We get by with a little help from our friends: shared adaptive variation provides a bridge to novel ecological specialists during adaptive radiation

Emilie J. Richards[1,2] ,Christopher H. Martin[1,2]

[1]Department of Integrative Biology, University of California, Berkeley

[2]Museum of Vertebrate Zoology, University of California, Berkeley

Correspondence: *chmartin@berkeley.edu*

Keywords: ecological speciation, adaptive radiation, novelty, speciation genomics, adaptive introgression, trophic innovation

Abstract

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 an intermediate ‘wide-mouth’ scale-eating ecomorph in a sympatric radiation of *Cyprinodon* pupfishes, illuminating the 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* on San Salvador Island, Bahamas, but is genetically differentiated, morphologically distinct, and often consumes scales. We then compared the timing of selective sweeps on shared and unique adaptive variants in trophic specialists to characterize their adaptive walk. Shared adaptive regions swept first in both the specialist *desquamator* and the intermediate ‘wide-mouth’ ecomorph, followed by unique sweeps of introgressed variation in ‘wide-mouth’ and de novo variation in *desquamator*. The two scale-eating populations additionally shared 9% of their hard selective sweeps with molluscivores *C. brontotheroides*, despite no single common ancestor among specialists. Our work provides a new microevolutionary framework for investigating how major ecological transitions occur and illustrates how both shared and unique genetic variation can provide a bridge for multiple species to access novel ecological niches.

Introduction

Rapid bursts of diversification and repeated bouts of speciation like those seen in adaptive radiations contradict current mechanistic speciation models that predict diversification should slow with time as available niche space becomes increasingly subdivided and disruptive selection becomes weaker with each recurrent speciation event (e.g. [1–3]). Diversification on complex adaptive landscapes with multiple empty fitness peaks corresponding to different niches provides an alternative mechanism to niche subdivision [4–6]. However, these landscapes present a new problem to our mechanistic understanding of adaptive radiations: How do populations manage to escape local optima, cross fitness valleys, and access new fitness peaks [7–10]? Colonizing new fitness peaks on the adaptive landscape presents challenges because it requires transitions in behaviors, morphological traits, or a combination of the two that allow organisms to adapt to new ecological niches [11]. Spectacular ecological transitions do often occur during adaptive radiation, such as blood-drinking [12] or plant carnivory [13,14], yet it is still poorly understood how such seemingly discontinuous transitions occur.

Recent conceptual frameworks for understanding adaptation to novel fitness peaks suggest that these major ecological transitions likely occur in stages of potentiation, actualization and refinement [15,16]. The initial emergence of a novel trait likely requires further refinement to become successfully incorporated into the functional ecology of an organism. Several experimental lab studies suggest that novel ecological transitions are highly contingent on accruing a series of mutations that incrementally refine adaptations to colonize new fitness peaks [16,17]. This idea that genetic background is important in setting the stage for adaptation also underlies many hypotheses for adaptive radiation, such as the hybrid swarm and syngameon hypotheses – in which radiations are driven by acquiring novel combinations of alleles through the exchange of genetic variation either from distinct lineages outside the radiation or within the radiation itself [18]. However, we are only just beginning to explore how gene flow and shared genetic variation gives recipient lineages access to new fitness peaks in the wild and generates adaptive radiations [6].

An adaptive radiation of trophic specialist pupfishes on San Salvador Island (SSI) in the Bahamas is an excellent system for understanding how the rapid evolution of major ecological transitions occurs in nature. This radiation contains a widespread generalist pupfish species (*Cyprinodon variegatus*) that occurs in sympatry with two previously described trophic specialists that are endemic to the hypersaline lakes on the island: a molluscivore (*C. bronotheroides*) with a novel nasal protrusion which is an oral-sheller of gastropods [19] and a scale-eating specialist (*C. desquamator*) with two-fold larger oral jaws [20]. The evolutionary novelties in this system originated recently; the lakes on SSI were dry during the last glacial maximum 6-20 kya years ago [21,22]. Intriguingly, we recently discovered a fourth species of pupfish living in sympatry with the two specialists and generalist on SSI [23]. This species exhibits intermediate jaw morphology between *C. desquamator* and *C. variegatus* (figure 1). Here we refer to this new ecomorph as the 'wide-mouth' because its mouth is wider than any other species in the radiation. The multi-peak fitness landscape driving this radiation suggests that *C. desquamator* is isolated by a large fitness valley from *C. variegatus* and *C. brontotheroides* [9] and this intermediate 'wide-mouth' may provide clues about the microevolutionary processes underlying how the observed novel fitness peaks are traversed in the wild.

Here we first investigate the position of the 'wide-mouth' on the ecological spectrum from generalist to scale-eating specialist using a combination of morphological, behavioral, dietary, and genomic data. We then estimated the demographic history of the ‘wide-mouth’ and explored the spatial origins and timing of selection on shared and unique genetic variation involved in adaptation to scale-eating to better understand this ecological transition. Our results suggest that while intermediate in jaw length, which is known to be relevant for the highly specialized scale-eater *C. desquamator*, *C* sp. ‘wide-mouth’ demonstrates transgressive morphology and a distinct genetic background. Our investigation of the timing of selection and genetic origins of the adaptive alleles shared and unique between the two scale-eating species indicates divergent adaptive walks that are highly dependent on their genetic background. Despite shared origins, access to unique genetic variation in each of the two scale-eating sister species likely resulted in distinct adaptive walks and ultimately contributed to the diversity of ecological specialists observed in this radiation.

