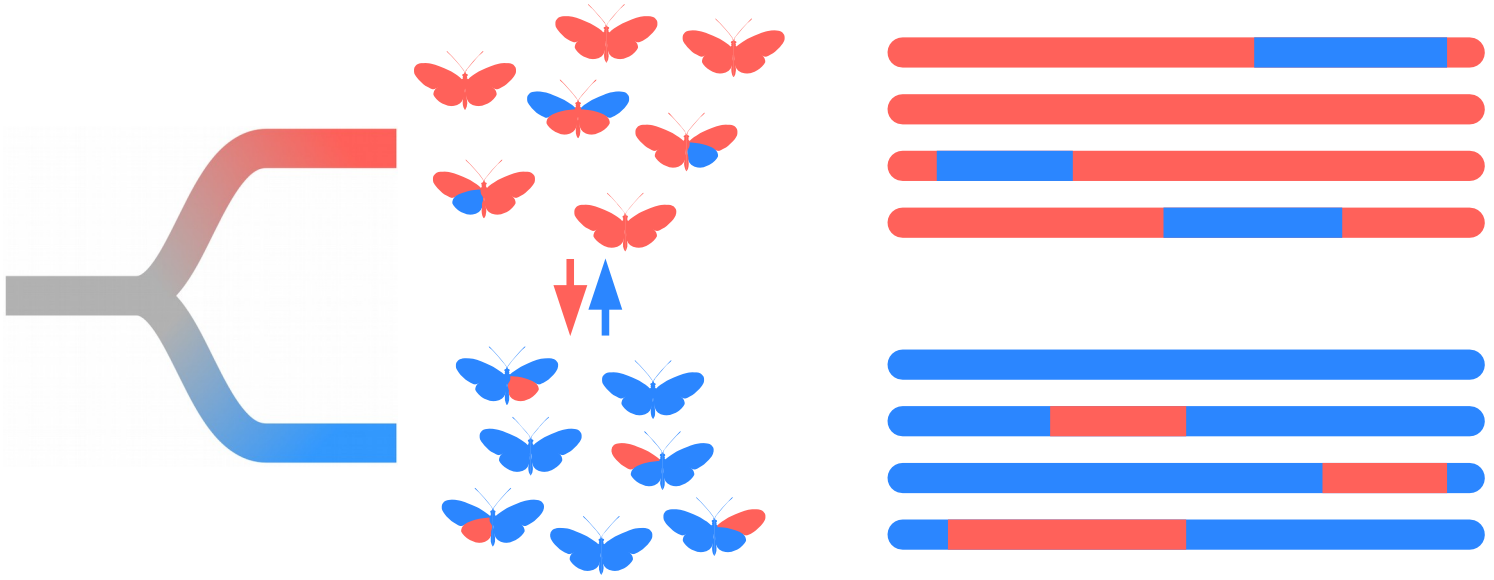
Speciation and hybridisation

Part 2: All that is gold does not glitter

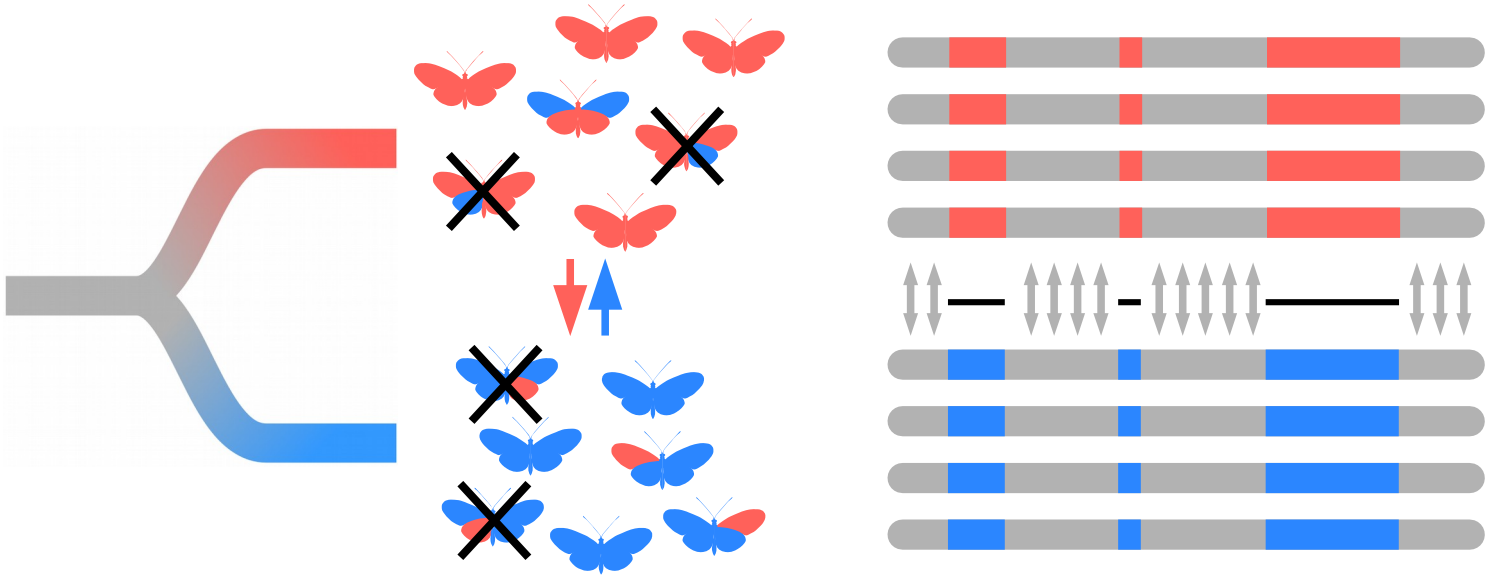
Simon Martin

Institute of Evolutionary Biology
University of Edinburgh

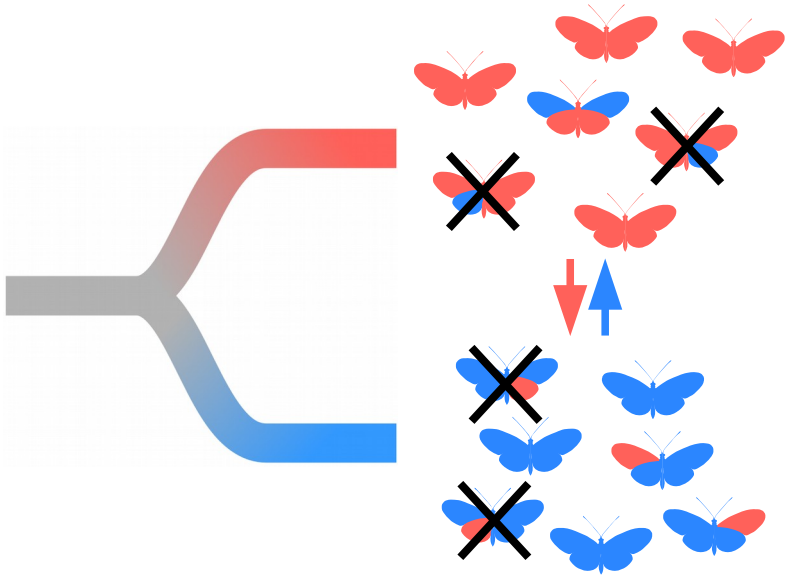
January 2020

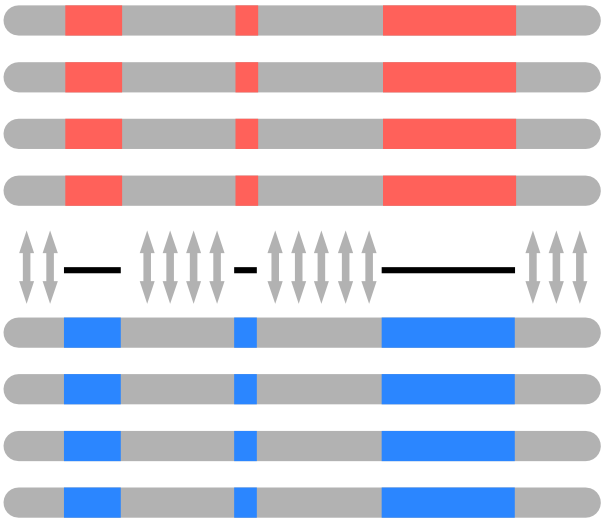


Barton & Bengtson 1986 Heredity, Mallet 1995 TREE, Wu 2001 JEB etc.

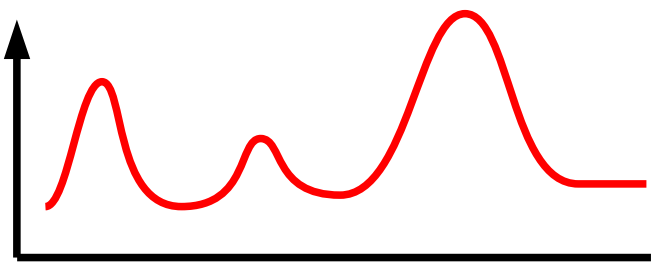


Barton & Bengtson 1986 Heredity, Mallet 1995 TREE, Wu 2001 JEB

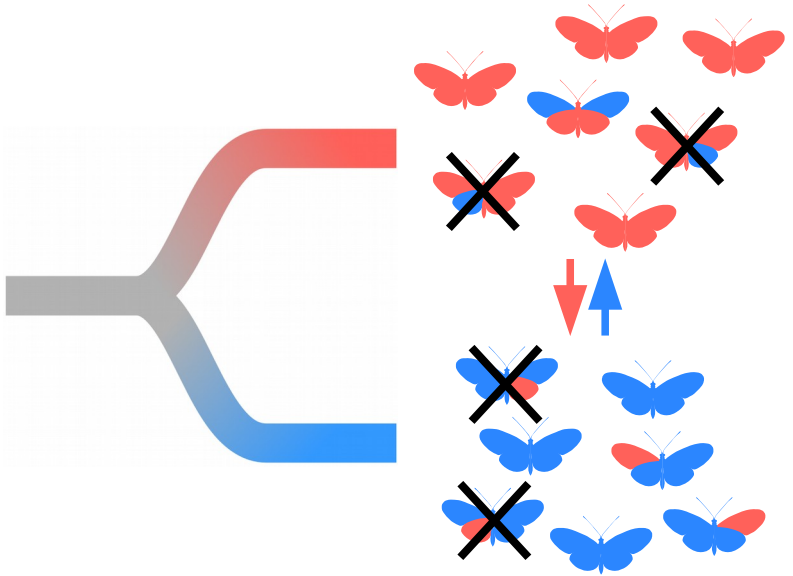


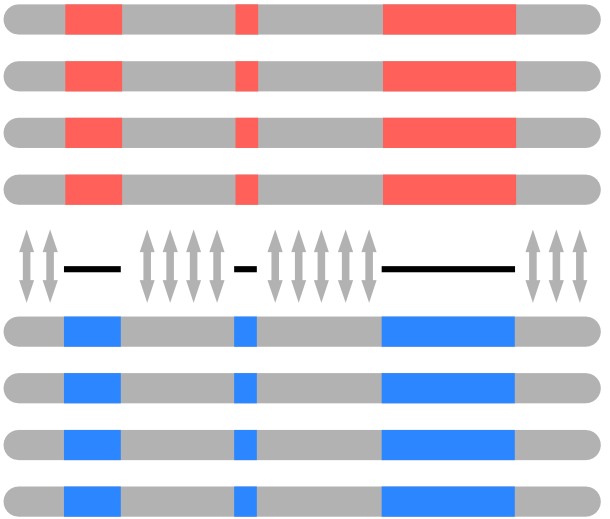


Strength of species barrier

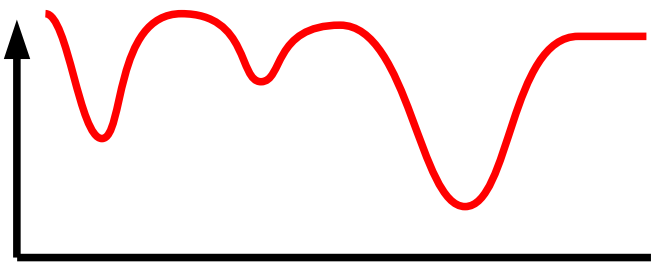


Barton & Bengtson 1986 Heredity, Mallet 1995 TREE, Wu 2001 JEB

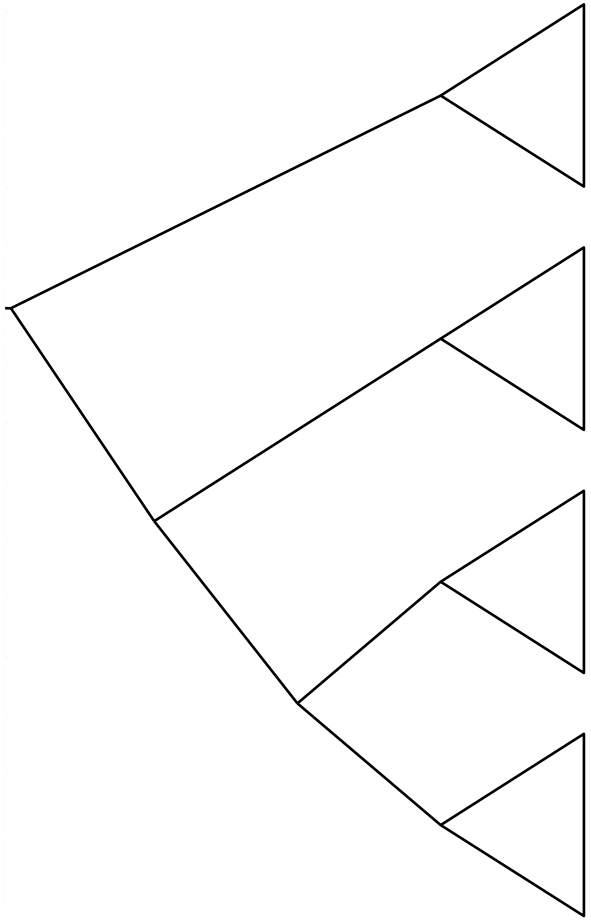




effective migration rate (m_e)



