

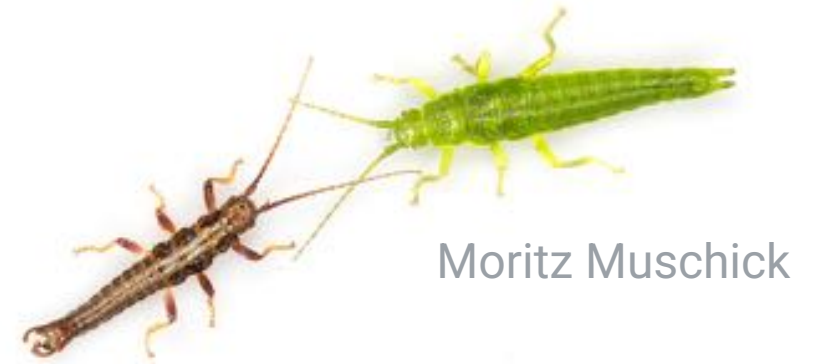
Confounding factors

Day 3 - Lecture 2

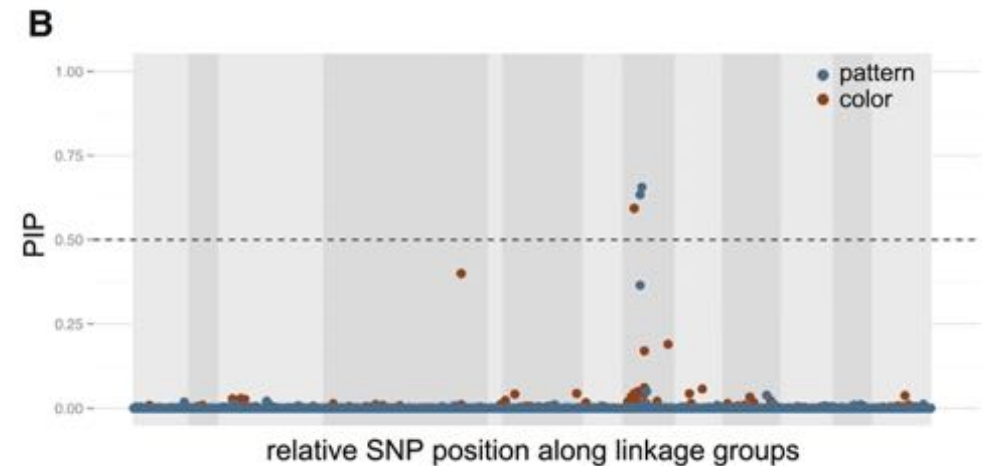
Anna Tigano, Ph.D.

Expectation

Selection increases the frequency of beneficial mutations, so that when we compare two populations adapted to different local conditions we should see a difference in frequency at the locus underpinning that adaptive trait.



Moritz Muschick



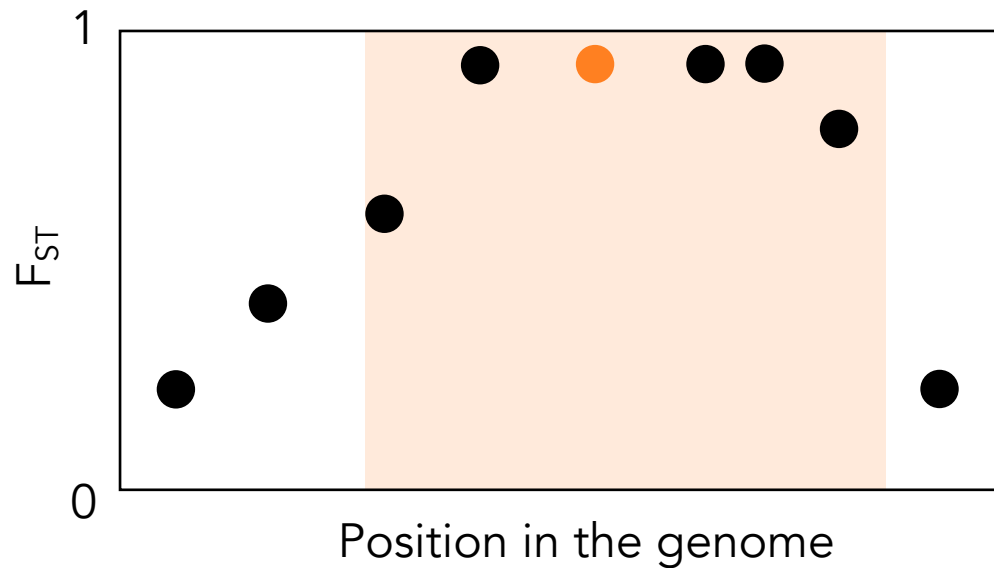
Comeault et al. 2015, Current Biology

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.

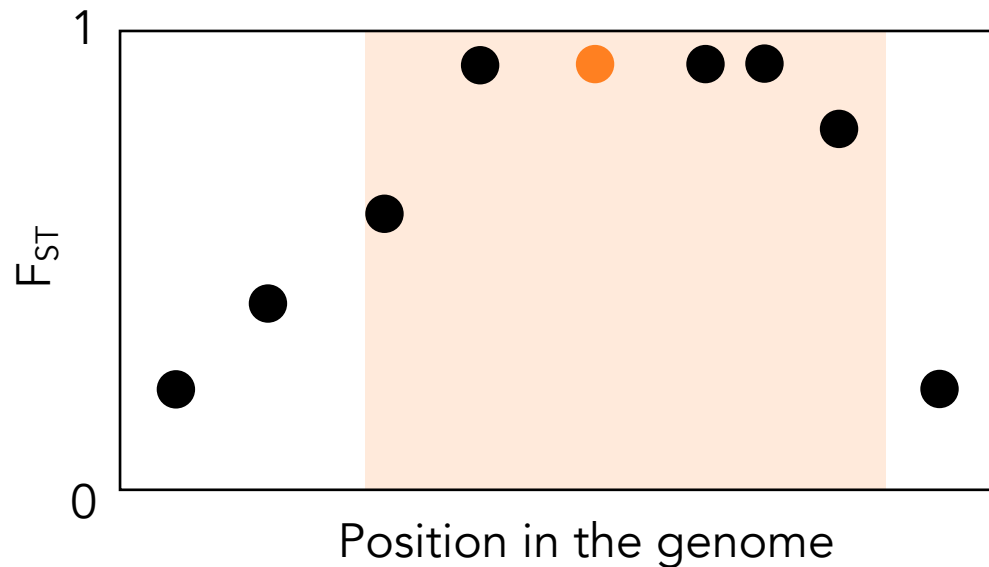
Linked selection

Due to linked selection, several sites in close proximity to the real target of selection could show high F_{ST} .



Linked selection

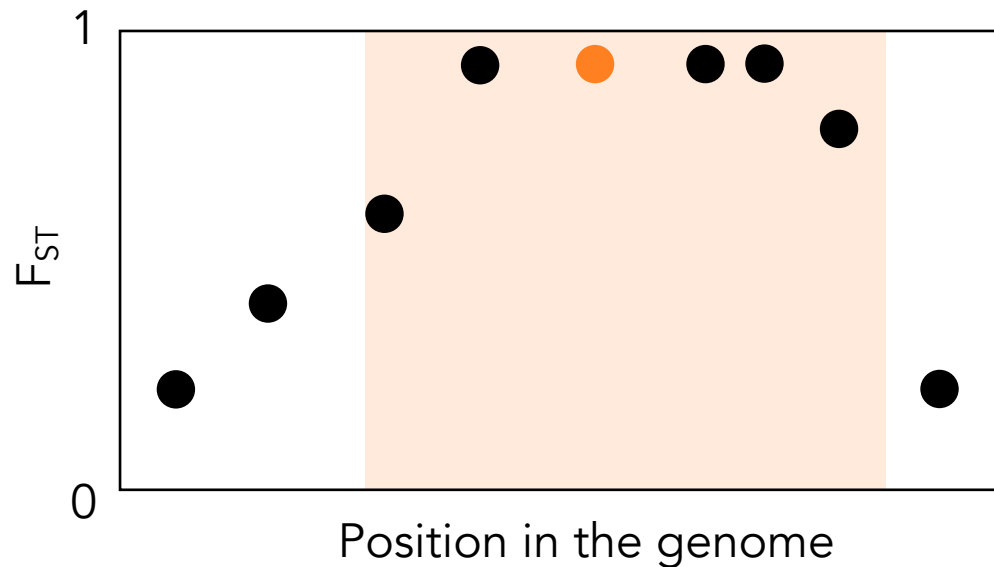
Due to linked selection, several sites in close proximity to the real target of selection could show high F_{ST} .



With whole genome data, it is likely that you will genotype all, or most, of these sites with high F_{ST} .

Linked selection

Due to linked selection, several sites in close proximity to the real target of selection could show high F_{ST} .

