Confounding factors

Day 3 - Lecture 2

(adapted from Anna Tigano, PhD)

Expectation

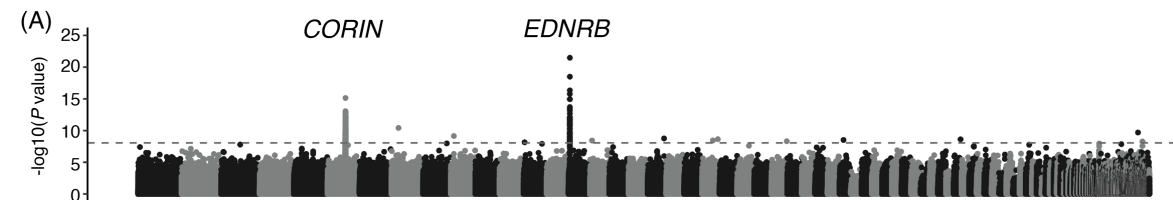Selection increases the frequency of beneficial mutations

⬇

When comparing lineages adapted to different local conditions

⬇

Expect difference in frequency at the adaptive loci

Association with winter color variation
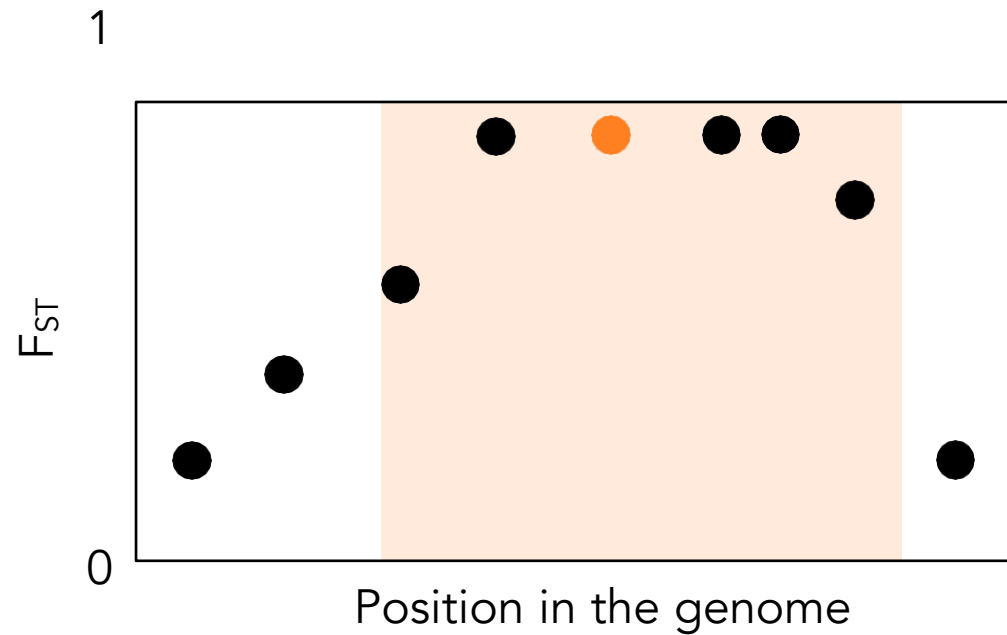

Ferreira et al 2023, Science

Yesterday I presented on the factors that contribute to our ability to detect signatures of a selective sweep in the genome.

However, that was a simple scenario, only one of many possible, and often more complex, scenarios.
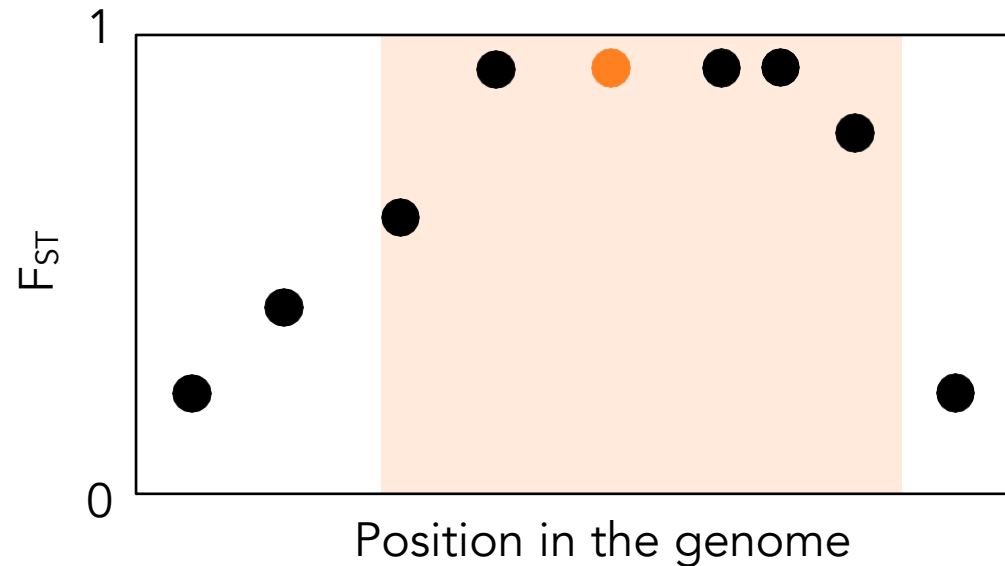
Linked selection

Sites in close proximity to the real target of selection could show high F_{ST}.

Linked selection

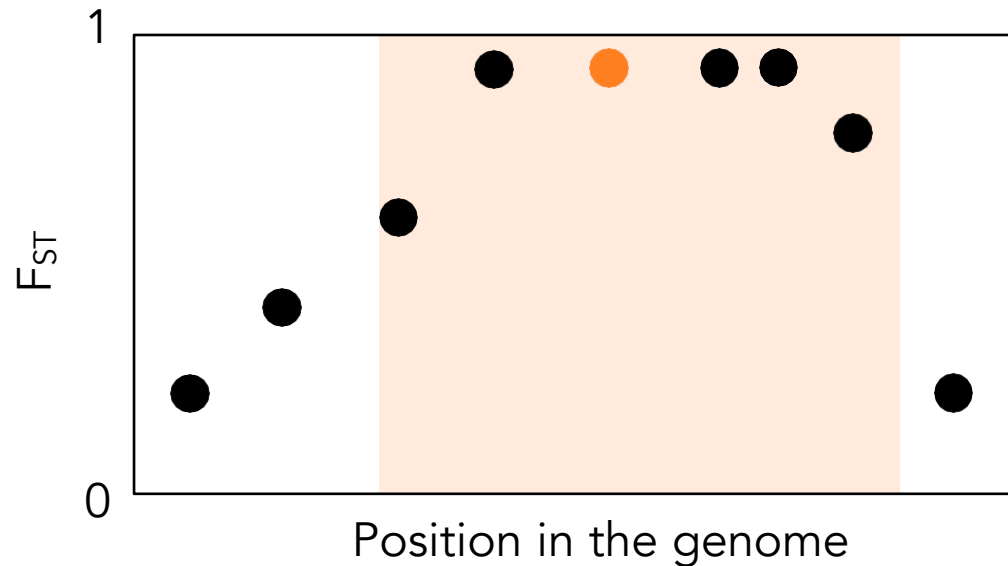Sites in close proximity to the real target of selection could show high F_{ST}.