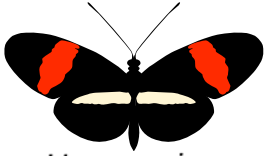
Aeschbacher et al. 2017



Outgroups

H. cydno

Panama

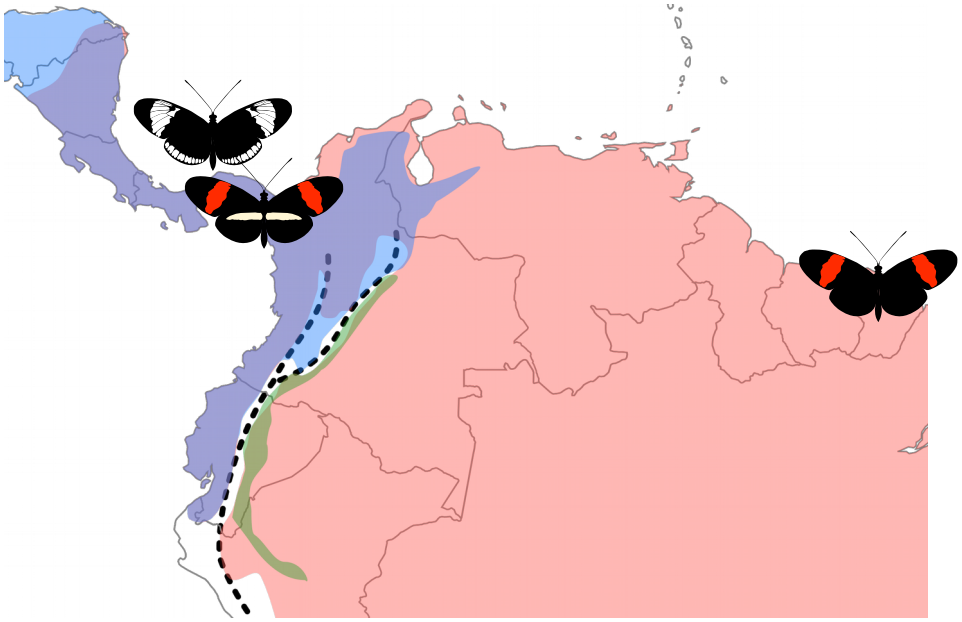


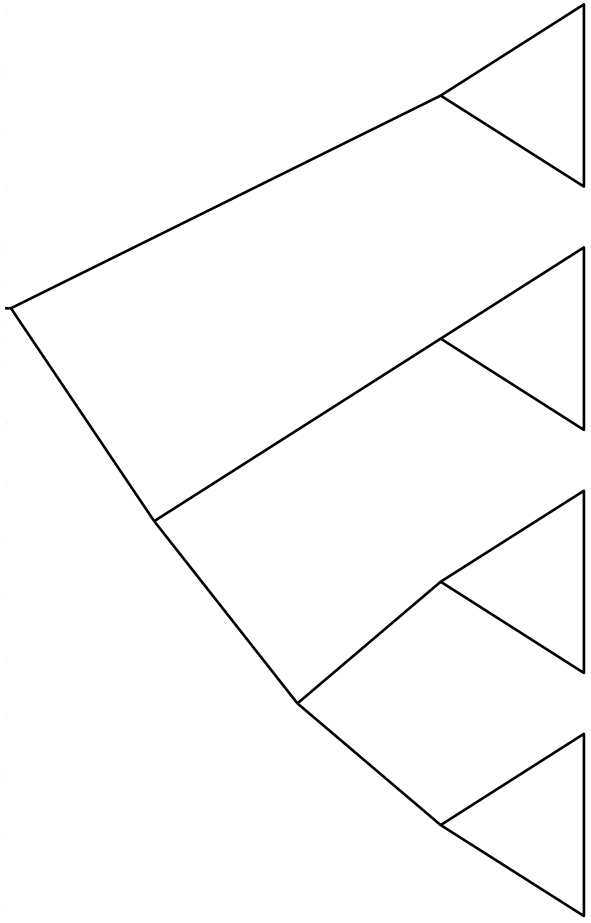
H. m. rosina

Panama

H. m. melpomene

French Guiana

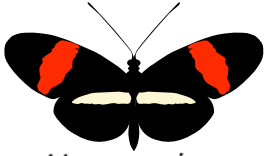




Outgroups

H. cydno

Panama

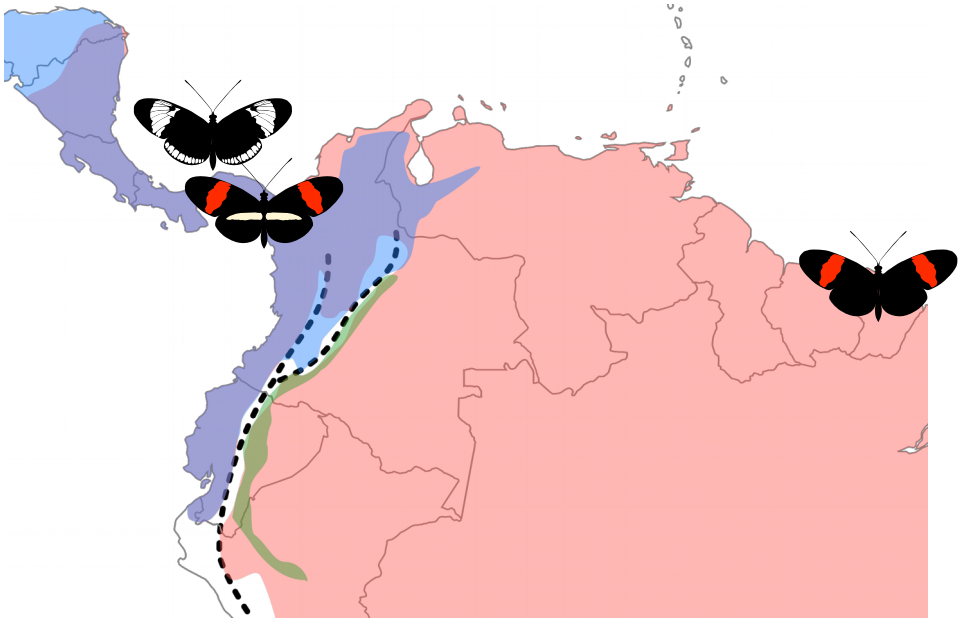


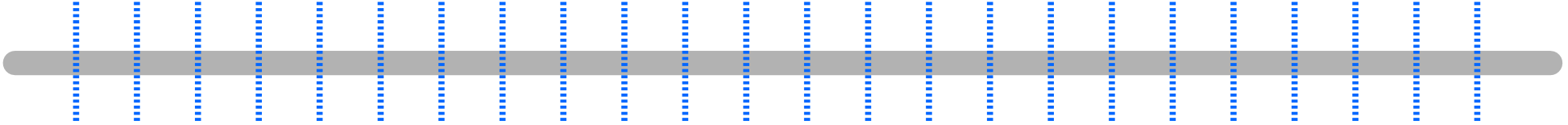
H. m. rosina

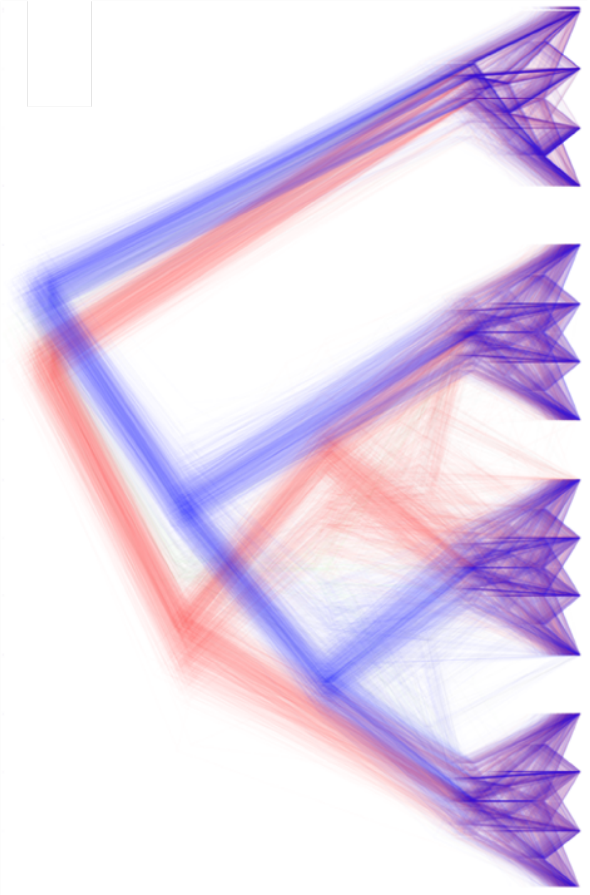
Panama

H. m. melpomene

French Guiana



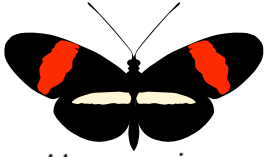




Outgroups

H. cydno

Panama

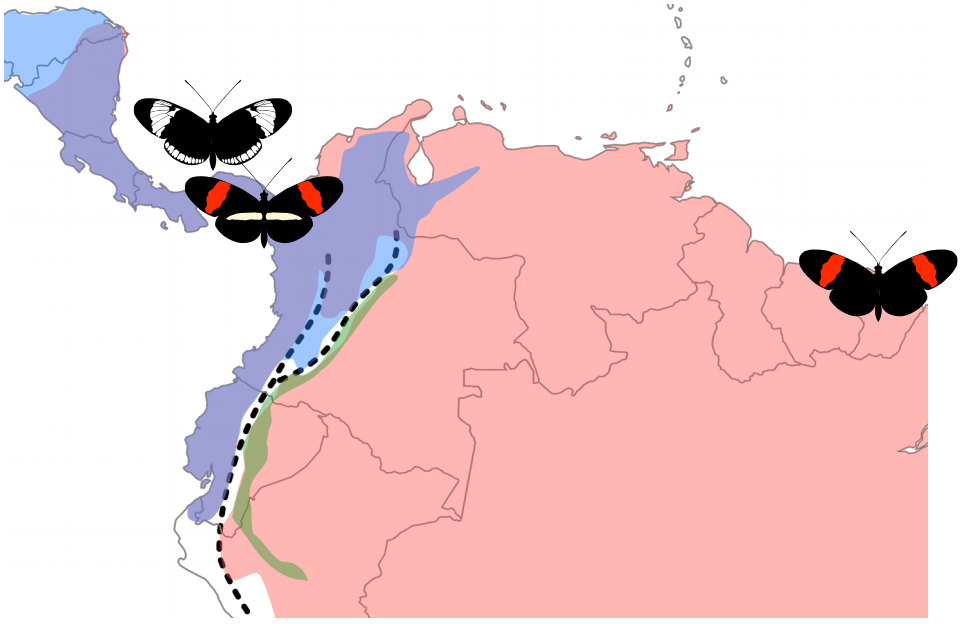


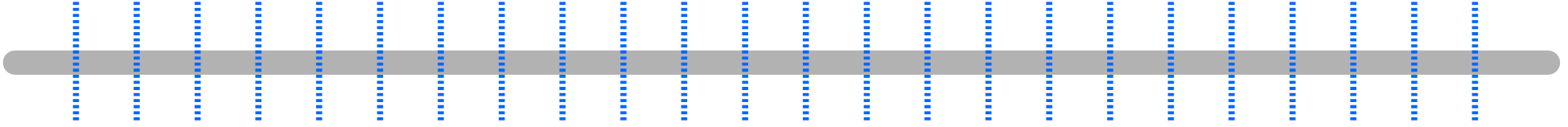
H. m. rosina

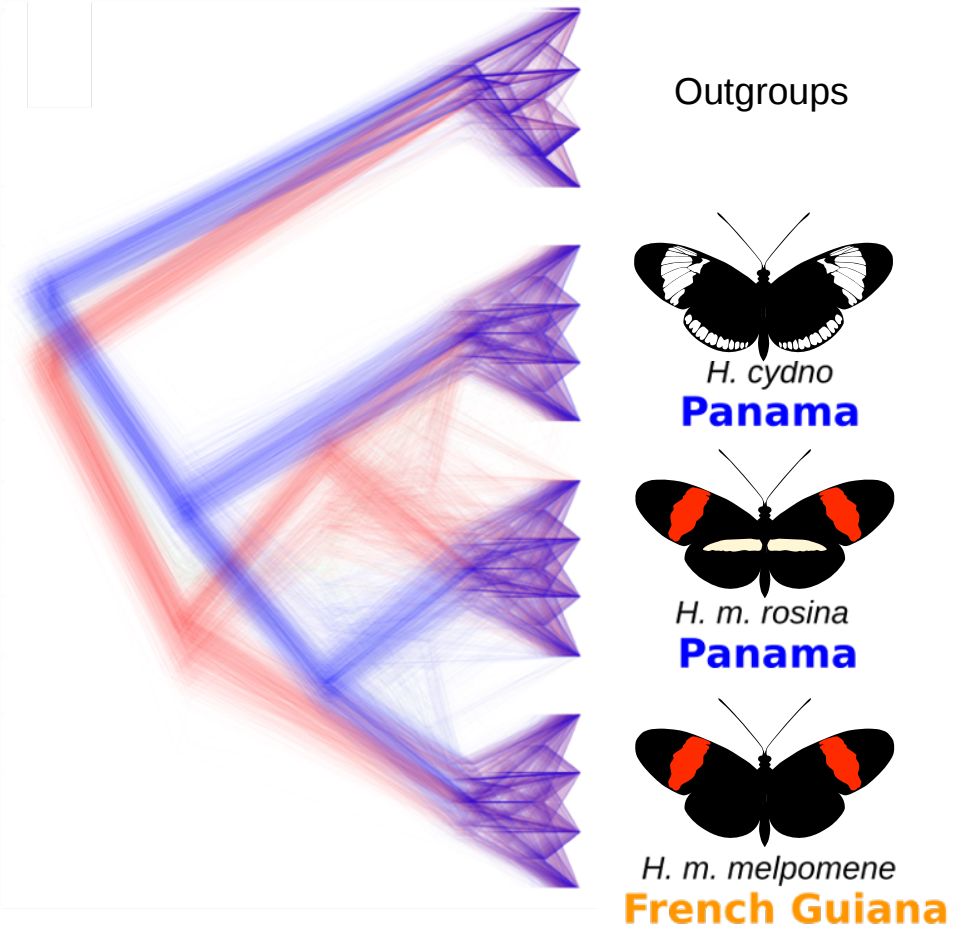
Panama

H. m. melpomene

French Guiana







Outgroups

H. cydno

Panama

H. m. rosina

Panama

H. m. melpomene

French Guiana

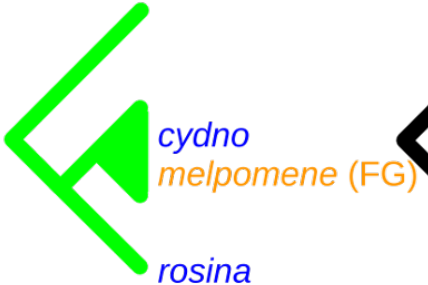
"species" topology

1510 (53.0%)

"geography" topology

1201 (42.2%)

"control" topology

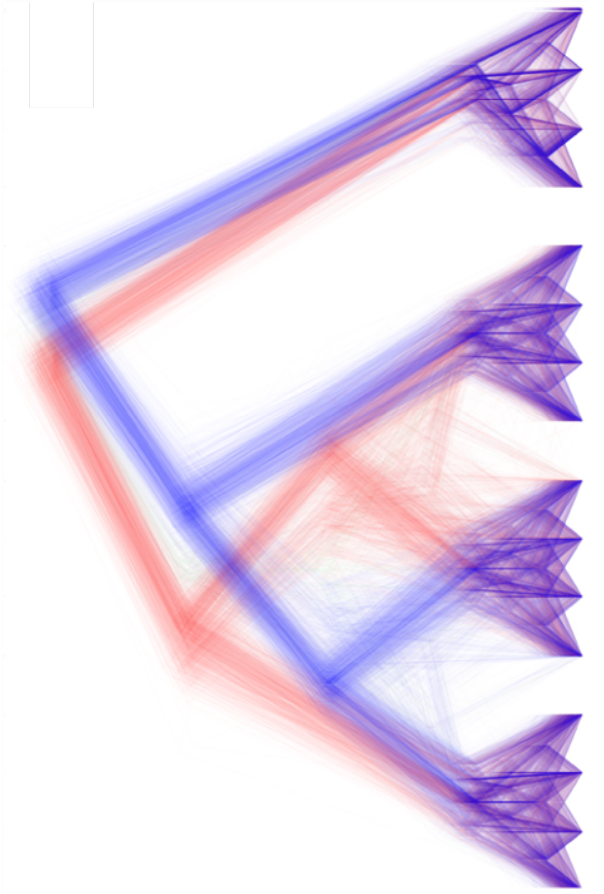


32 (1.1%)

"unresolved" topology

105 (3.7%)

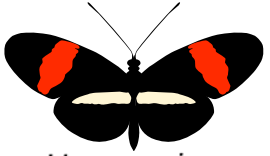
Martin et al. 2013 Genome Research



Outgroups

H. cydno

Panama



H. m. rosina

Panama

H. m. melpomene

French Guiana

"species" topology

1510 (53.0%)

"geography" topology

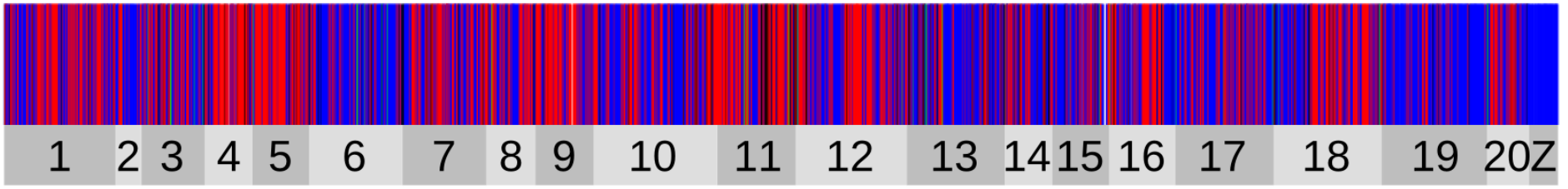
1201 (42.2%)

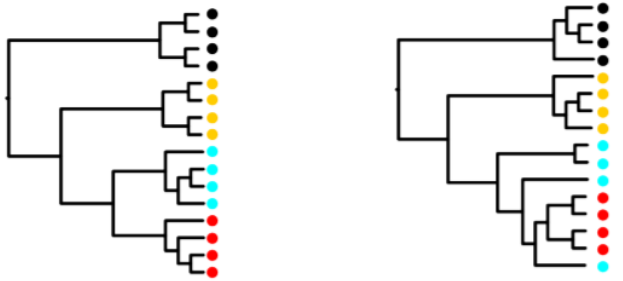
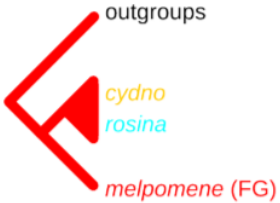
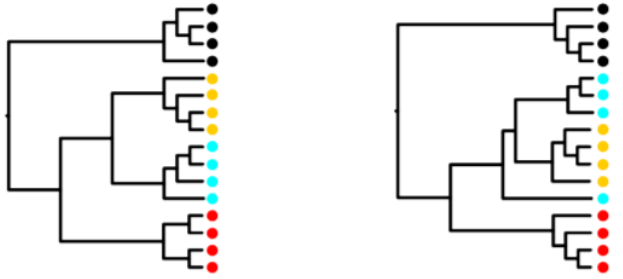
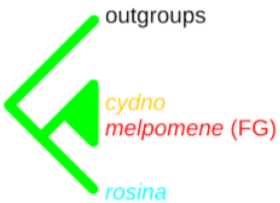
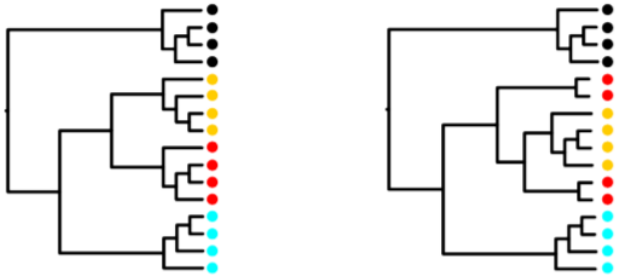
"control" topology

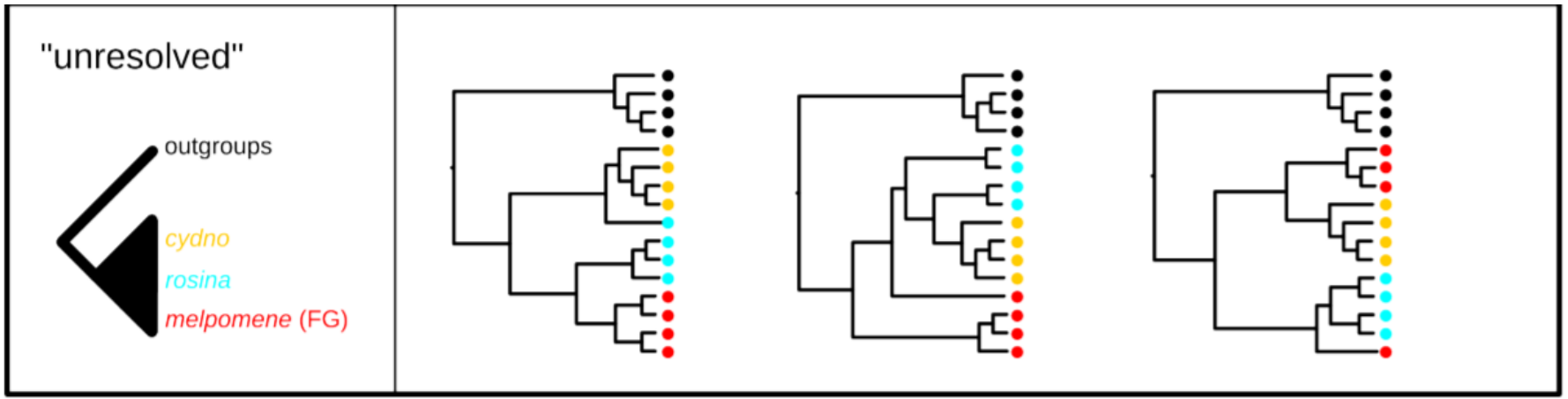
32 (1.1%)

"unresolved" topology

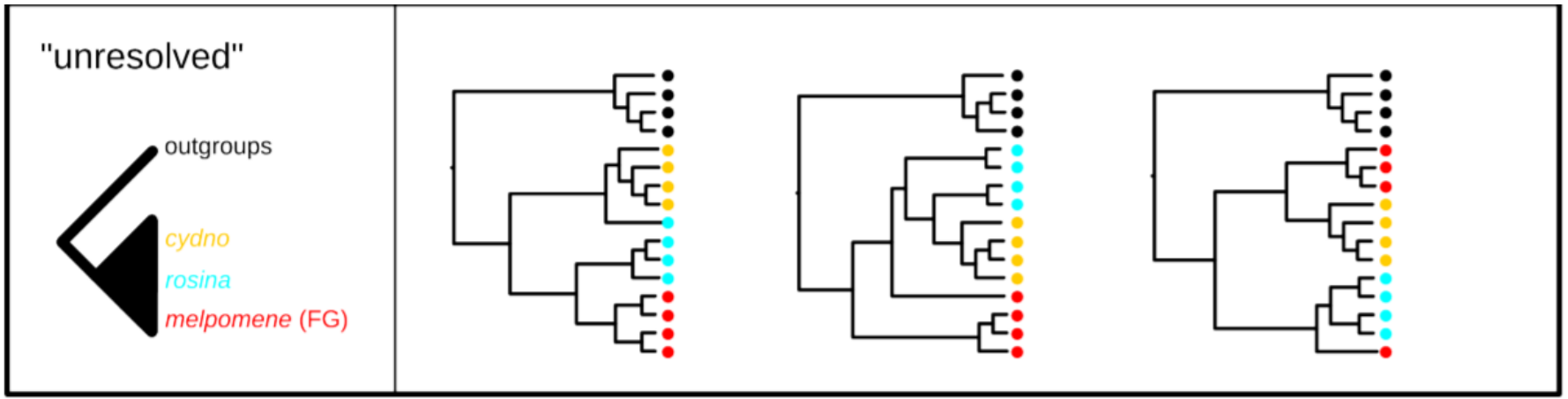
105 (3.7%)



"species tree" 	
"geography tree" 	
"control tree" 	

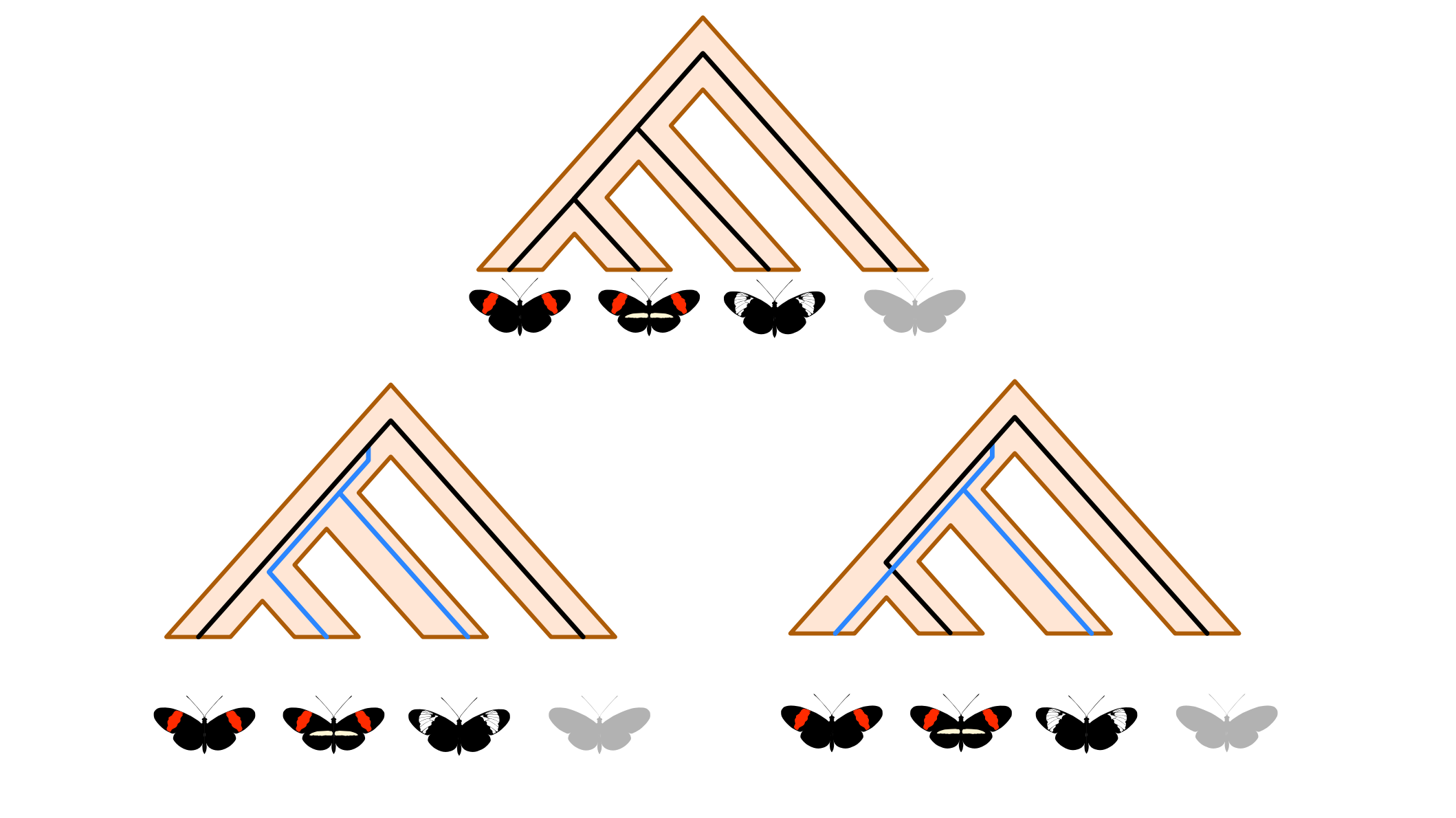

"unresolved"
outgroups
cydno
rosina
melpomene (FG)

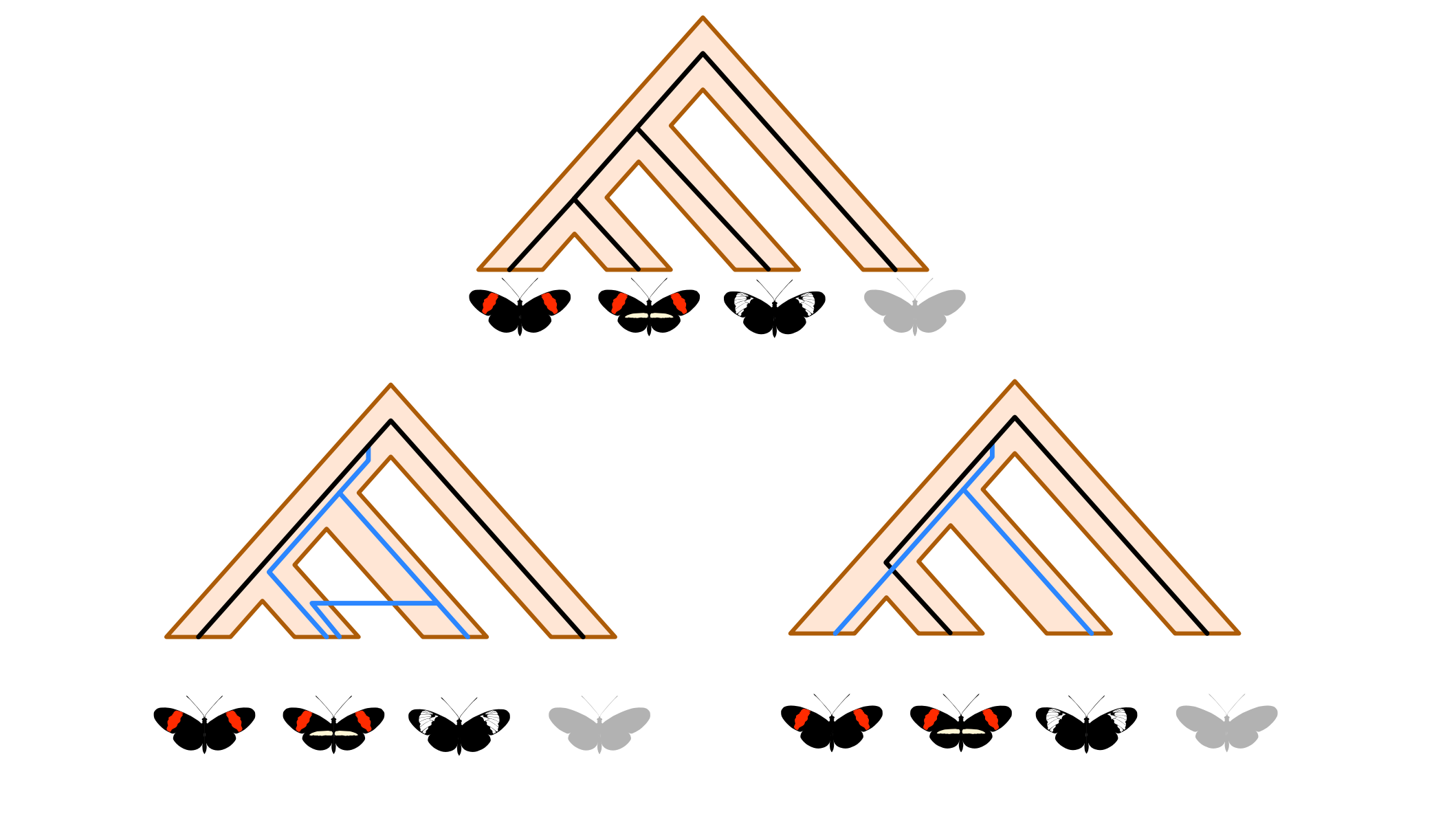
Window Size (kb)	Species topology	Geography topology	Control topology	Unsorted
10	38%	31%	2%	30%
20	45%	37%	2%	17%
50	51%	41%	1%	6%
100	53%	42%	1%	4%
200	56%	41%	1%	2%

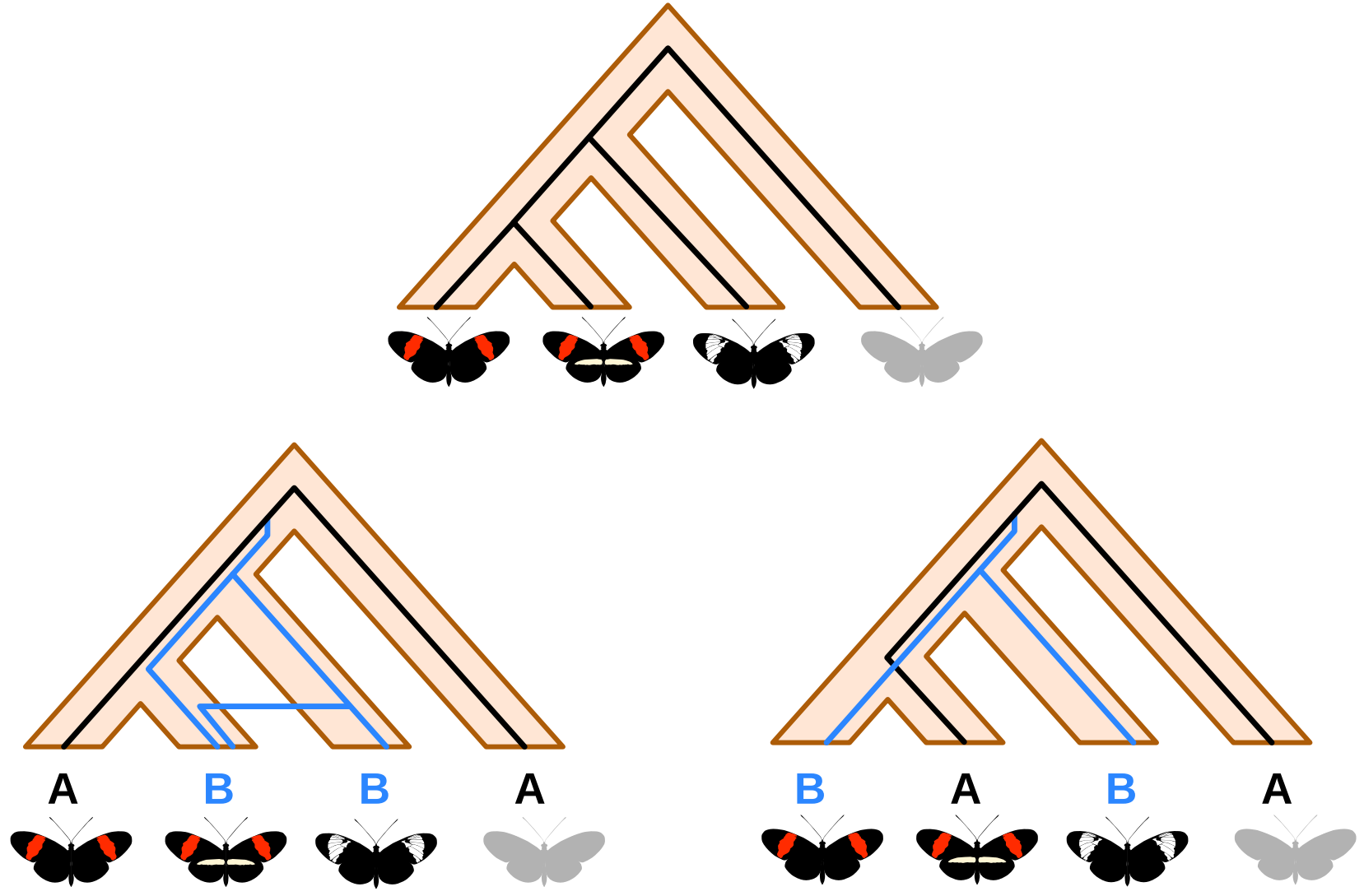

"unresolved"
outgroups
cydno
rosina
melpomene (FG)

