Science Advances
AAAS

Supplementary Materials for

Ecological interactions and genomic innovation fueled the evolution of ray-finned fish endothermy

Fernando Melendez-Vazquez *et al.*

Corresponding author: Dahiana Arcila, dkarcila@ucsd.edu

Sci. Adv. **11**, eads8488 (2025)
DOI: 10.1126/sciadv.ads8488

This PDF file includes:

Supplementary Text

Expanded sample collection and taxon selection

Phylogenomic analyses were conducted on two distinct datasets encompassing all main lineages that exhibit independent instances of endothermic evolution in marine vertebrates. The first dataset, referred to as the ray-finned fishes dataset (RFFD), uses target capture data and a comprehensive array of 1,051 ray-finned fishes designed for phylogenetic comparative analyses (table S1). The second dataset, the marine vertebrate dataset (MVD), integrates whole-genome sequencing data from 205 marine vertebrates, facilitating in-depth comparative genomic analyses (table S17).

For the RFFD, our taxonomic sampling comprised 1,051 species and up to 1,099 single-copy nuclear exon markers (table S1, dataset S1). For this study, we generated additional genomic data for 763 species of ray-finned fishes (dataset S1). This dataset includes 22 endothermic teleosts, representing four independent events of endothermic evolution in the families Lampridae (*Lampris incognitus*), Xiphiidae (*Xiphias gladius*), and Scombridae (e.g., *Thunnus orientalis, and Gasterochisma melampus*). Tissue samples were obtained from voucher specimens in museum collections (table S1), including The Australian Museum (AMS), The Smithsonian National Museum of Natural History (NMNH), The Hokkaido University Museum Fisheries Science Center (HUMZ), the University of Kansas Biodiversity Institute and Natural History Museum (KU), the Burke Museum (UW), and the Scripps Institution of Oceanography (SIO). We conducted DNA extractions at the Laboratory of Analytical Biology (LAB) at the NMNH. The extractions were carried out in a 96-well plate format using a GenePrep system. We used a 1% agarose gel stained with GelRed (Biotium) to identify high molecular DNA weight. Additionally, 15 DNA extractions were performed using the QUIAGEN DNeasy Blood & Tissue Kit following the manufacturer's instructions. QUIAGEN DNA extraction quality was assessed using fluorometric quantification via Qubit (Thermo Fisher Scientific). All DNA extractions were shipped to Arbor Biosciences for library preparation and target enrichment using the Fishlife project exon probe sets: Backbone 2, Eupercaria, and Pelagiaria (*70*). Our taxonomic sampling was expanded by adding previously published genomic data for 285 species of ray-finned fishes (*30, 71*) and three outgroups (*Xenopus tropicalis, Protopterus aethiopicus*, and *Latimeria chalumnae*) for a total of 1,051 species.

The marine vertebrate dataset (MVD) comprises whole genomes from 205 different marine species (table S17). This dataset includes 138 long-read genomes from NCBI and 69 additionally generated short-read genomes of ray-finned fishes for comparative genomic analyses (table S17 and dataset S9). Our taxonomic sampling comprised seven species of marine mammals, including whales (*Eschrichtius robustus, Balaenoptera musculus,* and *Megaptera novaeangliae*) and seals (*Arctocephalus gazella*, *Callorhinus ursinus*, *Mirounga angustirostris,* and *Mirounga leonine*), three species of penguins (*Eudyptes chrysolophus*, *Eudyptes moseleyi*, and *Eudyptes filholi*), two species of marine reptiles (*Dermochelys coriacea* and *Chelonia mydas*), one species of Chondrichthyes (*Carcharodon carcharias*), 172 species of ectothermic ray-finned fishes and 20 species of endothermic ray-finned fishes. For the additionally generated short-read genomes, we conducted DNA extractions as previously described. The samples were sent to the Oklahoma Medical Research Foundation Clinical Genomics Center for library preparation and sequencing. The short-read assemblies exhibited a read coverage ranging from 12X–300X, with some species having multiple individuals per species represented.

Expanded data assembly and quality control

Data assembly for the ray-finned fishes dataset followed the previously proposed bioinformatic pipeline by Hughes *et al*. (*70*). We first removed low-quality bases and any traces of adapter contamination using Trimmomatic v0.36 (*72*). Subsequently, raw reads were mapped to the respective reference sequence based on the three different Fishlife probe sets using Python (*73*), the Burrows-Wheeler Alignment tool (*74*), and SAMtools v.1.7-1.9 (*75*). We then created our initial assemblies using Velvet v1.2.10 (*76*). We implemented the automated target-restricted assembly method (aTRAM) software (*77*) with Trinity v2.8.5 (*78*) as the assembler. Exonic reading frames were collapsed and generated using CD-HIT (*79*) and exonerate (*80*), respectively. The assembled alignments were combined with previously published sequence data (*30*, *71*) and were subsequently aligned using MACSE v2.05 (*81*), which accounts for frameshifts and stop codons. To ensure the quality of our alignments, we removed any instance of single-taxon insertions, trimmed gappy edges, and filtered short sequences using a custom Python script (AlignmentCleanerCodons.py, https://github.com/lilychughes/FishLifeExonCapture). Additionally, we manually inspected approximately 20% of the gene alignments to check for reading frame modifications, potential misalignment, and data vetting using Geneious Prime v2022.2.2.

Comparative genomic data assembly followed two approaches for all 205 marine vertebrate species (MVD). The initial method focused on estimating a robust phylogenetic framework from the 1,099 single-copy nuclear loci detailed earlier (*70*). This step was achieved by mining whole genome sequencing data via the NHMMER framework in HMMER v3.1 (*82*). The *in silico* data mining process generates hidden Markov model (HMMs) profiles for each genetic marker, followed by a query search between the markers and the genomic sequences. Assembly was conducted in Trinity v2.8.5 (*78*), and all mined sequences were aligned in MACSE v2.05 (*81*). Gene alignment quality control followed the same procedure as above. The second approach aimed to assemble a set of single-copy orthologous candidate genes to pinpoint selection signatures linked to the evolution of endothermy. This step involved compiling sequencing data for 61 previously identified candidate genes associated with either endothermy or endothermic-mediated physiological processes stemming from a comprehensive literature review (*1*, *4*, *25–27*, *83*, *84*). Orthologs yielding high query similarity for each gene were downloaded from the Ensembl genome browser.

Moreover, 22 of the 205 species with annotated genomes in NCBI were selected from our genomic dataset to identify additional single-copy orthologous sequences using OrthoFinder v2.5.4 (*85*). Since OrthoFinder's input data are amino acid sequences, genome annotations in nucleotide and amino acid sequences spanning the diversity of marine vertebrates were downloaded from NCBI. After running OrthoFinder, we identified 833 single-copy orthologous amino acid sequences. Given that translation from amino acids to nucleotides is not feasible, we developed a custom Python script (dataset S11) that cross-referenced and matched the NCBI accession numbers (species and respective gene) from our identified orthologous amino acid sequences to our nucleotide sequence reference genome. This strategy facilitated the identification of the corresponding nucleotide sequence for every one of the 833 orthologous amino acid sequences, thereby yielding 833 single-copy orthologs in nucleotide sequences. Following the mining pipeline, these nucleotide sequences were used as references to map sequences for the remaining 183 unannotated species (dataset S9).

Expanded phylogenomic analyses and divergence time estimation

Phylogenetic analyses were conducted using both Maximum Likelihood (ML) and multispecies coalescent approaches for the RFFD and MVD datasets. For each dataset, individual gene alignments (ML) were concatenated into a supermatrix using AMAS (*86*) (datasets S1 and S10). We estimated preliminary trees using FastTree2 (*87*) for quality control and to detect potential contamination or species misidentification. Additionally, we visually inspected the gene alignments to address phylogenetic inconsistencies arising from alignments with limited taxonomic representation (e.g., species featured in fewer than ten markers) and to remove taxonomic duplicates that resulted from integrating previously published datasets (*30*, *71*).

The best-scoring ML trees for the RFFD and MVD datasets were obtained using IQTREE v2.0(*88*) under a mixture model partitioning scheme "MIX (JC, K2P, HKY, GTR)" (dataset S2). Mixture models offer the advantage of incorporating multiple substitution models and assessing the fit for each mixture category at every site. As a result, this approach is less computationally demanding than alternative techniques in determining the most suitable partitioning scheme for datasets (*71*, *89*). We estimated branch support on all ML datasets using the ultra-fast bootstrapping approach with the "-bnni" argument on 1,000 replicates (*88*). Gene trees were estimated from single-gene alignments using IQ-TREE and mixture models, as explained above, including one individual per species. Lastly, we conducted species-tree analyses using ASTRAL-III v5.1.1 (*90*) using gene trees as input after collapsing branches with <30% bootstrap support (dataset S2).

We accounted for phylogenetic uncertainty in our RFFD by estimating four randomly sampled genomic subsets, each comprising approximately 220 largely non-overlapping exon markers but including seven "anchor" legacy nuclear markers (*37*, *41*). Phylogenetic estimation for the subsets was conducted in IQTREE v2.0 (*88*) as indicated above. Although multispecies coalescent approaches are robust to low and moderate levels of missing data (*91*), we avoided using ASTRAL-III for subsets with high levels of missing data (>90%) because substantial missing data in some gene trees substantially increases gene tree error, thereby misleading multispecies coalescent tree inference(*91*, *92*).

Divergence time estimation was performed only for the ray-finned fish dataset (RFFD). Considering the computational challenges associated with estimating divergence times for large phylogenies within a Bayesian framework (*93–97*), we used a previously proposed MCMC time-calibrated tree of ray-finned fishes (*30*), which includes 303 species, as our reference tree. We employed a congruification approach using *geiger* v.2.0 (*98*). This method aligns divergence times from a calibrated reference tree to an uncalibrated target phylogeny, leveraging shared taxa as anchor points (*31*). For our uncalibrated target phylogeny, we selected both our IQ-TREE and ASTRAL trees, as well as trees generated from four independent subsets in IQ-TREE, resulting in a total of six trees. To account for uncertainty in divergence time, we employed the three alternative calibration schemes presented by Hughes *et al*. (*30*) for each of the six trees. These schemes consisted of 31, 33, and 34 fossil calibrations (datasets S2–S6), resulting in a total of 18 alternative time-calibrated trees (dataset S2). Upon completing the congruification step, we conducted an extensive comparison of our divergence times against other recently published time-calibrated phylogenies of ray-finned fishes at the family, order, and series level (*30*, *32–40*). Owing to the compelling evidence of additional fossil species informing the time of evolution of Pelagiaria (*38*), Carangaria (*36*), Anabantaria (*36*), and Syngnatharia (*37*), we included eight additional fossil

calibrations (table S2 and dataset S6) using treePL (*99*). These were summarized in our IQ-TREE tree based on 1,051 species and 1099 genes (Fig. 1 and fig. S1). In total, 19 time-calibrated trees based on the RFFD were used for subsequent phylogenetic comparative analyses (dataset S2).

Expanded compilation of ecological and morphological traits

We compiled a multifactorial matrix comprising five key ecological traits (table S3). These traits have been previously hypothesized to influence the evolution of fish endothermy: average standard length (cm), length-weight relationship (g), swim mode, diet, and depth range (m)(*1, 4, 13, 17, 84, 100, 101*). The matrix was derived from all 1,051 ray-finned fish species. Trait data were obtained from various repositories, including FishBase (*102*), The Smithsonian's Fishes: Greater Caribbean: A Guide to the Shorefishes of the Caribbean, and adjacent areas (*103*), Fishes of the World (*104*), the Fishes of Australia database (*105*), the Food and Agriculture Organization (*106*), the Discover Life Database (*107*), and the Reeflex Database (*108*). Body size metrics included average standard fish length (cm) and length-weight relationships, defined by the relational formula W=*a*L^*b*(*138*), where *a* describes the rate of change of weight with length (intercept), *b* represents weight at unit length (slope), and L refers to the length of the fish.

Dietary preferences were categorized into 30 distinct groups, ranging from zooplankton and zoobenthic organisms to combinations such as nekton-zoobenthos. We accounted for the impacts of fish vertical movement across the water column (*5*) by utilizing a depth range (m) metric. This metric represents the difference between each species' minimum and maximum depth, providing an estimated value of a species' theoretical vertical range. Recognizing the potential role of swimming styles in endothermy (*1, 17*), we identified 11 distinct swimming modes, including carangiform, sub-carangiform, thunniform, amiiform, labriform, diodontiform, and anguilliform. We employed alternative schemes to address ambiguous data coding for the abovementioned categories, such as corroborating with multiple databases and replicating tests when encountering ambiguous traits.

Additionally, we utilized the FishShapes (*46*) database to explore relationships between endothermy and various morphometric factors (table S4). Variables from this database included average body depth (cm), average body width (cm), average head depth (cm), average mouth width corrected for size (cm), average caudal peduncle width (cm), average weight (g), body depth-length ratio (cm), and body depth-length-weight ratio (cm^2/g). The body depth-length ratio was calculated using the following formula ((standard length/body depth) x standard length), while the body depth-length-weight ratio was calculated ((standard length/body depth) x weight). To ensure the reliability of our results, we re-evaluated the average standard length (cm) factor using FishShapes values to check for consistency with the average standard length obtained from multiple sources. This validation step was essential for confirming the robustness of our findings.

Expanded correlated trait evolutionary analyses

To elucidate potential factors associated with the evolution of endothermy in fishes–such as multi-state discrete characters (e.g., diet and swimming mode) or continuous characters (e.g., average length, length/weight relationship, and depth range)–we employed a Phylogenetic Generalized Logistic Regression (PGLR) analysis (*48*). The PGLR framework allows for the analysis of binary dependent variables where the distribution of values is not independent among species but is phylogenetically structured(*48, 139*). The 'phyloglm' function from the *phylolm* (*109*) R package

was used for these analyses. Analyses were structured into univariate and multiple logistic regression frameworks for each factor with 1000 bootstrap replicates across all our estimated chronograms (19 total trees). Bootstraping was used to assess the variability in our parameter estimates. This variability is then used to construct confidence intervals around these estimates, providing a measure of uncertainty for our model's predictions. The univariate approach explored the interaction between individual factors, such as diet type and endothermy status. In contrast, the multivariate framework incorporated interactions among all investigated factors (tables S3-S4) and their relationship with endothermy status. Given the inherent challenges in designating swim modes to istiophorids, especially when differentiating between sub-carangiform and thunniform classifications (*1*, *17*, *110–112*), we conducted additional univariate and multivariate regressions for each proposed swim mode (tables S5–S8). All results, including single-value outliers, can be found in Fig. 2A and tables S5–S8.

Expanded ancestral reconstructions of endothermy

We used stochastic character mapping of discrete traits (SIMMAP) (*42*) to estimate ancestral character states for the presence or absence of endothermy for all nodes of our RFFD phylogeny. The SIMMAP algorithm samples character states proportional to their conditional likelihood estimated under a Markov process for transitions between character states at each node of the. Conditional on these simulated values, it stochastically maps histories along the branches tree (*42*, *113*). Summarizing a large draw of stochastic mappings allows for approximating the marginal ancestral state. We first assessed whether these transitions fit a model where shifts between state 0 and state 1, and vice versa, exhibit equal rates (ER) or if, instead, these transitions are different (ARD) (*114*). We performed stochastic character mapping (100 simulations) across all of our estimated chronograms (19 total trees) using the 'make.simmap' function to account for phylogenetic and divergence time uncertainty in *phytools* v2.2-1 (*114*).

Expanded state-dependent diversification

Models for state-dependent speciation and extinction (SSE) have been recently developed (*115*). These models offer a robust toolkit for inferring the processes controlling the patterns of species richness and trait diversity (*45*, *116*), given that they jointly account for speciation and extinction rates for each character of interest alongside state transition rates(*45*, *117*). A form of these SSE models known as the Hidden State Speciation and Extinction (HiSSE) allows for the association of an unobserved state with multiple diversification rates (*43*). The HiSSE model acknowledges two traits classified as 0 or 1, which can be associated without (A) or with (B) a hidden state (*43*, *45*). Lastly, it associates speciation (λ) and extinction (μ) rates with these character state transitions (*43*, *45*). The objective of these models is to accommodate potential alterations in the transition rates among these character states by representing these shifts using hidden variables.

To discern the potential evolutionary impacts resulting from the origin of endothermy, we evaluated the suitability of various SSE models on our RFFD tree using the *hisse* (*43*) package in R. We assigned ectothermic fishes as "0" and endothermic fishes as "1" and estimated turnover rates, extirpation rates, state transitions, and rates of diversification (*45*). We estimated HiSSE rate parameters for five types of models. The null BiSSE model is a restricted form of the full BiSSE model, in which the birth and death rates are the same across different states (*118*). A full BiSSE (BiSSE like HiSSE) where we allowed all parameters to vary independently, assuming that

diversification rates are associated with the presence or absence of the trait. The HiSSE model looked at differential diversification associated with endothermy and any hidden states. CID-2 and CID-4 models ignored endothermy and focused on diversification associated with two (CID-2) or four (CID-4) hidden states. The Akaike information criterion (AIC) was then used to identify the best-fit model. Our sampling fraction consisted of 77% endothermic ray-finned fishes and 3% ectothermic ray-finned fishes. To account for this imbalance, we also ran an additional analysis in which we reduced our number of species to two per transition, giving us a sampling balance of 23% to 3% endothermic to ectothermic ray-finned fishes (dataset S7). To further assess for any potential biases in our results caused by imbalance, we also ran our SSE analyses using a previously published phylogeny of ray-finned fishes (*34*) based on 11,638 species of fishes with a sampling fraction of 93% to 36% endothermic to ectothermic ray-finned fishes and a completely balanced analysis with 36% to 36% endothermic to ectothermic ray-finned fishes. Boxplots were used to assess differences in diversification rates between two states for all tips and nodes across the phylogeny in the R package *ggplot2* v.3.3.2 (*121*) (dataset S7). Finally, we statistically compared tip-associated lineage diversification rates between ectothermic and endothermic species by conducting a phylogenetic ANOVA in *RRPP* with Bonferroni correction (dataset S7) (*119*, *120*).

Expanded explanation of model to assess time-variable associations in the origin of discrete traits

Past climatic events have notably driven many evolutionary tendencies in organisms (*6*, *13*, *122–129*). While some studies have alluded to the potential impact that past climatic changes may have on the evolution of endothermy (*6*, *7*, *13*, *26*, *124*, *125*, *127*), there have been no empirical methods suitable to test this hypothesis. To address this, we developed a generalized maximum likelihood version of the threshold model for binary characters proposed by Felsenstein(*50*). The threshold model assumes that the character transitions across the phylogeny are controlled by an underlying latent (unobserved) continuous process. A threshold value is chosen to assign character states to the values of the latent trait. The original implementation of the threshold model assumes that the latent trait evolves by Brownian Motion along the tree and uses Bayesian MCMC for integrating across all possible paths of that process conditional on the observed discrete states at the tips (*50*). Here, we extend these models by implementing maximum likelihood approaches using numerical integration – Gaussian quadratures – techniques to integrate across the latent factor following the algorithm described in a previous study (*130*). This maximum likelihood-based implementation allows us to consider the estimation of alternative evolutionary processes for the latent trait. Here, we implement a BM-based model where the rates of evolution depend on a time-dependent function (e.g., global climatic changes through time) (*131*). In this model, the transition between discrete states is thus controlled by changes in the dynamic of the latent factor, which is conditioned on some known – e.g., climatic – quantities changing through time. In our analyses, this model is used to assess whether cold/warm periods have spurred transitions from ectothermy to endothermy.

