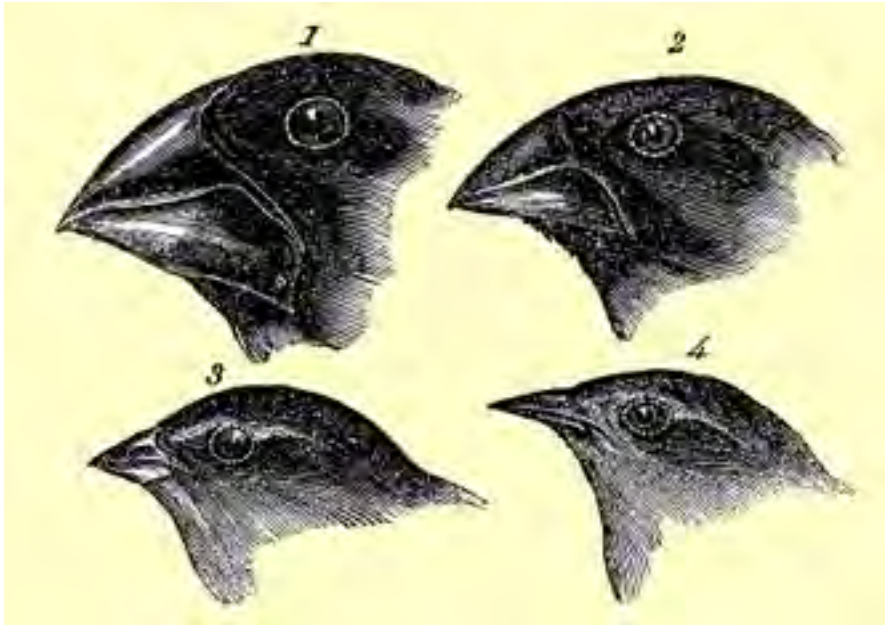
# Genomic studies of speciation and gene flow





# Why study speciation genomics?

# Why study speciation genomics?

Long-standing questions (role of geography/gene flow)

# Why study speciation genomics?

Long-standing questions (role of geography/gene flow)

How do genomes diverge?

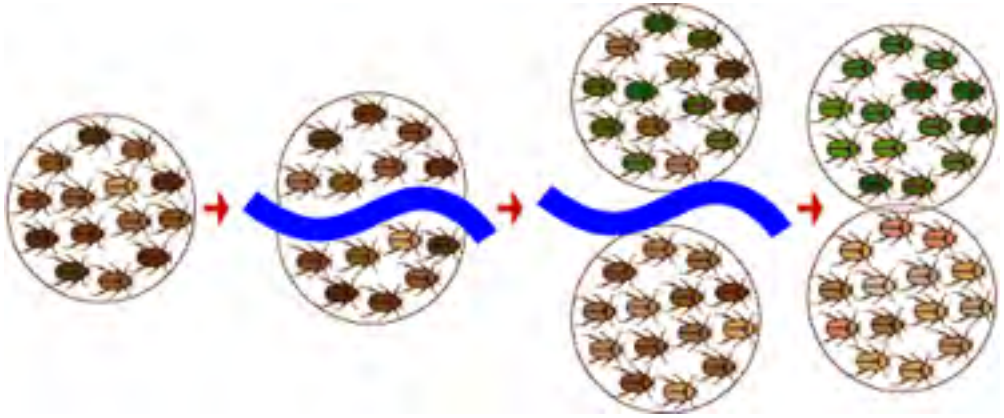
# Why study speciation genomics?

Long-standing questions (role of geography/gene flow)

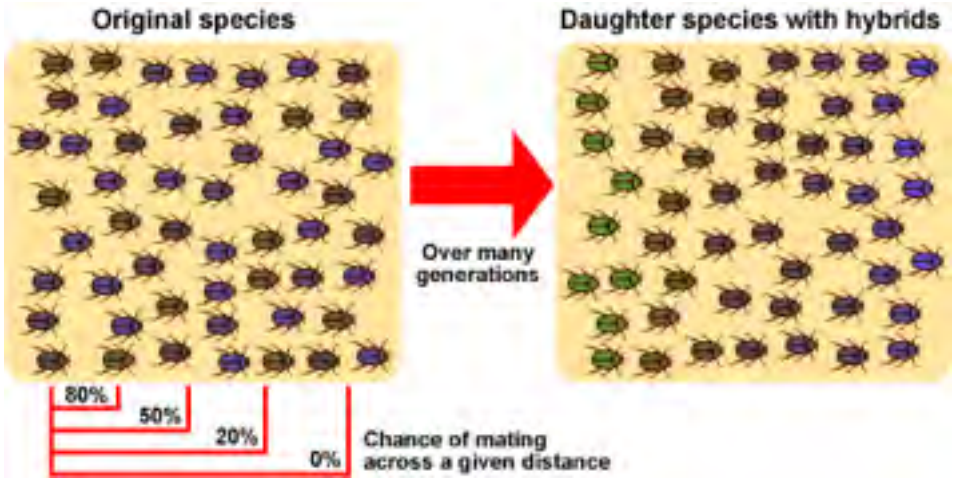
How do genomes diverge?

Find speciation genes

# Genomic divergence during speciation

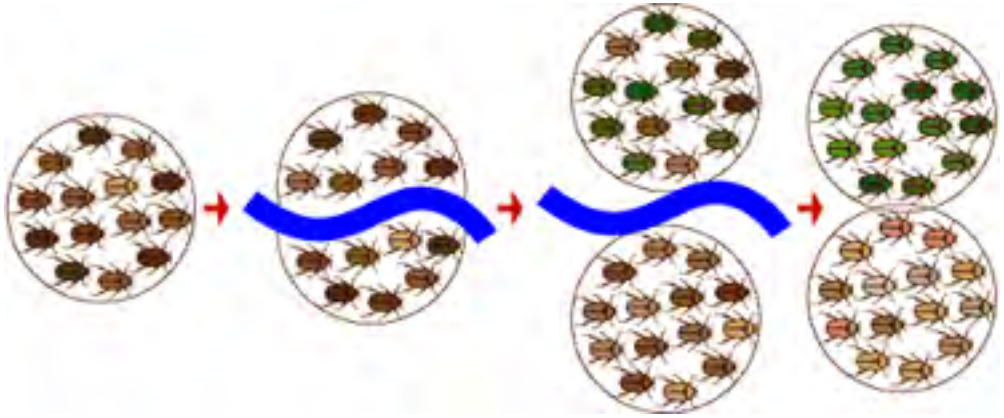


1. Speciation as a bi-product of physical isolation

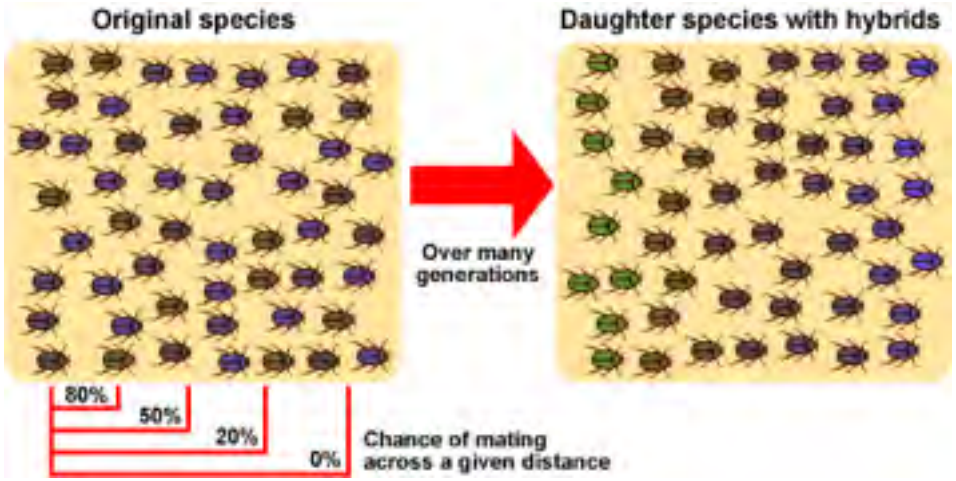



2. Speciation due to selection – without isolation

# Genomic divergence during speciation

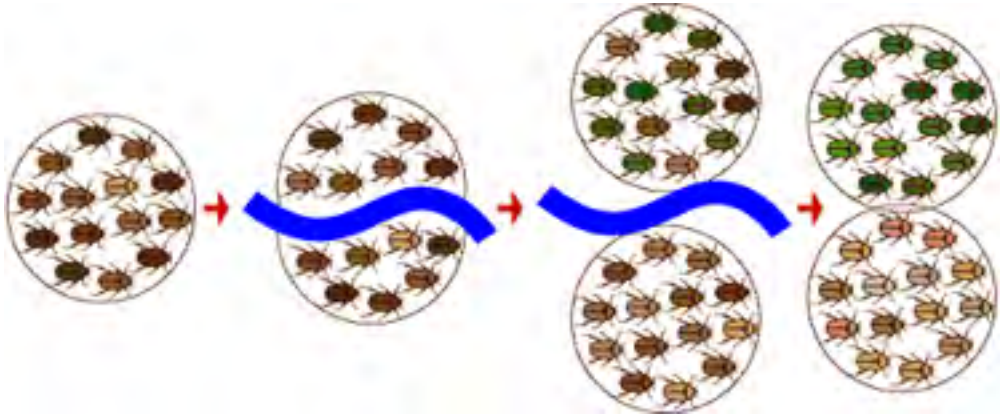


1. Speciation as a bi-product of physical isolation

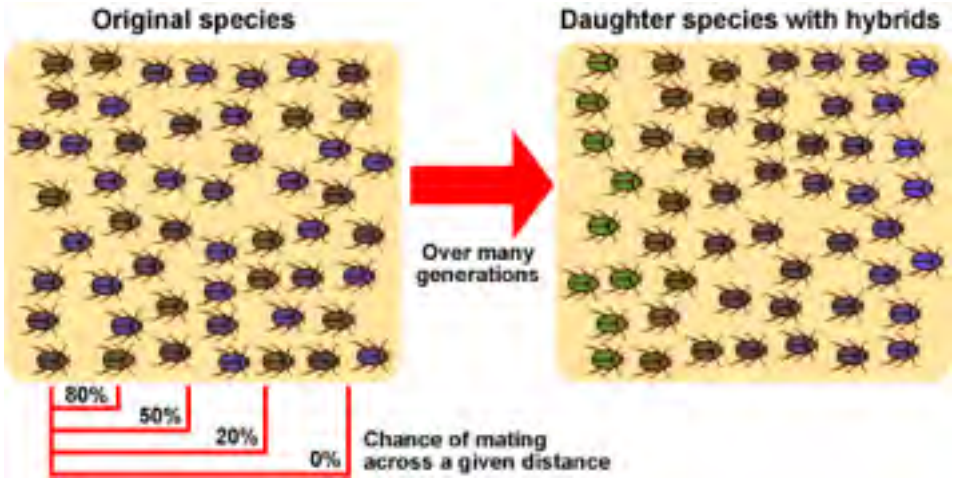


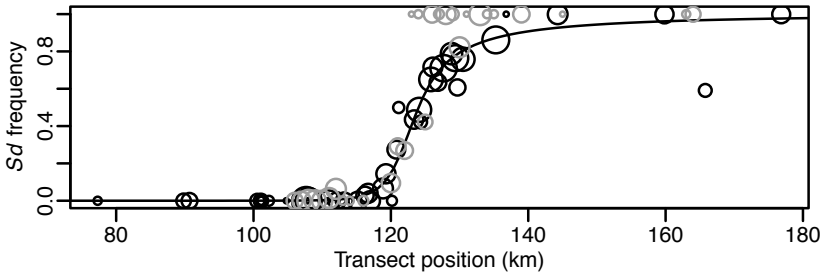

2. Speciation due to selection – without isolation

# Genomic divergence during speciation



1. Speciation as a bi-product of physical isolation

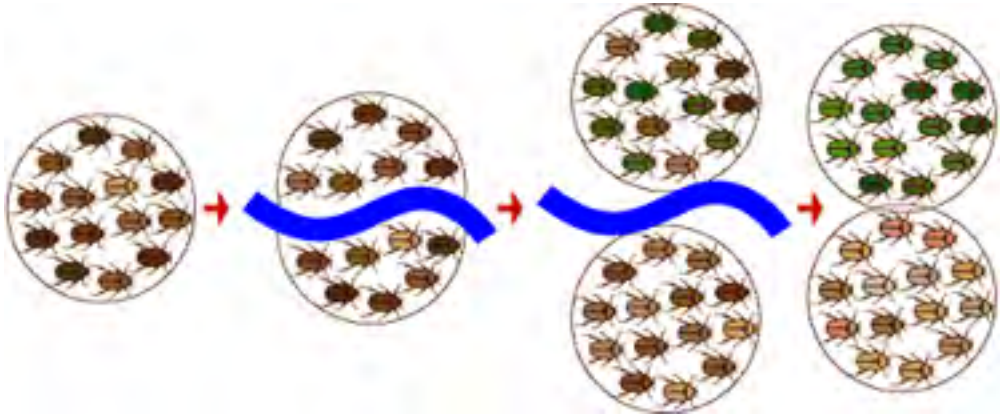


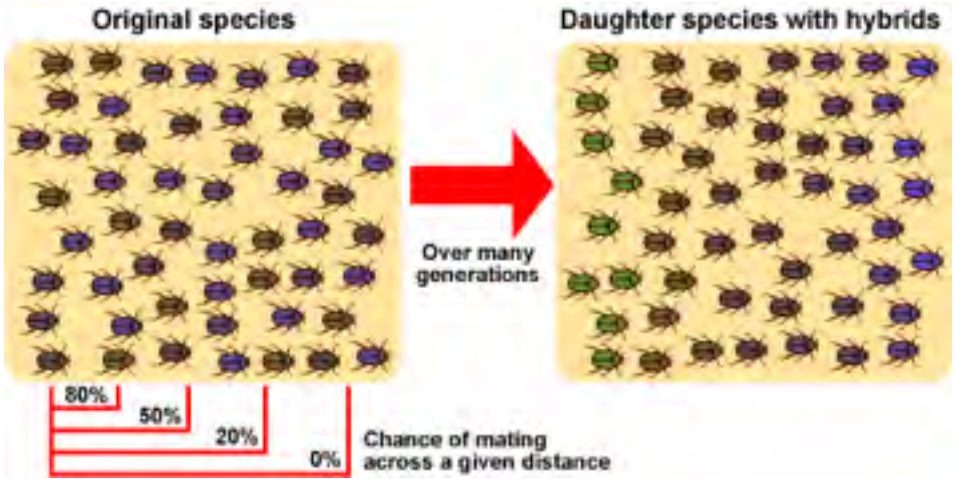

Cline theory - e.g. Barton and Gale 1993

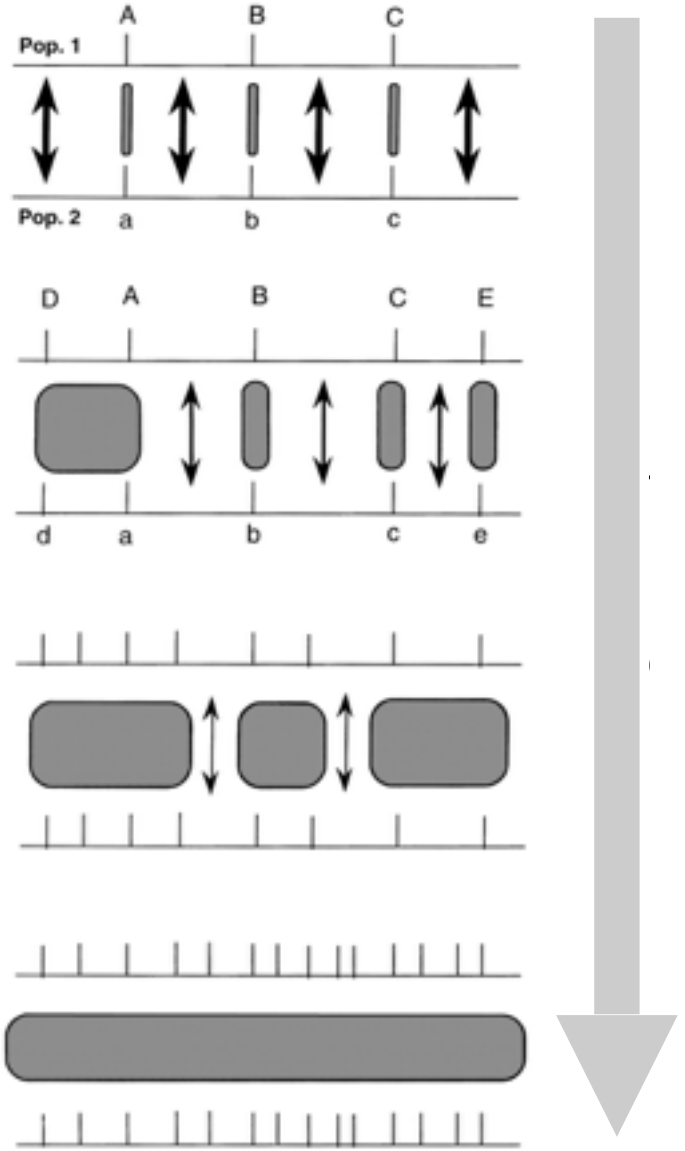
2. Speciation due to selection – without isolation

# Genomic divergence during speciation



1. Speciation as a bi-product of physical isolation



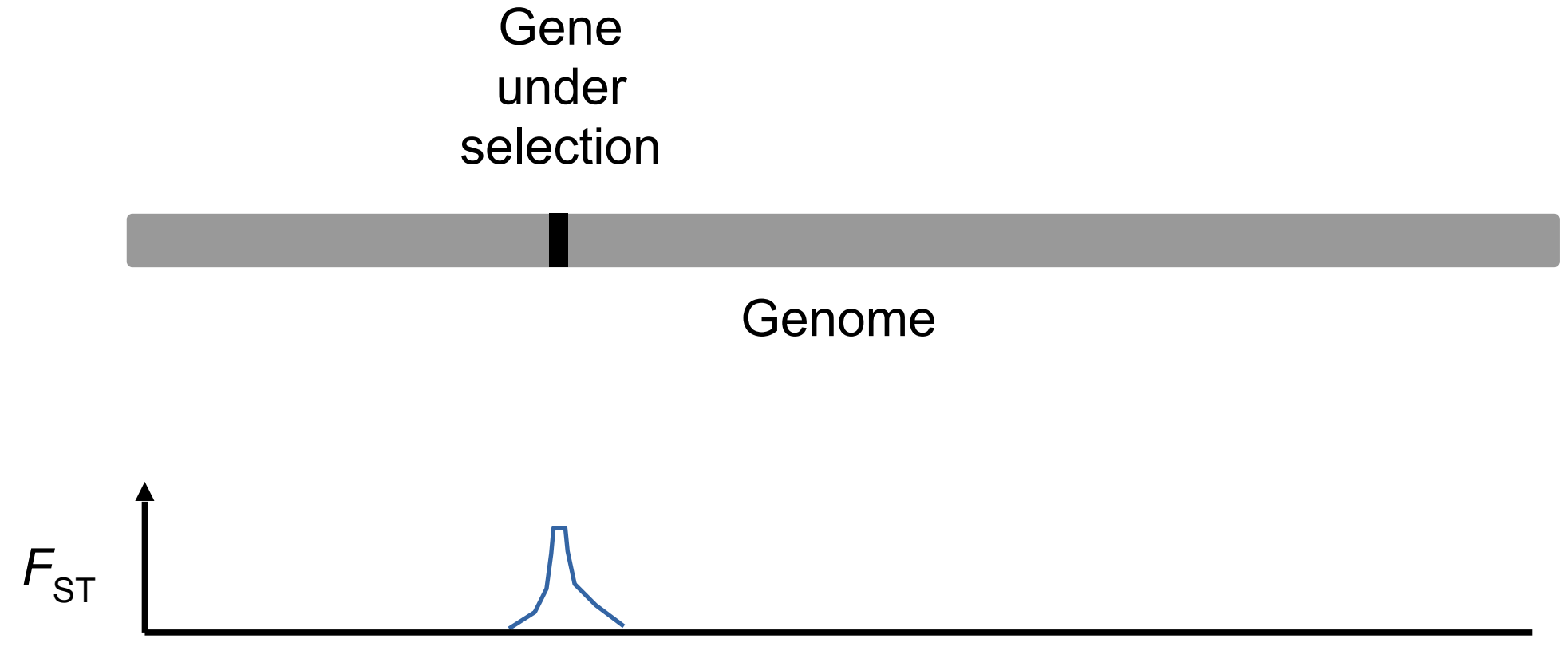




2. Speciation due to selection – without isolation



Wu 2001, JEB

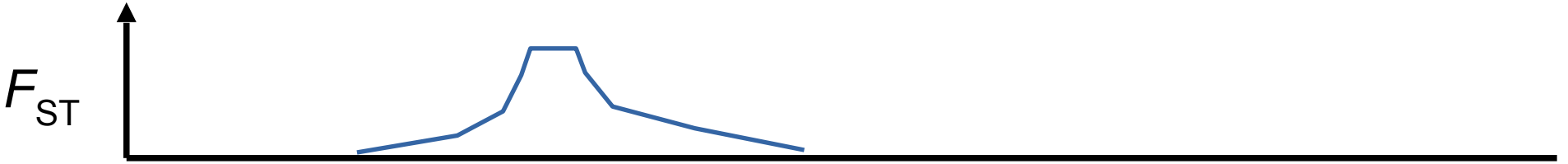
# Stage 1 - one or few loci under disruptive selection




Feder, Egan and Nosil TiG

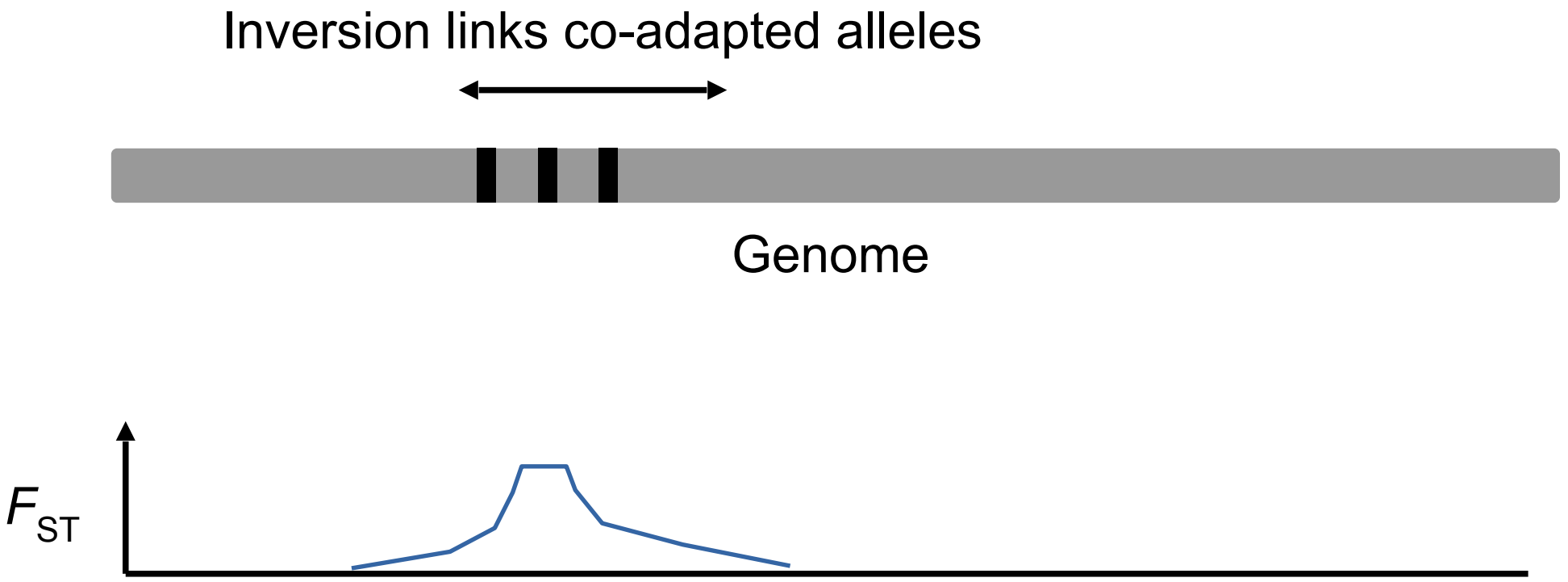
# Stage 2 - Divergence hitchhiking

Genome



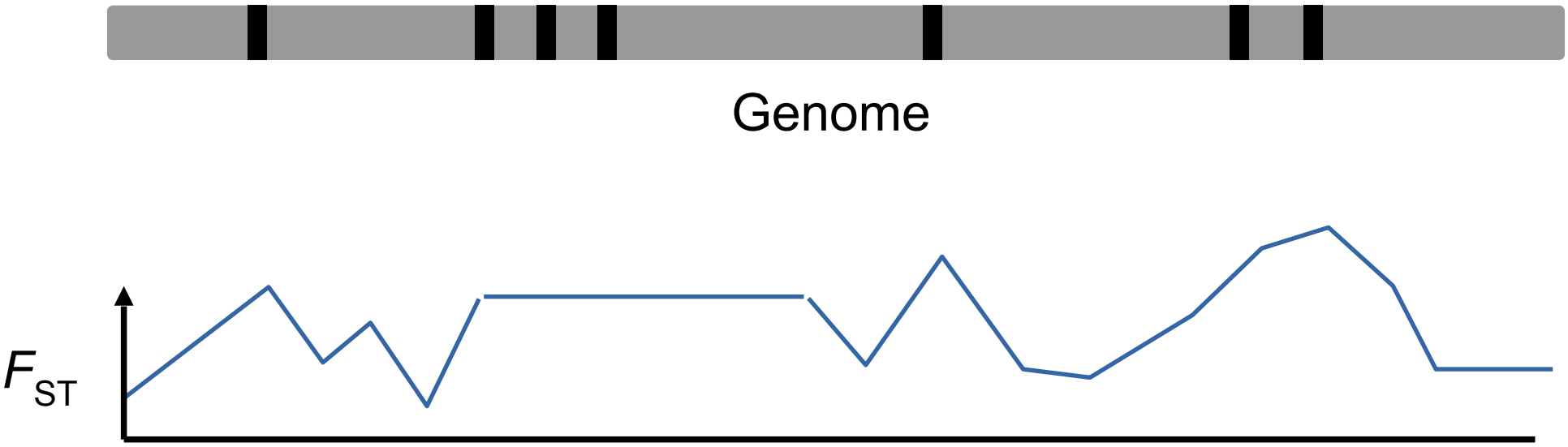

Feder, Egan and Nosil TiG

# Stage 2b - Inversion



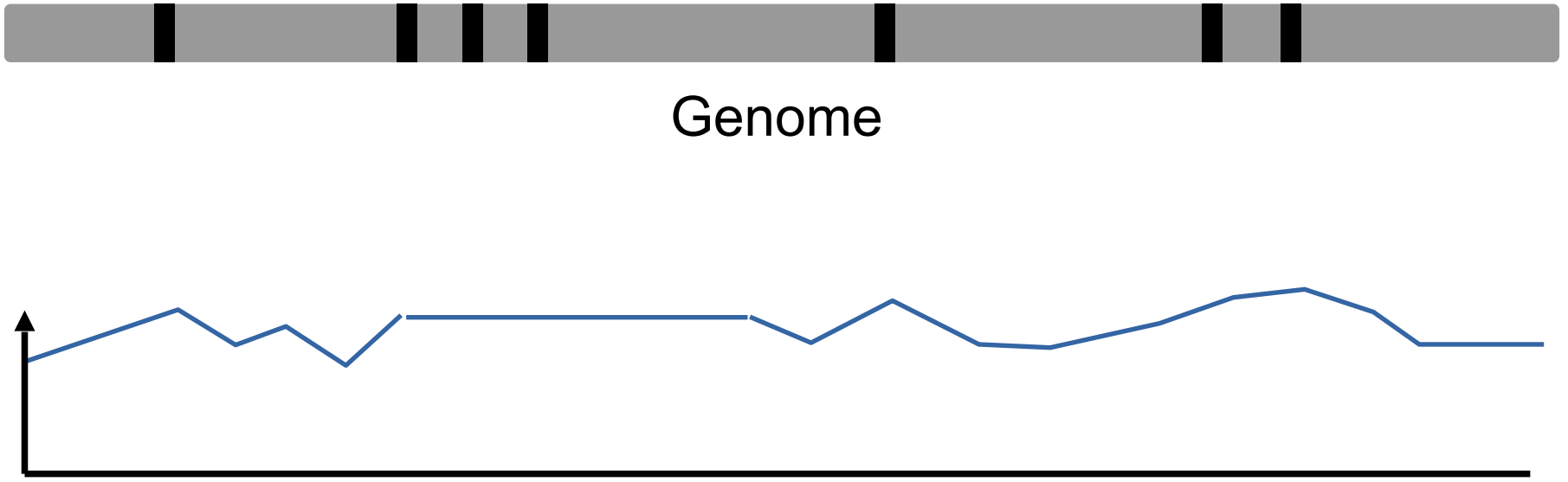

Feder, Egan and Nosil TiG

# Stage 3 - Genome hitchhiking




Feder, Egan and Nosil TiG

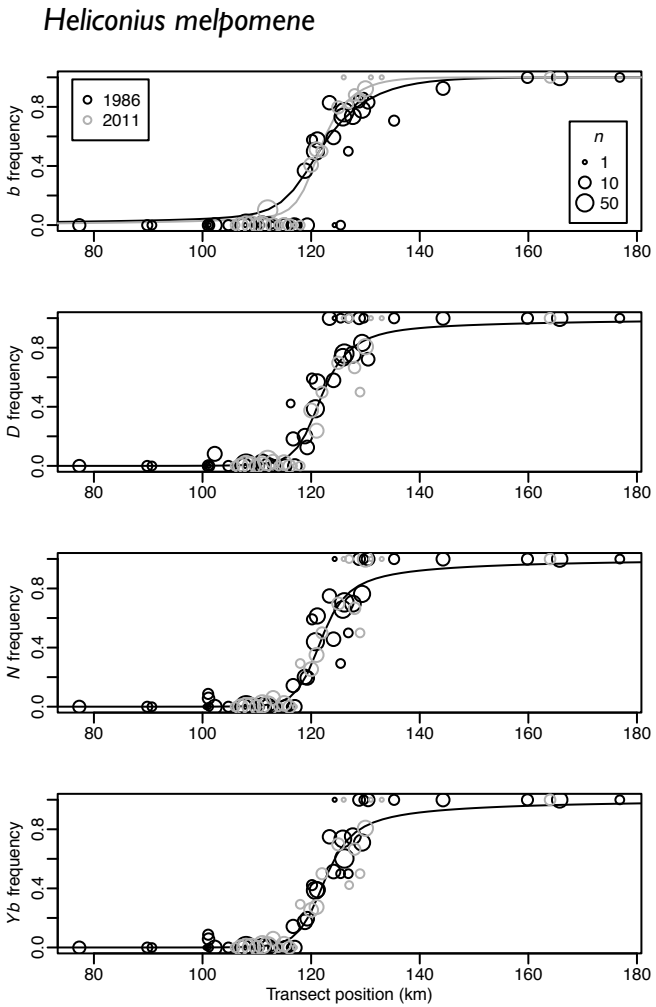
# Stage 4 - Genome wide isolation

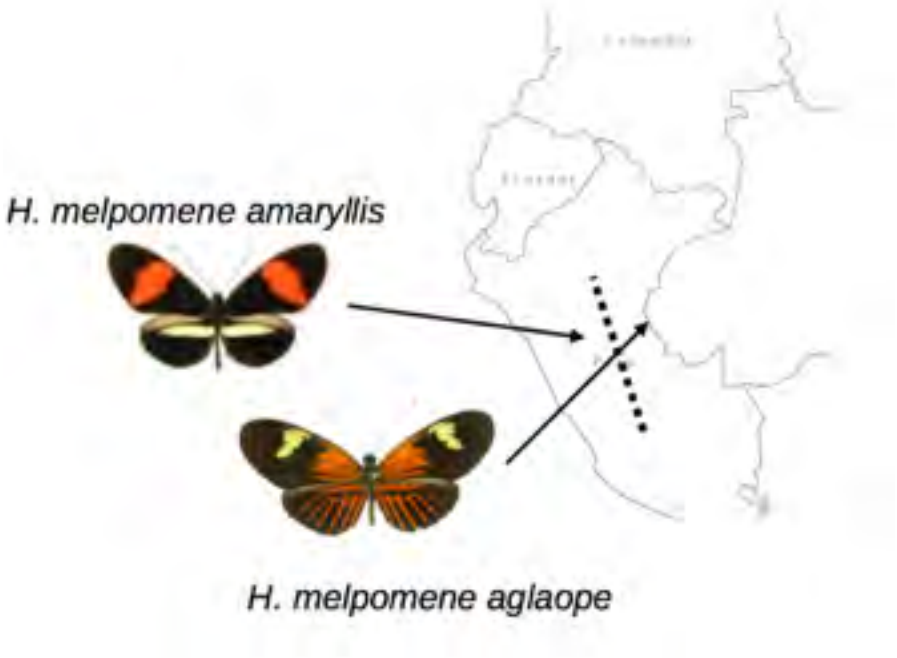



Feder, Egan and Nosil TiG

# Some sub-species clearly in stage 1

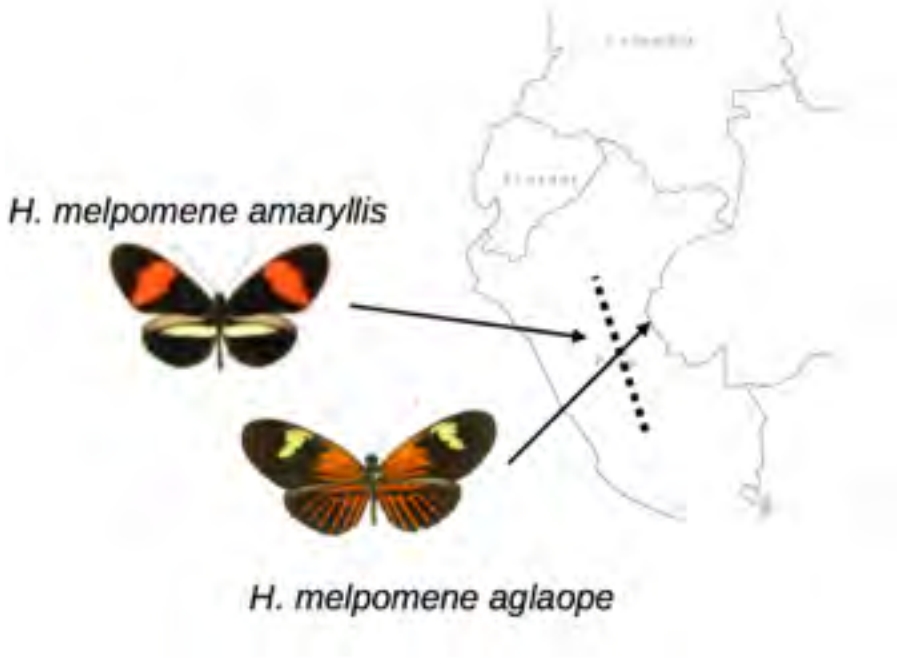
- Wing pattern “races” of *Heliconius melpomene*

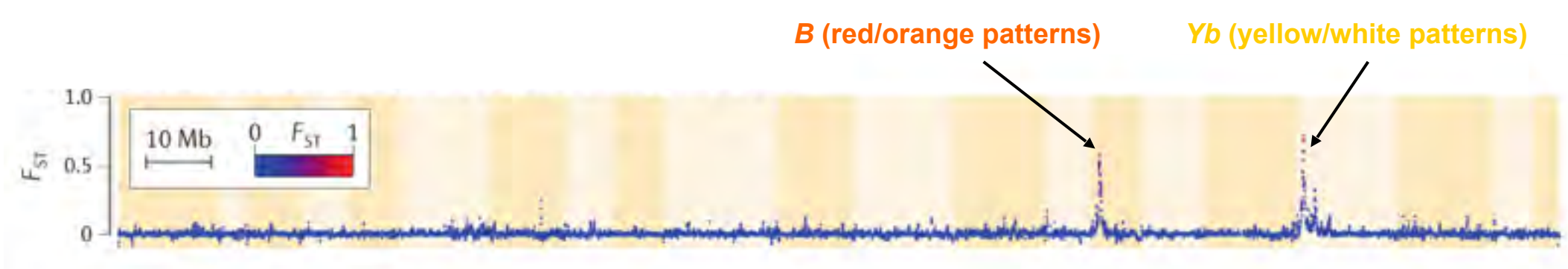


# Some sub-species clearly in stage 1

- Wing pattern “races” of *Heliconius melpomene*

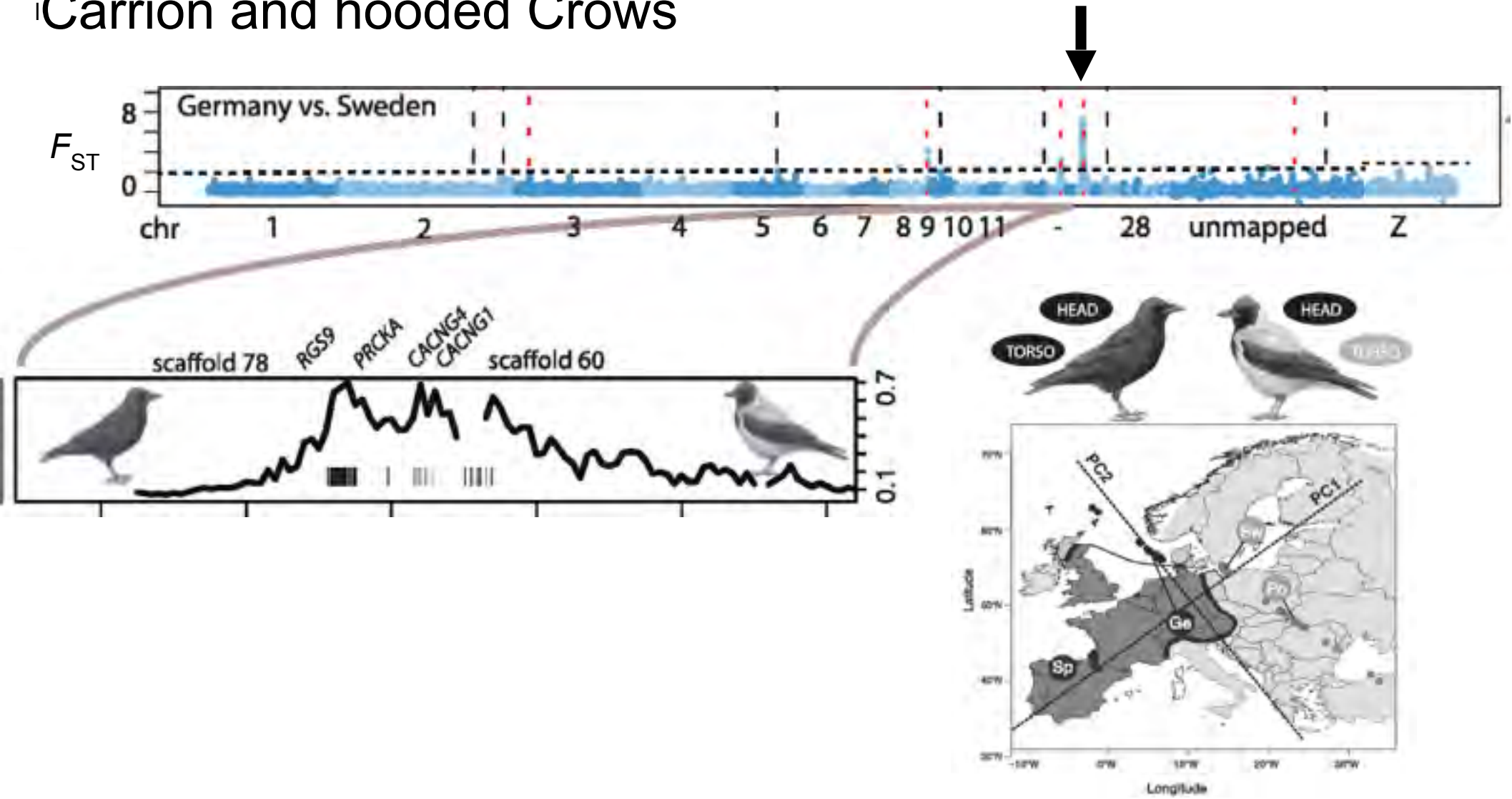



S. H. Martin et al. Genome Res. 23, 1817–1828 (2013).
O. Seehausen et al. Nat. Rev. Genet. 15, 176–92 (2014).

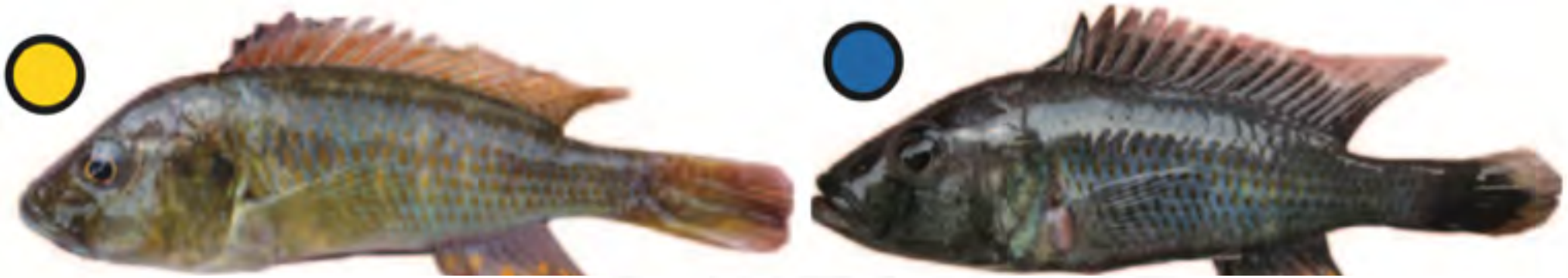
# Some sub-species clearly in stage 1

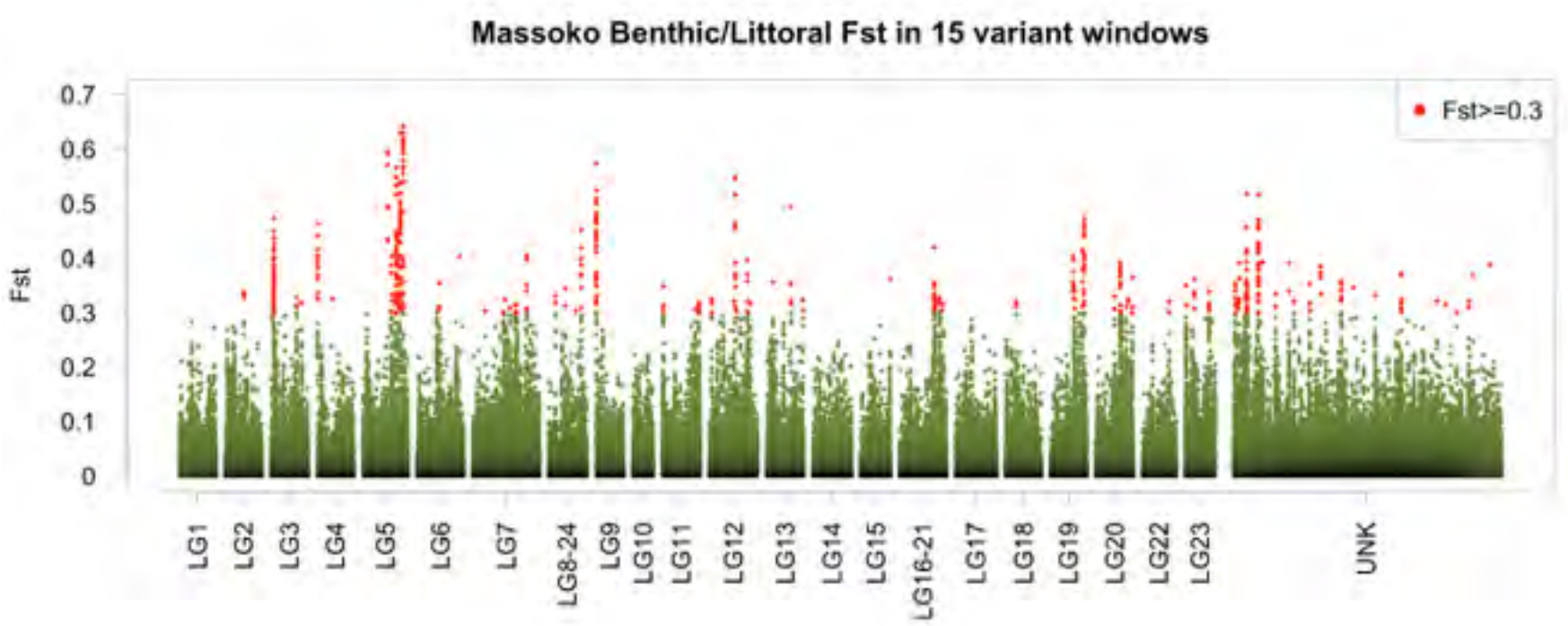
- Carrion and hooded Crows




Poelstra, J. W. et al. Science 344, 1410–4 (2014).

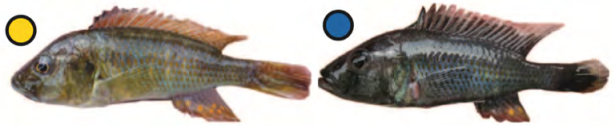
# And here is a recent example with multiple islands

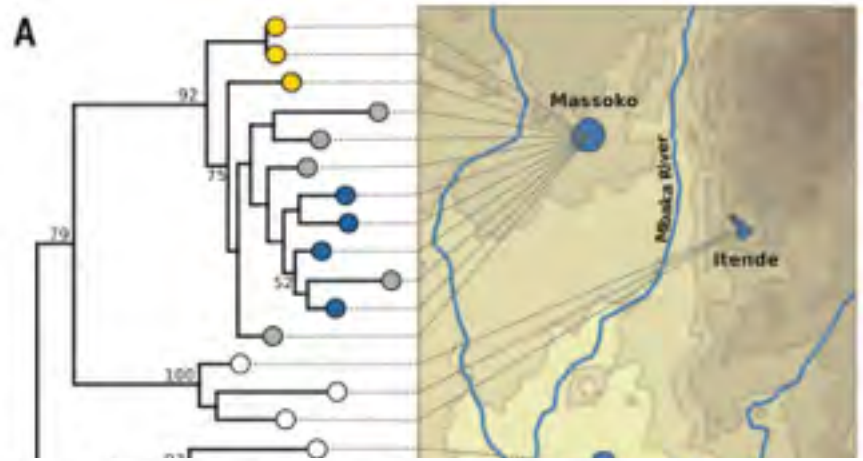


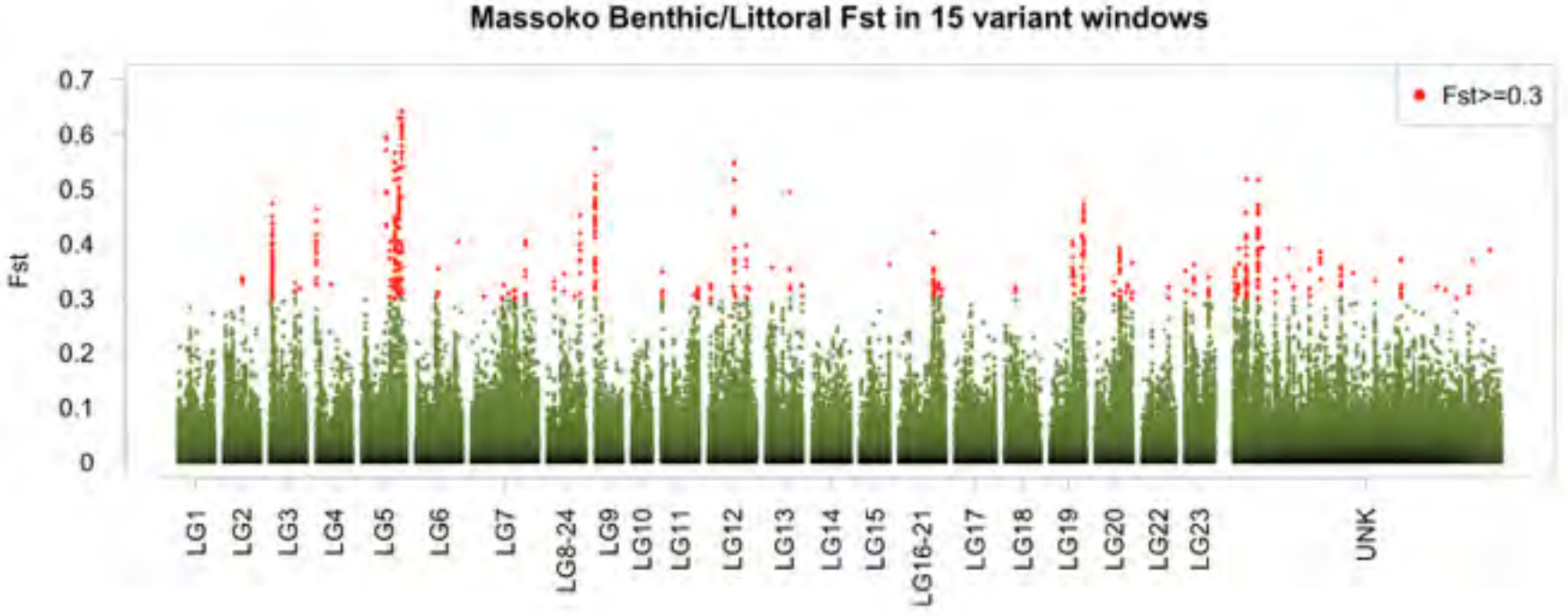



Malinsky et al., Science 350, 1493 (2015).

