

# Archaic human introgression

**Population genetics course**

**Martin Petr**

[mp@bodkan.net](mailto:mp@bodkan.net)

February 28, 2024

# 1856: Discovery of a ‘new man’



<https://www.donsmaps.com/neanderthaloriginal.html>



<https://twitter.com/Qafzeh/status/805339276334333953>

# 1856: Discovery of a ‘new man’



<https://www.donsmaps.com/neanderthaloriginal.html>



<https://twitter.com/Qafzeh/status/805339276334333953>

**Neander's valley**

# 1856: Discovery of a ‘new man’



<https://www.donsmaps.com/neanderthaloriginal.html>



<https://twitter.com/Qafzeh/status/805339276334333953>

**Neander's valley**  
valley = Thal in German

# 1856: Discovery of a ‘new man’



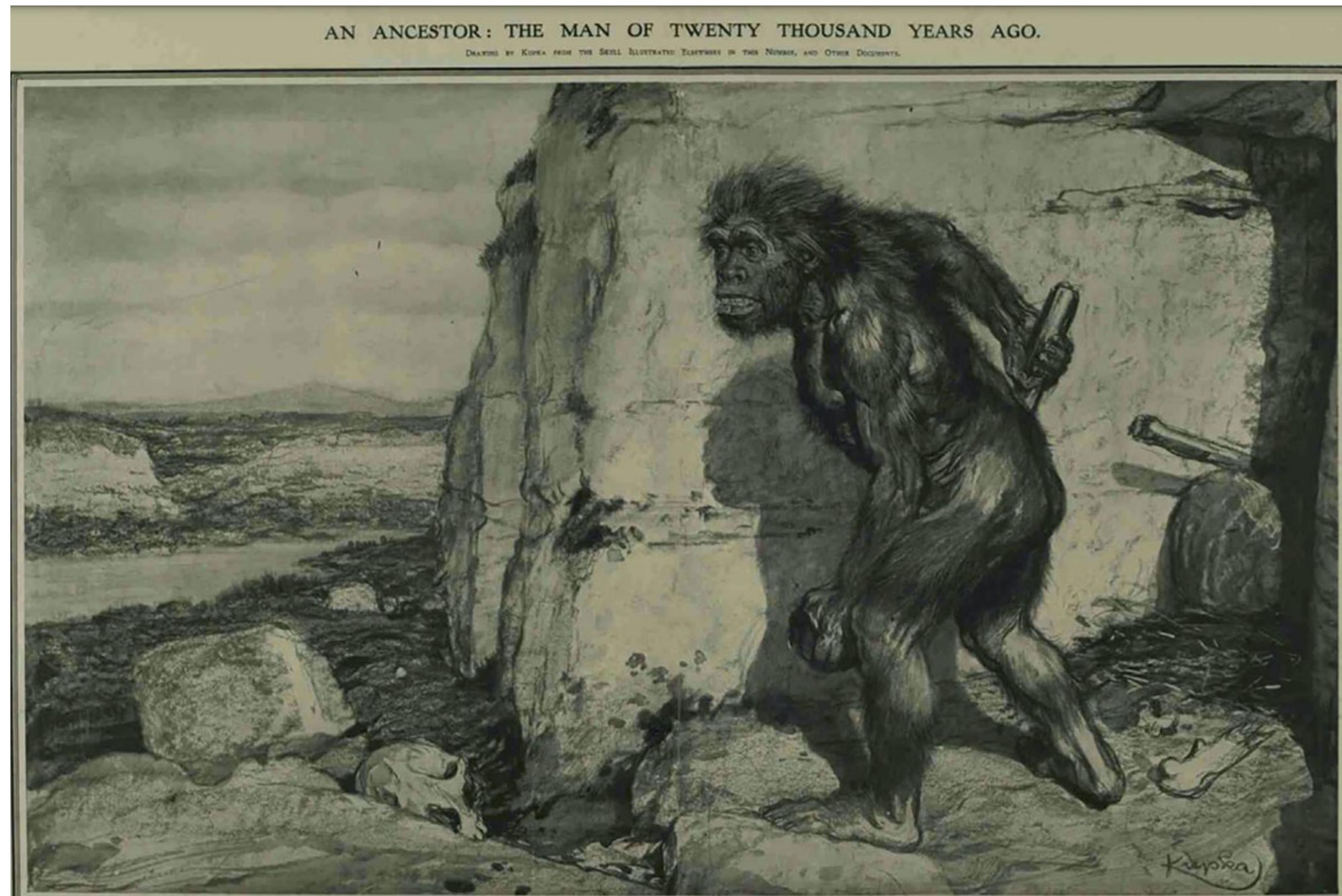
<https://www.donsmaps.com/neanderthaloriginal.html>

**Neander's valley**  
valley = Thal in German



<https://twitter.com/Qafzeh/status/805339276334333953>

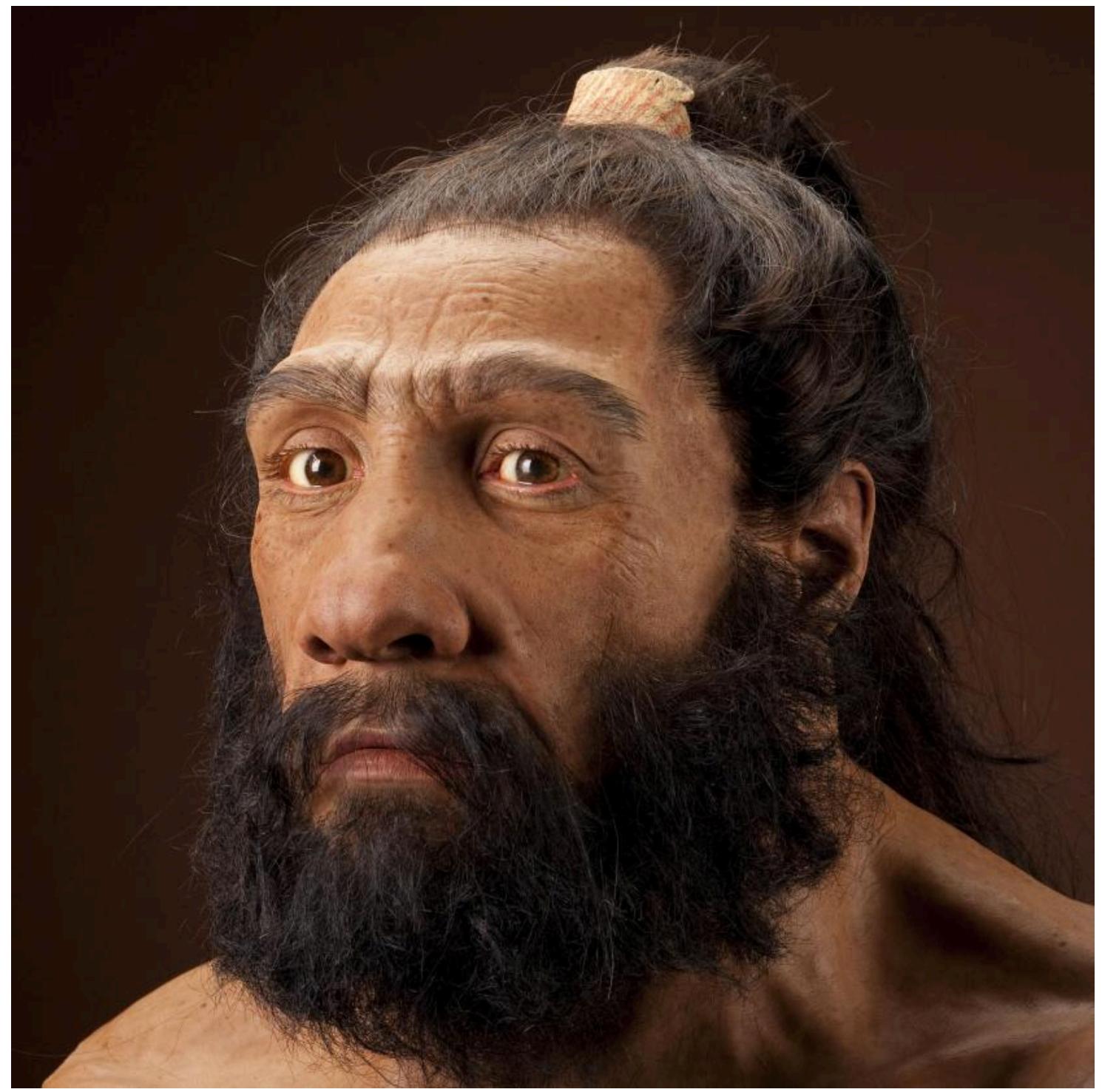
**Neanderthal**



The Illustrated London News, 1909



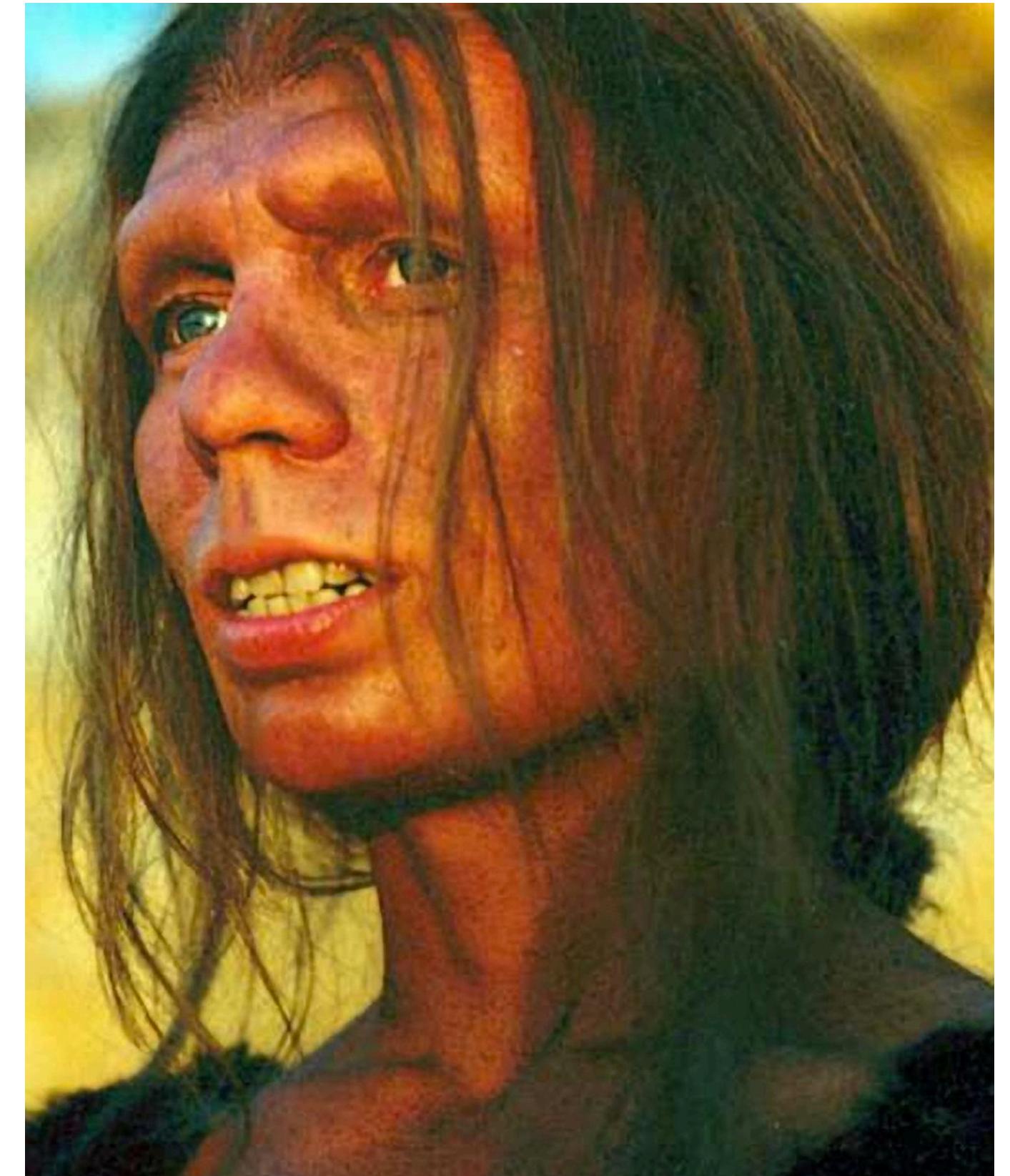
The Neanderthal Man movie poster, 1953



<http://humanorigins.si.edu/evidence/human-fossils/species/homo-neanderthalensis>



<https://www.neanderthal.de>



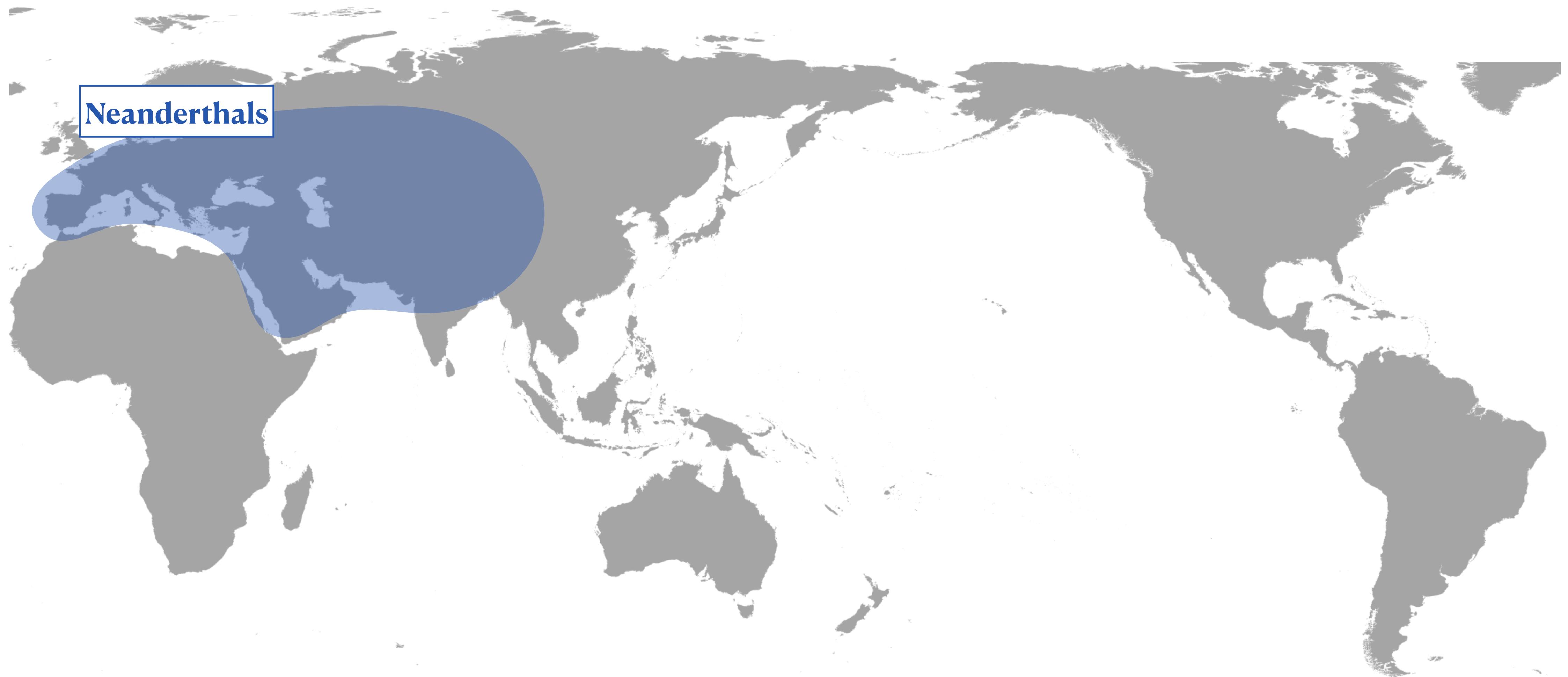
<https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.0020449>



Neanderthal Museum, Mettmann, Germany

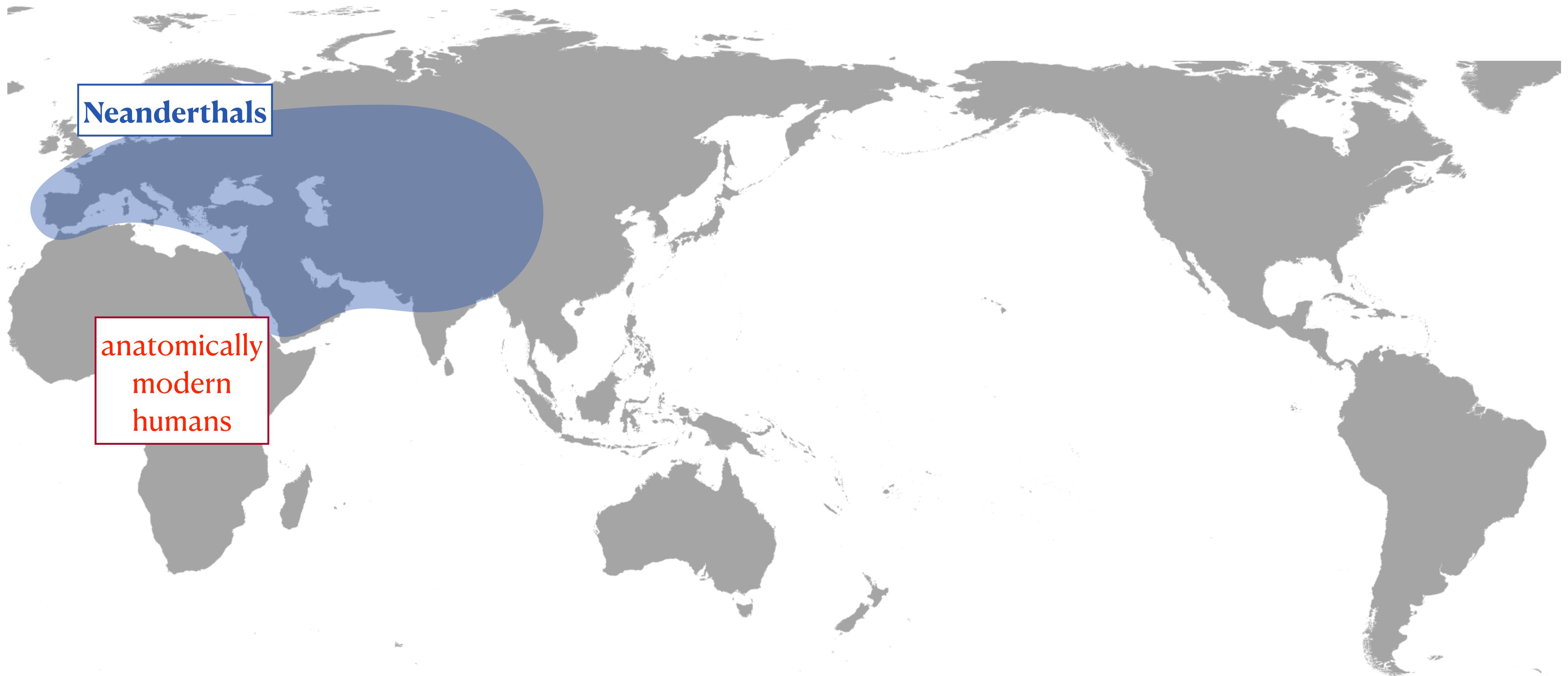


# ~600 kya B.P.: Neanderthals in Eurasia



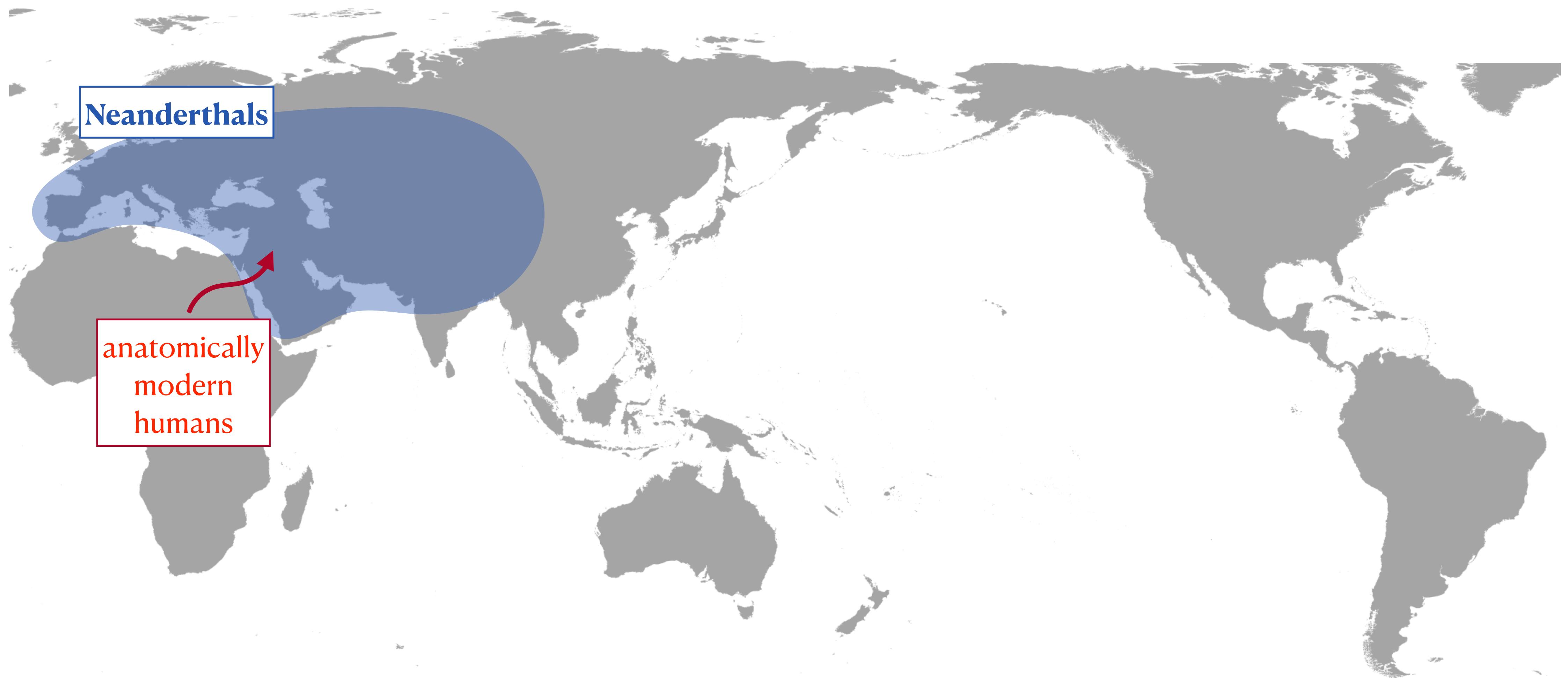
B.P. = "before present"

# **~300 kya B.P.: anatomically modern humans in Africa**



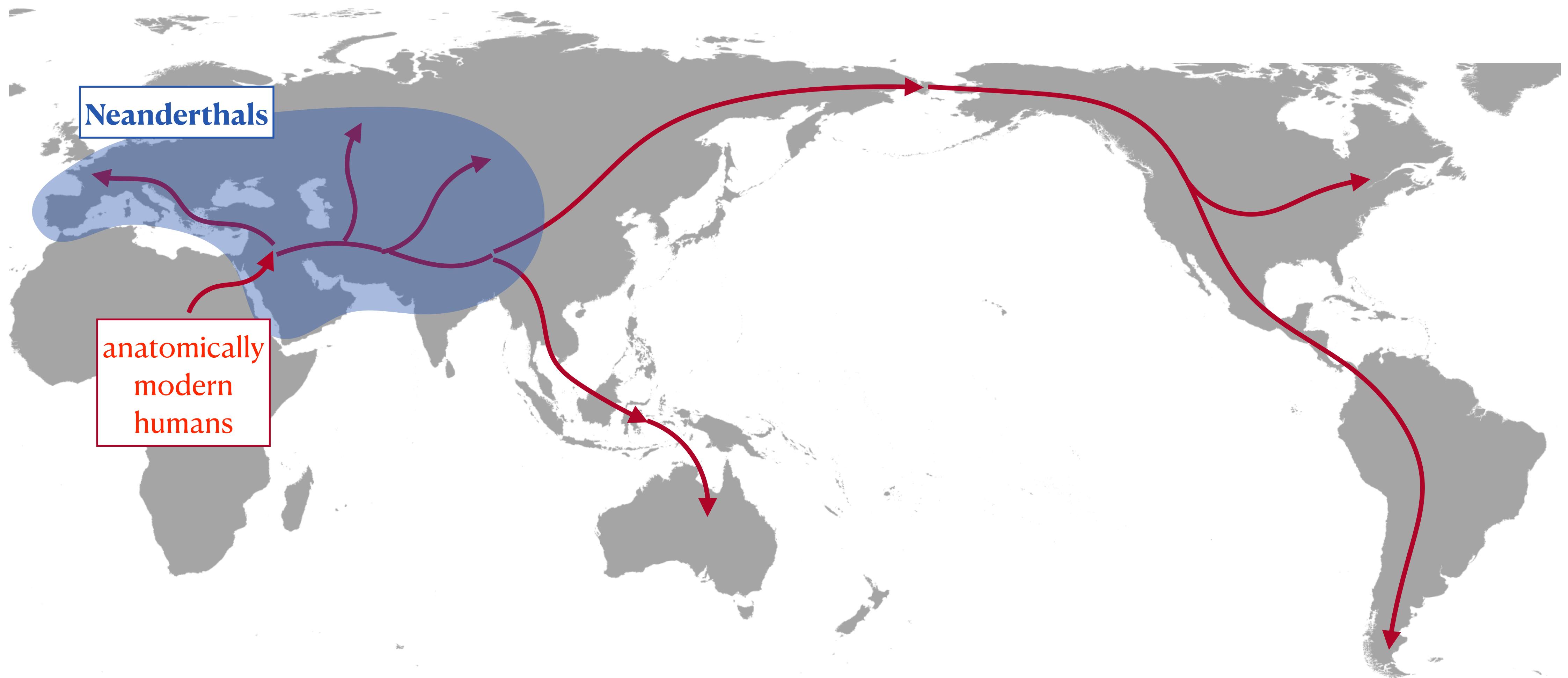
B.P. = "before present"

# **~70 kya B.P.: modern humans migrated out of Africa**



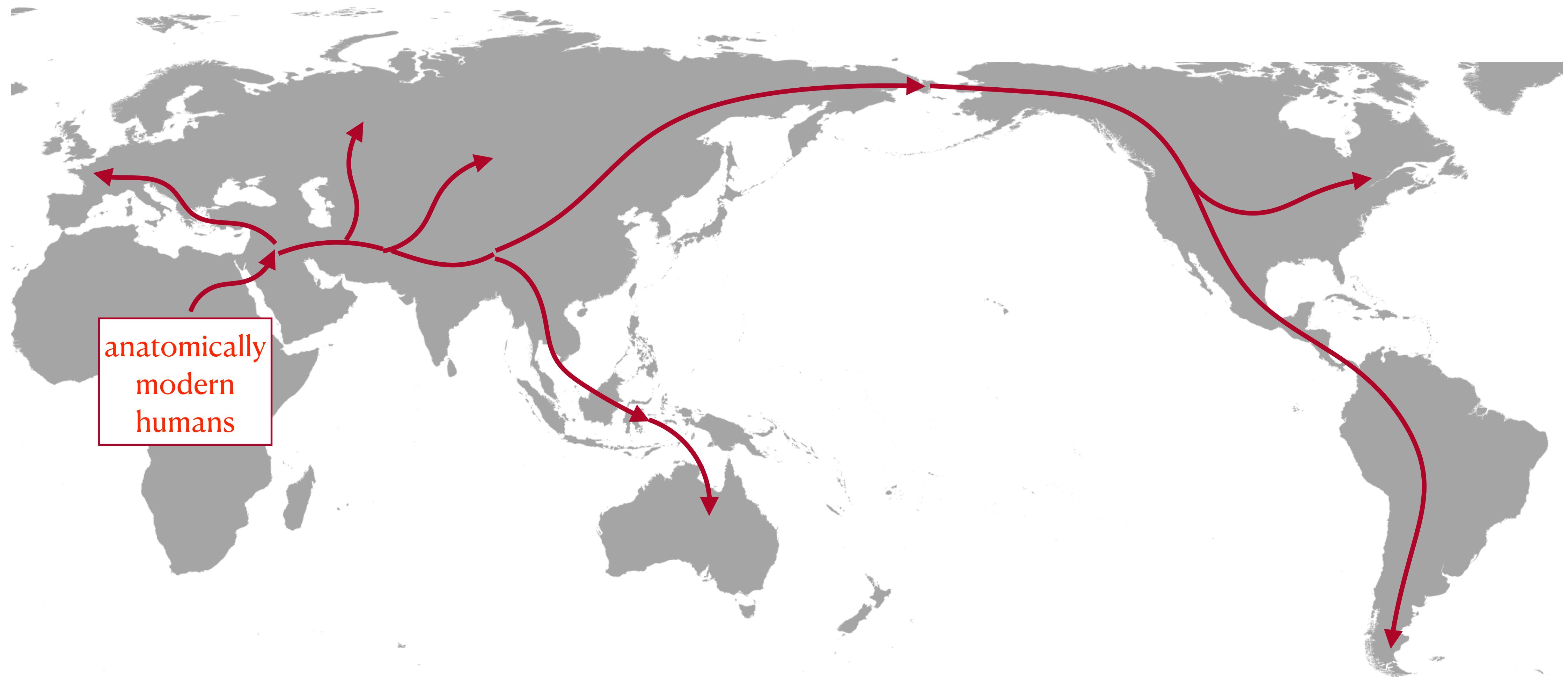
B.P. = "before present"

# **~70 kya B.P.: modern humans migrated out of Africa ...and colonized the entire world**



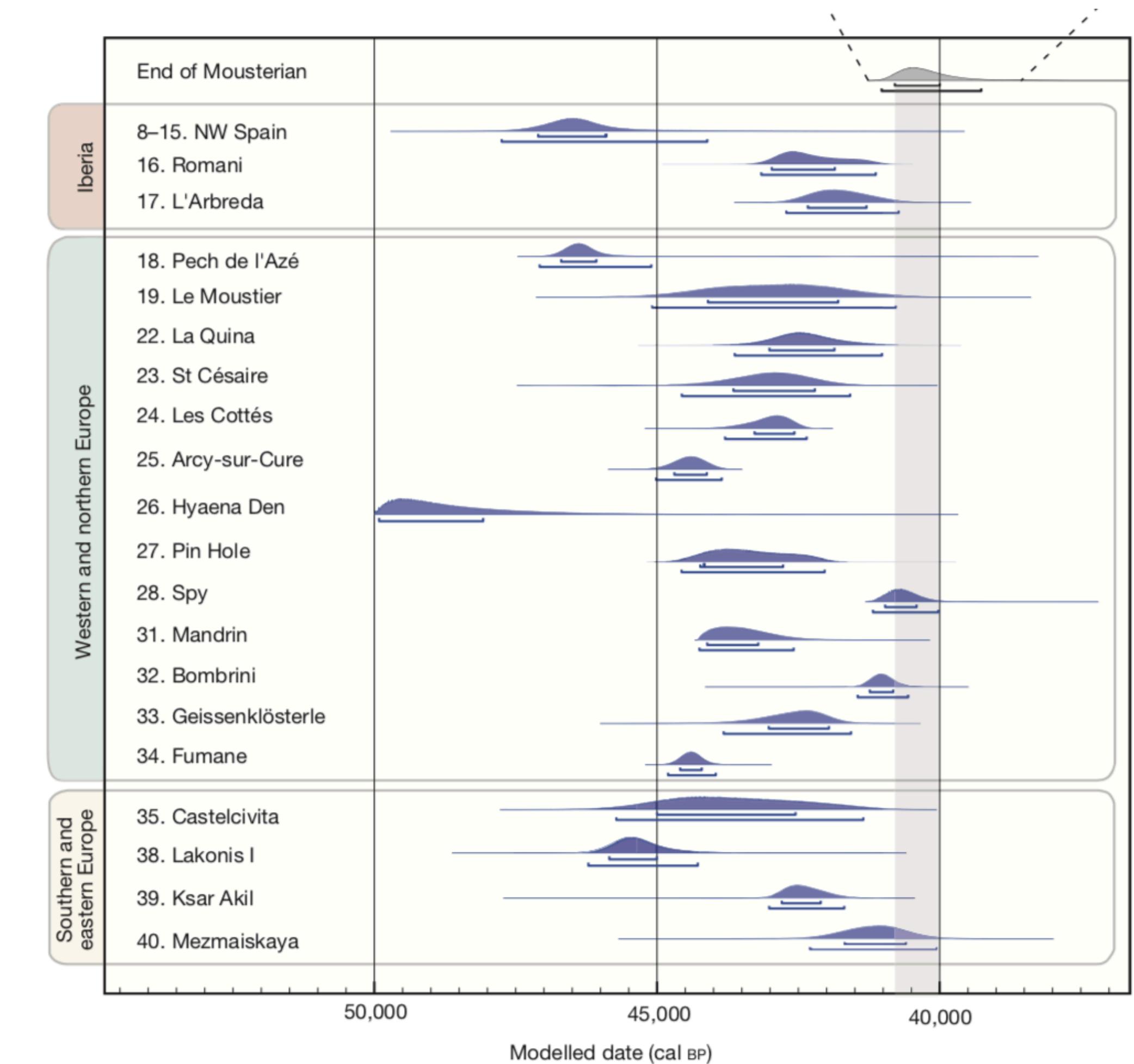
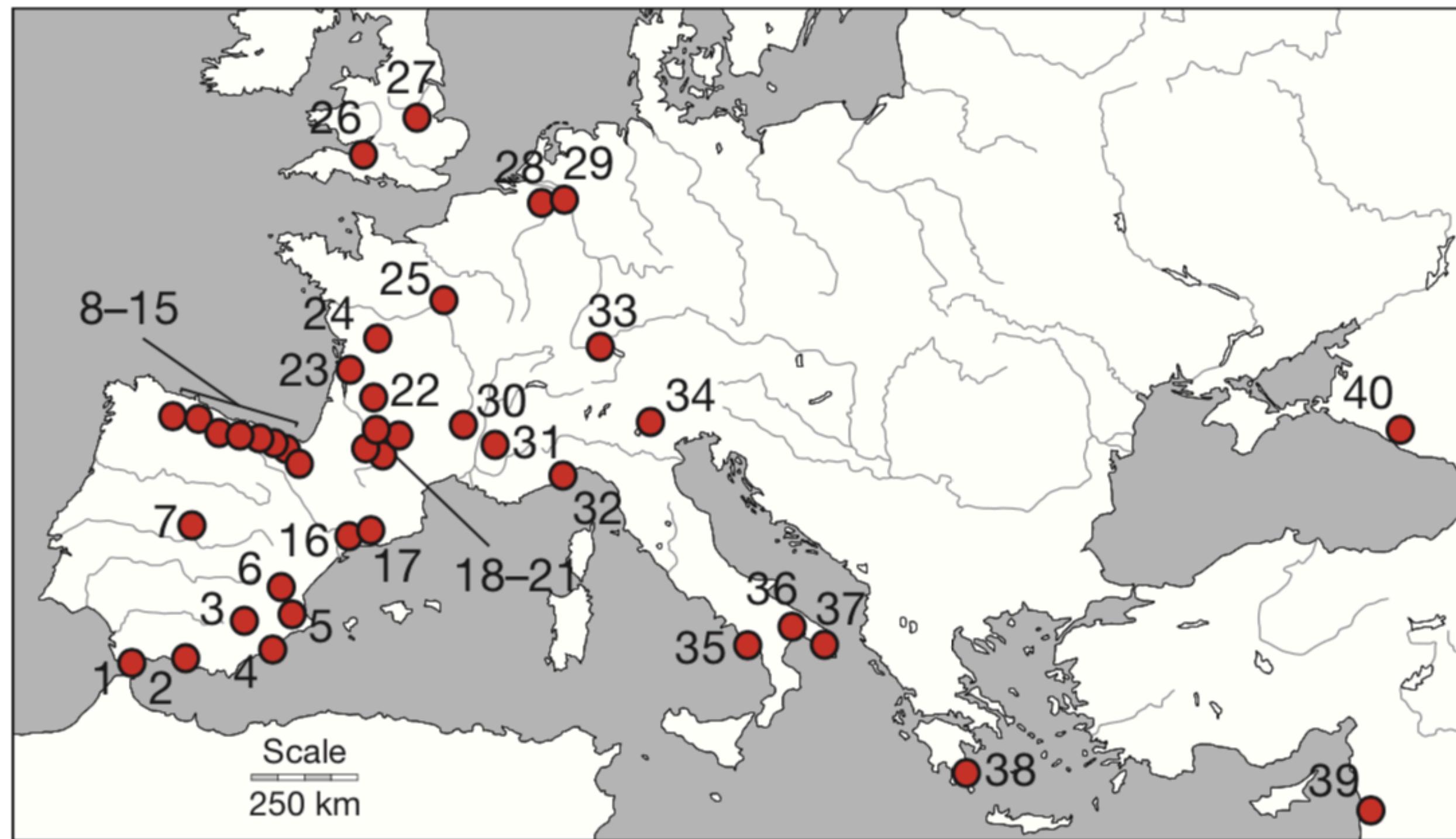
B.P. = "before present"

# $\sim$ 40 kya: Neanderthals vanished

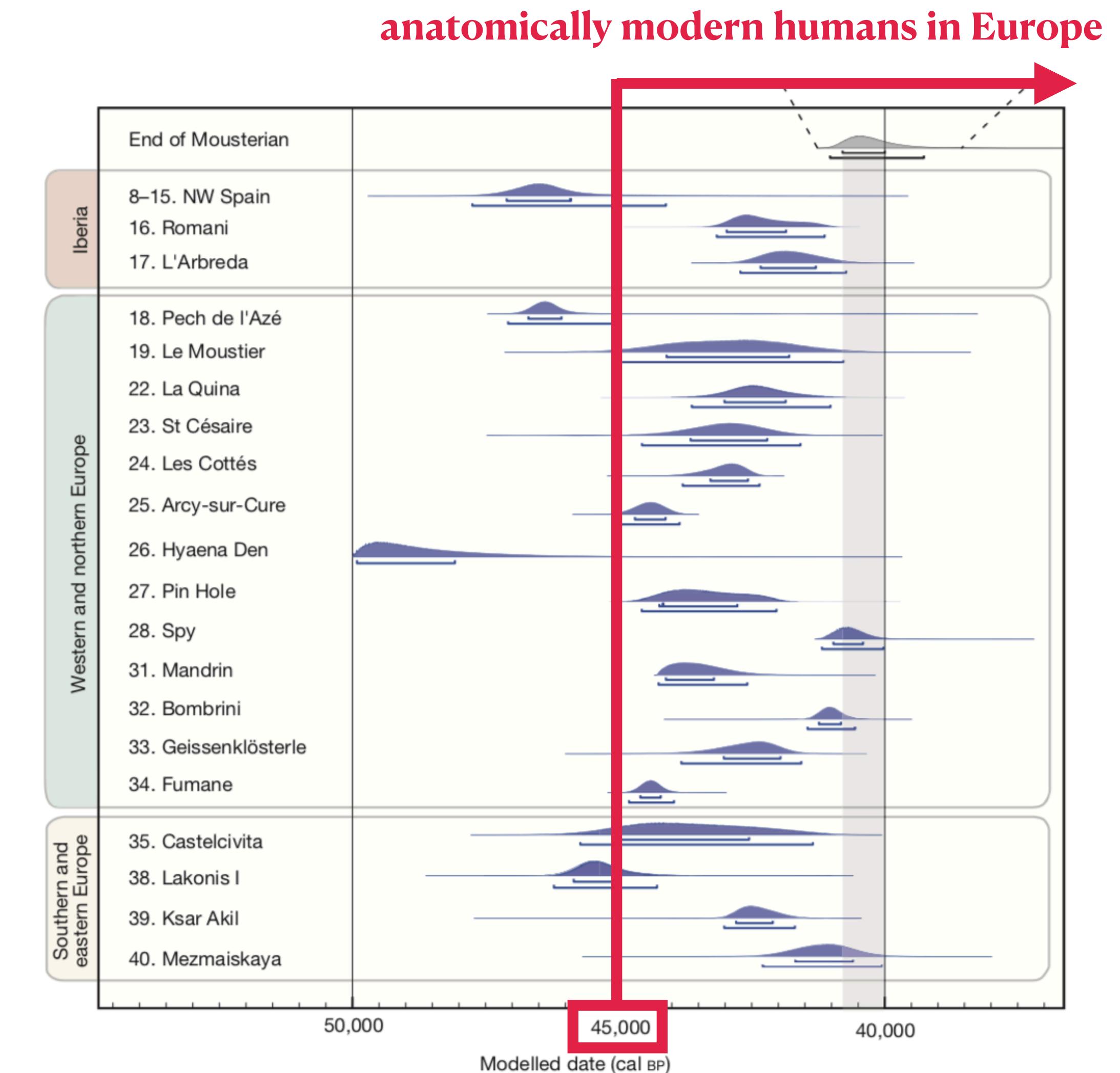
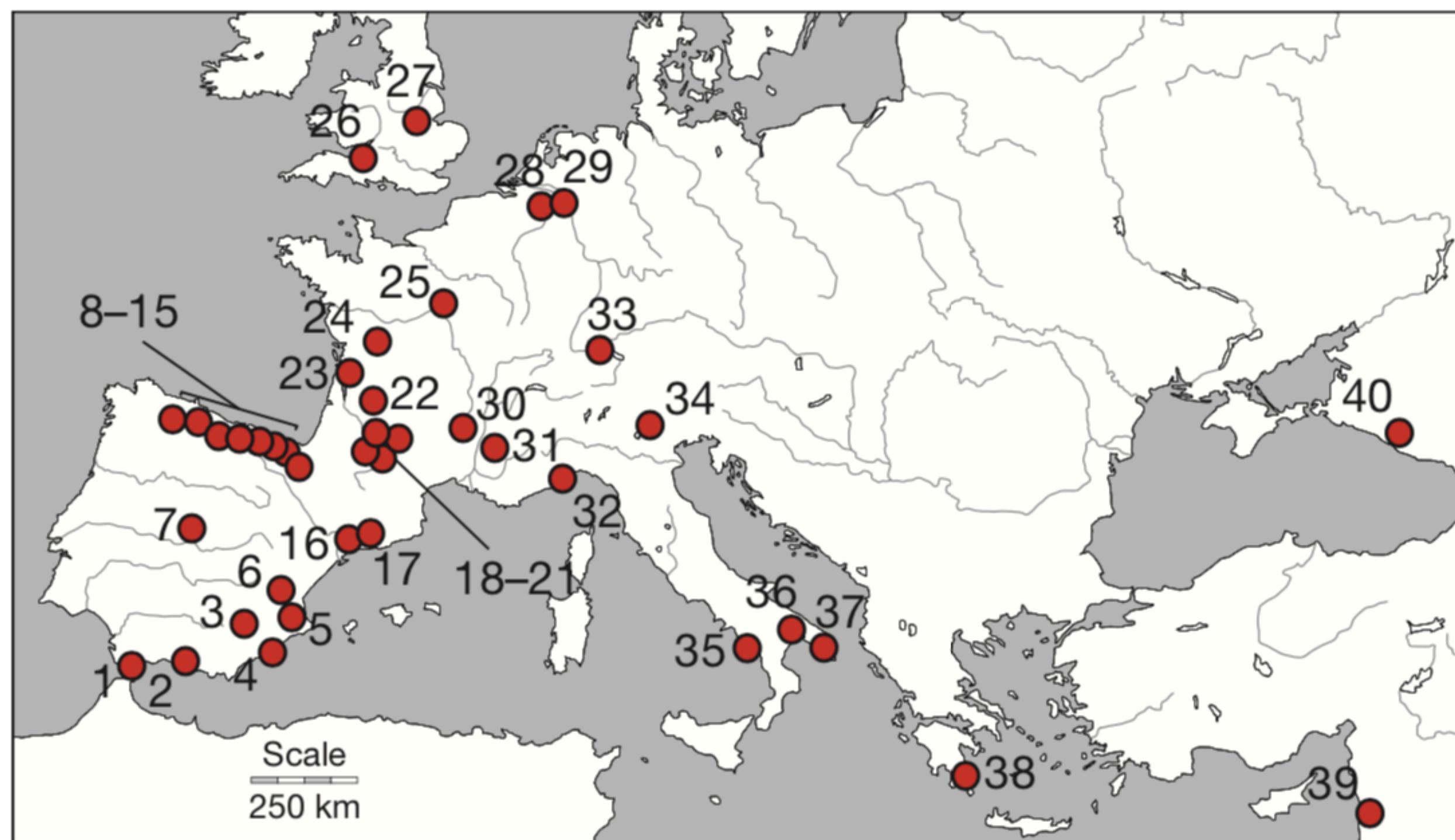


B.P. = "before present"

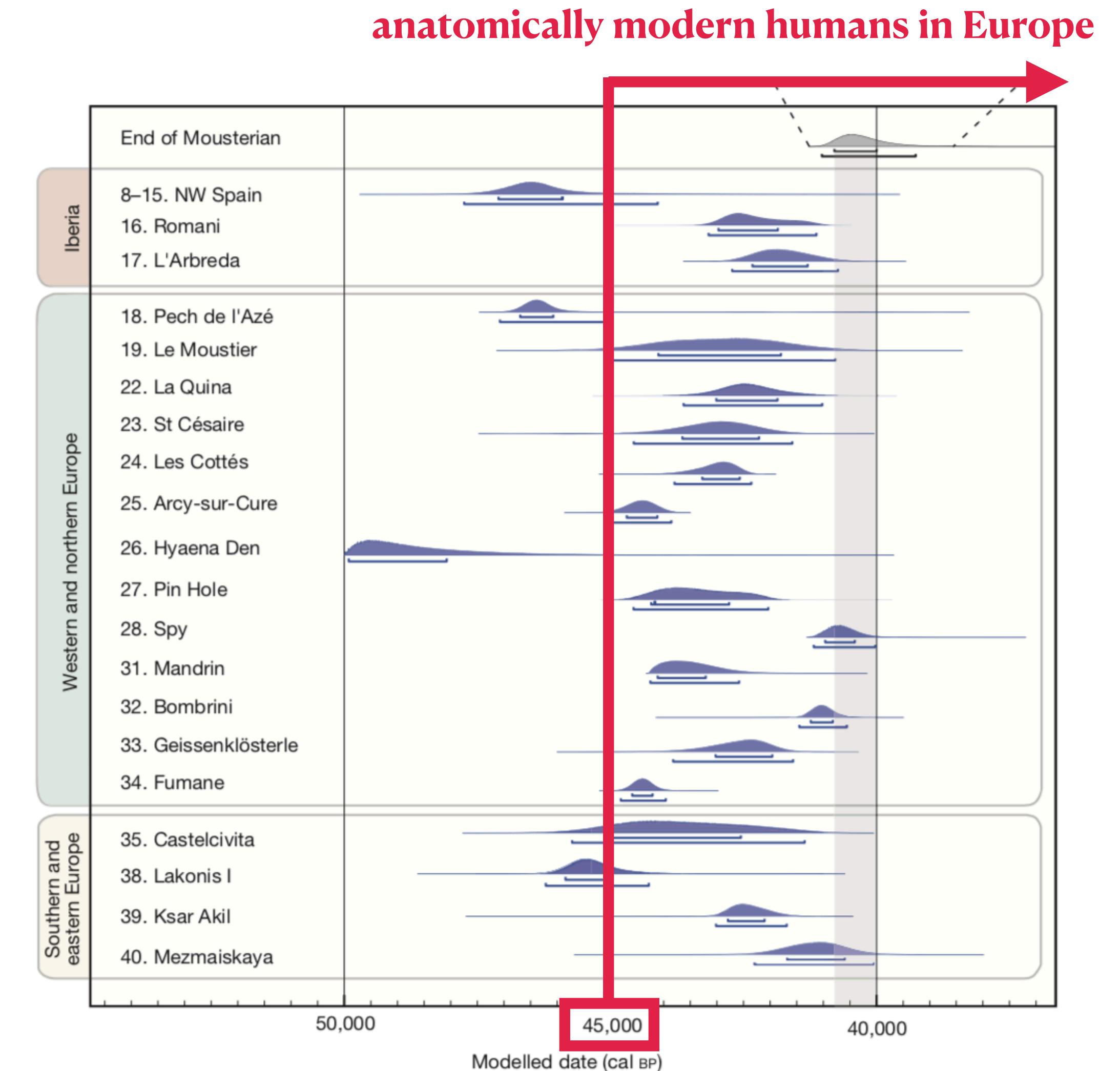
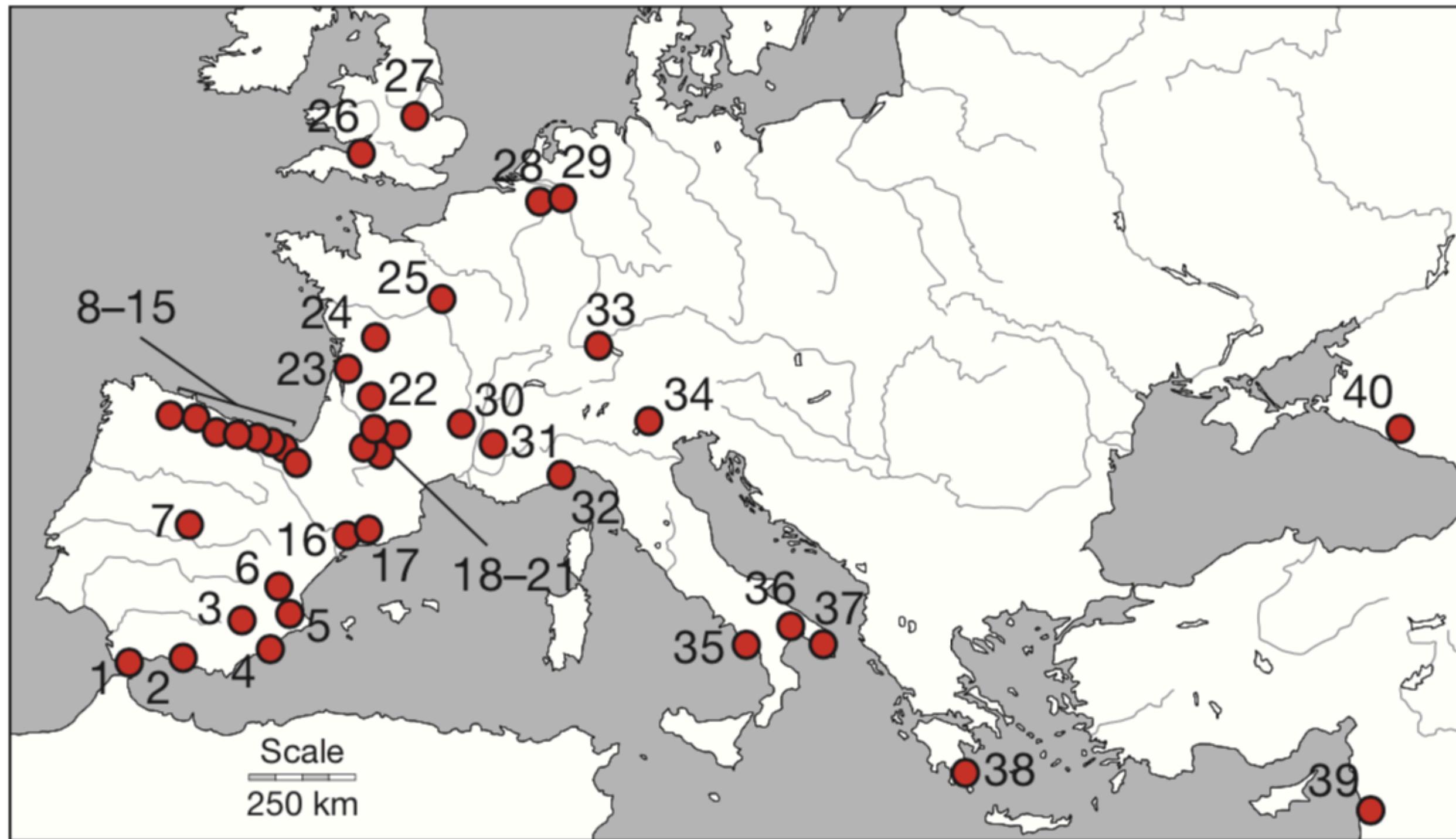
# ~40 kya: Neanderthals vanished



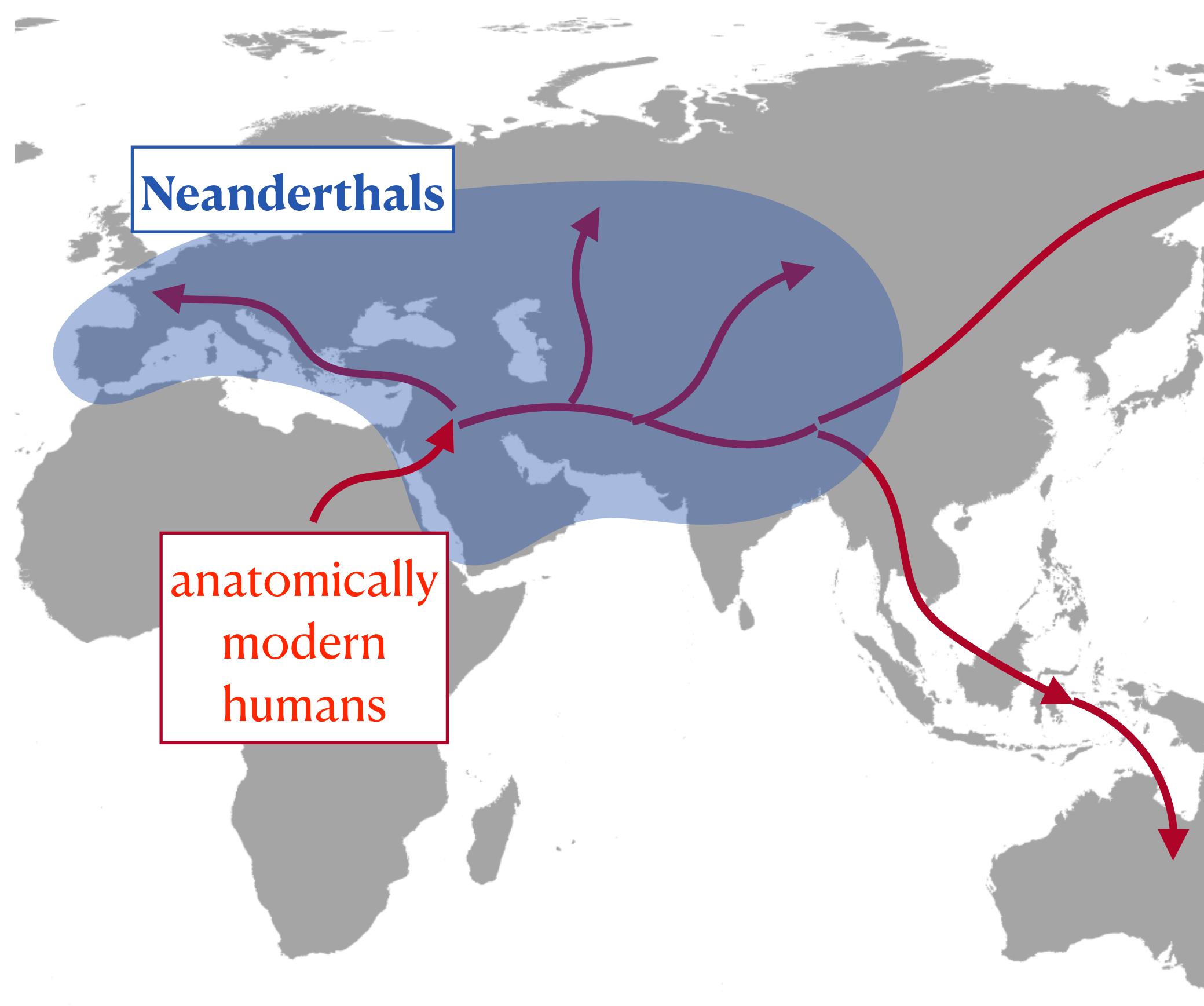
# ~40 kya: Neanderthals vanished



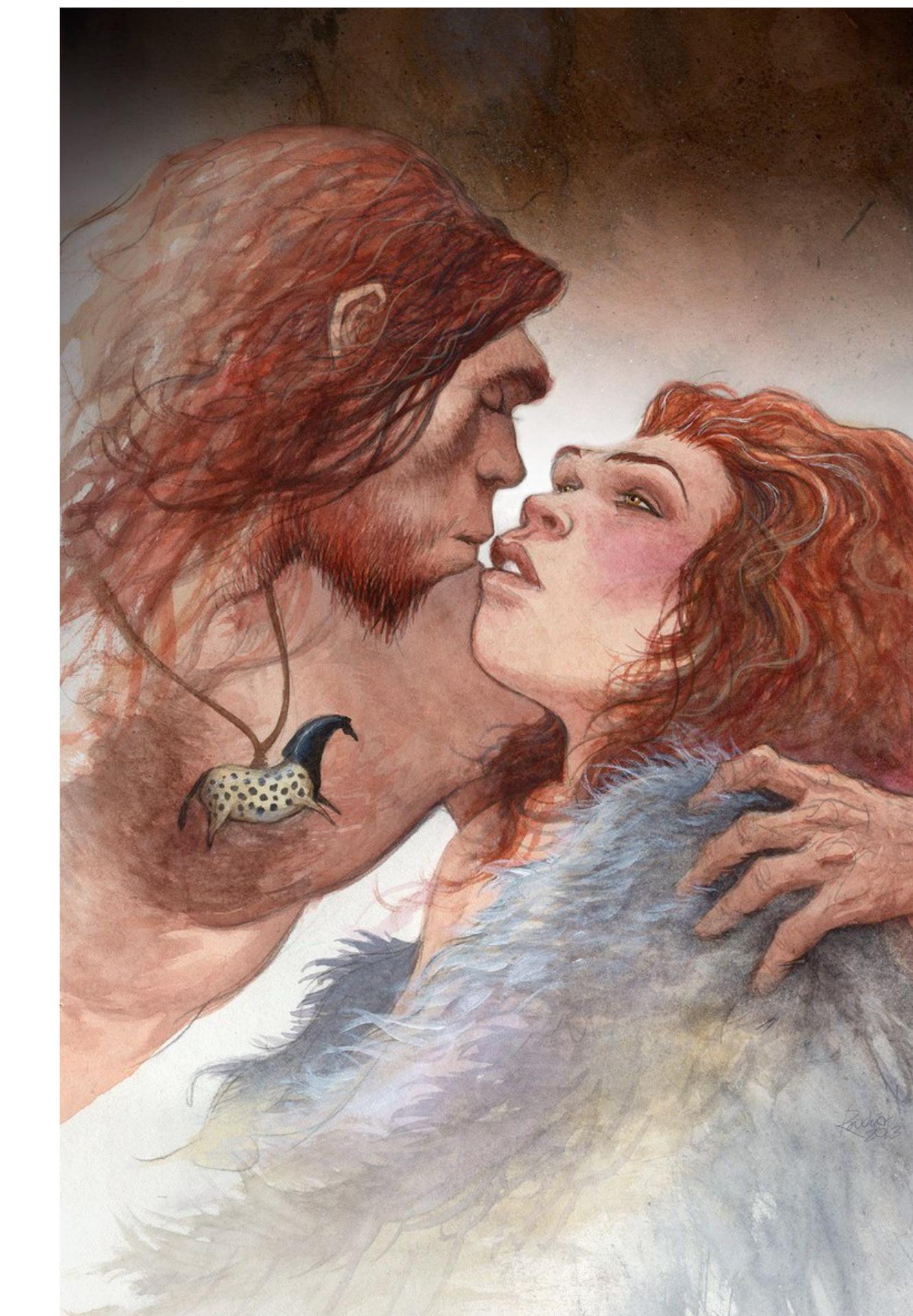
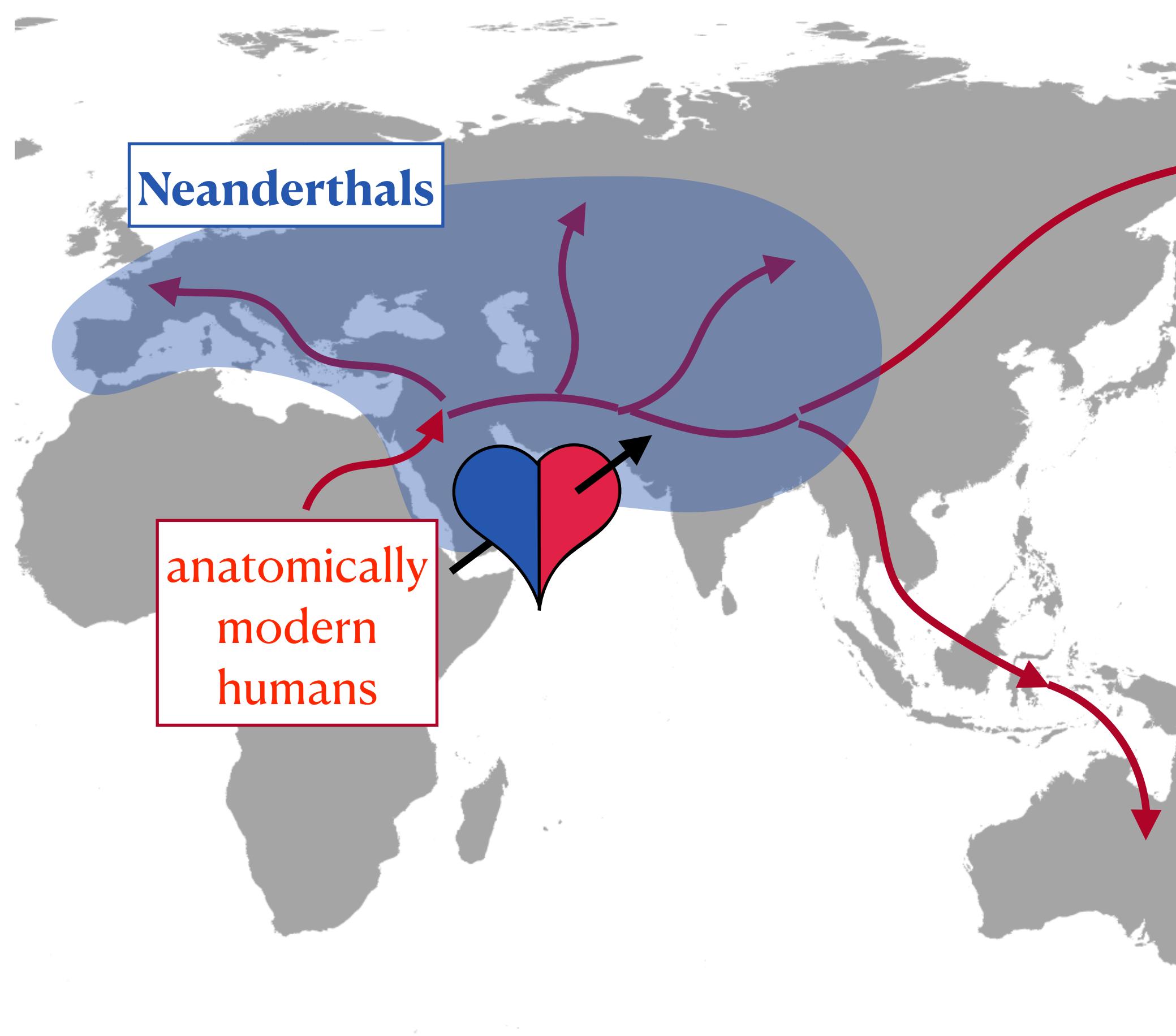
# What happened in the few thousand years of overlap?



