Evolutionary Physiology and Genomics in the Highly Adaptable Killifish (*Fundulus heteroclitus*)

Douglas L. Crawford,[*1] *Patricia M. Schulte,*[2] *Andrew Whitehead,*[3] *and Marjorie F. Oleksiak*[1]

ABSTRACT

By investigating evolutionary adaptations that change physiological functions, we can enhance our understanding of how organisms work, the importance of physiological traits, and the genes that influence these traits. This approach of investigating the evolution of physiological adaptation has been used with the teleost fish *Fundulus heteroclitus* and has produced insights into (i) how protein polymorphisms enhance swimming and development; (ii) the role of equilibrium enzymes in modulating metabolic flux; (iii) how variation in DNA sequences and mRNA expression patterns mitigate changes in temperature, pollution, and salinity; and (iv) the importance of nuclear-mitochondrial genome interactions for energy metabolism. *Fundulus heteroclitus* provides so many examples of adaptive evolution because their local population sizes are large, they have significant standing genetic variation, and they experience large ranges of environmental conditions that enhance the likelihood that adaptive evolution will occur. Thus, *F. heteroclitus* research takes advantage of evolutionary changes associated with exposure to diverse environments, both across the North American Atlantic coast and within local habitats, to contrast neutral versus adaptive divergence. Based on evolutionary analyses contrasting neutral and adaptive evolution in *F. heteroclitus* populations, we conclude that adaptive evolution can occur readily and rapidly, at least in part because it depends on large amounts of standing genetic variation among many genes that can alter physiological traits. These observations of polygenic adaptation enhance our understanding of how evolution and physiological adaptation progresses, thus informing both biological and medical scientists about genotype-phenotype relationships. © 2020 American Physiological Society. *Compr Physiol* 10:637-671, 2020.

Didactic Synopsis

Major teaching points

Evolutionary studies in the small teleost fish *Fundulus heteroclitus* have revealed how genetic variation influences biochemical and physiological function.

- These findings include
  - Genetic variation between LDH-B alleles that alters swimming abilities and development.
  - DNA sequence variation that changes LDH-B enzyme expression.
  - Quantitative differences in glycolytic enzyme expression related to thermal environment and how they modify cardiac metabolism.
  - mRNA expression changes that are influenced by physiological acclimation and alter metabolism.
  - Population-specific variation in DNA sequences and mRNA expression that mitigates variation in temperature, pollution, and osmotic environments.
  - DNA sequence variation within and between nuclear and mitochondrial genomes and their interactions that alter mitochondrial metabolism.
- Based on research contrasting neutral and adaptive hypotheses, we suggest adaptive evolution in *Fundulus* is common because local populations have significant standing genetic variation and adaptive evolution typically proceeds by polygenic selection.
- If evolutionary adaptation often proceeds by polygenic selection involving small changes in allele frequencies in many genes, as is the case in *Fundulus*, then results from

*Correspondence to dcrawford@rsmas.miami.edu

[1]Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, Florida, USA

[2]Department of Zoology, The University of British Columbia, Vancouver, Canada

[3]Department of Environmental Toxicology, University of California, Davis, California, USA

Published online, April 2020 (*comprehensivephysiology.com*)

DOI:10.1002/cphy.c190004

this system can enhance our general understanding of biochemical, physiological, and evolutionary processes.

Introduction: Evolutionary Physiology

Insights into physiological processes have benefited from comparisons among species in an evolutionary framework (15, 17, 66, 72, 162). Much of evolutionary physiology focuses on comparisons among species, with fewer detailed studies within species (72, 73). The power of studies among species is that they provide the widest range of environments to assay how organisms have evolved physiological traits to thrive in these environments. Yet, studies of populations within species that are exposed to different environments provide a similar perspective allowing an examination of how trait variation relates to Darwinian fitness in nature (17, 184) as well as addressing the genetics of adaptation (9, 10, 184).

In this article of the small estuarine teleost *F. heteroclitus*, we focus primarily on differences among populations within the species with some data among other closely related species to provide phylogenetic context for adaptation (*sensu*) (71, 72). Importantly, in this article, the term "adaptation" strictly refers to the evolution of a trait by natural selection. We do this to avoid the confusion with physiological plasticity (acclimation or acclimatization), where a trait is adjusted in an organism's lifetime by physiological processes (72).

Fundulus heteroclitus as a Model for Studying Physiological Adaptation

In *Spartina* saltmarsh estuaries along the eastern North American coast, the tide rolls in and then drains the estuary through sets of interconnecting tidal creeks teeming with small teleost killifish (*F. heteroclitus)*. In these saltmarsh estuaries, *F. heteroclitus* is a dominant and ecologically important species (101) with biomasses that can exceed a dry mass of 160 kg/ha (190) and with effective population sizes that can exceed 100,000 (60). For vertebrates, this large effective population size makes *F. heteroclitus* an outstanding model system for investigating adaptive evolution in physiological traits and the ways that animals have evolved to a changing environment.

The estuarine environments that *F. heteroclitus* inhabit are highly variable for a wide range of abiotic factors. Diurnal tidal cycles produce dramatic variations in the physical environment (1, 31, 161): temperatures can increase daily to greater than 30°C in upper estuaries or plunge to 12 to 15°C with incoming cold tides (172); salinity can vary from nearly freshwater due to heavy rains to salinities greater than seawater in desiccating ponds (1, 76, 207); oxygen concentrations can vary from anoxic to supersaturated (177). *F. heteroclitus*' ability to mount physiological responses to tolerate these conditions is well documented (31, 44, 48, 139, 161, 168, 199, 207).

In addition to variation within a local environment, habitats only kilometers apart can experience large differences in temperature, salinity, pollution levels, and other environmental parameters. In these environmentally diverse habitats, *F. heteroclitus* populations demonstrate significant variation in molecular, biochemical, and physiological traits (25, 31, 40, 152, 200). This makes *F. heteroclitus* an intriguing model in which to assess the potentially adaptive variation among populations and thus define the functionally important variation in these traits. Importantly, *F. heteroclitus* have several key characteristics that make this species particularly amenable to determining whether the observed variation in physiological traits among populations is adaptively significant. These characteristics include their distribution across environments, high site fidelity and small home range, large effective population sizes, and populations with well-defined population structure and demographics.

Distribution across environments

F. heteroclitus is distributed across a wide range of environmental conditions that are likely to impose a strong selection of physiological traits. One of these conditions is temperature: changes in temperature alter the rate of biochemical and physiological processes and thus impose strong selection to offset these effects (89, 180). The steep thermal cline along the North American east coast from Georgia (USA) to Nova Scotia (Canada) is the most well-known environmental variation for *F. heteroclitus* (43, 141, 148, 161, 200), and populations at the extremes of the species' distribution experience greater than 14°C difference in mean annual temperatures at 1 m depth (148). Additional temperature differences result from power plant thermal effluents (TEs) where thermally impacted local populations are 4 to 12°C warmer than neighboring populations (49). Extreme temperature variation can also be observed within a single estuary where shallow ponds are approximately 4°C warmer on average than deeper nearby basins.

In addition to temperature differences among populations, some *F. heteroclitus* populations experience extremely high levels of anthropogenic pollution (132, 157, 204). Both organic and inorganic pollutants exert strong selection because of their toxicity. Polluted sites are surrounded by relatively clean sites at distances of tens of kilometers or less, and separate polluted sites are located thousands of kilometers away. These separate polluted sites allow populations to independently adapt to the pollution challenges and provide biological replicates to better understand evolutionary adaptation (16, 31).

A third important environmental factor that varies among *F. heteroclitus* populations is salinity (199, 202). Salinity variation also imposes strong selection, having shaped the evolutionary history of aquatic organisms (202). The average salinity and variation in salinity across *F. heteroclitus*' distribution are geographically complex and are often related to freshwater input from rivers and streams.

High site fidelity and small home ranges

F. heteroclitus lay their eggs at the highest tides in the upper tidal regions in mussel shells or among *Spartina alterniflora* leaves. The eggs adhere to the substrate via a number of sticky "egg hairs." These eggs hatch with re-submersion approximately 2 weeks later on the next spring tide (2, 186). The observations of *F. heteroclitus* moving in and out with the tides, combined with a common breeding area within an estuary suggest a panmictic (freely interbreeding) population within a saltmarsh. Yet, *F. heteroclitus* has a home range that is much smaller than a single drainage system and has high site fidelity (5, 113). In a study of 1499 marked fish over 60 days, *F. heteroclitus* exhibited a home range of 36 m with the greatest distance moved being 375 m by just three fish (113). These fish also displayed strong site fidelity [returning to the same side of a creek after release over a 3-month period (113)]. Similarly, in a separate study, 97% of tagged individuals were found within 200 m of initial marking over two seasons (173), and stable isotopes indicate very few *F. heteroclitus* (3.4%) move more than 200 m in their lifetime (174). For the young of the year, 44% were recaptured within 5 m of the initial tagging site between August and December (4). High site fidelity was also supported by a remarkable mark-recapture study of greater than 14,000 individuals over 17 months, where the authors concluded that despite physical connectivity of the watershed within and among different creek systems, fish had almost complete fidelity to a single creek (5). These small home ranges and high site fidelities suggest that local populations may be genetically isolated from each other to some extent, which is likely to promote evolutionary adaptation to local environmental conditions by reducing the in-flow of maladapted alleles into a local population.

Large effective population size

Local *F. heteroclitus* populations are very large. In fact, the effective population size in *F. heteroclitus* may exceed 10^5 within a single creek (6, 60). Large populations are minimally impacted by neutral evolutionary processes and thus large local *Fundulus* populations are less likely to be influenced by random genetic drift, making adaptive evolution more effective. Population geneticists have quantified this effect and have shown that the selection coefficient (where 0 is the absence of selection and 1.0 is 100% death for the deleterious allele) only has to exceed $1/2N_e$ for selection to outweigh the effects of neutral drift. Thus, within large populations, even very weak selection can result in adaptive evolution (80, 170).

Closely related taxa

The characteristics of *F. heteroclitus* described above imply that this species is made up of numerous, somewhat interconnected, populations that are exposed to different environments and that are likely to have undergone local adaptation to their environmental conditions. Their habitat and distribution result in a mosaic of closely related taxa that have independent demographics that enhance evolutionary analyses. Evolutionary analyses are more readily accomplished among closely related taxa (populations within a species, or closely related species) than among distantly related taxa because it is simpler to define ancestral traits and distinguish random neutral changes from selectively advantageous ones. More specifically, isolated taxa will evolve by neutral processes, and the rate and scale of change are dependent on the size of the population and the divergence time (200). This means that distantly related species isolated by long time scales are likely to exhibit large genetic, biochemical, or physiological differences that could just as easily evolve by either neutral or adaptive evolution. Among closely related taxa, such as populations within a species, demographic effects are small, and thus few neutral changes evolve to obscure selectively advantageous ones. Thus, statistically rejecting the null hypothesis of neutral evolution is more likely among closely related taxa than among distantly related taxa because closely related taxa have fewer and smaller neutral divergences (40, 41, 71, 200). This makes evolutionary analyses easier among closely related taxa than among distantly related taxa, especially those where speciation involved a transitory small population (bottleneck).

Demography in *Fundulus heteroclitus*

Large local populations of *F. heteroclitus* subjected to biologically important environmental variation provide a valuable resource to investigate the physiological and genetic mechanisms responsible for evolutionary adaptation. However, all evolutionary analyses must consider demographic parameters that can cause evolutionary changes by neutral processes. That is, it is necessary to reject demographic effects, so we do not misinterpret neutral evolved differences as adaptive evolution (69, 81, 98, 99, 130). Demography, like species' historical relationships, influences evolutionary changes such that more closely related populations will tend to share evolved differences, both adaptive and neutral. Additionally, rapid reduction in population size or isolation of local populations can enhance neutral divergence. Thus, when assessing evolutionary adaptive change, demography can be a nuisance parameter: a factor that interferes with or obfuscates analyses that attempt to identify the effect of natural selection. For *F. heteroclitus,* neutral processes should be minimal because large populations have little random drift. Yet, *F. heteroclitus*' extensive geographic distribution across long time scales alters gene flow and population size, and thus neutral divergence contributes to population differentiation across their distribution (60, 175, 176) (Figure 1).

Although population sizes are large and evolutionarily stable, neutral divergence or adaptive divergence is influenced by an important historical break between the northern and southern *F. heteroclitus populations* (70). *F. heteroclitus*' genomes have extensive polymorphisms dominated by a sharp break (large change in allele frequencies) between

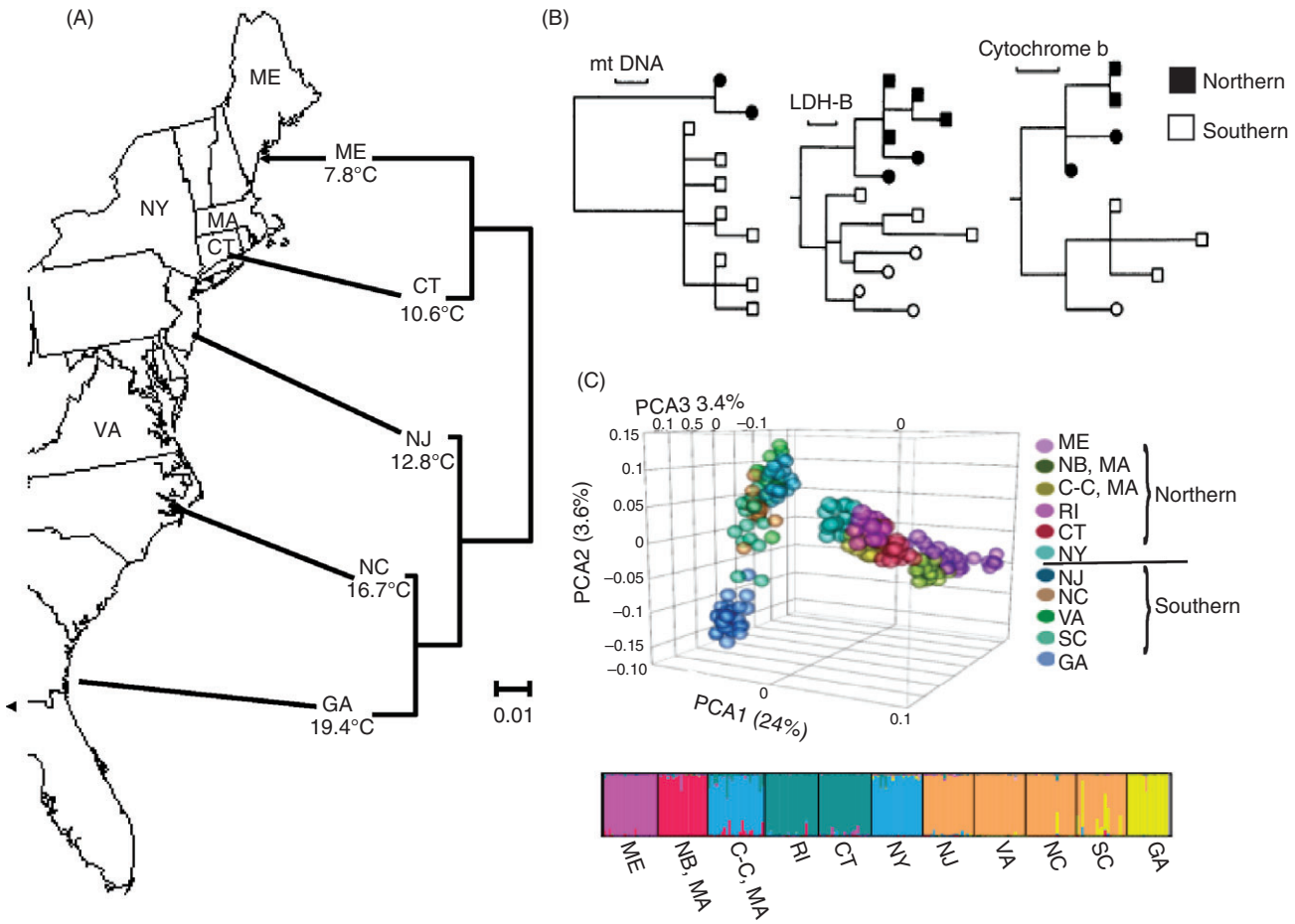


Figure 1 Demographic patterns among *F. heteroclitus* populations. (A) Neighbor-joining tree based on microsatellites (6, 60, 200). (B) Maximum parsimony tree for RFLP in mitochondria, LDH-B coding region, and 309 bp of Cytochrome B (19). Dark circles and squares are northern populations (north of Hudson River), unfilled are southern populations (south of Hudson River). (C) PCA and population structure ($k = 9$) based on 354 SNP, 30 individuals per population and principal component analysis (203, 204).

populations on both sides of the Hudson River associated with the southern extent of the Laurentide Ice Sheet during the last glaciation approximately 20,000 years ago (19, 74, 79, 148, 178). During this last glaciation, sea levels were approximately 150 m shallower, extending estuaries into the continental shelf. Yet the Hudson River potentially formed a migration barrier as it was a torrent due to the draining glacial ice melt (21, 22, 39, 56, 92, 102, 119, 123, 185). After the last glaciation, the admixture of northern and southern populations produced a peak of allozyme heterozygosity south of the Hudson near Atlantic City (148, 152). Combining mitochondrial sequences and microsatellite data, the best explanation for the patterns of DNA sequence variation is that a large *F. heteroclitus* refugium existed north of the Hudson River, and individuals from this refugium reinvaded locations just north of the Hudson (6, 79). These historical events most likely enhanced the evolutionary divergences among *F. heteroclitus* populations (6, 60, 148, 152, 179) (Figure 1).

The historical break between northern and southern populations would enhance any diverging evolutionary forces (neutral or adaptive) by minimizing migration (70). For example, it is clear that the evolved divergence among populations is related to the demographic isolation of northern and southern populations as seen among neutral microsatellites (Figure 1A). While the microsatellite data provides evidence to support the expected historical break (6, 60), the pattern of microsatellite variation is more gradual than if northern and southern populations remained isolated, indicating migration across previous boundaries. The evolved divergence between northern and southern populations is also seen among mitochondrial RFLPs (restriction fragment length polymorphisms), the nuclear gene lactate dehydrogenase B (heart-type, LDH-B), and the mitochondrial cytochrome b gene (19, 74) (Figure 1B). Additionally, 3-principal components utilizing 354 nuclear single nucleotide polymorphisms (SNPs) show a clear population structure (203) (Figure 1C). These SNPs are shared among both northern and southern populations, yet their frequency provides readily identifiable population structure within clades north or south of the Hudson River and between these two clades.

Demographic histories of populations are an important consideration in evolutionary physiology because they make

it more difficult to distinguish adaptive versus neutral divergence between populations. For example, the LDH-B pattern of allelic divergence (Figure 1B) could be due to natural selection, yet it is indistinguishable from neutral microsatellite patterns (Figures 1A and 1B). This suggests that simple patterns of genetic variation in DNA sequence alone are insufficient to distinguish adaptive from neutral divergence. This article of *Fundulus*' adaptation describes evolutionary studies that are enhanced by physiological measures, linking functional consequences of nucleotide divergence that influence fitness in traits from cellular metabolism up to organismal functions. We suggest that the combination of physiological data with statistical analyses of nucleotide divergence more strongly defines adaptive evolution than studies using only nucleotide divergence or physiology separately.

Adaptive Evolution

To appreciate adaptation in *F. heteroclitus*, it is important to define adaptation and how polygenic traits evolve by natural selection. Adaptation is the evolution by natural selection of biological functions that improve the performance, health, longevity, or reproduction of an organism. That is, adaptation is the evolution of genetic changes that alter phenotypic traits that improve Darwinian fitness. Investigating and determining which biological traits have evolved by natural selection and thus are adaptive is critical because it informs us that the variations in these traits are functionally important. However, it is difficult to demonstrate that variation in phenotypic traits among populations, species, or other phylogenetic groups is adaptive, as observed differences may arise due to random, neutral processes in addition to arising by natural selection. Despite the temptation to assume that all observed trait differences are adaptive, it is critical to test and reject the null hypothesis of neutral divergence prior to concluding that variation in a trait represents an adaptive change.

To define traits as adaptive, it is necessary to demonstrate that they are derived, heritable, and evolving by natural selection (vs. evolving by neutral processes) (10, 17, 72, 184). This is particularly powerful if coupled with a direct demonstration of the genetic basis of trait variation and identification of how this variation alters organismal performance. Although it is desirable to identify the specific genetic variants that are responsible for the variation in physiological traits, it can be particularly difficult to do so because physiological traits are generally quantitative traits with continuous phenotypic variation (e.g. body mass) rather than discrete traits, which have only a small number of phenotypic states (e.g. alternative color morphs). Often, variation in discrete traits is the result of variation in only one or a few genes, each with large phenotypic effects. In contrast, continuous phenotypic traits are typically polygenic—dependent upon many genes (tens to thousands), where each gene has a small phenotypic impact (18, 86, 153, 194). Selection of polygenic traits often involves only small changes in allele frequencies at many polymorphic loci (170). Thus, identifying the genetic basis for the variation in polygenic traits is a challenge because changes in allele frequencies and allele effect sizes are small and thus statistically difficult to detect (8, 18, 87, 194, 208). An additional complexity is that the genes influencing adaptive variation of polygenic traits (polygenic adaptation) need not be the same in all populations within a species or in all individuals within a population.

Individuals that share the same adaptive phenotype may not share the same set of underlying adaptive genetic variants. This is because many different combinations of genetic variants can produce the same adaptive phenotype (209). For example, metabolic flux through a pathway can be altered by many of the polymorphic enzymes in a pathway; yet, adaptation may only require change in a few of the many variable enzymes to enhance metabolic flux. The reason why multiple solutions from polygenic adaptation may evolve in a population is that in a population with sufficient standing genetic variation there are more variable genes that influence an adaptive phenotype than are needed for adaptation: $n_{tot} \gg n_{opt}$, where n_{tot} is the number of genes that can influence an adaptive phenotype and n_{opt} is the number required. (209). These functionally important variable genes could occur in different pathways or unrelated gene complexes, and each of these genes could have more than one functional variant (>2 alleles). When many different genes have a favorable allele and only a subset is needed to create the same phenotypic change (i.e. they are redundant), natural selection need only change the frequencies of some of the functionally important alleles to achieve an adaptive phenotype. As a result, with many hundreds of functionally equivalent variable genes, selection for the adaptive phenotype could be accomplished by changing the allele frequencies at some but not necessarily all functionally important genes because there are many variable genes that enhance the adaptive trait similarly (209). Thus, with sufficient standing genetic variation at the outset, polygenic adaptation will result in adaptive evolution of a trait influenced by many (tens to thousands) of variable genes of small effect, where many of the different genes are redundant (cause the same phenotypic advantage) (18, 86, 153, 189). It is this functional redundancy that allows individuals to have the same adaptive phenotype, without the population becoming fixed for any single adaptive locus.

The consequences of polygenic adaptation that occurs through standing genetic variation are that many members in natural populations will have the same adaptive phenotype; yet, not all adapted individuals will share the same combination of underlying adaptive alleles. Another predicted consequence of polygenic adaptation is that selection in different populations will cause adaptive changes in the frequencies of different sets of genes; though populations may converge on the same adaptive phenotype, it will be likely be underlain by frequency changes in different loci in different populations. Thus, the genotypes that alter an adaptive phenotype are not necessarily shared among all adapted individuals because many different allelic combinations,

across multiple loci, can yield the same adaptive phenotype. This concept of polygenic adaptation from standing genetic variation is important because it changes our expectation of easily detecting a gene, or set of genes, responsible for adaptive change in a physiological trait (153, 155, 170, 209).

Defining the genetic basis for polygenic adaptation is difficult because the analyses require being able to statistically identify many genes, each with small phenotypic effect, where there is more than one functionally important gene that can alter the adaptive phenotype (158). Furthermore, studies on inbred strains or populations with limited genetic diversity will have inconsistent results when trying to identify the genes underlying polygenic adaptation. We can address these difficulties in identifying adaptive variation by integrating functional and evolutionary analyses among populations subjected to ecologically and evolutionary important environmental variation. This approach identifies environmental factors that influence an organism's success and defines adaptive physiological function that enhances Darwinian fitness when organisms are stressed by these environmental factors. We then can examine natural populations adapted to these different environments to enhance our ability to identify the genetic basis for the adaptive variation in physiological function—populations that are physiologically adapted to their local environments should be enriched for the genetic changes that result in polygenic adaptation. It is the enrichment of polygenic adaptive genes among populations that makes this approach successful. This is the evolutionary approach taken in the examples of *Fundulus* evolutionary adaptations provided in this article.

