

1 Salinity Transitions in Ray-finned Fishes: The Roles of

2 Diadromy and Euryhalinity

3 William P. O'Connor Christopher P. Kenaley*

4 Department of Biology, Boston College, Chestnut Hill, MA 02465, USA

5 *Author for correspondence (kenaley@bc.edu)

6

7 *Keywords:* Halotolerance, Actinopterygii, ancestral states, comparative methods, machine
8 learning

Abstract

Ray-finned fishes inhabit a broad range of aquatic habitats and the evolutionary history of this lineage is marked by repeated transitions between marine and freshwater systems. Euryhalinity and diadromy have often been invoked as key mediators of marine-freshwater transitions. To date, no macroevolutionary study of stenohaline transitions that incorporates the roles of euryhalinity and diadromy has been undertaken. We sought to assess whether euryhalinity and diadromy were key innovations required for stenohaline transitions by evaluating the historical patterns of marine-freshwater transitions and undertaking state-dependent diversification analysis. Based on stochastic character mapping, we outlined the transition patterns of halohabitat-diadromy states. We found a complicated pattern of marine-freshwater transitions in which (1) transitions from marine to fresh waters were mediated both by diadromy and, more commonly, non-diadromous euryhalinity and (2) transitions from fresh to marine waters were driven solely by euryhalinity. We also found that net speciation rates varied significantly between each state, with euryhaline and diadromous lineages having the lowest and highest values, respectively. Based on these results, we assert that euryhalinity was the preeminent innovation that permitted ray-finned fishes to access new biomes and that competition in new salinity niches likely plays an important role in promoting or limiting transitions.

Introduction

27 Ray-finned fishes (Actinopterygii), the most diverse vertebrate lineage, have invaded nearly every
28 aquatic habitat, from deep and shallow marine systems, to lacustrine and riverine systems and
29 estuaries alike. The evolutionary history of this group is marked by repeated transitions between
30 marine and freshwater habitats (Bloom and Lovejoy, 2012; de Brito et al., 2022; Carrete Vega
31 and Wiens, 2012; Corush, 2019; Davis and Betancur-R, 2017; Rabosky, 2020). Each one of these
32 transitions represents an extreme shift in terms a new ecological niche (de Brito et al., 2022;
33 Carrete Vega and Wiens, 2012; Lee and Bell, 1999; Seehausen and Wagner, 2014) and drastically
34 different physiological requirements (Evans and Claiborne, 2008; Kültz, 2015).

35 Ray-finned fishes maintain the osmolality of their extracellular body fluids at a relatively
36 constant level of at approximately 300 mosmol kg⁻¹ and largely independent of environmental
37 salinity. Therefore, fishes inhabiting stenohaline environments must maintain an osmotic balance
38 in the face of a stark, but consistent, osmotic pressure to either achieve a hyposmotic gradient
39 in the case of stenohaline freshwater fishes or hyperosmotic gradient in the case of stenohaline
40 marine fishes. Euryhaline fishes must cope with substantial changes in environmental osmolal-
41 ity, often over short temporal time scales. They may do so either as diadromous species that
42 migrate between stenohaline systems or non-diadromous species that inhabit estuaries where
43 tidal dynamics and freshwater systems influence salinity regimes. The physiological shifts that
44 mediate stenohaline transitions—either marine to fresh or the reverse—have been studied ex-
45 tensively (Edwards and Marshall, 2012; Kültz, 2015; McCormick, 2001; McCormick et al., 2013;
46 McCormick and Saunders, 1987). These include complicated functional changes in tissues of the
47 gill, kidney, skin and intestine that span several scales from the epigenetic to tissue level. Given
48 the osmoregulatory challenges of stenohaline species and the drastic physiological changes that
49 must accompany transitions between stenohaline environments, euryhalinity must be seen as an
50 important capability mediating these transitions.

51 In addition to physiological shifts, ecological and macroevolutionary dynamics are key drivers

52 of major transitions in habitat. As a theory, ecological opportunity has often been invoked to ex-
53 plain macroevolutionary transitions between major biogeographical and physical regimes across
54 the tree of life (Simpson, 1953, 1984; Stroud and Losos, 2016; Yoder et al., 2010), and in partic-
55 ular, marine-freshwater transitions in fishes (Betancur-R et al., 2012; Davis et al., 2012; Santini
56 et al., 2013). This theory states that a new resource niche can be exploited by ecologically capable
57 lineages where species richness of ecologically similar taxa or competition is low. Ecological op-
58 portunity can be mediated by a combination of biogeographical processes, including dispersal,
59 ecological processes, including extinction of species of similar ecology, and the evolution of key
60 innovations that permit the use the novel resources.

61 In the context of marine-freshwater transitions, the evolution of the ability to withstand a
62 wide range of salinity, i.e., euryhalinity, may be seen as one such innovation that permits these
63 transitions by expanding the potential niche and ecological range of a lineage. In addition, di-
64 adromy, the cyclic migration between stenohaline systems, may be seen as a complementary
65 behavioral adaption that, when coupled with euryhalinity mediates marine-freshwater transi-
66 tions, and the reverse, by expanding and partitioning resource use across the life history of a
67 lineage (McDowall, 1997, 2001).

68 Gross (1987) was the first to codify the important relationship between euryhalinity and di-
69 adromy in shaping transitions between marine and freshwaters (McDowall, 1997), suggesting
70 that euryhalinity is an important preliminary state and diadromy was a key final state before
71 permanent invasion of either fresh or marine waters from marine or freshwaters. McDowall
72 (1997) remarked on the intuitive nature of this framework and posited that the predominant
73 forms of diadromy—anadromy and catadromy—were logical outcomes of such a process. For
74 example, in the case of marine-to-fresh transitions, a lineage of marine origin explores freshwa-
75 ters in adulthood to spawn after the evolution of euryhalinity, and transitions into a diadromous
76 species after this process reaches a steady sate. If this new environment offers some new advan-
77 tage that permits increased reproductive success, then this marine-to-fresh transition becomes
78 complete. This potential process is, of course, similar for freshwater to marine transitions, but

79 involves different starting and final salinity habitats.

80 The modest body of research devoted to the broader macroevolutionary dynamics of salinity-
81 habitat transitions represents several interesting trends. Betancur-R et al. (2015) found that transi-
82 tion dynamics vary asymmetrically between habitats with marine lineages colonizing freshwaters
83 rivers more frequently than freshwater lineage invading marine habitats. This important and illumi-
84 nating study focused on transitions between stenohaline habitats (fresh-to-marine and marine-to-
85 fresh) exclusively, however, leaving open the question of what role euryhalinity or diadromous
86 behavior play in salinity transitions. In addition, Betancur-R et al. (2015) limited their phyloge-
87 netic framework to a super tree of 1822 extant and extinct taxa. In a more expansive analysis of
88 over 13000 species, Rabosky (2020) also found that marine-to-freshwater transitions were much
89 more common than the reverse. This analysis implemented a stepwise gain and loss model
90 whereby all transitions between marine and freshwater environments moved through a brackish
91 intermediate stage. Like Betancur-R et al. (2015), Rabosky (2020) did not explicitly address the
92 role euryhalinity or diadromy played in evolutionary transitions between stenohaline environ-
93 ments. Corush (2019) found that transitions to diadromy were rare but that transitions out of
94 diadromy into both fresh and marine habitats were common. This study focused on the role of di-
95 adromy in shaping stenohaline transitions, but did not evaluate the influence of non-diadromous
96 euryhalinity in mediating transitions between marine and freshwater states. Taken as a whole,
97 these studies reveal the need to integrate a more comprehensive assessment of salinity tolerance
98 and migration behavior in macroevolutionary studies of shifts in major salinity habitat.

99 The primary goal of our study was to empirically assess the intermediate state hypothesis,
100 specifically that euryhalinity and diadromy, either alone or together, are key innovations required
101 for stenohaline transitions. Specifically, we sought to evaluate the historical patterns of marine-
102 freshwater transitions by exploring two models: (1) these transitions require the evolution of a
103 euryhaline state *and*, the behavioral strategy of diadromy, or (2) they require only euryhalinity.
104 Secondarily, we also sought to evaluate the role euryhalinity and diadromy play in marine-
105 freshwater transitions as transitional states or adaptive life histories. To address these aims,

106 we undertook several comparative phylogenetic analyses, including historical reconstructions of
107 halotolerance and diadromy and state-dependent diversification analysis. Through our historical
108 reconstructions, we sought to uncover the ancestral patterns of halohabitat-diadromy transitions
109 between them. We hypothesized that if both euryhalinity and diadromy are required as ordered
110 transitions to make marine-to-fresh or fresh-to-marine transitions, we would uncover a pattern
111 in which transitions are similarly common from marine or freshwater to euryhalinity, from eury-
112 halinity to diadromy, and diadromy to marine or freshwater. Alternatively, we hypothesized that
113 if marine-freshwater transitions are mediated by euryhalinity alone, we would uncover a pattern
114 in which marine or fresh water transitions to euryhalinity and transitions from euryhalinity to
115 fresh or marine waters were similar but much more common than transitions into and out of di-
116 adromy. Through our diversification analysis, we sought to uncover whether or not euryhalinity
117 and diadromy represent transient conditions, stepping stones on the path to marine-freshwater
118 transitions or adaptive strategies. We assert that relatively low and hight net diversification rates
119 represent a transient and adaptive strategies, respectively.

