

# *Historical warming consistently decreased size, dispersal and speciation rate of fish*

Article

Accepted Version

Avaria-Llautureo, J. ORCID: <https://orcid.org/0000-0002-8610-7428>, Venditti, C., Rivadeneira, M. M. ORCID: <https://orcid.org/0000-0002-1681-416X>, Inostroza-Michael, O., Rivera, R. J. ORCID: <https://orcid.org/0000-0001-7903-0314>, Hernández, C. E. ORCID: <https://orcid.org/0000-0002-9811-2881> and Canales-Aguirre, C. B. ORCID: <https://orcid.org/0000-0002-8468-6139> (2021) Historical warming consistently decreased size, dispersal and speciation rate of fish. *Nature Climate Change*. ISSN 1758-678X doi: <https://doi.org/10.1038/s41558-021-01123-5> Available at <https://centaur.reading.ac.uk/99750/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <http://dx.doi.org/10.1038/s41558-021-01123-5>

To link to this article DOI: <http://dx.doi.org/10.1038/s41558-021-01123-5>

Publisher: Nature Publishing Group

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

1 *Title: Warmer temperatures decrease size, dispersal ability and speciation rate in Clupeiform fish*  
2

3 **There is an ongoing debate as to whether fish body size will decrease with global warming**  
4 **and how changes in body may impact dispersal ability and speciation rate. Theory predicts**  
5 **that, when fish face warmer temperatures, they grow to smaller adult sizes, undergo a**  
6 **reduction in their ability to move, and increase their probability of speciation. However,**  
7 **evaluations of such predictions are hampered owing to the lack of empirical data spanning**  
8 **both wide temporal and geographical scales. Here, using phylogenetic methods,**  
9 **temperature, and 21,795 globally distributed occurrences for 158 Clupeiform fish species,**  
10 **we show that smaller fish have occurred in warmer waters for over 150 million years,**  
11 **across marine and freshwater realms. Smaller fish have historically moved the shortest**  
12 **distances and at low speeds. In addition, small fish display the lowest probability of giving**  
13 **rise to new species. Further, we found that fish species that displayed high speeds of**  
14 **geographical movement and rates of size evolution experienced higher rates of**  
15 **temperature change in their lineage. These results together with global warming predicts**  
16 **a future where smaller Clupeiform fish that have reduced ability to move over aquatic**  
17 **systems will be more prevalent. In turn, this will result in fewer species contributing to**  
18 **global biodiversity.**

19  
20 A great deal of scientific research seeks to anticipate the impact of human-induced global  
21 warming on Earth's biodiversity<sup>1-5</sup>. Compelling evidence suggests that global warming will  
22 increase species extinction risk<sup>6-8</sup>, but there are hints in the literature pointing to the idea that  
23 species have several alternative strategies which might enable them to survive such  
24 adversity<sup>2,3,9,10</sup>. Local adaptive changes to decrease body size or tracking of suitable  
25 environmental conditions over geographic space have emerged as common responses allowing  
26 species survival, especially in fish<sup>8,11-19</sup>. However, it is unknown to what extent fish get smaller  
27 with warming<sup>20</sup> and how these climate-induced changes in size will impact the ability of species  
28 to track optimal environmental conditions over aquatic systems, i.e., species dispersal ability<sup>4,5,10</sup>.  
29 Furthermore, the consequences that the interaction between temperature, size, and dispersal  
30 ability may have on speciation is less explored, even though speciation is the principal buffer  
31 preventing biodiversity loss in the face of species extinction<sup>21</sup>.

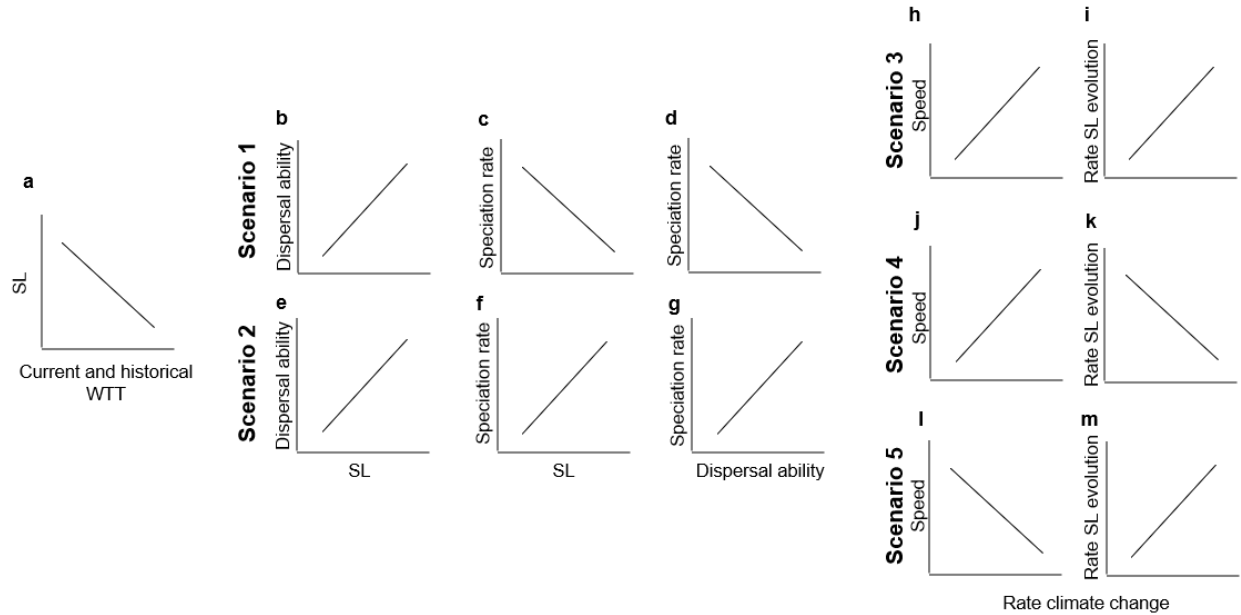
32  
33 Based on previous knowledge, we expect a positive association between fish size and dispersal  
34 ability given that bigger species are more efficient in terms of consuming energy for long-distance  
35 dispersals<sup>22</sup>, and their correlated life history strategies promote resilience under unpredictable  
36 environments<sup>23</sup>. Moreover, population genetics theory postulates that organisms with a high  
37 capacity to move can increase the gene flow within species; predicting a low probability of  
38 population divergence and speciation<sup>24</sup>. When these predictions are taken together, it is expected  
39 that the evolution of smaller fish under global warming (Fig. 1a) will decrease their dispersal ability  
40 (Fig 1b) but increase the rate at which they contribute with new species to biodiversity by local  
41 genetic differentiation (Fig. 1c and d). Nevertheless, there is a big gap between theoretical  
42 expectations and evidence owing to the lack of combined data on size evolution, temperature  
43 change, species dispersal ability and speciation rates. This patchy evidence comes from the fact  
44 that, first, the relationships between size, dispersal, and temperature change have only been  
45 evaluated across small temporal scales (i.e. decades)<sup>12,13,17-20,25,26</sup>, where the process of  
46 speciation cannot be observed. Second, species movement is notoriously difficult to quantify<sup>27-29</sup>  
47 so that most studies use data from extremely few individuals within species, measured in recent  
48 decades<sup>19</sup>.

49  
50 Here, for the first time, we test these predictions (and potential alternatives; Fig. 1) in Clupeiformes  
51 - a highly diverse order of fishes with a worldwide distribution, inhabiting the marine and  
52 freshwater realms<sup>30</sup> (Supplementary Figure 1). Clupeiformes include some of the most important  
53 species for fisheries<sup>31</sup>, such as the anchovy (*Engraulis ringens*), Atlantic herring (*Clupea*

54 *harengus*), Japanese pilchard (*Sardinops melanostictus*), Pacific herring (*Clupea pallasii*), and the  
55 South American pilchard (*Sardinops sagax*). We evaluated the relationship between water  
56 temperature tolerance (WTT) and standard length (SL) across the nodes of the Clupeiformes  
57 phylogenetic tree spanning ~150 Myr of evolutionary history (Supplementary Figure 2), and  
58 across their full global distribution in the present (Supplementary Figure 1). We estimated the  
59 *posterior* distribution of WTT values at phylogenetic nodes, which represent the subset of  
60 temperatures to which each species was adapted at the given node age. This does not represent  
61 a climate model-based or proxy-based measurement of paleotemperature *per se*. To evaluate the  
62 relationship between WTT, SL and the species ability to move over aquatic systems we inferred  
63 the historical distance and speed of fish historical movement in a three-dimensional space, using  
64 the Geo (Geographical) model<sup>32</sup> (Methods). This phylogenetic model estimates the *posterior*  
65 distribution of the estimated ancestral geographical locations for all nodes in a time-calibrated  
66 phylogenetic tree – allowing us to have a measure of the distance each species moved per-time  
67 unit (speed). Then, we evaluated the effect of SL and dispersal ability on Clupeiformes tip  
68 speciation rates.

69  
70 Our approach also provides information on the rate at which WTT has changed over lineages  
71 evolutionary history (phylogenetic branch). Thus, we can uniquely seek to know the range of rates  
72 at which the thermal environment of fish has changed (how fast) which, in turn, can reveal how  
73 quickly a species adapts. Studying species responses to the rate at which their thermal  
74 environmental change is now more pertinent than ever given the alarming accelerating-rates of  
75 heating of the oceans<sup>33</sup> and because species and populations respond differently when faced with  
76 a fast or slow change in their environment<sup>34,35</sup>.

77  
78 If higher temperatures select smaller fish, we expect to observe a negative relationship between  
79 SL and WTT over both evolutionary history and across extant species (Fig. 1a). If size reductions  
80 under global warming decrease the ability to move and increase the probability of speciation, we  
81 expect to observe a positive relationship between dispersal ability and SL (Fig. 1b) and a negative  
82 effect of SL and dispersal ability on speciation rate (Fig. 1c, d; Scenario 1). We evaluated an  
83 alternative scenario in which SL reductions and low dispersal ability decrease the probabilities of  
84 speciation so that we expect a positive effect of SL and dispersal ability on speciation rate (Fig.  
85 1f, g, Scenario 2). This alternative scenario can have support if dispersal promotes geographical  
86 expansion which increase the probabilities of vicariant speciation (range fragmentation by a  
87 physical barrier)<sup>36</sup>. Finally, if the rate of climate change can additionally modulate species  
88 dispersal and adaptation, the rate of WTT change should have a significant effect of on both the  
89 speed of movement and the rate of SL evolution (Fig. 1h-m). The slope of the relationships  
90 between rates will differ depending on how species respond when climate changes faster. Both  
91 slopes should be positive if species move faster and evolve rapidly (Fig. 1h, i; Scenario 3) –  
92 indicating rapid evolution away from its original location. The slopes can be positive and negative  
93 if species move faster and evolve slowly (Fig. 1j, k; Scenario 4) – indicating geographic tracking  
94 of optimal environmental conditions; and the slopes can be negative and positive if species move  
95 slowly and evolve rapidly (Fig. 1l, m; Scenario 5) – indicating rapid local adaptation.



**Figure 1. Global warming can impact fish species in multiple ways.** a, a negative relationship between standard length (SL) and water temperature tolerance (WTT), across the phylogeny and the extant global distribution of fish, support the idea that warmer temperatures have selected small fish over million years and at wide geographical scales. b - d, if small fish are less likely to disperse but more prone to speciate we should observe a positive relationship between dispersal ability and SL (b) and a negative effect of dispersal ability and SL on speciation rate (c, d). e - g, if small fish with lower dispersal ability are less prone to speciate we should observe a positive effect of dispersal ability and SL on speciation rates (h, j). h - m, additionally, species can respond differently to the rate at which temperature changes. When temperature changes faster species can move faster and adapt rapidly (h, i); move faster and adapt slowly (j, k); or move slowly and adapt rapidly (i, m).

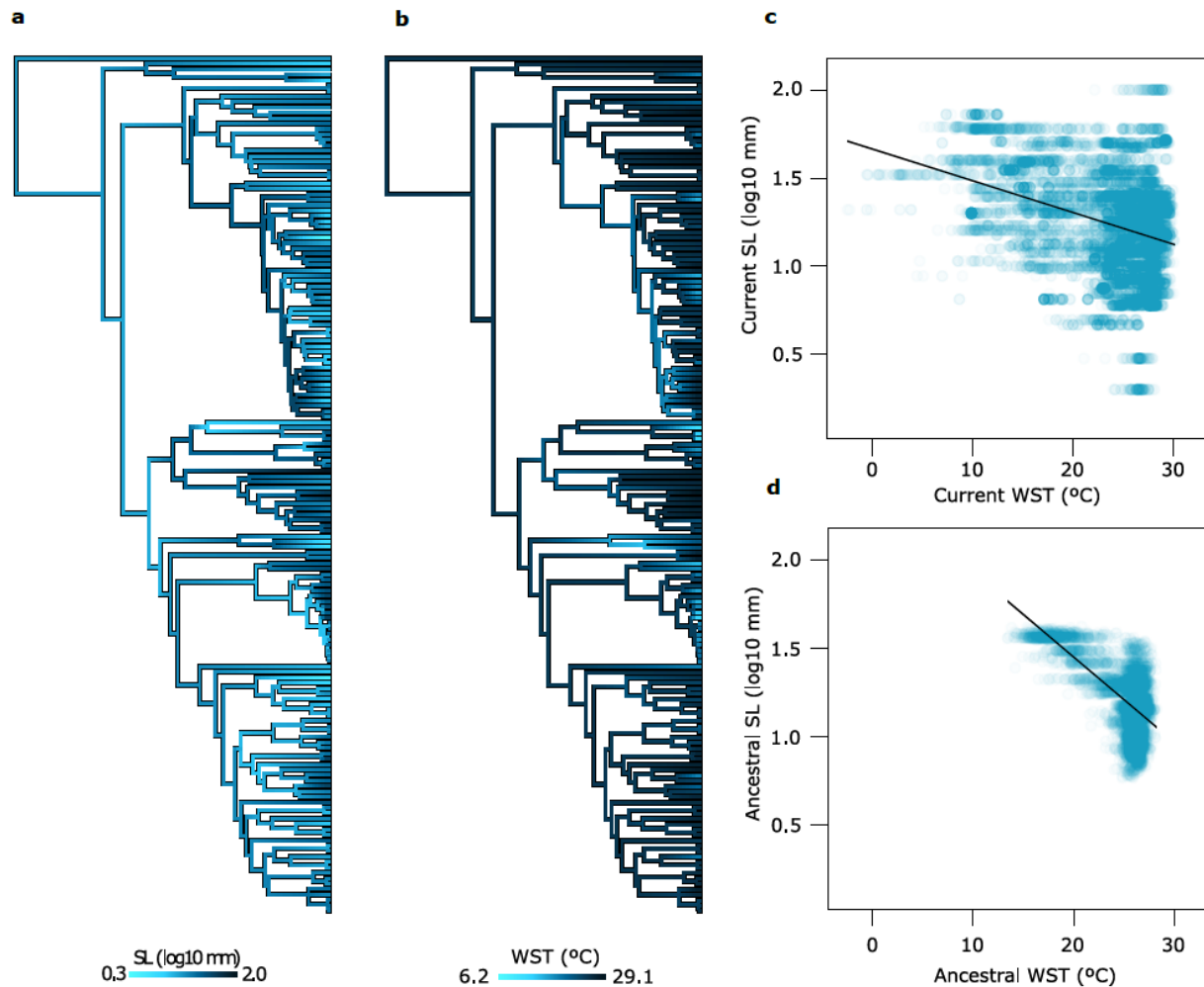
### SL and WTT over current and historical time

We studied the relationship between fish SL and WTT over their extant geographic distribution using the phylogenetic variable rates regression model<sup>37</sup> (Methods). This approach enables the simultaneous estimation of both an overall relationship between SL as a function of WTT across extant species, and any significant shifts in the rate of SL evolution that apply to the phylogenetically structured residual variance in the relationship. We also included the type of migration (diadromous and non-diadromous) as an additional binary variable in the regression, as previous studies show that diadromous fish are larger on average<sup>31</sup>. We used a Bayesian approach, that allows the estimation of regression coefficients while sampling the WTT data within each species. With this approach we can effectively evaluate the effect of WTT on SL while considering the temperature variability over the entire native distributional range of each species (Methods).

