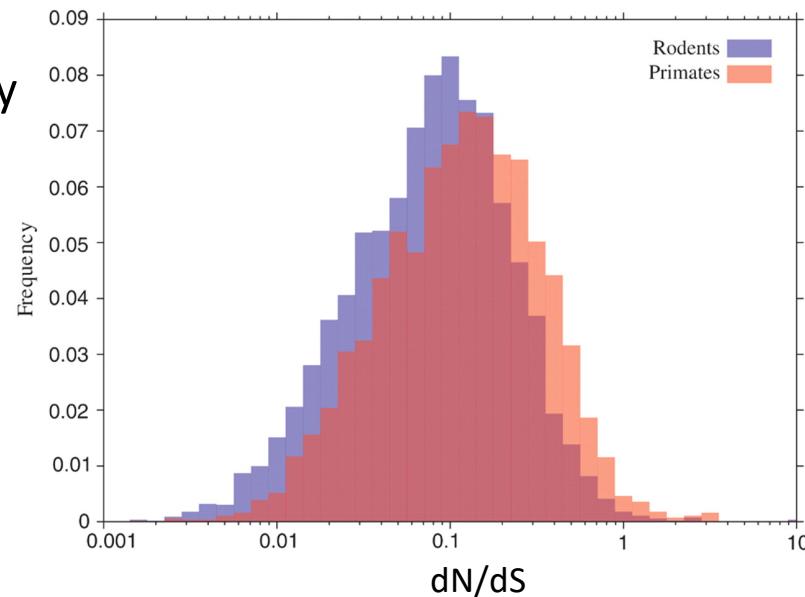




# Levels of constraint

- Neutral theory claims most new mutations are deleterious and are lost immediately. Only neutral mutations contribute to substitution.
- $dN = \text{nonsynonymous subst./site}$
- $dS = \text{synonymous subst./site}$
- Expected  $dN/dS$  for a “typical” gene.  $< 1$

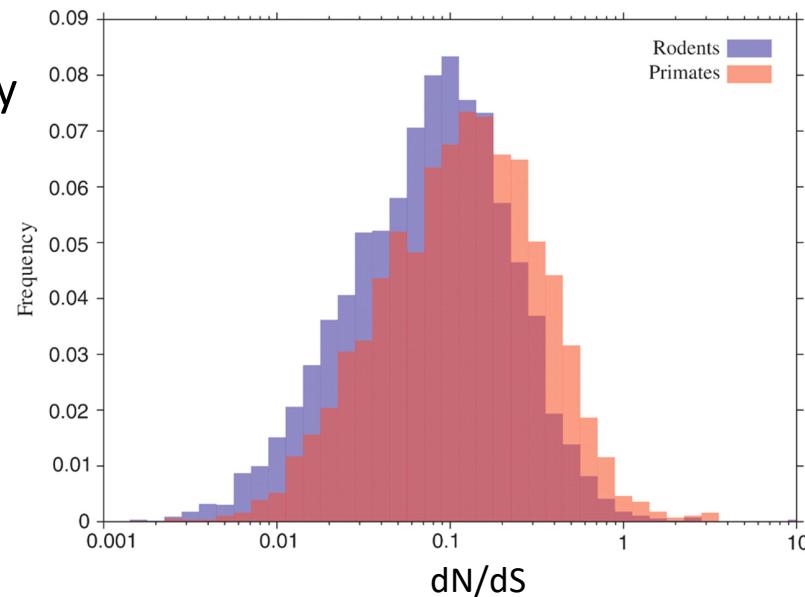
Consistent with neutral theory  
slower rate of substitution at  
more constrained sites.



# Levels of constraint

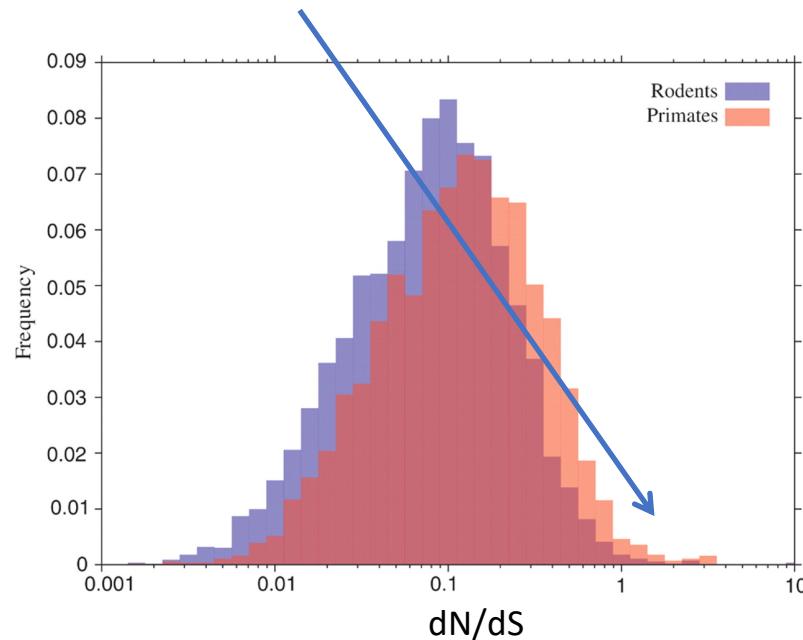
- Neutral theory claims most new mutations are deleterious and are lost immediately. Only neutral mutations contribute to substitution.
- $dN = \text{nonsynonymous subst./site}$
- $dS = \text{synonymous subst./site}$
- Expected  $dN/dS$  for a “typical” gene.  $< 1$

Consistent with neutral theory  
slower rate of substitution at  
more constrained sites.



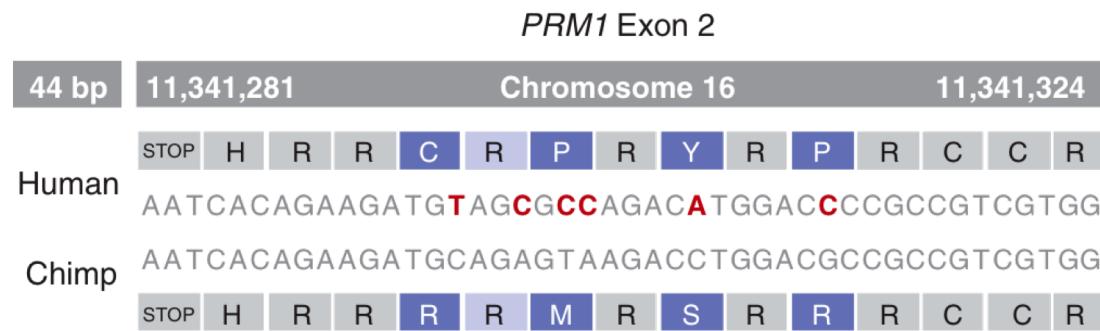
## Adaptive protein divergence and the dN/dS ratio

- dN = nonsynonymous subst./site
- dS = synonymous subst./site
- Expected dN/dS for a “typical” gene. < 1
- Expected dN/dS for a gene coding for an unconstrained protein = 1
- Recurrent directional selection, dN/dS > 1 \*



\* This is a very conservative test.