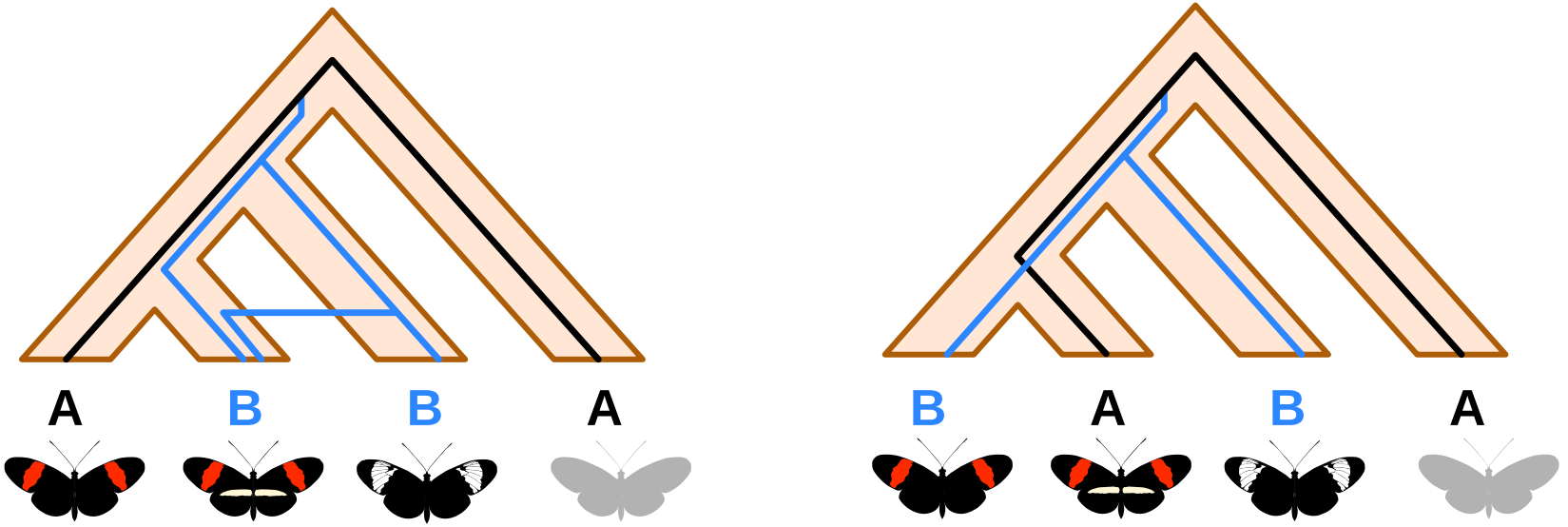
1. We need to work at much finer resolution

2. We need to ***quantify*** genealogy shapes

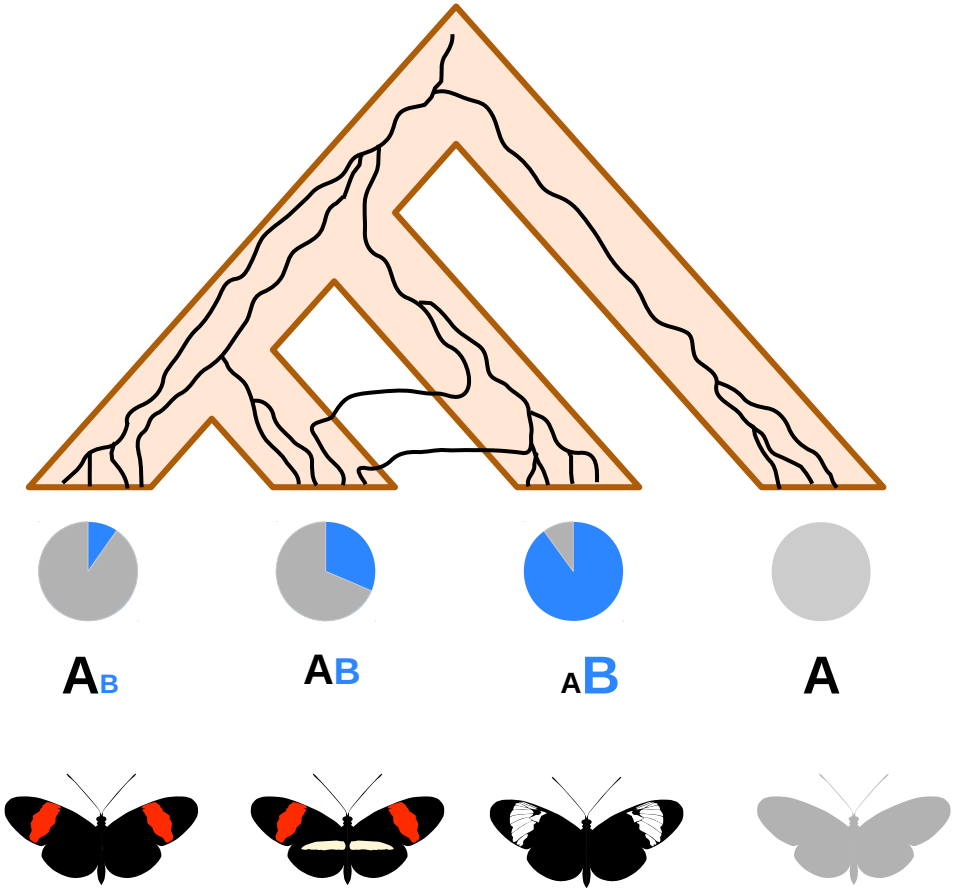





A
B
B
A
B
A
B
A



$$D(P_1, P_2, P_3, O) = \frac{\sum C_{ABBA}(i) - C_{BABA}(i)}{\sum C_{ABBA}(i) + C_{BABA}(i)}$$



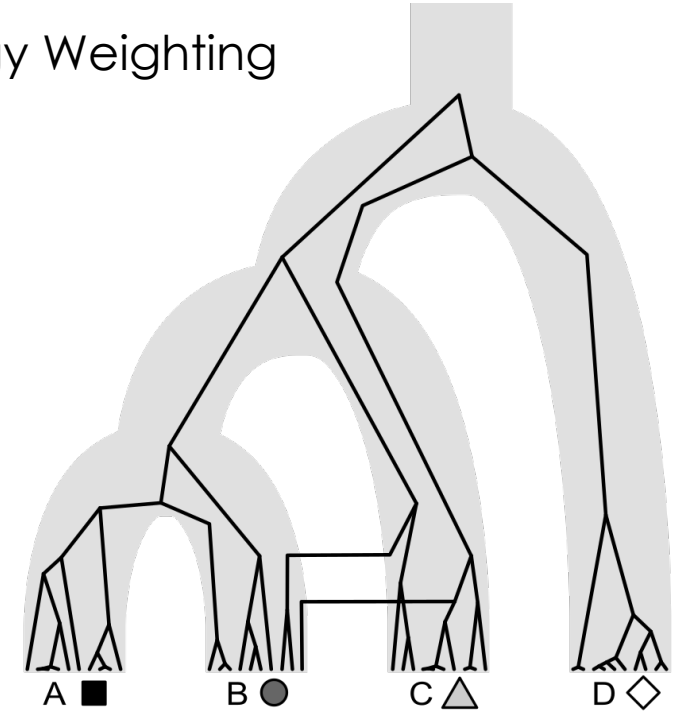
$$D(P_1, P_2, P_3, O) = \frac{\sum C_{ABBA}(i) - C_{BABA}(i)}{\sum C_{ABBA}(i) + C_{BABA}(i)}$$

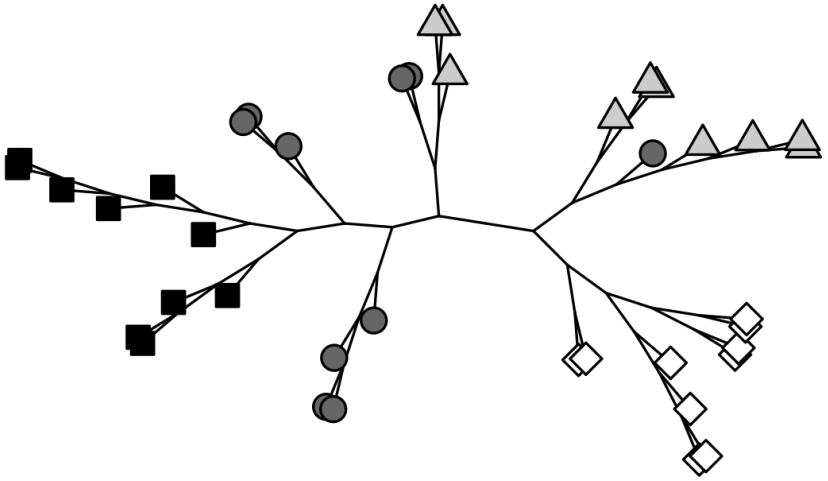
$$C_{ABBA}(i) = (1 - \hat{p}_{i1})\hat{p}_{i2}\hat{p}_{i3}(1 - \hat{p}_{i4})$$

$$C_{BABA}(i) = \hat{p}_{i1}(1 - \hat{p}_{i2})\hat{p}_{i3}(1 - \hat{p}_{i4})$$

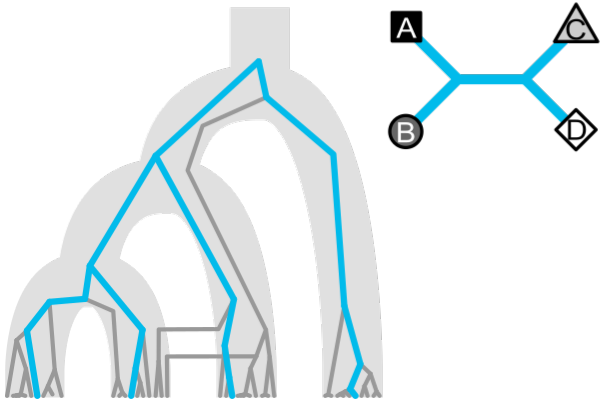
Can we quantify the shape of **trees** in the same way?

