

LETTER

Male competition fitness landscapes predict both forward and reverse speciation

Jason Keagy,* Liliana Lettieri and Janette W. Boughman

Department of Integrative Biology and BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, MI 48824, USA

*Correspondence: E-mail: keagy@msu.edu

Abstract

Speciation is facilitated when selection generates a rugged fitness landscape such that populations occupy different peaks separated by valleys. Competition for food resources is a strong ecological force that can generate such divergent selection. However, it is unclear whether intrasexual competition over resources that provide mating opportunities can generate rugged fitness landscapes that foster speciation. Here we use highly variable male F2 hybrids of benthic and limnetic threespine sticklebacks, *Gasterosteus aculeatus* Linnaeus, 1758, to quantify the male competition fitness landscape. We find that disruptive sexual selection generates two fitness peaks corresponding closely to the male phenotypes of the two parental species, favouring divergence. Most surprisingly, an additional region of high fitness favours novel hybrid phenotypes that correspond to those observed in a recent case of reverse speciation after anthropogenic disturbance. Our results reveal that sexual selection through male competition plays an integral role in both forward and reverse speciation.

Keywords

fitness landscape, fitness peak, male competition, multivariate, reverse speciation, sexual selection, speciation, species collapse, stickleback.

Ecology Letters (2016) 19: 71–80

INTRODUCTION

Rugged fitness landscapes with multiple peaks promote diversification, as populations evolve to occupy different peaks and intermediate individuals fall into fitness valleys, causing reproductive isolation between the diverging populations (Svensson & Calsbeek 2012). Although the fitness landscape metaphor has been valuable for guiding research on adaptive evolution and speciation (Svensson & Calsbeek 2012), fitness landscapes are typically conceived and quantified with respect to natural selection. Sexual selection fitness landscapes have rarely been investigated, especially in the context of speciation (Stelkens *et al.* 2008; Lackey & Boughman 2013a; Oh & Shaw 2013). Investigating fitness landscapes generated by sexual selection could address the currently unresolved role of sexual selection in the speciation process (West-Eberhard 1983; Panhuis *et al.* 2001; Ritchie 2007; Kraaijeveld *et al.* 2011; Rodríguez *et al.* 2013; Mendelson *et al.* 2014).

Ecological speciation research has focused on how ecological factors such as trade-offs produced by abiotic factors or diet generate divergent selection (Nosil 2012). In addition, competitive interactions over food can be critical in generating divergent selection favouring morphs specialising on different food resources, leaving maladapted hybrids at a competitive disadvantage (Schluter 1995; Nosil 2012; Martin & Wainwright 2013). However, competition for resources critical to reproduction is another important source of divergent selection (West-Eberhard 1983). Individuals compete over many resources that have large effects on reproductive fitness, such as territories, nests, nuptial gifts and display locations (Andersson 1994). When this strong intrasexual (often male-male) selection favours alternative phenotypic combinations, e.g. different morphs or reproductive strategies (Qvarnström *et al.*

2012), a rugged fitness landscape is created that could drive speciation forward, especially when a deep fitness valley exists between alternative fitness peaks. Negative frequency-dependent selection is the most often considered source of disruptive selection from male competition and is generated by aggression directed at individuals with similar phenotypes (van Doorn *et al.* 2004; Seehausen & Schluter 2004; Qvarnström *et al.* 2012). However, simply favouring particular phenotypic combinations over others could easily generate disruptive selection irrespective of morph frequency. Although male competition is ubiquitous (Andersson 1994) and theory suggests it could drive divergence (van Doorn *et al.* 2004), empirical study of its effects on speciation is scant and evidence is equivocal (Seehausen & Schluter 2004; Dijkstra *et al.* 2007; Qvarnström *et al.* 2012; Lackey & Boughman 2013a).

Typically as speciation progresses, successive barriers to gene flow evolve, increasing reproductive isolation. However, the speciation process can be quite dynamic and 'reverse speciation' can occur if reproductive barriers between formerly distinct species erode, causing the two species to collapse into one hybrid population (Seehausen 2006). Reverse speciation is typically explained by a shift in ecological selection (Seehausen 2006), such that selection against hybrids is reduced and the fitness landscape is flattened. However, total fitness is affected by multiple sources of selection, often acting in a sequence. For example, hybrids between benthic and limnetic sticklebacks have low fitness in nature because of poor feeding performance, and are unlikely to reach adulthood (Schluter 1995). Thus, even if certain types of hybrids were to have high fitness in a later life stage, this cannot be realised under normal circumstances. A novel hypothesis for how reverse speciation proceeds is that a reduction in ecological selection against hybrids allows them to reach

adulthood, and then sexual selection actually favours adult hybrid phenotypes.

Here, we test for male competition's role in both forward and reverse speciation in threespine sticklebacks, a classic model system for studying ecological speciation (McPhail 1994; Rundle *et al.* 2000). The role of natural selection (McPhail 1994; Schluter 1995) and female choice (Rundle *et al.* 2000; Boughman 2001) in promoting divergence between benthic and limnetic sticklebacks has been well studied, but little is known about male competition's role (Lackey & Boughman 2013a). For example, does male competition generate divergent selection, favouring alternate phenotypes and facilitating speciation? If yes, is this divergent selection synergistic with ecological selection? Does male competition influence reverse speciation? Male sticklebacks compete intensely for territories on which they build nests. Males also actively destroy each other's nests (van den Assem 1967), which are required for reproduction. Benthic and limnetic sticklebacks nest near each other within the range of female sampling (McPhail 1994), making them potential competitors (Lackey & Boughman 2013a). Thus the presence of a nest provides an excellent measure of male competitive fitness. The relationship between male phenotype and competitive fitness could have important effects on speciation dynamics.

Fitness landscapes quantify relationships between phenotypic (or genotypic) combinations and fitness (Svensson &

Calsbeek 2012). A complication in studying phenotypic fitness landscapes is that extant populations often reside on or near fitness peaks and have limited phenotypic variation in individual traits with strong intertrait correlations. Thus, extant populations cannot give information about the landscape beyond these peaks (Mitchell-Olds & Shaw 1987; Lexer *et al.* 2003). To circumvent this limitation, we take a powerful but underutilised approach using genetic crosses to create hybrids with novel and highly variable phenotypes (Schluter 1995; Schemske & Bradshaw 1999; Lexer *et al.* 2003; Stelkens *et al.* 2008; Martin & Wainwright 2013). Critically, this approach allowed us to quantify fitness for many intermediate and recombinational phenotypes not observed in pure species, permitting us to more completely visualise the fitness landscape.

We allowed F2 hybrid males to compete in naturalistic outdoor mesocosms at two different natural densities that varied the intensity of male competition. The mesocosms incorporated natural substrate, nesting material and food resources. We collected data on many male traits that are divergent between benthic and limnetic sticklebacks (Fig. 1) and recorded whether males successfully established and defended nests. We then quantified the multivariate fitness landscapes generated by male competition at both densities. In our analyses, we use the standard approach of defining peaks as local maxima within the bounds of our observed fitness landscapes (Martin & Wainwright 2013).