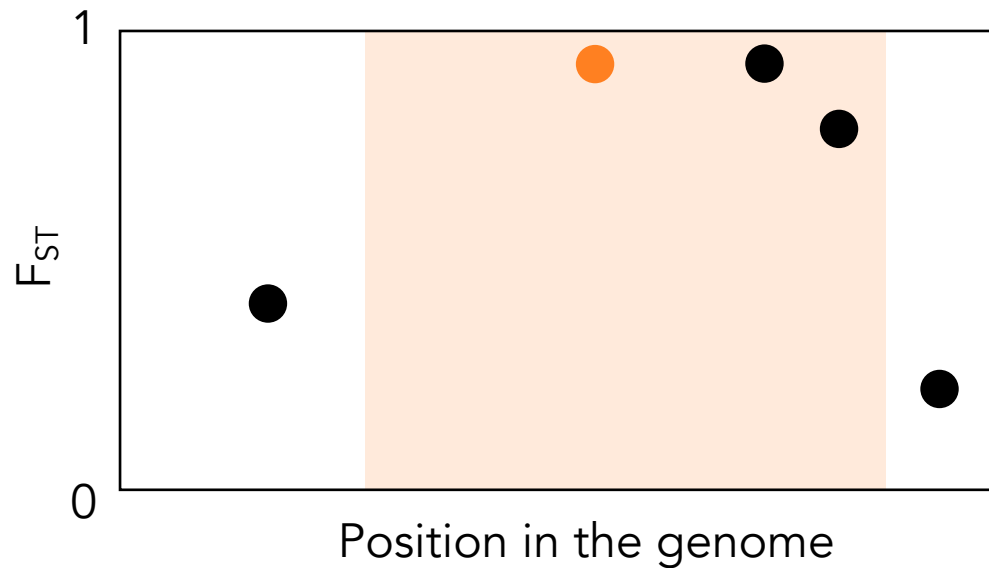


With whole genome data, it is likely that you will genotype all, or most, of these sites with high F_{ST} .

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

Linked selection

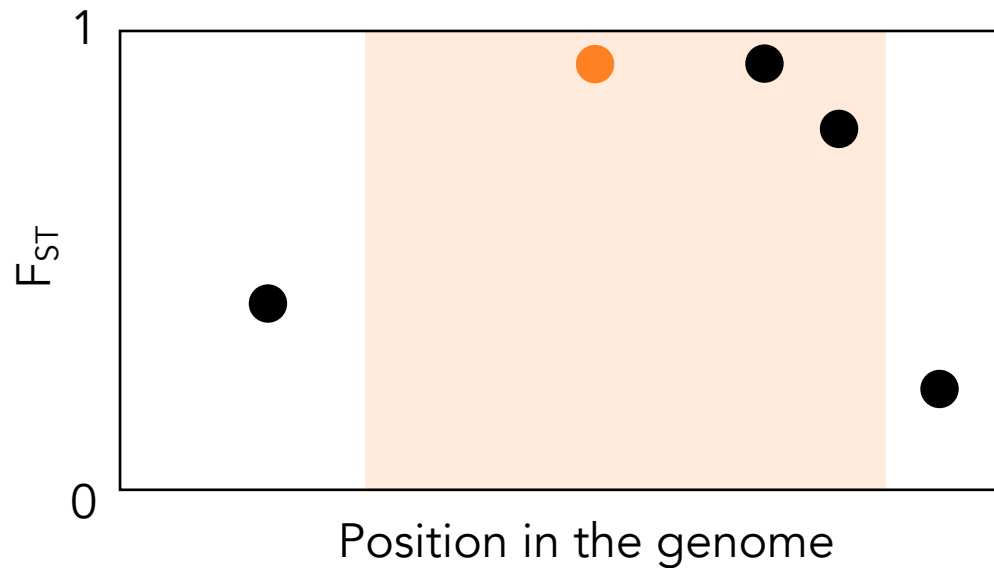
Due to linked selection, several sites in close proximity to the real target of selection could show high F_{ST} .



With reduced-representation approaches, you won't.

Linked selection

Due to linked selection, several 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 sites linked to it.