Results show that WTT has a significant negative effect on SL across the current geographic distribution of Clupeiformes (Fig. 2a;  $P_{\text{MCMC}} = 0.001$ ). This reveals that smaller Clupeiformes are found in warmer WTT, supporting the “*temperature-size rule*”<sup>38</sup>. Diadromous species were significantly larger than non-diadromous species on average (Supplementary Table 1;  $P_{\text{MCMC}} = 0$ ). Additionally, the variable rate regression did not detect any significant shifts in the rate of SL evolution, and fish SL was better explained by Brownian motion on the scaled phylogeny according to the Pagel’s Lambda ( $\lambda$ ) parameter (Fig. 2a; Supplementary Table 1).

129 To study the relationship between fish size and temperature in the deep past, we evaluated the  
130 relationship between the *posterior* sample of SL and WTT reconstructed at phylogenetic nodes,  
131 which comprises a temporal window of ~150 Myr. To conduct this analysis, we, firstly, inferred  
132 the *posterior* distribution of ancestral states of SL across nodes of the  $\lambda$ -scaled phylogeny (Fig.  
133 2b; Methods). Secondly, we inferred the *posterior* distribution of ancestral WTT across nodes of  
134 the rate-scaled phylogeny (Fig. 2c) obtained from the variable rate regression between WTT and  
135 absolute latitude across the 21,795 occurrence records (Methods; Supplementary Table 2). We  
136 found a significant negative association between the *posterior* sample of ancestral SL and WTT  
137 (Fig. 2d;  $P_{\text{MCMC}} = 0$ ), which support that Clupeiformes evolved smaller sizes under warmer WTT  
138 for over 150 Myr (Fig. 2d). Our results agree with the theoretical expectations (Fig. 1a), supporting  
139 that warmer temperatures select for smaller Clupeiform fish across large temporal and spatial  
140 scales.

141  
142 Finally, the variable rate regression for WTT indicates that the lower rate of temperature change  
143 at which Clupeiformes have adapted is  $0.0014\text{ }^{\circ}\text{C Myr}^{-1}$ , while the upper rate is  $0.79\text{ }^{\circ}\text{C Myr}^{-1}$   
144 ( $0.000000014$  and  $0.0000079\text{ }^{\circ}\text{C per decade}$ , respectively). These historical rates of change of  
145 WTT, given our data and approach, are far lower than the average rates of global warming that  
146 the planet is experiencing in the last decades;  $0.07\text{ }^{\circ}\text{C per decade}$  since 1880 to 1981, and  $0.18$   
147  $^{\circ}\text{C per decade}$  since 1981 (according to the NOAA 2019 Global Climate Summary). These results  
148 are comparable to that observed in terrestrial vertebrates<sup>39</sup>. The difference in rates of thermal  
149 change we observe might be because of the difference in time scale – millions of years vs  
150 decades. However, what is relevant in our results is that the estimated rates of WTT change per  
151 phylogenetic branch represent the rates of thermal change to which all species have adapted  
152 during their entire life. The thermal environments where species live are highly heritable at  
153 phylogenetic scales<sup>35,39</sup>, so if some species kept pace with rates of thermal change equal or faster  
154 than actual rates, over their entire life, then our methodology is highly likely to detect it. Together,  
155 if species are not able to track optimal environmental conditions, then a great part of biodiversity  
156 will not be able to adapt to the actual rates of local temperature change.



157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170

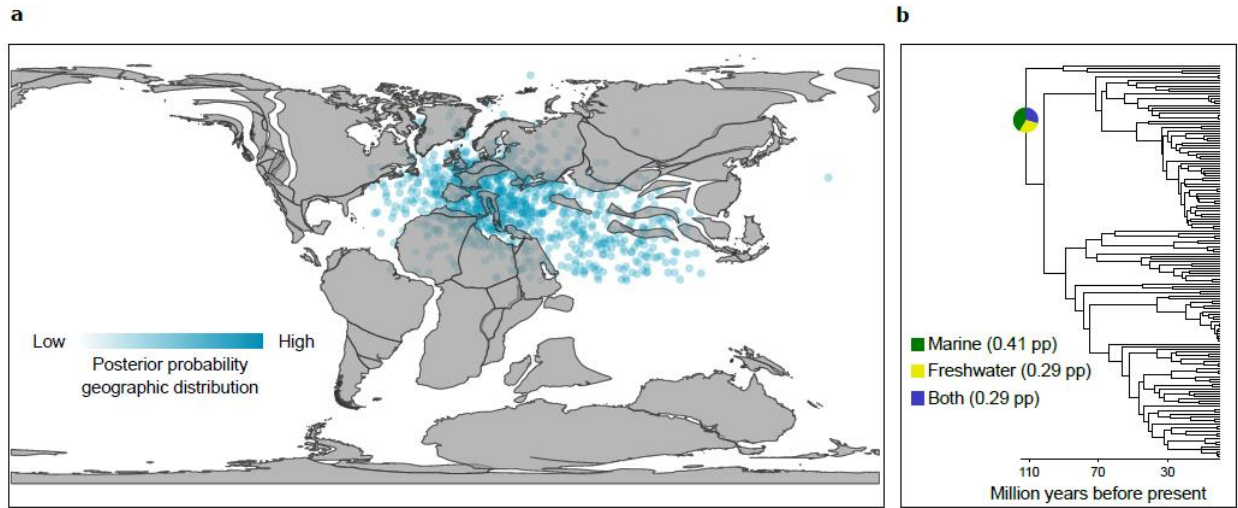
**Figure 2. Clupeiformes evolved smaller size in warmer temperatures for million years and in recent times.** **a, b.** Clupeiformes phylogenetic time tree with branches coloured according to the ancestral states for SL (**a**) and WTT (**b**). Ancestral states were estimated using the  $\lambda$ -model and the variable rate regression model for SL and WTT, respectively. **c.** Bayesian phylogenetic generalized least squares sustain that SL and WTT are negatively correlated across extant species ( $P_{\text{MCMC}} = 0.001$ ;  $n = 158,000$  observations sampled from extant species data). The black line represents the posterior mean slope of the phylogenetic regression, which was estimated while sampling within species WTT data. **d.** Bayesian generalized least squares shows a significant negative correlation between the ancestral SL and WTT values across nodes ( $P_{\text{MCMC}} = 0$ ;  $n = 157,000$  observations sampled from the *posterior* distribution of the estimated ancestral states across phylogenetic nodes). These results support the prediction in Fig. 1a. Line equation in **c**:  $y = 1.3 + 0.15(\text{Diadromus}) - 0.0077(\text{WTT})$ . Line equation in **d**:  $y = 2.44 - 0.047(\text{WTT})$ .

### SL and dispersal ability

171  
172 The geographic analyses support a model with significant variation in the speed of fish movement  
173 across phylogenetic branches (Supplementary Table 3). This implies that the current spatial  
174 diversity of Clupeiformes has been assembled by species dispersal at variable speed from the  
175 location of the most recent common ancestor (MRCA) of the group (Clupeoidei, after excluding  
176 *Denticeps clupeoides*; see Methods). The highest *posterior* density for the geographic distribution  
177 of the MRCA indicates that this ancestral species was distributed between the western Tethys  
178 Ocean and eastern of Proto Atlantic Ocean, mainly between Eurasia and Africa, around 111 Mya  
179 (Fig. 3a). The ancestral reconstruction of the habitat type for the MRCA supports that it was more

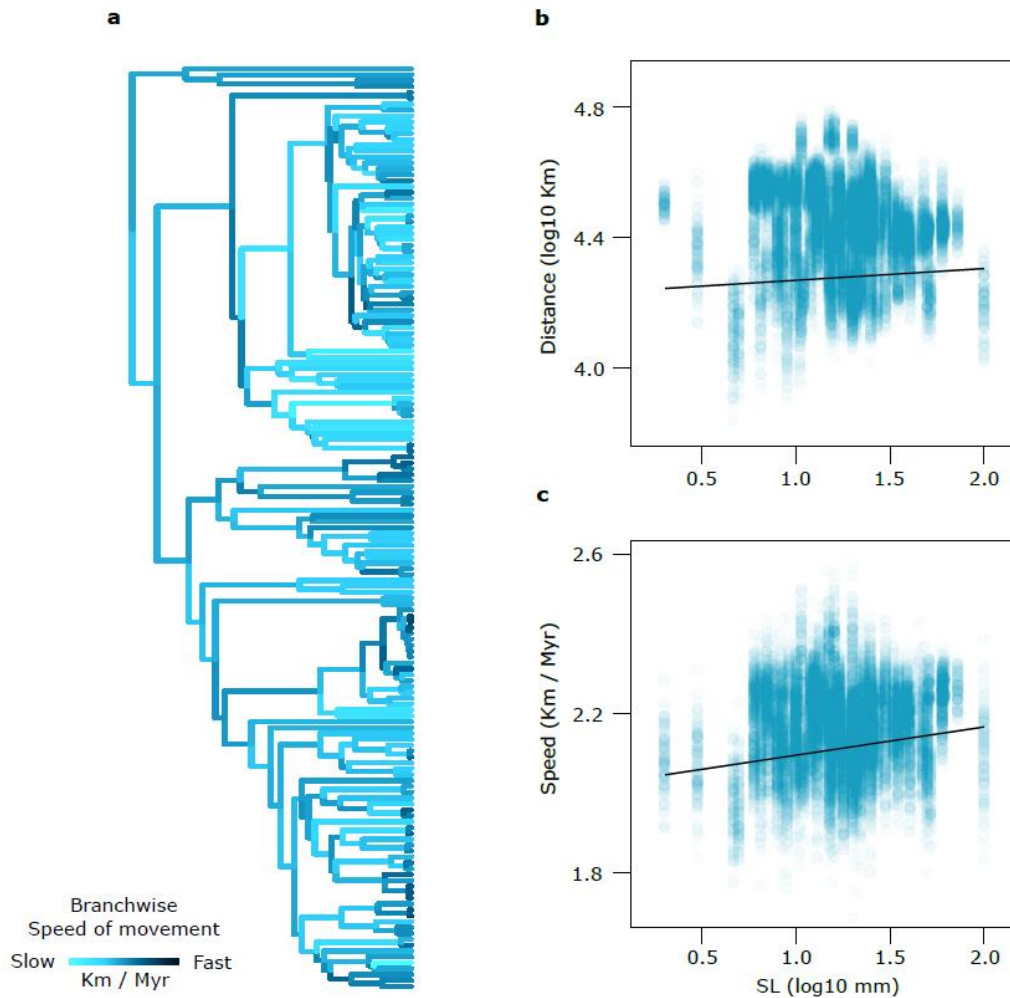


180 likely a marine species (Fig. 3b). However, there are some *posterior* coordinates that fall on  
 181 continents, and there is also some *posterior* probability indicating that the MRCA inhabited  
 182 freshwater realms (0.29; Fig. 3b) or both freshwater and marine realms (0.29; Fig. 3b). This  
 183 degree of uncertainty in the estimation of location and the habitat type (see Supplementary Figure  
 184 3 for more node locations) suggests that the MRCA was a marine species with ability to occupy  
 185 the freshwater space. The biology of the living species of Clupeoidei support this notion because  
 186 there are living species adapted to live in both freshwater (rivers, lakes, swamps) and saline  
 187 waters (estuaries, bays, sea).  
 188



189  
 190 **Figure 3. The ancestor of Clupeoidei was distributed across the western Tethys Ocean and the**  
 191 **eastern of the Proto Atlantic Ocean 111 million years ago. a.** *posterior* geographic distribution of the  
 192 phylogenetic node representing the ancestor of Clupeoidei. **b.** habitat type for the ancestor of Clupeoidei.  
 193 pp: *posterior* probability.  
 194

195 When we calculate the total distance that each species dispersed - along the lineage leading from  
 196 the MRCA to the living species (Supplementary Figure 4) - we observe that the shortest distance  
 197 was taken by the lineage of *Chirocentrus dorab* (9,608 km) while the largest distance by the  
 198 lineage of *Engraulis australis* (53,885 km). Note that this total distance was calculated across the  
 199 geographic centroids of the *posterior* locations at each phylogenetic node. However, the distances  
 200 dispersed vary due to the uncertainty in the estimation of species at each phylogenetic node  
 201 (Figure 3; Supplementary Figure 3). Thus, this uncertainty should be considered when studying  
 202 the correlates of species movement. We evaluated the effect of SL on the total distance moved  
 203 for each species from the MRCA (pathwise distance; Methods), and the median of the branch-  
 204 specific speed of movement along the path that links the MRCA with extant species (pathwise  
 205 speed; Methods) considering the uncertainty in the estimated ancestral locations. These  
 206 relationships were evaluated using Bayesian phylogenetic regression models that include the  
 207 *posterior* sample of 1,000 pathwise distances and speeds for each species in the estimation of  
 208 regression coefficients (Methods). Results show that SL correlates positively with both the  
 209 pathwise distances and the pathwise speed of movement (Fig. 4b and c; Supplementary Table 4  
 210 and 5, respectively). There were no significant differences in either the mean pathwise distances  
 211 or the pathwise speed of movement travelled by diadromous and non-diadromous species  
 212 (Supplementary Table 4 and 5). These results agree with theoretical expectations (Fig. 1b and  
 213 e). Smaller fish have had a reduced ability to disperse through water bodies over their  
 214 evolutionary history. They may find it hard to track suitable temperatures over geological time,  
 215 thus making them more prone to extinction if they cannot keep pace with the actual rates of local  
 216 heating of the oceans.