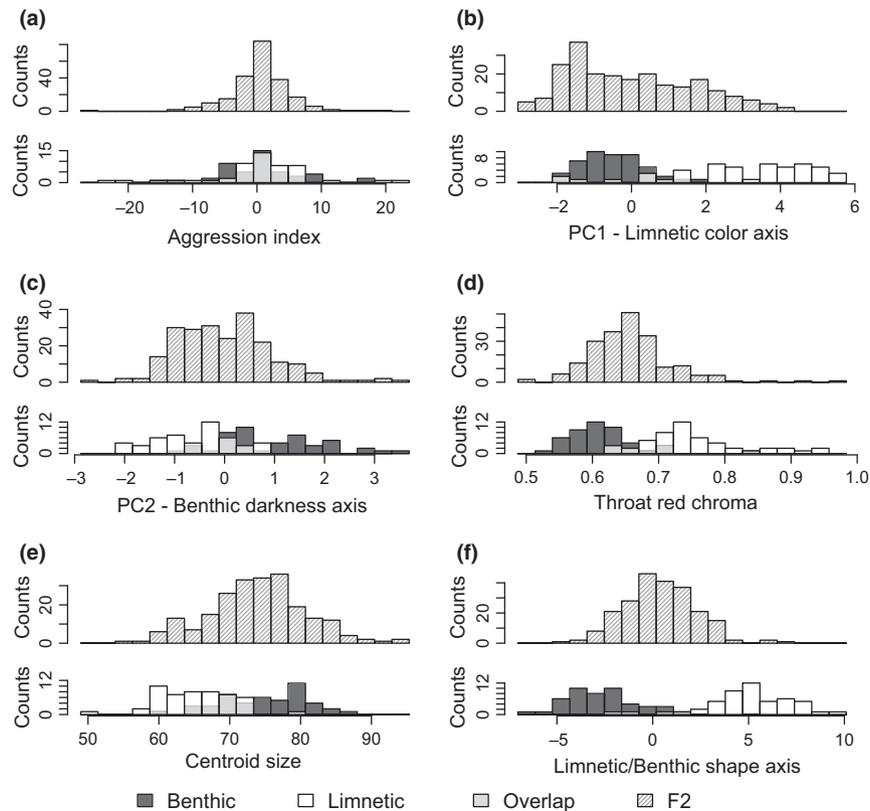


Figure 1 Trait distributions with F2 fish on the top half of each panel and pure species on the bottom. The pure species distributions are partially transparent to allow their full distributions to be visible, resulting in an intermediate shade of grey where they overlap. The two species, benthic and limnetic, differ significantly in all of these variables except for aggression index: (a) Aggression Index: $t_{92.65} = 0.16$, $P = 0.88$; (b) PC1 – Limnetic Color Axis: $t_{65.51} = -11.19$, $P < 2.2 \times 10^{-16}$; (c) PC2 – Benthic Darkness Axis: $t_{81.73} = 8.09$, $P = 4.79 \times 10^{-12}$; (d) Throat Red Chroma: $t_{86.61} = -10.29$, $P < 2.2 \times 10^{-16}$; (e) Centroid Size: $t_{90.23} = 8.45$, $P = 3.87 \times 10^{-13}$; (f) Limnetic/Benthic Shape Axis: $t_{91.18} = -18.46$, $P < 2.2 \times 10^{-16}$.

METHODS

Generating F2 hybrid fish

As part of a larger QTL project, we generated an F2 mapping population. We selected the parental benthic and limnetic populations based on their extreme phenotypic differences: Enos Lake benthics (Vancouver Island, British Columbia) resemble other benthics, but are also darker to the point of being black and Paxton Lake limnetics (Texada Island, British Columbia) are very colourful and slender. Using the most divergent populations is advantageous for quantifying fitness landscapes because it generates a larger range of hybrid phenotypes. Fish collected from Enos Lake and bred in our lab were those whose shape and colouration matched historical samples of benthic fish (McPhail 1994). F1 families were generated from every possible combination of 7 sets of benthic parents and 5 sets of limnetic parents; this cross design was used to retain segregating variation in the original outbred, natural populations and account for sex chromosome effects. We generated F2 families by using F1 brother-sister matings. This study used 72 F2 fish in 2011 and 153 F2 fish in 2012 for a total of 225 F2 fish. F2 fish had substantially reduced intertrait correlations (Tables S1 and S2) and displayed novel phenotypic combinations compared to pure species.

Pure benthic and limnetic fish

We also gathered phenotypic and competitive fitness data on groups of either pure benthic or pure limnetic fish, using the same methods as with F2 groups (described below). These data allowed us to plot pure species individuals in the same multivariate phenotypic spaces as F2 fish to see whether they occupied peaks predicted by the F2-generated fitness landscapes.

Since 2008, we have maintained lines generated from benthics caught in Enos Lake and limnetics caught in Paxton Lake. In 2012, we used 47 males from the benthic lines. Our limnetic lines were not large enough to use in this experiment; instead we used 50 Paxton Lake limnetics caught in 2013. Analysis of the phenotypic data (Fig. 1) clearly shows large differences between benthic and limnetic fish used in this study.

Field experiment

Adult fish were transported from our lab to Michigan State University's Kellogg Biological Station and housed there at 17–18°C on a 14 : 10 light : dark cycle. Prior to the experiment, we fed fish brine shrimp (*Artemia* spp.) and bloodworms (*Chironomus* spp.) twice daily.

A complete replicate consisted of 6 male fish in a high-density treatment and 3 male fish in a low-density treatment. Reproductive fish were selected randomly with the additional constraint that for each replicate no more than one fish was selected from a particular line. Replicates were composed of all F2 fish, all benthic fish, or all limnetic fish; tanks never contained mixtures of these categories.

Each male was placed in his own nesting tank (76 × 31 cm footprint, water depth: 28 cm). Each nesting tank had a

refuge built of 3 bricks, a substrate of fine gravel and sand, an artificial plant, and nest-building material (*Chara* spp. and grass). Starting 1 day later, for 2–3 days, a randomly selected gravid F2 female was placed in a clear jar in each nesting tank for 10 min daily to encourage reproductive behaviour. Using a series of randomly chosen highly phenotypically variable F2 females makes it unlikely that male reproductive behaviour was primed differentially based on phenotype. During their time in nesting tanks, 1/3 of each male's caudal fin was removed using a razor blade and stored in 95% ethanol for future DNA extraction and fish were marked with elastomer (Northwest Marine Technology, Shaw Island, WA, USA) in a unique colour/location pattern for identification.

After approximately 4 days, males were released into competitive arenas (Rubbermaid cattle stock tanks, diameter: 160 cm, water depth: 45 cm). Each tank had 3–4 cm of fine sand, 3 bricks arrayed equidistant from each other with an artificial plant on one side of each brick to provide structure and refuge, and substantial nest-building material. Zooplankton collected from nearby ponds provided the main food source (e.g. copepods, *Daphnia* spp., insect larvae). Once daily for the first 3 days, a gravid F2 female was placed in a clear jar for 15 min and then allowed to swim freely for an additional 15 min to entice males to nest. Starting on the fourth day post-release, we conducted 5 min focal follows of each male twice daily using JWatcher 1.0 (<http://www.jwatcher.ucla.edu/>), recording all aggressive interactions between the focal male and other males. We also recorded territorial boundaries and size (small, medium, or large) and nest locations. We added a gravid F2 female in a jar on the morning of the fifth and seventh days to maintain male reproductive status. On the eighth day post-release, we conducted female choice trials, which will be described in a separate paper. This schedule was extended to accommodate weather conditions or if gravid females were not available for female choice trials. Therefore, males generally had competitive behaviour scored twice on the fourth, fifth, sixth and seventh days post-release. Male F2 fish were in competitive arenas prior to female choice for a mean ± SD of 8.56 ± 2.30 days (benthics: 8.85 ± 1.61 days, limnetics: 9.18 ± 1.21 days). After female choice trials, males were colour scored, measured with a spectrometer, photographed, and weighed.