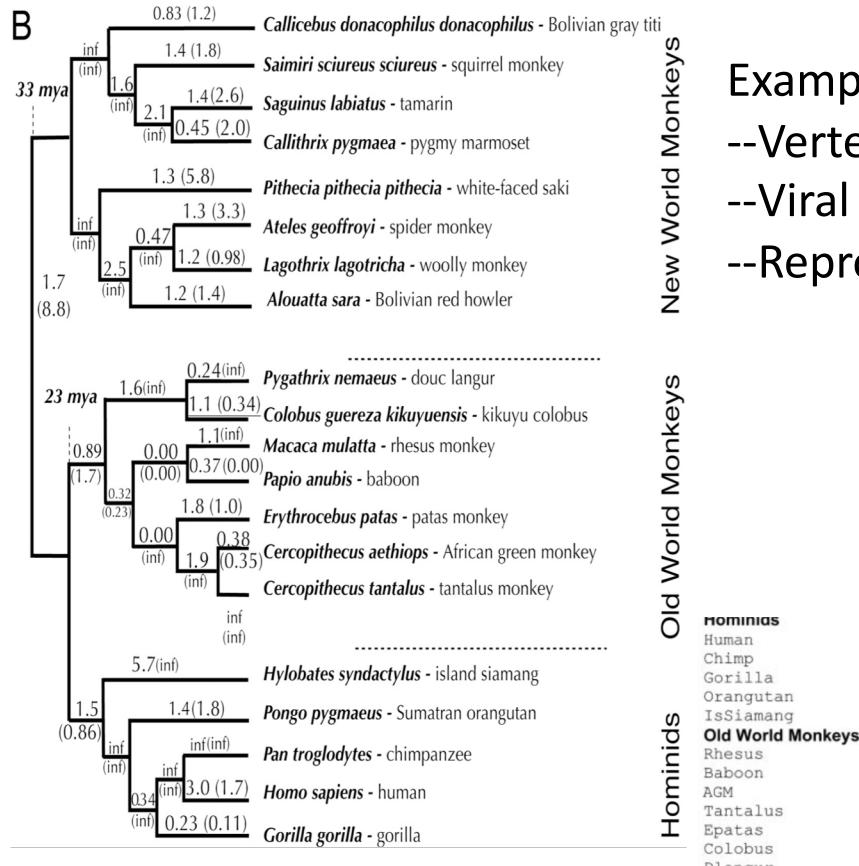
# An example



**Fig. 2.** Excess of function-altering mutations in *PRM1* exon 2. The *PRM1* gene exon 2 contains six differences between humans and chimpanzees, five of which alter amino acids (7, 8).

Sabeti et al 2007

## TRIM5 $\alpha$ gene involved in primate retrovirus defense. dN/dS > 1 for the majority of branches in phylogeny



Sawyer et al 2005

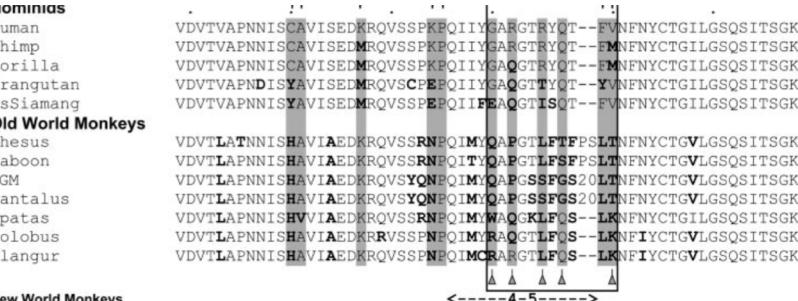
Examples of genes showing dN/dS > 1

- Vertebrate immune system
- Viral coat proteins in viruses
- Reproduction related genes