With whole genome data

We may target most of these sites with high F_{ST}.

Linked selection

Sites in close proximity to the real target of selection could show high F_{ST}.

1

F_{ST}

0

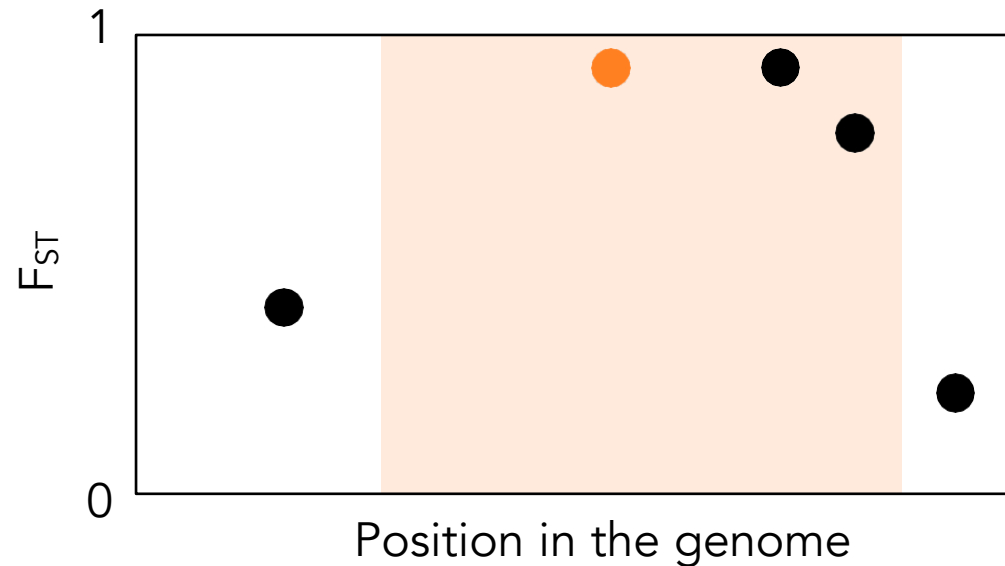Position in the genome

With whole genome data

We may target most of these sites with high F_{ST}.

The **tighter** linkage is, the harder is to **disentangle target from linked sites.**
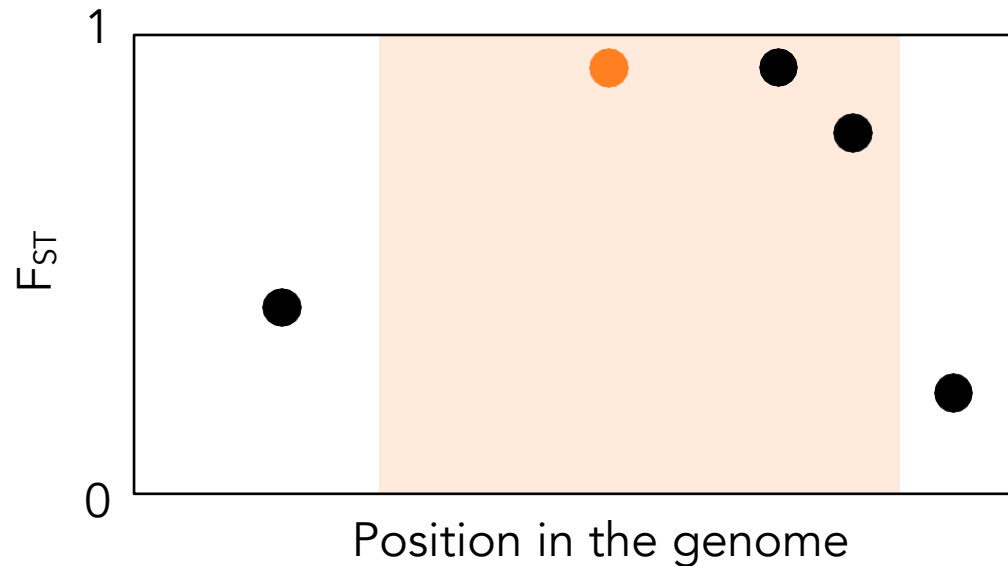
Linked selection

Sites in close proximity to the real target of selection could show high F_{ST}.

With reduced-representation approaches, you won't.

Linked selection

Sites in close proximity to the real target of selection could show high F_{ST}.

With reduced-representation approaches, you won't.

Your F_{ST} outlier could be the real target of selection, or a neutral site linked to it.
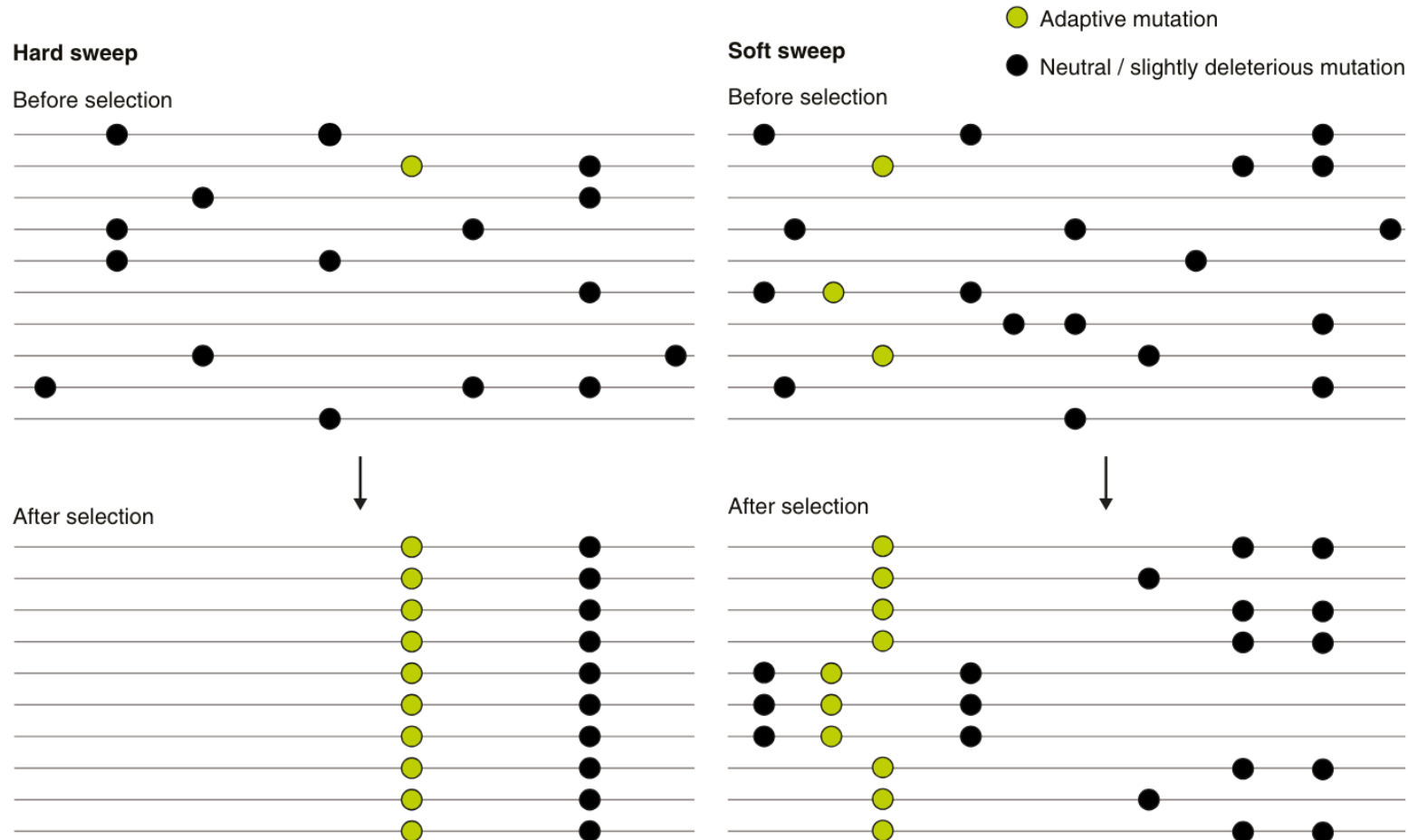
Positive selection

