www.sciencemag.org/cgi/content/full/1159978/DC1

Supporting Online Material for

Natural Selection on a Major Armor Gene in Threespine Stickleback

Rowan D. H. Barrett,* Sean M. Rogers, Dolph Schluter

*To whom correspondence should be addressed. E-mail: rbarrett@zoology.ubc.ca

Published 28 August 2008 on *Science* Express
DOI: 10.1126/science.1159978

This PDF file includes

MATERIALS AND METHODS

Collection of experimental fish

We collected stickleback in April and May of 2006 from Oyster Lagoon on the Sechelt peninsula in western British Columbia (49°36'48.6" N, 124°1'46.88" W). Oyster lagoon is a saltwater inlet with salinity ranging from 28-32 ppt, in which phenotypically partially armored fish occur at an approximate frequency of 0.01. This population breeds in saltwater and the rare sticklebacks with reduced plate number are marine in all other phenotypic characteristics (shape, size, color, spine length; *S1-3*). We sampled approximately 35,000 fish using unbaited minnow traps. We brought all fish not possessing the full number of lateral plates (n = 354) back to the lab, including those missing only one or two lateral plates, and released all other captured fish. In the lab we injected individual fish subepidermally with a fluorescent visible implant elastomer tag (Northwest Marine Technology) using a 29 gauge syringe.

A small fin clip taken from each fish allowed us to genotype the fish at diagnostic loci within the *Eda* gene that distinguish between low and complete alleles (*S4*). We also genotyped fish at a SNP within an α1 subunit of Na+-K+-ATPase for freshwater versus marine residency (*S5*) to confirm that they were marine residents rather than recent freshwater or stream migrants. Almost all alleles (96%) at this marker were consistent with those observed in marine populations.

Pond sampling

Of the 354 partial morphs collected from Oyster Lagoon, 182 were heterozygous at *Eda*. On June 1, 2006, we released the heterozygous fish into 4 artificial ponds located at the University of British Columbia, Vancouver, British Columbia, Canada (pond 1 N = 45, pond 2 N = 46, pond 3 N = 45, pond 4 N = 46). These ponds measure 23 m × 23 m and have a maximum depth of 3 m in the centre, as described (*S6*). Like many coastal lakes in British Columbia, the ponds are lined with sand and bordered with limestone. All ponds had been previously drained, cleaned and refilled in 2001, allowing plant and invertebrate communities to re-establish, but remaining free of fish until this experiment. The plants and invertebrates used to seed the ponds were collected from Paxton Lake, Texada Island, British Columbia, an 11-ha lake that contains wild sticklebacks. Apart from their construction, initialization, and use in prior experiments, the ponds are unmanipulated environments. In previous experiments these ponds have sustained large populations of sticklebacks over multiple generations, with life cycles and diets characteristic of their wild source populations (S*7*). Growth rates of fish in the ponds are similar to those of wild fish in freshwater lakes (*S8*). We observed F1 progeny in the ponds in August 2006. We were able to distinguish between the F1 fish and their parents because parents retained their elastomer tags and were significantly larger in size. The colonizing fish accounted for a very small proportion of the overall populations following the F1 generation, and we only observed three colonizers out of ~1500 individuals sampled over the full duration of the experiment. No colonizers were caught after the October 2006 sample. Sampling of the separate cohorts was not possible after July 2007 because F1 and F2 generations were no longer distinguishable, as their body size

distributions had merged. Fish were sampled with dip nets, traps, or a seine net. A variety of sampling methods were necessary during different times of the year because small juveniles cannot be caught in traps, the seine net disturbs nesting sites, and large adults cannot be easily caught with dip nets. We detected no difference in genotype frequencies using these different sampling methods as genotype ratios in samples from the same ponds, in the same month, with different methods were not significantly different (pond 1: $\chi^2 = 0.37$, df = 2, $P = 0.83$, pond 2: $\chi^2 = 1.73$, df = 2, $P = 0.42$, pond 3: $\chi^2 = 2.08$, df = 2, $P = 0.35$, pond 4: $\chi^2 = 0.04$, df = 2, $P = 0.98$). We recorded total length, phenotype, and breeding condition for all fish sampled, and then preserved each individual at -80^{o}C.

