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, Felipe Carrillo Puerto, Quintana Roo, Mexico.
[10]Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, United States

*Corresponding author: Matthew Kustra

Email: matthewckustra@berkeley.edu

Keywords: Speciation, De novo mutations, Introgression, Genomics, Adaptation

This PDF file includes:

Main Text
Figures 1 to 5
Supporting Information (Supplemental Methods; Tables S1-S4; Figures S1-S10)

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 compare the adaptive radiation of *Cyprinodon* pupfishes in Lake Chichancanab, Mexico, to an independent radiation of San Salvador Island (SSI) pupfishes in the Bahamas, and assess the repeatability of adaptive 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 (0.5%; 2/426) 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 geographic isolation of this inland lake. Instead, de novo mutations provided a substantial source of adaptive variation (30.6%) for the endemic zooplanktivore, 15 times higher than the endemic scale-eater on SSI. However, in parallel with SSI, we found strong evidence that adaptive divergence occurred in stages, first on regulatory and standing genetic variation, then on de novo and nonsynonymous mutations. Consistent with adaptive variants near opsin and spermatogenesis genes, functional categories unique to Chichancanab, we found greater visual acuity and divergent sperm morphology in lab-reared zooplanktivores relative to generalists using laboratory assays. Consistent with extensive adaptive de novo mutations in *WNT10A* and rapid diversification of tooth size in the zooplanktivore, we found that experimental inhibition of the Wnt pathway in generalists resulted in narrower oral teeth. We conclude that de novo mutations, not introgression, can drive rapid adaptive radiations 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 on the adaptive landscape with similar levels of genetic diversity, gene flow, and ecological opportunity (10, 11).

Beneficial de novo mutations can provide novel variation to reach new fitness peaks on the adaptive landscape during adaptive radiation (12, 13). However, the longer time needed for de novo beneficial mutations to arise contradicts the short timescales of rapid divergence (11, 14, 15). Hybridization can introduce large amounts of genetic variation and novel phenotypes (transgressive segregation) through adaptive introgression and hybrid swarms, enabling populations to reach previously inaccessible fitness peaks (16). Support for an important 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 relative importance of different sources of genetic variation for triggering and sustaining adaptive radiations across time and space.

Caribbean pupfishes are a remarkable system for assessing the repeatability of adaptive radiations due to two independent, recent adaptive radiations, each containing a generalist and multiple trophic specialists within isolated lake environments that have historically lacked predatory fishes and most competitors except *Gambusia*. The San Salvador Island (SSI), Bahamas radiation proceeded predominantly via the reassembly of pre-existing standing genetic variation found throughout the Caribbean, with substantial contributions of adaptive introgression to both the molluscivore (durophage) and scale-eating (lepidophage) specialists.

Only 0% and 2% of adaptive mutations in the molluscivore and scale-eater, respectively, were de novo mutations unique to SSI (26, 28–32). However, it was previously unknown how a second radiation in Lake Chichancanab, Mexico, drawing from the same Caribbean pool of standing genetic variation, unfolded over the same timeframe.

Here, we use whole-genome resequencing of 123 individuals to provide insight into the genomic architecture and genetic origins of the Lake Chichancanab *Cyprinodon* pupfish radiation (33). This recent adaptive radiation (~8000 years old (34)) consisted of at least five species (or up to seven (35)) historically (33, 35, 36), including a detritivore (*C. beltrani*), piscivore (*C. maya*), rocky substrate specialist (*C. verecundus*), bivalve specialist (*C. labiosus*), and zooplanktivore (*C. simus*). At the time of description (33), the relative abundances of these endemic species were 68-85% *C. beltrani*, 6-18% *C. maya*, 2-13% *C. labiosus*, and <1% *C. simus* and *C. verecundus*. However, due to recent colonization by invasive African tilapia (*Oreochromis niloticus x mossambicus*) (37, 38) and by several species native to surrounding basins but not Chichancanab (*Astyanax angustifrons, Mayaheros urophthalmus, Poecilia mexicana*; pers. obs.), all trophic specialist species now appear to be extinct since our collections in 2022, with only the generalist, *C. beltrani*, remaining in 2024. Despite exhibiting one of the fastest rates of trait diversification of any vertebrate in tooth size (130 times faster than background rates across *Cyprinodon* (29)) and rapid divergence in other craniofacial morphology (Figure 1A; S1), previous genetic studies of microsatellite data and mtDNA suggest that this radiation has low genetic differentiation (39, 40), enabling the identification of highly differentiated loci among these species.

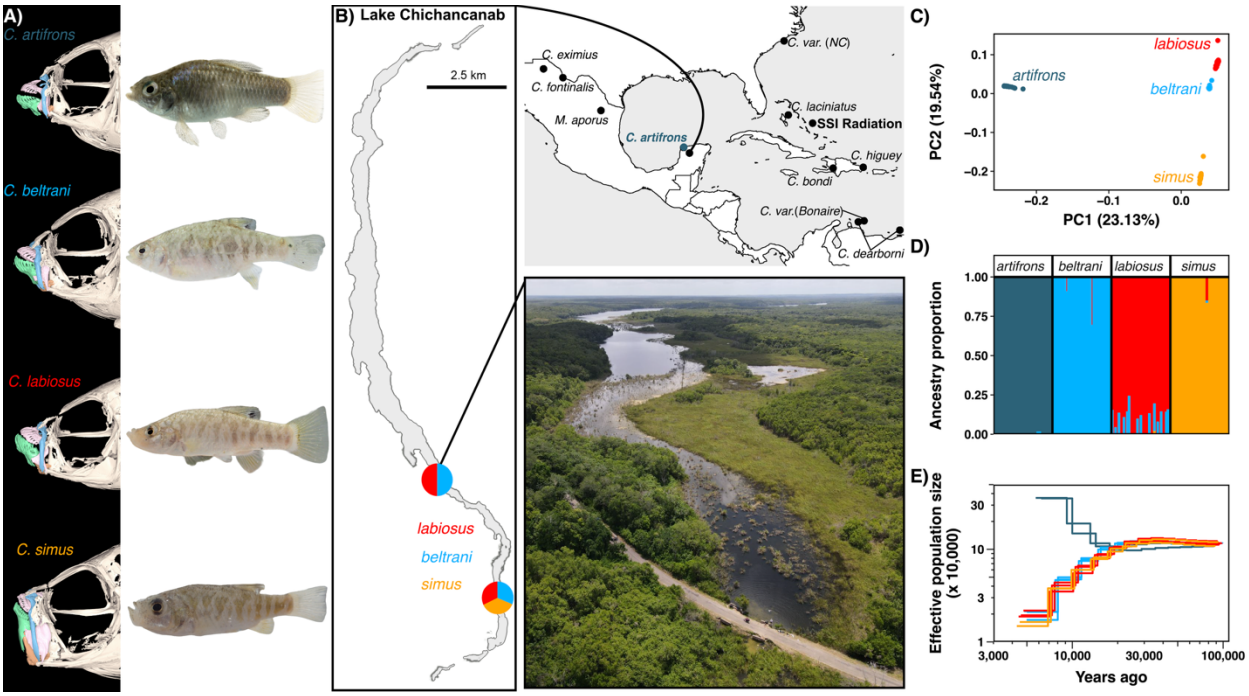


Figure 1. Yucatán *Cyprinodon* pupfishes show strong genetic differentiation. (A) µCT scans of Yucatán pupfish with maxilla colored in blue, premaxilla in pink, dentary in green, and articular in orange (see Figure S1 for different views). (B) Map of the Caribbean showing the sampling sites for all species in our dataset with an inset map of Lake Chichancanab and an aerial view image from one of the sampling sites. (C) Principal component analysis using a set of linkage disequilibrium pruned SNPs. (D) Estimated ancestry proportions for each individual inferred using *ADMIXTURE* with K = 4 (see Figure S3 for K=2-3) from the linkage-pruned dataset. (E) Estimated demographic history with *MSMC2* on a log scale. Each colored line represents an individual, with colors corresponding to different species.

Results

Genetic differentiation of the species flock

To measure population structure of the Yucatán *Cyprinodon* pupfishes, we 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*) using specimens collected from the wild in 2022 (Figure 1B). We aligned these samples, as well as eleven other previously sequenced Caribbean pupfish species, to the high-quality *Cyprindon nevadensis mionectes* reference genome (GCA_030533455.1). After calling and filtering variants, our dataset included 23,960,536 SNPs (12.4x mean depth). Using a linkage disequilibrium pruned set of SNPs (*n* = 5,476,855 SNPs), we used principal component analysis and *ADMIXTURE* to visualize population structure within Yucatán pupfishes. 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 1C). Within the Chichancanab species flock, *C. simus* was more genetically divergent from *C. labiosus* and *C. beltrani* (Figure 1C). These results were also supported by genome-wide F_{ST} estimates and a phylogenetic tree constructed with *ADMIXTOOLS2* (Table S1; Figure S2). Four clusters corresponded to the four different species, as determined by *ADMIXTURE* (Figure 1D). All *C. artifrons* individuals showed minimal evidence of admixture with Chichancanab species (Figure 1D; Figure S3).

Demographic history

Using *MSMC2* for Chichancanab species with a mutation rate of 1.56×10^{-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

145 species *C. artifrons* at the end of the Pleistocene 10-15 kya, consistent with receding sea levels
146 resulting in the formation of Lake Chichancanab (34). After divergence from *C. artifrons*, the
147 Chichancanab species flock showed a similar demographic history, characterized by a sharp
148 decline in effective population sizes, consistent with endemic sympatric species within an
149 isolated lake (Figure 1E).
150
151 *Adaptive loci within trophic specialists*
152 Although genome-wide differentiation among Chichancanab species was moderately low
153 (F_{ST} ranging from 0.032 to 0.09; Figure 2A; Table S1), we were able to identify highly
154 differentiated regions of the genome within the trophic specialist zooplanktivore *C. simus*. We
155 identified 19 fixed and 1,127 nearly-fixed (F_{ST} >0.95) SNPs in *C. simus* relative to other
156 Chichancanab pupfishes (Tables S2-S3). In contrast, we found zero fixed or nearly-fixed SNPs
157 within the generalist *C. beltrani* or bivalve-specialist C. *labiosus*. However, we were able to
158 identify some moderately differentiated SNPs within *C. labiosus* (n = 97) and *C. beltrani* (n =
159 13) relative to the other Chichancanab species at an F_{ST} threshold of 0.8 (Table S2).
160 To search for signals of selection, we calculated the normalized four-taxon population
161 branch statistic (PBS_{nj}) across 10-kb non-overlapping windows because linkage disequilibrium
162 decayed at this distance (Figure S4). Simulation studies demonstrate that normalized *PBS* is
163 effective for identifying both hard and soft selective sweeps (41). Out of the 1,146 SNPs fixed or
164 nearly fixed within *C. simus*, 661 were within a strongly supported selective sweep window (top
165 1% PBS_{nj} outlier (41)), including 425 SNPs within 20 kb of the first or last exon of an annotated
166 gene. These 425 candidate adaptive SNPs ("adaptive variants") were located in the proximity of,
167 or within, 55 genes (Figure 2A), distributed across 19 out of 24 chromosomes (Figure 2A).

168 Adaptive variants associated with the same gene were often distributed across multiple introns or
169 exons and spaced apart by 1kb or more (Figure S5, Table S4).
170
171 *Source and function of adaptive genetic variation in the zooplanktivore*
172 We next characterized the location of the 425 candidate adaptive variants near or within a
173 gene in the zooplanktivore *C. simus*. More than 95% were in intronic or flanking regions (20-kb
174 upstream or downstream of a gene; Figure 2B). Within coding regions, 13 out of 20 adaptive
175 variants were nonsynonymous mutations (Figure 2B). The number of adaptive variants per gene
176 followed an exponential distribution, with a handful of genes containing many variants and most
177 genes with only a single adaptive variant (Figure 2B). We functionally characterized these genes
178 by manually checking online databases and doing a literature search (Table S3; Figure 2B). 32
179 out of the 55 candidate genes were associated with craniofacial development (n =11; 20%) or
180 sensory/neural function (n = 21; 38.2%). We manually checked if any of these genes overlapped
181 with the 426 unique candidate genes identified in the SSI radiation (26, 28, 30, 32, 42–46), and
182 we found only two overlapping candidate genes (*DNM1* and *SPTLC3*), both of which are
183 associated with motor neuron development (47, 48) and are scale-eater-specific candidate genes.
184 Next, we characterized the source of the variants and found that 63% of the adaptive
185 SNPs in *C. simus* also existed as standing genetic variation in the coastal Yucatan population of
186 *C. artifrons* ("standing variation"), 31% were *de novo* variants found only in Chichancanab, and
187 6% were detected as standing genetic variation in other pupfish species but not *C. artifrons*
188 (Figure 2C).
189