Adaptation in *Fundulus heteroclitus*

The prescription for identifying adaptation is defining populations subjected to different environments with potential evolutionary impact, defining how variation in physiological traits allows organisms to cope with these environmental differences and exploring the genetic basis for the variable phenotypes. For *F. heteroclitus*, the habitats with different environments are hundreds of meters to less than 10 km apart. The physiological traits include metabolism, swimming speed, development, osmotic regulation, pollutant resistance, and thermal performance, among others. The genetic basis for the variation in these physiological traits includes a single gene or hundreds of independent genetic changes. We start our description of adaptation in *F. heteroclitus* with the oldest studies on the importance of a single gene (LDH-B) along the entire eastern seacoast of North America and end with a genetic analysis that suggests that evolutionary adaptation is rampant as illustrated by the many genetic changes among microhabitats within a single population. The evolved differences due to LDH-B are a classic example of a single gene causing adaptive phenotypes and thus serves as a contrast to the many other examples of polygenic adaptation in this species.

Adaptive variation in LDH-B along the thermal cline

The most renowned and well-characterized study of adaptation in *F. heteroclitus* involves the enzyme lactate dehydrogenase-B (*LDH-B)*, which is the "heart-type" isoform of this enzyme. In *F. heteroclitus*, many genes encoding enzymatic proteins have two or more alleles and populations of *F. heteroclitus* along the steep thermal cline of the North American Atlantic coast have genetic divergences among these allelic enzymes (allozymes, Figure 2) (32, 137, 143, 148, 152, 159, 191). The pattern of allozyme variation along the Atlantic thermal cline is complex (150, 152, 159). Some allozymes have a steep frequency change at the Hudson River, with complete fixation (100%) of one allele in the north and the alternative allele in the south [e.g. MDH-A, (159) Figure 1A] while others have smaller allele frequency changes [e.g. IDH-B, Figure 1A, 6PGDH-A, Ap-A, Est-B, (152, 159)]. LDH-B has a relatively gradual transition from nearly fixed for the northern allele (LDH-B^b) in Maine to nearly fixed for the southern allele (LDH-B^a) in Georgia (149, 152).

LDH-B is an exceptionally well-studied allozyme in *F. heteroclitus* (148). Mitton and Koehn (124) showed that a *F. heteroclitus* population resident in locally heated estuaries in New York (near TE from power plants) had a greater frequency of the LDH-B southern allele type compared to other northern populations not exposed to TE. This local change in allele frequencies suggested that natural selection alters the frequencies of LDH-B allozymes in response to temperature.

Functional studies for both LDH-B alleles initially examined enzyme kinetics across a wide range of temperatures and pH levels (144, 145) (Figures 2B and 2C). Enzymes catalyze the conversion of substrates to products and are involved in metabolic functions that influence energy production and substrate usage. The catalytic rate is dependent on kinetic constants (e.g. K_m and k_{cat}) and the amount of enzyme. LDH-B enzyme kinetics were determined using purified enzymes from a single population, which had both LDH-B alleles. The kinetic parameter k_{cat}/K_m is related to the reaction rate: with equal concentration of enzyme and substrate, the allelic enzyme with higher k_{cat}/K_m will have a faster reaction. Figures 2B and 2C highlight that (i) at lower temperatures the northern LDH-B (b/b) genotype has a greater reaction rate than the southern genotype (a/a), (ii) this difference disappears at warmer temperatures, and (iii) the differential response to temperature among the three LDH-B genotypes is pH dependent (Figure 2C). At pH 7 and below, reactions for the southern homozygote (a/a) and the heterozygote (a/b) are faster at higher temperatures than the northern homozygote (b/b). However, relative to the southern homozygote, the northern homozygote (b/b) has a higher reaction rate at lower temperatures. These differences in enzyme reaction rates suggest that the two amino acids that are distinct between the northern (LDH-B^b) and southern (LDH-B^a) allozymes (106, 148) are functionally significant.

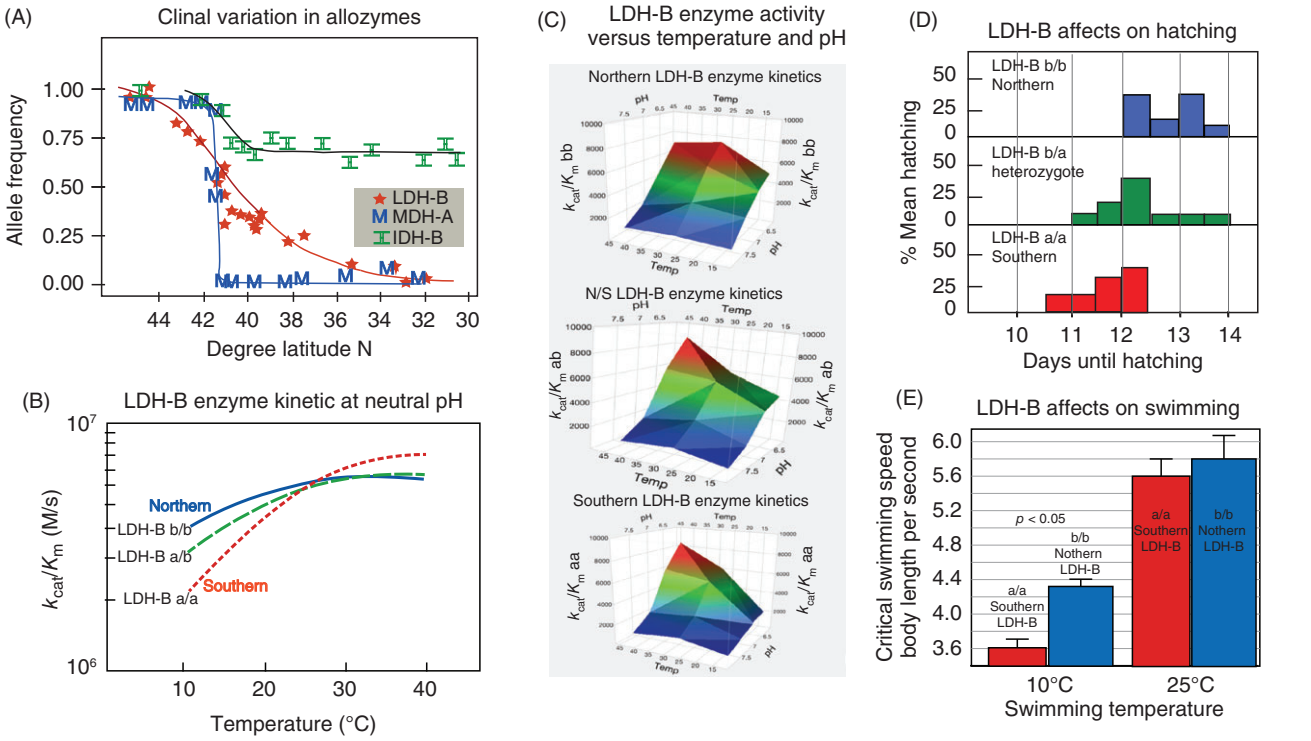


Figure 2 Allozymes in *F. heteroclitus*. (A) Allelic variations of protein enzymes (allozymes) for three different genes: LDH-B*, Malate dehydrogenase A (M, MDH-A), and Isocitrate dehydrogenase-B (I, IDH-B) (redrawn based on data in Refs 148, 159). The frequencies of the northern type alleles are plotted versus latitude along the eastern seacoast of North America (degrees latitude N). (B) Enzyme kinetics for the three genotypes of LDH-B [catalytic turn-over (k_{cat}) divided by the K_m] measured at different temperatures at neutral pH ([OH] = [H]) (redrawn based on data in Ref. 144). (C) Enzyme kinetics (k_{cat}/K_m) for the three LDH-B genotypes plotted against temperature and pH (redrawn based on data in Ref. 144). (D) Hatching time at 20°C for fish from single populations for the three LDH-B genotypes (redrawn based on data in Ref. 51). Hatching times were defined among 20 randomly crossed pairs, and larvae were genotyped. Data represent larvae genotypes. Similar results were obtained by 4 replicate mass crosses with 40 male and 40 female heterozygotes in each cross ($n > 1000$/cross). (E) Critical swimming speed (maximum sustainable swimming speeds in body lengths per second) for the two LDH-B homozygotes; all fish were from the same population (drawn from data in Ref. 52).

Differences in enzyme catalysis among populations, as is observed for LDH-B, do not allow us to conclude that the clinal allelic variation is adaptive because the difference in enzyme catalysis may not enhance fitness or other important physiological processes (148). Demonstration of adaptation requires evidence that the variations in LDH-B catalysis change organismal functions and that these variations enhance fitness. To determine whether variation in enzyme catalysis impacts higher levels of organismal function, DiMichele and colleagues measured hatching rates (51), swimming speeds (52), developmental metabolic rates (53, 138), and survival (54) as related to LDH-B genotypes. Hatching time (Figure 2D), a measure of developmental rates, is longer for the northern LDH-B homozygote than for either the heterozygote or the southern LDH-B homozygotes. While genotypes are referred to as "northern" and "southern," the fish used in these experiments were from the same population. This is important because other genes that could alter hatching should be randomly associated with 20 random crosses. Thus, the most parsimonious explanation is that the LDH-B genotype influences hatching rates (Figure 2D). These data are supported by four replicate mass crosses (40 males and 40 females per cross) between LDH-B heterozygotes. With 1000s of individuals assayed per cross, hatching rates were dependent on the LDH-B genotype (51).

The variation in hatching rate reflects differences in developmental metabolic rates (53, 138). Metabolism was specifically measured among different LDH-B genotypes. The northern LDH-B genotype exhibited lower metabolism over the course of development. Similar to the hatching experiment (above), the LDH-B genotype appears to be responsible for the difference in developmental rates because metabolism was measured with a randomized genetic background in offspring created from experimental crosses. In addition, to experimentally test the specific role of LDH-B, purified LDH-B enzyme was injected into eggs resulting in metabolic rates, which were dependent on which LDH-B was injected: higher metabolic rates were measured if the southern LDH-B was injected than if the northern LDH-B was injected regardless of the original eggs' genotype (50). This clearly demonstrates that the difference in LDH-B alleles causes significant physiological differences that change development. Despite this, phenotypic differences among LDH-B genotypes could be classified as adaptive only if

this developmental variation was proven to alter individual fitness.

Further evidence for the physiological importance of the allelic variation at the LDH-B gene is its association with critical swimming speed where the northern genotypes swim faster at the lower temperature associated with these populations [Figure 2E, (52)]. Here, data on swimming speeds reflect the difference in LDH-B reaction rates at different temperatures (Figure 2B): with critical swimming speed measured at 10°C, individuals with the northern LDH-B genotype were able to sustain a faster maximum swimming speed than those with the southern LDH-B genotype. These fish were collected from a single population and thus there should be a random genetic background suggesting that the LDH-B genotype was responsible for the swimming differences. The swimming differences are most likely related to LDH-B's allozymes altering ATP production and hemoglobin oxygen delivery in red blood cells. Individuals with the northern LDH-B genotype have greater ATP production, which changes oxygen binding in fish red blood cells (52). Specifically, the northern LDH-B genotypes with higher ATP concentrations display a pronounced Root-effect, which causes hemoglobin to release oxygen to tissues more readily at higher oxygen concentrations at the cost of lower total oxygen binding.

Quantitative difference in the amount of LDH-B and glycolytic enzymes

While studies by DiMichele, Place, and Powers focused on the variation in LDH-B enzyme kinetic parameters, others have demonstrated that the quantity of LDH-B also varies among populations, such that colder North Atlantic coast populations had a greater amount of LDH-B (40, 42, 134, 168). Additionally, this variation in enzyme amount is physiologically flexible. Specifically, when individuals are acclimated to different temperatures, LDH-B concentrations change (44, 168). These differences are tissue specific: in livers, but not cardiac tissue, the amount of LDH-B has an evolved and physiological acclimation difference (41, 140, 142, 168). What is apparent from these studies is that the difference in the amount of LDH-B between northern and southern *F. heteroclitus* is due to both intrinsic (evolved) and physiologically reversible changes. Because the difference occurs in individuals after long-term acclimation, it is considered to be an evolved, heritable difference between populations; however, irreversible physiological responses (e.g. irreversible developmental or epigenetic effects) cannot be ruled out.

In addition to the intrinsic difference between populations, individuals reversibly increase the amount of LDH-B when acclimated to colder temperatures for several weeks. Thus, *F. heteroclitus* in northern populations have greater LDH-B reaction rates due to the combination of enhanced enzyme kinetics, greater adaptive differences in enzyme quantities, and physiological induction in response to colder temperatures at ecologically relevant pH levels (43). These data indicate that individuals living in colder northern waters have essentially equivalent LDH-B activity as individuals living in warmer southern waters even though these populations have approximately 12°C difference in environmental temperatures. This combination of enzyme kinetics, protein expression, and reversible physiological induction highlights the diversity of mechanisms that enhance performance in different environments (43).

These functional physiological data that demonstrate an association between LDH-B genotype and performance strongly suggest that natural selection is responsible for the functional divergence between the northern and southern LDH-B alleles such that the LDH-B reaction rates are essentially the same between the populations inhabiting naturally occurring colder northern and warmer southern locations. However, the original evolutionary forces underlying LDH-B allelic variation could still be related to neutral processes. That is, even though LDH-B has significant functional effects, the increase in the frequency of the northern LDH-B allele could have been due to isolation and neutral demographic processes. However, two additional data sets allow us to reject this null hypothesis of neutral evolution.

Did LDH-B evolve by natural selection due to temperature? To address this and inquire about the relative importance of LDH-B versus other enzymes, the quantitative variation for all glycolytic enzymes among many *Fundulus* taxa was determined (Figure 3A). This study (141) quantified the glycolytic enzyme activities in 15 *Fundulus* taxa (2 populations in each of 7 species and 1 outgroup population) that inhabit different thermal environments to define significant changes that were consistent among taxa in similar thermal environments. The neutral expectation is that phylogenetically more closely related taxa should be more similar than genetically more distant taxa (72). Alternatively, if changes in glycolytic enzymes were evolving by natural selection, the enzyme activities would be independent or unrelated to phylogeny and instead be related to the local thermal environment. Two approaches were used to remove the effect of phylogeny (similarity between species due to closer evolutionary relationships) and thus inquire if thermal environment and glycolytic activities are significant related (Figure 3B). This study showed that independent of phylogeny, taxa inhabiting lower seawater temperatures had greater amounts of LDH-B, GAPDH (glyceraldehyde phosphate dehydrogenase), and PYK (pyruvate kinase) (141) (Figure 3B). Finding this pattern in LDH-B and the other two enzymes is most parsimoniously explained by natural selection favoring the increase in these enzymes in lower temperature environments. Thus, these data indicate the variation in LDH-B enzyme activity among many different *Fundulus* species is indicative of thermal adaptation (141).

The phylogenetic analysis of enzyme amounts predicted that these adaptive divergences in LDH-B, GAPDH, and PYK would have to alter a biologically important trait (Figure 3B). As a test of this hypothesis, variation in cardiac metabolism was measured within and among northern and

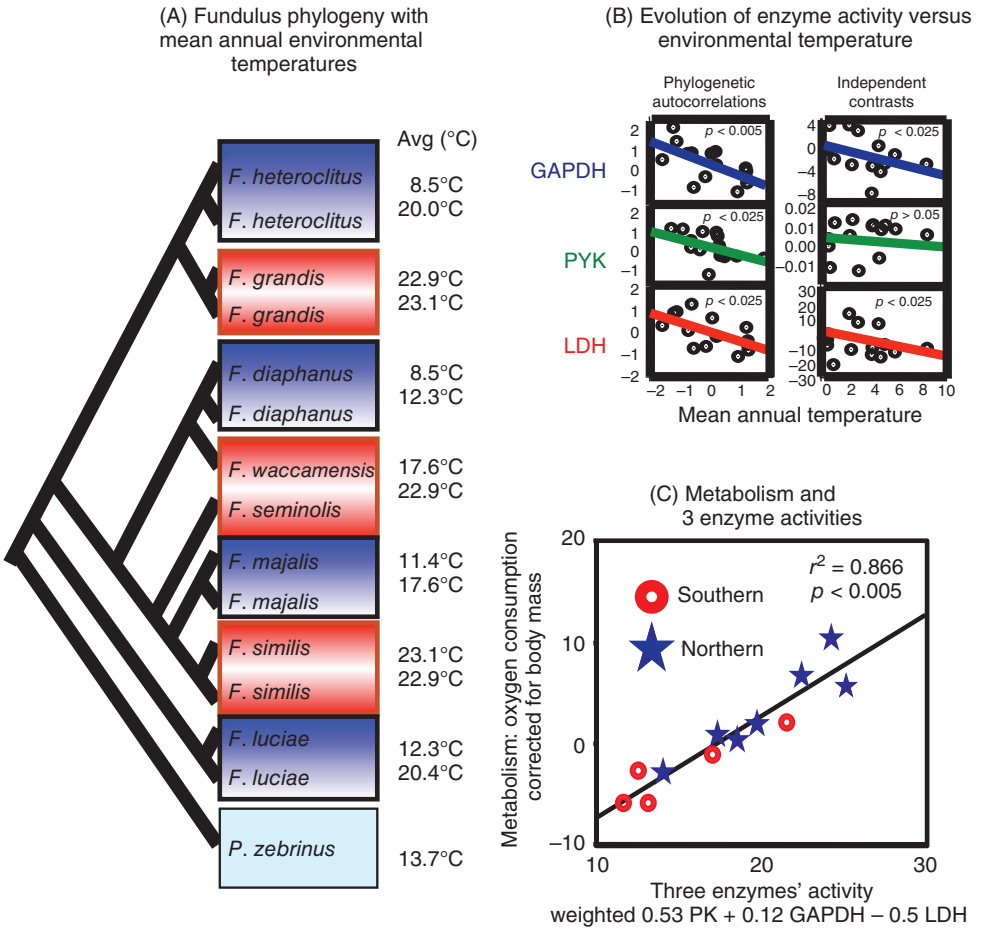


Figure 3 Phylogenetic analyses of enzyme amounts. (A) *Fundulus* phylogeny (141). There are 15 taxa: 2 populations from 7 species and a single population from the outgroup. Boxes represent taxa with similar environmental temperature variation: Blue–geographic variation in temperature with northern taxa being colder. Red–lack of geographic variation in temperature. (B) Two phylogenetic methods for correcting for species similarities in enzyme amount among 15 *Fundulus* taxa versus naturally occurring mean annual environmental temperatures. Only the three enzymes were significantly related to environmental temperature after correcting for phylogeny: glyceraldehyde-3-phosphate dehydrogenase (GAPDH), pyruvate kinase (PYK), and LDH-B. (C) *F. heteroclitus* glucose-dependent metabolism versus multiple factor equation using three phylogenetically important enzymes (GAPDH, PYK, and LDH-B). $r^2 = 0.866$ ($p < 0.005$) (146).

southern *F. heteroclitus* populations (146). Northern populations have a greater glucose-dependent cardiac metabolism than do southern populations. Examining the variation in all the glycolytic enzymes among these individuals indicated that the enzymes identified in the phylogenetic study (141) (LDH-B, GAPDH, and PYK) also explained the variation in *F. heteroclitus* cardiac metabolism (Figure 3C) (146). Thus, both a phylogenetic study among species and metabolic physiological study within species suggest that changes in LDH-B enzyme are adaptively important (Figure 3).

Evolutionary analyses on LDH-B DNA sequence variation

Another type of data indicating that variation in LDH-B is adaptively important is from the analysis of LDH-B DNA sequence variation. Specifically, if the variation in LDH-B activity is due to evolution by natural selection, then patterns of DNA sequence variation should support this hypothesis. Sequence comparisons (19, 106) reveal two amino acid substitutions that distinguish the two LDH-B alleles. One of these polymorphisms is located in exon 7, encoding amino acid position 311 (northern serine, southern alanine). The other polymorphism is located in exon 4, encoding amino acid position 185 (northern alanine and southern asparagine). This 4th exon amino acid substitution was found to be responsible for the difference in the thermal stability of the two LDH-B allelic enzymes (106). One way to determine if natural selection is responsible for the differences in DNA sequences between the northern and southern LDH-B alleles is to apply the McDonald-Kreitman neutrality test (118). This test uses a 2 × 2 contingency table, comparing the nonsynonymous

Table 1 LDH-B McDonald-Kreitman test

	Shared	Fixed
A: All polymorphic sites		
Syn	13	0
Non-Syn	1	2
		$p = 0.025$
B: Polymorphic sites found >1 individual		
Syn	8	0
Non-Syn	0	2
		$p = 0.022$

Synonymous (Syn) variation does not alter amino acids, and nonsynonymous (Non-Syn) codes for different amino acids.

(changes in genetic code that result in a different amino acid) to synonymous substitutions for polymorphic sites shared between taxa versus fixed differences between taxa. If neutral evolutionary mechanisms are responsible for amino acid polymorphisms, then the ratio of fixed to polymorphic sites should be similar for both synonymous and nonsynonymous sites. Typically, this test is used to examine variation within and between species (197); yet, the properties of this test hold if one compares two populations in which shared polymorphisms and fixed differences are not confounded (46). Previously published (19) LDH-B cDNA sequences from 11 individuals [from 2 northern populations (Nova Scotia and Maine) and 2 southern populations (Georgia and Florida)], revealed 3 nonsynonymous and 13 synonymous substitutions between northern and southern fish. All of the synonymous but only 1 of 3 nonsynonymous substitutions are shared among northern and southern *F. heteroclitus* (Table 1A). This is improbable ($p = 0.025$; Fisher's exact test) and thus violates the null hypothesis of neutral evolutionary models. If one only examines nucleotide changes shared by more than one individual, all eight synonymous polymorphic sites are shared among taxa and both nonsynonymous changes are fixed differences (Table 1B). This too is improbable ($p = 0.022$, Fisher's exact test). These data allow us to reject the null-hypothesis that the *Fundulus* LDH-B locus has evolved by neutral processes and provide evidence that it has evolved by natural selection.

To summarize the evidence for adaptive divergence in LDH-B along the North American Atlantic seacoast

- The clinal variation in LDH-B alleles is related to differences in enzyme catalysis, which is temperature and pH dependent.
- The variation in enzyme catalysis is functionally important because it alters hatching times, physiology, and critical swimming speeds.
- When tested at their native temperatures, the combination of the differences in enzyme amount and enzyme catalysis allows the predominant genotype in the cold northern populations to have equivalent enzyme activities as the genotype in the warm southern populations even though these populations experience approximately 12°C difference in environmental temperatures.
- Similarly, across many species, the amount of LDH-B that is associated with glucose-dependent cardiac metabolism is temperature dependent.
- Analysis of nonsynonymous and synonymous DNA sequence variation supports the concept that LDH-B is evolving by natural selection.

These data show an evolutionarily derived change in enzyme kinetics and concentrations as well as DNA sequence variation that is evolutionarily improbable unless they are evolving by natural selection.

Adaptation in mRNA expression along the thermal cline

Adaptive variation in LDH-B mRNA and transcription

It is clear that LDH-B enzyme concentration is important and may be more important than differences in amino acids that are responsible for the variation in the enzyme's kinetic constants (e.g. K_m and k_{cat}) (40, 43, 141, 142, 168). For LDH-B, the amount of LDH-B mRNA and its protein are significantly correlated (Figure 4A). Additionally, mRNA concentration and thus LDH-B enzyme concentration is dependent on the transcription rate (45). Thus, because of the significant difference in LDH-B expression between populations is regulated by the transcription rate, the molecular mechanisms that control transcription should also be different. The transcriptional mechanisms potentially responsible for the variation in LDH-B transcription rates include the LDH-B promoter sequence and the protein transcription factors that bind and regulate transcription. Much of the difference in LDH-B expression is due to two promoter regions: the proximal promoter and upstream hormonal responsive element (46, 163, 167, 169).