We rigorously tested the performance and efficacy of our model with extensive simulations focusing on accurately detecting the generating model and inferring a relationship with climatic changes (e.g., positive or negative association)(*131*). For our simulations and empirical analyses, we used spline interpolation of a global average paleo-climatic curve (*132*). We performed simulations on pure-birth trees of increasing size (100, 500, and 1,000 taxa) scaled to a root age of

540 myr using the 'pbtree' function found in the *phytools* R package (*114*). We then simulated a set of binary discrete traits across our tree based on varying parameters; this helped us determine if our function had enough power to correctly detect the generating model used to simulate the traits and parameters (*131*). For each parameter combination (detailed below), we simulated 100 datasets. In total, there were 16,200 simulations used to assess the statistical performances of the approach under various scenarios (figs. S2-S11 and table S9)

Following Clavel & Morlon 2017 (*131*), we used an exponential link function to relate the time-varying rate of a latent BM process (**X**) to climate change in our climatic-threshold model (i.e., $\sigma^2(t) = \sigma_0^2 e^{\beta T(t)}$; where σ_0^2 is the basal overall rate, and β is the parameter controlling the association to the time-dependent function T(t) – e.g., the interpolated climatic function described above). To test the identifiability of our threshold model, we simulated traits on our trees with varying levels of association with the climatic curve ranging from no association (β=0), to negative association (β= -4 and -8), and positive association (β= 2, 4, and 8). We ensured that our climatic model was not selected when it was not the generating model by considering alternative threshold models where the latent trait evolved under different processes, including Brownian motion (BM – with time-homogeneous rate), Pagel's Lambda, Early Burst (EB), and Ornstein-Uhlenbeck (OU) (*131*). For our simulations, we considered a fixed diffusion rate of one (σ^2=1) for BM, EB, Lambda, and OU latent processes. Given that the scale of the latent factor is arbitrary, the choice of this value has no incidences on the estimation. For the OU latent process, we used an α value that yielded a phylogenetic half-life of 20%, 30%, 50%, and 70% of the tree height (*131*). For EB simulations, we used values for the decay rate, resulting in identical half-lives to the ones used for the OU model (*131*). Traits simulated under the Lambda model were generated under three different values (λ= 0.3, 0.5, and 0.8) representing low to high phylogenetic signals (note that λ= 1 corresponds to BM). We used a starting value at the root for each latent process of zero and thus used a threshold value of zero to assign characters to either binary states conditional on the latent factor for each tip in the tree (i.e., latent value X_i<0 is assigned state "0" and X_i>=0 is assigned "1" for species *i*). We simulated 100 datasets for each parameter set. Finally, we replicated the model fit on each dataset using various numbers of subintervals for the numerical integrations (Gaussian quadrature) used to approximate of the log-likelihood of the threshold model (N= 100, 200, and 500).

The simulations show that the proposed models perform well in recovering parameters and in model selection (figs. S2–S11). We saw no discernable differences among the various levels of integration (N). However, we did see some substantial improvements in our model's performance as we increased the taxonomic sampling, indicating that higher levels of taxonomic representation will increase the accuracy of our model. Additionally, our model seemed to perform better under stronger parameter values, whereas lower parameter values yielded more conservative results (table S9). We thus fitted the model to our empirical dataset using the paleo-climatic curve (*132*) and our phylogeny based on 1,051 taxa and 1,099 single-copy nuclear exon markers (dataset S1). We assigned character states based on the presence (1) or absence (0) of endothermy for each taxon. We conducted a model selection analysis in which we compared the fit of our climactic threshold model to alternative models (with the latent factor model by BM, OU, EB, and Pagel's Lambda). These analyses were conducted using (N=90) for the subinterval of the numerical integration since there were no noticeable differences in the simulations with larger N. The climatic spline interpolation function was obtained with a degree of freedom (df) set to 300. To account for uncertainty in phylogenetic placement and divergence time estimation, we ran our model selection analyses using all of our previously estimated phylogenies and divergence times (see topological

and divergence time uncertainty section above). Due to the circumglobal distribution of our endothermic species (*1*, *4*, *5*), we favored the global average paleo-temperature curve (*132*) for our dataset. However, we also used paleo-climatic curves based on polar, tropical, and deep-sea temperatures (*132*) to ensure that evolutionary resolution was not compromised.

Expanded tests of biotic interactions in the origin of endothermy

The relatively recent evolution of endothermy in ray-finned fishes during the last ~50-30 myr has not been previously linked to the diversification dynamics of cetaceans, the major apex predator in the marine realm, which occurred during the Eocene (*4*, *51*). To test for an evolutionary relationship between the emergence of endothermy in ray-finned fishes and the evolutionary trends of modern cetaceans, we revisited our previously proposed modified threshold model. By leveraging previously published tip-dated phylogenies for the whole cetacean group (Neoceti), Mysticeti, and Odontoceti individually (*51*), we generated Lineage Through Time (LTT) plots for each group using the 'ltt' function found in the *phytools* R package (*114*) followed by a logarithmic transformation of those values. We then re-fitted our threshold model by substituting the paleo-climatic curve with the estimated LTT curve, allowing us to assess any evolutionary correlation between the diversification patterns of modern cetaceans and endothermic ray-finned fishes. By incorporating tip-dated phylogenies, we accounted for the lack of extinction consideration when estimating LTTs using only extant species (*134*). We also calculated the Diversification Through Time (DTT) plots for the cetaceans (*135*). By estimating both birth and extinction rates, DTTs provide a less biased evolutionary picture compared to non-tip-dated phylogenies (*134*, *135*). Because DTT estimation is based on an ultrametric tree, we estimated DTT values for all extant cetaceans (Odontoceti + Mysticeti) combined (*51*). Given that DTT will only accept ultrametric trees as input, we avoided doing separate analyses for the Odontoceti and Mysticeti groups, respectively, as we did with the LTT values, because when removing the fossil tips, the remaining taxonomic sampling was so low (<100 species) that it could have compromised the accuracy of our model as observed in our simulations (figs. S2–S8). We used the 'find.mle' and 'fit_bd' functions in the *Diversitree* (*136*) and *RPANDA* (*133*) packages to determine DTT values.

We also estimated the LTT and DTT values for all elasmobranchs using a previously published phylogeny (*53*) and the aforementioned approaches to further validate our results. Despite the elasmobranch phylogeny by Stein *et al*. (*53*) being one of the most taxonomically well-sampled ones available, it lacks fossils as terminal taxa, risking biased LTTs results (*134*). While using DTT values helps account for these potential issues (*135*), we also conducted additional analyses using LTT values for a tip-dated phylogeny of Carchariniformes and DTT estimates for them as well (*54*). We performed a model selection analysis, comparing the goodness of fit of our LTT and DTT-adapted threshold model to other models such as BM, OU, EB, and Pagel's Lambda. These analyses involved 90 integrations (N=90) for optimal computing power-to-accuracy ratio, and a spline interpolation function with 300 degrees of freedom (df=300) (dataset S8).

Given that our threshold model fits a time-varying function, using different calibrations and altering the dates when endothermy appeared in ray-finned fishes should significantly affect our results. For this reason, our time-dependent threshold model was tested against various artificially adjusted dates that altered the origin of endothermy by 50% and 25%, respectively, to further validate our results and avoid any potential instance of a false positive. If the observed effect remained constant across all dating alterations, it would indicate an issue with our model or the

curve used, suggesting false positives. Results using different dates revealed drastic changes, with the effect disappearing (tables S11–S16) in other cases (see also main text). These findings confirm that our model and curves function as intended and that the observed effect is closely related to the most reliable origins of endothermy in ray-finned fishes according to the fossil record (table S2). We will implement the model developed for this study in the next update of *RPANDA (133)*.

Expanded biogeographic occupancy overlap procedure

To delve deeper into the macroevolutionary history of potential ecological interactions between modern cetaceans and endothermic ray-finned fish lineages, we looked at fossil records from the past 40 myr by examining their historical biogeographic occupancy overlap. If no overlap was observed, it would indicate a lack of interactions between the groups, suggesting that modern cetaceans did not influence the emergence of endothermy in ray-finned fishes. We analyzed 12 geological formations, encompassing fossil lineages for three primary groups: Neoceti, endothermic chondrichthyans (Alopiidae and Lamnidae), and ray-finned fishes (Scombriformes, Istiophoriformes, and Lampridae). Furthermore, we incorporated fossil data for two ray-finned fish lineages, Clupeiformes and Myctophiformes, which are well-represented in geological formations and serve as prey for these marine predators (Fig. 3). These prey fossils were included to establish a historical dietary correlation to complement the geographical correlation observed, thereby reinforcing the plausibility of such ecological interactions occurring. We extracted information from The Paleobiology Database (paleobiodb.org), including taxonomic group, type of interaction (prey, predator, or both), early interval, maximum age, minimum age, paleo-longitude, and paleo-latitude (table S27). This information was extrapolated on several maps to better visualize the overlaps (Fig. 3).

Expanded description of tests of positive selection, rates of evolution, and molecular convergence in endothermic lineages

To test for signatures of selection across marine vertebrates, we analyzed 894 single-copy orthologs (see Data Assembly and Quality Control section) across our estimated Marine Vertebrate Dataset (MVD) tree, which includes 205 marine vertebrate species (table S17 and datasets S9–S10). We implemented two additional iterations of the Branch-Site Unrestricted Statistical Test for Episodic Diversification (BUSTED) analysis in the HyPhy package (*58*): BUSTED-E and BUSTED-PH. Each model is based on BUSTED[S], which accounts for synonymous site-to-site substitution rate variation (*59*, *60*). BUSTED-E is an alignment error filtering method that uses the sensitivity of branch-site models and categorizes specific regions of a sequence alignment as potential sources of error, grouping these problematic regions into a specific error-sink category (fig. S12). This model takes the standard BUSTED ω distribution and adds an error sink category where $\omega > 100$ (*60*). Codons that are categorized into this error sink category are masked without affecting the remaining ω distribution.

BUSTED-PH is a branch-site model using the random effects likelihood framework (*60*, *140*), applied to specific subsets of tree branches representing the partitioning of the tree by a discrete trait. This allows for the detection of signals of positive diversifying selection linked to the presence of a discrete trait (in this case, endothermy) (*137*). BUSTED-PH tests for positive selection associated with a specific binary phenotype and distinguishes phylogeny-wide positive selection from convergently evolving genes (*60*). BUSTED-PH fits both an unrestricted dN/dS

distribution and three different null models with the likelihood ratio test to identify: 1) foreground-specific positive selection; 2) background-specific positive selection; and 3) differences between foreground and background branches (*60*).

In the tests for positive selection, the null model restricts either the foreground or background branches to $dN/dS \leq 1$, while the other taxa remain unconstrained. To test for the difference between foreground and background, a null model is fitted with a single tree-wide unrestricted dN/dS distribution. We defined signatures of selection according to which class they are a member of using the following classification: C000: No selection on FG, no selection on BG, no difference between FG and BG selective regimes; C100: Selection on FG, no selection on BG, no difference between FG and BG selective regimes; C010: No selection on FG, selection on BG, no difference between FG and BG selective regimes; C001: No selection on FG, no selection on BG, statistically significant difference between FG and BG selective regimes; C101: Selection on FG, no selection on BG, statistically significant difference between FG and BG selective regimes; C110: Selection on FG, selection on BG, no difference between FG and BG selective regimes; C011: No selection on FG, selection on BG, statistically significant difference between FG and BG selective regimes; and C111: Selection on FG, selection on BG, statistically significant difference between FG and BG selective regimes .

Additionally, BUSTED-PH allows for an optional third subset of "nuisance" branches not directly used for inference (or without a phenotype designation). This approach helped segregate the signals of selection observed in our foreground species by allowing us to split our analyses into four scenarios: endotherms (fg) vs. ectotherms (bg), regional endotherms (fg) vs. ectotherms (bg), eye/brain endotherms (fg) vs. ectotherms (bg), and full-body endotherms (fg) vs. ectotherms (bg). Unlike other programs that prune "nuisance" species, negatively affecting taxonomic sampling (*25–27*) and increasing the probability of a type I error, BUSTED-PH maintains statistical power consistency (*59*, *60*).

We discerned the functional classifications of the identified genes using the Protein Analysis Through Evolutionary Relationships (PANTHER) database (*61*). PANTHER allowed us to infer gene function by leveraging their evolutionary relationships (*61*). Furthermore, the PANTHER system categorizes genes into families and subfamilies based on their evolutionary relationships, offering a comprehensive framework to visualize and succinctly summarize the functional roles of each gene (*61*). This organizing scheme enables efficient comparisons across diverse functional contexts, enhancing our understanding of gene function. Software used for selection analysis is available in the dedicated GitHub repository at: https://github.com/veg/Endothermy.

Finally, we complemented the selection results obtained over our previously mentioned four scenarios by examining the genes showing the highest relative rates of evolution and using the HyPhy method 'RER', which is based on the *RERconverge* R package (*62*). 'RER' allows for the detection of protein-coding sequences whose evolutionary rates were associated with endothermy. This method uses four nested models: (1) Proportional- A single branch length scaling parameter (Rshared) is inferred, scaling all branches relative to the input tree. (2) Two scaling parameters (Rforeground and Rbackground) are inferred, one for the foreground branches and one for the background branches. (3) Unconstrained foreground- Background branches share a scaling factor (Rbackground), while each foreground branch has its own independent branch length. The set of nested models was compared with AIC-c used to determine the best model fit. 'RER' also considers scenarios where a single branch drives the evolutionary signal; it tests the significance of shifts in relative evolutionary rates on both the "single branch only" and "all but

this branch" for each respective foreground branch. This test for relative evolutionary rates detects genes that converge towards faster evolutionary rates, not just those under positive selection. This analysis provides further substantiation for the observed signals of convergence in positive selection.

Comparative genomics reveals convergent molecular adaptations for endothermy

Our PGLR analyses (Fig. 2S and tables S5–S8) found a significant correlation between increased body length and endothermy. While a larger body plan helps maintain metabolic processes, like endothermy, efficiently across a more extensive mass of tissue (*141*, *142*), this finding by itself cannot help explain why endothermy is present in such a small amount of fishes, given that many other species have similar or longer body plans than endothermic ray-finned fishes (table 3AB). Given the intricate interplay between the different body features, hydrodynamics, and metabolism in fishes (*143*) coupled with the swim modes observed in endothermic ray-finned fishes (table 3A), our findings appear to align with the notion that endothermy originated in ray-finned fishes due in part to ecological interactions. Furthermore, by elucidating the evolutionary correlations among swim mode, endothermy, and potentially swim speed, we gain deeper insights into the distinctive approaches to endothermy observed in various fish species. There appears to be a correlation between swim mode and the endothermic strategy implemented: less stiff swimmers, like sailfishes and marlins, exhibit endothermy in their eye/brain region; stiffer swimmers, like tunas, have a higher degree of heat production; and the stiffest swimmers, such as opahs, possess full-body endothermy. Our results suggest an interplay between faster swim speed and the intensity of muscular and cardiovascular exertion (*55*, *142*) needed to produce and sustain elevated propulsion levels using a stiffer and less hydrodynamic swim mode coupled with the high metabolic demand that endothermy represents (*1*, *4*) and the different endothermic strategies employed by fishes.

Our study examined 894 single-copy orthologs to identify selection patterns within the Marine Vertebrate Dataset (MVD), which encompasses 205 species (table S17 and datasets S9–S10). Our runs used the Protein Analysis Through Evolutionary Relationships (PANTHER) database to classify the functions of the genes we identified (*61*). It organizes genes into families and subfamilies, creating a structured framework that facilitates the visualization and summarization of each gene's role (*61*). This systematic categorization allowed us to make efficient functional comparisons, substantially broadening our comprehension of how these genes contribute to the trait of endothermy in marine vertebrates. We also looked at the relative rates of evolution for each gene. This was achieved using the used the HyPhy method 'RER', based on the R package *RERconverge* (*62*, *63*). By incorporating this approach, we added an additional layer to our phenotype-to-genotype association analysis, enabling us to not only focus on genes under positive selection but also to consider those evolving at the fastest rates.

For HyPhy analyses, our first test scenario (All endotherms–S1) identified signals of positive selection across all endothermic species (foreground or fg) compared to ectothermic species (background or bg). In total, 51 (22 in class C111 and 29 in class C101) genes showed signs of positive selection in this scenario, with 122 genes filtered by BUSTED-E (Fig. 4A, fig. S13 and S17, and tables S18-S19), meaning these genes experienced selection in the foreground and little to no selection in the background. Our 'RER' results for our positively selected genes (table S20) indicated substantial accelerated evolution in genes associated with stimuli response (GO:0050896), such as *ddx54*, which exhibited the highest evolutionary rate (table S20), *adcy8*, and *slc38a10* (Fig. 4A). Moreover, genes associated with muscular function and localization

(GO:0051179), like *ryr1*, *ift122*, and *usp25*, also demonstrated signs of positive selection and accelerated evolution, along with genes for neural development (GO:0032502) and metabolic processes (GO:0008152) like *cnep1r1, pyroxd2, mknk1, ephb2, aptx, nfya,* and *ttll11*.

In our second scenario (Regional endotherms–S2), we compared selection signals between regional endotherms (i.e., tunas and the great white shark) in our foreground, ectotherms in our background, and the other endothermic representatives as our nuisance species. The analysis revealed that 17 (3 in class C111 and 14 in class C101) genes exhibited signs of positive selection in this scenario, with 121 genes filtered by BUSTED-E (Fig. 4B, fig. S14, and tables S18 and S21), including genes responsible for metabolic processes (GO:0008152) like *pyroxd2,* a gene involved in oxidoreductase processes crucial to energy generation, which exhibited the highest level of expression (table S22) and was also observed under selection in the previous scenario. The genes *nek5*, *rasa3*, and *igdcc4,* which are primarily involved in the regulation of biological processes (GO:0050789), including movement and muscle function, and cellular metabolism and communication (GO:0007154 and GO:0044237), were positively selected as well. We also saw some selection and accelerated rates of evolution in the *map3ks* and *bbs9* genes, which are associated with stimuli response (GO:0050896). Lastly, for this scenario, we also saw scenario-exclusive selection on the genes *afap1l2* and *sufu,* which are involved in cytoskeletal signaling and the negative regulation of signaling pathways, respectively, impacting cellular proliferation, survival, motility, gene expression, and general biological processes (GO:0050789).

The third scenario (Eye/brain endotherms–S3) placed eye/brain endotherms (i.e., sailfishes, marlins, and the butterfly mackerel) in our foreground, ectotherms in our background, and the other endothermic representatives as our nuisance species. Here, only 7 (5 in class C111 and 2 in class C101) genes exhibited signs of positive selection, with 120 genes filtered by BUSTED-E (Fig. 4C, fig. S15 and S17, and tables S18 and S23). The gene with the highest evolutionary rate for this scenario (table S24) was *elapor1,* which is involved in catabolic processes (GO:0044248) and the organization of endosomes and lysosomes. In addition to *elapor1*, another gene that was exclusively found under selection in this scenario (table S18) includes *lamp2,* which is involved in muscular function and localization (GO:0051179) by maintaining muscular function and health by facilitating autophagy processes through its isoform *lamp-2b*. Two other genes responsible for muscle function and regulation that were found under selection in this and the previous scenario (table S18) were *ppp1r12a* and *carnmt1*, which are key to muscle contraction and metabolism (GO:0008152), with the former regulating myosin phosphatase activity and the latter involved in the methylation of carnosine, contributing to muscle function and health. Lastly, the gene *dcaf6* also showed signs of positive selection in other scenarios (table S18). This gene is involved primarily in stimuli response (GO:0050896), more specifically in the primary cellular response to stimuli by acting as a coactivator to nuclear receptors.

Lastly, the fourth scenario (Full-body endotherms–S4) had a foreground composed of full-body endotherms (i.e., marine mammals, penguins, the leatherback sea turtle, and opahs), ectotherms in our background, and the other endothermic representatives as our nuisance species (Fig. 4D and fig. S17). Results for this scenario indicate 75 (38 in class C111 and 37 in class C101) genes under positive selection, with 108 genes filtered by BUSTED-E (tables S18 and S25), with the gene *pdlim7,* which is involved in actin organization (GO:0030036), heart development (GO:0007507), and muscle structure (GO:0061061), showing the highest evolutionary rate in our foreground (fig. S16 and table S26). Other notable scenario-exclusive genes include *dpys*, which is primarily involved in catabolic processes (GO:0006208), and *aco2*, which is involved in

metabolic processes (GO:0008152), more specifically in the generation of precursor metabolites and energy (GO:0006091). Additionally, the genes mt-nd1 and ndufb1 were identified under selection, both of which are integral to metabolic processes (GO:0008152) related to energy generation within the mitochondria, with *mt-nd1* being a part of complex I's oxidative phosphorylation pathway and *ndufb1* involved in the electron transfer process. The gene *sik3* was also found under selection, known for its role in mTOR signaling and chondrocyte differentiation, implicating it in metabolic processes and cellular communication (GO:0008152 and GO: 0007154).

