MOLECULAR ECOLOGY

Molecular Ecology (2012) doi: 10.1111/j.1365-294X.2012.05660.x

Admixture mapping of male nuptial colour and body shape in a recently formed hybrid population of threespine stickleback

TIFFANY B. MALEK,*†[1] JANETTE W. BOUGHMAN,‡[1] IAN DWORKIN‡ and CATHERINE L. PEICHEL*

*Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA, †Graduate Program in Molecular and Cellular Biology, University of Washington, Seattle, WA 98195, USA, ‡Department of Zoology & BEACON, Michigan State University, East Lansing, MI 48824, USA

Abstract

Despite recent progress, we still know relatively little about the genetic architecture that underlies adaptation to divergent environments. Determining whether the genetic architecture of phenotypic adaptation follows any predictable patterns requires data from a wide variety of species. However, in many organisms, genetic studies are hindered by the inability to perform genetic crosses in the laboratory or by long generation times. Admixture mapping is an approach that circumvents these issues by taking advantage of hybridization that occurs between populations or species in the wild. Here, we demonstrate the utility of admixture mapping in a naturally occurring hybrid population of threespine sticklebacks (*Gasterosteus aculeatus*) from Enos Lake, British Columbia. Until recently, this lake contained two species of sticklebacks adapted to divergent habitats within the lake. This benthic–limnetic species pair diverged in a number of phenotypes, including male nuptial coloration and body shape, which were previously shown to contribute to reproductive isolation between them. However, recent ecological disturbance has contributed to extensive hybridization between the species, and there is now a single, admixed population within Enos Lake. We collected over 500 males from Enos Lake and found that most had intermediate nuptial colour and body shape. By genotyping males with nuptial colour at the two extremes of the phenotypic distribution, we identified seven genomic regions on three chromosomes associated with divergence in male nuptial colour. These genomic regions are also associated with variation in body shape, suggesting that tight linkage and/or pleiotropy facilitated adaptation to divergent environments in benthic–limnetic species pairs.

Keywords: body shape, *Gasterosteus aculeatus*, genetic mapping, hybridization, male breeding colour, species pair

Received 13 January 2012; revision received 10 April 2012; accepted 18 April 2012

Introduction

A major challenge in evolutionary biology is to understand the genetic mechanisms by which organisms adapt to divergent environments. Despite recent progress, many questions about the genetics of adaptation still remain and addressing these questions will require empirical data from a variety of species and systems (Hoekstra & Coyne 2007; Stern & Orgogozo 2008). The development of genetic and genomic resources for a number of organisms has made it possible to use methods such as quantitative trait locus (QTL) mapping to investigate the genetic changes that occur during adaptation (reviewed by Orr 2005). However, QTL mapping

Correspondence: Janette W. Boughman, Fax: (517) 432-2789; E-mail: boughman@msu.edu

[1]These authors contributed equally to this work.

has not been feasible in most natural populations of vertebrate species or in any taxa with a large physical size and a long generation time. As an alternative, population genomic approaches that use genome-wide panels of molecular markers to identify regions of the genome with genetic signatures of selection have been used (Luikart *et al.* 2003; Schlotterer 2003; Storz 2005; Stinchcombe & Hoekstra 2008; Nosil *et al.* 2009; Hohenlohe *et al.* 2010; Stapley *et al.* 2010). An obstacle in these studies is the difficulty of correlating molecular signatures of selection with the actual phenotypic traits under selection (Stinchcombe & Hoekstra 2008; Nielson 2009; Nadeau & Jiggins 2010).

Alternative mapping approaches circumvent these limitations by taking advantage of historical recombination events that have occurred within a population to identify genotype–phenotype associations in naturally interbreeding populations (Slate 2005; Buerkle & Lexer 2008; Stinchcombe & Hoekstra 2008). This linkage disequilibrium (LD) or association mapping approach has now been widely used in human genetic studies (Hindorff *et al.* 2011). Whole-genome association studies generally require a very dense map of genetic markers when populations have been interbreeding for hundreds to thousands of generations (Stinchcombe & Hoekstra 2008). Although the recent development of genotyping strategies based on next-generation sequencing suggests that association mapping will be feasible in any organism (Nadeau & Jiggins 2010; Stapley *et al.* 2010; Davey *et al.* 2011), the number of individuals required and the associated costs of genotyping can still be prohibitive (Winkler *et al.* 2010). As an alternative, it is possible to use 'admixture mapping' in recently interbreeding populations. Although the regions of LD are much larger in recently interbreeding populations, resulting in lower resolution than can be obtained by association mapping, fewer individuals and markers are needed for an initial mapping study (Winkler *et al.* 2010). Admixture mapping has now been widely used to identify loci for human disease susceptibility (Winkler *et al.* 2010), and to identify genomic regions that contribute to reproductive isolation and phenotypic divergence in nature (Rieseberg & Buerkle 2002; Buerkle & Lexer 2008; Lexer *et al.* 2010).

In this study, we use admixture mapping to study the genetics of adaptive traits that also contribute to reproductive isolation in natural populations of threespine sticklebacks (*Gasterosteus aculeatus*). These fish are known for their rapid morphological, physiological and behavioural divergence, which occurred after marine ancestors invaded newly created freshwater lakes and streams at the end of the last ice age 10–15 000 years ago (Bell & Foster 1994). Divergent populations can be crossed in the laboratory, and recently available genetic and genomic tools (Peichel *et al.* 2001; Kingsley *et al.* 2004; Kingsley & Peichel 2007; Hohenlohe *et al.* 2010; Jones *et al.* 2012) can be used to investigate the genetic mechanisms of adaptation in these fish. In particular, QTL mapping has been performed to identify regions of the genome and, in some cases, the actual genes that underlie variation in a number of morphological and behavioural traits (Peichel *et al.* 2001; Colosimo *et al.* 2004, 2005; Cresko *et al.* 2004; Shapiro *et al.* 2004; Kimmel *et al.* 2005; Miller *et al.* 2007; Albert *et al.* 2008; Kitano *et al.* 2009; Chan *et al.* 2010; Greenwood *et al.* 2011). Although QTL mapping with laboratory crosses of sticklebacks has been successful, there are some limitations that can be overcome by admixture mapping. First, taking advantage of historical recombination events that have occurred in a natural population circumvents the time and space required to rear the large number of fish required for QTL mapping. Second, some phenotypes such as colour and behaviour are more robustly expressed in natural habitats (J. Boughman and C. Peichel, unpublished), broadening the types of phenotypic traits that can be productively studied. Third, recombination over the successive generations of hybridization breaks up the chromosomal blocks. Thus, even recently hybridizing natural populations are likely to harbour smaller blocks than those that are obtained after two generations of intercrossing, which is what is typically used for QTL mapping studies in sticklebacks.

Here, we focused on divergence in two traits, male nuptial colour and body shape, in benthic and limnetic species from Enos Lake. This is one of seven lakes in British Columbia that has contained two species of sticklebacks: a bottom dwelling, deep-bodied benthic species that feeds on macroinvertebrates and an open water, streamlined limnetic species that feeds on plankton (McPhail 1984, 1992, 1994; Gow *et al.* 2008). Divergence in male nuptial colour and body shape are both adaptations to the different environments inhabited by limnetic and benthic sticklebacks and contributes to reproductive isolation in the species pairs (Schluter & McPhail 1992; Boughman 2001; Boughman *et al.* 2005; Head & Boughman unpublished). Differences in body shape facilitate specialized feeding and are heritable (McPhail 1984, 1992, 1994; Schluter & McPhail 1992). Differences in male nuptial colour are conspicuous in the distinct photic environments found within Enos Lake. Limnetic males mate in shallow areas where the water is clear, and these males display breeding coloration consisting of a red throat, bright blue eyes and a blue or green back. By contrast, benthic males mate in deeper vegetated areas where the water is red-shifted, and they develop a dark black colour throughout the body (McPhail 1984; Boughman 2001). Divergence in male nuptial colour contributes to reproductive isolation between the

benthics and limnetics in Enos Lake, suggesting that it is adaptive and has contributed to speciation in these fish (Boughman 2001; Boughman *et al.* 2005). These differences in male nuptial colour have a genetic basis (Lewandowski & Boughman 2008), but males do not express these colours as robustly in the laboratory as in the wild, making nuptial colour a particularly appealing trait for admixture mapping in sticklebacks.