> Increases frequency of beneficial mutations

One mutation on one genomic background ⟵ Hard vs. soft sweeps ⟶ Different mutations and/or different genomic backgrounds

Adaptive mutation
Neutral / slightly deleterious mutation

Hard sweep
Before selection
After selection

Soft sweep
Before selection
After selection

McCoy & Akey, 2017, Genome Biology

Balancing selection

Balancing selection maintains multiple alleles in the gene pool of a population at frequencies larger than expected from genetic drift alone.

Mechanisms:

- Heterozygote advantage

- Frequency-dependent selection

- Spatially and/or temporally varying selection

Signature of balancing selection

Diversity (pi)
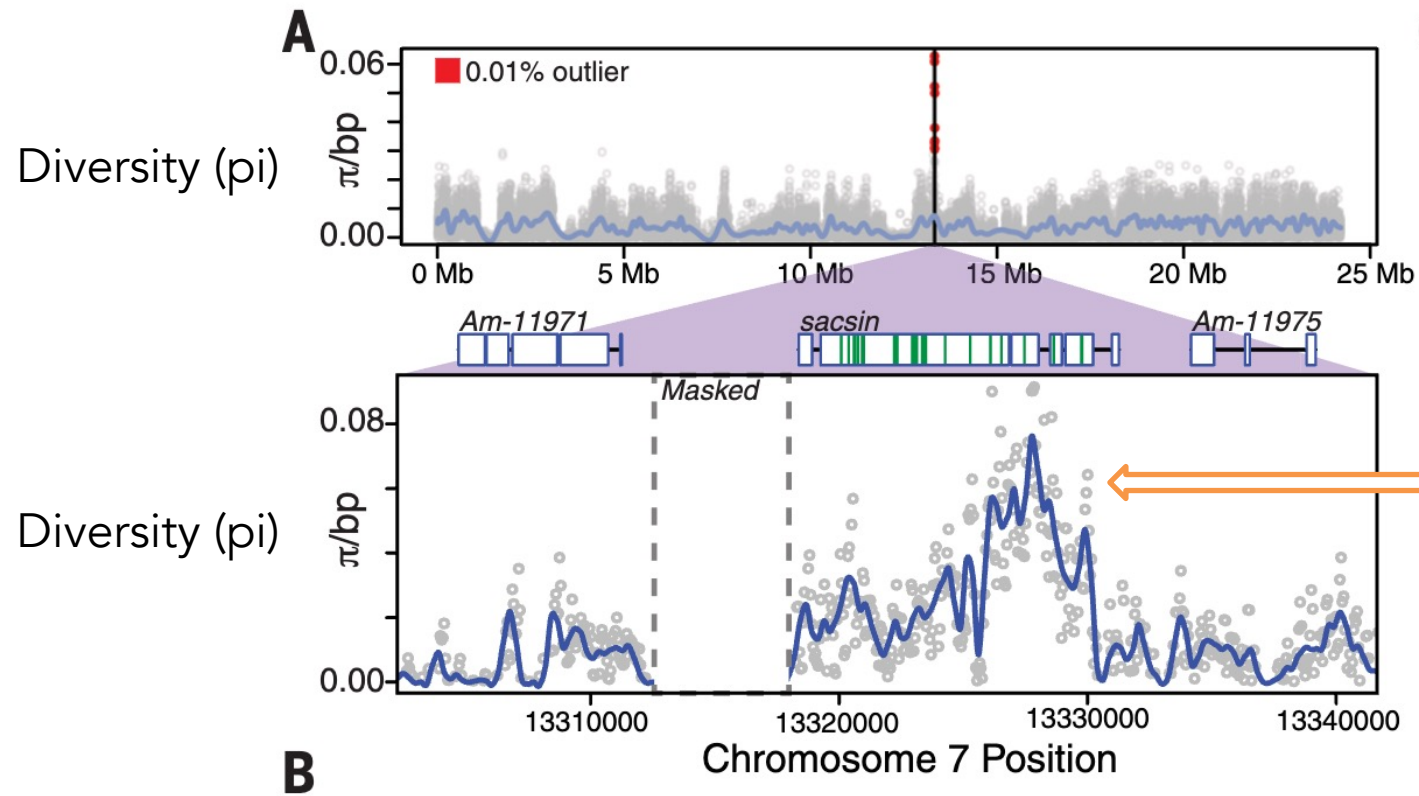
Diversity (pi)

