

Surprising spatiotemporal stability of a multi-peak fitness landscape revealed by independent field experiments measuring hybrid fitness

Christopher H. Martin^{1,2,3}  and Katelyn J. Gould⁴ 

¹Department of Integrative Biology, University of California, Berkeley, Berkeley, California 94720

²Museum of Vertebrate Zoology, University of California, Berkeley, Berkeley, California 94720

³E-mail: chmartin@berkeley.edu

⁴Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27515

Received December 9, 2019

Accepted September 14, 2020

The effect of the environment on fitness in natural populations is a fundamental question in evolutionary biology. However, experimental manipulations of both environment and phenotype at the same time are rare. Thus, the relative importance of the competitive environment versus intrinsic organismal performance in shaping the location, height, and fluidity of fitness peaks and valleys remains largely unknown. Here, we experimentally tested the effect of competitor frequency on the complex fitness landscape driving adaptive radiation of a generalist and two trophic specialist pupfishes, a scale-eater and molluscivore, endemic to hypersaline lakes on San Salvador Island (SSI), Bahamas. We manipulated phenotypes, by generating 3407 F4/F5 lab-reared hybrids, and competitive environment, by altering the frequency of rare transgressive hybrids between field enclosures in two independent lake populations. We then tracked hybrid survival and growth rates across these four field enclosures for 3–11 months. In contrast to competitive speciation theory, we found no evidence that the frequency of hybrid phenotypes affected their survival. Instead, we observed a strikingly similar fitness landscape to a previous independent field experiment, each supporting multiple fitness peaks for generalist and molluscivore phenotypes and a large fitness valley isolating the divergent scale-eater phenotype. These features of the fitness landscape were stable across manipulated competitive environments, multivariate trait axes, and spatiotemporal heterogeneity. We suggest that absolute performance constraints and divergent gene regulatory networks shape macroevolutionary (interspecific) fitness landscapes in addition to microevolutionary (intraspecific) competitive dynamics. This interplay between organism and environment underlies static and dynamic features of the adaptive landscape.

KEY WORDS: Adaptive dynamics, adaptive radiation, competition, key innovation, lepidophagy, negative frequency-dependent disruptive selection, novelty, trophic specialization.

The adaptive landscape, the complex mapping of fitness onto phenotype or genotype, is both a central unifying concept in evolutionary biology and an empirical measurement (Wright 1932; Lande 1979; Gavrillets 2004; Carneiro and Hartl 2010; Svensson and Calsbeek 2012) that links the microevolutionary processes of natural and sexual selection in wild populations with macroevolutionary patterns of speciation, novelty, and adaptive radiation

(Lande and Arnold 1983; Arnold et al. 2001; Kingsolver et al. 2001; Martin and Richards 2019).

Despite its central importance, it remains unclear what factors shape the fitness landscape across space and time. In classical views arising from both Wright's (1932) and Simpson's (1944) original conceptions of genotypic and phenotypic fitness landscapes, respectively, and Fisher's (1930) geometric model, fitness

optima occur on a high-dimensional landscape with both static and dynamic features. Static features of fitness landscapes arise due to negative epistasis within genotypic networks (Whitlock et al. 1995; Weinreich et al. 2005; Wright 1932) and functional trade-offs between different ecological niches or collections of similar niches known as adaptive zones (Simpson 1944; Roughgarden 1972; Higham et al. 2016; Evans et al. 2017). Dynamic features, emphasized by the more recent paradigm of adaptive dynamics originating in game theory (Abrams et al. 1993; Abrams 2001; Bolnick 2004; Doebeli and Dieckmann 2005), arise because the fitness landscape resembles a trampoline: as the relative frequency of phenotypically similar individuals competing for the same limited resources increases, their fitness decreases due to increased competition, whereas rare phenotypes at the extremes of the phenotype distribution should have a fitness advantage due to reduced competition for additional resources that they are able to exploit (Rosenzweig 1978). This process is known as negative frequency-dependent disruptive selection and can lead to “competitive” speciation, even when adapting to a unimodal resource distribution (Rosenzweig 1978; Dieckmann and Doebeli 1999; Kirkpatrick and Ravigné 2002; Matessi et al. 2002; Doebeli et al. 2005; Bürger and Schneider 2006; Otto et al. 2008). Laboratory and field studies of natural populations provide extensive support for negative frequency-dependent disruptive selection (Pfennig 1992; Hori 1993; Sinervo et al. 2000; Schluter 2003; Bolnick 2004; Kassen et al. 2004; Olendorf et al. 2006; Bolnick and Lau 2008; Weeks and Hoffmann 2008; Koskella and Lively 2009; Svensson and Calsbeek 2012; Kusche et al. 2012; Bolnick and Stutz 2017; Nosil et al. 2018), to the extent that some investigators assert its universality in all natural populations (Haller and Hendry 2014). Due to its elegance and mathematical tractability, frequency dependence has also been widely adopted by speciation theorists as the sole mechanism for disruptive selection in most models (Gavrilets 2004; Polechová and Barton 2005; Otto et al. 2008; Servedio and Boughman 2017; Martin and Richards 2019).

However, the relative contributions of the dynamic competitive environment versus stable fitness optima in shaping the broader topography of fitness landscapes across multiple species remain unclear, despite the importance of these interactions in providing the bridge between microevolutionary processes within a population and macroevolutionary scale patterns of diversification and adaptive radiation (Arnold et al. 2001; Higham et al. 2016; Hendry 2017; Martin and Richards 2019; Martin et al. 2019; Gillespie et al. 2020). For example, although negative frequency-dependent disruptive selection may be ubiquitous within populations, the threshold of phenotypic similarity necessary for intraspecific competition within the same ecological niche is rarely measured and this competition kernel can have major impacts on speciation (Dieckmann and

Doebeli 1999; Baptestini et al. 2009). Similarly, shifts in the location of fitness optima due to environmental stochasticity are often assumed to follow a Brownian motion process (Grant and Grant 2002; Hansen et al. 2008) without accounting for hard boundaries imposed by functional trade-offs or biophysical constraints (but see Enquist and Niklas 2002; Boucher and Demery 2016). Conversely, over broader macroevolutionary timescales the role of stable fitness optima in shaping trait diversification across a radiation is frequently assumed and tested by fitting Ornstein-Uhlenbeck models of trait diversification using phylogenetic comparative methods (Butler and King 2004; Harmon et al. 2010; Beaulieu and O’Meara 2012; O’Meara 2012; Uyeda and Harmon 2014; Burrell and Tan 2017) while not accounting for the effects of density- and frequency-dependent selection due to competition within a community (but see Rabosky 2009; Harmon et al. 2019; Landis et al. 2019).

Previous field experiments testing frequency-dependent selection have focused on only a single population or pair of ecotypes (Pfennig 1992; Sinervo et al. 2000; Kingsolver et al. 2002; Schluter et al. 2003; Bolnick 2004; Olendorf et al. 2006; Bleay et al. 2007; Pfennig et al. 2007; Weeks and Hoffmann 2008; Mappes et al. 2008; Calsbeek et al. 2009; Kusche et al. 2012; Bolnick and Stutz 2017) and there are still few studies spanning multiple habitats, traits, time periods, and species. In the few studies at these larger scales, rare transgressive hybrid phenotypes appear to suffer a fitness cost, not an advantage (Keagy et al. 2015; Martin 2016a). For example, in a hybrid mesocosm experiment investigating male-male competition in stickleback, the rarest transgressive phenotypes experienced the lowest reproductive success (Keagy et al. 2015), in contrast to predictions of sexual selection as a diversifying force (Seehausen and Schluter 2004; but see Servedio and Burger 2014; Kopp et al. 2018; Puebla et al. 2012). We previously estimated multiple fitness peaks driving adaptive radiation in the San Salvador Island (SSI) pupfishes by measuring the growth and survival of lab-reared F2 intercross and backcross hybrids placed in high- and low-density field enclosures (Martin and Wainwright 2013b). Hybrid phenotypes resembling the widespread generalist species were isolated by a local fitness peak, separated by a fitness valley from a higher fitness peak corresponding to hybrid phenotypes resembling the molluscivore specialist, whereas hybrid phenotypes resembling the scale-eating specialist suffered the lowest growth and survival. Interestingly, rare hybrid phenotypes in this experiment did not experience a survival advantage as predicted by competitive speciation theory. Instead, only hybrids resembling the generalist phenotype experienced a frequency-dependent survival advantage; rare transgressive hybrids suffered the lowest survival (Martin 2016a). These few experimental studies spanning multiple species suggest that negative frequency-dependent disruptive selection may primarily operate within a population,

rather than among species occupying divergent ecological niches (Martin 2016a; Martin and Richards 2019); however, experimental tests are needed.

Here, we experimentally manipulated the frequency of rare hybrid phenotypes to investigate the static and dynamic features of fitness landscapes. We use a tractable system for empirical measurements of the fitness landscape, an adaptive radiation of pupfishes endemic to San Salvador Island, Bahamas. We first generated transgressive hybrid phenotypes by multiple rounds of backcrossing and intercrossing between a generalist and two trophic specialist *Cyprinodon* pupfish species. We then experimentally manipulated the frequency, but not the overall density, of rare and transgressive hybrid phenotypes between two semi-natural field enclosures. This manipulation was performed independently in two ecologically distinct lakes over different exposure periods and seasons using independent hybrid populations originating from each lake. To assess fitness, we individually tagged hybrids, tracked their survival and growth rates in each enclosure, and measured their phenotype from 30 external traits, resulting in estimates of the fitness landscape in four independent field enclosures, two in each lake. This experimental procedure created a normally distributed but highly variable distribution of hybrid phenotypes, recreating the early stages of diversification before adaptive radiation began on San Salvador Island, which originated predominately from Caribbean-wide standing genetic variation that predates the age of the radiation (McGirr and Martin 2017, 2020; Richards and Martin 2017; Richards et al. 2020). We found negligible effects of the frequency of rare transgressive phenotypes on the survival fitness landscape in both lakes. However, combined inferences across all four field enclosures supported multiple fitness peaks for generalists and molluscivore phenotypes and a large fitness valley isolating the novel scale-eating specialist, strikingly similar to survival patterns observed in a previous independent field experiment. These results provide strong empirical field support for stasis of major features of a complex fitness landscape, which helps explain the rare evolution of scale-eating among fishes and the origins of ecological novelty.