The unique evolutionary history of the Enos Lake species pair has also created an ideal population for admixture mapping. Although the benthic and limnetic species in Enos Lake were once reproductively isolated (McPhail 1984; Ridgway & McPhail 1984), both morphological and genetic data demonstrate that these populations have been interbreeding for at least ten generations (Kraak *et al.* 2001; Gow *et al.* 2006; Taylor *et al.* 2006). Fish sampled prior to 1994 can be classified both morphologically and genetically as benthics or limnetics, but by 2002, fish are neither morphologically nor genetically classifiable to species, suggesting that interbreeding between the two Enos Lake species has created one admixed population (Taylor *et al.* 2006). Ecological disturbance is thought to have contributed to the increased hybridization, both by reducing sexual isolation (Lackey & Boughman unpublished) and by ecologically dependent postmating isolation (Behm *et al.* 2010). This reverse speciation event in Enos Lake has made it difficult to obtain individuals from the pure species for standard QTL mapping. However, recent hybridization does make it possible to use admixture mapping to identify genomic regions associated with colour and shape differences, and to determine whether these two key components of reproductive isolation are linked to the same regions of the genome.

Materials and methods

Fish sampling and colour measurements

We collected a total of 508 males with minnow traps and dip nets throughout Enos Lake, Vancouver Island, British Columbia, during the breeding seasons of 2006 (June 1–4; $n = 227$ males) and 2007 (May 21–24; $n = 281$ males). We scored seven traits relating to male nuptial colour on live males in the field. Five traits (red throat area, red throat intensity, eye colour intensity, body darkness and body brightness) were numerically scored on a scale from 0 (no colour) to 5 (maximal colour), with 11 values possible (e.g. 0, 0.5, 1, 1.5; Fig. S1, Supporting information). The presence or absence of erythrophores and melanophores was scored as two independent traits.

The left side of each individual was photographed under a canopy with a digital camera (Canon G7). To account for fluctuating natural light conditions, we used a constant neutral grey background in the photographs. This allowed for levelling of each digital image using Adobe Photoshop 7.0.1. After each image was taken, we clipped half of the caudal fin and stored it in 95% ethanol. Each male was released into the lake after the tissue sample was obtained.

Once the images were processed, two independent people examined all photographs from the June 2006 and May 2007 collections to select individuals with the most extreme nuptial coloration for bulk segregant analysis. Males with a red throat colour and no visible melanophores on the throat were placed in the extreme red group, males that were black throughout the entire body with no visible erythrophores on the throat were placed in the extreme black group, and males that displayed both red throats and black coloration throughout their bodies were placed in the intermediate group (Fig. 1). For each of these categories, each person independently selected fewer than 5% of the total individuals (person 1: $n = 14$ red males; $n = 21$ black males; $n = 19$ intermediate males; person 2: $n = 18$ red males; $n = 16$ black males; $n = 18$ intermediate males). A majority of these males were independently chosen by both parties ($n = 10$ red males; $n = 16$ black males; $n = 15$ intermediate males). To create DNA pools with an equal number of red and black males ($n = 17$ males in red pool; $n = 17$ males in black pool), seven additional red males and one additional black male were chosen from the combined list of males originally chosen by each person. Spearman's correlation coefficients were used to test whether there were significant correlations between each of the five numerically scored colour traits (red throat area, red throat intensity, eye colour intensity, body darkness and body brightness) and colour category (red or black) in these 34 males.

Microsatellite genotyping

Genomic DNA was isolated from fin clips of all samples by phenol–chloroform extraction, followed by ethanol precipitation. For bulk segregant analysis (Michelmore *et al.* 1991), two DNA pools containing equal concentrations (10 ng/μL) of each individual DNA were prepared: the red pool contained DNA from the 17 reddest males and the black pool contained DNA from the 17 blackest males (Fig. 1). These two DNA pools were each genotyped with 576 microsatellite markers distributed throughout the stickleback genome on an ABI 3100 (Applied Biosystems) as previously described (Peichel *et al.* 2001). Peak sizes for each marker were viewed with GENEMAPPER 3.7 software (Applied Biosystems). Only seven microsatellites on three linkage groups (*Stn5* on LG1, *Stn43*, *Stn46* and *Stn292* on LG4, and *Stn71*, *Stn74* and *Stn321* on LG7; Table S1, Support-

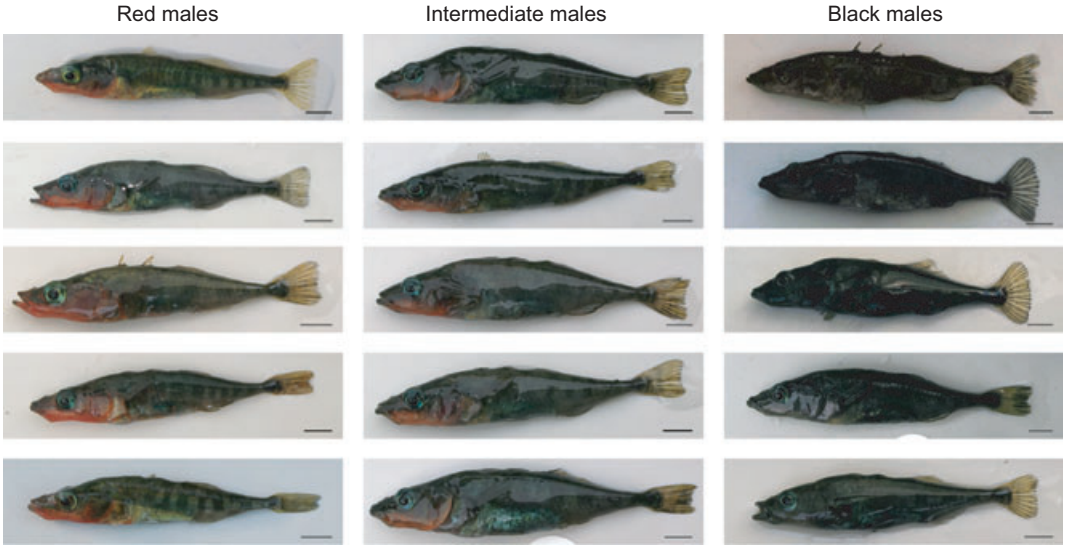

Fig. 1 Individuals representing the diversity of colour phenotypes present within the recently formed hybrid population of Enos Lake. Representative males shown are a sample of those used in each colour pool: red phenotype (red with no presence of melanophores on the throat), intermediate phenotype (red throats and black coloration throughout the body) and black phenotype (black with no presence of erythrophores on the throat). Scale bars = 0.5 cm.

ing information) displayed greater than a threefold difference in the fluorescence intensity of an allele between the red and black pools. Differences in allele frequency at these seven markers were first verified by genotyping the 34 individuals used in the red and black pools. Then, these seven markers were used to genotype all 508 individuals to analyse associations between genotype and phenotype.

Genetic ancestry analysis

To test for the possibility that individuals in the extreme colour pools simply represent genetically pure benthics or limnetics, individuals from the black (n = 17), red (n = 17) and intermediate pools (n = 13) were genotyped with five unlinked microsatellite markers (*Stn386* on LG2, *Stn43* on LG4, *Stn254* on LG12, *Stn295* on LG19 and *Stn216* on LG20; Table S1, Supporting information) that have previously been used to classify individuals from British Columbia species-pair lakes as benthic, limnetic or hybrid (Gow *et al.* 2006, 2008). STRUCTURE 2.2 (Pritchard *et al.* 2000; Falush *et al.* 2003, 2007) was then used to cluster individuals based on likelihood of ancestry. Analysis was performed under the admixture and correlated allele frequency models with a 'burn-in' period of 50 000 iterations followed by 200 000 Markov chain Monte Carlo (MCMC) repetitions. To determine the most likely number of genetic clusters (K), five runs of each probable K value (one through five) were carried out with the same number of 'burn-in' iterations and MCMC repetitions as earlier. The optimal value of K = 2 was determined both by the estimated natural logarithm probability of data (L(K)) and by the *ad hoc* statistic ΔK (Evanno *et al.* 2005).

