Title:

Diversity begets diversity through shared adaptive genetic variation: discovery of a cryptic wide-mouthed scale-eater

Authors: Emilie J. Richards[1] and Christopher H. Martin[1]

Affiliations: [1]University of California Berkeley

Correspondence: *chmartin@berkeley.edu*

Keywords: ecological speciation, ecomorph, adaptive radiation, novelty, innovation, selective sweep

Summary

Adaptive radiations involve astounding bursts of phenotypic, ecological, and species diversity. However, the microevolutionary processes that underlie the origins of these bursts are still poorly understood. We report the discovery of a cryptic intermediate 'wide-mouth' scale-eating ecomorph in a recent radiation of *Cyprinodon* pupfishes which provides crucial information about the evolutionary and ecological transition from a widespread algae-eating generalist to a novel microendemic scale-eating specialist. We first show that this ecomorph occurs in sympatry with generalist *C. variegatus* and scale-eating specialist *C. desquamator* across several hypersaline lakes on San Salvador Island, Bahamas, but is genetically differentiated, morphologically distinct when reared in a common garden, and sometimes consumes scales. We then compared the timing of selective sweeps on shared and unique adaptive variants in both scale-eating species to characterize the evolutionary path to scale-eating. We predicted that adaptation to the intermediate 'wide-mouth' scale-eating niche aided in the rapid divergence of the more specialized scale-eater *C. desquamator*. Therefore, selection for shared adaptive variants would occur first in 'wide-mouth'. Contrary to our prediction, four of the six sets of shared adaptive alleles in both scale-eating species swept significantly earlier in *C. desquamator*. Adaptive introgression from the specialist into the 'wide-mouth' ancestor may have resulted in parallel evolution of their dietary niche. Conversely, no adaptive alleles for scale-eating were reused in a third sympatric specialist *C. brontotheriodes*, despite sharing 9% of hard selective sweeps. Our work provides a microevolutionary framework for investigating how 'diversity begets diversity' during adaptive radiation.

Main Text

Introduction

Rapid bursts of diversification and repeated bouts of speciation like those seen in adaptive radiations contradict many of our existing mechanistic models for understanding speciation (reviewed in (Martin and Richards 2019)). For example, many speciation models predict that diversification should slow with time as available niche space becomes increasingly subdivided and disruptive selection becomes weaker with each recurrent speciation event (e.g. (Dieckmann and Doebeli 1999; Polechová et al. 2005; Bolnick 2006; Ito and Dieckmann 2007)). Diversification on complex adaptive landscapes with multiple empty fitness peaks corresponding to different niches provides an alternative mechanism to niche subdivision and might better explain how a rapid burst of radiation can occur (Kondrashov and Kondrashov 1999; Gavrilets 2004, 2014; Gavrilets and Vose 2005). However, empirical fitness landscapes present the new problem of how populations manage to escape local optima, cross fitness valleys, and access new fitness peaks (Wright 1932; Arnold et al. 2001; Burch and Chao 2004; Svennson and Calsbeek 2012; Martin and Wainwright 2013b; Martin 2016; Martin and Gould 2020).

Colonizing new fitness peaks on the adaptive landscape often requires transitions in behaviors, morphological traits, or a combination of the two that allow organisms to adapt to new ecological niches (Calsbeek and Irschick 2007). Some of the most spectacular ecological transitions occur during adaptive radiations, such as blood-drinking (Grant and Grant 2008) or plant carnivory(Givnish et al. 1984, 1997) .Adaptation to novel fitness peaks is a hallmark of adaptive radiation, yet it is still poorly understood how such seemingly discontinuous transitions occur. An early hypothesis for how these transitions occurred was through ‘macromutational’ leaps that resulted in ‘hopeful monsters’ (Goldschmidt 1933, 1940). Although these ideas were

eventually discounted by the insights of Modern Synthesis (Dobzhanksy 1937; Simpson 1953; Mayr 1960; Charlesworth et al. 1982), recent empirical and theoretical research on hybrid speciation has revived an appreciation for the sudden appearance of evolutionary novelty through transgressive phenotypes resulting from hybridization. For example, hybridization between two species of sunflowers led to new hybrid species with suites of transgressive traits that allowed them to colonize novel sand dune and salt marsh habitats (Rieseberg et al. 2003; Gross and Rieseberg 2005). However, the transgressive effects of hybridization are so myriad (Moran et al. 2020; Brice et al. 2021) that conditions for hybridization triggering the evolution of new species or new ecological niches are still poorly understood and contentious in diploid taxa (Nieto Feliner et al. 2017; Schumer et al. 2018).

Recent conceptual frameworks for understanding adaptation to novel fitness peaks suggest their underlying ecological transitions are more likely to occur in stages (Erwin 2019, 2021). The initial emergence of a novel trait is likely to require further adaptive refinement to become successfully incorporated into the ecology of an organism. One of the best examples of this described so far comes from the sequences of mutations tracked during the evolution of the novel niche of aerobic citrate metabolism in *E. coli*, in which evolution of citT promotor capture in one intermediate lineage actualized the phenotype of aerobic citrate metabolism while further mutations at dctA in subsequent lineages refined the efficiency of citrate utilization that allowed it to fully occupy this niche (Blount et al. 2012, 2020; Quandt et al. 2014; Turner et al. 2015). This important series of studies suggests that novel ecological transitions are highly contingent on a complex series of mutations that potentiate, actualize and refine the adaptations to colonize new fitness peaks (Jagdish et al. 2020).

This idea that adaptive alleles from one lineage can set the stage for further diversification of more lineages underlies some general hypothesis about rapid radiations as well, like the hybrid swarm and syngameon hypothesis – in which radiations are driven by the exchange of genetic variation either from distinct lineages outside the radiation or within the radiation itself. There is a wealth of evidence for extensive histories of hybridization in radiations (e.g.(The Heliconius Genome Consortium et al. 2012; Lamichhaney et al. 2015a; Meier et al. 2017; Richards and Martin 2017; Poelstra et al. 2018; Richards et al. 2018; Meier et al. 2019), that suggests radiations are fueled by diversity that was generated across many lineages. Well documented cases of exchanges of adaptive variation aiding further diversification within a radiation include the exchange of beak shape alleles across Darwin's finches (Lamichhaney et al. 2015b) and wing pattern alleles across *Heliconius* species (The Heliconius Genome Consortium et al. 2012) that allow multiple species to diverge into similar ecological niches. However, we are only just beginning to explore how the introduction of genetic diversity from such groups gives recipient lineages access to the distant fitness peaks that generate the wealth of ecological, morphological and species diversity that hallmark radiations and that such access would be impossible without it to fully understand the extents to which diversity can self-perpetuate in radiations.

An adaptive radiation of trophic specialist pupfishes on San Salvador Island in the Bahamas is an excellent system for understanding how major and rapid ecological shifts occur in nature. This radiation contains a wide-spread generalist pupfish species (*Cyprinodon variegatus*) that occurs in sympatry with two previously described trophic specialists that are endemic to many of the hypersaline lakes on the island: a molluscivore (*C. bronotheroides*) with a novel nasal protrusion which is an oral-sheller of gastropods (St. John et al. 2020) and a scale-eating

specialist (*C. desquamator*) with two-fold larger oral jaws (Martin and Wainwright 2013a). The evolutionary novelties in this system originated recently; the radiation is only about 6-20 kya years old, before which the hypersaline lakes on San Salvador Island were periodically dry during glacial maxima (Hagey and Mylroie 1995; Turner et al. 2008; Clark et al. 2009). Intriguingly, we recently discovered species of pupfish living sympatrically with the two specialists and generalist on San Salvador Island. This pupfish population first attracted our attention for their noticeably intermediate jaw morphology between *C. desquamator* and *C. variegatus* (Fig 1). Furthermore, in a previous whole genome resequencing study, individuals with this intermediate jaw morphology unexpectedly formed a monophyletic group across lake populations distinct from *C. desquamator* using the unsupervised machine-learning approach Saguaro (citation; (Richards and Martin 2017). Here we refer to this new ecomorph as the 'wide-mouth' for its smaller jaws but wider mouth than the scale-eating specialist *C. desquamator*. The complex empirical fitness landscape driving this radiation suggests that *C. desquamator* is isolated by a large fitness valley from *C. variegatus* and *C. brontotheroides* (Martin and Wainwright 2013b; Martin et al. 2019) and this intermediate 'wide-mouth' may provide clues as to how this valley was traversed.

Here we investigate the position of this 'wide-mouth' sits on the ecological spectrum from generalist to scale-eating specialist using a combination of morphological, behavioral, dietary, and genomic data. We estimated the demographic history of the 'wide-mouth' and explore the timing of selection on shared and unique genetic variation involved in adaptation to scale-eating to better understand the ecological transition from generalist to scale-eater. Additionally we explore the extent to which all trophic specialist pupfish on this island share adaptive genetic variation, despite occupying divergent ecological niches. Our results show that

diversification of one species in a radiation can promote further diversification by sharing adaptive genetic variation useful for specialization.

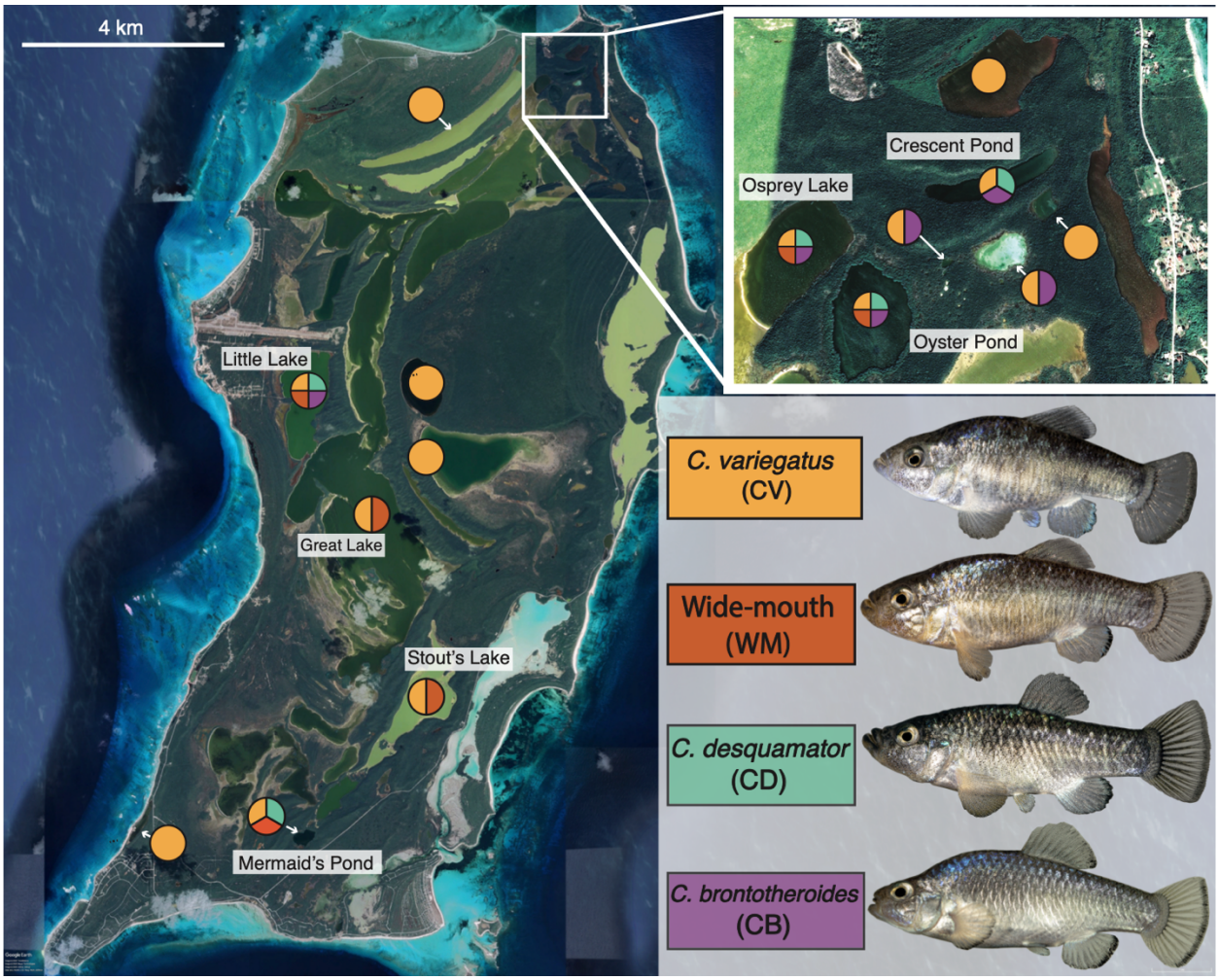

Fig. 1: Distribution of generalist, *C. brontotheriodes*, scale-eater specialist *C. desquamator*, and wide-mouth populations in hypersaline lakes on San Salvador Island. Pie charts indicate the presence of sympatric *Cyprinodon* species in each lake and are color-coded with representative pictures of generalist *C. variegatus* (gold) with typical trophic morphology of *Cyprinodon*, recently discovered 'wide-mouth' ecomorph (red-orange) with intermediate jaws, scale-eater *C. desquamator* (teal) with the largest oral jaws, and molluscivore *C. brontotheroides* (purple) with

a characteristic nasal protrusion. Labeled lakes contain all known ‘wide-mouth’ populations sampled for this study. Satellite image from Google Earth.

Results

Wide-mouth ecomorph is ecologically intermediate and morphologically distinct

We found the ‘wide-mouth’ ecomorph to be phenotypically distinct from *C. desquamator* and *C. variegatus* across a suite of traits (Fig. 2A-B). The lower jaw length of the ‘wide-mouth’ ecomorph was intermediate between *C. desquamator* and *C. variegatus* (Fig. 2C), while the buccal width and adductor mandibulae height were 8% larger in ‘wide-mouth’ than *C. desquamator* (Fig. 2E-F). These morphological differences were consistent across lakes that contained all three species (Osprey and Oyster) and within Great Lake which contains only wide-mouth and generalists (Fig S1). Trait differences in the ‘wide-mouth’ were heritable in a common garden laboratory environment after one generation. The ‘wide-mouth’ displayed 4% larger buccal width and 2% larger adductor mandibulae height in the lab environment than *C. desquamator* (Fig S2).

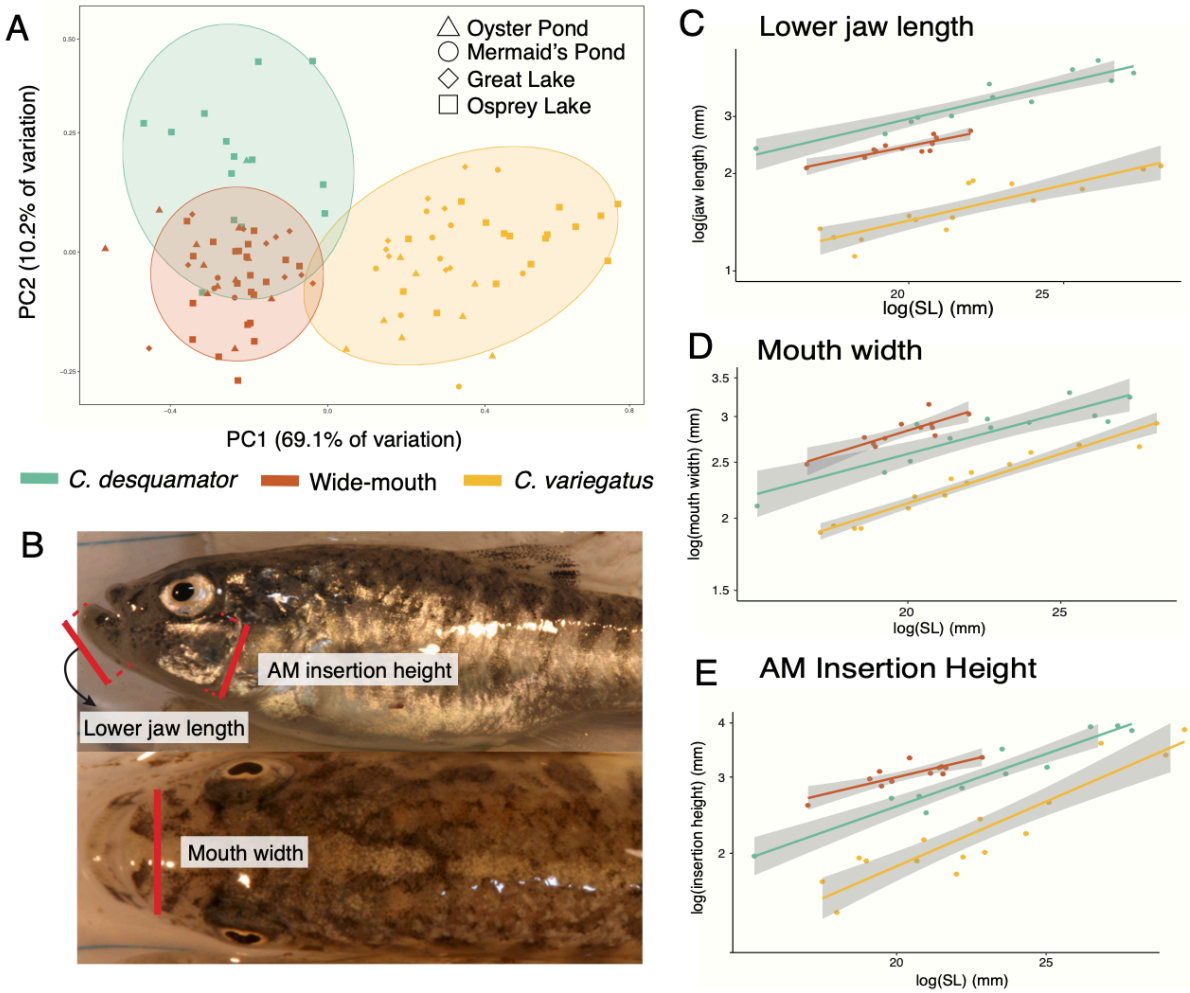

Fig 2. The wide-mouth ecomorph has distinct morphology within the San Salvador Island adaptive radiation across multiple traits. A) First two principal components of morphological diversity for 8 size-corrected traits and 95% confidence ellipses by species (*C. variegatus* : gold; wide-mouth: red-orange; *C. desquamator*: teal; *C. brontotheriodes* not shown). Shapes of points represent different lakes. PC1 is mainly described by lower jaw length and PC2 by adductor mandibulae insertion height, buccal and neurocranium width. B) External measurements of three morphological traits that best distinguish the wide-mouth ecomorph from both *C. desquamator* and generalists. C-E) The relationship between standard length (mm) of individuals and their C) lower jaw length, D) buccal cavity width, and E) adductor mandibulae muscle insertion height

(AM insertion) across individuals the three species in Osprey Lake. 95% confidence bands for linear models in gray.

