The Journal of Experimental Biology 215, 3587-3596

doi:10.1242/jeb.071951

RESEARCH ARTICLE

Correlates of prolonged swimming performance in F2 hybrids of migratory and non-migratory threespine stickleback

Anne C. Dalziel* and Patricia M. Schulte

Department of Zoology and Biodiversity Research Center, 6270 University Blvd, University of British Columbia, Vancouver, BC, Canada, V6T 1Z4

*Author for correspondence (dalziel@zoology.ubc.ca)

SUMMARY

Determining which underlying traits contribute to differences in whole-animal performance can be difficult when many traits differ between individuals with high and low capacities. We have previously found that migratory (anadromous marine) and non-migratory (stream-resident) threespine stickleback (*Gasterosteus aculeatus*) populations have genetically based differences in prolonged swimming performance (U_{crit}) that are associated with divergence of a number of candidate morphological and physiological traits (pectoral fin size and shape, body shape, pectoral muscle and heart size, and pectoral muscle metabolic enzyme activities). Here, we use F2 hybrid crosses to determine which traits are correlated with U_{crit} when expressed in a largely randomized genetic background and a range of trait values for other divergent traits. We found that four of our 12 candidate traits were positively correlated with U_{crit} in F2 hybrids and that the combined effects of ventricle mass, pectoral adductor mass and adductor citrate synthase activity accounted for 17.9% of the variation in U_{crit}. These data provide additional support for a causal role of muscle and heart size in mediating intraspecific differences in U_{crit}, but indicate that many candidate morphological and biochemical traits do not have a strong effect on U_{crit} when disassociated from other divergent traits. However, the limited variation in U_{crit} in our F2 hybrid families may have decreased our ability to detect correlations among these candidate traits and U_{crit}. These data suggest that many traits, interactions among traits and traits not measured in this study affect prolonged swimming performance in threespine stickleback.

Supplementary material available online at http://jeb.biologists.org/cgi/content/full/215/20/3587/DC1

Key words: critical swimming speed, common garden experiment, pectoral muscle size, ventricle size, citrate synthase activity.

Received 28 February 2012; Accepted 23 June 2012

INTRODUCTION

The ability to complete a whole-animal performance task, such as running from a predator or courting a mate, is determined by the integration of multiple traits at underlying levels of biological organization (reviewed by Arnold, 1983; Kingsolver and Huey, 2003). This underlying complexity [termed many-to-one mapping (Wainwright et al., 2005)] makes it possible for the same performance traits to evolve in a variety of ways. Determining which of the many possible underlying traits have evolved to cause differences in performance among populations and species provides insights into the physiological mechanisms by which ecologically relevant traits evolve, and the functional tradeoffs and facilitations that influence the evolution of performance in natural populations (reviewed by Walker, 2007; Walker, 2010). Variation in the capacity for endurance exercise is predicted to influence survival and reproduction in a number of species (reviewed by Husak and Fox, 2008; Irschick et al., 2008), including anadromous fishes (reviewed by Fraser et al., 2011). In this study we use an advanced generation (F2) hybrid cross between migratory (anadromous marine) and non-migratory (stream-resident) threespine stickleback (*Gasterosteus aculeatus* Linnaeus 1758) populations to identify the morphological and physiological traits that are correlated with the capacity for prolonged swimming performance.

Many of the morphological (reviewed by Webb, 1982; Weihs and Webb, 1983; Blake, 2004; Langerhans and Reznick, 2009) and physiological (reviewed by Jones and Randall, 1978; Kolok, 1999; Bernal et al., 2001; Farrell, 2002) traits that can improve prolonged swimming capacity have been identified by comparing species with exceptionally high prolonged swimming capacities with less-exceptional species (reviewed by Bernal et al., 2001). Manipulative studies (e.g. Pearson and Stevens, 1991; Gallaugher et al., 1995; Brauner et al., 1993; Brauner et al., 2011) and studies on naturally occurring variation within populations (e.g. Kolok, 1992; Reidy et al., 2000; Claireaux et al., 2005) have also been used to determine which traits correlate with differences in prolonged swimming (reviewed by Kolok, 1999). However, little is known about the specific traits that have evolved to cause differences in prolonged swimming capacity within and among natural populations. This is in large part because swimming performance (e.g. Lee et al., 2010) and the underlying traits that can influence prolonged swimming capacity (e.g. Hoffmann and Borg, 2006; Sharpe et al., 2008; Anttila et al., 2008) are phenotypically plastic. Therefore, experiments that control for environmental differences (i.e. 'common garden' experiments that can determine whether traits are genetically based) are needed to identify the traits that cause differences in swimming capacity.

Threespine sticklebacks are small teleost fish that live in fresh and salt waters throughout the northern hemisphere (reviewed by Wootton, 1984; Bell and Foster, 1994; Ostlund-Nilsson et al., 2007). On the northern Pacific Coast of North America, marine fish

established a number of freshwater populations when these habitats were uncovered after the Cordilleran Ice Sheet receded, approximately 10,000 to 12,000 years ago (McPhail, 1994). We have previously shown that differences in prolonged swimming capacity between wild stream-resident and anadromous marine (hereafter referred to as 'marine') stickleback populations (Schaarschmidt and Jürss, 2003; Tudorache et al., 2007) are genetically based (Dalziel et al., 2012a). In addition, we have found that a number of morphological, physiological and biochemical traits predicted to impact prolonged swimming have evolved in conjunction with differences in performance in stream-resident threespine stickleback populations (Dalziel et al., 2012a; Dalziel et al., 2012b). For example, stream-resident fish from Bonsall Creek (Vancouver Island, BC, Canada) have smaller pectoral fins, a less streamlined body shape, a lower maximal metabolic rate, smaller hearts, and smaller and more glycolytic pectoral muscles than migratory marine sticklebacks (Dalziel et al., 2012a; Dalziel et al., 2012b). Any of these traits could, in principle, cause reductions in swimming performance in stream-resident fish.

In natural populations of stream-resident and marine sticklebacks, all of the traits that have diverged between populations will co-vary. Therefore, comparing stream-resident with marine populations cannot conclusively show that any particular trait is causally associated with the reduced swimming performance of stream-resident fish. To address the issue of covariation among traits we have generated F2 hybrid crosses between stream and marine stickleback populations from Bonsall Creek. In advanced generation hybrid crosses, recombination breaks down much of the linkage disequilibrium found among loci (and phenotypic traits) in the parental populations, allowing us to test the effect of a candidate trait (and any tightly linked traits) on prolonged swimming performance in fish with a largely randomized genetic background. Significant correlations between candidate traits and swimming performance suggest that variation in the selected trait is sufficient to affect critical swimming speed (U_{crit}) in a fish with a range of trait values for all of the other un-linked phenotypic traits that vary between stream-resident and marine fish.

In this study we specifically tested for correlations between several morphological (pectoral fin area, pectoral fin shape and body shape as proxies for thrust generation and drag), physiological (ventricle and pectoral muscle mass as proxies for cardiac output and swimming muscle aerobic capacity) and biochemical [activity of citrate synthase (CS), cytochrome *c* oxidase (COX) and lactate dehydrogenase (LDH) as proxies for mitochondrial content and muscle fibre-type composition] traits and U_{crit} in F2 hybrid sticklebacks. The goal of this study was to identify the underlying mechanistic causes of variation in U_{crit} between stream-resident and marine stickleback populations.

MATERIALS AND METHODS

Experimental animals

The grandparents of the F2 hybrid fish used in this study were collected from wild populations living in Bonsall Creek on Vancouver Island (Dalziel et al., 2012a; Dalziel et al., 2012b) (BC Ministry of Environment Fish Collection Permits NA/SU06-26169 and NA/SU07-38414). First generation (F1) pure and hybrid crosses between wild marine and stream-resident sticklebacks were bred in the spring and summer of 2006 and 2007. Full details of our breeding protocols are presented in Dalziel et al. (Dalziel et al., 2012a). A single marine × stream (MS) F1 cross from 2007 was used to produce the three second-generation (F2) hybrid families used in the present study in April of 2009, so that the three MS F2 families all share the same grandparents, but have different parents. We used this crossing design to limit the number of alleles at each locus to facilitate future quantitative trait locus (QTL) mapping studies. The first MS F2 family was composed of 69 fish (34 females, 38 males, six fish of indeterminate sex), the second F2 family of 75 fish (41 females, 29 males, six fish of indeterminate sex) and the third family of 78 fish (31 females, 27 males, 11 fish of indeterminate sex), for a total of 222 fish. We raised fish in dechlorinated Vancouver tap water brought to 2±0.5‰ with Instant Ocean sea salt (Aquarium Systems, Mentor, OH, USA), and fed fish live brine shrimp twice per day for their first month, *Daphnia* and bloodworms daily for the next 3 months, and *Mysis* shrimp and bloodworms (chironomid larvae) from 4 months on. We reared fish at a natural photoperiod and laboratory temperatures (~11–17°C) until March (~11 months of age). At this age we transferred fish to a 15°C environmental chamber with a controlled 12 h:12 h light:dark photoperiod (the natural photoperiod for our collection sites in March) to prevent fish from entering the reproductive state. When fish reached a size of ~3.5 cm we individually tagged each fish and removed the mesh to give the fish additional space. These tags remained visible throughout the experiment and were used to individually identify each fish during subsequent measurements. The University of British Columbia animal care committee approved all breeding and experimental procedures (A07-0288).

Measurement of maximum prolonged swimming speed: U_{crit}

We used a U_{crit} test to assess prolonged swimming performance in our sticklebacks (Brett, 1964). In this test, water speed is increased in a step-wise manner until a fish can no longer maintain its position in the current. U_{crit} correlates with migratory difficulty among populations of salmonids (e.g. Lee et al., 2003), and is also predicted to be an ecologically relevant measure of prolonged swimming for other species of fish that migrate, forage in the open ocean or live in high-flow streams (Kolok, 1999; Plaut, 2001), as is the case for migratory marine threespine sticklebacks. In threespine sticklebacks, U_{crit} is also tightly correlated with the gait transition from pectoral fin rowing to a combination of pectoral fin rowing and caudal bursts (Dalziel et al., 2012a), which is suggested to be a good measure of sustained swimming speed (e.g. Korsmeyer et al., 2002; Svendsen et al., 2010). Oxygen consumption also continues to increase until U_{crit} is reached (Dalziel et al., 2012a), as has been observed in striped surfperch, another labriform swimmer (Svendsen et al., 2010). Together, these observations suggest that U_{crit} is a good measure of maximal aerobic swimming performance in sticklebacks.

To measure U_{crit}, we swam six individually labeled siblings in a Brett-style 10-l swim tunnel (SWIM-10; Loligo Systems, Hobro, Denmark) at a water temperature of 15±1°C and salinity of 2‰, and calibrated water speed with a vane wheel flow sensor (Höntzch ZSR25, Waiblingen, Germany). The U_{crit} trial generally followed the methods of Dalziel et al. (Dalziel et al., 2012a), with some modifications. In particular, we gave fish less time to recover after an initial training test [1 *versus* 3 h (Dalziel et al., 2012a)], a reduced time increment (2 *versus* 10 min) and decreased step-wise speed increase [0.3 *versus* 0.5 body lengths (BL) s^{-1}]. Changes in the time increment and step-wise speed increase can affect the measured U_{crit} (reviewed by Kolok, 1999), so we conducted preliminary studies to examine the effects of varying these parameters. We found no significant differences in U_{crit} values collected using different time increments (8.16±1.34 BL s^{-1} for 2 min and 8.37±1.18 BL s^{-1} for 10 min; *t*-test, P=0.083) and U_{crit} values collected on the same fish using these two methods were tightly correlated (Pearson product

moment correlation, r=0.911, P<0.0001, N=95) in the F1 crosses used by Dalziel et al. (Dalziel et al., 2012a) (A.C.D., unpublished data). Because a 2 min increment and 0.5 BL s^{-1} step-wise speed increase significantly reduced the time needed to run the trials, and caused only minor differences in absolute U_{crit} values, we selected this protocol. This was important as we needed to measure U_{crit}, and photograph and sample all 222 F2 fish in less than 1 month, to be sure that we measured all traits during the period within which we knew U_{crit} was highly repeatable in these populations (Dalziel et al., 2012a) and to minimize the possibility of age-related effects. We also re-swam a subset of 34 fish 2 weeks after the initial U_{crit} trial to further test the repeatability of U_{crit} in F2 hybrid sticklebacks, and found that swimming performance was significantly repeatable over this time period (supplementary material Fig. S1). At the time of U_{crit} trials our F2 fish were 1 year and 2–3 months of age. We found that there was an effect of size on U_{crit} in F2 hybrids (despite the incorporation of standard length in the calculation for U_{crit}), so all analyses were performed using the residuals of regressions of U_{crit} against standard length (U_{crit}*; see Statistical analysis for details).

