Parallel and non-parallel features of adaptive radiation in Yucatán pupfishes

Matthew C. Kustra[1,2,3,*], David Tian[1,2], M. Fernanda Palominos[1,2], Feifei Guo[1,2], Dylan Chau[1,2], Oskar Golwala[1,2], HoWan Chan[4], Andrés Alvarez Zapata[5], Reyna Guadalupe Cetz Paredes[5], Frida Ximena Cortés Sánchez[6], Sonia Gabriela Hernández[6], Adan Fernando Mar-Silva[6], Fernando Mex[7], Charles Tralka[1,2], Maribel Badillo-Alemán[7], Juan J. Schmitter-Soto[8], Carlos A. Gracida-Juárez[9], Christopher M. Martinez[10], Jairo Arroyave[6], and Christopher H. Martin[1,2]

[1]Department of Integrative Biology, University of California, Berkeley, CA, United States

[2]Museum of Vertebrate Zoology, University of California, Berkeley, CA, United States

[3]Miller Institute for Basic Research in Science, University of California, Berkeley, CA, United States

[4]Department of BioSciences, Rice University, Houston, TX, United States

[5]Instituto Tecnologico Superior de Conkal, Yucatan, Mexico

[6]Instituto de Biología, Universidad Nacional Autónoma de México, Mexico City, Mexico

[7]Unidad Multidisciplinaria de Docencia e Investigación, Unidad Sisal, Universidad Nacional Autónoma de México, Sisal, Yucatán, Mexico

[8] Departamento de Sistemática y Ecología Acuática, El Colegio de la Frontera Sur, Chetumal, Quintana Roo, Mexico

[9]Subdirección de Posgrado e Investigación, Tecnológico Nacional de México, Campus Felipe Carrillo Puerto.Carretera a Vigía Chico km 1.5, 77200 Felipe Carrillo Puerto, Quintana Roo.

[10]Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, United States

*Corresponding author: matthewckustra@berkeley.edu

Abstract

Understanding the extent of parallelism across adaptive radiations remains a central problem in evolutionary biology. We used whole-genome resequencing of 123 individuals to study an adaptive radiation of *Cyprinodon* pupfishes in Laguna Chichancanab, Mexico, compared to previous research on an independent radiation of San Salvador Island (SSI) pupfishes in the Bahamas, to assess the repeatability of their genetic architecture, sources of adaptive variation, and stages of selection. Despite rapid craniofacial divergence of trophic specialists within 8-15 kya, only two candidate genes (3.6%) were shared between Caribbean radiations. Although adaptive introgression played a major role in SSI, we found minimal evidence of adaptive introgression in Chichancanab, likely due to the higher geographical isolation of this inland lake. Instead, de novo mutations provided a substantial source of adaptive variation (30.6%) for the endemic zooplanktivore, nearly 10 times higher than the scale-eater in SSI. We found strong evidence that selection occurred in stages, first on regulatory and standing genetic variation, then on de novo and nonsynonymous mutations, parallel to SSI. Unique to Chichancanab, adaptive variants near opsin and spermatogenesis genes are consistent with our findings of greater visual acuity and differences in sperm morphology, respectively, within the zoooplanktivore. Our study suggests that de novo mutations may contribute more to rapid adaptive radiations than previously appreciated, particularly in isolated environments.

Introduction

Much of biodiversity results from adaptive radiation—when a clade diversifies rapidly into three or more new species (1–5). Young adaptive radiations provide excellent systems for investigating speciation and adaptation (3, 6–9). However, it remains unclear to what degree radiations occur in parallel given similar starting positions of genetic diversity, gene flow, and environments (10, 11).

Beneficial de novo mutations may provide novel variation to reach new fitness peaks on the adaptive landscape during adaptive radiations (12, 13); however, the time needed for de novo beneficial mutations to arise contrasts with the rapid divergence of adaptive radiations (11, 14, 15). Alternatively, hybridization could introduce large amounts of genetic variation and novel phenotypes (transgressive segregation) through hybrid swarms, enabling populations to reach previously inaccessible fitness peaks (16). Support for the role of hybridization in adaptive radiations has been demonstrated both theoretically (17, 18) and empirically in many systems (3, 19–27). However, we still do not know the importance of different sources of genetic variation across time and space.

Caribbean pupfishes are a remarkable system for assessing the repeatability of adaptive radiations due to two independent, recent adaptive radiations that contain a generalist and multiple trophic specialists within isolated lake environments that have historically lacked predators. The San Salvador Island (SSI), Bahamas radiation proceeded predominantly via the reassembly of pre-existing standing genetic variation found throughout the Caribbean, with additional contributions of adaptive introgression in both molluscivore (durophage) and scale-eating (lepidophage) specialists, and an additional 2% of de novo mutations within the scale-eater (26, 28–32). However, it is unclear how a second radiation in Laguna Chichancanab,

Mexico, drawing from the same Caribbean pool of standing genetic variation over the same timeframe unfolded, allowing us to assess the degree to which radiations are parallel.

Here, we use whole-genome resequencing of 123 individuals to provide insight into the genomic architecture and genetic origins of the Laguna Chichancanab *Cyprinodon* pupfish radiation (33). This recent adaptive radiation (~8000 years old (34)) consisted of at least five species historically (Strecker, 2005; Humphries and Miller 1981), including a detritivore (*C. beltrani*), piscivore (*C. maya*), rocky substrate specialist (*C. verecundus*), bivalve specialist (*C. labiosus*), and zooplanktivore (*C. simus*). However, due to recent colonization by native fishes not historically present in the lake and invasive colonization by African tilapia (*Oreochromis* sp.) (36, 37), all trophic specialist species now appear to be extinct since our collections in 2022, with only the generalist, *C. beltrani*, remaining in 2024 (MK pers. obs.) and 2025 (CG, JSS pers. obs.). Despite exhibiting one of the fastest rates of trait diversification of any vertebrate in tooth length (130 times faster than background rates across *Cyprinodon* (29)) and rapid divergence in other craniofacial morphology (Figure 1A), previous genetic studies of microsatellite data and mtDNA suggest that this radiation has low genetic differentiation (38, 39), potentially enabling the identification of highly differentiated loci among species that may be involved in trophic adaptation.

Results

Demographic history

To understand the demographic history and population structure of the Yucatán *Cyprinodon* pupfishes, we collected specimens from the wild in 2022 (Figure 1B) and conducted whole genome resequencing of the species flock and the coastal sister species (n = 15 *C.*

artifrons, *n* = 35 *C. beltrani*, *n* = 35 *C. labiosus*, and *n* = 17 *C. simus*). We aligned these samples, as well as eleven other previously sequenced Caribbean pupfish species, to the *Cyprindon nevadensis mionectes* reference genome (GCA_030533445.1). After calling and filtering variants, our dataset included 23,960,536 SNPs (12.4x mean depth).

Using MSMC2 on Chichancanab species on a linkage disequilibrium pruned set of SNPs (*n* =5,476,855 SNPs) and a mutation rate of 1.56 x 10^{-8} substitutions per base pair (estimated from a Caribbean *Cyprinodon* species (26)), we found that the three Chichancanab species (*C. beltrani*, *C. labiosus*, and *C. simus*) diverged from the coastal sister species *C. artifrons* at the end of the Pleistocene, 10-15 kya, consistent with receding sea levels resulting in the formation of the Chichancanab lake (34). Since divergence from *C. artifrons*, the Chichancanab species flock shows a similar demographic history, characterized by a sharp decline in effective population sizes, consistent with endemic sympatric species within an isolated lake (Figure 1C).

Genetic differentiation of the species flock

Using a linkage-pruned set of SNPs and the subset of data restricted to Yucatán pupfish, we used principal component analysis and ADMIXTURE to visualize population structure. We found that all species formed distinct, non-overlapping genetic clusters, with *C. artifrons* being the most genetically divergent from and sister to the Chichancanab species (Figure 1D). Within the Chichancanab species flock, *C. simus* was more genetically divergent from *C. labiosus* and *C. beltrani* (Figure 1D). These results were also supported by genome-wide F_{ST} estimates and a phylogenetic tree constructed with *ADMIXTOOLS2* (Table S1; Figure S1). Four clusters corresponded to the four different species, as determined by ADMIXTURE (Figure 1E). All *C.*

artifrons individuals showed minimal evidence of admixture with Chichancanab species (Figure 1E).

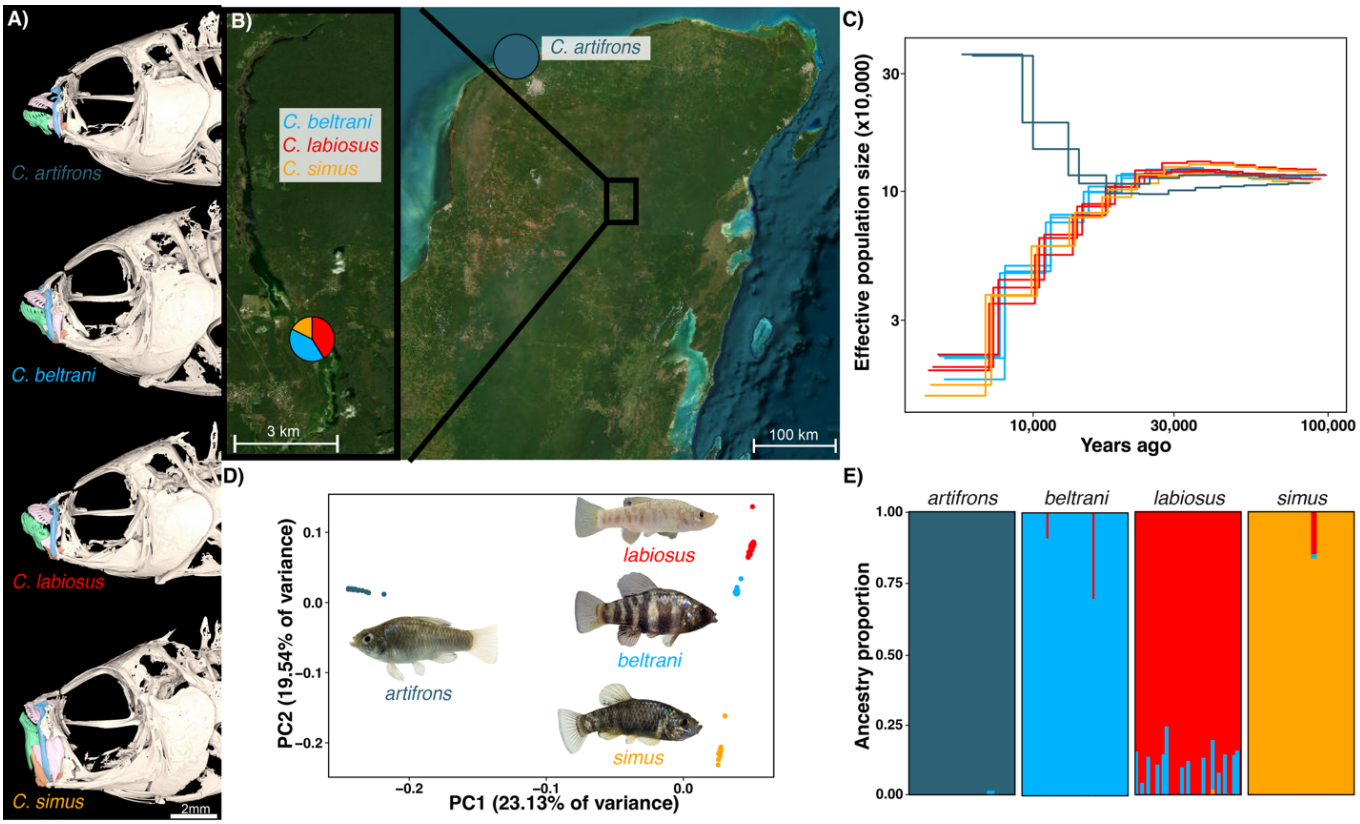


Figure 1. Yucatán *Cyprinodon* pupfishes show strong genetic differentiation. (A) µCT scans of Yucatán pupfish. (B) Google Earth image of the Yucatán Peninsula showing the sampling site for *C. artifrons*, with an inset map of Laguna Chichancanab. The pie chart represents the relative frequency of the different Chichancanab species in our dataset. (C) Estimated demographic history with *MSMC2* using linkage disequilibrium pruned SNPs. Each colored line represents an individual, with colors corresponding to different species. (D) Principal component analysis using a set of linkage disequilibrium pruned SNPs. (E) Estimated ancestry proportions for each individual inferred using *ADMIXTURE* with K = 4 (see Figure S2 for K=2-3) from the linkage-pruned dataset.

138 *Adaptive loci within trophic specialists*

139 Although genome-wide differentiation among Chichancanab species was moderately low
140 (F_{ST} ranging from 0.032 to 0.09; Figure 2A; Table S1), we were able to identify highly
141 differentiated regions of the genome within the trophic specialist zooplanktivore *C. simus*. We
142 identified 19 fixed and 1,127 nearly-fixed (F_{ST} >0.95) SNPs in *C. simus* relative to other
143 Chichancanab pupfishes (Table S2, 3). In contrast, we found zero fixed or nearly-fixed SNPs
144 within the generalist *C. beltrani* or bivalve-specialist C. *labiosus*. However, we were able to
145 identify some moderately differentiated SNPs within *C. labiosus* (n = 97) and *C. beltrani* (n =
146 13) relative to the other Chichancanab species at an F_{ST} threshold of 0.8 (Table S2).

147 To search for signals of selection, we calculated the normalized four-taxon population
148 branch statistic (PBS_{nj}; Schmidt et al., 2019) across 10-kb non-overlapping windows because
149 linkage disequilibrium decayed at this distance (Figure S3). Simulation studies demonstrate that
150 normalized *PBS* is effective in identifying both hard and soft selective sweeps (40). Out of the
151 1,146 SNPs fixed or nearly fixed within *C. simus*, 661 were within a strongly supported selective
152 sweep window (top 1% PBS_{nj} outlier (40)), including 425 SNPs within 20 kb of the first or last
153 exon of an annotated gene. These 425 candidate adaptive SNPs ("adaptive variants") were
154 located in the proximity of or within 55 genes (Figure 2A), distributed across 19 out of 24
155 chromosomes (Figure 2A). Adaptive variants associated with the same gene were often
156 distributed across multiple introns or exons and spaced apart by 1kb or more (Figure S4, Table
157 S4).