Methods

FREQUENCY MANIPULATION WITHIN FIELD ENCLOSURES

We individually tagged and photographed 3407 F4-F5 outbred juvenile hybrids resulting from crosses of all three species, generalist (*Cyprinodon variegatus*), molluscivore (*Cyprinodon bronthoeroides*), and scale-eater (*Cyprinodon desquamator*; Martin and Wainwright 2013a), from two different isolated lake populations on San Salvador Island, Bahamas. Tagged hybrids were then released into one of two field enclosures in each lake containing either a high- or low-frequency of rare transgressive hy-

brid phenotypes (Fig. 1; Table S1). Due to logistical constraints, hybrids could not be measured before the experiment and were sorted by eye into treatments by selectively choosing more transgressive/rare phenotypes for high-frequency enclosures and more generalist-like phenotypes for low-frequency enclosures from a common pool of hybrids from each lake population during the tagging procedure. This was performed in an alternating and haphazard manner over several days (see methods in Supporting Information) so that no treatment group was tagged in a single consecutive block. Our manual frequency manipulation procedure successfully increased the frequency of rare transgressive hybrid phenotypes in the high-frequency treatments and abundant generalist-type hybrids in the low-frequency treatments, respectively, resulting in a significant reduction in phenotypic variance on discriminant axis 2 (predominantly nasal protrusion) in lake 1 (Levene's test, $P < 0.0001$) and discriminant axis 1 (predominantly oral jaw size) in lake 2 (Levene's test, $P < 0.0001$) within the bivariate discriminant morphospace separating all three parental species. This shift in significant discriminant axes between lake populations likely reflects intraspecific differences in genetic architecture, nasal protrusion distance, and gene expression between these lake populations (Martin and Feinstein 2014; McGirr and Martin 2019, 2020). The total density of hybrids was held approximately constant between high- and low-frequency treatments (lake 1: high/low: 923/833 individuals; lake 2: high/low: 842/819 individuals; Table S1; note that only a random subset of 451 and 404 individuals were measured from the total released into lake 2 field enclosures due to low survivorship, which included all survivors plus a random subsample of deaths; see methods in Supporting Information).

SURVIVAL IN FIELD ENCLOSURES AND LABORATORY CONTROL ENVIRONMENTS

Exposure period was confounded with lake environment in this study, but separating these effects was not our aim. We measured hybrid survival after 3 months in lake 1 (high-frequency: 77.1% survival; low-frequency: 75% survival) and after 11 months in lake 2 (high-frequency: 1.4% survival; low-frequency: 1.2% survival; Table S1), in the latter case spanning most of the adult pupfish lifespan (Martin et al. 2016) but avoiding mortality due to senescence. There were no differences in survival probability between treatments in each lake (two-way logistic regression, treatment effect: $P = 0.237$). For a control comparison, additional hybrids from each lake population ($N = 199$ total individuals) were simultaneously tagged, raised in two laboratory aquaria, and their deaths and growth rates were tracked over 1 year concurrently with field experiments.

MORPHOMETRICS

Phenotypic similarity of each hybrid to the three parental species in each lake ($n = 236$ total lab-reared parental individuals

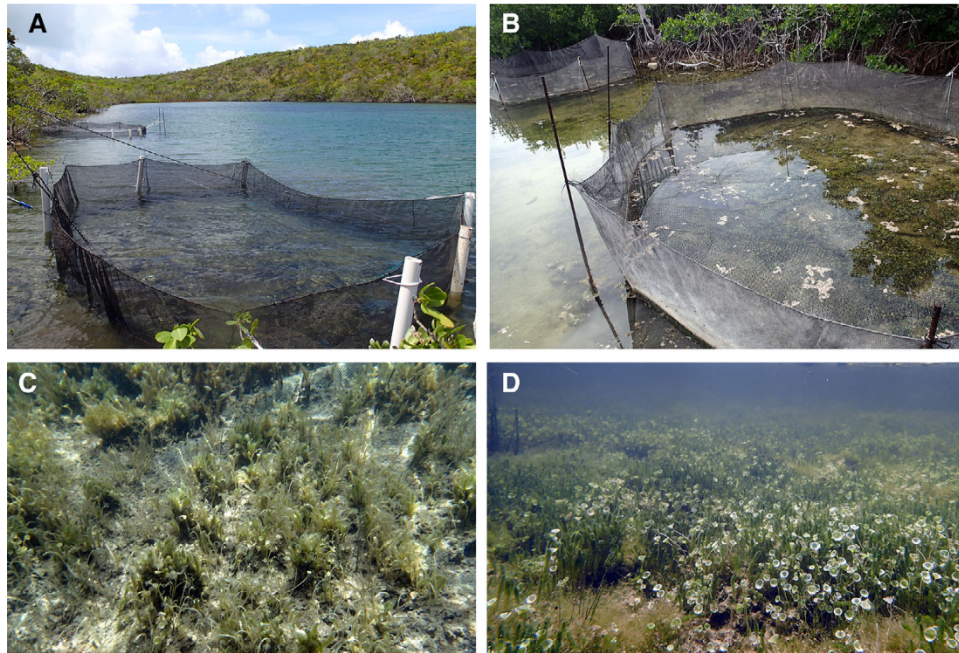


Figure 1. High- and low-frequency field enclosures and the associated benthic macroalgae communities inside each enclosure typical of surrounding littoral zone habitats in lake 1 (A and C: Crescent Pond) and lake 2 (B and D: Little Lake) after 3 month and 11 month field exposure periods, respectively. Each lake contained a high- and low-frequency field enclosure (four total) in which the frequencies of rare transgressive hybrid phenotypes versus common generalist phenotypes were altered by eye during tagging while maintaining comparable overall densities.

measured across both lakes; 33 scale-eaters, 20 molluscivores, and 80 generalists from lake 1; 48 scale-eaters, 25 molluscivores, and 30 generalists from lake 2) was calculated from 30 linear traits and angles measured from three prerelease photographs of each fish (Fig. S2). These traits were used to estimate two linear discriminant (LD) axes with major loadings of oral jaw size (LD axis 1) and nasal protrusion (LD axis 2; Table S2), diagnostic traits of each specialist species and major axes of trait diversification within this radiation (Martin and Wainwright 2011; Martin 2016b). Indeed, after correcting for standard length, residual jaw length variation within our hybrid populations exceeded the range of variation observed across allopatric *Cyprinodon* species and outgroup Cyprinodontidae species spanning over 20 million years since their most recent common ancestor (data from Martin and Wainwright 2011; Fig. 2).

SUMMARY OF STATISTICAL ANALYSES

Fitness landscapes were visualized in each enclosure by fitting thin-plate splines to the survival (binomial) data for each hybrid using generalized cross-validation, which minimizes residual prediction error of the spline surface. Splines were estimated using the Fields package (Nychka et al. 2017) in R (R Development Core Team 2016). We also used generalized projection pursuit regression to estimate the two major phenotypic axes most strongly associated with survival across the 30-trait morphospace

(Tables S3 and S4; Mitchell-Olds and Shaw 1987; Schluter and Nychka 1994; Morrissey 2014).

We formally tested for experimental treatment effects on fitness landscapes using generalized additive modeling (GAM) in the mgcv package (Wood 2017). This semiparametric modeling framework enables incorporation of spline terms into generalized linear models and tests of heterogeneity in selection surfaces across treatments, further discussed in Reynolds et al. (2016). We tested the power of this approach to detect frequency-dependent selection by simulating 1%, 10%, and 20% nonlinear survival increases on one edge of the observed low-frequency fitness landscapes (Fig. S8) and counting the number of times that a true frequency-dependent effect was detected in our GAM at an $\alpha = 0.05$ significance threshold over 1000 binomial samples of survivors and deaths from these surfaces. In each sample, we held our observed sample sizes and hybrid phenotype distributions constant.

To examine the local scale of frequency-dependence within each enclosure, we measured the frequency of similar hybrid phenotypes in local regions of morphospace (Martin 2016a). Within each enclosure, we estimated the Mahalanobis distance for each hybrid individual, the distance to the mean hybrid phenotype in the 30-dimensional morphospace correcting for trait covariances. We also calculated the sum of the Euclidean distance to the 10 nearest neighbors (i.e., the 10 most similar hybrid phenotypes) in 30-dimensional morphospace.

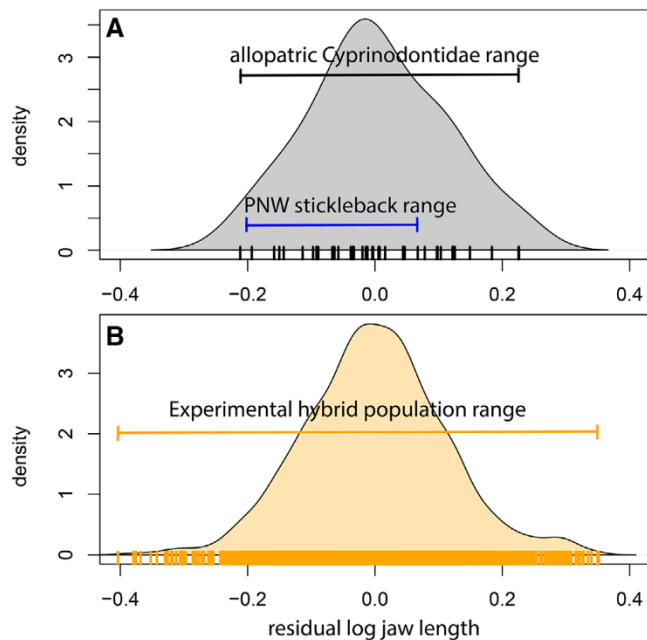


Figure 2. Residual jaw length variation is comparable between (A) allopatric trophic generalist *Cyprinodontidae* species (gray) that diverged 20 million years ago (Furness and Reznick 2015; data from Martin and Wainwright 2011) and (B) lab-reared F4/F5 hybrid populations (orange) measured from the field experiment in this study. Lab-reared hybrid variation exceeds the minimum and maximum residual upper jaw lengths across allopatric stickleback populations in the Pacific Northwest (PNW; data from fig. 1 reported in Lavin and McPhail 1986), which have frequently been the focus of previous field experiments supporting negative frequency-dependent disruptive selection (Schluter et al. 2003; Bolnick 2004; Bolnick and Stutz 2017). Each (A) species mean or (B) individual F4/F5 hybrid is represented by tick marks on the x-axis of the density plots. Residuals were calculated from a single pooled linear regression of log-transformed lower jaw length on log-transformed standard length without mean and variance standardization for comparison on the same absolute scale across all three systems.

