribution crossed zero, this indicate that the coefficient is significantly different from zero.

Phylogenetic prediction of ancestral range size. To better understand the evolution of range size under historical climate change, we conducted a phylogenetic predictive approach to estimate unknow values of range size at phylogenetic nodes. Phylogenetic prediction refers to estimating unknown species (tip) values based on the know values of other species, leveraging the structure of the phylogenetic tree and assumption of an evolutionary model[72]. We followed the approach in reference [73] where ancestral states are estimated by placing zero branch-length "false tips" at each internal node. We based our phylogenetic prediction on the PGLS regression model with variable rate so we can account for the variation in the rate of range size evolution when estimating ancestral states (model comparison showed that the variable rate model fit the data better than other evolution model). We ran the PGLS regression with variable rate in BayesTraits, using the known values of range size as response variable, and the known values of LT and LT_{RATE} as predictors (as we found those predictors to be significant in our regression models, see the Results section). Then, we estimated the maximum likelihood unknown values of range size at each “false tip” based on the “known” values of LT and LT_{RATE} at each phylogenetic node. These known values of LT and LT_{RATE} at phylogenetic nodes were estimated based on both the node location that were obtained from the Geo model with map restrictions and the palaeo-climate simulation.

Robustness of results to the uncertainty in phylogenetic relationships, ancestral locations, paleo temperature, and regression models. We tested the robustness of all our results to the phylogenetic uncertainty of Procellariiformes in terms of divergence times (branch lengths) and topological structure (Extended Data Fig. 2).

First, we ran the Geo model with map restrictions across each of the 500 phylogenetic trees in the posterior distribution. We, therefore, inferred ancestral locations across each of the 500 phylogenetic trees. Then, we obtained the posterior distribution of species Pathwise distance from each of these 500 Geo model analyses (n = 500 Pathwise distance from each of the 500 tree). We also obtained the LT at phylogenetic nodes, and the LT_{RATE} metric, from each of the 500 phylogenetic trees (n = 500 LT and LT_{RATE}, from each of the 500 tree).

Second, we evaluated the effect of phylogenetic uncertainty across all the final PGLS regressions models for each response variable. The final regression models were those regression models that have the statistically significant predictors only, using the MCC tree. Then, we ran the PGLS regression models across each of the 500 phylogenetic trees in BayesTraits, using samples of trait data. This means that, for each of the 500 PGLS, we used samples of trait data for LT (n ~ 500 per species), LT_{RATE} (n = 500 per species), and Pathwise Distance (n = 500 per species). We also used samples of trait data for the interacting terms between predictors, when necessary.

Acknowledgements

This work was supported by The Center for Ecology and Sustainable Management of Oceanic Island ESMOI, ANID-ATE 220044 BioDUCCT, ANID/Centros Regionales R20F0008 (CLAP), and ANID/FONDECYT 1251475.