Wide-mouth occupies a distinct intermediate scale-eating niche

We found that wide-mouth does appear to ingest scales, but at a significantly lower frequency than C. *desquamator* (Wilcoxon Rank Sum test, $P = 0.004$; Fig. 3A). We did not detect any scales in *C. variegatus* guts (Fig 3A). A previous gut content analyses of more individuals also found nearly zero occurrence of scales in the *C. variegatus* diet (Martin and Wainwright 2013c). All three sympatric species from Osprey Lake collected on the same day from the same site differed in δ15N levels (ANOVA, $P > 1.29\text{x}10^{-7}$; Fig. 3B) The 'wide-mouth' δ15N was intermediate between *C. variegatus* and *C. desquamator* (Tukey HSD; P=0.04 & 0.03 respectively). Similarly, high-speed scale-eating strikes on euthanized fish prey indicate an intermediate lower jaw angle in the 'wide-mouth' relative to generalist and scale-eating strikes (Fig. 3C; also see St. John et al. 2019 for larger analysis of scale-eating strikes relative to generalist).

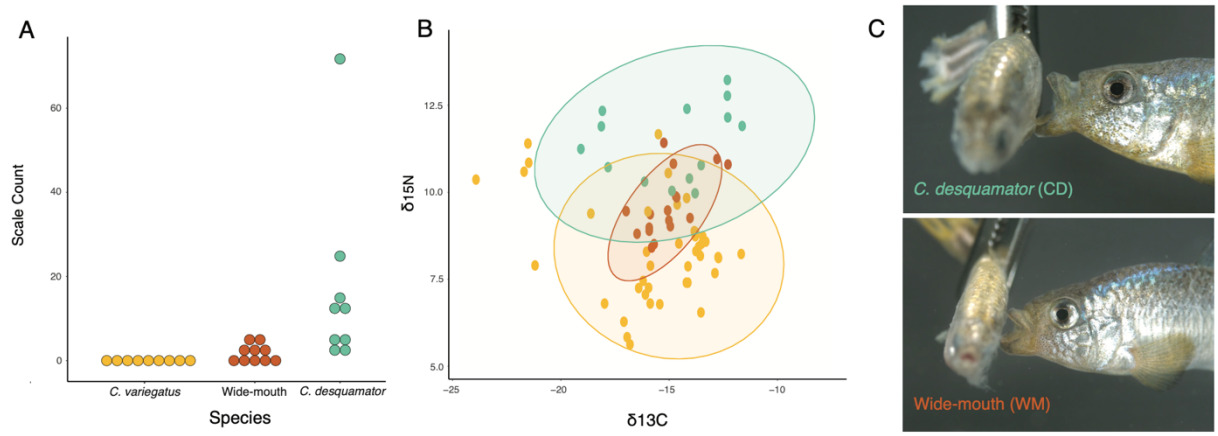

Fig. 3: The 'wide-mouth' individuals ingests scales. A) Dot plot of individual scale counts from gut content analysis of the hindgut of Osprey pupfish populations (10 individuals per species). B) Individual variation in relative trophic position (δ15N stable isotope ratio) and dietary carbon source (δ13C stable isotope ratio) with 95% confidence ellipses for the three species. C) Still images from Phantom VEO high-speed video footage of scale-eating strikes in *C. desquamator* and wide-mouth individuals. Note the distinct lower jaw angles with the suspensorium at the point of impact with the prey (also see (St. John et al. 2020)).

Wide-mouth did not result from hybridization between variegatus and desquamator

We tested whether 'wide-mouth' was the product of recent hybridization. First, extensive multi-year sampling efforts show that wide-mouth populations are not found in every lake where both *C. variegatus* and *C. desquamator* are present (e.g. Crescent Pond; Fig. 1). However, we explored this possibility in more depth by resequencing 24 genomes of 'wide-mouths' across lake populations and comparing these to 86 resequenced genomes of generalists, *C. brontotheriodes*, and *C. desquamator* from a previous study (Richards et al. 2021).

Several lines of genomic evidence support the 'wide-mouth' ecomorph as a distinct species. First, wide-mouth individuals do not occupy an intermediate position between *C. desquamator* and *C. variegatus* along either of the first two principal components that represent the major axes of genetic variation in the radiation. Instead, 'wide-mouths' occupy a distinct cluster relative to the other three species (Fig. 4A). This is further corroborated by the model-based approach of ADMIXTURE, in which wide-mouth individuals share more ancestry with each other than with either *C. desquamator* or *C. variegatus* under the two most likely K values based on cross-validation error estimates (K=6 and 7; Fig. 4B). This pattern holds when we

decrease or increase the K value used in the analysis to model more or less population structure within groups (Fig. S3). Lastly, we used formal tests for introgression and admixed populations, f_3 and f_4-statistics, to assess the evidence that 'wide-mouths' are the byproduct of recent admixture. The significant positive f_3-statistics and the non-significant f_4-statistics support our inference that the wide-mouth is not a recent admixed population of any pairwise combinations of the *C. brontotheriodes*, *C. variegatus*, or *C. desquamator* populations (Table S1).

Wide-mouth and C. desquamator populations are sister species which diverged shortly after the onset of adaptive radiation on San Salvador Island

To better understand the history of divergence and gene flow among the four species on San Salvador Island, we used *fastsimcoal2* (Excoffier et al. 2013) to run coalescent simulations and compare 28 different demographic models of divergence and gene flow on San Salvador Island (Table S2). These models represented all possible topologies among the four species and explored zero, early, and current gene flow scenarios. We evaluated their fit against our empirical site frequency spectrum (SFS) calculated from populations of the four species that occur in sympatry in Osprey Pond.

In the best supported model, wide-mouth was sister to desquamator with current gene flow and a divergence time estimate of 11,658 years ago (95 CI: 8,257-20,113 years; Fig 4C-D) (Table 1 and S2). This model indicates that the ancestor to wide-mouth and *C. desquamator* populations first diverged from *C. variegatus* 15,390 years ago (95% CI: 10,722-23,927 years;Fig 4D). This estimate of the origin of the radiation overlaps well with geological age estimates based on the filling of hypersaline lakes on San Salvador Island after the end of the last glacial maximum period (~6-19K years ago; (Hagey and Mylroie 1995; Turner et al. 2008; Clark

et al. 2009)). *C. variegatus* and *C. brontotheriodes* populations in Osprey Lake diverged much more recently than the two scale-eating populations 462 years ago (95% CI: 411-1,121 years; Fig 4D).

Table 1. Support for the five best-fitting demographic models for the evolution of the 'wide-mouths' from the site frequency spectrum. The likelihood and AIC scores for the top five best-fitting models estimated in *fastsimcoal2* for the Osprey populations of *C. variegatus* (CV), *C. brontotheroides* (CB), *C. desquamator* (CD), and 'wide-mouths' (WM) are presented here with a complete list of all models tested reported in Table S2. Change in likelihood (ΔLnL) represents the difference in likelihood from that of a simulated SFS expected by the demographic model tested. Change in AIC (ΔAIC) represents the difference in AIC scores from that of the model with the smallest ΔLnL. All models presented here represent different divergence scenarios with recent gene flow allowed, which had better support from the data than models with no gene flow or early gene flow. Visual representations of the top five models are depicted in Figure S5.

Divergence Model	LnL	ΔLnL	AIC	ΔAIC
(1) ((CD,WM),(CB,CV)) + (CD,WM) diverges first + recent gene flow	-11701	1077	54133	0
(2) ((CD,CB),CV),WM)) + recent gene flow	-11780	1112	54296	35
(3) ((CD,CB),WM),CV)) + recent gene flow	-11806	1138	54418	61
(4) ((CB,WM),CD),CV)) + recent gene flow	-11808	1140	54423	63
(5) ((CD,WM),(CB,CV)) + (CB,CV) diverges first + recent gene flow	-11846	1178	54602	101

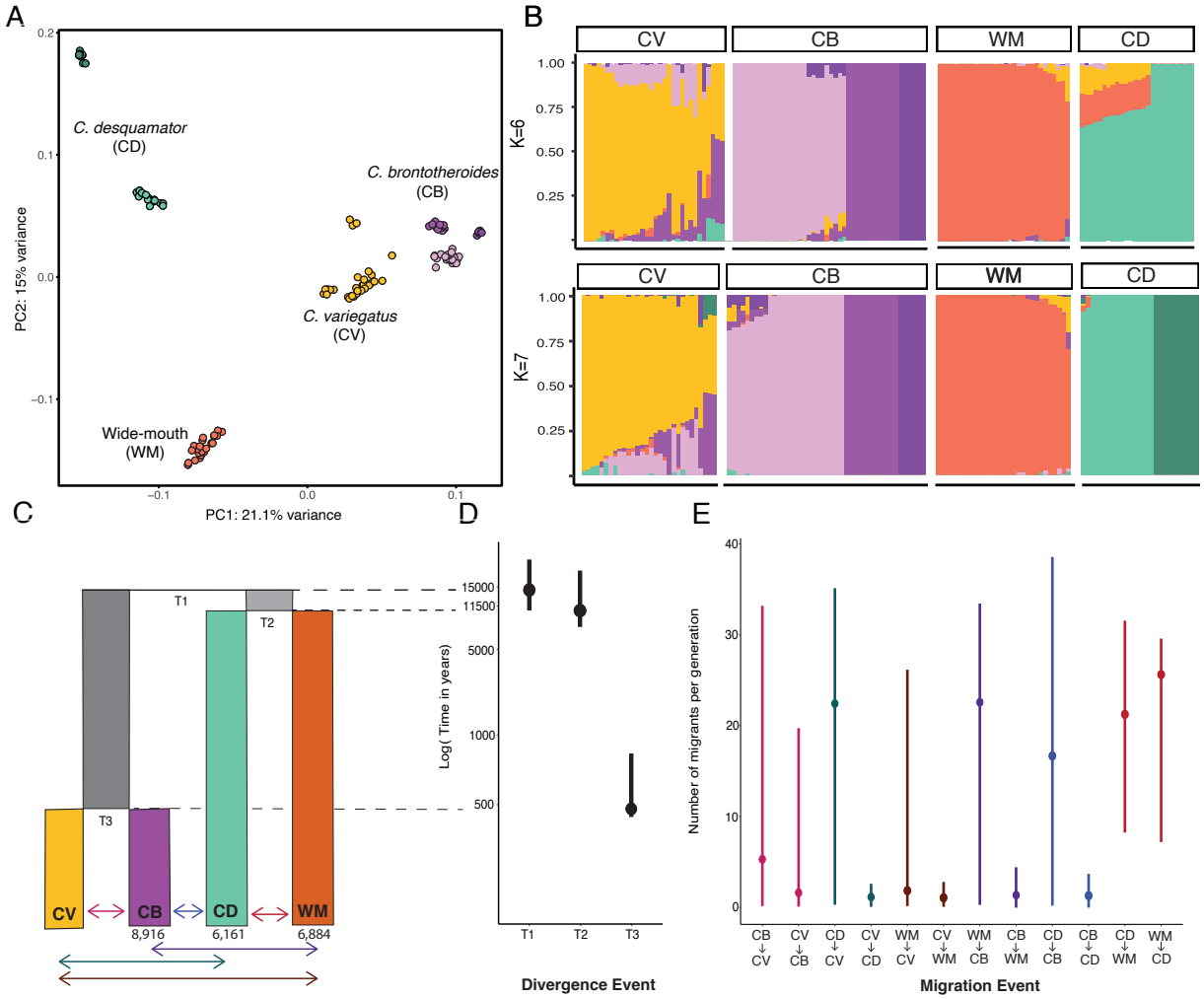

Fig 4. 'Wide-mouth' population did not result from recent hybridization. A) Principal components analysis of the four focal groups on San Salvador Island based on an LD-pruned subset of genetic variants (78,840 SNPs). B) Ancestry proportions across individuals of the four focal groups. Proportions were inferred from ADMIXTURE analyses with 2 values of K with the highest likelihood on the same LD-pruned dataset in A. C) Best supported demographic model in which *C. desquamator* diverged from the wide-mouth based on set of LD-pruned and no missing information across individuals set of variants (67,400 SNPs) of all Osprey Lake individuals. Divergence time is shown with the maximum likelihood point estimate from the run with the best fit and 95% confidence interval for that parameter estimate based on 100 bootstrap

replicates. D) Maximum likelihood point estimate and 95% confidence intervals for migration rate parameters involved in the best fitting model depicted in units of the number of migrants per generation.

Shared signatures of selection across the three specialists in the radiation despite trophic divergence

Next we investigated whether the evolution of new trophic specialists on San Salvador Island may have been aided by reuse of the same underlying genetic variation. We characterized regions of the genome under selection in the Osprey Lake populations of the four species using the sliding window SFS-based approach SweeD and neutral simulation thresholds based on demographic changes in effective population size over time (Fig S4; Table S3). A higher percentage of the genome was under positive selection in all specialist populations compared to *C. variegatus* population (Fig 5A). This pattern might be expected given the divergent selection pressures the specialists face to adapt to different trophic niches. Of all the specialists, the *C. brontotheriodes* exhibited both the most selective sweeps and the longest selective sweeps (Fig 5A). This suggests that the *C. brontotheriodes* have undergone selection most recently of all the species in Osprey Lake and is supported by the recent divergence time between the *C. brontotheriodes* and *C. variegatus* populations (Fig 4D).

Additionally, we assessed patterns of genetic divergence across the four species in Osprey Lake. Allowing for some level of gene flow between species, we calculated the number of fixed and nearly-fixed SNPs (Fst $\geq$ 0.95) between all pairwise combinations of populations in

Osprey Lake. All specialists, including the two scale-eating species, were more genetically diverged from each other than to *C. variegatus* species based on the number of SNPs fixed or nearly-fixed between them (Fig 5B). This pattern of stronger genetic divergence between specialists also held across fixed SNPs, the top 1% of F_{st}, and genome-wide average F_{st} (Table S4). Next, we characterized a set of candidate adaptive alleles for each specialist by focusing on those SNPs that were fixed or nearly fixed (Fst ≥ 0.95) against *C. variegatus* and found within a hard selective sweep region in the focal specialist. From this perspective including fixed or nearly fixed variation against the generalist species, the *C. brontotheriodes* had the fewest candidate adaptive alleles while *C. desquamator* had the most (Fig 5D).

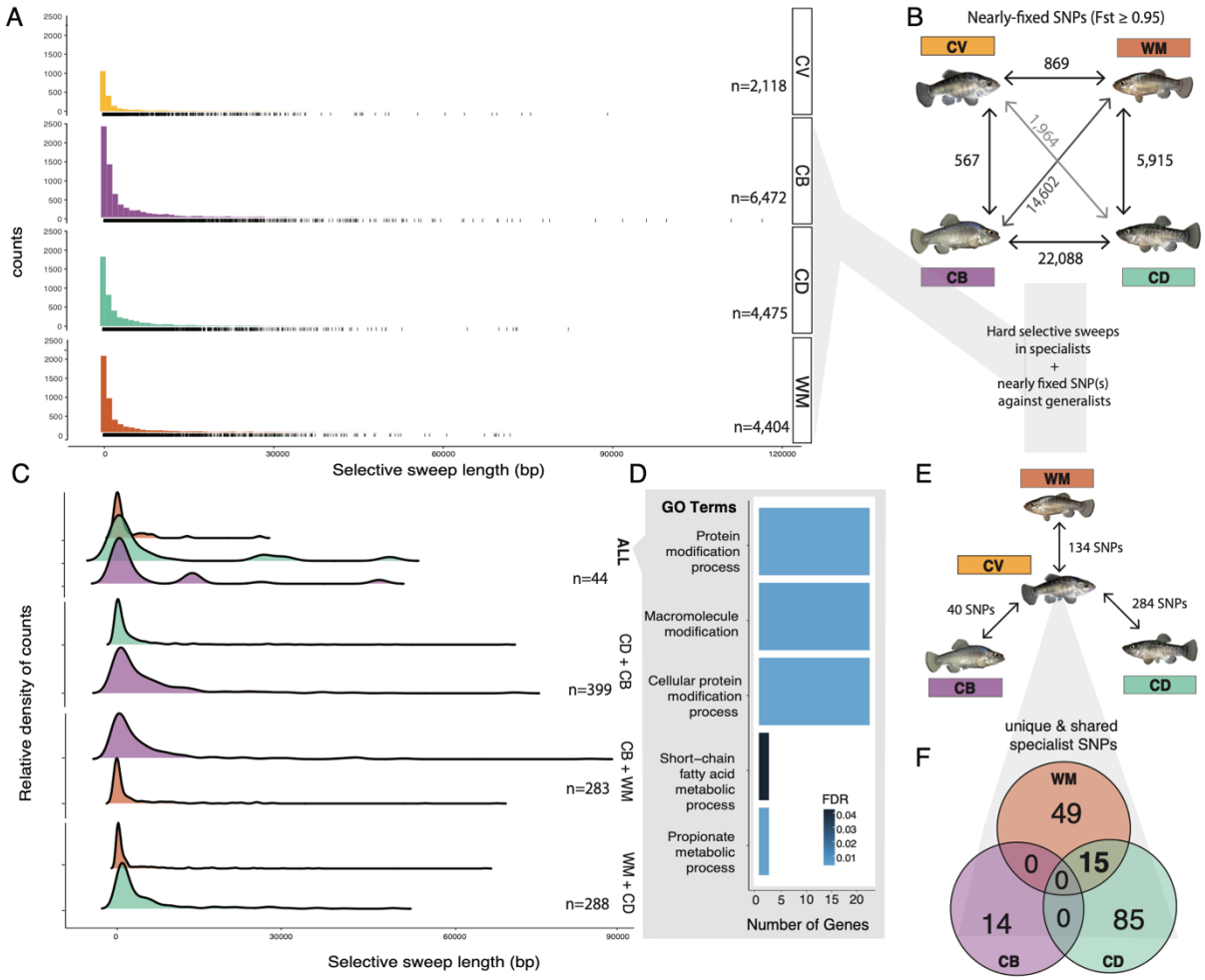

Figure 5: Patterns of selection and genetic divergence in specialist genomes. A) Histogram of selective sweep length distribution across all four San Salvador Island species. Rug plot below each histogram representing the counts of selective sweeps of particular lengths is included to highlight the presence of several larger selective sweeps in the tails of the distributions. B) The total number of nearly-fixed SNPS ($F_{st} \geq 0.95$) between each group in Osprey Pond. C) Ridgeline plots for length distributions of selective sweeps shared between different combinations of specialists (*C. brontotheriodes*: purple, *C. desquamator*: teal, 'wide-mouths': red-orange). Full results of the GO enrichment analysis of the 57 genes contained in the shared selective sweep regions across all specialists. D) The number of adaptive alleles (nearly-fixed SNPS [$F_{st} \geq 0.95$] relative to *C. variegatus* and under selection [CLR thresholds in Table S3]) in each population of specialists in Osprey Lake. Venn diagram highlights those adaptive alleles that are unique to each specialist and shared with other specialists.