It is possible that some regions of the high-dimensional trait space may still connect scale-eater phenotypes to other regions of the morphospace through a fitness ridge (Gavrilets 1997, 1999). To further explore the relative fitness of scale-eater hybrids, we visualized selection across all directions in the 30-trait morphospace by repeatedly sampling a random subset of 15 traits, calculating a discriminant axis for scale-eaters relative to generalists within this subspace, and estimating a survival spline for hybrid phenotypes on each arbitrary multivariate axis (Fig. 6). Although we did not measure all traits potentially affecting fitness, this procedure creates a random set of discriminant morphospaces to visualize all fitness pathways across our set of thirty morphological traits. Importantly, this procedure aligns these multivariate linear axes in the same direction from generalist to scale-eater phenotype enabling com-

parison of survival curves across random subsets of trait data. Further details on rearing conditions, field enclosures, morphometrics, and statistical analyses are provided in the Supporting Information.

Results

STRONG DIRECTIONAL AND NONLINEAR SELECTION FOR GENERALIST AND MOLLUSCIVORE PHENOTYPES

Selection surfaces for survival in lake 1 were flat and generally favored hybrid phenotypes most similar to generalists in the high-frequency treatment or generalist and scale-eater phenotypes in the low-frequency treatment in lake 1, but this relationship was not significant (binomial GAM with effect of discriminant morphospace spline on survival: $P > 0.05$; Fig. 3). In contrast, we found significant evidence of nonlinear divergent selection for hybrid phenotypes resembling the molluscivore in both high-frequency (binomial GAM with effect of discriminant morphospace thin-plate spline on survival: $P = 0.045$) and low-frequency ($P = 0.0002$) treatments in lake 2, despite lower survival rates over the longer 11 month exposure period in this lake (Fig. 3). Fitness landscapes in lake 2 were estimated from few survivors ($n = 22$); however, random permutation of survival in lake 2 showed that the probability of observing nonlinear selection equal to or more extreme than we observed (GAM $\chi^2 \geq 7.19$) due to chance was $P = 0.014$ in the high-frequency enclosure and $P < 0.000$ ($\chi^2 \geq 21.36$) in the low-frequency enclosure. Patterns of survival in the wild were contrasted by strong directional selection for hybrids resembling the scale-eater in laboratory control populations from both lakes (poisson GAM with effect of discriminant morphospace thin-plate spline on days survived in lab aquaria: $P < 0.00001$; Fig. 3).

We also visualized and analyzed selection on the survival fitness landscapes across all 30 traits in our dataset by estimating the two strongest axes of nonlinear selection using projection pursuit regression. The most transgressive hybrid phenotypes (i.e., least similar to parental phenotypes) suffered the lowest survival probability across treatments in both lakes along these axes (Fig. 4; Table S5: $P < 0.00001$). In contrast, survival in laboratory control populations was shifted away from the direction of selection in field enclosures (Fig. 4). We found significant evidence of directional selection on the strongest nonlinear axis in both lakes ($P < 0.00001$) and marginal evidence of directional selection on the second strongest axis only in lake 2 (Table S5). The traits with the highest loadings on axis 1 were (a) lower jaw length and (b) distance from the jaw joint to the orbit. The highest loadings on axis 2 were (a) angle between the premaxilla and orbit and (b) distance from the premaxilla to the pectoral girdle (Tables S3 and S4), further supporting strong selection on craniofacial trait diversification within this radiation.

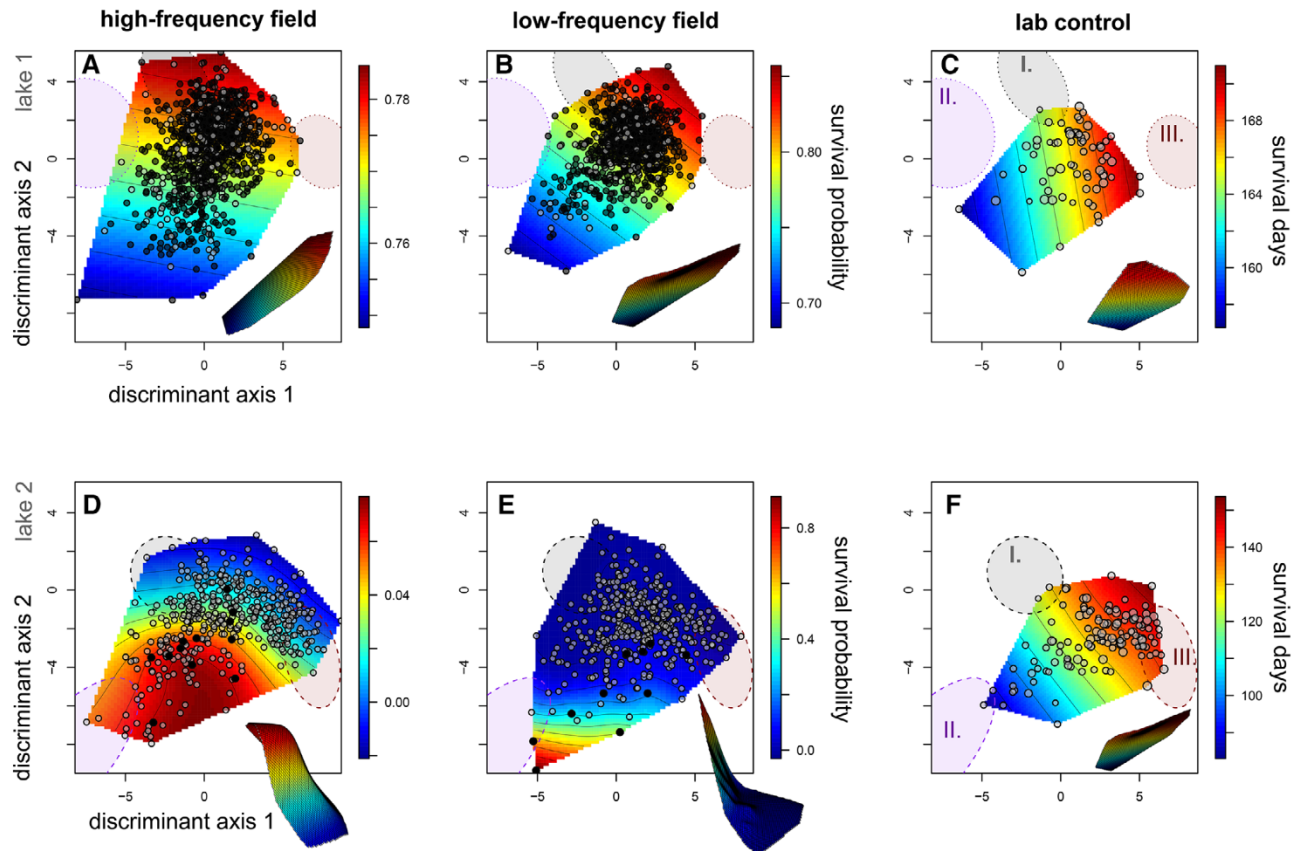


Figure 3. Survival fitness landscapes for hybrid populations in high- (first column) and low-frequency (second column) field enclosures and laboratory controls (third column). Thin-plate splines predict the probability of survival (heat color) across a single linear discriminant morphospace separating generalist and scale-eater phenotypes (x -axis: LD1) and generalist and molluscivore phenotypes (y -axis: LD2). Survivors in field enclosures are depicted in black relative to deaths over the 3-month and 11-month exposure periods, respectively. Laboratory control points are proportional to the number of days each hybrid survived within 151-L aquaria. All hybrids are plotted within a shared linear discriminant morphospace for parental species' phenotypes calculated from lab-reared F1 individuals of all three parental populations from both lakes (first row: lake 1; second row: lake 2); 95% confidence ellipses for each parental population in each lake are shown for generalists (I. gray), molluscivores (II. purple), and scale-eaters (III. red); note that molluscivores in lakes 1 and 2 show large differences in mean nasal protrusion distance (Martin and Feinstein 2014) leading to their separation along discriminant axis 2. Survival surfaces are depicted in three dimensions in the offset images.

NO EVIDENCE OF FREQUENCY-DEPENDENT SURVIVAL DIFFERENCES BETWEEN TREATMENTS

We found no evidence that manipulating the frequency of rare transgressive hybrid phenotypes between field enclosures in each lake affected either survival probability or the overall topography of the survival fitness landscapes within the discriminant morphospace (Table S6). Adding a fixed effect for this frequency treatment did not improve the fit to the survival data in any of the generalized additive models examined. Instead, models without the effect of frequency treatment were strongly favored (Table S6; $\Delta\text{AIC} = 11$). The power of our GAM modeling framework to detect a significant effect of frequency manipulation on the fitness landscape was 97% for an absolute survival increase of 20% in one region of the low-frequency field enclosures, 64% for

a 10% survival increase, and 16% for a 1% increase in survival (calculated from 1000 binomial sampling draws from simulated fitness surfaces depicted in Fig. S8). Twenty percent survival differences are still relatively minor compared to the 80% survival difference observed in lake 2 low-frequency enclosure (Fig. 3) and these power simulations were robust among the best generalized additive models compared in Table S6. Thus, given the steep clines in survival observed, we conclude that there was a minimal effect of the frequency manipulation on the survival fitness landscapes measured.