Invertebrate gamete

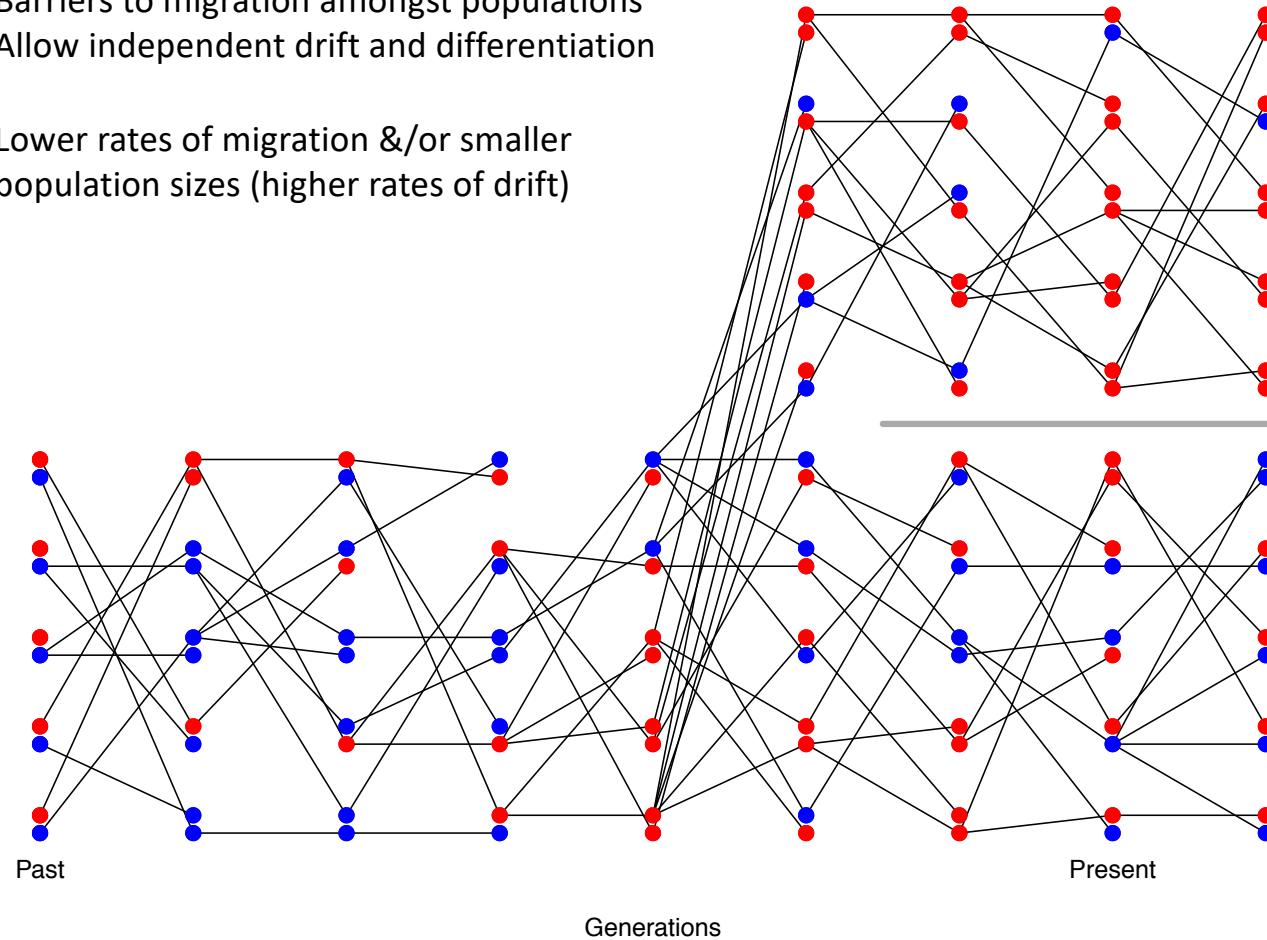
recognition proteins

Vertebrate gametogenesis

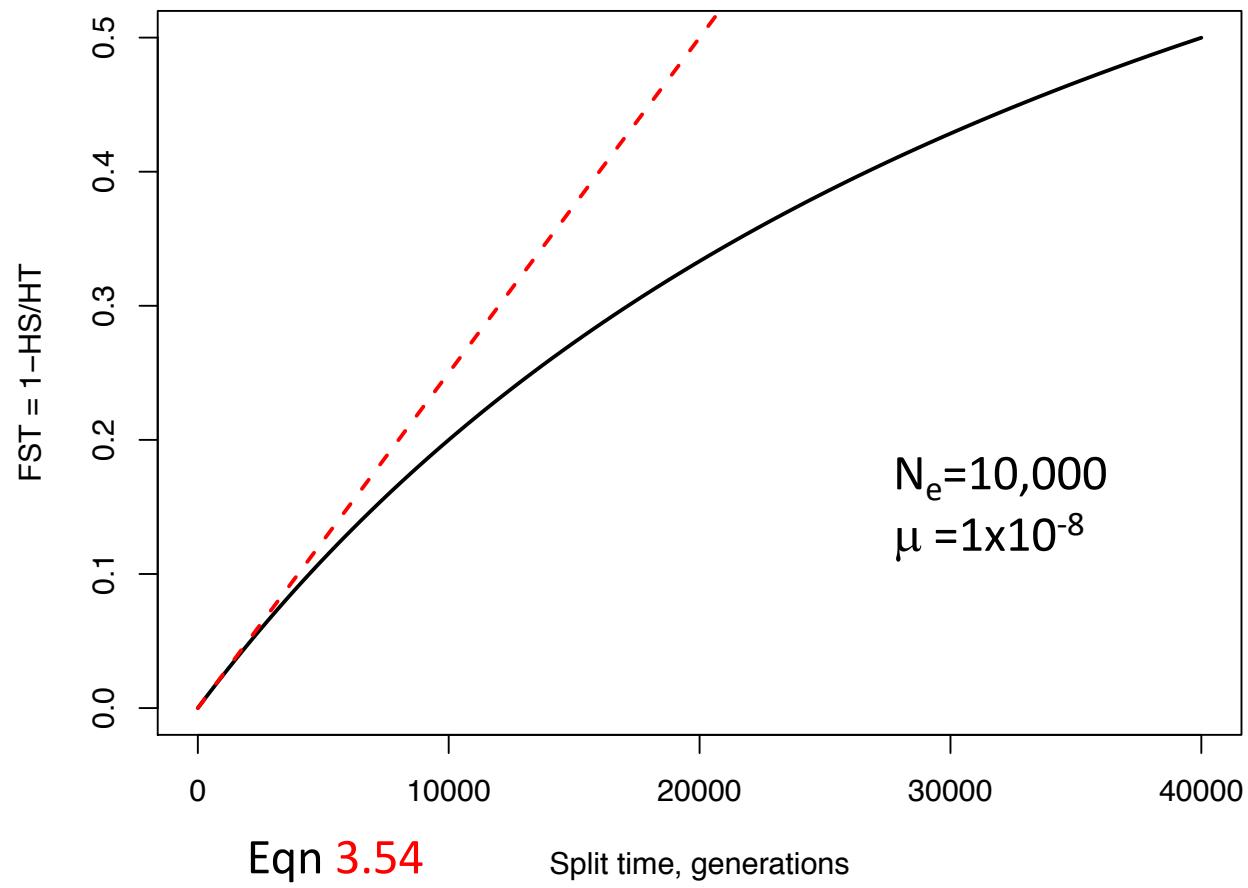


Barriers to migration amongst populations  
Allow independent drift and differentiation

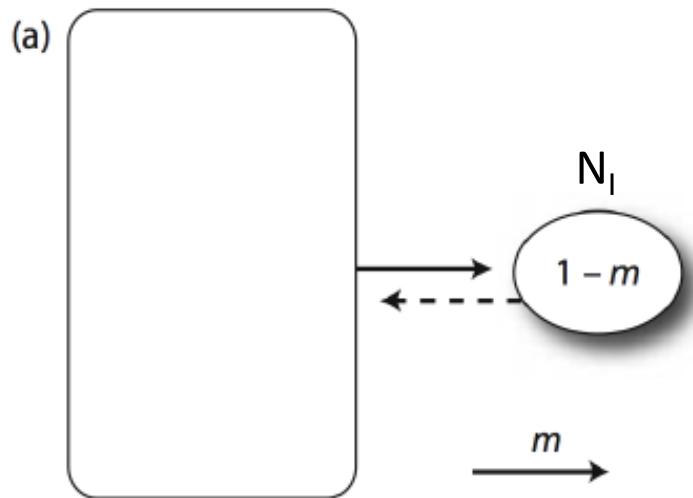
Lower rates of migration &/or smaller  
population sizes (higher rates of drift)