We found instances of convergent genes among various marine endothermic lineages, suggesting shared genetic strategies for acquiring warm-bloodedness despite the differences in the implementation of this strategy and their evolutionary histories (Fig. 4E and fig. S18). More specifically, when looking at two of our scenarios (S2 and S3), we see that the genes *carnmt1*, involved in the methylation of carnosine contributing to muscle function, and *dcaf6*, which acts as a coactivator to nuclear receptors, were positively selected and showed signs of accelerated relative rates regardless of the difference in endothermy type and differences in species considered in those scenarios. These findings highlight the shared genetic strategies, primarily regarding muscle function, for acquiring warm-bloodedness. Additionally, our individual scenario findings provide further evidence that the genomic origins of endothermy in marine vertebrates predominantly lie in metabolic processes, muscular development, regulatory pathways, and neural function/stimuli response. When comparing our results to the 61 previously identified candidate genes associated with endothermy and endothermic-mediated physiological processes (*1*, *4*, *25–27*), only five (*ryr1*, *aco2*, *ephb2*, *mt-nd2*, and *pyroxd2*) were confirmed under the more rigorous BUSTED-PH approach (Methods). The remaining genes showed selection but were not statistically significant or showed selection in the background with no major differences compared to the foreground. Furthermore, when analyzing the relative evolutionary rates without accounting exclusively for genes under positive selection, we see that for S1, S2, and S4, the genes with the highest rate (*brk1*, *ap2s1*, and *smyd3*, respectively) were not under positive selection. While not accounting for signals of positive selection and only looking at relative rates of evolution would yield a notably higher number of convergent genes, it is important to note that these higher rates for non-positively selected genes might be the result of unique phenotypes for that specific foreground (other than endothermy) that could influence the rates observed in the analyses (*63*). Hence, we favored the more conservative approach by highlighting genes that were both positively selected and had a higher relative rate of evolution. Future studies examining gene duplications and losses and exploring regulatory or non-coding regions could offer additional insights into the genomic basis of the evolution of endothermy.

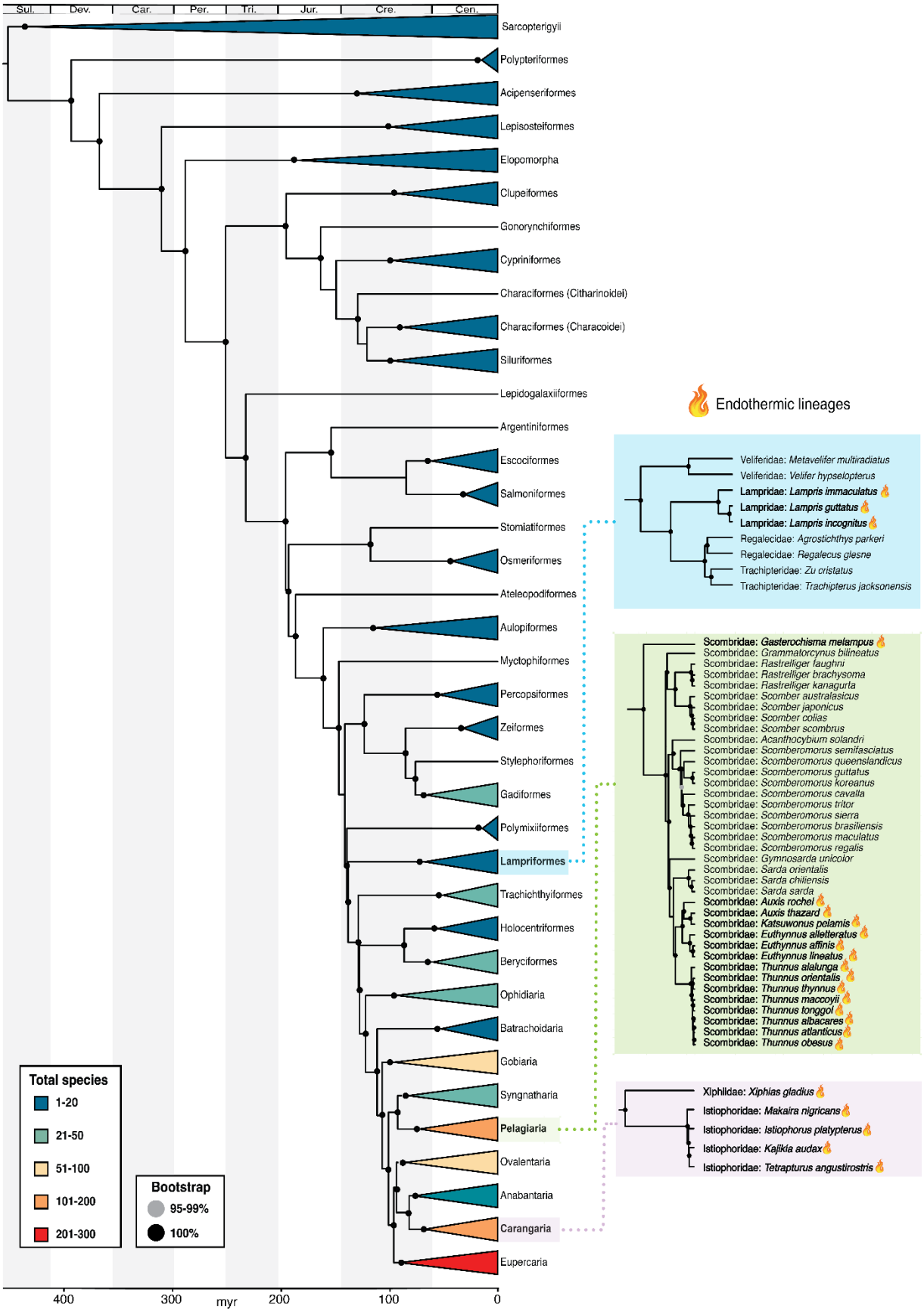
Sul.
Dev.
Car.
Per.
Tri.
Jur.
Cre.
Cen.
Sarcopterigyii
Polypteriformes
Acipenseriformes
Lepisosteiformes
Elopomorpha
Clupeiformes
Gonorynchiformes
Cypriniformes
Characiformes (Citharinoidei)
Characiformes (Characoidei)
Siluriformes
Lepidogalaxiiformes
Argentiniformes
Escociformes
Salmoniformes
Stomiatiformes
Osmeriformes
Ateleopodiformes
Aulopiformes
Myctophiformes
Percopsiformes
Zeiformes
Stylephoriformes
Gadiformes
Polymixiiformes
Lampriformes
Trachichthyiformes
Holocentriformes
Beryciformes
Ophidiaria
Batrachoidaria
Gobiaria
Syngnatharia
Pelagiaria
Ovalentaria
Anabantaria
Carangaria
Eupercaria
400
300
myr
200
100
0
Endothermic lineages
Veliferidae: Metavelifer multiradiatus
Veliferidae: Velifer hypselopterus
Lampridae: Lampris immaculatus
Lampridae: Lampris guttatus
Lampridae: Lampris incognitus
Regalecidae: Agrostichthys parkeri
Regalecidae: Regalecus glesne
Trachipteridae: Zu cristatus
Trachipteridae: Trachipterus jacksonensis
Scombridae: Gasterochisma melampus
Scombridae: Grammatorcynus bilineatus
Scombridae: Rastrelliger faughni
Scombridae: Rastrelliger brachysoma
Scombridae: Rastrelliger kanagurta
Scombridae: Scomber australasicus
Scombridae: Scomber japonicus
Scombridae: Scomber colias
Scombridae: Scomber scombrus
Scombridae: Acanthocybium solandri
Scombridae: Scomberomorus semifasciatus
Scombridae: Scomberomorus queenslandicus
Scombridae: Scomberomorus guttatus
Scombridae: Scomberomorus koreanus
Scombridae: Scomberomorus cavalla
Scombridae: Scomberomorus tritor
Scombridae: Scomberomorus sierra
Scombridae: Scomberomorus brasiliensis
Scombridae: Scomberomorus maculatus
Scombridae: Scomberomorus regalis
Scombridae: Gymnosarda unicolor
Scombridae: Sarda orientalis
Scombridae: Sarda chiliensis
Scombridae: Sarda sarda
Scombridae: Auxis rochei
Scombridae: Auxis thazard
Scombridae: Katsuwonus pelamis
Scombridae: Euthynnus alletteratus
Scombridae: Euthynnus affinis
Scombridae: Euthynnus lineatus
Scombridae: Thunnus alalunga
Scombridae: Thunnus orientalis
Scombridae: Thunnus thynnus
Scombridae: Thunnus maccoyii
Scombridae: Thunnus tonggol
Scombridae: Thunnus albacares
Scombridae: Thunnus atlanticus
Scombridae: Thunnus obesus
Xiphiidae: Xiphias gladius
Istiophoridae: Makaira nigricans
Istiophoridae: Istiophorus platypterus
Istiophoridae: Kajikia audax
Istiophoridae: Tetrapturus angustirostris
Total species
1-20
21-50
51-100
101-200
201-300
Bootstrap
95-99%
100%

Fig. S1. Summarized maximum likelihood phylogeny of 1,051 species of ray-finned fishes. Collapsed tips represent either an order or a series of ray-finned fishes (except the outgroup). Clades containing endothermic species are expanded, and the endothermic representatives appear in boldface. The color of the collapsed branches represents the number of species obtained from that order/series. Dates were congruified from an MCMC time-calibrated phylogeny and supplemented with eight additional calibrations, informing divergence times for major clades, including Pelagiaria, Carangaria, Anabantaria, and Syngnatharia.

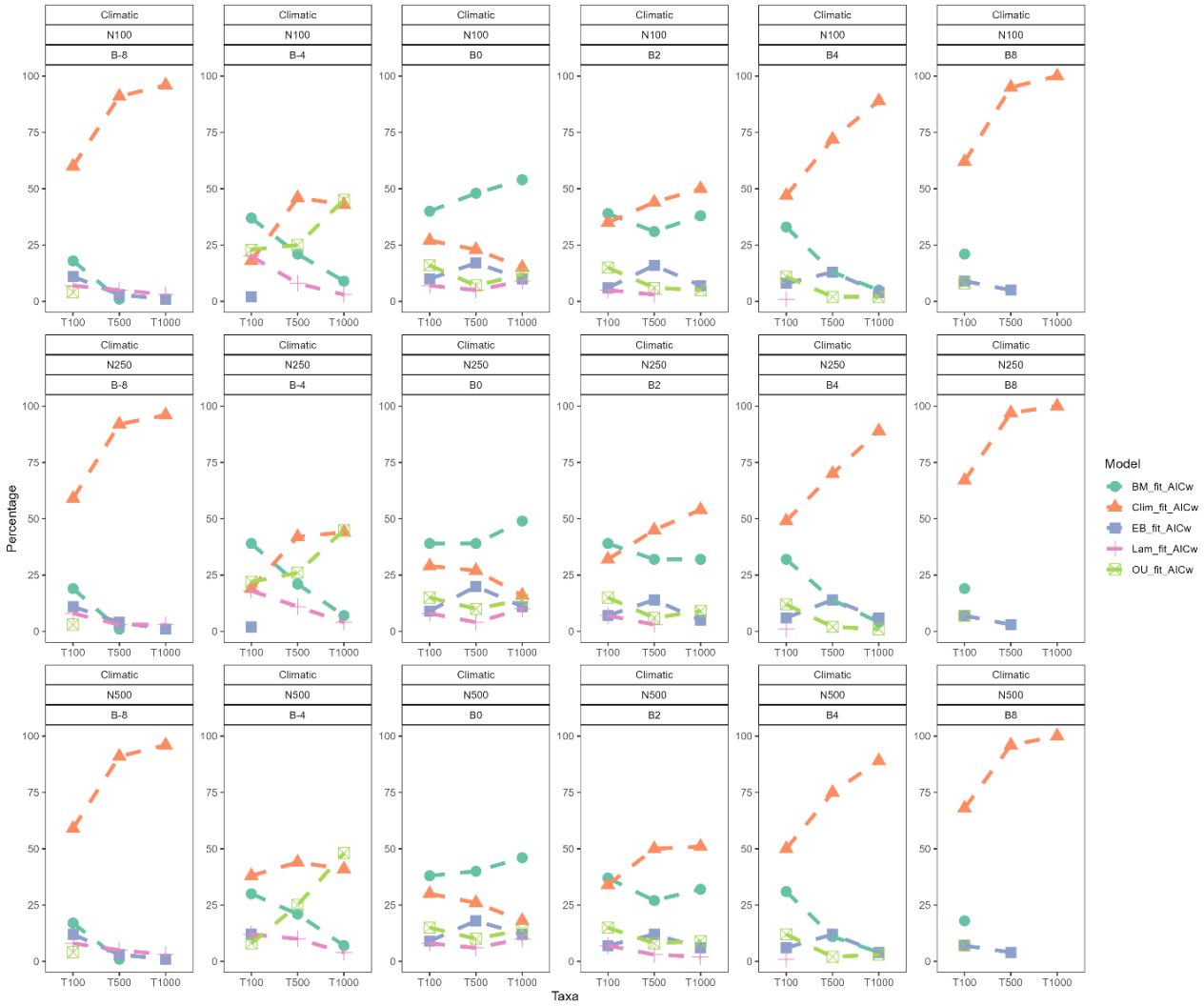

Fig. S2. Model fitting results simulated under our model with Climatic data based on the weighted Akaike scores for each model (AICw). Each plot represents a different integration level (N100, N250, N500) and the parameter estimated (β = -8, -4, 0, 2, 4, and 8). The simulation illustrates how the percentage (y-axis) of the false discovery rates vary in terms of different taxonomic sampling (x-axis) over different integration levels.

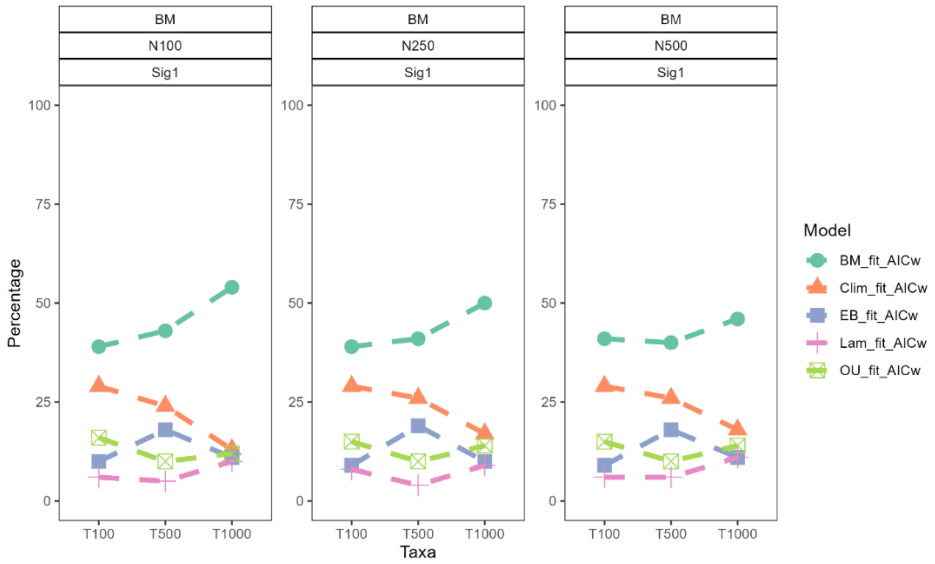

Fig. S3. Model fitting results simulated under Brownian Motion (BM) based on the weighted Akaike scores for each model (AICw). Each plot represents a different integration level (N100, N250, N500) and the parameter simulated (σ^2 =1). The simulation illustrates how the percentage (y-axis) of the false discovery rates vary in terms of different taxonomic sampling (x-axis) over different integration levels.

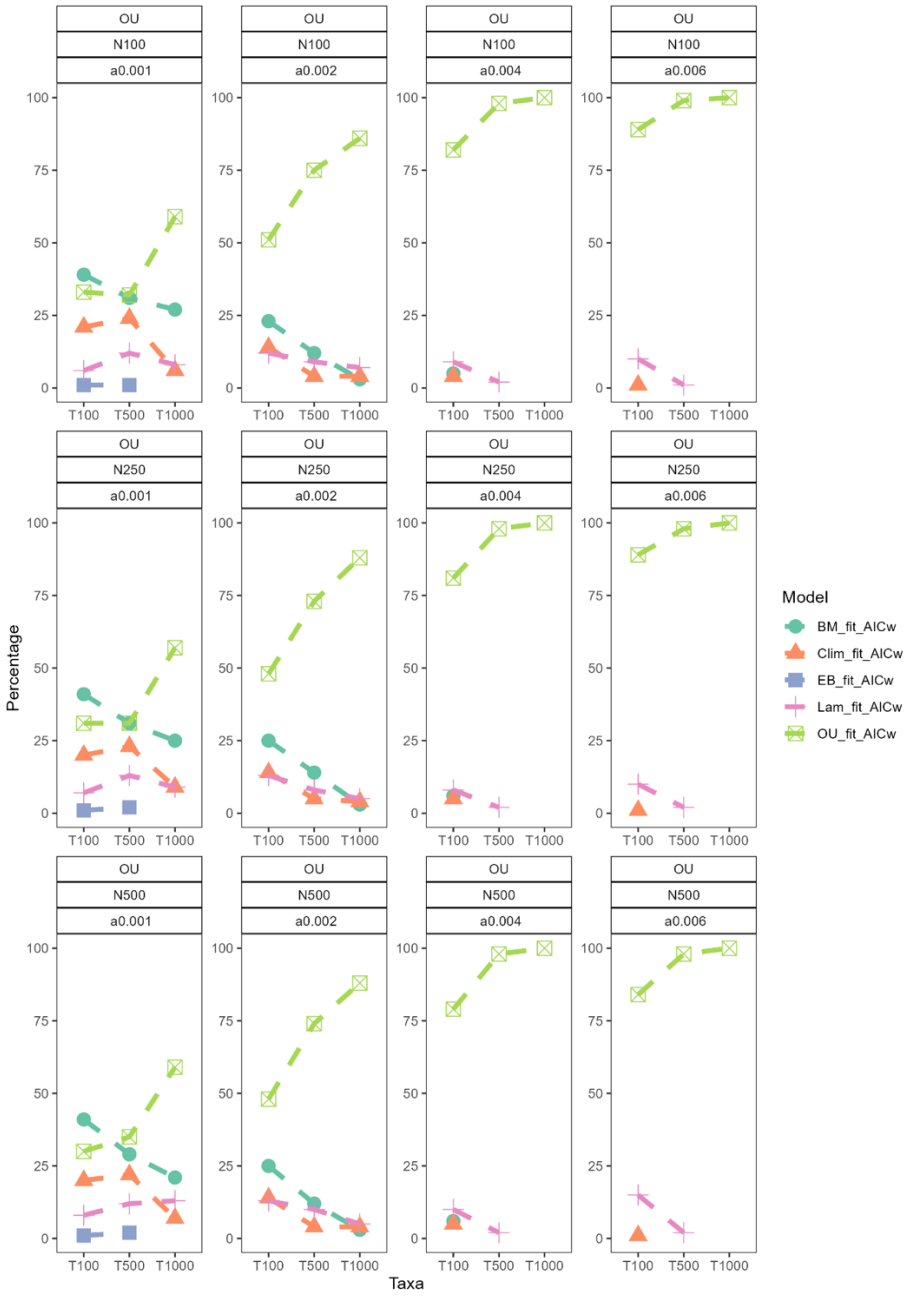

Fig. S4. Model fitting results simulated under the Ornstein-Uhlenbeck model (OU) based on the weighted Akaike scores for each model (AICw). Each plot represents a different integration level (N100, N250, N500) and the parameter simulated ($\alpha = 0.001, 0.002, 0.004, 0.006$). The simulation illustrates how the percentage (y-axis) of the false discovery rates vary in terms of different taxonomic sampling (x-axis) over different integration levels.