Similarly, we found no evidence for treatment effects on survival along the two major axes of nonlinear selection in either lake estimated from generalized projection pursuit regression (Table S7). Models without the effect of treatment provided

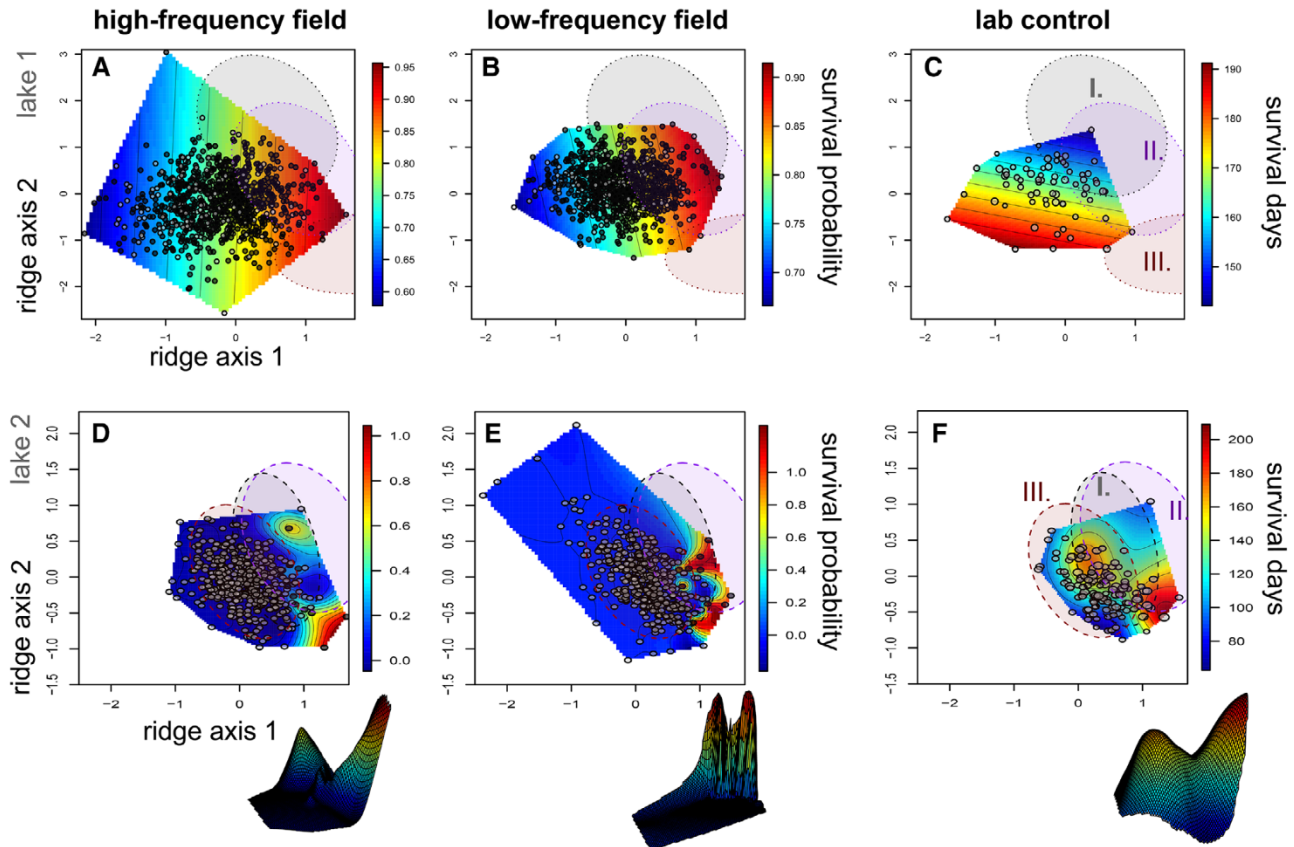


Figure 4. Survival fitness landscapes for hybrid populations in high- (first column) and low-frequency (second column) field enclosures and laboratory controls (third column) across the two major ridge axes of nonlinear selection within the 30-trait morphospace. Thin-plate splines predict the probability of survival (heat color) across the two major ridge axes associated with survival (Tables S3 and S4), estimated separately for lake 1 (first row) and lake 2 (second row) hybrid populations using generalized projection pursuit regression (Tables S3 and S4). Survivors in field enclosures are depicted in black relative to deaths over the 3-month (first row) and 11-month (second row) exposure periods, respectively (Table S1). Laboratory control points (third column) are proportional to the number of days each hybrid survived within 151-L aquaria; 95% confidence ellipses for each parental population in each lake are shown for generalists (I. gray), molluscivores (II. purple), and scale-eaters (III. red). Nonlinear survival surfaces are depicted in three dimensions in the offset images.

a comparable fit to the survival data on the two major axes of nonlinear selection in lake 1 ($\Delta\text{AIC} = 1$) and were marginally supported in lake 2 (Table S7; $\Delta\text{AIC} = 1.8$).

LOCAL SCALE OF FREQUENCY DEPENDENCE FOR GROWTH RATE BUT NOT SURVIVAL

Hybrids may only compete with other hybrid phenotypes most similar to their own. The frequency of hybrids with similar phenotypes in neighboring regions of the discriminant morphospace was not significantly associated with residual variation in survival (Fig. S4; Table S6). Generalized additive models including the fixed effect of rare hybrid frequency (distance to mean phenotype) did not meaningfully improve the fit to the survival data (Table S6; $\Delta\text{AIC} = 1$; similar results were found for nearest-neighbor Euclidean distance: Fig. S5).

In contrast, generalized additive models including the fixed effect of rare hybrid frequency (both Mahalanobis and nearest-

neighbor distance) strongly improved the fit to the growth data in lake 1 within the discriminant morphospace (Table S6; Fig. S6; $\Delta\text{AIC} = 12$; lake 2 was excluded from all growth rate analyses due to the low number of survivors). Similarly, including the fixed effect of rare hybrid frequency substantially improved the fit to the growth rate data along the two major axes of nonlinear selection in lake 1 (Table S7; $\Delta\text{AIC} = 3.85$).

JOINT INFERENCE OF A COMBINED FITNESS LANDSCAPE SUPPORTS STRIKING SIMILARITY TO A PREVIOUS INDEPENDENT FIELD EXPERIMENT

GAM enables estimation and visualization of a joint fitness landscape while controlling for the effects of lake/exposure period and frequency treatment in our experiment. Strikingly, the best combined model for survival (across all four treatments in both lakes: Table S2) supported an isolated fitness peak for hybrids resembling the generalist separated by a small fitness valley from

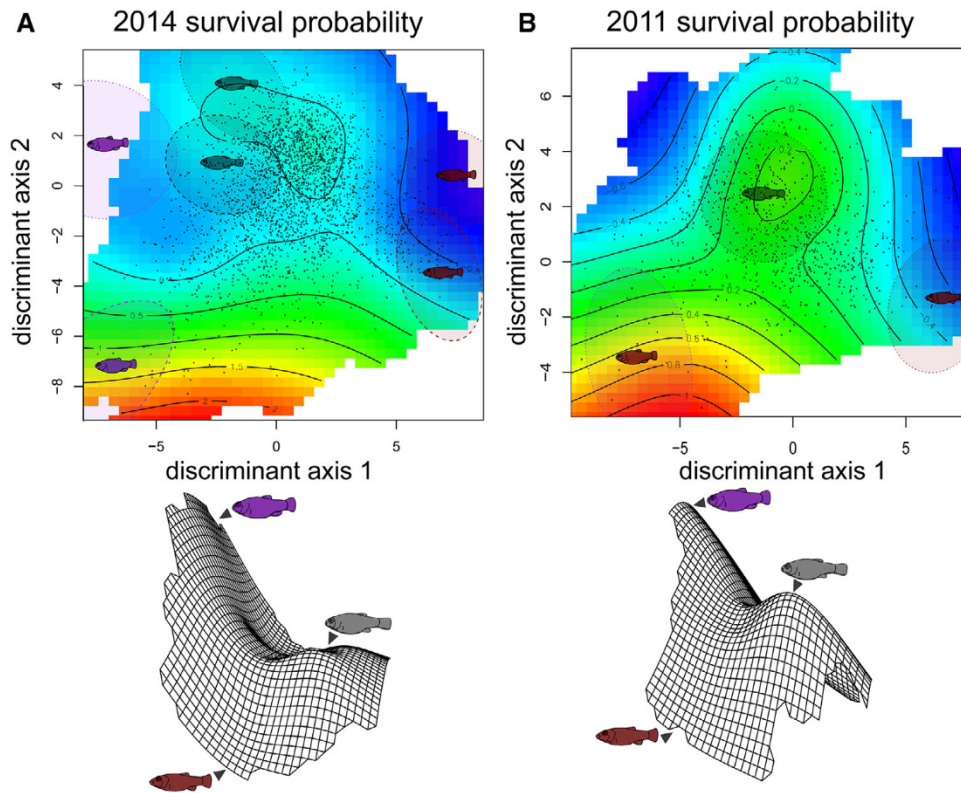


Figure 5. Survival fitness landscapes compared between independent field experiments in (A) 2014 (this study) and (B) 2011 estimated from thin-plate splines using generalized additive models. Thin-plate splines in two and three dimensions estimate hybrid survival probability (A) integrating across lake/field exposure period and high-/low-frequency treatment for all four enclosures in the 2014–2015 field experiment using a generalized additive model ($\text{survival} \sim \text{tps}(\text{LD1}, \text{LD2}) + \text{lake}$; Table S6); and (B) in only the high-density field enclosure in lake 1 from the 2011 field experiment (Martin and Wainwright 2013b), calculated from the original data using a generalized additive model ($\text{survival} \sim \text{tps}(\text{LD1}, \text{LD2})$). The discriminant morphospace separating generalist and scale-eater phenotypes (x -axis: LD1) and generalist and molluscivore phenotypes (y -axis: LD2) is calculated from lab-reared individuals of all three parental populations in both lakes in panel A and lake 1 only in panel B; 95% confidence ellipses show the location of generalist (gray), molluscivore (purple), and scale-eater (red) parental populations from lake 1 (small-dashed line) and lake 2 (large-dashed line). Discriminant morphospaces were estimated from independent sets of phenotypic measurements in each panel. The 2014–2015 morphospace was estimated from 30 external traits measured for six parental populations, three from each lake; the 2011 morphospace was estimated independently from a set of 16 external traits (partially overlapping with the 30-trait set) independently measured for three out of six parental populations (all three lake 1 populations).

a region of higher fitness corresponding to the molluscivore phenotype and a larger region of low fitness corresponding to the scale-eater phenotype (Fig. 5), similar to a previous independent field experiment in 2011 using F2 hybrids in these same lakes (Martin and Wainwright 2013b). However, in the previous field experiment few of the F2 hybrids fell within the phenotypic range of lab-reared scale-eaters (Martin and Wainwright 2013b). In our current experiment, over 70 hybrids occurred within the 95% confidence ellipse of lab-reared F1 scale-eater phenotypes in the discriminant morphospace. This provides evidence for environment-independent features of a complex fitness landscape across years, seasons, divergent lake environments, and experimentally manipulated frequencies of rare transgressive hybrid phenotypes.

