Adaptive divergence and the evolution of hybrid trait mismatch in threespine stickleback

Avneet K. Chhina,[1,2]* Ken A. Thompson,[1,3]* (iD) and Dolph Schluter[1] (iD)

[1]Department of Zoology & Biodiversity Research Centre, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

[2]E-mail: chhina.k.avneet@gmail.com

[3]E-mail: ken.thompson@zoology.ubc.ca

Selection against mismatched traits in hybrids is the phenotypic analogue of intrinsic hybrid incompatibilities. Mismatch occurs when hybrids resemble one parent population for some phenotypic traits and the other parent population for other traits, and is caused by dominance in opposing directions or from segregation of alleles in recombinant hybrids. In this study, we used threespine stickleback fish (*Gasterosteus aculeatus* L.) to test the theoretical prediction that trait mismatch in hybrids should increase with the magnitude of phenotypic divergence between parent populations. We measured morphological traits in parents and hybrids in crosses between a marine population representing the ancestral form and twelve freshwater populations that have diverged from this ancestral state to varying degrees according to their environments. We found that trait mismatch was greater in more divergent crosses for both F_1 and F_2 hybrids. In the F_1, the divergence–mismatch relationship was caused by traits having dominance in different directions, whereas it was caused by increasing segregating phenotypic variation in the F_2. Our results imply that extrinsic hybrid incompatibilities accumulate as phenotypic divergence proceeds.

KEY WORDS: Ecological speciation, hybridization, opposing dominance.

Impact Summary

Researchers interested in speciation aim to identify general processes that cause branching on the tree of life, or speciation. When different species encounter each other and interbreed, they might form hybrids that co-occur in the parental habitat. If these hybrids fail to persist, then they cannot interbreed with parent lineages and therefore cannot facilitate the exchange of genetic material. This barrier to gene flow promotes speciation. Therefore, it is critical for speciation researchers to understand the factors that affect the fitness of hybrids. It is becoming increasingly appreciated that some hybrids inherit "mismatched" combinations of parental traits, and that this mismatch might render them a poor fit in parental habitats. Our article asks whether the extent of trait mismatches in hybrids increases with the magnitude of adaptive phenotypic differences between parent lineages, which is predicted by theory. We used threespine stickleback fish to test

this prediction. Stickleback in freshwater lakes all originated approximately 10,000 years ago from a common marine ancestor. Although the marine ancestor eats zooplankton in the open water, freshwater populations have adapted to a remarkable diversity of habitats—some retain the ancestral zooplanktivorous habitat while others primarily eat large macroinvertebrates in shallow water. We find that hybrids between a marine population and zooplankitivorous freshwater populations have little mismatch, and that the extent of mismatch grows as the freshwater cross parent is increasingly different from the marine. We identify the broad genetic mechanisms that cause this pattern and find that they largely conform to theory. Because mismatch has been linked to reduced fitness in stickleback and other organisms, our study provides new evidence for a potentially general mechanism linking the process of adaptation to the evolution of reproductive isolation
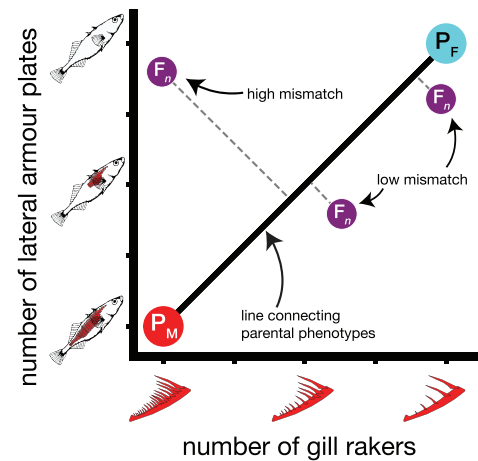
*equal contribution

The evolution of reduced hybrid fitness—postzygotic isolation—is a crucial component of the speciation process (Bateson 1909; Dobzhansky 1937; Mayr 1942, 1963; Muller 1942; Coyne and Orr 2004). Postzygotic isolation is often associated with the buildup of intrinsic genetic incompatibilities that accumulate as populations adapt and diverge (Coyne and Orr 1989; Coughlan and Matute 2020). Yet, many young species lack strong intrinsic incompatibilities and hybrid fitness is instead determined by how the phenotype of hybrids facilitates their interactions with the extrinsic ecological environment to influence performance (Grant 1981; Schluter 2000; Nosil 2012). Hybrids can have poor fitness if they have an intermediate phenotype in an environment where there is no intermediate niche (e.g., insects on distinct host plants; Matsubayashi et al. 2010; Bendall et al. 2017; Zhang et al. 2021). Also, the fitness of hybrids can be reduced if they inherit mismatched combinations of traits from parental species (Arnegard et al. 2014; Thompson et al. 2021). Extrinsic selection against hybrids could grow as parent populations diverge if the extent of trait mismatch increases over the course of divergence between populations, in a manner similar to the growth of intrinsic genetic incompatibilities (Coyne and Orr 1989, 1997; Edmands 1999).
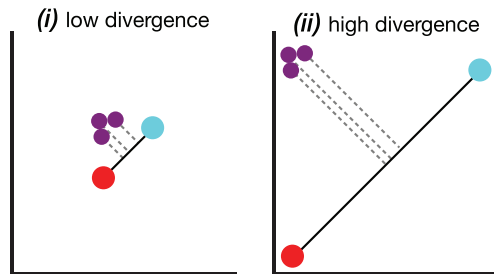
Hybrid mismatch occurs in a pair of traits if the hybrid resembles one parent species in one trait and the other parent species in the other trait (Fig. 1A). Dominance can cause mismatch in hybrids if some traits are dominant toward one parent species and other traits are dominant toward the other parent species. A recent synthesis study suggests that mismatch via dominance is common in F_1 hybrids (Thompson et al. 2021). Segregation variance can also cause mismatch, when individual recombinant hybrids deviate from an intermediate phenotype in opposite ways for different traits (East 1916; Castle 1921; Wright 1931; Schemske and Bradshaw 2002). Because dominance affects the F_1 more than the F_2 (Lynch and Walsh 1998), mismatch in F_1s is primarily expected to be a result of dominance whereas mismatch in recombinant hybrids might be due to one or both dominance and segregation variance.

Theory predicts that the magnitude of trait mismatch in hybrids should be positively associated with divergence between parent populations. Mismatch can be defined geometrically for a pair of traits as the distance between individual hybrid phenotypes and the line connecting parental mean phenotypes (Fig. 1A). For given dominance coefficients, hybrid mismatch will be low when parent species are phenotypically similar (Fig. 1B[i]). If the parents differ substantially, however, the same dominance will generate a greater magnitude of mismatch (Fig. 1B[ii]). The second reason divergence and mismatch should be associated is that the amount of phenotypic variation in the
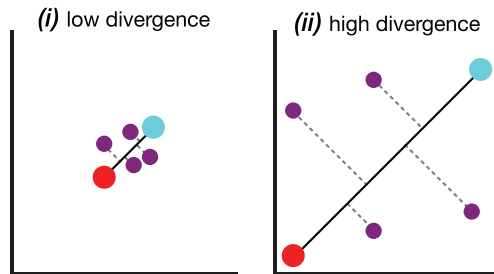

Figure 1. Visual overview of mismatch for two traits. All panels show parent mean phenotypes as large red (marine parent; P_M) and blue points (freshwater parent; P_F) and individual hybrids (F_n) as smaller purple points. Panel (A) shows three individual hybrids—one with high mismatch and two with low mismatch. The length of the dashed lines is the "mismatch" quantity. Panels (B) and (C) show the two mechanisms linking mismatch to phenotypic divergence—dominance and variation. Axes and colors are the same as in panel (A) but labels are omitted for clarity. Both divergence–mismatch panels show (i) a case of low phenotypic divergence between parents and (ii) a case of high phenotypic divergence between parents.

traits of recombinant hybrids—the segregation variance—is expected to increase with the magnitude of phenotypic divergence between parent populations (Slatkin and Lande 1994; Barton 2001; Chevin et al. 2014) (Fig. 1C). Greater segregation variance results in more extreme mismatched trait combinations appearing in hybrids, and therefore this mechanism is expected to generate greater mismatch as the magnitude of divergence between parents increases.