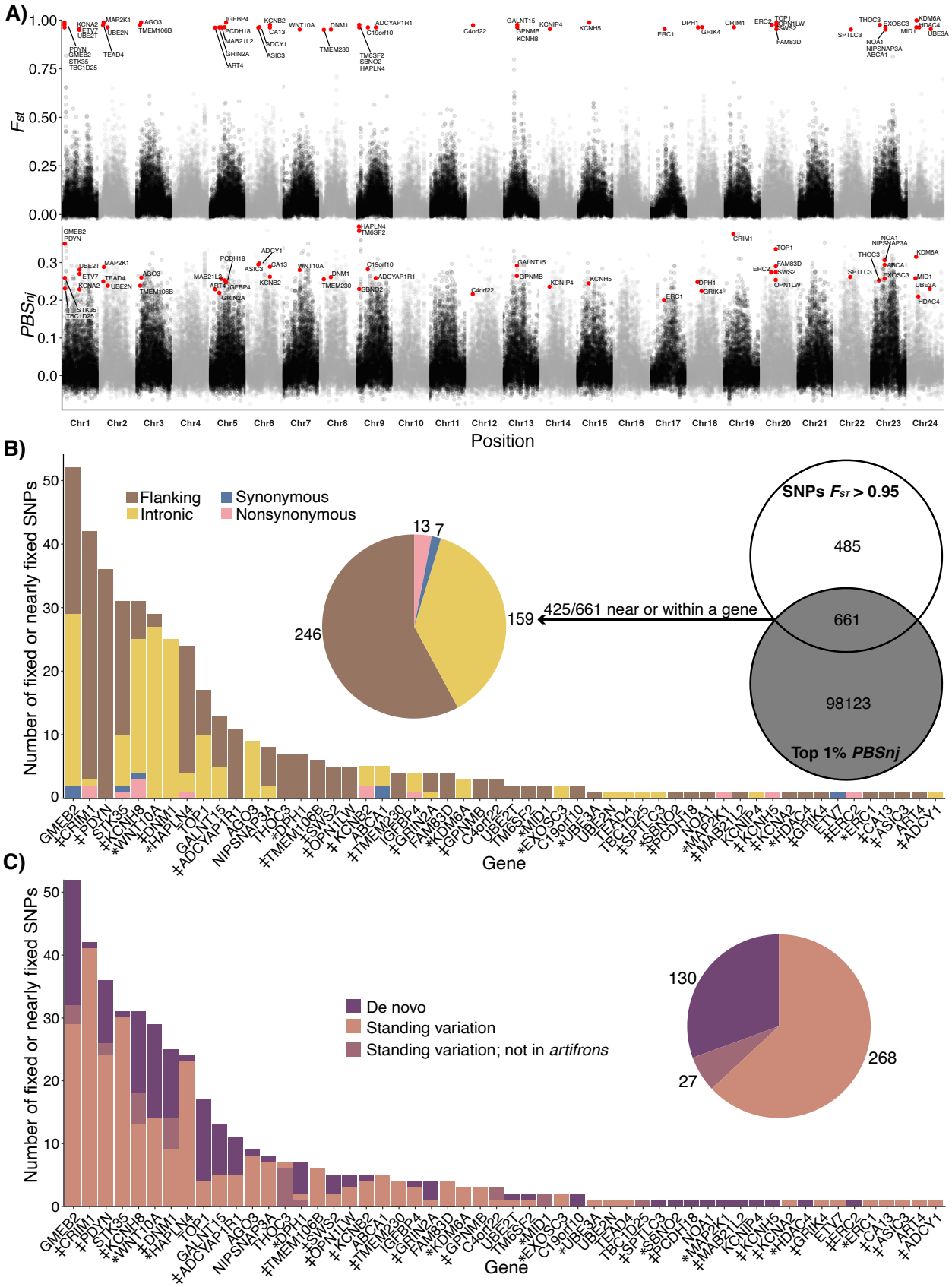

A)
F_{ST}
$PBSnj$
Chr1 Chr2 Chr3 Chr4 Chr5 Chr6 Chr7 Chr8 Chr9 Chr10 Chr11 Chr12 Chr13 Chr14 Chr15 Chr16 Chr17 Chr18 Chr19 Chr20 Chr21 Chr22 Chr23 Chr24
Position
B)
Number of fixed or nearly fixed SNPs
Flanking
Intronic
Synonymous
Nonsynonymous
13
7
246
159
425/661 near or within a gene
SNPs F_{ST} > 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: purple), as standing genetic variation in coastal sister species *C. artifrons* (pink), or in other Caribbean species, but not *C. artifrons*. * Indicates genes with a craniofacial annotation; ‡ indicates genes with neural or sensory annotation.

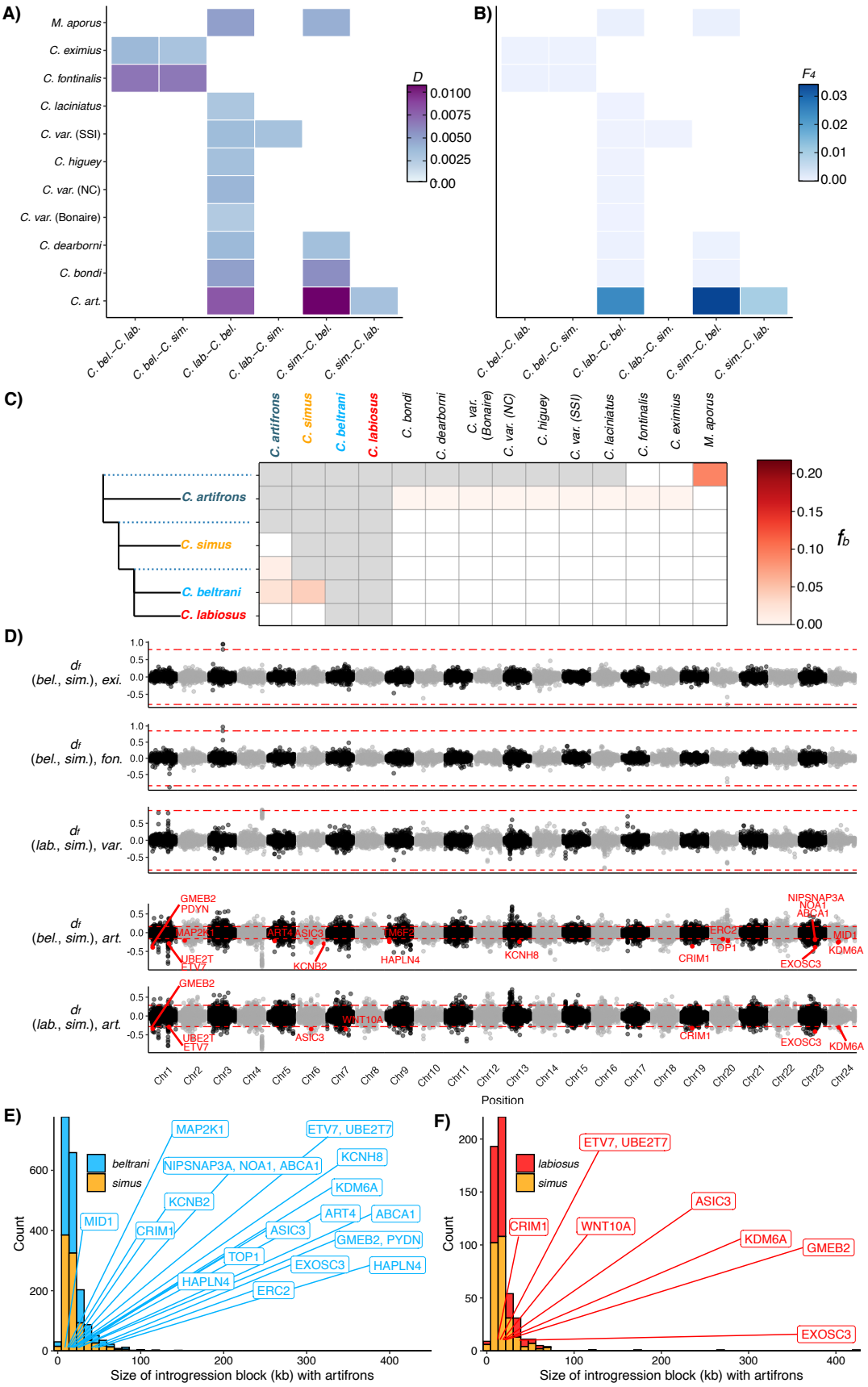
221 *Limited adaptive introgression within Chichancanab pupfishes*

222 To test for introgression into the Chichancanab radiation, we used *f*-statistics calculated in
223 *Dsuite* (49) to examine gene flow from 11 additional *Cyprinodon* species from across the clade
224 and the sister species to the genus *Megupsilon aporus*. Within Yucatán pupfishes, we found
225 evidence of introgression between *C. artifrons* and both *C. beltrani* and *C. labiosus*, with the
226 strongest signal in *C. beltrani* (Figure 3A, B). We also found evidence for introgression between
227 *C. beltrani* and most outgroup species examined, including the sister species to *Cyprinodon*,
228 *Megupsilon aporus*. Relative to *C. beltrani*, *C. labiosus* and *C. simus* also showed evidence of
229 introgression with pupfishes endemic to the Chihuahua desert (*C. eximius* and *C. fontinalis*;
230 Figure 3A, B). Furthermore, we also found evidence of introgression between the two sympatric
231 radiations (*C. variegatus* from SSI) and *C. simus* (Figure 3A, B). However, additional *f*-branch
232 statistics indicate that all patterns of excess allele sharing between these outgroup species besides
233 *C. artifrons* and the Chichancanab species flock likely resulted from ancient introgression events
234 before the colonization of Lake Chichancanab (Figure 3C; S6).

235 To test for adaptive introgression into *C. simus*, we next examined whether the set of
236 adaptive variants (Figure 2) occurred within outlier introgression windows using sliding window
237 scans of distance fraction (d_f(50)) for species that had significant genome-wide introgression
238 into *C. simus* (Figure 3A , B). We found that none of the outlier introgression windows
239 overlapped with *C. simus* adaptive variants (Figure 3D).

240 Next, we performed d_f scans for introgression between *C. artifrons* and Chichancanab
241 species. Although there were outlier introgression regions between *C. simus* and *C. artifrons*
242 (positive values d_f), these regions did not overlap with *C. simus* adaptive variants (Figure 3D).
243 However, the position of several of these *C. simus* adaptive variants overlapped with outlier

244 introgression windows for both *C. labiosus* and *C. beltrani* but not *C. simus* (Figure 3D). The
245 number of outlier introgression windows was much lower in *C. simus* relative to *C. labiosus* and
246 *C. beltrani* (Figure 3E, F). Furthermore, both *C. labiosus* and *C. beltrani* populations contained
247 several introgression tracts that were much longer than *C. simus* tract lengths, indicating much
248 more recent introgression events in these species (Figure 3E, F).
249


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
MID1
CRIM1
ART4
ABCA1
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

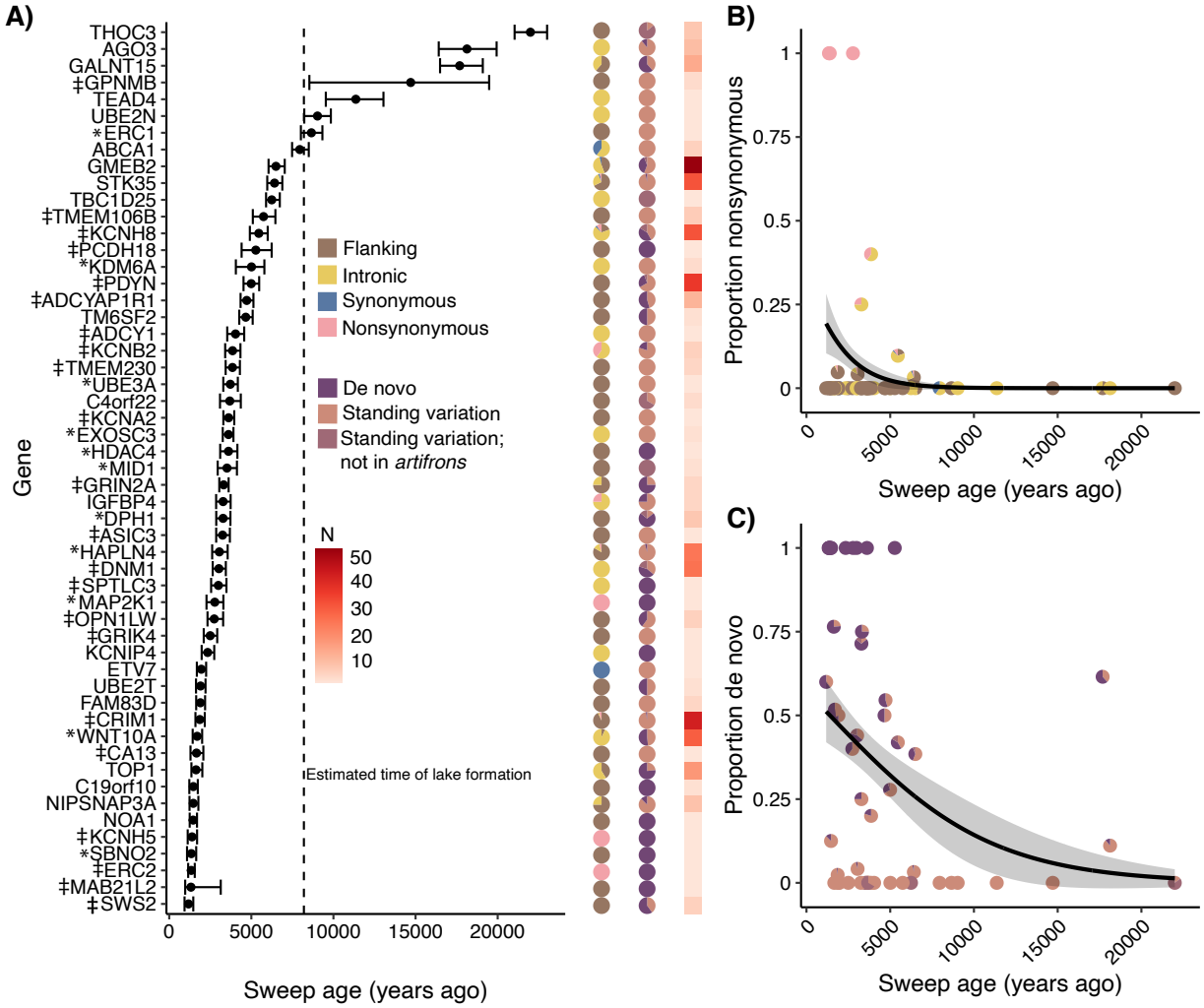
250

Figure 3. Minimal evidence of adaptive introgression into Chichancanab pupfishes. 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 S6 for full *f*-branch graphs. (D) Genome-wide plot of the distance fraction (d_f, an introgression test statistic (50)) 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 overall percentage of genome-wide introgression estimated from the 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 *starTMRCA* (51) to estimate the timing of selective sweeps for the set of candidate adaptive variants near genes within the zooplanktivore *C. simus*. We estimated that most selective sweeps occurred in a staggered sequence and not simultaneously after colonization of Lake Chichancanab (Figure 4A)(34). There were also a few older sweeps that we estimate may predate the age of Lake Chichancanab and divergence from the coastal sister species *C. artifrons* (Figure 1B). Overall, we observed no clear temporal stages of adaptive phenotypic divergence related to craniofacial genes (asterisks) or neural/sensory genes (‡; Figure 4A). Instead, we found that selection on these two broad functional categories occurred throughout the period of divergence of *C. simus* following the colonization of Lake Chichancanab, contrary to behavior-first, then craniofacial stages of adaptation (52) found in the SSI radiation (26).