# What happened in the few thousand years of overlap?



# Introgression / gene flow / admixture?



<https://twitter.com/ijhublin/status/739866080764628993>

# Morphological evidence?



**~40 thousand years old  
remains of a modern human  
Peștera cu Oase, Romania**

# Neanderthal DNA? (ancient DNA, aDNA)



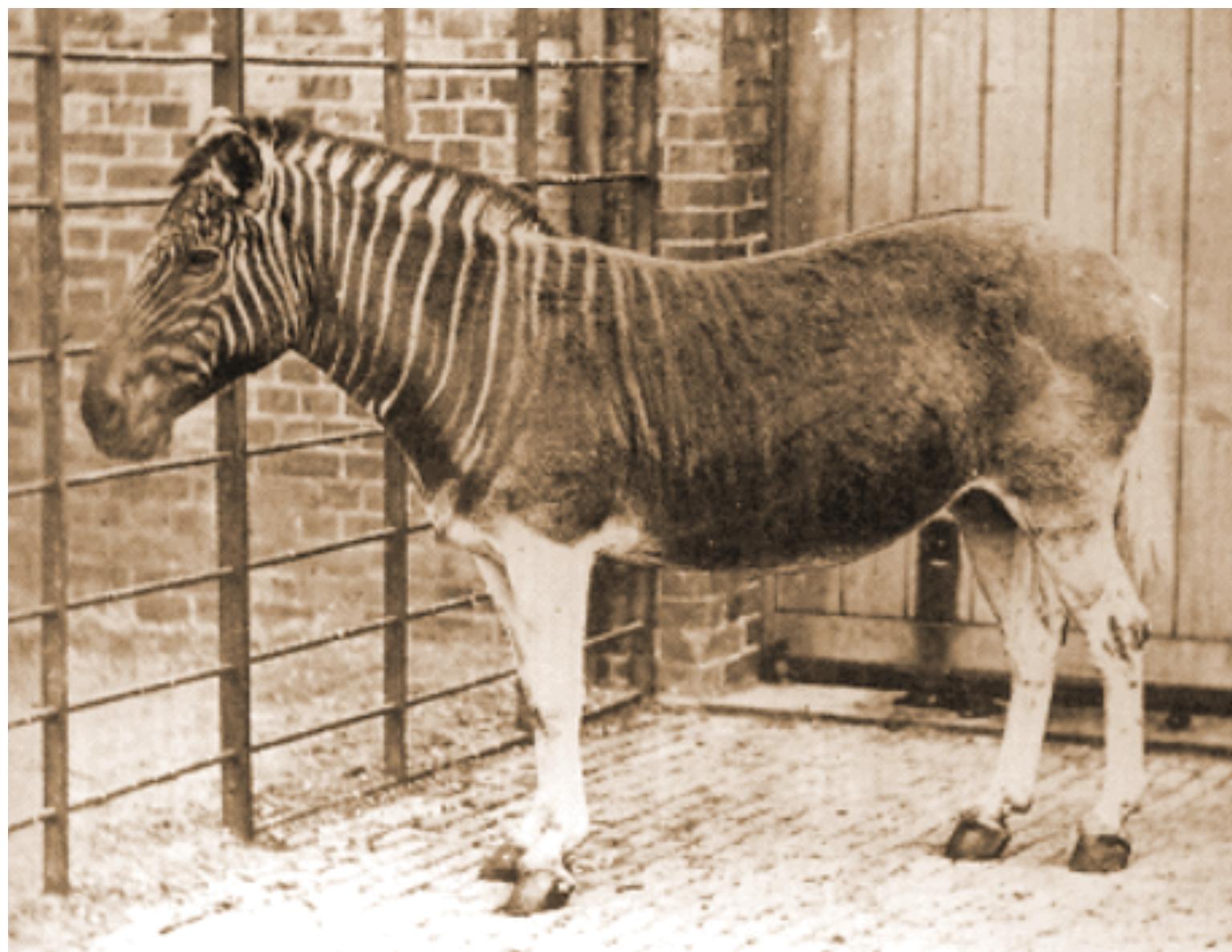
# 1984

## DNA sequences from the quagga, an extinct member of the horse family

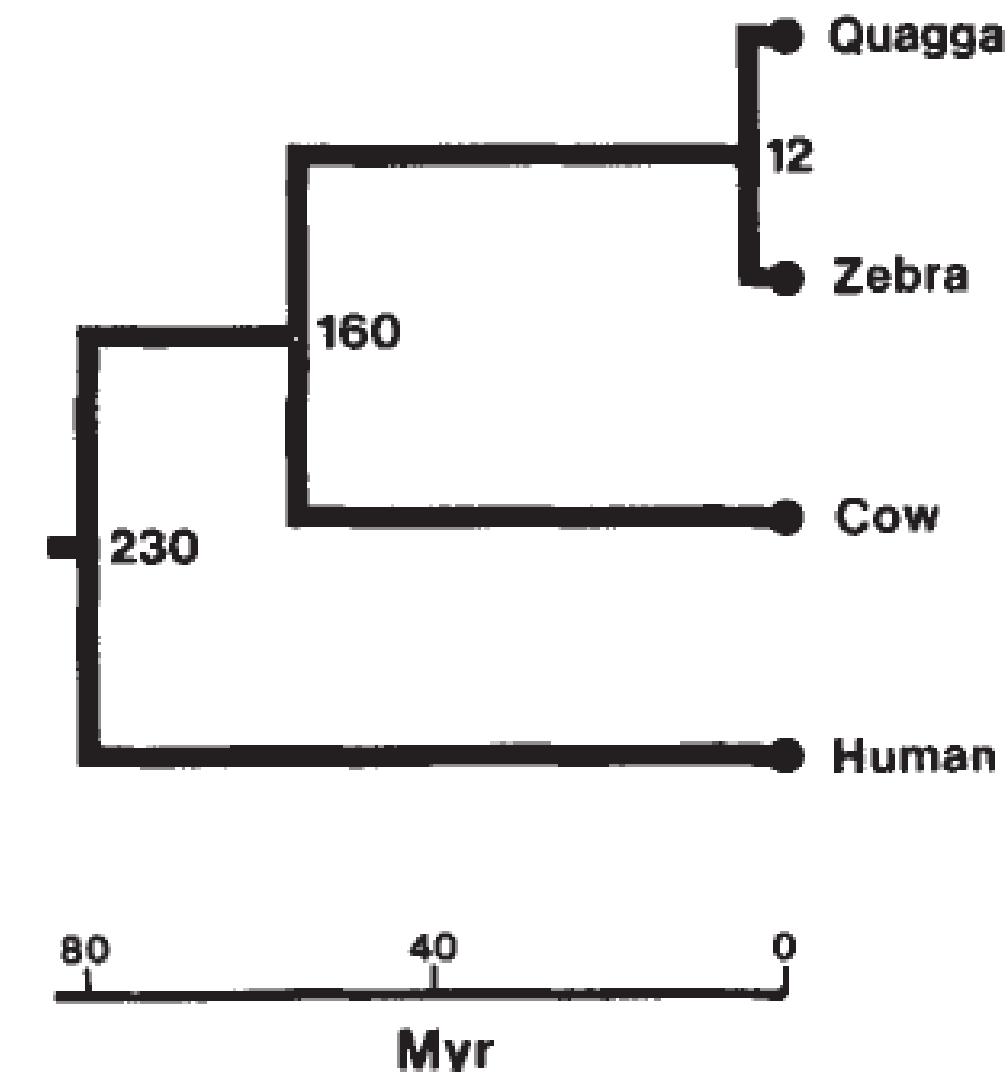
Russell Higuchi\*, Barbara Bowman\*, Mary Freiberger\*,  
Oliver A. Ryder† & Allan C. Wilson\*

\* Department of Biochemistry, University of California, Berkeley,  
California 94720, USA

† Research Department, San Diego Zoo, San Diego,  
California 92103, USA



- 150 years old tissue from a museum specimen
- 229 bp mitochondrial DNA



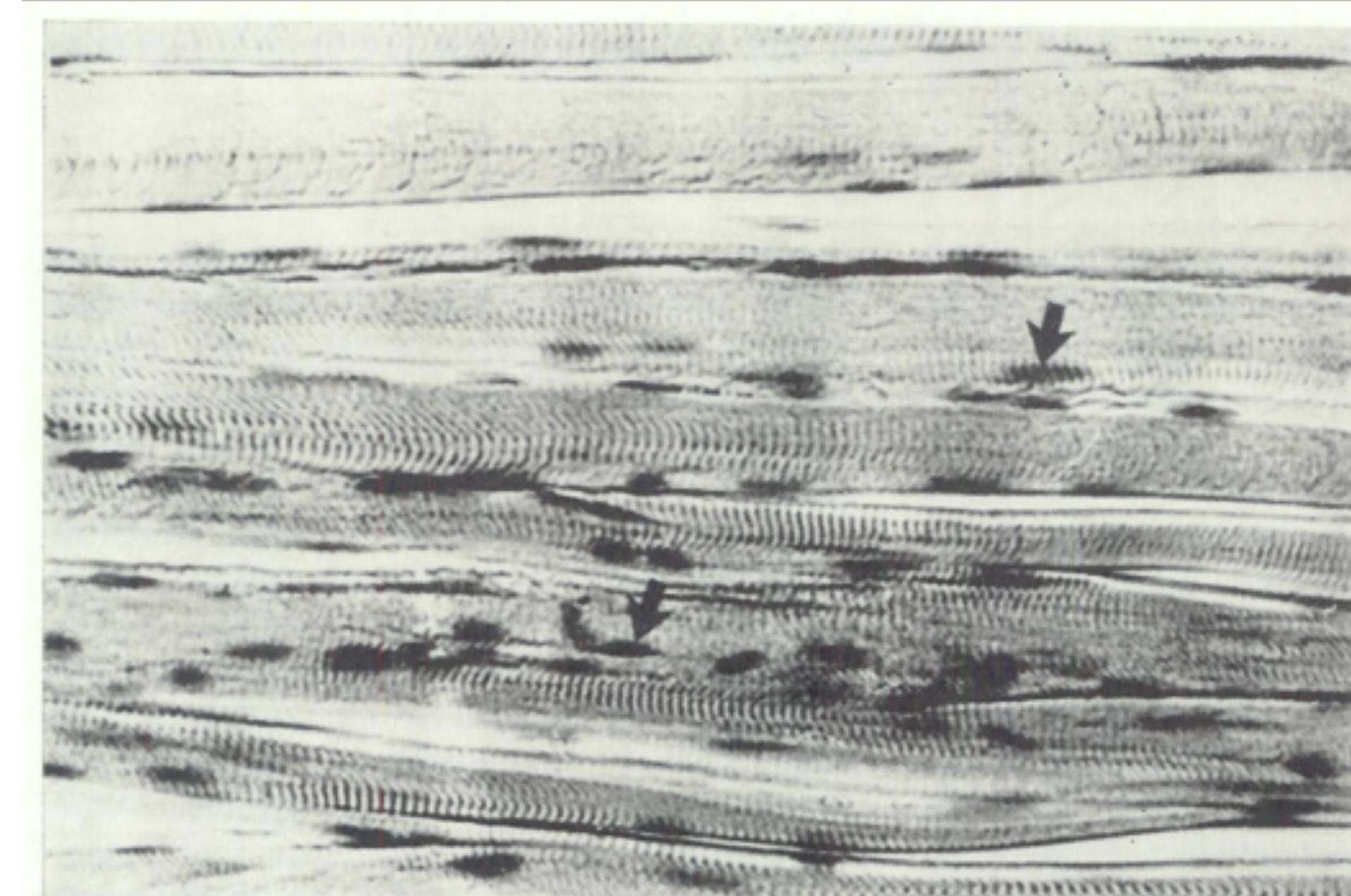
# 1985

## Molecular cloning of Ancient Egyptian mummy DNA

Svante Pääbo

Department of Cell Research, The Wallenberg Laboratory,  
University of Uppsala, Box 562, S-75122 Uppsala, Sweden and  
Institute of Egyptology, Gustavianum, University of Uppsala,  
S-75120 Uppsala, Sweden

- ~2400 years old mummy
- 3400 bp nuclear DNA



# PCR "revolution" in the aDNA field

1994

## DNA Sequence from Cretaceous Period Bone Fragments

Scott R. Woodward,\* Nathan J. Weyand, Mark Bunnell

DNA was extracted from 80-million-year-old bone fragments found in strata of the Upper Cretaceous Blackhawk Formation in the roof of an underground coal mine in eastern Utah. This DNA was used as the template in a polymerase chain reaction that amplified and sequenced a portion of the gene encoding mitochondrial cytochrome b. These sequences differ from all other cytochrome b sequences investigated, including those in the GenBank and European Molecular Biology Laboratory databases. DNA isolated from these bone fragments and the resulting gene sequences demonstrate that small fragments of DNA may survive in bone for millions of years.

SCIENCE • VOL. 266 • 18 NOVEMBER 1994



# PCR "revolution" in the aDNA field

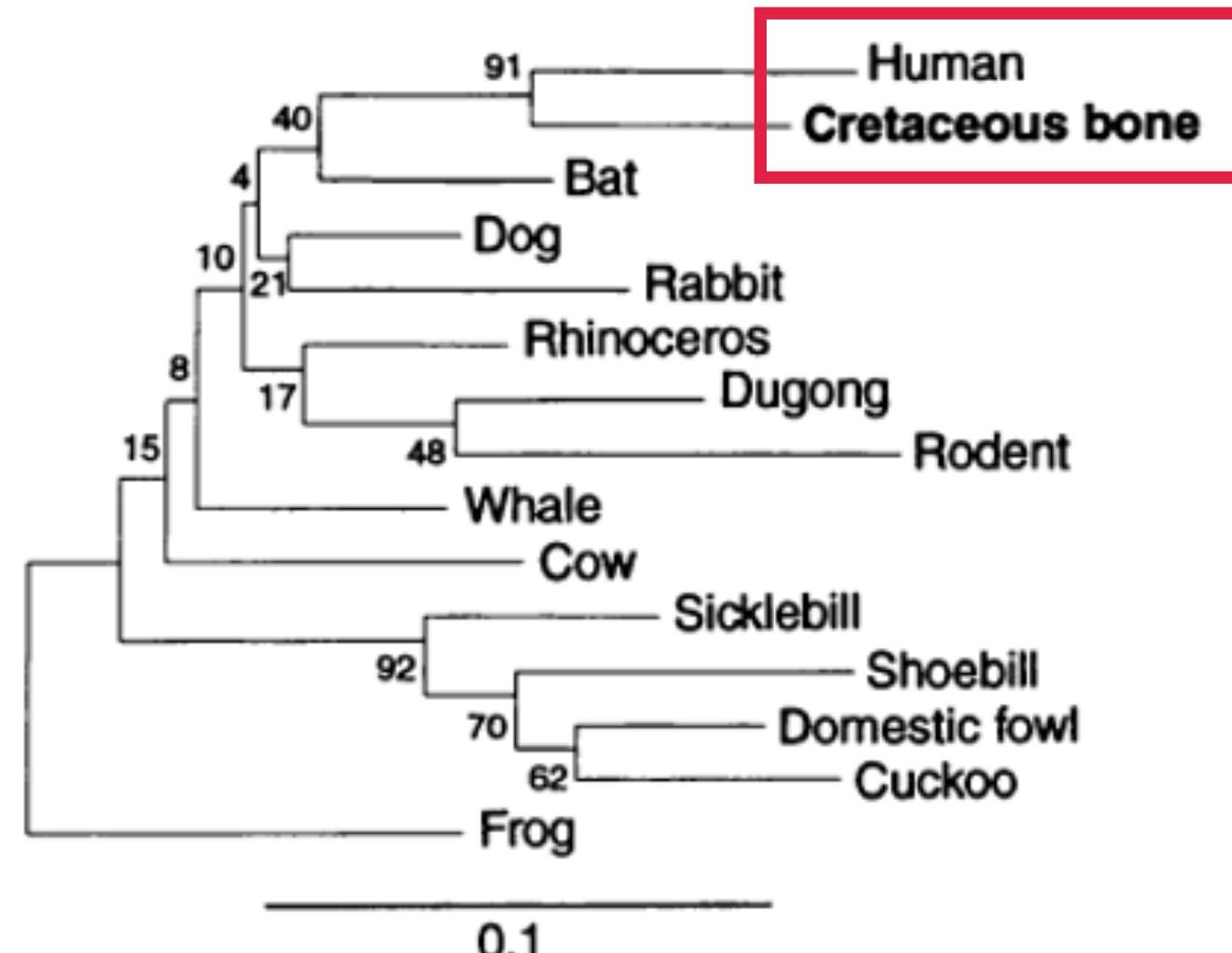
1994

## DNA Sequence from Cretaceous Period Bone Fragments

Scott R. Woodward,\* Nathan J. Weyand, Mark Bunnell

DNA was extracted from 80-million-year-old bone fragments found in strata of the Upper Cretaceous Blackhawk Formation in the roof of an underground coal mine in eastern Utah. This DNA was used as the template in a polymerase chain reaction that amplified and sequenced a portion of the gene encoding mitochondrial cytochrome b. These sequences differ from all other cytochrome b sequences investigated, including those in the GenBank and European Molecular Biology Laboratory databases. DNA isolated from these bone fragments and the resulting gene sequences demonstrate that small fragments of DNA may survive in bone for millions of years.

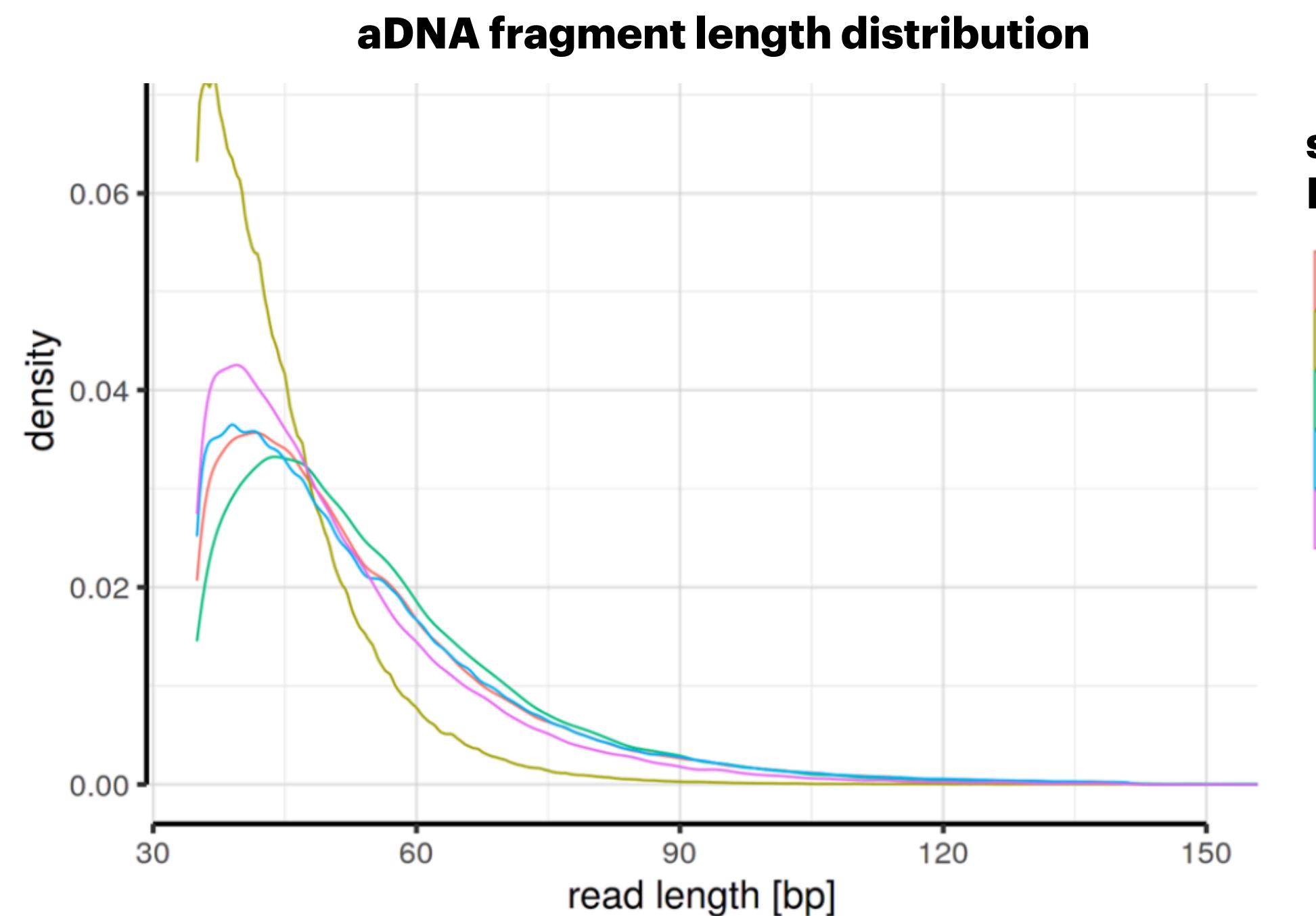
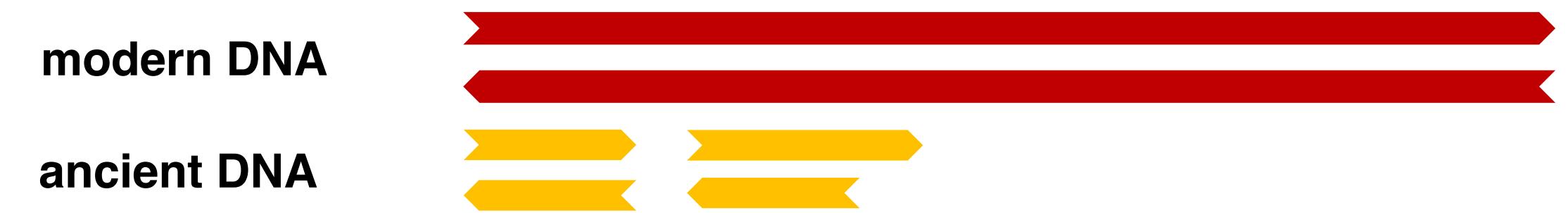
SCIENCE • VOL. 266 • 18 NOVEMBER 1994



# Molecular characteristics of aDNA

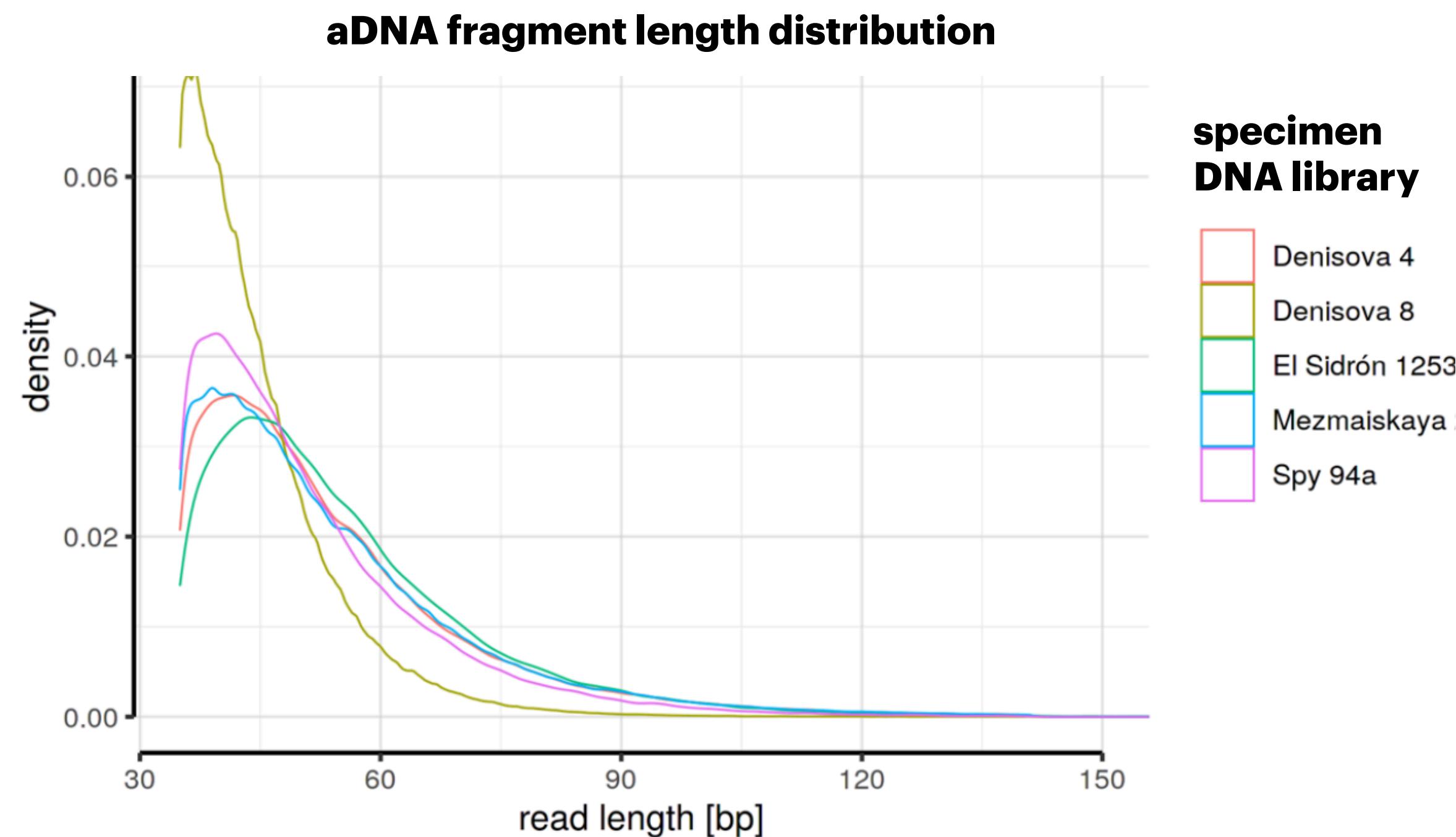
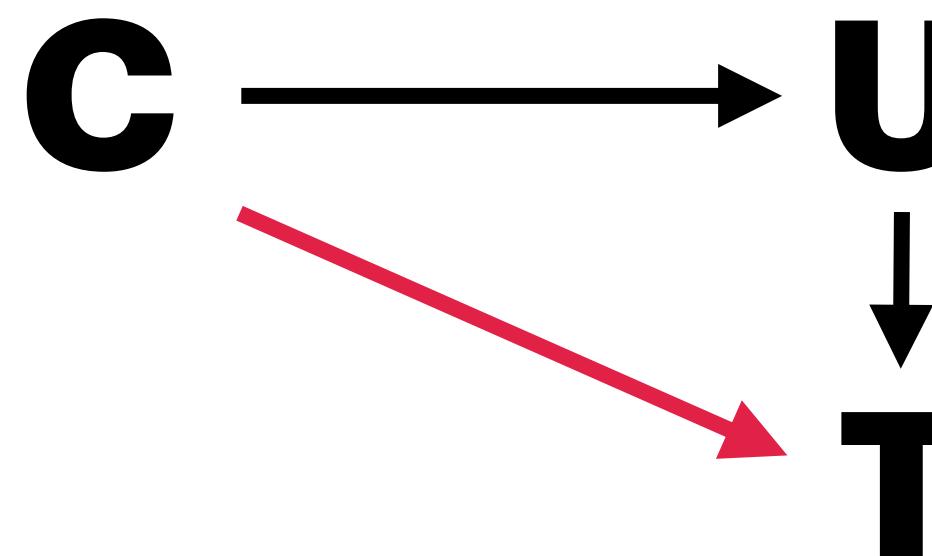
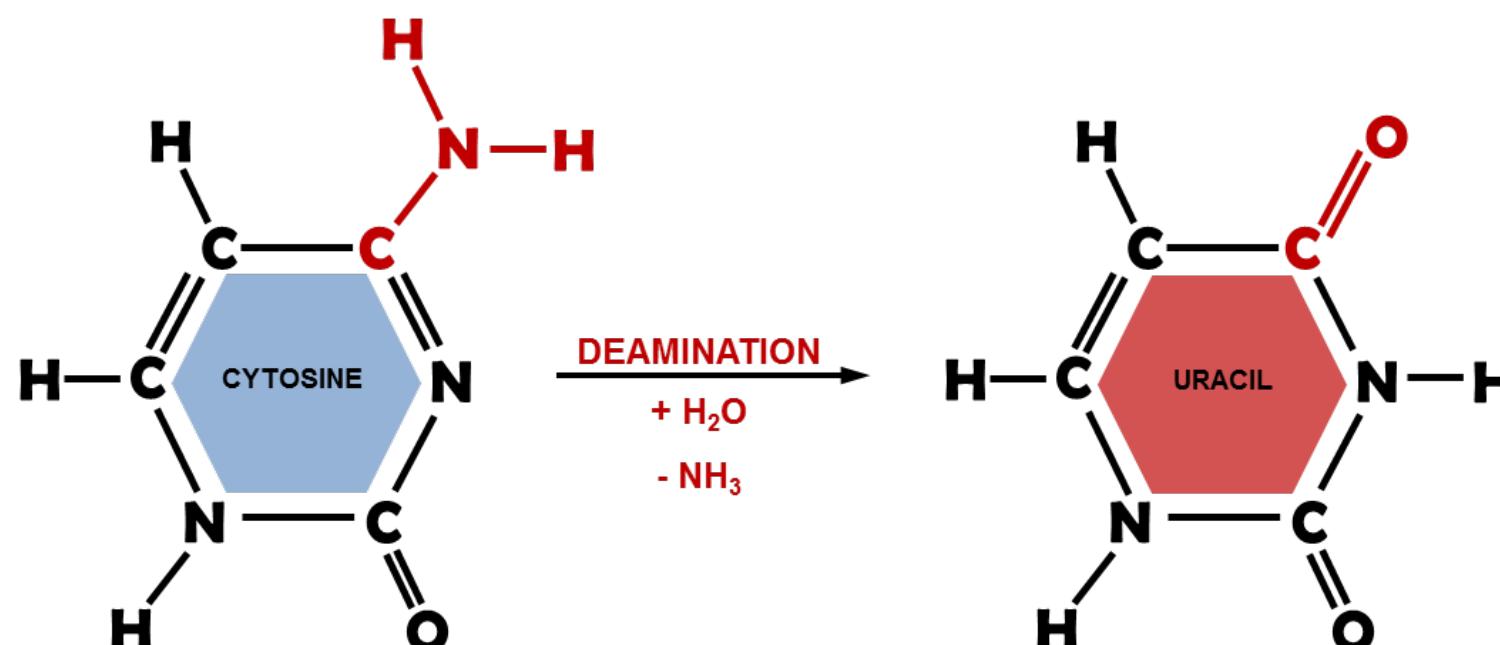
# Molecular characteristics of aDNA

- highly fragmented



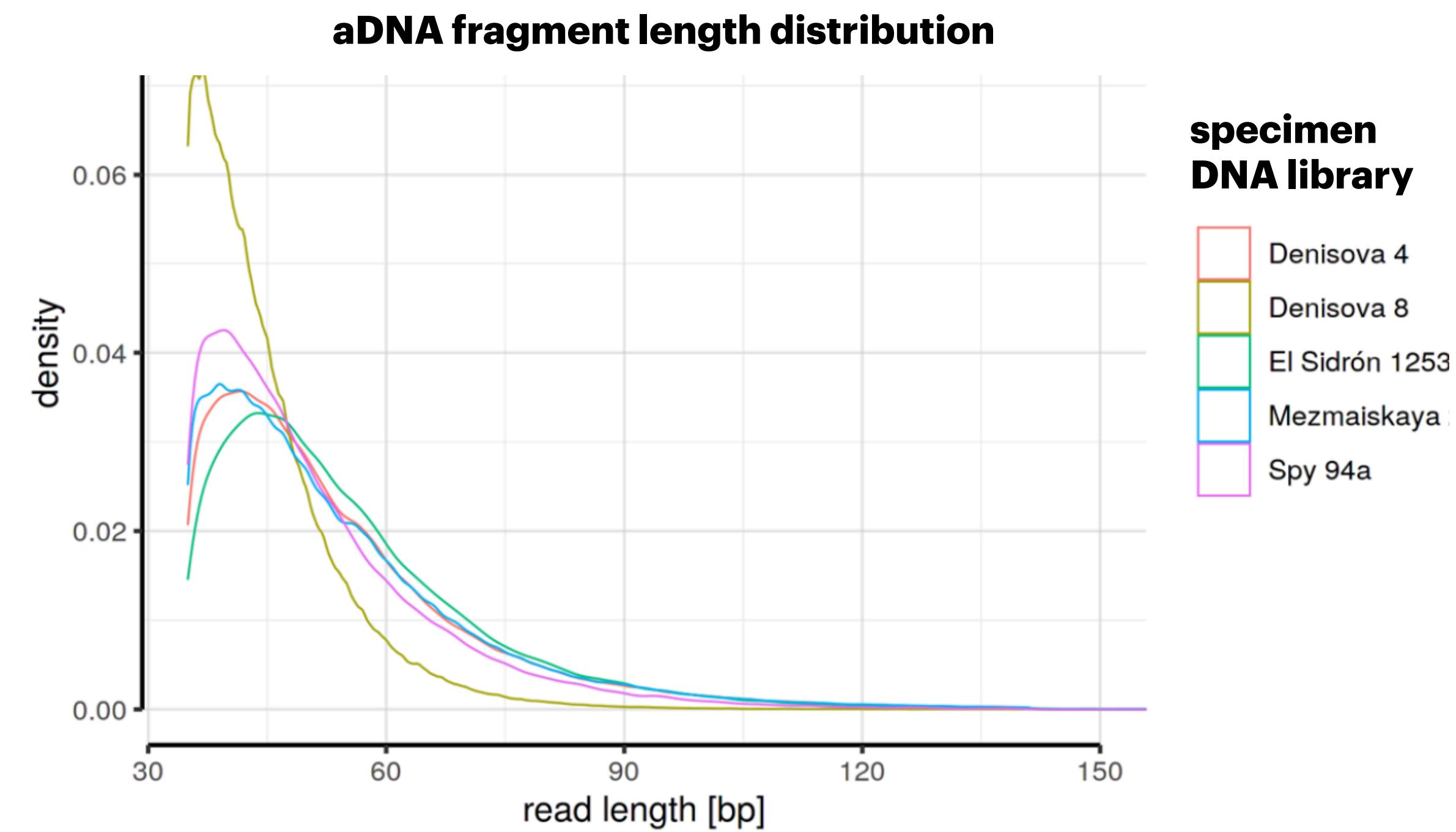
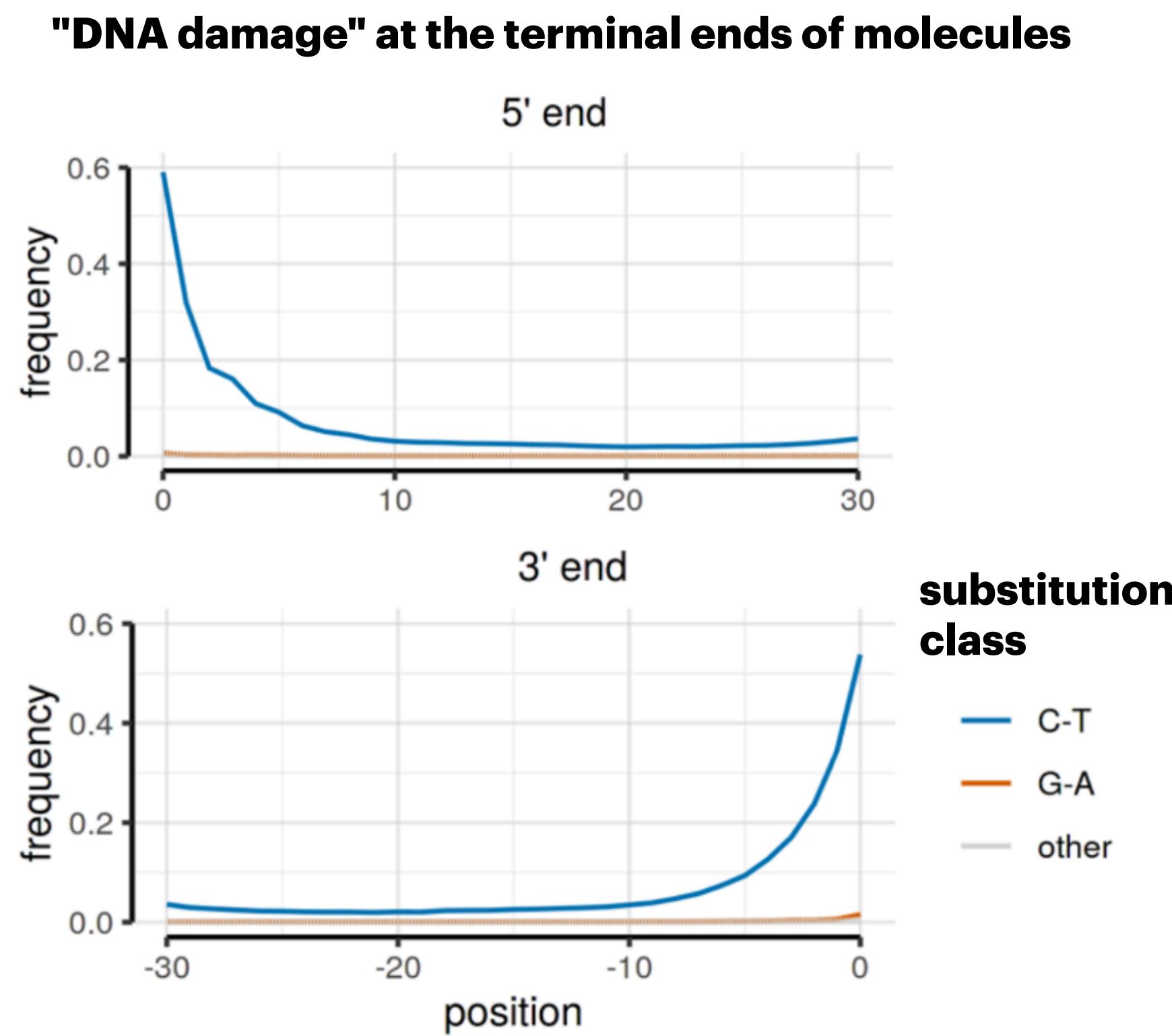
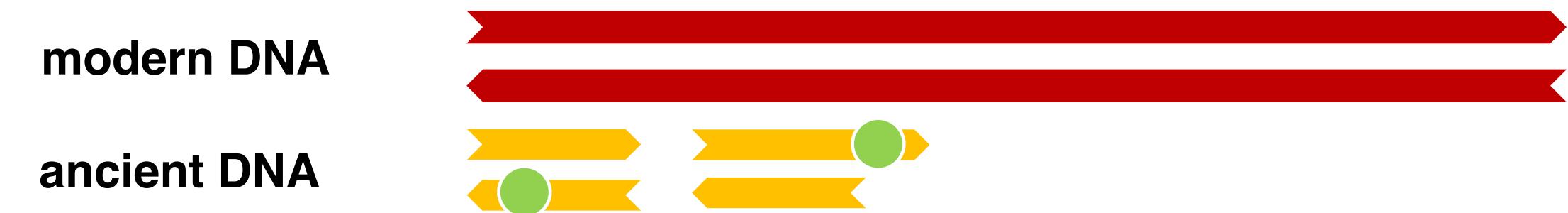
# Molecular characteristics of aDNA

- highly fragmented
- post-mortem chemical modifications



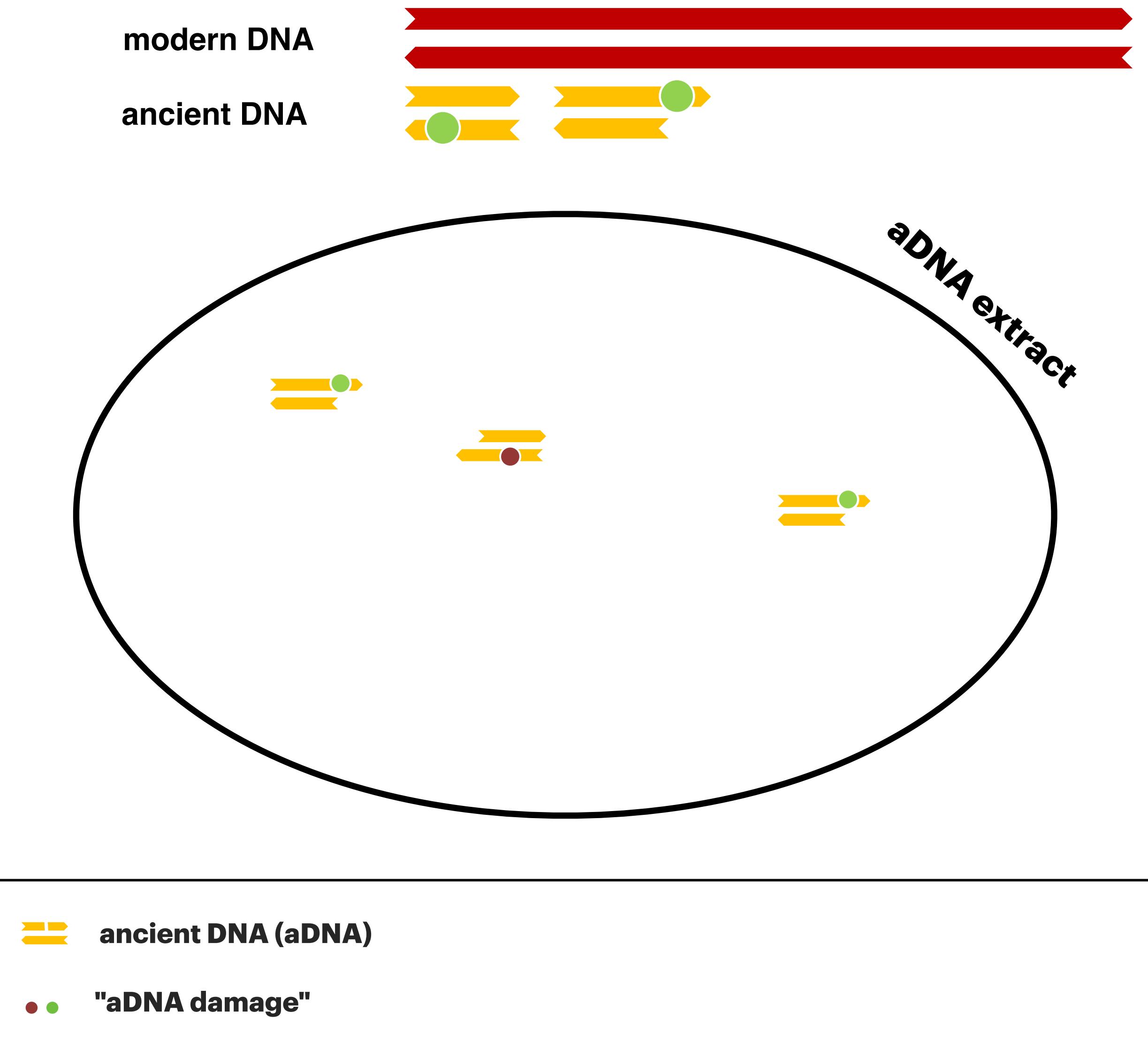
# Molecular characteristics of aDNA

- highly fragmented
- post-mortem chemical modifications



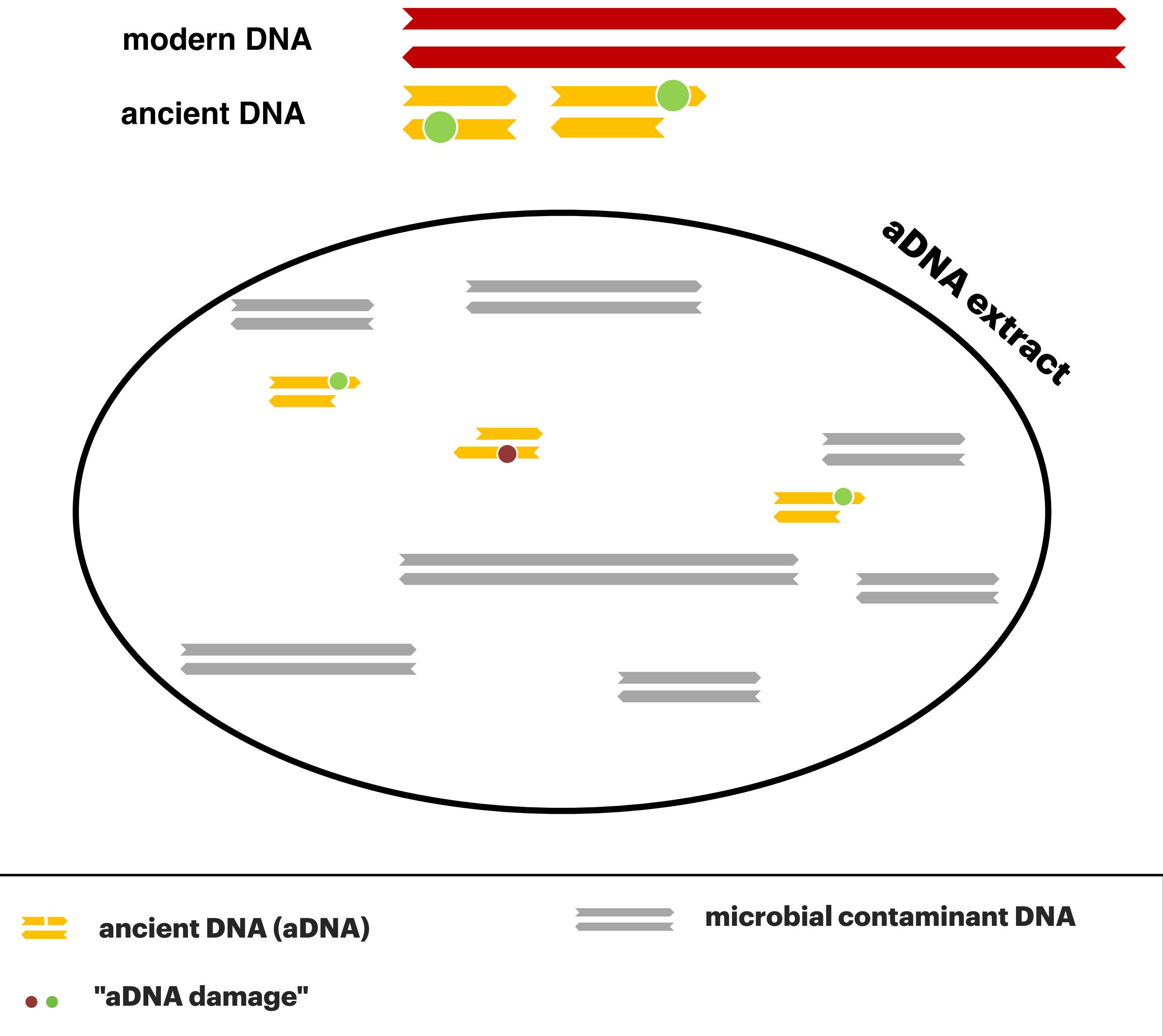
# Molecular characteristics of aDNA

- highly fragmented
- post-mortem chemical modifications
- contamination:
  - microbial DNA
  - human DNA (excavation, museum, lab)



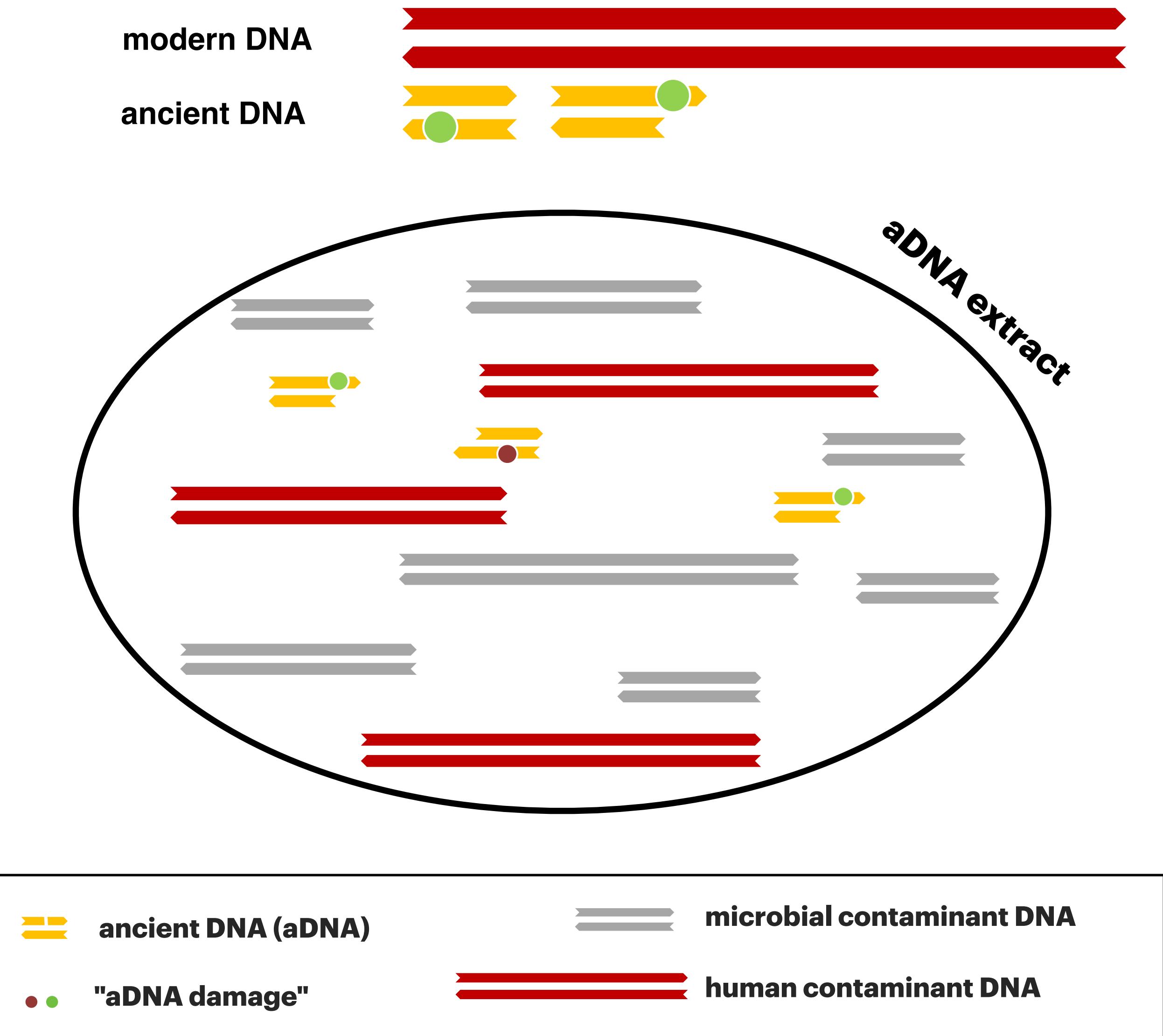
# Molecular characteristics of aDNA

- highly fragmented
- post-mortem chemical modifications
- contamination:
  - microbial DNA
  - human DNA (excavation, museum, lab)



# Molecular characteristics of aDNA

- highly fragmented
- post-mortem chemical modifications
- contamination:
  - microbial DNA
  - human DNA (excavation, museum, lab)

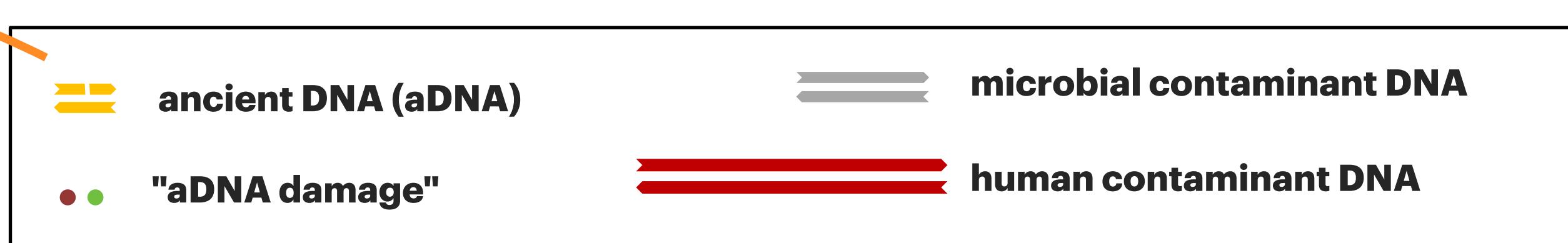
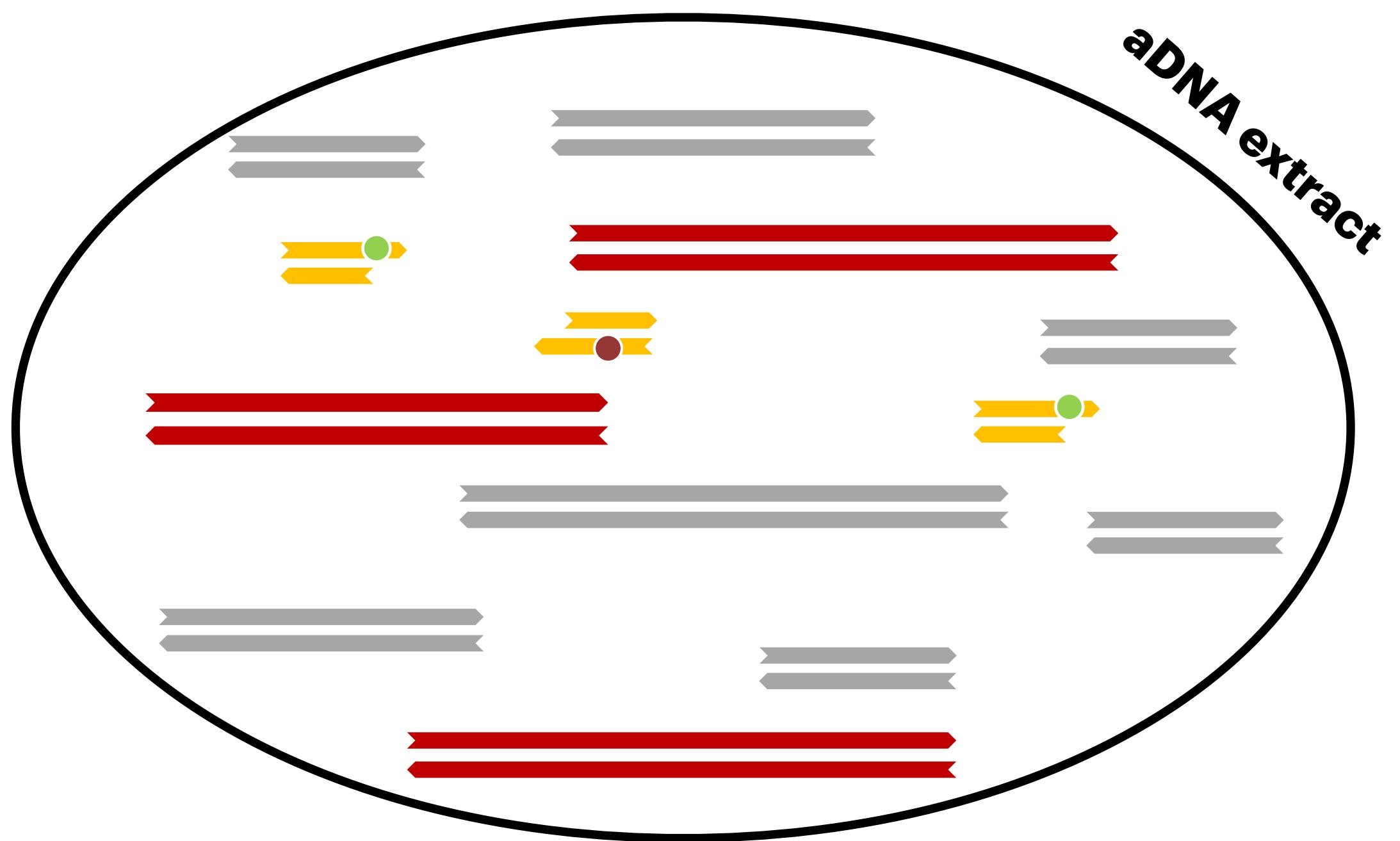


# Molecular characteristics of aDNA

- highly fragmented
- post-mortem chemical modifications
- contamination:
  - microbial DNA
  - human DNA (excavation, museum, lab)

aDNA extract

often < 1% endogenous DNA



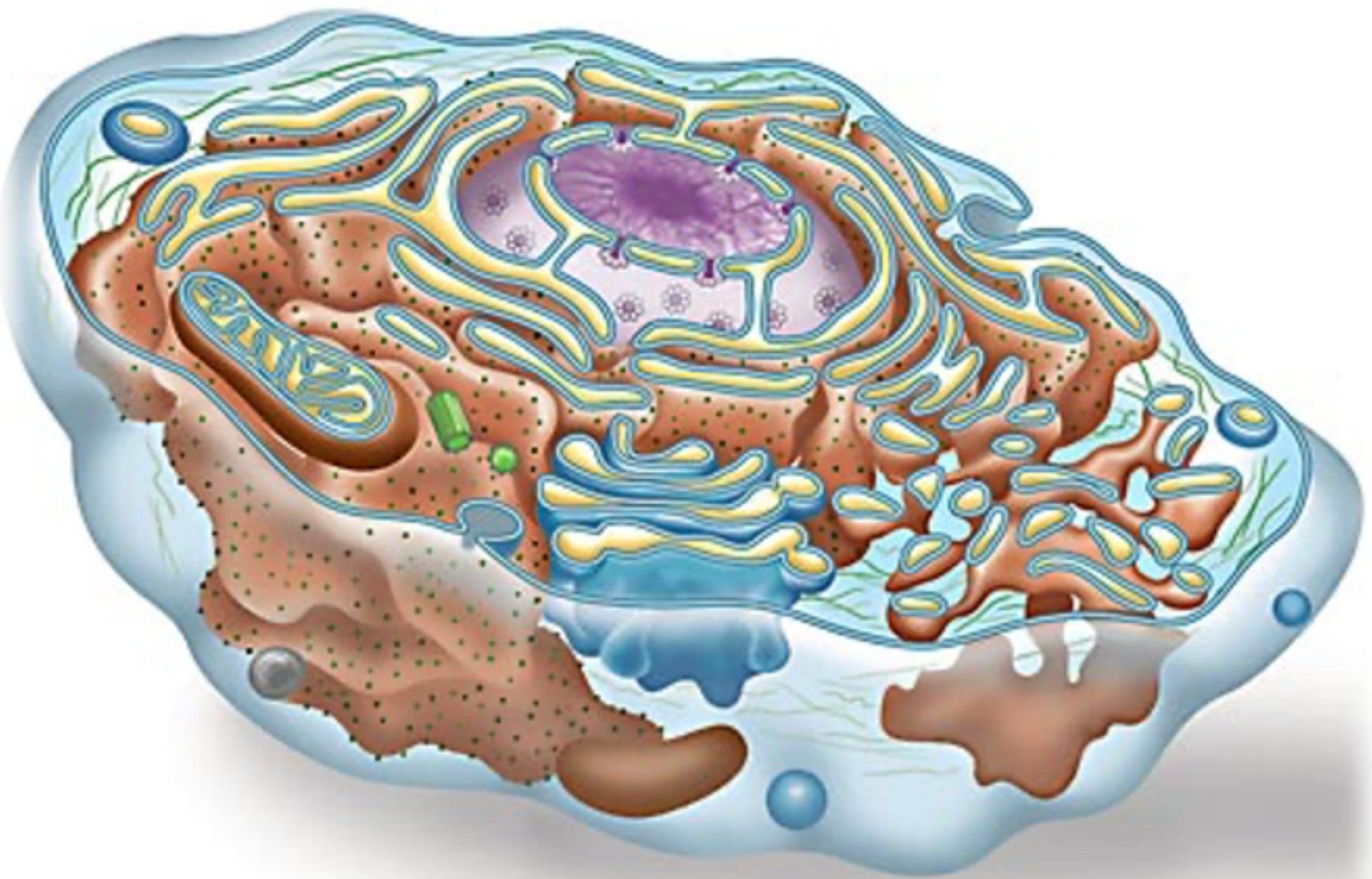
# Neanderthal DNA?



