Structural and Functional Differences in the Promoter and 5′ Flanking Region of *Ldh-B* Within and Between Populations of the Teleost *Fundulus heteroclitus*

Patricia M. Schulte, Marta Gómez-Chiarri and Dennis A. Powers

Hopkins Marine Station, Stanford University, Pacific Grove, California 93950

Manuscript received June 10, 1996
Accepted for publication November 22, 1996

ABSTRACT

We have investigated the mechanisms underlying differences in the transcriptional regulation of lactate dehydrogenase-B (*Ldh-B*) between northern and southern populations of a teleost fish, *Fundulus heteroclitus*. A 1-kb region immediately 5′ of the gene was sequenced from populations throughout the species range. There were two major allele classes in the sample, one containing alleles from Maine and another containing those from Florida. Populations from intermediate localities contained both allele classes. Some individuals from Georgia had sequences intermediate between the two classes, representing either ancestral alleles or recombinants. Tests of neutrality were applied to determine whether observed variation was consistent with neutral expectations. Significant deviations from neutral expectations were detected for the 5′ flanking region, but not for other loci. The functional consequences of flanking sequence variation were assessed by transfection of reporter gene constructs into cultured cells and injection into living fish. Consistent with observed variation in *Ldh-B* transcription rate between populations, significant differences in reporter gene activity were driven by flanking regions from northern and southern populations both in cell culture and *in vivo*. This functional differentiation, coupled with departures from neutral expectations, suggests that selection may have acted on the regulation of *Ldh-B* in *F. heteroclitus*.

POPULATIONS of the killifish, *Fundulus heteroclitus*, occur in bays, estuaries and salt marshes along the Atlantic coast of North America from Newfoundland to Florida. Although the species is sometimes split into two subspecies based on the genetic, morphological and behavioral differences between northern and southern populations, there is a wide zone at intermediate latitudes characterized by clinal variation in these characters (reviewed in POWERS *et al.* 1993). It has been suggested that concordances in the geographic distribution of these characters may be the result of secondary contact between previously isolated forms (POWERS and PLACE 1978; ABLE and FELLEY 1986; GONZALEZ-VILLASENOR and POWERS 1990; BERNARDI *et al.* 1993).

When two differentiated populations of a species come into contact, the cline formed at the intergrade zone will begin to erode due to gene flow unless selection acts to maintain the differentiation. Estimates of gene flow between populations of *F. heteroclitus* suggest that while the characteristics that distinguish the northern and southern forms may have arisen in isolation, the clines are most likely maintained by selection (BROWN and CHAPMAN 1991). Since some allelic isozymes are not clinal, while others are clinally distributed with slopes varying from gradual to steep (reviewed in POWERS *et al.* 1993), it seems likely that the intensity and type of selection, if any, varies from locus to locus.

Corresponding author: Patricia M. Schulte, Hopkins Marine Station, Stanford University, Oceanview Blvd., Pacific Grove, CA 93950. E-mail: schulte@leland.stanford.edu

Allelic isozymes of the glycolytic enzyme lactate dehydrogenase-B (LDH-B) vary from homozygous fixation for one allele at the northern extreme of the species range to fixation for an alternate allele in the south (POWERS and PLACE 1978). These alleles differ in their kinetic properties and are associated with differences in hatching time (DIMICHELE and POWERS 1982a), developmental rate (DIMICHELE *et al.* 1986), swimming performance (DIMICHELE and POWERS 1982b) and differential mortality in selection experiments conducted at elevated temperatures (DIMICHELE and POWERS 1991). Taken together, these studies strongly suggest that natural selection is operating on allelic variants of *Ldh-B*.

In addition to differences in LDH-B kinetics, northern populations have twice as much LDH-B in liver tissue as their southern counterparts (CRAWFORD and POWERS 1989). This twofold difference in LDH-B enzyme concentration between northern and southern populations is associated with twofold differences in both the abundance of the transcript and the transcription rate of the gene (CRAWFORD and POWERS 1992), suggesting that the differences in LDH-B enzyme concentration are the result of changes at the level of transcription. It has been suggested that the differences in enzyme concentration are functionally important (CRAWFORD and POWERS 1989) because twice as much enzyme is required to maintain the same reaction velocity for every 10° decrease in temperature, and the mean annual water temperature at the northern and southern extremes of the species range differs by ~12°.

Although detailed mechanistic analyses of the regulation of transcription have been undertaken for many genes, there have been relatively few attempts to study the evolution of differences in transcriptional regulation within or between species (reviewed in DICKINSON 1991; but see ODGERS 1995). In this report we describe structural and functional variation in the promoter and 5′ flanking regions of *Ldh-B* within and between populations of *F. heteroclitus*, and address some of the genetic and evolutionary mechanisms that may be responsible for this variation.

MATERIALS AND METHODS

DNA amplification and sequencing: *F. heteroclitus* were obtained from St. Andrew's, New Brunswick; Wiscasset, Maine; Woods Hole, Massachusetts; Stone Harbor, New Jersey; Sapelo Island, Georgia; and Whitney Island, Florida. Spleens were dissected, and genomic DNA was isolated by proteinase K digestion and phenol/chloroform extraction. Approximately 1 kb of the *Ldh-B* 5′ flanking sequence was PCR amplified from genomic DNA with *Taq* DNA polymerase (Perkin-Elmer) using primers (FLK) AGATATATTCCAGTTTTGCT and (UTR) CTGGGATCAGAGACTGAG (94° 1 min; 55° 1 min; 72° 2.5 min; 35 cycles). Primers were designed from a sequence of the *Ldh-B* 5′ flanking region obtained by screening a *F. heteroclitus* genomic library (SEGAL *et al.* 1996).

PCR products were end-polished with DNA polymerase I (Klenow fragment) and blunt-end cloned into pBluescript KS$^-$ (Stratagene). At least five clones per individual were sequenced with internal primers on both strands using Sequenáse 2.0 (Stratagene). Occasionally, three substantially different sequences were present in the original sample of five clones. In most cases this was the result of length variation in microsatellites contained within the cloned sequences. The probable cause of these errors is slippage of *Taq* polymerase during PCR, which has been shown to result in artefactual length variation in microsatellites (HOLSTEGE *et al.* 1994). If more than three alleles were detected, the PCR was repeated and treated as before. Majority rule consensus was used to choose the "correct" sequence or sequences.

The sequence of the *Ldh-B* flanking region was obtained from a closely related congeneric species, *F. grandis* (obtained from Galveston Bay, Texas) for use as an outgroup. We were able to obtain the 5′ flanking region of *F. grandis* using the primers designed for *F. heteroclitus*. All sequences have been deposited in GenBank under accession numbers U59833-U59856. Sequence alignments are available from the authors on request.

Phylogenetic analyses: Sequences were aligned using Geneworks (Intelligenetics) followed by alignment by eye. Because of the difficulty involved in aligning microsatellite sequences and the possibility of artefactual length variation (see above), all microsatellites and regions containing gaps were eliminated from genealogical analyses. Tree topologies were obtained using both maximum parsimony (MP) and neighbor-joining (NJ) methods. NJ trees were determined using the computer program MEGA (KUMAR *et al.* 1993). MP trees were determined using both MEGA and PAUP (SWOFFORD 1991). The degree of support for internal branches of the tree was further assessed by bootstrapping with 500 replicates (FELSENSTEIN 1985) using PAUP for the MP trees and MEGA for NJ trees.