In this article, we use threespine stickleback fish to test the prediction that trait mismatch in hybrids increases with the magnitude of morphological divergence between parent populations. Freshwater stickleback populations have independently diverged to varying degrees from a common marine ancestor since the last glacial maximum (approx. 10 kya). Contemporary marine populations remain abundant in the sea today and are readily crossed with derived forms. Variation among freshwater populations occurs primarily along a limnetic (i.e., zooplanktivorous) to benthic (i.e., consuming large macroinvertebrates living among the vegetation or lake sediments) axis (Schluter and McPhail 1992; Bell and Foster 1994). Although all have adapted to the freshwater habitat, the more limnetic freshwater populations tend to be phenotypically similar to marine populations whereas the more benthic populations are dissimilar. Because more benthic populations have undergone more phenotypic divergence from the marine ancestor, we hypothesized that their hybrids (in crosses with an extant marine population) will exhibit greater mismatch than those produced from crosses with less divergent populations. To test this hypothesis, we measured morphological traits in hybrids from crosses between the ancestral marine form and 12 derived freshwater populations, quantified mismatch, and investigated its causes.

Methods

STUDY SYSTEM

The threespine stickleback is a teleost fish species distributed throughout the coastal areas of the northern hemisphere (Bell and Foster 1994). Stickleback are a longstanding model system for studying the ecological basis of adaptation and speciation (Hagen 1967; McPhail 1969) due in large part to the remarkable phenotypic diversity of populations (Hubbs 1929). Marine stickleback colonized an array of postglacial lakes and have rapidly adapted to prevailing ecological conditions (Schluter 1996). Stickleback that live in lakes containing predators and other competitor fish species (e.g., prickly sculpin) remain similar to the marine population in many morphological traits (Ingram et al. 2012; Miller et al. 2019). By contrast, populations that have evolved in small lakes with few or no predators and competitors often have more derived phenotypes specialized for foraging on large benthic in-

vertebrates. Three lakes (formerly five) contain "species pairs" with reproductively isolated limnetic and benthic populations (McPhail 1984, 1992; Schluter and McPhail 1992)—the limnetics resemble the marine ancestor for many traits, whereas the benthics are among the most derived.

Because adaptive divergence between marine and freshwater populations occurred recently, populations can be readily crossed and typically have few if any "intrinsic" incompatibilities (Hatfield and Schluter 1999; Rogers et al. 2012; Lackey and Boughman 2017). Extant marine populations, within a particular geographic location, are phenotypically similar to the ancestral populations that founded present-day freshwater populations (Morris et al. 2018). We leveraged this continuum of phenotypic divergence using crosses to test the prediction that hybrid mismatch will be greater when more divergent benthic-feeding populations are crossed with a marine ancestral population than when this ancestral population is crossed with less divergent zooplanktivorous populations.

We focus specifically on phenotypic divergence as a predictor in our study but previous work in this system allows us to hypothesize about what might be expected if we focused on genetic divergence instead. Studying whole genomes from stickleback populations (single-species lakes) in coastal British Columbia, Miller (2019) found that mean phenotype (body shape) and was correlated with the major axis of genomic variation among populations, such that populations with similar phenotypes were similar in non-neutral regions of the genome (i.e., those regions evolving in parallel among independent populations). More zooplanktivorous populations (including Pachena and North lakes) were more similar to the marine populations (including the Little Campbell River) in these genomic regions than were more benthic-specialized populations (including Bullock and Cranby lakes). Similarly, Jones (2012a) used single nucleotide polymorphism (SNP) data to study genetic variation in the three species pair lakes considered here and the Little Campbell River marine population, as well as in other marine and single-species lake populations worldwide. The authors found that SNPs under selection were most similar between phenotypically similar populations whereas neutral SNPs largely grouped populations by geography (also see Taylor and McPhail 1999). Wang (2018) replicated these broad pattern using whole genomes. Thus, we would expect estimates of genetic divergence using non-neutral regions of the genome to parallel our estimates of phenotypic divergence of freshwater populations from the marine form.

FISH COLLECTION AND HUSBANDRY

Wild fish were collected in British Columbia, Canada, in April–June 2017 and 2018. We sampled twelve freshwater populations from nine lakes (Fig. 2A). Three lakes (Paxton, Priest, and Little Quarry) contain reproductively isolated benthic-
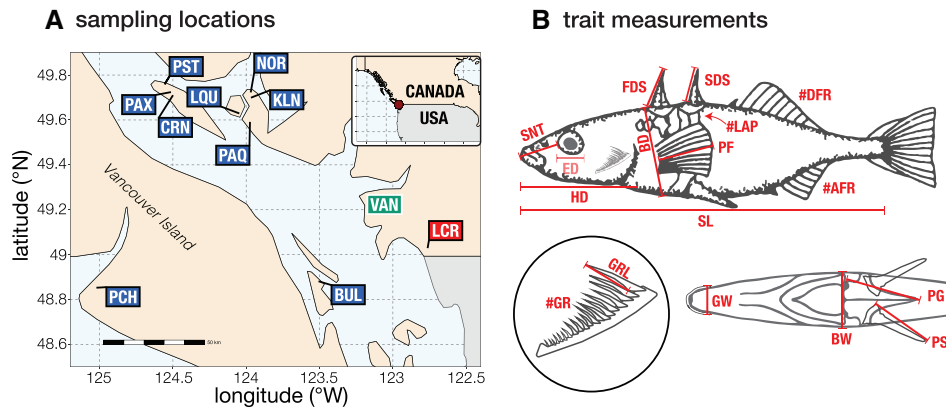
Figure 2. Overview of sampling locations and trait measurements. Panel (A) shows locations where the source populations were collected in British Columbia, Canada. Boxes show collection locations of the marine population (red box; LCR—Little Campbell River) and freshwater populations (blue boxes; left to right: PCH—Pachena Lake; PAX—Paxton Lake; CRN—Cranby Lake; PST—Priest Lake; LQU—Little Quarry Lake; PAQ—Paq (Lily) Lake; NOR—North Lake; KLN—Klein Lake; BUL—Bullock Lake). The green label indicates the location of Vancouver (VAN). Panel (B) shows the measurements of all 16 traits in the dataset and standard length. The upper section of the panel shows the lateral view (traits left to right: SNT—snout length; ED—eye diameter [the transparent shade of red indicates this trait was measured but not analyzed—see "Repeatability" section of methods]; HD—head length; FDS—length of first dorsal spine; BD—body depth; SL—standard length; SDS—length of second dorsal spine; PF—pectoral fin length; #LAP—number of lateral armor plates; #DFR—number of dorsal fin rays; #AFR—number of anal fin rays). The bottom left section of the panel shows a zoomed in drawing of the upper arm of the outer gill raker arch (#GR—number of gill rakers; GRL—length of longest gill raker). The lower right section shows an anteroventral view of the body (GW—gape width; BW—body width; PG—length of pelvic girdle; PS—length of pelvic spine). The upper drawing was originally published by Bell and Foster (1994) and is reused with permission from M. Bell.