120 **Materials and Methods**

121 *Salinity Habitat Classification*

122 To classify salinity habitat (i.e., halohabitat) requires a clear understanding of what this pheno-
123 type encompasses and precise definitions of each state. Here we treat halohabitat as ecological
124 halotolerance, the realized niched of a species. This contrasts physiological halotolerance, or the
125 potential niche, defined as the physiological range of a species with regards to its osmotic toler-
126 ance (Schultz and McCormick, 2012). We classified halohabitat of each species in our study as
127 one of three states, either stenohaline fresh, stenohaline marine, or euryhaline. The definition of
128 stenohalinity is rather simple and generally agreed upon: stenohaline fresh species only occur
129 in fresh water of very low salinity close to 0 ppt and stenohaline marine species occur in marine
130 waters of high salinity, greater than 30 ppt. However, the definition of euryhalinity is more prob-

131 lematic. Here we follow Schultz and McCormick (2012), Kültz (2015), and Lee and Bell (1999)
132 in defining euryhalinity as the ability to live in both stenohaline habitats or to live in brackish
133 habitats of considerably varying salinity over time and space. The ability to traverse both steno-
134 haline environments is generally indicative of diadromy while living in brackish environment
135 generally indicates estuarine habitat utilization. These two life histories may require different or
136 similar physiological mechanisms (Kültz, 2015); however, underlying both is the ability to main-
137 tain osmotic balance across a wide range of salinity. In our comparative analyses, we make the
138 distinction between euryhaline-nondiadromous and euryhaline diadromous states (see below).

139 We attempted to assign a halohabitat state to each species of the Actinopterygii included in the
140 phylogenetic hypothesis published by Rabosky et al. (2018). To do this, we applied a multistep
141 framework that included cross-referencing halohabitat categories reported by FishBase (Froese
142 and Pauly, 2023) and vetting these designations a series of machine learning (ML) operations to
143 predict halohabitat.

144 We began our assignment process by assembling salinity habitat data reported by FishBase
145 (Froese and Pauly, 2023) using the `rfishbase` package (Boettiger et al., 2012) written for R com-
146 puting environment (R Core Team, 2022). FishBase classifies species as marine, freshwater, or
147 brackish. We initially applied the euryhaline state to all brackish species. After compiling this in-
148 formation, we performed a preliminary qualitative assessment of assigned categories and noted
149 several errors, e.g., species known to be stenohaline but categorized by FishBase as brackish and
150 euryhaline. To address these issues and evaluate all of the salinity habitat designations for our
151 study species in FishBase, we undertook an ML-based classification analysis. For this, we be-
152 gan by assembling a list of marine and freshwater species from the World Register of Marine
153 Species (WoRMS) (Costello et al., 2013) using the R package `worms` (Holstein, 2018) and a list of
154 freshwater fish species downloaded directly from the Freshwater Biodiversity Data Portal (FBDP)
155 (Schmidt-Kloiber et al., 2019). This process resulted in three variables: FBDP freshwater, WoRMS
156 freshwater, and WoRMS marine. Each was coded as "0" or "1", with "1" representing presenting
157 freshwater, freshwater, and marine, respectively.

158 Next, we assembled a maximum of 1000 georeferenced museum records for each species by
159 querying the Global Biodiversity Information Facility (GBIF) using the R package `rgbif` (Cham-
160 berlain and Boettiger, 2017), totaling 1,967,067 records or 178 ± 262 records per species. Each
161 individual record was evaluated for common spatial errors in biological collections using the R
162 package `CoordinateCleaner`. For this we flagged and removed any coordinates that represented
163 country capitals and geographic centroids, those corresponding to the locality of collection in-
164 stitutions or GBIF headquarters, and any outliers. Outliers were defined as those records that
165 were outside five times the interquartile range of minimum distance to the next neighbor of the
166 species. In addition, we removed any records that were potentially acquired at a market by as-
167 sessing whether they occurred within the boundaries dense urbanization according to (Patterson
168 and Kelso, 2012). For this, we used the R package `sf` to evaluate the distance of each record to
169 the closest urban area and removed any for which the distance was 0 km. After this screening
170 process, our georeferenced dataset included 1,370,996 records or 124 ± 191 records per species

171 We then evaluated the occurrence of each record as residing in marine, freshwater, terrestrial,
172 or estuarine localities. For this we assembled polygon shape files from Freshwater Ecosystems
173 of the World (FEOW) (Abell et al., 2008) downloaded at <https://www.feow.org/download>, Ma-
174 rine Ecosystems of the World (MEOW) from The Nature Conservancy (Spalding et al., 2007)
175 downloaded at <https://geospatial.tnc.org/maps>, country shape files using the R package
176 `rnaturrearth` (Massicotte and South, 2023), and shape files for the global distribution of es-
177 tuaries (Alder, 2003) downloaded at <https://data.unep-wcmc.org/datasets/23>. We also re-
178 tried mean yearly (1981–2010) sea-surface salinity of each record using the Levitus dataset
179 (Garcia et al., 2013) using the R package `ocedata` (Kelley and Richards, 2022). From these five
180 variables—marine, freshwater, terrestrial, and estuarine occurrence and salinity, we constructed
181 five species-level variables to include in our ML classification dataset: the proportion of records
182 classified as marine, freshwater or estuarine, mean salinity (in ppt), and median distance to the
183 coastline (in km). A record was considered marine only if the record did not occur over land its
184 salinity was greater than 30 ppt. A record was considered freshwater only if it was more than 5

185 km inland, was not within a MEOW nor estuary, was within a FEOW, and it had no value for
186 salinity. A record was consider estuarine only if it occurred within an estuary boundary. For
187 each species, the total number of marine, freshwater, and estuarine records was divided by the
188 total number of records and arcsine transformed.

189 In addition to FishBase halohabitat designation, Our ML classification dataset therefore in-
190 cluded 9 predictive variables: proportion of marine, freshwater and estuarine records, median dis-
191 tance to the coast, log mean salinity, freshwater status occurring to FBDP, and marine and fresh-
192 water status according to WoRMS. We undertook supervised ML analysis to make revised halo-
193 habitat designations using two methods, a Breiman's random-forest algorithm (RF) (Breiman,
194 2001) and a gradient boosted classification machine (GBM) (Friedman, 2001). In each case, Fish-
195 Base halohabitat designation was the response variable. For our random forest analysis, we used
196 the `randomForest` function from the `randomForest` package (R Core Team, 2022) written for R.
197 We produced a forest of 10,000 trees using the default training size of two-thirds of the data to
198 train each tree and the number of variables randomly sampled as candidates at each split set to 2
199 ("mtry=2"). We performed our GBM analysis using the `H2o` package written for R (Fryda et al.,
200 2023). For this, we used the `h2o.gbm` function to produce 10,000 trees using 20% of the data as a
201 training set. We set the learning rate to 0.1, row sample rate to 1, maximum tree depth to 5, and
202 column sample rate per tree to 0.8. We used the `h2o.predict` function to then make GBM-based
203 predictions of halohabitat on the full dataset.

204 To make final halohabitat predictions we compared the FishBase designations with the RF
205 and GBM predictions for each species. For any species for which the three designations were
206 unanimous, final halohabitat was designated as that unanimous state. For each species for which
207 there was disagreement between the original FishBase designation and either the RF or GBM
208 predictions, we performed a literature search to evaluate the final halohabitat state. For this
209 process we took a conservative approach in making the final determination in regards to the
210 euryhaline state. If any of the FishBase, RF, or GBM designations were euryhaline, but we could
211 not establish through our literature search that the species spanned a broad range of salinities,

212 we deferred to the stenohaline state designated by the other sources. Final predictions along with
213 all ML dataset values are contained in Supplementary Table 1.

214 *Diadromy Classification*

215 To classify each of our study species as diadromous or non diadromous, we continued with an
216 approach similar to our halohabitat classification. We began with compiling a list of diadromous
217 species according to Corush (2019) and, if their final halohabitat designation was euryhaline, we
218 preliminarily classified these species at "euryhaline-diadromous". All others were classified as
219 euryhaline, fresh, or marine according to our final halohabitat designations. We then added a
220 single species-level variable to our ML data set, the span of the distance from the coastline. Here
221 we assumed that those with higher values would make more expansive migrations between
222 marine and freshwaters or the reverse and would therefore be indicative of diadromy. We then
223 repeated our RF and GBM machine learning analyses with preliminary diadromy classification as
224 the response variable using the same parameters outlined above. For any species for which there
225 was either disagreement between the Corush (2019) list of diadromous species and RF or GBM
226 predictions or between the original diadromy designation and either the RF or GBM predictions,
227 we performed a subsequent literature search to make a final designation. Final diadromous
228 predictions are contained in Supplementary Table 1.