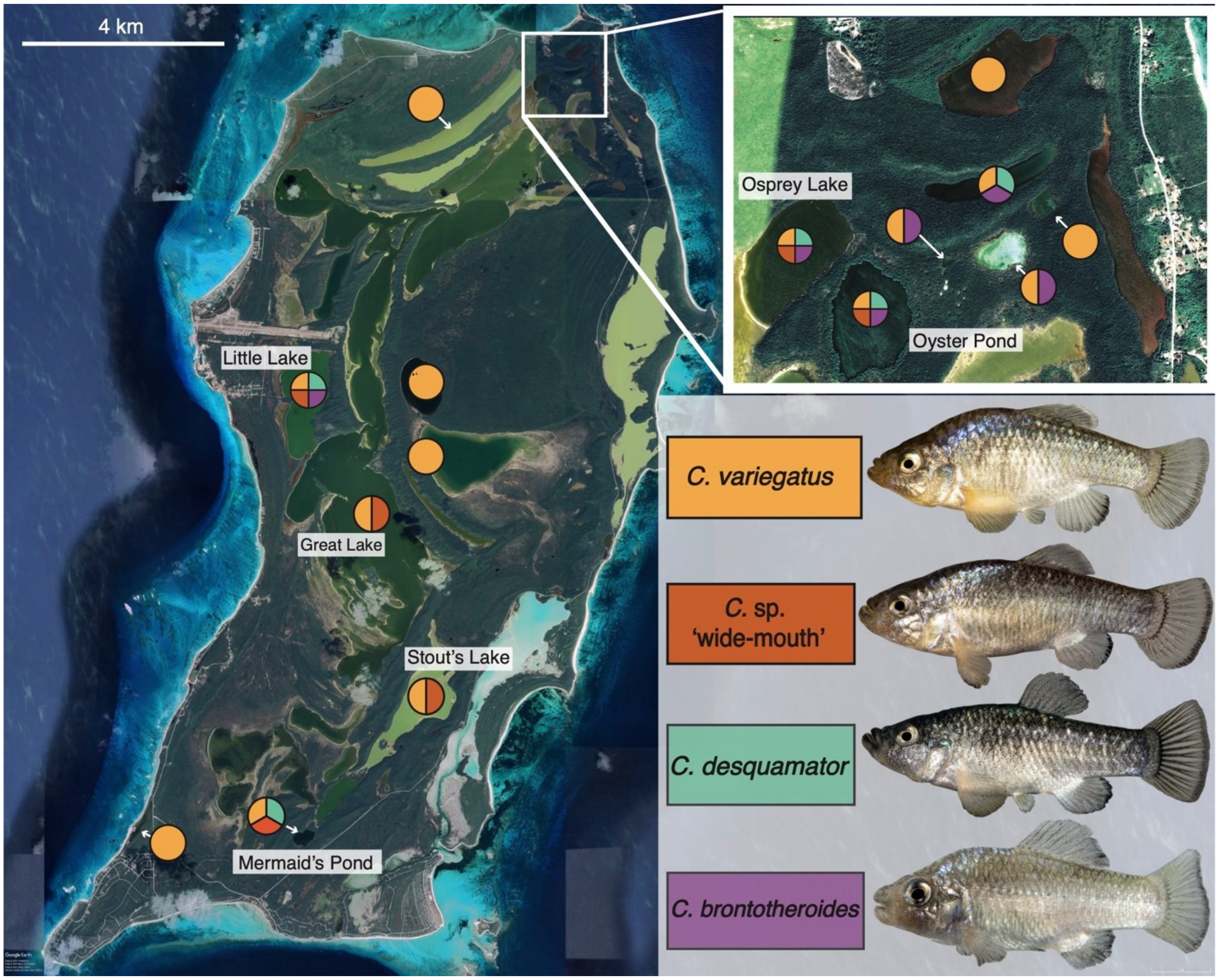

Figure 1. The San Salvador Island radiation of pupfish**.** Pie charts indicate the presence of sympatric *Cyprinodon* species in each lake and are color-coded with representative pictures of generalist *C. variegatus* (gold), recently discovered *C.* sp. 'wide-mouth' ecomorph (red-orange) with intermediate jaws, scale-eater *C. desquamator* (teal) with the largest oral jaws, and molluscivore *C. brontotheroides* (purple) with characteristic nasal protrusion. Labeled lakes contain all known *Cyprinodon* populations on the island that were sampled for this study. Satellite image from Google Earth.

Methods

Ecological and morphological characterization of ‘wide-mouth’ scale-eater

C. variegatus, C. desquamator, and *C.* sp. ‘wide-mouth’ individuals from 3 lake populations (Osprey Lake, Great Lake, and Oyster Pond) in which we had sufficient specimens (n=84; *C. brontotheroides* not shown) were measured for 9 external morphological traits using digital calipers. Traits were selected for specific connections to foraging performance which differed across the three species in a previous study [9]. We also characterized diet for *C. variegatus, C. desquamator,* and ‘wide-mouth’ in Osprey Lake from stomach content analyses (*n*=10 per species) and stable isotope analyses of muscle tissue from wild-collected samples (*n*=75). Dietary overlap was characterized by comparison of population mean scale count from gut contents using ANOVA, ellipse areas, and bivariate means on isotope biplots using SIBER [24]. See supplemental methods for more details on sample sizes and analyses.

Genomic library preparation and variant filtration

To explore the evolutionary history of *C.* sp. ‘wide-mouth’, we sequenced whole genomes of 22 individuals following protocols used in a previous study [25] that included genomes from *C. variegatus, C. desquamator*, and *C. brontotheroides*. Our final genetic dataset after filtering contained 6.4 million variants across 110 individuals from the four species (7.9x median coverage). See supplemental methods for the full sequencing and genotyping protocol.

Genomic origins of the *C.* sp. ‘wide-mouth’ scale-eater

We first tested whether these *C*. sp. ‘wide-mouth’ individuals represented recent (e.g. F1/F2) hybrids of *C. variegatus* and *C. desquamator* in the wild using principal component and ADMIXTURE analyses to look for the genome-wide pattern expected in PCAs when recent hybrids between two populations are included. We also used formal tests for introgression and admixed populations, f_3 and f_4-statistics [26], to assess whether ‘wide-mouth’ are the byproduct of recent admixture. Finally, we used *fastsimcoal2* (v2.6.0.3;[27]), a demographic modeling approach based on the folded minor allele frequency spectrum (mSFS), to discriminate among alternative evolutionary scenarios for the origin of ‘wide-mouth’ and estimated divergence times among all four species based on the best model fit from an AIC test (see supplementary methods for more detail).

Characterization of unique and shared adaptive alleles among specialists

Across all four populations in Osprey Lake, we looked for regions that showed evidence of a hard selective sweep using SweeD (v.3.3.4;[28]). The composite likelihood ratio (CLR) for a hard selective sweep was calculated in 50-kb windows across scaffolds that were at least 100-kb in length (99 scaffolds; 85.6% of the genome). Significance thresholds were determined using CLR values from neutral sequences simulated under MSMC inferred demographic scenarios of historical effective population size changes (Supplemental methods; figure S1; table S1).

Next, we searched for candidate adaptive alleles associated with species divergence by overlapping selective sweep regions with regions of high genetic divergence based on fixed or nearly fixed SNPs between species. We chose to also look at regions with nearly fixed SNPs (F_{st} ≥ 0.95) to accommodate ongoing gene flow between these young species. F_{st} between the populations and species was calculated per variant site using –weir-pop-fist function in vcftools (v.0.1.15;[29]).

Timing of selection on candidate adaptive alleles

We also determined the relative age of candidate adaptive alleles by generating estimates of coalescent times using starTMRCA (v0.6.1;[30]). 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 each species. These sets of variants were then analyzed in starTMRCA with a mutation rate of 1.56 x 10^{-8} substitutions per base pair (from Caribbean pupfishes [25]) and a recombination rate of $3.11\text{x}10^{-8}$ (from stickleback; [31]). 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 did not converge after an additional 4000 steps and were discarded from further analysis. See supplementary methods for more details on timing analyses.

Characterization of adaptive introgression adaptive alleles in *C.* sp. 'wide-mouth'

Lastly, we investigated the spatial origins of adaptive alleles shared and unique to the two scale-eating specialists by searching in our previous study spanning Caribbean-wide outgroup populations for these alleles [25]. Adaptive alleles were assigned as standing genetic variation if observed in any population outside SSI or de novo if they were only observed within populations on SSI. Additionally, we investigated signatures of introgression across the genome of *C.* sp. 'wide-mouth' and *C. desquamator* to determine if they showed evidence of adaptive introgression from outgroup generalist populations as observed previously [25]. See supplementary methods for more details on introgression analyses.