Statistical analyses: From the aligned sequences of the 23 *F. heteroclitus* alleles we estimated the parameter θ (TAJIMA 1983) excluding insertions and deletions. Tests of goodness of fit to an infinite sites models were applied (TAJIMA 1989; FU and LI 1993), as well as a test of the concordance between polymorphism and divergence (HKA test, HUDSON *et al.* 1987; KREITMAN and HUDSON 1991). To compare the levels of variation in mitochondrial and autosomal genes, the HKA test was modified so that the effective population size of mitochondrial genes was one-quarter that of autosomal genes (see *e.g.*, BEGUN and AQUADRO 1991; BALLARD and KREITMAN 1994). Modifications were also made to take account of the different number of alleles sampled for each locus (see *e.g.*, BERRY *et al.* 1991). Variation along the sequences was also investigated by making use of a "sliding window" approach (HUDSON and KAPLAN 1988). The recombination rate in the 5′ flanking sequence was estimated by calculating the recombination parameter (C) based on the variance in the distribution of numbers of sites at which pairs of alleles differ, using the method of HUDSON (1987). Linkage disequilibrium between sites was estimated using the Fisher's exact test (WEIR 1990).

Transfection assays: Portions of the 5′ flanking sequence were obtained by PCR amplification from the cloned 5′ flanking sequences *ME4, FL1, GA2, GA4* and *GA5* using primer UTRb (AGATCAGAAGAAAAGTCTGGGATCAGAGACTGAG), in the 5′ untranslated region of the mRNA, and one of the following primers: FLK; f500 (ATAATGAAAGTTTGTGCTC; sites 585–603 on the aligned sequences); f400 (CCTTGGCACCCTCATTTAT; sites 709–727) (Figure 1). These amplifications resulted in products containing the 5′ untranslated region of *Ldh-B* up to the first intron and ~1000, 500 or 400 bp of 5′ flanking sequence, respectively. The PCR products were blunt end cloned into pBluescript KS$^-$ and their sequences confirmed. They were then subcloned in the appropriate orientation into the luciferase plasmid pluc (all reporter gene plasmids kindly provided by Dr. MANFRED SCHARTL) and transfected into a human embryonic kidney cell line (HEK 293; ATCC) to assess transcriptional activity.

Cells were plated at a density of 1×10^6 per 10-cm dish and transfected with plasmid DNA by the calcium phosphate method (SAMBROOK *et al.* 1989). All transfections utilized 10 μg of the luciferase reporter gene plasmid (pluc)/*Ldh-B* flanking sequence construct and 5 μg of the plasmid pCMVtklacZ (containing the β-galactosidase reporter gene) as an internal control for transfection efficiency. Three days after transfection, cells were lysed and luciferase activity in the lysate was detected using the Enhanced Luciferase Assay Kit (Analytical Luminescence Laboratory). Light emission was measured with a single photon count program in a TM Analytic (Mark V) liquid scintillation counter. Galactosidase activity was measured spectrophotometrically using standard techniques (SAMBROOK *et al.* 1989). Results are expressed as fg luciferase per unit β-galactosidase activity. Differences in activity between constructs were tested using a two-way ANOVA followed by multiple comparisons among means (SOKAL and ROLF 1981).

The functional activity of representative full length constructs from Maine and Florida was also assessed by direct injection of reporter gene constructs into *F. heteroclitus in vivo*. Similar approaches have been utilized to analyze the expression of genes in both large and small mammals (WOLFF *et al.* 1990; VON HARSDORF *et al.* 1993). *F. heteroclitus* obtained from Whitney island Florida were acclimated to 50% seawater at 20° for a minimum of 4 months. Fish were lightly anesthetized with 3-aminobenzoic acid ethyl ester (MS-222) and 50 μl of transfection solution was injected into the ventral trunk musculature between the dorsal fin and lateral line. Transfection solution consisted of phosphate buffered saline containing 50 μg of test plasmid (full length constructs ME4 or FL1 see above), 10 μg of control plasmid (CMVtk β-galactosidase), and India ink to localize the site of injection. After 24 hr the fish were sacrificed by rapid decapitation and the region

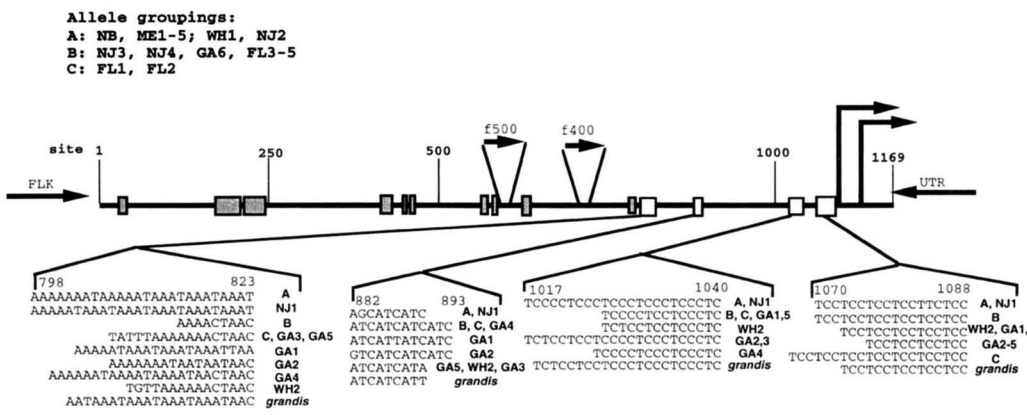

FIGURE 1.—Structure of the *Ldh-B* 5′ flanking region and location of PCR primers. Solid bar represents the flanking sequence numbered from the first base downstream of the 5′ PCR primer (FLK, indicated by arrow) to the last base before the 3′ PCR primer (UTR), including gaps introduced into the alignment. The locations of the major transcription start sites are indicated by bent arrows. Microsatellites are indicated by open boxes and the sequences for all alleles in the sample are shown. Alleles are named for sampling site: *NB,* New Brunswick; *ME,* Maine; *WH,* Woods Hole; *NJ,* New Jersey; *GA,* Georgia; *FL,* Florida. For clarity, the following groupings have been defined: A, all *NB* and *ME* alleles plus *WH1* and *NJ2;* B, *NJ3, NJ4, GA6, FL3-5;* C, *FL1, FL2.* Other insertion/deletions (ins/del), arbitrarily defined relative to the majority of sequences in the sample, are indicated by gray bars and are found in the alleles as follows: sites 30–43, del *F. grandis;* sites 174–210, del *FL1, FL2;* 211–247, ins *GA2, GA5;* 422–434, del *NJ2;* 448–450, del *FL2, FL5, NJ4, GA6;* 464–464, del *WH2, NJ3, FL1, FL4;* 565–572, del group A, *GA3, GA1;* 582–583, ins *F. grandis;* 624–633, del *F. grandis;* 897–899, ins *WH3.*

marked with India ink was excised. A region of equivalent weight was removed from the other side of the fish to control for endogenous β-galactosidase activity, which was never >10% of the total detected. The tissue was homogenized in lysis buffer and the homogenate clarified by microcentrifugation for 15 min at 10,000 × *g.* Luciferase and β-galactosidase activity in the lysate were detected as described above. Three separate experiments were performed, each with three fish injected per construct.

RESULTS AND DISCUSSION

Structure of the *Ldh-B* flanking region: Figure 1 shows the general structure of the *F. heteroclitus Ldh-B* 5′ flanking region. The transcription start sites for the gene are located in a cluster between sites 1095 and 1128, with the two major start sites at 1122 and 1128 (SEGAL *et al.* 1996). The proximal promoter is located between sites 950 and 1150, and contains several highly conserved sequence motifs (sites 1040–1069) that are similar to the Sp1 transcription factor binding site (DYNAN and TJIAN 1983). The major transcription start sites are associated with consensus initiator motifs (ROEDER 1991). Both the initiator and Sp1 binding sites are functionally important components of the *F. heteroclitus* proximal promoter (SEGAL *et al.* 1996). No TATA box is present, typical of the promoters of many housekeeping genes (DYNAN 1986).