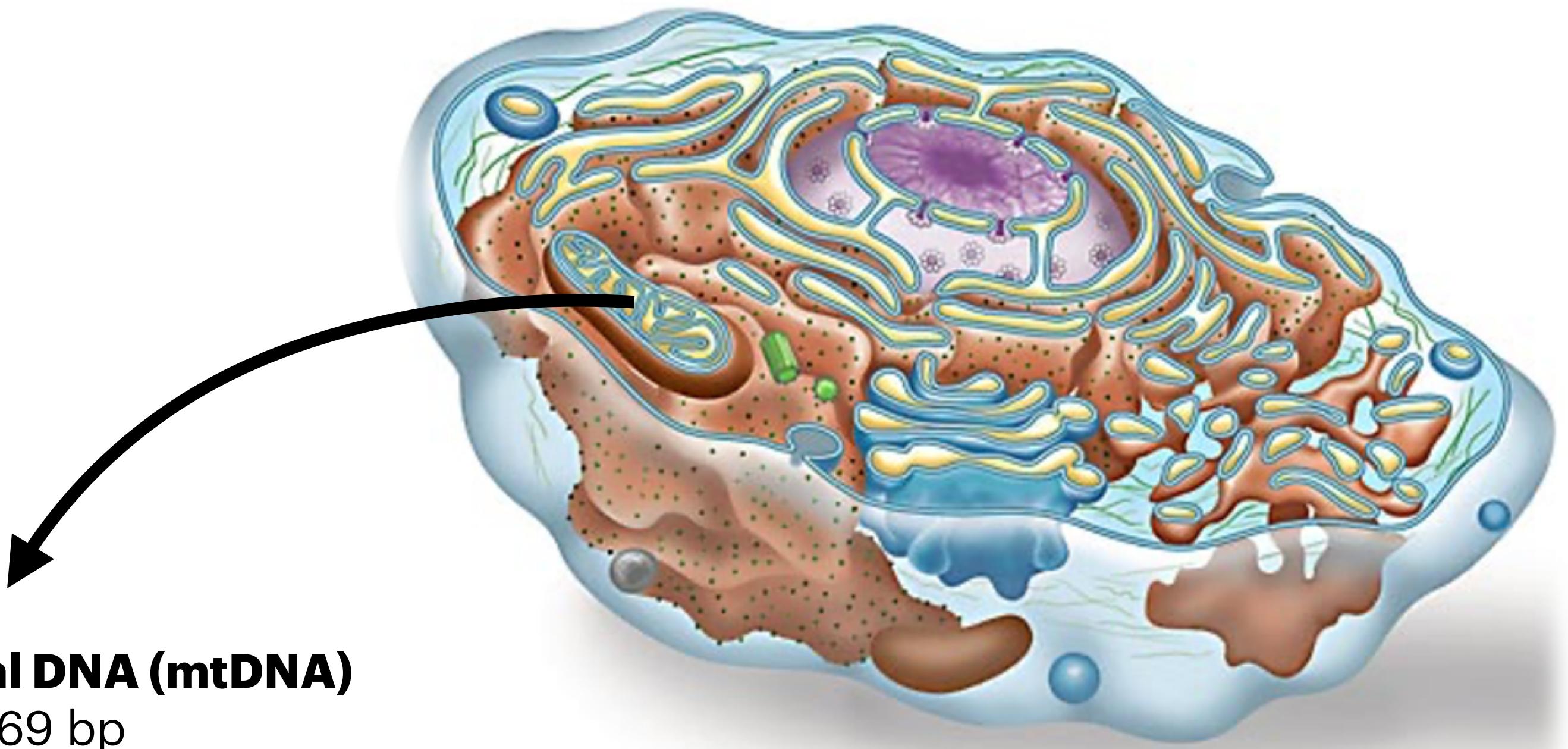
# Svante Pääbo



# Two sources of DNA



# Two sources of DNA

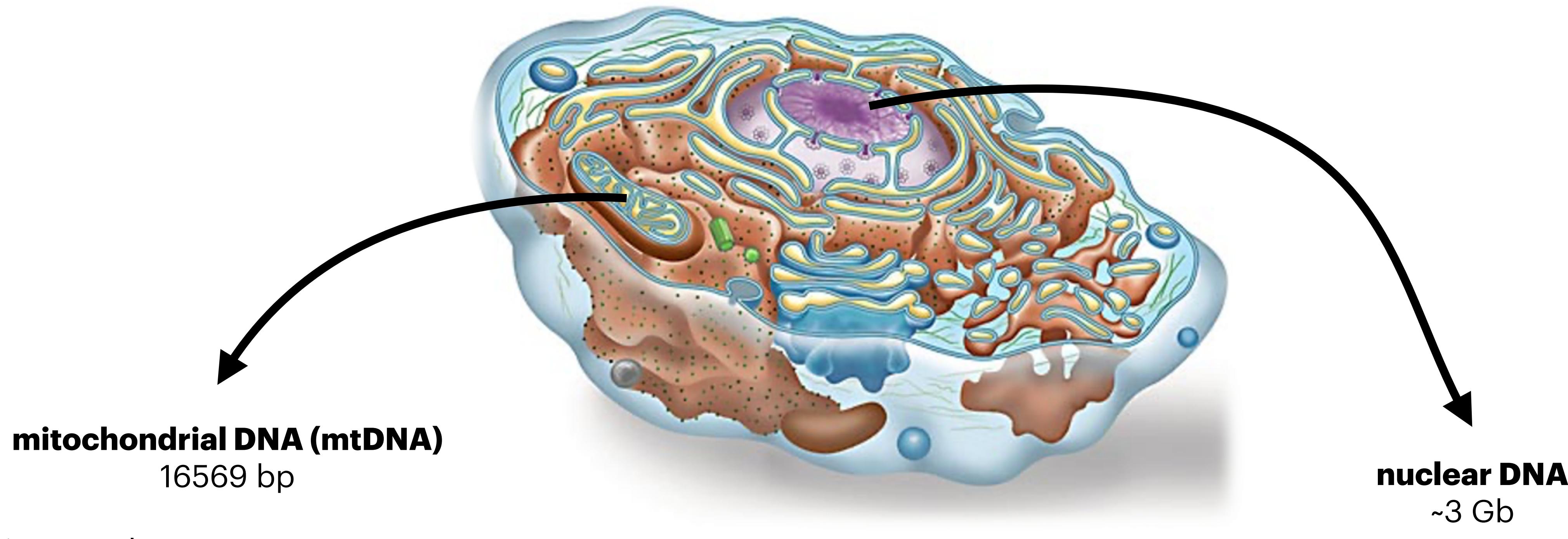


## mitochondrial DNA (mtDNA)

16569 bp

- just one locus
- transmitted along maternal line
- fast mutation rate
- **thousands of copies in a cell**

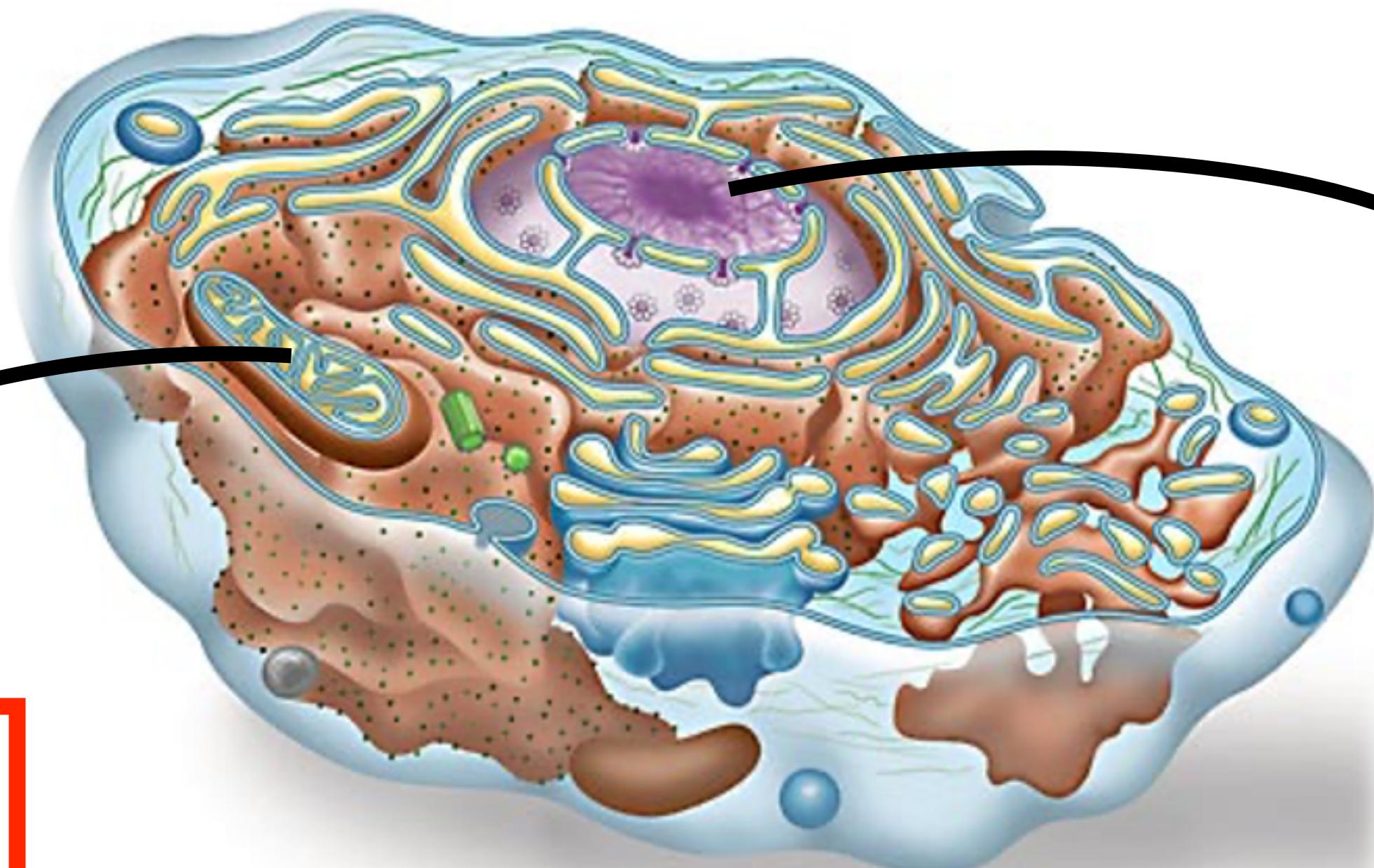
# Two sources of DNA



- just one locus
- transmitted along maternal line
- fast mutation rate
- **thousands of copies in a cell**

- thousands of loci due to recombination
- complex mosaic of genetic history
- **two copies in a human cell**

# Two sources of DNA



## mitochondrial DNA (mtDNA)

16569 bp

- just one locus
- transmitted along maternal line
- fast mutation rate
- **thousands of copies in a cell**

## nuclear DNA

~3 Gb

- thousands of loci due to recombination
- complex mosaic of genetic history
- **two copies in a human cell**

# 1997: Neanderthal mitochondrial DNA!

Cell, Vol. 90, 19–30, July 11, 1997, Copyright ©1997 by Cell Press

## Neandertal DNA Sequences and the Origin of Modern Humans

Matthias Krings,\* Anne Stone,† Ralf W. Schmitz,‡  
Heike Krainitzki,§ Mark Stoneking,† and Svante Pääbo\*



refseq		ACAGCAATCAGCCCTCAACTATCACACATCAACTGCAACTCCAAAGCCACCCCT-CACCCACTAGGGATACCAACAAACCTAACAGTACATAGTACATAAACCCATTACCGTACATAGCACATTACAGTC*
A12.1	L16,269	T...G...T.....A.....A.G...T.A.....T.....T.G.....G.....C.....T.....
A12.2		T...G...T.....A.....A.G...T.A.....T.....T.G.....G.....C.....T.....
A12.3		T...G...T.....A.....A.G...T.A.....T.....T.G.....G.....CG.....T.....
A12.4		T...G...T.....A.....A.G...T.A.....T.....T.G.....G.....C.....T.....
A12.5		T...G...T.....A.....A.G...T.A.....T.....T.G.....G.....C.....T.....
A12.6		T...G...T.....A.....A.G...T.A.....T.....T.G.....G.....C.....T.....
A12.7		T...G...T.....A.....A.G...T.A.....T.....T.G.....G.....C.....T.....
A12.8		T...G...T.....A.....A.G...T.A.....T.....T.G.....G.....C.....T.....
A12.9		T...G...T.....A.....A.G...T.A.....T.....T.G.....G.....C.....T.....
A12.10		T...G...T.....A.....A.G...T.A.....T.....T.G.....G.....C.....T.....
A12.11		T...G...T.....A.....A.G...T.A.....T.....T.G.....G.....C.....T.....
A12.12		T...G...T.....A.....A.G...T.A.....T.....T.G.....G.....C.....T.....
A12.13		T...G...T.....A.....A.G...T.A.....T.....T.G.....G.....C.....T.....
B13.1	CACACATCAACTGCAACTCCAA	A.G...T.A.....T.....G.....G.....C.....T.....
B13.2	L16,254	A.G...T.A.....T.....G.....G.....C.....T.....
B13.3		A.G...T.A.....T.....G.....G.....C.....T.....
B13.4		A.G...T.A.....T.....G.....G.....C.....T.....
B13.5		A.G...T.A.....T.....G.....G.....C.....T.....
B13.6		A.G...T.A.....T.....G.....G.....C.....T.....
B13.7		A.G...T.A.....T.....G.....G.....C.....T.....
B13.8		A.G...T.A.....T.....G.....G.....C.....T.....
B14.1	ACTACAACTCCAAAGRCGCCCTTA	.....G.....C.....T.....
B14.2	ML16,263/264	.....T.....G.....C.....T.....
B14.3		.....T.....G.....C.....T.....
B14.4		.....T.....G.....C.....G.....T.....
B14.5		.....T.....G.....C.....T.....
B14.6		.....T.....G.....C.....T.....
B14.7		.....T.....G.....C.....T.....
B14.8		.....T.....G.....C.....T.....
B14.9		.....T.....G.....C.....T.....
B14.10		.....T.....G.....C.....T.....
B14.11		.....T.....G.....C.....T.....
B14.12		.....T.....G.....C.....T.....
B14.13		.....T.....G.....C.....T.....
A15.1		CGTACATAGCACATTACAGT
A15.2	L16,347	.....T.....T.....
A15.3		.....C.....T.....
A15.4		.....C.....T.....
A15.5		.....C.....T.....
A15.6		.....C.....T.....
A15.7		.....C.....AA.....T.....
A15.8		.....C.....AA.....T.....
A15.9		.....C.....AA.....T.....
A15.10		.....C.....AA.....T.....
Neandertal (16,210-16,400)		.....T.....G.....T.....A.....A.G...T.A.....T.....G.....C.....T.....

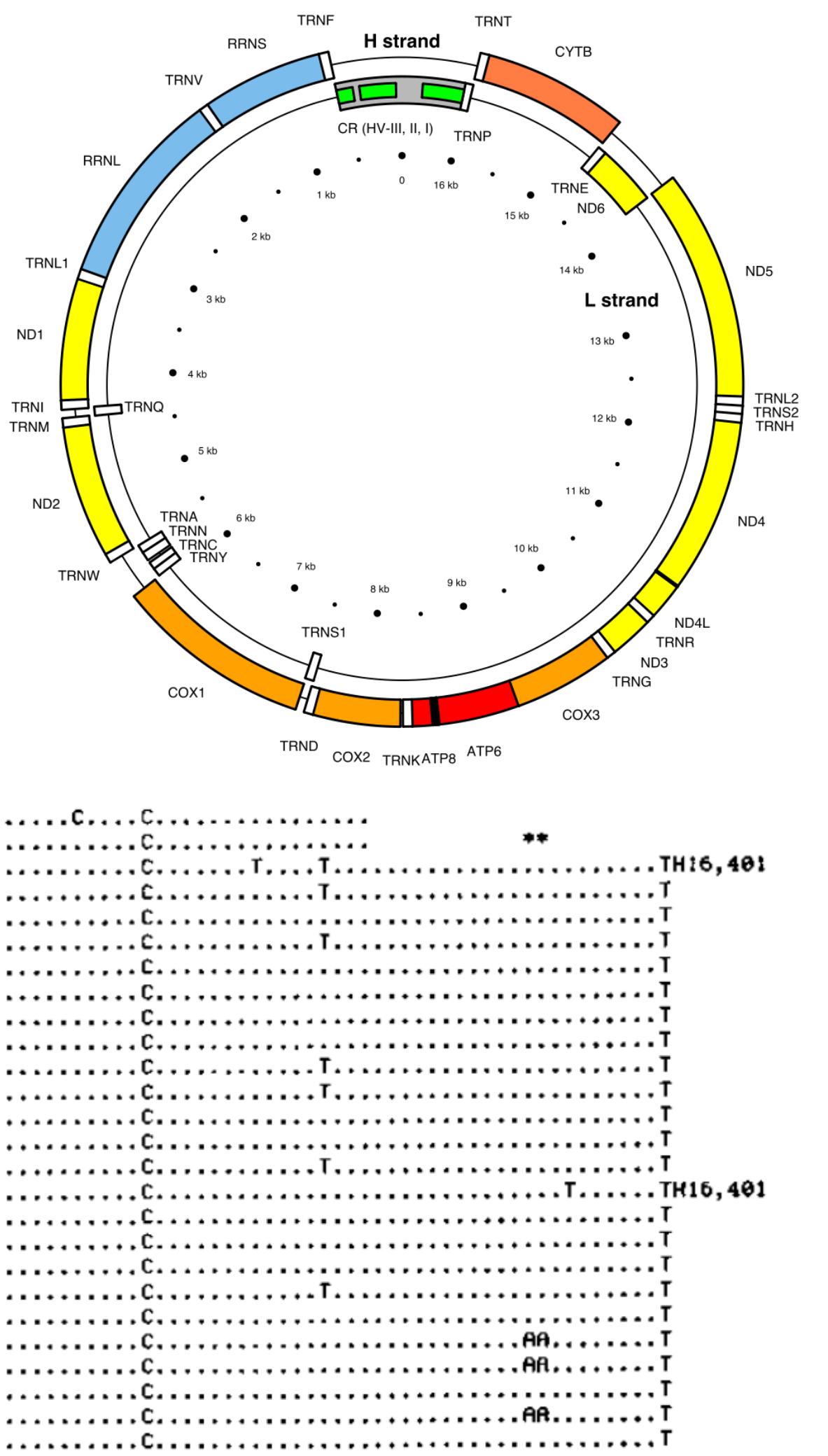
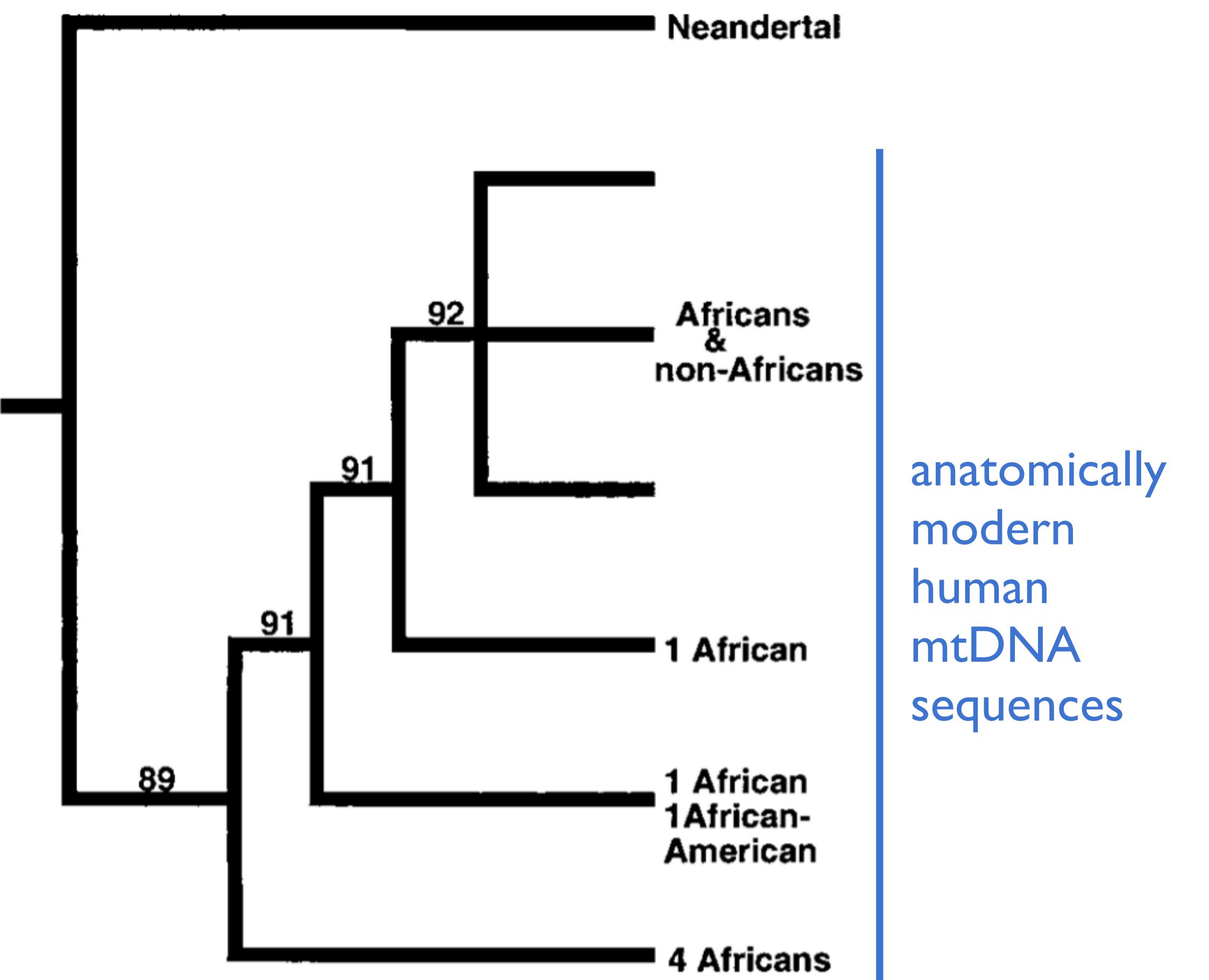
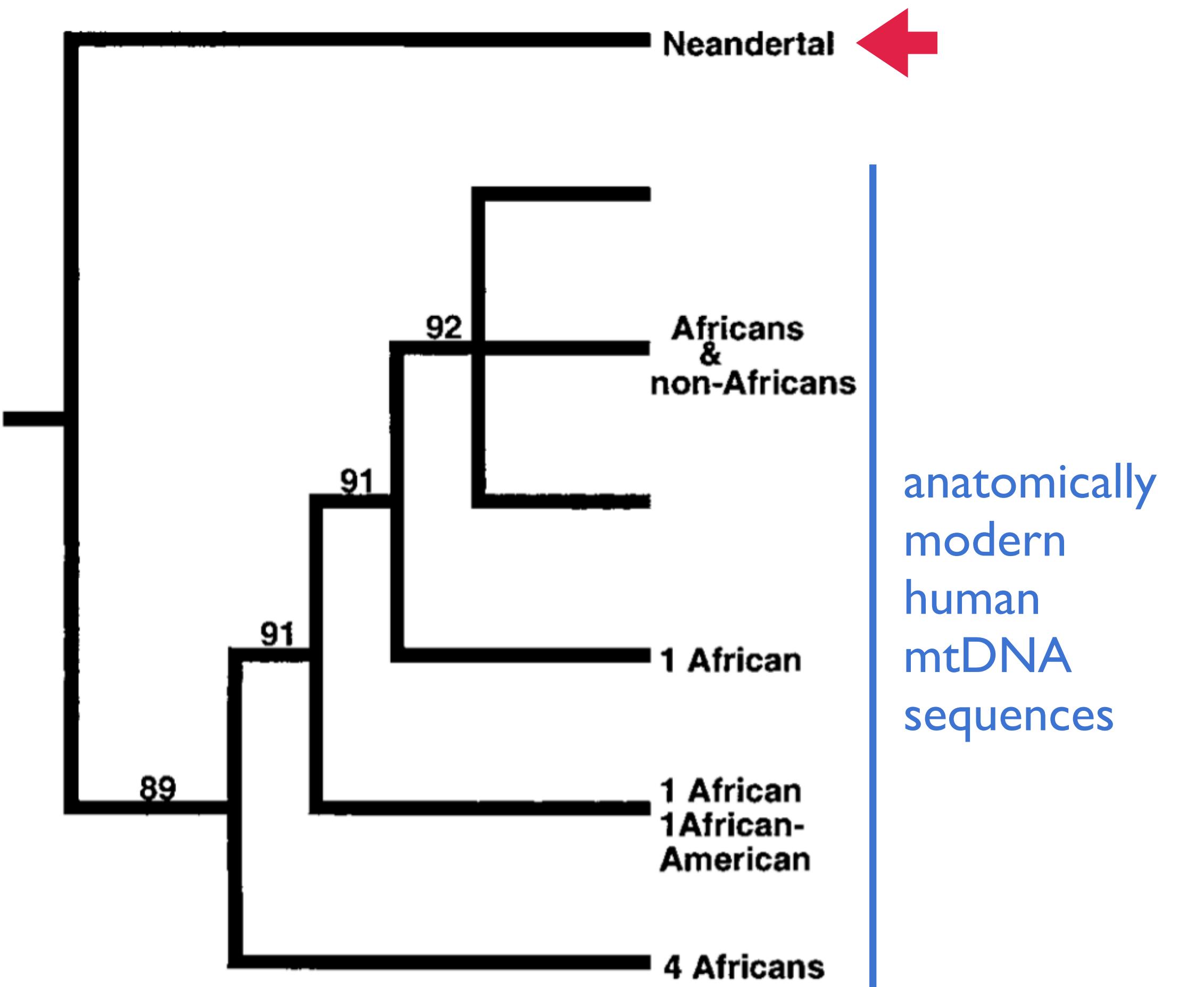
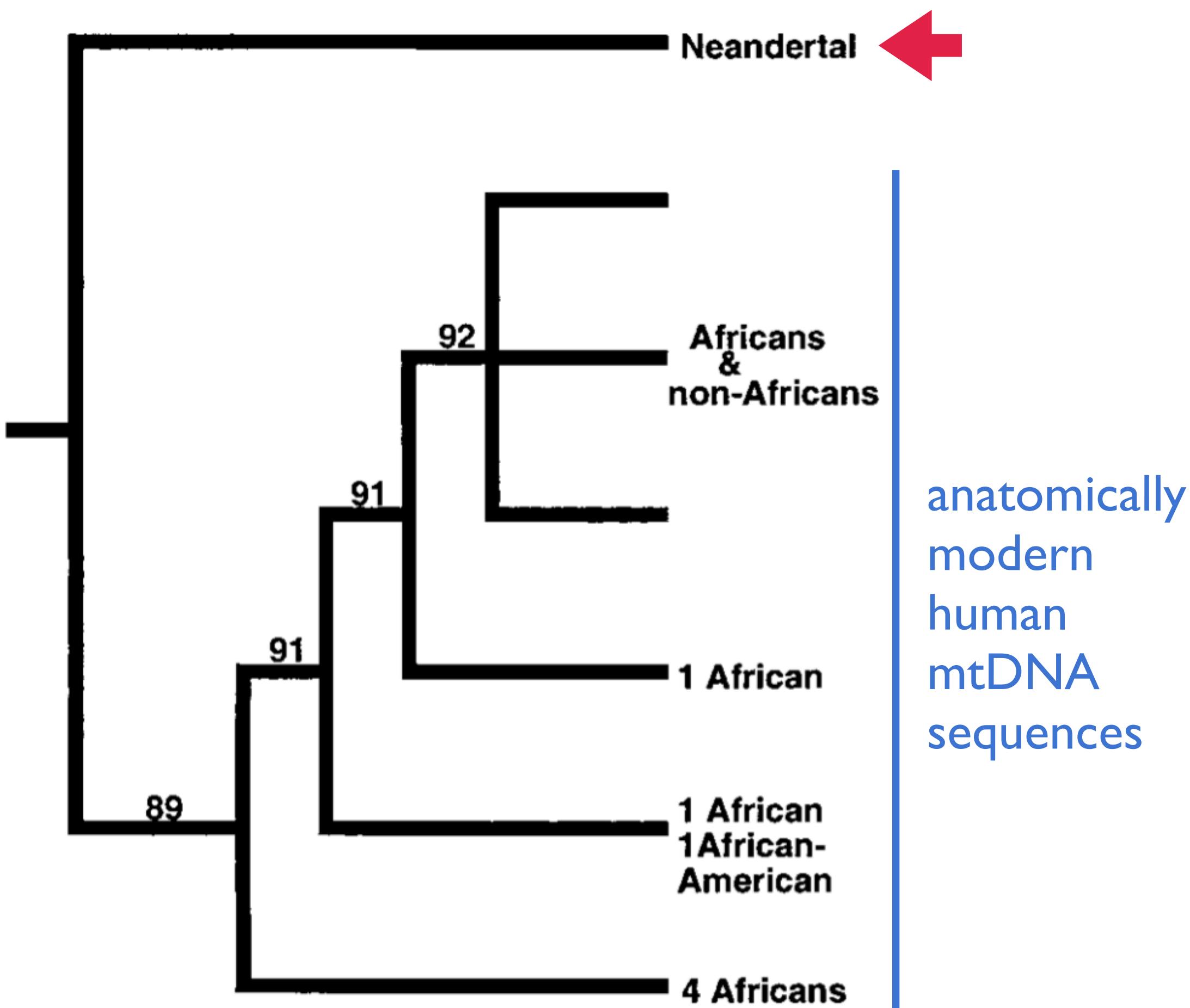


Figure 5. The DNA Sequences of Clones Used to Infer the Sequence of the Hypervariable Region I of the Neandertal Individual



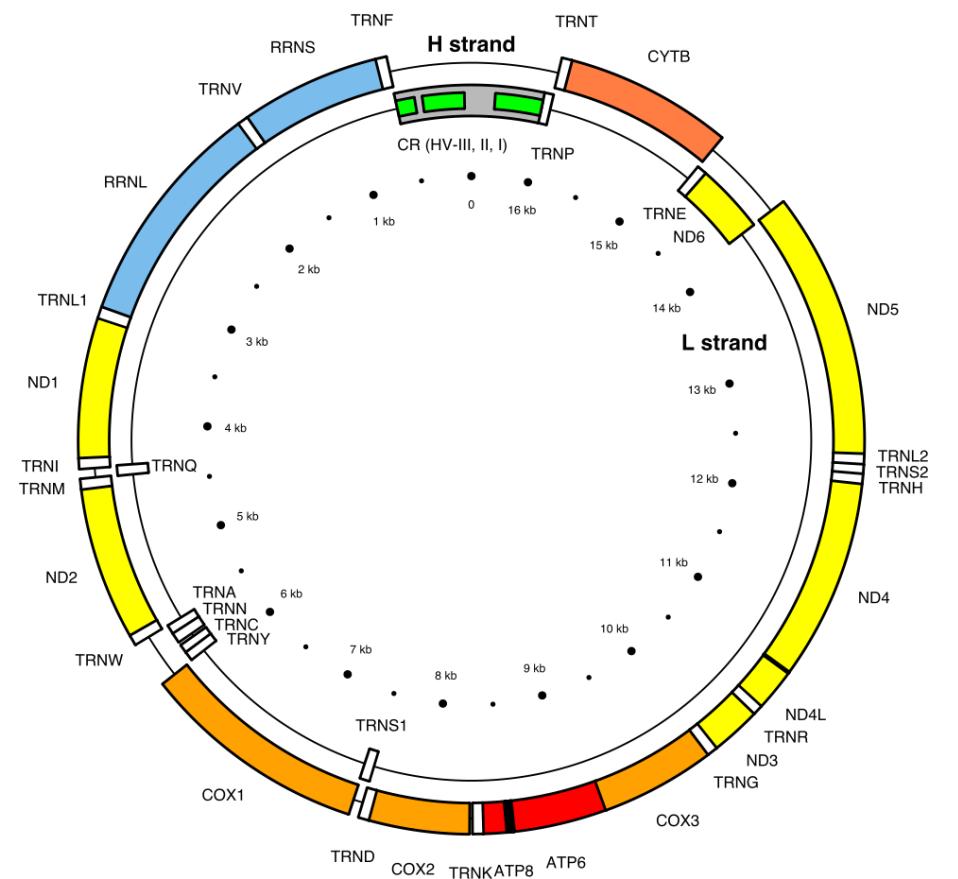


# No introgression?

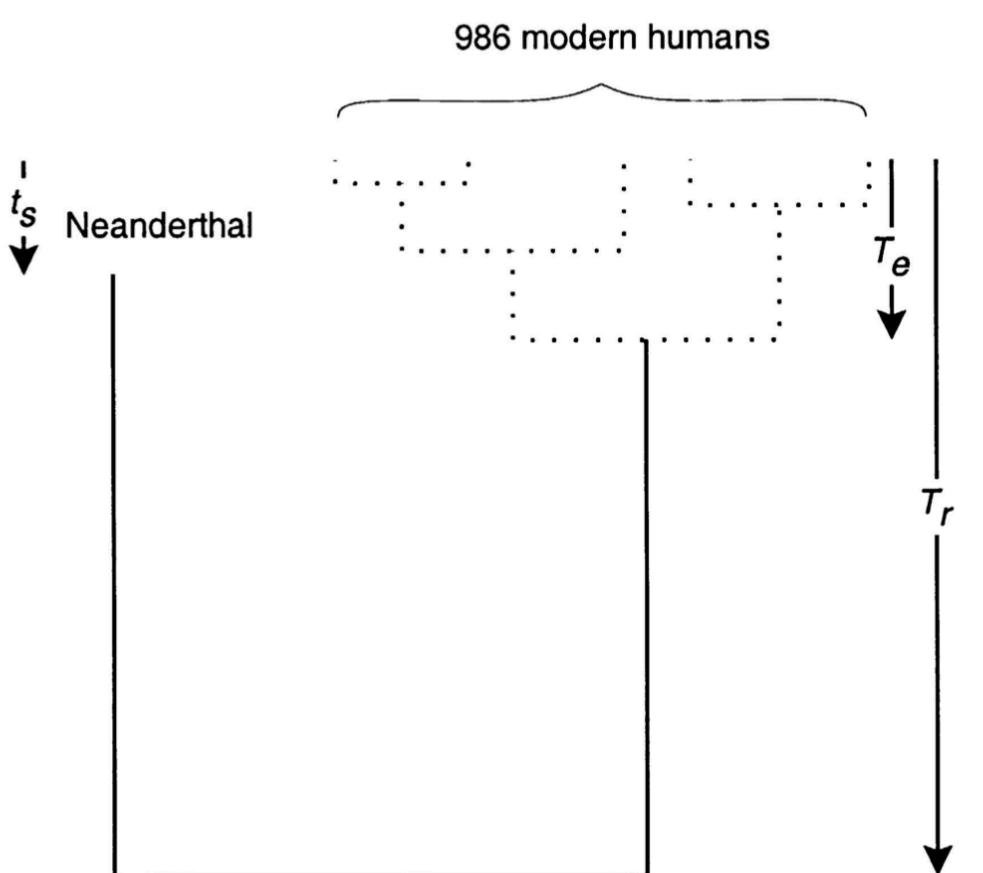


When the Neandertal mtDNA sequence is used to root a neighbor joining tree of modern human mtDNA sequences (Figure 7a), the first three branches consist exclusively of African sequences. The Neandertal mtDNA sequence thus supports a scenario in which modern humans arose recently in Africa as a distinct species and replaced Neandertals with little or no interbreeding.

# mitochondrial DNA



16 kb

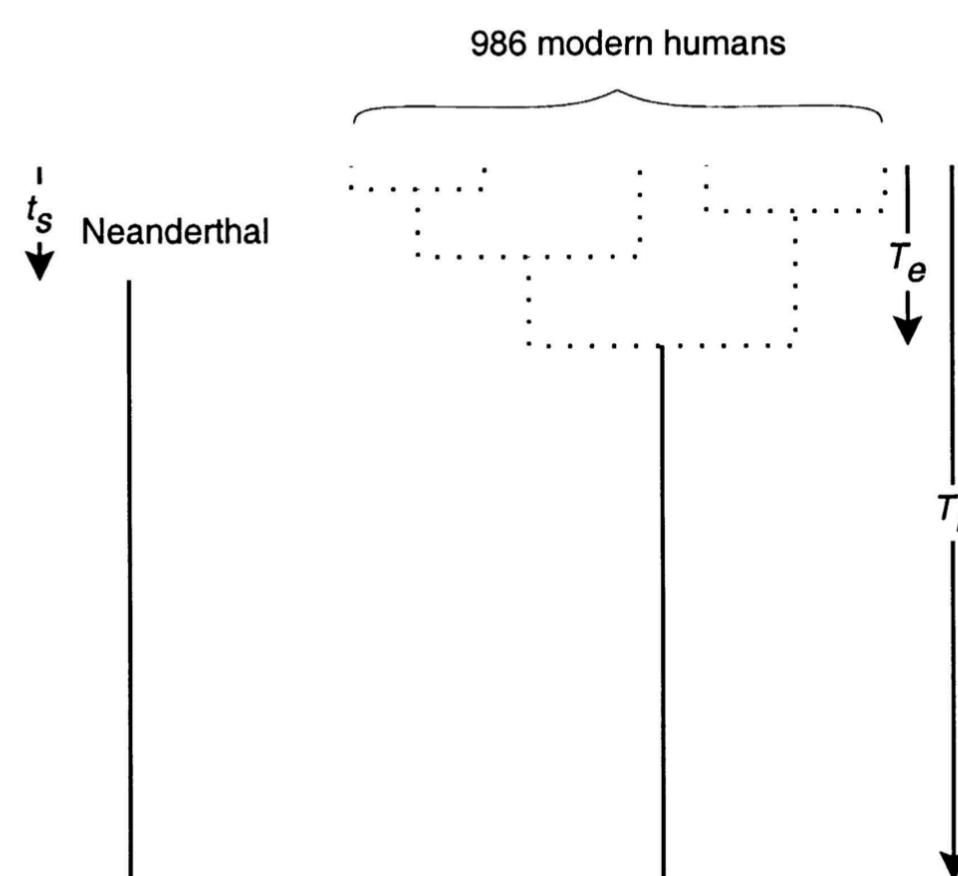
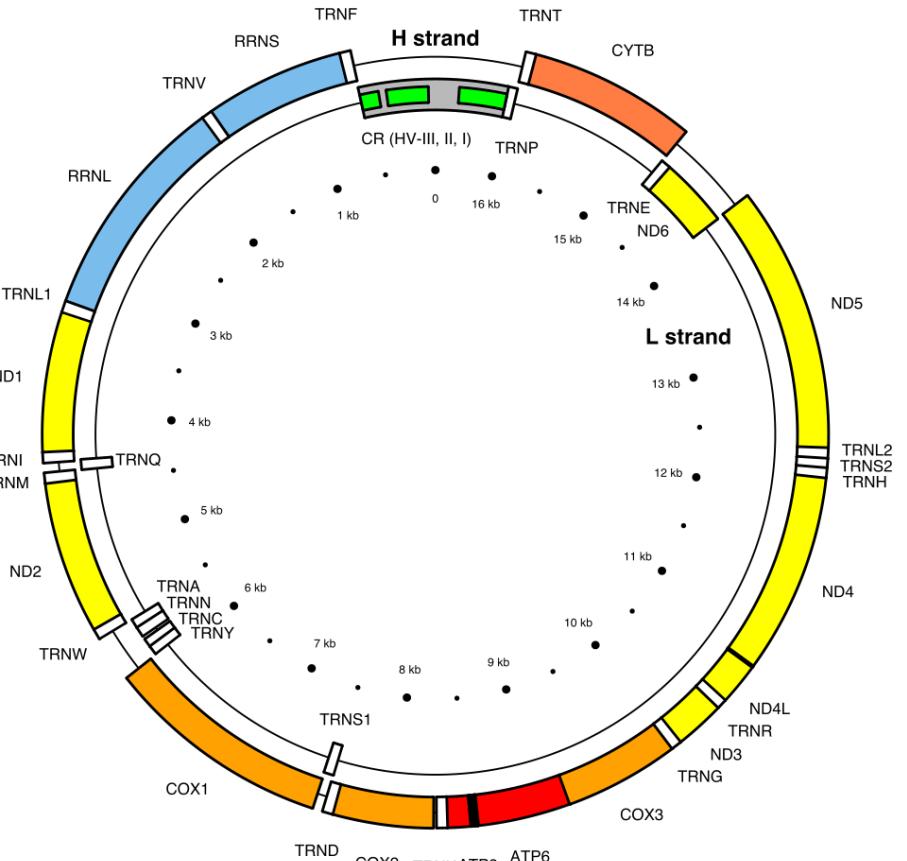


**one mtDNA phylogenetic tree**  
(maternal history)

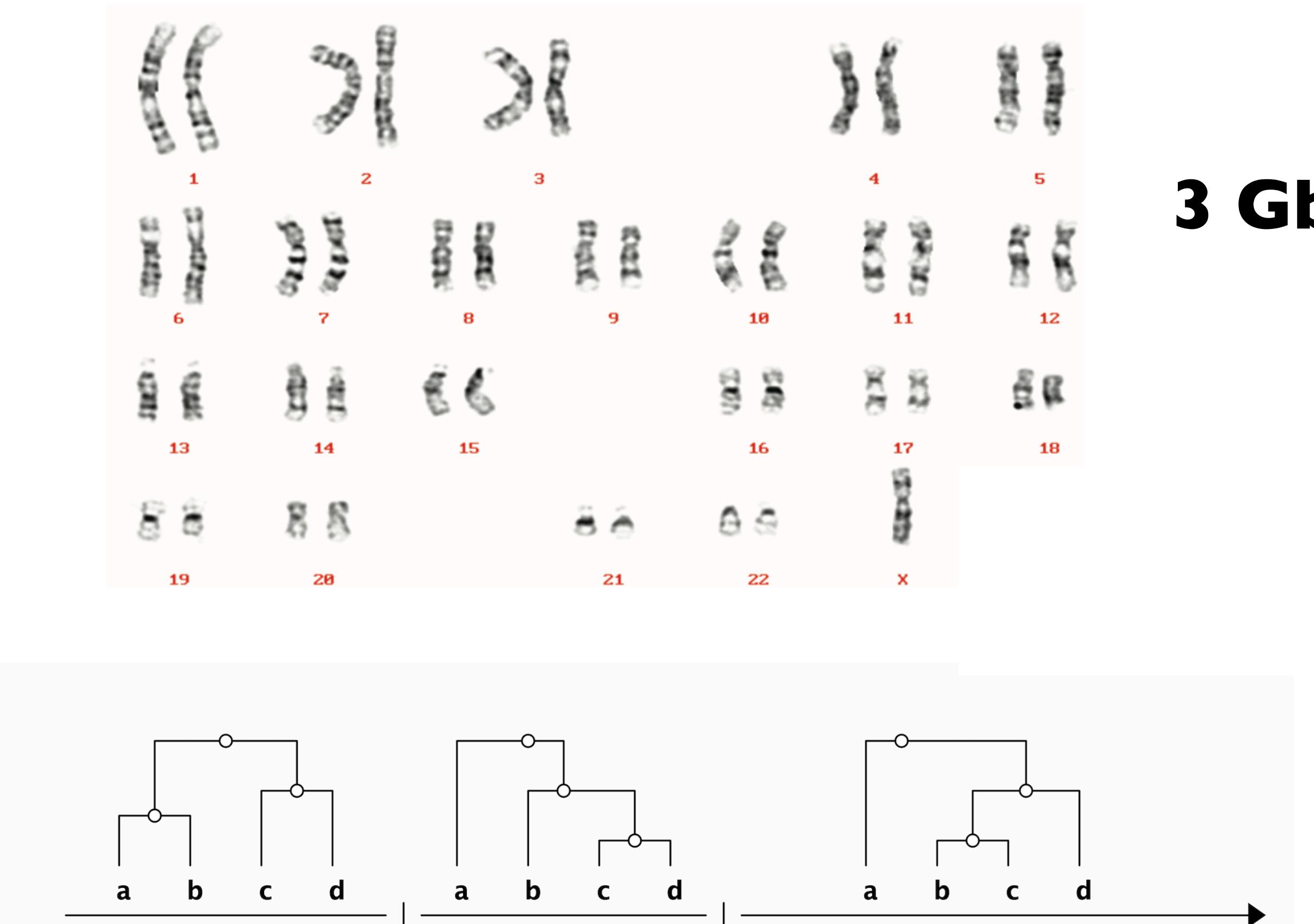
# mitochondrial DNA

**vs**

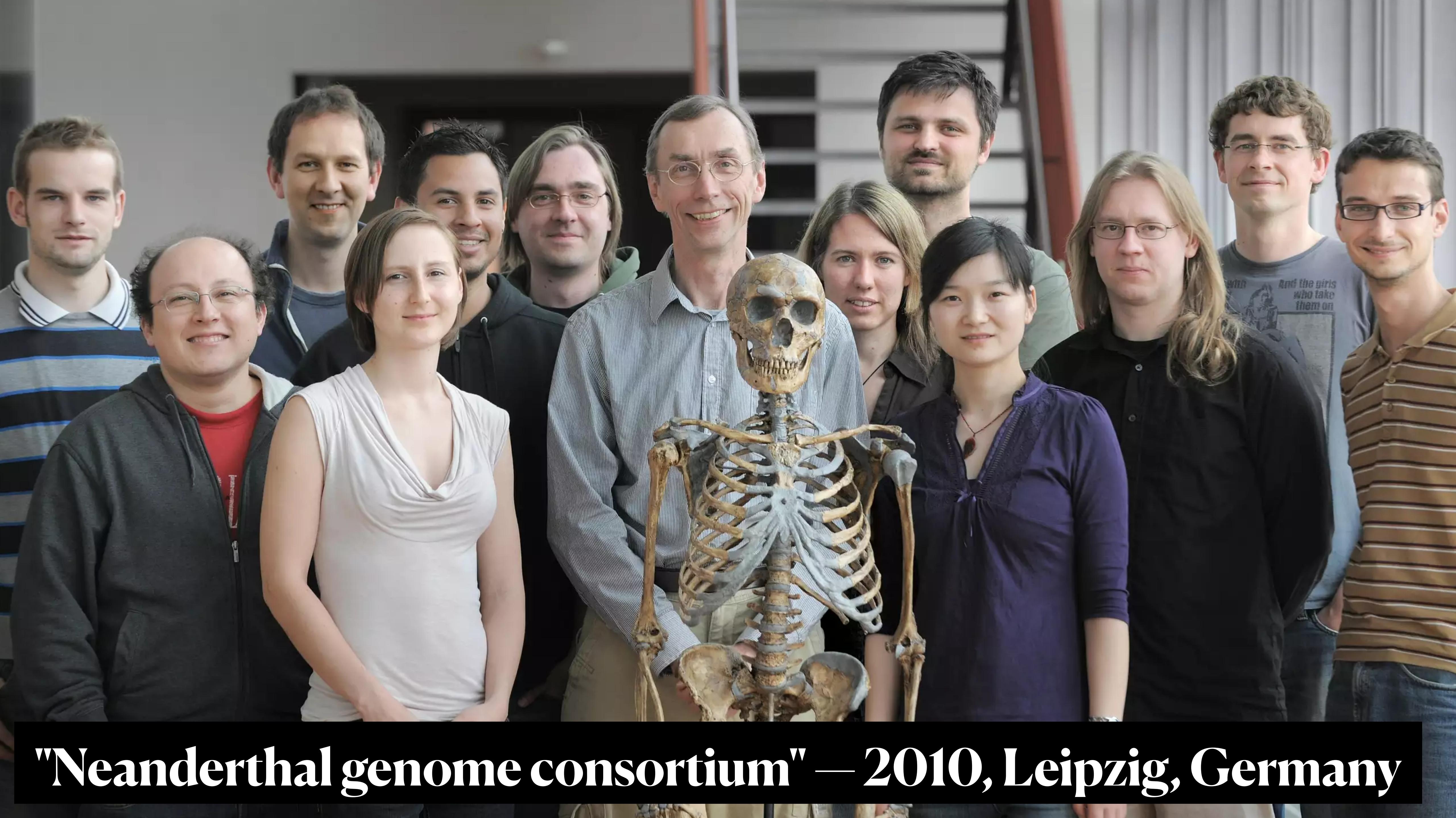
# nuclear DNA



**one mtDNA phylogenetic tree**  
(maternal history)

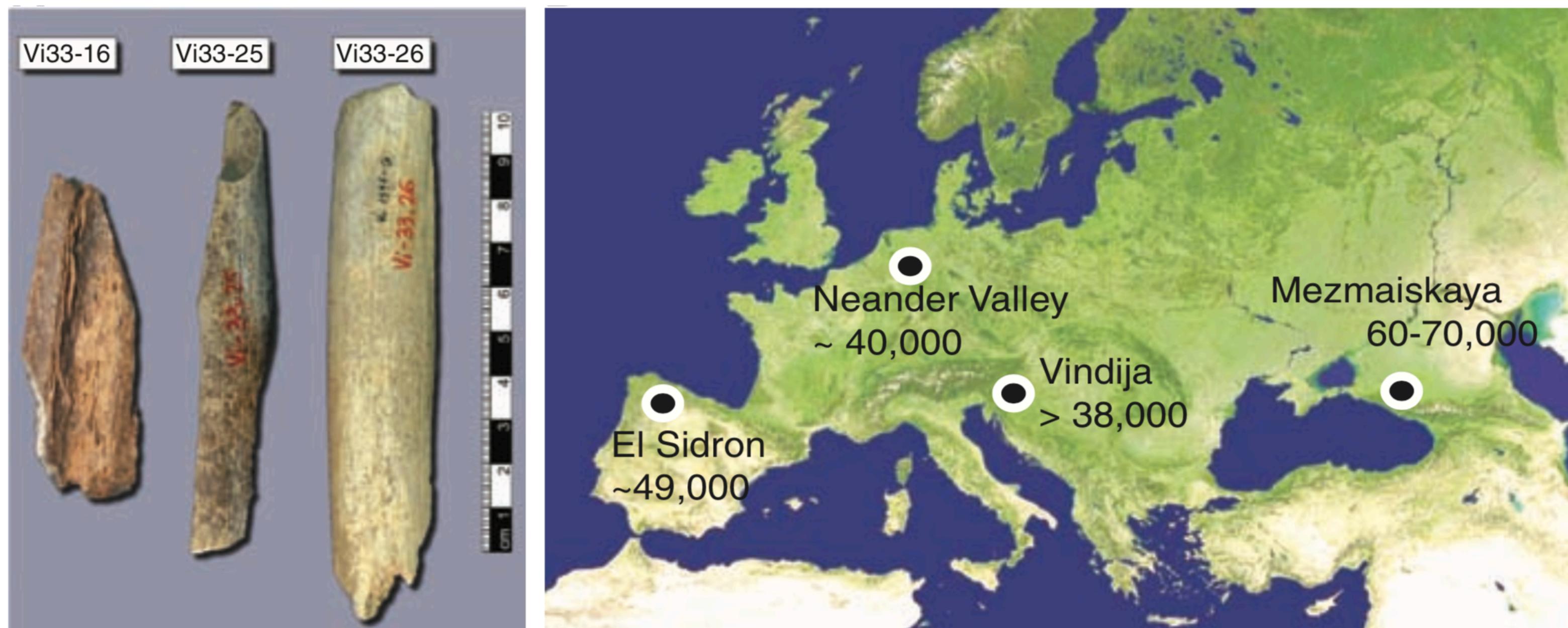


**thousands of trees along the nuclear genome**  
(result of recombination)



"Neanderthal genome consortium" – 2010, Leipzig, Germany

# 2010: "draft" Neanderthal genome



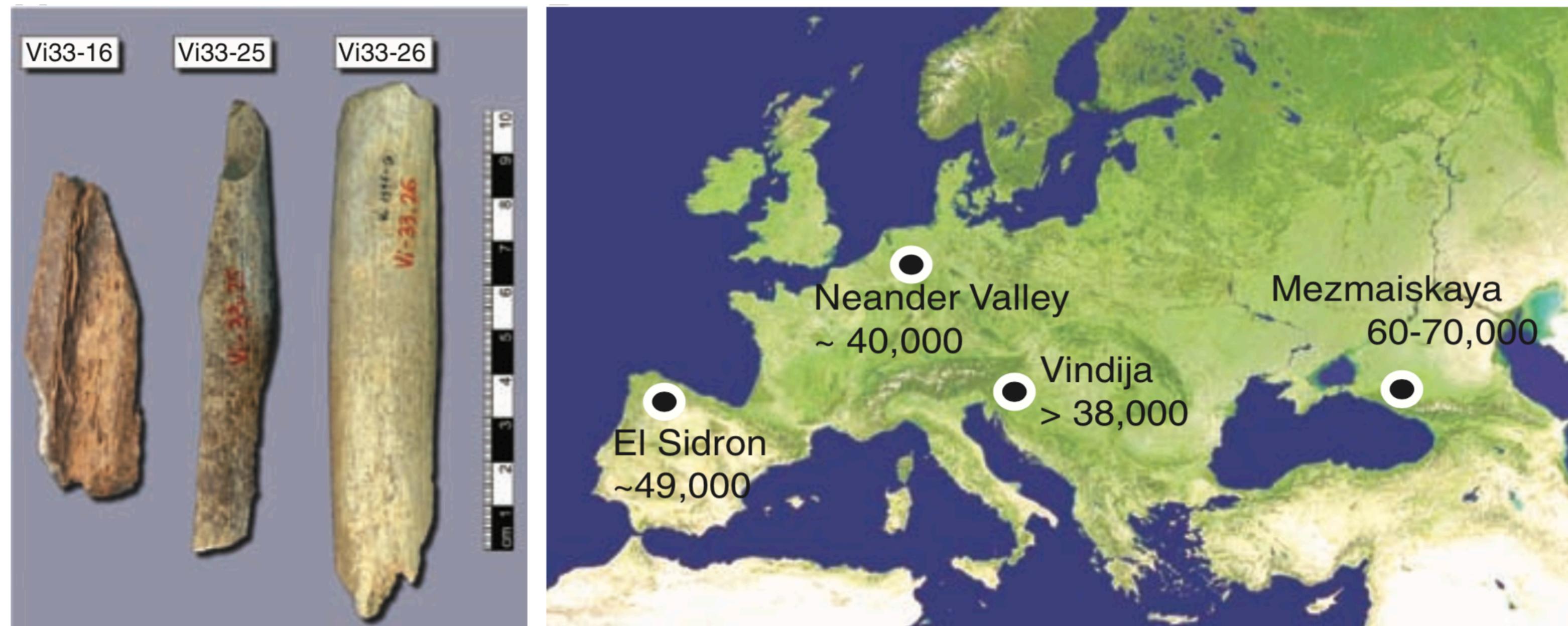
- DNA extracted from bone powder of three bones

# 2010: "draft" Neanderthal genome



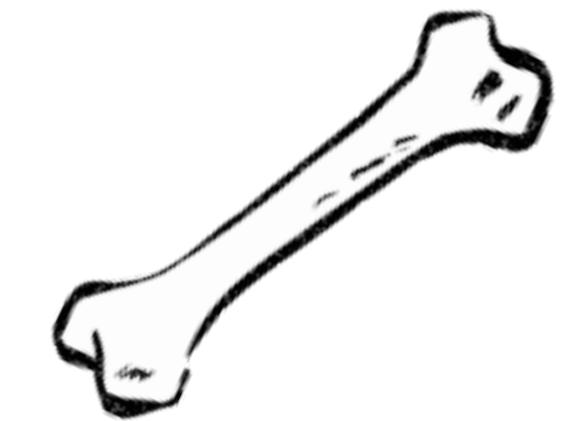
- DNA extracted from bone powder of three bones
- ~1.3X coverage genome

# 2010: "draft" Neanderthal genome



- DNA extracted from bone powder of three bones
- ~1.3X coverage genome
  - to this date four more high-coverage archaic genomes up to ~40X