Topology Weighting

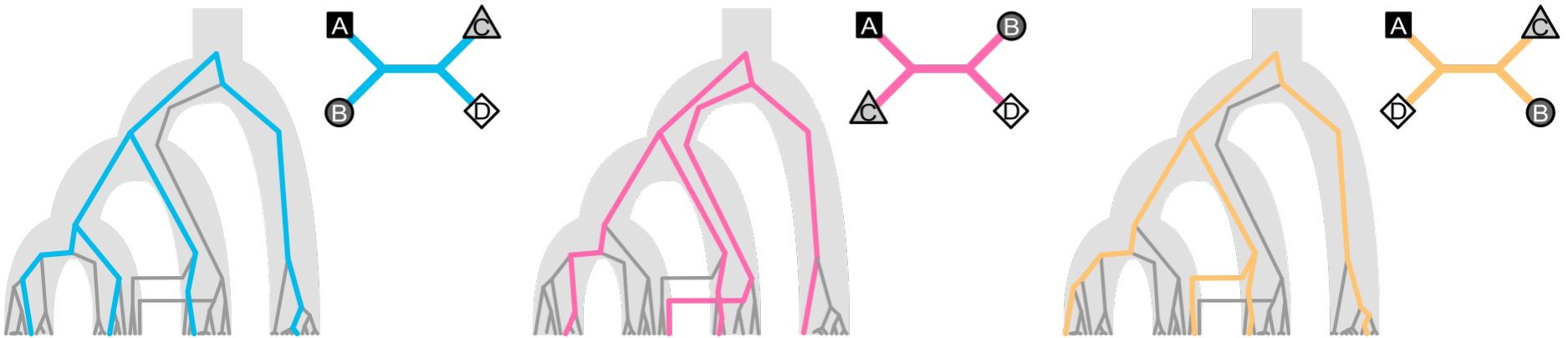




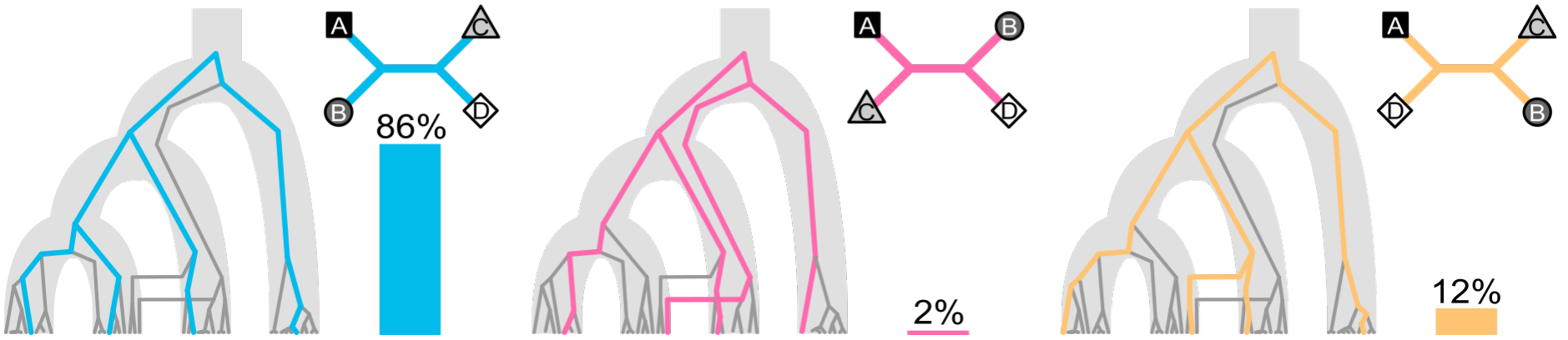
Topology Weighting

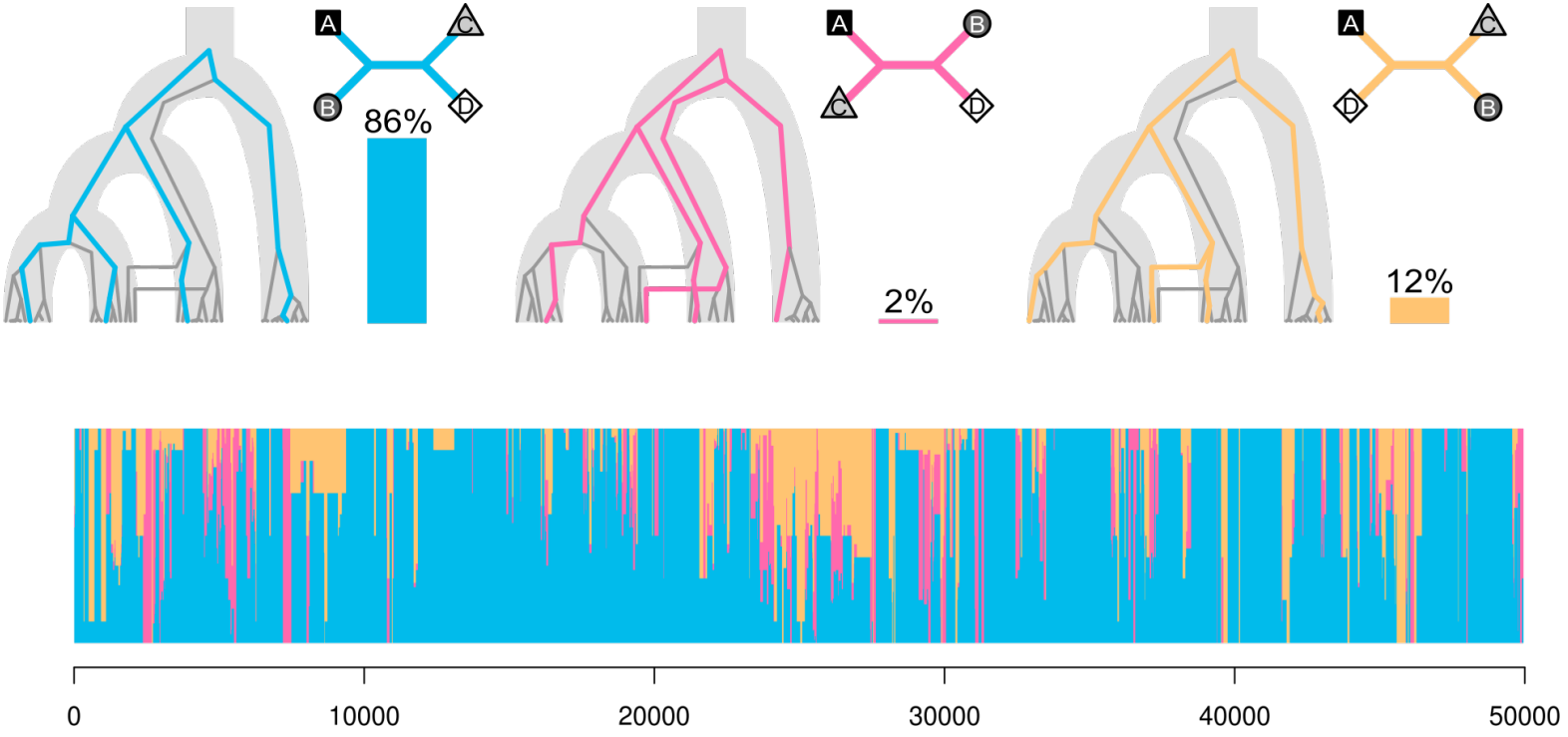


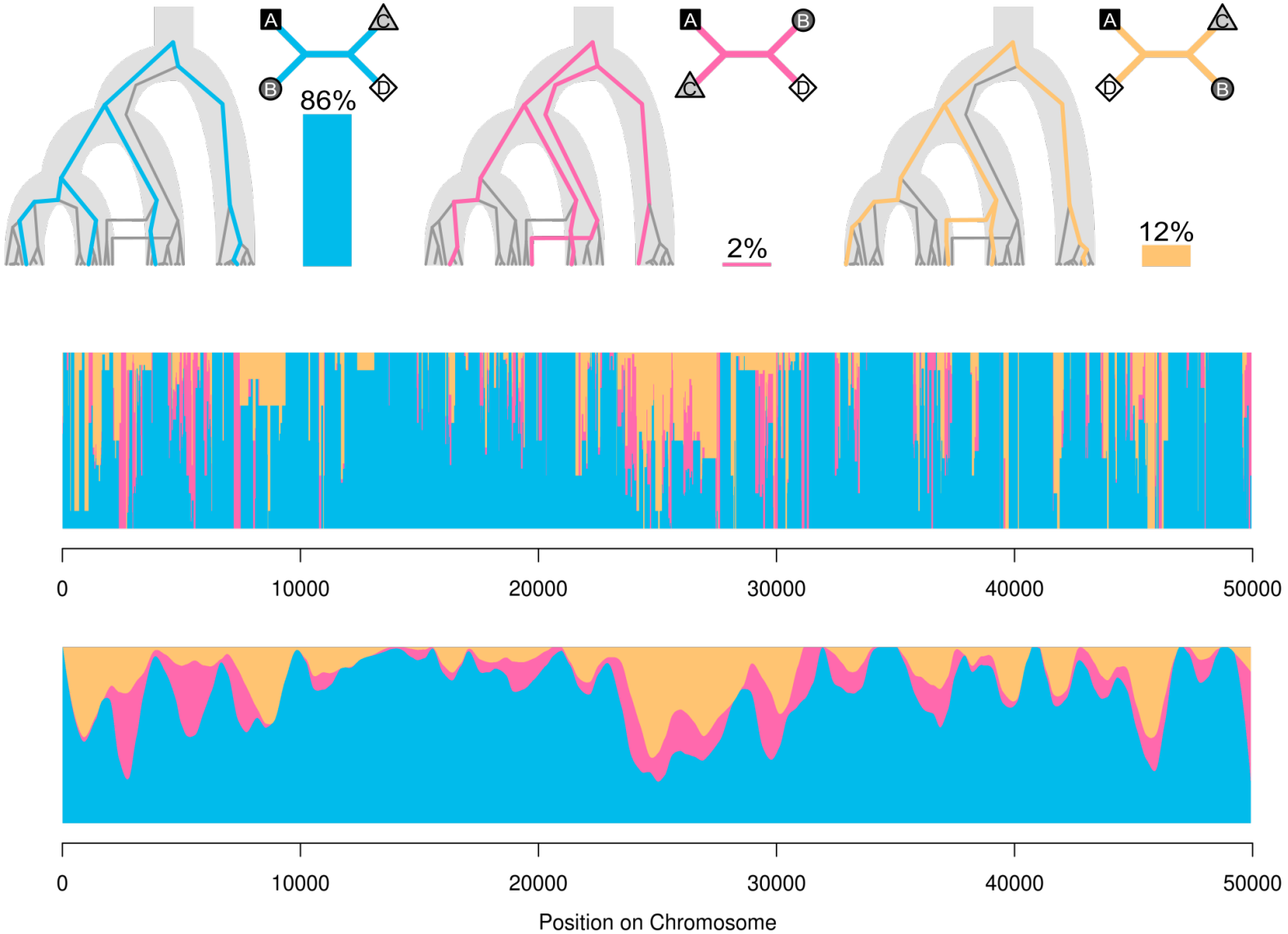
Topology Weighting

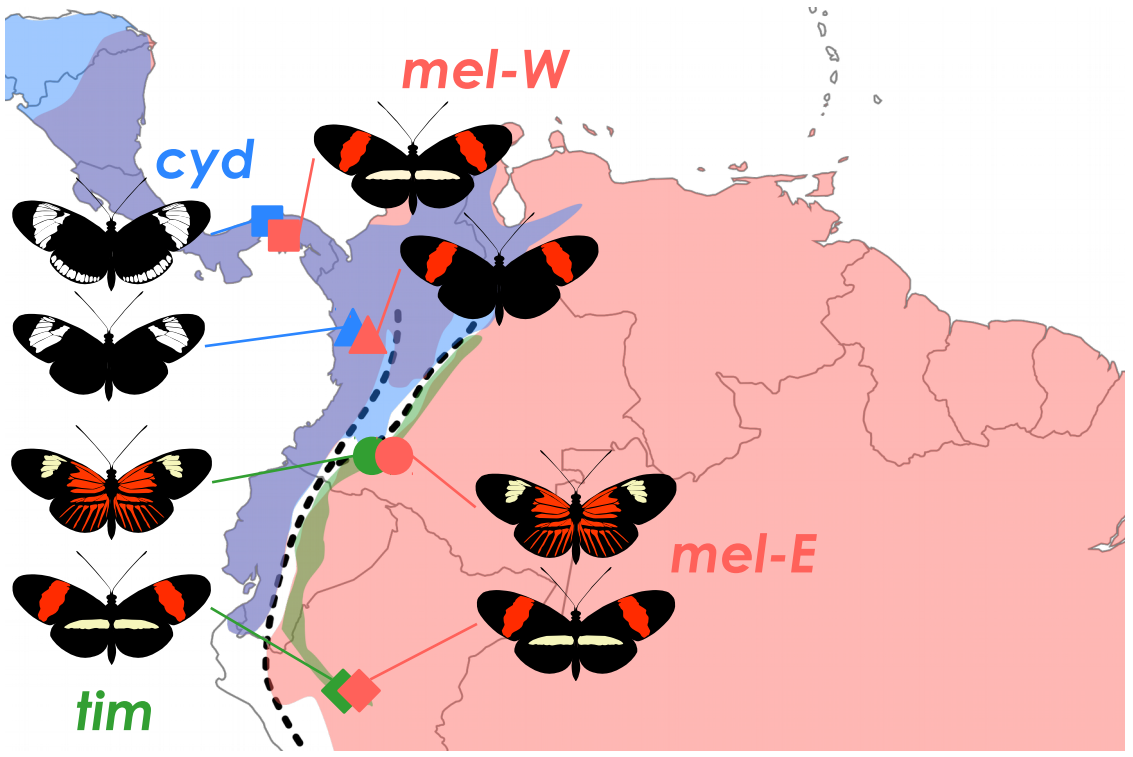


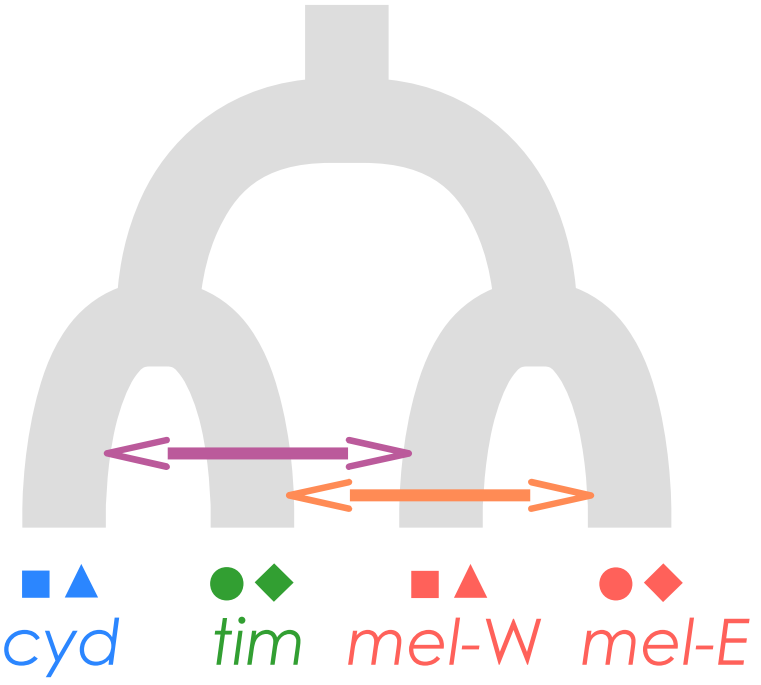
Topology Weighting

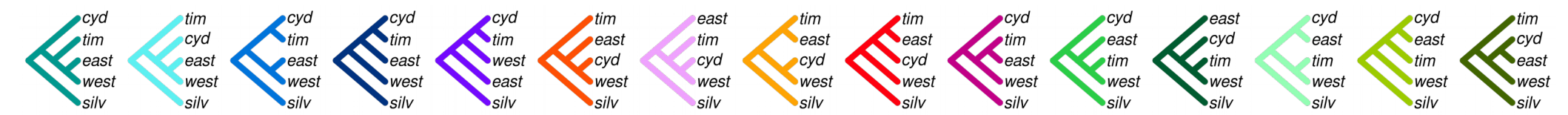



A
C
B
D
86%
A
B
C
D
2%
A
C
D
B
12%
0
10000
20000
30000
40000
50000


A
C
B
D
86%
A
B
C
D
2%
A
C
D
B
12%
0
10000
20000
30000
40000
50000
0
10000
20000
30000
40000
50000
Position on Chromosome


mel-W
cyd
mel-E
tim


cyd
tim
mel-W
mel-E



Unfused Autosomes

2

3

4

5

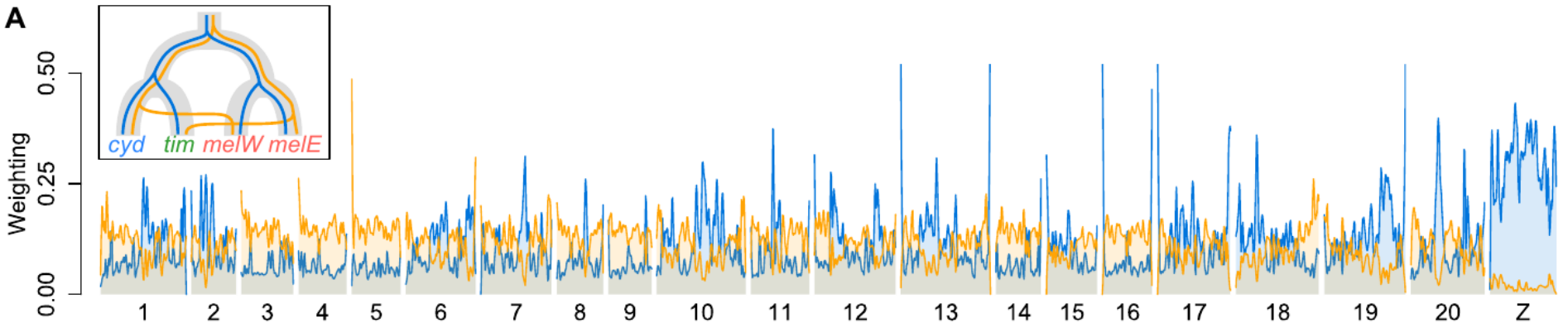
Fused Autosomes

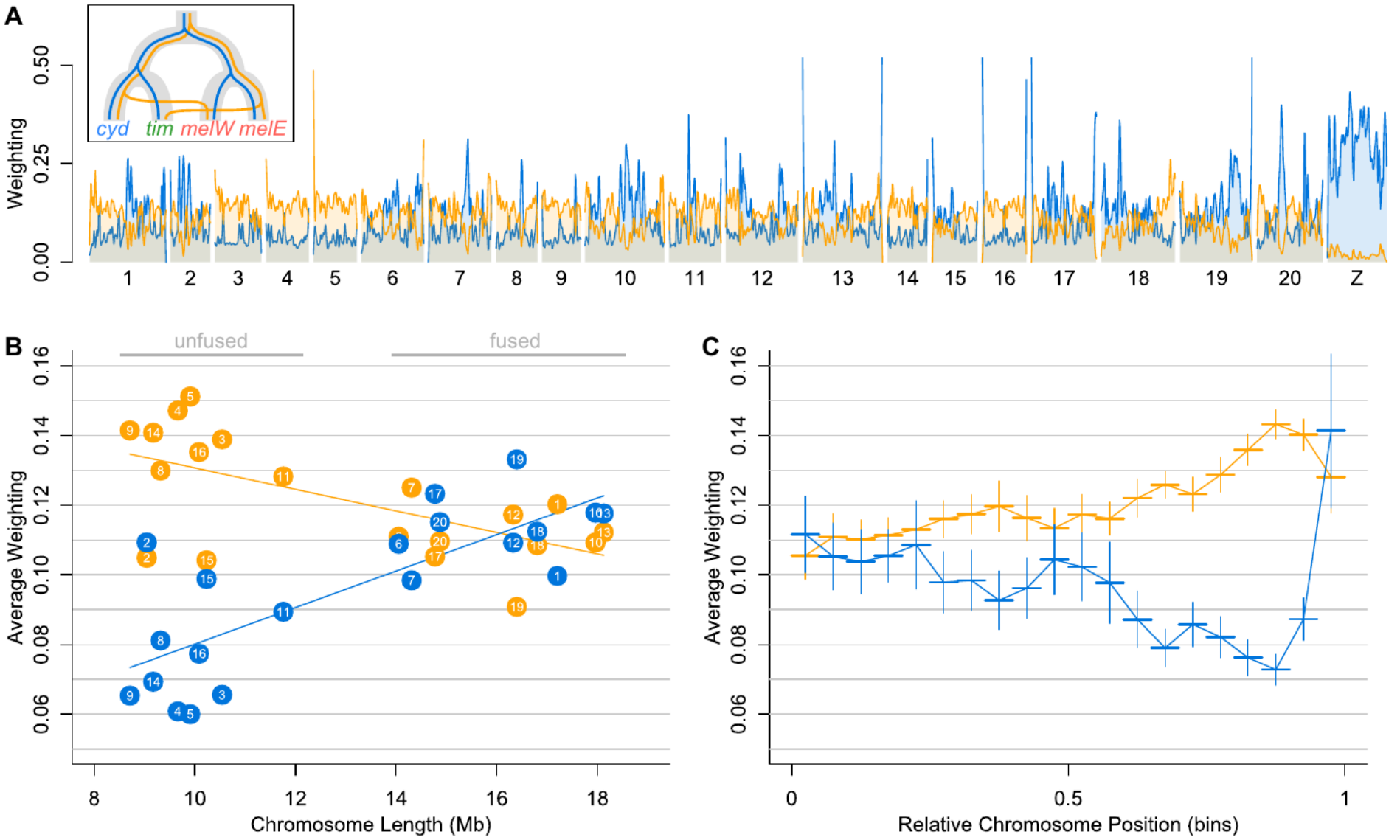
1

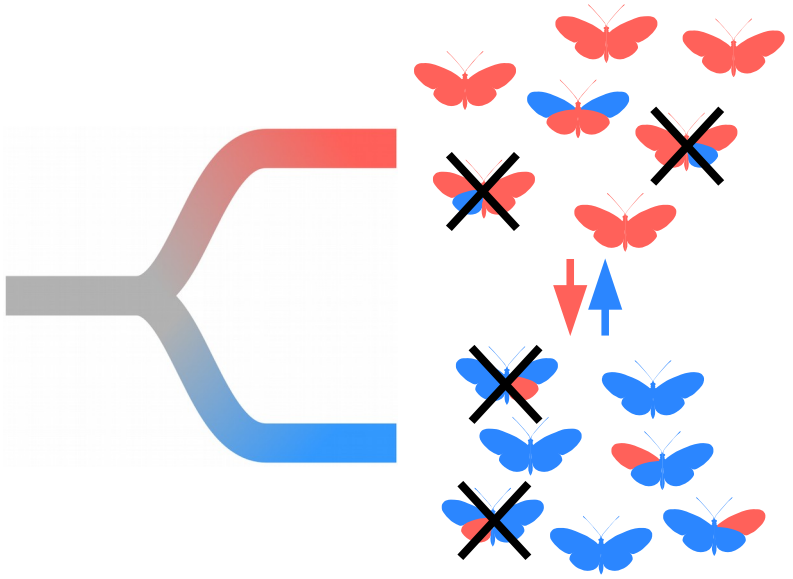
6

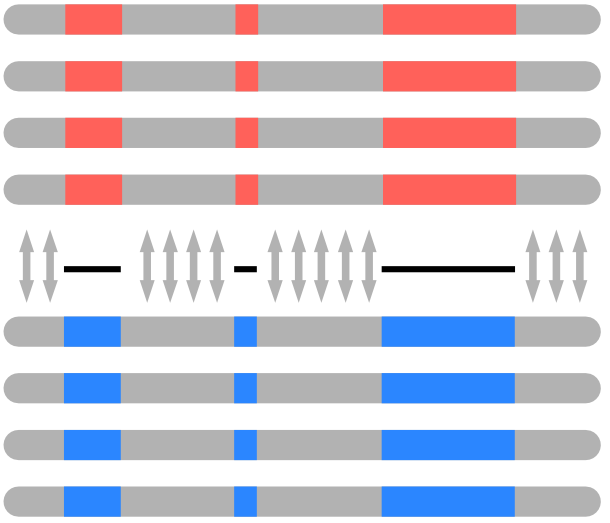
Z chromosome

21

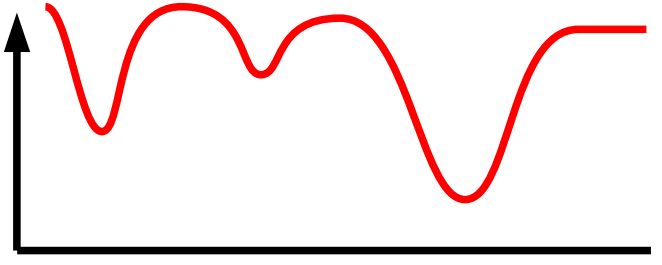

A
cyd
tim
melW
melE
Weighting
0.00
0.25
0.50
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Z


A
cyd
tim
melW
melE
Weighting
0.00
0.25
0.50
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Z
B
unfused
fused
Average Weighting
0.06
0.08
0.10
0.12
0.14
0.16
8
10
12
14
16
18
Chromosome Length (Mb)
C
Average Weighting
0.06
0.08
0.10
0.12
0.14
0.16
0
0.5
1
Relative Chromosome Position (bins)





effective migration rate (m_e)



We want a measure that is:

Proportional to the effective migration rate

Unaffected by confounding factors like N_e and mutation rate

We want a measure that is:

Proportional to the effective migration rate

Unaffected by confounding factors like N_e and mutation rate

$$D(P_1, P_2, P_3, O) = \frac{\sum C_{ABBA}(i) - C_{BABA}(i)}{\sum C_{ABBA}(i) + C_{BABA}(i)}$$

We want a measure that is:

Proportional to the effective migration rate

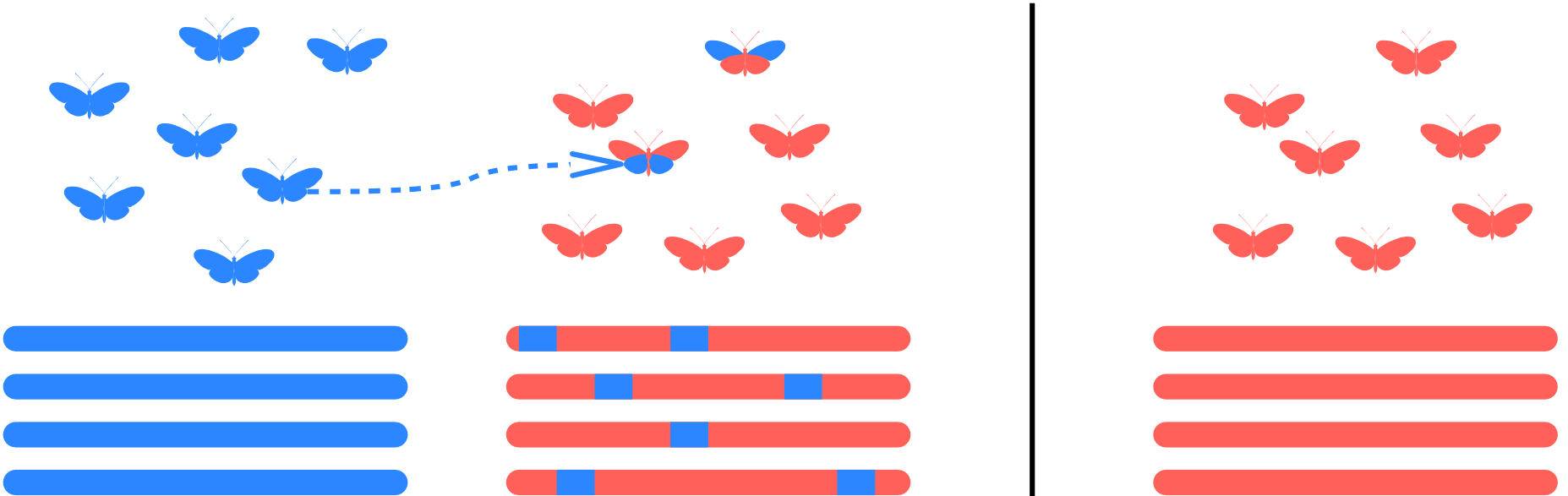
Unaffected by confounding factors like N_e and mutation rate

$$f_d \text{ ~~}B\text{~~}(P_1, P_2, P_3, O) = \frac{\sum C_{ABBA}(i) - C_{BABA}(i)}{\sum C_{ABBA}(i) + C_{BABA}(i)}$$

replace denominator with maximum possible value of numerator

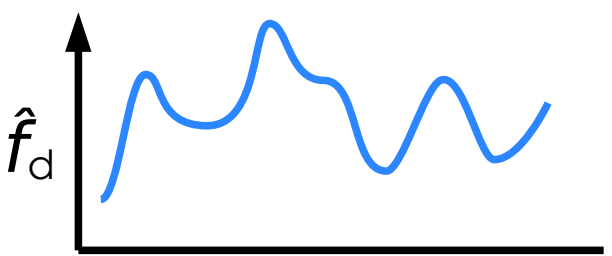
Martin et al. 2015 Mol. Biol. Evol.

$\hat{f}_d$: estimated admixture proportion

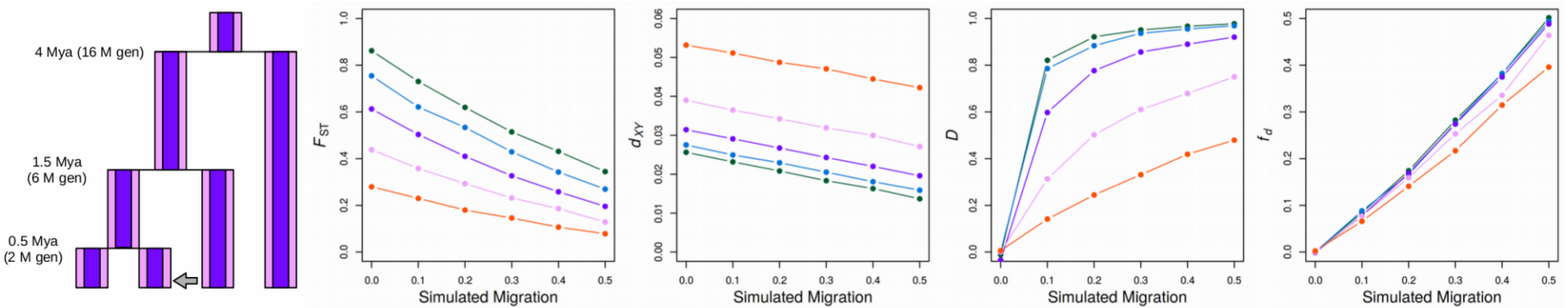


sympatric

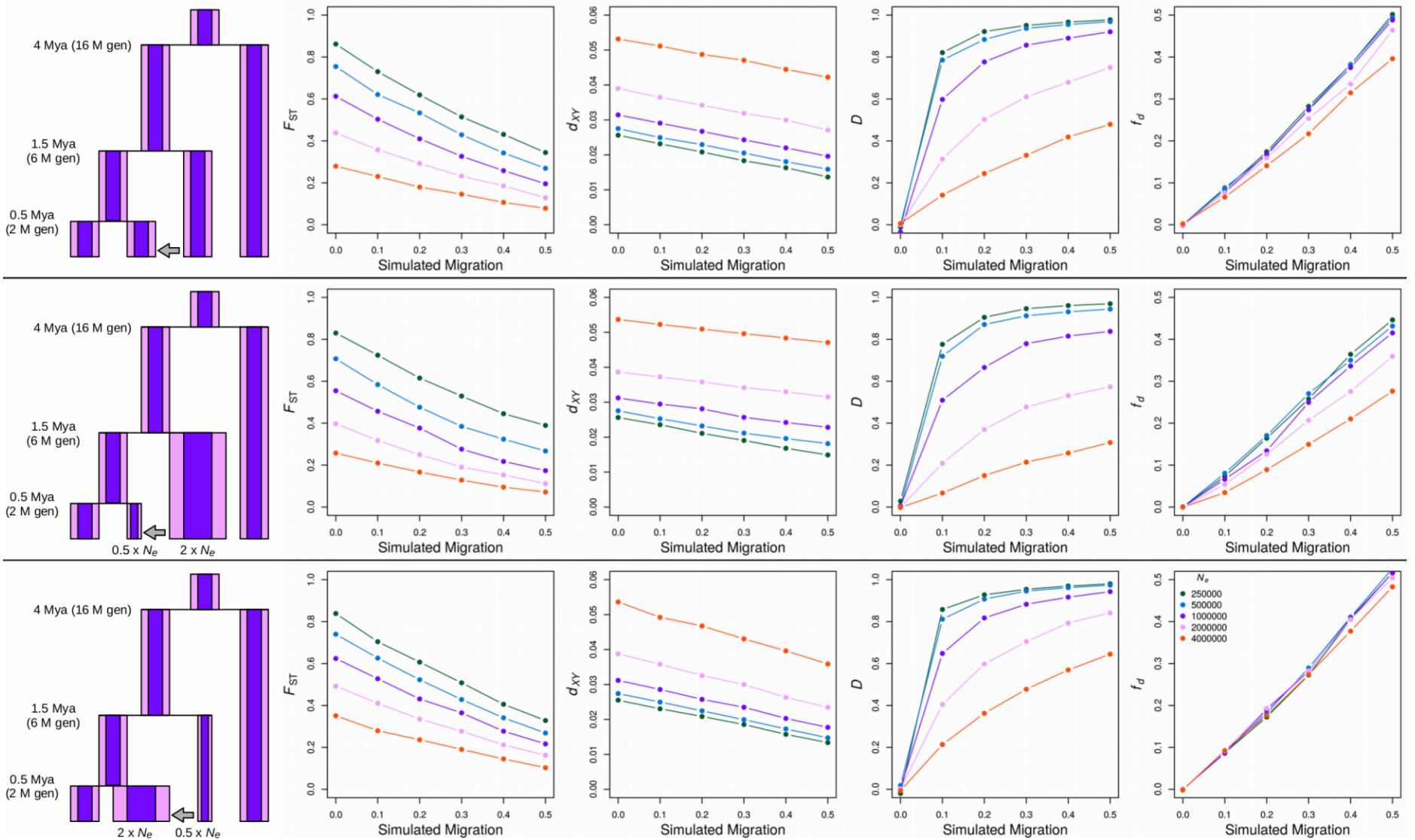
allopatric

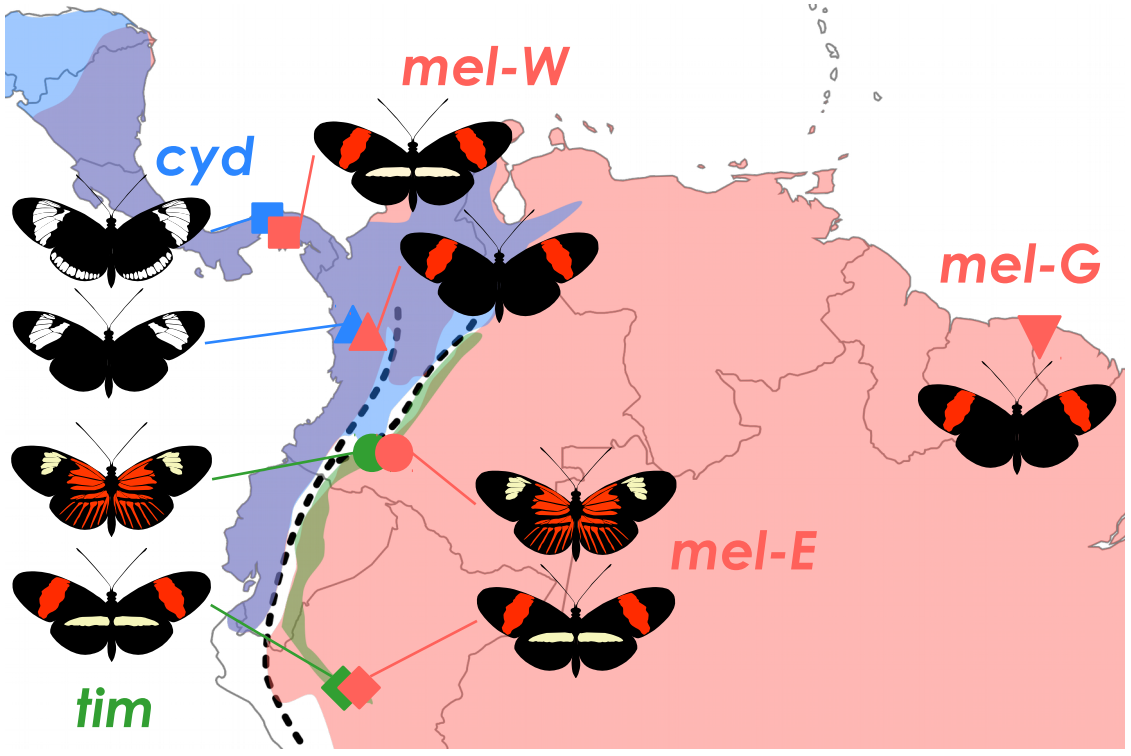


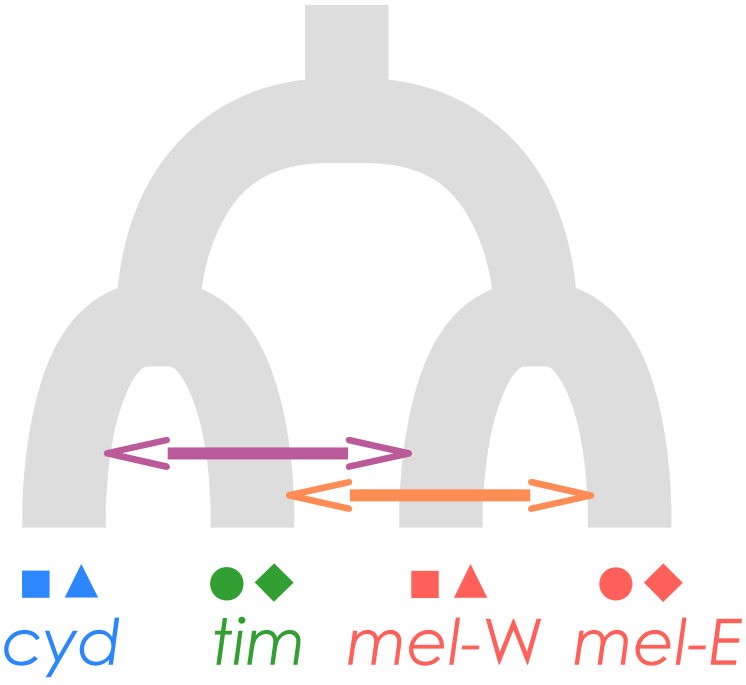
Martin et al. 2015 Mol. Biol. Evol.

