# Surprising spatiotemporal stability and frequency-independence across multiple fitness peaks driving adaptive radiation in the wild

Christopher H. Martin[1*], Katelyn Gould[2], Clare Bocklage[2]

[1]Department of Integrative Biology and Museum of Vertebrate Zoology, University of California, Berkeley, CA 94720, USA

[2]Department of Biology, University of North Carolina at Chapel Hill, NC, USA

[*]Corresponding Author: chmartin@berkeley.edu

Running title: Stability of the adaptive landscape

keywords: novelty, key innovation, lepidophagy, adaptive dynamics, negative frequency-dependent disruptive selection, competition

**Abstract**

The effect of the environment on fitness in natural populations is a fundamental question in evolutionary biology. However, most empirical field studies of fitness do not experimentally manipulate phenotypes or environmental conditions and rarely investigate more than a single species or population. Thus, the relative importance of the competitive environment versus intrinsic organismal performance in shaping the location, height, and fluidity of fitness peaks on the adaptive landscape remains largely unknown. We experimentally tested the effect of competitive environment on a multi-peak fitness landscape driving a microendemic adaptive radiation of generalist and trophic specialist pupfishes on San Salvador Island, Bahamas. We manipulated phenotypes, by generating lab-reared hybrid crosses, and competitive environment, by altering the frequency of rare phenotypes within field enclosures in their natural hypersaline lake environments on San Salvador. We tracked the growth and survival of 2,611 F4/F5 hybrids for 3 to 11 months in high- and low-frequency treatments replicated across two lake populations. We found strong evidence for frequency-dependent growth rates within and between enclosures, but no evidence for frequency-dependent survival differences. However, both fitness proxies supported a complex fitness landscape isolating generalist phenotypes on a local fitness peak separated by a small fitness valley from a higher fitness peak for molluscivores and a large fitness valley isolating the scale-eating phenotype in all trait dimensions. The striking consistency of this multi-peak fitness landscape across competitive environments, multivariate trait axes, and a previous field experiment provides experimental evidence for stasis, possibly due to fixed biomechanical constraints on organismal performance. These results challenge existing theory and highlight the interplay of organism and environment underlying the static and dynamic features of fitness landscapes.

## Introduction

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 [1–5] which links the microevolutionary processes of natural and sexual selection in wild populations with macroevolutionary patterns of speciation, novelty, and adaptive radiation [6–9].

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 [4] and Simpson's [10] original conceptions of genotypic and phenotypic fitness landscapes, respectively, and Fisher's geometric model [11], fitness optima are static on a high-dimensional fitness landscape. These static fitness peaks are due to epistasis within genotypic networks [12] or functional tradeoffs between different ecological niches or collections of similar niches known as adaptive zones [13]. However, a more recent paradigm originating from game theory [14–17] proposes that the fitness landscape is dynamic and resembles a trampoline: as the relative frequency of phenotypically similar individuals increases, their fitness decreases due to increased competition for resources, whereas rare phenotypes have a fitness advantage [5]. This is known as negative frequency-dependent disruptive selection and can lead to ecological speciation in sympatry, even when adapting to a unimodal resource distribution [18–22]. Laboratory and field studies of natural populations have provided extensive support for negative frequency-dependent disruptive selection [1,15,23–34] to the extent that some investigators assert its universality in all natural populations [35]. Due to its elegance and mathematical tractability, frequency-dependence has also been widely adopted by theorists as the sole mechanism for disruptive selection in most speciation models [5,9,36].

74 However, the relative contributions of the competitive environment (i.e. frequency-
75 dependence) versus static fitness optima (i.e. functional tradeoffs) in shaping the broader
76 topography of fitness landscapes across multiple species remains unknown. For example,
77 although negative frequency-dependent disruptive selection may be ubiquitous within
78 populations, the phenotypic scale of frequency-dependence is rarely measured (i.e. the
79 competition kernel), particularly across multiple species, and can have major impacts on
80 speciation [18,37]. Similarly, spatiotemporal shifts in the phenotypic location of a fitness
81 optimum for a population due to environmental stochasticity is often assumed to follow a
82 Brownian motion process [38,39] without accounting for hard boundaries imposed by functional
83 tradeoffs or biophysical constraints (but see: [40,41]). Conversely, at broader macroevolutionary
84 timescales the role of stable fitness optima in shaping trait diversification across a radiation is
85 frequently tested and supported by fitting Ornstein-Uhlenbeck models of trait diversification
86 using phylogenetic comparative methods [42–46] while ignoring the ecological effects of
87 competition and environmental stochasticity (but see [47,48]).

88 To our knowledge, all previous field experimental tests of frequency-dependent selection
89 involved only a single species or pair of ecomorphs [15,24–26,29,31,34,49,50] and there are still
90 few studies spanning multiple habitats, traits, time periods, or species. In the few examples of
91 studies at this larger scale, rare transgressive hybrid phenotypes appear to suffer a fitness cost,
92 not an advantage [51,52]. For example, in a hybrid mesocosm experiment investigating male-
93 male competition in stickleback, the rarest transgressive phenotypes experienced the lowest
94 reproductive success [52], in contrast to predictions of sexual selection as a diversifying force
95 ([53]; but see [54,55]). We also previously estimated multiple fitness peaks driving adaptive
96 radiation in a clade of trophic specialist pupfish species endemic to San Salvador Island,

Bahamas by measuring the growth and survival of laboratory-reared F2 intercross and backcross hybrids placed in field enclosures [56]. 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 phenotypes in this experiment did not experience a survival advantage and the scale of frequency-dependent survival appeared to operate only within the range of phenotypic diversity observed in the generalist population, not all three species [51]. These few experimental studies of multiple species suggest that frequency-dependent selection may have a limited phenotypic scale; however, experimental tests manipulating the frequency of rare hybrid phenotypes are needed.

Here we bridged these micro- and macroevolutionary scales by experimentally manipulating the frequency of rare hybrid phenotypes to test for an effect on the stability of multiple fitness optima in a tractable system for empirical measurements of the fitness landscape during nascent adaptive radiation. We created intermediate and transgressive hybrid phenotypes by multiple rounds of backcrossing and intercrossing between a generalist and two trophic specialist *Cyprinodon* pupfish species, manipulated the frequency of rare hybrid phenotypes between treatments in two independent lake populations on San Salvador Island, Bahamas, and then tracked individual hybrid survival and growth rates. We found negligible effects of competitor frequency on the survival of hybrids, but strong effects on growth. However, both fitness proxies supported a large and stable fitness valley isolating the novel scale-eating specialist across spatial and temporal environments. Our results lend strong empirical field support for stasis across multiple fitness peaks and valleys driving recent adaptive radiation.

## Results

### Field and laboratory survival of transgressive hybrid populations

We individually tagged and photographed 2,611 F4-F5 outbred juvenile hybrids resulting from crosses of all three species, generalist (*C. variegatus*), molluscivore (*C. brontotheroides*), and scale-eater (*C. desquamator*), from two different isolated lake populations on San Salvador Island, Bahamas before release into a high- and low-frequency field enclosure in each lake (Fig. 1; Table S1). The frequency-manipulation increased the frequency of transgressive hybrid phenotypes in the high-frequency treatment and generalist-type hybrids in the low-frequency treatment, 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 (Fig. S1). The total density of hybrids was held approximately constant between high and low-frequency treatments (lake 1: high/low: 923/823 individuals; lake 2: high/low: 842/819 individuals; Table S1).

To sample from a broader range of environmental variability, 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 half the pupfish reproductive lifespan (approximately 2 years) 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

142 individuals) were simultaneously tagged, raised in laboratory aquaria, and their deaths and
143 growth rates were tracked over one year.
144
145 **Hybrid phenotypic diversity**
146 Phenotypic similarity of each hybrid to the three parental species in each lake ($n$ = 236) was
147 calculated from 30 linear traits and angles measured from three pre-release photographs of each
148 fish (Fig. S3). These traits were used to estimate two linear discriminant (LD) axes with major
149 loadings (Table S2) of oral jaw size (LD axis 1) and nasal protrusion (LD axis 2), diagnostic
150 traits of each specialist species and major axes of rapid trait diversification within this radiation
151 [57,58]. Indeed, after correcting for standard length, residual jaw length variation within our
152 hybrid populations exceeded the range of variation observed across allopatric *Cyprinodon*
153 species and outgroup Cyprinodontidae species spanning over 20 million years since their most
154 recent common ancestor (data from [57]; Fig. 2).
155
156 **Strong directional and nonlinear selection for generalist and molluscivore phenotypes**
157 We fit thin-plate splines using generalized cross-validation or restricted maximum likelihood to
158 survival and growth rate data to visualize fitness landscapes across the discriminant morphospace
159 [59]. We found evidence for directional selection on the survival of hybrid phenotypes most
160 similar to generalist and molluscivore phenotypes in lake 1 (Fig. 3). We also found evidence of
161 stabilizing selection on the growth rates of generalist and molluscivore phenotypes in this lake
162 (Fig. S4). Despite low survival rates after 11 months in lake 2, we found strong evidence of
163 nonlinear selection for hybrid phenotypes resembling the molluscivore. These landscapes were
164 estimated from very few survivors ($n$ = 22); however, the density of non-survivors in the

surrounding regions of morphospace and similar patterns in both field enclosures provides robust support for nonlinear divergent selection for the molluscivore phenotype in this lake. 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 (Fig. 3).

We also used generalized projection pursuit regression to estimate the two multivariate linear phenotypic axes most strongly associated with survival across the 30-trait morphospace (Tables S3-S4) without making unfounded parametric assumptions about quadratic curvature as in canonical rotation analyses [60–62]. Visualization of survival fitness landscapes on the two major axes of selection indicated that the most transgressive hybrid phenotypes (i.e. least similar to parental phenotypes) suffered the lowest survival probability across treatments in both lakes (Fig. 4). In contrast, survival in laboratory control populations was shifted or opposite to the direction of selection in field enclosures along these two dominant axes of selection (Fig. 4).

We estimated the strength of multivariate selection gradients along these two major axes of selection and found significant evidence of directional selection ($P < 0.00001$) on ridge axis 1 in both lakes and marginal evidence of directional selection on ridge axis 2 in lake 2 (Table 1). The traits with the highest loadings on ridge axis 1 were a) lower jaw length and b) distance from the jaw joint to the orbit and on ridge axis 2 were a) angle between the premaxilla and orbit and b) distance from the premaxilla to the pectoral girdle (Tables S2-S3), further supporting strong selection on craniofacial trait diversification within this radiation [58].

**No evidence of frequency-dependent survival differences between treatments**

We found no evidence of significant treatment effects on either survival probability or the overall topography of the survival fitness landscape within the discriminant morphospace (Table 2). The

188 fixed effect of treatment did not improve the fit to the survival data in any of the generalized
189 additive models examined. Instead, models without the effect of frequency treatment were
190 strongly favored (Table 2; ΔAIC = 11). Similarly, we found no evidence for frequency-
191 dependent effects on survival between treatments on the two major axes of selection in either
192 lake estimated from generalized projection pursuit regression (Table 3). Models without the
193 effect of treatment provided a marginally better fit to the survival data on the two major axes of
194 selection in lake 1 (ΔAIC = 1) and were supported in lake 2 (Table 3; ΔAIC = 1.8).
195
196 **Strong evidence of frequency-dependent growth rates between treatments**
197 In contrast to survival, we found strong support for treatment effects on fitness landscapes
198 estimated from the independent fitness proxy of growth rate in lake 1 (Table 2; lake 2 was
199 excluded from all growth rate analyses due to the low number of survivors). Models including
200 the effect of frequency treatment on log-transformed growth rates in lake 1 were strongly favored
201 (ΔAIC = 68.6). This effect was robust to models including univariate splines and thin-plate
202 splines within the discriminant morphospace (Table 2).
203 Similarly, we found strong support for models including the effect of treatment for the
204 two major axes of selection across the entire morphospace in each lake estimated from
205 generalized projection pursuit regression (Table 3). The best supported model included the
206 effects of frequency treatment and univariate splines on the two major ridge axes of selection and
207 was strongly favored over a model without the effect of treatment (Table 3; ΔAIC = 46).
208
209 **Phenotypic scale of frequency-dependence for growth rate but not survival**

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, as an estimate of the rarity of each individual phenotype. We also calculated the Euclidean nearest neighbor distance to the ten most similar hybrid phenotypes in 30-dimensional morphospace for each hybrid as an estimate of the local frequency of competing phenotypes. Overall, these measures estimate the frequency of similar hybrid phenotypes relative to each hybrid to examine the phenotypic scale of frequency-dependence within each enclosure [51]. The frequency of similar phenotypes was not significantly associated with residual variation in survival not explained by hybrid phenotype along the two discriminant axes (Fig. 5; Table 2). Generalized additive models including the fixed effect of competitor frequency (distance to mean phenotype) only marginally improved the fit to the survival data (Table 2; $\Delta$AIC = 1; similar results were found when substituting nearest neighbor Euclidean distance for Mahalanobis distance: Fig. S1).

In contrast, generalized additive models including the fixed effect of competitor frequency (both Mahalanobis and nearest neighbor distance) strongly improved the fit to the growth data in lake 1 within the discriminant morphospace, even after accounting for the treatment effect (Table 2; $\Delta$AIC = 12). Similarly, for the two major axes of selection estimated from generalized projection pursuit progression, models including the fixed effect of competitor frequency substantially improved the fit to the growth data, even after accounting for the treatment effect (Table 3; $\Delta$AIC = 3.85).

**Both fitness proxies support multiple fitness peaks on the adaptive landscape**

Generalized additive modeling enables estimation and visualization of a joint fitness landscape after controlling for the effects of lake and treatment. The best supported models for survival included a fixed effect of lake and no effect of treatment, with either two univariate smoothing splines or a thin-plate plate and two smoothing splines modeling selection within the discriminant morphospace (Table 2). Models including spline terms were strongly supported over models including fixed linear effects of the discriminant axes (Table 2; $\Delta$AIC = 18.7). The best supported models for growth rate in lake 1 included a fixed effect of treatment (lake 2 was omitted from these analyses due to the low number of survivors available for growth rate estimates) and two univariate smoothing splines within the discriminant morphospace (Table 2).

Strikingly, the best combined models for survival (across both treatments in both lakes) and growth rate (both lake 1 treatments) each independently supported an isolated fitness peak for hybrids resembling the generalist separated by a fitness valley from a region of higher fitness corresponding to the molluscivore phenotype (Fig. 6). This multi-peak landscape, consistent across both survival and growth rate fitness proxies and different exposure periods in each lake, was also striking similar to a previous independent field experiment using F2 hybrids in these same lakes [56]. This supports the surprising spatiotemporal stability of a complex fitness landscape topography spanning a recent adaptive radiation of trophic specialists across years, seasons, divergent lake environments, and manipulated frequencies of hybrid transgressive phenotypes.

**Scale-eating trophic specialists are isolated by a highly stable fitness minimum**

Across both treatments and lakes, hybrids resembling the scale-eater phenotype suffered the lowest survival and growth rates, except in laboratory control populations (Figs. 3-4, 6). This

pattern is consistent with previous observations of consistent low survival and growth rates across density treatments and lakes in scale-eater F2 hybrids; however, in the previous experiment few of the F2 hybrids fell within the phenotypic range of lab-reared scale-eaters [56]. In our current experiment, over 70 hybrids occurred within the 95% confidence ellipse of lab-reared F1 scale-eater phenotypes within the discriminant morphospace.

However, 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 [63,64]. 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. 7). This results in a visualization of all possible fitness paths between generalist and scale-eater hybrid phenotypes for all subspaces within the 30-trait morphospace and, importantly, aligns these multivariate linear axes in the same direction from generalist to scale-eater phenotype for comparison of fitness curves across random subsets of the trait data. In three out of four field enclosures (with no relationship in the fourth), hybrids resembling scale-eaters suffered the lowest survival across nearly all visualized fitness paths, supporting their position in a high-dimensional fitness valley (Fig. 7). These analyses support a robust fitness minimum or 'hole' within the adaptive landscape isolating the scale-eater phenotype from other species which is consistent across different subsets of traits measured, lake environment, field exposure time, and frequency of competitors.

## Discussion

**No evidence of frequency-dependent survival in a multi-peak fitness landscape**

We conducted an experimental field test of frequency-dependent selection in a nascent adaptive radiation of trophic specialist pupfishes. We found negligible evidence of frequency-dependent survival between treatments manipulating the frequency of rare hybrid phenotypes in two different lake environments nor any relationship between survival and the frequency of competitors (Figs. 3-5; Tables 2-3). In contrast, growth rate of survivors in each lake showed strong evidence of frequency-dependence between treatments and increased as the frequency of hybrids with similar phenotypes decreased in high-frequency enclosures (Tables 2-3, Fig. S6). 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) and the two strongest axes of nonlinear selection estimated from generalized projection pursuit regression (Fig. 4). The lack of any signal of frequency-dependent survival suggests that differences in survival among hybrid phenotypes are robust to competitive conditions and reflect intrinsic viability and performance constraints. These results complement most previous experimental studies of frequency-dependence which measured only growth rates [15,28,31].