# And here is a recent example with multiple islands

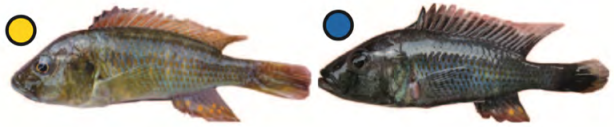


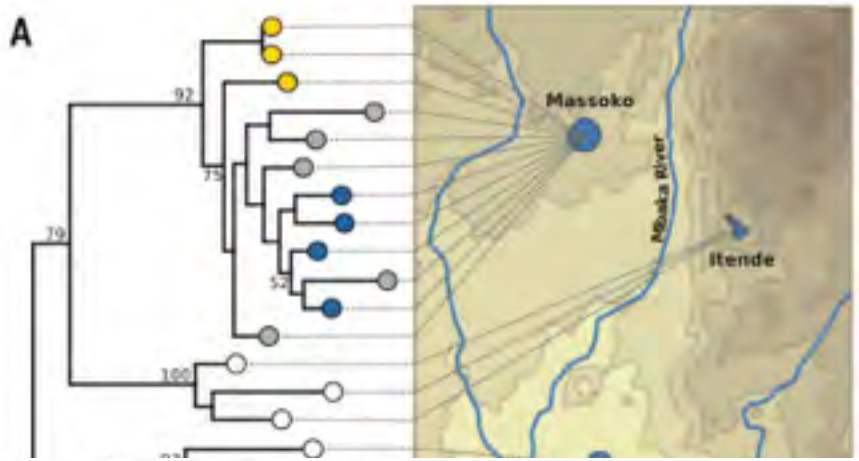


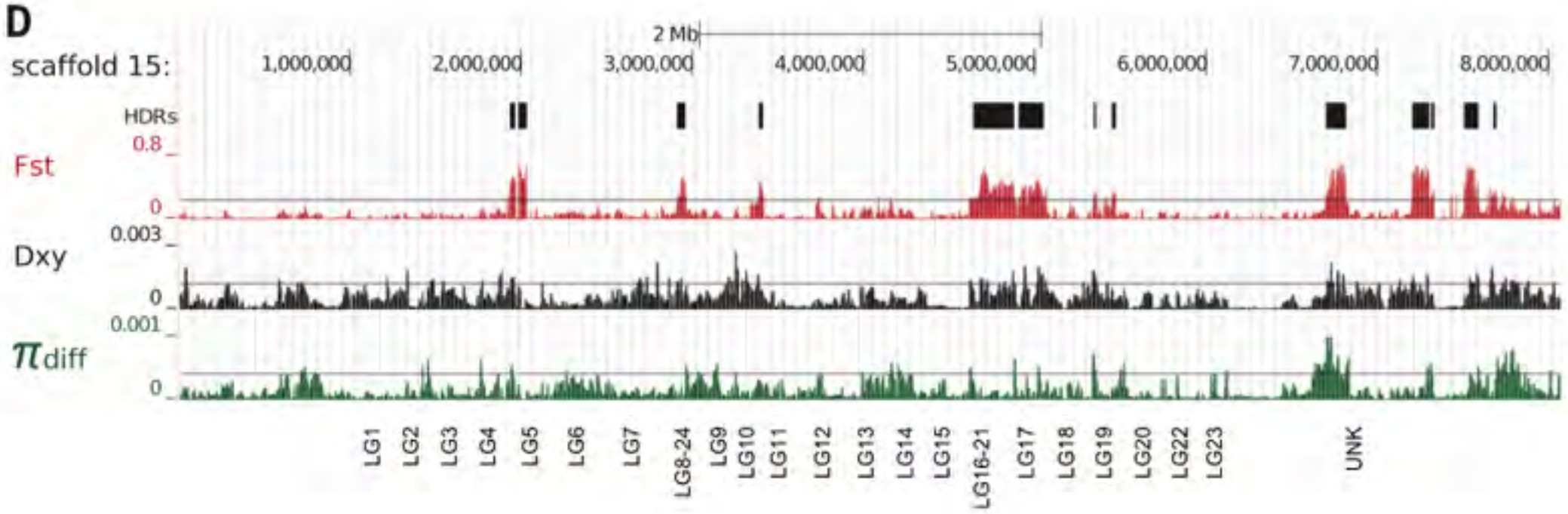




Malinsky et al., Science 350, 1493 (2015).

# And here is a recent example with multiple islands

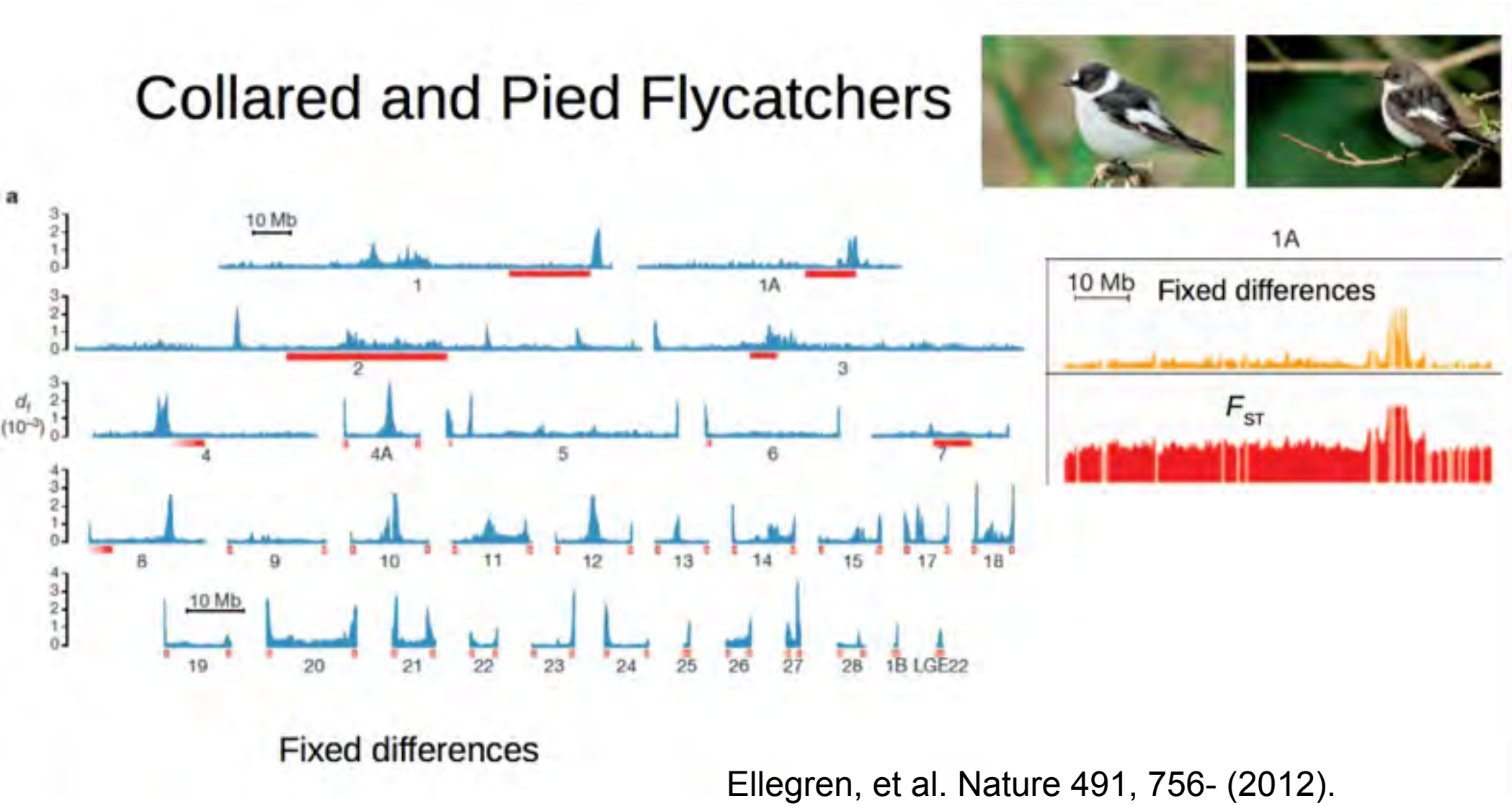





Malinsky et al., Science 350, 1493 (2015).

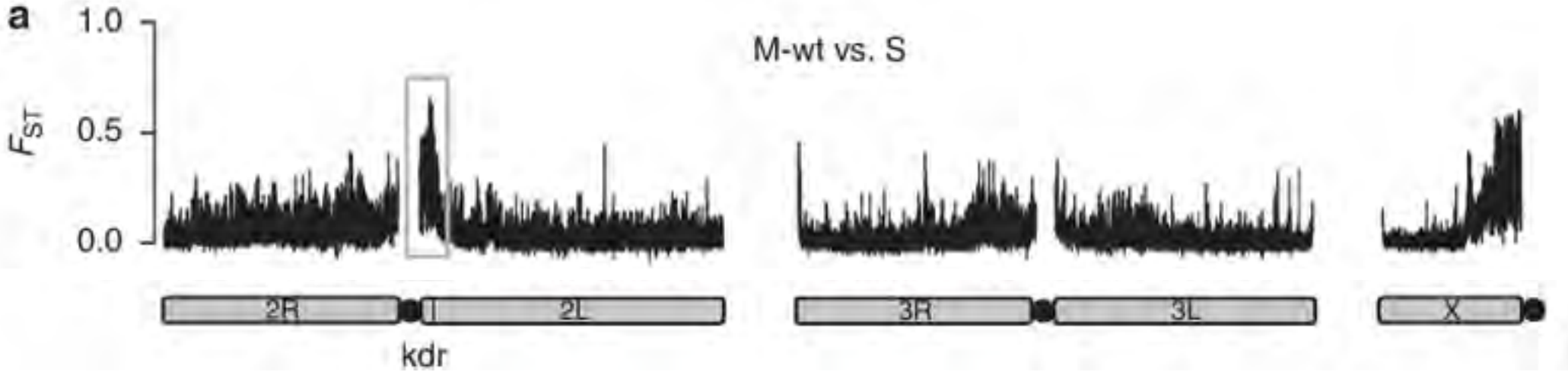
# Other species have islands…but are they real?




Ellegren, et al. Nature 491, 756- (2012).

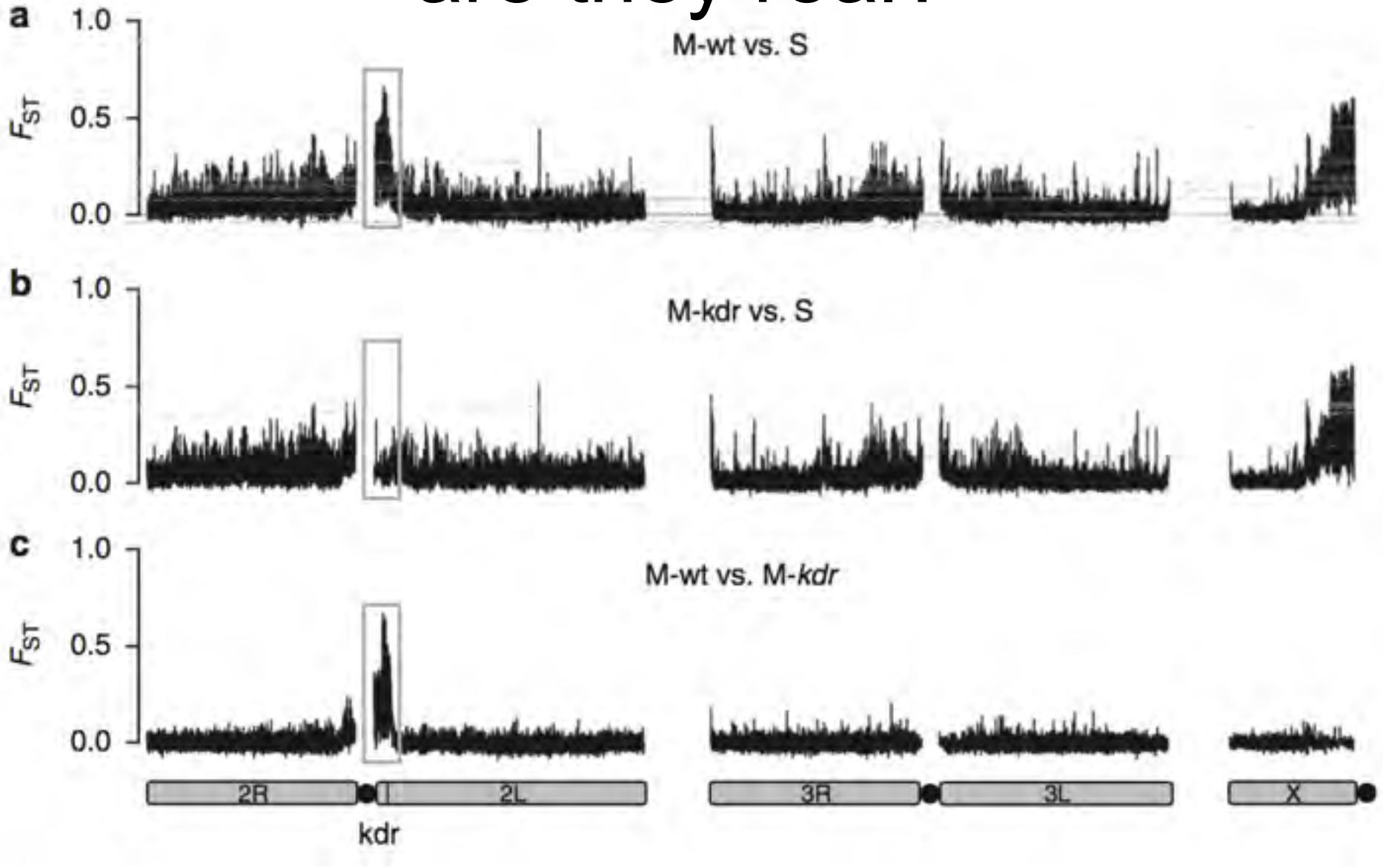
# Other species have islands…but are they real?

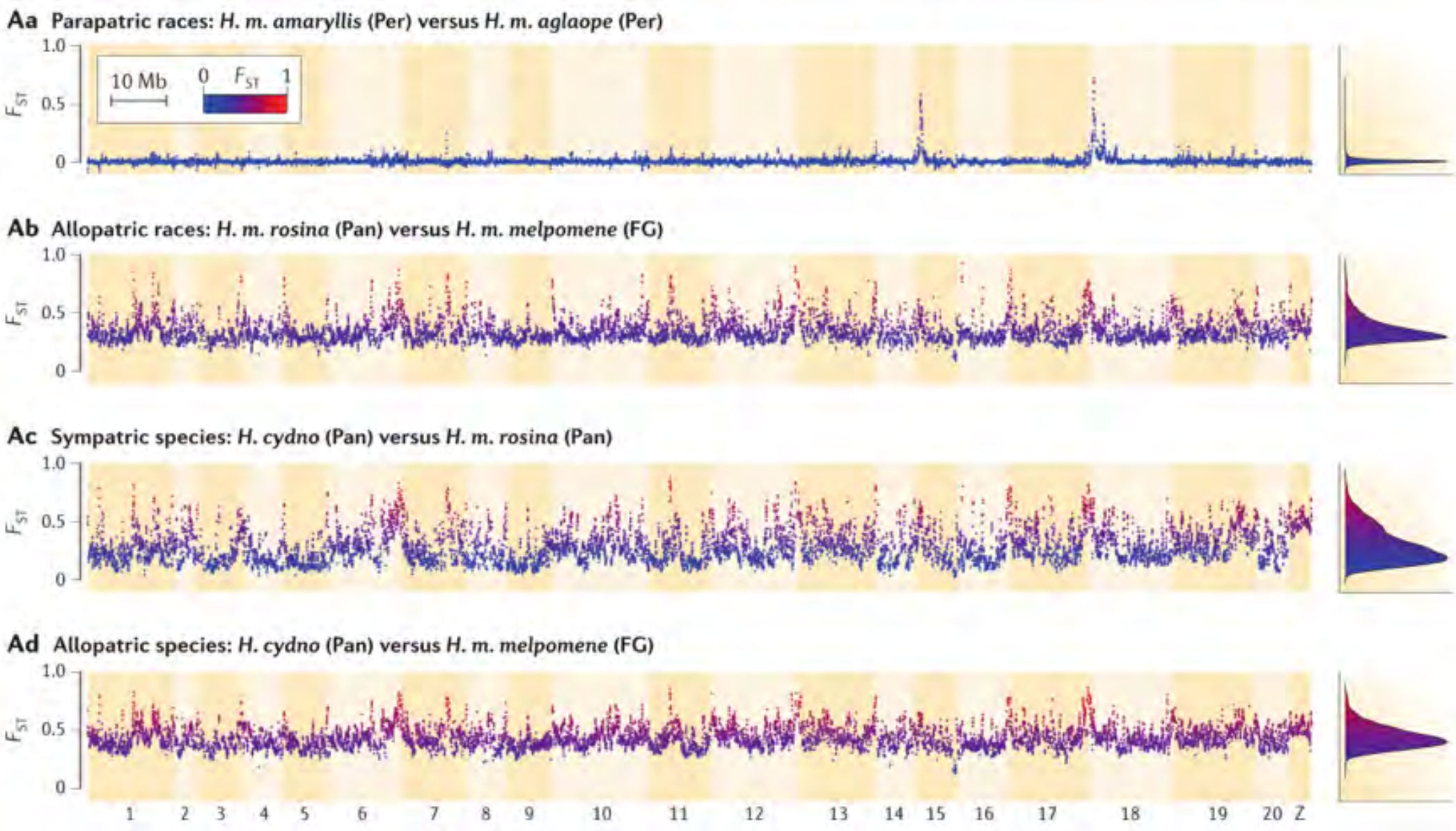
*Anopheles gambiae* and *A. coluzzi*
*Formerly M and S forms of A. gambiae*




Clarkson et al. 2014 Nature Communications

# Other species have islands…but are they real?






Seehausen et al., Nature Reviews Genetics, 2014

# What do patterns of $F_{st}$ really mean?

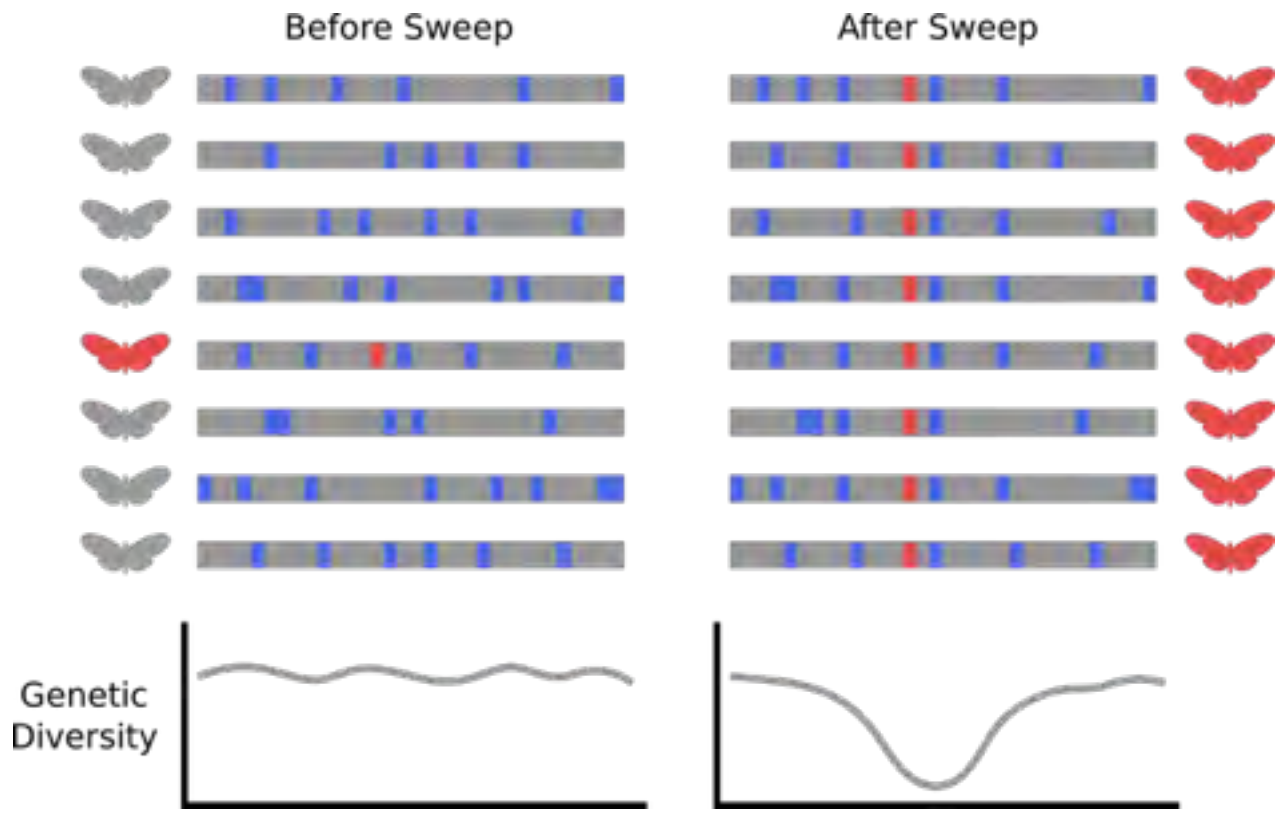
- $F_{st}$ measures relative divergence

$$F_{ST} = \frac{H_T - H_S}{H_T},$$

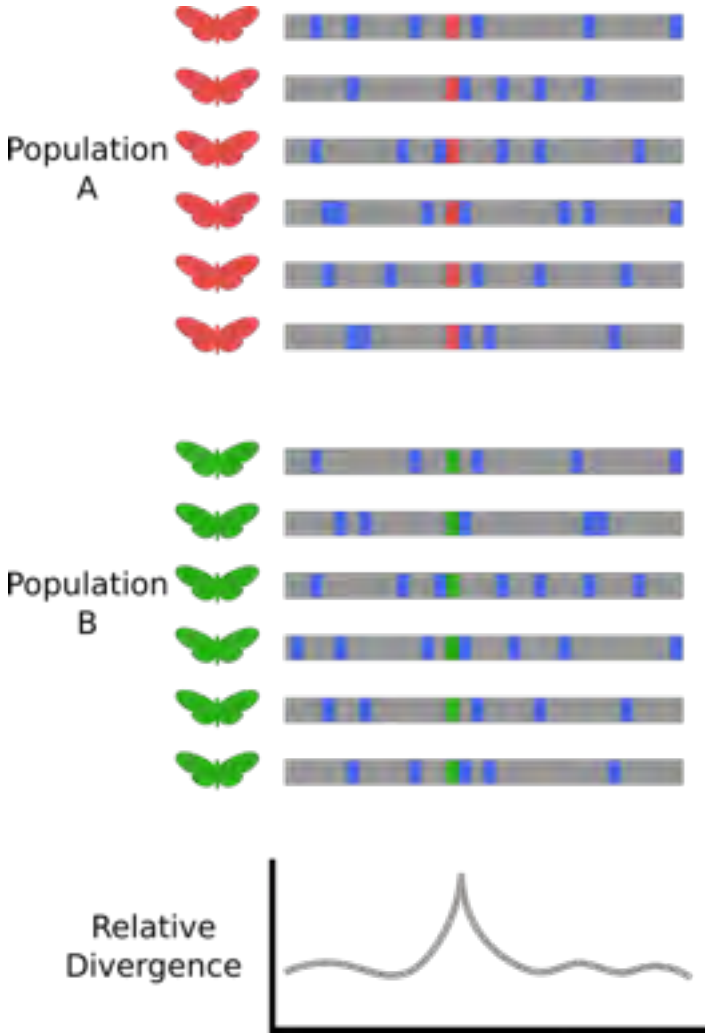
- Peaks indicate regions of higher than expected between population divergence, given the within population divergence

- Peaks can therefore result from reduced diversity within species

- This could be due to lower Ne within species (selective sweeps, background selection)
- So peaks NOT NECESSARILY due to reduced gene flow




Note that sometimes sweeps within species = speciation genes




Sweeps across the species barrier can also lead to Fst peaks

Double peaks??

Nicolas Bierne, Daniel Berner and others

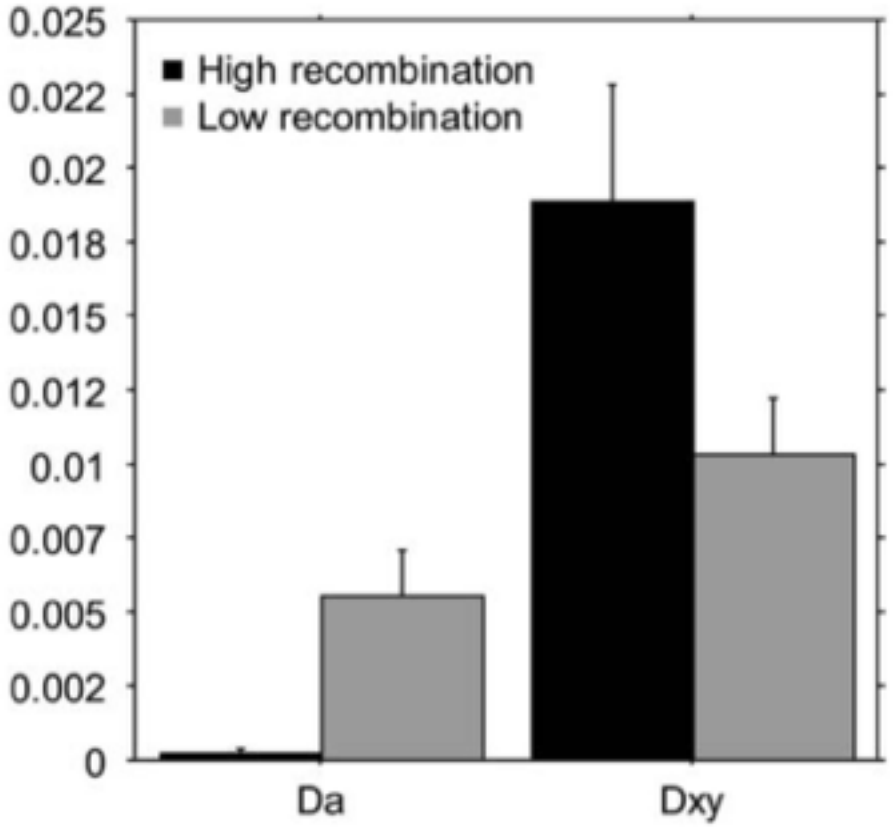
Heredity (2009) 103, 439–444

www.nature.com/hdy

REVIEW

# Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species

MAF Noor and SM Bennett
*Biology Department, Duke University, Durham, NC, USA*




Anopheles M-S divergence

Relative divergence higher in low recombination regions - not significant for absolute divergence

see also: Charlesworth 1998 MBE Measures of divergence…

MOLECULAR ECOLOGY

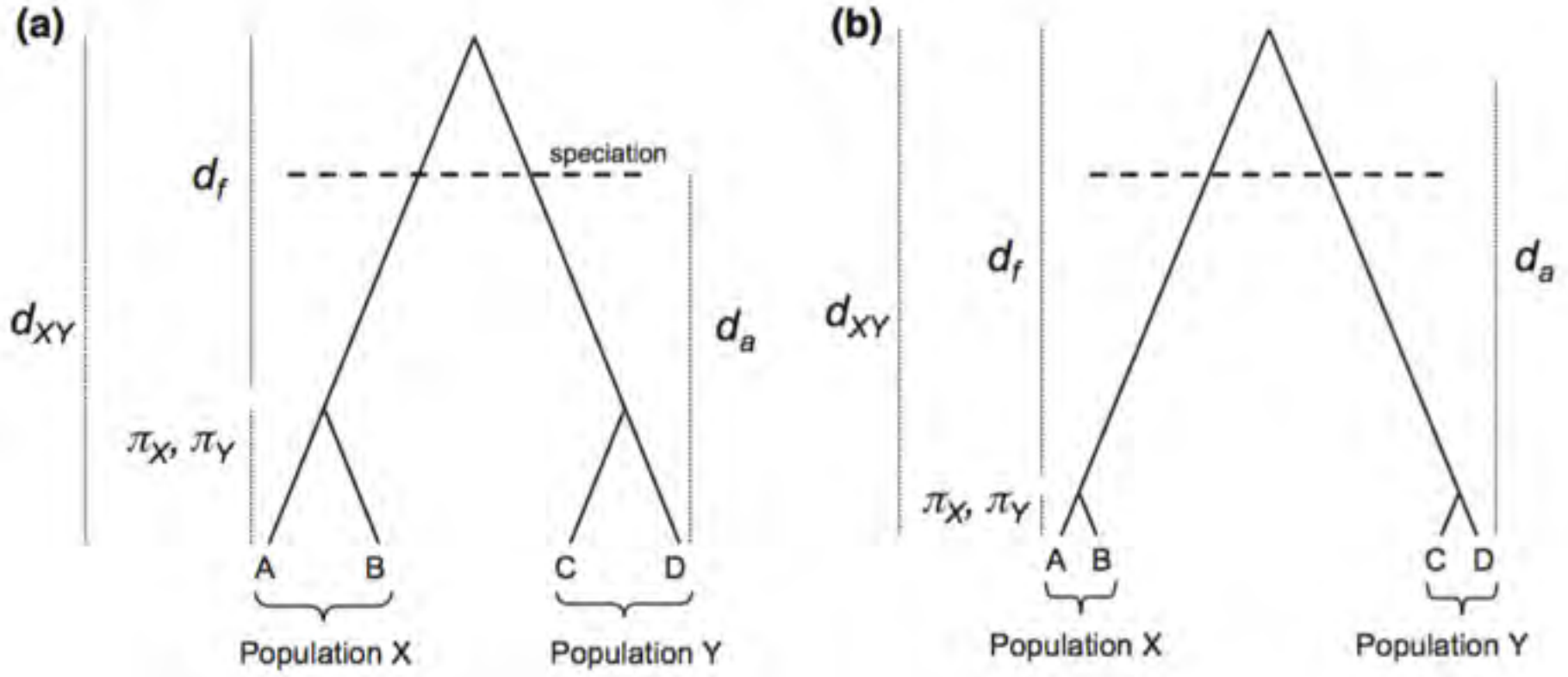
Molecular Ecology (2014) 23, 3133–3157 doi: 10.1111/mec.12796

INVITED REVIEWS AND SYNTHESES

# Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow

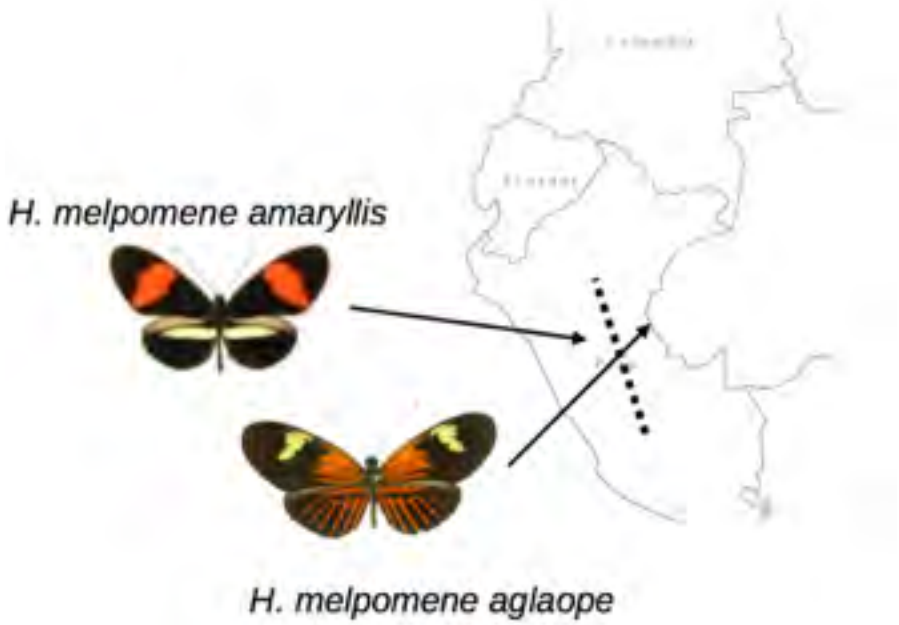
TAMI E. CRUICKSHANK* and MATTHEW W. HAHN*†

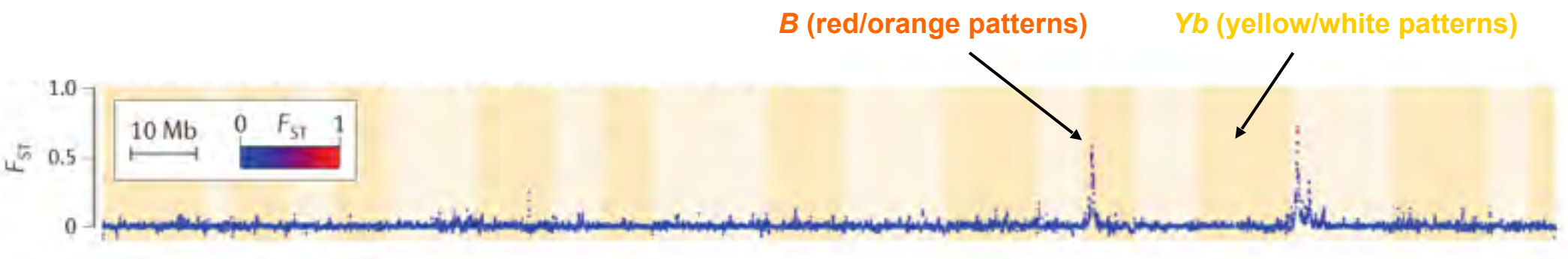
*Department of Biology, Indiana University, Bloomington, IN 47405, USA, †School of Informatics and Computing, Indiana University, Bloomington, IN 47405, USA



# No evidence for higher Dxy in wing pattern loci

- Wing pattern “races” of *Heliconius melpomene*



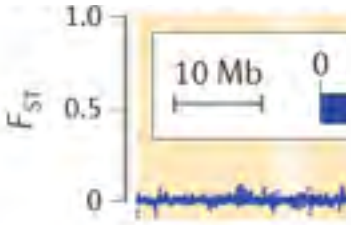




S. H. Martin et al. Genome Res. 23, 1817–1828 (2013).
O. Seehausen et al. Nat. Rev. Genet. 15, 176–92 (2014).

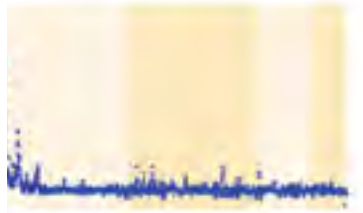
# No evidence for higher Dxy in wing pattern loci

- Wing pat
*Heliconiu*

One further issue with interpreting the data from these two races is whether this comparison relates to speciation at all. There is strong geographic structure involving the wing colour patterns that define these morphs as races, largely due to selection determined by colour morphs in the Müllerian mimic, *H. erato* (Mallet *et al.* 1990). But the races are not separate species: they do not show evidence of hybrid sterility or inviability and appear to be randomly mating in the narrow zone where the colour morphs overlap (Mallet *et al.* 1990). This raises the possibility that the colour-patterning loci contain locally adapted alleles within a largely panmictic (or at least continuously distributed) population and that gene flow outside of these regions represents nothing more than the normal movement of alleles within a species. In this case, there should be

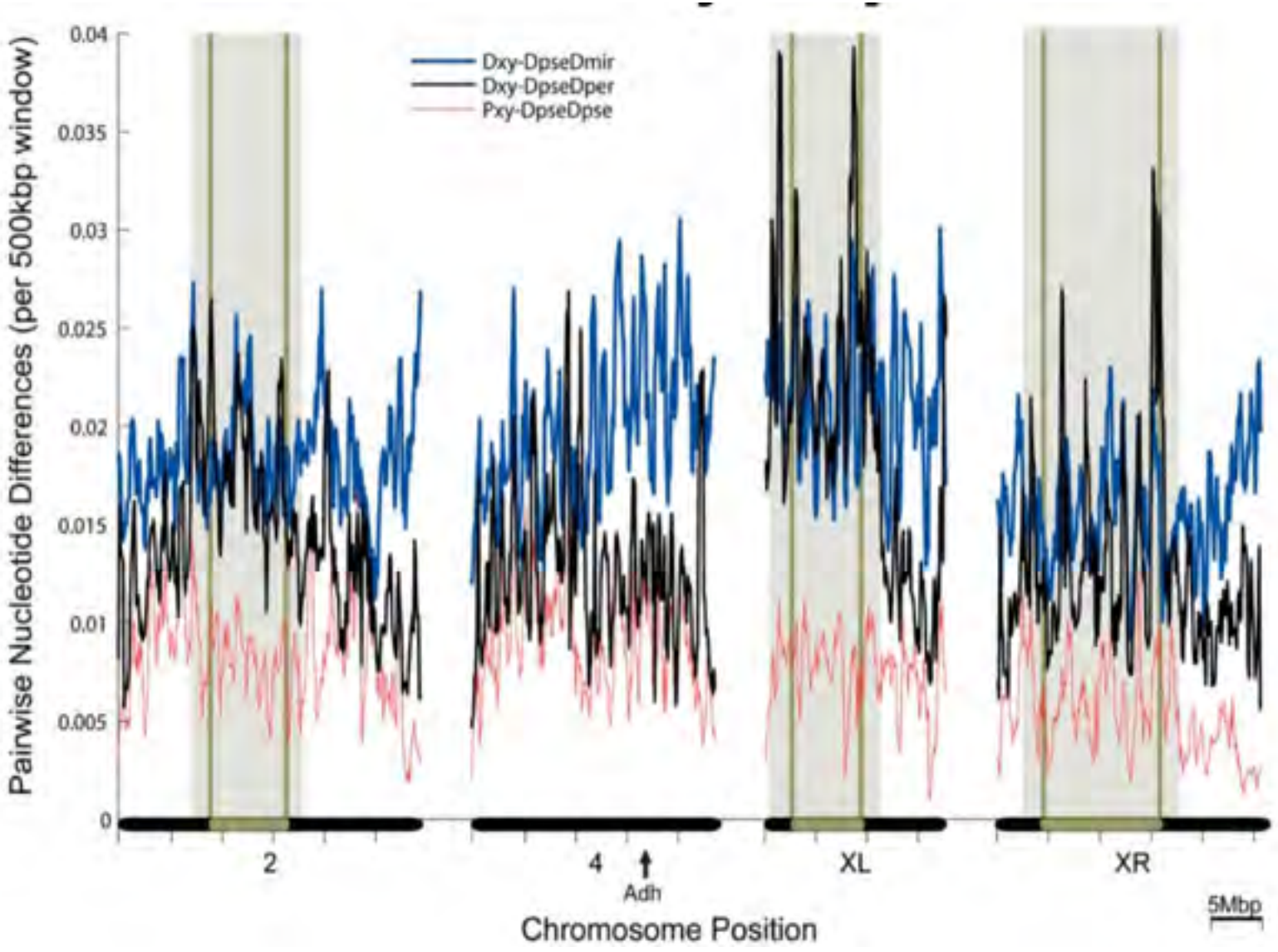



v/white patterns)



8 (2013).
-92 (2014).

# Suggestion that we use absolute measures of divergence?



# Understanding genomic divergence

No single statistic will capture the complex history of mutation, migration and selection

Patterns need to be interpreted in the specific context of the study species

# Much better to use explicit tests for gene flow

Need to design sampling so the expectations in the absence of gene flow are clear and testable

# Much better to use explicit tests for gene flow

Need to design sampling so the expectations in the absence of gene flow are clear and testable

The key is to identify ‘control’ populations that are not influenced by admixture

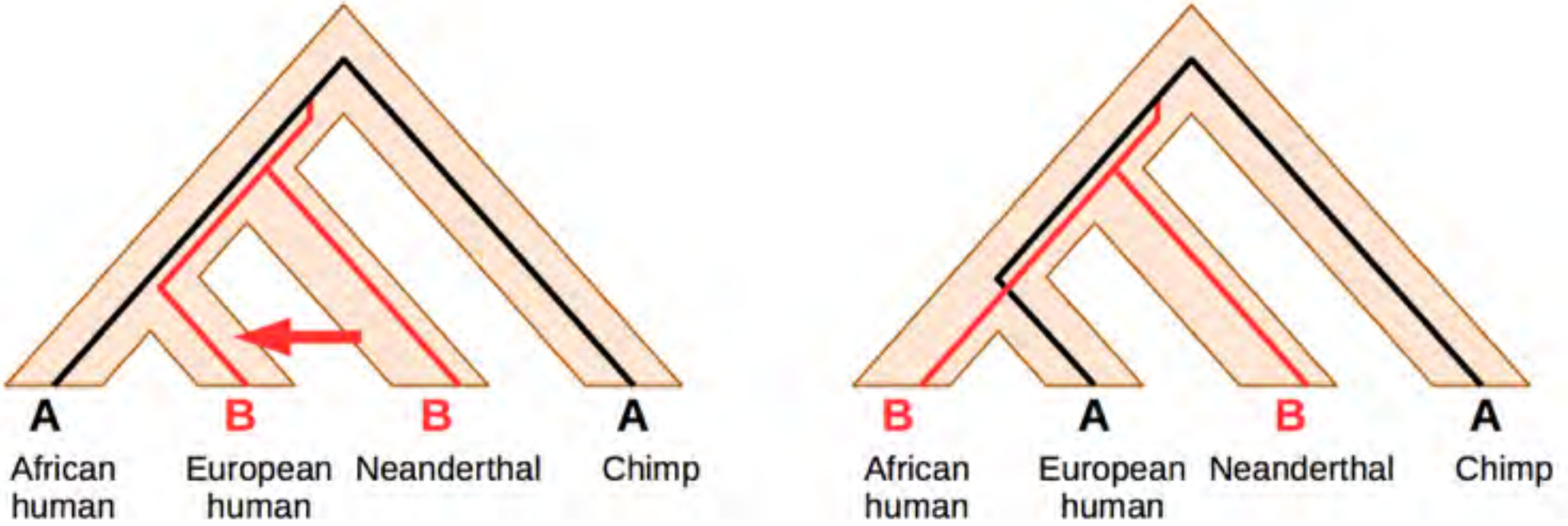
# Explicit tests for gene flow: Neanderthal genome

- Isolated DNA from bones 38,000 yrs old in Croatia
- We diverged from Neanderthals around 270-440,000yrs ago

- Evidence for gene exchange with humans (1-4% of genome?)

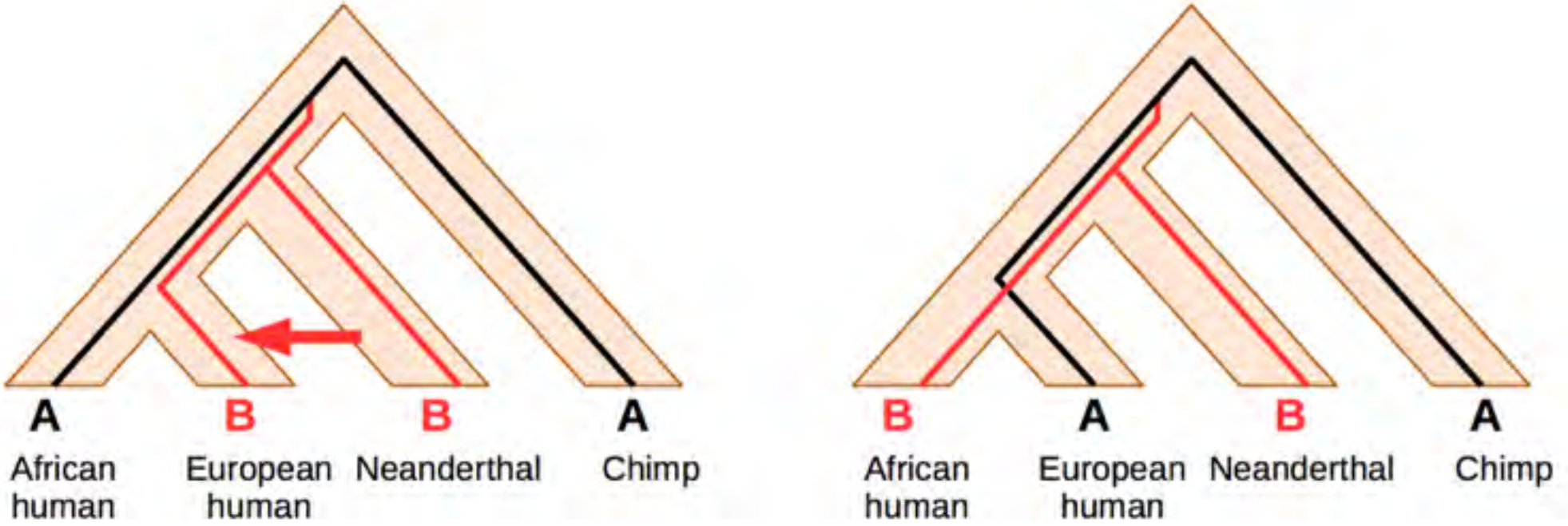
Green et al., 328:710 Science 2010

# Explicit tests for gene flow: ABBA-BABA test




$$D(P_1, P_2, P_3, O) = \frac{\sum C_{ABBA}(i) - C_{BABA}(i)}{\sum C_{ABBA}(i) + C_{BABA}(i)} \qquad (1)$$

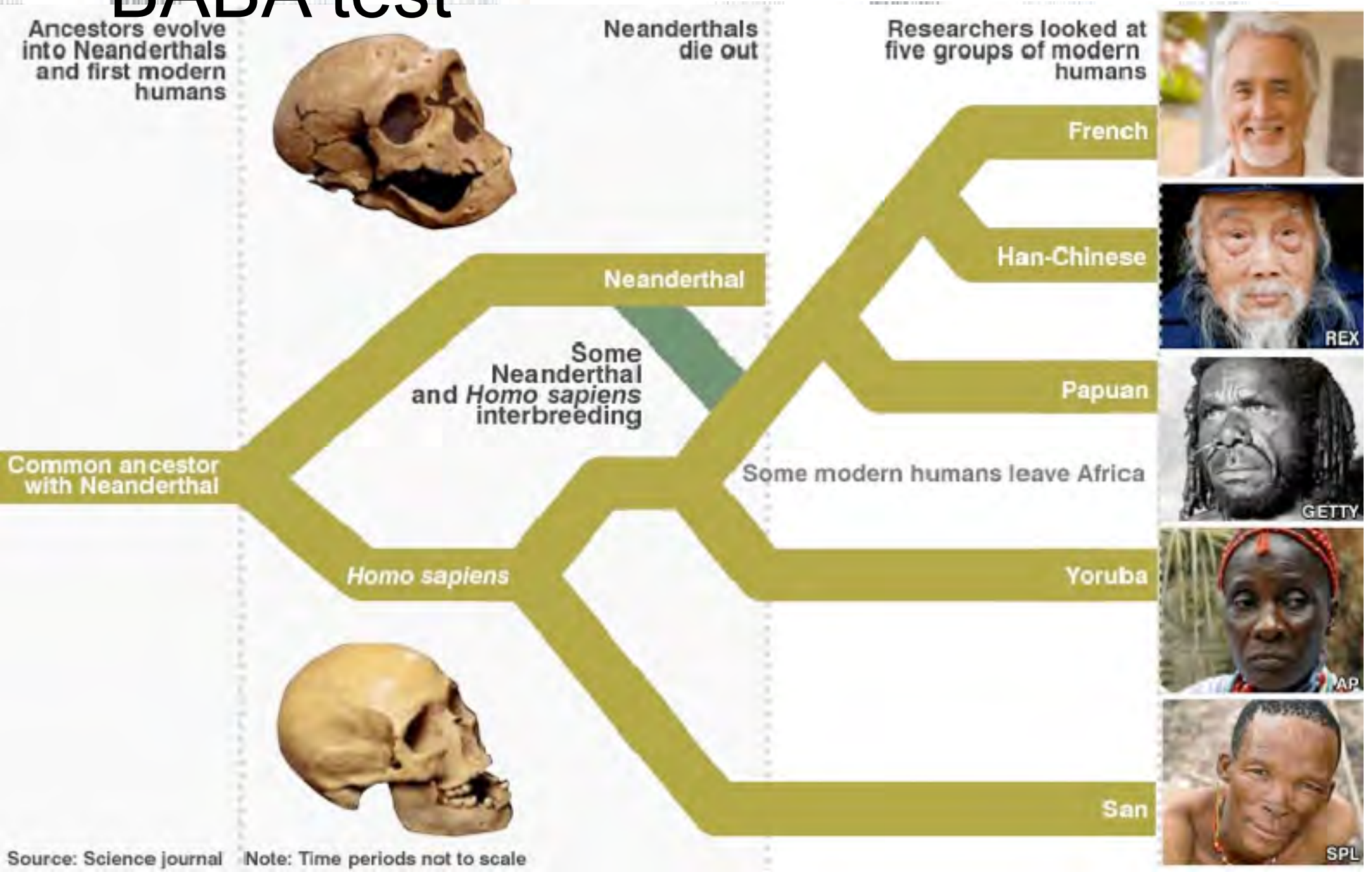
# Explicit tests for gene flow: ABBA-BABA test



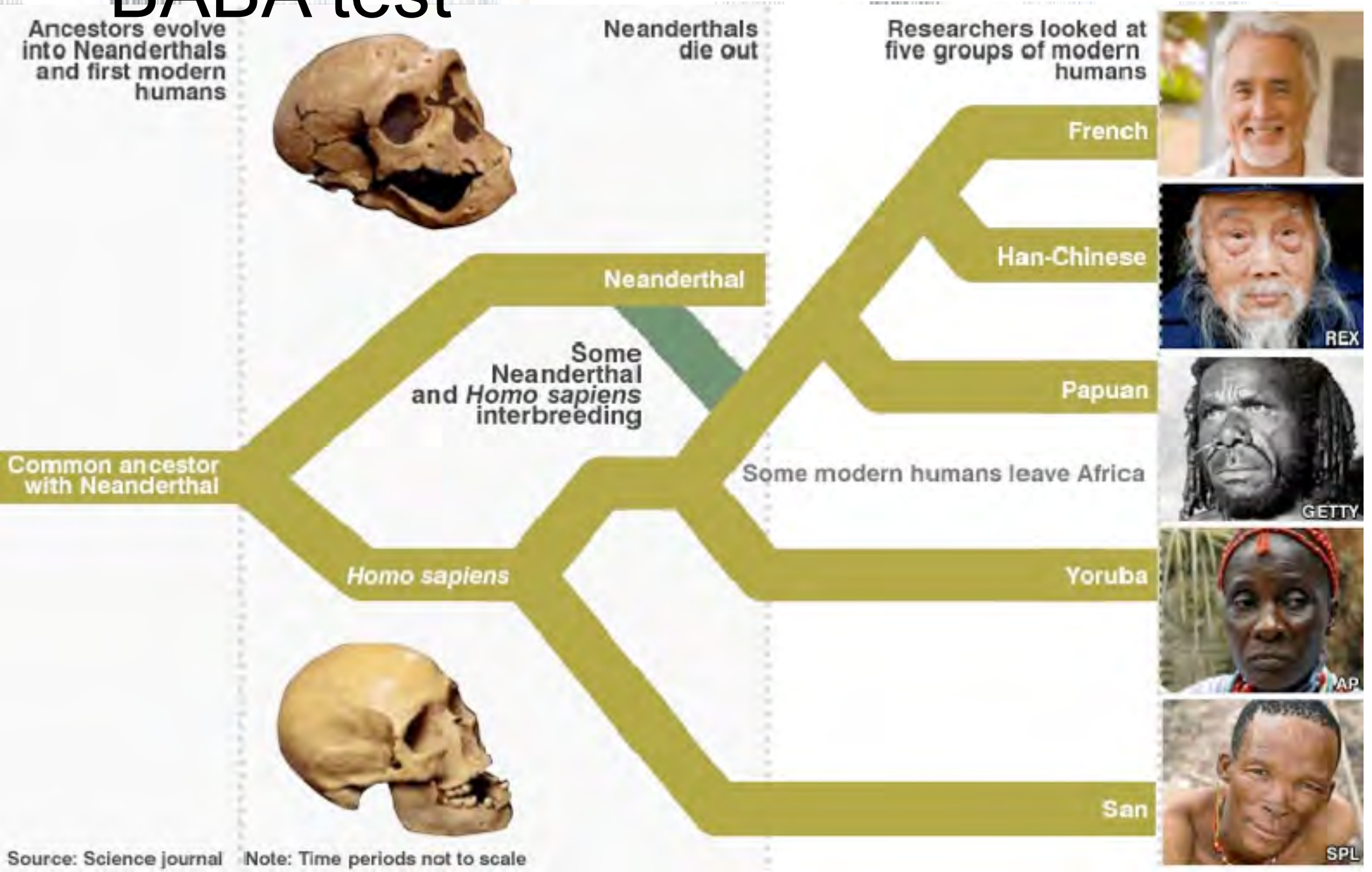

$$D(P_1, P_2, P_3, O) = \frac{\sum C_{ABBA}(i) - C_{BABA}(i)}{\sum C_{ABBA}(i) + C_{BABA}(i)} \qquad (1)$$