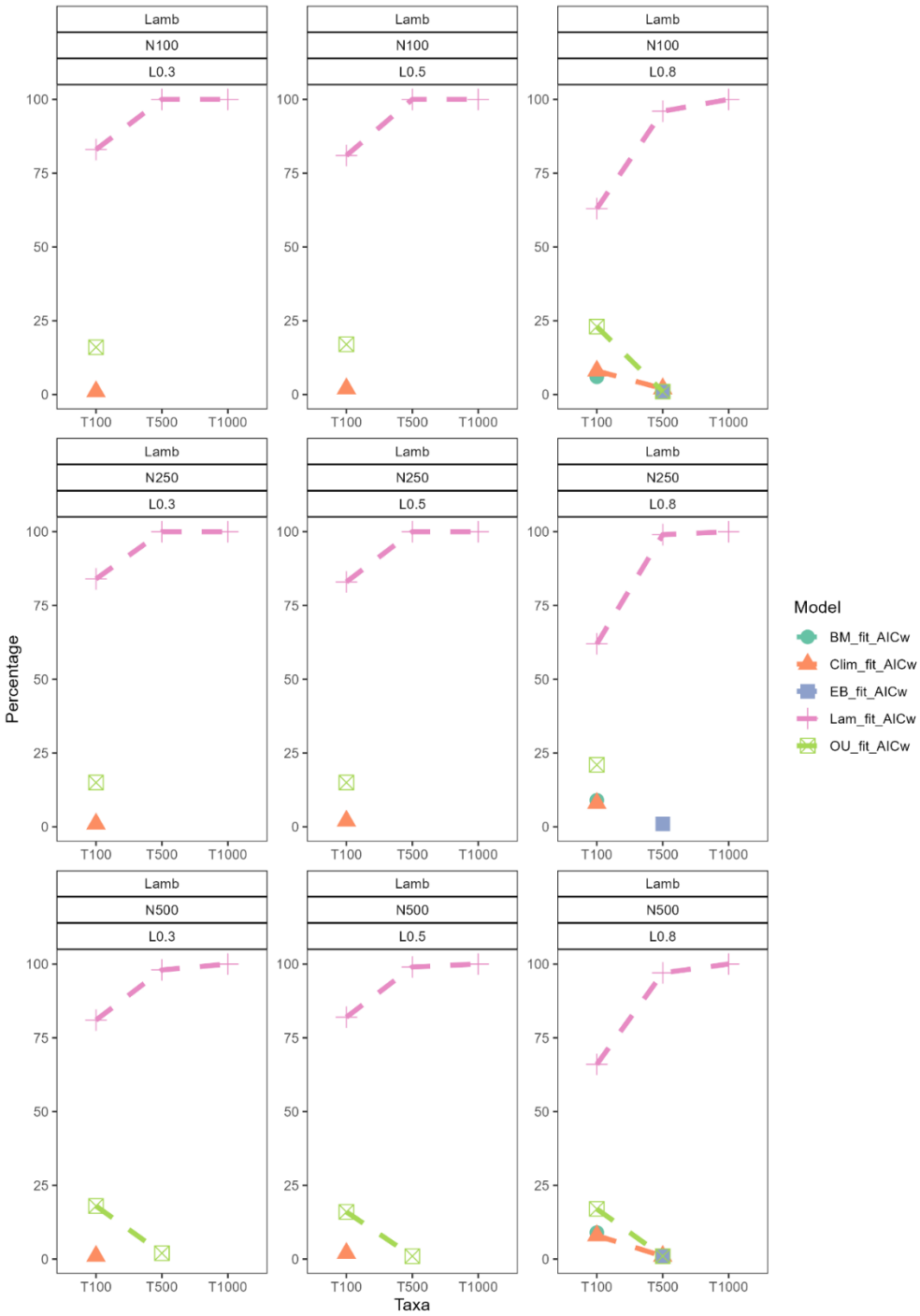

Fig. S5. Model fitting results simulated under the Lambda model (Lamb) based on the weighted Akaike scores for each model (AICw). Each plot represents a different integration level (N100, N250, N500) and the parameter simulated (λ = 0.3, 0.5, 0.8). The simulation illustrates how the percentage (y-axis) of the false discovery rates vary in terms of different taxonomic sampling (x-axis) over different integration levels.

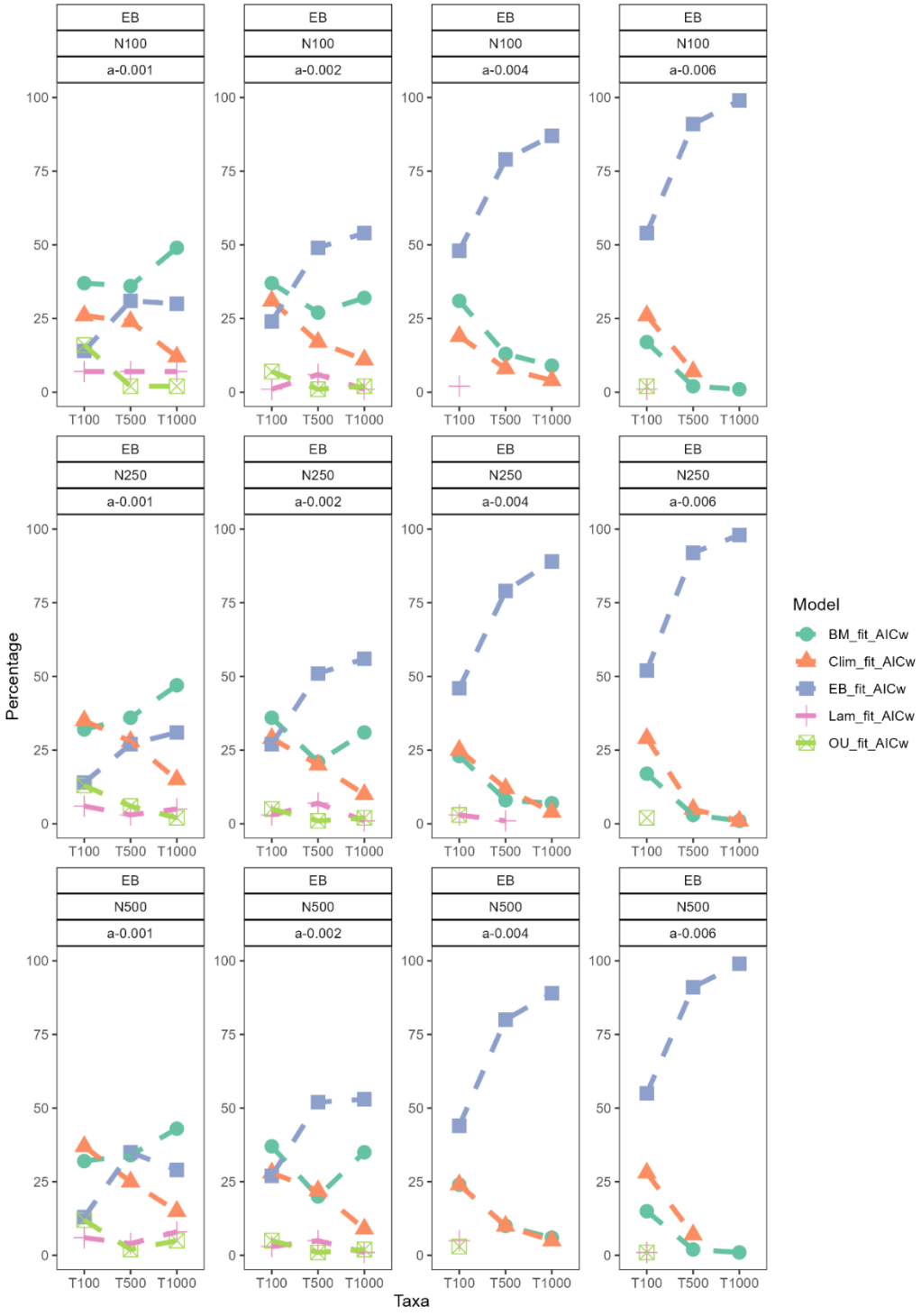

Fig. S6. Model fitting results simulated under the Early Burst model (EB) based on the weighted Akaike scores for each model (AICw). Each plot represents a different integration level (N100, N250, N500) and the parameter simulated (α = -0.001, -0.002, -0.004, -0.006). The simulation illustrates how the percentage (y-axis) of the false discovery rates vary in terms of different taxonomic sampling (x-axis) over different integration levels.

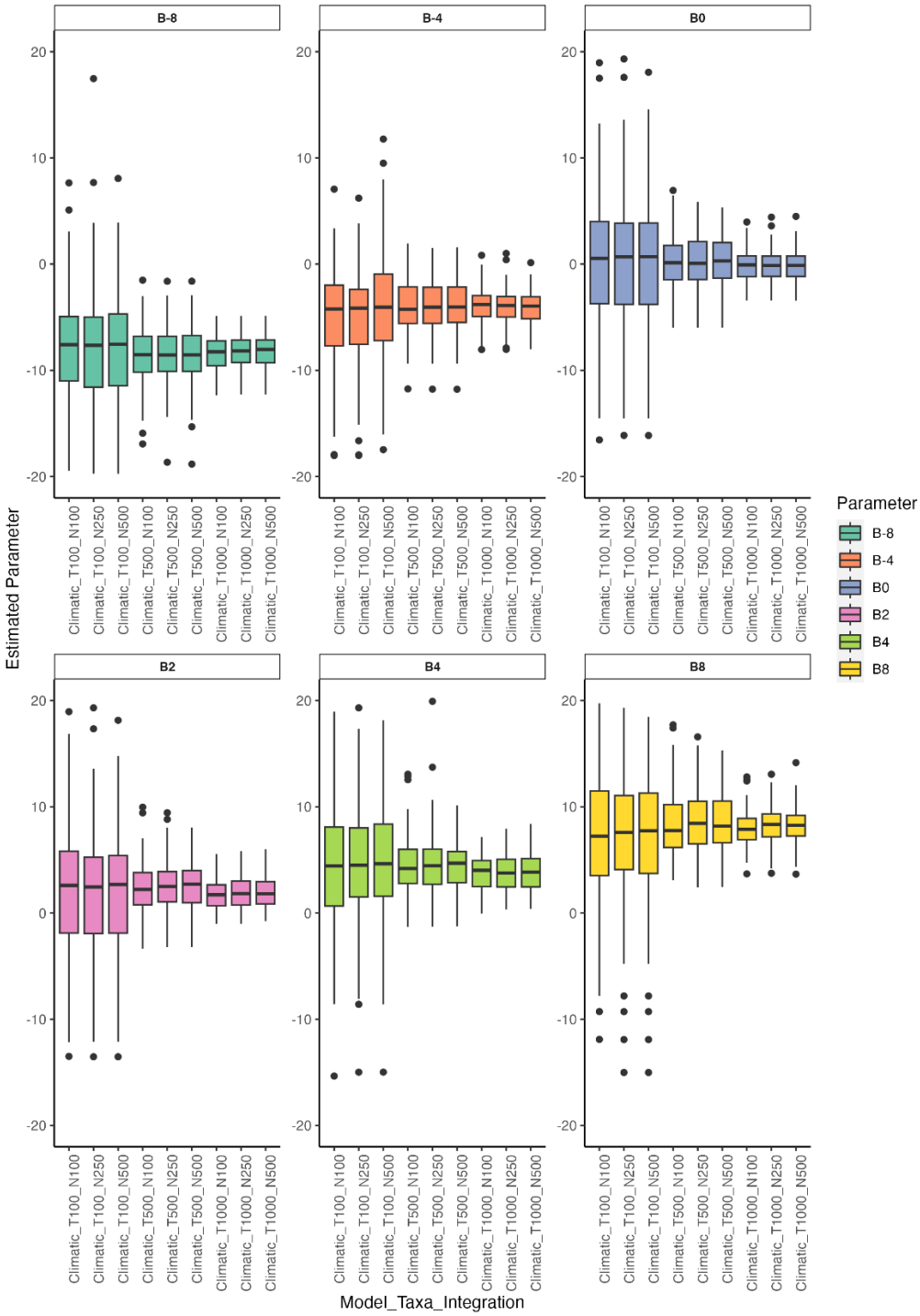

Fig. S7. Distribution of estimated β values for our climatic model. Each plot represents a different value of β and its changes over increasing taxonomic sampling (T100, T500, and T1000) and integration values (N100, N250, and N500).

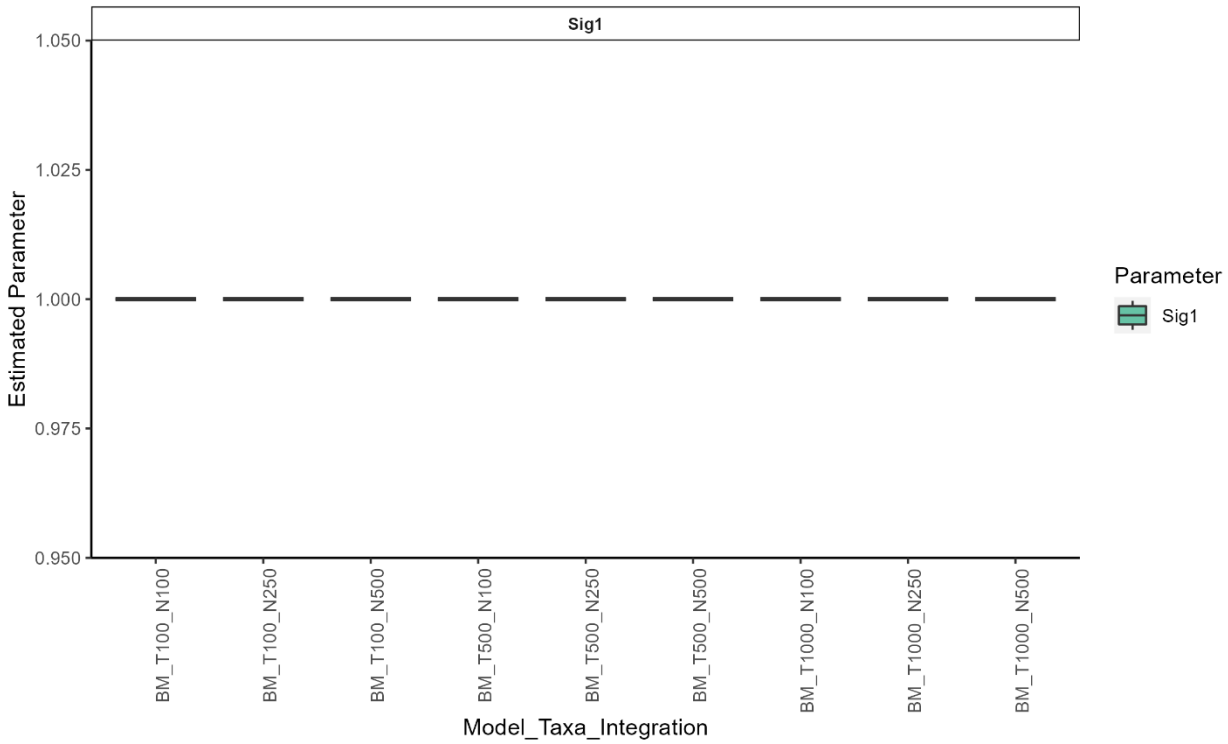

Fig. S8. Distribution of estimated σ values for the Brownian Motion (BM) model. For this plot we only used one value for sigma (σ=1). There is no observable distribution due to the Brownian motion estimation σ =1 being estimated correctly on all completed runs.

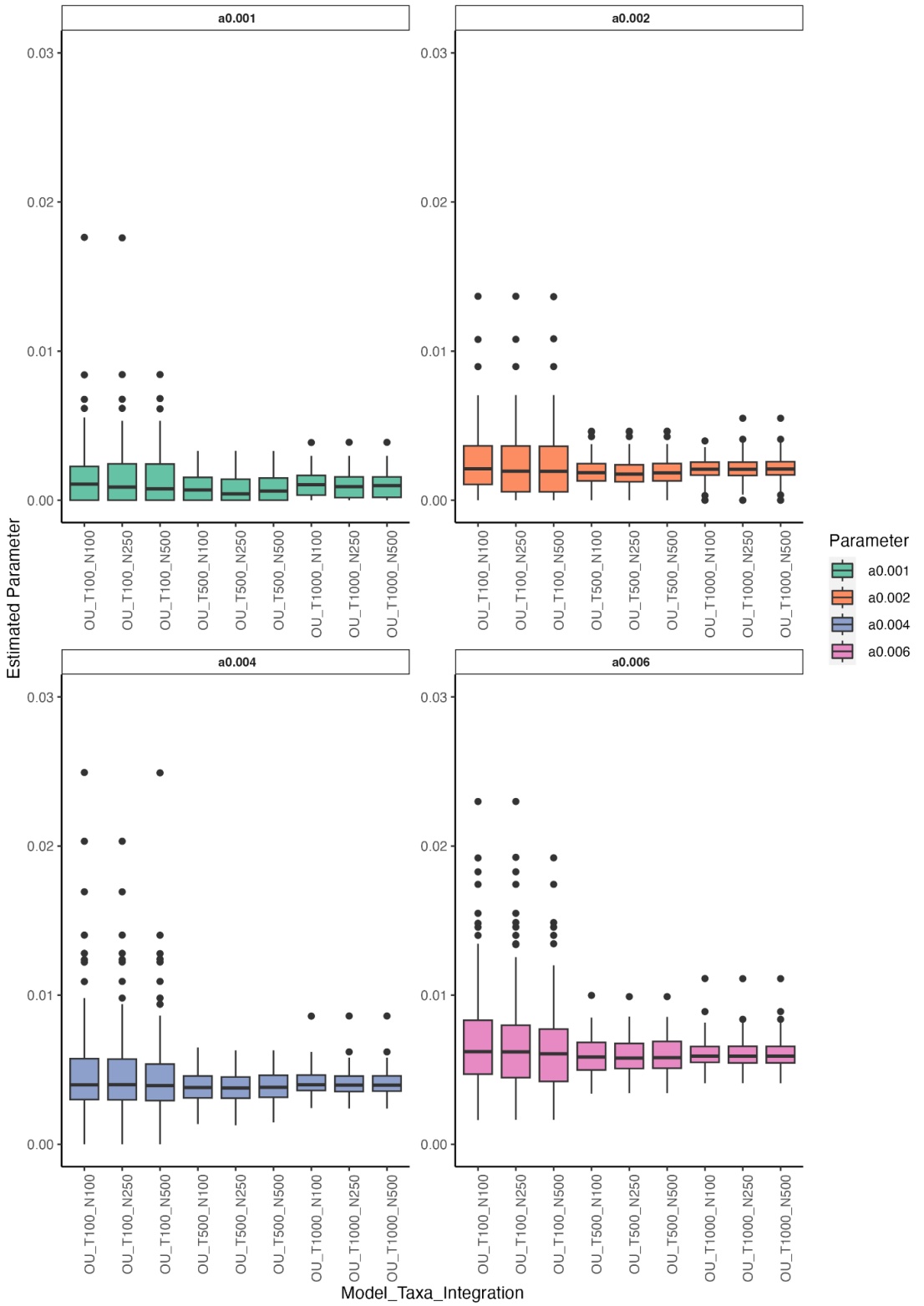

Fig. S9. Distribution of estimated α values for the Ornstein-Uhlenbeck (OU) model. Each plot represents a different value of α and its changes over increasing taxonomic sampling (T100, T500, and T1000) and integration values (N100, N250, and N500).

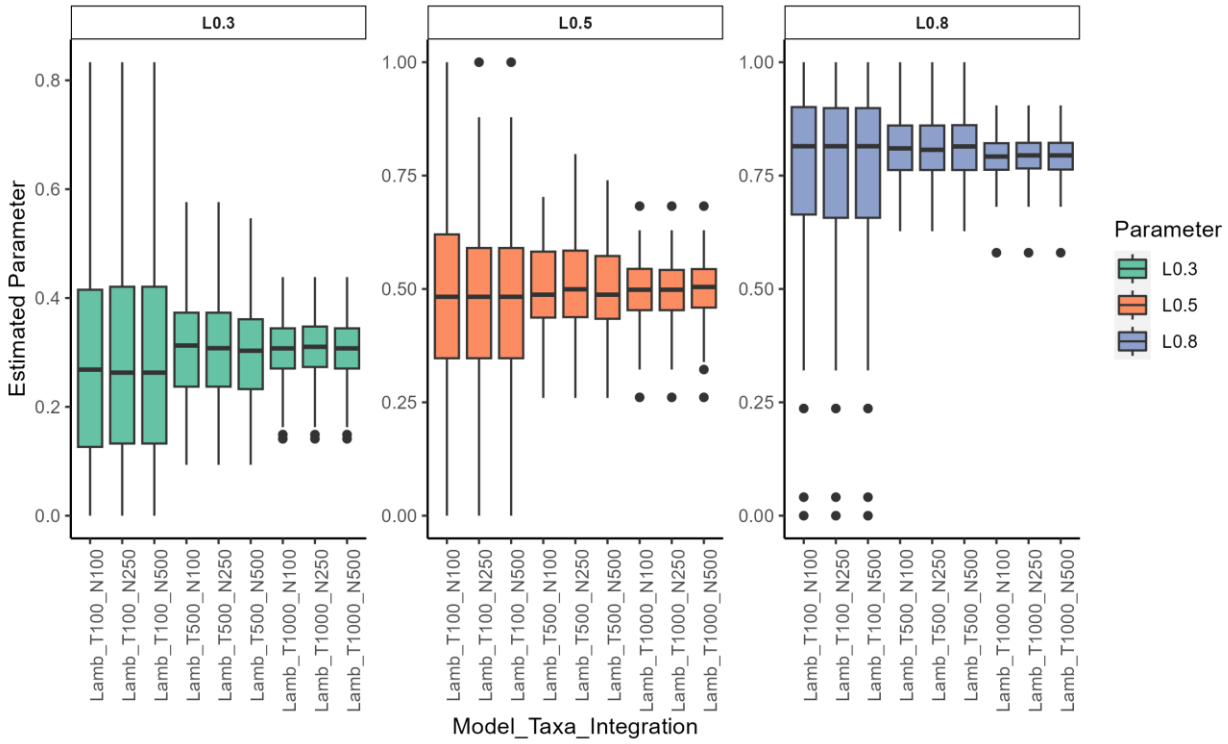

Fig. S10. Distribution of estimated λ values for the Lambda (Lamb) model. Each plot represents a different value of λ and its changes over increasing taxonomic sampling (T100, T500, and T1000) and integration values (N100, N250, and N500).

