



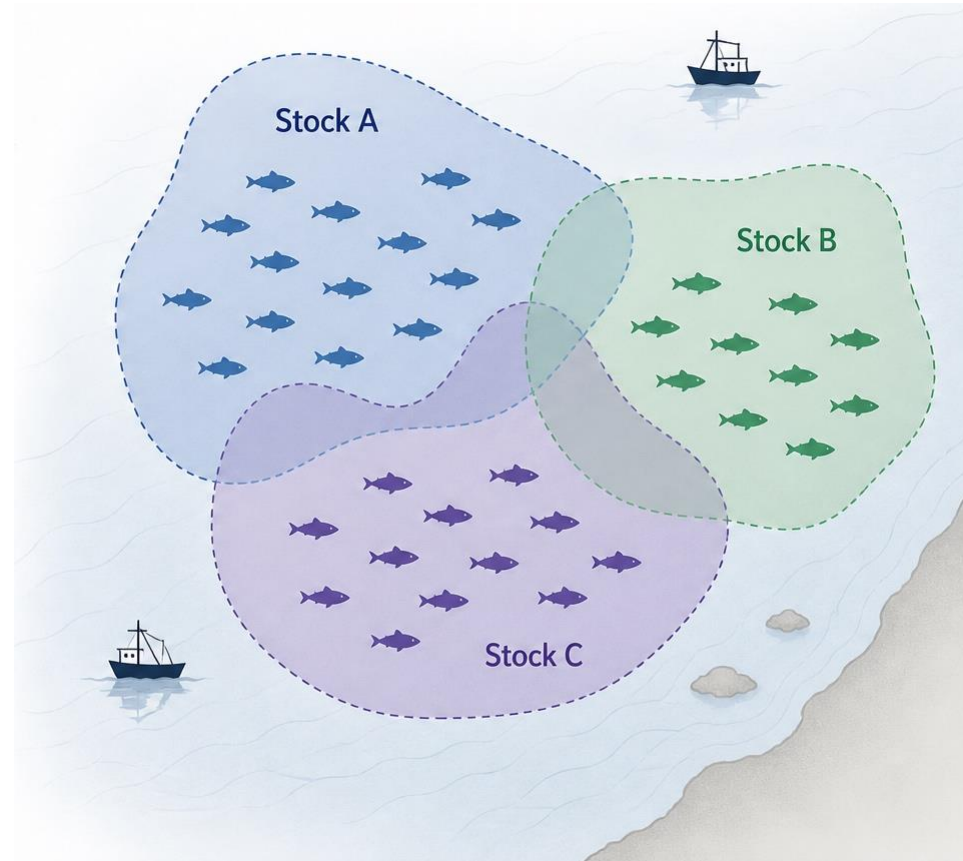
Hidden Diversity, Moving Stocks: Genomic Insights on Connectivity and Adaptation in Small Pelagics

Nina Overgaard Therkildsen

Department of Natural Resources and the Environment, Cornell University

Stock structure underpins key decisions

- Defines management units
- Determines how catch is allocated
- Shapes responses to environmental change



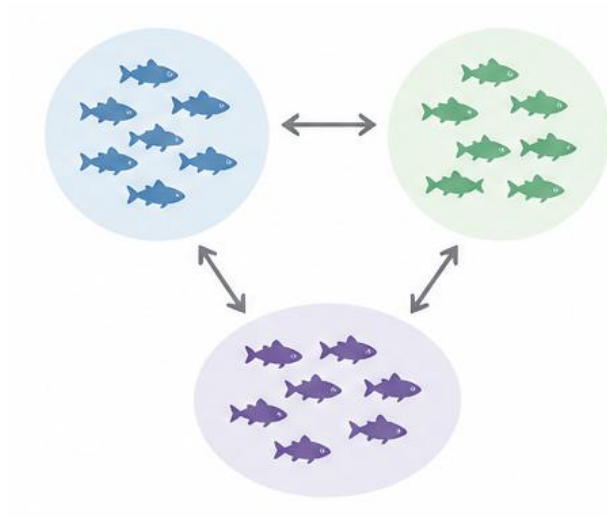
Key processes are hard to observe directly

Especially in small pelagics

- Highly mobile species
- Variable recruitment
- Shifting distributions

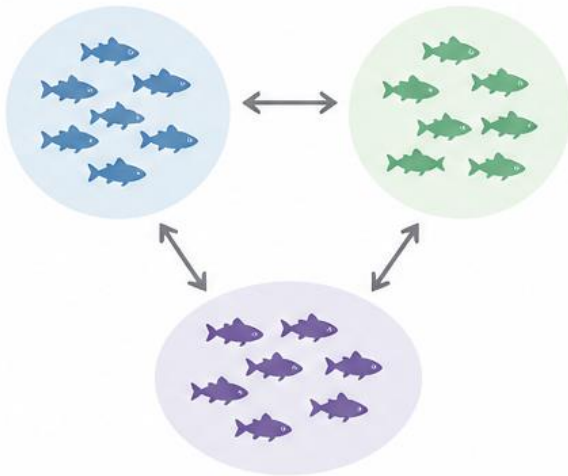
What we need to assess

**Who interbreeds
with whom**
(stock structure)

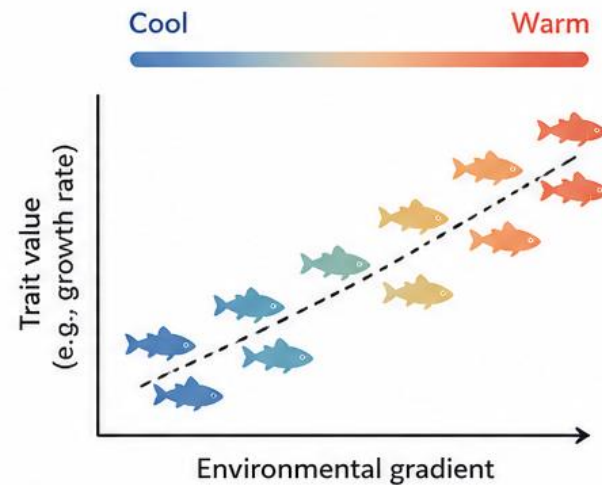


What we need to assess

**Who interbreeds
with whom**
(stock structure)

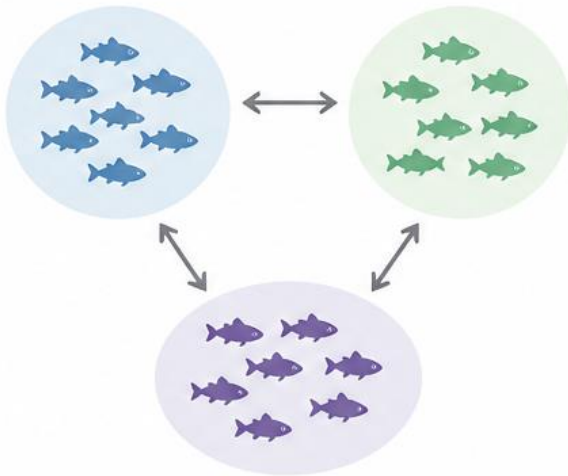


**How environmental
responses vary**
(adaptation)

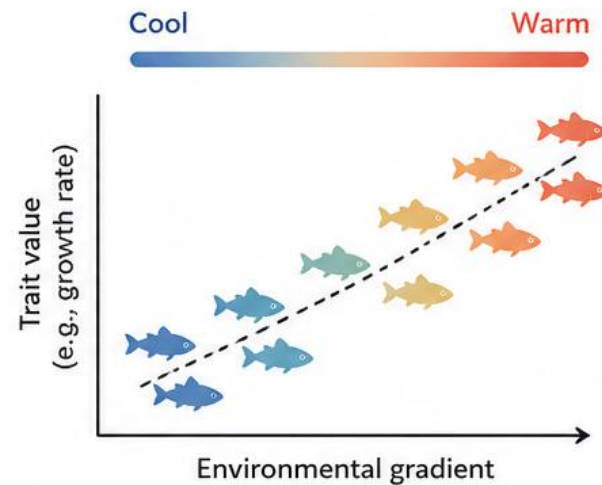


What we need to assess

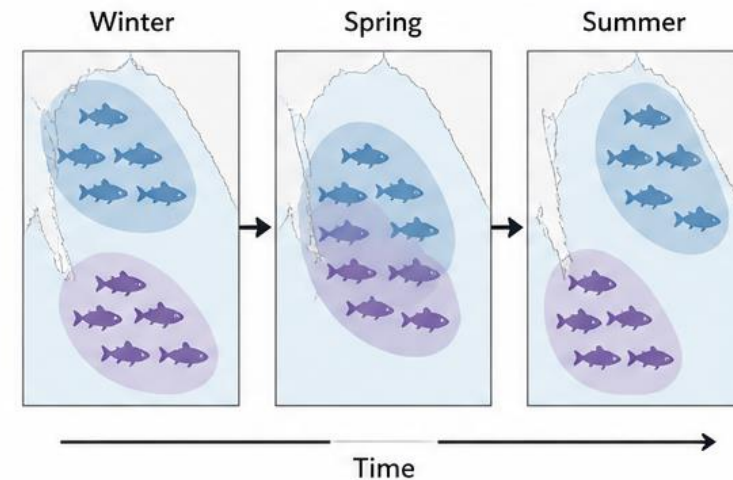
**Who interbreeds
with whom**
(stock structure)



**How environmental
responses vary**
(adaptation)

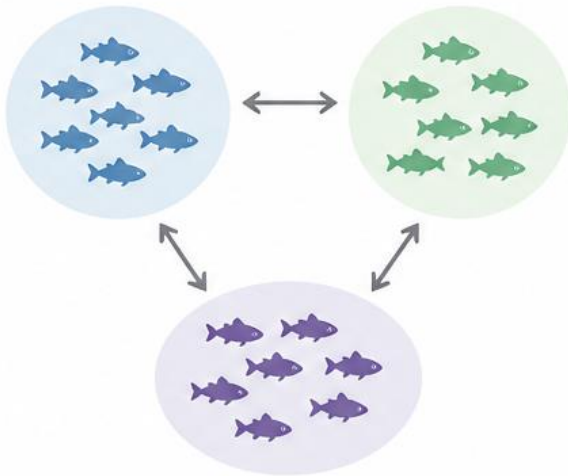


**Where different stocks occur
through time and space**
(mixing dynamics)

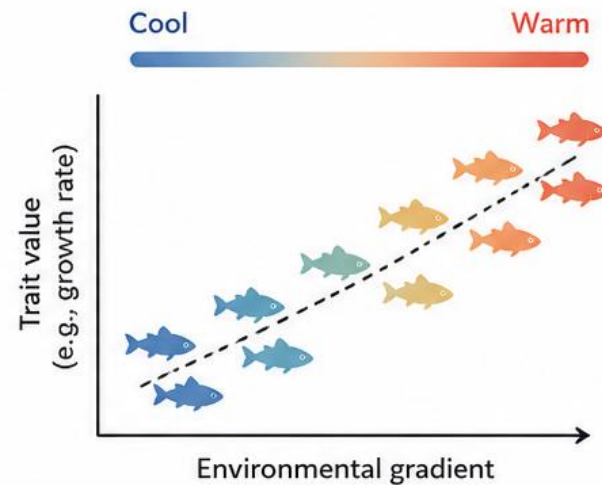


What we need to assess

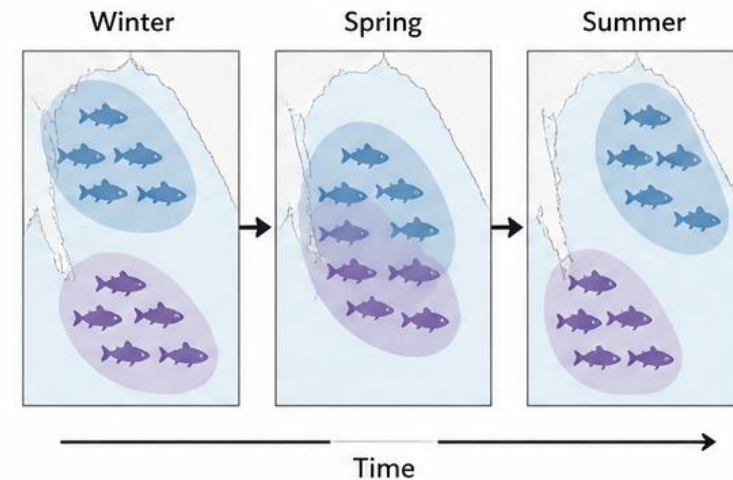
**Who interbreeds
with whom**
(stock structure)



**How environmental
responses vary**
(adaptation)



**Where different stocks occur
through time and space**
(mixing dynamics)

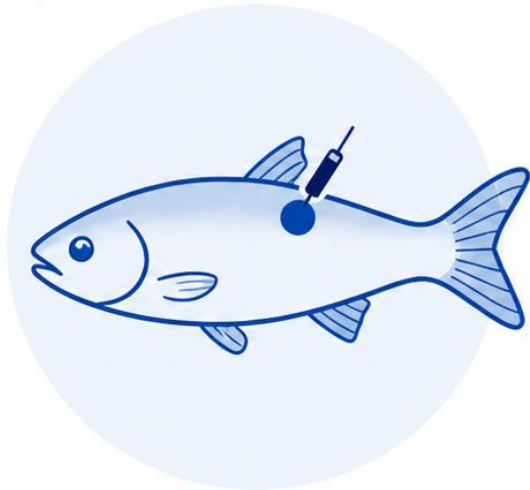


These underlying processes are essential for effective fisheries management

Multiple complementary tools

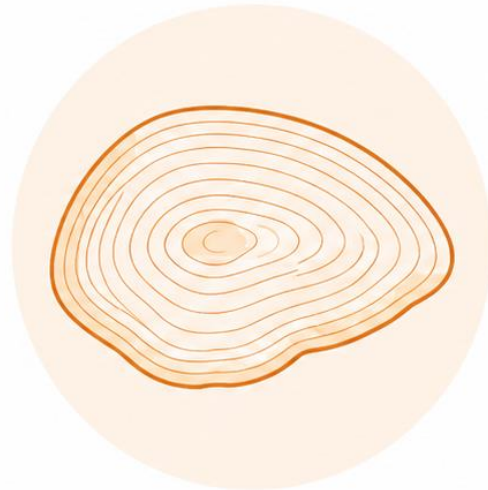
Provide complementary views of connectivity and life history

Tagging



Tracks movement and connectivity of individuals

Otolith analysis



Reconstructs environmental history and life history

Genetics



Reveals genetic variation and population structure



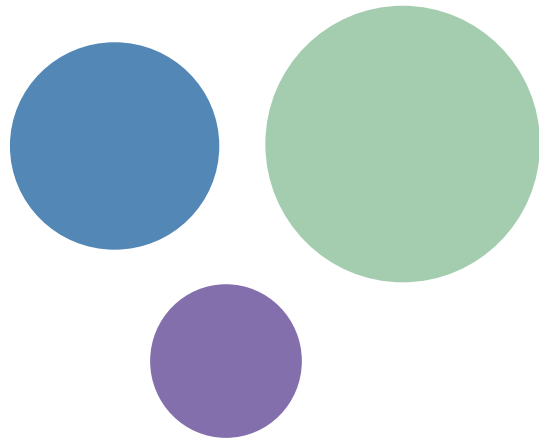
- 
- A satellite-style map of the Great Lakes region in North America, showing the five large lakes (Superior, Michigan, Huron, Erie, and Ontario) and the surrounding landmasses. The water bodies are highlighted in a light blue color. A semi-transparent dark green rectangular box is overlaid on the lower-left portion of the map, containing a bulleted list of three points.
- Small population sizes
 - Low dispersal/gene flow
 - Strong differentiation



- 
- Large population sizes
 - High dispersal/gene flow
 - Weak differentiation

Genomics adds the most value where signal is weakest

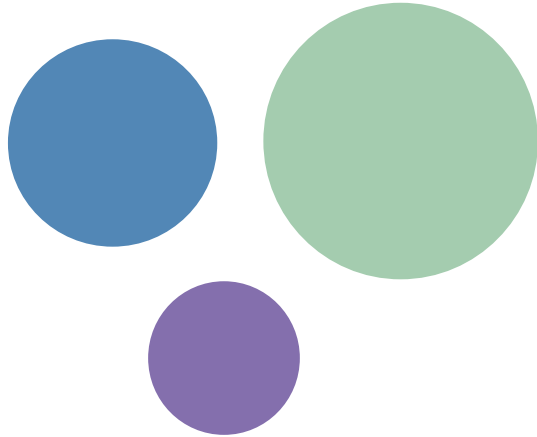
Small populations - low gene flow



Structure easy to detect

Genomics adds the most value where signal is weakest

Small populations - low gene flow



Structure easy to detect

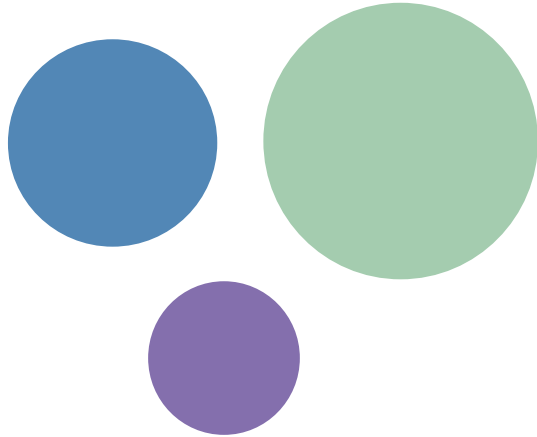
Large populations - high gene flow



Traditional markers often show “no structure”

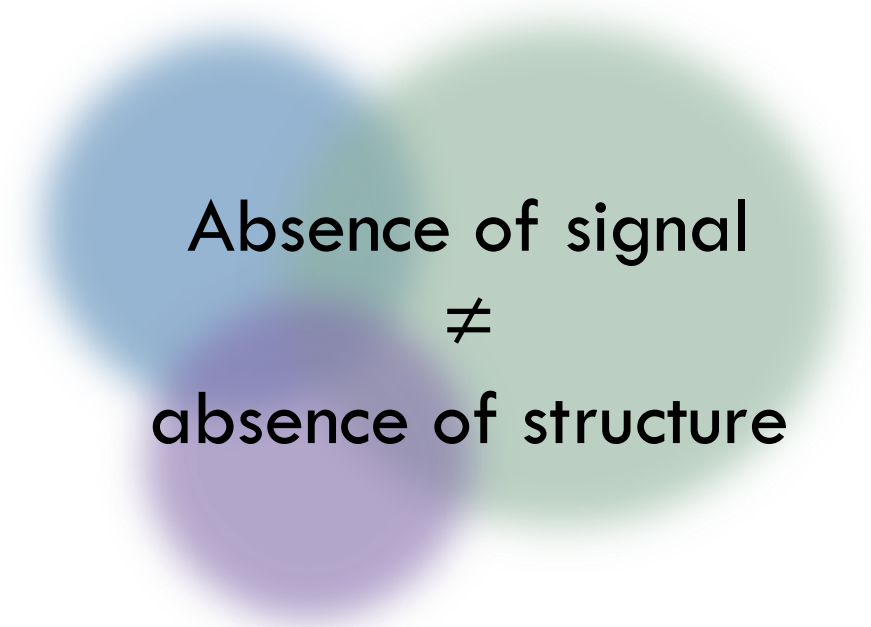
Genomics adds the most value where signal is weakest

Small populations - low gene flow



Structure easy to detect

Large populations - high gene flow

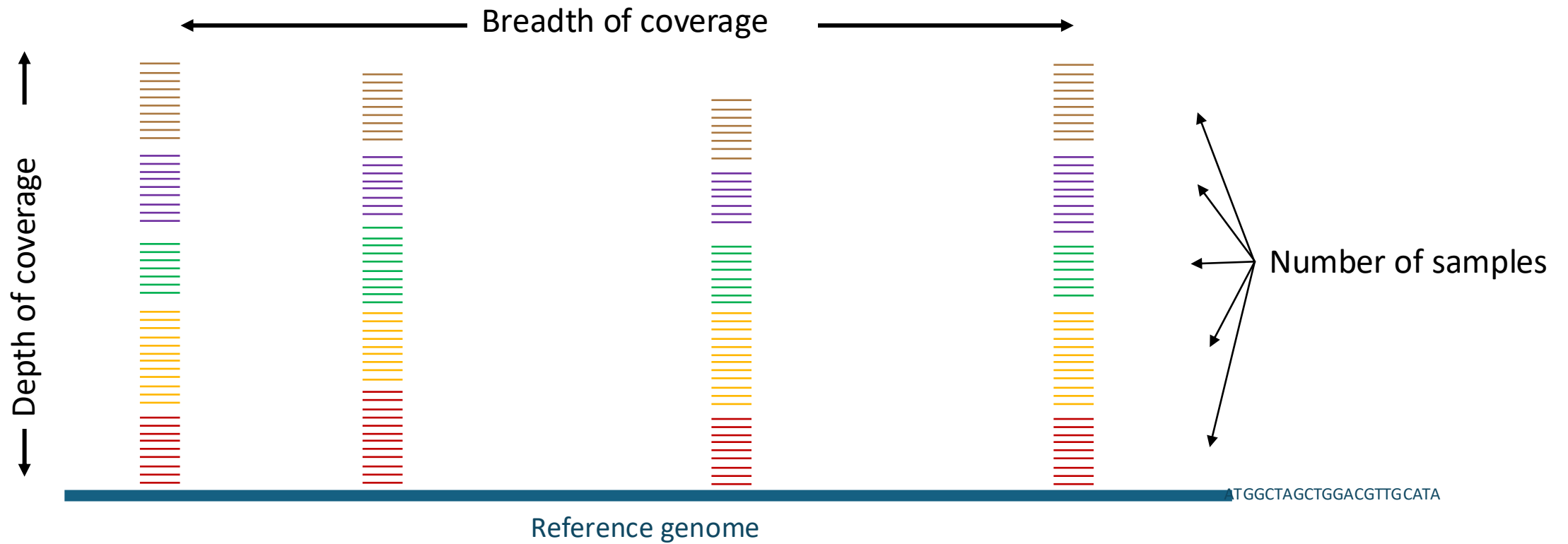


Absence of signal
 \neq
absence of structure

Traditional markers often show “no structure”

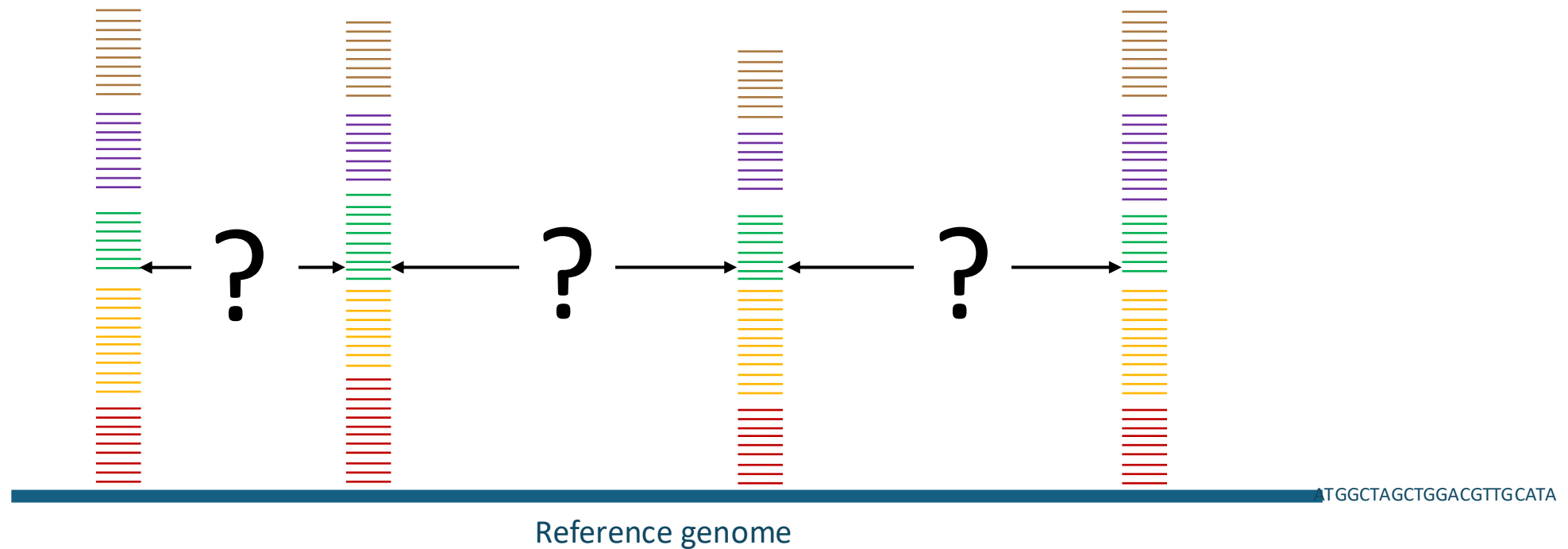
GTTCGCCAGATGATTTTATCAATTTATGCCACATGTGTTGATCTGTATCCTTGCAGAAAAGCAAACAGCAGAAGTGGGGTCAGGAACAGATAGTTCGGGGGAACAATAAGAC
ACTTCTGAGCCTAAGTGAATCCAGAGTTGCATCTTTGATTCTGACCAGCACATTTTAACTTTTCTGAACATCTTCTTTAAAGAAGTTTCATTGTTTTCCAAGACAAATGATGCAAATAA
ACTGTATCAAGGTTTTTTATTTGATCCATGAGACAACCTGTACAGAATTAGACAAGATTGTGTGCAAATAATCAAACCTGACATAGGCATTGATAATAATGACACAGTAATCCATAAGAT
TGTTTTACTTCTG GATTTATTAGCCTCAGATTGAATTAGGCTCTATAAGCCTGAGCATCAGTACATTTCTTT CTTTTATATATGAAAACAACCAAAGCTTGAGGTGGCATTTCATGACAGG
CCATTTATAAAACAACCCATTTCGGCATGTTTAAGGATGCCTTCAACAACATTCAAATAGGTATCAGCAGATCCTGACAATCTTAAAGGGGAAGTTCTCCCAAATATCCACAATTTAGGG
CCAGTGGATTTTTATTGGTCATTACTCATACCATTTATGTTCCACCTCCAGCAGACAGGGATGTCGTCTGCTGGAGGTGGTAACCCAGCGATTACCAACTACCAGGCTAACCACCAG
GATGCTGCCCTCATTGTGGCGTGAGATTTGCCAGCACAAACCTGAGGAAAGCGAGAATGAAGTTCAAGCAACACATCACCTGCCCCACTTTATGAAAGAGGACCCTAGAACACT
GCTCCTCTCTTCTCAATGGCAGCTACAAAGCCACACCTAAGCCCCATTGGCTGATCCGGACATACCACCATCCATCCATTATCAATACCGCTTATTCCCTGATGGGGTCGCGGGG
GAACTGAAGCCTATCCCAGTGGTCAACGGGCAAGAGGCAAGGTAAACCCTGGACAGTATCCGTCCATACCACCATGCTTCTTATTCCAAAACATAAAAAACAGCTCAGGCAGGCA
GCGCCAGCTGTCAAAGAATGACAC A GTGGTCCAGACCATCAGAGGATGCTCTGGTTCAAGGGCTGCGTTGCAAATTAGCCGACTTGACAAGCTCTGGACAGCCCCTAGATCCT
CACTACTGCTTACGACACAGGGCTTTTGACCAGGGATGTGAAGGCATACAAGGCTGCCCTCTGATTGGAGGGTCAGTGGTTCAATTCATGTTCCAGGCCACATGCGGAAGAGT
TTTTGGCAAGGCTCTGAACCCCAAGTTGCCTACGGGTCTATCTGTTGCGCCTACT TGGTGCATGACCTGTAGTAGAAAGCGCTTTGAGTGGTCAGTTTGACTGGAAAAGCACTATG
AAGTGAACGTAGATTCACATTTACATCTCCTAGGCTGCTTCCTAGAGCCTCTGCAGAGTAGTAAAGGAGGCTAAACAGTGCTATGGCATGAATACTGGAGTCACAGTTACAGATCT
TGAGTCCGAGGGACTGTGGCAGGAACTTTGTACCACTACGGTCTATGAAACGAAAAGTGTGTGTGCCATTTATAGTCCTTAATGGTGCATGCTGTTGATTCTCCGGCTGACAGTT
AGATACTTTTTTGCCTGCTTCGAGTCTGACAACAATCAAGCTGCTGTATTTCCAGTTAGAGAGGATAAAGTTGTGACTTGTAGACACACTACGGTATCAAAGCAGGTCATGTTTGGTCT
GGTGGATAAAAAAACTCCAGGATTGCCACAAAATGATATAGTTTTCTTTTATCTTTGATATATTTGAACAATAACAAAGAAAAAAACATATATTCCAACATTGCACCCAGCTATTACAAG
CATATCTTACTGTTTATCGGCAAAAAATAAATAAAATTACATCACACAACCTTATGAACGGCTAACGTTGAGTCTGCTGTTTTATAAAAGGTGAGTCTCAGTCCTATGACTTTTTGTTGCTCA
GATAAGGTGGGATCAAACCTGGCACGATATGTGTCTTGATTCTTTATTGTATTTTT GATTAAAAAAACACAACAACATGGGTTATAACTGTTACAGTGTCAGAAGCTGACCGAGAGTC
TTTCATCACTATTGTTTCATGTAGTCAGAGTCATTTAGGCCTATGAGCTGCGAAATGTGGGAGGAGTGGAAGGAGTTAAAGC CTGCCCTGAGCAACTGGCATCTGTGCTTACCTCA
TATTAACCTCTCCCTCGCTCGATGTGTAATTTGCAACTGTTTTAAAAAGGCTGTGAAGAAAAACCCAGCCCAGCTGTCCAATGACAACCGCCCCATAGCCCTCAGTTTCATGGTGATTA
AATGTTTTGAGAGACAATTTATAACCTTCATCACCTCCTCCCTGCCTGCTACCCTGCACCGGCTACACTTTTCGTACCACCCAGGAAGATAAACAGATGACACCACAGCCCTCACGCTC
ACAAAACCTCTTTTCTACGTGAACAGTAGAAGGATGTACAATGAATGACTGCTCTTCATGGATTACAGCTAAGCATTCAACGCTAAAGTGCCCTCAAGGCTGGTCACCATTGTGTGGTA
CTGAAGTGAACAAACTGGTAGGTGCCAGAAGAATATGTCTGCTGAATGTCAATAAGCAGAGCATCTTTCATACACTATCAGTGGGTGATAAGCACAGTTTTAAATATCTTCGTGTGA
CAATCACGACCCATCTTGGTCACTTCACATTAATCCACTGTGAGGAAAGCAGTTTTAAACCTGCACCACTGACAGCATCCTGTGTGGGAGAATCACCACCTGGATGGGCACCTGTA
CGAGCAGGACTTCTCGGCCTTCAAAGGGGTGGTCTGTTTCAGCT

RAD-seq approach or SNP chips

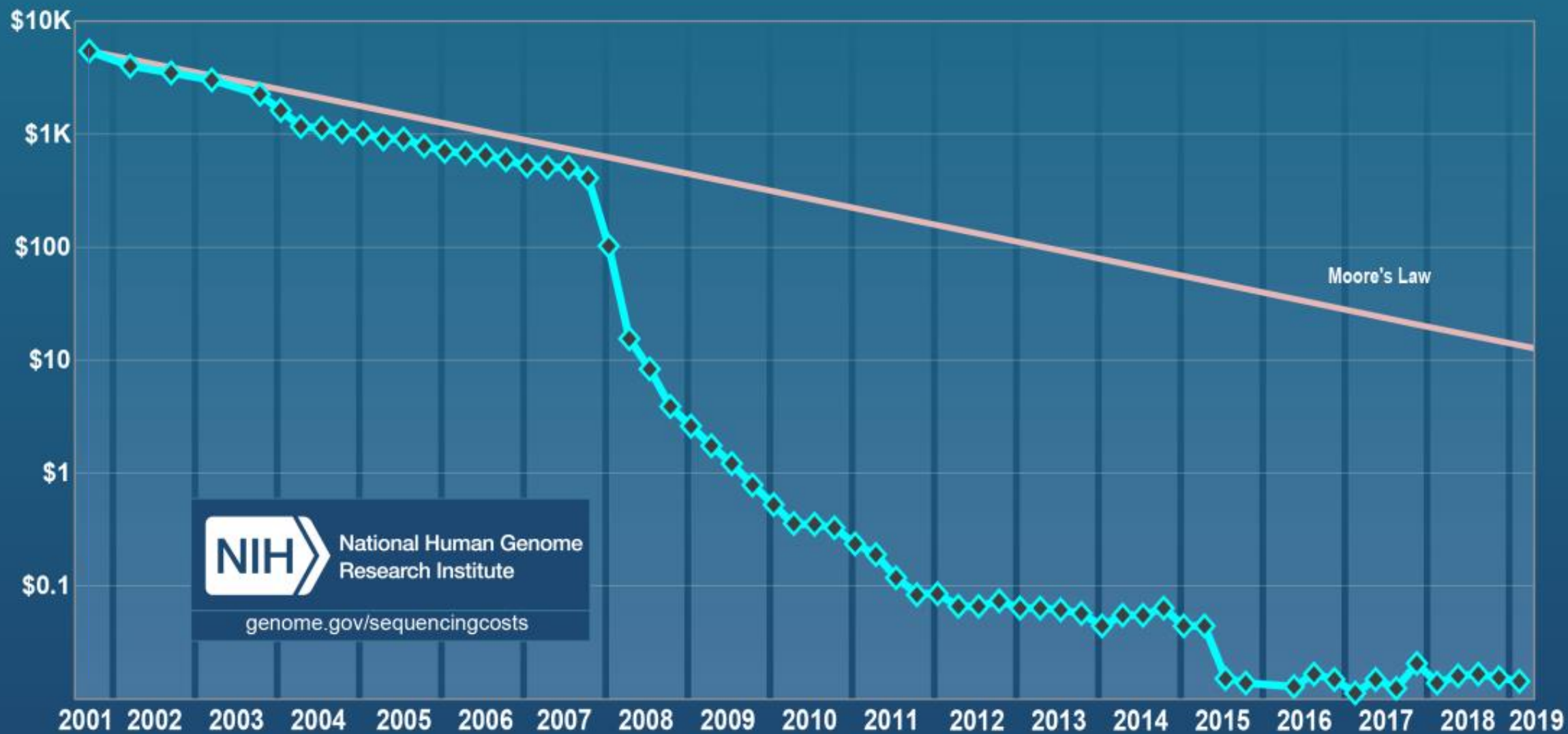


RAD-seq approach or SNP chip

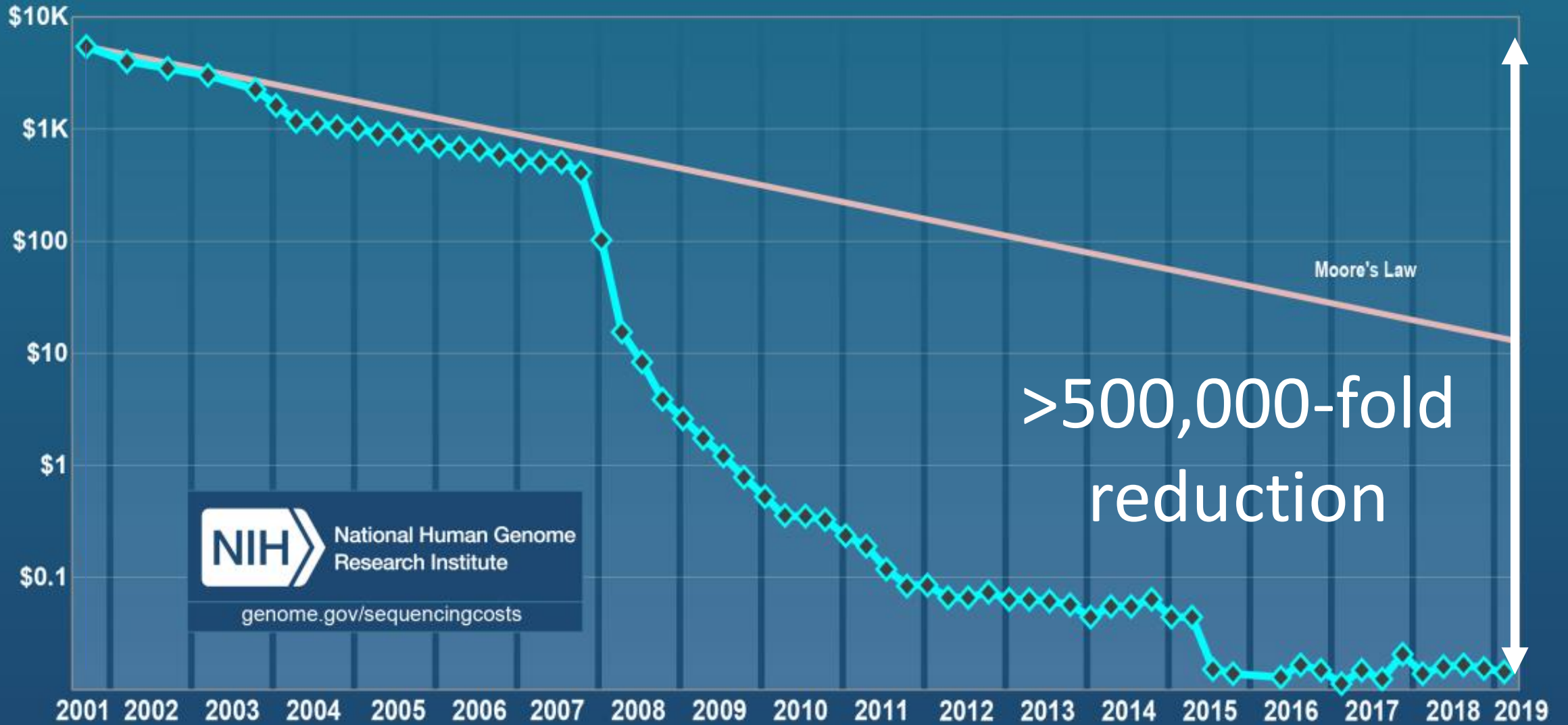
What are we missing?



