HHS Public Access

Author manuscript

Nature. Author manuscript; available in PMC 2015 January 17.

Published in final edited form as:

Nature. 2014 July 17; 511(7509): 307–311. doi:10.1038/nature13301.

Genetics of ecological divergence during speciation

Matthew E. Arnegard[1,2], **Matthew D. McGee**[3], **Blake Matthews**[4], **Kerry B. Marchinko**[2], **Gina L. Conte**[2], **Sahriar Kabir**[2], **Nicole Bedford**[2], **Sara Bergek**[5], **Yingguang Frank Chan**[6], **Felicity C. Jones**[6], **David M. Kingsley**[6], **Catherine L. Peichel**[1], and **Dolph Schluter**[2]

[1]Human Biology and Basic Sciences Divisions, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA [2]Biodiversity Research Centre and Zoology Department, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada [3]Department of Evolution and Ecology, University of California, Davis, CA 95616, USA [4]Department of Aquatic Ecology, EAWAG, 79 Seestraße, Kastanienbaum, Switzerland [5]Department of Aquatic Resources, Swedish University of Agricultural Sciences, Drottningholm, Sweden [6]Department of Developmental Biology and Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA 94305, USA

Abstract

Ecological differences often evolve early in speciation as divergent natural selection drives adaptation to distinct ecological niches, leading ultimately to reproductive isolation. Though this process is a major generator of biodiversity, its genetic basis remains poorly understood. Here we investigate the genetic architecture of niche differentiation in a sympatric species pair of threespine stickleback fish by mapping the environment-dependent effects of phenotypic traits on hybrid feeding and performance under semi-natural conditions. We show that multiple, unlinked loci act largely additively to determine position along the major niche axis separating these recently diverged species. We also find that functional mismatch between phenotypic traits reduces growth of some stickleback hybrids beyond that expected from an intermediate phenotype, suggesting a role for epistasis between the underlying genes. This functional mismatch might lead to hybrid incompatibilities that are analogous to those underlying intrinsic reproductive isolation but that depend on the ecological context.

Correspondence and requests for materials should be addressed to M.E.A. (marnegar@fhcrc.org).

Supplementary Information is available in the online version of the paper.

Author Contributions M.E.A., C.L.P., and D.S. designed, planned, and oversaw the project. M.E.A. made the crosses, set up the experimental pond, and coordinated all field and laboratory research. M.E.A., K.B.M., S.K., N.B., and S.B. conducted field work and stable isotope analysis. M.D.M. measured functional morphological traits. B.M. and M.E.A. measured and analysed gut contents. S.K., D.S., and M.E.A. performed landmark-based morphometric analyses. M.E.A. analysed relationships between all traits and trophic variation. F.C.J., Y.F.C., and D.M.K. designed the SNP genotyping array. M.E.A., G.L.C., C.L.P, and D.S. analysed SNP genotypes. D.S. determined genealogy of the mapping population based on SNP genotypes. M.E.A., C.L.P., and D.S. performed linkage and QTL analysis. M.E.A., C.L.P., and D.S. tested the genetic architecture of niche divergence. M.E.A., C.L.P., D.S., M.D.M., B.M., and G.L.C. interpreted the results. M.E.A. wrote the paper with input from C.L.P. and D.S., who are co-senior authors. All other authors provided editorial comments and approved of the final version of the manuscript.

Author Information: The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper.

Adaptation of populations to contrasting environments is a key mechanism for the origin of species[1-4]. Under this process, divergent selection leads to high performance of individuals exploiting alternative ecological niches via cumulative changes in potentially many traits[5]. Such traits may include morphological phenotypes involved in locomotion and prey capture, behavioural traits that affect encounter rates with different prey types, and phenotypes conferring defence against niche-specific enemies[2]. The complex phenotypic basis of niche use and classic genetic models of adaptation predict that divergence in niche use will have a multilocus genetic architecture with a substantial additive component[6,7]. On the other hand, ecological divergence is often rapid and repeatable, and may occur with gene flow[4], raising the possibility that niche divergence might be accomplished by a few key genomic regions[8,9]. Although the genetics of putatively adaptive traits have been widely investigated, testing these alternative predictions requires understanding of how genetic changes combine to determine whole-organism performance in different ecological niches[10,11].

Because feeding success in different trophic niches depends on an individual's phenotype and environment, we designed a novel approach to evaluate predictions about its genetic basis. First, we used a semi-natural setting that contained a resource distribution resembling the natural environment and allowed individuals to move freely between trophic niches. We then identified the morphological traits contributing to niche use and feeding performance and genetically mapped these traits. To confirm that detected loci underlie trophic variation, we fit the relationship between niche use and genotypes underlying the traits. Finally, we tested the fit of alternative genetic hypotheses of additive, dominance, and epistatic effects to axes of feeding variation.

We mapped the genetic basis of niche divergence between the ‘benthic’ and ‘limnetic’ species of threespine stickleback fish (*Gasterosteus aculeatus* complex) coexisting in Paxton Lake, British Columbia, Canada. This pair is one of several that independently evolved in postglacial lakes in as few as 12,000 generations by adaptation to alternative niches and frequency-dependent natural selection from resource competition[12-14]. Benthic and limnetic stickleback exhibit nearly complete assortative mating[15] and differ in multiple morphological traits that adapt them to contrasting inshore and pelagic lake habitats, respectively[14,16-19]. Each species pair likely arose from a double lake invasion via the sea[12], followed by further divergence with gene flow[16,20]. Hybrids are intermediate in morphology and are outperformed by each parental species in the preferred parental habitats[14,21-23]. Little intrinsic postzygotic isolation has evolved between the species, as laboratory-reared hybrids are viable and fertile[16,21].

Niche use and hybrid feeding performance

Just before the breeding season in spring, we introduced 40 F1 hybrids to an outdoor experimental pond approximating the environmental conditions and contrasting habitats of Paxton Lake (Extended Data Fig. 1, Supplementary Discussion 1). We retrieved 633 F2 hybrid juveniles prior to their first winter and quantified diet variation among them using stable isotopes ($\delta^{13}C$ and $\delta^{15}N$; Fig. 1a). In nature, use of open water resources by limnetic individuals gives them a lower $\delta^{13}C$ and higher $\delta^{15}N$ than the more littoral-feeding benthics,

and isotope variation is correlated with foraging trait morphology[17]. Body size (length in mm) was our measure of F2 hybrid feeding performance, reflecting how successfully the juveniles acquired food resources and grew during the experiment (Supplementary Discussion 2). Rapid attainment of adult body sizes often confers fitness advantages to stickleback via the effects of size on insect predator avoidance[24], overwinter survival[25], male resource holding potential[26], and female fecundity[14].

Under our experimental conditions, the major axis of bivariate isotope variation among F2 hybrids (PC1, hereafter 'niche score'; Fig. 1a) was consistent with the primary axis of limnetic–benthic niche divergence based on isotope data from multiple stickleback species pairs in nature[17] (Supplementary Discussion 3). A secondary axis of feeding variation (PC2) was also identified. To illustrate variation in phenotype and diet across isotope space, we compared recently consumed prey items among F2 hybrids from three regions of isotope space (Fig. 1a), which we delineated using loess-predicted body size contours surrounding individuals with the largest (groups 'L' and 'B') or smallest ('A') average body sizes. 'B' individuals had isotope signatures resembling those of the benthic species in nature and consumed significantly more larval chironomids (Fig. 1b), on which wild benthics specialise[14,18,19]. In contrast, 'L' individuals had a pelagic $\delta^{13}C$ signature and preyed most heavily on the calanoid copepod, *Skistodiaptomus oregonensis* (Fig. 1c), a key planktonic prey item on which limnetics are specialised[14,17,19]. The small F2 hybrids in group 'A' fed predominantly on a symphypleonan springtail species (Fig. 1d), which is not a major dietary component of benthics or limnetics in the native lakes[14,17]. Thus, we refer to PC2 as 'diet deviation score' because it reflects variation independent of the typical limnetic–benthic feeding axis. The groups did not differ in their consumption of *Chydorus* sp., a littoral cladoceran (Fig. 1e). Additional analyses of consumed prey using all F2 individuals confirmed these feeding patterns (Extended Data Fig. 2, Supplementary Discussion 4).

Analysis of juvenile size variation across the entire isotope space revealed a saddle-shaped landscape (Fig 1a). F2 hybrids exploiting either the limnetic ('L') or benthic ('B') extremes of isotope distribution grew more than the other F2 hybrids, which either had intermediate niche scores and diets or exhibited an alternative feeding pattern ('A'). In nature, benthics grow to larger adult size than limnetics[14,16] in part because they differ in the age of sexual maturity[27], but in our experiment, mean body size was similar between the 'B' and 'L' F2 hybrids (Fig. 1a). This finding might have resulted from sampling the experimental fish as juveniles or from resource abundance differences between the experimental pond and Paxton Lake. The body size valley at intermediate niche scores (Fig. 1a) persists when F2 family identity is included as a covariate, which controls for variation in F2 hatching date and hence fish age (Fig. 1f). Considering the 20 largest F2 families, niche score was reasonably well fit by a quadratic regression model including the family covariate (R^2 = 33.2%; $F_{21,416} = 9.847$; $P < 2.20 \times 10^{-16}$). Although we found only suggestive evidence for a positive quadratic term in this model (coefficient estimate = 0.173 ± 0.101 s.e.; $P = 0.086$; Supplementary Discussion 5), within-family regression revealed that 16 families individually exhibited positive quadratic coefficients, indicating that the dip in body size at intermediate niche score is statistically significant ($P = 0.012$; two-sided binomial test).

Overall, these results support the hypothesis that F2 hybrids with an intermediate trophic phenotype suffered a growth disadvantage.

Morphological basis of niche divergence

A large number of phenotypic traits contribute to niche score variation. To determine this we measured nine functional morphological traits that play important roles in prey capture and retention, including craniofacial traits affecting the capacity to generate suction pressure, the speed and extent of jaw protrusion, and the retention of ingested prey items (Fig. 2)[16-19]. We additionally measured the x- and y- coordinates of 19 morphological landmarks indicating body and head shape[28] (Extended Data Fig. 3), which are expected to influence feeding performance. We used all-subsets linear regression to test effects of functional morphological traits and body shape coordinates, separately, on niche score. The best functional trait models ($0 \leq \Delta AIC \leq 2$) fitting niche score contained terms for three of the five components of the suction feeding index[18], two key oral jaw traits[19], and both gill raker counts[16,17] (Supplementary Table 1). The best models fitting niche score to body shape contained terms for 22 of 38 landmark coordinates[28] (Supplementary Table 2). Hereafter, we consider the traits included in the best models to be 'component traits' of niche divergence between Paxton benthics and limnetics.

Genetic architecture of niche divergence

We conducted quantitative trait locus (QTL) mapping on all measured morphological traits and found 76 significant QTL, including 41 QTL for 19 of the 29 component traits. The QTL exhibit small- to moderate phenotypic effect sizes (Supplementary Table 3). Component trait QTL occur on 14 of the 21 linkage groups (LG) in the threespine stickleback genome[29] (Extended Data Fig. 3), suggesting that multiple genetic factors contribute to niche divergence between Paxton benthics and limnetics. Both among LG and within certain LG, we find significant clusters of co-localized QTL (Extended Data Table 1, Supplementary Discussion 6), indicating close linkage of genes underlying different niche use traits or pleiotropic genetic effects. Nearly all QTL for the component traits occur in known regions of repeated genomic differentiation between sympatric benthic and limnetic species in multiple lakes[30].

To determine how these QTL contribute to benthic–limnetic niche divergence, we fit multiple-QTL mapping (MQM) models of niche score to genotypes at QTL for component traits. We selected only the single morphological QTL having the strongest estimated effect on niche score from each of the 14 LG containing QTL for component traits. Although this method is conservative and likely underestimates the number of loci underlying niche score, it avoids overly complex models involving multiple, linked loci within LG. We found additive allelic effects across multiple loci (Fig. 3). Seven of the 14 selected QTL significantly affected niche score (Extended Data Table 2). Two of these loci reside within clusters of co-localized QTL on LG 4 and 16 (Extended Data Fig. 3, Extended Data Table 1, Supplementary Discussion 6). However, effect sizes are distributed approximately evenly among the seven significant loci (percentage variance explained [PVE] = 1.16–3.74%; Extended Data Table 2). Next, we allowed all significant pairwise QTLxQTL interactions to

enter the model and followed this by backward elimination of non-significant terms. The resulting 'full' MQM model contained four pairwise interactions in addition to main effects representing eleven of the 14 morphological QTL (Extended Data Table 3).