References

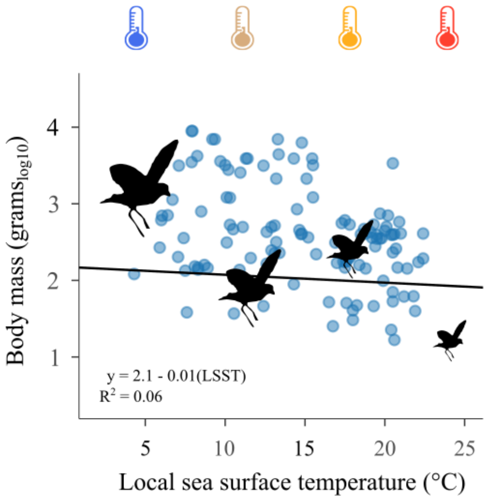
1. Otero, X. L., De La Peña-Lastra, S., Pérez-Alberti, A., Ferreira, T. O. & Huerta-Diaz, M. A. Seabird colonies as important global drivers in the nitrogen and phosphorus cycles. *Nat Commun* **9**, 246 (2018).
2. Sydeman, W., Thompson, S. & Kitaysky, A. Seabirds and climate change: roadmap for the future. *Mar Ecol Prog Ser* **454**, 107–117 (2012).
3. He, J., Tu, J., Yu, J. & Jiang, H. A global assessment of Bergmann's rule in mammals and birds. *Glob Chang Biol* **29**, 5199–5210 (2023).
4. Oswald, S. A. & Arnold, J. M. Direct impacts of climatic warming on heat stress in endothermic species: Seabirds as bioindicators of changing thermoregulatory constraints. *Integr Zool* **7**, 121–136 (2012).
5. Bozinovic, F., Calosi, P. & Spicer, J. I. Physiological correlates of geographic range in animals. *Annu Rev Ecol Evol Syst* **42**, 155–179 (2011).
6. Constanti Crosby, L., Sayol, F. & Horswill, C. Relative brain size is associated with natal dispersal rate and species' vulnerability to climate change in seabirds. *Oikos* **2023**, 1–12 (2023).
7. Lenoir, J. *et al.* Species better track climate warming in the oceans than on land. *Nat Ecol Evol* **4**, 1044–1059 (2020).
8. Sol, D. *et al.* Adaptations to marine environments and the evolution of slow-paced life histories in endotherms. *Nat Commun* **16**, 4265 (2025).
9. Avaria-Llautureo, J. *et al.* Historical warming consistently decreased size, dispersal and speciation rate of fish. *Nat Clim Chang* **11**, 787–793 (2021).
10. Boyles, J. G., Seebacher, F., Smit, B. & McKechnie, A. E. Adaptive thermoregulation in endotherms may alter responses to climate change. *Integr Comp Biol* **51**, 676–690 (2011).
11. Fristoe, T. S. *et al.* Metabolic heat production and thermal conductance are mass-independent adaptations to thermal environment in birds and mammals. *Proc Natl Acad Sci U S A* **112**, 15934–15939 (2015).
12. Clarke, A. & Johnston, N. M. Scaling of metabolic rate with body mass and temperature in teleost fish. *Journal of Animal Ecology* **68**, 893–905 (1999).
13. Clarke, A. *Principles of Thermal Ecology. Temperature, Energy and Life*. (Oxford University Press, 2017). doi:10.1093/oso/9780199551668.001.0001.
14. Gardner, J. L., Peters, A., Kearney, M. R., Joseph, L. & Heinsohn, R. Declining body size: A third universal response to warming? *Trends Ecol Evol* **26**, 285–291 (2011).
15. White, C. R., Blackburn, T. M., Martin, G. R. & Butler, P. J. Basal metabolic rate of birds is associated with habitat temperature and precipitation, not primary productivity. *Proceedings of the Royal Society B: Biological Sciences* **274**, 287–293 (2007).
16. Song, S. & Beissinger, S. R. Environmental and ecological correlates of avian field metabolic rate and water flux. *Funct Ecol* **34**, 811–821 (2020).
17. Sutherland, G. D. *et al.* Scaling of natal dispersal distances in terrestrial birds and mammals. *Conservation Ecology* **1**, 1–36 (2000).
18. Ferrer Obiol, J. *et al.* Palaeoceanographic changes in the late Pliocene promoted rapid diversification in pelagic seabirds. *J Biogeogr* **49**, 171–188 (2022).
19. Brown, J. H. & Maurer, B. A. Macroecology: The Division of Food and Space Among Species on Continents. *Science (1979)* **243**, 1145–1150 (1989).
20. Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, A. O. The global diversity of birds in space and time. *Nature* **491**, 444–448 (2012).
21. O'Donovan, C., Meade, A. & Venditti, C. Dinosaurs reveal the geographical signature of an evolutionary radiation. *Nat Ecol Evol* **2**, 452–458 (2018).
22. Pagel, M., Meade, A. & Barker, D. Bayesian estimation of ancestral character states on phylogenies. *Syst Biol* **53**, 673–684 (2004).
23. Avaria-Llautureo, J. *et al.* The radiation and geographic expansion of primates through diverse climates. (2024) doi:10.32942/X2FS77.