limnetic "species pairs" (McPhail 1992) and thus contributed two populations each). Fish from remaining lakes were single-species populations presenting a range of intermediate phenotypes from more zooplanktivorous and marine-like to more benthic. The marine population was collected from the Little Campbell River (Fig. 2A); the population is anadromous—living in the sea and breeding in freshwater, though we refer to them as "marine" here for consistency with previous studies (e.g., Schluter et al. 2021). Wild fish were caught using minnow traps or dip nets. We crossed six gravid marine females with six males from each freshwater population to generate six unique F_1 hybrid families per population, and also generated four to six nonhybrid (i.e., "pure") families for each freshwater parental population and the marine ancestor. All offspring were raised in the lab under common conditions (see Supporting Information "Methods"). Crosses were conducted in only one direction (marine as dam) to standardize cytoplasm among hybrid crosses and also because obtaining a sufficient number of wild gravid females for some populations was prohibitively difficult. When lab-raised fish reached reproductive maturity, F_1 hybrids from unrelated families were crossed to make three F_2 families within each cross-population (with the exception of Paxton Lake benthics which, due to aquarium space constraints in 2018, had only two F_2 families from the same two F_1 parent families).

Fish were lethally sampled when individuals in the tank reached a mean standard length of approximately 40 mm. For F_1s, tanks were subsampled and remaining individuals were raised to produce F_2s. For F_2s, entire tanks were lethally sampled. Fish had not reached reproductive maturity at the time of sampling, and we therefore could not determine their sex. Fish were preserved in formalin, stained with alizarin red, and then stored permanently in 40% isopropyl alcohol. See Table S1 and Table S2 for information on sample sizes.

PHENOTYPE MEASUREMENTS
We measured 16 traits and standard length on stained fish (Fig. 2B). For all traits, we measured at least 100 pure marine parents, and 30 pure freshwater parents, 30 F_1 hybrids, and 60 F_2 hybrids from each population and marine-freshwater cross (all lab-raised)for trait means, standard deviations (SDs), and sample sizes for all populations). We used a dissecting microscope to count the number of dorsal fin rays, anal fin rays, lateral armor plates, and gill rakers. We also measured the length of the longest gill raker using an ocular micrometer. We photographed the left and ventral sides of each fish with a Nikon D300 camera and used ImageJ (Abramoff et al. 2004) to make linear measurements of body dimensions and bones. Pectoral fins were dissected, immersed in a more concentrated alizarin red stain for

at least 24 hours, then photographed. We measured the length of the longest fin ray as pectoral fin length. All measurements with the exception of eye diameter were highly repeatable ($r \geq 0.9$; see Fig S1), and as a result all traits except eye diameter were used for subsequent analysis. A small number ($n = 9$) of fish had missing second dorsal spines, which caused them to be extreme outliers. We excluded these fish from the analysis. One additional fish that failed to inflate its swim bladder was also excluded.

We size-corrected all linear measurements by replacing raw measurements with the residuals from simple log-log (ln-transformation) linear regressions with standard length conducted across the entire dataset. We controlled for as much of the variation as possible among populations by sampling them at a relatively consistent mean size. Log-transformation of linear measurements renders trait variances comparable across fish of different sizes (Hatfield 1997). Some measurements are affected if fish are fixed with an open gape, so we further corrected for fixation position by assigning all fish a number (0, 1, or 2) depending on the extent to which the mouth was open and then performing a further correction that removed the effect of fixation position on gape width, snout length, and head length. Trait measurements for missing spines (first dorsal spine or pelvic spine) or pelvic girdle were given a raw value of 0.1 mm before log-transformation (the log of 0 is undefined). Unlike the second dorsal spine, variation in the presence of these traits is common and does not result in extreme outliers.

Following size-correction, traits were standardized across the entire dataset to a mean of 0 and an SD of 1. This was done because trait divergence has very different magnitudes for different traits (e.g., 30 plates or 1 mm).

DATA ANALYSIS

We first investigated whether trait mismatch was associated with the magnitude of phenotypic divergence between parent populations. Following this, we quantify the role of dominance and trait variation in driving this relationship.

Software

All data processing and model-fitting was done using R (R Core Team 2019) using the `tidyverse` (Wickham 2017). Mixed models were fit using `lme4` (Bates et al. 2014) and analysed using `lmerTest` (Kuznetsova et al. 2014) with the Kenward–Roger approximation for the denominator degrees of freedom (Kenward and Roger 1997). The "map" function in `purrr` (Henry and Wickham 2019), and associated functions in `broom` (Robinson et al. 2020), were used to streamline code for iterating models over grouping variables. Partial residuals were plotted using `visreg` (Breheny and Burchett 2017). `for` loop code was streamlined with the functions in `magicfor` (Makiyama 2016). We used the

`emmeans` package (Lenth et al. 2020) and the "cld" function in `multcomp` (Hothorn et al. 2008) to assist with post hoc comparisons. The functions in the "correlation" package (Makowski et al. 2019) produced correlation matrices.

Quantifying phenotypic divergence

We quantified the magnitude of phenotypic divergence between pure marine and freshwater populations as our main predictor of mismatch. To do this, we calculated the multivariate Euclidean distance between each freshwater population's mean phenotype and the marine mean phenotype based on 15 standardized traits. For all estimates of population mean phenotypes we use the unweighted mean of family means, though our conclusions are unchanged if we average across individuals rather than families.

Trait mismatch

We quantified trait mismatch as the extent to which individual hybrids deviate from the line connecting parental mean phenotypes (Fig. 1; Thompson et al. 2021). We calculated mismatch between pairs of traits or for all traits together, and because conclusions are broadly similar between the approaches we primarily consider the latter multivariate metric in the main text (see Supporting Information "Results" for pairwise mismatch methods and results). Correlations between pairs of traits in F_2 hybrids were low (median $|r_{\text{Pearson}}| = 0.2$), and most (87%) were not statistically significant at $P = 0.05$, and for this reason we do not use dimensionality-reduction techniques such as principle components analysis.

Mismatch is the shortest (i.e., perpendicular) Euclidean distance between a hybrid's phenotype and the line that connects the two parental mean phenotypes (Fig. 1). Mismatch (d_{mm}) was calculated using the standardized traits as:

$$d_{mm} = \left\| (\vec{F_n} - \vec{P_M}) - (\vec{P_F} - \vec{P_M}) \times \frac{(\vec{F_n} - \vec{P_M}) \cdot (\vec{P_F} - \vec{P_M})}{\|\vec{P_F} - \vec{P_M}\|^2} \right\|, (1)$$

where $\vec{F_n}$, $\vec{P_M}$, and $\vec{P_F}$ are the vectors of individual hybrid ($F_n = F_1$ or F_2), mean marine, and mean freshwater trait values. Individuals from parent populations exhibit deviations from the line connecting parent population means due to their phenotypic variation. The average deviation did not vary among freshwater parent populations ($F_{11, 33.7} = 1.45$; $P = 0.20$) and accounting for parent deviation in our model does not change our conclusions (see archived analysis code).