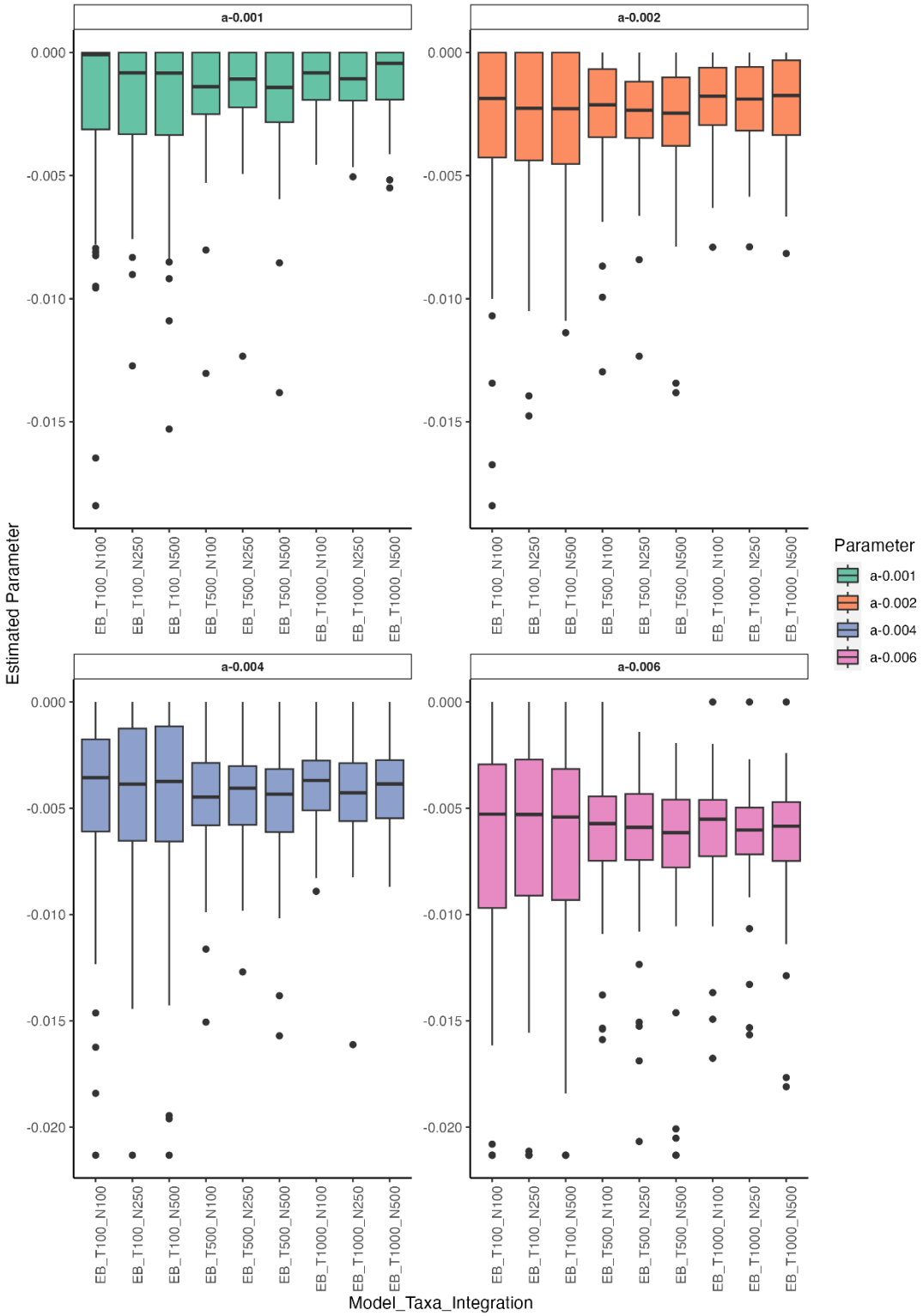

Fig. S11. Distribution of estimated α values for the Early Burst (EB) model. Each plot represents a different value of α and its changes over increasing taxonomic sampling (T100, T500, and T1000) and integration values (N100, N250, and N500).

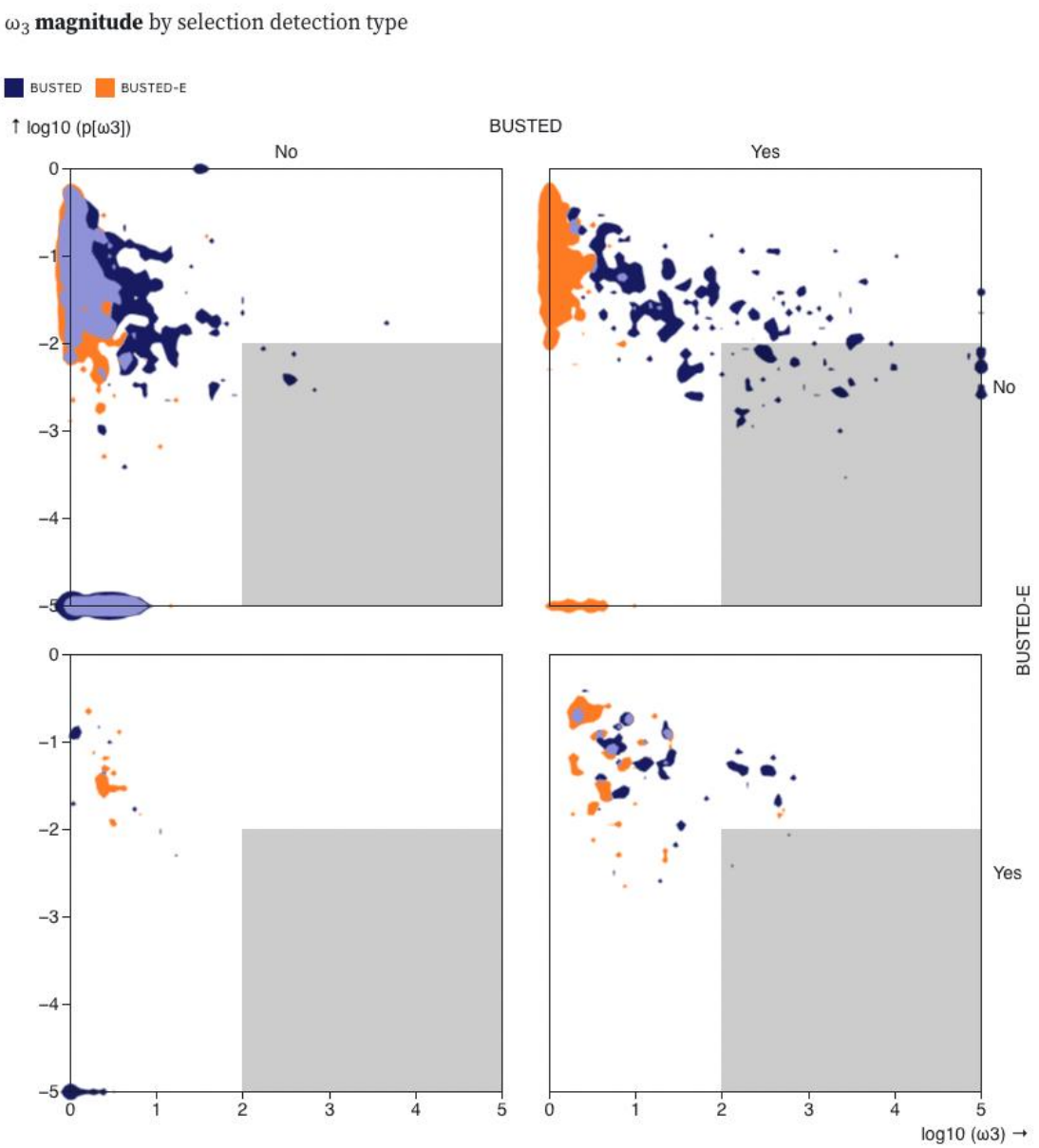

Fig. S12. Plot representing the ω3 magnitude by selection detection type, which shows that in cases where BUSTED detected positive selection, but BUSTED-E did not (top right, yes/no), the omega values for BUSTED-E shifted much lower than they were for BUSTED making them far more reliable and resistant against false positives.

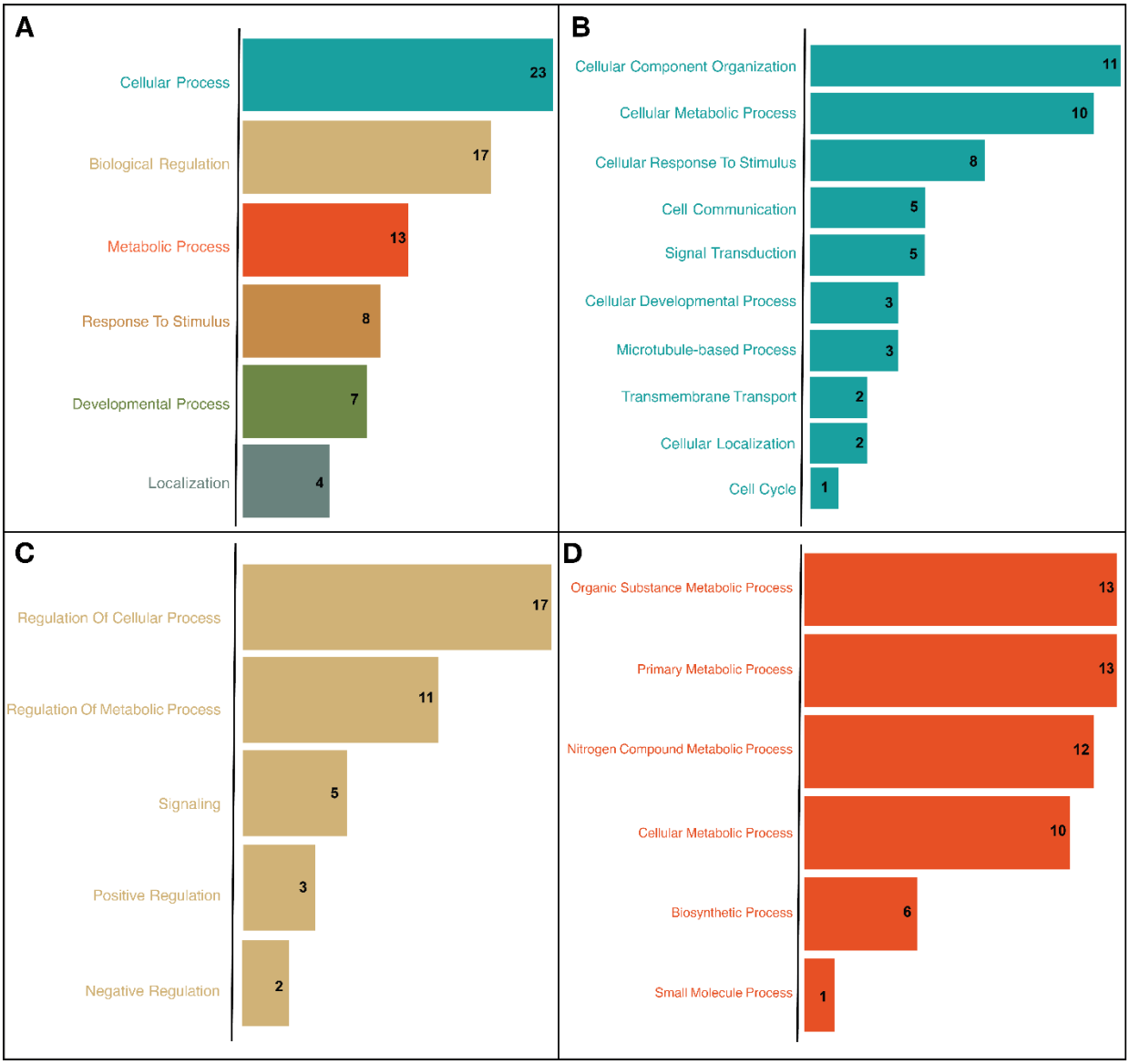

Fig. S13. Panther GO term categories based on Biological Process for genes under positive selection on our first scenario which included all endothermic species in the foreground and all ectothermic species in the background. The labels **(A)** represent the general functions associated with the Biological Process category, **(B)** shows a more detailed look into our dominant function, Cellular Process, **(C)** represents our second most dominant function, Biological Regulation, and **(D)** shows our third most abundant function, Metabolic Process. A total of 17 genes were not included in this figure because Panther was unable to find a proper category for them.

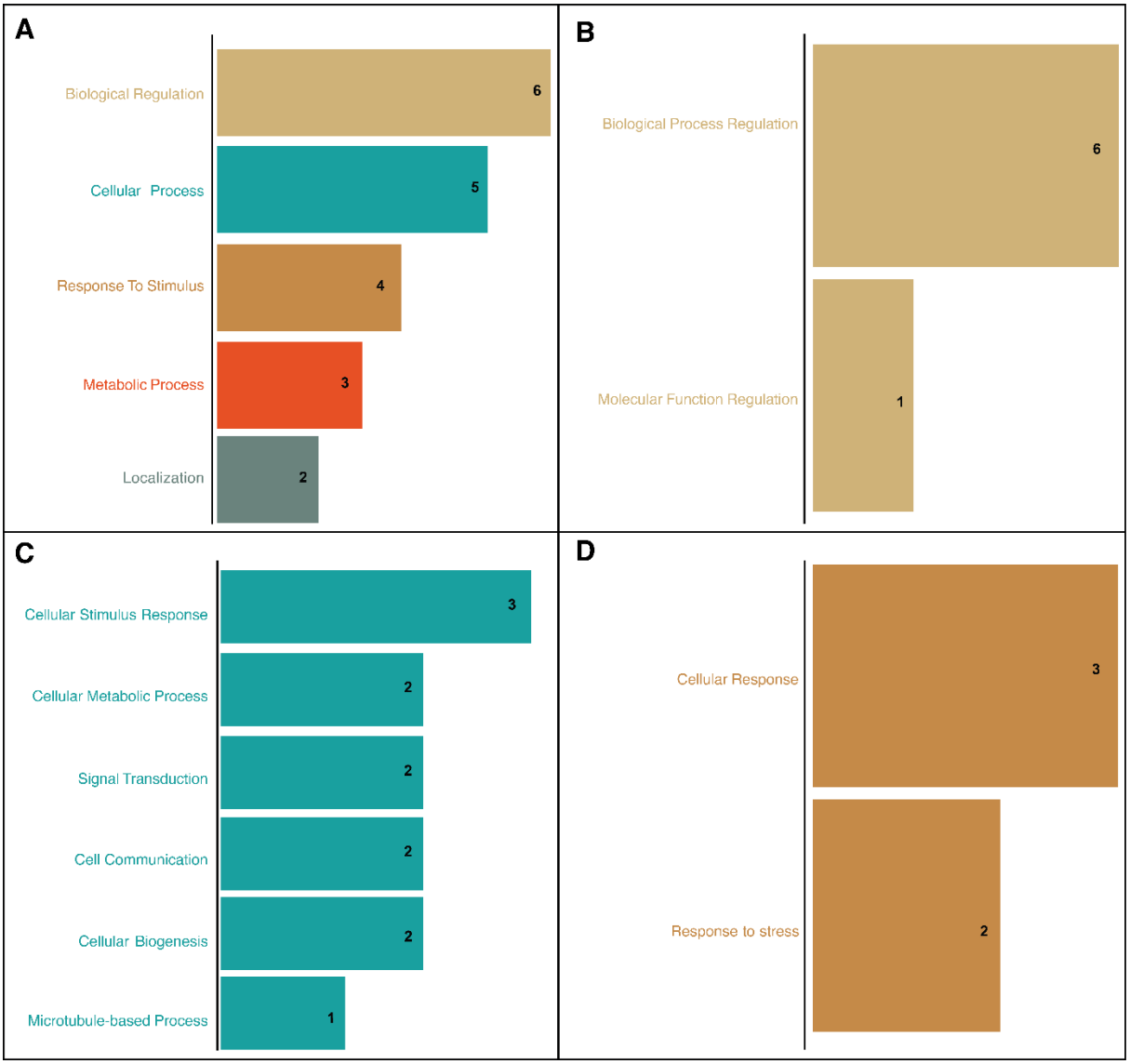

Fig. S14. Panther GO term categories based on Biological Process for genes under positive selection on our second scenario, which included regional endotherms (tunas and great white shark) in the foreground, the rest of the endothermic taxa as nuisance species, and all ectothermic species in the background. The labels **(A)** represent the general functions associated with the Biological Process category, **(B)** shows a more detailed look into our dominant function, Biological Regulation, **(C)** represents our second most dominant function, Cellular Process, and **(D)** shows our third most abundant function, Metabolic Process. A total of 6 genes were not included in this figure because Panther was unable to find a proper category for them.

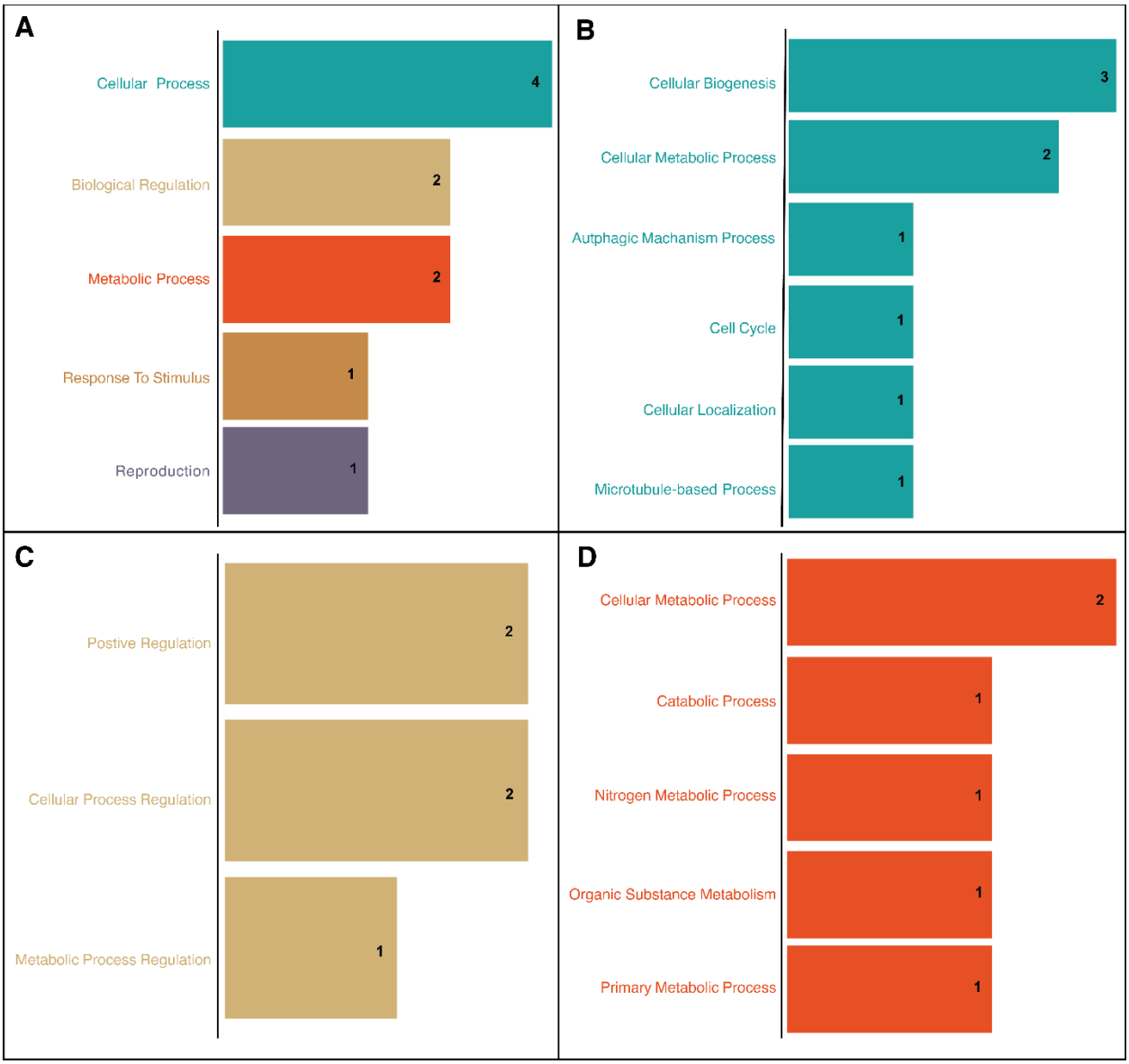

Fig. S15. Panther GO term categories based on Biological Process for genes under positive selection on our second scenario, which included eye/brain endotherms (billfishes and the butterfly mackerel) in the foreground, the rest of the endothermic taxa as nuisance species, and all ectothermic species in the background. The labels **(A)** represent the general functions associated with the Biological Process category, **(B)** shows a more detailed look into our dominant function, Cellular Process, **(C)** represents our second most dominant function, Biological Regulation, and **(D)** shows our fourth most abundant function, Metabolic Process. A total of 2 genes were not included in this figure because Panther was unable to find a proper category for them.