Positive selection

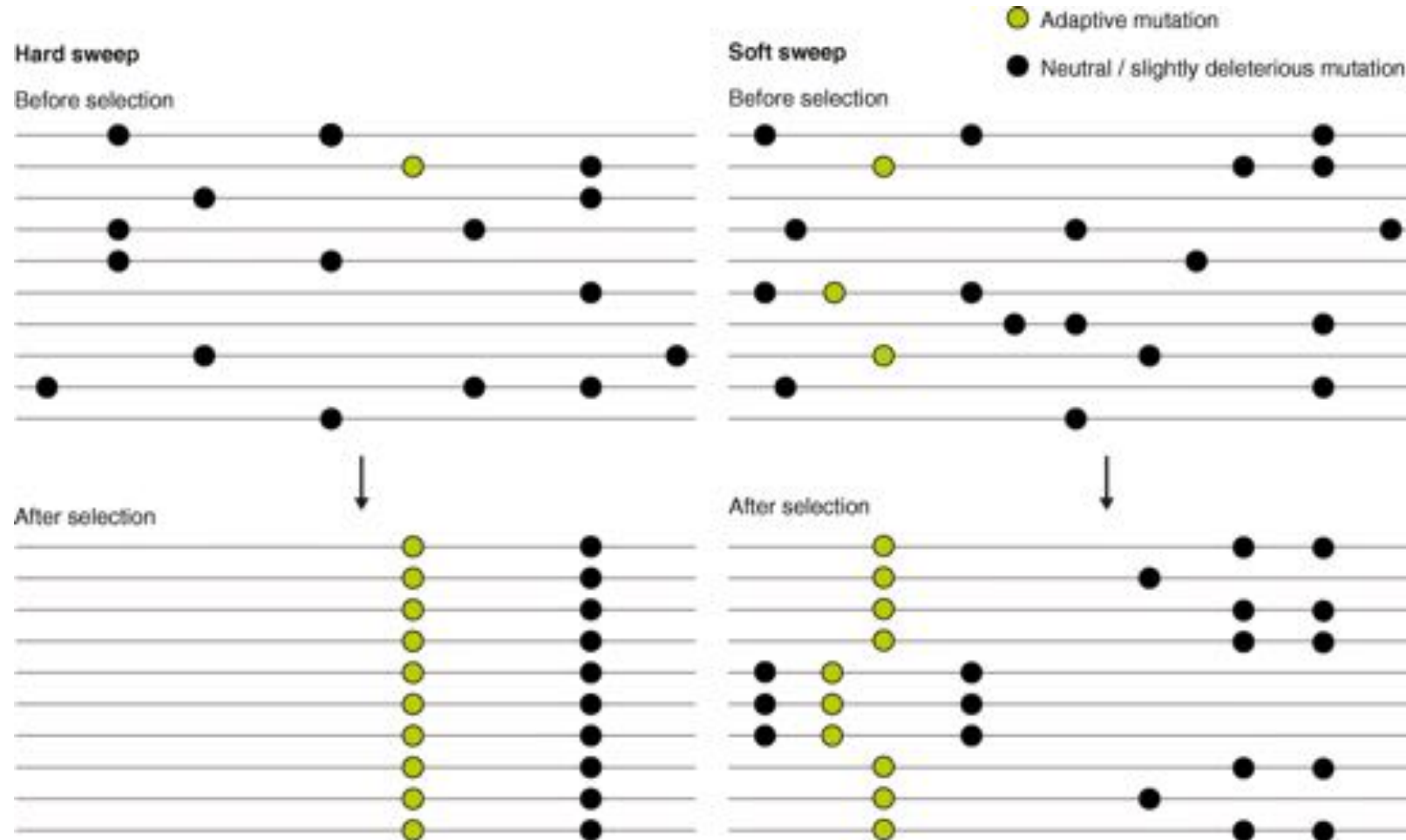
Natural selection = positive selection

→ Increases frequency of beneficial mutations

One mutation on one
genomic background

Hard vs. soft sweeps

Different mutations
and/or different
genomic backgrounds



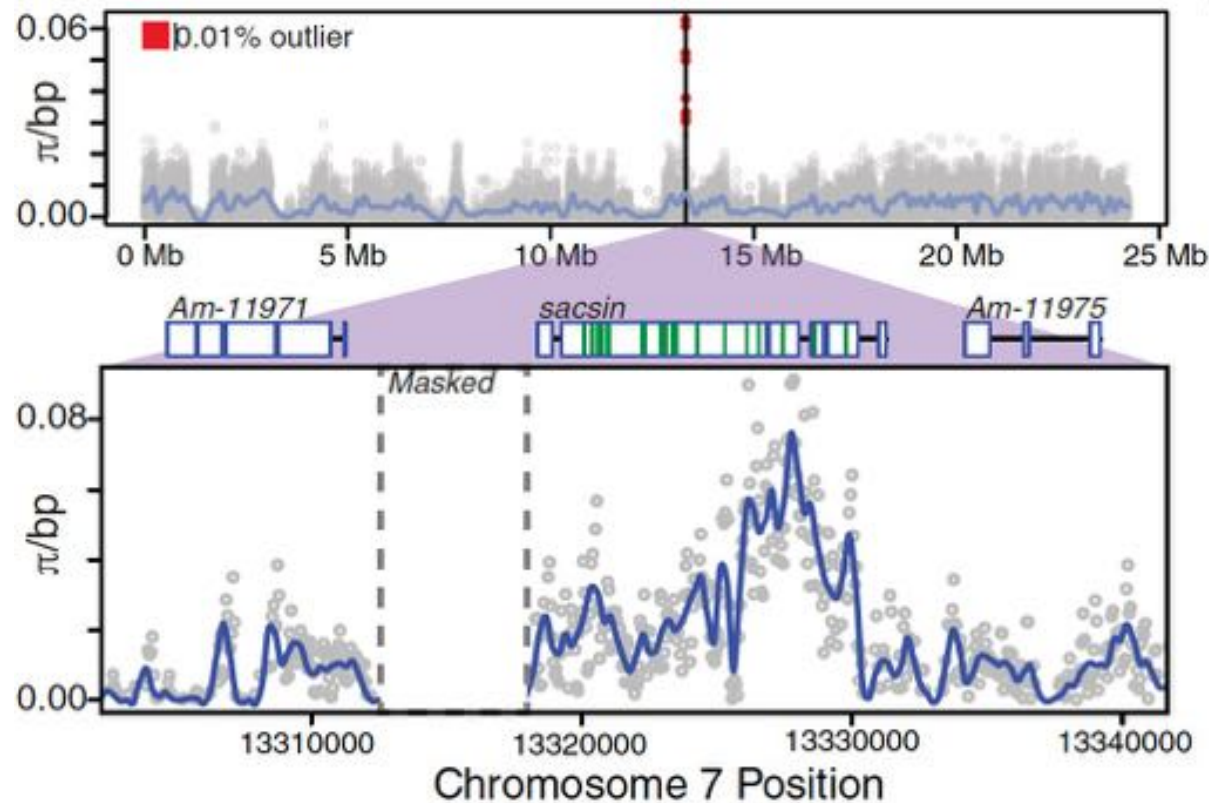
Balancing selection

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



Higher genetic diversity at the locus under balancing selection

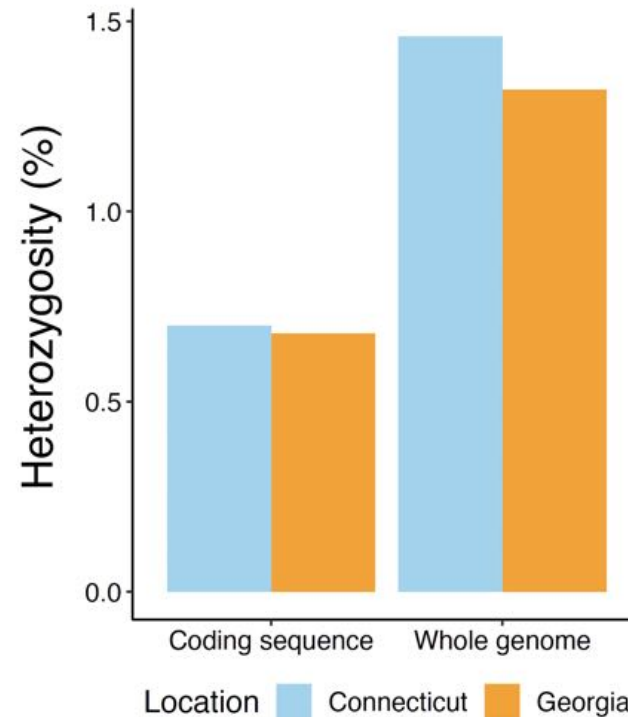
Background selection

Negative selection = purifying selection = background selection

→ Removes alleles that are deleterious

Similarly to positive selection, negative selection reduces diversity around the deleterious mutations under selection.

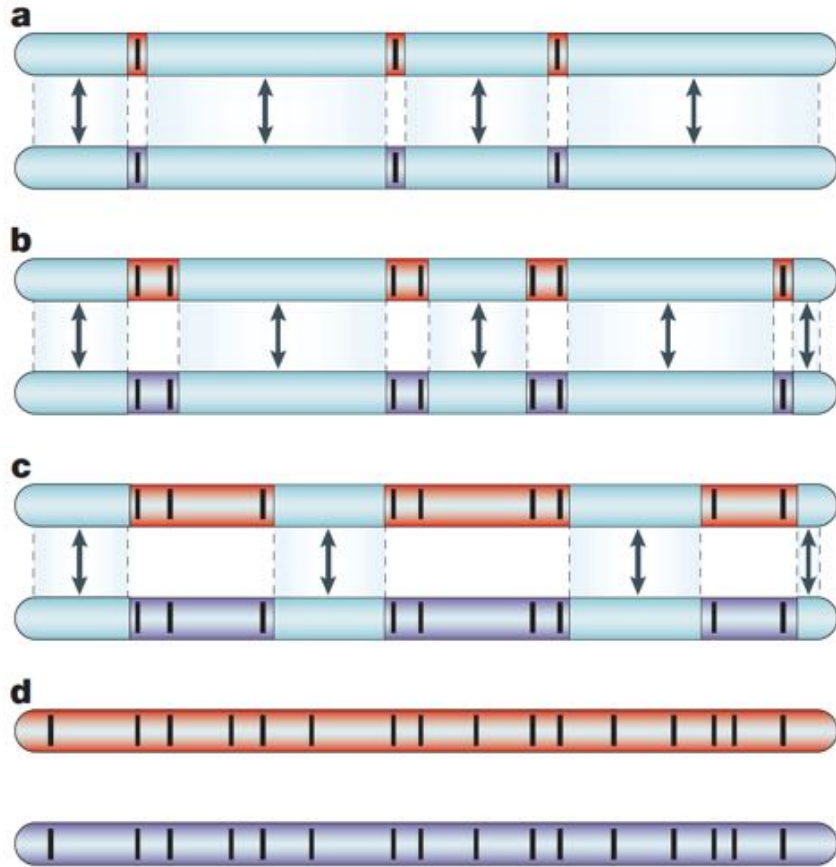
Protein-coding regions, for example, are under strict functional constraints and purifying selection is what keeps them functional.



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

Genomic islands of differentiation

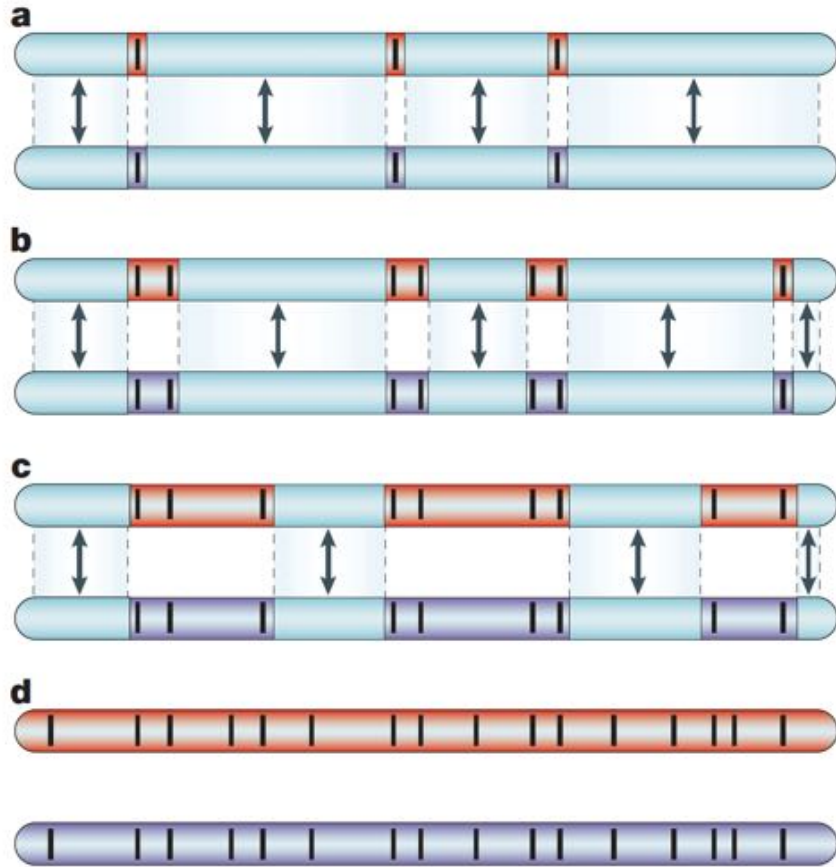
The speciation-with-gene-flow model is based on the genic view of speciation by Wu (2001)



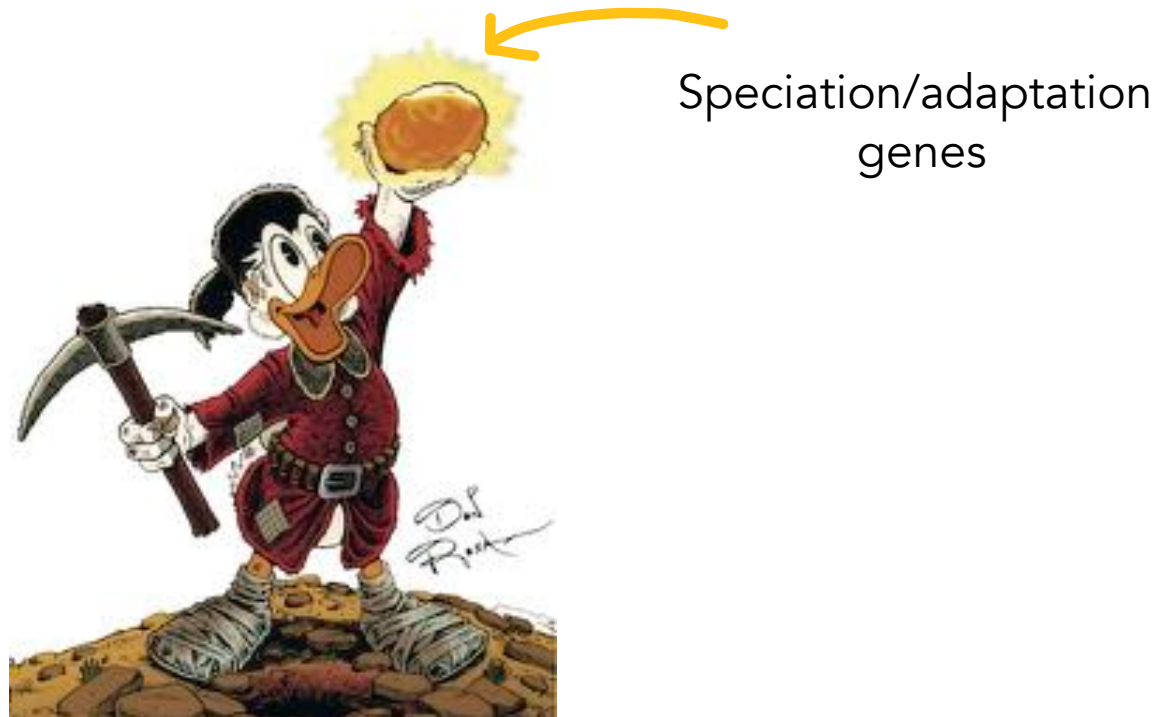
Wu and Ting 2004, Nat. Gen. Rev.