$$F_{ST} = 1 - H_S / H_T = (H_T - H_S) / H_T$$

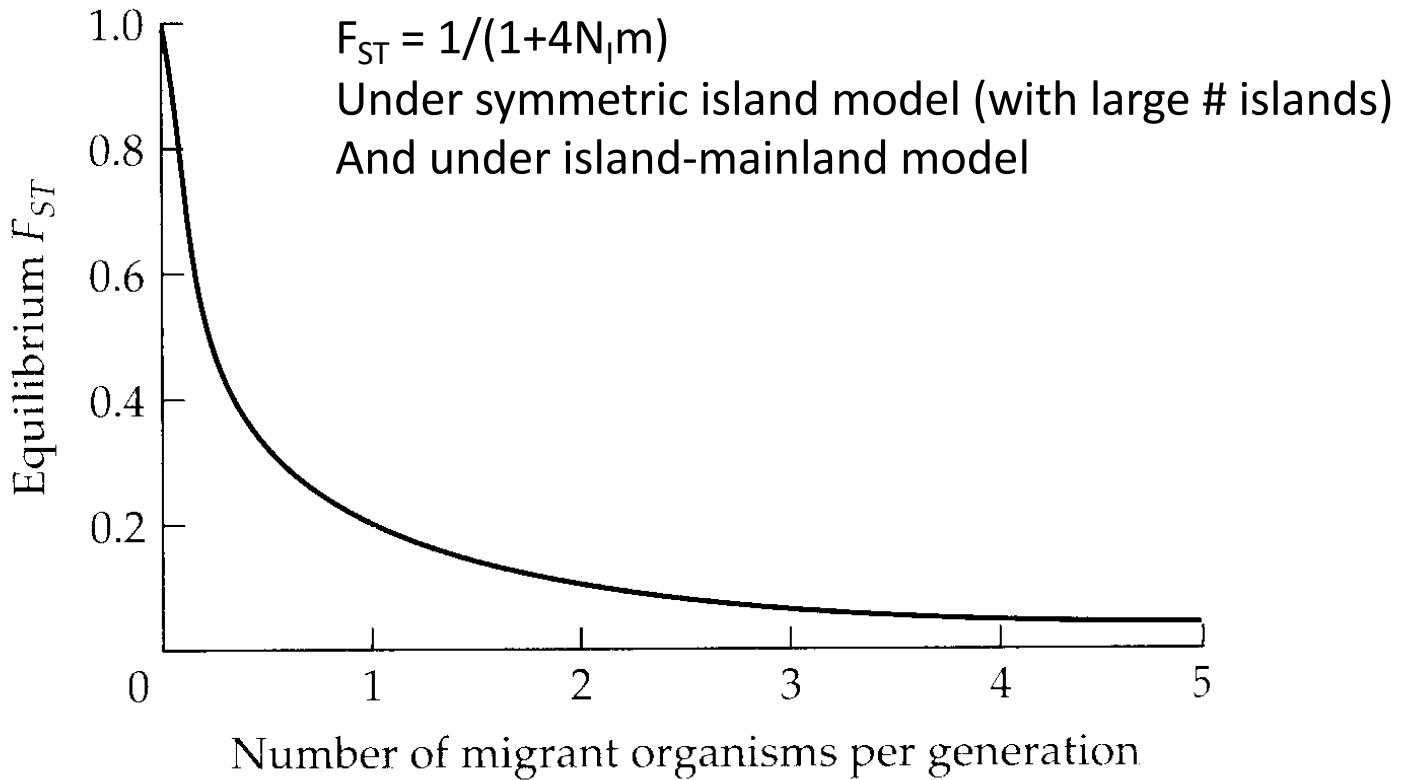


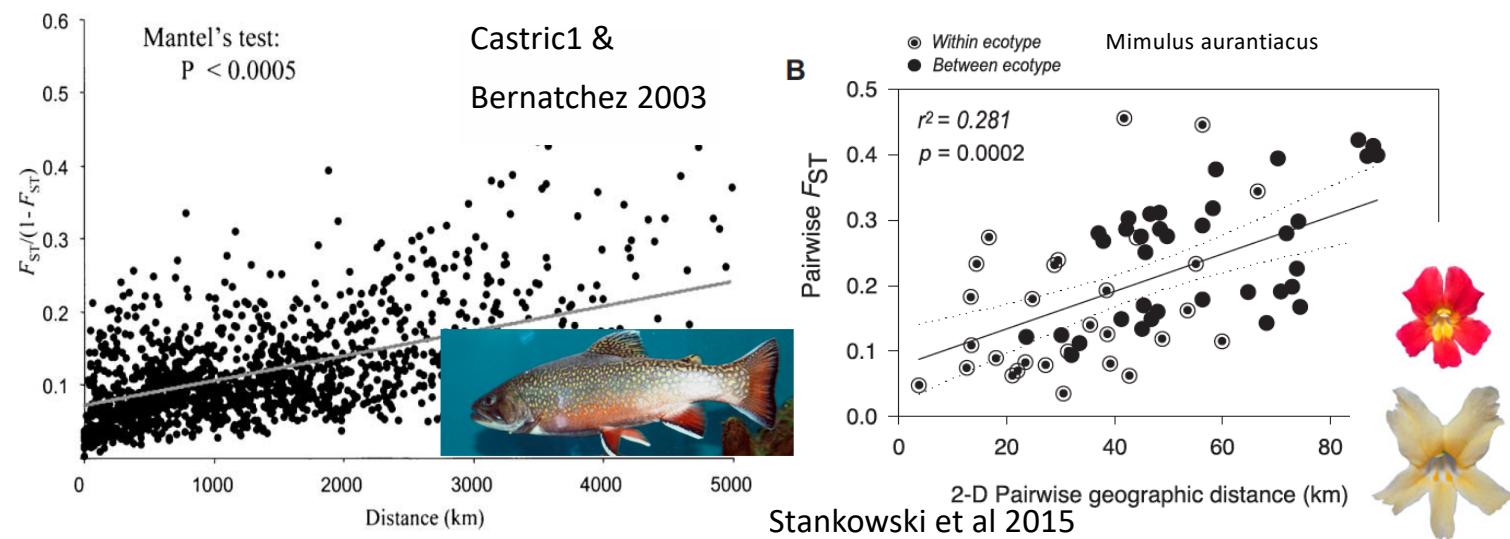
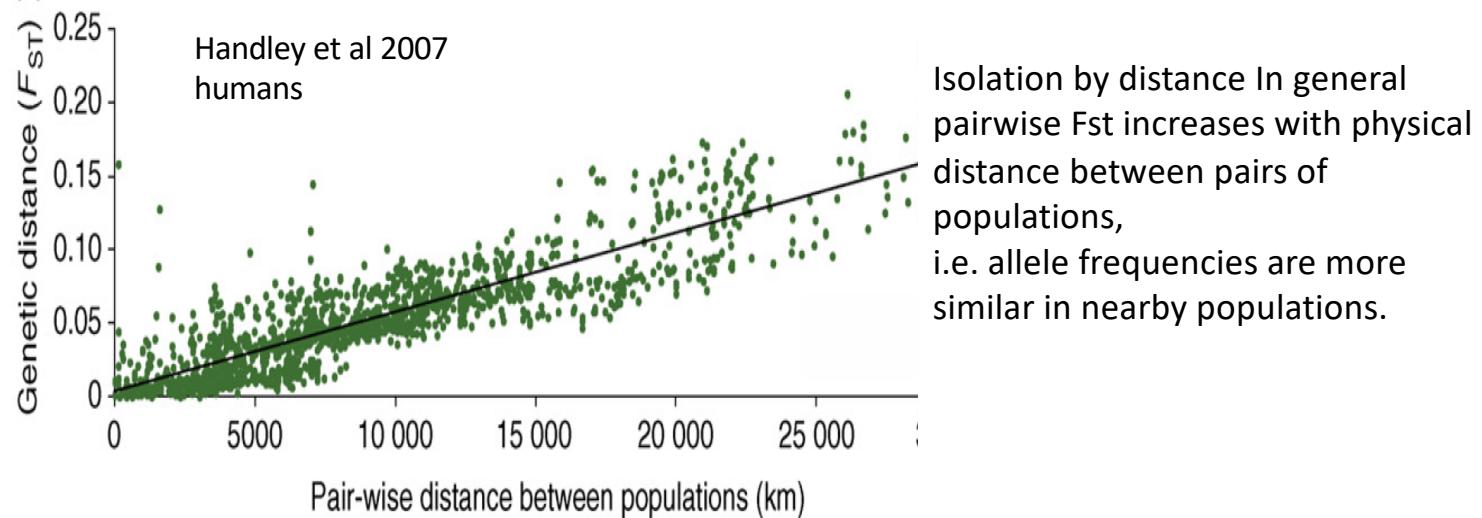
# Island-Mainland model