Phenotypic variables used in selection analyses

Aggression index

During male focal follows we recorded how often the focal male (1) approached other males at territorial boundaries, (2) bit other males, and (3) chased other males. We also recorded the number of times these behaviours were done to the focal male. We calculated a composite aggression index for each male using a custom script in R (v3.0.2, R Core Team 2013). For each focal observation, the script first weighted each interaction in terms of its aggressiveness (the order described above increases in aggressiveness). It then subtracted the sum of the weighted aggression received by the focal male from the sum of the weighted aggression given by the focal male. Finally, the script averaged this measure across all focal follows of that male for a final aggression index.

Colour scores

Using a 0–5 scale with 0.5 increments, we assessed (1) the area of the throat that was red, (2) the intensity of this red colouration, (3) the intensity of the blue eye colouration, (4) the intensity of the body colouration, and (5) the darkness of the body (Boughman 2001; Lewandowski & Boughman 2008). We performed a principal components analysis (PCA) of standardised colour score data as a variable reduction technique (Lande & Arnold 1983). The first 2 principal components (PCs) together explained 75% of the variance in colour score variables and were used in further analyses (Table S3). All variables except for body darkness loaded relatively highly and evenly on PC1. PC2 describes differences in body darkness. Therefore we refer to PC1 as the ‘limnetic colour axis’ (larger scores indicate fish that have colouration more like reproductive limnetic males, Fig. 1) and PC2 as the ‘benthic darkness axis’ (larger scores indicate fish that are darker like reproductive benthic males, Fig. 1).

Spectrometric measurement of throat colour

We used an S2000 spectrometer and PX-2 pulsed xenon light source (Ocean Optics, Dunedin, FL, USA) to measure reflectance of 3 areas on each fish. A custom-built holder maintained the probe at a consistent 45° angle to the fish. We used a Spectralon diffuse reflectance white standard. Using OOI-Base32 v2.0.1.4 (Ocean Optics, Dunedin, FL, USA), we took two measurements (each time moving the probe slightly) of each fish’s throat, dorsal and ventral surface. For this study, we focused on the throat because of its known importance in male competition (Bakker & Sevenster 1983; Baube 1997). We calculated ‘red chroma’, the proportion of light reflected between 575–700 nm relative to light reflected between 300–700 nm (Rick *et al.* 2011). This quantifies spectral saturation, or how ‘pure’ a colour is (Montgomerie 2006). We averaged the 2 measurements of red chroma. 12 fish in the high-density and 2 fish in the low-density treatments did not have spectrometric data because of technical problems and were excluded from fitness analyses.

Shape and size

We photographed each male’s left side using a digital camera (Canon G10; Canon U.S.A., Inc., Melville, NY, USA) inside a custom-built camera stand/light box. We digitised 22 landmarks used in previous studies (Taylor *et al.* 2006; Malek *et al.* 2012; Head *et al.* 2013; Lackey & Boughman 2013b; Fig. S6) on these photographs using tpsDig2 (Rohlf 2004). We also landmarked 27 limnetic and 76 benthic fish bred in our laboratory and not used in our experiment; the use of more benthic fish was due to their greater variability (e.g. limnetic/benthic shape axis scores, this study, Levene’s test: $F_{1,101} = 4.14$, $P = 0.045$). We aligned landmark x and y coordinates from all images using Generalised Procrustes Alignment (*gpa* function in the R package *geomorph* v1.1-0, Adams & Otárola-Castillo 2013). This removes non-shape variation (isometric aspects of size, position, and orientation). We then calculated PCs of the superimposed x and y coordinates. We reduced these 40 PCs to a manageable number of meaningful variables by using the PC scores of the 76 benthic and 27 limnetic lab-reared fish in a linear discriminant analy-

sis to produce an axis which best discriminated between these two groups. This linear discriminant axis correctly classified 41 benthics (87%) and 48 limnetics (96%) used in our experiment (Fig. S6). We projected PC scores of fish used in the male competition study onto this linear discriminant axis (referred to here as the ‘limnetic/benthic shape axis’); resulting scores that were larger describe males whose shape looks more ‘limnetic’. We estimated each male’s size by calculating the square root of the sum of squared distances between each landmark and the centroid (the centroid size).

Statistical analysis

Our data are available on the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.b786c>. All analyses were done in R (v3.0.2, R Core Team 2013). Colour scores were log transformed and red chroma was square root transformed to improve normality. In analyses with the high-density treatment, we only included tanks that had 5–6 fish throughout the trial.

Male competitive fitness was quantified as presence of a nest on the last day in the competitive arena. For this study we were interested in population-level selection on traits rather than selection relative to the mean within individual tanks (which only contained 3–6 fish). Therefore, we pooled data from all tanks within a treatment. This is analogous to quantifying how selection acts on an entire population in nature rather than a single small neighborhood.

All character traits were first standardised with a mean of zero and unit variance. Relative fitness was calculated as individual fitness divided by mean fitness. We estimated β , the vector of standardised linear selection gradients, from the partial regression coefficients of a multiple regression of relative fitness against the standardised trait values (Lande & Arnold 1983). We estimated γ , the matrix of standardised nonlinear selection gradients, using a full second-order polynomial regression (Lande & Arnold 1983). Nonlinear selection can be convex (stabilising if there is a local fitness maximum, Mitchell-Olds & Shaw 1987), concave (disruptive if there is a local fitness minimum, Mitchell-Olds & Shaw 1987), or correlational (particular trait combinations are favoured). We multiplied partial regression coefficients of quadratic terms by 2 to appropriately calculate quadratic selection gradients (Stinchcombe *et al.* 2008). While this technique makes no assumptions about departure from normality in the distribution of fitness, tests of significance can be impacted (Lande & Arnold 1983; Mitchell-Olds & Shaw 1987). In addition, because individuals are housed together and interact with each other in competitive arenas, they are not independent. We therefore calculated significance of selection gradients using a generalised linear mixed model with a binomial distribution and logit link function with tank as a random variable, using the *glmer* function in the R package *lme4* v1.0-5 (Bates *et al.* 2013).