Overall, the best supported models for survival included a fixed effect of lake, no treatment effect, and multiple univariate or bivariate thin-plate splines for the effect of selection (Table S2). Models including spline terms were strongly supported over models including fixed linear effects for the phenotypic discriminant axes (Table S2; $\Delta\text{AIC} = 18.7$).

VISUALIZATION OF HIGH-DIMENSIONAL FITNESS PATHWAYS

In three out of four field enclosures, hybrids resembling scale-eaters suffered the lowest survival across the majority of fitness pathways subsampled from the 30-trait morphospace (Fig. 6). This suggests that most fitness pathways, independent of the subsample of traits measured, lead to a fitness minimum isolating the

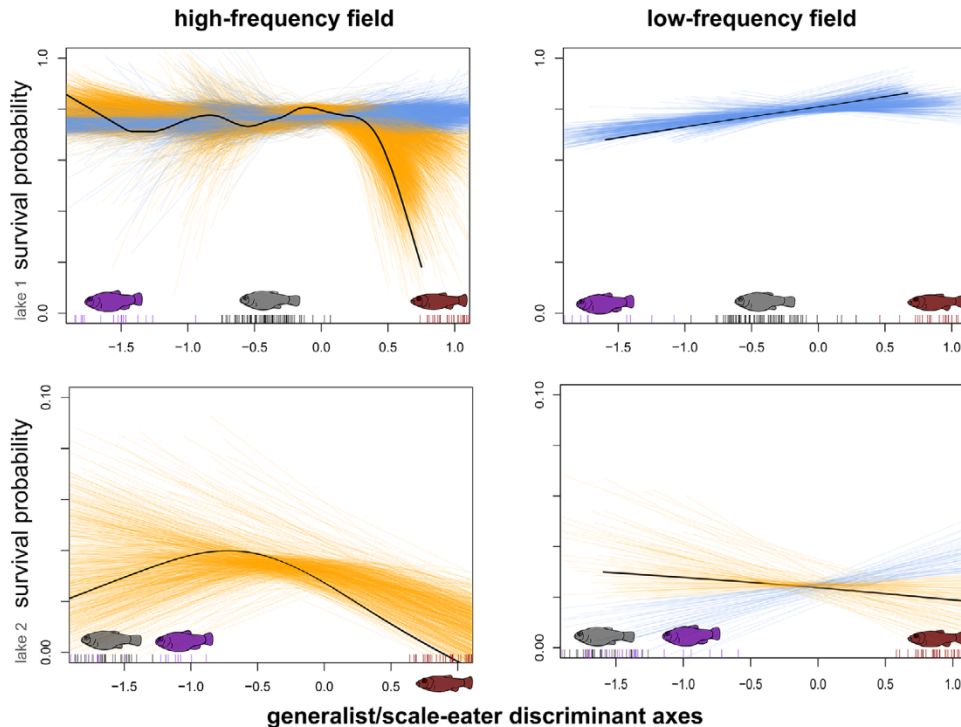


Figure 6. Spaghetti plots illustrate all possible fitness paths between generalist and scale-eater hybrid phenotypes in high- (first column) and low-frequency (second column) field enclosures. Each orange line depicts the relationship between survival and a random discriminant axis separating generalist and scale-eater phenotypes estimated from generalized cross-validation of a smoothing spline for 500 random subsets of 15 size-corrected traits (out of 30); the black line illustrates the smoothing spline estimated for a discriminant axis from all 30 traits, the orange lines illustrate the smoothing spline estimated for each subset. Each subsampled discriminant axis was rescaled so that the mean parental scale-eater phenotype = 1 (red arrows). Parental phenotypes are illustrated as black (generalist), purple (molluscivore), and red (scale-eater) tick marks on the x-axis.

scale-eater from all other phenotypes. There was no clear signal for dominant trait loadings underlying fitness pathways leading to higher scale-eater fitness (blue) versus lower scale-eater fitness (orange) when examining major functional traits such as jaw length, nasal protrusion, adductor height, or head height (Table S8).

Discussion

NO EVIDENCE OF RARE PHENOTYPE SURVIVAL ADVANTAGE ON A COMPLEX FITNESS LANDSCAPE

We conducted an experimental field test of frequency-dependent selection on rare hybrid phenotypes in an adaptive radiation of trophic specialist pupfishes endemic to San Salvador Island, Bahamas. We found no evidence of differences in survival or fitness landscape topography between treatments manipulating the frequency of rare transgressive hybrid phenotypes in two independent lake populations (Figs. 3, 4, and S8; Tables S6–S8; 97% power to detect 20% nonlinear survival differences). We also found no relationship between hybrid survival and the frequency of similar hybrid phenotypes in neighboring regions of

morphospace (Figs. S4 and S5; Tables S6 and S7). These patterns were consistent across two important cross-sections of the 30-trait morphospace: the two discriminant axes separating the three parental species (Fig. 3; Table S2) and the two major axes of nonlinear selection estimated from generalized projection pursuit regression (Fig. 4; Tables S3–S4). The lack of any signal of increased survival for rare hybrids suggests that consistent peaks and valleys on the survival fitness landscape are unaffected by the frequency of similar phenotypes and reflect intrinsic viability or organismal performance constraints rather than competitive dynamics. However, it is important to note that growth rate was affected by the frequency of competitors in lake 1 field enclosures (Figs. S4, S6, and S7) and we cannot rule out that additional fitness proxies, such as lifetime reproductive success, may also be affected by competitor frequency. Nonetheless, the stark contrast in frequency dependence between these key life history components, survival and growth, is intriguing and suggests that different selective agents control fitness dynamics and stasis across different life history stages (also see Shaw et al. 2008).

Combined estimates of the survival fitness landscape indicated a surprisingly consistent multipeak topography across two

independent field experiments (Fig. 5: 2014 and 2011). Evidence for multiple peaks across lakes resulted from integrating the higher survival of generalist and molluscivore phenotypes after 3 months in lake 1 with the much higher survival of molluscivore phenotypes after 11 months in lake 2 (Fig. 3). Growth rate in lake 1 also showed evidence of multiple peaks due to higher growth rates of generalist phenotypes in both enclosures combined with moderate molluscivore growth rates and low scale-eater growth rates in the low-frequency enclosure (Fig. S7). These joint landscapes combine different frequency treatments, exposure periods, lake environments, and potentially different selective regimes. However, we are most interested in estimating the fitness landscape driving adaptive radiation across lake environments, rather than lake-specific local adaptive regimes unrelated to species divergence. Furthermore, generalized additive models provided no evidence of different selective regimes between treatments or different fitness landscape topography between lakes (Table S6: very low support for models including a lake by thin-plate spline interaction term), supporting our inference of a single combined selective environment across lakes.

Across all four field enclosures and both fitness proxies, the most consistent feature of fitness landscape topography was a large fitness valley isolating hybrids resembling the scale-eater. The fitness landscape isolated this novel trophic specialist across spatiotemporal contexts: multiple treatments, lake environments, field exposure periods, and nearly all dimensions of the 30-trait hybrid morphospace (Figs. 3–6). This makes sense because this species is the most morphologically, ecologically, behaviorally, and genetically divergent specialist in the radiation (Martin and Wainwright 2013a; Lencer et al. 2016; McGirr and Martin 2018, 2020; Martin et al. 2019; St. John et al. 2019, 2020). This persistent multipeak fitness landscape within SSI's hypersaline lakes provides a striking explanation for the rarity of trophic specialization across the Caribbean because it suggests that generalist populations on neighboring islands are isolated on a local fitness optimum unable to colonize higher neighboring fitness peaks due to stabilizing selection opposing trait diversification (Martin and Wainwright 2013b). This is particularly intriguing given the availability of adaptive alleles for trophic specialization segregating in generalist populations on these neighboring islands (Richards et al. 2020).

EXTRINSIC AND INTRINSIC FACTORS COULD EXPLAIN LOW SURVIVAL OF TRANSGRESSIVE HYBRID PHENOTYPES

In contrast to predictions of negative frequency-dependent disruptive selection (Rosenzweig 1978; Gigord et al. 2001; Bolnick 2004; Haller and Hendry 2014), rare transgressive hybrid phenotypes suffered the lowest survival (Figs. 3 and 4). One explanation is that intrinsic viability and organismal performance were

impaired due to hybrid ancestry. Hybridization could result in mismatched craniofacial and behavioral traits leading to poor foraging performance (Parnell et al. 2012; Selz et al. 2014; Barreto et al. 2015). For example, oral jaw length variation of trophic specialist pupfish is not controlled by major effect quantitative trait loci (Martin et al. 2017), suggesting that the largest scale-eater jaws were not fully recovered in any of our F4/F5 hybrids. Second, F1 hybrid scale-eaters exhibit scale-biting strike kinematics typical of generalists, suggesting that hybrid scale-eating performance may be impaired (St. John, Holzman and Martin 2020). Impaired foraging performance of transgressive hybrids in our field enclosures is further supported because scale-eater hybrids fed only pellet foods exhibited the highest survival rates in lab aquaria (Figs. 3 and 4), demonstrating that the field environment contributed to the low survival of scale-eater hybrids. This is consistent with other laboratory studies of hybrid feeding performance (Parnell et al. 2008; Arnegard et al. 2014; Matthews and Albertson 2017; Selz and Seehausen 2019).