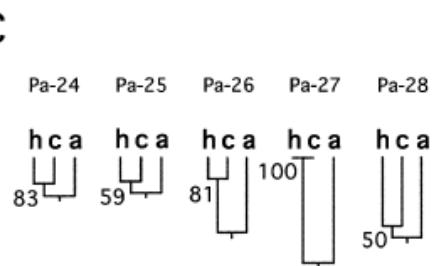
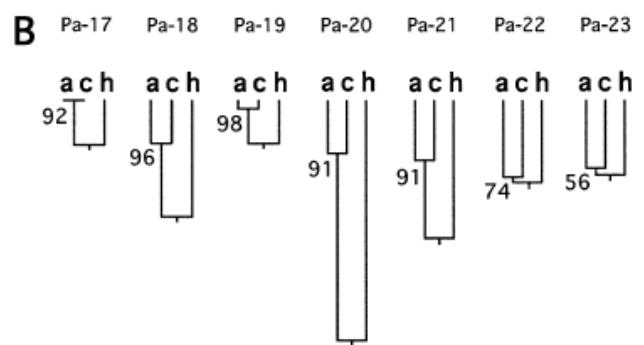
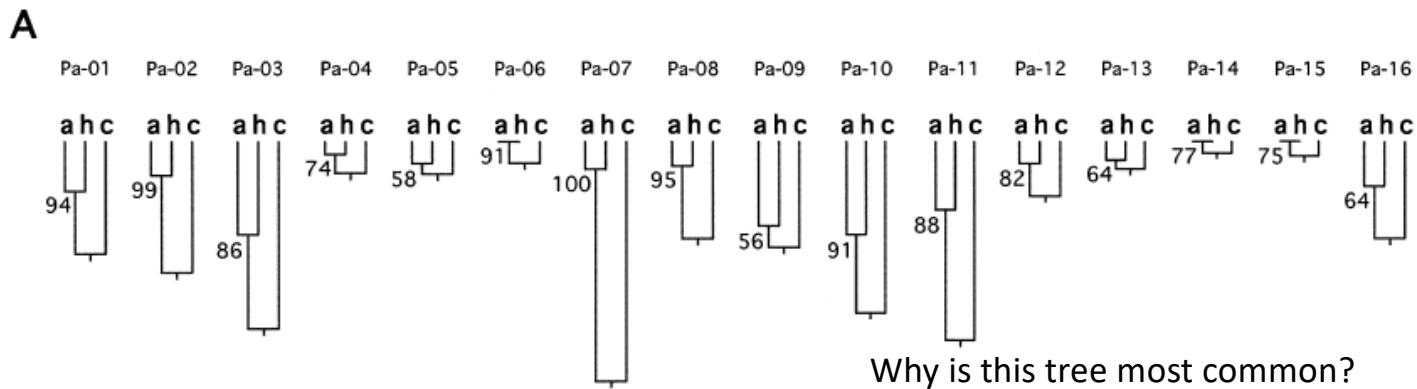


$$\begin{aligned} F_{ST} &= 1 - (H_{\text{island}} / H_T) \\ &= 1 / (1 + 4N_I m) \end{aligned}$$

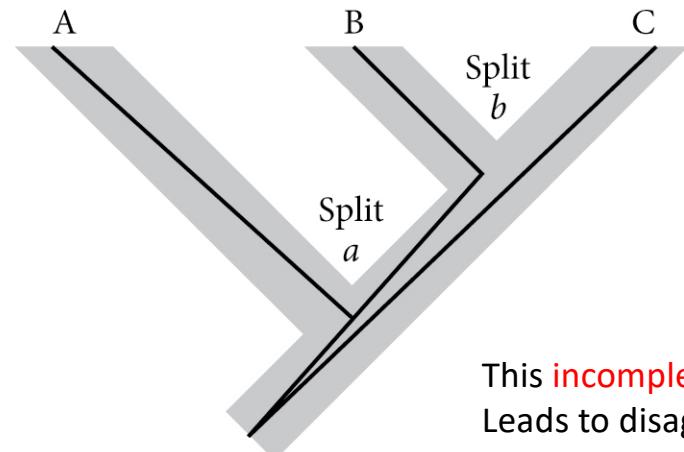
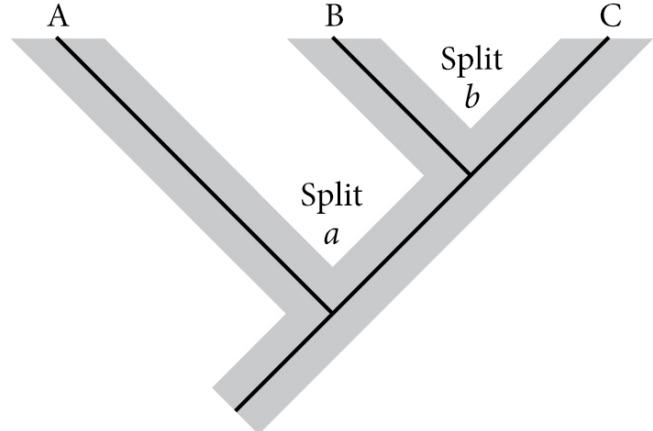
Eqn 3.59