To test the relative contributions of additive, dominance, and pairwise epistatic effects of these loci to niche score, we specified and compared three nested, general linear models at the markers nearest to the eleven significant QTL positions in the full MQM model. The 'additive' model contained only the additive effects of the eleven loci (Fig. 3b; adjusted R^2 = 15.8%; $F_{39,484}$ = 3.514; $P = 5.76 \times 10^{-11}$; AIC = 1533.42). By contrast, the 'additive +dominance' model contained both additive and dominance effects of these loci (adjusted R^2 = 14.4%; $F_{50,473}$ = 2.763; $P = 1.18 \times 10^{-8}$; AIC = 1551.80). Based on AIC, adjusted R^2, and results of a likelihood ratio (G) test, we conclude that dominance does not contribute significantly to the additive genetic model for niche score ($G_{(\text{add+dom, add})}$ = 3.625; P = 0.980). However, the 'full' model, with all additive and dominance effects across the eleven loci as well as the four significant pairwise epistatic interactions (Fig. 3c; adjusted R^2 = 20.6%; $F_{66,457}$ = 3.059; $P = 2.75 \times 10^{-12}$; AIC = 1526.37), provides a significantly better fit to the data than the additive model ($G_{(\text{full, add})}$ = 61.05; $P = 1.92 \times 10^{-4}$) or the additive +dominance model ($G_{(\text{full, add+dom})}$ = 57.43; $P = 1.41 \times 10^{-6}$). These results verify the prediction of a polygenic and largely additive basis to whole-organism niche use. F2 hybrids that grew the most during our study, reflecting high feeding performance, were either those individuals having the highest number of limnetic alleles across loci and the most limnetic-like phenotype and diet, or the highest number of benthic alleles and the most benthic-like phenotype and diet (Fig. 1b, c, Fig. 2a–d, Fig. 3a, Extended Data Fig. 2, 4–6). Pairwise genetic interactions also had a significant, albeit smaller, effect on niche score (*cf.* Fig. 3b, c). This is consistent with a role for epistasis in adaptation, though the importance of epistasis is likely underestimated because genetic interactions can be difficult to detect, particularly when they are weak or involve more than two interacting genetic factors[31,32].

Trait mismatch reduces growth

Analysis of the secondary axis of isotope variation, diet deviation score, provided additional evidence for non-additive effects. Inspection of phenotypes of group 'A' hybrids suggests that this subset of individuals experienced growth deficits (Fig. 1a) due to functional mismatch between certain traits. These 'A' individuals had distinctly limnetic-like lower jaw-opening inlevers (Fig. 2c, g), which contribute to the rapid jaw opening needed for successful strikes on evasive zooplankton such as *S. oregonensis*[19]. Yet, they also had reduced, or benthic-like, upper jaw protrusion (Fig. 2d, g), which is expected to decrease the efficiency of zooplanktivory[19]. The F2 hybrids in group 'A' also exhibited mismatches in other combinations of traits (Fig. 2a, Extended Data Fig. 4, 5, Supplementary Discussion 7). We predict that these conflicting trait combinations would reduce an individual's foraging success in both parental habitats, which could explain why these hybrids, as a group, were the smallest of any phenotypic class (Fig. 1a). This phenotypic interaction would imply epistasis for performance at underlying genes even if the phenotypic traits themselves have a largely additive genetic basis[33]. Such epistatic effects are expected to be manifest only in environments containing the divergent habitats to which the parental populations are adapted.

To further investigate, we applied the approach used in our genetic analysis of 'niche score' to 'diet deviation score'. Though many morphological traits underlie variation along this secondary feeding axis, MQM modeling revealed no statistically significant relationship between the QTL for these traits and diet deviation score. Consequently, we focused on strongest-effect QTL for the two traits that exhibited clear phenotypic mismatch in 'A' individuals, had strong effects on niche score (Extended Data Fig. 6, Supplementary Table 1), and are among the most divergent functional morphological phenotypes known in the species pair[19] (i.e., the QTL at 28.8 centimorgans on LG 4 for lower jaw-opening inlever, and 28.4 cM on LG 9 for upper jaw protrusion; Supplementary Table 3). Using genotypes at the marker nearest to each of these QTL, we found suggestive evidence of negative epistasis underlying diet deviation score in a two-way ANOVA of all F2 individuals (interaction term: $F_{4,598} = 2.254$; $P = 0.0621$).

Discussion

Though early in the speciation process, Paxton limnetic and benthic stickleback differ in many morphological traits. We show that many of these divergent traits contribute to variation in niche use and growth of juvenile F2 hybrids foraging freely in a semi-natural environment. Multiple genetic factors with largely additive effects, distributed across many chromosomes, underlie niche divergence along the limnetic–benthic resource gradient. Replacement of a limnetic allele by a benthic allele (or *vice versa*) at any of these loci shifts niche score in hybrids by roughly the same magnitude (Fig. 3a). We also found evidence for functional mismatch between phenotypic traits in hybrids that adopted an alternative feeding mode, accompanied by the slowest growth in the mapping population. This suggests that when multiple traits must function together, novel combinations of traits in hybrids might reduce performance below that expected for an intermediate phenotype. We predict that similar genetic architectures—involving multiple genomic regions each having a relatively small effect, coupled with the possibility of functional mismatch of some gene combinations—will be found for other complex, whole-organism phenotypes that depend on many component traits.

Our finding that niche divergence is determined by small-effect loci on more than half of the chromosomes in the threespine stickleback genome might not be expected in systems where gene flow still occurs. Theory indicates that loci with relatively large effect sizes under the strongest divergent selection will most readily resist gene flow[9], and these loci would be detected most readily by QTL mapping. In contrast, loci under weak divergent selection are less able to resist gene flow unless they are sufficiently tightly linked to other loci under sufficiently strong selection[9,34]. Nevertheless, our results are consistent with genome scans of ecologically divergent populations of several organisms, including threespine stickleback, which typically show differences in many regions distributed across the genome[29,30,35-38]. It is possible that the broadly-distributed genetic architecture of niche divergence in the Paxton Lake species pair has arisen from strong, multifarious divergent selection[39] acting simultaneously on the numerous traits that underlying adaptation to shallow benthic and open-water habitats. An intriguing possibility is that this broadly-distributed genetic architecture results from segregation of ancestral variation that arose during periods of allopatry[30,38,40].

Our results contribute to a better understanding of the genetics of environment-dependent reproductive isolation during ecological speciation, because divergence in traits underlying niche use reduces the fitness of intermediate phenotypes, including hybrids, when intermediate environments are uncommon or unprofitable[22,41,42]. Environment-dependent reproductive isolation accompanies the earliest stages of adaptation and may drive evolution of additional forms of reproductive isolation[2,43]. If rapid growth of certain threespine stickleback juveniles (e.g., those in groups 'L' and 'B') has positive fitness consequences, then disruptive selection found in the saddle-shaped landscape of body size (Fig. 1a) might reflect selection against intermediate hybrid phenotypes along the major axis of niche differentiation (Fig. 1f). This pattern corroborates the results from transplant experiments in the native lake showing a growth disadvantage in intermediate hybrids, relative to the two parental species in their native habitats[14,21,22]. Thus, our results on the genetics of divergence in niche use and whole-organism performance suggest that the underlying genetic basis of extrinsic postzygotic reproductive isolation between benthic and limnetic stickleback is largely additive. This contrasts with the genetics of environment-independent or 'intrinsic' postzygotic isolation, the evolution of which is well described by the Bateson-Dobzhansky-Muller model and is largely caused by negative epistatic interactions between loci[3,8,44]. Nevertheless, our results show that mismatch of oral jaw traits reduces feeding performance of some F2 hybrid stickleback beyond that expected from additive genetic effects alone. Functional mismatch between traits might therefore represent an environment-dependent counterpart to the deleterious inter-molecular interactions often associated with intrinsic postzygotic isolation[44]. As our results suggest, hybrids that are phenotypically mismatched for ecological performance traits may be produced inevitably as the process of ecological speciation unfolds, thereby contributing to the further evolution of reproductive isolation.

Methods

Experimental Pond and F2 Hybrid Population

We used an outdoor experimental pond at the University of British Columbia (UBC, Vancouver, Canada) containing shallow, littoral and deep, open-water habitats (Extended Data Fig. 1, Supplementary Discussion 1). Four *in vitro* crosses were made between unique, wild-caught F0 Paxton benthics and limnetics. F0 females were benthic for two crosses and limnetic for the other two. After raising the F1 families in separate aquariums[28], we stocked the pond on 17-March-2008 with five F1 adults per sex and family (n = 40). No food or nutrients were added to the pond after stocking (Supplementary Discussion 1). During these procedures, fin clips were removed from F0 and F1 individuals and stored in 95% ethanol for genetic analysis.

F1 adults mated freely in the pond to produce an F2 population. We collected 633 juvenile F2s in fall 2008 (October 5–21), when rapid stickleback growth begins to slow[49] and before any overwintering mortality[25]. F2 hybrids were captured using un-baited, fine-mesh minnow traps set in all parts of the pond. During fieldwork, we selected 99 F2 hybrids (in a blind manner) taken from traps deployed no longer than two hours. Each of these individuals was euthanized and preserved immediately for subsequent analysis of consumed food items in its

digestive tract. All other individuals were housed in tanks and processed within 24 hours. F1 adults were readily excluded by size.

Niche Use by F2 Juveniles

We euthanized F2 hybrids with an overdose of buffered tricaine methanesulfonate and rinsed them in distilled water. Caudal and left pectoral fins were removed and stored in 95% ethanol for genetic analysis. Using a clean scalpel, we sampled white skeletal muscle from the posterior left flank, excluding any skin or bone, and immediately lyophilised the samples in a BenchTop Manifold Freeze Dryer (Millrock Technology Inc.). Fish were fixed in 7.5% formalin (phosphate-buffered) for one month, and then transferred to 40% isopropanol.

We homogenized the freeze-dried muscle samples and took 0.8–1.2 mg subsamples, which were enclosed in tin capsules (Elemental Microanalysis Limited), placed in 96-well microplates, and stored in a vacuum-sealed desiccator. The subsamples were assayed for stable isotopes of carbon (^{12}C and ^{13}C) and nitrogen (^{14}N and ^{15}N) at the U.C. Davis Stable Isotope Facility in one continuous run in January 2009. Measurements were made with a PDZ Europa ANCA-GSL elemental analyser interfaced to a PDZ Europa 20-20 mass spectrometer (Sercon Ltd.); these are expressed as scaled isotope ratios, in parts per thousand (‰) relative to Pee Dee Belemnite or atmospheric N_2, using the standard delta notation ($\delta^{13}C$ or $\delta^{15}N$)[45,50]. We performed principal component analysis (PCA) on the bivariate isotope data using the function ‘prcomp’ in R (ver. 2.14.0)[51], after scaling both $\delta^{13}C$ and $\delta^{15}N$ to unit variance. The first PC axis (PC1, ‘niche score’) explained 56.5% of total variance in isotope space (first eigenvalue $\lambda_1 = 1.13$); PC2 (‘diet deviation score’) explained the remaining 43.5% of variance ($\lambda_2 = 0.87$).

The $\delta^{13}C$ and $\delta^{15}N$ signature of skeletal muscle is an integrative measure of an individual's long-term diet (i.e., several weeks to months)[17,45,50,52-57]. We compared these signatures to a direct measure of F2 hybrid feeding activity immediately (i.e., several hours) prior to capture, which we quantified via counts of ingested food items in 99 F2 hybrids. Food items in the digestive tract of each individual were counted after being sorted into the following 14 categories: adult aquatic snails (class Gastropoda); snail eggs; ostracods (class Ostracoda); calanoid copepods, all identified as *Skistodiaptomus oregonensis* (order Calanoida); cyclopoid copepods (order Cyclopoida); *Chydorus* sp. (order Cladocera); *Sida* sp. (Cladocera); *Gammarus* sp. (order Amphipoda); water mites (unranked taxon Hydrachnidiae, suborder Prostigmata); caddisfly larvae (order Trichoptera); chironomid larvae (family Chironomidae); beetle larvae (order Coleoptera); springtails (order Symphypleona, subclass Collembola); and all other terrestrial and surface-dwelling (i.e., neustonic) insects, combined. Four categories (chironomids, *S. oregonensis*, springtails, and *Chydorus*) accounted for >98% of all ingested food items across individuals.