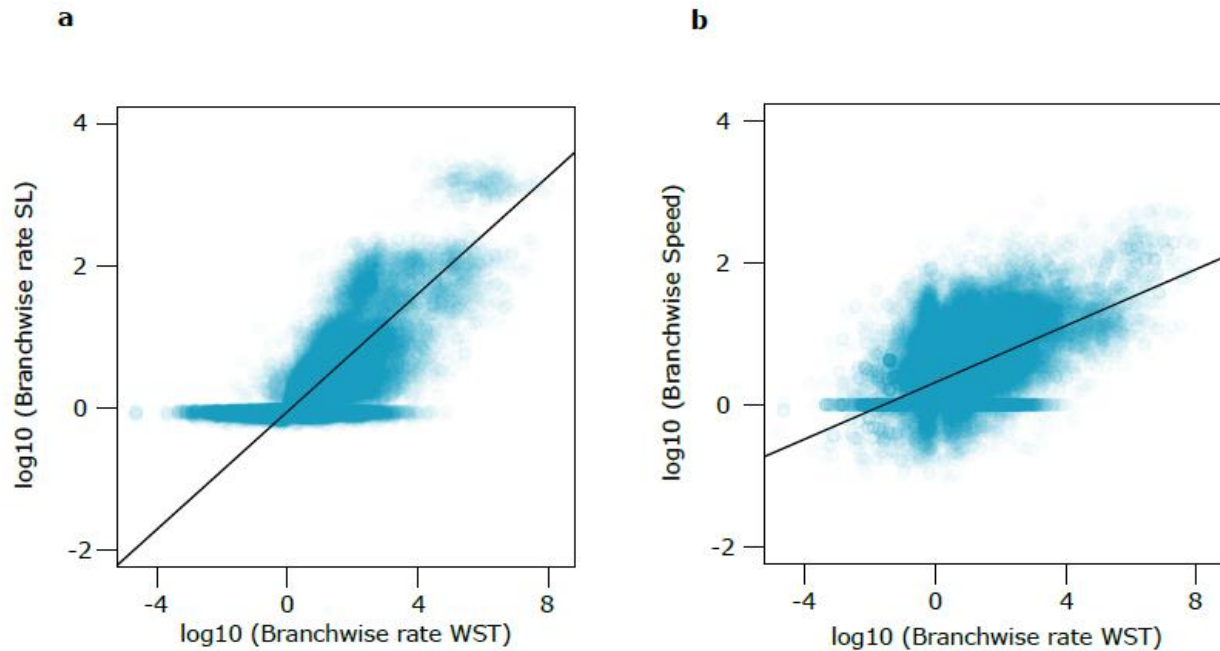


218 **Figure 4. Fish dispersal ability depend on body size.** a. Clupeiformes phylogenetic tree with branches  
 219 coloured according to the speed of movement. b. Bayesian phylogenetic generalized least squares show  
 220 that pathwise distance correlates positively with SL ( $BF > 5$ ;  $n = 157,000$  observations sampled from  
 221 species data). c. SL has also a significant positive effect on pathwise speed of movement ( $BF > 10$ ;  $n =$   
 222 157,000 observations sampled from species data). These results support the prediction in Fig. 1b and e.  
 223 Black lines (b and c) represent the mean slope estimated from the *posterior* distribution of phylogenetic  
 224 slopes. Line equation in b:  $y = 4.21 + 0.026(SL)$ . Line equation in c:  $y = 2.02 + 0.071(SL)$ .  
 225  
 226

### 227 **Fish response to the historical rate of WTT change**

228 We evaluated the effect that the rates of WTT change may have on both the rates of SL evolution  
 229 and the speed of movement across all branches of the Clupeiformes phylogeny, using Bayesian  
 230 GLS regressions that use samples of the data. We included the *posterior* sample of 1,000  
 231 branchwise rates estimated at each phylogenetic branch as sample data in regression analyses  
 232 (Methods). All branchwise rates were estimated by dividing the scaled branches (with the  $\lambda$ -model  
 233 for SL, the variable rate regression model for WTT, and the variable rate Geo model for speed)  
 234 with original branch lengths measured in time. The rate of WTT change had a positive effect on  
 235 both the rate of SL evolution and the speed of fish movement ( $P_{MCMC} = 0$ , Fig. 5a, b), meaning  
 236 that the SL of Clupeiformes have evolved rapidly, and they have dispersed faster when the  
 237 temperature of their aquatic environments changed at higher rates. These results agree with  
 238 theoretical expectation in Fig. 1h and b (Scenario 3), indicating that clupeiforms have evolved

239 rapidly, away from its original location, when climate changed faster. Under accelerated rates of  
 240 warming clupeiforms will evolve rapidly towards smaller sizes, concomitantly losing their ability to  
 241 move as size and temperature correlates negatively.  
 242

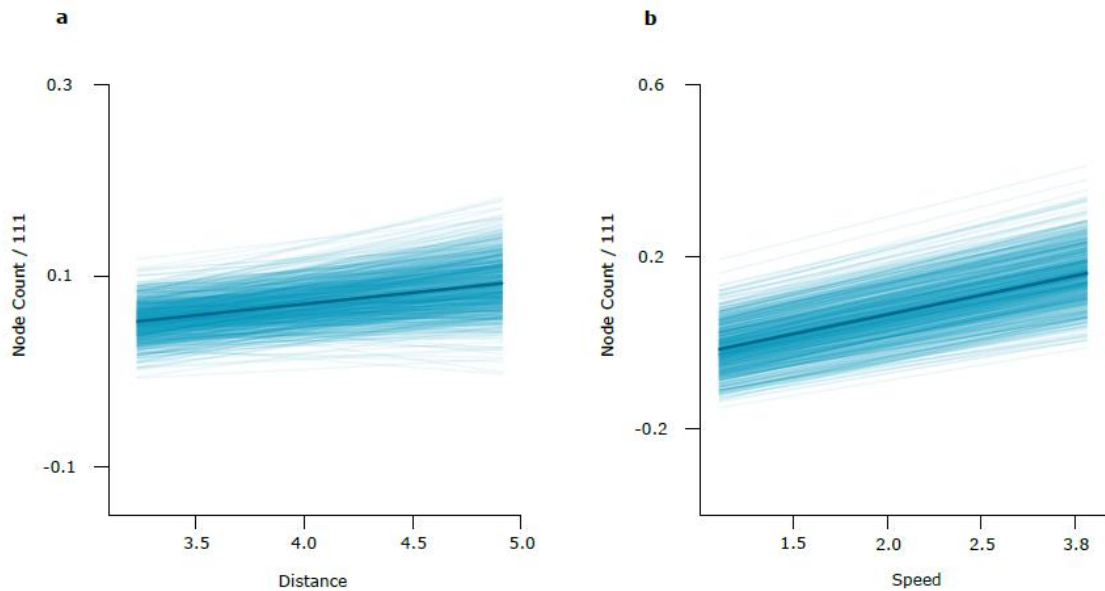


243 **Figure 5. Clupeiformes have evolved rapidly and moved faster when temperature changed at higher**  
 244 **rates.** **a.** Bayesian generalized least squares support that the branchwise rates of SL evolution are  
 245 positively correlated with the branchwise rates of WTT change ( $P_{\text{MCMC}} = 0$ ;  $n = 314,000$  phylogenetic  
 246 branches). **b.** The branchwise speed of fish movement are also positively correlated with the branchwise  
 247 rates of WTT change ( $P_{\text{MCMC}} = 0$ ;  $n = 312,000$  phylogenetic branches). These results support the predictions  
 248 in Fig. 1h and i. Black lines represent the mean slope estimated from the *posterior* distribution of  
 249 phylogenetic slopes. Line equation in **a**:  $y = -0.05 + 0.41(\text{Branchwise rate WTT})$ . Line equation in **b**:  $y =$   
 250  $0.31 + 0.2(\text{Branchwise rate WTT})$ .  
 251

252  
 253 **Effect of SL and dispersal ability on speciation rates**

254 We evaluated the relationship between speciation with dispersal ability and SL of Clupeiformes.  
 255 We used Bayesian phylogenetic regression models that include the uncertainty in parameter  
 256 estimation and samples of dispersal ability within species (Methods). Our results show that the  
 257 independent additive effect of pathwise distance and pathwise speed were significant ( $P_{\text{MCMC}} =$   
 258  $0.04$  and  $0$  respectively; Supplementary Table 6) – species that move longer distances and faster  
 259 were more likely to originate new species – supporting theoretical predictions in Fig. 1d. SL did  
 260 not have a significant effect on speciation when its independent additive effect or their interaction  
 261 with dispersal ability was evaluated (Supplementary Table 6), rejecting theoretical predictions in  
 262 Fig. 1c and f. These results suggest that fish SL, by its positive association with dispersal ability,  
 263 has an indirect effect on speciation rates. We speculate that SL by itself does not related to  
 264 speciation rate in clupeiform fish because natural selection on SL has not split populations in two  
 265 or more isolated groups (i.e., selection was not disruptive). Considering that warmer temperatures  
 266 have selected for smaller fish (and colder temperatures for bigger one) we also speculate that  
 267 changes in temperature generated a process of directional selection on fish SL – moving the  
 268 population mean towards smaller values when the temperature increased and vice versa.  
 269 Directional selection is unlikely to split populations, which can explain why we do not observe a  
 270 significant relationship between SL and speciation rates. The fact that higher dispersal ability does

271 correlate positively with speciation rates point to the idea that speciation in clupeiforms was  
 272 determined principally by geographic process - species moving further and faster could be more  
 273 likely to experience geographic isolation. Taking together, the speciation rates of smaller fish that  
 274 move slowly are lower than the speciation rates of their larger counterparts that moved faster and  
 275 larger distances. A scenario of smaller fish under global warming may cause a decrease of  
 276 speciation rate in fish, a phenomenon that can impoverish future biodiversity<sup>21</sup>.  
 277  
 278  
 279



280  
 281 **Figure 6. Clupeiformes with lower dispersal abilities have lower probabilities of originate new**  
 282 **species. a - b.** The Bayesian phylogenetic generalized least squares show that the pathwise distance of  
 283 movement and the pathwise speed of movement has a positive effect on speciation ( $P_{MCMC} = 0.04$  and 0,  
 284 respectively;  $n = 157,000$  observations sampled from species data). These results support the prediction in  
 285 Fig. 1g. Lighter lines show the *posterior* distribution of slopes and dark lines shows the *posterior* mean  
 286 slopes. These slopes were estimated while sampling the pathwise distance and speed within species  
 287 (Methods). Line equation for **a** and **b**:  $y = -0.21 + 0.023(\text{Distance}) + 0.09(\text{Speed})$ . The node count values  
 288 were divided by the tree length after excluding *D. clupeoides* (111 Myr).  
 289

290 **Conclusion**

291 Global change poses double jeopardy for fish body size, as both overfishing<sup>40</sup> and climate drive  
 292 populations towards smaller sizes. The phenomena of fish shrinking when facing hotter waters is  
 293 general in the evolutionary history of Clupeiformes and over their entire worldwide geographic  
 294 distribution. Provided that smaller Clupeiform fish adapted to warmer conditions are less capable  
 295 of disperse and in turn less able to originate new species, the scenario of global warming could  
 296 limit both their ability to find optimal environments to live and their capacity to buffer their  
 297 increasing extinction risk by the process of speciation. Furthermore, Clupeiform fish living in the  
 298 present are the survivors of a long evolutionary history under variable rates of temperature  
 299 change. They have responded to such historical changes by SL adaptation and dispersal at  
 300 variable rate and speed, respectively. However, such evolutionary processes have never involved  
 301 the current accelerating rates of heating of the water bodies. Clupeiformes will probably face an  
 302 increasing risk of extinction. This conclusion can be generalized to other fish if body size, dispersal  
 303 abilities, and speciation rates relates to each other as in Clupeiformes.  
 304  
 305

306 **References**

- 307
- 308 1. Parmesan, C. Ecological and Evolutionary Responses to Recent Climate Change. *Annu. Rev. Ecol. Evol. Syst.* **37**, 637–669 (2006).
- 309
- 310 2. Sheridan, J. A. & Bickford, D. Shrinking body size as an ecological response to climate
- 311 change. *Nat. Clim. Chang.* **1**, 401–406 (2011).
- 312 3. Gardner, J. L., Peters, A., Kearney, M. R., Joseph, L. & Heinsohn, R. Declining body size:
- 313 A third universal response to warming? *Trends Ecol. Evol.* **26**, 285–291 (2011).
- 314 4. McCauley, S. J. & Mabry, K. E. Climate change, body size, and phenotype dependent
- 315 dispersal. *Trends Ecol. Evol.* **26**, 554–555 (2011).
- 316 5. Norberg, J., Urban, M. C., Vellend, M., Klausmeier, C. A. & Loeuille, N. Eco-evolutionary
- 317 responses of biodiversity to climate change. *Nat. Clim. Chang.* **2**, 747–751 (2012).
- 318 6. Amigo, I. The Amazon’s fragile future. **578**, 505–507 (2020).
- 319 7. Reddin, C. J., Nätscher, P. S., Kocsis, Á. T., Pörtner, H. O. & Kiessling, W. Marine clade
- 320 sensitivities to climate change conform across timescales. *Nat. Clim. Chang.* **10**, (2020).
- 321 8. Comte, L. & Olden, J. D. Climatic vulnerability of the world’s freshwater and marine
- 322 fishes. *Nat. Clim. Chang.* **7**, 718–722 (2017).
- 323 9. Skelly, D. K. *et al.* Evolutionary responses to climate change. *Conserv. Biol.* **21**, 1353–
- 324 1355 (2007).
- 325 10. Chen, I., Hill, J. K., Ohlemüller, R., Roy, D. B. & Thomas, C. D. Rapid Range Shifts of
- 326 Species Associated with High Levels of Climate Warming. *Science (80-. ).* **1024**, 17–20
- 327 (2012).
- 328 11. Cheung, W. W. L. *et al.* Projecting global marine biodiversity impacts under climate
- 329 change scenarios. *Fish Fish.* **10**, 235–251 (2009).
- 330 12. Cheung, W. W. L. *et al.* Shrinking of fishes exacerbates impacts of global ocean changes
- 331 on marine ecosystems. *Nat. Clim. Chang.* **3**, 254–258 (2013).
- 332 13. Crozier, L. G. & Hutchings, J. A. Plastic and evolutionary responses to climate change in
- 333 fish. *Evol. Appl.* **7**, 68–87 (2014).
- 334 14. Travis, J. M. J. *et al.* Dispersal and species’ responses to climate change. *Oikos* **122**,
- 335 1532–1540 (2013).
- 336 15. Pauly, D. & Cheung, W. W. L. Sound physiological knowledge and principles in modeling
- 337 shrinking of fishes under climate change. *Glob. Chang. Biol.* **24**, e15–e26 (2018).
- 338 16. Tamario, C., Sunde, J., Petersson, E., Tibblin, P. & Forsman, A. Ecological and
- 339 Evolutionary Consequences of Environmental Change and Management Actions for
- 340 Migrating Fish. *Front. Ecol. Evol.* **7**, 1–24 (2019).
- 341 17. Ljungström, G., Claireaux, M., Fiksen, Ø. & Jørgensen, C. Body size adaptations under
- 342 climate change: zooplankton community more important than temperature or food
- 343 abundance in model of a zooplanktivorous fish. *Mar. Ecol. Prog. Ser.* **636**, 1–18 (2020).
- 344 18. Daufresne, M., Lengfellner, K. & Sommer, U. Global warming benefits the small in
- 345 aquatic ecosystems. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 12788–12793 (2009).
- 346 19. Lenoir, J. *et al.* Species better track climate warming in the oceans than on land. *Nat.*
- 347 *Ecol. Evol.* (2020) doi:10.1038/s41559-020-1198-2.
- 348 20. Audzijonyte, A. *et al.* Fish body sizes change with temperature but not all species shrink
- 349 with warming. *Nat. Ecol. Evol.* 1–6 (2020) doi:10.1038/s41559-020-1171-0.
- 350 21. Rosenzweig, M. L. Loss of speciation rate will impoverish future diversity. *Proc. Natl.*
- 351 *Acad. Sci. U. S. A.* **98**, 5404–5410 (2001).
- 352 22. Burns, M. D. & Bloom, D. D. Migratory lineages rapidly evolve larger body sizes than non-
- 353 migratory relatives in ray-finned fishes. *Proceedings. Biol. Sci.* **287**, 20192615 (2020).
- 354 23. Comte, L. & Olden, J. D. Evidence for dispersal syndromes in freshwater fishes. *Proc. R.*
- 355 *Soc. B Biol. Sci.* **285**, (2018).
- 356 24. Bohonak, A. J. Dispersal, gene flow, and population structure. *Q. Rev. Biol.* **74**, 21–45