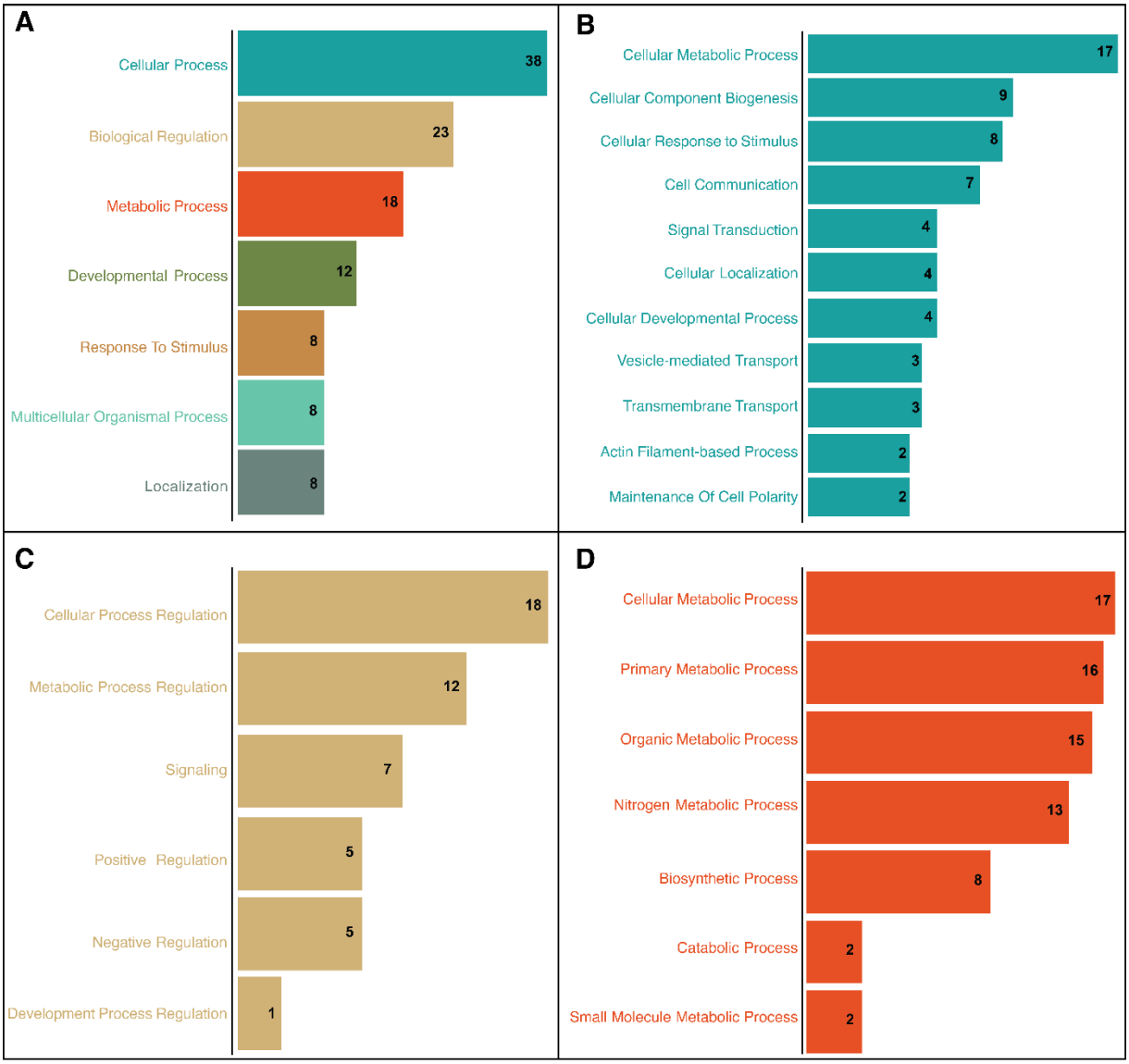

Fig. S16. Panther GO term categories based on Biological Process for genes under positive selection in our fourth scenario, which included full-body endotherms (tunas and great white shark) in the foreground, the rest of the endothermic taxa as nuisance species, and all ectothermic species in the background. The labels **(A)** represent the general functions associated with the Biological Process category, **(B)** shows a more detailed look into our dominant function, Cellular Process, **(C)** represents our second most dominant function, Biological Regulation, and **(D)** shows our third most abundant function, Metabolic Process. A total of 27 genes were not included in this figure because Panther was unable to find a proper category for them.

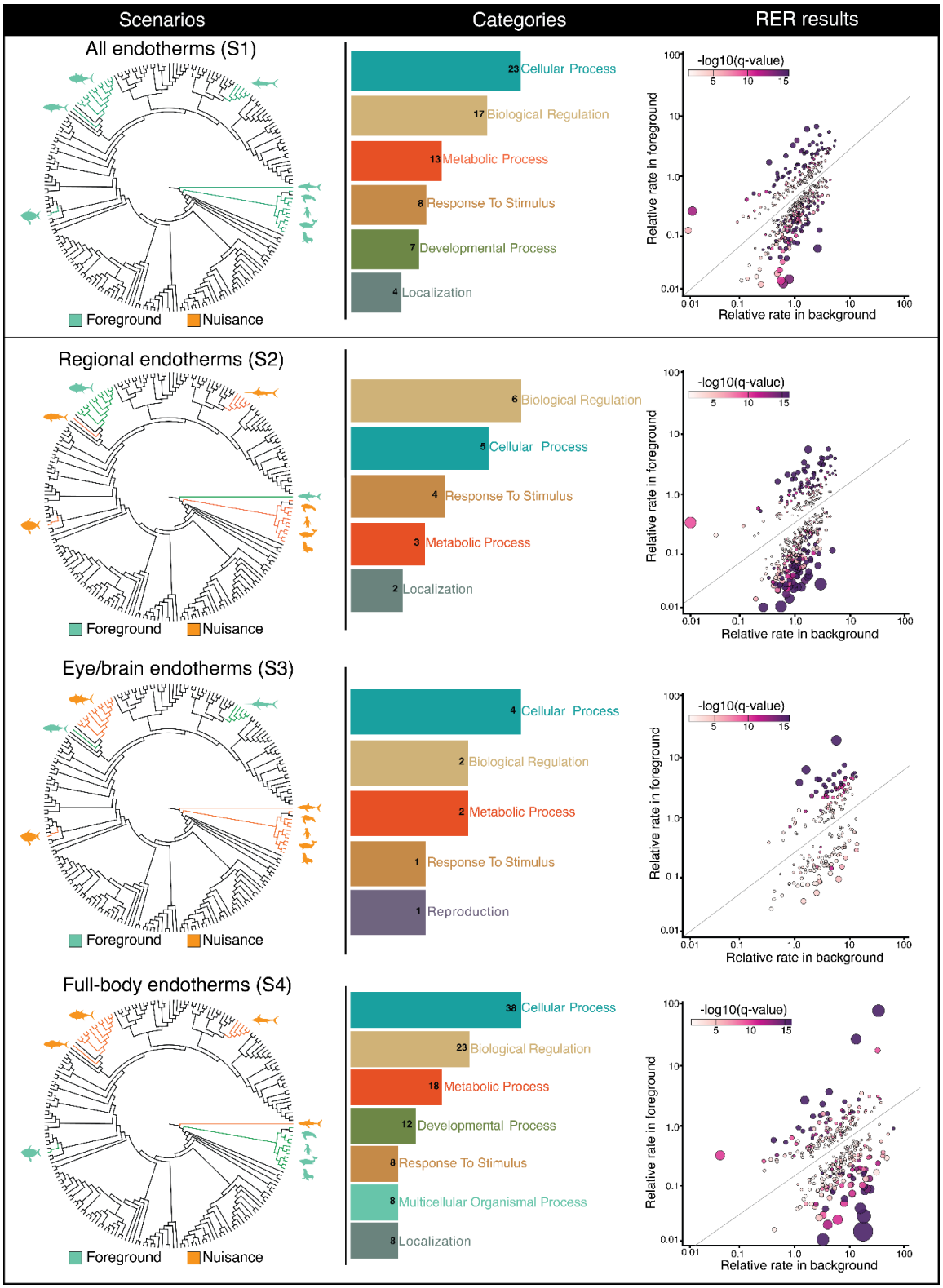

Fig. S17. Comparative genomic analyses for each of our studied scenarios**.** The first column represents the sampling scheme of each scenario. The second column contains the Panther GO categories for Biological Process in which our genes under selection fell based on each of our scenarios. The last column represents the relative amino acid rates (using a genome-average tree

as a reference) for the respective foreground (fg) and background (bg) groups for each scenario based on branches in the tree for individual genes. Points above the diagonal are accelerated on fg relative to bg. The color of each dot shows the FDR-corrected q-value (as -log10 (q)) for the rates being different between fg and bg. The bubble size is proportional to max [fg/bg, bg/fg].

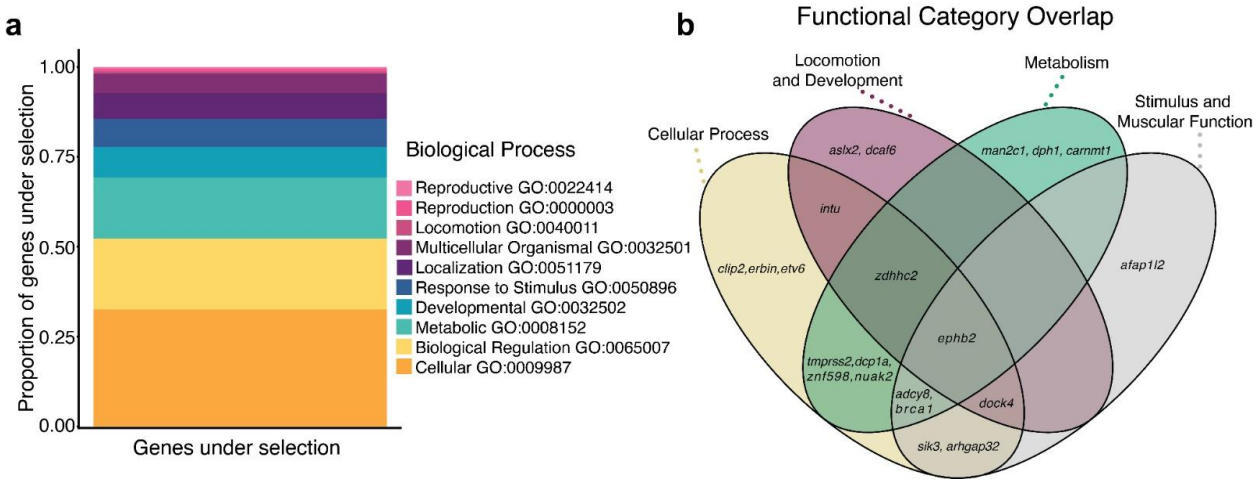

Fig. S18. Gene function and overlap based on Biological Process. **(A)** Plot representing the Panther GO term categories based on Biological Process for genes under positive selection across all of our scenarios and the proportion of genes that fall into each respective category. **(B)** Venn diagram depicting functional category overlap among positively selected genes across all scenarios. Each function represents a category associated with Panther GO terms based on Biological Processes.

Supplementary Tables (Separate file)

Table S1: List of specimens that compose the Ray-finned Fishes Database (RFFD). This table includes the family, genus, and species of the each studied individual alongside their respective museum catalog number, the institution affiliated with the sample, and the code of that institution. Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S2. Comparison of previously proposed ages for different higher-level clades. The second-to-last column shows the average age obtained from all previous studies, while the last column corresponds to the assigned age on our revised calibration file. Ages that appear in bold reflect the most probable and accurate date for the clade based on the nature of the study that proposed it, and were considered as the final ages regardless of the average value when accounting for other studies. Ages that have an asterisk (*) represent a potential outlier for the age. Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S3. Multi-factorial data matrix for 1051 species of ray-finned fishes. The table includes information about the species being analyzed, the depth range (m), average standard length (cm), length-weight relationship (g), diet, swimming mode, presence and absence of endothermy, and endothermy type for each species when applicable. Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S4. Multi-factorial data matrix for morphometrical characteristics for 542 species of ray-finned fishes obtained from the FishShapes database. The table includes information about the species being analyzed, the average standard length (cm), average body depth (cm), average width(cm), average head depth (cm), average mouth width corrected for size (cm), average caudal peduncle width (cm), average weight (g), body depth-length ratio (cm), body length-weight ratio, body length-depth-weight ratio (cm^2/g), and presence and absence of endothermy. Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S5. Outcomes of the univariate phylogenetic generalized logistic regression analyses conducted on distinct phylogenetic trees. The columns within the table represent the following: the analyzed factor, the FishShapes indicator, the number of taxa considered for analysis, the corresponding p-value, the p-value accounting for the false discovery rate (Q-Value), the phylogenetic signal (α), and the Akaike score. The two swimming mode variations correspond to the two possible assignments of swimming mode for istiophorids: sub-carangiform (Variation 1), and tunniform (Variation 2). Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S6. Outcomes of the multivariate phylogenetic generalized logistic regression analyses conducted on distinct phylogenetic trees. The columns within the table represent the following: the analyzed factor, the number of taxa considered for analysis, the corresponding p-value, he phylogenetic signal (α), and the Akaike score. The two swimming mode variations correspond to the two possible assignments of swimming mode for istiophorids: sub-carangiform (Variation 1), and tunniform (Variation 2). Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S7. Averaged p-values of the univariate phylogenetic generalized logistic regression analyses conducted on distinct phylogenetic trees. The columns within the table represent the analyzed factor and the corresponding avereaged p-value. The two swimming mode variations correspond to the two possible assignments of swimming mode for istiophorids: sub-carangiform

(Variation 1), and tunniform (Variation 2). Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S8. Averaged p-values of the multivariate phylogenetic generalized logistic regression analyses conducted on distinct phylogenetic trees. The columns within the table represent the swimming mode variation type (sub-carangiform (Variation 1) and tunniform (Variation 2)), the analyzed factor, and the corresponding averaged p-values. Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S9. Model fitting results for the 16,200 simulations conducted using our climatic model. The table includes the model generating the data (Model), the number of species (Taxa), the number of integrations (Integrations), the parameter being generated (Parameter), the Akaike scores for each model fit (AIC), the estimated parameter (Est. Param), the mu value estimated (μ), and the weighted Akaike scores for each model fit (AICw). Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S10. Model fitting results for each phylogenetic reconstruction of our RFFD. The table includes the phylogenetic tree being assessed (Tree), the paleoclimatic curve used in the analysis (Curve), Akaike scores for each model fit (AIC), weighted Akaike scores for each model fit (AICw), beta (β), mu (μ) , and lambda (λ). Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S11. Model fitting results for the lineages through time values (LTTs) for each of the cetacean groups studied based on each phylogenetic reconstruction of our RFFD. The table includes the phylogenetic tree being assessed (Tree), the cetacean group under study (Group), Akaike scores for each model fit (AIC), weighted Akaike scores for each model fit (AICw), beta (β), mu (μ) , and lambda (λ). Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S12. Model fitting results for the lineages through time values (DTTs) for each of the cetacean groups studied based on each phylogenetic reconstruction of our RFFD. The table includes the phylogenetic tree being assessed (Tree), the cetacean group under study (Group), Akaike scores for each model fit (AIC), weighted Akaike scores for each model fit (AICw), beta (β), mu (μ), and lambda (λ). Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S13. Model fitting results for the lineages through time values (LTTs) for the Chondrichthyes phylogeny studied based on each phylogenetic reconstruction of our RFFD. The table includes the phylogenetic tree being assessed (Tree), the group under study (Group), Akaike scores for each model fit (AIC), weighted Akaike scores for each model fit (AICw), beta (β), mu (μ) , and lambda (λ). Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S14. Model fitting results for the diversity through time values (DTTs) for the Chondrichthyes phylogeny studied based on each phylogenetic reconstruction of our RFFD. The table includes the phylogenetic tree being assessed (Tree), the group under study (Group), Akaike scores for each model fit (AIC), weighted Akaike scores for each model fit (AICw), beta (β), mu (μ) , and lambda (λ). Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S15. Model fitting results for the lineages through time values (LTTs) for the Carcharhiniformes phylogeny studied based on each phylogenetic reconstruction of our RFFD. The table includes the phylogenetic tree being assessed (Tree), the group under study (Group), Akaike scores for each model fit (AIC), weighted Akaike scores for each model fit (AICw), beta (β), mu (μ) , and lambda (λ). Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S16. Model fitting results for the net diversification through time values (DTTs) for the Carcharhiniformes phylogeny studied based on each phylogenetic reconstruction of our RFFD. The table includes the phylogenetic tree being assessed (Tree), the group under study (Group), Akaike scores for each model fit (AIC), weighted Akaike scores for each model fit (AICw), beta (β), mu (μ), and lambda (λ). Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S17. List of specimens the compose the Marine Vertebrates Dataset (MVD). This table includes the family, genus, and species of each studied individual alongside their respective museum catalog number or accesion number, the institution affiliated with the sample, and the code of that institution. Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S18. List of all of the genes across our four scenarios the exhibited signs of positive selection. An 'X' under the specified scenario indicates that the gene experienced positive selection in that particular scenario. Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S19. The results from our BUSTED-PH analysis for scenario 1, where all endotherms were included in the foreground and all ectotherms in the background. The columns in the table represent various parameters. The 'ID' column represents the ID of the alignment used. The 'Panther Code' is the code from the Panther Database associated with the alignment. 'Name' is the name of the gene. 'Sequences' is the number of sequences present in the alignment. 'Num FG' is the number of species in the foreground. 'LRT' indicates the likelihood ratio test values obtained for that specific gene. 'BUSTED-E Filtered' indicates if the alignment required filtering through BUSTED-E. 'ω1', 'ω2', and 'ω3' represent the three omega values estimated by the analysis. 'FG P-Value' is the p-value associated with the foreground group. 'BG P-Value' is the p-value associated with the background group. 'DF P-Value' is the difference between the p-values of the foreground and the background groups. 'FDR Class' inform the type of positive selection observed after the FDR correction with class values at 101 representing selection in the fg, no selection in the bg and significant differences between the degree of selection observed, and values at 111 representing selection in the fg, selection in the bg, but significant differences between the degree of selection observed, indicating a degree of positive selection in the fg. Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S20. The results from our relative evolutionary rates (RER) analysis for scenario 1, where all endotherms were included in the foreground and all ectotherms in the background. The columns in the table represent various parameters. The 'ID' column represents the ID of the alignment used. The 'Panther Code' is the code from the Panther Database associated with the alignment. 'Name' is the name of the gene. 'FG Rel. Rate' represents the relative evolutionary rates of our species in the foreground, whereas 'BG Rel. Rate' represents the relative evolutionary rates for our background group. 'P-Value FG=BG' and 'Q-Value FG=BG' represent

