A genetic attack on the defense complex

Greg Gibson

Summary

An increasing number of ''non-model'' organisms are becoming accessible to genetic analysis in the field, as evolutionary biologists develop dense molecular genetic maps. Peichel et al.'s recent study[1] provides a microsatellite-based map for threespine stickleback fish (_Gasterosteus aculeatus_), and the first evidence for QTL affecting feeding morphology and defensive armor. This species has undergone rapid and parallel morphological and behavioral evolution, and there is now hope that some of the genes responsible for the divergence may soon be identified. _BioEssays_ 24:487–489, 2002.
© 2002 Wiley Periodicals, Inc.

Introduction

Stickleback fish are among the most amazing and fascinating species that naturalists have chosen to study.[2] For one thing, their close cousins are bizarre little syngnathid seahorses and pipefish, noble in their silliness. For another, they exhibit a range of noteworthy behaviors, not least of which is nest-building by the colorful males who lure females to deposit eggs that they subsequently fertilize and nurture. They're also well known for their wide range of morphological divergence[3] particularly in the glacial lakes and streams of British Columbia, and more generally in coastal inlets on both sides of the Atlantic. Several examples of convergent evolution of two trophic morphs are thought to have been driven by the ecological potential for utilization of alternate food resources that expose the fish to alternate predator types. The resultant benthic (bottom dwelling) and limnetic (column feeding) species pairs that inhabit isolated lakes are reproductively isolated in the field, but can be crossed in the laboratory, facilitating genetic analysis. By comparison with African cichlids, which speciate profusely, sticklebacks thus provide an interesting model for morphological divergence that is not accompanied by classical speciation. For quantitative geneticists, though, one of their more endearing features is a preponderance of meristic variation: for example, _Gasterosteus aculeatus_ are threespine sticklebacks while _Pungitius pungitius_ are ninespined, and within each species there is

Department of Genetics, Gardner Hall, North Carolina State University, Raleigh, NC 27695-7614. E-mail: ggibson@unity.ncsu.edu
DOI 10.1002/bies.10104
Published online in Wiley InterScience (www.interscience.wiley.com).

ample variation for numbered attributes such as vertebrae and gill rakers.

These same features have begun to attract the attention of molecular biologists and geneticists who are interested in dissecting the ontogenic mechanisms of adaptive divergence. One objective has been to sort out the degree of relatedness between populations, particularly among the remarkable series of natural replicates in the British Columbian lakes.[4,5] Another has been to study variation in the expression of key developmental control genes: _Hox_ expression boundaries vary among lines in the laboratory, but the differences don't correlate with any obvious phenotypic differences.[6] A recently published study by Katie Peichel, David Kingsley and colleagues[1] now places sticklebacks at the forefront of efforts to map the quantitative trait loci (QTLs) that give rise to morphological polymorphism in the wild.

Genetic mapping in sticklebacks

Their efforts represent the state of the art in application of genome technology, moving from scratch to mapped QTL in just three years. It undoubtedly didn't hurt that Howard Hughes covered the bill, but one of the really encouraging messages conveyed by this study is that with application and sensible design, similar results might be expected from a moderate-sized grant for just about any species, not just flies, mice and select plants. The authors started simply by sequencing almost 200 kb of genomic clones, which led them to conclude that CA dinucleotide repeats occur once every 14 kb in the stickleback genome. They then pulled out 3,560 clones by screening with a poly-(GT) probe, identified 1,176 repeat-containing sequences, and designed oligonucleotide primer pairs to amplify 410 of these, 281 of which gave robust bands. 227 of these proved to be polymorphic in the 94 individuals scored in a particular cross involving four haploid genome equivalents, enough for an average of almost 10 markers on each of 26 linkage groups (slightly more than the expected 21 chromosomes, indicating that some gaps still need to be filled), or an overall marker density of 4 centiMorgans.