Previously, univariate quadratic selection (e.g. stabilising or disruptive) was detected surprisingly infrequently in nature (Kingsolver *et al.* 2001). A likely explanation is that selection often acts on trait combinations and correlational selection is not often quantified, or when it is, the large number of resulting coefficients can make interpreting overall patterns of selection difficult (Blows & Brooks 2003). Canonical rotation of the γ

matrix (Phillips & Arnold 1989; Blows & Brooks 2003) can overcome these limitations. Canonical rotation of the γ matrix results in the matrix \mathbf{M} containing normalised eigenvectors (\mathbf{m}_i) describing the major axes of nonlinear selection. Analogous to PCA, elements of \mathbf{m}_i describe the relationship between traits and transformed axes, and eigenvalues of \mathbf{M} , λ_i , indicate the strength of nonlinear selection along these axes (negative = convex, positive = concave). With this rotation, there is no correlational selection between canonical axes. Instead, correlational selection is reflected by how traits load relative to one another on each independent axis. Canonical axis scores were calculated by projecting each fish's original trait values onto the new canonical axes (analogous to PC scores). We calculated linear selection estimates (the vector θ) and standard errors for both the linear and nonlinear components with a full second-order polynomial regression of relative fitness against the canonical axis scores (the 'double-regression' method, Phillips & Arnold 1989; Bisgaard & Ankenman 1996). We determined statistical significance of these parameters using a randomisation test where we permuted fitness 10 000 times against the canonical axis scores (Chenoweth *et al.* 2012; Oh & Shaw 2013). This specifically addressed whether the relationship between fitness and particular trait combinations identified by canonical rotation analysis was non-zero (Chenoweth *et al.* 2012; Oh & Shaw 2013). We visualised fitness surfaces of significant canonical axis pairs by performing thin-plate splines, a non-parametric regression approach (Phillips & Arnold 1989), using the *Tps* function in the R package *fields* v6.8 (Fields Development Team 2006).

RESULTS

Different densities resulted in biologically meaningful effects (Fig. S1). In particular, males in high density were involved in

more aggressive interactions ($F_{1,47.48} = 17.63$, $P < 0.001$, Fig. S1-B) and perhaps as a consequence, had lower growth rates ($F_{1,48.86} = 4.24$, $P = 0.045$, Fig. S1-D). Fish in high density also had smaller territory sizes ($F_{1,36.45} = 6.74$, $P = 0.013$, Fig. S1-F). Thus, we analyzed fitness landscapes separately for the two density treatments. There was strong positive directional selection favouring higher aggression in high density ($\beta_{\text{aggression}} = 0.278 \pm 0.073$, $P < 0.001$, Table S4). There was no significant directional selection in low density (Table S5).

The γ matrix indicated significant nonlinear selection at both densities (Tables S4 and S5). We conducted a canonical rotation of the γ matrix to aid interpretation of multivariate nonlinear selection (Phillips & Arnold 1989; Blows & Brooks 2003). At high density, four multivariate selection axes had significant curvature (Table 1, Fig. 2, Fig. S2). Axis \mathbf{m}_1 produces strong disruptive selection favouring either all benthic or all limnetic multivariate phenotypes. Axis \mathbf{m}_6 is strongly stabilising and selects against both very colourful benthic-shaped fish and drab limnetic-shaped fish (Fig. S3). Thus, the two strongest axes of nonlinear selection (\mathbf{m}_1 and \mathbf{m}_6) in conjunction create fitness peaks corresponding to pure benthic or limnetic phenotypes. This is supported by overlap of the peaks by the 95% confidence ellipses of independently gathered pure species data (Fig. 2a). However, the next strongest axis, \mathbf{m}_5 , favours novel hybrid combinations of aggression, limnetic colour, benthic darkness, and limnetic shape. While limnetic fish tend to have phenotypes that predict high fitness on this axis, hybrids that are aggressive and dark (like benthic fish) but limnetic-like in shape and colour are most favoured. The final significant (but weakest) axis, \mathbf{m}_4 , favours individuals with intermediate to high aggression and red chroma, and intermediate to low size; phenotypes similar to limnetic fish.

Table 1 Canonical rotation of the standardised nonlinear selection gradient matrix

	M						θ_i	λ_i
	Aggression index	Limnetic colour axis	Benthic darkness axis	Throat red chroma	Centroid size	Limnetic/ Benthic shape axis		
Species diff	L = B	L > B	B > L	L > B	B > L	L > B		
High density ($n = 120$)								
\mathbf{m}_1	0.510	-0.626	0.356	-0.112	0.322	-0.326	-0.020	0.364**
\mathbf{m}_2	-0.351	0.012	0.828	0.067	-0.419	-0.104	-0.048	0.236
\mathbf{m}_3	0.407	0.018	-0.048	-0.667	-0.591	0.195	0.056	-0.040
\mathbf{m}_4	-0.381	0.157	0.180	-0.702	0.546	0.080	-0.178*	-0.219†
\mathbf{m}_5	-0.506	-0.476	-0.391	-0.200	-0.270	-0.502	-0.127	-0.299*
\mathbf{m}_6	-0.226	-0.597	0.002	0.074	-0.004	0.766	-0.158†	-0.468***
Low density ($n = 79$)								
\mathbf{m}_1	-0.062	-0.615	0.703	-0.140	-0.285	-0.154	0.074	0.426†
\mathbf{m}_2	0.669	0.465	0.238	-0.111	-0.274	-0.439	0.091	0.136
\mathbf{m}_3	-0.272	0.110	-0.289	-0.652	-0.628	0.103	-0.238†	0.111
\mathbf{m}_4	-0.403	-0.058	-0.268	0.327	-0.181	-0.789	0.047	0.047
\mathbf{m}_5	0.552	-0.624	-0.542	-0.069	-0.039	-0.072	-0.352*	-0.284
\mathbf{m}_6	0.088	0.021	-0.021	0.657	-0.644	0.381	0.198	-0.701**

† $P < 0.10$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Elements of matrix \mathbf{M} describe loadings of the original variables on the transformed axes (\mathbf{m}_i), which represent independent axes of nonlinear selection. The direction of species differences in each trait is given to aid interpretation (L = limnetic, B = benthic). Linear selection gradients (θ_i), and quadratic selection gradients (λ_i , the eigenvalue of each vector) on axes are shown; significance was determined through randomisation tests (10 000 permutations). Quadratic components were marginally significant for high-density axis \mathbf{m}_4 ($P = 0.054$) and low-density axis \mathbf{m}_1 ($P = 0.051$).