Higher genetic diversity at the locus under balancing selection

Fuller et al. 2020, Science

Background selection

Negative selection = purifying selection = background selection

Selection against deleterious mutations

Reduces diversity around the deleterious mutations under selection

Protein-coding regions are under strict functional constraints

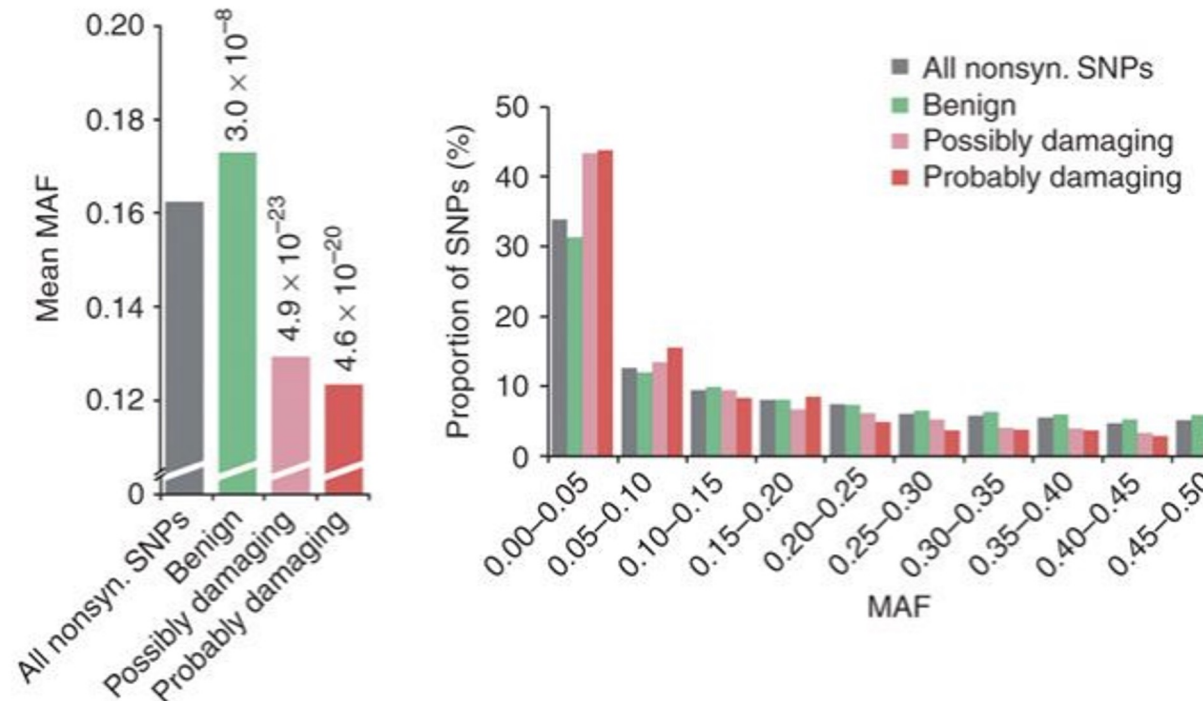# Background selection

Negative selection = purifying selection = background selection

Selection against deleterious mutations

Reduces diversity around the deleterious mutations under selection

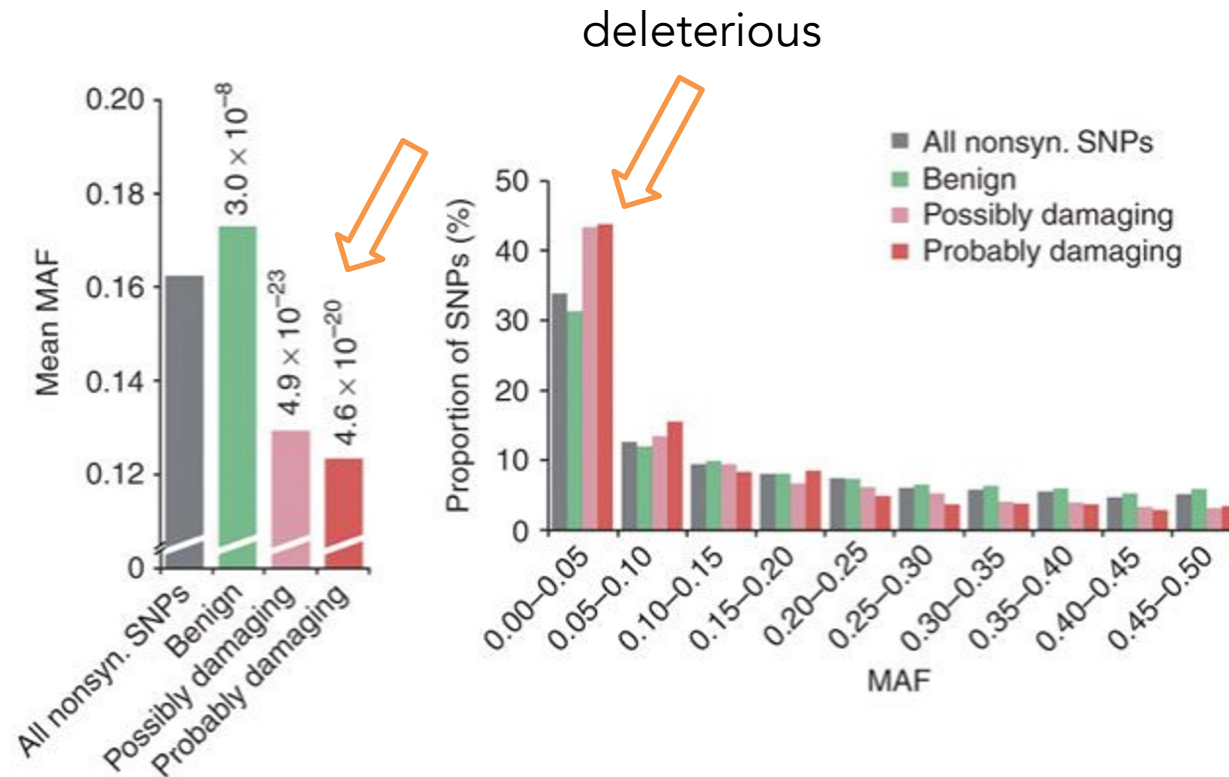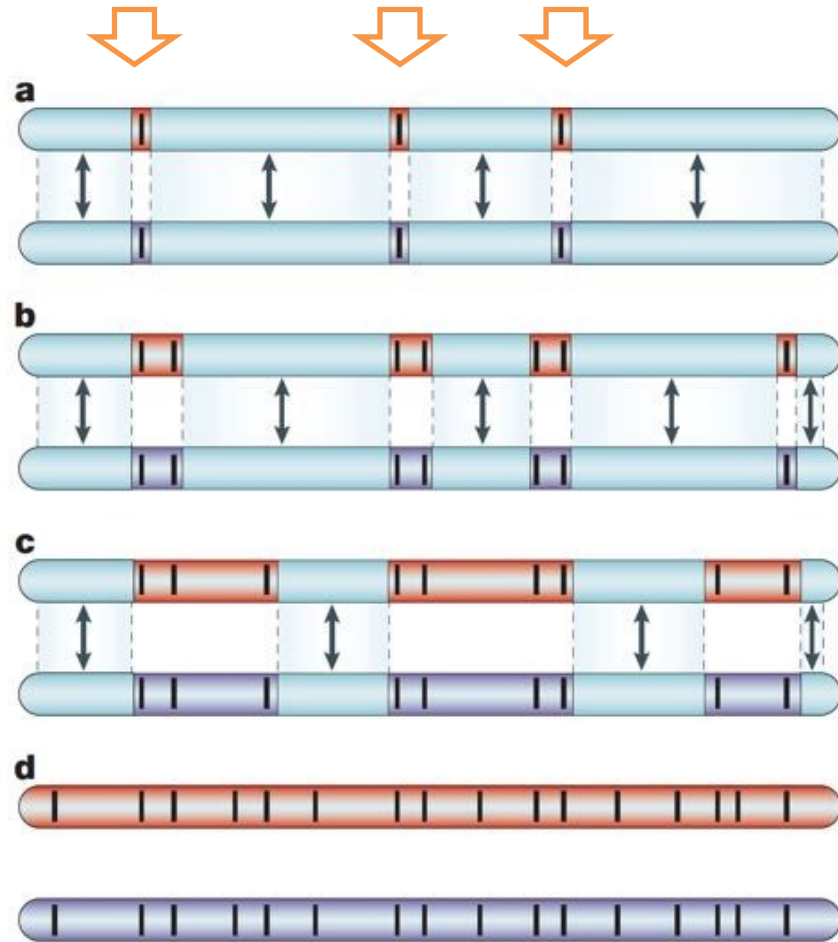Protein-coding regions are under strict functional constraints

Could background selection generate the same genomic islands of differentiation as positive selection?

Genomic islands of differentiation

Differentiation accumulates over time
Regions get longer

The speciation-with-gene-flow model is based on the genic view of speciation by Wu (2001)

Wu and Ting 2004, Nat. Gen. Rev.