Results

‘Wide-mouth’ ecomorph is ecologically intermediate and morphologically distinct

We found that the ‘wide-mouth’ ecomorph is morphologically distinct from *C. desquamator* and *C. variegatus* across a suite of craniofacial traits (figure 2A-B). The lower jaw length of ‘wide-mouth’ was intermediate between *C. desquamator* and *C. variegatus* (figure 2C), while the mouth width and adductor mandibulae height were 8% larger in ‘wide-mouth’ than *C. desquamator* (figure 2D-E). These morphological differences were consistent across multiple lakes (figure S2). Small modifications in craniofacial morphology among these species have major impacts on scale-eating performance in this system by altering kinematic traits such as peak gap size which is partially controlled by the length of the lower jaw, jaw protrusion distance, and the angle of the lower jaw relative to the suspensorium [32].

C. sp. ‘wide-mouth’ also did not show morphological divergence comparable to that observed in the molluscivore *C. brontotheroides.* The molluscivore specialist presents an opposing axis of morphological divergence to the scale-eating specialists, with shorter oral jaw length and larger eye diameter than even the generalist *C. variegatus*, in addition to a novel nasal protrusion of the maxilla not observed in any other Cyprinodontidae species [33].

Morphological traits were heritable in a common garden laboratory environment after one generation: lab-reared *C.* sp. ‘wide-mouth’ displayed significantly larger mouth width than *C. desquamator* (t-test; $P = 0.003$) and maintained their characteristic intermediate jaw lengths (ANOVA; $P = 0.03$, figure S3). There was also some evidence of phenotypic plasticity in both lab-reared *C. desquamator* and *C.* sp. 'wide-mouth' compared to wild individuals likely caused by the common lab diet. See supplementary results for more details.

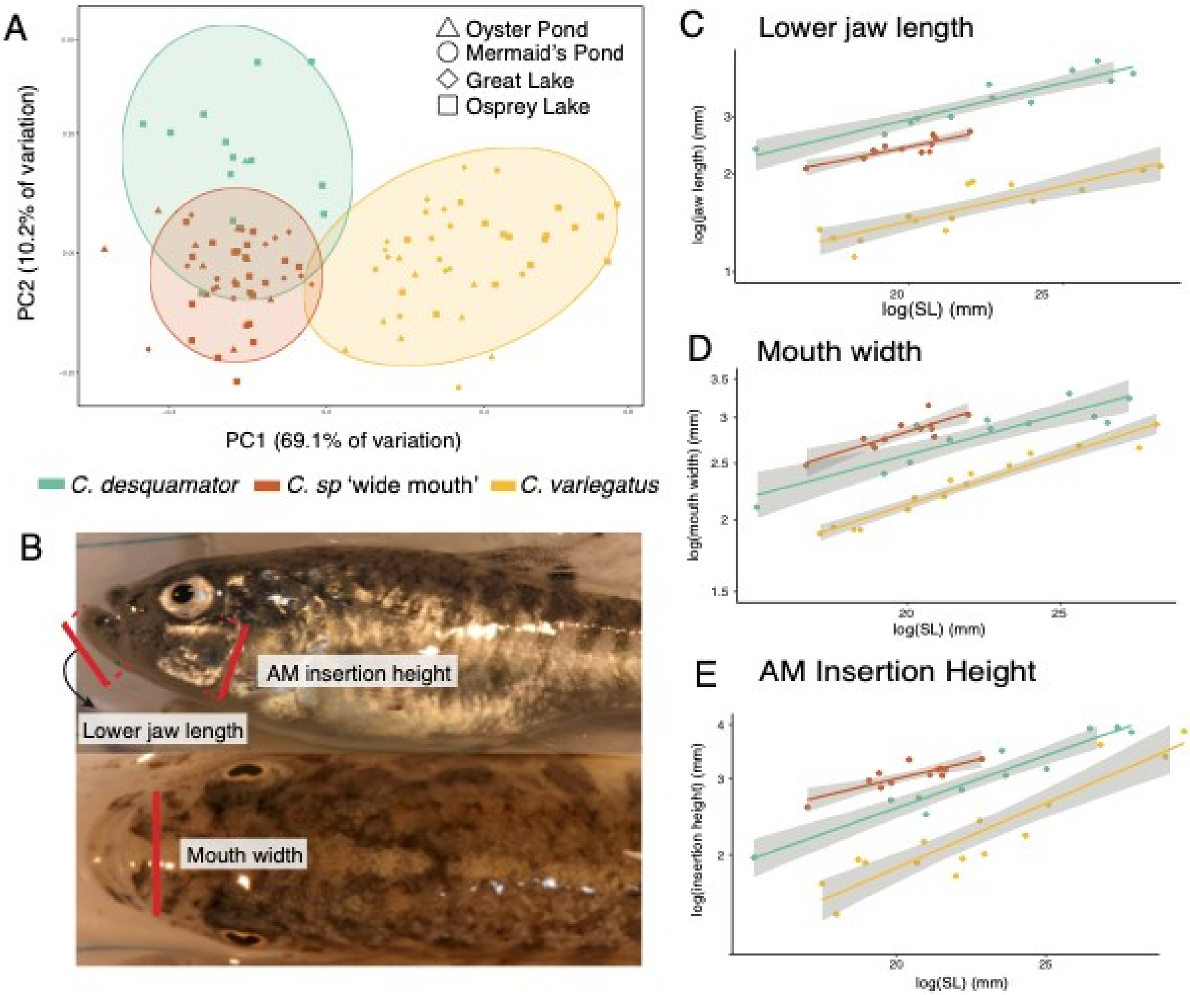

Figure 2. *C.* sp. 'wide-mouth' has distinct morphology within the San Salvador Island adaptive radiation. A) First two principal components of morphological diversity for 8 size-corrected traits and 95% confidence ellipses by species (*C. variegatus* : gold; *C.* sp. 'wide-mouth': red-orange; *C. desquamator*: teal; *C. brontotheriodes* not shown). PC1 is mainly described by lower jaw length and PC2 by adductor mandibulae insertion height, mouth width, and neurocranium width. B) Depictions of the three external measurements that best distinguished *C.* sp. 'wide-mouth' from both *C. desquamator* and *C. variegatus*, measured using digital calipers. C-E) The

relationship between standard length (mm) of individuals and their C) lower jaw length, D) buccal cavity width, and E) adductor mandibulae insertion height (AM insertion) across individuals of 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’ ingested scales, but at a significantly lower frequency than *C. desquamator* (Wilcoxon Rank Sum test, $P = 0.004$; figure 3A). We did not detect any scales in *C. variegatus* guts (figure 3A). Detritus made up the rest of the *C.* sp. ‘wide-mouth’ and *C. desquamator* diets and was the dominant component of *C. variegatus* gut contents, except for a single individual with one mollusk shell. A previous study that characterized contents of *C. variegatus*, *C. brontotheroides*, and *C. desquamator* populations across several ponds also found detritus to be the dominant component of each species’ diet (49-71%) and nearly zero scales in the gut contents of *C. variegatus* and *C. brontotheroides* [33].