Genomic islands of differentiation

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



Wu and Ting 2004, Nat. Gen. Rev.






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 Liedvogel¹

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 Ellegren¹




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 Ellegren¹  | Jochen B. W. Wolf^{1,3}

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¶

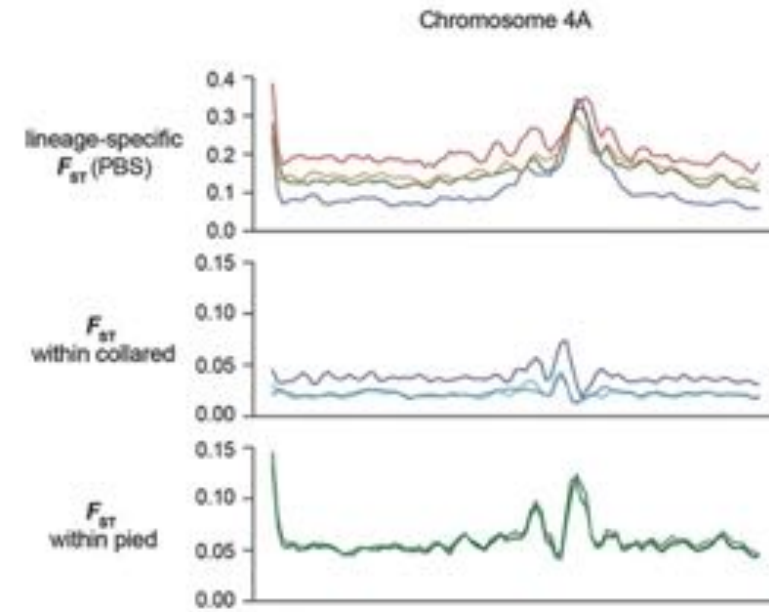
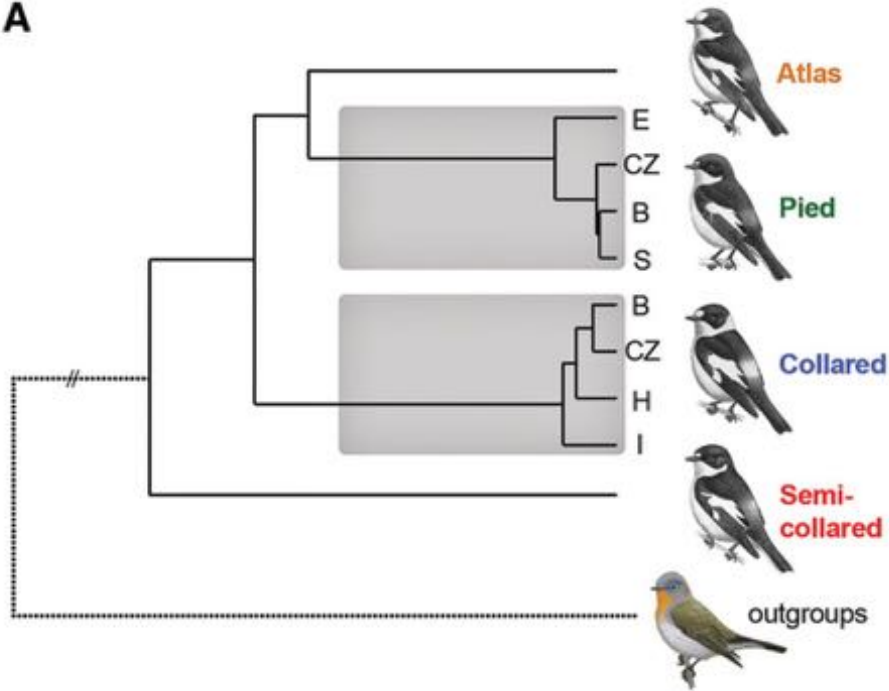
A comparison of genomic islands of differentiation across three young avian species pairs

Darren E. Irwin¹  | Borja Milá²  | David P. L. Toews^{1,3}  | Alan Brelsford^{1,4} | Haley L. Kenyon^{1,5} | Alison N. Porter¹ | Christine Gossen^{1,6} | Kira E. Delmore^{1,7} | Miguel Alcaide^{1,8} | Jessica H. Irwin¹

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 Ellegren¹

A

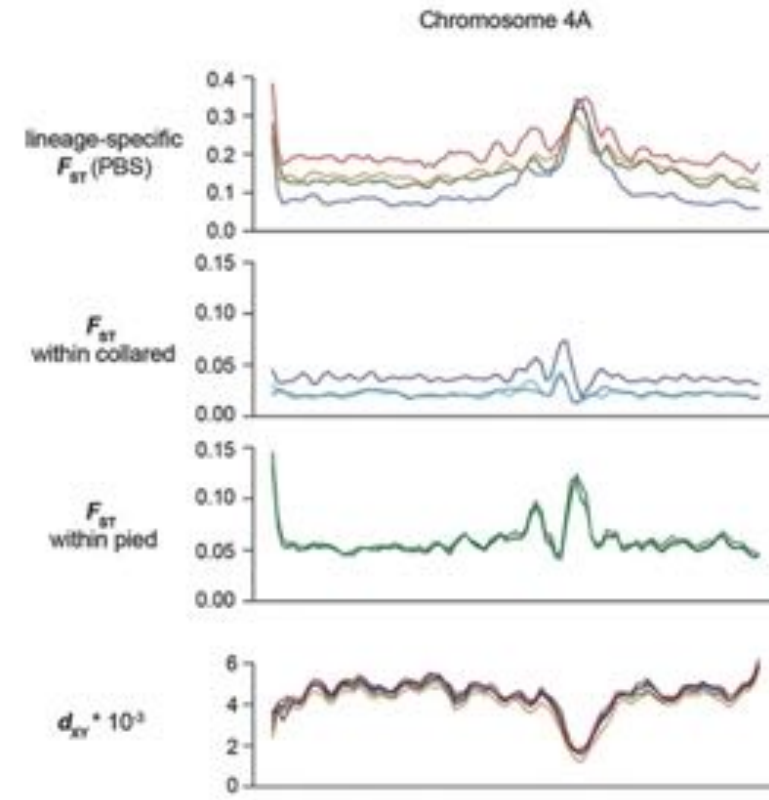
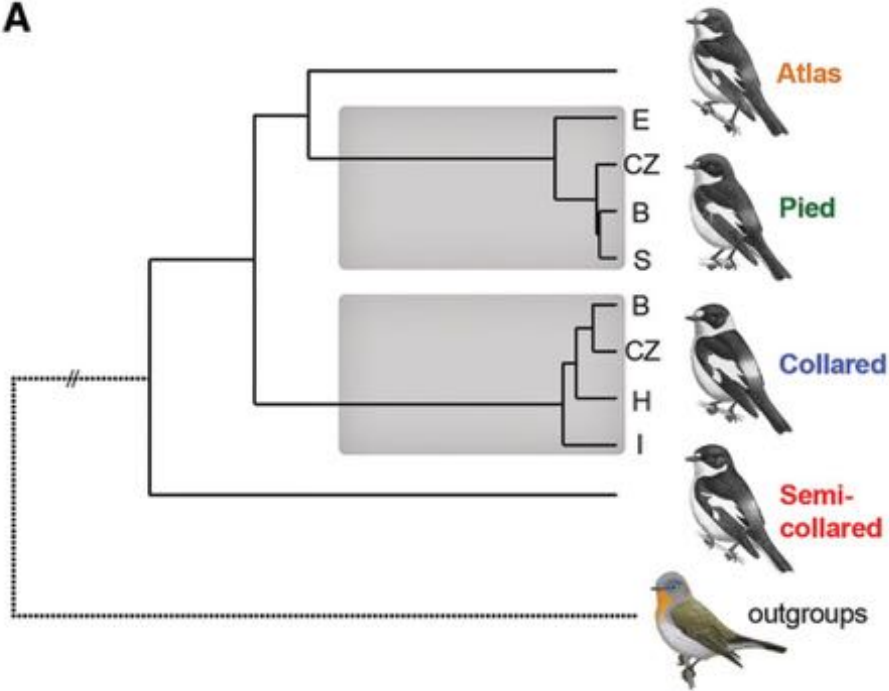


Does this mean that all these species/populations comparisons share the same 'speciation genes'?