We tested whether mismatch changed with the magnitude of phenotypic divergence between marine and freshwater parent populations. For simplicity, we analyze mean mismatch values for each cross type within a given marine × freshwater cross, treating population as the replicate. All of our qualitative conclusions are unchanged if we analyze individual-level data using

mixed models with family and population as nested random effects (see archived analysis code). Predictor variables were the (Euclidean) phenotypic divergence between the parental populations, hybrid category (F_1 or F_2), and their interaction. Freshwater population was a random effect.

Mechanisms underlying the divergence–mismatch relationship

We investigated the causal roles of dominance and phenotypic variation for generating the divergence–mismatch relationship. To determine how dominance affects mismatch for a given marine × freshwater cross, we calculated the mismatch of the mean hybrid phenotype for each population and hybrid generation—hereafter the effect of dominance on mismatch. To determine how phenotypic variation affects mismatch, we subtracted the mismatch of the mean hybrid phenotype (the effect of dominance calculated above) from each individual's unique mismatch value. We took the average of these differences for each family, then averaged these for a single estimate—the effect of variance on mismatch—per population. We used linear models to test whether the effects of dominance and variance on mismatch were associated with the magnitude of phenotypic divergence between parents.

We also quantified general patterns of dominance and phenotypic variation in both F_1 and F_2 hybrids to gain intuition about why they drive a divergence–mismatch relationship. We determined whether traits exhibited significant deviations from additivity for each population using linear mixed models where standardized trait values were the response, and the two predictor variables were (i) an additive term (fraction of the genome that is freshwater [$P_M = 0$, $F_n = 0.5$; $P_F = 1$]) and (ii) a dominance deviation (fraction of genome heterozygous [P_M and $P_F = 0$, $F_2 = 0.5$; $F_1 = 1$]) (Lynch and Walsh 1998). Family was a random effect. Dominance was only estimated for traits where the marine and freshwater parent populations were statistically distinguishable (*t*-test $P < 0.05$). To compare dominance among populations, we calculated dominance coefficients by quantifying where the mean (F_1 or F_2) hybrid phenotype fell when the marine parent's value was scaled to a value of 0 and the freshwater parent's value was scaled to a value of 1 (transgressive values [<0 and >1] are possible). We tested if dominance coefficients changed with the magnitude of phenotypic divergence between parents using linear models. To evaluate overall patterns of trait variation, we calculated the variance for each trait within each family, then took the average across families. We then fit linear models with mean population variance as the response and both phenotypic divergence between parent populations and hybrid category (and their interaction) as predictors.
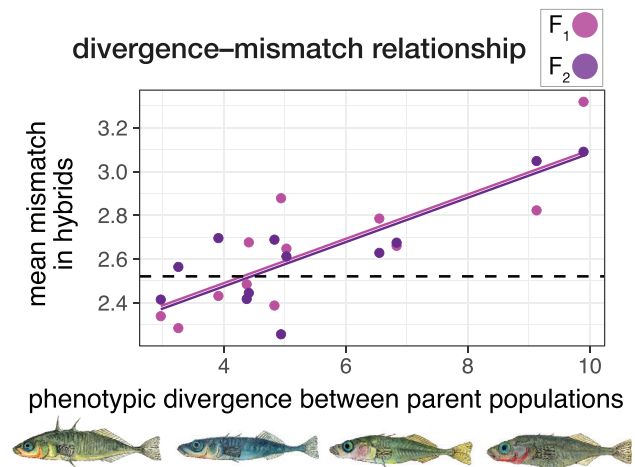

Figure 3. Trait mismatch in hybrids increases with the magnitude of phenotypic divergence between their parents. Each point is the mean mismatch (eq. 1) value across all F_1 (pink points and lines) or F_2 (purple points and lines) hybrids for a given marine × freshwater cross ($n = 12$ per hybrid type). The dashed horizontal line shows the mean "mismatch" value (eq. 1) calculated across the freshwater parent (i.e., nonhybrid) populations. Points and regression lines are partial residuals from mixed models. Fish images (by K. Chu) illustrate the range of phneotypes, from left to write: marine form, limnetic, intermediate single-species form, and benthic.

Results

PATTERNS OF PHENOTYPIC DIVERGENCE AMONG POPULATIONS

Marine × freshwater crosses differed substantially in the magnitude of phenotypic divergence between parent populations (main effect of "population": $F_{1, 41.8} = 34.1$; $P < 0.0001$; Fig. S2). Freshwater populations were between approximately 3–10 units diverged from the marine, based on 15 standardized traits. The benthic populations from the sympatric species pairs were among the most divergent from the marine ancestor, while two highly zooplanktivorous populations that co-exist with prickly sculpin were among the least diverged (Pachena Lake and North Lake). The number of traits that differed significantly between the freshwater and marine parents was positively correlated with the magnitude of divergence between parent populations (Fig. S3).

EVOLUTION OF TRAIT MISMATCH IN HYBRIDS

We found support for the prediction that hybrid trait mismatch increases with the magnitude of phenotypic divergence between parents. Considering all traits together, multivariate mismatch in hybrids was positively associated with the magnitude of phenotypic divergence between parents ($\hat{\beta} = 0.10 \pm 0.036$ [SE], $F_{1,10} = 7.71$, $P = 0.020$) (Fig. 3). The rate of increase of mismatch with divergence (regression slope) did not differ between F_1 and F_2 hybrids (divergence × category interaction $P = 0.61$). Thus, for every unit of multivariate phenotypic divergence between
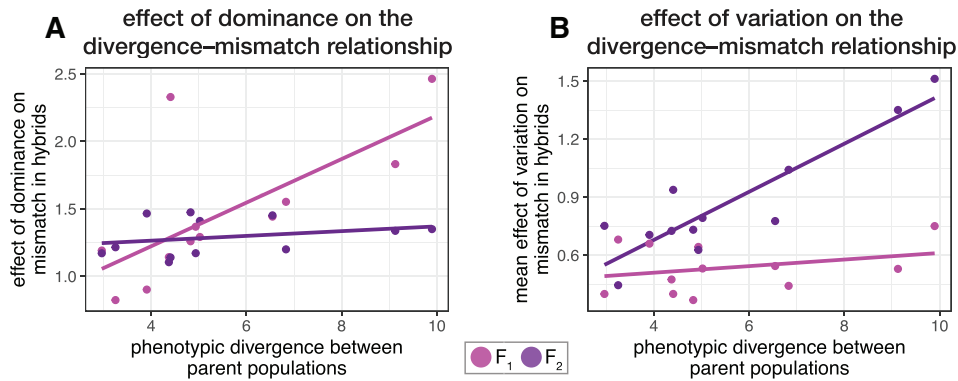
Figure 4. Dominance is the primary cause of the divergence–mismatch relationship in F_1 hybrids, while variance is the primary cause of the divergence–mismatch relationship in F_2 hybrids. Panel (A) depicts the mismatch of the mean hybrid phenotype (the *effect of dominance on mismatch*), which is only caused by dominance. The slope is significantly different from zero in F_1s but not in F_2s. Panel (B) depicts the mean difference between mismatch values for individual hybrids and the mismatch of the mean hybrid phenotype (the *effect of variance on mismatch*). The slope is significantly different from zero in F_2s but not in F_1s. One point is shown per marine × freshwater cross and category (i.e., F_1 or F_2; each $n = 12$). Points and regression lines are partial residuals from mixed models.

parents, mismatch in hybrids increases by approximately one-tenth that amount. Figure S4 gives an example of mismatch for two traits where parents have different magnitudes of divergence.