Next, we looked for shared patterns of genetic divergence and selection across the specialist genomes. Despite divergent specializations, we found evidence of 44 shared selective sweeps across all three specialist populations that were not shared with *C. variegatus* population (Fig 5C). These shared regions under selection in all specialists were significantly enriched for genes annotated for metabolic processes (Fig 5C). Shared selection in regions related to metabolic processes may stem from the shared factor of dietary specialization (also see (Mcgirr and Martin 2018)). We also found evidence of shared selection across all pairwise combinations of two specialists. The shared selective sweeps were shorter than any of the selective sweeps unique to each of the specialists, suggesting that the selection for these shared regions was not

the most recent or strongest in any of the specialists (Fig 5C). In contrast, we found no shared fixed or nearly fixed adaptive alleles across all three specialists. Although, there are shared adaptive alleles between them at lower frequencies (i.e. Fst < 0.62; Fig S6).

Wide-mouth does not contain unique adaptive alleles with known associations to craniofacial morphology

We did find that the two scale-eating populations shared a set of the same adaptive alleles that *C. brontotheriodes* did not have (Fig 5D). This consisted of 15 alleles in 6 shared hard selective sweeps in *C. desquamator* and 'wide-mouths': 10 of these variants were in unannotated regions, two were in the introns of the gene *daam2,* and three were in putative regulatory variants of the genes *usp50, atp8a1 and znf214* (one variant each).

Shared adaptive alleles in the gene *daam2,* a Wnt signaling regulator, are intriguing because knockdowns of the gene have been shown to cause abnormal snout morphology, osteoporosis, and changes in insulin and alkaline phosphate levels in mice (Dickinson et al. 2016), and abnormal cranial and skeletal development in zebrafish (Kida et al. 2007). Craniofacial morphology is one of the rapidly diversifying traits in this system, so these shared adaptive alleles in *daam2* are intriguing candidates shared craniofacial divergence in the two scale-eating species. Based on currently available data, none of the other adaptive alleles appear to be good candidate alleles for craniofacial morphology: *usp50* functions in protein metabolism and deubiquination (Lee et al. 2017), *atp8a1* an ion transporter (Soupene 2008), *znf214* is a transcription binding factor, and the 10 unannotated variants were not associated with lower jaw

size variation in a previous genome-wide association study of the radiation (Richards et al. 2021).

Despite the divergent craniofacial features of the 'wide-mouths', none of the adaptive alleles unique to the wide-mouth appear to be in or near genes annotated for craniofacial phenotypes in model organisms. In *C. desquamator,* three of the 13 sets of unique adaptive alleles are in or near genes annotated for craniofacial phenotypes: a *de novo* non-synonymous coding substitution in the gene *twist1*, several putative regulatory variants near the gene *gnaq,* and 8 variants in and near the gene *bri3bp*, which is located inside a QTL region for cranial height in pupfish (St. John et al. 2021). In *C. brontotheriodes*, there is also a candidate craniofacial adaptive allele: a non-synonymous coding substitution in the gene *kat6b*, which is annotated for abnormal craniofacial morphologies including shorter mandibles in mice (Thomas et al. 2000; Kraft et al. 2011) and Ohdo Syndrome and bulbous noses in humans (Clayton-Smith et al. 2011).

This pattern of unique alleles relevant to craniofacial phenotypes in the *C. brontotheriodes* and *C. desquamator*, but not the 'wide-mouths', holds if we lower the threshold to the top 1 percentile of F_{st} between specialists and generalist. 'Roof of mouth development', 'bone development' and 'skeletal system development' are in the top 20 most significantly enriched GO terms for the *C. desquamator* candidate adaptive alleles, with 6 significantly enriched terms relevant to cranial and skeletal development (out of 165 enriched terms total). Similarly, 'embryonic skeletal system development' and 'skeletal system development' were significantly enriched terms in the *C. brontotheriodes* candidate adaptive SNP set (out of 8 enriched terms total). However, the wide-mouth adaptive allele set at this lower F_{st} threshold is not significantly enriched for any GO terms related to craniofacial or skeletal morphology (n=52

terms; Fig S7). This suggests that 'wide-mouths' may not any unique adaptive alleles related to craniofacial morphology.

Selection for adaptive alleles shared in both scale-eating populations preceded selection for adaptive alleles unique to C. desquamator

Additionally, we estimated the timing of selection on unique and shared adaptive alleles in 'wide-mouths' and *C. desquamator* to explore the evolution of scale-eating novelty in this system. Previous studies indicate that *C. desquamator* scale-eater niche is separated from *C. variegatus* generalist niche by a deep fitness valley on the adaptive landscape, suggesting a challenging ecological transition. Adaptation to an intermediate scale-eating niche such as the 'wide-mouths' may have provided an easier first transition and the adaptive alleles involved may have set the stage for the more specialized scale-eater *C. desquamator* to evolve. Thus, we predicted that we would see A) selection on the shared adaptive alleles occurred before those unique to each species and B) selection for the shared adaptive alleles occurred earlier in the intermediate 'wide-mouths' than the more specialized scale-eater, *C. desquamator*. In order to test this, we estimated the age at which the adaptive alleles first started sweeping in the population using the coalescent-based approach starTMRCA (Smith et al. 2018).

Across the 30 sets of adaptive alleles with selective sweep ages, we found an overall difference in timing of selection among shared and unique adaptive alleles in the two scale-eater populations (ANOVA *P*-value = 0.00478). In the *C. desquamator* population, shared adaptive alleles were under selection before any unique adaptive alleles (Tukey HSD *P*-value = 0.003217; Fig. 6). We found a similar pattern in the 'wide-mouth', where shared adaptive alleles were generally under selection before those unique to the population (Fig. 6). However, this

difference was not significant due to one set of unique adaptive alleles near the gene *slitrk5* that was the oldest sweep of all those estimated (ANOVA, Tukey HSD; *P*-value=0.8367).

In contrast to our predictions, selection on 4 of the 6 shared sets of adaptive alleles occurred significantly earlier in *C. desquamator* than the 'wide-mouths'. Only a single set of these adaptive alleles had an older median age estimate in the 'wide-mouths' than *C. desquamator*, although the 95% HPD intervals surrounding these point estimates overlapped between the two population (Fig. 6). The other adaptive allele set that had overlapping 95% HPD intervals between the two populations were a set of unannotated alleles de novo to the scale-eating populations on San Salvador Island.

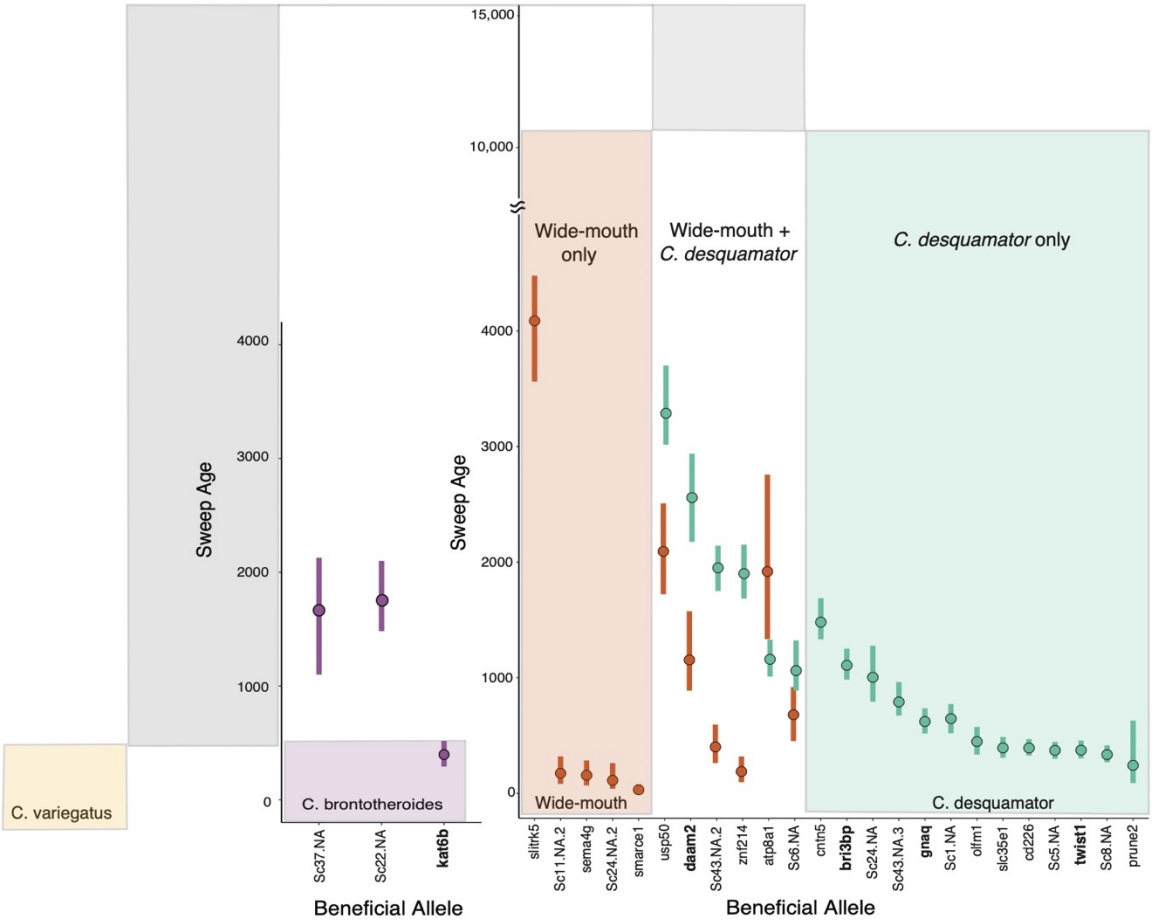

Fig 6. Timing of selection on adaptive alleles in trophic specialists. The mean and 95% HPD estimates for the timing of selection on sets of fixed or nearly fixed SNPs (named by the gene they are in or within 20-kb of) *C. desquamator*, 'wide-mouths', and *C. brontotheriodes* populations of Osprey Lake. The age of each beneficial allele is color coded by the species it swept in and the inferred demographic history is displayed in the background for comparison. Gene names highlighted in bold are associated with oral jaw size in model organisms of zebrafish and/or mice.

Shared adaptive variation may have resulted from introgression between C. desquamator and 'wide-mouth' scale-eaters

Finally, we explored several selection scenarios surrounding the shared adaptive alleles between 'wide-mouth' and *C. desquamator*. The most parsimonious answer for why these two populations share adaptive alleles is that these alleles swept in their common ancestor and were retained in both post divergence. However, the timing of selection on these alleles is much more recent in both populations than the inferred divergence time (Fig 6), suggesting that these alleles were not selected upon in their most recent common ancestor. The timing of selection on most of these shared adaptive alleles is significantly different between the populations (Fig 6), also suggesting that these alleles were not selected upon in a single event in the shared ancestor. This left us to explore two other scenarios that could explain the shared adaptive alleles: 1) these alleles were segregating in both populations after they initially diverged from each other and independently swept to fixation or 2) these adaptive alleles introgressed from one population to another after being selected for in one population.

To distinguish between these two scenarios, we assessed the regions surrounding these shared adaptive alleles for signatures of introgression. Introgression between recently diverged sister species is often more challenging to detect in comparison to introgression with a more distantly related outgroup because of the high genetic similarity between sister species. However, introgression between sister species can be detected from population level polymorphism data by looking for regions of the genome with lower absolute genetic divergence than expected given the genome-wide average (Joly et al. 2009; Geneva et al. 2015; Rosenzweig et al. 2016; Hibbins and Hahn 2021). We thus calculated D_{xy} in sliding windows across the genomes of the Osprey Lake populations and compared the distribution of D_{xy} values in the

regions surrounding the shared adaptive alleles to the distribution of D_{xy} values across the genome. The average D_{xy} of the shared adaptive alleles was 2X lower than average genome-wide values between *C. desquamator* and 'wide-mouths' (Fig 7A). It was also slightly lower (1.2X) than the average D_{xy} for all selective sweep regions in the wide-mouth genomes, suggesting it is not due to these regions under selection in general (Table S5). This lower D_{xy} was predominantly due to four out of the six regions containing shared adaptive alleles having ~2-5 times lower D_{xy} values than the genome-wide average. The other two regions had comparable (*znf214*) and higher (unannotated region on scaffold 43) D_{xy} values than the genome-wide average (Fig 7B). The lower genetic distance between *C. desquamator* and the 'wide-mouths' in these adaptive regions suggests they may have introgressed between the two groups rather than independent selection on ancestral polymorphisms since they diverged approximately 11,000 years ago.

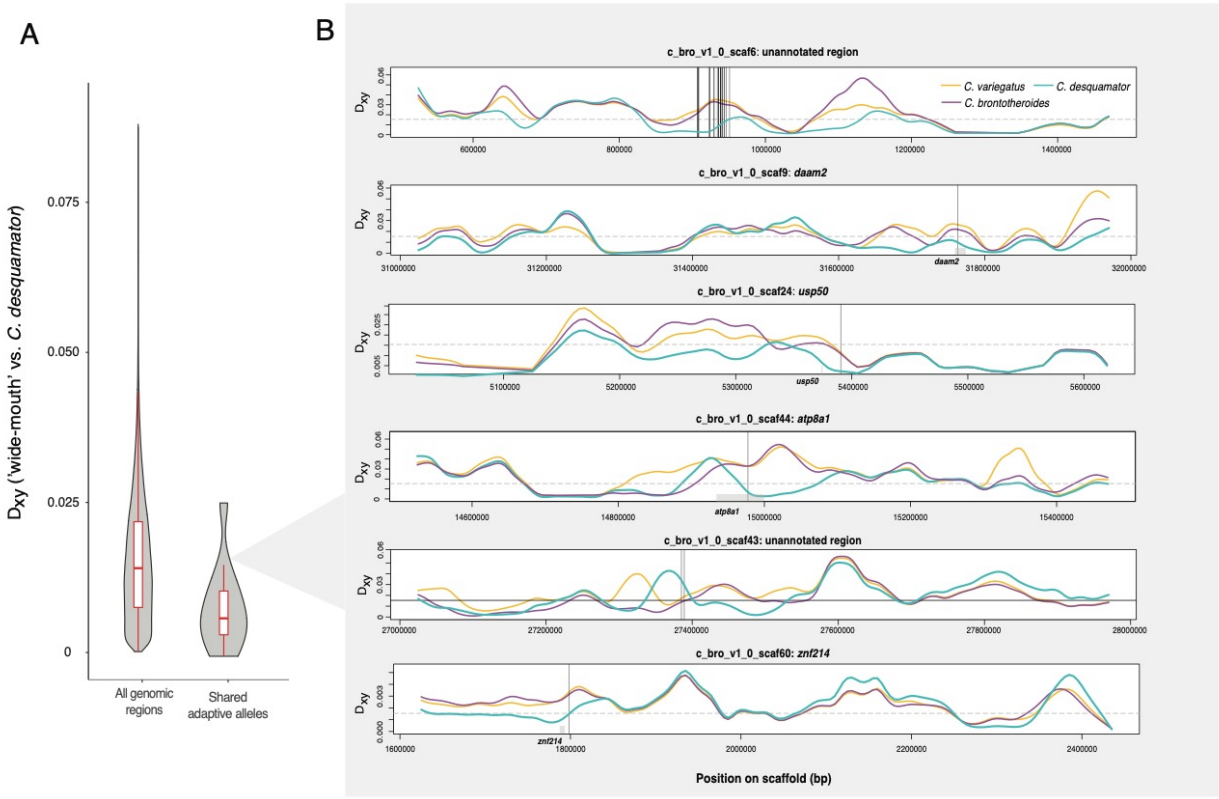

Fig 7: Absolute genetic divergence between the 'wide-mouth' and *C. desquamator*. A) Distribution of D_{xy} values across 20-kb regions of the genome for the whole genome, the selective sweep regions, regions with adaptive alleles shared between *C. desquamator* and 'wide-mouth' populations, and unique to the wide-mouth. B) D_{xy} values across the 1 million base pairs surrounding the six regions of shared genetic variation between *C. desquamator* and 'wide-mouth'. Vertical solid black lines indicate the position of the shared adaptive alleles between 'wide-mouth' and *C. desquamator*, horizontal dashed gray lines represent the genome-wide average D_{xy} between *C. desquamator* and 'wide-mouths'. The smoothed lines represent D_{xy} between *C. desquamator* and 'wide-mouths' (teal), *C. brontotheriodes* (purple) and *C. variegatus* (gold) across these portions of the scaffold. The D_{xy} values were summarized with smoothed splines using the smooth.spline() function in R for ease of visualization across the different comparisons. Gray boxes represent the nearest gene to the adaptive alleles and

unannotated adaptive alleles were those in which no gene was located within 20-kb of the SNP (an approximate LD estimate from a previous study;(Richards et al. 2021)).

Discussion

Discovery of a new cryptic scale-eater evolved through reuse of shared adaptive alleles

The traditional view and hallmark of adaptive radiations is that they occur in rapid bursts of diversification that slows down over time as resources for further diversification diminishes. An alternative possibility is that radiations can be self-propagating and that the diversity generated within the first stages of radiation helps beget further diversity (Whittaker 1977). This could happen in a variety of ways: through niche constructionism, such as exploitation of new trophic levels created by new species or physical alterations of environment by new species, that may create additional opportunities for further diversification and prolongment of radiation (reviewed in (Stroud and Losos 2016; Martin and Richards 2019). The diversity begets diversity hypothesis can be visualized as the evolution of the fitness landscape itself as species in the radiation colonize new peaks, providing access to new fitness peaks to fuel rapid radiation.