Cost per Raw Megabase of DNA Sequence



Cost per Raw Megabase of DNA Sequence



From one genome to thousands

ORDERS-OF-MAGNITUDE SHIFT IN SCALE

THEN

Human Genome Project
(2001–2003)



1 genome



~\$3 billion



>10 years

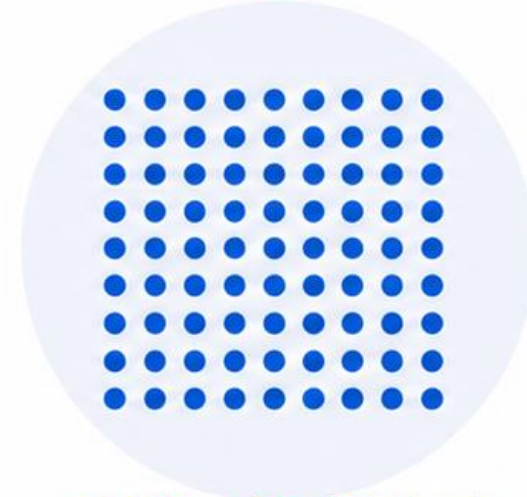


Global effort (2,800 researchers)



NOW

Today



1000s of individuals



~\$100–\$1000



Days–weeks



Routine effort

From one genome to thousands

ORDERS-OF-MAGNITUDE SHIFT IN SCALE

THEN

Human Genome Project
(2001–2003)



1 genome



~\$3 billion



>10 years



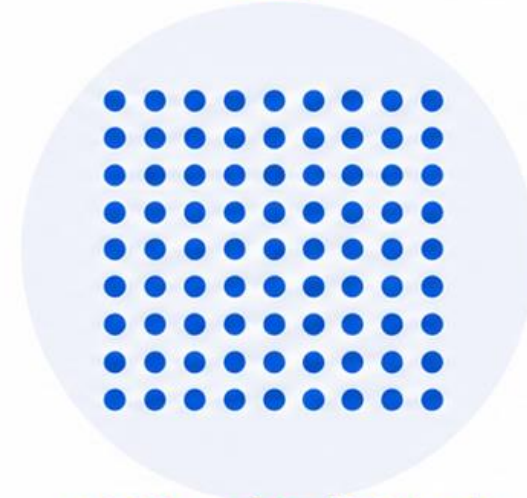
Global effort (2,800 researchers)



*The revolution is not
just cost—it's scale*

NOW

Today



1000s of individuals



~\$100–\$1000

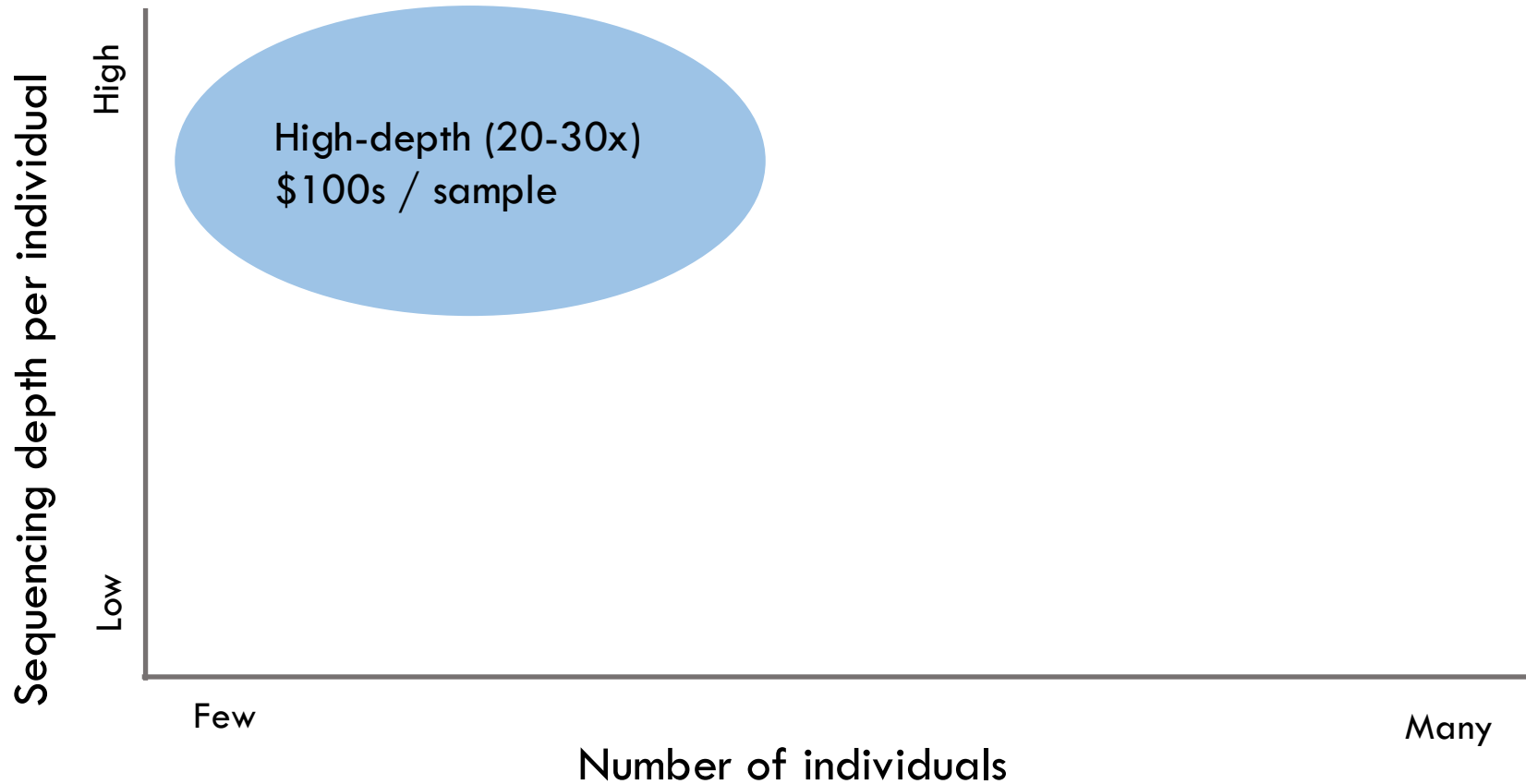


Days–weeks

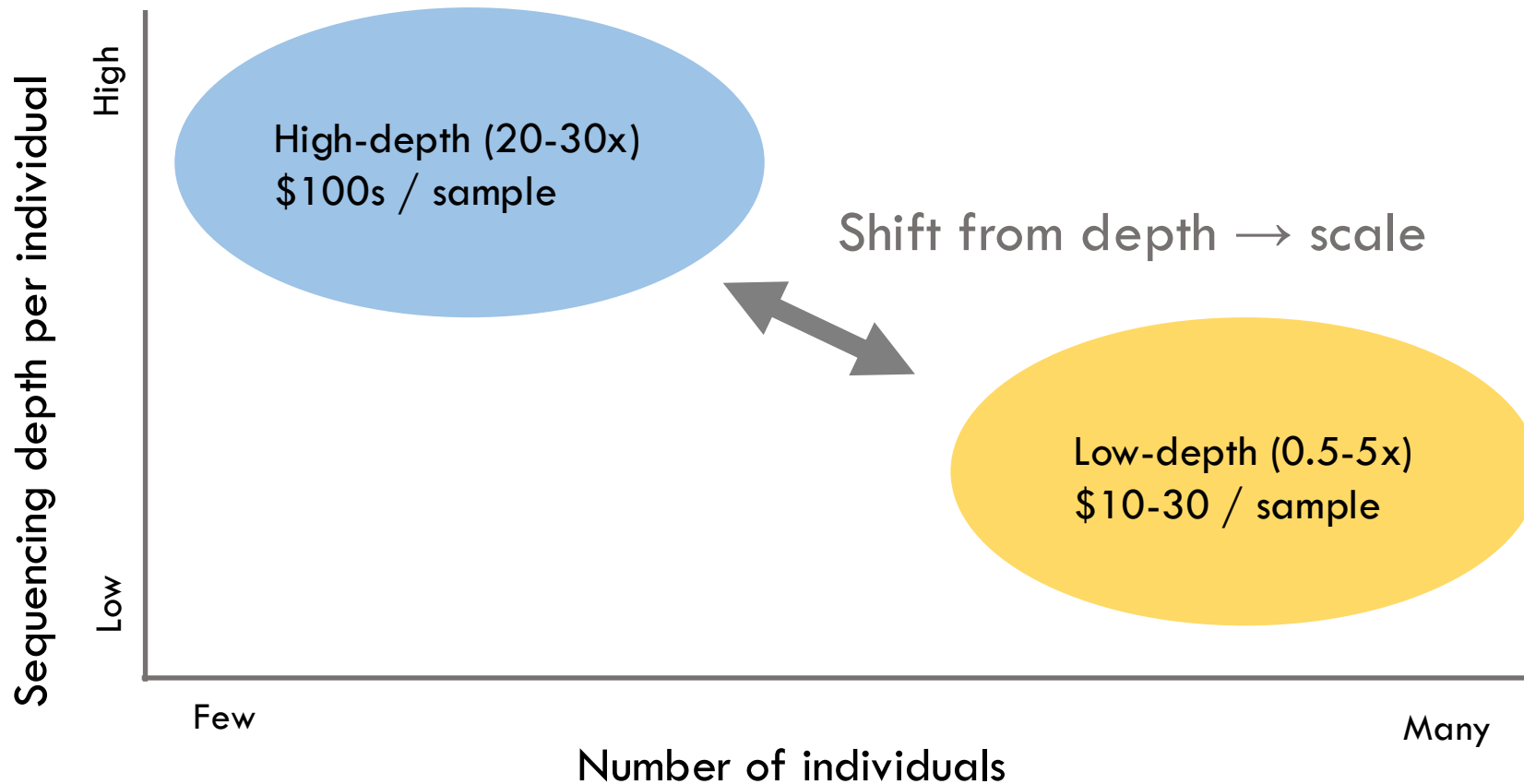


Routine effort

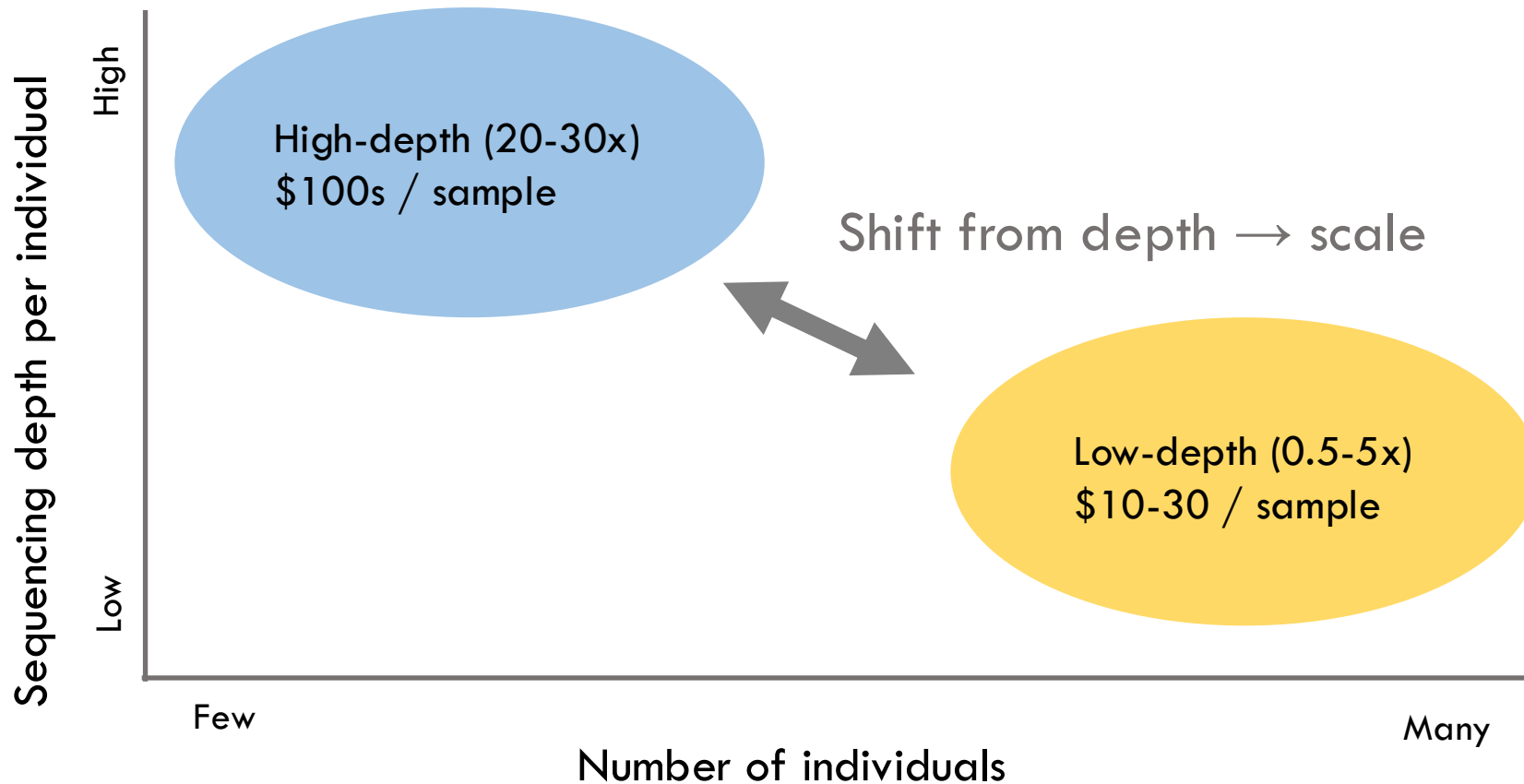
Sequencing depth vs number of individuals



Sequencing depth vs number of individuals



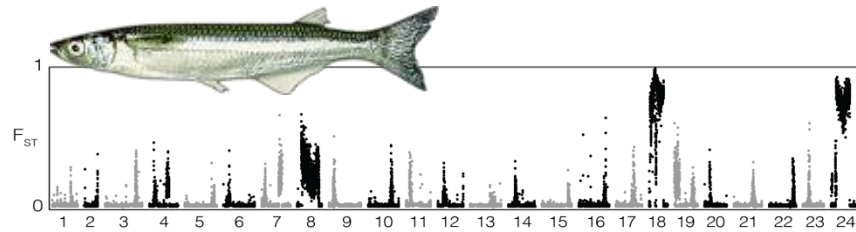
Sequencing depth vs number of individuals



Power comes from sampling more individuals
- not sequencing each one more deeply

Structure can be hidden in different ways

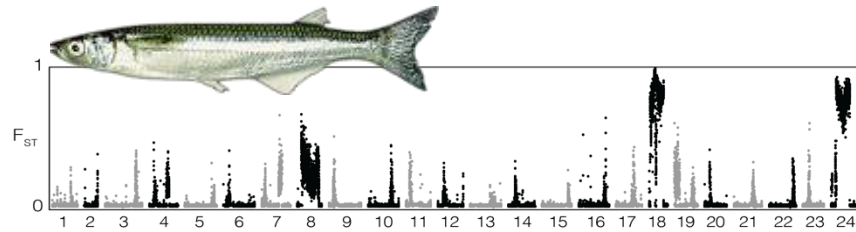
Strong adaptation despite homogeneous genomes



Atlantic silversides

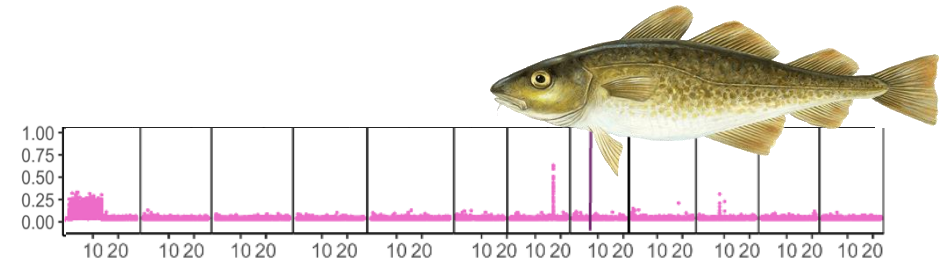
Structure can be hidden in different ways

Strong adaptation despite homogeneous genomes



Atlantic silversides

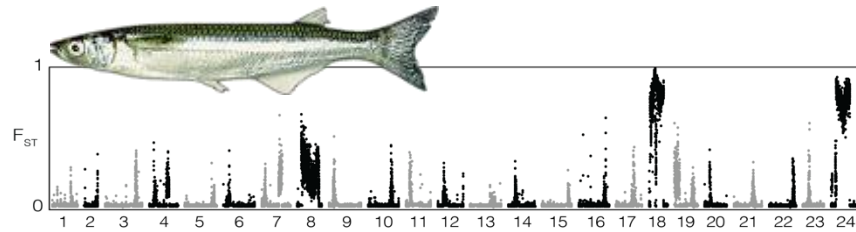
Differences confined to small genomic regions



Atlantic cod

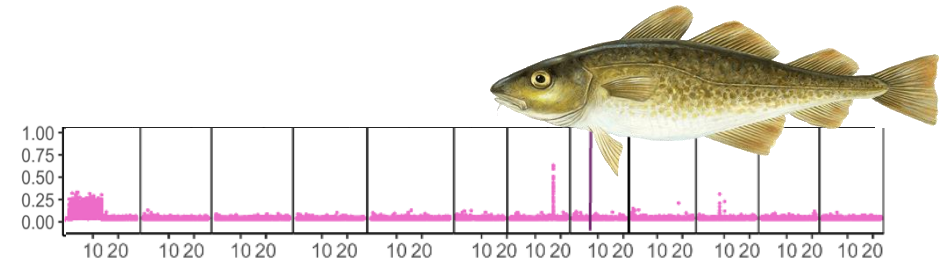
Structure can be hidden in different ways

Strong adaptation despite homogeneous genomes



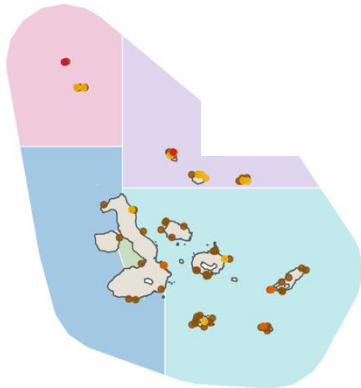
Atlantic silversides

Differences confined to small genomic regions



Atlantic cod

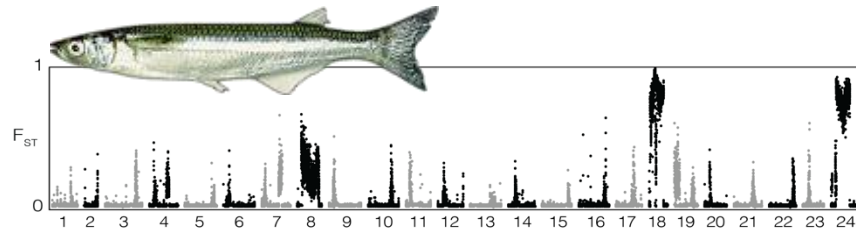
Cryptic mixing of distinct lineages



Brown sea cucumber

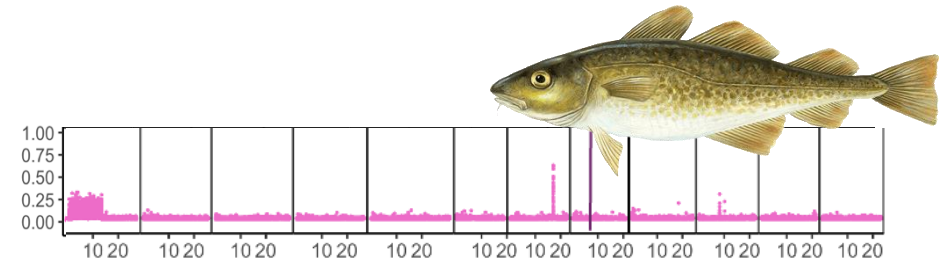
Structure can be hidden in different ways

Strong adaptation despite homogeneous genomes



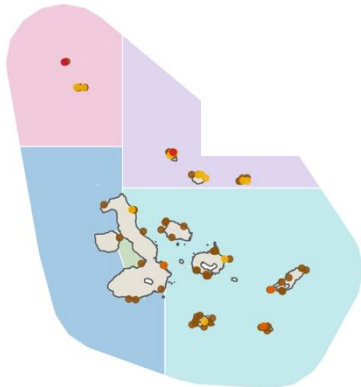
Atlantic silversides

Differences confined to small genomic regions



Atlantic cod

Cryptic mixing of distinct lineages



Brown sea cucumber

Tracking who mixes with whom—and when

		Population of Origin											
		MI	AW	KE	ME	CO	HU	DE	PO	BL	RK	SC	SJ
Population Assigned	MI	29	---	---	---	---	---	---	---	---	---	---	---
	AW	---	29	---	---	---	---	---	---	---	---	---	---
	KE	---	---	24	3	---	---	---	---	---	---	---	---
	ME	---	---	3	24	3	---	---	---	---	---	---	---
	CO	---	---	1	---	24	2	---	---	---	---	---	---
	HU	---	---	---	---	---	26	---	---	---	---	---	---
	DE	---	---	---	---	---	---	27	---	---	1	---	---
	PO	---	---	---	---	---	---	1	28	---	---	---	---
	BL	---	---	---	---	---	---	---	---	26	3	---	---
	RK	---	---	---	---	---	---	---	---	2	24	---	---
	SC	---	---	---	---	1	---	---	---	---	---	28	---
	SJ	---	---	---	---	---	---	---	---	---	---	---	28



American shad

Atlantic silverside *Menidia menidia*

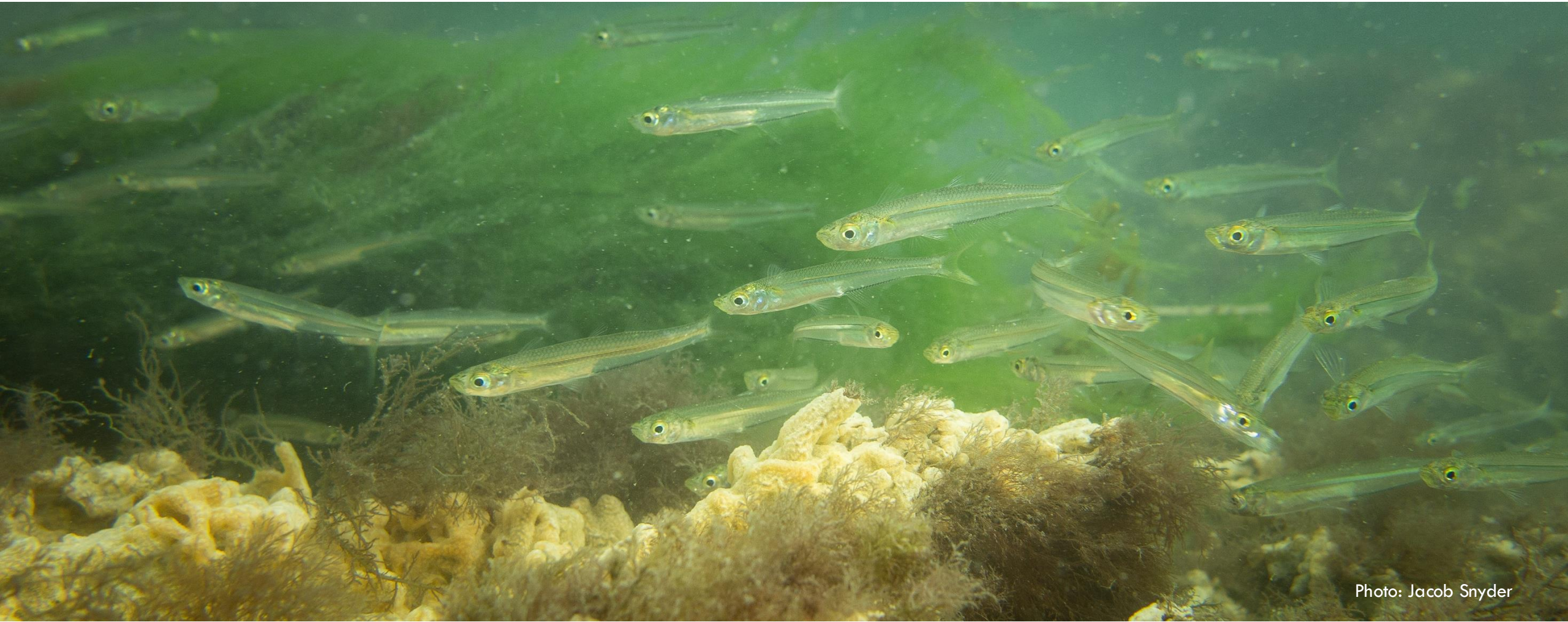
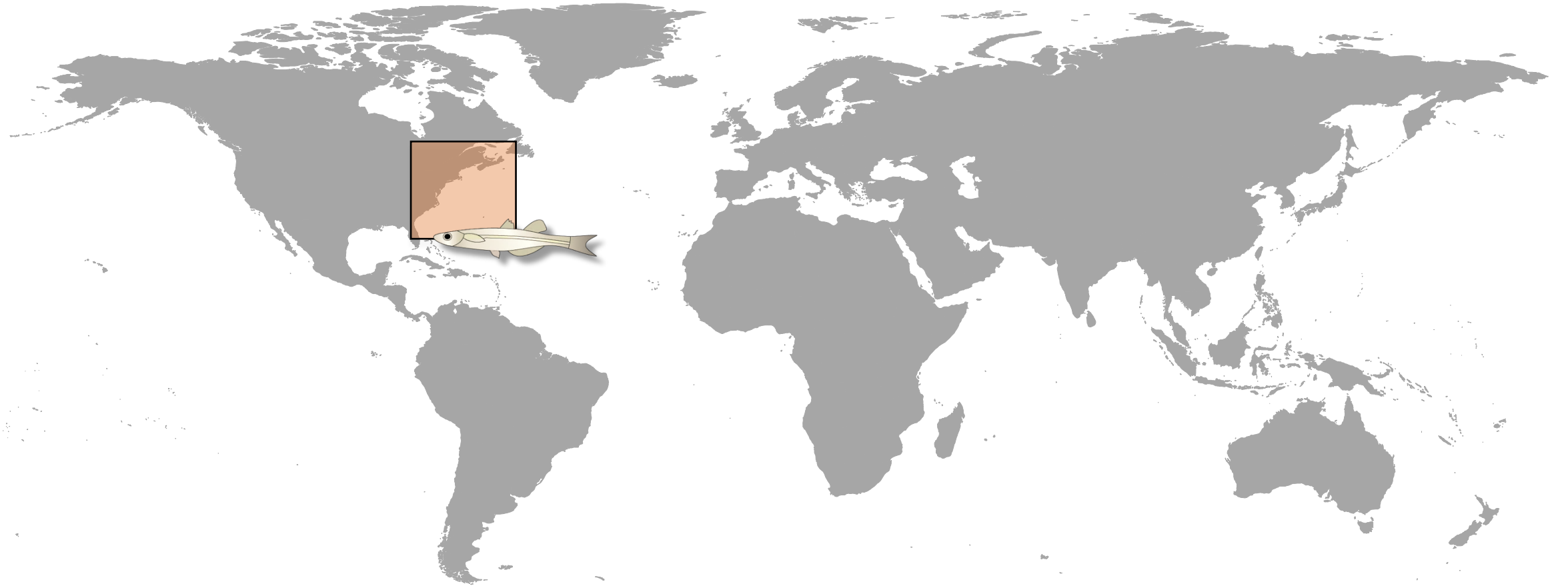
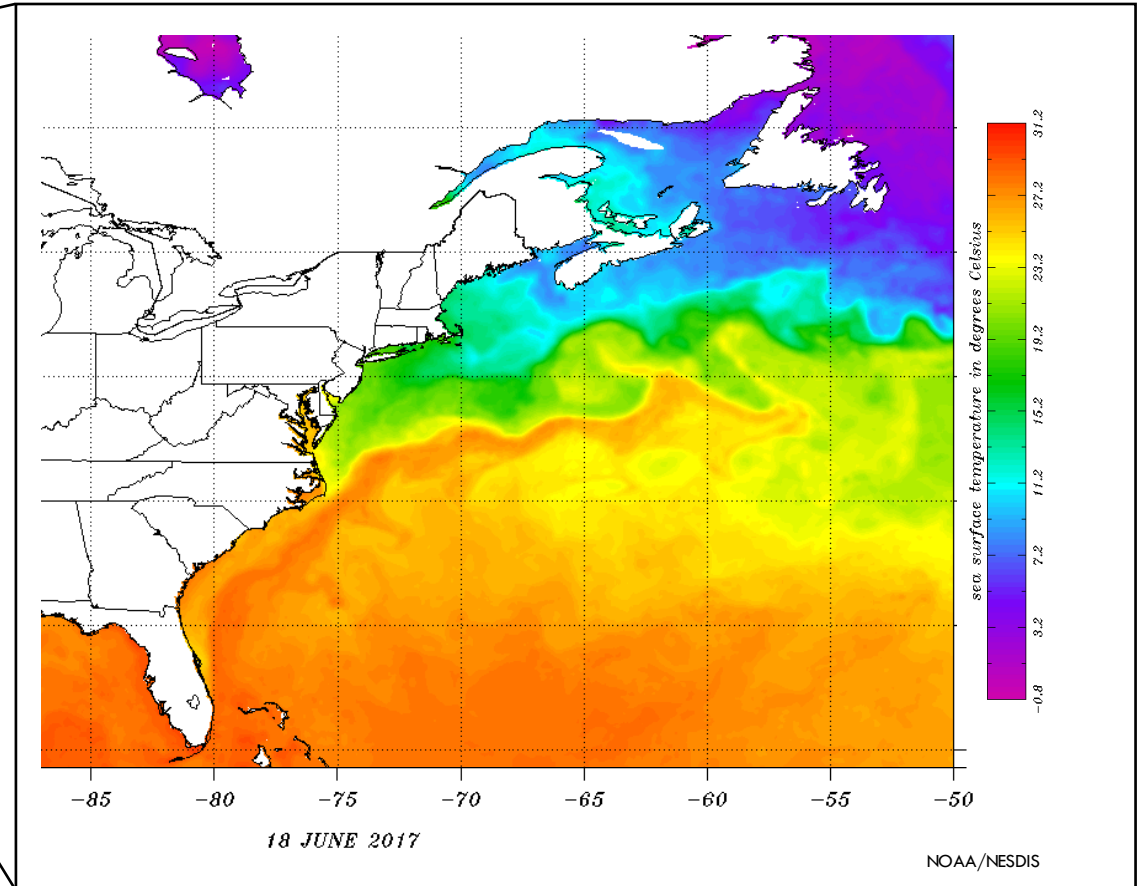
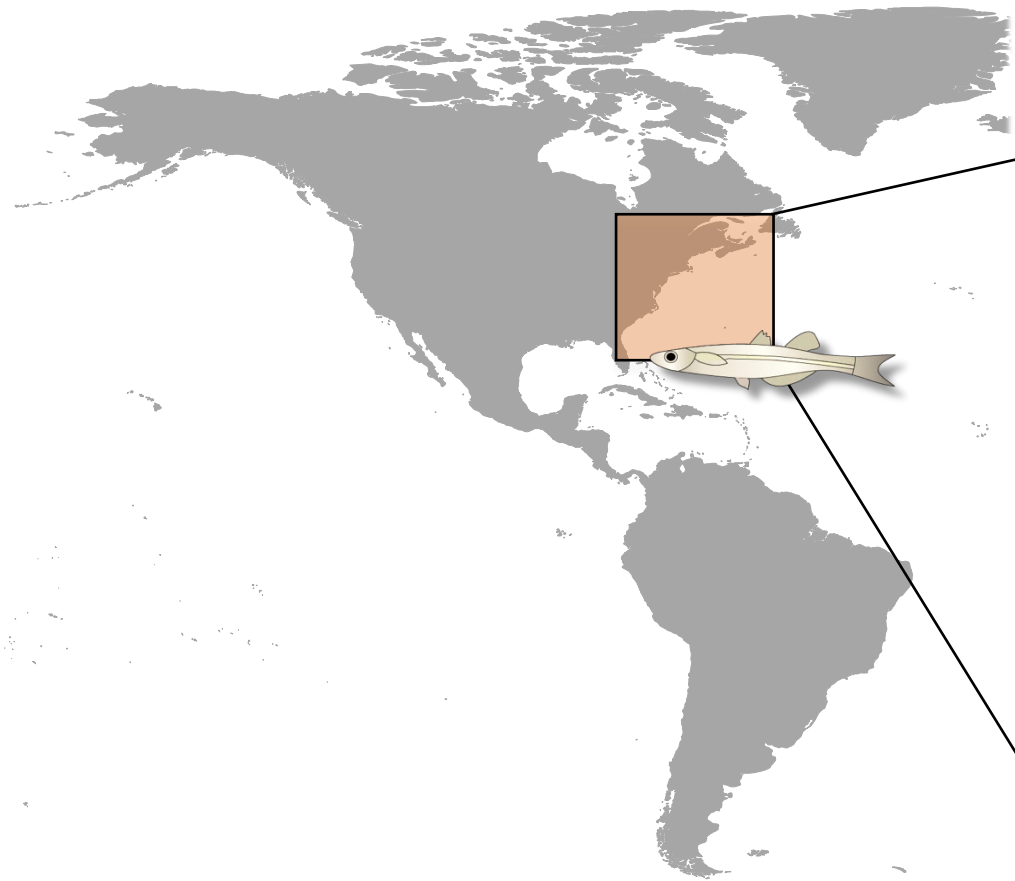
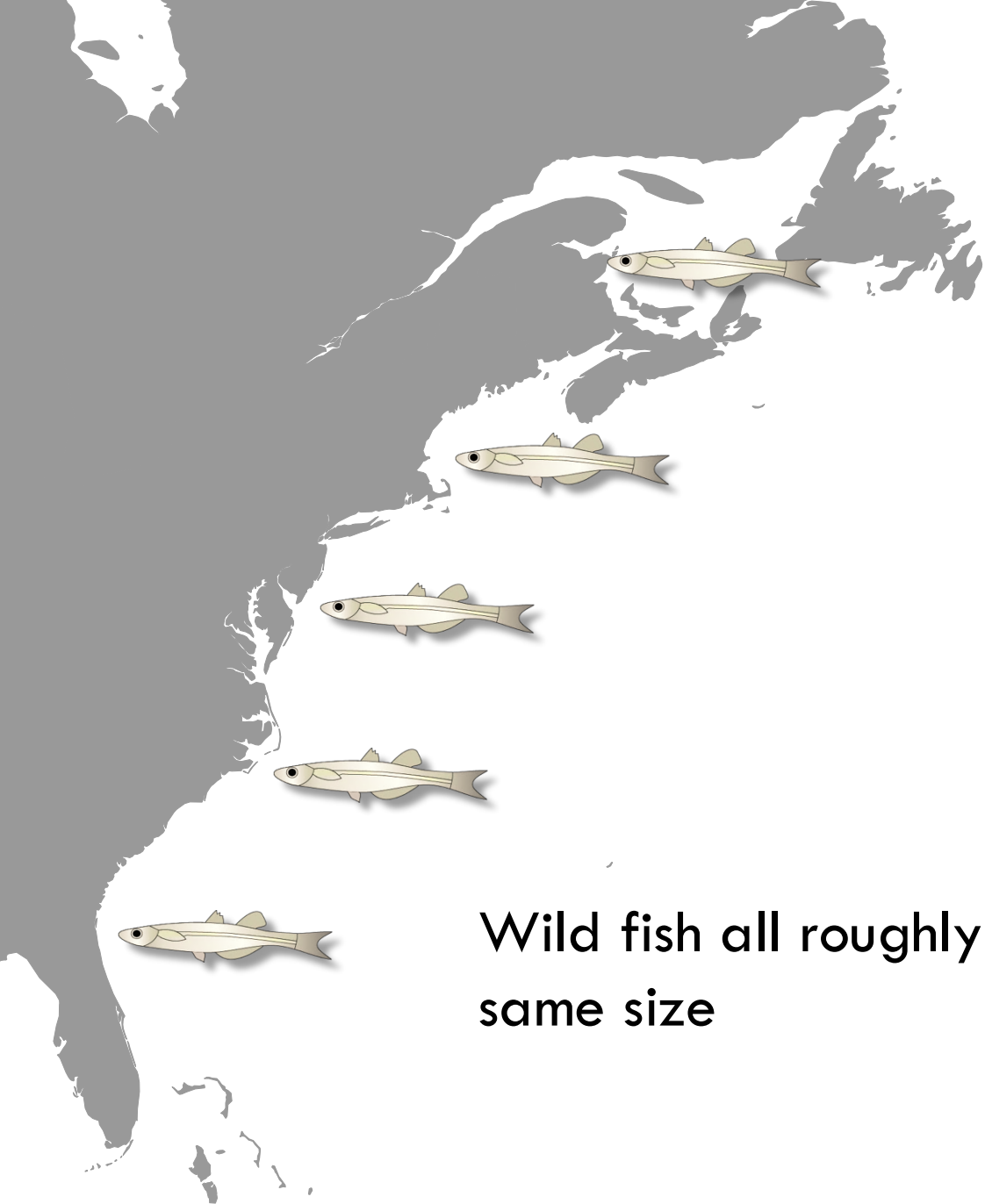


Photo: Jacob Snyder



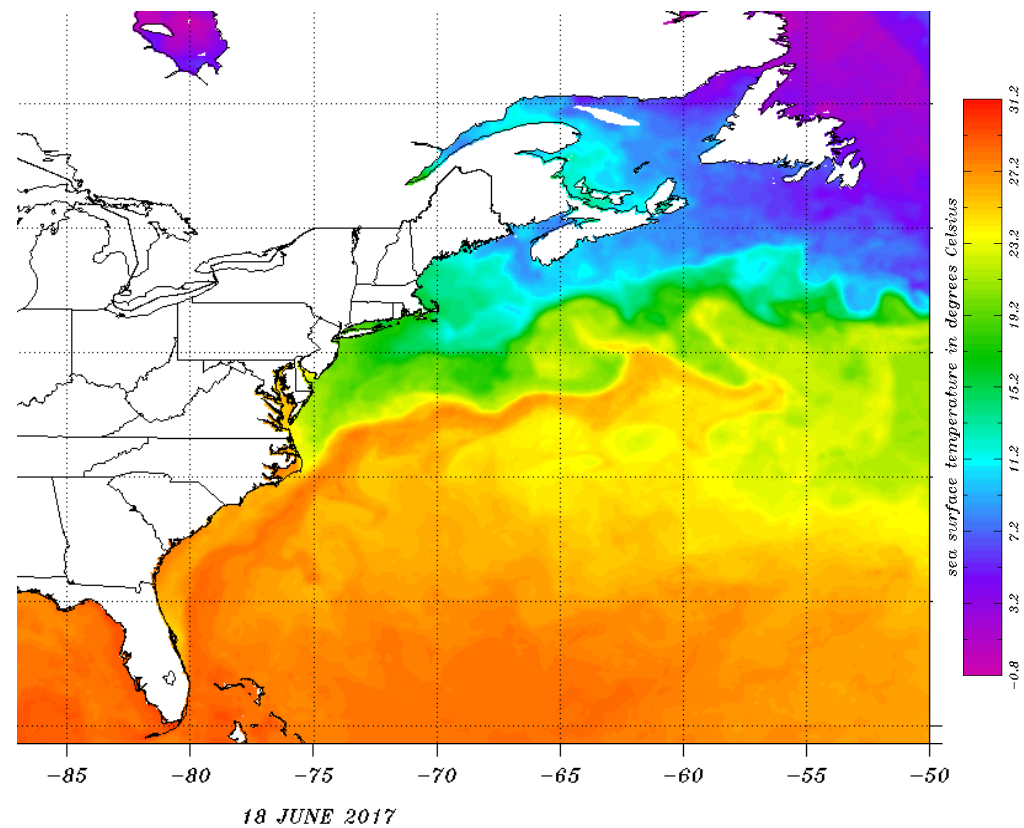
One of the world's steepest thermal gradients



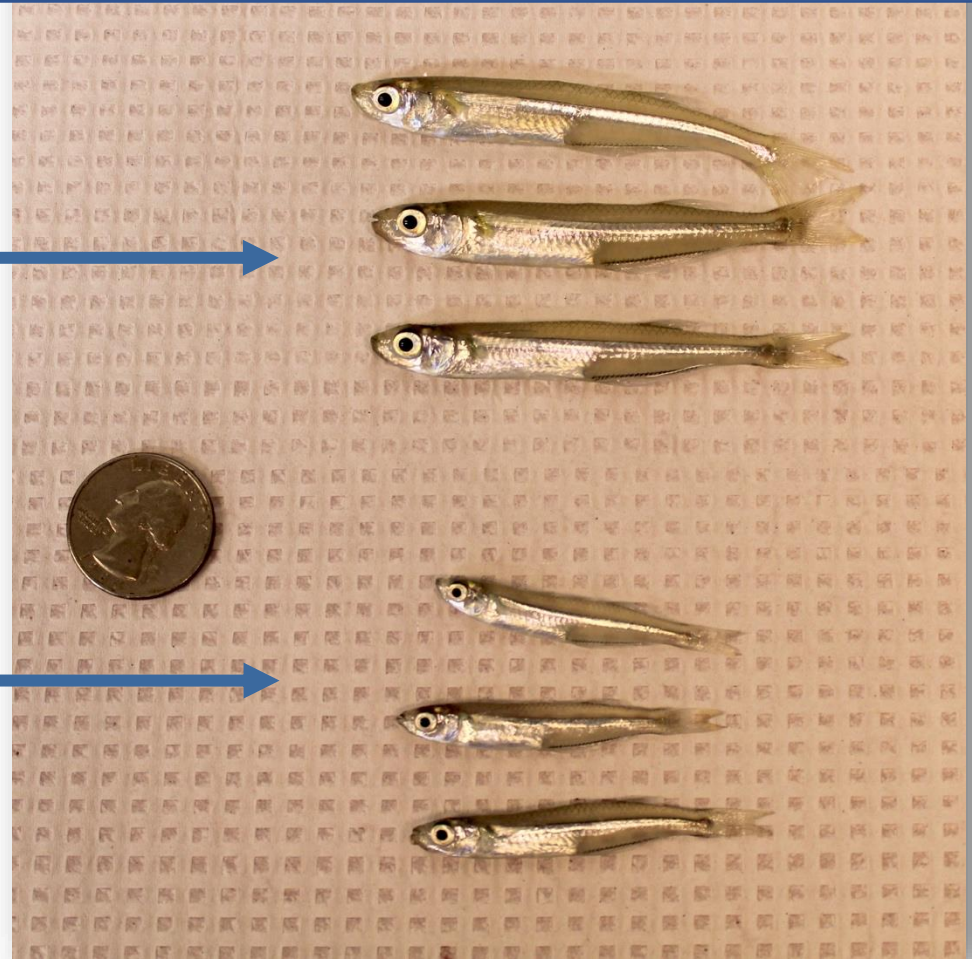


Wild fish all roughly the same size

One of the world's steepest thermal gradients



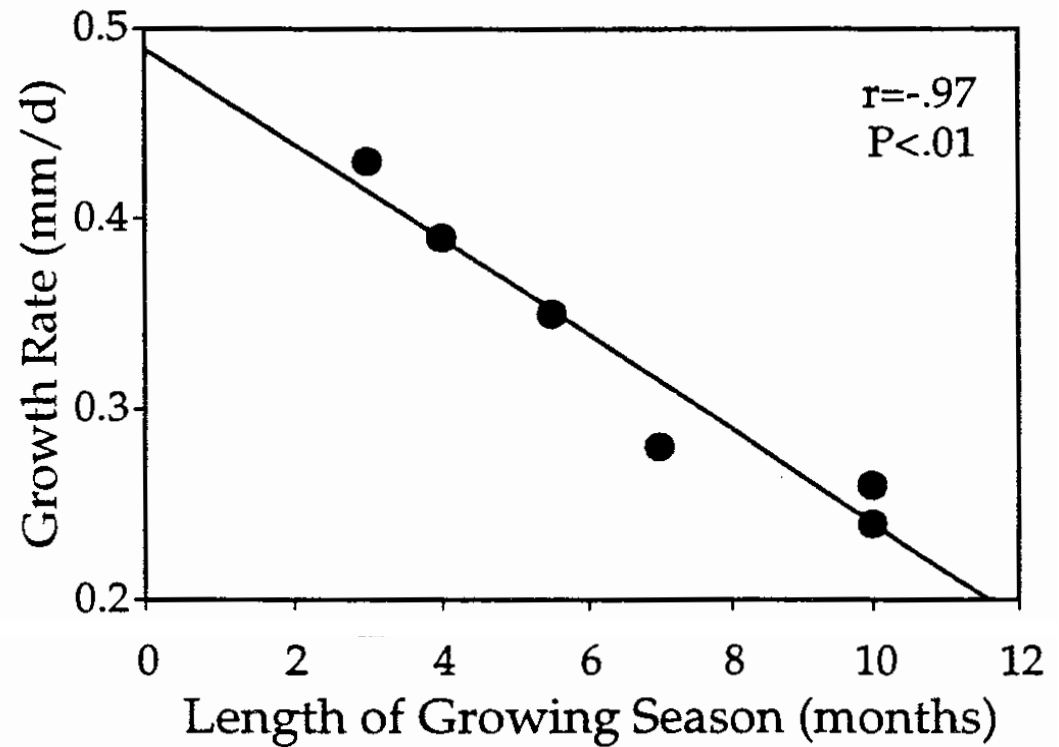
**Same age,
Common lab environment**



Growth capacity is tightly correlated with latitude



David Conover et al.