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 in more recent selective sweeps during the final stage of adaptive divergence in *C. simus* (Figure 4B, C). However, nonsynonymous mutations were not significantly more likely to be de novo mutations found only in Chichancanab species (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 ($\chi^2 = 0.111$, $p = 0.740$).



307

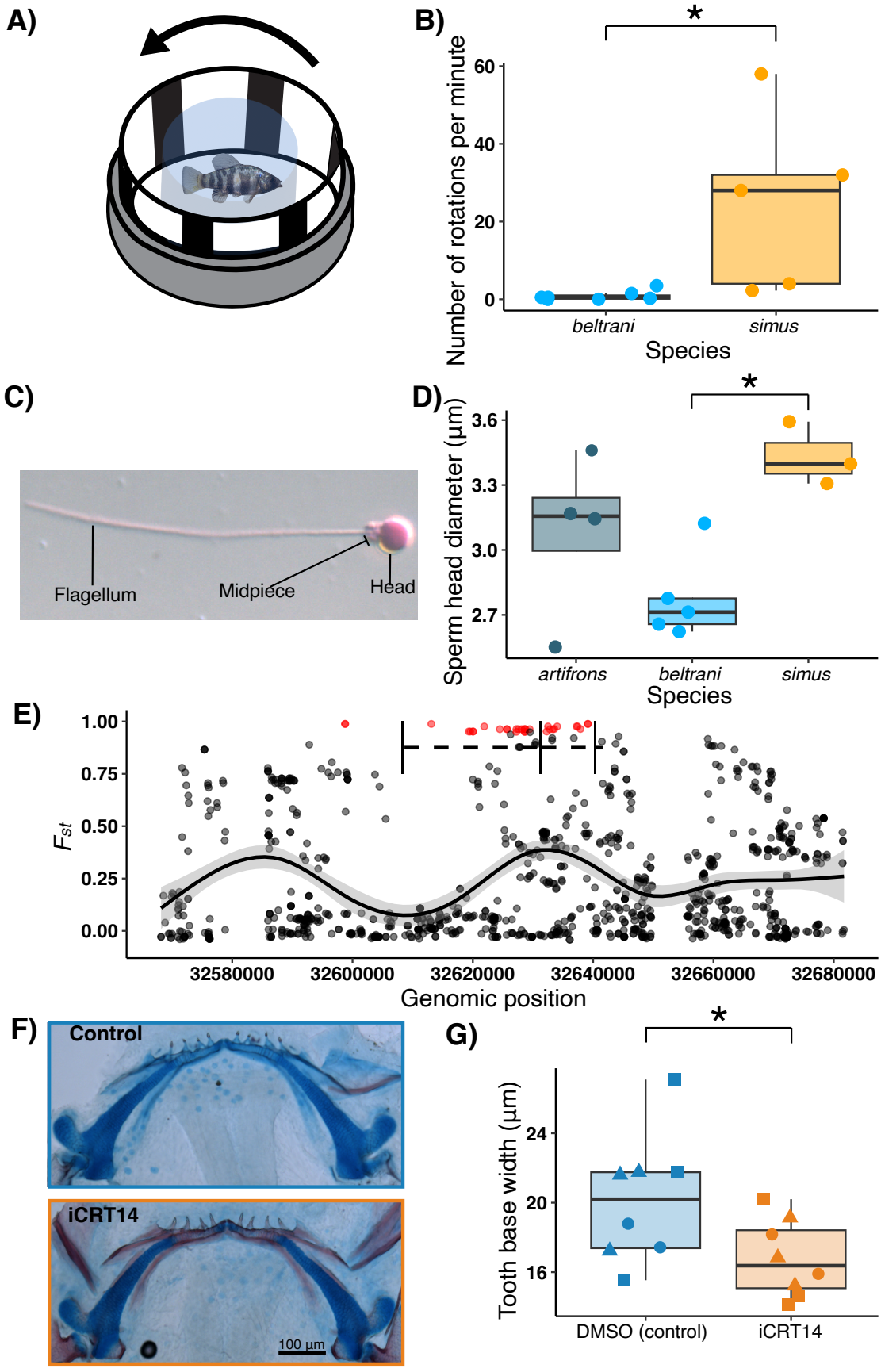
308 **Figure 4. Stages of adaptive divergence in the zooplanktivore.** (A) Points indicate median
309 sweep age with 95% credible intervals. The dashed line indicates the approximate age of Lake
310 Chichancanab (34). For each sweep, the first column of pie charts shows the proportions of
311 flanking/intronic/synonymous/nonsynonymous, second column of pie charts shows the
312 proportions of de novo/standing variation, and third column of squares shows the number of
313 candidate adaptive variants per gene (from Figure 2). For genes with multiple adaptive variants,
314 we used the variant with the median position as the focal position. * Indicates genes with a
315 craniofacial annotation; ‡ indicates genes with neural or sensory annotation. (B) The proportion
316 of nonsynonymous adaptive variants is significantly larger in more recent selective sweeps,
317 indicated by the best-fit quasibinomial regression line; each pie chart represents a single
318 candidate gene. (C) The proportion of de novo variants is significantly larger in more recent
319 selective sweeps; each pie chart represents a single candidate gene.

320

321

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

We identified several *C. simus* adaptive variants associated with eye development (*ASIC3*(53), *CRIM1*(54), *MAB21L2*(55)) and opsin (*OPN1LW*, *SWS2*) genes. Because the trophic specialization of *C. simus* requires visually locating zooplankton in the water column before precise suction-feeding strikes, we hypothesized that this species would have higher visual acuity than the detritivore, *C. beltrani*. We conducted optomotor response trials on lab-reared individuals from both species (F1 *C. beltrani* and long-term laboratory colony of F10+ *C. simus*) and found that *C. simus* has a much stronger optomotor response than *C. beltrani* in response to black-and-white vertical bars spinning at a constant rate of approximately 90 rotations per minute (Figure 5A, B; Wilcoxon signed-rank test: $W = 1$, $p = 0.009$). There was no significant difference in rotations per minute between the species for the same individuals during control observation periods with no spinning (Figure S7A; $W = 14$, $p = 0.3105$) or during positive control observation periods with an all-black background spinning at the same frequency (Figure S7B; $W = 9$, $p = 0.147$).


A)
B)
*
Number of rotations per minute
60
40
20
0
beltrani
simus
Species
C)
Flagellum
Midpiece
Head
D)
*
Sperm head diameter (µm)
3.6
3.3
3.0
2.7
artifrons
beltrani
simus
Species
E)
1.00
0.75
0.50
0.25
0.00
F_{st}
32580000
32600000
32620000
32640000
32660000
32680000
Genomic position
F)
Control
iCRT14
100 µm
G)
*
Tooth base width (µm)
24
20
16
DMSO (control)
iCRT14

339

Figure 5. Divergence in visual acuity, sperm morphology, and tooth development reflects candidate adaptive genes in the zooplanktivore. (A) Diagram of the optomotor response experimental apparatus (adapted from (56)). (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. (E) Manhattan plot showing per-site F_{st} of SNPs within 50 kb of *WNT10A*. Black rectangles indicate exons of *WNT10A*, and dashed lines indicate introns. Red points represent nearly fixed SNPs ($F_{st} > 0.95$), and black points represent all other SNPs. The black line is the smoothed average F_{st} in this region, with grey shading being the 95% confidence intervals. (F) Representative images of *C. beltrani* larvae that were treated with DMSO (control; top) or with iCRT14 Wnt inhibitor + DMSO (bottom) and stained with alizarin red (bone) and alcian blue (cartilage). (G) Boxplot and jittered points of average tooth base width of the left and right teeth closest to the mandibular symphysis for *C. beltrani* larvae treated with a DMSO control or Wnt inhibitor (iCRT14) + DMSO. Shapes represent three independent split-brood replicates.

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

Four candidate genes (*C4orf22*, *STK35*, *ERC2*, and *TBC1D25*) are known to affect spermatogenesis/fertility (57–60). 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 F10+ C. *simus*), differed in sperm morphology using differential interference contrast microscopy. We found that species significantly differed in sperm head length (Figure 5D; $F = 5.724$, $p = 0.025$), but not sperm midpiece length (Figure S8A; $F = 2.403$, $p = 0.147$) or flagellum length (Figure S8B; $F = 0.414$, $p = 0.673$). *C. simus* had significantly larger sperm heads than *C. beltrani* using Tukey's post hoc pairwise comparisons ($q = 3.367$; $p = 0.0207$; Fig. 5D).

Inhibition of WNT10A alters tooth morphology in Chichancanab pupfish

We identified a large number of highly differentiated SNPs surrounding the well-characterized *WNT10A*, which controls tooth, hair, and skin development within the Wnt signaling pathway (Figure 2; Figure 5E (61–63)). This is consistent with the most exceptional morphological difference between *C. simus* and *C. beltrani,* extreme divergence in tooth size (130 times faster than background rates (29)).

To test for the importance of *WNT10A* in tooth development within pupfishes, we used iCRT14 to inhibit the Wnt / β catenin pathway downstream of WNT10A during metamorphosis in *C. beltrani* lab-reared fry (64, 65). Chemical inhibition from 8 dpf to 22 dpf resulted in narrower teeth than control fry treated with DMSO ($\chi^2 = 4.840$, $p = 0.0278$; Figure 5F, G). There

were no significant differences in mean tooth length ($\chi^2 = 0.455$, $p = 0.500$; Figure S9) or number of ossified teeth when stained with alizarin red ($\chi^2 = 0.114$, $p = 0.7357$; Figure S10).

Discussion

Here, we investigated the genetic architecture, sources of adaptive variation, and temporal dynamics of selection underlying speciation and trophic specialization in a young adaptive radiation of *Cyprinodon* pupfishes endemic to Lake Chichancanab, Mexico. By comparing these patterns to our previous work on a recent radiation of *Cyprinodon* pupfishes in SSI, Bahamas (26, 28–32), we discovered parallel and non-parallel features of these independent radiations in similar environments. Despite replicate radiations of Caribbean pupfishes in large, isolated, predator-free, saline lakes with few competitors, we found little parallelism between the radiations. Although craniofacial morphology diverged rapidly within trophic specialists in both radiations, we found almost zero overlap in the adaptive genes underlying trophic specialization (0.5%), potentially reflecting the divergent specialists within each radiation. In contrast to SSI, there was no behavior-first stage of adaptive divergence in Chichancanab specialists and, with the exception of craniofacial development, candidate adaptive gene ontologies potentially relevant to speciation were not shared between radiations, 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 (66)), adaptive introgression likely played a minimal role in the Chichancanab radiation. We found no overlap between adaptive variants in the specialists and tracts of introgression. Instead, over 30% of the candidate adaptive variants in the zooplanktivore, *C. simus*, likely arose from de novo mutations compared to only 2% in the SSI scale-eater specialist and none in the molluscivore specialist. One outstanding parallel feature was a clear stage of refinement in the adaptive

415 divergence of the zooplanktivore, marked by increased rates of de novo and nonsynonymous
416 adaptive mutations occurring among the most recent selective sweeps in this species, similar to
417 the scale-eating specialist on SSI. In contrast to *C. simus*, we identified no sufficiently
418 differentiated SNPs within *C. labiosus*, the bivalve specialist, to identify candidate genes.
419 Nonetheless, minimal divergence of this specialist in Chichancanab parallels the minimal
420 divergence of the molluscivore on SSI, each in contrast to a more divergent trophic specialist.
421

422 *Adaptive introgression did not contribute to speciation of Chichancanab pupfishes*

423 Genomic analyses have revealed that hybridization and adaptive introgression played a
424 major role in many adaptive radiations (3, 11, 16, 19–27, 67, 68), including the *Cyprinodon* SSI
425 radiation (26, 31). However, in Chichancanab, we found minimal evidence of introgression
426 overall and did not detect any signatures of adaptive introgression following the colonization of
427 Lake Chichancanab. Although our analyses indicate extensive introgression among *Cyprinodon*
428 and *Megupsilon* lineages, we did not find any introgression events into the root of the radiation,
429 nor has there been secondary gene flow from outgroup species other than the sister species to the
430 radiation, *C. artifrons*. Secondary gene flow from *C. artifrons* was almost entirely detected with
431 *C. beltrani* and, to a lesser extent, *C. labiosus.* Gene flow may be higher with these species
432 because *C. artifrons,* a coastal benthic feeding detritivore, is less ecologically divergent from
433 both *C. beltrani,* a benthic feeding detritivore, and *C. labiosus*, a benthic feeding bivalve
434 specialist. Secondary gene flow between *C. artifrons* and both *C. beltrani* and *C. labiosus*
435 overlapped with many genes under selection in *C. simus*, suggesting that secondary gene flow is
436 helping increase similarities between the benthic detritivores within the lake and the coastal *C.*

437 *artifrons* population rather than by bringing in adaptive variation for trophic specialization in *C.*
438 *simus*.

439 These results are remarkably different from the *Cyprinodon* adaptive radiation of SSI,
440 where there is strong evidence of introgression into the root of the radiation and substantial
441 evidence of secondary adaptive introgression into both trophic specialists from at least three
442 different biogeographic regions of the Caribbean and Atlantic coast (26, 31). Lake Chichancanab
443 is an endorheic basin with no above-ground hydrological connections to the ocean and is
444 substantially more geographically isolated than SSI, a small island in the approximate center of
445 the Caribbean. Furthermore, the sister species to the Chichancanab species flock, *C. artifrons*, is
446 abundant along the entire coast of the Yucatán peninsula, suggesting that any gene flow into
447 Chichancanab would likely first have to pass through *C. artifrons* populations. An additional
448 factor is the water chemistry of Lake Chichancanab which is near-saturation with calcite and
449 gypsum and may be a greater physiological barrier to marine pupfish colonists than the slightly
450 hypersaline lakes of SSI (33-36). Thus, geography and ecology play an important role in
451 determining the relative importance of adaptive introgression in radiations.