Combined estimates of the fitness landscape for both survival and growth rate indicated a surprisingly consistent topography across space and time comprised of an isolated fitness peak corresponding to the generalist phenotype separated by a small fitness valley from a higher fitness peak corresponding to the molluscivore phenotype (Fig. 6). For survival, multiple peaks emerged from evidence of higher survival of generalist and molluscivore phenotypes after three months in lake 1 combined with evidence of much higher survival of the molluscivore phenotype after 11 months in lake 2 (Fig. 3). For growth rate in lake 1, multiple peaks emerged from higher growth rates of generalist phenotypes in both enclosures combined with moderate molluscivore

301 growth rates and very low scale-eater growth rates in the low-frequency enclosure (Fig. S4).
302 These joint landscapes are admittedly reflective of combining different frequency treatments,
303 exposure periods, lake environments, and potentially different selective regimes; however,
304 comparison of general additive models provided no evidence of different selective regimes
305 between treatments and only a fixed effect of lake environment, rather than a change in fitness
306 landscape topography between lake environments (e.g. Table 2: very low support for models
307 including a 'by lake' effect), thus supporting our inference of a single combined selective
308 environment.

309 Across all four field enclosures and both fitness proxies the most prominent and
310 consistent feature of fitness landscape topography was a large fitness valley isolating hybrids
311 resembling the scale-eater, the most morphologically, ecologically, and genetically divergent
312 specialist in the radiation [65–67], from all other hybrids across treatments, lake environments,
313 field exposure periods, and across nearly all dimensions of the 30-trait hybrid morphospace
314 (Figs. 3-4, 6-7). These empirical fitness landscape measurements were also consistent with
315 observations of low fitness in hybrids partially resembling the scale-eater in a previous
316 experiment [56].

317

318 **Static features of a complex fitness landscape underlie the rare origins of trophic novelty**

319 Overall, while there was strong selection against hybrid phenotypes driving species divergence
320 in this radiation, hybrid survival showed minimal sensitivity to the frequency of competitors,
321 supporting the classic view of static fitness peaks and valleys on the adaptive landscape. Indeed,
322 the combined estimate of the fitness landscape across all four field enclosures in this study was
323 strikingly similar to the original fitness landscape estimated for the high-density field enclosure

in lake 1 in our preview work [56]. This provides robust support for the spatiotemporal stability of an isolated generalist fitness peak separated by a large fitness valley from scale-eating and a smaller fitness valley and higher fitness peak for snail-eating. Our study also provides empirical field experimental data supporting a role for the stable fitness optima and minima frequently inferred from phylogenetic comparative studies [45,68] and observed in the fossil record [69,70].

This strikingly complex, frequency-independent, and persistent multi-peak fitness landscape within San Salvador Island's hypersaline lakes provides an explanation for the rarity of trophic specialization across the Caribbean if outgroup generalist populations are under widespread stabilizing selection opposing trait diversification [56]. Indeed, Caribbean generalist populations show very little diversity in their trophic skeletal morphology or dietary diversity relative to San Salvador populations [58]. Scale-eating populations on San Salvador Island also show strong phylogenetic comparative evidence of adaptation to a new adaptive zone relative to generalist populations on San Salvador and neighboring islands [58] and stronger pre-mating isolation and genetic divergence from other species [71–73], consistent with faster rates of trait divergence and speciation driven by adaptation to this highly novel trophic niche.

**What factors underlie spatiotemporally consistent fitness landscapes?**

The frequency of similar hybrid phenotypes strongly affected growth rates both between and within enclosures; however, the ultimate survival of hybrids over both 3- and 11-month exposure periods was unrelated to their relative frequency within enclosures. In contrast to predictions of the theory of negative frequency-dependent disruptive selection [15,35,74], rare and transgressive hybrid phenotypes outside parental ranges exhibited the lowest survival rates within field enclosures. One possible explanation is that the intrinsic viability and basic

347 performance of survival tasks by these hybrids was impaired due to their mosaic hybrid genetic
348 backgrounds. This may result in mismatched craniofacial traits leading to poor foraging
349 performance or impaired suction-feeding or scale-biting performance necessary for successful
350 foraging strikes. For example, the large oral jaws of scale-eaters appear to result from at least
351 four moderate effect quantitative trait loci on different linkage groups that each increase jaw size
352 [75], suggesting that the genetic basis of even this single trophic trait is moderately polygenic
353 and may not be fully recovered within F4/F5 hybrids. Furthermore, F1 hybrid scale-eaters
354 exhibit foraging kinematics during scale-biting strikes more similar to generalists which violates
355 expectations of additivity, suggesting that kinematic behaviors may be non-additive and severely
356 mismatched in more advanced hybrids [76]. Impaired foraging performance of hybrids similar to
357 scale-eaters in field enclosures is also supported by the observation that hybrids with scale-eater
358 morphologies in laboratory control aquaria showed the highest survival rates when fed only
359 pellet foods (Figs. 3-4). This indicates that the field environment is contributing to the low
360 survival of scale-eater hybrids.

361 An alternative non-mutually exclusive possibility is that intrinsic genetic
362 incompatibilities within hybrids are contributing to their low survival rates, particularly if more
363 transgressive hybrid phenotypes are associated with a greater number or more severe genetic
364 incompatibilities. Although the San Salvador radiation only diverged approximately 10,000 years
365 old, trophic specialists within the radiation contain ancient adaptive variants also found in
366 outgroups that diverged over 5 million years ago (Richards et al. in prep.). Genetic
367 incompatibility loci are known to segregate in wild populations ([77]; reviewed in [78]) and
368 hundreds of genetic incompatibility loci have also been found between swordtail fish species of
369 similar ages [79,80]. In support of this hypothesis, F1 hybrids of specialist species within the

San Salvador radiation show evidence of hybrid gene misregulation in approximately 10% of their differentially expressed genes, i.e. gene expression levels significantly different from parental expression levels, in whole larvae at 8 days post fertilization (dpf) and within craniofacial tissues at 17-20 dpf, respectively [81,82]. Although the fitness effects on hybrids are unknown, hybrid gene misregulation has been shown to affect hybrid viability and sterility in other systems [83–87] and misregulated genes in San Salvador Island pupfish species are enriched for developmental processes affecting ecological traits relevant to trophic specialization, including craniofacial morphology, muscle mass, and nitrogen metabolism [82]. However, the link between transgressive hybrid phenotypes and the extent of hybrid misexpression or genetic incompatibilities is still unknown.

**Does a large and stable fitness valley isolate scale-eating trophic specialization?**

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 [65,88–90] and across ontogenetic stages from juveniles only to obligate scale-eating adults [91–94]. In particular, it is unusually rare among cyprinodontiform fishes: lepidophagy is only known to have evolved once within the San Salvador Island radiation and is thus separated by 168 million years of evolutionary time from the most closely related scale-eating specialists (within all three radiations of East African haplochromine cichlids [65]).

There are multiple scale-eating strategies and multiple hypotheses for the origins of scale-eating [88,90], supporting the observation that different fish lineages have different evolutionary potential to evolve scale-eating (e.g. relatively common in East African haplochromine cichlids, rare in cyprinodontiforms). For example, scale-biting pursuit predators, including the scale-

393 eating pupfish, wimple piranha (*Cataprion mento*), and cookie-cutter shark (*Isistius brasiliensis*),
394 tear off scales and mucus from the side of their prey with perpendicular strikes using their entire
395 oral jaws, whereas scale-rasping specialists, such as *Roeboides* tetras and khavalchor catfish
396 (*Pachypterus khavalchor*), employ specialized externally protruding premaxillary teeth to
397 dislodge scales from the sides of their prey [92,95,96]. However, in nearly all cases, scale-eaters
398 are size-limited relative to their prey, unlike piscivorous fishes which generally grow much
399 larger [97]. This difference suggests that the energetic payoff from scales is low relative the
400 energetic demands of high-speed scale-eating or scale-rasping strikes. Scale-eating pupfish strike
401 approximately once per minute in the wild, resulting in only a few scales and a mouthful of
402 mucus per high-speed strike completed within 10 - 15 ms [98]. Overall, the rarity, high
403 performance demands, and low caloric payoffs of scale-eating suggest that a wide and deep
404 fitness valley isolates this niche and the necessary adaptive traits from all other ecological
405 niches.

406 Consistent with this biomechanical prediction, our fitness data suggest that hybrids
407 resembling scale-eaters suffered the highest fitness costs in growth and survival across all
408 treatments, lakes, trait subsets, and time periods, but not in laboratory control aquaria fed only
409 pellet foods. This provides an unexpected explanation for the rarity of scale-eating across
410 cyprinodontiforms fishes if adapting to this specialized trophic niches requires multiple
411 phenotypic traits that only provide fitness benefits in combination (i.e. fitness epistasis [12]).
412 This is supported by existing population genomic and quantitative genetic evidence which
413 suggests that adaptation to scale-eating is multifactorial and that multiple sources of adaptive
414 variation from across the Caribbean contributed to trait diversification in this radiation
415 [67,75,99]. The relevant adaptive alleles for scale-eating may be mismatched within hybrids,

such that all necessary morphological and behavioral traits never occurred within a single individual in our experimental hybrid populations. For example, aggressive and energetically demanding scale-eating behaviors may be highly deleterious in individuals with small oral jaws [76]. Although we included outbred backcrosses to scale-eaters in our hybrid populations, some generalist and molluscivore hybrid ancestry in these individuals may have impaired their scale-eating performance and behaviors so that scale-eating strikes were no longer energetically efficient. Conversely, algae-scraping with enlarged oral jaws or highly energetic strikes by hybrids resembling the generalist may also not provide a sufficient energy surplus for survival under field conditions.

**Evidence for a holey adaptive landscape**

One alternative metaphor to the fitness landscape is that the relationship between high-dimensional genotype or phenotype space and fitness more closely resembles a series of fitness ridges bypassing large regions of low fitness, or 'holes' [63,100]. The concept of holey adaptive landscapes emphasizes the possibility that two apparent fitness peaks in some low-dimensional phenotype or genotype space may be connected by a fitness ridge in higher-dimensional spaces. We tested this prediction by looking for fitness ridges connecting scale-eater phenotypes to generalist and molluscivore phenotypes across 500 random cross-sections in the 30-trait morphospace, rather than examining only a single two-dimensional cross-section. We found no evidence of a fitness ridge in any linear combination of these 30 traits connecting scale-eaters to other species in the radiation, instead nearly all survival curves appeared to decline in the region of the morphospace containing scale-eaters (Fig. 7). In one of the low-frequency treatments, survival curves showed no decline, but this treatment also included few hybrids closely

resembling the scale-eater phenotype and showed high survival overall (Fig. 7b). Thus, the entire volume of high-dimensional morphospace within our trait dataset containing the scale-eater phenotype appears to lead to a fitness minimum, or 'hole', isolating scale-eating phenotypes from the rest of the radiation. However, we failed to find evidence of any fitness ridges connecting these phenotypes to other species, one of the major predictions of the holey adaptive landscape hypothesis. Nonetheless, fitness ridges may still exist in genotype space or across unmeasured trait dimensions.

**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 the effect of competitor frequency on a fitness surface spanning diverse hybrid phenotypes within a nascent adaptive radiation, comparable to phenotypic divergence spanning over 20 million years of Cyprinodontiform evolution (Fig. 1). Growth rate exhibited strong dependence on the frequency of similar phenotypes within and among field enclosures as predicted. However, hybrid survival showed no signal of frequency-dependence, challenging existing theory and previous experiments on a single population or species pair. Furthermore, major features of the fitness landscape, including generalist and molluscivore fitness peaks and a large fitness valley isolating scale-eaters, were strikingly consistent across lake environments, competitor frequency, and field exposure period. This challenges our existing view of empirical fitness surfaces as highly sensitive to environmental perturbation. Instead, multi-peak fitness landscapes spanning macroevolutionary levels of phenotypic disparity display both static and dynamic features across

space, time, and fitness proxy. These empirical results strengthen the connection between microevolutionary dynamics and static features of macroevolutionary fitness landscapes.

## Materials and methods

### *Focal lakes*

Laboratory breeding colonies of all three species were collected from two different focal lakes (lake 1: Crescent Pond; lake 2: Little Lake) in 2008, 2013, and 2014 using seine nets and hand nets while snorkeling. Crescent Pond is a 500 m x 60 m x 2 m deep hypersaline (40 ppt) lake containing a 0.1 – 0.3 m layer of fine silt on hard carbonate bedrock and is isolated from surrounding lakes by a small limestone ridge. Generalists and molluscivores are morphologically more similar in this lake than others due to a shorter nasal protrusion and exhibit a sister relationship across most of their genome [71,99]. There is only one other fish species in Crescent Pond, *Gambusia hubbsi*. Little Lake is a 2.3 km x 1 km x 4 m deep hypersaline (40 ppt) lake with a substrate consisting of *Cerithium* spp. snail shells and other molluscs and is connected to the larger interior Great Lake system through a sand bar. The lake contains two fish species *Gambusia hubbsi* and *Atherinomorus stipes* and the amphibious *Kryptolebias marmoratus*. Aquatic flora in both lakes is dominated by thick mats of three macroalgal species (*Acetabularia crenulata*, *Batophora oerstedii*, and *Cladophora* sp.) and wigeon grass (*Ruppia maritima*); only Little Lake is surrounded by red mangrove (*Rhizophora mangle*) forests [58].

### *Experimental hybrid populations*

Independently for each focal lake, laboratories colonies of all three species were crossed in all directions resulting in both outbred F2 intercrosses and backcrosses (as originally described in

484 [56]). Hybrid populations from each lake were raised for an additional one to two generations in
485 a common laboratory environment before backcrossing to independently sampled colonies of
486 each parental species from each lake (collected in 2013 and 2014), resulting in an outbred
487 population of F4 and F5 intercross and backcross hybrids for each lake, each originating from
488 colonies of 10 – 20 wild-caught individuals from all three species.

489 Over six weeks, approximately 6,000 F4 and F5 hybrids were bred and raised in
490 recirculating 151-liter aquaria at 5-10 ppt salinity, 25 – 29° C, on a diet of newly hatched brine
491 shrimp, commercial pellet foods, frozen seafood, and dried seaweed. 95% water changes were
492 performed every other day. Juvenile hybrids were size-sorted using 1/8" mesh bags identical to
493 the field enclosure material resulting in a range of SL from 11.2 – 22.0 mm. Hybrids were
494 packed in oxygen-permeable shipping bags (Kordon, Inc.) and shipped to San Salvador Island,
495 Bahamas in packing crates (Rubbermaid ActionPacker Storage Box) as checked airline baggage.
496
497 ***Field enclosure experiments***
498 Hybrids were temporarily held in concrete holding tanks with flow-through seawater at the
499 Gerace Research Centre after arrival. Each individual was again size-sorted through 1/8" mesh
500 bags, anesthetized in a solution of buffered MS-222 (Finquel, Inc.), photographed on the left,
501 right, and dorsal sides using a Canon EOS 60D with an EF-S 60 mm f/2.8 USM macro lens
502 mounted on a tripod with external flash, and finally injected subcutaneously in the left dorsal
503 musculature with a 1 mm x 0.1 mm stainless steel sequential coded wire tag using a handheld
504 multi-shot injector (Northwest Marine Technologies, Inc.). A fin-clip from the caudal fin of each
505 fish was removed and stored with the archival tag in 100% ethanol at -20° C. Tagged hybrids
506 were allowed to recover for at least 4 days in flow-through holding tanks at the Gerace Research

Centre to fully regrow their caudal fins while fed on a diet of newly hatched brine shrimp and commercial pellet foods.

Field enclosures were 3.6 m x 4.9 m rectangular fully-enclosed bags with a mesh size of 0.318 cm (Christiansen's Net Company, Inc.) secured to either PVC pipe set in concrete (lake 1) or iron rebar hammered into the substrate (lake 2; Fig. 1). Two enclosures were deployed in the littoral zone of each lake after removing any debris, then the bottom mesh wall was weighted down with rocks and logs covered in macroalgae from the surrounding area and filled with benthic substrate, macroalgae, and wigeon grass from surrounding areas. Care was taken to avoid introduction of any adult fishes, but smaller pupfish and mosquitofish could still enter the enclosure through the mesh.

One enclosure in each lake was randomly selected as the high frequency treatment and the second enclosure was the low frequency treatment. Hybrids were individually selected for each treatment by eye, selecting more divergent phenotypes for the high-frequency treatments and selecting the most generalist-like hybrids for the low-frequency treatments. This resulted in reduced phenotypic variance and morphospace occupation within the low-frequency treatment within each lake. This also effectively reduced the frequency of scale-eater hybrids falling within the 95% confidence interval of parental scale-eater phenotypes. Hybrid densities within enclosures approximated the natural densities of 0.9% and 3% scale-eaters and 6 and 5% molluscivores in Crescent Pond and Little Lake, respectively [65].