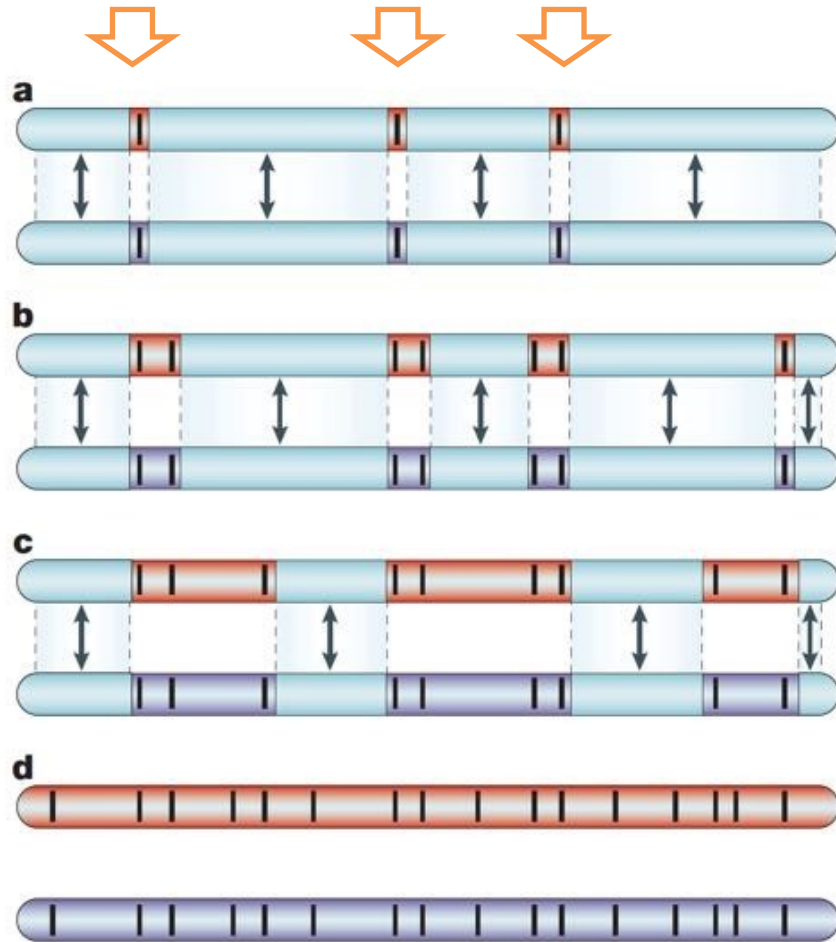# Genomic islands of differentiation

Differentiation accumulates over time
Regions get longer

Wu and Ting 2004, Nat. Gen. Rev.

The speciation-with-gene-flow model is based on the genic view of speciation by Wu (2001)

Speciation/adaptation genes

Comparative analysis examining patterns of genomic differentiation across multiple episodes of population divergence in birds

Kira E. Delmore,[1,2] Juan S. Lugo Ramos,[1] Benjamin M. Van Doren,[3] Max Lundberg,[4] Staffan Bensch,[4] Darren E. Irwin,[5] and Miriam Liedvogel[1]

Genomewide patterns of variation in genetic diversity are shared among populations, species and higher-order taxa

Nagarjun Vijay[1,2] | Matthias Weissensteiner[1,3] | Reto Burri[1,4] | Takeshi Kawakami[1,5] | Hans Ellegren[1] | Jochen B. W. Wolf[1,3]

Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of *Ficedula* flycatchers

Reto Burri,[1] Alexander Nater,[1] Takeshi Kawakami,[1] Carina F. Mugal,[1] Pall I. Olason,[2] Linnea Smeds,[1] Alexander Suh,[1] Ludovic Dutoit,[1] Stanislav Bureš,[3] Laszlo Z. Garamszegi,[4] Silje Hogner,[5,6] Juan Moreno,[7] Anna Qvarnström,[8] Milan Ružić,[9] Stein-Are Sæther,[5,10] Glenn-Peter Sætre,[5] Janos Török,[11] and Hans Ellegren[1]

Correlated patterns of genetic diversity and differentiation across an avian family

BENJAMIN M. VAN DOREN,*† LEONARDO CAMPAGNA,*† BARBARA HELM,‡ JUAN CARLOS ILLERA,§ IRBY J. LOVETTE*† and MIRIAM LIEDVOGEL¶
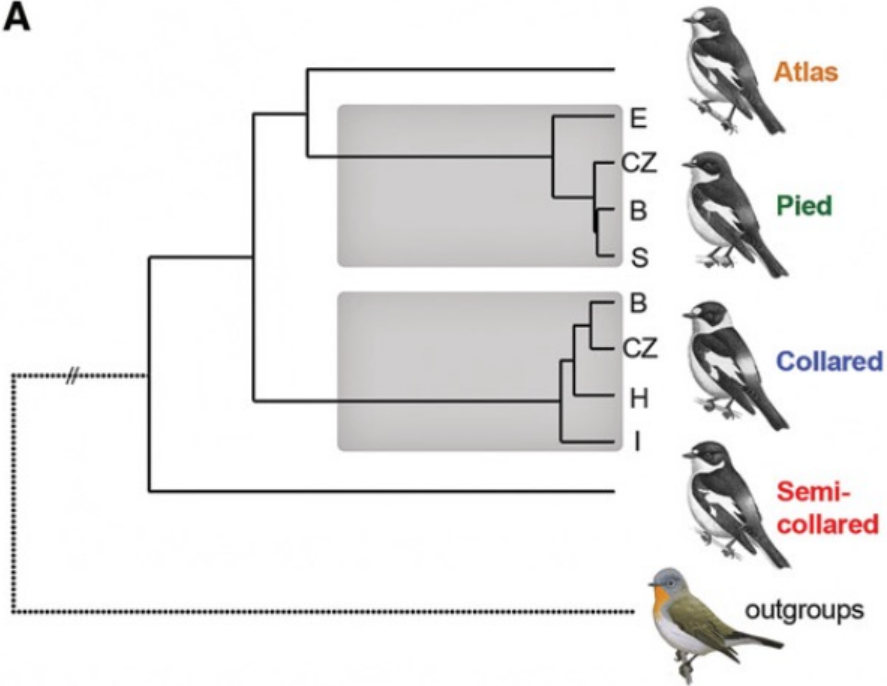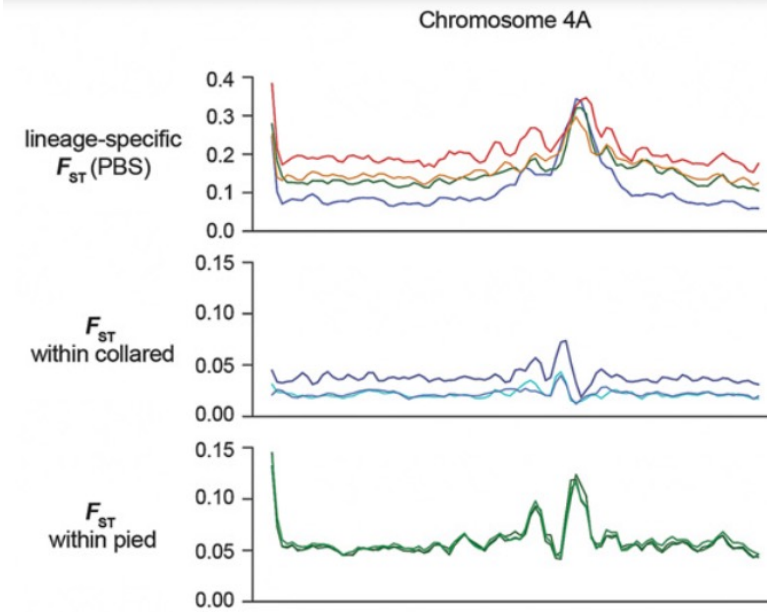
A comparison of genomic islands of differentiation across three young avian species pairs