229 *Ancestral State Estimation and Transitions*

230 We estimated the ancestral halohabitat-diadromy states across the ray-finned fish tree of life us-
231 ing stochastic character mapping (Bollback, 2006) as implemented in the R package phytools.
232 We pruned the phylogeny to include the 9738 species for which were able to assesses salin-
233 ity preference and diadromous behavior. To determine the best transition-rate model, we used
234 fitDiscrete in the R package geiger to fit three continuous-time Markov models of trait evo-
235 lution. These included an "equal rates" model (ER), where a single parameter governs all tran-

sition rates, a "symmetrical model" (SYM) in which forward and reverse transitions share the same parameter, and "all rates different" (ARD) model in which each transitions has its own parameter. The predictive performance of each model was assessed using Akaike Information Criterion (AIC) weights. We then used the model of discrete-trait evolution with the highest predictive performance to estimate a transition matrix, "Q", the was used in stochastic character mapping with the phytools function `make.simmap`, We repeated this producedre to produced a total of 1000 character maps. For each, we set the root prior ("pi") state to freshwater stenohaline, the likely state of the most common recent ancestor of the Actinopterygii (Betancur-R et al., 2015; Carrete Vega and Wiens, 2012; Evans and Claiborne, 2008; Schultz and McCormick, 2012). We summarized the character maps for each tree to generate a posterior estimate of mean number of transitions rate between each salinity preference state and visualized transitions out of and into each state using chord diagrams. In addition, we estimated the most likely state for each node in our study group's phylogeny according to highest posterior probability (PP) at each node under best-fitting model discrete character evolution.

250 *Trait-dependent Diversification*

Using the trimmed (Rabosky et al., 2018) tree, we estimated speciation, extinction, and transition rate parameters for halohabitat-diadromy groups using a multistate speciation and extinction model (MuSSE) in the R package `diversitree` (FitzJohn, 2012). For this we assembled a maximum-likelihood function using an unconstrained model. We then used the maximum-likelihood coefficients to perform a Bayesian analysis, sampling the posterior probability distribution of parameter values generated through a Markov Chain Monte Carlo (MCMC) sampling routine implemented in `diversitree`. We ran the MCMC chain for 2000 generations starting with exponential priors from an initial run of 100 generations. Our final distribution excluded 10% of the samples as burn-in and we assessed convergence using the effective sample sizes. We then calculated net diversification rates from our post-burning samples by subtracting extinction rates from speciation rates.

262

Results

263

Machine Learning Classification

264 Through our ML analysis supported by literature sources, we classified 4398 species as Marine
265 (45.1%), 4404 species as freshwater (45.2%), 832 as euryhaline (8.5%), and 106 as euryhaline-
266 diadromous (1.1%). RF and GBM predictions and FishBase halohabitat classifications were in
267 agreement for 86.8 % (8454) of our study species. Each ML procedure performed similarly in
268 making predictions congruent with FishBase states: Random forest predictions for halohabitat
269 states agreed with FishBase for 90% of our study species, while GMB predictions agreed with
270 89%. With 13.2% discordance between FishBase halohabitat and one of our ML routines, we
271 performed literature searches to confirm halohabitat for 1304 species. Most of these species were
272 classified by FishBase as brackish (our euryhaline state; Supplementary Table !).

273 After literature searches, we determined that only 0.16% of marine species and 0.057% of
274 freshwater species classified as such in FishBase could not be confirmed by our analysis. How-
275 ever, our literature-supported machine learning predictions suggest that 41.7% of the species
276 classified as euryhaline by FishBase were either marine or fresh. Specifically, our analysis sug-
277 gest 16.4% of the species classified by FishBase as euryhaline are freshwater and another 25.3%
278 of species are marine.

279 Both ML predictions of diadromy were in agreement for 89.0% (97) of the 109 species in our
280 study that were classified by Corush (2019) as diadromous. By augmenting this with literature
281 searches, we were able to confirm that 95 (87.2%) of these 109 species were diadromous and an
282 additional 11 were diadromous (Supplementary Table 1).

283

Ancestral State History and Transitions

284 Through fitting discrete models of halohabitat-diadromy states, we found that an all-rates-different
285 mode fit our comparative data best ($AIC_w=1.00$). Our ancestral state reconstructions reveal that

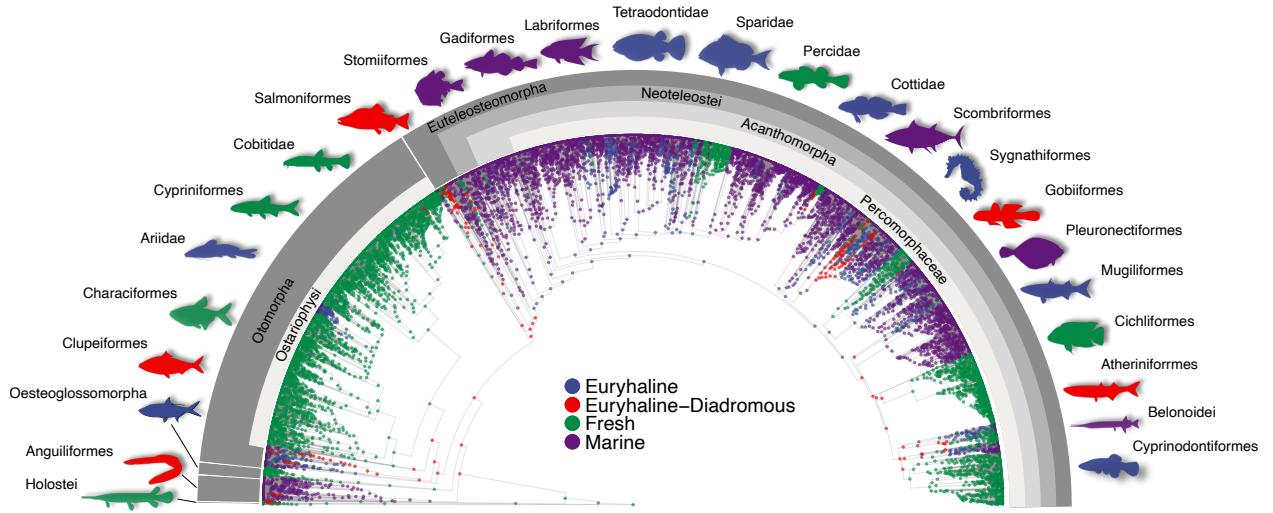


Figure 1: The phylogenetic relationships of 9738 actinopterygian fishes as inferred by Rabosky et al. (2018) and their halohabitat-diadromy history. Node color reflects the state with the highest posterior probability at each internal node under an all-rates-different model of discrete character evolution. Colors at the tip of the tree reflect machine-learning predictions for halohabitat-diadromy state. Colors for silhouettes represent the state for select groups across the ray-finned fish tree of life.

286 diadromy and non-diadromous euryhalinity are ancient states in the ray-finned fishes. The an-
 287 cestral condition of all teleosts was likely diadromous ($PP=0.724$) and this state continued in
 288 the most-recent common ancestor of the Euteleosteomorpha ($PP=0.839$), while the most recent
 289 common ancestor of the neoteleosts was likely non-diadromous euryhaline ($PP= 0.734$; Figure 1).
 290 Non diadromous-euryhalinity continued until the most common recent ancestor of the Ovalen-
 291 taria ($PP=0.664$, clade not highlighted in Figure 1, but see below).

292 Using this model to produce 1000 stochastic character maps resulted in a mean of 1119 and
 293 median of 1120 halohabitat-diadromy state transitions per tree. We found that there were no
 294 direct marine-to-fresh nor fresh-to-marine transitions in our stochastic maps. The most common
 295 transition was from a euryhaline to marine state (Figure 2). Marine to euryhaline transitions
 296 were 66 times more common than marine to euryhaline-diadromous transitions and euryhaline

297 to fresh and euryhaline-diadromous to fresh were about as equally common (Figure 2). Fresh
298 to euryhaline transitions were about 8 times more common than fresh to euryhaline-diadromous
299 transitions and transitions to marine states only occurred out of a euryhaline state (i.e., there
300 were no euryhaline-diadromous to marine transitions (Figure 2).

301 As a whole, these results suggest two important patterns concerning our initial hypotheses.
302 First marine-to-freshwater transitions are not mediated by stepwise transitions to euryhalinity,
303 euryhalinity to didaromy, and finally diadromy to freshwater. The marine-to-fresh transition
304 in ray-finned fishes seems to require euryhalinity, but not necessarily diadromy. We interpret
305 this first pattern because there appear to be no direct marine to fresh transitions in our study
306 group and there were many more euryhaline-fresh transitions than diadromous-fresh transitions.
307 Second, fresh-to-marine transitions do not require diadromy because we found no evidence of
308 diadromous-marine transitions but many euryhaline-marine transitions in our stochastic map-
309 ping.

310 *Diversification Patterns Across Halohabitat-diadromy Groups*

311 Our halohabitat-diadromy dependent diversification model revealed that, for all groups, speci-
312 ation rates outpaced extinction rates resulting in positive net diversification in each (Figure 3).
313 Net diversification was lowest in euryhaline lineages, highest in diadromous lineages, and inter-
314 mediate in fresh and marine lineages. Each of the state-depending net diversification rates were
315 significantly different: there was no overlap between their respective posterior 95% credibility
316 intervals. The modal net-diversification rate for diadromous lineages was nearly 18 time higher
317 than euryhaline lineages, while modal values for fresh and marine lineages were 3.4 and 2.9 times
318 higher than euryhaline lineages, respectfully. The modal net-diversification rate for freshwater
319 lineages was 1.2 times greater than marine lineages (Figure 3).