In Crescent Pond (lake 1), tagged hybrids were released in large batches into the high-frequency enclosure on May 15$^{th,}$ 19$^{th}$, and 28$^{th}$ and low-frequency enclosure on May 21$^{st}$, 26$^{th}$, and 28$^{th}$, 2014, respectively. In Little Lake (lake 2), all tagged hybrids were released into the high-frequency enclosure on May 18$^{th}$ and into the low-frequency enclosure on May 25$^{th}$ and

27th, 2014. Surviving hybrids were recovered from lake 1 on August 26th and 27th, 2014 after 3 months by carefully removing the substrate and sequentially lifting the entire mesh bottom, then photographed laterally and stored in 100% ethanol. To sample from a wider range of seasonal environments and recover the full time to reproductive maturity, surviving hybrids were recovered from lake 2 on April 28th and 29th, 2015 after 11 months in field enclosures. Tags were dissected from all survivors after preservation, read using a 100x tag-reading scope from Northwest Marine Technologies, Inc, and matched with archival tags to identify the survival status (0 or 1) of each tagged hybrid.

***Laboratory control***

Additional hybrids from each population (*n* = 199, Table S1) were raised in two 151-liter laboratory aquaria concurrent with the field experiment for 11 months. Control hybrids were raised on a diet of only commercial pellet foods to provide a uniform resource offering no advantages for specialized trophic morphology. Hybrids were fed once daily an amount of food that could be consumed in five minutes but not *ad libitum* and raised at 26 – 27° C in 5-10 ppt salinity (Instant Ocean) with weekly 95% water changes. Laboratory hybrids grew faster than fish placed in field enclosures and high densities within each aquarium population led to intense competition for food and high mortality rates, whereas survivors collected from field enclosures never reached maximum adult sizes or approached senescence. The day of each laboratory death was recorded, followed by removal of the tag to identify the pre-release photograph of that individual. Laboratory deaths were tracked from July 1st, 2014 until May 21st, 2015.

***Morphometrics***

553 Each hybrid used in field experimental ($n$ = 2,611) and laboratory control populations ($n$ = 199)
554 was measured for 21 landmarks (Fig. S3) on both left and right lateral sides and 10 landmarks on
555 the dorsal surface of the head plus a 4 mm size-standard grid for calibration of each image using
556 TpsDig2 [101]. F1 lab-reared individuals of each parental species from each lake ($n$ = 236) were
557 also measured in the same way. 27 linear distances and three angles were calculated from these
558 landmarks and then averaged for both lateral sides, resulting in 30 trait measurements plus
559 standard length. Traits were selected to capture phenotypic divergence in craniofacial
560 morphology and body shape (Fig. S3).

561 Linear trait measurements and angles were then imported into R (R Development Core
562 Team 2018) and size-corrected by taking the residuals from a standard major axis regression
563 relative to log-transformed standard length (SL) for each trait in the entire hybrid pool and F1
564 lab-reared parental individuals from both lakes (n = 110 generalists, 45 molluscivores, 81 scale-
565 eaters) using the sma function in the smatr package [103] in R. Standard major axis regression is
566 appropriate when trait measurement error is present on the x-axis and the y-axis and is equivalent
567 to the first principal component of phenotypic variance between these two axes [103]. Initial
568 inspection of size-correction plots indicated that ordinary least squares (OLS) regression tended
569 to overestimate the slope of the regression line, particularly for highly variable traits among
570 parental and hybrid populations such as oral jaw length. However, our results were robust to
571 OLS size-correction. No allometric scaling was observed among different species except for
572 nasal protrusion distance and nasal protrusion angle, which exhibited no association with log-
573 transformed SL and were not size-corrected. All size-corrected trait residuals and uncorrected
574 nasal protrusion distance and angle were standardized to a standard deviation of one and mean of
575 zero for comparisons across traits. There was no effect of standard length at introduction on

576 survival of hybrids (GLM logistic regression with effects of field enclosure and log-transformed
577 SL: $P = 0.709$)
578
579 ***Visualization of fitness landscapes***
580 Fitness landscapes were visualized in each enclosure by fitting thin-plate splines to the survival
581 (binomial) or growth rate (normal) data for each hybrid using generalized cross-validation
582 (GCV), which minimizes residual prediction error of the spline surface. Splines were estimated
583 using the Fields package [59] in R. When over-fitting was apparent, restricted estimation of
584 maximum likelihood (REML) was used to estimate the curvature of the spline instead of GCV
585 (used for both survival landscapes in lake 2; REML estimation of splines was identical to GCV
586 surfaces in lake 1). Growth rate was only examined in lake 1 due to the low number of survivors
587 with growth rate data from lake 2 (Table S1).
588 We focused on two different cross-sections of the 30-dimensional hybrid morphospace.
589 First, we examined selection on the two-dimensional linear discriminant morphospace
590 maximizing phenotypic separation among the three parental species from both lakes using the
591 lda function in the MASS package in R [104]. This morphospace provides a simple index of
592 hybrid similarity to each of the three parental species, comparable to multivariate hybrid indices
593 from similar studies of hybrid fitness between two species (e.g. [28,105]). Second, we examined
594 the two major axes of nonlinear selection within the 30-dimensional morphospace using
595 generalized projection pursuit regression for binomial data using the gppr function in the gsg
596 package in R [62,106], following the original recommendations of [61]. This approach avoids the
597 problematic quadratic assumptions of canonical rotation [107,108], particularly for highly
598 nonlinear data such as ours, and enables visualization of the strongest axes of nonlinear selection

599 within the dataset. We calculated the first two ridge axes using the gppr function with a binomial
600 family of response distributions and then projected hybrid phenotypes onto each ridge axis by
601 matrix multiplication. We performed generalized projection pursuit regression separately for lake
602 1 and lake 2 hybrid populations due to the large differences in survival (Table S1).
603
604 ***Generalized additive modeling***
605 We formally tested for experimental treatment effects on fitness landscapes using generalized
606 additive modeling in the mgcv package [109] in R. This modeling framework enables
607 incorporation of spline terms into generalized linear models and comparisons of models
608 containing spline, fixed, and random effect terms using AIC. For the survival data, we compared
609 models with the fixed effects of treatment and lake and all combinations of univariate smoothing
610 splines and thin-plate splines on both discriminant axes (Table 2) or both major ridge axes of
611 selection estimated from generalized projection pursuit regression (Table 3). We also explored
612 models allowing the thin-plate spline surface (i.e. the fitness landscape) to vary between lake
613 environments using the 'by' term within the thin-plate function. Finally, we included models
614 with a covariates including log-transformed standard length and distance measures of competitor
615 frequency within each enclosure based on either Mahalanobis distance or nearest-neighbor
616 Euclidean distance (see below). Models were compared using AIC. We examined a similar range
617 of models for the growth rate data, calculated from the difference in log-transformed SL between
618 pre-release photographs and surviving fish. However, we excluded lake 2 from all growth rate
619 analyses due to the low number of survivors in this lake.
620
621 ***Analyses of frequency-dependent selection within enclosures***

We used two approaches to measure the frequency of competitors within each enclosure. First, we calculated the Mahalanobis distance from each hybrid phenotype to the mean hybrid phenotype in the full 30-trait morphospace using the mahalanobis function in R. This distance estimates the disparity of each hybrid relative to the most abundant hybrid phenotypes while accounting for trait correlations. We also measured the frequency of competitors in the local region of morphospace surrounding each hybrid by calculating the sum of the Euclidean distance to the ten nearest neighbors in the full 30-trait morphospace, following the approach in [51]. We used the knn.dist function in the FNN package [110] to the calculate the matrix of distances among all hybrid pairs.

***Analyses of a fitness valley for scale-eater phenotypes***

To evaluate the stability of a survival fitness valley near scale-eaters across all dimensions in our 30-trait morphospace, we estimated and visualized smoothing splines for survival from random subsets of the trait dataset. We first estimated a smoothing spline for survival relative to the discriminant axis (LD1 in Fig. 3) for all 30 traits separating parental scale-eater phenotypes from generalist phenotypes. We then randomly drew 15 traits from the dataset and recalculated the generalist-scale-eater discriminant axis and a new survival spline for 500 trait subsets using a custom script in R. This enabled alignment of each discriminant vector from generalist to scale-eater regardless of the subset of traits sampled. Each of the 500 discriminant vectors was then rescaled to the mean parental scale-eater phenotype, so that each hybrid could be scored on the same scale proportional to their phenotypic similarity to the scale-eater (i.e. 1 = full match). Survival splines were estimated separately for each field enclosure and plotted on a shared,

644 rescaled generalist-scale-eater discriminant axis to visualize the overall evidence for a fitness
645 valley in the scale-eater region.
646
647 **Acknowledgments**
648 We thank the Miller Institute for Basic Research in Science, the University of North Carolina at
649 Chapel Hill, NSF CAREER award 1749764, and NIH 5R01DE027052-02 for funding to CHM.
650 The Bahamas Environmental Science and Technology Commission and the Ministry of
651 Agriculture kindly provided permission to export, import, and tag fish and conduct this research.
652 Erica Bree Rosenblum generously provided laboratory animal space at the University of
653 California, Berkeley. Rochelle Hanna, Velda Knowles, Thomas Rothfus, Markshaun Fields, and
654 the Gerace Research Centre provided logistical assistance; Alexander Payne, Christina Lim,
655 Qiongqiong Mei, Ivan Piedad, Sarah Bencuya, Stephanie Jeselson, Courtney Farge, Kristi Dixon,
656 and David Richard assisted with morphological measurements; and Erica Bree Rosenblum, Craig
657 Miller, Emilie Richards, Michelle St. John, and Joseph McGirr provided helpful comments on
658 the research and presentation. All animal care protocols were approved by the University of
659 California, Berkeley and the University of North Carolina at Chapel Hill Animal Care and Use
660 Committees.
661
662 **References**
663 1. Svensson EI, Calsbeek R. The adaptive landscape. Oxford: Oxford University Press;
664 2012.
665 2. Carneiro M, Hartl DL. Adaptive landscapes and protein evolution. Proc Natl Acad Sci.
666 2010;107: 1747–1751. doi:10.1073/pnas.0906192106

667 3. Lande R. Quantitative genetic analysis of multivariate evolution, applied to brain: body
668 size allometry. Evolution. 1979;33: 402–416. Available:
669 http://www.jstor.org/stable/10.2307/2407630
670 4. Wright S. The roles of mutation, inbreeding, crossbreeding and selection in evolution.
671 Proc Sixth Int Congr Genet. 1932;1: 356–366. Available:
672 http://scholar.google.com/scholar?cluster=10119462848045751904&hl=en&as_sdt=0,5#1
673 5. Gavrilets S. Fitness landscapes and the origin of species. Princeton University Press;
674 2004.
675 6. Kingsolver JG, Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hill CE, et al. The
676 strength of phenotypic selection in natural populations. Am Nat. 2001;157: 245–61.
677 doi:10.1086/319193
678 7. Arnold SJ, Pfrender ME, Jones AG. The adaptive landscape as a conceptual bridge
679 between micro- and macroevolution. Genetica. 2001;112–113: 9–32. Available:
680 http://www.ncbi.nlm.nih.gov/pubmed/11838790
681 8. Lande R, Arnold SJ. The Measurement of Selection on Correlated Characters. Evolution
682 (N Y). 1983;37: 1210. doi:10.2307/2408842
683 9. Martin CH, Richards EJ. The paradox behind the pattern of rapid adaptive radiation: how
684 can the speciation process sustain itself through an early burst? Annu Rev Ecol Evol Syst.
685 2019; In press.
686 10. Simpson G. Tempo and mode in evolution. 1944.
687 11. Fisher R. Genetical theory of natural selection. Clarendon Press; 1930.
688 12. Whitlock M, Phillips P, Moore F, Tonsor S. Multiple fitness peaks and epistasis. Annu
689 Rev Ecol Syst. 1995;26: 601–629. Available:

http://www.jstor.org/stable/10.2307/2097221

13. Simpson GG. Tempo and Mode in Evolution. 1944; 197–217.

14. Abrams P, Matsuda H, Harada Y. Evolutionarily unstable fitness maxima and stable fitness minima of continuous traits. Evol Ecol. 1993;

15. Bolnick DI. 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 (N Y). 2004;58: 608–618. doi:10.1554/03-326

16. Abrams P. Modelling the adaptive dynamics of traits involved in inter- and intarspecific interactions: An assesment of three methods. Ecol Lett. 2001;4: 166–175. doi:10.1046/j.1461-0248.2001.00199.x

17. Doebeli M, Dieckmann U. Adaptive dynamics as a mathematical tool for studying the ecology of speciation processes. J Evol Biol. 2005;18: 1194–200. doi:10.1111/j.1420-9101.2005.00912.x

18. Dieckmann U, Doebeli M. On the origin of species by sympatric speciation. Nature. 1999;400: 354–7. doi:10.1038/22521

19. Doebeli M, Dieckmann U, Metz JA, Tautz D. What we have also learned: adaptive speciation is theoretically plausible. Evolution (N Y). 2005;59: 691–699. doi:10.1111/j.0014-3820.2005.tb01028.x

20. Otto SP, Servedio MR, Nuismer SL. Frequency-dependent selection and the evolution of assortative mating. Genetics. 2008;179: 2091–112. doi:10.1534/genetics.107.084418

21. Matessi C, Gimelfarb A, Gavrilets S. Long-term buildup of reproductive isolation promoted by disruptive selection: how far does it go? Selection. 2002;2: 41–64. Available:

http://www.akademiai.com/index/P265T18035488188.pdf

22. Bürger R, Schneider KA. Intraspecific competitive divergence and convergence under assortative mating. Am Nat. 2006;167: 190–205. doi:10.1086/499375

23. Kassen R, Llewellyn M, Rainey PB. Ecological constraints on diversification in a model adaptive radiation. Nature. 2004;431: 984–988. doi:10.1038/nature02917.1.

24. Weeks A, Hoffmann A. Frequency-dependent selection maintains clonal diversity in an asexual organism. Proc Natl Acad Sci USA. 2008;105: 17872–17877.

25. Kusche H, Lee HJ, Meyer A. Mouth asymmetry in the textbook example of scale-eating cichlid fish is not a discrete dimorphism after all Subject collections Mouth asymmetry in the textbook example of scale-eating cichlid fish is not a discrete dimorphism after all. 2012; doi:10.1098/rspb.2012.2082

26. Pfennig D. POLYPHENISM IN SPADEFOOT TOAD TADPOLES AS A LOCALLY ADJUSTED EVOLUTIONARILY STABLE STRATEGY. Evolution (N Y). 1992;46: 1408–1420.

27. Bolnick DI, Lau OL. Predictable patterns of disruptive selection in stickleback in postglacial lakes. Am Nat. 2008;172: 1–11. doi:10.1086/587805

28. Schluter D. Frequency dependent natural selection during character displacement in sticklebacks. Evolution. 2003;57: 1142–50. Available: http://www.ncbi.nlm.nih.gov/pubmed/12836830

29. Olendorf R, Rodd FH, Punzalan D, Houde AE, Hurt C, Reznick DN, et al. Frequency-dependent survival in natural guppy populations. Nature. 2006;441: 633–6. doi:10.1038/nature04646

30. Koskella B, Lively CM. Evidence for negative frequency-dependent selection during

experimental coevolution of a freshwater snail and a sterilizing trematode. Evolution. 2009;63: 2213–21. doi:10.1111/j.1558-5646.2009.00711.x

31. Bolnick DI, Stutz WE. Frequency dependence limits divergent evolution by favouring rare immigrants over residents. Nature. 2017;546: 285–288. doi:10.1038/nature22351

32. Nosil P, Villoutreix R, de Carvalho CF, Farkas TE, Soria-Carrasco V, Feder JL, et al. Natural selection and the predictability of evolution inTimemastick insects. Science (80- ). 2018;359: 765–770. doi:10.1126/science.aap9125

33. Hori M. Frequency-Dependent Natural-Selection in the Handedness of Scale-Eating Cichlid Fish. Science (80- ). 1993;260: 216–219. doi:10.1126/science.260.5105.216

34. Sinervo B, Svensson E, Comendant T. Density cycles and an offspring quantity and quality game driven by natural selection. Nature. 2000;406: 985–8. doi:10.1038/35023149

35. Haller BC, Hendry AP. Solving the paradox of stasis: squashed stabilizing selection and the limits of detection. Evolution (N Y). 2014;68: 483–500. doi:10.1111/evo.12275

36. Polechová J, Barton NH. Speciation Through Competition□: a Critical Review Speciation Through Competition□: a Critical Review. 2005;59: 1194–1210. doi:10.1554/04-691

37. Baptestini EM, de Aguiar M a M, Bolnick DI, Araújo MS. The shape of the competition and carrying capacity kernels affects the likelihood of disruptive selection. J Theor Biol. 2009;259: 5–11. doi:10.1016/j.jtbi.2009.02.023

38. Hansen TF, Pienaar J, Orzack SH. A comparative method for studying adaptation to a randomly evolving environment. Evolution. 2008;62: 1965–77. doi:10.1111/j.1558-5646.2008.00412.x

39. Grant PR, Grant BR. Unpredictable Evolution in a 30-Year Study of Darwin ’ s Finches. 2002;296: 707–712.

759 40. Enquist B, Niklas K. Global allocation rules for patterns of biomass partitioning in seed
760 plants. Science (80- ). 2002;22: 1571–1520.
761 41. Boucher F, Demery V. Inferring bounded evolution in phenotypic characters from
762 phylogenetic comparative data. Syst Biol. 2016;In press. doi:10.1093/sysbio/syw015
763 42. Uyeda JC, Harmon LJ. A novel Bayesian method for inferring and interpreting the
764 dynamics of adaptive landscapes from phylogenetic comparative data. Syst Biol. 2014;63:
765 902–918. doi:10.1093/sysbio/syu057
766 43. O'Meara BC. Evolutionary Inferences from Phylogenies: A Review of Methods. Annu
767 Rev Ecol Evol Syst. 2012;43: 267–285. doi:10.1146/annurev-ecolsys-110411-160331
768 44. Harmon LJ, Losos JB, Davies TJ, Gillespie RG, Gittleman JL, Jennings WB, et al. Early
769 bursts of body size and shape evolution are rare in comparative data. Evolution. 2010;64:
770 2385–2396. doi:10.1111/j.1558-5646.2010.01025.x
771 45. Butler MA, King AA. Phylogenetic comparative analysis: a modeling approach for
772 adaptive evolution. Am Nat. 2004;164: 683–695. doi:10.1086/426002
773 46. Beaulieu J, O'Meara B. OUwie: an analysis of evolutionary rates in an OU framework. R
774 package version 1; 2012.
775 47. Rabosky DL. Ecological Limits on Clade Diversification in Higher Taxa. Am Nat.
776 2009;173: 662–674. doi:10.1086/597378
777 48. Harmon L, Andreazzi C, Debarre F, Drury J, Goldberg E, Martins A. Detecting the
778 Macroevolutionary Signal of Species Interactions. J Evol Biol. 2019;In press.
779 49. Calsbeek R, Bonvini L, Box R. GEOGRAPHIC VARIATION, FREQUENCY-
780 DEPENDENT SELECTION, AND THE MAINTENANCE OF A FEMALE-LIMITED
781 POLYMORPHISM. Evolution (N Y). 2009;64: 116–125.