We used body size of fish at capture (length in mm) as a measure of feeding performance (Supplementary Discussion 2). Body size was taken as distance between morphometric landmarks 1 and 13 (Extended Data Fig. 3). Size variation across the isotope landscape was visualized as the loess (local 2nd degree polynomial) regression surface of body size on $\delta^{13}C$ and $\delta^{15}N$, estimated using the R function ‘loess’[51] (span = 0.75). A plot of this surface suggested isotopically-distinct regions of extreme performance, reflected by especially large

or small average body size of the juveniles in each region (Fig. 1a). To facilitate comparison of diet and morphology among regions, we used contours of the loess regression fit to establish boundaries around individuals of largest or smallest predicted body size. Each boundary was the most extreme predicted size contour enclosing a unique set of individuals numbering approx. 15% of the distribution (n = 94–95 per region). Thus, region 'B' contained individuals of large average size near the performance peak at high δ^{13}C–low δ^{15}N (Fig. 1a), whereas region 'A' contained individuals with the smallest average size observed overall, at low δ^{13}C–low δ^{15}N. In these cases, simple use of appropriate contours allowed straightforward application of the 15% criterion. We wanted the third region ('L') to also include individuals that grew to large average size (like 'B') but were instead located around the performance peak at low δ^{13}C–high δ^{15}N. With 'L', however, a second criterion (minimization of PC1) was required to define a boundary that both contained an outer 'quantile' (15%) of the predicted performance distribution and retained isotopic distinctiveness from other regions. Specifically, the boundary of 'L' was the maximum loess-predicted size contour enclosing 15% of the distribution (around the low δ^{13}C–high δ^{15}N peak) after limiting this region to PC1 < 0.045. Next, we investigated variation in recent feeding activity among these categories of F2 hybrids with Kruskal-Wallis tests (R function 'kruskal.test'[51]) for differences in counts of ingested food items (Fig. 1b–e).

To test robustness of the performance valley at intermediate niche score (Fig. 1a), we fit body size to niche score with a cubic spline function including F2 family identity (indicating offspring of each unique F1×F1 pairing) as a covariate. Doing so accounts for among-family variation in F2 age at capture due to variable F1 breeding times, assuming unique F1 pairs mated only once. Supporting this assumption, we found no deviations from unimodal size frequency distributions in F2 families, judged by visual inspection and Hartigan's dip test[58] (R package 'diptest'[59]; 2,000 replicates per Monte Carlo simulation; each P > 0.175). Thus, cubic splines were estimated in 'glms' ver. 4.0 (www.zoology.ubc.ca/~schluter/wordpress/software/#spline)[48], using the 20 largest F2 families (full sibs per family: n = 12–48). Using the best smoothing parameter (i.e., λ with lowest cross-validation score), we obtained standard errors of predicted body sizes (1,000 bootstrap replicates). We also evaluated robustness of the performance valley by quadratic regression of body size on niche score, again using the 20 largest families and the family identity covariate (Supplementary Discussion 5). The regression was repeated using only individuals for which PC2<0 to ensure that presence of the size valley did not depend solely on unusually small individuals with PC2≥0, including those in region 'A'.

Morphological Trait Measurements

Three classes of morphological traits known to differ between wild Paxton benthics and limnetics (Supplementary Discussion 8) were measured: (1) morphometric traits reflecting body shape; (2) defensive armour traits; and (3) single or composite functional traits (head and jaw) with described roles in feeding[16-19,60]. We measured shape using the geometric morphometric approach of previous stickleback studies[28,61]. Fixed specimens were stained (48 h) in 1% aqueous KOH with 0.005% w/v Alizarin Red S (Merck KGaA) and de-stained in 40% isopropanol. A Nikon D1H camera and three strobe lights were used to make a digital image of the right side of each specimen alongside a ruler. We recorded x- and y-

coordinates of 19 morphological landmarks from these images using 'tpsDig' ver. 2.12[62] (Extended Data Fig. 3). We scaled, rotated, and superimposed landmark configurations using Generalized Procrustes analysis[63] (R package 'shapes'[64]), after which we used a standard approach to correct for a specimen bending artefact caused by fixation[28,61,65] (Supplementary Discussion 9). Resulting x- and y-coordinates were treated as individual traits when analysing relationships between shape and stable isotopes and performing quantitative trait locus (QTL) mapping.

Images enabled use of a simple ordinal scale for rapid characterization of three armour traits: pelvic girdle (right side of body) and first and second dorsal spines. These traits received a score of 0 when absent; 2 when present; and 1 when expressed at an intermediate size between these two categories. 'Well-developed' lateral bony plates[16] were also counted along the right flank (i.e., any plate whose height was judged to be at least 1/3 the individual's body depth at that plate).

We measured functional morphological traits using methods previously applied to stickleback[18,19,66]. Gill rakers on the left outer branchial arch were counted under a dissecting microscope after removing the arch and associated cartilage from the opercular cavity. Any stained protuberance was counted as either a long or short gill raker according to position (Fig. 2e). After clearing specimens by immersion in 30% w/v sodium borate with 1% w/v trypsin until translucent, we measured five component traits of the suction feeding index[18,19] (Fig. 2f): anterior epaxial muscle height (E_H) and width (E_W); neurocranium outlever length (N_{OL}); buccal cavity length (B_L); and gape (G). Suction index (S.I.) was calculated as $S.I. = (E_W \cdot E_H^2)/(3 \cdot B_L \cdot G \cdot [N_{OL} - (B_L/2)])$. Lastly, we measured upper jaw (premaxillary) protrusion length and lower jaw-opening inlever length (Fig. 2g)[19].

All functional morphological traits were corrected for body size (length) except long and short gill raker counts, which were uncorrelated with size. We used standardized major axis regression (function 'sma' in R package 'smatr'[67]) to test for differences in allometric scaling relationships of these traits between F2 hybrids in the mapping population and wild Paxton benthics and limnetics. This revealed no evidence of allometric differences between the experimental fish and natural populations (likelihood ratio tests, 2 d.f. each: $0.09 < P < 0.56$). Therefore, the traits were size-corrected by expressing them as residuals from ordinary least squares regression of each trait on body size[68].

Identifying 'Component Traits' of Niche Use

To determine which morphological traits predict variation along the primary trophic axis, we first regressed niche score on each trait separately using simple linear (least squares) regression. Armour traits were excluded, because we had no *a priori* evidence of specific influences on trophic variation (Supplementary Discussion 8). Similarly, we excluded suction feeding index, because each of its component traits was being considered. Scatterplots indicated that the data conformed reasonably well to parametric statistical assumptions. Only traits from the significant univariate regression models were considered further. Taking all such traits to be candidate explanatory variables, we performed all-subsets (multiple linear) regression[69] of niche score on the candidate traits using the R function 'leaps'[70]. This function returns and orders the best models based on Mallows' C_P[71],

which we converted to the Akaike Information Criterion (AIC)[72]. Due to partial redundancy between some of the functional traits and craniofacial landmarks, we considered functional morphology and shape trait classes separately when performing these exhaustive searches for trait subsets best predicting niche score.

The difference between AIC scores of the 'best' (AIC = 0) and 'i^{th} best' models is denoted ΔAIC. We considered all models with ΔAIC ≤ 2 to be statistically indistinguishable from the overall 'best' model identified for given class of traits[73]. Consequently, the full suite of morphological traits for which this approach found similarly strong, within-class evidence of an effect on niche score was the union of explanatory variables among all well supported models ($0 < AIC \leq 2$) across the two trait classes (functional morphology and shape; Supplementary Tables 1 and 2, respectively). Hereafter we refer to this full suite of traits as the 'component traits' of niche use. We repeated this entire procedure for identifying the morphological traits that influence a trophic axis using 'diet deviation score' as the response variable, instead of 'niche score'.

Genotyping and Pedigree Analysis

We isolated genomic DNA from fin tissue samples taken from the eight F0 founders, 40 F1 adults, and 633 F2 juveniles, using Proteinase K digestion, phenol-chloroform extraction, and ethanol precipitation[74]. We re-suspended DNA in 30μL of TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0) and diluted an aliquot of each sample to a concentration of approx. 25 ng/μL based on PicoGreen® assay (Life Technologies). All F0, F1, and F2 individuals were genotyped at 408 single nucleotide polymorphism (SNP) markers[30], which are distributed across the *G. aculeatus* genome and were polymorphic in our mapping population (Supplementary Table 4). Genotyping was performed with Illumina's GoldenGate assay at the Fred Hutchinson Cancer Research Center (Seattle, WA, USA), using GenomeStudio software (Illumina Inc.) to score genotypes.

We used a Bayesian parentage assignment algorithm[75] (R package 'MasterBayes'[76]) and all SNP genotypes to estimate F1 parentage of every F2 individual. Posterior probabilities of correct assignments of F2 hybrids to their estimated pair of F1 parents were high (mean ± s.d. = 0.999 ± 0.020). Assignments of F1 hybrids to known F0 parents were verified (posterior probability = 1 in every case).

Linkage Analysis and QTL Mapping

Among the 728 F2 hybrids collected in total from the experimental pond (n = 633 juveniles this study; n = 95 adult males collected in spring 2009), we used all individuals in F2 families with at least ten full sibs (range: 10–53) to construct a linkage map. This was done using JoinMap® ver. 3.0 ('cross pollinator' population code)[46]. All obtainable pairwise (between-SNP) recombination frequencies and associated $logarithm_{10}$ of odds (LOD) scores were computed for each F2 family, separately. We created a single pairwise data file by concatenating recombination frequencies and LOD scores across families and used this to produce the map (Supplementary Table 4).

We performed QTL mapping on all measured traits in R/qtl[47], using all F2 families from the linkage analysis. Retaining all families after excluding F2 hybrids collected in spring 2009 required a reduction in minimal acceptable family size (to eight full sibs). Accordingly, our dataset for QTL mapping consisted of 530 F2 hybrids in 29 F2 families (range: 8–48 sibs per family). Using R/qtl function 'scanone' we performed interval mapping on each trait with Haley-Knott regression and F2 family identity as a covariate. We conducted 10,000 permutations per trait to determine the genome-wide LOD threshold for significant QTL at $\alpha = 0.05$[47]. Resulting LOD thresholds ranged from 3.51–3.88 across traits (mean = 3.63). For every QTL, we estimated position of the peak LOD score in centimorgans (cM) and 1.5-LOD confidence interval around the peak[77]. R/qtl function 'fitqtl' was used to estimate percentage of phenotypic variance explained by each QTL, and 'find.marker'[47] was used to identify the nearest SNP.

Genetic Architecture of Niche Divergence

We investigated effects of the discovered morphological QTL on niche divergence between Paxton benthics and limnetics as follows. First, we considered only QTL underlying component traits of niche use. From these QTL we selected the single QTL per linkage group (LG) with the highest LOD score among niche use component traits mapping to that LG. This procedure identified 14 candidate morphological QTL (on different LG) with hypothesized genetic effects on niche score. Repeating this procedure for 'diet deviation score', we identified 15 QTL on different LG with hypothesized effects on this secondary trophic axis.

To model cumulative effects of the 14 candidate QTL on the niche score of F2 hybrids, we specified candidate QTL positions using R/qtl function 'makeqtl'. We then used 'fitqtl' to fit a multiple-QTL mapping (MQM) model of the main effects of the QTL on niche score (Extended Data Table 2). Next, we found all significant pairwise QTLxQTL interactions among the candidate loci by applying the 'addint' function[47] to the 14 candidate QTL. We added these interactions to the main-effects-only MQM model and performed backwards stepwise elimination of non-significant terms until arriving at the full MQM model (Extended Data Table 3). At every step, 'fitqtl' was used for model fitting with the Haley-Knott method and the F2 family covariate.

We repeated this modelling procedure for diet deviation score and its 15 candidate QTL. In this case, we found nine significant QTLxQTL interactions using 'addint', but these interactions were not accompanied by significant main effects (data not shown). Consequently, all further model comparisons were focussed on testing genetic effects on niche score.