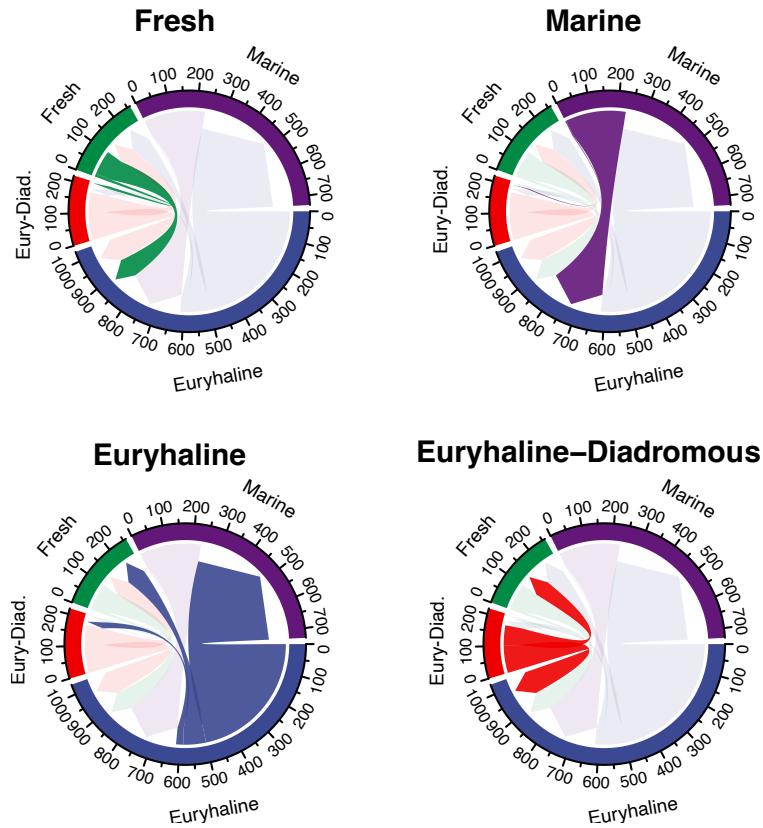


Figure 2: Directionality of transitions out of each halohabitat-diadromy state. Maximum chord width represents the mean number of transitions across 1000 stochastic character maps.

320

Discussion

321

Paths to Major Salinity Transitions

322 By using a combination of ancestral reconstructions and state-dependent diversification analysis,
 323 we uncovered complicated roles for euryhalinity and diadromy in marine-freshwater transitions.
 324 Specifically, we found that marine-freshwater transitions do not happen directly and an inter-
 325 mediate state is required; however, diadromy is not a requirement for these transitions. Our
 326 best-fitting discrete model describes a transition process whereby shifts to freshwater occur out
 327 of a euryhaline, nondiadromous state as frequently as out of a euryhaline-diadromous state
 328 (Figure 2). This model also describes a process in which transitions into marine water are con-

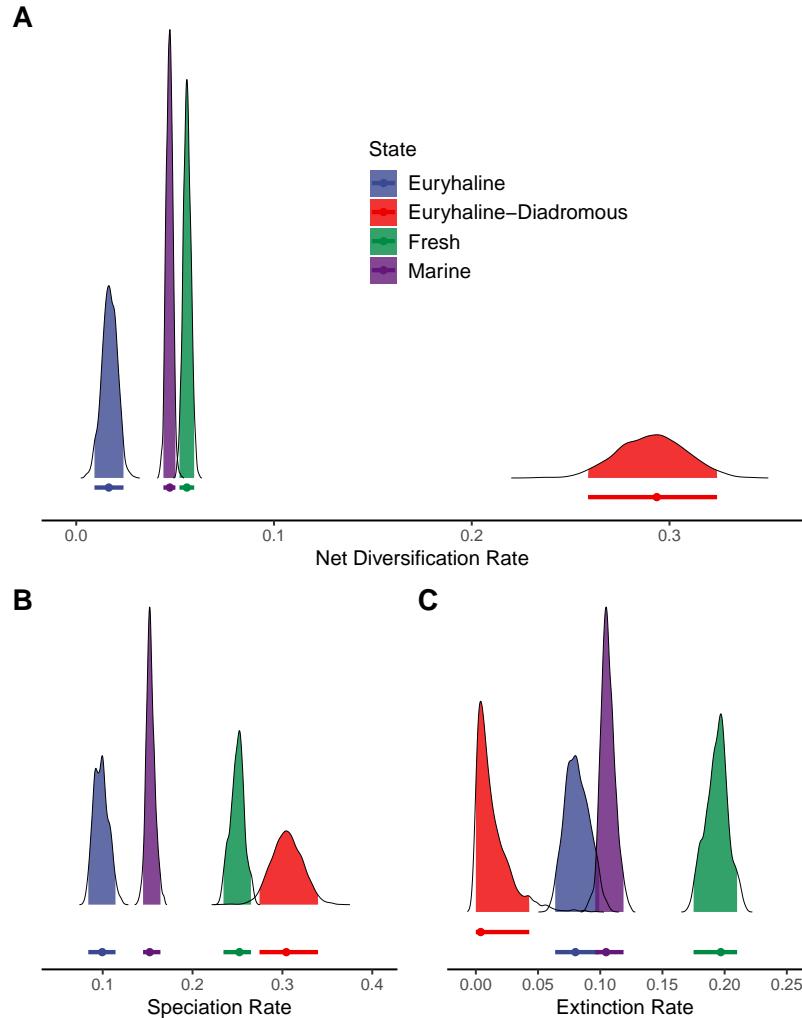


Figure 3: Posterior density distribution of historical net diversification, speciation, and extinction rate estimates (A,B, and C, respectively) for each halohabitat-diadromy state base on MuSSE analysis. Lines below the distributions indicate the 95% credibility intervals and dots the modal values.

329 considerably different, only occurring through nondiadromous euryhalinity. Taken together, these
 330 results challenge Gross' (1987) intermediate state hypothesis, specifically that diadromy is the
 331 final state before colonization of a new salinity habitat.

332 The pathways between major salinity transitions is nicely exemplified by the percomorph
 333 lineage that includes the Aanabantaria, Carangaria, and Ovelentaria, a lineage of marine, fresh-

334 water, euryhaline, and diadromous taxa. Figure 4A depicts the phylogenetic relationships of the
335 group according to (Rabosky et al., 2018), the states for each species and nodes of major lin-
336 eages. Figure 4B depicts the directionality out of each halohabitat-diadromy state. These results
337 symbolize the multiple pathways taken to freshwater transitions by fishes. For some groups,
338 there is overwhelming evidence that euryhalinity alone was the final state before a transition
339 to freshwater (e.g., Andrianichthyidae and Cichliformes). In others, namely the Anabantaria,
340 Atheriniformes and Cyprinodontiformes, diadromous and non-diadromous paths are common.
341 For marine transitions, we found no evidence of diadromy as the path, only euryhalinity (Figure
342 S1).

343 To our knowledge, this study is the first to include both euryhalinity and diadromy in a
344 empirical analysis of marine-freshwater transitions in fishes. However, other studies have found
345 that diadromy is not required for either marine-to-fresh or fresh-to-marine transition. In a broad
346 study of diadromy in ray-finned fishes, Corush (2019) found that marine-to-fresh transitions
347 often included a diadromous ancestor, however, this was not always the case. Corush (2019) also
348 found that direct marine-to-fresh transitions or the reverse, although rare, were possible. This
349 contrasts our results—we inferred no direct transitions from marine to fresh or vice versa (Figure
350 2))—and suggest that this may be due to the fact that Corush (2019) did not include euryhalinity
351 as a state in that study’s analysis.

352 Betancur-R (2010) and Betancur-R et al. (2012) found that marine-to-freshwater transitions
353 were common in ariid catfishes, a group in which diadromy is conspicuously absent. Considered
354 within the larger context of our broader comparative study, we found a similar pattern in the
355 Arioidea, the siluriform clade containing the preponderance of marine catfishes. Within the
356 group, transitions to a marine state occurred only through a euryhaline state (Figure S2). This is
357 not surprising given that diadromy is absent not just within the the Arioidea (Figure S2A), but
358 within the entire Ostariophysi as well (Figure 1).

359 In a study that focused on clupeiform fishes, Bloom and Lovejoy (2012) found that diadromy
360 played only a minor role in transitions between marine and freshwater biomes. Like our general

361 trend, they found that diadromy is only rarely an intermediate condition between marine and
362 freshwater lineages and that no marine species descended from a diadromous ancestor. Our
363 results specific to the clupeiforms contrast this previous study. Within the two major clades of
364 freshwater clupeiforms, lineages within the Engraulini and Dorosomatidae, we found each to
365 have a diadromous ancestor (Figure S3A) and that a transition to freshwater from diadromy
366 was a common path (Figure S3B).