The LDH-B proximal promoter (hundreds of base-pairs upstream from the start of transcription) has surprisingly high DNA sequence variation between populations at the transcription factor binding sites but much less variation at nonbinding sites (Figure 4B) (46, 167, 169). It is unusual to have this magnitude of variation at transcriptional binding sites among populations because they should be functionally constrained unless they evolved by natural selection. An insight into this pattern is that the promoter sequence

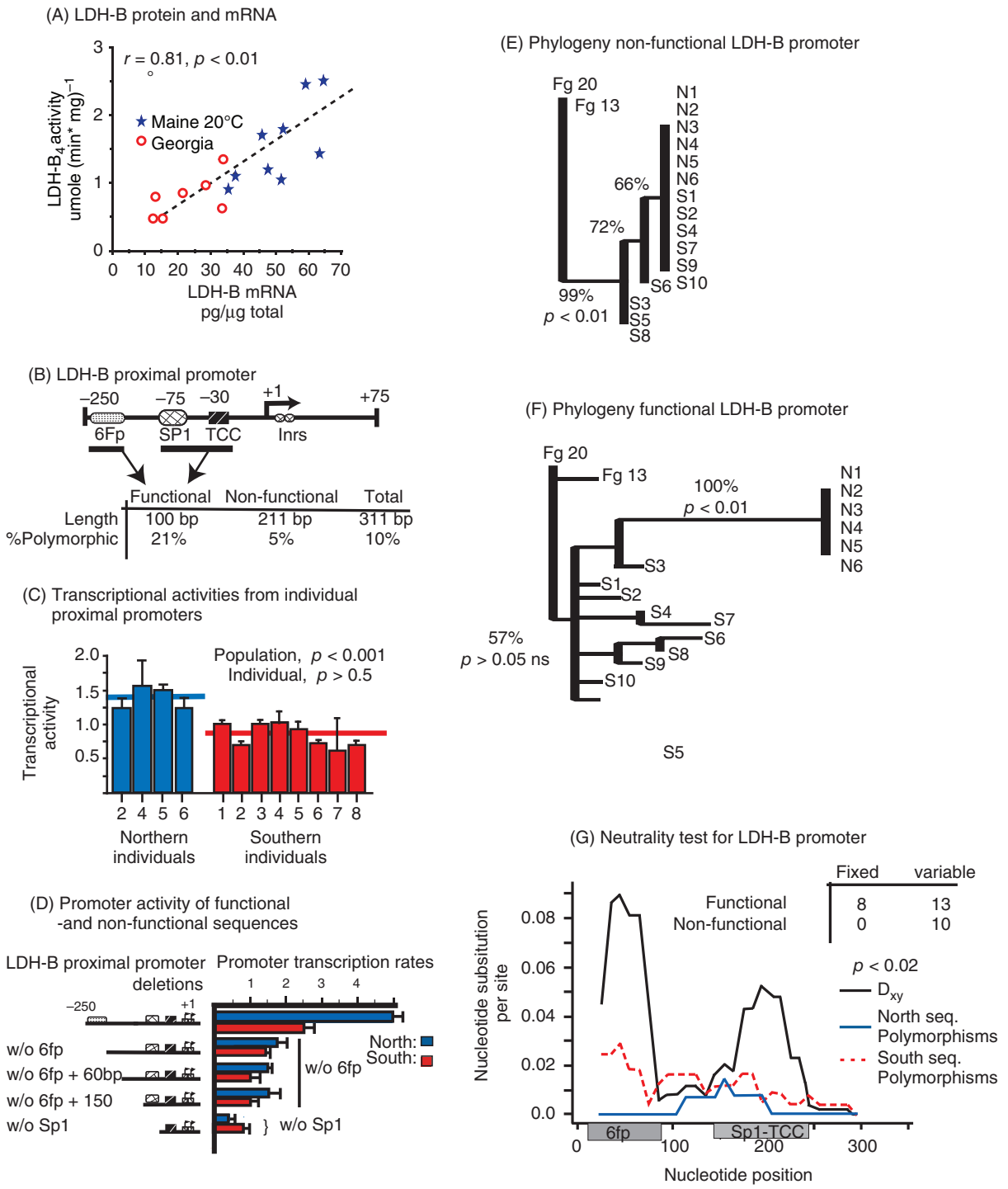


Figure 4 Adaptive evolution of LDH-B proximal promoter. (A) LDH-B mRNA versus protein for northern (Maine) and southern (Georgia) individuals. When fish are acclimated to 20°C, increasing amounts of mRNA are associated with larger amounts of LDH-B protein (measured as maximal enzyme activity ($r^2 = 0.81$, $p < 0.01$, data from Ref. 168). (B) LDH-B transcriptional binding sites and sequence variation. Functional sites are DNA sequences that bind protein transcriptional factors or affect transcription. (C) Individual promoter activity (line extending above columns are standard errors) defined by linking LDH-B proximal promoter to luciferase reporter gene and transfected into rainbow trout liver cell line (46). Promoter activities from northern individuals are significantly greater than promoter activities from southern individuals ($p < 0.001$). (D) Promoter activity with different proximal promoter elements. Binding site 6fp (but not intervening sequence), and SP1 reduce expression, and without SP1, the northern promoter activity is no longer greater than southern promoter activity. (E) Evolutionary relationship using nonfunctional sequences among *Fundulus* species and within *F. heteroclitus*. (F) Evolutionary relationship using functional DNA sequences (affect promoter activity). Northern *F. heteroclitus* functional sequences are derived and significantly different from southern and *F. grandis* promoter DNA sequences. (G) Sliding window of DNA sequence variation within and between northern and southern *F. heteroclitus*. Data derived from Refs 46, 169.

variation between populations, but not within a population, is related to the variation in mRNA expression when assayed in cell culture (Figure 4C) (46).

The functional importance of the LDH-B proximal promoter binding sites was defined experimentally by *in vivo* and *in vitro* DNA-binding assays, promoter assays in cell culture (Figures 4C and 4D), and molecular analysis of SP1 transcription factor (46, 167, 169). An *in vivo* DNA-binding assay identified DNA sequences bound to proteins in *F. heteroclitus* nuclei (169). An *in vitro* DNA-binding assay confirmed these *in vivo* studies and identified Sp1 as one of the proteins regulating transcription (46, 167). One of the consequences of the LDH-B DNA sequence variation is that it alters the binding affinity of the Sp1 transcription factor (167). Additionally, transfection of the entire LDH-B promoter or a part of the promoter into cell cultures indicates that several transcriptional binding sites are responsible for differences in mRNA expression (Figure 4D). Two of the promoter sequences (6fp and SP1) contribute to differences in promoter function when transfected into rainbow trout liver cell lines (46, 167). While cell culture results are dependent on the cell type (46), they all define similar functionally important sequences. Knowing the functional binding site enhances evolutionary analyses of LDH-B DNA sequence variation because we can contrast the patterns in functional (transcriptional binding) sites with nonfunctional sites. For example, in contrast to nonfunctional sites in the LDH-B promoter, the functional transcriptional binding sites are derived (Figures 4E and 4F). That is, the functional binding sites are unique to the northern *F. heteroclitus* in that they are different from both southern *F. heteroclitus* and *F. grandis* (sister species) (46, 167, 169). Thus, DNA sequence variance for a functional site is greater within a species than between species and is indicative of adaptive evolution (Figure 4G). This statistically significant difference in sequence variation at functional sites between populations rejects neutral evolution ($p < 0.02$, Figure 4G). Thus, the most parsimonious explanation for these patterns of DNA sequence variation is that the LDH-B proximal promoter has evolved by natural selection; when combined with the data above on LDH-B enzyme activity, the changes are adaptively important (46).

In addition to variation at the LDH-B proximal promoter, DNA sequence variation in another regulatory site is associated with changes in LDH-B expression (151, 163, 164). Specifically, further upstream from the proximal promoter is a steroid hormone (glucocorticoid) responsive element (GRE). This DNA sequence binds the hormone-receptor complex altering mRNA expression in southern but not northern populations. Only the southern GRE represses mRNA expression in *F. heteroclitus*, indicating that this DNA sequence is important for the adaptive divergence in LDH-B expression (163). The existence of sequence variation in the proximal promoter and GRE coupled with functional differences between northern and southern populations suggests that the adaptive regulation of LDH-B mRNA involves both regulatory elements: one enhances the probability of forming a transcriptional complex and the other increases its longevity.

Adaptive variation in genome-wide patterns of mRNA expression

At the beginning of the 21st century, microarrays made it possible to simultaneously measure the level of expression for thousands of different mRNAs that code for different proteins (30, 61). Microarrays are hundreds to thousands of 100 to 200 µm spots of DNA printed on a glass slide where each spot represents a different gene. These DNA spots quantitatively capture mRNA, and the amount of mRNA for each gene-spot is quantified by the amount of fluorescence from the dyes (Cy3 and Cy5) incorporated into the mRNA from two individuals. Research on mRNA expression in *F. heteroclitus* using this new technology applied ANOVA statistical analyses among individuals (vs. contrasted to a single control) to examine the variation within and among populations (133, 135) (Figure 5). This study and others (37, 97, 201) indicate the importance of replication within and among populations to properly define biologically and evolutionarily important changes in mRNA expression. For *F. heteroclitus,* using ANOVA among healthy males acclimated to common conditions produced two important discoveries: (i) that approximately 20% of mRNAs were significantly different among individuals within a population and (ii) approximately 3% of mRNAs had a derived non-neutral pattern of expression indicative of adaptive variation (Figure 5A). The variation among individuals was not due to environmental differences (all individuals were acclimated to a common condition) and was thus most likely a result of heritable differences. Adaptive differences were identified as derived in northern populations where the variation exceeded the variation among southern *F. heteroclitus* and its sister species *F. grandis* (Figure 5A) (133). This greater variance within a species than between species is indicative of adaptive evolution. In addition to the two main points, it also became clear that the magnitude of difference (fold-change) was not necessarily indicative of a significant change. For example, a fourfold difference in the mean between populations could be equally as large as the variation within the population and thus, the fourfold difference would be insignificant (Figure 5B). Similarly, several genes had adaptive patterns of mRNA expression yet had less than 1.5-fold differences between populations (Figure 5B, gray box). The statistical significance of these small changes is due to the very small variation within each population. This is both statistically and biologically important because the relatively conserved, invariant amount of mRNA within each population suggests stabilizing selection for a specific mRNA concentration, and thus the small change between populations suggests functional importance (133). What is clear from these studies is that evolution acts on the significant inter-individual variation to change mRNA expression. These patterns of mRNA expression are indicative of adaptive evolution.

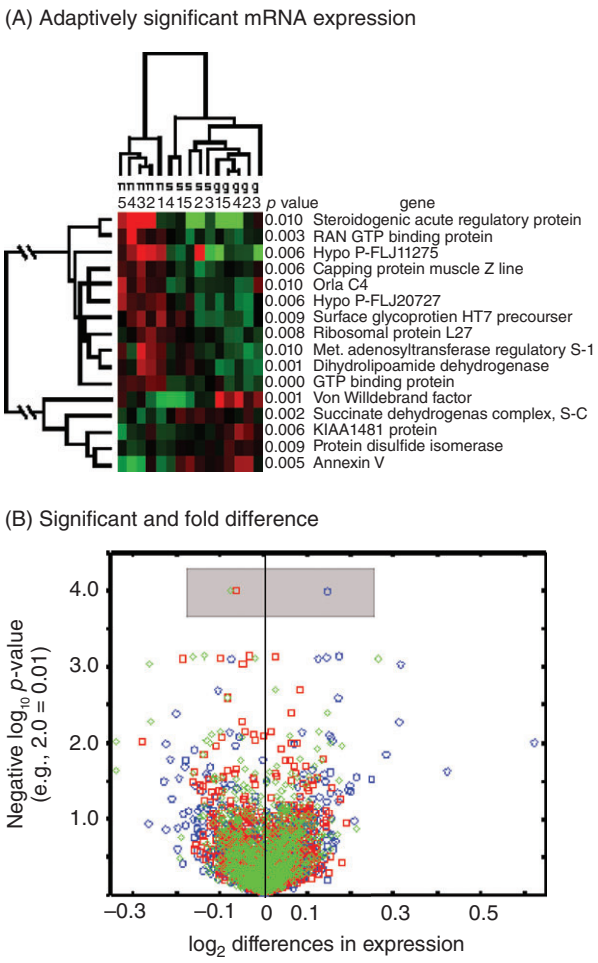


Figure 5 Microarray: genome-wide patterns of mRNA expression. (A) Heat map of adaptively significant mRNA expression. Red and green colors are the relative low or high expression. Notice northern individuals share similar expression patterns and are significantly different from southern *F. heteroclitus* and *F. grandis*. (B) Volcano plot: $\log_2$ expression relative to mean expression for each mRNA versus statistical significance as $-\log_{10}$ *p*-values. Gray box highlights the most significant mRNA where northern mRNA (blue circle) is statistically larger than both southern *F. heteroclitus* (red square) and *F. grandis* (green circle). Redrawn from Ref. 133.

In a separate study, adaptive differences, as well as reversible physiological acclimation effects, were studied to determine the number of mRNA transcripts influenced by these two processes (48). To compare adaptive differences and physiological effects, *F. grandis* and northern and southern *F. heteroclitus* were acclimated to three temperatures (12, 20, and 28°C) for more than 6 weeks and approximately 7000 unique mRNA transcripts were quantified using microarrays. Adaptive changes in mRNA expression were defined as those transcripts with significant derived expression (northern *F. heteroclitus vs*. both southern taxa—*F. heteroclitus* and the sister species *F. grandis*). These data revealed that more mRNA transcripts have significant adaptive differences than mRNA with significant acclimation effects. Furthermore, the mean adaptive difference was larger than the mean acclimation effects. Interestingly, few mRNAs were altered by both adaptive and acclimation effects and yet adaptive effects were more frequent at 12 or 28°C than at 20°C, and very few of these mRNAs had adaptive differences at two or more acclimation temperatures (48). When a significant interaction between adaptive and acclimation effects occurred, northern *F. heteroclitus* acclimated to 12°C were more similar to the southern taxa acclimated to 28°C. What this reveals is that northern individuals were most similar to southern individuals when acclimated to their mean summer temperatures. These data suggest that adaptive and physiological effects are independent of each other and that the genes of importance (adaptive mRNA transcripts) are dependent on the environment; that is, the mRNAs with adaptive expression were different at all three acclimation temperatures.

Further evidence of the adaptive importance of mRNA expression was demonstrated by examining five *F. heteroclitus* populations along the North America Atlantic coast (Figure 6A) (200). Among 329 metabolic mRNAs, 58 mRNAs (17.6%) significantly regressed with temperature. Yet, temperature and genetic similarity covary (i.e. along the coast, genetically more distant northern populations are also colder), and thus mRNA and temperature covariance could be due to neutral evolutionary processes. To identify adaptively significant patterns of mRNA expression along the cline requires quantifying differences in the amount of mRNA that is significantly associated with native temperature but independent of genetic similarity among populations (Figure 6B) (200). Thus, adaptive mRNA transcripts are those that regress significantly with habitat temperature following correction of the expected covariance due to genetic relatedness [phylogenetic generalized least squares (PGLS) analysis, (200)]. Among the 329 mRNAs, 13 (4%) have a significant temperature regression that is independent of genetic similarity among populations and are most parsimoniously described as evolving due to directional selection (Figure 6B).

In addition to PGLS analysis, ANOVAs were used to define neutral, stabilizing, or balancing selection. In Figure 6C, the 13 mRNAs under directional selection (defined above, pink circles) are identified as divergent along a habitat temperature gradient after correcting for variance due to phylogeny. Seven genes had significantly greater variation within a population than among populations (Figure 6C, blue circles), indicative of balancing selection (200). Twenty-four mRNAs had expression patterns indicative of stabilizing selection (Figure 6C, yellow circles). These 24 mRNAs had significantly lower variation both within and among populations than most genes (Bonferroni-corrected $p < 0.01$). Additionally, the mRNA with the least variation in expression levels is significantly biased for genes involved in oxidative phosphorylation (OxPhos) (10 of the 24 significant genes with low variance, $p < 0.05$; Fisher's exact test).

Overall, based on mRNA variation among many individuals and several populations and after identifying the neutral and adaptive portions of the variation, these data indicate that natural selection is acting on the expression of 44 out of

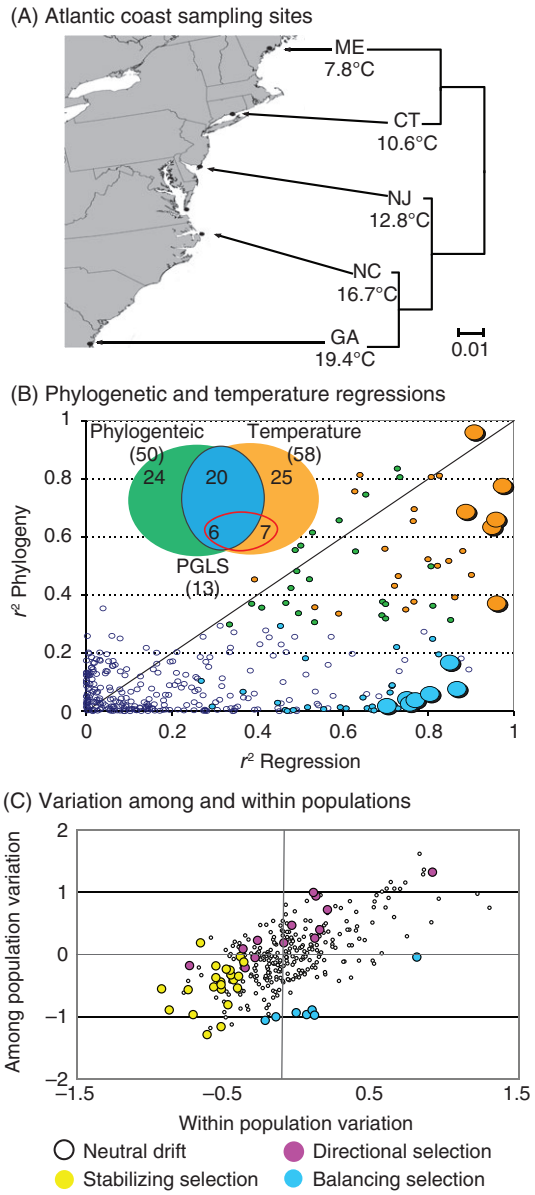


Figure 6 Clinal adaptive variation in mRNA expression. (A) Five sample sites and their phylogenetic relationship microsatellite-derived neighbor-joining tree with median annual temperatures (°C) averaged over 30 years. Branching pattern is a neighbor-joining tree constructed from pairwise Cavalli-Sforza and Edwards' chord distances (33) calculated from microsatellite allele frequencies. (B) Relationships between phylogenetic and ecological effects on variation in gene expression. For each gene, the explained variation (r^2) for phylogeny [genetic distance based on Cavalli-Sforza and Edwards' chord distances (33) calculated from microsatellite allele frequencies] versus the explained variation (r^2) for habitat temperatures. Venn diagram is for the numbers of genes that have significant regression with habitat temperature (orange), phylogeny (green), or both temperature and phylogeny (blue). Colors of spots in the graph correspond to Venn diagram. Enlarged spots are the 13 genes that regress significantly with habitat temperature after correcting for phylogeny (red circle; Venn diagram) using the phylogenetic generalized least squares (PGLS) approach, and thus appear to be evolving by natural selection. (C) Variation in mRNA expression within or among populations. Plotted are the log of variation. Ratio of variation is indicative of evolutionary processes (directional, stabilizing, balancing, or neutral). Redrawn from Ref. 200.

the 329 genes (~14%) with directional selection acting on 13 genes, stabilizing selection acting on 24 genes, and balancing selection acting on 7 genes.

Effect of gene expression on physiology

The functional significance of these *F. heteroclitus* adaptive patterns of mRNA expression was demonstrated by examining the relationship between cardiac metabolism and quantitative variation in mRNA for metabolic genes (136) (Figure 7). Previously in this manuscript (Figure 3), cardiac metabolism was considered adaptive because it was higher in northern *F. heteroclitus* due to adaptive increases in glycolytic enzymes (140, 141, 146). To examine how mRNA expression relates to cardiac metabolism, metabolism was measured in 16 *F. heteroclitus* individuals (eight each from Maine and Georgia, USA). Cardiac metabolism was measured using three different substrates: glucose, fatty acids, and secondary metabolites (LKA: lactate, ketones, and alcohols, Figure 7A). Cardiac metabolism is a function of body mass, so the body mass effect was removed by using the residuals from body mass-metabolic rate regression (i.e. residuals are the difference from the observed and predicted metabolic rate based on body mass).

Cardiac metabolism was highly variable among individuals (Figure 7A), with as much an 11-fold difference among individuals for fatty acid dependent cardiac metabolism and a twofold difference for glucose-dependent cardiac metabolism. Additionally, individuals vary in which substrate supports the greatest metabolic rate. For example, fatty acids supported one of the highest metabolic rates in one individual (G3, Figure 7A) while glucose supported one of the lowest metabolic rates in this same individual (Figure 7A). Thus, variation in the substrates that support metabolism arises because of the variation in the use of each substrate among individuals. Even with large variation in metabolism within populations, northern populations had significantly greater metabolic rates for glucose and fatty acid ($p < 0.02$) and nearly significantly greater rates for LKA ($p \sim 0.06$).

To examine if mRNA expression is related to cardiac metabolism, mRNA expression for 119 metabolic genes was measured in all 16 individuals (8 per population) at 16-fold replication. Surprisingly, most (94%) of the mRNA expression was significantly different among individuals within each population ($p < 0.01$; Figure 7B). This seemingly incredible level of significant differences in mRNA expression among individuals was discernable because of the high number of technical replicates for each individual. Despite the high variation among individuals in mRNA expression, individuals with more similar patterns of expression could be grouped (Figure 7C). Specifically, individuals clustered into three groups where individuals within a group were more similar and significantly different from individuals in other groups. For example, for the mRNAs in Figure 7C, if group 2 individuals had very high relative levels of expression, group 3 individuals would have very low levels of expression

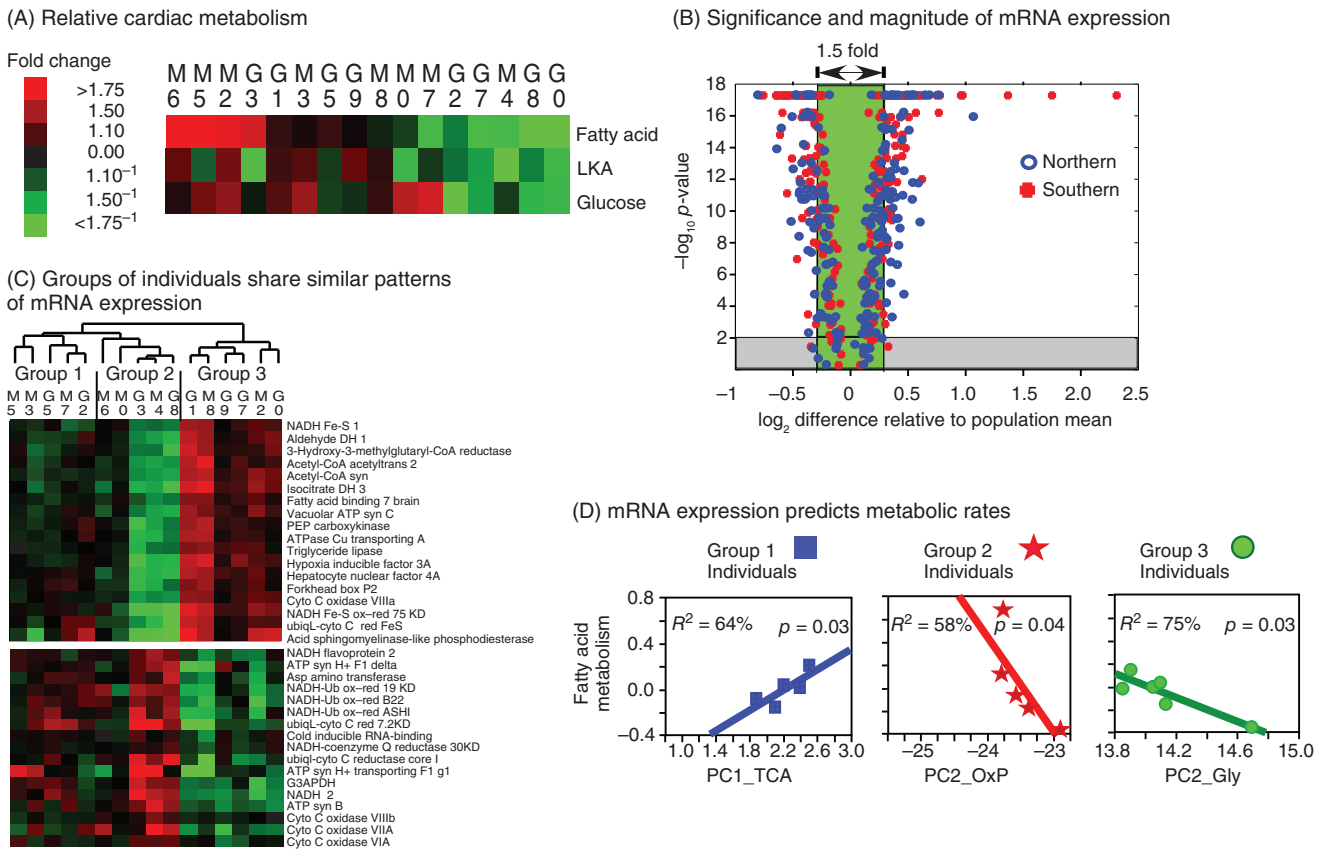


Figure 7 mRNA expression and cardiac metabolism. (A) Relative levels of cardiac metabolism for 16 individuals (8 per Maine and Georgia population). Cardiac metabolism was measured using glucose, fatty acid, and LKA (lactate, ketones, and ethanol) as substrates (136). Red is at least 1.75-fold greater and green is at least 1.75-fold lower than the overall mean. (B) Significant mRNA expression differences between individuals within a population (negative $\log_{10}$ values, thus 2 is equal to a *p*-value of 1%) versus the fold difference ($\log_2$ values, thus 1 = twofold difference). Fold differences are relative to the overall mean for each mRNA. Green background shadowing shows mRNA with 1.5-fold or less differences. *p*-Values are truncated at values more than 10^{-17}. (C) Patterns of mRNA expression among all 16 individuals (green is relatively low, red is relatively high). A subset of mRNAs coding for metabolic genes that show shared expression within groups that is significantly different among groups. (D) Fatty acid metabolic rates are relative to the mRNA expression. mRNA expression summarized as one of three primary biochemical pathways (two principal components each for glycolysis, TCA cycle, and oxidative phosphorylation). Similar patterns occur for glucose and LKA supported cardiac metabolism (136).