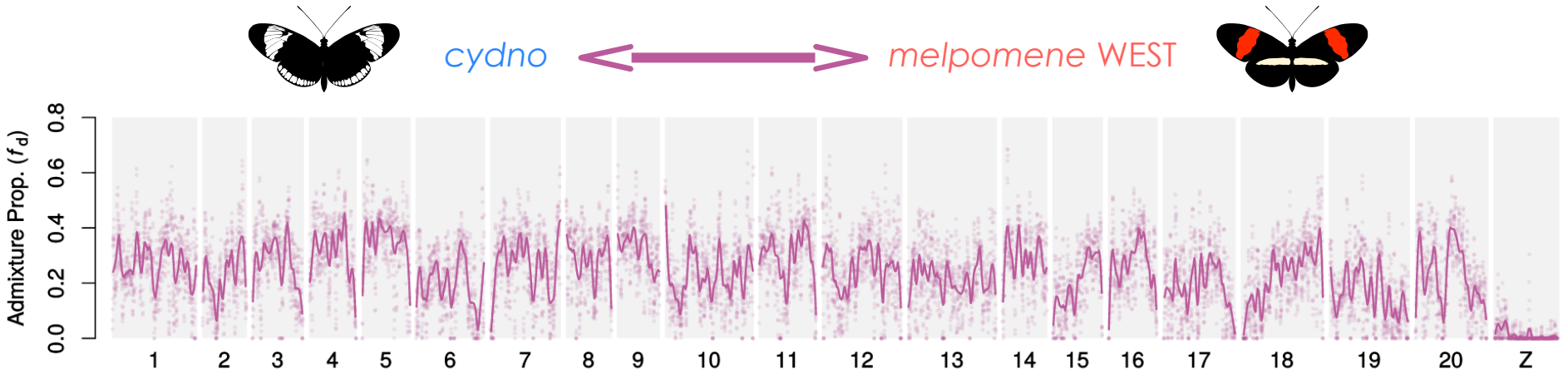


Martin et al. 2015 Mol. Biol. Evol., Martin et al. 2019 PLOS Biol.

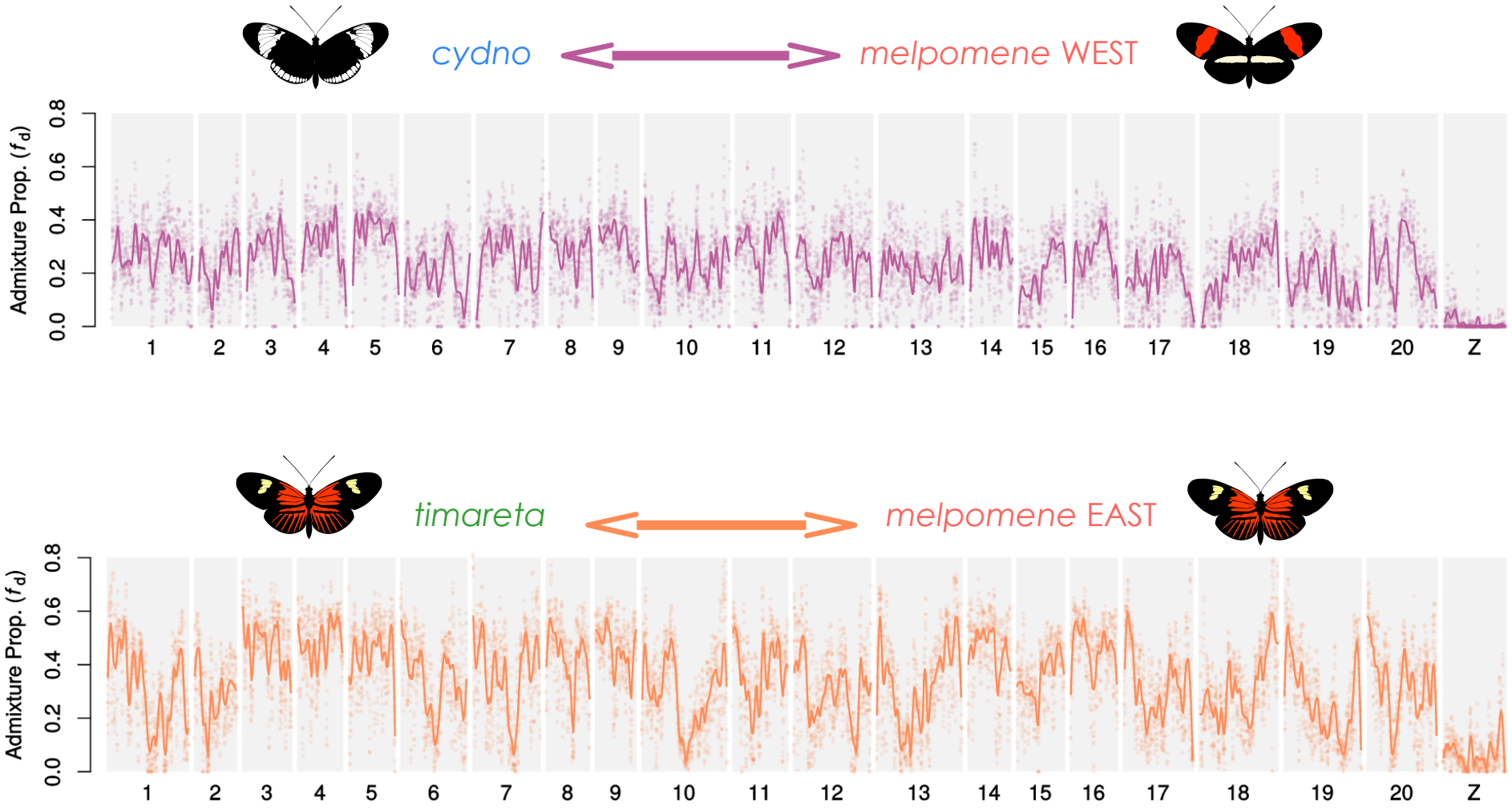

4 Mya (16 M gen)
1.5 Mya (6 M gen)
0.5 Mya (2 M gen)
0.5 x N_e
2 x N_e
F_{ST}
d_{XY}
D
f_d
Simulated Migration
N_e
250000
500000
1000000
2000000
4000000


mel-W
cyd
mel-G
mel-E
tim

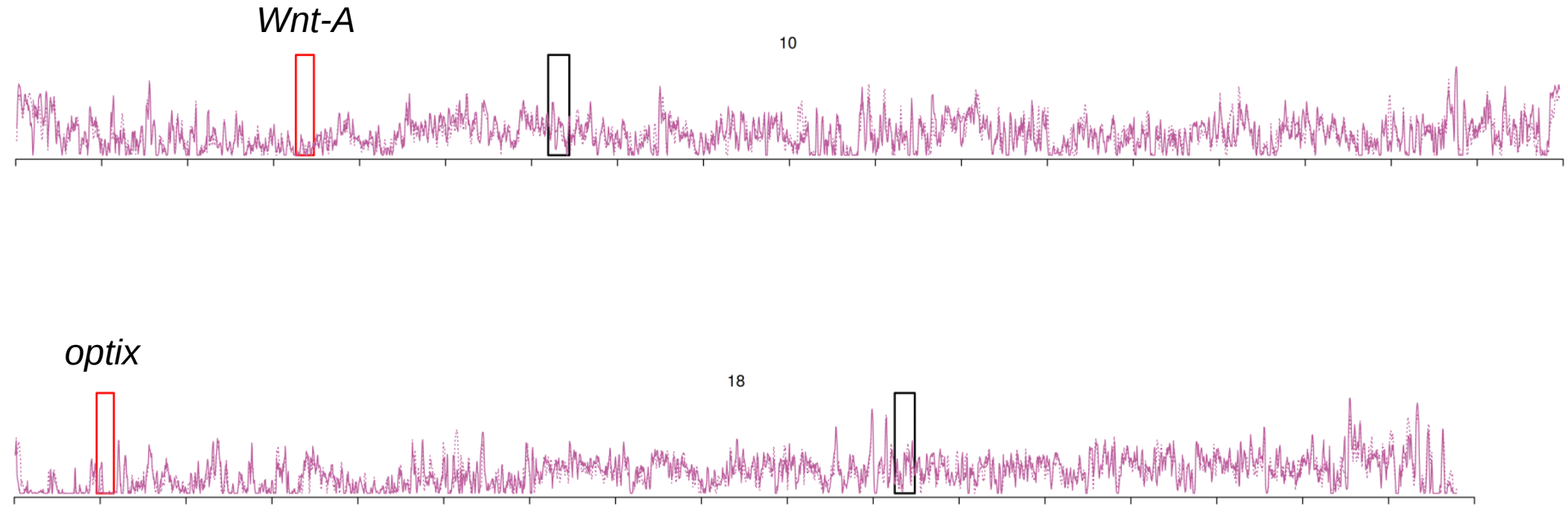

cyd
tim
mel-W
mel-E

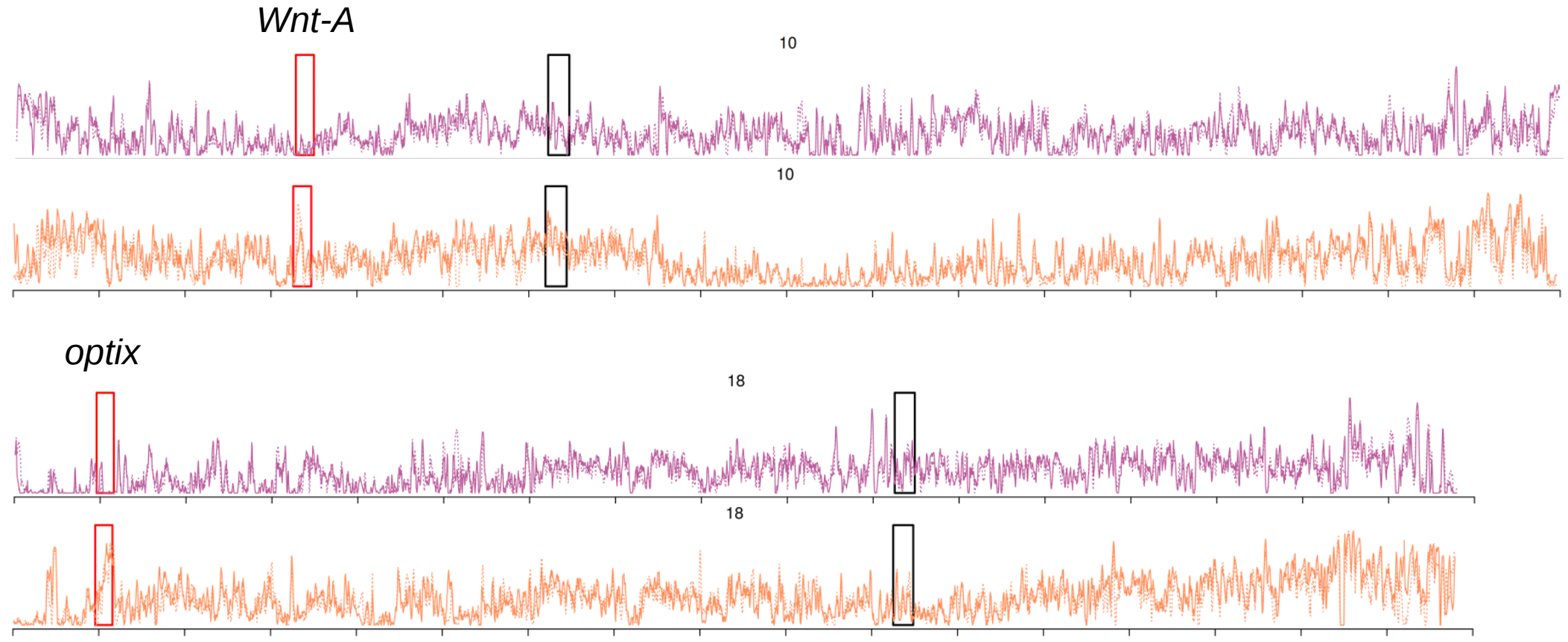


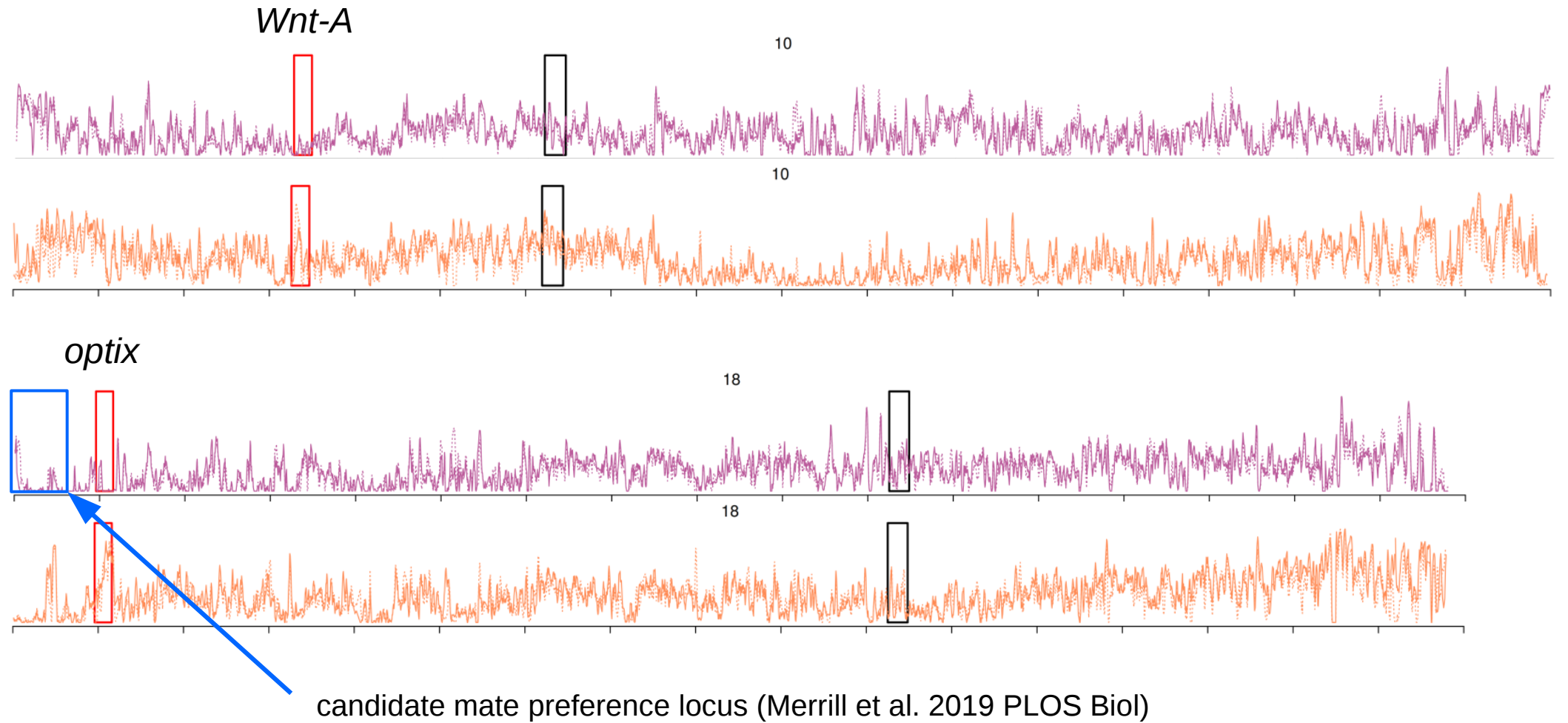
Martin et al. 2019 Plos Biology

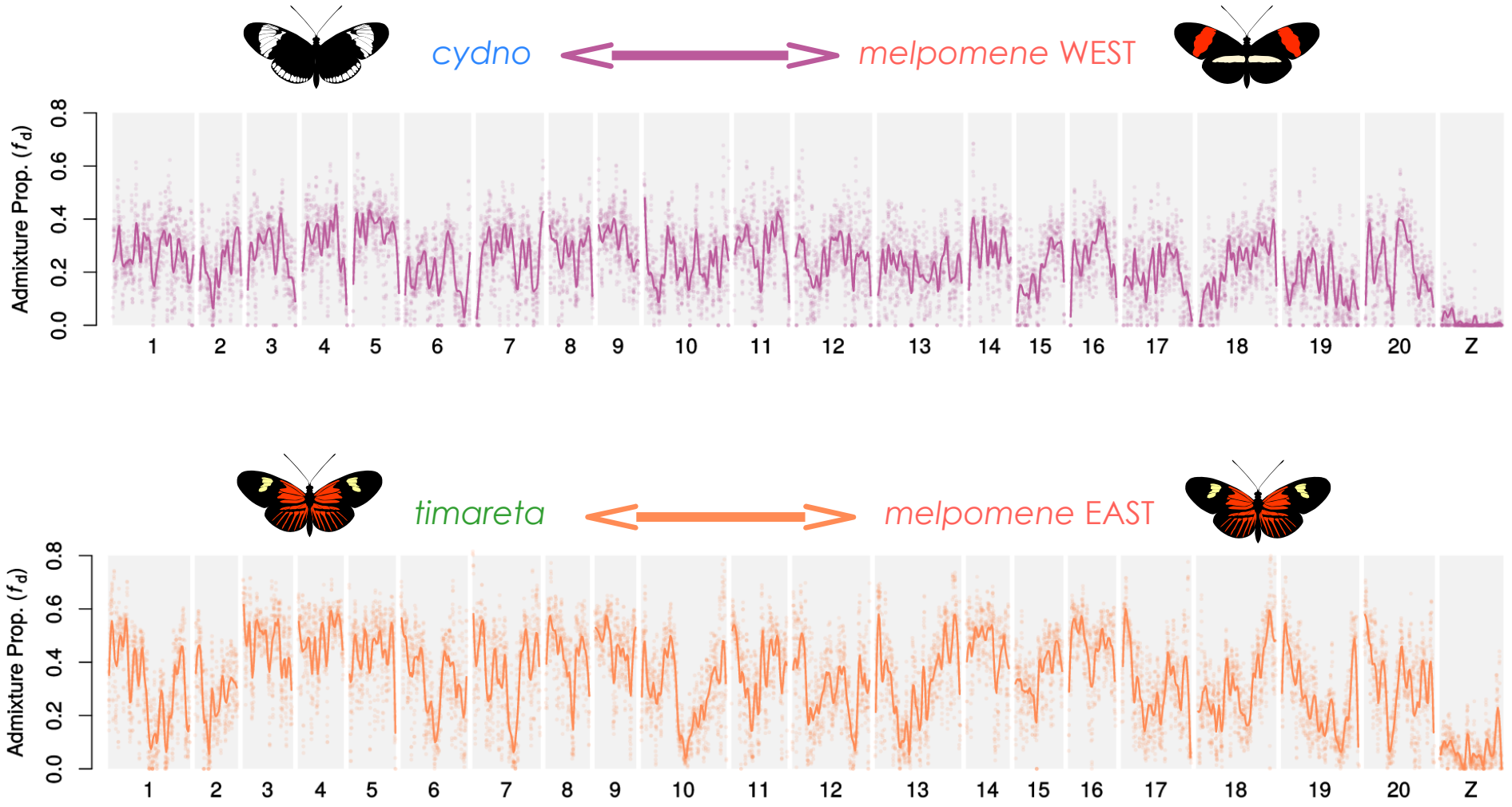


Martin et al. 2019 Plos Biology


Wnt-A
10
optix
18


Wnt-A
10
10
optix
18
18


Wnt-A
10
10
optix
18
18
candidate mate preference locus (Merrill et al. 2019 PLOS Biol)



Martin et al. 2019 Plos Biology

THE QTN PROGRAM AND THE ALLELES THAT MATTER FOR EVOLUTION: ALL THAT'S GOLD DOES NOT GLITTER

Matthew V. Rockman[1,2]

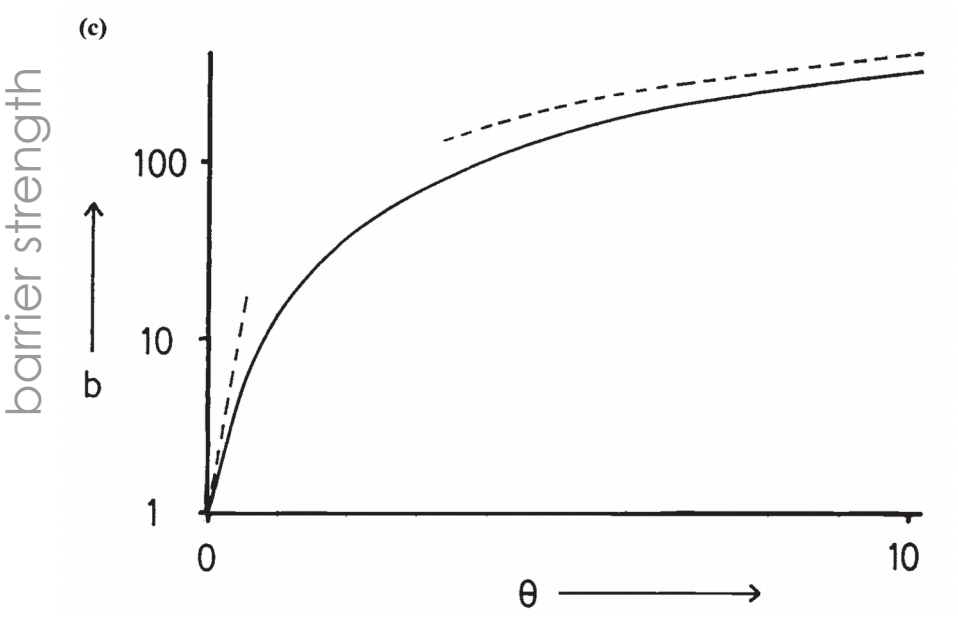
[1]*Department of Biology and Center for Genomics and Systems Biology, New York University, 12 Waverly Place, New York, NY 10003*

[2]*E-mail: mrockman@nyu.edu*

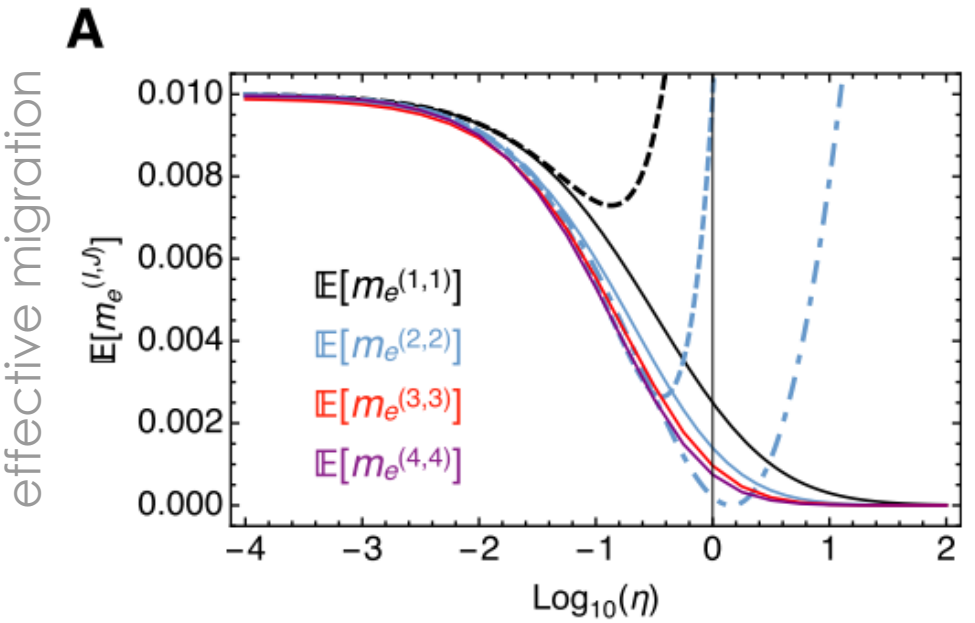
Received April 20, 2011
Accepted September 30, 2011

The search for the alleles that matter, the quantitative trait nucleotides (QTNs) that underlie heritable variation within populations and divergence among them, is a popular pursuit. But what is the question to which QTNs are the answer? Although their pursuit is often invoked as a means of addressing the molecular basis of phenotypic evolution or of estimating the roles of evolutionary forces, the QTNs that are accessible to experimentalists, QTNs of relatively large effect, may be uninformative about these issues if large-effect variants are unrepresentative of the alleles that matter. Although 20th century evolutionary biology generally viewed large-effect variants as atypical, the field has recently undergone a quiet realignment toward a view of readily discoverable large-effect alleles as the primary molecular substrates for evolution. I argue that neither theory nor data justify this realignment. Models and experimental findings covering broad swaths of evolutionary phenomena suggest that evolution often acts via large numbers of small-effect polygenes, individually undetectable. Moreover, these small-effect variants are different in kind, at the molecular level, from the large-effect alleles accessible to experimentalists. Although discoverable QTNs address some fundamental evolutionary questions, they are essentially misleading about many others.