UNDERLYING CAUSES OF THE DIVERGENCE–MISMATCH RELATIONSHIP

Dominance was the main cause of the divergence–mismatch relationship in F_1 hybrids, whereas phenotypic variation was the main cause of this relationship in F_2 hybrids (Fig. 4).

Mismatch of the mean hybrid phenotype—the effect of dominance on mismatch—increased with the magnitude of phenotypic divergence between parents in F_1 hybrids ($\hat{\beta}$ [95% CI] = 0.16 [0.052, 0.271]) but not in F_2s ($\hat{\beta} = 0.018$ [−0.92, 0.13]; parent divergence × hybrid category interaction term $F_{1,10} = 5.14$, $P = 0.047$) (Fig. 4A). More than three-fourth (77%) of traits were inherited nonadditively when tested within populations and two-thirds of traits exhibiting deviations from additivity tended toward recessivity (i.e., F_1s were more marine-like). The average dominance of traits differed among traits—for example, most F_1 hybrids had long heads, similar to the freshwater populations, but also had large pectoral fins, similar to the marine population (Fig. S5). Dominance coefficients typically did not change consistently with the magnitude of phenotypic divergence between parents, although such a pattern was evident for the number of lateral armor plates and the length of pelvic spines (Fig. S6).

We found that the mismatch caused by deviation from the mean hybrid phenotype—the effect of variation on mismatch—increased with the magnitude of phenotypic divergence between parents in F_2 hybrids ($\hat{\beta} = 0.12$ [0.069, 0.18]) but had no effect in F_1s ($\hat{\beta} = 0.017$ [−0.039, 0.073]) (Fig. 4B; parent divergence × hybrid category interaction term, $F_{1,10} = 3.57$, $P = 0.0051$).

The relationship between the variance-effect and divergence was caused by trait variation increasing with phenotypic divergence between parents in F_2 hybrids but not in F_1s (Fig. S7). Together, these analyses indicate that the quantitative genetic basis of the divergence–mismatch relationship differs between hybrid generations.

As a result of dominance, mismatch, and divergence between parents itself, the distance of 316 F_1 and F_2 hybrids to both the marine and the freshwater parent mean phenotypes—potentially fitness optima—grows with increasing divergence between populations (Fig. S8).

Discussion

In this study, we used experimental hybridization in stickleback to test whether the extent of trait mismatch in hybrids grows as parent populations diverge phenotypically. Trait mismatch has been associated with reduced individual fitness in recombinant stickleback (Arnegard et al. 2014) and sunflower (Thompson et al. 2021) hybrids, and a growing number of studies have used indirect inference to link mismatch in F_1 hybrids to reproductive isolation (Vinšálková and Gvoždík 2007; Matsubayashi et al. 2010; Cooper et al. 2018). Studies of fitness landscapes in hybrids (Martin and Wainwright 2013; Keagy et al. 2016) and correlational selection within species (Schluter 1994), also show patterns consistent the hypothesis that selection acts against mismatched trait combinations.

Our study was motivated by the fact that, although previous studies have documented a seemingly general relationship between ecological divergence and barriers to gene flow (Shafer and Wolf 2013), predictions of hypotheses that link adaptive divergence to the evolution of potentially maladaptive hybrid

phenotypes remain untested. Comparative studies of speciation (Matute and Cooper 2021) typically study divergence over time, whereas here we consider populations from different lakes that have diverged from a common ancestor to varying degrees in roughly the same amount of time—thus isolating the effect of phenotypic divergence. In support of our prediction, more derived and divergent freshwater parental populations tend to produce hybrids with increasingly mismatched phenotypes when each is crossed to the same marine population representing the ancestral form. The quantitative genetic underpinnings of the divergence–mismatch relationship—dominance in F_1s and segregation variance in F_2s—follow from theory and first principles. Below, we discuss these results in the context of speciation research and the genetics of adaptation.

RELATION TO "INTRINSIC" INCOMPATIBILITIES

Mismatch might be thought of as an ecological and phenotypic analogue to classic "intrinsic" Bateson–Dobzhansky–Muller (BDM; Bateson 1909; Dobzhansky 1937; Muller 1942) hybrid incompatibilities (Coyne and Orr 2004; Arnegard et al. 2014). In the BDM model, populations diverge at multiple loci each of whose transitions is advantageous (or neutral) in the genetic background where fixation occurred. Divergent alleles at loci having separate evolutionary histories might interact negatively when combined in a hybrid, causing reduced fitness. In the phenotypic analogue, populations diverge in multiple traits each of whose changes is advantageous in its own environment and on a given phenotypic background. When combined in a hybrid, (net) dominance relationships and segregation variance might produce mismatch between traits in hybrids, reducing performance and fitness. Such phenotypic incompatibilities imply negative interactions between the underlying genes, as in the BDM model. A phenotypic perspective provides additional insights. Phenotypic incompatibilities are inevitably environment-dependent. They may be detectable in one environment and not in another (Arnegard et al. 2014). Phenotypic incompatibilities will often be frequency- and density-dependent, at least for traits involved in resource exploitation, implying a more dynamic adaptive landscape for hybrid fitness than in the classic BDM model.

Because mismatch increases with parental divergence, this implies that phenotypic incompatibilities might evolve in a similar "clock-like" manner to intrinsic post-zygotic isolation. Coyne and Orr (1989, 1997) were the first to demonstrate that reproductive isolation between populations evolves as a function of divergence time. They found that both premating and intrinsic postzygotic isolation increased with genetic distance (a proxy for time) in *Drosophila*. This work spawned a small industry (Coughlan and Matute 2020; Matute and Cooper 2021) reporting similar patterns in groups as diverse as orchids and fishes (Bolnick and Near 2005; Scopece et al. 2013). In the present study, however,

freshwater stickleback populations varying in their degree of divergence from the ancestral marine form are similar in age, all having adapted to varying lake conditions since the end of the last ice age (Wang 2018; Miller et al. 2019). Hence, mismatch is decoupled from neutral genetic divergence and time in this specific case.