(Figure 7C). No other random sets of individuals shared these significant patterns. That is, the three groups of individuals are statistically different (136).

The level of mRNA expression, although highly variable, appears to be functionally important because it can be used to predict substrate-specific cardiac metabolic rates (Figure 7D). That is, within each of the three groups of individuals (Figure 7C), the variation in mRNA expression predicts the substrate-specific metabolic rates (Figure 7D). However, the mRNAs that statistically explain the variation in cardiac metabolism were different among the three groups (136). Specifically, mRNA expression was summarized as a linear equation for mRNAs encoding proteins for each of the three major biochemical pathways: glycolysis, TCA-cycle, and OxPhos. This linear equation was defined by a principal component analysis (PCA) across all 16 individuals, and the first two PCA axes were used to predict metabolic rates. Figure 7D shows the significant PCA and the explained variance (R^2) for each group of individuals for fatty acid metabolism. Notice the sets of mRNAs that predict fatty acid metabolism are different among the three groups. Similar significant relationships between mRNA and metabolism (not shown) were observed for glucose and LKA dependent metabolism (136). For example, glucose-dependent metabolism is related to glycolytic mRNAs expression ($R^2 = 81\%$) for group 1 individuals but OxPhos mRNAs for groups 2 and 3 ($R^2 = 65$–70%). Overall, mRNAs from different metabolic pathways are related to different substrate-specific metabolisms within a group.

These data demonstrate that mRNA expression variation can predict the level of cardiac metabolism, suggesting that mRNA expression differences lead to metabolic differences. Equally important is the observation that the large variation in the amount of mRNA expression (where 94% of mRNA are significantly different among individuals, Figure 7B) contributes to the large variation in cardiac metabolism (2- to

11-fold difference among individuals, Figure 7A). The observation that among different groups of individuals, cardiac metabolism is explained by different sets of mRNAs, which are part of different metabolic pathways (Figures 7C and 7D), suggests that multiple molecular pathways can achieve similar physiological outcomes. That is, if the heritable variation in mRNA expression is independent (i.e. not due to change in a single transcription factor) then these mRNA expression variations represent different genetic bases for the adaptive variation in cardiac metabolism. Thus, the adaptive variation in cardiac metabolism is polygenic in that the physiological variation depends on changes in many independent sets of genes.

Elucidating the relationship between mRNA expression and a key physiological trait like cardiac metabolism is important because these data suggest that altering metabolism can be achieved by changes in a diversity of mRNAs. This highlights the importance of standing genetic variation for biochemical adaptation and demonstrates that only some of the diverse variants need to be favored by natural selection for the evolution of physiological adaptations. Specifically, the large variation in fatty acid metabolism among the three groups of individuals is related to changes in mRNAs involving one of the three metabolic pathways: TCA, OxPhos, or glycolytic pathways. These observations support the concept of polygenic adaptation with sufficient standing genetic variation where more than one evolutionary solution exists; for some individuals, it is the variation for mRNA expression in OxPhos, in others, glycolysis. This concept has important consequences. For example, the fact that many different sets of polymorphic genes can influence adaptive metabolic change means that selection is more robust. It is more robust because polygenic adaptation involves (i) more than one gene with multiple alleles that cause an adaptive change, (ii) these alleles already exist at reasonable frequency, and (iii) these alleles need not go to fixation (100%). Thus, with many genetic targets that occur in many individuals, natural selection is more likely to occur than if few, rare, genetic variations occur (154, 155, 170).

Additionally, adaptive evolution can be achieved without too much genetic load for the same reason: adaptive evolution need not involve large changes in allele frequencies. The importance of polygenic selection is becoming recognized in many organisms, including humans, and is significantly altering our understanding of the genetic basis of adaptive evolution (59, 90, 93, 103, 105, 153, 182, 192).

To summarize the adaptive divergence in mRNA expression across the thermal cline

- LDH-B has greater enzyme concentrations in northern populations due to increased transcription rates related to adaptive divergence in DNA sequences that bind transcription factors.
- Within populations, the variation in mRNA expression is considerable: among hundreds of genes, approximately 20% have statistically significant differences in mRNA expression among individuals within a population.
- Evolutionary analyses suggest 3% to 4% of genes exhibit adaptive divergence in mRNA expression.
- Physiological acclimation has less of an influence on mRNA expression than adaptive divergence, and few genes have interactions between adaptive and physiological effects.
- Variation in mRNA expression explains up to 81% of the variation in cardiac metabolism, but more importantly, mRNAs encoding proteins from distinct pathways are important for different individuals.
- The patterns and extent of variation in mRNA expression suggest adaptive variation is due to standing genetic variation and is polygenic—many mRNAs have significant expression changes that influence adaptation with different sets of mRNAs being responsible for the variation in metabolic phenotypes.

Clinal Variation in Whole-organism Traits

Northern and southern *F. heteroclitus* populations differ in a variety of traits at the whole organism level, including morphological traits in both embryos and adults (3, 125), physiological traits in embryos such as developmental rate and metabolic rate (55), and physiological traits in adults, such as metabolic rate (12, 13, 65, 84, 136), thermal tolerance (64), hypoxia tolerance (116, 117), and salinity tolerance (166). For many of these traits, the differences between the northern and southern populations align with what would be predicted based on adaptation to their local thermal conditions. For example, fish from southern populations have greater tolerance of high temperatures and low oxygen concentrations than do fish from northern populations (64, 117), consistent with the warmer temperatures and lower oxygen saturation typical of more southern habitats. However, as discussed previously, differences between populations can result from either neutral or adaptive processes and determining whether a particular trait has evolved as a result of natural selection in response to a particular environmental factor is challenging.

If trait variation between populations at the extremes of the species range is the result of thermal adaptation, then geographic patterns of variation in these physiological traits should be consistent with the patterns of temperature variation along the coast. Thermal tolerance varies linearly with latitude, consistent with the latitudinal pattern in water temperature, whereas hypoxia tolerance undergoes a steep transition from the northern phenotype of low hypoxia tolerance to the southern phenotype of high hypoxia tolerance along the New Jersey coast (82). These data suggest that thermal and hypoxia tolerances are not genetically or functionally associated with each other and that the hypoxia tolerance difference between northern and southern

F. heteroclitus populations may not represent an adaptation to temperature.

The genetic basis of variation in thermal and hypoxia tolerances in *F. heteroclitus* has been examined using genome-wide association approaches. A traditional association analysis failed to detect any SNP associated with variation in thermal tolerance within a central New Jersey population and detected only four SNPs associated with variation in hypoxia tolerance (82). However, this type of association analysis considers each SNP as an independent site and generally only has power to detect SNPs that have a large effect on the phenotype. This may not be the most appropriate approach if trait variation is polygenic—resulting from interactive effects of SNPs in many different genes, each having a small effect on phenotype. Alternative methods, such as random forest analyses, that enable the identification of suites of genes that act together to explain phenotypic variation, may be more appropriate for the analysis of complex physiological traits (29). Using this type of analysis, 43.4% of the variation in thermal tolerance can be explained by variation in 47 SNPs, and 51.9% of the variation in hypoxia tolerance can be explained by variation in 35 SNPs. These data demonstrate that both thermal and hypoxia tolerance are polygenic traits that result from variation in multiple genes of small effect. However, none of the SNPs overlapped between the two analyses, suggesting that the genetic basis of these two traits is distinct.

Although identifying the genetic basis of variation in these two traits does not directly address the question of whether the traits themselves have evolved via natural selection, these data can be combined with sequence-based analyses of selection. Over 500 SNPs show significant evidence of departure from neutral expectations in *F. heteroclitus* along the Atlantic coast (82). When examining a single population in NJ, one of these SNPs was among those associated with variation in thermal tolerance. This SNP is located in a gene encoding a ubiquitin ligase that is involved in targeting thermally damaged proteins for destruction in the lysosome. Similarly, one SNP associated with variation in hypoxia tolerance overlapped with the SNPs showing significant evidence of selection. This SNP is located in a gene that acts as a coactivator of a protein that interacts with the hypoxia-inducible factor HIF-1a, which is a major transcriptional regulator of the response to hypoxia.

These results illustrate two key points: (i) physiological traits are often polygenic, and thus adaptive evolution in these traits may involve variation in multiple genes each with relatively small effects on the phenotype, and (ii) demonstrating the action of natural selection in these traits may require combining multiple types of evidence to support the hypothesis of adaptive evolution.

Local Adaptation

The studies described above examined the variation within and among populations across a large geographic distance—the eastern coast of North America. Unexpectedly, we have also discovered rapid local adaptation at scales from within the Chesapeake Bay to within single saltmarsh estuaries. Rapid local adaptation seems improbable when the cost of selection is considered (47, 100, 196). That is, natural selection requires a difference in survival or reproduction related to genetic variation, and thus selective death is linked to genotype. Because of variation in survival or selective death, selection of many genes can reduce the number of individuals to such an extent that the species becomes unviable (77). The greatest cost of selection occurs when an adaptive allele arises from a new mutation and goes to fixation (77). That is, starting from a very low frequency equal to $1/2N$ (i.e. one allele among N number of individuals in a population) the allele evolves to fixation (100% frequency). Through this type of selective sweep, allelic variation is lost from the population. Because of the cost of selection, or genetic load, and the identification of unexpectedly large amounts of allelic variation (110), much of natural standing genetic variation is attributed to neutral processes (98). Thus, the classic evolutionary expectation is few adaptive changes that evolve slowly.

If evolutionary adaptation is indeed slow and dependent on a few genes of large effect, we might not expect to find local adaptation among well connected *F. heteroclitus* populations where alleles from adjacent, nonlocal populations should swamp the effect of selection. Yet, recent data using the power of high-throughput sequencing indicates an abundance of genetic variation within and among populations. This high standing genetic variation may be enough to influence the success of individuals in local environments allowing frequent rapid adaptation to local environments. Given below is genomic evidence for local adaptation among *F. heteroclitus* populations separated by a few 100 km to less than 100 m.

Local adaptation to the osmotic environment

Intracellular solute concentrations that alter the water balance or osmolarity of cells are carefully controlled in vertebrate animals (181). For teleost fishes, intracellular solute concentration (osmolality) is maintained at levels that are higher than the environment for freshwater species, but lower than the environment for marine species. Regulation of constant plasma osmolality relative to the environment is crucial for maintaining physiological function and is one of the parameters that defines the fundamental niche of teleost fish. Most fish species are either freshwater specialists or marine specialists and have impaired fitness in high or low salinities, respectively. Osmoregulatory physiologies that are specialized for living in marine or freshwater habitats have evolved multiple times among fishes (20, 193). In contrast, relatively few fish species live in estuarine environments that have large environmental salinity fluctuations (165). Those fish that do live in estuaries typically have very wide physiological flexibility to maintain osmotic homeostasis across a wide range of salinities. These diverse physiological

abilities tend to suit each species to the osmotic environment in which they reside and have convergently evolved multiple times, implying that osmoregulation is a physiological trait that is adaptively important.

The evolutionary convergence of osmoregulatory physiologies between different osmotic environments is exemplified among *Fundulus* species (199). *F. heteroclitus* exhibits one of the broadest ranges of osmotic acclimation abilities known within the fishes (165). They are estuarine specialists that occupy the entire continuum of osmotic niches from marine to freshwater and are the most abundant fish species in salt marsh habitats along the Atlantic coast of the United States (101, 190). In contrast, congener *F. majalis* (Figure 3) is coastal but occupies the marine end of the salinity continuum within estuaries, exhibiting limited abilities to acclimate to dilute water (75). Approximately half of *Fundulus* species occupy a coastal niche and can tolerate a wide range of salinities (199). The remaining species are distributed throughout inland waterways, can tolerate extremely dilute water, but cannot acclimate to salty water. These freshwater *Fundulus* specialists have independently evolved at least three different times. Importantly, these freshwater *Fundulus* can survive and thrive in dilute conditions unlike their marine relatives. Such evolutionary convergence within *Fundulus* is strong evidence that osmoregulatory abilities within *Fundulus* are adaptively important for different osmotic environments.

In addition to these macro-evolutionary patterns among *Fundulus* species, *F. heteroclitus* populations occupy the entire salinity gradient from marine to freshwater in the Chesapeake Bay region (Figure 8A). As one moves upstream along the salinity gradient in the Potomac River, neighboring groups of fish are genetically very similar. However, fish populations found in dilute freshwater in the extreme upper estuary (FW-native) are genetically distinct from those just a few miles downstream in brackish water (BW-native) (Figure 8B) (202). In parallel, freshwater populations in the James River are genetically distinct from brackish populations just a few miles downstream. This suggests FW-native fish are locally adapted to their dilute freshwater environments, and this limits migration across the freshwater boundary. Indeed, for FW-native or BW-native populations, swim performance is highest in salinities that match their native environment and poorest in salinities that mismatch their native environment (28). Marine-native and BW-native populations are as genetically distinct as BW-native and FW-native populations (Figure 8B) (202). As such, traits that are evolving by neutral processes should be equally divergent between marine-native and BW-native fish as they are between BW-native and FW-native fish. However, this is not the case for osmoregulatory physiology, where marine and BW-native fish are similar, but FW-native fish are distinct, in their abilities to maintain osmotic homeostasis in dilute freshwater (202). That is, the physiological osmoregulatory divergence between FW-native and BW-native populations is greater than expected based on neutral genetic divergence, and thus this divergence is likely adaptive. These data are evidence for microevolutionary adaptation of osmoregulatory abilities along natural salinity gradients in *F. heteroclitus*.

The molecular basis for adaptive osmoregulatory divergence between marine and freshwater *F. heteroclitus* populations has been defined by comparative analyses of gill mRNA expression and genome-wide DNA sequence analyses. Gills are the primary organ responsible for osmoregulation in teleost fishes (62) and patterns of gill mRNA expression provide insights into the molecular pathways and physiological mechanisms that matter for adaptive osmoregulatory physiology. Along the Chesapeake salinity gradient, population differences in mRNA expression reveal changes in many genes and pathways in FW-native populations that are greater than expected based on neutral genetic distance (202). These mRNA expression differences are, therefore, putatively adaptive, mirroring patterns of physiological differences between populations. Furthermore, the expression of physiologically flexible mRNAs shows less variation among individuals than other mRNAs, and this expression variation is lowest in FW-native fish, which is further evidence that they are adaptive (171). Importantly, mRNAs that were physiologically responsive (up- or downregulated) to experimental salinity challenges (Figure 8D) were more likely to show patterns of adaptive divergence between populations than mRNA for genes that were not physiologically responsive (Figure 8C) (202), especially for mRNAs that respond quickly after salinity challenge. These data reveal the molecular mechanisms that underpin physiological adjustments to fluctuating environments and suggest that both physiology and underlying mRNA expression are collectively shaped by natural selection to suit different osmotic environments.

In *F. heteroclitus,* transcriptional responses to osmotic challenges, especially those showing patterns of adaptive divergence, provide insights into the biochemical and molecular physiology of osmoregulation. These mRNA expression patterns implicate pathways involved in cell volume regulation, polyamine synthesis, and immediate early signaling, among others, which are all crucial for adaptive osmoregulatory divergence (202). The evolutionary importance of these *F. heteroclitus* responses was supported by data showing that expression of these same genes and pathways also differed when compared to other *Fundulus* species with distinct osmoregulatory abilities occupying different salinity niches. The same molecular pathways and physiological processes were modified across osmotic environments on both macro (among species) and micro (within species) evolutionary timescales; this repeated modification is indicative of adaptive evolution.

The DNA sequence variation responsible for the evolved variation in *F. heteroclitus* osmoregulatory abilities has been examined through genome-wide association studies (GWAS) and genome scans for signatures of natural selection. GWAS seeks to identify DNA sequence variation (alleles) that is associated with the physiological variation. In *F. heteroclitus*, GWAS revealed that variation at 26 genes accounted for

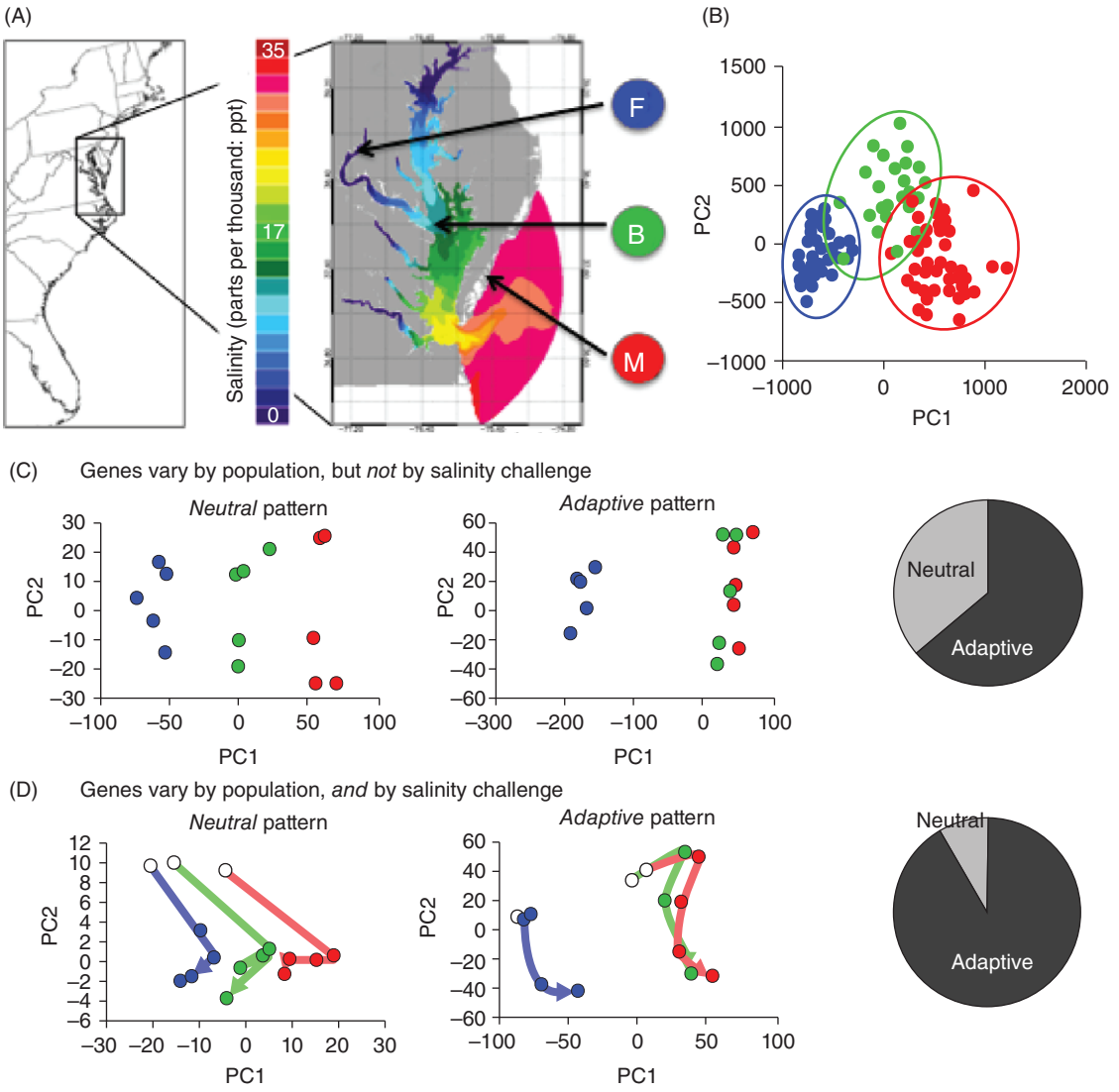


Figure 8 Local osmotic adaptation. *F. heteroclitus* population variation along a salinity gradient in the Chesapeake Bay. (A) Map of salinity gradient in the Chesapeake Bay, where experiments contrasted physiology and genomics of marine-native (M), brackish-native (BW), and freshwater-native (FW) populations. (B) Plot of genetic similarity of individuals collected from the three Chesapeake populations, where neighboring populations were equally genetically distant from each other. (C) Principal component analysis of genes that are differentially expressed between populations but not affected by salinity challenge. Genes where the pattern of population divergence matches the neutral expectation [e.g. as established by pattern of genetic relatedness shown in (B)] are included in the left panel and genes where the patterns of population divergences consistent with adaptation in the freshwater population (blue) are included in the right panel. Pie chart shows the proportion of genes within this set that show the neutral or adaptive pattern. (D) Principal component analysis of genes that are differentially expressed between populations and that are differentially expressed during salinity challenge. Genes where the pattern of population divergence matches the neutral expectation [e.g. as established by the pattern of genetic relatedness shown in (B)] are included in the left panel and genes where the patterns of population divergence consistent with adaptation in the freshwater population (blue) are included in the right panel. Pie chart shows the proportion of genes within this set that show the neutral or adaptive pattern. A greater proportion of genes that are transcriptionally responsive to salinity show the adaptive pattern than genes that are not responsive to salinity. Principal component analyses are redrawn from Ref. 202. Salinity gradient heatmap of the Chesapeake Bay was generated from the NOAA Chesapeake Bay Operational Forecast System (https://tidesandcurrents.noaa.gov/).

56% of the variation in osmoregulatory abilities among individuals (27). Not all individuals with the best osmoregulatory abilities had the same set of variants at each of these loci. This indicates that osmoregulatory ability is a polygenic trait; it involves allelic variation at multiple genes, and many combinations of these variants can support osmoregulatory abilities. These physiology-associated variants also tended to evolve by adaptive rather than neutral processes between BW-native and FW-native populations (27). This demonstrates that the genetic variation for osmoregulatory physiology is

important for adaptation to different osmotic environments and that this adaptation is polygenic.

Together, GWAS, selection scans, quantification of mRNA expression, and comparative physiology provide insight into the complex and polygenic adaptations that enable evolutionary transitions between osmotic niches. These integrative studies also demonstrate that natural selection and reversible physiological acclimation in fresh and brackish populations involve the same biochemical and molecular pathways. This relationship between physiological acclimation and adaptive divergence is different for salinity than for temperature; for temperature, adaptive mRNAs were different from the mRNAs that were responsive to acclimation (48). The adaptive divergence in response to temperature was based on differences between geographically distant populations (Maine and Georgia), whereas adaptive osmoregulatory divergence was measured among geographically proximate populations along a salinity gradient within the Chesapeake Bay. It is unclear if the fundamental differences where physiological and adaptive genes are shared (osmotic regulation) or are different (temperature) are due to the differences in the physiological systems or are due to the evolutionary time scale. Perhaps the difference between salinity and temperature adaptation is because osmotic adaptation and regulation are restricted to similar pathways, whereas adaptive temperature divergence may recruit variants in pathways that do not overlap with those that contribute to thermal acclimation. Also, many adaptive differences among populations in osmoregulatory mRNA expression were for genes that were differentially expressed very quickly (e.g. hours) after exposure to osmotic challenge, whereas temperature-regulated genes were compared between populations 6 weeks after thermal challenge. It is plausible that quickly responding gene expression pathways (e.g. sensing and signaling) contribute to adaptive divergence in gene expression along thermal gradients. Alternatively, more recent adaptation along salinity gradients, in contrast to presumably older adaptation to thermal gradients, may favor adaptive divergence that recruits physiologically inducible pathways. Defining the causes of this fundamental difference is important and merits further study.

To summarize adaptive osmoregulation in F. heteroclitus

- Among *Fundulus* species, similar osmoregulatory physiologies have evolved multiple times, in parallel with multiple independent species' radiations into freshwater environments, indicating that osmoregulatory physiology is an adaptive trait.
- Within *F. heteroclitus*, populations that inhabit extreme upper estuary habitats have an enhanced ability to tolerate very low salinity and perform better in low salinity compared to other populations that inhabit saltier habitats, indicating adaptation to local osmotic niches.
- Derived abilities to acclimate and perform well in low salinity exceed neutral expectations and are, therefore, also likely adaptive.
- Gene expression that is physiologically responsive to experimental osmotic challenges is more likely to show patterns of adaptive divergence than genes that are not responsive to challenge.
- The suite of physiological and mRNA responses to salinity, coupled with DNA sequence variation from multiple genes associated with variation in osmoregulatory abilities, indicates that osmoregulatory physiology is a complex polygenic trait that evolves by natural selection in alternate osmotic environments.
- Osmotic adaptation is different from thermal adaptation in that the genes involved in osmotic adaptation and reversible physiological responses are similar, but in temperature adaptation, few genes show both adaptive divergence and physiological responsiveness.