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 Ellegren¹

A

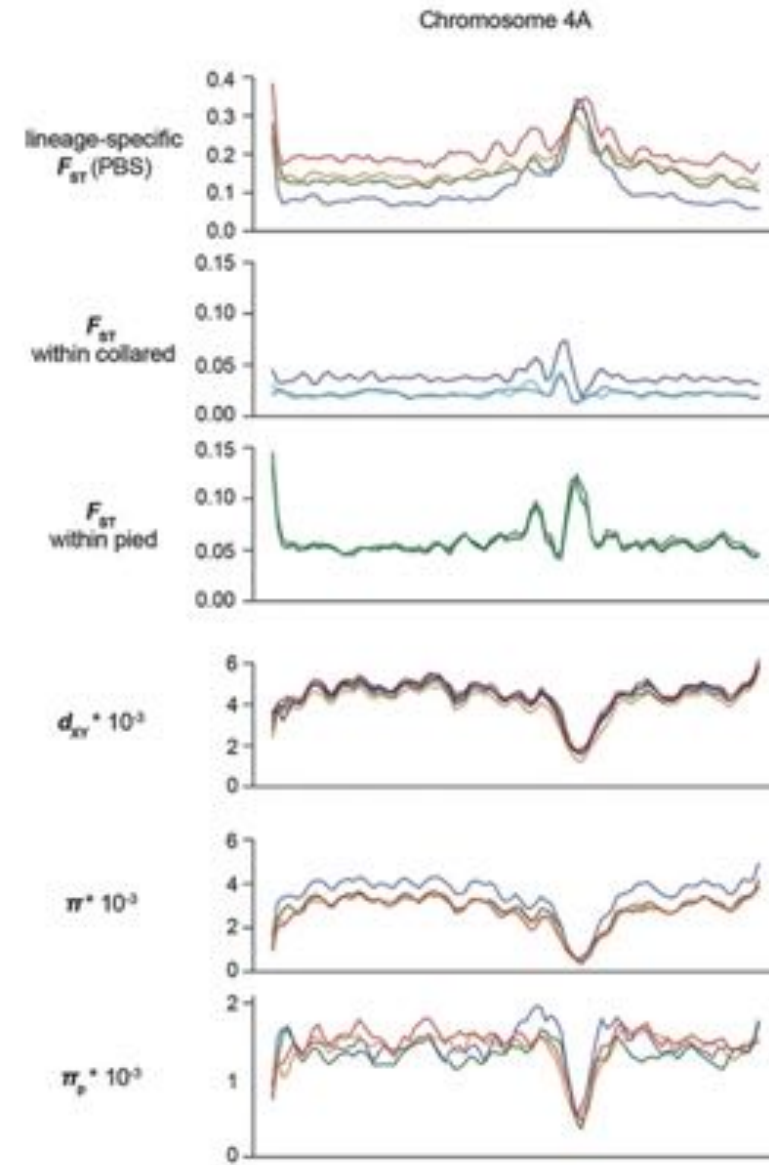
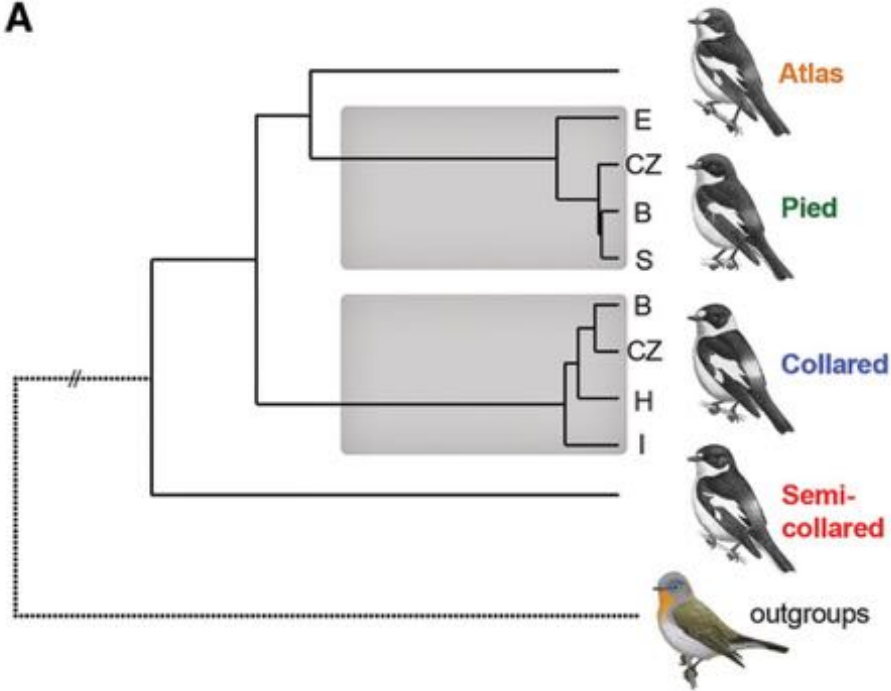


But sequence divergence is actually reduced here...

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 Ellegren¹

A

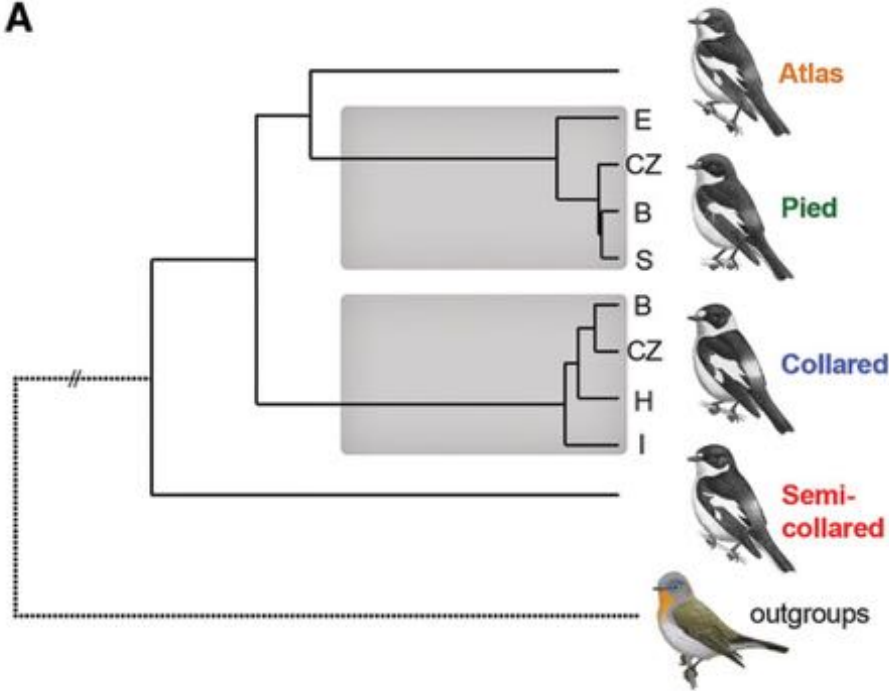


And nucleotide diversity, too.

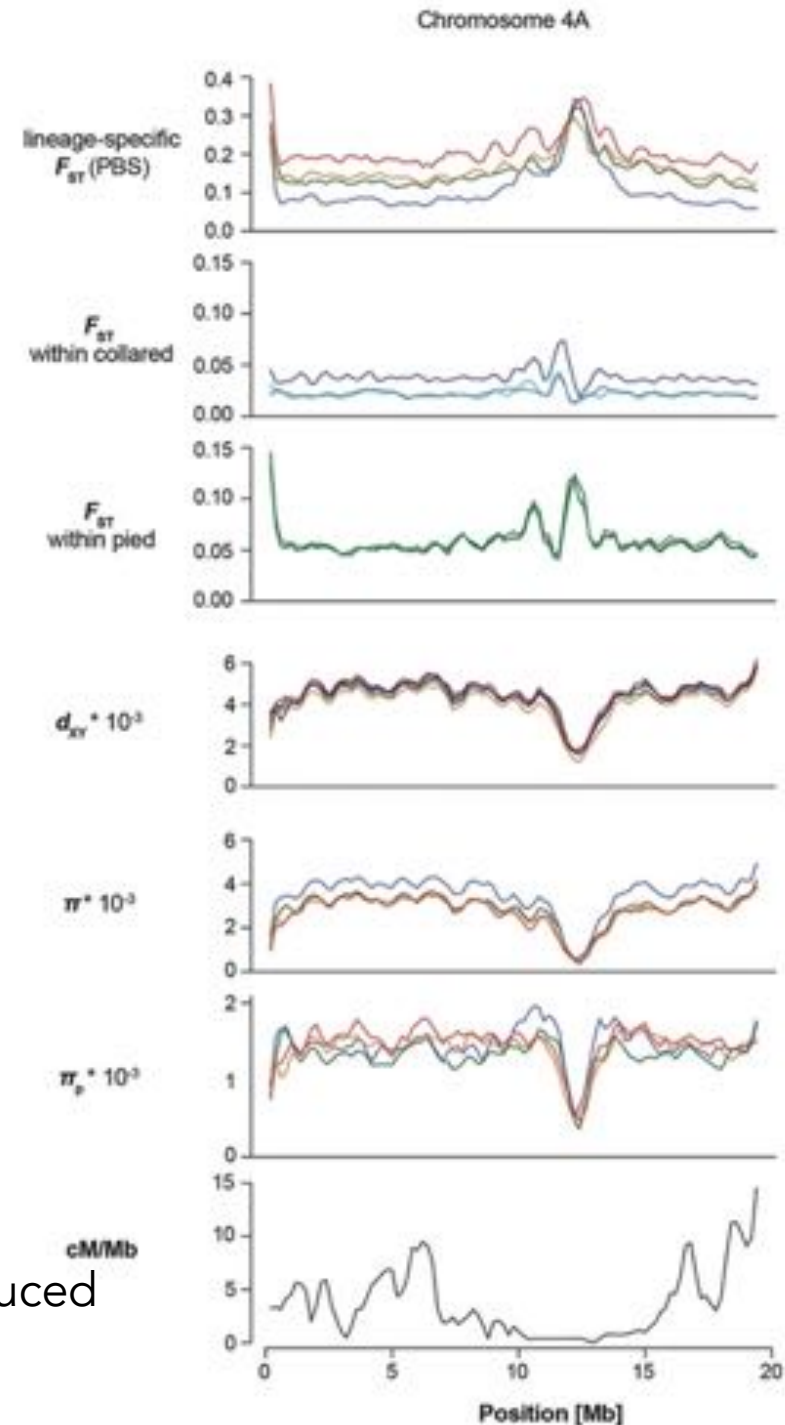
Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of *Ficedula* flycatchers