Darren E. Irwin[1] | Borja Milá[2] | David P. L. Toews[1,3] | Alan Brelsford[1,4] | Haley L. Kenyon[1,5] | Alison N. Porter[1] | Christine Grossen[1,6] | Kira E. Delmore[1,7] | Miguel Alcaide[1,8] | Jessica H. Irwin[1]

Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of *Ficedula* flycatchers

Reto Burri,[1] Alexander Nater,[1] Takeshi Kawakami,[1] Carina F. Mugal,[1] Pall I. Olason,[2]
Linnea Smeds,[1] Alexander Suh,[1] Ludovic Dutoit,[1] Stanislav Bureš,[3]
Laszlo Z. Garamszegi,[4] Silje Hogner,[5,6] Juan Moreno,[7] Anna Qvarnström,[8]
Milan Ružić,[9] Stein-Are Sæther,[5,10] Glenn-Peter Sætre,[5] Janos Török,[11]
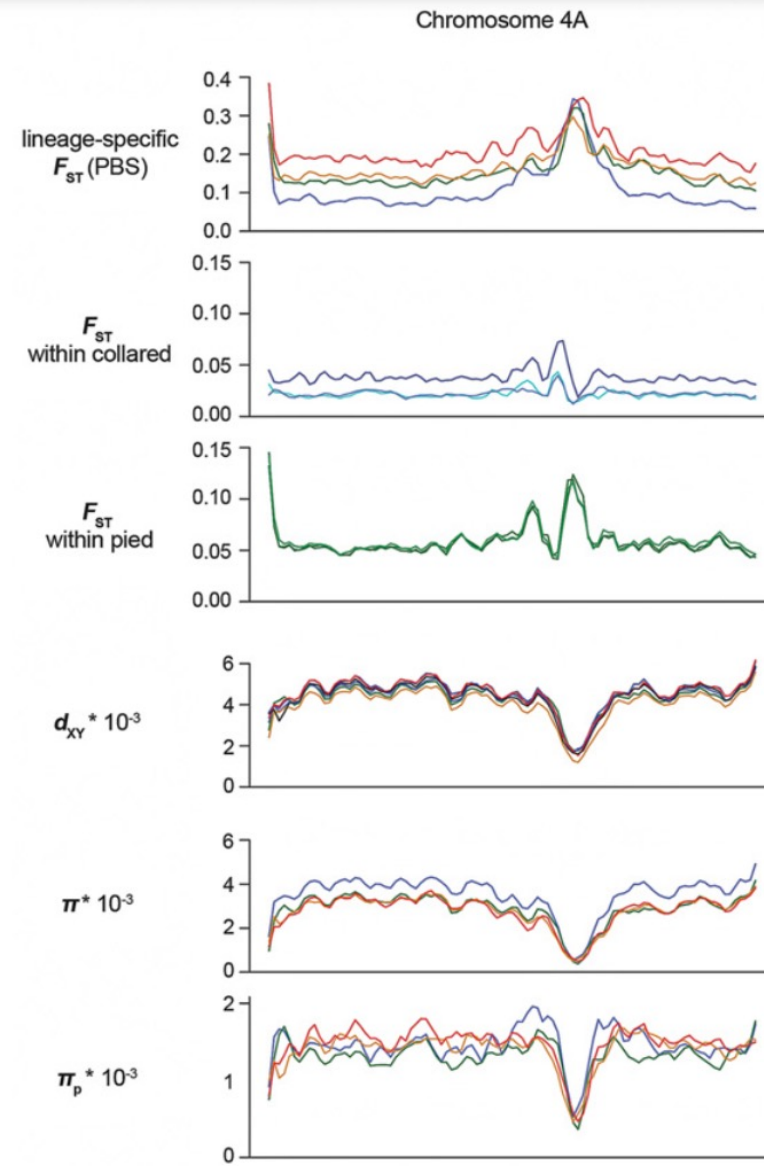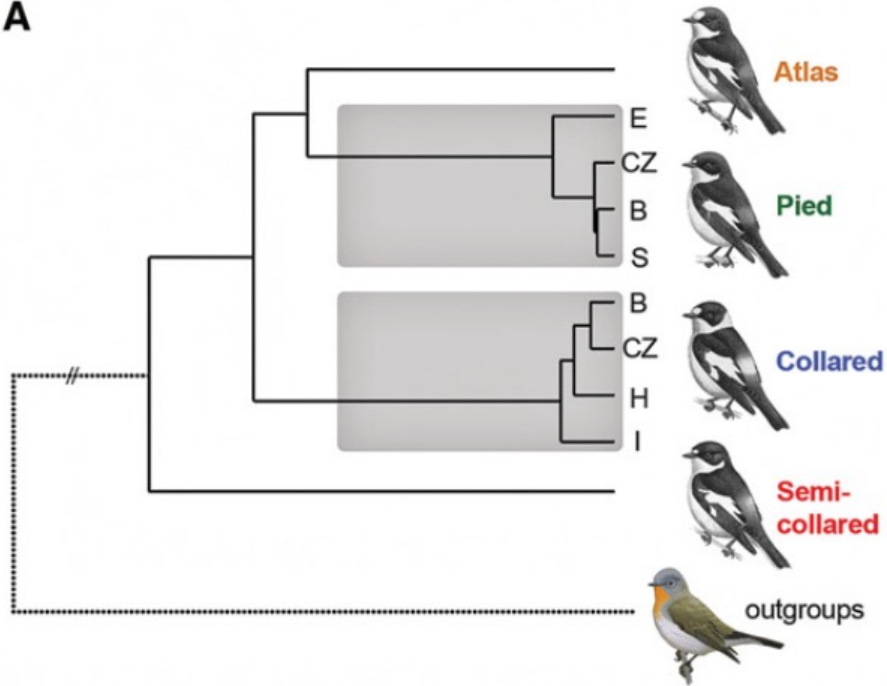and Hans Ellegren[1]

Does this mean that all these species/populations comparisons share the same 'speciation genes"?

Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of *Ficedula* flycatchers

Reto Burri,[1] Alexander Nater,[1] Takeshi Kawakami,[1] Carina F. Mugal,[1] Pall I. Olason,[2]
Linnea Smeds,[1] Alexander Suh,[1] Ludovic Dutoit,[1] Stanislav Bureš,[3]
Laszlo Z. Garamszegi,[4] Silje Hogner,[5,6] Juan Moreno,[7] Anna Qvarnström,[8]
Milan Ružić,[9] Stein-Are Sæther,[5,10] Glenn-Peter Sætre,[5] Janos Török,[11]
and Hans Ellegren[1]

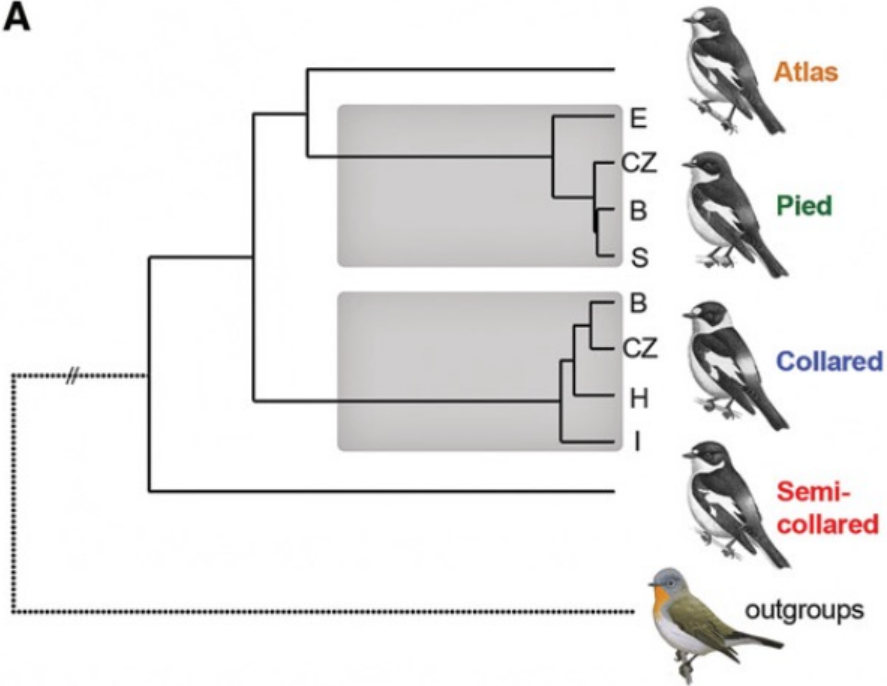But sequence divergence (d_{xy}) is actually reduced here…

And nucleotide diversity, too

Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of *Ficedula* flycatchers

Reto Burri,[1] Alexander Nater,[1] Takeshi Kawakami,[1] Carina F. Mugal,[1] Pall I. Olason,[2] Linnea Smeds,[1] Alexander Suh,[1] Ludovic Dutoit,[1] Stanislav Bureš,[3] Laszlo Z. Garamszegi,[4] Silje Hogner,[5,6] Juan Moreno,[7] Anna Qvarnström,[8] Milan Ružić,[9] Stein-Are Sæther,[5,10] Glenn-Peter Sætre,[5] Janos Török,[11] and Hans Ellegren[1]
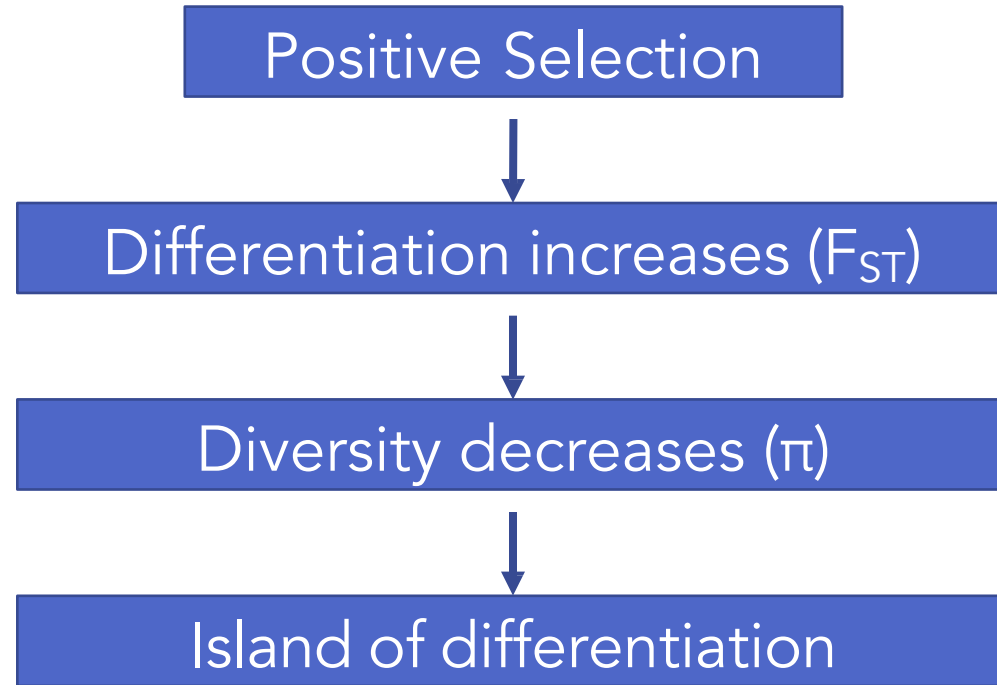
A

Recombination is reduced in the same area.
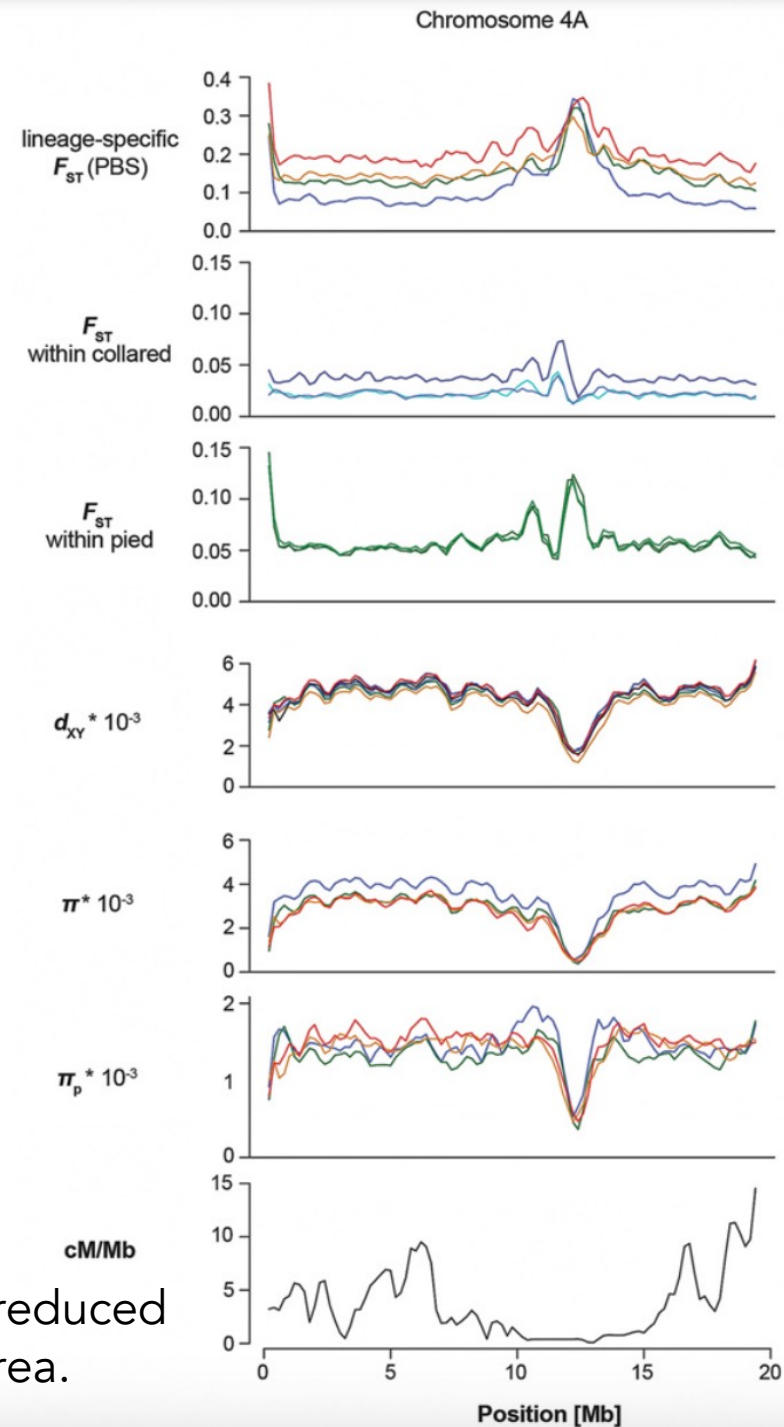
Adaptation/speciation with gene flow