Sequence variation in the flanking region: One striking difference between alleles is the presence of a number of microsatellites and small insertion-deletions in the promoter region (Figure 1). Two sets of TCC repeats of variable length (sites 1017–1040 and 1070–1088; Figure 1) are located surrounding the putative Sp1 sites (sites 1041–1069). The basic repeat unit of the Sp1 site (CCGGCC) is present once in alleles *NB, ME1-ME5, WH1, NJ1* and *NJ2;* twice in alleles *WH2, NJ3, NJ4, GA1, GA2, GA3, GA5, GA6, FL1-FL5* and *F. grandis;* and three times in allele *GA4.* We have previously shown that these microsatellites are functionally important components of the *F. heteroclitus Ldh-B* promoter, as assessed by transfection into a heterologous cell line (SEGAL *et al.* 1996). Two additional microsatellites are located within 200 bp upstream of the promoter: a small ATC repeat (sites 882–893) and a highly variable sequence composed almost entirely of thymine and adenine bases (sites 798–823). While no other microsatellites are located within 1 kb upstream of the transcription start sites, there are at least 10 additional small insertions or deletions (Figure 1).

As illustrated in Figure 2, sites that are variable within populations are fairly evenly distributed throughout the flanking region, whereas sites that vary between populations are unevenly distributed. The promoter (sites 950–1150) is the least variable part of the flanking region, both within and between populations. It has only 0.89 variable sites per 100 bp, and is less variable than silent sites within the coding region (which has 1.4 variable sites per 100 bp; BERNARDI *et al.* 1993). Low variability has been reported for the promoter regions of other genes (KREITMAN and HUDSON 1991; ODGERS *et al.* 1995), consistent with functional constraints. In contrast to the low variability in the promoter region, there is a peak of variability between the Maine and Florida populations near the central region of the 5′ flanking sequences (Figure 2). In this central region, the average pairwise difference between alleles from Maine and Florida populations is approximately five times that of the fourfold degenerate sites in the coding region (or 7.46 per 100 bp), but this region is not highly variable within populations (Table 1; Figure 2).

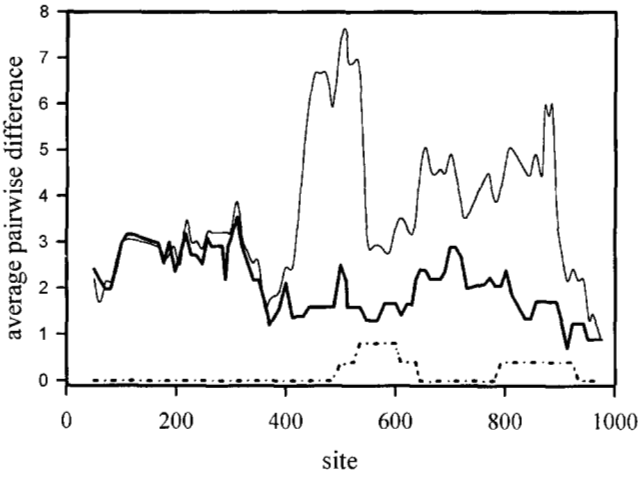

FIGURE 2.—Sliding window comparisons of average pairwise difference between *Ldh-B* 5′ flanking sequences. Fine solid line represents average pairwise difference between Maine and Florida alleles, bold line represents average pairwise difference within the Florida population, dashed line at bottom, average pairwise difference within the Maine population. A window of width 100 bp was placed at the 5′ end of the aligned sequences and the average number of nucleotide differences (excluding all insertions and deletions) between pairs of sequences was calculated for sites contained within the window. The resulting value was assigned to the nucleotide site at the center of the window. The window was advanced in 25-bp increments. The actual width of the window was adjusted to contain 100 sites not including deletions and unalignable regions.

Geographical population structure: Most alleles of the 5′ flanking region fall into one of two distinct groups, one containing all of the alleles from Maine and another containing all of the alleles from Florida (Figure 3). Alleles from Maine and Florida populations are almost as divergent from each other as they are from the outgroup, *F. grandis*. The average pairwise number of variable sites is 39.2 between Maine and Florida populations of *F. heteroclitus*, whereas the average pairwise number of variable sites between *F. grandis* and *F. heteroclitus* from Maine or Florida is 41.6 and 42.6, respectively. There is higher variability within the Florida population than within the Maine population (average pairwise difference within Florida and Maine populations is 12.9 and 0.64 sites, respectively; Figure 3; Table 1). Geographically intermediate populations from New Jersey and Woods Hole contain both Maine-like and Florida-like alleles. As a result, these populations have the highest level of sequence variability (average pairwise difference within the New Jersey sample is 22.9 sites).

Results from previous studies using both nuclear and mitochondrial markers suggest a sharp disjunction between northern and southern *F. heteroclitus* genotypes (GONZALEZ-VILLASENOR and POWERS 1990; BERNARDI *et al.* 1993). However, in phylogenetic analyses of the 5′ flanking sequences, several of the Georgia alleles (*GA1, GA2, GA3, GA4*) appear to be intermediate between Maine and Florida alleles (Figure 3). These alleles could represent ancestral forms or recombinants between Maine and Florida alleles. Putative recombinants have previously been observed in samples of the *Ldh-B* coding region from geographically intermediate populations (POWELL *et al.* 1992; BERNARDI *et al.* 1993).

Recombination: The relative rates of recombination and mutation in a gene region can be estimated based on the variance in the distribution of numbers of sites at which pairs of alleles differ (HUDSON 1987). Taking into account only point mutations, we estimate the recombination parameter (C) to be 22.0 for the 5′ flanking region ($\theta = 30.87$; $S_k^2 = 242$; $N = 23$; HUDSON 1987). Since $C = 4Nc$ (where c is recombination rate) and $\theta = 4N\mu$ (μ = mutation rate), the ratio $C{:}\theta$ should be an index of the relative rates of recombination and mutation. The $C{:}\theta$ ratio for the *Ldh-B* flanking region in *F. heteroclitus* is ~0.71, suggesting that recombination and mutation are about equally frequent on a per nucleotide basis. However, this estimate assumes that the alleles were drawn from a single randomly mating population. Because populations of *F. heteroclitus* throughout the species range may not satisfy this assumption, we also estimated the parameter C using only alleles from the Georgia and Florida populations. In this case, the $C{:}\theta$ is ~1.3, suggesting again that recombination and mutation are about equally frequent in the 5′ flanking region ($\theta = 30.58$; $S_k^2 = 206$; $n = 11$; $C = 40$).

The results of the Fisher's exact test (WEIR 1990) suggest that the variable sites in the *Ldh-B* flanking sequence are not in strong linkage disequilibrium (Figure 4). However, there is some indication that the flanking region can be divided into three major recombining blocks: one containing the promoter region, another in the central region of the sequence, and a third group upstream. We subdivided the sequences into these blocks, and constructed three separate phylogenetic trees (Figure 5). Trees constructed for the upstream and central regions suggest different but reasonably well supported genealogies, which is consistent with recombination. For example, alleles *GA1–GA4*, which appear intermediate between Maine and Florida alleles in Figure 3, group with some of the Florida alleles when only the upstream region is considered (Figure 5) and with the Maine alleles and *F. grandis* when only the central region is considered (Figure 5). While the hypothesis of recombination between Maine and Georgia alleles is reasonable, the possibility of recombination with a *F. grandis* allele cannot be dismissed *a priori*, since the ranges of *F. heteroclitus* and *F. grandis* overlap in northeastern Florida (LEE *et al.* 1980), and hybridization within the genus is common (NEWMAN 1908).