Analysis of association between colour and genotype

Nonparametric Kruskal–Wallis tests were used to test for associations between colour phenotypes and genotypes in the sample of all 508 individuals for each of the seven markers first identified in the bulk segregant analysis. The false discovery rate (FDR; Benjamini *et al.* 2001) was used to adjust significance thresholds for multiple tests. In addition, we performed single marker linear models, as well as a joint model with all markers examined simultaneously (full model). Unless specified, all analyses are based on the full model incorporating all markers, size and year. These analyses were performed in R (V2.12.1) using custom scripts (available on DRYAD).

Shape analysis

Morphometric analysis was carried out on the digital images of the 508 wild-caught individuals. The 20 landmarks (Fig. 2a) were chosen based on a previous analysis of benthic and limnetic shape in Enos Lake (Taylor *et al.* 2006), and a QTL mapping study of shape differences between Paxton Lake benthic and marine stickleback (Albert *et al.* 2008). Landmark coordinates were created using tpsDig2 (Rohlf 2004) and used for relative warp analysis using tpsRelw (Rohlf 2005) as previously

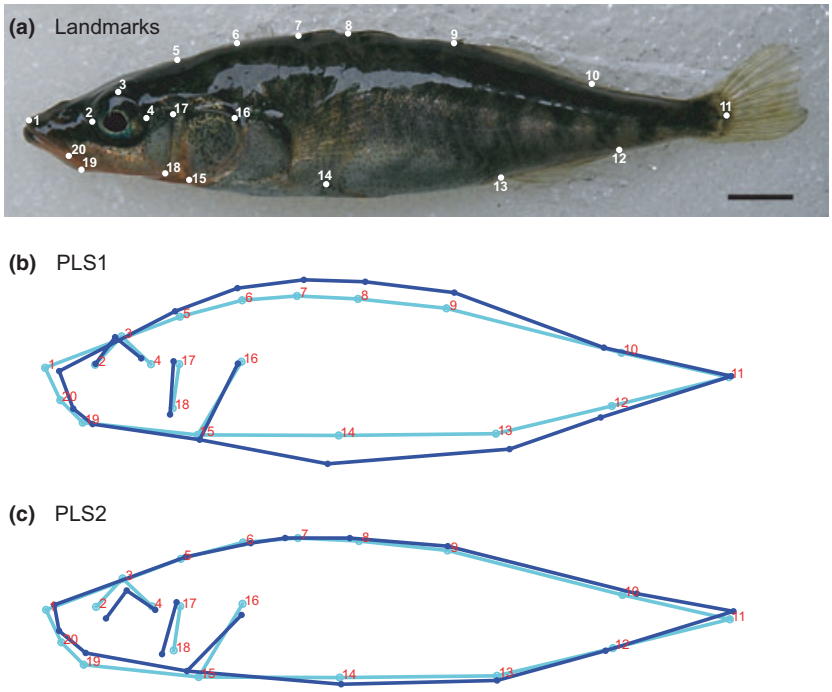

Fig. 2 Morphometric analysis of body shape, and its association with colour. (a) The 20 numbered landmarks are shown on a randomly chosen Enos Lake male: (1) anterior tip of upper lip; (2) anterior extent of eye; (3) dorsal extent of eye; (4) posterior extent of eye; (5) supraoccipital notch lateral to the dorsal midline; (6) point of intersection between the dorsal midline and the line starting in landmark 5; (7) anterior insertion of first dorsal spine; (8) anterior insertion of second dorsal spine; (9) anterior insertion of third dorsal spine; (10) posterior insertion of dorsal fin; (11) posterior extent of caudal peduncle; (12) posterior insertion of anal fin; (13) anterior insertion of anal fin; (14) insertion of pelvic spine into pelvic girdle; (15) ventral extent of operculum; (16) posteriodorsal extent of operculum; (17) dorsal extent of preopercular; (18) posterioventral extent of preopercular; (19) anterioventral extent of preopercular; (20) posterior tip of upper lip. The shape changes that maximally covary with colour in the sample of 508 individuals were determined using partial least squares (PLS) analysis. (b) PLS1 accounts for 79.99% of the covariation between shape and colour and shows covariation of overall body depth with red throat area and intensity, and to a lesser degree with body brightness and darkness (Fig. S3 and Table S3, Supporting information). Deeper-bodied fish (dark blue outline) are darker and have less intense red throats. (c) PLS2 accounts for 17.35% of the covariation between shape and colour and largely shows covariation between head shape and the brightness–darkness axis (Fig. S3 and Table S3, Supporting information). Fish with smaller heads (dark blue outline) are brighter. In (b) and (c), the light blue outline represents the population mean.

described (Taylor *et al*. 2006). Given that different scaling factors were used for the camera between the two sampling years, we adjusted by centring and scaling by the medians for the samples. This allows for a suitable correction for scaling; however, any overall differences in size between years will be confounded in the analysis.

Covariation between shape and colour

To assess covariation between shape and colour in the entire sample of 508 males, two-block partial least squares (PLS) analysis (Klingenberg & Zaklan 2000; Rohlf & Corti 2000) was used. This method performs a singular value decomposition on the matrix of covariances between two blocks of variables, in this study represented by colour and shape, respectively. This can be represented as follows:

$$\begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix}$$

where $\mathbf{S}_{11}$ and $\mathbf{S}_{22}$ represent the variance–covariance matrices for shape (Procrustes residuals) and colour, respectively. $\mathbf{S}_{12}$, the matrix of interest, contains the covariances between shape and colour variables. Singular value decomposition is used on $\mathbf{S}_{12}$ to extract vectors representing linear combinations of the original variables that maximally covary (between shape and colour). The first singular value represents the maximal degree of covariation between shape and colour, with each additional set of new variables representing statistically independent sets of shape and colour variables that covary (each successively smaller than the last). For these analyses, it was first confirmed that the residuals of the numerically scored colour traits conformed to a normal distribution. The analysis and visualizations

were generated using MORPHOJ (Klingenberg 2011), with 1000 permuted data sets to generate an empirical sampling distribution under the null model of no covariation. These results were confirmed using custom R scripts (available on DRYAD). In addition to using these vectors to assess and visualize the relationship between colour and shape, we additionally used them as covariates in the mapping analyses (below) to adjust for the potential pleiotropic effects of the markers on both shape and colour.

A scalar measure of covariation between shape and colour, the RV coefficient, was used (Escoufier 1973; Robert & Escoufier 1976; Klingenberg 2009):

$$\mathrm{RV} = \frac{\mathrm{trace}\,(\mathbf{S}_{12}\mathbf{S}_{21})}{\sqrt{\mathrm{trace}\,(\mathbf{S}_{11}\mathbf{S}_{11})\,\mathrm{trace}\,(\mathbf{S}_{22}\mathbf{S}_{22})}}$$

where the numerator can be interpreted as the total amount of covariation between the shape and colour (Rohlf & Corti 2000), and the denominator as the total amount of variation in the two sets of variables, thus scaling the numerator (Klingenberg 2009). For the bivariate case, this is equivalent to the Pearson correlation coefficient. The RV coefficient was estimated in MORPHOJ (Klingenberg 2011) and confirmed with custom R scripts (available on DRYAD).

Multivariate analysis of the association between shape and genotype

To test the association between genotype at each of the seven microsatellite markers identified in the bulk segregant analysis and shape phenotypes in the entire sample of 508 males, two approaches were utilized. First, single marker analysis was performed, in which just the marker of interest plus covariates of size and year of collection were each used in a separate MANOVA. For the shape data, both the Procrustes residuals for all landmarks and the partial warps and uniform components were examined (Rohlf & Bookstein 2003). These two methods were used to confirm that the results using the Procrustes residuals were stable given the need to use a generalized inverse for the calculations and to provide comparison. Because both methods give similar results, we report just the results from the analysis of partial warps and uniform components. For statistical inference, individual marker level permutation tests (with 10 000 permutations), as well as experiment-wise error rates (Churchill & Doerge 1994), were performed using custom R scripts (available on DRYAD).

Second, a model was examined that included all seven markers, as well as size and year as covariates, to allow for adjustments due to any lack of independence among markers. To evaluate how colour and shape are jointly influenced by the same loci, the same models as described earlier (both single marker and all markers) were used, this time including the PLS scores for the colour variables (both PLS1 and PLS2).