Using the full MQM model, we tested the importance of additive, dominance, and pairwise epistatic effects in the genetic architecture of niche divergence between Paxton benthics and limnetics. In R/qtl, we imputed genotypes at the SNP marker nearest to each QTL in this model using the Kosambi mapping function[78]. Subsequent model comparisons of QTL effects were performed with the linear model fitting function 'lm'[51] in R. We specified the 'full' genetic model by coding genotypes as categorical data and including all additive, dominance, and pairwise epistatic effects detected in R/qtl for the full model (Extended Data

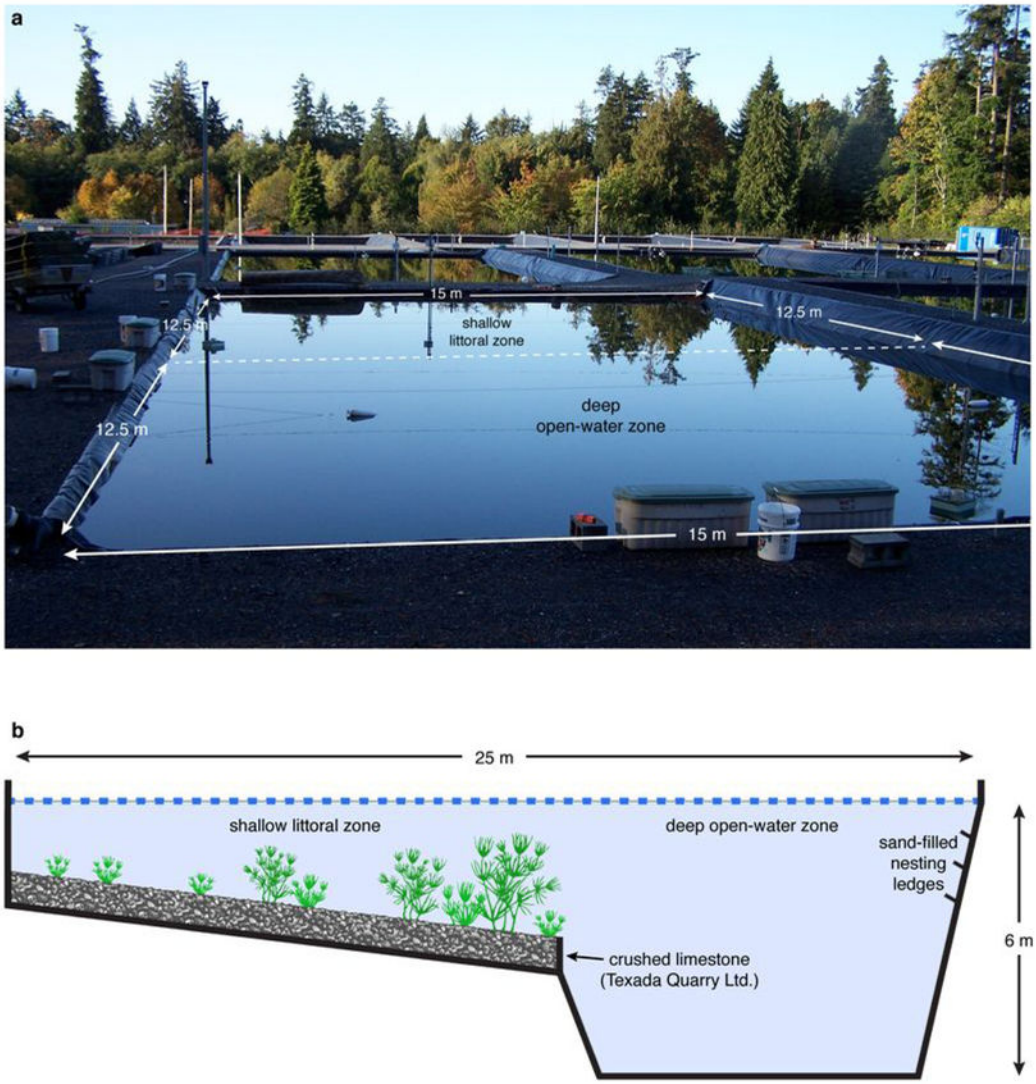
Table 3). Using this genotype coding, an ‘additive+dominance’ model was specified by including only main effects from the ‘full’ model. An ‘additive’ model was specified by coding genotypes as integer number of benthic alleles and including only main effects in the ‘full’ model. Each model again included the F2 family covariate. Models were compared using AIC and adjusted R^2, which penalize models with excessive numbers of terms[72,73,79,80], and using likelihood ratio tests (R function ‘lrtest’[81]).

Animal care, sample size determination, and data blinding

All field and laboratory procedures were approved by the University of British Columbia Animal Care Committee (protocols A07-0293 and A11-0402) and the Fred Hutchinson Cancer Research Center Institutional Animal Care and Use Committee (protocol 1797). Target sample size of F2 hybrids was determined to minimize bias when detecting QTL (the “Beavis effect”) and reduce sampling error for estimated QTL effect sizes. Realized sample sizes for QTL mapping (n = 473–530 F2 hybrids) were sufficient to minimize QTL detection bias and sampling error for effect sizes for every trait considered[47,82]. All reductions in sample size (from n = 633 juveniles collected) occurred in an unbiased fashion, because sample exclusion was based solely on missing phenotype or genotype data, or having too few full sibs in the collection. To avoid sampling and measurement biases, sample identities were not revealed to the authors and technicians who performed phenotypic measurements or genotyping until after all data collection was completed. The 99 F2 hybrids allocated for analysis of consumed food items were also selected in a blind manner during fieldwork.

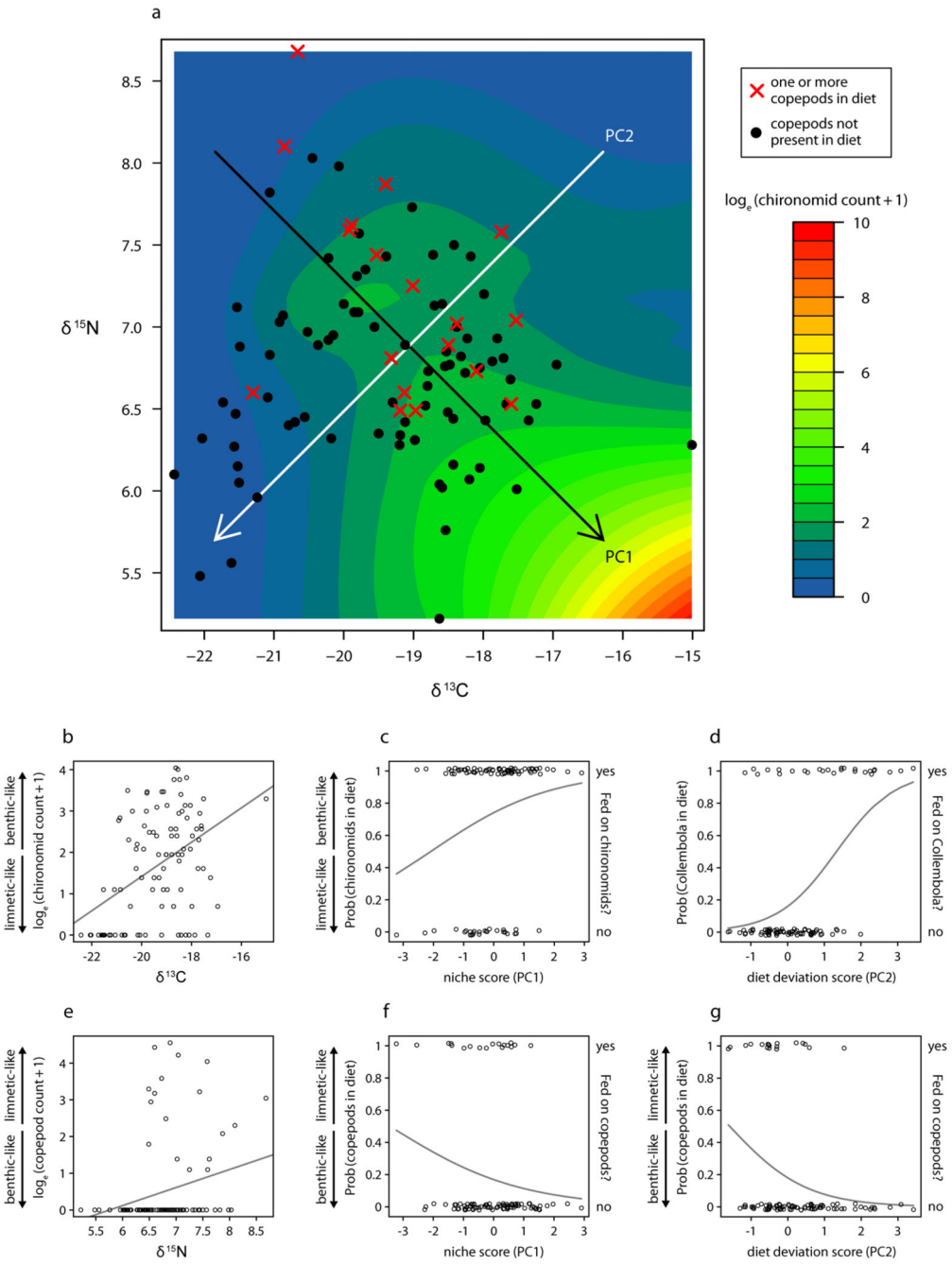
Constructed to help clarify how the two-dimensional isotope distribution of F2 hybrids related to other patterns of variation, F2 groups ‘B’, ‘L’, and ‘A’ each contained approximately 15% of all F2 hybrids collected. With application of a second criterion minimizing PC1 scores in just group ‘L’, the 15% inclusion criterion yielded the largest number of individuals per group without compromising group distinctiveness. Patterns of feeding and morphological variation among groups were robust to alternative body size thresholds (Supplementary Discussion 4), and they were confirmed by analyses using all individuals (Extended Data Fig. 2 and 6). Moreover, all available data were used for stable isotope PCA, component trait determination, QTL analysis, and genetic modelling, so results of these analyses did not depend on how F2 hybrids were categorized to understand the performance landscape.

Extended Data



Extended Data Figure 1. Experimental pond used in the study

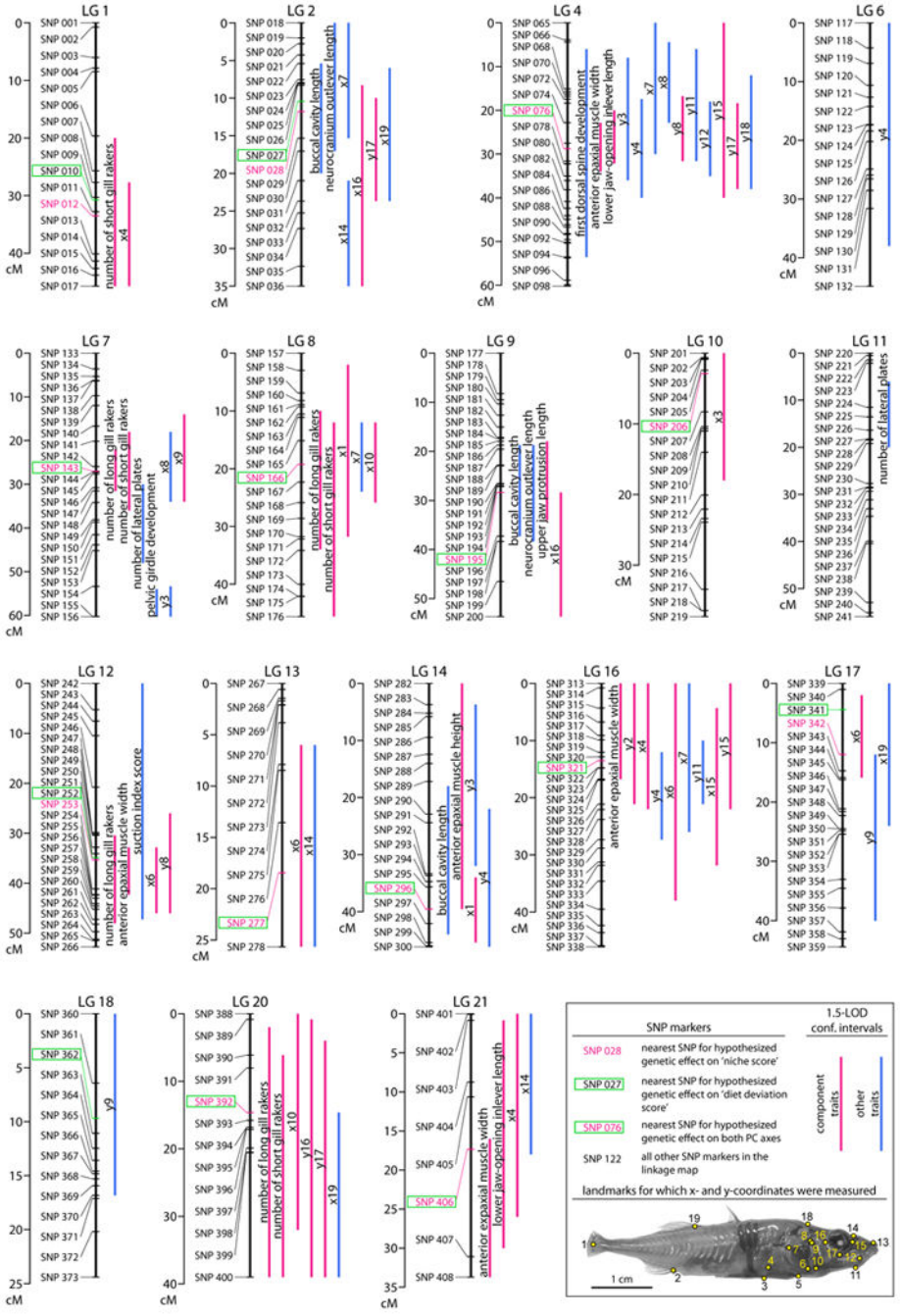
(a) Photograph of Pond #4 at the experimental pond facility of the University of British Columbia (Vancouver, B.C., Canada), taken in fall of 2008, during the collection of F2 juveniles. **(b)** Schematic of the pond profile. See Supplementary Discussion 1 for details on pond history prior to this study.