452

453 *De novo mutations primarily drive adaptive divergence of the zooplanktivore*

454 The importance of de novo mutations in adaptive radiations remains contentious, in part
455 due to the time needed for de novo beneficial mutations to arise (11, 15). Interestingly, we found
456 that 30.6% of the zooplanktivore adaptive variants were detected only within Chichancanab
457 species, suggesting their de novo origins within the basin. This proportion is 15.3 times greater
458 than the contribution of de novo mutations to adaptive divergence of the scale-eater in the SSI
459 radiation. A primary factor could be that the estimated effective population size for the

460 Chichancanab radiation is approximately twice that of the SSI radiation (26), indicating that
461 selection on de novo mutations is more effective, with a higher probability of fixation, all else
462 being equal (14). There may also be a lack of existing standing genetic variation within
463 Caribbean pupfish populations that is adaptive for zooplanktivory. In support of this hypothesis,
464 although radiations should proceed along genetic lines of least resistance (69), the primary axis
465 of craniofacial morphological variation along which the zooplanktivore diverged is far removed
466 from the first two major axes of craniofacial variation within Caribbean pupfish species (29).
467 Evolving along the line of greatest resistance instead of least resistance may require de novo
468 mutations. While the SSI scale-eater is also highly divergent, it is closer to existing craniofacial
469 variation within Caribbean pupfish species and in the direction of greatest variance along the first
470 two principal component axes (29). Similarly, mutational target size for the craniofacial features
471 of a zooplanktivore may be much smaller than the features of a scale-eater, resulting in less
472 standing genetic variation available for adaptation to this trophic niche. Alternatively, we may
473 not have detected these de novo mutations in other populations, but this is unlikely to explain all
474 de novo mutations given our extensive sampling of *Cyprinodon* species across the clade and
475 spanning the Caribbean (Figure 1B). Thus, our results highlight that de novo mutations may play
476 a more important role in driving young adaptive radiations than currently appreciated (12, 13),
477 especially when ecotypes are highly divergent from other species within the radiation.
478
479 *Minimal parallelism in the genes underlying adaptive divergence of trophic specialists*
480 We found only two candidate genes (*DNM1* and *SPTLC3*) among the 55 zooplanktivore
481 candidate genes in the Chichancanab radiation that overlapped with the hundreds of candidate
482 genes identified so far in the SSI radiation (26, 28, 30, 32, 42–46). Consistent with trophic

483 specialization requiring the pursuit of mobile prey, both genes are involved in motor neuron
484 development. *DNM1* is a dynamin with nervous tissue-specific expression that is crucial for the
485 proper formation of motor neuron axons (47). *SPTLC3* is involved in the biosynthesis of
486 sphingolipids (70), and suppression of *SPTLC3* in zebrafish results in motor neuron axon defects
487 (48).

488 The extremely limited number of shared adaptive genes is surprising given the high
489 conservation of gene function in both morphological development (71, 72) and behavior (73),
490 and the frequent parallelism of candidate genes across vertebrate adaptive radiations, particularly
491 for craniofacial genes. For example, *BMP4* is involved in diversification in both Galápagos
492 finches (74) and African cichlids (75, 76). Furthermore, within the SSI radiation, many of the
493 same genes show parallel patterns of differential expression even within divergent trophic
494 specialists (26, 43). Nonetheless, parallelism of shared regulatory networks, rather than specific
495 genes, is still a possibility.

496

497 *Rapid divergence in tooth morphology of the zooplanktivore through WNT10A*

498 Another reason for non-parallelism is that trophic specialists in each radiation have
499 diverged in different sets of traits during adaptation to different ecological niches (29).
500 Specifically, the Chichancanab radiation has diversified in tooth size over 130 times faster than
501 background rates, driven by the extremely short and narrow teeth of the zooplanktivore (Figure
502 S1; (29). We identified a well-characterized tooth development gene in the Wnt pathway,
503 *WNT10A*, with many de novo mutations in flanking (n = 2) and intronic (n = 13) regions.
504 Mutations in *WNT10A* are associated with tooth developmental disorders in humans (62) and
505 both knockout and overexpression experiments in sticklebacks and zebrafish have demonstrated

506 *WNT10A*'s role in teeth development and regeneration (61, 63). Consistent with *WNT10A*
507 playing a role in rapid divergence of tooth morphology, we found that inhibiting the Wnt
508 pathway in the *C. beltrani* generalist during pre-metamorphosis development resulted in
509 narrower teeth, similar to the reduced tooth sizes in the zooplanktivore *C. simus* (Figure 5C).
510 This indicates that the substantial number of de novo mutations in *WNT10A* may be suppressing
511 the Wnt pathway, resulting in a substantial reduction in tooth size in the zooplanktivore.
512
513 *Increased visual acuity in the zooplantkivore is consistent with unique selection on vision-related*
514 *genes*
515 One of the more striking features of both *Cyprinodon* radiations is the evolution of novel
516 trophic specialists from a generalist ancestor: the scale-eater (*C. desquamator*) in the SSI
517 radiation and the zooplanktivore (*C. simus*) in the Chichancanab radiation. Although these are
518 distinct dietary specializations, they share a similarity in that they require tracking free-
519 swimming prey. Indeed, four of the top 15 enriched GO terms of candidate genes in the SSI
520 scale-eater were related to eye development (26). Similarly, several of the candidate genes we
521 identified in the zooplanktivore were related to eye development, including *ASIC3*, *MAB21L2*,
522 and *CRIM1*. *CRIM1* is the most promising candidate because it is one of three candidate genes
523 with multiple nonsynonymous mutations (Figure 2). In zebrafish, a two-base pair deletion in
524 *CRIM1* resulted in abnormal lenses as well as a large eye-to-head ratio (54). Interestingly, one of
525 the defining morphological features of *C. simus* is its large eye-to-head ratio (33) (Figure 1A).
526 Consistent with candidate genes related to eye development, we confirmed that *C. simus* exhibits
527 a much stronger optomotor response than *C. beltrani* (Figure 5B), indicating that this specialist
528 has greatly improved visual acuity, even when reared in a common laboratory environment.

529 Although both radiations contained many vision-related candidate genes, selection on
530 opsin genes was unique to Chichancanab. Most notably, we found several adaptive variants
531 associated with *SWS2* (blue-sensitive), and *OPN1LW* (red-sensitive) in the zooplanktivore.
532 Interestingly, both opsins contain many candidate de novo mutations, and *SWS2* is the youngest
533 selective sweep in the zooplanktivore. One explanation could be that most Caribbean
534 *Cyprinodon* species are benthic and live in similar habitats, resulting in limited standing genetic
535 variation in opsins. Additionally, most environments (including SSI and Chichancanab) are
536 relatively shallow, so differences in depth—a common selective pressure on opsin evolution
537 (77)—are likely not a strong selective force on SSI. Thus, the main selective force acting on
538 vision in Chichancanab may be prey capture efficiency. In many fish species, zooplanktivory is
539 associated with higher expression of shorter wavelength opsins (e.g., *SWS1* (77)), whereas
540 piscivory is not often associated with changes in opsin expression (78, 79).
541
542 *Divergent sperm morphology in Chichancanab species is consistent with selection on*
543 *spermatogenesis genesis*
544 Although the role that premating sexual selection plays in speciation has historically been
545 the main focus (80, 81), there is both theoretical and empirical support for the role of postmating
546 sexual selection in maintaining reproductive isolation via postmating prezygotic reproductive
547 barriers (82–87). In the SSI radiation, species differ in pigmentation, and a handful of the
548 candidate adaptive variants fall in regulatory regions of genes associated with pigmentation (26),
549 but not spermatogenesis. Furthermore, there is evidence for premating isolation by female
550 preference in the scale-eater (88). However, in Chichancanab, the two species tested in our data

551 set (*C. labiosus* and *C. beltrani*) were not found to visually distinguish between different species
552 in the radiation in laboratory mate choice trials (89).
553 Unlike the SSI radiation, we did not find any candidate adaptive variants associated with
554 pigmentation in Chichancanab; instead, we identified adaptive variants associated with four
555 spermatogenesis genes, *ERC2*, *TBC1D25*, *C4orf22*, and *STK35*. Consistent with previous work
556 demonstrating the importance of these genes in spermatogenesis (57–60), we found that *C. simus*
557 does indeed have larger sperm heads than *C. beltrani*, suggesting a unique role for post-mating
558 prezygotic reproductive barriers in this radiation (Figure 5D).
559
560 *Parallelism in the stages of adaptation: refinement with de novo and nonsynonymous mutations*
561 Adaptation may occur in stages, with less pleiotropic cis-regulatory changes occurring
562 before potentially more pleiotropic amino acid changes (26, 90, 91). Selection is also predicted
563 to act on standing variation first (14, 92). In both radiations, we observed strong evidence for a
564 refinement stage in which coding substitutions were selected on last. In Chichancanab, these
565 late-sweeping coding substitutions were also de novo mutations (Figure 4). The simplest
566 explanation for parallelism in more recent ages of adaptive de novo mutations is the waiting time
567 required for beneficial mutations to (1) arise and (2) reach a substantial frequency within the
568 population (14, 92). These variants may only be selected in the later stages of an adaptive walk
569 because nonsynonymous mutations are often pleiotropic and likely deleterious in the ancestral
570 genetic background (14, 93). Thus, they may not be advantageous until later in the adaptive
571 walk, when sign epistasis may reverse negative pleiotropic effects (94–98). Our results highlight
572 that this final refinement stage is likely a general pattern during adaptive divergence.
573

Materials and Methods

Sampling and genotyping

In 2022, we collected Lake Chichancanab specimens (*C. labiosus*, *C. beltrani*, and *C. simus*) and *C. artifrons* (Figure 1C). Additional species were collected from the wild or sourced from captive breeding populations in previous years. 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. All specimens used in this study are catalogued in the Museum of Vertebrate Zoology Fishes collection (MVZ:Fish:1410-1499).

Individual DNA samples were sequenced using Illumina HiSeq 4000 and NovaSeq. We mapped reads to the UCB_CyNevMio_1.0 *Cyprindon nevadensis mionectes* reference genome (GCA_030533455.1). Variants were called using *GATK* (v.4.5.0.0) and filtered according to best practices (99), resulting in 23,960,536 SNPs. See SI for more details.

Population structure and demographic history of Yucatán Cyprinodon *pupfishes*

To assess population structure, we first pruned SNPs in linkage disequilibrium using *PLINK* (v. 1.9), resulting in 5,476,855 SNPs. With this linkage disequilibrium pruned data set, we visualized population structure with a PCA (*PLINK*) and *ADMIXTURE* (v1.3) (100).

We estimated demographic history using *MSMC2* (101) using three individuals with the highest mean depth. See SI for more details.

Identifying and characterizing candidate genes

We calculated F_{st} genome-wide, per site, and in non-overlapping 10-kb windows in *VCFtools* (v0.1.16) for all pairwise combinations of Yucatán *Cyprinodon* species (102). With the 10-kb windowed F_{st} values, we calculated a modified, normalized population branch statistic (PBS), PBS_{nj} (103). We classified SNPs as candidate adaptive variants 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 (41).

We classified variants within 20 kb of a gene (upstream or downstream) as flanking. For variants within a gene, we further classified them into intronic, synonymous, or nonsynonymous using *SnpEff* (104). To characterize the origin of candidate variants, we extracted the alleles from every species for each candidate variant. We classified variants that were found in any other populations outside of Lake Chichancanab as standing genetic variation. We further subdivided this group into variants that were found in other species but not in *C. artifrons*. Finally, if the alternative variant was only detected within Chichancanab species, we classified it as de novo. See SI for more details.

Introgression

We tested for evidence of introgression in Yucatán *Cyprinodon* species using *Dsuite* with *Cualac tessellatus* as the outgroup, calculating f_4-ratios, *D* statistics, and *f*-branch statistics (49). For species trios that showed evidence of introgression where *C. simus* was either P1or P2, we then looked for introgression regions across the genome by calculating the distance fraction (d_f (50)), in non-overlapping windows of 91 informative SNPs. To estimate introgression block

619 sizes, we merged neighboring windows that were introgression outliers and summed the total
620 size of the adjacent block. See SI for more details.
621
622 *Timing of selective sweeps*
623 For each candidate gene, we used *starTMRCA* (51) to estimate the age of selective
624 sweeps, extracting a 1-Mb region surrounding the candidate variant (51). We used a fixed
625 recombination rate of $2x10^{-8}$ (swordtail fishes; (105)) and a mutation rate of 1.56×10^{-8}
626 substitutions per base pair (Caribbean *Cyprinodon* species; 26). We then conducted separate
627 generalized linear models to test if the sweep age was correlated with the proportion of
628 nonsynonymous SNPs, proportion of de novo SNPs, and the number of SNPs. See SI for more
629 details.
630
631 *Optomotor response, sperm morphology, and Wnt inhibition experiment*
632 We tested for differences in visual acuity between the generalist/detritivore, *C. beltrani*,
633 and the zooplanktivore, *C. simus*, by conducting an optomotor response behavioral assay (56) on
634 lab-reared *C. beltrani* (F1; $n = 7$) and lab-reared *C. simus* (F10+; $n = 5$). Because the data were
635 nonparametric, we used a Wilcoxon rank-sum test to assess differences in optomotor response.
636 To test for differences in sperm morphology, we collected sperm samples from five
637 individuals per species in the lab (F10+*C. simus,* F1 *C. artifrons,* F1 *C. beltrani*). After fixing the
638 sperm sample in a 4% PFA solution stained with Rose Bengal, we plated and imaged sperm
639 under oil immersion with a 63x objective. We measured 10-30 sperm per individual in ImageJ
640 (106). We fit separate linear mixed-effects models for each trait of interest (sperm head,

641 midpiece, and flagellum lengths) with species as a fixed effect and a random intercept for
642 individual ID.

643 To confirm the role of the Wnt pathway in proper jaw and teeth development, we
644 conducted an experiment chemically inhibiting the Wnt pathway using iCRT14 during
645 metamorphosis (8dpf to 22 dpf) in *C. beltrani*. To control the exact time of development, we first
646 performed in vitro fertilizations and incubated developing embryos at ~26 °C until hatching
647 (~8dpf). On the day of hatching, we haphazardly split broods into control (DMSO) and
648 experimental groups (100 nanomolar iCRT14) to control for batch and family effects. After
649 euthanizing and staining fish with alizarin red (bone) and alcian blue (cartilage; (107)), we
650 quantified the number of ossified teeth and measured the tooth length and base width for the left
651 and right teeth closest to the mandibular symphysis using ImageJ (106). We fit separate linear
652 mixed-effects models for average tooth length and base width with experimental treatment as a
653 fixed effect and a random intercept for brood/replicate. For the number of ossified teeth, we used
654 a generalized linear mixed-effect model with a Poisson family and log-link function. See SI for
655 more details.