782 50. Schluter D, Columbia B, E-mail C. Frequency dependent natural selection during
783 character displacement in sticklebacks. Evolution. 2003;57: 1142–50. Available:
784 http://www.ncbi.nlm.nih.gov/pubmed/12836830
785 51. Martin CH. Context dependence in complex adaptive landscapes□: frequency and trait-
786 dependent selection surfaces within an adaptive radiation of Caribbean pupfishes.
787 Evolution (N Y). 2016; 1–18. doi:10.1111/evo.12932
788 52. Keagy J, Lettieri L, Boughman JW. Male competition fitness landscapes predict both
789 forward and reverse speciation. Ecol Lett. 2015;19: 71–80. doi:10.1111/ele.12544
790 53. Seehausen O, Schluter D. Male-male competition and nuptial-colour displacement as a
791 diversifying force in Lake Victoria cichlid fishes. Proc Biol Sci. 2004;271: 1345–53.
792 doi:10.1098/rspb.2004.2737
793 54. Servedio MR, Burger R. The counterintuitive role of sexual selection in species
794 maintenance and speciation. Proc Natl Acad Sci. 2014;111: 8113–8118.
795 doi:10.1073/pnas.1316484111
796 55. Kopp M, Servedio MR, Mendelson TC, Safran RJ, Rodríguez RL, Hauber ME, et al.
797 Mechanisms of Assortative Mating in Speciation with Gene Flow: Connecting Theory and
798 Empirical Research. Am Nat. 2017;191: 000–000. doi:10.1086/694889
799 56. Martin CH, Wainwright PC. Multiple fitness peaks on the adaptive landscape drive
800 adaptive radiation in the wild. Science. 2013;339: 208–211. doi:10.1126/science.1227710
801 57. Martin CH, Wainwright PC. Trophic novelty is linked to exceptional rates of
802 morphological diversification in two adaptive radiations of *Cyprinodon* pupfishes.
803 Evolution. 2011;65: 2197–212. doi:10.1111/j.1558-5646.2011.01294.x
804 58. Martin CH. The cryptic origins of evolutionary novelty: 1000□fold faster trophic

805 diversification rates without increased ecological opportunity or hybrid swarm. Evolution
806 (N Y). 2016;70.11: 2504–2519. doi:https://doi.org/10.1101/053140
807 59. Nychka D, Furrer R, Paige J, Sain S. fields: Tools for spatial data. R Packag version 96.
808 2017; doi:10.5065/D6W957CT
809 60. Mitchell-Olds T, Shaw. Regression analysis of natural selection: statistical inference and
810 biological interpretation. Evolution (N Y). 1987;41: 1149–1161. Available:
811 http://www.jstor.org/stable/10.2307/2409084
812 61. Schluter D, Nychka D. Exploring fitness surfaces. Am Nat. 1994;143: 597–616.
813 Available: http://www.jstor.org/stable/10.2307/2462902
814 62. Morrissey MB. In search of the best methods for multivariate selection analysis. Methods
815 Ecol Evol. 2014; n/a-n/a. doi:10.1111/2041-210X.12259
816 63. Gavrilets S. Evolution and speciation on holey adaptive landscapes. Trends Ecol Evol.
817 1997;12: 307–312. doi:10.1016/S0169-5347(97)01098-7
818 64. Gavrilets S. A Dynamical Theory of Speciation on Holey. 1999;154.
819 65. Martin CH, Wainwright PC. On the measurement of ecological novelty: scale-eating
820 pupfish are separated by 168 my from other scale-eating fishes. PLoS One. 2013;8:
821 e71164. doi:10.1371/journal.pone.0071164
822 66. Martin CH, Wainwright PC. A remarkable species flock of *Cyprinodon* pupfishes
823 endemic to San Salvador Island, Bahamas. Bull Peabody Museum Nat Hist. 2013;54:
824 231–240. Available: http://www.bioone.org/doi/abs/10.3374/014.054.0201
825 67. McGirr JA, Martin CH. Novel candidate genes underlying extreme trophic specialization
826 in Caribbean pupfishes. Mol Biol Evol. 2016; msw286. doi:10.1093/molbev/msw286
827 68. Uyeda JC, Hansen TF, Mcpeek a. The million-year wait for macroevolutionary bursts

Author ( s ): Josef C . Uyeda , Thomas F . Hansen , Stevan J . Arnold and Jason Pienaar Source : Proceedings of the National Academy of Sciences of the United States of America , Stable URL : http://www.j. 2011;108: 15908–15913. doi:10.5061/dryad.7d580

69. Landis MJ, Schraiber JG. Pulsed evolution shaped modern vertebrate body sizes. Proc Natl Acad Sci USA. 2017;114: 13224–13229. doi:10.1073/pnas.1710920114

70. Estes S, Arnold SJ. Resolving the Paradox of Stasis: Models with Stabilizing Selection Explain Evolutionary Divergence on All Timescales. Am Nat. 2007;169: 227–244. doi:10.1086/510633

71. Martin CH, Feinstein LC. Novel trophic niches drive variable progress towards ecological speciation within an adaptive radiation of pupfishes. Mol Ecol. 2014;23: 1846–62. doi:10.1111/mec.12658

72. West RJD, Kodric-Brown A. Mate Choice by Both Sexes Maintains Reproductive Isolation in a Species Flock of Pupfish ( Cyprinodon spp) in the Bahamas. Ethology. 2015;121: 793–800. doi:10.1111/eth.12394

73. Kodric-Brown a., West RJD. Asymmetries in premating isolating mechanisms in a sympatric species flock of pupfish (Cyprinodon). Behav Ecol. 2013;25: 69–75. doi:10.1093/beheco/art087

74. Gigord LD, Macnair MR, Smithson a. Negative frequency-dependent selection maintains a dramatic flower color polymorphism in the rewardless orchid Dactylorhiza sambucina (L.) Soo. Proc Natl Acad Sci U S A. 2001;98: 6253–5. doi:10.1073/pnas.111162598

75. Martin CH, Erickson PA, Miller CT. The genetic architecture of novel trophic specialists: larger effect sizes are associated with exceptional oral jaw diversification in a pupfish adaptive radiation. Mol Ecol. 2017;26: 624–638. doi:10.1111/mec.13935

851 76. St. John ME, Martin CH. Scale-eating specialists evolved adaptive feeding kinematics
852 within a microendemic radiation of San Salvador Island pupfishes. bioRxiv. 2019;648451.
853 77. Fishman L, Stathos A, Beardsley PM, Williams CF, Hill JP. Chromosomal
854 rearrangements and the genetics of reproductive barriers in mimulus (monkey flowers).
855 Evolution (N Y). 2013;67: 2547–2560. doi:10.1111/evo.12154
856 78. Cutter AD. The polymorphic prelude to Bateson-Dobzhansky-Muller incompatibilities.
857 Trends Ecol Evol. 2012;27: 209–18. doi:10.1016/j.tree.2011.11.004
858 79. Schumer M, Cui R, Powell DL, Dresner R, Rosenthal GG, Andolfatto P. High-resolution
859 mapping reveals hundreds of genetic incompatibilities in hybridizing fish species. Elife.
860 2014;3: 1–21. doi:10.7554/eLife.02535
861 80. Schumer M, Brandvain Y. Determining epistatic selection in admixed populations. Mol
862 Ecol. 2016;25: 2577–2591. doi:10.1111/mec.13641
863 81. McGirr J, Martin C. Hybrid misexpression in multiple developing tissues within a recent
864 adaptive radiation of Cyprinodon pupfishes. bioRxiv. 2019;372912.
865 82. McGirr J, Martin C. Ecological divergence in sympatry causes gene misregulation in
866 hybrids. bioRxiv. 2019;717025.
867 83. Mack K, Nachman M. Gene regulation and speciation. Trends Genet. 2017;33: 68–80.
868 84. Mack K, Campbell P, Nachman M. Gene regulation and speciation in house mice.
869 Genome Res. 2016;26: 451–461.
870 85. Ortiz-Barrientos D, Counterman B, Noor M. Gene expression divergence and the origin of
871 hybrid dysfunctions. Genetica. 2007;129: 71–81.
872 86. Renaut S, Bernatchez L. Transcriptome-wide signature of hybrid breakdown associated
873 with intrinsic reproductive isolation in lake whitefish species pairs (Coregonus spp.

874 Salmonidae). Heredity (Edinb). 2011;106: 1003–1011. doi:10.1038/hdy.2010.149

875 87. Renaut S, Nolte A, Bernatchez L. Gene expression divergence and hybrid misexpression
876 between lake whitefish species pairs (Coregonus spp. Salmonidae). Mol Biol Evol.
877 2009;26: 925–936.

878 88. Sazima I. Scale-eating in characoids and other fishes. Environ Biol Fishes. 1983;9: 87–
879 101. doi:10.1007/BF00690855

880 89. Kolmann MA, Huie JM, Evans K, Summers AP. Specialized specialists and the narrow
881 niche fallacy: A tale of scale-feeding fishes. R Soc Open Sci. 2018;5.
882 doi:10.1098/rsos.171581

883 90. St. John ME, McGirr JA, Martin CH. The behavioral origins of novelty: did increased
884 aggression lead to scale-eating in pupfishes? Behav Ecol. 2018;
885 doi:10.1093/beheco/ary196

886 91. Grubh AR, Winemiller KO. Ontogeny of Scale Feeding in the Asian Glassfish, Chanda
887 nama (Ambassidae). Copeia. 2004;2004: 903–907. doi:10.1643/CE-02-095R1

888 92. Janovetz J. Functional morphology of feeding in the scale-eating specialist Catoprion
889 mento. J Exp Biol. 2005;208: 4757–4768. doi:10.1242/jeb.01938

890 93. Koblmüller S, Egger B, Sturmbauer C, Sefc KM. Evolutionary history of Lake
891 Tanganyika's scale-eating cichlid fishes. Mol Phylogenet Evol. 2007;44: 1295–305.
892 doi:10.1016/j.ympev.2007.02.010

893 94. Raffini F, Meyer A. A comprehensive overview of the developmental basis and adaptive
894 significance of a textbook polymorphism: head asymmetry in the cichlid fish Perissodus
895 microlepis. Hydrobiologia. 2018;0123456789. doi:10.1007/s10750-018-3800-z

896 95. Peterson CC, Winemiller KO. Ontogenic diet shifts and scale-eating in Roeboides dayi, a

897 Neotropical characid. Environ Biol Fishes. 1997;49: 111–118.
898 doi:10.1023/A:1007353425275
899 96. Gosavi S, Kharat S, Kumkar P, Navarange S. Interplay between behavior, morphology
900 and physiology supports lepidophagy in the catfish Pachypterus khavalchor (Siluriformes:
901 Horabagridae). Zoology. 2018;126: 185–191.
902 97. Sazima I, Zoologia D De, Campinas UE De, Paulo S. Scale-eating in characoids and other
903 fishes. 1983;9.
904 98. St. John M, Martin C. Scale-eating specialists evolved adaptive feeding kinematics within
905 a microendemic radiation of San Salvador Island pupfishes. bioRxiv. 2019;
906 99. Richards EJ, Martin CH. Adaptive introgression from distant Caribbean islands
907 contributed to the diversification of a microendemic adaptive radiation of trophic
908 specialist pupfishes. PLoS Genet. 2017;13: 1–35. doi:10.1371/journal.pgen.1006919
909 100. Gavrilets S, Li H, Vose MD. Rapid parapatric speciation on holey adaptive landscapes.
910 Proc Biol Sci. 1998;265: 1483–9. doi:10.1098/rspb.1998.0461
911 101. Rohlf FJ. Comparative methods for the analysis of continuous variables: geometric
912 interpretations. Evolution. 2001;55: 2143–60. Available:
913 http://www.ncbi.nlm.nih.gov/pubmed/11794776
914 102. R Development Core Team. R: A Language and Environment for Statistical Computing. R
915 Found Stat Comput Vienna Austria. 2016;0: {ISBN} 3-900051-07-0.
916 doi:10.1038/sj.hdy.6800737
917 103. Warton D, Duursma R, Remko A, Falster D, S T. smatr 3 - an R package for estimation
918 and inference about allometric lines. Methods Ecol Evol. 2012;3: 257–259.
919 104. Venables W, Ripley B. Modern Applied Statistics with S. Fourth Edi. Springer, New

York.; 2002.

105. Schemske DW, Bradshaw HD. Pollinator preference and the evolution of floral traits in monkeyflowers (*Mimulus*). Proc Natl Acad Sci U S A. 1999;96: 11910–5. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=18386&tool=pmcentrez&rendertype=abstract

106. Morrissey M, Sakrejda K. gsg: an R package for inference of selection gradients. R Foundation for Statistical Computing; 2013.

107. Phillips PCPC, Arnold SJSJ. Visualizing multivariate selection. Evolution (N Y). 1989;43: 1209–1222. doi:10.2307/2409357

108. Blows MW, Brooks R. Measuring Nonlinear Selection. 2003;162.

109. Wood S. Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall; 2017.

110. Beygelzimer A, Kakadet S, Langford J, Arya S, Mount D, Li S. FNN: Fast Nearest Neighbor Search Algorithms and Applications. R Foundation for Statistical Computing; 2019.

111. Stinchcombe JR, Agrawal AF, Hohenlohe P a, Arnold SJ, Blows MW. Estimating nonlinear selection gradients using quadratic regression coefficients: double or nothing? Evolution. 2008;62: 2435–40. doi:10.1111/j.1558-5646.2008.00449.x

112. Furness A, Reznick D. Convergent evolution of alternative developmental trajectories associated with diapause in African and South American killifish. Proceeding R Soc London Ser B. 2015;

113. Lavin A, McPhail J. Adaptive Divergence of Trophic Phenotype among Freshwater Populations of the Threespine Stickleback (Gasterosteus aculeatus). Can J Fish Aquat Sci.

1986;43: 2455–2463.