The "snowball" model of the accumulation of hybrid incompatibilities, first put forward by Orr (1995), suggests that the number of hybrid incompatibilities should increase faster-than-linearly with divergence time. This is because the number of potential pairwise interactions among divergent loci—and thus pairwise incompatibilities—increases at least as fast ("at least" because this does not account for anything above pairwise interactions) as the square of the number of substitutions separating species. Empirical work has found support for this snowball model (Matute et al. 2010; Moyle and Nakazato 2010; Wang et al. 2015). Using the pairwise mismatch data (see Supporting Information), we determined whether each trait pair was significantly mismatched for each cross. In this analysis, we find that the number of trait pairs that are significantly mismatched increases quadratically with the magnitude of phenotypic divergence between parents in F_1 hybrids (see Fig. S9). In F_2 hybrids the pattern is significant but linear. We view this analysis as purely heuristic because trait pairs are not independent, though this issue likely also affects other empirical studies of snowballing incompatibilities because the interacting genes are not known (Matute et al. 2010; Moyle and Nakazato 2010). Nevertheless, trait mismatches do seem to snowball in a similar manner to intrinsic incompatibilities (Matute et al. 2010; Moyle and Nakazato 2010). Even though the number of pairs of traits showing mismatch grows faster than linearly with increasing phenotypic divergence between parents, we showed that multivariate mismatch grows linearly with divergence.

DOMINANCE OF FRESHWATER-ADAPTIVE TRAIT VALUES

We found that derived traits are typically not dominant in stickleback (see Fig. S5). Although unexpected if adaptation were from new mutation (Haldane 1927), this finding is consistent with what is known about adaptation to freshwater habitats in stickleback proceeding predominantly via the sorting of existing standing genetic variation (Jones et al. 2012; Nelson and Cresko 2018; Roberts Kingman et al., 2021). We also found that deviations from additivity were in a recessive direction (i.e., toward marine) more often than in a dominant direction. These results are consistent with the findings of Miller et al. (2014), who used QTL mapping to measure dominance in an F_2 marine (from Japan) × freshwater (Paxton Lake benthic) cross, and found that most QTL were additive or partially additive with a slight but significant bias toward recessivity. Finally, with two notable exceptions,

dominance coefficients changed idiosyncratically with increasing phenotypic divergence. This finding that dominance is difficult to predict is in agreement with a recent synthesis study that analyzed data from over 100 crosses (Thompson et al. 2021).

Dominance coefficients of two traits—the number of lateral armor plates and the length of the pelvic spine—varied with the magnitude of phenotype divergence between parent populations. Plate number was recessive in crosses between the marine ancestral form and the least divergent freshwater populations and increasingly dominant when the marine was crossed with more derived freshwater populations. Plate reduction in freshwater stickleback populations is known to be largely caused by a large-effect variant at the *Eda* locus (Colosimo et al. 2005; Archambeault et al. 2020), which is likely to be fixed for the freshwater allele in all but one (North Lake) of the freshwater populations considered here. Previous studies have shown that additional known alleles that reduce plate number in freshwater stickleback populations also modify the dominance of *Eda*'s effect on the number of armor plates (Colosimo et al. 2004). Our result adds to the understanding of this trait by showing that the net contribution of alleles that modify dominance of lateral armor plates depends on the degree of phenotypic divergence. We also found that reduced length of pelvic spines—the freshwater phenotype—was largely dominant in crosses involving less divergent freshwater populations and largely recessive in crosses involving more derived populations. Less is known about dominance modifiers for the length of the pelvic girdle, although the complete loss of the pelvic girdle is governed by recurrent de novo mutation in *PitX1* (Chan et al. 2010). Differences in the direction of the evolution of dominance between these two different armour traits could be associated with their pathways to fixation. The freshwater allele at *Eda* was likely present at the time freshwater populations originated, so dominance modifiers would have fixed afterward. By contrast, the freshwater allele at *PitX1* likely fixed well after substantial reduction of the pelvis had already occurred in some derived populations—this possibility warrants theoretical investigation. Although hypotheses about the evolution of dominance abound (Fisher 1928; Wright 1934), we cannot determine here whether the evolution of dominance documented for these two traits is incidental or adaptive.

CAVEATS, FUTURE DIRECTIONS, AND CONCLUSIONS

The clearest limitation of our study is its lack of a direct link between mismatch and fitness. In stickleback, selection against mismatched trait combinations has been shown in crosses between freshwater benthic and limnetic populations (Arnegard et al. 2014; Keagy et al. 2016). Selection against natural marine × freshwater hybrids has been inferred in hybrid zones from the steepness of clines (Vines et al. 2016), as well as observations of heterozygote deficit and cytonuclear disequilibrium (Jones et al.

2006). In the Little Campbell River, the source of the marine population used here, Hagen (1967) inferred that selection against marine × freshwater hybrids is "very intense", although the specific mechanisms of selection were unclear. Clearly, we must begin to conduct comparative studies where mismatch can be linked to fitness directly. In nhybrid zones, biologists can estimate the strength of selection against hybrids by, among other methods, measuring cline width and back-crossing rates. Future studies could leverage areas of ongoing hybridization to evaluate whether phenotypic mismatch measured in crosses predicts the strength of selection against natural hybrids (Barton and Hewitt 1985; Harrison 1993). Experimental arrays with recombinant hybrids and nonhybrid progenitors could be used to relate mismatch to back-crossing rates to identify its effectiveness as a barrier to gene flow. Experimental evolution studies could also be used to robustly estimate the generality of the divergence–mismatch relationship and its effect on hybrid fitness. It would be particularly valuable to identify generalities about whether the mismatch–fitness relationship is linear, diminishing, or accelerating. Clearly, more work is necessary to solidify our general understanding of the fitness effects of mismatch.

Some of our findings are more likely to be general than others. In particular, our results regarding F_2 hybrid phenotypic variation increasing with the magnitude of divergence between parents were predicted from theory (Slatkin and Lande 1994; Barton 2001; Chevin et al. 2014). It is therefore a reasonable expectation that this particular finding might extend to other systems. Because dominance is somewhat idiosyncratic, it is less clear how general our finding that dominance causes a divergence–mismatch relationship in F_1s will be, though dominance in hybrids is the rule rather than the exception (Thompson et al. 2021). Because the divergence–mismatch relationship we document here might be an important mechanism driving progress toward speciation, establishing which aspects are idiosyncratic and which are general seems worth the effort.

Ultimately, there are many causes of speciation and trait mismatch will be one of many proximate causes of reproductive isolation. Any mismatch would likely operate alongside other well-documented processes such as assortative mating (Rundle et al. 2000; Coyne and Orr 2004; Jiang et al. 2013) and selection against immigrants (Nosil et al. 2005). Empirical estimates of the relationship between ecological divergence and hybrid fitness (Edmands 1999; Funk et al. 2006; Shafer and Wolf 2013), or neutral divergence and hybrid fitness (Coughlan and Matute 2020; Matute and Cooper 2021), invariably find that these relationships are noisy. Because F_1 hybrid mismatch is prevalent in many systems (Thompson et al. 2021), it could be an immediate and powerful barrier to gene flow between many diverging lineages. As shown above, it might even "snowball". Future

studies clarifying the importance of mismatch for speciation are sorely needed.

LITERATURE CITED

Abramoff, M. D., P. J. Magalhaes, and S. J. Ram. 2004. Image processing with ImageJ. Biophoton. Int. 11:36–42.

Archambeault, S. L., L. R. Bärtschi, A. D. Merminod, and C. L. Peichel. 2020. Adaptation via pleiotropy and linkage: association mapping reveals a complex genetic architecture within the stickleback *Eda* locus. Evol. Lett. 4:282–301.

Arnegard, M. E., M. D. McGee, B. Matthews, K. B. Marchinko, G. L. Conte, S. Kabir, et al. 2014. Genetics of ecological divergence during speciation. Nat. 511:307–311.