656

657 **Acknowledgments**

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

02-14062-1 and AUP-2021-07-14515) were approved by the University of California, Berkeley Animal Care and Use committee and the UC Davis Institutional Animal Care and Use committee. Specimens are deposited in the Museum of Vertebrate Zoology (MVZ:Fish:1410-1499).

References

1. R. G. Gillespie, *et al.*, Comparing adaptive radiations across space, time, and taxa. *J. Hered.* **111**, 1–20 (2020).

2. C. H. Martin, E. J. Richards, The paradox behind the pattern of rapid adaptive radiation: how can the speciation process sustain itself through an early burst? *Annu. Rev. Ecol. Evol. Syst.* **50**, 569–593 (2019).

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

4. D. Schluter, *The ecology of adaptive radiation* (OUP Oxford, 2000).

5. J. T. Stroud, J. B. Losos, Ecological opportunity and adaptive radiation. *Annu. Rev. of Ecol. Evol. Syst.* **47**, 507–532 (2016).

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

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

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

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

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

11. D. A. Marques, J. I. Meier, O. Seehausen, A combinatorial view on speciation and adaptive radiation. *Trends Ecol. Evol.* **34**, 531–544 (2019).

12. J. B. Pease, D. C. Haak, M. W. Hahn, L. C. Moyle, Phylogenomics reveals three sources of adaptive variation during a rapid radiation. *PLOS Biol.* **14**, e1002379 (2016).

13. B. W. Stone, C. A. Wessinger, Ecological diversification in an adaptive radiation of plants: The role of de novo mutation and introgression. *Mol. Biol. Evol.* **41**, msae007 (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. Biol. Sci.* **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. J. M. Humphries, *Cyprinodon verecundus,* n. sp., a fifth species of pupfish from Laguna Chichancanab. *Copeia* **1984**, 58–68 (1984).

37. 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. Fish.* **107**, 401–414 (2024).

38. 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).

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

40. 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. Phyl. Evol.* **6**, 143–149 (1996).

41. 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).

42. J. A. McGirr, C. H. Martin, Novel candidate genes underlying extreme trophic specialization in Caribbean pupfishes. *Mol. Biol. Evol.* **34**, 873 (2017).

43. 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).

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

45. 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).

46. M. F. Palominos, V. Muhl, C. H. Martin, Craniofacial-specific transcriptomics uncovers novel genes underlying jaw divergence in dietary specialist pupfishes. *Genetics* ***231***, iyaf207 (2025).

47. 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).

48. C. Gonzaga-Jauregui, *et al.*, Exome sequence analysis suggests that genetic burden contributes to phenotypic variability and complex neuropathy. *Cell Rep.* **12**, 1169–1183 (2015).

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

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

51. 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).

52. J. T. Streelman, P. D. Danley, The stages of vertebrate evolutionary radiation. *Trends Ecol. Evol.* **18**, 126–131 (2003).

53. W.-G. Li, T.-L. Xu, ASIC3 Channels in multimodal sensory perception. *ACS Chem. Neurosci.* **2**, 26–37 (2011).

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

55. B. N. Kennedy, *et al.*, Zebrafish *rx3* and *mab21l2* are required during eye morphogenesis. *Dev. Biol.* **270**, 336–349 (2004).

56. 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).

57. 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. of Androl.* **24** (2022).

58. 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).

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

60. 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).

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

62. S. Nawaz, *et al.*, WNT10A missense mutation associated with a complete odonto-onycho-dermal dysplasia syndrome. *Eur. J. Hum. Genet.* **17**, 1600–1605 (2009).

63. E. L. Benard, *et al.*, WNT10A is required for zebrafish median fin fold maintenance and adult unpaired fin metamorphosis. *Dev. Dyn.* **253**, 566–592 (2024).

64. J. L. Catrow, Y. Zhang, M. Zhang, H. Ji, Discovery of selective small-molecule inhibitors for the β-catenin/T-cell factor protein–protein interaction through the optimization of the acyl hydrazone moiety. *J. Med. Chem.* **58**, 4678–4692 (2015).

65. F. C. Gonsalves, *et al.*, An RNAi-based chemical genetic screen identifies three small-molecule inhibitors of the Wnt/wingless signaling pathway. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 5954–5963 (2011).

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

67. J. V. Peñalba, *et al.*, The role of hybridization in species formation and persistence. *Cold Spring Harb. Perspect. Biol.* **16**, a041445 (2024).

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

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

70. T. Hornemann, *et al.*, The SPTLC3 subunit of serine palmitoyltransferase generates short chain sphingoid bases. *J. of Biol. Chem.* **284**, 26322–26330 (2009).

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

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

73. C. J. Reaume, M. B. Sokolowski, Conservation of gene function in behaviour. *Philos. Trans. R. Soc. B.* **366**, 2100–2110 (2011).

74. A. Abzhanov, M. Protas, B. R. Grant, P. R. Grant, C. J. Tabin, Bmp4 and morphological variation of beaks in Darwin's finches. *Science* **305**, 1462–1465 (2004).

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

76. Y. Terai, N. Morikawa, N. Okada, The evolution of the pro-domain of bone morphogenetic protein 4 (Bmp4) in an explosively speciated lineage of East African cichlid fishes. *Mol. Biol. Evol*. **19**, 1628–1632 (2002).

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

78. C. M. Hofmann, *et al.*, The eyes have it: regulatory and structural changes both underlie cichlid visual pigment diversity. *PLOS Biol.* **7**, e1000266 (2009).

79. K. E. O'Quin, C. M. Hofmann, H. A. Hofmann, K. L. Carleton, Parallel evolution of opsin gene expression in African cichlid fishes. *Mol. Biol. Evol.* **27**, 2839–2854 (2010).

80. M. G. Ritchie, Sexual selection and speciation. *Annu. Rev. Ecol. Evol. Syst.* **38**, 79–102 (2007).

81. 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).

82. 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).

83. M. D. Garlovsky, *et al.*, Synthesis and scope of the role of postmating prezygotic isolation in speciation. *Cold Spring Harb. Perspect. Biol*. **16**, a041429 (2023).

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

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

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

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

88. 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).

89. U. Strecker, A. Kodric-Brown, Mate recognition systems in a species flock of Mexican pupfish. *Journal of Evolutionary Biology* **12**, 927–935 (1999).

90. L. Campagna, *et al.*, Repeated divergent selection on pigmentation genes in a rapid finch radiation. *Science Advances* **3**, e1602404 (2017).

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

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

93. 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).

94. 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).

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

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

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

98. 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).

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

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

101. 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.

896 102. P. Danecek, *et al.*, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158
897 (2011).

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

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

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

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

908 107. M. Walker, C. Kimmel, A two-color acid-free cartilage and bone stain for zebrafish larvae.
909 *Biotech. Histochem.* **82**, 23–28 (2007).

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

Supporting Information for

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

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

[*]**Corresponding author:** Matthew Kustra

Email: matthewckustra@berkeley.edu

The supporting information includes:

Supporting text (SI methods)

Tables S1 to S4

Figures S1 to S10

SI References

SI Methods

Sampling

We collected Lake 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 (Fig. 1C). *C. labiosus* and *C. beltrani* were sampled from both sites for this study, whereas *C. simus* was only found at La Presumida. *C. artifrons* were collected by cast net from a 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 (*Megupsilon aporus*), and Frans Vermuelen (*Cyprinodon dearborni*). 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. All specimens used in this study are catalogued in the Museum of Vertebrate Zoology Fishes collection (MVZ:Fish:1410-1499).

The London Zoological Society provided our live colony of *C. simus* in 2015, originating from a much earlier field collection, which we subsequently raised for ten generations. We collected live colonies of *C. beltrani* and *C. artifrons* in 2024 from La Presumida, Lake Chichancanab and Sisal, respectively, over four days using a 5 x 1.3 m seine net with 1.6 mm mesh size.

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) (1). Reads were mapped to the UCB_CyNevMio_1.0 *Cyprindon nevadensis mionectes* reference genome (GCA_030533455.1) with *bwa mem* (v0.7.17) (2). We followed the *GATK* (v.4.5.0.0) genotyping pipeline to call variants (119). 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) (4). 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 (7).

After hard-filtering and variant calling, we filtered the data set further by removing SNPs with minor allele frequency < 0.05, more than 10% missing, depth less than 10x, and GQ less than 20 using *BCFtools* (v1.16)(8). This resulted in a final data set that included 23,960,536 SNPs.

Population structure of Yucatán Cyprinodon *pupfishes*

To assess population structure, we first pruned SNPs in linkage disequilibrium using *PLINK* (v. 1.9) with the following parameters: "-indep-pairwise 50 5 0.2" (9). This filtered our data set from 23,960,536 SNPs to 5,476,855 SNPs. With this linkage disequilibrium pruned data set, we first filtered the full VCF file to only include relevant Yucatán *Cyprinodon* species (i.e., *C. artifrons, C. beltrani, C. labiosus,* and *C. simus*), then conducted a PCA using *PLINK*. Next,

999 we used *ADMIXTURE* (v1.3) to determine the optimal number of population clusters that best fit
1000 the data and to assign individuals to their corresponding population clusters (10). We performed
1001 this analysis and calculated cross-validation error for the number of clusters K = 1-6. Models in
1002 ADMIXTURE with K=2-4 were equally supported (Figure S2A). However, given morphological
1003 distinctness and the results of the PCA, we present the results for when K = 4 (Figure 1D).
1004
1005 *Demographic history of Yucatán* Cyprinodon *pupfish*
1006 We estimated demographic history using *MSMC2* (11). For each Yucatán pupfish species,
1007 we selected three individuals with the highest mean depth and generated single-sample VCF
1008 files, individual mask files, and "mapability" mask files. We then ran *MSMC2* with default
1009 parameters except we set phasing to "unphased" and "P_PAR=8*1+25*1+1*2+1*3." For
1010 plotting, we used a generation time of 1 year and a mutation rate of 1.56×10^{-8} substitutions per
1011 base pair, estimated from lab-reared pedigrees of Caribbean *Cyprinodon* species (12).
1012
1013 *Identifying candidate genes*
1014 We calculated F_{st} genome-wide, per site, and in non-overlapping 10-kb windows using
1015 the "--weir-fist-pop" function in *VCFtools*(v0.1.16) for all pairwise combinations of Yucatán
1016 *Cyprinodon* species (13). We chose 10-kb windows because linkage disequilibrium substantially
1017 decayed at this distance (Figure S3), and these windows allowed us to quantify fine-scale
1018 genomic variation. With the 10-kb windowed F_{st} values, we calculated a modified population
1019 branch statistic (PBS) for four taxa, PBS_{nj} (14). This statistic is useful because it makes no
1020 assumptions about species topology or polarization (14). We then normalized PBS_{nj} by the total
1021 tree length, as normalized *PBS* is most effective and specific at identifying selective sweeps (15).

1022 We classified SNPs as candidate adaptive variants if they were fixed or nearly fixed (per-site F_{st}
1023 >0.95) compared to other species within the radiation and were among the top 1% outliers out of
1024 all PBS_{nj} windows, following the threshold used in (15).
1025
1026 *Introgression*
1027 Using the full VCF file, which contained a wide range of species in the family
1028 Cyprinodontidae, we tested for evidence of introgression in Yucatán *Cyprinodon* species using
1029 *Dsuite* (16). We first excluded any samples/species that had an average depth < 4. We then
1030 calculated all possible trios with a species tree based on the most recent phylogeny (17) with
1031 *Cualac tessellatus* as the outgroup. We calculated Z-scores to assess the significance of *D*
1032 statistics using 1,000 jackknife blocks and considered a Z-score greater than three as evidence
1033 for significant introgression (18). We only focused our primary analysis on species trios where
1034 P1-P2 included species within the Lake Chichancanab radiation. We used the other trio
1035 calculations to calculate *f*-branch (f_b) statistics to account for correlated f_4-ratio scores and to
1036 inform the timing of introgression events (16).
1037 For species trios that showed evidence of introgression where *C. simus* was either P1or
1038 P2, we then calculated the distance fraction (d_f), a robust metric of introgression for windowed
1039 analyses (19), in non-overlapping windows of 91 informative SNPs. We chose this number of
1040 SNPs because it resulted in windows of approximately 10 kb. We considered windows to be
1041 introgressed if they were in the top X% of absolute d_f value, where X% was the percentage of
1042 estimated genome wide introgression (e.g, f_4-ratio) for that focal group, following (20, 21). To
1043 estimate introgression block sizes, we merged neighboring windows that were introgression
1044 outliers and summed the total size of the adjacent blocks.

1045

1046 *Characterizing candidate genes*

1047 After identifying candidate variants, we then used *Bedtools* (v2.31) to see if variants
1048 occurred near a gene (within 20 kb) or within a gene (22) with a GFF annotation file for the *C.*
1049 *nev. mionectes* reference genome (Tian et al. in prep). We classified variants within 20 kb of a
1050 gene (upstream or downstream) as flanking. For SNPs within a gene, we further classified them
1051 into intronic, synonymous, or nonsynonymous using *SnpEff* (23).

1052 To characterize the origin of candidate variants, we extracted the alleles from every
1053 species for each candidate variant. We classified variants that were found in any other
1054 populations outside of Lake Chichancanab as standing genetic variation. We further subdivided
1055 this group into variants that were found in other species but not in *C. artifrons*. If variants fell
1056 within an introgression window, we classified them as "introgressed." Finally, if the alternative
1057 variant was only detected within Chichancanab species, we classified it as de novo.

1058

1059 *Timing of selective sweeps*

1060 To gain insight into the timing of selective sweeps in *C. simus*, we used *starTMRCA* (24).
1061 Because many genes contained multiple candidate variants, for this analysis, we selected the
1062 variant that was in the median position of each gene. We then extracted a 1-Mb region
1063 surrounding that variant and removed all sites that contained missing data. For simplification, we
1064 only used the generalist/detritivore, *C. beltrani*, as the "reference" population and the
1065 zooplanktivore, *C. simus*, as the "selected" population. We used a fixed recombination rate of
1066 $2x10^{-8}$ (swordtail fishes; (25)) and a mutation rate of 1.56×10^{-8} substitutions per base pair
1067 (Caribbean *Cyprinodon* species; 26). We then ran five separate Markov chains for 30,000

1068 iterations with a proposal standard deviation of 150 (preliminary analyses changing this
1069 parameter had minimal influence on results). We then discarded the first 10,000 iterations of each
1070 chain as burn-in.