Green *et al.* 2010 *Science* 328:710-722

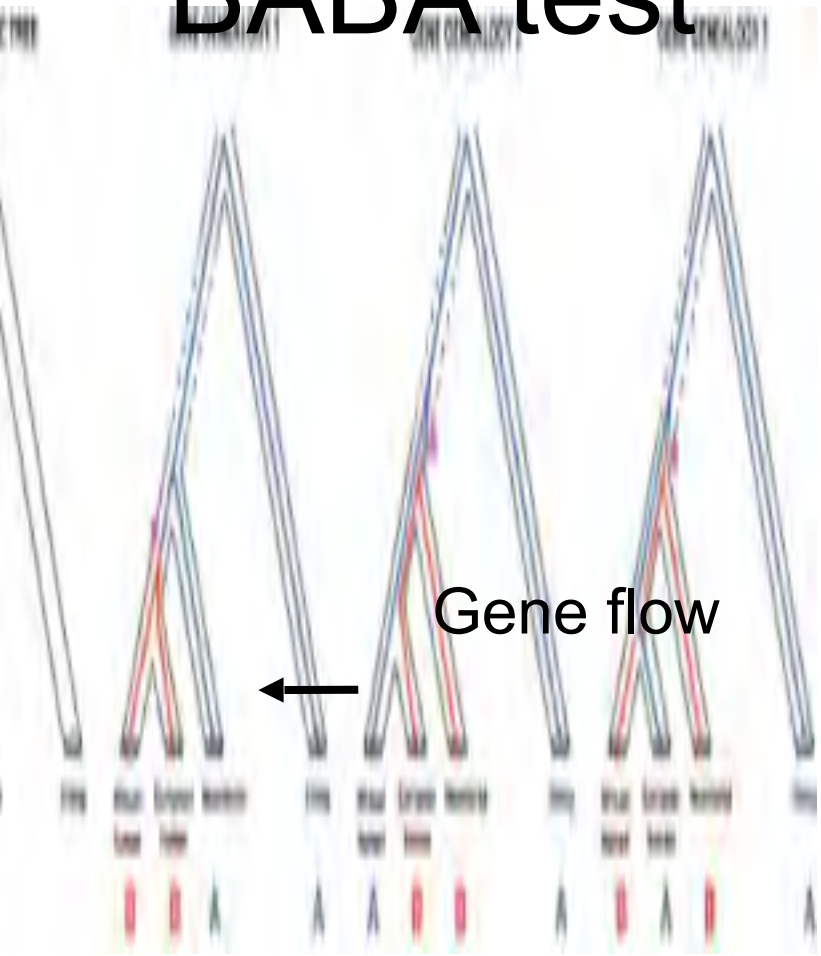
# Explicit tests for gene flow: ABBA-BABA test



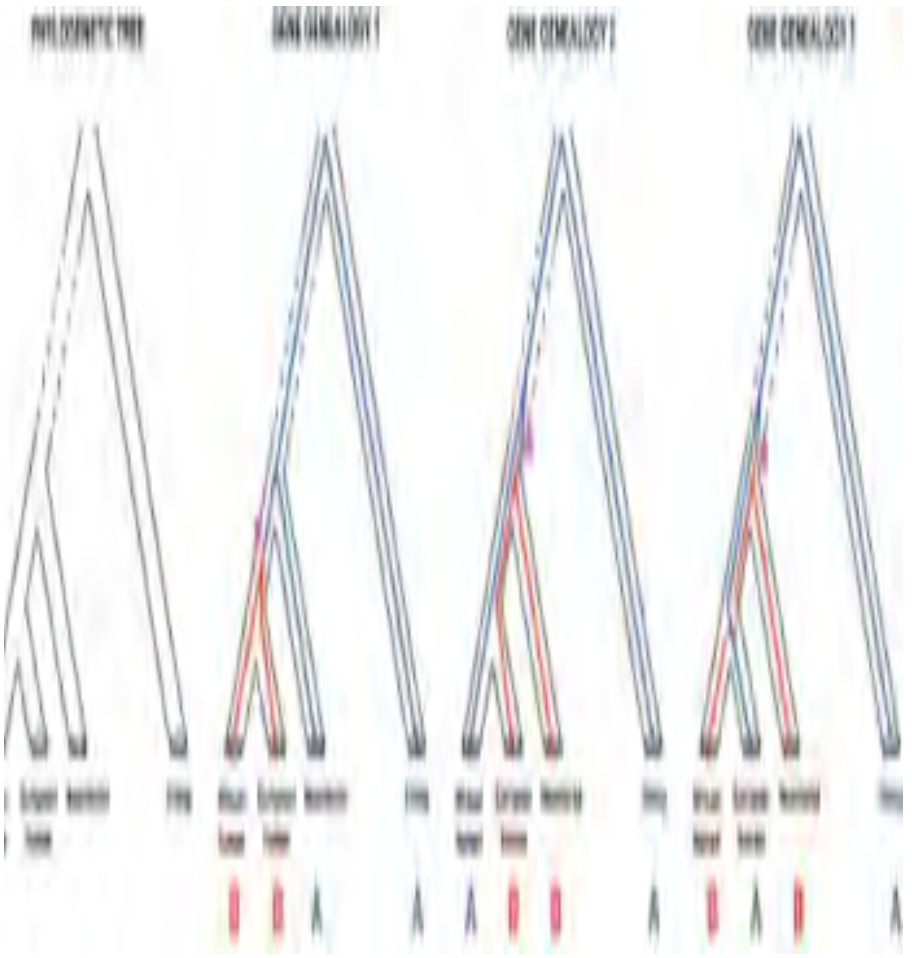
# Explicit tests for gene flow: ABBA-BABA test



# Explicit tests for gene flow: ABBA-BABA test



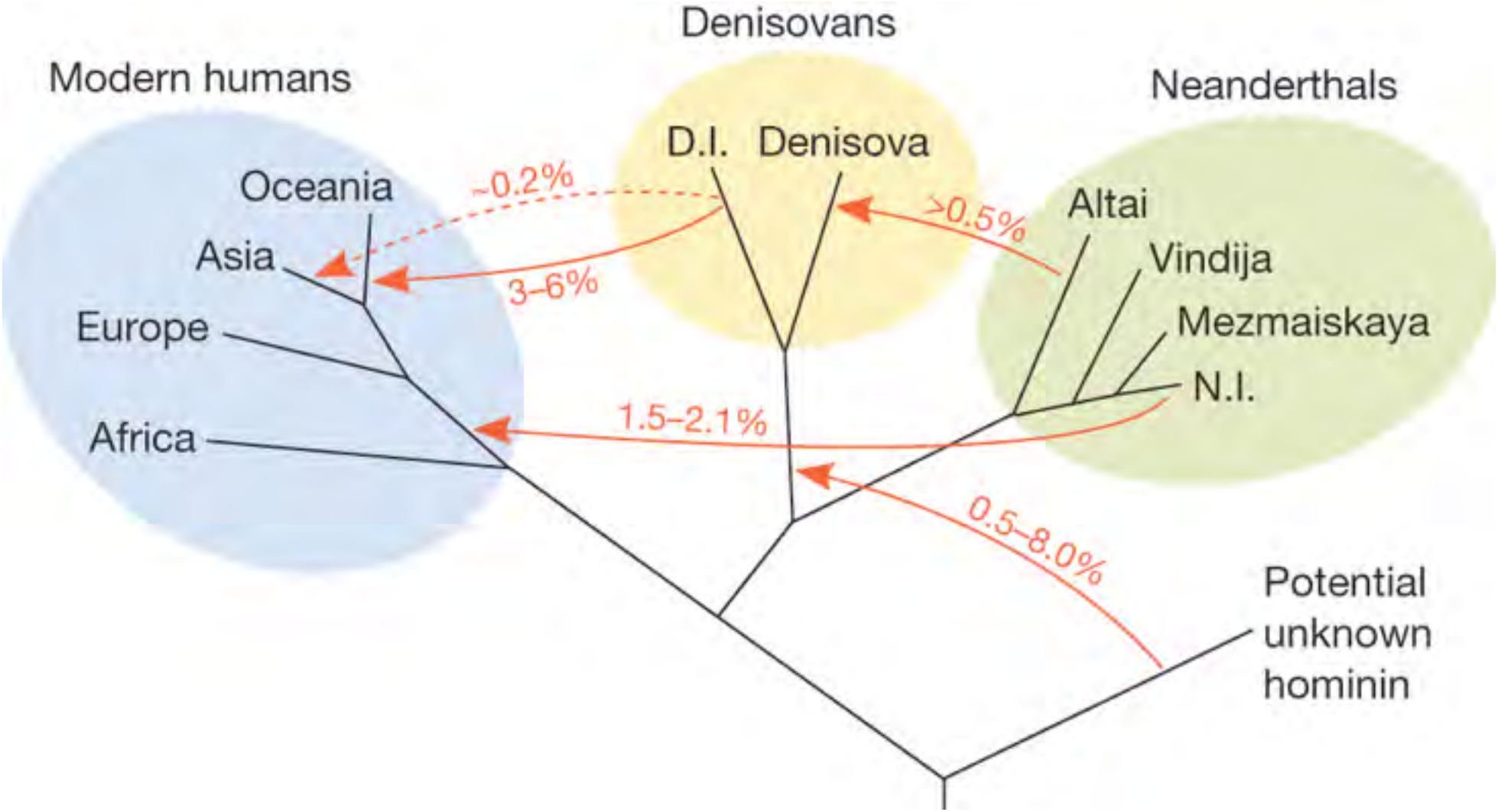

EXPECT:
50% ABBA
50% BABA



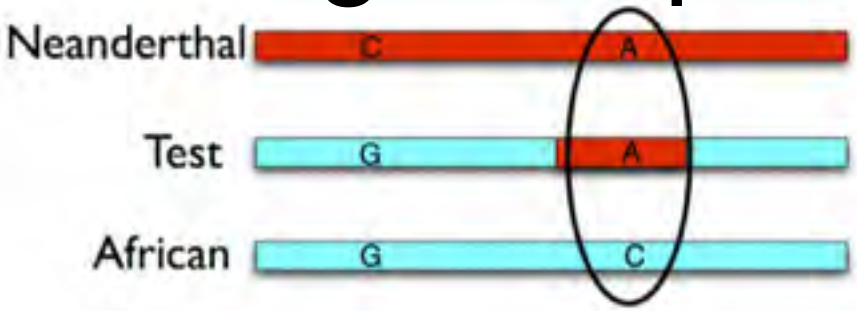
OBSERVE:
103612 ABBA
94029 BABA

Green *et al.* 2010
*Science* 328:710-722

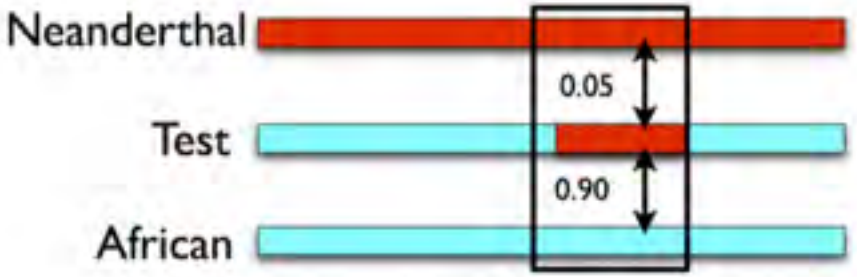
# Explicit tests for gene flow: ABBA-BABA test



# Explicit tests for gene flow: Combining multiple signals




Shared derived alleles

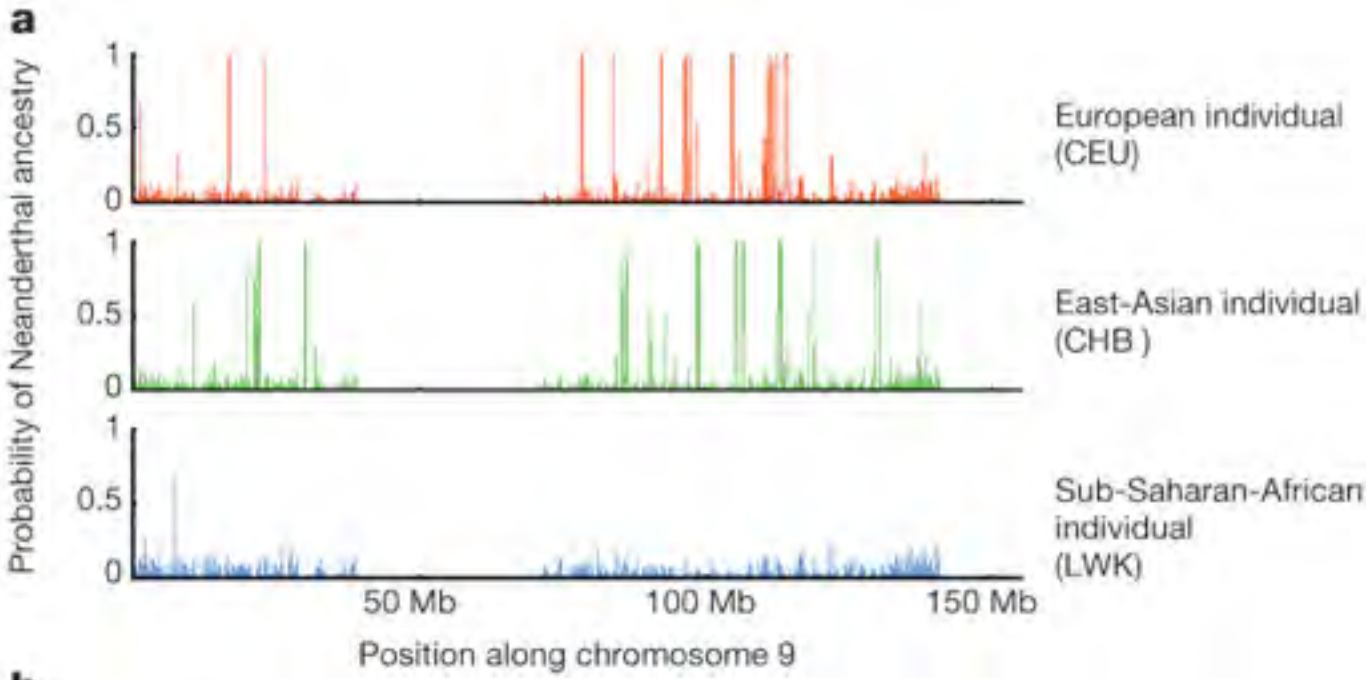


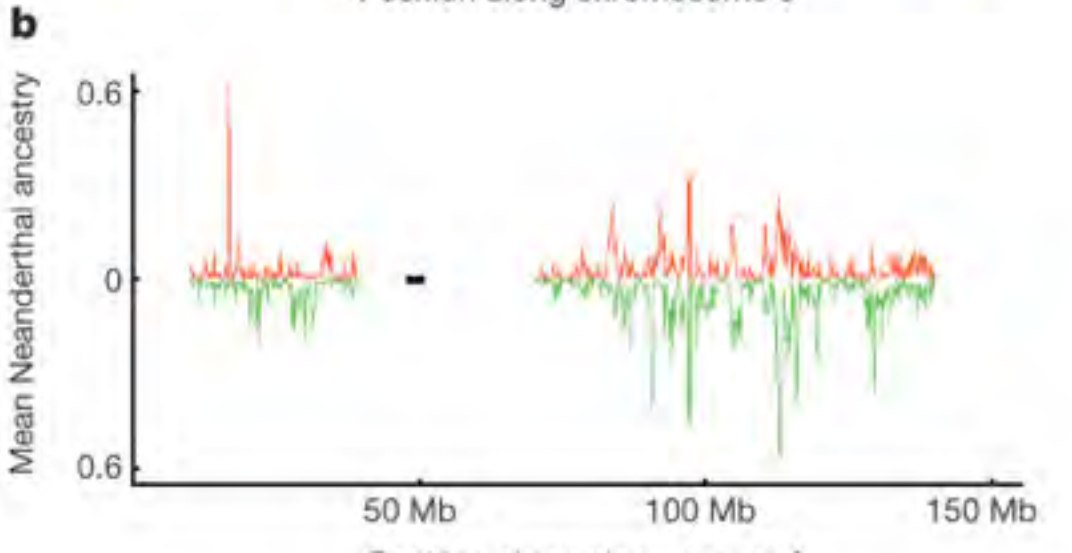

Segments similar to Neanderthal and divergent from African


Expected length of region given time since gene flow

Sankararaman, S. et al.
Nature 507, 354–7 (2014).

# Explicit tests for gene flow: Combining multiple signals

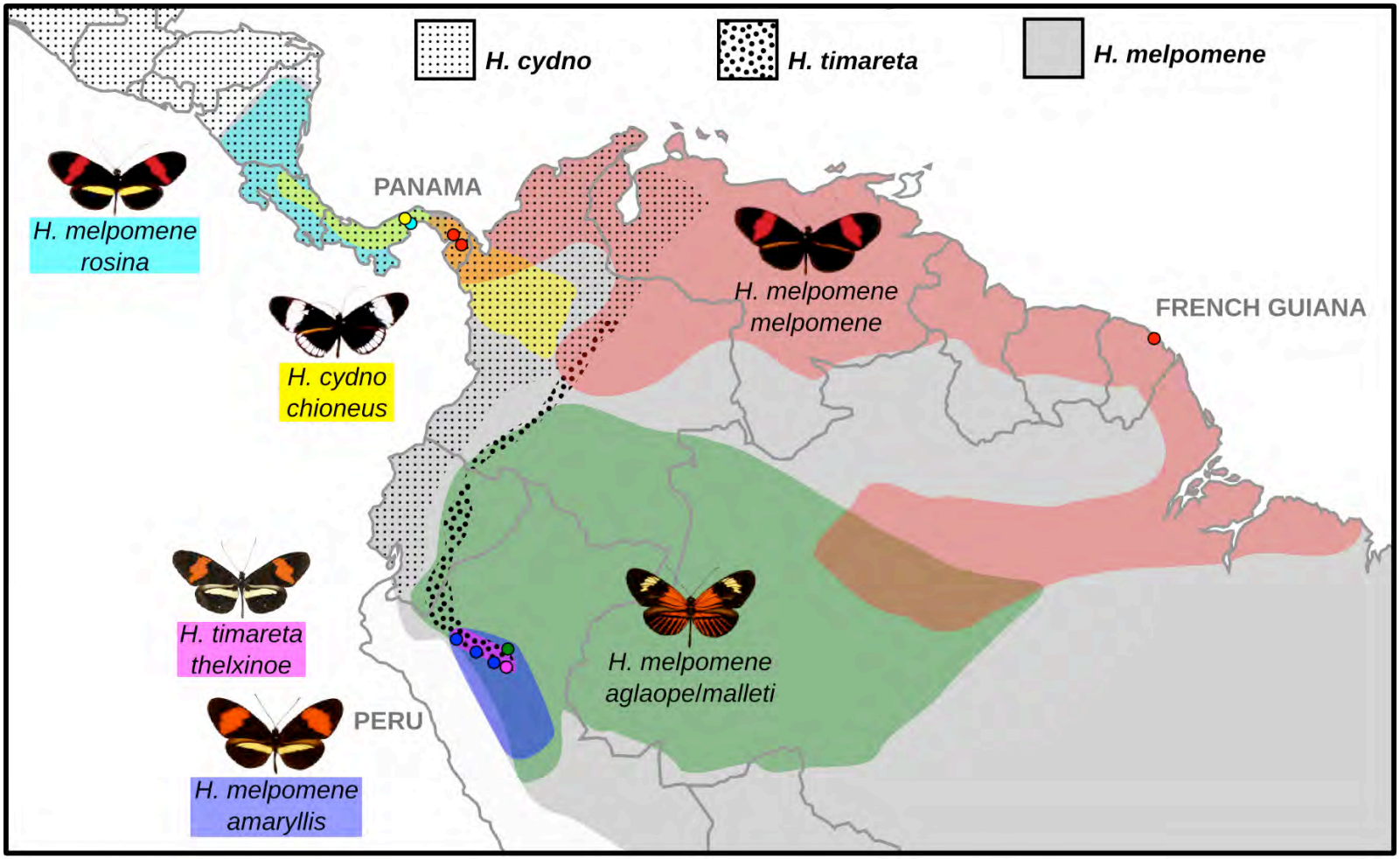



1) Derived alleles at high frequency shared with Neanderthal
2) High divergence to Africa but low to Neanderthal
3) Long haplotype blocks

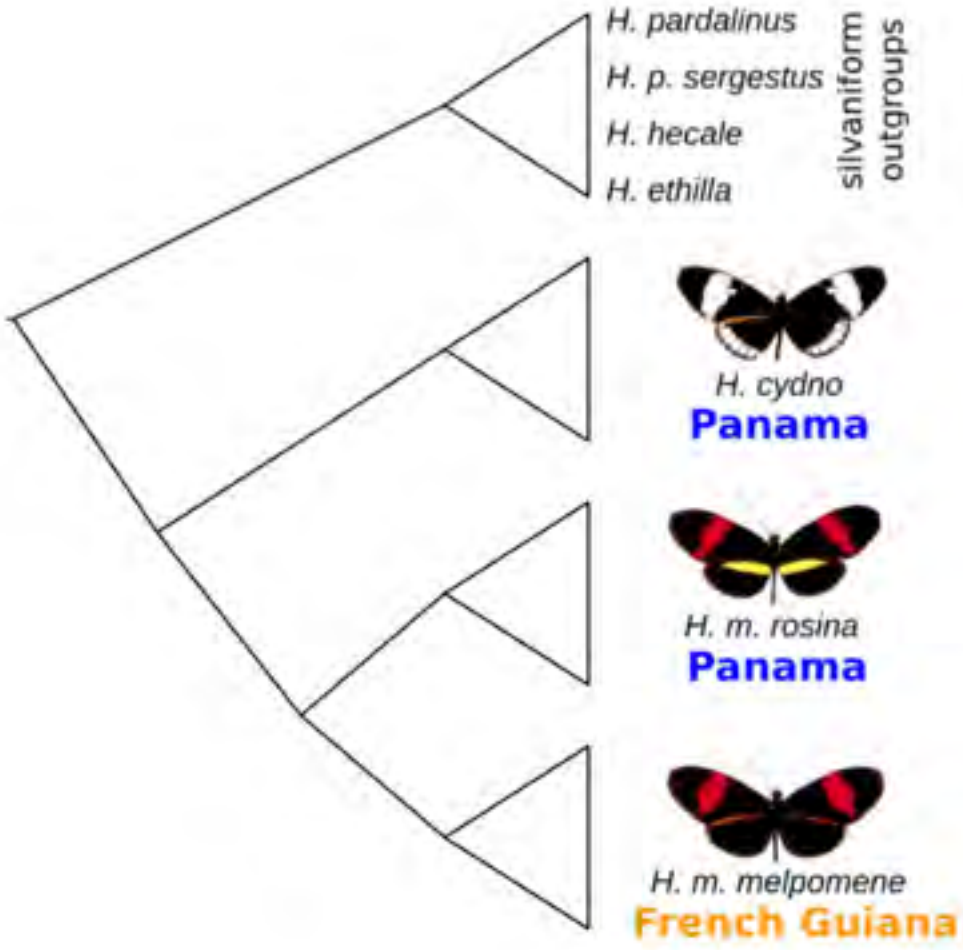
**The genomic landscape of Neanderthal ancestry in present-day humans - Sankararaman et al. Nature 2014**

# Explicit tests for gene flow: *Heliconius* butterflies




Martin et al., Genome Research 2013

# Explicit tests for gene flow: *Heliconius* butterflies

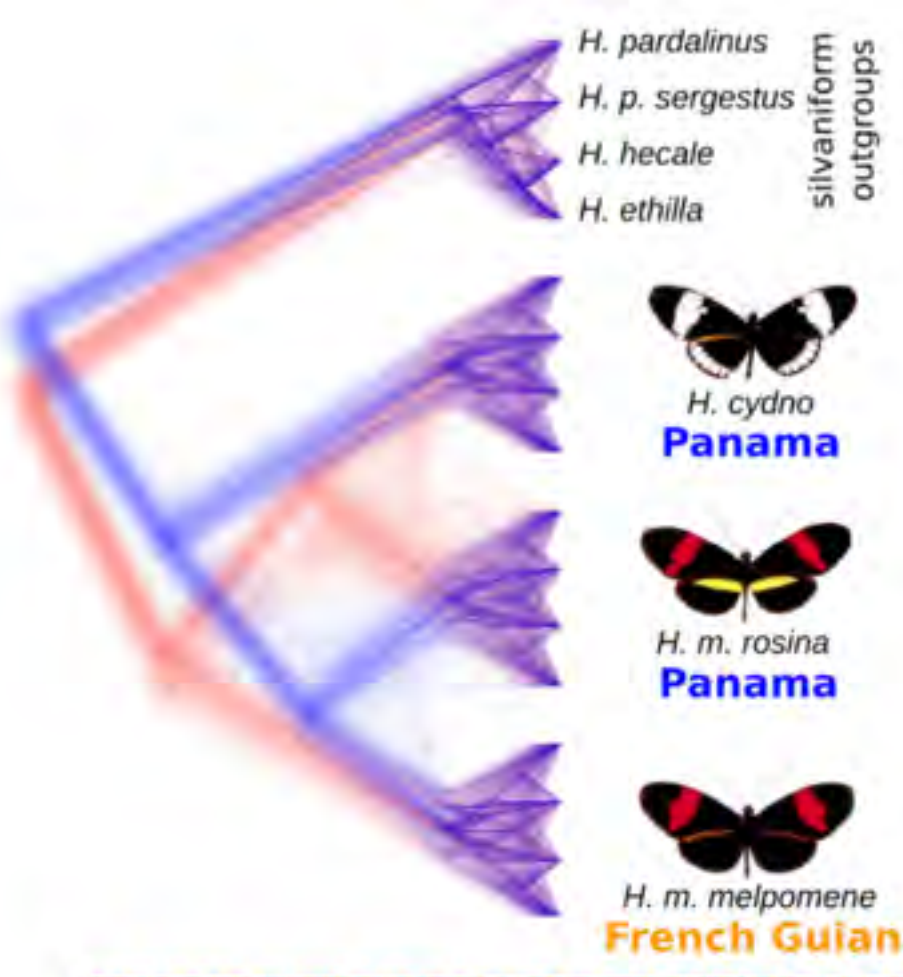



Whole-genome phylogeny supports grouping by species

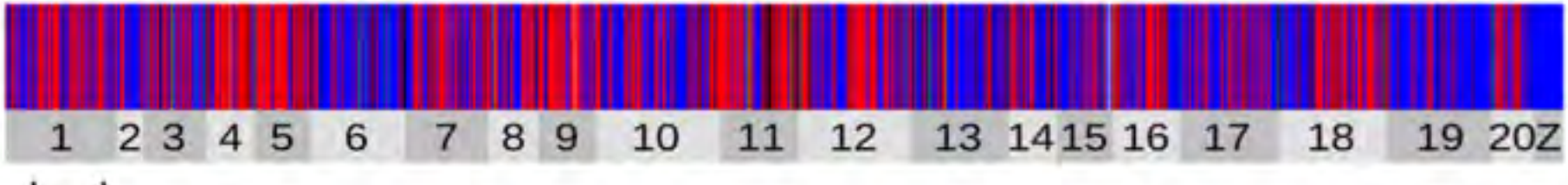
Many sources of reproductive isolation:

Female hybrids are sterile
Different host plant use
Different habitat preference
Strong assortative mating

# Explicit tests for gene flow: *Heliconius* butterflies



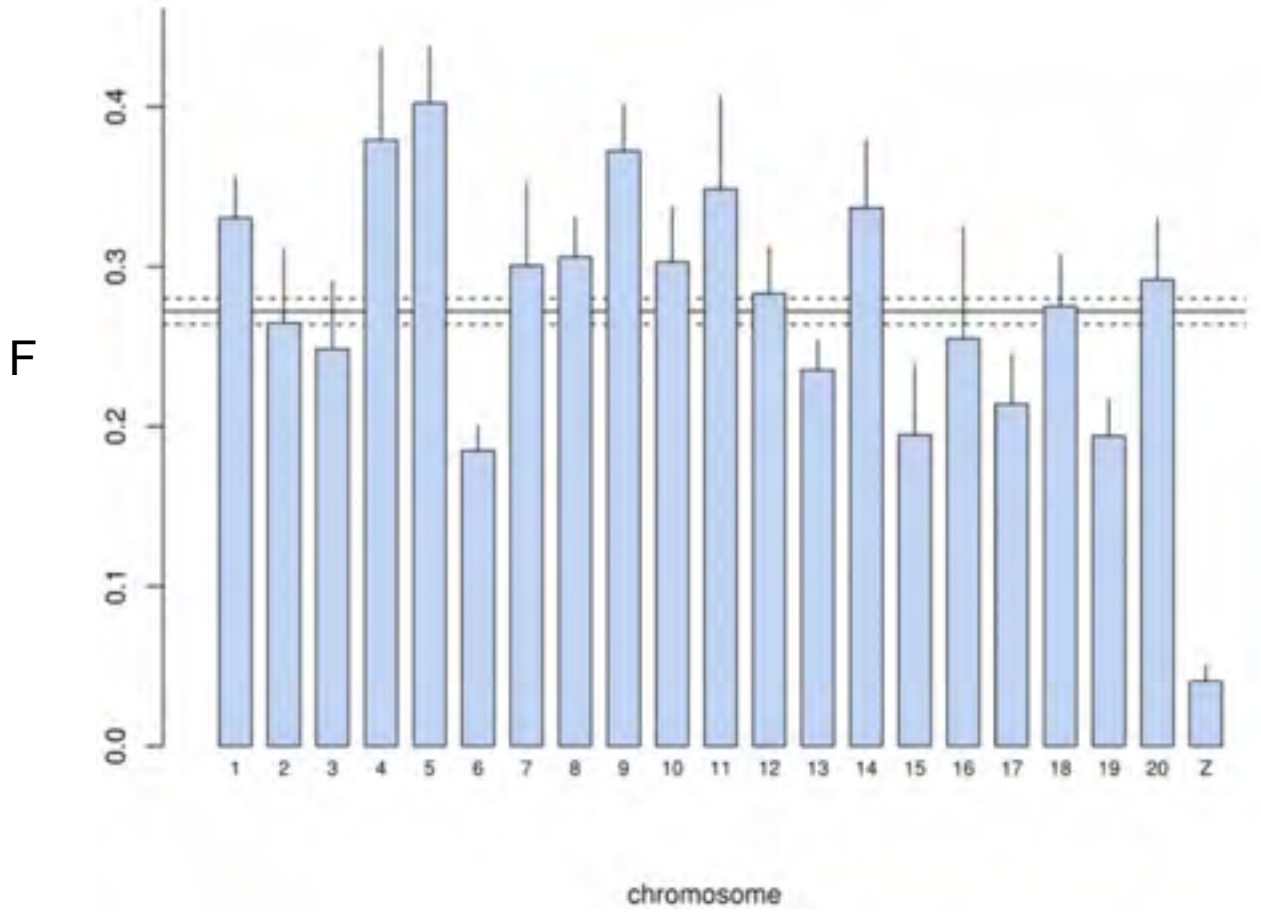

- 100 kb trees
- Only 53% group by species
- 42% group by geography!

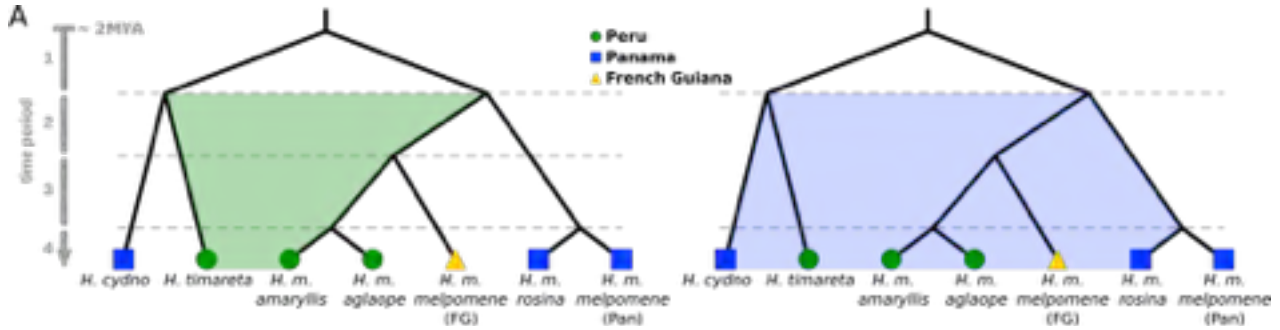


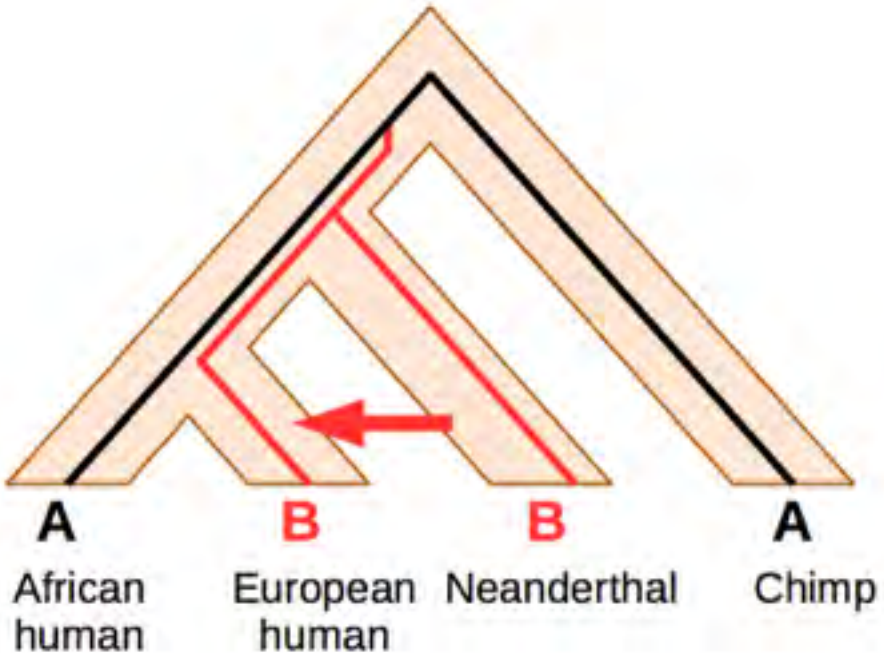

Martin et al. 2013 Genome Research

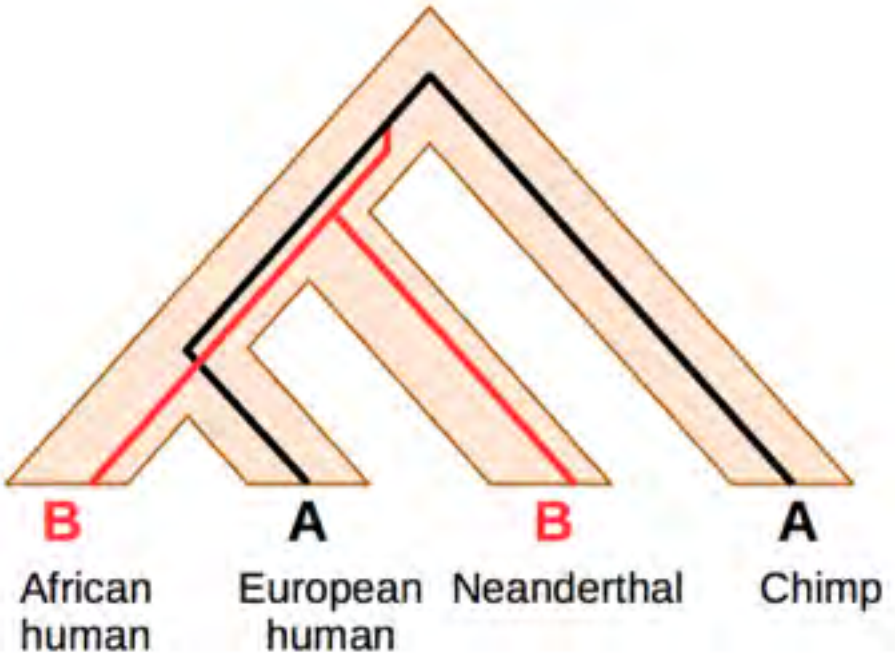
# Explicit tests for gene flow: *Heliconius* butterflies

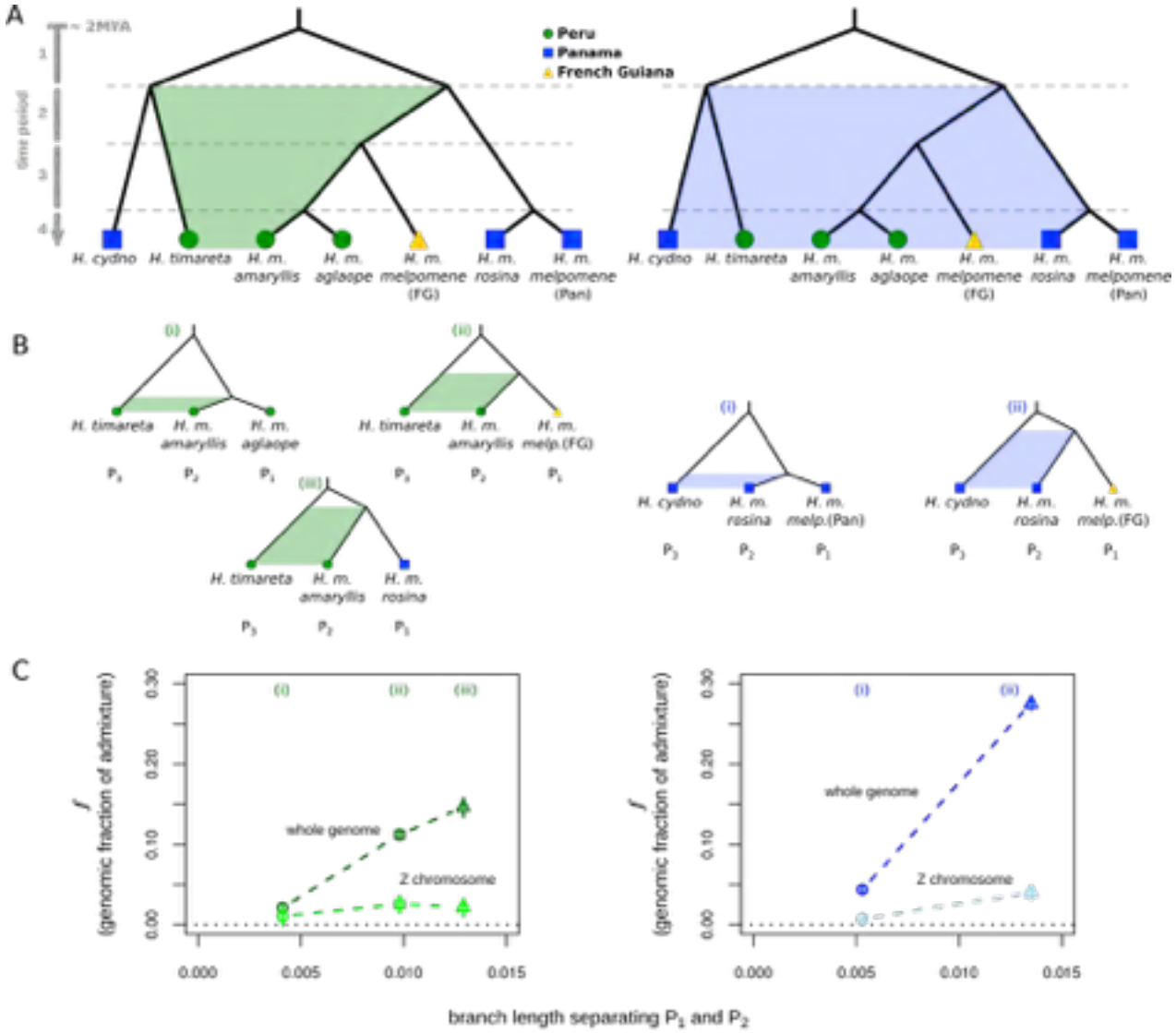
- Much larger proportion of genome is flowing as compared to Neanderthals
- Similarly strong effect on sex chromosome

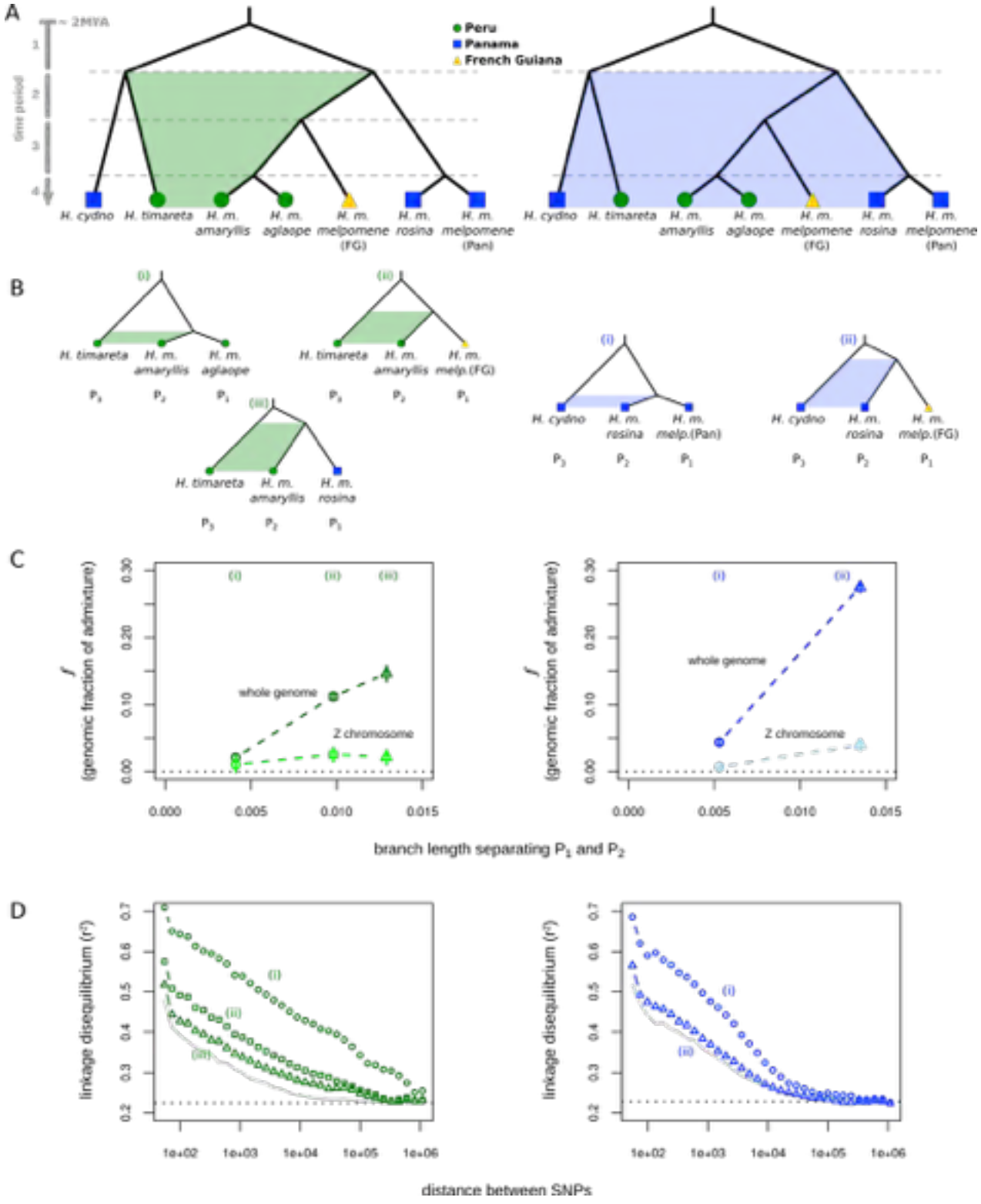


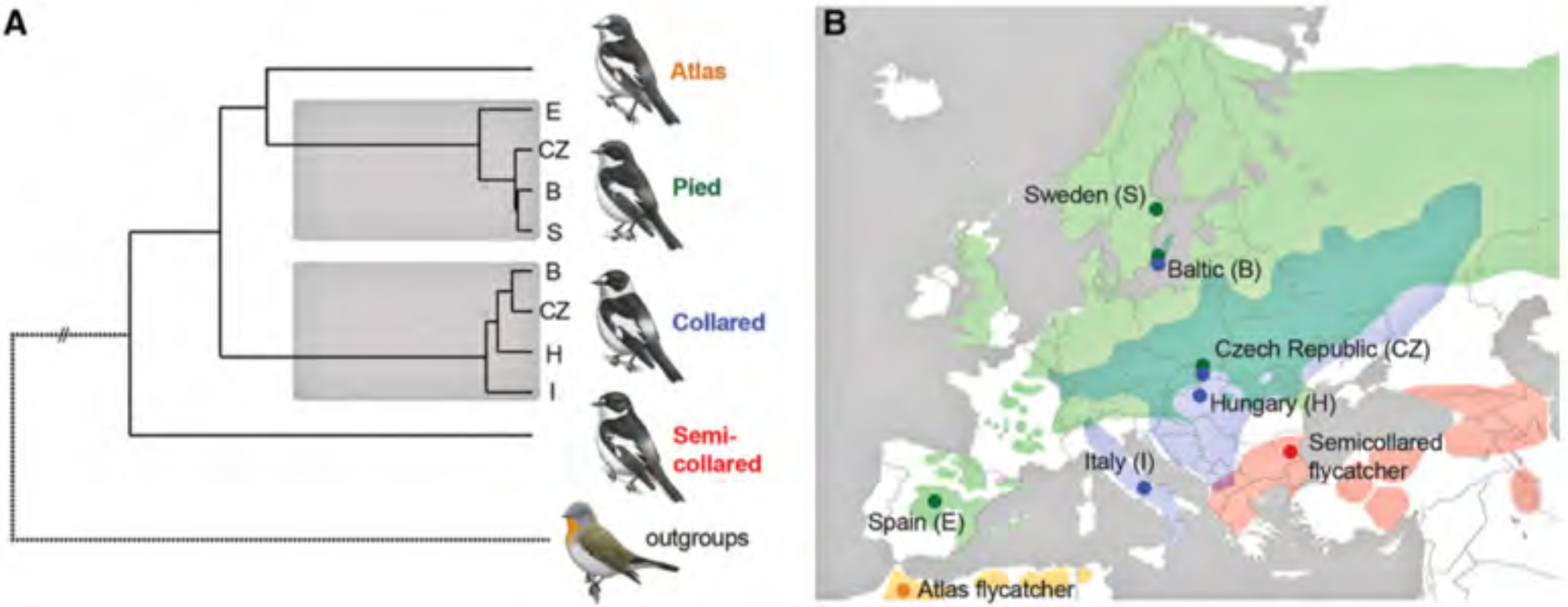

A
~ 2MYA
time period
Peru
Panama
French Guiana
H. cydno
H. timareta
H. m. amarylliis
H. m. aglaope
H. m. melpomene (FG)
H. m. rosina
H. m. melpomene (Pan)


A
B
B
A
African human
European human
Neanderthal
Chimp


B
A
B
A
African human
European human
Neanderthal
Chimp

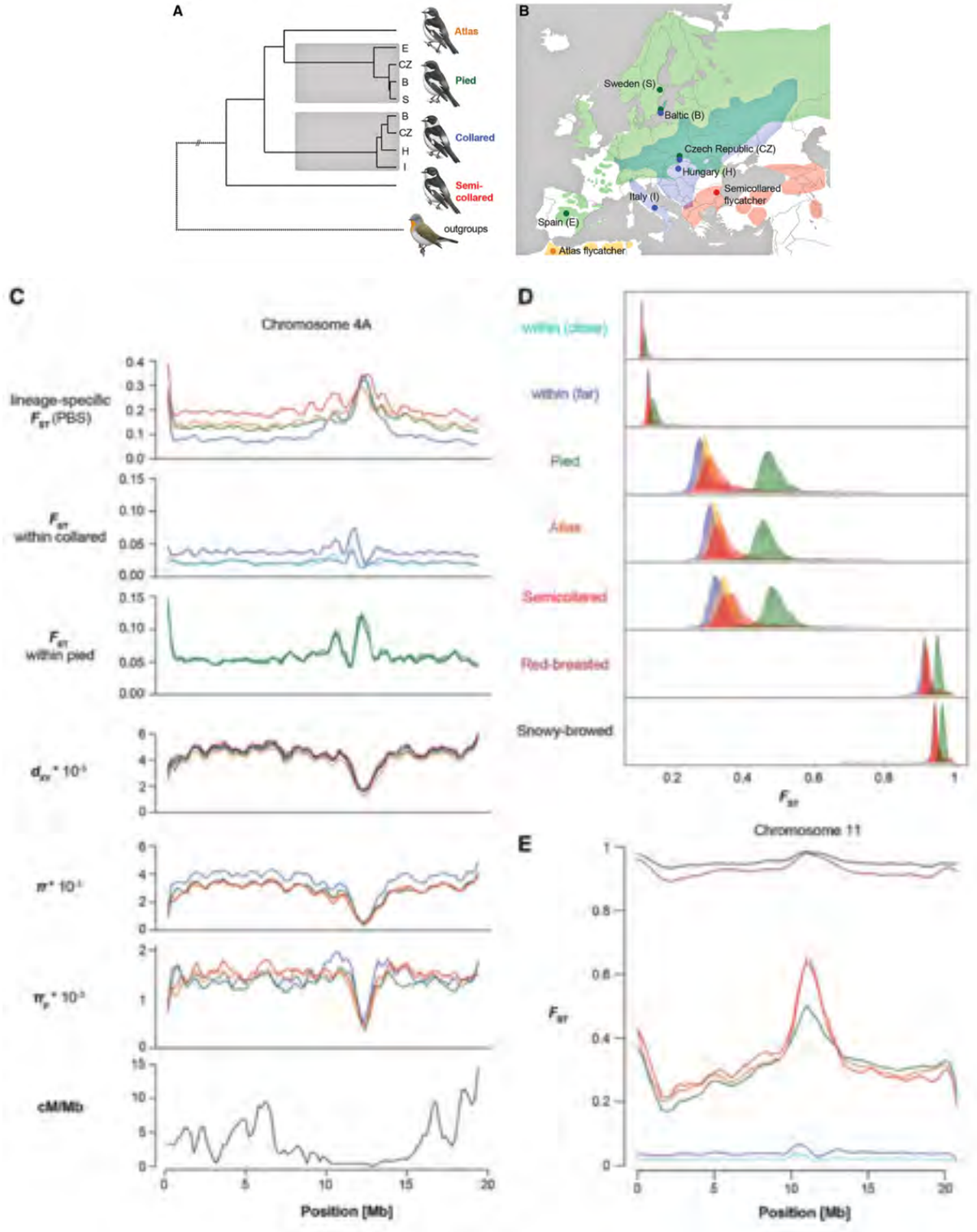

A
~ 2MYA
time period
Peru
Panama
French Guiana
H. cydno
H. timareta
H. m. amarylllis
H. m. aglaope
H. m. melpomene (FG)
H. m. rosina
H. m. melpomene (Pan)
B
(i)
(ii)
(iii)
H. timareta
H. m. amarylllis
H. m. aglaope
H. m. melp.(FG)
H. cydno
H. m. rosina
H. m. melp.(Pan)
P3
P2
P1
C
f (genomic fraction of admixture)
whole genome
Z chromosome
0.00
0.10
0.20
0.30
0.000
0.005
0.010
0.015
branch length separating P1 and P2