Reproduced from Conover. 1998. Bull. Mar. Sci.



Growing
season

Short



Long

Winter
duration

Long



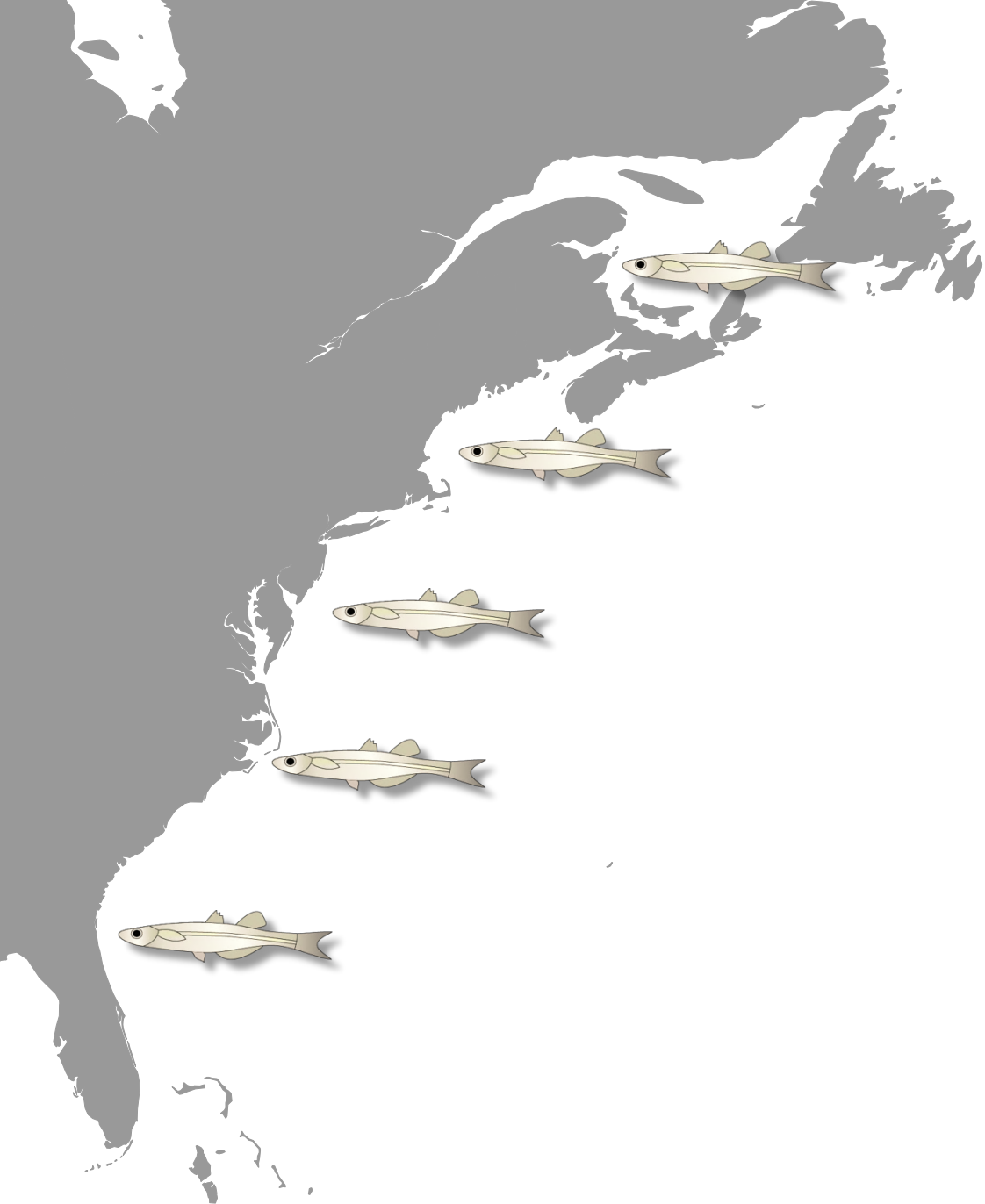
Short

Size-selective
winter mortality

Intense



Minor



Growing season

Short



Long

Winter duration

Long



Short

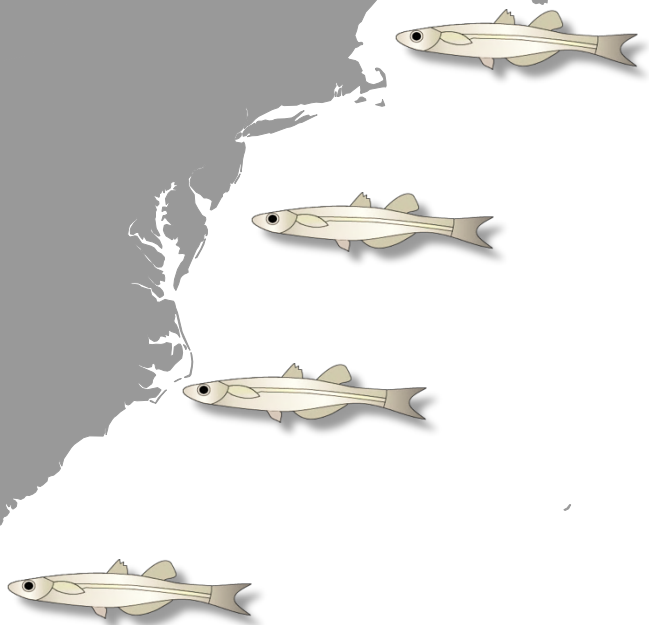
Size-selective winter mortality

Intense



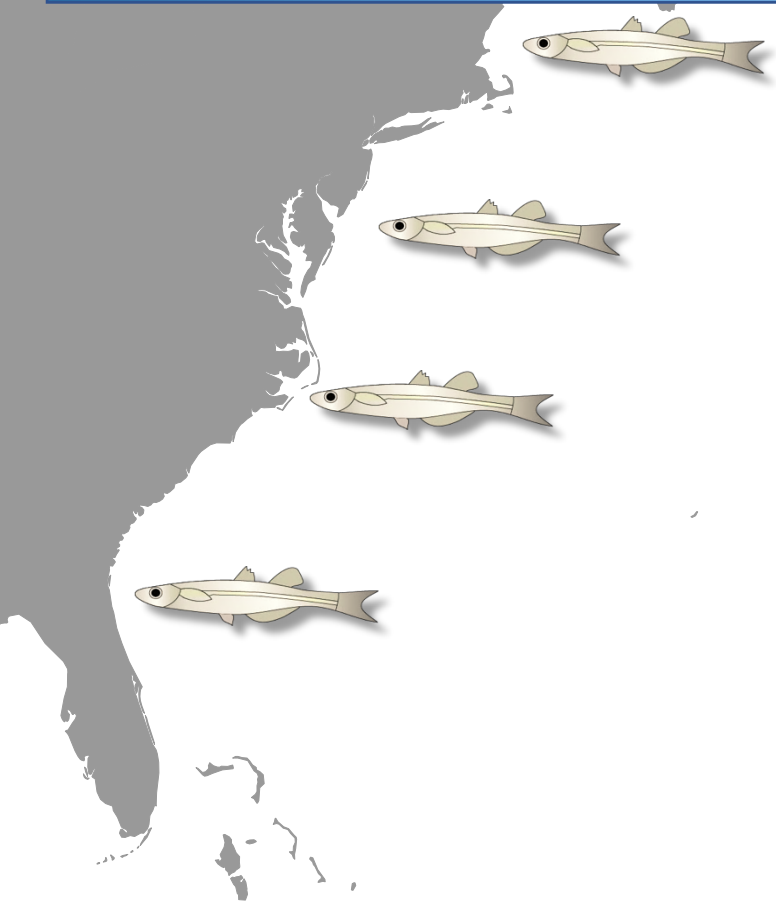
Minor

Counter-gradient variation:
Genetic influences on a trait oppose
environmental influences



Counter-gradient variation:

Genetic influences on a trait oppose environmental influences



Reduced phenotypic variation
→ **Cryptic adaptive divergence**

Local adaptation demonstrated in multiple correlated traits

Other traits that vary with latitude:

- Rates of energy consumption
- Metabolism
- Growth efficiency
- Lipid energy reserves
- Egg production rate
- Egg size and size of offspring at hatch
- Willingness to forage under threat of predation
- Number of vertebrae
- Levels of temperature-dep. sex determination



Local adaptation demonstrated in multiple correlated traits

Other traits that vary with latitude:

- Rates of energy consumption
- Metabolism
- Growth efficiency
- Lipid energy reserves
- Egg production rate
- Egg size and size of offspring at hatch
- Willingness to forage under threat of predation
- Number of vertebrae
- Levels of temperature-dep. sex determination



Maintained despite
extensive gene flow

Capacity for rapid evolution



REPORTS

Sustaining Fisheries Yields Over Evolutionary Time Scales

David O. Conover* and Stephan B. Munch

Fishery management plans ignore the potential for evolutionary change in harvestable biomass. We subjected populations of an exploited fish (*Menidia menidia*) to large, small, or random size-selective harvest of adults over four generations. Harvested biomass evolved rapidly in directions counter to the size-dependent force of fishing mortality. Large-harvested populations initially produced the highest catch but quickly evolved a lower yield than controls. Small-harvested populations did the reverse. These shifts were caused by selection of genotypes with slower or faster rates of growth. Management tools that preserve natural genetic variation are necessary for long-term sustainable yield.

It is well established that wild pest and pathogen populations may evolve in response to anthropogenic forces of mortality (I), but is the same true of fisheries? Fishing mortality is higher

ing behavior], with one major exception. The short generation time of *M. menidia* (1 year) coupled with the ease with which large populations can be maintained in captivity enable



initially produced the highest total yield and mean weight of fish but then declined. Small-harvested populations started with low yield and then increased. By the fourth generation of selection, the biomass harvested and the mean

Down

Questions

- What is the genomic basis for the adaptive cline?



Questions

- What is the genomic basis for the adaptive cline?
- Does the underlying genomic architecture help maintain local adaptation in spite of gene flow?





Maria Akopyan
UC Riverside



Arne Jacobs
U. Glasgow



Aryn Wilder
San Diego Zoo

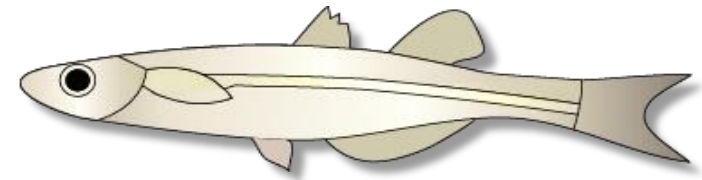


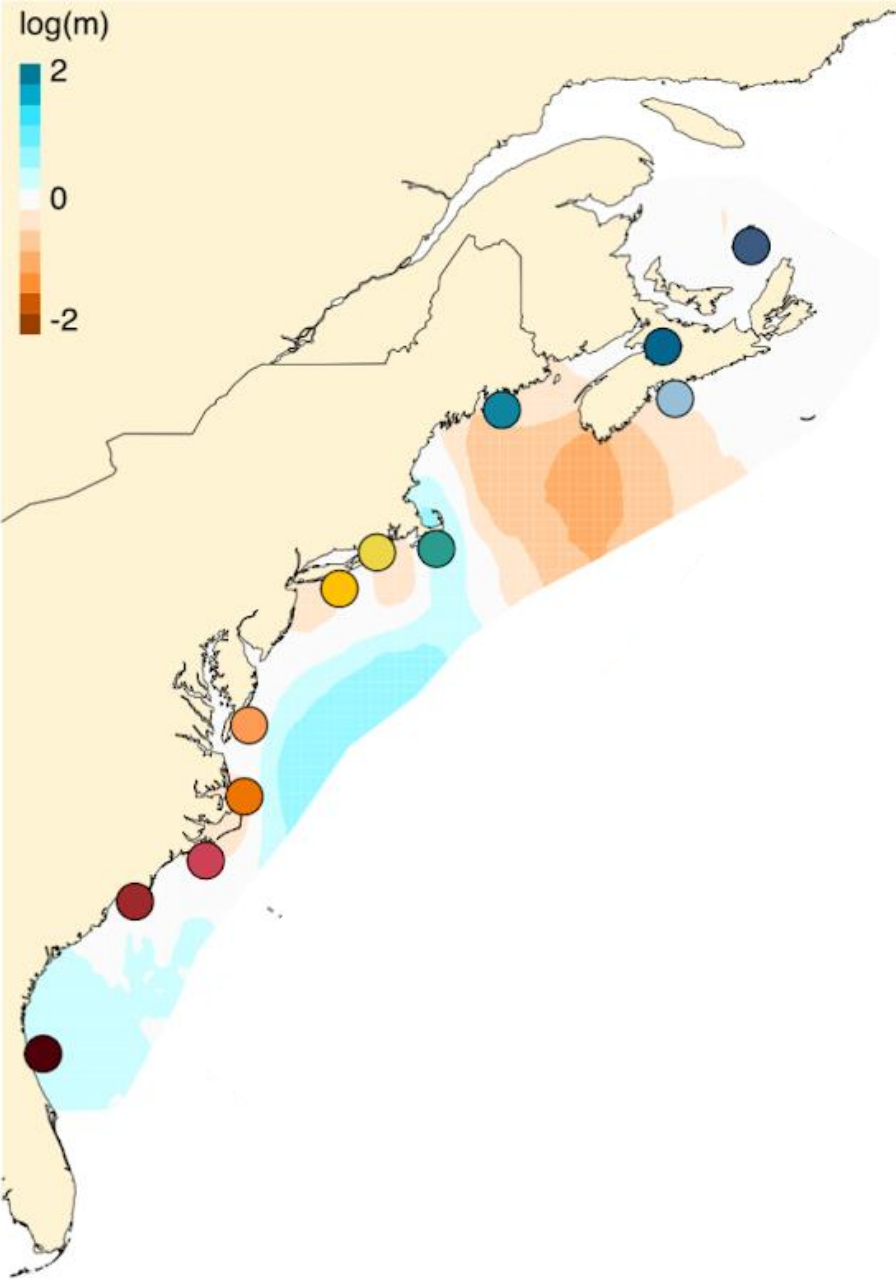
Anna Tigano
DFO Canada



Hannes Baumann
U. Conn

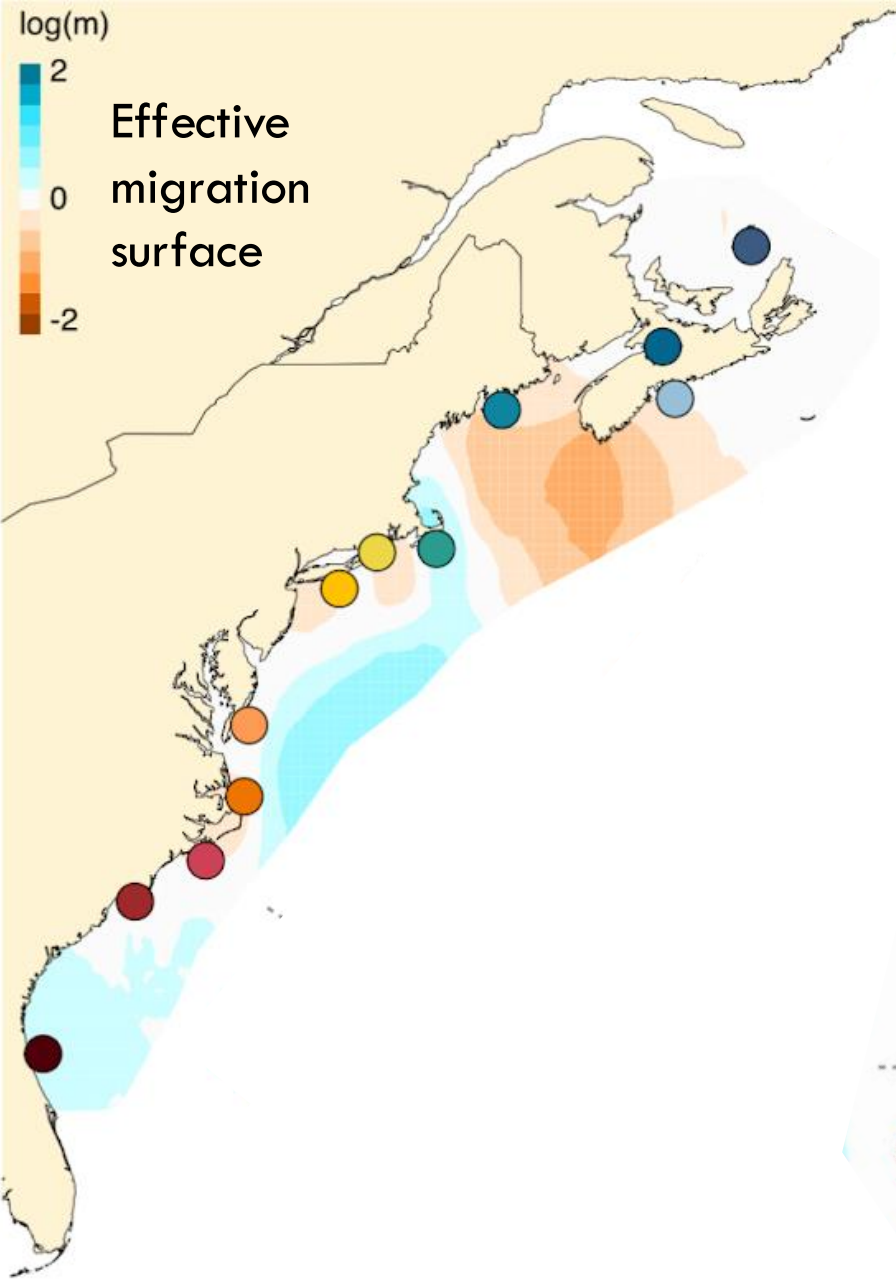
Team silverside





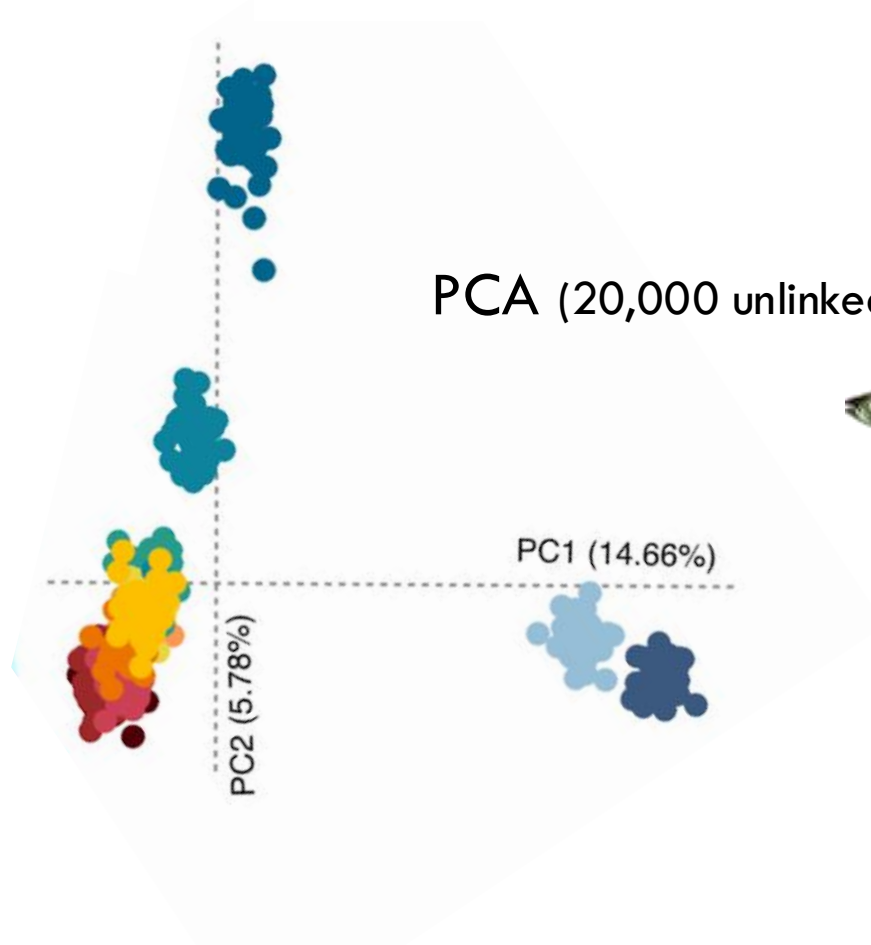
Population structure across latitudes

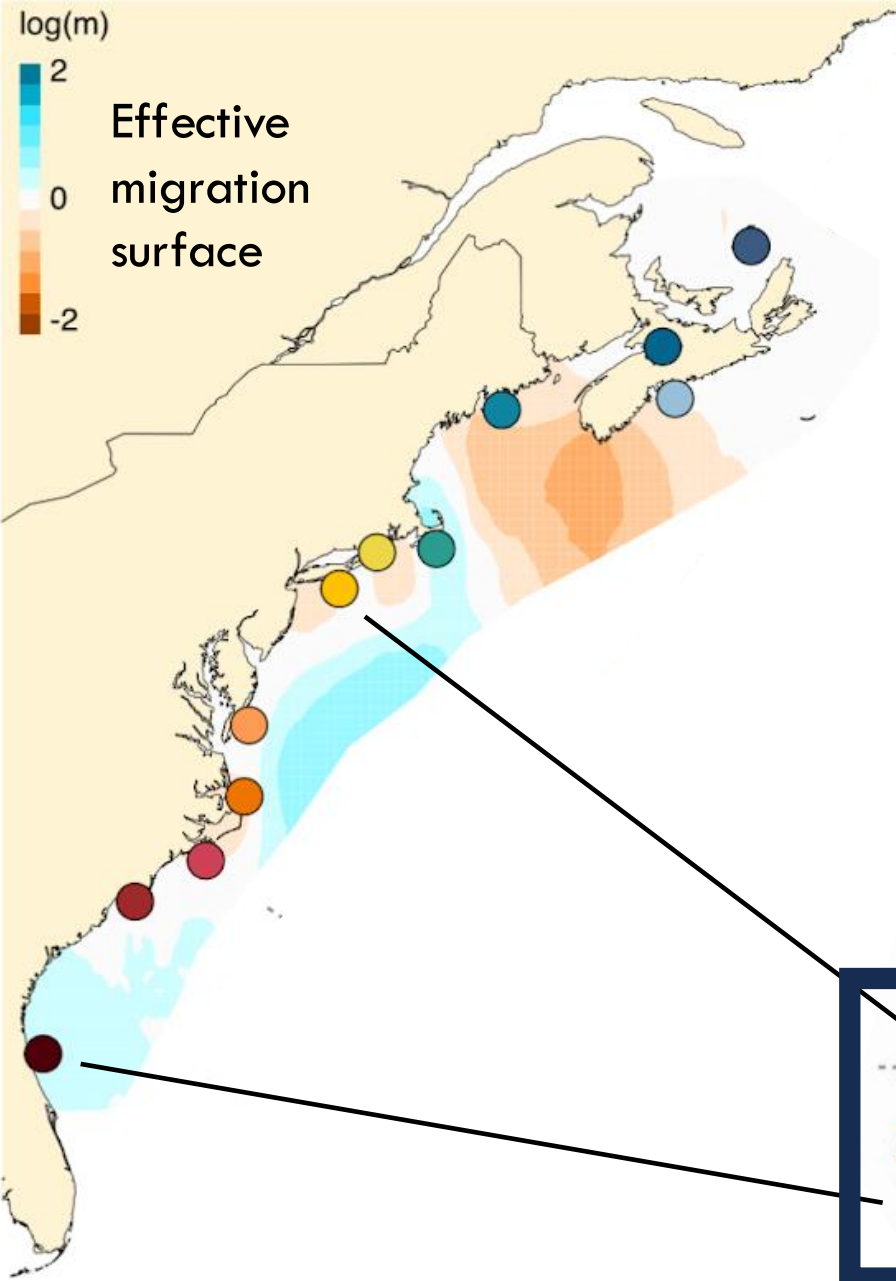
Low-coverage whole genome sequencing (lcWGS)
~1.3x for 50 individuals per population



Population structure across latitudes

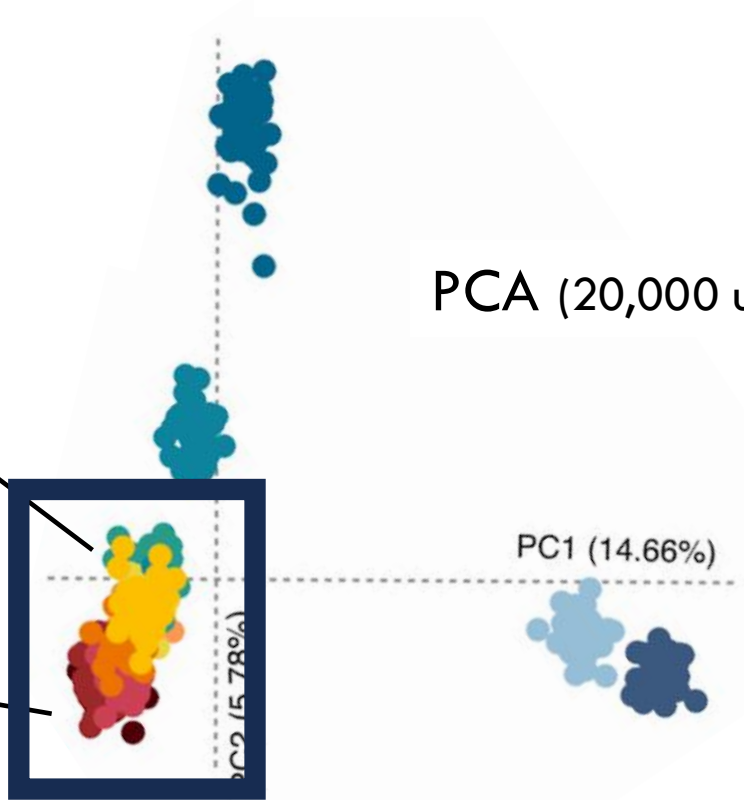
Low-coverage whole genome sequencing (lcWGS)
~1.3x for 50 individuals per population





Population structure across latitudes

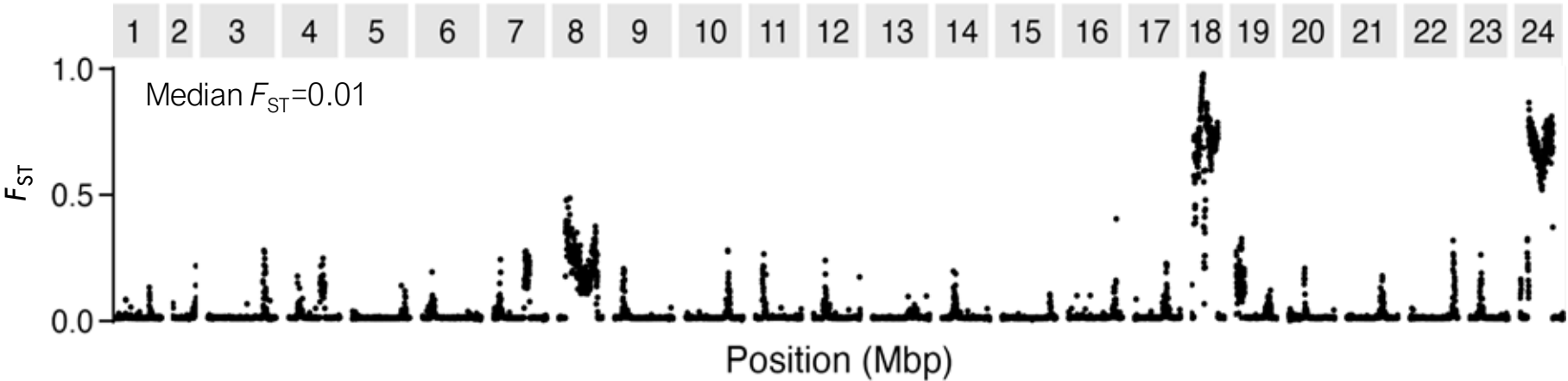
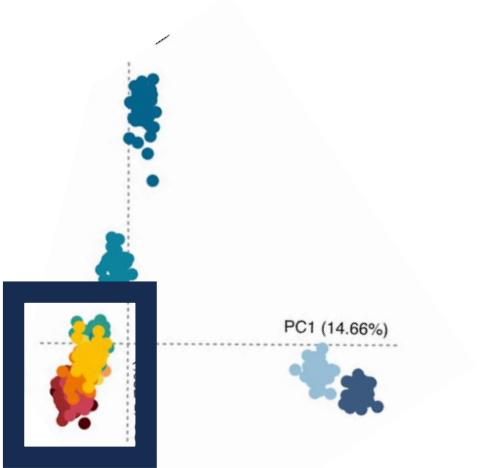
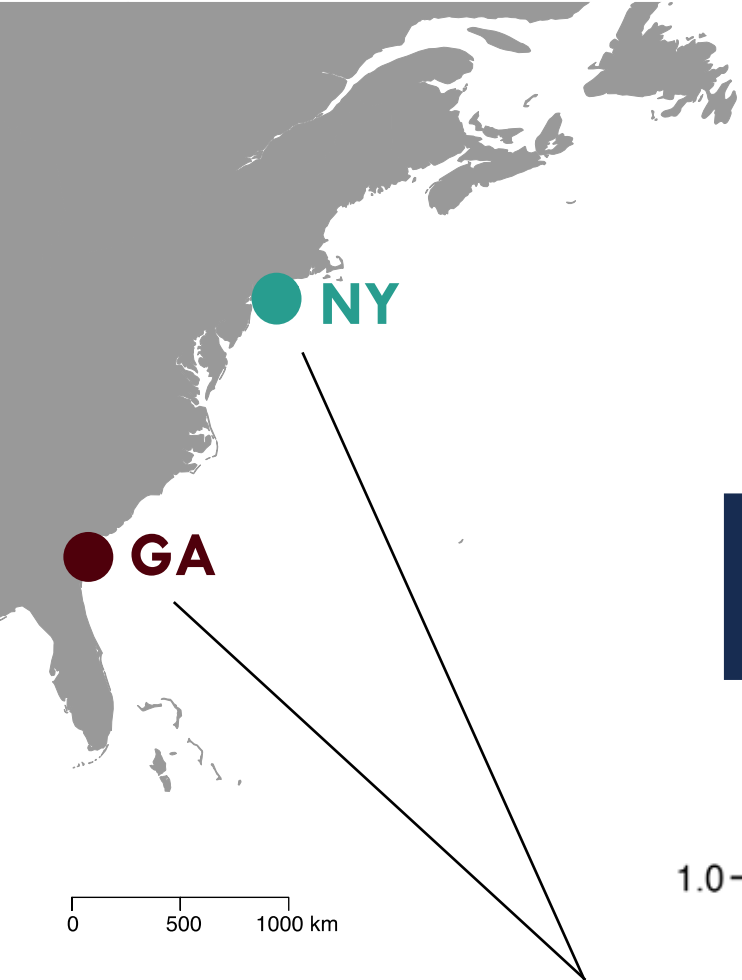
Low-coverage whole genome sequencing (lcWGS)
~1.3x for 50 individuals per population



PCA (20,000 unlinked SNPs)

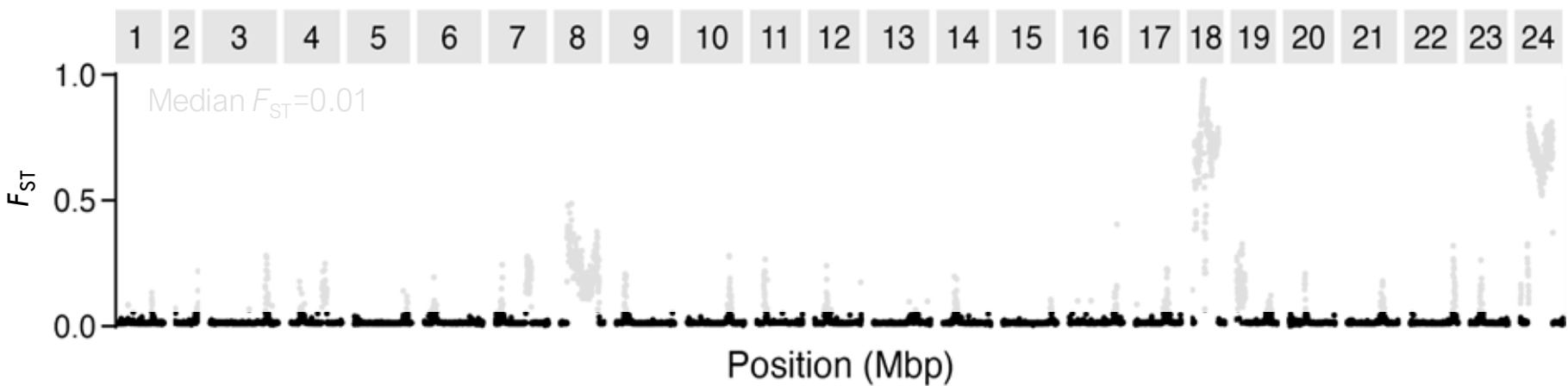
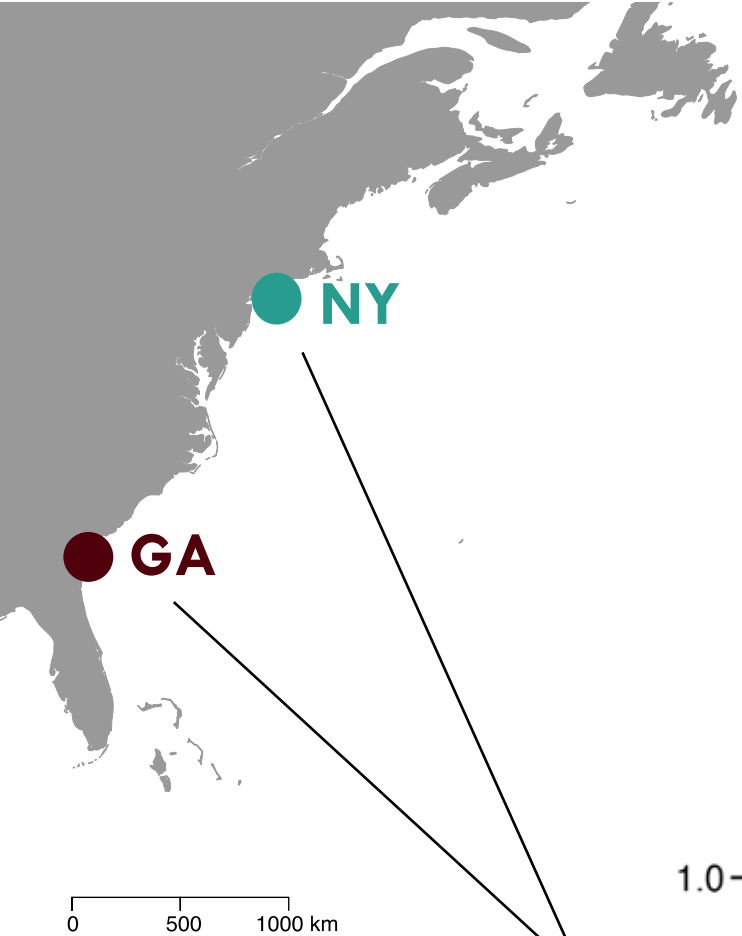


Weak structure hides adaptive differences

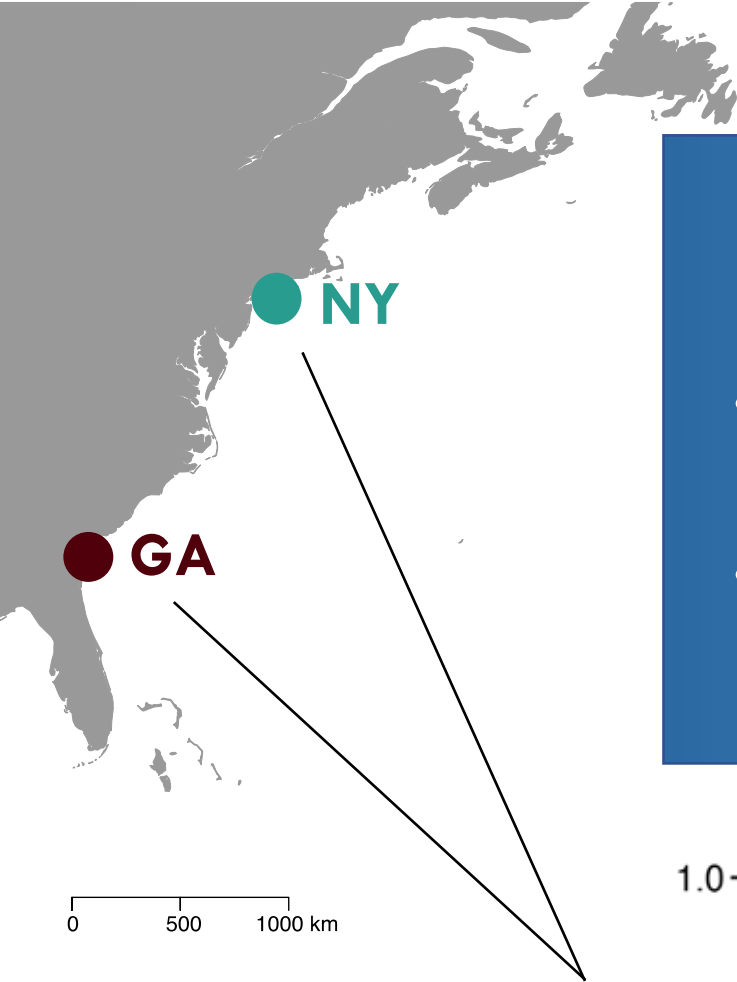


LcWGS data
~1.3x for 50 individuals per population

Weak structure hides adaptive differences

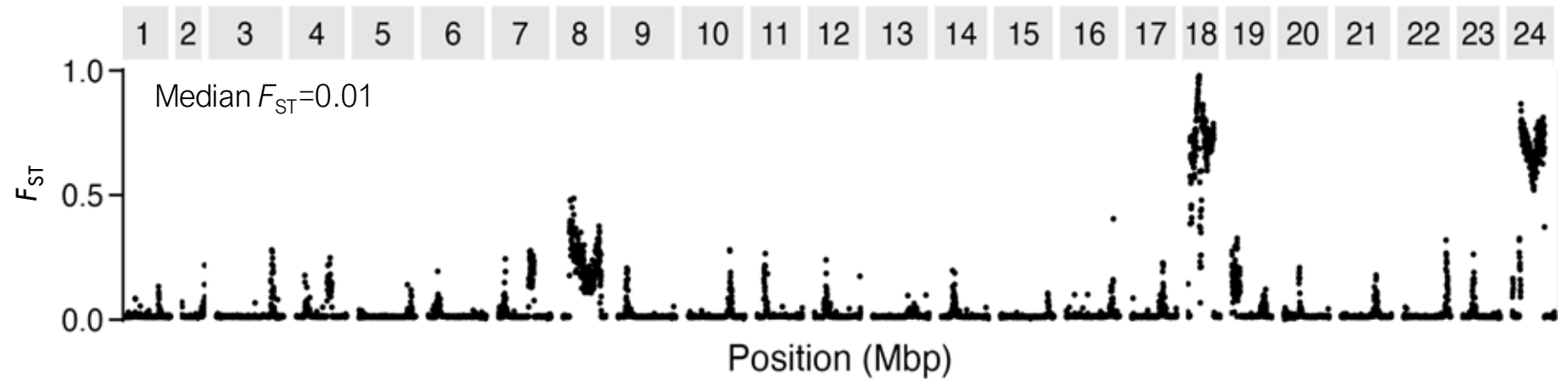


LcWGS data
~1.3x for 50 individuals per population



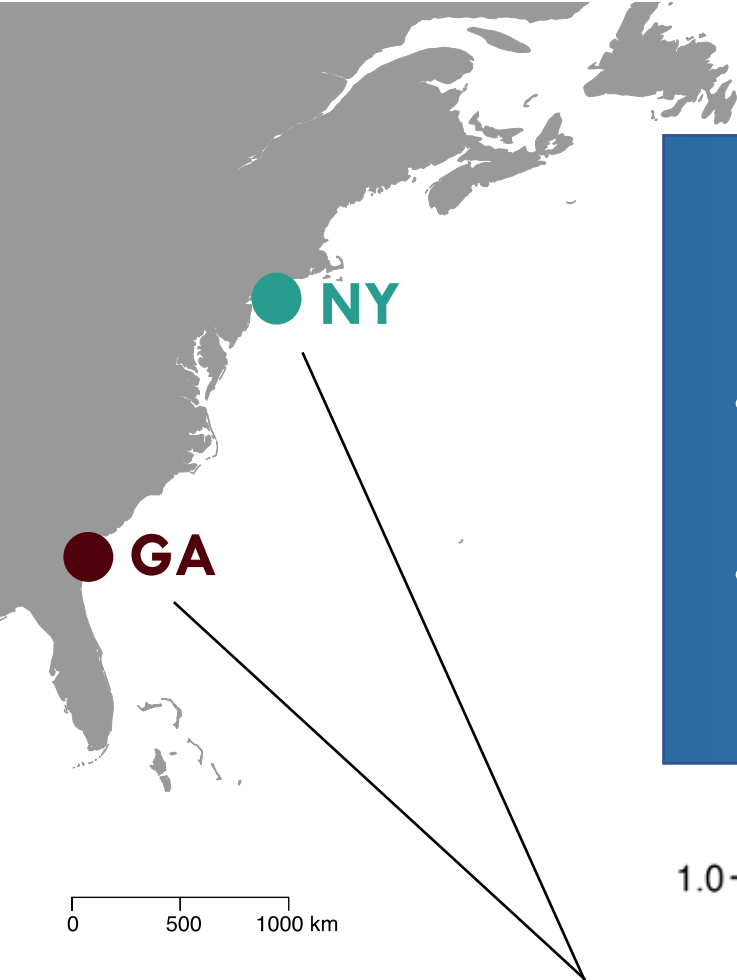
Extreme 'islands of divergence'