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 Ellegren¹

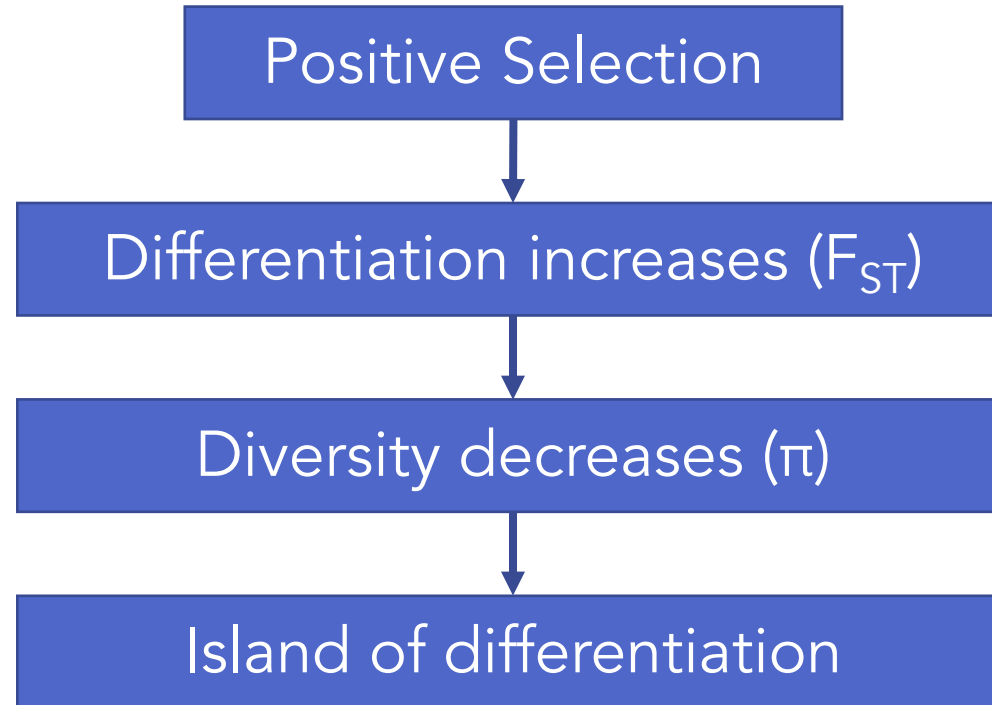
A



Recombination is reduced in the same area.



Adaptation/speciation with gene flow



Or background
selection without gene
flow?

Positive selection?

Differentiation increases (F_{ST})

Diversity decreases (π)

Island of differentiation

Low recombination
reduces diversity:

1. Recombination
introduces diversity
2. Recombination
reduces the
diversity-reducing
effect of positive and
background
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

Absolute measures of differentiation

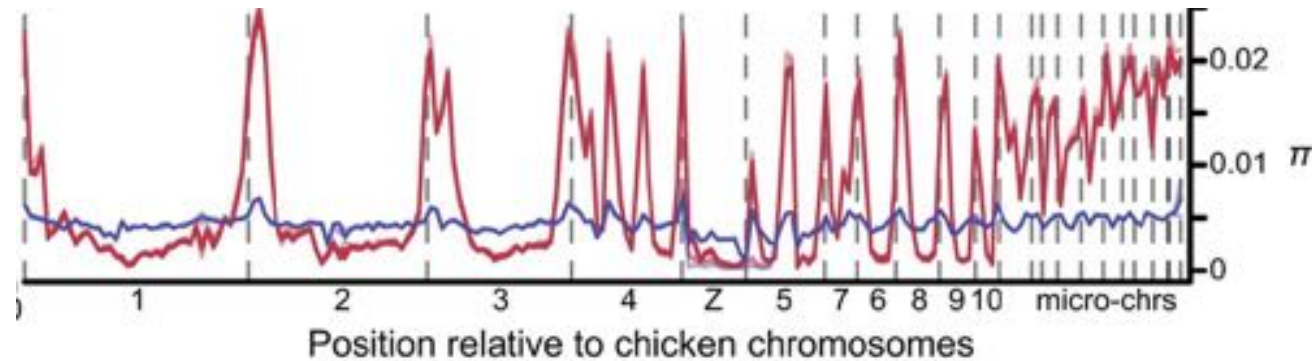
D_{xy} is the average number of pairwise differences between sequences from two populations, excluding all comparisons between sequences within populations.

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

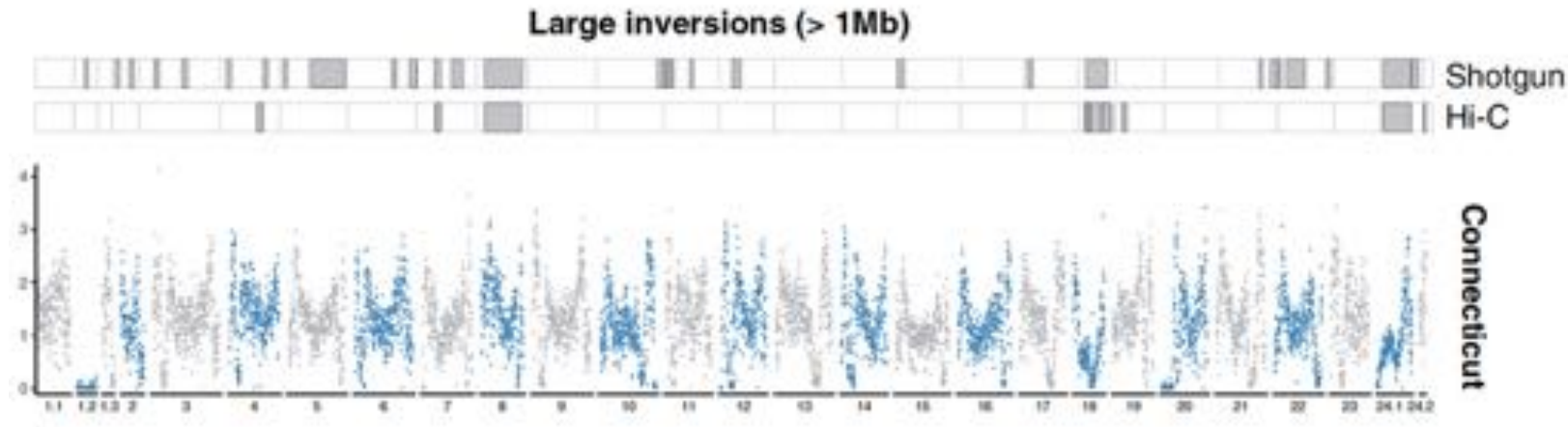


π varies

- **within** chromosome → higher towards the telomeres
- **across** chromosomes → higher in smaller chromosomes