367 Our analysis suggests that transitions to marine waters are mediated solely by the acquisition
368 of euryhalinity. This result corroborates the analysis of (McDowall, 2001) who found no marine
369 lineages of diadromous ancestry. Perhaps the logical pathway to marine invasions from a diadro-
370 mous lineage, if it were to happen at all, would be through catadromy (Gross, 1987), a life history
371 strategy in which reproduction occurs in marine waters and individuals migrate to freshwater
372 for feeding and growth. In this scenario, allopatric speciation due to reproductive isolation is
373 less likely because events that isolate populations, such as vacariance, in marine systems are rare
374 compared to freshwater systems (Albert and Reis, 2011; Bierne et al., 2003; Bloom et al., 2013;
375 Palumbi, 1994; Seehausen and Wagner, 2014) and therefore "sea-locked" catadromous lineages
376 are less likely to occur.

377 In uncovering that net diversification rates are significantly different across major halohabitat-
378 diadromy states, we suggest that euryhalinity and diadromy play different evolutionary roles in
379 major salinity transitions. The net diversification rate of non-diadromous euryhaline lineages,
380 although positive, is extraordinarily low compared to diadromous species (Figure 3A). This is
381 due in large part to the high extinction rate of non-diadromous euryhaline lineages (Figure 3C).
382 Because of this, we assert that the acquisition of euryhalinity is something an evolutionary dead
383 end and a true transitional state, while diadromy is an adaptive life history as well as a path
384 to potential new habitats, at least to freshwaters. Coupled with our stochastic mapping anal-
385 ysis which revealed that both euryhalinity and diadromy are intermediate states in transitions
386 to freshwater, but only euryhalinity is an intermediate state in transitions to marine water, we
387 assert that euryhaline lineages escape habitats of broad salinity habitats back to stenohaline en-

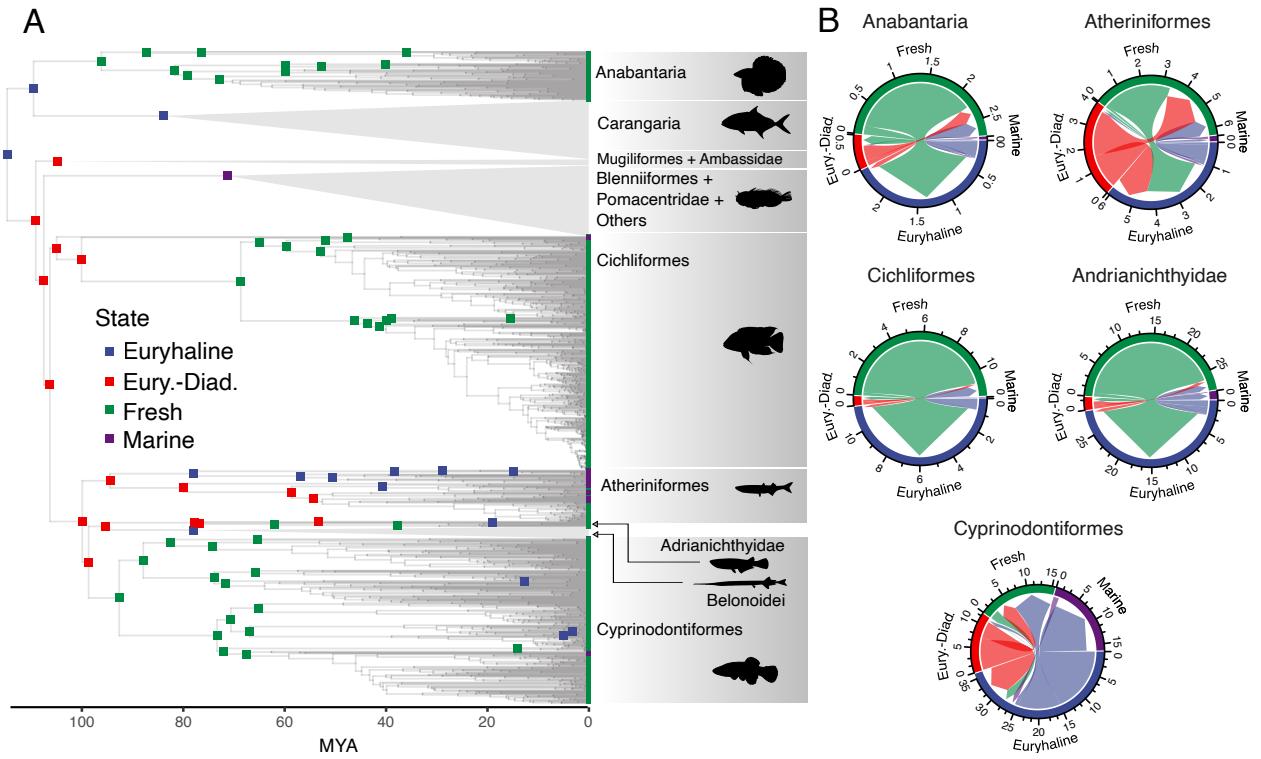


Figure 4: (A) The phylogenetic relationships of the Ovalentaria and Carangaria and other closely related lineages as inferred by Rabosky (2020) and superimposed halohabitat-diadromy states and transitions for predominately freshwater lineages. Collapsed clades represent predominately marine lineages. Square shapes at nodes and tips of the phylogeny represent maximum-likelihood derived state estimates and machine learning predictions for each species, respectively. (B) Directionality of transitions out of each halohabitat-diadromy state for predominately freshwater lineages. Maximum chord width represents the mean number of transitions across 1000 stochastic character maps.

388 environments under strong negative selective pressure, while diadromous lineages thrive between
 389 salinity environments. This contrast may be due to several biotic and abiotic factors. Diadromous
 390 lineages typically span large geographic ranges and this would tend to reduce diversification due
 391 to increased gene flow (Palumbi, 1994); however, by entering freshwaters to reproduce, popula-
 392 tions of diadromous lineages may potentially be subject to vicariance events that reduce gene

393 flow (McDowall, 2001; Seehausen and Wagner, 2014). These events may include several pro-
394 cesses linked to diversification in freshwaters, namely landlocking due to glacial activity (Bell
395 and Foster, 1994) or river capture and sea-level oscillations (Albert et al., 2020; Feutry et al.,
396 2013).

397 Most of the non-diadromous euryhaline species in our study are estuarine (Supplementary
398 Table 1). Many studies have demonstrated that estuarine habitats are subject to repeated ge-
399 ographic disturbances driven by oscillations in sea level (Marko et al., 2010; Shen et al., 2011)
400 and that this process drives diversification and high levels of species richness (Bilton et al., 2002;
401 Wilson, 2006). Our results appear to be at odds with this body of work. We note that the net
402 diversification rate of euryhaline estuarine lineages in our study is positive, even if low due to
403 relatively modest speciation, but slightly lower extinction rate (Figure 3). We suggest that this
404 contradiction can be reconciled if we see estuaries as areas of frequent diversification events by
405 resident lineages, contributing to a positive net diversification rate; however, due to the high
406 species richness, colonizing this biome from marine and especially freshwater is difficult due
407 to increased competition (Betancur-R et al., 2012; Brockhurst et al., 2007; Fukami et al., 2007;
408 Meyer and Kassen, 2007; Wiens and Graham, 2005). Thus, euryhaline lineages may flee estuarine
409 environments for marine or, less often, freshwaters in the face of this competition.

410 *Machine Learning and Comparative Data*

411 Overall, our ML halohabitat assignments were largely congruent with those taken from FishBase;
412 however, for euryhaline species, the halohabitat states assigned through our analysis differ sig-
413 nificantly from the states reported in FishBase. Taking the cypriniforms as an example, of the 102
414 species included in our study that FishBase classified as brackish, 88 of these were classified
415 as fresh according to our ML results. For all of these records, we could find no determinative
416 records of marine or estuarine occurrence in the literature or we found statements in the liter-
417 ature that unambiguously indicated a freshwater status (Supplementary Table 1). For many of
418 these euryhaline species, FishBase classifies them as brackish, despite explicitly reporting habitat

419 and behavioral information that does not support this classification or reports estuarine or brack-
420 ish residence with no supporting reference. For instance, the cyprinid *Acrossocheilus iridescent* is
421 classified as brackish, however, there is no supporting literature citation and statements about
422 habitat refer only to freshwaters. We would like to underscore that FishBase represents an ex-
423 tremely valuable tool for comparative and ecological research. Nonetheless, we also suggest that
424 researchers should scrutinize the ecological information compiled by this important database
425 before undertaking comparative analysis using these data.

426 *Conclusions*

427 Taking these results as a whole, we assert that euryhalinity was a key innovation that permitted
428 ray-finned fishes to transition between marine and freshwater biomes; however, the acquisition
429 of euryhalinity constrains diversification and must be seen a transitional state rather than an
430 adaptive phenotype. Diadromy, on the other hand, represents an adaptive life history strategy
431 in addition to being a transitional phenotype in shifts from marine to fresh waters only. Further-
432 more, ecological and abiotic factors likely play important roles in shaping transition dynamics.
433 Specifically, higher competition in new niches made available by the ability to withstand a broad
434 range of salinity (e.g., estuaries) may promote transitions to stenohaline biomes, while different
435 frequencies of vicariance events between marine and freshwater realms may promote diversi-
436 fication through diadromy in the path to freshwater and limit them in the path to the marine
437 realm.