- Degree of differentiation (1000s of fixed SNPs)
- Genomic breath (>2/3 the length of multiple chromosomes)



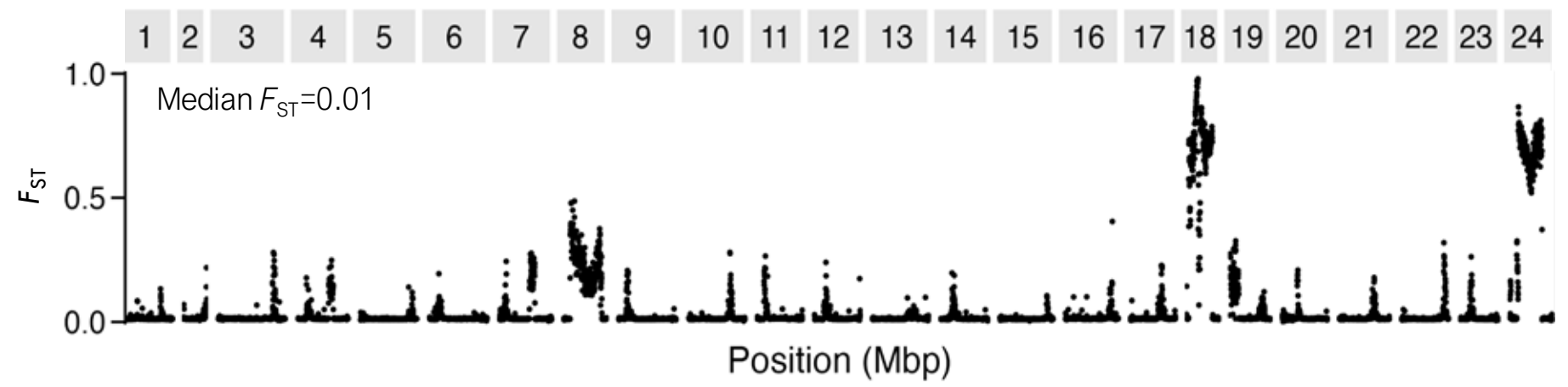
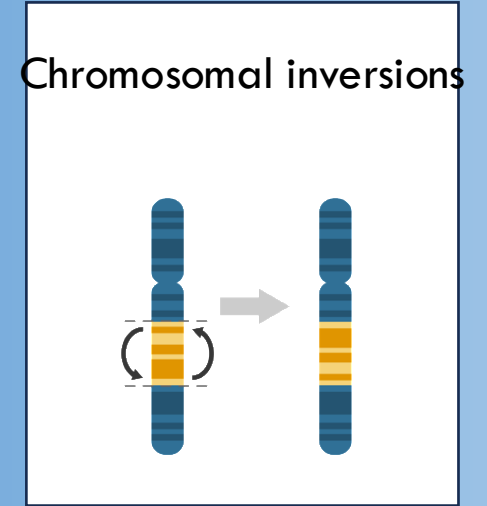
LcWGS data

~1.3x for 50 individuals per population



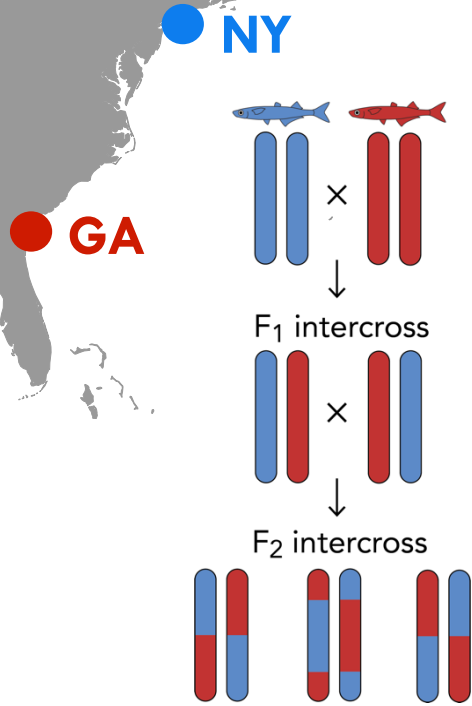
Extreme 'islands of divergence'

- Degree of differentiation (1000s of fixed SNPs)
- Genomic breath (>2/3 the length of multiple chromosomes)



LcWGS data
 ~1.3x for 50 individuals per population

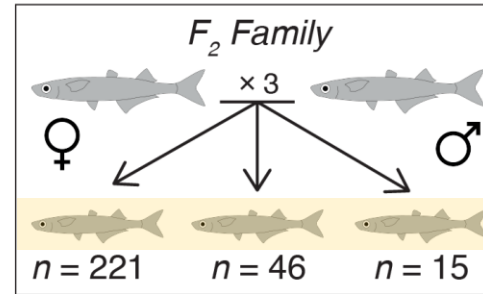
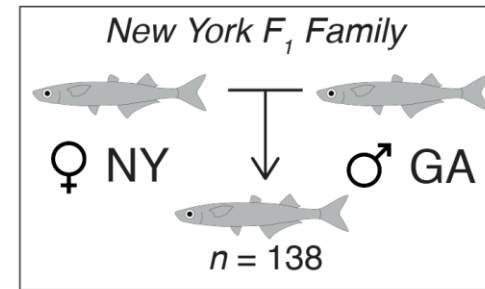
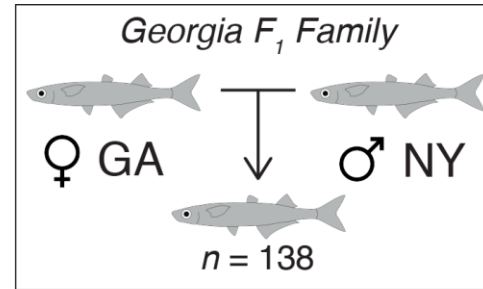
QTL mapping to link genotype to phenotype



wild caught
F₀ generation

F₁ generation

F₂ generation



Phenotyping

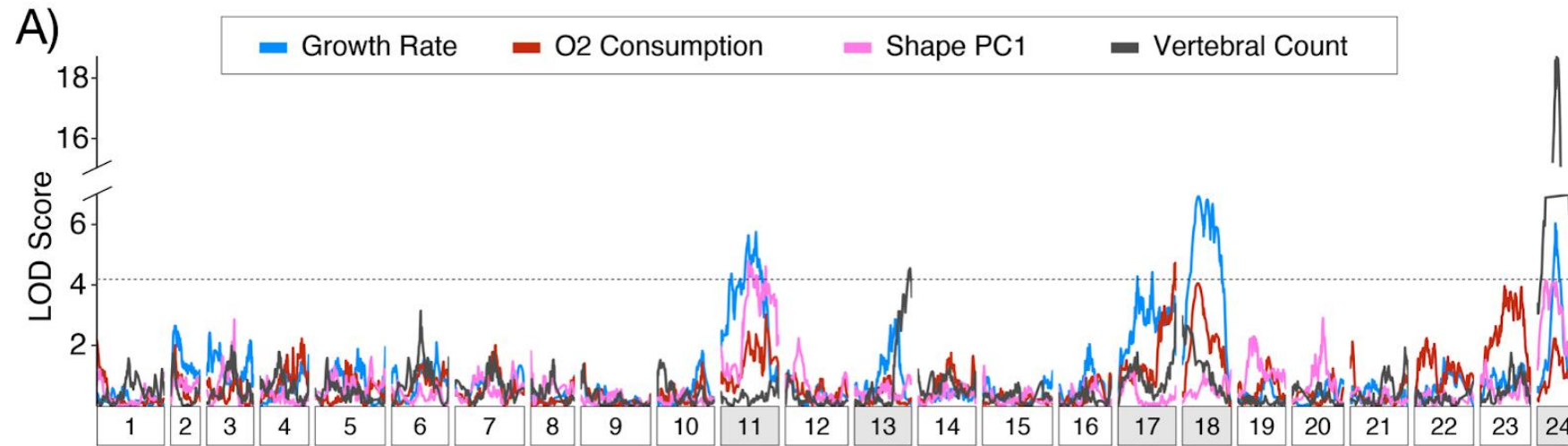
- Growth rate
- Vertebral number
- Body shape
- Resting metabolism
- Swimming speed

Genotyping

- ddRAD, $n=568$, 5 families
- ~50K SNPs @ 15X coverage
- ★ mapped to GA genome

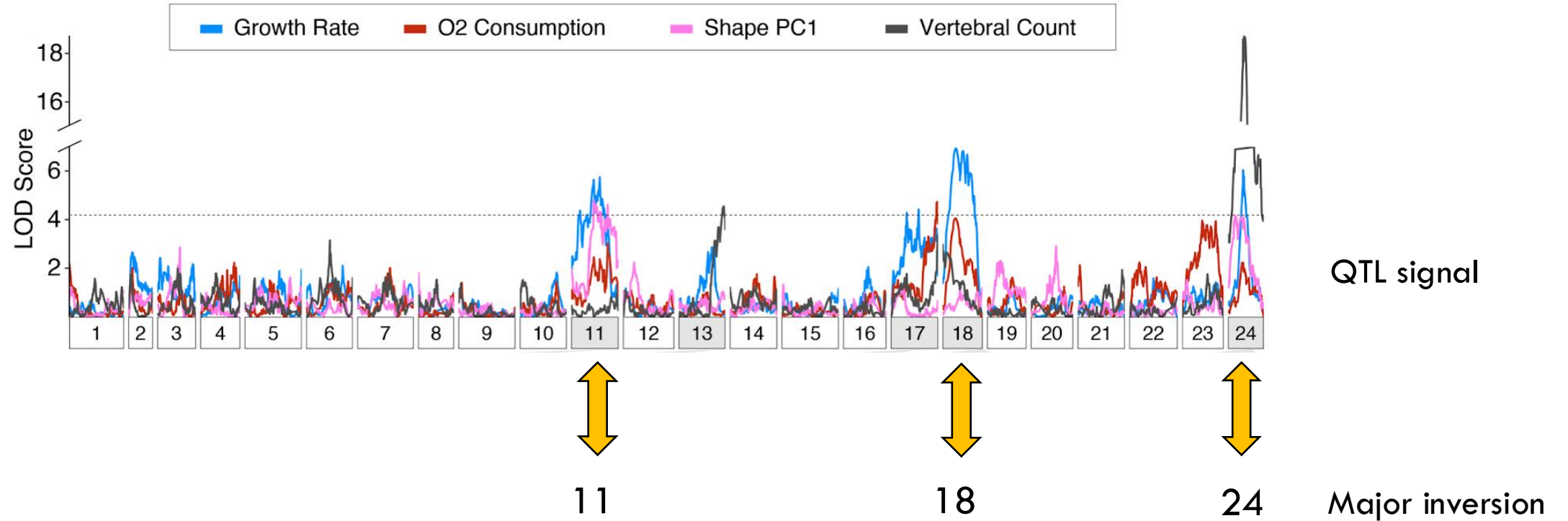


Multiple traits map to the same genomic locations

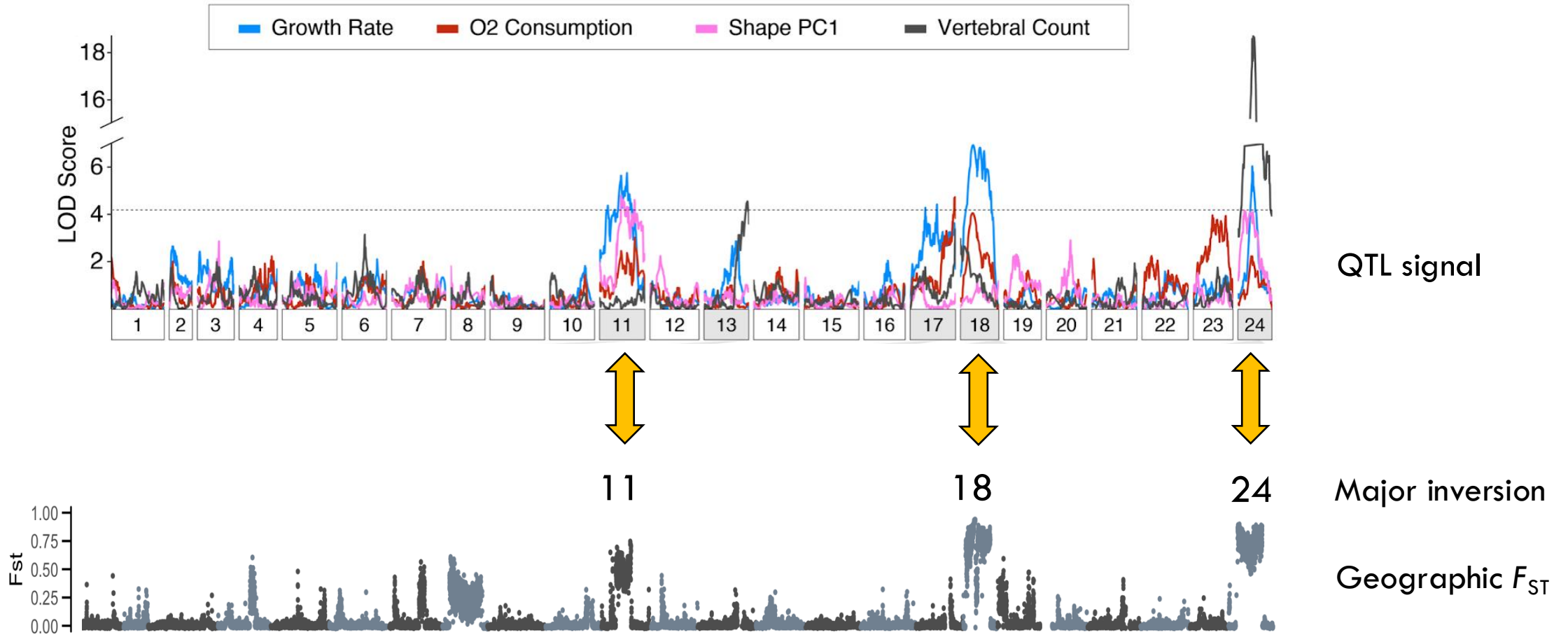


Eight significant QTL on five chromosomes for four traits

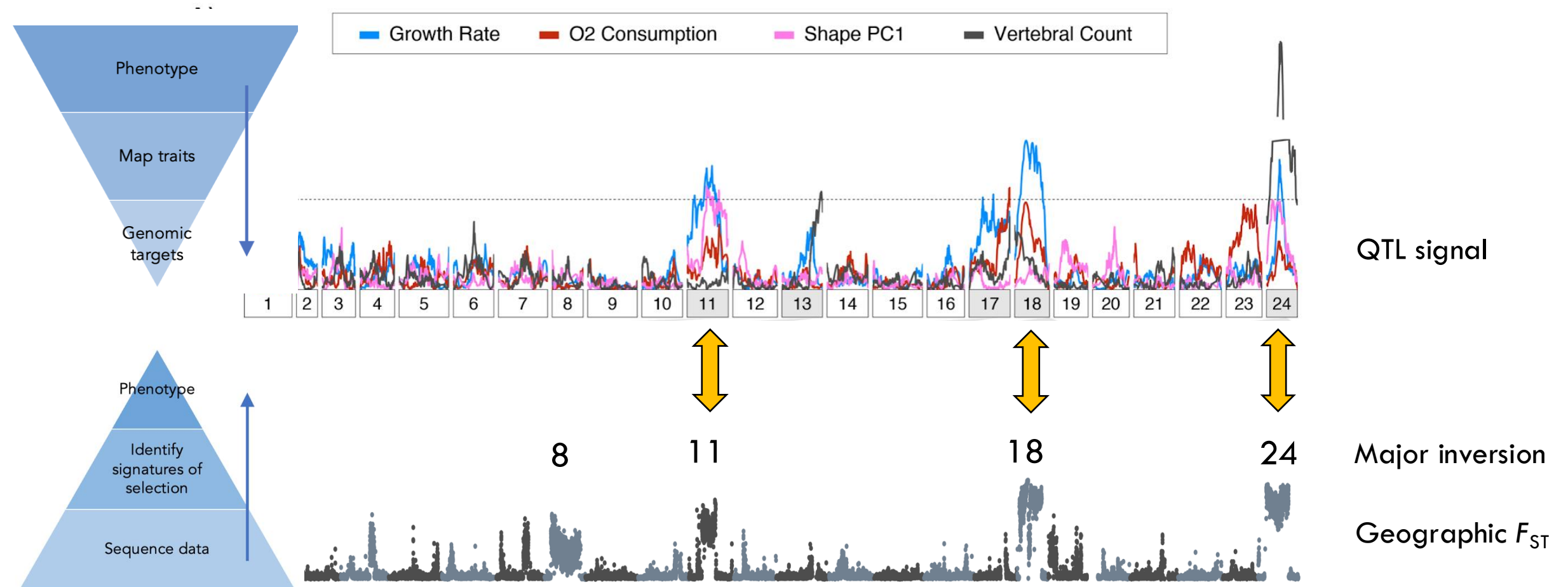
Strongest QTL signals coincide with inversions

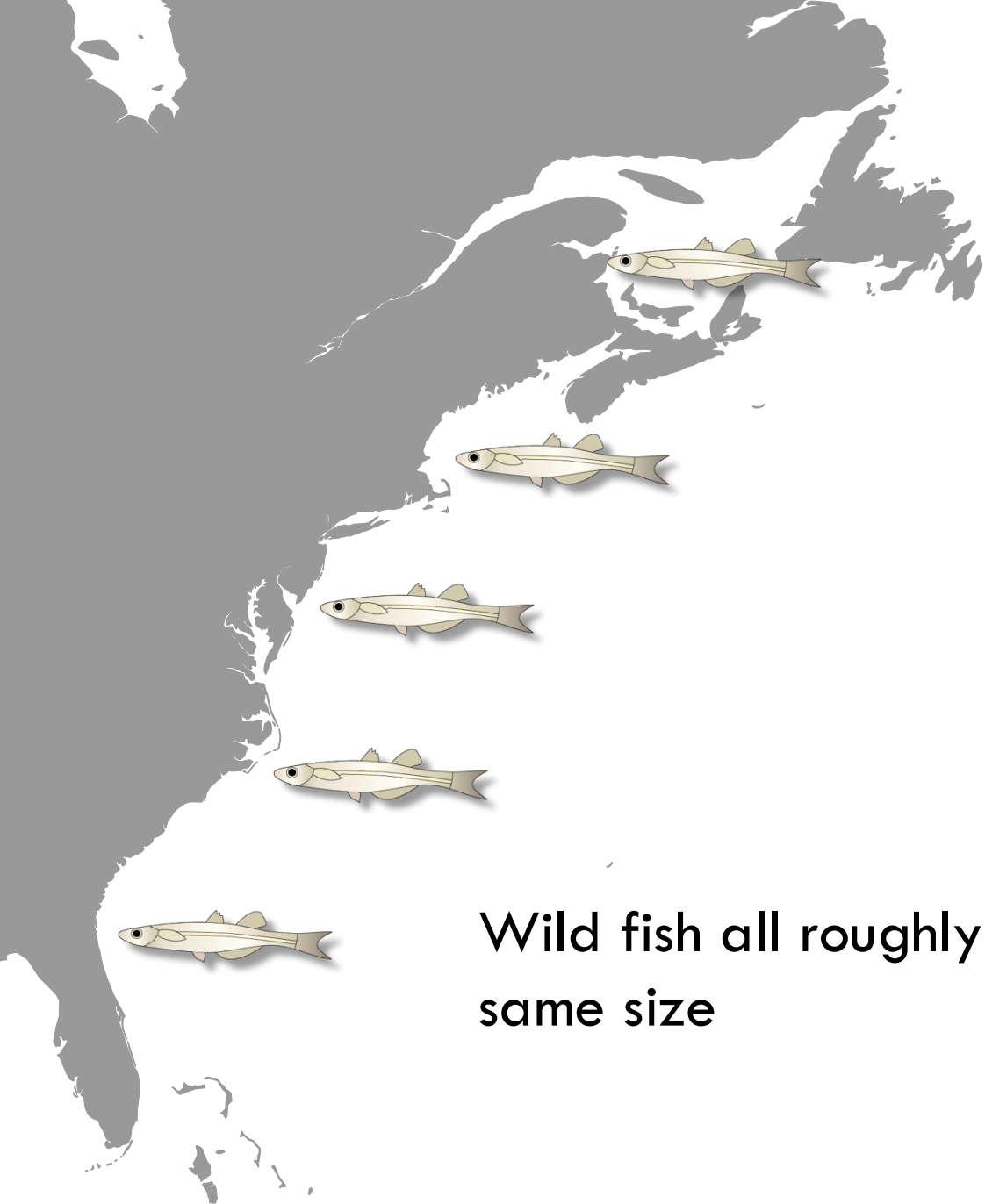


Strongest QTL signals coincide with inversions - and the blocks of strongest differentiation



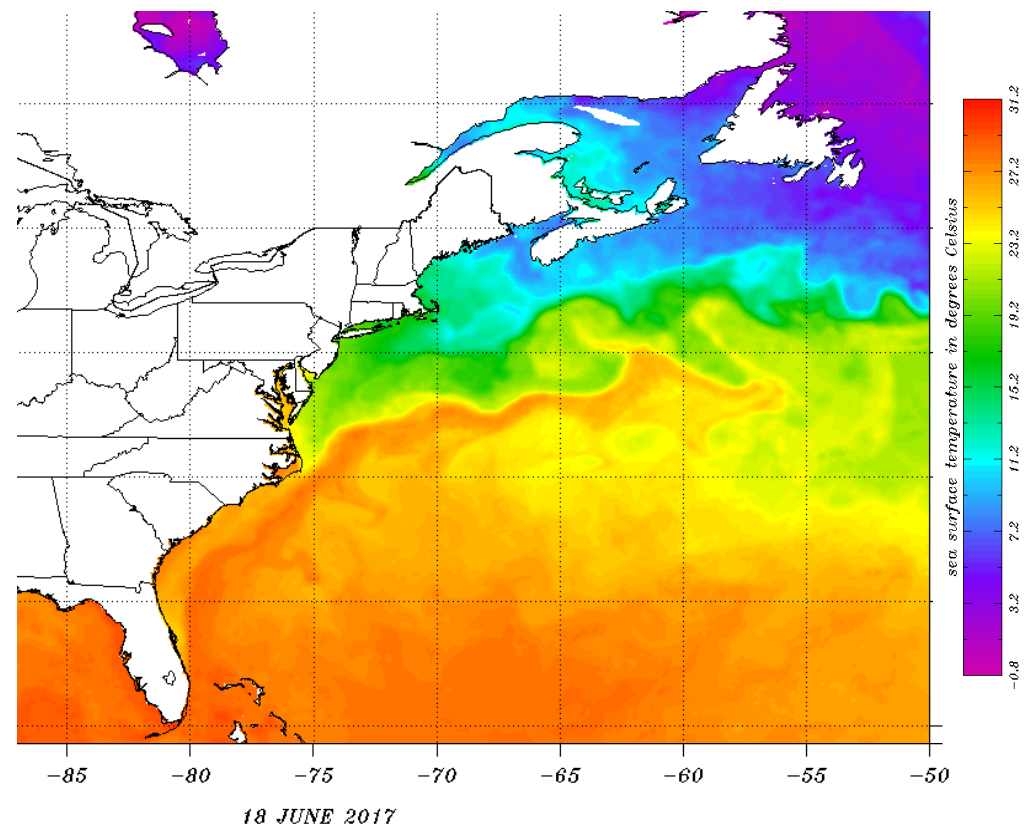
Inversions play a crucial role in local adaptation



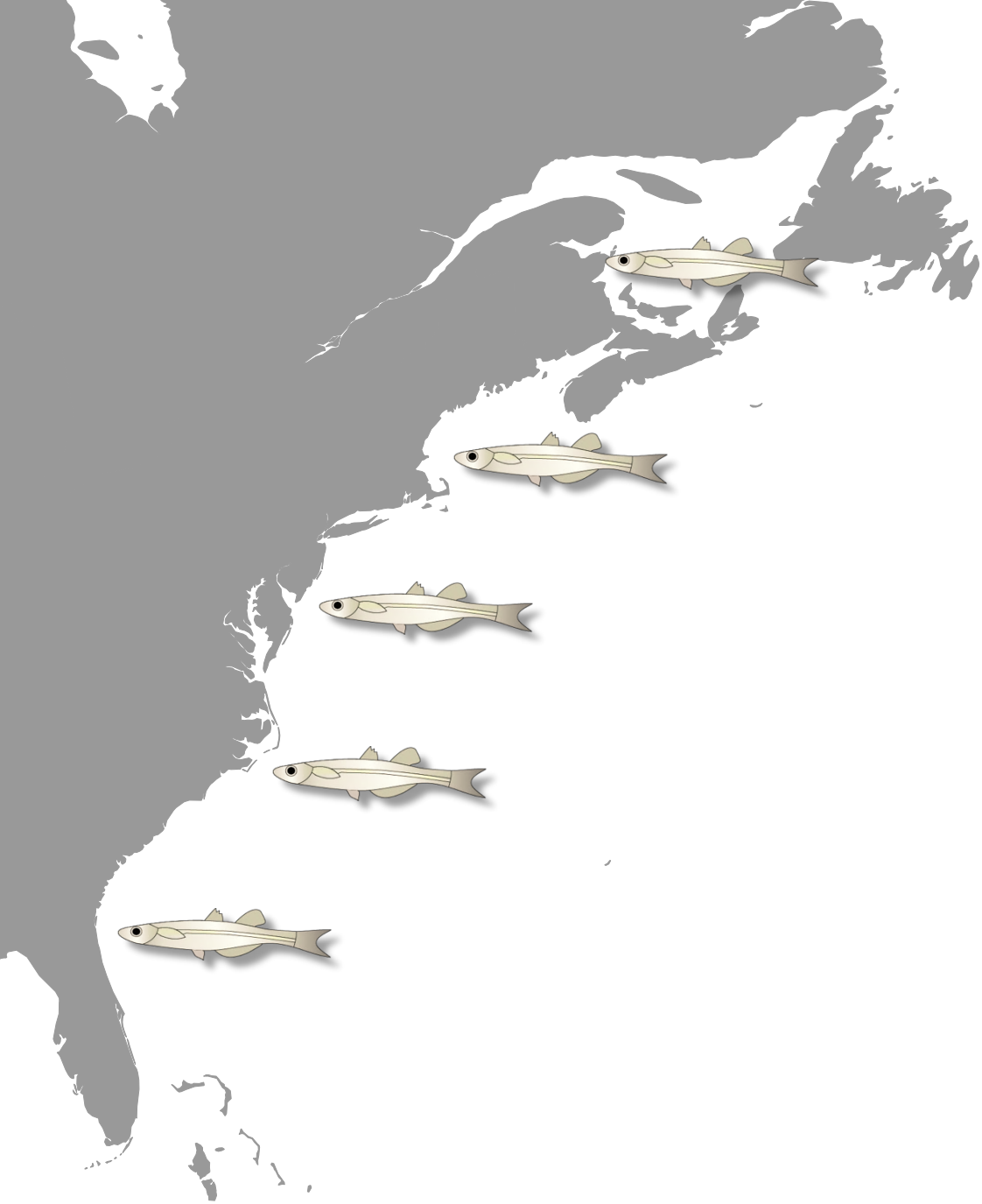


Wild fish all roughly the same size

One of the world's steepest thermal gradients



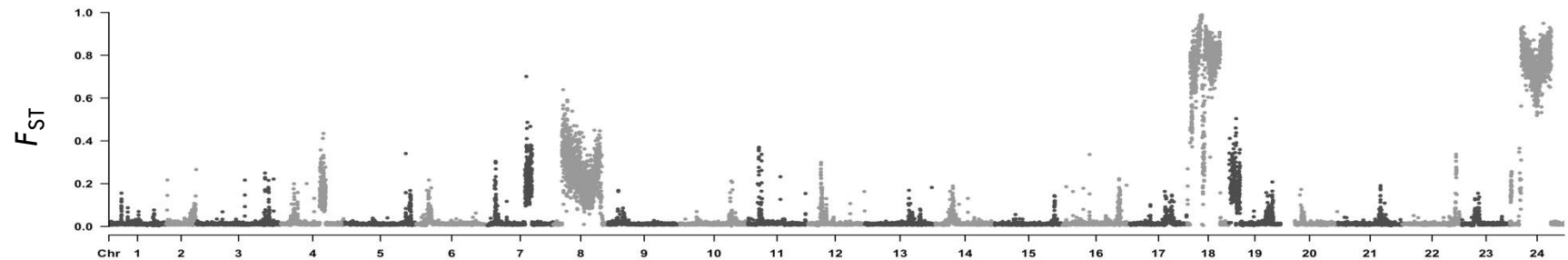
Common garden experiments



Common garden experiments



Whole genome sequencing

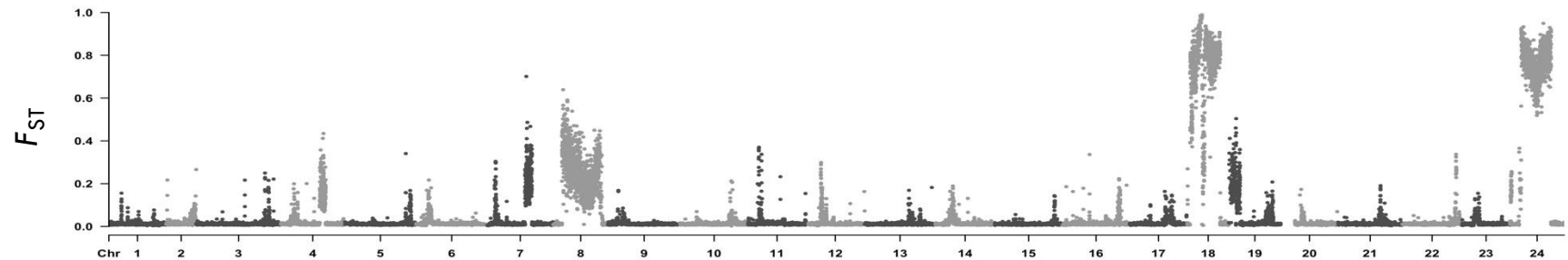


Striking cryptic
diversity
uncovered

Common garden experiments



Whole genome sequencing



Adaptive divergence in Atlantic cod in the Gulf of Maine

1. Use genomic analyses to characterize structure of active cod spawning populations in US waters
2. Identify adaptive genetic differences among populations