Local, rapid adaptation to temperature

Certain local estuaries are exceptionally warmer than the surrounding waters due to TE discharged from power plants (Figure 9). The increase in temperature by 4 to 12°C at these TE sites is less than 50 years old. Even though TE sites are young relative to nearby reference populations, *Fundulus* occupying these TE sites have greater critical thermal maxima (CTmax: maximum temperature at which an individual can no longer actively escape unfavorable conditions) (49). Genotyping by sequencing (GBS) was used to identify 5449 SNPs among 239 individuals comparing two triads: three populations with one TE population and two reference populations (one ~60 km north and one ~60 km south). This experimental design allows for a specific evolutionary analysis: SNPs evolving by natural selection should have significantly different allele frequencies at the TE site relative to both reference populations, but SNP allele frequencies should not differ between the reference sites. Using a triad design also controls for demography and other neutral processes that could create significant genetic divergence among populations.

Focusing on a single TE triad at Oyster Creek NJ, we can identify subtle demographic structure using a "discriminant analysis of principal components" (DAPC). This demographic structure is identified because DAPC maximizes the differences between populations while simultaneously minimizing the differences within populations (94). Basically, the DAPC analysis identifies a linear combination of SNPs producing the largest differences between populations. A DAPC of the Oyster Creek triad (north and south reference and TE sites, Figure 9A) shows that all three sites have subtle allele frequency differences (*X*-axis) and that the TE site can be distinguished from both reference sites (*Y*-axis). To identify the SNPs that contribute to this difference between

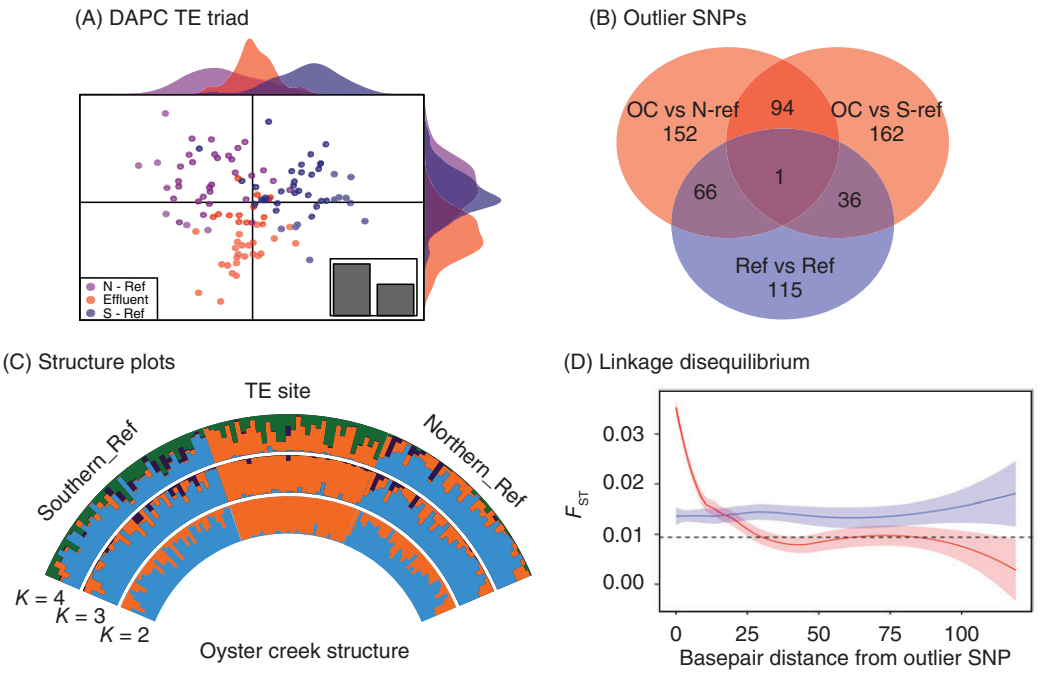


Figure 9 Local adaptation to warmer temperatures. Three populations (triads) were examined: a northern and southern reference population and a locally heated thermal effluent (TE) population. (A) Genetic structure among Oyster Creek TE site using all approximately 5400 SNPs. *X* and *Y* axes are the first and second principal components (linear equation maximizing the variation among populations). The first principal component separates all three sites, and the second separates the TE site (red) from both northern and southern reference sites. (B) Outlier SNPs with statistically large and unexpected F_{ST} values for paired comparisons between TE and references. SNPs evolving by natural selection are the 94 SNPs where TE differs from both reference populations but are not different between the pair of reference populations. (C) Structure plots using 94 outlier SNPs for 2, 3, or 4 groups of individuals ($k = 2$, 3, or 4). TE site is distinct in all comparisons. (D) Linkage disequilibrium as indicated by similar F_{ST} values relative to the DNA distance (base pair, bp). Dashed line is the mean, and shading is the 95% confidence bounds for the mean genome-wide F_{ST} value estimate for both TE versus reference comparisons. Red is the decline in F_{ST} value for outlier SNPs, and blue is the mean F_{ST} value when TE and references are randomly permutated.

the TE and reference sites, SNPs with significantly large F_{ST} values were selected. SNPs with significantly large F_{ST} values are unexpected because they greatly exceed neutral models based on random permutations (Figure 9B). These selectively important SNPs had to be an outlier for TE versus both references, but not between references. The rationale behind defining adaptive selection this way is threefold: (i) outlier SNPs have F_{ST} values that exceed the neutral expectation, (ii) they are significant for the TE sites relative to both cooler reference sites, and (iii) they do not have significant demographic effects because they are not different between the two cooler reference sites (49).

These data indicate that 94 SNPs in the Oyster Creek TE site display significantly large genetic distance (F_{ST}) that is most parsimoniously explained by adaptive selection. To verify that these SNPs have frequencies that distinguish the TE site from both references, the 94 outlier SNPs from Oyster Creek where used in a Structure analysis (154), which groups individuals based on shared and similar allele frequencies. The Structure analyses clearly distinguished the TE site from both reference sites (Figure 9C). Regardless of clustering into two, three, or four groups ($k = 2$, 3, or 4, Figure 9C), the Oyster Creek TE site segregates from both reference sites, which share similar allele frequencies.

These analyses of approximately 5400 SNPs show that many SNPs have unusually large F_{ST} values indicative of evolution by natural selection. The interesting observation is that none of the adaptive SNPs were rare in the reference populations, nor did they reach fixation in the TE population. Importantly, selection seems to be SNP-specific since linkage distance is minimal (Figure 9D) and there are no significant LDs among the 94 outlier SNPs in Oyster Creek.

In conclusion, the rapid adaptation to recent thermal warming at a TE site appears to be polygenic because it involves standing genetic variation in many independent genes that are not swept to fixation (49).

Local, rapid adaptation to pollution

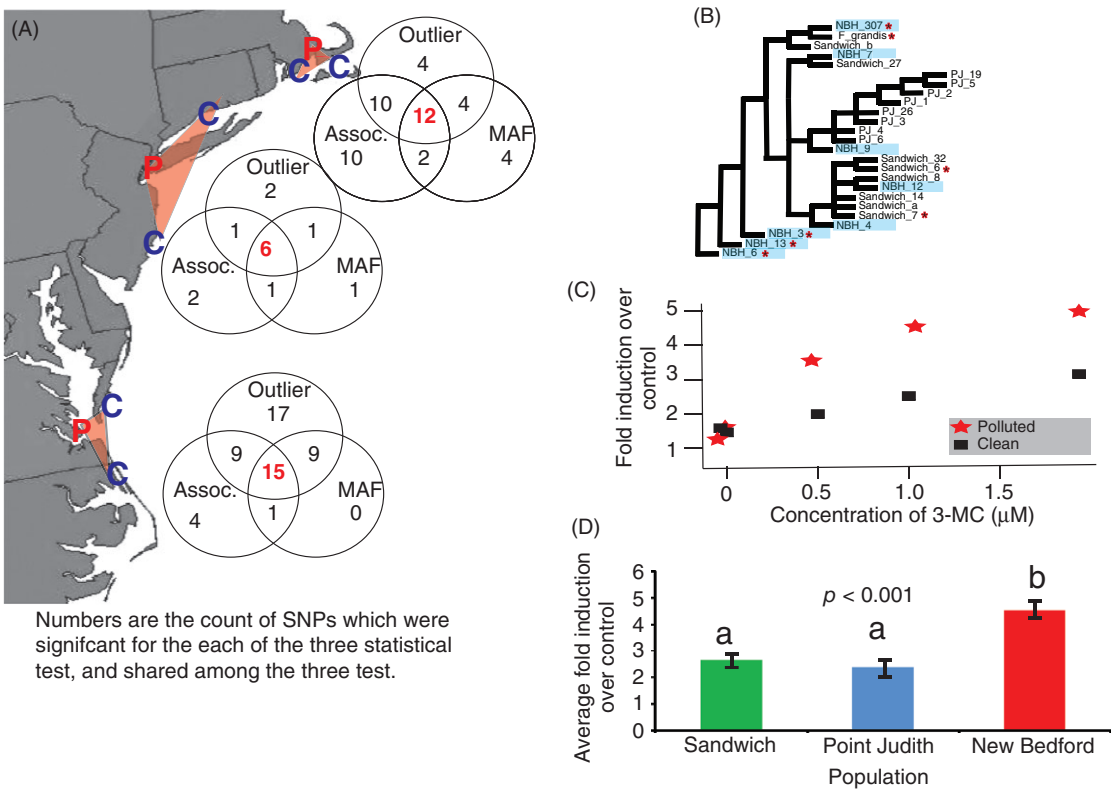
The production of persistent organic pollutants (POPs) including polycyclic aromatic hydrocarbons (PAHs) and polychlorinated biphenyls (PCBs) contribute to one of the more serious forms of pollution: exposure is associated with carcinogenicity and mutagenicity (11, 67, 109, 111, 160, 210)

and has been associated with metabolic diseases, including type 2 diabetes, obesity, and energy metabolism (7, 96, 107, 108, 112, 160). *F. heteroclitus* populations at polluted sites have adapted to a widely distributed class of POPs even though polluted habitats are only a few decades old (31, 38, 121, 126–129). The toxic effects of these POPs are mediated largely through the aryl hydrocarbon receptor (AHR) pathway (68, 122, 147). Importantly, *F. heteroclitus* embryos from the polluted populations are resistant to these POPs in the first and second generations after rearing in a common clean environment (127–129). Thus, the differential survival of fish from polluted habitats is due to genetic adaptation rather than reversible physiological effects.

Resistance to POPs in *F. heteroclitus* is associated with normal development and the lack of cardiac abnormalities that occur among individuals from reference populations (populations from relatively nonpolluted sites) when exposed to POPs during development (25, 26). The abnormal development is associated with changes in mRNA expression not seen in fish from polluted populations (24, 26). POPs also alter the function of metabolic pathways, specifically the OxPhos pathway that is responsible for the vast majority of ATP production. Similar to development and gene expression patterns, resistant populations show little change in OxPhos metabolism when exposed to POPs, whereas nearby clean reference populations have significant changes, which are heritable (57, 58). These phenotypic changes reduce POPs' influence on development and metabolism and enhance survival and are recent adaptations. A key challenge was to identify the molecular and genetic basis for this adaptation.

To identify the genetic basis for adaptation to pollution, one of the first genomic surveys used mass spectrometry, which sorts charged DNA molecules based on their mass-to-charge ratio, to identify 354 SNPs among three polluted populations [Figure 10A, (203, 205)]. For each of the 354 SNPs, the three polluted populations were compared to two references in a triad design [one clean population north and one clean population south of the resistant polluted population (Figure 10A)]. In the triad design, evolution by local adaptation requires the polluted population to be significantly different from both northern and southern reference sites without significant difference between the references. These requirements were applied to three separate evolutionary analyses: outlier F_{ST} distribution, significant association between allele frequencies and level of pollution, and significant differences in MAF (minor allele frequencies) (205). The union of three separate adaptive tests (i.e. all three tests were significant) identified approximately 2% to 4% adaptive SNPs among the three triads (Figure 10A). All three polluted populations are resistant to pollution with each having 6 to 15 SNPs significant for all three statistical tests (Figure 10A). Yet, only one SNP was shared among the three triads: a SNP in CYP1A, the enzyme cytochrome P4501A. CYP1A is a phase I xenobiotic-metabolizing enzyme integral to the detoxification pathway (205). The observation that the majority of putatively adaptive SNPs differed among the three polluted populations could reflect polygenic selection with a redundant genetic variation that results in a similar phenotype (resistance to pollution). Yet, there are two alternative interpretations of the lack of shared adaptive SNPs. One alternative is that each of the triads could have specific POPs and other environmental stressors and the selectively important SNPs represent unique solutions to these different selective forces. A second alternative is that with sparse sampling of the genome (as is the case for any GBS study) there can be variation in linkage among SNPs in highly polymorphic populations such that different SNPs are linked to the same adaptive variant in different populations. If this were to occur, one would risk incorrectly concluding that different loci are responsible for adaptation in different populations. Whether the different selectively important SNPs among populations represent redundant polygenic adaptation is thus unresolved.

The only selectively important SNP shared in all three polluted populations was the CYP1A SNP in the first intron. The frequency of this outlier SNP was approximately 60%, 22%, and 10% in polluted, southern reference and northern reference populations, respectively. To determine its functional significance, the promoter-intron was sequenced in 24 individuals from the polluted New Bedford population and the two reference populations [$n = 8$/population, Figure 10A (206)]. To test how these promoter-introns regulate mRNA transcription, these promoters were linked to a reporter gene, and the expression of the reporter gene was quantified in cell culture (Figures 10C and 10D). These hybrid gene complexes (CYP1A promoter-intron and reporter gene) were sensitive to the amount of added POP (Figure 10C) indicating that the promoter sequence was sufficient to regulate mRNA expression in response to pollution. Importantly, the promoters from the polluted New Bedford site had significantly greater gene expression than promoters from either clean, reference population (Figure 10D). Thus, polluted and both clean reference populations have functional differences due to DNA sequence variation (206). Interestingly, promoter-intron sequences exhibited extremely large sequence variation [~9% of the sites were polymorphic (206)] and the polluted promoter clades with both reference populations as well as with *F. grandis* suggesting they are independently evolved. Additionally, the outlier CYP1A SNP (red star Figure 10B) is found in both reference populations and is absent in approximately 16% of individuals from polluted New Bedford population (i.e. 16% are homozygous for the "clean reference allele") (205). Furthermore, none of the polymorphic sites were shared among all eight polluted promoters, and all of the variable DNA sites were found in the clean reference populations. Thus, even though the promoters from the polluted populations altered gene expression, this change was accomplished with distinct DNA sequence variations in each individual, and not parallel evolved changes. This pattern of multiple different genetic combinations influencing an adaptive change most likely arose because of the large amounts of standing genetic variation (Figure 10B), such that multiple variants may exist that affect function, but in a



Numbers are the count of SNPs which were signifcant for the each of the three statistical test, and shared among the three test.

Figure 10 Rapid local adaptation to pollution. Analyses of polluted populations using changes in allele frequencies for 354 SNPs defined by mass spectrophotometry. (A) In each of three comparisons a polluted population (P) was compared to two clean reference populations (C). The Venn diagram for these three triads (C-P-C) was used to identify statistically significant SNPs based on an outlier test (Outlier, unexpectedly large F_{ST}), environmental association of SNPs (Assoc.), and changes in minor allele frequencies (MAF). The red number is the number of SNPs that are significant in all three tests. (B) Maximum parsimony tree of the 24 CYP1A promoter-intron sequences use to test the effect of DNA sequence variation on gene expression. Blue highlights are sequences from polluted New Bedford populations. Red stars are for sequences with derived outlier SNP. (C) Induction of gene expression with exposure to persistent organic pollutants (POP) in cells in culture for CYP1A promoter from polluted and clean populations. (D) Average pollution-induced gene expression from CYP1A promoter from the two clean (green and blue) and the polluted New Bedford population (red). Letters represent post-hoc analysis indicating that the polluted New Bedford is significantly different from both clean populations, and there is no significant difference between the clean populations.

redundant manner. This is illustrative of how, among different individuals, different polymorphisms may achieve the same functional phenotype. With different sets of variants subjected to selection, it would lead to an incomplete sweep of any single mutation; this result is analogous to what happens during polygenic adaptation, such that small allele frequency changes underlie adaptive shifts in phenotype.

A limitation of the just discussed Williams and Oleksiak (205) study is that it only examined hundreds of SNPs. It is possible that much more shared DNA sequence variation exists in polluted populations. To investigate the presence of many shared DNA variants, whole genomes were sequenced in 384 individuals from 4 polluted and 4 clean reference populations [48 individuals from each population, Figure 11A (157)]. For all four polluted populations, *F. heteroclitus* showed little change in mRNA expression when exposed to POPs, whereas *F. heteroclitus* from clean reference populations were transcriptionally responsive to POP exposure (Figure 11B). That is, all polluted populations share a common phenotype: the absence of mRNA response to POP exposure, especially for genes involved in the AHR pathway (Figure 11C). These mRNA expression patterns were related to changes in the *F. heteroclitus* genome (157) (Figures 11C and 11D).

The AHR protein is a transcription factor that is activated by binding PAHs, PCBs, and other xenobiotics, which then changes mRNA expression and POP toxicity (Figure 11C). By binding PAHs, AHR is released from AIP (AHR interacting protein), moves into the nucleus, and associates with ARNT (aryl hydrocarbon receptor nuclear translocator) to form a complex that regulates the mRNA expression by interacting with DNA that binds AHR (AHR response elements or AHREs) (63, 183). One of the more common mRNAs regulated by AHR is CYP1A. A similar phenotypic response occurs in four polluted *F. heteroclitus* populations—they lack enhanced transcription of genes regulated by AHR

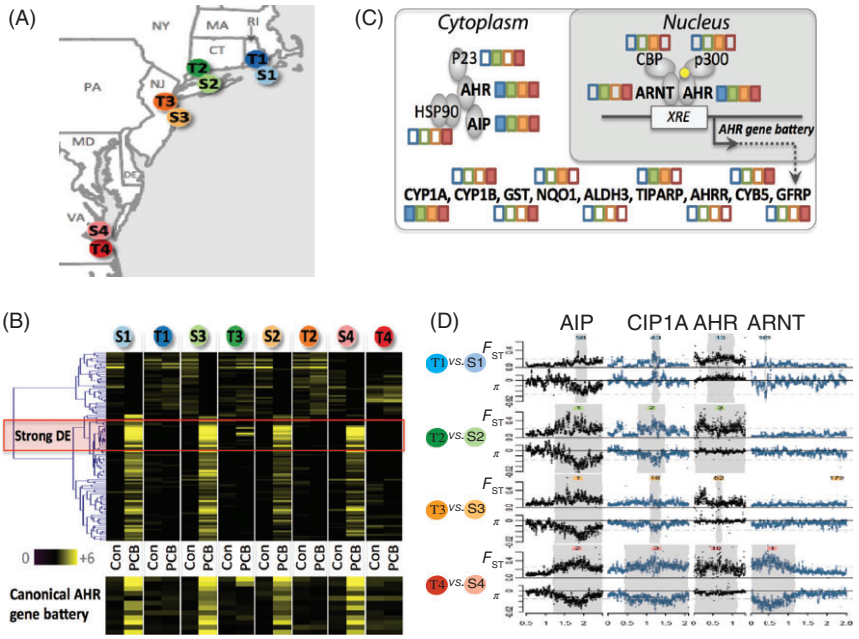


Figure 11 Genomics of adaptation to recent anthropogenic pollution. (A) Four pairs of populations were sampled: for each pair, one population inhabits highly polluted marine environments and individuals are tolerant to POPs (T), and the second population is in a clean, nonpolluted reference site and individuals are sensitive to POPs (S). (B) Pairs of mRNA expression for controls and POP exposure among tolerant (T) and sensitive (S) populations. Each population has mRNA expression for two sets of conditions: control and exposure to POP. In each row is the relative expression of an mRNA, with high expression as bright yellow. The lower panel highlights genes activated by ligand-bound AHR protein. (C) Diagram of AHR signaling pathway including co-regulators and transcriptional targets. Color boxes are color coded for location defined in (A). Filled boxes are genes identified as evolving by natural selection. (D) F_{ST} values and π (pi, nucleotide diversity) between tolerant (T) and sensitive (S) populations. Gray shading highlights DNA sequences with unusually large significant F_{ST} values, extreme pi values, or both. These regions of the genome also have significant Tajima's D (not shown).

(Figure 11B). While, thousands of DNA variants with signatures of natural selection occur in each of the four polluted populations, most are not shared among polluted populations. However, all four populations share changes in the DNA sequences for the AHR, AIP, ARNT, and CYP1A genes (Figure 11D). The DNA sequences in these regions exhibit significant changes indicative of evolution by natural selection: large F_{ST} values between polluted and clean reference populations, low nucleotide divergence, and significant negative Tajima's D (Figure 11D). These data showing locally polluted *F. heteroclitus* populations with greater survivorship to POPs have derived patterns of AHR-dependent mRNA expression and patterns of DNA sequence variation indicative of evolution by natural selection. Overall, these data are indicative of rapid adaptive evolution.

These data are similar to the original DNA sequence analyses using 354 SNPs: considering the genetic variants associated with pollution resistance, there are many more DNA variants associated with each polluted population than shared among all four populations. Additionally, while several genes are consistently involved in adaptive genomic changes among the four polluted populations (154), few of the selectively important variants within these genes are the same (Figure 11D). Instead, some populations have unique sets of genomic variants, even in genes that have shared signatures of selection, and adaptive alleles have not swept to 100% frequency. These patterns are consistent with many functionally redundant alleles offering a similar selective advantage segregating within adapted populations. The observation that there are many selectively important DNA sequence variants among many genes, that these sequence variants differ among polluted populations, and none of them are swept to 100% within a population, is indicative of polygenic adaptation. Furthermore, different functionally redundant unlinked variants that influence a single gene are segregating within populations and when adaptive they resemble polygenic adaptation: SNPs will not be sweep to 100% frequency.

One of the important lessons from these empirical results is that when multiple populations converge on the same adaptive polygenic phenotype, it is likely that the underlying adaptive alleles will be different in different populations. These results are consistent with a variety of theoretical studies discussed above (23, 170, 209). Furthermore, even if a trait has a simple

genetic basis, in different populations that converge on the same adaptive phenotype different allelic variants may be responsible because of genetic redundancy. Therefore, we suggest that in *F. heteroclitus* populations with high levels of genetic diversity, functional redundancy and polygenic adaptation commonly contribute to adaptive variation for physiological traits. This form of adaptive evolution results in incomplete sweeps of adaptive genetic variation within a population, and allele frequency changes at different sets of loci in different populations. Importantly, polygenic adaptation with redundancy maintains genetic polymorphism within and among populations, and thus provides the genetic variation necessary for future evolutionary change.

To summarize local, rapid adaptation in F. heteroclitus

- Rapid (<100 s of generations) adaptive evolution involves many genes.
- *F. heteroclitus* populations subjected to recent, anthropogenic changes in the thermal environment demonstrate enhanced survival to high temperatures associated with adaptive changes in mRNA expression and DNA sequence variation.
- *F. heteroclitus* populations living in highly polluted waters have evolved adaptive changes that enhance survivorship and are insensitive to the negative effects of pollution on development and metabolism relative to individuals from nonpolluted populations.
- Polluted populations exhibit adaptive changes in the AHR pathway that is observable in patterns of mRNA expression and DNA sequence variation across the genome.
- The adaptive responses to both recent changes in thermal and polluted environments arose from standing genetic variation and are best described as polygenic—involving many different genetic changes, not all of which are shared among all adapted individuals, such that allele frequency change at any given locus rarely proceeds to fixation within an adapted population.

Epistatic Evolution between Nuclear and Mitochondrial Genomes

Epistasis is when the functional influence of variation at one gene depends on the genetic variation at another gene. This limits the pace of adaptive evolution by requiring allelic variants in separate genes to occur together without recombination. Thus, epistatic evolution is the process by which the selection on an allele is influenced by the genetic variation at another gene. The study of epistatic evolution is challenging because of the large number of potential interactions and the difficulty of inferring phenotypes from the underlying additive effects of genetic variation (114, 115). One important example of the potential for epistatic interactions is the genes that form the OxPhos pathway. The OxPhos pathway involves proteins that are encoded in both the nuclear and mitochondrial genomes. As in nearly all animals, the mitochondrial genome is maternally inherited and thus different from the nuclear genome, which contains alleles from both parents. Importantly, the mitochondrial and nuclear genes that form the OxPhos pathway are responsible for most aerobic ATP production. The OxPhos pathway is comprised of 5 enzyme complexes with approximately 91 proteins; the mitochondrial genome encodes 13 of these proteins while the nuclear genome encodes 78 (Figure 12A). The interactions among these 91 proteins in the OxPhos pathway are sensitive to acute and chronic temperature exposures (161, 162). It is the DNA sequence variation between the two genomes (nuclear and mitochondrial) that ultimately is responsible for the observed variation in OxPhos (described below).