Barton, N., and G. Hewitt. 1985. Analysis of hybrid zones. Annu. Rev. Ecol. Systematics. 16:113–148.

H Barton, N.. 2001. The role of hybridization in evolution. Mol. Ecol. 10:551–568.

Bates, D., M. Maechler, B. Bolker, and W. S. 2014. lme4: linear mixed-effects models using Eigen and S4. R package version 1.1-7. Available at https://github.com/lme4/lme4/.

Bateson, W. 1909. Heredity and variation in modern lights. Pp. 85–101 *in* A. C. Seward, ed. Darwin and modern science. Cambridge Univ. Press, Cambridge, U.K.

Bell, M. A., and S. A. Foster. 1994. The evolutionary biology of the threespine stickleback. Oxford Univ. Press, Oxford, U.K.

Bendall, E. E., K. L. Vertacnik, and C. R. Linnen. 2017. Oviposition traits generate extrinsic postzygotic isolation between two pine sawfly species. BMC Evol. Biol. 17:1–15.

Bolnick, D. I., and T. J. Near. 2005. Tempo of hybrid inviability in centrarchid fishes (Teleostei: Centrarchidae). Evolution 59:1754–1767.

Breheny, P., and W. Burchett. 2017. Visualization of regression models using visreg. R J. 9:56–71.

E Castle, W.. 1921. An improved method of estimating the number of genetic factors concerned in cases of blending inheritance. Science 54:223.

Chan, Y. F., M. E. Marks, F. C. Jones, G. Villarreal, M. D. Shapiro, S. D. Brady, et al. 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. Science 327:302–305.

Chevin, L.-M., G. Decorzent, and T. Lenormand. 2014. Niche dimensionality and the genetics of ecological speciation. Evolution 68:1244–1256.

Colosimo, P. F., C. L. Peichel, K. Nereng, B. K. Blackman, M. D. Shapiro, D. Schluter, et al. 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. PLoS Biol. 2: e109.

Colosimo, P. F., K. E. Hosemann, S. Balabhadra, G. Villarreal, H. Dickson, J. Grimwood, et al. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. Science 307:1928–1933.

Cooper, B. S., A. Sedghifar, W. T. Nash, A. A. Comeault, and D. R. Matute. 2018. A maladaptive combination of traits contributes to the maintenance of a Drosophila hybrid zone. Curr. Biol. 28:2940–2947.

Coughlan, J. M., and D. R. Matute. 2020. The importance of intrinsic postzygotic barriers throughout the speciation process: intrinsic barriers throughout speciation. Philos. Trans. Roy. Soc. B 375: 20190533.

Coyne, J. A., and H. A. Orr. 1989. Patterns of speciation in Drosophila. Evolution 43:362–381.

———. 1997. "Patterns of speciation in Drosophila" Revisited. Evolution 51:295.

———. 2004. Speciation. Sinauer Associates, Sunderland, MA.

Dobzhansky, T. 1937. Genetics and the origin of species. Columbia Univ. Press, New York.

M East, E.. 1916. Inheritance in crosses between *Nicotiana langsdorffii* and *Nicotiana alata*. Genetics 1:311.

Edmands, S. 1999. Heterosis and outbreeding depression in interpopulation crosses spanning a wide range of divergence. Evolution 53:1757–1768.

A Fisher, R.. 1928. The possible modification of the response of the wild type to recurrent mutations. Amer. Nat. 62:115–126.

Funk, D. J., P. Nosil, and W. J. Etges. 2006. Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. Proc. Natl. Acad. Sci. U.S.A. 103:3209–3213.

Grant, V. 1981. Plant speciation. Columbia Univ. Press, New York, NY.

W Hagen, D.. 1967. Isolating mechanisrns in threespine stickleback (*Gasterosteus*). J. Fish. Res. Board Can. 24:1637–1692.

S Haldane, J. B.. 1927. A mathematical theory of natural and artificial selection, part V: selection and mutation. Math. Proc. Camb. Philos. Soc. 23:838–844.

Harrison, R. G. 1993. Hybrid zones and the evolutionary process. Oxford Univ. Press, New York.

Hatfield, T. 1997. Genetic divergence in adaptive characters between sympatric species of stickleback. Am. Nat. 149:1009–1029.

Hatfield, T., and D. Schluter. 1999. Ecological speciation in sticklebacks: environment-dependent hybrid fitness. Evolution 53:866–873.

Henry, L., and H. Wickham. 2019. purrr: Functional programming tools. Available at https://github.com/tidyverse/purrr/issues.

Torsten, Hothorn, Frank Bretz & Peter Westfall (2008). Simultaneous inference in general parametric models. Biometrical Journal 50(3):346–363.

L Hubbs, C.. 1929. The Atlantic American species of the fish genus, *Gasterosteus*. Occ. Pap. Mus. Zool. Univ. Michigan 200:1–9.

Ingram, T., R. Svanbäck, N. J. Kraft, P. Kratina, L. Southcott, and D. Schluter. 2012. Intraguild predation drives evolutionary niche shift in threespine stickleback. Evolution 66:1819–1832.

Jiang, Y., D. I. Bolnick, and M. Kirkpatrick. 2013. Assortative mating in animals. Am. Nat. 181:E125–E138.

Jones, F. C., C. Brown, J. M. Pemberton, and V. A. Braithwaite. 2006. Reproductive isolation in a threespine stickleback hybrid zone. J. Evol. Biol. 19:1531–1544.

Jones, F. C., M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, et al. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. Nature 484:55–61.

Keagy, J., L. Lettieri, and J. W. Boughman. 2016. Male competition fitness landscapes predict both forward and reverse speciation. Ecol. Lett. 19:71–80.

Kenward, M. G., and J. H. Roger. 1997. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 53:983–997.

Kingman, Garrett A. Roberts, Deven N. Vyas, Felicity C. Jones, Shannon D. Brady, Heidi I. Chen, Kerry Reid, Mark Milhaven et al. 2021 "Predicting future from past: The genomic basis of recurrent and rapid stickleback evolution." Science Advances 7, no. 25 eabg5285.

Kuznetsova, A., P. B. Brockhoff, and H. B. Christensen. 2014. lmerTest: tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). v. 2.0-11. Available at http://CRAN.R-project.org/package=lmerTest.

Lackey, A. C., and J. W. Boughman. 2017. Evolution of reproductive isolation in stickleback fish. Evolution 71:357–372.

Lenth, R., H. Singmann, J. Love, P. Buerkner, and M. Herve. 2020. Package emmeans. R topics documented. Am. Stat. 34:216–221.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. 1st ed. Sinauer, Sunderland, MA.

Makiyama, K. 2016. magicfor: magic functions to obtain results from for loops. Available at https://rdrr.io/cran/magicfor/.

Makowski, D., M. S. Ben-Shachar, I. Patil, and D. Lüdecke. 2019. Methods and algorithms for correlation analysis. R J. Open Source Soft. 5:2306.

Martin, C. H., and P. C. Wainwright. 2013. Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. Science 339:208–11.

Matsubayashi, K. W., I. Ohshima, and P. Nosil. 2010. Ecological speciation in phytophagous insects. Entomol. Exp. Appl. 134:1–27.

Matute, D. R., and B. S. Cooper. 2021. Comparative studies on speciation: 30 years since Coyne and Orr. Evolution: *In press*.75 764–778.