Given that most of the microsatellite markers had more than two alleles, pairwise comparisons on genotypes within markers were made using canonical variates analysis in MORPHOJ. Both Procrustes and Mahalanobis distance for the observed value were examined and compared to an empirical distribution generated using 10 000 permutations under the null model of no influence of genotype. Percent variance explained (PVE) was calculated by estimating the full and partial coefficients of determination from the individual marker models, as well as from the complete model (with all markers, size and colour PLS scores) by comparing the ratio of total variance for the fitted values over the total variance for the original data, using modified functions from (Claude 2008). In these analyses, it is possible that PVE is overestimated because these markers were first identified in the extreme colour pools. However, these analyses were conducted using the phenotype and genotype data from the full sample of 508 males and should reflect the relative effect sizes of these loci in the population sampled.

Results

Distribution of phenotypes among Enos Lake males

The majority of Enos Lake males collected in 2006 and 2007 had intermediate nuptial coloration, with aspects of both red and black nuptial coloration appearing in various combinations (Fig. 1). There was substantial variation in all the traits scored across the 508 males (Fig. S1, Supporting information). Approximately 3% (n = 17) of the males were categorized as having an extreme red nuptial colour, and 3% (n = 17) of the males were categorized as having an extreme black nuptial colour (Fig. 1). Within these red and black categories, red throat area, red throat intensity, body darkness and body brightness were significantly correlated with colour category and with each other, but not with eye intensity (Table S2, Supporting information).

Identification of genomic regions associated with male colour differences

Of the 576 microsatellite markers used to genotype the two extreme colour DNA pools, PCR products were obtained for 498 markers (86%) and 439 markers (76%) were informative (i.e. polymorphic). Only seven markers showed differences in allele frequencies between the red and black pools. These markers are located on three

linkage groups: *Stn5* on LG1; *Stn43*, *Stn46* and *Stn292* on LG4; *Stn71*, *Stn321* and *Stn74* on LG7 (Table S1, Supporting information). When the individuals from the pools were genotyped with these seven markers, there were significant differences in the relative frequencies of the alleles present in red and the black individuals ($P < 0.0001$ for all markers; Fig. 3). A single allele was primarily associated with black colour for all markers, while several alleles were associated with red colour at some markers (Fig. 3). However, when comparing the mean number of alleles between the black and red pools, this difference did not quite reach significance ($t = 1.84$; $P = 0.08$).

The full sample of 508 individuals was then genotyped with the markers identified in the bulked segregant analysis. When the genotypes of all 508 individuals at these seven markers were analysed, there were significant associations between colour phenotypes (red throat area, red throat intensity, body brightness and body darkness) and genotype (Kruskal–Wallis test; $P < 0.0005$ for all seven markers; Table 1).

Assessing benthic–limnetic ancestry

Given the small number of males with extreme red or black nuptial colour, it was possible that the black males are the remaining 'pure' benthic forms and that the red males are the remaining 'pure' limnetic forms in Enos Lake. To assess genetic ancestry, the 47 individuals from the red, black and intermediate categories were genotyped using five unlinked markers that have previously been used to distinguish sympatric benthic

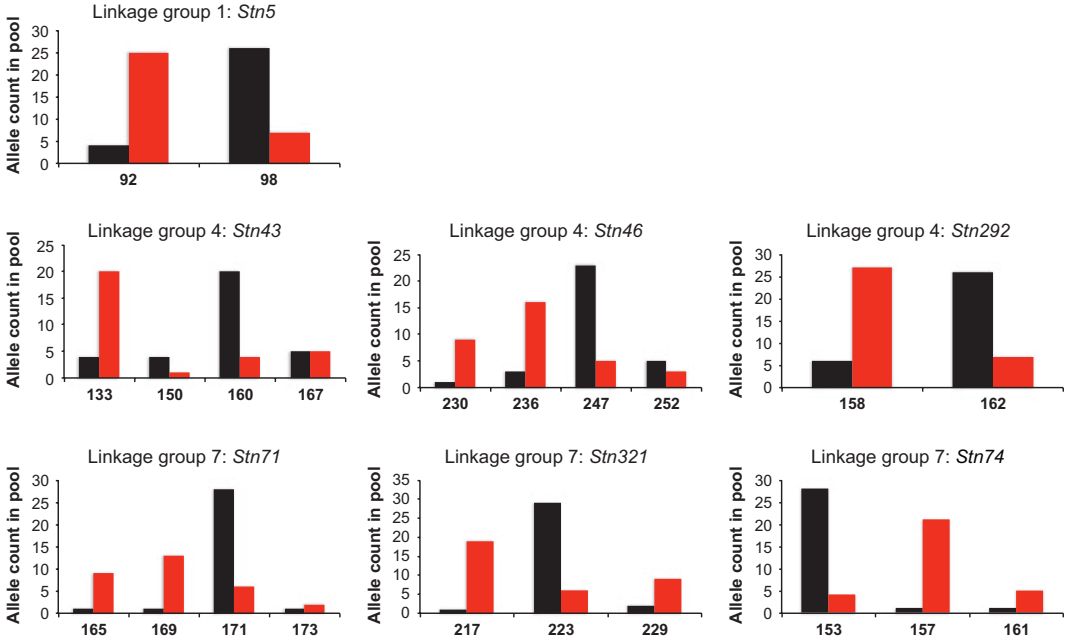

Fig. 3 Allele counts in males of the black pool (black bars) and males of the red pool (red bars) at seven markers (*Stn5* on LG1; *Stn43*, *Stn46*, *Stn292* on LG4; *Stn71*, *Stn321*, *Stn74* on LG7). The alleles present in males displaying extreme red colour differ significantly from the males displaying extreme black colour at all seven markers ($P < 0.0001$), using either a Pearson chi-square test or Fisher's exact test with a continuity correction when only two alleles were present.

Table 1 Associations between male colour phenotypes and genotypes in all 508 individuals sampled. *P*-values are shown and were calculated using the Kruskal–Wallis test. All associations are statistically significant using a false discovery rate of $\alpha = 0.05$

Marker	LG	Red throat area	Red throat intensity	Body brightness	Body darkness
Stn5	1	8.28E-08	9.11E-15	9.65E-06	1.39E-07
Stn43	4	2.88E-04	5.59E-10	1.63E-15	2.75E-13
Stn46	4	4.31E-06	9.56E-12	1.42E-15	1.78E-13
Stn292	4	9.88E-04	9.03E-07	1.35E-18	4.02E-19
Stn71	7	3.08E-16	2.21E-26	4.43E-08	1.09E-05
Stn321	7	8.20E-10	8.08E-25	1.16E-11	1.50E-10
Stn74	7	9.64E-11	1.44E-26	2.91E-09	9.43E-09

and limnetic forms in British Columbia (Gow *et al.* 2006, 2008; Table S1, Supporting information). When these data were analysed using STRUCTURE, the most likely number of population clusters was two (Fig. 4a; see Methods). Red males had a higher average probability (85 ± 0.04%) of belonging to one cluster, which likely represents limnetic ancestry. Conversely, black males had a higher average probability (71 ± 0.07%) of belonging to the other cluster, which likely represents benthic ancestry. Males that had an intermediate phenotype also had a higher average probability of belonging to the same cluster as those of the black phenotype (62 ± 0.09%).

To investigate whether any one of the five markers had a strong influence on these results, the STRUCTURE analysis was repeated, excluding each of the five markers from the analysis one at a time. When *Stn43* was excluded, the four remaining markers showed complete admixture; individuals had an average probability of 50 ± 1% (range, 45–55%) of belonging to either of the two genetic clusters (Fig. 4b). *Stn43* was one of the seven markers found to have a significant difference in allele frequency between the red and black pools in the bulk segregant analysis. Thus, with the exception of the marker linked to colour, the males used for bulk segregant analysis appear to be genetically admixed.