As well as the variable positioning of alleles *GA1–GA4*, there are additional differences between the phylogenies illustrated in Figure 5 that suggest other recombination events may have occurred. The genealogy of the region containing the promoter (sites 655–1169)

TABLE 1

Polymorphic sites within the 5′-flanking region and sites that have diverged between *F. heteroclitus* and *F. grandis*

Site #	11111 1111111111222222233333333334444444444444444555555555555566666666666666666777777888888888888888888900011 1222334557788899011344457766677880001445666714455778888889990111344556689013356667777889990134472333344455666777888899 56 6995678916769256123571238513279198914964663466808452301278958970267584801931545557912393756979206562346124180151382319560 5
ME1	TTACATGCGTGGGCATTGGTGCTTGCAAAAATTGAAACTCGACCTTGTCTGAGCACTGCTGATGTGGGCAAATTAGGGCCGTCTCGGTCGGGCCCCCCTGGCCTCGAACTCTCTAGGTGCAA
ME2	.....
ME4	.....
ME5	.....
ME3	.....T....C.....C.....
NB	.....G.....C.....
WH1	.....T.....
WH2	..A...AA.....AA..A.....TT....G....CCA.TT..GGCAA.CT.A..A..CT...G...TA.....A...AAT....TC...T.....T.CC...
NJ1	.....TC.....A.....
NJ2	.....A.....
NJ3	.C....AAC.....G.....A.....TT.GT.G....C.A.TT.TGGCAA.CT.A..A...C.....TAA...T.....AAT...ATC...T.....A...TACC..G
NJ4	......AAC.....G.....TT.GT.G....C.A..T.TGGCAACCT.A..A..CT.....TAA...T.....AAT...ATC...T.....A...TACC..G
GA1	.C....AAT..A.....T..T...T..GC.....T....G....A.A.....AA...T.....T..A...T.G.....A...TACC..G
GA2	C.....AAT.....AT.A....GC....G.T.....C.A...C.....C..A.....A....TT.G..AAT....TC...T....T....T.GT.C...G
GA3	.C....AAC.....G.A..CT.A.....C..A...T....G..A.C.A.....A.....CTA.....A.A.C.....G...A..T..TCA..TG.....A...T.C..G.
GA4	.C....AATA.....CT.....GC.....T...T.....C.A.T.....A.....A....TC...A..A...CT.....AAT.....A.CATGTA.....A...T.C....
GA5	.C....AA.....A..AT.A.T.GGC...TG.T.....CCA..T.TGGGAA.CT.A..AA.C.....TAA.T.T..GT.AA.....TCA..T.A.....T.C....
GA6	.C....AAC.....G.....A.....TT.G..G....C.A..T.TGGCAA.CT.A..A..CT.....TAA...T.....AAT...ATC...T.....A...TACC..G
FL1	.C....AA....AA...A..AT..A...G...A...T.....C.A.TT..GGCAA.CT.A...A..C.....AA.....AAT....TCA..TG.....A...T.C....
FL2	.C....AA....AA...A..AT..A...G...A...T.....C.A..T..GGCAA.CT.A...A..C.....AA.....AAT....TCA..TG.....A...T.C....
FL3	.C....AA.....G.....A.....TT.GT.G....C.A.TT.TGGCAA.CT.A..A...C.....TAA...T.....AAT...ATC...T.....A...TACC..G
FL4	.C....AAC.....G.....A.....TT.GT.G....C.A.TT.TGGCAA.CT.A..A...C.....TAA...T....AAAT...ATC...TA.....TACC..G
FL5	.C....AAC.....G.....A.....TT.GT.G....C.A..T.TGGCAA.CT.A..A..CT.....TAA...T.....AAT...ATC...T.....A...TACC..G
F. grandis	...GTG---.A.....A.AA.A...G.....A...TT....TT.GC.A.....T...A...G.C.-A.C.....T...A.C-.T..G..AATA...TC...T.....A.....T.C.A.G

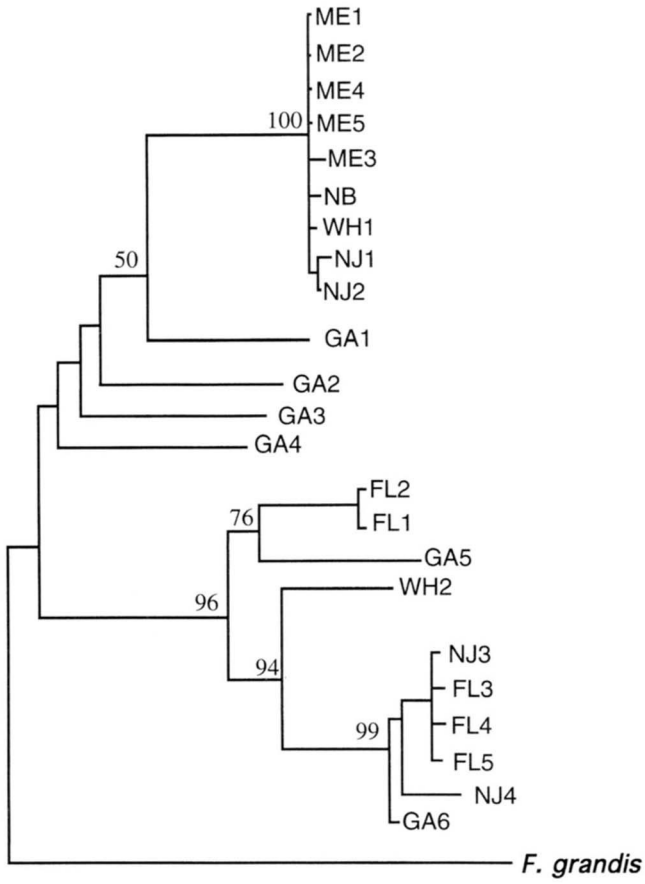

FIGURE 3.—Phylogenetic relationships among lactate dehydrogenase-B flanking sequences of *F. heteroclitus* from various populations along the Atlantic coast of North America. This tree was constructed using the neighbor-joining method using only transversions. All other methods of phylogenetic reconstruction resulted in similar trees. Numbers along the branches of the tree are the percentages of 500 bootstrap replicates supporting a particular clade. For clarity, only bootstrap values of 50% or greater are shown. *FL,* Florida; *GA,* Georgia; *NJ,* New Jersey; *WH,* Woods Hole; *ME,* Maine; *NB,* New Brunswick. Numbers identify alleles. The following alleles were drawn from heterozygous individuals: *ME2, ME3; NJ2, NJ4; FL2, FL4.*

is particularly difficult to reconstruct. Although the three major groups of alleles (Maine-like, *FL1*-like and *FL5*-like) are still apparent, no well supported genealogy can be determined for many of the sequences and a large number of conflicting phylogenies are equally good. Since results of this kind are expected when reconstructing the genealogies of sequences that have undergone recombination (HEIN 1993), these results are consistent with frequent recombination in the promoter region, although it should be noted that when the data are partitioned, statistical power to detect linkage disequilibrium may be very weak.

The genealogies in Figure 5 are based only on point mutations. However, the majority of the variation in the promoter is due to length variation in microsatellites. To further address the possibility of recombination within the promoter, we performed Fisher's exact tests to determine whether the alleles at the microsatellites are linked. There was no significant evidence of linkage disequilibrium (data not shown), suggesting that recombination may have occurred, although the possibility of parallel mutations cannot be ruled out. It has been suggested that the presence of microsatellites may increase the likelihood of polymerase slippage, gene conversion and recombination (see DOVER 1993 for review), which provides a potential mechanism for recombination in the microsatellite-rich *Ldh-B* promoter of *F. heteroclitus.*

Tests of neutrality: The hypothesis that all mutations in the 5′ flanking region are selectively neutral can be tested in a variety of ways. The Tajima test for data from a single locus (TAJIMA 1989) compares the estimate of θ based on the average number of pairwise differences to that based on the number of segregating sites. The test of FU and LI (1993) with outgroup compares estimates of the number of segregating sites to the number of mutations on external branches expected under neutrality. The results of these tests as applied to the flanking region and to previously published *Ldh-B* cDNA and