Northern *F. heteroclitus* populations have evolved a mitochondrial genome that differs from the southern type by five amino acid substitutions among the 13 mitochondrial encoded proteins (198). Yet, the DNA sequence changes between the northern and southern mitochondrial genome provide little evidence for adaptation (198). For quantitative physiological traits, *F. heteroclitus* individuals from northern and southern populations reveal subtle OxPhos metabolic differences that are most obvious at low temperatures (35, 36, 65, 83, 164). OxPhos metabolism is quantified by measuring mitochondrial oxygen consumption, or State 3 respiration (an integrative measure of ADP and substrate-dependent mitochondrial respiration). Importantly, OxPhos metabolic differences between *F. heteroclitus* with northern and southern mitotypes alter acute temperature responses (13) (Figure 12B): when acclimated to 28°C, individuals with southern mitotypes are much more sensitive to acute temperature changes than individuals with the northern mitotypes. Yet, interactions between nuclear and mitochondrial genomes are not addressed in these studies between individuals with northern and southern mitotypes.

To examine the epistatic evolution in *F. heteroclitus* involving the mitochondrial and nuclear genomes, we compared DNA sequence variation in nuclear genomes between individuals with northern or southern mitotypes. Fortunately, local populations in northern New Jersey, just south of the Hudson River, have nearly equal frequencies of both mitotypes (79). In these populations, the bimodal distribution of nuclear allelic variation with significantly lower number of heterozygotes is indicative of natural selection favoring a specific combination of nuclear and mitochondrial genes (120). The functional effect of this interaction and the role of specific genes involved were determined by measuring OxPhos physiology among individuals in a single population containing both mitotypes (13, 14).

Within and between the two mitotypes found in a single panmictic population the variation in nuclear genotypes alters

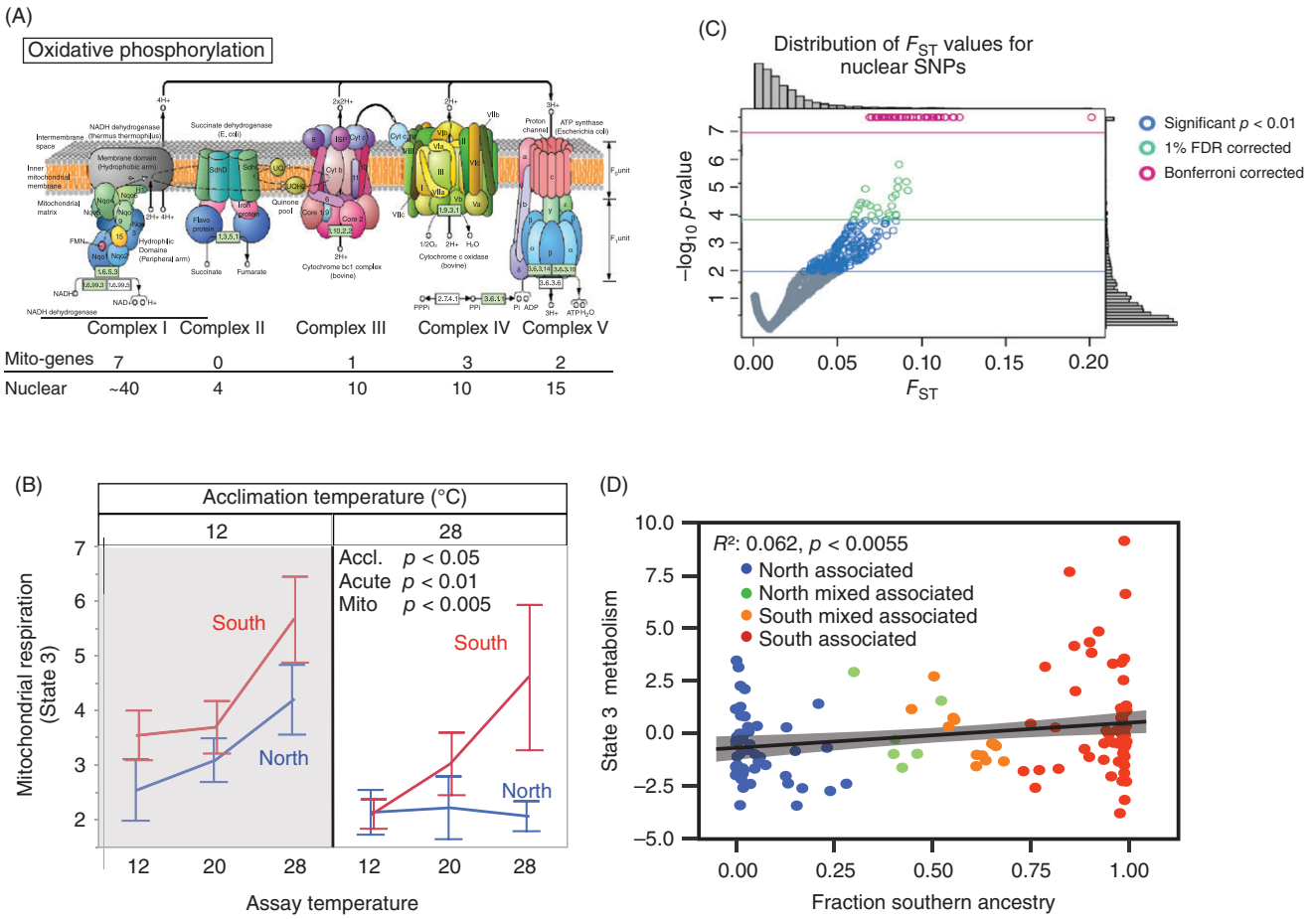


Figure 12 Epistatic adaptive evolution. (A) Oxidative Phosphorylation pathway and the number of protein subunits encoded by mitochondrial and nuclear genomes (95). (B) Mitochondrial OxPhos dependent respiration (State 3) measured in *Fundulus heteroclitus* from a single New Jersey population. Individuals were acclimated to either 12 or 28°C and had either the northern or southern mitochondrial haplotype. Acclimation, acute (assay temperature), and mitochondrial effects were all significant. (C) Distribution of wF_{ST} values for 11,705 nuclear SNPs calculated between the two mitochondrial haplotypes within the single population. Plot contains wF_{ST} values and corresponding negative log10 p-values (e.g. $-\log10(0.01) = 2$). Blue values are significant with a p-value less than 0.01, green values are significant with a 1% FDR correction, and purple values are significant with a Bonferroni correction. Histograms show wF_{ST} and p-value distributions. (D) Mitochondrial OxPhos dependent respiration (State 3) as a function of the fraction of southern nuclear alleles. State 3 is the residual from a mixed model with body mass, acclimation, and assay temperatures. Individuals with greater than 75% northern nuclear alleles and the northern mitochondria are blue, individuals with less than 75% northern nuclear alleles, and the northern mitochondria are green. Individuals with less than 75% southern nuclear alleles and with southern mitochondria are orange. Individuals with greater than 75% southern alleles and the southern mitochondria are red.

OxPhos (14). This result relied on evolutionary analyses that identify unexpectedly large, significant differences in nuclear allele frequencies between the two mitotypes (Figure 12C). In a single panmictic population (where individuals randomly breed), nuclear alleles should not vary between mitotypes. The variation in allele frequencies between individuals with different mitotypes relative to the total variation is F_{ST}, and because it is within a population, it was denoted as wF_{ST} (14). Examining approximately 11,000 different sites with DNA polymorphism (SNPs), revealed 349 outlier SNPs: SNPs with significantly large and statistically unlikely wF_{ST} values (Figure 12C). With little admixture and much support for random mating, the most parsimonious explanation for the allele frequency differences between the two mitotypes is that they are evolving by natural selection. These data suggest that the different mitotypes and nuclear genome are evolving epistatically. To test this hypothesis one can test whether these 349 outlier SNPs have functional effects.

Functionally, individuals with both "southern" nuclear and mitochondrial genomes have higher OxPhos metabolism (State 3) than individuals with both northern nuclear and mitochondrial genomes (Figure 12D). Yet, the evidence for epistatic evolution is among individuals with northern mitochondria and a high frequency of southern nuclear alleles, or *vice versa* (southern mitochondria with many northern alleles). We call these nuclear SNP alleles "northern" or "southern" because they are most frequently associated with the northern or southern mitochondrial haplotypes. Among individuals with a mixture of southern and northern mitochondrial

and nuclear genotypes, OxPhos metabolism is related to the frequency of "southern" nuclear alleles (Figure 12D): individuals with northern mitochondria and more southern nuclear genes had higher OxPhos metabolism while individuals with southern mitochondria and more northern nuclear genes had lower OxPhos metabolism (Figure 12D). These data (12–14, 83, 120) indicate that the interaction between the mitochondrial haplotype and nuclear genotype influences the survival of individuals as well as OxPhos physiology.

What is amazing about the epistatic evolution between the nuclear and mitochondrial genomes of *F. heteroclitus* is that it would require selection at every generation. Not surprisingly, no nuclear DNA sequence variation is fixed (at a frequency of 100%) for either mitochondrial haplotype. Instead, allele frequency differences are lower than 26%. Thus, significantly large wF_{ST} values are due to the small variance in nuclear genes between each haplotype. Additionally, because no individual has all SNPs associated with either the northern or southern mitochondrial haplotype, it suggests that only a subset of SNPs is necessary for epistatic adaptation. In conclusion, these data suggest high standing genetic variation influences mitochondrial function and only some of this variation is necessary to cause the adaptive change in OxPhos metabolism.

Assortative mating (88) provides an alternative explanation for the significant divergence between 349 outlier SNPs. For example, if the 349 SNPs modulate mating cue that contributed to assortative mating, then assortative mating could result in outlier wF_{ST}'s. Yet, this would not explain the physiological differences among individuals with different proportions of northern or southern nuclear 349 SNPs. Nor would assortative mating explain why the wF_{ST} values exceed the F_{ST} values for the same 349 SNPs between populations. It possible that an assortative mating cue is related to mitochondrial function and this would explain the data. That is, individuals with northern mitotype prefer individuals with both northern nuclear alleles and mitotypes because of a mating cue associated functional divergence in OxPhos function. This epistatic interaction would evolve due to sexual selection and would not necessarily be an adaptive physiology.

To summarize the evidence for F. heteroclitus *epistatic evolution*

- In northern New Jersey, populations have nearly equal frequencies of "northern" and "southern" type mitochondrial haplotypes that have 5 amino acid differences among the 13 mitochondrial-encoded proteins.
- In this population, individuals with the northern type mitochondria have different nuclear allele frequencies relative to individuals with the southern type mitochondria.
- The large significant differences in nuclear allele frequencies are indicative of natural selection altering allele frequencies in these genes due to epistatic evolution.
- Nuclear genes with large significant allele frequency differences alter OxPhos physiology depending on an individual's mitochondrial haplotype.
- These data indicate a significant epistatic interaction between mitochondrial and nuclear genomes that is evolving by natural selection, suggesting epistatic adaptation.

Fine-scale Evolution Within a Population to Microhabitats

Watching thousands of *F. heteroclitus* move in and out with the tides, where intertidal creeks are typically dry at low tide as they drain into basins, supports the concept that a local estuary is a single panmictic population. Yet, within these estuaries are shallow permanent ponds (0.5–1 m deep) where individuals are likely to be resident (5). These microhabitats (basins, inter-tidal creeks, and ponds) are less than a few 100 m apart and exhibit significant environmental differences (Figure 13A). In a panmictic *F. heteroclitus* population that lays and fertilizes eggs in the high upper-tidal zone of saltwater marsh estuaries, genetic variation among these physically close microhabitats is unlikely. The hypothesis of genetic similarity is supported by the observation that in populations from three geographically separate estuaries, most DNA sequences have similar allele frequencies among basins, intertidal creeks, and permanent ponds (195). Specifically, among approximately 4700 SNPs a vast majority have similar allele frequencies. Yet, surprisingly, in all three separate estuaries, individuals are significantly different for 1% to 2% of SNPs among microhabitats (Figure 13B). For each separate estuary, SNPs with significant and large differences in allele frequencies between microhabitats were determined using three statistical tests (Figure 13B). SNPs that were significant for all three tests were considered "outlier SNPs." The magnitude of the difference in allele frequencies is demonstrated in the F_{ST} values: F_{ST} values for the outlier SNPs are much greater than either most other SNPs or for random permutation of the data (e.g. only 0.4% of 413 million random F_{ST} values exceed the lower 95% CI of the outlier SNPs Figure 13C) (195). Importantly, among outlier SNPs, allele frequency differences between microhabitats (<100 m apart) exceed the differences among estuaries (100,000 m apart, Figure 13D). Based on these data, unexpected allele frequency differences among microhabitats are most likely due to evolution by natural selection.

The three New Jersey populations (Figure 13A) are replicate study sites for microhabitats, and each population has significant and large outlier SNPs among microhabitats. Yet, no outlier SNP was shared among all three populations. However, among all three populations outlier SNPs were present in common genes either by occurring (i) in the same gene at a different position, (ii) in a duplicate gene or paralog, or (iii) among genes with similar annotations. In fact, 11 outlier

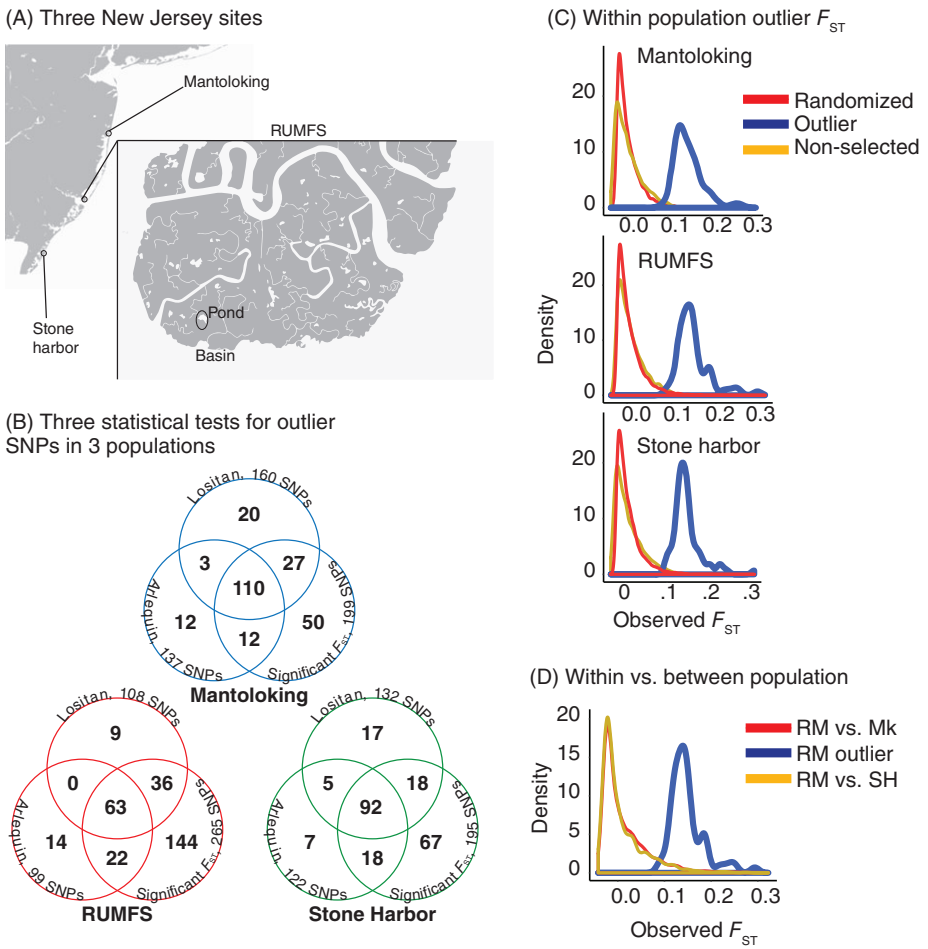


Figure 13 Fine-scale evolution among microhabitats. (A) Three New Jersey saltmarsh estuaries (Mantoloking, Rutgers University Marine Field Station, Stone Harbor) and an enlarged image of Rutgers University Marine Field Station with three microhabitats Basin (B), Creek (C), and Pond (P). The distance between microhabitats was never greater than 200 m and usually less than 50 m. (B) Evolutionary analyses among microhabitats for three populations, where each population has three analyses: (1) SNPs with significantly different F_{ST} values, (2) Lositan identified significant outlier SNPs, and (3) Arlequin identified significant outlier SNPs. Significant SNPs detected in all three analyses with joint FDR less than 1% were considered outlier SNPs. (C) Density of F_{ST} values within each of the three New Jersey saltmarsh populations. Plotted are large significant outlier SNPs (blue), 4352 nonoutlier SNPs (gold), and SNPs when population assignment is randomly permuted among microhabitats (red). (D) Density of outlier-SNP F_{ST} values within and among populations. Significant outlier SNP-specific F_{ST} values for within Rutgers University Marine Field Station (blue) and between Rutgers Marine Station and Stone Harbor (gold) or Mantoloking (red).

SNPs are in the same or paralogous gene. One example is outlier SNPs 33 bp apart in the same intron of vav guanine nucleotide exchange factor 2 (VAV2, signal transduction gene). Another example is outlier SNPs among different glutamate receptors found in all three replicate populations: (GRM4, GRM4′, and GRM5, where GRM4 and GRM4′ are two duplicate genes on different scaffolds). In addition to these examples, many more outlier SNPs share similar functions as defined by their gene ontology (195). These data suggest that many different and variable genes can respond to natural selection and distinct sets of adaptive genes occur in each population. Once again, these data support the concept of polygenic selection operating on standing genetic variation.

Polygenic selection from standing genetic variation also requires a lack of long-distance linkage (linkage disequilibrium, LD). By examining 15,259 SNPs, 8,180 (12% of all possible) have significant LD and 93% of these are less than 100 bp apart. Linkage between the 261 outlier SNPs with all the other 14, 998 SNPs reveal that 1192 have significant LDs, yet all but 67 or 94% of these are less than 100 bp apart. In the classic model of evolution, where new mutations quickly go to fixation when evolving by natural selection, outlier SNPs should be associated with large linkage blocks. Clearly, this is not what we find, nor would we expect to if the microhabitats within a population are inundated with admixed young of the year every spring. Instead, the short distances among all

SNPs and among 94% of outlier SNPs with significant LD is indicative of long-term standing genetic variation.

An alternative explanation to natural selection influencing these SNP allele frequencies is that individuals display homing behavior to specific microhabitats. Thus, offspring might choose their parent environments even though individuals randomly reproduce in the higher tidal zone. None of the data suggest widespread isolation (i.e. most SNPs have small F_{ST} values, Figure 13C). Yet, it is possible that the outlier SNPs modulate homing behavior. This would suggest natural selection favors alleles that alter behavioral choices, versus natural selection favoring these outlier SNPs because they enhance adult behavior. Either explanation still relies on polygenic selection but differs with respect to the life stage, physiological mechanism or functional trait evolving by selection.

The observation of significant and frequent allele frequency differences among microhabitats less than 100 m apart is surprising when at high tides individuals have access to the whole estuary. It is surprising because natural selection would have to exceed migration, which should be substantial. Yet, the strong selection is probable considering that ponds experience high daily temperature maxima, low nightly dissolved oxygen levels, and are more productive than basins or creeks (4, 5, 188). In contrast, tidal basins have lower diurnal fluxes in water temperature, salinity, and dissolved oxygen (78, 91, 187, 188). Large differences in SNP allele frequencies among microhabitats exceed those among geographically distant populations (Figure 13D) and are indicative of fine-scale genetic structure. The evolutionary importance of this observation depends on two points: (i) whether outlier F_{ST} values are statistical errors (type I errors) or (ii) whether neutral evolutionary processes are likely to explain the large F_{ST} outliers. A detailed analysis of SNP allele frequencies indicates that the genetic divergence among microhabitats is not due to these two points, and thus indicates adaptive divergence between microhabitats (195). What is missing are the functional consequences of genetic divergence. Functional data would greatly enhance this study, providing a separate dataset to support or reject the idea of rapid, fine-scale adaptation.

To summarize the data on fine-scale evolution

- Among three separate New Jersey estuaries, SNP allele frequencies are significantly different among microhabitats within each estuary.
- Among the outlier SNPs, none are shared in all three populations, but many of the outlier SNPs occur in the same or similar genes.
- Long-distance linkage is rare (nearly all SNPs exhibit LD of 100 bp or less).
- These patterns of divergence suggest polygenic selection from standing genetic variation on a fine ecological scale.

Discussion

F. heteroclitus has adapted to the thermal cline along the eastern seacoast of North America, along the salinity gradients in bays and estuaries and to locally heated, polluted, and variably saline environments. These adaptive changes have involved glycolytic or OxPhos enzymes, changes in genome-wide mRNA expression, and a large number of SNPs. The adaptive divergence in physiological traits and the genes that influence these traits occur at different spatial and temporal scales: along the eastern seacoast of North America or among bays evolving for thousands of years and in local environments evolving for a few decades or less.

One of the genes responsible for adaptation along the eastern seacoast is LDH-B. The two LDH-B alleles have different biochemical properties that alter reaction rates, glycolysis, swimming, and developmental rates (Figures 2 and 3). The different biochemical properties between the two LDH-B alleles are associated with nonsynonymous substitutions and analyses of sequence divergence among populations reject the null hypothesis of neutral divergence (Table 1). Along this continental coast are also promoter DNA sequence variations evolving by natural selection that alter LDH-B mRNA expression (Figure 4). Thus, the adaptations associated with LDH-B are related to both promoter variation influencing mRNA expression and two amino acid substitutions influencing enzyme catalysis. Neither the promoter with either amino acid substitution or the two amino acids are in LD. LDH-B evolving by natural selection is most similar to the classic theory of adaptive divergence where a single locus of large effect has large differences in allele frequencies creating a nearly fixed difference between populations.

Beyond the single LDH-B locus, genome-wide patterns of mRNA expression have a surprisingly high amount of variation among individuals within populations (~20% of mRNAs are significantly different among individuals). Two separate analyses indicate adaptive differences in mRNA expression among populations along the eastern seacoast (Figures 5 and 6). Importantly, the variation in mRNA expression alters cardiac metabolism (Figure 7) but in a complex way: the specific genes of importance responsible for physiological traits vary among individuals: in some individuals, mRNA expression in glycolysis is more important and in others, mRNA expression in OxPhos is more important. These data suggest that the expression of several different genes influence adaptive physiological traits and, in contrast to LDH-B, that evolution by natural selection occurs by polygenic adaptation from standing genetic variation (multiple genes of small effect where many polymorphisms result in similar adaptive phenotypic change).

In addition to adaptive patterns of mRNA expression (Figures 4–7), local populations have significant differences in SNP allele frequencies. Outlier SNPs, those evolving by natural selection, typically have allele frequency differences among populations with large, improbable F_{ST} values. In studies of local populations that are a few hundred kilometers

to less than 100 m apart, outlier SNPs have intermediate allele frequencies. That is, across environments with differences in temperature, pollution levels or salinity, outlier SNPs display allele frequency differences that are neither fixed nor nearly fixed (i.e. MAF > 10%). Importantly, among replicate polluted populations few of the outlier SNPs are shared; however, many evolutionarily significant SNPs are associated with the same gene, pathway, or metabolic function (Figures 10 and 11). The functional significance of this is seen in the similar adaptive patterns of mRNA expression where all four polluted populations share the same phenotypic pattern of mRNA expression even though few associated DNA sequence polymorphisms are shared among the adapted populations. On the finest scale, three replicate saltmarsh estuaries have outlier SNPs that are indicative of adaptation to microhabitats within each estuary, yet none of the outlier SNPs are shared among the three replicate saltmarshes (Figure 13). Yet among microhabitats and similar to studies on polluted populations, while none of the outlier SNPs are the same in all three estuaries, many occur in the same gene, in duplicate genes or in genes with a similar function. Based on this research we conclude that adaptive evolution can occur readily and rapidly when operating on large standing genetic variation among many genes that can influence physiological traits. These observations of genetic redundancy and polygenic adaptation enhance our understanding of evolution and physiological adaptation, thus informing both biological and medical scientists about genotype-phenotype relationships.

In addition to polygenic adaptation, changes in physiological function are associated with epistatic evolution (Figure 12), which adds another level of complexity to adaptive evolution. Natural selection is most effective via additive genetic variation. The reliance on the co-occurrence of alleles at two independent loci implies that many individuals will have the wrong combination of alleles. This reduces the effectiveness of natural selection. Pleiotropic effects (genetic variation that influences two or more traits) also occur. Thus, while polygenic adaptation can involve many different combinations of genes, some combinations may be more frequent because they epistatically work better together to create the adaptive phenotype and have few negative pleiotropic effects. This has the dual outcome of maintaining polymorphisms by selecting for combinations of epistatic genes and limiting the number of different adaptively important genotypes.