- (1999).
- 358 25. Perry, A. L., Low, P. J., Ellis, J. R. & Reynolds, J. D. Climate change and distribution  
359 shifts in marine fishes. *Science (80-. )*. **308**, 1912–1915 (2005).
  - 360 26. Pinsky, M. L., Worm, B., Fogarty, M. J., Sarmiento, J. L. & Levin, S. A. Marine taxa track  
361 local climate velocities. *Science (80-. )*. **341**, 1239–1242 (2013).
  - 362 27. Stevens, V. M. *et al.* A comparative analysis of dispersal syndromes in terrestrial and  
363 semi-terrestrial animals. *Ecol. Lett.* **17**, 1039–1052 (2014).
  - 364 28. Dieckmann, U., O'Hara, B. & Weisser, W. The evolutionary ecology of dispersal. *Trends*  
365 *Ecol. Evol.* **14**, 88–90 (1999).
  - 366 29. Kokko, H. & López-Sepulcre, A. From individual dispersal to species ranges:  
367 Perspectives for a changing world. *Science (80-. )*. **313**, 789–791 (2006).
  - 368 30. Lavoué, S., Miya, M., Musikasinthorn, P., Chen, W. J. & Nishida, M. Mitogenomic  
369 Evidence for an Indo-West Pacific Origin of the Clupeoidei (Teleostei: Clupeiformes).  
370 *PLoS One* **8**, (2013).
  - 371 31. Bloom, D. D., Burns, M. D. & Schriever, T. A. Evolution of body size and trophic position  
372 in migratory fishes: A phylogenetic comparative analysis of Clupeiformes (anchovies,  
373 herring, shad and allies). *Biol. J. Linn. Soc.* **125**, 302–314 (2018).
  - 374 32. O'Donovan, C., Meade, A. & Venditti, C. Dinosaurs reveal the geographical signature of  
375 an evolutionary radiation. *Nat. Ecol. Evol.* **2**, 452–458 (2018).
  - 376 33. Cheng, L. *et al.* Record-Setting Ocean Warmth Continued in 2019. *Adv. Atmos. Sci.* **37**,  
377 137–142 (2020).
  - 378 34. Pinek, L., Mansour, I., Lakovic, M., Ryo, M. & Rillig, M. C. Rate of environmental change  
379 across scales in ecology. *Biol. Rev.* **1**, (2020).
  - 380 35. Avaria-Llautureo, J., Hernández, C. E., Rodríguez-Serrano, E. & Venditti, C. The  
381 decoupled nature of basal metabolic rate and body temperature in endotherm evolution.  
382 *Nature* **572**, 651–654 (2019).
  - 383 36. Gaston, K. J. Species-range size distributions: products of speciation, extinction and  
384 transformation. *Philos. Trans. R. Soc. B Biol. Sci.* **353**, 219–230 (1998).
  - 385 37. Baker, J., Meade, A., Pagel, M. & Venditti, C. Positive phenotypic selection inferred from  
386 phylogenies. *Biol. J. Linn. Soc.* **118**, 95–115 (2016).
  - 387 38. Angilletta, M. J. & Dunham, A. E. The Temperature-Size Rule in Ectotherms: Simple  
388 Evolutionary Explanations May Not Be General. *Am. Nat.* **162**, 332–342 (2003).
  - 389 39. Quintero, I. & Wiens, J. J. Rates of projected climate change dramatically exceed past  
390 rates of climatic niche evolution among vertebrate species. *Ecol. Lett.* **16**, 1095–1103  
391 (2013).
  - 392 40. Pauly, D., Christensen, V., Dalsgaard, J., Froese, R. & Torres Jr, F. Fishing Down Marine  
393 Food Webs. *Science (80-. )*. **279**, 860–863 (1998).
  - 394 41. Rabosky, D. L. *et al.* An inverse latitudinal gradient in speciation rate for marine fishes.  
395 *Nature* **559**, 392–395 (2018).
  - 396 42. Whitehead, P. J. P. FAO Species Catalogue: Vol. 7 Clupeoid Fishes of the World. *FAO*  
397 *Fish. synopsis* **7**, 303 (1985).
  - 398 43. Charnov, E. L. & Berrigan, D. Evolution of life history parameters in animals with  
399 indeterminate growth, particularly fish. *Evol. Ecol.* **5**, 63–68 (1991).
  - 400 44. Önsoy, B., Tarkan, A. S., Filiz, H. & Bilge, G. Determination of the best length  
401 measurement of fish. *North. West. J. Zool.* **7**, 178–180 (2011).
  - 402 45. Mohseni, O. & Stefan, H. G. Stream temperature/air temperature relationship: A physical  
403 interpretation. *J. Hydrol.* **218**, 128–141 (1999).
  - 404 46. Morrill, J. C., Bales, R. C. & Conklin, M. H. Estimating stream temperature from air  
405 temperature: Implications for future water quality. *J. Environ. Eng.* **131**, 139–146 (2005).
  - 406 47. Sharma, S., Jackson, D. A., Minns, C. K. & Shuter, B. J. Will northern fish populations be  
407 in hot water because of climate change? *Glob. Chang. Biol.* **13**, 2052–2064 (2007).

- 408 48. Pagel, M., Meade, A. & Barker, D. Bayesian estimation of ancestral character states on  
409 phylogenies. *Syst. Biol.* **53**, 673–684 (2004).
- 410 49. Venditti, C., Meade, A. & Pagel, M. Multiple routes to mammalian diversity. *Nature* **479**,  
411 393–396 (2011).
- 412 50. Kocsis, Á. T. & Raja, N. B. chronosphere: Earth system history variables. (2020)  
413 doi:10.5281/zenodo.3530703.
- 414 51. Raftery, A. E. Hypothesis testing and model selection. in *Markov Chain Monte Carlo in*  
415 *Practice* (eds. Gilks, W., Richardson, S. & Spiegelhalter, D.) 163–187 (Chapman & Hall,  
416 1996).
- 417 52. Hijmans, R. J. geosphere: Spherical Trigonometry. R package version 1.5-10.  
418 <https://CRAN.R-project.org/package=geosphere>. (2019).
- 419 53. Harvey, M. G. & Rabosky, D. L. Continuous traits and speciation rates: Alternatives to  
420 state-dependent diversification models. *Methods Ecol. Evol.* **9**, 984–993 (2018).
- 421 54. Title, P. O. & Rabosky, D. L. Tip rates, phylogenies and diversification: What are we  
422 estimating, and how good are the estimates? *Methods Ecol. Evol.* **10**, 821–834 (2019).
- 423 55. Louca, S. & Pennell, M. W. Extant timetrees are consistent with a myriad of diversification  
424 histories. *Nature* **580**, 502–505 (2020).
- 425 56. Shafir, A., Azouri, D., Goldberg, E. E. & Mayrose, I. Heterogeneity in the rate of molecular  
426 sequence evolution substantially impacts the accuracy of detecting shifts in diversification  
427 rates. *Evolution (N. Y.)*. (2020) doi:<https://doi.org/10.1111/evo.14036>.
- 428 57. Ganzach, Y. Misleading Interaction and Curvilinear Terms. *Psychol. Methods* **2**, 235–247  
429 (1997).
- 430 58. Revell, L. J. phytools: An R package for phylogenetic comparative biology (and other  
431 things). *Methods Ecol. Evol.* **3**, 217–223 (2012).
- 432 59. Lunt, D. J. *et al.* Palaeogeographic controls on climate and proxy interpretation. *Clim.*  
433 *Past* **12**, 1181–1198 (2016).
- 434  
435  
436  
437

## 438 **Methods**

439 **Data.** Analyses were performed on the most recent time-calibrated phylogeny of 158  
440 Clupeiformes species (Supplementary Figure 2). This phylogeny was obtained from The Fish  
441 Tree of Life<sup>41</sup>. We used the maximum Standard Length (SL) in mm for these 158 species  
442 (Supplementary Table 7). We obtained the SL from FishBase and the FAO Species Catalogue  
443 for clupeoid fishes<sup>42</sup>. *Sensu* the FishBase System Glossary, the fish SL is the measurement from  
444 the most anterior tip of the body to the mid lateral posterior edge of the hypural plate (in fish with  
445 a hypural plate) or to the posterior end of the vertebral column (in fish lacking hypural plates). The  
446 maximum SL was used because of three reasons. First, maximum SL is preferred over mean SL  
447 because fishes have indeterminate growth<sup>43</sup>. Second, it is a more stable measure of size in  
448 teleosts to compare museum and collection samples<sup>44</sup>. Third, and most important, individuals that  
449 are commonly larger than the population average, and are outside the central distribution of size,  
450 are likely the individuals that allow the species to shift their geographic ranges<sup>4</sup>. 21,795  
451 georeferenced occurrences (Supplementary Figure 1; Supplementary Table 7) were obtained  
452 from marine and freshwater bodies (i.e., rivers and lakes) from Aquamaps  
453 (<https://www.aquamaps.org/>) and the IUCN (<https://www.iucnredlist.org/>) respectively. We  
454 obtained the geographic locations (within the native range) of 116 species available in Aquamaps,  
455 and locations within the polygon of distribution for 42 additional species available in the IUCN. To  
456 obtain the geographic locations from the IUCN, we sampled 100 random locations within each  
457 species polygon. All georeferenced occurrences were matched with information of water  
458 temperature, which represent water temperature tolerances for species (WTT; Supplementary

459 Table 7). For marine species, we used the mean annual sea surface temperature estimated from  
460 the Aquamaps database. For freshwater species, the mean annual air temperatures estimated  
461 from the WorldClim database (<https://worldclim.org/>) were used as a first-order proxy of the water  
462 surface temperature of the freshwater bodies<sup>45–47</sup>. By maximizing the number of locations and  
463 temperature records per species, instead of using single estimates (e.g., mean temperature at  
464 the geographic centroid of species distributional range) we can produce more precise estimates  
465 of both the ancestral locations and the ancestral thermal environments where Clupeiformes  
466 inhabited. Finally, information about the type of migration for each species (diadromous, non-  
467 diadromous) was obtained from Bloom et al<sup>31</sup> (Supplementary Table 7).

468  
469 **Inferring ancestral locations.** From the geographic locations within each species in the  
470 Clupeiformes phylogeny, we inferred the ancestral geo-distribution in a continuous, three-  
471 dimensional space. Ancestral locations were estimated for each phylogenetic node using the Geo  
472 model<sup>32</sup> in the computer program BayesTraits 3.0<sup>48</sup>. This model estimates the *posterior*  
473 distribution of ancestral locations measured in longitude and latitude, while sampling across all  
474 location-data within species, and considering the spherical nature of Earth. This natural  
475 assumption of the Earth as a spherical object avoids the erroneous calculation of distances  
476 between the inferred ancestral locations due to the non-continuity of the longitude scale. When  
477 based on a time-calibrated phylogeny, the Geo model simultaneously estimates the speed of  
478 species movement across each branch that links pairs of phylogenetic nodes (branchwise speed  
479 of movement). Additionally, the ancestral locations across phylogenetic nodes are estimated while  
480 considering the continuous variation in dispersal ability of each ancestral species – ranging from  
481 species quiescence (no movement), through constant movement in direct proportion of the  
482 passage of time, to fast species movement. Estimation of the branchwise speed of species  
483 movement are based on the variable rates model<sup>49</sup>, which detects shifts away from a background  
484 rate of evolution in continuous traits (expected under Brownian motion) in whole clades or  
485 individual branches. We also include data of the geographic locations of two Clupeiform fossils,  
486 one for the crown group of Engraulidae and another for the crown group of Dorosoma  
487 (Supplementary Figure 2). They were included as branches linked to the nodes where the two  
488 fossil belongs. We assigned ~zero branch-length (0.000001) to each fossil. The aim of assigning  
489 zero branch-length to each fossil is to ensure that the Geo model will not modify the branch so  
490 that the estimated location of the node will be at the fossil location with high accuracy and  
491 precision. Some variation will be present in the inference given the data of the remainder species  
492 in clade. The fossil data we used are those whose phylogenetic position at phylogenetic nodes  
493 are well known in The Fish Tree of Life<sup>41</sup>. The use of well-known node-fossils allowed us a more  
494 secure placing of paleo coordinates given that our methodological approach place fossil data at  
495 phylogenetic nodes. The geographic locations were extracted from the original papers describing  
496 the fossils. Then we reconstructed the paleo coordinates for the two fossils using the function  
497 *reconstruct* in the chronosphere R-package<sup>50</sup>. We used the age of the nodes for each fossil and  
498 the PALEOMAP model for the paleo coordinate reconstruction. Finally, we used the paleo  
499 coordinates as input in the Geo model analyses.

500  
501 We ran four MCMC chains for 250,000,000 iterations, sampling every 50,000 iterations, and  
502 discarding 200,000,000 as burn in. These procedures were conducted based on the Brownian  
503 motion (BM) model and the Variable Rates (VR) model (Supplementary Table 3). We checked for  
504 chain convergence using the Effective Sample Size (ESS) in Tracer v1.6, ensuring outputs with  
505 ESS > 200. The final sample includes 1,000 posterior locations for each phylogenetic node. We  
506 selected the model that fit the data better by means of Bayes factors (*BF*), using the marginal  
507 likelihoods estimated by stepping stone sampling. *BF* is calculated as the double of the difference  
508 between the log marginal likelihood of the complex model and the simple model. By convention,  
509 *BF* > 2 indicates positive support for the complex model, *BF* = 5–10 indicates strong support and



510 *BF* > 10 is considered very strong support<sup>51</sup>. We excluded the species *Denticeps clupeoides* from  
511 the Geo model analyses because its pathwise distance and speed of movement obtained from  
512 previous analyses were extreme outliers (Supplementary Figure 5), which can bias the inferences  
513 made from further regression analyses to evaluate the correlates of dispersal abilities.

514  
515 **Pathwise distances and speed of species movement.** We first define species dispersal as the  
516 movement of the species, considering its entire geographic range. We additionally define speed  
517 of species movement as the distance a species moves in an interval of time – kilometres per  
518 million year (see Supplementary Figure 4). In order to obtain the total distance that each species  
519 have historically dispersed through the oceans and rivers – starting from the location of the root  
520 of Clupeoidei (Clupeiformes without *D. clupeoides*) phylogenetic tree - we calculated the  
521 distances dispersed across each phylogenetic branch (branchwise distances) and then we  
522 summed these distances along the path that links the root with extant species (pathwise  
523 distances; Supplementary Figure 4). The branchwise distances were calculated using the  
524 *distCosine* function in the *geosphere* R package<sup>52</sup>. The *distCosine* function brings the shortest  
525 distance between two points, assuming a spherical earth. The distance is calculated according to  
526 the law of the cosines, and the method works at both large and small scales<sup>32</sup>. We calculated the  
527 branchwise distances for every location in the posterior sample, meaning that we have 1,000  
528 distances for every branch in the tree, and therefore, 1,000 pathwise distances for each species  
529 in the tree (Supplementary Table 7). With this approach we have the historical distance dispersed  
530 for each species, considering the uncertainty in ancestral locations estimates (Fig. 3 and  
531 Supplementary Figure 3). In order to have a measure of the speed at which each species in  
532 phylogeny have dispersed over historical time, we calculated the branchwise speed of movement  
533 in km per Myr - dividing the branchwise distances by the branch length of the time-calibrated tree.  
534 We also calculate the speed of movement for all the *posterior* sample of branchwise distances,  
535 and then we calculated the median speed of movement in the path that links the MRCA with  
536 extant species. Finally, we have 1,000 measures of the historical speed of movement for each  
537 species (Supplementary Table 7), which include the uncertainty in ancestral location estimates  
538 (Fig. 3 and Supplementary Figure 3).

539  
540 **Phylogenetic regressions.** To evaluate the expected relationships between SL, WTT, pathwise  
541 distance, pathwise speed of movement, and speciation rates, we performed Phylogenetic  
542 Generalized Least Squares regression models (PGLS) with Bayesian inference which allowed us  
543 to consider the uncertainty in both, parameters estimation and within species data. We consider  
544 the uncertainty within species by using the samples of data for WTT, georeferences, pathwise  
545 distances, and speed of movement. We also considered the uncertainty in ancestral states and  
546 branchwise rates of SL and WSL, and the Speed across phylogenetic branches. Under this  
547 approach, the MCMC samples the regression parameters and the sample data simultaneously,  
548 integrating the uncertainty of both factors in the results. All Bayesian regressions were done in  
549 the computer program *BayesTraits* 3.0.