24. Chown, S. L., Gaston, K. J. & Williams, P. H. Global patterns in species richness of pelagic seabirds: The Procellariiformes. *Ecography* **21**, 342–350 (1998).
25. Raftery, A. E. Hypothesis testing and model selection. in *Markov Chain Monte Carlo in Practice* (eds. Gilks, W., Richardson, S. & Spiegelhalter, D.) 163–187 (Chapman & Hall, London, 1996).
26. Mortimer, N. *et al.* Zealandia: Earth's Hidden Continent. *GSA Today* 27–35 (2017) doi:10.1130/GSATG321A.1.
27. Boeye, J., Travis, J. M. J., Stoks, R. & Bonte, D. More rapid climate change promotes evolutionary rescue through selection for increased dispersal distance. *Evol Appl* **6**, 353–364 (2013).
28. Quintero, I. & Wiens, J. J. Rates of projected climate change dramatically exceed past rates of climatic niche evolution among vertebrate species. *Ecol Lett* **16**, 1095–1103 (2013).
29. von Schuckmann, K. *et al.* The state of the global ocean. *State of the Planet* **4-osr8**, 1–30 (2024).
30. Loarie, S. R. *et al.* The velocity of climate change. *Nature* **462**, 1052–1055 (2009).
31. Simpson, G. G. *The Major Features of Evolution*. (Columbia University Press, New York, 1953).
32. Chen, I., Hill, J. K., Ohlemûller, R., Roy, D. B. & Thomas, C. D. Rapid Range Shifts of Species Associated with High Levels of Climate Warming. *Science (1979)* **1024**, 17–20 (2012).
33. NOAA. *NOAA National Centers for Environmental Information. Monthly Global Climate Report for Annual 2023*. (2024).
34. Sheard, C. *et al.* Ecological drivers of global gradients in avian dispersal inferred from wing morphology. *Nat Commun* **11**, 2463 (2020).
35. Tobias, J. A. *et al.* AVONET: morphological, ecological and geographical data for all birds. *Ecol Lett* **25**, 581–597 (2022).
36. Friedman, A. R., Hwang, Y.-T., Chiang, J. C. H. & Frierson, D. M. W. Interhemispheric Temperature Asymmetry over the Twentieth Century and in Future Projections. *J Clim* **26**, 5419–5433 (2013).
37. Quillfeldt, P. & Masello, J. F. Impacts of climate variation and potential effects of climate change on South American seabirds – a review. *Marine Biology Research* **9**, 337–357 (2013).
38. Torres, L. *et al.* Sea surface temperature, rather than land mass or geographic distance, may drive genetic differentiation in a species complex of highly dispersive seabirds. *Ecol Evol* **11**, 14960–14976 (2021).
39. Warheit, K. I. The seabird fossil record and the role of paleontology in understanding seabird community structure. in *Biology of marine birds* (eds. Schreiber, E. A. & Burger, J.) 17–56 (CRC Press, 2002).
40. Kennedy, M. & Page, R. D. M. Seabird supertrees: Combining partial estimates of procellariiform phylogeny. *Auk* **119**, 88–108 (2002).
41. BirdLife International & Handbook of the Birds of the World. Bird species distribution maps of the world. Version 2023.1. (2023).
42. Schreiber, E. A. & Burger, J. *Biology of Marine Birds*. (CRC Press, 2002). doi:10.1201/9781420036305.
43. Graham, N. A. J. *et al.* Seabirds enhance coral reef productivity and functioning in the absence of invasive rats. *Nature* **559**, 250–253 (2018).
44. Serratosa, J. *et al.* Environmental drivers of seabird at-sea distribution in the Eastern South Pacific Ocean: Assemblage composition across a longitudinal productivity gradient. *Front Mar Sci* **6**, (2020).
45. Croxall, J. P. *et al.* Seabird conservation status, threats and priority actions: a global assessment. *Bird Conserv Int* **22**, 1–34 (2012).
46. Dias, M. P. *et al.* Threats to seabirds: A global assessment. *Biol Conserv* **237**, 525–537 (2019).
47. Schliep, K. P. phangorn: phylogenetic analysis in R. *Bioinformatics* **27**, 592–593 (2011).