Genotyping

We isolated total genomic DNA from small caudal fin clips using a standard proteinase K phenol chloroform protocol (*S9*). We quantified DNA yield using spectrophotometry and then preserved DNA samples at -20°C. We used diagnostic in/del loci to identify low and complete *Eda* alleles (isolated from loci Stn380 and Stn381 within introns two and six of the *Eda* gene on linkage group 4, respectively (*S4*)). Four unlinked microsatellite loci, isolated and characterized by the Stanford Genome Research Centre, were selected for additional population genetic analyses (Linkage group and Genbank Accession numbers in parentheses): Stn224 (LG 11, BV678144), Stn314 (LG 8, BV678119), Stn387 (LG 2, BV678140), and Stn388 (LG 9, BV678141). Microsatellite and *Eda* alleles were amplified by PCR with a DNA Engine® Peltier Thermal Cycler (MJ research, Inc., Waltham, Massachusetts) in 10 µl reactions containing 5 to 15 ng of genomic DNA, 1uM of each forward and reverse primer, 1X PCR buffer, 0.25 mM of each dNTP, 1.5 mM MgCl2, and 0.25U of AmpliTaq Gold polymerase (Applied Biosystems, Foster city, California). Cycling conditions were standardized over all loci as follows: 93 ^{o}C for 3 min, 95 ^{o}C 30 s, 59 ^{o}C 30s, 72 ^{o}C 30 s, 5 cycles of 94 ^{o}C 30 s, 59 ^{o}C 30 s, 72 ^{o}C 30 s, 35 cycles of 90 ^{o}C 30 s, 60 ^{o}C 30 s, 72 ^{o}C 30 s, followed by 72 oC for 10 min and then cooled to 4 ^{o}C. Electrophoresis consisted of pooling PCR products with an internal size standard (LIZ 500bp, Applied Biosystems) and loading onto a 3730S Automated Sequencer (Applied Biosystems). Allelic sizes (in base pairs) were determined by reference to the internal sizing standard in the software GENEMAPPER (Applied Biosystems). We found no significant difference between genotype frequencies at Stn380 and Stn381 in the first four samples ($n = 698$, $\chi^2 = 0.275$, df = 2, $P = 0.872$) and therefore used only Stn381 for the remaining samples. All genotype data presented is for Stn381 only.

Calculating selection and dominance coefficients

We calculated *S* (the selection coefficient) for viability selection on the low and complete *Eda* alleles as the change in frequency of each allele relative to the change in frequency of the most fit allele, subtracted from 1 (*S10*). Similarly, *S* for *Eda* genotypes was calculated as the change in frequency of the homozygous low genotype relative to the change in frequency of the homozygous complete genotype, subtracted from 1 when selection favored the homozygous complete genotype. When there was selection against the homozygous complete genotype we calculated *S* as the change in frequency of the homozygous low genotype relative to the change in frequency of the homozygous

complete genotype minus 1. We calculated *h* (the dominance coefficient) as the change in frequency of the heterozygous genotype relative to the change in frequency of the homozygous complete genotype minus 1, divided by *S* (*S10*). Standard errors for *S* and *h* are from measurements of $n = 4$ ponds.

Testing Hardy-Weinberg proportions with combined neutral markers

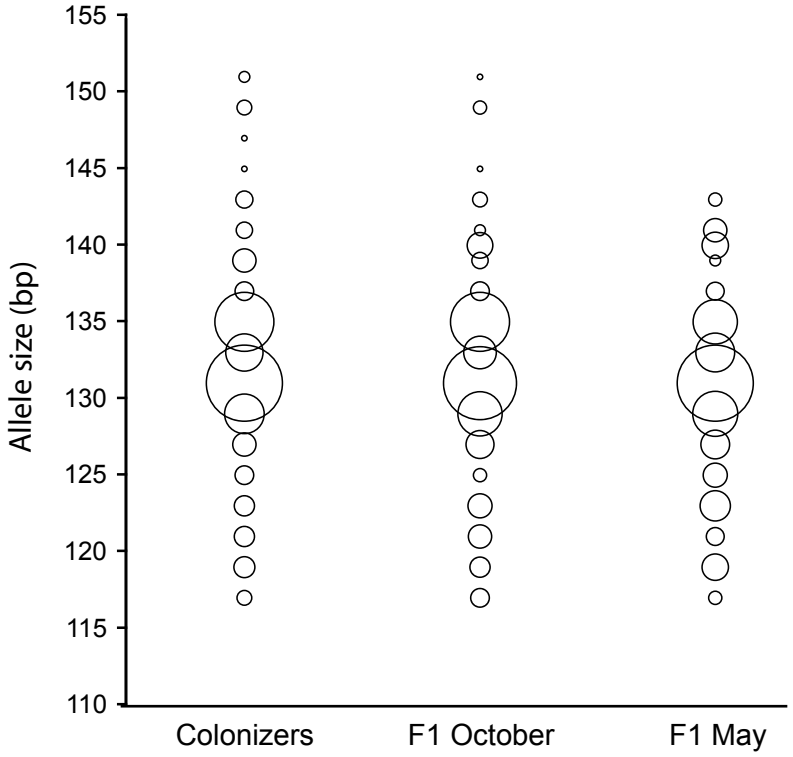
Departures from Hardy-Weinberg proportions were tested in each sample with an approximation of an exact test from a Markov chain iteration implemented in GENEPOP 4.0 (S*11*). Multilocus values of significance for HW tests were calculated following Fisher's method to combine probabilities from different tests (S*12*). Critical significance levels were corrected for multiple tests following the Bonferroni procedure (S*13*).