# Typical aDNA workflow



# Typical aDNA workflow



**dentist  
drill**

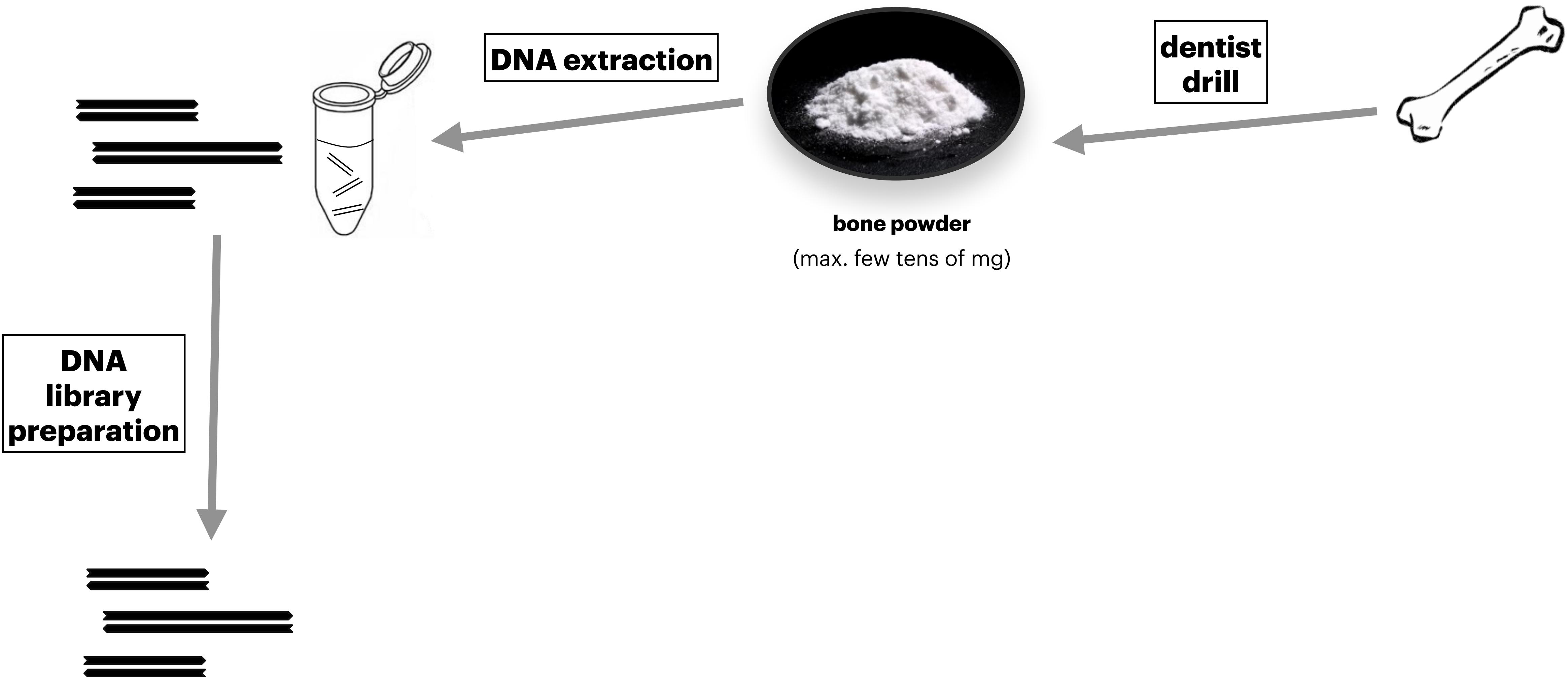


**bone powder**  
(max. few tens of mg)

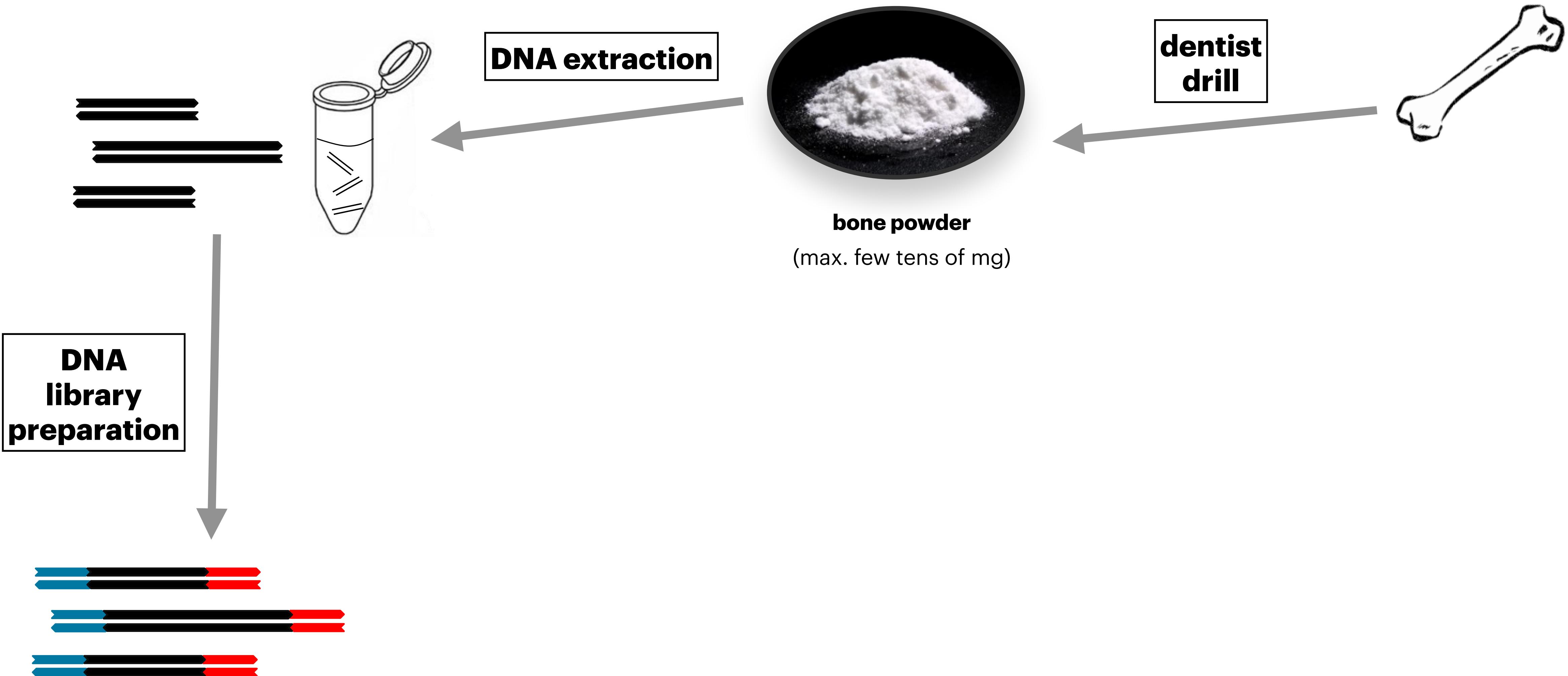
# Typical aDNA workflow



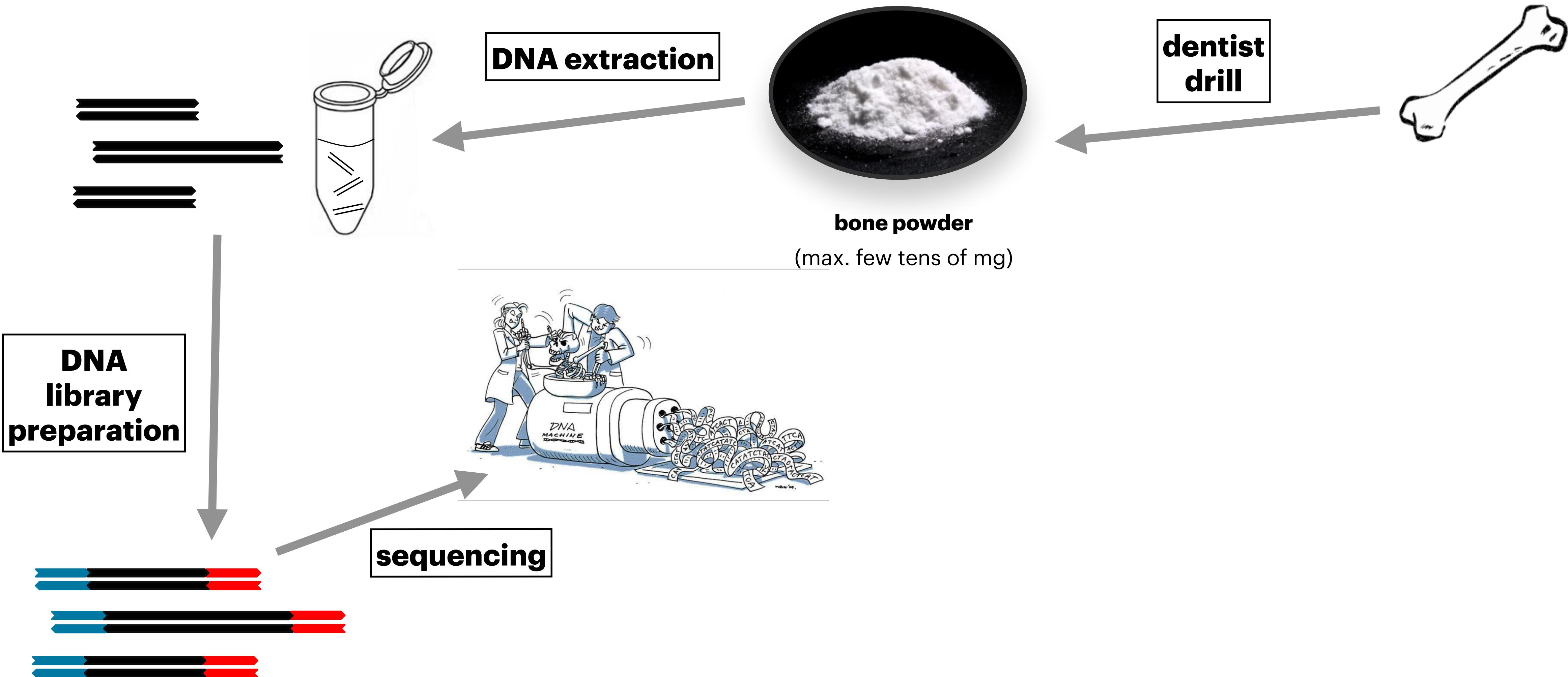
# Typical aDNA workflow



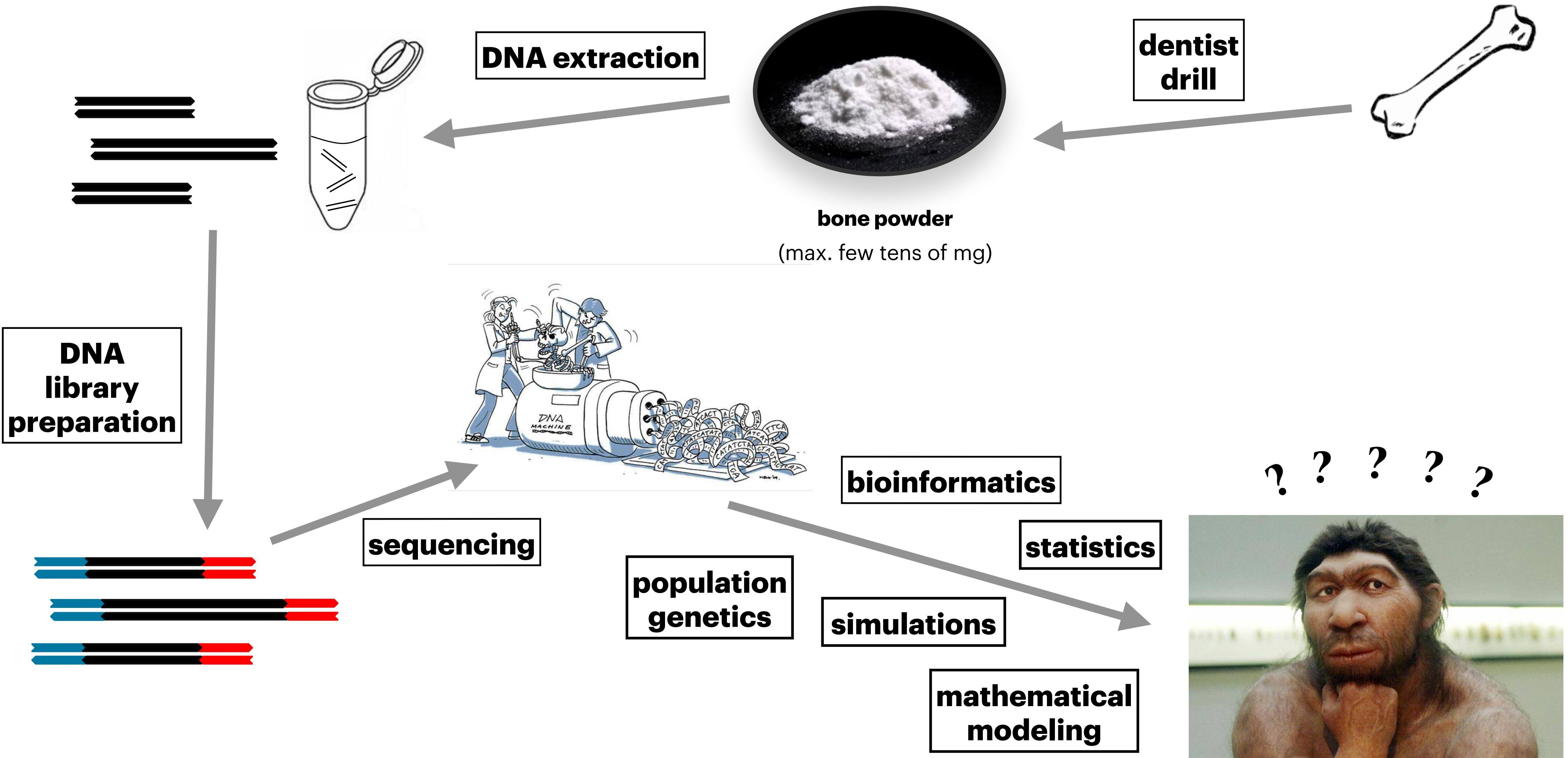
# Typical aDNA workflow



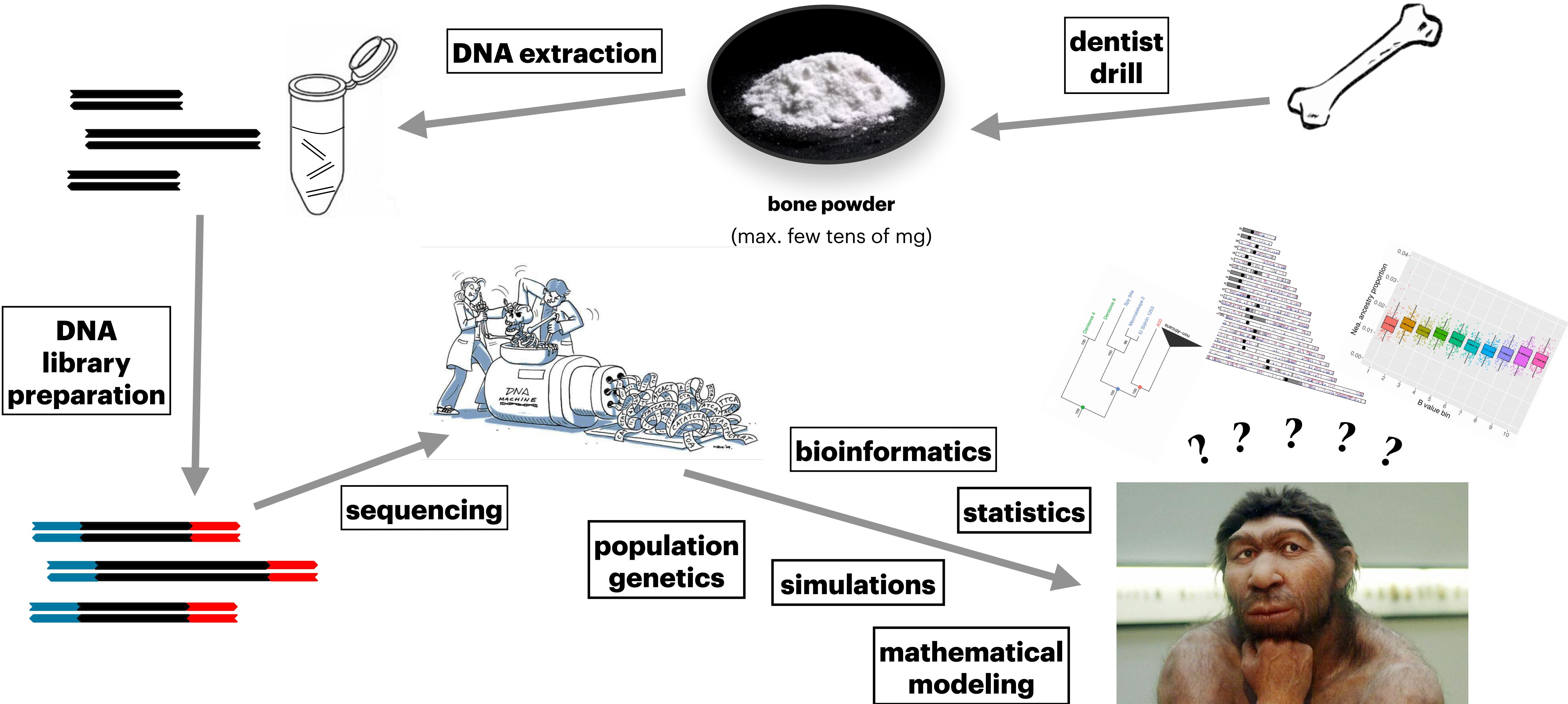
# Typical aDNA workflow



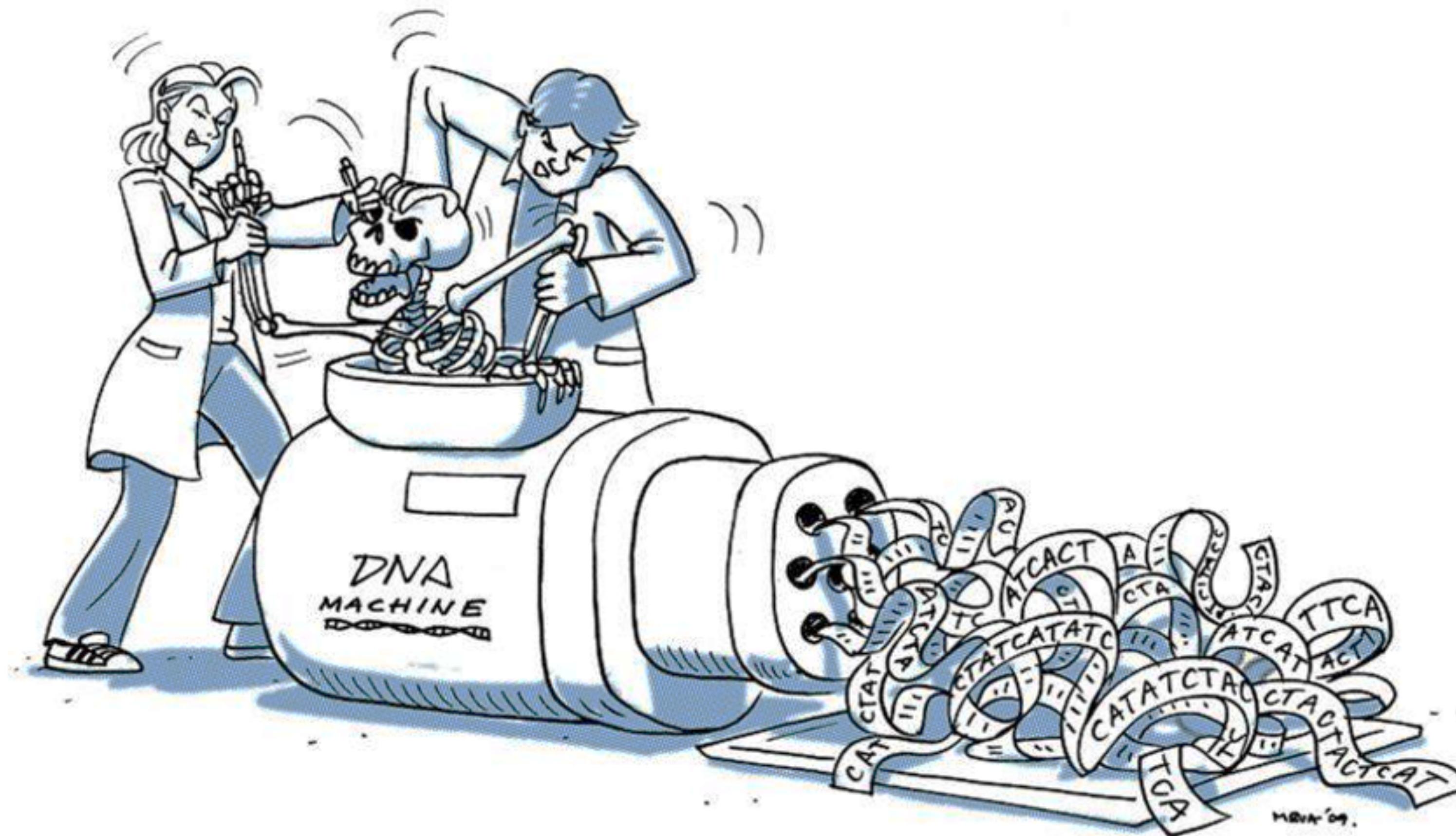
# Typical aDNA workflow



# Typical aDNA workflow



# So we have a Neanderthal genome... what now?

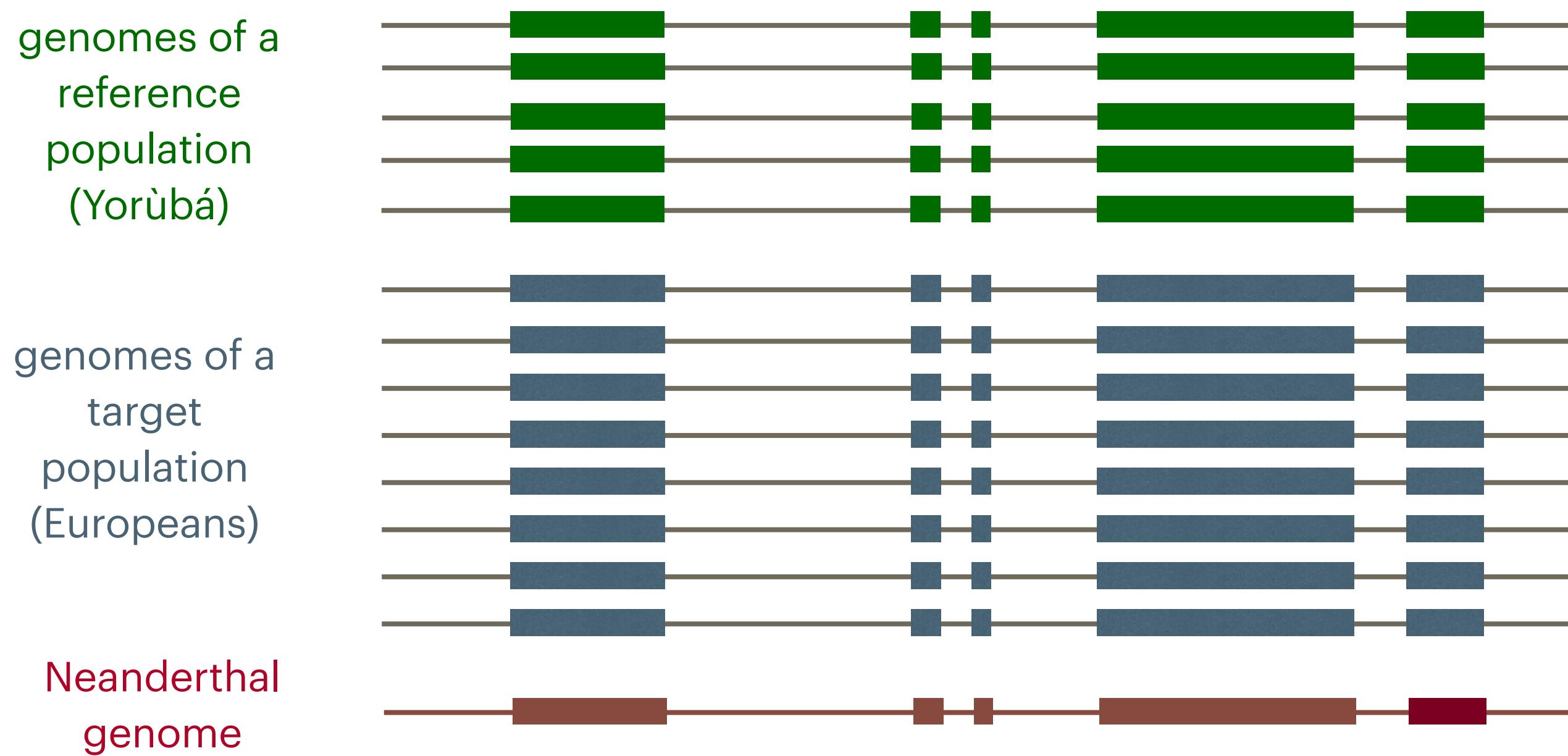


# How to test the gene-flow hypothesis?

---



# How to test the gene-flow hypothesis?



David Reich

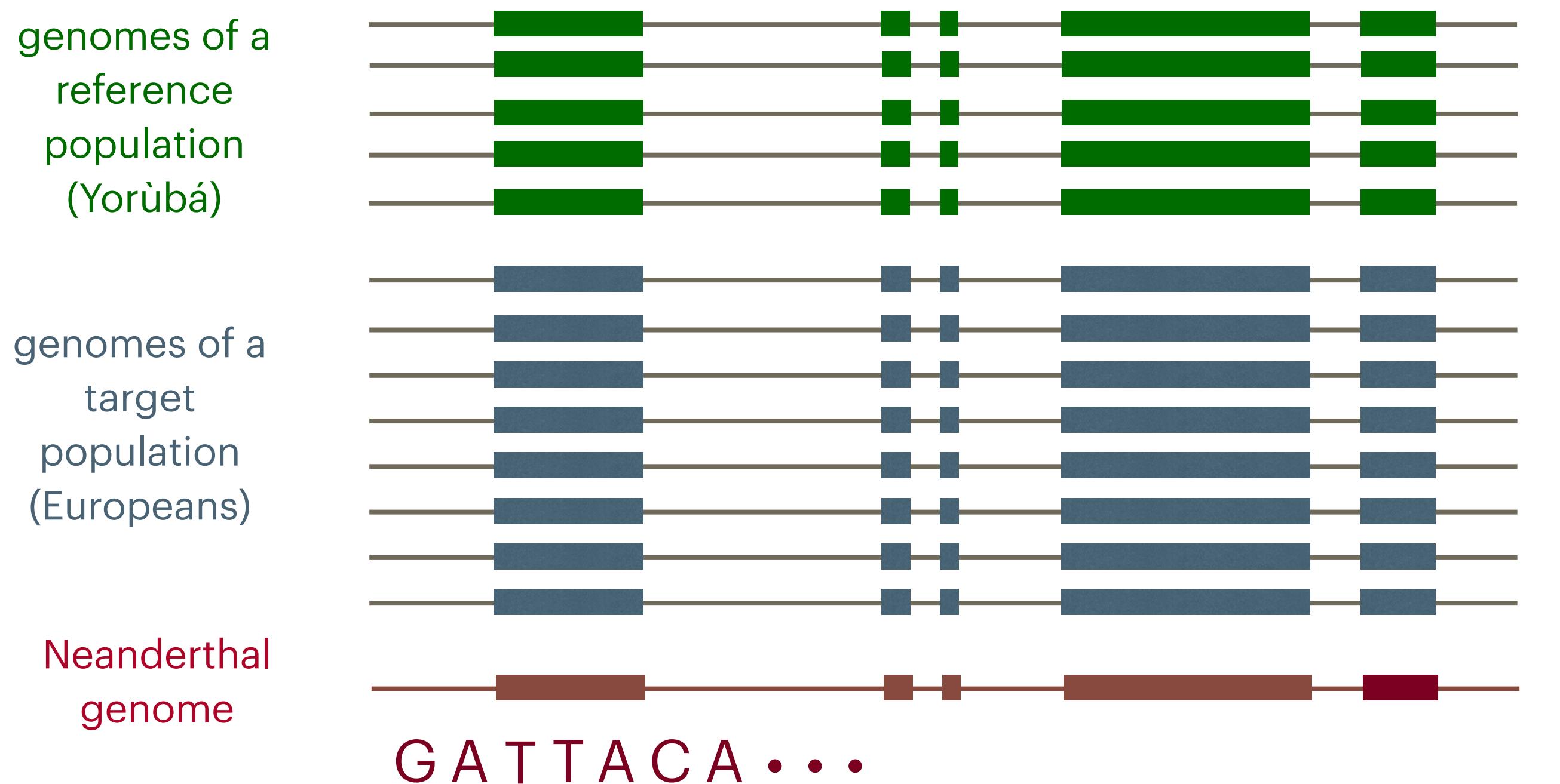


Nick Patterson



Green et al. (Science 2010)

# How to test the gene-flow hypothesis?



David Reich

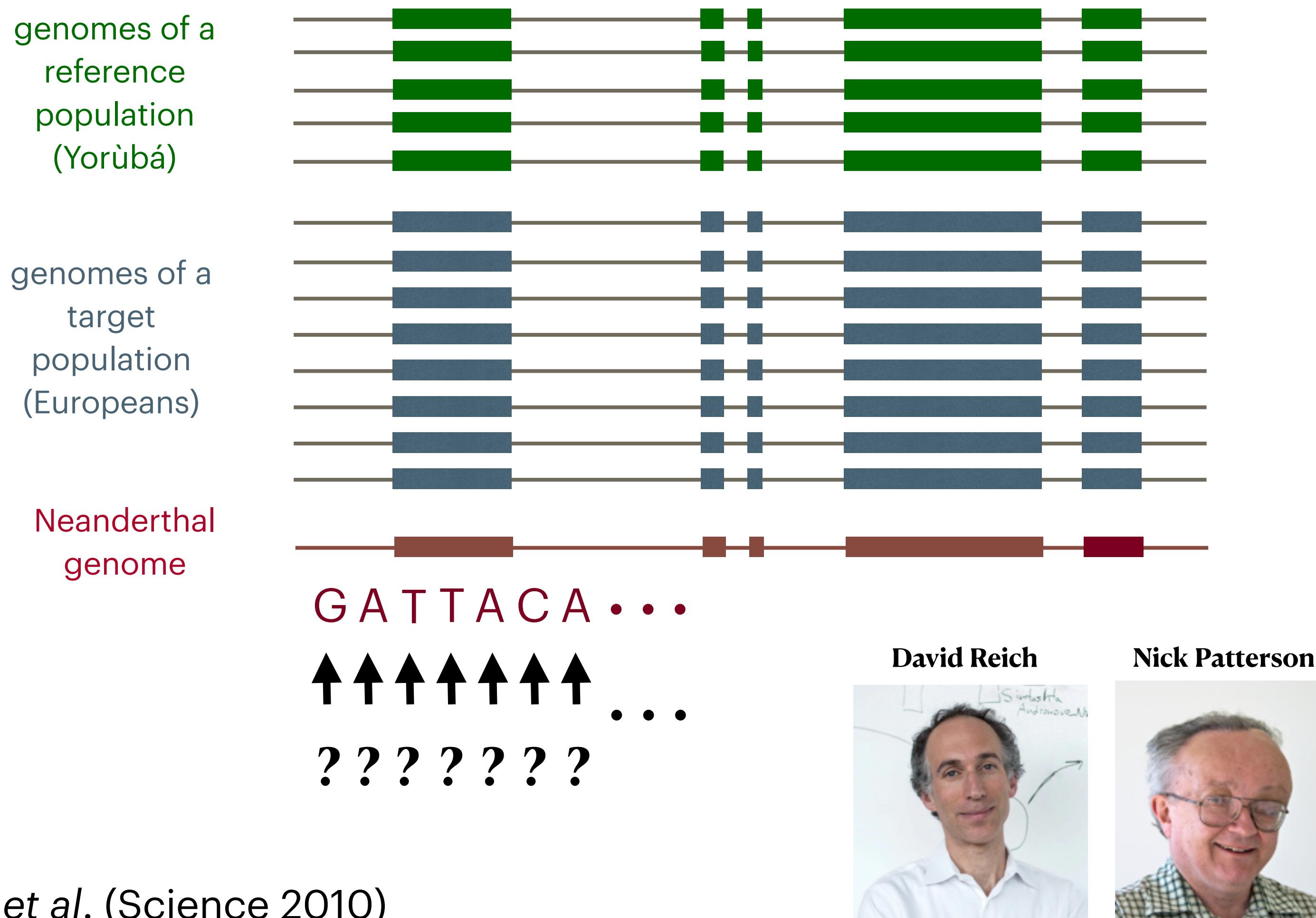


Nick Patterson



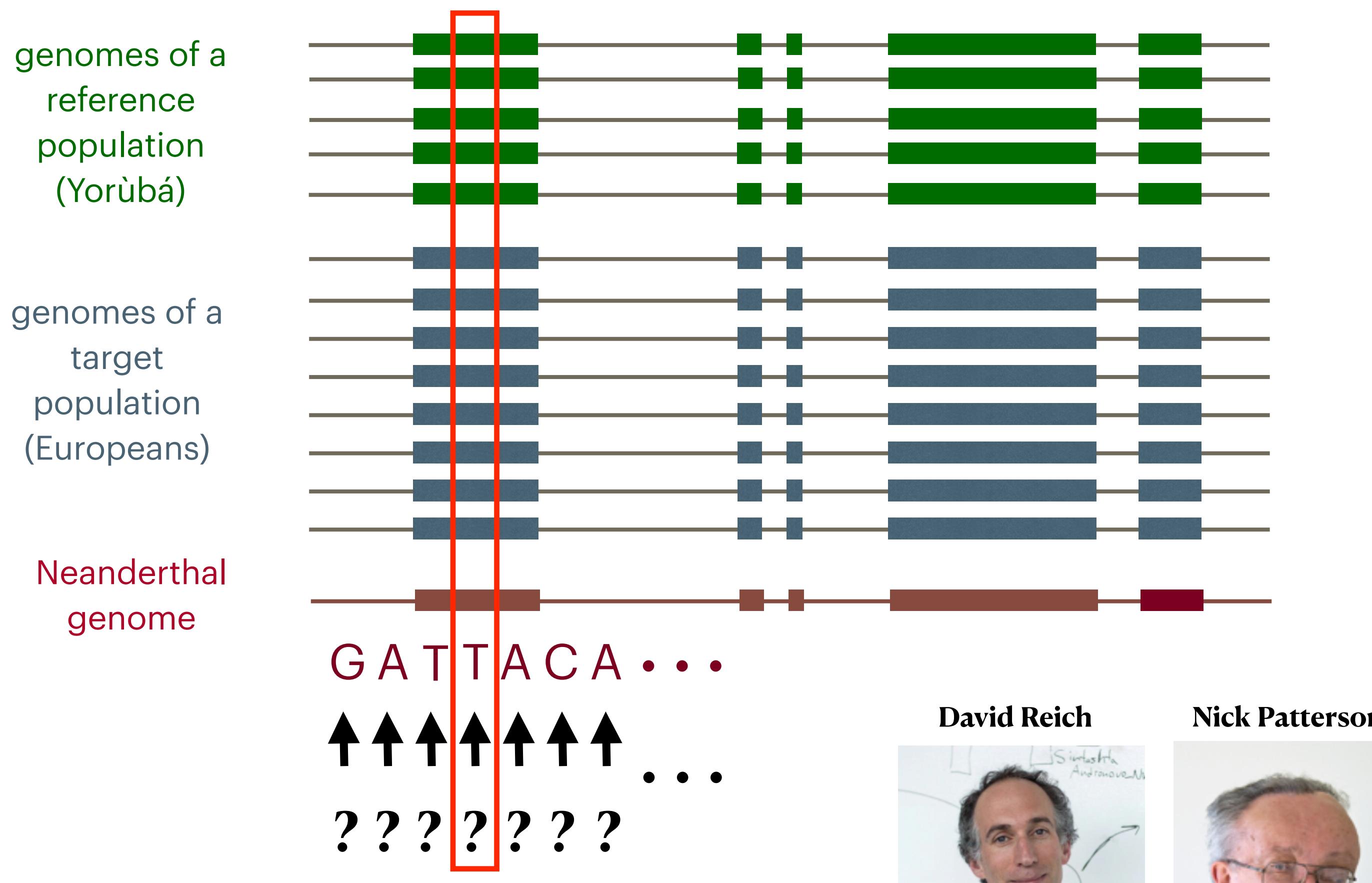
Green et al. (Science 2010)

# How to test the gene-flow hypothesis?



# How to test the gene-flow hypothesis?

# Situation at one locus...



Green et al. (Science 2010)

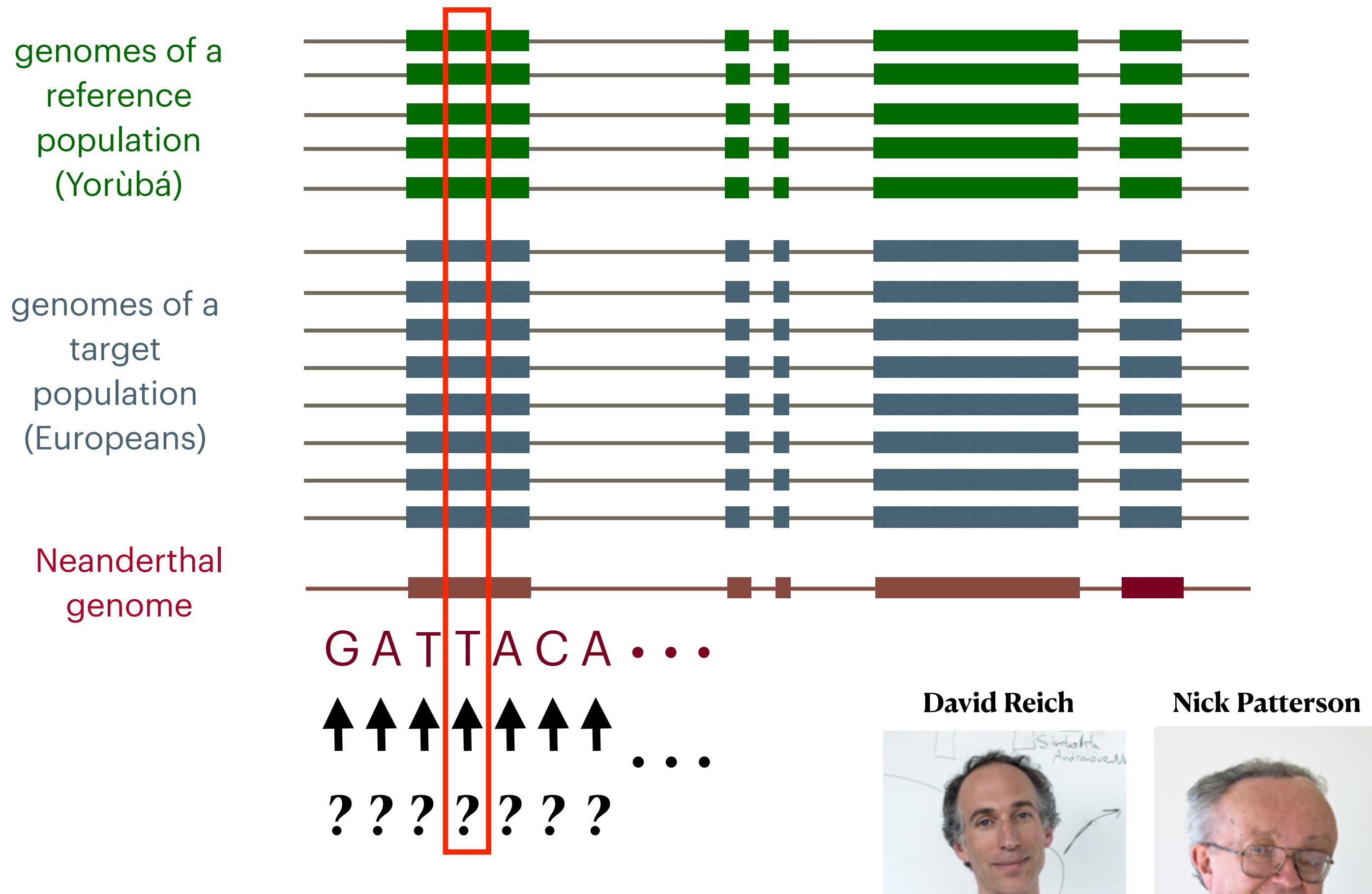


David Reic



Nick Patterson

# How to test the gene-flow hypothesis?

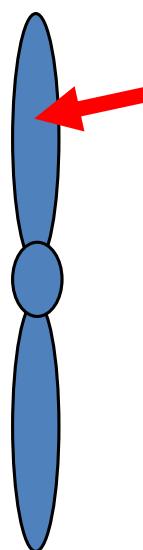


Situation at one locus...

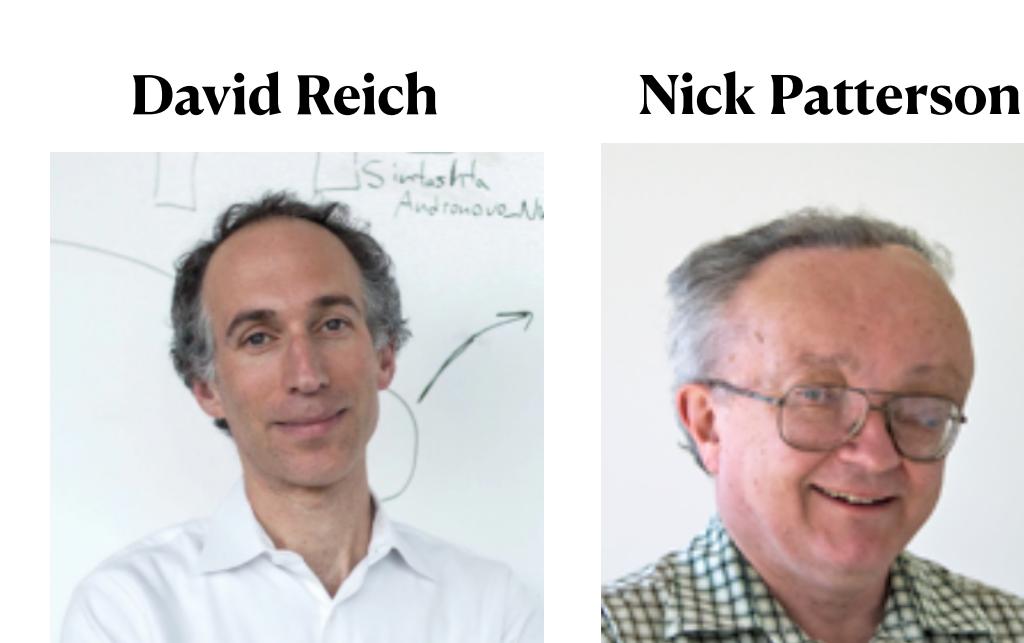
person of an African ancestry



person of a non-african ancestry (i.e. European)

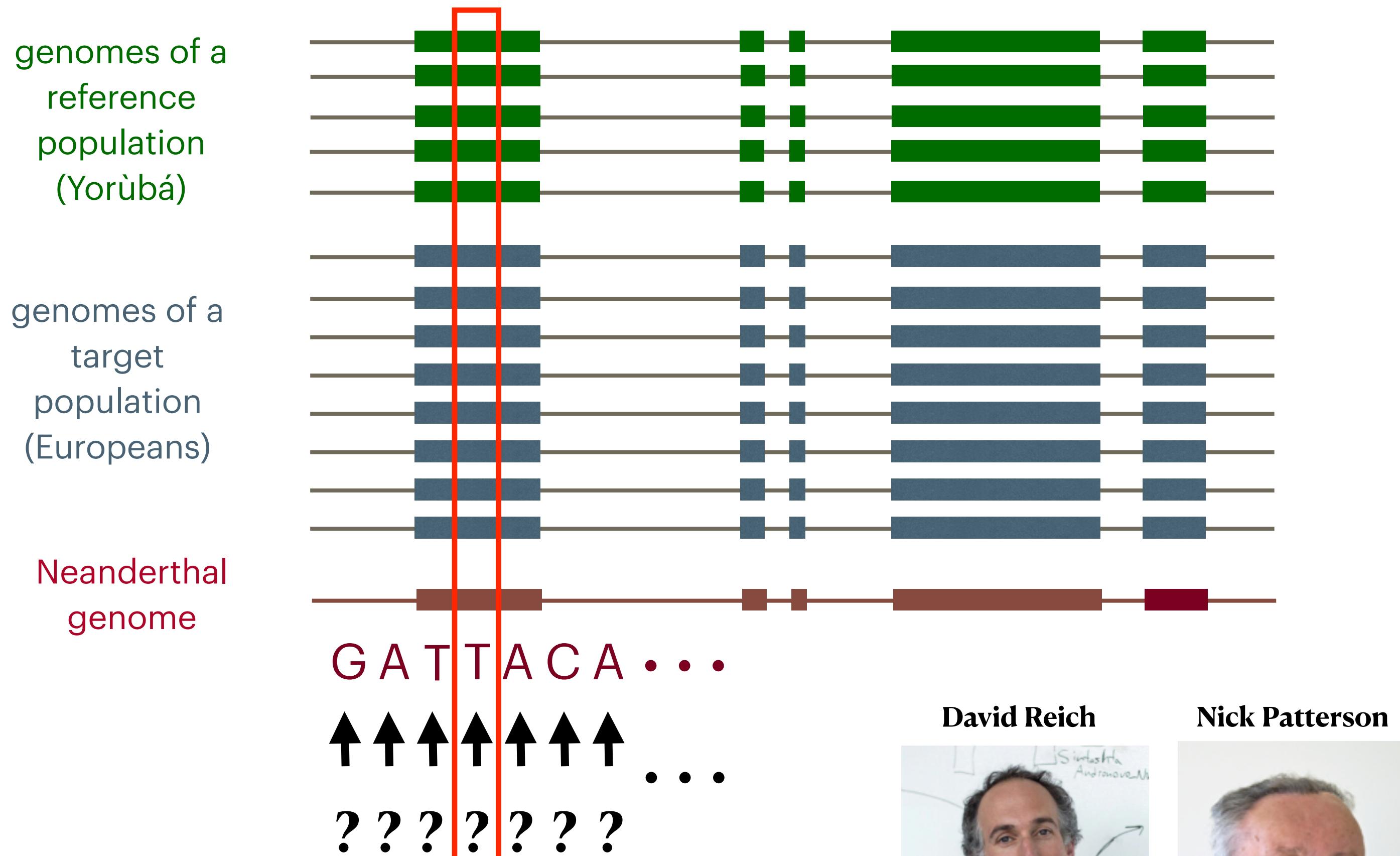


Neanderthal



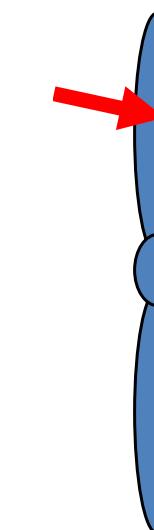
Green et al. (Science 2010)

# How to test the gene-flow hypothesis?

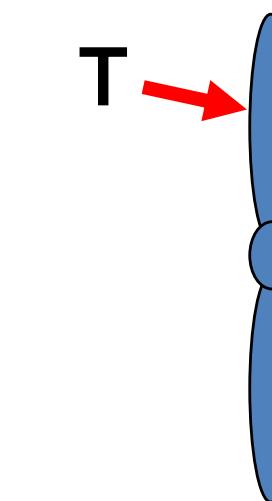
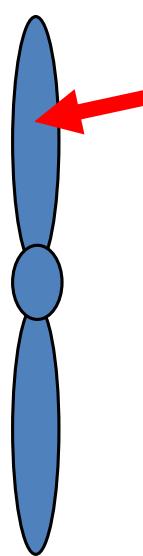


Situation at one locus...

person of an African ancestry

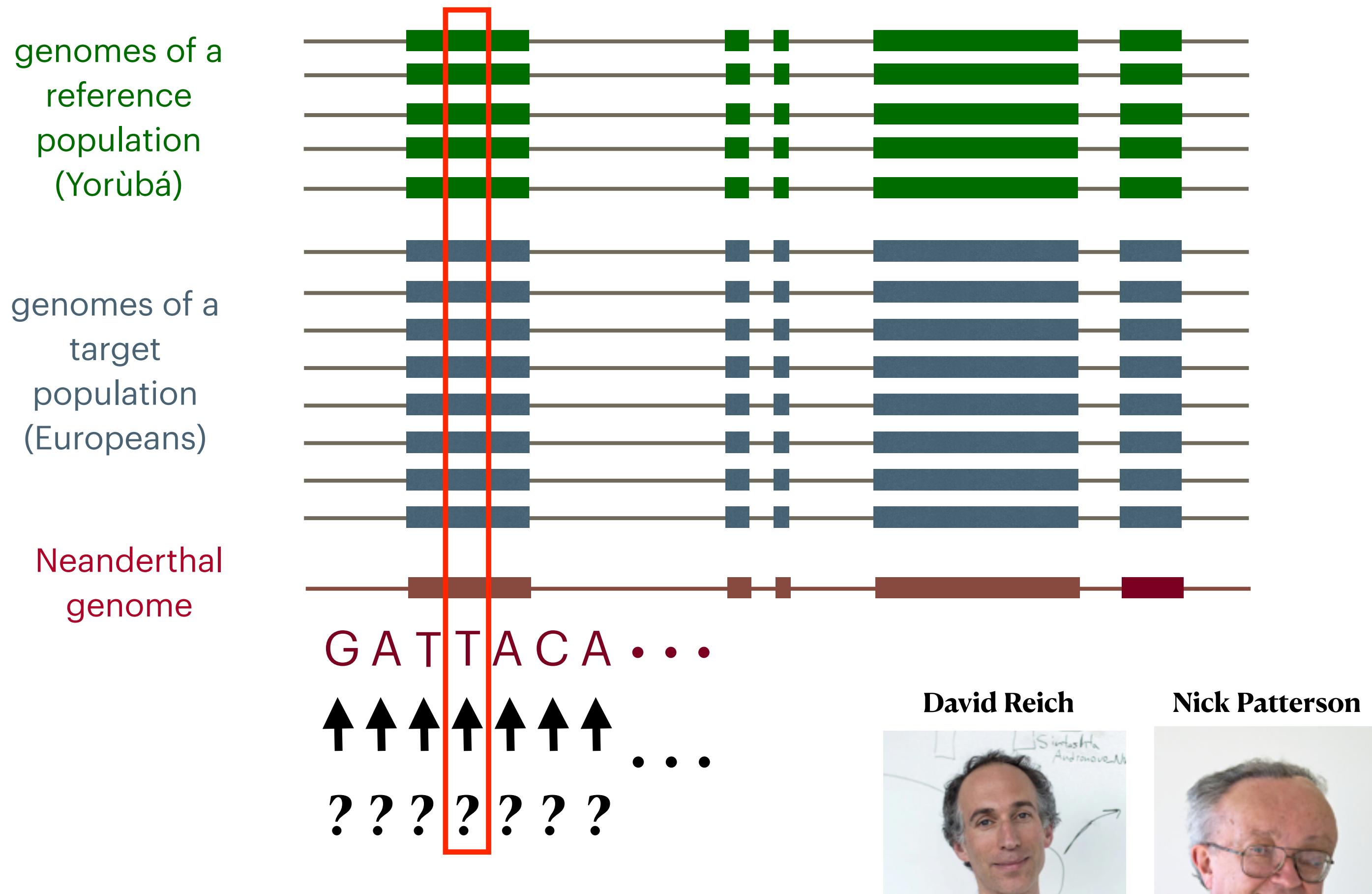


person of a non-african ancestry (i.e. European)



Green et al. (Science 2010)

# How to test the gene-flow hypothesis?

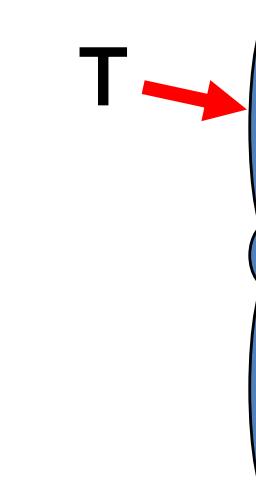
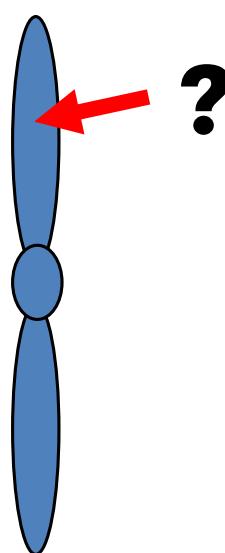


Situation at one locus...

person of an African ancestry



person of a non-african ancestry (i.e. European)



Neanderthal

Green et al. (Science 2010)

# Formal test of introgression

## $f_4$ statistic

Given a set of 4 samples,  $f_4$  compares counts of observed **BABA** vs **ABBA** site patterns.

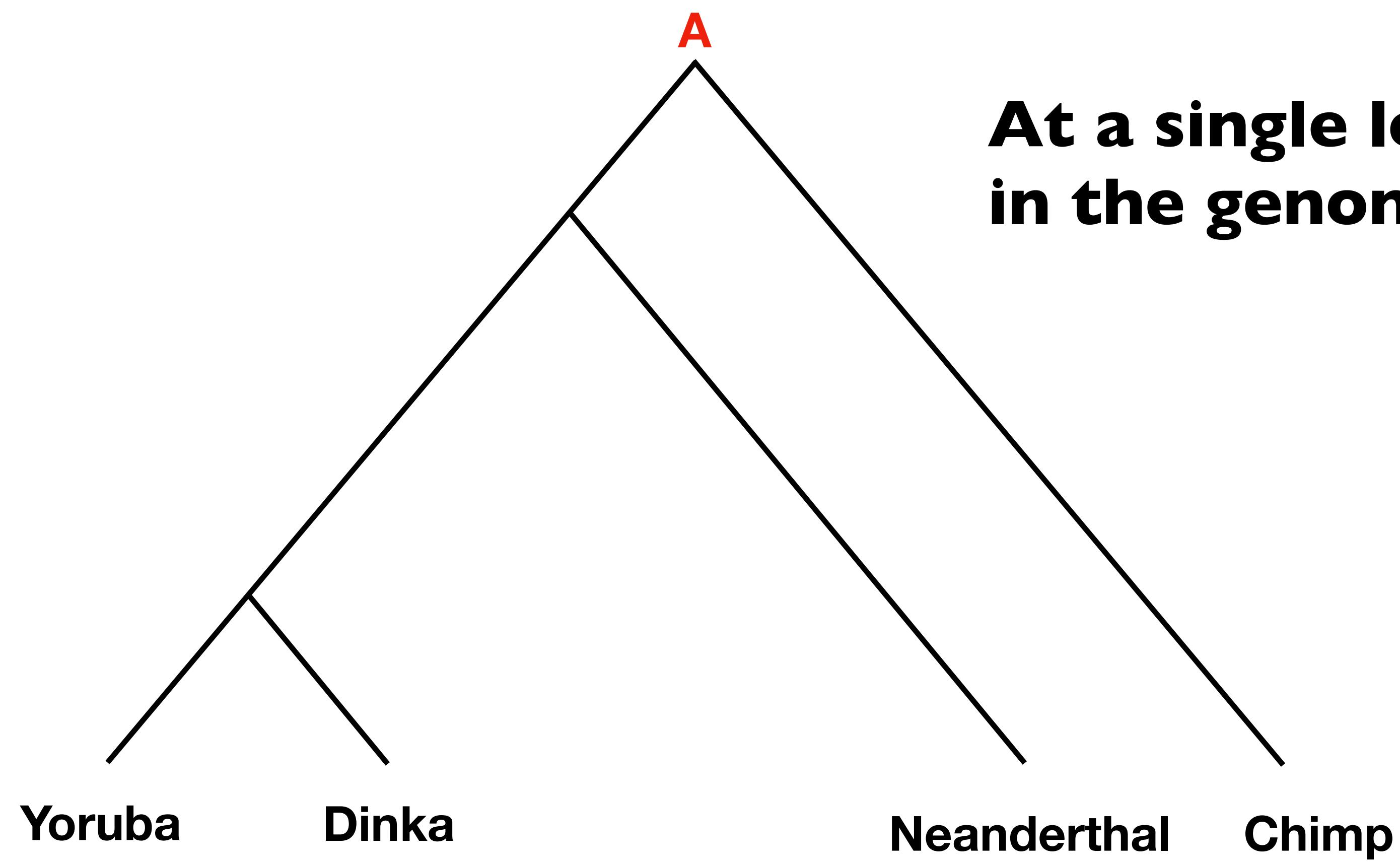
Nick Patterson



David Reich

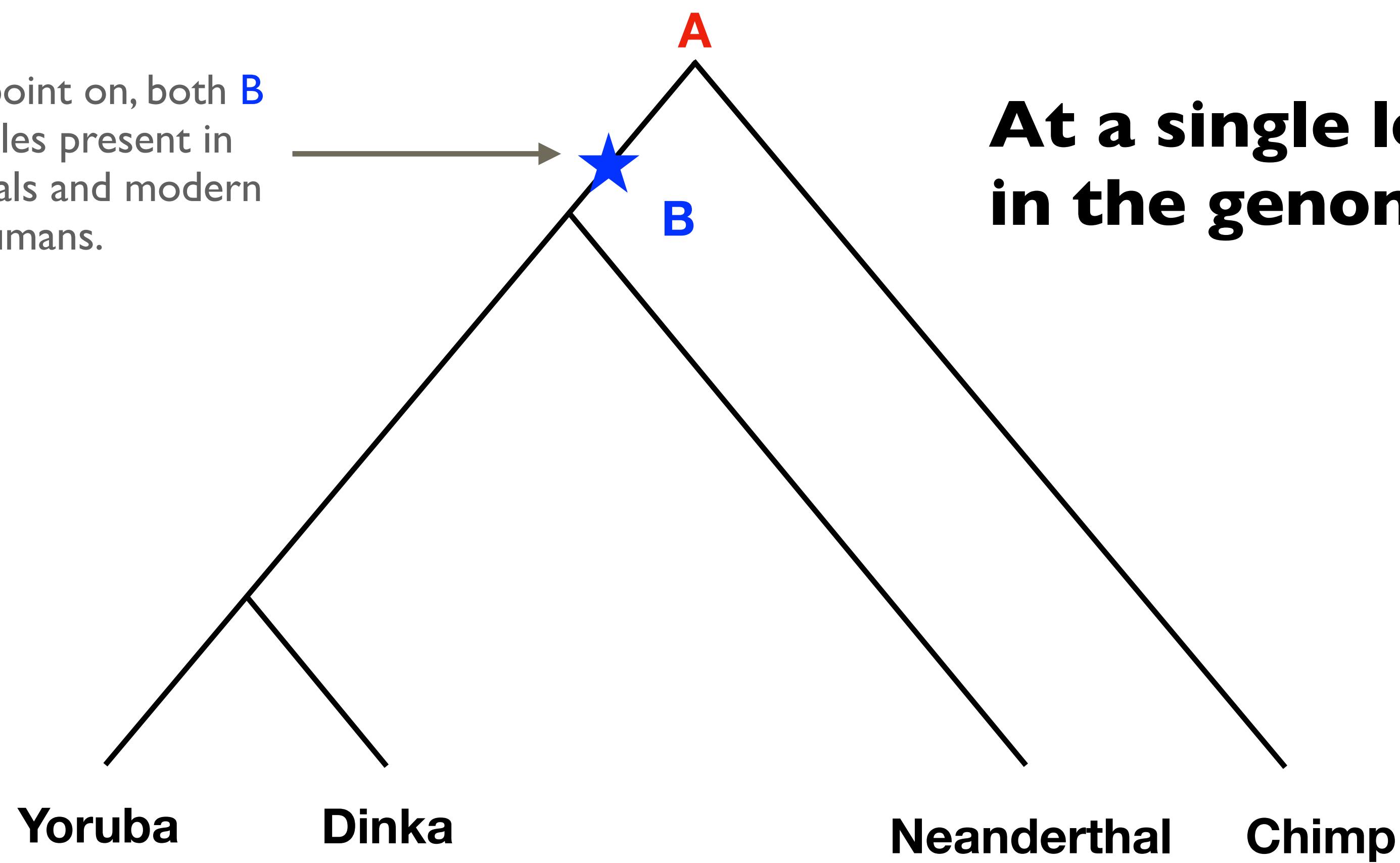


A variant of this statistic that is used equally often is known as **D statistic**.

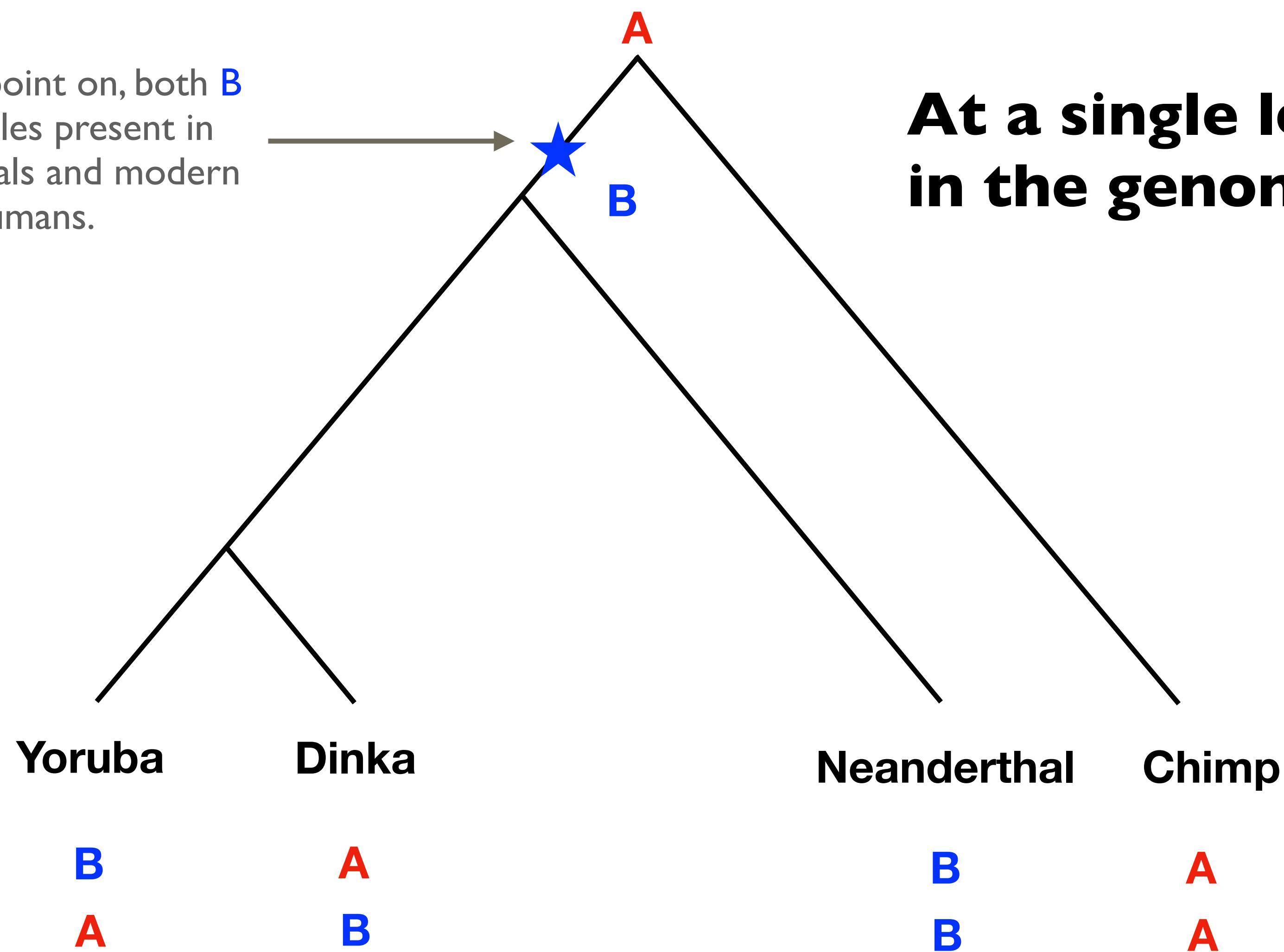


**At a single locus  
in the genome...**

From this point on, both **B** and **A** alleles present in Neanderthals and modern humans.

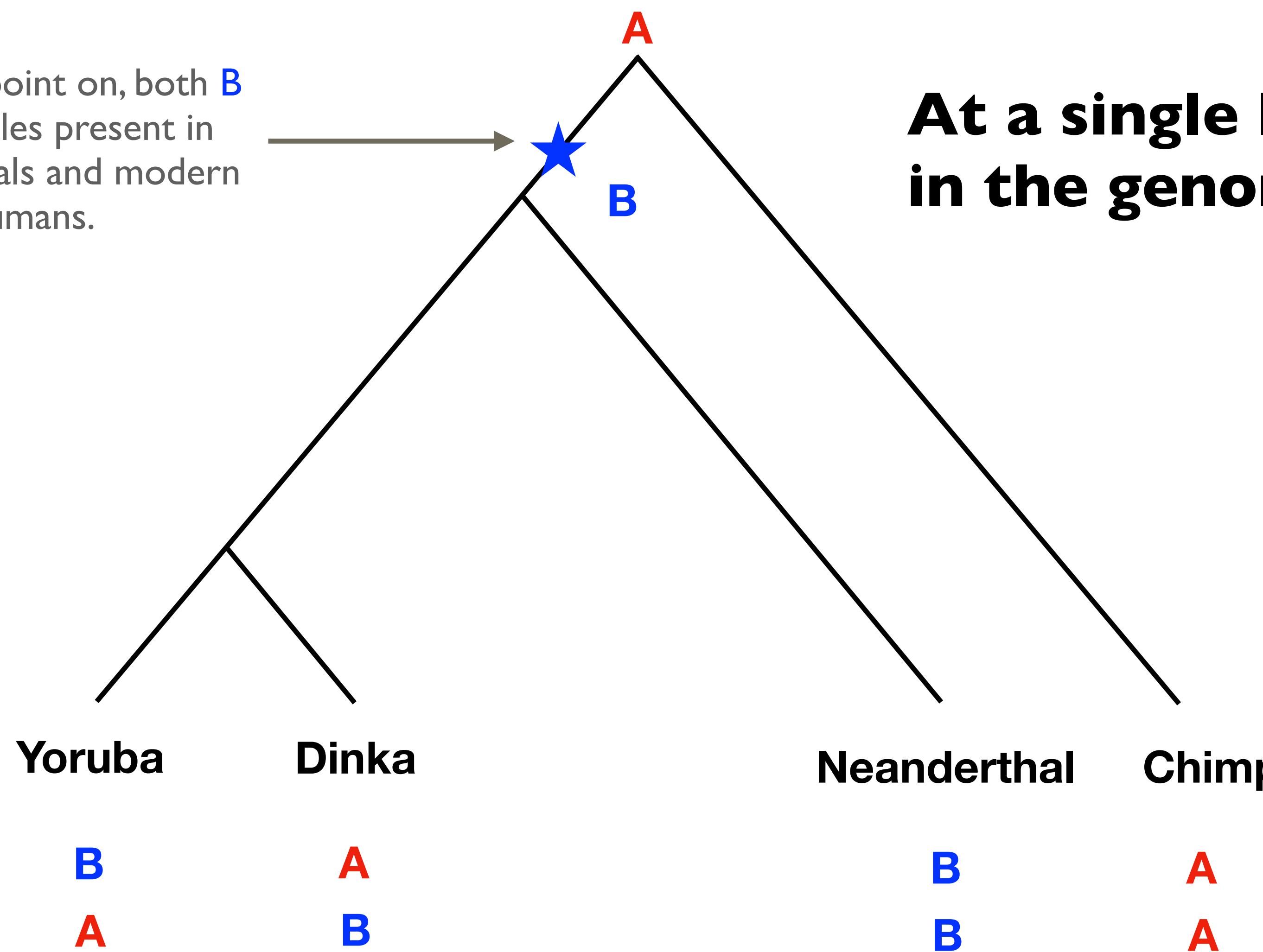


From this point on, both **B** and **A** alleles present in Neanderthals and modern humans.