Associations between body shape and male colour

Benthics and limnetics differ dramatically in their overall body shape: limnetics have streamlined bodies, while benthics have deeper bodies (Schluter & McPhail 1992). To determine whether shape is segregating independently of colour in the population, an analysis of body shape was performed by placing 20 landmarks on digital images of the complete set of individuals (n = 508; Fig. 2a). When the first and second relative warps of shape are plotted, a single cluster is observed, consistent with previous studies (Taylor *et al.* 2006; Behm *et al.* 2010). However, there is some evidence that males from the black and red pools show differences in shape (Fig. S2, Supporting information). Consistent with this analysis, using two-block partial least squares (PLS) analysis, significant covariation between colour and shape was observed in the complete set of 508 individuals (Fig. 2). The overall magnitude of covariation between shape and colour variables was 0.11, as measured using the RV coefficient, consistent with a modest degree of phenotypic covariation. The first PLS accounted for 79.99% of the explained covariance and is most strongly associated with a vector dominated by the influence of red throat intensity and area, and with decreasing darkness (Fig. S3 and Table S3, Supporting information). Fish with more negative values of PLS1 (i.e. darker) tend to be deeper-bodied than fish with positive values of PLS1 (i.e. greater red throat area and intensity) (Fig. 2b; Fig S3, Supporting information). The second PLS accounts for 17.35% of this covariance and is most strongly associated with variation in head shape, body darkness and body brightness. Fish with more positive values of PLS2 (i.e. brighter) tend to have smaller heads (Fig. 2c; Fig. S3 and Table S3, Supporting information). Interestingly, variation in eye intensity does not load substantially on either PLS.

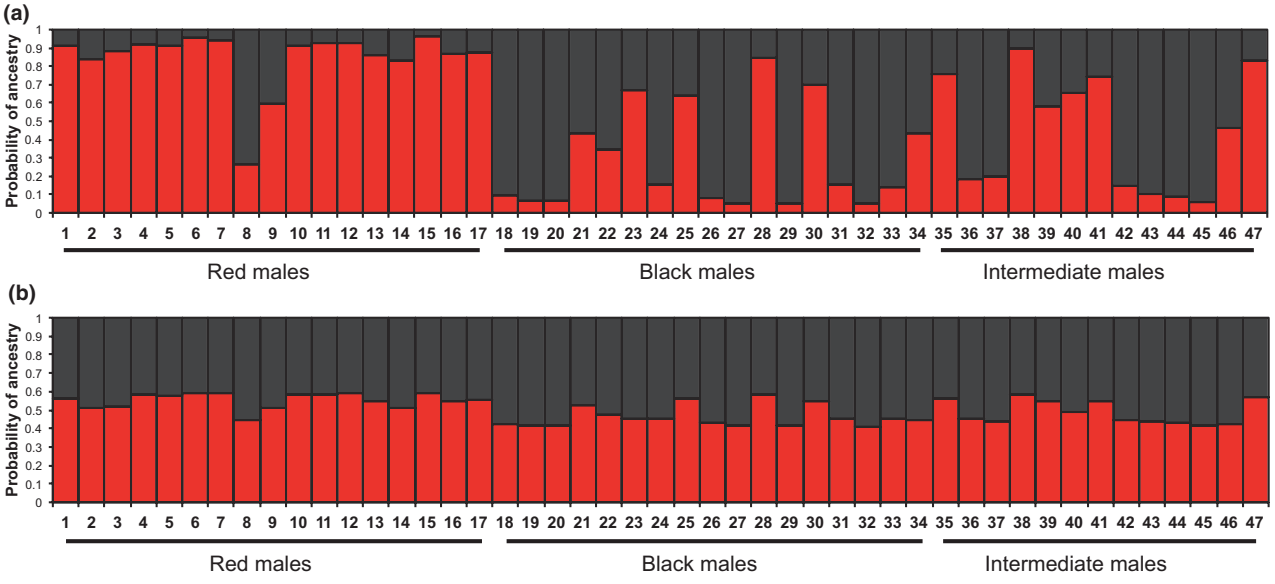

Fig. 4 STRUCTURE analysis with (a) all five markers included in the analysis and (b) *Stn43* omitted from the analysis. Each column represents a male used in the pool indicated below the bracket. Bars represent the probability that the individual belongs to cluster 1 (red) and cluster 2 (dark grey).

Linkage between genetic markers, male colour and body shape

Given the covariation between shape and colour in these males, we wanted to examine whether the associations between genotype and colour phenotype were simply the result of linkage between these markers and shape phenotypes. Thus, for both individual colour traits and the linear combinations of colour traits that explain the greatest covariation with shape (PLS1 and PLS2), we performed single marker analyses, in addition to a model including all seven genetic markers, size and year as explanatory covariates (Table 2). Including all markers in the analysis, the genotypic effects accounted for between 21.6% and 36.5% of the variation in the individual colour traits, 45.0% of the variation in PLS1 and 11.0% of the variation in PLS2. The addition of the other covariates only adds a few additional percent variation (with the exception of body brightness where size accounts for an additional ~8% of the phenotypic variance). Thus, there is still a significant overall association between colour and the genetic markers, even when covariation with shape is taken into account.

To examine the association of the markers with shape, we took an explicitly multivariate approach utilizing both single marker analyses and a model simultaneously accounting for all markers, in addition to body size and year of collection (Table 3). In general, these results were largely consistent across analyses, with all markers demonstrating a significant effect on shape, both for the single and joint marker analyses. From the single marker analyses, markers generally accounted for between 4% and 6.6% of the variation in shape. In the analysis of all markers simultaneously, the model accounted for 19.3% of the variation in shape, and the partial coefficients of determination accounted for between 1.3% and 4.4% of the variation for the individ-

Table 2 Effects of genotype on male colour phenotypes in all 508 individuals sampled. All seven markers, year and size were included simultaneously as predictors for each colour trait. The *P*-value associated with the effect of a given predictor, adjusted for all others (including all other markers), is given for each colour trait. PLS1 and PLS2 represent the linear combinations of colour traits that explain the greatest covariation with shape. The percent variance explained (PVE) represents the influence of all markers considered together (genotype) or all markers plus size and year of collection (full model). See Methods for full description of models.

Trait	LG1 *Stn5*	LG4 *Stn43*	LG4 *Stn46*	LG4 *Stn292*	LG7 *Stn71*	LG7 *Stn321*	LG7 *Stn74*	Year	Size	PVE genotype (%)	PVE full model (%)
Area	0.0479	0.279	0.327	0.167	3.60E-05	0.0198	1.36E-10	2.68E-04	0.0395	23.8	27.0
Intensity	9.34E-04	0.779	0.0157	0.0446	3.58E-05	6.15E-06	2.2E-16	6.87E-04	0.584	36.5	38.0
Brightness	0.531	0.328	5.03E-05	6.16E-04	0.0105	1.18E-03	4.66E-09	9.83E-07	1.79E-04	21.6	29.0
Darkness	0.0598	0.724	3.99E-04	4.29E-03	0.0460	1.19E-04	5.84E-09	0.262	5.36E-06	21.6	25.6
PLS1	1.46E-04	0.0397	1.62E-04	0.607	1.14E-06	2.43E-09	2.20E-16	0.0127	0.0961	45.0	45.6
PLS2	0.660	0.630	0.0996	2.41E-04	4.31E-03	0.376	8.69E-03	2.72E-06	2.11E-04	11.0	18.6

Table 3 Associations between shape phenotype and genotype in all 508 individuals sampled. For single marker analysis, the multivariate linear model was fit with the focal marker, including size and year as covariates. Wilks lambda and the *F* approximation were used to evaluate the probability of observing the data under the null model of no influence of the marker. Marker by marker permutation tests were performed (1000 permutations computed experiment-wise, not by marker), with $P < 0.001$ for each marker. The percent variance explained (PVE) is given for each marker without accounting for any other markers. We also examined all markers simultaneously (including year and size), evaluated the significance of each marker via permutation and present the partial PVE for each marker when accounting for other markers, year and size.