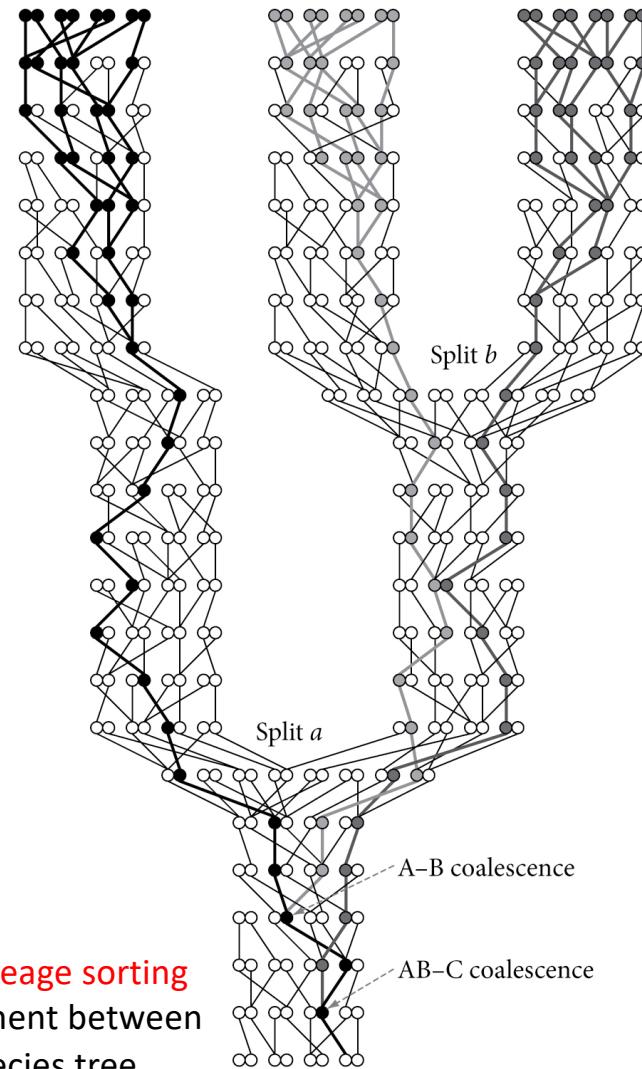




28 gene trees for 3  
AUSTRALIAN  
GRASS FINCHES (*POEPHILA*) Scale  
JENNINGS & EDWARDS 2005



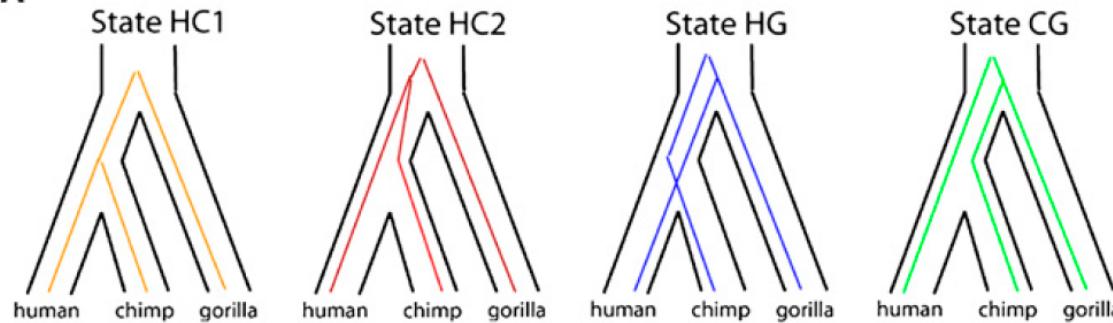
This **incomplete lineage sorting**  
Leads to disagreement between  
Gene trees and species tree.



Baum & Smith book

# Human-Chimp-Gorilla incomplete lineage sorting

**A**



Species tree:

Uninformative

Agrees

Two speciation events

$n = 3$



5 patterns

Disagrees

1	0	0	1	0
0	1	0	1	0
0	0	1	1	0

1 pattern

1

0

2 patterns

0

1

1

0

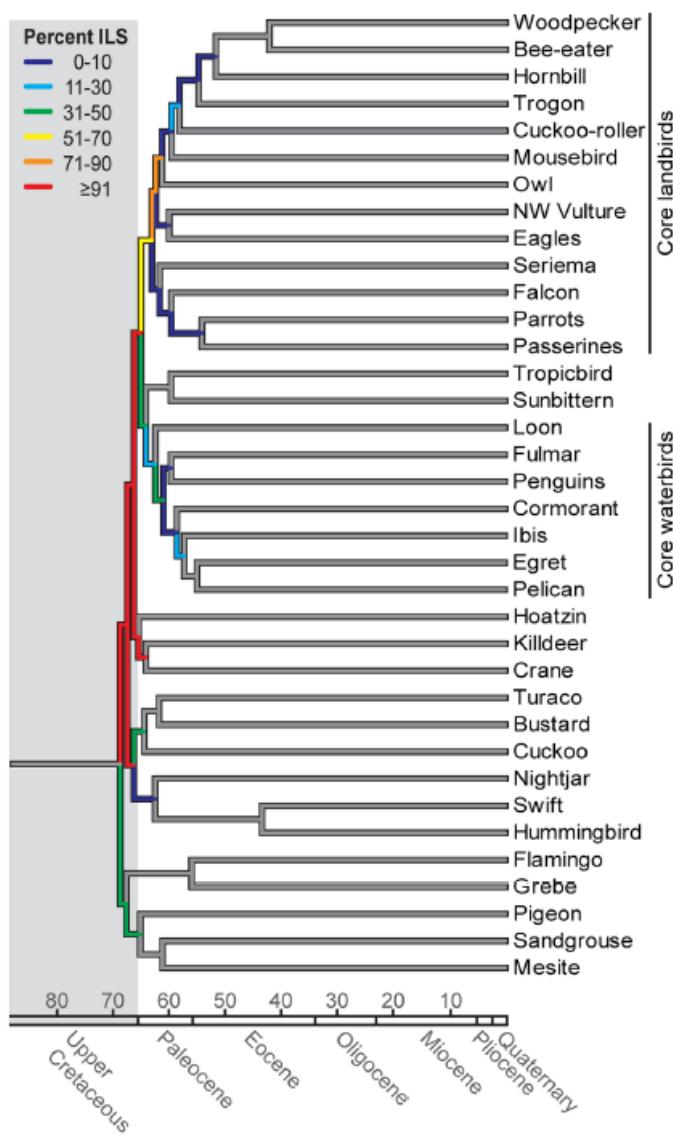
1

1

Project  
overview

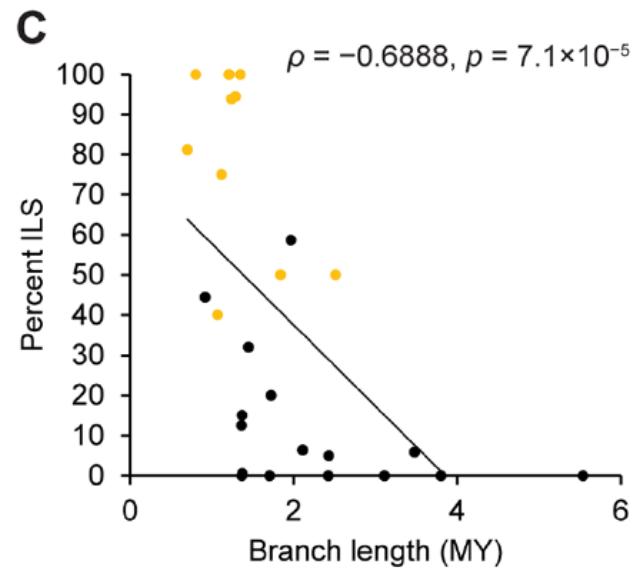


© Jon Fjeldså

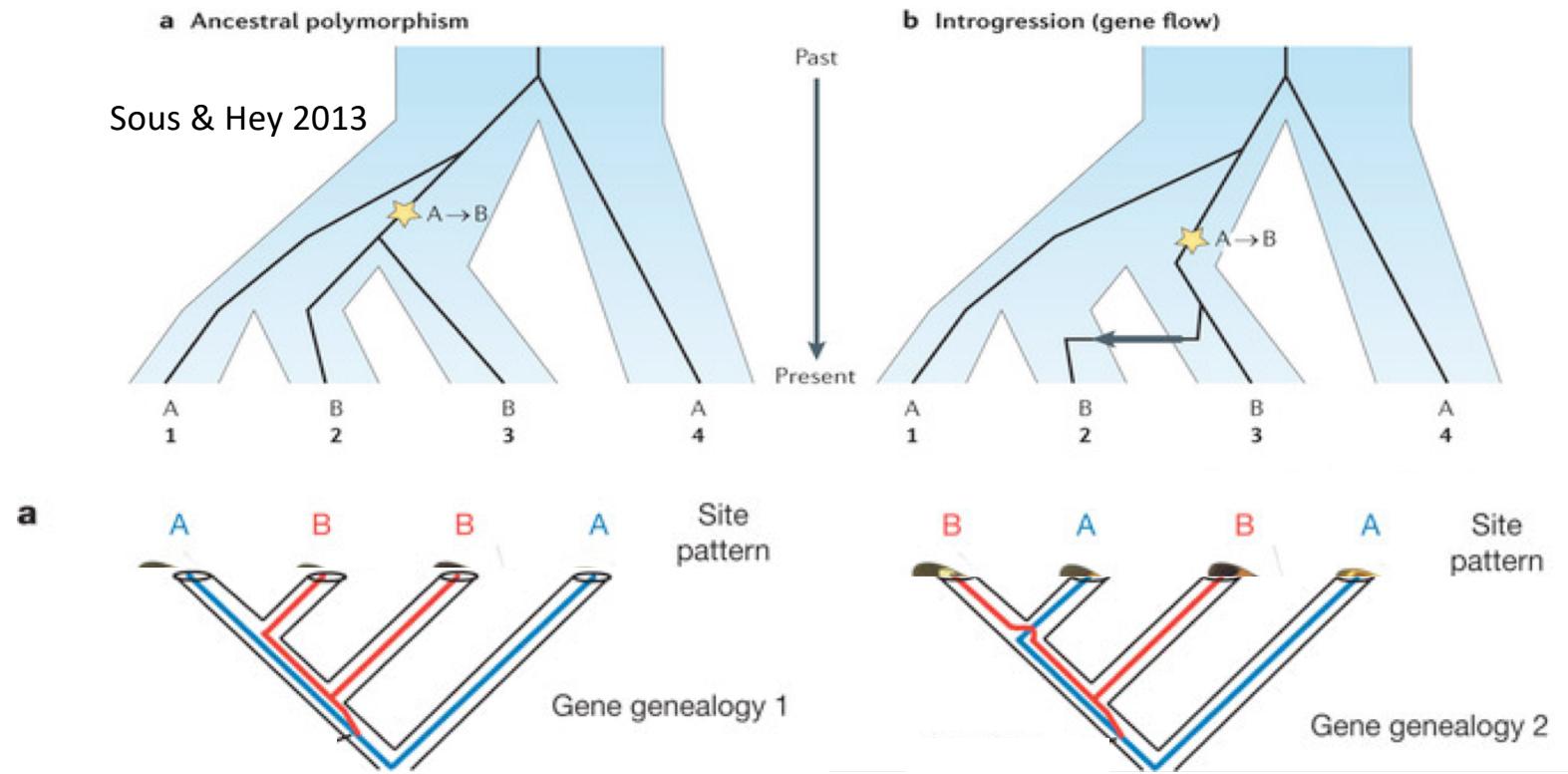


## The Dynamics of Incomplete Lineage Sorting across the Ancient Adaptive Radiation of Neoavian Birds

Alexander Suh\*, Linnéa Smeds, Hans Ellegren



Gene tree-species tree conflict can result from introgression or incomplete lineage sorting



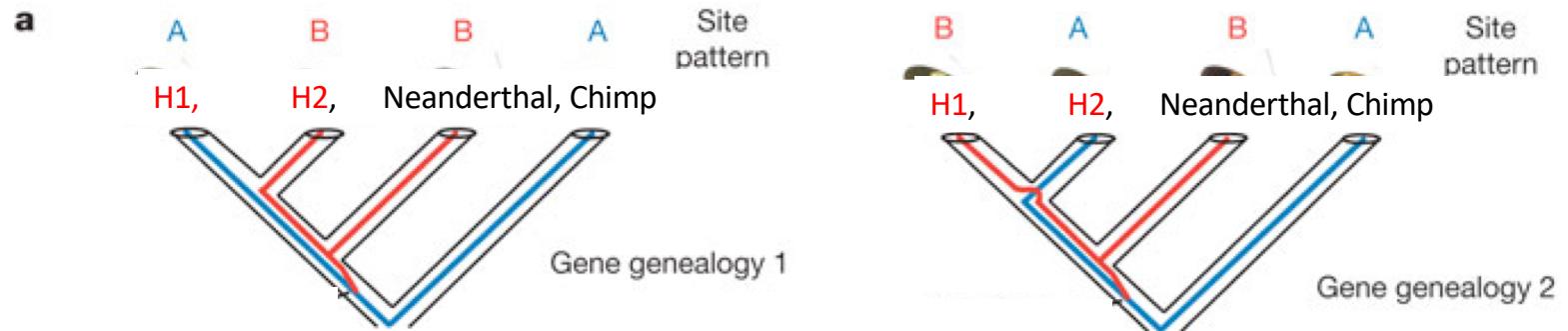
Distinguishing incomplete lineage sorting from introgression (ABBA-BABA statistic)

