Evolutionary physiology and genomics in the highly adaptable killifish (*Fundulus heteroclitus*).

Douglas L. Crawford[1], Patricia M. Schulte2, Andrew Whitehead3 and Marjorie F. Oleksiak1

1: University of Miami, Rosenstiel School of Marine and Atmospheric Sciences, Miami FL 33149,
2: Department of Zoology, The University of British Columbia, 6270 University Blvd, Vancouver, British Columbia, Canada, V6T 1Z4

3: Department of Environmental Toxicology, University of California, Davis, CA 95616, USA

Running Head
Evolution within and among populations.

Key Words
Natural selection, polygenic traits, polygenic adaptation, standing genetic variation, Fundulus, teleost, evolution, physiology, cardiac, osmotic, thermal.

Abstract

By investigating evolutionary adaptations that affect physiological functions we can enhance our understanding about how organisms work, the importance of physiological traits, and the genes that affect these traits. This approach of investigating physiological adaptation in the teleost fish *Fundulus heteroclitus* has produced insights into (i) how protein polymorphisms affect swimming and development, (ii) the role of equilibrium enzymes in modulating metabolic flux, (iii) how variations 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* is a model system for studying adaptive evolution because it has large local population sizes, significant standing genetic variation, and populations are subjected to large ranges of environmental conditions that enhance the likelihood that adaptive evolution will occur. Based on evolutionary analyses contrasting neutral and adaptive evolution, we conclude that adaptive evolution can occur readily and rapidly in this system because it is polygenic: it depends on large standing genetic variation among many genes that can affect physiological traits. This polygenic adaptation with considerable polymorphism illustrates an important finding: there are often more than one genetic solution for similar adaptive phenotypes. One of the consequences of polygenic adaptive is that select genes will not go to fixation but remain polymorphic. These observations of polygenic adaptation enhance our understanding of how evolution and physiological adaptation progress thus informing both biological and medical scientists about genotype-phenotype relationships.

Didactic Synopsis

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

- These findings include:
  - Genetic variation between LDH-B alleles affects swimming and development.
  - DNA sequence variation affects on LDH-B enzyme expression.
  - Quantitative differences in glycolytic enzymes related to thermal environment and their affect on cardiac metabolism.
  - mRNA expression changes that are influenced by physiological acclimation affects on metabolism.
  - Population specific variation in DNA sequences and mRNA expression mitigates variation in temperature, pollution, and osmotic environments.
  - DNA sequence variation within and between nuclear and mitochondrial genomes and their affects on 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 proceeds by polygenic selection.
- If adaptation often proceeds by polygenic selection, this enhances our understanding of biochemical, physiological and evolutionary processes because it suggests that there are many genes of importance (i.e., that can influence biologically important traits).

***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 (*Fundulus heteroclitus)*. In these saltmarsh estuaries, *F. heteroclitus* is a dominant and ecologically important species (92) with biomasses that can exceed a dry mass of 160kg/ha (178) and with effective population sizes that can exceed 100,000 (55). These large effective population sizes make *F. heteroclitus* an outstanding model system for investigating adaptive evolution in physiological traits and the ways that animals adapt to a changing environment.

The estuarine environments that *F. heteroclitus* inhabit are highly variable for a wide range of abiotic factors. With diurnal tidal cycles, there are dramatic variations in the physical environment (1, 26, 152): temperatures can increase daily to > 30°C in upper estuaries or plunge to 12-15°C with incoming cold tides (161); salinity can vary from nearly freshwater due to heavy rains to salinities greater than seawater in desiccating ponds (1, 68, 195); oxygen concentrations can vary from anoxic to supersaturated (166). *F. heteroclitus'* ability to mount physiological responses to tolerate these conditions is well documented (26, 39, 43, 130, 152, 158, 187, 195).

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 (20, 26, 35, 143, 188). This makes *F. heteroclitus* an intriguing model in which to assess the potentially adaptive variation among populations and thus define the functional importance variation in these traits. Importantly*, F. heteroclitus* have several key characteristics that make this species particularly amenable to determining whether 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 define demographies.

Distribution across environments

Fundulus heteroclitus are distributed across a wide range of environmental conditions that are likely to impose strong selection on physiological traits. One of these conditions is temperature, where temperature affects the rate of biochemical and physiological processes and thus imposes strong selection to offset these effects (80, 169). 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* (40, 131, 139, 152, 188), and populations at the extremes of the species distribution have > 14°C difference in mean annual temperatures at 1m depth (139). Additional temperature differences result from power plant thermal effluents where thermally impacted local populations are 4-12°C warmer than neighboring populations (44). Extreme temperature variation can also be observed within a single estuary where shallow ponds are ~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 (124, 147, 194). 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 1,000s of kilometers away. These separate polluted sites allow populations to independently adapt to the pollution challenges and provide biological replicates to better understand adaptation (13, 26).

A third important environmental factor that varies among *F. heteroclitus* populations is salinity (187, 190). Salinity variation also imposes strong selection, having shaped the evolutionary history of aquatic organisms (190). The average salinity and variation in salinity across *F. heteroclitus'* distribution is geographically complex and is 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 two weeks later on the next spring tide (2, 174). 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 much smaller than a single drainage system with high site fidelity (5, 104). In a study of 1,499 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 (104). These fish also displayed strong site fidelity (returning to the same side of a creek after release over a 3 month period (104). Similarly, in a separate study, 97% of tagged individuals were found within 200 m of initial marking over two seasons (162), and stable isotopes indicate very few *F. heteroclitus* (3.4%) move more than 200 m in their lifetime (163). For 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 > 14,000 individuals over 17 months, where the authors concluded that despite physical connectivity of this watershed within and among different creek systems, there was 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 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, 55). Large populations are minimally impacted by neutral processes and are thus less affected by random 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.

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 mosaic of closely related taxa that have independent demographic that enhances 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 the ancestral trait 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 is dependent on the size of the population and the divergence time (188). 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, there are minimal demographic effects, and thus there are few neutral changes to obscure selectively advantageous ones. Thus, among closely related taxa, there are fewer and smaller neutral divergences than among distantly related taxa making it more likely to statistically reject the null hypothesis of neutral evolution (35, 36, 65, 188). This makes evolutionary analyses easier among closely related taxa than among distantly related taxa, especially those where speciation involved a transitory small population (bottleneck).

Figure 1. Demographic Patterns among *F. heteroclitus* Populations. **A.** Neighbor-joining tree based on microsatellites (6, 55, 188). **B.** Maximum parsimony tree for RFLP in mitochondria, LDH-B coding region, and 309 bp of Cytochrome B (15). 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 (191, 194).

Didactic Figure Legend 1. Populations of *F. heteroclitus* along the eastern North Atlantic seacoast are more closely related on either side of the Hudson River. This pattern is seen in 1A with microsatellites (neutral markers) where Maine (ME) and Connecticut (CT) cluster separately from New Jersey (NJ); even though NJ and CT are geographically closer. Other molecular markers provide similar data. These include LDH-B sequences, mitochondrial DNA and 354 variable sites in the nuclear genome.

Demography in *Fundulus heteroclitus*

Large local populations of *F. heteroclitus* subjected to biologically important environmental variation provide a valuable resource to investigate evolutionary adaptation and its physiological consequences. However, all evolutionary analyses must consider demographic parameters that 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 (63, 72, 89, 90, 121). Demography, like species' phylogenies, affects evolutionary changes such that more closely related populations will share evolved differences, both adaptive and neutral. Thus, when assessing 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,* the effects of neutral processes should be minimized because large populations are less affected by random drift. Yet, *F. heteroclitus'* extensive geographic distribution across long time scales has affected gene flow and population size, and thus neutral divergence contributes to population differentiation across their distribution (55, 164, 165) (Fig 1).

Although population sizes are large and evolutionarily stable, neutral divergence or adaptive divergence is affected by an important historical break between the northern and southern *F. heteroclitus populations* (64)*. F. heteroclitus'* genomes have extensive polymorphisms dominated by a sharp break (large change in allele frequencies) between 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 (15, 66, 71, 139, 168). 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 (17, 18, 34, 51, 83, 93, 110, 114, 173). Combining mitochondrial sequences and microsatellite data, the best explanation for the patterns of DNA sequence variation is that a large *F. heteroclitus* refuge existed north of the Hudson River, and individuals from this refuge re-invaded locations just north of the Hudson (6, 71). These historical events most likely affected the evolutionary divergences among *F. heteroclitus* populations (6, 55, 139, 143, 167) (Fig. 1).

The historical break between northern and southern populations would enhance any diverging evolutionary forces (neutral or adaptive) by minimizing migration (64). For example, it is clear that the evolved divergence is related to the demographic isolation of northern and southern populations as seen among neutral microsatellites (Fig. 1A), While the microsatellite data provides evidence to support the expected historical break (6, 55), 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 (15, 66)(Fig. 1B). Additionally, among 354 nuclear single nucleotide polymorphisms (SNPs) there is clear population structure that can be captured by 3-principal components (191) (Fig 1C). These SNPs are shared among both northern and southern populations yet their frequency provides readily identifiable population structure both within populations 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 affect how readily we can discern adaptive *versus* neutral divergence between populations. For example, the LDH-B pattern of allelic divergence (Fig. 1B) could be due to natural selection, yet it is indistinguishable from neutral microsatellite patterns (Fig. 1A and B). This suggests that simple patterns of genetic variation in DNA sequence alone are insufficient to distinguish adaptive from neutral divergence. This review of *Fundulus'* adaptation describes evolutionary studies that are enhanced by physiological measures linking functional consequences of nucleotide divergence that affect fitness 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 based using only nucleotide divergence or physiology separately.

Adaptive Evolution

To appreciate adaptation in *Fundulus* heteroclitus it is important to define adaptation and how polygenic traits affect adaptation. Adaptation is the evolution by natural selection of biological functions that improve the performance, health, longevity or reproduction of an organism. That is, adaption is the evolution of genetic changes because these changes alter phenotypic traits that improve 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 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 (versus evolving by neutral processes). This is particularly powerful if coupled with a direct demonstration of the genetic basis of trait variation, and identification of their effects on organismal performance. While it is desirable to identify the genetic basis for the variation in physiological traits it can be particularly difficult because physiological traits are 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 (hundreds to thousands), where each gene has a small phenotypic impact (14, 78, 144, 182). Identifying the genetic basis for the variation in polygenic traits presents a challenge because they are statistically difficult to identify since they have small effect (8, 14, 79, 182, 196). An additional complexity is that the genes affecting adaptive variation of polygenic traits (polygenic adaptation) need not be the same in all individuals.

The genetic basis for polygenic adaptation may be dissimilar among individuals because there can be many different genetic solutions for the evolution of an adaptive phenotype. For example, metabolic flux through a pathway can be affected by many of the polymorphic enzymes in a pathway; yet adaption only requires change in a few of the many variable enzymes to enhance metabolic flux. The explanation for multiple solutions from polygenic adaptation is that in a population with sufficient standing genetic variation there will be many variable genes that could affect an adaptive phenotype. 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 variation (> 2 alleles). When these variable genes affect the same phenotypic change (i.e., they are redundant), natural selection need only change the frequencies of some of functionally important alleles to achieve an adaptive phenotype. As a result, selection for the adaptive phenotype could be accomplished by altering the allele frequencies at some but not necessarily all functionally important genes because many of the variable genes have similar effects on the adaptive trait. Thus, polygenic adaptation with sufficient standing genetic will result in adaptive evolution of a trait affected by many (hundreds to thousands) of variable genes of small effect where many of the different genes are redundant (effecting the same phenotypic advantage) (14, 78, 144, 177). It is the redundancy that allows individuals to reach the same phenotype, but the population does not become fixed for any one adaptive loci.

The consequences of polygenic adaptation with standing genetic variation are that many members in natural populations will have the same adaptive phenotype; yet, all adapted individuals may not have evolved changes in the same alleles or variable sites. To put it succinctly, the genotypes that affect an adaptive phenotype are not shared among all adapted individuals because there are many genes, pathways or gene complexes that can achieve the same enhanced functions. This concept of polygenic adaptation with sufficient standing genetic variation is important because it changes our expectation of finding a gene, or set of genes responsible for adaptive change in a physiological trait.

Defining the genetic basis for polygenic adaptation is difficult because the analyses requires being able to statistically identify genes with small effect and, among species with large populations (and thus high standing genetic variation), there may be more than one functionally important gene effecting the adaptive phenotype (148). Furthermore, studies on selected families or inbred systems 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 affect an organism's success and defines adaptive physiological function that enhances fitness when organisms are affected 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—population with physiological adaptation to local environment should be enriched for the genetic changes affecting 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 review.

Adaptation in *Fundulus heteroclitus*

The prescription for identifying adaptation is defining populations subjected to evolutionarily important different environments, defining how variation in physiological traits mitigate these environmental differences, and exploring the genetic basis for the variable phenotypes. For *Fundulus heteroclitus*, the habitats with different environments are 1,000s to less than 10 kilometer apart. The physiological traits include metabolism, swimming speed, development, osmotic regulation, resistance to pollution and thermal performance. 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 of the importance of a single gene (LDH-B) along the entire eastern seacoast of North America, and end with a genetic analysis that suggest that evolutionary adaptation is rampant effecting genetic changes among microhabitats within a single population.

Adaptive variation in LDH-B Along the Thermal Cline

The most renowned and well characterized study of adaption 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 *F. heteroclitus* along the steep thermal cline of the North American Atlantic coast have genetic divergences among these allelic enzymes (allozymes, Fig. 2) (27, 128, 134, 139, 143, 149, 179). The pattern of allozyme variation along the Atlantic thermal cline is complex (141, 143, 149). 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, (149) Fig. 1A) while others have smaller allele frequency changes (e.g, IDH-B, Fig. 1A, 6PGDH-A, Ap-A, Est-B, (143, 149). LDH-B has a relatively gradual transition from nearly fixed for the northern allele ($LDH\text{-}B^b$) in Maine to nearly fixed for the southern allele ($LDH\text{-}B^a$) in Georgia (140, 143).

LDH-B is an exceptionally well-studied allozyme in *F. heteroclitus* (139). Mitton and Koehn (115) showed that at locally heated estuaries in New York (near thermal effluent from power plants) the *F. heteroclitus* population had a greater frequency of the LDH-B southern allele type compared to other northern populations not exposed to thermal effluent. This local change in allele frequencies suggested that the LDH-B allozyme variation is affected by natural selection in response to temperature.

Functional studies for both LDH-B alleles initially examined the enzyme kinetics across a wide range of temperatures and pH levels (135, 136)(Fig 2B & C). Enzymes catalyze the conversion of substrates to products and are involved in metabolic functions that affect energy production and usage. The catalytic rate is dependent on kinetic constants (e.g. K_m, and kcat) 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 kcat/K_m is related to the reaction rate, such that equal concentrations of enzyme and substrate result in a higher kcat/K_m and thus a faster reaction. Figure 2B and C highlight that: 1) at lower temperatures the northern LDH-B (b/b) genotype has a

Figure 2. Allozymes in *F. heteroclitus*. **A**. Allelic variation 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 ((139, 149)). 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 (kcat) divided by the K_m) measured at different temperatures at neutral pH ([OH] = [H]) (redrawn based on data in (135)).**C**. Enzyme kinetics (kcat/K_m) for the three LDH-B genotypes plotted against temperature and pH (redrawn based on data in (135)). **D**. Hatching time at 20°C for fish from single populations for the three LDH-B genotypes (redrawn based on data in (48)). Hatching times were defined among twenty randomly crossed pairs, and larvae were genotyped. Data represent larvae genotypes. Similar results where obtained by 4 replicate mass crosses with 40 male and 40 female heterozygotes in each cross (n > 1,000/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 (49)).

Didactic Figure Legend 2: **A:** Along the eastern seacoast of North America, *F. heteroclitus* has a clinal variation in the genetic variants (alleles) for many different enzymes. Three of these patterns are shown here: MDH-A with a steep change of allele frequencies at the Hudson River, IDH-B, with a small change in allele frequencies and LDH-B with a gradual change in allele frequencies. **B &C**: LDH-B enzyme reaction rates (kinetic constants) with different genotypes (homozygotes for the northern allele type, heterozygotes and homozygotes for southern allele type). These different LDH-B genotypes affect both hatching times (**D**) and swimming speeds (**E**). For the enzyme kinetics, hatching and swimming speeds, all fish are from the same population, ensuring a random genetic background, thus differences can be attributed to LDH-B genotypes.

greater reaction rate than the southern genotype (a/a), 2) this difference disappears at warmer temperatures and 3) differential response to temperature among the three LDH-B genotypes is pH dependent (Fig. 2C). At pH 7 and below, reactions for the southern homozygote (a/a) and the heterozygote (a/b) are faster at higher temperatures, but the northern homozygote (b/b) has a higher reaction rate at lower temperatures. These differences in enzyme reactions suggest that the two amino acids that are distinct between the northern (LDH-B^b) and southern (LDH-B^a) allozymes (97, 139) are functionally significant. Yet, even though the difference in amino acids between the two LDH-B alleles effect a change in enzyme function, it does not necessarily mean they are adaptive. That is, the difference in LDH-B amino acids and their effect on enzyme catalysis may not affect performance, longevity or fitness, and thus the north-south divergence could be neutral (139).

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 observed changes in reaction rates may **not** cause a change in metabolism or other important physiological processes. Demonstrating of adaptation requires evidence that the variations in LDH-B catalysis affect organismal functions and that these variations affect fitness. To determine whether variation at the enzyme catalysis impacts higher levels of organismal function, DiMichele and colleagues measured hatching rates (48), swimming speeds (49), developmental metabolic rates (47, 129), and survival (46) as they related to LDH-B genotypes. Hatching time (Fig. 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 among 20 random crosses used to produce the different larval genotypes there should be a random association with the other genes that could affect hatching. Thus, the most parsimonious explanation is that the LDH-B genotype affects hatching rates (Fig. 2D). These data are supported by four replicate mass crosses (40 males and 40 females per cross) between LDH-B heterozygotes. With 1,000s of individuals assayed per cross, hatching rates were dependent on the LDH-B genotype (48).

The variation in hatching rate reflects the difference in developmental metabolic rates (47, 129). 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 crosses were made from individuals from the same population, and thus there should be a random genetic background suggesting that the LDH-B genotype is responsible for the difference in development. To test this idea, purified LDH-B enzyme was injected into eggs with the metabolic rates being 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 (45). This clearly demonstrates that the difference in LDH-B alleles causes significant physiological differences that affect development. Despite this, 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 (Fig. 2E, (49)). Here, data on swimming speeds reflect the difference in LDH-B reaction rates at different temperatures (Fig. 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. Similar to the other physiological studies, these fish were collected from the same population and thus there should be 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 effect on ATP production and hemoglobin oxygen delivery in red blood cells. Individuals with the northern LDH-B genotype have greater ATP production, which affects oxygen binding in fish red blood cells (49). 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, and is correlated with temperature differences along the

North Atlantic coast. There is a higher concentration of LDH-B enzyme in the liver among individuals from northern populations than from southern populations (35, 37, 123, 158), and this variation in enzyme amount is physiologically flexible. Specifically, when individuals are acclimated to different temperatures, LDH-B concentrations change (39, 158). These differences are tissue specific: there is an evolved and physiological acclimation difference in the amount of LDH-B in liver but not cardiac tissue (36, 132, 133, 158). What is apparent from these studies is that there is both an intrinsic (evolved) and a physiologically reversible difference in the amount of LDH-B between northern and southern *F. heteroclitus.* The intrinsic difference occurs in individuals after long-term acclimation and thus is considered to be an evolved, heritable difference between populations. 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 (40). 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 ~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 (40).

These functional physiological data that demonstrate an association between LDH-B genotype and performance strongly suggest that natural selection has effected an adaptive 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 effecting 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 (Fig. 3A). This study (131) 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 closer related taxa will be more similar that genetically more distant taxa. 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 (Fig. 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) (131) (Fig. 3B). Finding this pattern in LDH-B and the other two enzymes is inexplicable unless natural selection favors the increase in these enzymes in lower temperature environments. Thus these data indicate that among many different species the variation in LDH-B enzyme activity is indicative of thermal adaptation (131).