**At a single locus  
in the genome...**

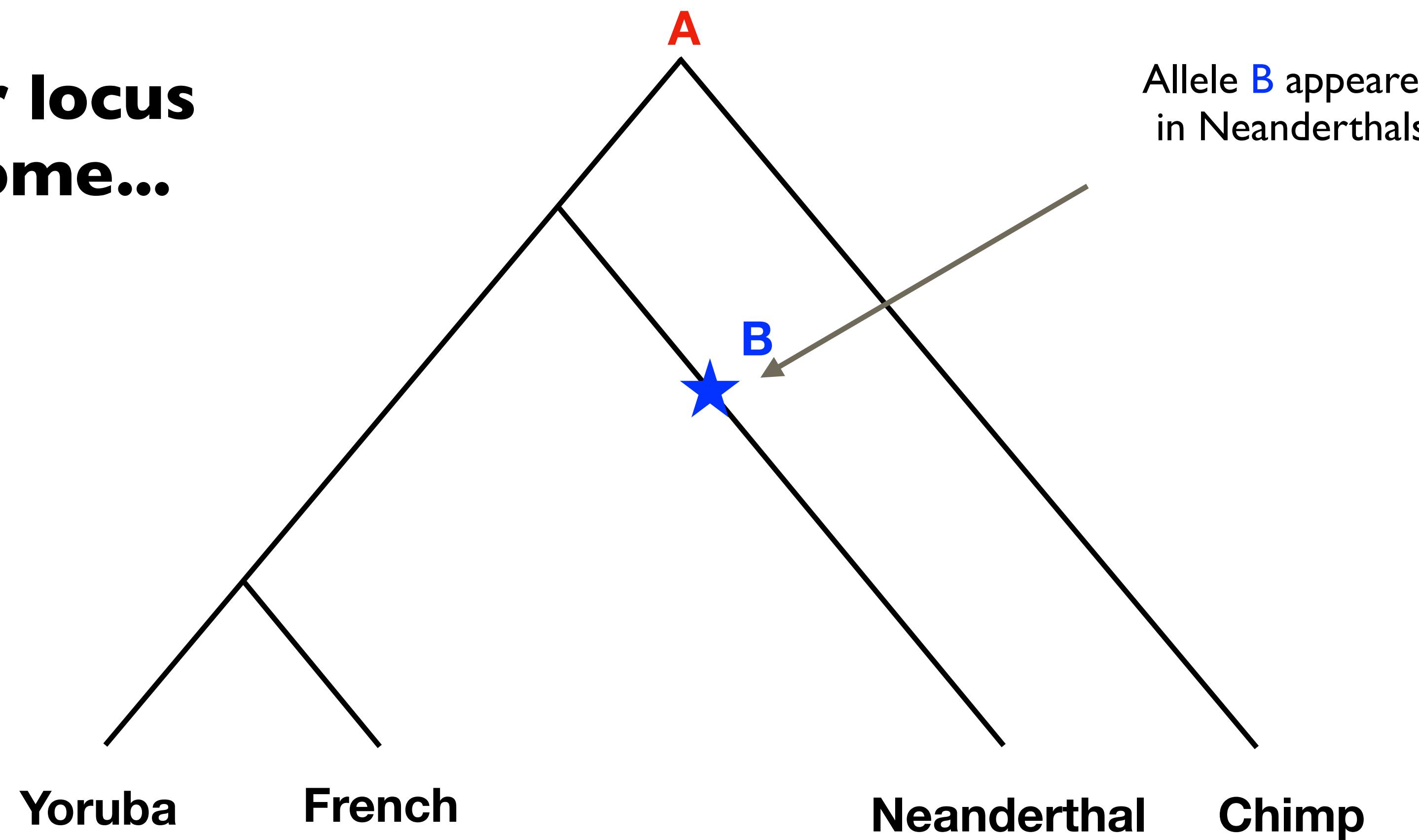
From this point on, both **B** and **A** alleles present in Neanderthals and modern humans.



**At a single locus  
in the genome...**

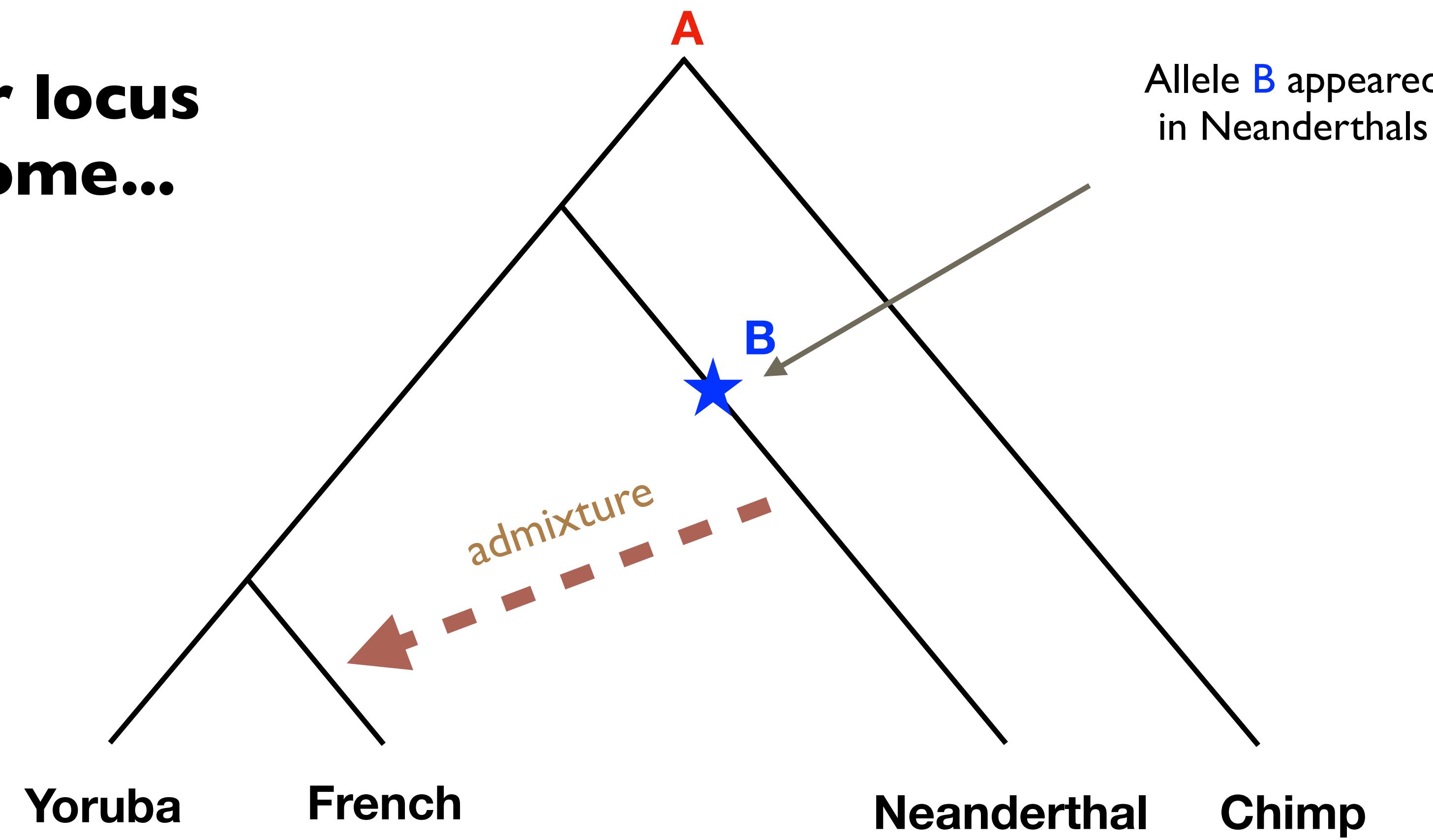
**No introgression: *f4 statistic* = (# BABA - #ABBA) / # SNPs ~ 0**

**At another locus  
in the genome...**

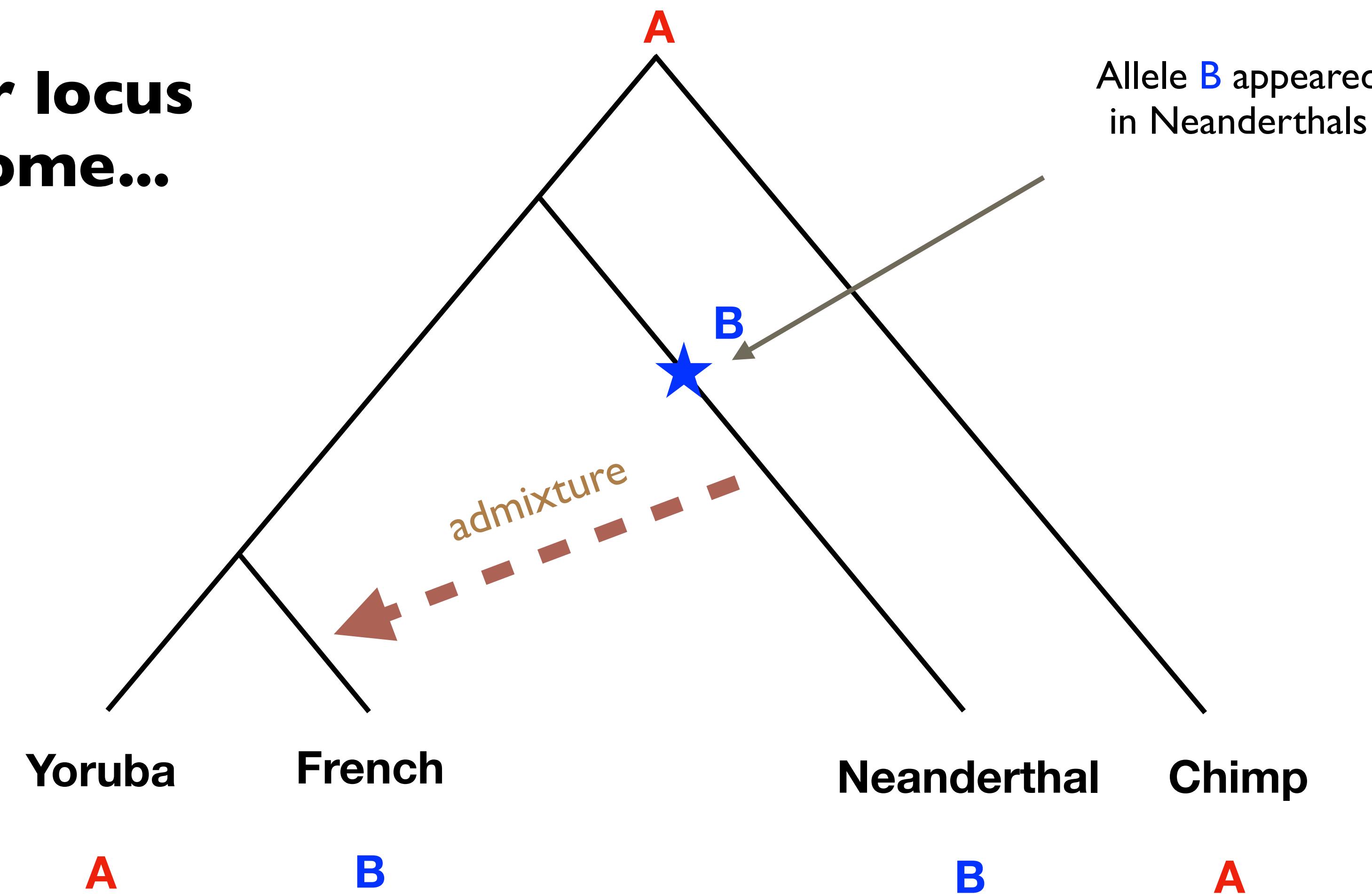


Allele B appeared  
in Neanderthals

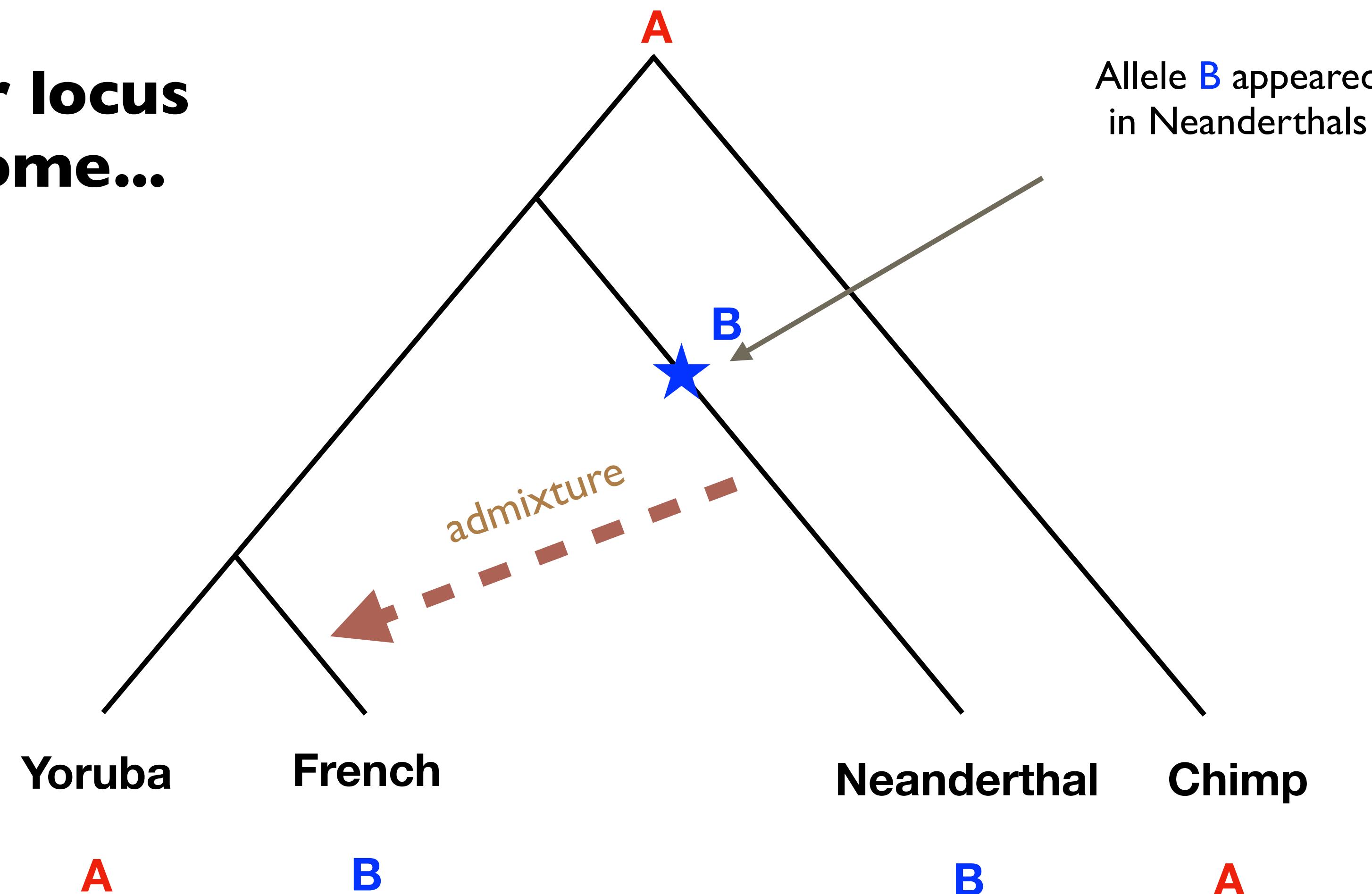
**At another locus  
in the genome...**



**At another locus  
in the genome...**

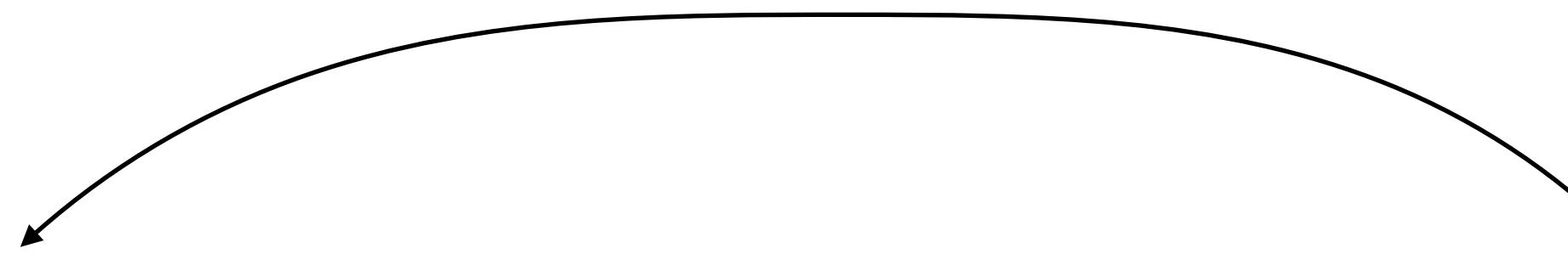


**At another locus  
in the genome...**



**Introgression: *f4 statistic* = (# BABA - # ABBA) / # SNPs < 0**

# Test of Neanderthal admixture in practice



$f_4(\text{Yoruba}, \mathbf{X}; \text{Neanderthal}, \text{Chimp}) \dots$  where  $\mathbf{X} = \text{Dinka or French}$

# Test of Neanderthal admixture in practice

$f_4(\text{Yoruba}, X; \text{Neanderthal}, \text{Chimp}) \dots$  where  $X = \text{Dinka or French}$

$$f_4 = (\# \text{ BABA} - \# \text{ ABBA}) / \# \text{ SNPs}$$

# Test of Neanderthal admixture in practice

$f_4(\text{Yoruba}, X; \text{Neanderthal}, \text{Chimp}) \dots$  where  $X = \text{Dinka or French}$

$$f_4 = (\# \text{BABA} - \# \text{ABBA}) / \# \text{SNPs}$$

X	#BABA	#ABBA	#total	f4
<b>French</b>	44409	46865	1436967	-0.0017
<b>Dinka</b>	43025	43182	1436978	-0.0001

# Test of Neanderthal admixture in practice

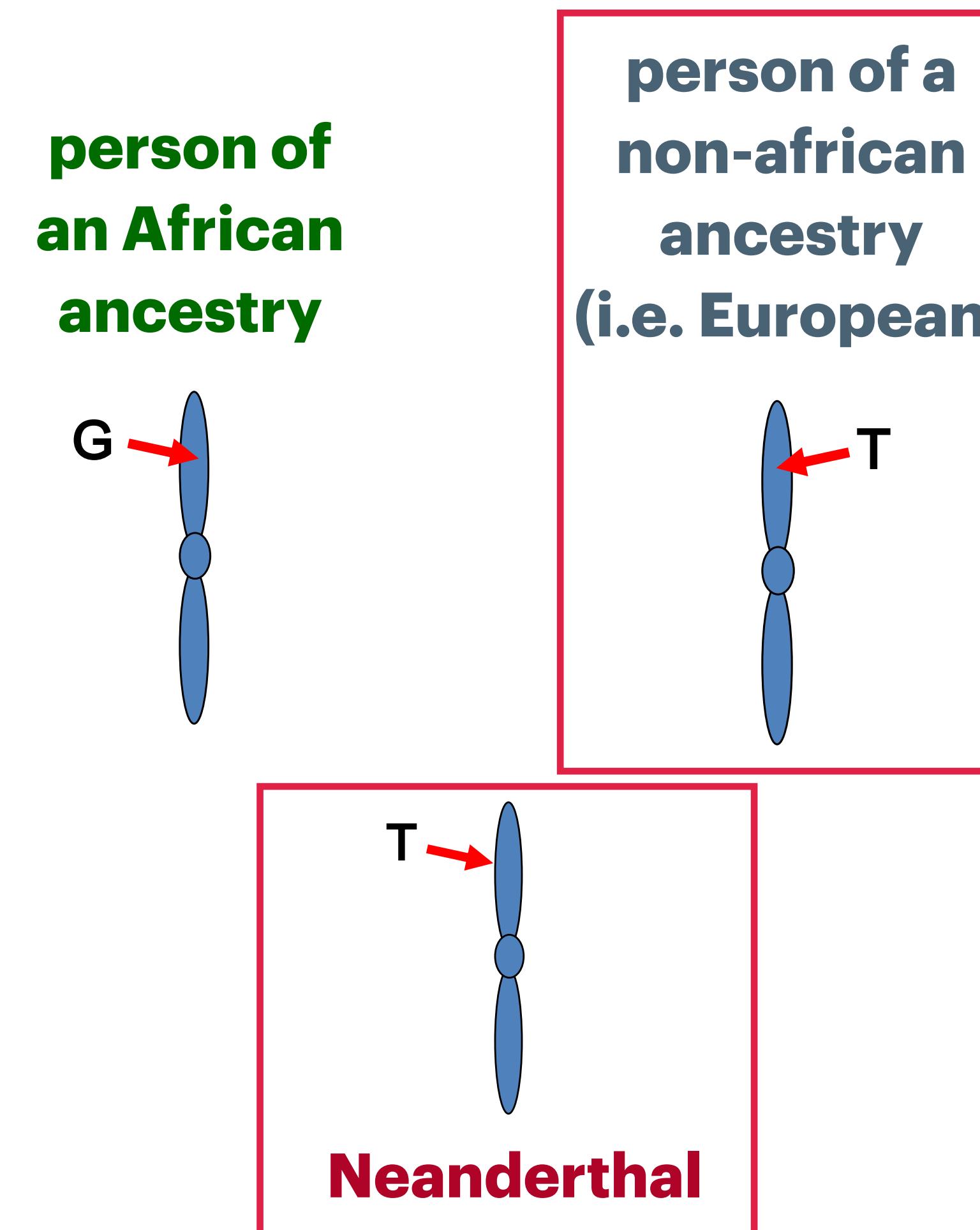
$f_4(\text{Yoruba}, X; \text{Neanderthal}, \text{Chimp}) \dots$  where  $X = \text{Dinka or French}$

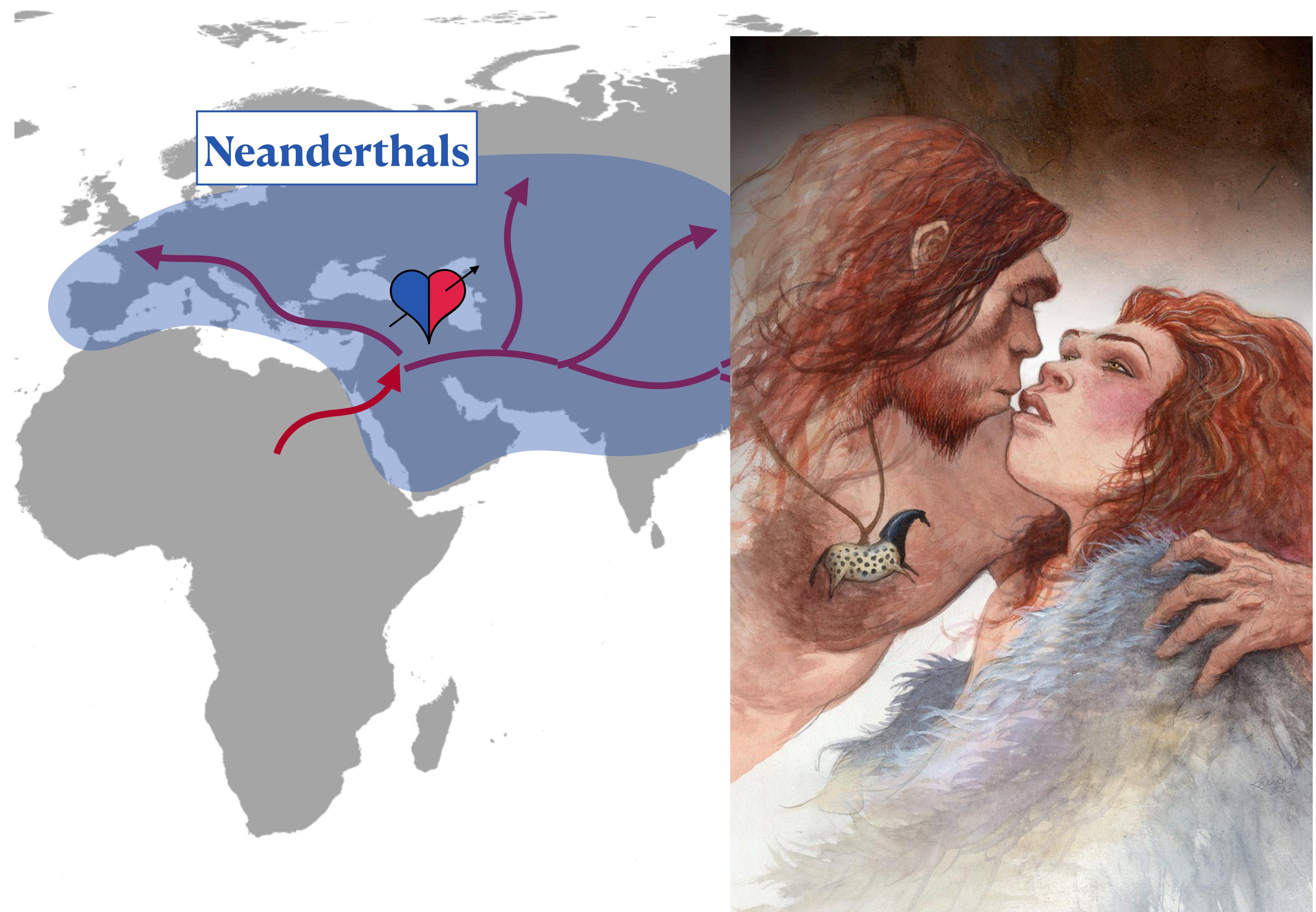
$$f_4 = (\# \text{BABA} - \# \text{ABBA}) / \# \text{SNPs}$$

X	#BABA	#ABBA	#total	f4	Z
<b>French</b>	44409	46865	1436967	-0.0017	-7.83
<b>Dinka</b>	43025	43182	1436978	-0.0001	-0.72

$|Z \text{ score}| > 3$  considered significant

# Test of Neanderthal admixture in practice





# What about the proportion of Neanderthal ancestry?

# What about the proportion of Neanderthal ancestry?

We compare the rate of allele sharing of sample X and a Neanderthal,  
with the rate of sharing between two Neanderthals.

# What about the proportion of Neanderthal ancestry?

We compare the rate of allele sharing of sample X and a Neanderthal, with the rate of sharing between two Neanderthals.

$f4(\text{Yoruba}, \mathbf{X}; \text{Altai}, \text{Chimp})$



X	#BABA	#ABBA	#sites	f4
<b>French</b>	44409	46865	1436967	-0.0017
<b>Dinka</b>	43025	43182	1436978	-0.0001

# What about the proportion of Neanderthal ancestry?

We compare the rate of allele sharing of sample X and a Neanderthal, with the rate of sharing between two Neanderthals.

$f4(\text{Yoruba}, \mathbf{X}; \text{Altai}, \text{Chimp})$



X	#BABA	#ABBA	#sites	f4
<b>French</b>	44409	46865	1436967	-0.0017
<b>Dinka</b>	43025	43182	1436978	-0.0001
<b>Vindija</b>	7337	118956	1436703	<b>-0.0777</b>

# What about the proportion of Neanderthal ancestry?

We compare the rate of allele sharing of sample X and a Neanderthal, with the rate of sharing between two Neanderthals.

$f4(\text{Yoruba}, \mathbf{X}; \text{Altai}, \text{Chimp}) / f4(\text{Yoruba}, \mathbf{Vindija}; \text{Altai}, \text{Chimp})$



X	#BABA	#ABBA	#sites	f4
<b>French</b>	44409	46865	1436967	-0.0017
<b>Dinka</b>	43025	43182	1436978	-0.0001
<b>Vindija</b>	7337	118956	1436703	<b>-0.0777</b>

# What about the proportion of Neanderthal ancestry?

We compare the rate of allele sharing of sample X and a Neanderthal, with the rate of sharing between two Neanderthals.

$f4(\text{Yoruba}, \mathbf{X}; \text{Altai}, \text{Chimp}) / f4(\text{Yoruba}, \mathbf{Vindija}; \text{Altai}, \text{Chimp})$

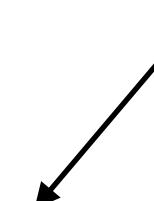


X	#BABA	#ABBA	#sites	f4
<b>French</b>	44409	46865	1436967	-0.0017
<b>Dinka</b>	43025	43182	1436978	-0.0001
<b>Vindija</b>	7337	118956	1436703	<b>-0.0777</b>

# What about the proportion of Neanderthal ancestry?

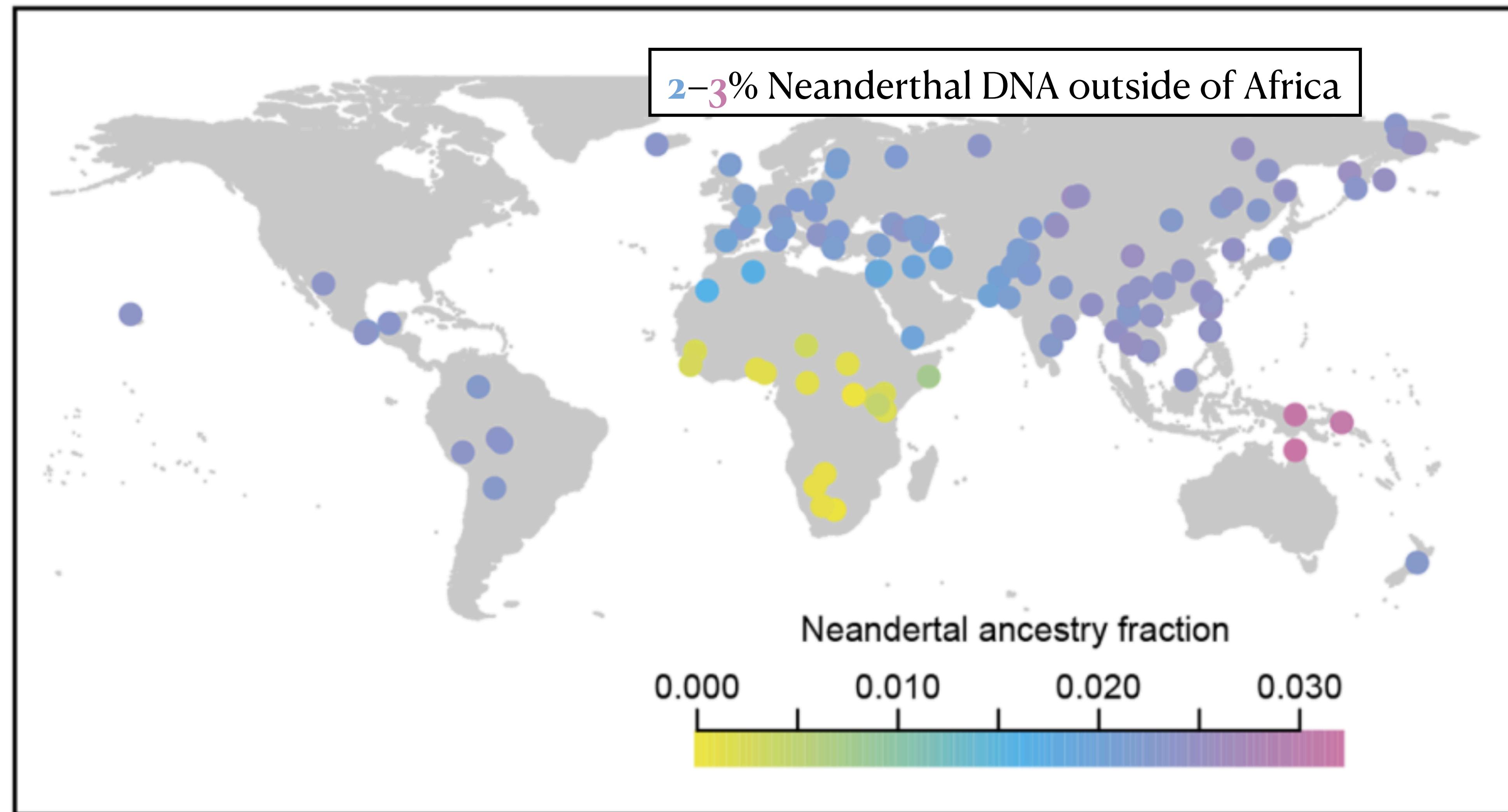
We compare the rate of allele sharing of sample X and a Neanderthal, with the rate of sharing between two Neanderthals.

$f4(\text{Yoruba}, \mathbf{X}; \text{Altai}, \text{Chimp}) / f4(\text{Yoruba}, \mathbf{Vindija}; \text{Altai}, \text{Chimp})$

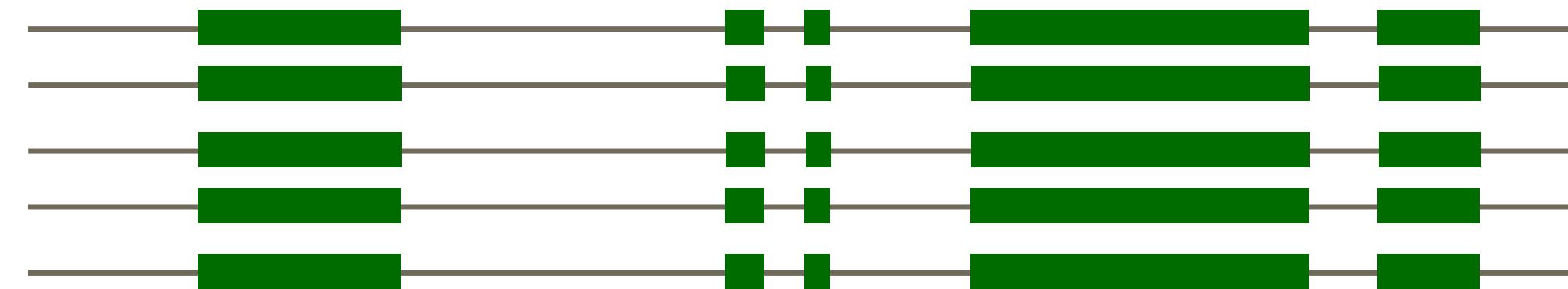


X	#BABA	#ABBA	#sites	f4	ratio
French	44409	46865	1436967	-0.0017	2.2%
Dinka	43025	43182	1436978	-0.0001	0.1%
Vindija	7337	118956	1436703	-0.0777	100%

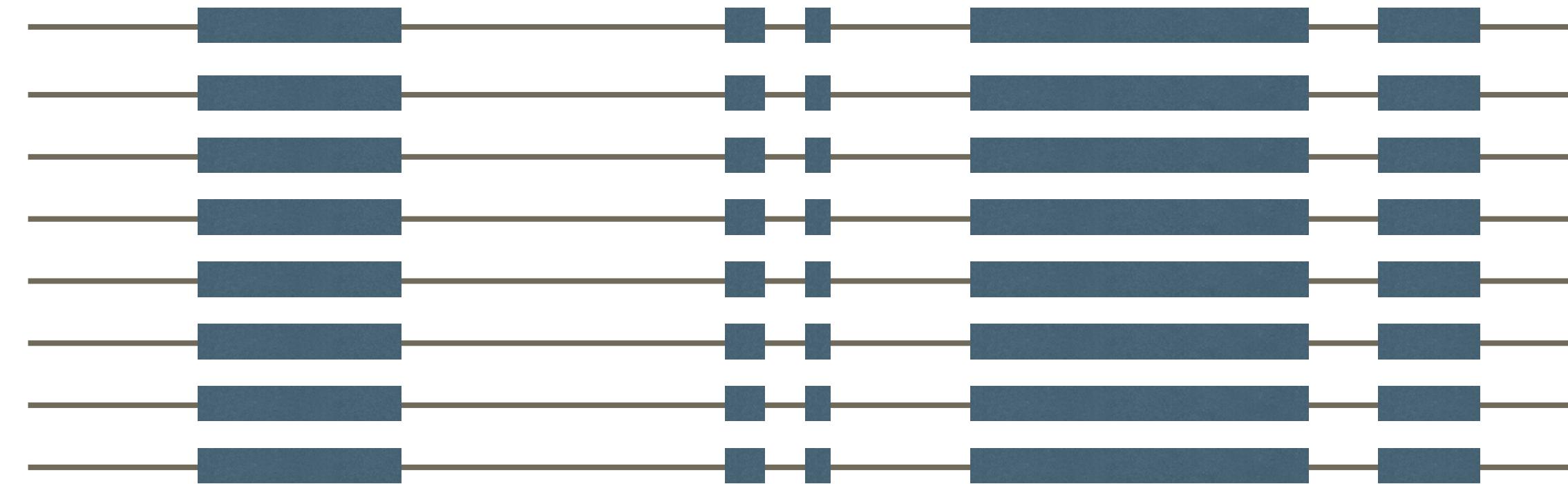
# Proportion of Neanderthal DNA in humans today



# We can go from detecting SNPs of Neanderthal origin...



genomes of a  
reference population  
(Yorùbá)



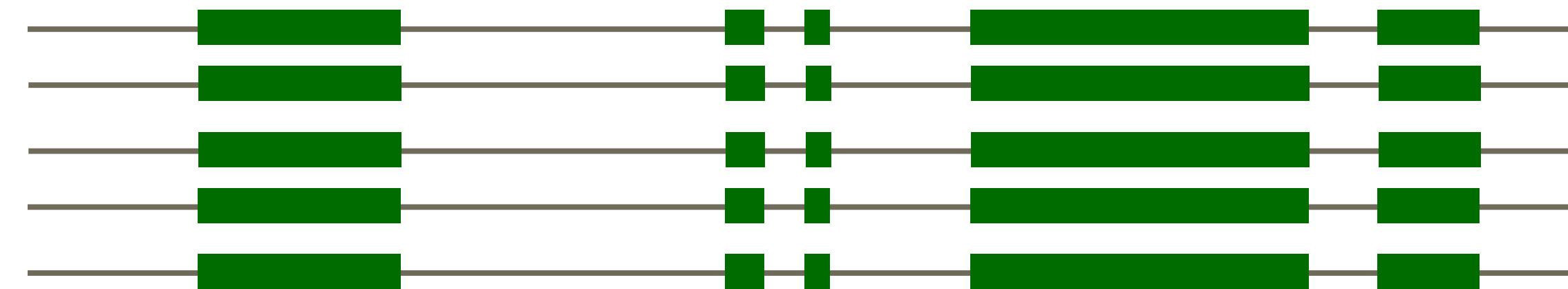
genomes of a  
target population  
(Europeans)



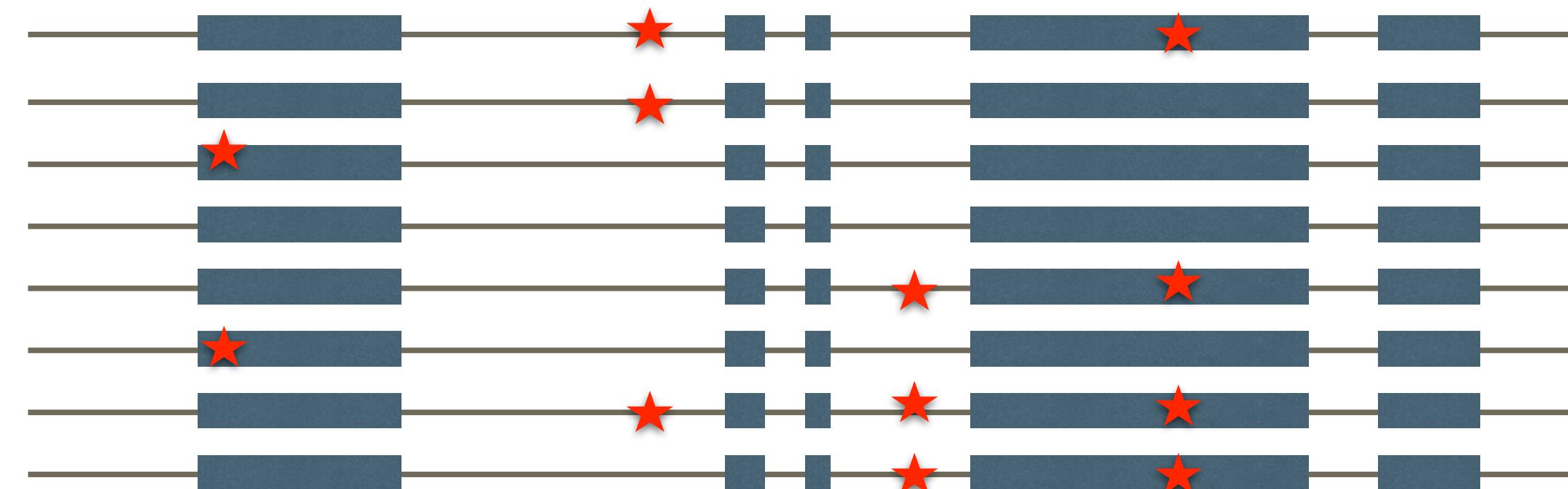
Neanderthal  
genome

★ = Neanderthal-derived mutation

# We can go from detecting SNPs of Neanderthal origin...



genomes of a  
reference population  
(Yorùbá)



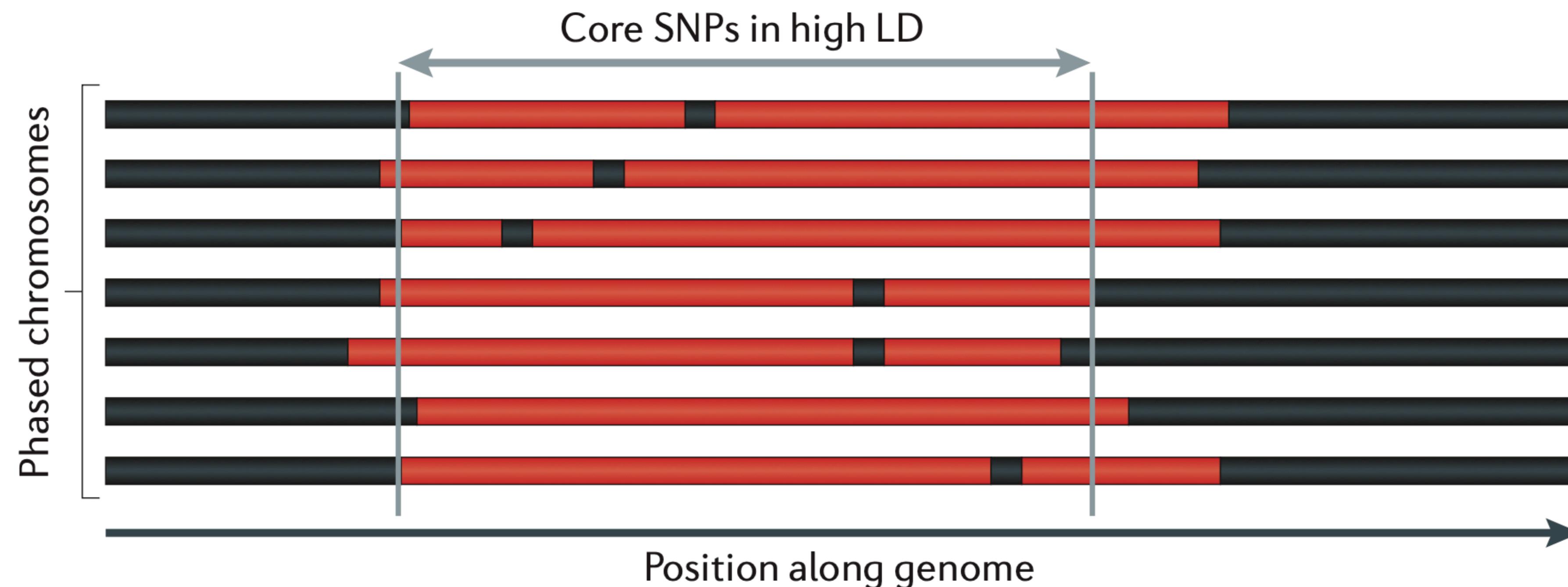
genomes of a  
target population  
(Europeans)



Neanderthal  
genome

★ = Neanderthal-derived mutation

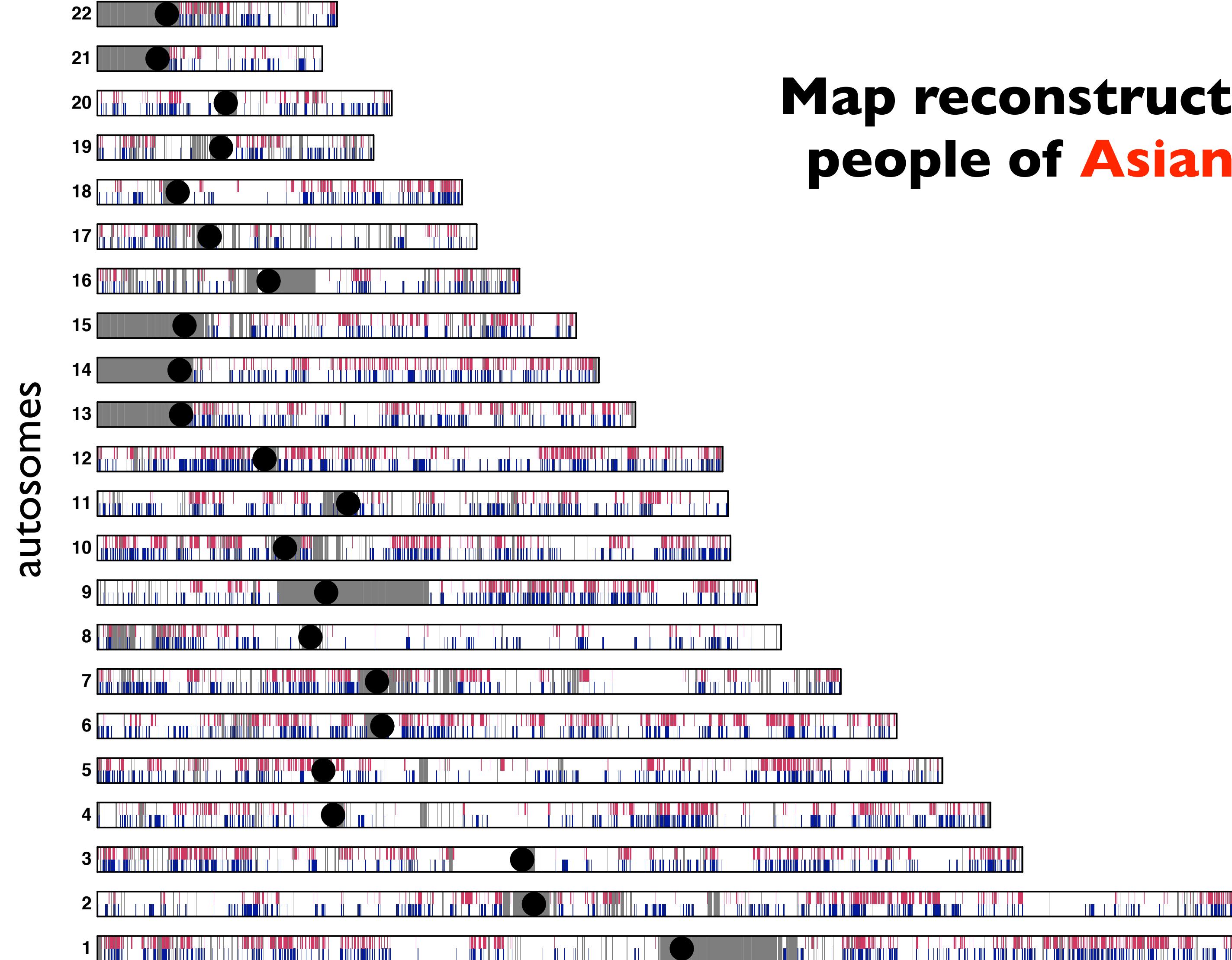
# ...to continuous segments of Neanderthal DNA



★ = Neanderthal-derived mutation

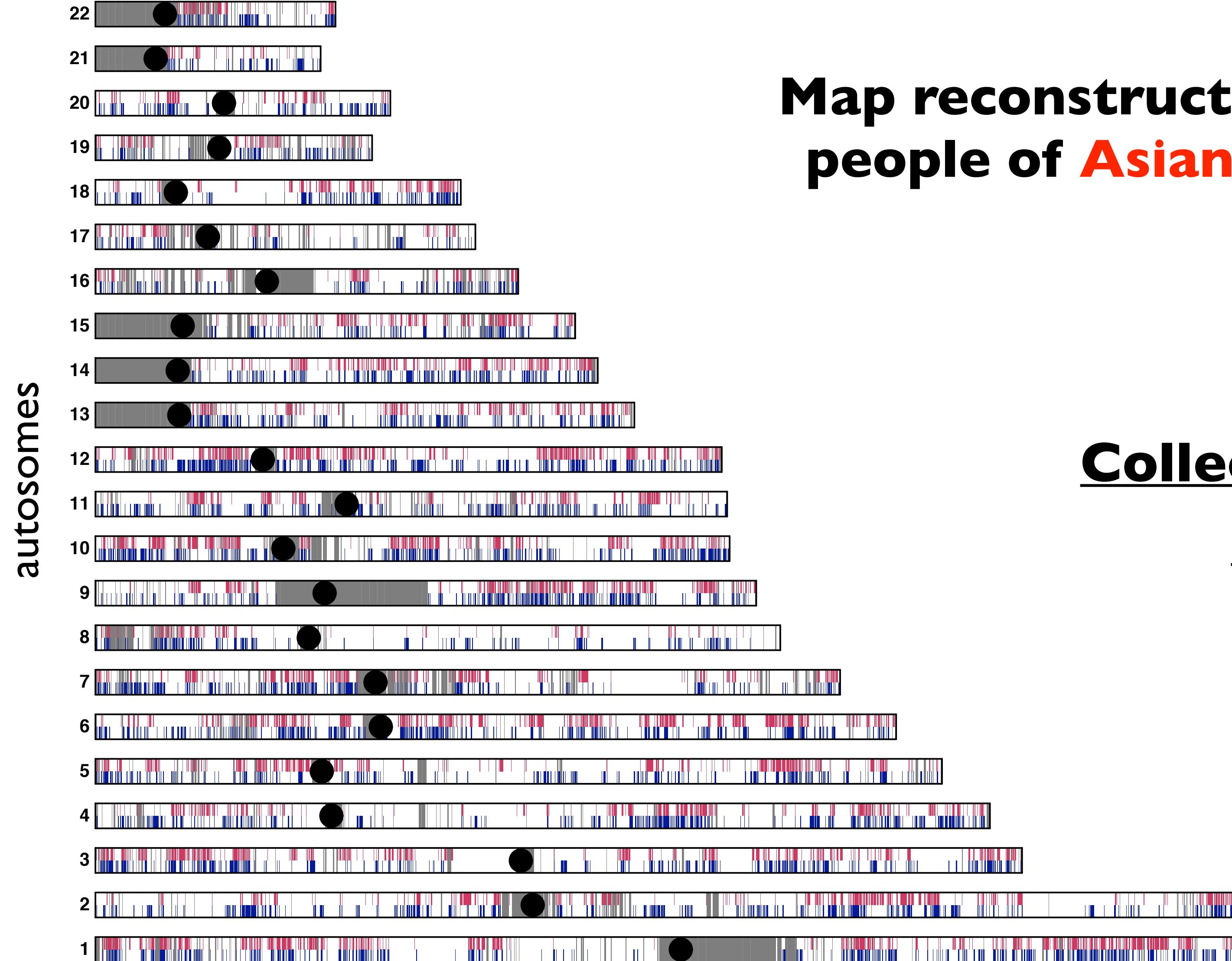
Racimo et al. (Nature Reviews 2015)

# Genomic maps of Neanderthal DNA today



**Map reconstructed from genomes of 700 people of Asian and European ancestry.**

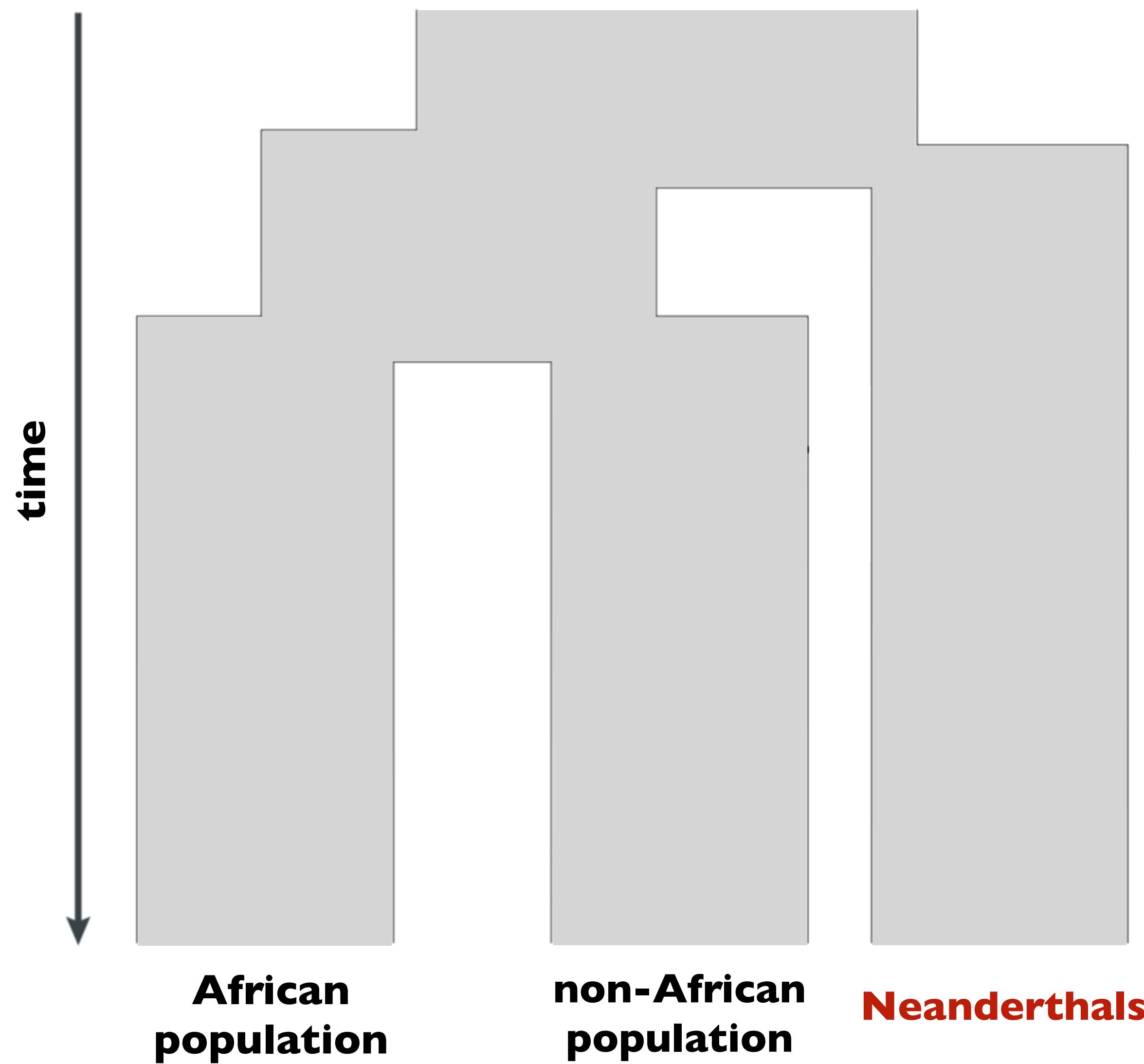
# Genomic maps of Neanderthal DNA today



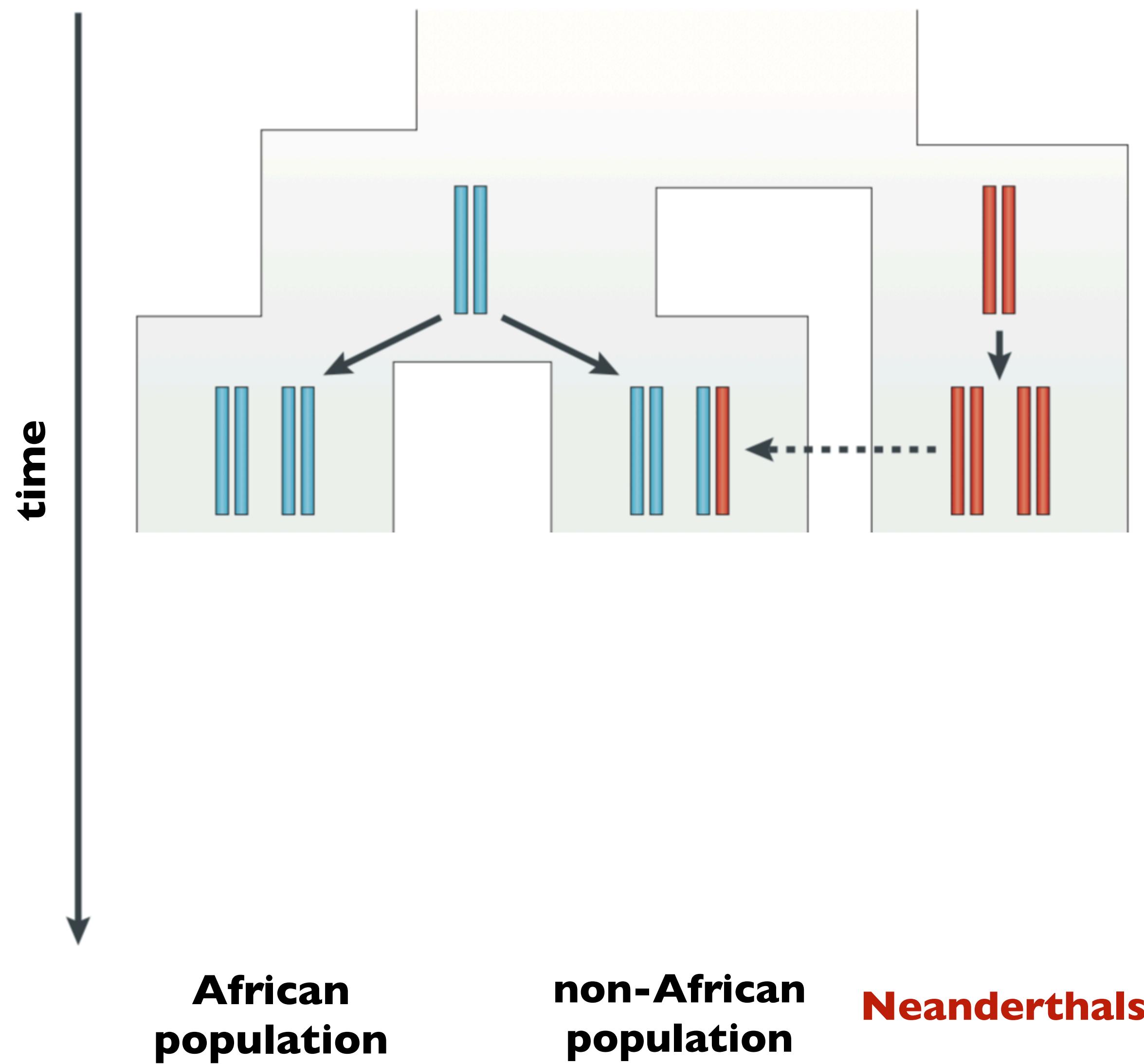
**Map reconstructed from genomes of 700 people of **Asian** and **European** ancestry.**

**Collectively we carry around**  
**40% Neanderthal DNA!**

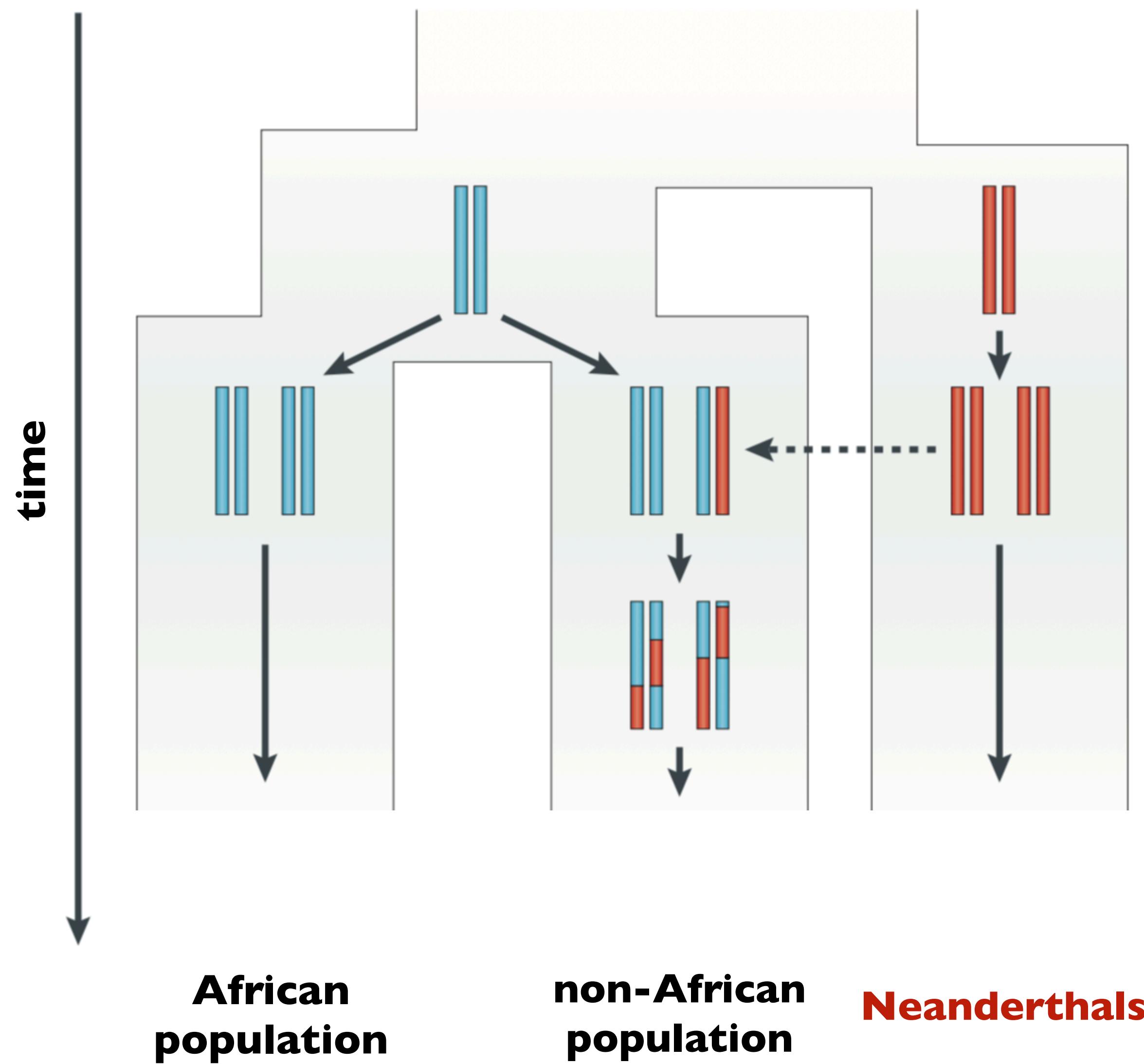
# When did introgression happen?