Another way we could think about the evolution of one species promoting the evolution of another is at the microevolutionary level through shared adaptive variants that can help other populations colonize new or ecologically similar areas of the fitness landscape. However, the genetic basis and microevolutionary processes underlying major ecological transitions that require colonizing new fitness peaks are still poorly understood in nature. Our discovery of cryptic new scale-eating species through morphological, dietary, and genomic analyses revealed shared selective sweeps among specialists in a radiation of pupfish despite the divergent

morphological and ecological adaptations among them. However, only in the two scale-eating populations, *C. desquamator* and 'wide-mouth', do we find shared nearly-fixed or fixed adaptive alleles. The 'wide-mouth' and *C. desquamator* are sister species, suggesting a single origin for scale-eating. Nonetheless, these shared adaptive alleles were not selected upon in the ancestral population and swept first in the *C. desquamator* population. This is in contrast to our initial prediction that the evolution of this intermediate scale-eater aided the evolution of the more specialized scale-eater *C. desquamator*. Instead, adaptive alleles from *C. desquamator* may have introgressed and then selected upon in the 'wide-mouth' population. Additionally, unlike the *C. brontotheriodes* and *C. desquamator*, the cryptic wide-mouth has little unique genetic variation relevant to craniofacial morphology, which is a major axis of diversification in the other trophic specialists (Martin and Wainwright 2011). An intriguing implication of these findings is that adaptive variants within one specialist may aid the evolution of another into a similar niche and promote further ecological divergence within radiations, in line with the diversity begets diversity hypothesis (Whittaker 1977).

An adaptive walk underlies the major ecological transition from generalist to scale-eating specialist

One of the foundational models of adaptation is that it proceeds in 'adaptive walks' towards fitness optima that involve the sequential fixation of adaptive alleles that move a population in the phenotypic direction of that optimal peak (Maynard Smith 1970; Gillespie 1984; Kauffman and Levin 1987; Orr 1998, 2005). The distinct timing of selection across different adaptive alleles suggests that the ecological transition from generalist to novel scale-eating specialist involved such an adaptive walk rather than a sudden burst of concurrent selection events after

some major environmental shift. These distinct, multiple bouts of selection could be caused by mutation-limited (i.e. waiting for new beneficial alleles; (Bell 2013; Lindsey et al. 2013; Rousselle et al. 2020)) or mutation-order processes (i.e. epistatic interactions; (Mani and Clarke 1990; Schluter 2009; Good et al. 2017).

All of the shared adaptive alleles between the scale-eating populations occur at low frequency in generalist populations on other Caribbean islands based on a Caribbean-wide genomic dataset from a previous study ((Richards et al. 2021); Table S7). For example, three copies of the shared *usp50* adaptive allele were found outside of San Salvador Island (Table S7). This indicates that initial bouts of selection occurred on available standing genetic variation and that staggered timing of hard selective sweeps in each trophic specialist most likely reflect mutation-order processes in which selection on a beneficial allele was contingent on prior fixation of other adaptive alleles in each specialists' genetic background. However, several adaptive alleles originated from introgression or de novo mutations found only on San Salvador Island (Table S7), so part of the adaptive walk may have also been mutation-limited.

Cryptic scale-eating species reveals features of the adaptive walk towards scale-eating specialization

Although the two scale-eating species shared a set of adaptive alleles and a recent common ancestor, 5 out 6 shared adaptive regions swept at significantly different times between the two species. This difference in timing may result from several different scenarios: 1) independent adaptive walks to the same scale-eating niche, 2) independent adaptive walks to different scale-

eating niches or 3) the adaptive walk of one population depended on the adaptive walk of the other population. We explore each of these scenarios in turn.

First, the difference in timing of selection on the same shared adaptive alleles could indicate independent adaptive walks to the same scale-eating peak that occurred at different times and/or by slightly different routes. *C. desquamator* and 'wide-mouth' populations have predominantly abutting ranges, with only a small amount of geographic overlap on the island in four lakes, Osprey, Oyster, Little Lake, and Mermaid's Pond (Fig 1). If this current distribution is representative of their historical ranges and the two lineages began diverging about 11,000 years ago, it is possible that the adaptive walks took place in different lakes and largely independent of one another. The differences in timing on the shared adaptive alleles and the presence of unique alleles observed in each population are reminiscent of mutation-order speciation(Schluter 2009; Schluter and Conte 2009). In mutation-order speciation the same alleles are favored in populations that are adapting to similar environments, yet by chance and/or epistatic interactions with different genetic backgrounds, similar adaptive alleles fix in just one population. At least one set of alleles near the gene *slitrk5* are unique to the wide-mouth scale-eater and fixed well before selection occurred on the shared adaptive variants between the two scale-eater populations (Fig 5). Epistatic interactions of these *slitrk5* alleles may have prevented the 'wide-mouths' from fixing additional segregating alleles that are uniquely fixed in *C. desquamator.*

Second, the difference in timing of selection may have occurred because the two scale-eating populations are adapting to two different scale-eating niches. The intermediate diet, distinctly sized morphological traits, and smaller body size in the wide-mouth scale-eater compared to *C. desquamator* may indicate the wide-mouth scale-eater is adapting to a different

scale-eating niche. While the two scale-eating populations do share some overlap in adaptive alleles and selective sweeps, the majority of selective sweeps are unique to each species (Fig 5 & 6). Unique neurogenesis, brain and nervous system development (*slitrk5,sema4g, and smarce1*), whereas unique adaptive alleles in desquamator population include craniofacial development annotations (*olfm1, gnaq, twist1*) as well. This difference in gene annotations can also be seen at the broader level of regions of the genome under selection (Fig S7). Therefore, the differences in timing at shared adaptive alleles might reflect differences in relative strength of selection due to different selective regimes experienced by the two scale-eating populations. While starTMRCA is fairly robust in its timing estimates to different strengths of the selection (Smith et al. 2018), we cannot rule out that differences in relative timing of selection might partially reflect differences in strength of selection occurring in the two scale-eating populations.

Finally, another scenario that could have led to the older timing of selection across shared adaptive alleles in desquamator than the wide mouth is one in which the adaptive walk of one population depended on adaptive alleles from another. The significantly older timing of selection in *C. desquamator* on shared adaptive alleles and the low absolute genetic divergence between these two scale-eating population in these regions (Fig 6) suggests that these alleles may have introgressed between the two populations. Given that these shared adaptive alleles appear as standing genetic variation across Caribbean populations, albeit at very low frequency in our sampling, it is likely that the shared alleles were present in the ancestor and were segregating in both populations after they diverged. There are several reasons why adaptive introgression may have been necessary for adaptive divergence of the 'wide-mouth' despite the allele initially segregating in the ancestral population. In line with the mutation-order processes (e.g. Mani and Clarke 1990; Schluter 2009; Good et al. 2017), these alleles might not have been adaptive until

the right genomic background was present in the 'wide-mouth'. Introgression from *C. desquamator* where the allele was already swept to high frequency could have raised the frequency of these alleles in 'wide-mouth' and increased the likelihood of fixation for the beneficial alleles (reviewed in (Patwa and Wahl 2008)). This is consistent with previously proposed hypotheses for how diversity begets further diversity in adaptive radiations (Whittaker 1977; Stroud and Losos 2016; Martin and Richards 2019).

Shared aspects genomic landscape of selection among all specialists additionally supports specialization promoting further specialization

Intriguingly we also find evidence of selective sweeps shared across all three specialists despite their divergent adaptations and a lack of a shared specialist ancestor. However, there are no commonly shared adaptive alleles fixed against *C. variegatus* in these shared selective sweep regions. Hard selective sweeps indicate a selection scenario in which a single beneficial allele of large effect on a trait is swept to high frequency (reviewed in (Pavlidis and Alachiotis 2017)). The lack of highly divergent alleles in these regions might indicate polygenic selection events underlie the shared signatures across specialists. The broad scale transition from dietary generalist to dietary specialist may involve polygenic selection with small shifts in the frequency of many alleles. Despite having very different expectations about the genetic basis of adaptation, polygenic selection events may have been detected as a hard selective sweep in this study as the two selection types can be challenging to distinguish based solely on patterns of genetic variation (Chevin and Hospital 2008; Höllinger et al. 2019; Thornton 2019; Barghi et al. 2020). Recently developed frameworks provide additional criteria to help distinguish between the two selective regimes (Barghi et al. 2020) but are beyond the scope of this current study.

These shared selective sweeps across all specialists are not the longest selective sweeps in any of the specialist genomes (> 50-kb; Fig 5C), indicating they are not the most recent or strongest selection events. However, these shared selective sweeps appear relevant dietary specialization as these regions are enriched for genes annotated for metabolic processes such as short chain fatty acid and propionate metabolism (Fig 5). Propionates and other short chain fatty acids are common microbiome metabolites (Flint et al. 2012). A recent microbiome study in lab-reared populations found that *C. desquamator* microbiomes appear enriched for *Burkholderiaceae* bacteria that can digest collagen, further supporting an adaptive role for microbiomes in the radiation (Heras and Martin 2021).

We also see pairwise shared selective sweeps between all combinations of two specialists. In contrast to the sweeps shared between all specialists, we see stronger genetic divergence among species in these regions and a proportion of selective sweeps shared between only two specialists do contain nearly-fixed alleles. However, the specialists that share selective sweeps in these regions are fixed for different adaptive alleles, consistent with previously observed patterns of parallel differential gene expression in specialists despite divergent genotype (Mcgirr and Martin 2018). These alleles were possibly under balancing selection in the ancestor (most of these alleles are segregating at intermediate frequency in *C. variegatus* currently) and the alternate alleles were driven to fixation between the specialists during their respective divergences. While the species have discrete phenotypes across several morphological traits important to divergence in this system, these traits often lie on a continuous axis (i.e. smaller oral jaw lengths in the *C. brontotheriodes*, intermediate jaw lengths in *C. variegatus*, and larger oral jaw lengths in 'wide-mouths' and *C. desquamator*). Some of these alleles that are

alternatively fixed between specialists at the same position may have incomplete dominance effects on such traits.

In line with adaptation to divergent trophic niches across the species, we do still see many unique signatures of selection among the specialists. *C. variegatus* population had the fewest selective sweep signatures. All three specialists had more selective sweeps potentially resulting from strong directional selection for adapting to new fitness peaks. However, while we might expect stronger signatures of selection in *C. desquamator* genomes given the deeper valley on the fitness landscape separating the scale-eater fitness peak (Martin and Wainwright 2013b; Martin and Gould 2020), C. *brontotheriodes* had the most selective sweeps detected in their genome. They also have the longest sweeps detected (e.g. 100-kb), suggesting that selection may have occurred most recently in the *C. brontotheriodes*. This matches with the most recent divergence event being between generalist and *C. brontotheriodes* in Osprey Lake. An intriguing implication of this result is the potentially long wait time for the *C. brontotheriodes* species compared to the novel scale-eating species despite it being separated by a shallower fitness valley. Although we cannot yet rule out that the recent divergence time we estimated in Osprey Lake reflects a more recent arrival of the *C. brontotheriodes* to that lake and initially diverged longer ago in another lake not incorporated into our demographic model.

Methods

Sampling

The wide-mouth individuals were collected on San Salvador Island in the Bahamas using hand and seine nets between 2017 and 2018. Wide-mouth individuals were collected from 5 lakes in which taxa occurs in sympatry with just *C. variegatus* (Great Lake, Mermaid's Pond, and Stout's Lake), with generalist and *C. desquamator* (Oyster Pond and Osprey Lake; Fig. 1). In Oyster Pond, *C. desquamator* is extremely rare (only a single Oyster Pond individual grouped genetically and morphologically with *C. desquamator*), so Osprey Lake appears to be the only lake where all three species and the recently discovered wide-mouth coexist at appreciable frequencies in sympatry. To the best of our knowledge, wide-mouth populations currently exist only in these lakes, but have been collected in Little Lake in the past (Bruce Turner, pers. obs 1990s).

Fishes were euthanized in an overdose of buffered MS-222 (Finquel, Inc.) following approved protocols from the University of North Carolina at Chapel Hill Animal Care and Use Committee (#18-061.0), and the University of California, Berkeley Animal Care and Use Committee (AUP-2015-01-7053) and preserved in 95-100% ethanol.

Ecological and morphological characterization of 'wide-mouth' scale-eater

Multiple individuals from 3 lake populations (Osprey Lake, Great Lake, and Oyster Pond) in which we had measurable specimens of the species (n=84) were measured for 9 traits using dial calipers. Traits were selected for specific connections to foraging performance (Wainwright et al. 2004; Patricia Hernandez et al. 2009) and that differed across the three groups in previous study

when externally measured (Martin and Wainwright 2013b). For each specimen we measured (1) standard length from the anterior tip of the premaxilla to the posterior margin on the midline of the caudal peduncle, (2) lower jaw length from the medial point of the dentary on the lower jaw to dorsal point of rotation in the quadrate-articular joint, (3) the width of the buccal from the distance between the proximal sides of the buccal cavity in dorsal view, (4) the height of the adductor mandibulae insertion from the vertex between the vertical and horizontal arms of the preopercule to the dorsal margin of the hyomandibula, (5) the distance from the insertion of the first ray of the dorsal fin to the insertion of the first ray of the anal fin ,(6) the height of the caudal peduncle from the dorsal and ventral procurrent rays on the caudal fin, (7) neurocranium width from the narrowest distance between the eyes from the dorsal view, (8) body width was distance from just behind the later margins of the opercula from the medial position down the body, and (9) diameter of the eye through the major axes of the iris. Measurements were made by a single observer (EJR). Repeatability of measurement was assessed by remeasuring random subset of 10 individuals per species twice and determining the closeness of agreement between the measurements as a percentage (100% being that multiple independent measurements were the exact same value). Measurements on the full dataset were not taken until repeatability was above 85% per trait. To remove the effects of body size on trait size, we used the residuals from a linear regression of the log-transformed trait on log transformed standard length. We also compared these morphological measurements between lab reared F2 and wild caught specimen to rule out strong plasticity in traits of interest.

For a subset of individuals collected from each population (*C. variegatus, C. desquamator,* and 'wide-mouth') in Osprey Lake, diet was characterized from stomach content analyses (n=10 each) and stable isotope analyses (n=75;42 *C. variegatus*, 16 *C. desquamator*, 17

‘wide-mouth’ individuals). These are two complementary approaches that reflect short-term and long-term snapshots of the dietary differences among these groups. For stomach content analysis, a single observer (EJR) dissected out a 1 cm portion of the gut from the distal end (hindgut region) and counted the number of scales found. For stable isotope analyses, dried muscle tissue samples were taken from the caudal peduncle region of individuals in the field on the day they were collected and euthanized. These tissues were then sent to the UC Davis Stable isotope facility and analyzed with a PDZ Europa ANCA-GSL elemental analyzer interfaced to a PDZ Europa 20-20 isotope ratio mass spectrometer to characterize the natural abundances of δ13C and δ15N amongst individuals of different groups.

Genomic library preparation

We sequenced 24 wide-mouth individuals following protocols used in a previous study (Richards et al. 2020) in which we sequenced genomes from *Cyprinodon variegatus*, *C. desquamator*, and *C. bronototheroides*. Raw reads were mapped from these 24 individuals to a de-novo assembly of *Cyprinodon brontotheroides* reference genome (v 1.0; total sequence length = 1,162,855,435 bp; number of scaffold = 15,698, scaffold N50 = 32 Mbp) with bwa-mem (v 0.7.12;(Li and Durbin 2011)). Duplicate reads were identified using MarkDuplicates and BAM indices were created using BuildBamIndex in the Picard software package (http://picard.sourceforge.net(v.2.0.1)). We followed the best practices guide recommended in the Genome Analysis Toolkit (v 3.5;(DePristo et al. 2011)) to call and refine our single nucleotide polymorphism (SNP) variant dataset using the program HaplotypeCaller. Variants in wide-mouth individuals were called jointly with the 202 individuals sequenced from a previous study (Richards et al. 2020). We filtered SNPs based on the recommended hard filter criteria (i.e.

QD < 2.0; FS > 60; MQRankSum < -12.5; ReadPosRankSum < -8;(DePristo et al. 2011; Marsden et al. 2014)) because we lacked high-quality known variants for these non-model species. After selecting for only individuals from San Salvador Island, variants were additionally filtered to remove SNPs with a minor allele frequency below 0.05, genotype quality below 20, or containing more than 10% missing data across all individuals at the site using vcftools (v.0.1.15;(Danecek et al. 2011)). Variants in poorly mapped regions were then removed using a mask file generated from the program SNPable (http://bit.ly/snpable; k-mer length =50, and 'stringency'=0.5). Our final dataset after filtering contained 6.4 million variants with 7.9x median coverage per individual.

Testing for signatures of recent hybrids

We first tested whether these wide-mouth individuals represented recent (e.g. F1/F2) hybrids of *C. variegatus* and *C. desquamator* in the wild. These two young species do reproduce with each other in the lab and occasional cross-species matings have been observed in the wild (unpublished data St. John). As a first visual assessment, we conducted principal component analysis on *C. desquamator*, generalist and wide-mouth individuals to look for the genome-wide pattern expected in PCAs in which recent hybrids between two populations are included. To perform the PCA, the genetic dataset was first pruned for SNPs in strong linkage disequilibrium using the LD pruning function (--indep-pairwise 50 5 0.5) in plink (v1.9;(Purcell et al. 2007)), leaving 2.6 million variants. We then ran a principle component analysis using the eigenvectors outputted by plink's pca function (--pca). The first two principal components were plotted in R (R Core Team 2018 v3.5.0).

As a complimentary assessment, we used an admixture model-based approach to estimate the proportion of shared ancestry among individuals in our dataset using ADMIXTURE (v.1.3.0)(Alexander et al. 2009). The number of populations (K) was decided upon using ADMIXTURE's cross-validation method (--cv) across 1-10 population values of K. A K of 4 populations was then chosen using the broken-stick method. Ancestry proportions estimated by ADMIXTURE were plotted in R for the K value with the highest likelihood and the two K values surrounding to explore whether the strong signatures of population structure in Crescent Pond *C. desquamator* individuals was masking hybridization signatures in any of the wide-mouth populations.