Measurement of morphological traits predicted to influence U_{crit}

Photographs of F2 stickleback (N=222 fish from three F2 families) were taken less than 3 weeks after we measured U_{crit}. We anesthetized fish with 0.2 g tricaine methanesulfonate buffered with 0.4 g sodium bicarbonate in 1 l of water and photographed the right side of the fish with a ruler in the field of view. We took a second photograph of the right pectoral fin maximally spread over a laminated sheet of paper to measure pectoral fin area and shape. Fin area was measured by tracing an outline of the fin in ImageJ (National Institutes of Health, Bethesda, MD, USA). We used TPSdig 2.1 (Rohlf, 2010) to digitize six landmarks onto the stickleback's pectoral fin (supplementary material Fig. S2A), and 12 landmarks onto the stickleback's body to measure fin and body shape traits (supplementary material Fig. S2B,C) (see Dalziel et al., 2012a). We chose to measure five body shape traits predicted to mediate evolutionary variation in prolonged swimming capacity in fishes (reviewed by Blake, 2004; Langerhans and Reznick, 2009) that varied between F1 stream-resident and marine stickleback crosses (Dalziel et al., 2012a). These traits were: (1) fineness ratio (standard length divided by maximum depth), (2) head depth, (3) posterior depth at third spine, (4) caudal peduncle depth and (5) caudal area. Linear measures were collected from landmarks with TMorphGen6c (IMP suite 2006) (Zelditch et al., 2004). We corrected measurements for overall body size by performing a least-squared regression against mass and using residuals in all subsequent analyses. Residuals were made positive by the addition of a constant, $\log_{10}$ transformed, and divided by two for linear measures and by three for caudal area in preparation for multivariate analyses.

To obtain a composite measure of body shape and pectoral fin shape, we performed linear discriminant function analyses on the five body shape traits and six pectoral fin landmarks with the MASS package in R (Venables and Ripley, 2002), and 'trained' the analysis to differentiate pure marine from pure stream-resident fish collected by Dalziel et al. (Dalziel et al., 2012a) (Table 1). Therefore, values for the first linear discriminant will differentiate 'marine-shaped' fish from 'stream-resident-shaped' fish. At the time photos were taken, both F2 and F1 fish [measured by Dalziel et al. (Dalziel et al., 2012a)] were less than 1 month older than they were in U_{crit} trials.

Table 1. Variance explained by, and factor loadings for, the linear discriminant (ld) function produced in the discriminant function analysis of six body shape traits in the threespine stickleback

Trait	Coefficient of linear discriminants
Fineness	–97.054
Caudal peduncle depth	61.704
Caudal area	32.557
Posterior depth	–37.250
Head depth	–36.315

Ld1 values for F2 fish were obtained by training with data from pure stream and pure marine crosses: marine fish have low ld1 scores and stream fish have high ld1 scores.
See Materials and methods for a full description of traits.

Measurement of physiological and biochemical traits predicted to influence U_{crit}

We terminally sampled F2 hybrid stickleback for tissue collection less than 1 month after we measured U_{crit} (fish were 1 year and 3–4 months of age). We killed fish by placing them in an overdose of anaesthetic (1 g l^{-1} tricaine methanesulfonate buffered with 2 g l^{-1} sodium bicarbonate), and as soon as a fish lost equilibrium (<30 s), it was blotted dry and weighed. We then removed the heart and pectoral adductor and abductor muscles with the aid of a dissecting microscope. All tissues were snap-frozen in liquid nitrogen and stored at –80°C. Prior to freezing our stickleback heart samples, we separated the bulbus arteriosis and atria from the ventricle, and any blood still remaining in the ventricle was removed by blotting the ventricle against a damp kimwipe. The F1 crosses measured in Dalziel et al. (Dalziel et al., 2012b) were approximately 1 year and 5–9 months of age at the time of biochemical sampling, so were between 1 and 6 months older than F2 fish.

To measure enzyme activities, we re-weighed frozen ventricles and pectoral muscles (g), and immediately added pectoral muscles to 20 volumes of chilled homogenization buffer (50 mmol l^{-1} hepes, 1 mmol l^{-1} EDTA and 0.1% Triton X-100; pH 7.4) in 4 ml Wheaton glass homogenizers kept on ice. We measured enzyme activities for COX (EC 1.9.3.1, complex IV in the electron transport chain, and found on the inner mitochondrial membrane), CS (EC 2.3.3.1, a citric acid cycle enzyme found in the mitochondrial matrix), LDH (EC 1.1.1.27, a glycolytic enzyme found in the cytosol, which catalyzes the inter-conversion of pyruvate and NADH to lactate and NAD^+, and allows for high glycolytic flux during cellular hypoxia) and pyruvate kinase (PK; another glycolytic enzyme found in the cytosol) on whole-cell extracts at 25°C using the non-limiting substrate concentrations listed in Dalziel et al. (Dalziel et al., 2012b). Pectoral muscle physiology is predicted to influence swimming performance in sticklebacks, which use these muscles to power prolonged swimming (Walker, 2004). We have previously found that a higher activity of LDH per gram abductor and adductor muscle indicated a higher proportion of pink (fast-oxidative glycolytic) fibres (Dalziel et al., 2012b), so we used this measure as a proxy for fibre-type composition in the pectoral muscles. We also chose to measure the activity of PK per gram adductor and abductor muscle because there was a trend towards lower PK in marine crosses and we found that males had lower PK activities than females, suggesting a possible interaction with sex (Dalziel et al., 2012b). CS and COX are indicators of mitochondrial content, and pectoral muscles with higher mitochondrial content should be able to produce more aerobically generated ATP to fuel prolonged swimming. Although we did not find significant differences in CS and COX among stream-resident and marine stickleback F1 crosses, we did see a

slightly higher activity of these enzymes in marine crosses (Dalziel et al., 2012b). Therefore, we measured these enzymes in F2 fish to further examine the impact of mitochondrial enzyme content on U_{crit}. As well, we measured COX activity because we have found a number of non-synonymous sequence differences in the mitochondrial DNA of stream-resident and marine populations in the genes for cytochrome *b*, NADH dehydrogenase subunit 2 and ATP synthase subunit 6 (A.C.D. and H. Kim, unpublished data); COX is composed of nuclear and mitochondrially encoded protein subunits, so low COX activity may be indicative of possible hybrid incompatibilities due to interactions between mitochondrial and nuclear genes (Burton et al., 2006).

Statistical analysis

All statistical analyses were conducted using R v2.11.1 (R Development Core Team, 2010). We studied the effect of fish size on all of our measurements by testing for a significant correlation with body mass (candidate traits) or body length (U_{crit}). For all measurements that were significantly correlated with mass, we corrected for size by calculating the residuals from a least-squared linear regression against mass (all candidate traits) or standard length (U_{crit}), which was the best fit to the data. To compare data collected for F1 line crosses in Dalziel et al. (Dalziel et al., 2012a; Dalziel et al., 2012b), we calculated residuals with the same linear equations as for our F2 individuals.

We examined the effect of each of our 12 candidate traits (fixed effects) on residual U_{crit} (U_{crit}*; response variable), with family and sex as nested random effects, using the nlme package in R (Pinheiro and Bates, 2000). Our 12 candidate traits were: residual pectoral fin area, pectoral fin shape linear discriminant 1 (ld1), body shape ld1, residual ventricle mass, residual adductor mass, residual abductor mass, residual COX activity per gram adductor, residual CS activity per gram adductor, LDH activity per gram adductor, residual COX activity per gram abductor, residual CS activity per gram abductor and LDH activity per gram abductor. We also modeled the effect of each candidate trait against U_{crit}* with a fixed-effect-only model (no nested random effects) to obtain an estimate for the fraction of variation in U_{crit}* explained by each explanatory variable (R^2). We also examined correlations among explanatory traits, and accounted for multiple comparisons by adjusting our cut-off for significant *P*-values to a corrected *P*-value based upon the false discovery rate calculated by the Brainwaver package in R (Achard, 2010).

We next conducted a multiple linear regression to examine the combined predictive power of our explanatory traits and interactions among these traits. We only included explanatory variables that were significantly associated with U_{crit}* (see preceding paragraph: adductor CS, abductor CS, ventricle mass and adductor mass). To examine the impact of these candidate explanatory traits on U_{crit}*, we followed a model selection approach suggested by Zuur et al. (Zuur et al., 2009) for linear mixed-effects models, which begins with a model with as many of our explanatory variables as possible (fixed effects) and their interactions, then finds the optimal random variable structure, next finds the optimal fixed-effects structure, and finally examines, interprets and validates the optimal model. We examined our fixed-effects structure by using step-wise multiple linear regression (forward and reverse) to determine which explanatory variables should be included in the model. We then examined the significance of the interaction between adductor mass and CS activity, which was strongly predicted by biological mechanism. We also examined all other possible models including all 12 traits, and all two- and three-way interactions using the stepAIC protocol implemented by the MASS package (Venables and Ripley, 2002) in R, and compared Akaike information criterion (AIC) values (data not shown). However, this preliminary analysis did not suggest that any other interactions were significant (data not shown).

Because we only had a small number of closely related F2 families (*N*=3), we could not directly test the contributions of additive and additive-dominance models of composite gene action to phenotypic variation in U_{crit} and candidate traits with the joint-scaling regression technique (Lynch and Walsh, 1998), or examine line variances to test for segregation variance (i.e. high variance in F2 hybrids). In addition, F1 and F2 lines were reared in different years, and thus under slightly different environments. However, we have displayed our data in a manner that allows for a visual comparison of line means in supplementary material Figs S3, S5 and S7.

RESULTS

U_{crit}* of F2 hybrids

U_{crit}* was significantly repeatable over a 2 week time period in F2 hybrid sticklebacks (t=8.27, P<0.001, r^2=0.671; supplementary material Fig. S1), and there was no effect of sex ($F_{1,187}$=1.754, P=0.187; data not shown), family ($F_{2,187}$=2.252, P=0.108; data not shown) or the combination of sex and family ($F_{5,185}$=1.813, P=0.112; data not shown) on U_{crit}* (residuals of regression against standard length). Second-generation F2 hybrids expressed a range of performance phenotypes similar to that of F1 hybrids (Fig. 1A). No F2 fish reached U_{crit}* values as high as the mean U_{crit}* for pure marine F1 crosses, and only 15 of 222 F2 fish had U_{crit}* values as

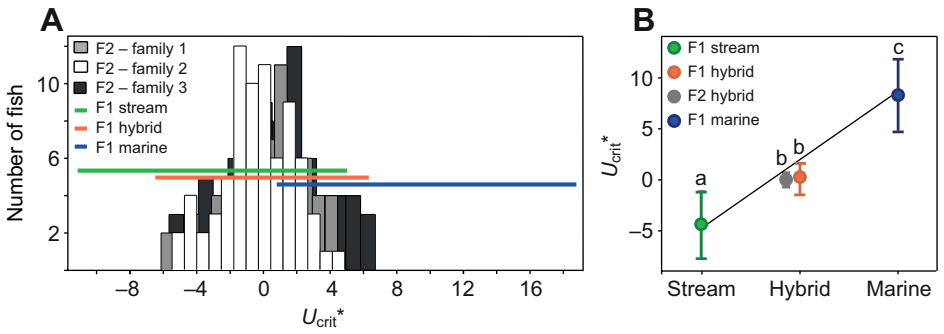

Fig. 1. (A) Histogram of residual critical swimming speed (U_{crit}*) for each F2 hybrid family with the range of U_{crit}* values reached by all F1 individuals represented by colored bars. (B) Grand means ± s.d. of F1 and F2 line crosses ($F_{3,22}$=25.10, P<0.0001). The line connecting F1 stream-resident and marine crosses represents the *a priori* expectation for trait values in hybrids if divergence among ecotypes is due to genes with only additive effects. F2 line means are compiled from data from the three F2 families (n=69–78 fish per family) and F1 line means are grand means compiled from family means (N=7 F1 stream families, 5 F1 marine families and 11 F1 hybrid families, with n=6 fish per family). Data for F1 crosses are from Dalziel et al. (Dalziel et al., 2012a).

low as the mean for pure F1 stream crosses (Fig. 1A). F2 hybrids also had mean U_{crit}* values that were similar to those of F1 hybrid sticklebacks (Fig. 1B), and F1 and F2 hybrid lines had U_{crit}* values that were slightly more similar to stream-resident than marine crosses.