the respective p and q-values of our foreground background comparison. Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S21. The results from our BUSTED-PH analysis for scenario 2, where only regional endotherms were included in the foreground, the rest of the endotherms were placed as nuisance species, and all ectotherms remained in the background. The columns in the table represent various parameters. The 'ID' column represents the ID of the alignment used. The 'Panther Code' is the code from the Panther Database associated with the alignment. 'Name' is the name of the gene. 'Sequences' is the number of sequences present in the alignment. 'Num FG' is the number of species in the foreground. 'LRT' indicates the likelihood ratio test values obtained for that specific gene. 'BUSTED-E Filtered' indicates if the alignment required filtering through BUSTED-E. 'ω1', 'ω2', and 'ω3' represent the three omega values estimated by the analysis. 'FG P-Value' is the p-value associated with the foreground group. 'BG P-Value' is the p-value associated with the background group. 'DF P-Value' is the difference between the p-values of the foreground and the background groups. 'FDR Class' inform the type of positive selection observed after the FDR correction with class values at 101 representing selection in the fg, no selection in the bg and significant differences between the degree of selection observed, and values at 111 representing selection in the fg, selection in the bg, but significant differences between the degree of selection observed, indicating a degree of positive selection in the fg. Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S22. The results from our relative evolutionary rates (RER) analysis for scenario 2, where only regional endotherms were included in the foreground, the rest of the endotherms were placed as nuisance species, and all ectotherms remained in the background. The columns in the table represent various parameters. The 'ID' column represents the ID of the alignment used. The 'Panther Code' is the code from the Panther Database associated with the alignment. 'Name' is the name of the gene. 'FG Rel. Rate' represents the relative evolutionary rates of our species in the foreground, whereas 'BG Rel. Rate' represents the relative evolutionary rates for our background group. 'P-Value FG=BG' and 'Q-Value FG=BG' represent the respective p and q-values of our foreground background comparison. Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S23. The results from our BUSTED-PH analysis for scenario 3, where only eye-brain endotherms were included in the foreground, the rest of the endotherms were placed as nuisance species, and all ectotherms remained in the background. The columns in the table represent various parameters. The 'ID' column represents the ID of the alignment used. The 'Panther Code' is the code from the Panther Database associated with the alignment. 'Name' is the name of the gene. 'Sequences' is the number of sequences present in the alignment. 'Num FG' is the number of species in the foreground. 'LRT' indicates the likelihood ratio test values obtained for that specific gene. 'BUSTED-E Filtered' indicates if the alignment required filtering through BUSTED-E. 'ω1', 'ω2', and 'ω3' represent the three omega values estimated by the analysis. 'FG P-Value' is the p-value associated with the foreground group. 'BG P-Value' is the p-value associated with the background group. 'DF P-Value' is the difference between the p-values of the foreground and the background groups. 'FDR Class' inform the type of positive selection observed after the FDR correction with class values at 101 representing selection in the fg, no selection in the bg and significant differences between the degree of selection observed, and values at 111 representing selection in the fg, selection in the bg, but significant differences between the degree

of selection observed, indicating a degree of positive selection in the fg. Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S24. The results from our relative evolutionary rates (RER) analysis for scenario 3, where only eye-brain endotherms were included in the foreground, the rest of the endotherms were placed as nuisance species, and all ectotherms remained in the background. The columns in the table represent various parameters. The 'ID' column represents the ID of the alignment used. The 'Panther Code' is the code from the Panther Database associated with the alignment. 'Name' is the name of the gene. 'FG Rel. Rate' represents the relative evolutionary rates of our species in the foreground, whereas 'BG Rel. Rate' represents the relative evolutionary rates for our background group. 'P-Value FG=BG' and 'Q-Value FG=BG' represent the respective p and q-values of our foreground background comparison. Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S25. The results from our BUSTED-PH analysis for scenario 4, where only full-bodied endotherms were included in the foreground, the rest of the endotherms were placed as nuisance species, and all ectotherms remained in the background. The columns in the table represent various parameters. The 'ID' column represents the ID of the alignment used. The 'Panther Code' is the code from the Panther Database associated with the alignment. 'Name' is the name of the gene. 'Sequences' is the number of sequences present in the alignment. 'Num FG' is the number of species in the foreground. 'LRT' indicates the likelihood ratio test values obtained for that specific gene. 'BUSTED-E Filtered' indicates if the alignment required filtering through BUSTED-E. 'ω1', 'ω2', and 'ω3' represent the three omega values estimated by the analysis. 'FG P-Value' is the p-value associated with the foreground group. 'BG P-Value' is the p-value associated with the background group. 'DF P-Value' is the difference between the p-values of the foreground and the background groups. 'FDR Class' inform the type of positive selection observed after the FDR correction with class values at 101 representing selection in the fg, no selection in the bg and significant differences between the degree of selection observed, and values at 111 representing selection in the fg, selection in the bg, but significant differences between the degree of selection observed, indicating a degree of positive selection in the fg. Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S26. The results from our relative evolutionary rates (RER) analysis for scenario 4, where only full-bodied endotherms were included in the foreground, the rest of the endotherms were placed as nuisance species, and all ectotherms remained in the background. The columns in the table represent various parameters. The 'ID' column represents the ID of the alignment used. The 'Panther Code' is the code from the Panther Database associated with the alignment. 'Name' is the name of the gene. 'FG Rel. Rate' represents the relative evolutionary rates of our species in the foreground, whereas 'BG Rel. Rate' represents the relative evolutionary rates for our background group. 'P-Value FG=BG' and 'Q-Value FG=BG' represent the respective p and q-values of our foreground background comparison. Access data at https://datadryad.org/downloads/file_stream/3956152.

Table S27. Data on the ecological interactions of various fish and mammal groups during the Late Miocene and Serravallian intervals, extracted from the Paleobiology Database. The 'Group' column identifies the taxonomic group involved in the interaction, while 'Interaction' specifies the type of ecological relationship observed. 'Early Interval' denotes the geological time frame, with 'Max Age' and 'Min. Age' indicating its temporal boundaries in million years ago

(Mya). The locations of these interactions are represented by coordinates in 'Paleo-longitude' and 'Paleo-latitude'. In this dataset, Thunnini is observed as a predator/prey during both intervals at distinct global positions. This information provides valuable insights into the ecological dynamics of Thunnini during these geological periods. Access data at https://datadryad.org/downloads/file_stream/3956152.

Supplementary Data (Separate File)

Dataset S1. All gene alignments for the RFFD. This dataset contains all gene alignments for the Ray-Finned Fish Dataset (RFFD), providing the foundational data for phylogenetic and evolutionary analyses. Access data at https://datadryad.org/downloads/file_stream/3956147.

Dataset S2. File containing all phylograms and chronograms used in this study. A comprehensive file with all phylograms and chronograms used in this study, supporting divergence time estimation and evolutionary modeling. Access data at https://datadryad.org/downloads/file_stream/3956147.

Dataset S3. RFFD calibration file using 31 fossils. The RFFD calibration file incorporating 31 fossil constraints, essential for time-calibrated phylogenetic reconstructions. Access data at https://datadryad.org/downloads/file_stream/3956147.

Dataset S4. RFFD calibration file using 33 fossils. The RFFD calibration file using 33 fossil constraints, refining divergence estimates within the dataset. Access data at https://datadryad.org/downloads/file_stream/3956147.

Dataset S5. RFFD calibration file using 34 fossils. The RFFD calibration file applying 34 fossil constraints, further improving calibration accuracy in evolutionary analyses. Access data at https://datadryad.org/downloads/file_stream/3956147.

Dataset S6. RFFD calibration file using averaged calibrations from previous studies. A calibration file for the RFFD using averaged calibrations from previous studies, integrating multiple sources for robust phylogenetic dating. Access data at https://datadryad.org/downloads/file_stream/3956147.

Dataset S7. Scripts and results for the HiSSE analyses. A collection of scripts and results from the HiSSE analyses, detailing the hidden state-dependent diversification models used in this study. Access data at https://datadryad.org/downloads/file_stream/3956147.

Dataset S8. ***Script for the modified threshold model, including the climatic, LTT, and DTT analyses.*** A script for the modified threshold model, including climatic, Lineage Through Time (LTT), and Disparity Through Time (DTT) analyses. Access data at https://datadryad.org/downloads/file_stream/3956147.

Dataset S9. List of gene IDs for the 894 single-copy orthologs for the MVD. A list of gene IDs for the 894 single-copy orthologs identified in the Marine Vertebrate Dataset (MVD), supporting comparative genomic analyses. Access data at https://datadryad.org/downloads/file_stream/3956147.

Dataset S10. Single-copy orthologs obtained from OrthoFinder and exon marker alignments for MVD. Single-copy orthologs derived from OrthoFinder and exon marker alignments for the MVD, providing key genomic data for evolutionary inference. Access data at https://datadryad.org/downloads/file_stream/3956147.

Dataset S11. Technique used to match amino acid to nucleotide sequences. A custom Python script designed to match amino acid sequences to their corresponding nucleotide sequences, facilitating molecular evolutionary analyses. Access data at https://datadryad.org/downloads/file_stream/3956147.

REFERENCES AND NOTES

1. B. A. Block, J. R. Finnerty, Endothermy in fishes: A phylogenetic analysis of constraints, predispositions, and selection pressures. *Environ. Biol. Fishes* **40**, 283–302 (1994).

2. W. J. Hillenius, J. A. Ruben, The evolution of endothermy in terrestrial vertebrates: Who? When? Why? *Physiol. Biochem. Zool.* **77**, 1019–1042 (2004).

3. L. J. Legendre, D. Davesne, The evolution of mechanisms involved in vertebrate endothermy. *Philos. Trans. R. Soc. London Ser. B Biol. Sci.* **375**, 20190136 (2020).

4. K. A. Dickson, J. B. Graham, Evolution and consequences of endothermy in fishes. *Physiol. Biochem. Zool.* **77**, 998–1018 (2004).

5. D. J. Madigan, A. B. Carlisle, L. D. Gardner, N. Jayasundara, F. Micheli, K. M. Schaefer, D. W. Fuller, B. A. Block, Assessing niche width of endothermic fish from genes to ecosystem. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 8350–8355 (2015).

6. J. Rolland, D. Silvestro, D. Schluter, A. Guisan, O. Broennimann, N. Salamin, The impact of endothermy on the climatic niche evolution and the distribution of vertebrate diversity. *Nat. Ecol. Evol.* **2**, 459–464 (2018).

7. L. Harding, A. Jackson, A. Barnett, I. Donohue, L. Halsey, C. Huveneers, C. Meyer, Y. Papastamatiou, J. M. Semmens, E. Spencer, Y. Watanabe, N. Payne, Endothermy makes fishes faster but does not expand their thermal niche. *Funct. Ecol.* **35**, 1951–1959 (2021).

8. D. Bernal, K. A. Dickson, R. E. Shadwick, J. B. Graham, Review: Analysis of the evolutionary convergence for high-performance swimming in lamnid sharks and tunas. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **129**, 695–726 (2001).

9. B. A. Block, J. R. Finnerty, A. F. Stewart, J. Kidd, Evolution of endothermy in fish: Mapping physiological traits on a molecular phylogeny. *Science* **260**, 210–214 (1993).

10. F. G. Carey, K. D. Lawson, Temperature regulation in free-swimming bluefin tuna. *Comp. Biochem. Physiol. A Comp. Physiol.* **44**, 375–392 (1973).

11. K. A. Dickson, N. M. Johnson, J. M. Donley, J. A. Hoskinson, M. W. Hansen, J. D'souza Tessier, Ontogenetic changes in characteristics required for endothermy in juvenile black skipjack tuna (*Euthynnus lineatus*). *J. Exp. Biol.* **203**, 3077–3087 (2000).

12. J. B. Thiebot, J. C. McInnes, Why do marine endotherms eat gelatinous prey? *ICES J. Mar. Sci.* **77**, 58–71 (2020).

13. H. G. Ferrón, Regional endothermy as a trigger for gigantism in some extinct macropredatory sharks. *PLOS ONE* **12**, e0185185 (2017).

14. C. Pimiento, J. L. Cantalapiedra, K. Shimada, D. J. Field, J. B. Smaers, Evolutionary pathways toward gigantism in sharks and rays. *Evolution* **73**, 588–599 (2019).

15. R. W. Brill, Selective advantages conferred by the high-performance physiology of tunas, billfishes, and dolphin fish. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **113**, 3–15 (1996).

16. J. B. Graham, K. A. Dickson, The evolution of thunniform locomotion and heat conservation in scombrid fishes: New insights based on the morphology of *Allothunnus fallai*. *Zool. J. Linn. Soc.* **129**, 419–466 (2000).

17. Y. Y. Watanabe, K. J. Goldman, J. E. Caselle, D. D. Chapman, Y. P. Papastamatiou, Comparative analyses of animal-tracking data reveal ecological significance of endothermy in fishes. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6104–6109 (2015).

18. V. Domínguez-García, S. Kéfi, The structure and robustness of ecological networks with two interaction types. *PLoS Comput. Biol.* **20**, e1011770 (2024).

19. M. J. Benton, The origin of endothermy in synapsids and archosaurs and arms races in the Triassic. *Gondw. Res.* **100**, 261–289 (2021).

20. K. A. Monsch, A revision of scombrid fishes (*Scombroidei, Perciformes*) from the Middle Eocene of Monte Bolca, Italy. *Palaeontology* **49**, 873–888 (2006).

21. F. Santini, G. Carnevale, L. Sorenson, First molecular scombrid timetree (*Percomorpha: Scombridae*) shows recent radiation of tunas following invasion of pelagic habitat. *Ital. J. Zool.* **80**, 210–221 (2013).

22. M. B. Santos, G. J. Pierce, Cetacean diet. *Encycl. Anim. Cogn. Behav.* **1**, 1–9 (2017).

23. M. S. Savoca, M. F. Czapanskiy, S. R. Kahane-Rapport, W. T. Gough, J. A. Fahlbusch, K. C. Bierlich, P. S. Segre, J. Di Clemente, G. S. Penry, D. N. Wiley, J. Calambokidis, D. P. Nowacek, D. W. Johnston, N. D. Pyenson, A. S. Friedlaender, E. L. Hazen, J. A. Goldbogen, Baleen whale prey consumption based on high-resolution foraging measurements. *Nature* **599**, 85–90 (2021).

24. R. Motani, K. Shimada, Skeletal convergence in thunniform sharks, ichthyosaurs, whales, and tunas, and its possible ecological links through marine ecosystem evolution. *Sci. Rep.* **13**, 16664 (2023).

25. B. Wu, C. Feng, C. Zhu, W. Xu, Y. Yuan, M. Hu, K. Yuan, Y. Li, Y. Ren, Y. Zhou, H. Jiang, Q. Qiu, W. Wang, S. He, K. Wang, The genomes of two billfishes provide insights into the evolution of endothermy in teleosts. *Mol. Biol. Evol.* **19**, 2413–2427 (2021).

26. A. G. Ciezarek, O. G. Osborne, O. N. Shipley, E. J. Brooks, S. R. Tracey, J. D. McAllister, L. D. Gardner, M. J. E. Sternberg, B. Block, V. Savolainen, Phylotranscriptomic insights into the diversification of endothermic *Thunnus* tunas. *Mol. Biol. Evol.* **36**, 84–96 (2019).

27. X. Wang, M. Qu, Y. Liu, R. F. Schneider, Y. Song, Z. Chen, H. Zhang, Y. Zhang, H. Yu, S. Zhang, D. Li, G. Qin, S. Ma, J. Zhong, J. Yin, S. Liu, G. Fan, A. Meyer, D. Wang, Q. Lin, Genomic basis of evolutionary adaptation in a warm-blooded fish. *Innovation* **3**, 100185 (2022).

28. J. Bo, W. Q. Lv, N. Sun, C. Wang, K. Wang, P. Liu, C. G. Feng, S. P. He, L. D. Yang, Opah (*Lampris megalopsis*) genome sheds light on the evolution of aquatic endothermy. *Zool. Res.* **43**, 26–29 (2022).

29. B. T. L. Nichio, J. N. Marchaukoski, R. T. Raittz, New tools in orthology analysis: A brief review of promising perspectives. *Front. Genet.* **8**, 165 (2017).

30. L. C. Hughes, G. Ortí, Y. Huang, Y. Sun, C. C. Baldwin, A. W. Thompson, D. Arcila, R. Betancur-R, C. Li, L. Becker, N. Bellora, X. Zhao, X. Li, M. Wang, C. Fang, B. Xie, Z. Zhou, H. Huang, S. Chen, B. Venkatesh, Q. Shi, Comprehensive phylogeny of ray-finned fishes (*Actinopterygii*) based on transcriptomic and genomic data. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 6249–6254 (2018).

31. J. M. Eastman, L. J. Harmon, D. C. Tank, Congruification: Support for time scaling large phylogenetic trees. *Methods Ecol. Evol.* **4**, 688–691 (2013).

32. R. Betancur-R, E. O. Wiley, G. Arratia, A. Acero, N. Bailly, M. Miya, G. Lecointre, G. Ortí, Phylogenetic classification of bony fishes. *BMC Evol. Biol.* **17**, 162 (2017).

33. T. J. Near, R. I. Eytan, A. Dornburg, K. L. Kuhn, J. A. Moore, M. P. Davis, P. C. Wainwright, M. Friedman, W. L. Smith, Resolution of ray-finned fish phylogeny and timing of diversification. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 13698–13703 (2012).

34. D. L. Rabosky, J. Chang, P. O. Title, P. F. Cowman, L. Sallan, M. Friedman, K. Kaschner, C. Garilao, T. J. Near, M. Coll, M. E. Alfaro, An inverse latitudinal gradient in speciation rate for marine fishes. *Nature* **559**, 392–395 (2018).

35. M. E. Alfaro, B. C. Faircloth, R. C. Harrington, L. Sorenson, M. Friedman, C. E. Thacker, C. H. Oliveros, D. Černý, T. J. Near, Explosive diversification of marine fishes at the Cretaceous–Palaeogene boundary. *Nat. Ecol. Evol.* **2**, 688–696 (2018).

36. E. Ribeiro, A. M. Davis, R. A. Rivero-Vega, G. Ortí, R. Betancur-R, Post-Cretaceous bursts of evolution along the benthic-pelagic axis in marine fishes. *Proc. Biol. Sci.* **285**, 20182010 (2018).

37. A. Santaquiteria, A. C. Siqueira, E. Duarte-Ribeiro, G. Carnevale, W. T. White, J. J. Pogonoski, C. C. Baldwin, G. Ortí, D. Arcila, B. R. Ricardo, Phylogenomics and historical biogeography of seahorses, dragonets, goatfishes, and allies (*Teleostei: Syngnatharia*):

Assessing factors driving uncertainty in biogeographic inferences. *Syst. Biol.* **70**, 1145–1162 (2021).

38. M. Friedman, K. L. Feilich, H. T. Beckett, M. E. Alfaro, B. C. Faircloth, D. Černý, M. Miya, T. J. Near, R. C. Harrington, A phylogenomic framework for pelagiarian fishes (*Acanthomorpha: Percomorpha*) highlights mosaic radiation in the open ocean. *Proc. R. Soc. London Ser. B Biol. Sci.* **286**, 20191502 (2019).

39. R. C. Harrington, B. C. Faircloth, R. I. Eytan, W. L. Smith, T. J. Near, M. E. Alfaro, M. Friedman, Phylogenomic analysis of carangimorph fishes reveals flatfish asymmetry arose in a blink of the evolutionary eye. *BMC Evol. Biol.* **16**, 224 (2016).

40. A. Ghezelayagh, R. C. Harrington, E. D. Burress, M. A. Campbell, J. C. Buckner, P. Chakrabarty, J. R. Glass, W. T. McCraney, P. J. Unmack, C. E. Thacker, M. E. Alfaro, S. T. Friedman, W. B. Ludt, P. F. Cowman, M. Friedman, S. A. Price, A. Dornburg, B. C. Faircloth, P. C. Wainwright, T. J. Near, Prolonged morphological expansion of spiny-rayed fishes following the end-Cretaceous. *Nat. Ecol. Evol.* **6**, 1211–1220 (2022).

41. M. Rincon-Sandoval, E. Duarte-Ribeiro, A. M. Davis, A. Santaquiteria, L. C. Hughes, C. C. Baldwin, L. Soto-Torres, A. P. Acero, H. J. Walker, K. E. Carpenter, M. Sheaves, G. Ortí, D. Arcila, R. R. Betancur-R, Evolutionary determinism and convergence associated with water-column transitions in marine fishes. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 33396–33403 (2020).