		Single marker analysis				All markers simultaneously	
Marker	LG	Wilks λ	*F*	*P*	PVE (%)	*P*	Partial PVE (%)
Stn5	1	0.453	2.213	1.59E-16	4.25	<0.001	2.10
Stn43	4	0.207	1.412	5.03E-09	6.56	0.003	3.60
Stn46	4	0.317	1.413	4.04E-07	5.33	<0.001	2.90
Stn292	4	0.482	2.027	3.61E-13	4.02	<0.001	1.30
Stn71	7	0.314	1.423	2.28E-07	4.74	<0.001	2.20
Stn321	7	0.352	1.412	1.32E-06	4.64	0.02	2.80
Stn74	7	0.205	1.425	1.93E-09	6.14	<0.001	4.40

ual markers. Including body size and year of collection into these models accounts for only an additional 4.5% of the variation. We also examined an additional model that included (in addition to markers, year and size) the colour PLS scores from the partial least squares analysis, to account for covariation between shape, colour and the markers. While these colour scores are clearly associated with shape (Fig. 2; Fig. S3 and Table S3, Supporting information), the inclusion of these colour vectors did not substantially change any results (with all markers remaining significant). This model only accounted for 26.1% of the overall variation as compared to 23.8% without the colour PLS scores, suggesting that colour is having relatively small influences on the association between the shape and the markers. Thus, we are able to independently map colour and shape to the same genomic regions.

Discussion

Selective DNA pooling strategy

In this study, we took advantage of a unique hybrid population and used an admixture mapping approach to identify regions of the genome that contribute to variation in male nuptial colour in the threespine stickleback. Using bulk segregant analysis (Michelmore *et al.* 1991), *that is*, pooling the individuals with the most extreme colour phenotypes, we were able to rapidly screen hundreds of genome-wide microsatellite markers to identify genomic regions that are associated with both male colour and body shape on three chromosomes. Although we may have lost some power to detect loci of small effect through this selective DNA pooling strategy, our preliminary screen was focused on detecting loci with relatively major effects on male nuptial coloration. Selective genotyping of pools of phenotypic extremes from laboratory crosses had been shown to result in only a slight loss of power compared to complete genotyping for detecting QTL of relatively large effect (Darvasi & Soller 1994). This selective DNA pooling strategy may also be useful for other admixture mapping studies in natural hybrid populations, as it greatly reduces the costs and time associated with genotyping a large number of individuals, even when using next-generation genotyping technologies (Davey *et al.* 2011).

Assessment of population structure in Enos Lake

A major concern in admixture mapping is that population structure can cause unlinked markers to appear linked, resulting in spurious genotype–phenotype associations (Pritchard & Rosenberg 1999; Pfaff *et al.* 2001). One way to determine whether population stratification exists in an admixed population is to assess the levels of background LD by genotyping with additional unlinked markers (Pritchard & Rosenberg 1999; Falush *et al.* 2003; Lexer *et al.* 2007; Gompert & Buerkle 2009). Microsatellite markers are ideal because they provide more power to detect any population structure compared to biallelic markers and, as a result, fewer are needed (Pritchard & Rosenberg 1999). For human populations, simulation studies suggested that genotyping with at least 15–20 unlinked microsatellites would be sufficient to detect any underlying population stratification (Pritchard & Rosenberg 1999). However, using only five markers is sufficient to detect population stratification if diagnostic microsatellite markers for the ancestral populations are available (Shriver *et al.* 1997; Pritchard & Rosenberg 1999).

Previous work had identified a set of highly discriminatory markers for the benthic–limnetic species pairs in Enos, Paxton, Priest and Little Quarry lakes (Gow *et al.* 2006, 2008). To assess the genetic structure of the Enos Lake sample, we genotyped the extreme individuals used in the pools with five of these diagnostic markers that are located on five independent linkage groups. This analysis revealed a genetically admixed population when we used four of five diagnostic markers (Fig. 4b). Each individual has an approximately equal probability of belonging to one of two ancestral groups, suggesting that there is not major population stratification, even within these individuals chosen for their extreme colour phenotypes. Although we did not perform this analysis on all 508 males in our sample, we thought it likely that the highest levels of background LD would occur in individuals with the most ancestral-like phenotypes (i.e. pure red or pure black nuptial colour). Thus, our STRUCTURE analysis with these four markers suggests that Enos Lake is an admixed population, with little population structure found in the individuals chosen for selective genotyping.

Interestingly, when a fifth diagnostic marker, *Stn43*, was included in the analysis, some population structure was apparent, with the red and black individuals having a higher probability of belonging to different populations (Fig. 4a). This same marker, *Stn43*, was also isolated in our bulk segregant analysis and shows a significant association with both male nuptial colour and body shape. Because these diagnostic markers were identified in what was essentially a population genomics approach to identify 'outlier' loci between populations (Luikart *et al.* 2003; Schlotterer 2003; Storz 2005; Stinchcombe & Hoekstra 2008; Nosil *et al.* 2009), our study demonstrates that these benthic–limnetic diagnostic markers are likely linked to traits under divergent selection, with *Stn43* particularly linked to colour and

shape traits. Our results also suggest that combining a population genomics approach with pools of individuals chosen to represent the extremes of phenotypic differentiation between the populations has a high probability of identifying loci linked to adaptive traits in natural populations.

Linkage of male colour and body shape

From our genetic analysis, we conclude that the blackest and reddest fish do not simply represent the remaining 'pure' benthics and 'pure' limnetics in Enos Lake. However, there remains substantial covariation between body shape and male colour across all 508 fish collected (Fig. 2; Figs S2 and S3, Supporting information). Furthermore, all the markers that are significantly associated with male nuptial colour are also significantly associated with body shape, even after adjusting for the patterns of phenotypic covariation (Tables 2 and 3). All of the genomic regions detected in our study were identified in previous QTL studies of morphological divergence between Paxton Lake benthic and Japanese Pacific marine sticklebacks, including divergence in body shape and skeletal traits (Colosimo *et al.* 2004; Shapiro *et al.* 2004; Albert *et al.* 2008; Table S1, Supporting information), for example, pelvic reduction and three-shape QTL map to *Stn5* on LG1, pelvic reduction, plate number and shape QTL map to the same region of LG4 as we have identified in this study, plate number QTL map to *Stn71* on LG7 and eleven-shape QTL map to *Stn321* on LG7 (Colosimo *et al.* 2004; Shapiro *et al.* 2004; Albert *et al.* 2008). These genomic regions were also identified as outliers in a recent population genomic study of three other benthic–limnetic species pairs (Jones *et al.* 2012). The concordant locations of these mapped traits in stickleback suggest that, at least in this system, there may be relatively few genomic regions that make a large contribution to adaptation.

The seven markers that we found to be associated with both male nuptial colour and body shape are found on three linkage groups, with one marker on LG1, three markers on LG4 separated by 9 megabases (Mb) and three markers on LG7 separated by 14 Mb (Table S1, Supporting information). As we have not yet genotyped additional markers on these linkage groups, we do not know the physical size of the genomic regions that we have detected. It is possible that each of the three loci that we have detected on LG4 or LG7 represents independent loci that each makes a separate contribution to variation in male nuptial colour or body shape. It is also possible that there is only a single locus that contributes to phenotypic variation and that these three loci might have been detected simply because there is a large block of linkage disequilibrium (LD) on each of these linkage groups. Much additional work is needed to establish the extent of LD in these regions, to discern whether the loci we have identified on LG4 and LG7 represent three distinct genomic regions or one large genomic interval and to ultimately narrow down these genomic regions to identify the actual genes that underlie divergence in male nuptial colour and body shape.

The blocks of LD we found might exist because they have not yet been disrupted by recombination in this recently admixed population. Alternatively, these blocks of LD might remain because they harbour multiple traits (such as colour and shape) that were important for adaptation and/or reproductive isolation between the benthic and limnetic species in Enos Lake, and are therefore maintained by selective or genetic mechanisms. Clustering of adaptive traits in few genomic regions may be the result of multidimensional adaptation. Organisms that colonize novel environments are likely to be subject to divergent selection from multiple sources acting on multiple traits. In our system, colonization of freshwater lakes by the ancestral oceanic stickleback would have exposed the fish to different prey and predator communities and different water chemistry, potentially selecting on many morphological, behavioural and physiological traits. We know that many traits differ for benthic and limnetic sticklebacks (Schluter & McPhail 1992). Correlational selection on suites of adaptive traits could cause clustering of QTL by favouring particular combinations of traits (i.e. red colour with streamlined shape), thus favouring physical proximity in the genome (Butlin 2005; McKinnon & Pierrotti 2010). Although such tight linkage could be facilitated by reduced recombination in inversions (Noor *et al.* 2001; Rieseberg 2001; Butlin 2005; Kirkpatrick & Barton 2006; Hoffmann & Rieseberg 2008), we do not yet have any evidence for inversions between benthics and limnetics in these regions of the genome. These regions may also represent 'genomic islands of divergence' (Nosil *et al.* 2009) whereby differentiation in regions under divergent selection accumulates (including in regions that cause reproductive isolation), spreading to regions linked to these QTL via hitchhiking (Barton 2000). Finally, it is possible that detecting linkage to the same genomic regions could arise because the same genes control multiple traits (i.e. pleiotropy). Additional work will allow us to disentangle the selective and genetic mechanisms that underlie multidimensional adaptation in this system.