158

159 *Source and function of adaptive genetic variation in the zooplanktivore*

We next characterized the location of the 425 candidate adaptive variants near or within a gene in the zooplanktivore *C. simus*. More than 95% were in intronic or flanking regions (20-kb upstream or downstream of a gene; Figure 2B). Within coding regions, 13 out of 20 adaptive variants were nonsynonymous mutations (Figure 2B). The number of adaptive variants per gene followed an exponential distribution, with a handful of genes containing many variants and most genes with only a single adaptive variant (Figure 2B). We functionally characterized these genes by manually checking online databases and doing a literature search (Table S3; Extended Table S1). 32 out of the 55 candidate genes were associated with craniofacial development (n =11; 20%) or sensory/neural function (n = 21; 38.2%). We manually checked if any of these genes overlapped with the 176 unique candidate genes identified in the SSI radiation (26), and we found only two overlapping candidate genes (*DNM1* and *SPTLC3*), both of which are associated with motor neuron development (41, 42) and are scale-eater-specific candidate genes.

Next, we characterized the source of the variants and found that 63% of the adaptive SNPs in *C. simus* also existed as standing genetic variation in the coastal Yucatan population of *C. artifrons* ("standing variation"), 31% were *de novo* variants found only in Chichancanab, and 6% were detected as standing genetic variation in other pupfish species but not *C. artifrons* (Figure 2C).

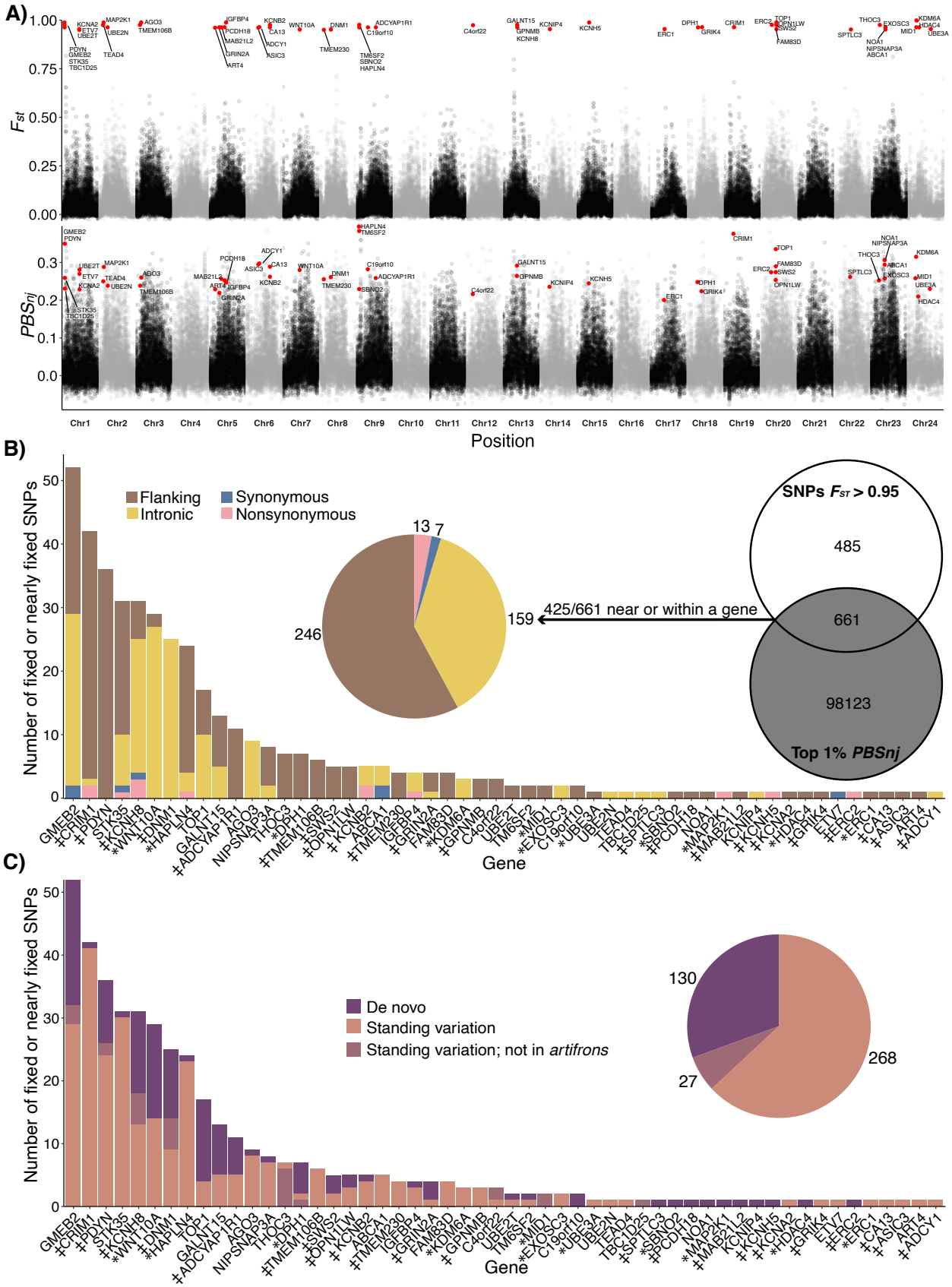

A)
F_{st}
$PBSnj$
Chr1 Chr2 Chr3 Chr4 Chr5 Chr6 Chr7 Chr8 Chr9 Chr10 Chr11 Chr12 Chr13 Chr14 Chr15 Chr16 Chr17 Chr18 Chr19 Chr20 Chr21 Chr22 Chr23 Chr24
Position
B)
Number of fixed or nearly fixed SNPs
Flanking
Intronic
Synonymous
Nonsynonymous
13
7
246
159
425/661 near or within a gene
SNPs F_{ST} > 0.95
485
661
98123
Top 1% PBSnj
Gene
C)
Number of fixed or nearly fixed SNPs
De novo
Standing variation
Standing variation; not in artifrons
130
27
268
Gene

Figure 2. Architecture and source of adaptive genetic variation in the zooplanktivore. (A, top row) F_{st} between *C. simus* and *C. beltrani* + *C. labiosus* calculated in non-overlapping 10-kb windows. F_{st} values for candidate SNPs (red circles) were calculated separately per locus to show the location of fixed variants. (A, lower row) Population branch statistics (PBS_{nj}) with *C. simus* as the target population versus *C. beltrani*, *C. labiosus*, and *C. artifrons* in 10-kb non-overlapping windows. Candidate genes within 20 kb of fixed or nearly-fixed (Fst > 0.95) SNPs and within PBS_{nj} 1% outlier windows are highlighted in red and labeled. Alternating colors represent different chromosomes. (B) Candidate adaptive variants near or within a gene were defined as SNPs that were fixed (n = 17) or nearly fixed (n = 1,129, F_{st} > 0.95) and occurred within a 1% population branch statistic (PBS_{nj}) outlier window for *C. simus* (Venn diagram). Pie chart and bar graphs indicate the proportion of these 425 adaptive variants that were found within the flanking, intronic, or coding regions. (C) Pie chart and bar graphs indicate the proportion of these 425 variants that were detected only within Chichancanab (de novo), as standing genetic variation in coastal sister species *C. artifrons*, or in other Caribbean species, but not *C. artifrons*. * Indicates genes with a craniofacial annotation; ‡ indicates genes with neural or sensory annotation.

208 *Limited adaptive introgression within Chichancanab pupfishes*

209 To test for introgression into the Chichancanab radiation, we used *f*-statistics calculated in
210 *Dsuite* (43) to examine gene flow from 11 additional *Cyprinodon* species from across the clade
211 and the sister species to the genus *Megupsilon aporus*. Within Yucatán pupfishes, we found
212 evidence of introgression between *C. artifrons* and both *C. beltrani* and *C. labiosus*, with the
213 strongest signal in *C. beltrani* (Figure 3A, B). We also found evidence for introgression between
214 *C. beltrani* and most outgroup species examined, including the sister species to *Cyprinodon*,
215 *Megupsilon aporus*. Relative to *C. beltrani*, *C. labiosus* and *C. simus* also showed evidence of
216 introgression with pupfishes endemic to the Chihuahua desert (*C. eximius* and *C. fontinalis*;
217 Figure 3A, B). Furthermore, we also found evidence of introgression between the two sympatric
218 radiations (*C. variegatus* from SSI) and *C. simus* (Figure 3A, B). However, additional *f*-branch
219 statistics indicate that all patterns of excess allele sharing between these outgroup species besides
220 *C. artifrons* and the Chichancanab species flock likely result from ancient introgression events
221 arising before the colonization of Laguna Chichancanab (Figure 3C; S5).

222 To test for adaptive introgression into *C. simus*, we next examined whether the set of
223 adaptive variants (Figure 2) occurred within outlier introgression windows using sliding window
224 scans of distance fraction (d_f; Pfeifer & Kapan, 2019) for species that had significant genome-
225 wide introgression into *C. simus* (Figure 3A , B). We found that none of the outlier introgression
226 windows overlapped with *C. simus* adaptive variants (Figure 3D).

227 Next, we performed *df* scans for introgression between *C. artifrons* and Chichancanab
228 species. Although there were outlier introgression regions between *C. simus* and *C. artifrons*
229 (positive values d_f), these regions did not overlap with *C. simus* adaptive variants (Figure 3D).
230 However, the position of several of these *C. simus* adaptive variants was found in outlier

231 introgression windows in *C. labiosus* and *C. beltrani* (Figure 3D). The number of outlier
232 introgression windows was much lower in *C. simus* relative to *C. labiosus* and *C. beltrani*
233 (Figure 3E, F). Furthermore, both *C. labiosus* and *C. beltrani* genomes contained several
234 introgression tracts that were much longer than *C. simus* tract lengths, indicative of much more
235 recent introgression events in these species (Figure 3E, F).

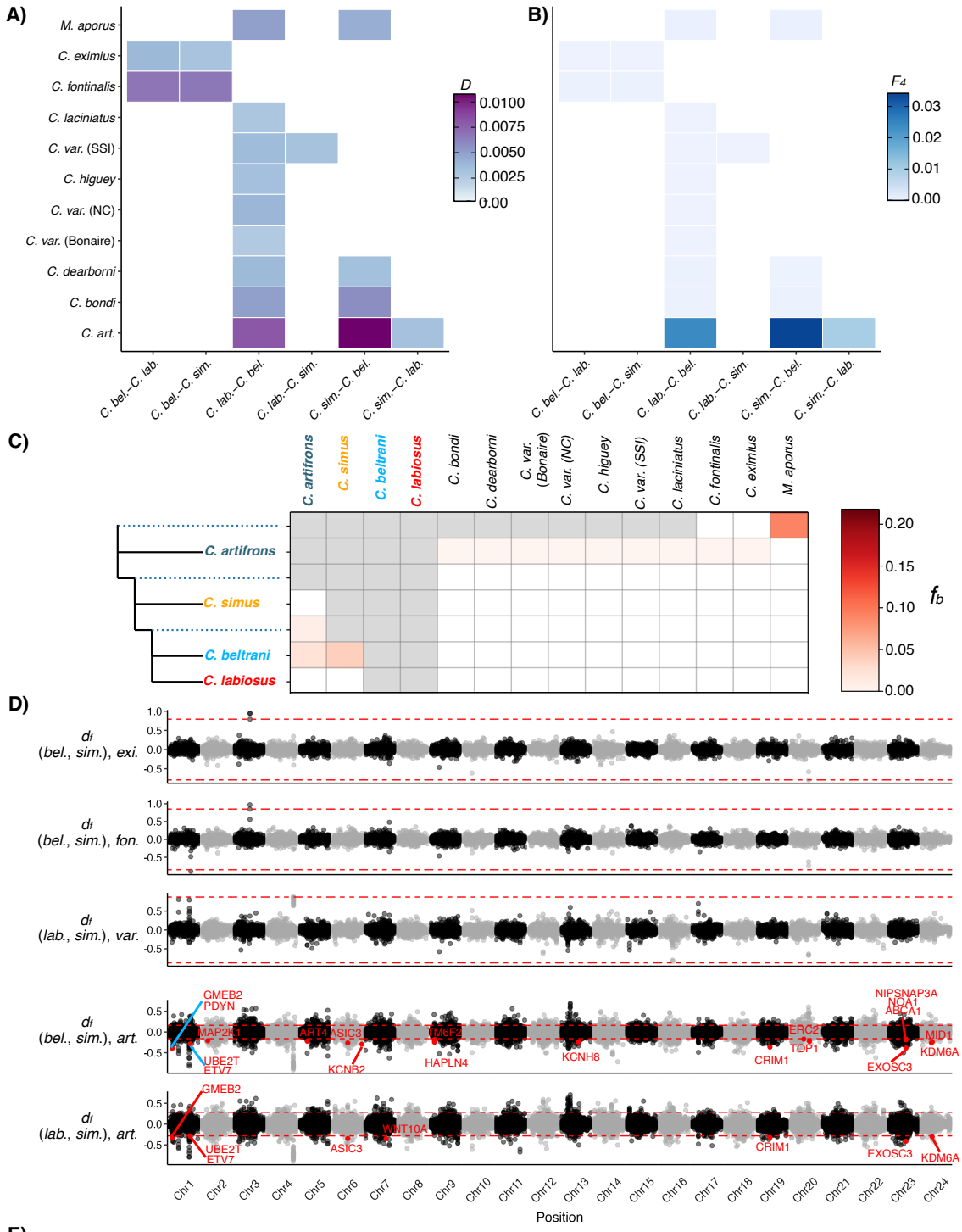
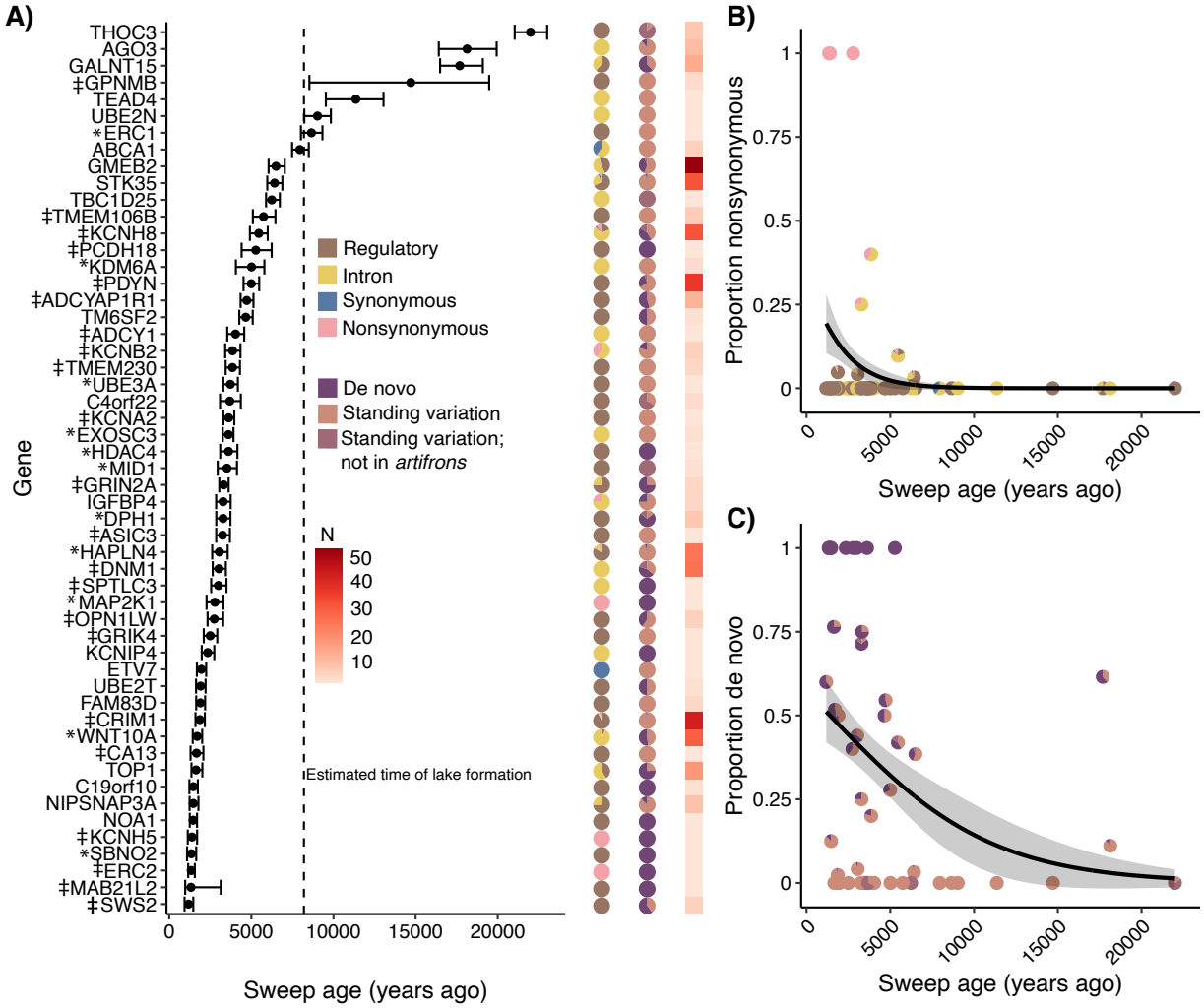