Figure 3. Phylogenetic Analyses of Enzyme Amounts. **A.** *Fundulus* phylogeny (131). 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 correction methods for correcting for species similarities in enzyme amount among fifteen *Fundulus* taxa *versus* naturally occurring mean annual environmental temperatures. Only the three enzymes were significantly related to environmental temperature after correcting for phylogeny: GAPDH- glyceraldehyde-3-phosphate dehydrogenase, PYK pyruvate kinase, and LDH-B. **C.** *F. heteroclitus* glucose dependent metabolism *versus* multiple factor equation using three phylogenetically important enzymes (GAPDH, PYK and LDH-B. $r^2 = 0.866$ ($p < 0.005$) (137).

Didactic Figure Legend 3: **A.** The evolutionary relationships among 15 populations and species of *Fundulus.* Blue boxes have large differences in natural environmental temperature between populations. **B.** Phylogenetically corrected enzyme amounts for the three enzymes that are related to environmental temperatures. Phylogenetic correction is analogous to correcting for body mass and provides measures of enzymes as a phylogenetically independent value for each population or species. **C.** Metabolic rates for northern and southern *F. heteroclitus* populations are a function of the amount of three enzymes GAPDH- glyceraldehyde-3-phosphate dehydrogenase; PYK pyruvate kinase and LDH-B.

The phylogenetic analysis of enzyme amounts predicted that these adapted divergences in LDH-B, GAPDH and pyruvate kinase would have to affect a biologically important trait (Fig. 3B). As a test of this hypothesis, variation in cardiac metabolism was measured within and among northern and southern *F. heteroclitus* populations (137). 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 (131) (LDH-B, GAPDH and pyruvate kinase) also explained the variation in *F. heteroclitus* cardiac metabolism (Fig. 3C) (137). Thus, both a phylogenetic study among species and metabolic physiological study within species indicate that changes in LDH-B enzyme are adaptively important (Fig. 3).

Evolutionary Analyses on LDH-B DNA Sequence Variation

A second data set 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, there should be patterns of sequence variation that support this hypothesis. Sequence comparisons (15, 97) reveal two amino acid substitutions that distinguish the two LDH-B alleles. One of these polymorphisms

TABLE 1. LDH-B MCDONALD-KREITMAN TEST

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

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 (97). 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 (109). This test uses a 2x2 contingency table, comparing the non-synonymous (changes in genetic code that results 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 non-synonymous sites. Typically, this test is used to examine variation within and between species (185), yet the properties of this test hold if one compares two populations in which shared polymorphisms and fixed differences are not confounded (41). Previously published (15) LDH-B cDNA sequences from 11 individuals {from 2 northern populations (Nova Scotia and Maine) and 2 southern populations (Georgia and Florida)}, revealed 3 non-synonymous and 13 synonymous substitutions between northern and southern fish. All of the synonymous but only only 1 of 3 non-synonymous substitutions are shared among northern and southern *F. heteroclitus* (Table 1A). This is improbable (p =0.025; Fisher 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 8 synonymous polymorphic sites are shared among taxa and both non-synonymous changes are fixed differences (Table 1B). This too is improbable (p = 0.022, Fisher 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 EASTERN NORTH 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 affects hatching times, physiology and critical swimming speeds.
- When tested at their native temperatures, the combination of the differences in enzyme amount and enzyme catalysis allow the predominate genotype in the cold northern populations to have equivalent enzyme activities as the genotype in the warm southern populations even though these populations have ~12°C difference in environmental temperatures.

- Across many species there are similar temperature dependent changes in the amount of LDH-B, and the variation in LDH-B is associated with changes in glucose dependent cardiac metabolism.
- Analysis of non-synonymous 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 affected 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

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 (158), **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 (41). 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 non-functional 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 (41, 159).

Didactic Figure Legend 4: A. LDH-B mRNA expression defines the amount of LDH-B enzyme, and the variation in mRNA expression is related to the LDH-B proximal promoter. (**B)** DNA sequence variation where there is surprisingly more variation in functional sites that bind transcription factors than in non-functional sites. **C.** LDH-B proximal promoters from northern and southern individuals demonstrate that the DNA sequence variation between *F. heteroclitus* populations effect a difference in mRNA expression. **D.** Deleting functional DNA but not non-functional sites affects promoter activity and mRNA expression. **E and F.** Evolutionary relationships within and between species demonstrating that non-functional DNA sequence are similar to neutral expectation, but the functional sites demonstrate much greater difference between populations than between species. This pattern is indicative of evolution by natural selection. **G.** Analysis of DNA sequence variation indicating non-neutral patterns for functional DNA sequences.

than differences in amino acids that affect the kinetic constants of an enzyme (e.g. K_m, and kcat) (35, 40, 131, 132, 158). For LDH-B, there is significant correlation between the amount of LDH-B mRNA and its protein (Fig. 4A). Additionally, mRNA concentration is dependent on transcription rate ($R^2 = 0.71$, $p < 0.01$) (38). These data indicate that the concentration of LDH-B enzyme is dependent on LDH-B mRNA transcription rate (38). Thus, because LDH-B expression is regulated by the rate of transcription and there is a significant difference in LDH-B expression between populations, the molecular mechanisms that control transcription should be different. The transcriptional mechanisms that could affect LDH-B transcription include the LDH-B promoter sequence and the protein transcription factors that bind and effect a change in transcription. Much of the difference in LDH-B expression is due to two promoter regions: the proximal promoter and upstream hormonal responsive element (41, 153, 157, 159).

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 non-binding sites (Fig. 4B) (41, 157, 159). 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 variation between populations, but not within a population, effects a significant change in LDH-B expression when assayed in cell culture (Fig. 4C) (41).

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 (Fig. 4C & D), and molecular analysis of SP1 transcription factor (41, 157, 159). An *in vivo* DNA binding assay identified DNA sequences bound to proteins in *F. heteroclitus* nuclei (159). An *in vitro* DNA binding assay confirmed these *in vivo* studies and identified Sp1 as one of the proteins affecting transcription (41, 157). One of the consequences of the LDH-B DNA sequence variation is that it alters the binding affinity of the Sp1 transcription factor (157). Additionally, transfection of the LDH-B promoter or part of the promoter into cell cultures indicates that several transcriptional binding sites affect mRNA expression (Fig. 4D). Two of the promoter sequences (6fp and SP1) contribute to the difference in promoter function when transfected into Rainbow Trout liver cell lines (41, 157). While cell culture results are dependent on the cell type (41), 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 non-functional sites. For example, in contrast to non-functional sites in the LDH-B promoter, the functional transcriptional binding sites are derived (Fig. 4E &F). 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) (41, 157, 159). This shows that there is greater DNA sequence variance for a functional site between populations within a species than between species and is indicative of adaptive evolution (Fig. 4G). This statistically significant difference in sequence variation at functional sites between populations rejects neutral evolution ($p < 0.02$, Fig. 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 and in combination with data above on LDH-B enzyme activity, the changes are adaptively important (41).

In addition to variation at the LDH-B proximal promoter, there is another explanation for the changes in LDH-B transcription rates (142, 153, 154). 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 affects mRNA expression in *F. heteroclitus*, indicating that this DNA sequence is important for the adaptive divergence in LDH-B expression (153). 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 processes: 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 (25, 56). Microarrays are simply 100-200µm spots of DNA distributed across 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 (versus contrast to a single control) to examine the variation within and among populations (125, 126) (Fig. 5). This study and others (32, 88, 189) 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: 1) that ~20% of mRNAs were significantly different among individuals within a population and 2) ~3% of

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 circle) and *F. grandis* (green circle). Redrawn from (125).

Didactic Figure Legend 5: Genome wide patterns of mRNA expression measured with microarrays. A. Relative expression of adaptively important genes showing that northern individuals have different expression than both southern and *F. grandis* individuals. **B.** Plot of relative expression level (as $\log_2$ values, thus 1.0 = 2 fold difference) and probability of being significantly different (as negative $\log_{10}$ values, thus 2 = 0.01 for northern *versus* both southern and *F. grandis.* Gray box highlights p-value ~0.0001, with less than 1.5 fold difference.

mRNAs had a derived non-neutral pattern of expression indicative of adaptive variation (Fig. 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* (Fig. 5A) (125). 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 indicative of significant change. For example, while there could be a 4-fold difference in the mean between populations, the variation within the population could be equally large and thus, the four-fold difference would be insignificant (Fig. 5B). Similarly, several genes had adaptive patterns of mRNA expression yet had less than 1.5-fold differences between populations (Fig. 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 (125). What is clear from these studies is that there is significant inter-individual variation that evolution can act on to effect changes in mRNA expression. These patterns of mRNA expression are indicative of adaptive evolution.

In a separate study, adaptive differences as well as reversible physiological acclimation effects were studied to determine the number of mRNA transcripts affected (43). 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 ~ 7,000 unique mRNA transcripts were quantified using microarrays. Adaptive changes in mRNA expression were defined as those transcripts with significant derived expression (northern *F. heteroclitus versus* both southern taxa -- *F. heteroclitus* and the sister species *F. grandis*). These data revealed that there are more mRNA transcripts with a significant adaptive difference than there are mRNA with significant acclimation effect. Furthermore, the mean adaptive difference was larger than acclimation effects. Interestingly, few mRNAs were affected by both adaptive and acclimation effects, and yet adaptive effects were more frequent at 12°C or 28°C than at 20°C and very few of these mRNAs had adaptive differences at two or more acclimation temperatures (43). When there was a significant interaction between adaptive and acclimation effects, 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 (Fig 6A) (188). 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 difference in the amount of mRNA that is significantly associated with native temperature but independent of genetic similarity among populations (Fig. 6B) (188). Thus, adaptive mRNA transcripts are those that regress significantly with habitat temperature following correction of the expected covariance due to genetic relatedness (PGLS analysis, (188)). 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 (Fig. 6B).

Figure 6. Clinal Adaptive Variation in mRNA expression, A. Five sample sites and the phylogenetic relationship among populations based on a microsatellite-derived neighbor-joining tree with median annual temperatures (ºC) averaged over 30 years. Branching pattern is a neighbor-joining tree constructed from pairwise Cavalli-Sforza and Edwards' chord distances (28) 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 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 logs of variation. Ratios of variation are indicative of evolutionary processes (directional, stabilizing, balancing or neutral). Redrawn from (188).

Didactic Figure Legend 6. mRNA expression was measured among five *F. heteroclitus* population distributed along the eastern seacoast of North America (A). Because demography and thus neutral evolutionary processes can create difference among these five populations, the effect of demography was compared (Y-axis) to the effect of habitat temperature (B). The enlarged spots are adaptively important because there is a significant relationship between mRNA expression and habitat after removing demographic effects. (C) Patterns of variation within and between populations are indicative of different evolutionary effects.

In addition to PGLS analysis, the ratios of variation (ANOVA) are 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 (Fig. 6C, blue circles), indicative of balancing selection (188). Twenty-four mRNAs had expression patterns affected by stabilizing selection (Fig. 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, are significantly biased for genes involved in oxidative phosphorylation (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 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 (127) (Fig. 7).

Figure 7. mRNA Expression and Cardiac Metabolism. **A.** Relative levels of cardiac metabolism for sixteen individuals (8 per Maine and Georgia population). Cardiac metabolism was measured using glucose, fatty acid and LKA (lactate, ketones and ethanol) as substrates (127). 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 = 2-fold 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 relative to the mRNA expression. mRNA expression summarized as of 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 (127).

Didactic Figure Legend 7. A. mRNA expression predicts heart metabolic rates when supported by glucose, fatty-acid or mixture of secondary metabolites (lactate, ketones and ethyl alcohol). Individuals differ in their overall metabolism and which substrate supports the highest metabolic rates. **B.** Significant difference in mRNA expression among individuals within each population (Maine and Georgia). Most mRNAs have a significant difference between individuals ($p < 0.01$ to 10^{-17}) even though the magnitude of the difference (fold-change) is relatively small (few genes have 2-fold or more differences among individuals). **C.** There are three groups of individuals that share similar patterns of expression such that individuals in one group are up (red in group 3) and individuals in another group are down (green group 2). **D.** mRNA expression was summarized as a linear combination of genes in three primary biochemical pathways: glycolysis, TCA-cycle and oxidative phosphorylation. The mRNA from different biochemical pathway predict fatty-acid metabolic rates for different groups of individuals. For example, TCA mRNAs explain 64% of the variation in fatty-acid dependent cardiac metabolism for group 1, while oxidative phosphorylation mRNAs or glycolytic mRNAs explain group 2 and 3 cardiac metabolism (respectively).

Previously in this manuscript (Fig. 3), cardiac metabolism was considered adaptive because it was higher in northern *F. heteroclitus* due to adaptive increases in glycolytic enzymes (131, 133, 137). 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, Fig. 7A). Cardiac metabolism is a function of body-mass, and the effect of body mass 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 (Fig. 7A): as much an 11-fold difference among individuals for fatty-acid dependent cardiac metabolism and a 2-fold difference for glucose dependent cardiac metabolism. Additionally, within an individual there is variation in which substrate supports the greatest metabolic rate. For example, fatty acids supported one of the highest metabolic rates in G3 (one individual), but glucose supported one of the lowest metabolic rates in this individual (Fig 7A). Thus, there was variation among individuals in metabolism for each substrate and consequently variation in substrate-supported metabolism. Even with large variation in metabolism within populations, there were significant differences between 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 ($p < 0.01$; Fig. 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 (Fig. 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 mRNA in figure 7C, if group 2 individuals had very high relative levels of expression, group 3 individuals would have very low levels of expression (Fig. 7C). No other random sets of individuals shared these significant patterns. That is, the three groups of individuals are statistically unique (127).

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 (Fig. 7D). That is, within each of the three groups of individuals (Fig. 7C), the variation in mRNA expression predicts the substrate specific metabolic rates (Fig. 7D). However the mRNAs that statistically explain the variation in cardiac metabolism were different among the three groups (127). 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 oxidative phosphorylation (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 the glucose and LKA dependent metabolism (127). 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% to 70%). Overall, mRNAs from different metabolic pathways are related to different substrate specific metabolisms within a group.

These data demonstrate that the variation in mRNA expression can predict the level of cardiac metabolism, suggesting that differences in mRNA expression lead to differences in metabolism. Equally important is the observation that the large variation in the amount of mRNA expression (where 94% of mRNA are significantly different among individuals, Fig. 7B) contributes to the large variation in cardiac metabolism (2-11 fold difference among individuals, Fig. 7A). The observation that among different groups of individuals, cardiac metabolism is explained by different sets of mRNA, which are part of different metabolic pathways (Fig. 7C & D), suggests that there are multiple genetic pathways to achieve similar physiological outcomes. That is, if the heritable variation in mRNA expression are independent (i.e., not due to change in a single transcription factor) than these variation in mRNA represent different genetic basis for the adaptive variation in cardiac metabolism. Thus, for the adaptive variation in cardiac metabolism, is polygenic in that this 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 variations need to be harnessed to effect a physiological adaptation. Specifically, within a population there is large variation in cardiac metabolic rates fueled by fatty acids, yet the mRNAs associated with this variation differ among individuals—the variation among the 3 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 occurs. This concept has important consequences. For example, the fact that many unique polymorphic sets of genes can effect an adaptive metabolic change means that selection is more robust. It is more robust because there are many polymorphisms that can effect an adaptive change, the genes that affect an adaptive change already exist at reasonably frequency and these genes 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 there are few rare genetic variations. Additionally, adaptive evolution can be achieve without too much genetic load for the same reason—there need not be large change in allele frequencies. The importance of polygenic selection is becoming recognized in many organisms including humans and is significantly altering our understanding of evolution (54, 81, 84, 94, 96, 144, 171, 180).

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 there is considerable amount of variation in mRNA expression: ~ 20% of hundreds of genes have a statistically significant difference in mRNA expression among individual within a population.
- Evolutionary analyses suggest 3-4% of genes exhibit adaptive divergence in mRNA expression.
- Physiological acclimation has less of an effect on mRNA expression than adaptive divergence, and, for most genes, there is little interaction 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—there are many significant mRNA expression changes that can effect an adaptive change with different sets of mRNAs affecting 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, 116), physiological traits in embryos such as developmental rate (50), physiological traits in adults, such as metabolic rate (10, 11, 60, 76, 127), thermal tolerance (59), hypoxia tolerance (107, 108), and salinity tolerance (156). 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 (60, 108), 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 (73). 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 **S**ingle **N**ucleotide **P**olymorphisms (SNPs) associated with variation in thermal tolerance within a central New Jersey population, and detected only four SNPs associated with variation in hypoxia tolerance (74). 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 - the result of SNPs in many different genes, each having a small effect on phenotype. Alternative methods, such as random forest analysis, 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 (24). 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 tolerance 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 (73) (Healy et al. 2018). 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: 1) 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 2) 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: from within Chesapeake Bay to within single saltmarsh estuaries. Rapid local adaptation seems improbable when the cost of selection is considered (42, 91, 184). 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 on many genes can reduce the numbers of individuals to such an extent that the species becomes unviable (69). The greatest cost of selection occurs when an adaptive allele arises from a new mutation and goes to fixation. 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) (69). Through this type of selective sweep, allelic variation is lost from the population. Because of this cost of selection, or genetic load, and the identification of unexpectedly large amounts of allelic variation (101), much of the natural standing genetic variation is attributed to neutral processes (89). 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 would not expect to find local adaptation among well connected *F. heteroclitus* populations where alleles from adjacent, non-local populations should swamp the effect of selection. Yet, recent data using the power of high-throughput sequencing indicates an abundance of genetic variation. This high standing genetic variation may be enough to affect 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 kilometers to less than 100 meters.

Local Adaptation to the Osmotic Environment

Intracellular solute concentrations that affect the water balance or osmolarity of cells is carefully controlled in vertebrate animals (170). 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 defines the fundamental niche of teleost fish. Most fish species are either freshwater specialists or marine specialists (Fig. 8A), 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 (16, 181). In contrast, relatively few fish species live in estuarine environments that have large environmental salinity fluctuations (155). 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 (187). *F. heteroclitus* exhibits one of the broadest ranges of osmotic acclimation abilities known within the fishes (155). 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 (92, 178). In contrast, closely related *F. majalis* (Fig. 3) is coastal, but occupies the marine end of the salinity continuum within estuaries, exhibiting limited abilities to acclimate to dilute water (67). Approximately half of *Fundulus* species occupy a coastal niche and can tolerate a wide range of salinities (187). 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 (Fig 8A). As one moves upstream along the salinity gradient in the Potomac River, neighboring groups of fish are genetically very similar. However, once the water becomes dilute freshwater in the extreme upper estuary, the fish populations found there (FW-native) are genetically distinct from those just a few

Figure 8 Local Osmotic Adaptation. *F. heteroclitus* population variation along a salinity gradient in Chesapeake Bay. **A**) Map of salinity gradient in 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 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 re-drawn from (190). Salinity gradient heatmap of Chesapeake Bay was generated from the NOAA Chesapeake Bay Operational Forecast System (https://tidesandcurrents.noaa.gov/).

Didactic Figure Legend 8. *F. heteroclitus* populations live in marine environments with different salinities that are a few hundred kilometers apart (A). These populations differ genetically due to both neutral and potentially adaptive genetic variation (B). The change in mRNA expression in response to different salinities depends on whether the genes evolve by neutral or adaptive processes.

miles downstream in brackish water (BW-native) (Fig. 8B) (190). 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 (23). Marine-native and BW-native populations are as genetically distinct as BW-native and FW-native populations (Fig 8B) (190). 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 fresh water (190). 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 micro-evolutionary 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 (57) 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 (190). These mRNA expression differences are therefore putatively adaptive, mirroring patterns of physiological differences between populations. Furthermore, genes who mRNA expression enable physiological flexibility show less variation in their expression among individuals than other expressed genes, and this expression variation is lowest in FW-native fish, which is further evidence that they are adaptive (160). Importantly, mRNAs that were physiologically responsive (up- or down-regulated) to experimental salinity challenges (Fig 8D) were more likely to show patterns of adaptive divergence between populations than mRNA for genes that were not physiologically responsive (Fig 8C) (190), 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 (190). The evolutionary importance of these *F. heteroclitus* responses were 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. Observations that modification of the same molecular pathways and physiological processes occurring on macro (among species) and micro (within species) evolutionary timescales is indicative that a complex suite of physiological systems is repeatedly modified by natural selection thereby supporting adaptation to different osmotic environments.