Collaborators



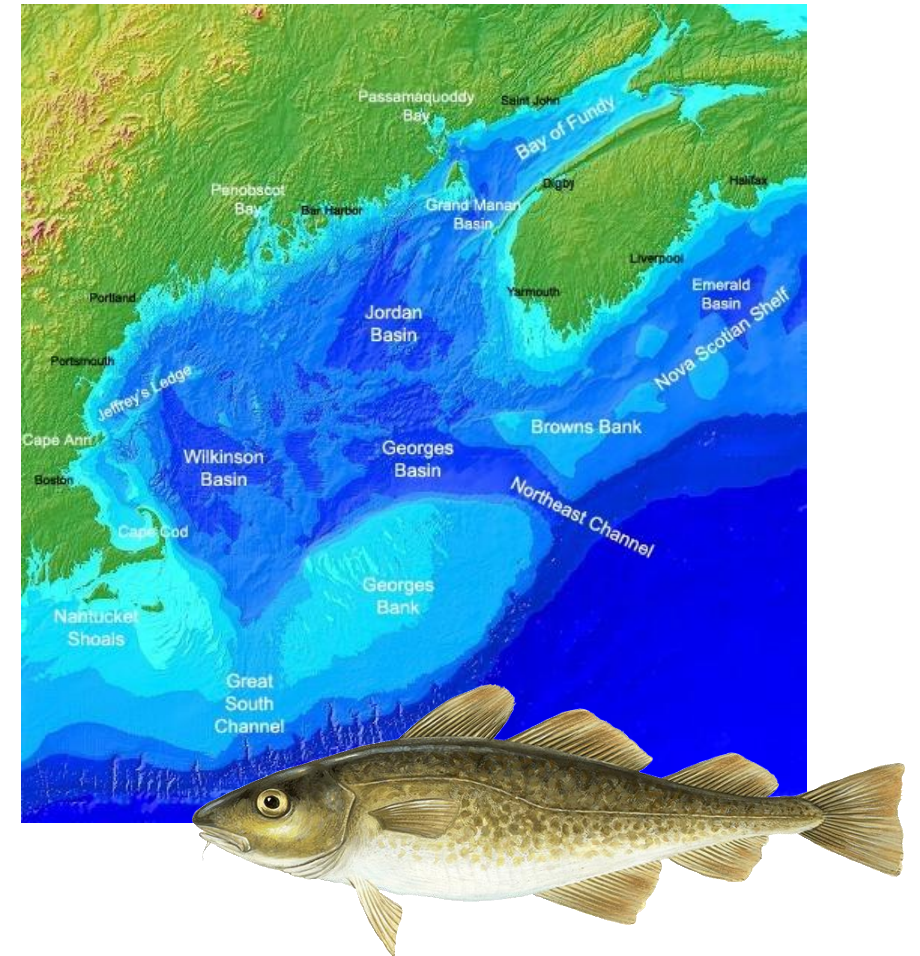
Adrienne Kovach



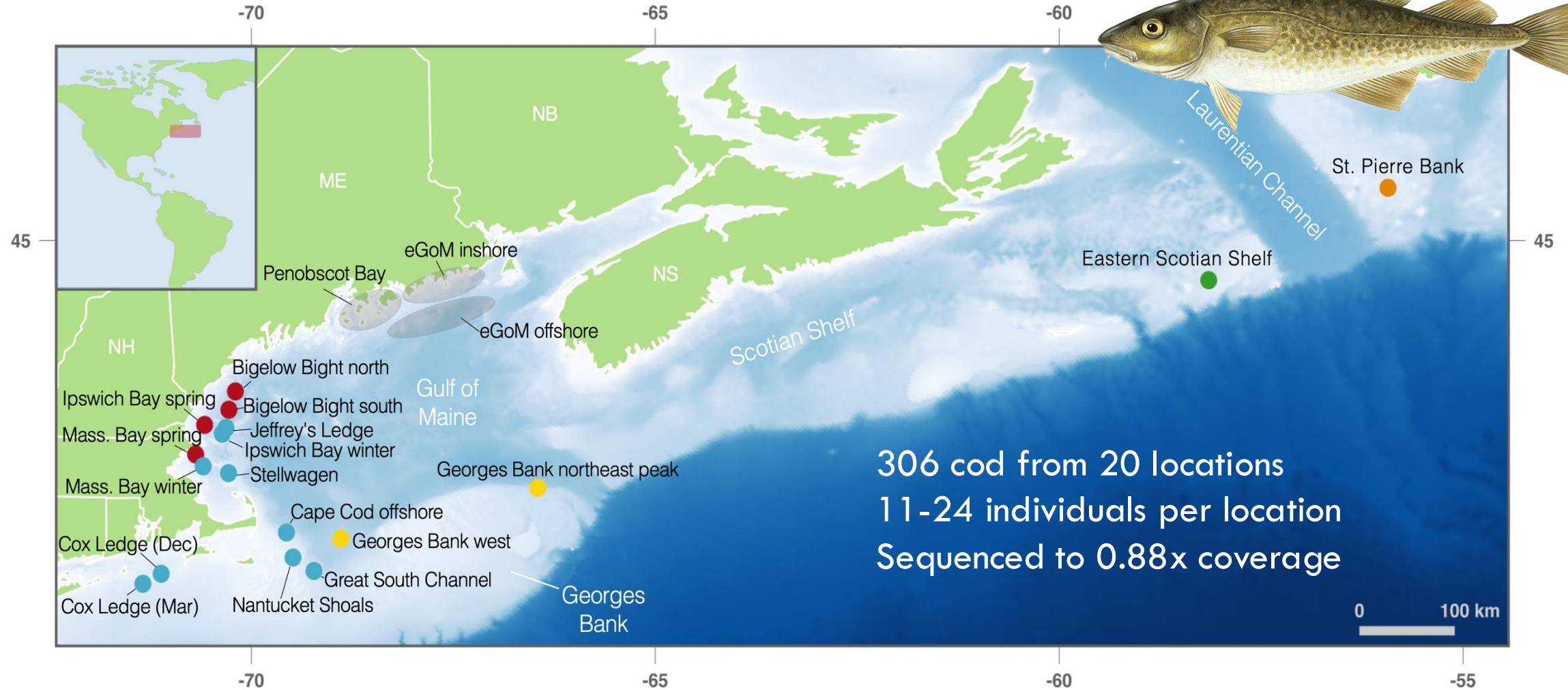
Gemma Clucas



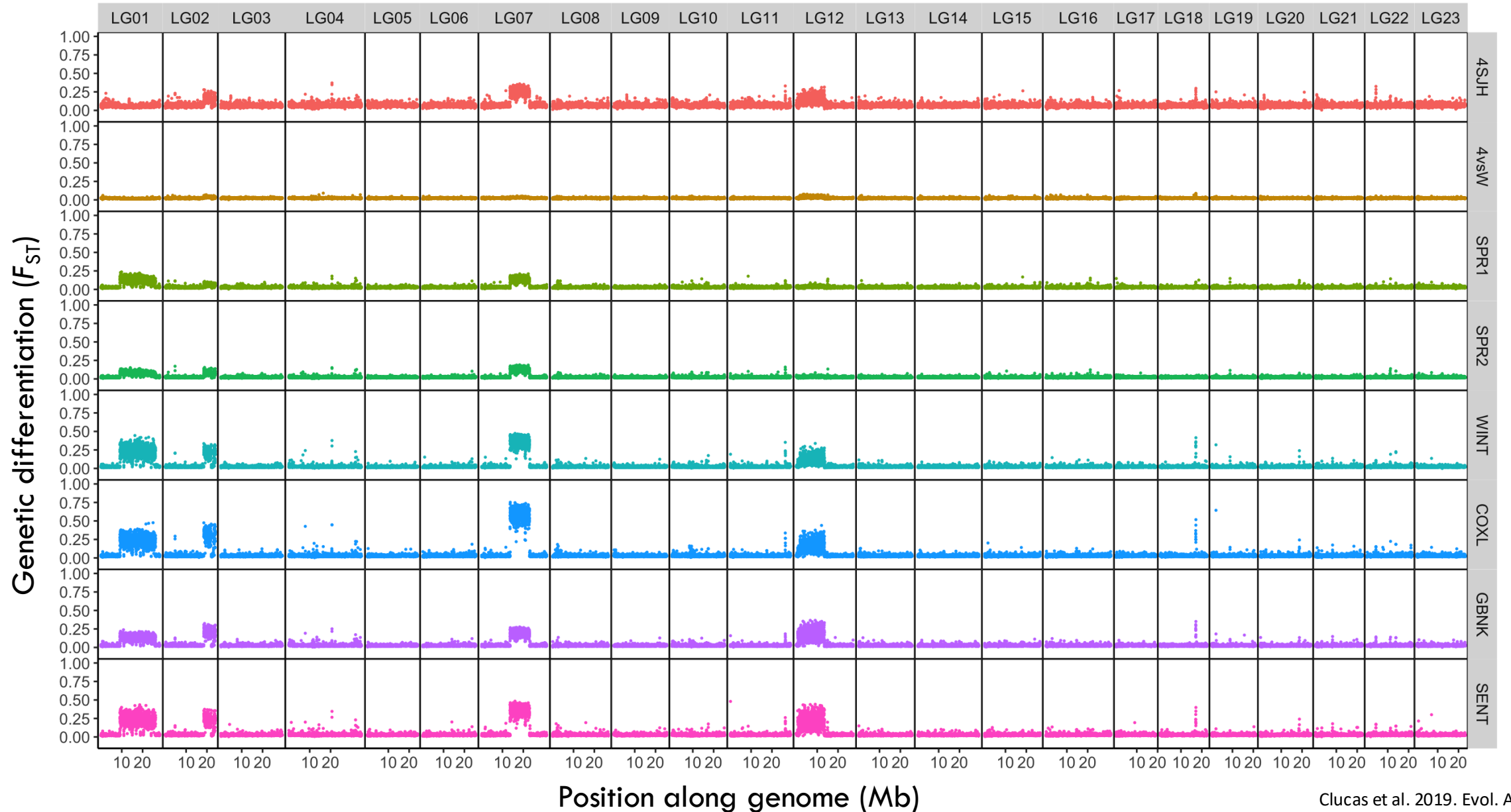
Nicolas Lou



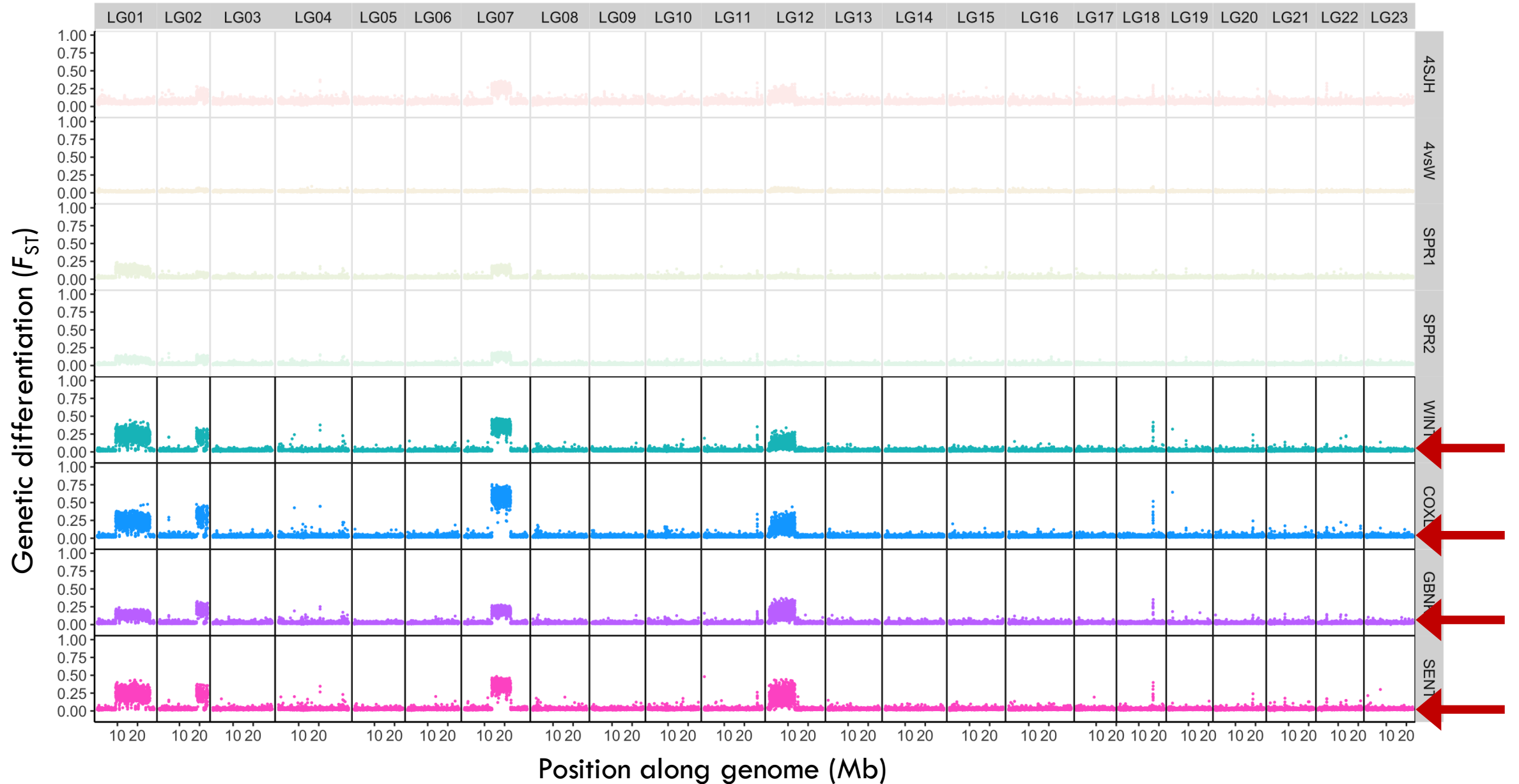
Sampling — actively spawning fish



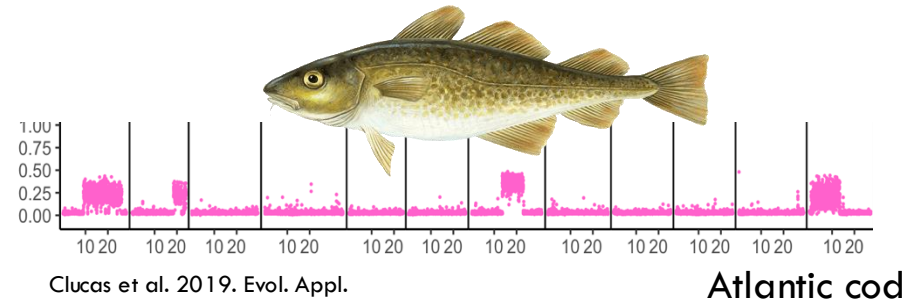
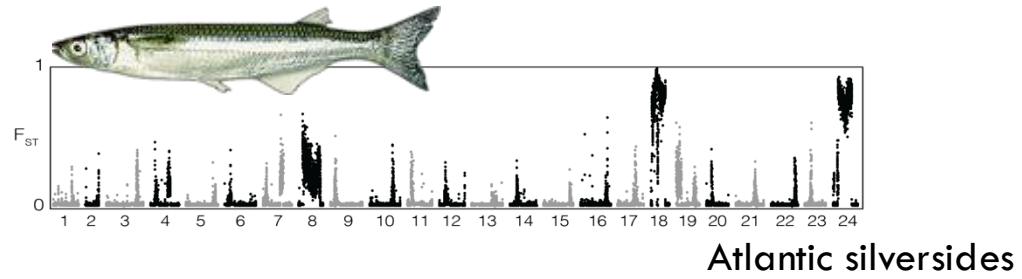
Pairwise comparisons to St. Pierre Bank, Canada (15 kb windows)



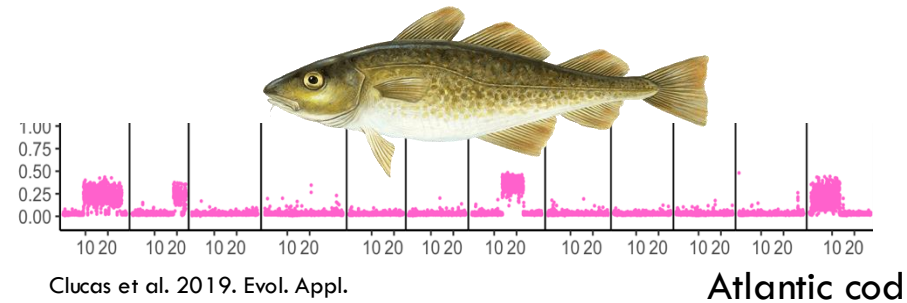
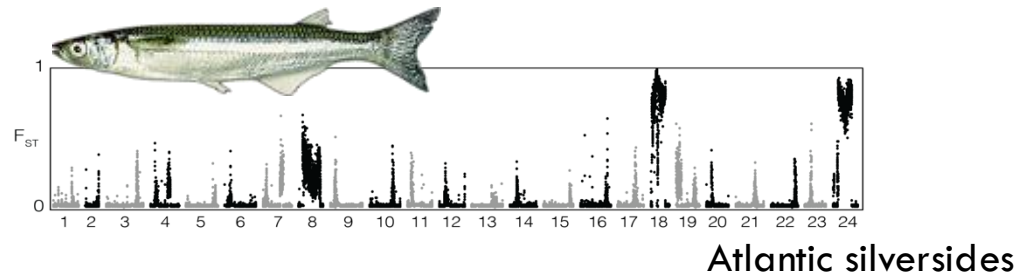
Pairwise comparisons to St. Pierre Bank, Canada (15 kb windows)



Chromosomal inversions everywhere!



Chromosomal inversions everywhere!



Clucas et al. 2019. *Evol. Appl.*

Locally Adaptive Inversions Modulate Genetic Variation at Different Geographic Scales in a Seaweed Fly

Claire Mérot,^{*1} Emma L. Berdan,² Hugo Cayuela,^{1,3} Haig Djambazian,⁴ Anne-Laure Ferchaud,¹ Ragoussis,⁴ Maren Wellenreuther,^{5,6} and Louis Bernatchez¹

¹Centre de Recherche en Biologie et des Systèmes (IBIS), Université Laval, Québec, Canada
²Department of Biology, Science for Life Laboratory, Stockholm University, Stockholm, Sweden
³Department of Biology, University of Lausanne, Lausanne, Switzerland
⁴Department of Biology, University of Toronto, Toronto, Canada
⁵Department of Biology, Nelson, Nelson, New Zealand
⁶Department of Biology, University of Auckland, Auckland, New Zealand

Received: 29 September 2021 | Revised: 25 October 2021 | Accepted: 26 October 2021
 DOI: 10.1111/mec.16258

NEWS AND VIEWS
 Perspective

MOLECULAR ECOLOGY | WILEY

Ready on arrival: Standing variation at a chromosomal inversion contributes to rapid adaptation in an invasive marine crab

Joshua A. Thia

Bio21 Institute, School of BioSciences, University of Melbourne, Victoria, Australia

Correspondence
 Joshua A Thia, Bio21 Institute, School of BioSciences, University of Melbourne, Victoria, Australia.
 Emails: joshua.thia@unimelb.edu.au; josh.thia@ive.com

In this issue of *Molecular Ecology*, Joshua A. Thia et al. show that standing variation at a chromosomal inversion contributes to rapid adaptation in an invasive marine crab. The authors demonstrate that the inversion is associated with a large effect on the fitness of the crab, and that the inversion is maintained in the population despite low diversity in the region. This finding highlights the importance of standing variation in adaptation and the role of chromosomal inversions in maintaining genetic diversity in populations.

Molecular Ecology

ORIGINAL ARTICLE

Inversions Dominate Evolution in the European Sardine (*Sardina pilchardus*) Amid Strong Gene Flow

Stephen J. Sabatino^{1,2} | M. Pilar Cabezas^{3,4} | Paulo Pereira^{1,2} | Susana Garrido^{5,6} | António M. Santos^{1,2,7} | Miguel Carneiro^{1,2} | Paulo T. Santos^{8,7} | Bruno Louro⁸ | Cymon J. Cox⁸ | Adelino V. M. Canário⁸ | Ana Verissimo^{1,2}

¹CIBIO-InBIO, Research Center in Biodiversity and Genetic Resources, University of Porto, Vairão, Portugal | ²BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Vairão, Portugal | ³Centre of Molecular and Environmental Biology (CBMA) and ARNET—Aquatic Research Network, Department of Biology, University of Minho, Gualtar Campus, Braga, Portugal | ⁴Institute of Science and Innovation for Bio-Sustainability (IB-S), University of Minho, Gualtar Campus, Braga, Portugal | ⁵Portuguese Institute for sea and Atmosphere (IPMA), Lisbon, Portugal | ⁶Marine and Environmental Sciences Centre (MARE), Faculty of Sciences, University of Lisbon, Lisbon, Portugal | ⁷Faculty of Sciences, University of Porto, Porto, Portugal | ⁸Algarve Centre for Marine Sciences (CCMAR/CIMAR), Gambelas Campus, University of Algarve, Faro, Portugal

Correspondence: Stephen J. Sabatino (sjsabatino@gmail.com)

Received: 27 January 2025 | Revised: 7 May 2025 | Accepted: 30 June 2025

Handling Editor: Nick Hamilton Barton

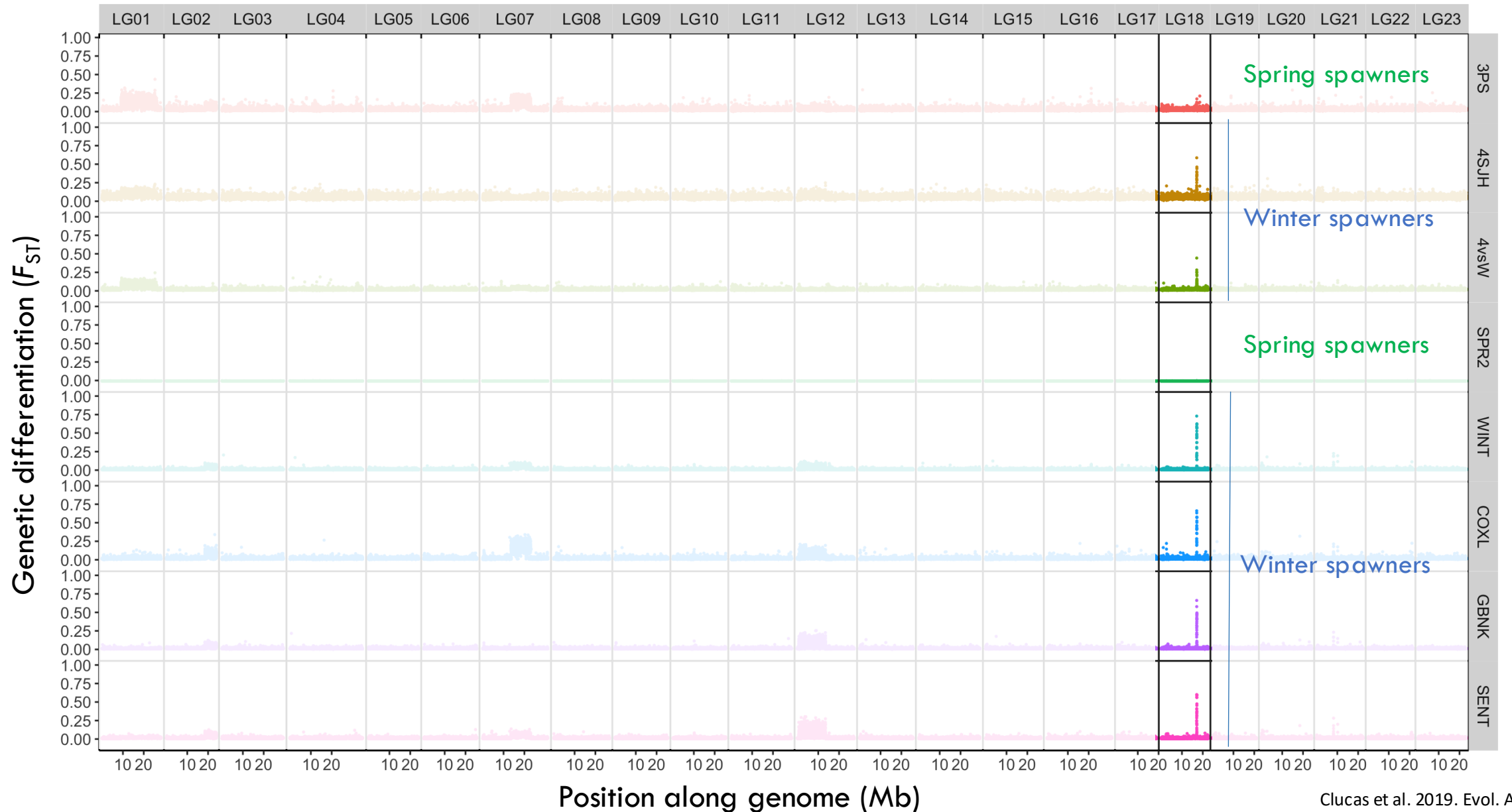
Funding: This work was supported by the Portuguese national funds from FCT—Fundação para a Ciência e Tecnologia (LA/P/0101/2020, UIDB/04326/2021, UIDP/04326/2022, 10.54499/2022.02937/CEECIND/CP1730/CT0011, CEECINST/00014/2018/CP1512/CT0002) and European Maritime and Fisheries Fund (EMFF) (MAR-01.04.02-FEAMP-0024, MAR-111.4.1-FEAMPA-00001).

Keywords: adaptation | climate change | life-history evolution | population genetics—empirical | wildlife management

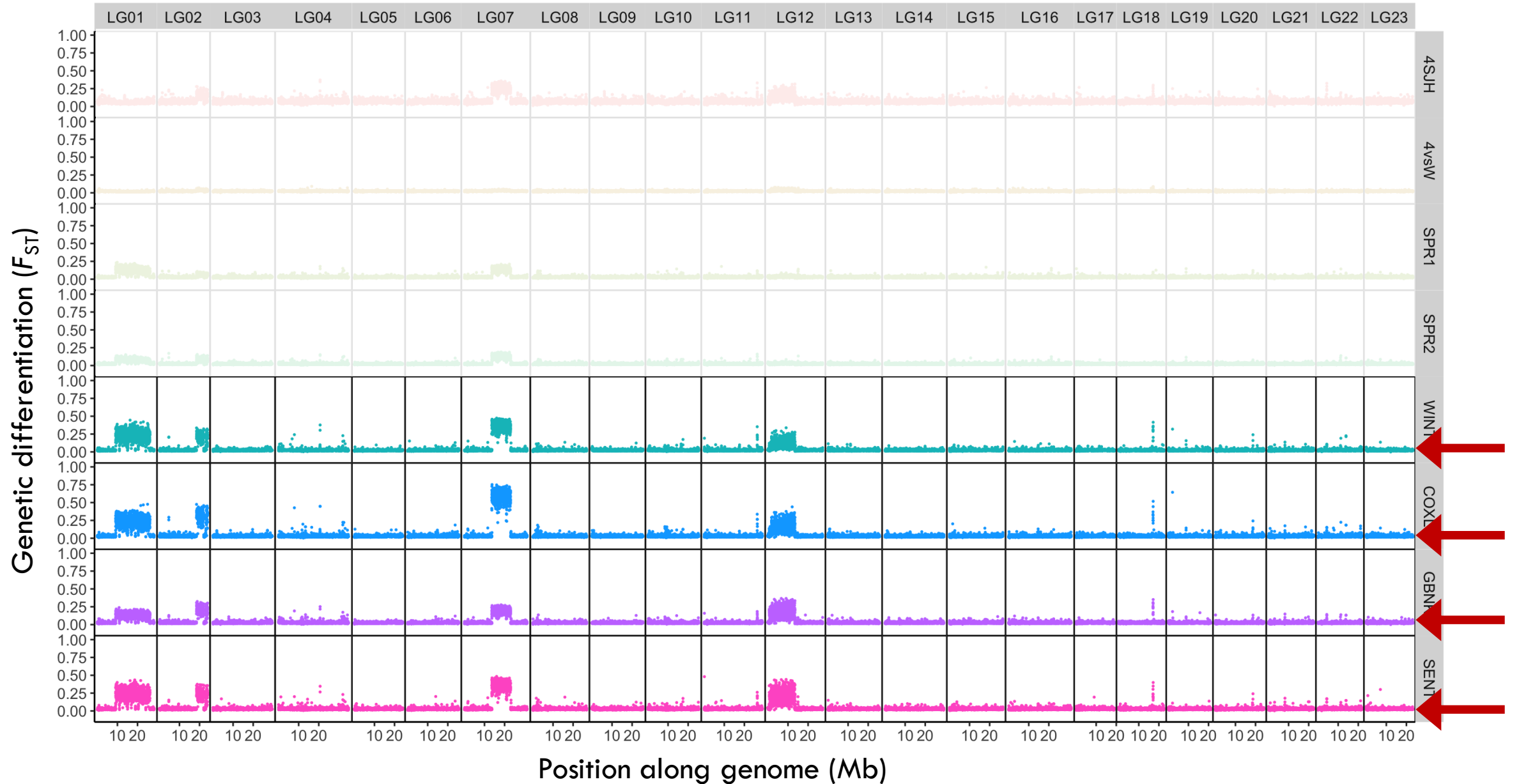
ABSTRACT

Inversions can play key roles in the genetic architecture of adaptation, but the scale of their effects across different species remains poorly understood. Here, we use whole-genome sequencing and demographic modelling to investigate the influence of inversions on the population genomics of the r-selected European sardine (*Sardina pilchardus*). Allele frequency differences from millions of SNPs across 34 populations spanning the species' range were analysed. Genomic scans identified several extreme outlier regions overlapping large inversions (29–52 Mbp), collectively representing over half the genome. Our findings suggest these inversions correlate with locally adapted life-history strategies. First, SNPs within outlier regions containing inversions exhibited striking allele frequency differences between Atlantic and Mediterranean sardines, which differ in key adaptive life-history traits. In the Atlantic, inversion allele frequencies varied latitudinally, while in the Mediterranean, they shifted longitudinally.

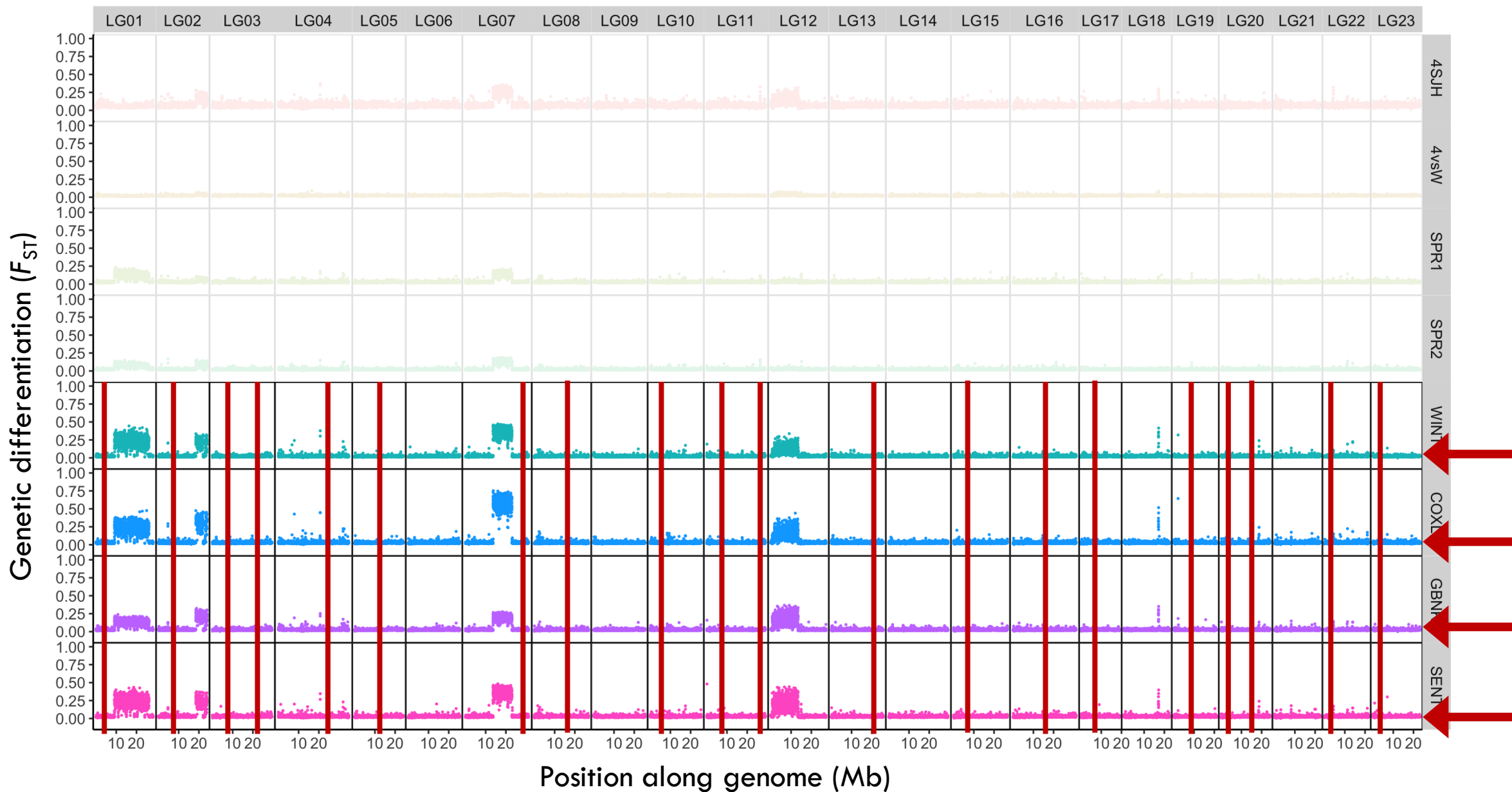
Pairwise comparisons to the wGoM spring spawners (5 kb windows)



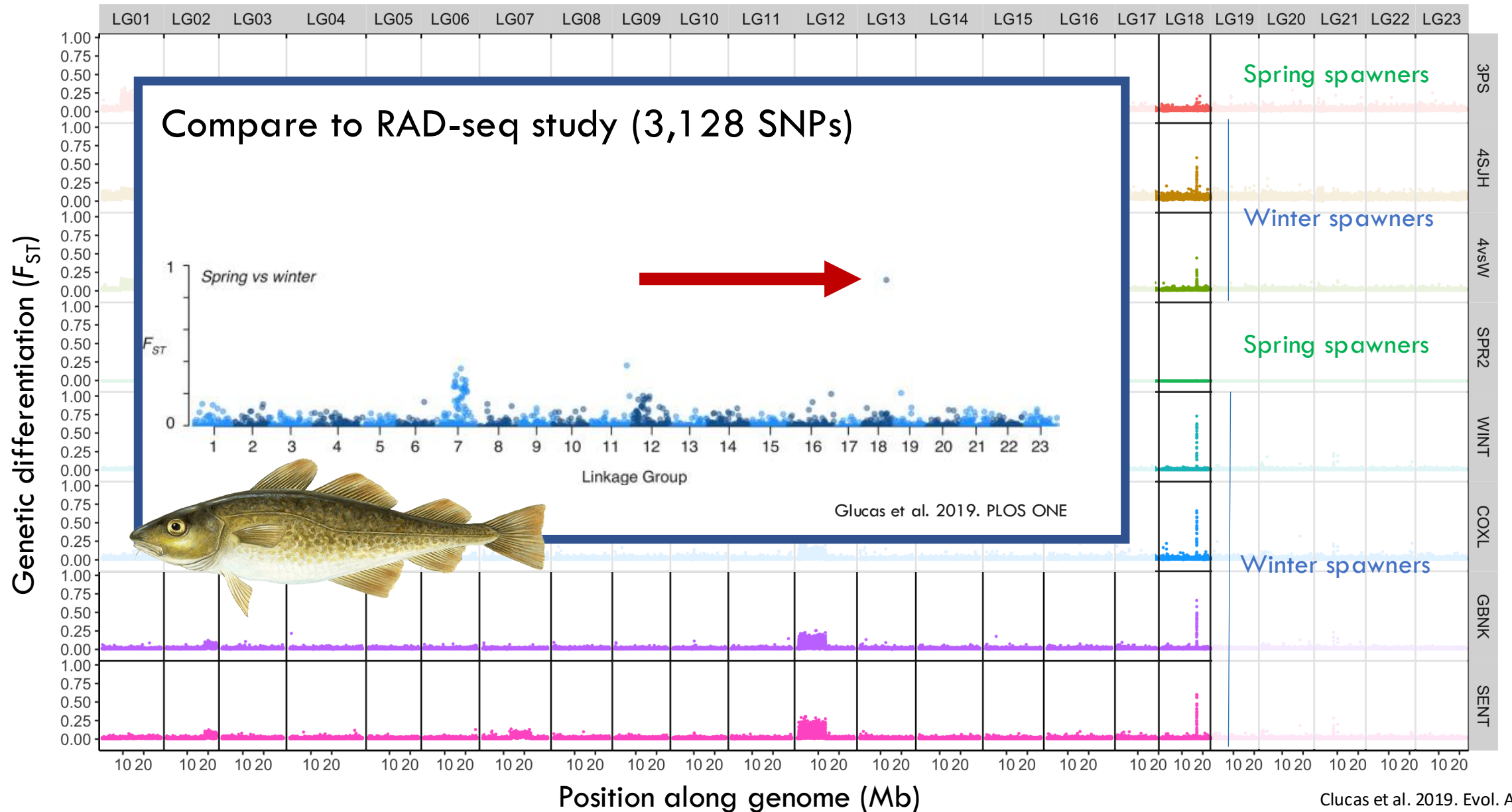
Pairwise comparisons to St. Pierre Bank, Canada (15 kb windows)



Sampling with 20 random markers could miss differentiated regions



Pairwise comparisons to the wGoM spring spawners (5 kb windows)





Atlantic Cod Stock Structure Working Group

Atlantic cod have been a mainstay of coastal economies around the Northern Atlantic for hundred of years. While cod in the North Atlantic are similar to one another in many ways, the population is actually made up of discrete stock components distinguished by differences in characteristics like size, growth rate, and even genetics. While it is well known that there are multiple cod stocks, fully identifying those stocks on a biological and ecological scale has proven elusive.

In the Northwest Atlantic, both the United States and Canada have long histories of harvesting and trading these fish, and each country is responsible for managing cod fisheries within their waters. While some cod stocks stay within a country's waters, some also straddle the international boundaries as well as those set up for management purposes. Understanding how the stocks that constitute the cod population are the same and how they are different is a key part of improving how we manage and use these fish and what can be done to restore their numbers.

Researchers in the U.S. and Canada regularly monitor and study cod. These scientists may be working for government, or universities, or non-profit organizations. To compose a complete and detailed picture of the cod stocks means mounting a truly international and interdisciplinary effort to bring together all of the data, studies, and knowledge about stock components.

In February 2018, a working group was convened to take on this challenge. Organizers plan a two-year effort: one year to synthesize information about stock structure and one year to look into the management implications of their findings. The NOAA Northeast Fisheries Science Center is hosting the working group and updates on their activities will be posted here.



Photo credit: [Staticflickr.com](#)

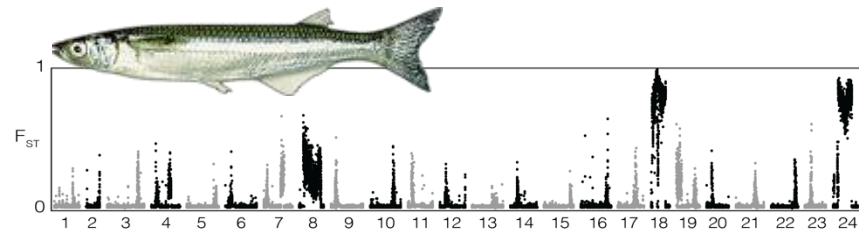
More on Atlantic cod

- [Consumer guide](#)
- [2017 U.S. assessment update](#)
- [2018 Transboundary Resource Assessment Committee Meeting](#)

<p>Contacts</p> <ul style="list-style-type: none"> • Co-Chairs 	<p>Meetings</p> <ul style="list-style-type: none"> • Upcoming 	<p>Updates</p> <ul style="list-style-type: none"> • June 19
--	---	---

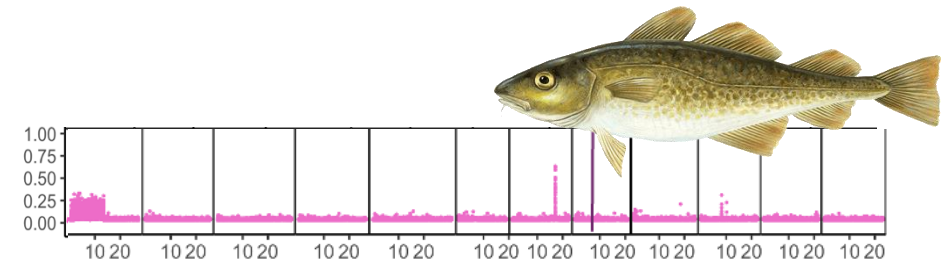
Differentiation often limited to parts of the genome

Strong adaptation despite homogeneous genomes



Atlantic silversides

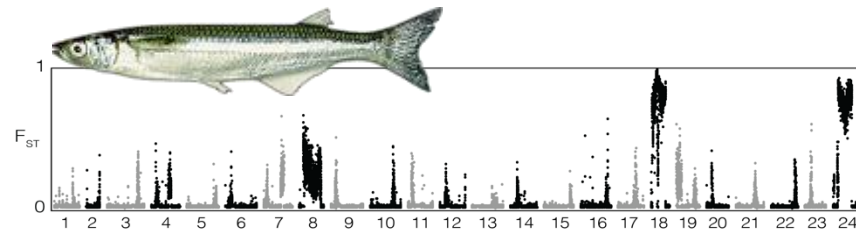
Differences confined to small genomic regions



Atlantic cod

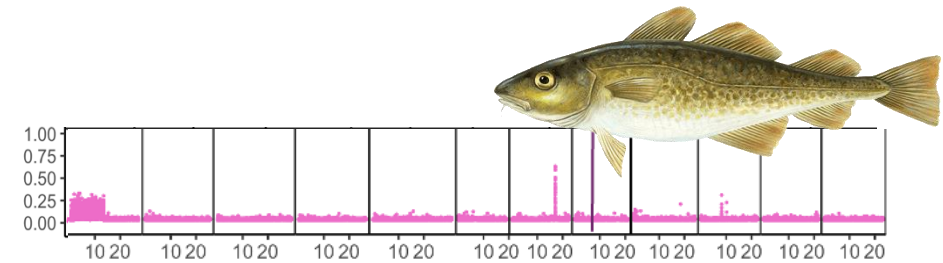
Differentiation often limited to parts of the genome

Strong adaptation despite homogeneous genomes



Atlantic silversides

Differences confined to small genomic regions



Atlantic cod

Surprising mixture of lineages can also occur

Sea cucumbers in the Galapagos islands

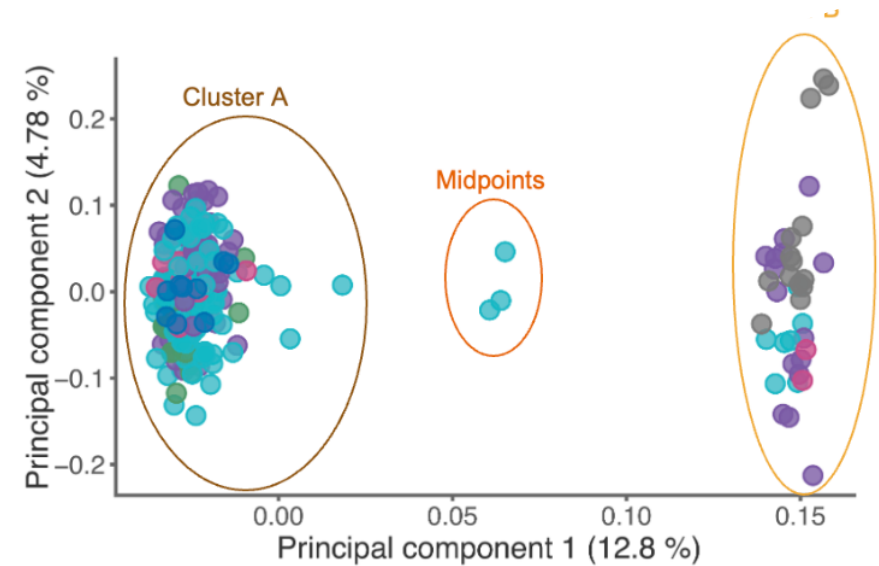
Mixture of local and mainland-connected lineage



Jaime Ortiz-Pachar

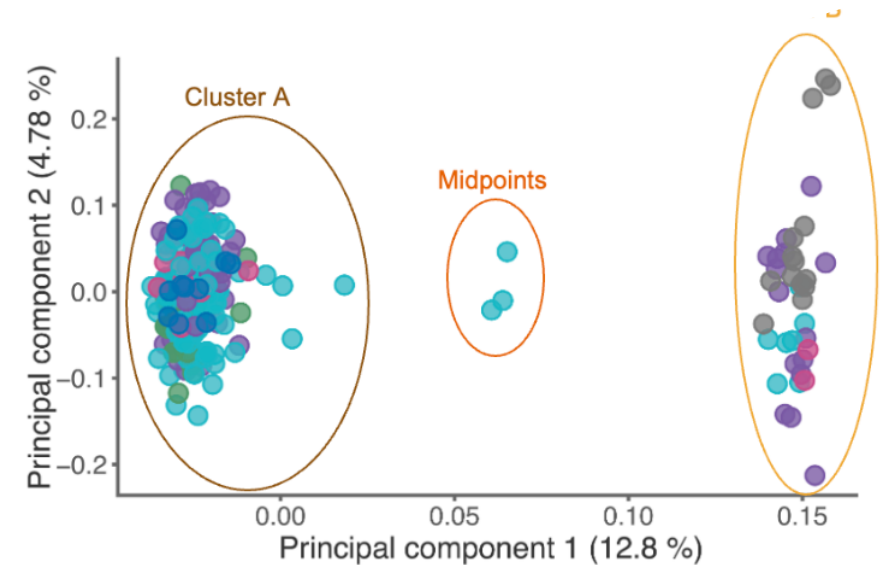
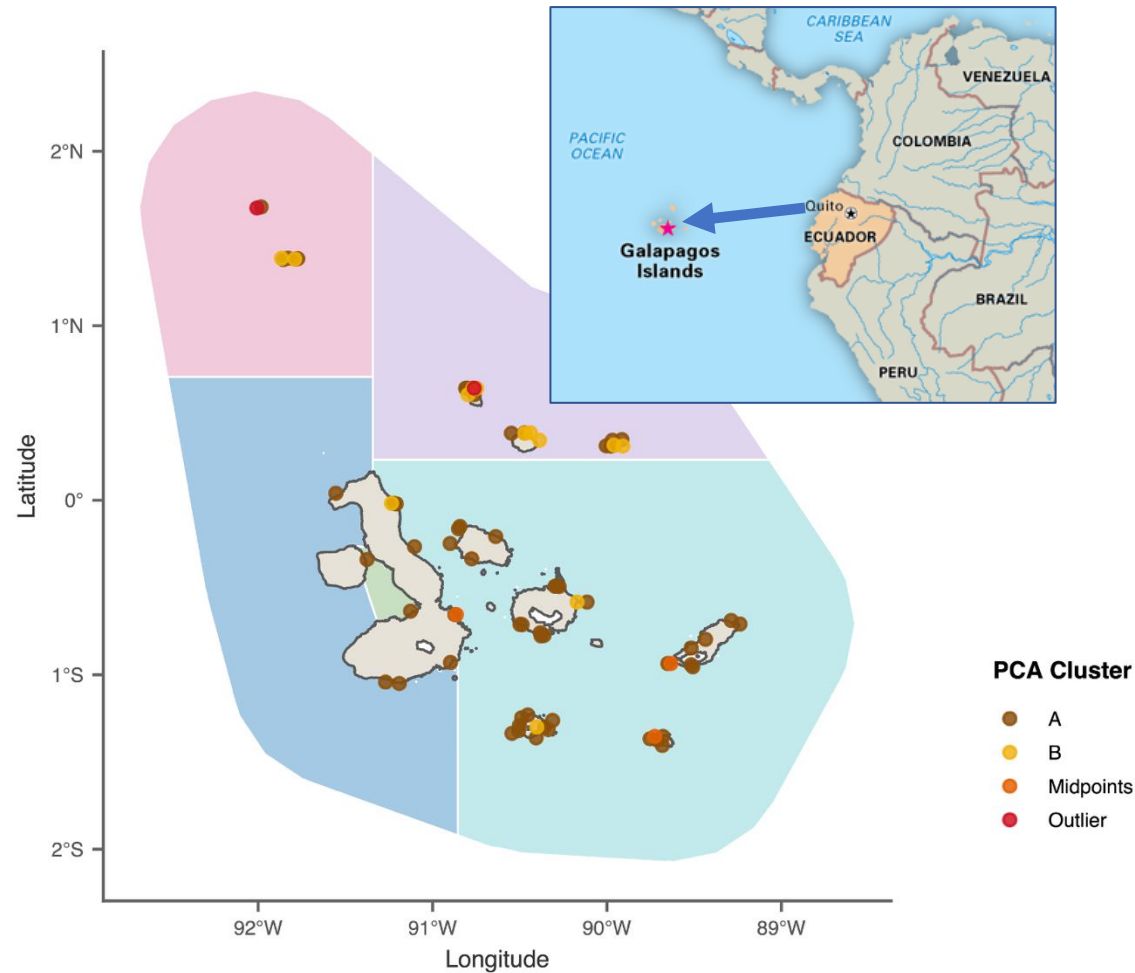
Sea cucumbers in the Galapagos islands