Under incomplete lineage sorting & no introgression

Number of ABBA loci = Number of BABA loci

$$D = (\#(BABA) - \#(ABBA))/C, \quad C = \#(ABBA) + \#(BABA)$$

Green et al 2010, Nature



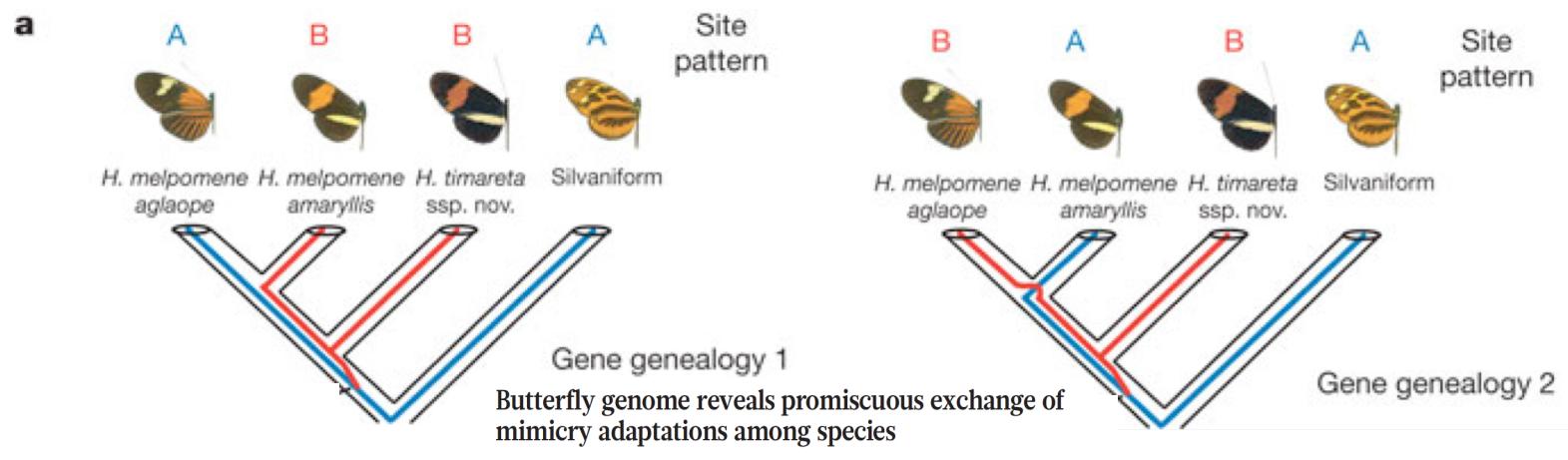
Under incomplete lineage sorting & no introgression  
 Number of ABBA loci = Number of BABA loci  
 $D = (\#(BABA) - \#(ABBA))/C; C = \#(ABBA) + \#(BABA)$

### Genetic history of an archaic hominin group from Denisova Cave in Siberia

David Reich<sup>1,2\*</sup>, Richard E. Green<sup>3,4\*</sup>, Martin Kircher<sup>5\*</sup>, Johannes Krause<sup>3,5\*</sup>, Nick Patterson<sup>6\*</sup>, Eric V. Durand<sup>6\*</sup>, Bence Viola<sup>3,7\*</sup>, Adrian W. Briggs<sup>3,7</sup>, Udo Stenzel<sup>3</sup>, Philipp L. F. Johnson<sup>8</sup>, Tomislav Maricic<sup>9</sup>, Jeffrey M. Good<sup>10</sup>, Tomas Marques-Bonet<sup>10,11</sup>, Can Alkan<sup>10</sup>, Qiaomei Fu<sup>3,12</sup>, Swapna Mallick<sup>1,2</sup>, Heng Li<sup>13</sup>, Matthias Meyer<sup>14</sup>, Evan E. Eichler<sup>15</sup>, Mark Stoneking<sup>16</sup>, Michael Richards<sup>17,18</sup>, Sahra Talamo<sup>19</sup>, Michael V. Shunkov<sup>14</sup>, Anatoli P. Derevianko<sup>14</sup>, Jean-Jacques Hublin<sup>19</sup>, Janet Kelso<sup>3</sup>, Montgomery Slatkin<sup>6</sup> & Svante Pääbo<sup>3</sup>

**Table 1 | Sharing of derived alleles between present-day and archaic hominins**

Sample H <sub>1</sub>	Sample H <sub>2</sub>	D(H <sub>1</sub> , H <sub>2</sub> , Neanderthal, chimpanzee)					D(H <sub>1</sub> , H <sub>2</sub> , Denisova, chimpanzee)					Z-score	
		n <sub>BABA</sub>	n <sub>ABBA</sub>	D (%)	s.e. (%)	Z-score	n <sub>BABA</sub>	n <sub>ABBA</sub>	D (%)	s.e. (%)	Z-score		
<b>African/African*</b>													
San	Yoruba	23,690	23,855	-0.3	0.6	-0.6	39,042	39,019	0.0	0.5	0.1		0.1
<b>Eurasian/African*</b>													
French	San	25,242	22,982	4.7	0.6	7.6†	39,838	38,495	1.7	0.5	3.4†		3.4†
French	Yoruba	21,794	19,890	4.6	0.7	6.9†	34,262	33,078	1.8	0.5	3.6†		3.6†
Han	San	25,081	22,470	5.5	0.6	8.5†	38,815	37,439	1.8	0.5	3.4†		3.4†
Han	Yoruba	21,741	19,412	5.7	0.7	7.9†	33,182	32,184	1.5	0.5	2.8		2.8
<b>Eurasian/Melanesian*</b>													
French	Papuan	15,523	15,548	-0.1	0.8	-0.1	23,509	25,470	-4.0	0.7	-5.7†		-5.7†
Han	Papuan	15,059	14,677	1.3	0.9	1.5	22,262	24,198	-4.2	0.7	-5.8†		-5.8†
<b>Melanesian/African*</b>													
Papuan1	San	21,985	20,366	3.8	0.7	5.1†	35,923	32,841	4.5	0.6	7.2†		7.2†
Papuan1	Yoruba	19,107	17,646	4.0	0.8	4.9†	30,995	28,186	4.7	0.6	7.4†		7.4†



The Heliconius Genome Consortium\*

### Under incomplete lineage sorting

$$D = (\#(BABA) - \#(ABBA))/C$$

$$D = 0.037 \pm 0.003 \text{ (two-tailed } D = 0, P = 1 \times 10^{-40})$$