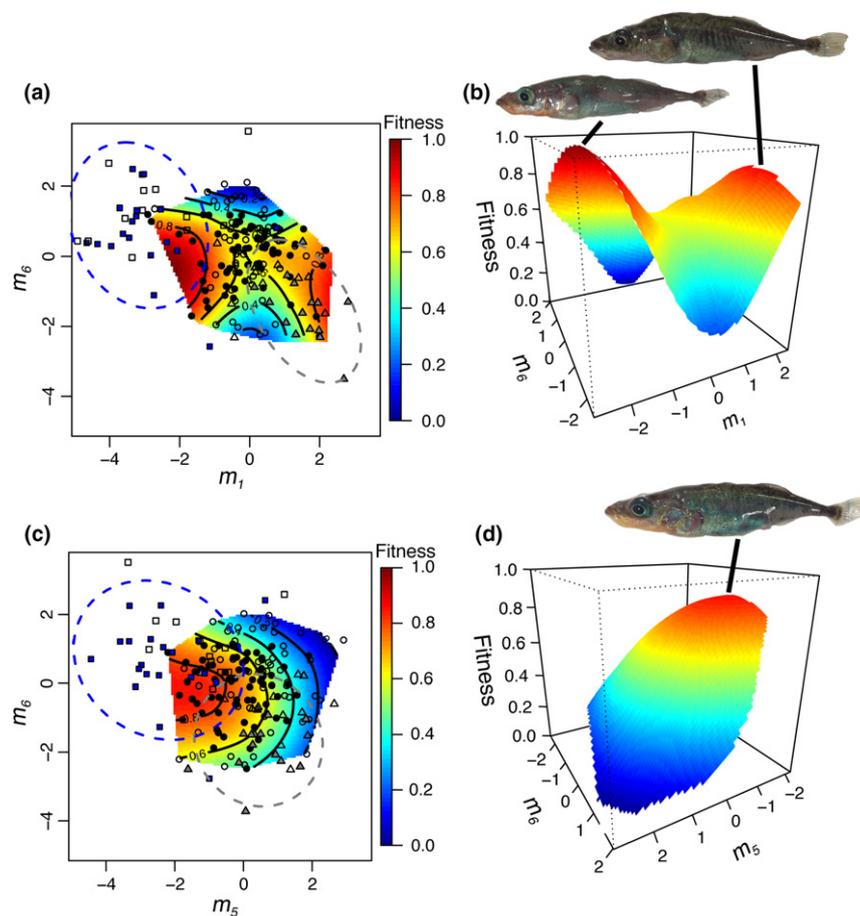


Figure 2 Fitness surfaces generated from canonical rotation of high-density γ matrix. These are non-parametric thin-plate spline approximations of the relationship between two canonical axes and absolute fitness (0 = no nest, 1 = nest). Photographs are representative F2 hybrid fish from each peak. (a and c) Contour plots of the fitness surfaces generated by F2 hybrid males (circles). Squares represent limnetic males and triangles represent benthic males, which were not used to generate the fitness surfaces, but are plotted to assist interpretation. Dashed lines indicate 95% confidence ellipses for the phenotypes of pure species fish that had nests. Closed symbols indicate individuals that had nests and open symbols those that did not. (b and d) 3-D representations of the fitness surfaces shown in (a and c), rotated to increase visibility of features. Animated rotating versions of these fitness surfaces are included in Video S1.

At low density, canonical rotation analysis indicated two multivariate axes with significant curvature and one with significant linear selection (Table 1, Fig. 3, Fig. S4). Similar to high density, a disruptive selection axis, m_1 , selects for individuals with either benthic or limnetic colouration, promoting species differences. A significant stabilising selection axis, m_6 , favours novel hybrid combinations of red chroma, size, and shape. Thus, compared to high density, peaks formed by the joint action of axes m_1 and m_6 in low density are not as well aligned with pure species phenotypic distributions. The axis with significant linear selection, m_5 , favours both limnetic colour and benthic darkness as well as low aggression, favouring particular hybrids.

At both densities, axis m_5 produces a high fitness region indicating positive correlational selection for limnetic colouration and benthic darkness, favouring hybrids. This colour combination surprisingly matches the unique colouration of reproductive males currently found in Enos Lake, in which a formerly distinct species pair (McPhail 1994) has recently and rapidly collapsed into a hybrid swarm following anthro-

pogenic disturbance (Taylor *et al.* 2006). Previously in this lake, limnetic males were colourful and benthic males nearly black, occupying opposite ends of the limnetic colour and benthic darkness axes (McPhail 1994). After introduction of crayfish, sexual isolation and ecological selection against hybrids was weakened, resulting in the creation of many hybrids who survived to adulthood (Behm *et al.* 2010; Lackey & Boughman 2013b). Contrasting with the historical colour pattern and with the pure species used in this experiment, hybrid males recently collected from Enos Lake (Malek *et al.* 2012) were both colourful and dark, suggesting they have evolved to occupy the novel high fitness region predicted by axis m_5 (Fig. 4, Fig. S5).

We initially visualised fitness landscapes by plotting pairs of canonical axes, as suggested by Phillips & Arnold (1989). This approach has been standard practice when utilising the canonical rotation method (Blows & Brooks 2003; Chenoweth *et al.* 2012; Oh & Shaw 2013). However, fitness is best predicted by a model that includes all axes with non-zero linear or quadratic components. Phillips & Arnold (1989) suggest an alternative for

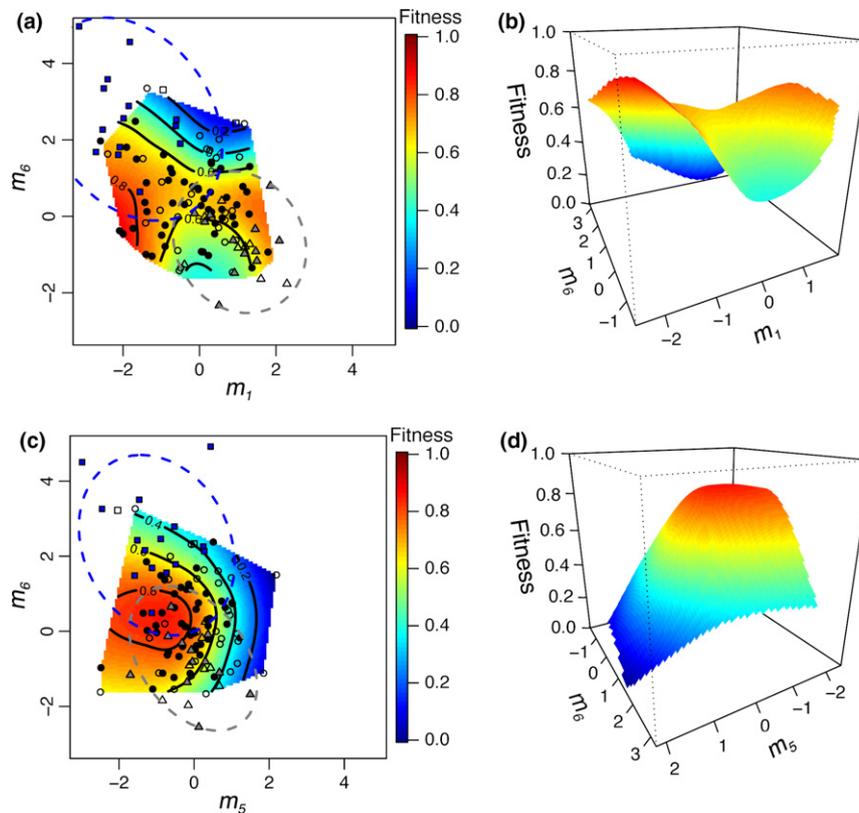


Figure 3 Fitness surfaces generated from canonical rotation of low-density γ matrix. These are non-parametric thin-plate spline approximations of the relationship between two canonical axes and absolute fitness (0 = no nest, 1 = nest). (a and c) Contour plots of the fitness surfaces generated by F2 hybrid males (circles). Squares represent limnetic males and triangles represent benthic males, which were not used to generate the fitness surfaces, but are plotted to assist interpretation. Dashed lines indicate 95% confidence ellipses for the phenotypes of pure species fish that had nests. Closed symbols indicate individuals that had nests and open symbols indicate individuals that did not. (b and d) 3-D representations of the fitness surfaces shown in (a and c), rotated as in Fig. 2. Animated rotating versions of these fitness surfaces are included in Video S1.

visualising fitness effects of three canonical axes at the same time: contour representations of four-dimensional selection surfaces, with fitness as the fourth axis. Plotting the predicted fitness values for all combinations of m_1 , m_5 , and m_6 yields novel insights about the male competition fitness landscapes (Fig. 5). This visualisation shows two peaks formed primarily by m_1 and m_6 that correspond roughly to limnetic and benthic multivariate phenotypes, especially in high density. In addition, m_5 creates another high fitness region that bridges these two peaks, encouraging hybridisation (Fig. 5a and c). That m_5 creates this bridge can be verified by eliminating the influence of m_5 on fitness (i.e., setting the linear and quadratic coefficients of m_5 to zero), which removes the high fitness region between the peaks (Fig. 5b and d). This effect is weaker in the low-density landscape because it is flatter than the high-density landscape, as expected with reduced male competition. The plots in Figs 2 and 3 are roughly equivalent to averaging fitness in Fig. 5 across whichever axis is not plotted.