Genome structure

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

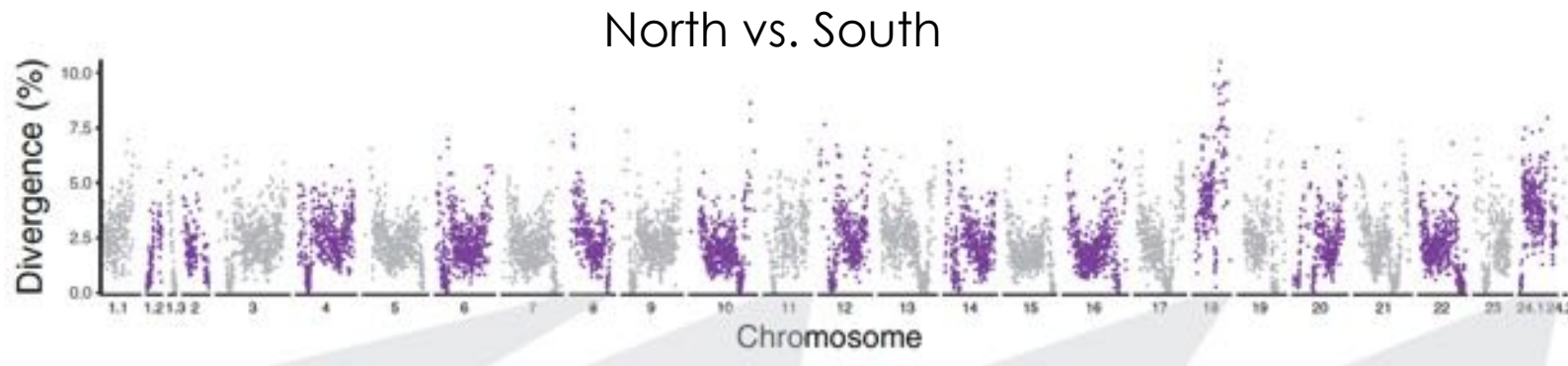


π varies

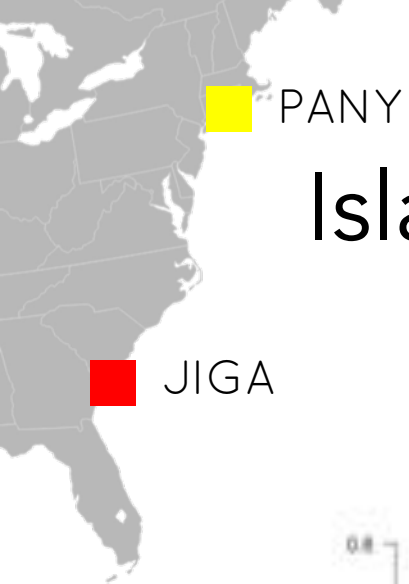
- **within** chromosome → higher towards the telomeres
- **across** chromosomes → higher in smaller chromosomes
- due to **structural variants** → for their effect on recombination

Genome structure

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

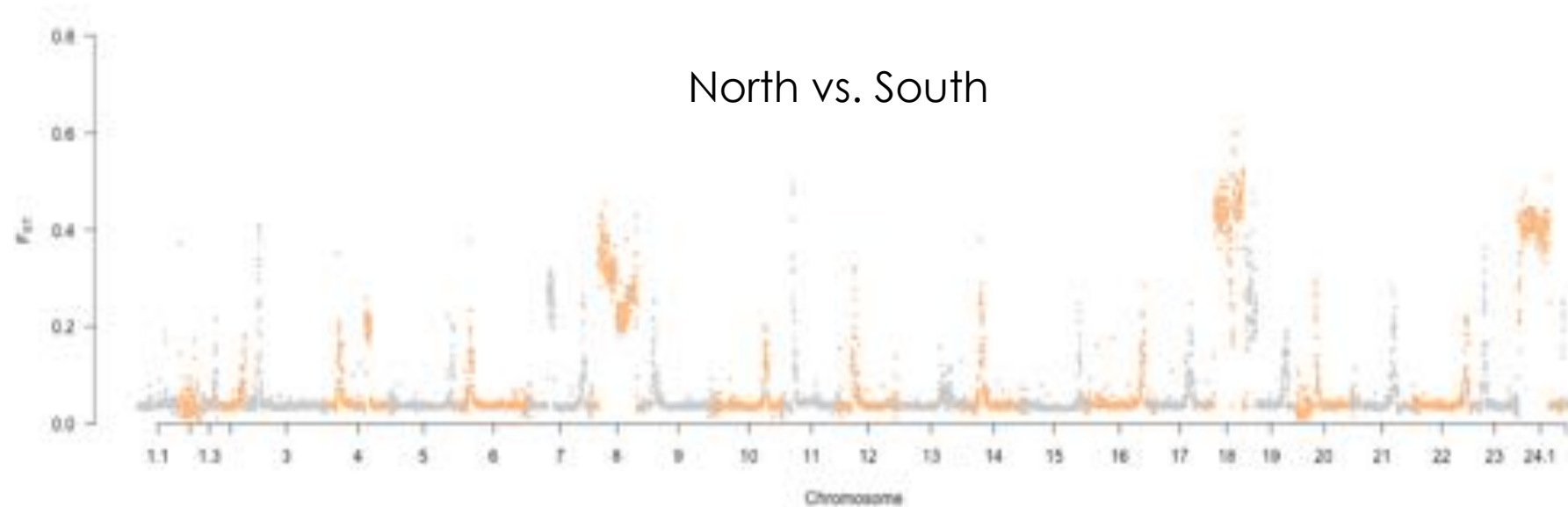


Effect of variation in diversity on divergence → elevated divergence in areas of high diversity, and viceversa