Effect of individual candidate traits on U_{crit}* in F2 hybrids

We found that four of the 12 candidate traits had a significant relationship with U_{crit}* using a mixed-effects linear model (Fig. 2, Table 2, supplementary material Figs S4, S6 and S8), including residual ventricle mass (Fig. 2A), residual adductor mass (Fig. 2B), and residual CS per gram adductor (Fig. 2C) and abductor (Fig. 2D).

Effect of multiple candidate traits on U_{crit}* in F2 hybrids

We next assessed the ability of our explanatory variables, and interactions among variables, to predict U_{crit}* by conducting a multiple linear regression. We only included explanatory variables that had a significant effect on U_{crit}* and, to prevent multicollinearity (reviewed by Slinker and Glantz, 1985), we removed abductor CS activity because it was tightly correlated with adductor CS activity (Table 3) and because the latter had stronger predictive power in our single variable analyses. After removing correlated traits we were left with three explanatory variables in our full multivariate model: ventricle mass, adductor mass and adductor CS. We found that the optimal structure for our model did not include the random effects of family of origin or sex [data not shown, following procedures outlined by Zuur et al. (Zuur et al., 2009)]. Therefore, we used a multiple linear regression including fixed effects only. As expected from our single trait analyses, stepwise multiple linear regressions (forward and reverse) found that all three explanatory should be included in the model. We also included the interaction between adductor mass and CS activity in one of our models, because this interaction is expected to determine the overall 'aerobic capacity' of the adductor muscle. We found that a model including the explanatory variables of residual ventricle mass, residual adductor mass, residual adductor CS activity, and the interaction between adductor mass and CS activity per gram adductor resulted in the best fit (Table 4, model 1), but that this model was not significantly different from a model without the interaction term (Table 4, model 2). Therefore, the reduced model (model 1: adductor CS + ventricle mass + adductor mass) was selected as the best fit model, and explained 17.9% of the variation in U_{crit}* in F2 hybrids.

DISCUSSION

Determining which morphological, physiological and biochemical traits cause differences in whole-animal performance capacity can be a difficult task if many traits vary among populations with high and low capacities. In such populations, all candidate traits will co-vary so it is not possible to determine the effect of any single trait on performance using the comparative method. However, there are a number of experimental designs that can isolate the effect of a particular candidate trait on performance, including physiological manipulations (e.g. applying pharmacological agents or surgical manipulations), reverse genetics (e.g. RNAi or gene insertions/deletions) and classical genetic techniques (i.e. producing controlled crosses) (reviewed by Dalziel et al., 2009). In this study,

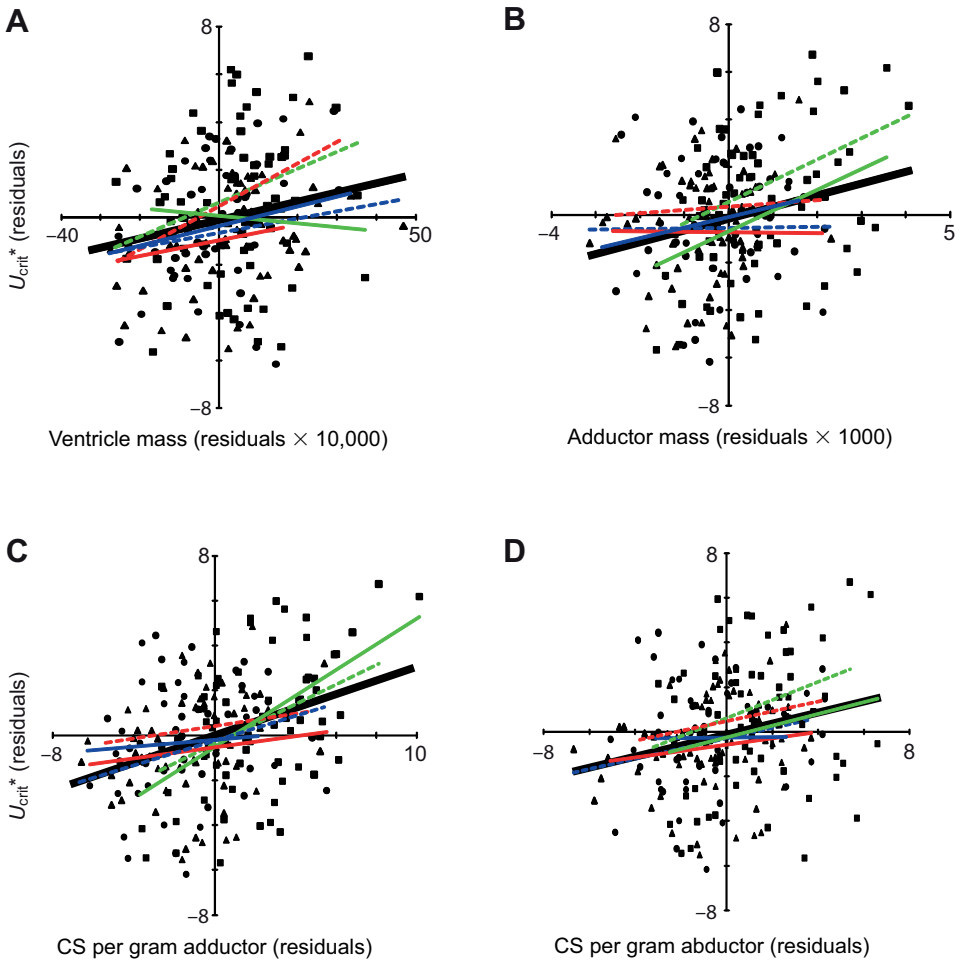

Fig. 2. Relationship between U_{crit}* and residual (A) ventricle mass, (B) adductor mass, (C) citrate synthase (CS) per gram adductor and (D) CS per gram abductor in F2 fish. Thick black lines represent the fitted values for residual ventricle mass ($F_{1,177}$=7.456, P=0.007), adductor mass ($F_{1,184}$=11.250, P=0.001), CS per gram adductor ($F_{1,184}$=28.101, P<0.001) and CS per gram abductor ($F_{1,183}$=10.66, P=0.001) for all data (with family and sex as nested random effects; Table 2). The colored lines represent the fitted values for each family and sex (red dashed line, family 1 females; solid red line, family 1 males; blue dashed line, family 2 females; solid blue line, family 2 males; green dashed line, family 3 females; solid green line, family 3 males. Circles symbolize values for family 1 fish, triangles symbolize family 2 fish, and squares symbolize family 3 fish.

Table 2. Results of linear mixed-model regression of candidate traits (fixed effects) *versus* residual critical swimming speed (U_{crit}*) in the threespine stickleback

Trait/parameter	R^2	Slope	*P*	*F*	d.f.
Pectoral fin surface area (residuals)	0.000	–3.020	0.487	0.485	1,179
Pectoral fin shape ld1	0.003	–0.223	0.227	1.471	1,182
Body shape ld1	0.000	–0.003	0.941	0.005	1,175
Ventricle mass (residuals)	**0.050**	**0.046**	**0.007**	**7.457**	**1,177**
Adductor mass (residuals)	**0.053**	**0.492**	**<0.001**	**11.250**	**1,184**
Abductor mass (residuals)	0.008	0.053	0.124	2.385	1,183
COX per gram adductor (residuals)	0.000	0.016	0.313	1.023	1,184
CS per gram adductor (residuals)	**0.125**	**0.303**	**<0.001**	**28.101**	**1,184**
LDH per gram adductor	0.002	–0.064	0.331	0.948	1,184
COX per gram abductor (residuals)	0.000	0.021	0.230	1.450	1,183
CS per gram abductor (residuals)	**0.051**	**0.246**	**0.001**	**10.660**	**1,183**
LDH per gram abductor	0.000	0.010	0.738	0.112	1,183

Family and sex were included as nested random effects. R^2 values are from models with fixed-effects only.
COX, cytochrome *c* oxidase; CS, citrate synthase; LDH, lactate dehydrogenase.
Values in bold identify the variables that are significantly correlated with U_{crit}*.

we made controlled crosses in the laboratory to identify the traits that have a significant effect on U_{crit}* in F2 hybrids between stickleback populations with high (marine) and low (stream-resident) U_{crit}* values.

We found that four of 12 candidate traits that we measured (i.e. ventricle mass, adductor mass, CS activity per gram adductor and CS activity per gram abductor) significantly regressed against U_{crit}* in F2 hybrids with a largely randomized genetic background. This does not mean that the other eight traits have no effect on U_{crit}*, it simply argues that their effects are highly dependent on the genetic background in which they are expressed. The lack of a significant relationship between many of the candidate traits we measured and U_{crit}* may also be due to the limited variation in U_{crit}* in our F2 hybrid families, which could reduce power to detect correlations between candidate traits and performance. The activities of CS per gram abductor and adductor were tightly correlated, so to prevent multicollinearity, we only included ventricle mass, pectoral adductor mass and the activity of CS per gram adductor muscle in our multiple linear regressions. The combined effects of these three traits accounted for 17.9% of variation in U_{crit}*. These data suggest that there may be traits which were not measured in this study but have a strong effect on U_{crit}*, and that many traits are likely necessary to achieve high swimming performance. Data from our three F2 crosses also suggest that many loci contribute to variation in swimming performance between stream-resident and marine threespine sticklebacks.

Which underlying morphological, biochemical and physiological traits influence U_{crit}?

The observation that eight of the 12 candidate traits that we measured did not have a significant relationship with U_{crit}* in F2 hybrids (pectoral fin size and shape, body streamlining, abductor

Table 3. Correlations among explanatory variables (fixed-effects only)

	Pectoral fin area	Pectoral fin shape	Body shape ld1	Ventricle mass	Adductor mass	Abductor mass	Adductor CS	Adductor COX	Adductor PK	Adductor LDH	Abductor CS	Abductor COX	Abductor PK	Abductor LDH
Pectoral fin area	–	*r*=0.179 *P*=0.0142	*r*=–0.181 *P*=0.0147	*r*=0.134 *P*=0.0729	*r*=0.052 *P*=0.480	*r*=0.114 *P* =0.121	*r*=0.019 *P*=0.792	*r*=0.171 *P*=0.0197	***r*=0.290** ***P*<0.0001**	*r*=0.066 *P*=0.367	*r*=–0.017 *P*=0.820	*r*=0.165 *P*=0.0247	***r*=0.339** ***P*<0.0001**	***r*=0.254** ***P*=0.0005**
Pectoral fin shape	–	–	*r*=–0.023 *P*=0.757	*r*=0.013 *P*=0.863	*r*=–0.201 *P*=0.0059	*r*=–0.170 *P*=0.0208	*r*=0.007 *P*=0.920	*r*=0.089 *P*=0.224	*r*=0.136 *P*=0.0636	*r*=0.197 *P*=0.0070	*r*=0.004 *P*=0.955	*r*=0.0835 *P*=0.258	*r*=0.202 *P*=0.0058	*r*=0.163 *P*=0.027
Body shape ld1	–	–	–	*r*=–0.0154 *P*=0.840	*r*=–0.0193 *P*=0.796	*r*=–0.0827 *P*=0.268	***r*=0.333** ***P*<0.0001**	*r*=0.0603 *P*=0.418	*r*=–0.078 *P*=0.295	*r*=–0.0635 *P*=0.394	***r*=0.349** ***P*<0.0001**	*r*=0.0203 *P*=0.786	*r*=–0.119 *P*=0.111	*r*=–0.133 *P*=0.0745
Ventricle mass	–	–	–	–	*r*=0.092 *P*=0.217	*r*=0.150 *P*=0.0434	*r*=0.128 *P*=0.0841	*r*=0.065 *P*=0.385	*r*=0.018 *P*=0.811	*r*=–0.144 *P*=0.0523	*r*=0.127 *P*=0.0884	*r*=0.053 *P*=0.482	*r*=0.054 *P*=0.468	*r*=0.002 *P*=0.974
Adductor mass	–	–	–	–	–	***r*=0.565** ***P*<0.0001**	***r*=0.237** ***P*<0.0001**	*r*=–0.075 *P*=0.301	*r*=–0.194 *P*=0.0071	*r*=–0.211 *P*=0.0034	***r*=0.276** ***P*=0.0001**	*r*=–0.064 *P*=0.381	*r*=–0.165 *P*=0.023	*r*=–0.202 *P*=0.0051
Abductor mass	–	–	–	–	–	–	*r*=0.219 *P*=0.0024	*r*=0.135 *P*=0.0624	*r*=–0.118 *P*=0.104	*r*=–0.119 *P*=0.102	***r*=0.251** ***P*=0.0005**	*r*=0.094 *P*=0.195	*r*=–0.091 *P*=0.211	*r*=–0.054 *P*=0.456
Adductor CS	–	–	–	–	–	–	–	***r*=0.399** ***P*<0.0001**	*r*=0.210 *P*=0.0035	*r*=–0.090 *P*=0.217	***r*=0.793** ***P*<0.0001**	***r*=0.313** ***P*<0.0001**	*r*=0.141 *P*=0.0528	*r*=0.072 *P*=0.325
Adductor COX	–	–	–	–	–	–	–	–	***r*=0.222** ***P*=0.0021**	*r*=0.161 *P*=0.026	***r*=0.376** ***P*<0.0001**	***r*=0.754** ***P*<0.0001**	*r*=0.193 *P*=0.0076	***r*=0.317** ***P*<0.0001**
Adductor PK	–	–	–	–	–	–	–	–	–	***r*=0.321** ***P*<0.0001**	*r*=–0.007 *P*=0.925	*r*=0.117 *P*=0.108	***r*=0.844** ***P*<0.0001**	***r*=0.648** ***P*<0.0001**
Adductor LDH	–	–	–	–	–	–	–	–	–	–	*r*=–0.121 *P*=0.0973	*r*=0.111 *P*=0.126	***r*=0.275** ***P*=0.0001**	***r*=0.611** ***P*<0.0001**
Abductor CS	–	–	–	–	–	–	–	–	–	–	–	***r*=0.413** ***P*<0.0001**	*r*=–0.025 *P*=0.731	*r*=–0.066 *P*=0.366
Abductor COX	–	–	–	–	–	–	–	–	–	–	–	–	*r*=0.134 *P*=0.065	*r*=0.224 *P*=0.0018
Abductor PK	–	–	–	–	–	–	–	–	–	–	–	–	–	***r*=0.720** ***P*<0.0001**
Abductor LDH	–	–	–	–	–	–	–	–	–	–	–	–	–	–