A)
M. aporus
C. eximius
C. fontinalis
C. laciniatus
C. var. (SSI)
C. higuey
C. var. (NC)
C. var. (Bonaire)
C. dearborni
C. bondi
C. art.
C. bel.–C. lab.
C. bel.–C. sim.
C. lab.–C. bel.
C. lab.–C. sim.
C. sim.–C. bel.
C. sim.–C. lab.
D
0.0100
0.0075
0.0050
0.0025
0.00
B)
F4
0.03
0.02
0.01
0.00
C)
C. artifrons
C. simus
C. beltrani
C. labiosus
C. bondi
C. dearborni
C. var. (Bonaire)
C. var. (NC)
C. higuey
C. var. (SSI)
C. laciniatus
C. fontinalis
C. eximius
M. aporus
fb
0.20
0.15
0.10
0.05
0.00
D)
df (bel., sim.), exi.
df (bel., sim.), fon.
df (lab., sim.), var.
df (bel., sim.), art.
df (lab., sim.), art.
GMEB2
PDYN
MAP2K1
UBE2T
ETV7
ART4
ASIC3
KCNB2
TM6F2
HAPLN4
KCNH8
ERC2
TOP1
CRIM1
NIPSNAP3A
NOA1
ABCA1
MID1
KDM6A
EXOSC3
WNT10A
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chr10
Chr11
Chr12
Chr13
Chr14
Chr15
Chr16
Chr17
Chr18
Chr19
Chr20
Chr21
Chr22
Chr23
Chr24
Position
E)
beltrani
simus
Count
MAP2K1
ETV7, UBE2T7
NIPSNAP3A, NOA1, ABCA1
KCNH8
KDM6A
KCNB2
ART4
ABCA1
MID1
CRIM1
ASIC3
GMEB2, PYDN
TOP1
EXOSC3
HAPLN4
HAPLN4
ERC2
Size of introgression block (kb) with artifrons
F)
labiosus
simus
Count
ETV7, UBE2T7
ASIC3
CRIM1
WNT10A
KDM6A
GMEB2
EXOSC3
Size of introgression block (kb) with artifrons

Figure 3. Limited adaptive introgression into *C. simus*. Heatmaps of (A) Patterson's *D*-statistics and (B) F_4 ratios for P1-P2 species combinations within the Chichancanab radiation along the x-axis and P3 focal species along the y-axis. All four-taxon trees used *Cualac tessellatus* as the outgroup (P4). Only significant (*Z*-score > 3) results are colored; combinations not relevant to the Chichancanab radiation are not shown. (C) *f*-branch (f_b) statistics indicate introgression occurring before the Chichancanab radiation. Red boxes indicate significant introgression between P3 species (top row) and the focal species or internal branch (dashed lines) on the left. Non-significant f_b statistics are indicated by white boxes. Grey boxes represent comparisons that cannot be made since f_b cannot be calculated for introgression between sister taxa. See Figure S5 for full *f*-branch graphs. (D) Genome-wide plot of the distance fraction (d_f, an introgression test statistic; Pfeifer & Kapan, 2019) calculated in non-overlapping windows of 91 SNPs (approximately 10 kb) for species trios that showed evidence of introgression with *C. simus*. d_f outlier windows that contained a candidate gene under selection in *C. simus* are colored red and labeled with the gene name. The y-axis label gives the introgression topology: (P1, P2), P3. Positive values indicate regions of introgression between *C. simus* and P3 (from top to bottom: *C. eximius*, *C. fontinalis*, *C. variegatus* (SSI), and *C. artifrons* in the bottom two subplots); negative values indicate regions of introgression between P1 (*C. beltrani* or *C. labiosus*) and P3. Red dashed lines indicate outlier cutoffs, which were determined empirically by the percentage of genome-wide introgression estimated by F_4 ratios in panel B. (E, F) Distribution of the size of d_f outlier introgression tract lengths (in 10 kb windows) between *C. artifrons* and Chichancanab species. The introgression tracts containing candidate adaptive genes in *C. simus* are labeled with gene names. SSI: San Salvador Island, Bahamas; NC: North Carolina, USA.

The stages of adaptive divergence of the zooplanktivore

We used *starTRMCA* (45) to estimate the timing of selective sweeps for the set of candidate adaptive variants near genes within the zooplanktivore *C. simus* (Fig. 4). We estimated that most selective sweeps occurred in a staggered sequence and not simultaneously after colonization of Laguna Chichancanab (Figure 5A)(34). There were also a few older sweeps that we estimate may predate the age of Laguna Chichancanab and divergence from the coastal sister species *C. artifrons* (Figure 1B). Overall, we observed no clear temporal stages of adaptive divergence related to craniofacial genes (asterisks: Figure 5) or neural/sensory genes (‡: Figure 5); instead, we found that selection on these two broad functional categories occurred throughout the period of divergence of *C. simus* following the colonization of Laguna Chichancanab.

Next, we used generalized linear regressions to test if sweep age was correlated with the proportion of nonsynonymous adaptive variants, de novo adaptive variants, or the overall number of adaptive variants per gene region. We found that both nonsynonymous adaptive variants ($\chi^2 = 5.852$, $p = 0.016$) and de novo mutations ($\chi^2 = 6.221$, $p = 0.013$) were significantly more likely to occur among the more recent selective sweeps during the final stage of adaptive divergence in *C. simus* relative to other members of the species flock (Figure 5B, C). However, nonsynonymous mutations were not significantly more likely to be de novo mutations (All adaptive variants: $\chi^2 = 0.102$, $P = 0.749$; singleton adaptive variants: $\chi^2 = 2.010$, $p = 0.156$), and there was no correlation between estimated sweep age and the number of candidate adaptive variants per gene region ($\chi^2 = 0.111$, $p = 0.740$).



291

292 **Figure 4. Stages of adaptive divergence in the zooplanktivore.** (A) Points indicate median
293 sweep age with 95% credible intervals. The dashed line indicates the approximate age of Laguna
294 Chichancanab (34). For each sweep, the first column of pie charts shows the proportions of
295 flanking/intronic/synonymous/nonsynonymous, second column of pie charts shows the
296 proportions of de novo/standing variation, and third column of squares shows the number of
297 candidate adaptive variants per gene (from Fig. 2). For genes with multiple adaptive variants, we
298 used the variant with the median position as the focal position. * Indicates genes with a
299 craniofacial annotation; ‡ indicates genes with neural or sensory annotation. (B) The proportion
300 of nonsynonymous adaptive variants is significantly larger in more recent selective sweeps,
301 indicated by the best-fit quasibinomial regression line; each pie chart represents a single
302 candidate gene. (C) The proportion of de novo variants is significantly larger in more recent
303 selective sweeps; each pie chart represents a single candidate gene.

304

305

306 *Increased visual acuity in the zooplanktivore is consistent with vision-related candidate adaptive*
307 *genes*

308 We identified many *C. simus* adaptive variants associated with eye development (*ASIC3*,
309 *CRIM1*, *MAB21L2*) and opsin genes (*OPN1LW*, *SWS2*). Because the trophic specialization of *C.*
310 *simus* requires visually locating zooplankton in the open water before targeted suction-feeding
311 strikes, we hypothesized that this species would have higher visual acuity than the detritivore, *C.*
312 *beltrani*. We conducted an optomotor response experiment on lab-reared individuals from both
313 species (F1 *C. beltrani* and long-term laboratory colony of *C. simus* (F10+)) and found that *C.*
314 *simus* has a much stronger optomotor response than *C. beltrani* in response to black-and-white
315 vertical bars spinning at a constant rate of ~89.4 rotations per minute (Figure 5A, B; Wilcoxon
316 signed-rank test: $W = 1$, $p = 0.009$). There was no significant difference in rotations per minute
317 between the species for the same individuals during control observation periods of no spinning
318 (Figure S6A; $W = 14$, $p = 0.3105$) or during a control treatment with an all-black background
319 spinning at the same frequency (Figure S6B; $W = 9$, $p = 0.147$).

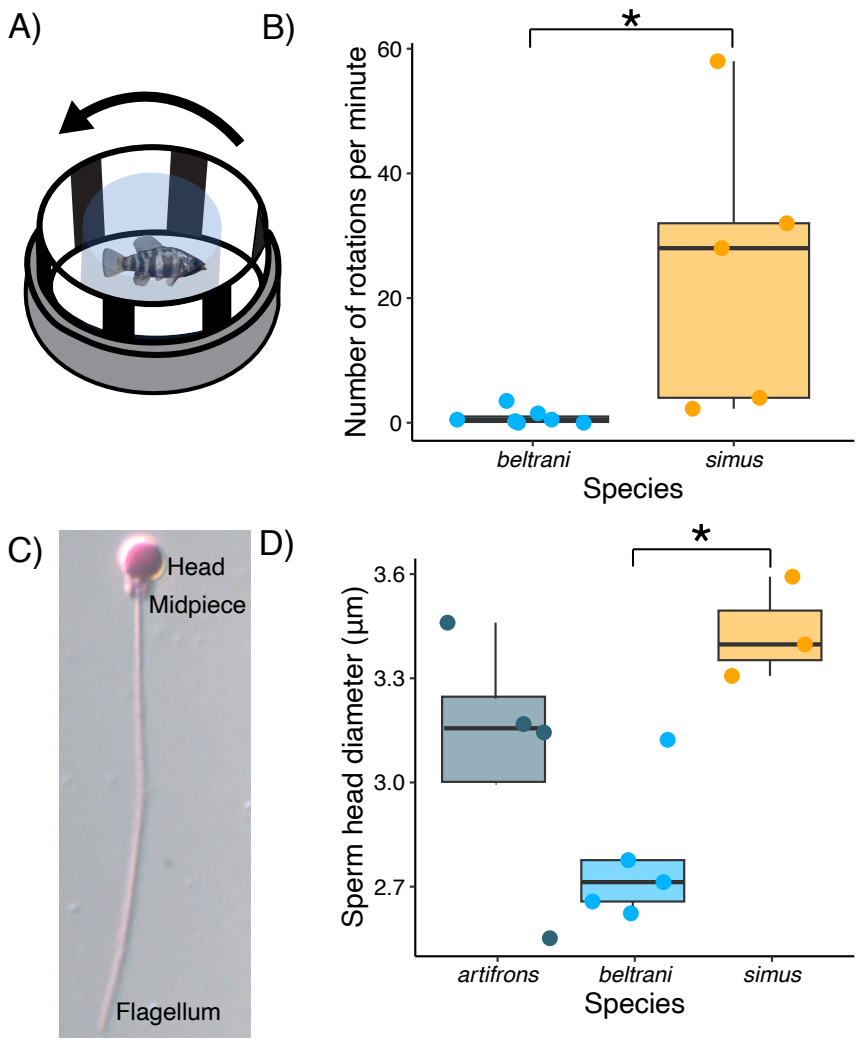


Figure 5. Zooplanktivore *C. simus* has significantly greater visual acuity and divergent sperm morphology relative to the generalist *C. beltrani*. (A) Diagram of the optomotor response experimental apparatus (adapted from Neuhauss et al., 1999). (B) Boxplot and jittered points of the number of complete rotations fish swam per minute, indicative of their optomotor response to black-and-white vertical bars spinning at a constant rate. (C) Representative labeled image of a single *C. beltrani* sperm at 100x stained with rose bengal. (D) Boxplots with jittered points of sperm head diameter for *C. artifrons, beltrani,* and *simus.* Significant Tukey's post hoc differences are shown with an asterisk.

Divergent sperm morphology of the zooplanktivore is consistent with spermatogenesis-related candidate adaptive genes

Four candidate genes (*C4orf22*, *STK35*, *ERC2*, and *TBC1D25*) showed annotations relating to spermatogenesis/fertility. We therefore examined whether three Yucatan species, raised in a common garden laboratory environment (F1 *C. artifrons,* F1 *beltrani,* and a long-term laboratory colony of C. *simus* F10+), differed in sperm morphology. We found that species significantly differed in sperm head length (Figure 5D; $F = 5.724$, $p = 0.025$), but not sperm midpiece length (Figure S7A; $F = 2.403$, $p = 0.147$) or flagellum length (Figure S7B; $F = 0.414$, $p = 0.673$). Tukey's post hoc pairwise comparisons indicated that *C. simus* had significantly larger sperm heads than *C. beltrani* ($q = 3.367$; $p = 0.0207$; Fig. 5D).