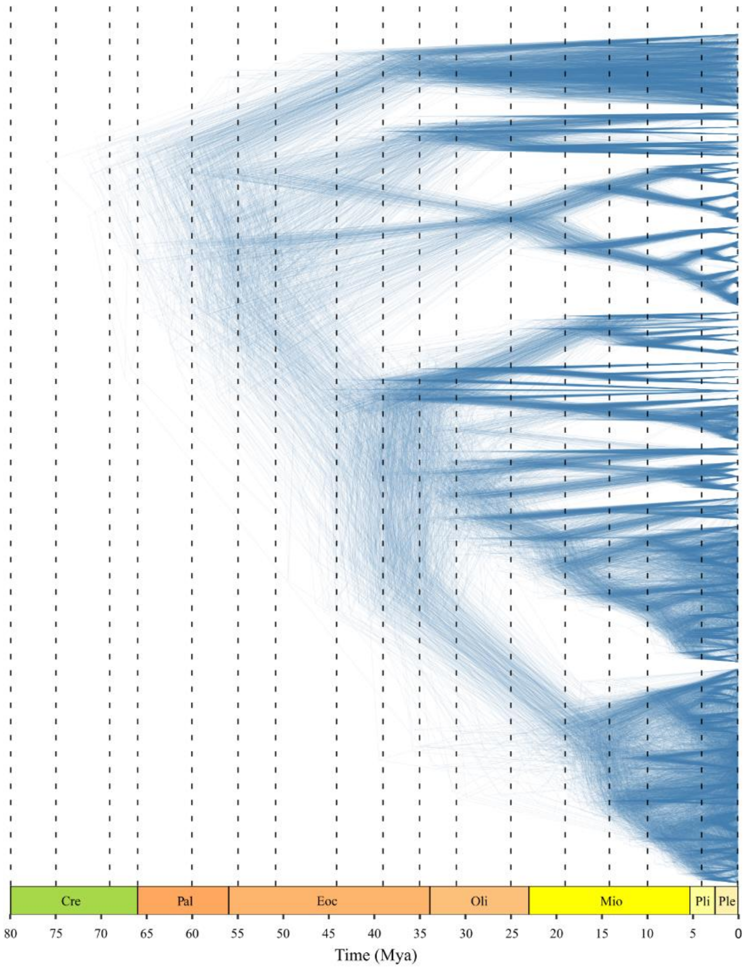
48. Hijmans, R. raster: Geographic Data Analysis and Modeling. R package. Preprint at (2024).
49. Venditti, C., Meade, A. & Pagel, M. Multiple routes to mammalian diversity. *Nature* **479**, 393–396 (2011).
50. Xie, W., Lewis, P. O., Fan, Y., Kuo, L. & Chen, M. H. Improving marginal likelihood estimation for bayesian phylogenetic model selection. *Syst Biol* **60**, 150–160 (2011).
51. Scotese, C. R. & Wright, N. PALEOMAP Paleodigital Elevation Models (PaleoDEMS) for the Phanerozoic. (2018).
52. Tyberghein, L. *et al.* Bio-ORACLE: A global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography* **21**, 272–281 (2012).
53. Hijmans, R. J., Barbosa, M., Ghosh, A. & Mandel, A. geodata: Download Geographic Data. R package. Preprint at (2024).
54. Pope, V. D., Gallani, M. L., Rowntree, P. R. & Stratton, R. A. The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3. *Clim Dyn* **16**, 123–146 (2000).
55. Gordon, C. *et al.* The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Clim Dyn* **16**, 16–147 (2000).
56. Valdes, P. J. *et al.* The BRIDGE HadCM3 family of climate models: HadCM3@Bristol v1.0. *Geosci Model Dev* **10**, 3715–3743 (2017).
57. Valdes, P. J., Scotese, C. R. & Lunt, D. J. Deep ocean temperatures through time. *Climate of the Past* **17**, 1483–1506 (2021).
58. Zachos, J., Pagani, H., Sloan, L., Thomas, E. & Billups, K. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science (1979)* **292**, 686–693 (2001).
59. Pohl, A., Wong Hearing, T., Franc, A., Sepulchre, P. & Scotese, C. R. Dataset of Phanerozoic continental climate and Köppen–Geiger climate classes. *Data Brief* **43**, 108424 (2022).
60. Korasidis, V. A., Wing, S. L., Shields, C. A. & Kiehl, J. T. Global Changes in Terrestrial Vegetation and Continental Climate During the Paleocene-Eocene Thermal Maximum. *Paleoceanogr Paleoclimatol* **37**, (2022).
61. Wong Hearing, T. W. *et al.* Quantitative comparison of geological data and model simulations constrains early Cambrian geography and climate. *Nat Commun* **12**, 3868 (2021).
62. Ontiveros, D. E. *et al.* Impact of global climate cooling on Ordovician marine biodiversity. *Nat Commun* **14**, 6098 (2023).
63. Wilson, L. N. *et al.* Global latitudinal gradients and the evolution of body size in dinosaurs and mammals. *Nat Commun* **15**, 2864 (2024).
64. Title, P. O. & Rabosky, D. L. Tip rates, phylogenies and diversification: What are we estimating, and how good are the estimates? *Methods Ecol Evol* **10**, 821–834 (2019).
65. Harvey, M. G. & Rabosky, D. L. Continuous traits and speciation rates: Alternatives to state-dependent diversification models. *Methods Ecol Evol* **9**, 984–993 (2018).
66. Baker, J., Meade, A., Pagel, M. & Venditti, C. Positive phenotypic selection inferred from phylogenies. *Biological Journal of the Linnean Society* **118**, 95–115 (2016).
67. Ganzach, Y. Misleading Interaction and Curvilinear Terms. *Psychol Methods* **2**, 235–247 (1997).
68. Alzate, A. & Onstein, R. E. Understanding the relationship between dispersal and range size. *Ecol Lett* **25**, 2303–2323 (2022).
69. Miller, T. E. X. *et al.* Eco-evolutionary dynamics of range expansion. *Ecology* **101**, e03139 (2020).
70. Smyčka, J., Toszogyova, A. & Storch, D. The relationship between geographic range size and rates of species diversification. *Nat Commun* **14**, 5559 (2023).
71. Skeels, A., Bach, W., Hagen, O., Jetz, W. & Pellissier, L. Temperature-Dependent Evolutionary Speed Shapes the Evolution of Biodiversity Patterns Across Tetrapod Radiations. *Syst Biol* **72**, 341–356 (2023).

72. Organ, C. L., Shedlock, A. M., Meade, A., Pagel, M. & Edwards, S. V. Origin of avian genome size and structure in non-avian dinosaurs. *Nature* **446**, 180–184 (2007).
73. Baker, J., Meade, A., Pagel, M. & Venditti, C. Adaptive evolution toward larger size in mammals. *Proc Natl Acad Sci U S A* **112**, 5093–5098 (2015).