When corrected for multiple comparisons, the cut-off for significant *P*-values is 0.00241. Significant correlations are in bold.

Table 4. Results of multiple linear regression analyses

	Adjusted R^2	F	d.f.	P	Estimate	s.e.m.	t	P
Model 1: U_{crit}* ~ ventricle mass + adductor mass + adductor CS + adductor mass:adductor CS	0.188	11.03	4,169	0.0000000568				
Coefficient								
(Intercept)					–0.148	0.190	–0.780	0.436
Ventricle mass					0.029	0.014	2.139	0.034
Adductor mass					0.250	0.151	1.653	0.100
Adductor CS					0.261	0.062	4.212	<0.001
Adductor mass:adductor CS					0.067	0.039	1.724	0.087
Model 2: U_{crit}* ~ ventricle mass + adductor mass + adductor CS	0.179	13.56	3,170	0.0000000566				
Coefficient								
(Intercept)					–0.075	0.186	–0.403	0.688
Ventricle mass					0.032	0.014	2.387	0.018
Adductor mass					0.274	0.151	1.816	0.071
Adductor CS					0.276	0.062	4.461	<0.001
Model comparison		2.971	2,169	0.087				

mass, and LDH and COX activity per gram adductor and abductor muscles), despite the fact that six of these eight traits were significantly different between populations in laboratory-reared F1 pure crosses (pectoral fin size and shape, body streamlining, abductor mass, and LDH activity per gram adductor and abductor muscles) (Dalziel et al., 2012a; Dalziel et al., 2012b), suggests that many traits that differ among stream-resident and marine sticklebacks do not have a strong effect on U_{crit}* when disassociated from other traits that differ between ecotypes (i.e. in a randomized genetic background). The effect of genetic background may be particularly important for traits such as body morphology, which all act to influence body streamlining. The weak correlations among all body and fin shape traits in F2 fish (Table 3, supplementary material Table S1) argue that these traits are not necessarily inherited together, which opens the possibility that high contributions to streamlining by one trait may be affected by complex interactions with another, thus resulting in no significant correlation among morphology and performance. We used a composite measure of overall body shape (body shape ld1), to partially address this issue, but this composite trait did not significantly regress against U_{crit}* either. However, this does not rule out the possibility that body and fin shape traits may cause some of the variation in U_{crit}* among stream-resident and marine sticklebacks, but does suggest that any single body/fin shape trait alone is not sufficient to significantly affect U_{crit}* when expressed in a randomized genetic background.

The correlations between ventricle and pectoral muscle mass and U_{crit}* observed here [and in Dalziel et al. (Dalziel et al., 2012b)] argue for a causal role of these traits in intraspecific differences in U_{crit}*, but may also be due to other loci in close physical linkage with the loci underlying these traits. Alternatively, our finding that CS activity per gram adductor muscle contributes to differences in U_{crit}* in F2 stickleback hybrids was not anticipated from our earlier studies of F1 pure crosses: we found that Bonsall Creek F1 marine fish had slightly higher mitochondrial enzyme activities (CS, COX) in their pectoral muscles, but these differences were not significant (Dalziel et al., 2012b). The findings that CS is a significant predictor of U_{crit}* in F2 hybrids may be due to our increased statistical power in the current study or may indicate that the effects of CS activity differ in pure and randomized genetic backgrounds. Although the activity of COX, another mitochondrial enzyme, was significantly correlated with CS activity (r=0.399 in adductor and r=0.413 in abductor muscle; Table 3), regressions between adductor and abductor COX and U_{crit}* were not significant. This discrepancy may result from our methods of measuring enzyme content *via* enzyme activity, because COX activity may not be as reflective of COX enzyme content as CS activity is of CS enzyme content. COX is a multimeric protein that is composed of nuclear and mitochondrial encoded subunits (reviewed by Capaldi, 1996), and is allosterically regulated (e.g. Kadenbach et al., 1997), whereas CS is encoded by a single nuclear gene, and is a homodimer with no known covalent modifications or allosteric regulators (Wiegand and Remington, 1986). Therefore, CS activity is predicted to be a better proxy for enzyme content, and thus mitochondrial content, oxidative capacity and aerobic performance.

It is clear from this study that no single candidate trait that we measured is sufficient to confer a high U_{crit}* in F2 hybrid stickleback, and that even a combination of our candidate traits can only explain ~17.9% of the variation in swimming performance. One possible reason for this limited predictive power is that we did not measure all of the traits that contribute to variation in U_{crit}* in F2 hybrid sticklebacks. For example, Taylor and McPhail (Taylor and McPhail, 1986) found that wild stream-resident fish had lower fin-beat frequencies than marine sticklebacks, and we found that this was also qualitatively true for fish reared in a common garden (A.C.D., personal observations). However, we did not quantitatively measure fin-beat frequency, or any of the other kinematic traits that can impact U_{crit}* (e.g. Walker, 2004). In addition, differences in the axial muscle, which powers caudal bursts, may also affect U_{crit}*. As well, many of the traits we measured are proxies for traits known to influence prolonged swimming in fish (reviewed by Kolok, 1999), such as maximal metabolic rate (e.g. Reidy et al., 2000), cardiac performance (e.g. Claireaux et al., 2005), skeletal muscle metabolic and contractile properties (e.g. Anttila et al., 2008), and drag (reviewed by Langerhans and Reznick, 2009). Although proxies for many of these traits have been themselves significantly correlated with prolonged swimming performance, such as COX activities of the cardiac and skeletal muscles in largemouth bass (Kolok, 1992), it is likely that these proxy traits are not fully representative of the physiological traits we aimed to study. For example, we measured ventricle mass as a proxy for cardiac output, but differences in heart rate (e.g. Eliason et al., 2011) and ventricle shape (e.g. Claireaux et al., 2005) may also influence this trait. In addition, pectoral muscle size and mitochondrial enzyme activity per gram muscle were used as proxies for muscle contractile properties [e.g. fatigue resistance and contraction rate

(e.g. Syme, 2006)], but differences in a number of other traits, such as muscle fuel storage (reviewed by Gibb and Dickson, 2002; Weber, 2011) and the content of calcium handling proteins (e.g. James et al., 2011; Seebacher and Walter, 2012), can also influence muscle power output and endurance.

In addition, many of our candidate traits are predicted to influence maximal metabolic rate ($\dot{V}_{O_2,max}$), but we have not directly measured this trait, or all of the other component traits in the oxygen transport and utilization cascade. Many other traits predicted to influence $\dot{V}_{O_2,max}$ were measured in F1 crosses, and did not differ significantly between stream-resident and marine crosses (e.g. gill surface area, hematocrit, mean cellular hemoglobin content, hemoglobin–oxygen binding affinity) (Dalziel et al., 2012b). However, our findings that CS per gram adductor was a significant predictor of U_{crit}* in F2 hybrids, despite the fact that it was not significantly different in F1 crosses, suggests that this could be the case for other traits (Dalziel et al., 2012b). Finally, we do not believe that behavioral differences contribute to variation in U_{crit}* in this forced, laboratory-based measure of performance (A.C.D., unpublished observations), but behavioral traits may be critical to performance in more ecologically relevant prolonged swimming tasks (e.g. migratory success). One way to find any additional traits that contribute to variation in U_{crit}* is to take an 'unbiased' approach, such as conducting a QTL mapping study. Such an experiment may uncover loci that contribute to variation in U_{crit}*, but for which there are no *a priori* predictions for the effect of a trait on performance (reviewed by Dalziel et al., 2009). Another factor that may have decreased our ability to detect an effect of candidate traits on performance is the relatively small range of U_{crit}* values, and of some of our candidate traits, in our F2 hybrid families relative to the differences between stream-resident and marine populations. Because we did find some variation among families in U_{crit}*, performing our experiments on a larger number of F2 families, and families bred from independent grandparental lines, might be one way to obtain a wider range of values for U_{crit}* and candidate traits.

In principle, it is also possible that the effects of a given candidate trait on U_{crit}* are be determined by complex interactions with other traits (i.e. epistasis). While our multiple linear regressions should detect simple interactions that are constant throughout the data range, they may not detect more complex interactions. Evidence for interactions among traits can come from experiments displaying that the effect of a focal trait varies depending on the genetic background in which is it expressed (reviewed by Demuth and Wade, 2006). To explore this possibility, we split our data into two groups based upon the value of a second 'background' trait, and then re-examined regressions of the first trait against U_{crit}*. We present this data exploration for the three candidate traits found to regress significantly against U_{crit}* in F2 hybrids: ventricle mass, adductor mass and adductor CS (supplementary material Fig. S9). We found that the relationship between candidate traits and U_{crit}* varied depending on the value of a second 'background' trait. For example, in the F2 fish with the largest (top 50%) and smallest (bottom 50%) adductors, the R^2 of adductor CS per gram regressed against U_{crit}* varied from 0.06 to 0.15, and the contributions of each family to this overall relationship varied, further suggesting that genetic background, which also varies among families, needs to be further examined (supplementary material Fig. S9). Other forward genetics crossing designs (reviewed by Dalziel et al., 2009), reverse genetic methods such as gene insertions or deletions (e.g. Colosimo et al., 2005; Chan et al., 2010), and physiological manipulations (e.g. Seebacher and Walter, 2012) may prove to be informative approaches to addressing this question.

U_{crit}* and candidate trait values in F1 and F2 lines

No F2 fish reached U_{crit}* values as high as those of marine fish (Fig. 1A). This is likely due to a dominance of stream-resident alleles, similar to our findings in F1 hybrids, and not because of intrinsic hybrid incompatibilities. Hybrids between stream-resident and marine sticklebacks are viable in the laboratory (e.g. Hagen, 1967; Schluter et al., 2004), and adult hybrids are commonly found in Bonsall Creek (Hagen, 1967) (T. H. Vines and A.C.D., unpublished observations) and throughout the species range (e.g. Jones et al., 2006). In the present study, we could not explicitly test for composite genetic effects (e.g. additive, dominance, epistasis) or increased variation in F2 lines when compared with F1 lines [i.e. segregation variance (see Schluter et al., 2004)] because we only have data for three closely related F2 families. However, our data can provide some insight into the genetic basis for differences in U_{crit}*. For example, we found that no F2 hybrids reached U_{crit}* values outside the range of parental lines (i.e. transgressive segregation was not observed; Fig. 1A), and we did not find any evidence for segregation variance in F2 hybrids (e.g. higher variance in F2 lines than in F1 pure and hybrid lines), indicating that the difference in U_{crit}* between marine and stream-resident stickleback is likely due to a relatively large number of genes with small effect sizes (Lande, 1981).