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 physiological variation. In *F. heteroclitus* GWAS revealed that variation at 26 genes accounted for 56% of the variation in osmoregulatory abilities among individuals (22). 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 (22). 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 (43). 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 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 evolutionary time scale. Perhaps the difference between salinity and temperature adaptation is because osmotic adaptation and regulation is 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 were compared between populations six 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 thermal effluent (TE) discharged from power plants (Fig. 9). The increase in temperature by 4-12°C at these TE sites is less than 50 years old. Even though TE sites are young, *Fundulus* occupying these TE sites have greater critical thermal maxima (CTmax: maximum temperature at which an individual can no longer actively escape unfavorable conditions)(44). Genotyping by sequencing (GBS) was used to identify 5,449 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 from the reference. 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 the Oyster Creek NJ, there is subtle demographic structure that can be identified by a discriminant analysis of principal components (DAPC) that maximizes the differences between populations while simultaneously minimizing the differences within populations (85). 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, Fig. 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 TE and reference sites statistical outliers, 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 (Fig. 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: 1) outlier SNPs have F_{ST} values that exceed the neutral expectation, 2) they are significant for the TE sites relative to both cooler reference sites and 3) they do not have significant demographic effects because they are not different between the two cooler reference sites (44).

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 (145), which groups individuals based on shared and similar allele frequencies. The Structure analyses clearly distinguished the TE site from both reference sites (Fig. 9C).

Figure 9. Local adaptation to warmer temperatures created by power plant thermal effluence. 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 ~5,400 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.

Didactic Figure Legend 9. With a triads (3 populations: one hot TE and two cool reference populations north and south of the TE), (**A)** we can distinguish between populations, and more importantly, TE is different from both references. We can identify the SNPs that have the greatest effect on genetic distance (F_{ST}) as “outlier SNPs”: SNPs with statistically unexpectedly large F_{ST} values. For this study (**B)**, the 94 “adaptive SNPs” or SNPs evolving by natural selection have statistically large F_{ST} **values** between the TE and both references but are not different between references. These 94 SNPs reveal population structure where the TE population is unique and different from both references (**C**). Surprisingly, there is little linkage among SNPs (**D**). That is, the F_{ST} values for outlier SNPs declines to random expectation after 25 bp of DNA.

Regardless of clustering into two, three or four groups (k = 2, 3, or 4, Fig. 9C), the Oyster Creek TE site segregates from both reference sites, which share similar allele frequencies.

These analyses of ~5,400 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 (Fig. 9D). That is, the 94 outlier SNPs in Oyster Creek are genetically unlinked: the F_{ST} value drops to insignificant values only 25bp from the outlier SNP (Fig. 9D). The rapid decrease in F_{ST} values is indicative of very small linkage groups, and thus each SNP is most likely evolving independently. In conclusion, rapid adaptation to recent thermal warming at a TE site is polygenic and due to allele frequency changes from standing genetic variation involves many independent genes that do not require allelic variation to go to fixation (44).

Local, Rapid Adaptation to Pollution

The production of persistent organic pollutants (POPs) including polycyclic aromatic hydrocarbons (PAHs) and polychlorinated biphenyl (PCB) contribute to one of the more serious forms of pollution: exposure is associated with carcinogenicity and mutagenicity (9, 61, 100, 102, 150, 197), and has been associated with metabolic diseases, including type 2 diabetes, obesity, and energy metabolism (7, 87, 98, 99, 103, 150). *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 (26, 33, 112, 117-120). The toxic effects of these POPs are mediated largely through the aryl hydrocarbon receptor (AHR) pathway (62, 113, 138). Importantly, *F. heteroclitus* embryos from the polluted populations are resistant to these POPs in first and second generations after rearing in common clean environment (118-120). Thus, the differential survival of fish from polluted habitats is due to genetic adaptation rather than physiological reversible 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 non-polluted sites) when exposed to POPs during development (20, 21). The abnormal development is associated with changes in mRNA expression not seen in fish from polluted populations (19, 21). POPs also affect metabolic pathways, specifically the oxidative phosphorylation pathway that is responsible for the vast majority of ATP production. Similar to development and gene expression patterns, resistant populations show little change in oxidative phosphorylation metabolism when exposed to POPs, whereas nearby clean reference populations have significant changes, which are heritable (52, 53). These phenotypic changes that reduce the effect of POPs on development and metabolism and enhance survival are recent adaptions. One of the questions is whether we can identify the molecular and genetic basis for this adaption.

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 (Fig. 10A (191, 192)). 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 (Fig. 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) (192). The union of three separate adaptive tests (i.e., all three tests were significant) identified ~2% to 4% adaptive SNPs among the three triads (Fig. 10A). All three polluted populations are resistant to pollution with each having 6-15 SNPs significant for all three statistical tests (Fig. 10A). Yet, only one SNP was shared among the three triads: CYP1A, the enzyme cytochrome P4501A. CYP1A is a phase I xenobiotic metabolizing enzyme integral to the detoxification pathway (192). The observation that only one SNP is shared among all three polluted populations could reflect polygenic selection with redundant genetic variation that results in a similar phenotype (resistance to pollution) with different

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 triad (C-P-C) were 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.** Induction of gene expression with exposure to persistent organic pollutants (POP) in cells in culture for CYP1A promoter from polluted and clean populations. **C.** 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.

Didactic Figure Legend 10. A. illustrates the three locations where each location has a polluted and two clean reference populations (triad). Each of these locations was used to define changes in DNA sequences for 354 separate SNPs. The Venn diagram illustrated the number of SNPs with DNA changes shared for three different statistical tests. The red number is the number of SNPs that are statistically significant for all three tests. **B** and **C**, show how DNA promoter sequences effect a change in gene expression when exposed to pollution. Each promoter was linked to a reporter gene and cultured cells were transformed with these DNA sequences. Thus, in the same genetic cell line the promoter sequences from polluted populations induce higher gene expression than promoter sequences from clean reference populations. In **C**, populations with different letters are statistically different.

adaptive genes. Alternatively, the different genetic solutions could be that the different selectively important SNPs represent unique solutions to the triad specific POPs and other pollutants. Whether the different selectively important SNPs are polygenic is unresolved, yet functional studies on CYP1A are indicative of polygenic adaptive evolution that results in multiple different genetic solutions.
The only selectively important SNP shared in all three polluted populations was the CYP1A SNP in the first intron. 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, Fig. 10A). To test how these promoter-intron affected mRNA transcription,these promoters were linked to a reporter gene and the expression of the reporter gene was quantified in cell culture (Fig. 10B & C). These hybrid gene complexes (CYP1A promoter-intron and reporter gene) were sensitive to the amount of added POP (Fig. 10B) indicating that the promoter sequence could affect 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 (Fig. 10C). Thus, there is a functional difference due to the DNA sequence variation between the polluted and both clean reference populations. Interestingly, the promoter-intron sequences exhibited extremely large sequence variation: ~9% of the sites were polymorphic (193). 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 population. Thus, even though the promoter from the polluted populations effected a change in gene expression, this was accomplished via distinct DNA sequence variations. This pattern of multiple different genetic divergences effecting an adaptive change is indicative of polygenic adaptation from sufficient standing genetic variation.

A limitation of the Williams and Oleksiak (193) 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, (Fig. 11A (147)). 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 sensitive to POP exposure (Fig. 11B). That is, there is a shared common phenotype (mRNA expression) in all polluted populations: the absence of mRNA response to POP exposure, especially for genes involved in the AHR pathway (Fig. 11C). These mRNA expression patterns were related to changes in the *F. heteroclitus* genome (147) (Fig. 11C & D).

The AHR protein is a transcription factor that is activated by binding PAH, PCB and other xenobiotics, which then affects mRNA expression and POP toxicity (Fig. 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 bind AHR (AHR response elements or AHREs) (58, 172). One of the more common mRNAs regulated by AHR is CYP1A. In the four polluted *F. heteroclitus* populations*,* there is a similar phenotypic response—lack of enhanced transcription to genes regulated by AHR (Fig. 11B). While, there are thousands of DNA variants with signature of natural selection 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 (Fig. 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 (Fig. 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.

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, non-polluted 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 a 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   pi, nucleotide diversity) between tolerant (T) and sensitive (S) populations. Gray shading highlights DNA sequences with unusually large significant F_{ST} values, extreme   i   values, or both. These regions of the genome also have significant Tajima's D (not shown).

Didactic Figure Legend 11. A. Shows the pairs of populations used in this study where one population inhabiting polluted water was compared with a population inhabiting clean water. One of the comparisons (**B**) was mRNA expression patterns revealing that sensitive populations, but not tolerant populations, responded to pollution exposure, and this was particularly relevant to AHR regulated genes (lower panel in **B**). The diagram (**C**) illustrates AHR regulatory pathway with filled boxes indicating genes evolving by natural selection. These change in mRNA expression, especially associated with the AHR pathway, have patterns of DNA sequence variation indicative of evolution by natural selection (**D**).

These data are similar to the original DNA sequence analyses using 354 SNPs. While there are consistent adaptive genomic changes among the four polluted populations for several genes (147), few of these variants are shared (Fig. 11D). Instead each population has a unique set of genomic variants. That is, similar to CYP1A where a polluted promoter effected a change in mRNA expression without sharing the same DNA sequence; the genomic variation is observed at the same genes, but distinct DNA variants are responsible in each of the tolerant populations. This is an example of polygenic adaptation, where many different genetic solutions to an external challenge lead to selection for the same adaptive phenotype. It is important that these genetic variants existed in the ancestral population at reasonable frequencies in order for polygenic adaptation to be possible.

TO SUMMARIZE LOCAL, RAPID ADAPTATION IN F. HETEROCLITUS:

- Rapid (<100s of generations) adaptive evolution involving many genes.
- *F. heteroclitus* populations subjected to recent, anthropogenic changes in the thermal environment demonstrate enhanced survival to high temperature 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 non-polluted 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.

Epistatic Evolution between Nuclear and Mitochondrial Genomes

Epistasis complicates adaptive evolution by requiring allelic variants in separate genes to co-occur (i.e., to avoid recombination). Epistasis describes the functional variation of one gene as a consequence of the genetic variation at another gene. Thus, epistatic evolution is the process by which the selection for an allele is affected by the genetic variation at another gene. The study of epistatic evolution is problematic because of the large number of potential interactions and the difficulty of inferring phenotypes from the underlying additive effects of genetic variation (105, 106). Yet, there are biological interactions among the genes that form the oxidative phosphorylation (OxPhos) pathway. The OxPhos pathway involves both the nuclear and mitochondrial genomes. As in nearly all animals, the mitochondrial genome (haplotype) 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 (Fig. 12A). The interactions among these 91 proteins in the OxPhos pathway are sensitive to acute and chronic temperature exposures (151, 152). It is the DNA sequence variation between the two genomes (nuclear and mitochondrial) that ultimately are 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 (186). Yet, between these two mitochondrial haplotypes (mitotypes) there is little evidence that DNA sequence changes between the northern and southern mitochondrial genome are adaptive (186). For quantitative physiological traits, *F.*

heteroclitus individuals from northern and southern populations reveal subtle OxPhos metabolic differences that are most obvious at low temperatures (30, 31, 60, 75, 154). 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 (11) (Fig. 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, these studies between individuals with northern and southern mitotypes do not address if there are interactions between nuclear and mitochondrial genomes.

To examine 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, there are local populations

Figure 12. Epistatic Adaptive Evolution. **A.** Oxidative Phosphorylation pathway and the number of protein subunits encoded by mitochondrial and nuclear genomes (86). **B.** Mitochondrial OxPhos dependent respiration (State 3) measured in *Fundulus heteroclitus* from a single New Jersey population. Individuals were acclimated to either 12°C 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 <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 > 75 % northern nuclear alleles and the northern mitochondria are blue, individuals with < 75% northern nuclear alleles and the northern mitochondria are green. Individuals with < 75% southern nuclear alleles and with southern mitochondria are orange. Individuals with >75% southern alleles and the southern mitochondria are red.

Didactic Figure Legend 12. A. Mitochondrial metabolism produces most of the cells ATP *via* the Oxidative Phosphorylation pathway. This pathway has five enzyme complexes with a total 91 proteins – 13 from the mitochondrial genome and 78 from the nuclear genome. **B**. In a single population the mitochondrial respiration is affected by acclimation, assay temperature (acute effects), and an individual's mitochondria. **C**. Within this population are 349 nuclear DNA sequence variants (SNPs) that have large and statistically unlikely differences in allele frequencies between mitochondrial haplotypes. This difference in nuclear allele frequencies is denoted as wF_{ST} value—F_{ST} values within a population between mitochondrial types. **D**. The difference in nuclear SNPs affect mitochondrial respiration: individuals with more "southern" nuclear alleles, regardless of which mitochondria have higher metabolic rates.

in northern New Jersey, just south of the Hudson river, with nearly equal frequencies of both mitotypes (71). In these populations there is a bimodal distribution of nuclear allelic variation: there is a significantly lower number of heterozygotes, indicative of natural selection favoring a specific combination of nuclear and mitochondrial genes (111). 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 (11, 12).

Within and between the two mitotypes found in a single panmictic population the variation in nuclear genotypes affects OxPhos (12). This result relied on evolutionary analyses that identify unexpectedly large, significant differences in nuclear allele frequencies between the two mitotypes (Fig. 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} (12). Examining ~11,000 different sites with DNA polymorphism (SNPs), revealed 349 outlier SNPs: SNPs with significantly large and statistically unlikely wF_{ST} values (Fig. 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 (Fig. 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 (Fig. 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 (Fig. 12D). These data (10-12, 75, 111) indicate that the interaction between the mitochondrial haplotype and nuclear genotype affects 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, there is no nuclear DNA sequence variation that is fixed (at 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 are necessary for epistatic adaptation. In conclusion, these data suggest high standing genetic variation that affects mitochondrial function with only some of this variation having changed due to natural selection to effect the adaptive change in OxPhos metabolism.

An alternative explanation for the significant divergence between 349 outlier SNPs is individuals mating assortatively rather than randomly. For example, if the 349 SNPs affected a 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 there are nearly equal frequencies of "northern" and "southern" type mitochondrial haplotypes with 5 amino acid differences among the 13 mitochondrial-encoded proteins.
- In this population, there is a bias in nuclear DNA sequences such that individuals with the northern type mitochondria have different nuclear allele frequencies from individuals with the southern type mitochondria.
- The large significant differences in nuclear allele frequencies are indicative of natural selection affecting these genes due to epistatic evolution.
- Nuclear genes with large significant allele frequency differences affect 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 meters apart and exhibit significant environmental differences (Fig. 13A). In a panmictic *F. heteroclitus* population that lays and fertilizes eggs in the high upper-tidal zone, it seems unlikely that there would or could be genetic variation associated with these physically close microhabitats in saltwater marsh estuaries. 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 (183). Specifically, among ~4,700 SNPs a vast majority have similar allele frequencies. Yet, surprisingly, in all three separate estuaries individuals are significantly different for 1-2% of SNPs among microhabitats (Fig. 13B). For each separate estuary, SNPs with significant and large differences in allele

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 200m and usually less than 50m. **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 3 analyses with joint FDR < 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), 4,352 non-outlier 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).

Didactic Figure Legend 13. Three saltmarsh estuaries were examined along the New Jersey shore. In each saltmarsh there are permanent ponds and intertidal creeks that drain into basins (**A**). To determine if there was significant, potentially adaptive divergence, three statistical tests were applied to SNP allele frequencies. SNPs that were significant for all three tests were considered outlier SNPs that are most likely evolving by natural selection (**B**). The F_{ST} values for outlier, non-outlier and randomized data sets is shown for each population (**C**) and compared between Rutgers NJ and either Stone Harbor, NJ or Mantoloking, NJ.

frequencies between microhabitats were determined using three statistical tests (Fig. 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 FST values exceed the lower 95% CI of the outlier SNPs Fig. 13C) (183). Importantly, among outlier SNPs, allele frequency differences between microhabitats (< 100m apart) exceed the differences among estuaries (100,000m apart, Fig. 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 (Fig. 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 a) in the same gene at a different position, b) in a duplicate gene or paralog, or c) among genes with similar annotations. In fact, 11 outlier SNPS are in the same or paralogous gene. One example is outlier SNPs 33bp apart in the same intron of vav guanine nucleotide exchange factor 2 (VAV2, signal transduction gene). Another example is outlier SNPs that affect all three replicate populations for glutamate receptors: (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 (183). 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 we find that there are 8,180 (12% of all possible) with significant LD and 93% of these are < 100bp apart. For the 261 outlier SNPs there are 1,192 significant LDs with 15,259 other SNPs, and all but 67 or 94% of these are <100bp apart. In the classic model of evolution, where new mutations quickly go to fixation when affected 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 affecting 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, Fig. 13C). Yet, it is possible that the outlier SNPs affect such homing behavior. This would suggest natural selection favors alleles that affect a behavioral choice, *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 affected by selection.

The observation of significant and frequent allele frequency differences among microhabitats less than 100 meters 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, 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, 176) (Teo and Able, 2003b; Able et al., 2006, 2012). In contrast, tidal basins have lower diurnal fluxes in water temperature, salinity and dissolved oxygen (70, 82, 175, 176). Large differences in SNP allele frequencies among microhabitats exceed those among geographically distant populations (Fig. 13D) and are indicative of fine-scale genetic structure. The evolutionary importance of this observation depends on two points: (1) whether outlier FST values are statistical errors (type I errors) or (2) whether neutral evolutionary processes are likely to explain the large FST 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 (183). 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 there are significant differences in SNP allele frequencies 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.
- There is little long distance linkage (nearly all SNPs exhibit LD of 100bp or less).
- These patterns of divergence suggest polygenic selection from standing genetic variation on a fine ecological scale.

Discussion

Fundulus heteroclitus have adapted to the thermal cline along the eastern seacoast of North America, along salinity gradients in bays and esturaries 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 affect these traits occur at different spatial and temporal scales: along the eastern seacoast of North America or among bays evolving for 1,000s 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 affect reaction rates, glycolysis, swimming, and developmental rates (Fig. 2 & 3). The different biochemical properties between the two LDH-B alleles are associate with non-synonymous substitutions and the difference in DNA sequences between Maine and Georgia populations reject the null hypothesis of neutral divergence (Table 1). Across this continental coast are also differences in LDH-B expression due to DNA sequence changes that alter promoter function and are evolving by natural selection (Fig 4). Thus, the adaptation associated with LDH-B are related to changes in promoter effecting mRNA and two amino acid substitutions that are show no linkage disequilibrium. LDH-B evolving by natural selection and affecting adaptive physiological traits is most similar to the classic theory of adaptive divergence where there are large, nearly fixed differences in allele frequencies with a single locus of large effect on phenotypic traits.

Beyond the single LDH-B locus, there are genome wide patterns of mRNA expression that show a surprisingly high amount of variation among individuals within populations (~20% of mRNAs are significantly different among individuals). Two separate analyses indicate that along the eastern seacoast there are adaptive differences in mRNA expression among populations (Figs. 5 & 6). Importantly, the variation in mRNA expression affects cardiac metabolism (Fig. 7) but in a complex way: the specific genes of importance that affect 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 can affect 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 can effect similar adaptive phenotypic change).

In addition to adaptive patterns of mRNA expression (Fig. 4, 5, 6 & 7), there are significant SNP allele frequency differences among local populations. 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 meters 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., minor allele frequencies > 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 (Fig. 10 & 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 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 affect physiological traits. These observations of 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 (Fig. 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.