Regardless of the underlying mechanism, our data are consistent with previous empirical work, demonstrating that genetic linkage of traits underlying ecological specialization and reproductive isolation might

facilitate ecological speciation (Hawthorne & Via 2001; Kronforst *et al.* 2006; Lowry & Willis 2010). Divergence in body shape is important for adaptation to benthic or limnetic habitats because of effects on foraging and locomotion, and body shape is an important component of mate choice in these species (Webb 1982, 1984; Schluter & McPhail 1992; Walker 1997; Head & Boughman *in revision*). Similarly, divergence in male colour is an adaptation to the different signalling environments of benthic and limnetic mating habitat and has previously been shown to contribute to reproductive isolation between the species pairs (Boughman 2001; Boughman *et al.* 2005). Thus, it is possible that genetic linkage between these traits might have facilitated the process of ecological speciation in Enos Lake.

Prospects for admixture mapping in sticklebacks and other organisms

There has been increasing interest in utilizing admixture mapping in natural populations to identify the genetic basis of adaptation and reproductive isolation (Buerkle & Lexer 2008). The current study has provided an example of using admixture mapping to identify genomic regions associated with divergence in specific traits that contribute to adaptation and speciation in sticklebacks. Because admixture mapping relies upon recent hybridization events, the power to more finely map these large genomic regions is limited. Thus, further work is necessary to determine which genes underlie divergence in stickleback male nuptial coloration and body shape. However, this study demonstrates the utility of this approach for the initial identification of genomic regions that contribute to adaptive divergence between natural populations. Because several morphological and behavioural traits differ between benthic and limnetic sticklebacks, we can now use the admixed Enos population to identify the genomic regions that contribute to other divergent phenotypes. Although laboratory crosses can be made in sticklebacks and provide a complement to admixture mapping in hybrid populations, there are many interesting systems in which laboratory crosses are simply not feasible. Therefore, pursuing a similar strategy to the one used in our study may be the only option for exploring the genetic basis of adaptation in many systems. Admixture mapping can also be used to identify regions of the genome under selection or that contribute to reproductive isolation (Gompert & Buerkle 2009, 2010). Due to the advent of next-generation sequencing, identifying molecular markers for any species is now feasible (Davey *et al.* 2011), suggesting that admixture mapping will be a powerful approach to reveal the genetics of adaptation and speciation in natural populations.

Acknowledgements

We are grateful to Megan Head for assistance in the field and to Alex Buerkle, Jen Gow, Bret Payseur and Sean Rogers for comments on a previous version of the manuscript. This study was funded by a Career Award in the Biomedical Sciences from the Burroughs Wellcome Fund (C.L.P.), National Institutes of Health grants T32 GM07270 (T.B.M.), P50 HG002568 (C.L.P.), R01 GM094424 (I.D.) and National Science Foundation grants IBN0416808 (J.W.B.), CAREER award DEB0952659 (J.W.B.) and MCB0922344 (I.D).

References

Albert AY, Sawaya S, Vines TH *et al.* (2008) The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. *Evolution*, **62**, 76–85.

Barton N (2000) Genetic hitchhiking. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **355**, 1553–1562.

Behm JE, Ives AR, Boughman JW (2010) Breakdown in postmating isolation and the collapse of a species pair through hybridization. *American Naturalist*, **175**, 11–26.

Bell MA, Foster SA (1994) *The Evolutionary Biology of the Threespine Stickleback*. Oxford University Press, Oxford.

Benjamini Y, Drai D, Elmer G *et al.* (2001) Controlling the false discovery rate in behavior genetics research. *Behavioural Brain Research*, **125**, 279–284.

Boughman JW (2001) Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature*, **411**, 944–948.

Boughman JW, Rundle HD, Schluter D (2005) Parallel evolution of sexual isolation in sticklebacks. *Evolution*, **59**, 361–373.

Buerkle CA, Lexer C (2008) Admixture as the basis for genetic mapping. *Trends in Ecology and Evolution*, **23**, 686–694.

Butlin R (2005) Recombination and speciation. *Molecular Ecology*, **14**, 2621–2635.

Chan YF, Marks ME, Jones FC *et al.* (2010) Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science*, **327**, 302–305.

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Genetics*, **138**, 963–971.

Claude J (2008) *Morphometrics with R*. Springer, New York.

Colosimo PF, Peichel CL, Nereng K *et al.* (2004) The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biology*, **2**, 635–641.

Colosimo PF, Hosemann KE, Balabhadra S *et al.* (2005) Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science*, **307**, 1928–1933.

Cresko WA, Amores A, Wilson C *et al.* (2004) Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proceedings of the National Academy of Sciences, USA*, **101**, 6050–6055.

Darvasi A, Soller M (1994) Selective DNA pooling for determination of linkage between a molecular marker and a quantitative trait locus. *Genetics*, **138**, 1365–1373.

Davey JW, Hohenlohe PA, Etter PD *et al.* (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, **12**, 499–510.

Escoufier Y (1973) Treatment of vector variables. *Biometrics*, **29**, 751–760.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multi-locus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Falush D, Stephans M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.

Gompert Z, Buerkle CA (2009) A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. *Molecular Ecology*, **18**, 1207–1224.

Gompert Z, Buerkle CA (2010) Bayesian estimation of genomic clines. *Molecular Ecology*, **20**, 2111–2127.

Gow JL, Peichel CL, Taylor EB (2006) Contrasting hybridization rates between sympatric three-spined sticklebacks highlight the fragility of reproductive barriers between evolutionarily young species. *Molecular Ecology*, **15**, 739–752.

Gow JL, Rogers SM, Jackson M, Schluter D (2008) Ecological predictions lead to the discovery of a benthic-limnetic species pair of threespine stickleback in Little Quarry Lake, British Columbia. *Canadian Journal of Zoology*, **86**, 564–571.

Greenwood AK, Jones FC, Chan YF *et al.* (2011) The genetic basis of divergent pigment patterns in threespine sticklebacks. *Heredity*, **107**, 155–166.

Hawthorne DJ, Via S (2001) Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature*, **412**, 904–907.

Hindorff LA, Junkins HA, Mehta JP *et al.* (2011) *A Catalogue of Published Genome-Wide Association Studies*. Available from <http://www.genome.gov/26525384> (accessed 5 April 2011).

Hoekstra HE, Coyne JA (2007) The locus of evolution: evo devo and the genetics of adaptation. *Evolution*, **61**, 995–1016.

Hoffmann AA, Rieseberg LH (2008) Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? *Annual Review of Ecology, Evolution, and Systematics*, **39**, 21–42.

Hohenlohe PA, Bassham S, Etter PD *et al.* (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics*, **6**, e1000862.

Jones FC, Chan YF, Schmutz J *et al.* (2012) A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Current Biology*, **22**, 83–90.

Kimmel CB, Ullmann B, Walker C *et al.* (2005) Evolution and development of facial bone morphology in threespine sticklebacks. *Proceedings of the National Academy of Sciences, USA*, **102**, 5791–5796.

Kingsley DM, Peichel CL (2007) The molecular genetics of evolutionary change in sticklebacks. In: *Biology of the Three-Spined Stickleback* (eds Östlund-Nilsson S, Mayer I and Huntingford F), pp. 41–81. CRC Press, Boca Raton, Florida.

Kingsley DM, Zhu B, Osoegawa K *et al.* (2004) New genomic tools for molecular studies of evolutionary change in threespine sticklebacks. *Behaviour*, **141**, 1331–1344.