In contrast to our findings for U_{crit}*, F2 hybrids expressed trait values outside of the range of parental lines for a number of candidate traits (supplementary material Figs S3, S5 and S7). Such transgressive segregation can be due to epistasis or the action of complementary genes, which occurs when genes with 'positive' effects are distributed among parental lines, causing the phenotype in parental lines to be less than the absolute maximum value because the 'positive' effects at some loci are negated by the 'negative' effects at other loci (reviewed by Lynch and Walsh, 1998; Rieseberg et al., 1999). In addition, we found that F2 hybrids displayed a large range of trait values for ventricle masses, abductor masses and body shapes, which might be indicative of segregation variation (supplementary material Figs S3, S5 and S7). However, these data must be interpreted with caution, because we reared F1 and F2 fish in different years (F2 fish were bred and raised 2 years later), and it is possible that environmental differences between years affected our results. We were able to control temperature, salinity, light:dark conditions and feeding between years [see methods in Dalziel et al. (Dalziel et al., 2012a)], but there were differences in fish density that could have affected growth rates and, thus, metabolic traits (e.g. Guderley, 1994; Guderley et al., 2001). It is also possible that differences in age at the time of sampling affected our biochemical measurements, as some F1 fish were up to 6 months older than F2 fish when these traits were measured (up to 1 year and 6–9 months of age *versus* 1 year and 3–4 months of age). In particular, the higher mitochondrial enzyme activities (CS and COX) in the pectoral muscles of F2 fish when compared with F1 fish may be due to senescence in the older F1 fish, as CS activity is known to decline in axial muscles by 2 years of age in migratory marine sticklebacks (Dufresne et al., 1990; Guderley et al., 2001).

Traits contributing to the capacity for whole-animal performance

Other studies that have attempted to correlate prolonged swimming in fishes with underlying morphological, physiological and biochemical traits have had mixed success (e.g. Kolok, 1992; Kolok and Farrell, 1994; Garenc et al., 1999; Gibb and Dickson, 2002; Odell et al., 2003; Claireaux et al., 2005). For example, intra-individual variation in $\dot{V}_{O_2,max}$ in Trinidadian guppies (*Poecilia reticulata*) was not significantly correlated with any of the six measured candidate traits (swimming muscle, heart and gill size

and muscle CS, LDH and myofibrillar ATPase activities) after corrections for multiple comparisons (Odell et al., 2003). Gibb and Dickson (Gibb and Dickson, 2002) found that muscle aerobic enzyme activities (red muscle, white muscle and heart CS activity, red muscle and heart 3-hydroxy-o-acylCoA dehydrogenase activity and myoglobin content) were not significantly correlated with swimming performance in two scombrid fishes [(kawakawa tuna (*Euthynnus affinis*) and chub mackerel (*Scomber japonicus*)]. However, Kolok (Kolok, 1992) was able to predict up to 73% of variation U_{crit} capacity in wild-caught largemouth bass (*Micropterus salmoides*) by measuring variation in red muscle COX, condition factor and gill filament density. Interestingly, in this experiment on wild-caught fish, the predictors of endurance swimming (a set velocity test) varied with season, suggesting that trait plasticity, and not just genetically based differences in mean values, may be critical to performance (Kolok, 1992).

Studies of burst swimming performance in fishes have also found that the predictors of swimming performance can vary within a population over time. For example, Garenc et al. (Garenc et al., 1999) found that axial muscle COX and PK activities were significant predictors of burst swimming capacity in adult threespine sticklebacks, but not in juvenile fish, suggesting that changes associated with reproduction in adults may result in muscle enzyme levels limiting performance. Many studies of intra-individual differences in endurance exercise capacity and $\dot{V}_{O2,max}$ in other vertebrate species have successfully linked underlying traits with performance (e.g. Garland, 1984; Garland and Else, 1987; Garland and Bennett, 1990; Longphre and Gatten, 1994; Chappell et al., 1999; Hammond et al., 2000), but other studies have been unable to link performance to any of the measured underlying candidate traits (e.g. Bennett et al., 1989; Chappell et al., 2007). These studies, in combination with our results, argue that determining the predictors of locomotory capacity is a complex problem, and that even when predictors can be identified they are often dependent upon species, genetic background, age, sex, reproductive status and season. These studies also highlight the importance of performing experiments under ecologically relevant conditions if differences in morphology/physiology and performance are to be linked to fitness (reviewed by Arnold, 1983; Kingsolver and Huey, 2003).

ACKNOWLEDGEMENTS

We thank W. E. Vandersteen and S. M. Rogers for their advice on crossing designs and line cross analyses, and D. Schluter, E. B. Taylor and J. G. Richards for helpful comments on earlier versions of this manuscript.

FUNDING

This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) through discovery grants to P.M.S. and a Canada Graduate Scholarship to A.C.D. A.C.D was also supported by a University of British Columbia University Graduate pre-doctoral fellowship.

REFERENCES

Achard, S. (2010). brainwaver: basic wavelet analysis of multivariate time series with visualization and parameratization using graph theory. R package version 1.5, available at http://cran.r-project.org/src/contrib/Archive/brainwaver/.

Anttila, K., Järvilehto, M. and Mänttäri, S. (2008). The swimming performance of brown trout and whitefish: the effects of exercise on Ca^{2+} handling and oxidative capacity of swimming muscles. *J. Comp. Physiol. B* **178**, 465-475.

Arnold, S. J. (1983). Morphology, performance, and fitness. *Am. Zool.* **23**, 347-361.

Bell, M. A. and Foster, S. A. (1994). The evolutionary biology of the threespine stickleback. Oxford: Oxford University Press.

Bennett, A. F., Garland, T., Jr and Else, P. L. (1989). Individual correlation of morphology, muscle mechanics, and locomotion in a salamander. *Am. J. Physiol.* **256**, R1200-R1208.

Bernal, D., Dickson, K. A., Shadwick, R. E. and Graham, J. B. (2001). Analysis of the evolutionary convergence for high performance swimming in lamnid sharks and tunas. *Comp. Biochem. Physiol.* **129A**, 695-726.

Blake, R. W. (2004). Fish functional design and swimming performance. *J. Fish Biol.* **65**, 1193-1222.

Brauner, C. J., Val, A. L. and Randall, D. J. (1993). The effect of graded methemoglobin levels on the swimming performance of chinook salmon (*Onchorhynchus tshawtscha*). *J. Exp. Biol.* **185**, 121-135.

Brauner, C. J., Matey, V., Zhang, W., Richards, J. G., Dhillon, R., Cao, Z. D., Wang, Y. and Fu, S. J. (2011). Gill remodeling in crucian carp during sustained exercise and the effect on subsequent swimming performance. *Physiol. Biochem. Zool.* **84**, 535-542.

Brett, J. R. (1964). The respiratory metabolism and swimming performance of young sockeye salmon. *J. Fish. Res. Board Can.* **21**, 1183-1226.

Burton, R. S., Ellison, C. K. and Harrison, J. S. (2006). The sorry state of F2 hybrids: consequences of rapid mitochondrial DNA evolution in allopatric populations. *Am. Nat.* **168 Suppl. 6**, S14-S24.

Capaldi, R. A. (1996). The complexity of a respiratory complex. *Nat. Struct. Biol.* **3**, 570-574.

Chan, Y. F., Marks, M. E., Jones, F. C., Villarreal, G., Shapiro, M. D., Brady, S. D., Southwick, A. M., Absher, D. M., Grimwood, J., Schmutz, J. et al. (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* **327**, 302-305.

Chappell, M. A., Bech, C. and Buttemer, W. A. (1999). The relationship of central and peripheral organ masses to aerobic performance variation in house sparrows. *J. Exp. Biol.* **202**, 2269-2279.

Chappell, M. A., Garland, T., Jr, Robertson, G. F. and Saltzman, W. (2007). Relationships among running performance, aerobic physiology and organ mass in male Mongolian gerbils. *J. Exp. Biol.* **210**, 4179-4197.

Claireaux, G., McKenzie, D. J., Genge, A. G., Chatelier, A., Aubin, J. and Farrell, A. P. (2005). Linking swimming performance, cardiac pumping ability and cardiac anatomy in rainbow trout. *J. Exp. Biol.* **208**, 1775-1784.

Colosimo, P. F., Hosemann, K. E., Balabhadra, S., Villarreal, G., Dickson, M., Grimwood, J., Schmutz, J., Myers, R. M., Schluter, D. and Kingsley, D. M. (2005). Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307**, 1928-1933.

Dalziel, A. C., Rogers, S. M. and Schulte, P. M. (2009). Linking genotypes to phenotypes and fitness: how mechanistic biology can inform molecular ecology. *Mol. Ecol.* **18**, 4997-5017.

Dalziel, A. C., Vines, T. H. and Schulte, P. M. (2012a). Reductions in prolonged swimming capacity following freshwater colonizatoin in multiple threespine stickleback populations. *Evolution* **66**, 1226-1239.

Dalziel, A. C., Ou, M. and Schulte, P. M. (2012b). Mechanisms underlying parallel reductions in aerobic capacity in non-migratory threespine stickleback (*Gasterosteus aculeatus*) populations. *J. Exp. Biol.* **215**, 746-759.

Demuth, J. P. and Wade, M. J. (2006). Experimental methods for measuring gene interactions. *Annu. Rev. Ecol. Evol. Syst.* **37**, 289-316.

Dufresne, F., FitzGerald, G. J. and Lachance, S. (1990). Age and size-related differences in reproductive success and reproductive costs in threespine sticklebacks (*Gasterosteus aculeatus*). *Behav. Ecol.* **1**, 140-147.

Eliason, E. J., Clark, T. D., Hague, M. J., Hanson, L. M., Gallagher, Z. S., Jeffries, K. M., Gale, M. K., Patterson, D. A., Hinch, S. G. and Farrell, A. P. (2011). Differences in thermal tolerance among sockeye salmon populations. *Science* **332**, 109-112.

Farrell, A. P. (2002). Cardiorespiratory performance in salmonids during exercise at high temperature: insights into cardiovascular design limitations in fishes. *Comp. Biochem. Physiol.* **132A**, 797-810.

Fraser, D. J., Weir, L. K., Bernatchez, L., Hansen, M. M. and Taylor, E. B. (2011). Extent and scale of local adaptation in salmonid fishes: review and meta-analysis. *Heredity* **106**, 404-420.

Gallaugher, P., Thorarensen, H. and Farrell, A. P. (1995). Hematocrit in oxygen transport and swimming in rainbow trout (*Oncorhynchus mykiss*). *Respir. Physiol.* **102**, 279-292.

Garenc, C., Couture, P., Laflamme, M. A. and Guderley, H. (1999). Metabolic correlates of burst swimming capacity of juvenile and adult threespine stickleback (*Gasterosteus aculeatus*). *J. Comp. Physiol. B* **169**, 113-122.

Garland, T., Jr (1984). Physiological correlates of locomotory performance in a lizard: an allometric approach. *Am. J. Physiol.* **247**, R806-R815.

Garland, T., Jr and Bennett, A. F. (1990). Quantitative genetics of maximal oxygen consumption in a garter snake. *Am. J. Physiol.* **259**, R986-R992.

Garland, T., Jr and Else, P. L. (1987). Seasonal, sexual, and individual variation in endurance and activity metabolism in lizards. *Am. J. Physiol.* **252**, R439-R449.

Gibb, A. C. and Dickson, K. A. (2002). Functional morphology and biochemical indices of performance: Is there a correlation between metabolic enzyme activity and swimming performance? *Integr. Comp. Biol.* **42**, 199-207.

Guderley, H. (1994). Physiological ecology and evolution of the threespine stickleback. In *The Evolutionary Biology of the Threespine Stickleback* (ed. M. A. Bell and S. A. Foster), pp. 85-113. Oxford: Oxford University Press.