These patterns of genetic divergence and their impact on physiological function as demonstrated in *F. heteroclitus* enhance our understanding of evolutionary genetics. In 1966, Lewontin and Hubby (110) provided data on 18 proteins and concluded that excessive genetic variation exists within a species, but they could not resolve the evolutionary forces responsible for this variation. Sixty years later, it is not clear that we have resolved this dilemma (34, 85, 104, 131). Yet, the studies described here for the teleost *F. heteroclitus* suggest that within a species many polymorphisms may exist due to natural selection acting on different spatial and temporal scales and among genomes (mitochondrial and nuclear) within individuals. Most importantly, multiple solutions for physiological adaptation occur on small geographic and small time scales, and depend on many polymorphic genes where combinations of different subsets of allelic variants in different individuals can cause equivalent adaptive phenotypes. To identify these polygenic solutions and thus address Lewontin's and Hubby's (110) fundamental problem about the maintenance and evolution of genetic polymorphism required studies that used selectively different environments and combined genomic analyses with physiological determinations. The attributes of closely related polymorphic populations living in diverse environments make the small teleost fish *F. heteroclitus* an excellent model for these studies and for understanding adaptive physiology.

Conclusion

For F. heteroclitus populations, evolutionarily adaptive divergence occurs for transcriptional, biochemical, metabolic, osmotic, and whole animal physiologies. The strength of the studies reviewed here is the use of an evolutionary approach contrasting large populations (large N, number of individuals) in habitats where individuals are well suited to their environment. Large population sizes mean that small selection coefficients ($>1/2N_e$) are effective and thus we can expect fine-tuning of adaptive phenotypes. The variation in the native environments with associated changes in physiological functions provides a "natural experimental" approach to identify genetic changes associated with environmental variation and then relate them to functionally important physiological traits. It is the combination of functional quantitative physiological analyses with evolutionary analyses that have provided insights into the genetics of adaptation in this system.

The evolutionary insights from the research on *F. heteroclitus* are many. For LDH-B, it is its pleiotropic effect on hemoglobin oxygen binding (and not the direct effect on metabolism *per se*) that is one of the advantages of northern LDH-B^b in cold waters. Additionally, the effect of the single LDH-B gene on adaptive physiology involves many independent evolved changes: nonsynonymous substitutions that alter enzyme kinetics, changes in DNA that alter transcription binding sites altering the amount LDH-B expression, and the acclimation response. Most importantly, in contrast to the importance of a single gene of large effect (like LDH-B), many evolved physiological traits appear to be polygenic adaptations that originated from standing genetic variation. Moreover, they involve many genes of small effect with many redundant adaptive solutions. For example, the adaptive variation in cardiac metabolism is associated with changes in the utilization of different substrates (glucose, fatty acid, or secondary metabolites) and variation among individuals in which substrate is most important. Furthermore, among different groups of individuals, the rate of

utilization for any single substrate is dependent on different pathways (glycolysis, TCA, or OxPhos). In general, results from *F. heteroclitus* are consistent with the general finding that evolutionary change that enhances Darwinian fitness to local environments are polygenic. Importantly, populations that have similar adaptive phenotypes associated with similar environments do not share all of the same adaptive genetic polymorphisms. The observation that individuals within a population that have adapted to the local environment do not share all of the same adaptive alleles at multiple unlinked loci is best explained by a past and ongoing process of polygenic adaptation. Polygenic adaptation is likely in populations with large amounts of standing genetic variation because multiple polymorphisms are likely to be functionally redundant and available for natural selection. Similarly, multiple polymorphisms within a single gene may also be functionally redundant, such that natural selection often does not fix any single genetic variant within a population. Thus, because different individuals can have different adaptive alleles it follows that different populations may not share the same genetic variation responsible for an adaptive phenotype, even if the same gene is involved.

The most important evolutionary insight from studies of adaptive evolution in *F. heteroclitus* may be the support for the hypothesis that polygenic adaptation in highly polymorphic populations is responsible for adaptive evolutionary changes in physiological traits. We define polygenic adaptation with sufficient standing genetic variation as the evolution by natural selection of a trait influenced by many (tens to thousands) of variable genes of small effect, where many of the different variants are functionally redundant. If polygenic adaptation is achieved by changes in diverse and redundant polymorphisms, then it may require large and highly polymorphic population characteristic of species like *F. heteroclitus* (156). Yet, highly polymorphic populations (populations with high amounts of standing genetic variation) and polygenic adaptation are not independent. They are not independent because polygenic adaptation both requires many genes that have similar effects on the phenotype (redundant variation) and provides a mechanism to maintain genetic variation within populations. Thus, with multiple solutions for an adaptive "problem" (ways to make an adaptive phenotype), few of the genetic polymorphisms will go to fixation—most genetic polymorphisms will remain polymorphic, even those with a selective advantage. That is, this high level of standing genetic variation occurs because with polygenic adaptation few alleles are swept to fixation during selection and thus, there is a great reduction in the rate or near elimination of purge genetic variation even under strong selection (153, 170, 209). Thus, under this scenario, one need not only invoke balancing selection for heterozygotes or epistasis to account for large reservoirs of standing variation in some species like *F. heteroclitus*; instead, natural selection favoring multiple solutions can retain functionally important variation.

Although polygenic adaptation is an attractive explanation for the patterns we observe in *F. heteroclitus* and can explain most of the data, it is not the only possible explanation. The variation in genetic response associated with adaptation within a population could arise because of pleiotropy, epistasis, or unaccounted gene-by-environment interactions. Similarly, the lack of the parallel genetic variation among populations adapted to similar environments could reflect subtle differences in the environment (e.g. greater salinity with higher temperatures, or different combinations of organic pollution), differences in ecological communities (e.g. fewer predators), or gene-by-environmental interactions. Thus, although we believe the existing evidence strongly suggests that polygenic adaptation is important in this system, we acknowledge the need for further research. In particular, future research should focus on testing the predictions of polygenic adaptation. *Fundulus* genomic research that examines equivalent adaptive physiological changes among populations should contribute to these objectives, and *Fundulus* should continue to serve as a model system to study evolutionary adaptation, particularly when studying complex physiological traits in natural populations.

References

1. Able KW, Balletto JH, Hagan SM, Jivoff PR, Strait K. Linkages between salt marshes and other nekton habitats in Delaware Bay, USA. *Rev Fish Sci* 15: 1-61, 2007.
2. Able KW, Castagna M. Aspects of an undescribed reproductive behavior in *Fundulus heteroclitus* (Pisces: Cyprinodontae) from Virginia. *Chesapeake Sci* 16: 282, 1975.
3. Able KW, Felley JD. Geographical variation in *Fundulus-heteroclitus* – Tests for concordance between egg and adult morphologies. *Am Zool* 26: 145-157, 1986.
4. Able KW, Hagan SM, Brown SA. Habitat use, movement, and growth of young-of-the-year *Fundulus* spp. in southern New Jersey salt marshes: Comparisons based on tag/recapture. *J Exp Mar Biol Ecol* 335: 177-187, 2006.
5. Able KW, Vivian DN, Petruzzelli G, Hagan SM. Connectivity among salt marsh subhabitats: Residency and movements of the mummichog (*Fundulus heteroclitus*). *Estuaries Coast* 35: 743-753, 2012.
6. Adams SM, Lindmeier JB, Duvernell DD. Microsatellite analysis of the phylogeography, Pleistocene history and secondary contact hypotheses for the killifish, *Fundulus heteroclitus*. *Mol Ecol* 15: 1109-1123, 2006.
7. Airaksinen R, Rantakokko P, Eriksson JG, Blomstedt P, Kajantie E, Kiviranta H. Association between type 2 diabetes and exposure to persistent organic pollutants. *Diabetes Care* 34: 1972-1979, 2011.
8. Anholt RRH, Mackay TFC. The road less traveled: From genotype to phenotype in flies and humans. *Mamm Genome* 29: 5-23, 2018.
9. Arnold SJ. Morphology, performance and fitness. *Am Zool* 23: 347-361, 1983.
10. Arnold SJ. Genetic correlation and the evolution of physiology. In: Feder ME, Bennett AE, Burggren WW, Huey RB, editors. *New Directions in Ecological Physiology*. Cambridge, New York, etc: Cambridge University Press, 1987, p. 189-215.
11. Arnot JA, Armitage JM, McCarty LS, Wania F, Cousins IT, Toose-Reid L. Toward a consistent evaluative framework for POP risk characterization. *Environ Sci Technol* 45: 97-103, 2011.
12. Baris TZ, Blier PU, Pichaud N, Crawford DL, Oleksiak MF. Gene by environmental interactions affecting oxidative phosphorylation and thermal sensitivity. *Am J Physiol Regul Integr Comp Physiol* 311: R157-R165, 2016.
13. Baris TZ, Crawford DL, Oleksiak MF. Acclimation and acute temperature effects on population differences in oxidative phosphorylation. *Am J Physiol Regul Integr Comp Physiol* 310: R185-R196, 2016.
14. Baris TZ, Wagner DN, Dayan DI, Du X, Blier PU, Pichaud N, Oleksiak MF, Crawford DL. Evolved genetic and phenotypic differences due to mitochondrial-nuclear interactions. *PLoS Genet* 13: e1006517, 2017.
15. Bartholomew GA. Interspecific comparison as a tool for ecological physiologists. In: Feder ME, Bennett AF, Burggren WW, Huey RB, editors. *New Directions in Ecological Physiology*. Cambridge, New York, etc: Cambridge University Press, 1987, p. 11-37.
16. Bello SM, Franks DG, Stegeman JJ, Hahn ME. Acquired resistance to Ah receptor agonists in a population of Atlantic killifish (*Fundulus*

heteroclitus) inhabiting a marine superfund site: In vivo and in vitro studies on the inducibility of xenobiotic metabolizing enzymes. *Toxicol Sci* 60: 77-91, 2001.
17. Bennett AF, Huey RB. Studying the evolution of physiological performance. *Oxford Surv Evol Biol* 7: 251-284, 1990.
18. Berg JJ, Coop G. A population genetic signal of polygenic adaptation. *PLoS Genet* 10: e1004412, 2014.
19. Bernardi G, Sordino P, Powers DA. Concordant mitochondrial and nuclear DNA phylogenies for populations of the teleost fish *Fundulus heteroclitus*. *Proc Natl Acad Sci U S A* 90: 9271-9274, 1993.
20. Betancur RR, Orti G, Pyron RA. Fossil-based comparative analyses reveal ancient marine ancestry erased by extinction in ray-finned fishes. *Ecol Lett* 18: 441-450, 2015.
21. Bloom AL. Sea level and coastal changes. In: Wright WEJ, editor. *The Holocene*. Minneapolis: University of Minnesota, 1983, p. 26-41.
22. Bloom AL. Sea level and coastal morphology of the United States through the late Wisconsin glacial maximum. In: Porter SC, editor. *The Late Pleistocene*. Minneapolis: University of Minnesota, 1983, p. 215-229.
23. Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: From polygenic to omnigenic. *Cell* 169: 1177-1186, 2017.
24. Bozinovic G, Oleksiak MF. Genomic approaches with natural fish populations from polluted environments. *Environ Toxicol Chem* 30: 283-289, 2011.
25. Bozinovic G, Sit TL, Di Giulio R, Wills LF, Oleksiak MF. Genomic and physiological responses to strong selective pressure during late organogenesis: Few gene expression changes found despite striking morphological differences. *BMC Genomics* 14: 779, 2013.
26. Bozinovic G, Sit TL, Hinton DE, Oleksiak MF. Gene expression throughout a vertebrate's embryogenesis. *BMC Genomics* 12: 132, 2011.
27. Brennan RS, Healy TM, Bryant HJ, La MV, Schulte PM, Whitehead A. Genome-wide selection scans integrated with association mapping reveal mechanisms of physiological adaptation across a salinity gradient in killifish. *bioRxiv*, 2018. DOI: 10.1101/254854.
28. Brennan RS, Hwang R, Tse M, Fangue NA, Whitehead A. Local adaptation to osmotic environment in killifish, *Fundulus heteroclitus*, is supported by divergence in swimming performance but not by differences in excess post-exercise oxygen consumption or aerobic scope. *Comp Biochem Physiol A Mol Integr Physiol* 196: 11-19, 2016.
29. Brieuc MSO, Waters CD, Drinan DP, Naish KA. A practical introduction to Random Forest for genetic association studies in ecology and evolution. *Mol Ecol Resour* 18: 755-766, 2018.
30. Brown PO, Botstein D. Exploring the new world of the genome with DNA microarrays. *Nat Genet* 21: 33-37, 1999.
31. Burnett KG, Bain LJ, Baldwin WS, Callard GV, Cohen S, Di Giulio RT, Evans DH, Gomez-Chiarri M, Hahn ME, Hoover CA, Karchner SI, Katoh F, Maclatchy DL, Marshall WS, Meyer JN, Nacci DE, Oleksiak MF, Rees BB, Singer TD, Stegeman JJ, Towle DW, Van Veld PA, Vogelbein WK, Whitehead A, Winn RN, Crawford DL. *Fundulus* as the premier teleost model in environmental biology: Opportunities for new insights using genomics. *Comp Biochem Physiol Part D Genomics Proteomics* 2: 257-286, 2007.
32. Cashon RE, Van Beneden RJ, Powers DA. Biochemical genetics of *Fundulus heteroclitus* (L.). IV. Spatial variation in gene frequencies of Idh-A, Idh-B, 6-Pgdh-A, and Est-S. *Biochem Genet* 19: 715-728, 1981.
33. Cavalli-Sforza LL, Edwards AWF. Phylogenetic analysis: Models and estimation procedures. *Evolution* 21: 550-570, 1967.
34. Charlesworth B. Causes of natural variation in fitness: Evidence from studies of Drosophila populations. *Proc Natl Acad Sci U S A* 112: 1662-1669, 2015.
35. Chung DJ, Bryant HJ, Schulte PM. Thermal acclimation and subspecies-specific effects on heart and brain mitochondrial performance in a eurythermal teleost (*Fundulus heteroclitus*). *J Exp Biol* 220: 1459-1471, 2017.
36. Chung DJ, Morrison PR, Bryant HJ, Jung E, Brauner CJ, Schulte PM. Intraspecific variation and plasticity in mitochondrial oxygen binding affinity as a response to environmental temperature. *Sci Rep* 7: 16238, 2017.
37. Churchill GA. Fundamentals of experimental design for cDNA microarrays. *Nat Genet* 32 (Suppl): 490-495, 2002.
38. Clark BW, Bone AJ, Di Giulio RT. Resistance to teratogenesis by F1 and F2 embryos of PAH-adapted *Fundulus heteroclitus* is strongly inherited despite reduced recalcitrance of the AHR pathway. *Environ Sci Pollut Res Int* 21: 13898-13908, 2014.
39. Cline RM, editor. *Late Quaternary Paleoceanography and Paleoclimatology*. Boulder: Geological Society of America, 1976.
40. Crawford DL, Oleksiak MF. The biological importance of measuring individual variation. *J Exp Biol* 210: 1613-1621, 2007.
41. Crawford DL, Pierce VA, Segal JA. Evolutionary physiology of closely related taxa: Analyses of enzyme expression. *Am Zool* 39: 389-400, 1999.
42. Crawford DL, Place AR, Powers DA. Clinal variation in the specific activity of lactate dehydrogenase-B. *J Exp Zool* 255: 110-113, 1990.
43. Crawford DL, Powers DA. Molecular-basis of evolutionary adaptation at the lactate dehydrogenase-B locus in the fish *Fundulus-heteroclitus*. *Proc Natl Acad Sci U S A* 86: 9365-9369, 1989.
44. Crawford DL, Powers DA. Molecular adaptation to the thermal environment: Genetic and physiological mechanisms. In: Clegg MT, O'Brien SJ, editors. *Molecular Evolution: Proceeding of a UCLA Colloquium 1989*. New York: Wiley-Liss, 1990, p. 213-222.
45. Crawford DL, Powers DA. Evolutionary adaptation to different thermal environments via transcriptional regulation. *Mol Biol Evol* 9: 806-813, 1992.
46. Crawford DL, Segal JA, Barnett JL. Evolutionary analysis of TATA-less proximal promoter function. *Mol Biol Evol* 16: 194-207, 1999.
47. Crow JF. Some possibilities for measuring selection intensities in man. *Am Anthropol* 60: 1-13, 1958.
48. Dayan DI, Crawford DL, Oleksiak MF. Phenotypic plasticity in gene expression contributes to divergence of locally adapted populations of *Fundulus heteroclitus*. *Mol Ecol* 24: 3345-3359, 2015.
49. Dayan DI, Du X, Baris TZ, Wagner DN, Crawford DL, Oleksiak MF. Population genomics of rapid evolution in natural populations: Polygenic selection in response to power station thermal effluents. *BMC Evol Biol* 19: 61, 2019.
50. DiMichele L, Paynter KT, Powers DA. Evidence of lactate dehydrogenase-B allozyme effects in the teleost, *Fundulus heteroclitus*. *Science* 253: 898-900, 1991.
51. DiMichele L, Powers DA. LDH-B genotype-specific hatching times of *Fundulus heteroclitus* embryos. *Nature* 296: 563-564, 1982.
52. Dimichele L, Powers DA. Physiological-basis for swimming endurance differences between Ldh-B genotypes of *Fundulus-heteroclitus*. *Science* 216: 1014-1016, 1982.
53. Dimichele L, Powers DA. Developmental and oxygen-consumption rate differences between lactate dehydrogenase-B genotypes of *Fundulus-heteroclitus* and their effect on hatching time. *Physiol Zool* 57: 52-56, 1984.
54. Dimichele L, Powers DA. Allozyme variation, developmental rate, and differential mortality in the teleost Fundulus-heteroclitus. *Physiol Zool* 64: 1426-1443, 1991.
55. DiMichele L, Westerman ME. Geographic variation in development rate between populations of the teleost *Fundulus heteroclitus*. *Mar Biol* 128: 1-7, 1997.
56. Donnelly JP, Driscoll NW, Uchupi E, Keigwin LD, Schwab WC, Thieler ER, Swift SA. Catastrophic meltwater discharge down the Hudson Valley: A potential trigger for the Intra-Allerod cold period. *Geology* 33: 89-92, 2005.
57. Du X, Crawford DL, Nacci DE, Oleksiak MF. Heritable oxidative phosphorylation differences in a pollutant resistant *Fundulus heteroclitus* population. *Aquat Toxicol* 177: 44-50, 2016.
58. Du X, Crawford DL, Oleksiak MF. Effects of anthropogenic pollution on the oxidative phosphorylation pathway of hepatocytes from natural populations of *Fundulus heteroclitus*. *Aquat Toxicol* 165: 231-240, 2015.
59. Duforet-Frebourg N, Luu K, Laval G, Bazin E, Blum MG. Detecting genomic signatures of natural selection with principal component analysis: Application to the 1000 genomes data. *Mol Biol Evol* 33: 1082-1093, 2016.
60. Duvernell DD, Lindmeier JB, Faust KE, Whitehead A. Relative influences of historical and contemporary forces shaping the distribution of genetic variation in the Atlantic killifish, *Fundulus heteroclitus*. *Mol Ecol* 17: 1344-1360, 2008.
61. Eisen MB, Brown PO. DNA arrays for analysis of gene expression. *Methods Enzymol* 303: 179-205, 1999.
62. Evans DH, Piermarini PM, Choe KP. The multifunctional fish gill: Dominant site of gas exchange, osmoregulation, acid-base regulation, and excretion of nitrogenous waste. *Physiol Rev* 85: 97-177, 2005.
63. Fader KA, Zacharewski TR. Beyond the aryl hydrocarbon receptor: Pathway interactions in the hepatotoxicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin and related compounds. *Curr Opin Toxicol* 2: 36-41, 2017.
64. Fangue NA, Hofmeister M, Schulte PM. Intraspecific variation in thermal tolerance and heat shock protein gene expression in common killifish, *Fundulus heteroclitus*. *J Exp Biol* 209: 2859-2872, 2006.
65. Fangue NA, Richards JG, Schulte PM. Do mitochondrial properties explain intraspecific variation in thermal tolerance? *J Exp Biol* 212: 514-522, 2009.
66. Feder ME, Bennett AF, Huey RB. Evolutionary physiology. *Annu Rev Ecol Syst* 31: 315-341, 2000.
67. Fisher B. Most unwanted. *Environ Health Perspect* 107: A18-A23, 1999.
68. Fraccalvieri D, Soshilov AA, Karchner SI, Franks DG, Pandini A, Bonati L, Hahn ME, Denison MS. Comparative analysis of homology models of the AH receptor ligand binding domain: Verification of structure-function predictions by site-directed mutagenesis of a nonfunctional receptor. *Biochemistry* 52: 714-725, 2013.