Lastly, to discriminate among alternative evolutionary scenarios for the origin of wide-mouth and to estimate divergence time among all four species, we used demographic modeling based on the folded minor allele frequency spectrum (mSFS). The observed mSFS was computed using SFStools script from D. Marques available on github (https://github.com/marqueda/SFS-scripts). For these demographic model comparisons, we used only individuals from Osprey Lake (*C. desquamator*: n=10; *C. variegatus* : n=12; 'wide-mouths': n=11; *C. brontotheriodes*: n=13) to avoid having to additionally model complicated population structure across ponds. We contrasted 28 demographic models of different topologies and gene flow scenarios across the four groups (Fig S5;Table 1 and S2). For each model, the fit to the observed multidimensional mSFS was maximized using the composite-likelihood method fastsimcoal (v2.6.0.3;(Excoffier et al. 2013) with 100,000 coalescent simulations, 40 expectation-maximization cycles, and pooling all entries with less than 20 SNPs. For parameter estimates, we used a wide search range with log-uniform distributions with the range of some being informed by previous estimates of effective population size from MSMC and the age of the last glacial maximum (~20kya) prior to

which the lakes would have been dry on SSI. These ranges were not upper bounded by these specified priors and so the simulations were free to explore parameter space that exceeded the priors.

For each demographic model, we ran 100 independent *fastsimcoal2* runs to determine the parameter estimates with the max likelihood. The best fitting demographic model was identified using Akaike information criteria (AIC) and the best supported model was the one with the lowest AIC. To get confidence intervals for the parameter estimates from the best fitting model, we simulated 100 mSFS based on the maximum likelihood parameter estimates from our best fitting model and ran 50 independent runs with this model on each simulated SFS. The parameter point estimates from the run with the highest likelihood (of the 50 independent runs) from each simulated SFS were then used to compute 95 percentile confidence intervals as a measure of uncertainty in the parameter estimates from the observed SFS.

Testing for shared adaptive alleles

Patterns of directional selection

Various demographic histories can shift the distribution of low- and high-frequency derived variants to falsely resemble signatures of hard selective sweeps. In a previous study, we used MSMC analyses to infer histories of population size changes in *C. variegatus*, *C. desquamator* and *C. brontotheriodes* (). In this study, we repeat the same analysis for the wide-mouth. In order to account for demography of the wide-mouth population in downstream analyses, we used the MSMC (v. 1.0.1;(Schiffels and Durbin 2014) to infer historical effective population size (N_e) changes in the wide-mouth. We ran MSMC on unphased GATK-called genotypes separately for a single individual wide-mouth from Osprey Lake with 16x mean coverage across its genome

(Fig S4). As recommended in the MSMC documentation, we masked out sites with less than half or more than double the mean coverage for that individual or with a genotype quality below 20. We also excluded sites with <10 reads as recommended by Nadachowska-Brzyska et al. (Nadachowska-Brzyska et al. 2016). To scale the output of MSMC to real time and effective population sizes, we used a one-year generation time (Martin et al. 2016) and the estimated spontaneous mutation rate of 1.56 $x10^{-8}$ estimated from high coverage sequencing of two independent crosses of San Salvador Island species from a previous study (Richards et al. 2020).

Across all four populations in Osprey Lake, we looked for regions that appeared to be under strong divergent selection in the form of a hard selective sweep from the site frequency spectrum calculated with SweeD (v.3.3.4;(Pavlidis et al. 2013)). In this calculation of the composite likelihood ratio (CLR) of a sweep, we incorporated our empirical estimate of the decrease in population size for each focal population estimated from MSMC analyses in 50-kb windows across scaffolds that were at least 100-kb in length (99 scaffolds; 85.6% of the genome). We also calculated CLRs across 100,000 scaffolds consisting of neutrally evolving sequences simulated with ms-move (Garrigan and Geneva 2014), controlling for the impact of the inferred population size decreases over time for each population from MSMC runs mentioned above (Fig. S4; Table S3). The CLR ratios for the simulated datasets were then used to assess outlier CLR ratios from the empirical dataset. Regions with CLR ratios above the 95th percentile value of CLR from the neutral simulated dataset were considered candidate hard selective sweep regions (Table S3). We compared the selective sweeps across *C. variegatus, C. desquamator, C. brontotheriodes*, and the 'wide-mouth' to look for shared and unique selective sweeps among the four groups.

Characterization of candidate adaptive alleles

We searched for candidate adaptive alleles underlying divergent traits among the species by overlapping selective sweeps regions with regions of high genetic divergence in form of nearly fixed SNPs between groups. We chose to look at nearly fixed SNPs over only fixed variation to accommodate the ongoing geneflow occurring between these young species. We took two approaches to finding nearly-fixed variants: 1) a fixed threshold of $F_{st} \geq 0.95$ across all comparisons and 2) and threshold of the 99.9th percentile of F_{st}, which varies between comparisons (range of F_{st} 0.73-0.83).

We look at different combinations of groups for F_{st} calculations including a) those highly divergent between *C. variegatus* and each of the specialists, b) those unique to each specialist against all other groups and c) those shared between two or three specialists against *C. variegatus*. For nearly fixed variants with the $F_{st} \geq 0.95$ threshold, we looked for putative function of the candidate adaptive alleles by looking at gene annotations of any gene the variant was in or near (within 20-kb of the gene, which is within the 50-kb LD decay estimate). We used available gene annotations from model organisms of mice and zebrafish from MGI, ZFIN, and we checked other annotation databases and studies for verification of putative function, including Phenoscape Knowledgebase (https://kb.phenoscape.org/#/home), NCBI's PubMed (https://www.ncbi.nlm.nih.gov/pubmed), and the Gene Ontology database using AMIGO2 (Balsa-Canto et al. 2016) .

For the 99.9th percentile F_{st} variants: We performed gene ontology (GO) enrichment analyses for genes near candidate adaptive variants using ShinyGo (v.0.51;(Ge and Jung 2018)). In the *C. brontotheroides* reference genome annotations (described in *de novo* genome assembly and annotation section), gene symbols largely match human gene symbols. Thus, the best matching

species when we searched for enrichment across biological process ontologies curated mostly human gene functions.

Timing of selection on candidate adaptive alleles

Lastly, we also determined the relative age of candidate adaptive alleles by generating estimates of coalescent times using the R package starTMRCA (v0.6-1;(Smith et al. 2018)). For each candidate adaptive allele that was unique to the three specialists and the 16 shared alleles between *C. desquamator* and wide-mouth, a 1-Mb window surrounding the variant was extracted into separate vcfs for individuals within each species. The program requires complete genotype data so we first filtered out any individuals with more than 10% missing data (1 *C. brontotheriodes* and 2 'wide-mouth' individuals). With Tassel5 (Bradbury et al. 2007), we then used the LD KKNI command to infer missing sites based on LD for remaining individuals with less than 10% missing data. Subsequently we removed the small number of sites with any missing data across individuals within each population.

These sets of variants were then analyzed in starTMRCA with the mutation rate of 1.56×10^{-8} substitutions per base pair , and a recombination rate of 3.11×10^{-8} (from genome-wide recombination estimate for stickleback; (Roesti et al. 2013)). Each analysis was run three times per focal adaptive allele and all runs were checked for convergence between and within runs. Most runs rapidly converged within the initial 6000 steps, but 5 runs that didn't not converge after an additional 4000 steps were discarded from further analysis.

Acknowledgements

We thank Priya Moorjani, Michelle St John, Joe McGirr, Jackie Galvez, David Tian, Austin Patton, and Joe Heras for helpful comments on the manuscript; the Gerace Research Centre and Troy Day for logistical support; the governments of the Bahamas for permission to collect and export samples; the Vincent J. Coates Genomics Sequencing Center and Functional Genomics Laboratory at UC Berkeley for performing whole-genome library preparation and sequencing (supported by NIH S10 OD018174 Instrumentation Grant), and the University of California Berkeley for computational resources. **Funding:** This work was funded by the National Science Foundation DEB CAREER grant #1749764, National Institutes of Health grant 5R01DE027052-02, the University of North Carolina at Chapel Hill, and the University of California, Berkeley to CHM, and graduate research funding from the Museum of Vertebrate Zoology to EJR.

Author Contributions

Conceptualization – EJR,CHM; Data Curation: EJR,CHM; Statistical analyses: EJR; Resources: CHM; Visualization: EJR,CHM; Writing – original draft: EJR; Writing – review & editing: EJR, CHM

Declaration of Interests

The authors declare no competing interest.

References

Alexander, D. H., J. Novembre, and K. Lange. 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19:1655–1664. Cold Spring Harbor Lab.

Arnold, S. J., M. E. Pfrender, and J. A. G. 2001. The Apaptive Landscape as a conceptual bridge between mico- and macroevolution. Genetica 112–113:9–32.

Balsa-Canto, E., D. Henriques, A. Gabor, and J. R. Banga. 2016. AMIGO2, a toolbox for dynamic modeling, optimization and control in systems biology. Bioinformatics 32:1–2.

Barghi, N., J. Hermisson, and C. Schlötterer. 2020. Polygenic adaptation: a unifying framework to understand positive selection. Nat. Rev. Genet. 21:769–781. Springer US.

Bell, G. 2013. Evolutionary rescue and the limits of adaptation. Philos. Trans. R. Soc. B Biol.

Sci. 368:20120080. Royal Society.
Blount, Z. D., J. E. Barrick, C. J. Davidson, and R. E. Lenski. 2012. Genomic analysis of a key innovation in an experimental *Escherichia coli* population. Nature 489:513–518.
Blount, Z. D., R. Maddamsetti, N. A. Grant, S. T. Ahmed, T. Jagdish, J. A. Baxter, B. A. Sommerfeld, A. Tillman, J. Moore, J. L. Slonczewski, J. E. Barrick, and R. E. Lenski. 2020. Genomic and phenotypic evolution of escherichia coli in a novel citrate-only resource environment. Elife 9:1–64.
Bolnick, D. I. 2006. Multi-species outcomes in a common model of sympatric speciation. J. Theor. Biol. 241:734–744.
Bradbury, P. J., Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss, and E. S. Buckler. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635.
Brice, C., Z. Zhang, D. Bendixsen, and R. Stelkens. 2021. Hybridization Outcomes Have Strong Genomic and Environmental Contingencies. Am. Nat. 198:E000–E000.
Burch, C. L., and L. Chao. 2004. Epistasis and its relationship to canalization in the RNA virus φ6. Genetics 167:559–567.
Calsbeek, R., and D. J. Irschick. 2007. The quick and the dead: Correlational selection on morphology, performance, and habitat use in island lizards. Evolution (N. Y). 61:2493–2503.
Charlesworth, B., R. Lande, and M. Slatkin. 1982. A Neo-Darwinian Commentary on Macroevolution. Evolution (N. Y). 36:474–498. [Society for the Study of Evolution, Wiley].
Chevin, L. M., and F. Hospital. 2008. Selective sweep at a quantitative trait locus in the presence of background genetic variation. Genetics 180:1645–1660.
Clark, P. U., A. S. Dyke, J. D. Shakun, A. E. Carlson, J. Clark, B. Wohlfarth, J. X. Mitrovica, S. W. Hostetler, and A. M. McCabe. 2009. The Last Glacial Maximum. Science (80-.). 325:710–714.
Clayton-Smith, J., J. O'Sullivan, S. Daly, S. Bhaskar, R. Day, B. Anderson, A. K. Voss, T. Thomas, L. G. Biesecker, P. Smith, A. Fryer, K. E. Chandler, B. Kerr, M. Tassabehji, S. A. Lynch, M. Krajewska-Walasek, S. McKee, J. Smith, E. Sweeney, S. Mansour, S. Mohammed, D. Donnai, and G. Black. 2011. Whole-exome-sequencing identifies mutations in histone acetyltransferase gene KAT6B in individuals with the say-barber-biesecker variant of Ohdo syndrome. Am. J. Hum. Genet. 89:675–681. The American Society of Human Genetics.
Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, and 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158.
DePristo, M. A., E. Banks, R. Poplin, K. V Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernytsky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, and M. J. Daly. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43:491–8.
Dickinson, M. E., A. M. Flenniken, X. Ji, L. Teboul, M. D. Wong, J. K. White, T. F. Meehan, W. J. Weninger, H. Westerberg, H. Adissu, C. N. Baker, L. Bower, J. M. Brown, L. Brianna Caddle, F. Chiani, D. Clary, J. Cleak, M. J. Daly, J. M. Denegre, B. Doe, M. E. Dolan, S.

M. Edie, H. Fuchs, V. Gailus-Durner, A. Galli, A. Gambadoro, J. Gallegos, S. Guo, N. R. Horner, C. wei Hsu, S. J. Johnson, S. Kalaga, L. C. Keith, L. Lanoue, T. N. Lawson, M. Lek, M. Mark, S. Marschall, J. Mason, M. L. McElwee, S. Newbigging, L. M. J. Nutter, K. A. Peterson, R. Ramirez-Solis, D. J. Rowland, E. Ryder, K. E. Samocha, J. R. Seavitt, M. Selloum, Z. Szoke-Kovacs, M. Tamura, A. G. Trainor, I. Tudose, S. Wakana, J. Warren, O. Wendling, D. B. West, L. Wong, A. Yoshiki, D. G. MacArthur, G. P. Tocchini-Valentini, X. Gao, P. Flicek, A. Bradley, W. C. Skarnes, M. J. Justice, H. E. Parkinson, M. Moore, S. Wells, R. E. Braun, K. L. Svenson, M. Hrabe de Angelis, Y. Herault, T. Mohun, A. M. Mallon, R. Mark Henkelman, S. D. M. Brown, D. J. Adams, K. C. Kent Lloyd, C. McKerlie, A. L. Beaudet, M. Bucan, S. A. Murray, M. McKay, B. Urban, C. Lund, E. Froeter, T. LaCasse, A. Mehalow, E. Gordon, L. R. Donahue, R. Taft, P. Kutney, S. Dion, L. Goodwin, S. Kales, R. Urban, K. Palmer, F. Pertuy, D. Bitz, B. Weber, P. Goetz-Reiner, H. Jacobs, E. Le Marchand, A. El Amri, L. El Fertak, H. Ennah, D. Ali-Hadji, A. Ayadi, M. Wattenhofer-Donze, S. Jacquot, P. André, M. C. Birling, G. Pavlovic, T. Sorg, I. Morse, F. Benso, M. E. Stewart, C. Copley, J. Harrison, S. Joynson, R. Guo, D. Qu, S. Spring, L. Yu, J. Ellegood, L. Morikawa, X. Shang, P. Feugas, A. Creighton, P. C. Penton, O. Danisment, N. Griggs, C. L. Tudor, A. L. Green, C. Icoresi Mazzeo, E. Siragher, C. Lillistone, E. Tuck, D. Gleeson, D. Sethi, T. Bayzetinova, J. Burvill, B. Habib, L. Weavers, R. Maswood, E. Miklejewska, M. Woods, E. Grau, S. Newman, C. Sinclair, E. Brown, S. Ayabe, M. Iwama, and A. Murakami. 2016. High-throughput discovery of novel developmental phenotypes. Nature 537:508–514. Nature Publishing Group.

Dieckmann, U., and M. Doebeli. 1999. On the origin of species by sympatric speciation. Nature 400:354. Macmillan Magazines Ltd.

Dobzhanksy, T. 1937. Genetics and the origin of species. Columbia University Press, New York.

Erwin, D. H. 2021. A conceptual framework of evolutionary novelty and innovation. Biol. Rev. 96:1–15.

Erwin, D. H. 2019. Prospects for a General Theory of Evolutionary Novelty. J. Comput. Biol. 26:735–744.

Excoffier, L., I. Dupanloup, E. Huerta-Sánchez, V. C. Sousa, and M. Foll. 2013. Robust Demographic Inference from Genomic and SNP Data. PLOS Genet. 9:e1003905. Public Library of Science.

Flint, H. J., K. P. Scott, P. Louis, and S. H. Duncan. 2012. The role of the gut microbiota in nutrition and health. Nat. Rev. Gastroenterol. Hepatol. 9:577–589.

Garrigan, D., and A. Geneva. 2014. msmove: A modified version of Hudson's coalescent simulator ms allowing for finer control and tracking of migrant genealogies.

Gavrilets, S. 2004. Fitness landscape and origin of species. Princeton University Press, Princeton, NJ.

Gavrilets, S. 2014. Models of speciation: where are we now? J. Hered. 105:743–755.

Gavrilets, S., and A. Vose. 2005. Dynamic patterns of adaptive radiation. Proc. Natl. Acad. Sci. U. S. A. 102:18040–18045.

Ge, S. X., and D. Jung. 2018. ShinyGO: a graphical enrichment tool for animals and plants. bioRxiv 2.

Geneva, A. J., C. A. Muirhead, S. B. Kingan, and D. Garrigan. 2015. A New Method to Scan Genomes for Introgression in a Secondary Contact Model. PLoS One 10:e0118621. Public Library of Science.

Gillespie, J. H. 1984. Molecular Evolution Over the Mutational Landscape. Evolution (N. Y).