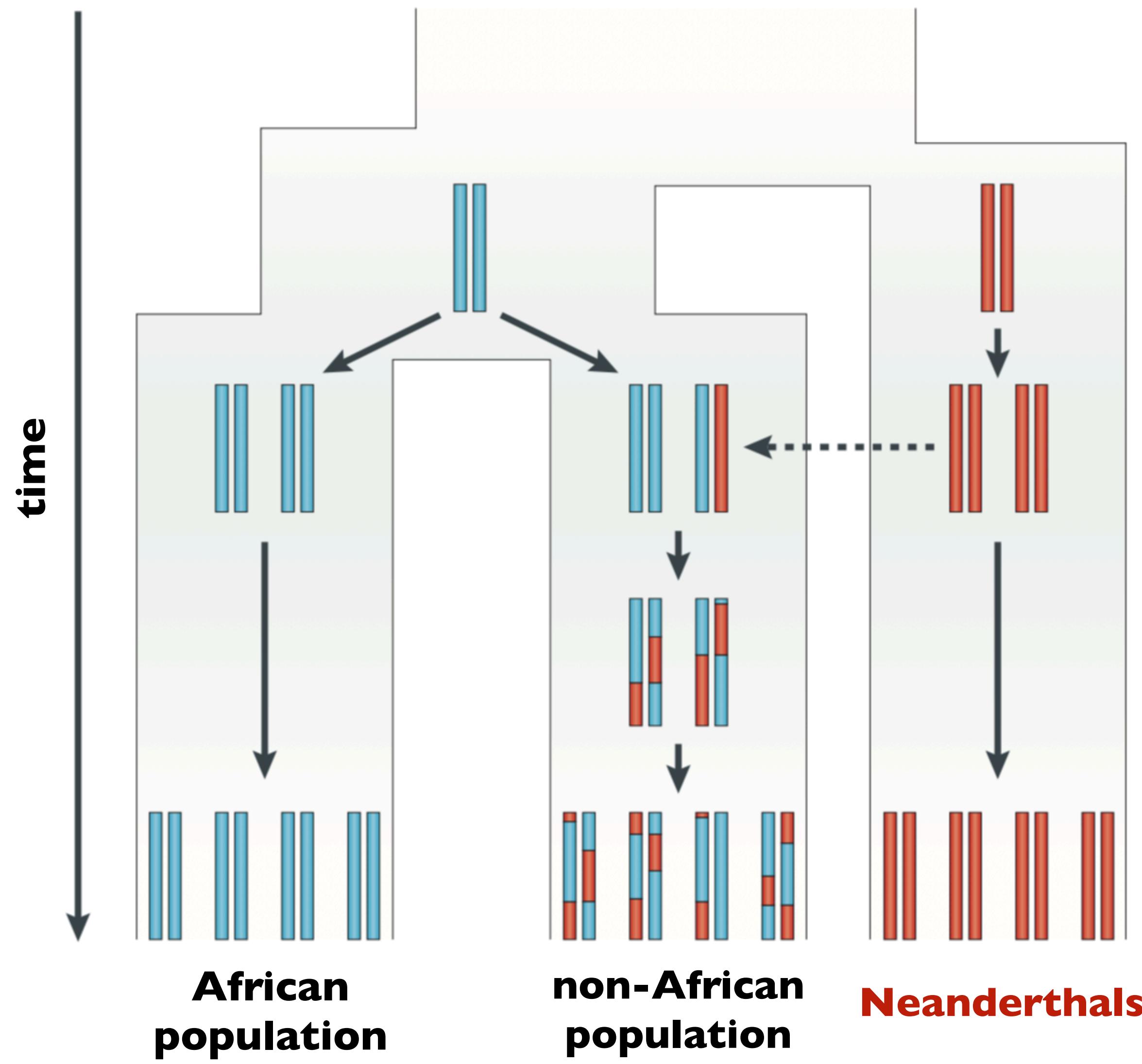
# When did introgression happen?



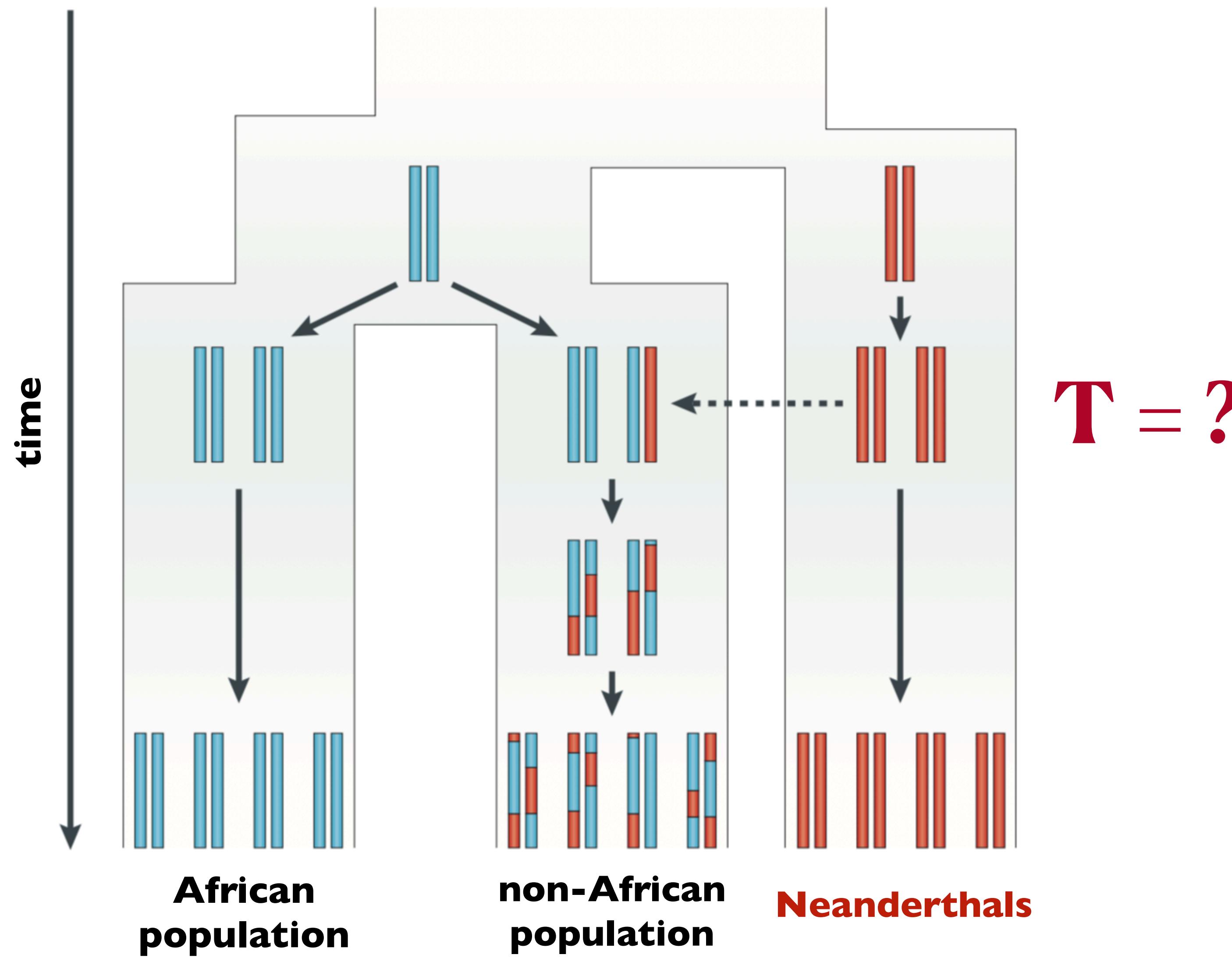
# When did introgression happen?



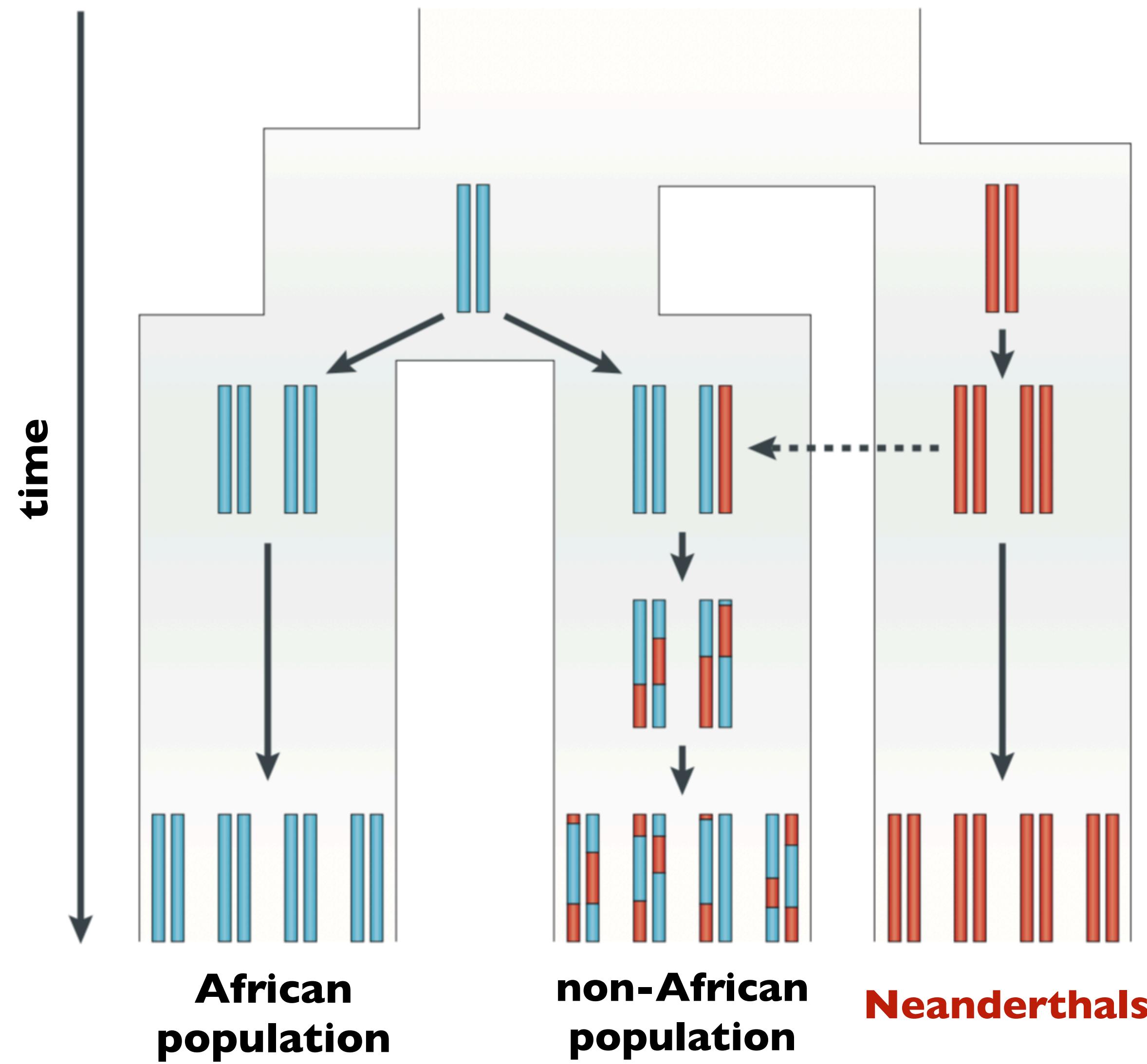
# When did introgression happen?



# When did introgression happen?



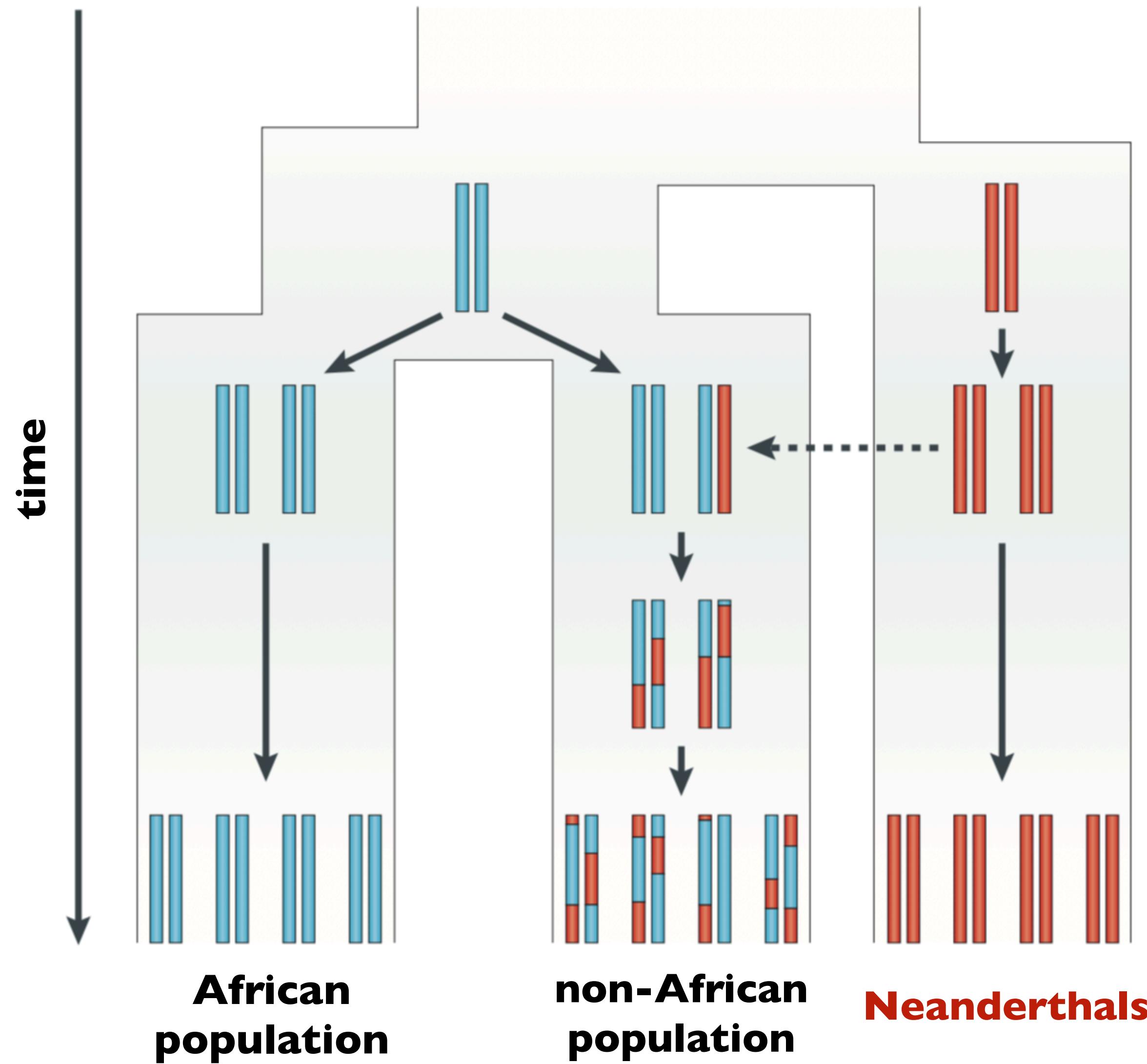
# When did introgression happen?



Haplotypes carrying **Neanderthal alleles** get shorter over time due to recombination (statistically “regularly”).

**T = ?**

# When did introgression happen?

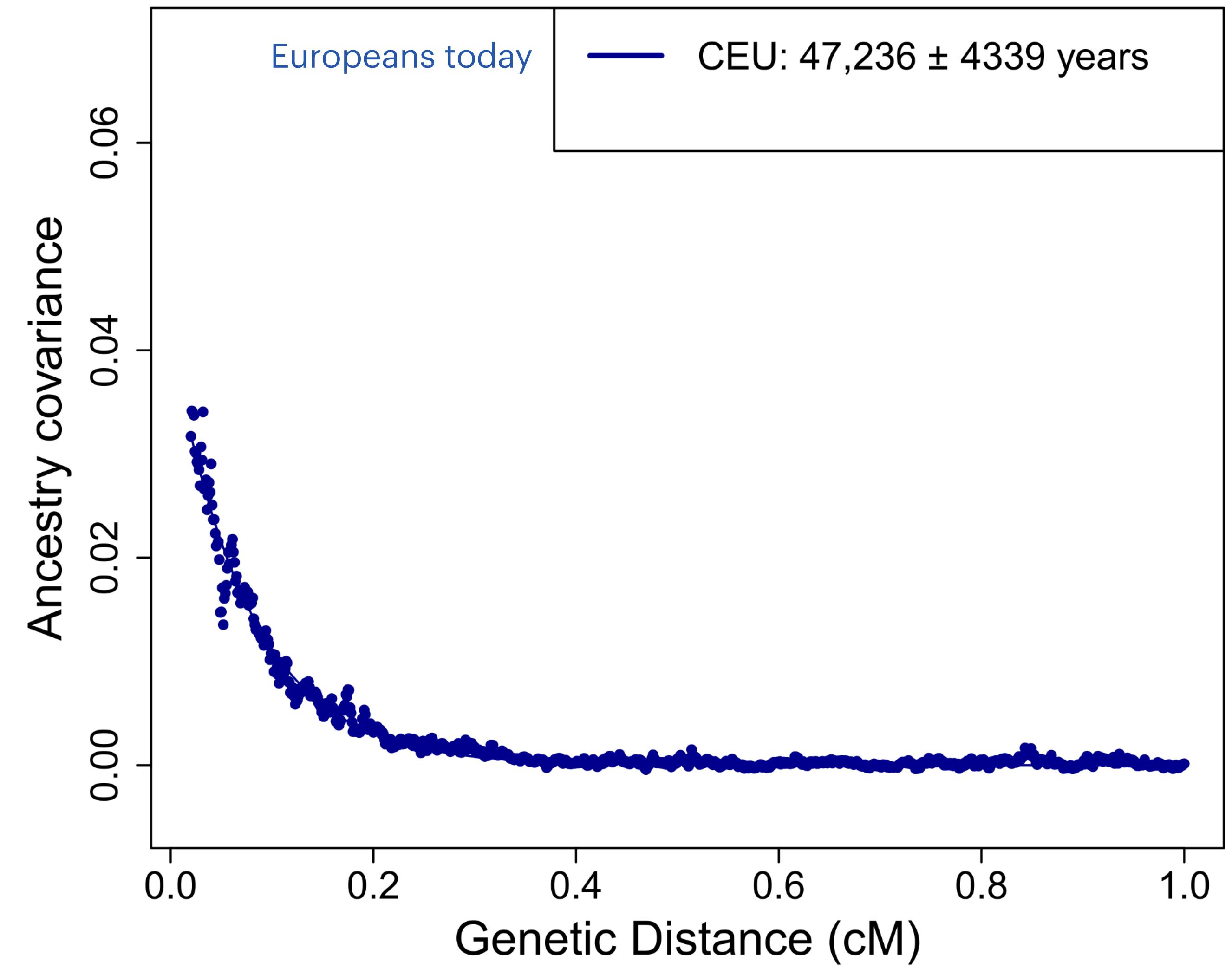
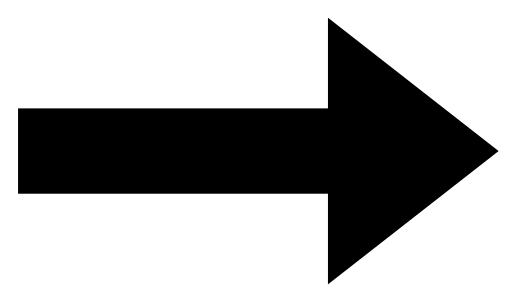


Haplotypes carrying **Neanderthal alleles** get shorter over time due to recombination (statistically “regularly”).

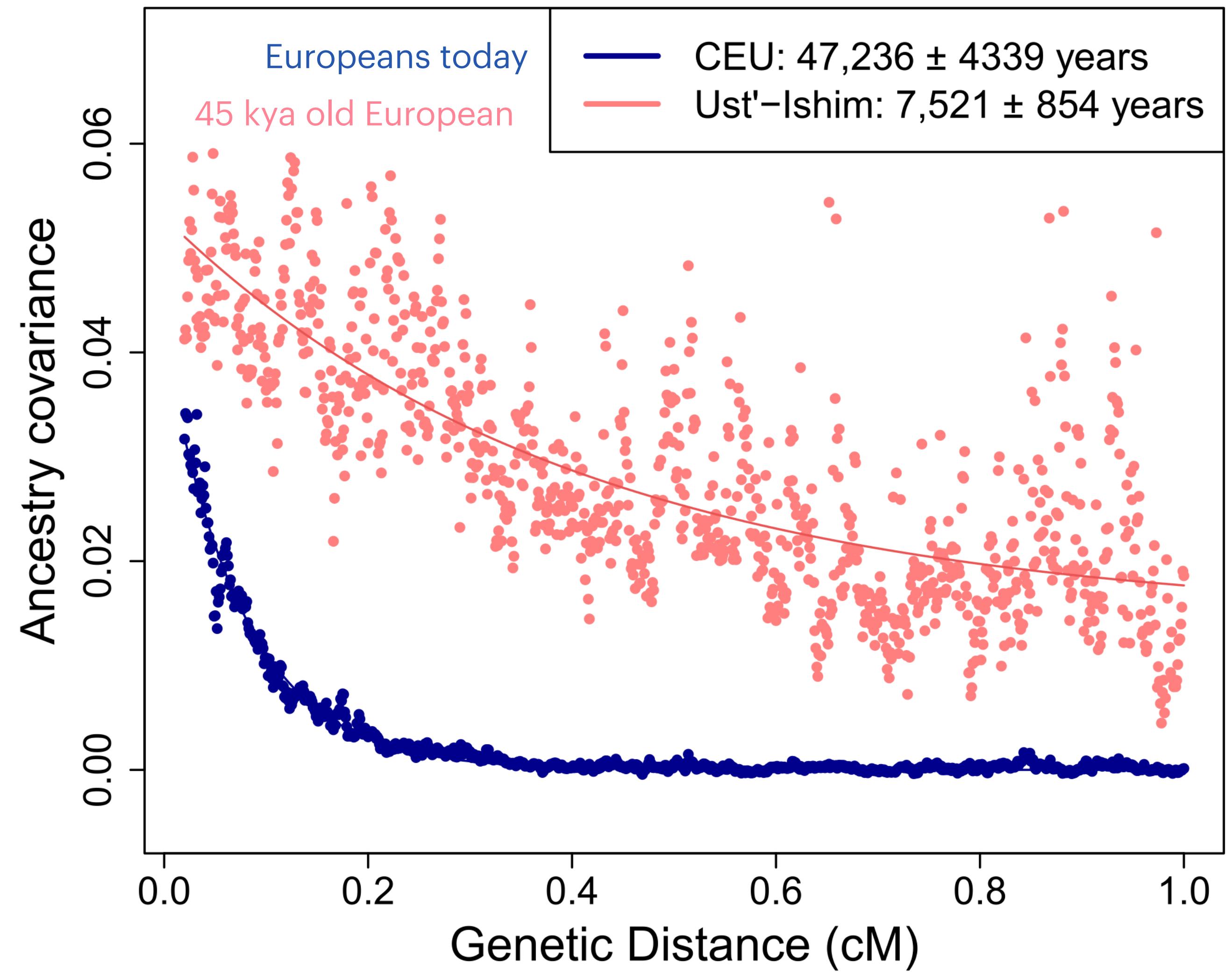
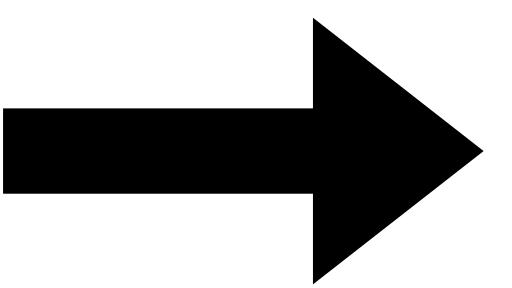
$T = ?$

Expected length of these haplotypes follows the principle of exponential decay.

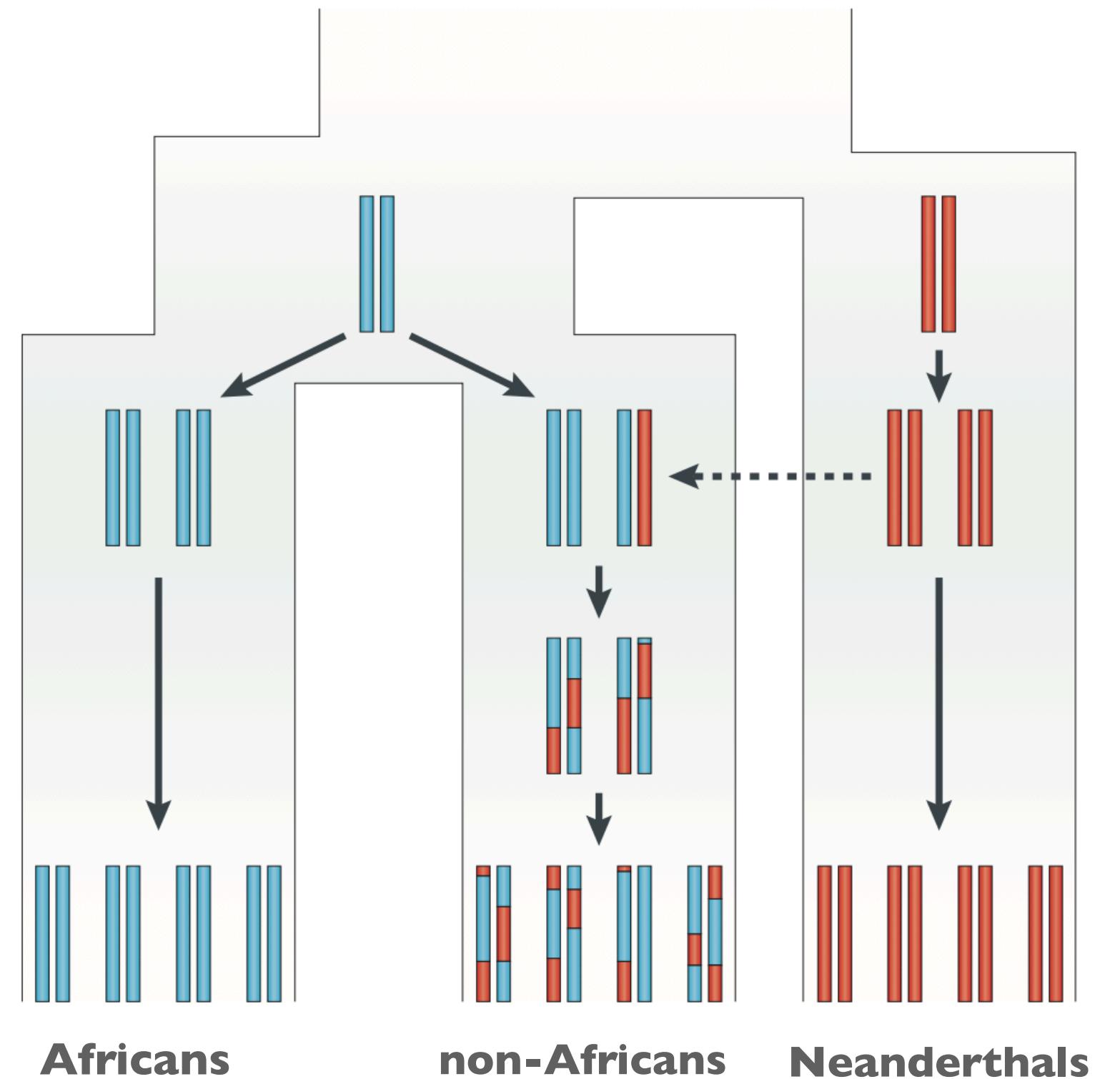
# Introgression happened around 50 thousand years ago



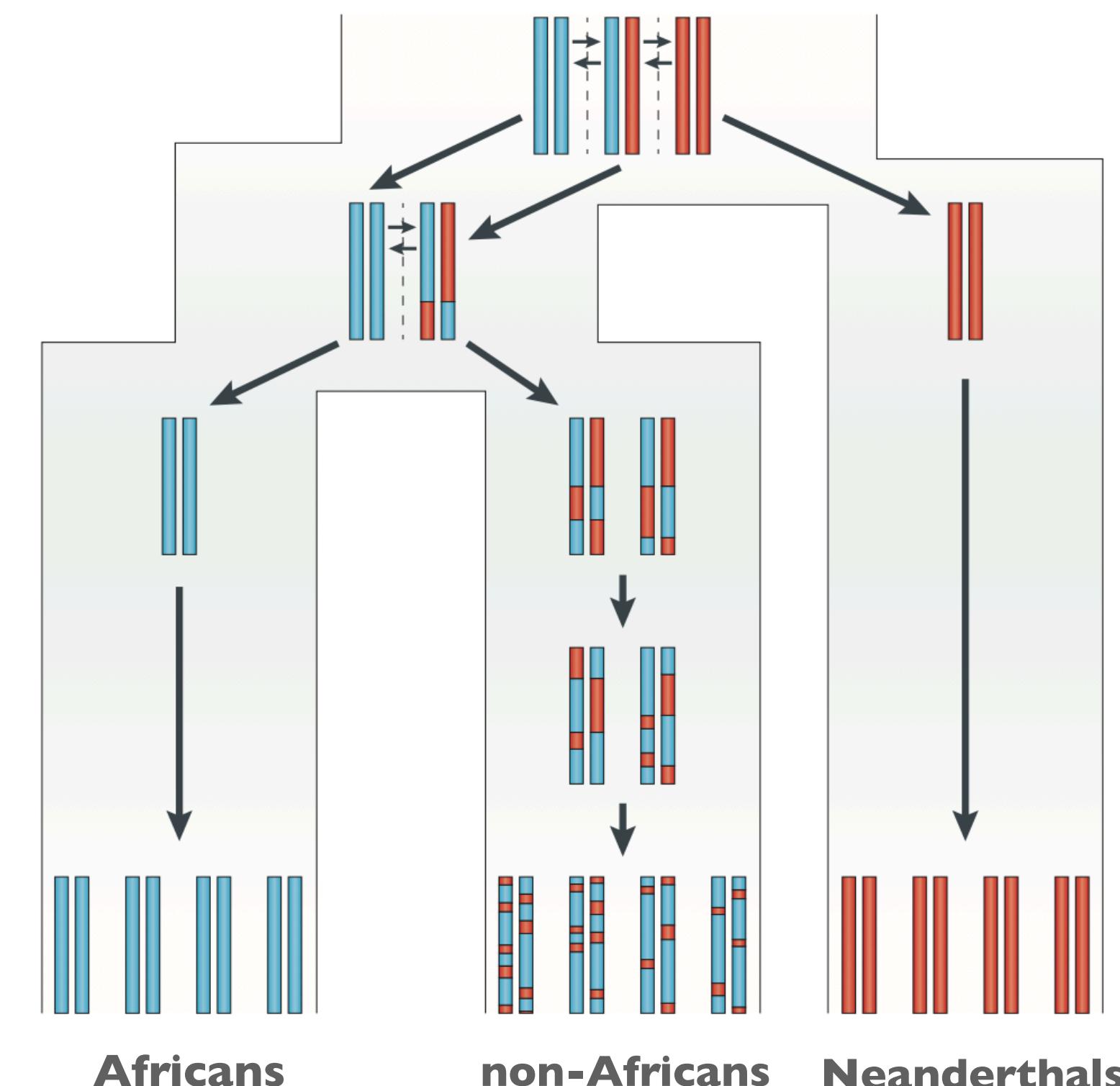
# Introgression happened around 50 thousand years ago



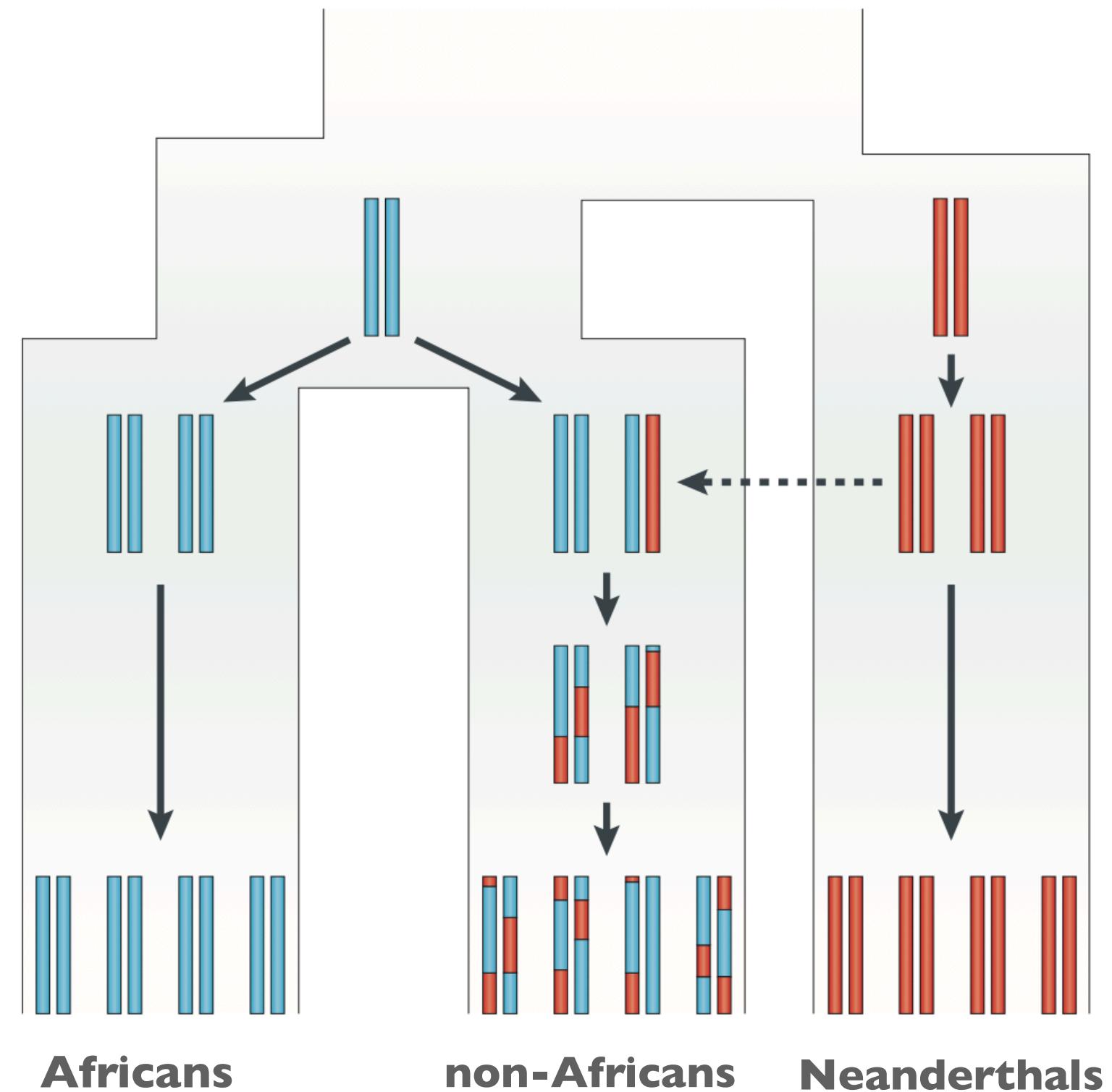
# Introgression?



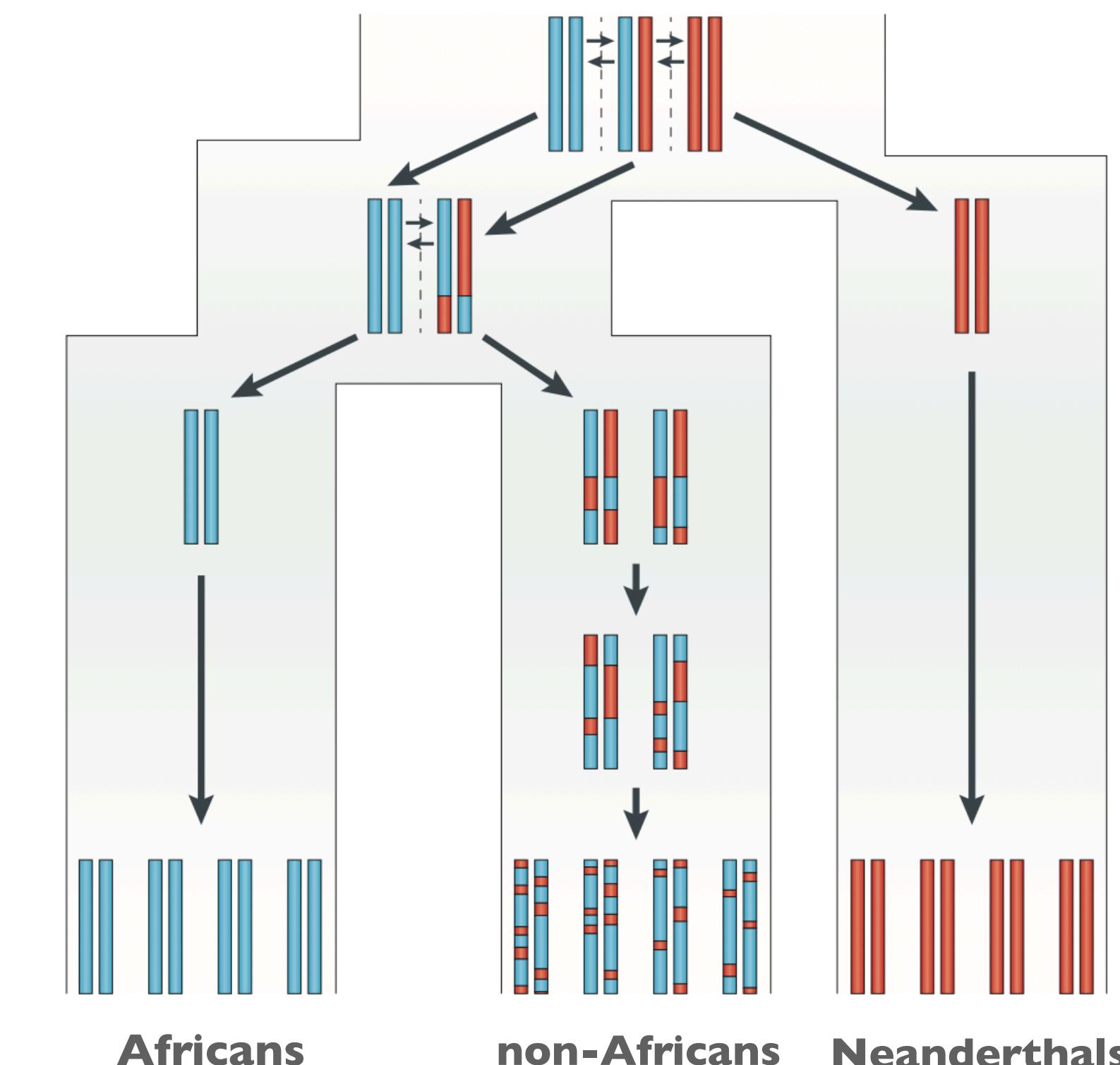
# Structure?



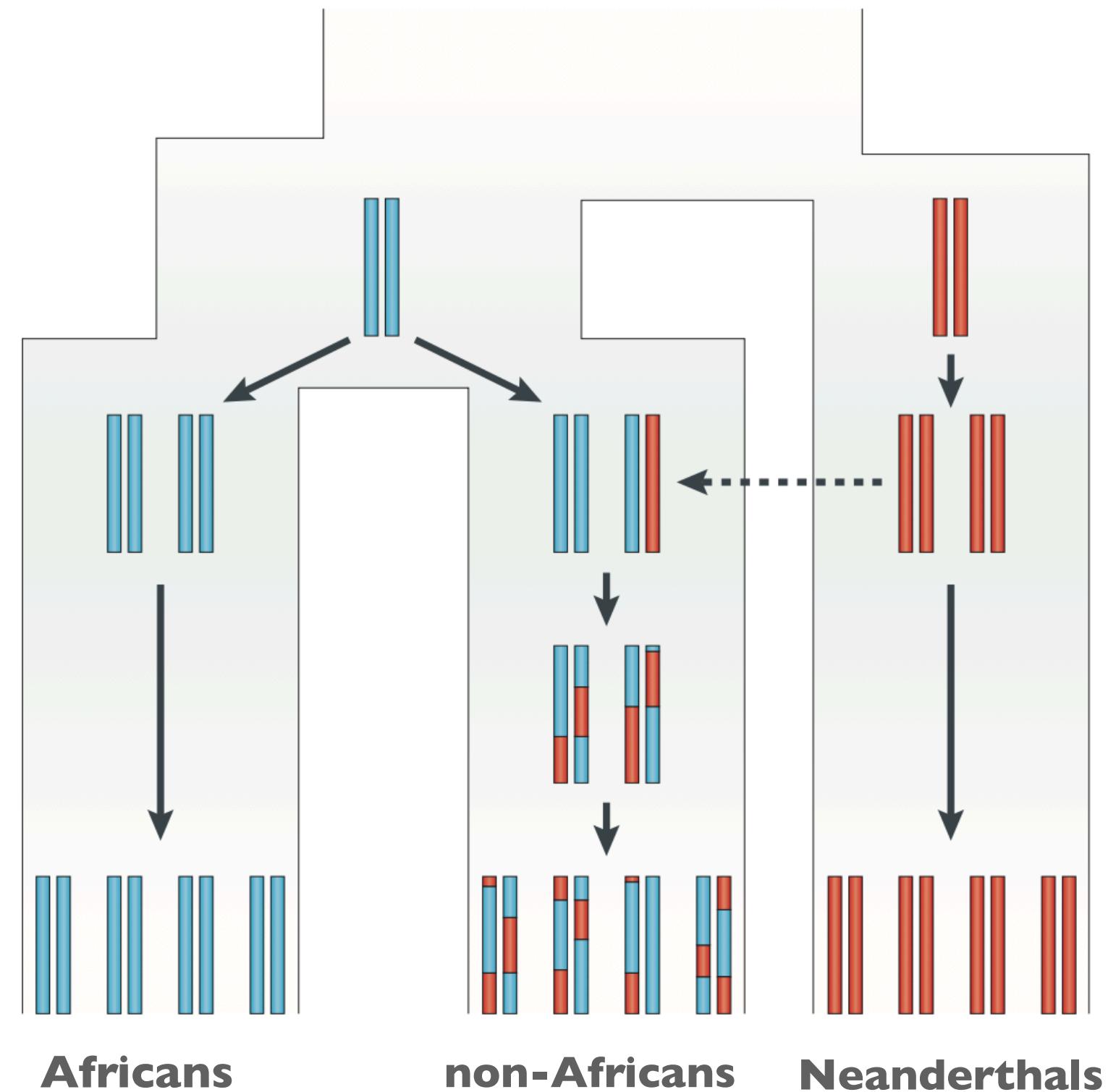
# Introgression?



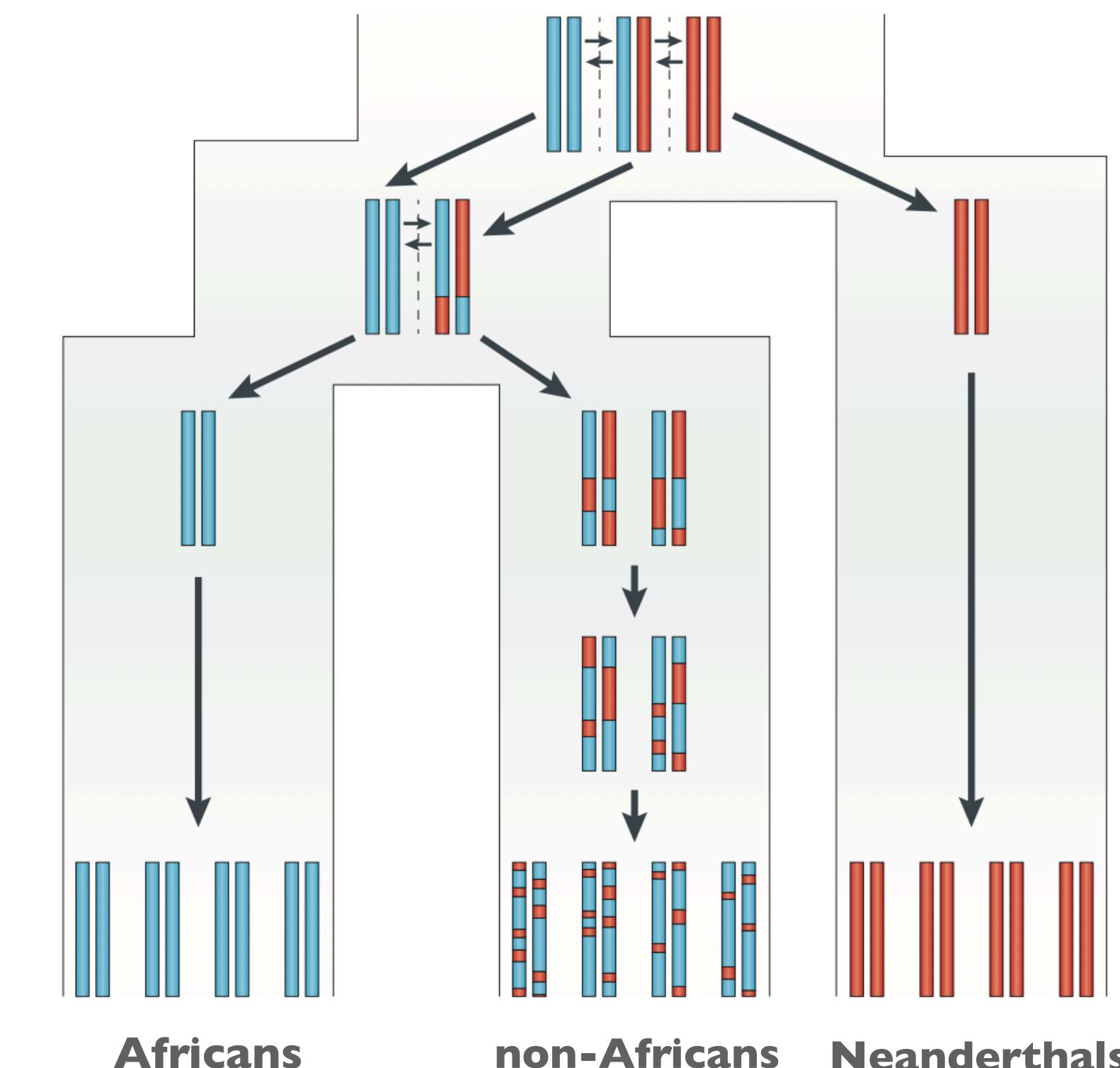
# Structure?



# Introgression?

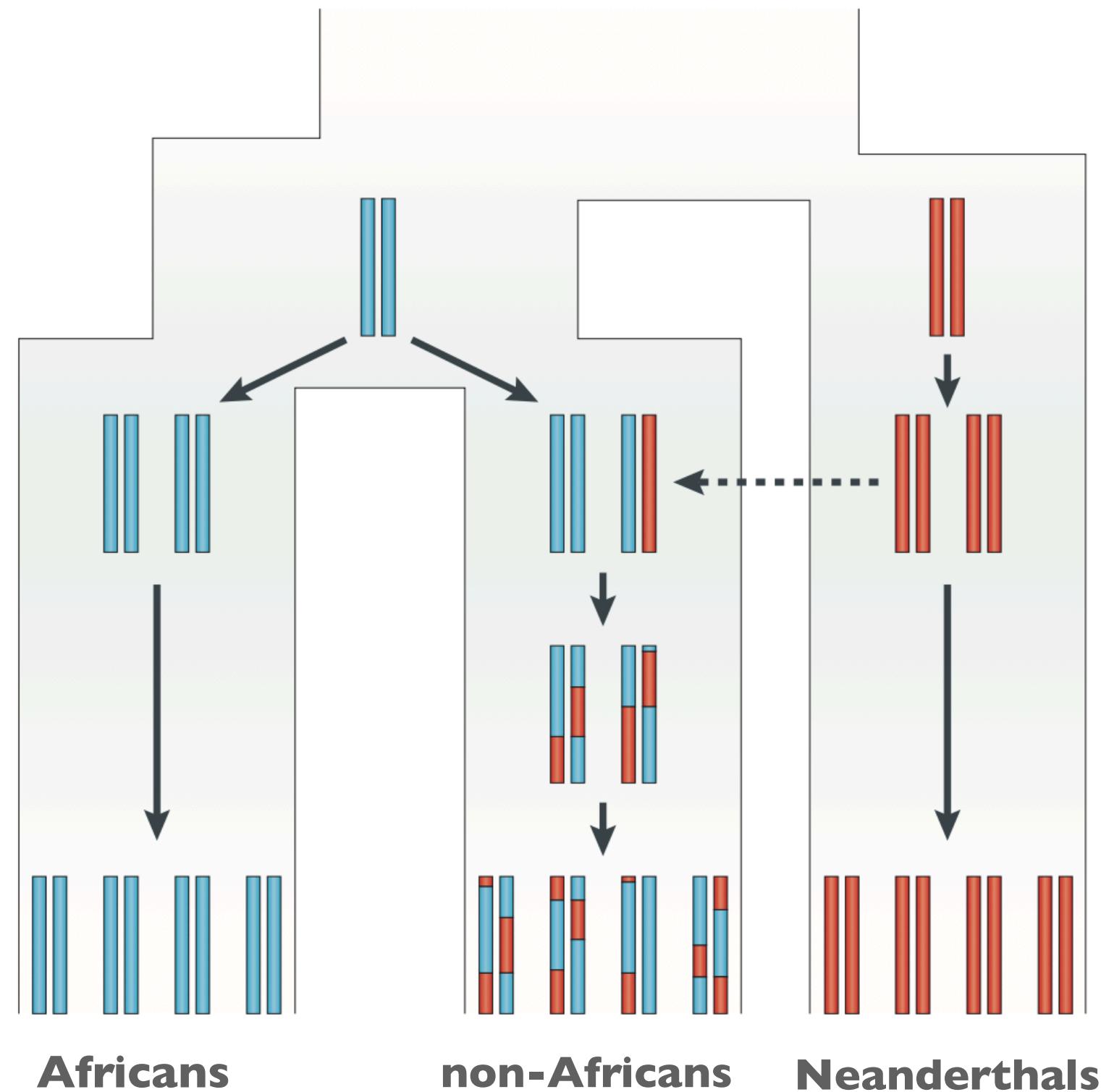


# Structure?

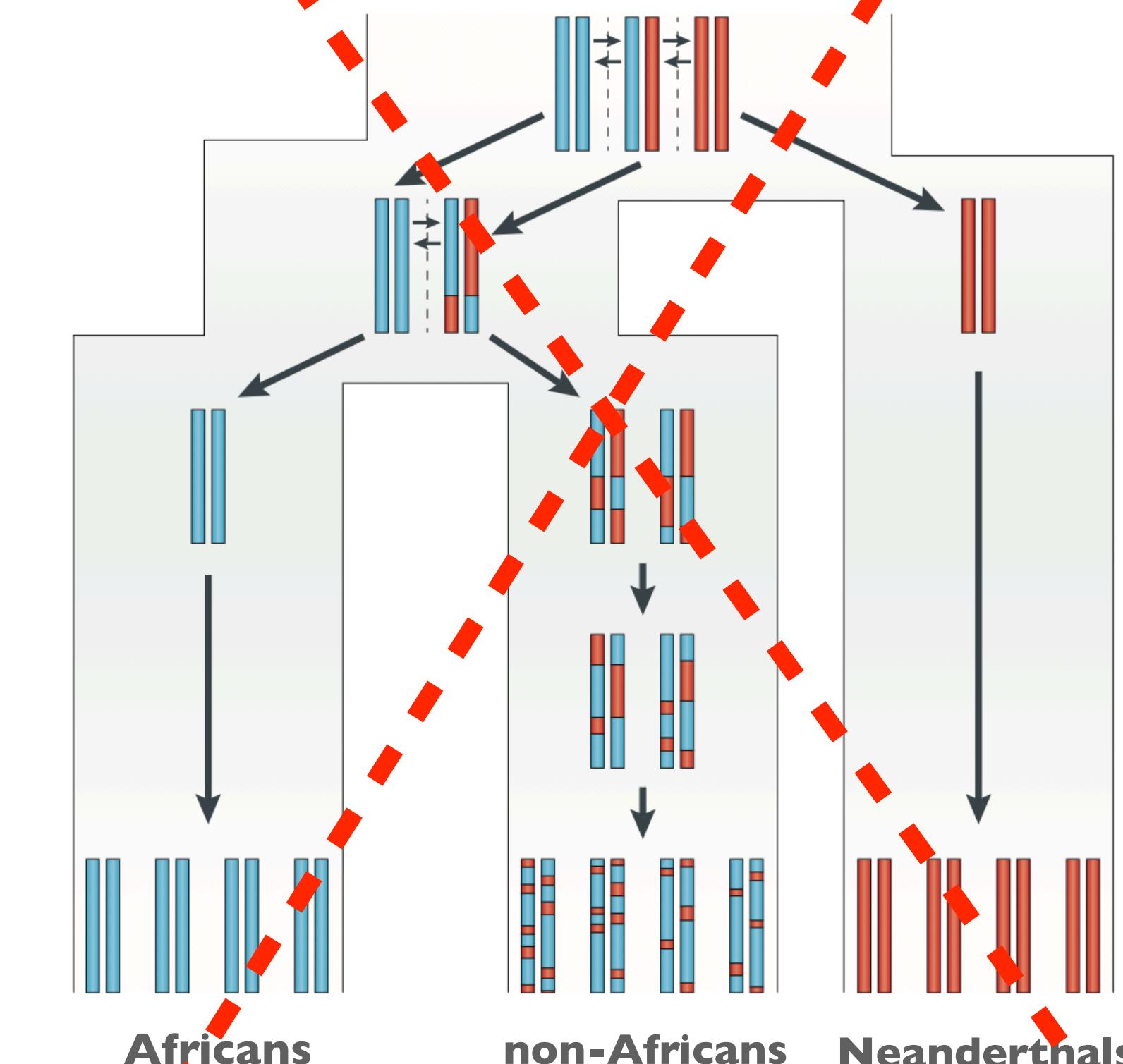


**Under structure, “Neanderthal-looking”  
haplotypes will be much shorter!**

# Introgression?



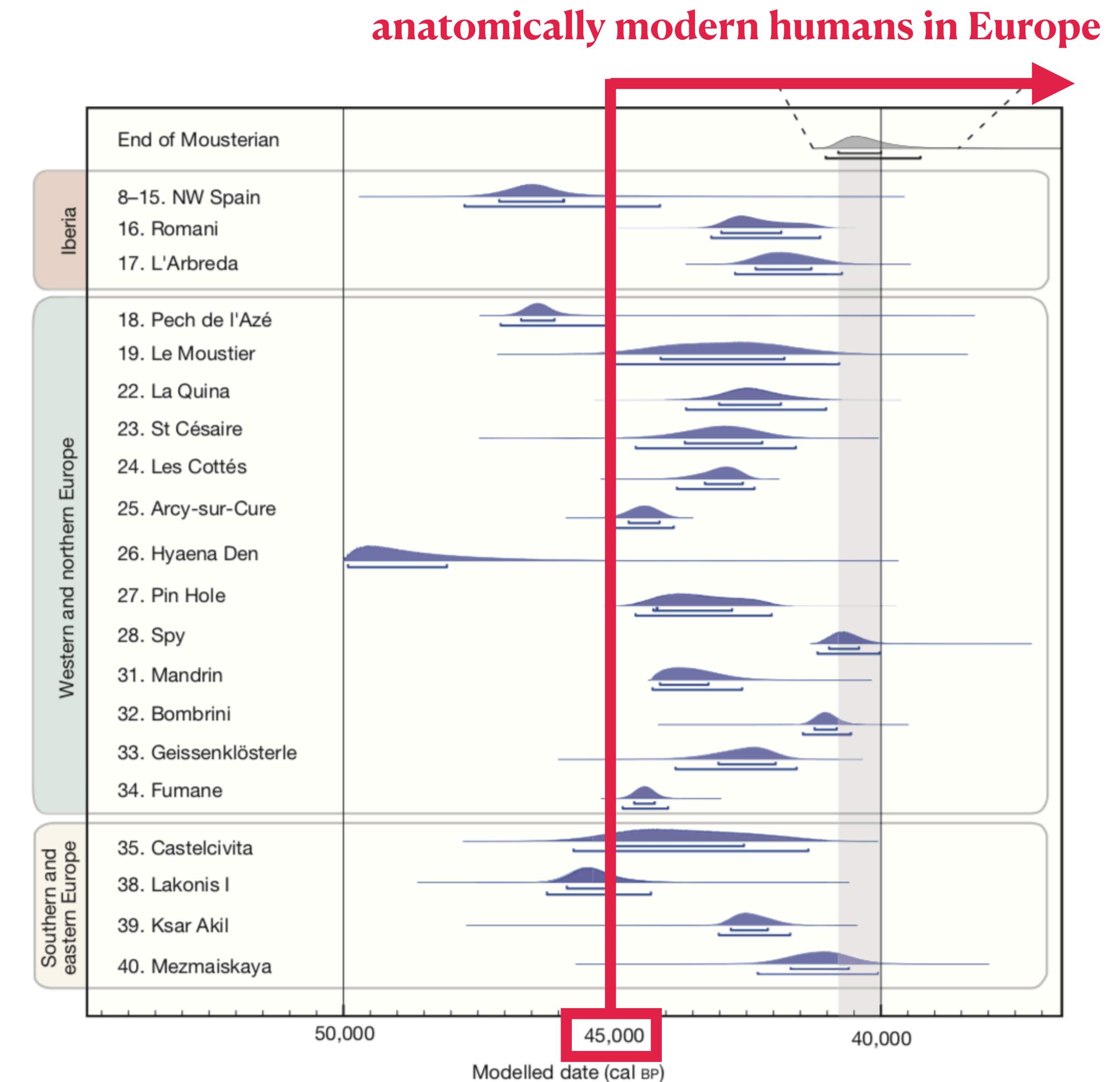
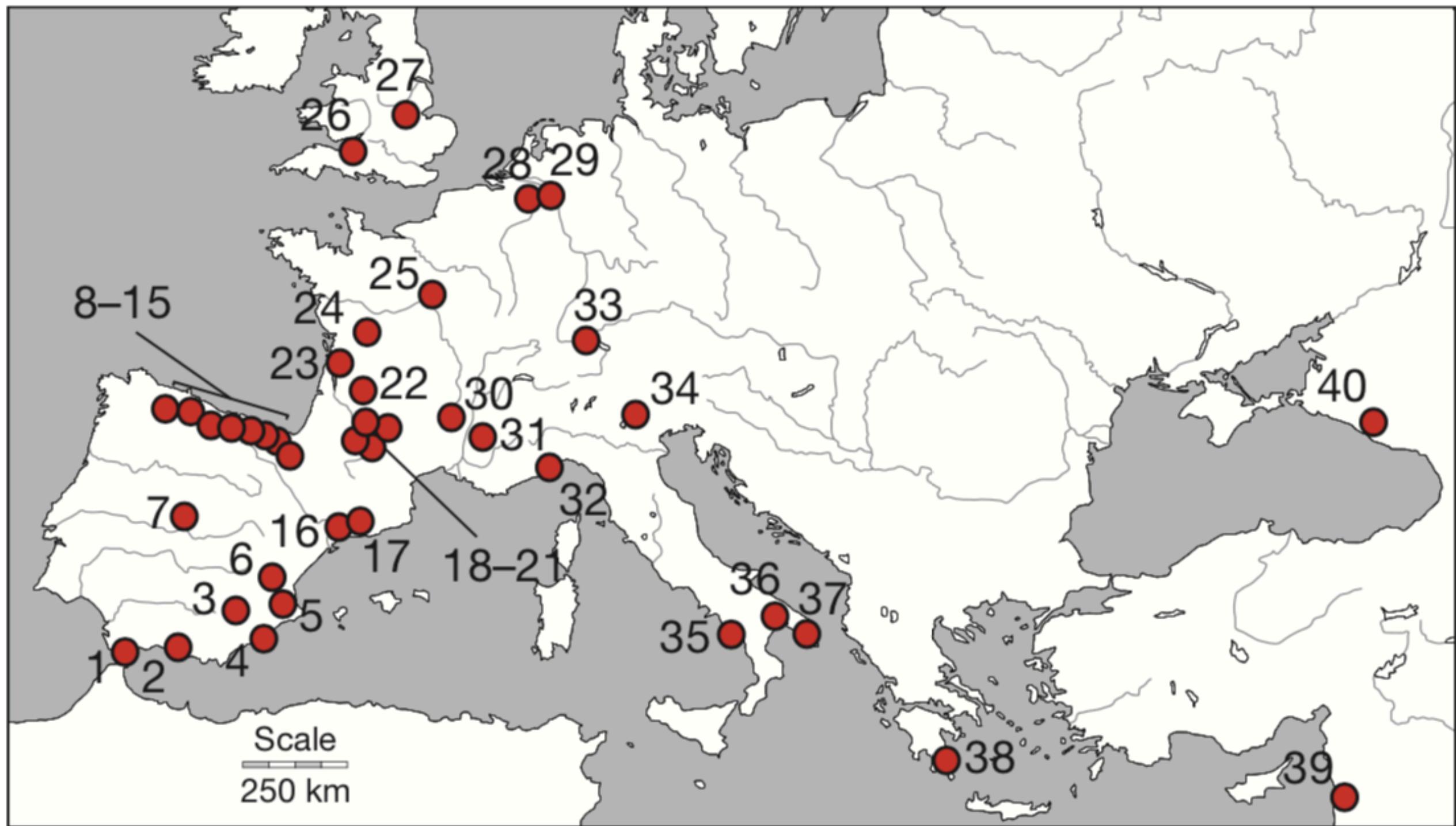
# Structure?