Intrinsic genetic incompatibilities in hybrids may also be contributing to their low survival rates, particularly if more transgressive hybrid phenotypes are associated with a greater number or more severe genetic incompatibilities (McGirr and Martin 2019, 2020). Genetic incompatibility loci are known to segregate in wild populations (Fishman et al. 2013; reviewed in Cutter 2012; Coughlan and Matute 2020) and hundreds of genetic incompatibility loci have been found in hybrid zones between swordtail fish species (Schumer et al. 2014; Schumer and Brandvain 2016). In support of this hypothesis, F1 hybrids of SSI specialist species misexpressed 9% of genes in larval tissues, much greater than expected given their sympatry and equivalent in magnitude to crosses between generalist pupfish populations separated by 1000 km (McGirr and Martin 2019, 2020). Although the fitness effects on hybrids are unknown, hybrid gene misexpression has been shown to affect hybrid viability and sterility in other systems (Ortiz-Barrientos et al. 2007; Renaut et al. 2009; Renaut and Bernatchez 2011; Mack et al. 2016; Mack and Nachman 2017) and misexpressed genes in SSI hybrids were enriched in networks affecting craniofacial morphology, muscle mass, and nitrogen metabolism (McGirr and Martin 2020). However, the link between transgressive hybrid phenotypes, divergent ecological selection, and the amount of gene misexpression and misregulation is still unknown (Kulmuni and Westram 2017).

BIOMECHANICAL AND PHYSIOLOGICAL CONSTRAINTS OF SCALE-EATING UNDERLIE THE LARGE AND STABLE FITNESS VALLEY ISOLATING THIS ECOLOGICAL NICHE

The rare evolution, high performance demands, and low caloric payoffs of scale-eating suggest that a wide and deep fitness valley may isolate this niche from all other ecological niches.

First, scale-eating (lepidophagy) is a particularly rare trophic niche among fishes and has evolved independently only 19 times across diverse marine, coastal, riverine, and lacustrine environments (Sazima 1983; Martin and Wainwright 2013c; Kolmann et al. 2018; St John et al. 2019; Grubh and Winemiller 2004; Janovetz 2005; Koblmüller et al. 2007; Raffini and Meyer 2018; Hori 1993). Scale-eating only evolved once in Cyprinodontiformes and is separated by 168 million years of evolutionary time from the most closely related African cichlid scale-eating specialists, providing a quantitative measure of the ecological novelty of this niche (Martin and Wainwright 2013c). Second, scale-eating pupfish perform frequent, rapid strikes resulting in only a few scales and mouthful of skin mucus (St. John and Martin 2019). Thus, the energetic payoff from scale-eating is likely low relative to performance demands, consistent with the observation that all scale-eaters are size limited relative to their prey, unlike piscivorous fishes that generally grow much larger than their prey (Sazima et al. 1983).

Conclusion: On the Origins of Novelty During Adaptive Radiation on Multiple Fitness Peaks

The sensitivity of fitness landscape topography to the environment is rarely measured beyond a single population or fitness peak. Here, we experimentally tested whether manipulating the frequency of rare hybrids phenotypes altered a complex fitness surface driving a diverse radiation of trophic specialist pupfishes, comparable in phenotypic diversity to outgroup species spanning 20 million years. Hybrid survival showed no signal of frequency dependence, in contrast to predictions of competitive speciation theory and previous experiments conducted within a single population or species pair. Furthermore, multiple fitness peaks and a large fitness valley isolating the scale-eating phenotype were surprisingly stable features shared across different lake environments, competitor frequencies, and two independent field experiments. This challenges our existing view of empirical fitness surfaces as highly sensitive to environmental perturbation (e.g., Grant and Grant 2002). Instead, a multi-peak fitness landscape spanning macroevolutionary levels of phenotypic disparity displays both dynamic and surprisingly static features across space, time, and competitive environment. These empirical results strengthen the connection between microevolutionary dynamic processes and macroevolutionary stasis by investigating the complex interplay between organismal phenotype and environment within a rapidly diversifying radiation.

ACKNOWLEDGMENTS

We thank the Miller Institute for Basic Research in Science, the University of California, Berkeley, the University of North Carolina at Chapel

Hill, National Science Foundation 1749764 CAREER award and National Institutes of Health 5R01DE027052-02 for funding to CHM. The Bahamas Environmental Science and Technology Commission and the Ministry of Agriculture kindly provided permission to export, import, and tag fish and conduct this research. Erica Bree Rosenblum generously provided laboratory animal space at the University of California, Berkeley. R. Hanna, V. Knowles, and the Gerace Research Centre provided logistical assistance; C. Bocklage, A. Payne, C. Lim, Q. Mei, I. Piedad, S. Bencuya, S. Jeselson, C. Farge, K. Dixon, and D. Richard assisted with morphological measurements; and E. B. Rosenblum, C. Miller, E. Richards, M. St. John, and J. McGirr provided helpful comments on the project. All animal care protocols were approved by the University of California, Berkeley and the University of North Carolina at Chapel Hill Animal Care and Use Committees.

DATA ARCHIVING

All data used for this study will be deposited in the Dryad Digital Repository upon acceptance of the manuscript. <https://doi.org/10.5061/dryad.95x69p8hz>.

ETHICS STATEMENT

The Bahamas Environmental Science and Technology Commission and the Ministry of Agriculture kindly provided permission to export, import, and tag fish and conduct this research (research permits renewed annually from 2011 to 2018). All animal care protocols were approved by the University of California, Berkeley and the University of North Carolina at Chapel Hill Animal Care and Use Committees.

AUTHOR CONTRIBUTIONS

CHM designed the study, performed field experiments, analyzed the data, wrote and edited the manuscript, and funded the study. KG contributed substantially to laboratory data collection.

LITERATURE CITED

- Abrams, P. 2001. Modelling the adaptive dynamics of traits involved in inter- and intraspecific interactions: an assessment of three methods. *Ecol. Lett.* 4:166–175.
- Abrams, P., H. Matsuda, and Y. Harada. 1993. Evolutionarily unstable fitness maxima and stable fitness minima of continuous traits. *Evol. Ecol.* 7:465–487.
- Arnegard, M. E., M. D. McGee, B. Matthews, K. B. Marchinko, G. L. Conte, S. Kabir, et al. 2014. Genetics of ecological divergence during speciation. *Nature* 511:307–311.
- Arnold, S. J., M. E. Pfrender, and A. G. Jones. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* 112–113:9–32.
- Baptistini, E. M., M. A. M. de Aguiar, D. I. Bolnick, and M. S. Araújo. 2009. The shape of the competition and carrying capacity kernels affects the likelihood of disruptive selection. *J. Theor. Biol.* 259:5–11.
- Barreto, F. S., R. J. Pereira, and R. S. Burton. 2015. Hybrid dysfunction and physiological compensation in gene expression. *Mol. Biol. Evol.* 32:613–622.
- Beaulieu, J. and B. O'Meara. 2012. OUwie: an analysis of evolutionary rates in an OU framework.
- Bleay, C., T. Comendant, and B. Sinervo. 2007. An experimental test of frequency-dependent selection on male mating strategy in the field. *Proc. R. Soc. B Biol. Sci.* 274:2019–2025.
- Bolnick, D. I. 2004. Can Intraspecific competition drive disruptive selection? an Experimental test in natural populations of sticklebacks can