1071 To test for stages of adaptation, we conducted separate generalized linear models to test if
1072 the sweep age was correlated with the proportion of nonsynonymous SNPs, proportion of de
1073 novo SNPs, and the number of SNPs. For the generalized linear regressions of proportion of
1074 nonsynonymous SNPs and proportion of de novo SNPs, we used a quasi-binomial family to
1075 account for overdispersion with a logit link function. For the number of SNPs, we used a quasi-
1076 Poisson family with a log link function. To test for the significance of sweep age, we used a
1077 likelihood ratio test comparing a model with the estimated sweep age as an effect compared to a
1078 null model.

1079

1080 *Optomotor response*

1081 We tested for differences in visual acuity between the generalist/detritivore, *C. beltrani*,
1082 and the zooplanktivore, *C. simus*, by conducting an optomotor response behavioral assay (26).
1083 We tested lab-reared *C. beltrani* (F1, $n = 7$) and lab-reared *C. simus* (long-term laboratory colony
1084 exceeding ten generations in the lab, $n = 5$). Each fish was tested consecutively with each control
1085 or treatment for 1-minute observation periods, with the control observation occurring first,
1086 followed by presentation of all-black (positive control) and black-and-white spinning bars
1087 (treatment: acuity trial). The order of presentation for positive control and treatment was
1088 alternated in each trial. Fish were placed in a suspended cylindrical plastic container 0.3 m in
1089 diameter that remained stationary during the course of the trials. Spinning black-and-white or all-

1090 black paper was rotated outside the clear plastic container at ~90 RPM during each observation
1091 period.
1092 Due to the data being nonparametric (count, zero-inflated, unequal variance between
1093 groups), we conducted a Wilcoxon rank sum test for the on-banded (all black) and on-banded
1094 portions separately. Because we had an a priori hypothesis that *C. simus* would display a stronger
1095 optomotor response (greater visual acuity), we calculated *p*-values with a one-tailed test.
1096
1097 *Sperm morphology*
1098 To test for differences in sperm morphology, we collected sperm samples from five
1099 individuals per species in the lab (*C. simus, C. artifrons, C. beltrani*). We first anesthetized the
1100 fish with a solution of MS-222. After the fish was anesthetized, we collected 1μL of milt (fish
1101 semen) with a 1μL glass microcapillary tube and fixed the sample overnight in a 4% PFA
1102 solution stained with Rose Bengal. We then plated sperm on a positively charged slide with a
1103 sealed coverslip. We imaged sperm with oil immersion with a 63x objective and measured 10-30
1104 sperm per individual in ImageJ (27). Due to poor sperm samples from some individuals (not
1105 sufficient quantities), we ended up with the following sample sizes: three *simus*, four *artifrons*,
1106 and five *beltrani*.
1107 To retain information about variation within individuals, we fit separate linear mixed-
1108 effects models for each trait of interest (sperm head, midpiece, and flagellum lengths) with
1109 population as a fixed effect and a random intercept for individual ID, using *LME4* (28). To assess
1110 the significance of the model, we used a Type II ANOVA with a Wald *F* test using Kenward-
1111 Roger degrees of freedom. If there was a significant effect of population, we conducted post hoc
1112 pairwise tests using the *Emmeans* package (29) and corrected for multiple comparisons using the

1113 Tukey method. We did not correct for phylogeny due to how closely related the species are, the
1114 limited sample size, and the fact that the only significant differences were between sister species.
1115
1116 *WNT experiment*
1117 To confirm the role of the WNT pathway in proper jaw and teeth development, we
1118 conducted an experiment chemically inhibiting the WNT pathway using iCRT14 during
1119 metamorphosis in *C. beltrani*. To control the exact time of development, we first performed in
1120 vitro fertilizations and incubated developing embryos at ~26 until hatching (~8dpf). On the day
1121 of hatching, we haphazardly split broods into control and experimental groups to control for
1122 batch and family effects. We raised experimental groups in tank water with a 100 nanomolar
1123 concentration of iCRT14 and control groups in a 100 nanomolar concentration of DMSO
1124 (solution used to dilute stock iCRT14). We ran the experiment for two weeks until individuals
1125 were 22dpf, changing the solution every 3-4 days. We performed this experiment on four
1126 separate broods (~12 individuals/treatment). One brood was excluded due to improper
1127 development (no ossification of the spine) in the controls.
1128 After the experiment, we euthanized fish with MS-222 and stained fish with alizarin red
1129 (bone) and alcian blue (cartilage) following a modified published protocol (30). After staining,
1130 we prepared whole mounts and imaged fish with fluorescent microscopy (Zeiss LSM880 FCS)
1131 to quantify ossification of the jaw and teeth (31). We then dissected out the lower jaw and
1132 prepared flat mounts to measure teeth morphology. Specifically, we measured the tooth length
1133 and base width for the left and right teeth closest to the mandibular symphysis using ImageJ (27).
1134 For statistical analyses, we took the mean of the left and right tooth length and base width.

We fit separate linear mixed-effects models for tooth length and base width with experimental treatment as a fixed effect and a random intercept for brood/replicate, using *LME4* (28). For the number of ossified teeth, we used a generalized linear mixed-effect model with a Poisson family and log-link function. To assess the significance of these models, we used a likelihood ratio test comparing it to a null model without treatment as a fixed effect.

SI Tables

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

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

1146 **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

1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159

1160 **Table S3.** Fixed SNPs between *C. simus* and *C. beltrani* + *C. labiosus*. Bold rows indicate SNPs within top 1% of PBS windows.

1161 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 (32).
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 (33).
Chr 3: 32286638	RBMS3	14339	22.93	
Chr 3: 32286614	RBMS3	14363	22.93	
Chr 11: 27324192	GRIA3	Intron	2.85	
Chr 11: 27336469	GRIA3	2772	3.18	Glutamate receptor
Chr 11: 27336472	GRIA3	2775	3.18	
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 (34)
Chr 24: **12693158**	**KDM6A**	**Intron**	**0.05**	**Kabuki syndrome (craniofacial development; 35)**

1162

1163 **Table S4.** Summary statistics of the distance between closest adaptive candidate SNPs within the same
1164 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

1165

SI Figures

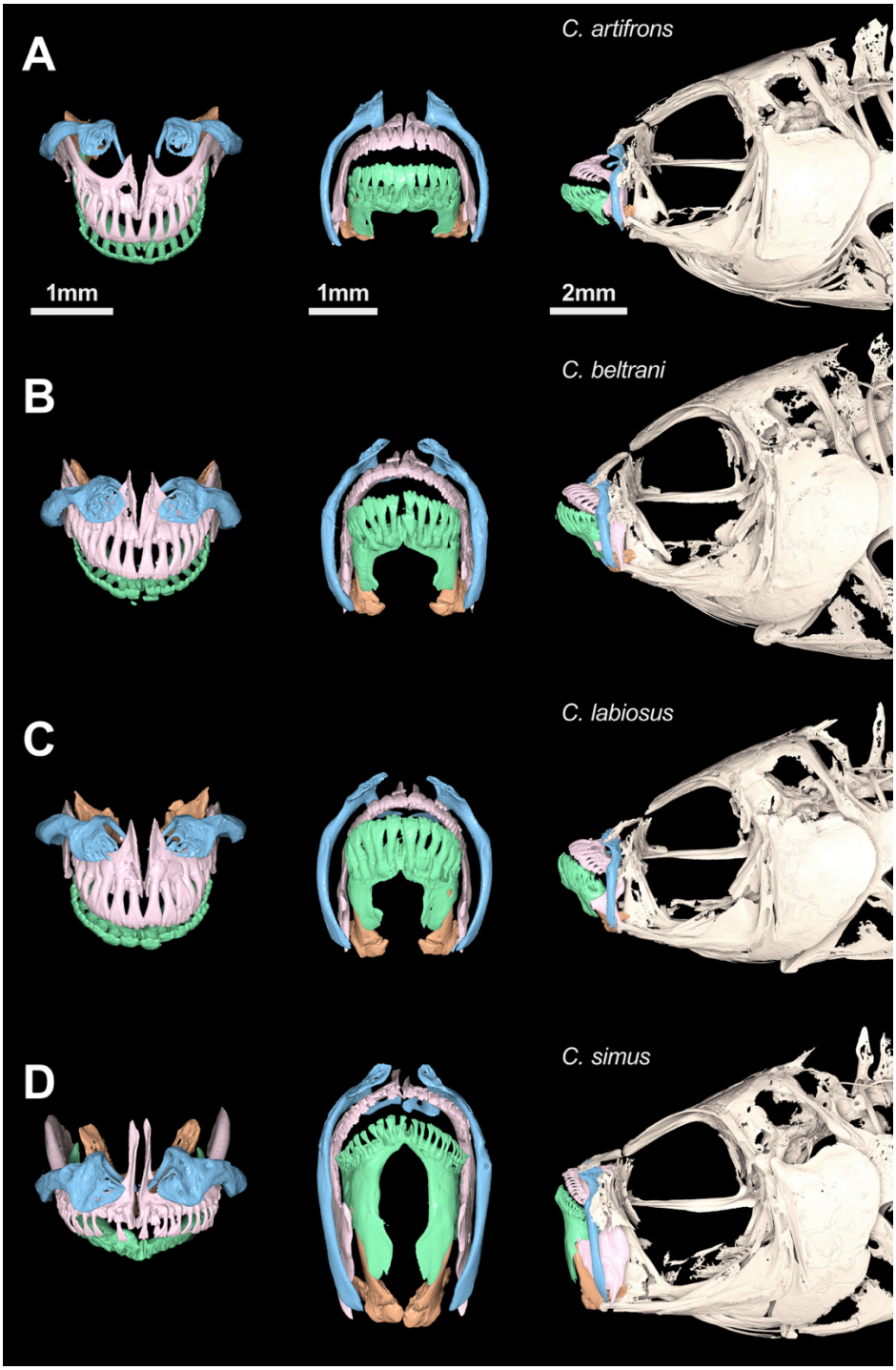


Figure S1. µCT scans of Yucatán pupfish with maxilla colored in blue, premaxilla in pink, dentary in green, and articular in orange.

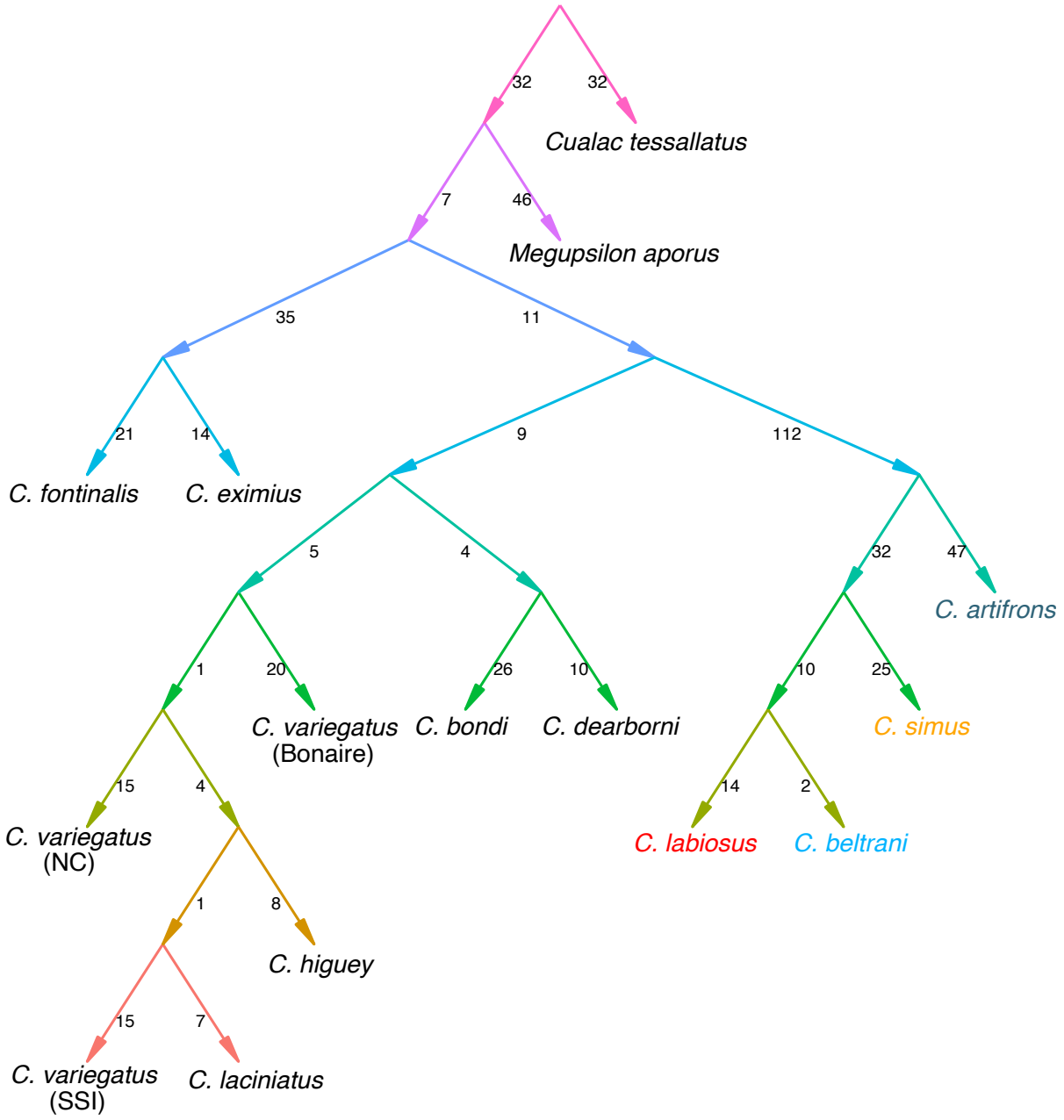


Figure S2. Best fitting admixture graph assuming zero admixture events. Admixture graph was found using the *findGraphs* function in *ADMIXTOOLS2* (36). Number on edges represent drift weights.