A
~ 2MYA
time period
Peru
Panama
French Guiana
H. cydno
H. timareta
H. m. amarylllis
H. m. aglaope
H. m. melpomene (FG)
H. m. rosina
H. m. melpomene (Pan)
B
(i)
(ii)
(iii)
H. timareta
H. m. amarylllis
H. m. aglaope
H. m. melp.(FG)
H. cydno
H. m. rosina
H. m. melp.(Pan)
P3
P2
P1
C
f (genomic fraction of admixture)
whole genome
Z chromosome
branch length separating P1 and P2
D
linkage disequilibrium (r²)
distance between SNPs




Sequenced 20 individuals per population at 20x coverage

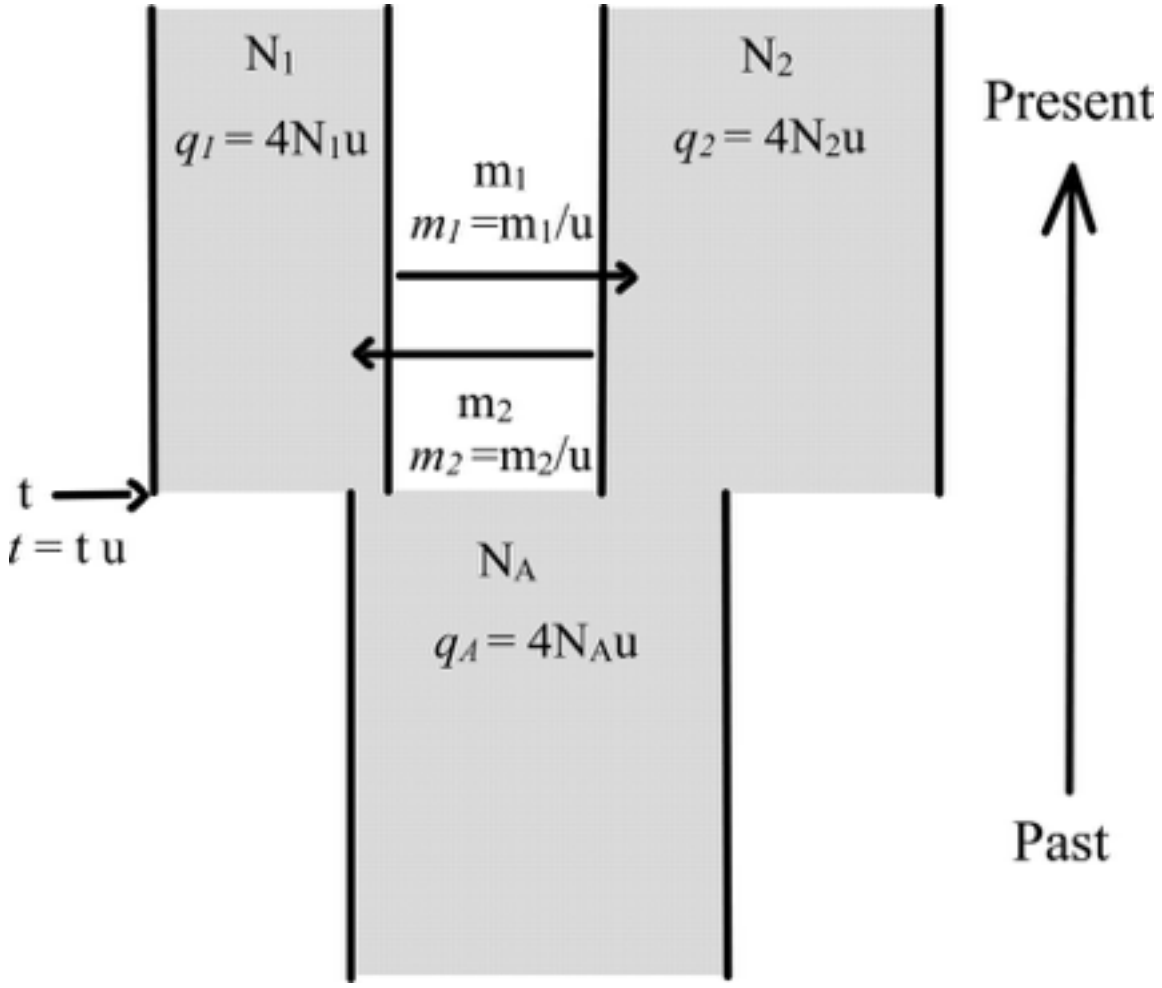
Burri et al., Genome Research 2015


A
Atlas
E
CZ
Pied
B
S
B
CZ
Collared
H
I
Semi-
collared
outgroups
B
Sweden (S)
Baltic (B)
Czech Republic (CZ)
Hungary (H)
Italy (I)
Semicollared
flycatcher
Spain (E)
Atlas flycatcher
C
Chromosome 4A
lineage-specific
$F_{ST}$ (PBS)
$F_{ST}$
within collared
$F_{ST}$
within pied
$d_{xy}$ * $10^{-3}$
$\pi$ * $10^{-3}$
$\pi_s$ * $10^{-3}$
cM/Mb
Position [Mb]
D
within (close)
within (far)
Pied
Atlas
Semicollared
Red-breasted
Snowy-browed
$F_{ST}$
E
Chromosome 11
$F_{ST}$
Position [Mb]

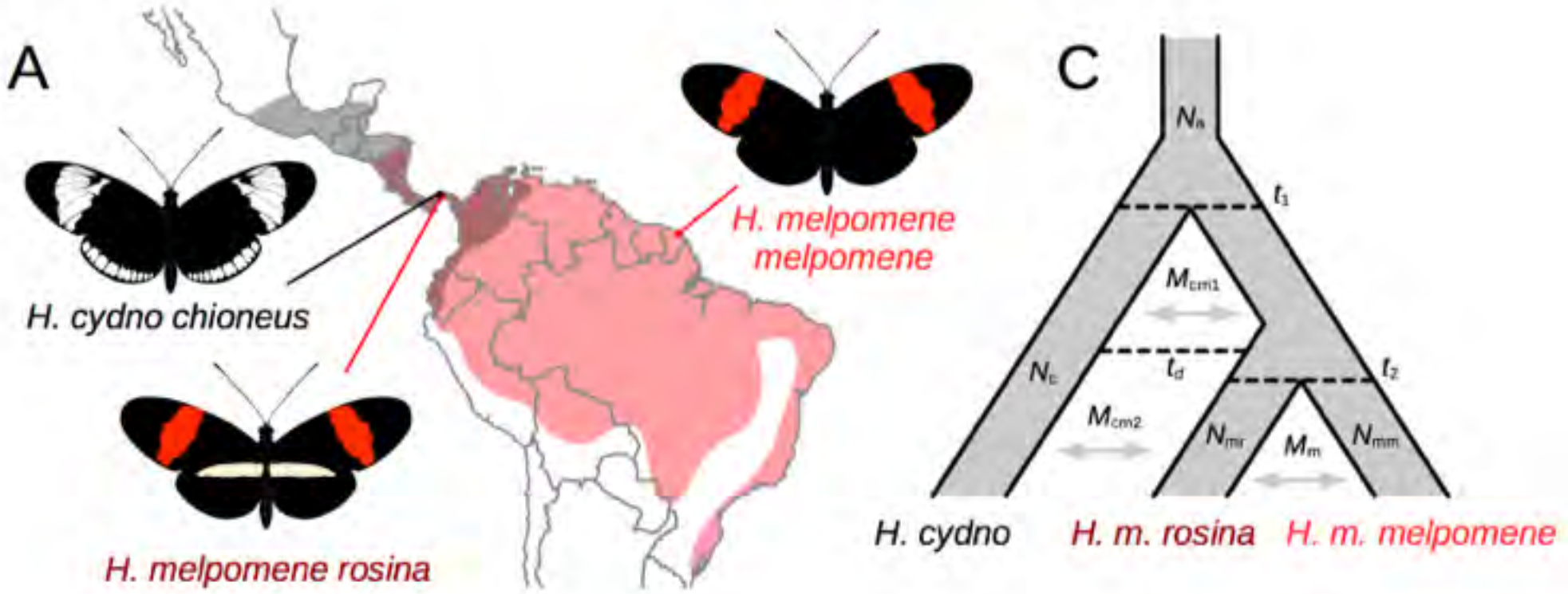
**Supplemental Table S4**. ABBA-BABA tests for gene flow. Populations/species among which the test indicates gene flow are highlighted in bold.

| 1. Inner | 2. Inner | 1. Outgroup | Mean(D) | SE(D) | p-value |
|---|---|---|---|---|---|
| collared Italy | collared CZ | pied CZ | 0.0010 | 0.0010 | 0.3344 |
| pied Spain | pied CZ | collared CZ | 0.0004 | 0.0005 | 0.4186 |
| **pied Spain** | Atlas | **collared Italy** | -0.1648 | 0.0027 | $<10^{-4}$ |
| **pied Spain** | Atlas | **semicollared** | -0.0108 | 0.0016 | $<10^{-4}$ |
| pied Spain | **collared Italy** | **semicollared** | 0.1162 | 0.0018 | $<10^{-4}$ |
| Atlas | **collared Italy** | **semicollared** | 0.1242 | 0.0016 | $<10^{-4}$ |

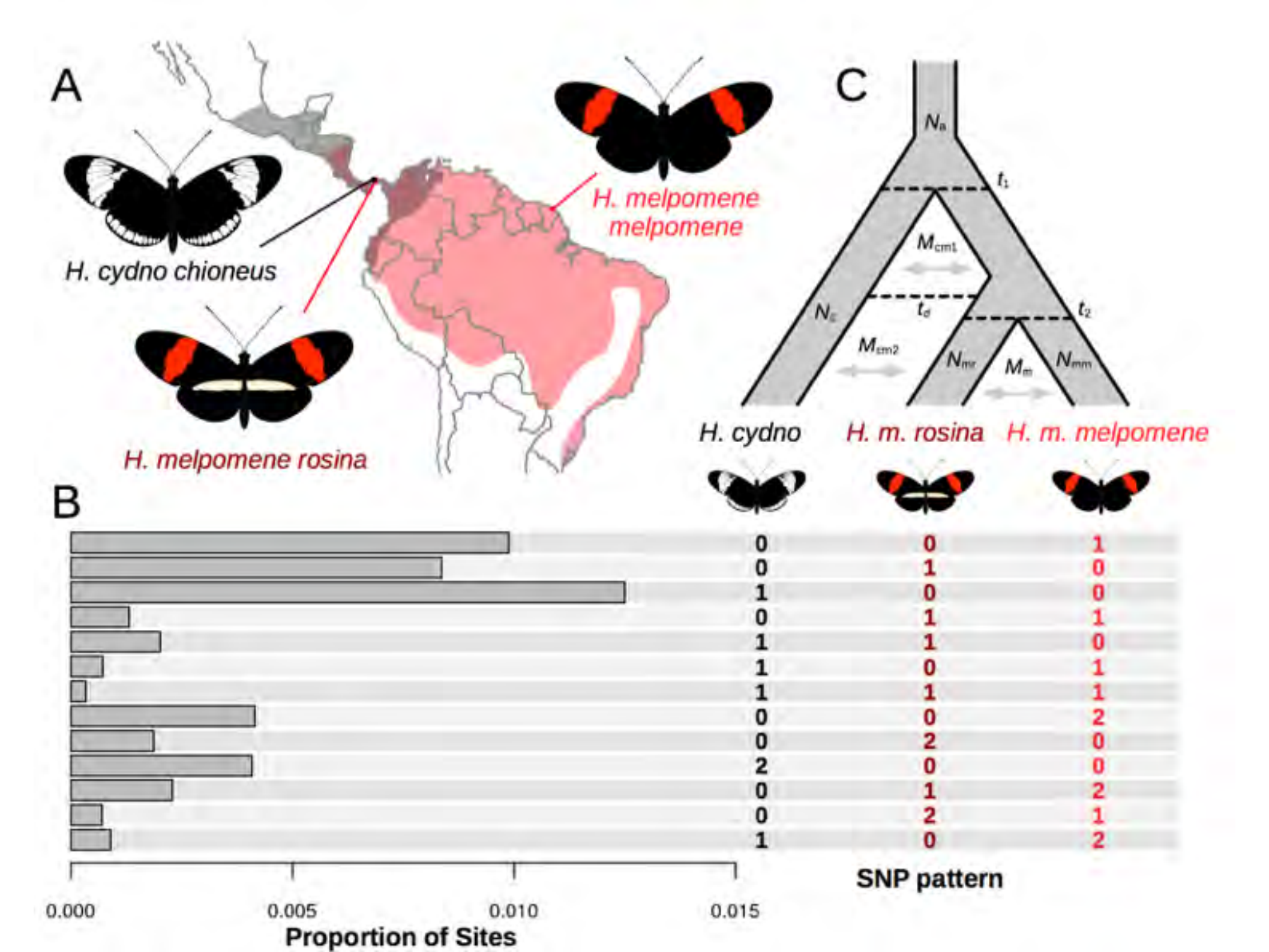
# An alternative is to take an explicit modelling approach

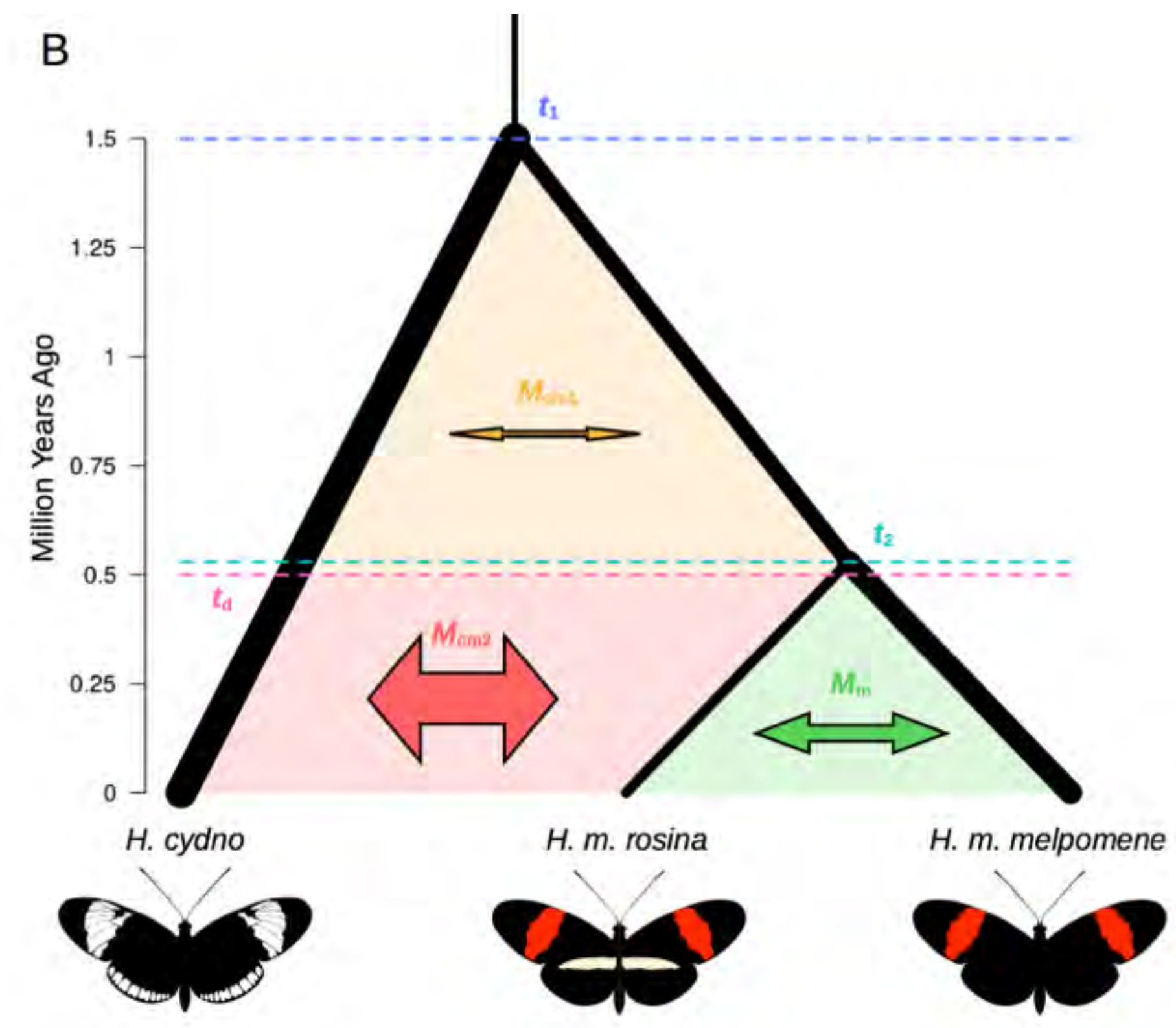


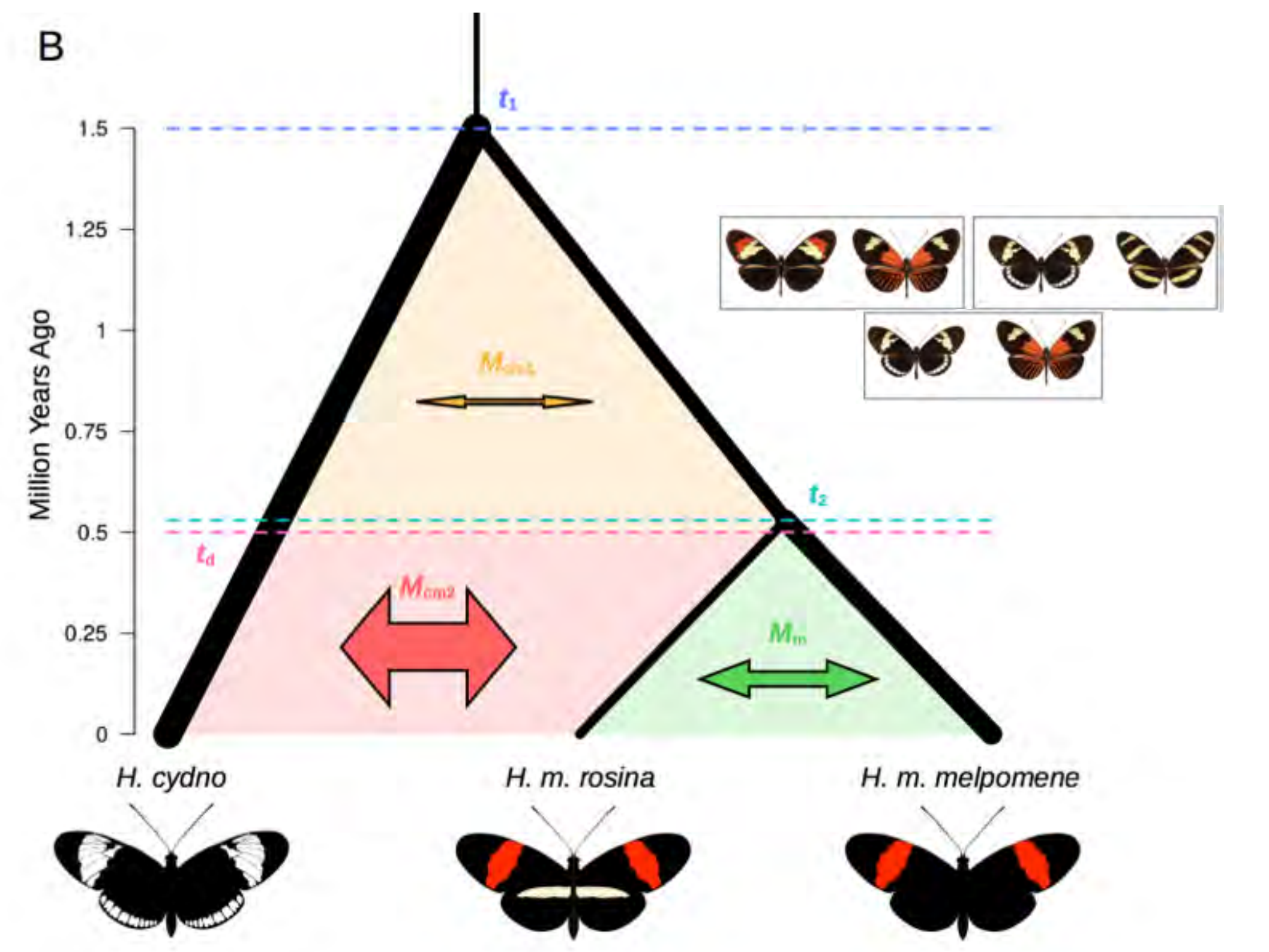

IM and IMa
Jody Hey




Martin et al., Biorxiv 2015


A
H. cydno chioneus
H. melpomene melpomene
H. melpomene rosina
C
$N_a$
$t_1$
$M_{cm1}$
$N_c$
$t_d$
$t_2$
$M_{cm2}$
$N_{mr}$
$M_m$
$N_{mm}$
H. cydno
H. m. rosina
H. m. melpomene
B
0 0 1
0 1 0
1 0 0
0 1 1
1 1 0
1 0 1
1 1 1
0 0 2
0 2 0
2 0 0
0 1 2
0 2 1
1 0 2
SNP pattern
0.000
0.005
0.010
0.015
Proportion of Sites


B
$t_1$
1.5
1.25
1
0.75
0.5
0.25
0
Million Years Ago
$M_{cm1}$
$t_2$
$t_d$
$M_{cm2}$
$M_m$
H. cydno
H. m. rosina
H. m. melpomene


B
$t_1$
1.5
1.25
1
0.75
0.5
0.25
0
Million Years Ago
$M_{cm1}$
$t_2$
$t_d$
$M_{cm2}$
$M_m$
H. cydno
H. m. rosina
H. m. melpomene

So far models have mostly just estimated genome-wide parameters…assuming the genome is homogenous

Where we need to go next is to incorporate genome heterogeneity in selection and recombination

So far models have mostly just estimated genome-wide parameters…assuming the genome is homogenous

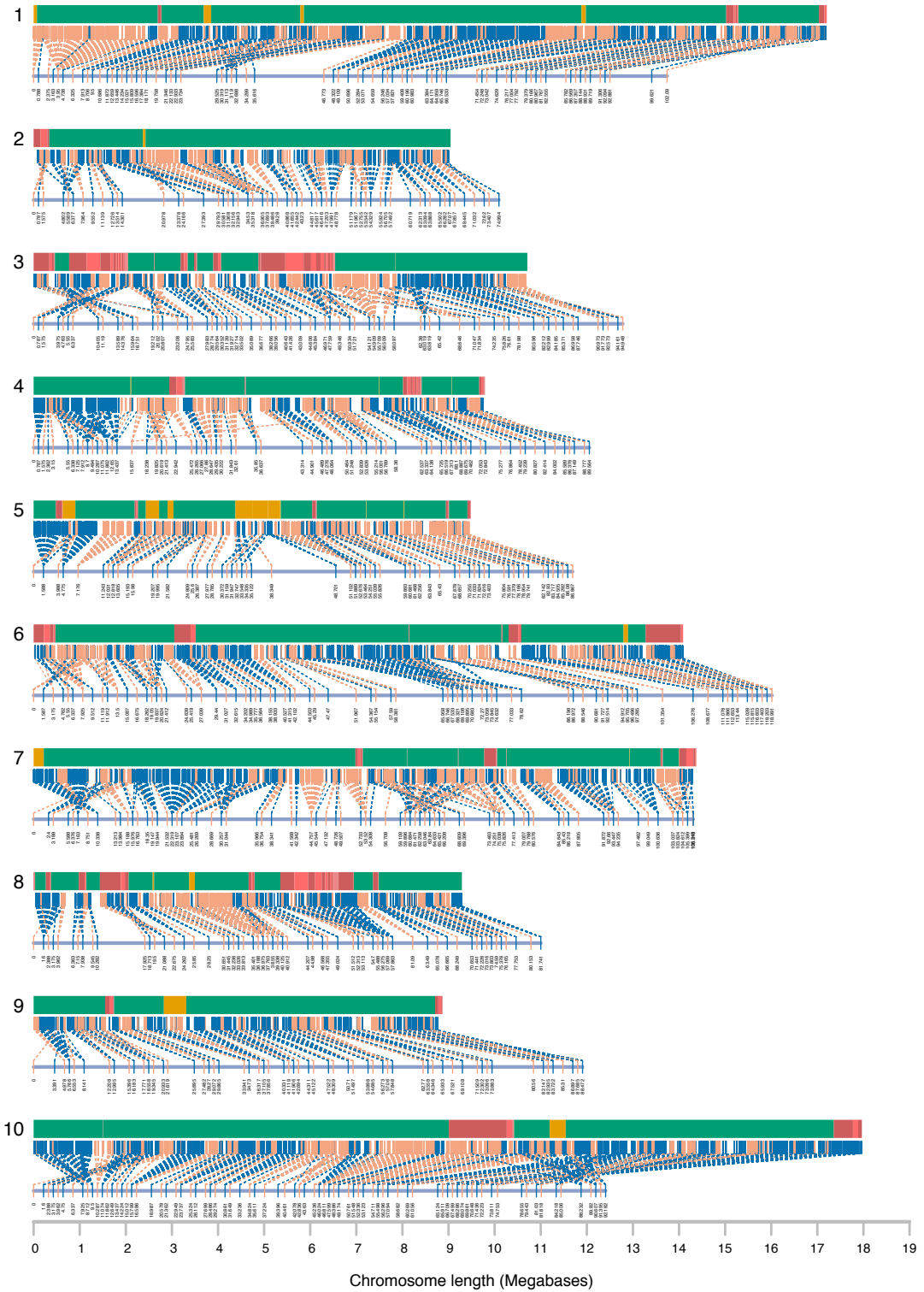
Where we need to go next is to incorporate genome heterogeneity in selection and recombination

# Identifying Loci Under Selection Against Gene Flow in Isolation-with-Migration Models

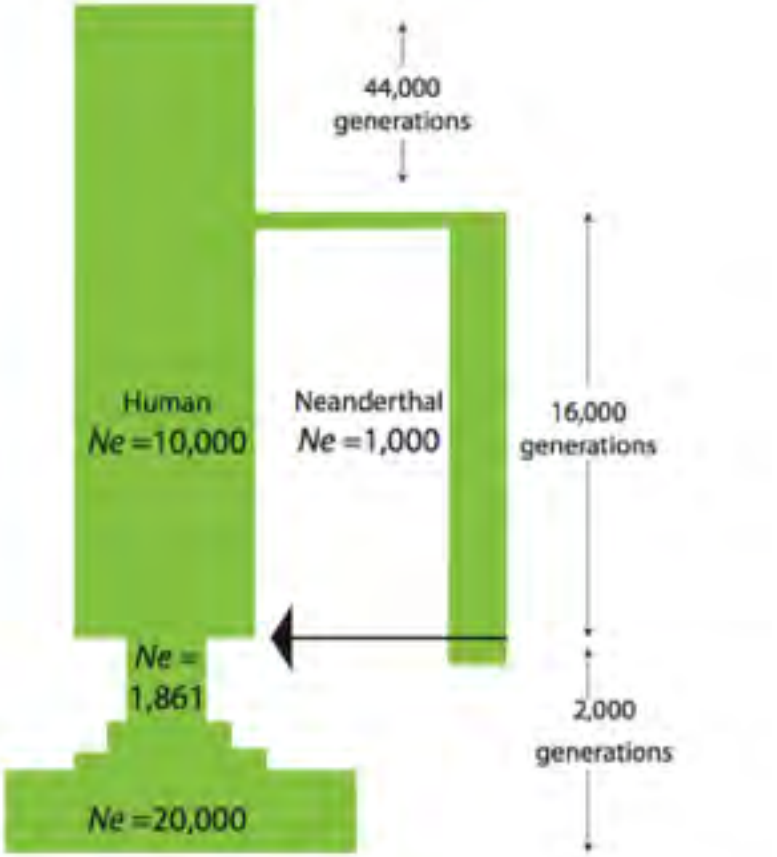
**Vitor C. Sousa,*,1,2 Miguel Carneiro,† Nuno Ferrand,† and Jody Hey*,1**

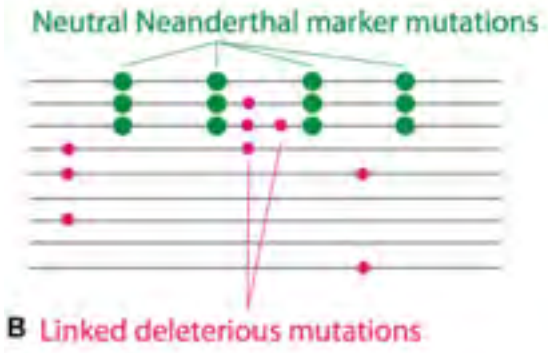
*Department of Genetics, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854, and †CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, 4099-002 Porto, Portugal

# High density linkage maps to map the recombination landscape



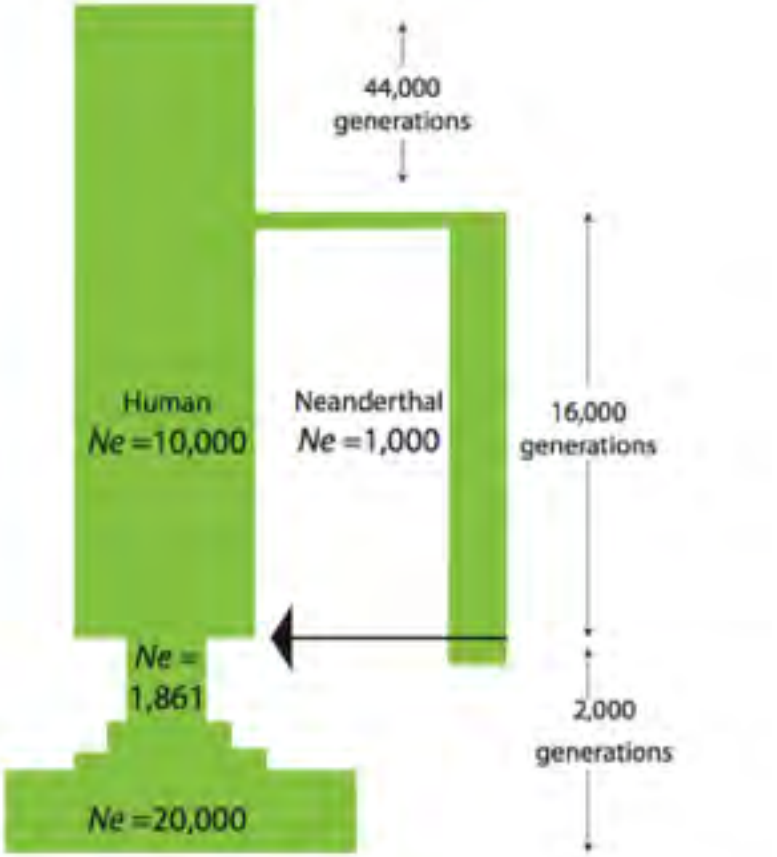
# The effect of background selection on introgression in humans

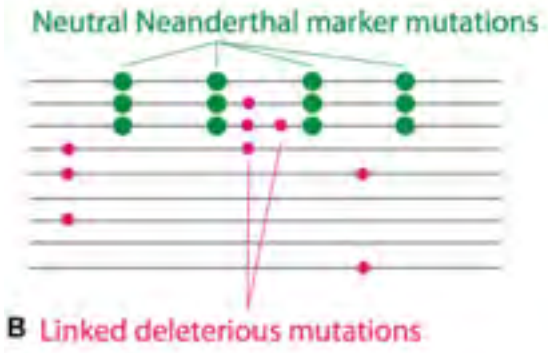



Harris and Nielson Biorxiv 2015

# The effect of background selection on introgression in humans




Admixture is less in gene rich regions supporting this model…..

Harris and Nielson Biorxiv 2015

# Population and speciation genomics: Conclusions

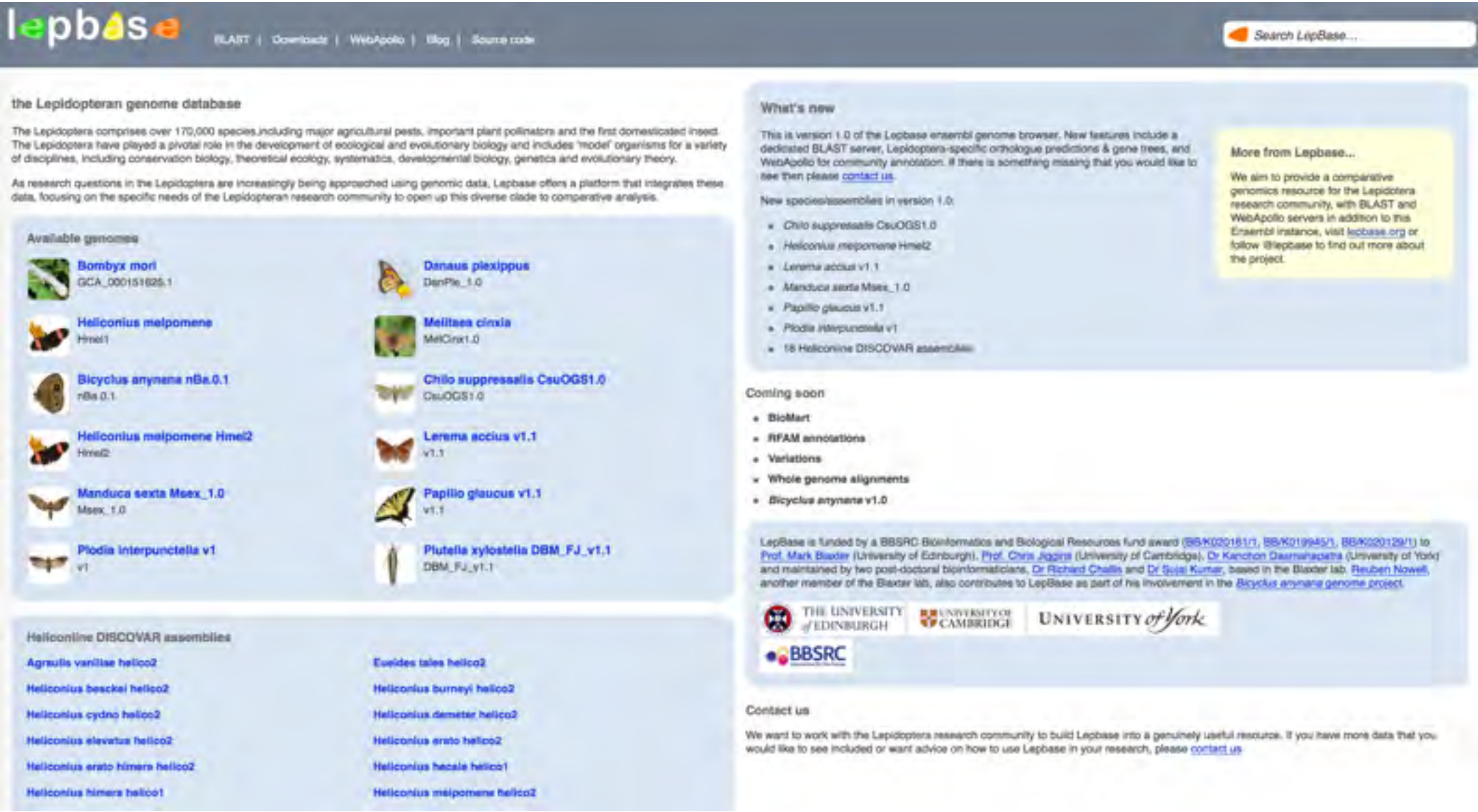
- Great power to detect subtle signals of selection and gene flow
- Can make more general observations about genes and regions involved in adaptation
- BUT genomic processes complicate the picture

- Best approaches combine multiple signals to infer process

- Eventually we need to combine background selection, recombination, positive selection

# And finally a shameless plug….

# And finally a shameless plug….



# Adaptive introgression

Photo credit
Andrei Sourakov

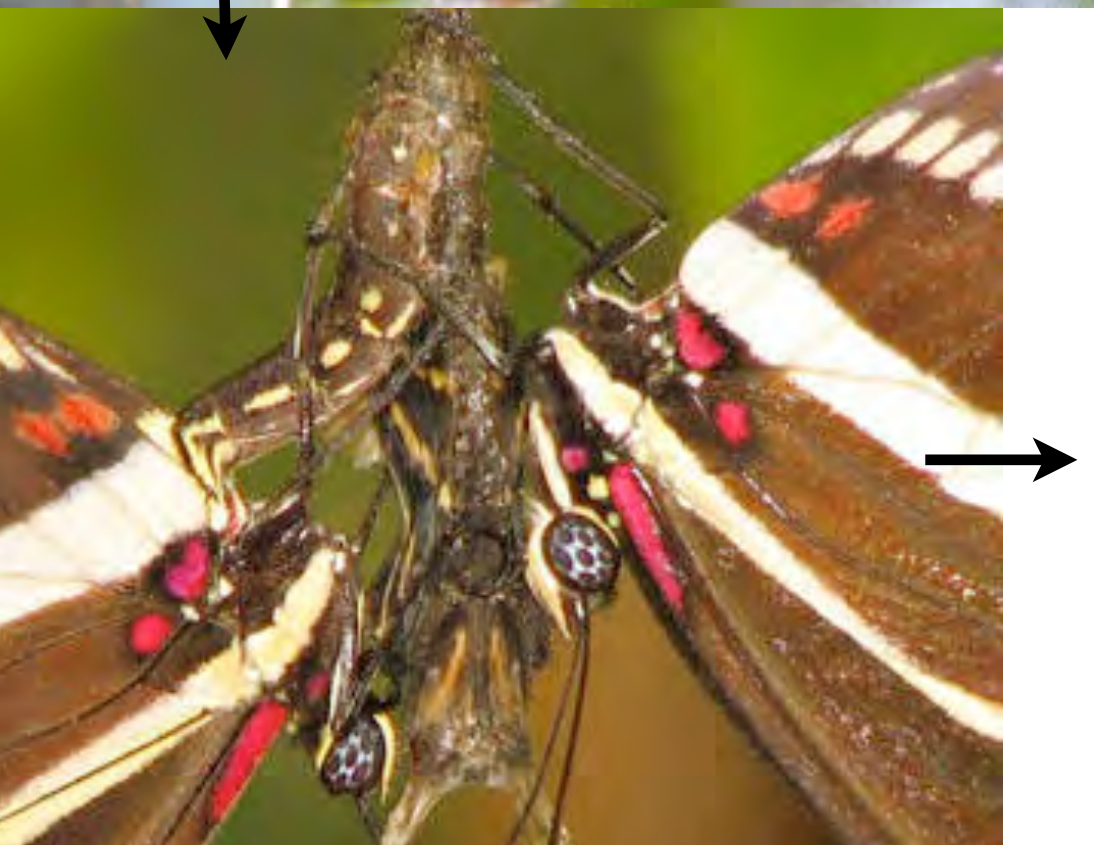
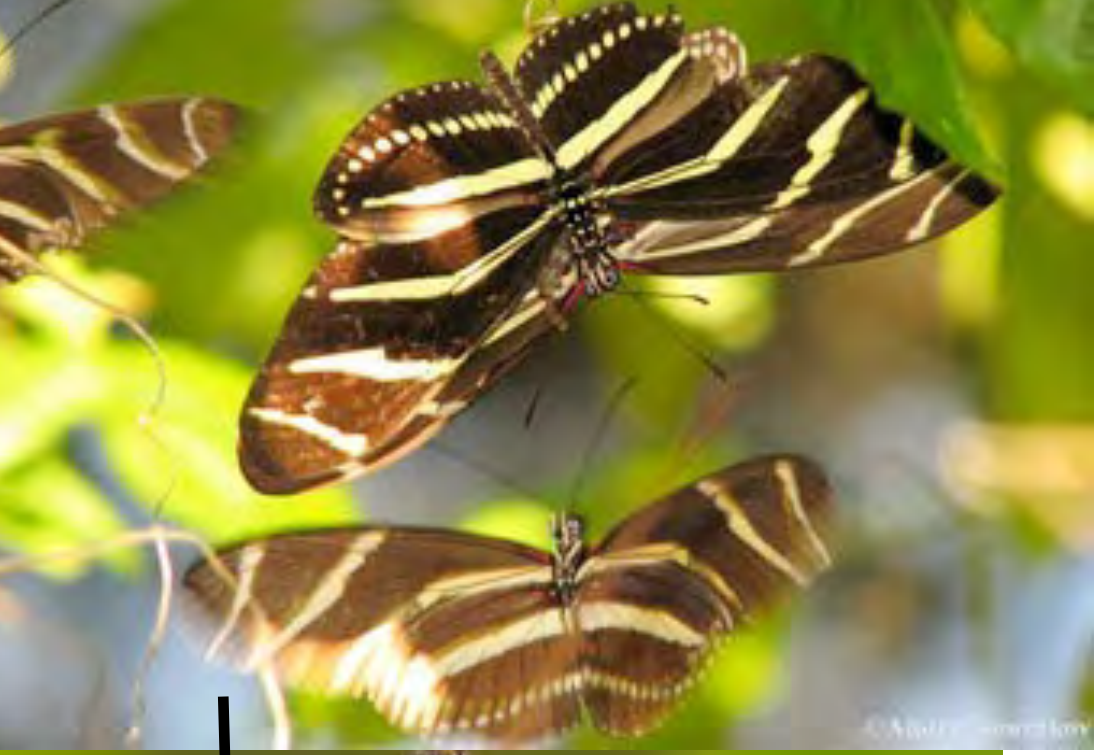
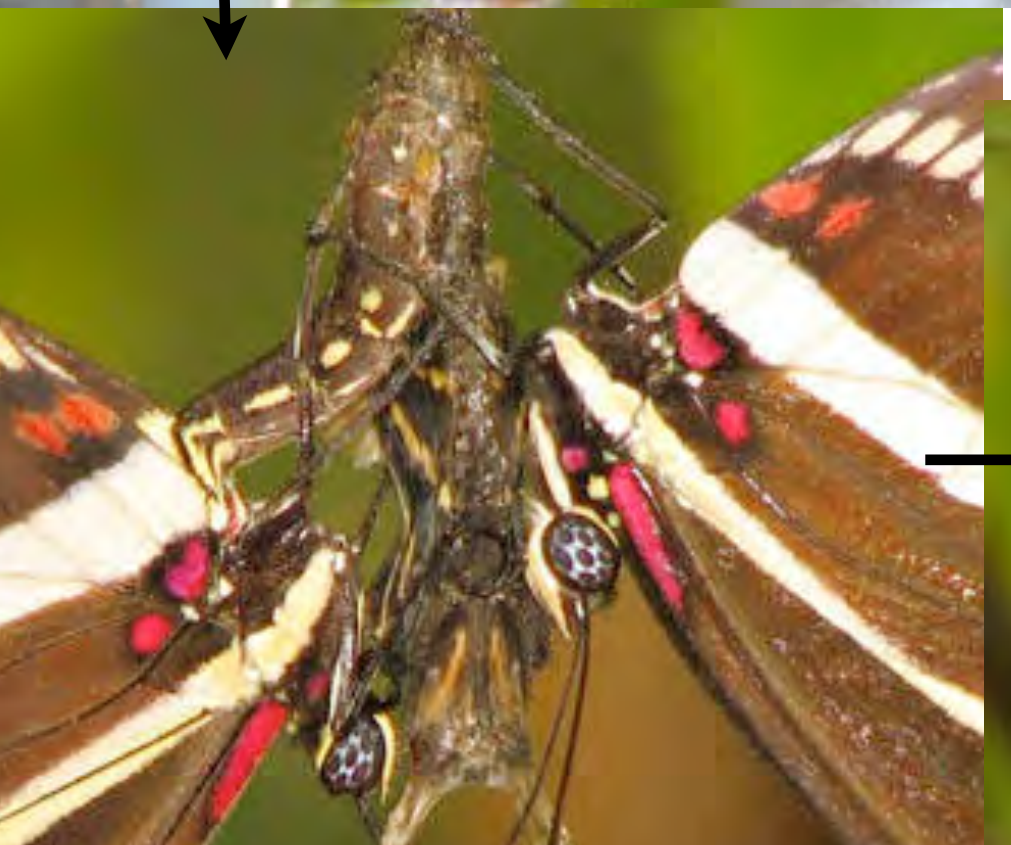