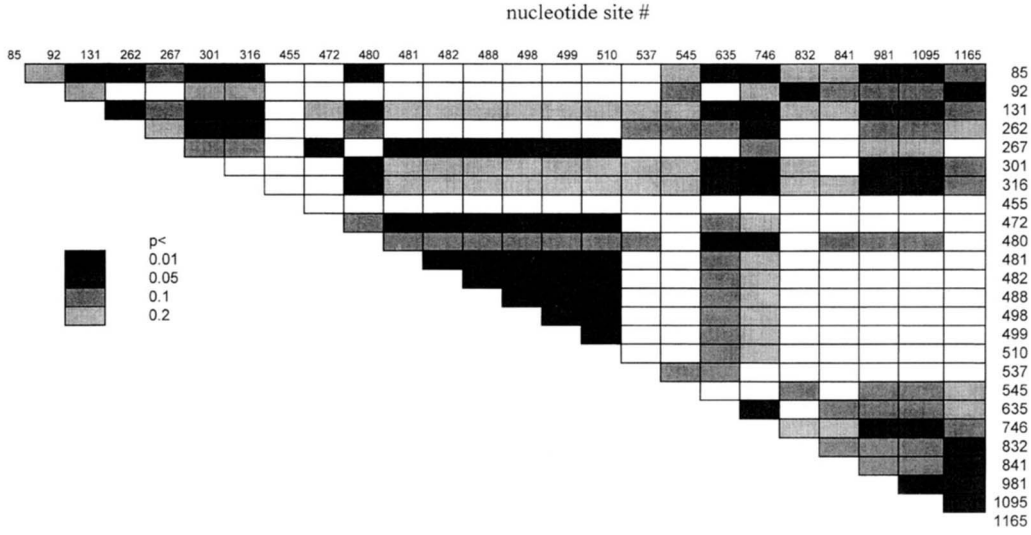

FIGURE 4.—Fisher's exact test for linkage disequilibrium for the Georgia and Florida populations. Because of the very large number of variable sites, we show only those sites with two alleles and for which the rare allele is found in more than three individuals.

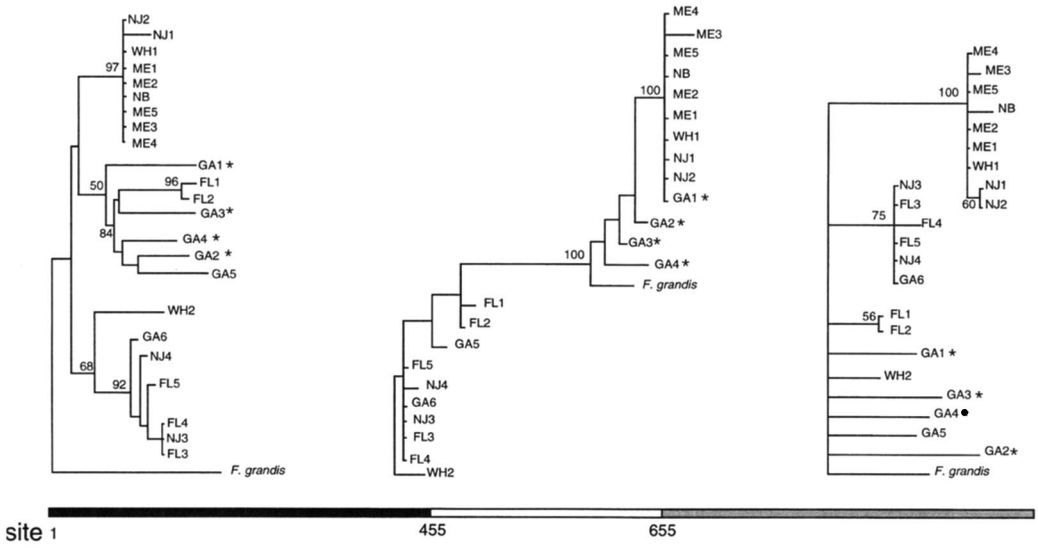

FIGURE 5.—Maximum parsimony trees for three regions of the *Ldh-B* 5′ flanking sequence: sites 1–454, 455–654,;655–1169. Sites 1–454 50% majority rule consensus of 10 trees; sites 455–654 generate a single most parsimonious tree; sites 655–1169 50% majority rule consensus of 53 equally parsimonious trees. Bar below represents the flanking sequence, major transcription start sites are at 1122 and 1128. Numbers above branches indicate bootstrap values for 500 replicates (expressed as percentages). All values above 50% are shown.

mitochondrial cyt-b (cyt-b) data (BERNARDI *et al.* 1993) are presented in Table 2. There is no evidence for departure from neutrality when these tests are applied to pooled data from all populations, although both the FU and LI (1993) and TAJIMA (1989) tests tend toward slightly negative values. Slightly negative values are to be expected when one or a few sequences are drawn from several different geographic regions, since such a sample is likely to have an excess of unique or low frequency polymorphisms. When the approach of TAJIMA (1989) is applied to each population separately, there is no evidence for departures from neutrality for any of the loci (Table 2). On the other hand, the FU and LI (1993) test rejects the null hypothesis that all sites are neutral in the 5′ flanking region for the Maine and New Jersey populations, but not for the *Ldh-B* coding region or cyt-b (Table 2). The difference between the two tests may reflect the low sensitivity of the TAJIMA (1989) analysis when sample sizes are small (see SIMONSEN *et al.* 1995 for discussion).

The positive D-statistic in the FU and LI (1993) test for the flanking region in New Jersey population (Table 2) suggests that there are two ancient alleles within the population. This result is consistent with either secondary intergradation between northern and southern forms, or balancing selection (FU and LI 1993). Because data are not available for the other loci in the New Jersey population, we cannot distinguish between these two possibilities, but previous studies (GONZALEZ-VILLASENOR and POWERS 1990) have also suggested that clines in *F. heteroclitus* may be the result of secondary intergradation.

The negative D-statistic for the flanking region of the Maine population (Table 2) suggests that the alleles in this population are of recent origin, which is consistent with a number of interpretations including purifying selection, background selection (CHARLESWORTH *et al.* 1993), selective sweep, population bottleneck, or rapid population expansion (FU and LI 1993). Purifying selection and background selection are unlikely in this case because they would be expected to influence variation in all populations. Since Maine populations have lower variability than southern populations for most molecular markers (POWERS *et al.* 1993), a population bottleneck or rapid population expansion are more reasonable explanations for the departures from neutrality in

TABLE 2

D statistics for various region of the *F. heteroclitus* genome

	5′ flanking region		*Ldh-B* cDNA		Cytochrom-b	
Population	Tajima	Fu and Li	Tajima	Fu and Li	Tajima	Fu and Li
All	−0.82	−0.72	−0.18	−0.17	−0.29	−0.73
North	−0.62	−1.99*	−0.59	−0.85	−0.54	0.12
New Jersey	0.27	1.91*	NA	NA	NA	NA
South	−0.12	−0.75	−0.44	−0.77	−0.80	−0.89

Tajima D statistics are calculated as in TAJIMA (1989), Fu and Li D-statistics (with outgroup) are calculated as in FU and LI (1993) using *F. grandis* as outgroup for the 5′ flanking region and *F. parvipinnus* as outgroup for *Ldh-B* cDNA and cytochrome-b. *Ldh-B* cDNA and cytochrome-b sequences from BERNARDI *et al.* (1993). Populations are as follows: North includes Maine and New Brunswick for the 5′ flanking region; Maine and Nova Scotia for cDNA and cytochrome-b. South includes all individuals from Georgia and Florida for all loci. *Significant D-statistics, $P < 0.05$. NA indicates data not available.

TABLE 3

HKA tests for *Ldh-B* 5′ flanking region, coding region and cytochrome-b (cyt-b)

	Ldh-B 5′ flanking	*Ldh-B* coding	cyt-b	χ^2
Polymorphism (N)	5 (6)	6 (6)	2 (4)	Flanking and coding
Polymorphism (S)	29 (4)	6 (4)	5 (4)	4.95*
				Flanking and cyt-b
Divergence (N *vs.* S)	41.8	5.8	4.5	11.24***
Polymorphism (N)/divergence	0.12	1.03	0.44	Cyt-b and coding
Polymorphism (S)/divergence	0.69	1.03	1.11	2.9

The HKA test of HUDSON *et al.* (1987) has been modified to take into account the difference in effective population size between nuclear and mitochondrial genomes and the fact that there are different sample sizes for the two loci. Variation within populations is calculated as number of polymorphic sites within Maine, New Brunswick and Nova Scotia (N) or Florida (S). Divergence between populations is average pairwise difference between N and S. *Significant, $P < 0.05$; ***Significant, $P < 0.001$. Sample size indicated in parentheses.

the FU and LI tests. Based on results from a variety of other markers, we have previously suggested that the northern populations may have undergone a population expansion (CASHON *et al.* 1981; POWERS *et al.* 1986, 1993). However, there is no indication of significant departures from neutral expectations for the *Ldh-B* coding region or cyt-b in the Maine population (Table 2) as would be expected if a population bottleneck or rapid population expansion had occurred. If the FU and LI test is sufficiently sensitive to detect a population bottleneck at cyt-b and the coding region, this discrepancy between loci suggests the possibility of a selective sweep at the 5′ flanking region in the northern population.