**Table 1.** 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). β and γ were estimated in separate regressions. Directional selection gradients in bold were significant in a one-way logistic regression model with fixed effects of ridge terms. Quadratic coefficients from the multiple regression model were doubled to estimate quadratic selection gradients [111]. However, please note that a quadratic approximation is inappropriate to model these highly nonlinear survival data [60].

| Lake 1 (Crescent Pond) | | | | γ |
|---|---|---|---|---|
| **ppr ridge axis** | **β** | ***N* = 1756** | **A1** | **A2** |

| | | | | |
|---|---|---|---|---|
| A1 | 0.095***** | | -0.050 | |
| A2 | 0.005 | | -0.026 | 0.030 |
| Lake 2 (Little Lake) | | | γ | |
| **ppr ridge axis** | **β** | ***N* = 855** | **A1** | **A2** |
| A1 | 0.126***** | | .408 | |
| A2 | 0.052* | | .185 | .081 |

*P < 0.05, **P < 0.01, ***P < 0.001, *****P < 0.00001

**Table 2.** Model selection comparison of general additive models for survival and growth rate of hybrids placed in field enclosures in both lakes. The best supported model is indicated first. Notation is adopted from the mgcv package in R: s(LD1) indicates a smoothing spline fit to discriminant axis one; tps (LD1, LD2) indicates a thin-plate spline fit to the two discriminant axes; competitor distance indicates the Mahalanobis distance from each hybrid phenotype to the mean phenotype within the 30-trait morphospace while accounting for trait correlations. Growth rate models were only analyzed for lake 1 due to the low survival rates in lake 2.

| **model** | **AIC** | **ΔAIC** |
|---|---|---|
| survival ~ s(LD1) + s(LD2) + lake | 2033.095 | - |
| survival ~ tps(LD1, LD2) + s(LD1) + s(LD2) + lake | 2033.096 | 0.001 |
| survival ~ tps(LD1, LD2) + lake | 2042.247 | 9 |
| survival ~ tps(LD1, LD2) + lake + competitor distance | 2043.28 | 10 |
| survival ~ tps(LD1, LD2) + lake + **treatment** + logSL | 2044.308 | 11 |
| survival ~ tps(LD1, LD2) + lake + logSL | 2044.120 | 11 |

| | | |
|---|---|---|
| survival ~ tps(LD1, LD2) + lake + **treatment** + lake***treatment** + logSL | 2051.536 | 18 |
| survival ~ LD1 + LD2 + lake | 2051.765 | 19 |
| survival ~ tps(LD1, LD2, by: lake) | 2298.805 | 266 |
| survival ~ tps(LD1, LD2, by: **treatment**) | 2993.312 | 960 |
| survival ~ tps(LD1, LD2) | 3034.863 | 1002 |
| log(growth) ~ s(LD1) + s(LD2) + logSL + **treatment** + competitor distance | -3226.596 | - |
| log(growth) ~ tps(LD1, LD2) + s(LD1) + s(LD2) + logSL + **treatment** + competitor distance | -3226.596 | - |
| log(growth) ~ s(LD1) + s(LD2) + logSL + **treatment** | -3214.575 | 12 |
| log(growth) ~ tps(LD1, LD2) + logSL + **treatment** | -3208.659 | 18 |
| log(growth) ~ s(LD1) + s(LD2) + logSL + competitor distance | -3160.196 | 67 |
| log(growth) ~ s(LD1) + s(LD2) + logSL | -3145.952 | 81 |
| log(growth) ~ tps(LD1, LD2) + s(LD1) + s(LD2) + logSL | -3145.952 | 81 |
| log(growth) ~ LD1 + LD2 + logSL | -3131.850 | 95 |
| log(growth) ~ LD1 + LD2 | -2830.445 | 396 |

**Table 3.** 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 projection. The best supported model is indicated first. Notation is adopted from the mgcv package in R: s(A1) indicates a smoothing spline fit to ridge axis one; tps (A1, A2) indicates a thin-plate spline fit to the two ridges; competitor distance indicates the Mahalanobis distance from each hybrid phenotype to the mean phenotype within the 30-trait morphospace while accounting for trait correlations.

| **lake** | **model** | **AIC** | **ΔAIC** |
|---|---|---|---|
| lake 1 (Crescent Pond) | survival ~ s(A1) + s(A2) | 1819.724 | - |
| | survival ~ s(A1) + s(A2) + **treatment** | 1820.595 | 0.87 |
| | survival ~ tps(A1,A2) + **treatment** | 1820.595 | 0.87 |
| | survival ~ A1 + A2 + **treatment** | 1820.594 | 0.87 |
| | survival ~ s(A1) + s(A2) + tps(A1,A2) + **treatment** | 1821.649 | 2 |
| | survival ~ s(A1) + s(A2) + **treatment** + competitor distance | 1822.595 | 3 |
| | log(growth) ~ s(A1) + s(A1) + logSL + **treatment** + competitor | -3230.104 | - |

| | | | |
|---|---|---|---|
| | distance | | |
| | log(growth) ~ s(A1) + s(A1) + logSL + **treatment** | -3226.259 | 3.8 |
| | log(growth) ~ s(A1) + s(A1) + logSL + competitor distance | -3185.980 | 44 |
| | log(growth) ~ s(A1) + s(A1) + logSL | -3180.214 | 50 |
| lake 2 (Little Lake) | survival ~ s(A1) + s(A2) | 108.7232 | - |
| | survival ~ s(A1) + s(A2) + **treatment** | 110.4928 | 1.8 |
| | survival ~ s(A1) + s(A2) + tps(A1,A2) + **treatment** | 110.4925 | 1.8 |
| | survival ~ tps(A1,A2) + **treatment** | 110.4927 | 1.8 |
| | survival ~ A1 + A2 + **treatment** | 110.4921 | 1.8 |
| | survival ~ s(A1) + s(A2) + **treatment** + competitor distance | 110.6482 | 1.9 |

**Fig. 1**

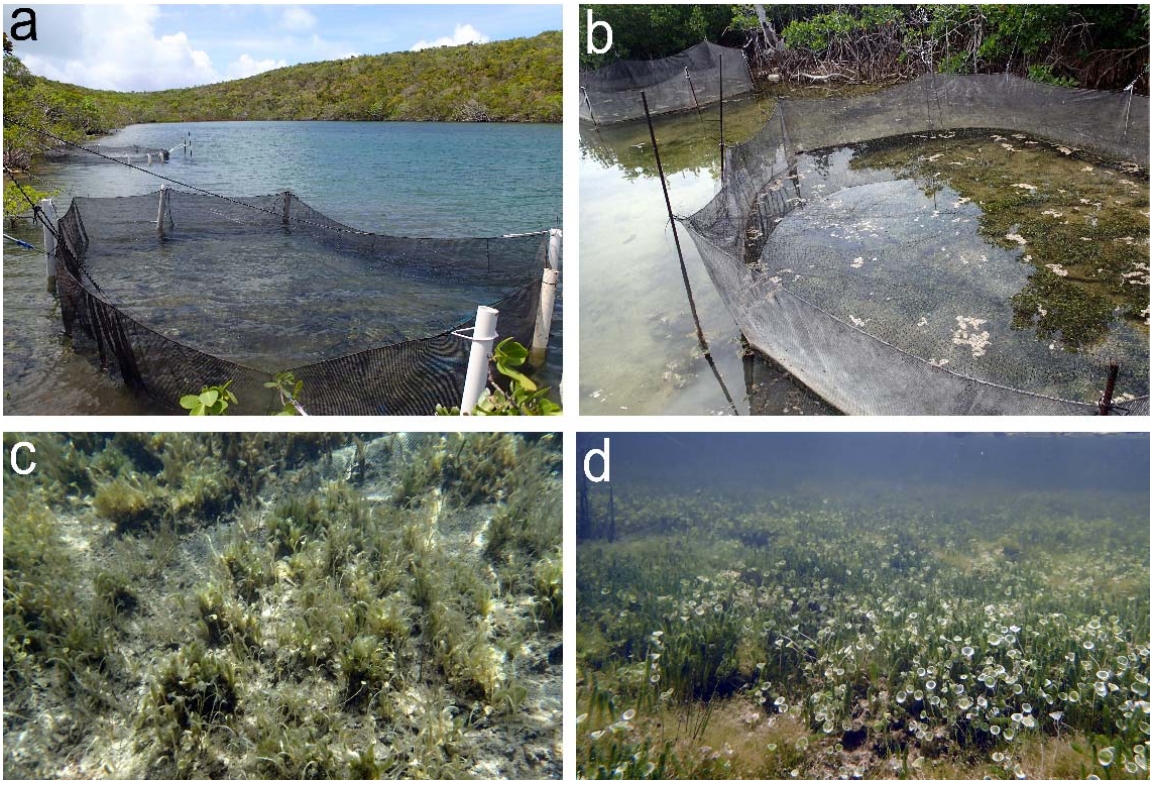


**Fig. 1 High- and low-frequency field enclosures** and the associated benthic macroalgae communities inside each enclosure typical of surrounding littoral zone habitats in lake 1 (*a*,*c*: Crescent Pond) and lake 2 (*b*,*d*: Little Lake) after 3 month and 11 month field exposure periods, respectively.

**Fig. 2**

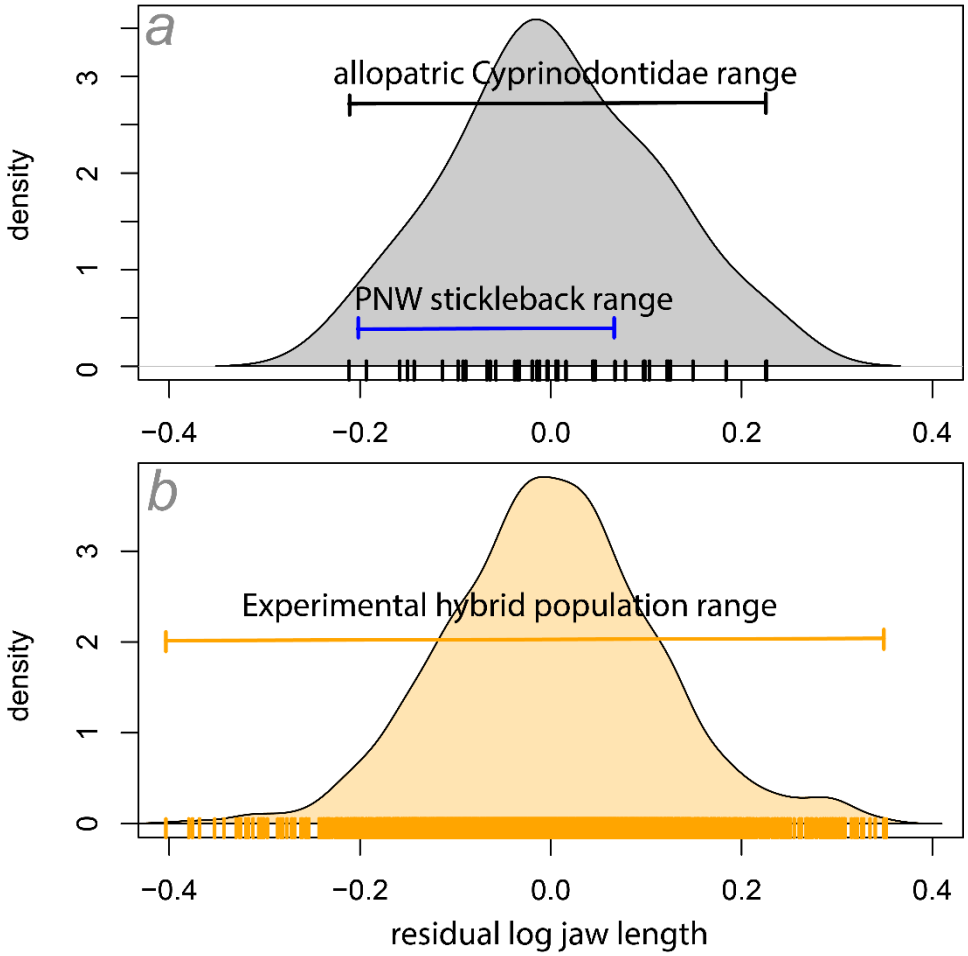



**Fig. 2 Residual jaw length variation** in *a)* **allopatric Cyprinodontidae species** (grey) diverging over 20 million years ago ([112]; data from [57]) and *b)* lab-reared F4/F5 hybrid populations (orange) measured from the field experiment in this study. The minimum and maximum residual upper jaw lengths across allopatric stickleback populations in the Pacific Northwest (PNW; data from Fig. 1 reported in [113]) are also included for comparison. Each *a)* species mean or *b)* individual F4/F5 hybrid is represented as tick marks on the x-axis plus density plots as an estimate of residual variation. Residuals were calculated from a 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 studies.

**Fig. 3**

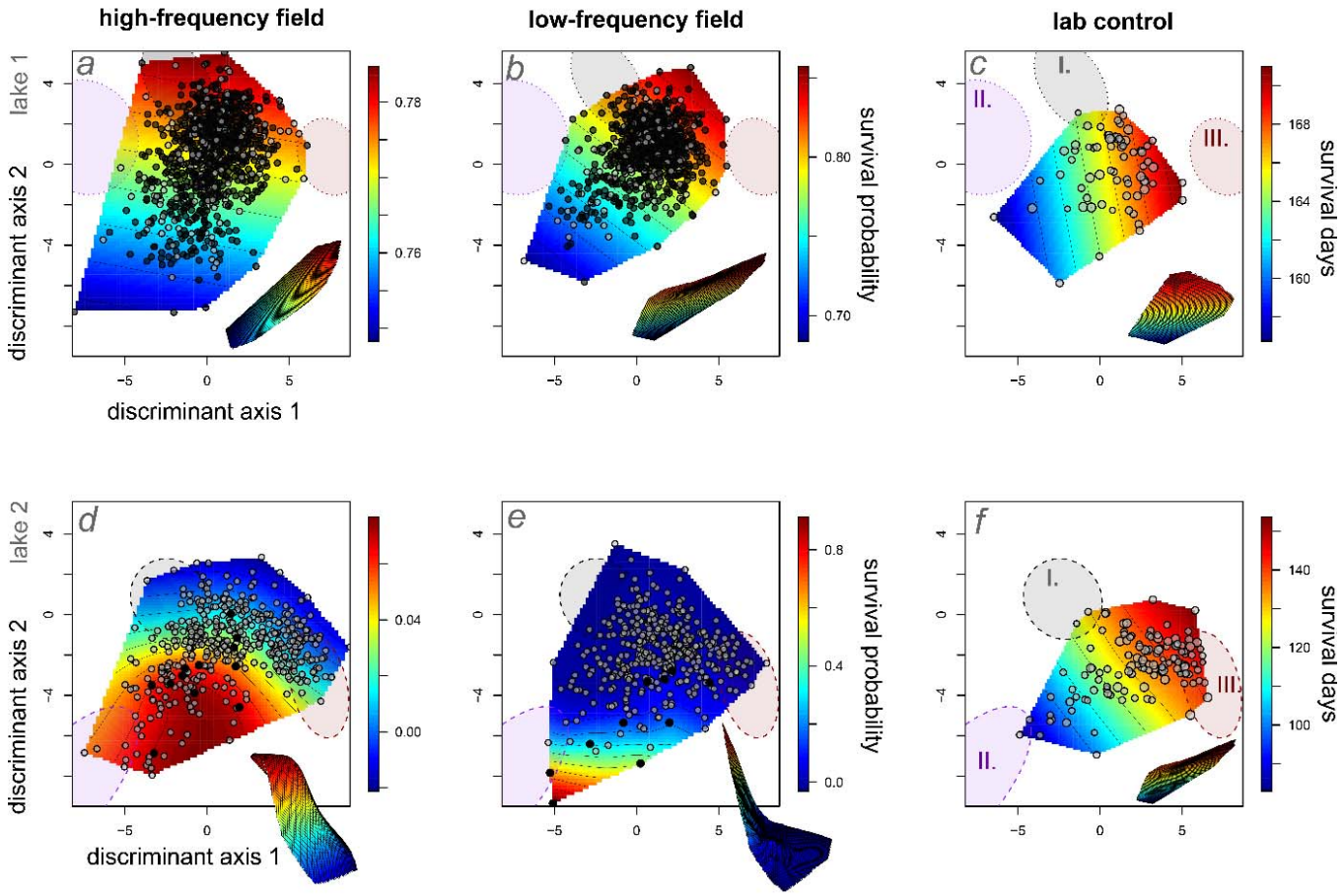



**Fig. 3 Survival fitness landscapes for hybrid populations in high- (first column) and low-frequency (second column) field enclosures and laboratory controls (third column).** Thin-plate splines predict the probability of survival (heat color) across a single linear discriminant morphospace separating 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 1-year exposure periods, respectively. Laboratory control points are proportional to the number of days each hybrid survived within 151-liter aquaria. All hybrids are plotted within a shared linear discriminant morphospace calculated from lab-reared F1 individuals of parental populations in both lakes (first row: lake 1; second row: lake 2). 95%

1033 confidence ellipses for each parental population in each lake are shown for generalists (I. grey),

1034 molluscivores (II. purple), and scale-eaters (III. red).