38:1116–1129. [Society for the Study of Evolution, Wiley].
Givnish, T. J., E. L. Burkhardt, R. E. Happel, and J. D. Weintraub. 1984. Carnivory in the Bromeliad Brocchinia reducta , with a Cost / Benefit Model for the General Restriction of Carnivorous Plants to Sunny , Moist , Nutrient-Poor Habitats. Am. Nat. 124:479–497.
Givnish, T. J., K. J. Sytsma, J. Smith, W. Hahn, B. DH, and E. Burkhardt. 1997. Molecular evolution and adaptive radiation in *Brocchinia* (Bromeliaceae: Pitcairnioideae) atop tepuis of the Guyana Shield. Pp. 259–311 *in* Molecular evolution and adaptive radiation. Cambridge University Press, Cambridge.
Goldschmidt, R. 1933. Some Aspects of Evolution. Science (80-.). 78:539–547. American Association for the Advancement of Science.
Goldschmidt, R. 1940. The material basis of evolution. Yale University Press.
Good, B. H., M. J. McDonald, J. E. Barrick, R. E. Lenski, and M. M. Desai. 2017. The dynamics of molecular evolution over 60,000 generations. Nature 551:45–50.
Grant, P. R., and R. B. Grant. 2008. How and Why Species Multiply. Princeton University Press.
Gross, B. L., and L. H. Rieseberg. 2005. The ecological genetics of homoploid hybrid speciation. J. Hered. 96:241–252.
Hagey, F. M., and J. . Mylroie. 1995. Pleistocene lake and lagoon deposits, San Salvador Island, Bahamas. Geol. Soc. Am. 30:77–90.
Heras, J., and C. Martin. 2021. Nonadaptive radiation of the gut microbiome in an adaptive radiation of Cyprinodon pupfishes with minor shifts for scale-eating. bioRxiv. Cold Spring Harbor Laboratory.
Hibbins, M. S., and M. W. Hahn. 2021. Phylogenomic approaches to detecting and characterizing introgression. ecoevorxiv, doi: 10.32942/osf.io/uahd8.
Höllinger, I., P. S. Pennings, and J. Hermisson. 2019. Polygenic adaptation: From sweeps to subtle frequency shifts. PLOS Genet. 15:e1008035. Public Library of Science.
Ito, H. C., and U. Dieckmann. 2007. A new mechanism for recurrent adaptive radiations. Am. Nat. 170.
Jagdish, T., J. J. Morris, B. D. Wade, and Z. D. Blount. 2020. Probing the Deep Genetic Basis of a Novel Trait in Escherichia coli* BT - Evolution in Action: Past, Present and Future: A Festschrift in Honor of Erik D. Goodman. Pp. 107–122 *in* W. Banzhaf, B. H. C. Cheng, K. Deb, K. E. Holekamp, R. E. Lenski, C. Ofria, R. T. Pennock, W. F. Punch, and D. J. Whittaker, eds. Springer International Publishing, Cham.
Joly, S., P. A. McLenachan, and P. J. Lockhart. 2009. A statistical approach for distinguishing hybridization and incomplete lineage sorting. Am. Nat. 174.
Kauffman, S., and S. Levin. 1987. Towards a general theory of adaptive walks on rugged landscapes. J. Theor. Biol. 128:11–45.
Kondrashov, A. S., and F. A. Kondrashov. 1999. Interactions among quantitative traits in the course of sympatric speciation. Nature 400:351–354. Macmillan Magazines Ltd.
Kraft, M., I. C. Cirstea, A. K. Voss, T. Thomas, I. Goehring, B. N. Sheikh, L. Gordon, H. Scott, G. K. Smyth, M. R. Ahmadian, U. Trautmann, M. Zenker, M. Tartaglia, A. Ekici, A. Reis, H. G. Dörr, A. Rauch, and C. T. Thiel. 2011. Disruption of the histone acetyltransferase MYST4 leads to a noonan syndrome-like phenotype and hyperactivated MAPK signaling in humans and mice. J. Clin. Invest. 121:3479–3491.
Lamichhaney, S., J. Berglund, M. S. Almén, K. Maqbool, M. Grabherr, A. Martinez-Barrio, M. Promerová, C.-J. Rubin, C. Wang, N. Zamani, B. R. Grant, P. R. Grant, M. T. Webster, and L. Andersson. 2015a. Evolution of Darwin's finches and their beaks revealed by genome

sequencing. Nature 518:371–375.
Lamichhaney, S., J. Berglund, M. S. Almén, K. Maqbool, M. Grabherr, A. Martinez-Barrio, M. Promerová, C.-J. Rubin, C. Wang, N. Zamani, B. R. Grant, P. R. Grant, M. T. Webster, and L. Andersson. 2015b. Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature 518:371–375.
Lee, J. Y., D. Seo, J. You, S. Chung, J. S. Park, J. H. Lee, S. M. Jung, Y. S. Lee, and S. H. Park. 2017. The deubiquitinating enzyme, ubiquitin-specific peptidase 50, regulates inflammasome activation by targeting the ASC adaptor protein. FEBS Lett. 591:479–490.
Li, H., and R. Durbin. 2011. Inference of human population history from individual whole-genome sequences. Nature 475:493–496. Nature Publishing Group.
Lindsey, H. A., J. Gallie, S. Taylor, and B. Kerr. 2013. Evolutionary rescue from extinction is contingent on a lower rate of environmental change. Nature 494:463–467.
Mani, G. S., and B. C. Clarke. 1990. Mutational order: a major stochastic process in evolution. Proc. R. Soc. London. B. Biol. Sci. 240:29–37. Royal Society.
Marsden, C. D., Y. Lee, K. Kreppel, A. Weakley, A. Cornel, H. M. Ferguson, E. Eskin, and G. C. Lanzaro. 2014. Diversity, differentiation, and linkage disequilibrium: prospects for association mapping in the malaria vector Anopheles arabiensis. G3 (Bethesda). 4:121–31.
Martin, C. H. 2016. Context dependence in complex adaptive landscapes: frequency and trait-dependent selection surfaces within an adaptive radiation of Caribbean pupfishes. Evolution (N. Y). 70:1265–1282.
Martin, C. H., J. E. Crawford, B. J. Turner, and L. H. Simons. 2016. Diabolical survival in Death Valley: recent pupfish colonization, gene flow and genetic assimilation in the smallest species range on earth. Proc. R. Soc. B Biol. Sci. 283:20152334.
Martin, C. H., K. Gould, and C. Bocklage. 2019. Surprising spatiotemporal stability and frequency-independence across multiple fitness peaks driving adaptive radiation in the wild.
Martin, C. H., and K. J. Gould. 2020. Surprising spatiotemporal stability of a multi-peak fitness landscape revealed by independent field experiments measuring hybrid fitness. Evol. Lett. 530–544.
Martin, C. H., and E. J. Richards. 2019. 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.
Martin, C. H., and P. C. Wainwright. 2013a. A remarkable species flock of *Cyprinodon* pupfishes endemic to San Salvador Island, Bahamas. Bull. Peabody Museum Nat. Hist. 54:231–240.
Martin, C. H., and P. C. Wainwright. 2013b. Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. Science 339:208–11.
Martin, C. H., and P. C. Wainwright. 2013c. On the measurement of ecological novelty: scale-eating pupfish are separated by 168 my from other scale-eating fishes. PLoS One 8:e71164.
Martin, C. H., and P. C. Wainwright. 2011. Trophic novelty is linked to exceptional rates of morphological diversification in two adaptive radiations of cyprinodon pupfish. Evolution (N. Y). 65:2197–2212.
Maynard Smith, J. 1970. Natural Selection and the Concept of a Protein Space. Nature 225:563–564.
Mayr, E. 1960. The emergence of evolutionary novelties. Evol. after Darwin.
Mcgirr, J. A., and C. H. Martin. 2018. Parallel evolution of gene expression between trophic specialists despite divergent genotypes and morphologies. Evol. Lett. 2:62–75.

Meier, J. I., D. A. Marques, S. Mwaiko, C. E. Wagner, L. Excoffier, and O. Seehausen. 2017. Ancient hybridization fuels rapid cichlid fish adaptive radiations. Nat. Commun. 8:14363. Nature Publishing Group.

Meier, J. I., R. B. Stelkens, D. A. Joyce, U. K. Schliewen, O. M. Selz, C. E. Wagner, S. Mwaiko, N. Phiri, C. Katongo, and O. Seehausen. 2019. Hybridization explains rapid adaptive radiation in Lake Mweru cichlid fi shes. Nat. Commun., doi: 10.1038/s41467-019-13278-z. Springer US.

Moran, B. M., C. Payne, Q. Langdon, D. L. Powell, Y. Brandvain, and Schumer. 2020. The genomic consequences of hybridization.

Nadachowska-Brzyska, K., R. Burri, L. Smeds, and H. Ellegren. 2016. PSMC analysis of effective population sizes in molecular ecology and its application to black-and-white Ficedula flycatchers. Mol. Ecol. 25:1058–1072.

Nieto Feliner, G., I. Álvarez, J. Fuertes-Aguilar, M. Heuertz, I. Marques, F. Moharrek, R. Piñeiro, R. Riina, J. A. Rosselló, P. S. Soltis, and I. Villa-Machío. 2017. Is homoploid hybrid speciation that rare? An empiricist's view. Heredity (Edinb). 118:513–516.

Orr, H. A. 2005. The genetic theory of adaptation: A brief history. Nat. Rev. Genet. 6:119–127.

Orr, H. A. 1998. The Population Genetics of Adaptation: The Distribution of Factors Fixed during Adaptive Evolution. Evolution (N. Y). 52:935.

Patricia Hernandez, L., A. C. Gibb, and L. Ferry-Graham. 2009. Trophic apparatus in cyprinodontiform fishes: Functional specializations for picking and scraping behaviors. J. Morphol. 270:645–661.

Patwa, Z., and L. M. Wahl. 2008. The fixation probability of beneficial mutations. J. R. Soc. Interface 5:1279–1289.

Pavlidis, P., and N. Alachiotis. 2017. A survey of methods and tools to detect recent and strong positive selection. J. Biol. Res. 24:7.

Pavlidis, P., D. Živković, A. Stamatakis, and N. Alachiotis. 2013. SweeD: Likelihood-based detection of selective sweeps in thousands of genomes. Mol. Biol. Evol. 30:2224–2234.

Poelstra, J. W., E. J. Richards, and C. H. Martin. 2018. Speciation in sympatry with ongoing secondary gene flow and a potential olfactory trigger in a radiation of Cameroon cichlids. Mol. Ecol. 27:1–19.

Polechová, J., N. H. Barton, and S. Gavrilefs. 2005. Speciation thrrough competition: a critical review. Evolution (N. Y). 59:1194–1210. The Society for the Study of Evolution.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, and P. C. Sham. 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am. J. Hum. Genet. 81:559–575.

Quandt, E. M., D. E. Deatherage, A. D. Ellington, G. Georgiou, and J. E. Barrick. 2014. Recursive genomewide recombination and sequencing reveals a key refinement step in the evolution of a metabolic innovation in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. 111:2217–2222.

Richards, E. J., and C. H. Martin. 2017. Adaptive introgression from distant Caribbean islands contributed to the diversification of a microendemic adaptive radiation of trophic specialist pupfishes. PLoS Genet. 13:1–35.

Richards, E. J., J. A. McGirr, J. R. Wang, M. E. St. John, J. W. Poelstra, M. J. Solano, D. C. O'Connell, B. J. Turner, and C. H. Martin. 2021. A vertebrate adaptive radiation is assembled from an ancient and disjunct spatiotemporal landscape. Proc. Natl. Acad. Sci. U.

S. A. 118.
Richards, E. J., J. W. Poelstra, and C. H. Martin. 2018. 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.
Richards, E., J. McGirr, J. Wang, M. St. John, J. Poelstra, M. Solano, D. O'Connell, B. Turner, and C. Martin. 2020. Major stages of vertebrate adaptive radiation are assembled from a disparate spatiotemporal landscape. , doi: 10.1101/2020.03.12.988774.
Rieseberg, L. H., O. Raymond, D. M. Rosenthal, Z. Lai, K. Livingstone, T. Nakazato, J. L. Durphy, A. E. Schwarzbach, L. A. Donovan, and C. Lexer. 2003. Major Ecological Transitions in Wild Sunflowers Facilitated by Hybridization. Science (80-.). 301:1211–1216.
Roesti, M., D. Moser, and D. Berner. 2013. Recombination in the threespine stickleback genome - Patterns and consequences. Mol. Ecol. 22:3014–3027.
Rosenzweig, B. K., J. B. Pease, N. J. Besansky, and M. W. Hahn. 2016. Powerful methods for detecting introgressed regions from population genomic data. Mol. Ecol. 25:2387–2397. Wiley/Blackwell (10.1111).
Rousselle, M., P. Simion, M.-K. Tilak, E. Figuet, B. Nabholz, and N. Galtier. 2020. Is adaptation limited by mutation? A timescale-dependent effect of genetic diversity on the adaptive substitution rate in animals. PLOS Genet. 16:e1008668. Public Library of Science.
Schiffels, S., and R. Durbin. 2014. Inferring human population size and separation history from multiple genome sequences. Nat. Genet. 46:919. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.
Schluter, D. 2009. Evidence for Ecological Speciation and Its Alternative. Science (80-.). 323:737 LP – 741.
Schluter, D., and G. L. Conte. 2009. Genetics and ecological speciation. Proc. Natl. Acad. Sci. 106:9955–9962.
Schumer, M., G. G. Rosenthal, and P. Andolfatto. 2018. What do we mean when we talk about hybrid speciation? Heredity (Edinb). 120:379–382.
Simpson, G. G. 1953. The Major Features of Evolution. Columbia University Press.
Smith, J., G. Coop, M. Stephens, and J. Novembre. 2018. Estimating Time to the Common Ancestor for a Beneficial Allele. Mol. Biol. Evol. 35:1003–1017.
Soupene, E. 2008. ATP8A1 activity and phosphatidylserine transbilayer movement. J. Receptor. Ligand Channel Res. Volume 1:1–10.
St. John, M. E., R. Holzman, and C. H. Martin. 2020. Rapid adaptive evolution of scale-eating kinematics to a novel ecological niche. J. Exp. Biol. 223:jeb217570.
Stroud, J. T., and J. B. Losos. 2016. Ecological Opportunity and Adaptive Radiation. Annu. Rev. Ecol. Evol. Syst. 47:507–532.
Svennson, E., and R. Calsbeek. 2012. The adaptive landscape in evolutionary biology. Oxford University Press.
The Heliconius Genome Consortium, K. K. Dasmahapatra, J. R. Walters, A. D. Briscoe, J. W. Davey, A. Whibley, N. J. Nadeau, A. V. Zimin, D. S. T. Hughes, L. C. Ferguson, S. H. Martin, C. Salazar, J. J. Lewis, S. Adler, S.-J. Ahn, D. a. Baker, S. W. Baxter, N. L. Chamberlain, R. Chauhan, B. a. Counterman, T. Dalmay, L. E. Gilbert, K. Gordon, D. G. Heckel, H. M. Hines, K. J. Hoff, P. W. H. Holland, E. Jacquin-Joly, F. M. Jiggins, R. T. Jones, D. D. Kapan, P. Kersey, G. Lamas, D. Lawson, D. Mapleson, L. S. Maroja, A.

Martin, S. Moxon, W. J. Palmer, R. Papa, A. Papanicolaou, Y. Pauchet, D. A. Ray, N. Rosser, S. L. Salzberg, M. a. Supple, A. Surridge, A. Tenger-Trolander, H. Vogel, P. a. Wilkinson, D. Wilson, J. a. Yorke, F. Yuan, A. L. Balmuth, C. Eland, K. Gharbi, M. Thomson, R. A. Gibbs, Y. Han, J. C. Jayaseelan, C. Kovar, T. Mathew, D. M. Muzny, F. Ongeri, L.-L. Pu, J. Qu, R. L. Thornton, K. C. Worley, Y.-Q. Wu, M. Linares, M. L. Blaxter, R. H. ffrench-Constant, M. Joron, M. R. Kronforst, S. P. Mullen, R. D. Reed, S. E. Scherer, S. Richards, J. Mallet, W. Owen McMillan, and C. D. Jiggins. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature 487:94–98. Nature Publishing Group.

Thomas, T., A. K. Voss, K. Chowdhury, and P. Gruss. 2000. Querkopf, a MYST family histone acetyltransferase, is required for normal cerebral cortex development. Development 127:2537–2548. England.

Thornton, K. R. 2019. Polygenic adaptation to an environmental shift: Temporal dynamics of variation under Gaussian stabilizing selection and additive effects on a single trait. Genetics 213:1513–1530.

Turner, B. J., D. D. Duvernell, T. M. Bunt, and M. G. Barton. 2008. Reproductive isolation among endemic pupfishes (Cyprinodon) on San Salvador Island, Bahamas: Microsatellite evidence. Biol. J. Linn. Soc. 95:566–582.

Turner, C., Z. Blount, D. Mitchell, and R. Lenski. 2015. Evolution and coexistence in response to a key innovation in a long-term evolution experiment with Escherichia coli. bioRxiv 020958.

Wainwright, P. C., D. R. Bellwood, M. W. Westneat, J. R. Grubich, and A. S. Hoey. 2004. A functional morphospace for the skull of labrid fishes: Patterns of diversity in a complex biomechanical system. Biol. J. Linn. Soc. 82:1–25.

Whittaker, R. 1977. Evolution of species diversity in land communities. Evol. Biol.

Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proc. Sixth Int. Congr. Genet. 1:356–366.

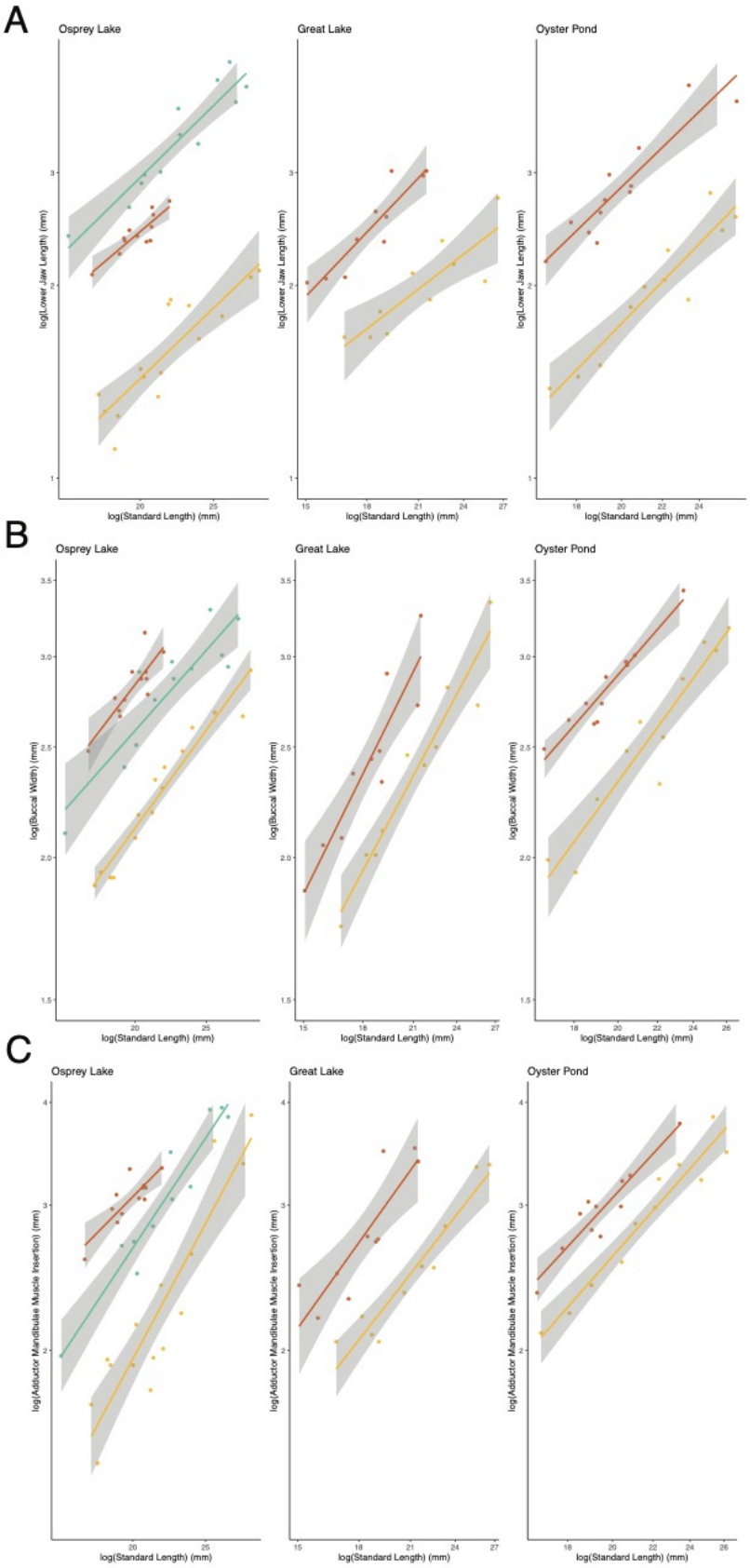
A
Osprey Lake
Great Lake
Oyster Pond
log(Lower Jaw Length) (mm)
log(Standard Length) (mm)
B
Osprey Lake
Great Lake
Oyster Pond
log(Buccal Width) (mm)
log(Standard Length) (mm)
C
Osprey Lake
Great Lake
Oyster Pond
log(Adductor Mandibulae Muscle Insertion) (mm)
log(Standard Length) (mm)

Figure S1. Relationships between three divergent traits and standard length across ponds. The relationship between standard length (mm) of individuals and their D) lower jaw length, E) buccal cavity width, and F) adductor mandibulae muscle insertion height across individuals the three species *C. desquamator* (teal), *C. variegatus* (gold), and 'wide-mouths' (red-orange) across different ponds (Osprey,Oyster and Great Lake). Colored lines represent linear model of these relationships for each species with their 95% confidence bands in gray.