Extended Data Figure 2. Feeding patterns in relation to isotope signatures

Plots show relationships between ingested prey counts from all available F2 hybrids (n = 99) and stable isotope data. **(a)** Loess-smoothed surface (span = 0.75, 2nd degree polynomials) of predicted chironomid counts plotted on original isotope axes ($\delta^{13}C$, $\delta^{15}N$). As with all other count data plotted here, counts were transformed as $\log_e$(chironomids+1) and mapped according to the coloured scale. PC1 (black arrow) and PC2 (white) are based on the entire isotope distribution (Fig. 1a). Individuals are plotted as points according to presence (×) or absence (•) of calanoid copepods in their digestive tracts. **(b–g)** Linear or logistic regression,

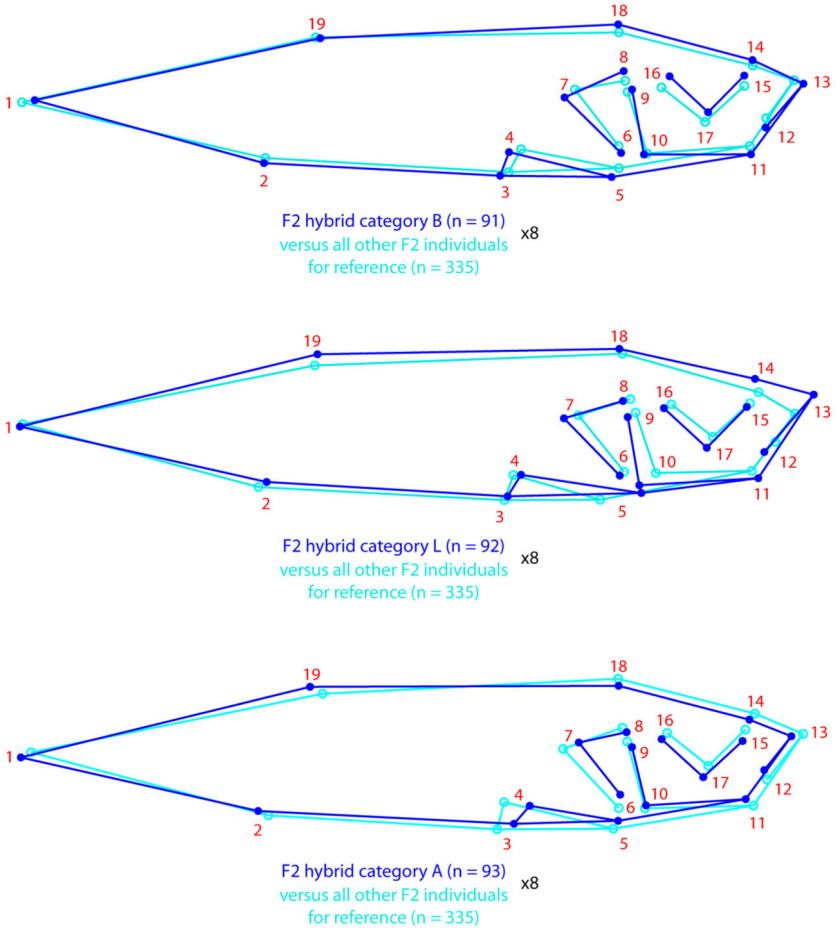
accordingly, of ingested prey count or presence/absence data (transformed as above) on the different axes through isotope space. Results: **(b)** chironomid count against $\delta^{13}C$, linear regression, slope est. = 0.415, $R^2 = 0.199$, $F_{1,97} = 24.1$, $P = 3.70\times10^{-6}$; **(c)** chironomid presence against niche score, logistic regression, slope coeff. = 0.504, z = 2.23, $P = 0.0255$; **(d)** Collembola presence against diet deviation score, logistic regression, slope coeff. = 1.25, z = 4.26, $P = 2.03\times10^{-5}$; **(e)** calanoid count against $\delta^{15}N$, linear regression, slope est. = 0.492, $R^2 = 0.0608$, $F_{1,97} = 6.28$, $P = 0.0139$; **(f)** calanoid presence against niche score, logistic regression, slope coeff. = -0.463, z = -1.84, $P = 0.0651$; **(g)** calanoid presence against diet deviation score, logistic regression, slope coeff. = -0.958, z = -2.67, $P = 0.00766$.



Extended Data Figure 3. Linkage map showing QTL for all traits

All *G. aculeatus* chromosomes are represented by linkage groups (LG) in the complete linkage map for this study (LG and chromosomes use same numbering[29]; LG with no mapped QTL are omitted here). Map distances indicated with scale at left of each LG in centimorgans (cM). Coloured bars (at right) are 1.5-LOD confidence intervals for QTL position (red bars, 'component traits of niche use'; blue bars, 'other traits'; Supplementary Table 3 provides LOD scores, map positions of LOD peaks, and effect sizes). The given SNP identifiers (IDs) are only for reference to Supplementary Table 4, which provides

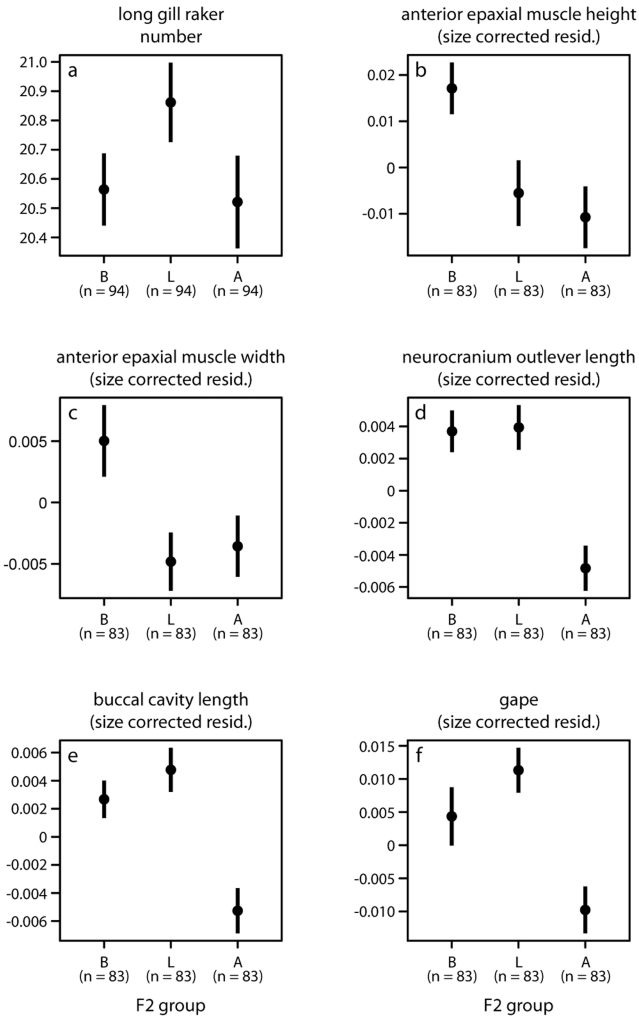
published SNP data[30]. For clarity every other ID is omitted for SNP 066–098, even though these markers are present in the map. Markers closest to candidate QTL for genetic model comparisons are highlighted: red text, nearest to candidate QTL for niche score; green boxes, diet deviation score. Numbered traits are x- and y- coordinates of morphometric landmarks (indicated on fish photo): 1, posterior midpoint caudal peduncle; 2, anterior insertion anal fin at first soft ray; 3, posteroventral corner ectocoracoid; 4, posterodorsal corner ectocoracoid; 5, anterior-most corner ectocoracoid; 6, anteroventral corner opercle; 7, posterodorsal corner opercle; 8, dorsal edge opercle-hyomandibular boundary; 9, dorsal-most extent preopercle; 10, posteroventral corner preopercle; 11, anterior-most extent preopercle along ventral silhouette; 12, posteroventral extent maxilla; 13, anterodorsal extent maxilla; 14, suture between nasal and frontal bones along dorsal silhouette; 15, anterior margin orbit; 16, posterior margin orbit; 17, ventral margin orbit (landmarks 15–17 placed in line with vertical or horizontal midpoint of eye); 18, posterior extent supraoccipital along dorsal silhouette; 19, anterior insertion dorsal fin at first soft ray.



Extended Data Figure 4. Shape variation among F2 hybrid groups

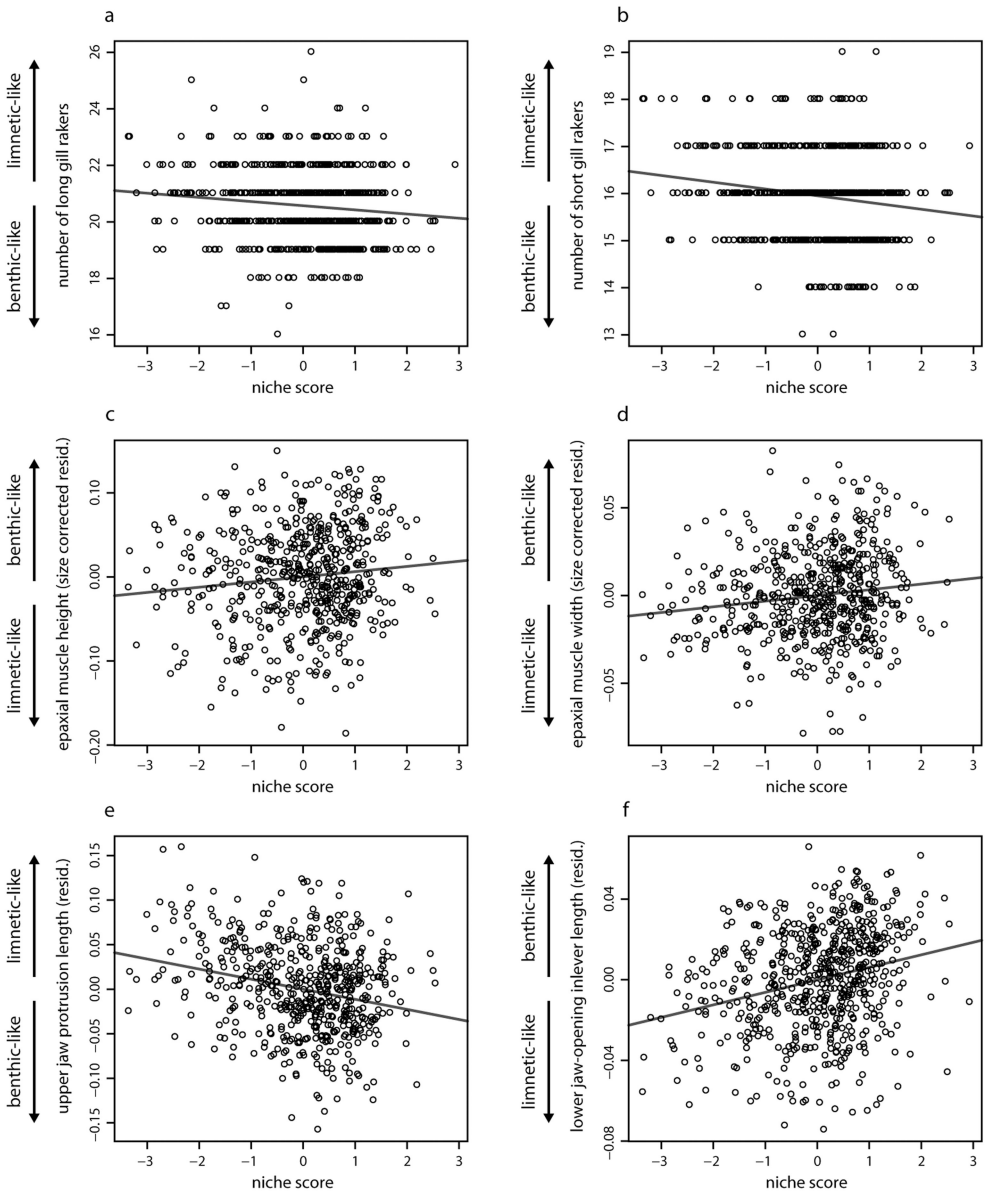
Each panel compares mean body shape of individuals in one of three categories of F2 hybrids (‘B’, ‘L’, or ‘A’; shown in dark blue) with the mean shape of all other F2 hybrids (category membership shown in Fig. 1a). Using data for 19 Procrustes-superimposed and