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 (101) provided data on 18 proteins and concluded that excessive genetic variation exists within a species, but they could not resolve the evolutionary forces affecting this variation. Sixty years later, it is not clear that we have resolved this dilemma (29, 77, 95, 122). 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, on smaller geographic and smaller time scales there appear to be many solutions for physiological adaptation: there are many polymorphic genes with only a subset required to effect an adaptive physiological change. To identify these polygenic solutions and thus address Lewontin's and Hubby's (101) 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 *Fundulus heteroclitus* an excellent model for these studies and for understanding adaptive physiology.

Conclusion

There is adaptive divergence in transcriptional, biochemical, metabolic, osmotic and whole animal physiologies within and among *F. heteroclitus* populations. The strength of these studies 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 populations 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 provide an experimental approach to identify genetic changes associated with environmental habitats 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.

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 involve many independent evolve changes: non-synonymous substitutions that affect enzyme kinetics, changes in DNA that alter transcription binding sites altering the amount LDH-B expression and the evolution of an 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 from standing genetic variation – involving many genes of small effect where there are many redundant adaptive solutions. For example, the adaptive variation in cardiac metabolism is associated with changes in the utilization of different substrate (glucose, fatty acid or secondary metabolites) and the 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, evolutionary change that enhances fitness in local environments, which are exposed to different temperatures, concentration of pollutant, osmotic conditions or ecological habitats; are polygenic. Importantly, in populations that have similar adaptive phenotypes associated similar environments do not shared the same evolutionary important genetic polymorphisms. The observation that in different populations exposed to the same environmental there are variable evolutionary important genetic changes (i.e., they occur in different variable sites within the same gene or in different genes) is best explained by polygenic adaptation in populations with sufficient standing genetic variation such that there are multiple polymorphism which affect the same adaptive phenotype .

The most important evolutionary insights from studies of adaptive evolution in *F. heteroclitus* may be the hypothesis that polygenic adaptation in highly polymorphic populations affecting the adaptive evolutionary changes in physiological traits. We define polygenic adaptive with sufficient standing genetic variation as the evolution by natural selection of a trait affected by many (hundreds to thousands) of variable genes of small effect where many of the different genes are redundant (effecting the same phenotypic advantage). If this hypothesis is correct (if polygenic adaptation is achieved by changes in diverse and redundant polymorphisms) it depends on the large and highly polymorphic populations characteristic of species like *Fundulus heteroclitus* (146). Yet, highly polymorphic populations (populations with high standing genetic variation) and polygenic adaptation are not independent. Yes, polygenic adaptation requires many genes that have similar effects on the phenotype (redundant variation) but also polygenic adaptation provides a mechanism to maintain these polymorphisms. That is, because there are many solutions for an adaptive phenotype, few of the genetic polymorphism will go to fixation—most genetic polymorphism will remain polymorphic even with a selective advantage. This is an intriguing adaptive explanation of the large amount of genetic polymorphisms observed in *F. heteroclitus*: the large amount of standing genetic variation is not due to balancing selection favoring heterozygotes, but instead due to natural selection maintaining polymorphisms because there are multiple genetic solutions underlying adaptively significant trait variation.

While polygenic adaptation is an attractive explanation for the patterns we observe in *F. heteroclitus* and explains most of the data, it is not the only possible explanation. The variation in genetic response associate with adaptation within a population could arise because of pleiotropy, epistasis or unaccounted gene by environmental 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), change in ecologically communities (e.g., fewer predators), or gene by environmental interactions. Thus, while we tentatively suggest that polygenic adaptation is important, the consequences of polygenic adaptation are too significant to jump to this conclusion. Instead, further research should focus on testing the predictions of polygenic adaptation and the theoretical consequences on genetic polymorphisms should be examined. *Fundulus* genomic research examining equivalent adaptive physiological changes among populations should contribute to these objectives and *Fundulus* should continue to serve as model system to study evolutionary adaptation.

References

1. **Able KW, Balletto JH, Hagan SM, Jivoff PR, and Strait K**. Linkages between salt marshes and other nekton habitats in Delaware Bay, USA. *Rev Fish Sci* 15: 1-61, 2007.
2. **Able KW, and Castagna M**. Aspects of an Undescribed Reproductive Behavior in Fundulus heteroclitus (Pisces: Cyprinodontae) from Virginia. *Chesapeake Science* 16: 282, 1975.
3. **Able KW, and Felley JD**. Geographical variation in Fundulus heteroclitus: tests for concordance between egg and adult morphologies. *American Zoologist* 26: 145-157, 1986.
4. **Able KW, Hagan SM, and 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. *Journal of Experimental Marine Biology and Ecology* 335: 177-187, 2006.
5. **Able KW, Vivian DN, Petruzzelli G, and Hagan SM**. Connectivity Among Salt Marsh Subhabitats: Residency and Movements of the Mummichog (Fundulus heteroclitus). *Estuar Coast* 35: 743-753, 2012.
6. **Adams SM, Lindmeier JB, and 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, and Kiviranta H**. Association Between Type 2 Diabetes and Exposure to Persistent Organic Pollutants. *Diabetes Care* 34: 1972-1979, 2011.
8. **Anholt RRH, and Mackay TFC**. The road less traveled: from genotype to phenotype in flies and humans. *Mamm Genome* 29: 5-23, 2018.
9. **Arnot JA, Armitage JM, McCarty LS, Wania F, Cousins IT, and Toose-Reid L**. Toward a consistent evaluative framework for POP risk characterization. *Environmental science & technology* 45: 97-103, 2011.
10. **Baris TZ, Blier PU, Pichaud N, Crawford DL, and Oleksiak MF**. Gene by environmental interactions affecting oxidative phosphorylation and thermal sensitivity. *American journal of physiology Regulatory, integrative and comparative physiology* 311: R157-165, 2016.
11. **Baris TZ, Crawford DL, and Oleksiak MF**. Acclimation and acute temperature effects on population differences in oxidative phosphorylation. *American journal of physiology Regulatory, integrative and comparative physiology* 310: R185-196, 2016.
12. **Baris TZ, Wagner DN, Dayan DI, Du X, Blier PU, Pichaud N, Oleksiak MF, and Crawford DL**. Evolved genetic and phenotypic differences due to mitochondrial-nuclear interactions. *PLoS Genet* 13: e1006517, 2017.
13. **Bello SM, Franks DG, Stegeman JJ, and 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. *Toxicological sciences : an official journal of the Society of Toxicology* 60: 77-91, 2001.
14. **Berg JJ, and Coop G**. A population genetic signal of polygenic adaptation. *PLoS Genet* 10: e1004412, 2014.

15. **Bernardi G, Sordino P, and 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.
16. **Betancur-R R, Orti G, and Pyron RA**. Fossil-based comparative analyses reveal ancient marine ancestry erased by extinction in ray-finned fishes. *Ecol Lett* 18: 441-450, 2015.
17. **Bloom AL**. Sea Level and Coastal Changes. In: *The Holocene*, edited by Wright WEJ. Minneapolis: University of Minnesota, 1983, p. 26-41.
18. **Bloom AL**. Sea Level and Coastal Morphology of the United States through the Late Wisconsin Glacial Maximum. In: *The Late Pleistocene*, edited by Porter SC. Minneapolis: University of Minnesota, 1983, p. 215-229.
19. **Bozinovic G, and Oleksiak MF**. Genomic approaches with natural fish populations from polluted environments. *Environmental toxicology and chemistry / SETAC* 30: 283-289, 2011.
20. **Bozinovic G, Sit TL, Di Giulio R, Wills LF, and 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.
21. **Bozinovic G, Sit TL, Hinton DE, and Oleksiak MF**. Gene expression throughout a vertebrate's embryogenesis. *BMC Genomics* 12: 132, 2011.
22. **Brennan RS, Healy TM, Bryant HJ, La MV, Schulte PM, and Whitehead A**. Genome-wide selection scans integrated with association mapping reveal mechanisms of physiological adaptation across a salinity gradient in killifish. *bioRxiv* 2018.
23. **Brennan RS, Hwang R, Tse M, Fangue NA, and 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.
24. **Brieuc MSO, Waters CD, Drinan DP, and Naish KA**. A practical introduction to Random Forest for genetic association studies in ecology and evolution. *Mol Ecol Resour* 18: 755-766, 2018.
25. **Brown PO, and Botstein D**. Exploring the new world of the genome with DNA microarrays. *Nat Genet* 21: 33-37, 1999.
26. **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, and 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.
27. **Cashon RE, Van Beneden RJ, and 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.
28. **Cavalli-Sforza LL, and Edwards AWF**. Phylogenetic analysis: Models and estimation procedures. *Evolution* 32: 550-570, 1967.

29. **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.
30. **Chung DJ, Bryant HJ, and 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.
31. **Chung DJ, Morrison PR, Bryant HJ, Jung E, Brauner CJ, and Schulte PM**. Intraspecific variation and plasticity in mitochondrial oxygen binding affinity as a response to environmental temperature. *Sci Rep* 7: 16238, 2017.
32. **Churchill GA**. Fundamentals of experimental design for cDNA microarrays. *Nat Genet* 32 Suppl: 490-495, 2002.
33. **Clark BW, Bone AJ, and 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. *Environmental science and pollution research international* 21: 13898-13908, 2014.
34. **Cline RM** editor. *Late Quaternary Paleoceanography and Paleoclimatology*. Boulder: Geological Society of America, 1976.
35. **Crawford DL, and Oleksiak MF**. The biological importance of measuring individual variation. *J Exp Biol* 210: 1613-1621, 2007.
36. **Crawford DL, Pierce VA, and Segal JA**. Evolutionary physiology of closely related taxa: Analyses of enzyme expression. *American Zoologist* 39: 389-400, 1999.
37. **Crawford DL, Place AR, and Powers DA**. Clinal Variation in the Specific Activity of Lactate Dehydrogenase-B. *J Exp Zool* 255: 110-113, 1990.
38. **Crawford DL, and Powers DA**. Evolutionary adaptation to different thermal environments via transcriptional regulation. *Mol Biol Evol* 9: 806-813, 1992.
39. **Crawford DL, and Powers DA**. Molecular Adaptation to the Thermal Environment: Genetic and Physiological Mechanisms. In: *Molecular Evolution: Proceeding of a UCLA Colloquium 1989*, edited by Clegg MT, and O'Brien SJ. New York: Wiley-Liss, 1990, p. 213-222.
40. **Crawford DL, and Powers DA**. Molecular-Basis of Evolutionary Adaptation at the Lactate Dehydrogenase-B Locus in the Fish Fundulus-Heteroclitus. *Proceedings of the National Academy of Sciences of the United States of America* 86: 9365-9369, 1989.
41. **Crawford DL, Segal JA, and Barnett JL**. Evolutionary analysis of TATA-less proximal promoter function. *Mol Biol Evol* 16: 194-207, 1999.
42. **Crow JF**. Some Possibilities for Measuring Selection Intensities in Man. *American Anthropologist* 60: 1-13, 1958.
43. **Dayan DI, Crawford DL, and Oleksiak MF**. Phenotypic plasticity in gene expression contributes to divergence of locally adapted populations of Fundulus heteroclitus. *Mol Ecol* 24: 3345-3359, 2015.
44. **Dayan DI, Du X, Baris TZ, Wagner DN, Crawford DL, and Oleksiak MF**. Population genomics of rapid evolution in natural populations: polygenic selection in response to power station thermal effluents. *submitted* 2017.
45. **DiMichele L, Paynter KT, and Powers DA**. Evidence of lactate dehydrogenase-B allozyme effects in the teleost, *Fundulus heteroclitus*. *Science* 253: 898-900, 1991.

46. **Dimichele L, and Powers DA**. Allozyme Variation, Developmental Rate, and Differential Mortality in the Teleost Fundulus-Heteroclitus. *Physiological Zoology* 64: 1426-1443, 1991.
47. **Dimichele L, and Powers DA**. Developmental and Oxygen-Consumption Rate Differences between Lactate Dehydrogenase-B Genotypes of Fundulus-Heteroclitus and Their Effect on Hatching Time. *Physiological Zoology* 57: 52-56, 1984.
48. **DiMichele L, and Powers DA**. LDH-B genotype-specific hatching times of *Fundulus heteroclitus* embryos. *Nature* 296: 563-564, 1982.
49. **Dimichele L, and Powers DA**. Physiological-Basis for Swimming Endurance Differences between Ldh-B Genotypes of Fundulus-Heteroclitus. *Science* 216: 1014-1016, 1982.
50. **DiMichele L, and Westerman M**. Geographic variation in development rate between populations of the teleost Fundulus heteroclitus. *Marine Biology* 128: 1-7, 1997.
51. **Donnelly JP, Driscoll NW, Uchupi E, Keigwin LD, Schwab WC, Thieler ER, and Swift SA**. Catastrophic meltwater discharge down the Hudson Valley: A potential trigger for the Intra-Aller√∏d cold period. *Geology* 33: 89-92, 2005.
52. **Du X, Crawford DL, Nacci DE, and Oleksiak MF**. Heritable oxidative phosphorylation differences in a pollutant resistant Fundulus heteroclitus population. *Aquat Toxicol* 177: 44-50, 2016.
53. **Du X, Crawford DL, and 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.
54. **Duforet-Frebourg N, Luu K, Laval G, Bazin E, and 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.
55. **Duvernell DD, Lindmeier JB, Faust KE, and 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.
56. **Eisen MB, and Brown PO**. DNA arrays for analysis of gene expression. *Methods Enzymol* 303: 179-205, 1999.
57. **Evans DH, Piermarini PM, and Choe KP**. The multifunctional fish gill: Dominant site of gas exchange, osmoregulation, acid-base regulation, and excretion of nitrogenous waste. *Physiological Reviews* 85: 97-177, 2005.
58. **Fader KA, and 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.
59. **Fangue NA, Hofmeister M, and Schulte PM**. Intraspecific variation in thermal tolerance and heat shock protein gene expression in common killifish, Fundulus heteroclitus. *Journal of Experimental Biology* 209: 2859-2872, 2006.
60. **Fangue NA, Richards JG, and Schulte PM**. Do mitochondrial properties explain intraspecific variation in thermal tolerance? *J Exp Biol* 212: 514-522, 2009.
61. **Fisher B**. Most unwanted. *Environ Health Perspect* 107: A18 - A23, 1999.
62. **Fraccalvieri D, Soshilov AA, Karchner SI, Franks DG, Pandini A, Bonati L, Hahn ME, and 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.
63. **Gaggiotti OE, Bekkevold D, Jorgensen HB, Foll M, Carvalho GR, Andre C, and 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.
64. **Garcia Ramos G, and Kirkpatrick M**. Genetic models of adaptation and gene flow in peripheral populations. *Evolution* 51: 21-28, 1997.
65. **Garland T, Jr., and Adolph SC**. Why not to do two-species comparative studies: limitations on inferring adaptation. *Physiological Zoology* 67: 797-828, 1994.
66. **Gonzalez-Villasenor LI, and Powers DA**. Mitochondrial-DNA Restriction-Site Polymorphisms in the Teleost Fundulus Heteroclitus Support Secondary Intergradation. *Evolution* 44: 27-37, 1990.
67. **Griffith RW**. Environment and salinity tolerance in the genus *Fundulus*. *Copeia* 319-331, 1974.
68. **Haas HL, Freeman CJ, Logan JM, Deegan L, and Gaines EF**. Examining mummichog growth and movement: Are some individuals making intra-season migrations to optimize growth? *Journal of Experimental Marine Biology and Ecology* 369: 8-16, 2009.
69. **Haldane JBS**. The cost of natural selection. *J Genetics* 55: 511-524, 1957.
70. **Halpin PM**. Habitat use patterns of the mummichog, *Fundulus heteroclitus,* in New England. I. Intramarsh variation. *Estuaries* 20: 618-625, 1997.
71. **Haney RA, Dionne M, Puritz J, and 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.
72. **Hartl DL, Dykhuizen DE, and Dean AM**. Limits of adaptation: the evolution of selective neutrality. *Genetics* 111: 655-674, 1985.
73. **Healy TM, Brennan RS, Whitehead A, and Schulte PM**. Tolerance traits related to climate change resilience are independent and polygenic. *Glob Chang Biol* 2018.
74. **Healy TM, Brennan RS, Whitehead A, and Schulte PM**. Tolerance traits related to climate change resilience are independent and polygenic. *Glob Chang Biol* 24: 5348-5360, 2018.
75. **Healy TM, Bryant HJ, and Schulte PM**. Mitochondrial genotype and phenotypic plasticity of gene expression in response to cold acclimation in killifish. *Mol Ecol* 26: 814-830, 2017.
76. **Healy TM, and Schulte PM**. Factors affecting plasticity in whole-organism thermal tolerance in common killifish (Fundulus heteroclitus). *J Comp Physiol B* 182: 49-62, 2012.
77. **Henn BM, Botigue LR, Bustamante CD, Clark AG, and Gravel S**. Estimating the mutation load in human genomes. *Nat Rev Genet* 16: 333-343, 2015.
78. **Hermisson J, and Pennings PS**. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* 169: 2335-2352, 2005.
79. **Hermisson J, and Wagner GP**. The population genetic theory of hidden variation and genetic robustness. *Genetics* 168: 2271-2284, 2004.

80. **Hochachka PW, and Somero GN**. *Biochemical Adaptation, Mechanism and Process in Physiological Evolution*. New York, NY: Oxford University Press, 2002, p. 408.
81. **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, and Mackay TF**. Epistasis dominates the genetic architecture of Drosophila quantitative traits. *Proc Natl Acad Sci U S A* 109: 15553-15559, 2012.
82. **Hunter KL, Fox MG, and Able KW**. Habitat influences on reproductive allocation and growth of the mummichog (Fundulus heteroclitus) in a coastal salt marsh. *Marine Biology* 151: 617-627, 2007.
83. **Imbrie J, McIntyre A, and Moore TCJ**. The Ocean around North America at the Last Glacial Maximum. In: *The Late Pleistocene*, edited by Porter SC. Minneapolis: University of Minnesota, 1983, p. 230-238.
84. **Jha AR, Zhou D, Brown CD, Kreitman M, Haddad GG, and White KP**. Shared Genetic Signals of Hypoxia Adaptation in Drosophila and in High-Altitude Human Populations. *Mol Biol Evol* 33: 501-517, 2016.
85. **Jombart T, Devillard S, and Balloux F**. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* 11: 94, 2010.
86. **Kanehisa M, and Goto S**. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28: 27-30, 2000.
87. **Karami-Mohajeri S, and 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.
88. **Kerr MK, and Churchill GA**. Statistical design and the analysis of gene expression microarray data. *Genet Res* 77: 123-128, 2001.
89. **Kimura M**. Evolutionary Rate at the Molecular Level. *Nature* 217: 624-626, 1968.
90. **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.
91. **Kimura M, and Crow JF**. The number of alleles that can be maintained in a finite population. *Genetics* 49: 725-738, 1964.
92. **Kneib RT**. The Role of Fundulus-Heteroclitus in Salt-Marsh Trophic Dynamics. *American Zoologist* 26: 259-269, 1986.
93. **Knox JC**. Responses of River Systems to Holocene Climates. In: *The Holocene*, edited by Wright WEJ. Minneapolis: University of Minnesota, 1983, p. 26-41.
94. **Kong SW, Lee IH, Leshchiner I, Krier J, Kraft P, Rehm HL, Green RC, Kohane IS, MacRae CA, and MedSeq P**. Summarizing polygenic risks for complex diseases in a clinical whole-genome report. *Genetics in medicine : official journal of the American College of Medical Genetics* 17: 536-544, 2015.
95. **Kreitman M**. The neutral theory is dead. Long live the neutral theory. *Bioessays* 18: 678-683; discussion 683, 1996.
96. **Laporte M, Pavey SA, Rougeux C, Pierron F, Lauzent M, Budzinski H, Labadie P, Geneste E, Couture P, Baudrimont M, and 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.
97. **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 Biology*Johns Hopkins University, 1990.
98. **Lee D, Lee I, Song K, Steffes M, Toscano W, Baker B, and 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.
99. **Lee DH, Lind L, Jacobs DR, Jr., Salihovic S, van Bavel B, and 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.
100. **Lee SH, Ra JS, Choi JW, Yim BJ, Jung MS, and Kim SD**. Human health risks associated with dietary exposure to persistent organic pollutants (POPs) in river water in Korea. *The Science of the total environment* 470-471: 1362-1369, 2014.
101. **Lewontin RC, and 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.
102. **Li QQ, Loganath A, Chong YS, Tan J, and Obbard JP**. Persistent organic pollutants and adverse health effects in humans. *Journal of toxicology and environmental health Part A* 69: 1987-2005, 2006.
103. **Lim S, Cho YM, Park KS, and Lee HK**. Persistent organic pollutants, mitochondrial dysfunction, and metabolic syndrome. *Ann N Y Acad Sci* 1201: 166-176, 2010.
104. **Lotrich VA**. Summer Home Range and Movements of Fundulus Heteroclitus (Pisces: Cyprinodontidae) in a Tidal Creek. *Ecology* 56: 191-198, 1975.
105. **Mackay TF**. Epistasis for quantitative traits in Drosophila. *Methods Mol Biol* 1253: 47-70, 2015.
106. **Mackay TF, and Moore JH**. Why epistasis is important for tackling complex human disease genetics. *Genome Med* 6: 124, 2014.
107. **McBryan TL, Anttila K, Healy TM, and Schulte PM**. Responses to Temperature and Hypoxia as Interacting Stressors in Fish: Implications for Adaptation to Environmental Change. *Integrative and Comparative Biology* 53: 648-659, 2013.
108. **McBryan TL, Healy TM, Haakons KL, and Schulte PM**. Warm acclimation improves hypoxia tolerance in Fundulus heteroclitus. *J Exp Biol* 219: 474-484, 2016.
109. **McDonald JH, and Kreitman M**. Adaptive Protein Evolution at the Adh Locus in Drosophila. *Nature* 351: 652-654, 1991.
110. **McIntyre A, Kipp NG, Be AWH, Crowley T, Kellogg T, Gardner JV, Prell W, and Ruddiman WF**. Glacial North America 18,000 years ago: a climate reconstruction. In: *Late Quaternary Paleoceanography and Paleoclimatology*, edited by Cline RM, and Hays JD. Boulder: Geological Society of America, 1976, p. 43-76.
111. **McKenzie JL, Dhillon RS, and 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.