Positive Selection

↓

Differentiation increases (F_{ST})

↓

Diversity decreases (π)

↓

Island of differentiation

Or background selection without gene flow?

Positive selection?

$F_{ST} = (\pi_T - \pi_S)/\pi_T$

F_{ST} is affected by levels of standing genetic variation:
- It is inflated in areas of low diversity
- It is low when diversity is high

Differentiation increases (F_{ST})

Low recombination reduces diversity:
1. Recombination introduces diversity
2. Recombination reduces the diversity-reducing effect of positive and background selection

Diversity decreases (π)

Island of differentiation

# Absolute measures of divergence

$D_{xy} =$
average number of pairwise differences between sequences from two populations, excluding all comparisons between sequences within populations.

It is independent of the levels of diversity within the two populations examined.

However, it is affected by ancestral levels of diversity and mutation rate.

Genome structure

Variation in recombination rate across the genome affects patterns of diversity, differentiation, and divergence


Large inversions (> 1Mb)
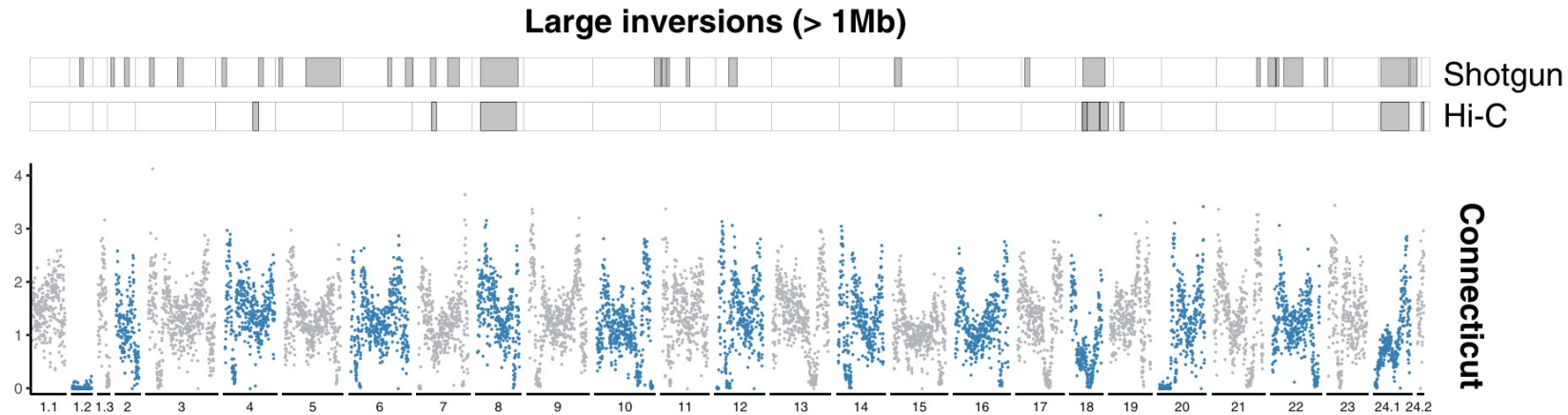
π varies
- **within** chromosome > higher towards the telomeres
- **across** chromosomes > higher in smaller chromosomes
- due to **structural variants >** for their effect on recombination

Tigano et al. 2022, GBE

Genome structure

Variation in recombination rate across the genome affects patterns of diversity, differentiation, and divergence

North vs. South
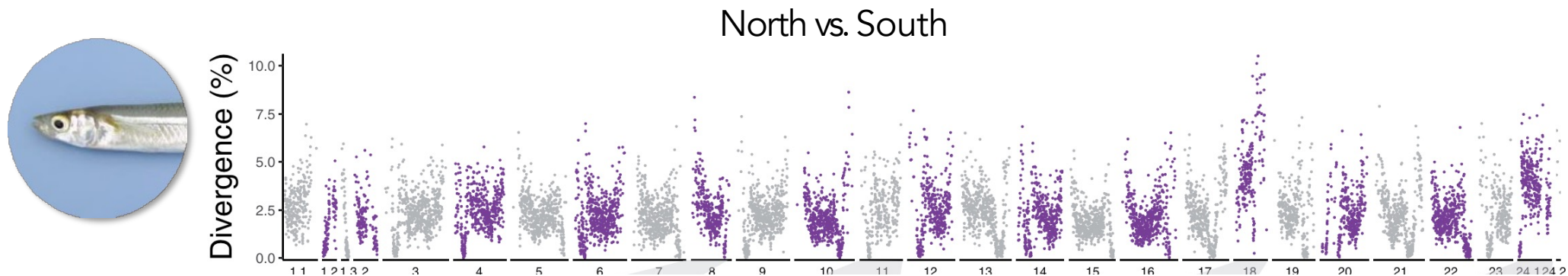
Effect of variation in diversity on divergence > elevated divergence in areas of high diversity, and vice-versa

Tigano et al. 2022, GBE

Structural variation

Clinal analysis and phenotypic to genotype associations support inversions underlying adaptation

But...

SVs (including inversions) can be maintained as polymorphism within populations or species by neutral processes.

Structural variation

How structural variation can confound signatures of selection.

- Recombination suppression > elevated differentiation due to low recombination rather than selection

Structural variation

How structural variation can confound signatures of selection.

• Recombination suppression > elevated differentiation due to low   recombination rather than selection

• Collapse of paralog sequences (Copy Number Variants) > detection of excess of polymorphism and heterozygosity

Structural variation

How structural variation can confound signatures of selection.

- Recombination suppression > elevated differentiation due to low recombination rather than selection

- Collapse of paralog sequences (Copy Number Variants) > detection of excess of polymorphism and heterozygosity

- Low quality assembly and mapping > bias genotype calls

The end.