Using this map, Peichel et al.[1] then managed to map 9 QTL affecting a suite of six traits relating to feeding morphology and defensive armor that differ between the benthic and limnetic morphs in Priest Lake B.C., as summarized in Table 1. The first striking result is that just one or two QTL account for up to two-thirds of the phenotypic variance for

Table 1. Summary of QTL effects

| Trait	No. of QTL	PVE (%)[*]	Comment
Number of long gill rakers	0	0	Many small effect loci?
Number of short gill rakers	2	63	Between morphs, with ion-additivity
Lateral plate number	2	48	Between morphs, with ion-additivity
Dorsal spine 1 length	2	38	Between benthics and limnetics
Dorsal spine 2 length	2	39	Within benthics
Pelvic spine length	1	25	Within benthics

[*]Percent of variance explained (by the QTL markers, in progeny of the backcross).

each trait in their cross. The only exception was the number of long gill rakers, an adaptation that is thought to increase feeding performance on zooplankton in limnetics (benthics prefer small invertebrates), but which appears to be influenced by a large number of loci of small effect.[2,3] Measurable genetic loci were detected for several aspects of the stickleback defense complex[7,8] including the number of lateral plates, and lengths of the dorsal and pelvic spines. Variation in these features helps to protect the fish from predation either by birds and bigger fish in open-water environments, or by insects near-shore—long spines either occlude small gapes or provide something to grasp. Other aspects of the skeleton such as the robustness of the pelvic girdle, are also known to be vary greatly among lakes, and it will be interesting to see whether homologs of genes known from mouse genetics to affect skeletogenesis map near any of the stickleback QTLs.

Do these results imply that major effect substitutions are the predominant genetic factors controlling evolutionary divergence? The answer, consistent with recent results from *Drosophila* and several plant systems, is a qualified yes.[9,10] The problem with mapping QTL is that, unlike recombination mapping of Mendelian factors, the associations between genetic markers and phenotypes are probabilistic rather than fully penetrant. Even in a controlled environment, genetic effects are masked or enhanced by chance. In a relatively small cross involving about 100 individuals, and for traits that only differ by a few standard deviation units between the parents, chance can both mask and enhance significant allelic effects. The enhancement effect was first pointed out by William Beavis in 1994,[11] and has the consequence that the observed magnitude of QTL effects tend to overestimate the true genetic contribution. That is, as shown in Fig. 1, while statistical tools are used to ensure that most identified QTL truly exceed a threshold that filters out false positives, the Beavis effect tells us that the loci that make it past the threshold often do so with a little help from chance, and that chance component is included in the estimate of the percentage of the variance that is explained by the QTL. Under some circumstances, the overestimate can be as much as ten-fold or more.[12]