Recombination predicts barrier strength

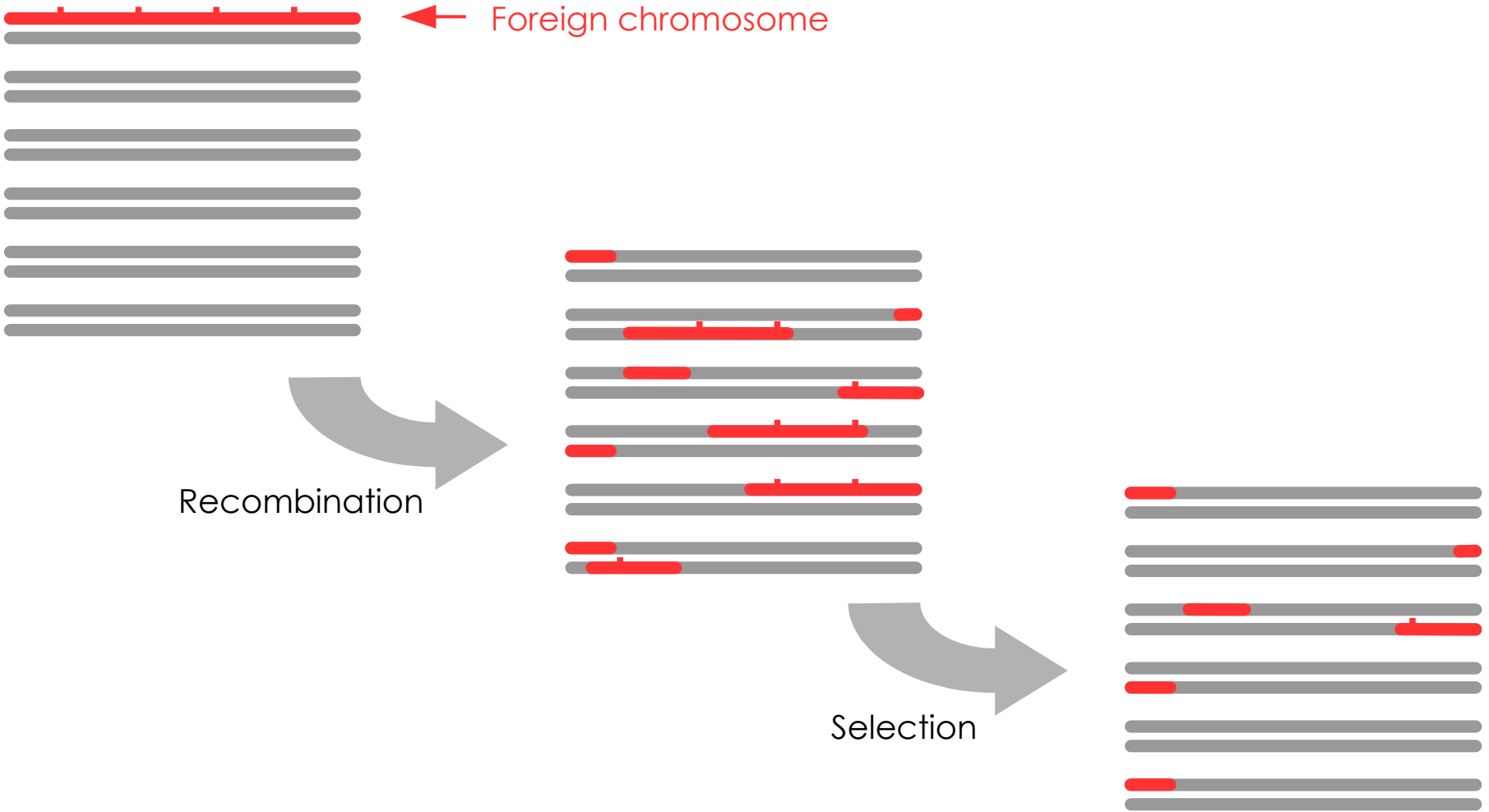


Barton & Bengtsson 1986

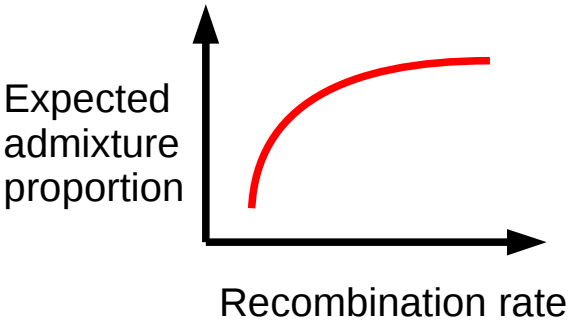


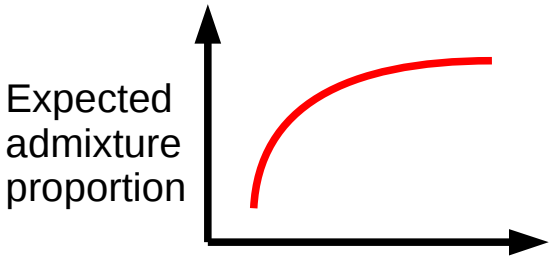
Aeschbacher et al. 2017

A simple simulation example

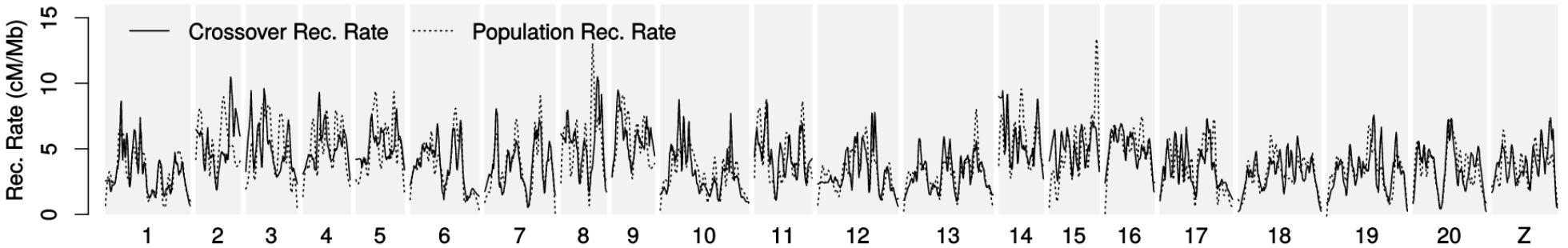


Barton & Bengtsson 1986, Aeschbacher et al. 2017

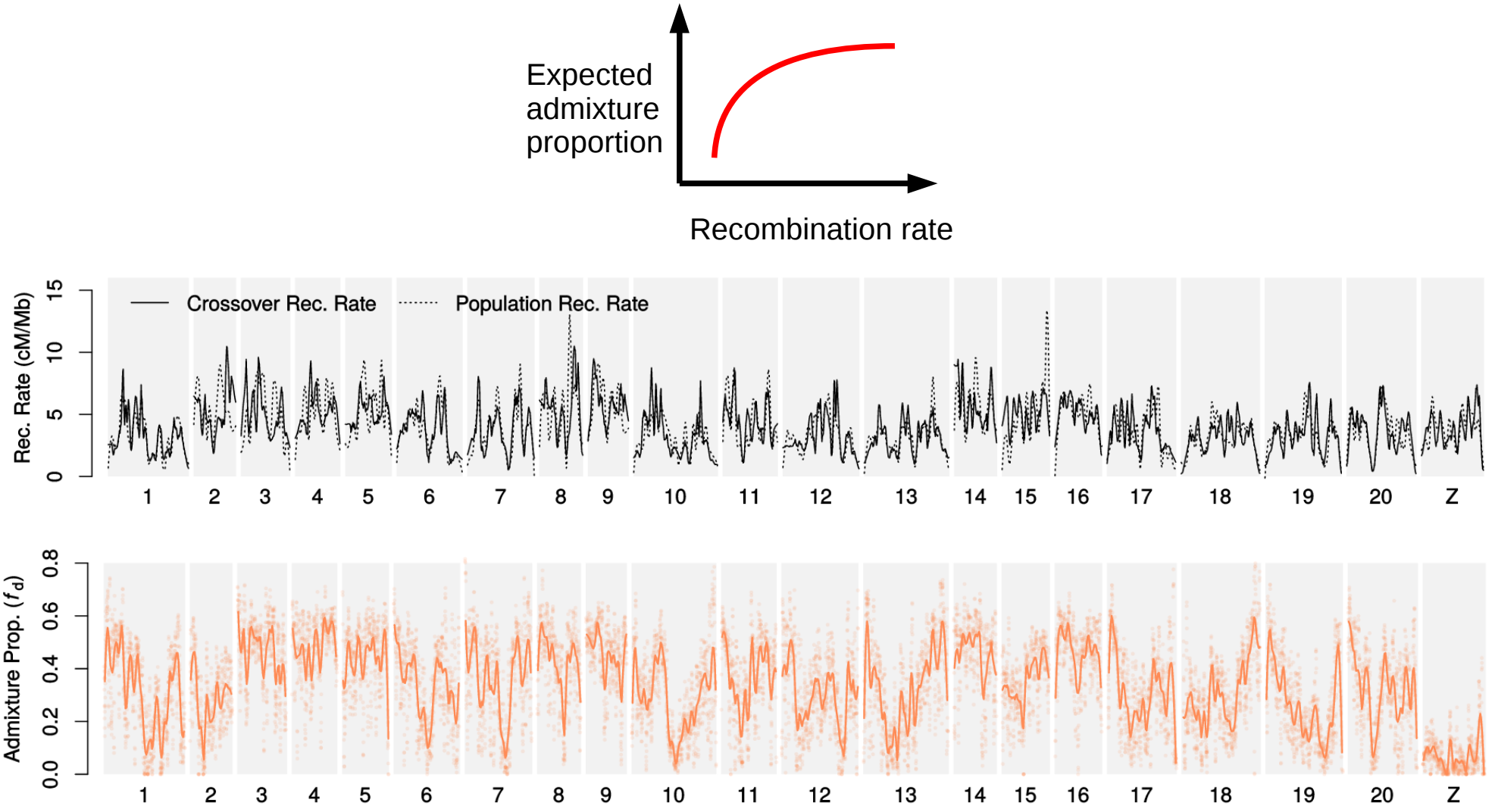

Expected
admixture
proportion
Recombination rate

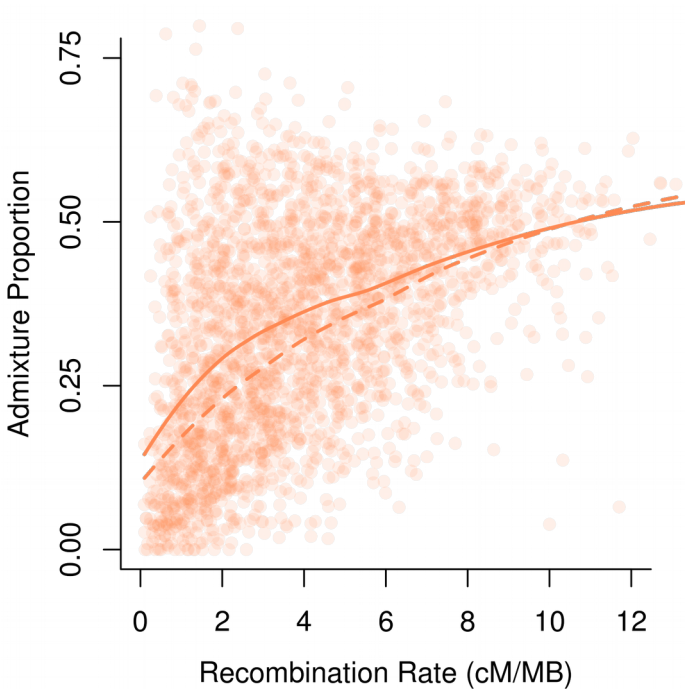


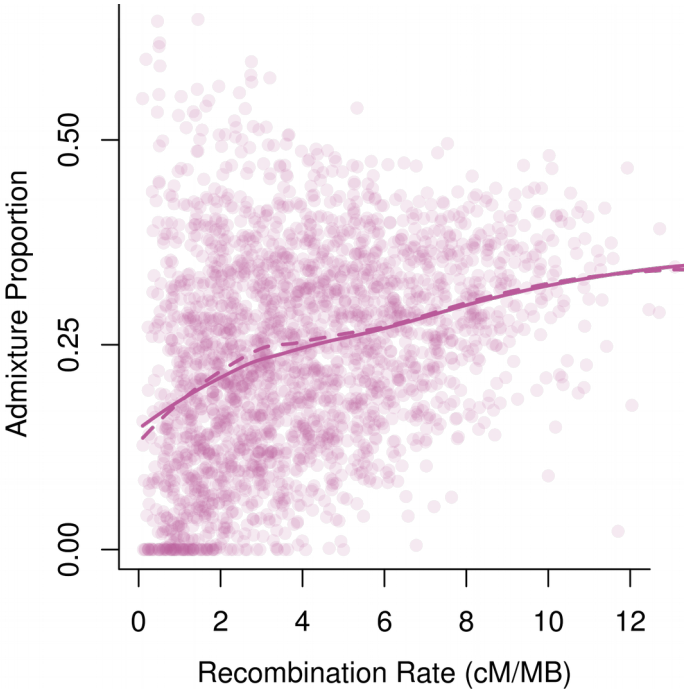
Recombination rate



Davey et al. 2017, Martin et al. 2019 Plos Biology

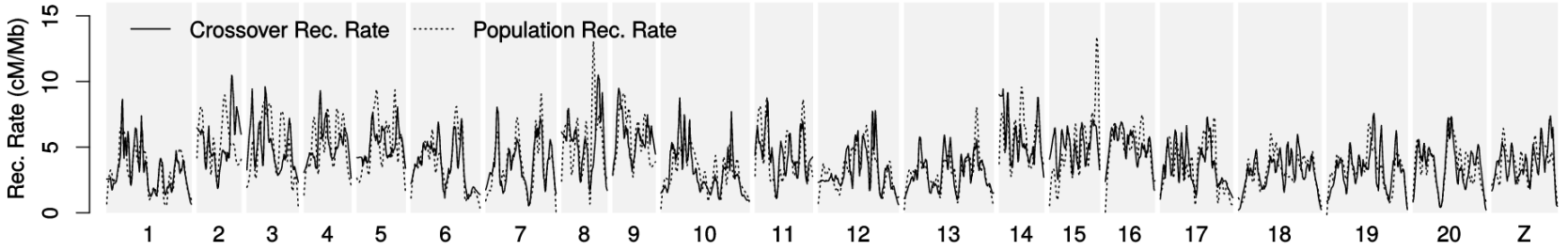

Expected
admixture
proportion
Recombination rate
Rec. Rate (cM/Mb)
0
5
10
15
Crossover Rec. Rate
Population Rec. Rate
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Z
Admixture Prop. (f_d)
0.0
0.2
0.4
0.6
0.8

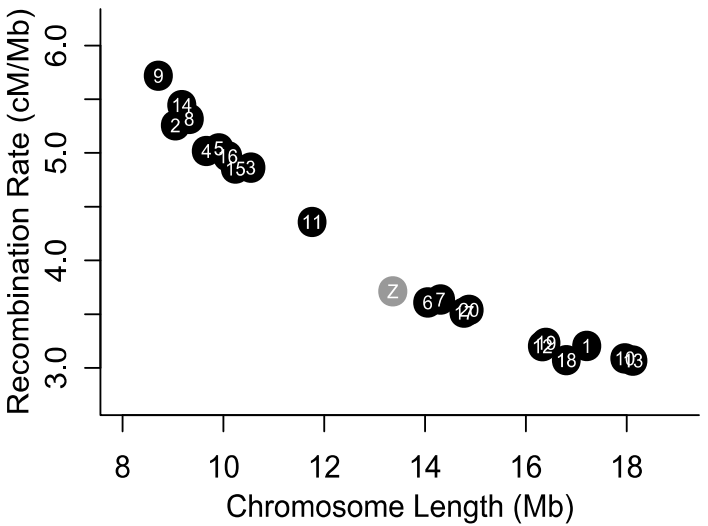




The species barrier is more porous where recombination rates are higher

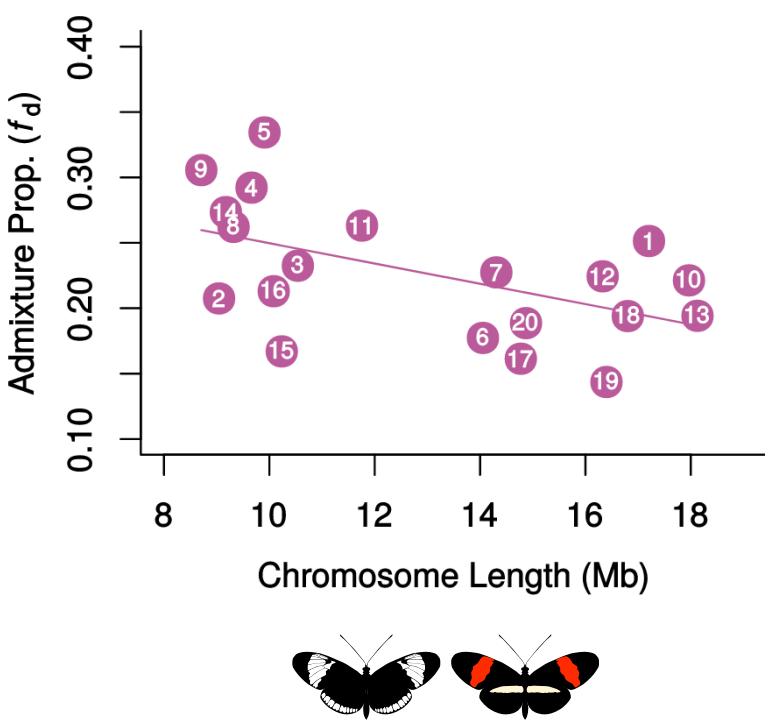
Recombination also varies at the chromosome scale

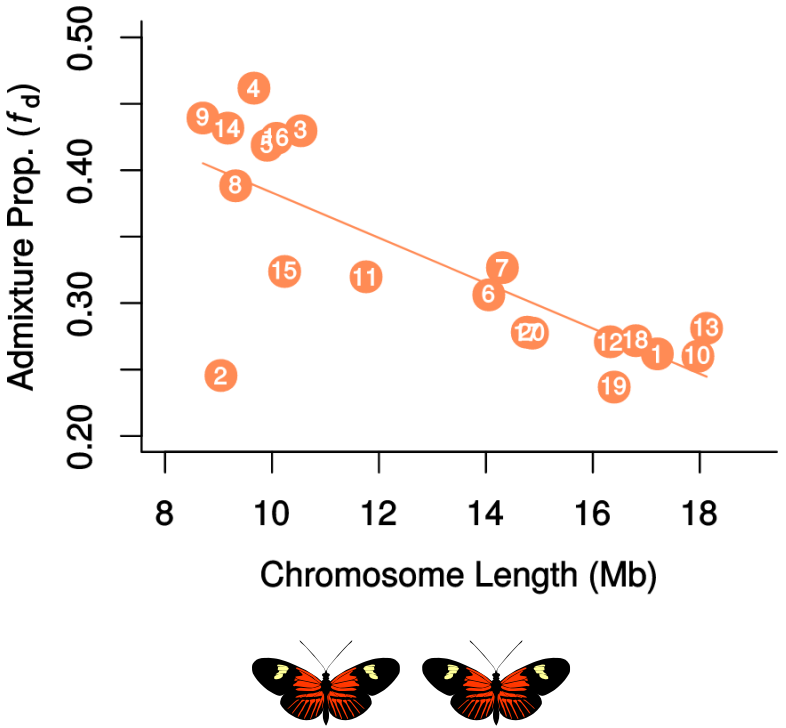




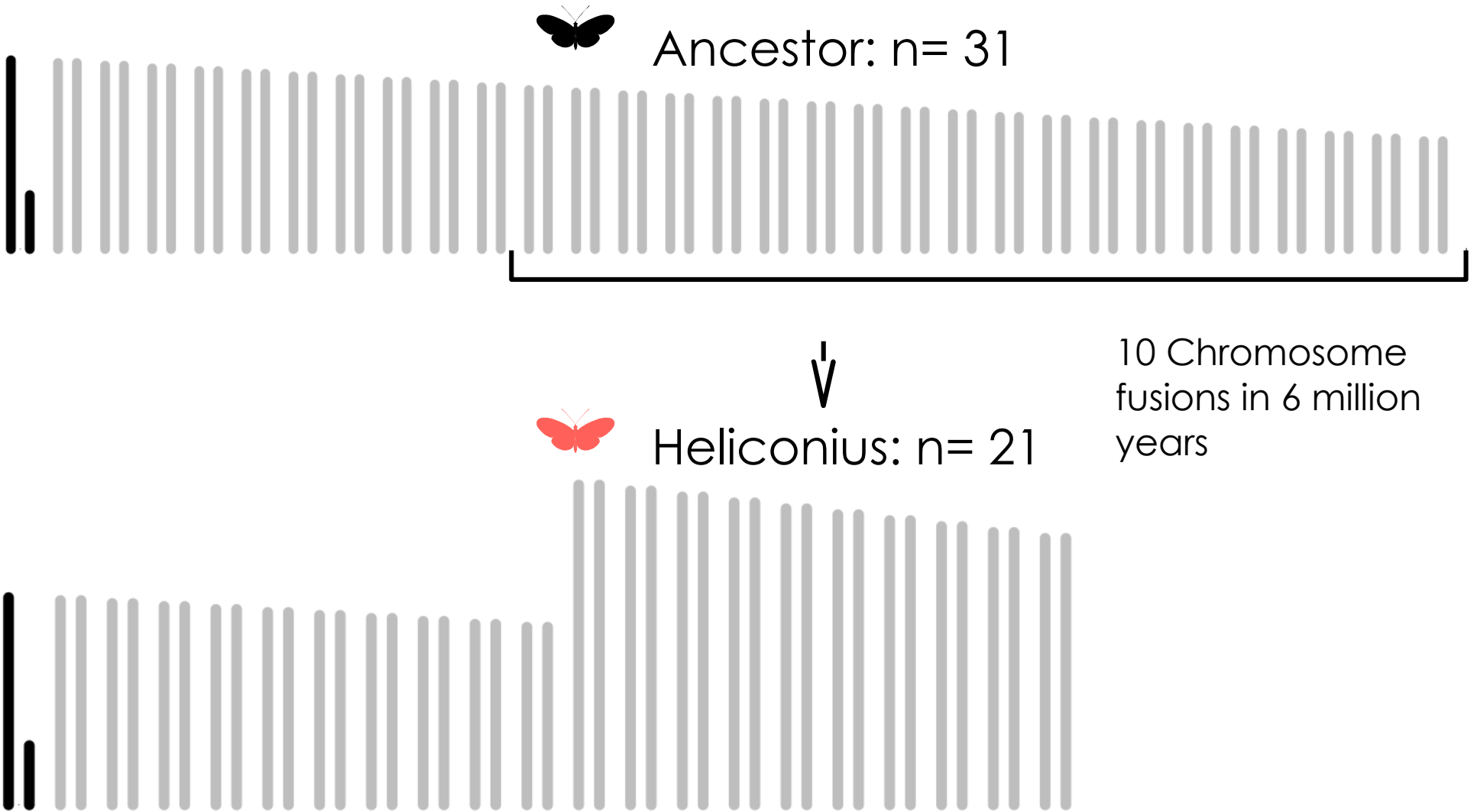
Martin et al. 2019 Plos Biol.