Mixture of local and mainland-connected lineage



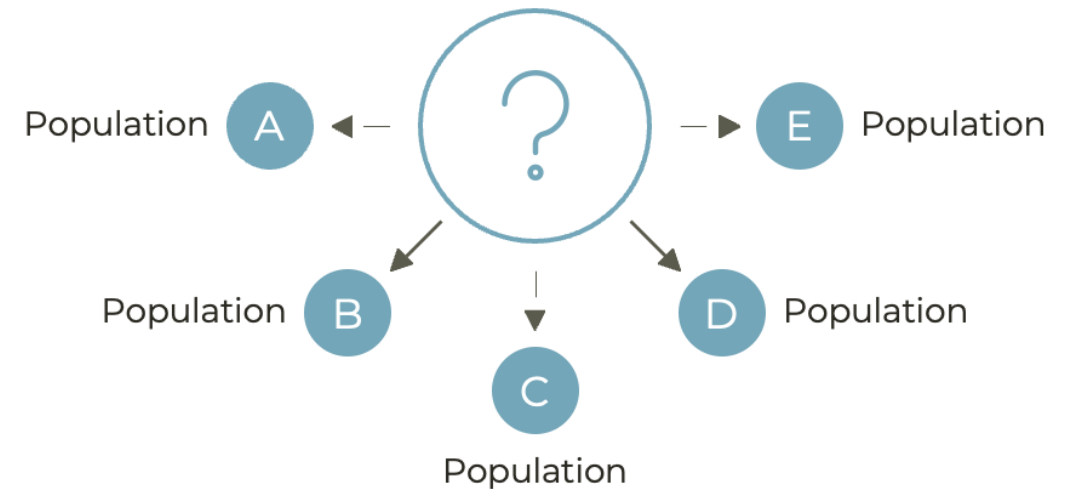
Sea cucumbers in the Galapagos islands

Mixture of local and mainland-connected lineage



From hidden structure to measurable mixing

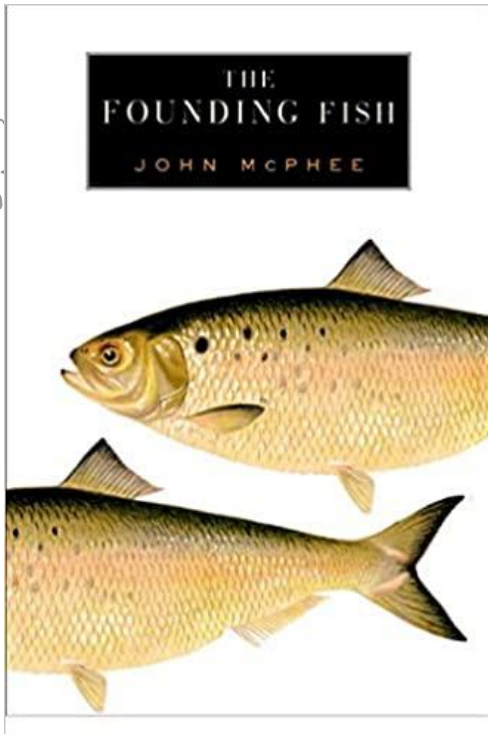
- Subtle and cryptic structure is common
- Management decisions depend on who mixes with whom
 - and when
- We need to assign individuals to their origin



American shad



- Called “the founding” fish because of historical abundance
- Now decimated in many places, but some remaining fisheries

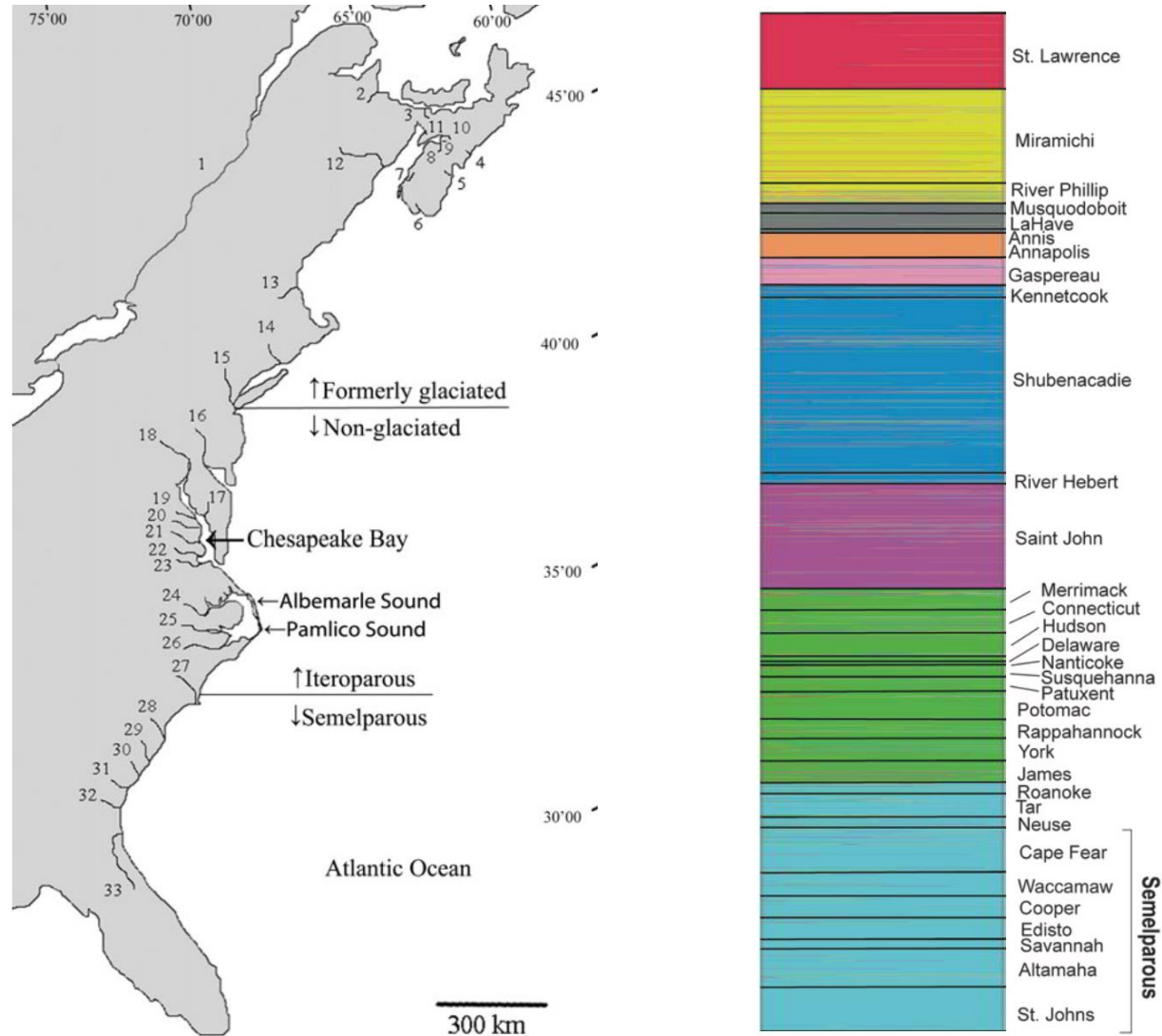


American shad

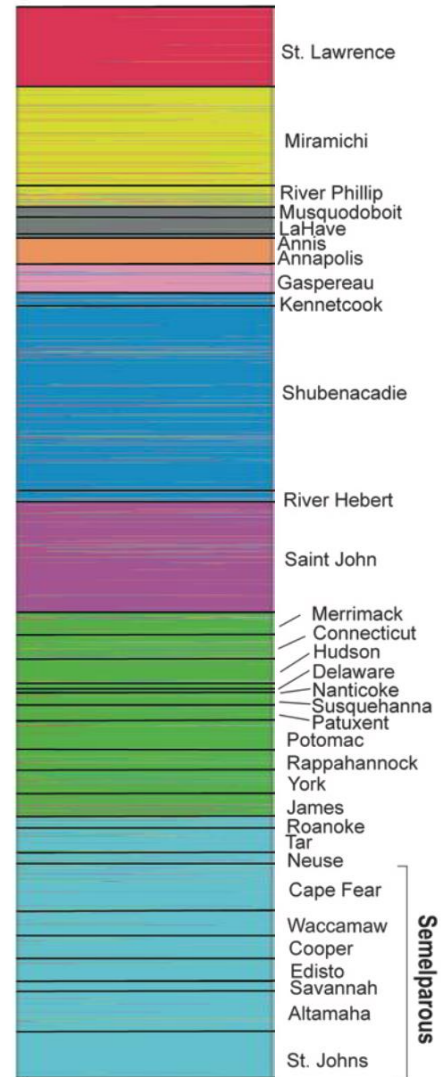
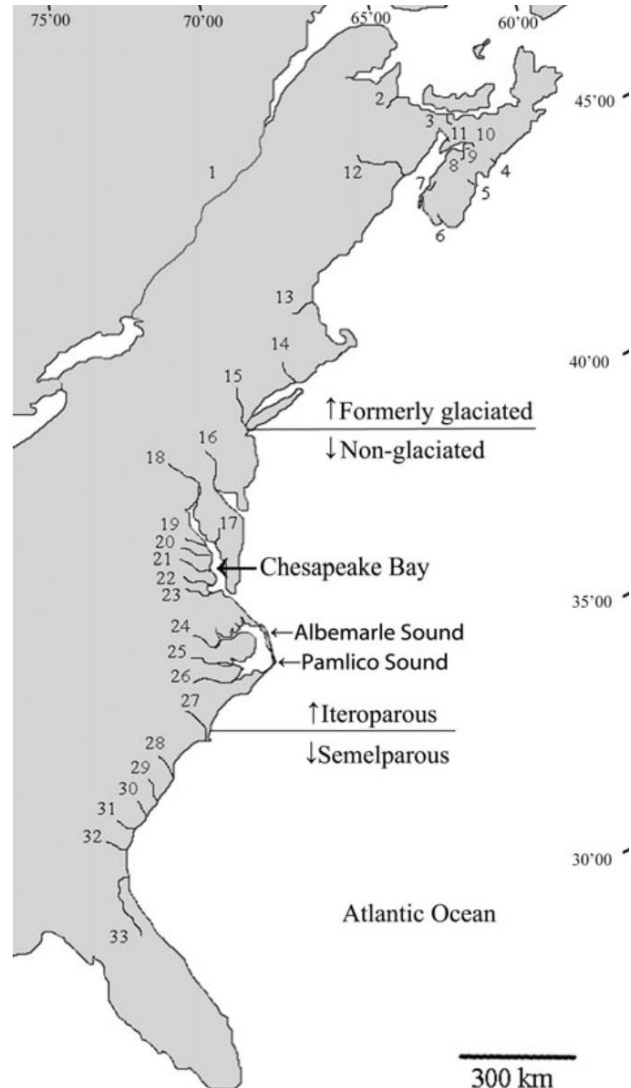
- Called “the founding” fish because of historical abundance
- Now decimated in many places, but some remaining fisheries
- Interest in assigning marine bycatch to population



Limited population differentiation in US waters with microsatellites



Limited population differentiation in US waters with microsatellites

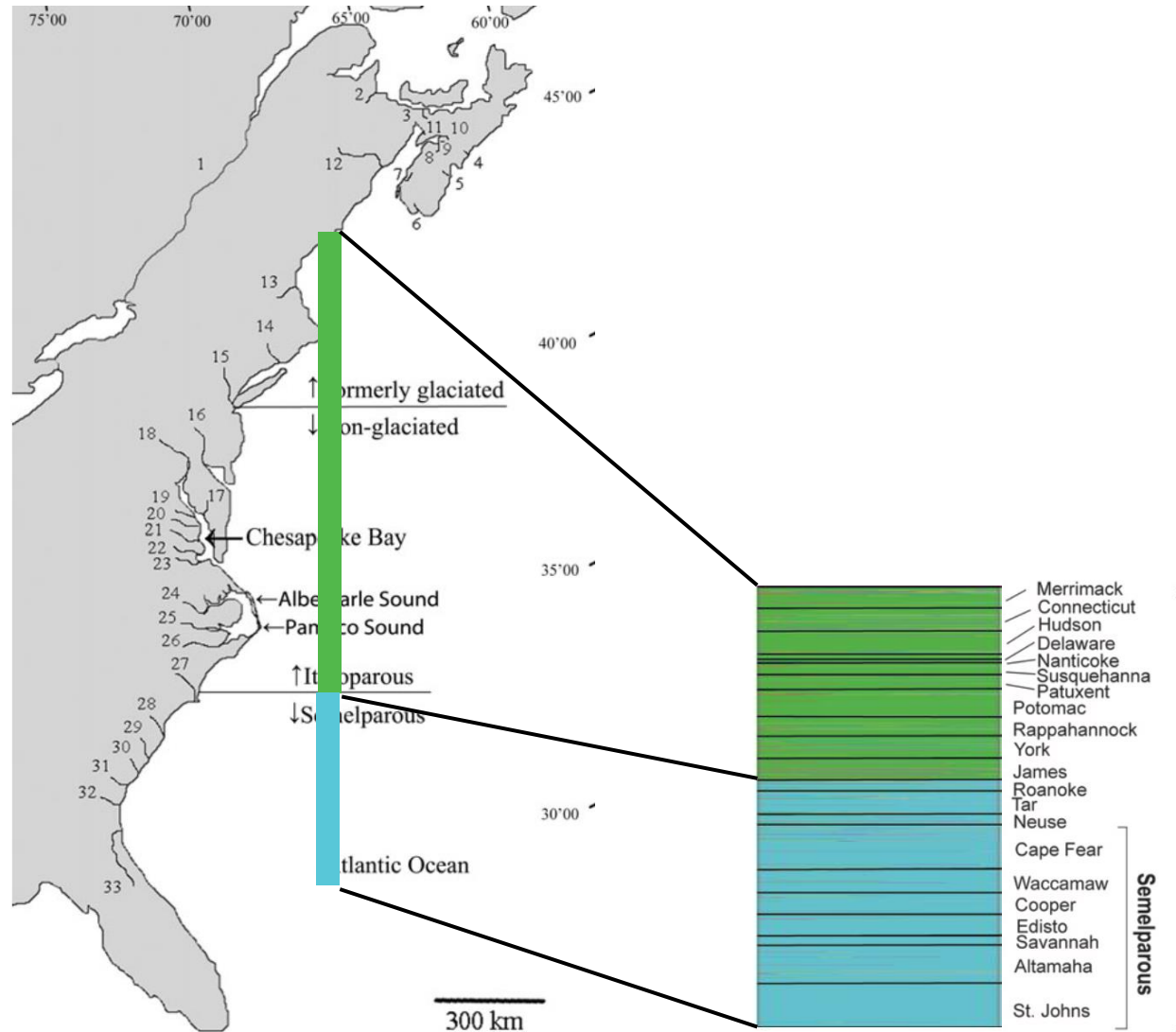


Canadian populations

U.S. populations

Semelparous

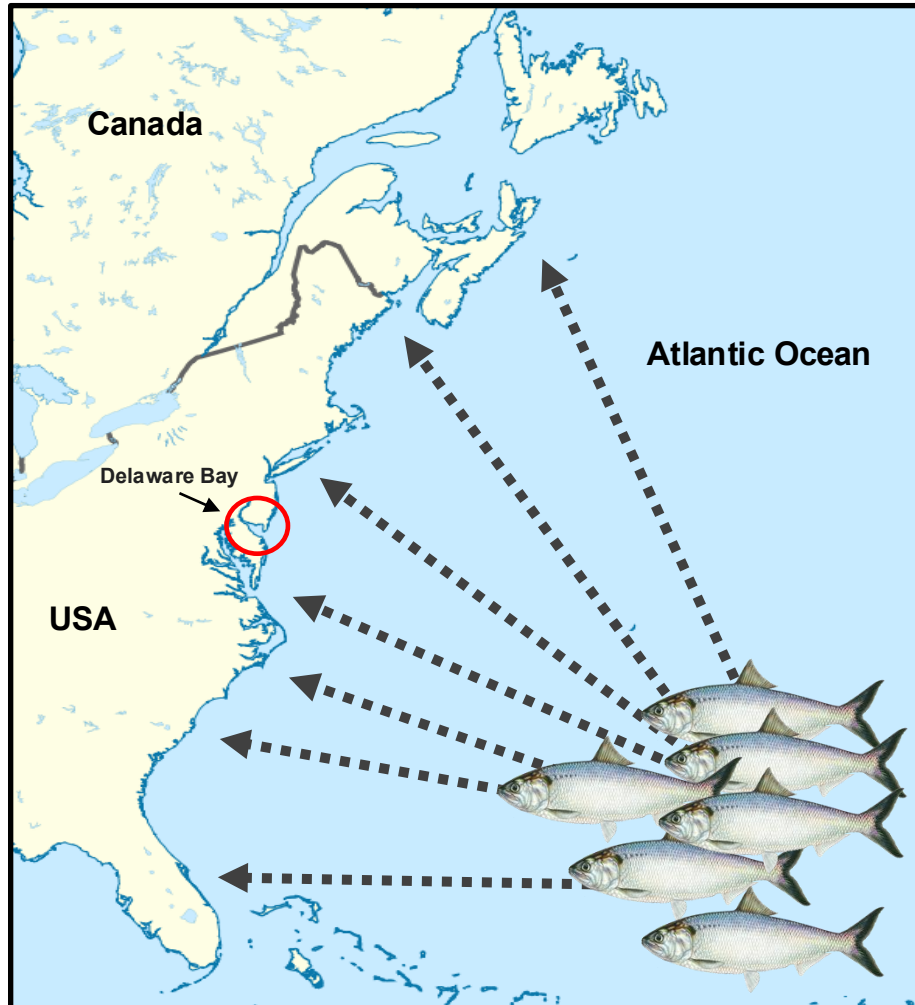
Limited population differentiation in US waters with microsatellites



U.S. populations

Semelparous

Microsatellite assignment

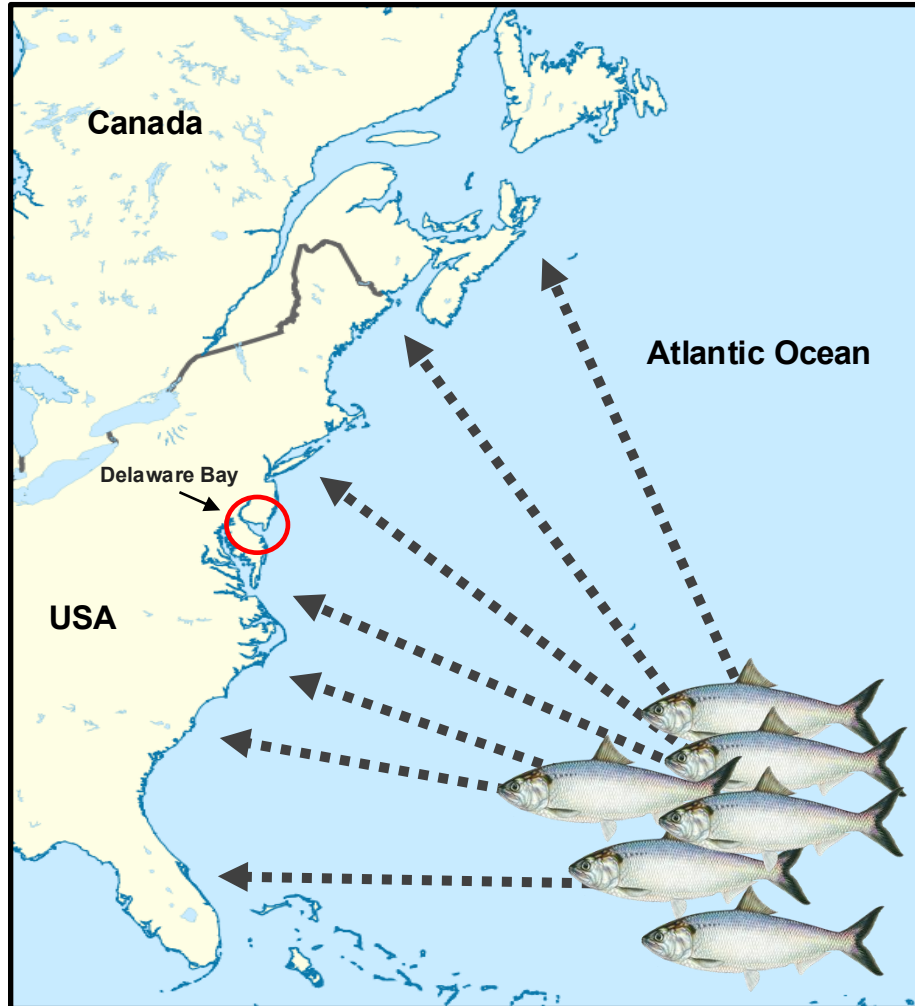


Resolving Contribution of American Shad Mixed Stock Fishery in Lower Delaware River and Bay: 2017-2020 collection evaluation

Meredith L. Bartron and Lauren M. Prasko
USFWS Northeast Fishery Center, Lamar, PA

Collection location	Sample size	Correct assignment
Narraguagus	50	0.88
Penobscot	93	0.37
Androscoggin	48	0.19
Kennebec	24	0.04
Sheepscot	23	0.04
Saco	50	0.10
Merrimack	49	0.10
Connecticut	54	0.15
Hudson	200	0.47
Delaware	198	0.43
Susquehanna	48	0.19
Potomac	49	0.22
Tar	38	0.16
Cooper	48	0.92

Microsatellite assignment



Resolving Contribution of American Shad Mixed Stock Fishery in Lower Delaware River and Bay: 2017-2020 collection evaluation

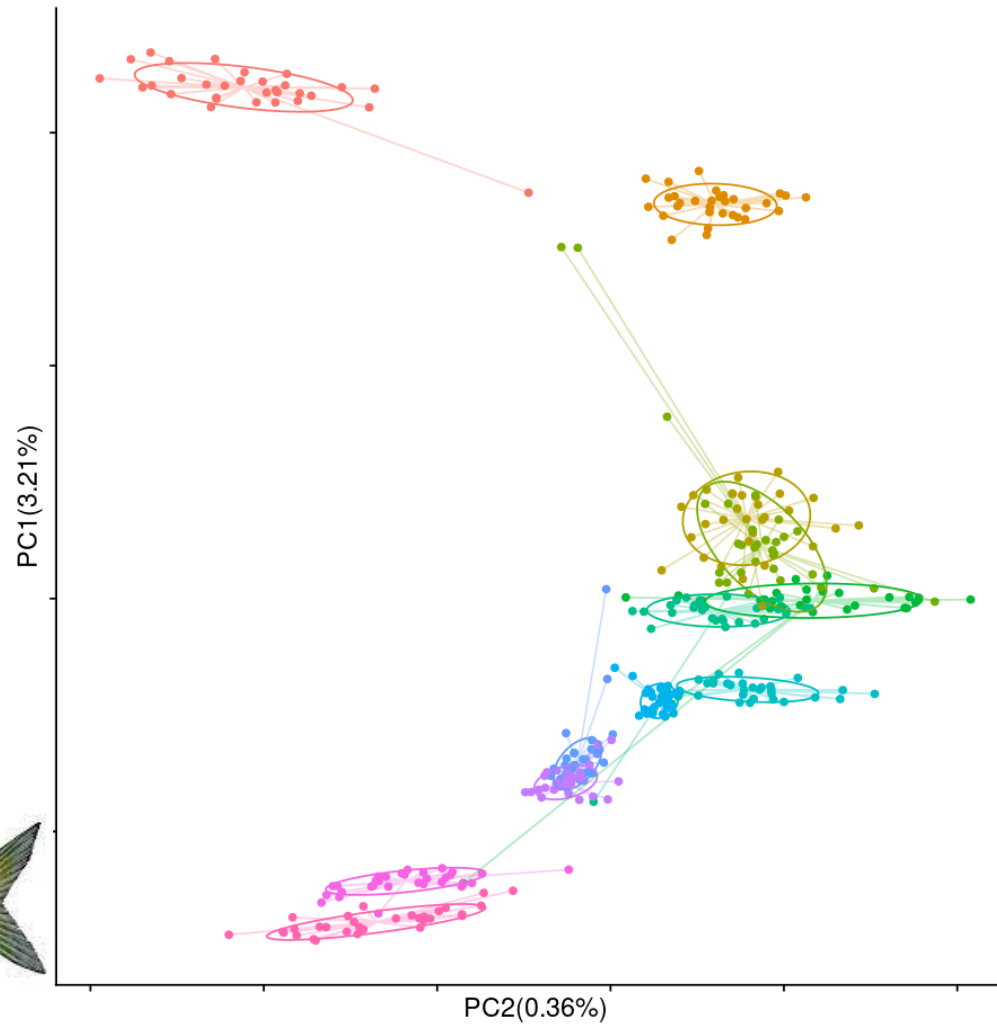
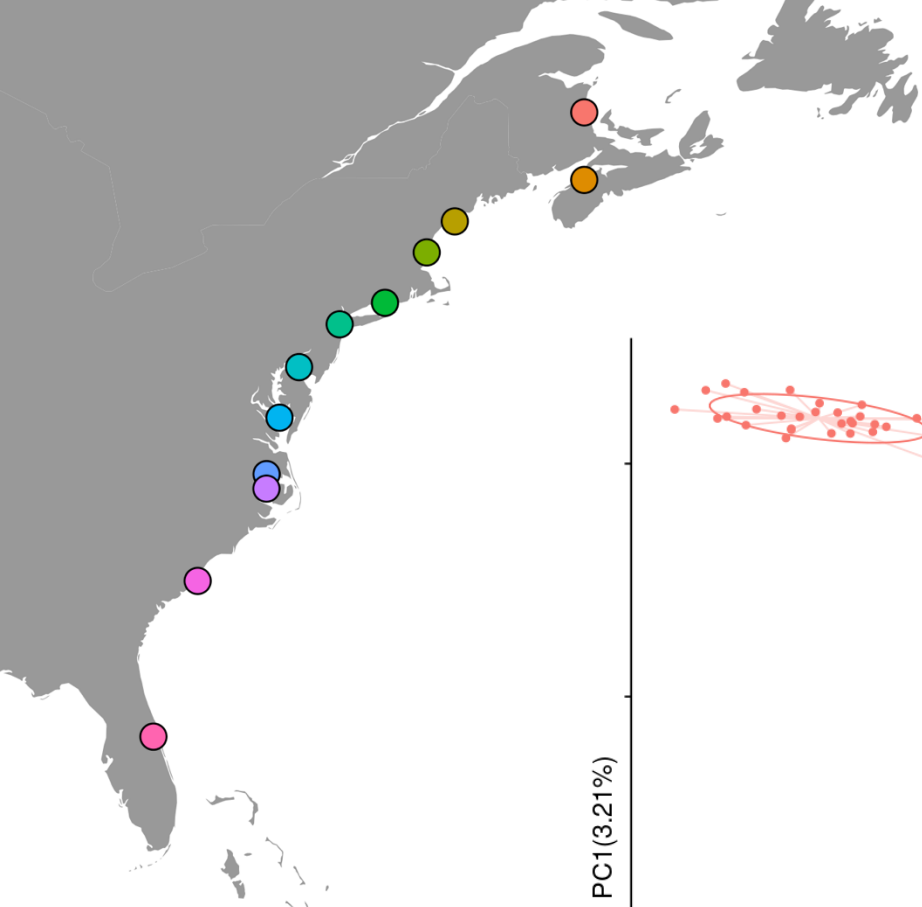
Meredith L. Bartron and Lauren M. Prasko
USFWS Northeast Fishery Center, Lamar, PA

Collection location	Sample size	Correct assignment
Narraguagus	50	0.88
Penobscot	93	0.37
Androscoggin	48	0.19
Kennebec	24	0.04
Sheepscot	23	0.04
Saco	50	0.10
Merrimack	49	0.10
Connecticut	54	0.15
Hudson	200	0.47
Delaware	198	0.43
Susquehanna	48	0.19
Potomac	49	0.22
Tar	38	0.16
Cooper	48	0.92

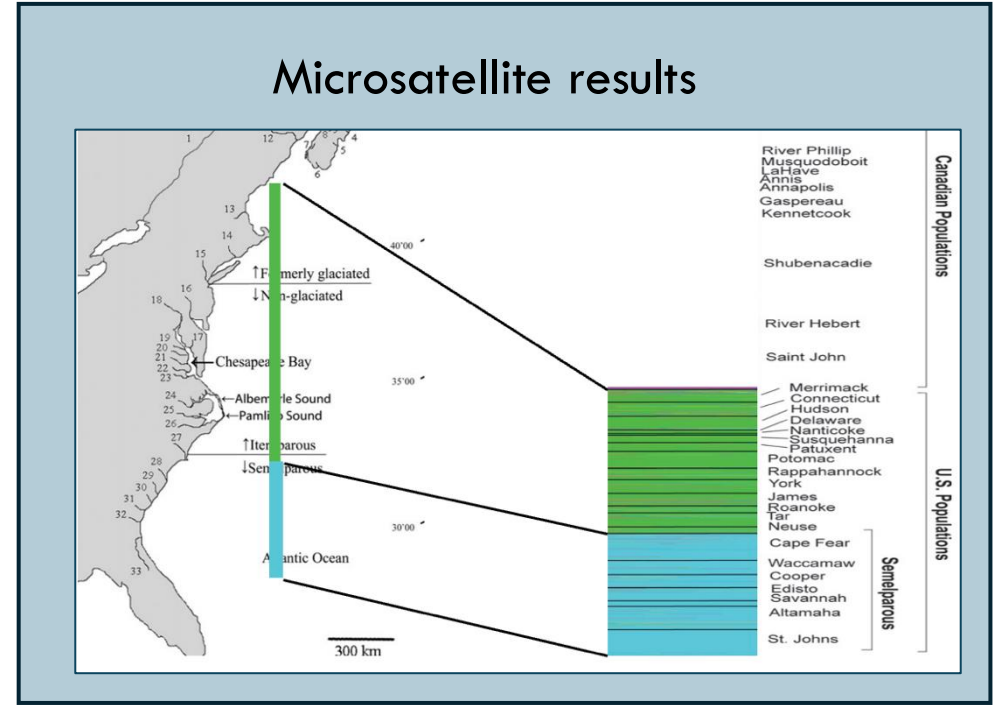
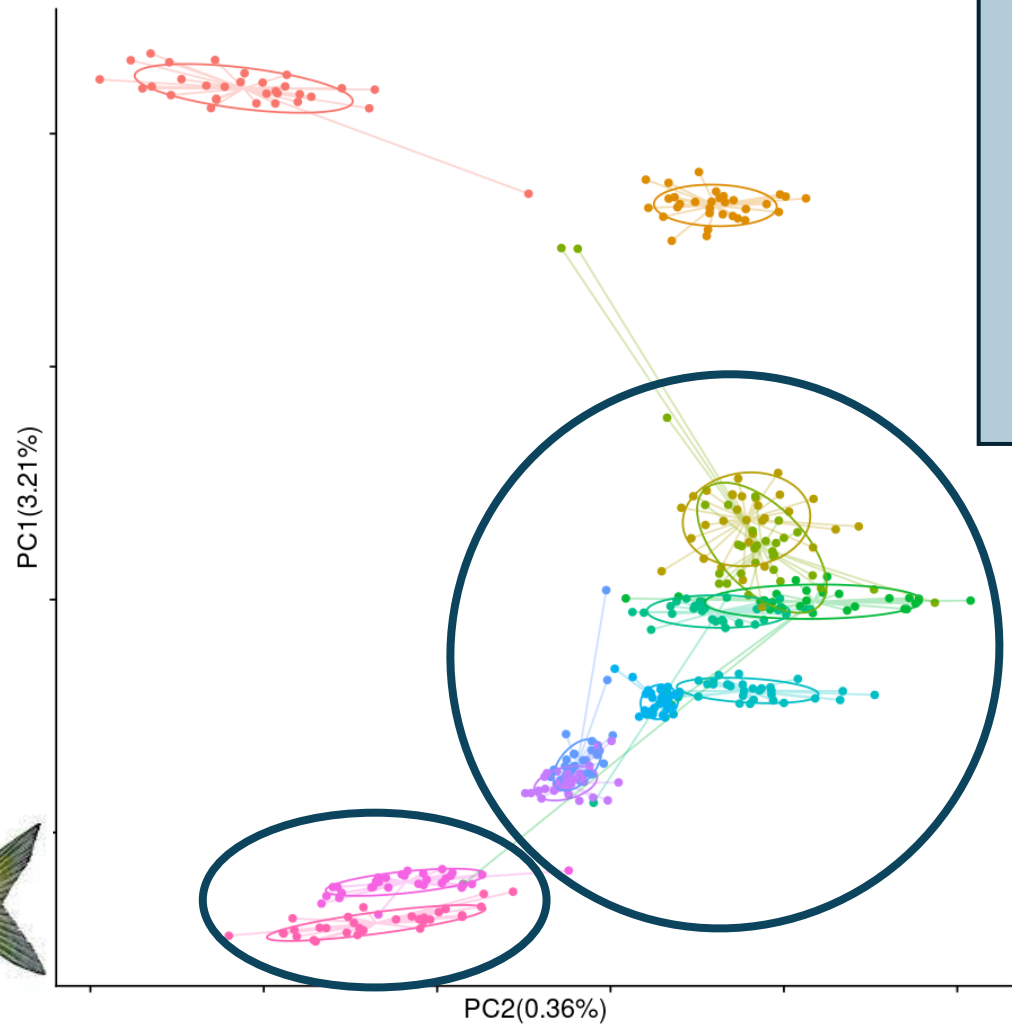
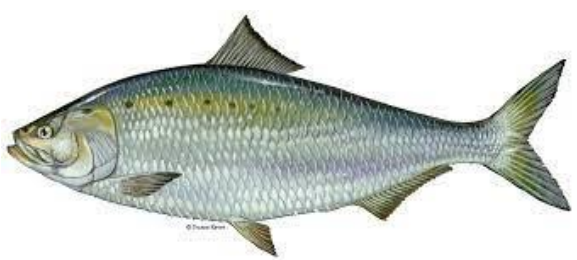
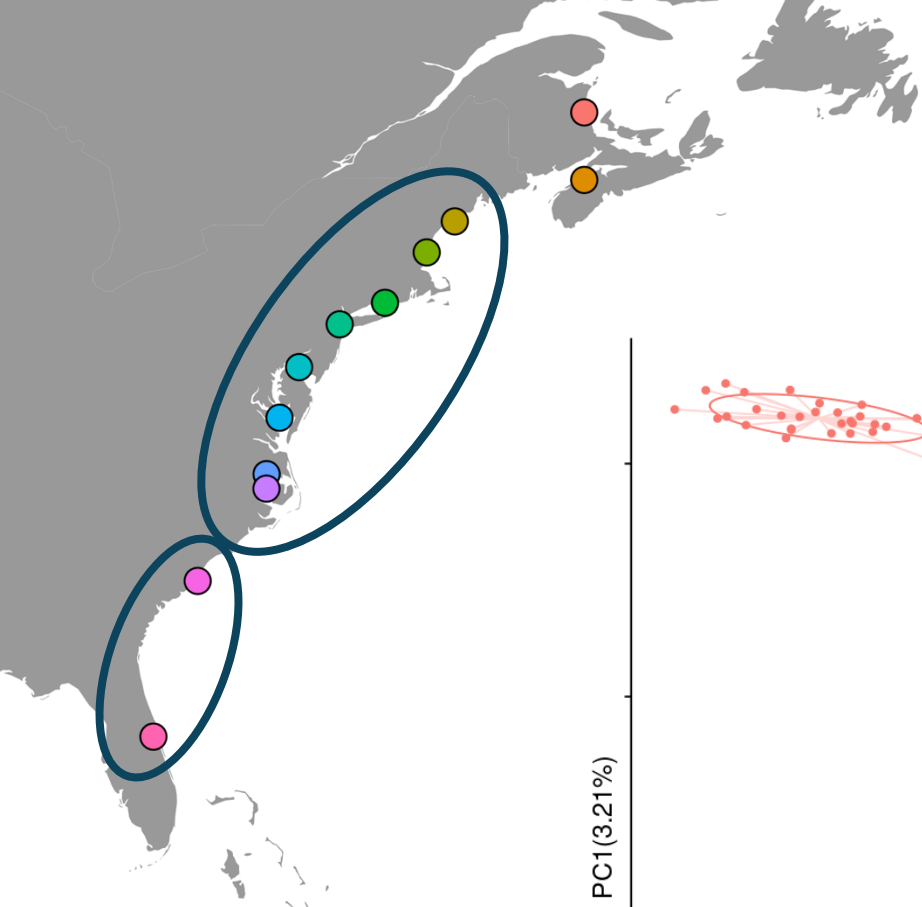
Mean: 30.2%
Range: 4 - 92%