Kirkpatrick M, Barton N (2006) Chromosome inversions, local adaptation and speciation. *Genetics*, **173**, 419–434.

Kitano J, Ross JA, Mori S *et al.* (2009) A role for a neo-sex chromosome in stickleback speciation. *Nature*, **461**, 1079–1083.

Klingenberg CP (2009) Morphometric integration and modularity in configurations of landmarks: tools for evaluating *a priori* hypotheses. *Evolution and Development*, **11**, 405–421.

Klingenberg CP (2011) MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources*, **11**, 353–357.

Klingenberg CP, Zaklan SD (2000) Morphological integration between development compartments in the *Drosophila* wing. *Evolution*, **54**, 1273–1285.

Kraak SBM, Mundwiler B, Hart PJB (2001) Increased number of hybrids between benthics and limnetic threespined sticklebacks in Enos Lake, Canada: the collapse of a species pair? *Journal of Fish Biology*, **58**, 1458–1464.

Kronforst MR, Young LG, Kapan DD *et al.* (2006) Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. *Proceedings of the National Academy of Sciences, USA*, **103**, 6575–6580.

Lewandowski E, Boughman J (2008) Effects of genetics and light environment on color expression in threespine sticklebacks. *Biological Journal of the Linnean Society*, **94**, 663–673.

Lexer C, Buerkle CA, Joseph JA, Heinze B, Fay MF (2007) Admixture in European *Populus* hybrid zones makes feasible the mapping of loci that contribute to reproductive isolation and trait differences. *Heredity*, **98**, 74–84.

Lexer C, Joseph JA, van Loo M *et al.* (2010) Genomic admixture analysis in European Populus spp. reveals unexpected patterns of reproductive isolation and mating. *Genetics*, **186**, 699–712.

Lowry DB, Willis JH (2010) A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biology*, **8**, e1000500.

Luikart G, England PR, Tallmon D *et al.* (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981–994.

McKinnon JS, Pierrotti MER (2010) Colour polymorphism and correlated characters: genetic mechanisms and evolution. *Molecular Ecology*, **19**, 5101–5125.

McPhail JD (1984) Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): Morphological and genetic evidence for a species pair in Enos Lake, British Columbia. *Canadian Journal of Zoology*, **62**, 1402–1408.

McPhail JD (1992) Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): evidence for a species pair in Paxton Lake, Texada Island, British Columbia. *Canadian Journal of Zoology*, **70**, 361–369.

McPhail JD (1994) Speciation and the evolution of reproductive isolation in the sticklebacks (*Gasterosteus*) of south-western British Columbia. In: *The Evolutionary Biology of the Threespine Stickleback* (eds Bell MA and Foster SA), pp. 399–437. Oxford University Press, Oxford.

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences, USA*, **88**, 9828–9832.

Miller CT, Beleza S, Pollen AA *et al.* (2007) *cis*-Regulatory changes in *Kit ligand* expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell*, **131**, 1179–1189.
Nadeau NJ, Jiggins CD (2010) A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations *Trends in Genetics*, **26**, 484–492.
Nielson R (2009) Adaptionism – 30 years after Gould and Lewontin. *Evoution*, **63**, 2487–2490.
Noor MAF, Grams KL, Bertucci LA, Reiland J (2001) Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences, USA*, **98**, 12084–12088.
Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, **18**, 375–402.
Orr HA (2005) The genetic theory of adaptation: a brief history. *Nature Reviews Genetics*, **6**, 119–127.
Peichel CL, Nereng KS, Ohgi KA *et al.* (2001) The genetic architecture of divergence between threespine stickleback species. *Nature*, **414**, 901–905.
Pfaff CL, Parra EJ, Bonilla C *et al.* (2001) Population structure in admixed populations: effect of admixture dynamics on the pattern of linkage disequilibrium. *American Journal of Human Genetics*, **68**, 198–207.
Pritchard JK, Rosenberg NA (1999) Use of unlinked genetic markers to detect population stratification in association studies. *American Journal of Human Genetics*, **65**, 220–228.
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
Ridgway MS, McPhail JD (1984) Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): mate choice and reproductive isolation in the Enos Lake species pair. *Canadian Journal of Zoology*, **62**, 1813–1818.
Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends in Ecology and Evolution*, **16**, 351–358.
Rieseberg LH, Buerkle CA (2002) Genetic mapping in hybrid zones. *American Naturalist*, **159**, S36–S50.
Robert P, Escoufier Y (1976) Unifying tool for linear multivariate statistical methods – Rv-coefficient. *Journal of the Royal Statistical Society Series C-Applied Statistics*, **25**, 257–265.
Rohlf FJ (2004) *tpsDig, Digitize Landmarks and Outlines, version 2.0*. Department of Ecology and Evolution, State University of New York at Stony Brook. Available from http://life.bio.sunysb.edu/morph
Rohlf FJ (2005) *tpsRelw, Relative Warps Analysis, version 1.42*. Department of Ecology and Evolution, State University of New York at Stony Brook. Available from http://life.bio.sunysb.edu/morph
Rohlf FJ, Bookstein FL (2003) Computing the uniform component of shape variation. *Systematic Biology*, **52**, 66–69.
Rohlf FJ, Corti M (2000) Use of two-block partial least-squares to study covariation in shape. *Systematic Biology*, **49**, 740–753.
Schlotterer C (2003) Hitchhiking mapping – functional genomics from the population genetics perspective. *Trends in Genetics*, **19**, 32–38.
Schluter D, McPhail JD (1992) Ecological character displacement and speciation in sticklebacks. *American Naturalist*, **140**, 85–108.
Shapiro MD, Marks ME, Peichel CL *et al.* (2004) Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature*, **428**, 717–723.
Shriver MD, Smith MW, Jin L *et al.* (1997) Ethnic-affiliation estimation by use of population-specific DNA markers. *American Journal of Human Genetics*, **60**, 957–964.
Slate J (2005) Quantitative trait locus mapping in natural populations: progress, caveats, future directions. *Molecular Ecology*, **14**, 363–379.
Stapley S, Reger J, Feulner PGD *et al.* (2010) Adaptation genomics: the next generation. *Trends in Ecology and Evolution*, **25**, 705–712.
Stern DL, Orgogozo V (2008) The loci of evolution: how predictable is genetic evolution? *Evolution*, **62**, 2155–2177.
Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, **100**, 158–170.
Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14**, 671–688.
Taylor EB, Boughman JW, Groenenboom M, Sniatynski M, Schluter D, Gow JL (2006) Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Molecular Ecology*, **15**, 343–355.
Walker JA (1997) Ecological morphology of lacustrine threespine stickleback *Gasterosteus aculeatus* L. (Gasterosteidae) body shape. *Biological Journal of the Linnean Society*, **61**, 3–50.
Webb PW (1982) Locomotor patterns in the evolution of actinopterygian fishes. *American Zoologist*, **22**, 293–302.
Webb PW (1984) Body form, locomotion and foraging in aquatic vertebrates. *American Zoologist*, **24**, 107–120.
Winkler CA, Nelson GW, Smith MW (2010) Admixture mapping comes of age. *Annual Review of Genomics and Human Genetics*, **11**, 65–89.

J.B. studies the selective causes and genetic basis of speciation, focusing on sexual and natural selection on colour, morphology, and behaviour in threespine sticklebacks. I.D. studies non-additive genetic effects of complex phenotypes. C.P. is interested in the genetics of behavior and sex chromosome evolution, again using stickleback fish. T.M. studies the genetics of adaptation. This work is part of her Ph.D. thesis.

Data accessibility

All genotypic and phenotypic data, as well as custom R scripts are deposited in the DRYAD database: doi: 10.5061/dryad.55j7t.

Supporting information

Additional supporting information may be found in the online version of this article:

Table S1 Genomic positions and quantitative trait locus linked to microsatellite markers used in this study.

Table S2 Correlations between individual colour traits in the 34 males from the red and black pools.

Table S3 Contribution (weights) of individual colour traits to covariances between shape and colour, using partial least squares.

Fig. S1 Distribution of colour phenotypes in 508 Enos Lake males collected in 2006 and 2007.

Fig. S2 Body shape of males from the red and black pools relative to all 508 males collected.

Fig. S3 Biplot of covariance among colour traits in relationship to shape.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.