112. **Meyer J, and Di Giulio R**. Heritable adaptation and fitness costs in killifish (Fundulus heteroclitus) inhabiting a polluted estuary. *Ecological Applications* 13: 490-503, 2003.
113. **Meyer JN, Wassenberg DM, Karchner SI, Hahn ME, and Di Giulio RT**. Expression and inducibility of aryl hydrocarbon receptor pathway genes in wild-caught killifish (Fundulus heteroclitus) with different contaminant-exposure histories. *Environmental toxicology and chemistry / SETAC* 22: 2337-2343, 2003.
114. **Mickelson DM, Clayton L, Fullerton DS, and Borns HWJ**. The Late Wisconsin Glacial Record of Laurentide Ice Sheet in the United States. In: *The Late Pleistocene*, edited by Porter SC. Minneapolis: University of Minnesota, 1983, p. 3-37.
115. **Mitton JB, and Koehn RK**. Genetic organization and adaptive response of allozymes to ecological variables in Fundulus heteroclitus. *Genetics* 79: 97-111, 1975.
116. **Morin RP, and Able KW**. Patterns of geographic variation in the egg morphology of the fundulid fish, Fundulus heteroclitus. *Copeia* 726-740, 1983.
117. **Nacci D, Champlin D, Coiro L, McKinney R, and Jayaraman S**. Predicting the occurrence of genetic adaptation to dioxin like compounds in populations of the estuarine fish Fundulus heteroclitus. *Environmental Toxicology and Chemistry* 21: 1525-1532, 2002.
118. **Nacci D, Coiro L, Champlin D, Jayaraman S, McKinney R, Gleason TR, Munns WR, Specker JL, and Cooper KR**. Adaptations of wild populations of the estuarine fish Fundulus heteroclitus to persistent environmental contaminants. *Marine Biology* 134: 9-17, 1999.
119. **Nacci D, Huber M, Champlin D, Jayaraman S, Cohen S, Gauger E, Fong A, and 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. *Environmental pollution* 157: 857-864, 2009.
120. **Nacci DE, Champlin D, and Jayaraman S**. Adaptation of the Estuarine Fish Fundulus heteroclitus (Atlantic Killifish) to Polychlorinated Biphenyls (PCBs). *Estuar Coast* 33: 853-864, 2010.
121. **Nei M, Suzuki Y, and Nozawa M**. The neutral theory of molecular evolution in the genomic era. In: *Annual Review of Genomics and Human Genetics*2010, p. 265-289.
122. **Ohta T**. The Nearly Neutral Theory of Molecular Evolution. *Annual Review of Ecology and Systematics* 23: 263-286, 1992.
123. **Oleksiak M, F., and Crawford Douglas L**. Functional Genomics in Fishes, Insights into Physiological Complexity. In: *The Physiology of Fishes*, edited by Evan D, and Claiborne J. Boca Raton: CRC Press, 2006, p. 523-550.
124. **Oleksiak MF**. Marine genomics: insights and challenges. *Brief Funct Genomics* 15: 331-332, 2016.
125. **Oleksiak MF, Churchill GA, and Crawford DL**. Variation in gene expression within and among natural populations. *Nature Genetics* 32: 261-266, 2002.
126. **Oleksiak MF, Kolell KJ, and Crawford DL**. Utility of natural populations for microarray analyses: isolation of genes necessary for functional genomic studies. *Mar Biotechnol (NY)* 3: S203-211, 2001.
127. **Oleksiak MF, Roach JL, and Crawford DL**. Natural variation in cardiac metabolism and gene expression in Fundulus heteroclitus. *Nat Genet* 37: 67-72, 2005.

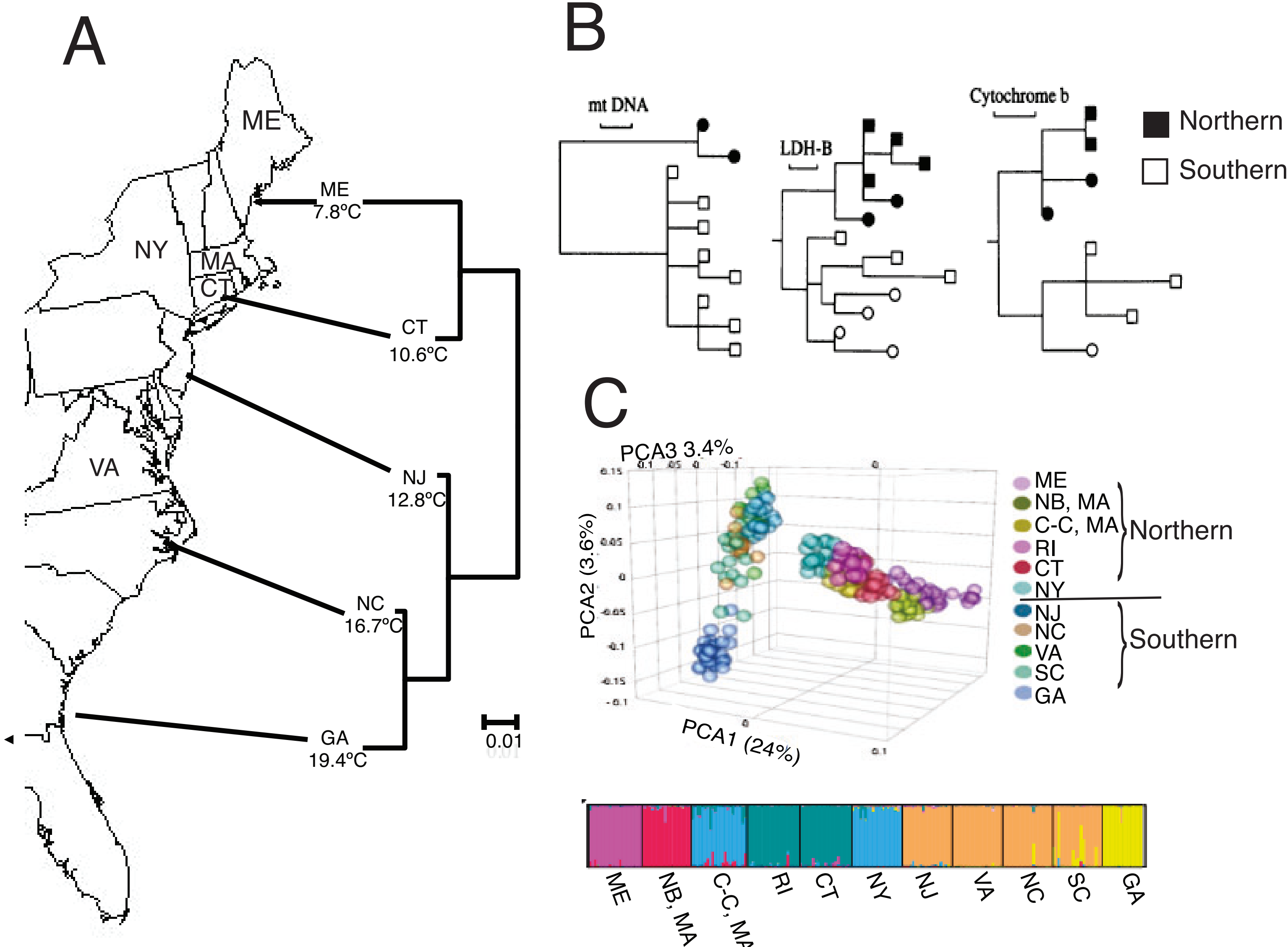
128. **Palumbi SR, Sidell BD, Vanbeneden R, Smith GD, and Powers DA**. Glucosephosphate Isomerase (Gpi) of the Teleost Fundulus-Heteroclitus (Linnaeus) - Isozymes, Allozymes and Their Physiological Roles. *Journal of Comparative Physiology* 138: 49-57, 1980.
129. **Paynter KT, Dimichele L, Hand SC, and Powers DA**. Metabolic Implications of Ldh-B Genotype during Early Development in Fundulus-Heteroclitus. *J Exp Zool* 257: 24-33, 1991.
130. **Pierce VA, and Crawford DL**. Effect of acclimation on glycolytic enzymes in Atlantic and Gulf Coast Fundulus species. *American Zoologist* 35: 99A-99A, 1995.
131. **Pierce VA, and Crawford DL**. Phylogenetic analysis of glycolytic enzyme expression. *Science* 276: 256-259, 1997.
132. **Pierce VA, and Crawford DL**. Phylogenetic analysis of thermal acclimation of the glycolytic enzymes in the genus Fundulus. *Physiol Zool* 70: 597-609, 1997.
133. **Pierce VA, and Crawford DL**. Variation in the glycolytic pathway: The role of evolutionary and physiological processes. *Physiological Zoology* 69: 489-508, 1996.
134. **Place AR, and 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.
135. **Place AR, and 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.
136. **Place AR, and Powers DA**. Kinetic characterization of the lactate dehydrogenase (LDH-B4) allozymes of *Fundulus heteroclitus*. *J Biol Chem* 259: 1309-1318, 1984.
137. **Podrabsky JE, Javillonar C, Hand SC, and Crawford DL**. Intraspecific variation in aerobic metabolism and glycolytic enzyme expression in heart ventricles. *American journal of physiology Regulatory, integrative and comparative physiology* 279: R2344-2348, 2000.
138. **Powell WH, Bright R, Bello SM, and Hahn ME**. Developmental and tissue-specific expression of AHR1, AHR2, and ARNT2 in dioxin-sensitive and -resistant populations of the marine fish Fundulus heteroclitus. *Toxicological sciences : an official journal of the Society of Toxicology* 57: 229-239, 2000.
139. **Powers DA, Lauerman T, Crawford D, and DiMichele L**. Genetic mechanisms for adapting to a changing environment. *Annu Rev Genet* 25: 629-659, 1991.
140. **Powers DA, and Place AR**. Biochemical genetics of Fundulus heteroclitus (L.). I. Temporal and spatial variation in gene frequencies of Ldh-B, Mdh-A, Gpi-B, and Pgm-A. *Biochemical Genetics* 16: 593-607, 1978.
141. **Powers DA, Ropson I, Brown DC, Vanbeneden R, Cashon R, Gonzalezvillasenor LI, and Dimichele JA**. Genetic-Variation in Fundulus-Heteroclitus - Geographic-Distribution. *American Zoologist* 26: 131-144, 1986.
142. **Powers DA, and 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.

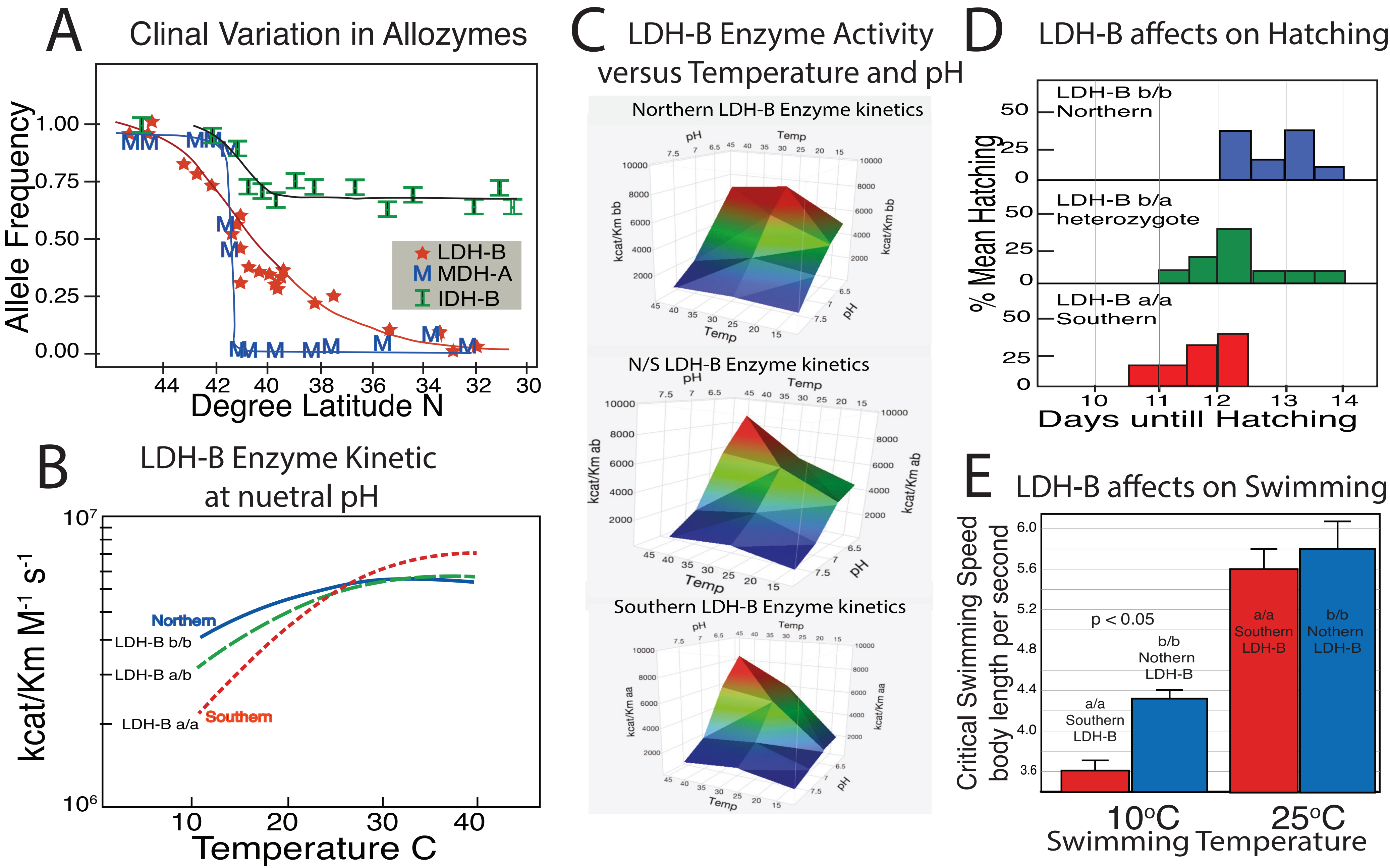
143. **Powers DA, Smith M, Gonzalez-Villasenor I, DiMichele L, Crawford DL, Bernardi G, and Lauerman T**. A multidisciplinary approach to the selectionist/neutralist controversy using the model teleost, *Fundulus heteroclitus*. In: *Oxford Surveys in Evolutionary Biology*, edited by Futuyma D, and Antonovics J. New York, NY: Oxford University Press, 1993, p. 43-108.
144. **Pritchard JK, Pickrell JK, and Coop G**. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr Biol* 20: R208-215, 2010.
145. **Pritchard JK, Stephens M, and Donnelly P**. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959, 2000.
146. **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, and Whitehead A**. The landscape of extreme genomic variation in the highly adaptable Atlantic killifish. *Genome Biol Evol* 2017.
147. **Reid NM, Proestou DA, Clark BW, Warren WC, Colbourne JK, Shaw JR, Karchner SI, Hahn ME, Nacci D, Oleksiak MF, Crawford DL, and Whitehead A**. The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science* 354: 1305-1308, 2016.
148. **Rockman MV**. The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* 66: 1-17, 2012.
149. **Ropson IJ, Brown EC, and Powers DA**. Biochemical Genetics of Fundulus Heteroclitus (L.). Vi. Geographical Variation in the Gene Frequencies of 15 Loci. *Evolution* 44: 16-26, 1990.
150. **Ruzzin J**. Public health concern behind the exposure to persistent organic pollutants and the risk of metabolic diseases. *BMC Public Health* 12: 298, 2012.
151. **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.
152. **Schulte PM**. What is environmental stress? Insights from fish living in a variable environment. *J Exp Biol* 217: 23-34, 2014.
153. **Schulte PM, Glemet HC, Fiebig AA, and 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.
154. **Schulte PM, GomezChiarri M, and 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.
155. **Schultz ET, and McCormick SD**. Euryhalinity in an Evolutionary Context. *Euryhaline Fishes* 32: 477-533, 2013.
156. **Scott GR, Rogers JT, Richards JG, Wood CM, and 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.
157. **Segal JA, Barnett JL, and Crawford DL**. Functional analyses of natural variation in Sp1 binding sites of a TATA-less promoter. *J Mol Evol* 49: 736-749, 1999.
158. **Segal JA, and Crawford DL**. LDH-B enzyme expression: the mechanisms of altered gene expression in acclimation and evolutionary adaptation. *Am J Physiol* 267: R1150-1153, 1994.