unbent landmarks (Extended Data Fig. 3), wireframe diagrams were produced and plotted in MorphoJ ver. 1.04a, based on discriminant function analysis (Supplementary Discussion 7). Shape differences are magnified eightfold for easier visual comparison. Sample sizes: n = 91 ('B'), n = 92 ('L'), n = 93 ('A'), n = 335 (all other F2 hybrids). Supplementary Discussion 7 describes patterns of variation in several specific features of shape, which can be interpreted from these data.



Extended Data Figure 5. Variation of additional traits among F2 hybrid groups

Means (± 1 s.e.m.) of F2 hybrids in groups 'B', 'L', and 'A' (Fig. 1a) are shown for the following traits: **(a)** number of long gill rakers (ANOVA, $F_{2,279}$ = 1.756, P = 0.175); **(b)** residual epaxial muscle height ($F_{2,246}$ = 5.219, P = 0.00603); **(c)** residual epaxial muscle width ($F_{2,246}$ = 4.223, P = 0.0157); **(d)** neurocranium outlever length ($F_{2,246}$ = 13.36, P = 3.10×10^{-6}); **(e)** residual buccal cavity length ($F_{2,246}$ = 12.26, P = 8.42×10^{-6}); **(f)** residual gape ($F_{2,246}$ = 7.974, P = 4.41×10^{-4}). Traits are illustrated in Fig. 2 e–g. The data conformed reasonably well to parametric statistical assumptions, so ANOVA was used to test trait variation among categories.



Extended Data Figure 6. Relationships between F2 hybrid functional morphology and niche score

For key functional morphological traits known to differ between Paxton benthics and limnetics, trait data from all available F2 hybrids are plotted against niche score and fitted with linear regression (raw data for gill raker counts; size-corrected data for other traits): **(a)** number of long gill rakers (R^2 = 0.0146; $F_{1,629}$ = 9.32; P = 0.00236); **(b)** number of short gill rakers (R^2 = 0.0253; $F_{1,629}$ = 16.30; P = 6.06×10^{-5}); **(c)** residual epaxial muscle height (R^2 = 0.0125; $F_{1,552}$ = 7.00; P = 0.00804); **(d)** residual epaxial muscle width (R^2 = 0.0189; $F_{1,552}$ = 10.61; P = 0.00119); **(e)** residual upper jaw protrusion length (R^2 = 0.0580; $F_{1,552}$ = 34.00; P = 9.40 × 10^{-9}); **(f)** residual lower jaw-opening inlever length (R^2 = 0.0660; $F_{1,615}$ = 43.43; P = 9.45 × 10^{-11}). Traits illustrated in Fig. 2 e–g. Directions of benthic–limnetic divergence in Paxton Lake (arrows at left of plots, here and Fig. 2 a–d) are based on

previously published studies[16,18,19], combined with validating counts of long and short gill rakers for this study (data not shown).

Extended Data Table 1
Goodness-of-fit tests for genomic distribution of QTL

Expected numbers of QTL on LG, under a random distribution null hypothesis (simple proportional model), were based on the known size (2nd column from left) and gene content (4th col.; predicted number of coding + non-coding genes) of corresponding chromosomes (obtained from Ensembl genome browser 17-July-2013; based on initial *G. aculeatus* assembly, Broad S1, Feb. 2006). Observed numbers and percentages of QTL for all measured traits or only component traits given in last 4 columns (at right). Results of all tests support the alternate hypothesis of QTL clustering:, $\chi^2_{20}=45.17$ P = 0.0016 (for all traits, with a null expectation based on chromosome size); $\chi^2_{20}=55.76$, P = 0.0002 (all traits, based on gene number); $\chi^2_{20}=34.87$, P = 0.0219 (component traits, based on chromosome size); $\chi^2_{20}=39.12$, P = 0.0083 (component traits, based on gene number); P-values estimated by Monte Carlo simulation (10,000 replicates each) due to small expected counts for many LG. Standardised residuals were used to identify LG with QTL counts deviating from random expectation (Supplementary Discussion 6): *P < 0.05 (expectation based on chromosome size); †P < 0.05 (gene content). Sums for size (‡) and gene content (§) exclude unassembled regions of the genome.

LG	Size (b.p.)	Percentage of assembled genome's size	Coding + non-coding genes	Percentage of total count of coding + non-coding genes	QTL for all traits	Percentage of total QTL count (all traits)	QTL for component traits only	Percentage of total QTL count (component traits)
1	28,185,914	7.033 %	1,257	6.581 %	2	2.632 %	2	4.878 %
2	23,295,652	5.812%	860	4.502 %	7	9.210%	2	4.878 %
3	16,798,506	4.191 %	932	4.880 %	0	0%	0	0%
4	32,632,948	8.142%	1,323	6.926 %	13*,†	17.105%	5	12.195%
5	12,251,397	3.057 %	732	3.832 %	0	0%	0	0%
6	17,083,675	4.263 %	721	3.775 %	1	1.316%	0	0%
7	27,937,443	6.970 %	1,320	6.911 %	7	9.210%	3	7.317%
8	19,368,704	4.833 %	881	4.612%	5	6.579 %	4	9.756 %
9	20,249,479	5.052 %	1,012	5.298 %	4	5.263 %	2	4.878 %
10	15,657,440	3.907 %	815	4.267 %	1	1.316%	1	2.439 %
11	16,706,052	4.168%	1.058	5.539 %	1	1.316%	0	0%
12	18,401,067	4.591 %	1.003	5.251 %	5	6.579 %	4	9.756 %
13	20,083,130	5.011 %	970	5.078 %	2	2.632 %	1	2.439 %
14	15,246,461	3.804 %	736	3.853 %	5	6.579 %	2	4.878 %
15	16,198,764	4.042 %	778	4.073 %	0	0%	0	0%
16	18,115,788	4.520 %	801	4.194%	9*,†	11.842 %	6*,†	14.634%
17	14,603,141	3.644 %	702	3.675 %	3	3.947 %	1	2.439 %
18	16,282,716	4.063 %	762	3.989 %	1	1.316%	0	0%

LG	Size (b.p.)	Percentage of assembled genome's size	Coding + non-coding genes	Percentage of total count of coding + non-coding genes	QTL for all traits	Percentage of total QTL count (all traits)	QTL for component traits only	Percentage of total QTL count (component traits)
19	20,240,660	5.050 %	1,044	5.466 %	0*,†	0%	0	0%
20	19,732,071	4.923 %	931	4.874 %	6	7.895 %	5*,†	12.195%
21	11,717,487	2.924 %	463	2.424 %	4	5.263 %	3t	7.317%
Sum	400,788,495‡	100%	19,101§	100 °o	76	100%	41	100%

Extended Data Table 2
MQM model of only main effects of 14 candidate morphological QTL on niche score

The table summarises a main effects-only MQM model enforced to contain all the selected candidate morphological QTL for niche score (run in R/qtl: niche score as response variable, Haley-Knott regression, with F2 family covariate). Model terms (at left) are named according to LG locations of the candidate QTL, which were limited to the one best candidate per LG prior to modelling (see Methods). For each QTL (model term), the table also gives: map position in centimorgans (cM); nearest SNP marker; percentage of total variance explained (PVE) for niche score; *F*-test statistic; corresponding degrees of freedom (d.f.); and *P*-value. Significant model terms indicated as follows: * $0.01 \leq P < 0.05$; † $0.001 \leq P < 0.01$; and ‡ $P < 0.001$. Overall model results (abbreviations: SS, sum of squares; LOD, logarithm$_{10}$ of odds score): $SS_{model} = 169.34$; $d.f._{model} = 56$; $SS_{error} = 464.09$; $d.f._{error} = 473$; $LOD_{model} = 35.80$; $PVE_{model} = 26.73\%$; *P*-value (F) = 2.91×10^{-11}.

Model term	Map location (cM)	Nearest SNP	PVE	*F*	d.f.	*P*-value (*F*)
LG1	36.00	chrI:25560380	0.52 %	1.688	2	0.185991
LG2†	12.00	chrII:10092618	1.69 %	5.470	2	0.004480†
LG4‡	28.76	ChrIV:10997988	2.34 %	7.542	2	0.000596‡
LG7‡	26.99	chrVII:19857837	3.28 %	10.58	2	3.21×10^{-05}‡
LG8‡	18.00	chrVIII:16299555	3.74 %	12.08	2	7.63×10^{-06}‡
LG9	28.38	chrIX:6126845	0.84 %	2.719	2	0.066968
LG10	4.00	chrX:1275840	0.66 %	2.121	2	0.121029
LG12	38.00	chrXII:15046849	0.57 %	1.846	2	0.159010
LG13*	20.00	chrXIII:17392141	1.16%	3.733	2	0.024619*
LG14	39.55	chrXIV:4632223	0.24 %	0.776	2	0.460631
LG16†	13.52	chrXVI:9981125	2.16%	6.989	2	0.001020†
LG17	12.00	chrXVII:2232080	0.03 %	0.109	2	0.896894
LG20*	14.62	ChrXX:9279241	1.43 %	4.630	2	0.010198*
LG21	18.00	chrXXI:11060209	0.33 %	1.080	2	0.340247

Extended Data Table 3

Full MQM model of main QTL effects and effects of pairwise QTL interactions on niche score

The table summarises the final MQM model of candidate QTL effects on niche score, obtained by the stepwise model selection procedure described in the Methods. At all steps, model fitting was performed in R/qtl (niche score as response variable, Haley-Knott regression, with F2 family covariate). Model terms (at left) are named according to LG location of candidate morphological QTL (map positions in Extended Data Table 2). For each term (QTL), the table also gives: percentage of total variance explained (PVE) for niche score; *F*-test statistic; corresponding degrees of freedom (d.f.); and *P*-value. Significant model terms indicated as follows: * $0.01 \leq P < 0.05$; † $0.001 \leq P < 0.01$; and ‡ $P < 0.001$. Overall model results (abbreviations: SS, sum of squares; LOD, logarithm_{10} of odds score): $SS_{model} = 209.75$; $d.f._{model} = 66$; $SS_{error} = 423.68$; $d.f._{error} = 463$; $LOD_{model} = 46.28$; $PVE_{model} = 33.11\%$; *P*-value (*F*) = 3.44×10^{-15}

Model term	PVE	*F*	d.f.	*P*-value (*F*)
LG1†	3.02 %	3.486	6	0.00222†
LG2†	1.50 %	5.200	2	0.00584†
LG4‡	4.30 %	4.958	6	6.21×10^{-05}‡
LG7‡	4.63 %	5.341	6	2.40×10^{-06}‡
LG8‡	3.48 %	12.04	2	7.97×10^{-06}‡
LG9*	0.96 %	3.328	2	0.0367[2]*
LG13*	1.18%	4.099	2	0.01720*
LG14*	1.90 %	2.192	6	0.04268*
LG16‡	5.57 %	3.858	10	4.87×10^{-05}‡
LGl7†	3.95 %	2.737	10	0.00278†
LG20†	1.43 %	4.956	2	0.00742†
LGl×LGl6†	2.48 %	4.284	4	0.00205†
LG4×LG17†	1.99 %	3.440	4	0.00873†
LG7×LG17*	1.54 %	2.658	4	0.03233*
LG14×LG16*	1.36 %	2.414	4	0.04817*

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding was provided by grants from the U.S. National Institutes of Health (F32 GM086125 to M.E.A., R01 GM089733 to C.L.P. and D.S., and P50 HG002568 to C.L.P. and D.M.K.), the Natural Sciences and Engineering Research Council of Canada (to D.S.), and the Canada Foundation for Innovation (to D.S.). We thank Jessica Perez for counting gill rakers, and Cassandra Sather for performing lab work for SNP genotyping. Stable isotopes were analysed at the UC Davis Stable Isotope Facility. Karl Broman, Graham Coop, Anna Greenwood, Peter Wainwright, Michael White, Matthew Wund, and four anonymous referees provided instructive comments, which helped us improve the paper.