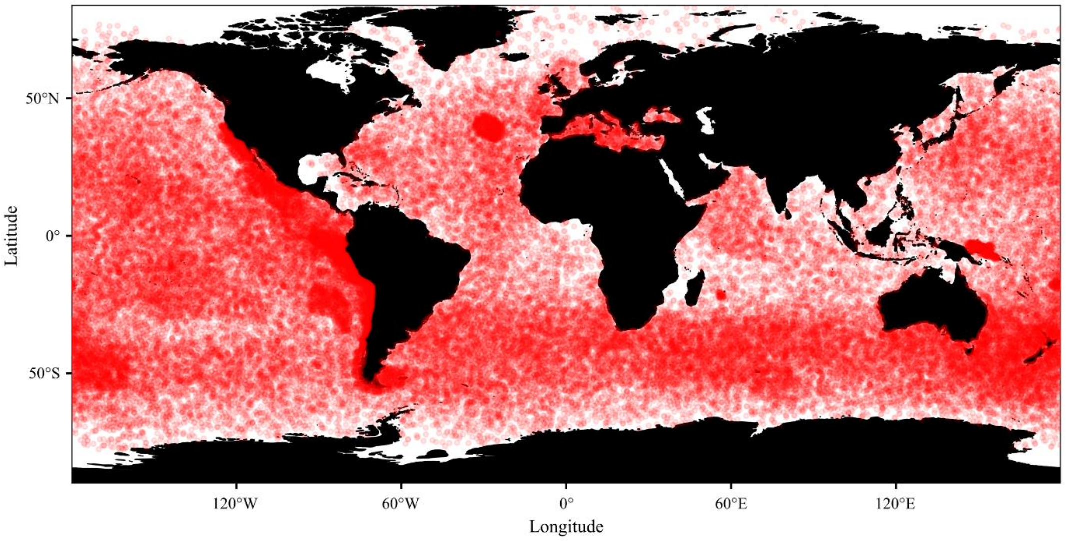
Extended Data Figures



Extended Data Fig. 1: Bivariate phylogenetic regression between body mass and local sea surface temperature (LT). Despite LT has a significant negative effect of body mass, it explains 6% of their variance only. This suggest there can be additional, more relevant factors associated to seabirds' mass diversity. We estimated the regression coefficients by mean of a Bayesian phylogenetic regression, estimating Pagel's lambda. The regression coefficients and the R^2 displayed in the figure are the mean values from the posterior distribution. The solid black line indicates the mean regression slope estimated from the posterior distribution (n = 120).

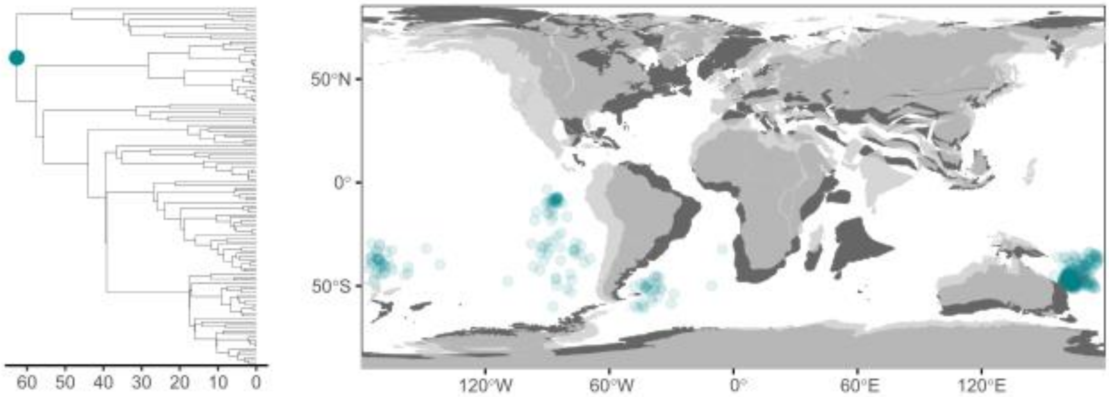


Extended Data Fig. 2: DensiTree obtained from the sample of 500 dated phylogenetic trees and ages of the simulated climate data. The dated trees are stacked on top of each other and the topology of the trees are rotated to ensure the consistency of the tip order. We ran our downstream analyses on each of these 500 phylogenetic to test the robustness of results to phylogenetic uncertainty. Vertical segmented lines indicate the age of each of the climatic layers we used to extract the sea surface temperature from the geographic location reconstructed at phylogenetic nodes.