The intermediate scale-eating dietary niche of the wide-mouth ecomorph is complemented by our stable isotope analyses, which provide long-term snapshots of the carbon sources and relative trophic levels in these species. Osprey Lake individuals collected on the same day from the same site differed in δ15N levels across species (ANOVA, $P = 4.55\text{x}10^{-6}$; figure 3B and S4); ‘wide-mouth’ δ15N was intermediate between *C. variegatus* and *C. desquamator* (Tukey HSD; $P = 1.34 \text{ x}10^{-5}$ & $1.11 \text{ x}10^{-4}$ respectively), supporting its intermediate trophic position. Additionally, SIBER analyses indicate this distinct trophic positioning based on the lack of extensive overlap in niche space measured by standard ellipse areas and bivariate means with 95% confidence intervals of isotope values among the species (figure 3B).

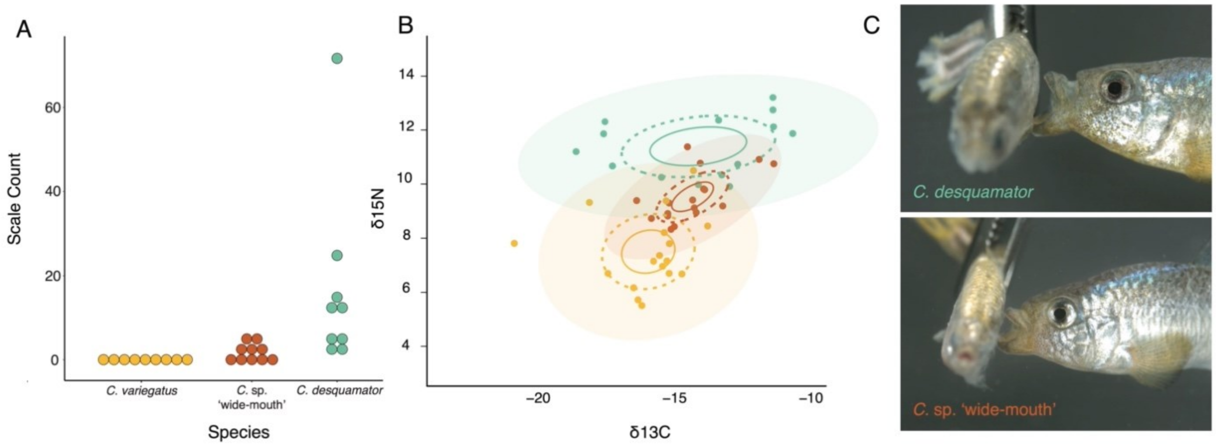

Figure 3. *C.* sp. ‘wide-mouth’ ingests scales. A) Scale counts from gut content analysis of the hindgut of Osprey pupfish populations (10 individuals per species). B) Relative trophic position (δ15N stable isotope ratio) and dietary carbon source (δ13C stable isotope ratio) with 95% confidence ellipses for generalist and scale-eating species. Solid lines represent 95% confidence intervals around bivariate mean, dotted lines represent standard ellipse areas corrected for sample size (contain 40% of data; SEAc), shaded circles represent ellipse area that contain 95% of the data calculated using the R package SIBER. C) Still images of scale-eating strikes in *C. desquamator* and *C.* sp. ‘wide-mouth’ filmed at 1100 fps on a Phantom VEO 440S camera.

<u>*C.* sp. ‘wide-mouth’ did not result from hybridization between *C. variegatus* and *C. desquamator*</u>

Several lines of genomic evidence from PCA, ADMIXTURE, and *f-statistic* analyses support the ‘wide-mouth’ ecomorph as a genetically distinct species rather than a recent hybrid between *C. desquamator* and *C. variegatus* on SSI (figure 4A-C & S5-6; see Supplementary material for more details). Demographic modeling of divergence and gene flow on SSI places *C.* sp. ‘wide-mouth’ as sister to *C. desquamator*, supporting previous phylogenetic inference [34]. In the best supported model of 28 demographic models tested (table S2), ‘wide-mouth’ and *C. desquamator* diverged 11,658 years ago (95 CI: 8,257-20,113 years; figure S7; table S2) with ongoing gene flow. Additionally, *C. desquamator* and *C.* sp. ‘wide-mouth’ are more genetically diverged from each other than to the generalist *C. variegatus* (e.g. F_{st} in figure 4C).