- intraspecific competition drive disruptive selection? An experimental test in natural populations of sticklebacks. *Evolution* 58:608–618.
- Bolnick, D. I. and O. L. Lau. 2008. Predictable patterns of disruptive selection in stickleback in postglacial lakes. *Am. Nat.* 172:1–11.
- Bolnick, D. I. and W. E. Stutz. 2017. Frequency dependence limits divergent evolution by favouring rare immigrants over residents. *Nature* 546:285–288.
- Boucher, F. and V. Demery. 2016. Inferring bounded evolution in phenotypic characters from phylogenetic comparative data. *Syst. Biol.* 65:651–661.
- Bürger, R. and K. A. Schneider. 2006. Intraspecific competitive divergence and convergence under assortative mating. *Am. Nat.* 167:190–205.
- Burruss, E. D. and M. Tan. 2017. Ecological opportunity alters the timing and shape of adaptive radiation. *Evolution* 71:2650–2660.
- Butler, M. A. and A. A. King. 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *Am. Nat.* 164:683–695.
- Calsbeek, R., L. Bonvini, and R. Box. 2009. Geographic variation, frequency-dependent selection, and the maintenance of a female-limited polymorphism. *Evolution* 64:116–125.
- Carneiro, M. and D. L. Hartl. 2010. Adaptive landscapes and protein evolution. *Proc. Natl. Acad. Sci.* 107:1747–1751.
- Coughlan, J. M. and D. R. Matute. 2020. The importance of intrinsic postzygotic barriers throughout the speciation process. *Philos. Trans. R. Soc. B.* 375:20190533.
- Cutter, A. D. 2012. The polymorphic prelude to Bateson-Dobzhansky-Muller incompatibilities. *Trends Ecol. Evol.* 27:209–218.
- Dieckmann, U. and M. Doebeli. 1999. On the origin of species by sympatric speciation. *Nature* 400:354–357.
- Doebeli, M. and U. Dieckmann. 2005. Adaptive dynamics as a mathematical tool for studying the ecology of speciation processes. *J. Evol. Biol.* 18:1194–200.
- Doebeli, M., U. Dieckmann, J. A. Metz, and D. Tautz. 2005. What we have also learned: adaptive speciation is theoretically plausible. *Evolution* 59:691–699.
- Enquist, B. and K. Niklas. 2002. Global allocation rules for patterns of biomass partitioning in seed plants. *Science* 22:1571–1520.
- Evans, K., W. Crampton, and J. Albert. 2017. Taxonomic revision of the deep channel electric fish genus *Sternarchella* (Teleostei: Gymnotiformes: Apteronotidae), with descriptions of two new species. *Neotrop. Ichthyol.* <https://doi.org/10.1590/1982-0224-20160168>
- Fisher, R. 1930. *Genetical theory of natural selection*. Clarendon Press, Oxford, U.K.
- Fishman, L., A. Stathos, P. M. Beardsley, C. F. Williams, and J. P. Hill. 2013. Chromosomal rearrangements and the genetics of reproductive barriers in *Mimulus* (monkey flowers). *Evolution* 67:2547–2560.
- Furness, A. and D. Reznick. 2015. Convergent evolution of alternative developmental trajectories associated with diapause in African and South American killifish. *Proc. R. Soc. Lond. Ser. B.* 282:20142189.
- Gavrilets, S. 1997. Evolution and speciation on holey adaptive landscapes. *Trends Ecol. Evol.* 12:307–312.
- . 1999. A dynamical theory of speciation on holey adaptive landscapes. *Am. Nat.* 154:1–22.
- . 2004. *Fitness landscapes and the origin of species*. Princeton Univ. Press, Princeton, NJ.
- Gigord, L. D., M. R. Macnair, and A. Smithson. 2001. Negative frequency-dependent selection maintains a dramatic flower color polymorphism in the rewardless orchid *Dactylorhiza sambucina* (L.) Soo. *Proc. Natl. Acad. Sci. USA* 98:6253–6255.
- Gillespie, R. G., G. M. Bennett, L. De Meester, J. L. Feder, R. C. Fleischer, L. J. Harmon, et al. 2020. Comparing adaptive radiations across space, time, and taxa. *J. Hered.* 111:1–20.
- Grant, P. R. and B. R. Grant. 2002. Unpredictable evolution in a 30-year study of Darwin's finches. *Science* 296:707–712.
- Grubh, A. R. and K. O. Winemiller. 2004. Ontogeny of scale feeding in the Asian glassfish, *Chanda nama* (Ambassidae). *Copeia* 2004:903–907.
- Haller, B. C. and A. P. Hendry. 2014. Solving the paradox of stasis: squashed stabilizing selection and the limits of detection. *Evolution* 68:483–500.
- Hansen, T. F., J. Pienaar, and S. H. Orzack. 2008. A comparative method for studying adaptation to a randomly evolving environment. *Evolution* 62:1965–1977.
- Harmon, L. J., J. B. Losos, T. J. Davies, R. G. Gillespie, J. L. Gittleman, W. B. Jennings, et al. 2010. Early bursts of body size and shape evolution are rare in comparative data. *Evolution* 64:2385–2396.
- Harmon, L., C. Andreatzi, F. Debarre, J. Drury, E. Goldberg, and A. Martins. 2019. Detecting the macroevolutionary signal of species interactions. *J. Evol. Biol.* 32:769–782.
- Hendry, A. 2017. *Eco-evolutionary dynamics*. Princeton Univ. Press, Princeton, NJ.
- Higham, T. E., S. M. Rogers, R. B. Langerhans, H. A. Jamniczky, G. V. Lauder, W. J. Stewart, et al. 2016. Speciation through the lens of biomechanics: locomotion, prey capture and reproductive isolation. *Proc. R. Soc. B Biol. Sci.* 283:20161294.
- Hori, M. 1993. Frequency-dependent natural-selection in the handedness of scale-eating Cichlid Fish. *Science* 260, 216–219.
- Janovetz, J. 2005. Functional morphology of feeding in the scale-eating specialist *Catoprin mento*. *J. Exp. Biol.* 208:4757–4768.
- Kassen, R., M. Llewellyn, and P. B. Rainey. 2004. Ecological constraints on diversification in a model adaptive radiation. *Nature* 431:984–988.
- Keagy, J., L. Lettieri, and J. W. Boughman. 2015. Male competition fitness landscapes predict both forward and reverse speciation. *Ecol. Lett.* 19:71–80.
- Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, et al. 2001. The strength of phenotypic selection in natural populations. *Am. Nat.* 157:245–261.
- Kingsolver, J. G., D. W. Pfennig, and M. R. Servedio. 2002. Migration, local adaptation and the evolution of plasticity. *Trends Ecol. Evol.* 17:540–541.
- Kirkpatrick, M. and V. Ravigné. 2002. Speciation by natural and sexual selection: models and experiments. *Am. Nat.* 159:S22–S35.
- Kobl Müller, S., B. Egger, C. Sturmbauer, and K. M. Sefc. 2007. Evolutionary history of Lake Tanganyika's scale-eating cichlid fishes. *Mol. Phylogenet. Evol.* 44:1295–305.
- Kolmann, M. A., J. M. Huie, K. Evans, and A. P. Summers. 2018. Specialized specialists and the narrow niche fallacy: a tale of scale-feeding fishes. *R. Soc. Open Sci.* 5:171581.
- Kopp, M., M. R. Servedio, T. C. Mendelson, R. J. Safran, R. L. Rodríguez, M. E. Hauber, et al. 2018. Mechanisms of assortative mating in speciation with gene flow: connecting theory and empirical research. *Am. Nat.* 191:1–20.
- Koskella, B. and C. M. Lively. 2009. Evidence for negative frequency-dependent selection during experimental coevolution of a freshwater snail and a sterilizing trematode. *Evolution* 63:2213–21.
- Kulmuni, J. and A. M. Westram. 2017. Intrinsic incompatibilities evolving as a by-product of divergent ecological selection: considering them in empirical studies on divergence with gene flow. *Mol. Ecol.* 26:3093–3103.
- Kusche, H., H. J. Lee, and A. Meyer. 2012. Mouth asymmetry in the textbook example of scale-eating cichlid fish is not a discrete dimorphism after all. *Proc. R. Soc. B* 279:4715–4723.
- Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* 33:402–416.

- Lande, R. and S. J. Arnold. 1983. The measurement of selection on correlated characters. *Evolution* 37:1210–1226.
- Landis, M., E. Edwards, and M. Donoghue. 2019. Modeling phylogenetic biome shifts on a planet with a past. *bioRxiv*. <https://doi.org/10.1101/832527>.
- Lavin, A. and J. McPhail. 1986. Adaptive divergence of trophic phenotype among freshwater populations of the threespine stickleback (*Gasterosteus aculeatus*). *Can. J. Fish. Aquat. Sci.* 43:2455–2463.
- Lencer, E. S., M. L. Riccio, and A. R. McCune. 2016. Changes in growth rates of oral jaw elements produce evolutionary novelty in Bahamian pupfish. *J. Morphol.* 277:935–947.
- Mack, K. and M. Nachman. 2017. Gene regulation and speciation. *Trends Genet.* 33:68–80.
- Mack, K., P. Campbell, and M. Nachman. 2016. Gene regulation and speciation in house mice. *Genome Res.* 26:451–461.
- Mappes, T., M. Koivula, E. Koskela, T. A. Oksanen, T. Savolainen, and B. Sinervo. 2008. Frequency and density-dependent selection on life-history strategies - a field experiment. *PLoS One* 3:e1687.
- Martin, C. H. 2016a. Context dependence in complex adaptive landscapes: frequency and trait-dependent selection surfaces within an adaptive radiation of Caribbean pupfishes. *Evolution* 70:1265–1282.
- . 2016b. The cryptic origins of evolutionary novelty: 1000-fold faster trophic diversification rates without increased ecological opportunity or hybrid swarm. *Evolution* 70(11):2504–2519.
- 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. Syst.* 50:569–593.
- Martin, C. H. and L. C. Feinstein. 2014. Novel trophic niches drive variable progress towards ecological speciation within an adaptive radiation of pupfishes. *Mol. Ecol.* 23:1846–1862.
- Martin, C. H. and P. C. Wainwright. 2011. Trophic novelty is linked to exceptional rates of morphological diversification in two adaptive radiations of *Cyprinodon* pupfishes. *Evolution* 65:2197–2212.
- 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.* 283:23–34.
- Martin, C. H., P. A. Erickson, and C. T. Miller. 2017. The genetic architecture of novel trophic specialists: larger effect sizes are associated with exceptional oral jaw diversification in a pupfish adaptive radiation. *Mol. Ecol.* 26:624–638.
- Martin, C. H., J. A. McGirr, E. J. Richards, and M. E. St. John. 2019. How to investigate the origins of novelty: insights gained from genetic, behavioral, and fitness perspectives. *Integr. Org. Biol.* 1:obz018.
- . 2013a. A remarkable species flock of *Cyprinodon* pupfishes endemic to San Salvador Island, Bahamas. *Bull. Peabody Mus. Nat. Hist.* 54:231–240.
- . 2013b. Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. *Science* 339:208–211.
- . 2013c. On the measurement of ecological novelty: scale-eating pupfish are separated by 168 my from other scale-eating fishes. *PLoS One* 8:e71164.
- Matessi, C., A. Gimelfarb, and S. Gavrilets. 2002. Long-term buildup of reproductive isolation promoted by disruptive selection: how far does it go? *Selection* 2:41–64.
- Matthews, D. G. and R. C. Albertson. 2017. Effect of craniofacial genotype on the relationship between morphology and feeding performance in cichlid fishes. *Evolution* 71:2050–2061.
- McGirr, J. A. and C. H. Martin. 2020a. Ecological divergence in sympatry causes gene misexpression in hybrids. *Mol. Ecol.* 29:2707–2721.
- . 2020b. Conspicuous candidate alleles point to cis-regulatory divergence underlying rapidly evolving craniofacial phenotypes. *bioRxiv*. <https://doi.org/10.1101/2020.03.23.003947>.
- . (2017). Novel candidate genes underlying extreme trophic specialization in Caribbean pupfishes. *Mol. Biol. Evol.* 34:873–888.
- . 2018. Parallel evolution of gene expression between trophic specialists despite divergent genotypes and morphologies. *Evol. Lett.* 2:62–75.
- . 2019. Hybrid gene misregulation in multiple developing tissues within a recent adaptive radiation of *Cyprinodon* pupfishes. *PLoS One* 14:e0218899.
- Mitchell-Olds, T. and Shaw. 1987. Regression analysis of natural selection: statistical inference and biological interpretation. *Evolution* 41:1149–1161.
- Morrissey, M. B. 2014. In search of the best methods for multivariate selection analysis. *Methods Ecol. Evol.* 5:1095–1109.
- Nosil, P., R. Villoutreix, C. F. de Carvalho, T. E. Farkas, V. Soria-Carrasco, J. L. Feder, et al. 2018. Natural selection and the predictability of evolution in Timemaster insects. *Science* 359:765–770.
- Nychka, D., R. Furrer, J. Paige, and S. Sain. 2017. fields: tools for spatial data. R Packag. version 9.6.
- O’Meara, B. C. 2012. Evolutionary inferences from phylogenies: a review of methods. *Annu. Rev. Ecol. Syst.* 43:267–285.
- Olendorf, R., F. H. Rodd, D. Punzalan, A. E. Houde, C. Hurt, D. N. Reznick, et al. 2006. Frequency-dependent survival in natural guppy populations. *Nature* 441:633–636.
- Ortiz-Barrientos, D., B. Counterman, and M. Noor. 2007. Gene expression divergence and the origin of hybrid dysfunctions. *Genetica* 129:71–81.
- Puebla, O., E. Bermingham, and F. Guichard. 2012. Pairing dynamics and the origin of species. *Proc. R. Soc. B* 279:1085–1092. <https://doi.org/10.1098/rspb.2011.1549>
- Otto, S. P., M. R. Servedio, and S. L. Nuismer. 2008. Frequency-dependent selection and the evolution of assortative mating. *Genetics* 179:2091–2112.
- Parnell, N. F., C. D. Hulsey, and J. T. Strelman. 2008. Hybridization produces novelty when the mapping of form to function is many to one. *BMC Evol. Biol.* 8:122.
- . 2012. The genetic basis of a complex functional system. *Evolution* 66:3352–3366.
- Pfennig, D. 1992. Polyphenism in spadefoot toad tadpoles as a logically adjusted evolutionarily stable strategy. *Evolution* 46:1408–1420.
- Pfennig, D. W., A. M. Rice, and R. A. Martin. 2007. Field and experimental evidence for competition’s role in phenotypic divergence. *Evolution* 61:257–271.
- Polechová, J. and N. H. Barton. 2005. Speciation through competition: a critical review. *Evolution* 59:1194–1210.
- R Development Core Team. 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna.
- Rabosky, D. L. 2009. Ecological limits on clade diversification in higher taxa. *Am. Nat.* 173:662–674.
- Raffini, F. and A. Meyer. 2018. A comprehensive overview of the developmental basis and adaptive significance of a textbook polymorphism: head asymmetry in the cichlid fish *Perissodus microlepis*. *Hydrobiologia* 832:65–84.
- Renaut, S. and L. Bernatchez. 2011. Transcriptome-wide signature of hybrid breakdown associated with intrinsic reproductive isolation in lake whitefish species pairs (*Coregonus* spp. Salmonidae). *Heredity* 106:1003–1011.
- Renaut, S., A. Nolte, and L. Bernatchez. 2009. Gene expression divergence and hybrid misexpression between lake whitefish species pairs (*Coregonus* spp. Salmonidae). *Mol. Biol. Evol.* 26:925–936.