Matute, D. R., I. A. Butler, D. A. Turissini, and J. A. Coyne. 2010. A test of the snowball theory for the rate of evolution of hybrid incompatibilities. Sci. 329:1518–1521.

Mayr, E. 1942. Systematics and the origin of species. Columbia Univ. Press, New York.

———. 1963. Animal species and evolution. The Belknap Press, Cambridge, MA.

McPhail, J. D.. 1969. Predation and the evolution of a stickleback (*Gasterosteus*). J. Fish. Res. Board Can. 26:3183–3208.

———. 1984. Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): morphological and genetic evidence for a species pair in Enos Lake, British Columbia. Can. J. Zool. 62:1402–1408.

———. 1992. Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): evidence for a species-pair in Paxton Lake, Texada Island, British Columbia. Can. J. Zool. 70:361–369.

Miller, C. T., A. M. Glazer, B. R. Summers, B. K. Blackman, A. R. Norman, M. D. Shapiro, et al. 2014. Modular skeletal evolution in sticklebacks is controlled by additive and clustered quantitative trait loci. Genetics 197:405–420.

Miller, S. E., M. Roesti, and D. Schluter. 2019. A single interacting species leads to widespread parallel evolution of the stickleback genome. Curr. Biol. 29:530–537.

Morris, M. R., E. Bowles, B. E. Allen, H. A. Jamniczky, and S. M. Rogers. 2018. Contemporary ancestor? Adaptive divergence from standing genetic variation in Pacific marine threespine stickleback. BMC Evol. Biol. 18:113.

Moyle, L. C., and T. Nakazato. 2010. Hybrid incompatibility "snowballs" between *Solanum* species. Science 329:1521–1523.

J Muller, H.. 1942. Isolating mechanisms, evolution and temperature. Biol. Symp. 6:71–125.

Nelson, T. C., and W. A. Cresko. 2018. Ancient genomic variation underlies repeated ecological adaptation in young stickleback populations. Evol. Lett. 2:9–21.

Nosil, P. 2012. Ecological speciation. Oxford Univ. Press, Oxford, U.K.

Nosil, P., T. H. Vines, and D. J. Funk. 2005. Reproductive isolation caused by natural selection against immigrants from divergent habitats. Evol. Int. J. Org. Evol. 59:705–719.

A Orr, H.. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics 139:1805–1813.

R Core Team. 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at https://www.R-project.org/.

Robinson, D., A. Hayes, and S. Couch. 2020. broom: convert statistical objects into tidy tibbles. Available at https://broom.tidymodels.org/.

Rogers, S. M., P. Tamkee, B. Summers, S. Balabahadra, M. Marks, D. M. Kingsley, et al. 2012. Genetic signature of adaptive peak shift in threespine stickleback. Evolution 66:2439–2450.

Rundle, H. D., L. Nagel, J. Boughman, D. Schluter, and J. Wenrick Boughman. 2000. Natural selection and parallel speciation in sympatric sticklebacks. Science 287:306–308.

Schemske, D. W., and H. D. Bradshaw. 2002. Pollinator preference and the evolution of floral traits in monkeyflowers (Mimulus). Proc. Natl. Acad. Sci. 96:11910–11915.

Schluter, D. 1994. Experimental evidence that competition promotes divergence in adaptive radiation. Science 266:798–801.

———. 1996. Ecological speciation in postglacial fishes. Philos. Trans. R. Soc. B 351:807–814.

———. 2000. The ecology of adaptive radiation. Oxford Univ. Press, New York.

Schluter, D., and J. D. McPhail. 1992. Ecological character displacement and speciation in sticklebacks. Am. Nat. 140:85–108.

Schluter, D., K. B. Marchinko, M. E. Arnegard, H. Zhang, S. D. Brady, F. C. Jones, et al. 2021. Fitness maps to a large-effect locus in introduced stickleback populations. Proc. Natl. Acad. Sci. 118: e1914889118.

Scopece, G., A. Croce, C. Lexer, and S. Cozzolino. 2013. Components of reproductive isolation between Orchis Mascula and Orchis Pauciflora. Evolution. 67:2083–2093.

Shafer, A. B. A., and J. B. W. Wolf. 2013. Widespread evidence for incipient ecological speciation: a meta-analysis of isolation-by-ecology. Ecol. Lett. 16:940–950.

Slatkin, M., and R. Lande. 1994. Segregation variance after hybridization of isolated populations. Genet. Res. 64:51–56.

Thompson, K. A., M. Urquhart-Cronish, K. D. Whitney, L. H. Rieseberg, and D. Schluter. 2021. Patterns, predictors, and consequences of dominance in hybrids. Am. Nat. 197 E72–E88.

Vines, T., A. Dalziel, A. Albert, T. Veen, P. Schulte, and D. Schluter. 2016. Cline coupling and uncoupling in a stickleback hybrid zone. Evolution 70:1023–1038.

Vinšálková, T., and L. Gvoždík. 2007. Mismatch between temperature preferences and morphology in F1 hybrid newts (*Triturus carnifex* × *T. dobrogicus*). J. Thermal Biol. 32:433–439.

Wang, M. 2018. The genetic basis of adaptation and speciation in benthic and limnetic threespine stickleback (*Gasterosterus aculeatus*). Ph.D. thesis, Max Planck Campus Tübingen.

Wang, R. J., M. A. White, and B. A. Payseur. 2015. The pace of hybrid incompatibility evolution in house mice. Genetics 201:229–242.

Wickham, H. 2017. Tidyverse: easily install and load the "Tidyverse." R package version 1.2.1. Available at https://tidyverse.tidyverse.org/.

Wright, S. 1931. Evolution in mendelian populations. Genetics 16:97–159.

———. 1934. Physiological and evolutionary theories of dominance. Am. Nat. 68:24–53.

Zhang, L., G. R. Hood, A. M. Roush, S. A. Shzu, M. S. Comerford, J. R. Ott, et al. 2021. Asymmetric, but opposing reductions in immigrant viability and fecundity promote reproductive isolation among host-associated populations of an insect herbivore. Evolution 75:476–489.

Associate Editor: R. Snook

Supporting Information

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

Table S1: Sample sizes that we strove for in this study; see Table S2 for realized sample sizes.

Table S2: Sample sizes realized in this study (total *n* fish in study: 1658). Some analyses of individual traits had more individuals than we list here, but individuals needed to have all traits measured to be included in mismatch analyses

Figure S1: Repeatability data for all measured traits.

Figure S2: Among-population variation in the phenotypic divergence to the marine ancestor.

Figure S3: Phenotypic divergence between parents is positively associated with the number of traits that differ between them.

Figure S4: Visualization of pairwise mismatch for two traits using our empirical trait data.

Figure S5: Dominance of the freshwater phenotype in hybrids for different traits.

Figure S6: Evolution of dominance.

Figure S7: Phenotypic variation increases with the magnitude of phenotypic divergence between parents in F2 (purple) hybrids but not in F1 hybrids (pink).

Figure S8: The phenotypic distance between hybrids and parent mean phenotype (potentially fitness optima) increases with the magnitude of divergence between parents.

Figure S9: The number of mismatched trait pairs 'snowballs' with the magnitude of phenotypic divergence between parents, but only in F1 hybrids.

Figure S10: Distribution of divergence–mismatch slopes ($\hat{\beta}$) for pairwise analyses.