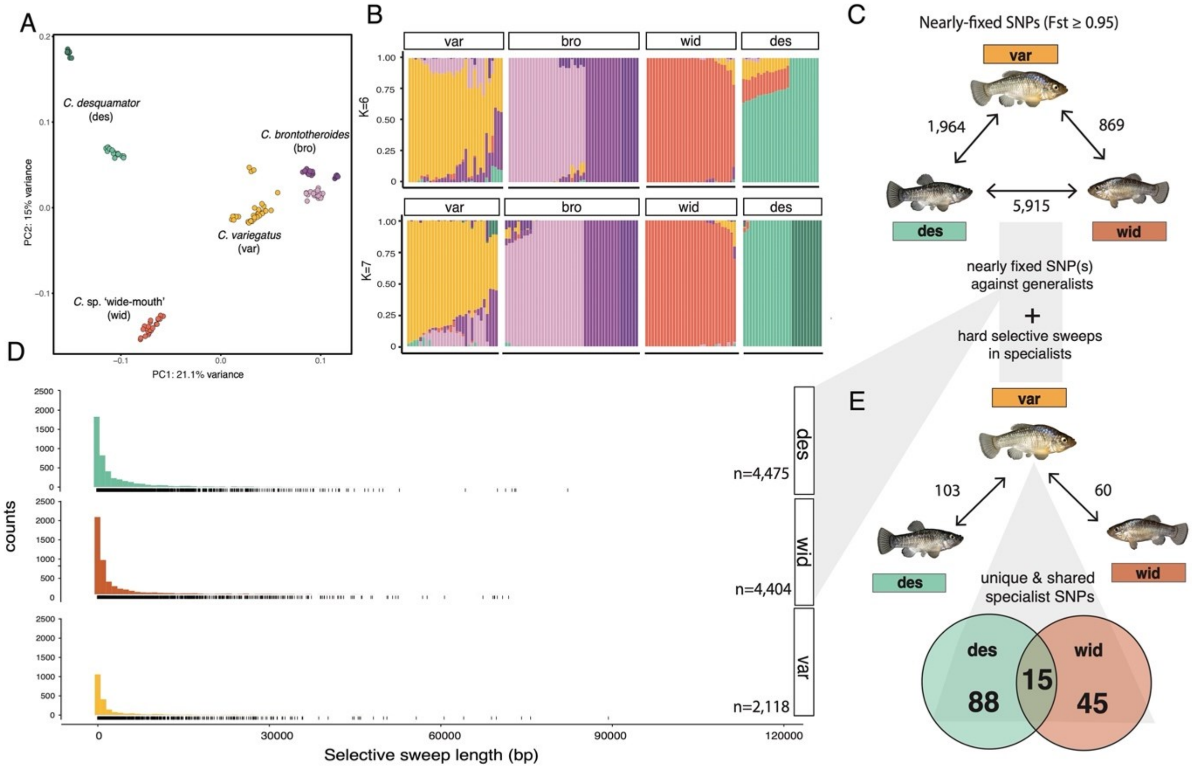

Figure 4. Patterns of selection and genetic divergence in specialist genomes**.** A) Principal components analysis of the four focal groups on San Salvador Island based on an LD-pruned subset of genetic variants (78,840 SNPs). B) Ancestry proportions across individuals of the four focal groups. Proportions were inferred from ADMIXTURE analyses with 2 values of K with the highest likelihood on the same LD-pruned dataset in A. C) Selective sweep length distributions across generalist and scale-eating species. Rug plot below each histogram represents the counts of selective sweeps in different length bins. D) The total number of fixed or nearly-fixed SNPs ($F_{st} \geq 0.95$) between each group in Osprey Pond. E) The number of adaptive alleles (fixed or nearly-fixed SNPS [$F_{st} \geq 0.95$] relative to *C. variegatus* and under selection in the scale-eating

specialists in Osprey Lake. Venn diagram highlights those adaptive alleles that are unique to each specialist and shared with the other specialist.

C. sp. 'wide-mouth' is comprised of both shared and unique adaptive alleles

Next we looked at regions of the genome in both *C. desquamator* and *C*. sp. wide-mouth that showed strong evidence of hard selective sweeps. We found 6 shared hard selective sweeps in both species containing a total of 15 SNPS that were fixed or nearly-fixed compared to the sympatric generalist *C. variegatus* (figure 4E): 10 SNPs were in unannotated regions, two were in the introns of the gene *daam2,* and three were in putative regulatory regions (with 20-kb) of the genes *usp50, atp8a1,* and *znf214* (one variant each). Shared adaptive alleles in the gene *daam2,* a wnt signaling regulator*,* are intriguing because knockdown of this gene causes abnormal snout morphology in mice [35] and abnormal cranial and skeletal development in zebrafish [36].

We also found unique sets of adaptive alleles in *C*. sp. 'wide-mouth' and *C. desquamator* (figure 4E). None of the adaptive alleles unique to *C*. sp. 'wide-mouth' were in or near genes annotated for craniofacial phenotypes in model organisms, despite its distinctive craniofacial morphologies. In *C. desquamator,* three of 12 unique adaptive alleles were in or near genes associated with or known to affect 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 or near the gene *bri3bp*, which is located inside a QTL region for cranial height in pupfish [37]. *C. brontotheriodes* also contained at least one unique candidate craniofacial adaptive allele: a non-synonymous coding substitution in the gene *kat6b* (figure 5), which is associated with abnormal craniofacial morphologies, including shorter mandibles, in mice [38].

This pattern of unique alleles relevant to craniofacial phenotypes in specialists *C. brontotheriodes* and *C. desquamator*, but not *C.* sp. 'wide-mouth', holds even if we lower the threshold to the top 1 percentile of F_{st} outliers between specialists and generalist (see supplemental results; figure S9-10).

The origins of adaptive alleles in *C.* sp. *'wide-mouth' and desquamator*

The adaptive alleles shared by *C. desquamator* and *C.* sp. 'wide-mouth' occurred as low frequency standing genetic variation in the Caribbean, with the exception of a single de novo allele on SSI located in an unannotated region on scaffold 6 (figure 5;table S3). The adaptive alleles unique to *C. desquamator* and *C.* sp. 'wide-mouth' also predominantly came from standing genetic variation (84% and 81%, respectively). 14% of adaptive alleles unique to *C. desquamator* were de novo mutations to SSI and 2% occurred in candidate introgression regions (table S4). We found the opposite in *C.* sp. 'wide-mouth': only 4% of their unique adaptive alleles were de novo mutations whereas 15% occurred in candidate introgression regions (table S4). This adaptive introgression occurs with generalist populations sampled from North Carolina and Laguna Bavaro in the Dominican Republic (table S5; figure S11). Using the Relative Node Depth (RND) statistic, we also discovered that 5 of the 6 shared adaptive alleles (all except for the unannotated region on scaffold 43;table S6) appear introgressed between *C. desquamator* and *C.* sp. 'wide-mouth', suggesting a substantial contribution of introgression to the adaptive alleles observed in scale-eating specialists.

Timing of selection on adaptive alleles reveals features of the adaptive walk to scale-eating

Selective sweeps occurred much more recently in both populations than their inferred divergence times (figure 5). Intriguingly, selection on 4 of the 6 adaptive alleles occurred significantly earlier in *C. desquamator* than 'wide-mouth'. Only a single adaptive allele had an older median age estimate in 'wide-mouth' than *C. desquamator*, although the 95% HPD intervals overlapped between the species (figure 5). Additionally, overall we found a significant difference in timing of selection between shared and unique adaptive alleles in the two scale-eater populations (ANOVA *P*-value = 0.00478). In *C. desquamator*, shared adaptive alleles swept before any unique adaptive alleles (Tukey HSD *P*-value = 0.003217; figure 5). For the 'wide-mouth', shared adaptive alleles with *C. desquamator* also generally swept before those unique to the species, despite these unique alleles being standing and introgressed variation from the Caribbean (figure 5). However, this difference in timing between shared and unique adaptive alleles in C. sp 'wide-mouth' was not significant due to one unique adaptive allele (*slitrk5)* that swept first (figure 5; ANOVA, Tukey HSD; *P* = 0.8367). This adaptive allele resides in a region that appears to be introgressed with the Laguna Bavaro generalist population in the Dominican Republic where this allele also show signs of a hard selective sweep [25]. The older age estimate of this sweep in *C.* sp 'wide-mouth' might be due to older shared selection for the alleles in other Caribbean populations before introgression with *C.* sp 'wide-mouth'. All other introgressed adaptive alleles in *C.* sp. 'wide-mouth' swept more recently than shared sweeps with *desquamator*, including the shared de novo allele, and were not under selection in outgroup generalist populations.

Intriguingly, all but one of the de novo adaptive alleles in *C. desquamator* swept at the same time in the recent past (figure 5). Only one of these adaptive alleles in *olfm1* region has 95% HPD age range that overlaps with the next oldest selective sweep on standing genetic variation (*gnaq*; figure 5), suggesting a discrete stage of selection on new mutations in *C. desquamator*.