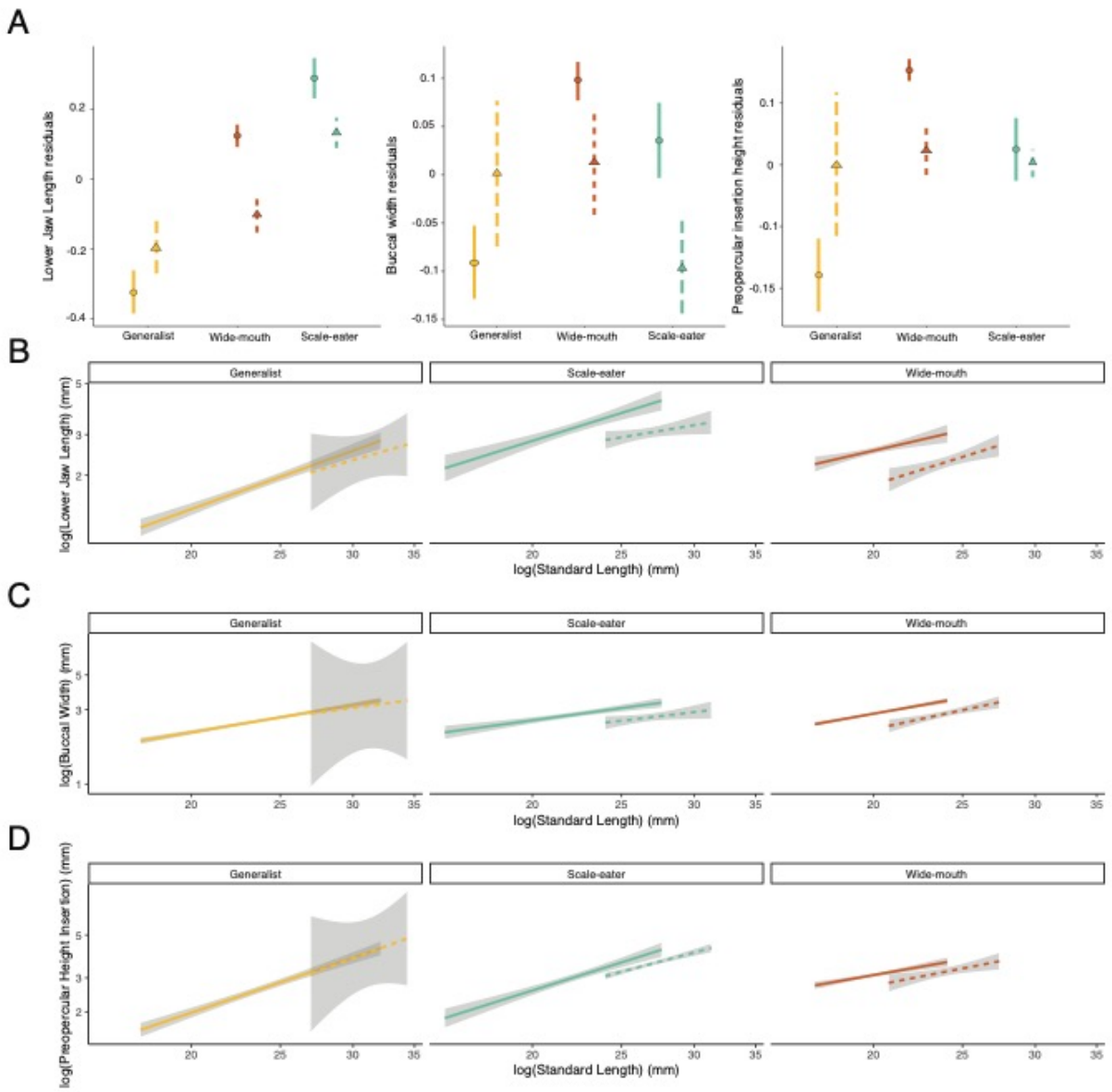

Figure S2. Comparison of the three focal divergent traits in lab and wild populations. A. 95% CI of the standardized sizes of lower jaw length (left), buccal width (center) and preopercular insertion height (right) for *C. variegatus* (gold), 'wide-mouths' (red-orange) and *C. desquamator* (teal). Lines (wild: solid; lab:dashed) are the 95% CI interval and symbols (wild: circle; lab: triangle) represent the mean values of traits from each population. Lab raised individuals include F1 and F2 generations.

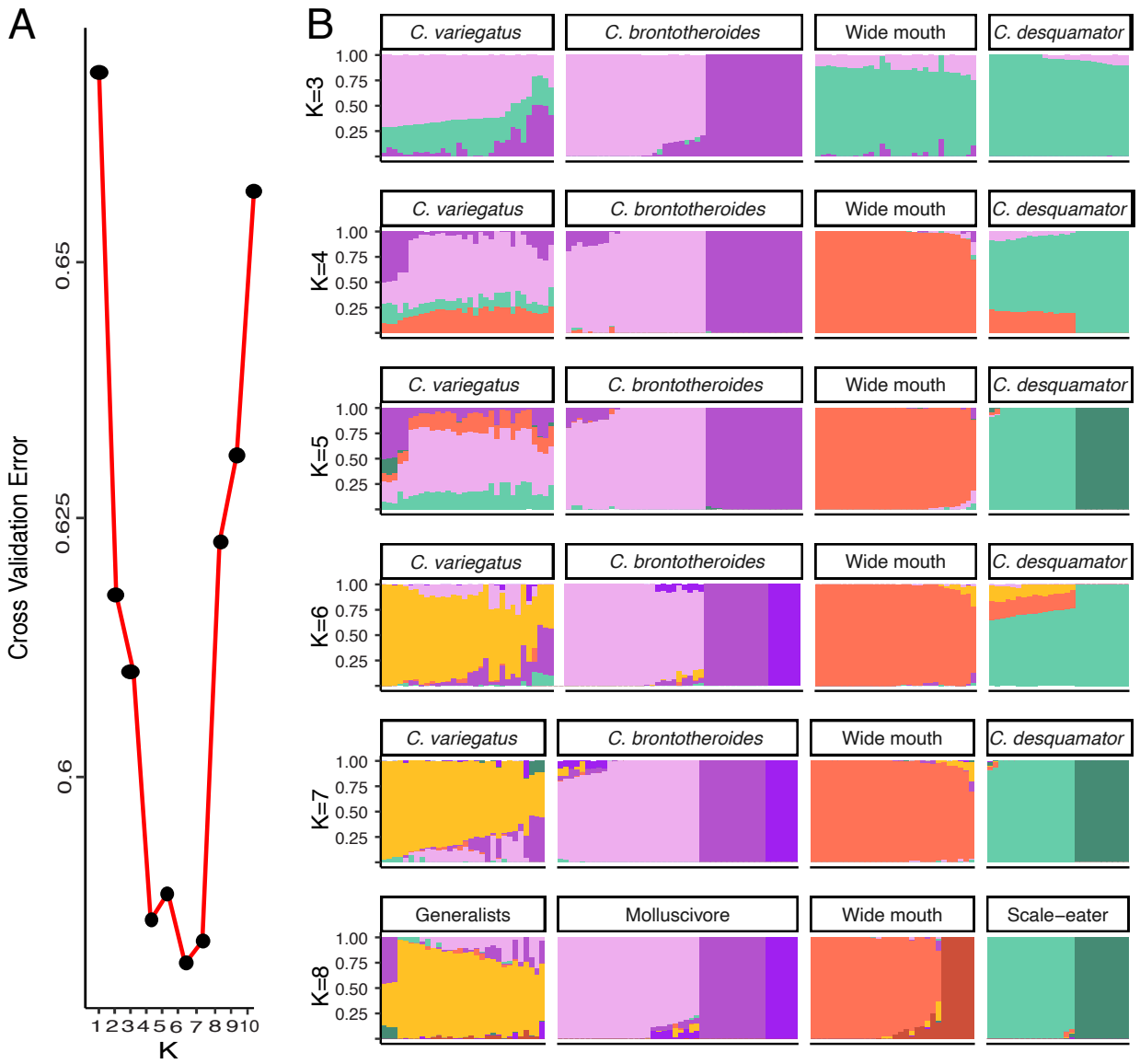

Figure S3. 'Wide-mouths' are not an admixed population. ADMIXTURE analysis of San Salvador Island populations using an LD-pruned subset of genetic variants from XXX individuals (78,840 variants). A) Cross validation error estimates from ADMIXTURE that indicate a K model of 6-7 were the best fit. B) Ancestry proportions across individuals of the three focal groups. Proportions were inferred from ADMIXTURE analyses with 2 values of K with the highest likelihood on the same LD-pruned dataset in A.

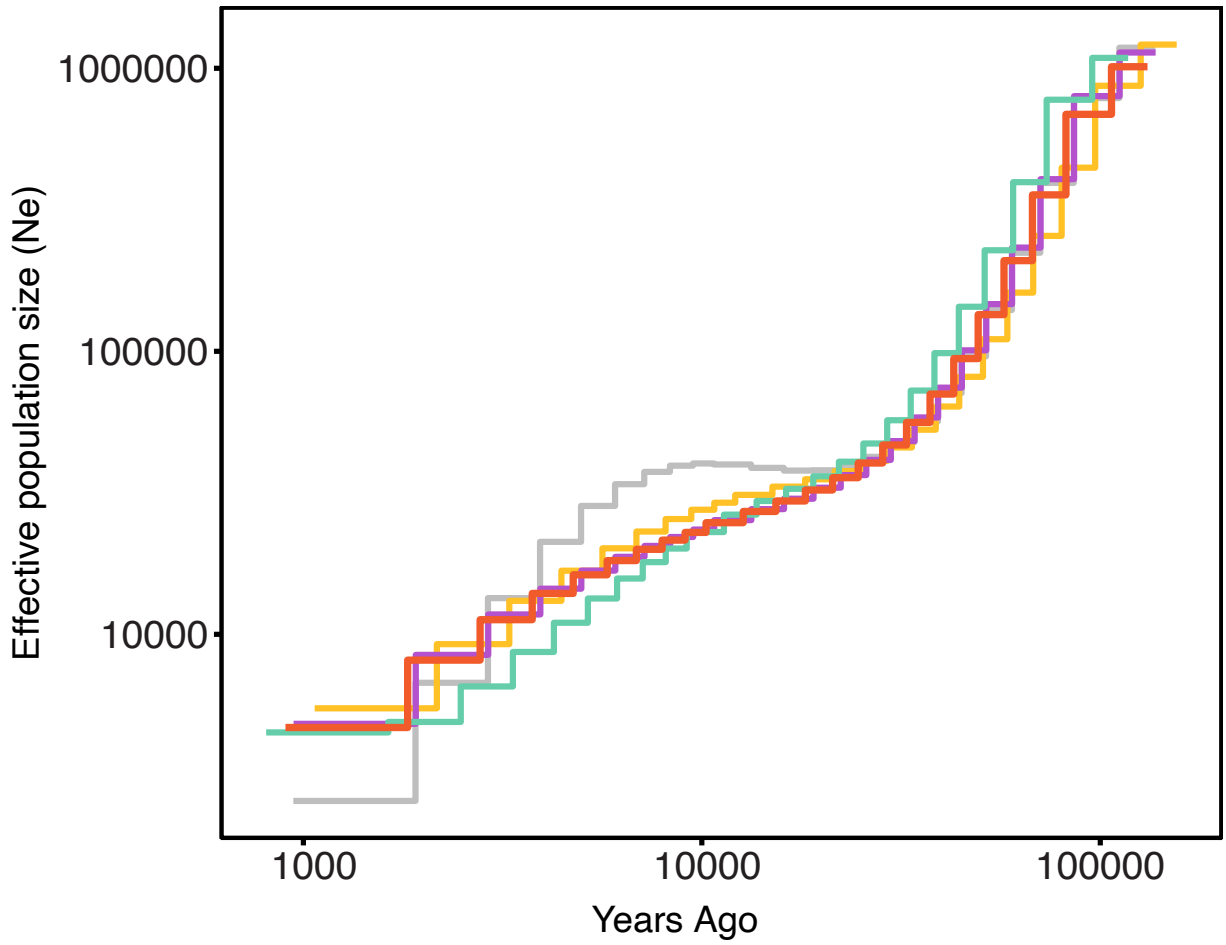

Fig S4. Changes in effective population size over time for Osprey populations inferred using MSMC (Schiffels and Durbin 2014) on high-coverage (18-36X) genomes from single individuals of each of the four species in Osprey Lake, San Salvador Island Bahamas: *C. variegatus* (gold), *C. brontotheroides* (purple), *C. desquamator* (teal), and 'wide-mouths' (red-orange). Time was scaled using a mutation rate of 1.56x10-8 mutations/basepair/gen and a generation time of 1 year.

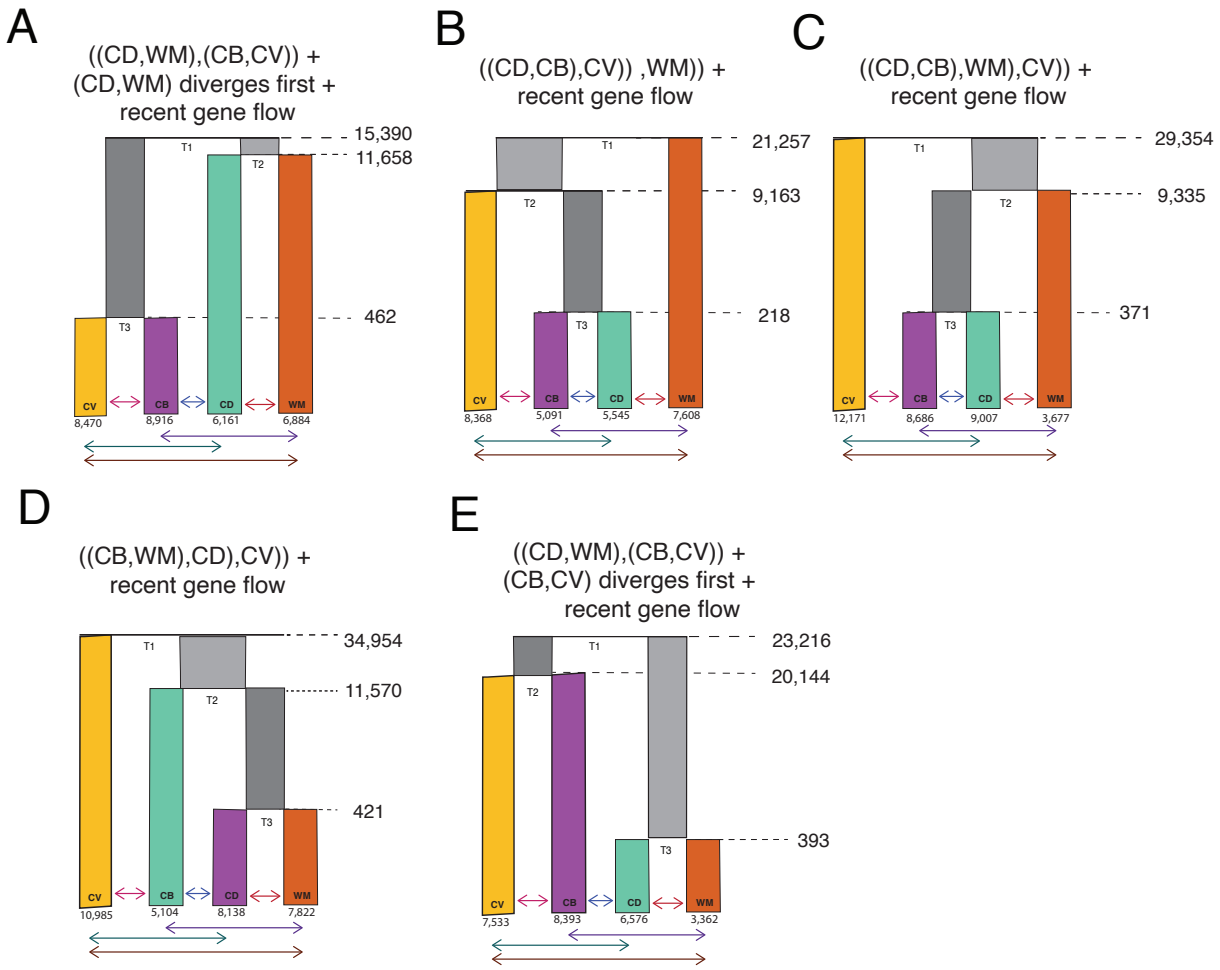

Fig S5. Top five best fitting demographic models for Osprey Lake. From a genetic dataset that was LD-pruned and had no missing information across individuals comprised of 67,400 SNPs. Listed in alphabetical order based on AIC scores (Table 1 and S2) with best supported model in A. All top models were of recent bidirectional gene flow but varied in topology among the four species.