Photo credit
Andrei Sourakov





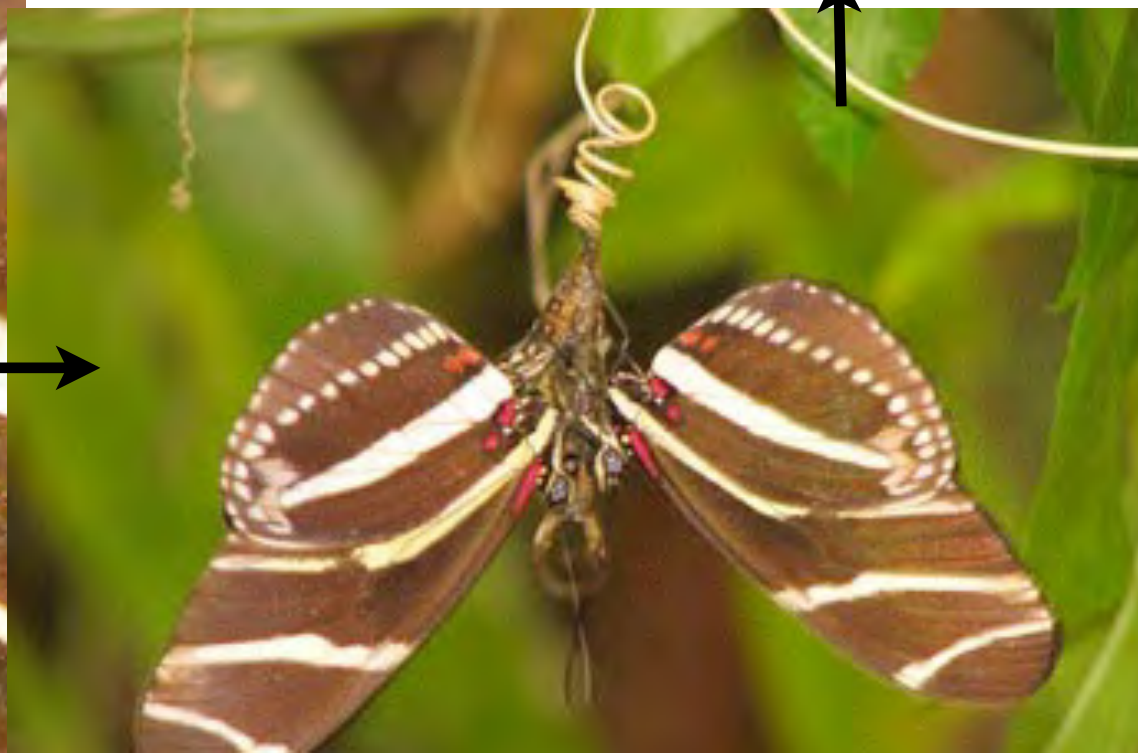
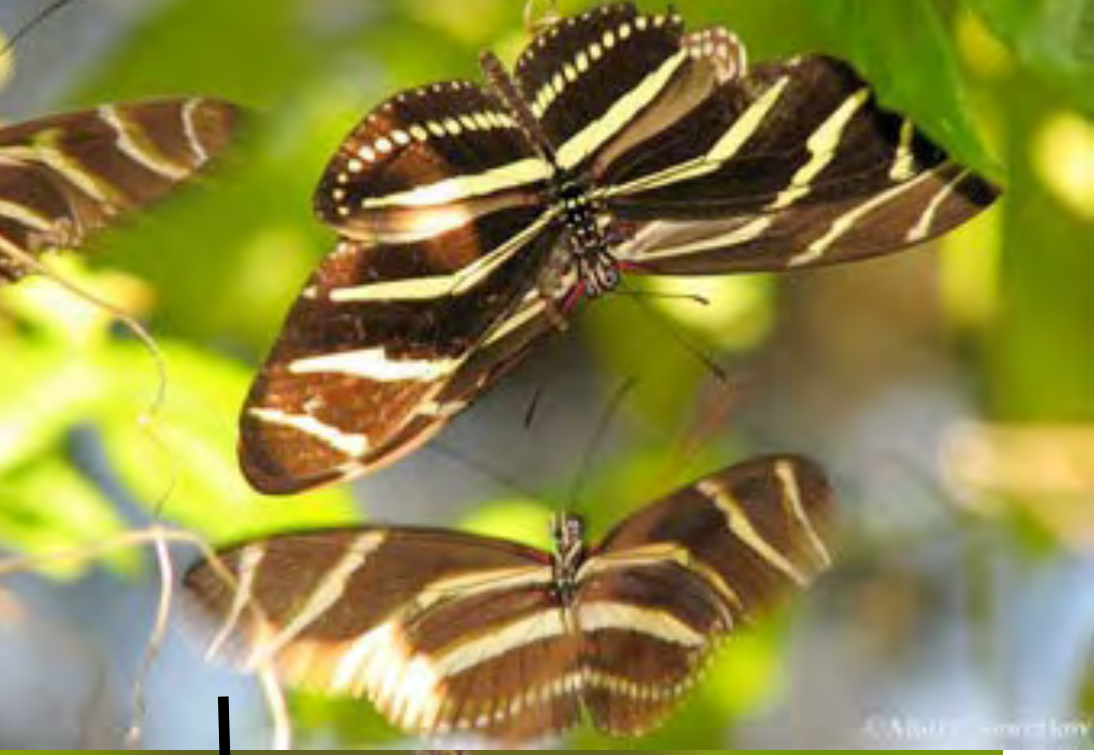
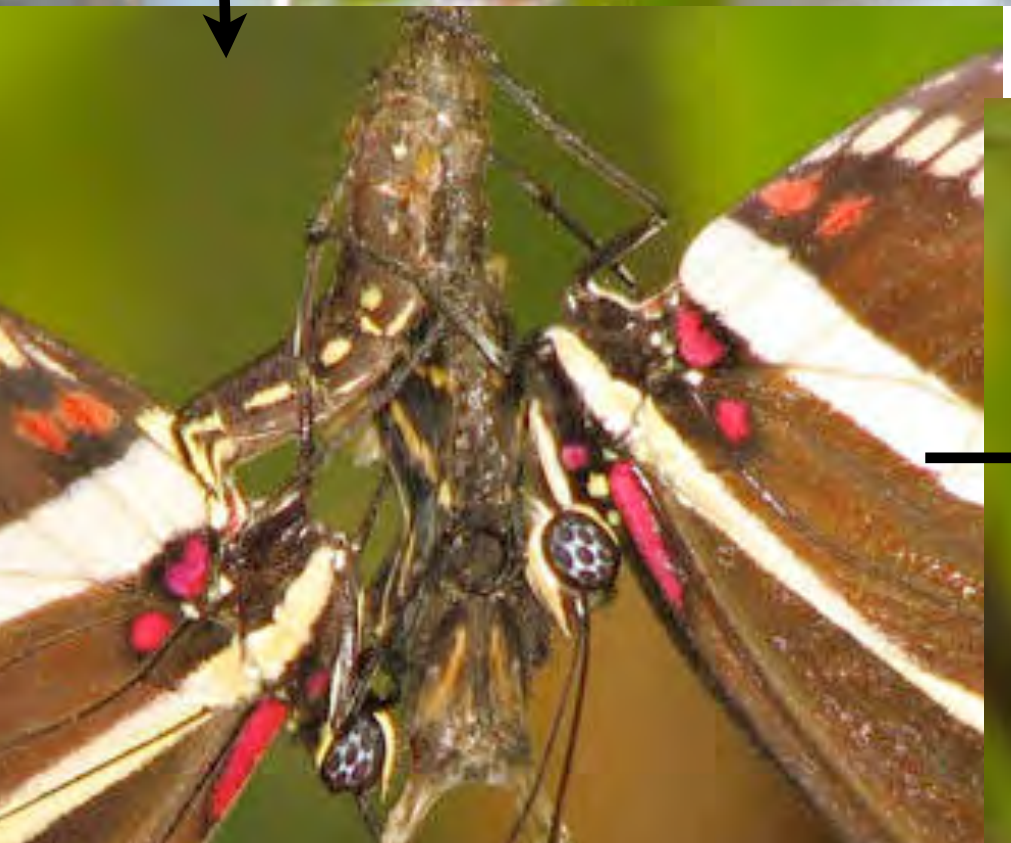
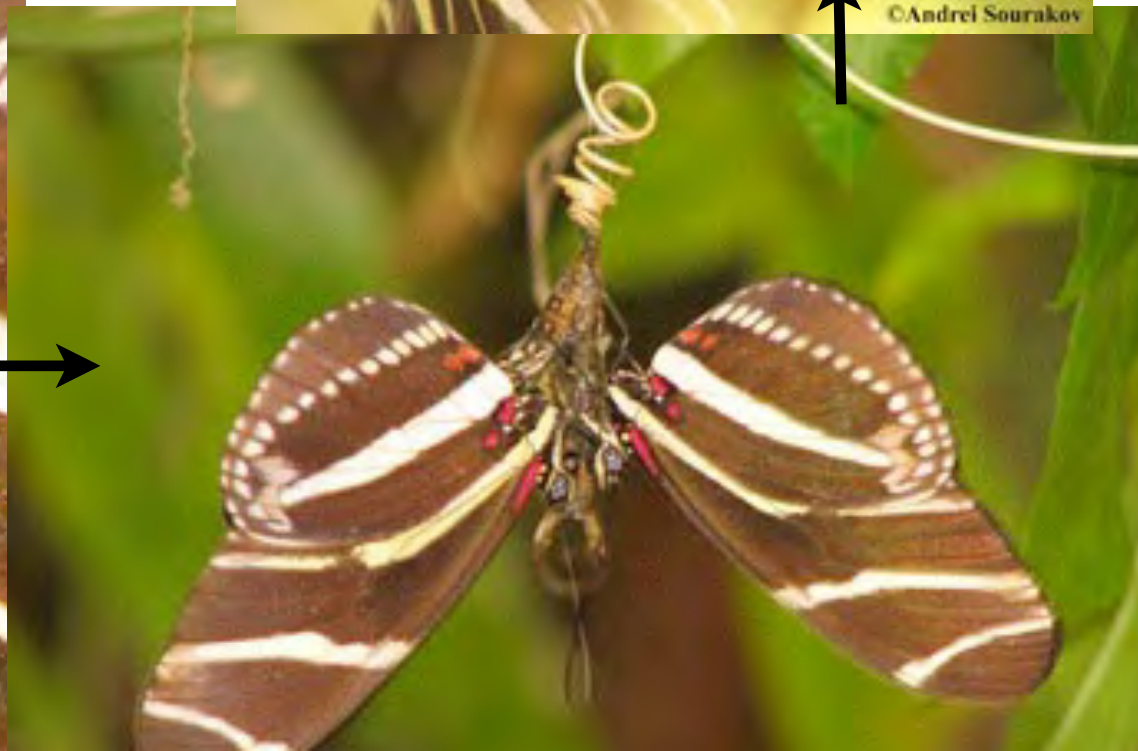


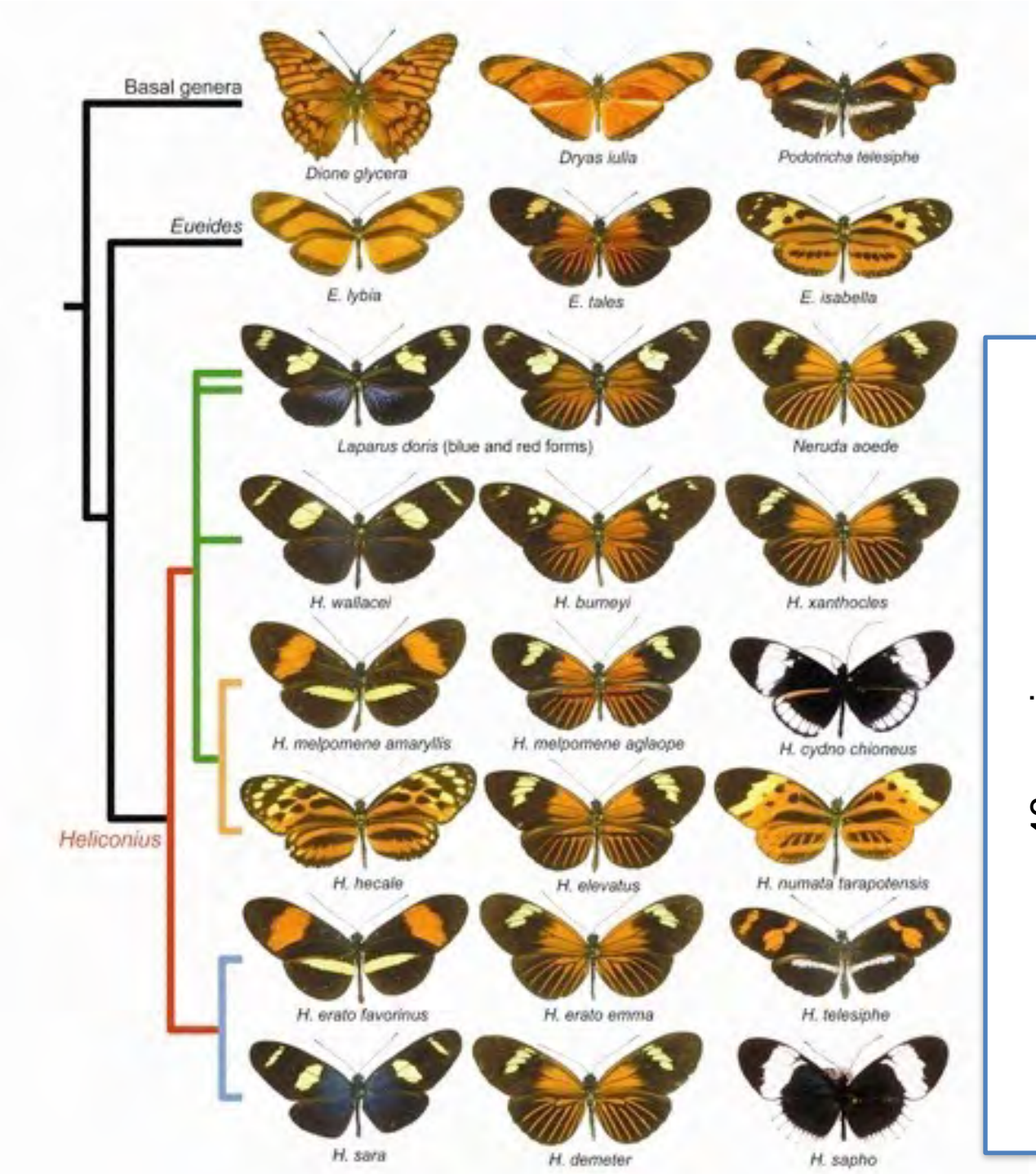
Photo credit
Andrei Sourakov




©Andrei Sourakov





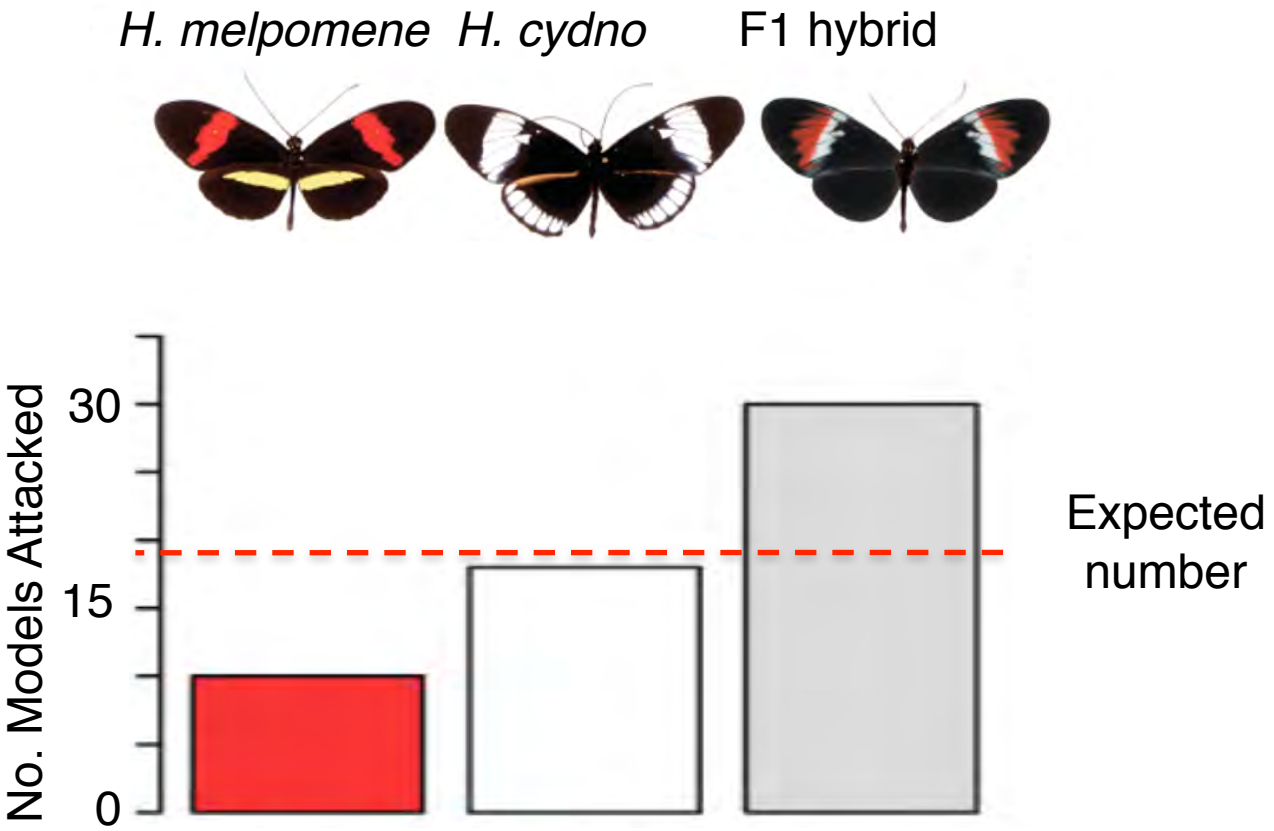



43 species

77 species

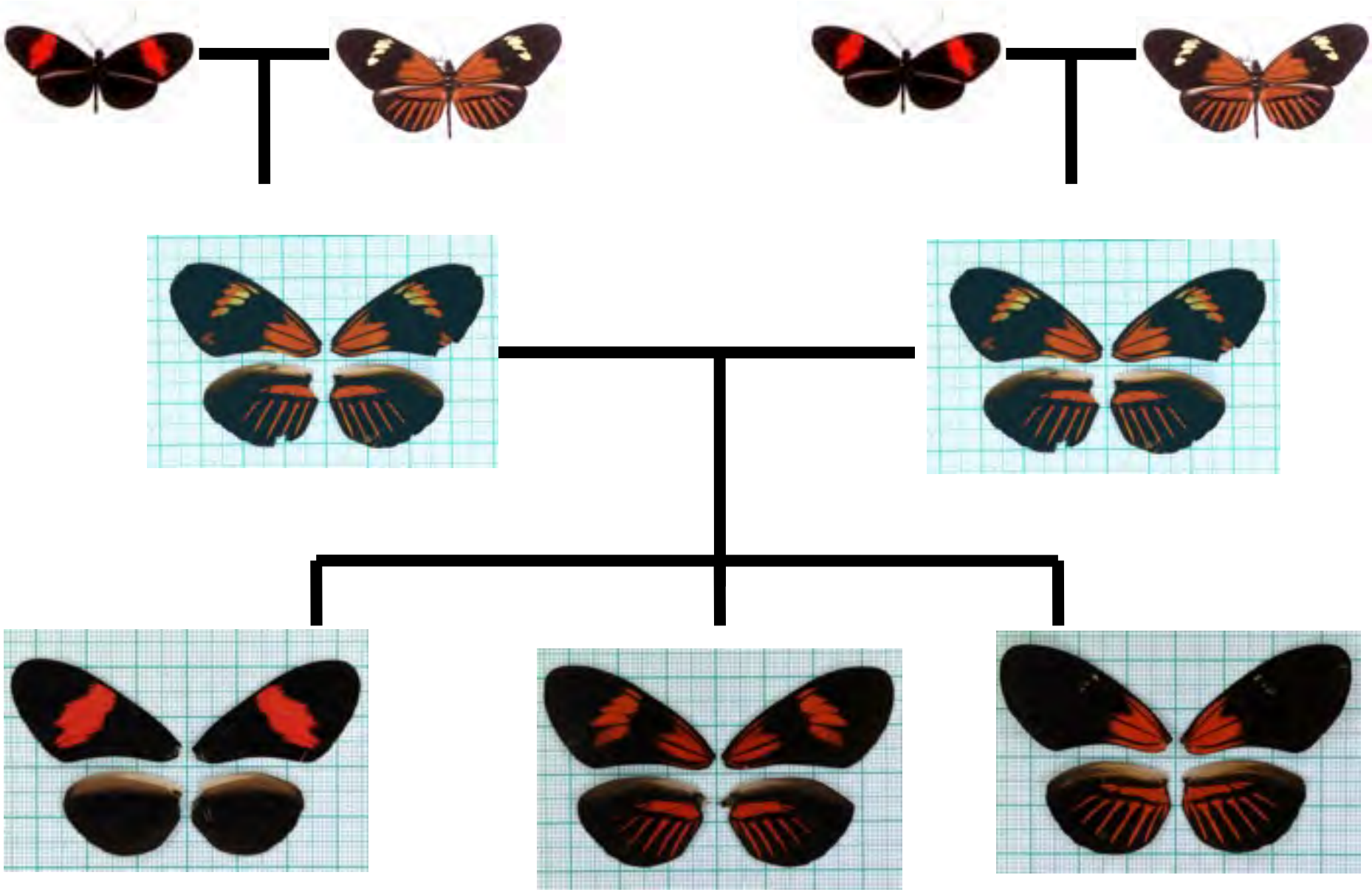
Rufous-tailed Jacamar

Fritz Müller

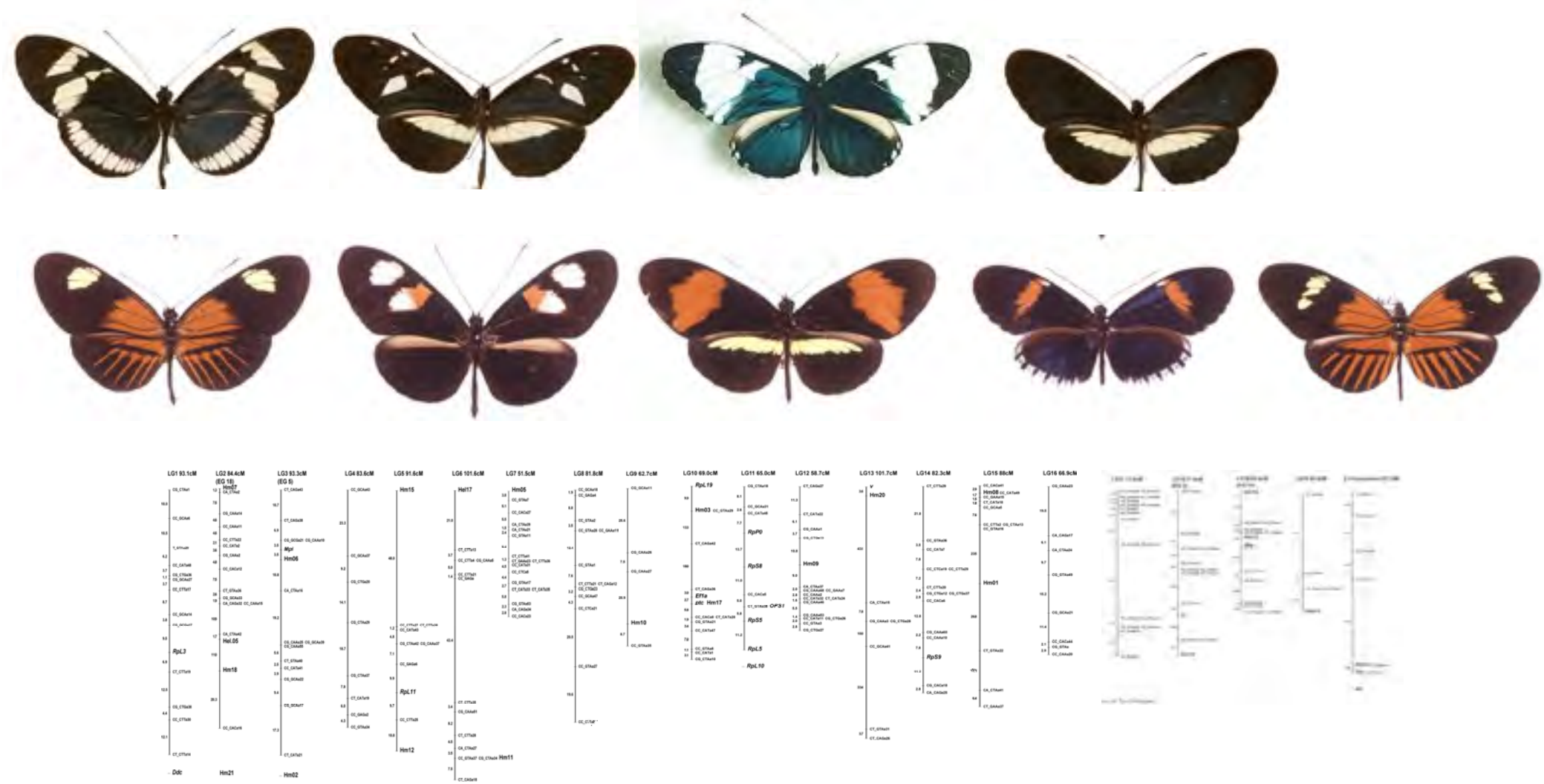



– *G*-test: $G = 7.25$, d.f. = 1, $p = 0.007$

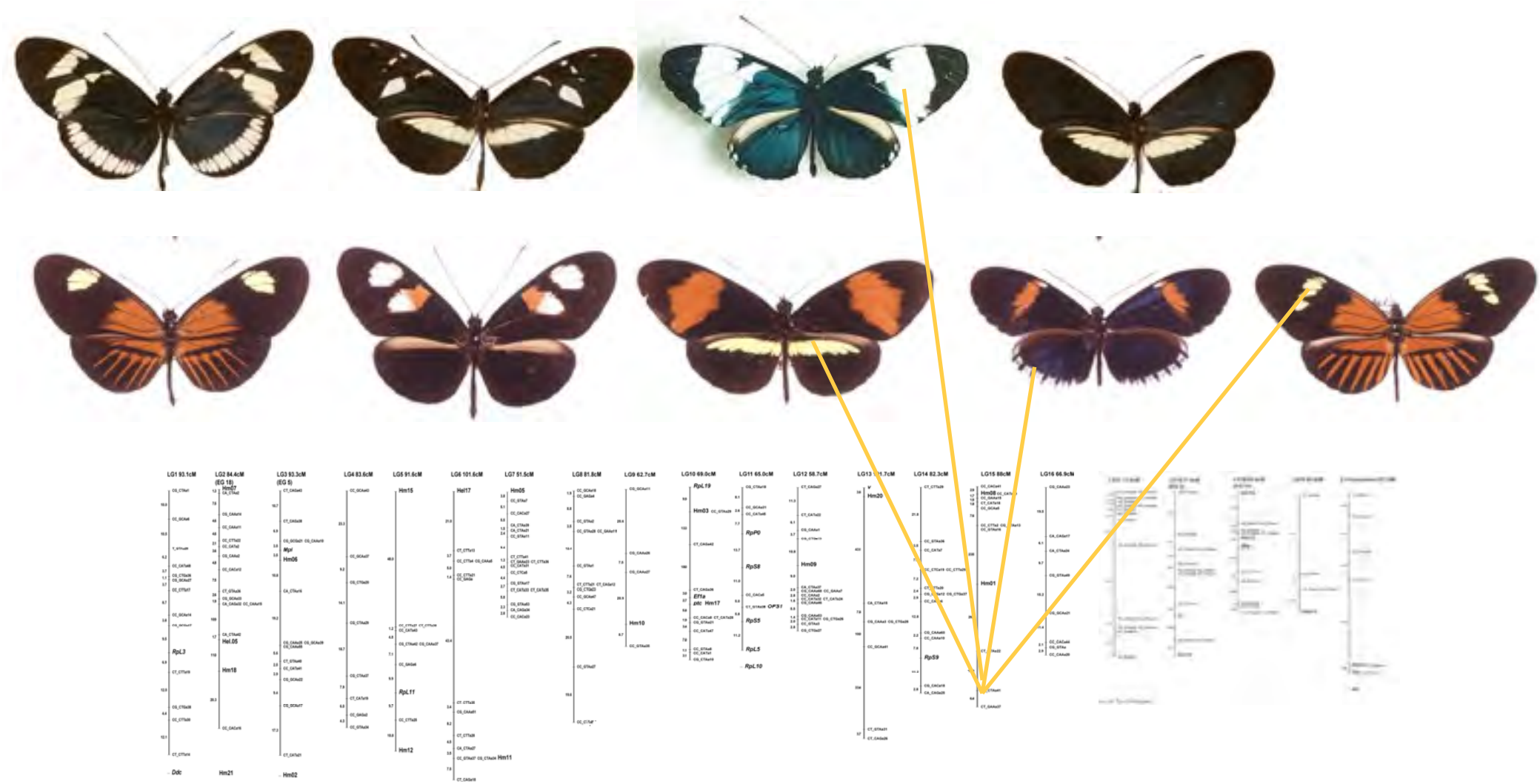
Merrill et al., Proc. Roy. Soc 2012


102
204
76

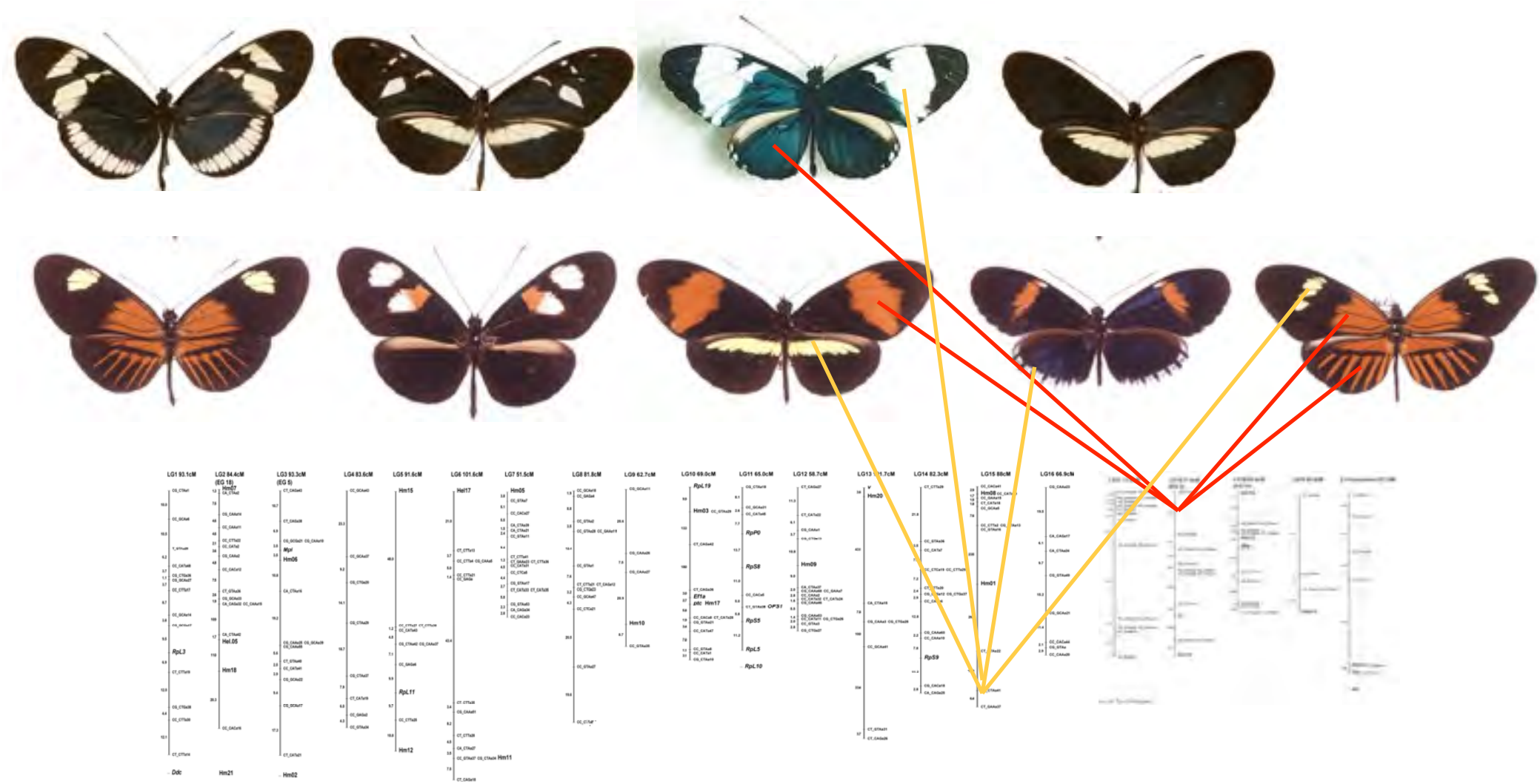
# Several major loci control *Heliconius* patterns



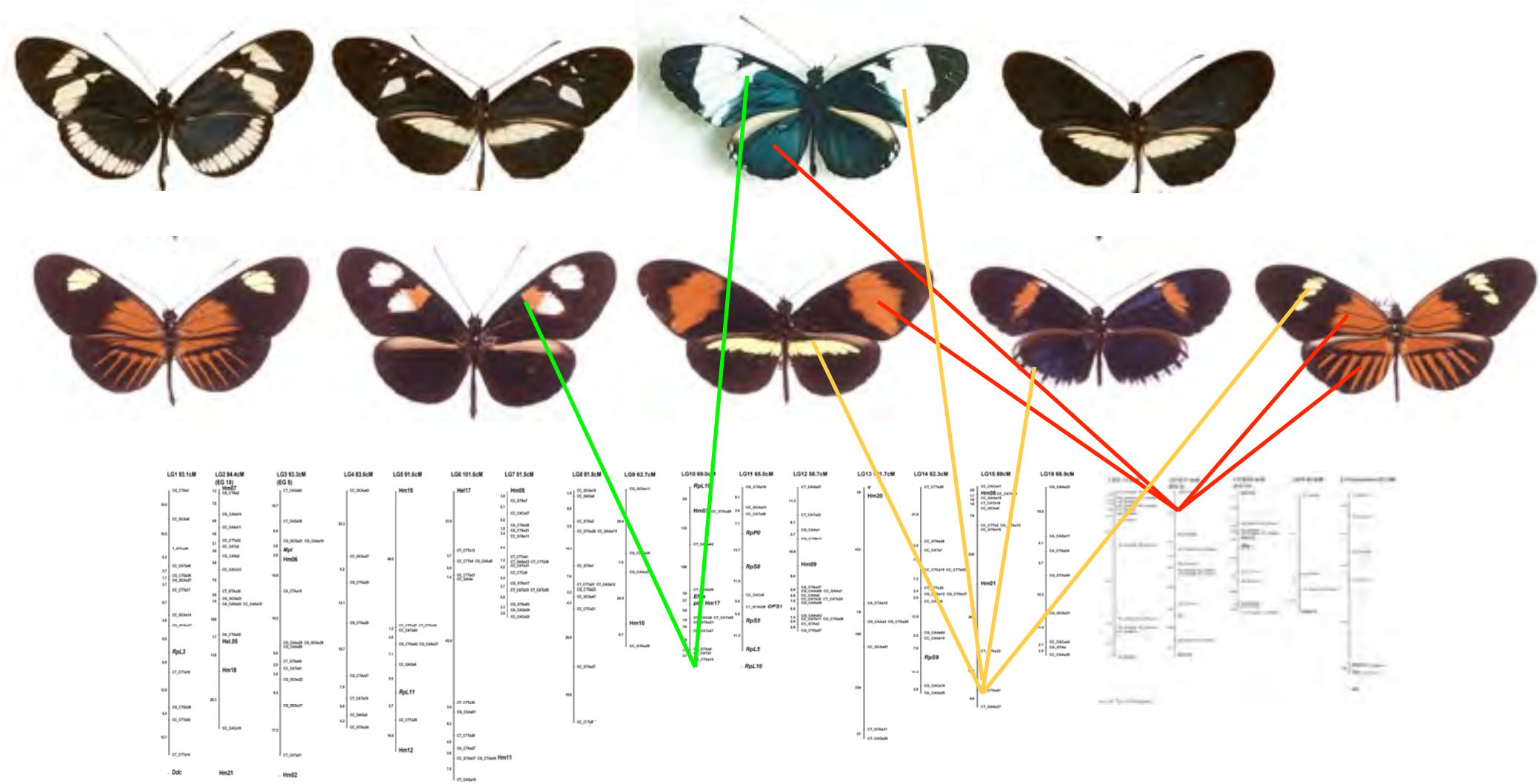
# Several major loci control *Heliconius* patterns



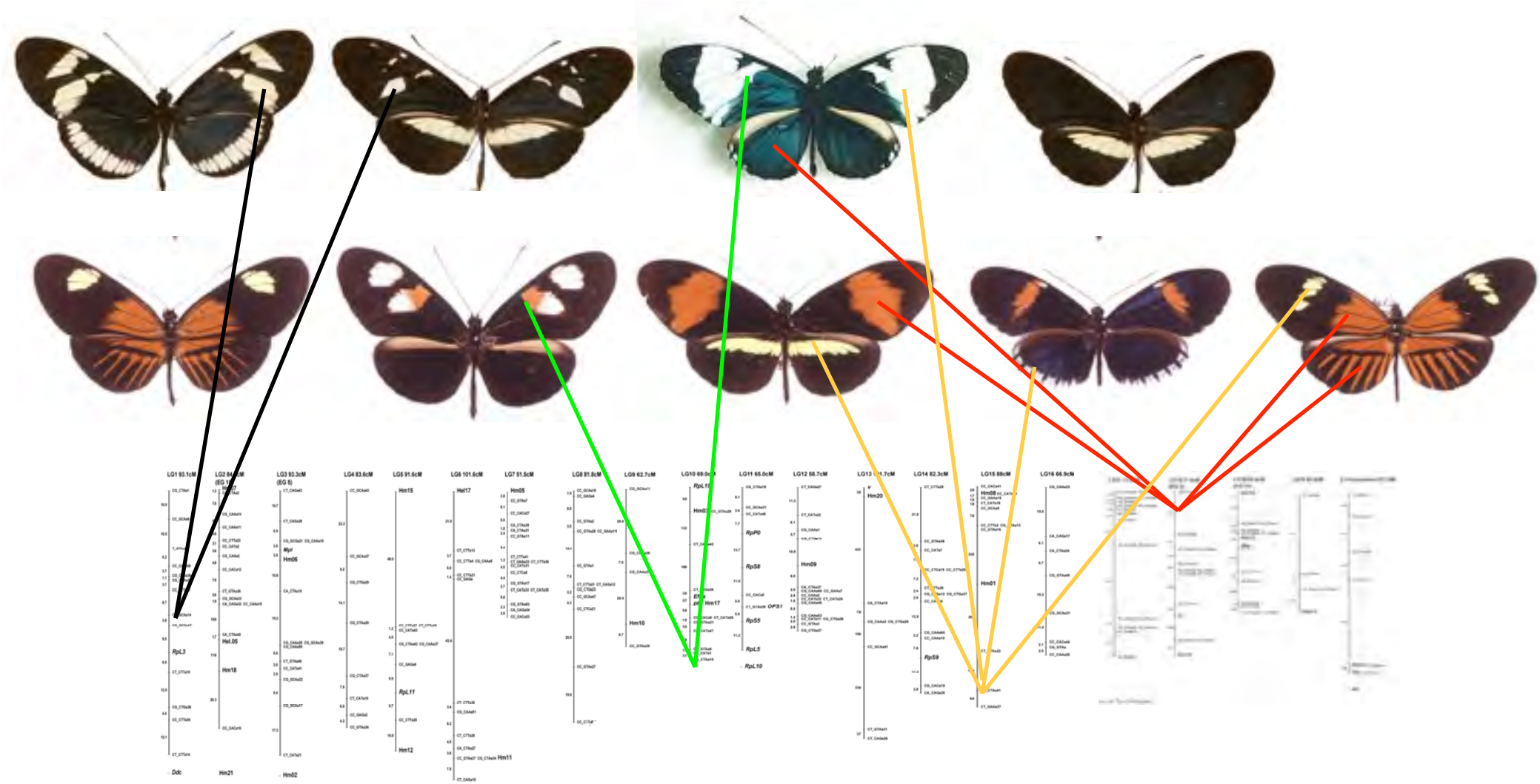
# Several major loci control *Heliconius* patterns



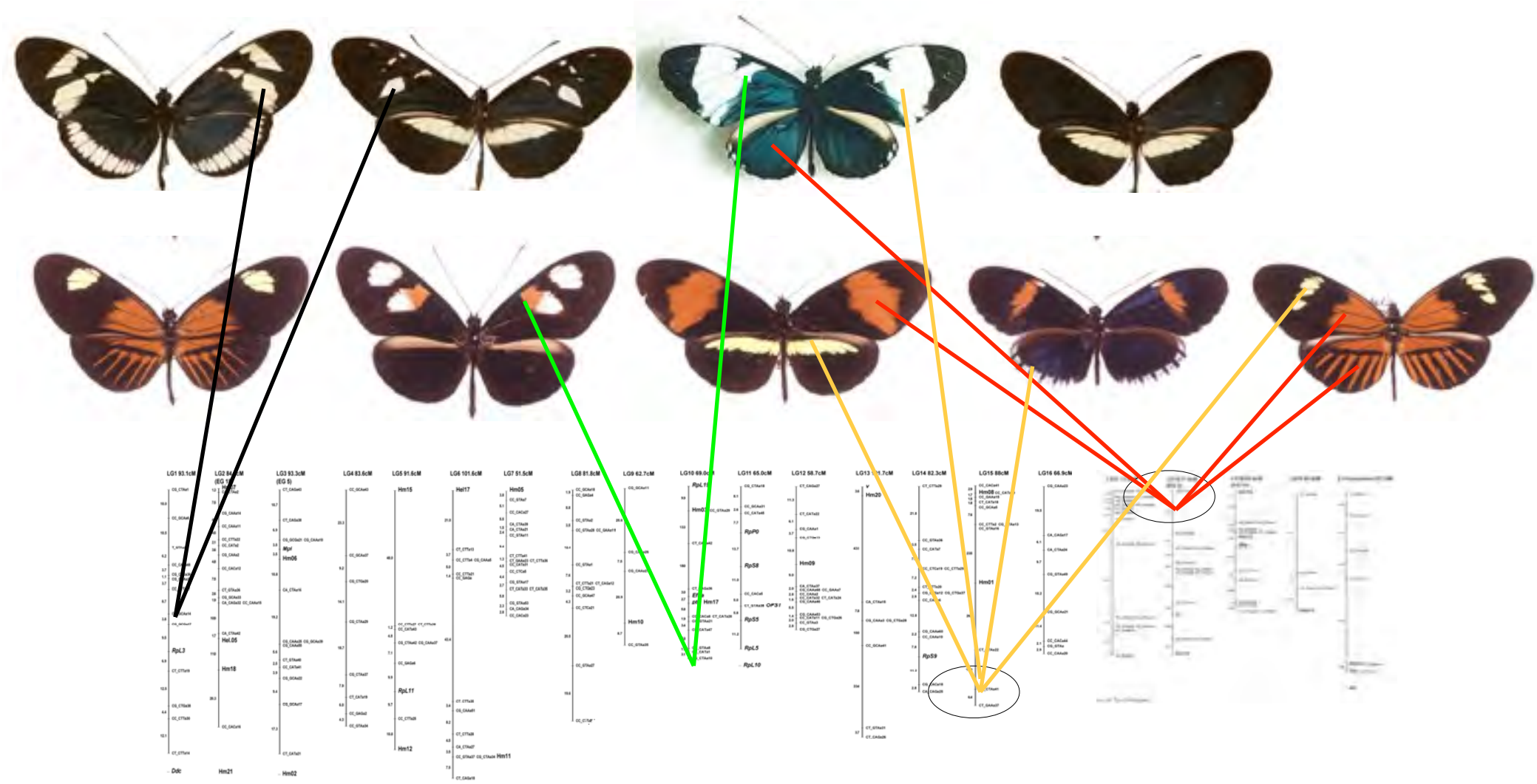
# Several major loci control *Heliconius* patterns

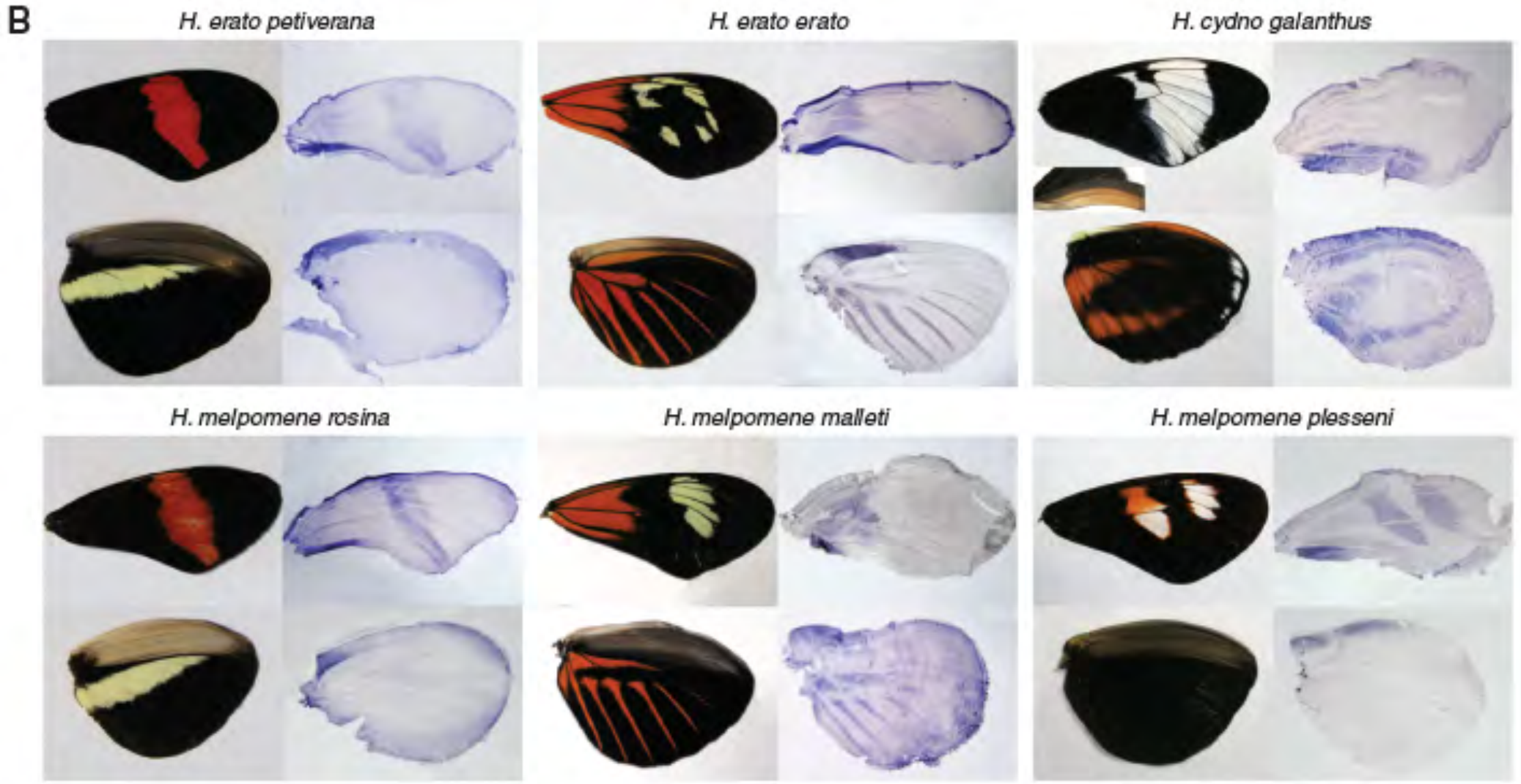


# Several major loci control *Heliconius* patterns

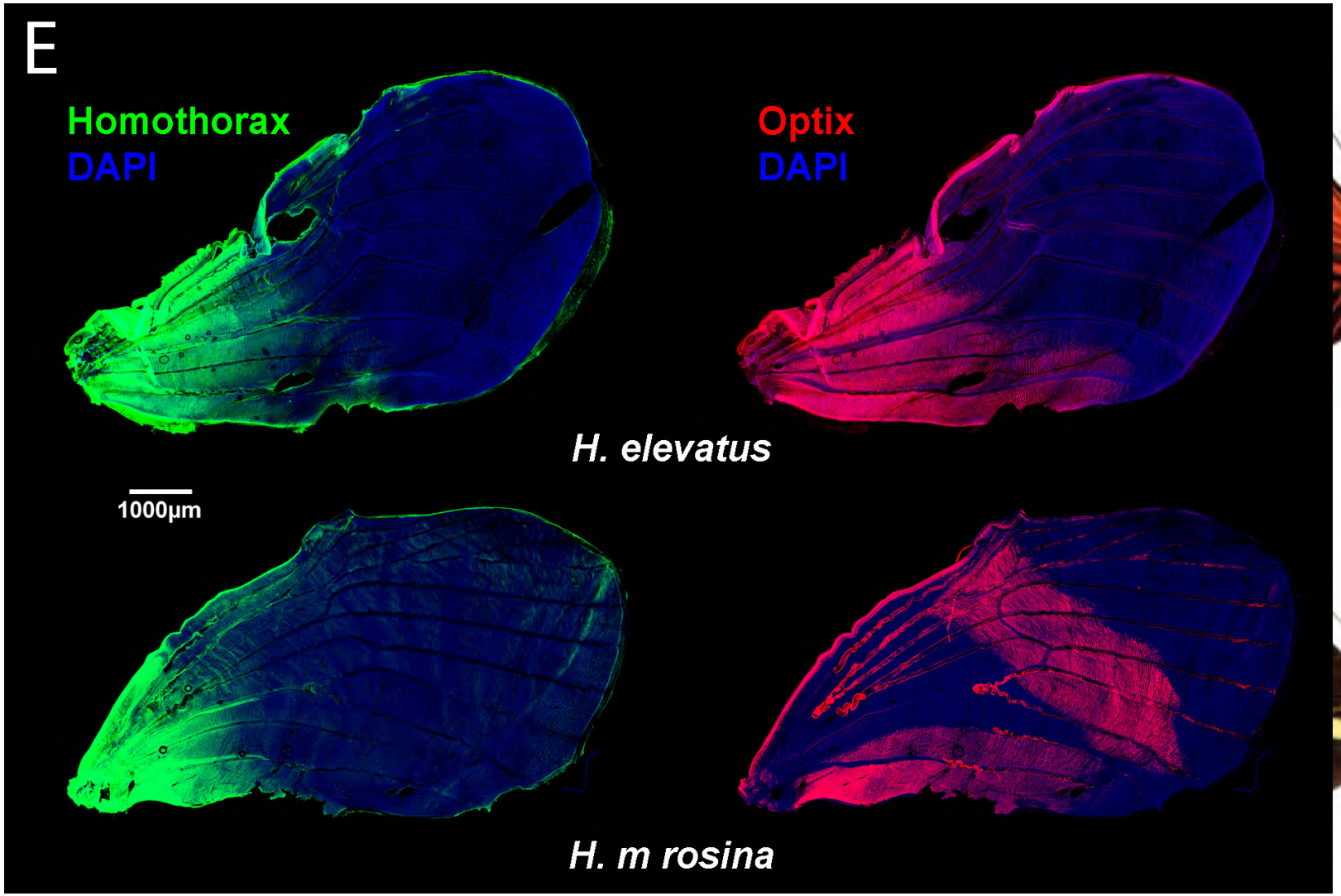


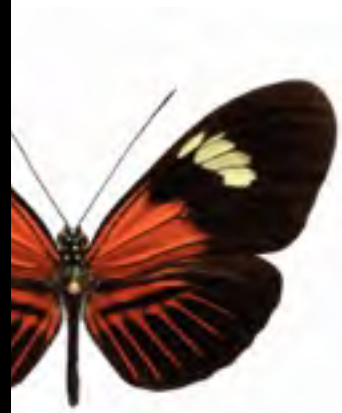
# Several major loci control *Heliconius* patterns