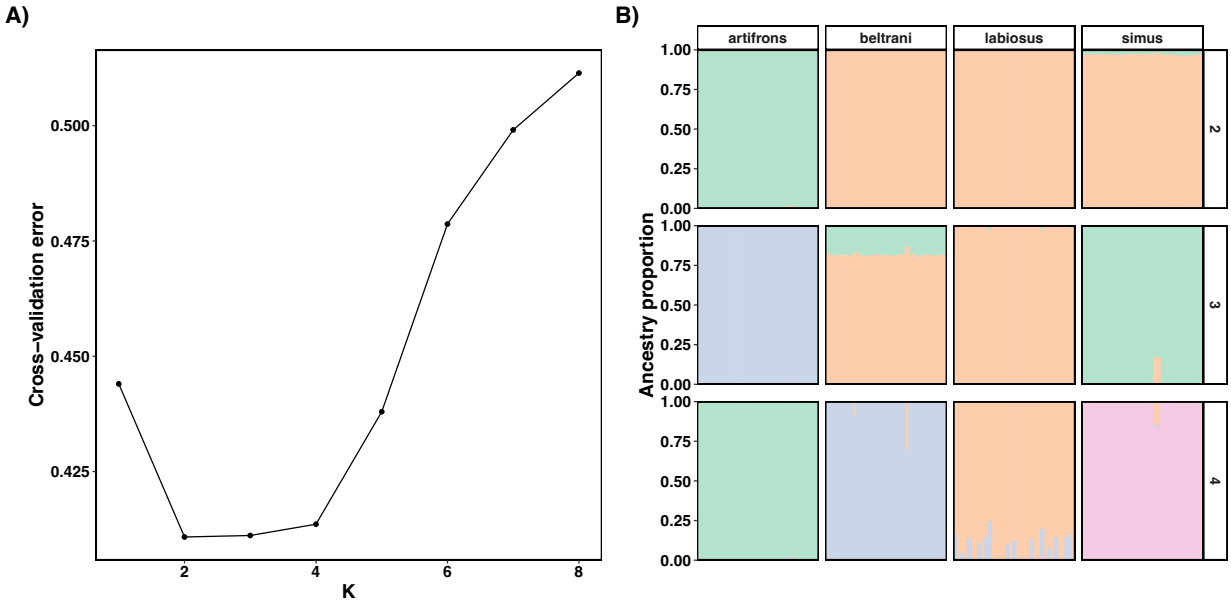


Figure S3. ADMIXTURE results support genetic differentiation of Yucatán Cyprinodon pupfish species. (A) Cross-validation error for different values of K clusters, equally support 2, 3, and 4 clusters. (B) *ADMIXTURE* plots showing the clusters at different values of K.

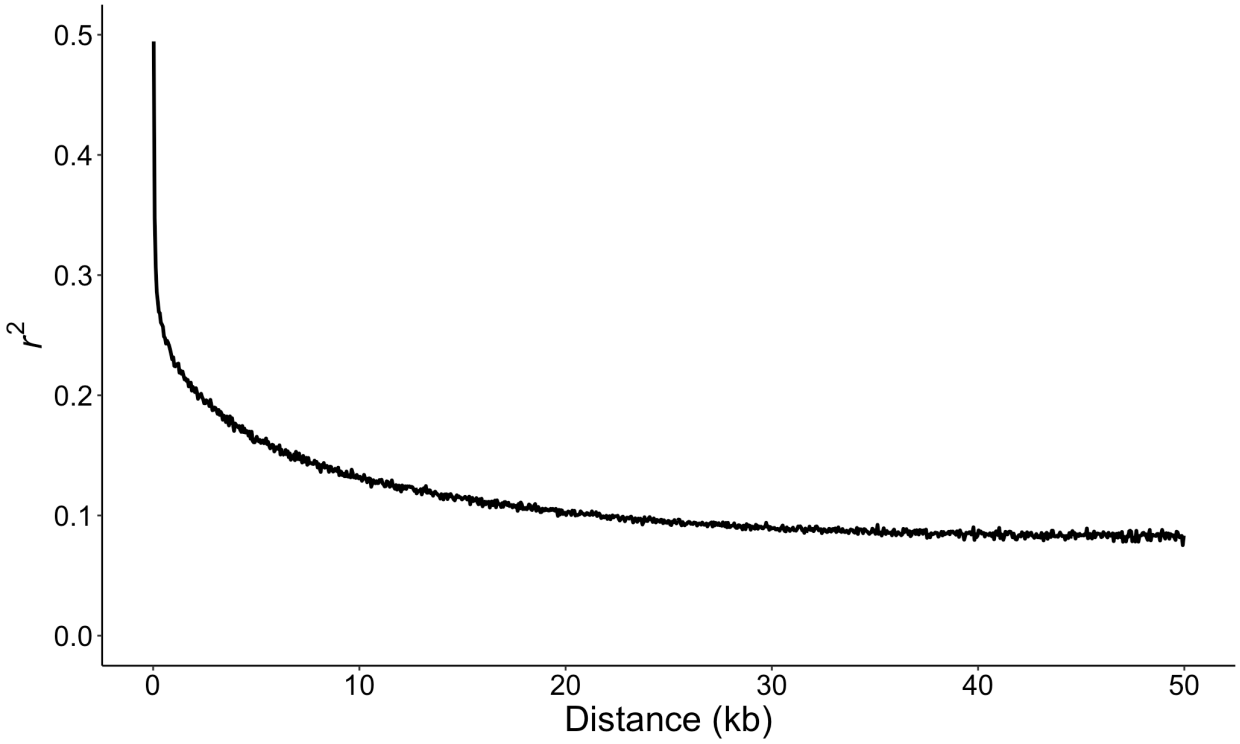
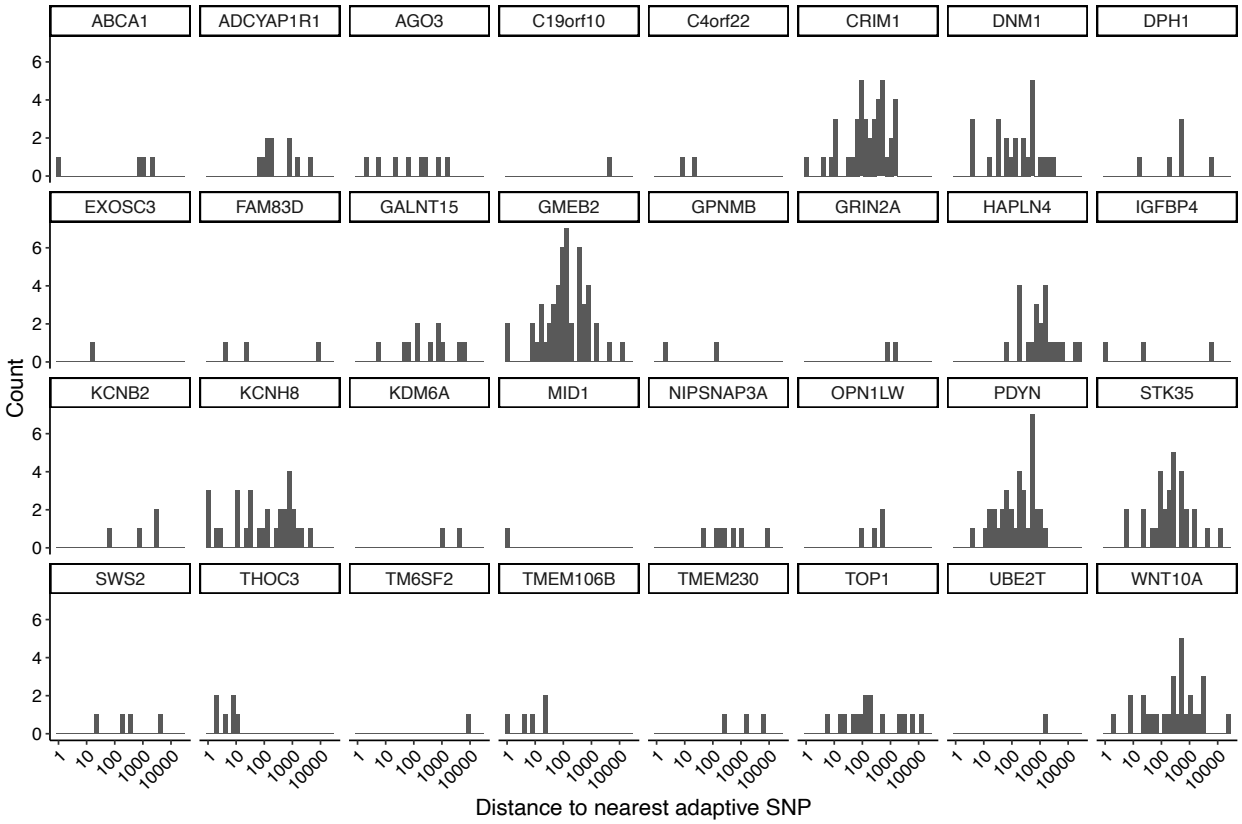


Figure S4. Linkage decay occurs rapidly in *Cyprinodon simus*. Linkage decay was calculated in windows of 1000 kb using *PLINK* (v1.9). The results graphed are a random sample of all calculated r^2 values taken for plotting purposes (n = 9,645 r^2 values).



1211

1212 **Figure S5. Adaptative SNPs associated with the same gene are generally spaced apart.**

1213 Distribution of the distance to nearest adaptive SNP within a gene. Note the log scale of the x-

1214 axis.

1215

1216

1217

1218

1219

1220

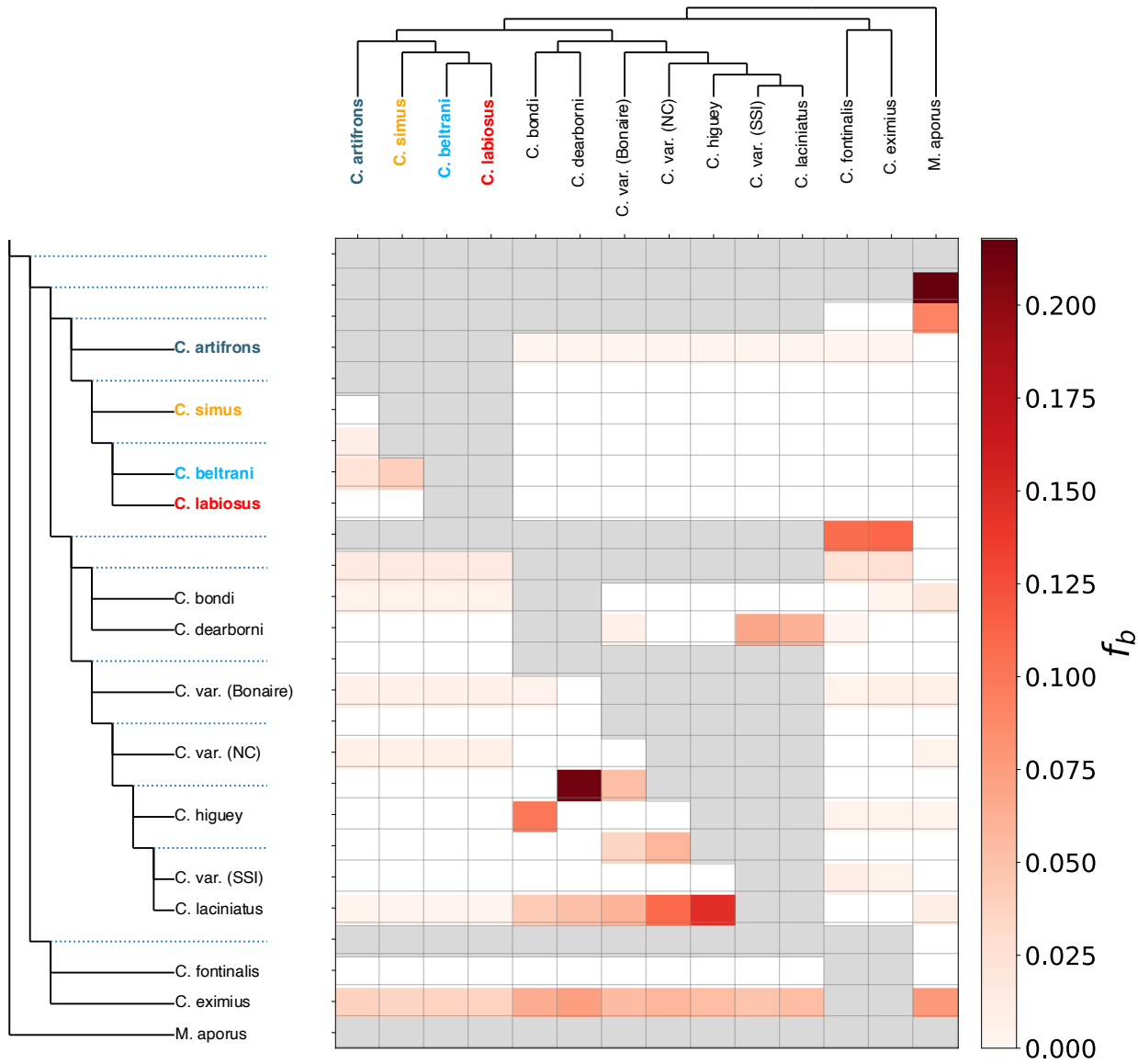


Figure S6. *f*-branch (f_b) show that introgression with desert pupfish occurred before the Chichancanab radiation. Boxes with red color represent significant introgression between P3 species (top) and the species/internal branch on the left. For visualization, non-significant f_b were changed to zero 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.