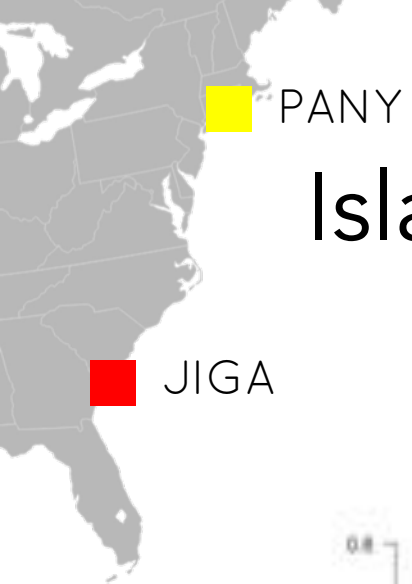


Islands of divergence in the Atlantic silverside

Relative differentiation - F_{ST}

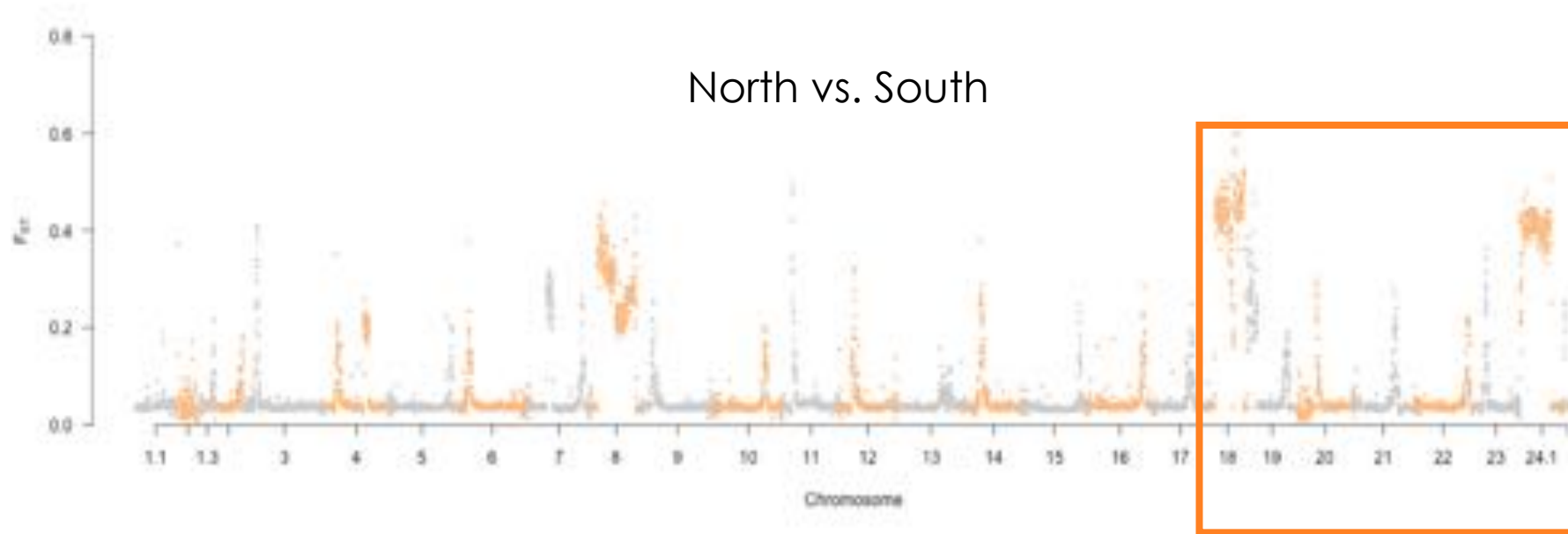


In preparation



Islands of divergence in the Atlantic silverside

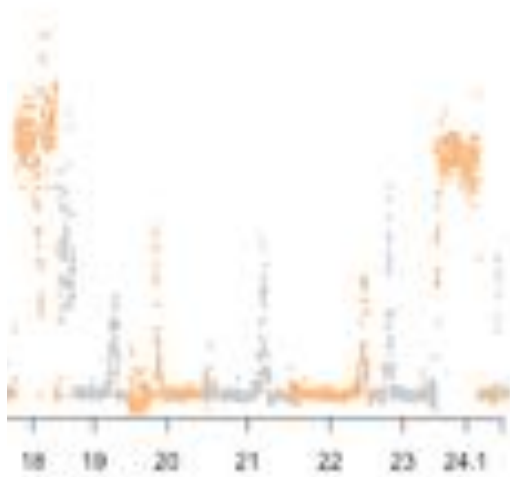
Relative differentiation - F_{ST}



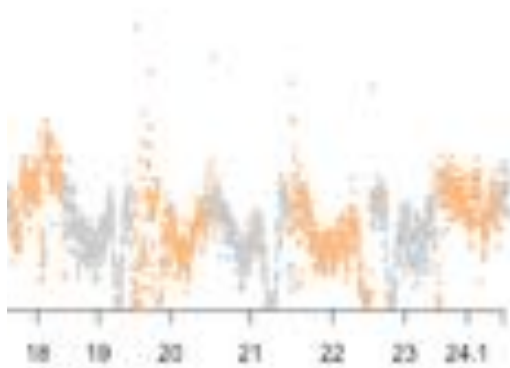
In preparation

Differentiation

F_{ST}

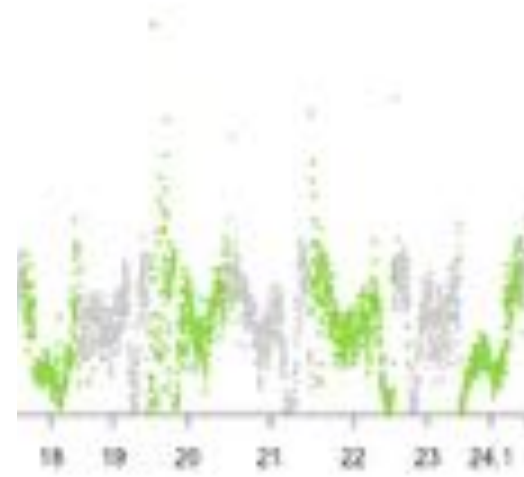


D_{xy}

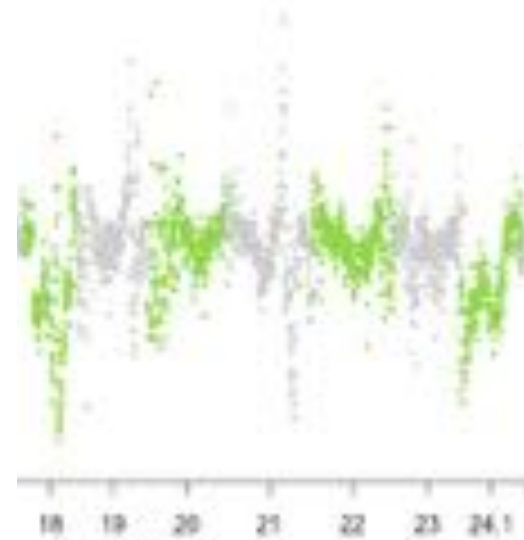


Diversity

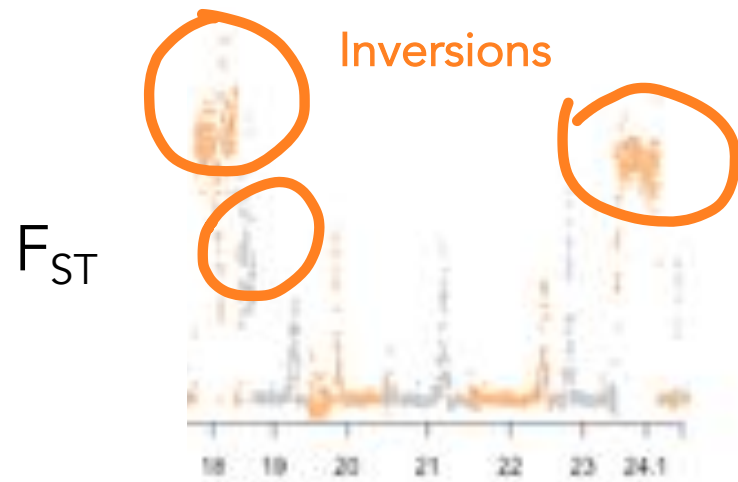
π



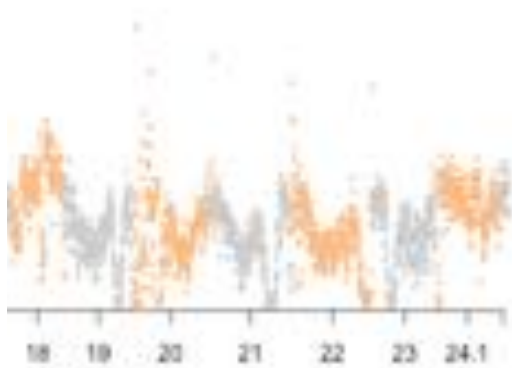
Tajima's D



Differentiation

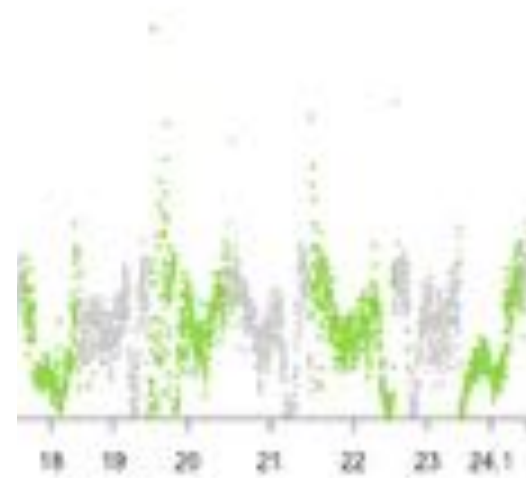


D_{xy}

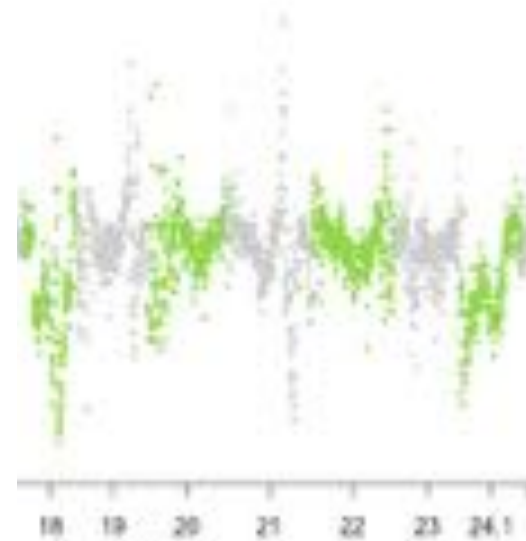


Diversity

π



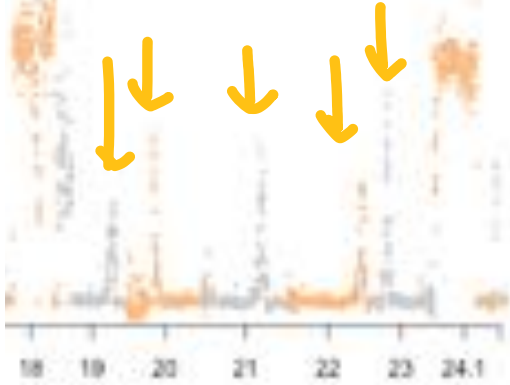
Tajima's D



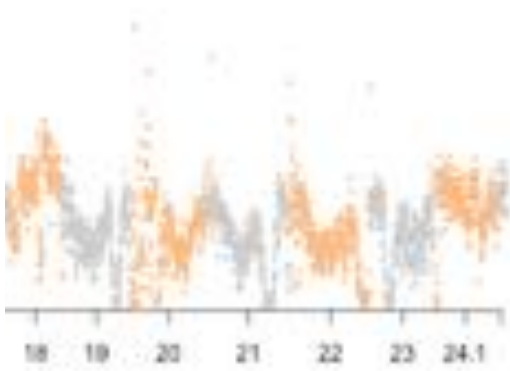
Differentiation

Centromeres?

F_{ST}

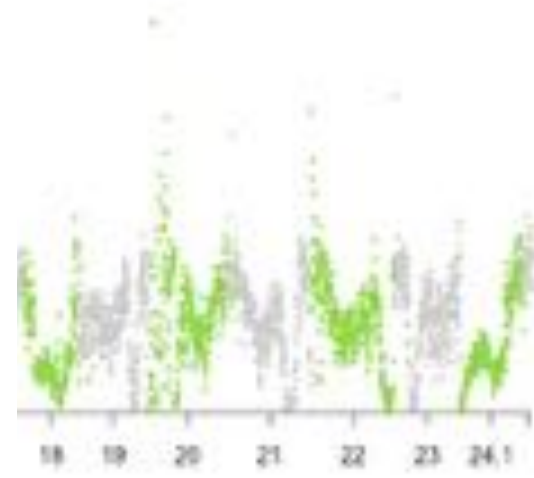


D_{xy}

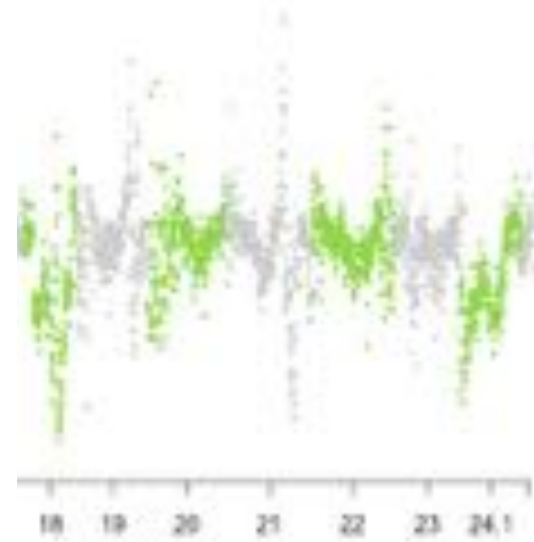


Diversity

π



Tajima's D



In preparation

Structural variation

Though we have evidence from clinal analyses and correspondence between phenotypic differences and frequency of those inversions, 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
- Important variants could get filtered out → low quality assembly and mapping

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
- Important variants could get filtered out → low quality assembly and mapping
- Missing heritability → if only sequence variation is screened but the variant associated with phenotype/environment in a SV

The end!