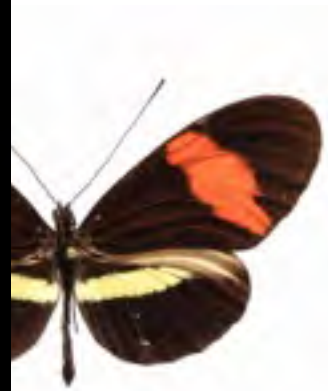





Reed et al., 2011 Science

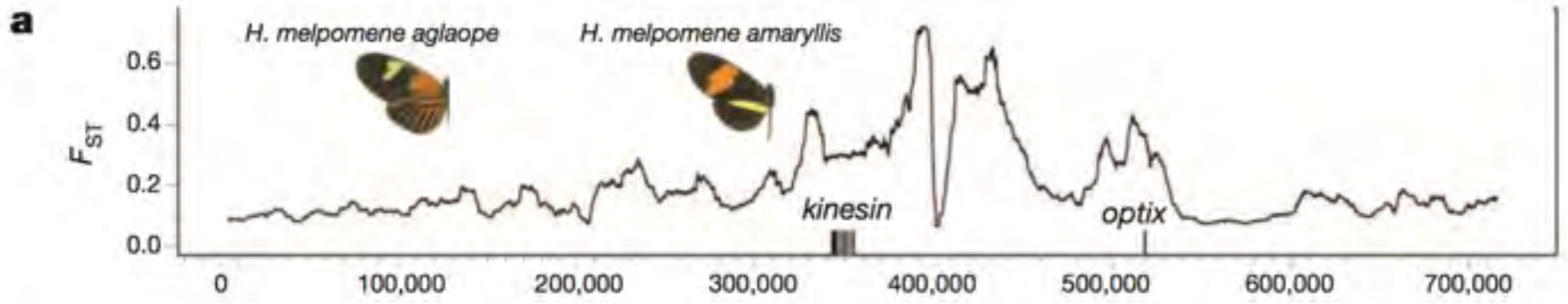




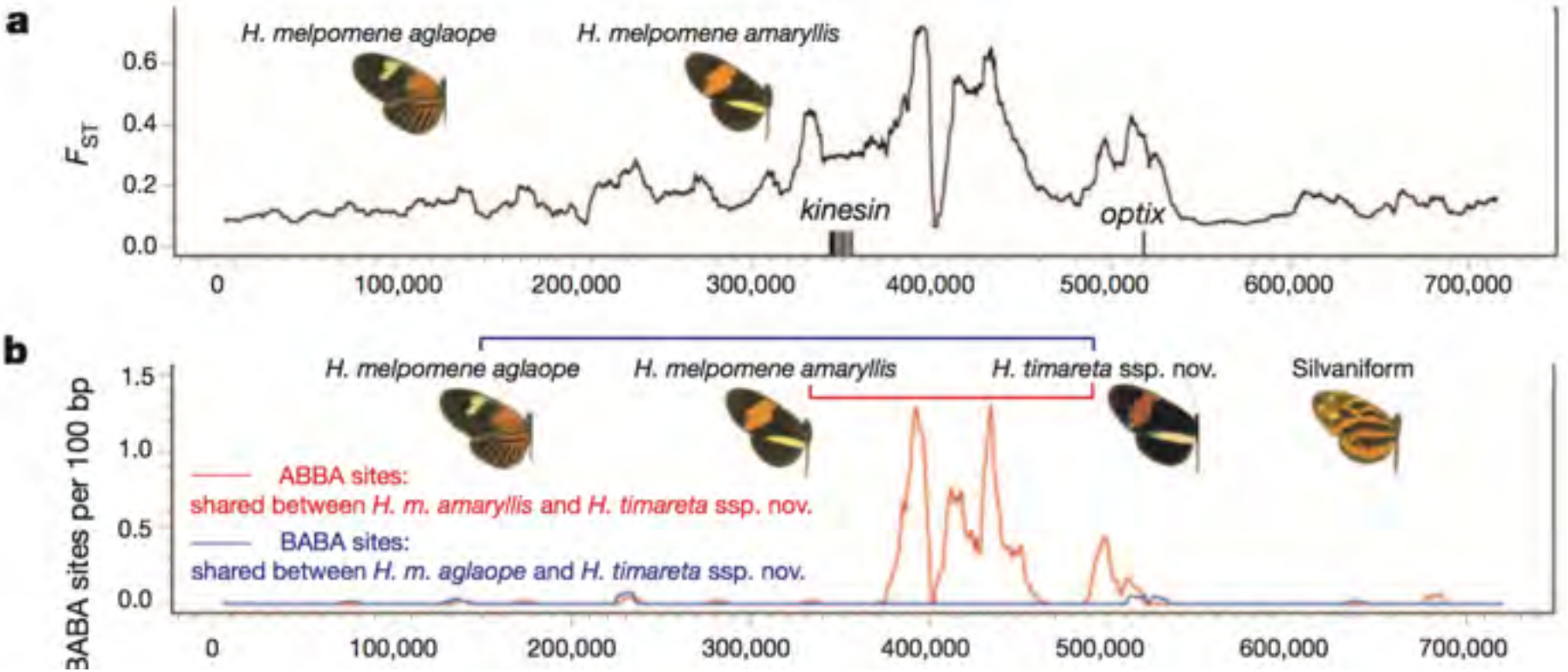
Richard Wallbank

# Cross-species sharing



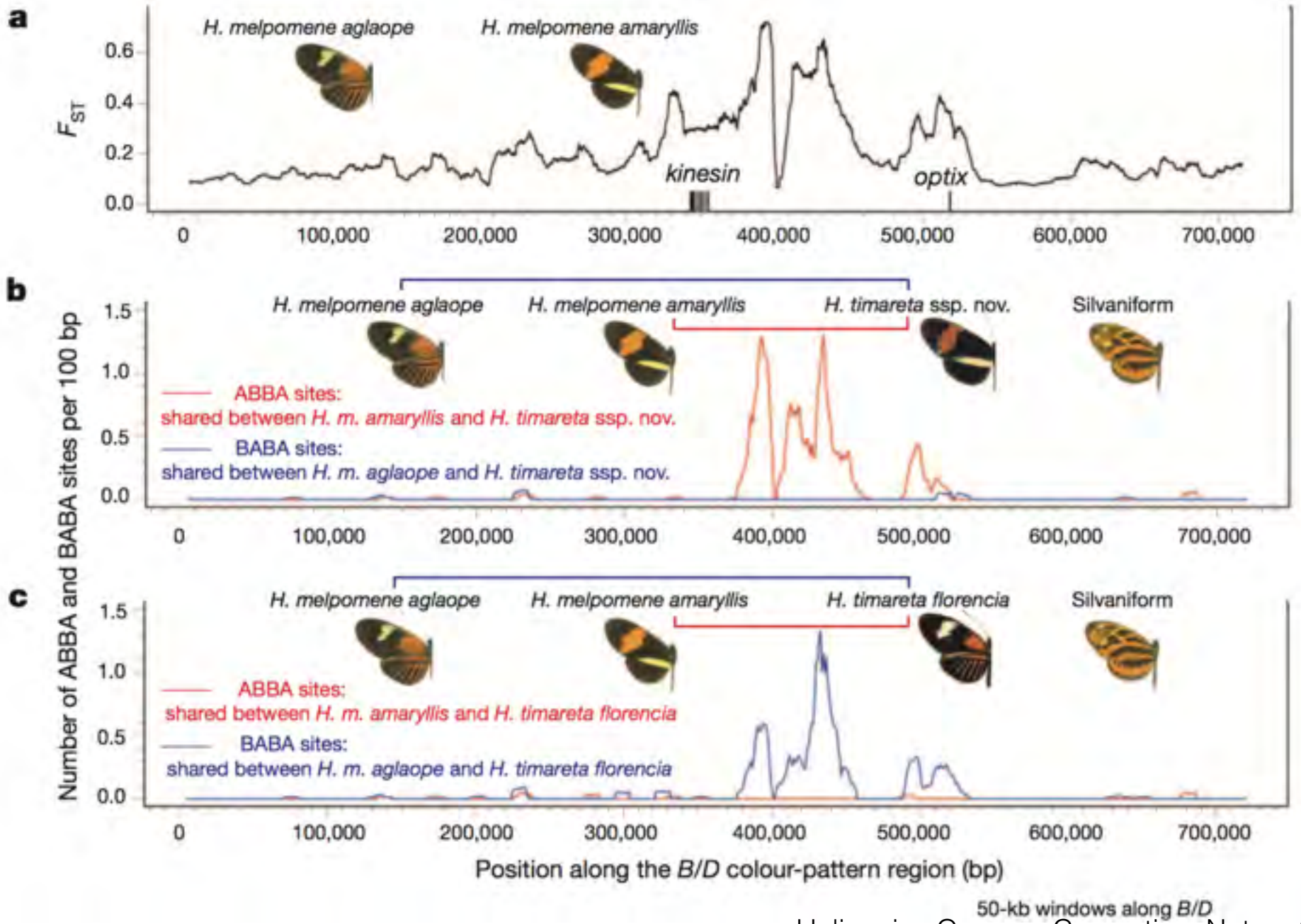

Heliconius Genome Consortium Nature 2012

# Cross-species sharing



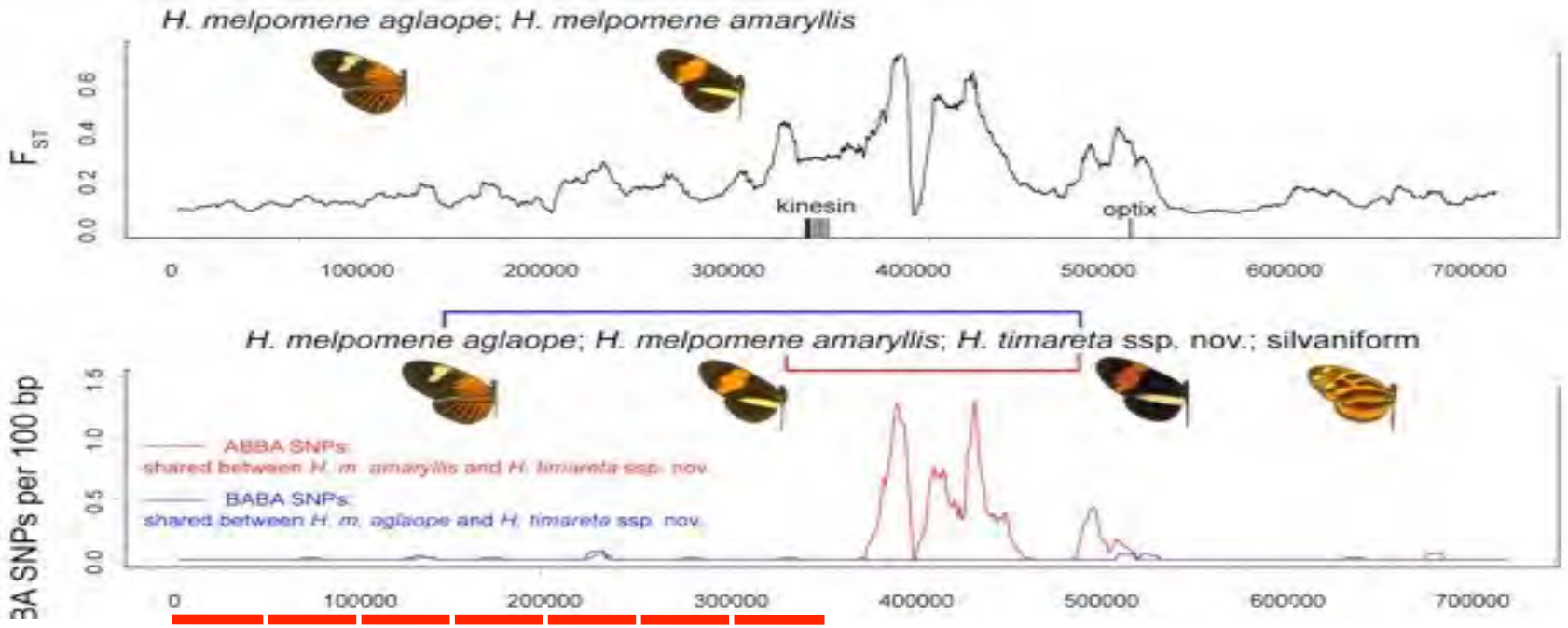

Heliconius Genome Consortium Nature 2012

# Cross-species sharing




Heliconius Genome Consortium Nature 2012

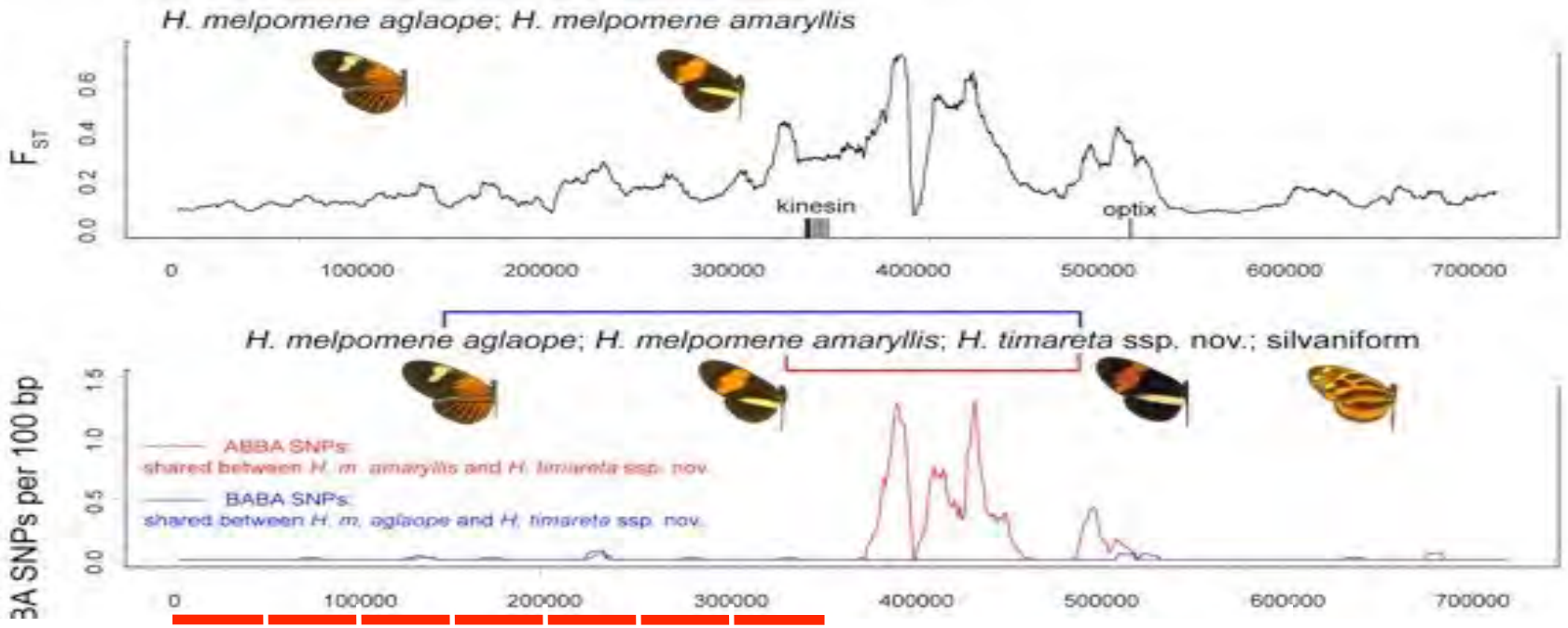
# Phylogenies across *B*/*D*



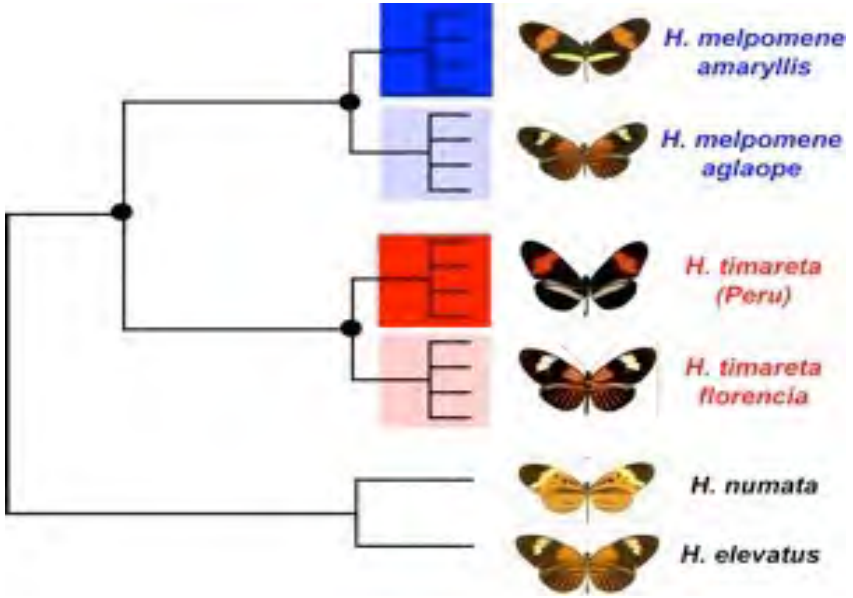

ML tree based
on
50,000 bp

Heliconius Genome Consortium Nature 2012

# Phylogenies across *B/D*

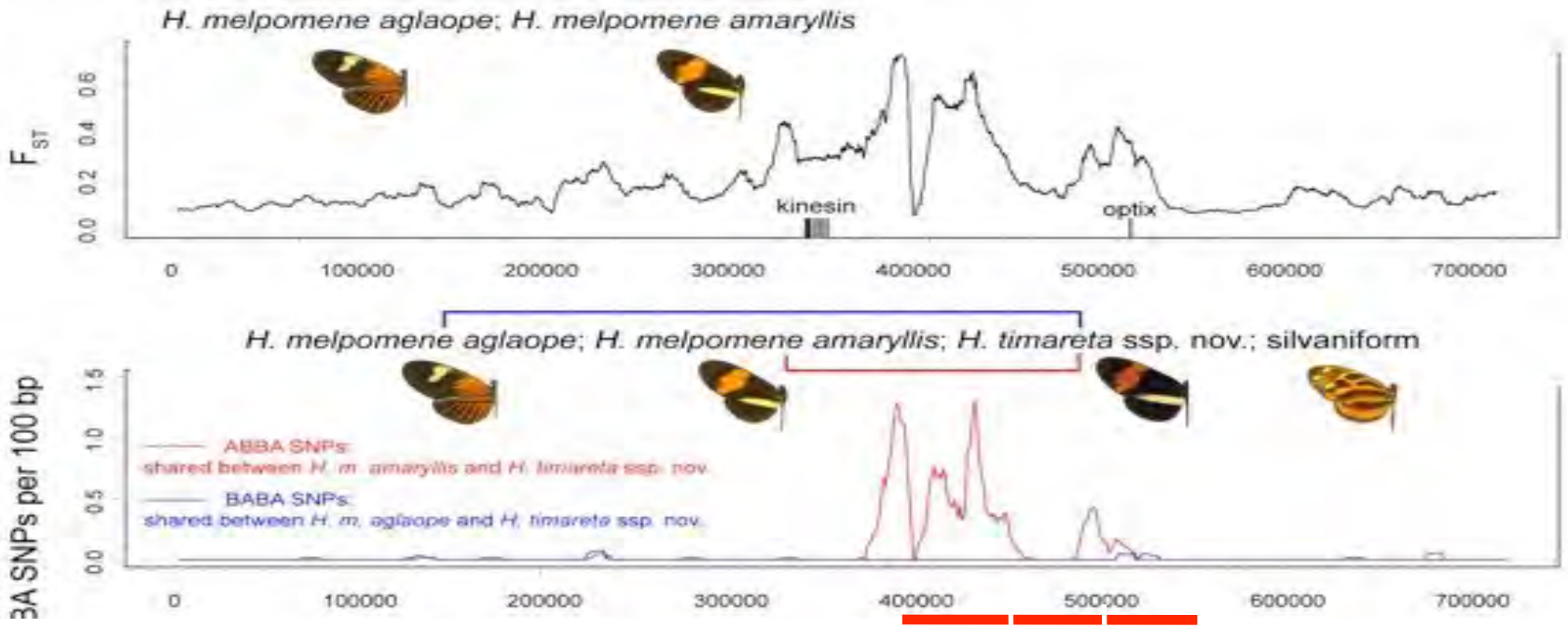



ML tree based on 50,000 bp



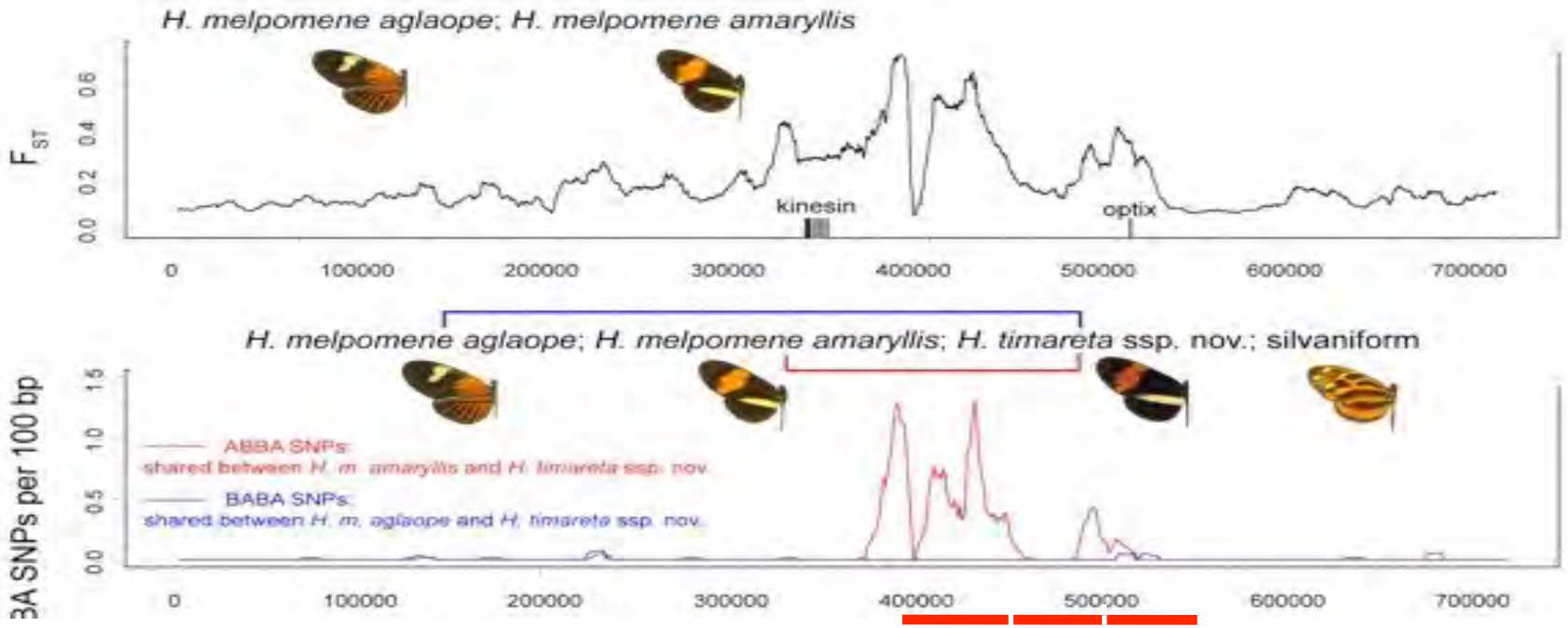

Heliconius Genome Consortium Nature 2012

# Phylogenies across *B*/*D*

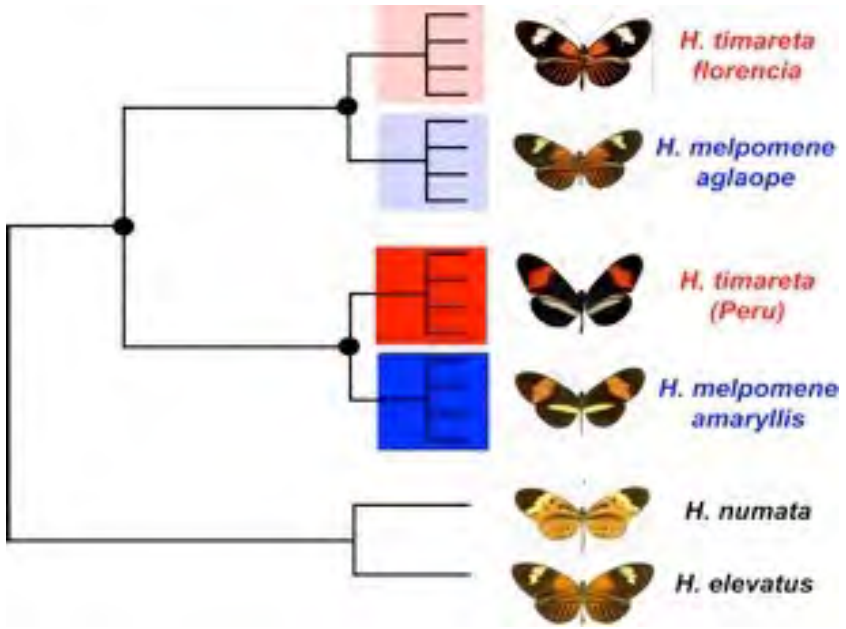



ML tree based
on
50,000 bp

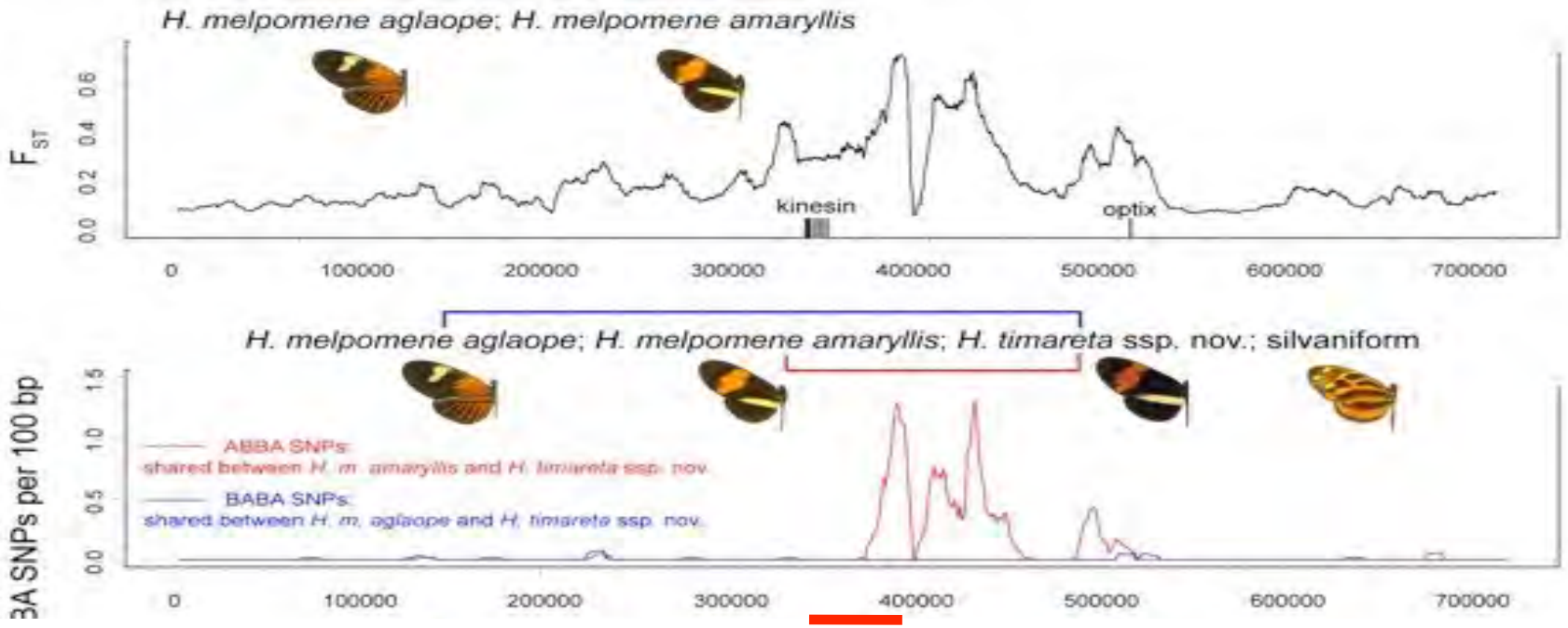
# Phylogenies across *B/D*




ML tree based on 50,000 bp

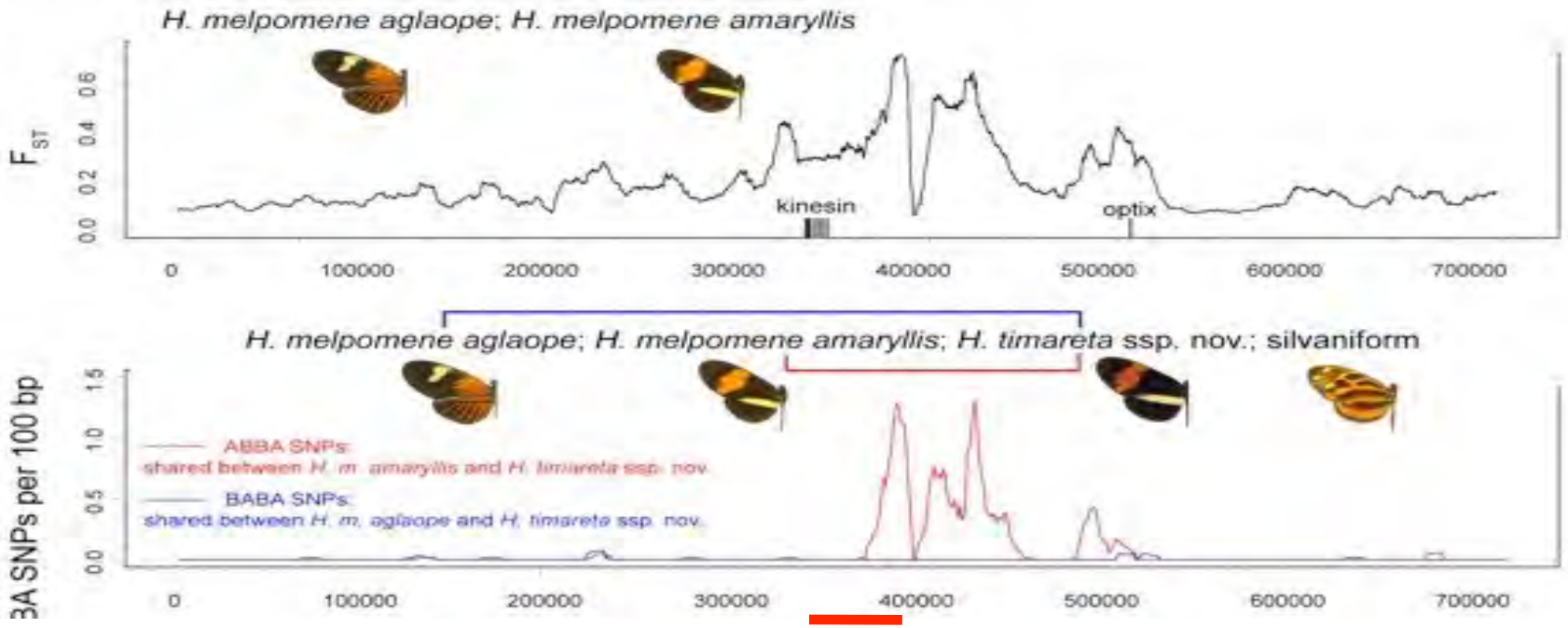


# Phylogenies across *B/D*

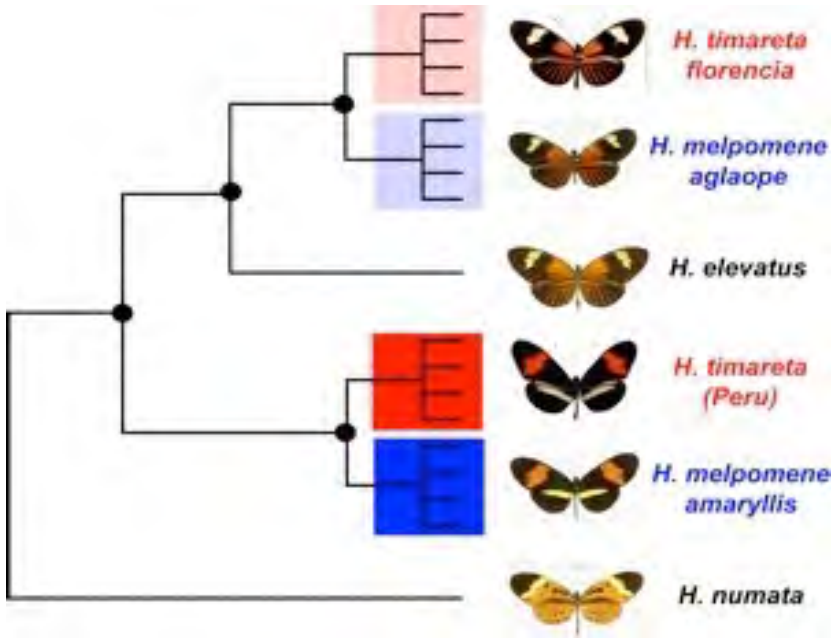



ML tree based
on
50,000 bp

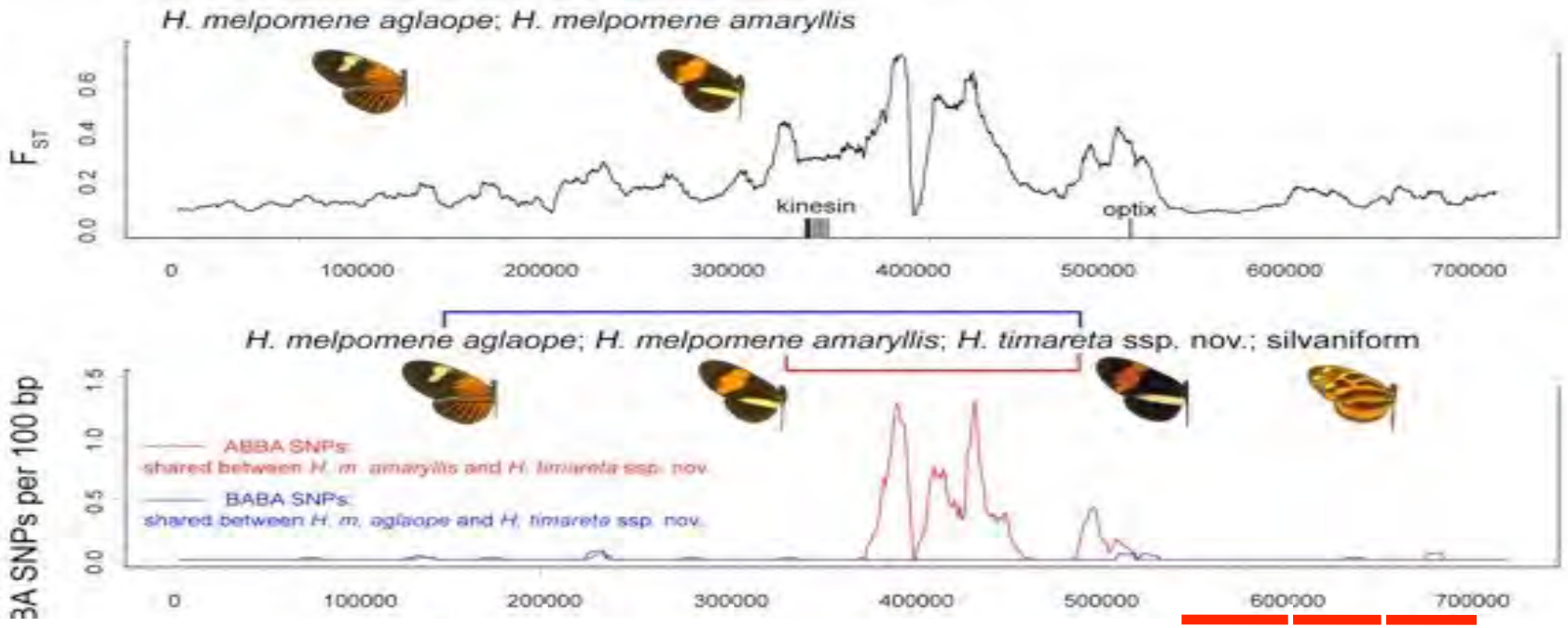
# Phylogenies across *B/D*




ML tree based on 50,000 bp

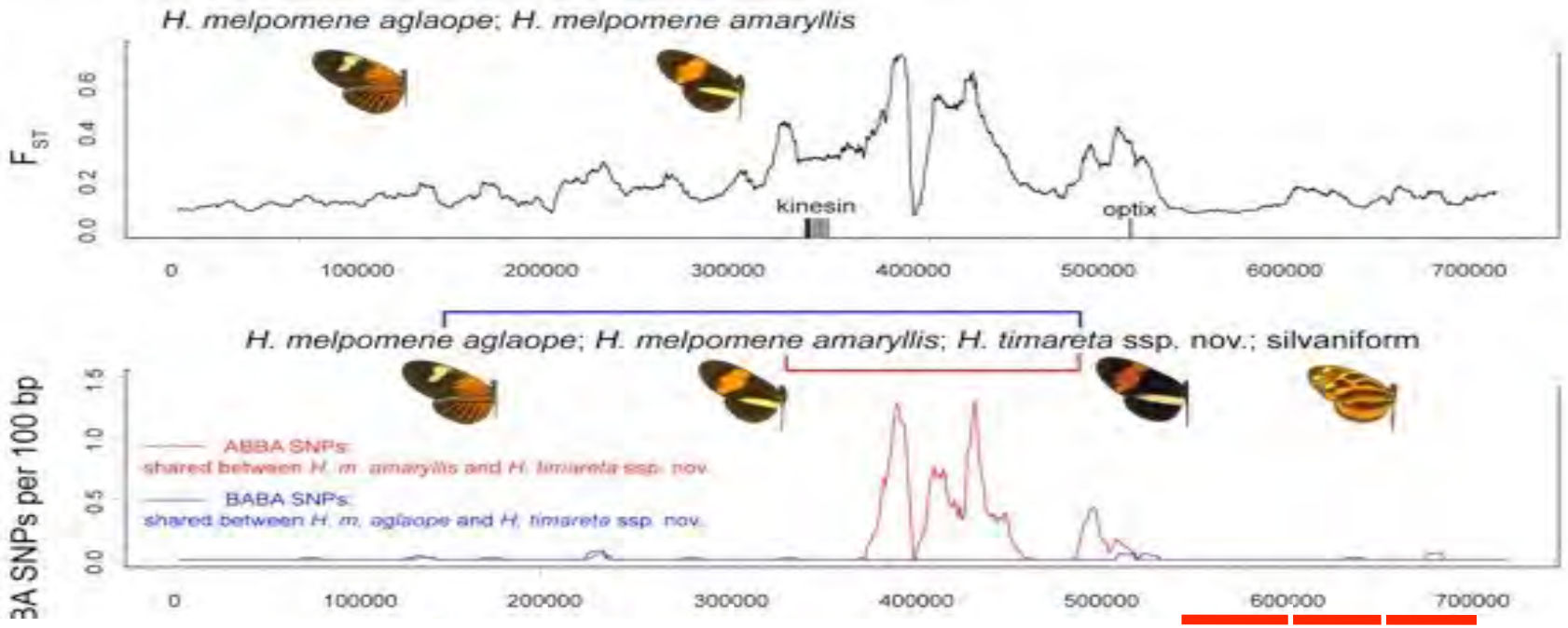


# Phylogenies across *B/D*

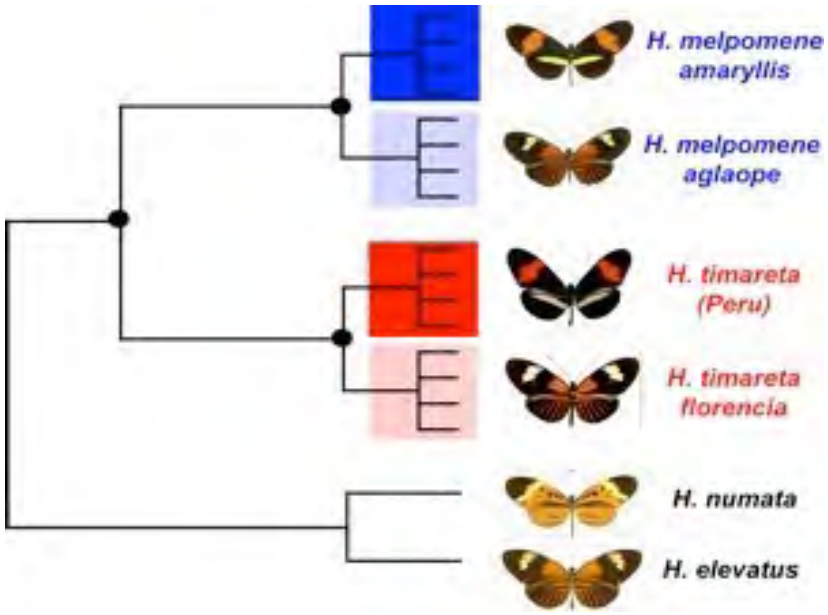



ML tree based
on
50,000 bp

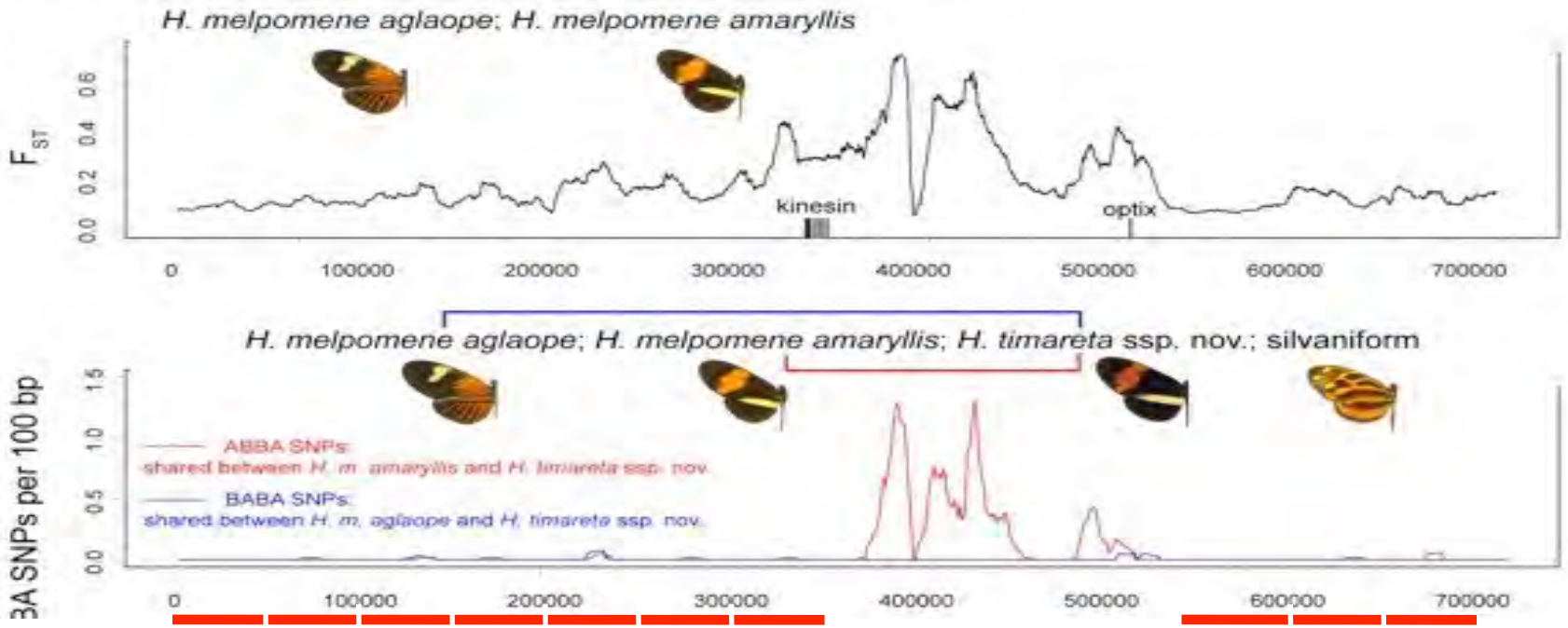
# Phylogenies across *B/D*




ML tree based on 50,000 bp

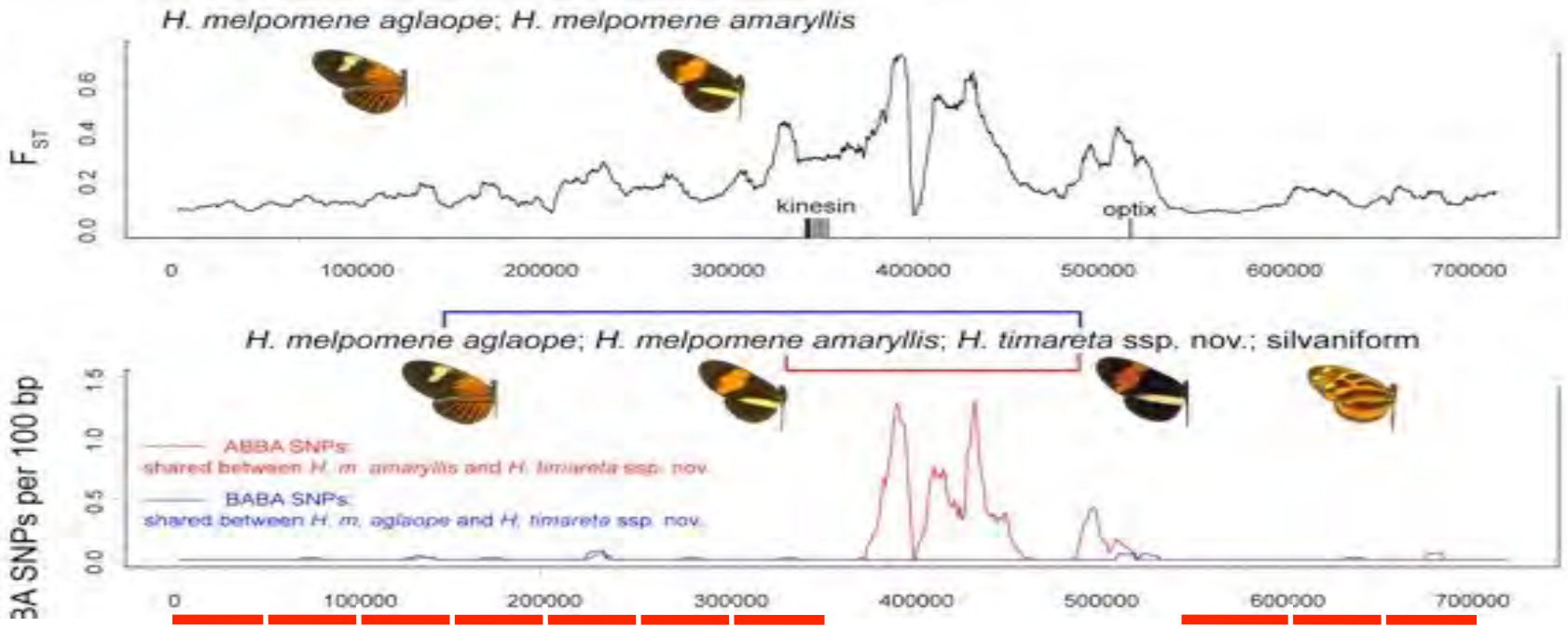


# Phylogenies across *B/D*

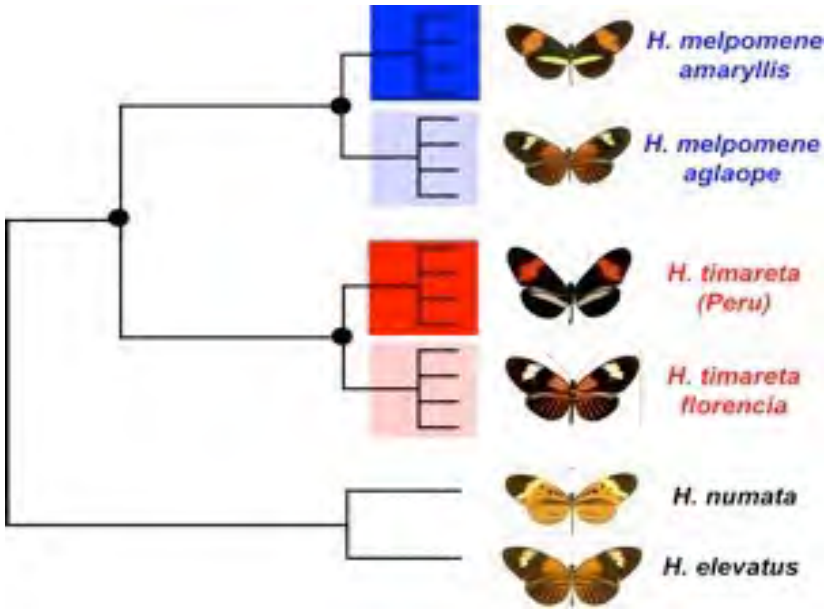



ML tree based
on
50,000 bp

# Phylogenies across *B/D*




ML tree based on 50,000 bp



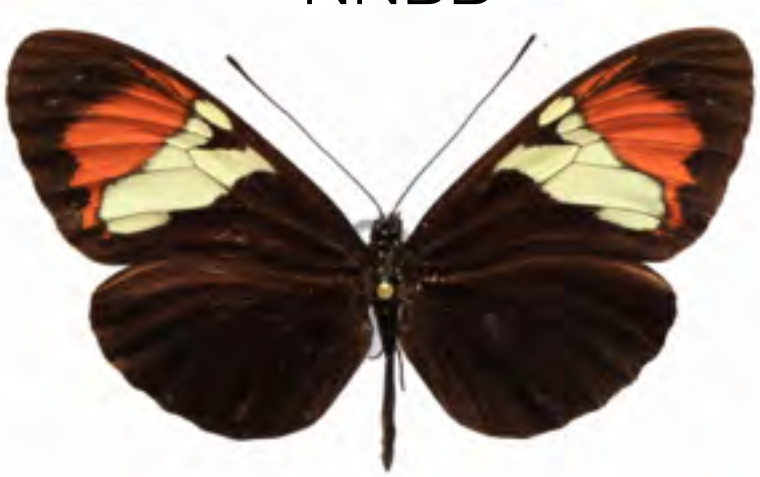
# Okay, so introgression causes mimicry

# Okay, so introgression causes mimicry

But mimicry is weird, right?