Low-coverage whole genome sequencing

~0.8x coverage for 30 individuals per location

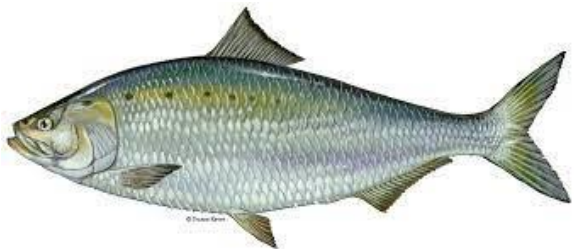
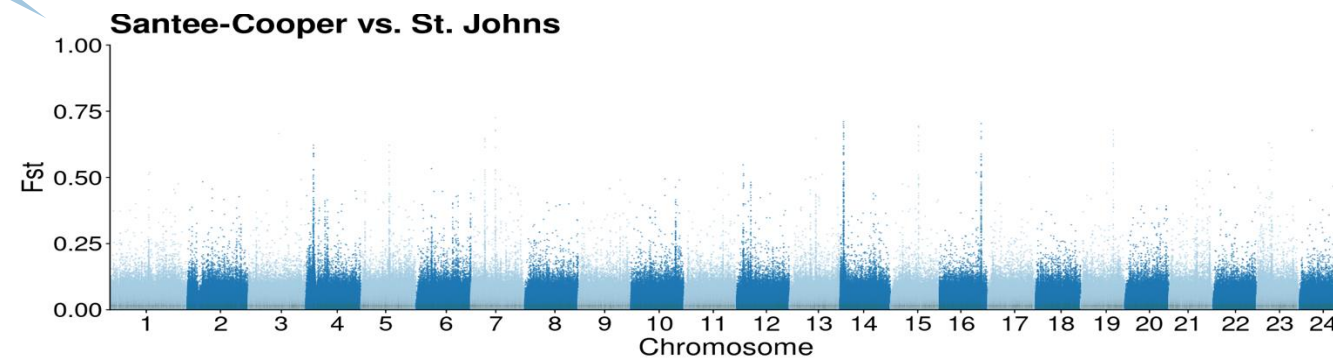
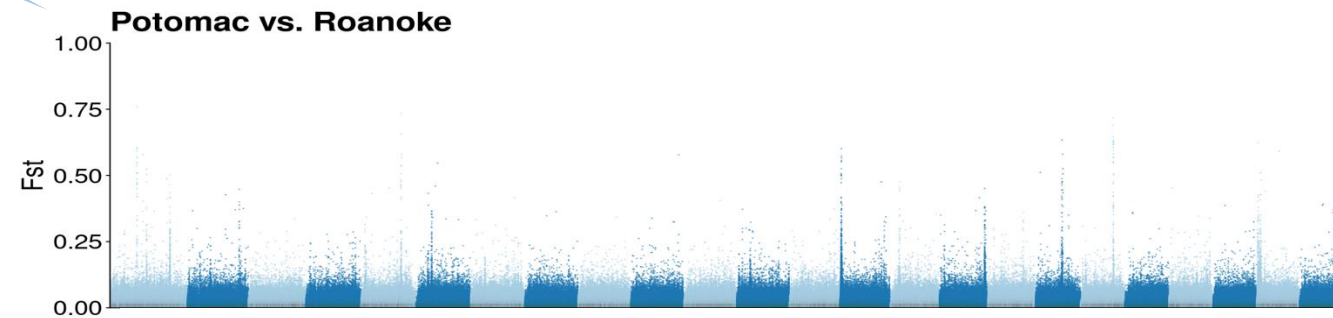
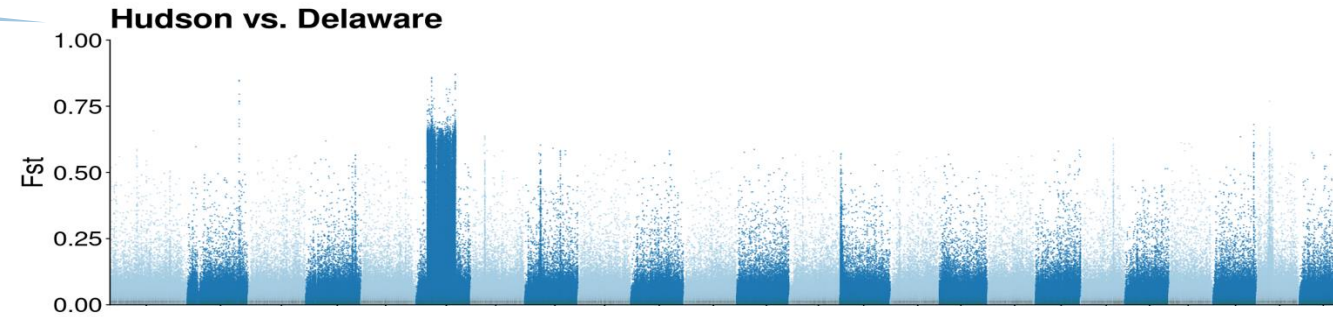
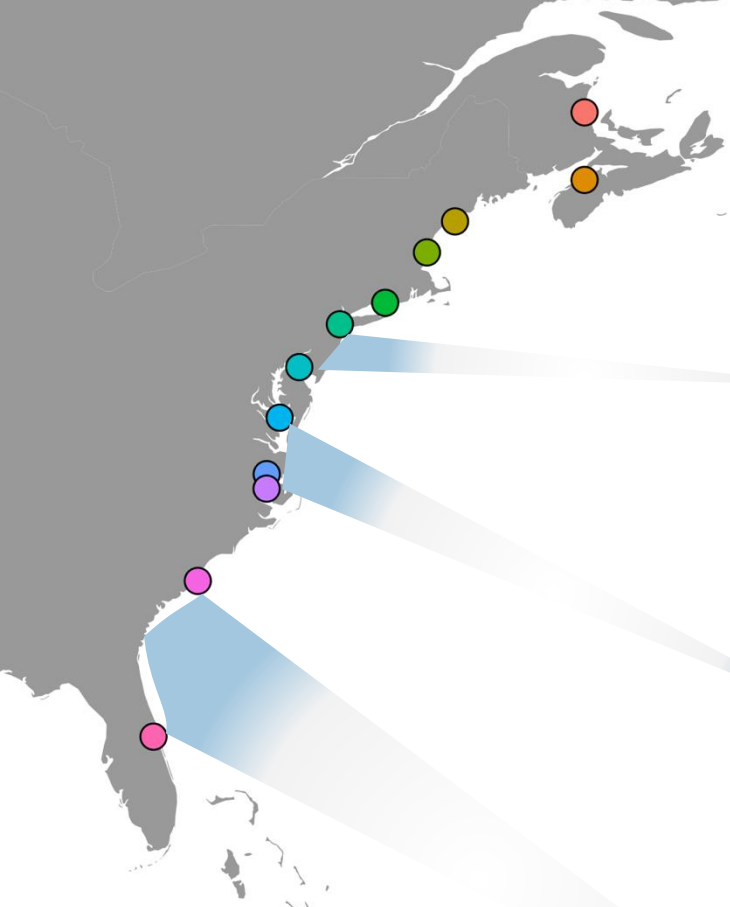


Ryan Franckowiak

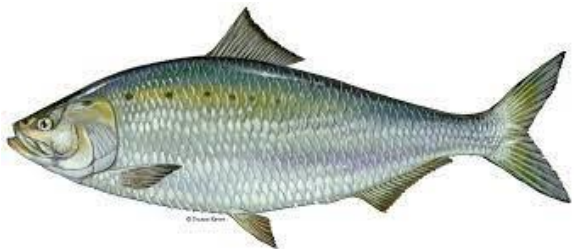
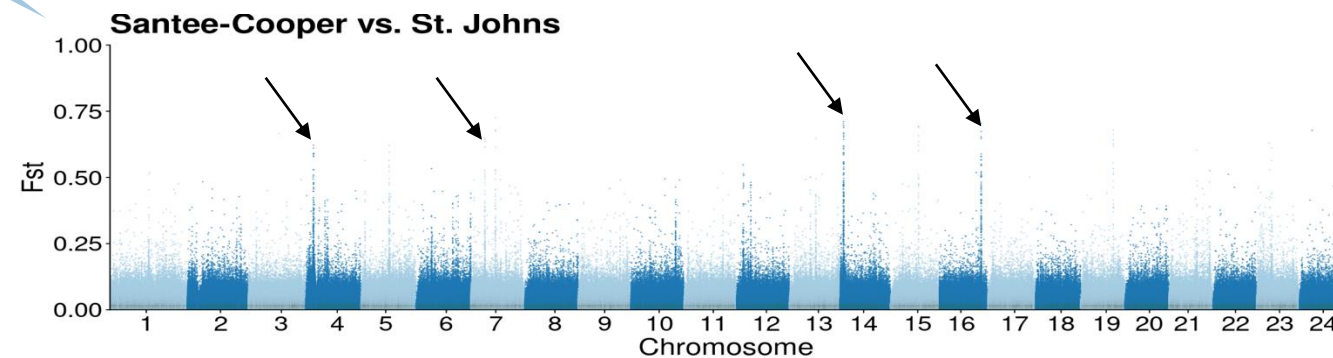
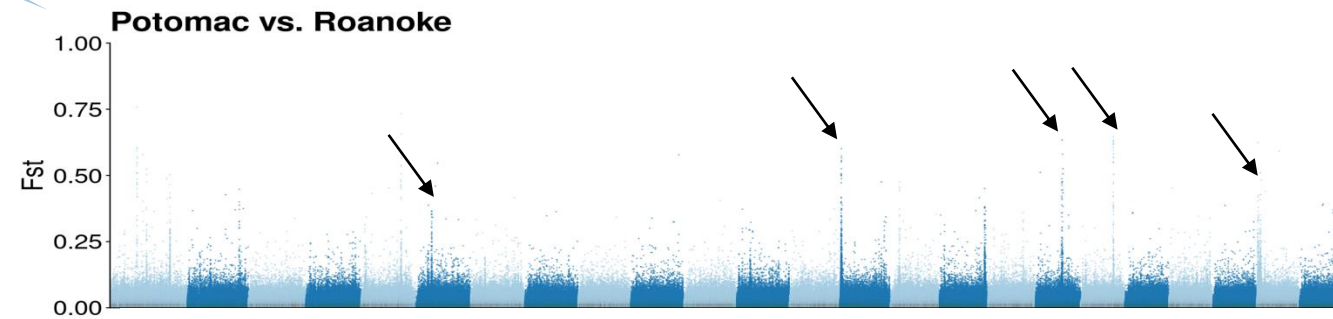
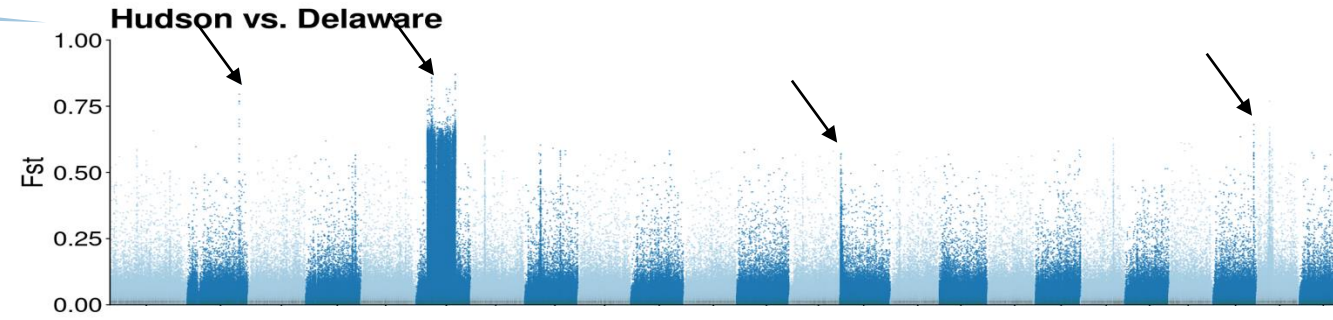
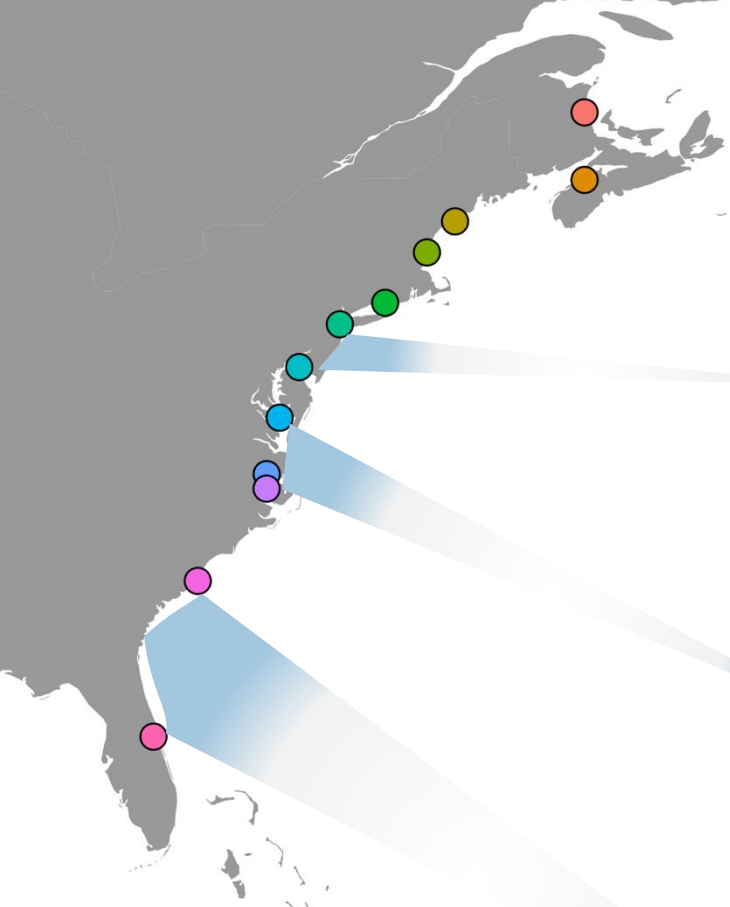


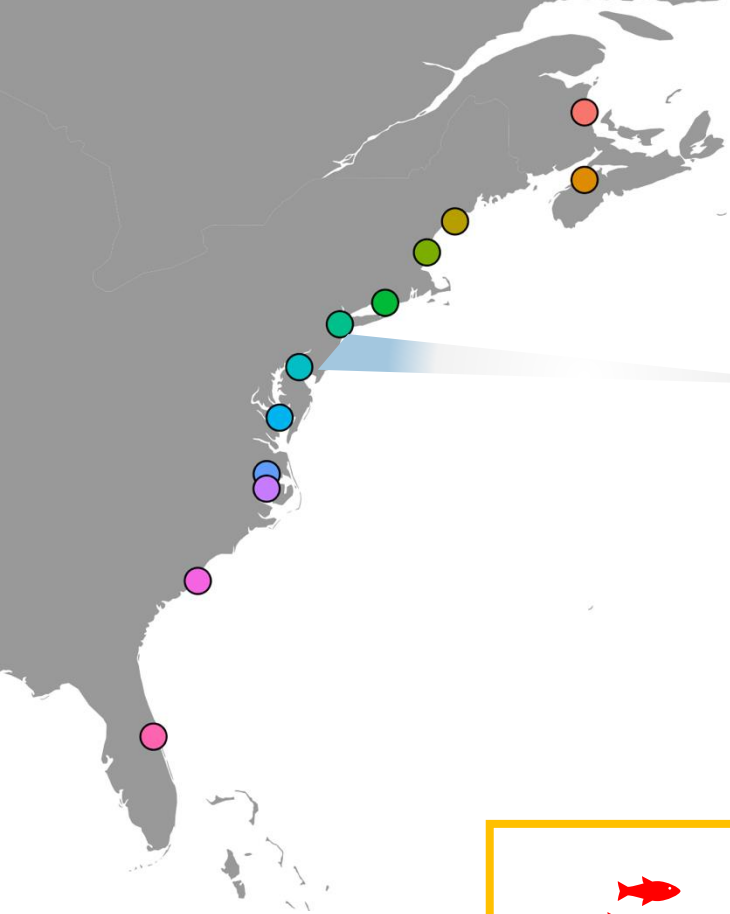
Increased population resolution
in US waters

Pairwise comparisons of differentiation among neighboring populations

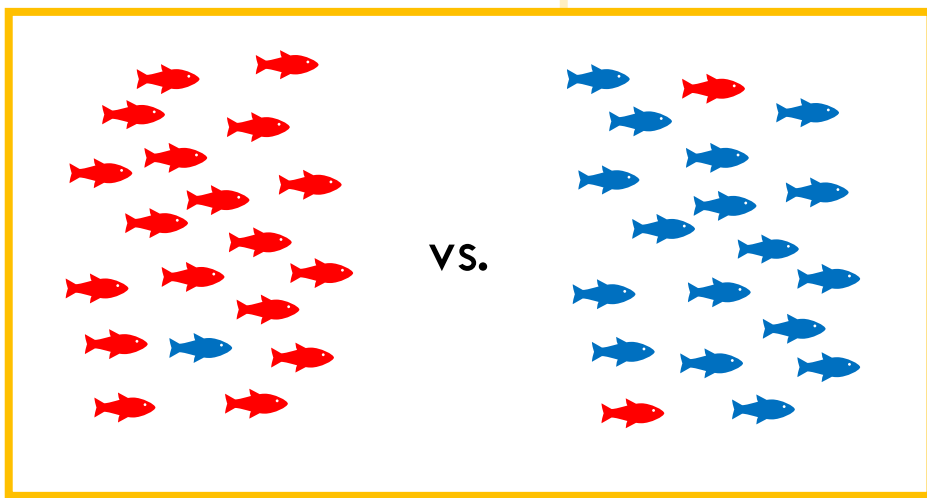
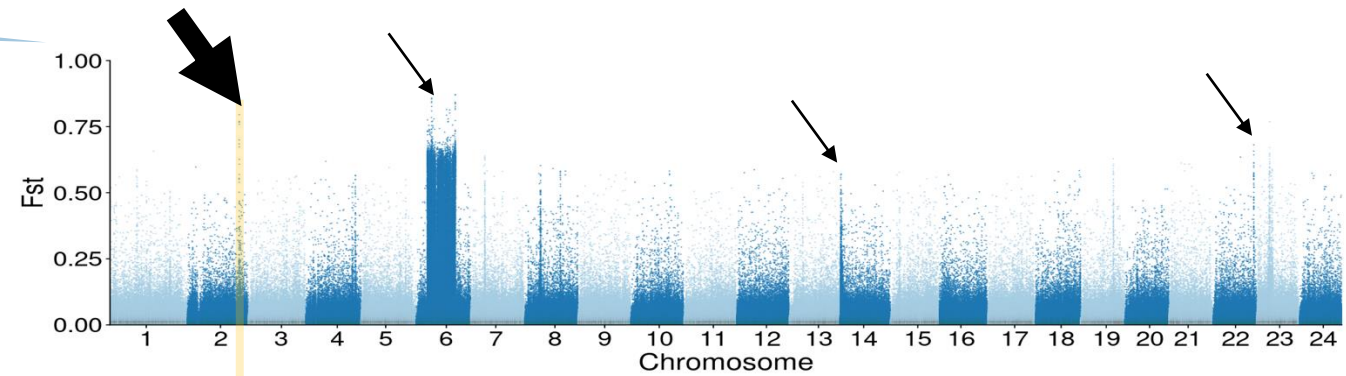


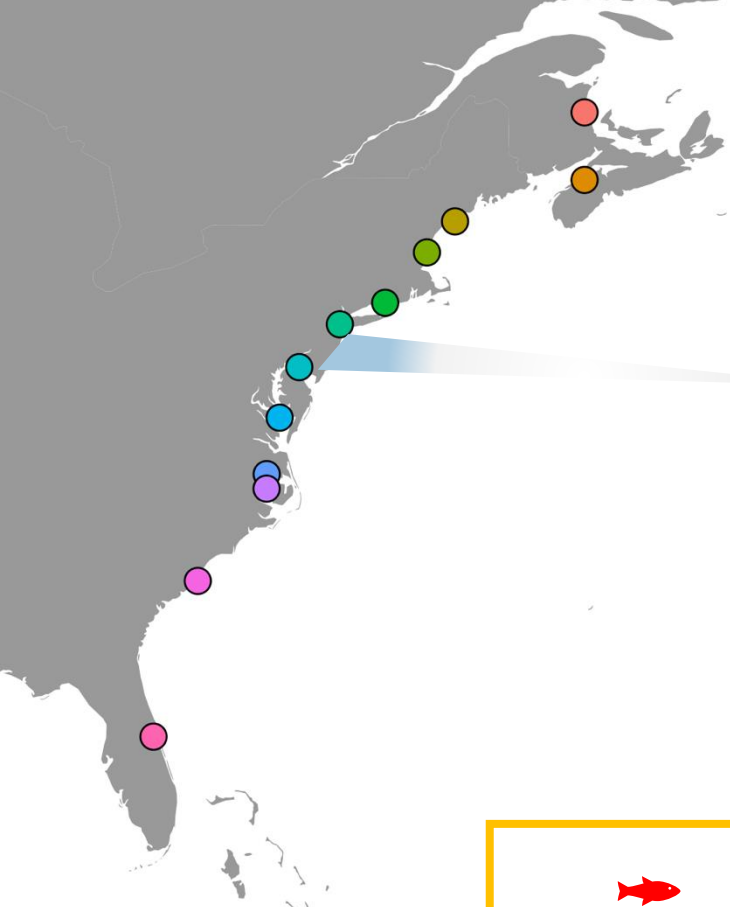
Selecting markers from the biggest peaks should improve power for assignment



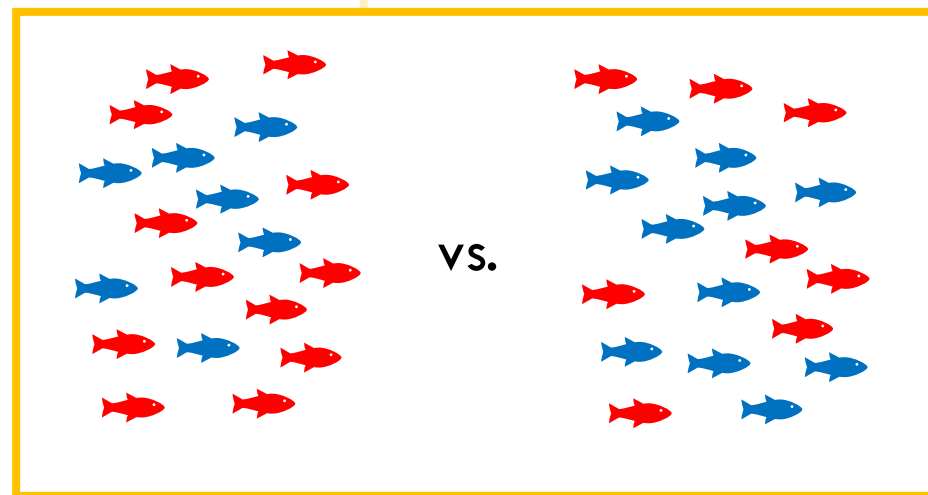
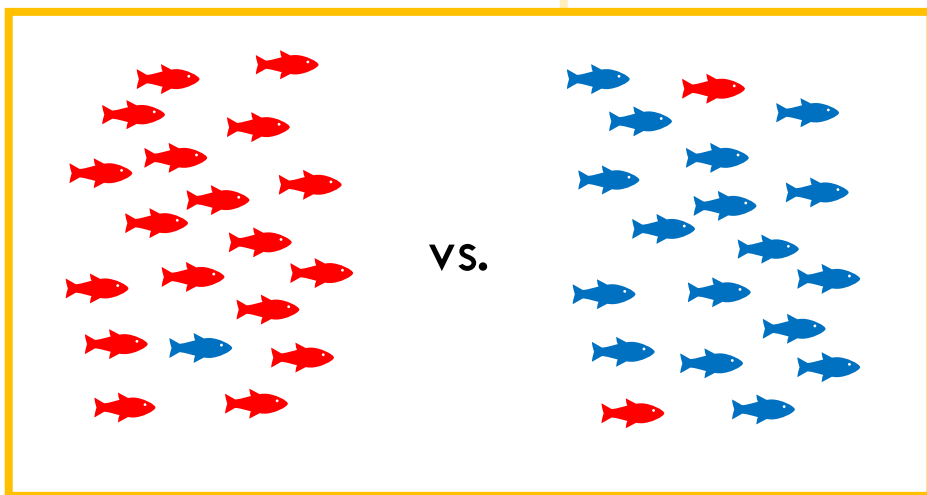
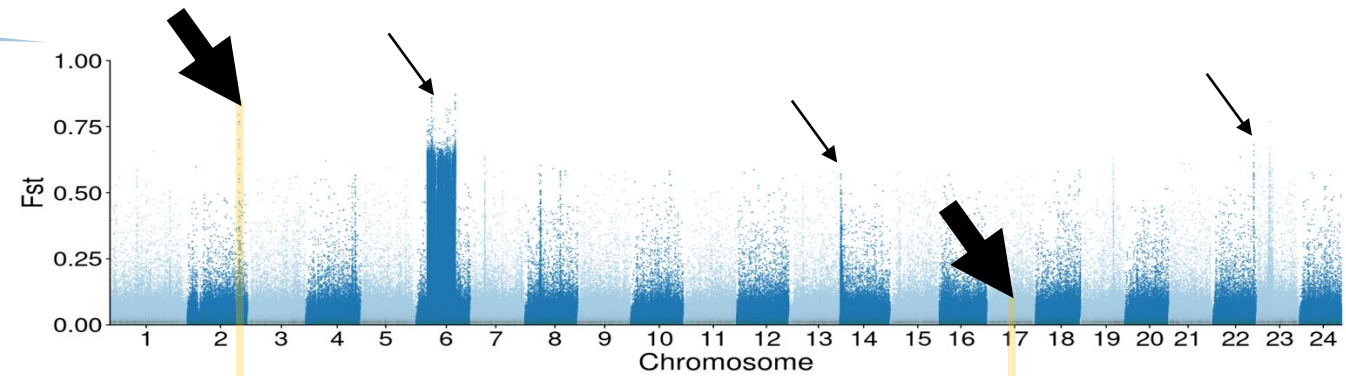


Selecting markers from the biggest peaks should improve power for assignment





Selecting markers from the biggest peaks should improve power for assignment



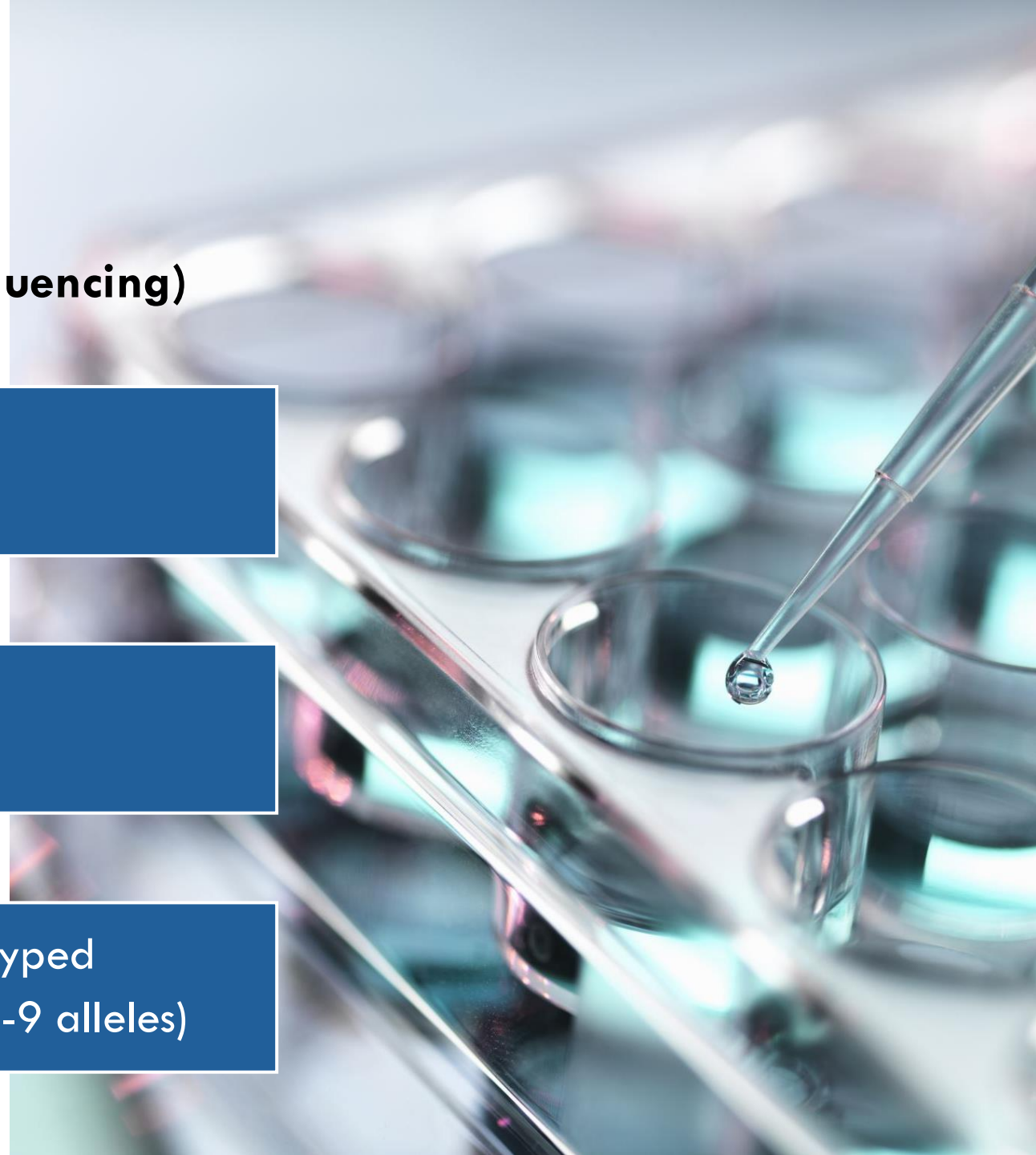
Marker panel

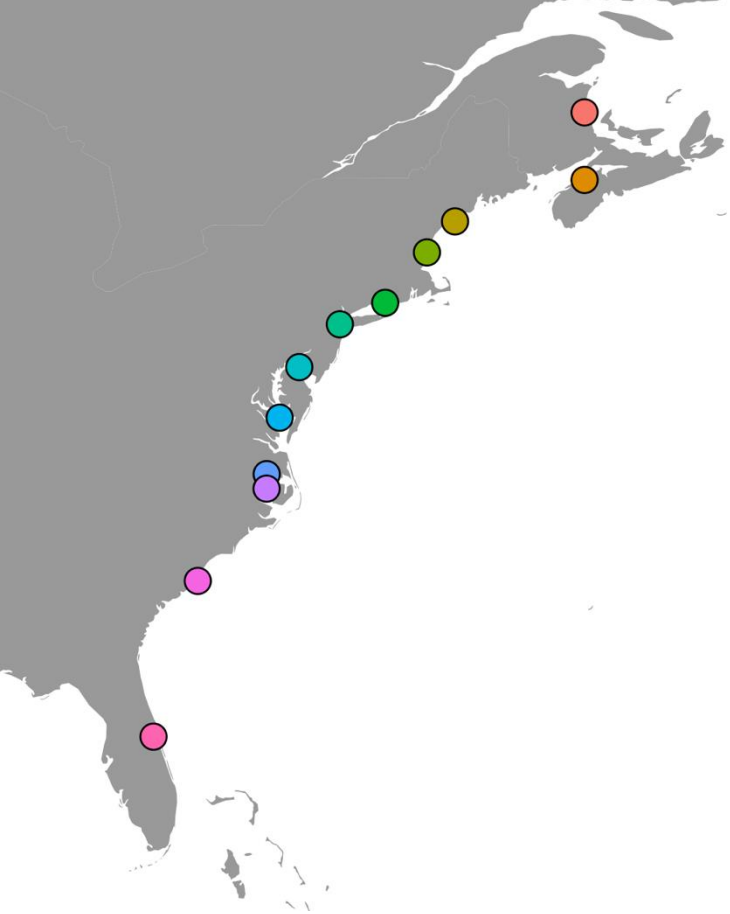
GT-seq (highly multiplexed amplicon sequencing)

578 top-ranked SNPs screened

248 retained for genotyping

107 microhaplotypes (100bp) genotyped
(average 4.2 SNPs per microhaplotype, 2-9 alleles)

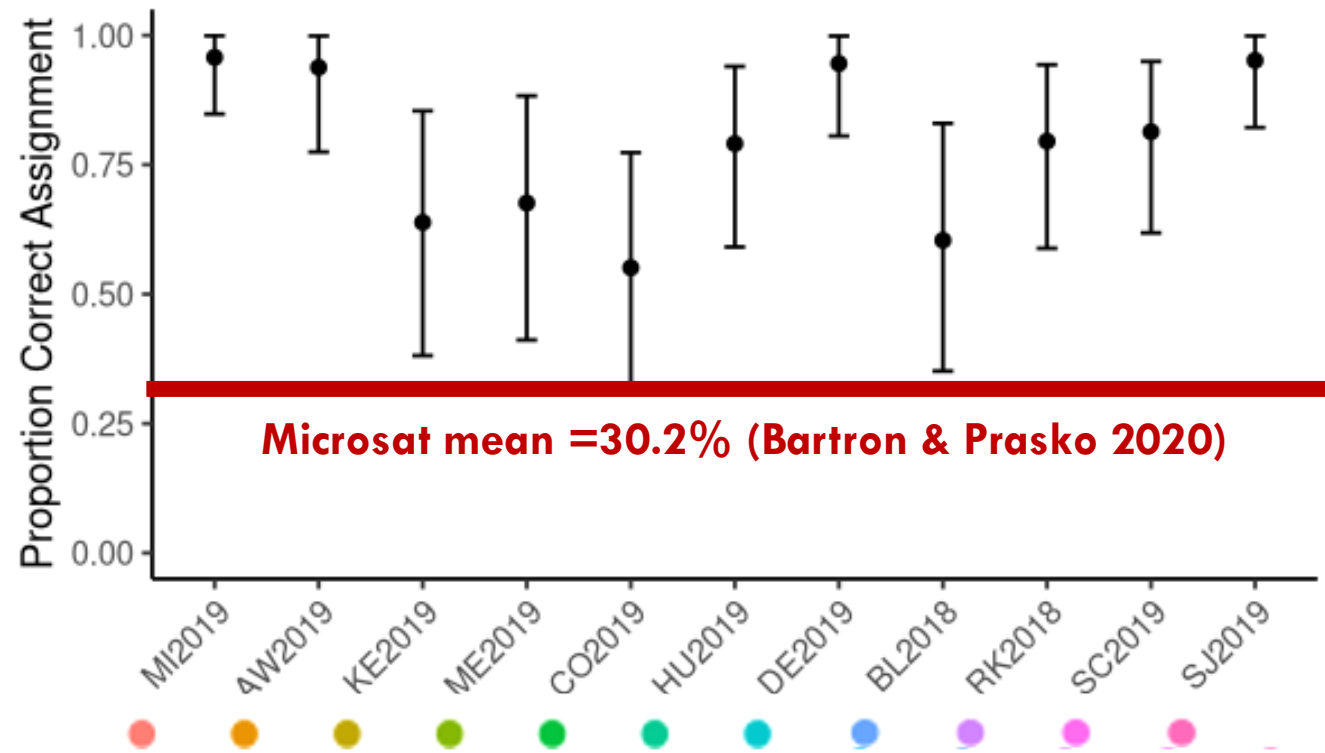


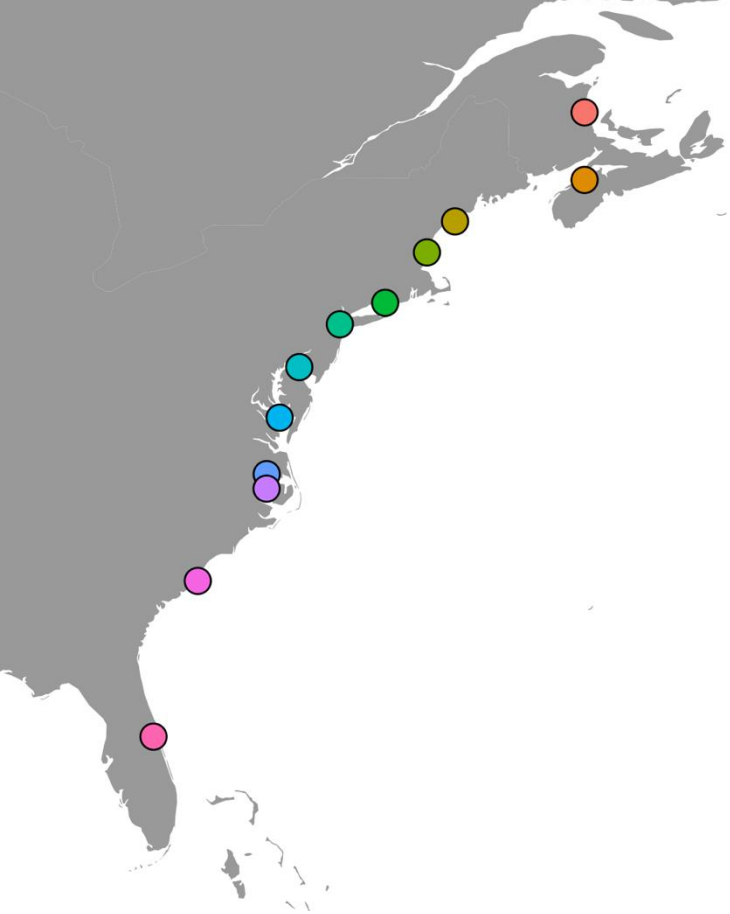


Performance of GT-seq panel

SNP selection based on 30 training individuals from each population
Assignment success assessed for 20 separate in-year test individuals

Leave-one-out assignment with different simulated mixture proportions

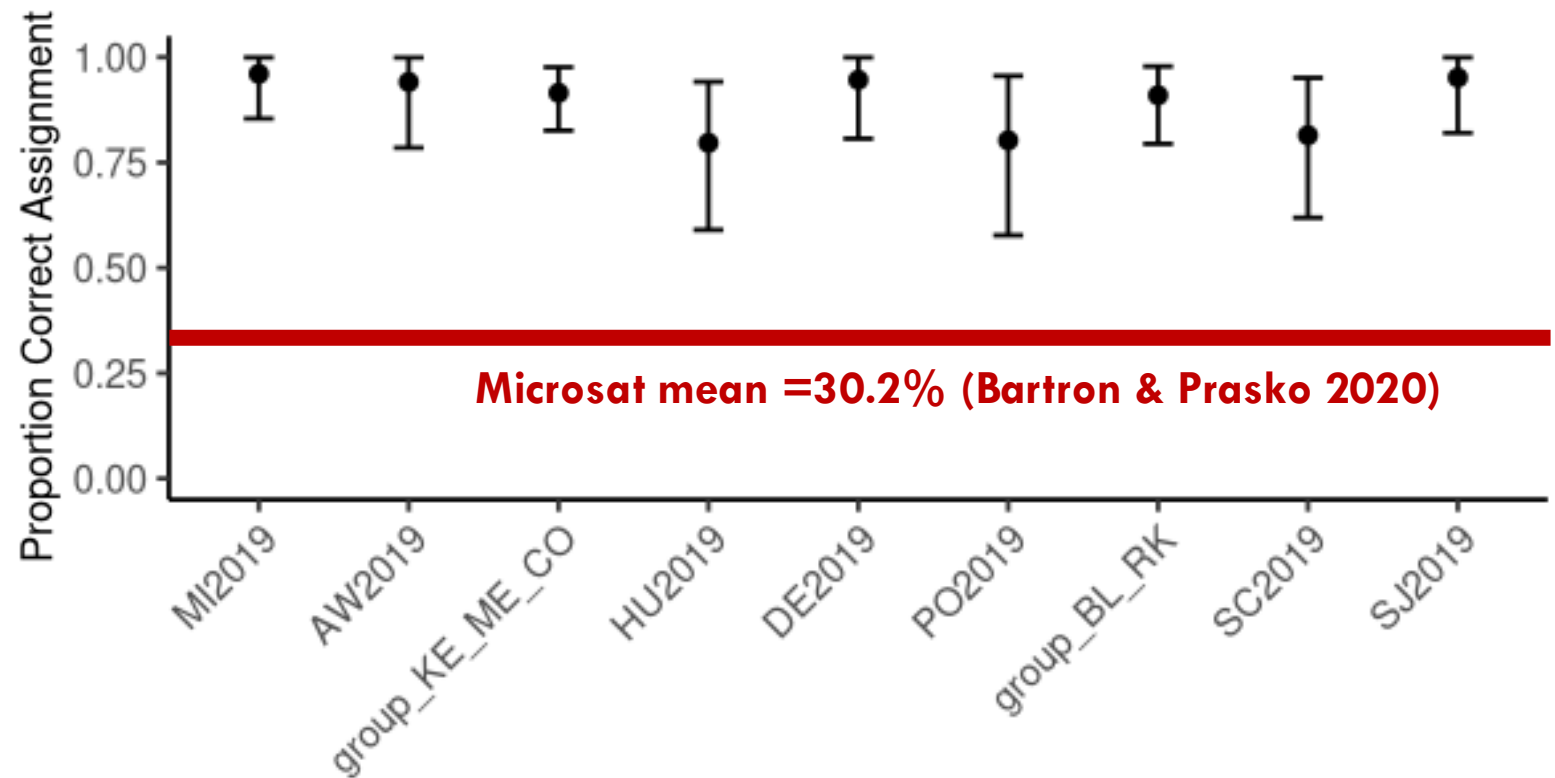




Performance of GT-seq panel

SNP selection based on 30 training individuals from each population
Assignment success assessed for 20 separate in-year test individuals

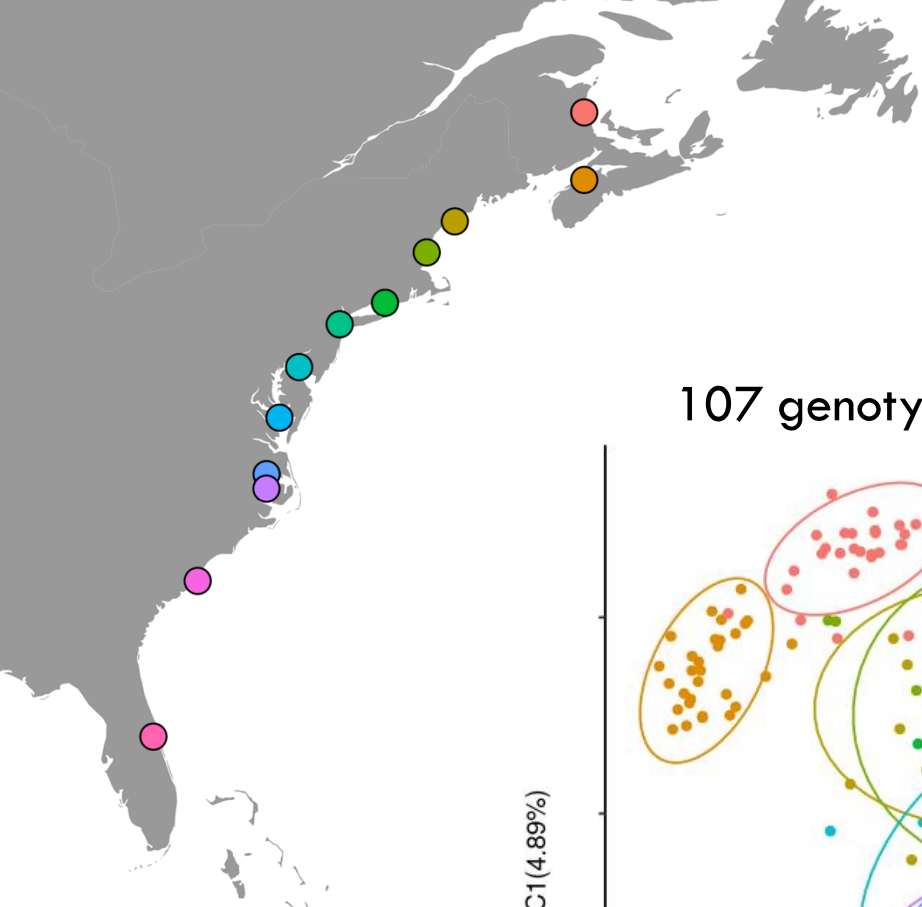
Leave-one-out assignment with different simulated mixture proportions