The corollary of this, which surely ought to be called the Butthead effect, is that some QTL go undetected because chance operated in the opposite direction to actual genetic effects, masking these QTL from exceeding the statistical threshold. The result is that the estimate of the number of loci affecting any given quantitative trait is almost by definition an underestimate. The problem is that with small genetic differences between parents and small sample sizes, it is very difficult to obtain a clear picture of the genetic architecture of the variation, so caution must be exercised before jumping to the conclusion that mutations that are responsible for more than a quarter of the differences between two populations dominate the adaptive landscape.
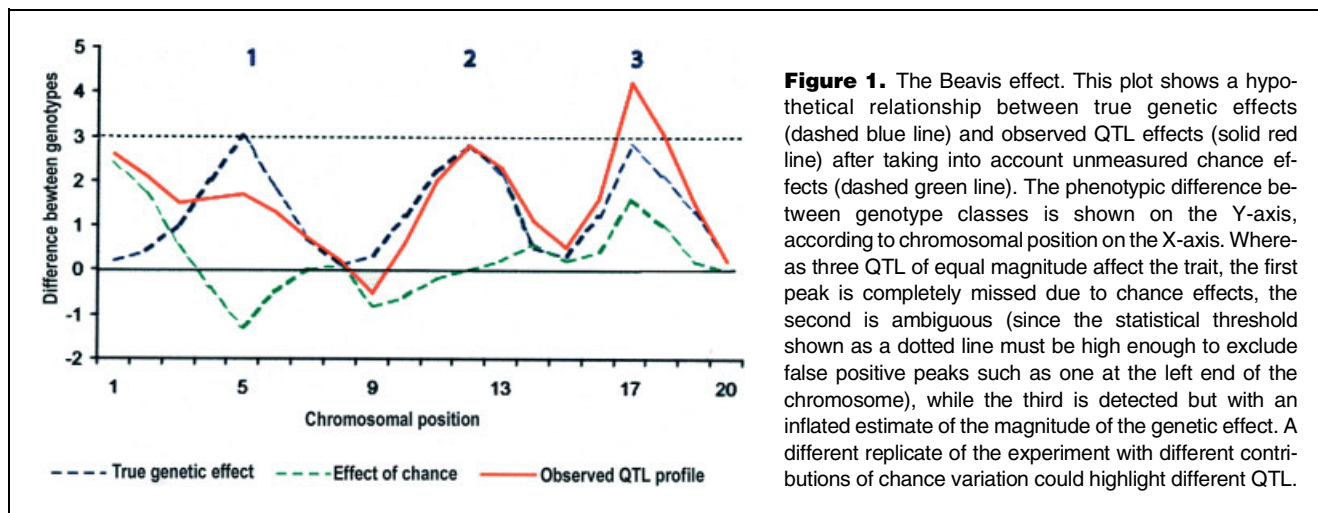
Figure 1. The Beavis effect. This plot shows a hypothetical relationship between true genetic effects (dashed blue line) and observed QTL effects (solid red line) after taking into account unmeasured chance effects (dashed green line). The phenotypic difference between genotype classes is shown on the Y-axis, according to chromosomal position on the X-axis. Whereas three QTL of equal magnitude affect the trait, the first peak is completely missed due to chance effects, the second is ambiguous (since the statistical threshold shown as a dotted line must be high enough to exclude false positive peaks such as one at the left end of the chromosome), while the third is detected but with an inflated estimate of the magnitude of the genetic effect. A different replicate of the experiment with different contributions of chance variation could highlight different QTL.

Interpreting the architecture of genetic variation

A few other notable features of the study need emphasis. The first is that several of the QTL effects were actually observed between genotypes within the benthic class. Because they started their study with outbred wild fish, most of their markers were polymorphic within the parents, and the fact that they adopted a backcross design allowed them to contrast both within and between species differences. That is, an F_1 individual from a cross between a benthic (B1B1) and a limnetic (L1L1) parent was backcrossed to a second benthic fish (B2B3) so that the progeny were all L1B2, L1B3, B1B2, or B1B3 at each marker. The MapQTL software[13] used to perform the QTL analysis allows estimation of the mean phenotype of each class, and hence whether an effect is within or between morphs. That six of the nine identified QTL differ among benthic genotypes (three of which <u>only</u> differed among benthics), indicates that the major effect QTL are not restricted to between morph comparisons and hence are not solely associated with morphological divergence. It will be fascinating to learn whether the same loci appear in similar crosses set up among individuals from different lakes, and hence whether or not the QTL are maintained within populations, possibly as a result of balancing selection.

The study also provides good evidence that interactions among QTL alleles may influence morphological divergence. For example, three of the QTL, including the two loci affecting the number of small gill rakers, showed a significant difference between the phenotype of L1B2 and L1B3 individuals that was not attributable to the difference between the B2 and B3 alleles. That is to say, the magnitude of the effect of the limnetic allele was modified by the nature of the benthic allele on the opposite chromosome. A possible explanation for such effects is that subtle differences in the amount of activity of gene products can lead to a threshold response to a continuous underlying liability, resulting in relatively large meristic changes, but there is still ample room for development of theory on this matter.[14] By contrast, interactions between pairs of loci were generally additive and no synergistic interaction among loci was reported. Nevertheless, summation of effects can produce dramatic differences, such as possession of limnetic alleles at both QTL giving rise to more than 10-fold longer first dorsal spines than those observed in benthic individuals (1.4 versus 0.1 mm). Such results suggest that the morphological divergence may be explained simply by differences in allele frequency of a relatively small number of loci between limnetic and benthic populations.