# Novelty can arise through introgression and recombination

NNBB



Heliconius heurippa

Camilo Salazar

Mavarez et al., Nature 2006

# Novelty can arise through introgression and recombination

NNBB

Heliconius heurippa

NNbb

Heliconius cydno cordula

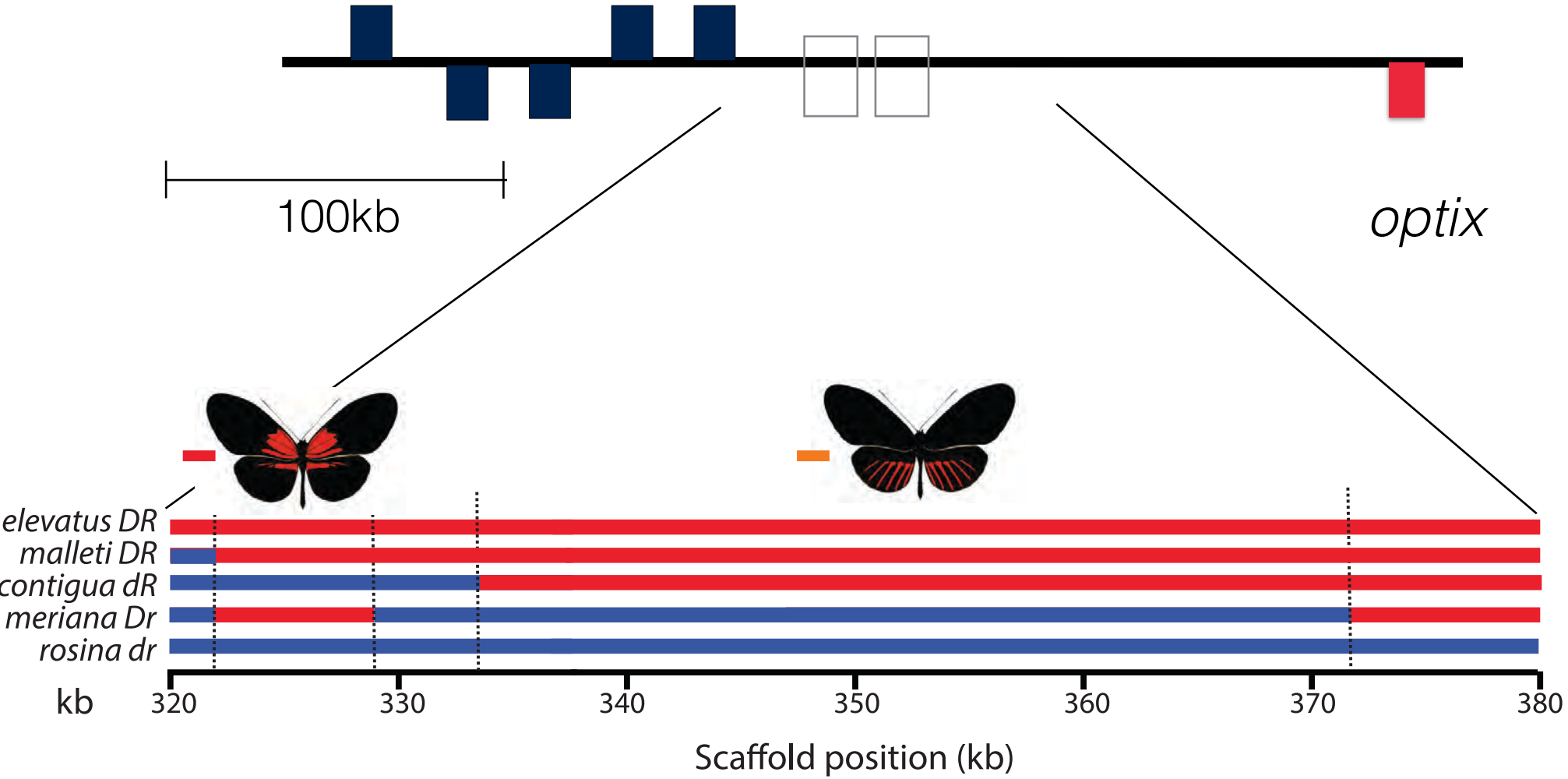
nnBB

Heliconius melpomene melpomene

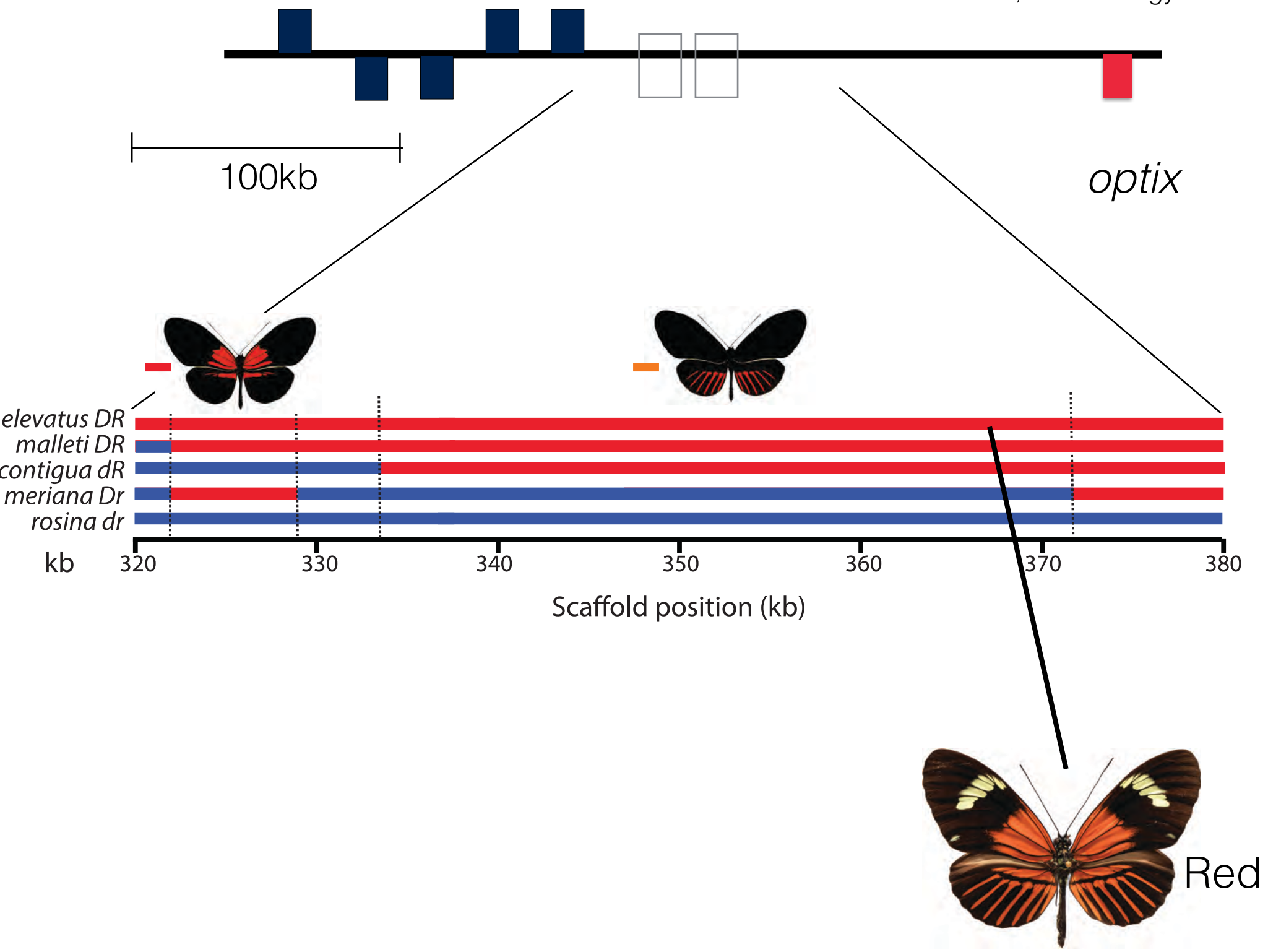
Camilo Salazar

Mavarez et al., Nature 2006

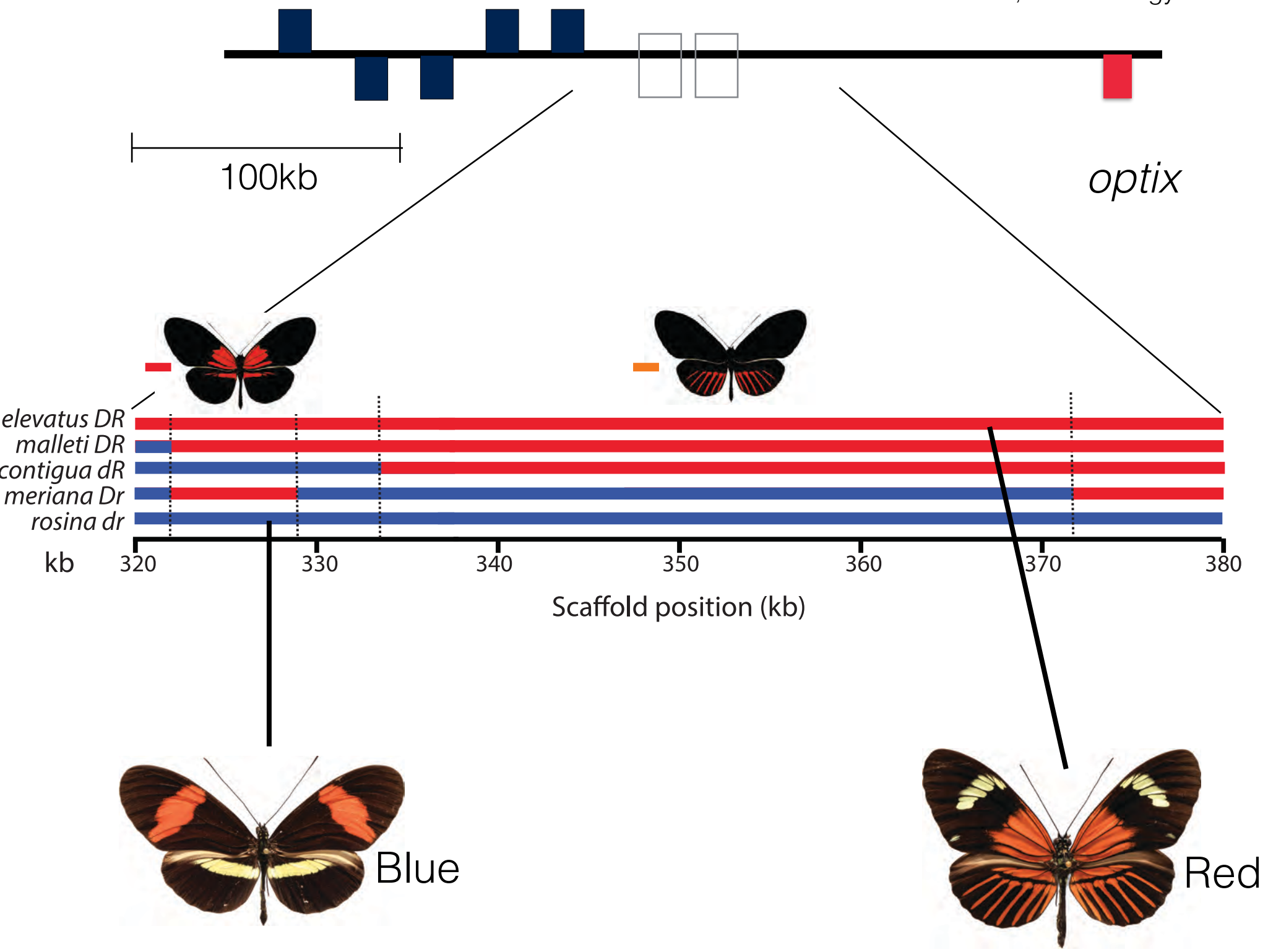
Wallbank et al., PLoS Biology 2016



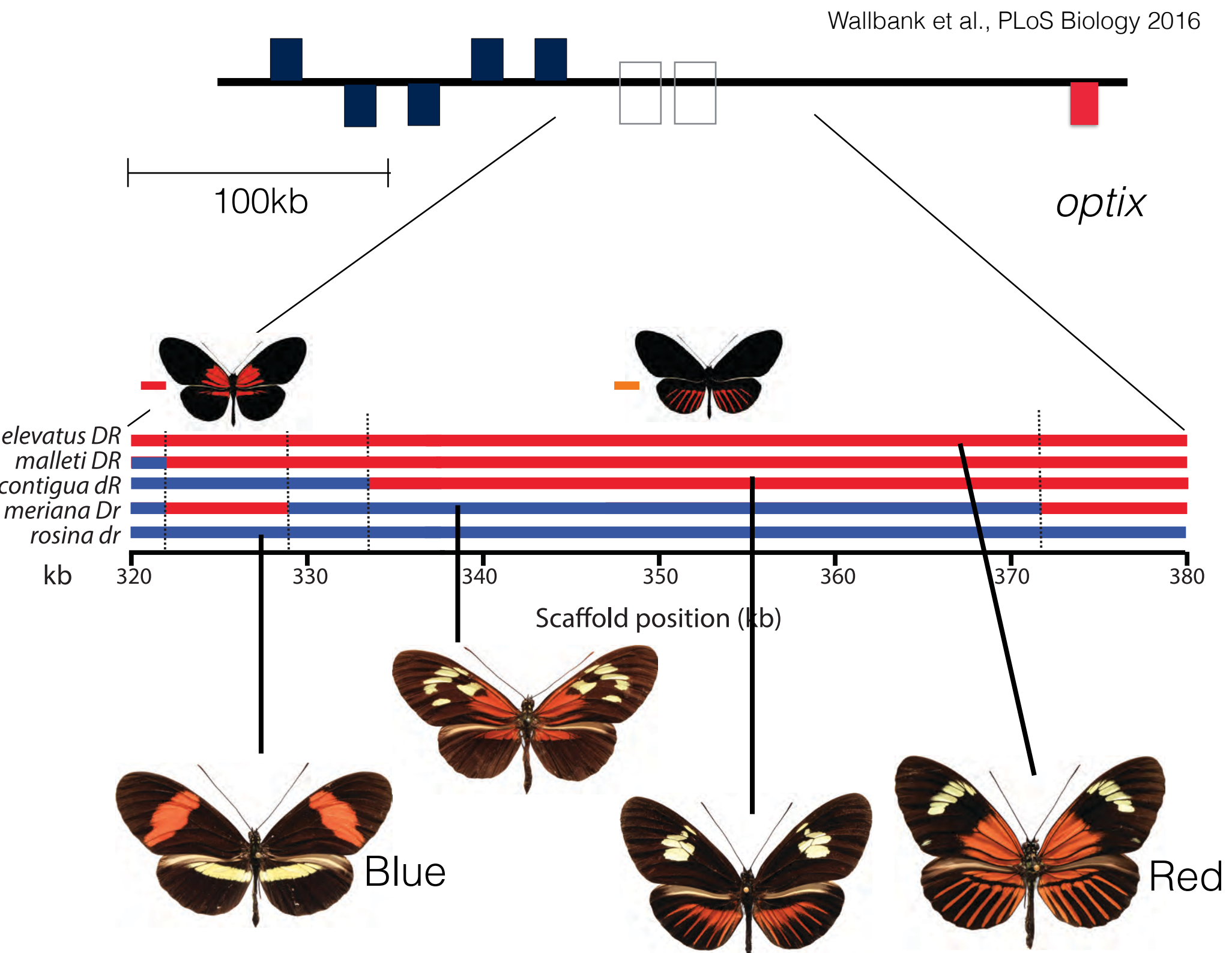
Wallbank et al., PLoS Biology 2016



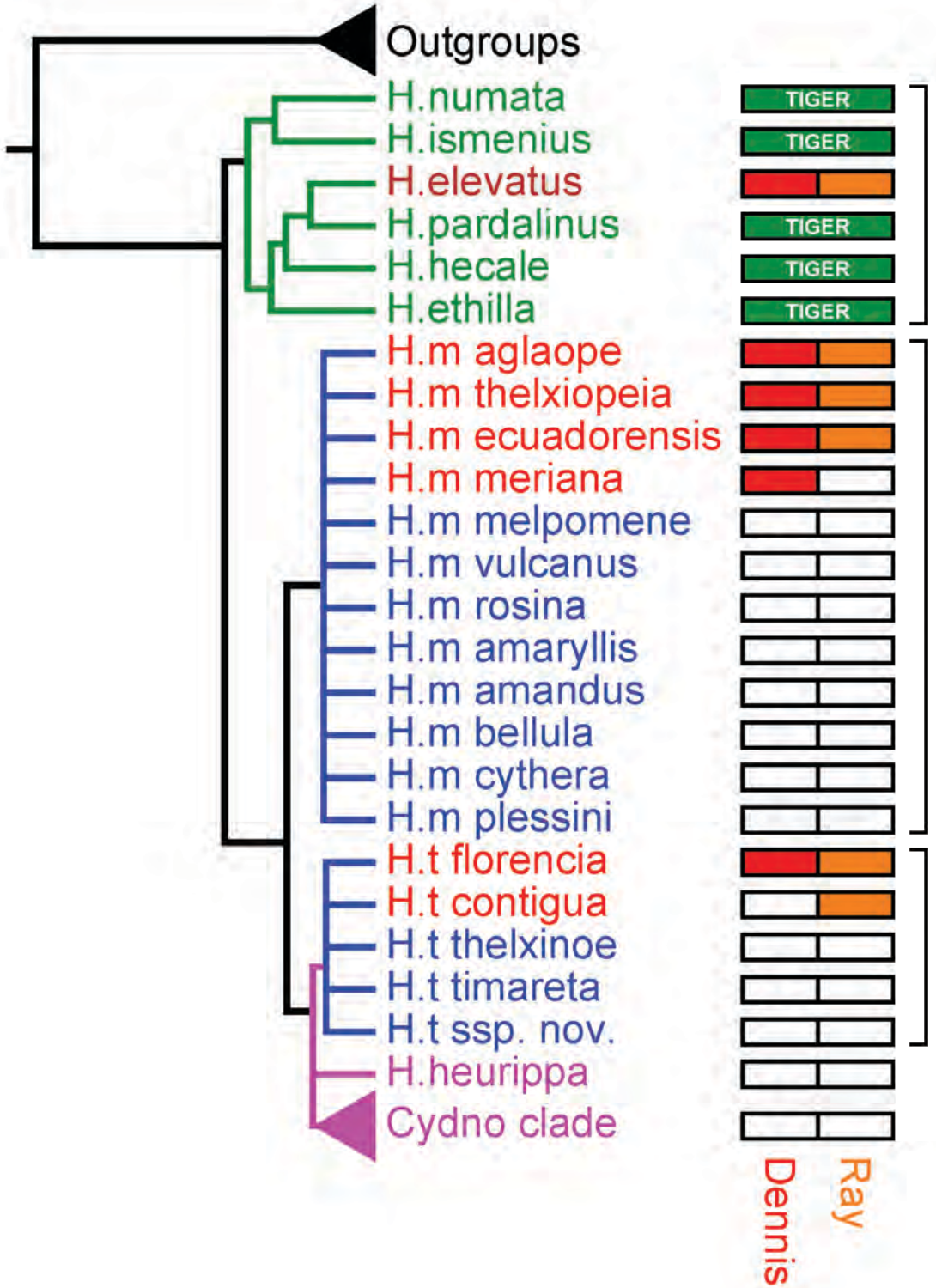
Wallbank et al., PLoS Biology 2016



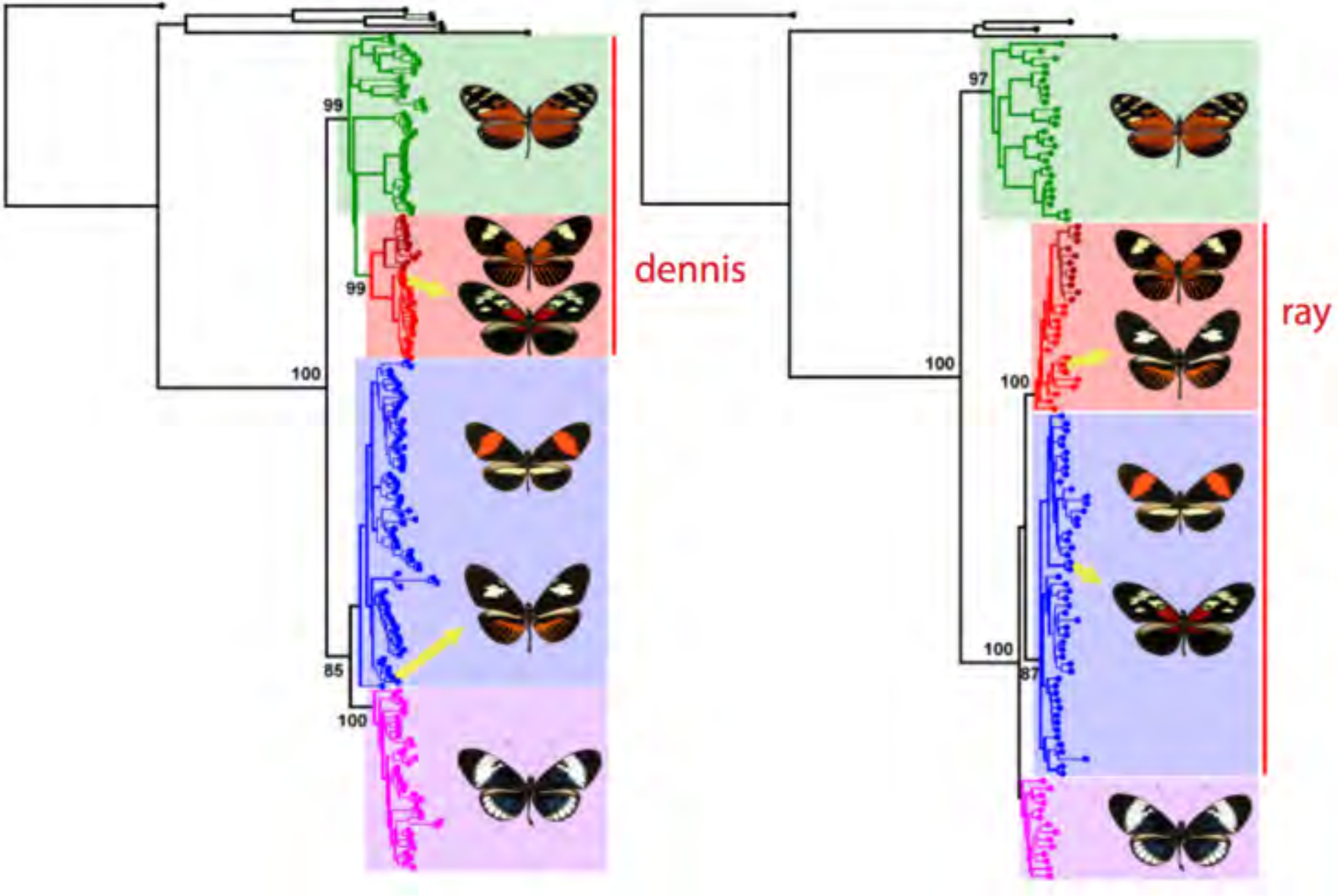
Wallbank et al., PLoS Biology 2016



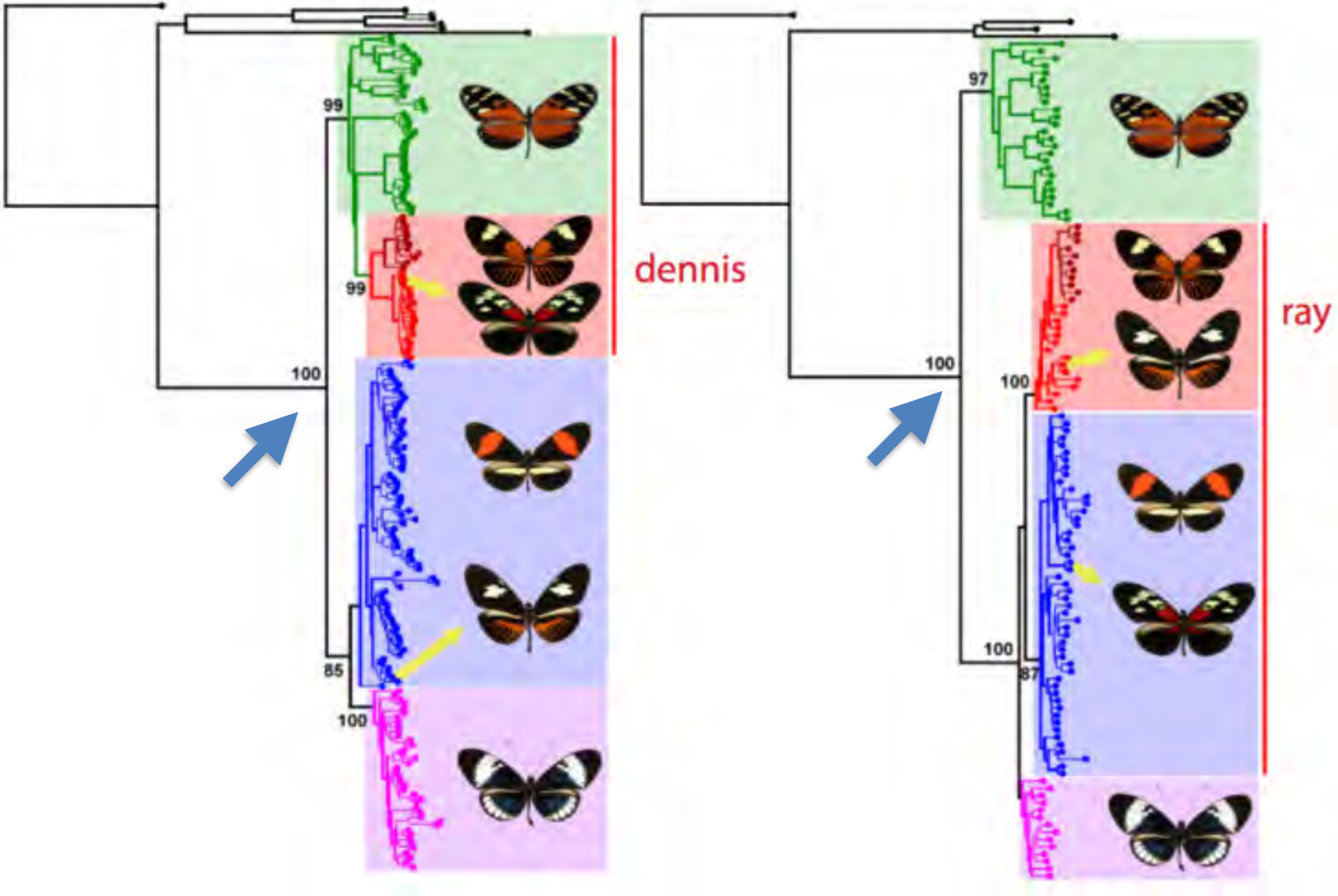
Wallbank et al., PLoS Biology 2016




Wallbank et al., PLoS Biology 2016

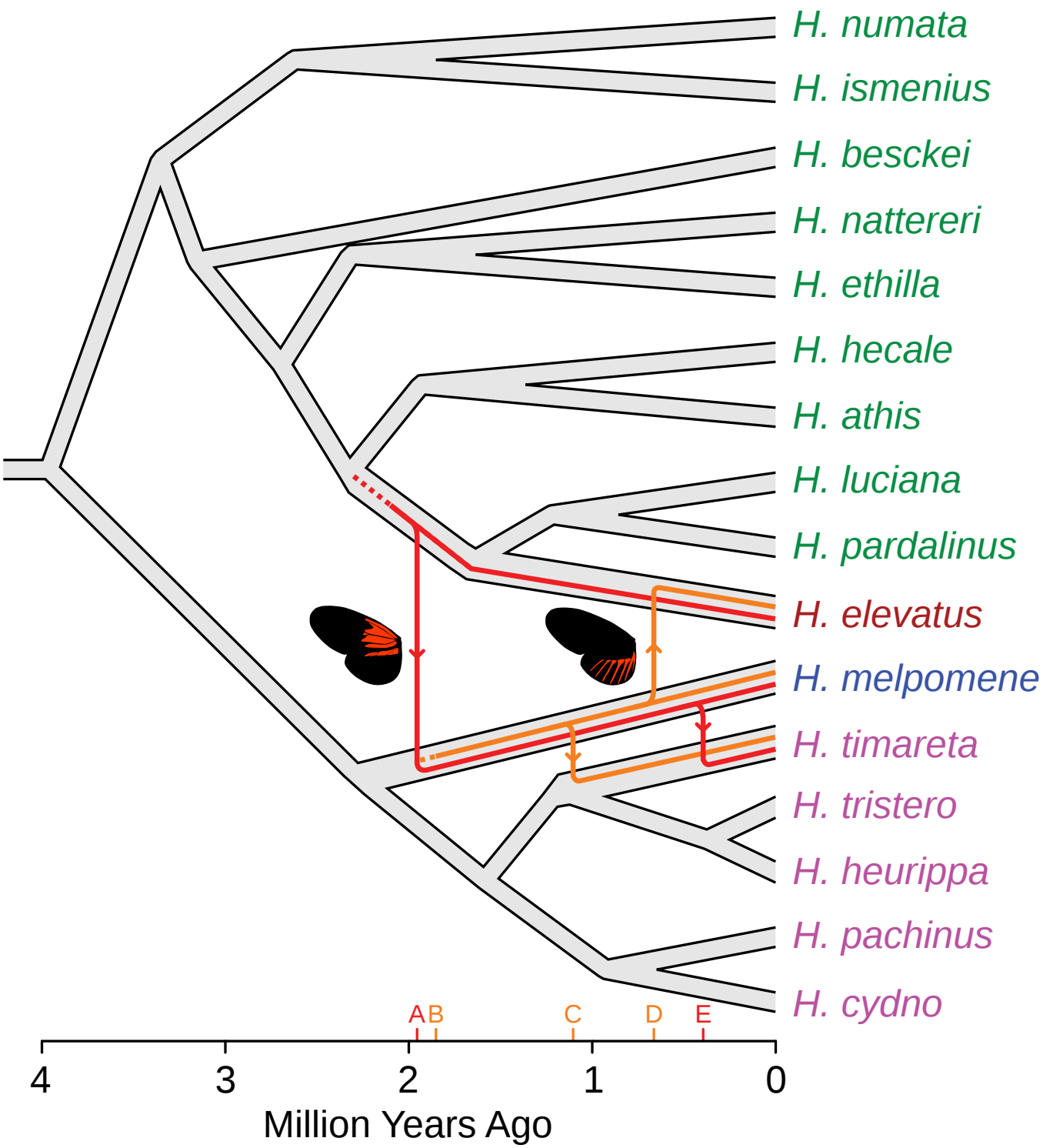


Wallbank et al., PLoS Biology 2016




Generate dated trees using this node as a reference point

Wallbank et al., PLoS Biology 2016

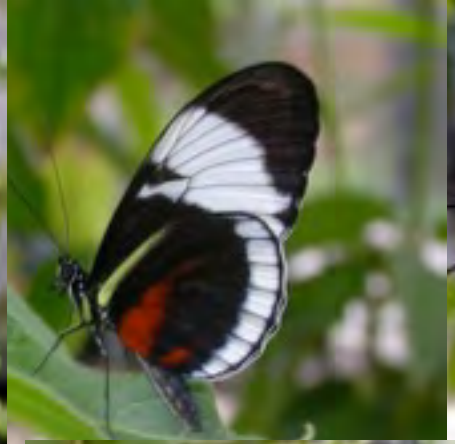
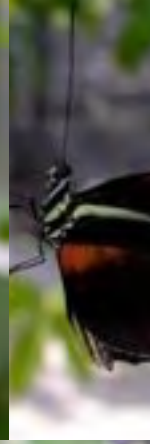
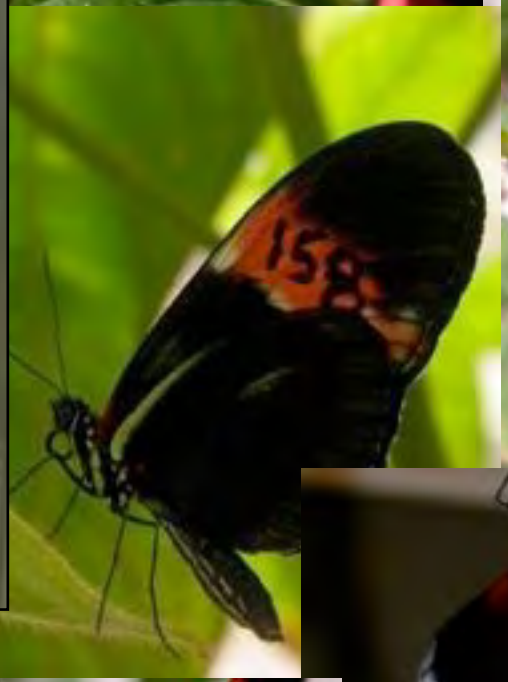


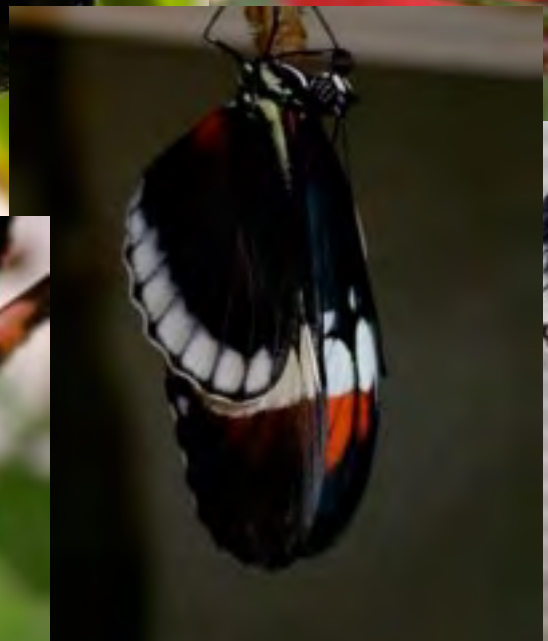
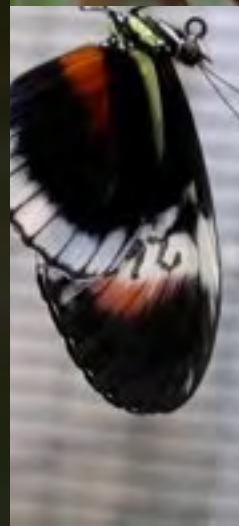
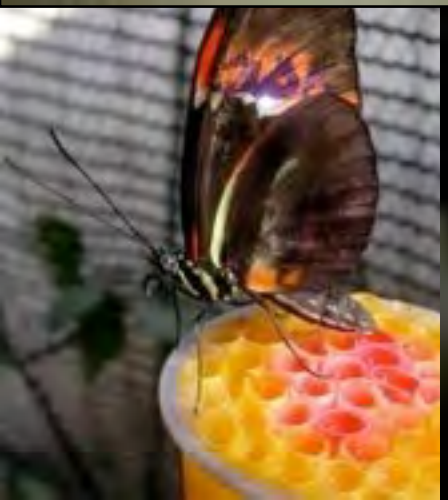
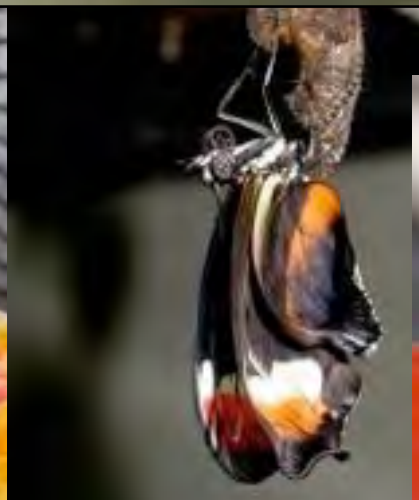
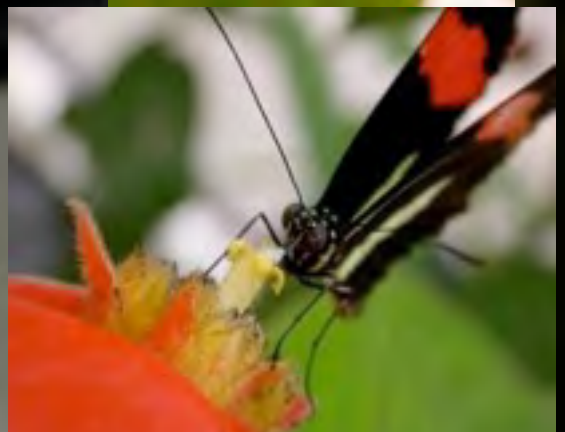
# What about behaviour?

# What about behaviour?

# H. melpomene X H. cydno



Richard Merrill

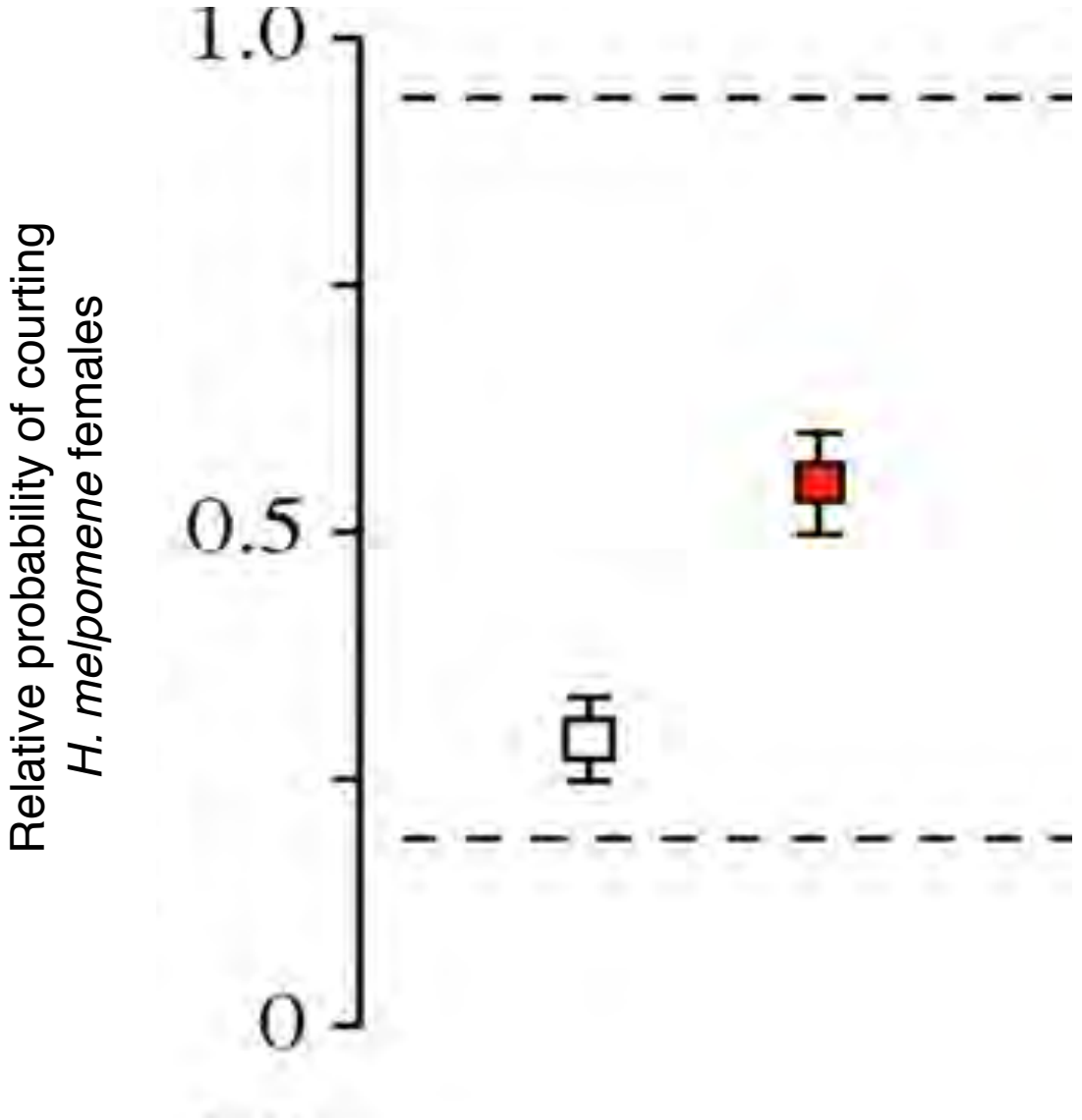
Difference between approaches to cydno and melponmene