Literature Cited

1. Darwin, C. The Origin of Species by Means of Natural Selection. John Murray; 1859.
2. Schluter, D. The Ecology of Adaptive Radiation. Oxford University Press; 2000.
3. Coyne, JA.; Orr, HA. Speciation. Sinauer Associates, Inc.; 2004.
4. Nosil, P. Ecological Speciation. Oxford University Press; 2012.
5. Chase, JM.; Leibold, MA. Ecological Niches: Linking Classical and Contemporary Approaches. The University of Chicago Press; 2003.
6. Fisher, RA. The Genetical Theory of Natural Selection. Oxford University Press; 1930.
7. Orr HA. The genetic theory of adaptation: a brief history. Nat Rev Genet. 2005; 6:119–127. [PubMed: 15716908]
8. Gavrilets, S. Fitness Landscapes and the Origin of Species. Vol. 41. Princeton Univ. Press; 2004.
9. Yeaman S, Whitlock MC. The genetic architecture of adaptation under migration-selection balance. Evolution. 2011; 65:1897–1911. [PubMed: 21729046]
10. Mackay TFC, Stone EA, Ayroles JF. The genetics of quantitative traits: challenges and prospects. Nat Rev Genet. 2009; 10:565–577. [PubMed: 19584810]
11. Barrett RDH, Hoekstra HE. Molecular spandrels: tests of adaptation at the genetic level. Nat Rev Genet. 2011; 12:767–780. [PubMed: 22005986]
12. Taylor EB, McPhail JD. Historical contingency and ecological determinism interact to prime speciation in sticklebacks, Gasterosteus. Proc R Soc Lond B. 2000; 267:2375–2384.
13. Schluter D. Experimental evidence that competition promotes divergence in adaptive radiation. Science. 1994; 266:798–801. [PubMed: 17730400]
14. Schluter D. Adaptive radiation in sticklebacks: trade-offs in feeding performance and growth. Ecology. 1995; 76:82–90.
15. Rundle HD, Nagel LN, Boughman JW, Schluter D. Natural selection and parallel speciation in sympatric sticklebacks. Science. 2000; 287:306–308. [PubMed: 10634785]
16. McPhail JD. Ecology and evolution of sympatric sticklebacks (Gasterosteus): evidence for a species pair in Paxton Lake, Texada Island, British Columbia. Can J Zool. 1992; 70:361–369.
17. Matthews B, Marchinko KB, Bolnick DI, Mazumder A. Specialization of trophic position and habitat use by sticklebacks in an adaptive radiation. Ecology. 2010; 91:1025–1034. [PubMed: 20462117]
18. McGee MD, Wainwright PC. Convergent evolution as a generator of phenotypic diversity in threespine stickleback. Evolution. 2013; 67:1204–1208. [PubMed: 23550768]
19. McGee MD, Schluter D, Wainwright PC. Functional basis of ecological divergence in sympatric stickleback. BMC Evol Biol. 2013; 13:277. [PubMed: 24380474]
20. Gow JL, Peichel CL, Taylor EB. Contrasting hybridization rates between sympatric three-spined sticklebacks highlight the fragility of reproductive barriers between evolutionarily young species. Mol Ecol. 2006; 15:739–752. [PubMed: 16499699]
21. Hatfield T, Schluter D. Ecological speciation in sticklebacks: environment-dependent hybrid fitness. Evolution. 1999; 53:866–873.
22. Rundle HD. A test of ecologically dependent postmating isolation between sympatric sticklebacks. Evolution. 2002; 56:322–329. [PubMed: 11926500]
23. Gow JL, Peichel CL, Taylor EB. Ecological selection against hybrids in natural populations of sympatric threespine sticklebacks. J Evol Biol. 2007; 20:2173–2180. [PubMed: 17887972]
24. Marchinko KB. Predation's role in repeated phenotypic and genetic divergence of armor in threespine stickleback. Evolution. 2009; 63:127–138. [PubMed: 18803682]
25. Carlson SM, Kottas A, Mangel M. Bayesian analysis of size-dependent overwinter mortality from size-frequency distributions. Ecology. 2010; 91:1016–1024. [PubMed: 20462116]
26. Candolin U, Voigt HR. Correlation between male size and territory quality: consequence of male competition or predation susceptibility? Oikos. 2001; 95:225–230.

27. MacColl ADC. Parasites may contribute to ‘magic trait’ evolution in the adaptive radiation of three-spined sticklebacks, Gasterosteus aculeatus (Gasterosteiformes: Gasterosteidae). Biol J Linn Soc. 2009; 96:425–433.
28. Rogers SM, et al. Genetic signature of adaptive peak shift in threespine stickleback. Evolution. 2012; 66:2439–2450. [PubMed: 22834743]
29. Jones FC, et al. The genomic basis of adaptive evolution in threespine sticklebacks. Nature. 2012; 484:55–61. [PubMed: 22481358]
30. Jones FC, et al. A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. Curr Biol. 2012; 22:83–90. [PubMed: 22197244]
31. Phillips PC. Epistasis - the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet. 2008; 9:855–867. [PubMed: 18852697]
32. Mackay TFC. Epistasis and quantitative traits: using model organisms to study gene-gene interactions. Nat Rev Genet. 2014; 15:22–33. [PubMed: 24296533]
33. Whitlock MC, Phillips PC, Moore FBG, Tonsor SJ. Multiple fitness peaks and epistasis. Annu Rev Ecol Syst. 1995; 26:601–629.
34. Via S. Natural selection in action during speciation. Proc Natl Acad Sci USA. 2009; 106:9939–9946. [PubMed: 19528641]
35. Hohenlohe PA, et al. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genetics. 2010; 6 e1000862.
36. Feder JL, Egan SP, Nosil P. The genomics of speciation-with-gene-flow. Trends Genet. 2012; 28:342–350.
37. Strasburg JL, et al. What can patterns of differentiation across plant genomes tell us about adaptation and speciation? Phil Trans R Soc B. 2012; 367:364–373. [PubMed: 22201166]
38. Seehausen O, et al. Genomics and the origin of species. Nat Rev Genet. 2014; 15:176–192. [PubMed: 24535286]
39. Rice WR, Hostert EE. Laboratory experiments on speciation: what have we learned in 40 years? Evolution. 1993; 47:1637–1653.
40. Schluter D, Conte GL. Genetics and ecological speciation. Proc Natl Acad Sci USA. 2009; 106:9955–9962. [PubMed: 19528639]
41. Egan SP, Funk DJ. Ecologically dependent postmating isolation between sympatric host forms of Neochlamisus bebbianae leaf beetles. Proc Natl Acad Sci USA. 2009; 106:19426–19431. [PubMed: 19875692]
42. McBride CS, Singer MC. Field studies reveal strong postmating isolation between ecologically divergent butterfly populations. PLoS Biol. 2010; 8 e1000529.
43. Rundle HD, Nosil P. Ecological speciation. Ecol Lett. 2005; 8:336–352.
44. Presgraves DC. The molecular evolutionary basis of species formation. Nat Rev Genet. 2010; 11:175–180. [PubMed: 20051985]
45. Fry, B. Stable Isotope Ecology. Springer; 2006.
46. van Ooijen, JW.; Voorrips, RE. JoinMap® 30: Software for the Calculation of Genetic Linkage Maps. Plant Research International B. V.; 2001.
47. Broman, KW.; Sen, S. A Guide to QTL Mapping with R/qtl. Springer Science+Business Media, LLC; 2009.
48. Schluter D. Estimating the form of natural selection on a quantitative trait. Evolution. 1988; 42:849–861.
49. Wootton, RJ. A Functional Biology of Sticklebacks. Vol. 265. Univ. of California Press; 1984.
50. Post DM. Using stable isotopes to estimate trophic position: models, methods, and assumptions. Ecology. 2002; 83:703–718.
51. The R Foundation for Statistical Computing; Vienna, Austria: 2011. R: A language and environment for statistical computing v. 2.14.0 Leopard build 64-bit (5933). available from: http://www.R-project.org/
52. Vander Zanden MJ, Vadeboncoeur Y. Fishes as integrators of benthic and pelagic food webs in lakes. Ecology. 2002; 83:2152–2161.

53. Bolnick DI, et al. The ecology of individuals: incidence and implications of individual specialization. Am Nat. 2003; 161:1–28. [PubMed: 12650459]
54. McIntyre PB, Flecker AS. Rapid turnover of tissue nitrogen of primary consumers in tropical freshwaters. Oecologia. 2006; 148:12–21. [PubMed: 16456686]
55. Harmon LJ, et al. Evolutionary diversification in stickleback affects ecosystem functioning. Nature. 2009; 458:1167–1170. [PubMed: 19339968]
56. Behm JE, Ives AR, Boughman JW. Breakdown in postmating isolation and the collapse of a species pair through hybridization. Am Nat. 2010; 175:11–26. [PubMed: 19916869]
57. Bolnick DI. Sympatric speciation in threespine stickleback: Why not? International Journal of Ecology. 2011 2011, Article ID 942847.
58. Hartigan JA, Hartigan PM. The dip test of unimodality. Ann Stat. 1985; 13:70–84.
59. Package ‘diptest’: Hartigan's dip test statistic for unimodality - corrected code v. 0.75-4. 2011. (CRAN: Comprehensive R Archive Network, http://CRAN.R-project.org/package=diptest
60. Schluter D. Adaptive radiation in sticklebacks: size, shape, and habitat use efficiency. Ecology. 1993; 74:699–709.
61. Albert AYK, et al. The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. Evolution. 2008; 62:76–85. [PubMed: 18005154]
62. State University of New York; Stony Brook, NY: 2006. tpsDig v. 2.12 (Department of Ecology and Evolution. available from: http://life.bio.sunysb.edu/morph/
63. Rohlf FJ, Slice D. Extensions of the Procrustes method for the optimal superimposition of landmarks. Syst Zool. 1990; 39:40–59.
64. CRAN: Comprehensive R Archive Network; 2009. Package ‘shapes’: statistical shape analysis v. 1.1-3. http://CRAN.R-project.org/package=shapes
65. Arnegard ME, et al. Sexual signal evolution outpaces ecological divergence during electric fish species radiation. Am Nat. 2010; 176:335–356. [PubMed: 20653442]
66. Peichel CL, et al. The genetic architecture of divergence between threespine stickleback species. Nature. 2001; 414:901–905. [PubMed: 11780061]
67. Warton DI, Duursma RA, Falster DS, Taskinen S. SMATR 3 – an R package for estimation and inference about allometric lines. Methods in Ecology and Evolution. 2012; 3:257–259.
68. Smith RJ. Use and misuse of the reduced major axis for line-fitting. Am J Phys Anthropol. 2009; 140:476–486. [PubMed: 19425097]
69. Miller, A. Subset Selection in Regression. 2nd edn. Vol. Vol. 95. Chapman & Hall/CRC; 2002.
70. CRAN: Comprehensive R Archive Network; 2009. Package ‘leaps’: regression subset selection v. 2.9. http://CRAN.R-project.org/package=leaps
71. Mallows CL. Some comments on C_P. Technometrics. 1973; 15:661–675.
72. Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 1974; 19:716–723.
73. Burnham, KP.; Anderson, DR. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. 2nd edn. Springer-Verlag; 2002.
74. Sambrook, J.; Russell, DW. Molecular Cloning: A Laboratory Manual. 3rd edition. Cold Spring Harbor Laboratory Press; 2001.
75. Hadfield JD, Richardson DS, Burke T. Towards unbiased parentage assignment: combining genetic, behavioural and spatial data in a Bayesian framework. Mol Ecol. 2006; 15:3715–3730. [PubMed: 17032269]
76. CRAN: Comprehensive R Archive Network; 2012. Package ‘MasterBayes’: ML and MCMC methods for pedigree reconstruction and analysis v. 2.50. http://cran.r-project.org/web/packages/MasterBayes/
77. Manichaikul A, Dupuis J, Sen S, Broman KW. Poor performance of bootstrap confidence intervals for the location of a quantitative trait locus. Genetics. 2006; 174:481–489. [PubMed: 16783000]
78. Kosambi DD. The estimation of map distance from recombination values. Ann Eugen. 1944; 12:172–175.
79. Theil, H. Economic Forecasts and Policy. 2nd edn. North-Holland Pub. Co; 1961.