Discussion

Here, we investigated the genetic architecture, source of adaptive variation, and temporal dynamics of selection underlying speciation and trophic specialization in a young radiation of *Cyprinodon* pupfishes endemic to Laguna Chichancanab, Mexico. By comparing these patterns to our previous work on a recent radiation of *Cyprinodon* pupfishes in SSI, Bahamas (26, 28–32), we can examine the parallel and non-parallel features of these independent radiations in similar environments. Despite replicate radiations of Caribbean pupfishes in predator-free, large, isolated saline lakes with few competitors, we found little parallelism between the radiations. Although craniofacial morphology evolved rapidly within trophic specialists in both radiations, we found almost zero overlap in the adaptive genes underlying trophic specialization, potentially reflecting the divergent specialists within each radiation. There was no behavior-first stage of adaptive divergence in Chichancanab species and gene categories potentially relevant to speciation were not shared, such as pigmentation genes in SSI versus opsin and spermatogenesis

genes in Chichancanab. Unlike the SSI radiation and most other well-studied adaptive radiations (e.g., Malawi cichlids (20), Galápagos finches (47)), adaptive introgression likely played a minimal role in the Chichancanab radiation. Instead, over 30% of the candidate adaptive variants in the zooplanktivore, *C. simus*, likely arose from de novo mutations compared to 2% in the SSI scale-eater specialist. The one parallel genetic feature was a clear final stage of refinement in the adaptive divergence of the zooplanktivore, marked by increased rates of de novo and nonsynonymous adaptive mutations occurring later in the radiation.

Adaptive introgression did not contribute to adaptive radiation of Chichancanab pupfishes

Genomic analyses have shown that hybridization and introgression played a major role in many adaptive radiations (3, 3, 16, 19–27, 48, 49), including the *Cyprinodon* SSI radiation (26, 31). However, in the Chichancanab radiation, we find little evidence for introgression playing a major role, and did not detect any evidence of adaptive introgression following the colonization of Laguna Chichancanab. Although our analyses indicate extensive introgression among Cyprinodontidae lineages, we did not find any introgression events into the root of the radiation, nor has there been gene flow from species other than the sister-species to the radiation, *C. artifrons*. This secondary gene flow with *C. artifrons* is mostly between *C. beltrani* and, to a lesser extent, *C. labiosus.* Gene flow may be higher with these species because *C. artifrons,* a benthic feeding detritivore, is less ecologically divergent from both *C. beltrani,* a benthic feeding detritivore, and *C. labiosus*, a benthic feeding bivalve specialist. This secondary gene flow between *C. artifrons* and *C. beltrani* and *C. labiosus* overlaps with many genes under selection in *C. simus*, suggesting that secondary gene flow is potentially helping maintain divergence

379 between *C. simus* and the benthic Chichancanab species not by bringing in new adaptive
380 variation for trophic specialization, but by maintaining the generalist/omnivore species.
381 These results are remarkably different from the *Cyprinodon* adaptive radiation of SSI,
382 where there is strong evidence of introgression at the root of the radiation, evidence of ongoing
383 introgression, and the presence of many introgressed adaptive variants in both trophic specialists
384 (26, 31). Laguna Chichancanab is a landlocked lake with no above-ground river connections and
385 is geographically more isolated compared to SSI, a small island in the approximate center of the
386 Caribbean. Furthermore, the sister species, *C. artifrons*, is abundant along the entire coast of the
387 Yucatán, meaning that any gene flow into Chichancanab would likely first have to pass through
388 *C. artifrons.* An additional, non-mutually exclusive explanation could be that the water chemistry
389 of Laguna Chichancanab, with near-saturation of calcite and gypsum, is a greater barrier to non-
390 adapted pupfish species than the hypersaline lakes of SSI. Thus, geography and ecology may
391 play an important role in determining the relative role that introgression may play in radiations.
392
393 *De novo mutations play the dominant role in driving adaptive divergence of the zooplanktivore*
394 The importance of de novo mutations in adaptive radiations remains contentious, in part
395 due to the time needed for de novo beneficial mutations to arise (11, 15). Interestingly, we find
396 that over 30% of the zooplanktivore adaptive variants were detected only within Chichancanab
397 species, suggesting their de novo origins within the basin, approximately 10 times greater than
398 the role of de novo mutations in the adaptive divergence of the scale-eater in the SSI radiation. A
399 primary factor could be that the estimated effective population size for the Chichancanab
400 radiation is approximately twice that of the SSI radiation (26), meaning that selection on de novo
401 mutations would be more effective, with a higher probability of fixation, all else being equal

402 (14). There may also be a lack of existing standing genetic variation relevant to the trophic
403 specialization of zooplanktivory. Although radiations should proceed along the genetic lines of
404 least resistance (50), the morphological axis along which the zooplanktivore has diverged is
405 orthogonal to almost all morphological variation that exists in Caribbean pupfish species (29).
406 Evolving along the line of greatest resistance instead of the line of least resistance may require de
407 novo mutations. While the SSI scale-eater is also highly divergent, it is closer to the existing
408 morphological space of Caribbean pupfish species and in the direction of greatest variance, at
409 least within the two phenotypic dimensions of greatest variance (29). Finally, an alternative
410 possibility could be that many of these de novo mutations do exist as standing variation and were
411 missed in our dataset, but this is unlikely to explain all de novo mutations, given our extensive
412 sampling of *Cyprinodon* species across the clade and spanning the Caribbean. Thus, our results
413 highlight that de novo mutations may play a more important role in driving young adaptive
414 radiations than currently appreciated (12, 13), especially when ecotypes are highly divergent
415 from other species within the radiation.
416
417 *Extremely limited overlap in the genes underlying adaptive radiations of trophic specialists*
418 We found only two candidate genes (*DNM1* and *SPTLC3*) among the 55 zooplanktivore
419 candidate genes in the Chichancanab radiation that overlapped with the hundreds of candidate
420 genes identified so far in the SSI radiation (26, 28, 30, 32, 51–55). Consistent with both trophic
421 specializations requiring the pursuit of mobile prey, both genes are involved in motor neuron
422 development. *DNM1* is a dynamin with nervous tissue-specific expression that is crucial for the
423 proper formation of motor neuron axons (41). *SPTLC3* is involved in the biosynthesis of

424 sphingolipids (56), and suppression of *SPTLC3* in zebrafish results in motor neuron axon defects
425 (42).
426 The limited number of shared adaptive genes is somewhat surprising given the high
427 conservation of gene function in both morphological development (57, 58) and behavior (59),
428 and the frequent reuse of the same candidate genes among vertebrate adaptive radiations,
429 particularly for craniofacial genes. For example, *BMP4* is involved in diversification in both
430 Galápagos finches (60) and African cichlids (61, 62). Further, within the SSI radiation, many of
431 the same genes are involved in craniofacial divergence between divergent trophic specialists (26,
432 53). Selection may still be acting on the same genetic pathways and genetic regulatory networks
433 missed by our direct candidate comparisons.
434 Another potential reason for this non-parallelism could be that trophic specialists in each
435 radiation have diverged in different sets of traits (29). Specifically, the Chichancanab radiation
436 has diversified in tooth length over 130 times faster than background rates, driven by the
437 extremely short teeth of the zooplanktivore (29). Consistent with divergence in tooth
438 morphology, one of the candidate craniofacial genes with many de novo mutations in flanking
439 and intronic regions was *WNT10A*. Mutations in *WNT10A* are associated with tooth
440 developmental disorders in humans (63). Furthermore, knockout and overexpression experiments
441 in sticklebacks and zebrafish have demonstrated *WNT10A*'s role in teeth development and
442 regeneration (64, 65).
443
444 *The parallel role of vision but the non-parallelism of adaptive candidate genes*
445 One of the more striking features of both *Cyprinodon* radiations is the evolution of novel
446 trophic specialists from a generalist ancestor: the scale-eater (*C. desquamator*) in the SSI

447 radiation and the zooplanktivore (*C. simus*) in the Chichancanab radiation. Although these are
448 distinct dietary specializations, they share a similarity in that they require tracking free-
449 swimming prey. Indeed, four of the top 15 enriched GO terms of candidate genes in the SSI
450 scale-eater were related to eye development (26). Similarly, several of the candidate genes we
451 identified in the zooplanktivore were related to eye development (*ASIC3*, *MAB21L2*, and
452 *CRIM1*). In zebrafish, a two-basepair deletion of *CRIM1* results in misshapen lenses as well as a
453 large eye-to-head ratio (66). Interestingly, one of the defining morphological features of *C. simus*
454 is its large eye-to-head ratio (33) (Figure 1A). *CRIM1* is also a promising candidate because it is
455 one of the few candidate genes with multiple nonsynonymous mutations (Figure 2B).

456 Consistent with candidate genes related to eye development and trophic specialization,
457 *C. simus* does indeed exhibit a stronger optomotor response compared to *C. beltrani*, indicating
458 that this specialist has greatly improved visual acuity to detect rapidly rotating black-and-white
459 bars.

460 Although both radiations contained many vision-related candidate genes, there was zero
461 overlap in these genes. Most notably, we found several adaptive variants associated with opsins
462 (*SWS2*, blue-sensitive, and *OPN1LW*, red-sensitive) in the zooplanktivore. However, in the SSI
463 radiation, opsins do not appear to be under selection in any species (26). This result is somewhat
464 surprising given that opsins appear to be evolutionarily labile and rapidly evolve in other fish
465 lineages (e.g., sticklebacks (67), cichlids (68, 69), wrasses (70), reviewed in (71)). One plausible
466 explanation could be that SSI lake environments are relatively shallow, so differences in depth—
467 a common selective pressure on opsin evolution (71)—are likely not a strong selective force on
468 SSI. Alternatively, the main selective force acting on opsins may be prey capture efficiency. In
469 many fish species, zooplanktivory is associated with higher expression of shorter wavelength

470 opsins (e.g., *SWS1* (71)), whereas piscivory is not often associated with changes in opsin
471 expression (68, 69). Interestingly, both opsins contain many candidate de novo mutations, and
472 *SWS2* is the youngest estimated sweep in the zooplanktivore. One explanation could be that most
473 Caribbean *Cyprinodon* species are benthic and live in similar habitats, resulting in limited
474 standing genetic variation in opsins.
475
476 *Non-parallelism of sexual selection: the relative role of pre- and post-mating sexual selection*
477 Although the role that premating sexual selection plays in speciation has historically been
478 the main focus (72, 73), there is both theoretical and empirical support for the role of postmating
479 sexual selection in maintaining reproductive isolation via postmating prezygotic reproductive
480 barriers (74–79). In the SSI radiation, species differ in pigmentation, and a handful of the
481 candidate adaptive variants fall in regulatory regions of genes associated with pigmentation (26).
482 Furthermore, there is evidence for premating isolation by female preference in the scale-eater
483 (80). However, in Chichancanab, *C. labiosus* does not visually distinguish between different
484 species in the radiation (81). Unlike the SSI radiation, we did not find any candidate adaptive
485 variants associated with pigmentation; instead, we identified adaptive variants associated with
486 spermatogenesis (*ERC2*, *TBC1D25*, *C4orf22*, *STK35*). Consistent with previous work
487 demonstrating the importance of these genes in spermatogenesis (82–85), we find that *C. simus*
488 does indeed have larger sperm heads than *C. beltrani*, suggesting a potential role for post-mating
489 prezygotic reproductive barriers.
490
491 *Parallelism in the final stages of adaptation: refinement with de novo and nonsynonymous*
492 *mutations*

Adaptation may occur in stages, with less pleiotropic cis-regulatory changes occurring before more pleiotropic coding changes (26, 86, 87). Selection is also predicted to act on standing variation first due to the waiting time for de novo mutations to arise within a population (14, 88). In both radiations (Richards et al., 2021), we observe strong evidence for these stages of adaptation.

The simplest explanation for why we see parallelism in de novo mutations sweeping later could be due to the waiting time required for a beneficial mutation to (1) arise and (2) reach a substantial frequency within the population (11, 14). Although de novo and nonsynonymous mutations were not statistically associated across all candidate variants, many of the late sweeping adaptive variants are both de novo and nonsynonymous. These variants may also sweep late because nonsynonymous mutations are often pleiotropic and deleterious in an ancestral genetic background (14, 89). Thus, they may not be advantageous until later in the adaptive walk, where compensatory epistasis may mediate negative pleiotropic effects (90–94). Our results highlight that this final refinement stage is likely a general pattern to be expected across adaptive radiations.

Methods

Sampling

We collected Laguna Chichancanab specimens in 2022 over two days using a 5 x 1.3 m seine net with 1.6 mm mesh size. We sampled two sites, an entrance road by La Presumida, and a bridge crossing the narrow section of the lake on the road to San Diego. *C. labiosus* and *C. beltrani* were sampled from both sites for this study, whereas *C. simus* was only collected at La Presumida. *C. artifrons* were collected by cast net from the coastal estuary at Sisal. Additional

specimens were collected from the Bahamas (Lake Cunningham, New Providence Island and Crescent Pond, San Salvador Island), the Dominican Republic (Laguna Bavaro), or Fort Fisher, North Carolina as described previously (Martin 2016) or sourced from the American Killifish Association, the London Zoological Society, the Dallas Children's Aquarium, and Frans Vermuelen. All specimens were euthanized in an overdose of buffered MS-222 (Fiquel, Inc.) following approved animal care and use protocols from the University of California, Berkeley and the University of California, Davis.

Sequencing, Variant calling and filtering

Individual DNA samples were resequenced using Illumina Hiseq 4000 and Novaseq. Raw reads were trimmed with *fastp* (v0.23.4) (95). Reads were mapped to the UCB_CyNevMio_1.0 *Cyprindon nevadensis mionectes* reference genome (GCA_030533445.1) with *bwa mem* (v0.7.17) (96). Duplicate reads in the bam files were marked with *Picard MarkDuplicates* (GATK v4.5.0.0) (118). Coverage and mapping quality were assessed with *qualimap* (v2.2.2) (97). We followed the *GATK* (v.4.5.0.0) genotyping pipeline to call variants (119). We used *HaplotypeCaller* (-ERC GVCF) v4.5.0.0 to call variants for each individual and stored variants in the GenomicsDB datastore format prior to using *GenotypeGVCFs* to perform joint genotyping. We restricted our analyses to biallelic SNPs and applied the recommended *GATK* hard filters (QD < 2, QUAL < 30, SOR > 3, FS > 60, MQ < 40, MQRankSum < -12.5, ReadPosRankSum < -8) to filter SNPs (118, 120). Missing genotypes were reset to ./. from 0/0 due to how *GATK* represents missing genotypes in version 4.5.0.0 (101).