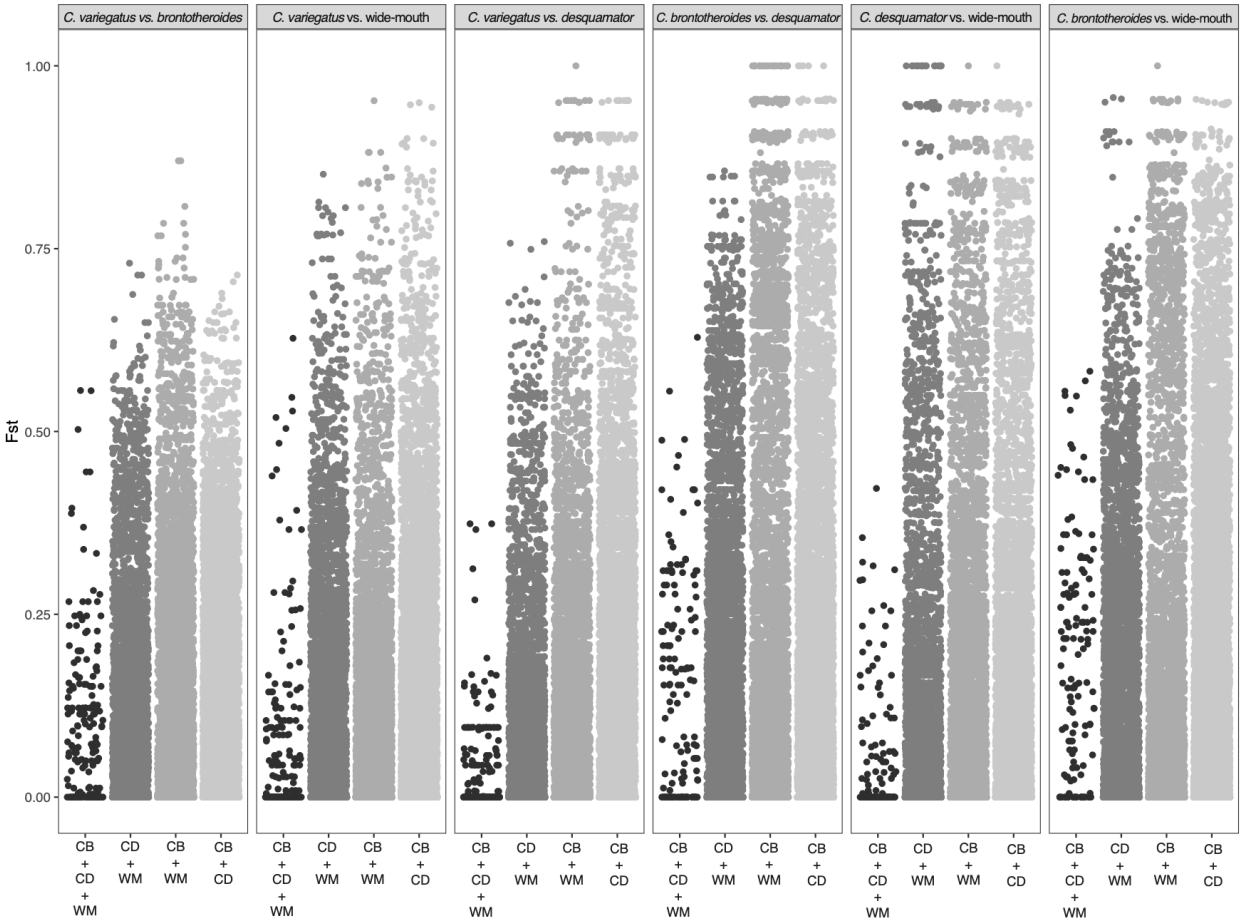

Fig S6. Genetic divergence among population in Osprey Lake in regions of shared selection across two or more specialist. Each panel represents pairwise F_{st} comparison between two populations for every SNP found in regions under shared selection across two or more specialist species on San Salvador Island: C. brontotheroides (CB); C. desquamator (CD), 'wide-mouths' (WM). Regions shared across all three specialists (CB+CD+WM) do not contain highly divergent SNPs between any of the species in the radiation. Highly divergent SNPs (Fst > 0.75) are only found in regions shared between one or two specialists.

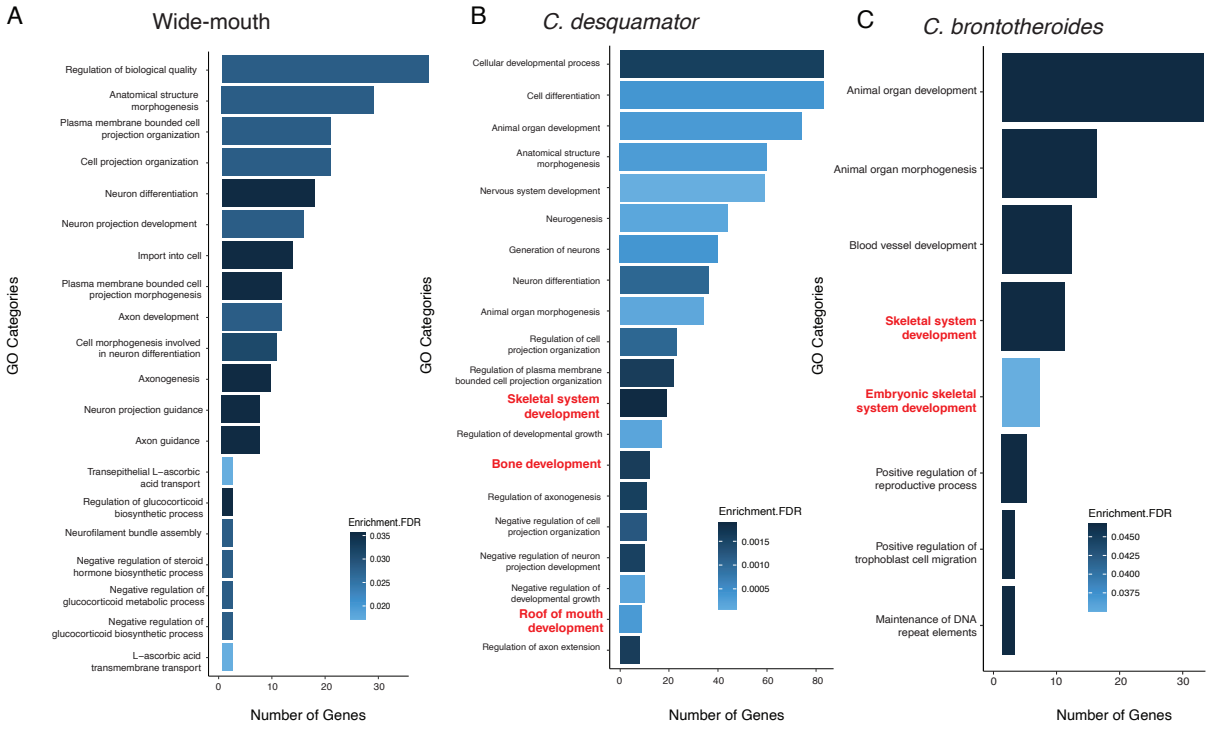

Fig S7. Top 20 enriched GO categories for divergent alleles in each of the three specialists. GO enrichment analyses were performed on genes in or near (within 20-kb) of a SNP that was under a hard selective sweep and strongly diverged from generalist species (top 1% of Fst values across genome) for a) 'wide-mouth' b) *C. desquamator*, and C) *C. brontotheriodes*. GO categories that were significantly enriched for relevant terms corresponding to craniofacial development, the major axes of morphological divergence in this radiation, are highlighted in red. All terms included were significant at an FDR < 0.05 and full list of terms in Data S1-3.

Table S1. No significant genome-wide signature of admixture among populations on San Salvador Islands. Results from three formal tests for introgression to assess whether 'wide-mouth' populations are hybrids between generalist *C. variegatus* and scale-eater *C. desquamator*: D-, f_3- and f_4-statistics across all possible combinations of Osprey Lake populations of generalist species. Significant f_3-statistic has Z-score > -2, significant f_4 and D-statistic Z-scores smaller than -2 and greater than 2. The only significant signature of admixture (bolded Z-scores) comes from D- and f_4-statistics based on relationships that violate the expected tree (((C. desquamator; 'wide-mouth'),Generalist),Outgroup) and therefore should not be interpreted as evidence of 'wide-mouths' being an admixed population. *C. variegatus* population from Rum Cay, the nearest neighbor island in the Bahamas to San Salvador Island was used as an outgroup population for D- and f_4-statistics.

f_3-statistic						
A	B	C	--	f_3	stderr	*Z*
C. variegatus	*C. desquamator*	Wide-mouth		0.004626	0.00088	5.257
f_4- statistic						
P1	P2	P3	P4	f_4	stderr	*Z*
Wide-mouth	*C. desquamator*	*C. variegatus*	Rum Cay	-0.000196	0.000164	-1.2
Wide-mouth	*C. variegatus*	*C. desquamator*	Rum Cay	0.003971	0.000268	**14.836**
Wide-mouth	*C. variegatus*	*C. brontotheroides*	Rum Cay	0.000186	0.000201	0.927
D-statistic						
P1	P2	P3	Outgroup	D	stderr	*Z*
Wide-mouth	*C. desquamator*	*C. variegatus*	Rum Cay	-0.0018	0.001498	-1.199
Wide-mouth	*C. variegatus*	*C. desquamator*	Rum Cay	0.035	0.002326	**15.068**
Wide-mouth	*C. variegatus*	*C. brontotheroides*	Rum Cay	0.0017	0.001792	0.931

Table S2. Support for the 28 demographic models for the evolution of the 'wide-mouths' from the site frequency spectrum. The likelihood and AIC scores all demographic models estimated in *fastsimcoal2* for the Osprey populations of *C. variegatus* (CV), *C. brontotheroides* (CB), *C. desquamator* (CD), and 'wide-mouths' (WM) are presented here with a complete list of all models tested reported in Table S2. Change in likelihood (ΔLnL) represents the difference in likelihood from that of a simulated SFS expected by the demographic model tested. Change in AIC (ΔAIC) represents the difference in AIC scores from that of the model with the smallest ΔLnL. All models presented here represent different divergence scenarios with recent gene flow allowed, which had better support from the data than models with no gene flow or early gene flow. Visual representations of the top five models are depicted in Figure S5.

Divergence Model	LnL	ΔLnL	AIC	ΔAIC
(1) ((CD,WM),(CB,CV)) + (CD,WM) diverges first + recent gene flow	-11701	1077	54133	0
(2) ((CD,CB),CV)) ,WM)) + recent gene flow	-11780	1112	54296	35
(3) ((CD,CB),WM),CV)) + recent gene flow	-11806	1138	54418	61
(4) ((CB,WM),CD),CV)) + recent gene flow	-11808	1140	54423	63
(5) ((CD,WM),(CB,CV)) + (CB,CV) diverges first + recent gene flow	-11846	1178	54602	101
(6) ((CD,WM), M),CV)) + recent gene flow	-11855	1187	54642	110
(7) ((CD,CB),WM),CV)) + early gene flow	-12016	1348	55364	271
(8) ((CD,CV),CB),WM)) + recent gene flow	-12024	1356	55426	279
(9) ((WM,M),CD),CV)) + early gene flow	-12086	1418	55688	341
(10) ((WM,CD),CB),CV)) + early gene flow	-11855	1419	55692	342
(11) ((CD,WM),(CB,CV)) + early gene flow	-12139	1471	55929	394
(12) ((CD,WM), CB),CV)) + no gene flow	-12315	1648	56738	570
(13) ((WM,M),CD),CV)) + no gene flow	-12326	1658	56787	581
(14) ((CD,CB),WM),CV)) + no gene flow	-12334	1667	56826	589
(15) ((CD,WM),(CB,CV)) + recent sister gene flow only	-12394	1726	57106	649
(16) ((CD,WM),(CB,CV)) + no gene flow	-12396	1728	57109	651
(17) ((WM,CV)),CB),CD)) + recent gene flow	-12399	1731	57148	654
(18) ((M,CV)) ,WM),CD)) + recent gene flow	-12401	1733	57157	656
(19) ((WM,M),CV)),CD)) + recent gene flow	-12401	1734	57163	657
(20) ((CD,WM),CV),CB)) + recent gene flow	-12420	1752	57244	675
(21) ((WM,CV)),CD),CB)) + recent gene flow	-12425	1757	57265	680
(22) ((CD,CV),WM),CB)) + recent gene flow	-12425	1757	57267	680
(23) ((WM,CV)),(CB,CD)) + no gene flow	-23245	12577	107070	11500
(24) ((WM,CV)),(CB,CD)) + early gene flow	-23658	12990	108976	11913
(25) ((WM,CV)),(CB,CD)) + recent gene flow	-23659	12991	109003	11914

(26) ((CD,CV),(CB,WM)) + early gene flow	-34983	24315	161129	23238
(27) ((CD,CV),(CB,WM)) + no gene flow	-34983	24315	161125	23238
(28) ((CD,CV),(CB,WM)) + recent gene flow	-34983	24315	161149	23238

Table S3. Parameters for selective sweep analyses in SweeD.
The 95th percentile of composite likelihood ratio threshold based on neutral simulations under the demographic scenario of decreasing population size through time inferred with MSMC (Fig S4), and the population size change and haplotype number information parameters required by SweeD for each species.

Species	CLR threshold	SweeD Commands
C. variegatus	0.40320	-folded -strictPolymorphic -G 0.407 -eN 5.45 181.8 -s 24
C. brontotheroides	0.43473	-folded -strictPolymorphic -G 0.389 -eN 5.88 196 -s 32
C. desquamator	0.40537	-folded -strictPolymorphic -G 0.218 -eN 8.11 270 -s 16
'wide-mouth'	0.39262	-folded -strictPolymorphic -G 0.203 -eN 8.57 276 -s 22

Table S4. Genetic divergence among the four populations in Osprey Lake, San Salvador Island across different thresholds of F_{st}. Relative measure of genetic divergence was calculated in pairwise combinations of the different species as the number of fixed SNPs between them, the number of fixed or nearly fixed SNPs between them, the top 1% of F_{st} between them from the distribution of F_{st} between SNPs and the genome-wide average F_{st} across all SNPs in the genome.

F_{st} Comparison	Number of fixed SNPs (F_{st} =1)	Number of nearly fixed SNPS ($F_{st} \geq$ 0.95)	Top 1%	Genome-wide average
'wide-mouth' vs *C. desquamator*	5,212	5,915	0.84	0.131
'wide-mouth' vs *C. brontotheroides*	5238	14,602	0.85	0.15
'wide-mouth' vs *C. variegatus*	247	869	0.68	0.104
C. desquamator vs *C. brontotheroides*	6941	22,088	0.86	0.17
C. desquamator vs *C. variegatus*	414	1,964	0.71	0.116
C. variegatus vs *C. brontotheroides*	110	567	0.637	0.09

Table S5. Comparison of the average absolute genetic divergence (Dxy) between the four species in Osprey Lake 'wide-mouth', *C. variegatus, C. desquamator and C. brontotheriodes* and the shared adaptive alleles between scale-eater populations.

Comparison	Mean *Dxy*
C. variegatus vs *C. brontotheroides*	0.167
C. variegatus vs *C. desquamator*	0.169
C. variegatus vs 'wide-mouths'	0.165
C. brontotheroides vs *C. desquamator*	0.162
C. brontotheroides vs 'wide-mouths'	0.156
C. desquamator vs 'wide-mouths'	0.157
Selective sweeps in 'wide-mouths'	0.123
Shared genetic variants with *C. desquamator*	0.0814
Unique genetic variants in 'wide-mouths'	0.24

Table S7. The distribution of shared adaptive alleles between the two scale-eating species across the Caribbean. Numbers represent the copies of the C. desquamatory and wide-mouth allele present in outgroup populations collected from
San Salvador Island (SS), Rum Cay (RC), Cat Island (CT), Exumas Islands (EX), Long Island (LG), Mayaguana (MG), New Providence Island (NPI) in the Bahamas. Lagunas Bavaros (BAV) and Etang Saumautre (ETA) in Dominican Republic, Sarasota Estuary in Florida (FL), Curacoa (CUR), Caicos Island (CAI) and Isla Margarita (MAR) in Venezuela.

			C. variegatus								*C. laciniatus*	*C. higuey*	*C. bondi*	*C. dearborni*		
Allele	Gene	Distribution	SS	RC	CT	EX	LG	MG	FL	NC	NPI	BAV	ETA	CUR	CAI	MAR
HiC_scaffold_6:907101	NA	SGV	0	0	1	0	0	0	0	12	0	0	0	0	0	0
HiC_scaffold_6:907425	NA	SGV	1	0	0	0	0	0	0	7	0	0	0	0	0	0
HiC_scaffold_6:907665	NA	SGV	0	0	0	0	0	0	0	7	0	0	0	0	0	0
HiC_scaffold_6:908229	NA	SGV	1	0	1	0	0	0	0	6	3	0	0	0	0	0
HiC_scaffold_6:908889	NA	SGV	1	0	0	0	0	0	0	6	0	0	0	0	0	0
HiC_scaffold_6:923182	NA	SGV	0	5	0	1	0	0	1	0	0	0	0	0	0	0
HiC_scaffold_6:923590	NA	SGV	0	5	0	2	0	0	0	0	0	0	0	0	0	0
HiC_scaffold_6:924203	NA	de novo	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HiC_scaffold_6:929621	NA	SGV	1	2	1	0	0	0	0	0	0	0	0	0	0	0
HiC_scaffold_9:31763392	daam2	SGV	6	3	0	1	0	1	1	1	6	1	0	0	0	0
HiC_scaffold_9:31763400	daam2	SGV	8	5	0	1	1	0	1	8	10	8	0	0	0	0
HiC_scaffold_24:5390450	usp50	SGV	1	1	0	0	0	0	0	0	1	0	0	0	1	0
HiC_scaffold_43:27385265	NA	SGV	3	1	0	2	1	1	0	5	1	0	1	2	0	11
HiC_scaffold_43:27389200	NA	SGV	5	1	0	0	1	1	1	12	12	2	1	3	1	12
HiC_scaffold_44:14963656	atp8a1	SGV	3	1	0	1	0	1	1	12	12	2	1	3	1	12
HiC_scaffold_60:1798737	znf214	SGV	3	0	0	0	0	1	0	0	2	0	0	0	0	0