438 *Data Availability*

439 Code and data for our analyses are available at <https://github.com/ckenaley/SalinityTransitions>
440 and Dryad Repository DOI: 10.5061/dryad.bzkh189g4.

441 *Author Contributions*

442 CK conceived the study, WO and CK performed the analyses and wrote the manuscript.

443 *Funding*

444 Funding for this study was provided Morrissey College of Arts and Sciences at Boston College
445 (both CK and WO).

446 *Acknowledgments*

447 We wish to thank the members of the Kenaley Lab past and present, the staff at Boston College's
448 Information Technology Services for assistance in using their Linux Cluster, and xx anonymous
449 reviewers whose suggestions greatly improved the manuscript

450

Literature Cited

451 Abell, R., M. L. Thieme, C. Revenga, M. Bryer, M. Kottelat, N. Bogutskaya, B. Coad, N. Mandrak,
452 S. C. Balderas, W. Bussing, et al., 2008. Freshwater ecoregions of the world: a new map of
453 biogeographic units for freshwater biodiversity conservation. BioScience 58:403–414.

454 Albert, J. S. and R. Reis, 2011. Historical Biogeography of Neotropical Freshwater Fishes. Univ
455 of California Press.

456 Albert, J. S., V. A. Tagliacollo, and F. Dagosta, 2020. Diversification of Neotropical freshwater
457 fishes. Annual Review of Ecology, Evolution, and Systematics 51:27–53.

458 Alder, J., 2003. Putting the coast in the sea around us project. The Sea Around Us Newsletter
459 15:1–2.

460 Bell, M. A. and S. A. Foster, 1994. Introduction to the evolutionary biology of the threespine
461 stickleback. The Evolutionary Biology of the Threespine Stickleback Pp. 1–27.

462 Betancur-R, R., 2010. Molecular phylogenetics supports multiple evolutionary transitions from
463 marine to freshwater habitats in ariid catfishes. Molecular Phylogenetics and Evolution 55:249–
464 258.

- 465 Betancur-R, R., G. Ortí, and R. A. Pyron, 2015. Fossil-based comparative analyses reveal ancient
466 marine ancestry erased by extinction in ray-finned fishes. *Ecology Letters* 18:441–450.
- 467 Betancur-R, R., G. Ortí, A. M. Stein, A. P. Marceniuk, and R. Alexander Pyron, 2012. Appar-
468 ent signal of competition limiting diversification after ecological transitions from marine to
469 freshwater habitats. *Ecology Letters* 15:822–830.
- 470 Bierne, N., F. Bonhomme, and P. David, 2003. Habitat preference and the marine-speciation
471 paradox. *Proceedings of the Royal Society B: Biological Sciences* 270:1399–1406.
- 472 Bilton, D., J. Paula, and J. Bishop, 2002. Dispersal, genetic differentiation and speciation in
473 estuarine organisms. *Estuarine, Coastal and Shelf Science* 55:937–952.
- 474 Bloom, D. D. and N. Lovejoy, 2012. Molecular phylogenetics reveals a pattern of biome con-
475 servatism in New World anchovies (family Engraulidae). *Journal of Evolutionary Biology*
476 25:701–715.
- 477 Bloom, D. D., J. T. Weir, K. R. Piller, and N. R. Lovejoy, 2013. Do freshwater fishes diversify
478 faster than marine fishes? A test using state-dependent diversification analyses and molecular
479 phylogenetics of New World silversides (Atherinopsidae). *Evolution* 67:2040–2057.
- 480 Boettiger, C., D. T. Lang, and P. Wainwright, 2012. rfishbase: Exploring, manipulating and
481 visualizing fishbase data from r. *Journal of Fish Biology* 81:2030–2039.
- 482 Bollback, J. P., 2006. SIMMAP: stochastic character mapping of discrete traits on phylogenies.
483 *BMC Bioinformatics* 7:1–7.
- 484 Breiman, L., 2001. Random forests. *Machine Learning* 45:5–32.
- 485 de Brito, V., R. Betancur-R, M. D. Burns, T. J. Buser, K. W. Conway, J. P. Fontenelle, M. A. Kolmann,
486 W. T. McCraney, C. E. Thacker, and D. D. Bloom, 2022. Patterns of phenotypic evolution
487 associated with marine/freshwater transitions in fishes. *Integrative and Comparative Biology*
488 62:406–423.

- 489 Brockhurst, M. A., N. Colegrave, D. J. Hodgson, and A. Buckling, 2007. Niche occupation limits
490 adaptive radiation in experimental microcosms. PLoS One 2:e193.
- 491 Carrete Vega, G. and J. J. Wiens, 2012. Why are there so few fish in the sea? Proceedings of the
492 Royal Society B: Biological Sciences 279:2323–2329.
- 493 Chamberlain, S. and C. Boettiger, 2017. R python, and ruby clients for gbif species occurrence
494 data. PeerJ PrePrints .
- 495 Corush, J. B., 2019. Evolutionary patterns of diadromy in fishes: More than a transitional state
496 between marine and freshwater. BMC Evolutionary Biology 19:1–13.
- 497 Costello, M. J., P. Bouchet, G. Boxshall, K. Fauchald, D. Gordon, B. W. Hoeksema, G. C. Poore,
498 R. W. van Soest, S. Stöhr, T. C. Walter, et al., 2013. Global coordination and standardisation
499 in marine biodiversity through the World Register of Marine Species (WoRMS) and related
500 databases. PloS one 8:e51629.
- 501 Davis, A. M. and R. Betancur-R, 2017. Widespread ecomorphological convergence in multiple
502 fish families spanning the marine–freshwater interface. Proceedings of the Royal Society B:
503 Biological Sciences 284:20170565.
- 504 Davis, A. M., P. J. Unmack, B. Pusey, J. Johnson, and R. G. Pearson, 2012. Marine–freshwater
505 transitions are associated with the evolution of dietary diversification in terapontid grunters
506 (Teleostei: Terapontidae). Journal of Evolutionary Biology 25:1163–1179.
- 507 Edwards, S. L. and W. S. Marshall, 2012. Principles and patterns of osmoregulation and eury-
508 halinity in fishes. Pp. 1–44, *in* Fish physiology, vol. 32. Elsevier.
- 509 Evans, D. H. and J. B. Claiborne, 2008. Osmotic and ionic regulation in fishes. Pp. 295–366, *in*
510 Osmotic and ionic regulation. CRC Press.
- 511 Feutry, P., M. Castelin, J. R. Ovenden, A. Dettai, T. Robinet, C. Cruaud, and P. Keith, 2013.

- 512 Evolution of diadromy in fish: Insights from a tropical genus (*Kuhlia* species). The American
513 Naturalist 181:52–63.
- 514 FitzJohn, R. G., 2012. Diversitree: Comparative phylogenetic analyses of diversification in R.
515 Methods in Ecology and Evolution 3:1084–1092.
- 516 Friedman, J. H., 2001. Greedy function approximation: A gradient boosting machine. Annals of
517 Statistics Pp. 1189–1232.
- 518 Froese, R. and D. Pauly, 2023. Fishbase. URL www.FishBase.org.
- 519 Fryda, T., E. LeDell, N. Gill, S. Aiello, A. Fu, A. Candel, C. Click, T. Kraljevic, T. Nykodym,
520 P. Aboyoun, M. Kurka, M. Malohlava, S. Poirier, and W. Wong, 2023. h2o: R Interface for the
521 'H2O' Scalable Machine Learning Platform. URL <https://CRAN.R-project.org/package=h2o>. R package version 3.42.0.2.
- 523 Fukami, T., H. J. Beaumont, X.-X. Zhang, and P. B. Rainey, 2007. Immigration history controls
524 diversification in experimental adaptive radiation. Nature 446:436–439.
- 525 Garcia, H. E., R. A. Locarnini, T. P. Boyer, J. I. Antonov, O. K. Baranova, M. M. Zweng, J. R.
526 Reagan, D. R. Johnson, A. V. Mishonov, and S. Levitus, 2013. World Ocean Atlas 2013. Volume
527 4, Dissolved inorganic nutrients (phosphate, nitrate, silicate).
- 528 Gross, M. R., 1987. Evolution of diadromy in fishes. *in* American Fisheries Society Symposium,
529 vol. 1, Pp. 14–25.
- 530 Holstein, J., 2018. WORMS: Retrieving Aphia Information from World Register of Marine Species.
531 URL <https://CRAN.R-project.org/package=worms>. R package version 0.2.2.
- 532 Kelley, D. and C. Richards, 2022. ocedata: Oceanographic Data Sets for 'oce' Package. R package
533 version 0.2.2.
- 534 Kültz, D., 2015. Physiological mechanisms used by fish to cope with salinity stress. The Journal
535 of Experimental Biology 218:1907–1914.