159. **Segal JA, Schulte PM, Powers DA, and Crawford DL**. Descriptive and functional characterization of variation in the *Fundulus heteroclitus Ldh-B* proximal promoter. *J Exp Zool* 275: 355-364, 1996.
160. **Shaw JR, Hampton TH, King BL, Whitehead A, Galvez F, Gross RH, Keith N, Notch E, Jung D, Glaholt SP, Chen CY, Colbourne JK, and Stanton BA**. Natural Selection Canalizes Expression Variation of Environmentally Induced Plasticity-Enabling Genes. *Molecular Biology and Evolution* 31: 3002-3015, 2014.
161. **Sidell BD, Johnston IA, Moerland TS, and 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.
162. **Skinner MA, Courtenay SC, Parker WR, and Curry RA**. Site fidelity of mummichogs (Fundulus heteroclitus) in an Atlantic Canadian estuary. *Water Quality Research Journal of Canada* 40: 288-298, 2005.
163. **Skinner MA, Courtenay SC, Parker WR, and Curry RA**. Stable isotopic assessment of site fidelity of mummichogs, Fundulus heteroclitus, exposed to multiple anthropogenic inputs. *Environmental Biology of Fishes* 94: 695-706, 2012.
164. **Slatkin M**. Gene flow and population structure. In: *Ecological Genetics*, edited by Real LA. Princeton, NJ: Princeton University Press, 1994, p. 3-17.
165. **Slatkin M**. Gene flow and the geographic structure of natural populations. *Science* 236: 787-792, 1987.
166. **Smith KJ, and Able KW**. Dissolved oxygen dynamics in salt marsh pools and its potential impacts on fish assemblages. *Marine Ecology Progress Series* 258: 223-232, 2003.
167. **Smith MW, C. GM, and A. PD**. Genetic introgression of nuclear alleles between populations of the teleost Fundulus heteroclitus. *Molecular Marine Biology & Biotechnology* 1: 226-238, 1992.
168. **Smith MW, Chapman RW, and 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. *Molecular Marine Biology and Biotechnology* 7: 79-87, 1998.
169. **Somero GN**. Proteins and temperature. *Annual Review of Physiology* 57: 43-68, 1995.
170. **Somero GN, Lockwood BL, and Tomanek L**. *Biochemical Adaptation: Response to Environmental Challenges from Life's Origins to the Anthropocene*. Oxford University Press, 2017.
171. **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, and Andreassen OA**. Genetic Markers of Human Evolution Are Enriched in Schizophrenia. *Biological psychiatry* 80: 284-292, 2016.
172. **Stockinger B, Di Meglio P, Gialitakis M, and Duarte JH**. The aryl hydrocarbon receptor: multitasking in the immune system. *Annu Rev Immunol* 32: 403-432, 2014.
173. **Strange RM, and Burr BM**. Intraspecific Phylogeography of North American Highland Fishes: A Test of the Pleistocene Vicariance Hypothesis. *Evolution* 51: 885-897, 1997.

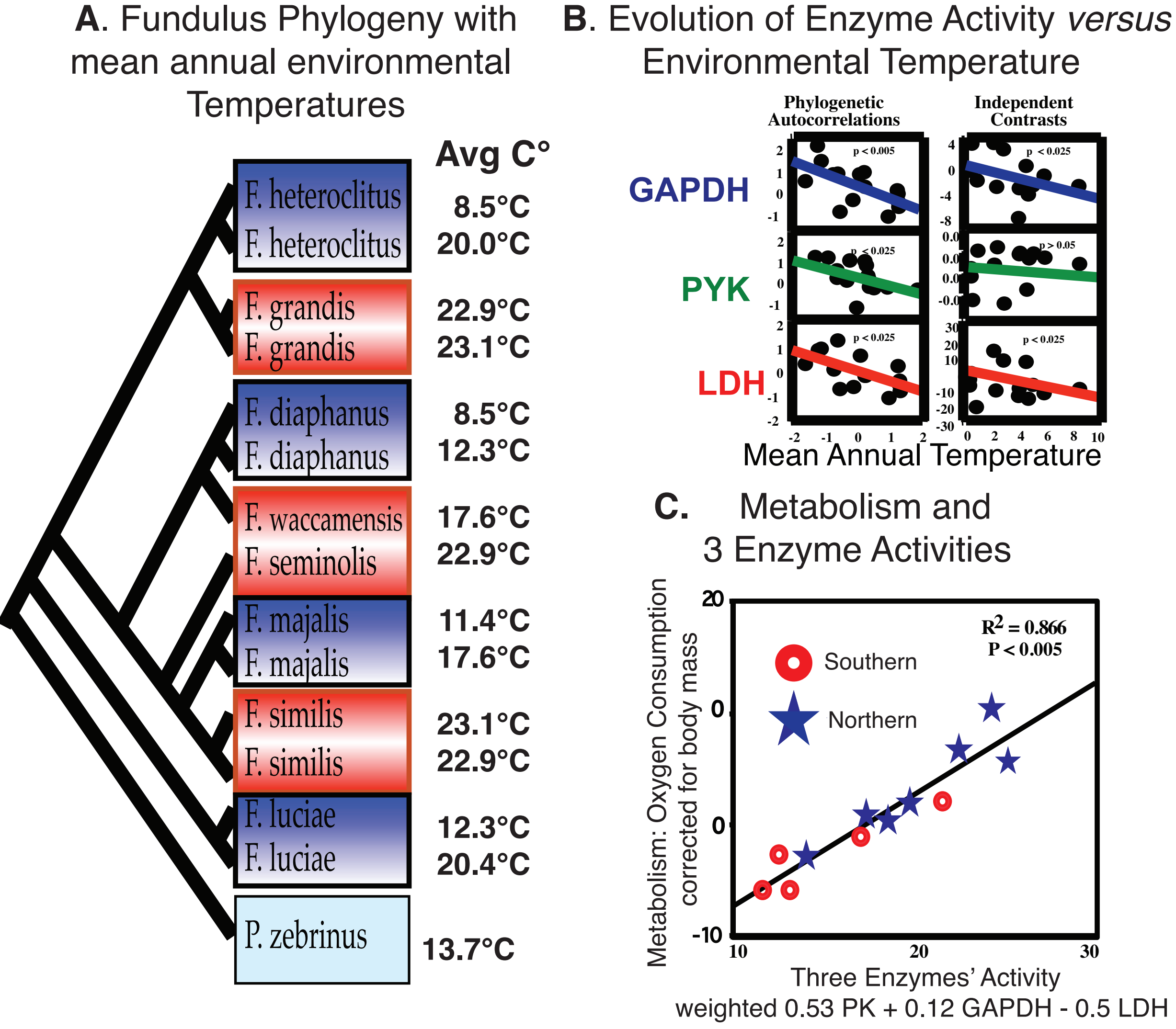
174. **Taylor MH, Michele LD, and Leach GJ**. Egg Stranding in the Life Cycle of the Mummichog, Fundulus heteroclitus. *Copeia* 1977: 397, 1977.
175. **Teo SLH, and Able KW**. Growth and production of the mummichog (Fundulus heteroclitus) in a restored salt marsh. *Estuaries* 26: 51-63, 2003.
176. **Teo SLH, and Able KW**. Habitat use and movement of the mummichog (Fundulus heteroclitus) in a restored salt marsh. *Estuaries* 26: 720-730, 2003.
177. **Turchin MC, Chiang CWK, Palmer CD, Sankararaman S, Reich D, and Hirschhorn JN**. Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nat Genet* 44: 1015-1019, 2012.
178. **Valiela I, Wright JE, Teal JM, and Volkmann SB**. Growth, production and energy transformations in the salt-marsh killifish Fundulus heteroclitus. *Marine Biology* 40: 135-144, 1977.
179. **Van Beneden RJ, Cashon RE, and 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.
180. **Vassy JL, Hivert MF, Porneala B, Dauriz M, Florez JC, Dupuis J, Siscovick DS, Fornage M, Rasmussen-Torvik LJ, Bouchard C, and Meigs JB**. Polygenic type 2 diabetes prediction at the limit of common variant detection. *Diabetes* 63: 2172-2182, 2014.
181. **Vega GC, and Wiens JJ**. Why are there so few fish in the sea? *P Roy Soc B-Biol Sci* 279: 2323-2329, 2012.
182. **Visscher PM, Brown MA, McCarthy MI, and Yang J**. Five years of GWAS discovery. *Am J Hum Genet* 90: 7-24, 2012.
183. **Wagner DN, Baris TZ, Dayan DI, Du X, Oleksiak MF, and Crawford DL**. Fine-scale genetic structure due to adaptive divergence among microhabitats. *Heredity (Edinb)* 118: 594-604, 2017.
184. **Wallace B**. *Fifty years of genetic load: an odyssey*. Ithaca, NY: Cornell University Press, 1991, p. 174.
185. **Wayne ML, and Simonsen KL**. Statistical tests of neutrality in the age of weak selection. *Trends Ecol Evol* 13: 236-240, 1998.
186. **Whitehead A**. Comparative mitochondrial genomics within and among species of killifish. *BMC Evol Biol* 9: 11, 2009.
187. **Whitehead A**. The evolutionary radiation of diverse osmotolerant physiologies in killifish (Fundulus sp.). *Evolution* 64: 2070-2085, 2010.
188. **Whitehead A, and Crawford DL**. Neutral and adaptive variation in gene expression. *Proc Natl Acad Sci U S A* 103: 5425-5430, 2006.
189. **Whitehead A, and Crawford DL**. Variation within and among species in gene expression: raw material for evolution. *Mol Ecol* 15: 1197-1211, 2006.
190. **Whitehead A, Roach JL, Zhang S, and 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.
191. **Williams LM, Ma X, Boyko AR, Bustamante CD, and Oleksiak MF**. SNP identification, verification, and utility for population genetics in a non-model genus. *BMC Genet* 11: 32, 2010.

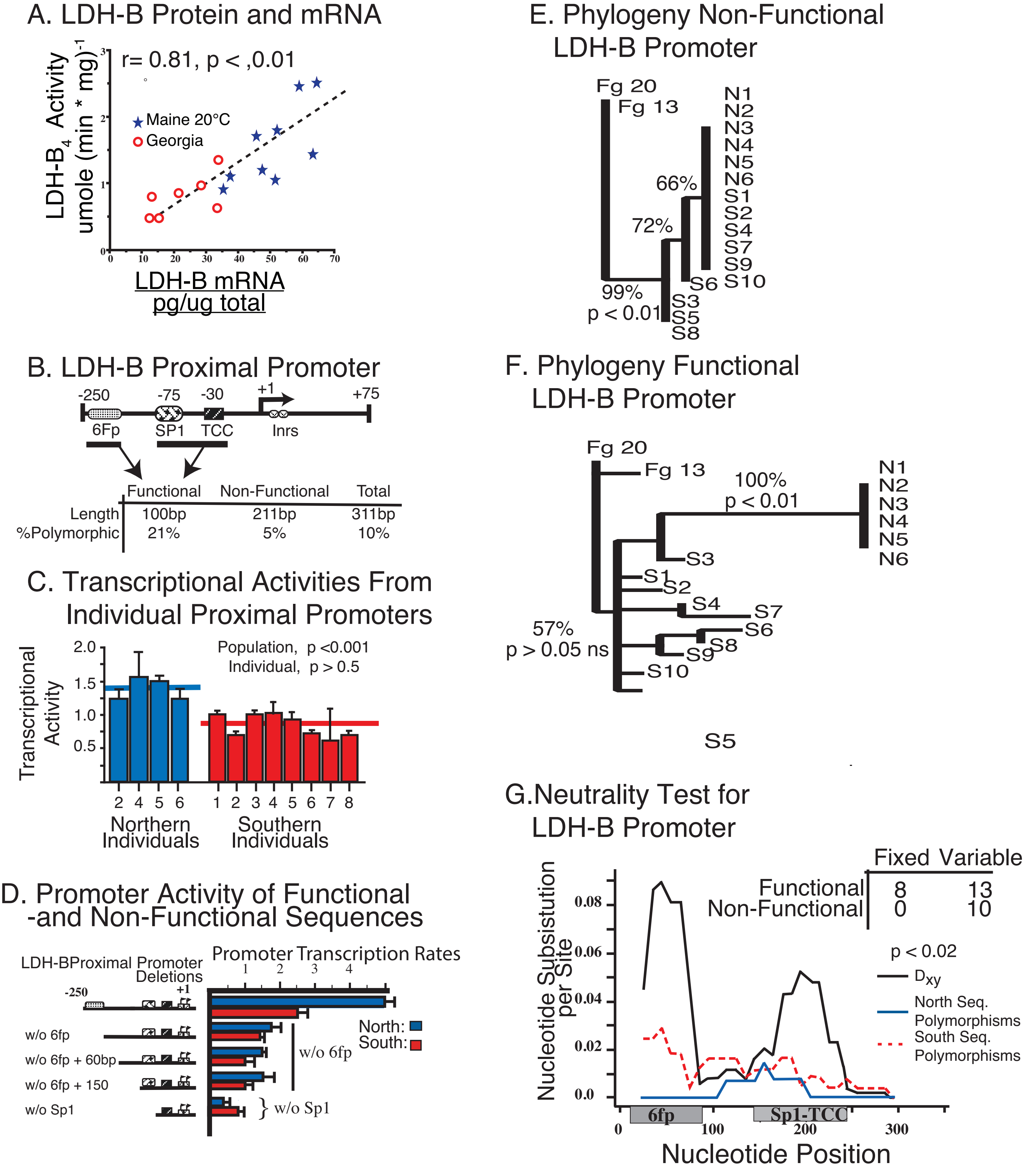
192. **Williams LM, and Oleksiak MF**. Ecologically and evolutionarily important SNPs identified in natural populations. *Mol Biol Evol* 28: 1817-1826, 2011.
193. **Williams LM, and Oleksiak MF**. Evolutionary and functional analyses of cytochrome P4501A promoter polymorphisms in natural populations. *Mol Ecol* 20: 5236-5247, 2011.
194. **Williams LM, and Oleksiak MF**. Signatures of selection in natural populations adapted to chronic pollution. *BMC Evol Biol* 8: 282, 2008.
195. **Wood CM, and 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.
196. **Yang J, Lee T, Kim J, Cho MC, Han BG, Lee JY, Lee HJ, Cho S, and Kim H**. Ubiquitous polygenicity of human complex traits: genome-wide analysis of 49 traits in Koreans. *PLoS Genet* 9: e1003355, 2013.
197. **Yu HY, Guo Y, and Zeng EY**. Dietary intake of persistent organic pollutants and potential health risks via consumption of global aquatic products. *Environmental toxicology and chemistry / SETAC* 29: 2135-2142, 2010.


A
ME
NY
MA
CT
VA
ME
7.8ºC
CT
10.6ºC
NJ
12.8ºC
NC
16.7ºC
GA
19.4ºC
0.01
B
mt DNA
LDH-B
Cytochrome b
Northern
Southern
C
PCA3 3.4%
PCA2 (3.6%)
PCA1 (24%)
ME
NB, MA
C-C, MA
RI
CT
NY
Northern
NJ
NC
VA
SC
GA
Southern
ME
NB, MA
C-C, MA
RI
CT
NY
NJ
VA
NC
SC
GA



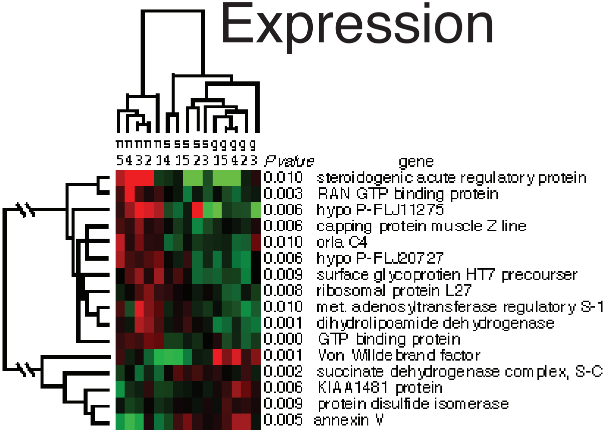
A Clinal Variation in Allozymes
Allele Frequency
1.00
0.75
0.50
0.25
0.00
44 42 40 38 36 34 32 30
Degree Latitude N
LDH-B
MDH-A
IDH-B
B LDH-B Enzyme Kinetic at nuetral pH
kcat/Km M-1 s-1
Northern
LDH-B b/b
LDH-B a/b
LDH-B a/a
Southern
10 20 30 40
Temperature C
C LDH-B Enzyme Activity versus Temperature and pH
Northern LDH-B Enzyme kinetics
N/S LDH-B Enzyme kinetics
Southern LDH-B Enzyme kinetics
pH
Temp
kcat/Km bb
kcat/Km ab
kcat/Km aa
D LDH-B affects on Hatching
% Mean Hatching
LDH-B b/b
Northern
LDH-B b/a
heterozygote
LDH-B a/a
Southern
10 11 12 13 14
Days untill Hatching
E LDH-B affects on Swimming
Critical Swimming Speed
body length per second
p < 0.05
a/a
Southern
LDH-B
b/b
Nothern
LDH-B
10°C
25°C
Swimming Temperature


A. Fundulus Phylogeny with mean annual environmental Temperatures
Avg C°
F. heteroclitus 8.5°C
F. heteroclitus 20.0°C
F. grandis 22.9°C
F. grandis 23.1°C
F. diaphanus 8.5°C
F. diaphanus 12.3°C
F. waccamensis 17.6°C
F. seminolis 22.9°C
F. majalis 11.4°C
F. majalis 17.6°C
F. similis 23.1°C
F. similis 22.9°C
F. luciae 12.3°C
F. luciae 20.4°C
P. zebrinus 13.7°C
B. Evolution of Enzyme Activity versus Environmental Temperature
Phylogenetic Autocorrelations
Independent Contrasts
GAPDH
p < 0.005
p < 0.025
PYK
p < 0.025
p > 0.05
LDH
p < 0.025
p < 0.025
Mean Annual Temperature
C. Metabolism and 3 Enzyme Activities
R2 = 0.866
P < 0.005
Southern
Northern
Metabolism: Oxygen Consumption corrected for body mass
Three Enzymes' Activity
weighted 0.53 PK + 0.12 GAPDH - 0.5 LDH

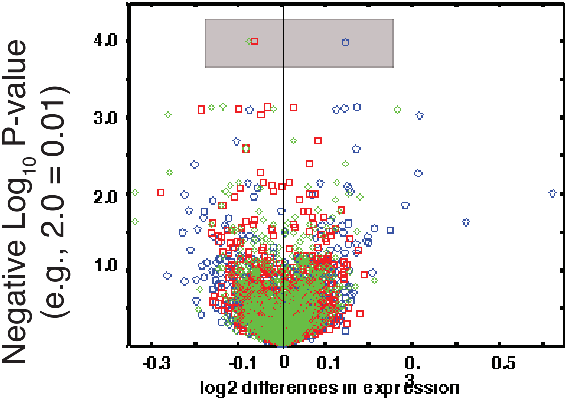


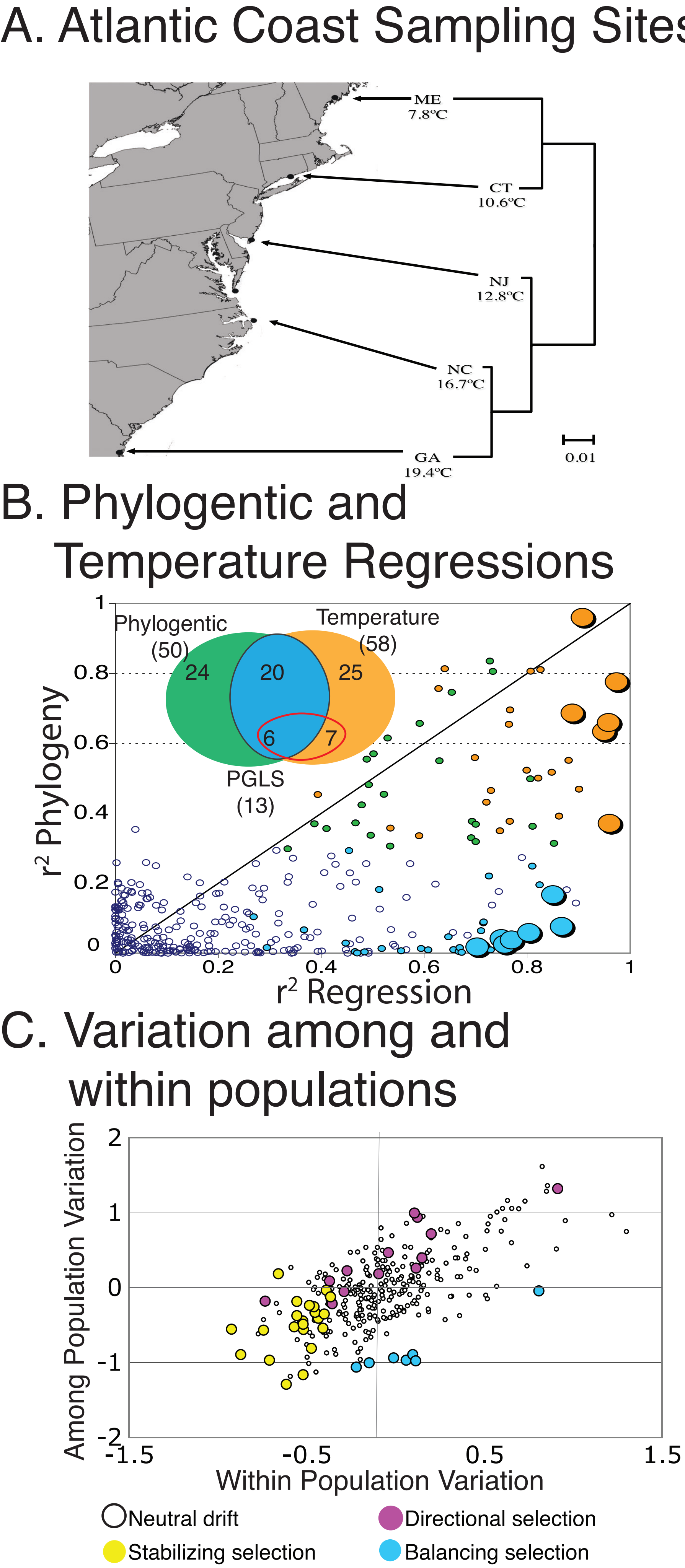
A. LDH-B Protein and mRNA
r= 0.81, p < ,0.01
Maine 20°C
Georgia
LDH-B4 Activity umole (min * mg)-1
LDH-B mRNA pg/ug total
B. LDH-B Proximal Promoter
-250
-75
-30
+1
+75
6Fp
SP1
TCC
Inrs
Functional
Non-Functional
Total
Length
100bp
211bp
311bp
%Polymorphic
21%
5%
10%
C. Transcriptional Activities From Individual Proximal Promoters
Population, p <0.001
Individual, p > 0.5
Transcriptional Activity
Northern Individuals
Southern Individuals
D. Promoter Activity of Functional -and Non-Functional Sequences
LDH-BProximal Promoter Deletions
Promoter Transcription Rates
w/o 6fp
w/o 6fp + 60bp
w/o 6fp + 150
w/o Sp1
North:
South:
E. Phylogeny Non-Functional LDH-B Promoter
Fg 20
Fg 13
66%
72%
99%
p < 0.01
F. Phylogeny Functional LDH-B Promoter
100%
p < 0.01
57%
p > 0.05 ns
G.Neutrality Test for LDH-B Promoter
Fixed
Variable
Functional
8
13
Non-Functional
0
10
p < 0.02
Dxy
North Seq. Polymorphisms
South Seq. Polymorphisms
Nucleotide Subsistution per Site
6fp
Sp1-TCC
Nucleotide Position