80. Engle RF, Brown SJ. Model selection for forecasting. Appl Math Comput. 1986; 20:313–327.
81. CRAN: Comprehensive R Archive Network; 2012. Package ‘lmtest’: testing linear regression models v. 0.9-30. http://CRAN.R-project.org/package=lmtest
82. Erickson DL, Fenster CB, Stenøien HK, Price D. Quantitative trait locus analyses and the study of evolutionary process. Mol Ecol. 2004; 13:2505–2522. [PubMed: 15315666]

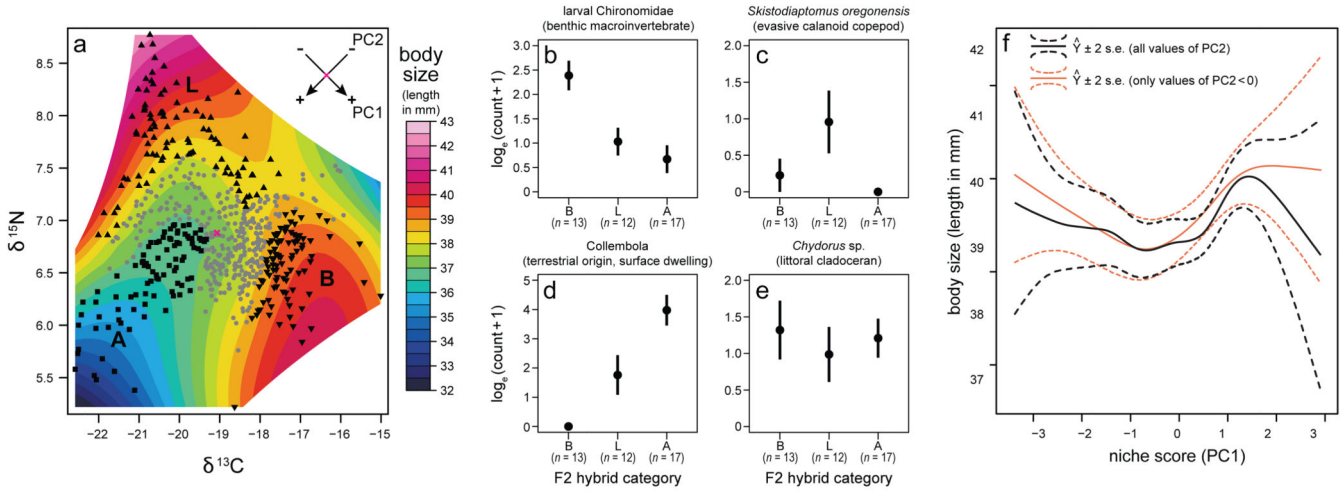


Figure 1. Niche use and body size

(a) Stable isotopes ($\delta^{13}C$, $\delta^{15}N$) for 625 F2 hybrids, showing contours of loess-smoothed body size (mm). Individuals with extreme loess-predicted size shown as black points (▼ for group 'B', ▲ for 'L', ■ for 'A'; each contains 15% of individuals sampled from the pond; 'L' restricted to PC1 < 0.045 to preserve group distinctiveness). Other individuals shown as grey circles. Arrows indicate principal components of isotope distribution (PC1, 'niche score'; PC2, 'diet deviation score'; origin, red cross). **(b–e)** Counts of common food items (mean ± 1 s.e.m.) in digestive tracts of 'B', 'L', and 'A' individuals. Kruskal-Wallis tests for differences among groups: larval Chironomidae ($\chi^2_2=13.52$, $P = 0.001$); *S. oregonensis* ($\chi^2_2=7.547$, $P = 0.023$); Collembola ($\chi^2_2=18.67$, $P = 8.82\times10^{-5}$); *Chydorus* sp. ($\chi^2_2=0.629$, $P = 0.730$). **(f)** Cubic splines[48] of mean body size against 'niche score' (predicted values ± 2 s.e.m.) estimated using the 20 largest F2 families ($n = 438$ individuals), 1,000 bootstrap replicates, and F2 family as a covariate (black, all individuals; orange, individuals with PC2<0).

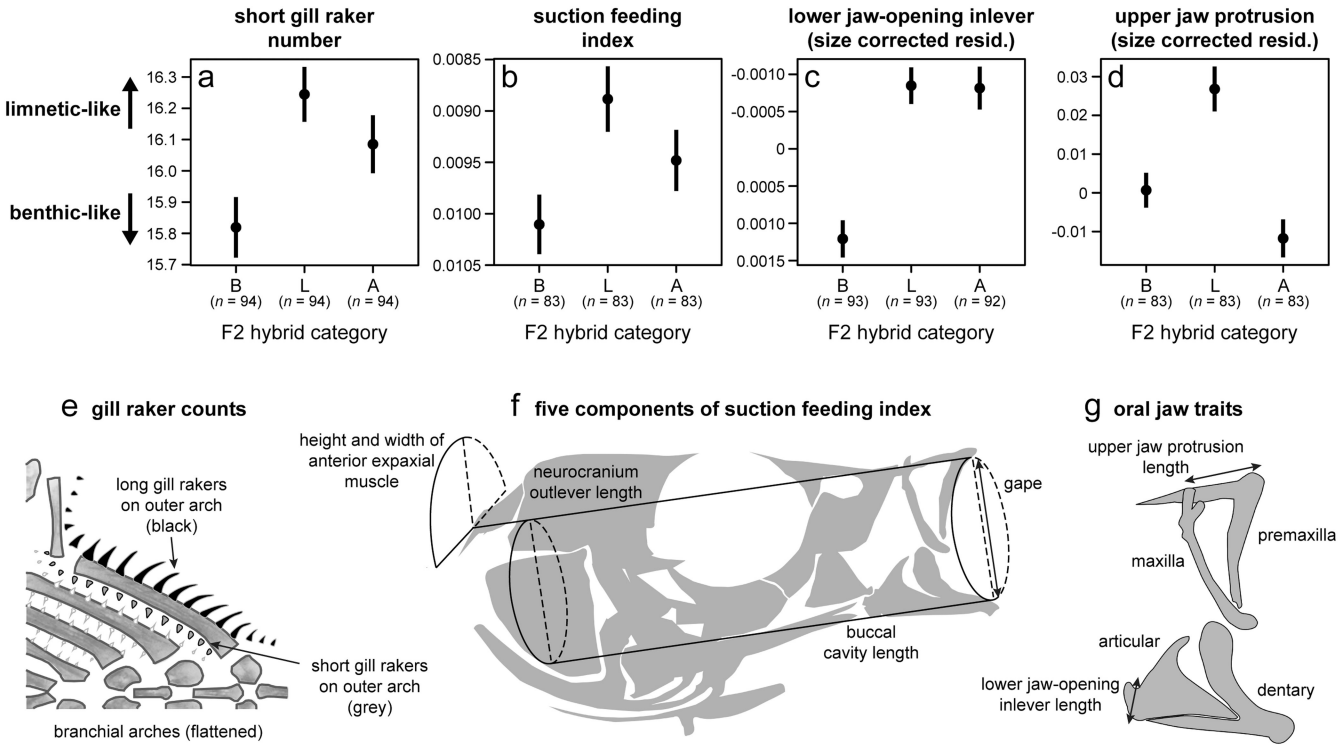


Figure 2. Trait variation among F2 hybrid groups

Trait means (± 1 s.e.m.) of F2 hybrids in categories 'B', 'L', and 'A' (Fig. 1a): **(a)** number of short gill rakers (ANOVA, $F_{2,279} = 5.396$, $P = 0.005$); **(b)** suction feeding index ($F_{2,246} = 4.080$, $P = 0.018$); **(c)** residual lower jaw-opening inlever length ($F_{2,275} = 20.36$, $P = 5.65\times10^{-9}$); **(d)** residual upper jaw protrusion length ($F_{2,246} = 14.94$, $P = 7.54\times10^{-7}$). Trait illustrations: **(e)** gill rakers, functioning in prey retention[16,17]; **(f)** five components of suction index[18,19]; **(g)** two oral jaw traits, which influence efficiency of capturing evasive zooplankton[19]. Arrows indicate directions of benthic–limnetic divergence (vertical axes of b and c inverted to facilitate visual comparisons).

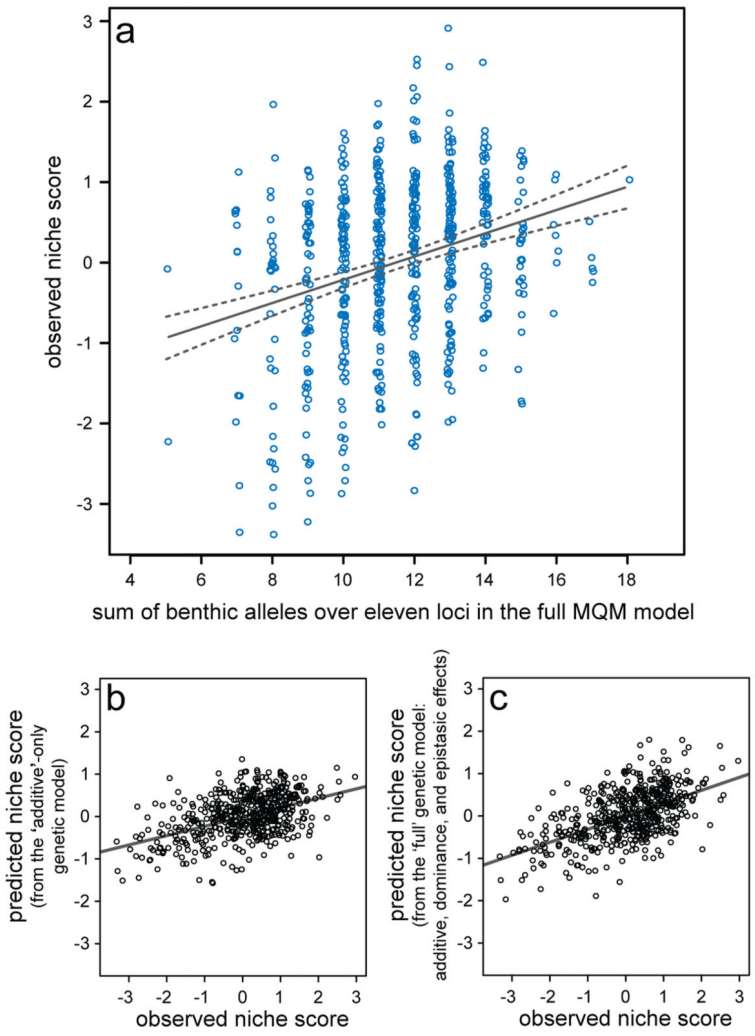


Figure 3. Genetic architecture of niche divergence

(a) Niche scores of F2 hybrids are predicted from the number of benthic alleles summed across eleven unlinked loci (R^2 = 0.081; $F_{1,605}$ = 53.52; P = 8.18×10^{-13}). Dashed lines are 95% confidence intervals of regression line (solid). **(b)** Observed niche score compared with that predicted by the additive-only genetic model. **(c)** Observed niche score compared with that predicted by the full genetic model. Statistics for **(b)** and **(c)** are provided in the text.