The possibility that various regions of the genome are subject to different evolutionary forces can be examined using the HKA test (HUDSON *et al.* 1987). By comparing the ratio of intraspecific to interspecific nucleotide variation at several loci, the HKA approach tests a prediction of the neutral theory that levels of polymorphism within a species should be correlated with levels of divergence between species across all loci. This test should be able to distinguish between a population bottleneck, which would affect all loci, and a selective sweep, which should only affect the locus under selection. We compared variation at three loci: the *Ldh-B* coding region, mitochondrial cyt-b (BERNARDI *et al.* 1993) and the *Ldh-B* 5′ flanking region, using a modified HKA test that compares average pairwise difference within and between species (KREITMAN and HUDSON 1991). When polymorphism in northern and southern populations is compared to divergence between them, there is significant deviation from neutral expectations only in comparisons involving the *Ldh-B* 5′ flanking region (Table 3). This suggests that the deviations from neutrality are the result of variation in the 5′ flanking region rather than in one of the other two loci, possibly as a result of selection. However, the results of HKA tests should be interpreted with caution since the violation of any one of the underlying assumptions may result in rejection of the null hypothesis. The assumptions that may be violated by the data presented here are: (1) no selection, (2) large constant population size, (3) no linkage between loci, (4) no recombination within a locus, and (5) no ancestral polymorphism. Violations of assumptions 3–5 are thought to make the test more conservative (see for example HUDSON *et al.* 1987; FORD *et al.* 1994 for discussion) and thus are unlikely to account for the deviations from neutrality. Therefore these results are most likely accounted for by either changes in population size or selection.

Functional analyses: If selection is acting on the 5′ flanking region, the minimal condition that must be met is evidence of some form of functional differentiation between alleles. To determine whether any of the sequence variability in the 5′ flanking region of *F. heteroclitus* has functional consequences for transcriptional regulation, we compared the ability of several randomly selected alleles from northern and southern populations to promote the transcription of a luciferase reporter gene in human embryonic kidney (HEK 293) cells. We chose to perform the initial functional experiments using HEK 293 due to the lack of cell lines from *F. heteroclitus* and the fact that most fish cell lines and primary cultures tested showed low and variable transfection efficiencies (M. GÓMEZ-CHIARRI, unpublished results). HEK 293 has extremely good growth and high transfection efficiencies that maximize the sensitivity of the functional assays. This may be critical since activity differences in natural populations appear to be small (for example, see ODGERS *et al.* 1995), and the twofold differences expected for *F. heteroclitus Ldh-B* might be essentially undetectable under many conditions.

We first assessed the luciferase activity promoted by one of the Maine alleles (*ME4*) and several of the most divergent alleles from southern populations (*GA2, GA4, GA5, FL1*), including two of the putative recombinants. Length variation in the microsatellites that alter the spacing between conserved, functionally important sites in the promoter does not appear to have strong effects on transcription since there are no differences in the activity of some constructs that differ in microsatellite length (*e.g., GA5* and *ME4; GA2* and *GA4;* Figure 6). In

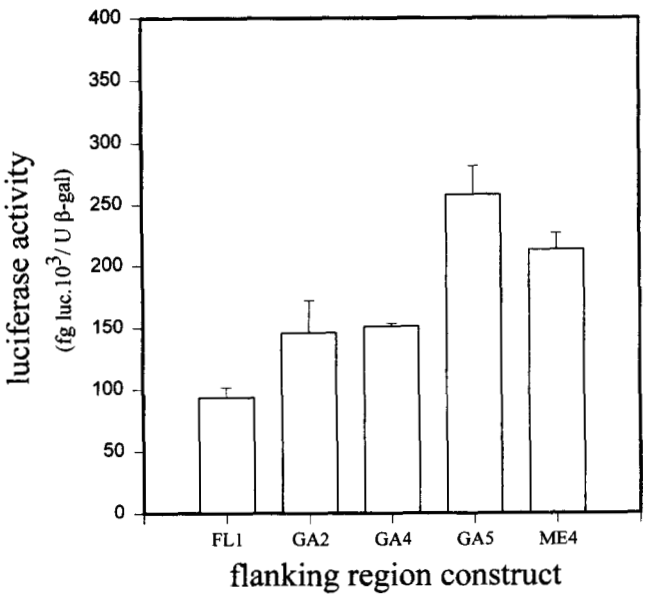

FIGURE 6.—Activity of *Ldh-B* flanking sequence constructs *ME4, GA2, GA4, GA5* and *FL1* when transfected into cells in culture in fg luciferase per unit β-galactosidase. Bars, mean ± SEM ($n = 3$). *ME4* and *GA5* are significantly different from all others $P < 0.05$.

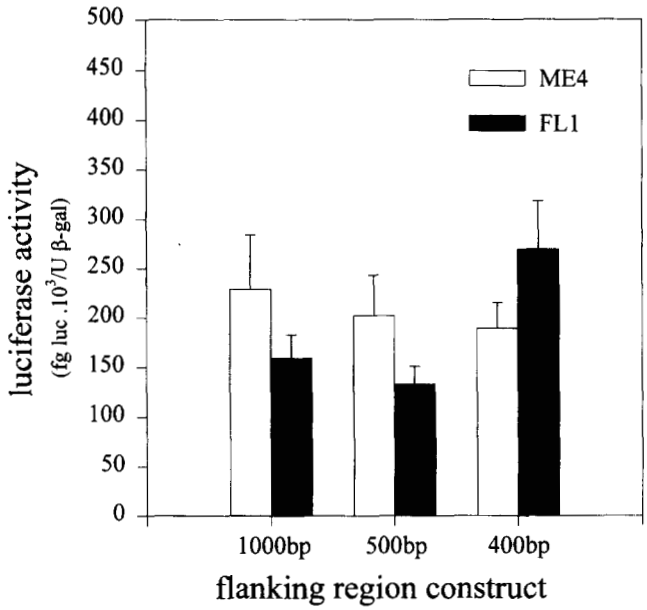

FIGURE 7.—Activity of *Ldh-B* deletion constructs for representative northern (*ME4*) and southern (*FL1*) sequences when transfected into cells in culture. Activity is fg luciferase per unit β-galactosidase. Bars, mean ± SEM ($n = 5$). Three constructs of different length were tested. Each construct contained 70 bp of *Ldh-B* 5′ untranslated sequence plus ~1000, 500 or 400 bp of 5′ flanking sequence.

contrast, the Maine allele (*ME4*) promotes significantly greater transcription than the mean level promoted by the southern alleles (*GA2, GA4, GA5, FL1*) (20.62 ± 2.00 *vs.* 15.21 ± 7.41 fg luc/Uβgal; $P < 0.05$) and significantly greater transcription than all but one southern allele (*GA5*). This is consistent with the differences in *Ldh-B* transcription rate between populations (CRAWFORD and POWERS 1992). To localize the sequences responsible for the observed differences in transcriptional activity between alleles, we compared the activity of various deletions of *FL1* and *ME4*. As can be seen in Figure 7, there are significant differences between *ME4* and *FL1* 1000- and 500-bp constructs, but no differences between *ME4* and *FL1* for the 400-bp construct. There are no significant differences in activity among northern constructs, but for southern constructs the shortest construct gives significantly greater activity than either of the two longer constructs (Figure 7). These results suggest that there are sequences repressing transcription in the region deleted from the shortest of the constructs (sites 585–709) in FL1 and that this repressor activity is absent in *ME4*.