A. Adaptively Significant mRNA Expression

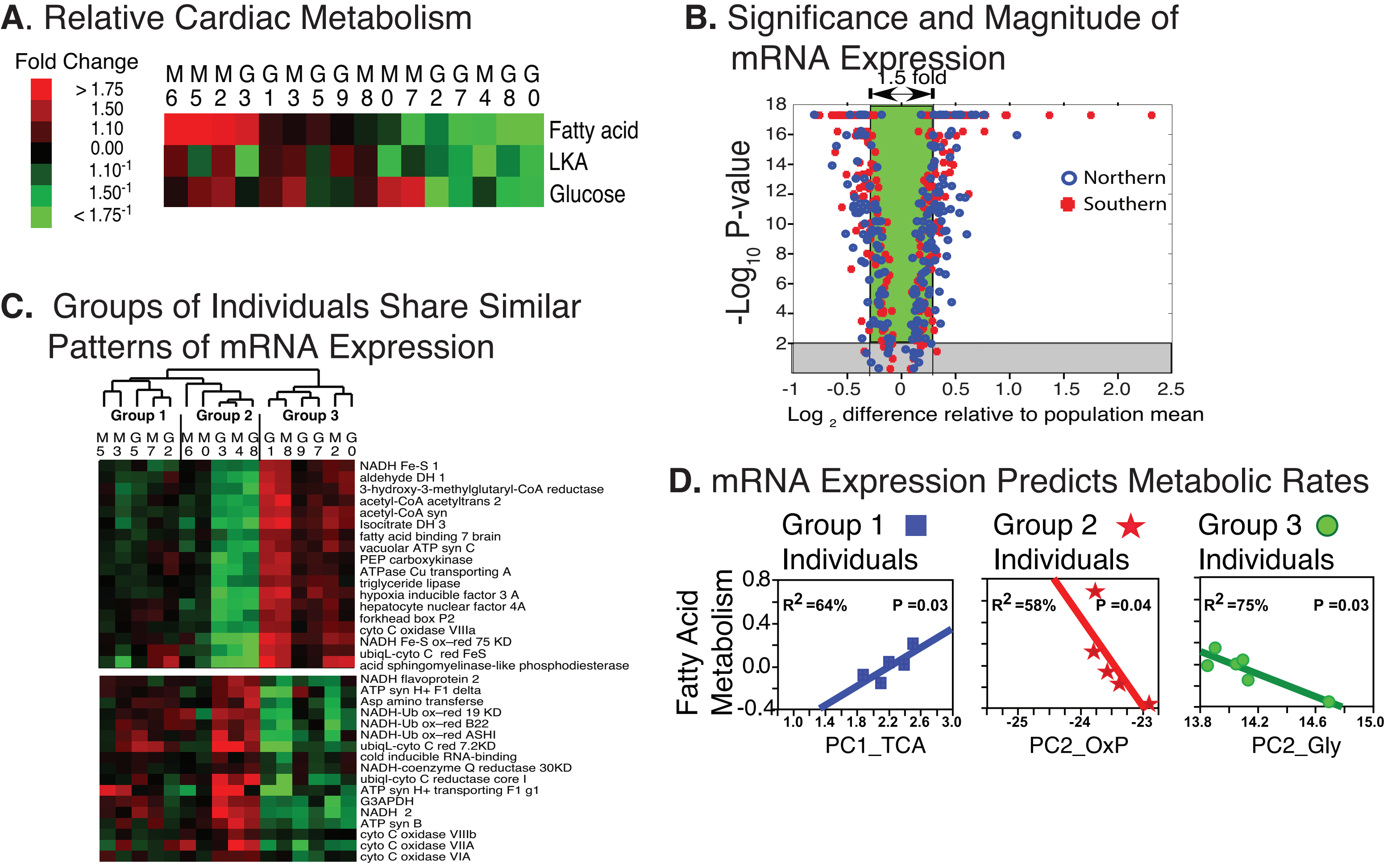


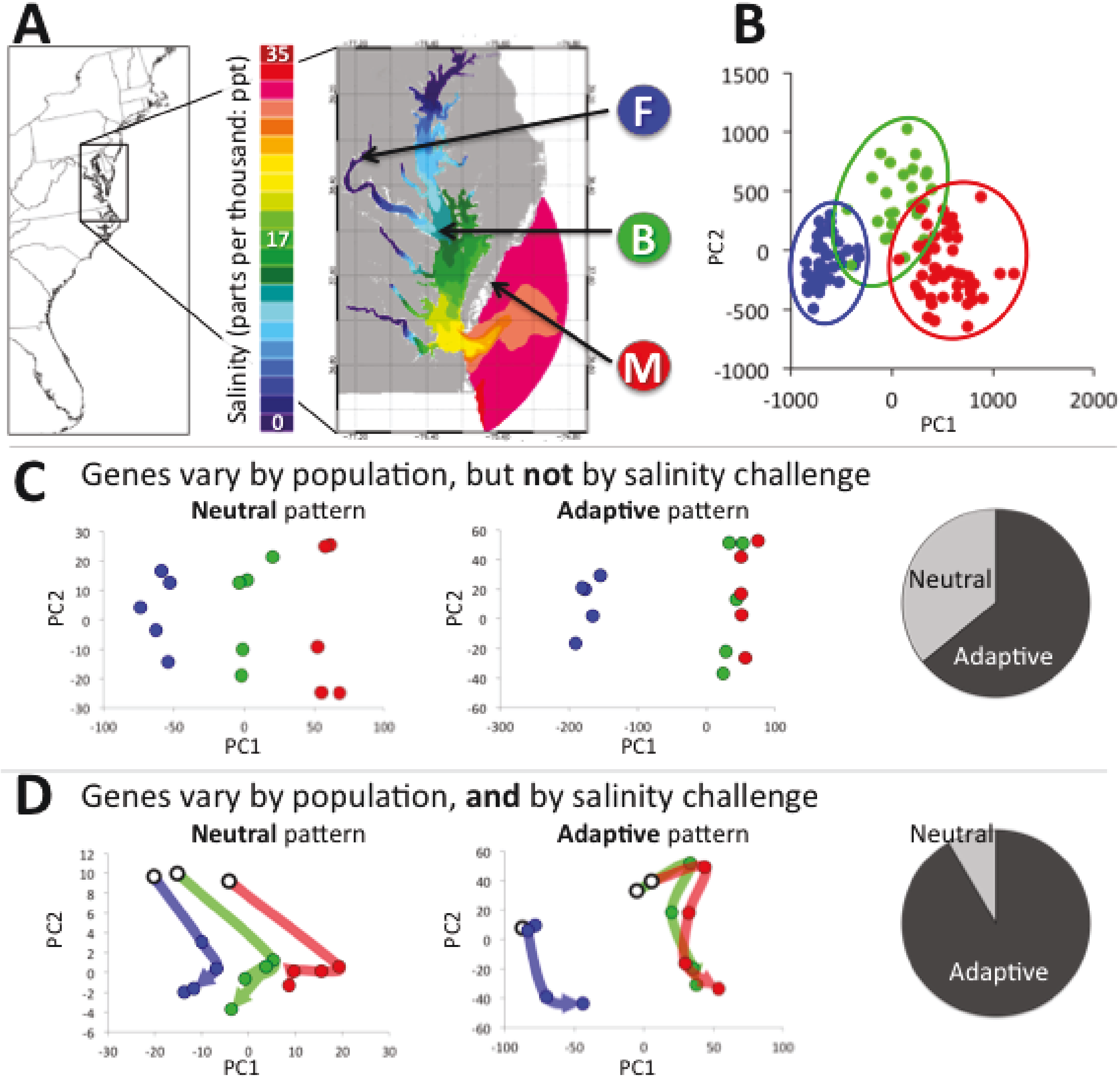
B. Significant and Fold Difference



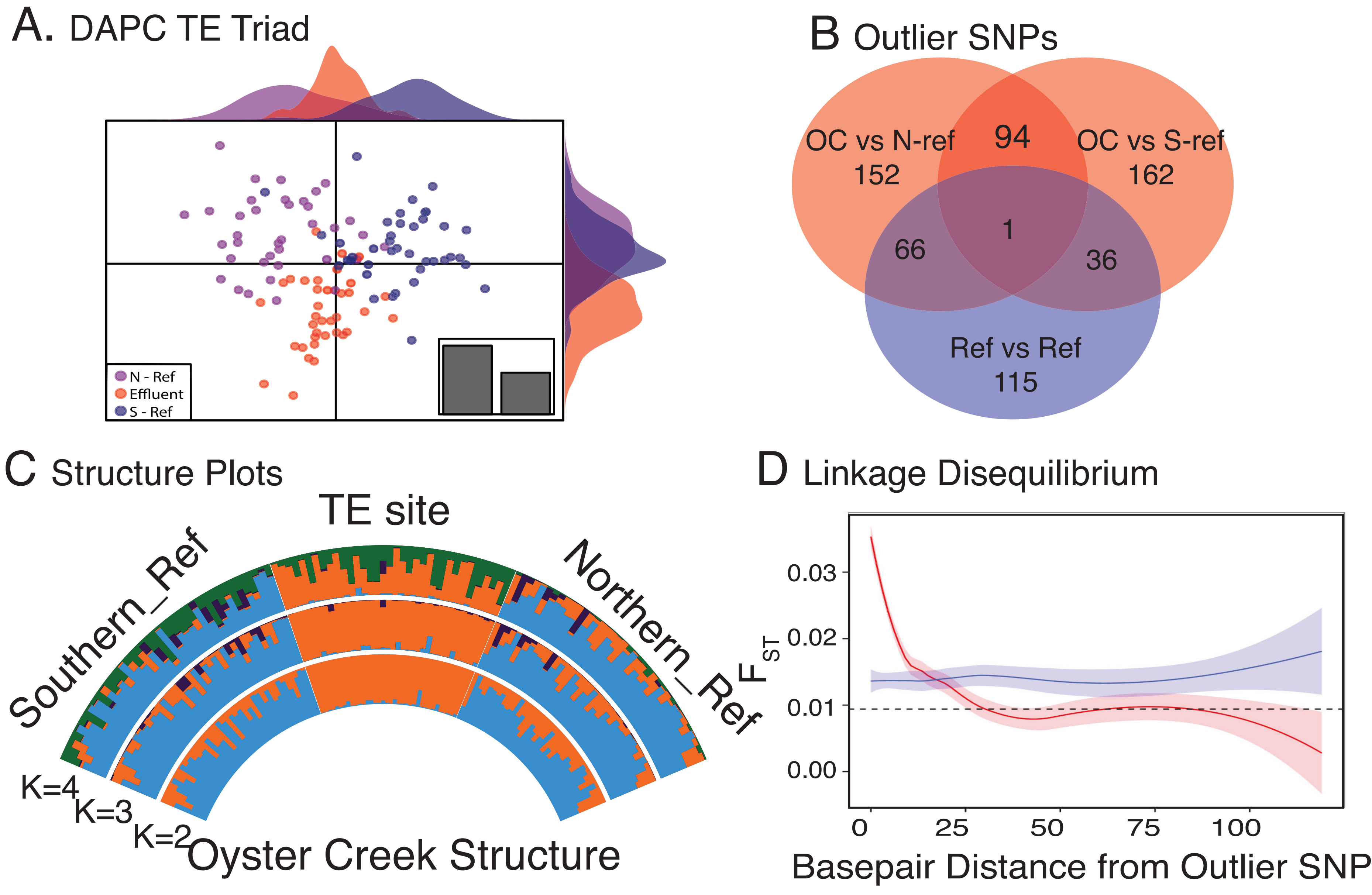


A. Atlantic Coast Sampling Sites
ME
7.8ºC
CT
10.6ºC
NJ
12.8ºC
NC
16.7ºC
GA
19.4ºC
0.01
B. Phylogentic and Temperature Regressions
Phylogentic
(50)
Temperature
(58)
24
20
25
6
7
PGLS
(13)
r2 Phylogeny
r2 Regression
0
0.2
0.4
0.6
0.8
1
C. Variation among and within populations
Among Population Variation
Within Population Variation
2
1
0
-1
-2
-1.5
-0.5
0.5
1.5
Neutral drift
Directional selection
Stabilizing selection
Balancing selection

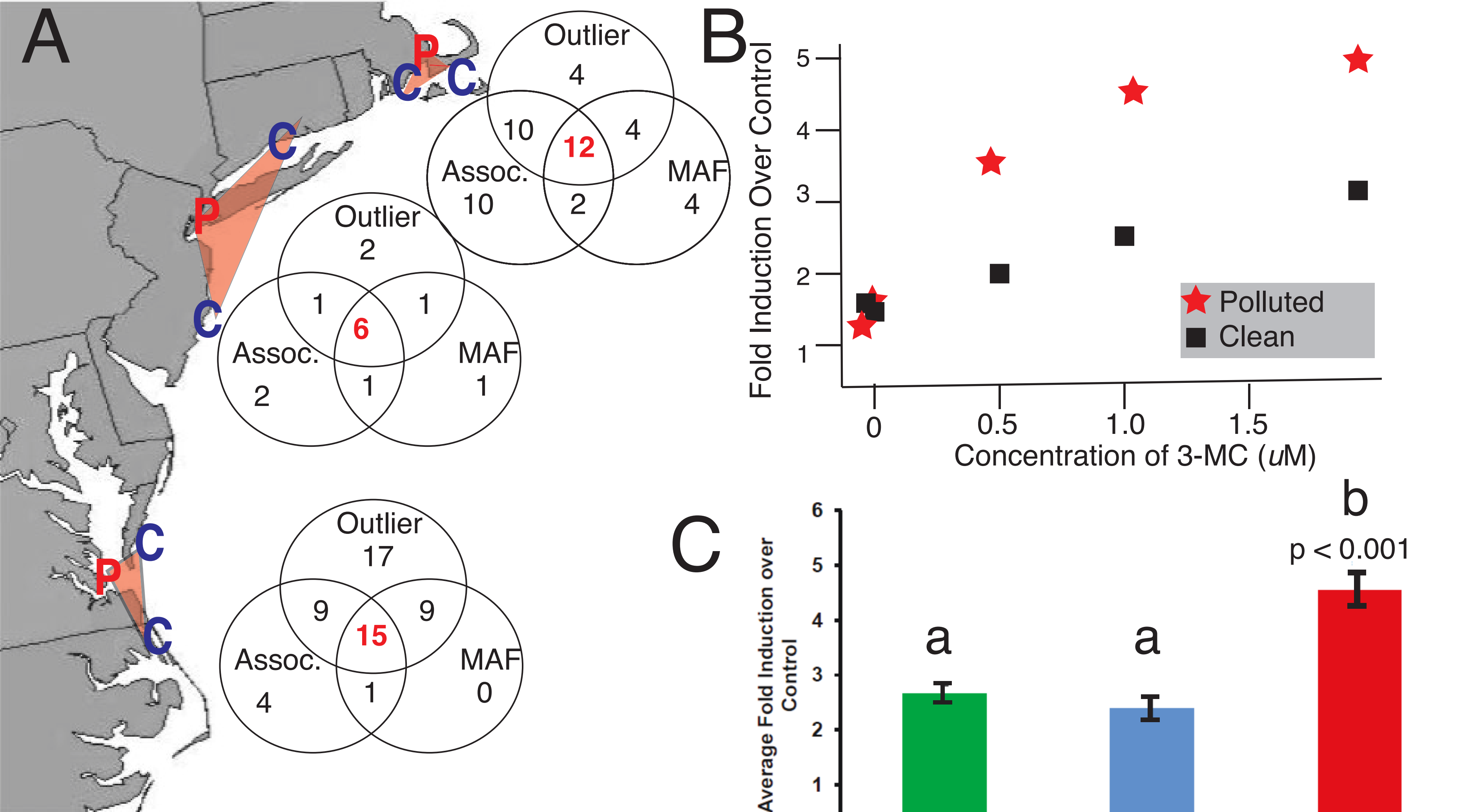

A. Relative Cardiac Metabolism
Fold Change
> 1.75
1.50
1.10
0.00
1.10-1
1.50-1
< 1.75-1
M6 M5 M2 G3 G1 M3 G5 G9 M8 M0 M7 G2 G7 M4 G8 G0
Fatty acid
LKA
Glucose
B. Significance and Magnitude of mRNA Expression
1.5 fold
-Log10 P-value
Northern
Southern
Log 2 difference relative to population mean
C. Groups of Individuals Share Similar Patterns of mRNA Expression
Group 1
Group 2
Group 3
M5 M3 G5 M7 G2 M6 M0 G3 M4 G8 G1 M8 G9 G7 M2 G0
NADH Fe-S 1
aldehyde DH 1
3-hydroxy-3-methylglutaryl-CoA reductase
acetyl-CoA acetyltrans 2
acetyl-CoA syn
Isocitrate DH 3
fatty acid binding 7 brain
vacuolar ATP syn C
PEP carboxykinase
ATPase Cu transporting A
triglyceride lipase
hypoxia inducible factor 3 A
hepatocyte nuclear factor 4A
forkhead box P2
cyto C oxidase VIIIa
NADH Fe-S ox–red 75 KD
ubiqL-cyto C red FeS
acid sphingomyelinase-like phosphodiesterase
NADH flavoprotein 2
ATP syn H+ F1 delta
Asp amino transferse
NADH-Ub ox–red 19 KD
NADH-Ub ox–red B22
NADH-Ub ox–red ASHI
ubiqL-cyto C red 7.2KD
cold inducible RNA-binding
NADH-coenzyme Q reductase 30KD
ubiql-cyto C reductase core I
ATP syn H+ transporting F1 g1
G3APDH
NADH 2
ATP syn B
cyto C oxidase VIIIb
cyto C oxidase VIIA
cyto C oxidase VIA
D. mRNA Expression Predicts Metabolic Rates
Group 1 Individuals
Group 2 Individuals
Group 3 Individuals
Fatty Acid Metabolism
R2 =64% P =0.03
R2 =58% P =0.04
R2 =75% P =0.03
PC1_TCA
PC2_OxP
PC2_Gly