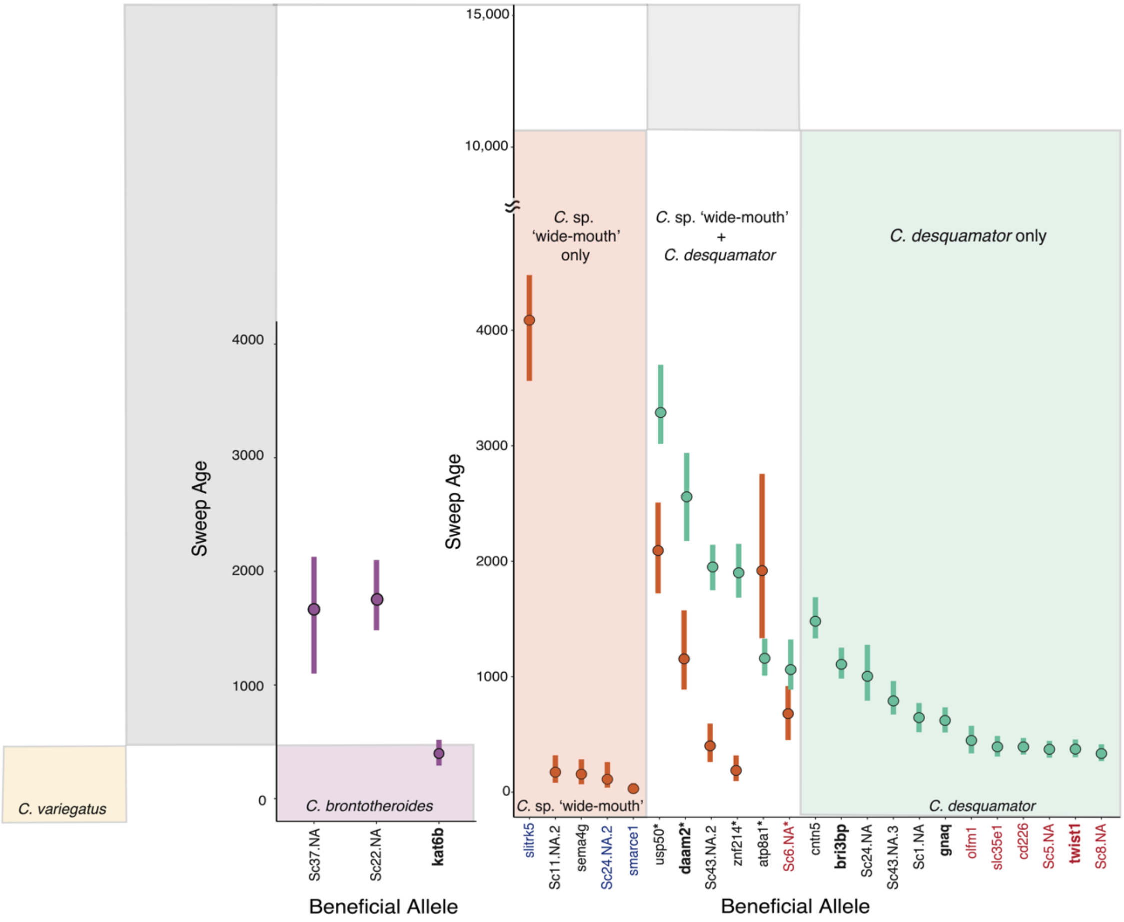

Figure 5. Timing of selection on adaptive alleles in trophic specialists nested within the demographic history of the radiation. The median and 95% HPD estimates for the timing of selection on sets of fixed or nearly fixed SNPs (named by the gene they are in or within 20-kb of) for the three specialist populations found in sympatry in Osprey Lake (sweeps in *C. variegatus* not shown). The age of each beneficial allele is color coded by the species and the inferred demographic history is displayed in the background for comparison. Gene names highlighted in bold are associated with oral jaw size. Gene names are colored by source of genetic variation (de novo: red; introgressed with outgroup: blue; standing genetic variation:

black). Gene names with asterisk indicate those inferred as introgressed between *C. desquamator* and *C.* sp. 'wide-mouth'.

Shared signatures of selection across the three specialists in the radiation

Lastly, we compared the genetic divergence and selection patterns observed in the two scale-eating specialists to the divergent molluscivores specialist *C. brontotheroides* to investigate the extent of allele sharing among all three trophic specialists in this adaptive radiation. We found that no fixed or nearly-fixed alleles relative to the generalist *C. variegatus* were shared across all three specialists (figure S9-S10; supplementary results). However, we did find evidence of 44 shared selective sweeps across all three specialist populations that were not shared with *C. variegatus* populations (figure S12). These shared regions were significantly enriched for genes annotated for metabolic processes (figure S12), suggesting shared selection for metabolizing the more protein-rich diet across all three trophic specialists (also see [39]).

Discussion

Discovery of a new cryptic intermediate scale-eater highlights the power of reusing genetic variation to access novel niches

The hallmark of adaptive radiation is a rapid burst of diversification which is predicted by theory to slow down over time as niche subdivision increases [6]. 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 [40]. This could happen through exploitation of new trophic levels created by new species or physical alterations of the environment by new species that may create additional opportunities for speciation (reviewed in [6,41]). The diversity begets diversity hypothesis can also be visualized as the exploration of a complex multi-peaked fitness landscape; as species in the radiation colonize new peaks, this provides access to additional neighboring fitness peaks to fuel rapid radiation. Our discovery of a cryptic new scale-eating species through morphological, dietary, and genomic analyses revealed shared nearly-fixed or fixed adaptive alleles in both scale-eating species relative to the generalist *C. variegatus*. While *C.* sp. 'wide-mouth' is ecologically intermediate in its scale-eating behavior, our estimates of the relative timing of selective sweeps suggest that these shared alleles were first selected upon in the more specialized scale-eater *C. desquamator*, although unaccounted for demographic differences may also be contributing to differences in estimated timing between species.

Intriguingly, the shared adaptive alleles between *C. desquamator* and *C.* sp. 'wide-mouth' have potentially introgressed recently rather than selected upon in their shared common ancestor. Five of the six regions surrounding these shared adaptive alleles showed patterns of high genetic similarity consistent with introgression (table S6). Alternatively, this genetic similarity may also be caused by strong background selection on shared ancestral variation. Effective population sizes are not drastically different between the two species and exon density is not in the upper tail of the genome wide-distribution (figure S1;table S6), two conditions in which background selection tends to confound adaptive introgression inferences [42,43]. However, we do not have extensive knowledge of recombination breakpoints in this non-model system to distinguish between strong background selection on shared ancestral variation and adaptive introgression scenarios for each candidate adaptative introgression region.

We also found strong signatures of introgression in *C.* sp. 'wide-mouth' genomes from outgroup generalist populations that were not shared with *C. desquamator* (figure S11; table S7). Craniofacial morphology is a major axis of diversification between trophic specialists in this system [44], yet *C.* sp. 'wide-mouth' appears to have little unique genetic variation relevant for craniofacial traits compared to the other two specialists (figure S10). Despite this, they do exhibit transgressive craniofacial phenotypes not seen in the other specialists. An intriguing implication of these findings is that hybridization may allow different species to share many of the same adaptive alleles to occupy distinct but similar niches, in line with the syngameon and 'diversity begets diversity' hypotheses of adaptive radiation [18,40].