- 536 Lee, C. E. and M. A. Bell, 1999. Causes and consequences of recent freshwater invasions by
537 saltwater animals. *Trends in Ecology & Evolution* 14:284–288.
- 538 Marko, P. B., J. M. Hoffman, S. A. Emme, T. M. McGovern, C. C. Keever, and L. Nicole Cox, 2010.
539 The ‘Expansion–Contraction model’ of Pleistocene biogeography: Rocky shores suffer a sea
540 change? *Molecular Ecology* 19:146–169.
- 541 Massicotte, P. and A. South, 2023. rnaturalearth: World Map Data from Natural Earth. URL
542 <https://CRAN.R-project.org/package=rnaturalearth>. R package version 0.3.2.
- 543 McCormick, S. D., 2001. Endocrine control of osmoregulation in teleost fish. *American Zoologist*
544 41:781–794.
- 545 McCormick, S. D., A. P. Farrell, and C. J. Brauner, 2013. *Fish Physiology: Euryhaline Fishes*.
546 Academic Press.
- 547 McCormick, S. D. and R. L. Saunders, 1987. Preparatory physiological adaptations for marine
548 life of salmonids: Osmoregulation, growth, and metabolism. *in* American Fisheries Society
549 Symposium, vol. 1(21), Pp. 1–229.
- 550 McDowall, R. M., 1997. The evolution of diadromy in fishes (revisited) and its place in phyloge-
551 netic analysis. *Reviews in Fish Biology and Fisheries* 7:443–462.
- 552 ———, 2001. Diadromy, diversity and divergence: Implications for speciation processes in fishes.
553 *Fish and Fisheries* 2:278–285.
- 554 Meyer, J. R. and R. Kassen, 2007. The effects of competition and predation on diversification in a
555 model adaptive radiation. *Nature* 446:432–435.
- 556 Palumbi, S. R., 1994. Genetic divergence, reproductive isolation, and marine speciation. *Annual
557 Review of Ecology and Systematics* 25:547–572.
- 558 Patterson, T. and N. Kelso, 2012. World urban areas, Landscan, 1:10 million (2012). North
559 American Cartographic Information Society .

- 560 R Core Team, 2022. R: A Language and Environment for Statistical Computing. R Foundation
561 for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- 562 Rabosky, D. L., 2020. Speciation rate and the diversity of fishes in freshwaters and the oceans.
563 Journal of Biogeography 47:1207–1217.
- 564 Rabosky, D. L., J. Chang, P. O. Title, P. F. Cowman, L. Sallan, M. Friedman, K. Kaschner, C. Gari-
565 lao, T. J. Near, M. Coll, et al., 2018. An inverse latitudinal gradient in speciation rate for marine
566 fishes. Nature 559:392–395.
- 567 Santini, F., M. T. Nguyen, L. Sorenson, T. B. Waltzek, J. Lynch Alfaro, J. M. Eastman, and M. E.
568 Alfaro, 2013. Do habitat shifts drive diversification in teleost fishes? An example from the
569 pufferfishes (Tetraodontidae). Journal of Evolutionary Biology 26:1003–1018.
- 570 Schmidt-Kloiber, A., V. Bremerich, A. De Wever, S. C. Jähnig, K. Martens, J. Strackbein, K. Tock-
571 ner, and D. Hering, 2019. The freshwater information platform: A global online network
572 providing data, tools and resources for science and policy support. Hydrobiologia 838:1–11.
- 573 Schultz, E. T. and S. D. McCormick, 2012. Euryhalinity in an evolutionary context. Fish physiol-
574 ogy 32:477–533.
- 575 Seehausen, O. and C. E. Wagner, 2014. Speciation in freshwater fishes. Annual review of Ecology,
576 Evolution, and Systematics 45:621–651.
- 577 Shen, K.-N., B. W. Jamandre, C.-C. Hsu, W.-N. Tzeng, and J.-D. Durand, 2011. Plio-Pleistocene
578 sea level and temperature fluctuations in the northwestern Pacific promoted speciation in the
579 globally-distributed flathead mullet *Mugil cephalus*. BMC Evolutionary Biology 11:1–17.
- 580 Simpson, G. G., 1953. The major features of evolution. Columbia University Press.
- 581 ———, 1984. Tempo and mode in evolution. Columbia University Press.

- 582 Spalding, M. D., H. E. Fox, G. R. Allen, N. Davidson, Z. A. Ferdaña, M. Finlayson, B. S. Halpern,
583 M. A. Jorge, A. Lombana, S. A. Lourie, et al., 2007. Marine ecoregions of the world: A biore-
584 gionalization of coastal and shelf areas. BioScience 57:573–583.
- 585 Stroud, J. T. and J. B. Losos, 2016. Ecological opportunity and adaptive radiation. Annual Review
586 of Ecology, Evolution, and Systematics 47:507–532.
- 587 Wiens, J. J. and C. H. Graham, 2005. Niche conservatism: Integrating evolution, ecology, and
588 conservation biology. Annual Review of Ecology, Evolution, and Systematics 36:519–539.
- 589 Wilson, A., 2006. Genetic signature of recent glaciation on populations of a near-shore marine
590 fish species (*Syngnathus leptorhynchus*). Molecular Ecology 15:1857–1871.
- 591 Yoder, J., E. Clancey, S. Des Roches, J. Eastman, L. Gentry, W. Godsoe, T. Hagey, D. Jochim-
592 sen, B. Oswald, J. Robertson, et al., 2010. Ecological opportunity and the origin of adaptive
593 radiations. Journal of Evolutionary Biology 23:1581–1596.

594

Supplementary Content

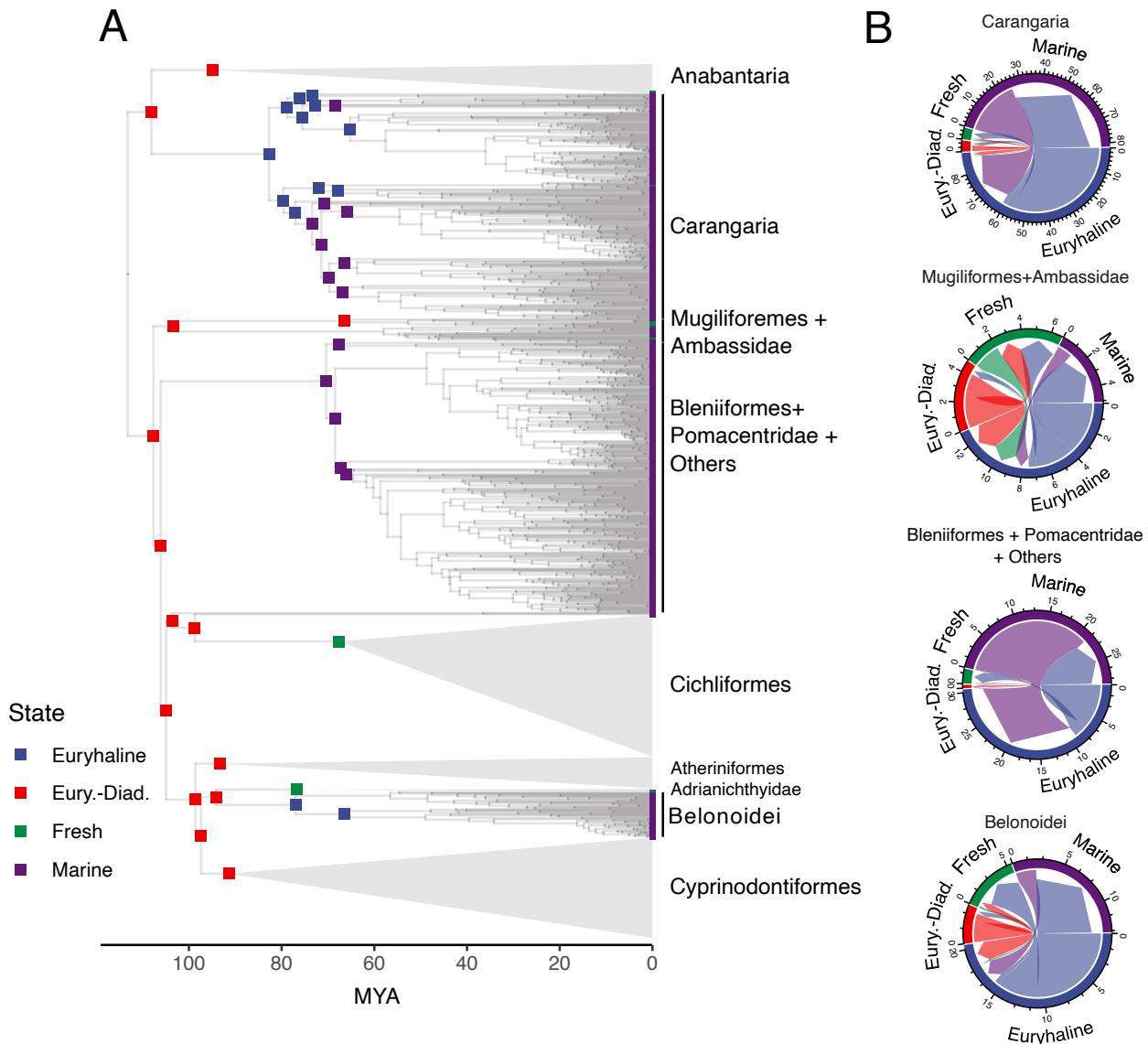


Figure S1: (A) The phylogenetic relationships of the Ovalentaria and Carangaria and other closely related lineages as inferred by Rabosky (2020) and superimposed halohabitat-diadromy states and transitions for predominately marine lineages. Collapsed clades represent predominantly freshwater lineages. Square shapes at nodes and tips of the phylogeny represent maximum-likelihood derived state estimates and machine learning predictions for each species, respectively. (B) Directionality of transitions out of each halohabitat-diadromy state for predominantly marine lineages. Maximum chord width represents the mean number of transitions across 1000 stochastic character maps.