550  
551 First, we conducted a multiple phylogenetic regression to evaluate the relationship between SL,  
552 WTT and type of migration, including the sample of WTT within species. We compared the BM,  
553 Lambda model (LA), and Ornstein-Uhlenbeck model (OU) for these regressions. We also  
554 evaluated the variation in the SL evolution rate using the variable rates (VR) regression model<sup>37</sup>,  
555 and model that integrate both the VR and LA model (VRLA). The VR regression model enable  
556 the simultaneous estimation of both an overall relationship between SL as a function of WTT and  
557 type of migration, and any shift in the rate that applies to the phylogenetically structured residual  
558 variance in the relationship. The VR regression model identifies heterogeneity in the rate of  
559 evolution along phylogenetic branches (branchwise rates) by dividing the rate into two  
560 parameters: a background rate parameter ( $\sigma^2_b$ ), which assumes that changes in the trait of interest

561 are drawn from an underlying BM process, and a second parameter,  $r$ , which identifies a branch-  
562 specific rate shift. A full set of branchwise rates are estimated by adjusting the lengths of each  
563 branch in a time-calibrated tree (stretching or compressing a branch is equivalent to increasing  
564 or decreasing the phenotypic rate of change relative to the underlying Brownian rate of evolution).  
565 Branchwise rates are defined by a set of branch-specific scalars  $r$  ( $0 < r < \infty$ ) that scale each  
566 branch to optimize the phenotypic rate of change to a BM process ( $\sigma_b^2 \times r$ ). If phenotypic change  
567 occurred at rates faster than the background rate, along a specific branch of the tree, then  $r > 1$   
568 and the branch is stretched. Rates slower than the background rate are detected by  $r < 1$  and the  
569 branch is compressed. If the trait evolves at a constant rate along a branch, then the branch will  
570 not be modified (that is,  $r = 1$ ). There is no limit or prior expectation in the number of the  $r$  branch  
571 scalars,  $r$  numbers vary from zero (no branch is scaled) to  $n$ , in which  $n$  is the number of branches  
572 in the phylogenetic tree. Regarding the values of each  $r$  parameter, we used a gamma prior, with  
573  $\alpha = 1.1$  and a  $\beta$  parameter that is rescaled such that the median of the distribution is equal to  
574  $1^{37,49}$ . With this setting, the numbers of the rate increases and decreases that are proposed are  
575 balanced<sup>49</sup>. We ran four MCMC chains for 151,000,000 iterations, sampling every 50,000  
576 iterations, and discarding 101,000,000 as burn-in. We checked for chain convergence using the  
577 ESS in Tracer v1.6, ensuring of using outputs with ESS > 200.

578  
579 Second, in order to estimate the rates of WTT change through the Clupeiformes phylogeny, we  
580 conducted a Bayesian VRLA regression between WTT and latitude (comparing it with the BM,  
581 LA, OU, and VR regression models; Supplementary Table 2). We included the sample of WTT  
582 and latitude within each species in regression analyses. We ran four MCMC chains for  
583 300,000,000 iterations, sampling every 250,000 iterations, and discarding 150,000,000 as burn-  
584 in. We checked for chain convergence using the ESS in Tracer v1.6, ensuring of using outputs  
585 with ESS > 200.

586  
587 Third, we evaluated the relationship between the pathwise distance with SL and the type of  
588 migration, and between the pathwise speed with SL and type of migration. We included in the  
589 phylogenetic regressions the sample of species data for the pathwise distance and speed of  
590 movement, comparing regressions fitted with the BM, LA, OU, VR, and VRLA model  
591 (Supplementary Table 4 and 5). We ran MCMC chains with different number of iterations,  
592 sampling, and burn-in, in order to ensure of using outputs with ESS > 100. Regressions for  
593 pathwise distance had all ESS > 100 (Supplementary Table 8). Regressions for Speed had ESS  
594 > 100 for the BM and LA model, and ESS < 100 for the OU, VR, and VRLA model (Supplementary  
595 Table 9). However, regressions for pathwise speed based on all models (including those with  
596 ESS < 100) give the same result: SL had a positive effect on pathwise speed.

597  
598 Fourth, we evaluated the relationship between speciation rates with pathwise speed, SL, pathwise  
599 distance, and WTT - including the sample of data for pathwise distances and speed of movement.  
600 We used tip-specific estimates of speciation rates to evaluate the regression between speciation  
601 rates and the multiple explanatory variables. Among the recommended non-model-based tip-rate  
602 metrics to study the correlates of speciation rates (i.e. inverse of equal splits [ES], node density  
603 [ND] and the inverse of terminal branch length [TB])<sup>53</sup> we based our interpretations on the node  
604 density along the phylogenetic paths, divided by the age of the phylogeny (111 Myr after excluding  
605 *D. clupeoides*). Our choice is based on the fact that ND is the least influenced metric by potential  
606 biases and sources of uncertainty associated with branch length estimation from empirical data<sup>54</sup>  
607 – ND capture the average speciation rate over the entire phylogenetic path and weight equally all  
608 branch lengths along the paths. We did not use the tip-rate speciation metric estimated from time-  
609 varying birth-death diversification models owing to the striking uncertainty in the speciation rates  
610 values when they are estimated from phylogenies with extant species only<sup>55</sup>, and due to the

611 erroneous inference of the general diversification patterns when the variation in rates of sequence  
612 evolution are not properly considered in time-tree inference<sup>56</sup>.

613  
614 Additionally, we used PGLS regression models to evaluate regression-coefficients-significance  
615 because PGLS-ND has the highest statistical power when compared with PGLS-ES and PGLS-  
616 TB<sup>53</sup>. Furthermore, PGLS allow us to evaluate the simultaneous effect of multiple explanatory  
617 variables whose effect on speciation rates can be modelled as a linear or non-linear function. This  
618 last point is of utmost importance for our objective because there are expected interactions  
619 between the main explanatory variables (e.g. pathwise speed and SL, WTT and SL) and also  
620 because there are statistical complications associated with estimating interactions without  
621 including quadratic terms (i.e. non-linear functions between the independent and explanatory  
622 variables)<sup>57</sup>. Our full PGLS-ND regression model is described by the following equation:  $ND \sim$   
623  $Speed + SL + Distance + WTT + Speed^2 + SL^2 + Distance^2 + WTT^2 + (Speed * SL) + (Distance *$   
624  $SL) + (WTT * SL)$ . Then, we reached the simpler reduced PGLS-ND regression model based on  
625 strict criteria: we removed the single most non-significant regression-coefficient from the full  
626 regression model, then we reiterated this procedure across every simpler regression until we get  
627 the regression with significant covariates only. We conducted these regression analyses  
628 comparing the BM and LA model. The final regression is in Supplementary Table 6. We ran  
629 51,000,000 iterations, sampling every 50,000 iterations, and discarding the first 10,000,000  
630 iterations as burn in. Regression coefficients were judged to be significant according to a  
631 calculated  $P_{MCMC}$  value for each posterior of regression coefficients. For cases in which <5% of  
632 samples in the posterior distribution crossed zero, this indicates that the coefficient is significantly  
633 different from zero.

634  
635 **Nonphylogenetic regressions.** We applied Bayesian GLS regressions to evaluate the  
636 relationship between the branchwise rates of SL evolution, the branchwise speed of movement  
637 and the branchwise rates of WTT change. We obtained these branchwise rates and speed of  
638 movement using the rate-scaled branches as dividend and the original branch lengths (measured  
639 in time) as the divisor. Specifically, we divided the branches from the LA-scaled posterior sample  
640 of trees for SL, the VRLA-scaled posterior sample of trees for WTT, and the VR-scaled posterior  
641 sample of trees for geographic occurrences. We used 1,000 scaled trees so that we had 1,000  
642 observations of rates per phylogenetic branch. The use of the posterior sample of rate-scaled  
643 branches allows us to include the uncertainty of rates estimation in regression analyses (Fig. 4;  
644 Supplementary Table 9).

645  
646 Additionally, we regressed the *posterior* sample of ancestral SL on the *posterior* sample of  
647 ancestral WTT, inferred at each node of the Clupeiformes phylogeny. Ancestral states were  
648 inferred with the fastAnc function of the phytools R-package<sup>58</sup>, which assumes a constant-rate  
649 Brownian motion model for the evolution of continuous traits. We used the *posterior* sample of  
650 scaled trees, obtained from the model outputs that fit the data better, i.e., LA for SL and the VRLA  
651 model for WTT. The use of rate scaled trees allow us to include the variation in the rate of evolution  
652 when estimating the ancestral states at each phylogenetic node. We used a sample of 1,000  
653 scaled trees, which also allow us to include the uncertainty of ancestral states estimation in  
654 regression analyses (Fig. 2d; Supplementary Table 10). We validated the WTT inferred at nodes  
655 with phylogenetic models by comparing them to model-based temperature reconstructions. We  
656 randomly selected eight nodes (plus the MRCA of Clupeoidei) and matched the median  
657 temperature estimated from the phylogenetic approach with the environmental temperatures  
658 reconstructed from the output of the HadCM3L Earth-System-Model encompassing from the  
659 Jurassic to the Eocene<sup>59</sup>. We matched the age of each node with the respective geologic stage.  
660 We extracted both air (mainland) temperature and sea surface temperature, available at a 3.75 x  
661 2.5° longitude-latitude resolution, based on the 95% of the posterior density of coordinates at each

662 node. It is important to note that the HadCM3L model is based on the Getech model as boundary  
663 conditions, while we used a different model (the PALEOMAP model) to reconstruct the paleo  
664 coordinates of the two fossils. As such, the PALEOMAP reconstruction for the fossils will be  
665 different to what their position in the Getech palaeogeography would be. However, as the  
666 difference between Getech and PALEOMAP from the Cretaceous onwards is small and given  
667 that we used a 3.75x2.5 resolution, the points will probably still fall in the same grid cell under  
668 either reconstruction. On the other hand, the difference will also be negligible as the vast majority  
669 of the *posterior* density of coordinates came from the Geo model.  
670

671 The results show that the phylogenetic estimation of temperatures is positively correlated with  
672 both the estimation of air temperature ( $t = 5.4$ ,  $p = 0.0009$ ) and the estimation of sea temperature  
673 ( $t = 3.17$ ,  $p = 0.01$ ; Supplementary Figure 6) which are based on the HadCM3L Earth-System-  
674 Model. The phylogenetic estimations fall also within the range of air temperature or sea  
675 temperature, depending on whether the phylogenetic node was more likely a freshwater or marine  
676 species. There were three phylogenetic nodes (node 4, node 6, and node 7; Supplementary  
677 Figure 3), with the highest *posterior* distribution for marine habitat, in which the phylogenetic  
678 estimations fall outside the range of sea temperature. However, these three nodes are  
679 reconstructed with high precision around islands (Supplementary Figure 3) which suggest that  
680 those species occupied the inland waters. When considering the range of air temperature, the  
681 phylogenetic estimations of these three nodes fall within the range of the HadCM3L Earth-  
682 System-Model.  
683

684 Finally, we conducted the Bayesian nonphylogenetic GLS regressions in BayesTraits by setting  
685 the Pagel's Lambda parameter to zero, which discards the phylogenetic covariance of the data  
686 values. To sample from the posterior distribution of rates per phylogenetic branch, and from the  
687 posterior distribution of ancestral states at each phylogenetic node, we ran Bayesian regressions  
688 that sample within tips data. We ran 51,000,000 iterations, sampling every 50,000 iterations, and  
689 discarding the first 1,000,000 iterations as burn in. All chains had ESS > 200. Regression  
690 coefficients were judged to be significant according to a calculated  $P_{MCMC}$  value for each posterior  
691 of regression coefficients. For cases in which <5% of samples in the posterior distribution crossed  
692 zero, this indicates that the coefficient is significantly different from zero. We used a uniform prior  
693 for regression coefficients (slopes) as we do not know what the relationship between the response  
694 and predictor variable is. The prior ranged from -100 to 100, to ensure that all possible slope  
695 values are sampled.  
696

#### 697 **Code availability**

698 All analyses in this study were done using BayesTraits version 3 available at  
699 <http://www.evolution.rdg.ac.uk/BayesTraitsV3/BayesTraitsV3.html>  
700

#### 701 **Competing interests**

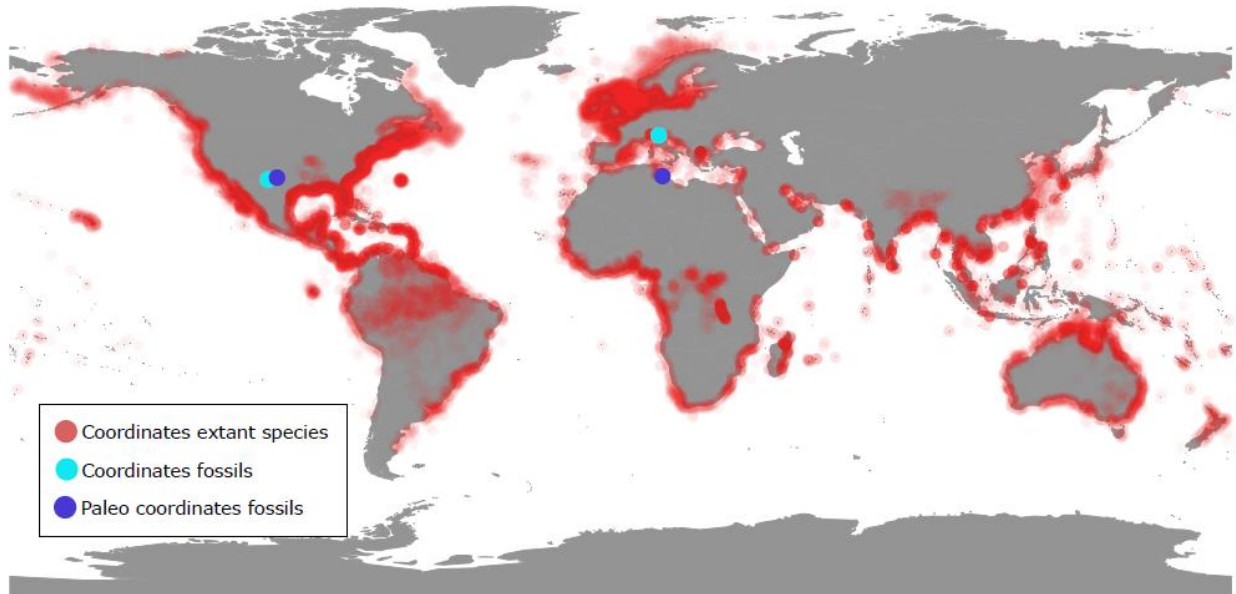
702 The authors declare no competing interests.  
703