A photograph of four ABBA wax figures standing against a dark background with purple lighting. From left to right: Björn Ulvaeus in a white zip-up top, Agnetha Fältskog in a white and red cheerleader-style vest with 'ABBA' on it, Anni-Frid Lyngstad in a white cardigan over a striped top, and Benny Andersson in a blue jacket over a striped shirt.

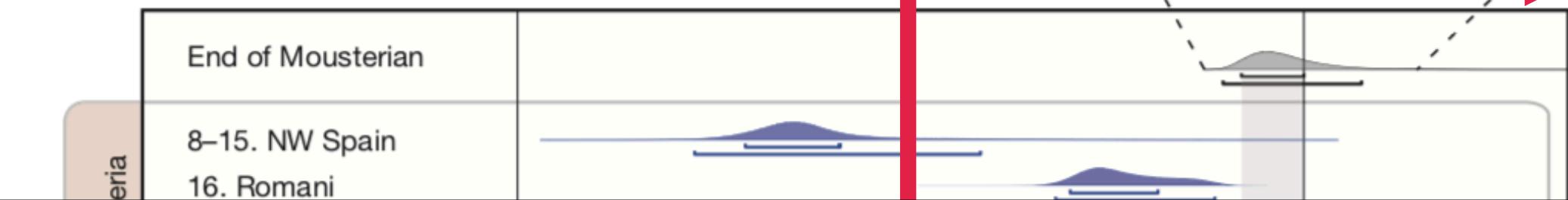
# Exercises

[github.com/bodkan/ku-introgression2024](https://github.com/bodkan/ku-introgression2024)

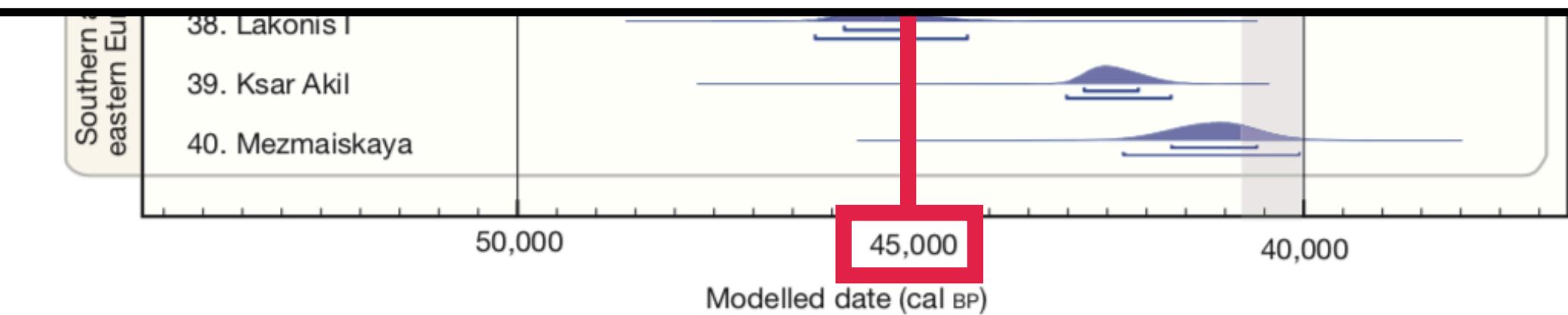


Higham et al. (Nature, 2014)

anatomically modern humans in Europe



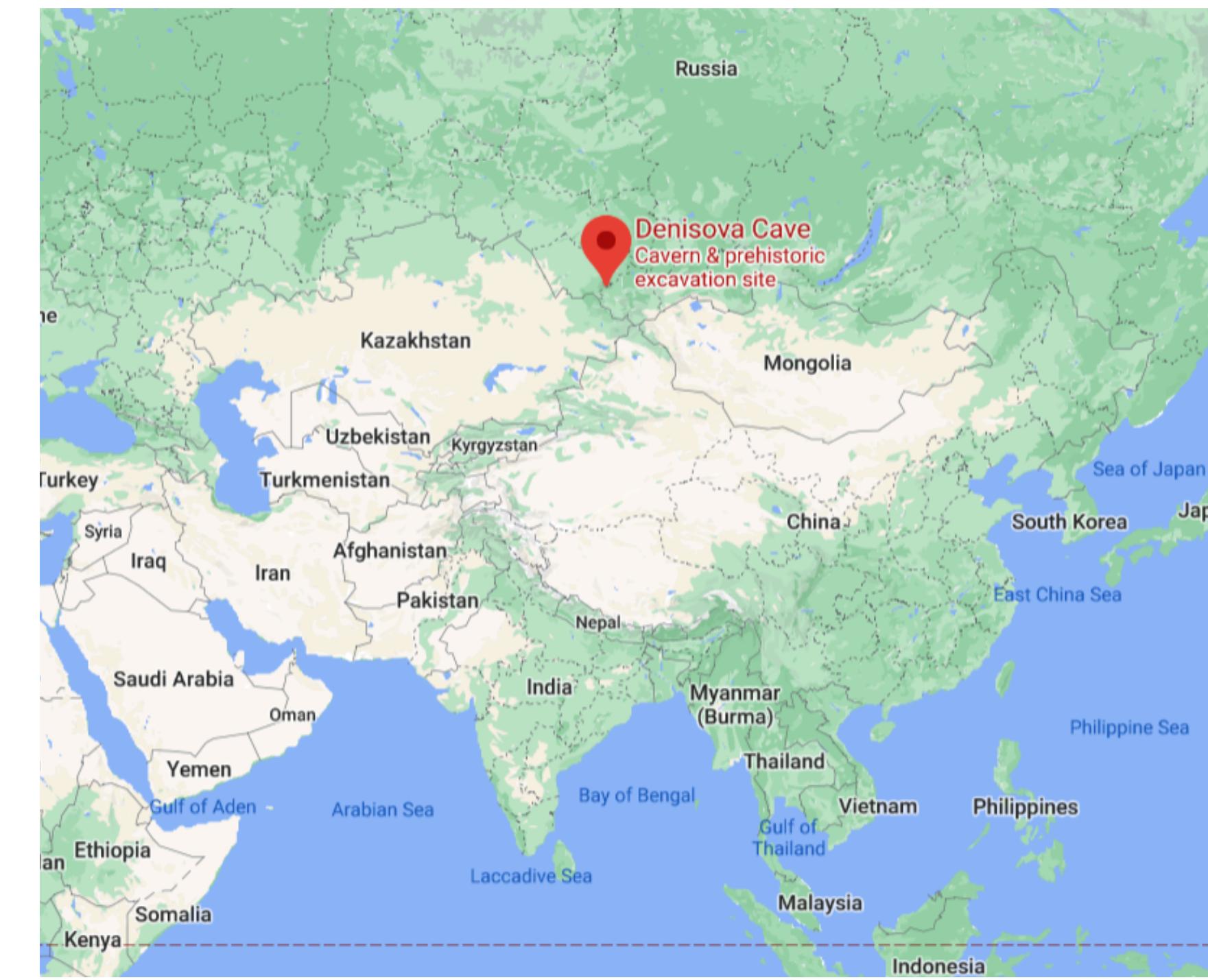
# Neanderthals were not the only archaics around at the time



Higham et al. (Nature, 2014)



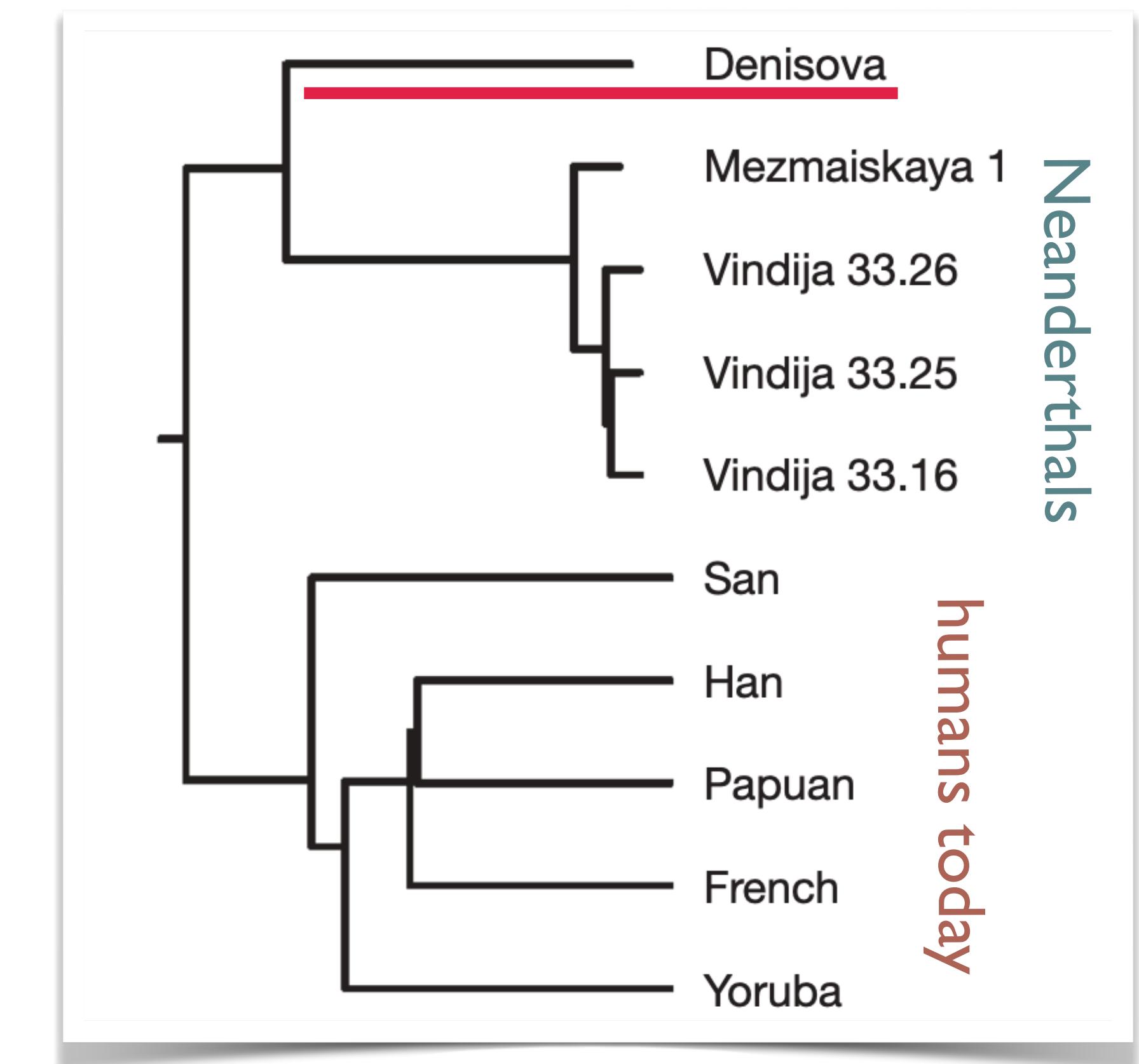
By Демин Алексей Барнаул - Own work, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=48890364>





Denisova cave

By Демин Алексей Барнаул - Own work, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=48890364>



## nature

Open Access | Published: 22 December 2010

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

David Reich , Richard E. Green, Martin Kircher, Johannes Krause, Nick Patterson, Eric Y. Durand, Bence Viola , Adrian W. Briggs, Udo Stenzel, Philip L. F. Johnson, Tomislav Maricic, Jeffrey M. Good, Tomas Marques-Bonet, Can Alkan, Qiaomei Fu, Swapan Mallick, Heng Li, Matthias Meyer, Evan E. Eichler, Mark Stoneking, Michael Richards, Sahra Talamo, Michael V. Shunkov, Anatoli P. Derevianko, Jean-Jacques Hublin, Janet Kelso, Montgomery Slatkin & Svante Pääbo

— Show fewer authors

*Nature* 468, 1053–1060 (2010) | [Cite this article](#)

# Traces of Denisovans in genomes of present-day humans

$f4(\text{human 1}, \text{human 2}; \text{Denisovan}, \text{Chimp})$

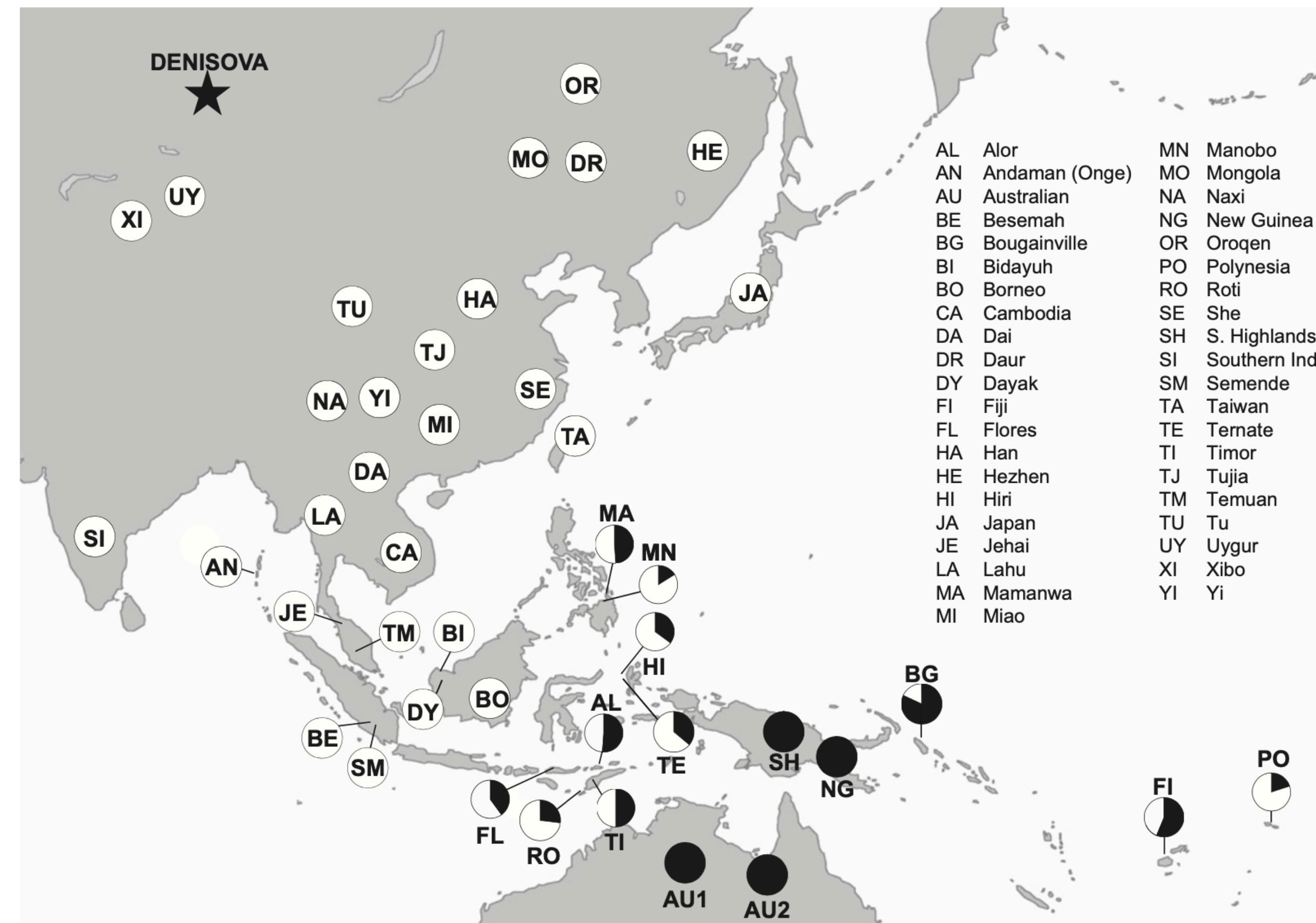
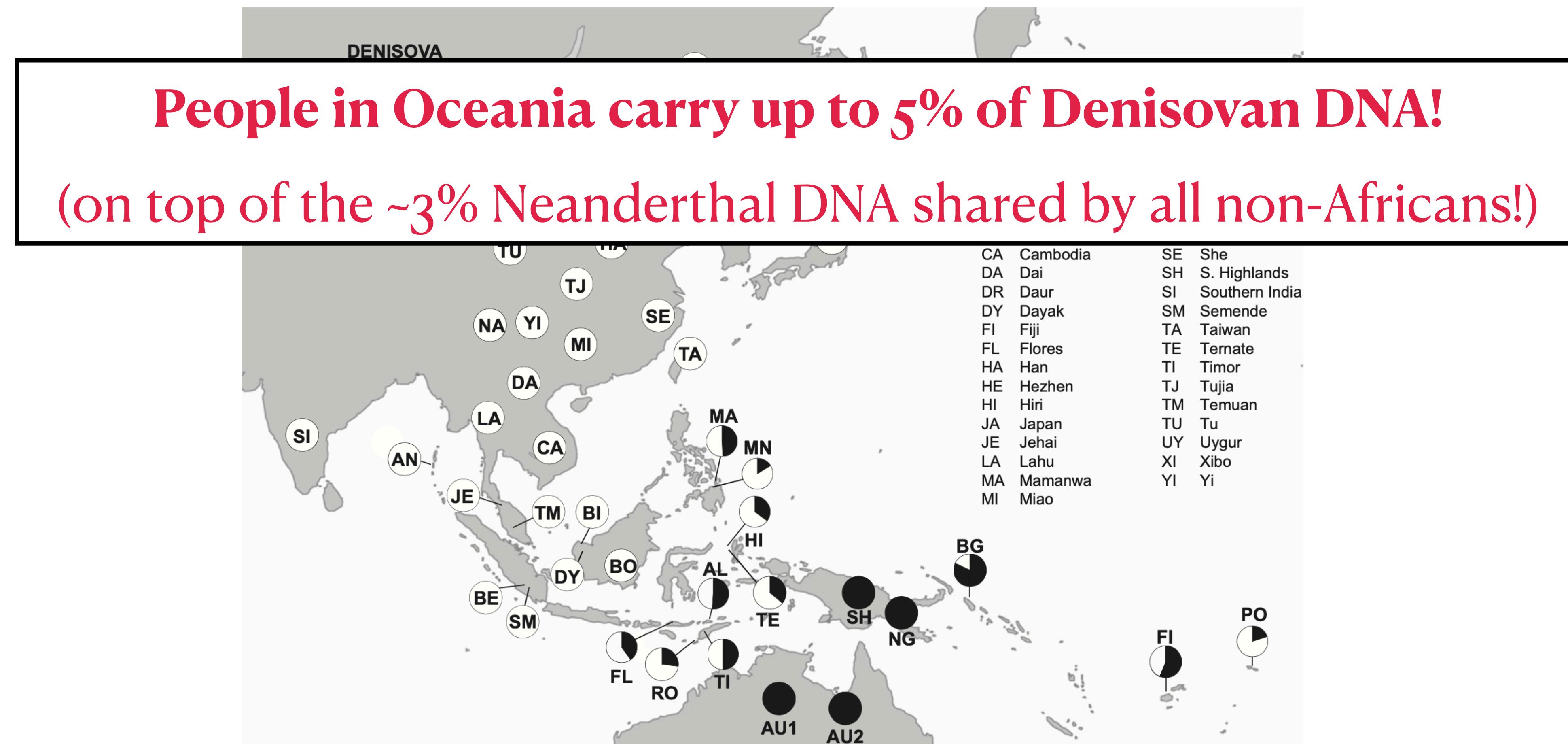


Figure 1. Denisovan Genetic Material as a Fraction of that in New Guineans

Reich et al., 2011

# Traces of Denisovans in genomes of present-day humans

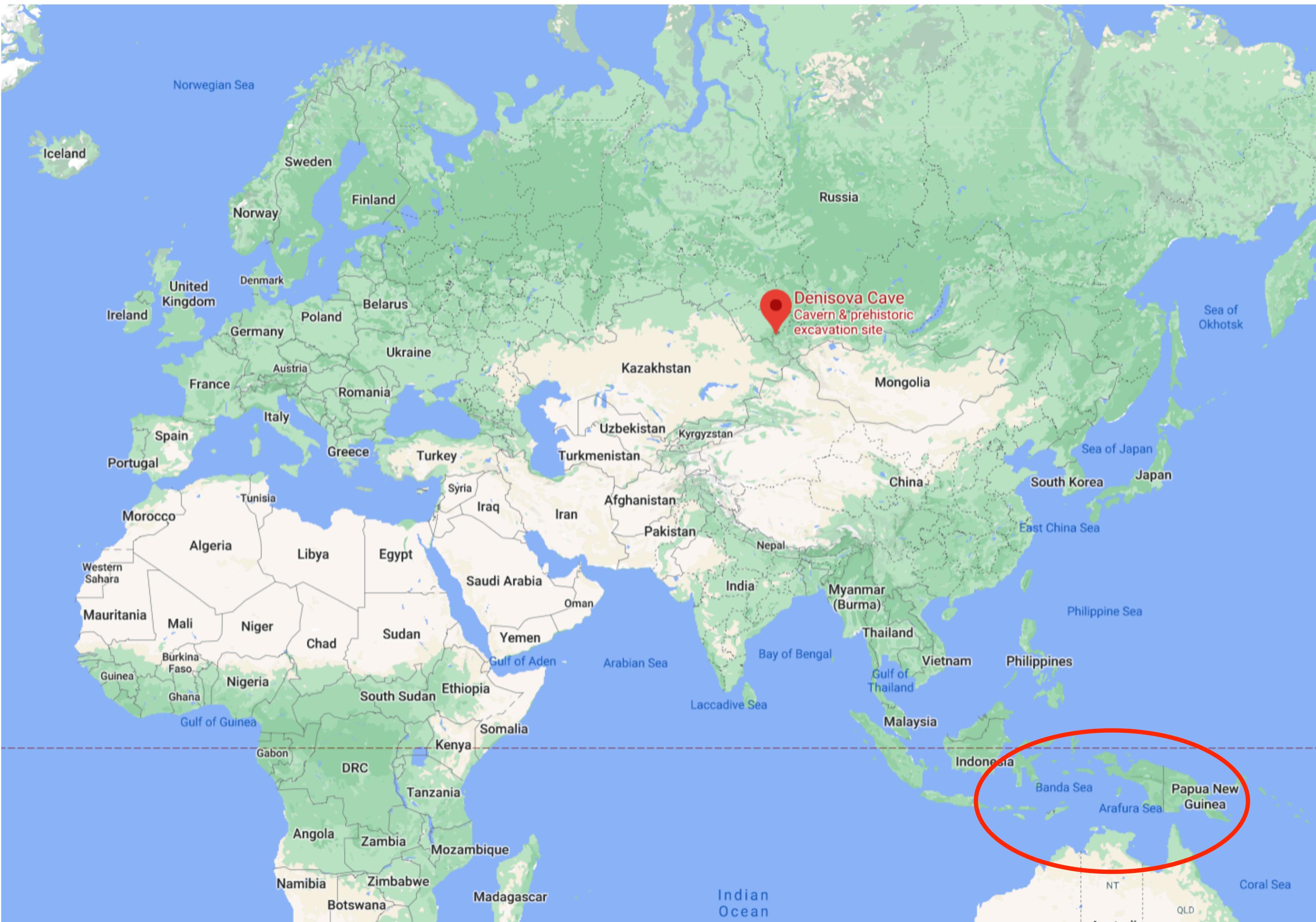
$f4(\text{human 1}, \text{human 2}; \text{Denisovan}, \text{Chimp})$



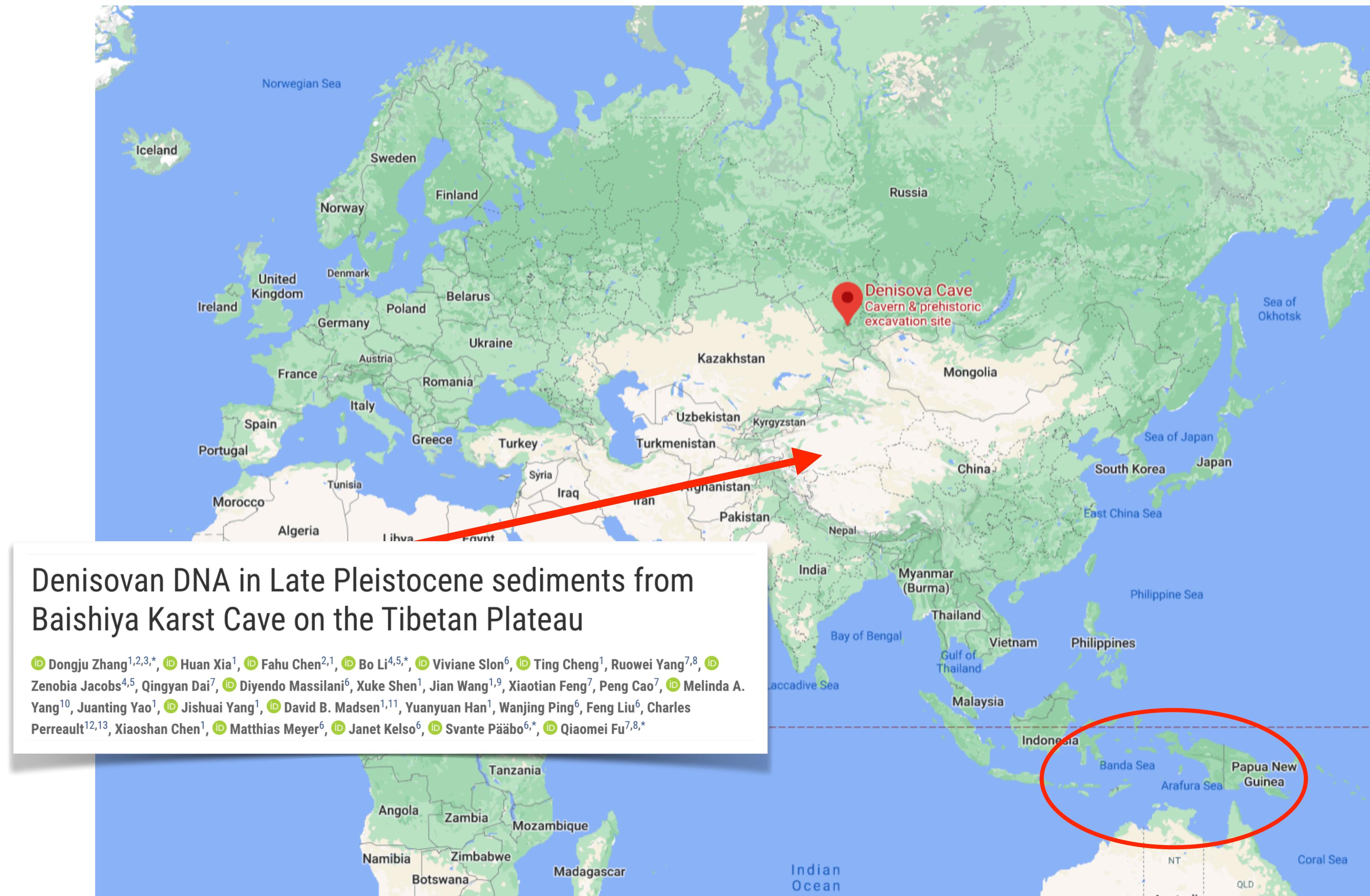
Reich et al., 2011

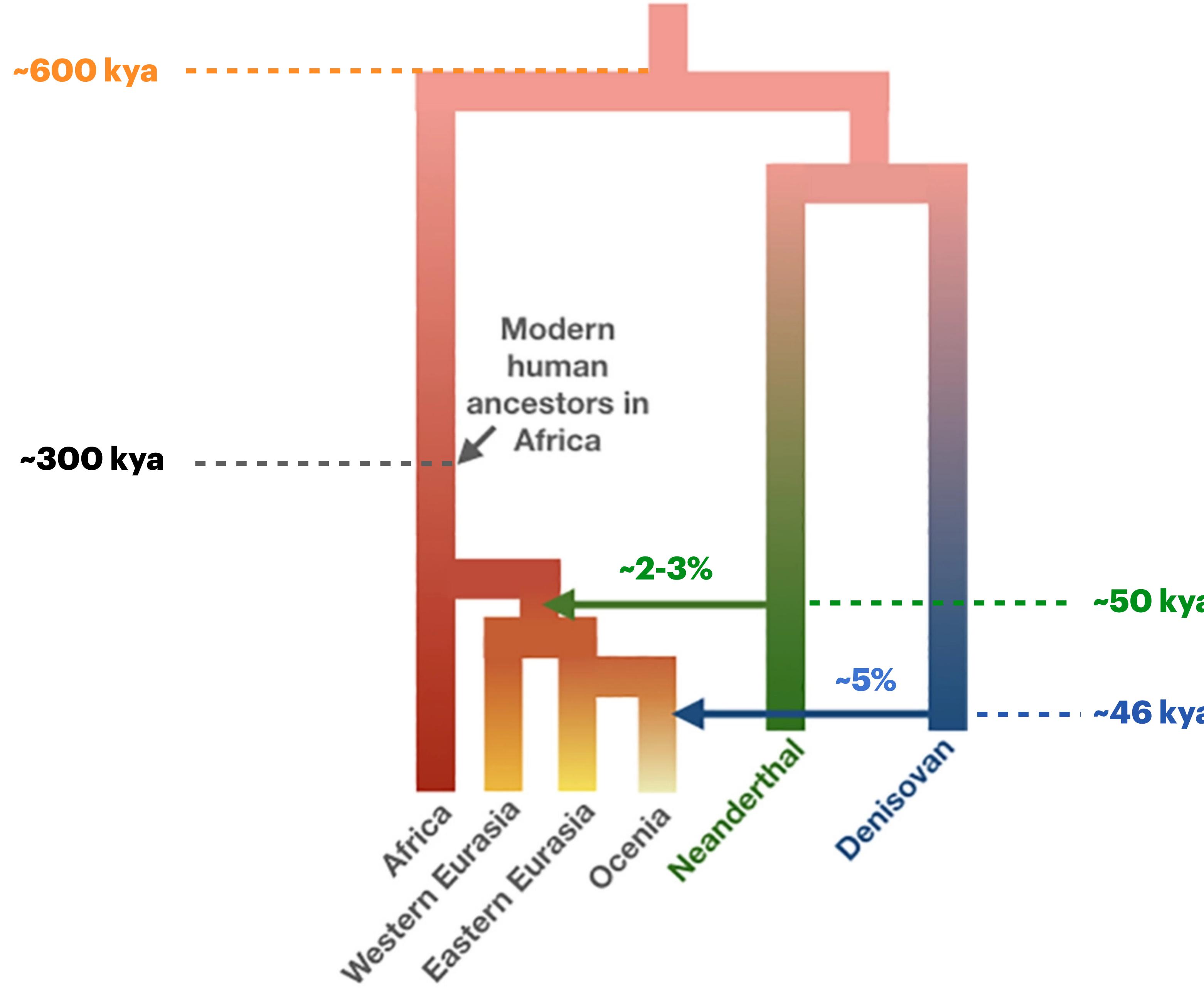
Figure 1. Denisovan Genetic Material as a Fraction of that in New Guineans

# Geographic range of Denisovans?



# Geographic range of Denisovans?

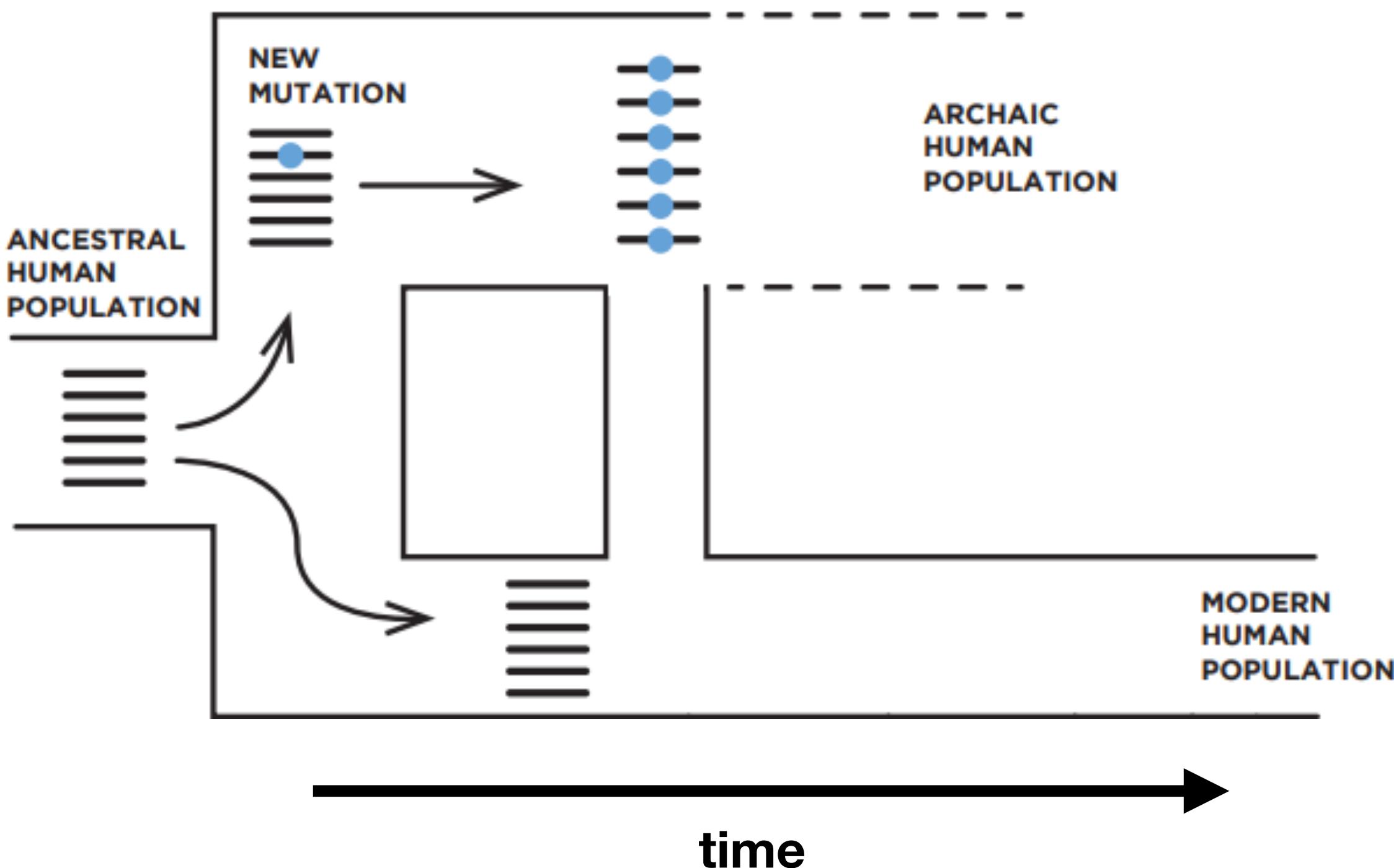




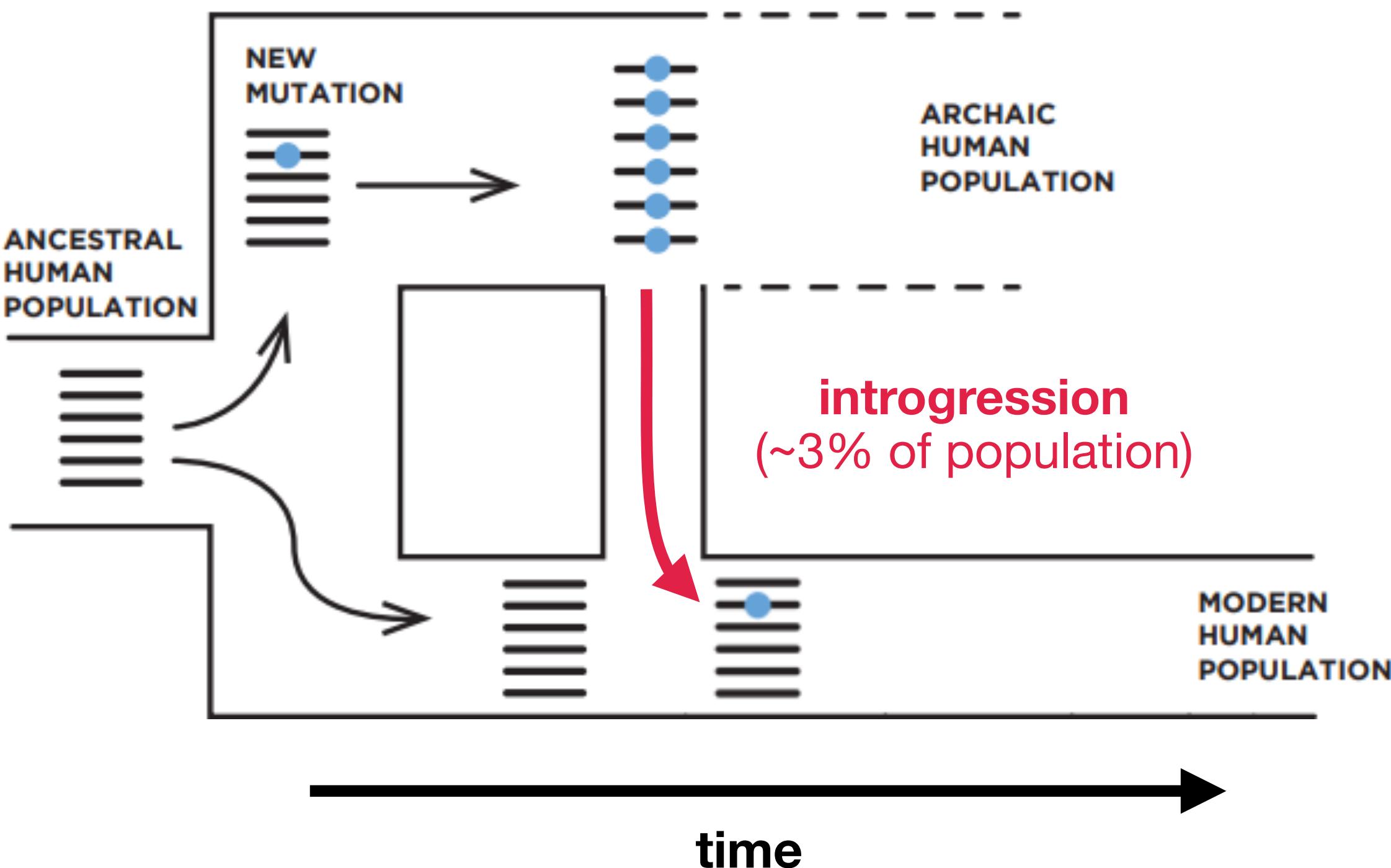
adapted from the Yearbook of Physical Anthropology (2019) by Gokcumen

# Consequences of introgression on human functional biology?

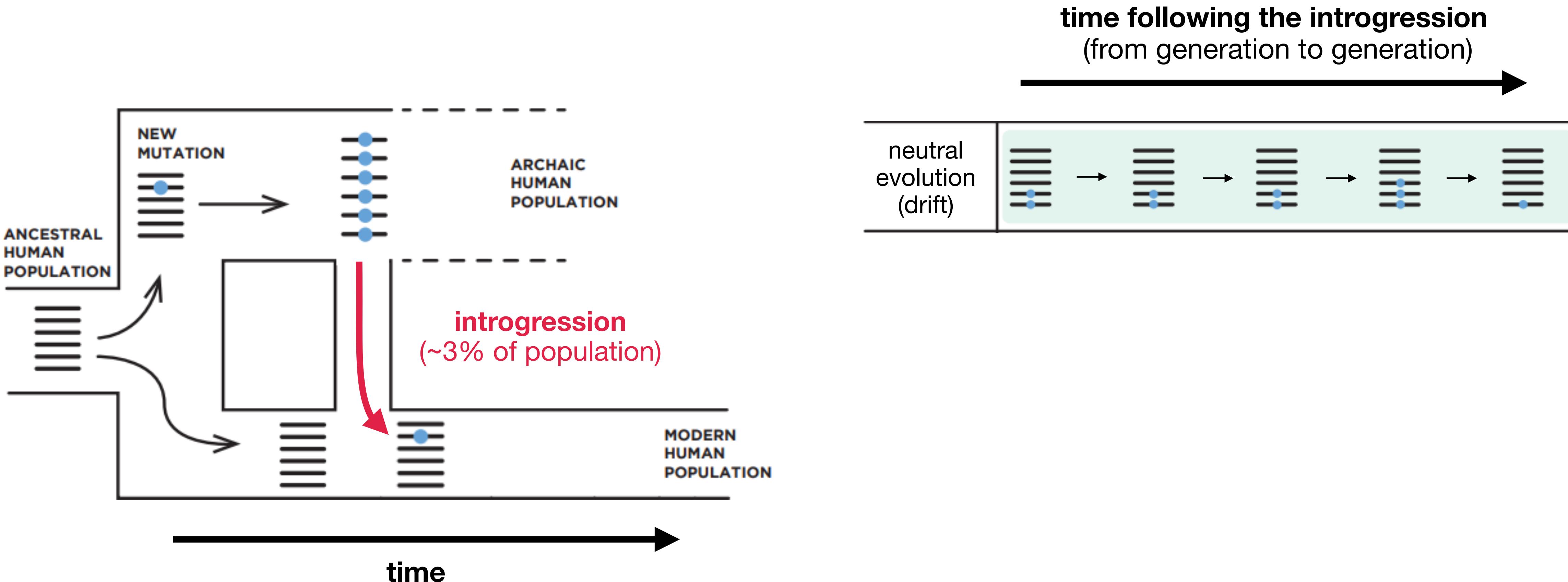
# What are the possible fates of introgressed DNA?



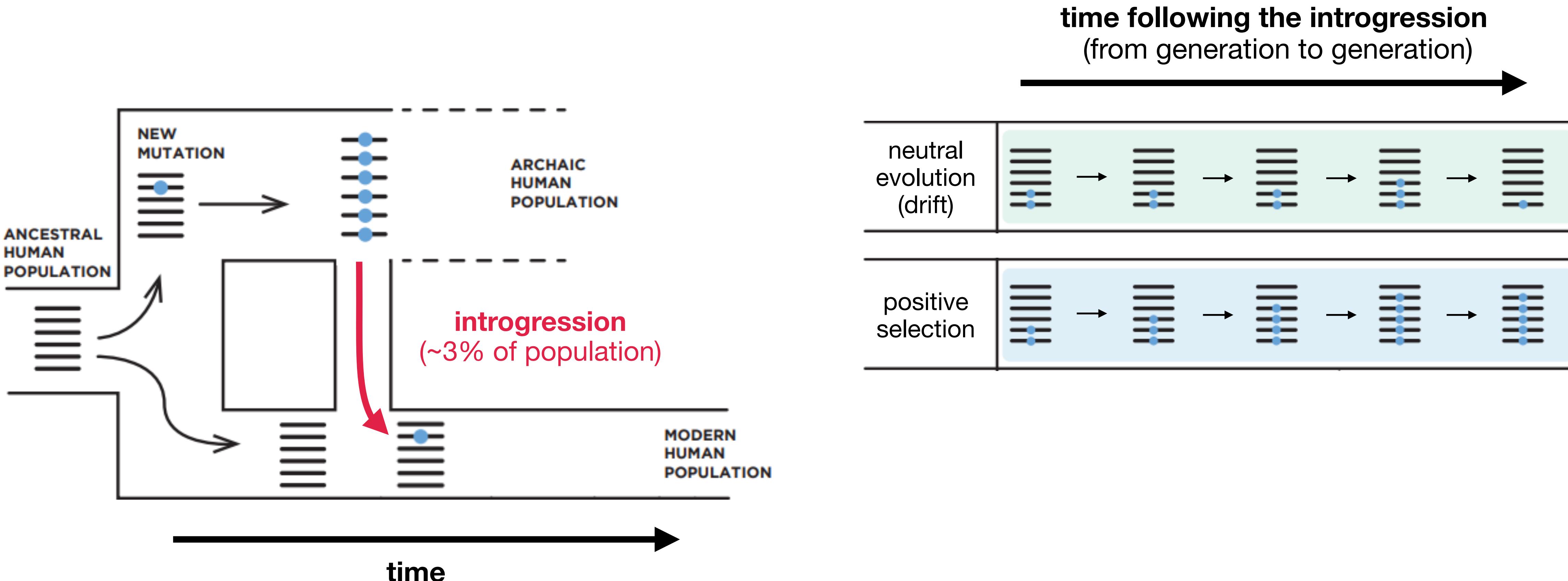
# What are the possible fates of introgressed DNA?



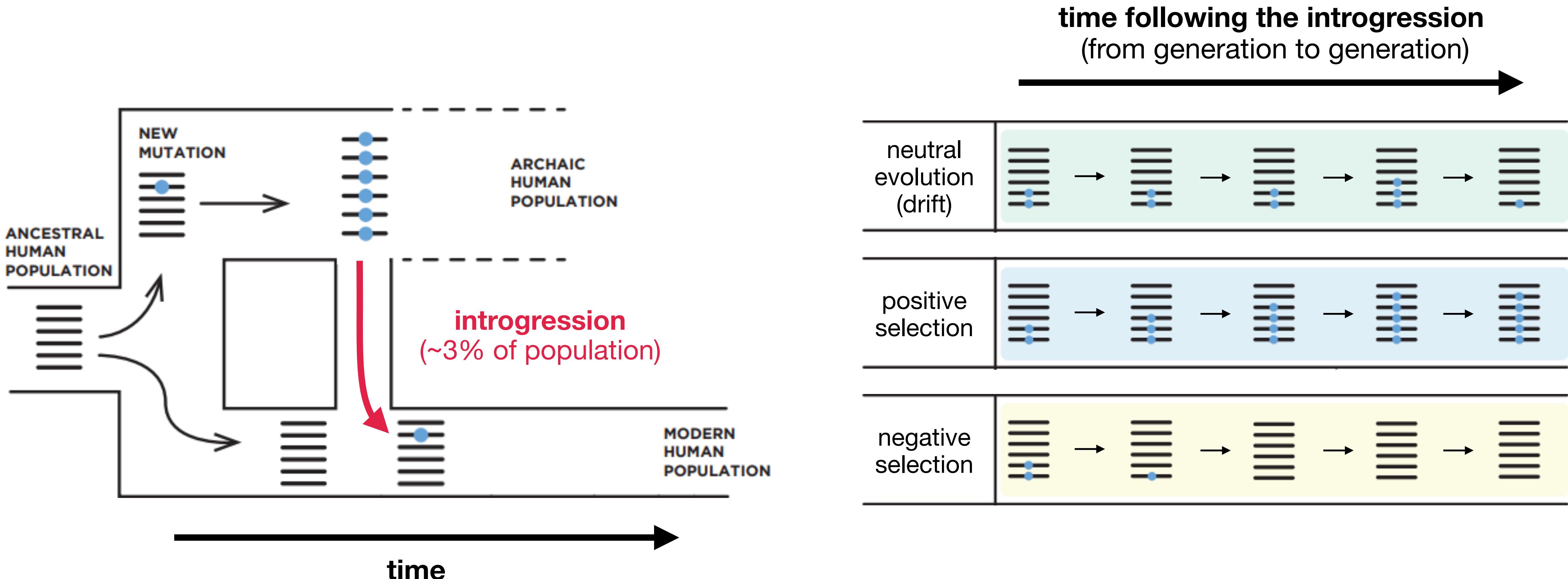
# What are the possible fates of introgressed DNA?

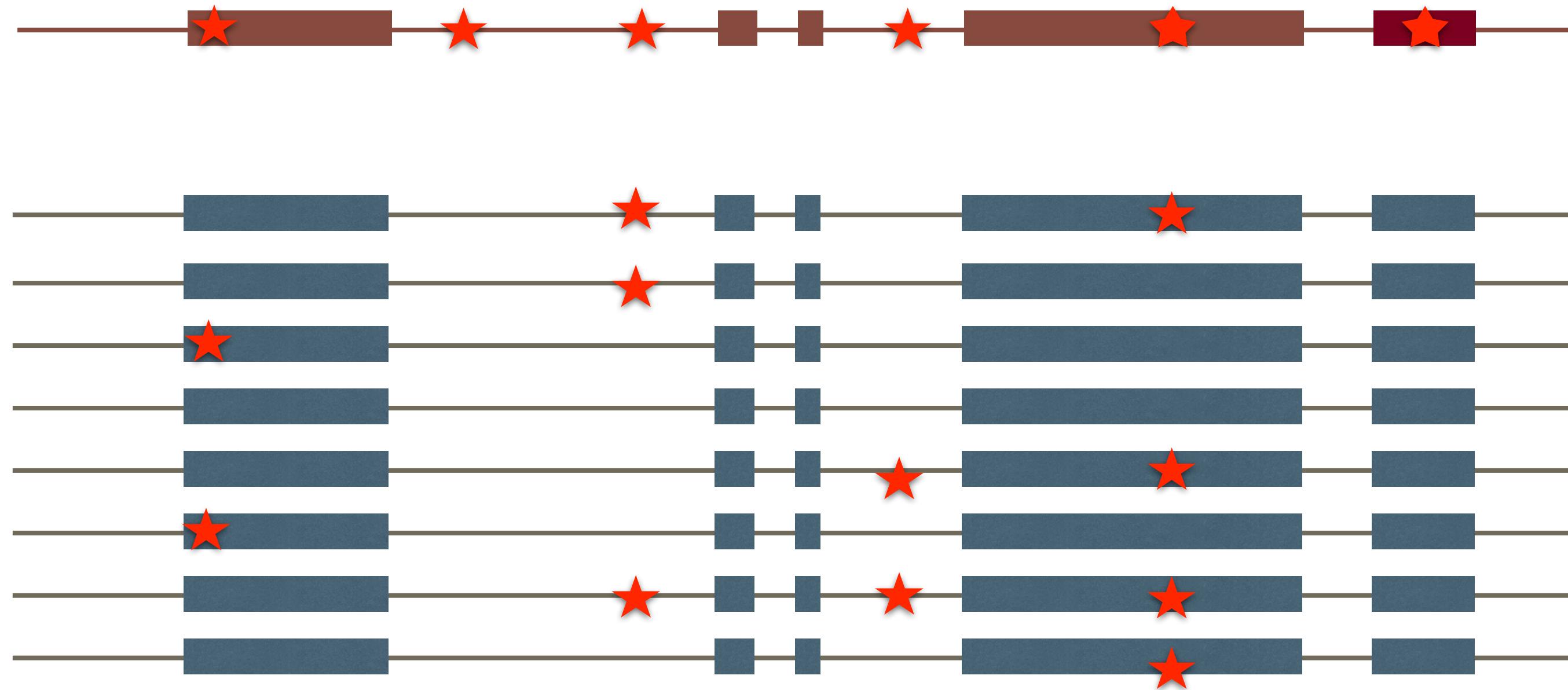


# What are the possible fates of introgressed DNA?



# What are the possible fates of introgressed DNA?





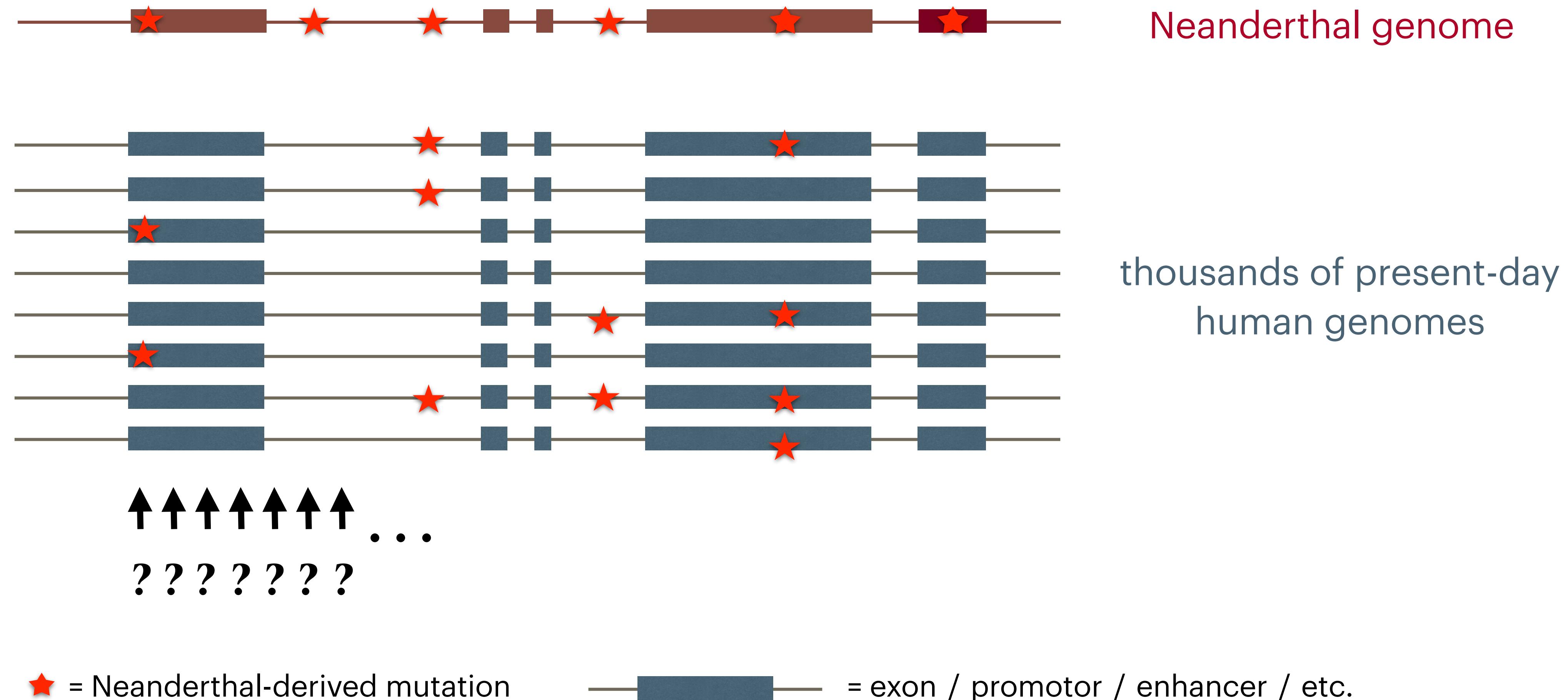
Neanderthal genome

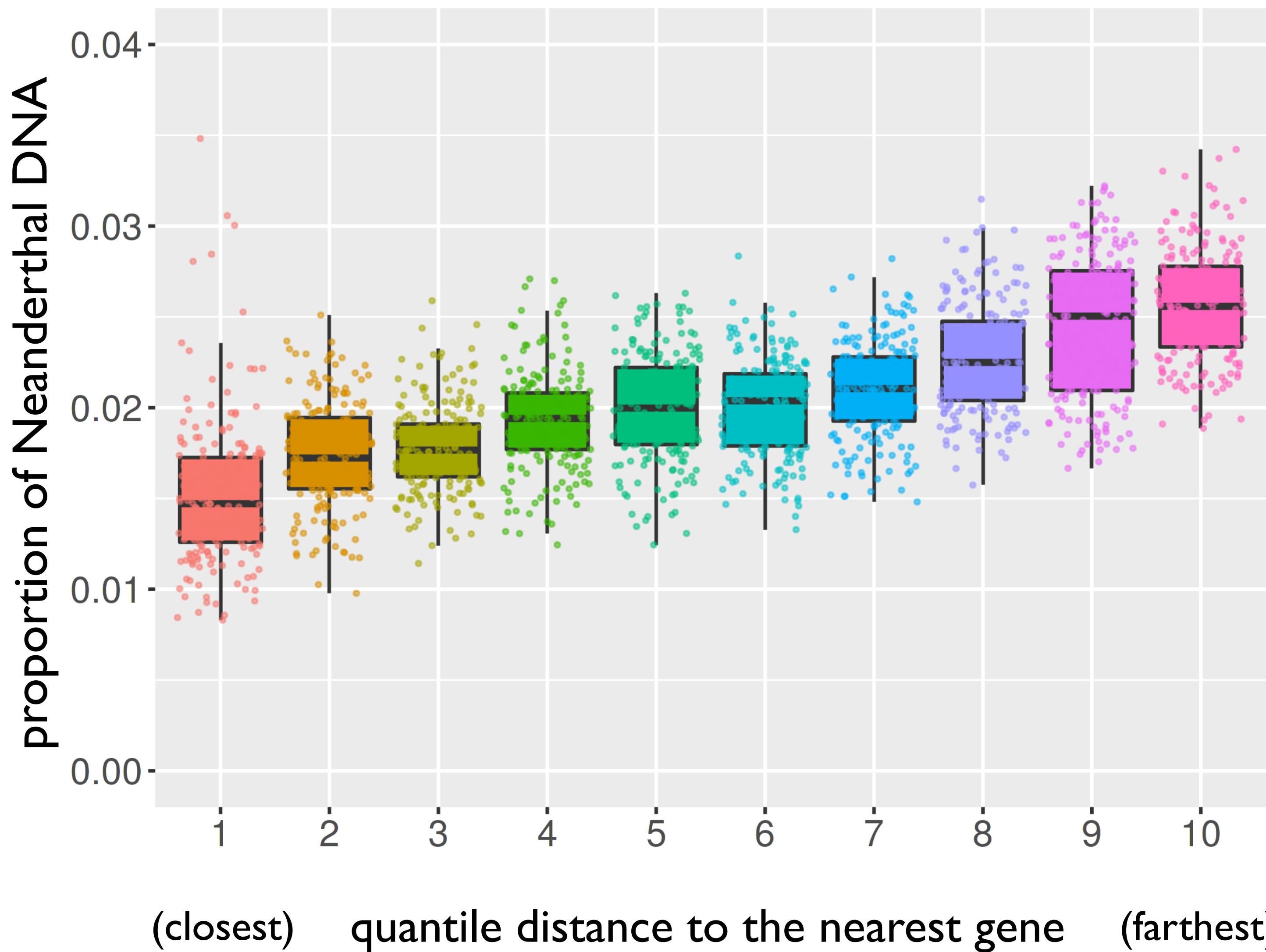
thousands of present-day  
human genomes

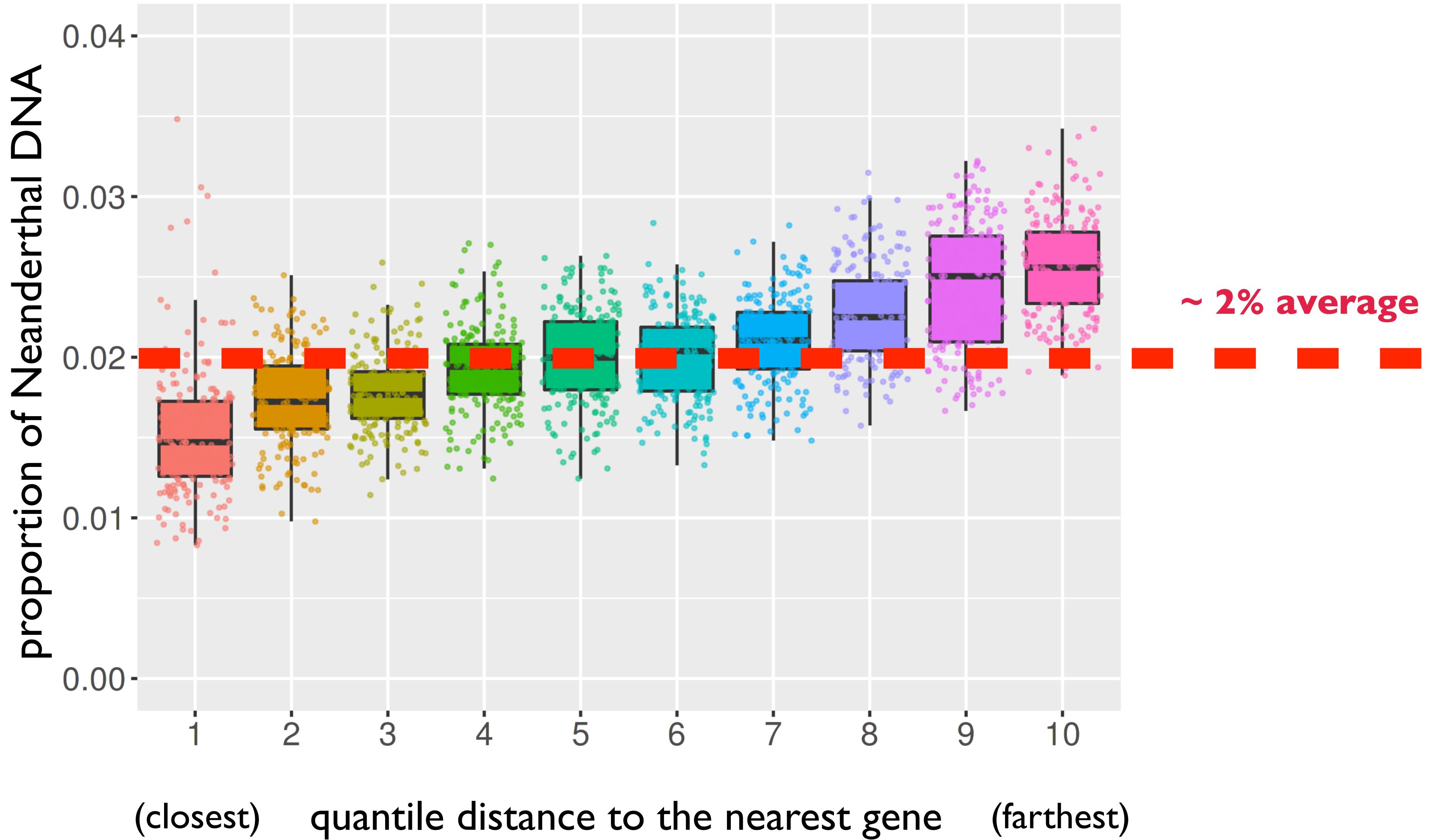
★ = Neanderthal-derived mutation

—■— = exon / promotor / enhancer / etc.