We examined the region between sites 585 and 709 in FL1 for known transcription factor binding sites using the program Signal Scan (PRESTRIDGE 1993). A sequence located between sites 704 and 710 in the *FL1* allele is identical to the MTV-GRE negative regulatory motif (LANGER and OSTROWSKI 1988). This putative repressor motif is altered in ME4 due to a polymorphism at site 707 (Table 1). *GA5* also lacks this motif (Table 1), which might account for the high level of transcription promoted by this allele (Figure 6). However, further experiments are required to assess this possibility since *GA5* also differs from FL1 at positions other than the putative repressor.

Because these functional assays were performed using mammalian cells in culture, these results should be interpreted with caution. However, the known behavior of this repressor is consistent with that of *F. heteroclitus Ldh-B*. The MTV-GRE repressor is known to function only at low glucocorticoid concentration (LANGER and OSTROWSKI 1988), and differences in *Ldh-B* gene regulation between populations of *F. heteroclitus* are observed only at higher temperatures (SEGAL and CRAWFORD 1994) when glucocorticoid levels are low (LEACH and TAYLOR 1977). This suggests that a similar steroid-mediated mechanism might be responsible for the differences in *Ldh-B* regulation between northern and southern *F. heteroclitus*.

Because the action of regulatory sequences in heterologous cell cultures is not always equivalent to their action *in vivo*, we also assayed the function of the reporter gene constructs from Maine and Florida in *F. heteroclitus* muscle. When injected into fish obtained from the Florida population, the Maine construct promotes twofold greater luciferase activity than the Florida construct (Figure 8). While we have yet to directly test the function of these constructs in liver, the results of both the cell culture and intramuscular injections are consistent with the differences in *Ldh-B* transcription observed between northern and southern populations (CRAWFORD and POWERS 1992).

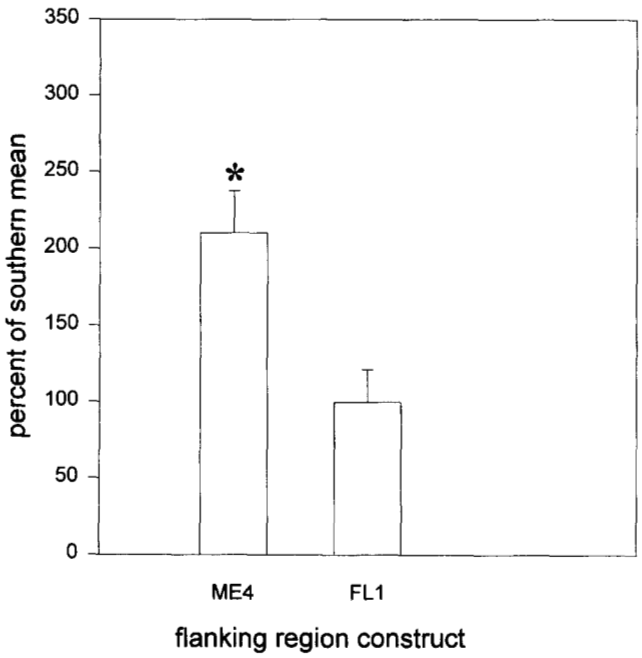

FIGURE 8.—Activity of *Ldh-B* 5′ flanking region constructs for representative northern (*ME4*) and southern (*FL1*) sequences when injected into *F. heteroclitus* muscle *in vivo*. Bars are mean ± SEM from three separate experiments expressed relative to the southern mean from each experiment. Each construct contained 70 bp of *Ldh-B* 5′ untranslated sequence plus ~1000 bp of 5′ flanking sequence.

Conclusions: There is extensive geographic variation in the 5′ flanking sequences of the *Ldh-B* gene in different populations of *F. heteroclitus*. Maine and Florida populations can be clearly distinguished, based on phylogenetic analyses of their alleles. Geographically intermediate populations are the most variable, and contain both northern and southern allele classes, which is consistent with a secondary intergradation model of cline formation in *F. heteroclitus*. Various tests of neutrality indicate that the 5′ flanking region is not behaving according to neutral expectations. These deviations from neutrality could be the result of a number of processes, including changes in population size and selection. *In vitro* transfection assays and *in vivo* injections show that there are functional differences between northern and southern alleles consistent with the differences in *Ldh-B* transcription between populations (CRAWFORD and POWERS 1992). This functional differentiation, coupled with departures from neutral expectations in the 5′ flanking region, suggests the possibility that selection has acted on the regulation of *Ldh-B* in *F. heteroclitus*.

We thank D. BUSHEK, D. CRAWFORD, L. DIMICHELE, B. MARSHALL, V. PIERCE, J. SEGAL and T. SINGER for help in obtaining various fish samples. Thanks to J. QUATTRO, G. BERNARDI, K. CALLICOTT, S. CLENDENNEN, C. REEB, W. EANES and the anonymous reviewers of this manuscript for their helpful comments and suggestions. Particular thanks to M. KREITMAN for his assistance with the HKA analyses. This work was supported by National Science Foundation grants BSR-9022648 and DEB-96-29572 to D.P. and National Science Foundation dissertation improvement grant DEB-95-20567 to P.S.

LITERATURE CITED

ABLE, K. W., and J. D. FELLEY, 1986 Geographical variation in *Fundulus heteroclitus:* tests for concordance between egg and adult morphologies. Am. Zool. **26:** 145–157.

BALLARD, J. W. O., and M. KREITMAN, 1994 Unraveling selection in the mitochondrial genome of Drosophila. Genetics **138:** 757–772.

BEGUN, D. J., and C. F. AQUADRO, 1991 Directional selection and genetic hitchhiking of the *yellow-achaete* region in *Drosophila*. Genetics **129:** 1147–1158

BERNARDI, G., P. SORDINO and D. A. POWERS, 1993 Concordant mitochondrial and nuclear DNA phylogenies for populations of the teleost fish *Fundulus heteroclitus*. Proc. Natl. Acad. Sci. USA **90:** 9271–9274.

BERRY, A. J., J. W. AJIOKA and M. KREITMAN, 1991 Lack of polymorphism on the *Drosophila* fourth chromosome resulting from selection. Genetics **129:** 1111–1117.

BROWN, B. L., and R. W. CHAPMAN, 1991 Gene flow and mitochondrial DNA variation in the killifish, *Fundulus heteroclitus*. Evolution **45:** 1147–1161.

CASHON, R. E., R. J. VAN BENEDEN and D. A. POWERS, 1981 Biochemical genetics of *Fundulus heteroclitus*. IV: Spatial variation in gene frequencies of Idh-A, Idh-B, 6-Pgdh-A, and Est-S. Biochem. Genetics **19:** 715–728.

CHARLESWORTH, B., M. T. MORGAN and D. CHARLESWORTH, 1993 The effect of deleterious mutations on neutral molecular variation. Genetics **134:** 1289–1303.

CRAWFORD, D. L., and D. A. POWERS, 1989 Molecular basis of evolutionary adaptation at the lactate dehydrogenase-B locus in the fish *Fundulus heteroclitus*. Proc. Natl. Acad. Sci. USA **86:** 9365–9369.

CRAWFORD, D. L., and D. A. POWERS, 1992 Evolutionary adaptation to different thermal environments via transcriptional regulation. Mol. Biol. Evol. **9:** 806–813.

DICKINSON, W. J., 1991 The evolution of regulatory genes and patterns in Drosophila, pp. 127–173 in *Evolutionary Biology*, Vol. 25, edited by M. K HECHT, B. WALLACE and R. J. MACINTYRE. Plenum Press, NY.

DIMICHELE, L., and D. A. POWERS, 1982a LDH-B genotype specific hatching times of *Fundulus heteroclitus* embryos. Nature **296:** 563–564.

DIMICHELE, L., and D. A. POWERS, 1982b Physiological basis for swimming endurance differences between LDH-B genotypes of *Fundulus heteroclitus*. Science **216:** 1014–1016.