#### 704 **Data availability statement**

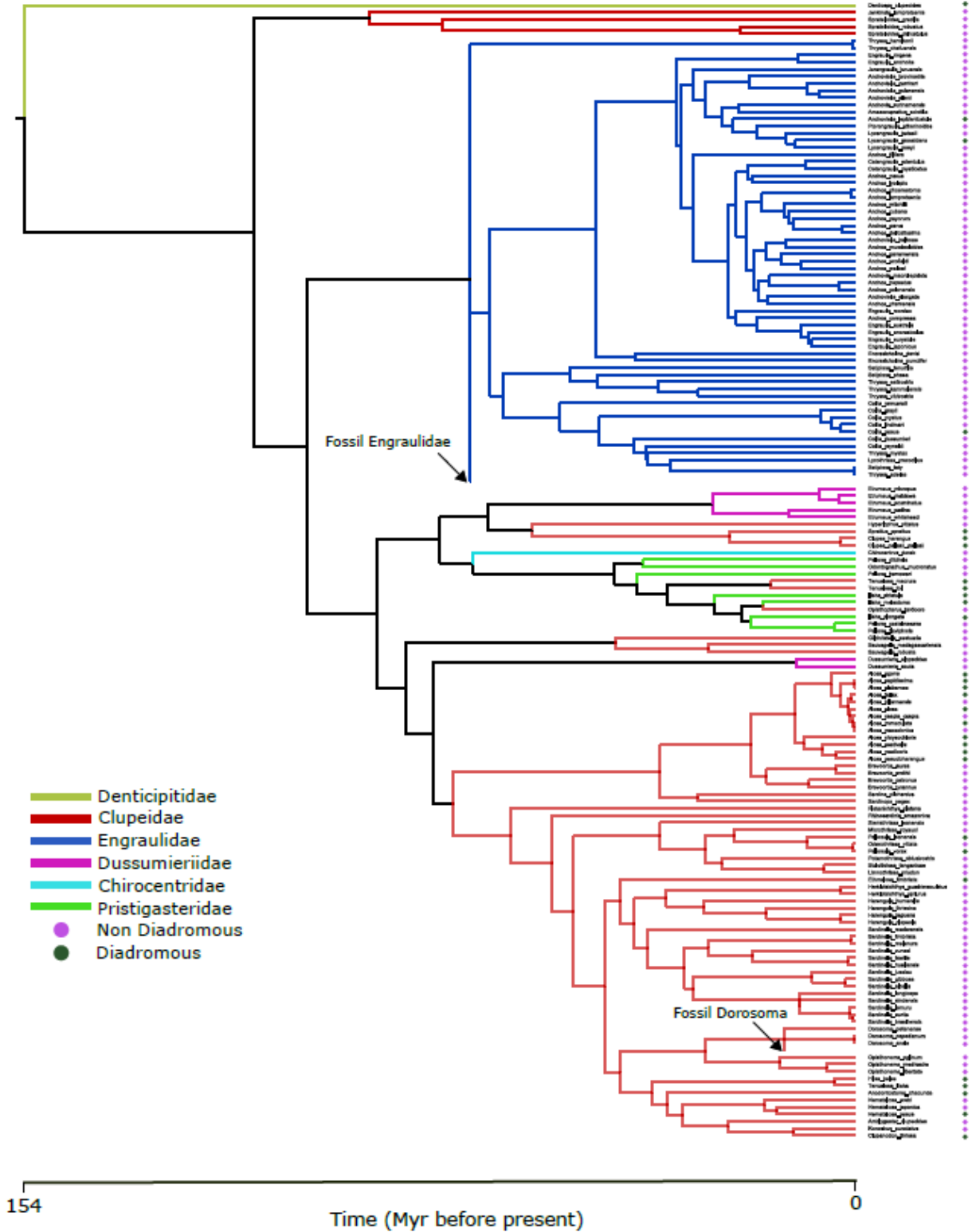
705 No new data were generated for this study. The data used for this paper are available from the  
706 original sources cited in the Methods and Supplementary Information.  
707

708 **Correspondence and request for materials** should be addressed to the corresponding author.  
709

710 **Supplementary Figures**  
711



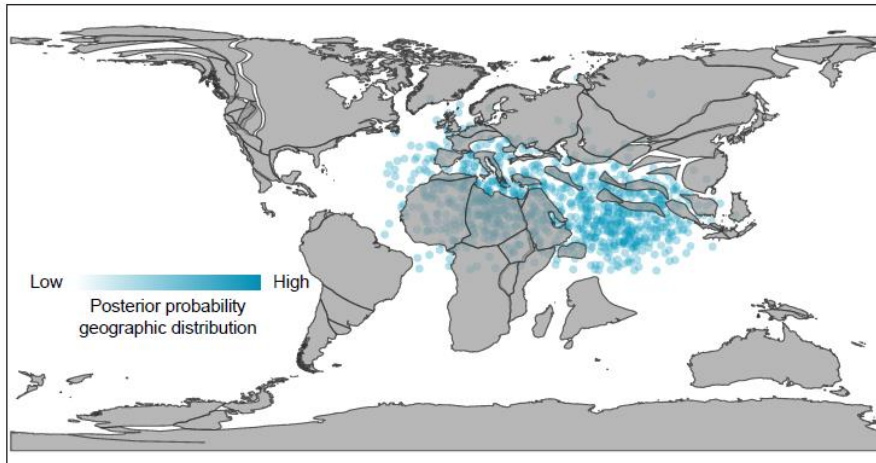
712 **Supplementary Figure 1. Geographic distribution of Clupeiformes species used in this study.** Red  
713 dots represent the geographic occurrences obtained from Aquamaps and the random sample within IUCN  
714 polygons, which comprises 21,795 datapoints for 158 species. The paleo coordinates for the fossils of  
715 *Dorosoma* (America) and *Engraulidae* (Europe) were estimated using the PALEOMAP model in the  
716 chronosphere R package. The coordinates of extant species plus the two paleo coordinates were used as  
717 input data to reconstruct ancestral locations across phylogenetic nodes.  
718  
719



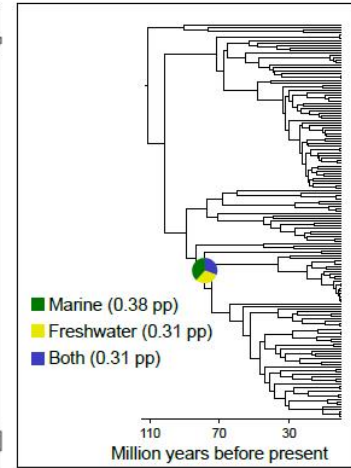
720  
721  
722  
723  
724  
725

**Supplementary Figure 2. Clupeiformes phylogenetic tree used in this study.** The phylogenetic tree was obtained from the Fish Tree of Life and represent the most updated topology and divergence times of the group. Note that branch colours represent the taxonomic arrangement of the group and are used for reference only. Fossils added, and type of migration are indicated. For the Geo model analyses we excluded *Denticiceps clupeoides* (Methods). Nevertheless, we included *D. cupleoides* in all other analyses.

1a

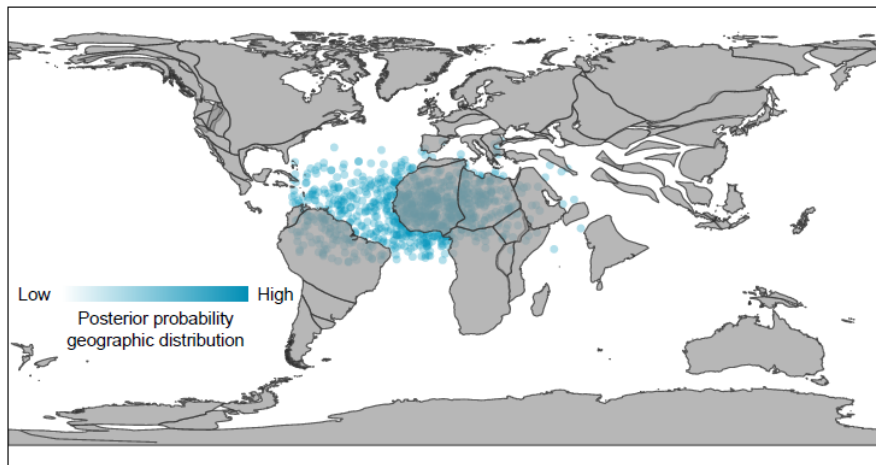


1b

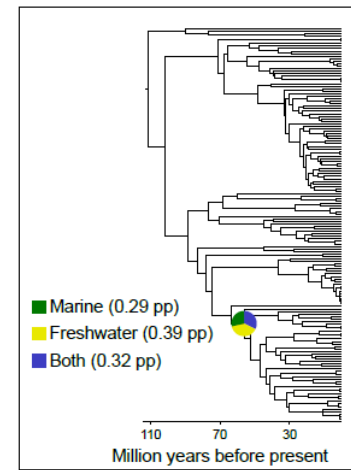


726

2a

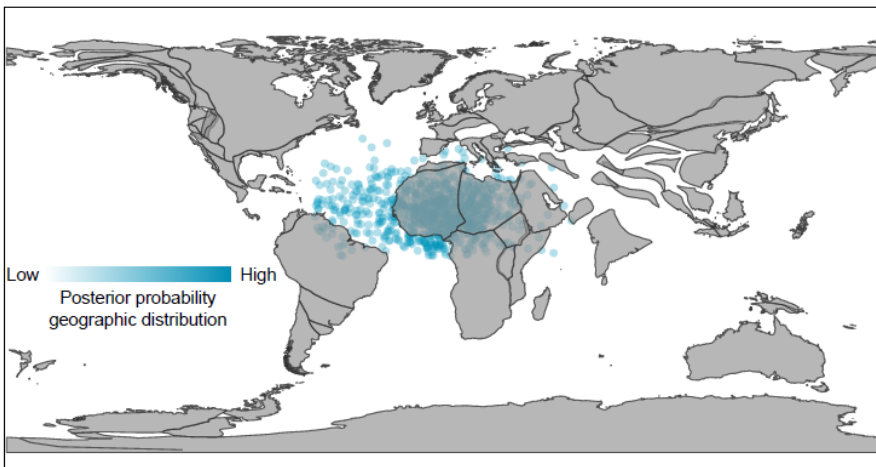


2b

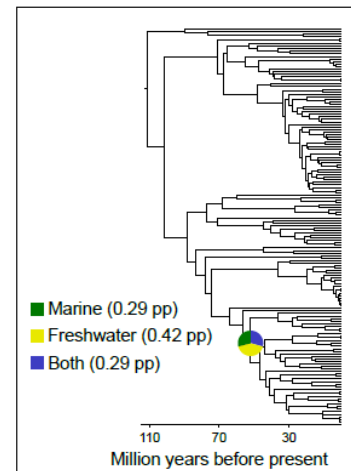


727

3a

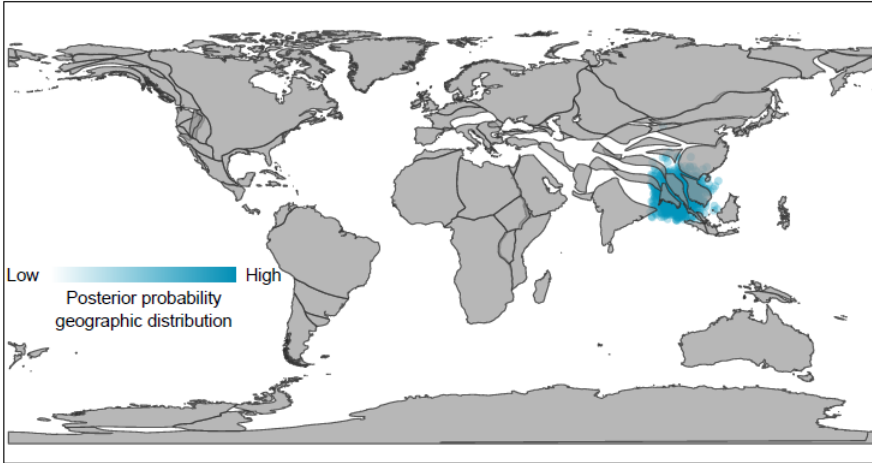


3b

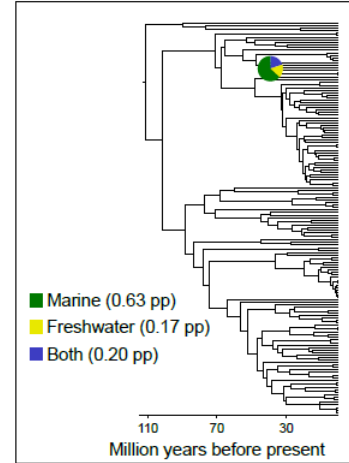


728

4a

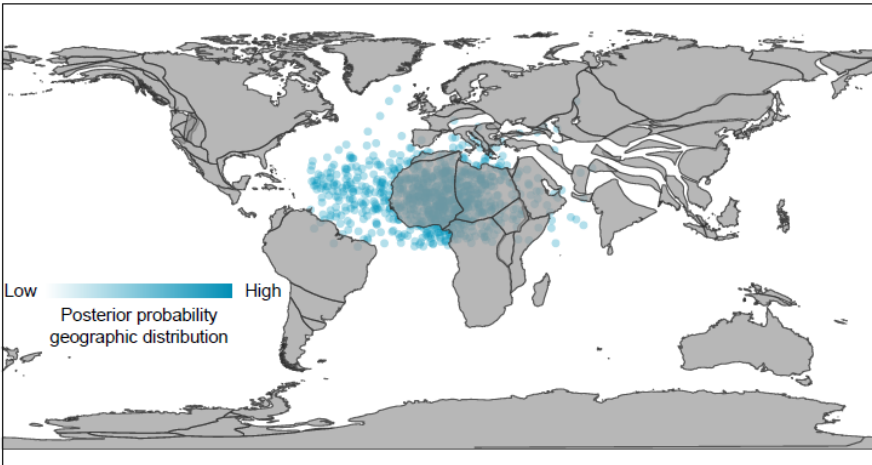


4b

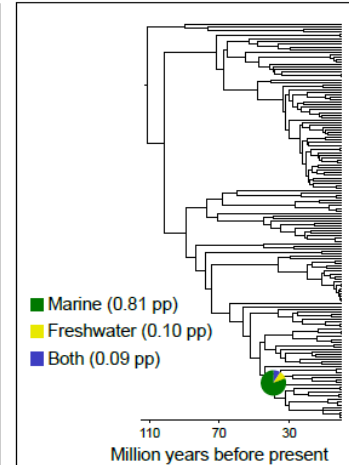


729

5a



5b

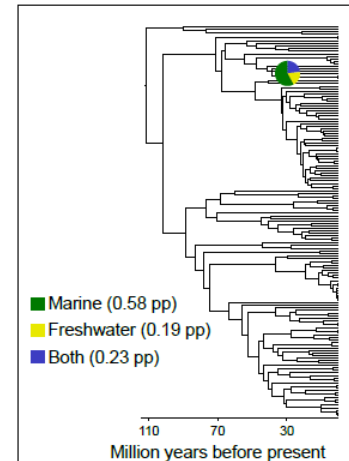


730

6a



6b



731

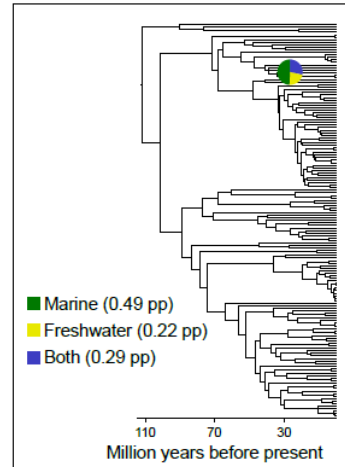
732



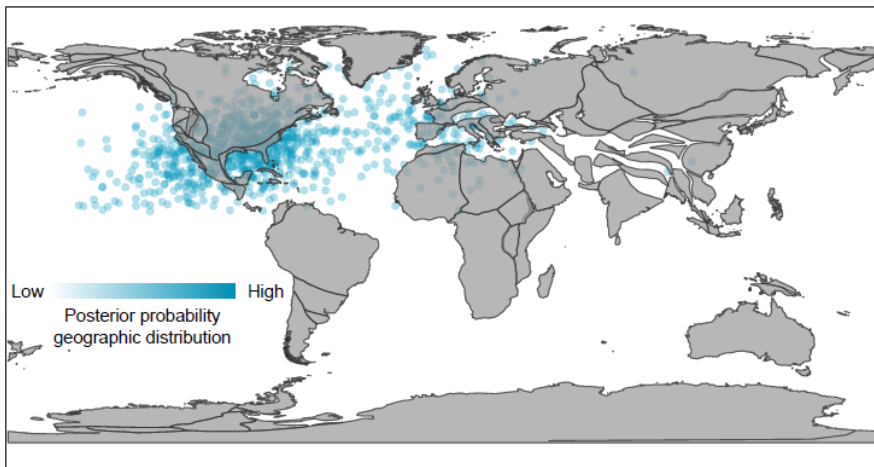
733 7a



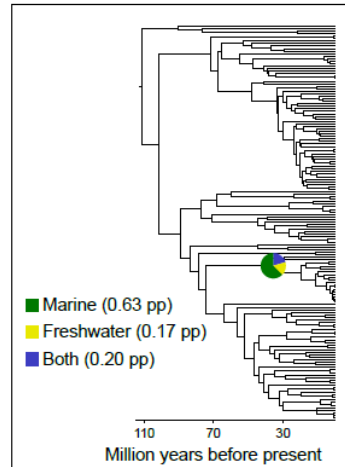
7b



8a



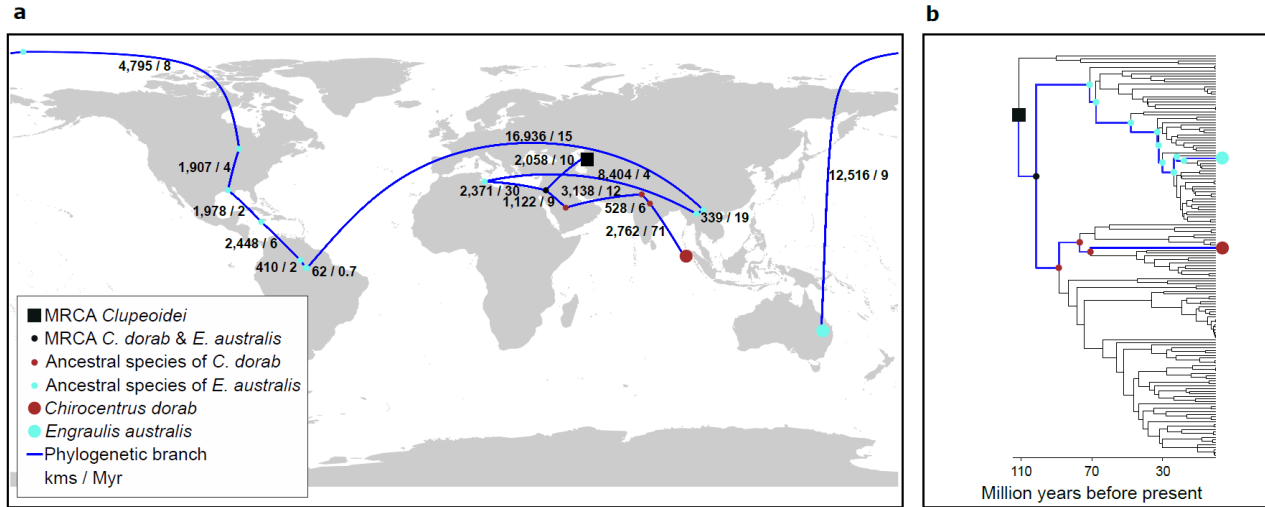
8b



733

734  
 735  
 736  
 737  
 738  
 739  
 740  
 741  
 742  
 743  
 744  
 745  
 746  
 747  
 748  
 749  
 750  
 751  
 752  
 753  
 754