- Richards, E. and C. Martin. 2017. Adaptive introgression from distant Caribbean islands contributed to the diversification of a microendemic radiation of trophic specialist pupfishes. *PLoS Genet.* 13:e1006919.
- Richards, E. J., J. A. McGirr, J. R. Wang, M. E. St. John, J. W. Poelstra, M. J. Solano, et al. 2020. Major stages of vertebrate adaptive radiation are assembled from a disparate spatiotemporal landscape. *bioRxiv*. <https://doi.org/10.1101/2020.03.12.988774>.
- Reynolds, R. J., G. de los Campos, S. P. Egan, and J. R. Ott. 2016. Modelling heterogeneity among fitness functions using random regression. *Methods in Ecology and Evolution* 7:70–79. <https://doi.org/10.1111/2041-210X.12440>
- Rosenzweig, M. L. 1978. Competitive speciation. *Biol. J. Linn. Soc.* 10:275–289.
- Roughgarden, J. 1972. Evolution of niche width. *Am. Nat.* 106:683–718.
- Sazima, I. 1983. Scale-eating in characoids and other fishes. *Environ. Biol. Fishes* 9:87–101.
- Sazima, I., D. De Zoologia, U.E. De Campinas, and S. Paulo. 1983. Scale-eating in characoids and other fishes. Pp. 9–23 in T. M. Zaret, ed. *Evolutionary ecology of neotropical freshwater fishes. Developments in environmental biology of fishes. Vol. 3.* Springer, Dordrecht, the Netherlands.
- Schluter, D. 2003. Frequency dependent natural selection during character displacement in sticklebacks. *Evolution* 57:1142–1150.
- Schluter, D. and D. Nychka. 1994. Exploring fitness surfaces. *Am. Nat.* 143:597–616.
- Schluter, D., B. Columbia, and C. E-mail. 2003. Frequency dependent natural selection during character displacement in sticklebacks. *Evolution* 57:1142–1150.
- Schumer, M. and Y. Brandvain. 2016. Determining epistatic selection in admixed populations. *Mol. Ecol.* 25:2577–2591.
- Schumer, M., R. Cui, D. L. Powell, R. Dresner, G. G. Rosenthal, and P. Andolfatto. 2014. High-resolution mapping reveals hundreds of genetic incompatibilities in hybridizing fish species. *Elife* 3:1–21.
- Seehausen, O. and D. Schluter. 2004. Male-male competition and nuptial-colour displacement as a diversifying force in Lake Victoria cichlid fishes. *Proc. Biol. Sci.* 271:1345–1353.
- Selz, O. and O. Seehausen. 2019. Interspecific hybridization can generate functional novelty in cichlid fish. *Proceeding R. Soc. London Ser. B* 286:20191621
- Selz, O. M., K. Lucek, K. A Young, and O. Seehausen. 2014. Relaxed trait covariance in interspecific cichlid hybrids predicts morphological diversity in adaptive radiations. *J. Evol. Biol.* 27:11–24.
- Servedio, M. R. and J. W. Boughman. 2017. The role of sexual selection in local adaptation and speciation. *Annu. Rev. Ecol. Evol. Syst.* 48:85–109.
- Servedio, M. R. and R. Burger. 2014. The counterintuitive role of sexual selection in species maintenance and speciation. *Proc. Natl. Acad. Sci.* 111:8113–8118.
- Shaw, R. G., C. J. Geyer, S. Wagenius, H. H. Hangelbroek, and J. R. Eterson. 2008. Unifying life-history analyses for inference of fitness and population growth. *Am. Nat.* 172:E35–47.
- Simpson, G. G. 1944. *Tempo and mode in evolution.* Columbia Univ. Press, New York.
- Sinervo, B., E. Svensson, and T. Comendant 2000. Density cycles and an offspring quantity and quality game driven by natural selection. *Nature* 406:985–988.
- St. John, M. and C. Martin. 2019. Scale-eating specialists evolved adaptive feeding kinematics within a microendemic radiation of San Salvador Island pupfishes. *bioRxiv*. <https://doi.org/10.1101/648451>
- St. John, M. E., J. A. McGirr, and C. H. Martin. 2019. The behavioral origins of novelty: did increased aggression lead to scale-eating in pupfishes? *Behav. Ecol.* 30:557–569.
- St. John, M. E., R. Holzman, C. H. Martin. 2020. Rapid adaptive evolution of scale-eating kinematics to a novel ecological niche. *Journal of Experimental Biology* 223:jeb217570
- Svensson, E. I. and R. Calsbeek. 2012. *The adaptive landscape.* Oxford Univ. Press, Oxford, U.K.
- Uyeda, J. C. and L. J. Harmon. 2014. A novel Bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. *Syst. Biol.* 63:902–918.
- Weeks, A. and A. Hoffmann. 2008. Frequency-dependent selection maintains clonal diversity in an asexual organism. *Proc. Natl. Acad. Sci. USA* 105:17872–17877.
- Weinreich, D. M., R. A. Watson, and L. Chao. 2005. Perspective: sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* 59:1165–1174.
- Whitlock, M., P. Phillips, F. Moore, and S. Tonsor. 1995. Multiple fitness peaks and epistasis. *Annu. Rev. Ecol. Syst.* 26:601–629.
- Wood, S. 2017. *Generalized additive models: an introduction with R.* 2nd ed. Chapman and Hall, Lond.
- Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proc. Sixth Int. Congr. Genet.* 1:356–366.

Associate Editor: Z. Gompert

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Sample sizes for high-frequency and low-frequency field enclosures and laboratory controls for each lake hybrid population.

Table S2. Trait loadings on the two linear discriminant axes maximizing phenotypic separation among F1 lab-reared individuals of the three parental species from both lakes.

Table S3. Trait loadings on the two ridge axes most strongly associated with survival probability within the Crescent Pond high and low-frequency field enclosures estimated using generalized projection pursuit regression.

Table S4. Trait loadings on the two ridge axes most strongly associated with survival probability within the Little Lake (lake 2) high and low-frequency field enclosures estimated using generalized projection pursuit regression.

Table S5. Directional selection gradients (β) and matrix of quadratic and correlational selection gradients (γ) on the two main ridge axes of selection estimated using generalized projection pursuit regression on the 30-trait morphological dataset for each lake (frequency treatments pooled based on model selection evidence).

Table S6. Model selection comparison of general additive models for survival and growth rate of hybrids placed in field enclosures in both lakes.

Table S7. Model selection comparison of general additive models for survival and growth rate of hybrids placed in field enclosures in both lakes for the first two major axes of selection (A1 and A2) estimated using generalized projection pursuit regression.

Table S8. Proportion of major trait loadings for dominant functional traits along survival fitness pathways leading to lower (orange) or higher (blue) fitness for scale-eater phenotypes in Fig. 6.

Figure S1. Log-transformed lower jaw length versus log-transformed standard length for *a*) allopatric Cyprinodontidae species (black) and *b*) hybrid populations (one color per treatment) used in this study.

Figure S2. Histograms depicting the phenotypic variance of hybrid populations in high- (gray bars) and low-frequency (orange/blue) treatments in lake 1 (first row) and lake 2 (second row) on the first and second discriminant axes (LD1 and LD2 from Fig. 3).

Figure S3. Morphometric landmarks indicating the 28 linear distances, 3 angles (25-27), and standard length (SL) for *a*) lateral, *b*) close-up of the craniofacial region, *c*) dorsal view, and *d*) close-up of the injected coded wire tag in the dorsal musculature including injection site (note different hybrid image used here for clarity).

Figure S4. Growth rate fitness landscapes for *a*) high-frequency and *b*) low-frequency treatments in lake 1 (Crescent Pond).

Figure S5. Residual survival probability relative to the density of similar hybrid phenotypes in high- (first column) and low-frequency (second column) field enclosures.

Figure S6. Residual growth rate relative to the density of similar hybrid phenotypes in high- (first column) and low-frequency (second column) field enclosures.

Figure S7. Joint survival (first column) and growth (second column) fitness landscapes estimated across treatments and lake environments using generalized additive modeling.

Figure S8. Simulated 1%, 10%, and 20% effect sizes in the upper region (top fifth) of the survival fitness landscapes for both low-frequency field enclosures.