30
20
10
0
-10

bb
Bb

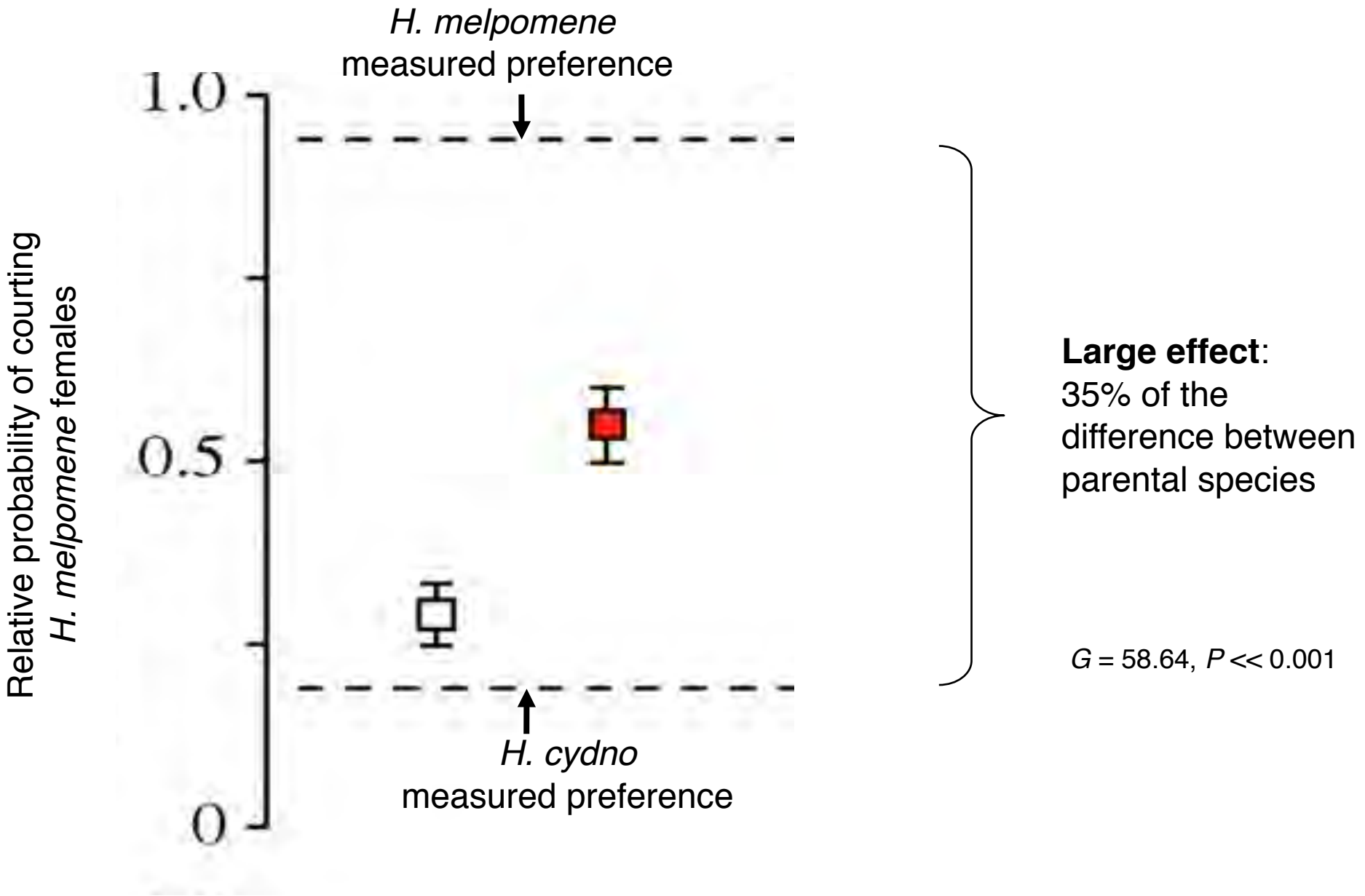
Genotype

# Mate preference segregates with forewing colour in backcross hybrids




Merrill *et al.* Proc Roy Soc B, 2011

# Mate preference segregates with forewing colour in backcross hybrids




Merrill *et al.* Proc Roy Soc B, 2011

NNbb

Heliconius cydno cordula

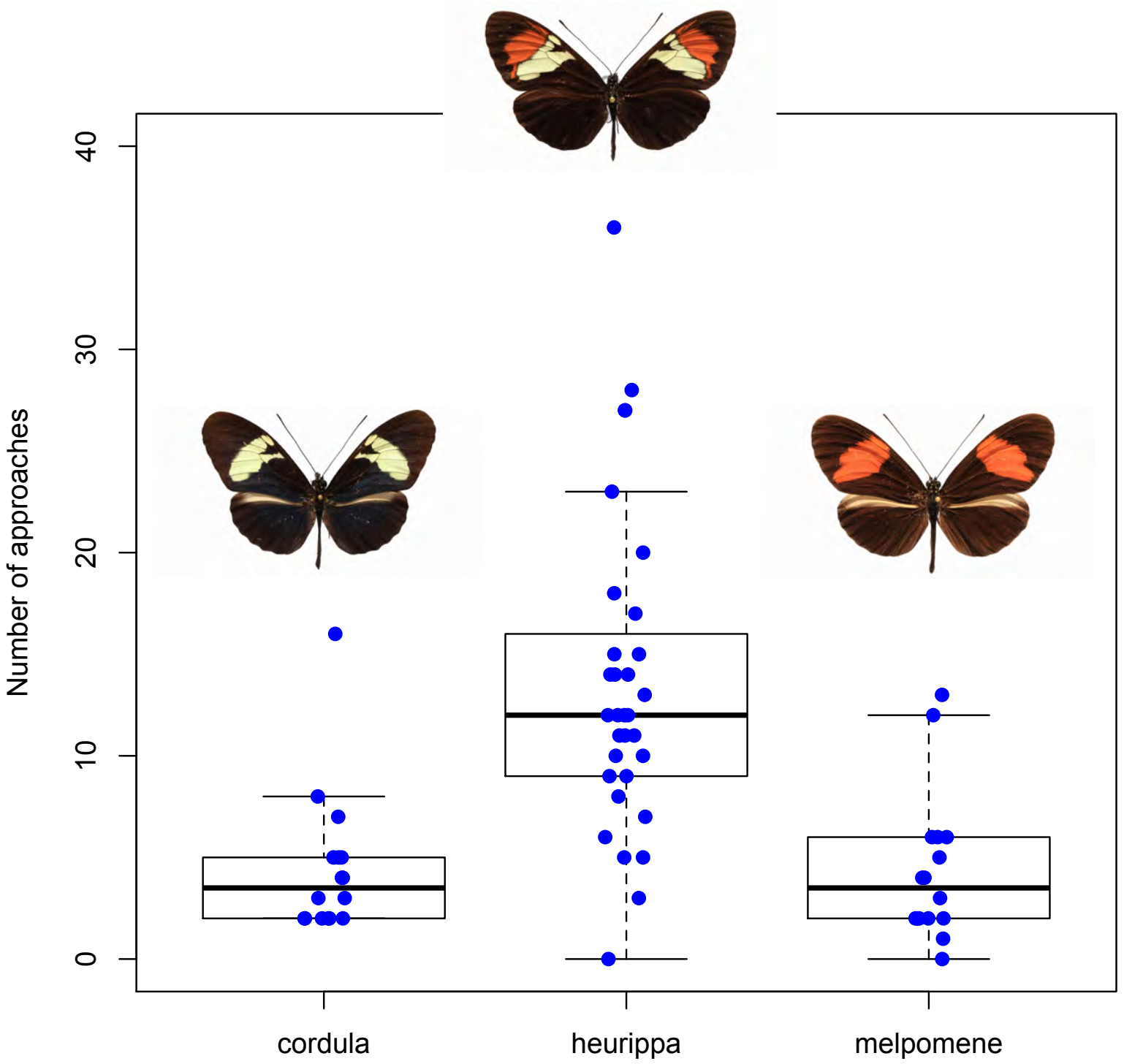
NNBB

Heliconius heurippa

nnBB

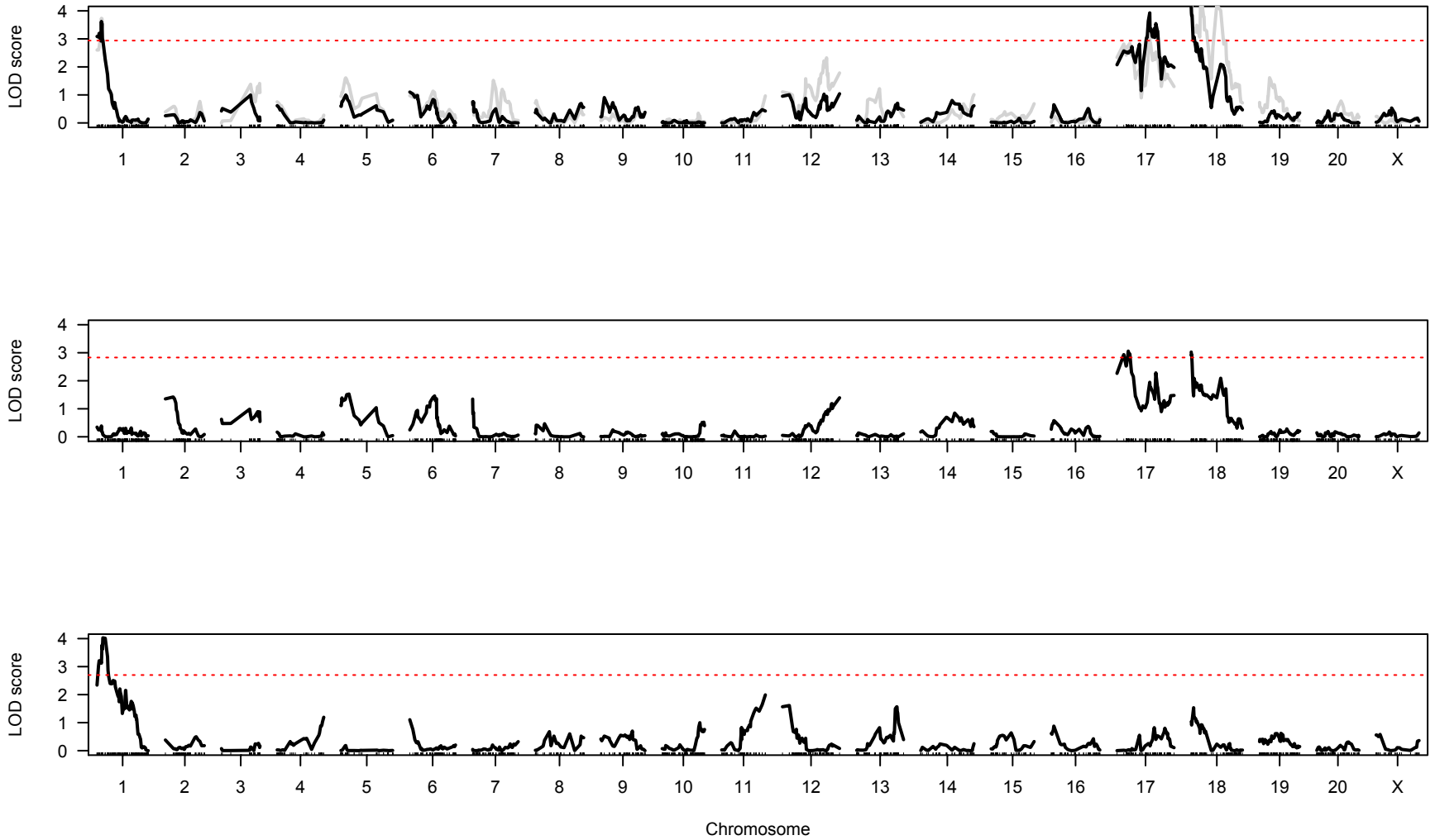
Heliconius melpomene melpomene

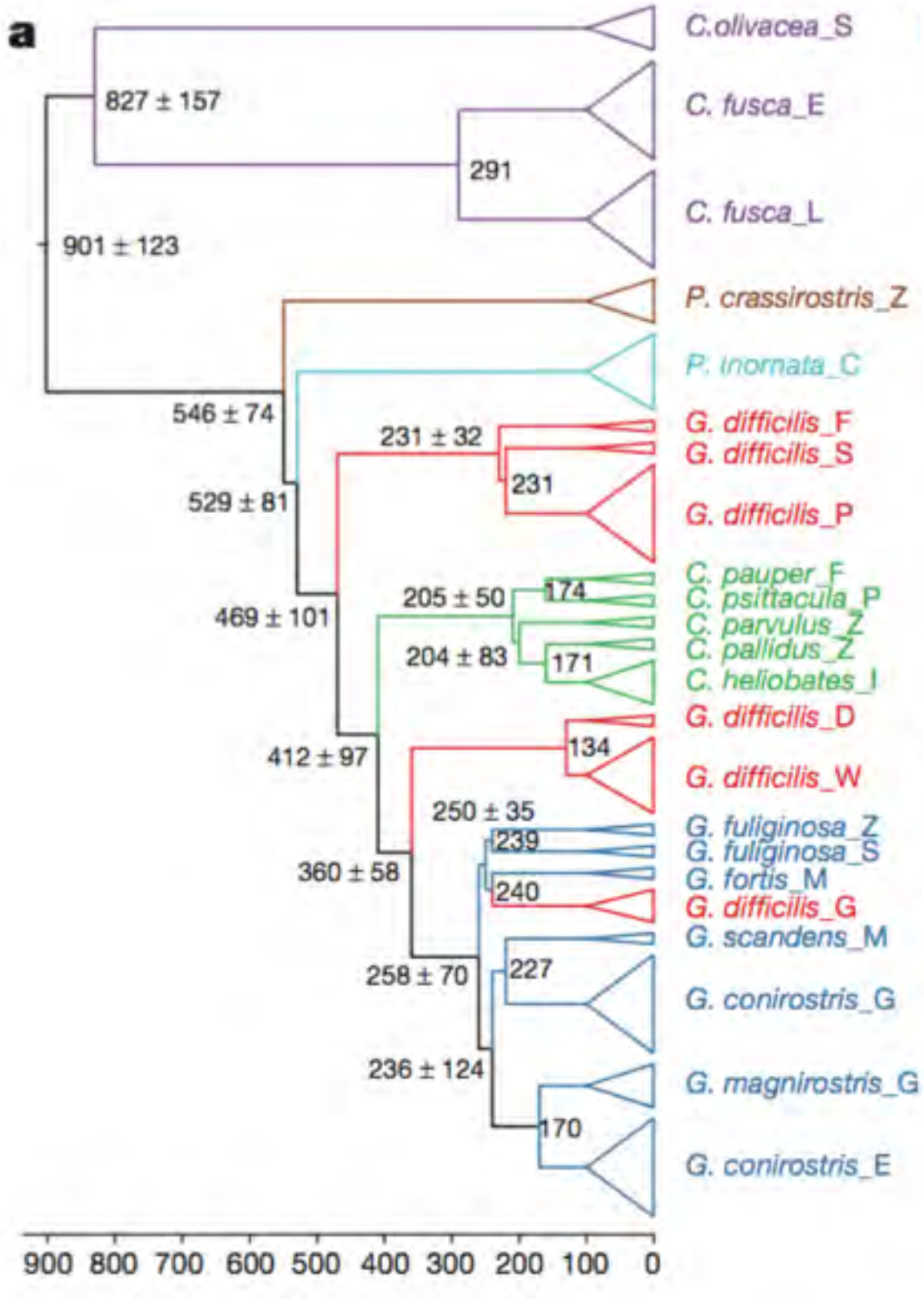
Mavarez et al., Nature 2006



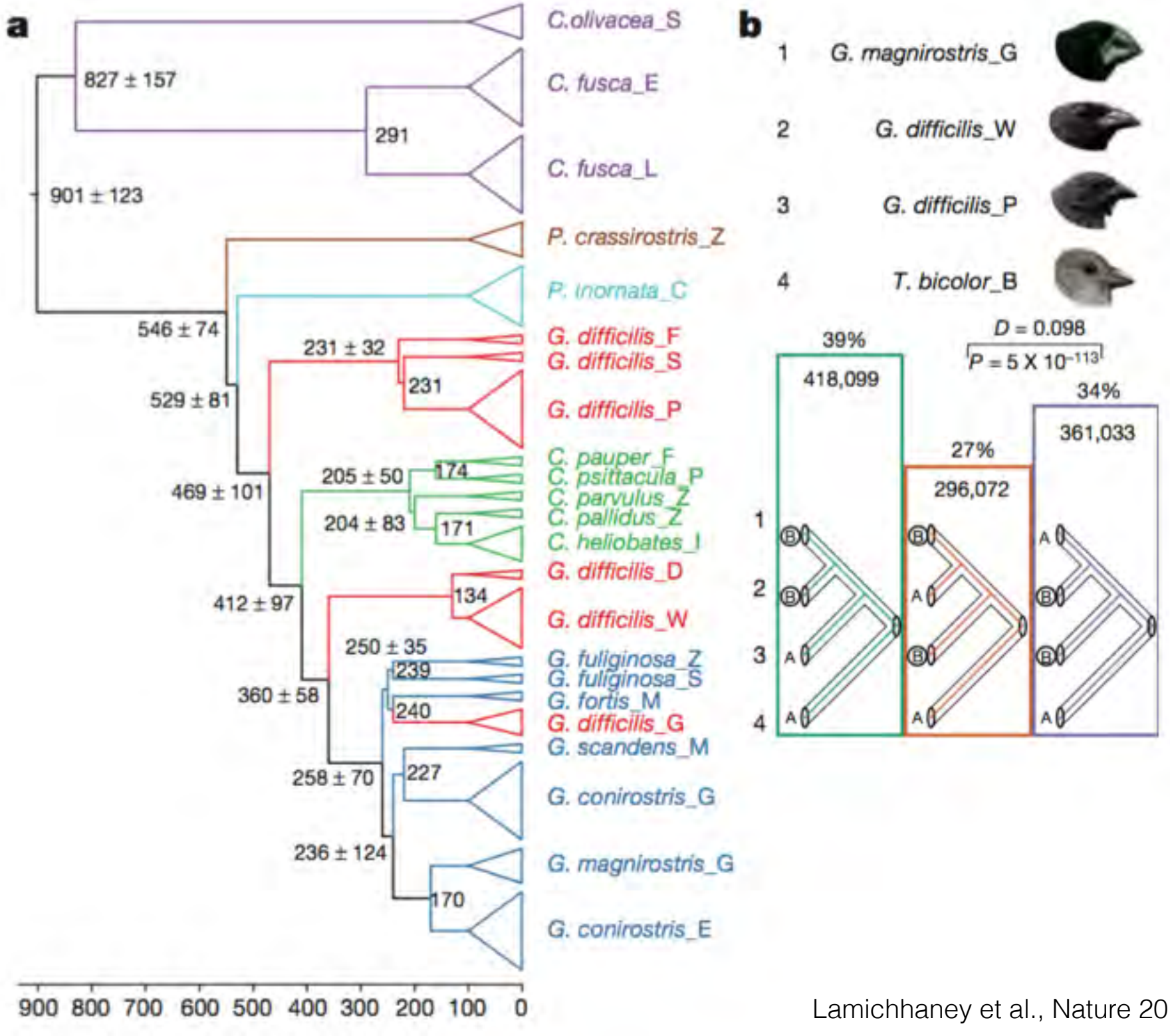

Melo et al., Evolution 2009

# Now working on QTL maps of species differences in behaviour:



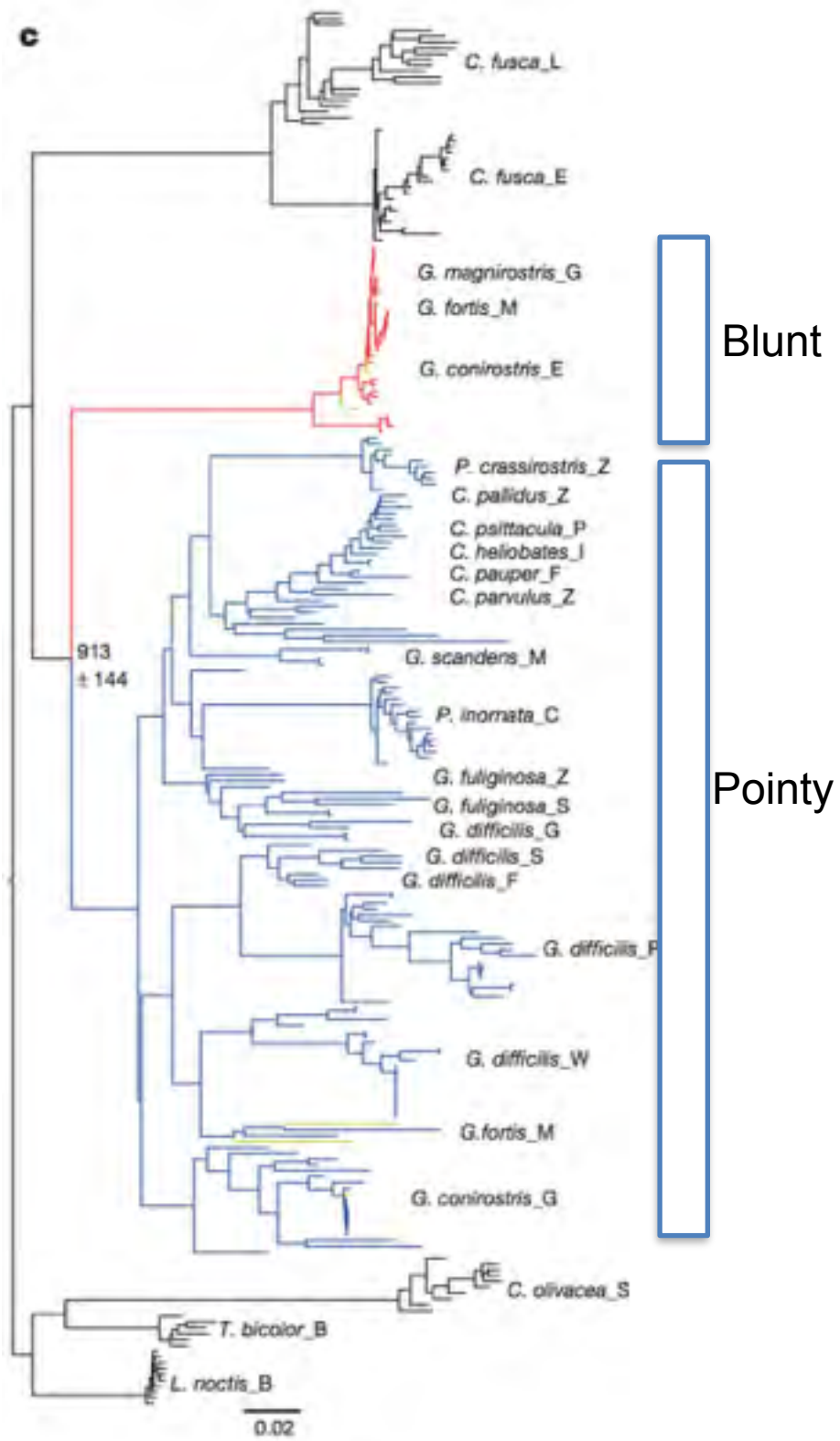


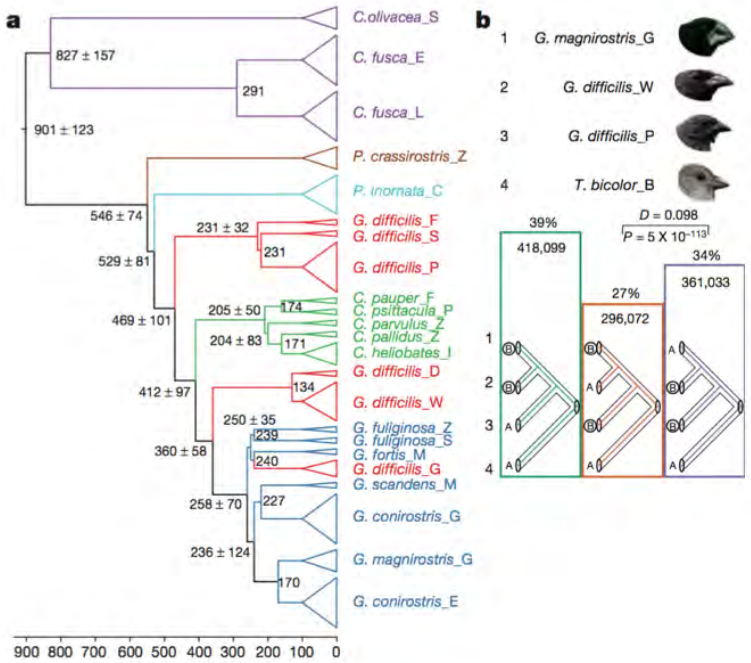

Lamichhaney et al., Nature 2015




Lamichhaney et al., Nature 2015

# ALX1 associated with beak shape

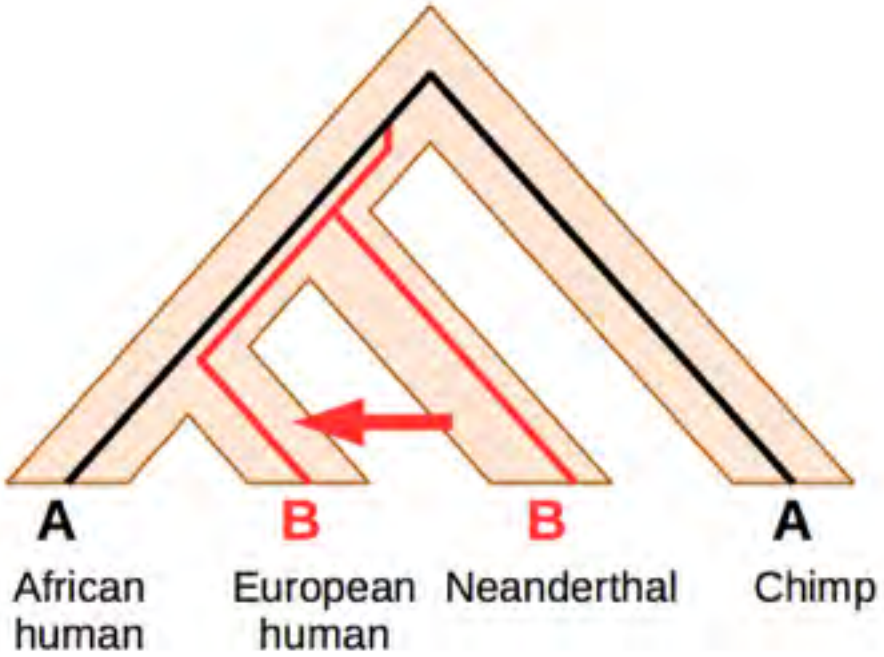


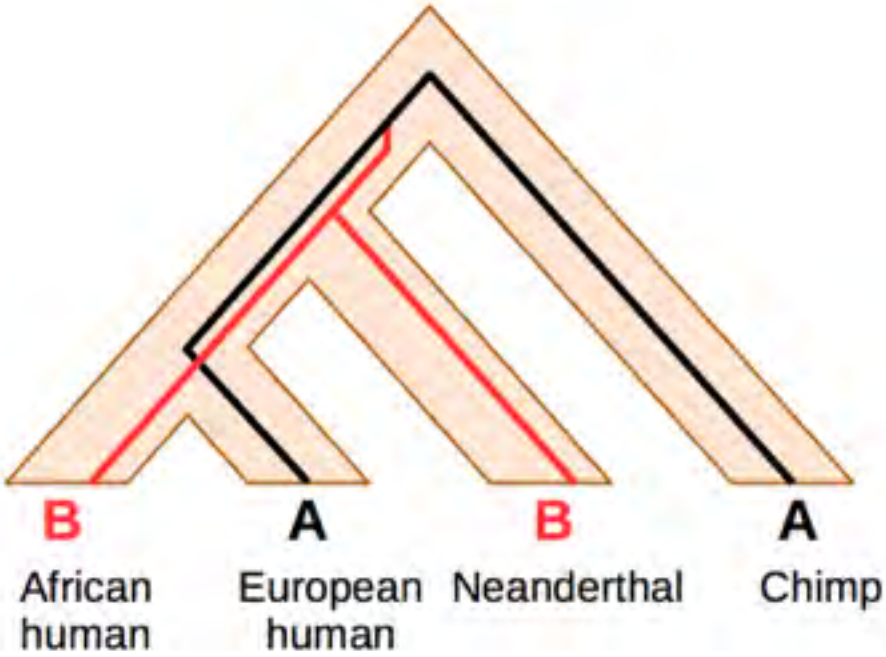



Most of these studies use phenotype associations to identify introgressed loci

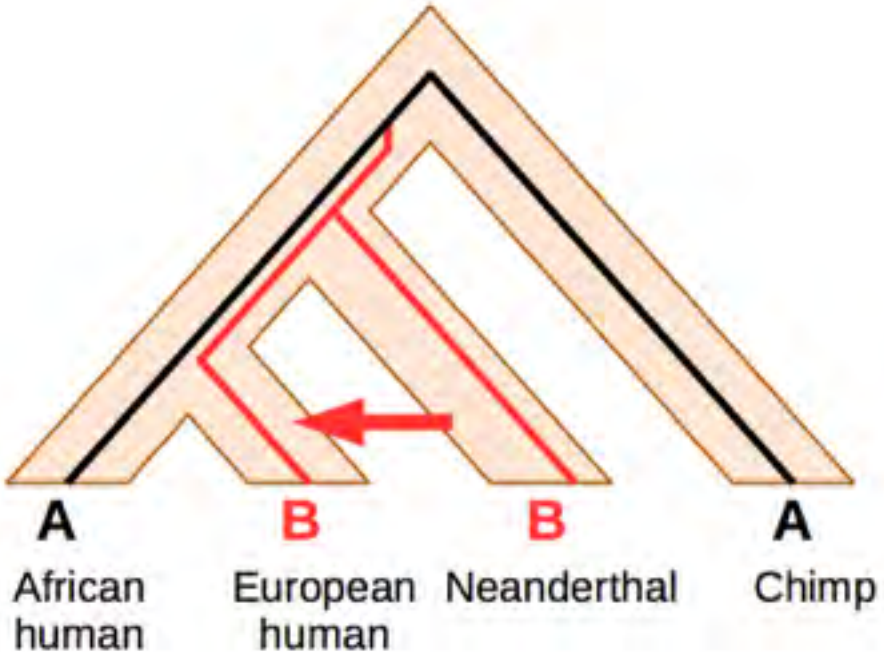
But can we identify them a priori using the ABBA-BABA method?

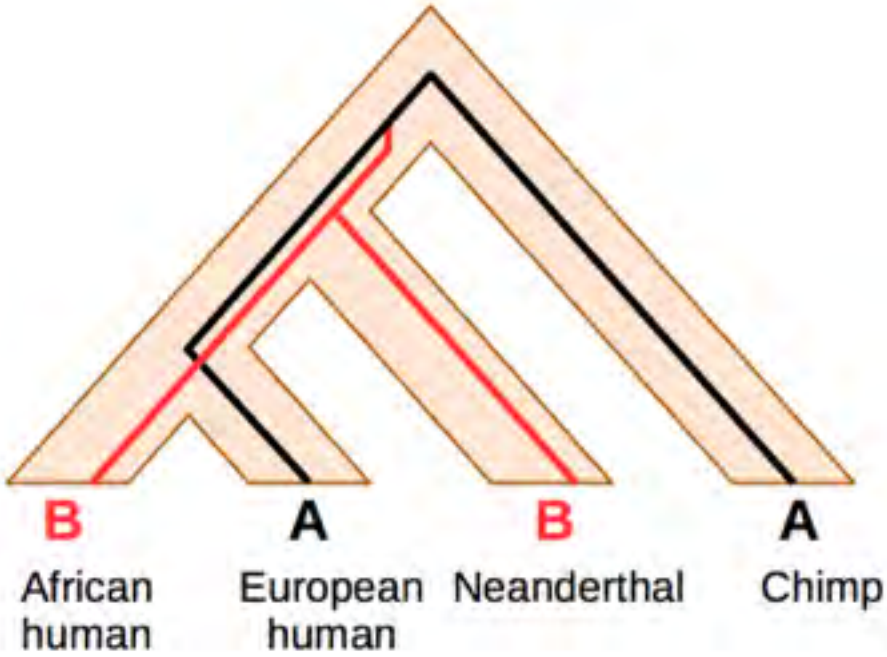
# Explicit tests for gene flow: ABBA-BABA test



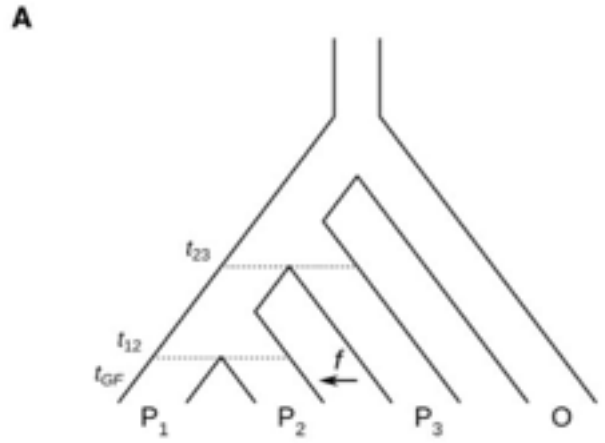



# Explicit tests for gene flow: ABBA-BABA test



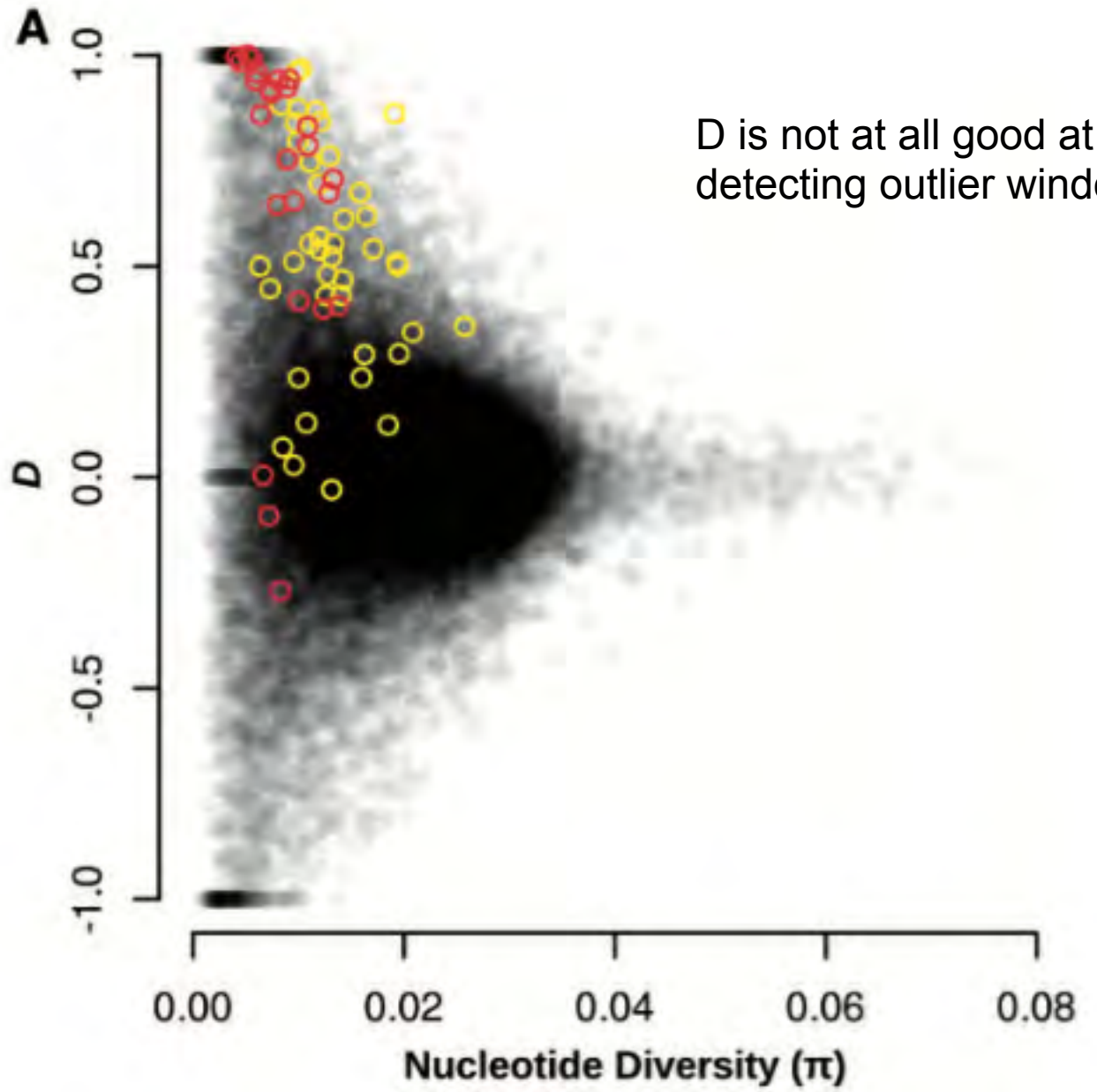

Green *et al.* 2010 *Science* 328:710-722




$$D(P_1, P_2, P_3, O) = \frac{\sum C_{ABBA}(i) - C_{BABA}(i)}{\sum C_{ABBA}(i) + C_{BABA}(i)} \qquad (1)$$

D is quite dependent on the number of informative sites (the denominator)

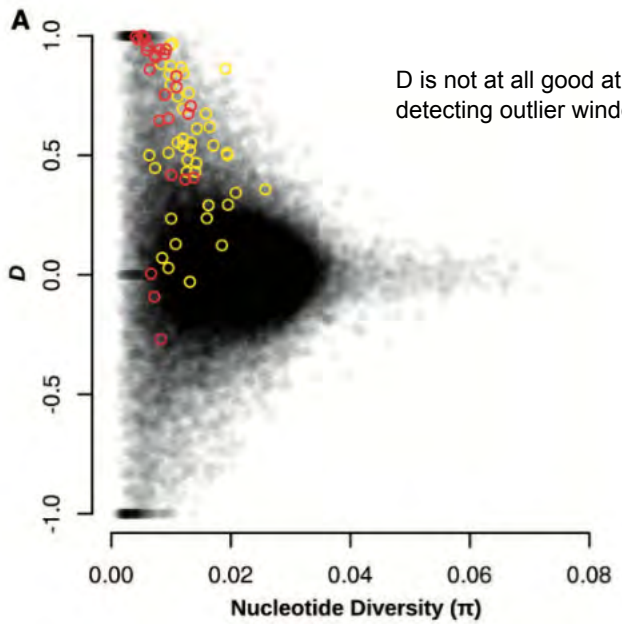
Martin et al., MBE 2014



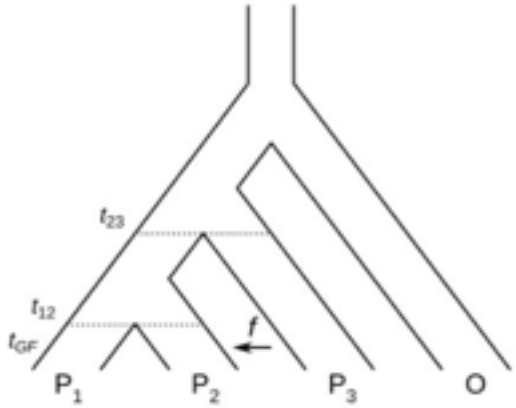

D is not at all good at detecting outlier windows

ites (the denominator)

Martin et al., MBE 2014



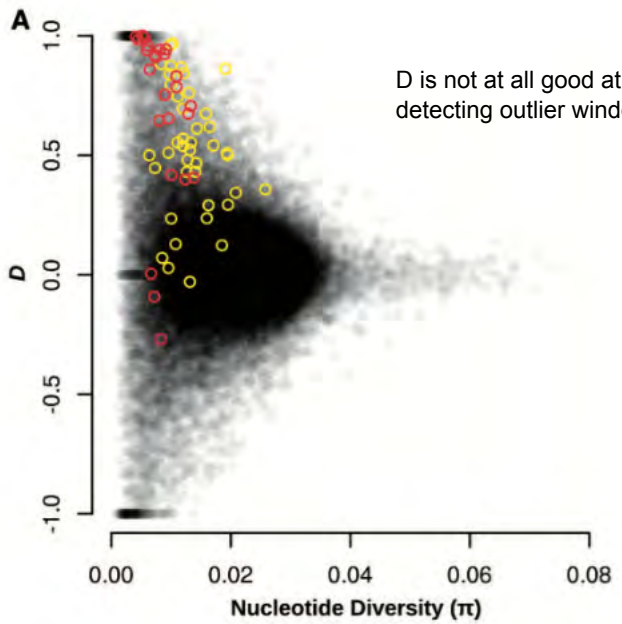

D is not at all good at detecting outlier windows



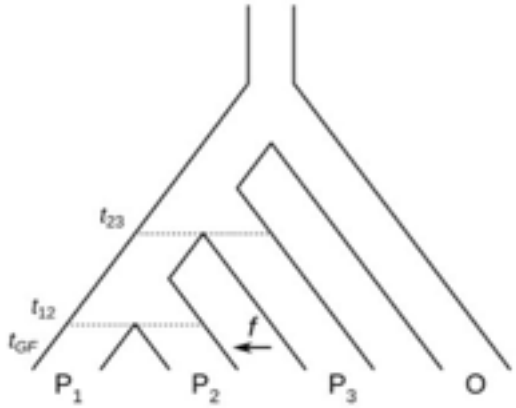

$$D(P_1, P_2, P_3, O) = \frac{\sum C_{ABBA}(i) - C_{BABA}(i)}{\sum C_{ABBA}(i) + C_{BABA}(i)} \qquad (1)$$

D is quite dependent on the number of informative sites (the denominator)

Martin et al., MBE 2014




D is not at all good at detecting outlier windows



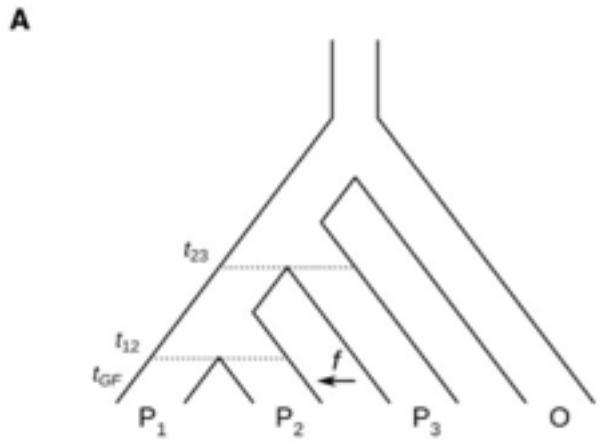

$$D(P_1, P_2, P_3, O) = \frac{\sum C_{ABBA}(i) - C_{BABA}(i)}{\sum C_{ABBA}(i) + C_{BABA}(i)} \qquad (1)$$

D is quite dependent on the number of informative sites (the denominator)

$$\hat{f}_G = \frac{S(P_1, P_2, P_3, O)}{S(P_1, P_{3a}, P_{3b}, O)}$$

Where s is numerator from the D equation
f is the fraction of introgression compared to maximum possible

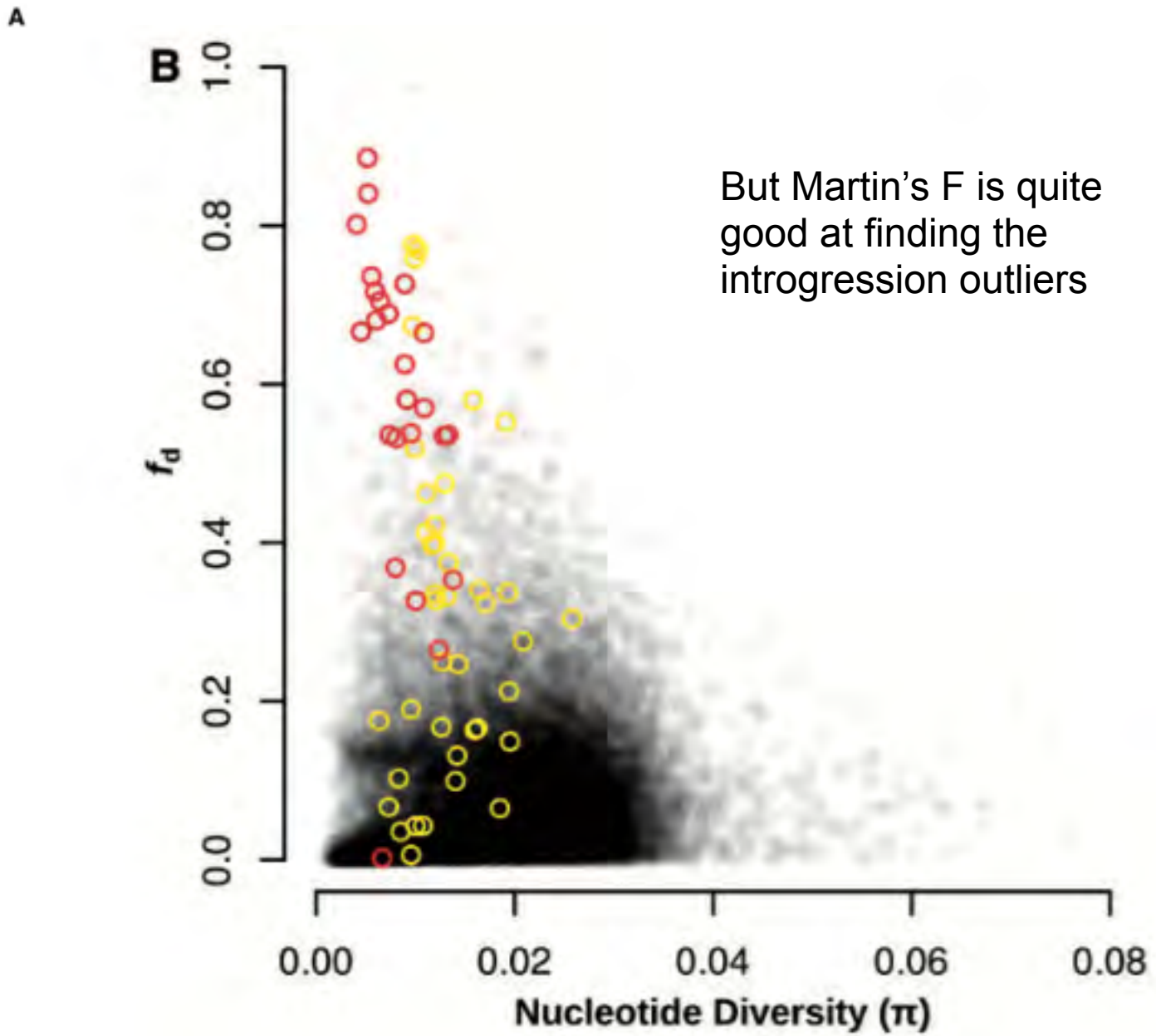
Martin et al., MBE 2014




Martin's F

$$\hat{f}_d = \frac{S(P_1, P_2, P_3, O)}{S(P_1, P_D, P_D, O)}$$

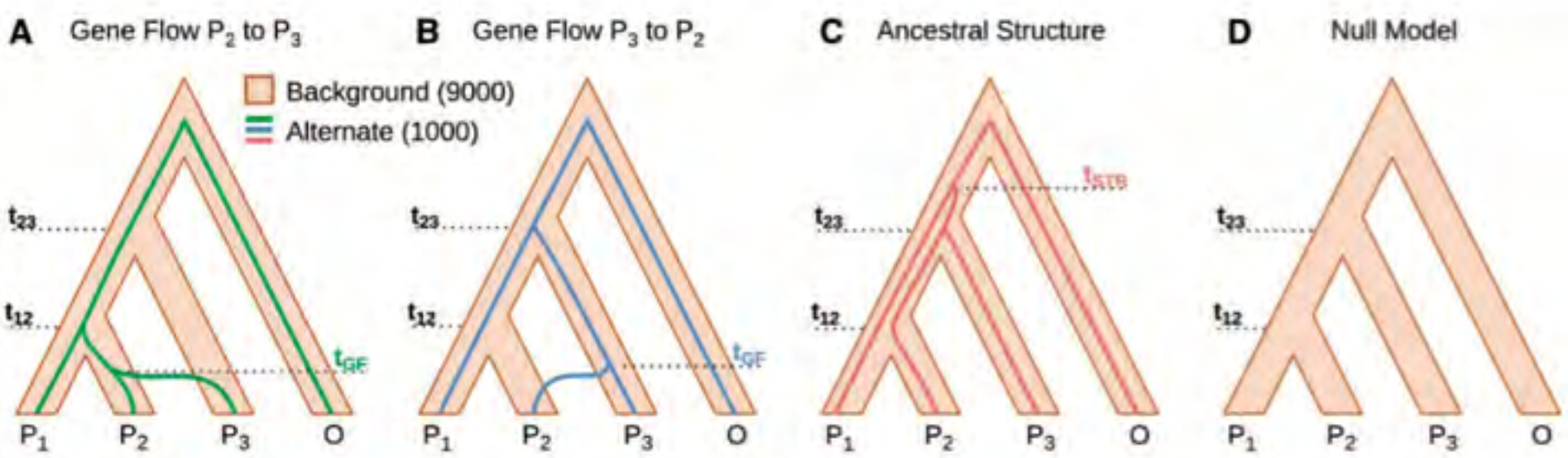
Martin et al., MBE 2014

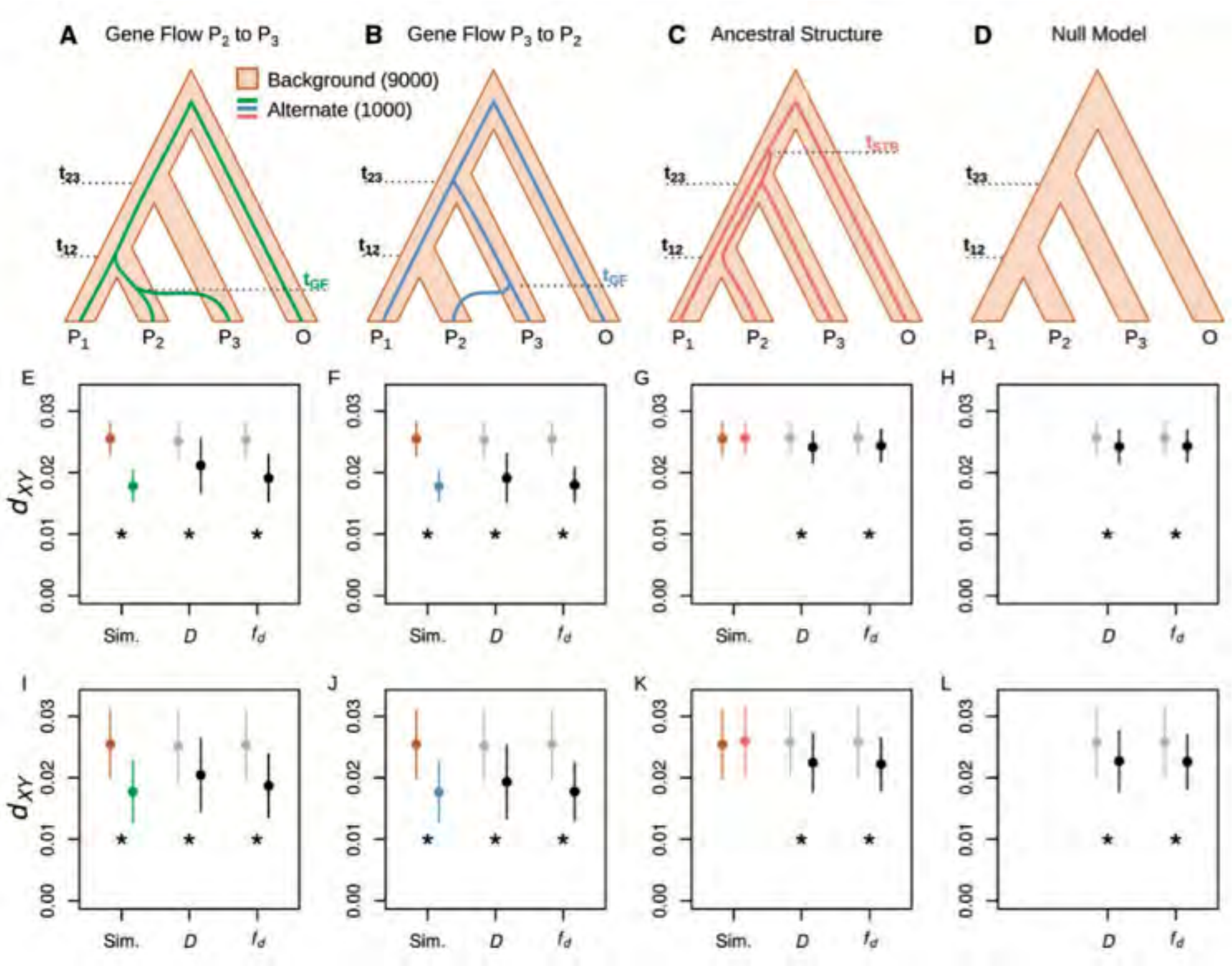



But Martin's F is quite good at finding the introgression outliers

Martin et al., MBE 2014

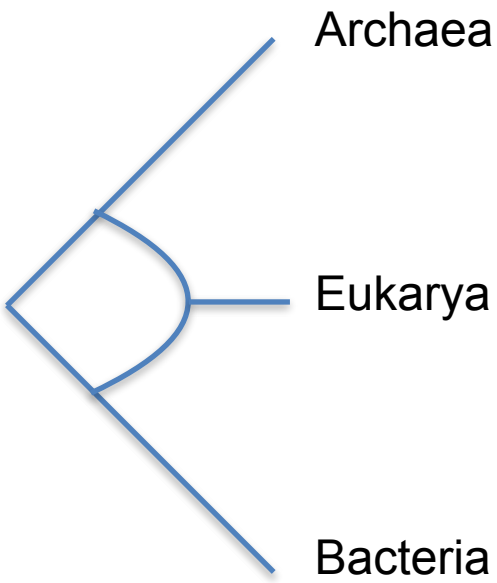

- Smith and Kronforst argued that introgression could be inferred where ABBA-BABA outliers showed lower Dxy compared to genome-wide average


A
Gene Flow P2 to P3
B
Gene Flow P3 to P2
C
Ancestral Structure
D
Null Model
Background (9000)
Alternate (1000)
t23
t12
tGF
tSTR
P1
P2
P3
O


A Gene Flow P2 to P3
B Gene Flow P3 to P2
C Ancestral Structure
D Null Model
Background (9000)
Alternate (1000)
t23
t12
tGF
tSTR
P1
P2
P3
O
E
F
G
H
I
J
K
L
dXY
0.00
0.01
0.02
0.03
Sim.
D
fd

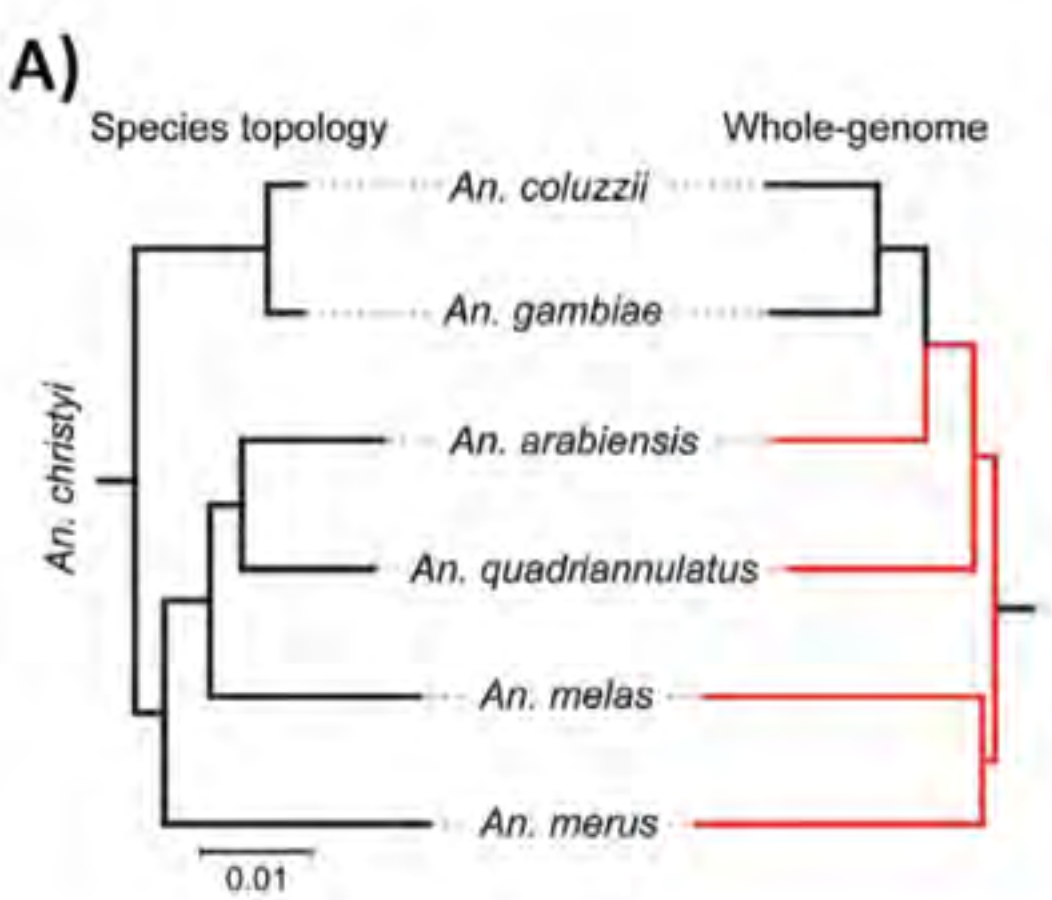
- Be wary of window based D statistics
- F is better than D…
- Sampling design is very important!

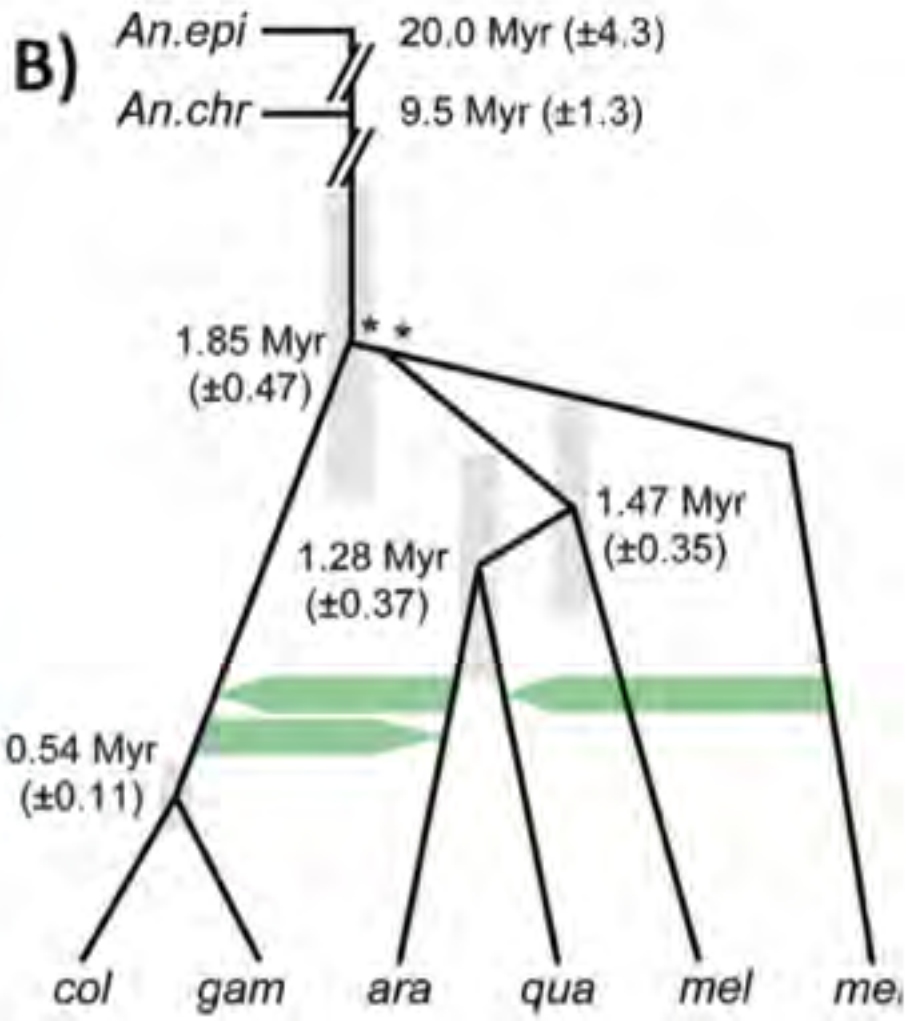
# Implications for tree-thinking



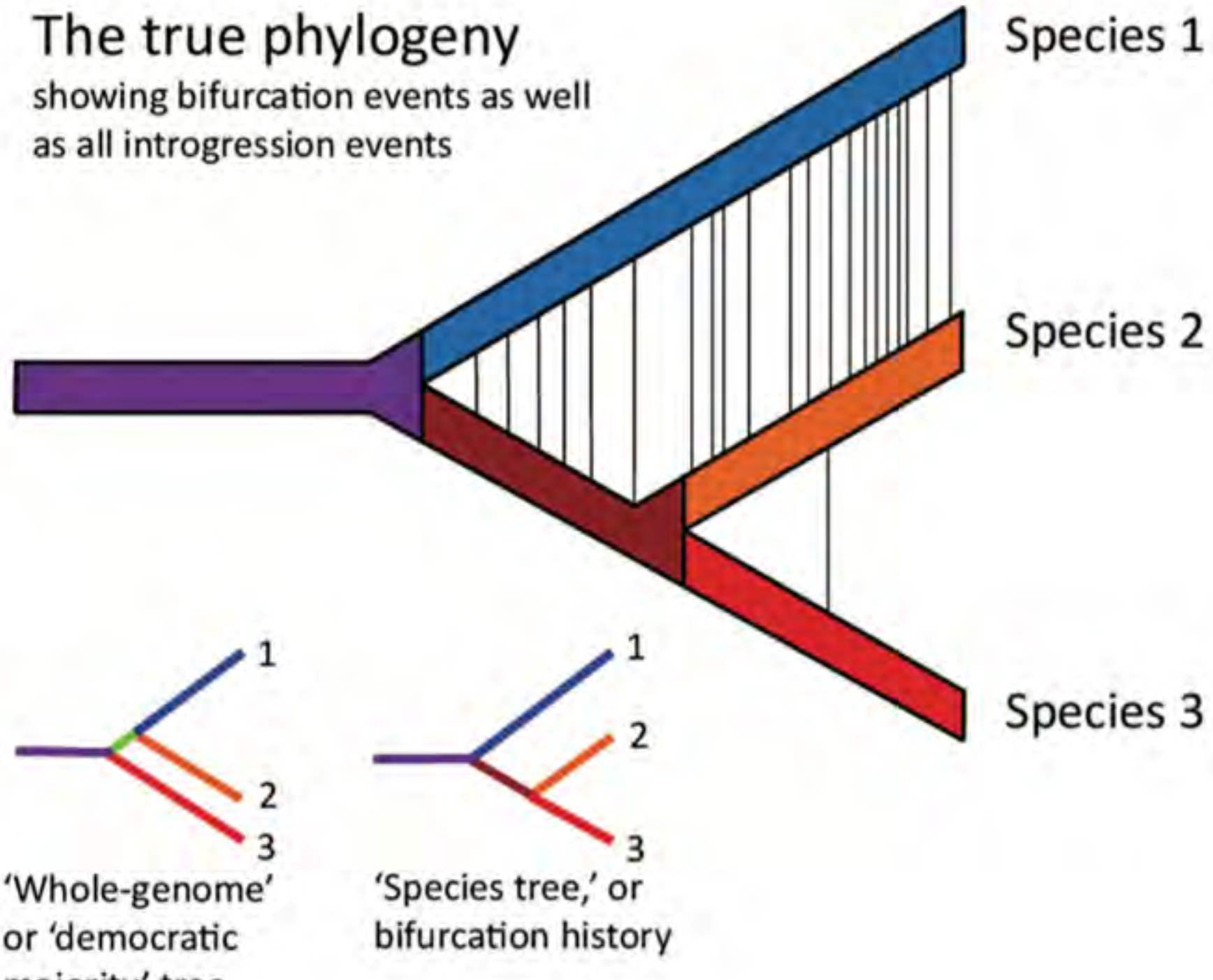

**The tree of life is reticulated**

# Implications for tree-thinking




Mallet, Hahn and Besansky BioEssays 2015
Hahn and Nakhleh Evolution 2015


The true phylogeny
showing bifurcation events as well
as all introgression events
Species 1
Species 2
Species 3
1
2
3
'Whole-genome'
or 'democratic
majority' tree
1
2
3
'Species tree,' or
bifurcation history

Okay, so what have we learnt and where do we go from here?