1035

1036

1037

1038

1039

1040

1041

1042

1043

1044

1045

1046

1047

1048

1049

1050

1051

1052

1053

1054

1055

**Fig. 4**

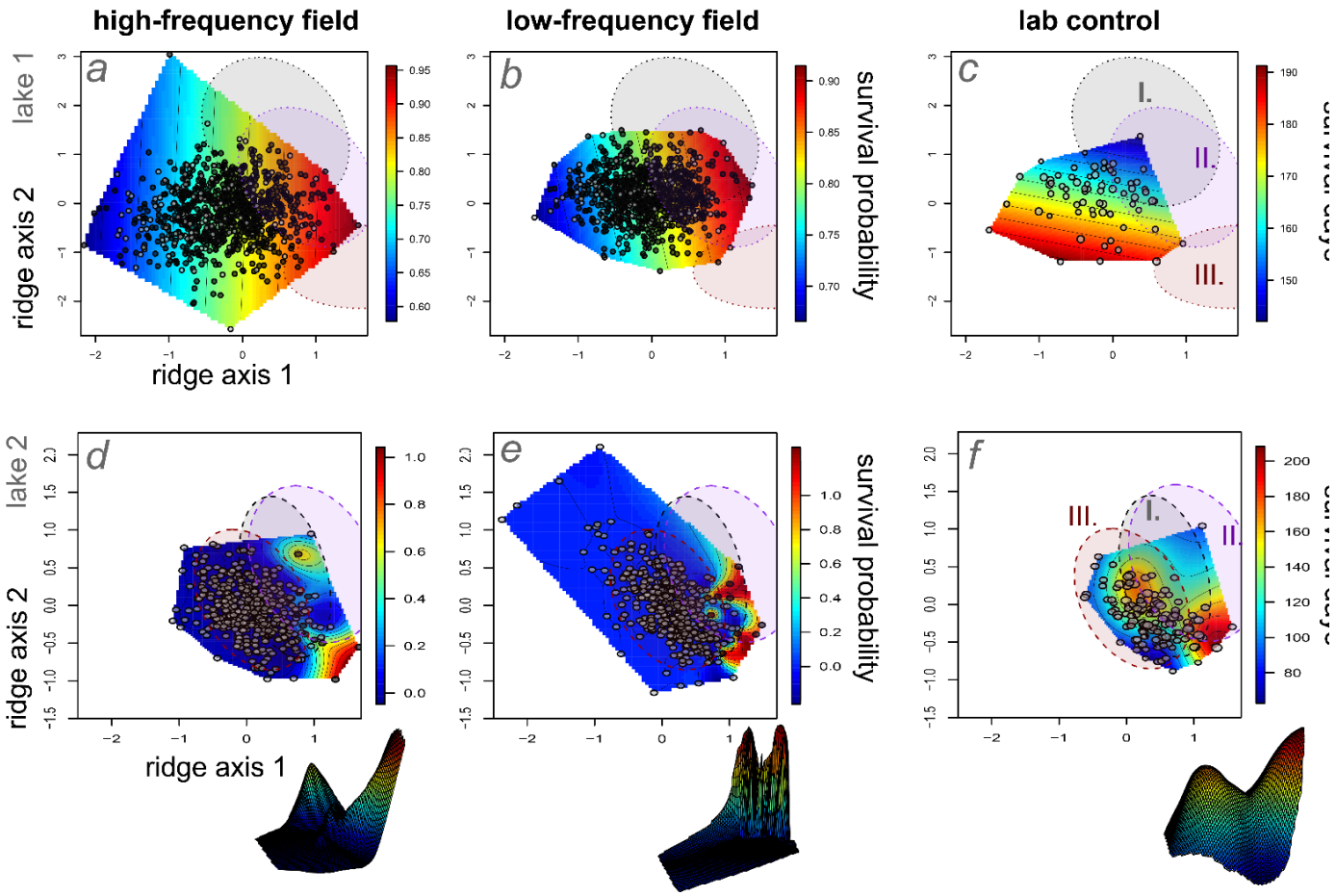



**Fig. 4 Survival fitness landscapes for hybrid populations in high- (first column) and low-frequency (second column) field enclosures and laboratory controls (third column) across the two major ridge axes of selection within the 30-trait morphospace.** Thin-plate splines predict the probability of survival (heat color) across the two major ridge axes associated with survival (Table S3-4), estimated separately for lake 1 (first row) and lake 2 (second row) hybrid populations using generalized projection pursuit regression. Survivors in field enclosures are depicted in black relative to deaths over the 3-month (first row) and 1-year (second row) exposure periods, respectively. Laboratory control points (third column) are proportional to the number of days each hybrid survived within 151-liter aquaria. 95% confidence ellipses for each

1067 parental population in each lake are shown for generalists (I. grey), molluscivores (II. purple),
1068 and scale-eaters (III. red).
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089

**Fig. 5**

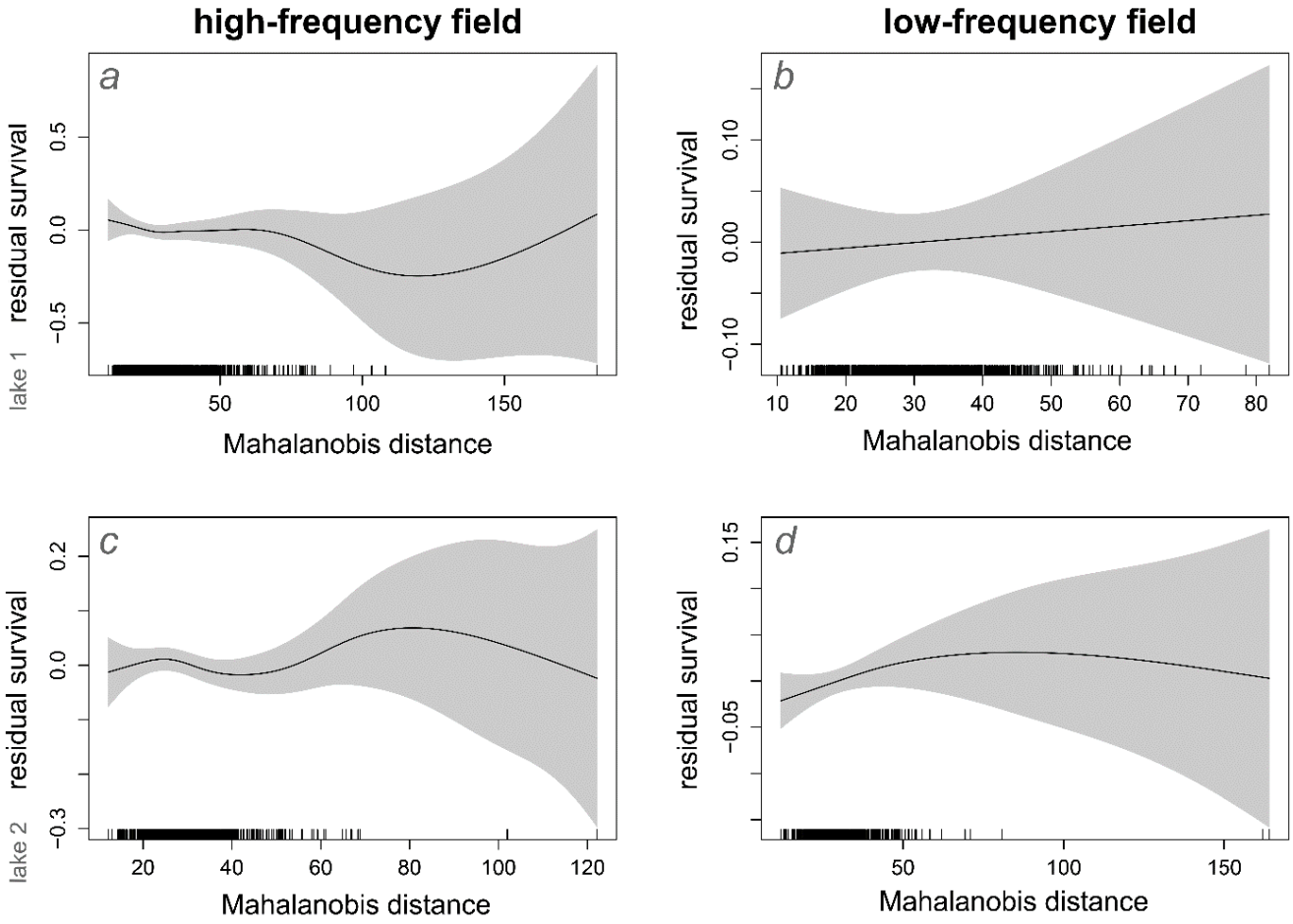



**Fig. 5 Residual survival probability relative to the density of similar hybrid phenotypes in high- (first column) and low-frequency (second column) field enclosures in each lake.** Residuals calculated from the survival probability not explained by the thin-plate splines estimated for hybrid phenotype in the discriminant morphospace (Fig. 2). The frequency of similar hybrid phenotypes was calculated for each hybrid from Mahalanobis distance to the mean hybrid phenotype within the full 30-trait hybrid morphospace account for trait correlations. Similar patterns were found using sum of the nearest-neighbor Euclidean distances to the ten most similar hybrid phenotypes in the full 30-trait morphospace (Fig. S5).

**Fig. 6**

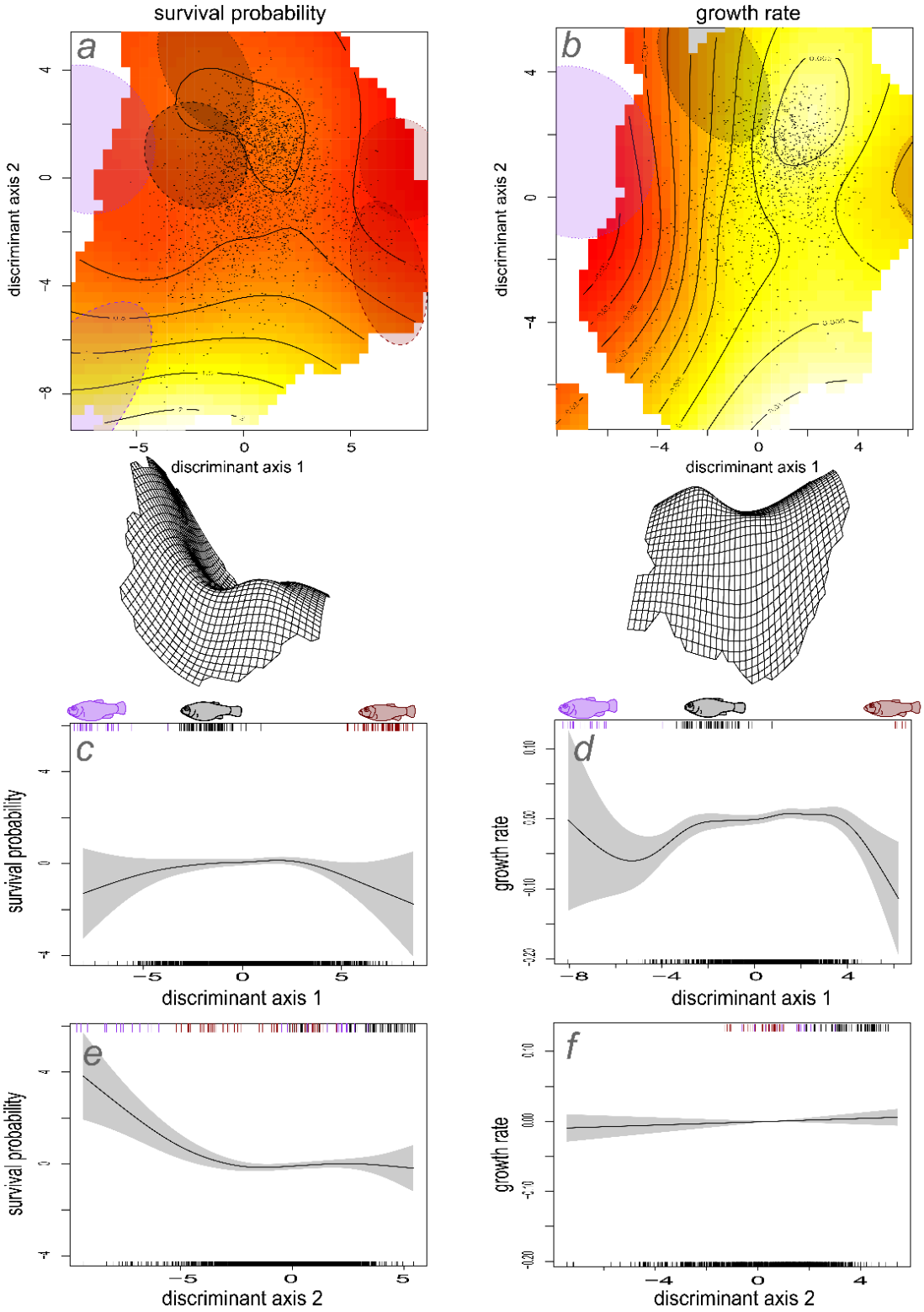
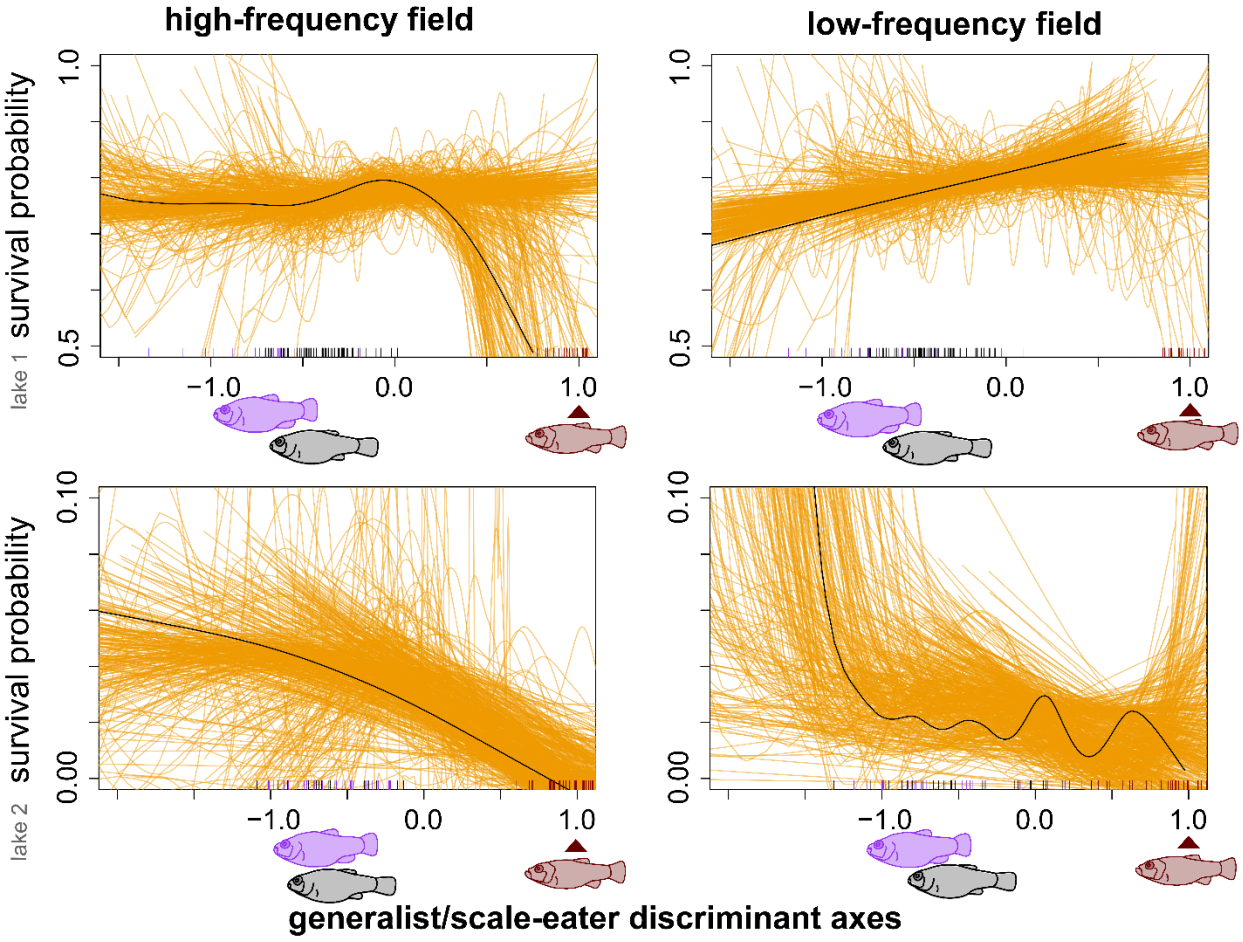


**Fig. 6 Joint survival (first column) and growth (second column) fitness landscapes estimated across treatments and lake environments using generalized additive modeling.** Thin-plate splines estimate the probability of *a)* survival controlling for lake and treatment effects across all four enclosures and *b)* growth rate controlling for treatment effects in lake 1 (lake 2 was excluded due to low survival rates). Thin-plate splines and smoothing splines (± 1 SE) are depicted within the linear discriminant morphospace separating generalist and scale-eater phenotypes (x-axis: LD1) and generalist and molluscivore phenotypes (y-axis: LD2) calculated from laboratory-reared individuals of parental populations in both lakes. 95% confidence ellipses show the location of generalist (grey), molluscivore (purple), and scale-eater (red) parental populations from lake 1 (small dashed line) and lake 2 (large dashed line). All hybrids are represented by points or tick marks on the x-axis and parental individuals are represented by tick marks on the upper margin.