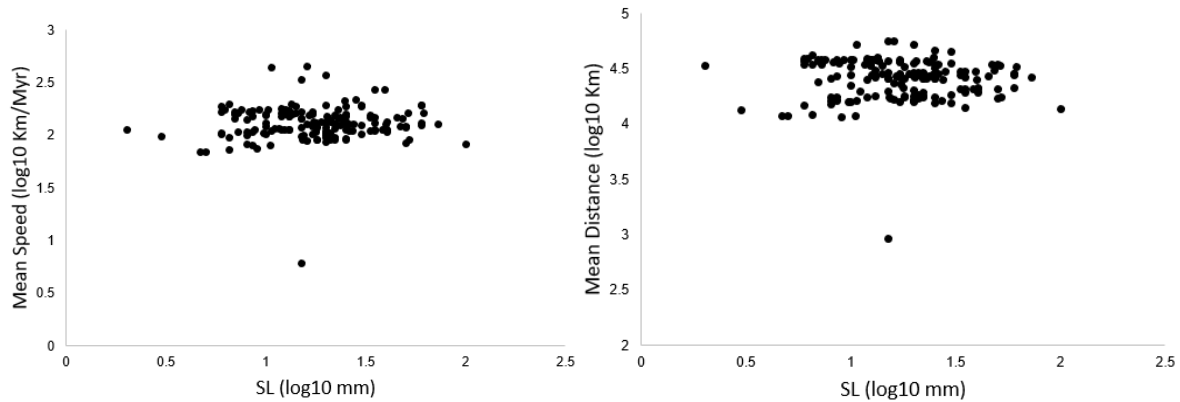
**Supplementary Figure 3. Posterior geographic distribution and posterior probability of habitat type for eight phylogenetic nodes.** We selected eight random nodes ranging from 111 to 33 Mya. 1 = 78 Mya; 2 = 56 Mya; 3 = 52 Mya; 4 = 41 Mya; 5 = 40 Mya; 6 = 38.8 Mya; 7 = 38.4 Mya; 8 = 36 Mya. **a.** The posterior coordinates were estimated with Geo model. **b.** The ancestral habitat type for these eight random nodes was estimated using phylogenetic models for discrete trait evolution.



755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770

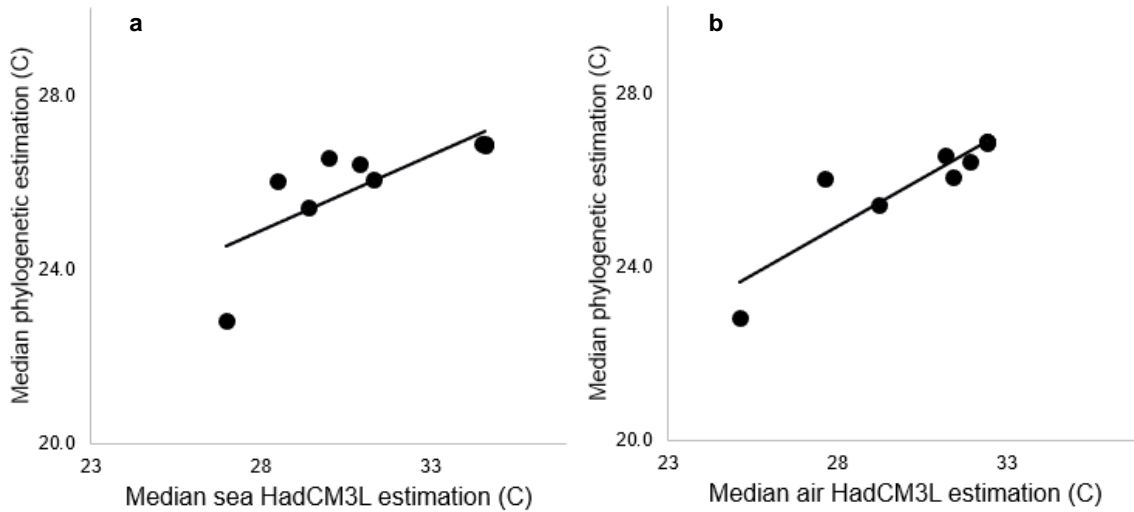
**Supplementary Figure 4. Two continuous geographic routes for the lines of descent leading to *Chirocentrus dorab* and *Engraulis australis*.** The Geo model estimate the *posterior* probability of ancestral species locations (phylogenetic nodes) from geo-referenced occurrences of individuals within extinct and extant species. Ancestral locations are estimated while allowing the speed of species movement to vary across phylogenetic branches. The circles and squares are the geographic centroid estimated from the posterior distribution of coordinates (phylogenetic nodes) and the sample of coordinates from extant species. Note that the geographic centroids are used to obtain an example of the average route travelled for each species. However, we used 1,000 values of total distance and speed (using the full posterior distribution of estimated locations) for each species in all the analyses of this study. Note also that the map represents the actual location of continents – which is included as reference only.

771  
772  
773



774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787

**Supplementary Figure 5. The estimated location for *Denticeps clupeioides* made their speed and distance of movement to be an outlier in regression analyses.** We removed *D. clupeioides* from the Geo analyses because that species descends directly from the MRCA of Clupeiformes and its location is estimated near to the location of the MRCA. This means that species has dispersed a short distance in an exceptionally long time period of 150 million years. This causes the speed and distance of movement for that species to be extremely low and far away from the rest of data when evaluating the correlates of speed and distance. We plot the mean speed and distance for all species.



788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798

**Supplementary Figure 6. Comparison between median temperatures inferred independently from the phylogenetic approach and the HadCM3L Earth-System-Model.** We selected eight random nodes plus the MRCA of Clupeoidei for comparison. Line equation in **a**:  $y = 15 + 0.35(x)$ . Line equation in **b**:  $y = 12 + 0.45(x)$ .

799 **Supplementary Tables**

800

801 **Table 1.** Evolutionary model fitting for the regression that evaluate the effect of type of migration  
 802 and water surface temperature (WTT) on fish standard length (SL). Data analysed includes the  
 803 maximum SL and samples of WTT, within the native range, for each species. The log Marginal  
 804 Likelihood (Marginal Lh) estimated by stepping stone sampling, provides the models support  
 805 given the data and priors. More positive values support a given model, where differences >1  
 806 indicates positive evidence; differences between 2.5 - 5 indicates strong support; and differences  
 807 > 5 indicates very strong support for a model over the other. BM = Brownian Motion, LA = Lambda,  
 808 OU = Ornstein-Uhlenbeck, VR = Variable Rate, VRLA = Variable Rate and Lambda.

809

SL Phylogenetic Regression Model	Marginal Lh. BM	Marginal Lh. LA	Marginal Lh. OU	Marginal Lh. VR	Marginal Lh. VRLA
SL ~ $\alpha + \beta_1(\text{Diadromous}) + \beta_2(\text{WTT})$	-59.11	<b>8.09</b>	-19.84	-16.29	8.13

810

811

812 **Table 2.** Evolutionary model fitting for the regression that evaluates the effect of absolute latitude  
 813 on WTT. Data analysed includes a sample of WTT and absolute latitude (AbsLat) within the native  
 814 range of each species. The log Marginal Likelihood (Marginal Lh) estimated by stepping stone  
 815 sampling, provides the models support given the data and priors. More positive values support a  
 816 given model, where differences >1 indicates positive evidence; differences between 2.5 - 5  
 817 indicates strong support; and differences > 5 indicates very strong support for a model over the  
 818 other. BM = Brownian Motion, LA = Lambda, OU = Ornstein-Uhlenbeck, VR = Variable Rate,  
 819 VRLA = Variable Rate and Lambda.

820

WTT Phylogenetic Regression Model	Marginal Lh. BM	Marginal Lh. LA	Marginal Lh. OU	Marginal Lh. VR	Marginal Lh. VRLA
WTT ~ $\alpha + \beta_1(\text{AbsLat}) + \beta_2(\text{AbsLat})^2$	-421.8	-338.9	-340.1	-318.2	<b>-304.3</b>

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841 **Table 3.** Geographical model (Geo model) fitting for Clupeiformes georeferenced data. The Geo  
 842 model estimate the longitudes and latitudes across the nodes of the phylogenetic tree by means  
 843 of Bayesian inference. These coordinates are estimated onto a three-dimensional cartesian  
 844 coordinates system which were modelled using Brownian motion (BM) – the rate of location  
 845 change across the tree is constant. We also allowed the rate of location-change to vary across  
 846 phylogenetic branches by fitting the Variable Rate model (VR). The log Marginal Likelihood  
 847 (Marginal Lh) estimated by stepping stone sampling, provides the models support given the data  
 848 and priors. More positive values support a given model, where differences >1 indicates positive  
 849 evidence (Bayes Factor > 2); differences between 2.5 - 5 indicates strong support (Bayes Factor  
 850 5 – 10); and differences > 5 indicates very strong support for a model over the other (Bayes Factor  
 851 > 10).  
 852

Chain	Marginal Lh. Geographical model BM	Marginal Lh. Geographical model VR	Bayes Factor BM vs VR
1	-8545.36	-8008.62	1073.48
2	-8546.16	-8008.82	1074.68
3	-8546.74	-8002.93	1087.62
4	-8545.16	-8005.31	1079.70

853  
 854 **Table 4.** Evolutionary model fitting for the regression that evaluate the effect of SL and type of  
 855 migration on the speed of fish movement. The log Marginal Likelihood (Marginal Lh) estimated by  
 856 stepping stone sampling, provides the models support given the data and priors. More positive  
 857 values support a given model, where differences >1 indicates positive evidence (Bayes Factor >  
 858 2); differences between 2.5 - 5 indicates strong support (Bayes Factor 5 – 10); and differences >  
 859 5 indicates very strong support for a model over the other (Bayes Factor > 10). BM = Brownian  
 860 Motion, LA = Lambda, OU = Ornstein-Uhlenbeck, VR = Variable Rate, VRLA = Variable Rate and  
 861 Lambda.  
 862  
 863

	Marginal Lh. BM	Marginal Lh. LA	Marginal Lh. OU	Marginal Lh. VR	Marginal Lh. VRLA
Distance ~ $\alpha + \beta_1(\text{SL})$	137.83	131.89	133,33	<b>145.63</b>	142.58
Distance ~ $\alpha + \beta_1(\text{SL}) + \beta_2(\text{Diadromous})$	128.75	126.14	124.12	140.50	134.74
Distance	121.68	115.44	114.99	128.99	127.81

864  
 865  
 866  
 867  
 868  
 869  
 870  
 871  
 872  
 873  
 874  
 875  
 876

877 **Table 5.** Evolutionary model fitting for the regression that evaluate the effect of SL and type of  
 878 migration on the distance of fish movement. The log Marginal Likelihood (Marginal Lh) estimated  
 879 by stepping stone sampling, provides the models support given the data and priors. More positive  
 880 values support a given model, where differences >1 indicates positive evidence (Bayes Factor >  
 881 2); differences between 2.5 - 5 indicates strong support (Bayes Factor 5 – 10); and differences >  
 882 5 indicates very strong support for a model over the other (Bayes Factor > 10). BM = Brownian  
 883 Motion, LA = Lambda, OU = Ornstein-Uhlenbeck, VR = Variable Rate, VRLA = Variable Rate and  
 884 Lambda.  
 885  
 886

	Marginal Lh. BM	Marginal Lh. LA	Marginal Lh. OU	Marginal Lh. VR	Marginal Lh. VRLA
Speed ~ $\alpha + \beta_1(\text{SL})$	106.41	102.70	99.71	<b>135.16</b>	109.77
Speed ~ $\alpha + \beta_1(\text{SL}) + \beta_2(\text{Diadromous})$	97.90	99.33	89.61	113.02	101.47
Speed	93.24	85.39	89.13	94.13	92.18

887  
 888  
 889 **Table 6.** Phylogenetic regression model for Node Density (ND) obtained after reducing the full  
 890 model  $\text{ND} \sim \text{Speed} + \text{SL} + \text{Distance} + \text{WTT} + \text{Speed}^2 + \text{SL}^2 + \text{Distance}^2 + \text{WTT}^2 + (\text{Speed} * \text{SL})$   
 891  $+ (\text{Distance} * \text{SL}) + (\text{WTT} * \text{SL})$ . The log Marginal Likelihood (Marginal Lh) estimated by stepping  
 892 stone sampling, provides the models support given the data and priors. More positive values  
 893 support a given model, where differences >1 indicates positive evidence (Bayes Factor > 2);  
 894 differences between 2.5 - 5 indicates strong support (Bayes Factor 5 – 10); and differences > 5  
 895 indicates very strong support for a model over the other (Bayes Factor > 10).  
 896  
 897

	Marginal Lh. BM	Marginal Lh. LA
$\text{ND} \sim \alpha + \beta_1(\text{Speed}) + \beta_2(\text{Distance})$	<b>481.54</b>	473.75

898  
 899  
 900  
 901  
 902  
 903  
 904  
 905  
 906  
 907  
 908  
 909  
 910  
 911  
 912  
 913  
 914  
 915  
 916

917 **Table 7.** Data sample size (n) for each species used in this study. SL: Maximum standard length,  
 918 Lon-Lat: Longitude and latitude, WTT: Water temperature tolerance. Note that there is 1 standard  
 919 length for every species as we used one data point – the maximum standard length. The number  
 920 of coordinates and temperature are the same as temperature data was obtained for every  
 921 coordinate.  
 922