Longer chromosomes form stronger barriers



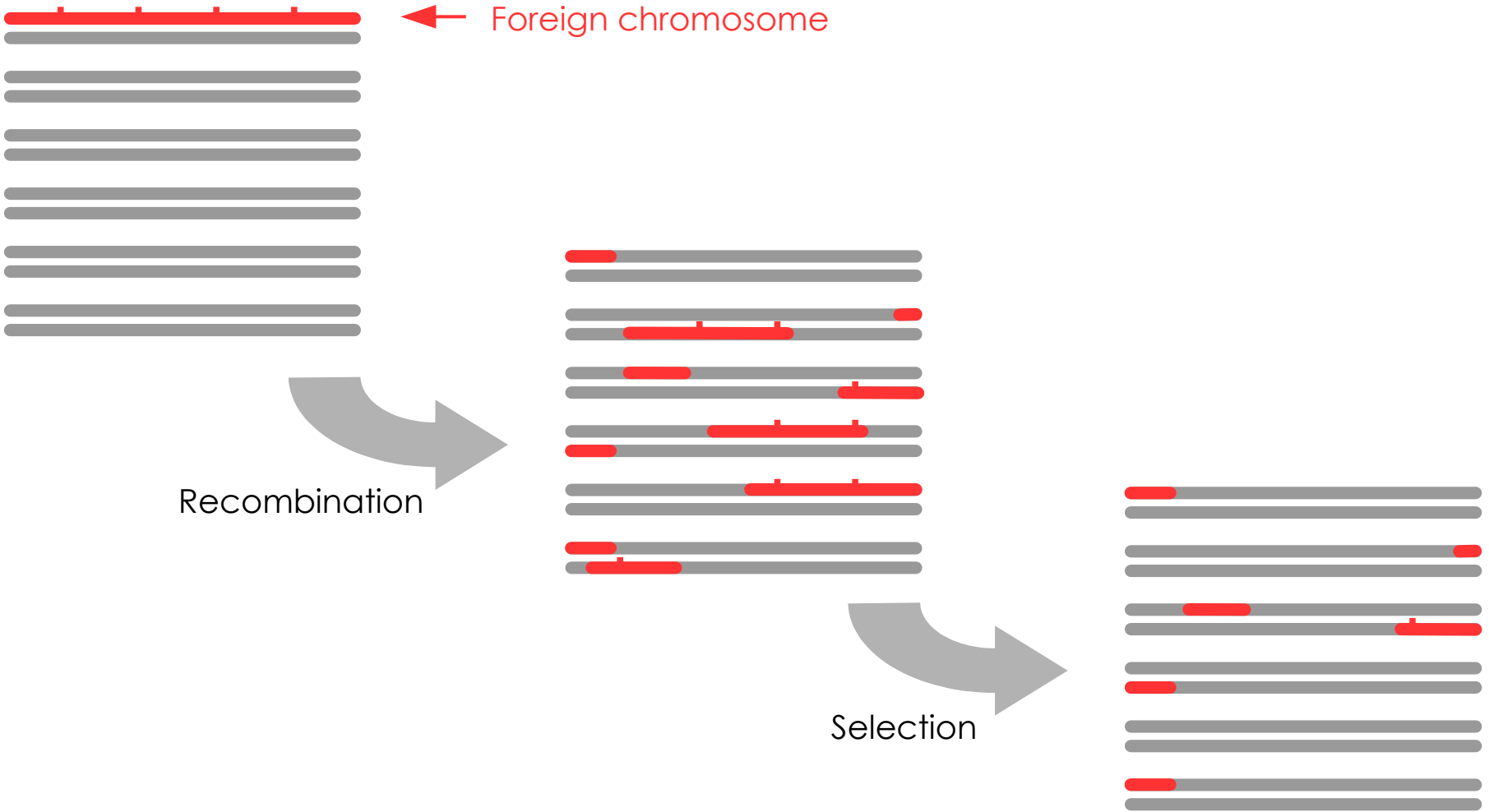


Martin et al. 2019 Plos Biol.



Davey et al. 2016 G3, Ahola et al. 2014 Nature Comm.

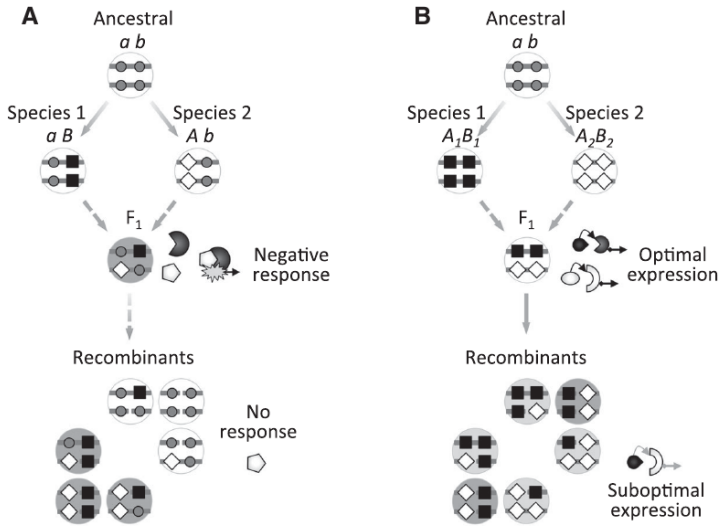
But what about incompatibilities?

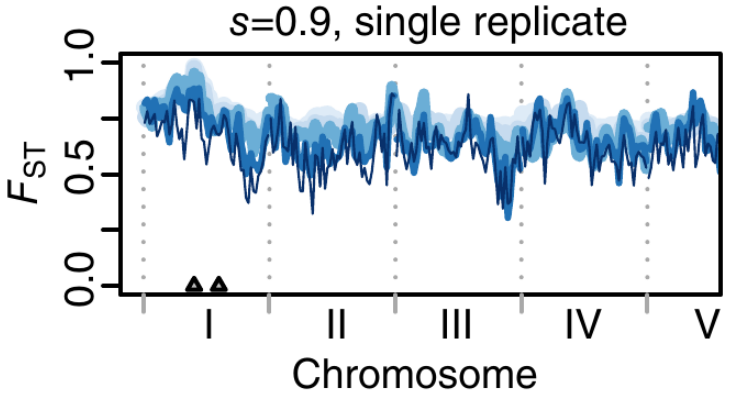


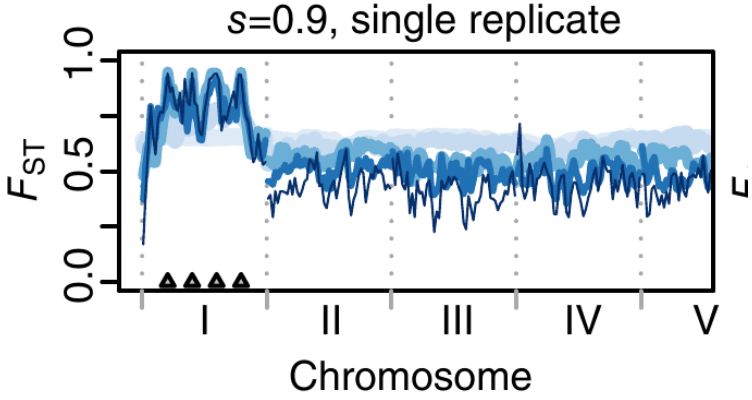
Barton & Bengtsson 1986, Aeschbacher et al. 2017

The genetic architecture of hybrid incompatibilities and their effect on barriers to introgression in secondary contact

Dorothea Lindtke[1,2,3] and C. Alex Buerkle[1]



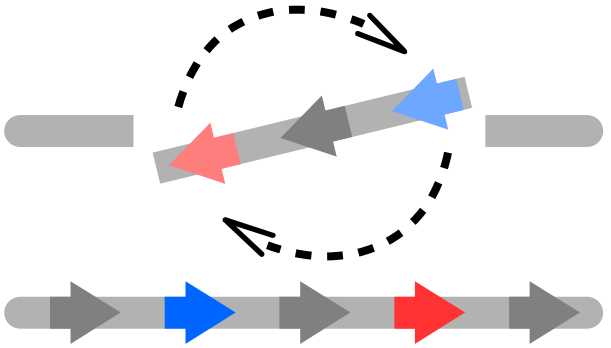


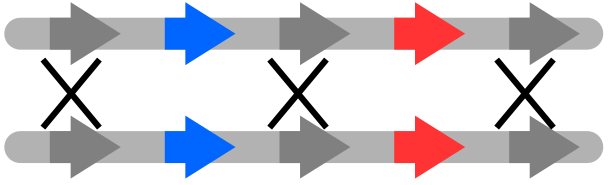


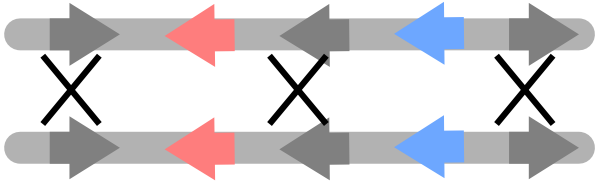
Sometimes gold does glitter

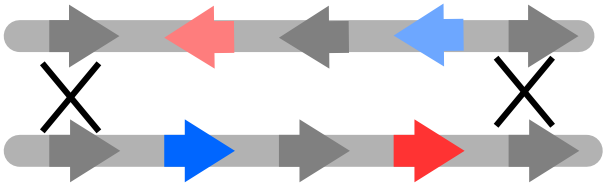
A pleasing ending

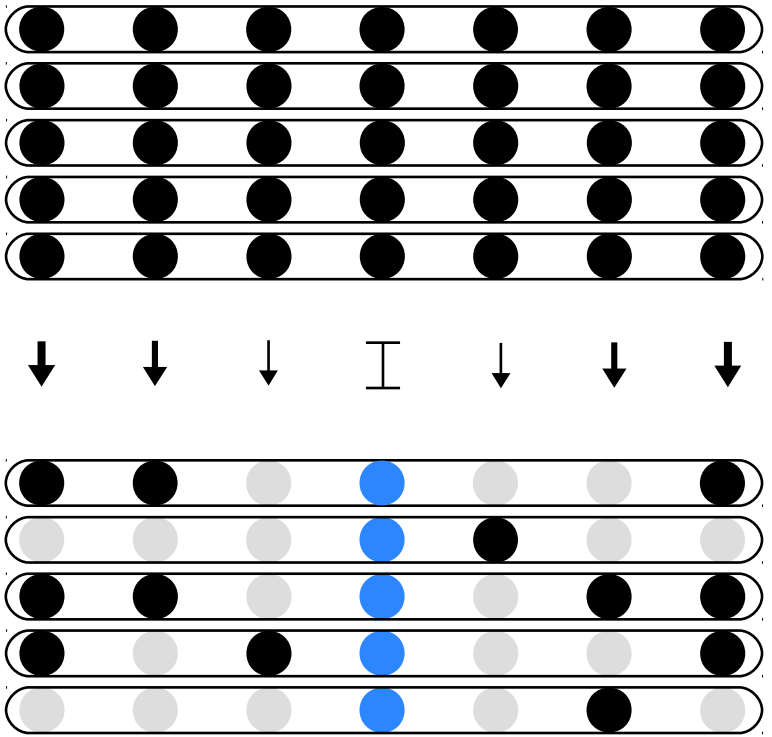
Localised recombination suppression



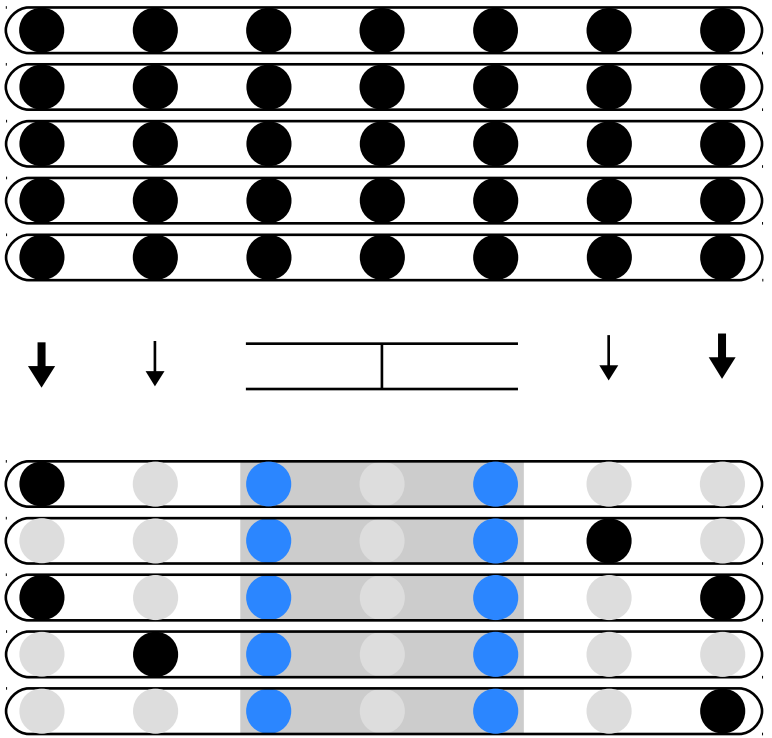






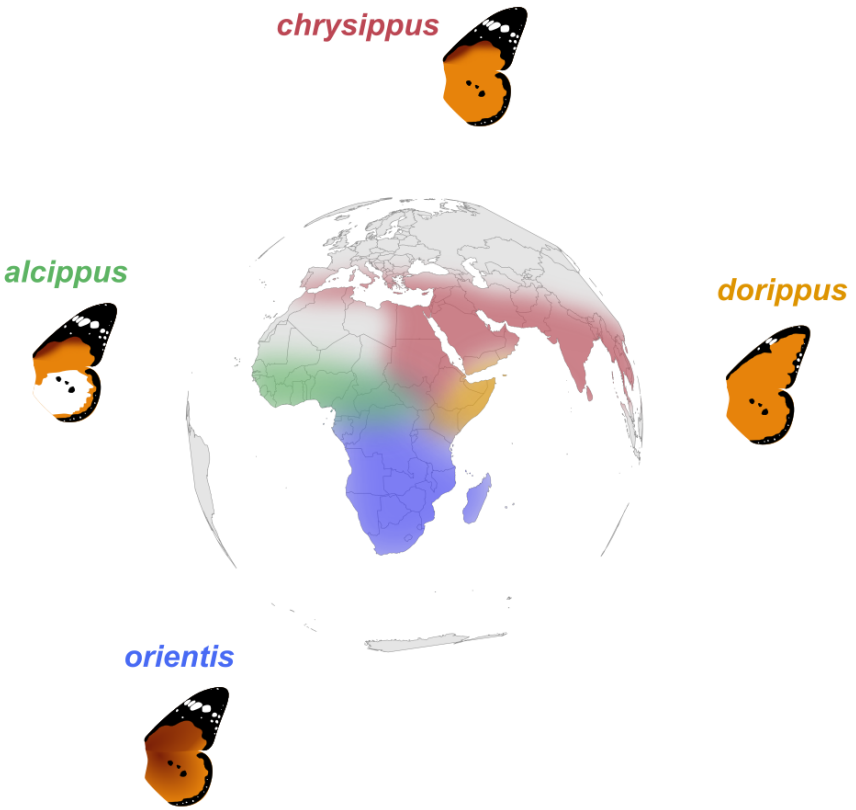


Barton 1979



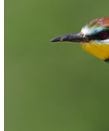
Kirkpatrick and Barton 2006


Danaus chrysippus

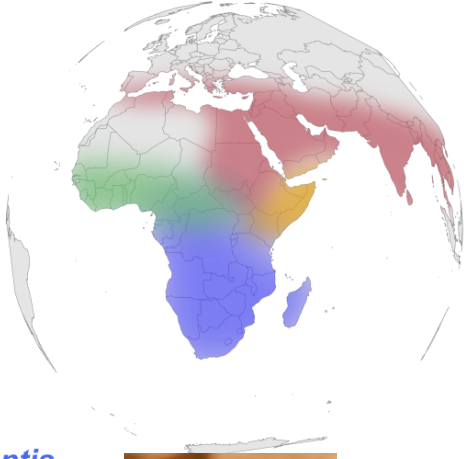

chrysippus
alcippus
dorippus
orientis

EASTERN DRC
POPULATIONS.

chrysippus



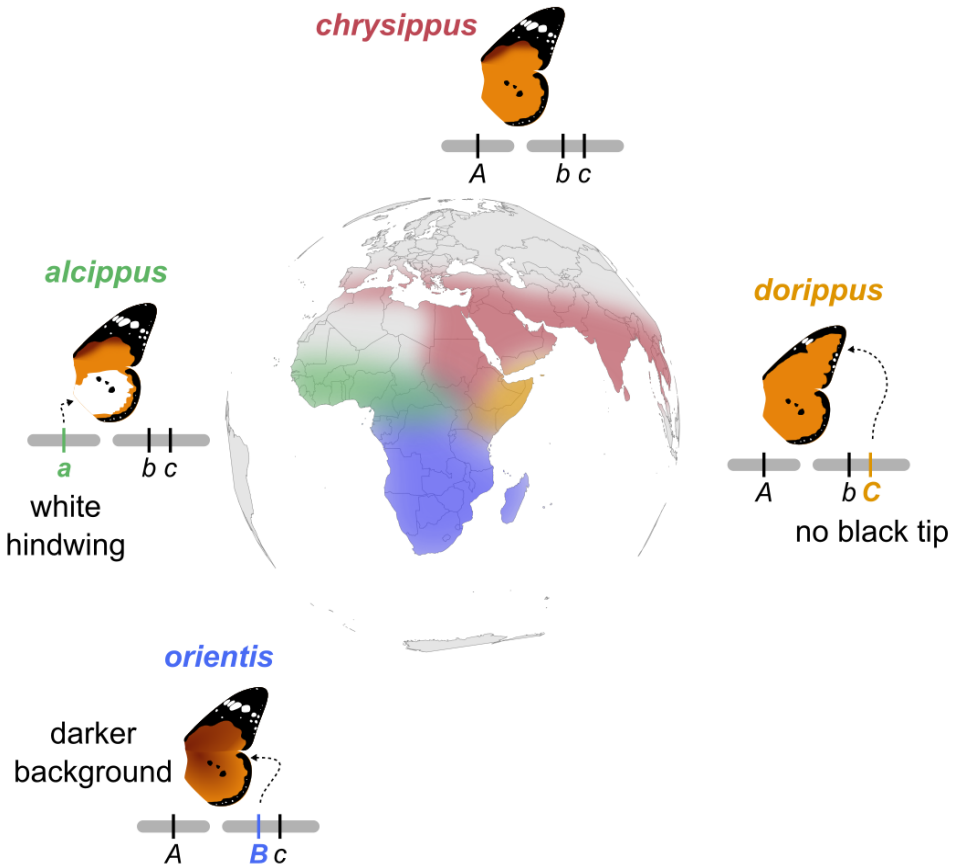
alcippus



dorippus

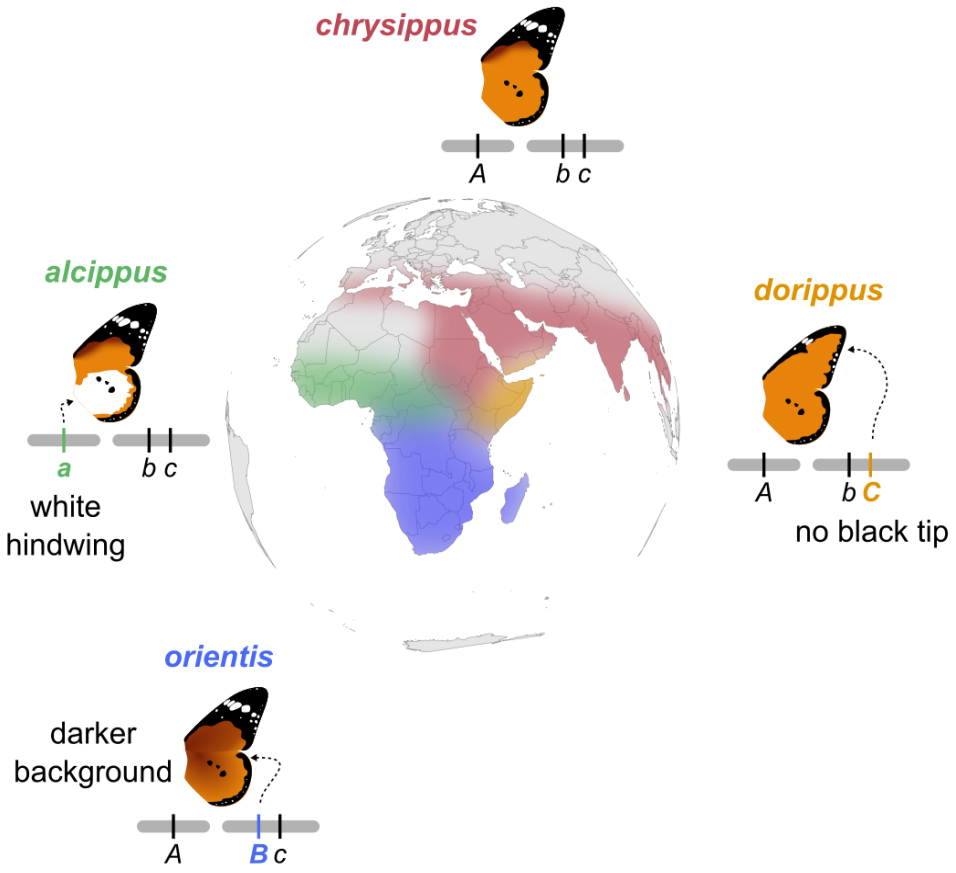
orientis

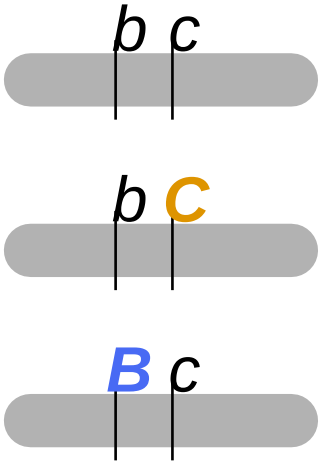
Large effect genes control differences



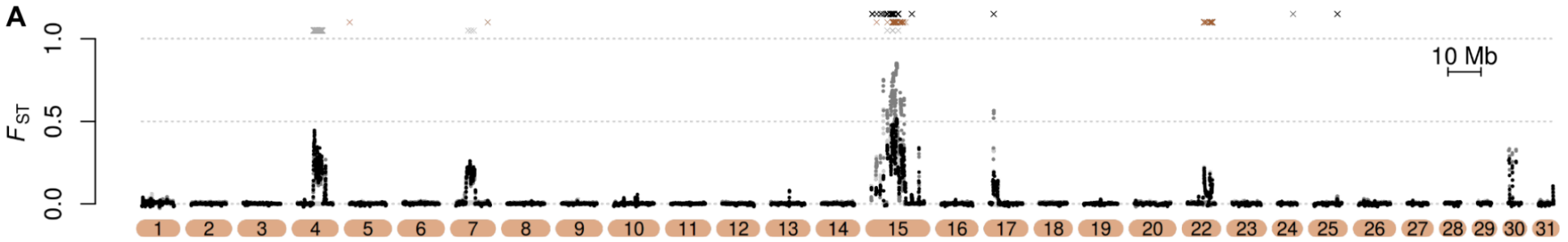
Clarke et al. 1973, Smith 1975,

A supergene?