Guderley, H., Leroy, P. H. and Gagné, A. (2001). Thermal acclimation, growth, and burst swimming of threespine stickleback: enzymatic correlates and influence of photoperiod. *Physiol. Biochem. Zool.* **74**, 66-74.

Hagen, D. W. (1967). Isolating mechanisms in threespine sticklebacks (*Gasterosteus*). *J. Fish. Res. Board Can.* **24**, 1637-1692.

Hammond, K. A., Chappell, M. A., Cardullo, R. A., Lin, R. S. and Johnsen, T. S. (2000). The mechanistic basis of aerobic performance variation in red junglefowl. *J. Exp. Biol.* **203**, 2053-2064.

Hoffmann, E. and Borg, B. (2006). Sex differences in pectoral muscles but not in pectoral fins in the three-spined stickleback, *Gasterosteus aculeatus*. *J. Fish Biol.* **68**, 1451-1459.

Husak, J. F. and Fox, S. F. (2008). Sexual selection on locomotor performance. *Evol. Ecol. Res.* **10**, 213-228.

Irschick, D. J., Meyers, J. J., Husak, J. F. and Le Galliard, J. F. (2008). How does selection operate on whole-organism functional performance capacities? A review and synthesis. *Evol. Ecol. Res.* **10**, 177-196.
James, R. S., Walter, I. and Seebacher, F. (2011). Variation in expression of calcium-handling proteins is associated with inter-individual differences in mechanical performance of rat (*Rattus norvegicus*) skeletal muscle. *J. Exp. Biol.* **214**, 3542-3548.
Jones, D. R. and Randall, D. J. (1978). The respiratory and circulatory systems during exercise. In *Fish Physiology: Locomotion*, Vol. 7 (ed. W. S. Hoar and J. D. Randall), pp. 101-189. New York: Academic Press.
Jones, F. C., Brown, C., Pemberton, J. M. and Braithwaite, V. A. (2006). Reproductive isolation in a threespine stickleback hybrid zone. *J. Evol. Biol.* **19**, 1531-1544.
Kadenbach, B., Frank, V., Rieger, T. and Napiwotzki, J. (1997). Regulation of respiration and energy transduction in cytochrome c oxidase isozymes by allosteric effectors. *Mol. Cell. Biochem.* **174**, 131-135.
Kingsolver, J. G. and Huey, R. B. (2003). Introduction: the evolution of morphology, performance, and fitness. *Integr. Comp. Biol.* **43**, 361-366.
Kolok, A. S. (1992). Morphological and physiological correlates with swimming performance in juvenile largemouth bass. *Am. J. Physiol.* **263**, R1042-R1048.
Kolok, A. S. (1999). Interindividual variation in the prolonged locomotor performance of ectothermic vertebrates: a comparison of fish and herpetofaunal methodologies and a brief review of the recent fish literature. *Can. J. Fish. Aquat. Sci.* **56**, 700-710.
Kolok, A. S. and Farrell, A. P. (1994). Individual variation in the swimming performance and cardiac performance of northern squawfish, *Ptychocheilus oregonensis*. *Physiol. Zool.* **67**, 706-722.
Korsmeyer, K. E., Steffensen, J. F. and Herskin, J. (2002). Energetics of median and paired fin swimming, body and caudal fin swimming, and gait transition in parrotfish (*Scarus schlegeli*) and triggerfish (*Rhinecanthus aculeatus*). *J. Exp. Biol.* **205**, 1253-1263.
Lande, R. (1981). The minimum number of genes contributing to quantitative variation between and within populations. *Genetics* **99**, 541-553.
Langerhans, R. B. and Reznick, D. N. (2009). Ecology and evolution of swimming performance in fishes: predicting evolution with biomechanics. In *Fish Locomotion: an Etho-ecological Perspective* (ed. P. Domenici and B. G. Kapoor), pp. 200-248. Enfield, NH: Science Publishers.
Lee, C. G., Farrell, A. P., Lotto, A., MacNutt, M. J., Hinch, S. G. and Healey, M. C. (2003). The effect of temperature on swimming performance and oxygen consumption in adult sockeye (*Oncorhynchus nerka*) and coho (*O. kisutch*) salmon stocks. *J. Exp. Biol.* **206**, 3239-3251.
Lee, W. S., Monaghan, P. and Metcalfe, N. B. (2010). The trade-off between growth rate and locomotor performance varies with perceived time until breeding. *J. Exp. Biol.* **213**, 3289-3298.
Longphre, M. and Gatten, R. E. (1994). Individual variability in sustained performance, aerobic metabolism, oxygen transport, and enzyme activity in toads (*Bufo marinus*). *Copeia* **1994**, 887-896.
Lynch, M. and Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sunderland, MA: Sinauer Associates.
McPhail, J. D. (1994). Speciation and the evolution of reproductive isolation in the sticklebacks (*Gasterosteus aculeatus*) of south-western British Columbia. In *The Evolutionary Biology of the Threespine Stickleback* (ed. M. A. Bell and S. A. Foster), pp. 399-437. Oxford: Oxford University Press.
Odell, J. P., Chappell, M. A. and Dickson, K. A. (2003). Morphological and enzymatic correlates of aerobic and burst performance in different populations of Trinidadian guppies *Poecilia reticulata*. *J. Exp. Biol.* **206**, 3707-3718.
Ostlund-Nilsson, S., Mayer, I. and Huntingford, F. (2007). Biology of the three-spined stickleback. New York: CRC Press.
Pearson, M. P. and Stevens, E. D. (1991). Splenectomy impairs aerobic swim performance in trout. *Can. J. Zool.* **69**, 2089-2092.
Pinheiro, J. C. and Bates, D. M. (2000). *Mixed-effects Models in S and S-PLUS*. New York: Springer.
Plaut, I. (2001). Critical swimming speed: its ecological relevance. *Comp. Biochem. Physiol. A* **131**, 41-50.
R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
Reidy, S. P., Kerr, S. R. and Nelson, J. A. (2000). Aerobic and anaerobic swimming performance of individual Atlantic cod. *J. Exp. Biol.* **203**, 347-357.
Rieseberg, L. H., Archer, M. A. and Wayne, R. K. (1999). Transgressive segregation, adaptation and speciation. *Heredity* **83**, 363-372.
Rohlf, J. F. (2010). tpsDig v2.16. Department of Ecology and Evolution, State University of New York, Stony Brook, NY, USA.
Schaarschmidt, T. and Jürss, K. (2003). Locomotory capacity of Baltic Sea and freshwater populations of the threespine stickleback (*Gasterosteus aculeatus*). *Comp. Biochem. Physiol.* **135A**, 411-424.
Schluter, D., Clifford, E. A., Nemethy, M. and McKinnon, J. S. (2004). Parallel evolution and inheritance of quantitative traits. *Am. Nat.* **163**, 809-822.
Seebacher, F. and Walter, I. (2012). Differences in locomotor performance between individuals: importance of parvalbumin, calcium handling and metabolism. *J. Exp. Biol.* **215**, 663-670.
Sharpe, D. M. T., Rasanen, K., Berner, D. and Hendry, A. P. (2008). Genetic and environmental contributions to the morphology of lake and stream stickleback: implications for gene flow and reproductive isolation. *Evol. Ecol. Res.* **10**, 849-866.
Slinker, B. K. and Glantz, S. A. (1985). Multiple regression for physiological data analysis: the problem of multicollinearity. *Am. J. Physiol.* **249**, R1-R12.
Svendsen, J. C., Tudorache, C., Jordan, A. D., Steffensen, J. F., Aarestrup, K. and Domenici, P. (2010). Partition of aerobic and anaerobic swimming costs related to gait transitions in a labriform swimmer. *J. Exp. Biol.* **213**, 2177-2183.
Syme, D. A. (2006). Functional properties of skeletal muscle. In *Fish Biomechanics*, Vol. 23 (ed. R. E. Shadwick and G. V. Lauder), pp. 179-240. San Diego, CA: Academic Press.
Taylor, E. B. and McPhail, J. D. (1986). Prolonged and burst swimming in anadromous and freshwater threespine stickleback, *Gasterosteus aculeatus*. *Can. J. Zool.* **64**, 416-420.
Tudorache, C., Blust, R. and De Boeck, G. (2007). Swimming capacity and energetics of migrating and non-migrating morphs of three-spined stickleback *Gasterosteus aculeatus* L. and their ecological implications. *J. Fish Biol.* **71**, 1448-1456.
Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S*. New York: Springer.
Wainwright, P. C., Alfaro, M. E., Bolnick, D. I. and Hulsey, C. D. (2005). Many-to-one mapping of form to function: a general principle in organismal design? *Integr. Comp. Biol.* **45**, 256-262.
Walker, J. A. (2004). Dynamics of pectoral fin rowing in a fish with an extreme rowing stroke: the threespine stickleback (*Gasterosteus aculeatus*). *J. Exp. Biol.* **207**, 1925-1939.
Walker, J. A. (2007). A general model of functional constraints on phenotypic evolution. *Am. Nat.* **170**, 681-689.
Walker, J. A. (2010). An integrative model of evolutionary covariance: a symposium on body shape in fishes. *Integr. Comp. Biol.* **50**, 1051-1056.
Webb, P. W. (1982). Locomotor patterns in the evolution of actinopterygian fishes. *Am. Zool.* **22**, 329-342.
Weber, J. M. (2011). Metabolic fuels: regulating fluxes to select mix. *J. Exp. Biol.* **214**, 286-294.
Weihs, D. and Webb, P. W. (1983). Optimization of locomotion. In *Fish Biomechanics* (ed. P. W. Webb and D. Weihs), pp. 339-371. New York: Praeger Publishers.
Wiegand, G. and Remington, S. J. (1986). Citrate synthase: structure, control, and mechanism. *Annu. Rev. Biophys. Biophys. Chem.* **15**, 97-117.
Wootton, R. J. (1984). *A Functional Biology of Sticklebacks*. London: Croom Helm.
Zelditch, M. L., Swiderski, D. L., Sheets, H. D. and Fink, W. L. (2004). *Geometric Morphometrics for Biologists: a Primer*. San Diego, CA: Academic Press.
Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A. and Smith, G. M. (2009). *Mixed Effects Models and Extensions in Ecology with R*. New York: Springer.

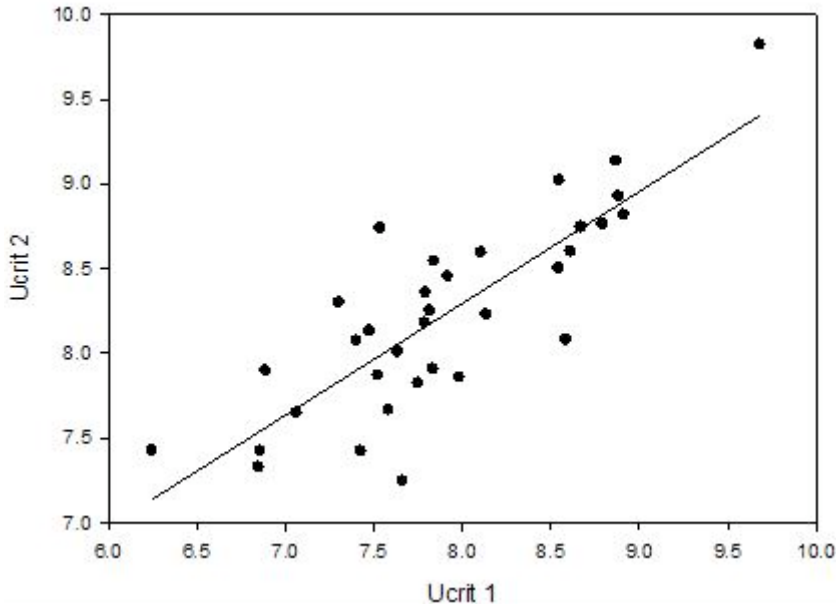

Fig. S1. Repeatability of U_{crit} in Bonsall Creek F2 hybrid sticklebacks. U_{crit} 2 was measured 2 weeks after U_{crit} 1 [$U_{crit}2=0.660\times(U_{crit}1)+3.015$, d.f.=33, $t=8.27$, $P<0.001$, $r^2=0.671$].