Supplementary References

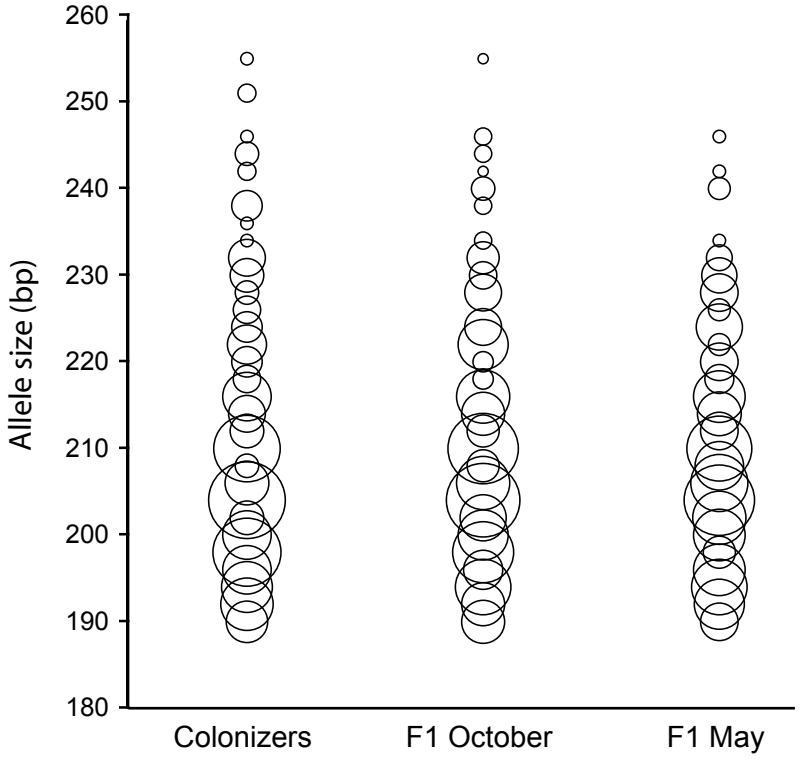
S1. S. A. Foster, *Behav. Ecol.* **5**, 114 (1994).
S2. S. A. Foster, K. A. Shaw, K. L. Robert, J. A. Baker, *Behaviour* **145**, 485 (2008).
S3. R. Siamoto, thesis, University of British Columbia (1995).
S4. P. F. Colosimo *et al.*, *Science* **307**, 1928 (2005).
S5. F. C. Jones, C. Brown, J. M. Pemberton, V. A. Braithwaite, *Journal of Evolutionary Biology* **19**, 1531 (2006).
S6. D. Schluter, *Science* **266**, 798 (1994).
S7. D. Schluter, *Evolution* **57**, 1142 (2003).
S8. T. Day, J. Pritchard, D. Schluter, *Evolution* **48**, 1723 (1994).
S9. J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular cloning: A laboratory manual.* (Cold Spring Harbour Laboratory Press, New York, 1989).
S10. D.L. Hartl, A.G. Clark, Principles of Population Genetics (Sinauer Associates, Inc., Sunderland, ed. 3rd, 1997).
S11. M. Raymond, F. Rousset, *Journal of Heredity* **86**, 248 (1995).
S12. R. R. Sokal, F. J. Rohlf, *Biometry* (W.H. Freeman, San Francisco, ed. 3rd, 1995).
S13. W. R. Rice, *Evolution* **43**, 223 (1989).

Fig. S1: Frequency distribution of alleles at two neutral microsatellite markers unlinked to *Eda,* Stn224 on linkage group 11 (A) and Stn 314 on linkage group 8 (B). Each circle along the vertical axis indicates a distinct allele (in base pair size) while the diameter of each circle is proportional to the relative allele frequency calculated for all ponds combined. An analysis of the allele frequencies at both loci with Fisher's combined probability test indicated that both the colonizers ($\chi^2 = 2.73$, $df = 4$, $P = 0.60$) and the individuals sampled in October 2006 ($\chi^2 = 8.06$, $df = 4$, $P = 0.09$) and May 2007 ($\chi^2 = 7.05$, $df = 4$, $P = 0.13$) were in Hardy-Weinberg Equilibrium.

A.


155
150
145
140
135
130
125
120
115
110
Allele size (bp)
Colonizers
F1 October
F1 May

B.


260
250
240
230
220
210
200
190
180
Allele size (bp)
Colonizers
F1 October
F1 May