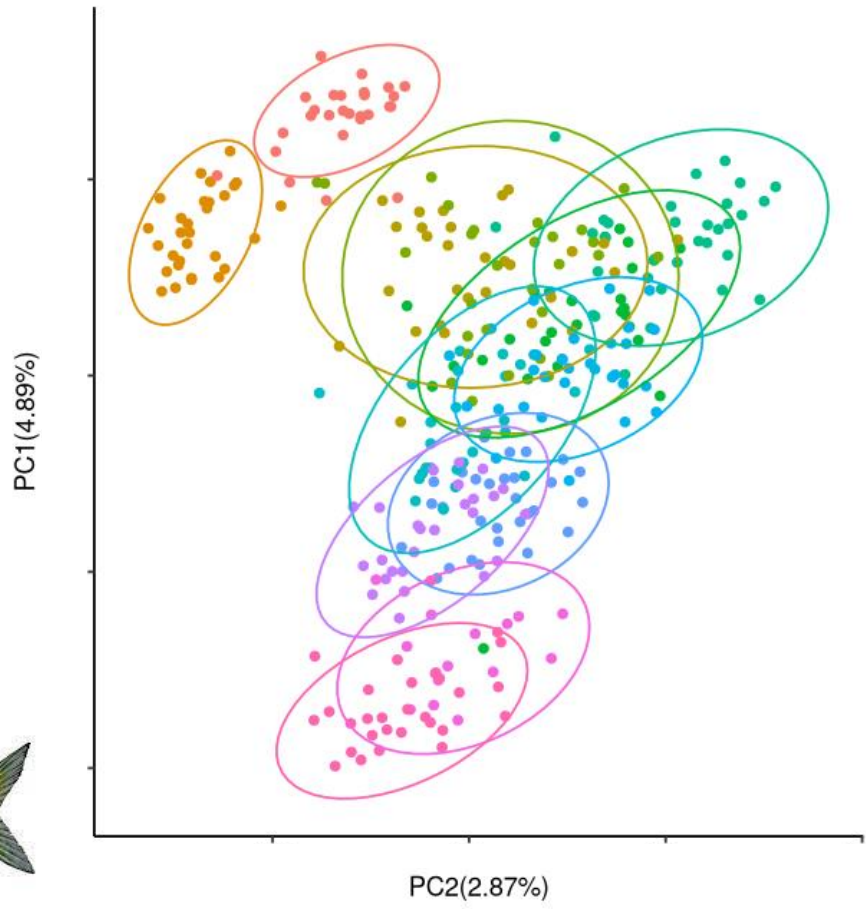
Downsides of genotyping panel

- Costly to develop
- Ascertainment bias based on which populations are included
- Large investment in targeted primers
- Does not capture the full variation

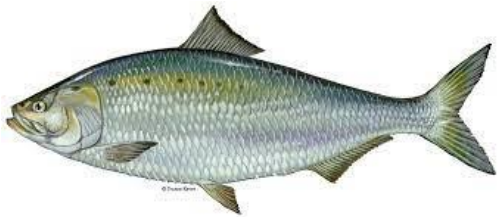
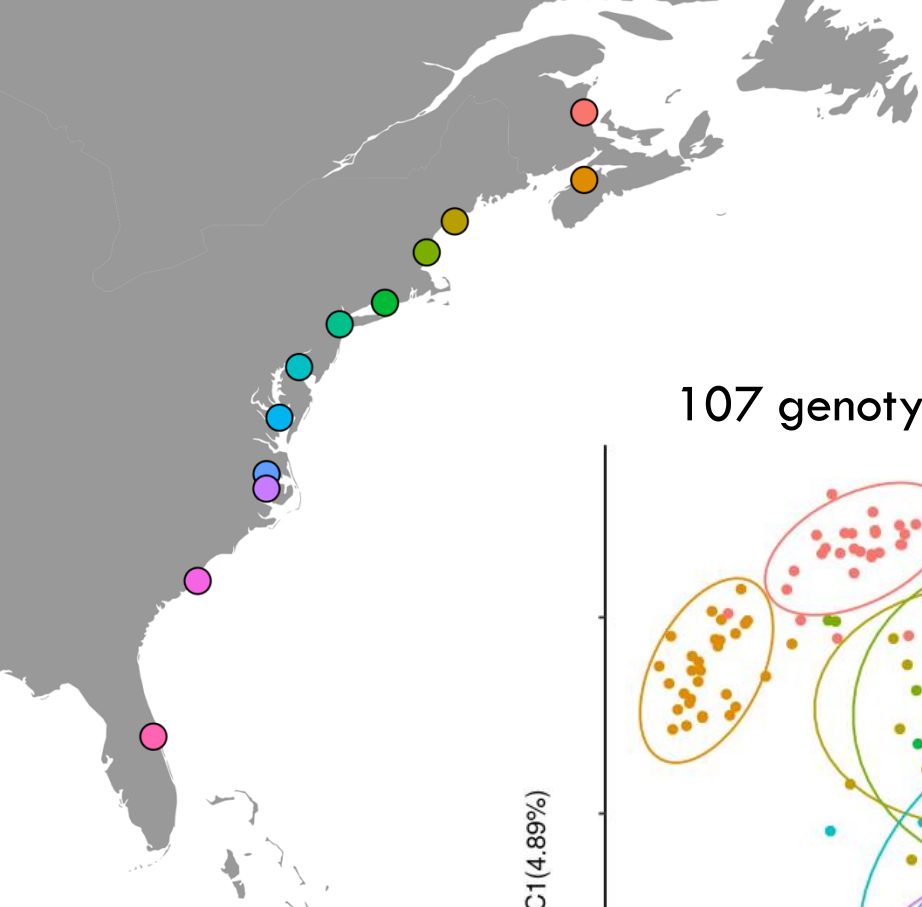
Principal Components Analysis



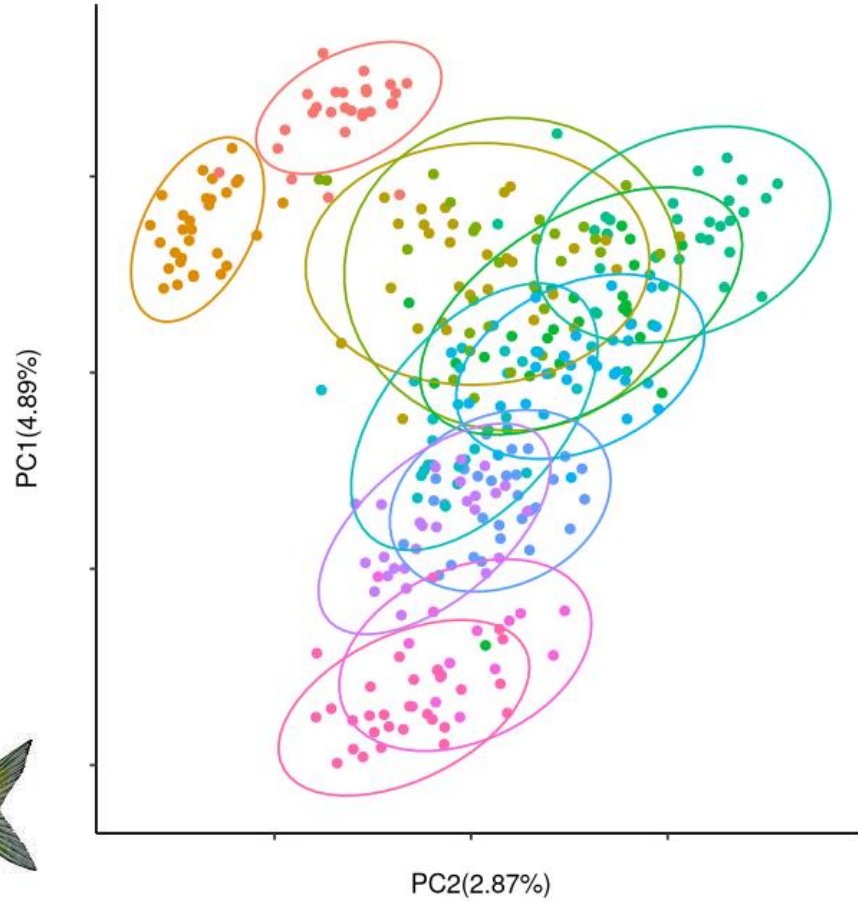
107 genotyped microhaplotypes



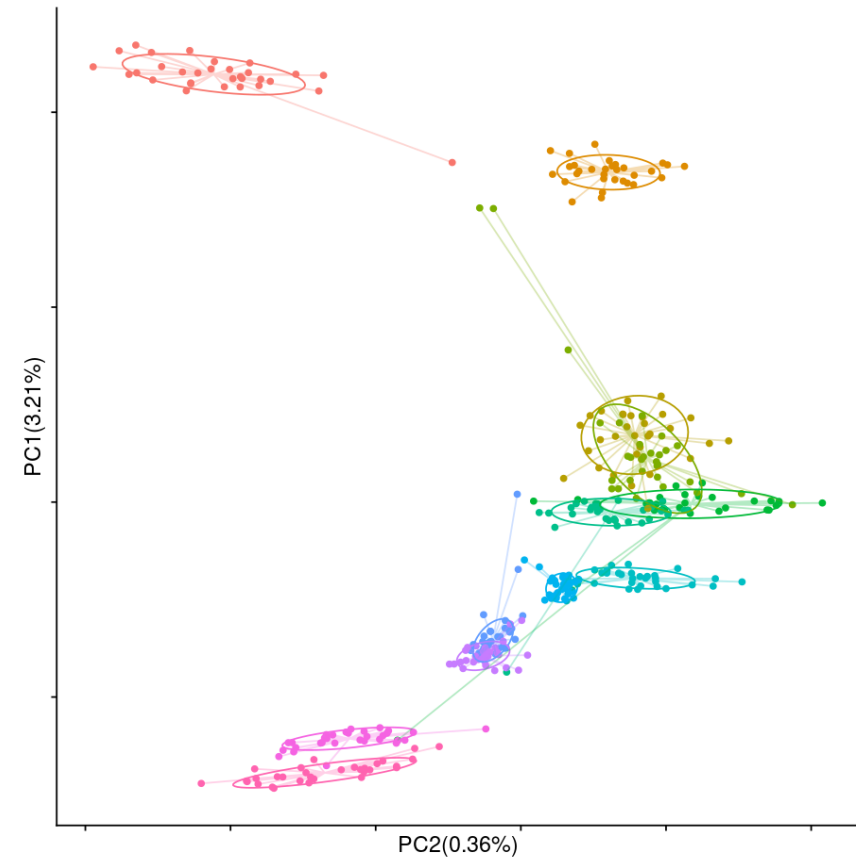
Principal Components Analysis



107 genotyped microhaplotypes



IcWGS (~0.8x), 12 million SNPs



Assignment based directly on low-coverage whole genome data



Received: 19 September 2023 | Accepted: 15 December 2023

DOI: 10.1111/2041-210X.14286

RESEARCH ARTICLE

Methods in Ecology and Evolution

Population assignment from genotype likelihoods for low-coverage whole-genome sequencing data

Matthew G. DeSaix¹ | Marina D. Rodriguez¹ | Kristen C. Ruegg¹ |
Eric C. Anderson^{1,2,3}

¹Department of Biology, Colorado State University, Fort Collins, Colorado, USA

²Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service NOAA, Santa Cruz, California, USA

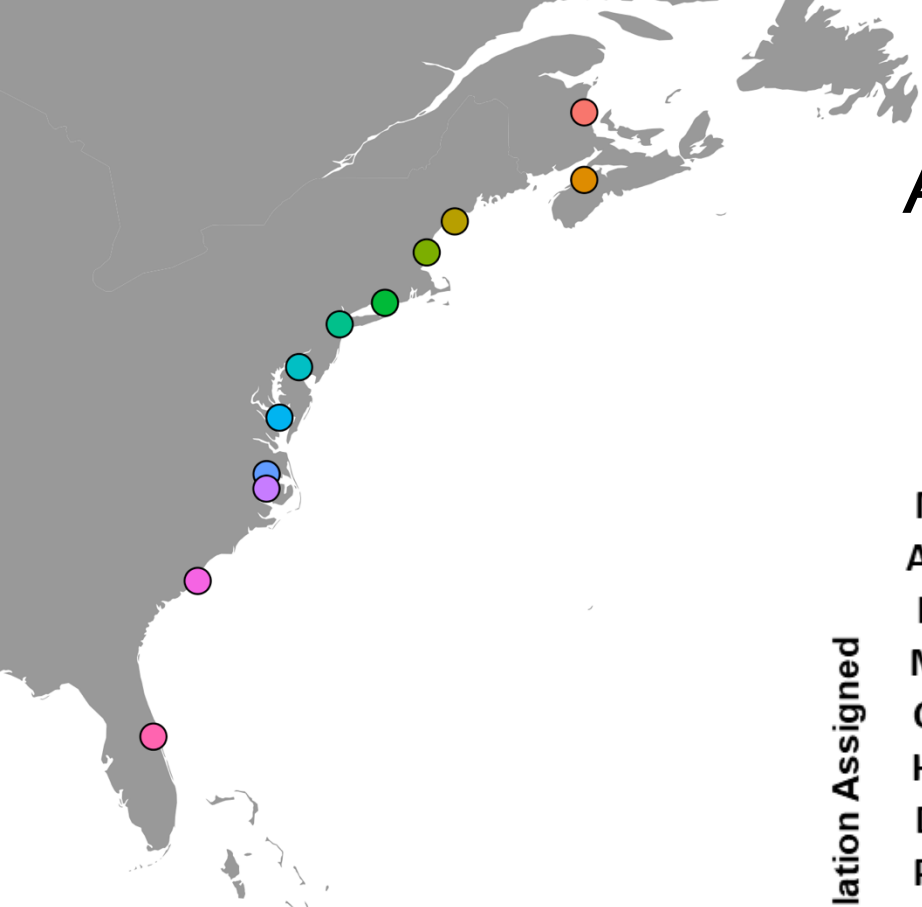
³Department of Fisheries, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, Colorado, USA

Correspondence

Matthew G. DeSaix
Email: mgdesaix@gmail.com

Abstract

1. Low-coverage whole-genome sequencing (WGS) is increasingly used for the study of evolution and ecology in both model and non-model organisms; however, effective application of low-coverage WGS data requires the implementation of probabilistic frameworks to account for the uncertainties in genotype likelihoods.
2. Here, we present a probabilistic framework for using genotype likelihoods for standard population assignment applications. Additionally, we derive the Fisher information for allele frequency from genotype likelihoods and use that to describe a novel metric, the *effective sample size*, which figures heavily in assign-

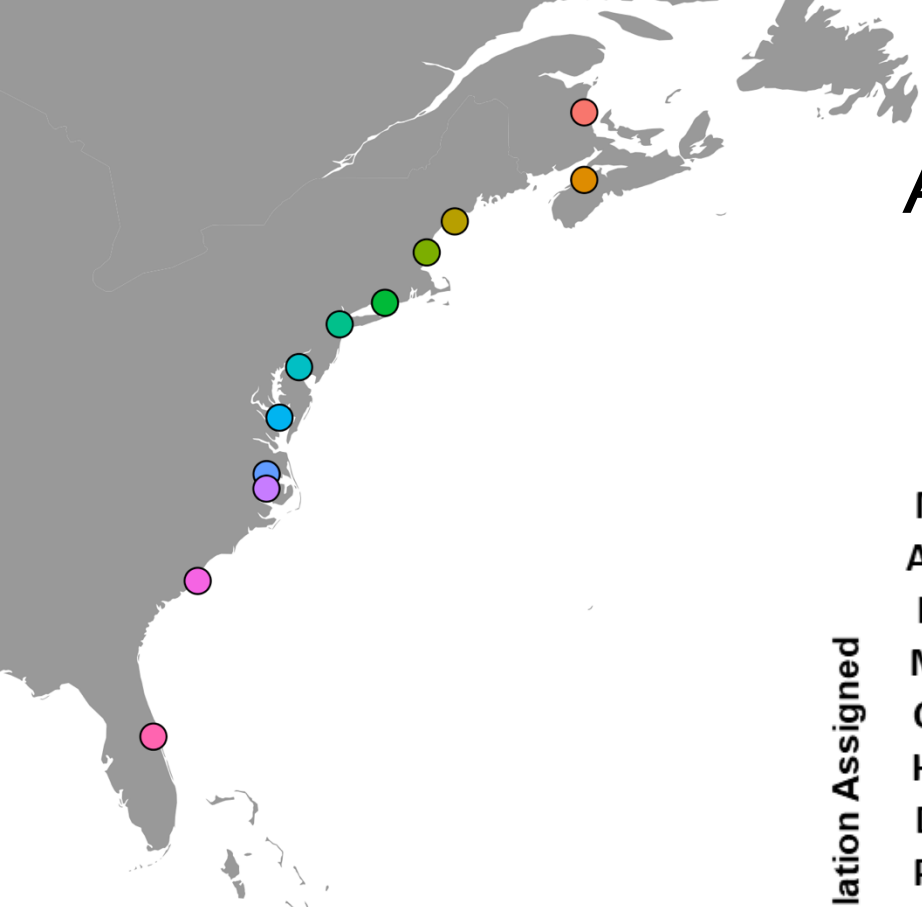


Assignment from low-coverage data

Standardized effective sample size

	Population of Origin											
	MI	AW	KE	ME	CO	HU	DE	PO	BL	RK	SC	SJ
MI	29	---	---	---	---	---	---	---	---	---	---	---
AW	---	29	---	---	---	---	---	---	---	---	---	---
KE	---	---	24	3	---	---	---	---	---	---	---	---
ME	---	---	3	24	3	---	---	---	---	---	---	---
CO	---	---	1	---	24	2	---	---	---	---	---	---
HU	---	---	---	---	---	26	---	---	---	---	---	---
DE	---	---	---	---	---	---	27	---	---	1	---	---
PO	---	---	---	---	---	---	1	28	---	---	---	---
BL	---	---	---	---	---	---	---	---	26	3	---	---
RK	---	---	---	---	---	---	---	---	2	24	---	---
SC	---	---	---	---	1	---	---	---	---	---	28	---
SJ	---	---	---	---	---	---	---	---	---	---	---	28

All individuals (~3.5x lcWGS) assigned based on genotype likelihoods at ~7million SNPs with leave-one-out



Assignment from low-coverage data

Standardized effective sample size

	Population of Origin											
	MI	AW	KE	ME	CO	HU	DE	PO	BL	RK	SC	SJ
MI	29	---	---	---	---	---	---	---	---	---	---	---
AW	---	29	---	---	---	---	---	---	---	---	---	---
KE	---	---	24	3	---	---	---	---	---	---	---	---
ME	---	---	3	24	3	---	---	---	---	---	---	---
CO	---	---	1	---	24	2	---	---	---	---	---	---
HU	---	---	---	---	---	26	---	---	---	---	---	---
DE	---	---	---	---	---	---	27	---	---	1	---	---
PO	---	---	---	---	---	---	1	28	---	---	---	---
BL	---	---	---	---	---	---	---	---	26	3	---	---
RK	---	---	---	---	---	---	---	---	2	24	---	---
SC	---	---	---	---	1	---	---	---	---	---	28	---
SJ	---	---	---	---	---	---	---	---	---	---	---	28

All individuals (~3.5x lcWGS) assigned based on genotype likelihoods at ~7million SNPs with leave-one-out

How low can we go?

- Analysis on prior data has retained near-perfect assignment when down-sampling to 0.01x sequencing depth

(DeSaix et al. 2023. *Molecular Ecology* , DeSaix et al. 2024. *Methods in Ecology and Evolution*)

How low can we go?

- Analysis on prior data has retained near-perfect assignment when down-sampling to 0.01x sequencing depth

(DeSaix et al. 2023. *Molecular Ecology* , DeSaix et al. 2024. *Methods in Ecology and Evolution*)

- With a 1 Gb genome, we then only need ~10Mb sequencing data to get 0.01x genome coverage

How low can we go?

- Analysis on prior data has retained near-perfect assignment when down-sampling to 0.01x sequencing depth

(DeSaix et al. 2023. *Molecular Ecology* , DeSaix et al. 2024. *Methods in Ecology and Evolution*)

- With a 1 Gb genome, we then only need ~10Mb sequencing data to get 0.01x genome coverage (**cost < \$0.1**)

Cost comparison

GT-seq (200-400 markers)

- Cost for library prep and sequencing ~\$6

LcWGS (millions of markers)

- Library prep: ~\$7
- Sequencing test sample: \$0.1

Cost comparison

GT-seq (200-400 markers)

- Cost for library prep and sequencing ~\$6

LcWGS (millions of markers)

- Library prep: ~\$7
- Sequencing test sample: \$0.1

Routine running costs very similar

Assignment from whole genome data often more powerful

Combating illegal fishing

INCREASING RESOLUTION IN ASSIGNMENT TESTING



Sample from catch

Fish harvested in unknown location



Genomic assignment

High-resolution genetic data assigns fish to its likely population of origin



Origin identified

Determine where the fish came from and whether the catch is authorized



Inform enforcement

Science-based evidence supports compliance and helps deter illegal fishing

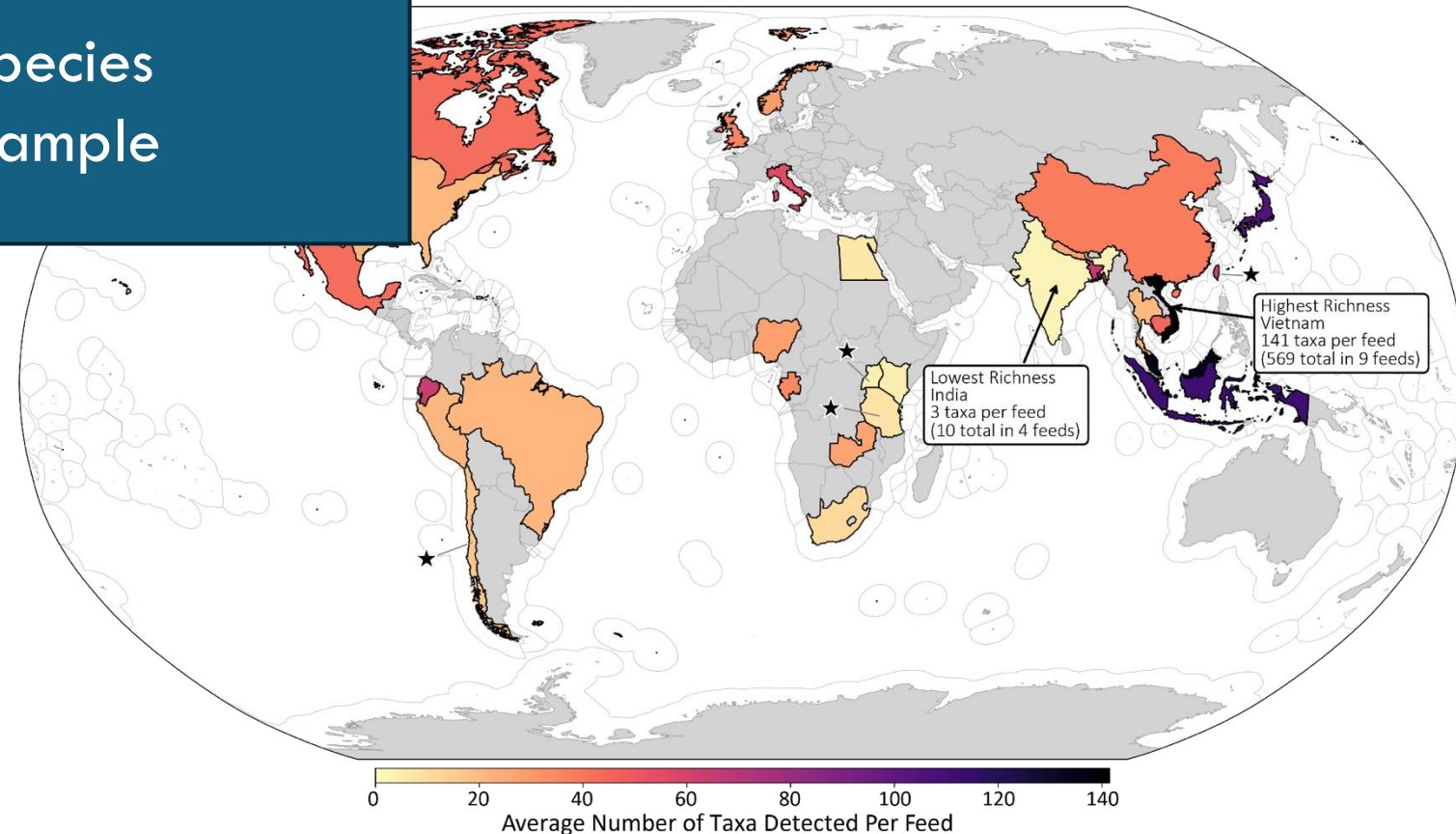


Higher resolution assignment provides the power to trace catch to its source—***helping ensure sustainable fisheries and protect ocean resources.***

Genomics as a forensic tool

- Trace biological material through supply chains

Average number of fish species
detected per aquafeed sample



Acknowledgements



Key Collaborators

Maria Akopyan
Hannes Baumann
Aryn Wilder
Arne Jacobs
Anna Tigano
Jessi Rick
David Conover
Nicolas Lou
Harmony Borchardt-Wier
Adrienne Kovach
Gemma Clucas

Ryan Franckowiak
David Kazyak
Miluska Hyde
Jessica Best
Greg Kenney
Eric Anderson
Jaime Ortiz-Pachar
Nick Locatelli
Diana Baetscher
Pete McIntyre
Pavel Dimens
... and many others

Funding:



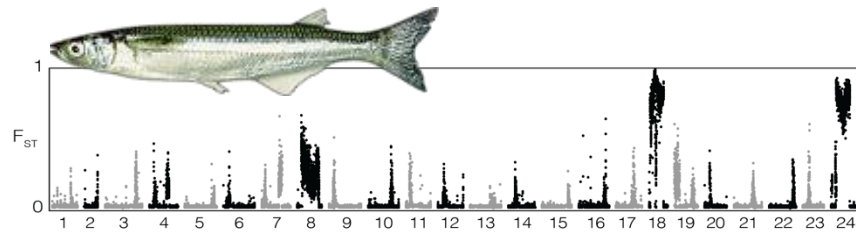
BIO-OCE 1434325
BIO-OCE 1756316



Cornell University

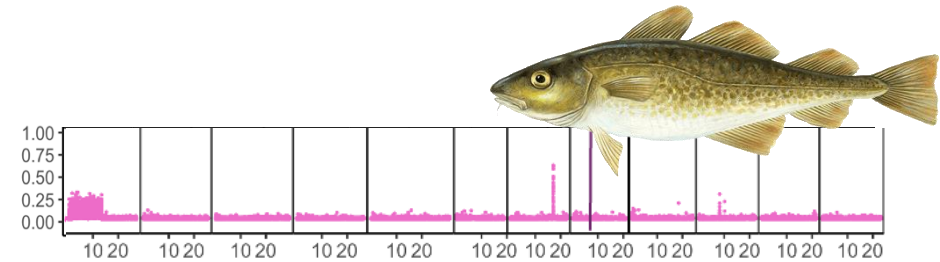
Structure can be hidden in different ways

Strong adaptation despite homogeneous genomes



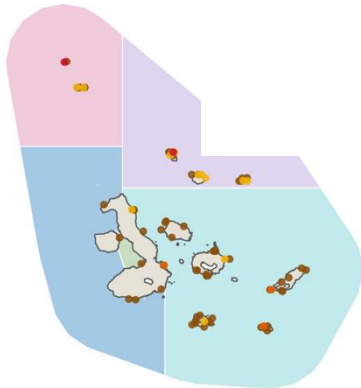
Atlantic silversides

Differences confined to small genomic regions



Atlantic cod

Cryptic mixing of distinct lineages



Brown sea cucumber

Tracking who mixes with whom—and when

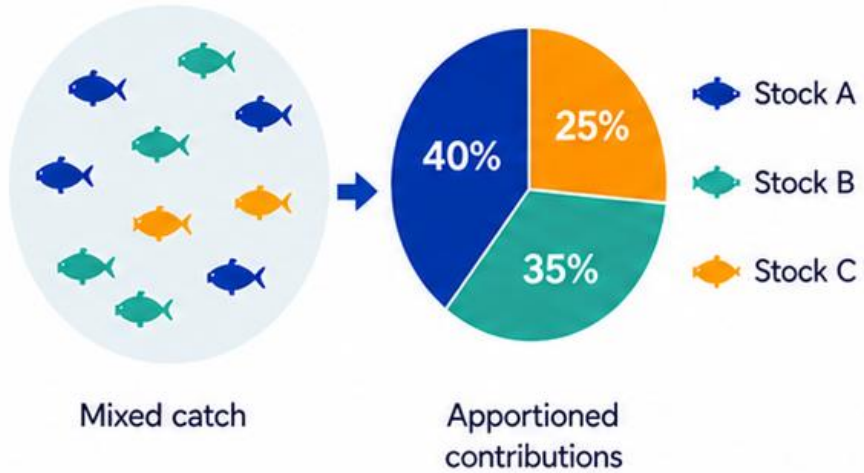
		Population of Origin											
		MI	AW	KE	ME	CO	HU	DE	PO	BL	RK	SC	SJ
Population Assigned	MI	29	---	---	---	---	---	---	---	---	---	---	---
	AW	---	29	---	---	---	---	---	---	---	---	---	---
	KE	---	---	24	3	---	---	---	---	---	---	---	---
	ME	---	---	3	24	3	---	---	---	---	---	---	---
	CO	---	---	1	---	24	2	---	---	---	---	---	---
	HU	---	---	---	---	---	26	---	---	---	---	---	---
	DE	---	---	---	---	---	---	27	---	---	1	---	---
	PO	---	---	---	---	---	---	1	28	---	---	---	---
	BL	---	---	---	---	---	---	---	---	26	3	---	---
	RK	---	---	---	---	---	---	---	---	2	24	---	---
	SC	---	---	---	---	1	---	---	---	---	---	28	---
	SJ	---	---	---	---	---	---	---	---	---	---	---	28



American shad

From science to decisions

1 Mixed-stock fisheries



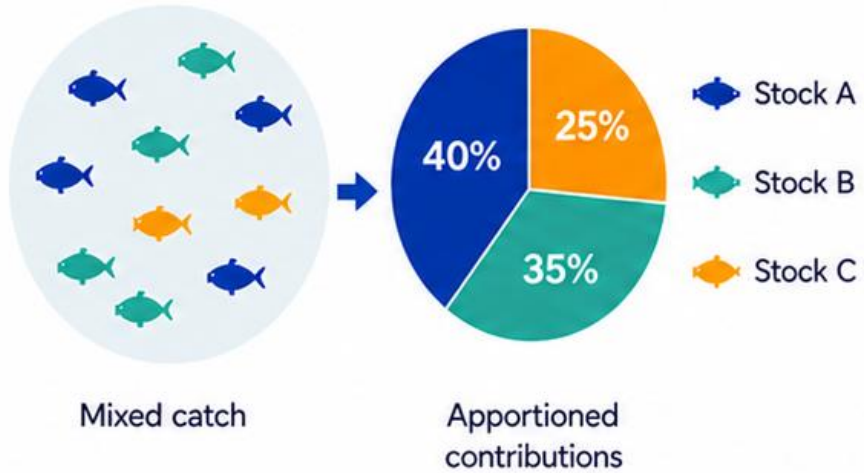
Quantify the contribution of each stock to mixed fisheries.



Genomic insights provide the resolution needed to turn complex data into **actionable information** for sustainable management.

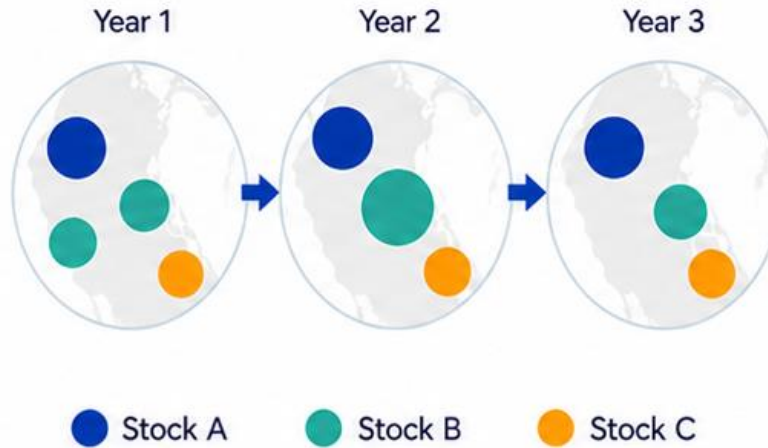
From science to decisions

1 Mixed-stock fisheries



Quantify the contribution of each stock to mixed fisheries.

2 Shifting distributions



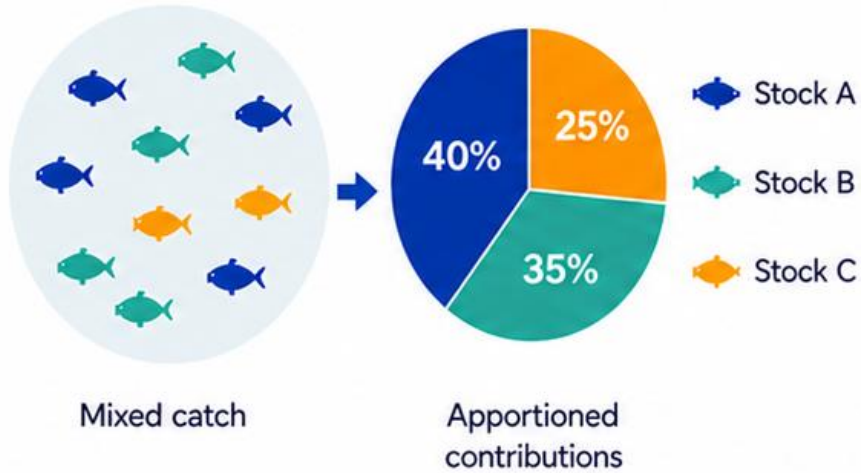
Monitor and adapt to changes in the composition of stocks through time.



Genomic insights provide the resolution needed to turn complex data into **actionable information** for sustainable management.

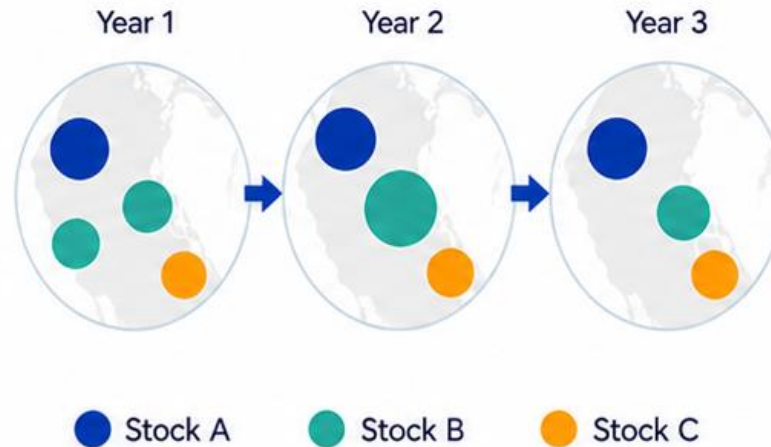
From science to decisions

1 Mixed-stock fisheries



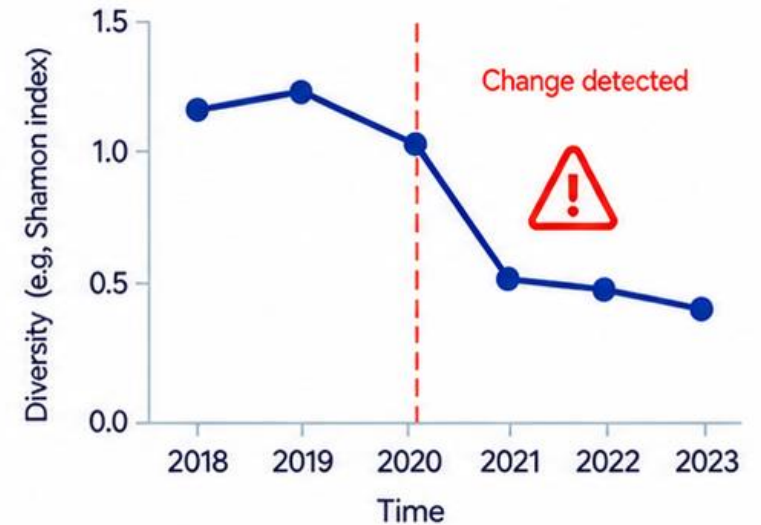
Quantify the contribution of each stock to mixed fisheries.

2 Shifting distributions



Monitor and adapt to changes in the composition of stocks through time.

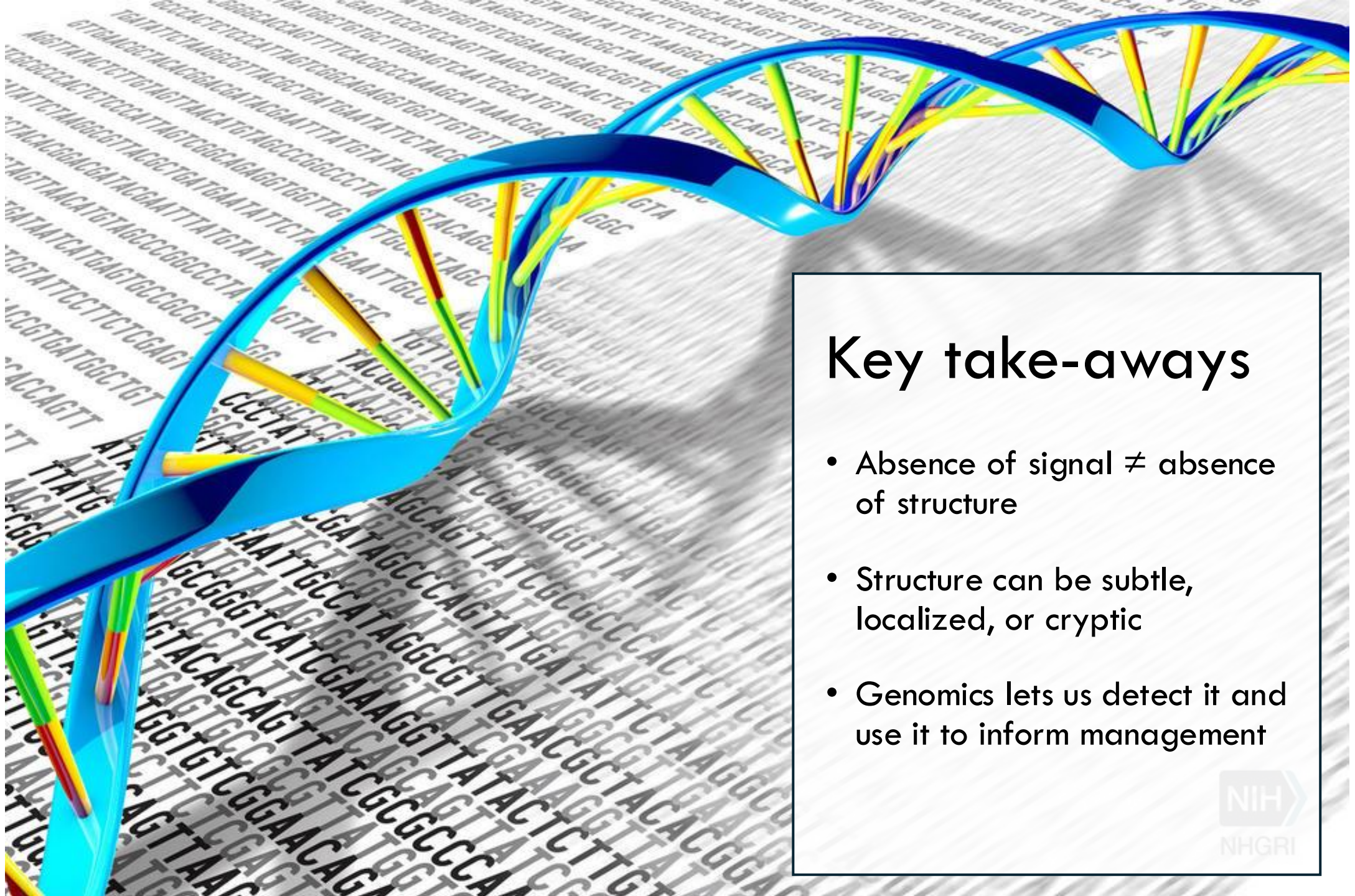
3 Early warning



Detect changes early to inform management and conservation.



Genomic insights provide the resolution needed to turn complex data into **actionable information** for sustainable management.



Key take-aways

- Absence of signal \neq absence of structure
- Structure can be subtle, localized, or cryptic
- Genomics lets us detect it and use it to inform management