69. Gaggiotti OE, Bekkevold D, Jorgensen HB, Foll M, Carvalho GR, Andre C, Ruzzante DE. Disentangling the effects of evolutionary, demographic, and environmental factors influencing genetic structure of natural populations: Atlantic herring as a case study. *Evolution* 63: 2939-2951, 2009.
70. Garcia-Ramos G, Kirkpatrick M. Genetic models of adaptation and gene flow in peripheral populations. *Evolution* 51: 21-28, 1997.
71. Garland T Jr, Adolph SC. Why not to do two-species comparative studies: Limitations on inferring adaptation. *Physiol Zool* 67: 797-828, 1994.
72. Garland T Jr, Carter PA. Evolutionary physiology. *Annu Rev Physiol* 56: 579-621, 1994.
73. Garland T, Adolph SC. Physiological differentiation of vertebrate populations. *Annu Rev Ecol Syst* 22: 193-228, 1991.
74. Gonzalez-Villasenor LI, Powers DA. Mitochondrial-DNA restriction-site polymorphisms in the teleost *Fundulus Heteroclitus* support secondary intergradation. *Evolution* 44: 27-37, 1990.
75. Griffith RW. Environment and salinity tolerance in the genus *Fundulus*. *Copeia* 1974: 319-331, 1974.
76. Haas HL, Freeman CJ, Logan JM, Deegan L, Gaines EF. Examining mummichog growth and movement: Are some individuals making intra-season migrations to optimize growth? *J Exp Mar Biol Ecol* 369: 8-16, 2009.
77. Haldane JBS. The cost of natural selection. *J Genet* 55: 511-524, 1957.
78. Halpin PM. Habitat use patterns of the mummichog, *Fundulus heteroclitus,* in New England. I. Intramarsh variation. *Estuaries* 20: 618-625, 1997.
79. Haney RA, Dionne M, Puritz J, Rand DM. The comparative phylogeography of east coast estuarine fishes in formerly glaciated sites: Persistence versus recolonization in Cyprinodon variegatus ovinus and *Fundulus heteroclitus* macrolepidotus. *J Hered* 100: 284-296, 2009.
80. Hartl DL, Clark AG. *Principles of Population Genetics*. Sunderland, MA: Sinauer Associates, 2007.
81. Hartl DL, Dykhuizen DE, Dean AM. Limits of adaptation: The evolution of selective neutrality. *Genetics* 111: 655-674, 1985.
82. Healy TM, Brennan RS, Whitehead A, Schulte PM. Tolerance traits related to climate change resilience are independent and polygenic. *Glob Chang Biol* 24: 5348-5360, 2018.
83. Healy TM, Bryant HJ, Schulte PM. Mitochondrial genotype and phenotypic plasticity of gene expression in response to cold acclimation in killifish. *Mol Ecol* 26: 814-830, 2017.
84. Healy TM, Schulte PM. Factors affecting plasticity in whole-organism thermal tolerance in common killifish (*Fundulus heteroclitus*). *J Comp Physiol B* 182: 49-62, 2012.
85. Henn BM, Botigue LR, Bustamante CD, Clark AG, Gravel S. Estimating the mutation load in human genomes. *Nat Rev Genet* 16: 333-343, 2015.
86. Hermisson J, Pennings PS. Soft sweeps: Molecular population genetics of adaptation from standing genetic variation. *Genetics* 169: 2335-2352, 2005.
87. Hermisson J, Wagner GP. The population genetic theory of hidden variation and genetic robustness. *Genetics* 168: 2271-2284, 2004.
88. Hill GE. Mitonuclear mate choice: A missing component of sexual selection theory? *Bioessays* 40, 2018. DOI: 10.1002/bies.201700191.
89. Hochachka PW, Somero GN. *Biochemical Adaptation, Mechanism and Process in Physiological Evolution*. New York, NY: Oxford University Press, 2002, p. 408.
90. Huang W, Richards S, Carbone MA, Zhu D, Anholt RR, Ayroles JF, Duncan L, Jordan KW, Lawrence F, Magwire MM, Warner CB, Blankenburg K, Han Y, Javaid M, Jayaseelan J, Jhangiani SN, Muzny D, Ongeri F, Perales L, Wu YQ, Zhang Y, Zou X, Stone EA, Gibbs RA, Mackay TF. Epistasis dominates the genetic architecture of Drosophila quantitative traits. *Proc Natl Acad Sci U S A* 109: 15553-15559, 2012.
91. Hunter KL, Fox MG, Able KW. Habitat influences on reproductive allocation and growth of the mummichog (*Fundulus heteroclitus*) in a coastal salt marsh. *Mar Biol* 151: 617-627, 2007.
92. Imbrie J, McIntyre A, Moore TCJ. The ocean around North America at the last glacial maximum. In: Porter SC, editor. *The Late Pleistocene*. Minneapolis: University of Minnesota, 1983, p. 230-238.
93. Jha AR, Zhou D, Brown CD, Kreitman M, Haddad GG, White KP. Shared genetic signals of hypoxia adaptation in Drosophila and in high-altitude human populations. *Mol Biol Evol* 33: 501-517, 2016.
94. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genet* 11: 94, 2010.
95. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28: 27-30, 2000.
96. Karami-Mohajeri S, Abdollahi M. Toxic influence of organophosphate, carbamate, and organochlorine pesticides on cellular metabolism of lipids, proteins, and carbohydrates: A systematic review. *Hum Exp Toxicol* 30: 1119-1140, 2011.
97. Kerr MK, Churchill GA. Statistical design and the analysis of gene expression microarray data. *Genet Res* 77: 123-128, 2001.
98. Kimura M. Evolutionary rate at the molecular level. *Nature* 217: 624-626, 1968.
99. Kimura M. Recent development of the neutral theory viewed from the Wrightian tradition of theoretical population genetics. *Proc Natl Acad Sci U S A* 88: 5969-5973, 1991.
100. Kimura M, Crow JF. The number of alleles that can be maintained in a finite population. *Genetics* 49: 725-738, 1964.
101. Kneib RT. The role of *Fundulus-Heteroclitus* in salt-marsh trophic dynamics. *Am Zool* 26: 259-269, 1986.
102. Knox JC. Responses of river systems to holocene climates. In: Wright WEJ, editor. *The Holocene*. Minneapolis: University of Minnesota, 1983, p. 26-41.
103. Kong SW, Lee IH, Leshchiner I, Krier J, Kraft P, Rehm HL, Green RC, Kohane IS, MacRae CA, MedSeq P. Summarizing polygenic risks for complex diseases in a clinical whole-genome report. *Genet Med* 17: 536-544, 2015.
104. Kreitman M. The neutral theory is dead. Long live the neutral theory. *Bioessays* 18: 678-683; discussion 683, 1996.
105. Laporte M, Pavey SA, Rougeux C, Pierron F, Lauzent M, Budzinski H, Labadie P, Geneste E, Couture P, Baudrimont M, Bernatchez L. RAD sequencing reveals within-generation polygenic selection in response to anthropogenic organic and metal contamination in North Atlantic Eels. *Mol Ecol* 25: 219-237, 2016.
106. Lauerman T. The Functional Significance of the Amino Acid Differences between Allelic Isozymes of Heart Type Lactate Dehydrogenase in *Fundulus heteroclitus*. Ph.D. Thesis. In: Department of BiologyJohns Hopkins University, 1990.
107. Lee DH, Lee IK, Song K, Steffes M, Toscano W, Baker B, Jacobs D. A strong dose-response relation between serum concentrations of persistent organic pollutants and diabetes: Results from the National Health and Examination Survey 1999-2002. *Diabetes Care* 29: 1638-1644, 2006.
108. Lee DH, Lind L, Jacobs DR Jr, Salihovic S, van Bavel B, Lind PM. Associations of persistent organic pollutants with abdominal obesity in the elderly: The Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) study. *Environ Int* 40: 170-178, 2012.
109. Lee SH, Ra JS, Choi JW, Yim BJ, Jung MS, Kim SD. Human health risks associated with dietary exposure to persistent organic pollutants (POPs) in river water in Korea. *Sci Total Environ* 470-471: 1362-1369, 2014.
110. Lewontin RC, Hubby JL. A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of Drosophila pseudoobscura. *Genetics* 54: 595-609, 1966.
111. Li QQ, Loganath A, Chong YS, Tan J, Obbard JP. Persistent organic pollutants and adverse health effects in humans. *J Toxicol Environ Health A* 69: 1987-2005, 2006.
112. Lim S, Cho YM, Park KS, Lee HK. Persistent organic pollutants, mitochondrial dysfunction, and metabolic syndrome. *Ann N Y Acad Sci* 1201: 166-176, 2010.
113. Lotrich VA. Summer home range and movements of *Fundulus Heteroclitus* (Pisces: Cyprinodontidae) in a tidal creek. *Ecology* 56: 191-198, 1975.
114. Mackay TF. Epistasis for quantitative traits in Drosophila. *Methods Mol Biol* 1253: 47-70, 2015.
115. Mackay TF, Moore JH. Why epistasis is important for tackling complex human disease genetics. *Genome Med* 6: 124, 2014.
116. McBryan TL, Anttila K, Healy TM, Schulte PM. Responses to temperature and hypoxia as interacting stressors in fish: Implications for adaptation to environmental change. *Integr Comp Biol* 53: 648-659, 2013.
117. McBryan TL, Healy TM, Haakons KL, Schulte PM. Warm acclimation improves hypoxia tolerance in *Fundulus heteroclitus*. *J Exp Biol* 219: 474-484, 2016.
118. McDonald JH, Kreitman M. Adaptive protein evolution at the Adh locus in Drosophila. *Nature* 351: 652-654, 1991.
119. McIntyre A, Kipp NG, Be AWH, Crowley T, Kellogg T, Gardner JV, Prell W, Ruddiman WF. Glacial North America 18,000 years ago: A climate reconstruction. In: Cline RM, Hays JD, editors. *Late Quaternary Paleoceanography and Paleoclimatology*. Boulder: Geological Society of America, 1976, p. 43-76.
120. McKenzie JL, Dhillon RS, Schulte PM. Steep, coincident, and concordant clines in mitochondrial and nuclear-encoded genes in a hybrid zone between subspecies of Atlantic killifish, *Fundulus heteroclitus*. *Ecol Evol* 6: 5771-5787, 2016.
121. Meyer J, Di Giulio R. Heritable adaptation and fitness costs in killifish (*Fundulus heteroclitus*) inhabiting a polluted estuary. *Ecol Appl* 13: 490-503, 2003.
122. Meyer JN, Wassenberg DM, Karchner SI, Hahn ME, Di Giulio RT. Expression and inducibility of aryl hydrocarbon receptor pathway genes in wild-caught killifish (*Fundulus heteroclitus*) with different contaminant-exposure histories. *Environ Toxicol Chem* 22: 2337-2343, 2003.

123. Mickelson DM, Clayton L, Fullerton DS, Borns HWJ. The late Wisconsin glacial record of laurentide ice sheet in the United States. In: Porter SC, editor. *The Late Pleistocene*. Minneapolis: University of Minnesota, 1983, p. 3-37.
124. Mitton JB, Koehn RK. Genetic organization and adaptive response of allozymes to ecological variables in *Fundulus heteroclitus*. *Genetics* 79: 97-111, 1975.
125. Morin RP, Able KW. Patterns of geographic-variation in the egg morphology of the fundulid fish, *Fundulus-heteroclitus*. *Copeia* 1983: 726-740, 1983.
126. Nacci D, Champlin D, Coiro L, McKinney R, Jayaraman S. Predicting the occurrence of genetic adaptation to dioxin like compounds in populations of the estuarine fish *Fundulus heteroclitus*. *Environ Toxicol Chem* 21: 1525-1532, 2002.
127. Nacci D, Coiro L, Champlin D, Jayaraman S, McKinney R, Gleason TR, Munns WR, Specker JL, Cooper KR. Adaptations of wild populations of the estuarine fish *Fundulus heteroclitus* to persistent environmental contaminants. *Mar Biol* 134: 9-17, 1999.
128. Nacci D, Huber M, Champlin D, Jayaraman S, Cohen S, Gauger E, Fong A, Gomez-Chiarri M. Evolution of tolerance to PCBs and susceptibility to a bacterial pathogen (Vibrio harveyi) in Atlantic killifish (*Fundulus heteroclitus*) from New Bedford (MA, USA) harbor. *Environ Pollut* 157: 857-864, 2009.
129. Nacci DE, Champlin D, Jayaraman S. Adaptation of the estuarine fish *Fundulus heteroclitus* (Atlantic Killifish) to polychlorinated biphenyls (PCBs). *Estuaries Coast* 33: 853-864, 2010.
130. Nei M, Suzuki Y, Nozawa M. The neutral theory of molecular evolution in the genomic era. *Annu Rev Genomics Hum Genet* 11: 265-289, 2010.
131. Ohta T. The nearly neutral theory of molecular evolution. *Annu Rev Ecol Syst* 23: 263-286, 1992.
132. Oleksiak MF. Marine genomics: Insights and challenges. *Brief Funct Genomics* 15: 331-332, 2016.
133. Oleksiak MF, Churchill GA, Crawford DL. Variation in gene expression within and among natural populations. *Nat Genet* 32: 261-266, 2002.
134. Oleksiak MF, Crawford Douglas L. Functional genomics in fishes, insights into physiological complexity. In: Evan D, Claiborne J, editors. *The Physiology of Fishes*. Boca Raton: CRC Press, 2006, p. 523-550.
135. Oleksiak MF, Kolell KJ, Crawford DL. Utility of natural populations for microarray analyses: Isolation of genes necessary for functional genomic studies. *Mar Biotechnol (NY)* 3: S203-S211, 2001.
136. Oleksiak MF, Roach JL, Crawford DL. Natural variation in cardiac metabolism and gene expression in *Fundulus heteroclitus*. *Nat Genet* 37: 67-72, 2005.
137. Palumbi SR, Sidell BD, Vanbeneden R, Smith GD, Powers DA. Glucosephosphate isomerase (Gpi) of the teleost *Fundulus-heteroclitus* (Linnaeus) – Isozymes, allozymes and their physiological roles. *J Comp Physiol* 138: 49-57, 1980.
138. Paynter KT, Dimichele L, Hand SC, Powers DA. Metabolic implications of Ldh-B genotype during early development in *Fundulus-heteroclitus*. *J Exp Zool* 257: 24-33, 1991.
139. Pierce VA, Crawford DL. Effect of acclimation on glycolytic enzymes in Atlantic and Gulf Coast *Fundulus* species. *Am Zool* 35: 99A, 1995.
140. Pierce VA, Crawford DL. Variation in the glycolytic pathway: The role of evolutionary and physiological processes. *Physiol Zool* 69: 489-508, 1996.
141. Pierce VA, Crawford DL. Phylogenetic analysis of glycolytic enzyme expression. *Science* 276: 256-259, 1997.
142. Pierce VA, Crawford DL. Phylogenetic analysis of thermal acclimation of the glycolytic enzymes in the genus *Fundulus*. *Physiol Zool* 70: 597-609, 1997.
143. Place AR, Powers DA. Genetic bases for protein polymorphism in *Fundulus heteroclitus* (L.). I. Lactate dehydrogenase (Ldh-B), malate dehydrogenase (Mdh-A), glucosephosphate isomerase (Gpi-B), and phosphoglucomutase (Pgm-A). *Biochem Genet* 16: 577-591, 1978.
144. Place AR, Powers DA. Genetic variation and relative catalytic efficiencies: Lactate dehydrogenase B allozymes of *Fundulus heteroclitus*. *Proc Natl Acad Sci U S A* 76: 2354-2358, 1979.
145. Place AR, Powers DA. Kinetic characterization of the lactate dehydrogenase (LDH-B4) allozymes of *Fundulus heteroclitus*. *J Biol Chem* 259: 1309-1318, 1984.
146. Podrabsky JE, Javillonar C, Hand SC, Crawford DL. Intraspecific variation in aerobic metabolism and glycolytic enzyme expression in heart ventricles. *Am J Physiol Regul Integr Comp Physiol* 279: R2344-R2348, 2000.
147. Powell WH, Bright R, Bello SM, Hahn ME. Developmental and tissue-specific expression of AHR1, AHR2, and ARNT2 in dioxin-sensitive and -resistant populations of the marine fish *Fundulus heteroclitus*. *Toxicol Sci* 57: 229-239, 2000.
148. Powers DA, Lauerman T, Crawford D, DiMichele L. Genetic mechanisms for adapting to a changing environment. *Annu Rev Genet* 25: 629-659, 1991.
149. Powers DA, Place AR. Biochemical genetics of *Fundulus* heterolitus (L.). I. Temporal and spatial variation in gene frequencies of Ldh-B, Mdh-A, Gpi-B, and Pgm-A. *Biochem Genet* 16: 593-607, 1978.
150. Powers DA, Ropson I, Brown DC, Vanbeneden R, Cashon R, Gonzalezvillasenor LI, Dimichele JA. Genetic-variation in *Fundulus-heteroclitus* – Geographic-distribution. *Am Zool* 26: 131-144, 1986.
151. Powers DA, Schulte PM. Evolutionary adaptations of gene structure and expression in natural populations in relation to a changing environment: A multidisciplinary approach to address the million-year saga of a small fish. *J Exp Zool* 282: 71-94, 1998.
152. Powers DA, Smith M, Gonzalez-Villasenor I, DiMichele L, Crawford DL, Bernardi G, Lauerman T. A multidisciplinary approach to the selectionist/neutralist controversy using the model teleost, *Fundulus heteroclitus*. In: Futuyma D, Antonovics J, editors. *Oxford Surveys in Evolutionary Biology*. New York, NY: Oxford University Press, 1993, p. 43-108.
153. Pritchard JK, Pickrell JK, Coop G. The genetics of human adaptation: Hard sweeps, soft sweeps, and polygenic adaptation. *Curr Biol* 20: R208-R215, 2010.
154. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959, 2000.
155. Przeworski M, Coop G, Wall JD. The signature of positive selection on standing genetic variation. *Evolution* 59: 2312-2323, 2005.
156. Reid NM, Jackson CE, Gilbert D, Minx P, Montague MJ, Hampton TH, Helfrich LW, King BL, Nacci DE, Aluru N, Karchner SI, Colbourne JK, Hahn ME, Shaw JR, Oleksiak MF, Crawford DL, Warren WC, Whitehead A. The landscape of extreme genomic variation in the highly adaptable Atlantic killifish. *Genome Biol Evol* 9: 659-676, 2017.
157. Reid NM, Proestou DA, Clark BW, Warren WC, Colbourne JK, Shaw JR, Karchner SI, Hahn ME, Nacci D, Oleksiak MF, Crawford DL, Whitehead A. The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science* 354: 1305-1308, 2016.
158. Rockman MV. The QTN program and the alleles that matter for evolution: All that's gold does not glitter. *Evolution* 66: 1-17, 2012.
159. Ropson IJ, Brown EC, Powers DA. Biochemical genetics of *Fundulus Heteroclitus* (L.). Vi. Geographical variation in the gene frequencies of 15 loci. *Evolution* 44: 16-26, 1990.
160. Ruzzin J. Public health concern behind the exposure to persistent organic pollutants and the risk of metabolic diseases. *BMC Public Health* 12: 298, 2012.
161. Schulte PM. What is environmental stress? Insights from fish living in a variable environment. *J Exp Biol* 217: 23-34, 2014.
162. Schulte PM. The effects of temperature on aerobic metabolism: Towards a mechanistic understanding of the responses of ectotherms to a changing environment. *J Exp Biol* 218: 1856-1866, 2015.
163. Schulte PM, Glemet HC, Fiebig AA, Powers DA. Adaptive variation in lactate dehydrogenase-B gene expression: Role of a stress-responsive regulatory element. *Proc Natl Acad Sci U S A* 97: 6597-6602, 2000.
164. Schulte PM, GomezChiarri M, Powers DA. Structural and functional differences in the promoter and 5' flanking region of Ldh-B within and between populations of the teleost *Fundulus heteroclitus*. *Genetics* 145: 759-769, 1997.
165. Schultz ET, McCormick SD. Euryhalinity in an evolutionary context. *Euryhaline Fishes* 32: 477-533, 2013.
166. Scott GR, Rogers JT, Richards JG, Wood CM, Schulte PM. Intraspecific divergence of ionoregulatory physiology in the euryhaline teleost *Fundulus heteroclitus*: Possible mechanisms of freshwater adaptation. *J Exp Biol* 207: 3399-3410, 2004.
167. Segal JA, Barnett JL, Crawford DL. Functional analyses of natural variation in Sp1 binding sites of a TATA-less promoter. *J Mol Evol* 49: 736-749, 1999.
168. Segal JA, Crawford DL. LDH-B enzyme expression: The mechanisms of altered gene expression in acclimation and evolutionary adaptation. *Am J Physiol* 267: R1150-R1153, 1994.
169. Segal JA, Schulte PM, Powers DA, Crawford DL. Descriptive and functional characterization of variation in the *Fundulus heteroclitus Ldh-B* proximal promoter. *J Exp Zool* 275: 355-364, 1996.
170. Sella G, Barton NH. Thinking about the evolution of complex traits in the era of genome-wide association studies. *Annu Rev Genomics Hum Genet* 20: 461-493, 2019.
171. Shaw JR, Hampton TH, King BL, Whitehead A, Galvez F, Gross RH, Keith N, Notch E, Jung D, Glaholt SP, Chen CY, Colbourne JK, Stanton BA. Natural selection canalizes expression variation of environmentally induced plasticity-enabling genes. *Mol Biol Evol* 31: 3002-3015, 2014.
172. Sidell BD, Johnston IA, Moerland TS, Goldspink G. The eurythermal myofibrillar protein complex of the mummichog (*Fundulus heteroclitus*): Adaptation to a fluctuating thermal enviroment. *J Comp Physiol B* 153: 167-173, 1983.
173. Skinner MA, Courtenay SC, Parker WR, Curry RA. Site fidelity of mummichogs (*Fundulus heteroclitus*) in an Atlantic Canadian Estuary. *Water Qual Res J Can* 40: 288-298, 2005.

174. Skinner MA, Courtenay SC, Parker WR, Curry RA. Stable isotopic assessment of site fidelity of mummichogs, *Fundulus heteroclitus*, exposed to multiple anthropogenic inputs. *Environ Biol Fishes* 94: 695-706, 2012.
175. Slatkin M. Gene flow and the geographic structure of natural populations. *Science* 236: 787-792, 1987.
176. Slatkin M. Gene flow and population structure. In: Real LA, editor. *Ecological Genetics*. Princeton, NJ: Princeton University Press, 1994, p. 3-17.
177. Smith KJ, Able KW. Dissolved oxygen dynamics in salt marsh pools and its potential impacts on fish assemblages. *Mar Ecol Prog Ser* 258: 223-232, 2003.
178. Smith MW, Chapman RW, Powers DA. Mitochondrial DNA analysis of Atlantic Coast, Chesapeake Bay, and Delaware Bay populations of the teleost *Fundulus heteroclitus* indicates temporally unstable distributions over geologic time. *Mol Mar Biol Biotechnol* 7: 79-87, 1998.
179. Smith MW, Glimcher MC, Powers DA. Genetic introgression of nuclear alleles between populations of the teleost *Fundulus heteroclitus*. *Mol Mar Biol Biotechnol* 1: 226-238, 1992.
180. Somero GN. Proteins and temperature. *Annu Rev Physiol* 57: 43-68, 1995.
181. Somero GN, Lockwood BL, Tomanek L. *Biochemical Adaptation: Response to Environmental Challenges from Life's Origins to the Anthropocene*. Oxford, UK: Oxford University Press, 2017, p. 108-136.
182. Srinivasan S, Bettella F, Mattingsdal M, Wang Y, Witoelar A, Schork AJ, Thompson WK, Zuber V, Schizophrenia Working Group of the Psychiatric Genomics Consortium TIHGC, Winsvold BS, Zwart JA, Collier DA, Desikan RS, Melle I, Werge T, Dale AM, Djurovic S, Andreassen OA. Genetic markers of human evolution are enriched in schizophrenia. *Biol Psychiatry* 80: 284-292, 2016.
183. Stockinger B, Di Meglio P, Gialitakis M, Duarte JH. The aryl hydrocarbon receptor: Multitasking in the immune system. *Annu Rev Immunol* 32: 403-432, 2014.
184. Storz JF, Bridgham JT, Kelly SA, Garland T Jr. Genetic approaches in comparative and evolutionary physiology. *Am J Physiol Regul Integr Comp Physiol* 309: R197-R214, 2015.
185. Strange RM, Burr BM. Intraspecific phylogeography of North American highland fishes: A test of the Pleistocene vicariance hypothesis. *Evolution* 51: 885-897, 1997.
186. Taylor MH, Michele LD, Leach GJ. Egg stranding in the life cycle of the Mummichog, *Fundulus heteroclitus*. *Copeia* 1977: 397, 1977.
187. Teo SLH, Able KW. Growth and production of the mummichog (*Fundulus heteroclitus*) in a restored salt marsh. *Estuaries* 26: 51-63, 2003.
188. Teo SLH, Able KW. Habitat use and movement of the mummichog (*Fundulus heteroclitus*) in a restored salt marsh. *Estuaries* 26: 720-730, 2003.
189. Turchin MC, Chiang CWK, Palmer CD, Sankararaman S, Reich D, Hirschhorn JN. Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nat Genet* 44: 1015-1019, 2012.
190. Valiela I, Wright JE, Teal JM, Volkmann SB. Growth, production and energy transformations in the salt-marsh killifish *Fundulus heteroclitus*. *Mar Biol* 40: 135-144, 1977.
191. Van Beneden RJ, Cashon RE, Powers DA. Biochemical genetics of *Fundulus heteroclitus* (L.). III. Inheritance of isocitrate dehydrogenase (Idh-A and Idh-B), 6-phosphogluconate dehydrogenase (6-Pgdh-A), and serum esterase (Est-S) polymorphisms. *Biochem Genet* 19: 701-714, 1981.
192. Vassy JL, Hivert MF, Porneala B, Dauriz M, Florez JC, Dupuis J, Siscovick DS, Fornage M, Rasmussen-Torvik LJ, Bouchard C, Meigs JB. Polygenic type 2 diabetes prediction at the limit of common variant detection. *Diabetes* 63: 2172-2182, 2014.
193. Vega GC, Wiens JJ. Why are there so few fish in the sea? *Proc Biol Sci* 279: 2323-2329, 2012.
194. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet* 90: 7-24, 2012.
195. Wagner DN, Baris TZ, Dayan DI, Du X, Oleksiak MF, Crawford DL. Fine-scale genetic structure due to adaptive divergence among microhabitats. *Heredity (Edinb)* 118: 594-604, 2017.
196. Wallace B. *Fifty Years of Genetic Load: An Odyssey*. Ithaca, NY: Cornell University Press, 1991, p. 174.
197. Wayne ML, Simonsen KL. Statistical tests of neutrality in the age of weak selection. *Trends Ecol Evol* 13: 236-240, 1998.
198. Whitehead A. Comparative mitochondrial genomics within and among species of killifish. *BMC Evol Biol* 9: 11, 2009.
199. Whitehead A. The evolutionary radiation of diverse osmotolerant physiologies in killifish (*Fundulus* sp.). *Evolution* 64: 2070-2085, 2010.
200. Whitehead A, Crawford DL. Neutral and adaptive variation in gene expression. *Proc Natl Acad Sci U S A* 103: 5425-5430, 2006.
201. Whitehead A, Crawford DL. Variation within and among species in gene expression: Raw material for evolution. *Mol Ecol* 15: 1197-1211, 2006.
202. Whitehead A, Roach JL, Zhang S, Galvez F. Genomic mechanisms of evolved physiological plasticity in killifish distributed along an environmental salinity gradient. *Proc Natl Acad Sci U S A* 108: 6193-6198, 2011.
203. Williams LM, Ma X, Boyko AR, Bustamante CD, Oleksiak MF. SNP identification, verification, and utility for population genetics in a non-model genus. *BMC Genet* 11: 32, 2010.
204. Williams LM, Oleksiak MF. Signatures of selection in natural populations adapted to chronic pollution. *BMC Evol Biol* 8: 282, 2008.
205. Williams LM, Oleksiak MF. Ecologically and evolutionarily important SNPs identified in natural populations. *Mol Biol Evol* 28: 1817-1826, 2011.
206. Williams LM, Oleksiak MF. Evolutionary and functional analyses of cytochrome P4501A promoter polymorphisms in natural populations. *Mol Ecol* 20: 5236-5247, 2011.
207. Wood CM, Marshall WS. Ion balance, acid-base regulation, and chloride cell-function in the common killifish, *Fundulus-heteroclitus* – A euryhaline estuarine teleost. *Estuaries* 17: 34-52, 1994.
208. Yang J, Lee T, Kim J, Cho MC, Han BG, Lee JY, Lee HJ, Cho S, Kim H. Ubiquitous polygenicity of human complex traits: Genome-wide analysis of 49 traits in Koreans. *PLoS Genet* 9: e1003355, 2013.
209. Yeaman S. Local adaptation by alleles of small effect. *Am Nat* 186 (Suppl 1): S74-S89, 2015.
210. Yu HY, Guo Y, Zeng EY. Dietary intake of persistent organic pollutants and potential health risks via consumption of global aquatic products. *Environ Toxicol Chem* 29: 2135-2142, 2010.