538 After hard-filtering and variant calling, we filtered the data set further by removing SNPs
539 with minor allele frequency < 0.05, more than 10% missing, depth less than 10x, and GQ less
540 than 20 using *BCFtools* (v1.16)(102). This resulted in a final data set that included 23,960,536
541 SNPs.
542
543 *Population structure of Yucatán Cyprinodon pupfishes*
544 To assess population structure, we first pruned SNPs in linkage disequilibrium using
545 *PLINK* (v. 1.9) with the following parameters: "-indep-pairwise 50 5 0.2" (103). This filtered our
546 data set from 23,960,536 SNPs to 5,476,855 SNPs. With this linkage disequilibrium pruned data
547 set, we first filtered the full VCF file to only include relevant Yucatán *Cyprinodon* species (i.e.,
548 *C. artifrons, C. beltrani, C. labiosus,* and *C. simus*), then conducted a PCA using *PLINK*. Next,
549 we used *ADMIXTURE* (v1.3) to determine the optimal number of population clusters that best fit
550 the data and to assign individuals to their corresponding population clusters (104). We performed
551 this analysis and calculated cross-validation error for the number of clusters K = 1-6. Models in
552 ADMIXTURE with K=2-4 were equally supported (Figure S2A). However, given morphological
553 distinctness and the results of the PCA, we present the results for when K = 4 (Figure 1D).
554
555 *Demographic history of Yucatán Cyprinodon pupfish*
556 We estimated demographic history using *MSMC2* (105). For each Yucatán pupfish
557 species, we selected three individuals with the highest mean depth and generated single-sample
558 VCF files, individual mask files, and "mapability" mask files. We then ran *MSMC2* with default
559 parameters except we set phasing to "unphased" and "P_PAR=8*1+25*1+1*2+1*3." For

plotting, we used a generation time of 1 year and a mutation rate of 1.56 x 10^{-8} substitutions per base pair, estimated from Caribbean *Cyprinodon* species (26).

Identifying candidate genes

We calculated F_{st} genome-wide, per site, and in non-overlapping 10-kb windows using the "--weir-fist-pop" function in *VCFtools*(v0.1.16) for all pairwise combinations of Yucatán *Cyprinodon* species (106). We chose 10-kb windows because linkage disequilibrium was substantially decayed at this distance (Figure S3), and these windows allowed us to quantify finer-scale genomic variation. With the 10-kb windowed F_{st} values, we calculated a modified population branch statistic (PBS) for four taxa, PBS_{nj} (107). This statistic is useful because it makes no assumptions about species topology or polarization (107). We then normalized PBS_{nj} by the total tree length, as normalized *PBS* is most effective and specific at identifying selective sweeps (40). We classified SNPs as candidates if they were fixed or nearly fixed (per-site F_{st} >0.95) compared to other species within the radiation and were among the top 1% outliers out of all PBS_{nj} windows, following the threshold used in (40).

Introgression

Using the full VCF file, which contained a wide range of species in the family Cyprinodontidae, we tested for evidence of introgression in Yucatán *Cyprinodon* species using *Dsuite* (43). We first excluded any samples/species that had an average depth < 4. We then calculated all possible trios with a species tree based on the most recent phylogeny (108) with *Cualac tessellatus* as the outgroup. We calculated Z-scores to assess the significance of *D* statistics using 1,000 jackknife blocks and considered a Z-score greater than three as evidence

583 for significant introgression (109). We only focused our primary analysis on species trios where
584 P1-P2 included species within the Laguna Chichancanab radiation. We used the other trio
585 calculations to calculate *f*-branch (f_b) statistics to account for correlated f_4-ratio scores and to
586 inform the timing of introgression events (43).

587 For species trios that showed evidence of introgression where *C. simus* was either P1or
588 P2, we then calculated the distance fraction (d_f), a robust metric of introgression for windowed
589 analyses (44), in non-overlapping windows of 91 informative SNPs. We chose this number of
590 SNPs because it resulted in windows of approximately 10 kb. We considered windows to be
591 introgressed if they were in the top X% of absolute d_f value, where X% was the percentage of
592 estimated genome wide introgression (e.g, f_4-ratio) (110, 111). To estimate introgression block
593 sizes, we merged neighboring windows that were both introgression outliers and took the total
594 size of the blocks.

595

596 *Characterizing candidate genes*

597 After identifying candidate SNPs, we then used *Bedtools* (v2.31) to see if SNPs occurred
598 near (20-kb) or within a gene (112) with a GFF annotation file for the *C. nev. mionectes*
599 reference genome (Tian et al. in prep). We classified SNPs within 20 kb of a gene (upstream or
600 downstream) as flanking. For SNPs within a gene, we further classified them into intronic,
601 synonymous, or nonsynonymous using *SnpEff* (113).

602 To characterize the origin of candidate SNPs, we extracted the alleles from every species
603 for each candidate SNP. We classified variants that were found in any other species as "standing
604 variation". We further subdivided this group into variants that were found in other species but not
605 in *C. artifrons*. If "standing variation" variants fell within an introgression window, we classified

606 them as “introgressed.” Finally, if the alternative variant was only found in Chichancanab
607 species, we classified it as de novo.
608
609 *Timing of selective sweeps*
610 To gain insight into the timing of selective sweeps in *C. simus*, we used *starTRMCA* (45).
611 Because many genes contained multiple candidate SNPs, for this analysis, we selected the
612 variant that was in the median position of each gene. We then extracted a 1-Mb region
613 surrounding that variant and removed all sites that contained missing data. For simplification, we
614 only used the generalist/detritivore, *C. beltrani*, as the “reference” population and the
615 zooplanktivore, *C. simus*, as the “selected” population. We used a fixed recombination rate of
616 $2x10^{-8}$ (swordtail fishes; (114)) and a mutation rate of $1.56x10^{-8}$ substitutions per base pair
617 (Caribbean *Cyprinodon* species; 26). We then ran five separate Markov chains for 30,000
618 iterations with a proposal standard deviation of 150 (preliminary analyses changing this
619 parameter had minimal influence on results). We then discarded the first 10,000 iterations of each
620 chain as burn-in.
621 To test for stages of adaptation, we conducted separate generalized linear regressions to
622 test if the sweep age was correlated with the proportion of nonsynonymous SNPs, proportion of
623 de novo SNPs, and the number of SNPs. For the generalized linear regressions of proportion of
624 nonsynonymous SNPs and proportion of de novo SNPs, we used a quasi-binomial family to
625 account for overdispersion with a logit link function. For the number of SNPs, we used a quasi-
626 Poisson family with a log link function. To test for the significance of sweep age, we used a
627 likelihood ratio test comparing a model with the estimated sweep age as an effect compared to a
628 null model.

Optomotor response

We tested for differences in visual acuity between the generalist/detritivore, *C. beltrani*, and the zooplanktivore, *C. simus*, by conducting an optomotor response behavioral assay. We tested lab-reared *C. beltrani* (F1, $n = 7$) and lab-reared *C. simus* (long-term laboratory colony exceeding ten generations in the lab, $n = 5$). Each fish was tested consecutively with each control or treatment for 1-minute observation periods, with the control observation occurring first, followed by presentation of all-black and black-and-white spinning bars. The order of presentation for the latter to treatments was alternated in each trial. Fish were placed in a suspended cylindrical plastic container 0.3 m in diameter that remained stationary during the course of the trials. Spinning black-and-white or all-black paper was rotated outside the clear plastic container at an RPM of ~90rpm during each observation period.

Due to the data being nonparametric (count, zero-inflated, unequal variance between groups), we conducted a Wilcoxon rank sum test for the on-banded (all black) and on-banded portions separately. Because we had an a priori hypothesis that *C. simus* would display a stronger optomotor response (greater visual acuity), we calculated *p*-values with a one-tailed test.

Sperm morphology

To test for differences in sperm morphology, we collected sperm samples from five individuals per species in the lab (*C. simus, C. artifrons, C. beltrani*). We first anesthetized the fish with a solution of MS-222. After the fish was anesthetized, we collected 1μL of milt (fish semen) with a 1μL glass microcapillary tube and fixed the sample overnight in a 4% PFA solution stained with Rose Bengal. We then plated sperm on a positively charged slide with a

sealed coverslip. We imaged sperm with oil immersion with a 63x objective and measured 10-30 sperm per individual in ImageJ (115). Due to poor sperm samples from some individuals (little sperm), we ended up with the following sample sizes: three *simus*, four *artifrons*, and five *beltrani*.

To retain information about variation within individuals, we fit separate linear mixed-effects models for each trait of interest (sperm head, midpiece, and flagellum lengths) with population as a fixed effect and a random intercept for individual ID, using *LME4* (116). To assess the significance of the model, we used a Type II ANOVA with a Wald *F* test using Kenward-Roger degrees of freedom. If there was a significant effect of population, we conducted post hoc pairwise tests using the *Emmeans* package (117) and corrected for multiple comparisons using the Tukey method. We did not correct for phylogeny due to how closely related the species are, the limited sample size, and the fact that the only significant differences were between sister species.

Acknowledgments

MCK was supported by a Miller Postdoctoral Research Fellowship from the Miller Institute for Basic Research in Science, University of California Berkeley. This research was funded by NSF CAREER 1749764, NIH 5R01DE027052-02, Daphne and Ted Pengelley Award from the Center for Population Biology, and ARCS Foundation awards to CHM. We thank the Gerace Research Center and Troy Day for logistical support in the Bahamas, and the governments of Mexico and the Bahamas BEST commission for permission to collect and export samples (Permit No. PPF/DGOPA-001/20). All research procedures and animal care protocols (AUP-2021-

674 02-14062-1 and AUP-2021-07-14515) were approved by the University of California, Berkeley
675 Animal Care and Use committee and the UC Davis Institutional Animal Care and Use committee.

676

677 **References**

678 1. R. G. Gillespie, *et al.*, Comparing Adaptive Radiations Across Space, Time, and Taxa. *J*
679 *Hered.* **111**, 1–20 (2020).

680 2. C. H. Martin, E. J. Richards, The Paradox Behind the Pattern of Rapid Adaptive
681 Radiation: How Can the Speciation Process Sustain Itself Through an Early Burst? *Annu.*
682 *Rev. Ecol. Evol. Syst.* **50**, 569–593 (2019).

683 3. Y. Naciri, H. P. Linder, The genetics of evolutionary radiations. *Biol. Rev.* **95**, 1055–1072
684 (2020).

685 4. D. Schluter, *The Ecology of Adaptive Radiation* (OUP Oxford, 2000).

686 5. J. T. Stroud, J. B. Losos, Ecological Opportunity and Adaptive Radiation. *Annu. Rev. Ecol.*
687 *Evol. Syst.* **47**, 507–532 (2016).

688 6. R. Butlin, *et al.*, What do we need to know about speciation? *Trends Ecol. Evol.* **27**, 27–39
689 (2012).

690 7. J. A. Coyne, H. A. Orr, *Speciation* (Oxford University Press, 2004).

691 8. D. R. Matute, B. S. Cooper, Comparative studies on speciation: 30 years since Coyne and
692 Orr. *Evolution* **75**, 764–778 (2021).

693 9. D. Schluter, L. H. Rieseberg, Three problems in the genetics of speciation by selection.
694 *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2122153119 (2022).

695 10. D. Berner, W. Salzburger, The genomics of organismal diversification illuminated by
696 adaptive radiations. *Trends Genet.* **31**, 491–499 (2015).

697 11. D. A. Marques, J. I. Meier, O. Seehausen, A Combinatorial View on Speciation and
698 Adaptive Radiation. *Trends Ecol. Evol.* **34**, 531–544 (2019).

699 12. J. B. Pease, D. C. Haak, M. W. Hahn, L. C. Moyle, Phylogenomics Reveals Three Sources
700 of Adaptive Variation during a Rapid Radiation. *PLOS Biol.* **14**, e1002379 (2016).

701 13. B. W. Stone, C. A. Wessinger, Ecological Diversification in an Adaptive Radiation of
702 Plants: The Role of De Novo Mutation and Introgression. *Mol. Biol. Evol.* **41**, msae007
703 (2024).

14. R. D. H. Barrett, D. Schluter, Adaptation from standing genetic variation. *Trends Ecol. Evol.* **23**, 38–44 (2008).

15. M. Rousselle, *et al.*, Is adaptation limited by mutation? A timescale-dependent effect of genetic diversity on the adaptive substitution rate in animals. *PLOS Genet.* **16**, e1008668 (2020).

16. O. Seehausen, Hybridization and adaptive radiation. *Trends Ecol. Evol.* **19**, 198–207 (2004).

17. K. Kagawa, G. Takimoto, O. Seehausen, Transgressive segregation in mating traits drives hybrid speciation. *Evolution* **77**, 1622–1633 (2023).

18. K. Kagawa, G. Takimoto, Hybridization can promote adaptive radiation by means of transgressive segregation. *Ecol. Lett.* **21**, 264–274 (2018).

19. R. De-Kayne, *et al.*, Genomic architecture of adaptive radiation and hybridization in Alpine whitefish. *Nat. Commun.* **13**, 4479 (2022).

20. D. A. Joyce, *et al.*, Repeated colonization and hybridization in Lake Malawi cichlids. *Curr. Biol.* **21**, R108–R109 (2011).

21. S. Lamichhaney, *et al.*, Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* **518**, 371–375 (2015).

22. C. H. Martin, *et al.*, Complex histories of repeated gene flow in Cameroon crater lake cichlids cast doubt on one of the clearest examples of sympatric speciation. *Evolution* **69**, 1406–1422 (2015).

23. J. I. Meier, *et al.*, Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nat. Commun*. **8**, 14363 (2017).

24. C. Pouchon, *et al.*, Phylogenomic Analysis of the Explosive Adaptive Radiation of the Espeletia Complex (Asteraceae) in the Tropical Andes. *Syst. Biol*. **67**, 1041–1060 (2018).

25. E. J. Richards, J. W. Poelstra, C. H. Martin, Don't throw out the sympatric speciation with the crater lake water: fine-scale investigation of introgression provides equivocal support for causal role of secondary gene flow in one of the clearest examples of sympatric speciation. *Evol. Lett.* **2**, 524–540 (2018).

26. E. J. Richards, *et al.*, A vertebrate adaptive radiation is assembled from an ancient and disjunct spatiotemporal landscape. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2011811118 (2021).

27. N. Rosser, *et al.*, Hybrid speciation driven by multilocus introgression of ecological traits. *Nature* **628**, 811–817 (2024).

28. M. E. S. John, J. C. Dunker, E. J. Richards, S. Romero, C. H. Martin, Parallel evolution of integrated craniofacial traits in trophic specialist pupfishes. *Ecol. Evol.* **14**, e11640 (2024).