Species	SL	Migration type	Lon-Lat	WTT	Pathwise Distance	Pathwise Speed	Node Density
<i>Alosa aestivalis</i>	1	1	219	219	1,000	1,000	1
<i>Alosa agone</i>	1	1	100	100	1,000	1,000	1
<i>Alosa alabamae</i>	1	1	100	100	1,000	1,000	1
<i>Alosa alosa</i>	1	1	243	243	1,000	1,000	1
<i>Alosa caspia caspia</i>	1	1	6	6	1,000	1,000	1
<i>Alosa chrysochloris</i>	1	1	42	42	1,000	1,000	1
<i>Alosa fallax</i>	1	1	466	466	1,000	1,000	1
<i>Alosa immaculata</i>	1	1	11	11	1,000	1,000	1
<i>Alosa killarnensis</i>	1	1	1	1	1,000	1,000	1
<i>Alosa macedonica</i>	1	1	100	100	1,000	1,000	1
<i>Alosa mediocris</i>	1	1	77	77	1,000	1,000	1
<i>Alosa pseudoharengus</i>	1	1	344	344	1,000	1,000	1
<i>Alosa sapidissima</i>	1	1	304	304	1,000	1,000	1
<i>Amazonsprattus scintilla</i>	1	1	100	100	1,000	1,000	1
<i>Amblygaster clupeioides</i>	1	1	26	26	1,000	1,000	1
<i>Anchoa cayorum</i>	1	1	69	69	1,000	1,000	1
<i>Anchoa chamensis</i>	1	1	4	4	1,000	1,000	1
<i>Anchoa choerostoma</i>	1	1	100	100	1,000	1,000	1
<i>Anchoa colonensis</i>	1	1	45	45	1,000	1,000	1
<i>Anchoa compressa</i>	1	1	18	18	1,000	1,000	1
<i>Anchoa cubana</i>	1	1	100	100	1,000	1,000	1
<i>Anchoa delicatissima</i>	1	1	8	8	1,000	1,000	1
<i>Anchoa filifera</i>	1	1	47	47	1,000	1,000	1
<i>Anchoa hepsetus</i>	1	1	299	299	1,000	1,000	1
<i>Anchoa lamprotaenia</i>	1	1	80	80	1,000	1,000	1
<i>Anchoa lyolepis</i>	1	1	186	186	1,000	1,000	1
<i>Anchoa mitchilli</i>	1	1	206	206	1,000	1,000	1
<i>Anchoa mundeoloides</i>	1	1	100	100	1,000	1,000	1
<i>Anchoa nasus</i>	1	1	97	97	1,000	1,000	1
<i>Anchoa panamensis</i>	1	1	5	5	1,000	1,000	1
<i>Anchoa parva</i>	1	1	38	38	1,000	1,000	1
<i>Anchoa scofieldi</i>	1	1	8	8	1,000	1,000	1
<i>Anchoa walkeri</i>	1	1	43	43	1,000	1,000	1
<i>Anchovia macrolepidota</i>	1	1	90	90	1,000	1,000	1
<i>Anchovia surinamensis</i>	1	1	100	100	1,000	1,000	1
<i>Anchoviella alleni</i>	1	1	100	100	1,000	1,000	1

<i>Anchoviella balboae</i>	1	1	100	100	1,000	1,000	1
<i>Anchoviella brevirostris</i>	1	1	15	15	1,000	1,000	1
<i>Anchoviella carrikeri</i>	1	1	100	100	1,000	1,000	1
<i>Anchoviella elongata</i>	1	1	100	100	1,000	1,000	1
<i>Anchoviella guianensis</i>	1	1	100	100	1,000	1,000	1
<i>Anchoviella lepidentostole</i>	1	1	44	44	1,000	1,000	1
<i>Anodontostoma chacunda</i>	1	1	163	163	1,000	1,000	1
<i>Brevoortia aurea</i>	1	1	20	20	1,000	1,000	1
<i>Brevoortia patronus</i>	1	1	66	66	1,000	1,000	1
<i>Brevoortia smithi</i>	1	1	49	49	1,000	1,000	1
<i>Brevoortia tyrannus</i>	1	1	153	153	1,000	1,000	1
<i>Cetengraulis edentulus</i>	1	1	115	115	1,000	1,000	1
<i>Cetengraulis mysticetus</i>	1	1	81	81	1,000	1,000	1
<i>Chirocentrus dorab</i>	1	1	294	294	1,000	1,000	1
<i>Clupanodon thrissa</i>	1	1	13	13	1,000	1,000	1
<i>Clupea harengus</i>	1	1	2,140	2,140	1,000	1,000	1
<i>Clupea pallasii pallasii</i>	1	1	647	647	1,000	1,000	1
<i>Coilia dussumieri</i>	1	1	37	37	1,000	1,000	1
<i>Coilia grayii</i>	1	1	13	13	1,000	1,000	1
<i>Coilia lindmani</i>	1	1	100	100	1,000	1,000	1
<i>Coilia mystus</i>	1	1	100	100	1,000	1,000	1
<i>Coilia nasus</i>	1	1	20	20	1,000	1,000	1
<i>Coilia ramcarati</i>	1	1	100	100	1,000	1,000	1
<i>Coilia reynaldi</i>	1	1	7	7	1,000	1,000	1
<i>Denticeps clupeoides</i>	1	1	100	100	-	-	1
<i>Dorosoma anale</i>	1	1	100	100	1,000	1,000	1
<i>Dorosoma cepedianum</i>	1	1	121	121	1,000	1,000	1
<i>Dorosoma petenense</i>	1	1	96	96	1,000	1,000	1
<i>Dussumieria acuta</i>	1	1	65	65	1,000	1,000	1
<i>Dussumieria elopsoides</i>	1	1	271	271	1,000	1,000	1
<i>Encrasicholina devisi</i>	1	1	60	60	1,000	1,000	1
<i>Encrasicholina punctifer</i>	1	1	120	120	1,000	1,000	1
<i>Engraulis anchoita</i>	1	1	93	93	1,000	1,000	1
<i>Engraulis australis</i>	1	1	251	251	1,000	1,000	1
<i>Engraulis encrasicolus</i>	1	1	863	863	1,000	1,000	1
<i>Engraulis eurystole</i>	1	1	187	187	1,000	1,000	1
<i>Engraulis japonicus</i>	1	1	125	125	1,000	1,000	1
<i>Engraulis mordax</i>	1	1	292	292	1,000	1,000	1
<i>Engraulis ringens</i>	1	1	63	63	1,000	1,000	1
<i>Ethmalosa fimbriata</i>	1	1	105	105	1,000	1,000	1
<i>Etrumeus acuminatus</i>	1	1	100	100	1,000	1,000	1
<i>Etrumeus makiawa</i>	1	1	100	100	1,000	1,000	1
<i>Etrumeus micropus</i>	1	1	100	100	1,000	1,000	1



<i>Etrumeus sadina</i>	1	1	258	258	1,000	1,000	1
<i>Etrumeus whiteheadi</i>	1	1	100	100	1,000	1,000	1
<i>Gilchristella aestuaria</i>	1	1	14	14	1,000	1,000	1
<i>Harengula clupeola</i>	1	1	129	129	1,000	1,000	1
<i>Harengula humeralis</i>	1	1	124	124	1,000	1,000	1
<i>Harengula jaguana</i>	1	1	296	296	1,000	1,000	1
<i>Harengula thrissina</i>	1	1	105	105	1,000	1,000	1
<i>Herklotsichthys quadrimaculatus</i>	1	1	207	207	1,000	1,000	1
<i>Herklotsichthys spilurus</i>	1	1	9	9	1,000	1,000	1
<i>Hilsa kelee</i>	1	1	59	59	1,000	1,000	1
<i>Hyperlophus vittatus</i>	1	1	88	88	1,000	1,000	1
<i>Ilisha elongata</i>	1	1	40	40	1,000	1,000	1
<i>Ilisha melastoma</i>	1	1	51	51	1,000	1,000	1
<i>Ilisha striatula</i>	1	1	4	4	1,000	1,000	1
<i>Jenkinsia lamprotaenia</i>	1	1	118	118	1,000	1,000	1
<i>Jurengraulis juruensis</i>	1	1	100	100	1,000	1,000	1
<i>Konosirus punctatus</i>	1	1	42	42	1,000	1,000	1
<i>Limnothrissa miodon</i>	1	1	100	100	1,000	1,000	1
<i>Lycengraulis batesii</i>	1	1	100	100	1,000	1,000	1
<i>Lycengraulis grossidens</i>	1	1	89	89	1,000	1,000	1
<i>Lycengraulis poeyi</i>	1	1	20	20	1,000	1,000	1
<i>Lycothrissa crocodilus</i>	1	1	100	100	1,000	1,000	1
<i>Microthrissa royauxi</i>	1	1	100	100	1,000	1,000	1
<i>Nematalosa erebi</i>	1	1	100	100	1,000	1,000	1
<i>Nematalosa japonica</i>	1	1	14	14	1,000	1,000	1
<i>Nematalosa nasus</i>	1	1	69	69	1,000	1,000	1
<i>Odaxothrissa vittata</i>	1	1	100	100	1,000	1,000	1
<i>Odontognathus mucronatus</i>	1	1	41	41	1,000	1,000	1
<i>Opisthonema libertate</i>	1	1	125	125	1,000	1,000	1
<i>Opisthonema medirastre</i>	1	1	59	59	1,000	1,000	1
<i>Opisthonema oglinum</i>	1	1	374	374	1,000	1,000	1
<i>Opisthopecterus tardoore</i>	1	1	32	32	1,000	1,000	1
<i>Pellona castelnaeana</i>	1	1	100	100	1,000	1,000	1
<i>Pellona ditchela</i>	1	1	290	290	1,000	1,000	1
<i>Pellona flavipinnis</i>	1	1	100	100	1,000	1,000	1
<i>Pellona harroweri</i>	1	1	66	66	1,000	1,000	1
<i>Pellonula leonensis</i>	1	1	38	38	1,000	1,000	1
<i>Pellonula vorax</i>	1	1	100	100	1,000	1,000	1
<i>Platanichthys platana</i>	1	1	100	100	1,000	1,000	1
<i>Potamothrissa obtusirostris</i>	1	1	100	100	1,000	1,000	1
<i>Pterengraulis atherinoides</i>	1	1	100	100	1,000	1,000	1
<i>Rhinosardinia amazonica</i>	1	1	100	100	1,000	1,000	1
<i>Sardina pilchardus</i>	1	1	714	714	1,000	1,000	1

<i>Sardinella albella</i>	1	1	190	190	1,000	1,000	1
<i>Sardinella aurita</i>	1	1	674	674	1,000	1,000	1
<i>Sardinella brasiliensis</i>	1	1	75	75	1,000	1,000	1
<i>Sardinella fimbriata</i>	1	1	56	56	1,000	1,000	1
<i>Sardinella gibbosa</i>	1	1	226	226	1,000	1,000	1
<i>Sardinella hualiensis</i>	1	1	11	11	1,000	1,000	1
<i>Sardinella jussieu</i>	1	1	7	7	1,000	1,000	1
<i>Sardinella lemuru</i>	1	1	61	61	1,000	1,000	1
<i>Sardinella longiceps</i>	1	1	38	38	1,000	1,000	1
<i>Sardinella maderensis</i>	1	1	168	168	1,000	1,000	1
<i>Sardinella melanura</i>	1	1	76	76	1,000	1,000	1
<i>Sardinella sindensis</i>	1	1	24	24	1,000	1,000	1
<i>Sardinella tawilis</i>	1	1	100	100	1,000	1,000	1
<i>Sardinella zunasi</i>	1	1	33	33	1,000	1,000	1
<i>Sardinops sagax</i>	1	1	745	745	1,000	1,000	1
<i>Sauvagella madagascariensis</i>	1	1	100	100	1,000	1,000	1
<i>Sauvagella robusta</i>	1	1	100	100	1,000	1,000	1
<i>Setipinna phasa</i>	1	1	100	100	1,000	1,000	1
<i>Setipinna taty</i>	1	1	47	47	1,000	1,000	1
<i>Setipinna tenuifilis</i>	1	1	86	86	1,000	1,000	1
<i>Sierrathrissa leonensis</i>	1	1	100	100	1,000	1,000	1
<i>Spratelloides delicatulus</i>	1	1	311	311	1,000	1,000	1
<i>Spratelloides gracilis</i>	1	1	169	169	1,000	1,000	1
<i>Spratelloides robustus</i>	1	1	69	69	1,000	1,000	1
<i>Sprattus sprattus</i>	1	1	825	825	1,000	1,000	1
<i>Stolothrissa tanganicae</i>	1	1	100	100	1,000	1,000	1
<i>Tenuالosa ilisha</i>	1	1	47	47	1,000	1,000	1
<i>Tenuالosa macrura</i>	1	1	7	7	1,000	1,000	1
<i>Tenuالosa toli</i>	1	1	19	19	1,000	1,000	1
<i>Thryssa adelae</i>	1	1	3	3	1,000	1,000	1
<i>Thryssa chefuensis</i>	1	1	100	100	1,000	1,000	1
<i>Thryssa hamiltonii</i>	1	1	166	166	1,000	1,000	1
<i>Thryssa kammalensis</i>	1	1	8	8	1,000	1,000	1
<i>Thryssa mystax</i>	1	1	33	33	1,000	1,000	1
<i>Thryssa setirostris</i>	1	1	211	211	1,000	1,000	1
<i>Thryssa vitrirostris</i>	1	1	65	65	1,000	1,000	1

923  
924  
925  
926  
927  
928  
929  
930

931  
 932  
 933  
 934  
 935  
 936  
 937

**Table 8.** Chain settings for regression analysis of pathwise distance. Number of iterations / burn-in / sampling frequency / and effective sample size (ESS). BM = Brownian Motion, LA = Lambda, OU = Ornstein-Uhlenbeck, VR = Variable Rate, VRLA = Variable Rate and Lambda. SL = Maximum standard length.

	BM	LA	OU	VR	VRLA
Distance ~ $\alpha + \beta_1(\text{Diadromous}) + \beta_2(\text{SL})$	150E6/50E6/ 10E4/176	150E6/50E6/ 10E4/424	300E6/200E6/ 10E4/414	600E6/500E6/ 10E4/150	150E6/50E6/ 10E4/229
Distance ~ $\alpha + \beta_1(\text{SL})$	200E6/100E6/ 10E4/332	51E6/1E6/ 5E4/537	51E6/1E6/ 5E4/271	600E6/500E6/ 10E4/129	200E6/100E6/ 10E4/296
Distance	51E6/1E6/ 5E4/262	51E6/1E6/ 5E4/605	100E6/50E6/ 5E4/276	300E6/200E6/ 10E4/153	51E6/1E6/5 E4/212

938  
 939  
 940  
 941  
 942  
 943

**Table 9.** Chain settings for regression analysis of pathwise speed. Number of iterations / burn-in / sampling frequency / and effective sample size (ESS). BM = Brownian Motion, LA = Lambda, OU = Ornstein-Uhlenbeck, VR = Variable Rate, VRLA = Variable Rate and Lambda. SL = Maximum standard length.

	BM	LA	OU	VR	VRLA
Speed ~ $\alpha + \beta_1(\text{diadromous}) + \beta_2(\text{SL})$	100E6/50E6/ 5E4/126	110E6/10E6/ 10E4/382	-	-	-
Speed ~ $\alpha + \beta_1(\text{SL})$	150E6/50E6/ 4E4/133	100E6/50E6/ 5E4/194	-	-	-
Speed	100E6/50E6/ 5E4/168	51E6/1E6/ 5E4/461	51E6/1E6/ 5E4/198	-	300E6/250E6/ 5E4/171

944  
 945  
 946  
 947  
 948

**Table 10.** Data sample size (n) of the estimated ancestral states, and the estimated branchwise rates. SL: Maximum standard length, WTT: Water temperature tolerance.

Variable	n
Ancestral SL	157,000
Ancestral WTT	157,000
Branchwise Rate SL	314,000
Branchwise Rate WTT	314,000
Branchwise Speed	312,000

949