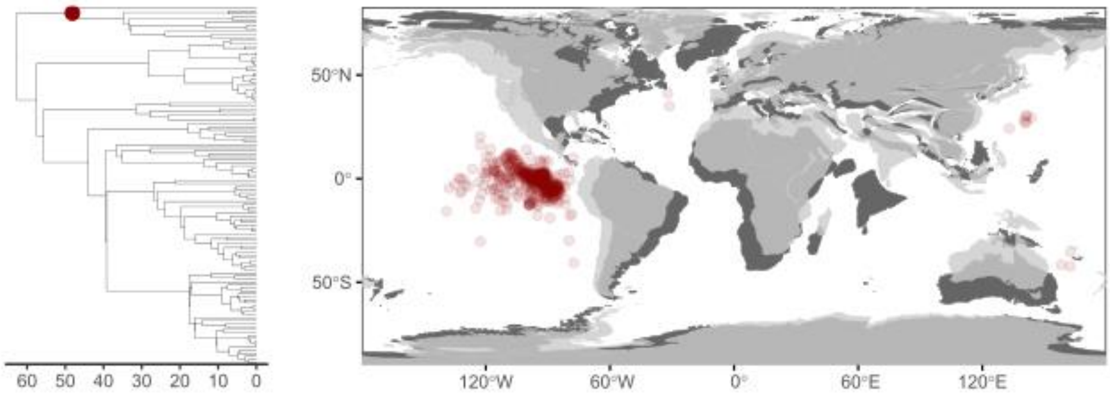


Extended Data Fig. 3: Extant distribution of Procellariiformes. Red-transparent dots represent the geographic coordinates generated at random within the BirdLife distribution polygons for 120 species. These data were used for downstream analyses (see Methods).