1126 **Fig. 7**




1127

1128 **Fig. 7 Spaghetti plots illustrate all possible fitness paths between generalist and scale-eater**

1129 **hybrid phenotypes in high- (first column) and low-frequency (second column) field**

1130 **enclosures.** Each orange line depicts the relationship between survival and a random

1131 discriminant axis separating generalist and scale-eater phenotypes estimated from generalized

1132 cross-validation of a smoothing spline for 500 random subsets of 15 size-corrected traits (out of

1133 30); the black line illustrates the smoothing spline estimated for a discriminant axis from all 30

1134 traits, the grey lines illustrate the smoothing spline estimated for each subset. Each subsampled

1135 discriminant axis was rescaled to 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.

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

| lake | exposure period | high-frequency treatment survivors / total | low-frequency treatment survivors / total | control |
|---|---|---|---|---|
| Crescent Pond | 3 months | 712 / 923; 77.1% | 662 / 883; 75.0% | 69 |
| Little Lake | 11 months | 12 / 842; 1.4% | 10 / 819; 1.2% | 130 |

1179 **Table S2.** Trait loadings on the two linear discriminant axes maximizing phenotypic separation
1180 among F1 lab-reared individuals of the three parental species from both lakes. Numbered
1181 landmarks correspond to illustrations of linear distances and angles in Fig. S3.

| trait | trait | LD1 | LD2 |
|---|---|---|---|
| | cranialwidth | -0.10 | 0.00 |
| | innereyetosnout | -0.33 | 0.28 |
| | suspensorium | 0.74 | 0.37 |
| | dorsalsnoutlen | 0.11 | -0.37 |
| | adductorht | 0.22 | -0.23 |
| | jawlen | 1.07 | -0.19 |
| | ad2pect | -0.07 | 0.16 |
| | pmxlen | 0.52 | -0.38 |
| | foreeyewidth | 0.17 | -0.42 |
| | bodydepth | 0.10 | -0.27 |
| | dorsaltocaudal | -0.12 | 0.01 |
| | headht | 0.55 | -0.53 |
| | analtocaudal | -0.29 | -0.07 |
| | caudalpedht | -0.19 | 0.29 |
| | pmx2add | -0.76 | 0.49 |
| | jaw2pect | -0.10 | -0.07 |
| | snoutlen | 0.69 | -0.19 |
| | foresnout | -0.64 | 0.26 |
| | eyewidth | 0.01 | 0.04 |
| | eyetosnout | 0.29 | 0.71 |
| | headwidth | -0.11 | 0.23 |
| | nose | -1.06 | -1.61 |
| | hindeyewidth | -0.12 | 0.03 |
| | eyeht | -0.76 | -0.11 |
| | topeyeangle | 0.35 | 0.43 |
| | lowereyeangle | -0.23 | 0.68 |
| | nasalangle | 0.49 | 0.01 |
| | headlen | 0.02 | -1.07 |
| | bellylen | -0.44 | 0.3 |
| | pectinsertion | -0.19 | -0.19 |
| | buccalwidth | 0.16 | 0.3 |

1182
1183
1184
1185

**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. Numbered landmarks correspond to illustrations of linear distances and angles in Fig. S3.

| trait | trait | ridge axis 1 (A1) | ridge axis 2 (A2) |
|---|---|---|---|
| 1. nasal protrusion | 1. nose | -0.02 | 0.06 |
| 2. interorbital width | 2. cranialwidth | 0.2 | 0.02 |
| 3. orbit to premaxilla | 3. innereyetosnout | 0.27 | 0.22 |
| 4. suspensorium length | 4. suspensorium | -0.24 | 0.23 |
| 5. dorsal facial length | 5. dorsalsnoutlen | -0.26 | -0.22 |
| 6. adductor height | 6. adductorht | 0.26 | -0.03 |
| 7. lower jaw length | 7. jawlen | 0.21 | 0.42 |
| 8. subopercle to pectoral girdle | 8. ad2pect | -0.11 | -0.26 |
| 9. premaxilla length | pmxlen | -0.25 | -0.31 |
| 10. jaw joint to orbit | foreeyewidth | 0.36 | 0.04 |
| 11. body depth | bodydepth | 0.15 | 0.09 |
| 12. dorsal to caudal distance | dorsaltocaudal | -0.18 | 0.01 |
| 13. head height | headht | 0.17 | -0.35 |
| 14. anal to caudal distance | analtocaudal | -0.02 | 0.1 |
| 15. caudal peduncle height | caudalpedht | -0.11 | 0.17 |
| 16. lateral skull length | pmx2add | -0.13 | -0.17 |
| 17. upper jaw to pectoral girdle | jaw2pect | -0.05 | 0.15 |
| 18. lateral facial length | snoutlen | 0.05 | -0.2 |
| 19. nasal length | foresnout | -0.02 | -0.06 |
| 20. horizontal orbit diameter | eyewidth | -0.07 | -0.08 |
| 21. adductor to premaxilla | eyetosnout | -0.09 | 0.14 |
| 22. max. neurocranium width | headwidth | 0.05 | 0.13 |
| 23. orbital neurocranium width | hindeyewidth | -0.12 | -0.02 |
| 24. vertical orbit diameter | eyeht | 0.13 | 0.23 |
| 25. premaxilla to orbit angle | topeyeangle | -0.23 | 0.14 |
| 26. premaxilla to adductor angle | lowereyeangle | -0.16 | 0.24 |
| 27. nasal protrusion anagle | nasalangle | 0.06 | 0.18 |
| 28. neurocranium to premaxilla | headlen | -0.2 | -0.15 |
| 29. orbit to anal fin insertion | bellylen | -0.22 | 0.03 |
| 30. pectoral fin insertion width | pectinsertion | -0.25 | -0.04 |
| 31. gape width | buccalwidth | -0.23 | -0.12 |

**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. Numbered landmarks correspond to illustrations of linear distances and angles in Fig. S3.

| trait | trait | ridge axis 1 (A1) | ridge axis 2 (A2) |
|---|---|---|---|
| 1. nasal protrusion | nose | 0.18 | -0.02 |
| 2. interorbital width | cranialwidth | -0.07 | 0.14 |
| 3. orbit to premaxilla | innereyetosnout | -0.08 | 0.24 |
| 4. suspensorium length | suspensorium | -0.19 | 0.12 |
| 5. dorsal facial length | dorsalsnoutlen | 0.2 | -0.33 |
| 6. adductor height | adductorht | 0 | 0.01 |
| 7. lower mandible length | jawlen | -0.26 | 0.11 |
| 8. subopercle to pectoral girdle | ad2pect | -0.13 | 0.02 |
| 9. premaxilla length | pmxlen | 0.13 | -0.04 |
| 10. jaw joint to orbit | foreeyewidth | -0.05 | 0.19 |
| 11. body depth | bodydepth | -0.03 | 0.15 |
| 12. dorsal to caudal distance | dorsaltocaudal | 0.03 | -0.17 |
| 13. head height | headht | 0.04 | -0.07 |
| 14. anal to caudal distance | analtocaudal | -0.07 | -0.18 |
| 15. caudal peduncle height | caudalpedht | -0.05 | -0.15 |
| 16. lateral skull length | pmx2add | -0.23 | 0.22 |
| 17. upper jaw to pectoral girdle | jaw2pect | 0.34 | -0.43 |
| 18. lateral facial length | snoutlen | -0.17 | 0.19 |
| 19. nasal length | foresnout | 0.03 | -0.05 |
| 20. horizontal orbit diameter | eyewidth | 0 | 0.08 |
| 21. adductor to premaxilla | eyetosnout | -0.23 | 0.22 |
| 22. max. neurocranium width | headwidth | 0.08 | -0.08 |
| 23. orbital neurocranium width | hindeyewidth | 0.15 | -0.12 |
| 24. vertical orbit diameter | eyeht | 0.25 | -0.19 |
| 25. premaxilla to orbit angle | topeyeangle | -0.51 | 0.47 |
| 26. premaxilla to adductor angle | lowereyeangle | -0.26 | 0.13 |
| 27. nasal protrusion anagle | nasalangle | 0.02 | 0.01 |
| 28. neurocranium to premaxilla | headlen | 0.21 | 0.02 |
| 29. orbit to anal fin insertion | bellylen | 0.03 | -0.12 |
| 30. pectoral fin insertion width | pectinsertion | 0.12 | 0.04 |
| 31. gape width | buccalwidth | -0.19 | 0.01 |

**Fig. S1**

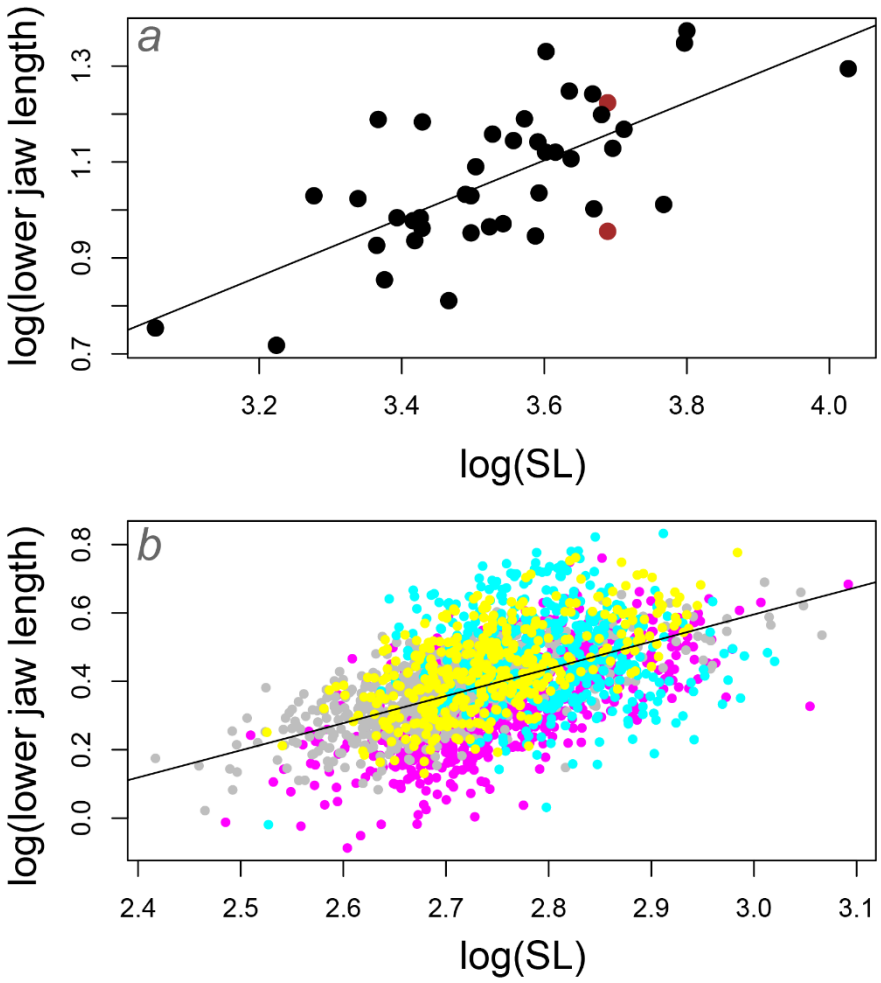



**Fig. S1** Log-transformed lower jaw length versus log-transformed standard length for *a)* allopatric Cyprinodontidae species (black) and *b)* hybrid populations (one color per treatment) used in this study. The minimum and maximum upper jaw lengths of allopatric stickleback populations in the Pacific Northwest are also shown for reference (brown). Analyses of data published in [57] and Fig. 1 reported in [113].

**Fig. S2**

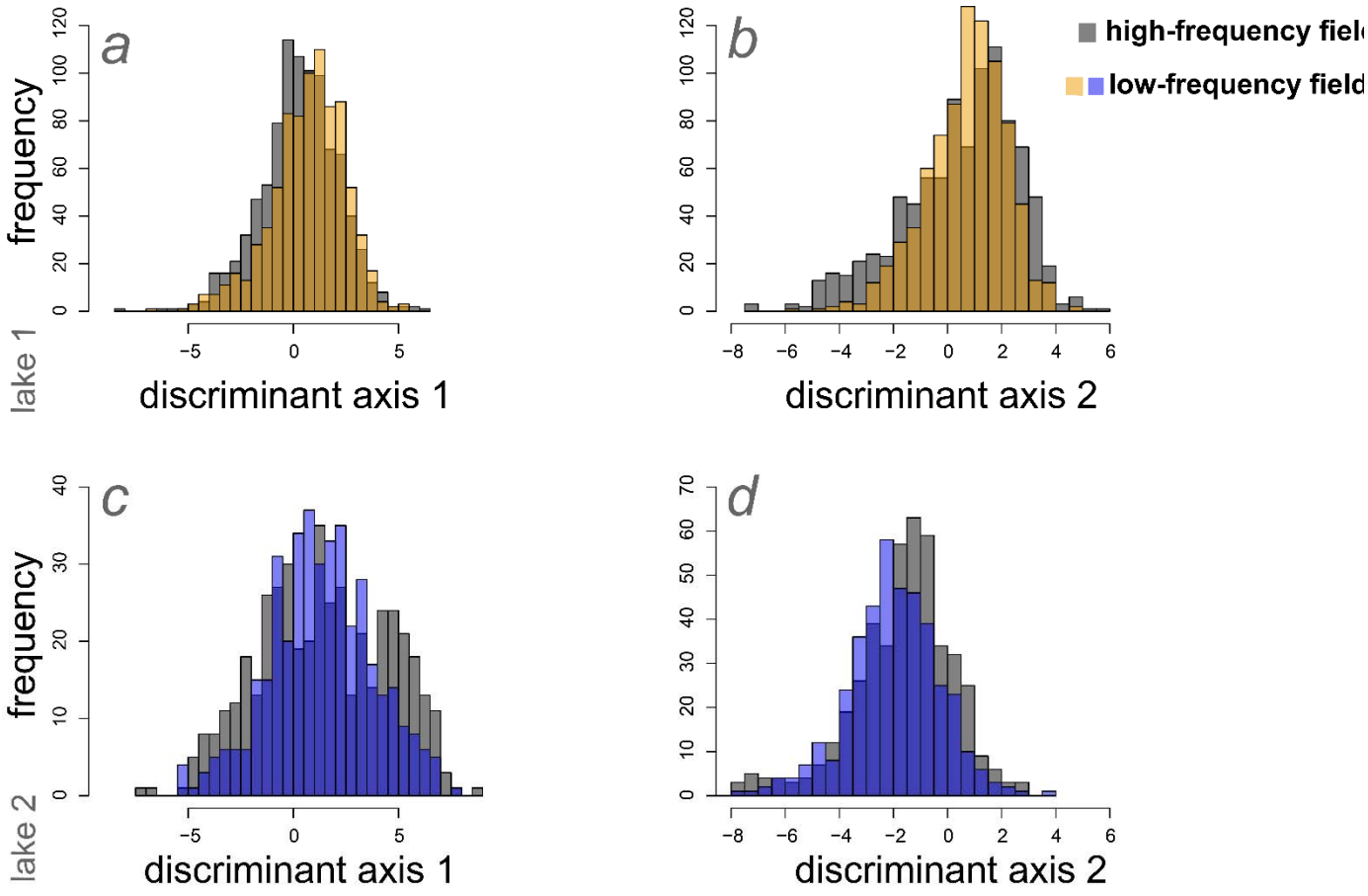
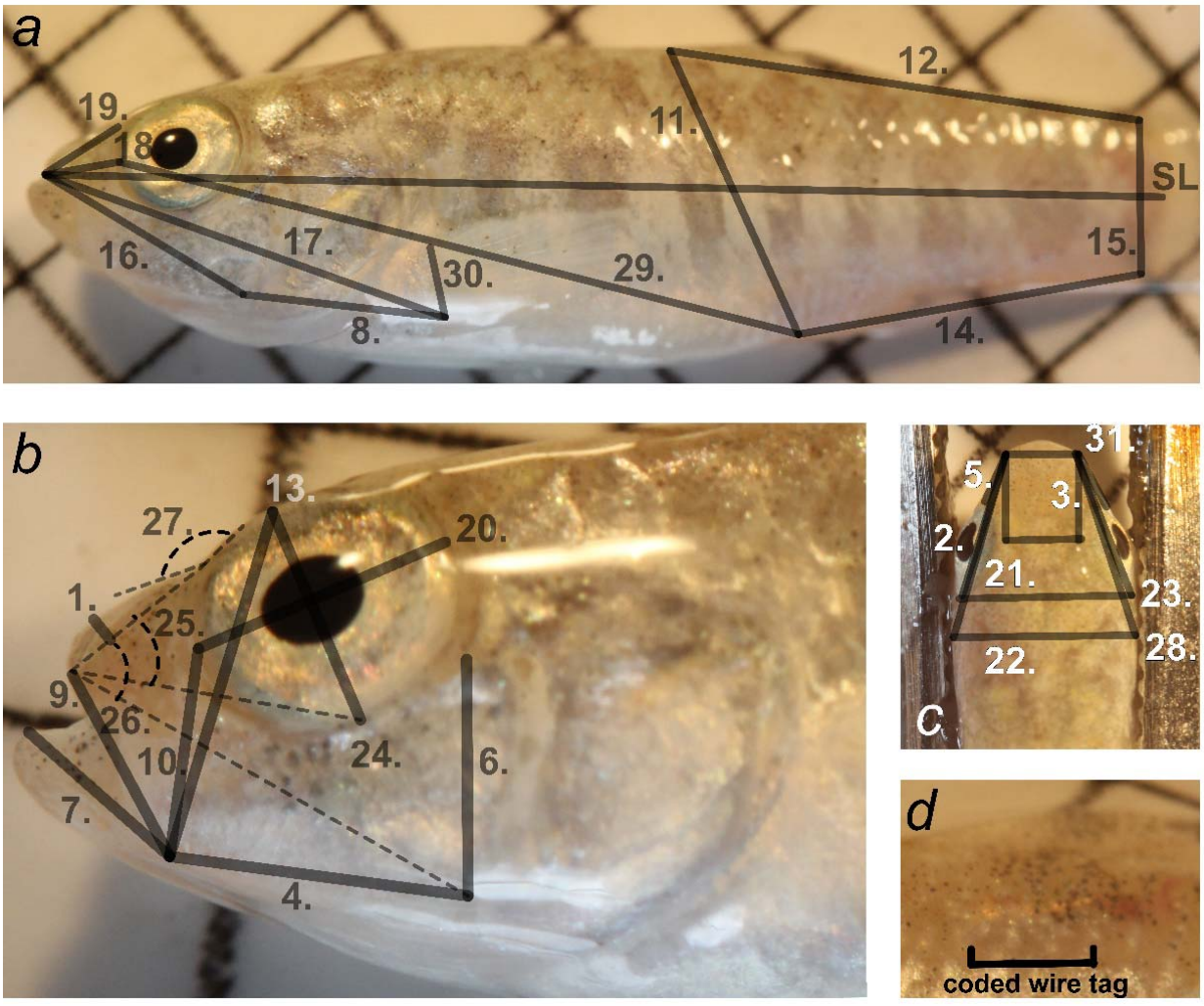



**Fig. S2** Histograms depicting the phenotypic variance of hybrid populations in high- (gray bars) and low-frequency (orange/blue) treatments in lake 1 (first row) and lake 2 (second row) on the first and second discriminant axes (LD1 and LD2 from Fig. 3).