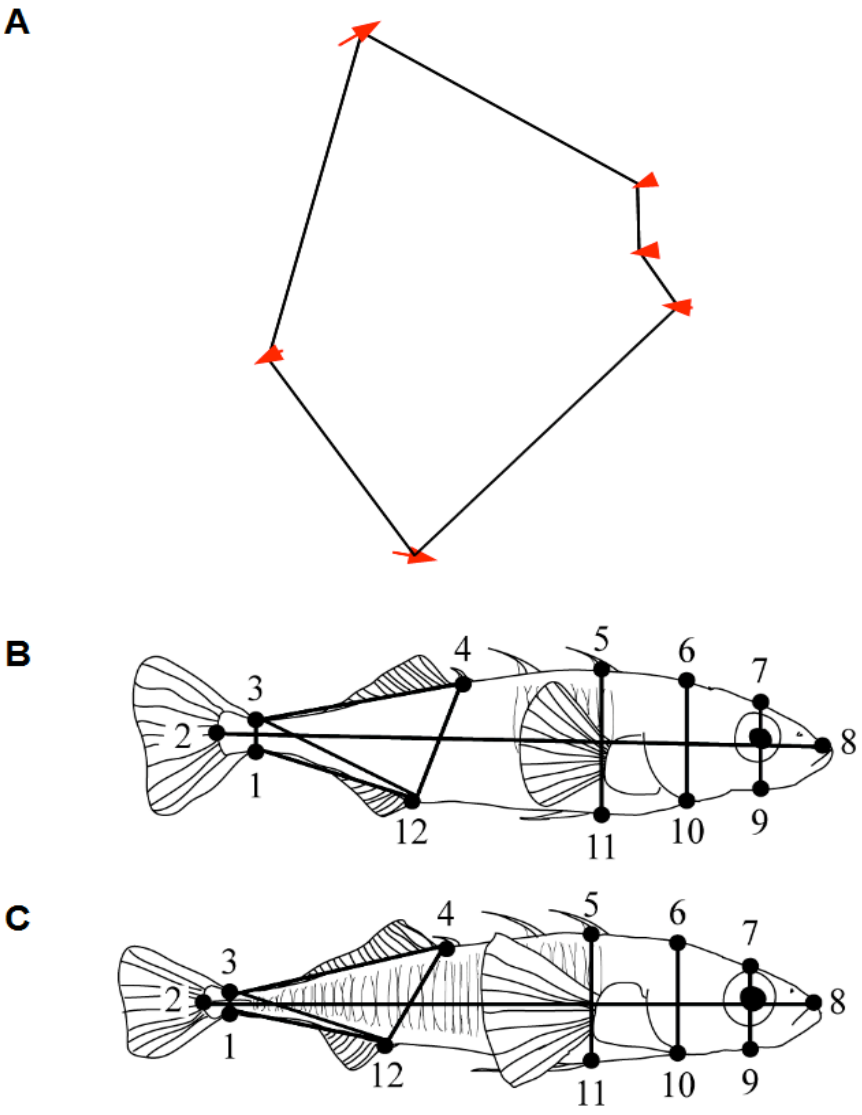

Fig. S2. (A) Landmarks used to measure pectoral fin shape are located at the center of the nearest arrow. Arrows multiply by four the changes in landmark position that occur among cross-types for pectoral fin shape linear discriminant 1 (ld1), and generally summarize changes in fin shape from a marine to a stream-resident fish. Representative stream-resident (B) and anadromous (C) sticklebacks showing landmarks (numbered circles) used to measure the five body shape variables used in this paper (from black lines connecting landmarks).

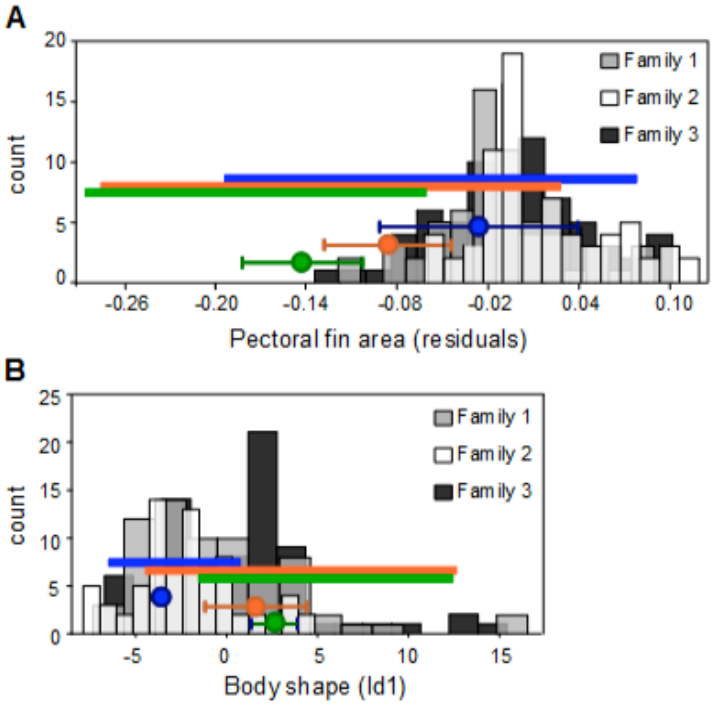

Fig. S3. Histograms of residual (A) pectoral fin area and (B) body shape ld1 values in F2 fish. The full range of values measured for all F1 individuals of a given cross type (green, pure stream-resident crosses; orange, F1 hybrid crosses; blue, pure marine crosses) are represented by thick colored bars, and the mean values ± s.d. for F1 line crosses are denoted with circles and thin lines (data from Dalziel et al. 2012a).

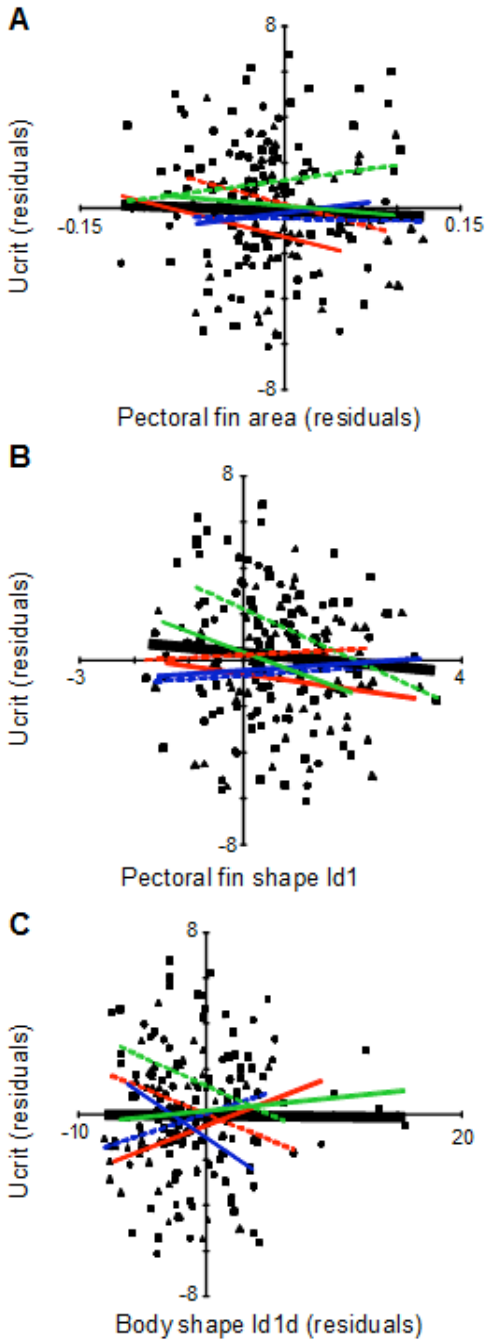

Fig. S4. The relationship between U_{crit} and (A) residual pectoral fin area, (B) pectoral fin shape ld1 (low values indicate marine shaped fins and high values indicate stream-resident shaped fins; see Fig. S2) and (C) body shape ld1 in F2 fish. The thick black line represents the fitted values for the whole population (with family and sex as nested random effects) for residual pectoral fin area ($F_{1,179}$=0.485, P=0.487), pectoral fin shape ld1 ($F_{1,182}$=1.471, P=0.2267) and body shape ld1($F_{1,175}$=0.005, P=0.941) (Table 2). The colored lines represent the fitted values for each family and sex (red dashed line, family 1 females; solid red line, family 1 males; blue dashed line, family 2 females; solid blue line, family 2 males; green dashed line, family 3 females; solid green line, family 3 males). Circles symbolize values for family 1 fish, triangles symbolize family 2 fish, and squares symbolize family 3 fish.

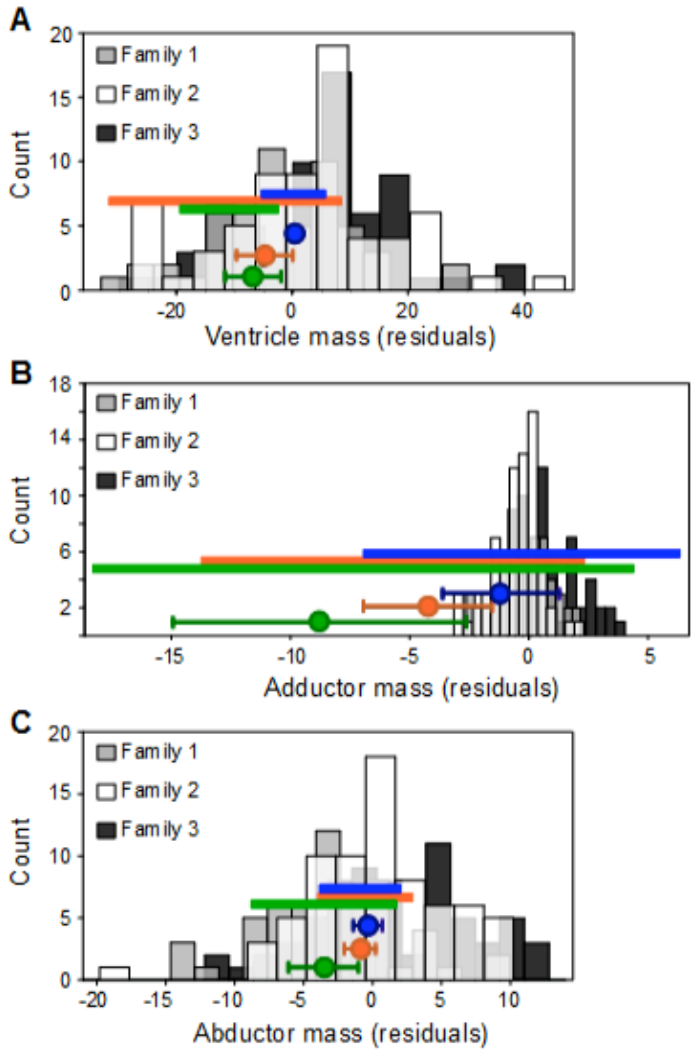

Fig. S5. Histograms of residual (A) ventricle mass, (B) residual pectoral adductor mass and (C) residual pectoral abductor mass of F2 fish. The full range of values measured for all F1 individuals of a given cross type (green, pure stream-resident crosses; orange, F1 hybrid crosses; blue, pure marine crosses) are represented by thick colored bars, and the mean values ± s.d. for F1 line crosses are denoted with circles and thin lines (data from Dalziel et al. 2012a).

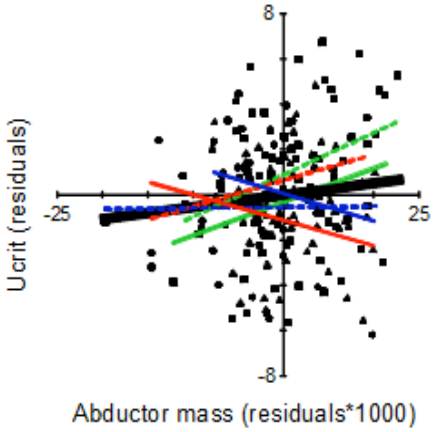

Fig. S6. The relationship between U_{crit} and residual abductor mass in F2 fish, with data presented as in Fig. S4. Thick black lines represent the fitted values for residual abductor mass ($F_{1,183}$=2.385, P=0.124) for all data (Table 2). Data for ventricle and adductor mass are presented in Fig. 2.