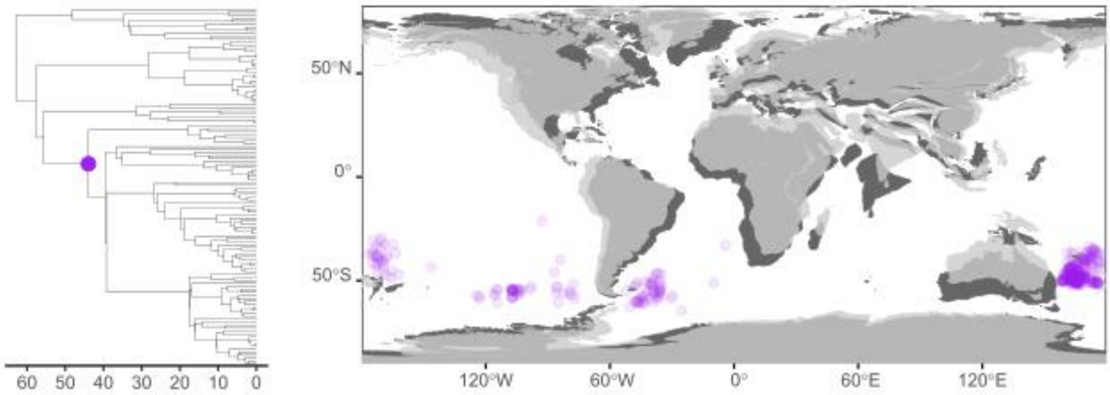
a Most recent common ancestor



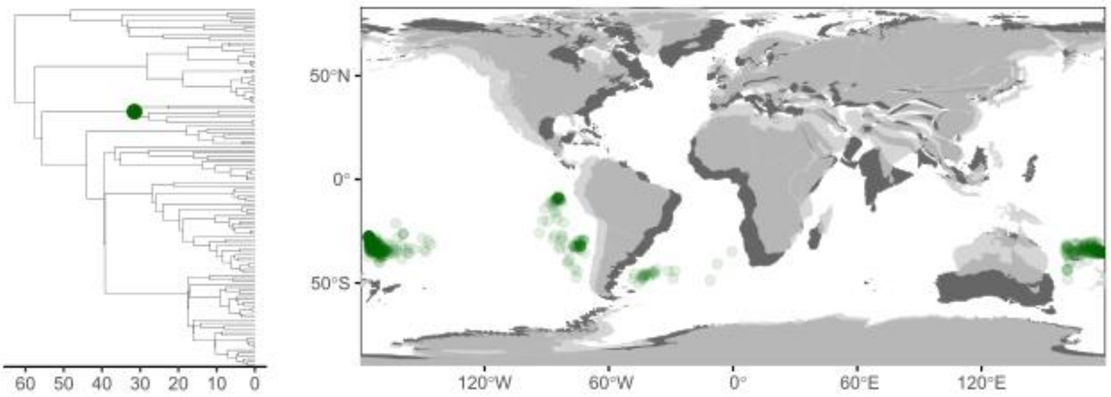
b Hydrobatidae



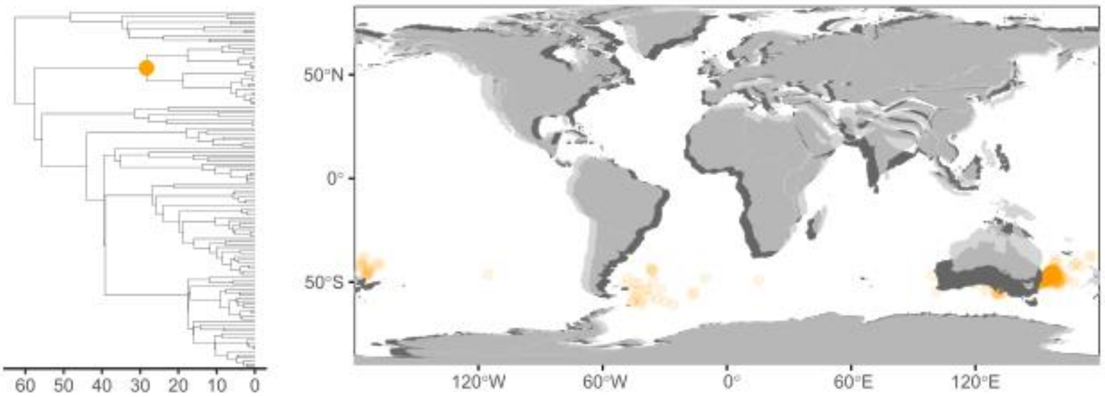
c Procellariidae



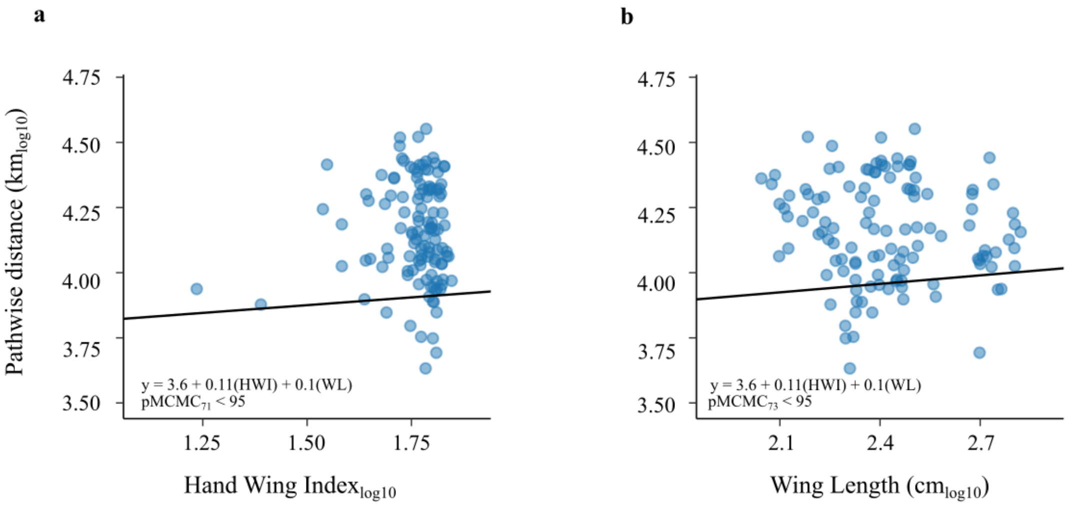
d Oceanitidae



e Diomedeidae



Extended Data Fig. 4. Ancestral locations estimated across each of the 500 phylogenetic trees. We ran the Geo model with palaeo-map restrictions on each tree. We show the phylogenetic nodes representing the most recent common ancestor of Procellariiformes, and four families. Dark palaeo-maps correspond to the oldest node-age in the posterior sample of trees. Light palaeo-maps correspond to the youngest node-age in the posterior sample of trees. Transparent points in the map correspond to the median coordinates from each of the 500 Geo analysis. **a**, the procellariform MRCA was in what is now the Coral Sea, east of Australia, in 82% of the analyses. **b**, the Hydrobatidae MRCA was in the Pacific Ocean, west of Central and South America, in 97.4% of the analyses. **c**, the Procellariidae MRCA was in the Coral Sea in 80% of the analyses. **d**, the Oceanitidae MRCA was in the Coral Sea in 79.8% of the analyses. Finally, the Diomedeidae MRCA was in the Coral Sea, in 96% of the analyses (Extended Data Fig. 5e).



Extended Data Fig. 5. The Pathwise distance is decoupled from commonly used metrics of bird's dispersal ability. Neither (**a**) the Hand Wing Index nor (**b**) the wing length are significantly associated with Pathwise distance (pMCMC < 95). We estimated the regression coefficients by means of a multiple Bayesian phylogenetic regression, estimating Pagel's lambda. The regression coefficients in the functions are the mean values from the posterior distribution. The solid black line indicates the mean regression slope estimated from the posterior distribution (n = 120).

Supplementary Text

In the following, we address commonly concerns related to our methodological approach and to its implication for seabirds' biogeography and evolution.