29. C. H. Martin, P. C. Wainwright, Trophic novelty is linked to exceptional rates of morphological diversification in two adaptive radiations of cyprinodon pupfish. *Evolution* **65**, 2197–2212 (2011).

30. A. H. Patton, E. J. Richards, K. J. Gould, L. K. Buie, C. H. Martin, Hybridization alters the shape of the genotypic fitness landscape, increasing access to novel fitness peaks during adaptive radiation. *eLife* **11**, e72905 (2022).

31. E. J. Richards, C. H. Martin, Adaptive introgression from distant Caribbean islands contributed to the diversification of a microendemic adaptive radiation of trophic specialist pupfishes. *PLOS Genet.* **13**, e1006919 (2017).

32. E. J. Richards, C. H. Martin, We get by with a little help from our friends: shared adaptive variation provides a bridge to novel ecological specialists during adaptive radiation. *Proc. R. Soc. B.* **289**, 20220613 (2022).

33. J. M. Humphries, R. R. Miller, A Remarkable Species Flock of Pupfishes, Genus Cyprinodon, from Yucatán, México. *Copeia* **1981**, 52–64 (1981).

34. D. A. Hodell, J. H. Curtis, M. Brenner, Possible role of climate in the collapse of Classic Maya civilization. *Nature* **375**, 391–394 (1995).

35. U. Strecker, Description of a new species from Laguna Chichancanab, Yucatan, Mexico: Cyprinodon suavium (Pisces: Cyprinodontidae). *Hydrobiologia* **541**, 107–115 (2005).

36. C. A. Gracida-Juárez, J. J. Schmitter-Soto, M. J. Genner, Community structure of indigenous fishes relative to habitat variation and invasive tilapia in lakes of Quintana Roo, Mexico. *Environ. Biol. Fishes* **107**, 401–414 (2024).

37. U. Strecker, The impact of invasive fish on an endemic Cyprinodon species flock (Teleostei) from Laguna Chichancanab, Yucatan, Mexico. *Ecol. Freshw. Fish* **15**, 408–418 (2006).

38. U. Strecker, Genetic differentiation and reproductive isolation in a *Cyprinodon* fish species flock from Laguna Chichancanab, Mexico. *Mol. Phylogenetics Evol.* **39**, 865–872 (2006).

39. U. Strecker, C. G. Meyer, C. Sturmbauer, H. Wilkens, Genetic Divergence and Speciation in an Extremely Young Species Flock in Mexico Formed by the Genus*Cyprinodon*(Cyprinodontidae, Teleostei). *Mol. Phylogenetics Evol.* **6**, 143–149 (1996).

40. M. Shpak, K. N. Lawrence, J. E. Pool, The Precision and Power of Population Branch Statistics in Identifying the Genomic Signatures of Local Adaptation. *Genome Biol. Evol.* **17**, evaf080 (2025).

41. C. Bragato, *et al.*, Zebrafish dnm1a gene plays a role in the formation of axons and synapses in the nervous tissue. *J. Neurosci. Res.* **101**, 1345–1359 (2023).

42. C. Gonzaga-Jauregui, *et al.*, Exome Sequence Analysis Suggests that Genetic Burden Contributes to Phenotypic Variability and Complex Neuropathy. *Cell Reports* **12**, 1169–1183 (2015).

43. M. Malinsky, M. Matschiner, H. Svardal, Dsuite - Fast D-statistics and related admixture evidence from VCF files. *Mol. Ecol. Res.* **21**, 584–595 (2021).

44. B. Pfeifer, D. D. Kapan, Estimates of introgression as a function of pairwise distances. *BMC Bioinform.* **20**, 207 (2019).

45. J. Smith, G. Coop, M. Stephens, J. Novembre, Estimating Time to the Common Ancestor for a Beneficial Allele. *Mol. Biol. Evol*. **35**, 1003–1017 (2018).

46. S. C. F. Neuhauss, *et al.*, Genetic Disorders of Vision Revealed by a Behavioral Screen of 400 Essential Loci in Zebrafish. *J. Neurosci.* **19**, 8603–8615 (1999).

47. E. D. Enbody, *et al.*, Community-wide genome sequencing reveals 30 years of Darwin's finch evolution. *Science* **381**, eadf6218 (2023).

48. J. V. Peñalba, *et al.*, The Role of Hybridization in Species Formation and Persistence. *Cold Spring Harb. Perspect. Biol.* **16**, a041445 (2024).

49. O. Seehausen, Conditions when hybridization might predispose populations for adaptive radiation. *J. Evol. Biol.* **26**, 279–281 (2013).

50. D. Schluter, Adaptive radiation along genetic lines of least resistance. *Evolution* **50**, 1766–1774 (1996).

51. J. A. McGirr, C. H. Martin, Few Fixed Variants between Trophic Specialist Pupfish Species Reveal Candidate Cis-Regulatory Alleles Underlying Rapid Craniofacial Divergence. *Mol. Biol. Evol.* **38**, 405–423 (2021).

52. J. A. McGirr, C. H. Martin, Novel Candidate Genes Underlying Extreme Trophic Specialization in Caribbean Pupfishes. *Mol. Biol. Evol.***34**, 873 (2017).

53. J. A. McGirr, C. H. Martin, Parallel evolution of gene expression between trophic specialists despite divergent genotypes and morphologies. *Evol. Lett.* **2**, 62–75 (2018).

54. J. A. McGirr, C. H. Martin, Ecological divergence in sympatry causes gene misexpression in hybrids. *Mol. Ecol.* **29**, 2707–2721 (2020).

55. M. F. Palominos, V. Muhl, C. H. Martin, Craniofacial-specific transcriptomics uncovers novel genes underlying jaw divergence in dietary specialist pupfishes. *Genetics* iyaf207 (2025).

56. T. Hornemann, *et al.*, The SPTLC3 Subunit of Serine Palmitoyltransferase Generates Short Chain Sphingoid Bases. *J. of Biol. Chem.* **284**, 26322–26330 (2009).

806 57. M. E. Bronner, N. M. LeDouarin, Development and evolution of the neural crest: An
807 overview. *Dev. Biol.* **366**, 2–9 (2012).

808 58. F. Hirth, H. Reichert, Conserved genetic programs in insect and mammalian brain
809 development. *BioEssays* **21**, 677–684 (1999).

810 59. C. J. Reaume, M. B. Sokolowski, Conservation of gene function in behaviour. *Phil. Tran.*
811 *R. Soc. B.* **366**, 2100–2110 (2011).

812 60. A. Abzhanov, M. Protas, B. R. Grant, P. R. Grant, C. J. Tabin, Bmp4 and Morphological
813 Variation of Beaks in Darwin's Finches. *Science* **305**, 1462–1465 (2004).

814 61. R. C. Albertson, J. T. Streelman, T. D. Kocher, P. C. Yelick, Integration and evolution of
815 the cichlid mandible: The molecular basis of alternate feeding strategies. *Proc. Nat. Acad.*
816 *Sci. U.S.A.* **102**, 16287–16292 (2005).

817 62. Y. Terai, N. Morikawa, N. Okada, The Evolution of the Pro-Domain of Bone
818 Morphogenetic Protein 4 (Bmp4) in an Explosively Speciated Lineage of East African
819 Cichlid Fishes. *Mol. Biol. Evol.* **19**, 1628–1632 (2002).

820 63. S. Nawaz, *et al.*, WNT10A missense mutation associated with a complete Odonto-
821 Onycho-Dermal Dysplasia syndrome. *Eur. J. Hum. Genet.* **17**, 1600–1605 (2009).

822 64. E. L. Benard, *et al.*, wnt10a is required for zebrafish median fin fold maintenance and
823 adult unpaired fin metamorphosis. *Dev. Dyn.* **253**, 566–592 (2024).

824 65. T. A. Square, *et al.*, Modulation of tooth regeneration through opposing responses to Wnt
825 and BMP signals in teleosts. *Development* **150**, dev202168 (2023).

826 66. T. Le, *et al.*, A zebrafish model of crim1 loss of function has small and misshapen lenses
827 with dysregulated clic4 and fgf1b expression. *Front. Cell Dev. Biol.* **13** (2025).

828 67. D. A. Marques, *et al.*, Convergent evolution of SWS2 opsin facilitates adaptive radiation
829 of threespine stickleback into different light environments. *PLOS Biol.***15**, e2001627
830 (2017).

831 68. C. M. Hofmann, *et al.*, The Eyes Have It: Regulatory and Structural Changes Both
832 Underlie Cichlid Visual Pigment Diversity. *PLOS Biol.***7**, e1000266 (2009).

833 69. K. E. O'Quin, C. M. Hofmann, H. A. Hofmann, K. L. Carleton, Parallel Evolution of
834 Opsin Gene Expression in African Cichlid Fishes. *Mol. Biol. Evol.* **27**, 2839–2854 (2010).

835 70. G. A. C. Phillips, K. L. Carleton, N. J. Marshall, Multiple Genetic Mechanisms Contribute
836 to Visual Sensitivity Variation in the Labridae. *Mol. Biol. Evol*. **33**, 201–215 (2016).

837 71. K. L. Carleton, D. Escobar-Camacho, S. M. Stieb, F. Cortesi, N. J. Marshall, Seeing the
838 rainbow: mechanisms underlying spectral sensitivity in teleost fishes. *J. Exp. Biol.* **223**,
839 jeb193334 (2020).

72. M. G. Ritchie, Sexual Selection and Speciation. *Annu. Rev. Ecol. Evol. Syst.* **38**, 79–102 (2007).

73. M. R. Servedio, J. W. Boughman, The Role of Sexual Selection in Local Adaptation and Speciation. *Annu. Rev. Ecol. Evol. Syst.* **48**, 85–109 (2017).

74. A. Devigili, *et al.*, Possible glimpses into early speciation: the effect of ovarian fluid on sperm velocity accords with post-copulatory isolation between two guppy populations. *J. Evol. Biol.* **31**, 66–74 (2018).

75. M. D. Garlovsky, *et al.*, Synthesis and Scope of the Role of Postmating Prezygotic Isolation in Speciation. *Cold Spring Harb. Perspect. Biol*. a041429 (2023).

76. M. C. Kustra, M. R. Servedio, S. H. Alonzo, Cryptic female choice can maintain reproductive isolation. *Evolution* **79**, 2259–2273 (2025).

77. M. C. Kustra, S. H. Alonzo, The coevolutionary dynamics of cryptic female choice. *Evol. Lett.* **7**, 191–202 (2023).

78. P. D. Lorch, M. R. Servedio, The evolution of conspecific gamete precedence and its effect on reinforcement. *J. Evol. Biol.* **20**, 937–949 (2007).

79. J. L. Marshall, M. L. Arnold, D. J. Howard, Reinforcement: the road not taken. *Trends Ecol. Evol.* **17**, 558–563 (2002).

80. A. Kodric-Brown, R. J. D. West, Asymmetries in premating isolating mechanisms in a sympatric species flock of pupfish (Cyprinodon). *Behav. Ecol.* **25**, 69–75 (2014).

81. U. Strecker, A. Kodric-Brown, Mate recognition systems in a species flock of Mexican pupfish. *J. Evol. Biol.* **12**, 927–935 (1999).

82. M. H. K. Chau, *et al.*, Investigation of the genetic etiology in male infertility with apparently balanced chromosomal structural rearrangements by genome sequencing. *Asian J. Androl.* **24** (2022).

83. H. Li, Y. Dai, Z. Luo, D. Nie, Cloning of a new testis-enriched gene C4orf22 and its role in cell cycle and apoptosis in mouse spermatogenic cells. *Mol. Biol. Rep.* **46**, 2029–2038 (2019).

84. Y. Miyamoto, *et al.*, The STK35 locus contributes to normal gametogenesis and encodes a lncRNA responsive to oxidative stress. *Biol. Open* **7**, bio032631 (2018).

85. S. Nawaz, S. Hussain, S. Basit, W. Ahmad, First evidence of involvement of *TBC1D25* in causing human male infertility. *Eur. J. Med. Genet.* **64**, 104142 (2021).

86. L. Campagna, *et al.*, Repeated divergent selection on pigmentation genes in a rapid finch radiation. *Sci. Adv.* **3**, e1602404 (2017).

87. F. C. Jones, *et al.*, The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**, 55–61 (2012).

88. Z. D. Blount, R. E. Lenski, J. B. Losos, Contingency and determinism in evolution: Replaying life's tape. *Science* **362**, eaam5979 (2018).

89. P. W. Hedrick, Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Mol. Ecol.* **22**, 4606–4618 (2013).

90. S. G. Das, M. Mungan, J. Krug, Epistasis-mediated compensatory evolution in a fitness landscape with adaptational tradeoffs. *Proc. Natl. Acad. Sci. U.S.A.* **122**, e2422520122 (2025).

91. M. J. Harms, J. W. Thornton, Historical contingency and its biophysical basis in glucocorticoid receptor evolution. *Nature* **512**, 203–207 (2014).

92. M. Karageorgi, *et al.*, Genome editing retraces the evolution of toxin resistance in the monarch butterfly. *Nature* **574**, 409–412 (2019).

93. R. D. Tarvin, *et al.*, Interacting amino acid replacements allow poison frogs to evolve epibatidine resistance. *Science* **357**, 1261–1266 (2017).

94. D. M. Weinreich, N. F. Delaney, M. A. DePristo, D. L. Hartl, Darwinian Evolution Can Follow Only Very Few Mutational Paths to Fitter Proteins. *Science* **312**, 111–114 (2006).

95. S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).

96. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. [Preprint] (2013). Available at: http://arxiv.org/abs/1303.3997.

97. K. Okonechnikov, A. Conesa, F. García-Alcalde, Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **32**, 292–294 (2016).

98. A. McKenna, *et al.*, The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).

99. M. A. DePristo, *et al.*, A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).

100. G. A. Van der Auwera, *et al.*, From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Curr. Prot. Bioinf.* **43**, 11.10.1-11.10.33 (2013).

101. D. Caetano-Anolles, GenotypeGVCFs and the death of the dot (obsolete as of GATK 4.6.0.0). *GATK* (2024).

102. P. Danecek, *et al.*, Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).

103. S. Purcell, *et al.*, PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Amer. J. Hum. Genet.* **81**, 559–575 (2007).

104. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome. Res.* **19**, 1655–1664 (2009).

105. S. Schiffels, K. Wang, "MSMC and MSMC2: The Multiple Sequentially Markovian Coalescent" in *Statistical Population Genomics*, J. Y. Dutheil, Ed. (Springer US, 2020), pp. 147–166.