In both densities there was a disruptive selection axis (m_1) promoting speciation. An alternative explanation for this disruptive selection is that intermediate genotypic combinations caused behavioural dysfunction. If this ‘intrinsic inviability’ hypothesis was correct, fish that were unable to build nests in nesting tanks, where there was no competition, would have m_1 scores corresponding to lower fitness in competitive arenas.

However, we find no evidence of this pattern (high density: $t_{108} = -1.56$, $P = 0.12$, low density: $t_{68} = -0.07$, $P = 0.94$). The ‘intrinsic inviability’ hypothesis also cannot explain axes of selection that promote hybridisation by conferring high fitness on males with novel trait combinations (e.g. m_5).

Loadings on canonical axes were often different between low and high density (Table 1). In fact, we found marginal support for the hypothesis that the fitness landscapes are different in form ($P = 0.1$, Supporting Information Supplemental Methods). Moreover, fitness was predicted to be on average higher in low density than high density ($t_{197} = 3.92$, $P = 0.0001$), supporting the observation that the low-density fitness landscape has wider areas of high fitness (e.g. compare Fig. 5a and c).

The prevailing explanation for how male competition might generate multiple fitness peaks and foster speciation is negative frequency-dependent selection (Seehausen & Schluter 2004; Qvarnström *et al.* 2012). This hypothesis predicts that males with rare phenotypes should have high fitness. To test this we quantified ‘rareness’ by calculating each male’s average Euclidean distance from the other males in their tank. Male rareness did not predict fitness (GLM, high-density: $\beta = -4.02 \pm 2.29$, $P = 0.08$; low-density: $\beta = -3.23 \pm 2.36$, $P = 0.17$); in fact rareness was selected against with both densities combined ($\beta = -4.29 \pm 1.58$, $P = 0.007$).

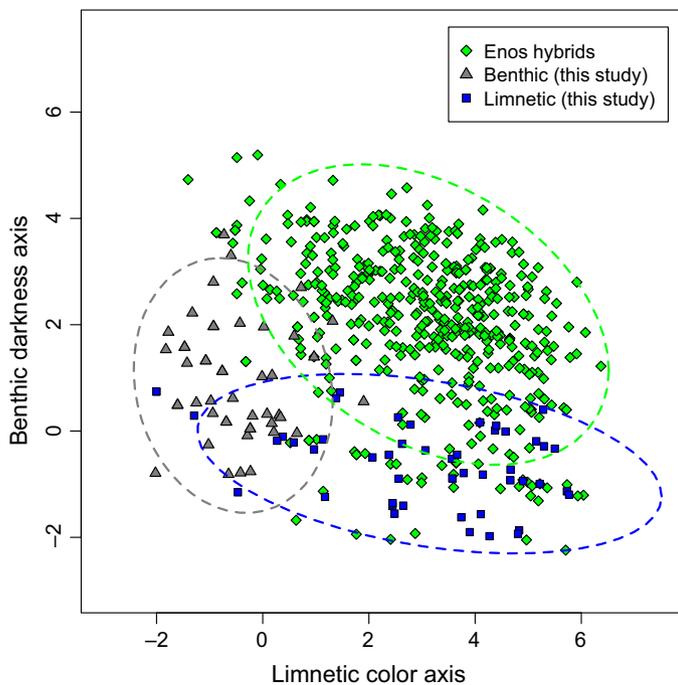


Figure 4 Distribution of nuptial colour for Enos Lake hybrid swarm compared to pure species. Naturally generated hybrid sticklebacks from Enos Lake (Malek *et al.* 2012) occupy a region in phenotypic space that maximises both limnetic colour and benthic darkness, which is not expressed by either pure species. This pattern is consistent with a response to correlational selection on limnetic colour and benthic darkness predicted by m_s in both densities. Dashed lines indicate 95% confidence ellipses. The differences between groups are statistically significant (MANOVA, Pillai: 0.46, approximate $F_{4,1112}$: 82.364, $P < 2.2 \times 10^{-16}$).

DISCUSSION

We assessed the competitive fitness of male stickleback multivariate phenotypes, including those not normally found in nature, using highly variable lab-generated F2 hybrids of benthic and limnetic sticklebacks. Sexual selection via male competition can be at least as fierce as competition over food, and leads to intense selection (Andersson 1994). At both densities, canonical rotation analysis revealed a complex fitness landscape with multiple peaks.

The correspondence between multivariate phenotypes favoured by disruptive selection (m_1 in both densities) and those of pure benthic and limnetic sticklebacks is striking. That male competition alone can predict adaptive phenotypic divergence in a classic example of ecological speciation suggests male competition is an important, albeit underappreciated, force in speciation. Surprisingly, in both density treatments, the fitness landscape also has a significant axis of selection that favours mismatched trait combinations seen in natural hybrids in a recent case of reverse speciation following anthropogenic disturbance. This suggests that male competition can also promote reverse speciation, at least temporarily, when other barriers to production of hybrid offspring break down.

Negative frequency-dependent selection is the prevailing explanation for how male competition might generate multiple fitness peaks (Seehausen & Schluter 2004; Qvarnström

et al. 2012). In contrast to one of its primary predictions, our data suggests that rare males were selected against. Thus, it was not rareness, *per se*, that gave males an advantage, but rather specific trait combinations. This finding also suggests male competition can generate divergent selection through mechanisms other than negative frequency-dependent selection. This broadens the conditions under which male competition might contribute to divergent adaptation and reproductive isolation.

Most theoretical and empirical studies of male competition and speciation have focused on single traits (e.g. colour or size), but this approach likely underestimates the strength of nonlinear selection (Blows & Brooks 2003). We did not detect disruptive selection on any single trait (Table S6). Selection in nature almost certainly acts on multiple interacting traits; the lack of univariate disruptive selection compared to the strong multivariate effects we found highlights the importance of analysing selection in a multivariate manner.

The fact that we detected divergent selection through male competition alone emphasises its importance to the speciation process. However, male competition is obviously not the only agent of selection affecting fitness of individual sticklebacks. For example, experimental and correlational evidence strongly establishes the importance of ecology in driving divergence between limnetic and benthic sticklebacks, including in traits we measured, such as size and shape (McPhail 1994; Head *et al.* 2013). Many of these traits have also been implicated in female choice (size and shape: Boughman *et al.* 2005; Conte & Schluter 2013; Head *et al.* 2013; colour: Boughman 2001; Boughman *et al.* 2005). Interestingly, the divergent selection from male competition we detected is synergistic with other sources of selection quantified in past studies. This consistent selection by multiple selective agents could explain the rapid and repeated speciation observed in sticklebacks. In addition, the complexity of the male competition fitness landscapes we quantified emphasises the fragility of reproductive barriers. When other sources of selection against hybridisation crumble, male competition fitness landscapes provide high fitness refuges for hybrids that further erode species differences. However, pure species phenotypes correspond to even higher fitness regions on these landscapes and so such reversals in reproductive isolation may be temporary.