Finally, it is now legitimate to ask whether it may be possible to actually clone the loci that are responsible for the QTL effects. A pessimist would point out that two LOD support intervals, a standard measure of the possible extent of a QTL peak, cover 30 or 40 cM (almost a chromosome arm) for most of the loci. However, there is no reason why greater marker densities cannot be scored in many more fish and, in this genomic era, we can expect syntenic relationships with other vertebrate genomes to quickly supply candidate genes, such as bone morphogenesis proteins.[15] Sticklebacks are members of the largest radiation of ray-finned fish, the Acanthopterygii, which occurred long after the split in teleost phylogeny from the clade that includes zebrafish, the ostareophysins,[16] probably well before the mammalian radiation. They may however be closer in gene order to pufferfish, for which draft genome sequence already exists.[17,18] Even so, further studies won't just be about finding genes. There is such a wealth of questions that can now be asked with this technology about convergent evolution in glacial lakes, clinal variation along replicate continental coastlines, divergence from sister species, the nature of sex determination, not to mention the quantitative genetics of behavioral variation. Furthermore, we can look forward to seeing similar tools developed for several more non-model organisms of evolutionary and ecological interest made accessible by the arrival of genome research.

References

1. Peichel CL, Nereng K, Ohgi K, Cole B, Colosimo P, Buerkle CA, Schluter D, Kingsley DM. The genetic architecture of divergence between three-espine stickleback species. Nature 2001;;414:901–905.
2. Bell MA, Foster SA. The evolutionary biology of the threespine stickleback. New York: Oxford Science; 1994.
3. Hagen DW, Gilbertson LG. Geographic variation and environmental selection in *Gasterosteus aculeatus* L. in the Pacific Northwest, America. Evolution 1972;;26:32–51.
4. Hatfield T. Genetic divergence in adaptive characters between sympatric species of stickleback. Am Nat 1997;149:1009–1029.
5. Taylor EB, McPhail JD. Evolutionary history of an adaptive radiation in species pairs of threespine sticklebacks (*Gasterosteus aculeatus*): insights from mitochondrial DNA. Biol J Linn Soc 1999;66:271–291.
6. Ahn D, Gibson G. Expression patterns of threespine stickleback *Hox* genes and insights into the evolution of the vertebrate body axis. Dev Genes Evol 1999;209:482–494.
7. Moodie GE. Predation, natural selection and adaptation in an unusual threespine stickleback. Heredity 1972;28:155–167.
8. Reimchen TE. Spine deficiency and polymorphism in a population of *Gasterosteus aculeatus*: an adaptation to predators? Can J Zool 1980;58: 1232–1244.
9. Mackay TFC. The genetic architecture of quantitative traits. Annu Rev Genet 2001;35:303–339.
10. Barton NH, Keightley PD. Understanding quantitative genetic variation. Nat Rev Genet 2002;3:11–21.
11. Beavis WD. The power and deceit of QTL experiments: lessons from comparative QTL studies. In 49th Annual Corn and Sorghum Research Conference. Washington DC: American Seed Trade Association; 1994. pp 252–268.
12. Lynch M, Walsh B. Genetics and analysis of quantitative traits. Sunderland MA: Sinauer Associates; 1998.
13. VanOoijen JW, Maliepard C. MapQTL, version 3.0. Software for the calculation of QTL positions on genetic maps. 1996. Center for Plant Breeding and Reproductive Research, Wageningen.
14. Rendel JM. Canalization and Gene Control. New York: Logos Press; 1967.
15. Kingsley DM. Genetic control of bone and joint formation. Novartis Found Symp 2001;232:213–222.
16. Nelson JS. Fishes of the World, 2nd ed. New York: Wiley-Interscience; 1984.
17. Venkatesh B, Gilligan P, Brenner S. *Fugu*: a compact vertebrate reference genome. FEBS Lett 2000;476:3–7.
18. Roest-Crollius H, et al. Estimate of human gene number provided by genome-wide analysis using *Tetraodon nigroviridis* DNA sequence. Nat Genet 2001;25:235–238.