106. P. Danecek, *et al.*, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).

107. J. M. Schmidt, M. de Manuel, T. Marques-Bonet, S. Castellano, A. M. Andrés, The impact of genetic adaptation on chimpanzee subspecies differentiation. *PLOS Genet.* **15**, e1008485 (2019).

108. S. G. Hernández-Ávila, C. W. Hoagstrom, W. A. Matamoros, Historical biogeography of North American killifishes (Cyprinodontiformes) recapitulates geographical history in the Gulf of México watershed. *Zool. J. Linn. Soc.* **202**, zlae105 (2024).

109. E. Y. Durand, N. Patterson, D. Reich, M. Slatkin, Testing for Ancient Admixture between Closely Related Populations. *Mol. Biol. Evol*. **28**, 2239–2252 (2011).

110. C. Feng, J. Wang, A. Liston, M. Kang, Recombination Variation Shapes Phylogeny and Introgression in Wild Diploid Strawberries. *Mol. Biol. Evol.* **40**, msad049 (2023).

111. A. Morales-Cruz, *et al.*, Introgression among North American wild grapes (Vitis) fuels biotic and abiotic adaptation. *Genome Biol.* **22**, 254 (2021).

112. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

113. P. Cingolani, *et al.*, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **6**, 80–92 (2012).

114. M. Schumer, *et al.*, Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science* **360**, 656–660 (2018).

115. C. T. Rueden, *et al.*, ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinform.* **18**, 529 (2017).

116. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).

117. R. V. Lenth, *emmeans: Estimated Marginal Means, aka Least-Squares Means* (2021).

118. G. A. Van der Auwere & B. D. O' Connor. *Genomics in the cloud: using docker, GATK, and WDL in Terra.* (Sebastopol, CA O'Reilly Media, 2020).

119. A. McKenna *et al.,* The genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297-1303 (2010).

120. M. A. DePristo *et al.,* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491-498 (2011).

967 **Supplementary materials**

968

969 **Table S1.** Mean genome-wide pairwise F_{ST} statistics within the Chichancanab radiation and with
970 sister outgroup (*artifrons*).

	labiosus	***beltrani***	***artifrons***
simus	0.090	0.068	0.182
labiosus		0.032	0.184
beltrani			0.168

971

Table S2. F_{ST} statistics within the Chichancanab radiation.

Comparison	**F_{ST} = 1 (fixed)**	**1> F_{ST} ≥ 0.95 (nearly fixed)**	**0.95> F_{ST} ≥ 0.9**	**0.9> F_{ST} ≥0.8**	**Other**	**Mean F_{ST}**
simus vs beltrani + *labiosus*	19	1127	3304	20173	10359605	0.07
labiosus vs *beltrani* + *simus*	0	0	3	94	10384131	0.04
beltrani vs *simus* + *labiosus*	0	0	0	13	10384215	0.02

986 **Table S3.** Fixed SNPs between *C. simus* and *C. beltrani* + *C. labiosus*. Bold rows indicate SNPs within top 1% of PBS windows.

987 Bolded genes only are genes that have regions that fall within the top 1% PBS percentile.

Position	Gene	Distance to gene	PBS percentile	Annotation
Chr 1: 5290600	**STK35**	12594	1.85	
Chr 1:5290602	**STK35**	12592	1.85	Fertility, eye development and oxidative stress (Miyamoto et al 2018).
Chr 1:5290604	**STK35**	12590	1.85	
Chr 2: 17498913	SLC17A8	Intron	1.24	Glutamate neurotransmitter, inner hair cell dysfunction
Chr 3: **11687002**	**AGO3**	**Intron**	**0.34**	**RNA regulator**
Chr 3: 32286614	RBMS3	14363	22.93	
Chr 3: 32286627	RBMS3	14350	22.93	Involved in chondrogenic craniofacial development of zebrafish
Chr 3: 32286638	RBMS3	14339	22.93	(Jayasena and Bronner 2012).
Chr 3: 32286614	RBMS3	14363	22.93	
Chr 11: 27324192	GRIA3	Intron	2.85	
Chr 11: 27336469	GRIA3	2772	3.18	Glutamate receptor strongly expressed in zebrafish retina
Chr 11: 27336472	GRIA3	2775	3.18	development.
Chr 11: 27336476	GRIA3	2779	3.18	
Chr 11: **28255536**	**NA**	**3026**	**0.98**	**NA**
Chr 11: **28255616**	**NA**	**3106**	**0.98**	**NA**
Chr 11: **28256343**	**NA**	**3833**	**0.98**	**NA**
Chr 11: **28257367**	**NA**	**4857**	**0.98**	**NA**
Chr 11: 28260047	NA	7537	2.87	NA
Chr 12: 13854139	**C4orf22**	11903	1.30	Spermatogenesis, cilia, flagellum
Chr 24: **12693158**	**KDM6A**	**Intron**	**0.05**	**Kabuki syndrome (craniofacial development)**

988

Table S4. Summary statistics of the distance between closest adaptive candidate SNPs within the same gene

Gene	Median Distance	Mean Distance	SD	
TM6SF2	7430	7430	NA	2
C19orf10	4789	4789	NA	2
KDM6A	2892.5	2892.5	2410.52702	3
KCNB2	1803	1738.75	1614.92299	5
UBE2T	1742	1742	NA	2
TMEM230	1734	3039	3599.52427	4
ABCA1	1037.5	1097.5	959.375665	5
HAPLN4	974	1608.21739	9927.05273	24
GRIN2A	904	1270.6667	4273.38008	4
DPH1	608	1224.5	1999.25934	7
WNT10A	402	1442.17857	3986.39626	29
OPN1LW	386.5	359.5	225.229513	5
NIPSNAP3A	253	1451.71429	2819.19065	8
GALNT15	246.5	675.166667	3156.86746	13
SWS2	246	1381.25	2423.03892	5
STK35	217	793.366667	1992.8543	31
CRIM1	201	47.829268	2176.05381	42
PDYN	187	326.371429	355.925561	36
ADCYAP1R1	186.5	833	1304.35501	11
DNM1	186.5	459.708333	677.261781	25
DNM1	186.5	459.708333	677.261781	25
KCNH8	131.5	584.366667	1030.65532	31
TOP1	124	332.375	6687.63393	17
AGO3	115	357.625	552.097284	9
GMEB2	112	319.490196	2485.05809	52
GPNMB	57	57	77.781746	3
IGFBP4	26	2254	3880.67997	4
FAM83D	23	2523.66667	4347.74658	4
EXOSC3	17	17	NA	2
C4orf22	15	15	9.899495	3
TMEM106B	7	11.4	10.358571	6
THOC3	6	5.833333	3.600926	7
MID1	1	1	NA	2

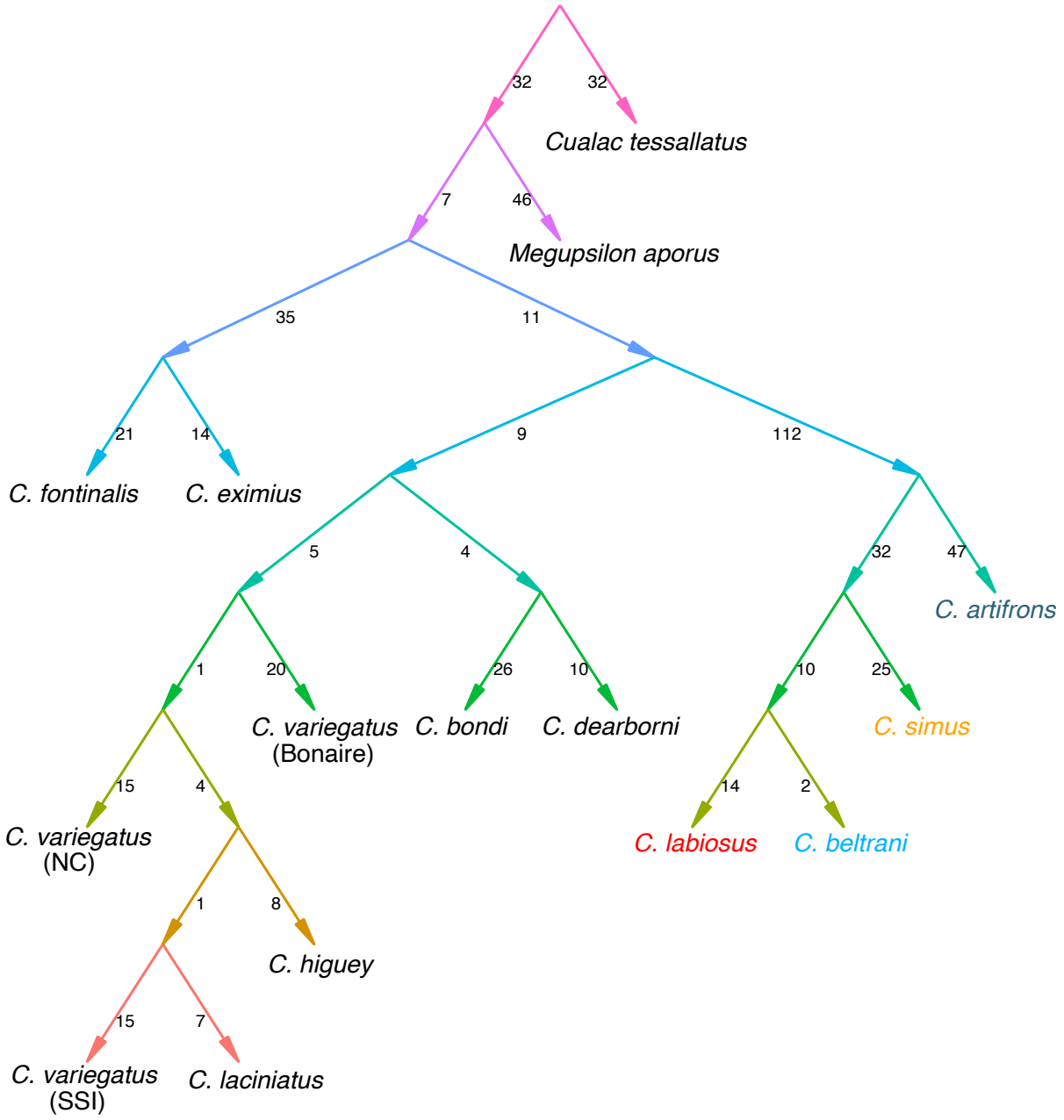


Figure S1. Best fitting admixture graph assuming zero admixture events. Admixture graph was found using the *findGraphs* function in *ADMIXTOOLS2* (Maier et al., 2023). Number on edges represent drift weights.

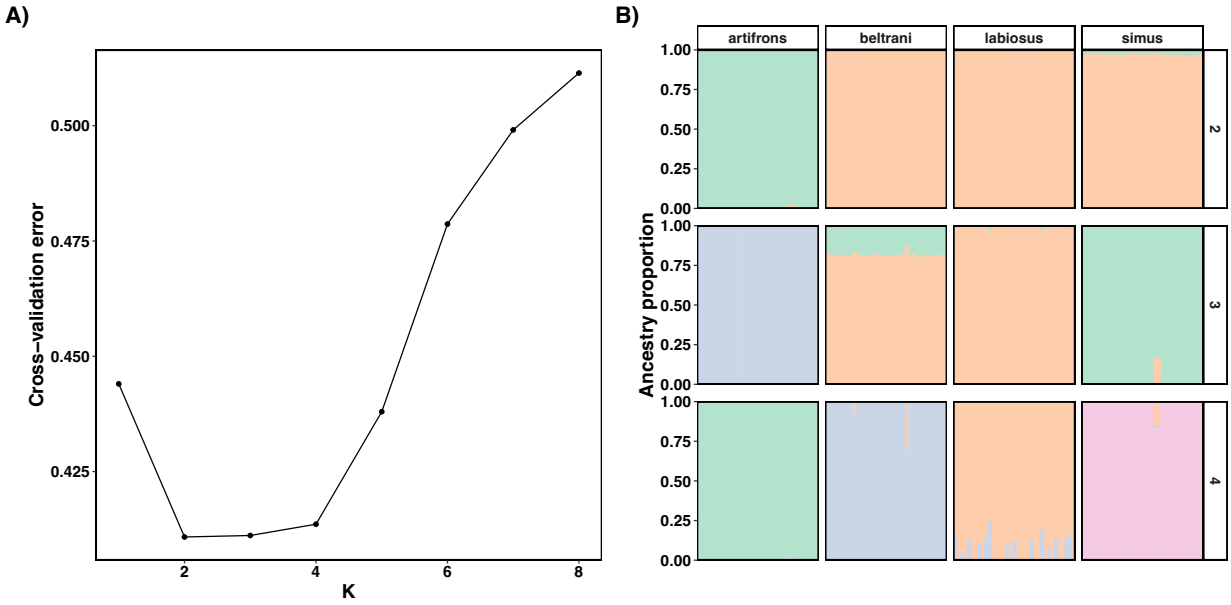


Figure S2. ADMIXTURE results support genetic differentiation of Yucatán Cyprinodon pupfish species. (A) Cross-validation error for different values of K clusters, equally support 2, 3, and 4 clusters. (B) ADMIXTURE plots showing the clusters at different values of K.