42. J. P. Bollback, SIMMAP: Stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics* **7**, 88 (2006).

43. J. M. Beaulieu, B. C. O'Meara, Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Syst. Biol.* **65**, 583–601 (2016).

44. P. C. Wainwright, S. J. Longo, Functional innovations and the conquest of the oceans by acanthomorph fishes. *Curr. Biol.* **27**, R550–R557 (2017).

45. J. Ng, S. D. Smith, Why are red flowers so rare? Testing the macroevolutionary causes of tippiness. *J. Evol. Biol.* **31**, 1863–1875 (2018).

46. S. A. Price, S. T. Friedman, K. A. Corn, O. Larouche, K. Brockelsby, A. J. Lee, M. Nagaraj, N. G. Bertrand, M. Danao, M. C. Coyne, J. R. Estrada, R. Friedman, E. Hoeft, M. Iwan, D. Gross, J. H. Kao, B. Landry, M. J. Linares, C. McGlinn, J. A. Nguyen, A. G. Proffitt, S. Rodriguez, M. R. Rupp, E. Y. Shen, V. Susman, A. J. Tovar, L. L. J. Vary, K. L. Zapfe, P. C. Wainwright, FishShapes v1: Functionally relevant measurements of teleost shape and size on three dimensions. *Ecology* **103**, e3829 (2022).

47. V. Di Santo, E. Goerig, D. K. Wainwright, O. Akanyeti, J. C. Liao, T. Castro-Santos, G. V. Lauder, Convergence of undulatory swimming kinematics across a diversity of fishes. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2113206118 (2021).

48. A. R. Ives, T. Garland, Phylogenetic logistic regression for binary dependent variables. *Syst. Biol.* **59**, 9–26 (2010).

49. B. A. Block, Endothermy in fish: Thermogenesis, ecology, and evolution. *Biochem. Mol. Biol. Fishes* **1**, 269–311 (1991).

50. J. Felsenstein, D. D. Ackerly, M. A. McPeek, A comparative method for both discrete and continuous characters using the threshold model. *Am. Nat.* **179**, 145–156 (2012).

51. G. Burin, T. Park, T. D. James, G. J. Slater, N. Cooper, The dynamic adaptive landscape of cetacean body size. *Curr. Biol.* **33**, 1787–1794.e3 (2023).

52. M. A. Silva, R. Feio, R. Prieto, J. M. Gonçalves, R. S. Santos, Interactions between cetaceans and the tuna fishery in the Azores. *Mar. Mamm. Sci.* **18**, 893–901 (2002).

53. R. W. Stein, C. G. Mull, T. S. Kuhn, N. C. Aschliman, L. N. K. Davidson, J. B. Joy, G. J. Smith, N. K. Dulvy, A. O. Mooers, Global priorities for conserving the evolutionary history of sharks, rays, and chimaeras. *Nat. Ecol. Evol.* **2**, 288–298 (2018).

54. B. Brée, F. L. Condamine, G. Guinot, Combining palaeontological and neontological data shows a delayed diversification burst of carcharhiniform sharks likely mediated by environmental change. *Sci. Rep.* **12**, 21906 (2022).

55. A. F. Bennett, J. A. Ruben, Endothermy and activity in vertebrates. *Science* **206**, 649–654 (1979).

56. L. J. Revell, L. J. Harmon, D. C. Collar, Phylogenetic signal, evolutionary process, and rate. *Syst. Biol.* **57**, 591–601 (2008).

57. A. Capobianco, M. Friedman, Vicariance and dispersal in southern hemisphere freshwater fish clades: A palaeontological perspective. *Biol. Rev. Camb. Philos. Soc.* **94**, 662–699 (2019).

58. S. L. Kosakovsky Pond, A. F. Y. Poon, R. Velazquez, S. Weaver, N. L. Hepler, B. Murrell, S. D. Shank, B. R. Magalis, D. Bouvier, A. Nekrutenko, S. Wisotsky, S. J. Spielman, S. D. W. Frost, S. V. Muse, HyPhy 2.5—A customizable platform for evolutionary hypothesis testing using phylogenies. *Mol. Biol. Evol.* **37**, 295–299 (2020).

59. S. R. Wisotsky, S. L. K. Pond, S. D. Shank, S. V. Muse, Synonymous site-to-site substitution rate variation dramatically inflates false positive rates of selection analyses: Ignore at your own peril. *Mol. Biol. Evol.* **37**, 2430–2439 (2020).

60. A. Selberg, N. L. Clark, T. B. Sackton, S. V. Muse, A. G. Lucaci, S. Weaver, A. Nekrutenko, M. Chikina, S. L. Kosakovsky Pond, Minus the error: Estimating dN/dS and testing for natural selection in the presence of residual alignment errors. bioRxiv 2024.11.13.620707 [Preprint] (2024). https://doi.org/10.1101/2024.11.13.620707.

61. P. D. Thomas, D. Ebert, A. Muruganujan, T. Mushayahama, L. P. Albou, H. Mi, PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci.* **31**, 8–22 (2022).

62. A. Kowalczyk, W. K. Meyer, R. Partha, W. Mao, N. L. Clark, M. Chikina, RERconverge: An R package for associating evolutionary rates with convergent traits. *Bioinformatics* **35**, 4815–4817 (2019).

63. A. Kowalczyk, M. Chikina, N. L. Clark, Complementary evolution of coding and noncoding sequence underlies mammalian hairlessness. *Elife* **11**, e76911 (2022).

64. W. H. Neill, R. K. Chang, A. E. Dizon, Magnitude and ecological implications of thermal inertia in skipjack tuna, *Katsuwonus pelamis* (Linnaeus). *Environ. Biol. Fishes* **1**, 61–80 (1976).

65. J. Parker, G. Tsagkogeorga, J. A. Cotton, Y. Liu, P. Provero, P. Stupka, S. J. Rossiter, Genome-wide signatures of convergent evolution in echolocating mammals. *Nature* **502**, 228–231 (2013).

66. S. Pan, Y. Lin, Q. Liu, J. Duan, Z. Lin, Y. Wang, X. Wang, S. M. Lam, Z. Zou, G. Shui, Y. Zhang, Z. Zhang, X. Zhan, Convergent genomic signatures of flight loss in birds suggest a switch of main fuel. *Nat. Commun.* **10**, 2756 (2019).

67. F. Seebacher, Is endothermy an evolutionary by-product? *Trends Ecol. Evol.* **35**, 503–511 (2020).

68. N. L. Payne, E. P. Snelling, R. Fitzpatrick, J. Seymour, R. Courtney, A. Barnett, Y. Y. Watanabe, D. W. Sims, L. Squire, J. M. Semmens, A new method for resolving uncertainty of energy requirements in large water breathers: The 'mega-flume' seagoing swim-tunnel respirometer. *Methods Ecol. Evol.* **6**, 668–677 (2015).

69. M. C. Arostegui, M. R. Shero, L. R. Frank, R. M. Berquist, C. D. Braun, An enigmatic pelagic fish with internalized red muscle: A future regional endotherm or forever an ectotherm? *J. Fish Biol.* **102**, 1311–1326 (2023).

70. L. C. Hughes, G. Ortí, H. Saad, C. Li, W. T. White, C. C. Baldwin, K. A. Crandall, D. Arcila, R. Betancur-R, Exon probe sets and bioinformatics pipelines for all levels of fish phylogenomics. *Mol. Ecol. Resour.* **21**, 816–833 (2021).

71. D. Arcila, L. C. Hughes, F. Meléndez-Vázquez, C. C. Baldwin, W. T. White, K. E. Carpenter, J. T. Williams, M. D. Santos, J. J. Pogonoski, M. Miya, G. Ortí, R. Betancur-R, Testing the utility of alternative metrics of branch support to address the ancient evolutionary radiation of tunas, stromateoids, and allies (*Teleostei: Pelagiaria*). *Syst. Biol.* **70**, 1123–1144 (2021).

72. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

73. G. Van Rossum, F. L. Drake, *Python 3 Reference Manual* (CreateSpace, 2009).

74. H. Li, R. Durbin, Fast and accurate short-read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

75. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).

76. D. R. Zerbino, E. Birney, Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **18**, 821–829 (2008).

77. J. M. Allen, R. LaFrance, R. A. Folk, K. P. Johnson, R. P. Guralnick, aTRAM 2.0: An improved, flexible locus assembler for NGS data. *Evol. Bioinform. Online* **14**, 1176934318774546 (2018).

78. M. G. Grabherr, B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. Di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, A. Regev, Trinity: Reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* **29**, 644–652 (2011).

79. L. Fu, B. Niu, Z. Zhu, S. Wu, W. Li, CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).

80. G. S. C. Slater, E. Birney, Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* **6**, 31 (2005).

81. V. Ranwez, E. J. P. Douzery, C. Cambon, N. Chantret, F. Delsuc, MACSE v2: Toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. *Mol. Biol. Evol.* **35**, 2582–2584 (2018).

82. T. J. Wheeler, S. R. Eddy, Nhmmer: DNA homology search with profile HMMs. *Bioinformatics* **29**, 2487–2489 (2013).

83. G. L. J. Galli, M. S. Lipnick, H. A. Shiels, B. A. Block, Temperature effects on Ca^{2+} cycling in scombrid cardiomyocytes: A phylogenetic comparison. *J. Exp. Biol.* **214**, 1068–1076 (2011).

84. J. C. Patterson, C. A. Sepulveda, D. Bernal, The vascular morphology and in vivo muscle temperatures of thresher sharks (*Alopiidae*). *J. Morphol.* **272**, 1353–1364 (2011).

85. D. M. Emms, S. Kelly, OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019).

86. M. L. Borowiec, AMAS: A fast tool for alignment manipulation and computing of summary statistics. *PeerJ* **28**, e1600 (2016).

87. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLOS ONE* **5**, e9490 (2010).

88. B. Q. Minh, H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. Von Haeseler, R. Lanfear, E. Teeling, IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).

89. P. Kapli, Z. Yang, M. J. Telford, Phylogenetic tree building in the genomic age. *Nat. Rev. Genet.* **21**, 428–444 (2020).

90. C. Zhang, M. Rabiee, E. Sayyari, S. Mirarab, ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**, 153 (2018).

91. M. Nute, J. Chou, E. K. Molloy, T. Warnow, The performance of coalescent-based species tree estimation methods under models of missing data. *BMC Genomics* **19**, 286 (2018).

92. X. Jiao, T. Flouri, Z. Yang, Multispecies coalescent and its applications to infer species phylogenies and cross-species gene flow. *Natl. Sci. Rev.* **8**, nwab127 (2021).

93. T. Britton, C. L. Anderson, D. Jacquet, S. Lundqvist, K. Bremer, Estimating divergence times in large phylogenetic trees. *Syst. Biol.* **56**, 741–752 (2007).

94. R. A. Pyron, J. J. Wiens, A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. *Mol. Phylogenet. Evol.* **61**, 543–583 (2011).

95. K. Tamura, F. U. Battistuzzi, P. Billing-Ross, O. Murillo, A. Filipski, S. Kumar, Estimating divergence times in large molecular phylogenies. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 19333–19338 (2012).

96. S. A. Smith, J. M. Beaulieu, A. Stamatakis, M. J. Donoghue, Understanding angiosperm diversification using small and large phylogenetic trees. *Am. J. Bot.* **98**, 404–414 (2011).

97. C. E. Hinchliff, E. H. Roalson, Using supermatrices for phylogenetic inquiry: An example using the sedges. *Syst. Biol.* **62**, 205–219 (2013).

98. M. W. Pennell, J. M. Eastman, G. J. Slater, J. W. Brown, J. C. Uyeda, R. G. Fitzjohn, M. E. Alfaro, L. J. Harmon, geiger v2.0: An expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics* **30**, 2216–2218 (2014).

99. S. A. Smith, B. C. O'Meara, treePL: Divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* **28**, 2689–2690 (2012).

100. Y. Y. Watanabe, I. Nakamura, W. C. Chiang, Behavioural thermoregulation linked to foraging in blue sharks. *Mar. Biol.* **168**, 161 (2021).

101. A. Stoehr, J. St. Martin, S. Aalbers, C. Sepulveda, D. Bernal, Free-swimming swordfish, *Xiphias gladius*, alter the rate of whole-body heat transfer: Morphological and physiological specializations for thermoregulation. *ICES J. Mar. Sci.* **75**, 858–870 (2018).

102. R. Froese, D. Pauly, Fishbase (2023); https://fishbase.se/search.php.

103. D. R. Robertson, J. Van Tassell, Shorefishes of the Greater Caribbean: Online Information System (2023); https://biogeodb.stri.si.edu/caribbean/en/pages.

104. J. S. Nelson, T. C. Grande, M. V. H. Wilson, *Fish of the World* (Wiley, 2016), pp. 1–752.

105. Museums Victoria, Fishes of Australia, https://fishesofaustralia.net.au/ [accessed 2023].

106. FAO. FISHSTAT Database, https://fao.org/fishery/en/fishstat [accessed 2023].

107. J. Pickering, Discover Life Database, https://www.discoverlife.org/ [accessed 2023].

108. Reeflex, Fishes, https://reeflex.net/ [accessed 2023].

109. L. S. Tung Ho, C. Ané, A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Syst. Biol.* **63**, 397–408 (2014).

110. G. J. Li, L. Zhu, X. Y. Lu, Numerical studies on locomotion performance of fish-like tail fins. *J. Hydrodynam.* **24**, 488–495 (2012).

111. D. H. Kühlmann, B. B. Collette, C. E. Nauen, *FAO Species Catalogue. Vol. 2, Scombrids of the World*. *FAO Fisheries Synopsis No. 125* (FAO, 1983), 137 pp.

112. R. W. Blake, L. M. Chatters, P. Domenici, Turning radius of yellowfin tuna (*Thunnus albacares*) in unsteady swimming manoeuvres. *J. Fish Biol.* **46**, 536–538 (1995).

113. A. Y. Hsiang, D. J. Field, T. H. Webster, A. D. B. Behlke, M. B. Davis, R. A. Racicot, J. A. Gauthier, The origin of snakes: Revealing the ecology, behavior, and evolutionary history of early snakes using genomics, phenomics, and the fossil record. *BMC Evol. Biol.* **15**, 87 (2015).

114. L. J. Revell, *phytools*: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).

115. K. Magnuson-Ford, S. P. Otto, Linking the investigations of character evolution and species diversification. *Am. Nat.* **180**, 225–245 (2012).

116. R. G. Fitzjohn, Quantitative traits and diversification. *Syst. Biol.* **59**, 619–633 (2010).

117. J. Ng, S. D. Smith, How traits shape trees: New approaches for detecting character state-dependent lineage diversification. *J. Evol. Biol.* **27**, 2035–2045 (2014).

118. W. P. Maddison, P. E. Midford, S. P. Otto, Estimating a binary character's effect on speciation and extinction. *Syst. Biol.* **56**, 701–710 (2007).

119. M. L. Collyer, D. C. Adams, *RRPP*: An R package for fitting linear models to high-dimensional data using residual randomization. *Methods Ecol. Evol.* **9**, 1772–1779 (2018).

120. D. C. Adams, M. L. Collyer, Phylogenetic ANOVA: Group-clade aggregation, biological challenges, and a refined permutation procedure. *Evolution* **72**, 1204–1215 (2018).

121. G. Hiscott, C. Fox, M. Parry, D. Bryant, Efficient recycled algorithms for quantitative trait models on phylogenies. *Genome Biol. Evol.* **8**, 1338–1350 (2016).

122. H. Wickham, ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* **3**, 180–185 (2011).

123. E. M. Troyer, R. Betancur-R, L. C. Hughes, M. Westneat, G. Carnevale, W. T. White, J. J. Pogonoski, J. C. Tyler, C. C. Baldwin, G. Ortí, A. Brinkworth, J. Clavel, D. Arcila, The impact of paleoclimatic changes on body size evolution in marine fishes. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2122486119 (2022).

124. M. J. Novacek, E. E. Cleland, The current biodiversity extinction event: Scenarios for mitigation and recovery. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5466–5470 (2001).

125. C. H. Basson, O. Levy, M. J. Angilletta, S. Clusella-Trullas, Lizards paid a greater opportunity cost to thermoregulate in a less heterogeneous environment. *Funct. Ecol.* **31**, 856–865 (2017).

126. L. B. Buckley, A. H. Hurlbert, W. Jetz, Broad-scale ecological implications of ectothermy and endothermy in changing environments. *Glob. Ecol. Biogeogr.* **21**, 873–885 (2012).

127. R. E. Buskirk, Zoogeographic patterns and tectonic history of Jamaica and the Northern Caribbean. *J. Biogeogr.* **12**, 445–461 (1985).

128. F. L. Condamine, J. Romieu, G. Guinot, Climate cooling and clade competition likely drove the decline of lamniform sharks. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 20584–20590 (2019).

129. A. Salces-Castellano, J. Patiño, N. Alvarez, C. Andújar, P. Arribas, J. J. Braojos-Ruiz, M. Arco-Aguilar, V. García-Olivares, D. N. Karger, H. López, I. Manolopoulou, P. Oromí, A. J. Pérez-Delgado, W. E. Peterman, K. F. Rijsdijk, B. C. Emerson, Climate drives community-wide divergence within species over a limited spatial scale: Evidence from an oceanic island. *Ecol. Lett.* **23**, 305–315 (2020).

130. N. C. Stenseth, A. Mysterud, G. Ottersen, J. W. Hurrell, K. S. Chan, M. Lima, Ecological effects of climate fluctuations. *Science* **297**, 1292–1296 (2002).

131. J. Clavel, H. Morlon, Accelerated body size evolution during cold climatic periods in the Cenozoic. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 4183–4188 (2017).

132. C. R. Scotese, H. Song, B. J. W. Mills, D. G. van der Meer, Phanerozoic paleotemperatures: The Earth's changing climate during the last 540 million years. *Earth Sci. Rev.* **215**, 103503 (2021).

133. H. Morlon, E. Lewitus, F. L. Condamine, M. Manceau, J. Clavel, J. Drury, RPANDA: An R package for macroevolutionary analyses on phylogenetic trees. *Methods Ecol. Evol.* **7**, 589–597 (2016).

134. A. J. Helmstetter, S. Glemin, J. Käfer, R. Zenil-Ferguson, H. D. S. Sauquet, H. De Boer, L. P. M. J. Dagallier, N. Mazet, E. L. Reboud, T. L. P. Couvreur, F. L. Condamine, Pulled diversification rates, lineages-through-time plots, and modern macroevolutionary modeling. *Syst. Biol.* **71**, 758–773 (2022).

135. C. Simpson, W. Kiessling, "Diversity of life through time" in *Encyclopedia of Life Sciences* (Wiley, 2010).

136. R. G. Fitzjohn, Diversitree: Comparative phylogenetic analyses of diversification in R. *Methods Ecol. Evol.* **3**, 1084–1092 (2012).

137. A. G. Lucaci, J. D. Zehr, D. Enard, J. W. Thornton, S. L. Kosakovsky Pond, Evolutionary shortcuts via multinucleotide substitutions and their impact on natural selection analyses. *Mol. Biol. Evol.* **40**, msad150 (2023).

138. O. Fafioye, O. P. Ayodele, Length-weight relationship and condition factor of four commercial fish species of Oyan Lake, Nigeria. *Exam. Mar. Biol. Oceanogr.* **2**, 10.31031/EIMBO.2018.02.000543 (2018).

139. J. B. Swann, S. J. Holland, M. Petersen, T. W. Pietsch, T. Boehm, The immunogenetics of sexual parasitism. *Science* **369**, 1608–1615 (2020).

140. S. L. Kosakovsky Pond, B. Murrell, M. Fourment, S. D. W. Frost, W. Delport, K. Scheffler, A random effects branch-site model for detecting episodic diversifying selection. *Mol. Biol. Evol.* **28**, 3033–3043 (2011).

141. N. J. Marra, V. P. Richards, A. Early, S. M. Bogdanowicz, P. D. Pavinski Bitar, M. J. Stanhope, M. S. Shivji, Comparative transcriptomics of elasmobranchs and teleosts highlight important processes in adaptive immunity and regional endothermy. *BMC Genomics* **18**, 87 (2017).

142. M. S. Hedrick, S. S. Hillman, What drove the evolution of endothermy? *J. Exp. Biol.* **219**, 300–301 (2016).

143. G. V. Lauder, E. G. Drucker, Forces, fishes, and fluids: Hydrodynamic mechanisms of aquatic locomotion. *News Physiol. Sci.* **17**, 235–240 (2002).