1222 **Fig. S3**




1223

1224 **Fig. S3 Morphometric landmarks** indicating the 28 linear distances, 3 angles (25-27), and

1225 standard length (SL) for *a)* lateral, *b)* close-up of the craniofacial region, *c)* dorsal view, and *d)*

1226 close-up of the injected coded wire tag in the dorsal musculature including injection site (note

1227 different hybrid image used here for clarity). Numbers correspond to named traits in Tables 2-4.

1228 Lateral measurements were collected from pre-release photographs of both the left and right

1229 sides of each hybrid and averaged for all analyses. Dorsal view anteroposterior measurements

1230 (2,3,5,23,28) were also measured on each side and averaged for analyses. The 2 mm grid boxes

1231 in each image were used for calibration.

**Fig. S4**

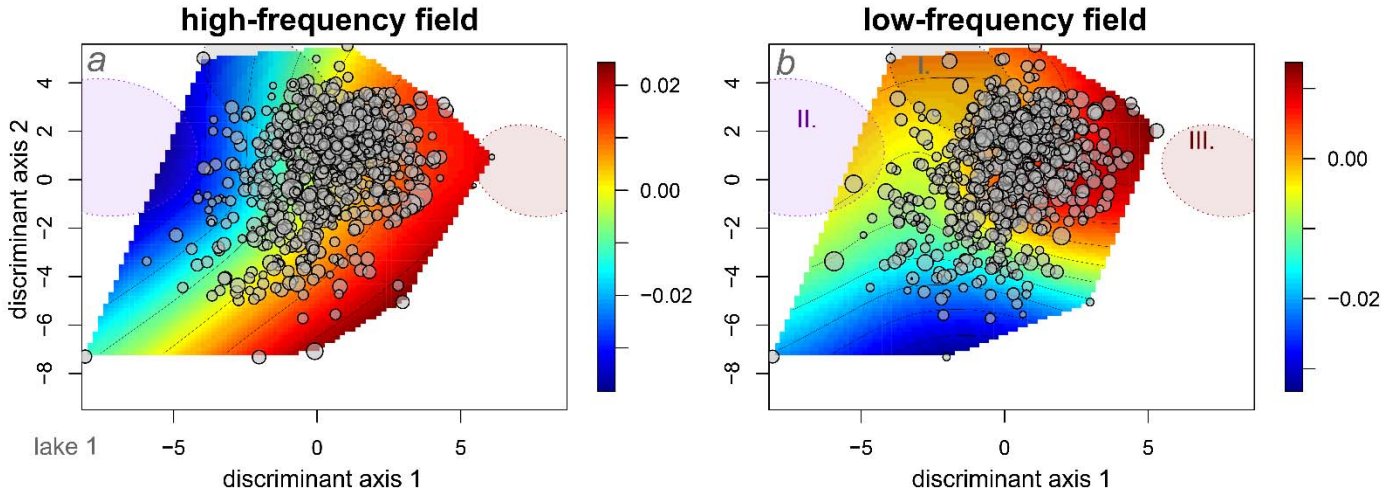



**Fig. S4 Growth rate fitness landscapes for a) high-frequency and b) low-frequency treatments in lake 1 (Crescent Pond).** Thin-plate splines predict the growth rate (heat color) across a single linear discriminant morphospace separating generalist and scale-eater phenotypes (x-axis: LD1) and generalist and molluscivore phenotypes (y-axis: LD2). Survivors in field enclosures are depicted in black relative to deaths over the 3-month exposure period. All hybrids are plotted within a shared linear discriminant morphospace calculated from lab-reared F1 individuals of parental populations in both lakes. 95% confidence ellipses indicate generalist (I. grey), molluscivore (II. purple), and scale-eater (III. red) regions of the morphospace.

**Fig. S5**

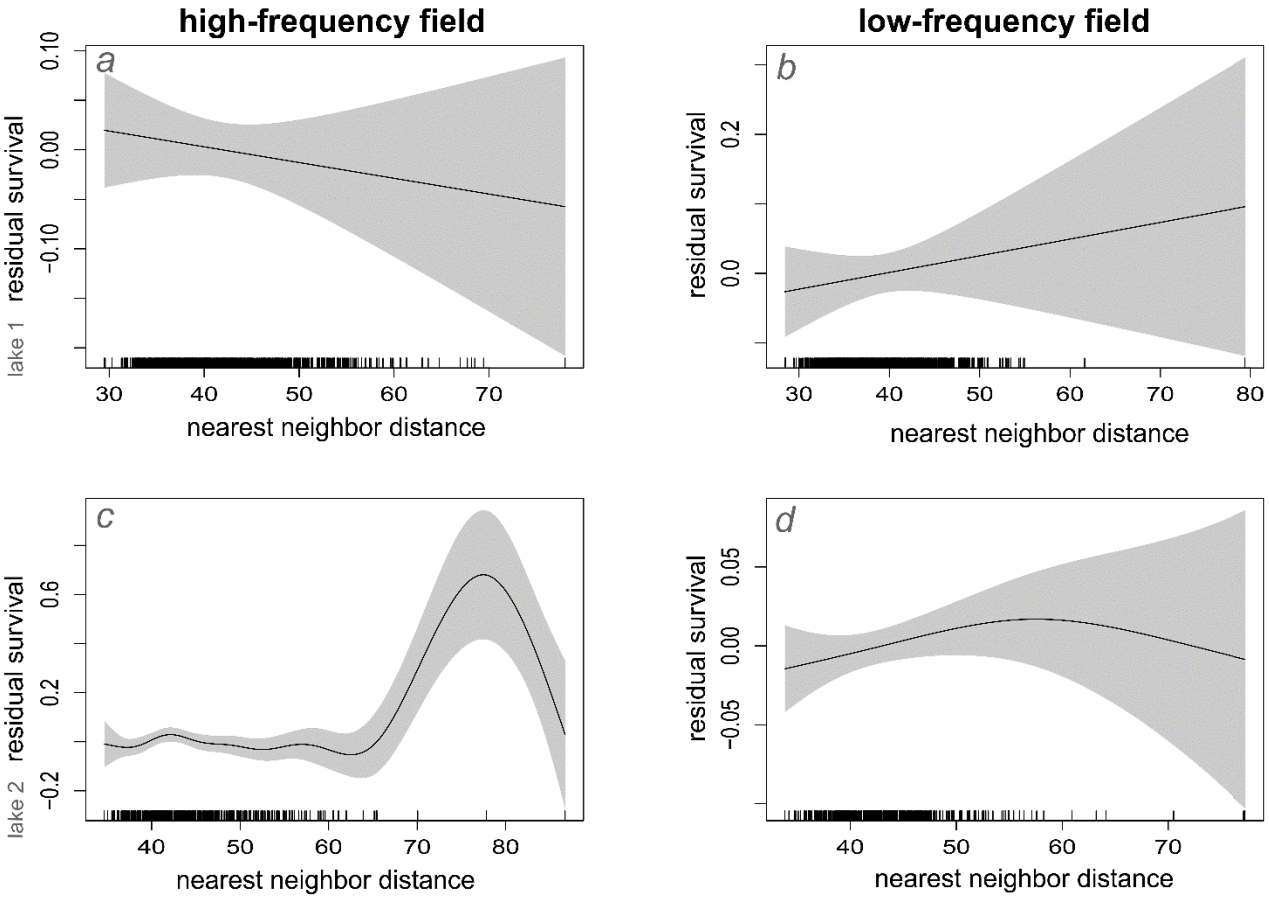



**Fig. S5 Residual survival probability relative to the density of similar hybrid phenotypes in high- (first column) and low-frequency (second column) field enclosures.** Residuals calculated from the excess survival probability not explained by the thin-plate splines estimated for hybrid phenotype in the discriminant morphospace (Fig. 3). The frequency of similar hybrid phenotypes was calculated for each hybrid from the nearest-neighbor Euclidean distances to the ten most similar hybrid phenotypes in the full morphospace.

**Fig. S6**

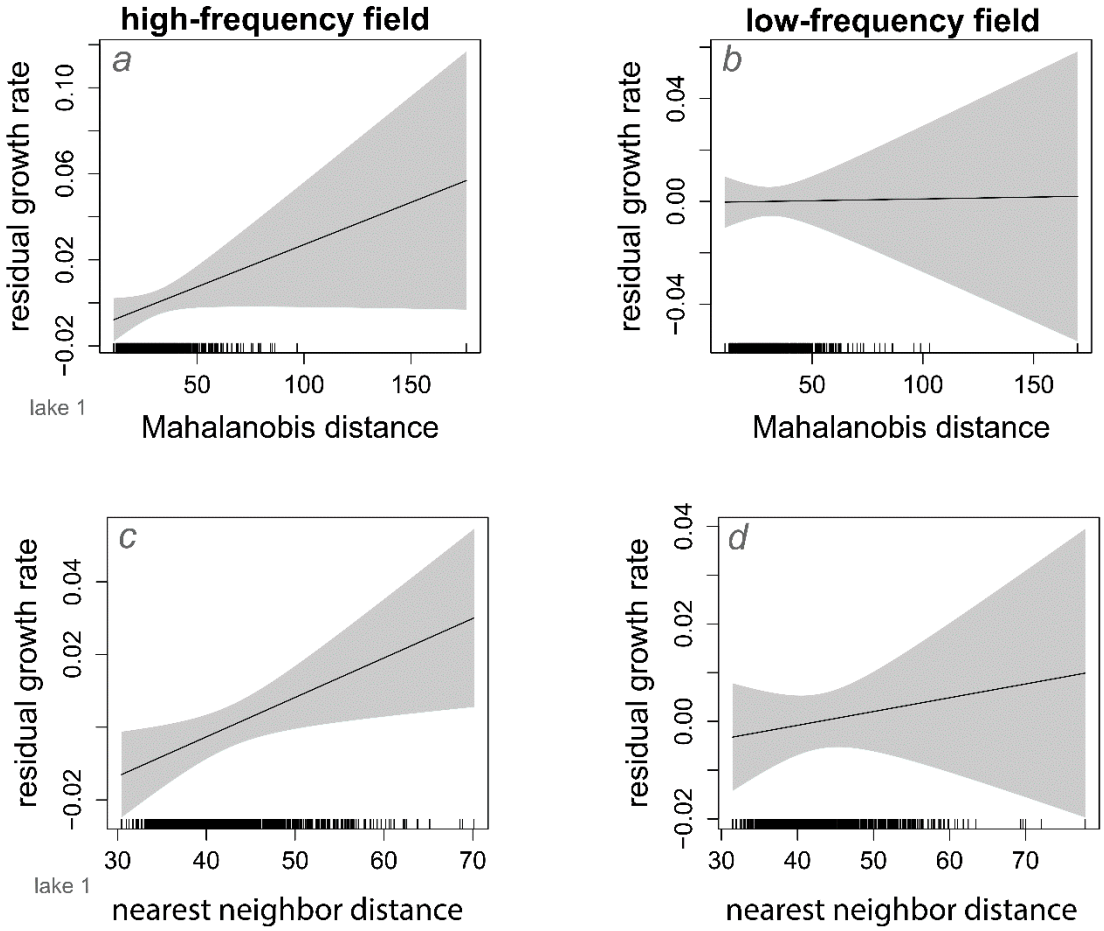



**Fig. S6 Residual growth rate relative to the density of similar hybrid phenotypes in high- (first column) and low-frequency (second column) field enclosures.** Residuals calculated from the growth rate variance not explained by the thin-plate splines estimated for hybrid phenotype in the linear discriminant morphospace (Fig. S5). The frequency of similar hybrid phenotypes was calculated for each hybrid from *a-b)* the Mahalanobis distance to the mean hybrid phenotype within each treatment and *c-d)* the sum of nearest-neighbor Euclidean distances to the ten most similar hybrid phenotypes in the full 30-trait morphospace.

**Fig. S7**

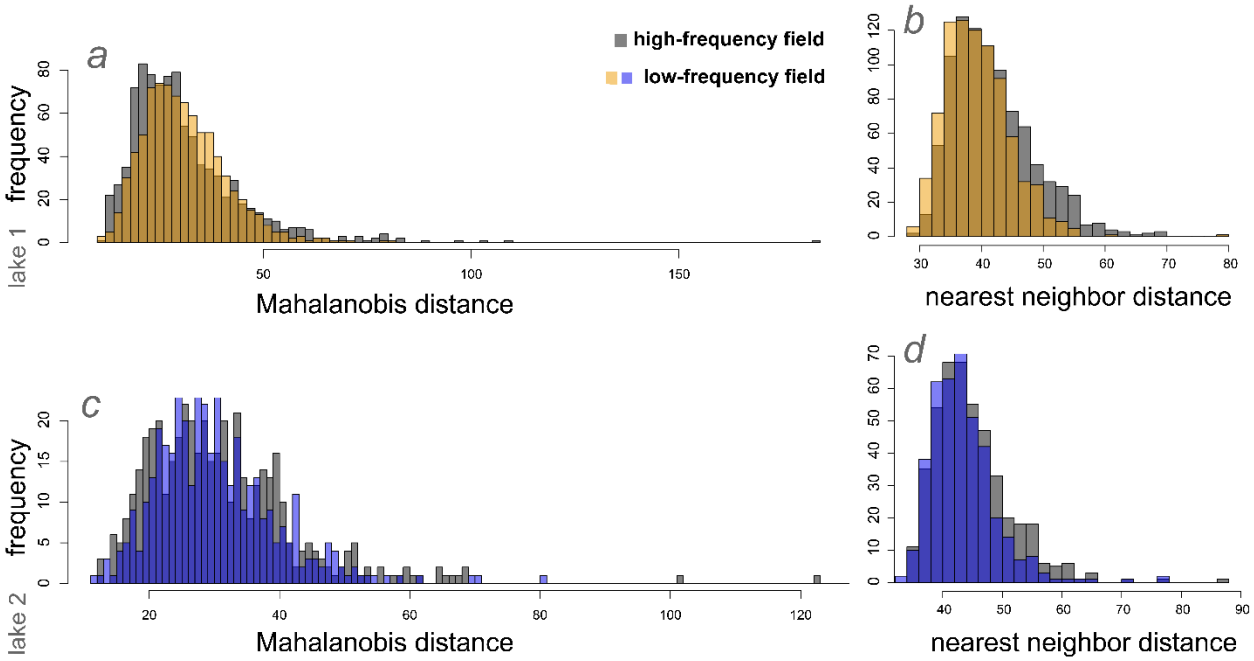



**Fig. S7 Histograms depicting the frequency of competitors** in high- (gray bars) and low-frequency (orange/blue) treatments in lake 1 (first row) and lake 2 (second row) for both Mahalanobis distance (first column) and nearest neighbor distance to the ten nearest phenotypes (second column).