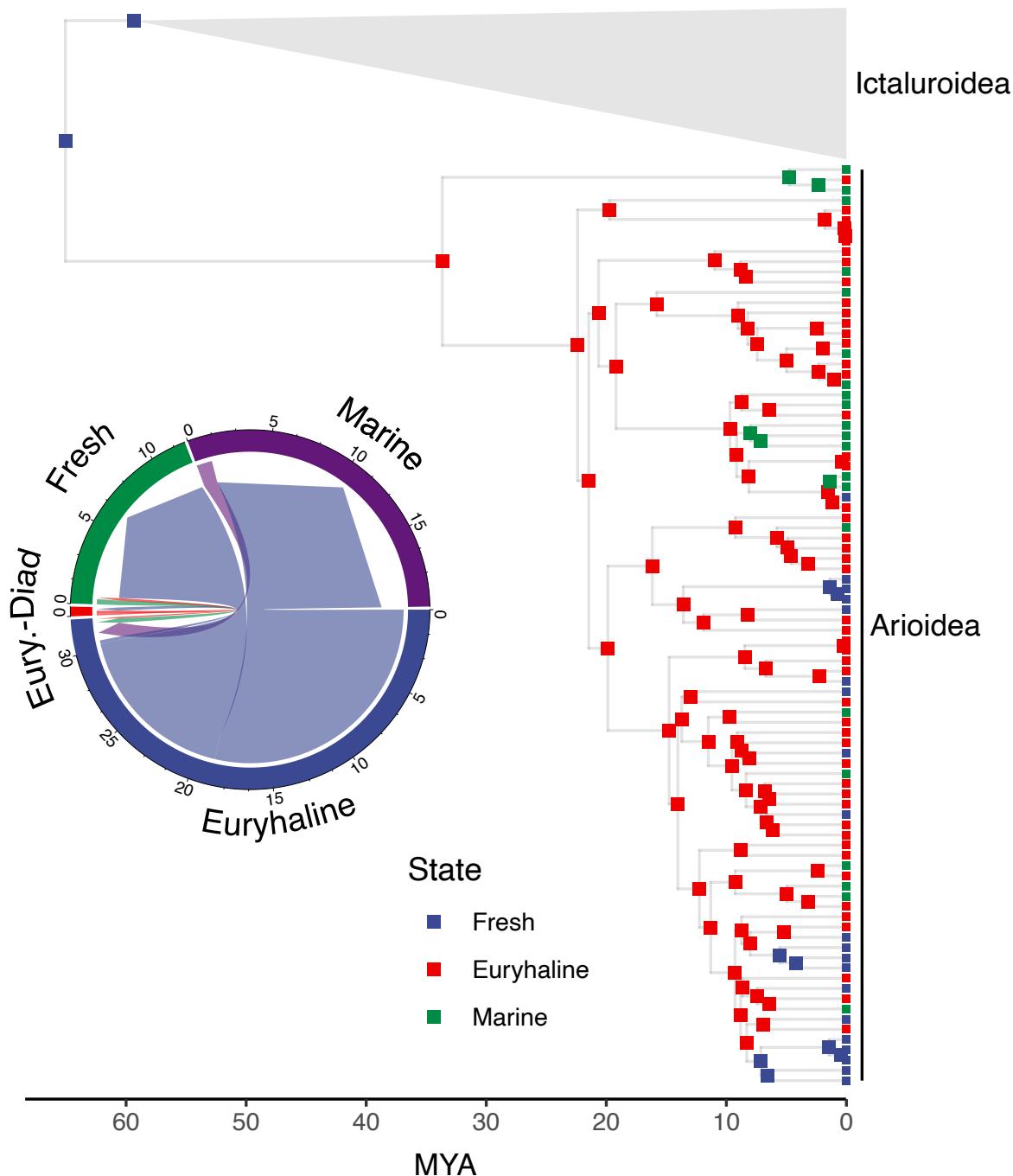


Figure S2: The phylogenetic relationships of arioid catfishes as inferred by Rabosky (2020) and superimposed halohabitat-diadromy states and transitions for this predominately marine lineages. Collapsed clades represent predominantly freshwater sister lineage Ictaluroidea. Square shapes at nodes and tips of the phylogeny represent maximum-likelihood derived state estimates and machine learning predictions for each species, respectively. Inset chord diagram represents the directionality of transitions out of each halohabitat-diadromy state for the Arioidea. Maximum chord width represents the mean number of transitions across 1000 stochastic character maps.

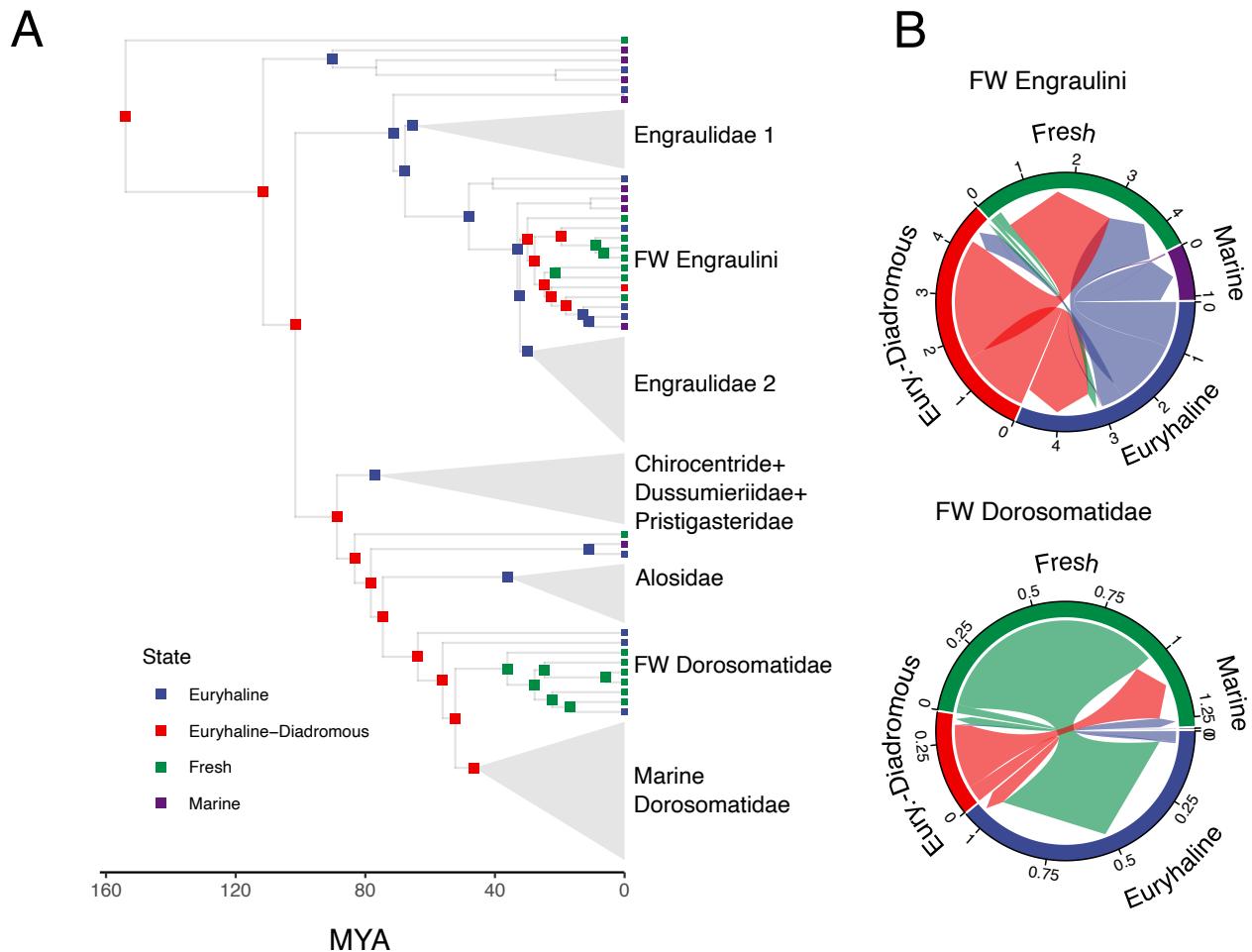


Figure S3: (A) The phylogenetic relationships of the Clupeiformes as inferred by Rabosky (2020) and superimposed halohabitat-diadromy states and transitions for predominately freshwater lineages. Collapsed clades represent predominantly marine lineages. Square shapes at nodes and tips of the phylogeny represent maximum-likelihood derived state estimates and machine learning predictions for each species, respectively. (B) Directionality of transitions out of each halohabitat-diadromy state for the two predominantly freshwater lineages. Maximum chord width represents the mean number of transitions across 1000 stochastic character maps.