The male competition fitness landscapes we quantified indicate strong selection. The difference between the predicted maximum and minimum fitness values is nearly five times that of a recent study that quantified a complex ecological fitness landscape promoting speciation in pupfishes (Martin & Wainwright 2013). Such strong multivariate sexual selection parallels previous findings that directional sexual selection on single traits can be stronger than natural selection on single traits (Kingsolver *et al.* 2001). Our results suggest that male competition has played a pivotal and unappreciated role in the dynamic speciation process. It likely helped push speciation forward by favouring two different male phenotypic combinations characterising pure species. It has also likely helped reverse speciation by favouring novel phenotypic combinations expressed by hybrids.

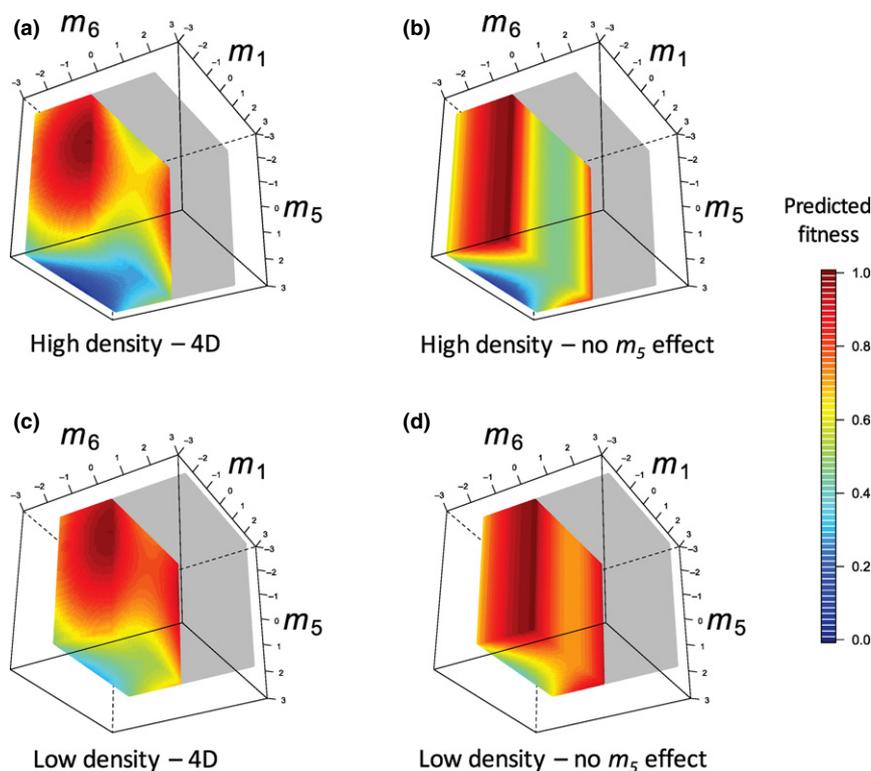


Figure 5 Further examination of multivariate fitness landscapes. These 4-D contour plots represent the predicted fitness (probability of holding a nest) for each combination of m_1 , m_5 and m_6 , with fitness as the fourth dimension represented by colour contours. Only one half of the resulting cube is displayed to allow better visualisation of the landscape features, with the remainder indicated by grey. The full visualisation can be seen built layer by layer in Video S2. This visualisation clearly shows two peaks formed primarily by each density's respective m_1 and m_6 axes. These peaks correspond roughly to limnetic and benthic multivariate phenotypes. In addition, m_5 creates another high fitness region that bridges these two peaks. This bridge could encourage hybridisation as hybrids with the corresponding phenotypes would have relatively high fitness. Further evidence that m_5 creates this bridge is seen by eliminating the influence of m_5 on fitness, which removes the high fitness region between the peaks (compare a to b and c to d). In b and d, because fitness is no longer affected by m_5 , the plot is identical for all values of m_5 . This visualisation also highlights qualitative differences between the two densities in their respective fitness landscapes such as steeper peaks in high density and higher average fitness in low density due to reduced competition (compare a to c).

ACKNOWLEDGEMENTS

This research was supported with funding to JWB by a CAREER grant from the National Science Foundation (DEB-0952659) and by the BEACON Center for the Study of Evolution in Action (NSF DBI-0939454). The authors have no conflicts of interest. Research was conducted under permits from the Ministry of the Environment, BC and approval of the Institutional Animal Care and Use Committee of Michigan State University. Many undergraduates helped generate hybrids and collect data, especially J. Blais, K. Keller, P. Lansky, W. Lehto, A. Marchand, N. Mark, J. Martinez, R. Minter, and R. Walquist. I. Cooper, A. Lackey and other members of the Boughman laboratory gave additional support. W. Pitchers provided valuable statistical advice. R. Lenski and three anonymous reviewers provided comments on the manuscript.

AUTHOR CONTRIBUTIONS

JWB conceived and secured funding for the project. JWB, JK, and LL generated hybrid fish and collected data. JK performed additional data processing. JK performed all statistical analyses. JK and JWB wrote the manuscript.

REFERENCES

- Adams, D.C. & Otárola-Castillo, E. (2013). *geomorph*: an R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.*, 4, 393–399.
- Andersson, M. (1994). *Sexual Selection*. Princeton University Press, Princeton.
- van den Assem, J. (1967). Territory in the three-spined stickleback *Gasterosteus aculeatus* L.: an experimental study in intra-specific competition. *Behaviour, Supplement* 16, III–164.
- Bakker, T.C.M. & Sevenster, P. (1983). Determinants of dominance in male sticklebacks (*Gasterosteus aculeatus* L.). *Behaviour*, 86, 55–71.
- Bates, D., Maechler, M., Bolker, B. & Walker, S. (2013). Package “*lme4*.” Linear mixed-effects models using Eigen and S4. Available at: <https://github.com/lme4/lme4/>. Last accessed 25 October 2013.
- Baube, C.L. (1997). Manipulations of signalling environment affect male competitive success in three-spined sticklebacks. *Anim. Behav.*, 53, 819–833.
- Behm, J.E., Ives, A.R. & Boughman, J.W. (2010). Breakdown in postmating isolation and the collapse of a species pair through hybridization. *Am. Nat.*, 175, 11–26.
- Bisgaard, S. & Ankenman, B. (1996). Standard errors for the eigenvalues in second-order response surface models. *Technometrics*, 38, 238–246.
- Blows, M.W. & Brooks, R. (2003). Measuring nonlinear selection. *Am. Nat.*, 162, 815–820.
- Boughman, J.W. (2001). Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature*, 411, 944–948.