DIMICHELE, L., and D. A. POWERS, 1991 Allozyme variation, developmental rate and differential mortality in the model teleost, *Fundulus heteroclitus*. Physiol. Zool. **64:** 1426–1443.

DIMICHELE, L., D. A. POWERS and J. A. DIMICHELE, 1986 Developmental and physiological consequences of genetic variation at enzyme synthesizing loci in *Fundulus heteroclitus*. Am. Zool. **26:** 201–208.

DIMICHELE, L., K. T. PAYNTER and D. A. POWERS, 1991 Lactate dehydrogenase-B allozymes directly effect development of *Fundulus heteroclitus*. Science **253:** 898–900.

DOVER, G. A., 1993 Evolution of genetic redundancy for advanced players. Curr. Opin. Genet. Dev. **3:** 902–910.

DYNAN, W. S., 1986 Promoters for housekeeping genes. Trends Biochem. **11:** 196–197.

DYNAN, W. S., and R. TJIAN, 1983 The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. Cell **35:** 79–87.

FELSENSTEIN, J., 1985 Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:** 783–791.

FORD, M. J., C. K. YOON and C. F. AQUADRO, 1994 Molecular evolution of the *Period* gene in *Drosophila athabasca*. Mol. Biol. Evol. **11:** 169–182.

FU, Y.-X., and W.-H. LI, 1993 Statistical tests of neutrality of mutations. Genetics **133:** 693–709.

GONZALEZ-VILLASENOR, L. I., and D. A. POWERS, 1990 Mitochondrial DNA restriction-site polymorphisms in the teleost *Fundulus heteroclitus* support secondary intergradation. Evolution **44:** 27–37.

HEIN, J., 1993 A heuristic method to reconstruct the history of sequences subject to recombination. J. Mol. Evol. **36:** 396–405.

HOLSTEGE, F. C. P., P. C. VAN DER VLIET and H. T. M. TIMMERS, 1994 Instability of human TATA-binding protein CAG triplet repeats during amplification by PCR. Biochem. Biophys. Acta **1219:** 157–159.

HUDSON, R. R., 1987 Estimating the recombination parameter of a finite population model without selection. Genet. Res. **17:** 245–250.

HUDSON, R. R., and N. L. KAPLAN, 1988 The coalescent process in models with selection and recombination. Genetics **120:** 831–840.

HUDSON, R. R., M. KREITMAN and M. AGUADE, 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116:** 153–159.

KRAFT, H. J., W. HENDRIKS, W. W. DE JONG, N. H. LUBSEN and J. G. SCHOENMAKERS, 1993 Duck lactate dehydrogenase B/epsilon-crystallin gene. Lens recruitment of a GC-promoter. J. Mol. Biol. **229:** 849–859.

KREITMAN, M., and R. R. HUDSON, 1991 Inferring the evolutionary histories of the *Adh* and *Adh-dup* loci in *Drosophila melanogaster* from patterns of polymorphism and divergence. Genetics **127:** 565–582.

KUMAR, S., K. TAMURA and M. NEI, 1993 *MEGA: Molecular Evolutionary Genetics Analysis, version 1.0.* The Pennsylvania State University, University Park, PA.

LANGER, S. J., and M. C. OSTROWSKI, 1988 Negative regulation of transcription *in vitro* by a glucocorticoid response element is mediated by a trans-acting factor. Mol. Cell. Biol. **8:** 3872–3881.

LEACH, G. L., and M. H. TAYLOR, 1977 Seasonal measurements of serum glucose and serum cortisol in a natural population of *Fundulus heteroclitus* L. Comp. Biochem. Physiol. **56A:** 217–223.

LEE, D. S., C. R. GILBERT, C. H. HOCUTT, R. E. JENKINS, D. E. MCALLISTER *et al.,* 1980 *Atlas of North American Freshwater Fishes.* North Carolina State Museum of Natural History.

NEWMAN, H. H., 1908 The process of heredity as exhibited by the development of *Fundulus* hybrids. J. Exp. Zool. **5:** 503–571.

ODGERS, W. A., M. J. HEALY and J. G. OAKESHOTT, 1995 Nucleotide polymorphism in the 5′ promoter region of esterase 6 in *Drosophila melanogaster* and its relationship to enzyme activity variation. Genetics **141:** 215–222.

POWELL, M. A., D. L. CRAWFORD, T. LAUERMAN and D. A. POWERS, 1992 Analysis of cryptic alleles of *Fundulus heteroclitus* lactate dehydrogenase by a novel allele-specific polymerase chain reaction. Mol. Mar. Biol. Biotechnol. **1:** 391–396.

POWERS, D. A., and A. R. PLACE, 1978 Biochemical genetics of *Fundulus heteroclitus.* I. Temporal and spatial variation in gene frequencies of: Ldh-B, Mdh-A, Gpi-B and Pgm-A. Biochem. Genet. **16:** 593–607.

POWERS, D. A., I. ROPSON, D. BROWN, R. J. VAN BENEDEN, R. CASHON *et al.,* 1986 Geographic variation in *Fundulus heteroclitus.* Am. Zool. **26:** 131–144.

POWERS, D. A., M. SMITH, I. GONZALEZ-VILLASENOR, L. DIMICHELE, D. CRAWFORD *et al.,* 1993 A multidisciplinary approach to the selectionist/neutralist controversy using the model teleost *Fundulus heteroclitus,* pp. 43–107 in *Oxford Survey of Evolutionary Biology,* Vol. 9, edited by D. FUTUYUMA and J. ANTONOVICS. Oxford University Press, Oxford.

PRESTRIDGE, D. S., and G. STORMO 1991 Signal Scan 3.0: new data base and program features. Computer Appl. Biosciences **9:** 43–108.

ROEDER, R. G., 1991 The complexities of eukaryotic transcription preinitiation complex assembly. Trends Biochem. Sci. **16:** 402–408.

SAMBROOK, J. E., E. F. FRITSCH and T. MANIATIS, 1989 *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

SEGAL, J. A., and D. L . CRAWFORD, 1994 LDH-B enzyme expression: the mechanisms of altered gene expression in acclimation and evolutionary adaptation. Am. J. Physiol. **267:** R1150-R1153.

SEGAL, J. A., P. M. SCHULTE, D. A. POWERS and D. L. CRAWFORD, 1996 Descriptive and functional characterization of variation in the *Fundulus heteroclitus Ldh-B* proximal promoter. J. Exp. Zool. **275:** 355–364.

SIMONSEN, K. L., G. A. CHURCHILL and C. F. AQUADRO, 1995 Properties of statistical tests of neutrality for DNA polymorphism data. Genetics **141:** 413–429.

SOKAL, R. R., and F. J. ROLF, 1981 *Biometry,* Ed. 2. W. H. Freeman and Co., New York.

SWOFFORD, D. L., 1991 *PAUP: Phylogenetic Analysis Using Parsimony. Version 3.0.* Illinois Natural History Survey. Champaign, IL.

TAJIMA, F., 1983 Evolutionary relationship of DNA sequences in finite populations. Genetics **105:** 437–460.

TAJIMA, F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123:** 585–595.

TAKANO, T., and S. S.-L. LI, 1989 Structure of the human lactate dehydrogenase B gene. Biochem. J. **257:** 921–924.

VON HARSDORF, R., R. J. SCHOTT, Y.-T. SHEN, S. F. VATNER, V. MAHDAVI *et al.,* 1993 Gene injection into canine myocardium as a useful model for studying gene expression in the heart of large mammals. Circ. Res. **72:** 688–695.

WEIR, B. S., 1990 *Genetic Data Analysis.* Sinauer Associates, Sunderland, MA.

WOLFF, J. A., R. W. MALONE, P. WILLIAMS, W. CHONG, G. ACSADI *et al.,* 1990 Direct gene transfer into mouse muscle *in vivo.* Science **247:** 1456–1468.

Communicating editor: W. F. EANES