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

The foundational model 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 the local optimum [45]. The distinct timing of selection across different adaptive alleles in both *C. desquamator* and *C.* sp 'wide-mouth' suggests that the ecological transition from generalist to novel scale-eating specialist involved an adaptive walk in which selection on a beneficial allele was contingent on prior fixation of other adaptive alleles in each specialists' genetic background (see supplemental materials for further discussion). This is best highlighted by the pattern observed in *C. desquamator* in which nearly all de novo mutations swept at the same time in a distinct selective stage from other adaptive variants rather than being selected upon as they originated (figure 5).

The (un)predictability of adaptive walks to novel ecological niches

A recent study characterizing genotypic fitness landscapes underlying the transition from *C. variegatus* and *C. desquamator* revealed a rugged landscape with many local fitness peaks, likely due to epistatic interactions among alleles [46]. The staggered timing of selection on alleles lends support to this finding. Epistasis can reduce the number of adaptive walks selection will promote [47], and might explain why the same adaptive alleles were the first to undergo hard selective sweeps in both 'wide-mouth' and *desquamator*.

We also found evidence for shared selective sweeps across all three specialists in regions that are enriched for genes annotated for metabolic processes such as short chain fatty acid and propionate metabolism (figure S12D). The lack of fixed alleles in these regions relevant to dietary specialization suggests polygenic selection (see supplemental for more discussion). Subtle shifts of allele frequencies across the genome can lead to divergent genomic backgrounds that give populations access to different ecological niches (e.g. [48,49]).

While both shared sweeps among all specialists and shared adaptive alleles among the two scale-eating species suggest constrained adaptive walks along overlapping genotypic pathways, we still see most selective sweeps are unique to each species in this radiation (figure 4; figure S11). Curiously, some adaptive standing genetic variation rose to high frequency in *C. desquamator* but did not similarly undergo selection in *C.* sp. 'wide-mouth', despite its adaptation to a similar ecological niche and the presence of these alleles segregating at low frequency in the 'wide-mouth' population. This highlights the dual influence of epistatic interactions on adaptive walks in rugged landscapes – epistasis reduces number of available paths but increases the number of local fitness peaks populations can get stuck on [50]. Selection on the same adaptive alleles may have allowed both scale-eating species access to the same area of the fitness landscape but epistatic interactions with private mutations and introgressed variation in each lineage may have resulted in divergent paths to scale-eating, ultimately contributing to diverse evolutionary outcomes even from a shared starting point.

The use of adaptive alleles from distinct spatial sources, the distinct morphologies and divergent genomic backgrounds, and potential introgression of adaptive alleles from the more specialized scale-eater *C. desquamator* into *C.* sp. 'wide-mouth' reveals a tangled path for novel ecological transitions in nature. The complex epistatic interactions at microevolutionary scales implicated in this study make it all the more fascinating that novel ecological transitions are a common macroevolutionary feature of adaptive radiation.

Acknowledgements

We thank Priya Moorjani, Michelle St. John, Joseph McGirr, Jacquelyn Galvez, David Tian, Austin Patton, Joseph Heras, and three anonymous reviewers for helpful comments on the manuscript; the Gerace Research Centre and Troy Day for logistical support; the government of the Bahamas for permission to collect and export samples. This research 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.

Author Contributions

Conceptualization: EJR,CHM; Data Collection: EJR,CHM; Statistical analyses: EJR; Resources: CHM; Visualization: EJR,CHM; Original draft: EJR; Revising: EJR, CHM

Declaration of Interests

The authors declare no competing interests.

Ethics Statement

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. All animals were collected and exported with 2017 and 2018 permits from the Bahamas Environmental Science and Technology commission and Ministry of Agriculture.

Data Accessibility

All genomic data are archived at the National Center for Biotechnology Information BioProject Database (Accessions: PRJNA833158, PRJNA690558; PRJNA394148, PRJNA391309; and PRJNA305422). Genomic analyses, ecological and morphological data are available through dryad repository: https://doi.org/10.6078/D18H9B. Scripts are available at https://github.com/emiliejrichards/Cyprinodon_sp_wide_mouth_ecomorph.

References

1. Dieckmann U, Doebeli M. 1999 On the origin of species by sympatric speciation. *Nature* **400**, 354.
2. Polechová J, Barton NH, Gavrilefs S. 2005 Speciation thrrough competition: a critical review. *Evolution (N. Y).* **59**, 1194–1210. (doi:10.1554/04-691)
3. Bolnick DI. 2006 Multi-species outcomes in a common model of sympatric speciation. *J. Theor. Biol.* **241**, 734–744. (doi:10.1016/j.jtbi.2006.01.009)
4. Kondrashov AS, Kondrashov FA. 1999 Interactions among quantitative traits in the course of sympatric speciation. *Nature* **400**, 351–354.
5. Gavrilets S. 2014 Models of speciation: where are we now? *J. Hered.* **105**, 743–755. (doi:10.1093/jhered/esu045)
6. Martin CH, Richards EJ. 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.*
7. Arnold SJ, Pfrender ME, G. JA. 2001 The Apaptive Landscape as a conceptual bridge between mico- and macroevolution. *Genetica* **112–113**, 9–32.
8. Svennson E, Calsbeek R. 2012 *The adaptive landscape in evolutionary biology*. Oxford University Press.
9. Martin CH, Wainwright PC. 2013 Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. *Science* **339**, 208–11. (doi:10.1126/science.1227710)

10. Martin CH, Gould KJ. 2020 Surprising spatiotemporal stability of a multi-peak fitness landscape revealed by independent field experiments measuring hybrid fitness. *Evol. Lett.* , 530–544. (doi:10.1002/evl3.195)
11. Calsbeek R, Irschick DJ. 2007 The quick and the dead: Correlational selection on morphology, performance, and habitat use in island lizards. *Evolution (N. Y).* **61**, 2493–2503. (doi:10.1111/j.1558-5646.2007.00206.x)
12. Grant PR, Grant RB. 2008 *How and Why Species Multiply*. Princeton University Press.
13. Givnish TJ, Sytsma KJ, Smith J, Hahn W, DH B, Burkhardt E. 1997 Molecular evolution and adaptive radiation in *Brocchinia* (Bromeliaceae: Pitcairnioideae) atop tepuis of the Guyana Shield. In *Molecular evolution and adaptive radiation*, pp. 259–311. Cambridge: Cambridge University Press.
14. Givnish TJ, Burkhardt EL, Happel RE, Weintraub JD. 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.
15. Erwin DH. 2021 A conceptual framework of evolutionary novelty and innovation. *Biol. Rev.* **96**, 1–15. (doi:10.1111/brv.12643)
16. Blount ZD, Barrick JE, Davidson CJ, Lenski RE. 2012 Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature* **489**, 513–518. (doi:10.1038/nature11514)
17. Quandt EM, Deatherage DE, Ellington AD, Georgiou G, Barrick JE. 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. (doi:10.1073/pnas.1314561111)
18. Seehausen O. 2004 Hybridization and adaptive radiation. *Trends Ecol. Evol.* **19**, 198–207. (doi:10.1016/j.tree.2004.01.003)
19. John MES, Dixon KE, Martin CH. 2020 Oral shelling within an adaptive radiation of pupfishes : Testing the adaptive function of a novel nasal protrusion and behavioural preference. , 1–9. (doi:10.1111/jfb.14344)
20. Martin CH, Wainwright PC. 2013 A remarkable species flock of *Cyprinodon* pupfishes endemic to San Salvador Island, Bahamas. *Bull. Peabody Museum Nat. Hist.* **54**, 231–240.
21. Turner BJ, Duvernell DD, Bunt TM, Barton MG. 2008 Reproductive isolation among endemic pupfishes (Cyprinodon) on San Salvador Island, Bahamas: Microsatellite evidence. *Biol. J. Linn. Soc.* **95**, 566–582. (doi:10.1111/j.1095-8312.2008.01079.x)
22. Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, Wohlfarth B, Mitrovica JX, Hostetler SW, McCabe AM. 2009 The Last Glacial Maximum. *Science (80-.).* **325**, 710–714. (doi:10.1126/science.1172873)
23. Richards EJ, Martin CH. 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. (doi:10.1371/journal.pgen.1006919)
24. Jackson AL, Inger R, Parnell AC, Bearhop S. 2011 Comparing isotopic niche widths among and within communities: SIBER - Stable Isotope Bayesian Ellipses in R. *J. Anim. Ecol.* **80**, 595–602. (doi:10.1111/j.1365-2656.2011.01806.x)
25. Richards EJ, McGirr JA, Wang JR, St. John ME, Poelstra JW, Solano MJ, O'Connell DC, Turner BJ, Martin CH. 2021 A vertebrate adaptive radiation is assembled from an ancient and disjunct spatiotemporal landscape. *Proc. Natl. Acad. Sci. U. S. A.* **118**. (doi:10.1073/pnas.2011811118)