Biased reconstruction of the dispersal process. Our Geo model infers ancestral locations and dispersal routes without accounting for key biological or ecological factors (that may determine variation in dispersal capacity) as parameters in the model. This omission could lead some researcher to think that our method brings unrealistic or biased inferences about the "real" dispersal process which is shaped by variation in species dispersal capacity. However, this is a misleading conclusion. This is because, whether the historical dispersal process was shaped by variation in species dispersal capacity (owing to ecological or biological correlates), the Geo model will detect such variation across phylogenetic branches by means of the variable rate model[1,2]. In fact, by conducting simulations, we demonstrated that our approach could recover dispersal patterns shaped by varying dispersal capacity with high accuracy and precision[3].

Our methodological approach thus brings the opportunity to explore the potential correlates of such variation in species dispersal capacity by conducting analyses a posteriori. For example, we found a strong effect of the rate of local climate change (LT_{RATE}), range size, and interaction between them, in explaining dispersal distance variation, which is in line with mechanistic explanations of species dispersal capacity[4]. Future work studying the effect of other important factors on dispersal distance such as ocean currents and primary productivity will improve our understanding about the additional factors that shaped the historical dispersal process of seabirds that we inferred in this study.

Lack of fossil in ancestral location reconstruction. One important concern in our study is the absence of fossil data in the Geo model analyses. The inclusion of fossil information close to the origin of a clade (the root of a phylogenetic tree) can drastically change the inference of its geographic location [5,6]. The fossil record of early Procellariiformes is scarce, fragmentary, and subject to high degrees of taxonomic uncertainty[7] which limit our ability to evaluate the robustness of our results to the inclusion of fossil information of early Procellariiformes. However, early procellariiforms' fossils, tentatively assigned to Diomedeidae and Procellariidae, from the late Eocene in Antarctic Peninsula, align with our inferred locations for those families (Fig. 4; Extended Data Fig. 5) which suggest that our results may not change drastically with the inclusion of early, taxonomically certain, procellariiforms' fossils.

Ancestral locations on land. We conducted a Geo model analysis, restricting the location to be found on land, and we evaluated whether this model fit the data better than the model restricting the locations to be found on the oceans. The Bayes Factor model comparison showed strong support for the model constraining location on the oceans rather than on land (Bayes Factor > 10) , indicating that the extant distribution of seabirds has been shaped by a geographic process occurring mainly through the palae-oceans.

Supplementary References

1. O'Donovan, C., Meade, A. & Venditti, C. Dinosaurs reveal the geographical signature of an evolutionary radiation. *Nat Ecol Evol* 2, 452–458 (2018).
2. Venditti, C., Meade, A. & Pagel, M. Multiple routes to mammalian diversity. *Nature* 479, 393–396 (2011).

3. Avaria-Llautureo, J. *et al.* The radiation and geographic expansion of primates through diverse climates. (2024) doi:10.32942/X2FS77.

4. Boeye, J., Travis, J. M. J., Stoks, R. & Bonte, D. More rapid climate change promotes evolutionary rescue through selection for increased dispersal distance. *Evol Appl* 6, 353–364 (2013).

5. Wisniewski, A. L., Lloyd, G. T. & Slater, G. J. Extant species fail to estimate ancestral geographical ranges at older nodes in primate phylogeny. *Proceedings of the Royal Society B: Biological Sciences* 289, (2022).

6. Faurby, S., Silvestro, D., Werdelin, L. & Antonelli, A. Reliable biogeography requires fossils: insights from a new species-level phylogeny of extinct and living carnivores. *Proceedings of the Royal Society B: Biological Sciences* 291, (2024).

7. Acosta Hospitaleche, C. & Gelfo, J. N. Procellariiform remains and a new species from the latest Eocene of Antarctica. *Hist Biol* 29, 755–769 (2017).

Supplementary Tables

Table 1. Evolutionary model fitting for the Bayesian PGLS regressions. The log Marginal Likelihood (Marginal Lh) estimated by stepping stone sampling, provides the models support given the data and priors. The best fitting model, based on the Bayes Factor, is highlighted in bold. BM = Brownian Motion, LA = Lambda, VR = Variable Rate, VRLA = Variable Rate and Lambda, LT = local temperature, LT_{RATE} = rate of local temperature change, Pathwise distance = extant species geographic distances travelled from the location of the common ancestor. Speciation rate = node count along the species phylogenetic paths.

Regression model	Marginal Lh. BM	Marginal Lh. LA	Marginal Lh. VR	Marginal Lh. VRLA
Body mass ~ $\alpha + \beta_1$(Range size)	8.8	4.5	**9.8**	4.8
Range size ~ $\alpha + \beta_1(LT_{RATE}) + \beta_2(LT) + \beta_3$(Body mass)	-148.1	-133.1	**-121.8**	-124.1
Pathwise distance ~ $\alpha + \beta_1(LT_{RATE}) + \beta_2$(Range Size) + $\beta_3(LT_{RATE}$ * Range size)	32.3	**36.7**	37.8	36.3
Speciation rate ~ $\alpha + \beta_1(LT_{RATE}) + \beta_2$(Body mass)	**172.3**	165.3	NA	NA