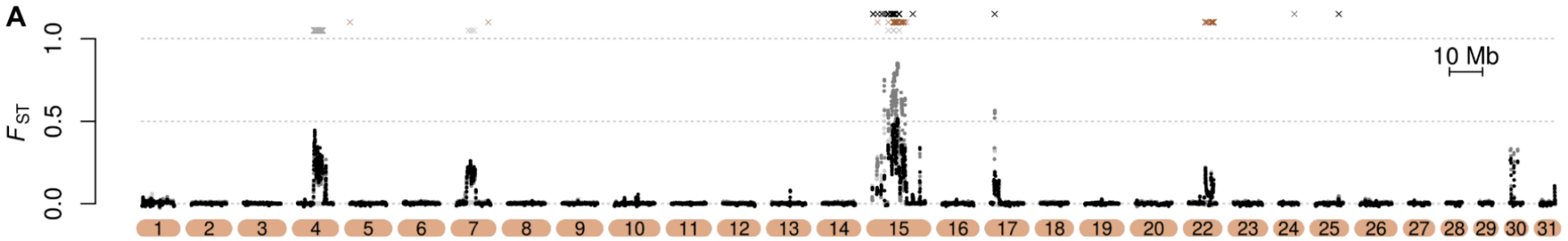


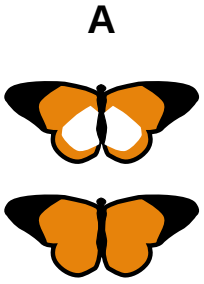


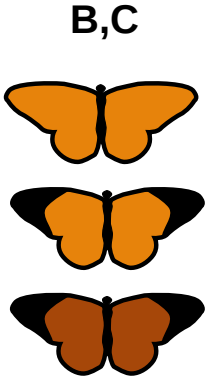
Smith et al. 2010



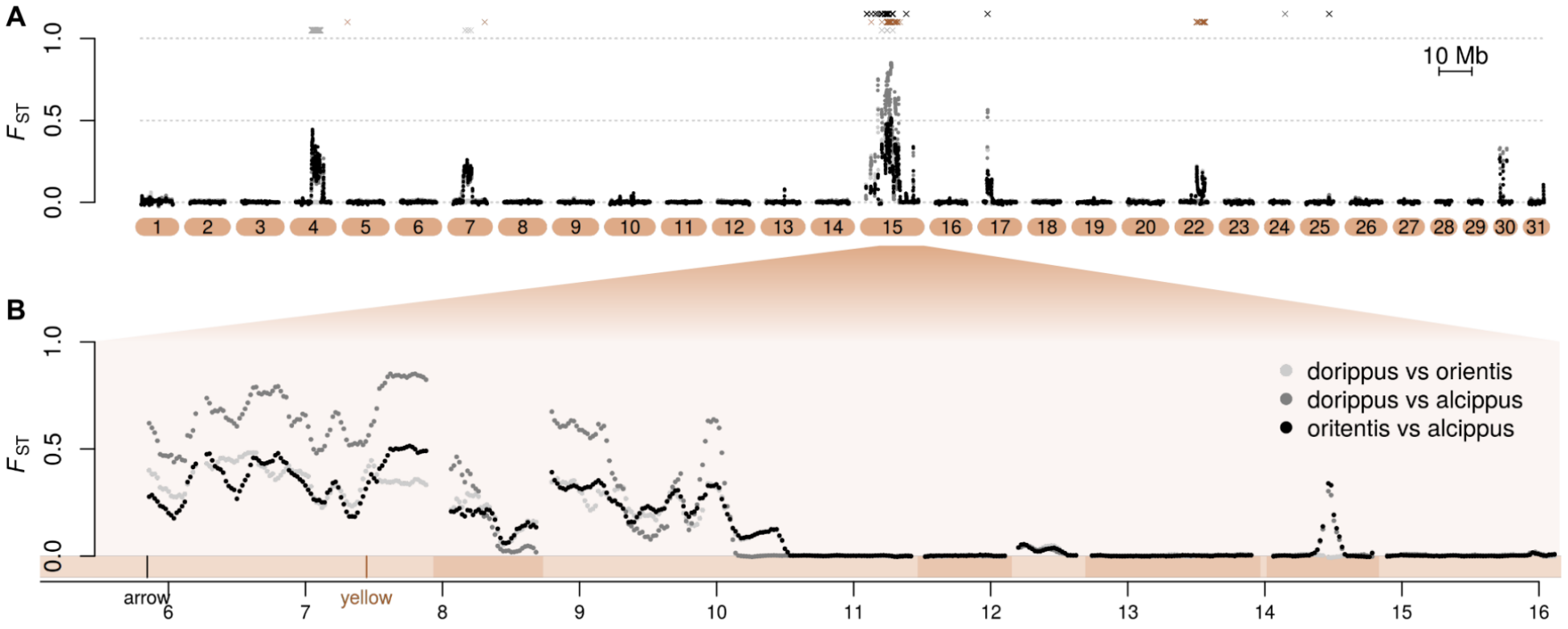
Martin et al. 2020 Plos Biol (in press)



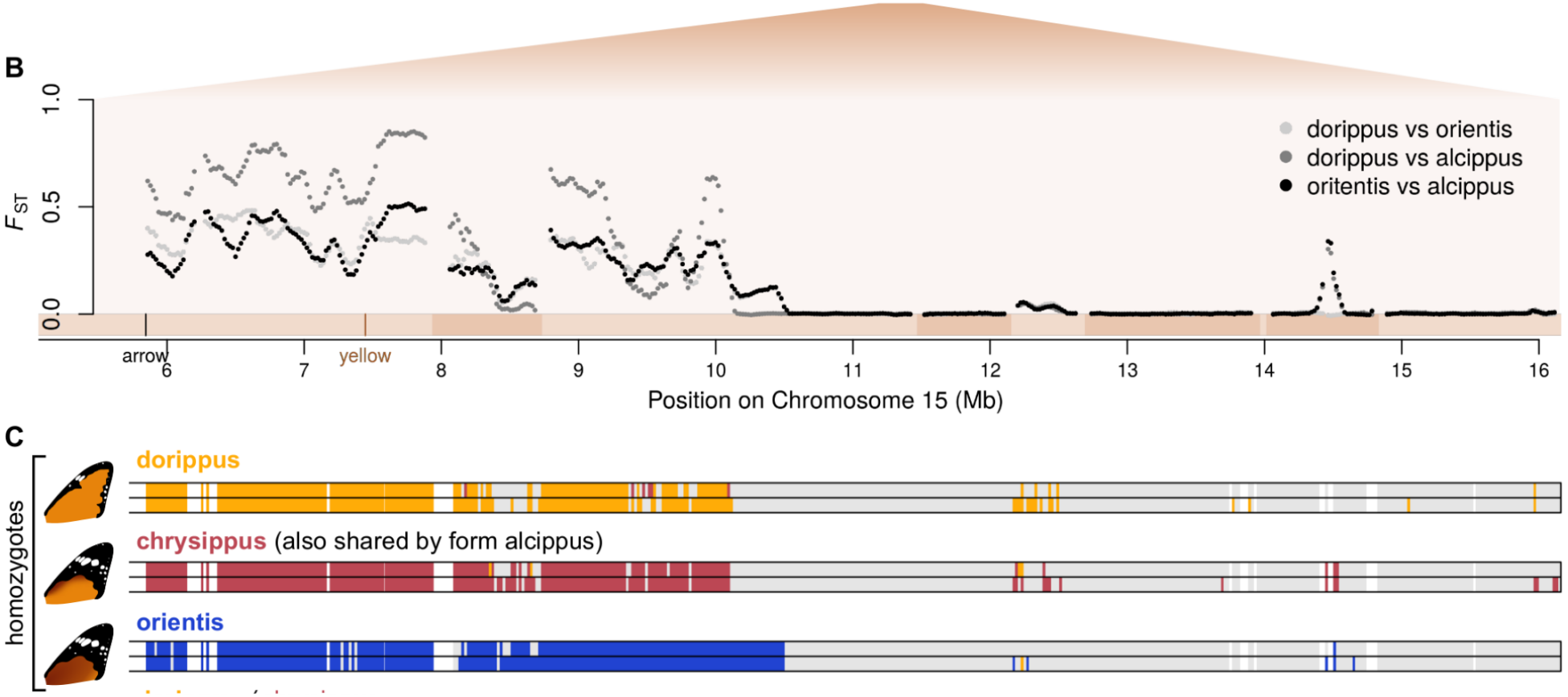




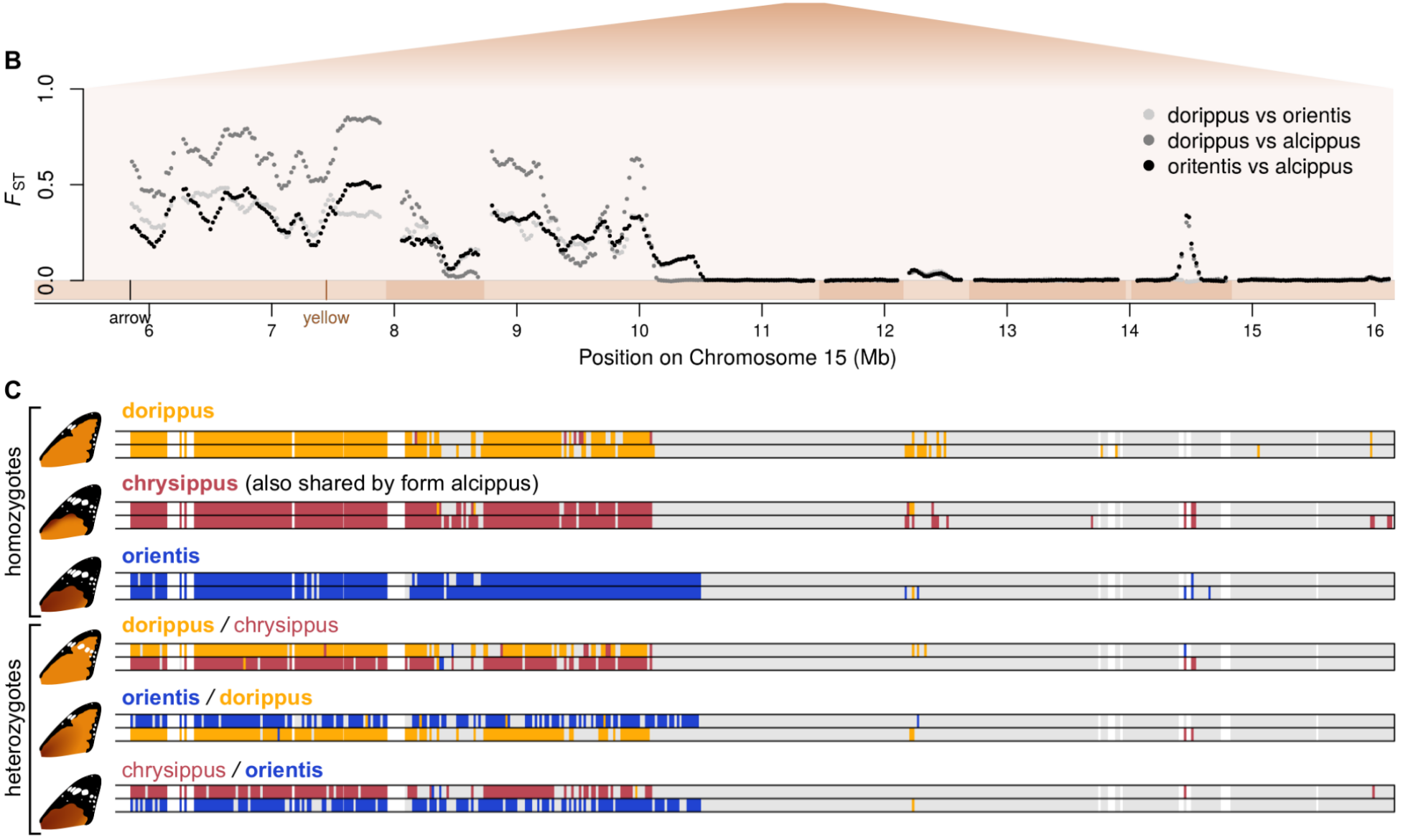
Martin et al. 2020 Plos Biol (in press)



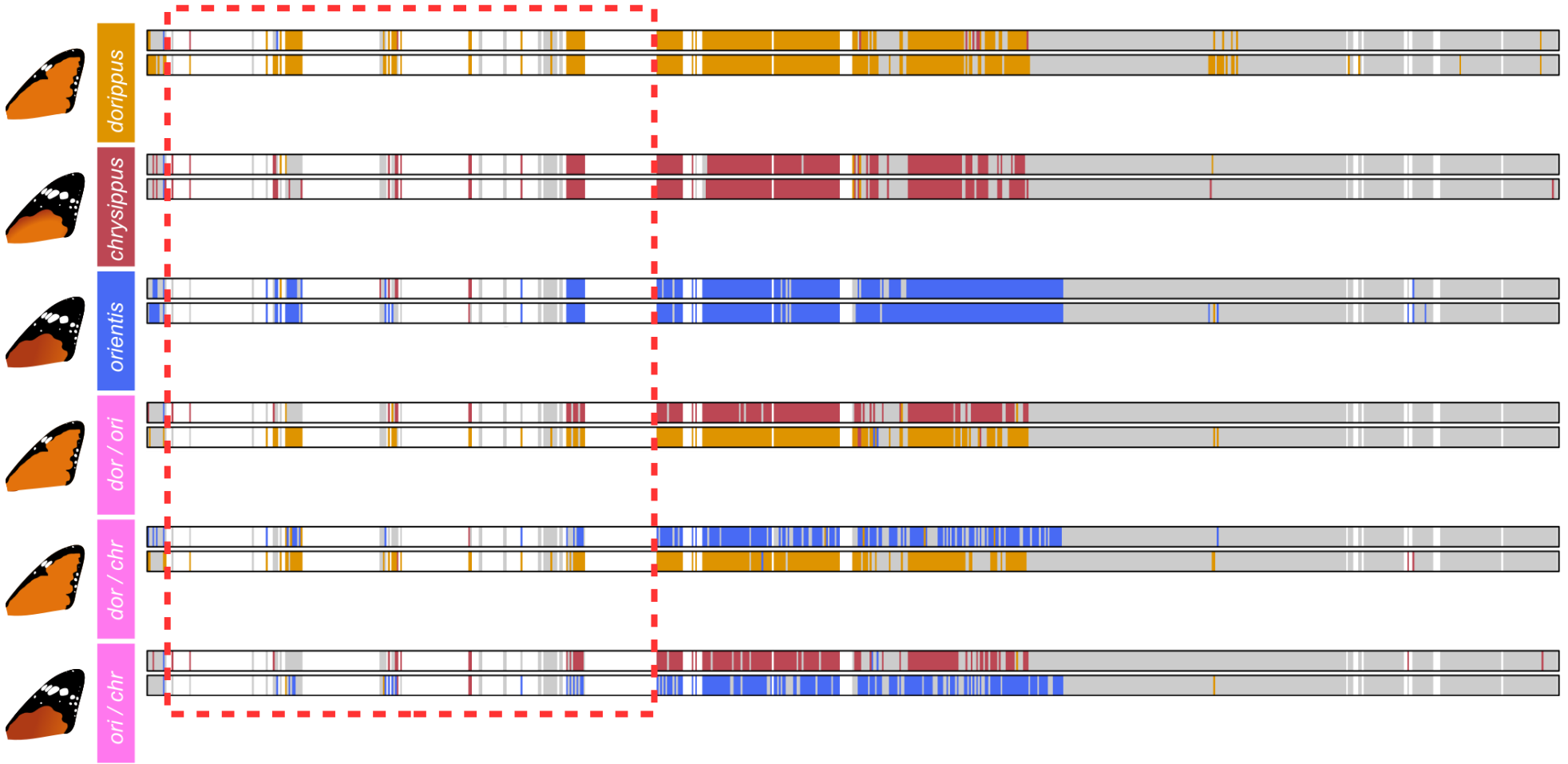
Martin et al. 2020 Plos Biol (in press)



Martin et al. 2020 Plos Biol (in press)

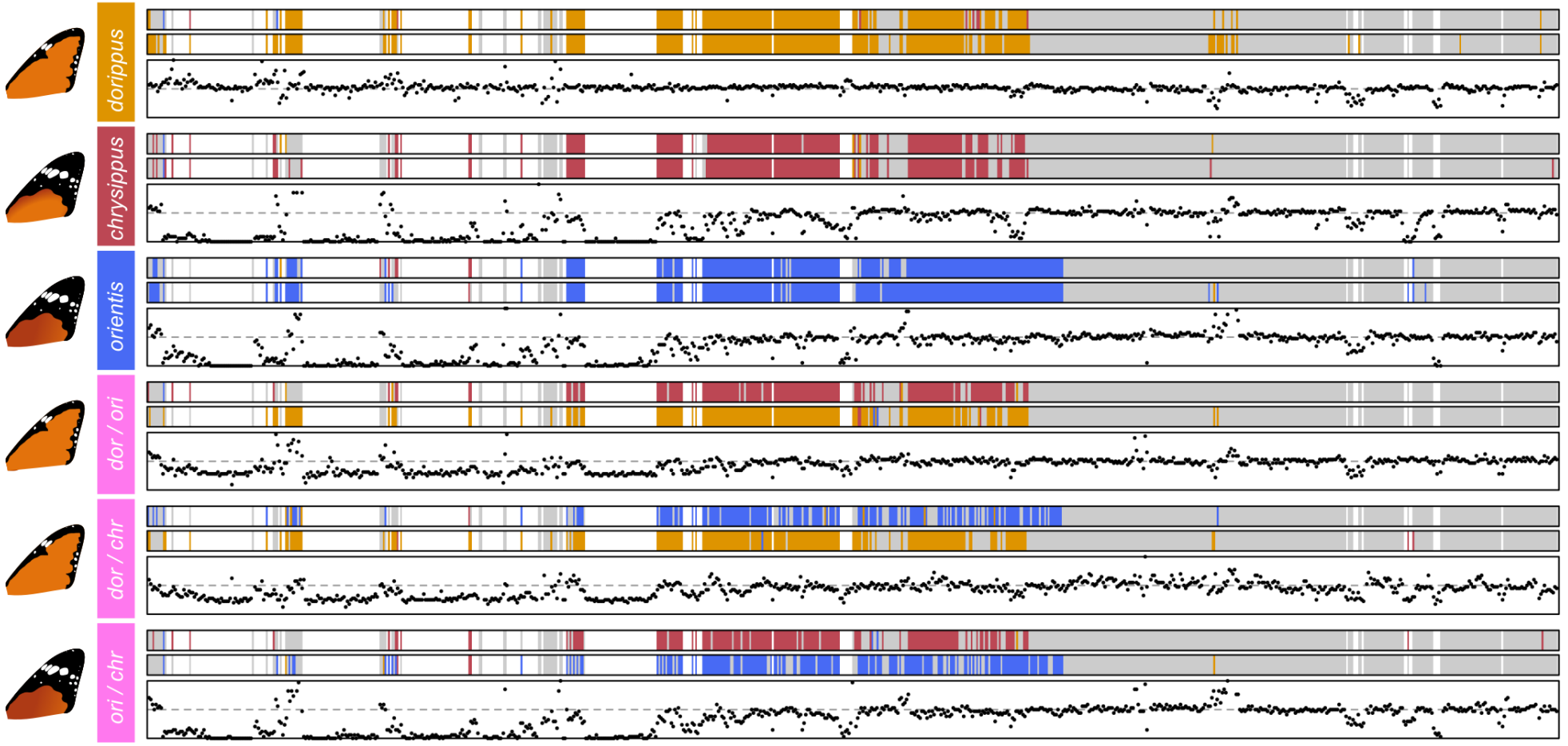

B
1.0
0.5
0.0
F_{ST}
dorippus vs orientis
dorippus vs alcippus
oritentis vs alcippus
arrow
yellow
6
7
8
9
10
11
12
13
14
15
16
Position on Chromosome 15 (Mb)
C
homozygotes
heterozygotes
dorippus
chrysippus (also shared by form alcippus)
orientis
dorippus / chrysippus
orientis / dorippus
chrysippus / orientis

Lots of sequencing “gaps” on one end...



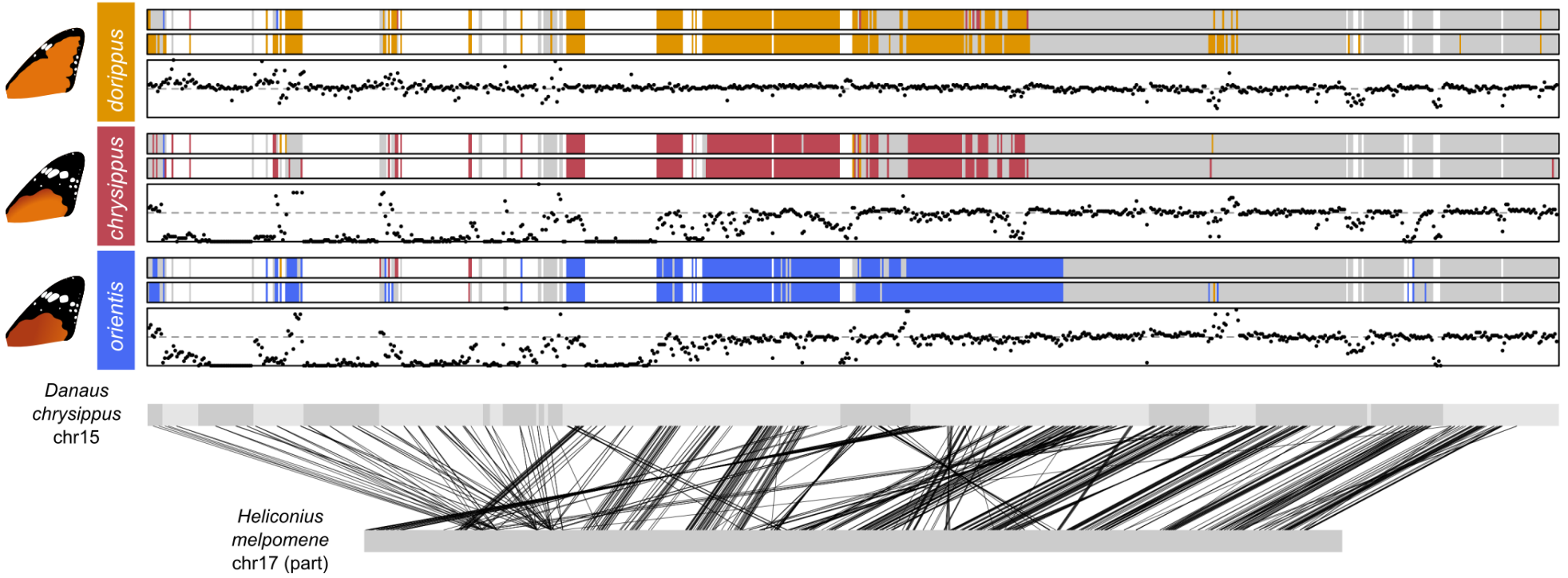
Martin et al. 2020 Plos Biol (in press)

Missing sequences in *chrysippus* and *orientis* alleles



Martin et al. 2020 Plos Biol (in press)

Tandem duplications specific to the ***dorippus*** allele



Martin et al. 2020 Plos Biol (in press)