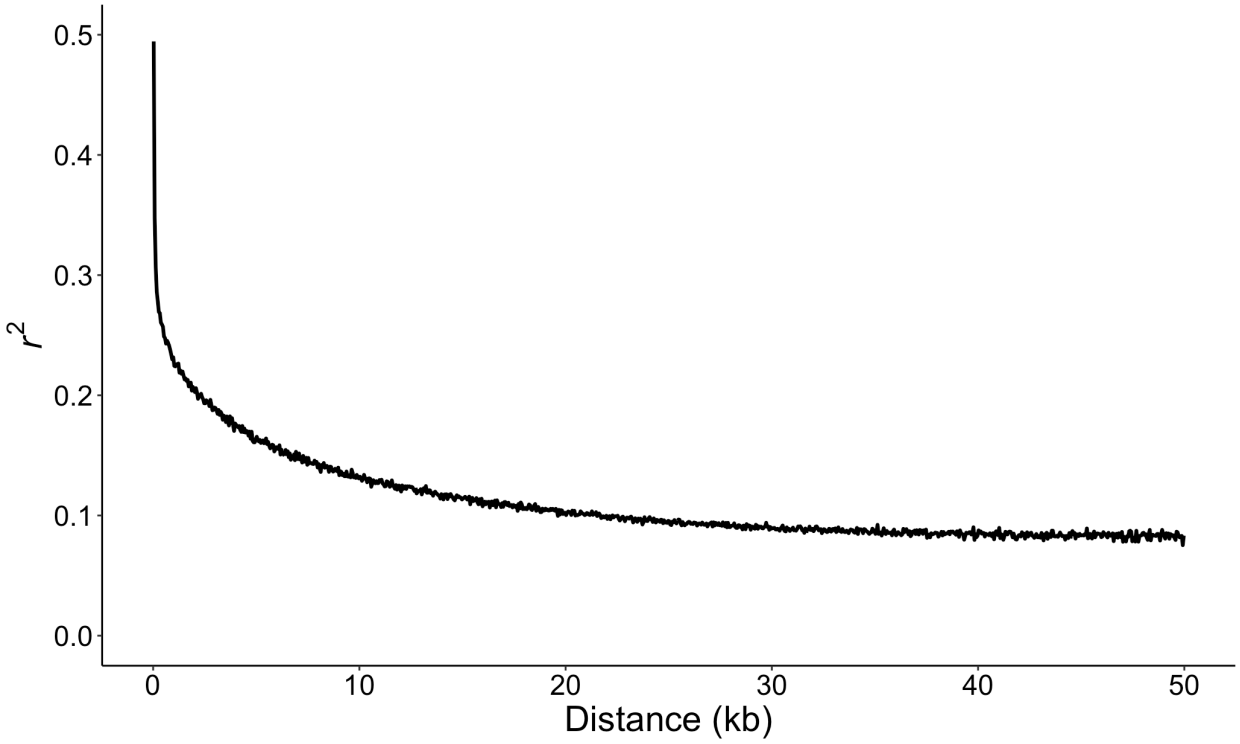
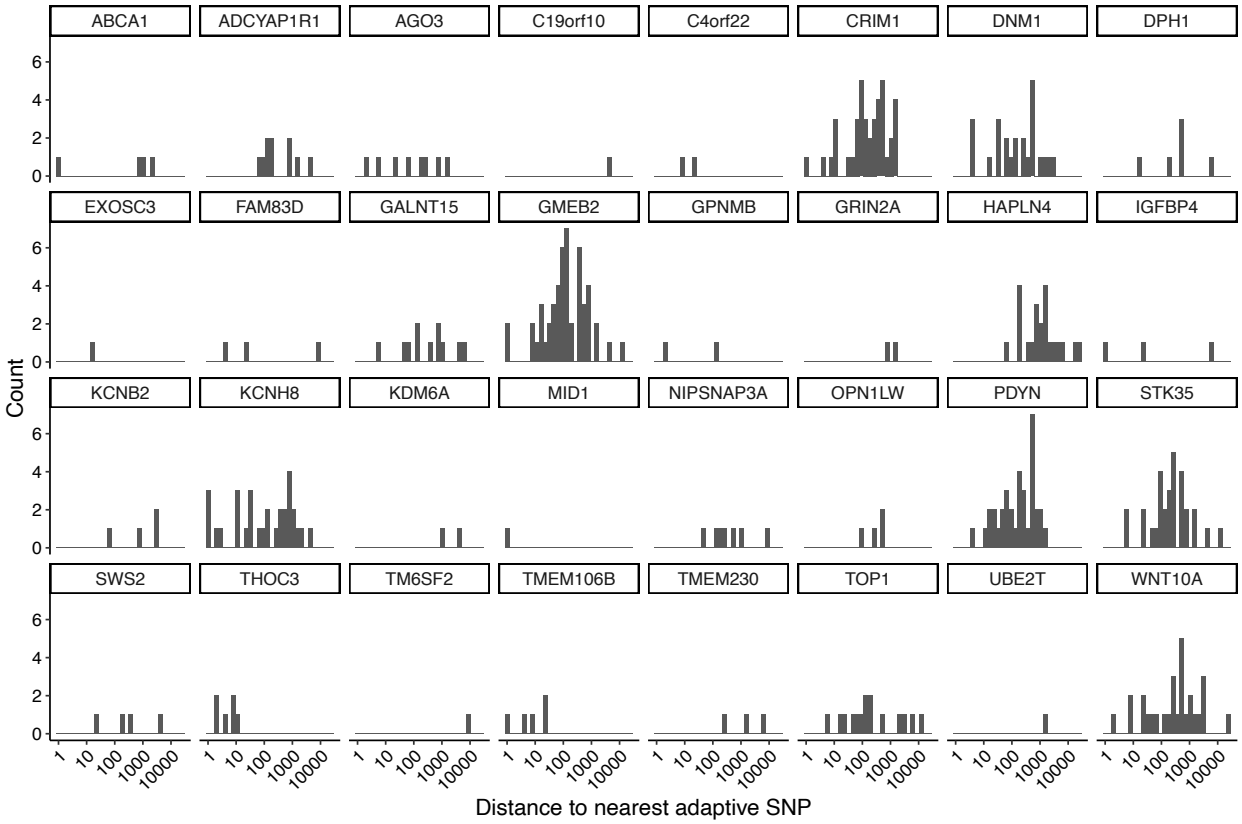


Figure S3. Linkage decay occurs rapidly in *Cyprinodon simus*. Linkage decay was calculated in windows of 1000KB using PLINK (v1.9). The results graphed are a random sample of all calculated r^2 values taken for plotting purposes (n = 9,645 r^2 values).



1033

1034 **Figure S4. Adaptative SNPs associated with the same gene are generally spaced apart.**

1035 Distribution of the distance to nearest adaptive SNP within a gene. Note the log scale of the x-

1036 axis.

1037

1038

1039

1040

1041

1042

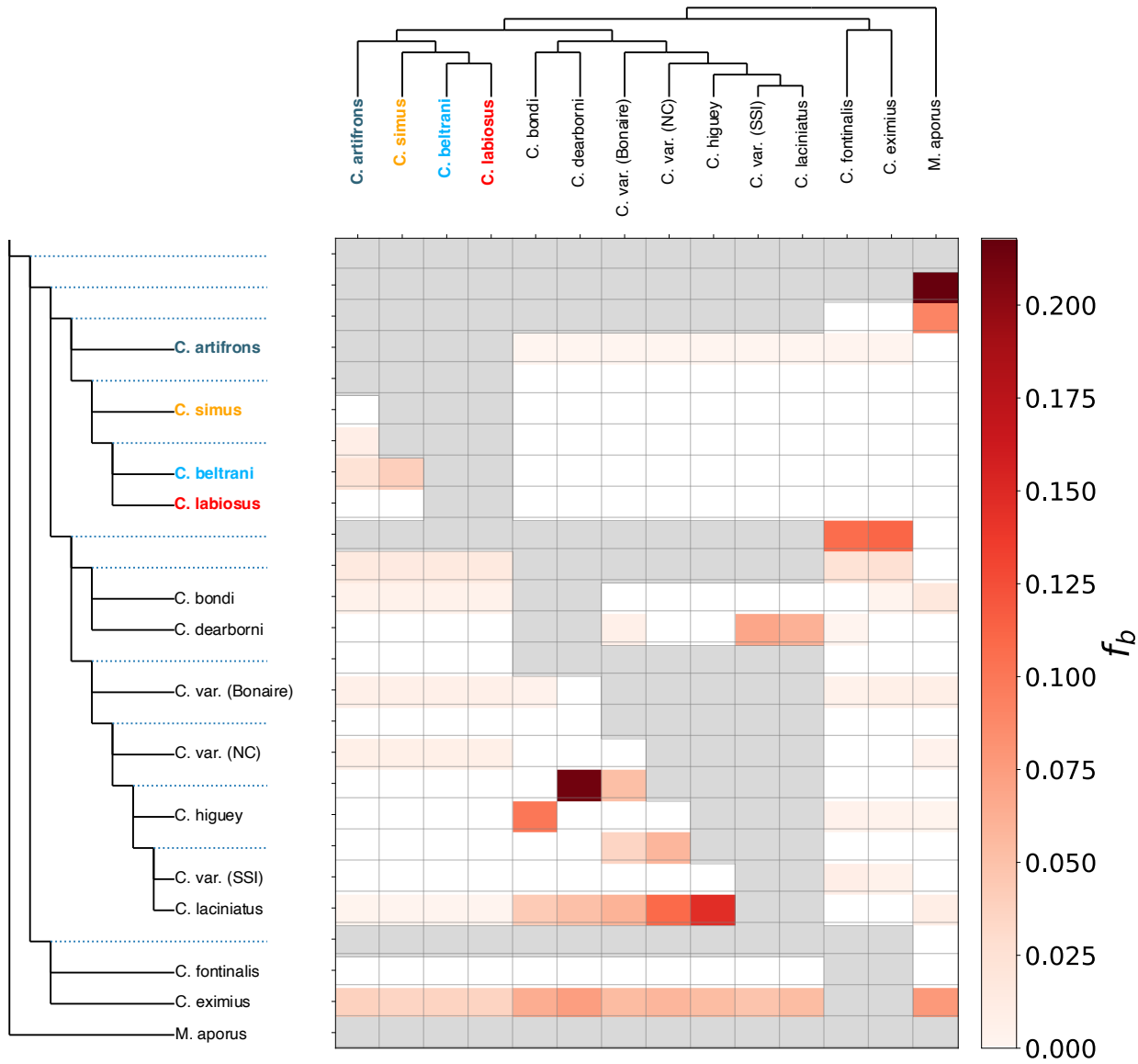


Figure S5. *f*-branch (f_b) show that introgression with desert pupfish occurred before the Chichancanab radiation. Boxes with red color represent significant introgression between P3 species (top) and the species/internal branch on the left. For visualization, non-significant f_b were changed to zero and are completely white boxes. Grey boxes represent comparisons that cannot be made, as f_b cannot be calculated for introgression between sister taxon. Dashed lines on the left are internal branches. SSI: San Salvador Island, Bahamas; NC: North Carolina, USA.

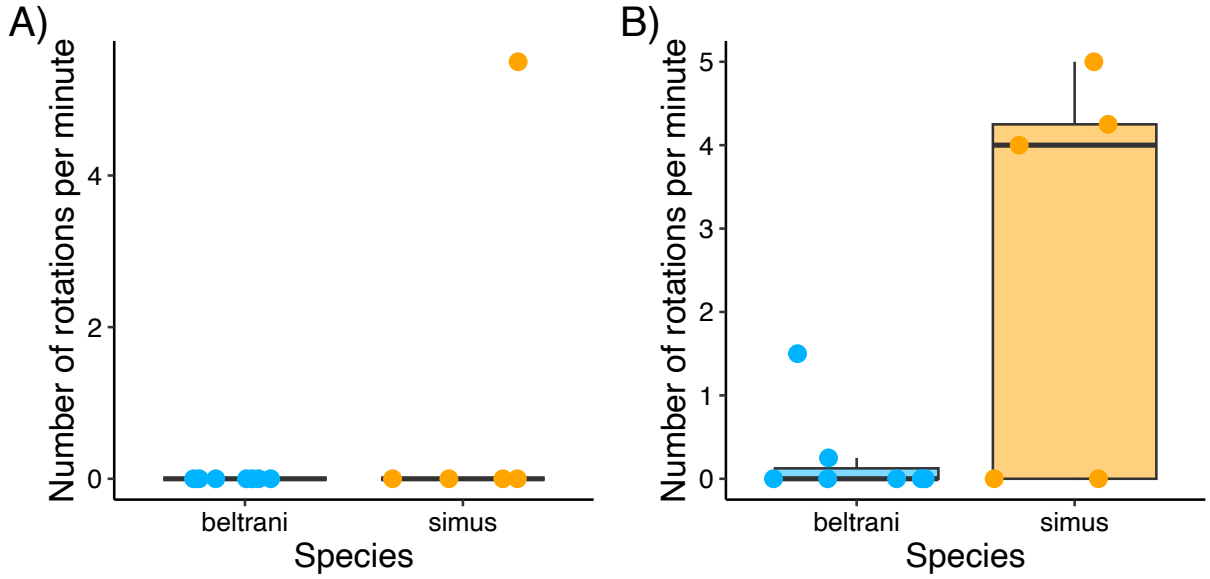
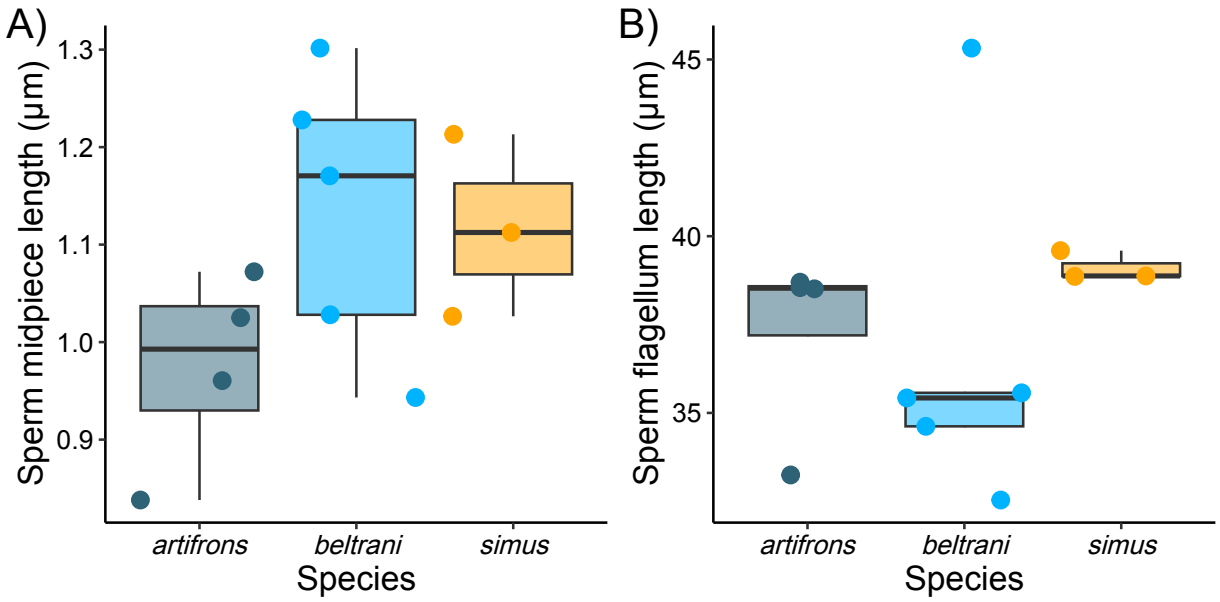


Figure S6. **Both species did not have a strong response to either control scenario in the optomotor response experiment. (A)** Number of complete rotations fish swam when the machine was off and **(B)** when the machine was on but the apparatus was all black**.** Boxplots and points of the number of complete rotations individuals performed during the "off-banded" portion of the experiment. Note that some boxplots are compressed to a single line because most individuals had no response.



1068

1069 **Figure S7**. **Species do not differ in sperm midpiece size or sperm flagellum length.** Boxplots

1070 and points of the mean sperm morphological estimate for an individual.

1071