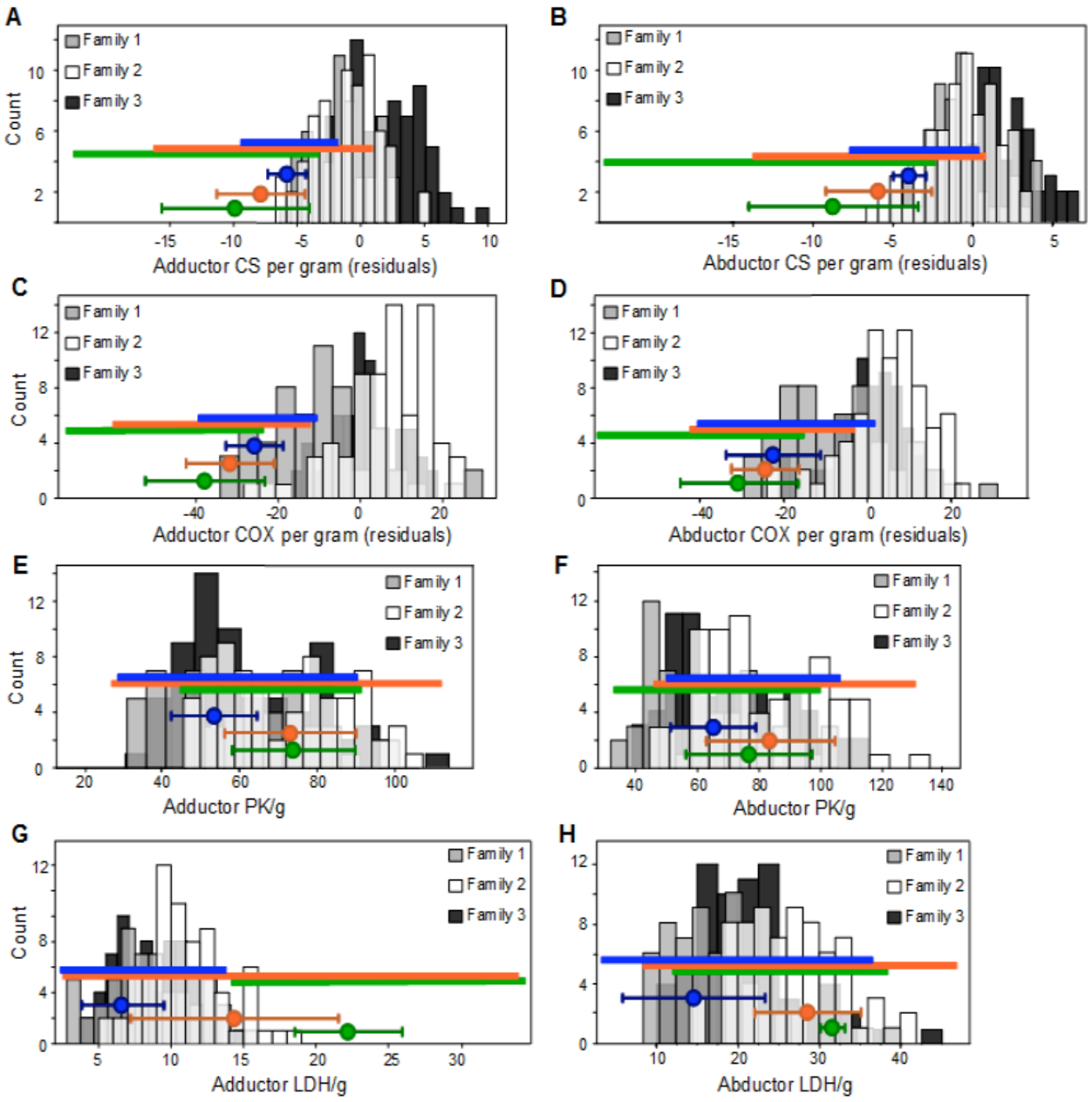

Fig. S7. Histograms of residual citrate synthase (CS) activity per gram adductor (A) and abductor (B) muscle, residual cytochrome *c* oxidase (COX) activity per gram adductor (C) and abductor (D) muscle, pyruvate kinase per gram adductor (E) and abductor (F) muscle, and lactate dehydrogenase (LDH) activity per gram adductor (G) and abductor (H) muscle in F2 fish. There is a strong effect of sex on PK activities (E,F), resulting in a bimodal distribution. The full range of values measured for all F1 individuals of a given cross type (green, pure stream-resident crosses; orange, F1 hybrid crosses; blue, pure marine crosses) are represented by thick colored bars, and the mean values ± s.d. for F1 line crosses are denoted with circles and thin lines (data from Dalziel et al. 2012b).

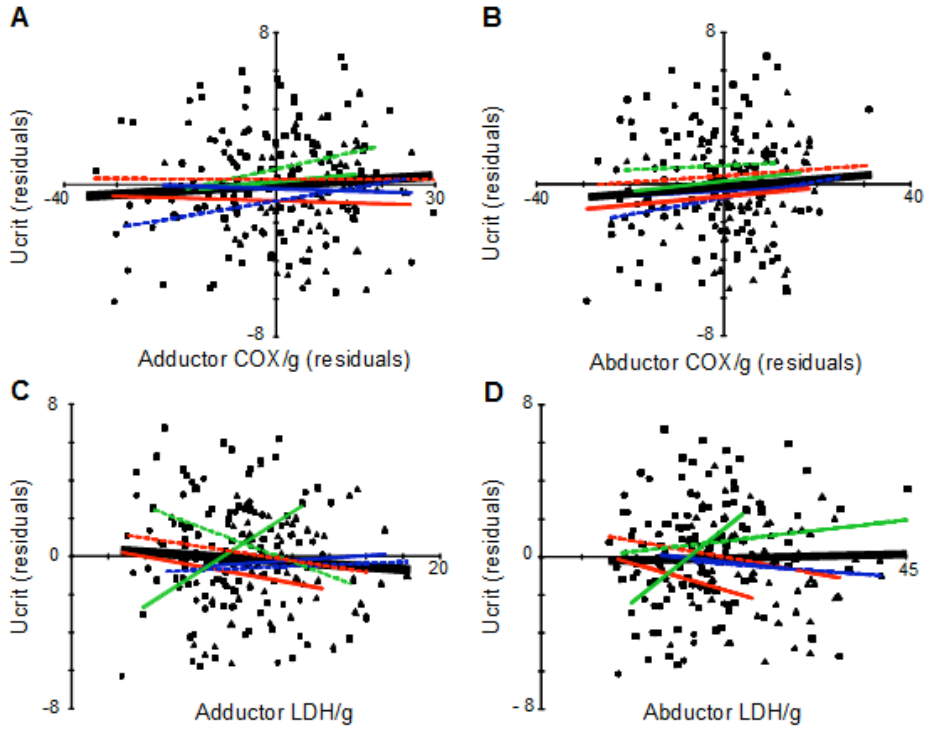

Fig. S8. The relationship between U_{crit} and (A) residual COX and (C) LDH activity per gram adductor muscle in F2 fish, and (B) residual COX and (D) LDH activity per gram abductor muscle in F2 fish. Data are presented as in Fig. S4. Thick black lines represent the fitted values for residual COX per gram adductor ($F_{1,184}$=1.023, P=0.313), LDH per gram adductor ($F_{1,184}$=0.948, P=0.332), residual COX per gram abductor ($F_{1,183}$=1.450, P=0.230) and LDH per gram abductor ($F_{1,183}$=0.112, P=0.738) for all data (Table 2). Data for CS are presented in Fig. 2.

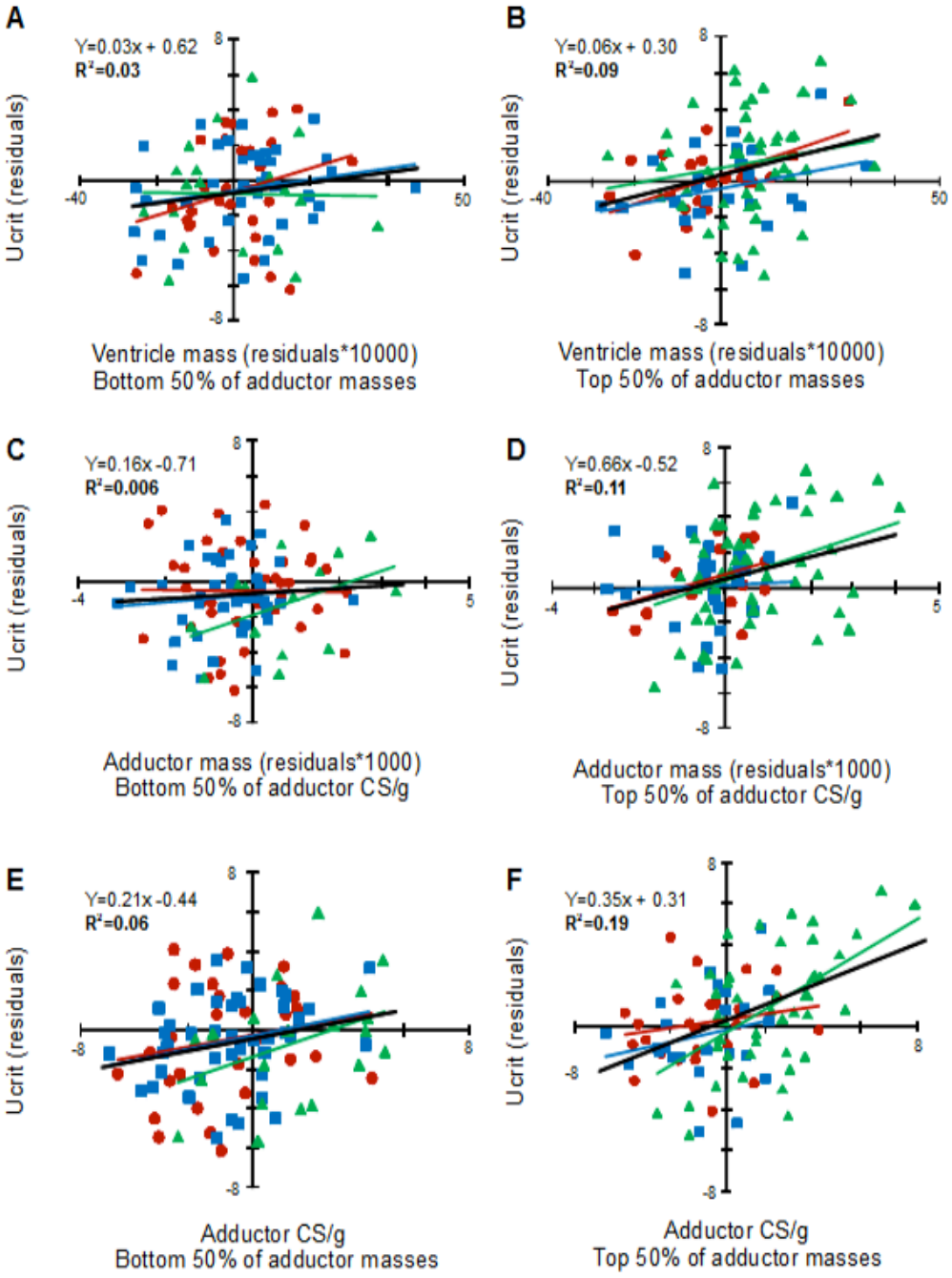

Fig. S9. The relationship between U_{crit} and (A,B) residual ventricle mass, (C,D) residual adductor mass and (E,F) residual adductor CS per gram for the highest 50% of values (B,D,F) and lowest 50% of values (A,C,E) of a second 'background' trait listed on the *x*-axis. The thick black line represents the fitted curve for all values, and colored lines represent the fitted curves for each individual F2 family (red, family 1; blue, family 2; green, family 3).

Table S1. Correlations among explanatory variables (fixed-effects only)

	U_{crit}*	Fineness	Caudal peduncle depth	Caudal area	Posterior depth	Head depth
U_{crit}*	–	r=–0.117 P=0.094	r=–0.058 P=0.409	r=–0.117 P=0.096	r=0.074 P=0.292	r=–0.021 P=0.769
Fineness	–	–	**r=–0.277** **P=0.00006**	r=0.082 P=0.242	**r=–0.269** **P=0.000094**	r=–0.035 P=0.617
Caudal peduncle depth	–	–	–	**r=0.350** **P<0.000001**	**r=0.333** **P<0.000001**	r=0.117 P=0.094
Caudal area	–	–	–	–	**r=0.522** **P<0.000001**	r=–0.029 P=0.682
Posterior depth	–	–	–	–	–	r=0.083 P=0.234
Head depth	–	–	–	–	–	–

When corrected for multiple comparisons, the cut-off for significant P-values is 0.000094. Significantly correlations are bolded.