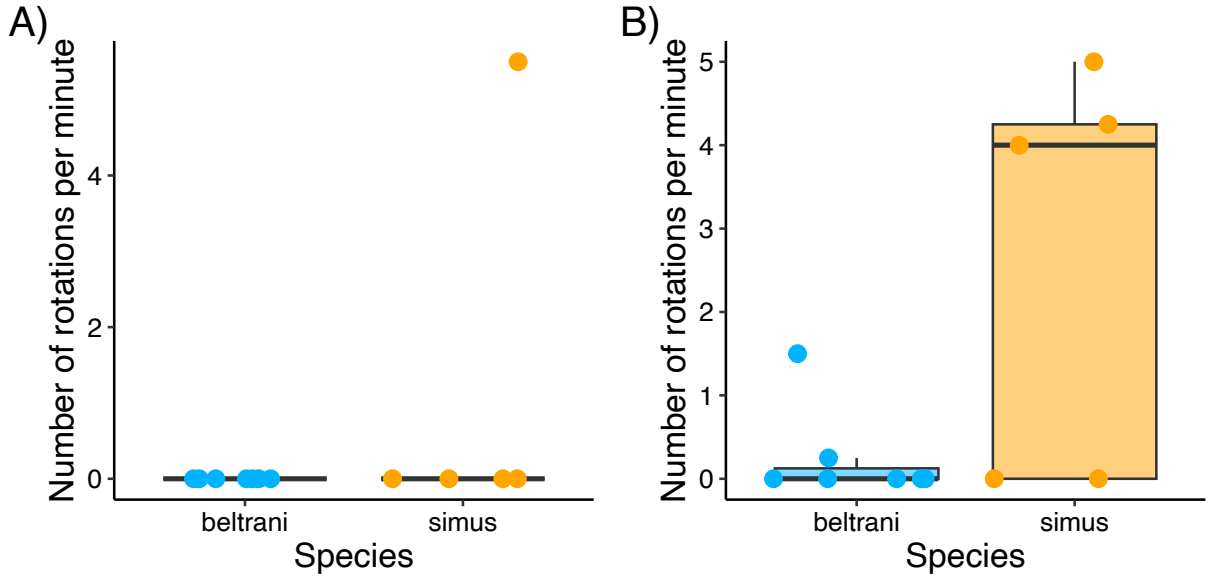


Figure S7. **Both species did not have a strong response to either control scenario in the optomotor response experiment. (A)** Number of complete rotations fish swam when the machine was off and **(B)** when the machine was on but the apparatus was all black**.** Boxplots and points of the number of 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.

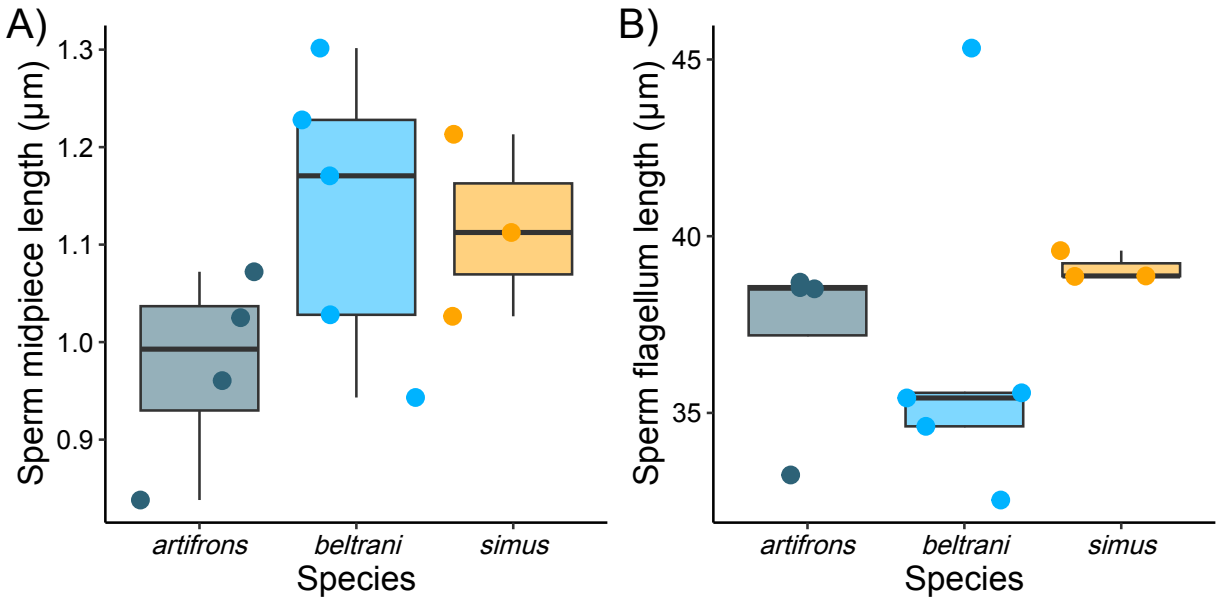


Figure S8. **Species do not differ in sperm midpiece size or sperm flagellum length.** Boxplots and points of the mean sperm morphological estimate for an individual.

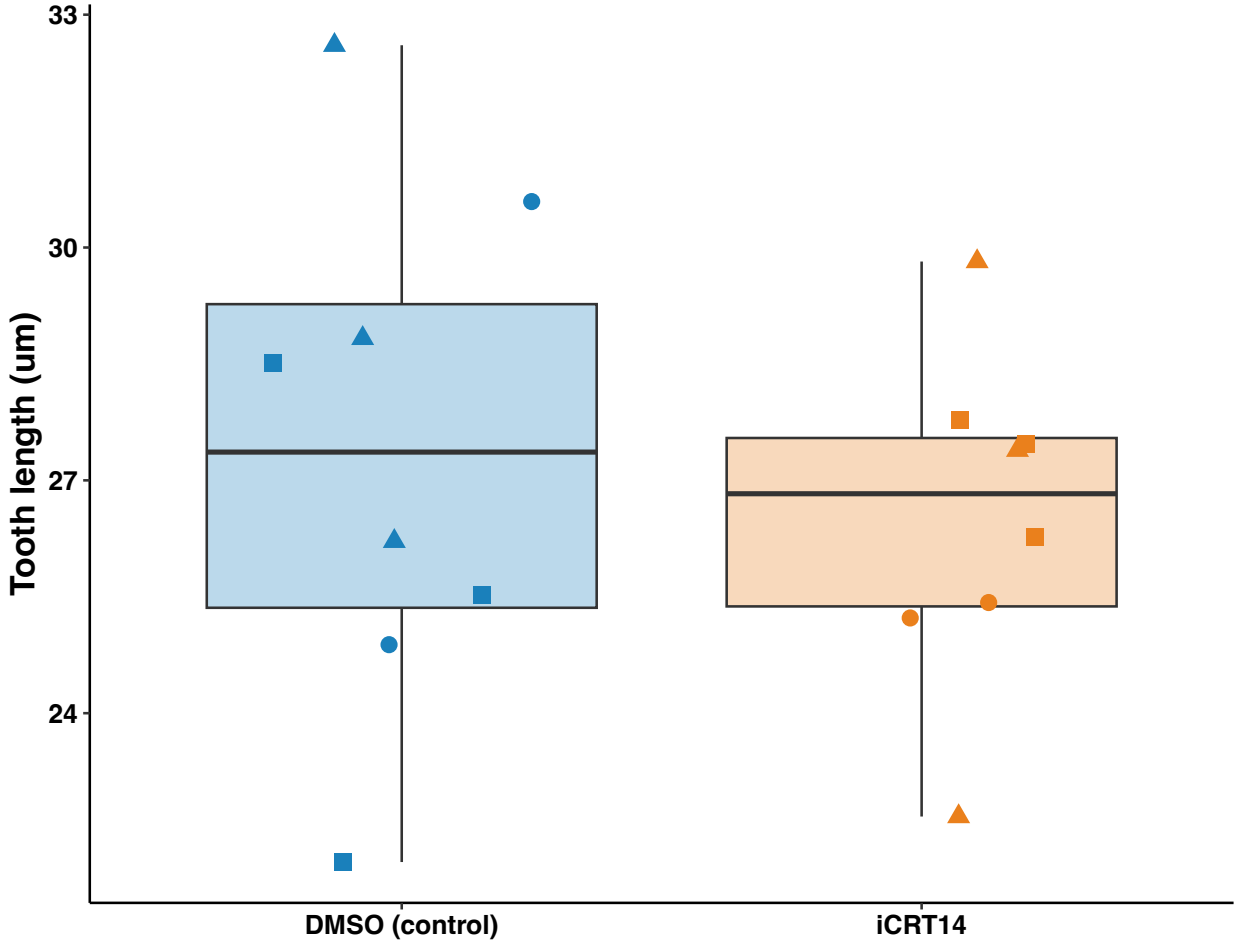


Figure S9. Tooth length was not affected by Wnt inhibition. Boxplot and jittered points of average tooth length of the left and right teeth closest to the mandibular symphysis for *C. beltrani* larvae treated with either a control or Wnt inhibitor (iCRT14). Shapes represent different split-brood replicates.

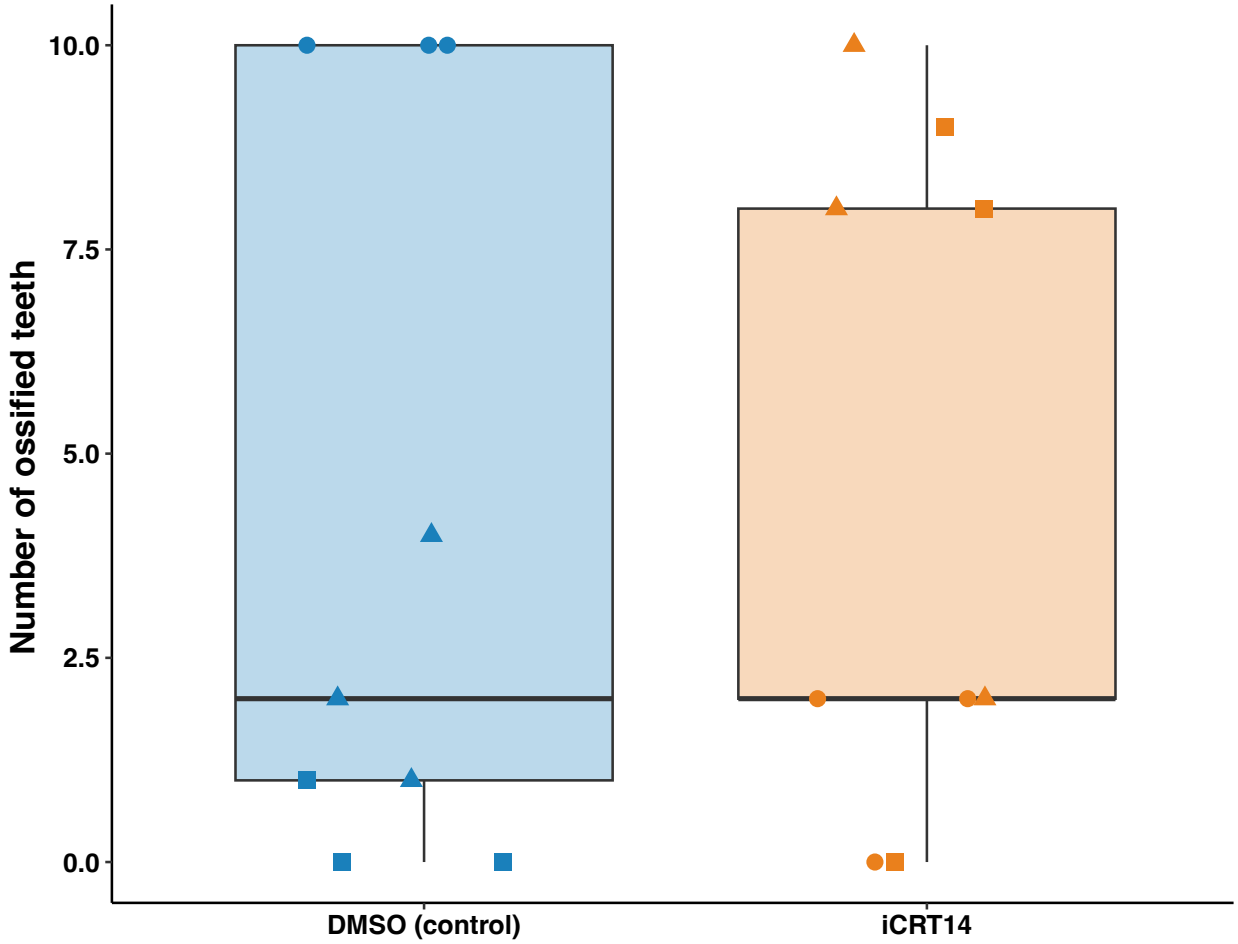


Figure S10. Number of ossified teeth was not affected by Wnt inhibition. Boxplot and jittered points of number of ossified teeth in *C. beltrani* larvae treated with either a control or Wnt inhibitor (iCRT14). Shapes represent different split-brood replicates.

SI References:

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

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

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

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

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

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

7. D. Caetano-Anolles, GenotypeGVCFs and the death of the dot (obsolete as of GATK 4.6.0.0). *GATK* (2024). Available at: https://gatk.broadinstitute.org/hc/en-us/articles/6012243429531-GenotypeGVCFs-and-the-death-of-the-dot-obsolete-as-of-GATK-4-6-0-0.

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

9. 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).

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

11. 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.

12. 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).

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

14. 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).

15. 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).

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

17. 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).

18. E. Y. Durand, N. Patterson, D. Reich, M. Slatkin, Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* **28**, 2239–2252 (2011).

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

20. C. Feng, J. Wang, A. Liston, M. Kang, Recombination variation shapes phylogeny and introgression in wild diploid strawberries. *Mol. Biol. Evol.* **40**, msad049 (2023).

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

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

23. 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).

24. 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).

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

26. 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).

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

28. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).

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

30. M. Walker, C. Kimmel, A two-color acid-free cartilage and bone stain for zebrafish larvae. *Biotech. Histochem.* **82**, 23–28 (2007).

31. R. Koita, S. Oikawa, T. Tani, M. Matsuda, A. Kawamura, Live visualization of calcified bones in zebrafish and medaka larvae and juveniles using calcein and alizarin red S. *Bio-protocol* **14** (2024).

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

33. C.S. Jayasena, M.E. Bronner, Rbms3 functions in craniofacial development by posttranscriptionally modulating TGF-β signaling. *J. Cell. Biol.* **199**(2012)

34. 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).

35. P.M. Van Laarhoven, Kabuki syndrome genes KMT2D and KDM6A: functional analyses demonstrate critical roles in craniofacial, heart and brain development. *Hum. Mol. Genet.* **24**, (2015).

36. R. Maier, *et al.*, On the limits of fitting complex models of population history to f-statistics. *eLife* **12**, e85492 (2023).