26. Peter BM. 2016 Admixture, population structure, and f-statistics. *Genetics* **202**, 1485–1501. (doi:10.1534/genetics.115.183913)
27. Excoffier L, Dupanloup I, Huerta-Sánchez E, Sousa VC, Foll M. 2013 Robust Demographic Inference from Genomic and SNP Data. *PLOS Genet.* **9**, e1003905. (doi:10.1371/journal.pgen.1003905)
28. Pavlidis P, Živković D, Stamatakis A, Alachiotis N. 2013 SweeD: Likelihood-based detection of selective sweeps in thousands of genomes. *Mol. Biol. Evol.* **30**, 2224–2234. (doi:10.1093/molbev/mst112)
29. Danecek P *et al.* 2011 The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158. (doi:10.1093/bioinformatics/btr330)
30. Smith J, Coop G, Stephens M, Novembre J. 2018 Estimating Time to the Common Ancestor for a Beneficial Allele. *Mol. Biol. Evol.* **35**, 1003–1017. (doi:10.1101/071241)
31. Roesti M, Moser D, Berner D. 2013 Recombination in the threespine stickleback genome - Patterns and consequences. *Mol. Ecol.* **22**, 3014–3027. (doi:10.1111/mec.12322)
32. St. John ME, Holzman R, Martin CH. 2020 Rapid adaptive evolution of scale-eating kinematics to a novel ecological niche. *J. Exp. Biol.* **223**, jeb217570. (doi:10.1242/jeb.217570)
33. Martin CH, Wainwright PC. 2013 On the measurement of ecological novelty: scale-eating pupfish are separated by 168 my from other scale-eating fishes. *PLoS One* **8**, e71164. (doi:10.1371/journal.pone.0071164)
34. Richards EJ, Martin CH. 2017 Adaptive introgression from distant Caribbean islands contributed to the diversification of a microendemic adaptive radiation of trophic specialist pupfishes. *PLOS Genet.* **13**, e1006919.
35. Dickinson ME *et al.* 2016 High-throughput discovery of novel developmental phenotypes. *Nature* **537**, 508–514. (doi:10.1038/nature19356)
36. Kida YS, Sato T, Miyasaka KY, Suto A, Ogura T. 2007 Daam1 regulates the endocytosis of EphB during the convergent extension of the zebrafish notochord. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 6708–6713. (doi:10.1073/pnas.0608946104)
37. St. John, Michelle E., Dunker JC, Richards, Emilie J., Romero S, Martin CH. 2021 Parallel genetic changes underlie integrated craniofacial traits in an adaptive radiation of trophic specialist pupfishes. , 1–46.
38. Thomas T, Voss AK, Chowdhury K, Gruss P. 2000 Querkopf, a MYST family histone acetyltransferase, is required for normal cerebral cortex development. *Development* **127**, 2537–2548.
39. Mcgirr JA, Martin CH. 2018 Parallel evolution of gene expression between trophic specialists despite divergent genotypes and morphologies. *Evol. Lett.* **2**, 62–75. (doi:10.1002/evl3.41)
40. Whittaker R. 1977 Evolution of species diversity in land communities. *Evol. Biol.*
41. Stroud JT, Losos JB. 2016 Ecological Opportunity and Adaptive Radiation. *Annu. Rev. Ecol. Evol. Syst.* **47**, 507–532. (doi:10.1146/annurev-ecolsys-121415-032254)
42. Kim BY, Huber CD, Lohmueller KE. 2018 Deleterious variation shapes the genomic landscape of introgression. *PLoS Genet.* **14**, 1–30. (doi:10.1371/journal.pgen.1007741)
43. Zhang X, Kim B, Lohmueller KE, Huerta-Sánchez E. 2020 The impact of recessive deleterious variation on signals of adaptive introgression in human populations. *Genetics* **215**, 799–812. (doi:10.1534/genetics.120.303081)
44. Martin CH, Wainwright PC. 2011 Trophic novelty is linked to exceptional rates of

morphological diversification in two adaptive radiations of cyprinodon pupfish. *Evolution (N. Y).* **65**, 2197–2212. (doi:10.1111/j.1558-5646.2011.01294.x)

45. Orr HA. 2005 The genetic theory of adaptation: A brief history. *Nat. Rev. Genet.* **6**, 119–127. (doi:10.1038/nrg1523)
46. Patton AH, Richards EJ, Gould KJ, Buie LK, Christopher H. 2021 Adaptive introgression and de novo mutations increase access to novel fitness peaks on the fitness landscape during a vertebrate adaptive radiation.
47. Weinreich DM, Delaney NF, DePristo MA, Hartl DL. 2006 Darwinian Evolution Can Follow Only Very Few Mutational Paths to Fitter Proteins. *Science (80-.).* **312**, 111 LP – 114. (doi:10.1126/science.1123539)
48. Otte KA, Nolte V, Mallard F, Schlötterer C. 2021 The genetic architecture of temperature adaptation is shaped by population ancestry and not by selection regime. *Genome Biol.* **22**, 1–25. (doi:10.1186/s13059-021-02425-9)
49. Hansen TF. 2013 Why epistasis is important for selection and adaptation. *Evolution (N. Y).* **67**, 3501–3511. (doi:10.1111/evo.12214)
50. Fragata I, Blanckaert A, Dias Louro MA, Liberles DA, Bank C. 2019 Evolution in the light of fitness landscape theory. *Trends Ecol. Evol.* **34**, 69–82. (doi:https://doi.org/10.1016/j.tree.2018.10.009)