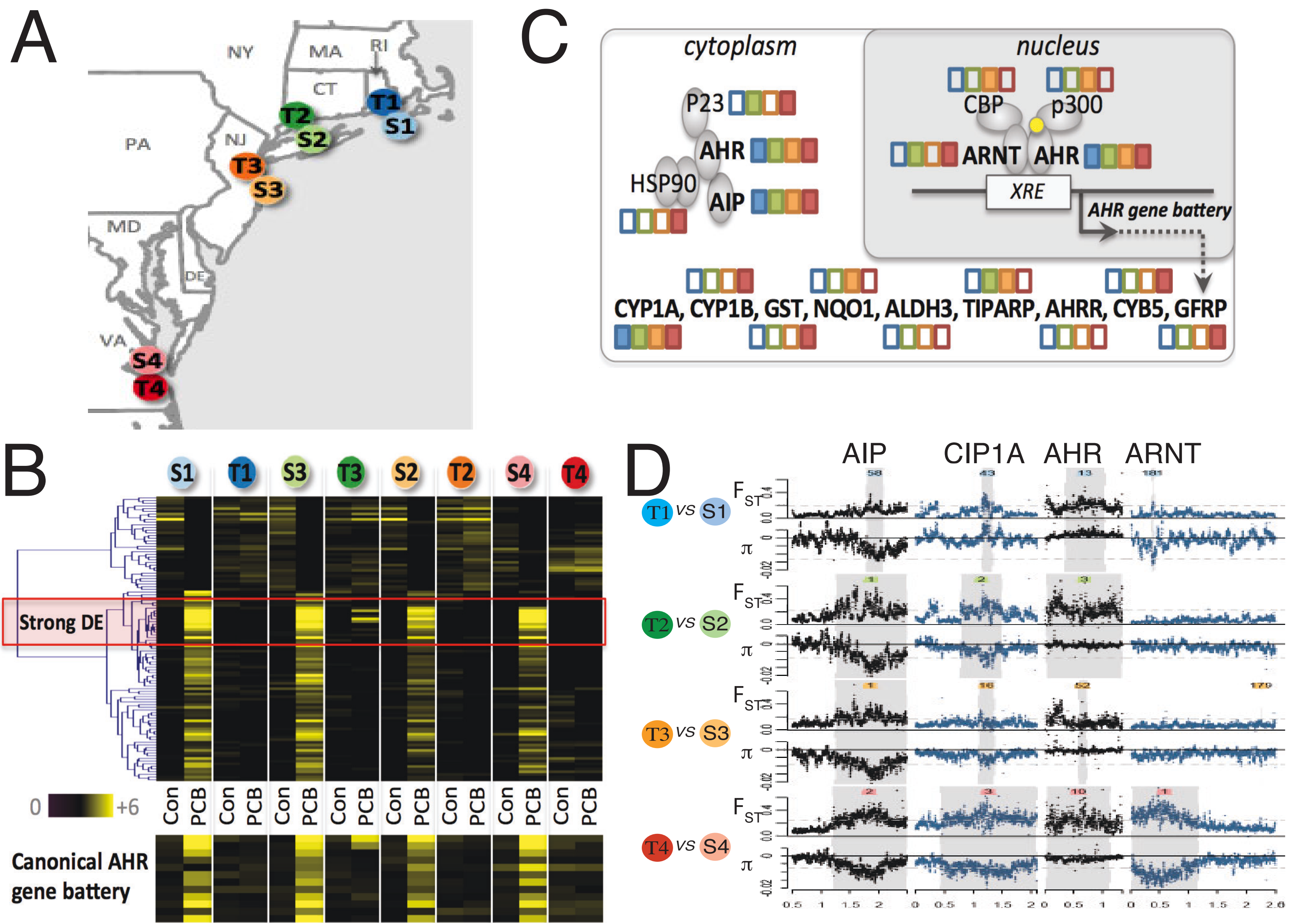
A
35
17
0
Salinity (parts per thousand: ppt)
F
B
M
B
PC2
PC1
C
Genes vary by population, but **not** by salinity challenge
Neutral pattern
Adaptive pattern
Neutral
Adaptive
D
Genes vary by population, **and** by salinity challenge
Neutral pattern
Adaptive pattern
Neutral
Adaptive



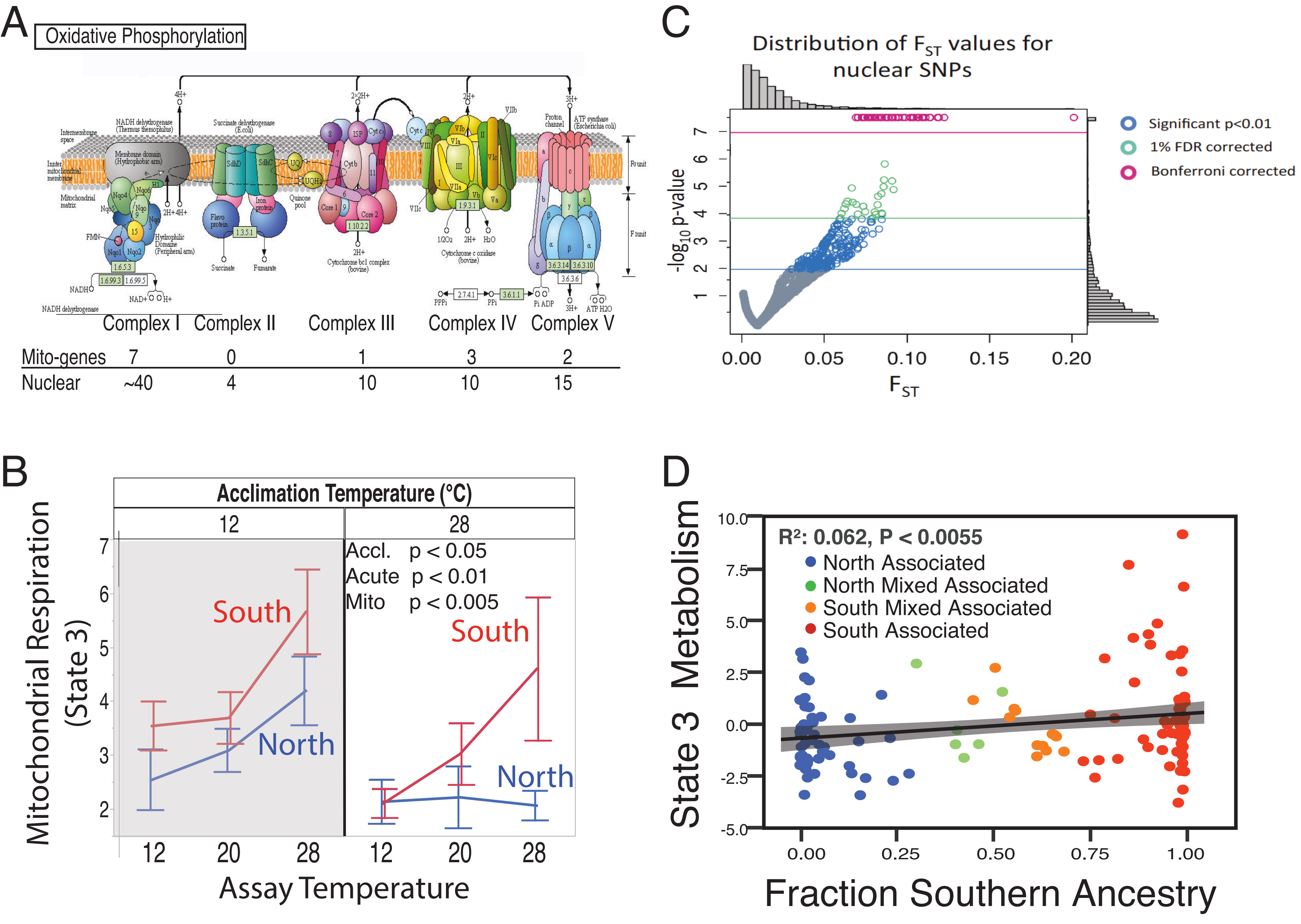
A. DAPC TE Triad
N - Ref
Effluent
S - Ref
B Outlier SNPs
OC vs N-ref
152
94
OC vs S-ref
162
66
1
36
Ref vs Ref
115
C Structure Plots
TE site
Southern_Ref
Northern_Ref
K=4
K=3
K=2
Oyster Creek Structure
D Linkage Disequilibrium
F
ST
0.03
0.02
0.01
0.00
0
25
50
75
100
Basepair Distance from Outlier SNP



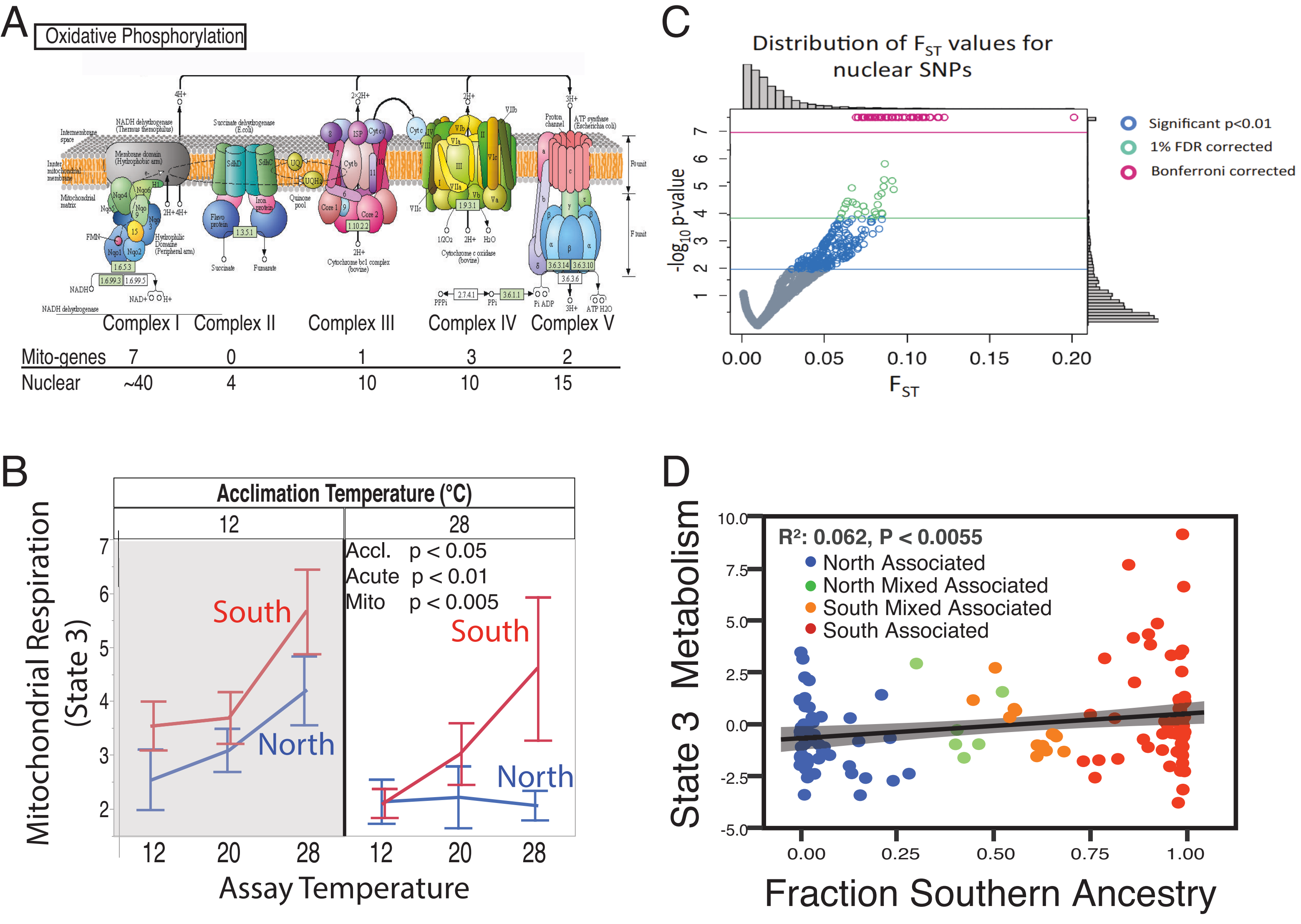
Numbers are the count of SNPs which were signifcant for the each of the three statistical test, and shared among the three test.

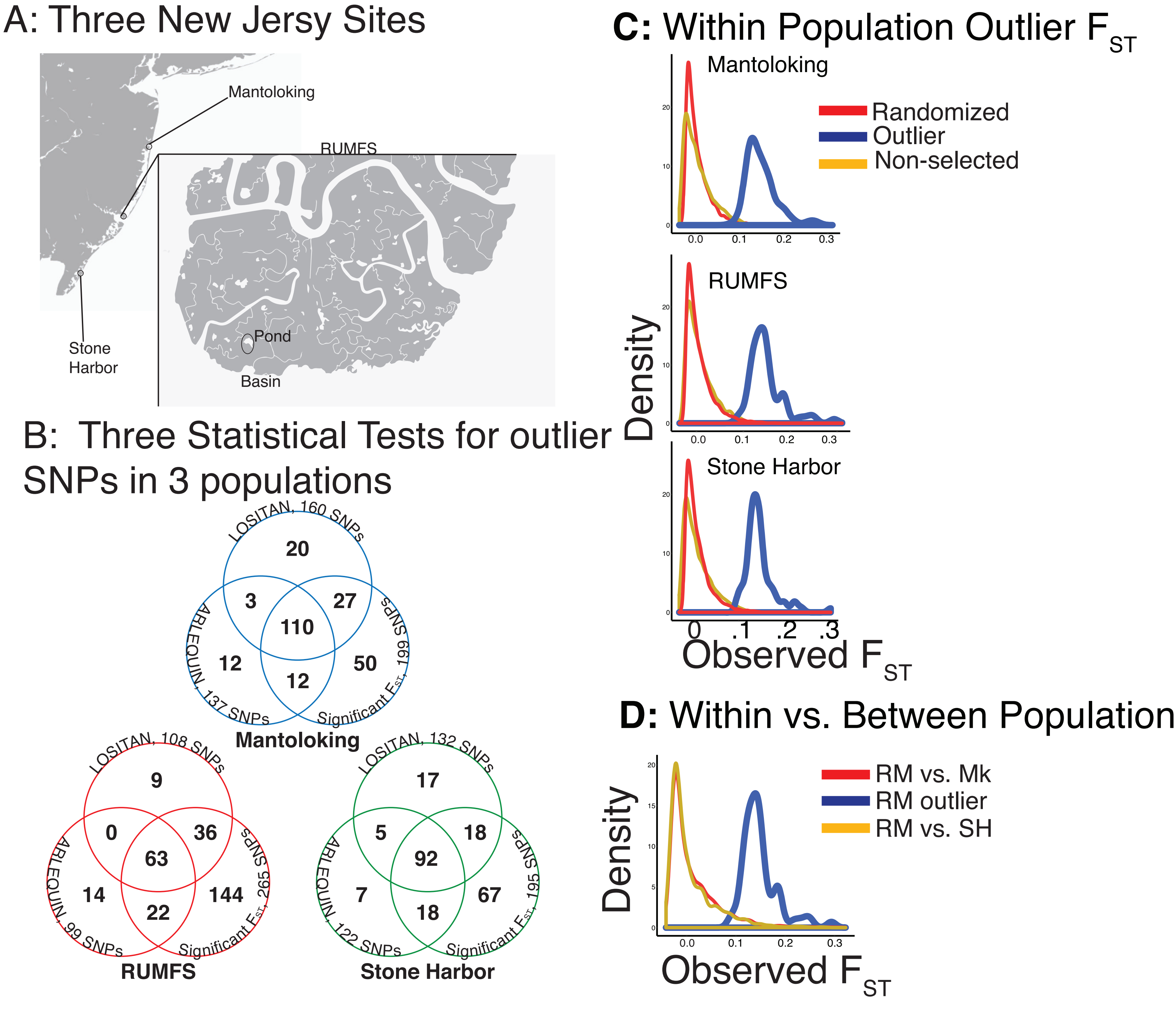


A
NY
MA
RI
CT
T1
S1
T2
S2
PA
NJ
T3
S3
MD
DE
VA
S4
T4
B
S1 T1 S3 T3 S2 T2 S4 T4
Strong DE
0
+6
Con PCB Con PCB Con PCB Con PCB Con PCB Con PCB Con PCB Con PCB
Canonical AHR gene battery
C
cytoplasm
nucleus
P23
AHR
HSP90
AIP
CBP
p300
ARNT
AHR
XRE
AHR gene battery
CYP1A, CYP1B, GST, NQO1, ALDH3, TIPARP, AHRR, CYB5, GFRP
D
AIP
CIP1A
AHR
ARNT
T1 VS S1
T2 VS S2
T3 VS S3
T4 VS S4
F_{ST}
π



A
Oxidative Phosphorylation
Complex I
Complex II
Complex III
Complex IV
Complex V
Mito-genes 7 0 1 3 2
Nuclear ~40 4 10 10 15
C
Distribution of F_{ST} values for nuclear SNPs
Significant p<0.01
1% FDR corrected
Bonferroni corrected
$-\log_{10}$ p-value
F_{ST}
B
Acclimation Temperature (°C)
12
28
Accl. p < 0.05
Acute p < 0.01
Mito p < 0.005
South
North
Mitochondrial Respiration (State 3)
Assay Temperature
D
R²: 0.062, P < 0.0055
North Associated
North Mixed Associated
South Mixed Associated
South Associated
State 3 Metabolism
Fraction Southern Ancestry


A
Oxidative Phosphorylation
Complex I
Complex II
Complex III
Complex IV
Complex V
Mito-genes 7 0 1 3 2
Nuclear ~40 4 10 10 15
B
Acclimation Temperature (°C)
12
28
Accl. p < 0.05
Acute p < 0.01
Mito p < 0.005
South
North
South
North
Mitochondrial Respiration (State 3)
Assay Temperature
C
Distribution of F_{ST} values for nuclear SNPs
Significant p<0.01
1% FDR corrected
Bonferroni corrected
$-\log_{10}$ p-value
F_{ST}
D
R²: 0.062, P < 0.0055
North Associated
North Mixed Associated
South Mixed Associated
South Associated
State 3 Metabolism
Fraction Southern Ancestry



A: Three New Jersey Sites
Mantoloking
RUMFS
Stone Harbor
Pond
Basin
B: Three Statistical Tests for outlier SNPs in 3 populations
LOSITAN, 160 SNPs
ARLEQUIN, 137 SNPs
Significant FST, 199 SNPs
20
3
27
110
12
12
50
Mantoloking
LOSITAN, 108 SNPs
ARLEQUIN, 99 SNPs
Significant FST, 265 SNPs
9
0
36
63
14
22
144
RUMFS
LOSITAN, 132 SNPs
ARLEQUIN, 122 SNPs
Significant FST, 195 SNPs
17
5
18
92
7
18
67
Stone Harbor
C: Within Population Outlier FST
Mantoloking
Randomized
Outlier
Non-selected
RUMFS
Stone Harbor
Density
Observed FST
D: Within vs. Between Population
RM vs. Mk
RM outlier
RM vs. SH
Density
Observed FST