- Boughman, J.W., Rundle, H.D. & Schluter, D. (2005). Parallel evolution of sexual isolation in sticklebacks. *Evolution*, 59, 361–373.
- Chenoweth, S.F., Hunt, J. & Rundle, H.D. (2012). Analyzing and comparing the geometry of individual fitness surfaces. In: *The Adaptive Landscape in Evolutionary Biology* (eds Svensson, E.I., Calsbeek, R.). Oxford University Press, Oxford, pp. 126–149.
- Conte, G.L. & Schluter, D. (2013). Experimental confirmation that body size determines mate preference via phenotype matching in a stickleback species pair. *Evolution*, 67, 1477–1484.
- Dijkstra, P.D., Seehausen, O., Pierotti, M.E.R. & Groothuis, T.G.G. (2007). Male-male competition and speciation: aggression bias towards differently coloured rivals varies between stages of speciation in a Lake Victoria cichlid species complex. *J. Evol. Biol.*, 20, 496–502.
- van Doorn, G.S., Dieckmann, U. & Weissing, F.J. (2004). Sympatric speciation by sexual selection: a critical reevaluation. *Am. Nat.*, 163, 709–725.
- Fields Development Team. (2006). *fields: Tools for Spatial Data*. National Center for Atmospheric Research, Boulder, CO. Available at: <http://www.cgd.ucar.edu/Software/Fields>. Last accessed 19 July 2013.
- Head, M.L., Kozak, G.M. & Boughman, J.W. (2013). Female mate preferences for male body size and shape promote sexual isolation in threespine sticklebacks. *Ecol. Evol.*, 3, 2183–2196.
- Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E. et al. (2001). The strength of phenotypic selection in natural populations. *Am. Nat.*, 157, 245–261.
- Kraaijeveld, K., Kraaijeveld-Smit, F.J.L. & Maan, M.E. (2011). Sexual selection and speciation: the comparative evidence revisited. *Biol. Rev.*, 86, 367–377.
- Lackey, A.C.R. & Boughman, J.W. (2013a). Divergent sexual selection via male competition: ecology is key. *J. Evol. Biol.*, 26, 1611–1624.
- Lackey, A.C.R. & Boughman, J.W. (2013b). Loss of sexual isolation in a hybridizing stickleback species pair. *Curr. Zool.*, 59, 591–603.
- Lande, R. & Arnold, S.J. (1983). The measurement of selection on correlated characters. *Evolution*, 37, 1210–1226.
- Lewandowski, E. & Boughman, J. (2008). Effects of genetics and light environment on colour expression in threespine sticklebacks. *Biol. J. Linn. Soc.*, 94, 663–673.
- Lexer, C., Randell, R.A. & Rieseberg, L.H. (2003). Experimental hybridization as a tool for studying selection in the wild. *Ecology*, 84, 1688–1699.
- Malek, T.B., Boughman, J.W., Dworkin, I. & Peichel, C.L. (2012). Admixture mapping of male nuptial colour and body shape in a recently formed hybrid population of threespine stickleback. *Mol. Ecol.*, 21, 5265–5279.
- Martin, C.H. & Wainwright, P.C. (2013). Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. *Science*, 339, 208–211.
- McPhail, J.D. (1994). Speciation and the evolution of reproductive isolation in the sticklebacks (*Gasterosteus*) of south-western British Columbia. In: *The Evolutionary Biology of the Threespine Stickleback* (eds Bell, M.A., Foster, S.A.). Oxford University Press, Oxford, pp. 399–437.
- Mendelson, T.C., Martin, M.D. & Flaxman, S.M. (2014). Mutation-order divergence by sexual selection: diversification of sexual signals in similar environments as a first step in speciation. *Ecol. Lett.*, 17, 1053–1066. doi:10.1111/ele.12313.
- Mitchell-Olds, T. & Shaw, R.G. (1987). Regression analysis of natural selection: statistical inference and biological interpretation. *Evolution*, 41, 1149–1161.
- Montgomerie, R. (2006). Analyzing colors. In: *Bird Coloration: Mechanisms and Measurements* (eds Hill, G.E. & McGraw, K.). Cambridge University Press, Cambridge, pp. 90–147.
- Nosil, P. (2012). *Ecological Speciation*. Oxford University Press, Oxford.
- Oh, K.P. & Shaw, K.L. (2013). Multivariate sexual selection in a rapidly evolving speciation phenotype. *Proc. R. Soc. B*, 280, 20130482.
- Panhuis, T.M., Butlin, R., Zuk, M. & Tregenza, T. (2001). Sexual selection and speciation. *Trends Ecol. Evol.*, 16, 364–371.
- Phillips, P.C. & Arnold, S.J. (1989). Visualizing multivariate selection. *Evolution*, 43, 1209–1222.
- Qvarnström, A., Vallin, N. & Rudh, A. (2012). The role of male contest competition over mates in speciation. *Curr. Zool.*, 58, 493–509.
- R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: <http://www.R-project.org/> Last accessed 25 September 2013.
- Rick, I.P., Mehli, M. & Bakker, T.C.M. (2011). Male red ornamentation is associated with female red sensitivity in sticklebacks. *PLoS ONE*, 6, e25554.
- Ritchie, M.G. (2007). Sexual selection and speciation. *Annu. Rev. Ecol. Evol. Syst.*, 38, 79–102.
- Rodríguez, R.L., Boughman, J.W., Gray, D.A., Hebets, E.A., Höbel, G. & Symes, L.B. (2013). Diversification under sexual selection: the relative roles of mate preference strength and the degree of divergence in mate preferences. *Ecol. Lett.*, 16, 964–974.
- Rohlf, F.J. (2004). tpsDig, Digitize Landmarks and Outlines, version 2.0. Available at: <http://life.bio.sunysb.edu/morph>. Last accessed date 31 August 2011.
- Rundle, H.D., Nagel, L., Boughman, J.W. & Schluter, D. (2000). Natural selection and parallel speciation in sympatric sticklebacks. *Science*, 287, 306–308.
- Schemske, D.W. & Bradshaw, H.D. (1999). Pollinator preference and the evolution of floral traits in monkeyflowers (*Mimulus*). *PNAS*, 96, 11910–11915.
- Schluter, D. (1995). Adaptive radiation in sticklebacks: trade-offs in feeding performance and growth. *Ecology*, 76, 82–90.
- Seehausen, O. (2006). Conservation: losing biodiversity by reverse speciation. *Curr. Biol.*, 16, R334–R337.
- Seehausen, O. & Schluter, D. (2004). Male-male competition and nuptial-colour displacement as a diversifying force in Lake Victoria cichlid fishes. *Proc. R. Soc. B*, 271, 1345–1353.
- Stelkens, R.B., Pierotti, M.E.R., Joyce, D.A., Smith, A.M., van der Sluijs, I. & Seehausen, O. (2008). Disruptive sexual selection on male nuptial coloration in an experimental hybrid population of cichlid fish. *Phil. Trans. R. Soc. B.*, 363, 2861–2870.
- Stinchcombe, J.R., Agrawal, A.F., Hohenlohe, P.A., Arnold, S.J. & Blows, M.W. (2008). Estimating nonlinear selection gradients using quadratic regression coefficients: double or nothing? *Evolution*, 62, 2435–2440.
- Svensson, E. & Calsbeek, R. (2012). *The Adaptive Landscape in Evolutionary Biology*. Oxford University Press, Oxford.
- Taylor, E.B., Boughman, J.W., Groenenboom, M., Sniatynski, M., Schluter, D. & Gow, J.L. (2006). Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Mol. Ecol.*, 15, 343–355.
- West-Eberhard, M.J. (1983). Sexual selection, social competition, and speciation. *Q. Rev. Biol.*, 58, 155–183.

SUPPORTING INFORMATION

Additional Supporting Information may be downloaded via the online version of this article at Wiley Online Library (www.ecologyletters.com).

Editor, Greg Grether

Manuscript received 29 June 2015

First decision made 1 August 2015

Second decision made 8 October 2015

Manuscript accepted 15 October 2015