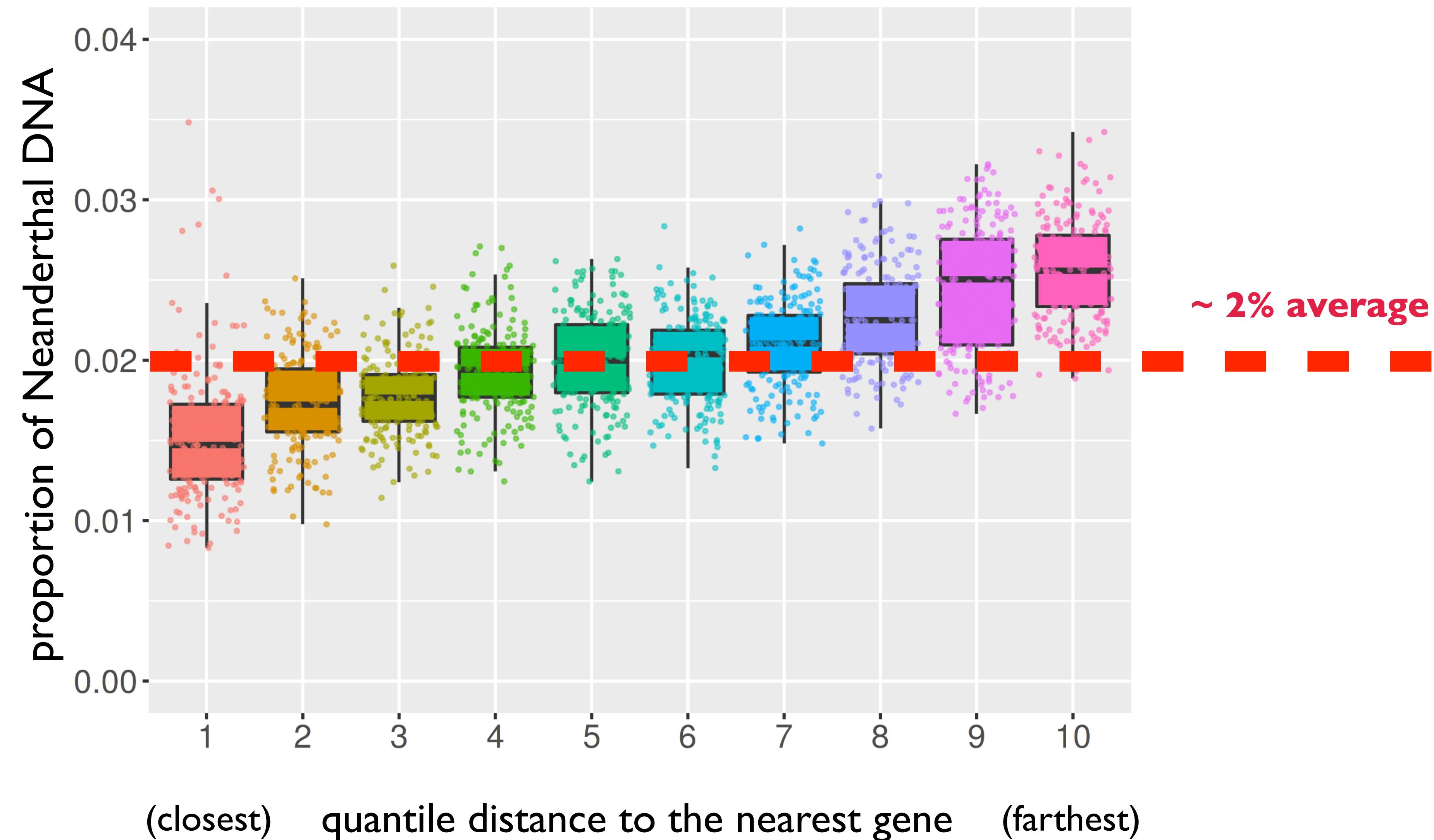
# Is there a relationship between the presence of a Neanderthal allele and its distance to the nearest important gene?



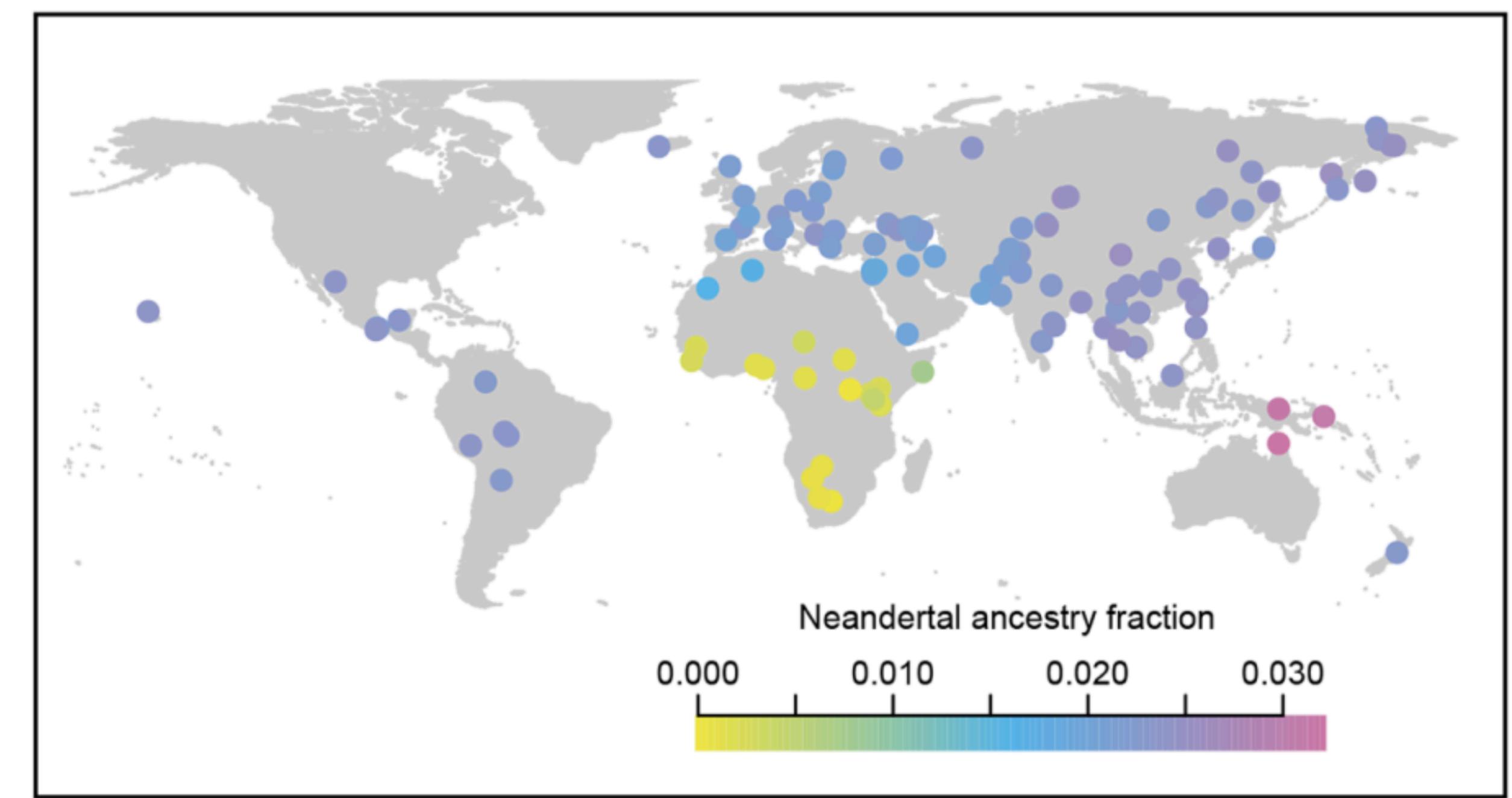
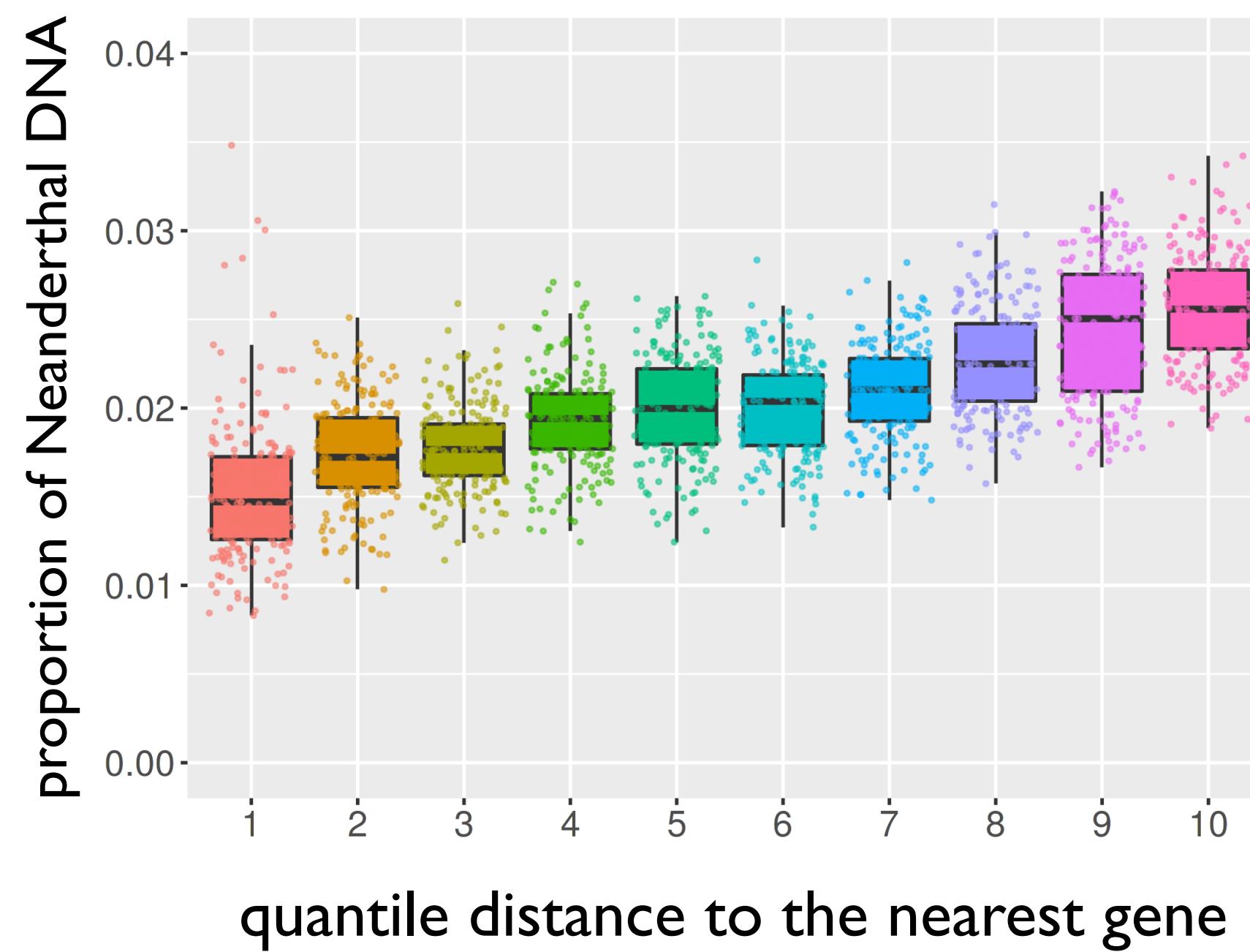




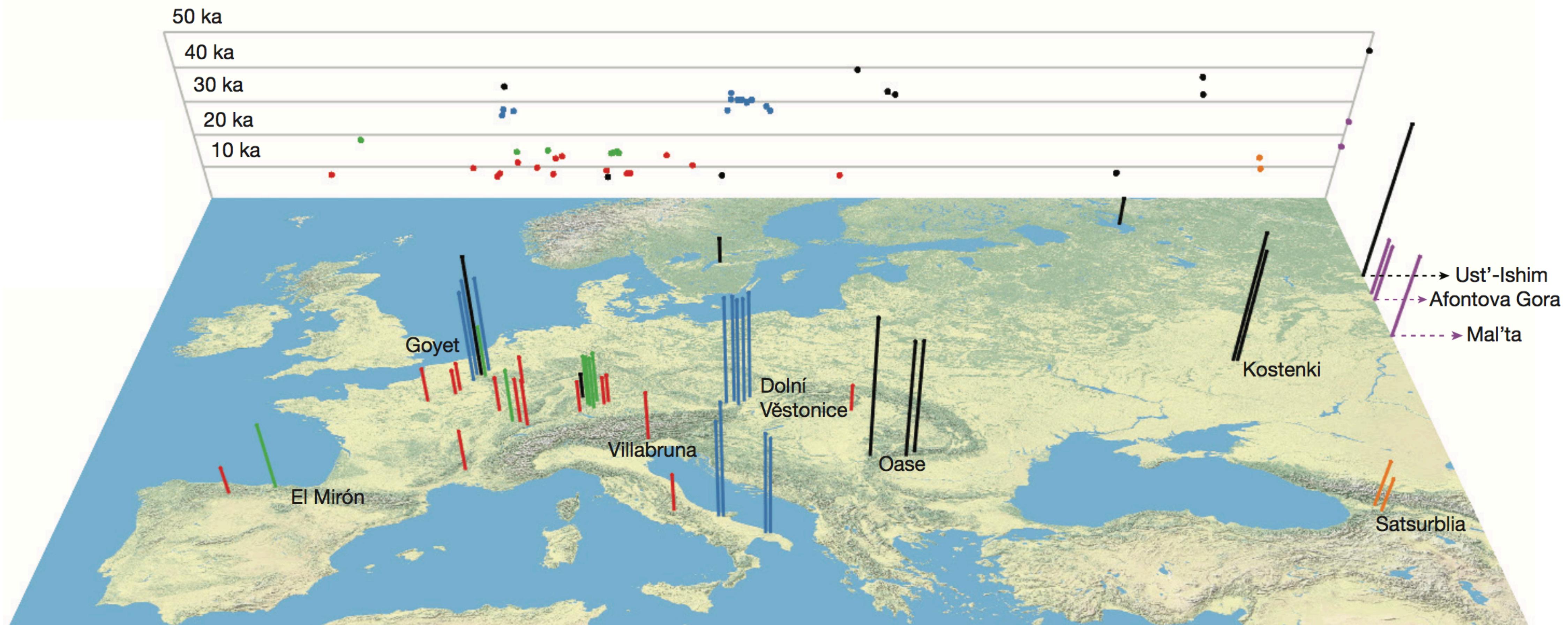
# Genome-wide selection *against* Neanderthal DNA!



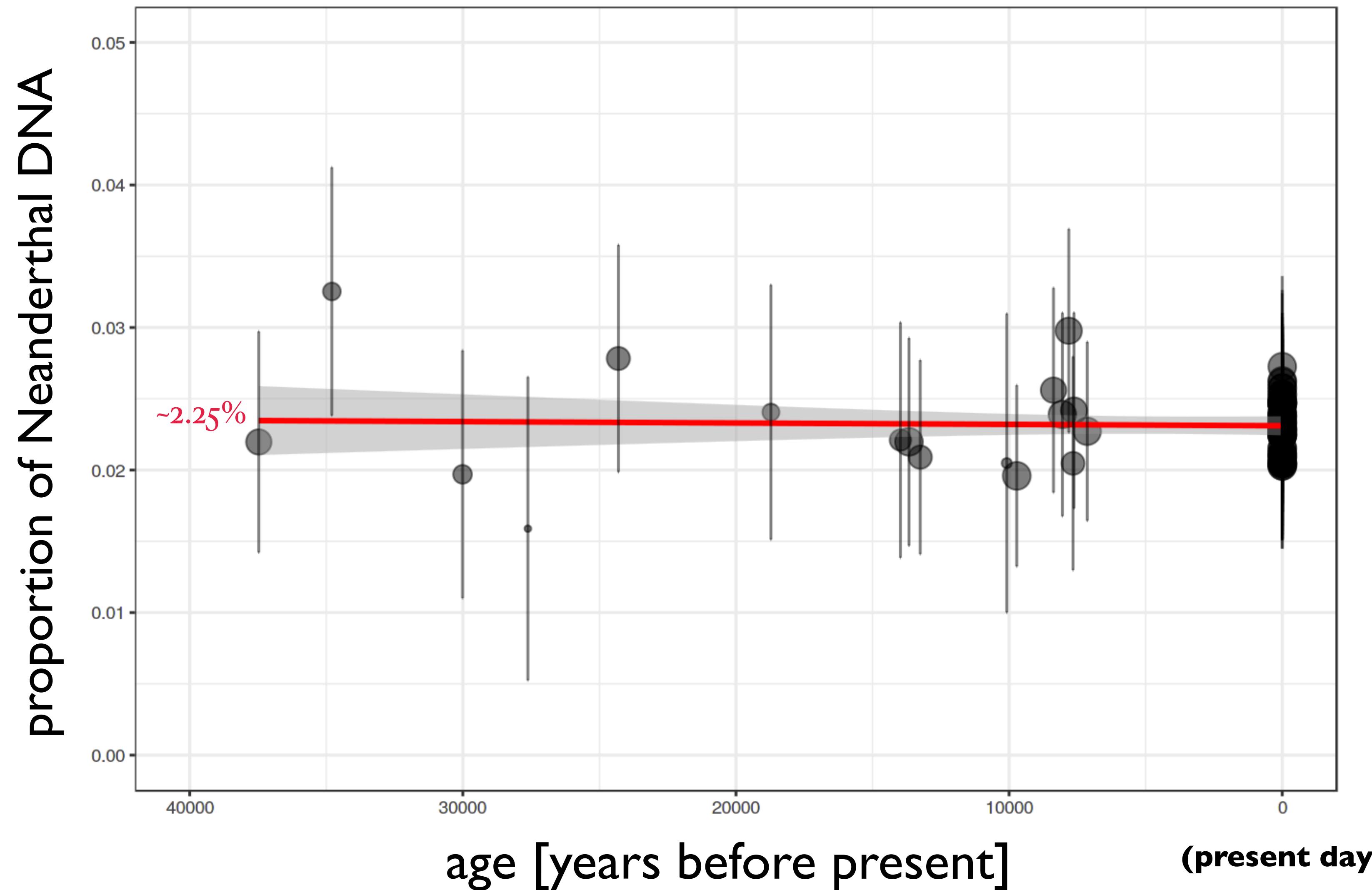
# On what time scale did this happen?



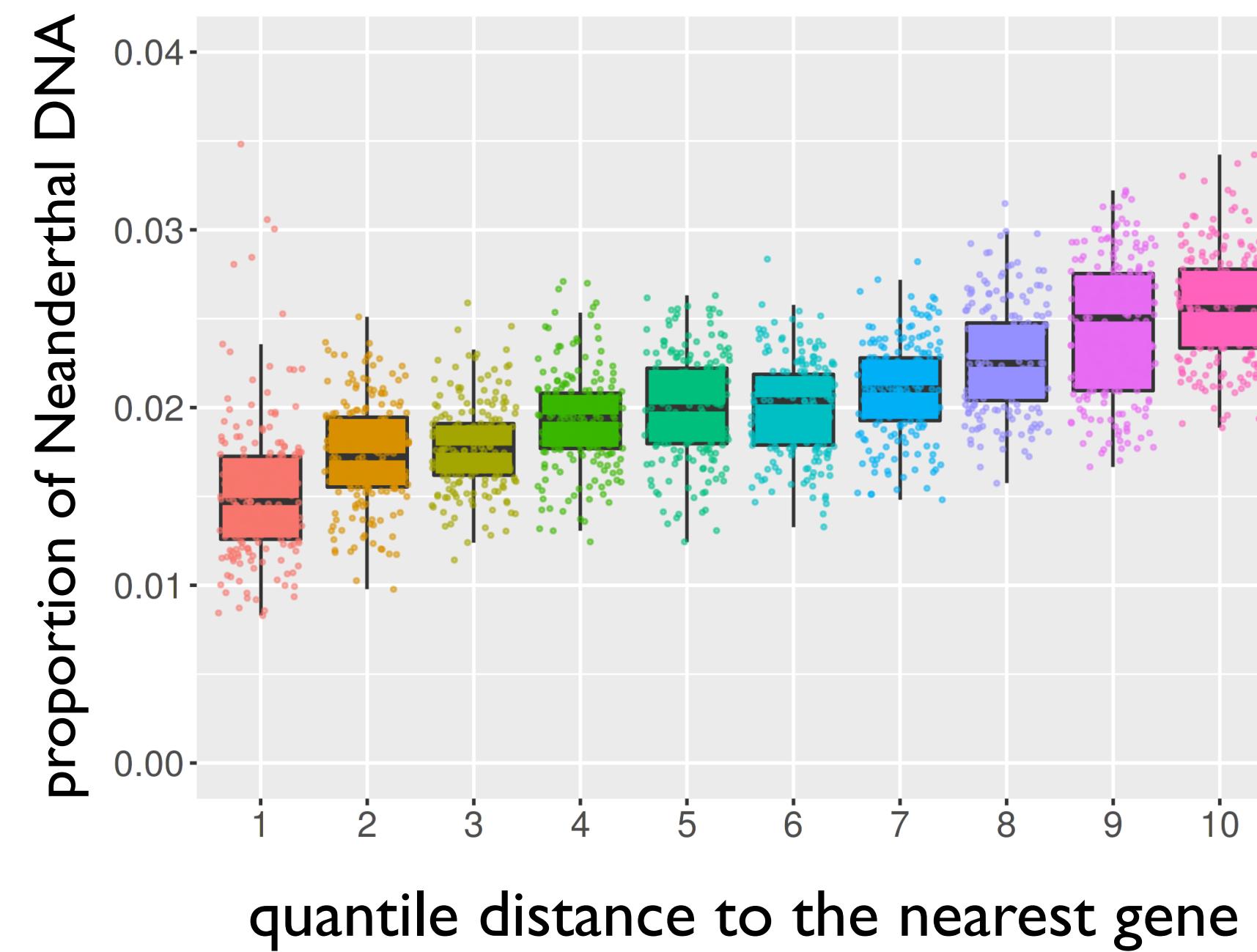
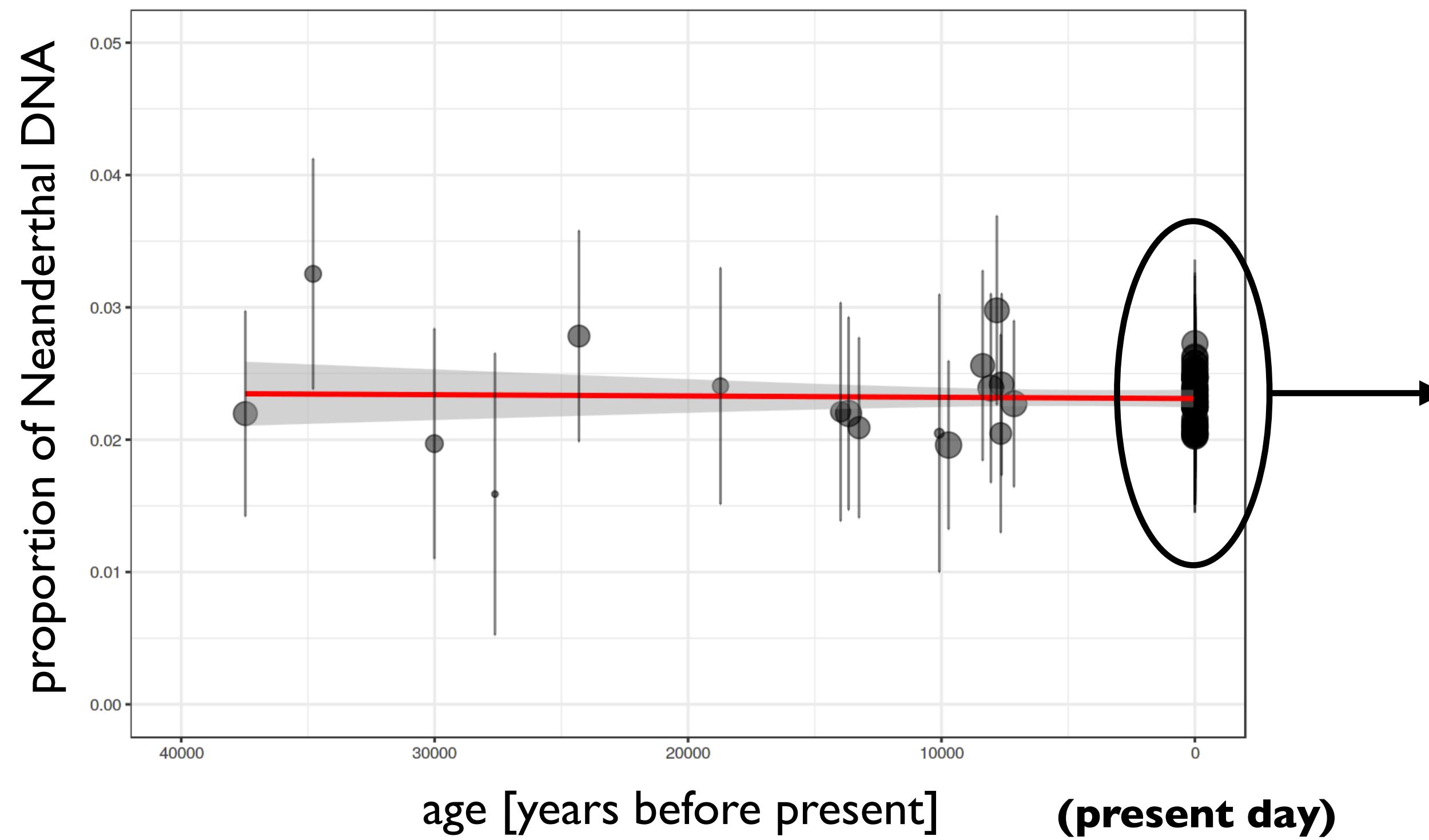
# aDNA of “early-modern humans” from Europe (45-10 kya BP)



# Proportion of Neanderthal ancestry in modern humans remained constant for tens of thousands of years

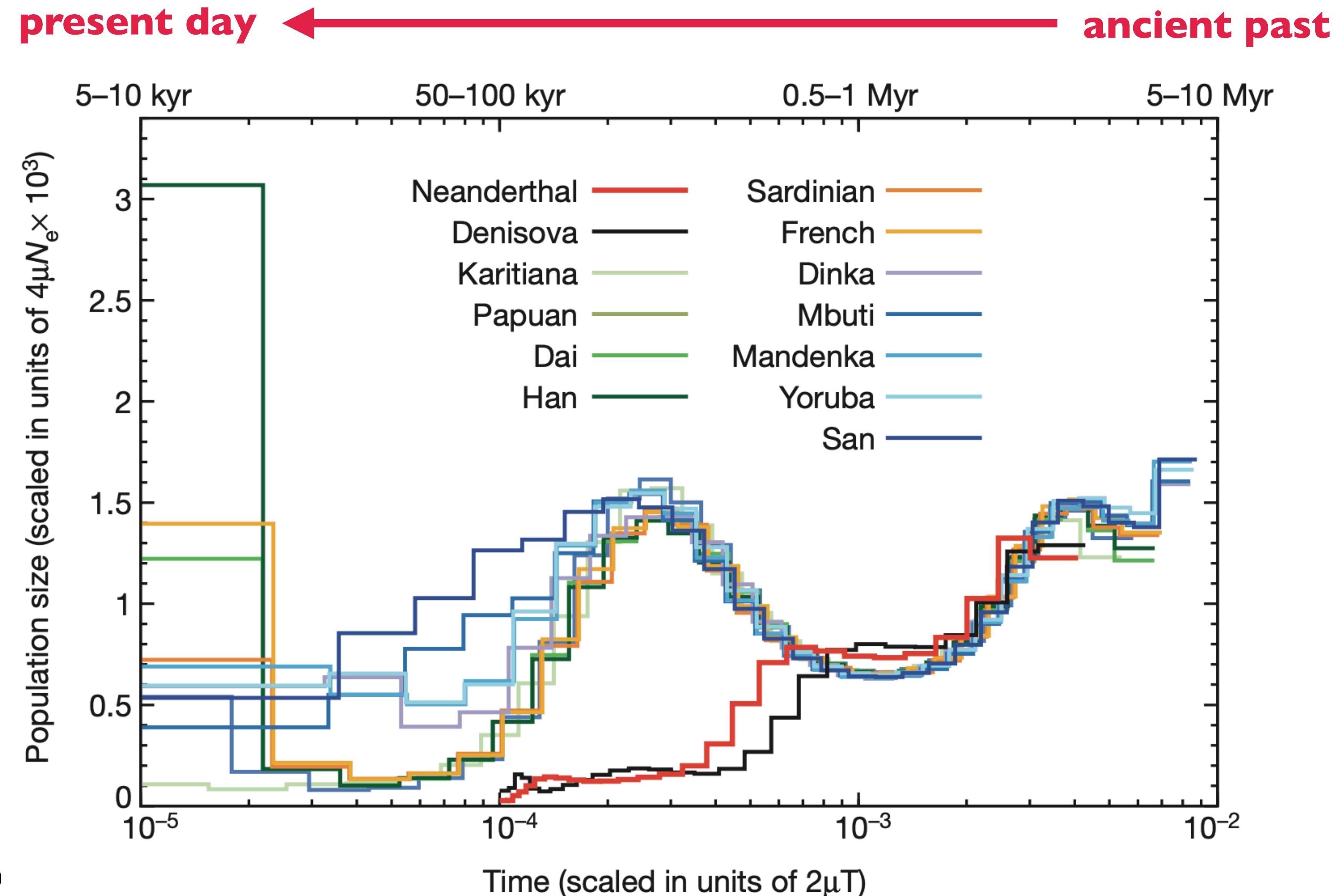


# Empirical trajectory of Neanderthal ancestry contradicts the signal of negative selection?



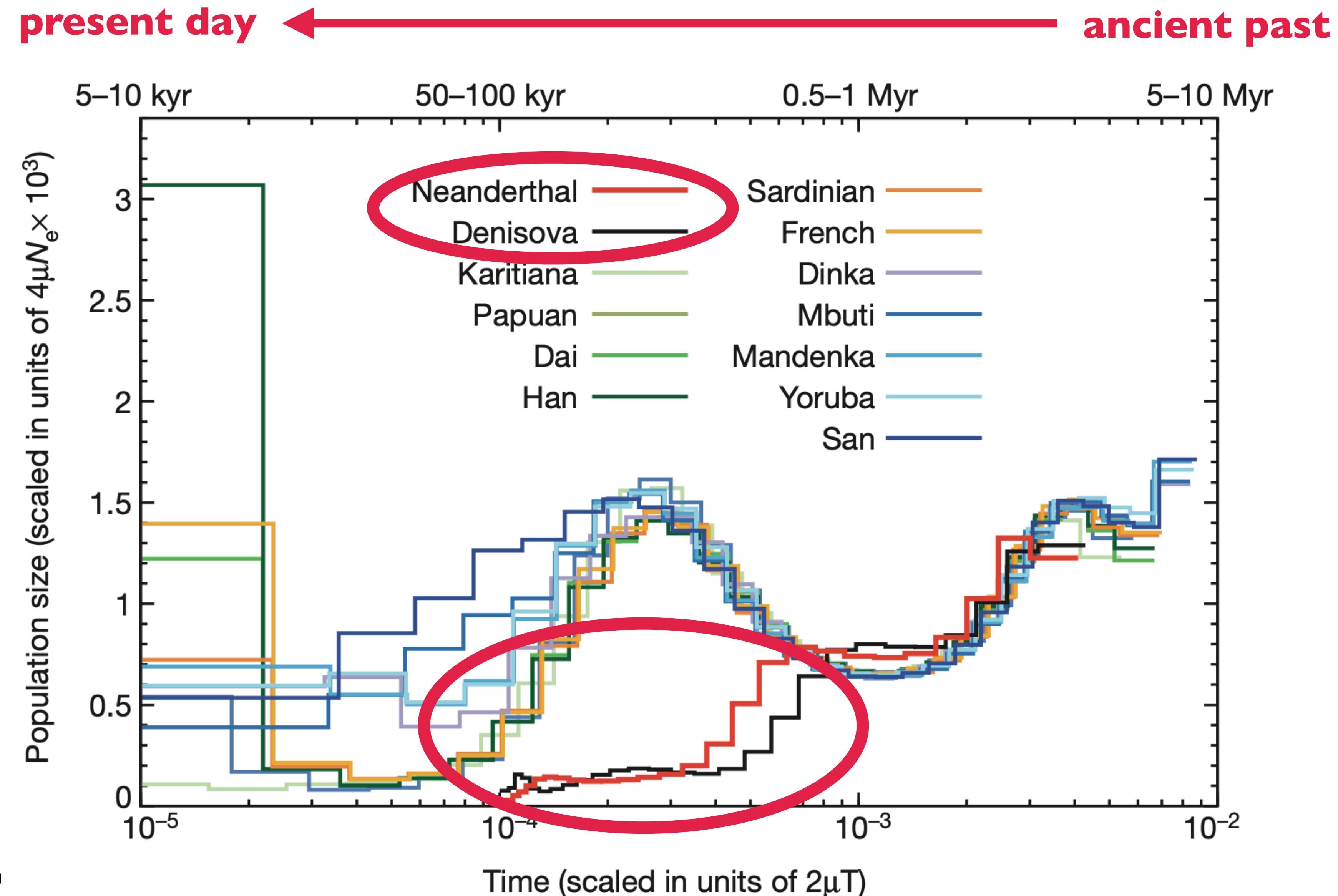
# Demographic collapse of archaic human populations

Pairwise Sequentially  
Markovian Coalescent  
(PSMC) model



# Demographic collapse of archaic human populations

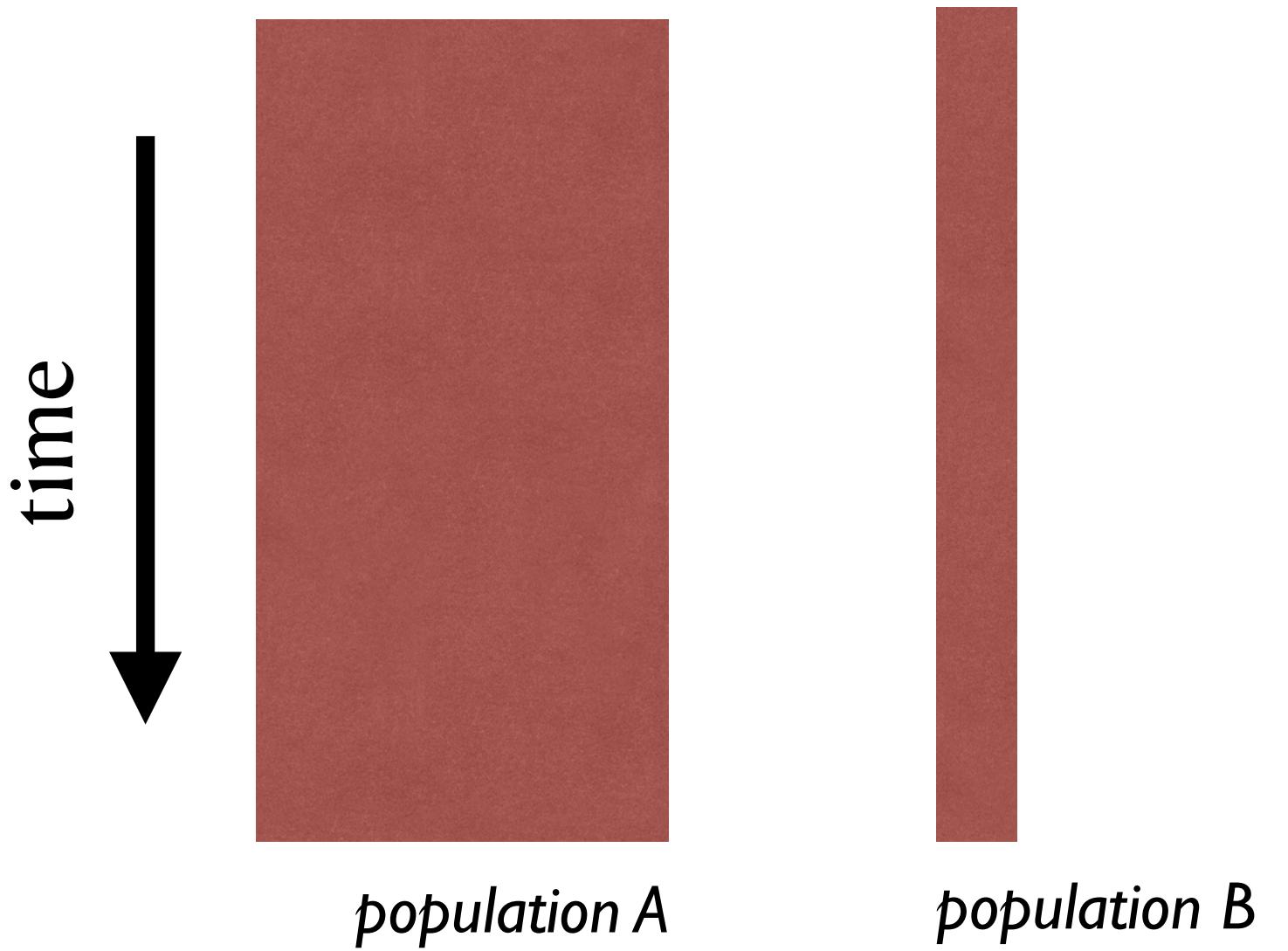
Pairwise Sequentially  
Markovian Coalescent  
(PSMC) model



**What was the impact of this on  
accumulation of deleterious mutations?**

# Nearly-neutral theory of molecular evolution

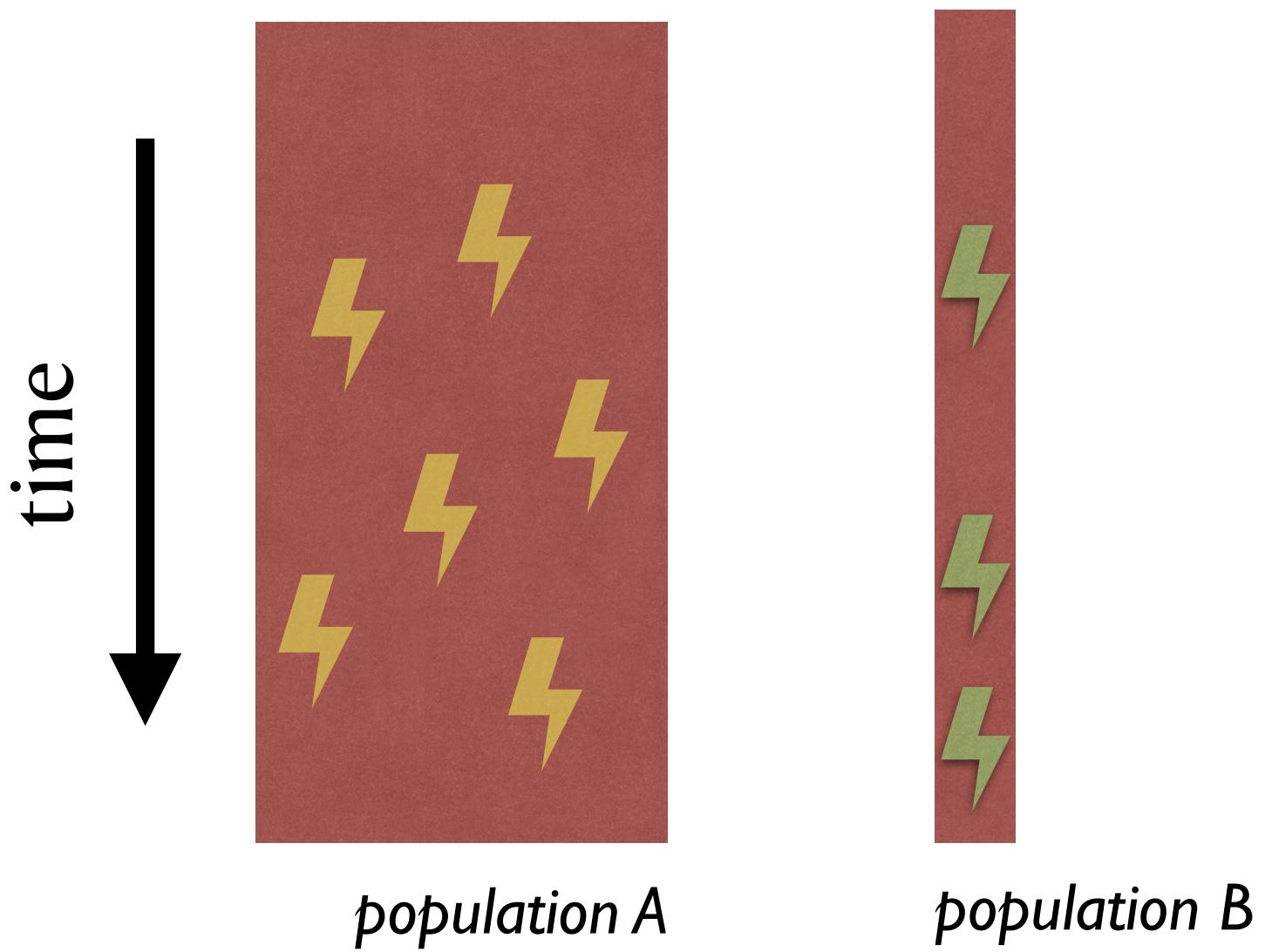
**Tomoko  
Ohta**



<https://womentyoushouldknow.net/population-genetics-tomoko-ohta/>

# Nearly-neutral theory of molecular evolution

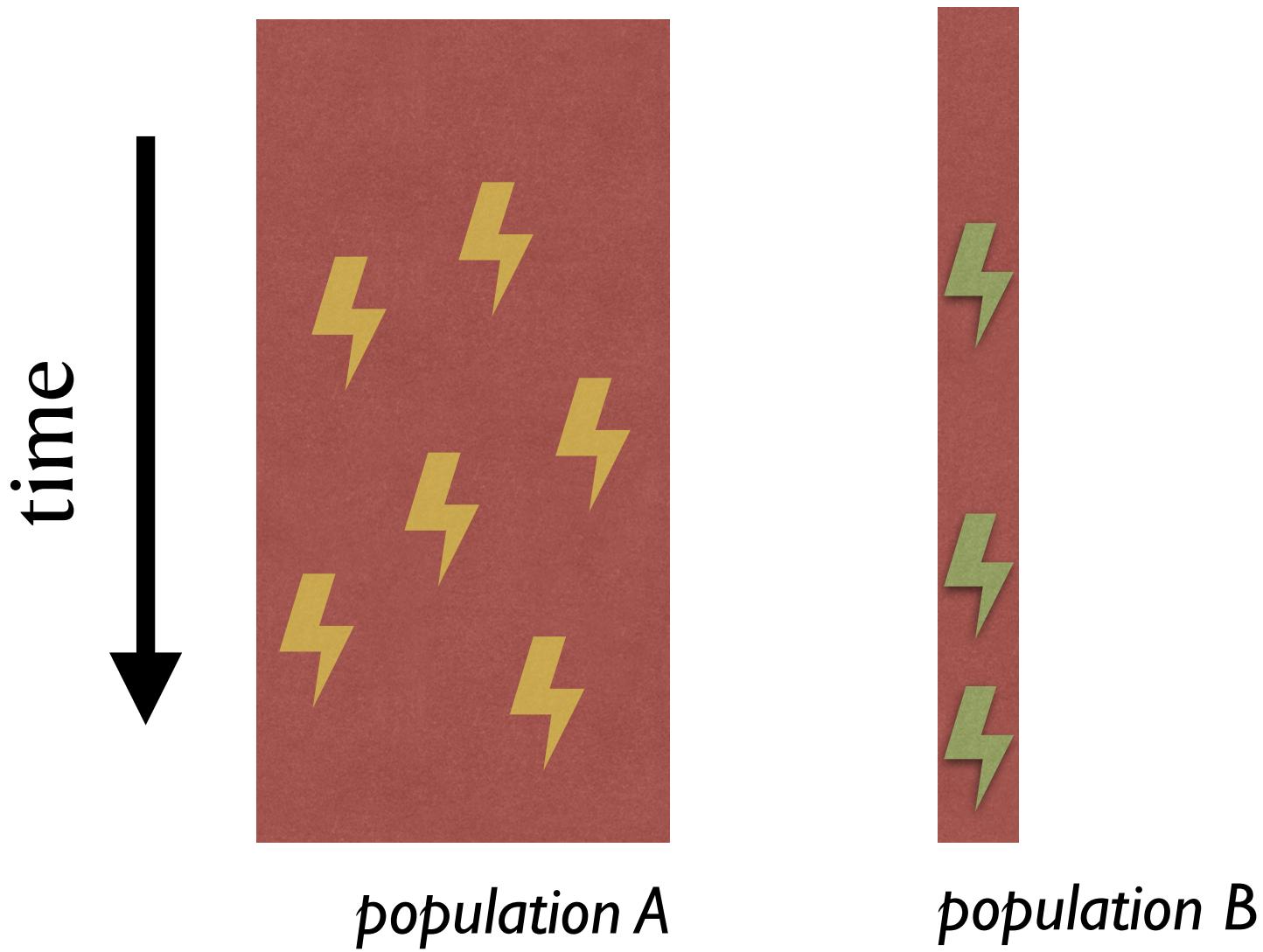
**Tomoko  
Ohta**



We know that genomes accumulate mutations

# Nearly-neutral theory of molecular evolution

**Tomoko  
Ohta**



**We know that genomes accumulate mutations**

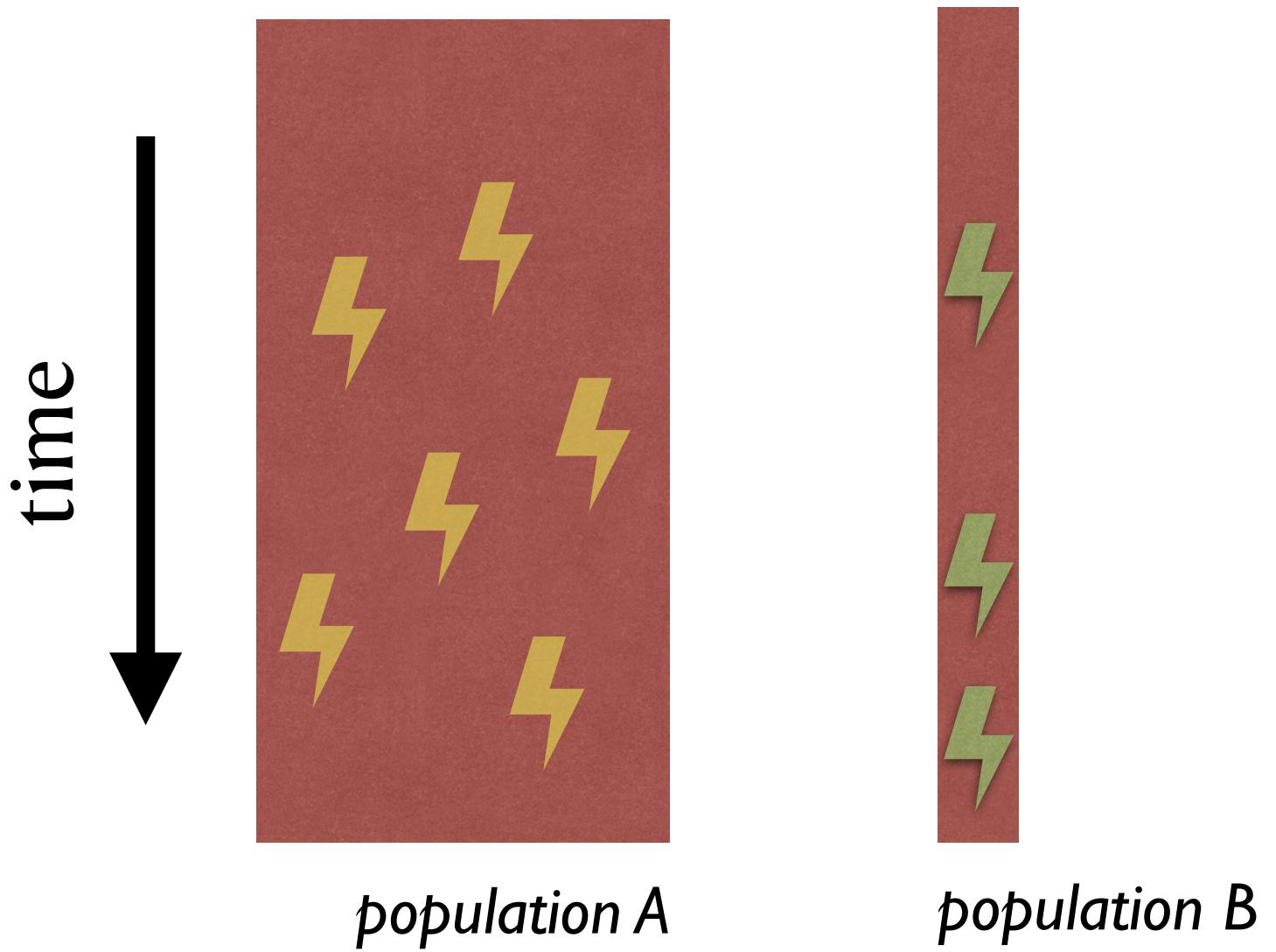
- a tiny number are extremely negative (**lethal**)
- a tiny number are positive (**adaptive**)
- a large number does nothing (**neutral**)
- many are “mildly deleterious” (**nearly neutral**)

# Nearly-neutral theory of molecular evolution

**Tomoko  
Ohta**



<https://womentyoushouldknow.net/population-genetics-tomoko-ohta/>



**We know that genomes accumulate mutations**

- a tiny number are extremely negative (**lethal**)
- a tiny number are positive (**adaptive**)
- a large number does nothing (**neutral**)
- many are “mildly deleterious” (**nearly neutral**)

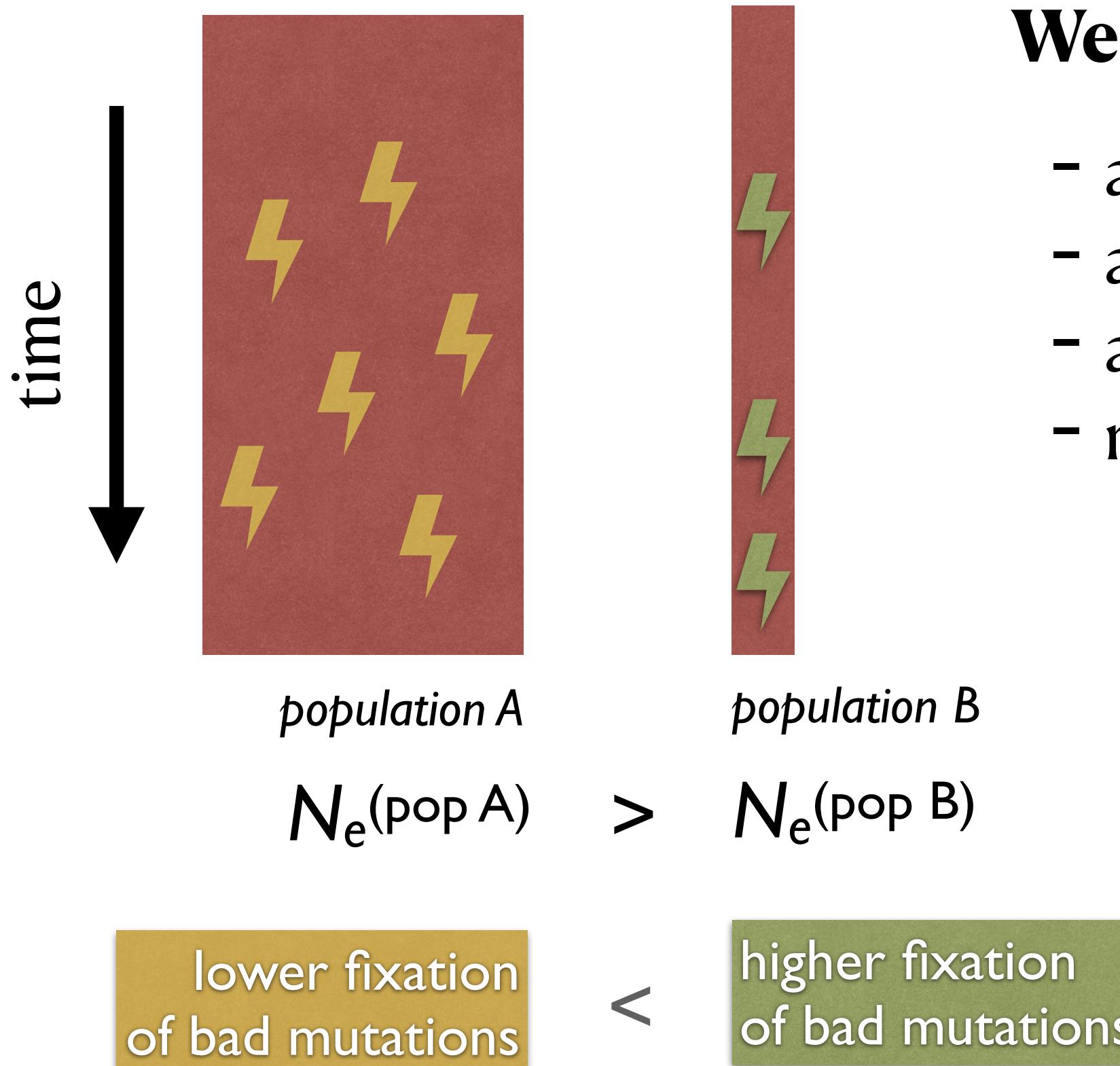
**Efficacy of negative selection  
depends on the effective  
population size ( $N_e$ ).**

# Nearly-neutral theory of molecular evolution

**Tomoko  
Ohta**



<https://womentyoushouldknow.net/population-genetics-tomoko-ohta/>



**We know that genomes accumulate mutations**

- a tiny number are extremely negative (**lethal**)
- a tiny number are positive (**adaptive**)
- a large number does nothing (**neutral**)
- many are “mildly deleterious” (**nearly neutral**)

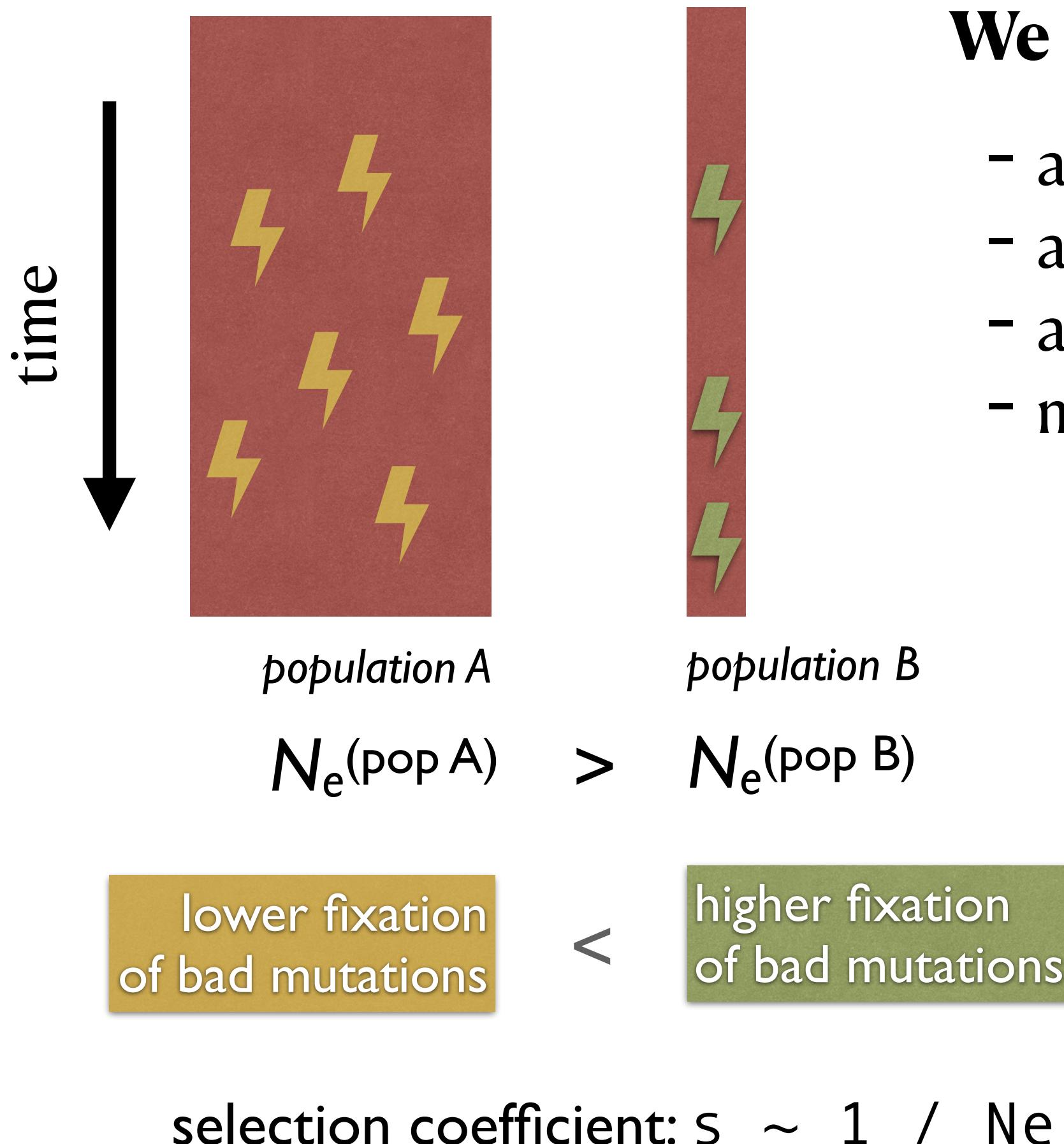
**Efficacy of negative selection  
depends on the effective  
population size ( $N_e$ ).**

# Nearly-neutral theory of molecular evolution

**Tomoko  
Ohta**



<https://womentyoushouldknow.net/population-genetics-tomoko-ohta/>



**We know that genomes accumulate mutations**

- a tiny number are extremely negative (**lethal**)
- a tiny number are positive (**adaptive**)
- a large number does nothing (**neutral**)
- many are “mildly deleterious” (**nearly neutral**)

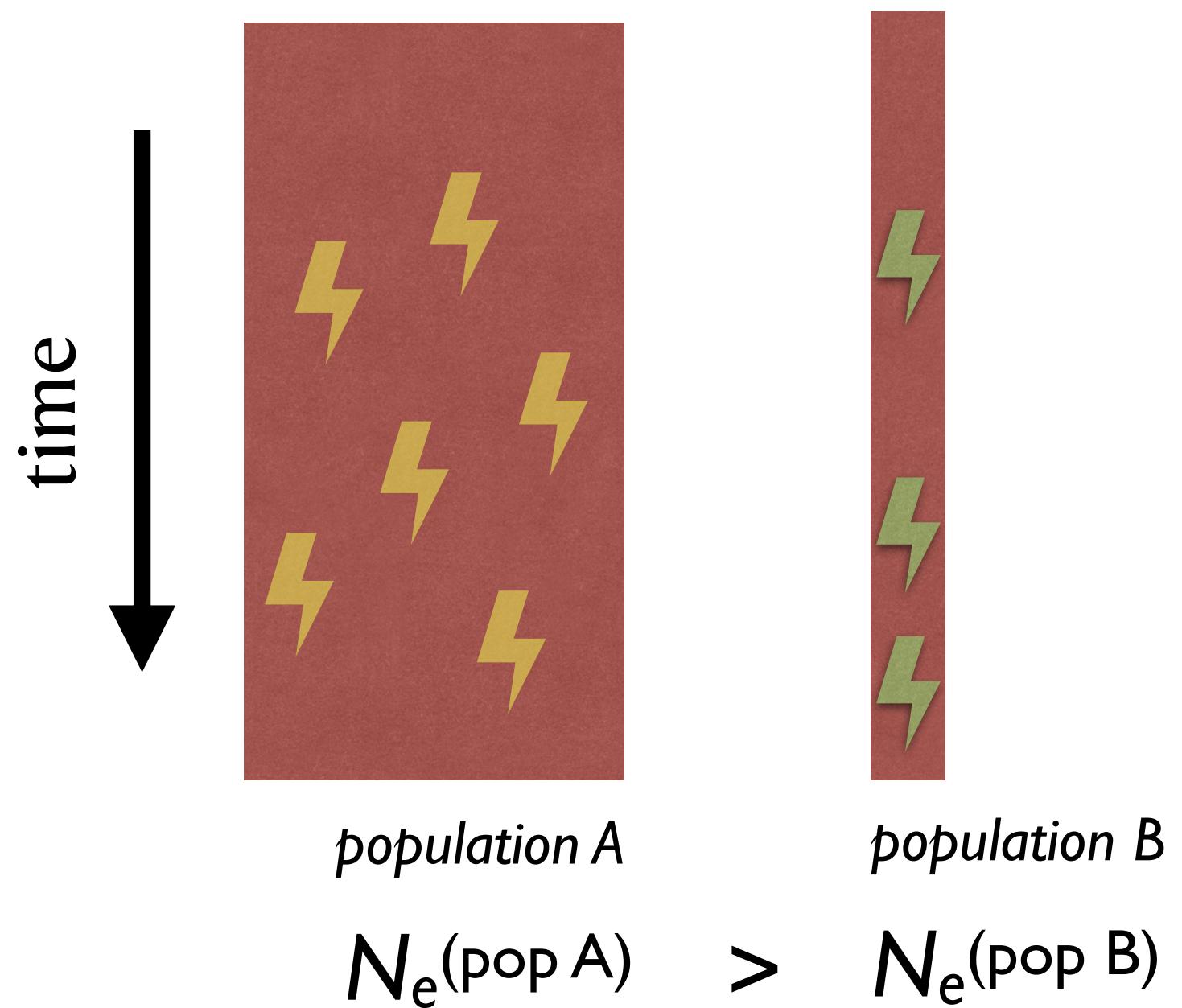
**Efficacy of negative selection depends on the effective population size ( $N_e$ ).**

# Nearly-neutral theory of molecular evolution

**Tomoko  
Ohta**

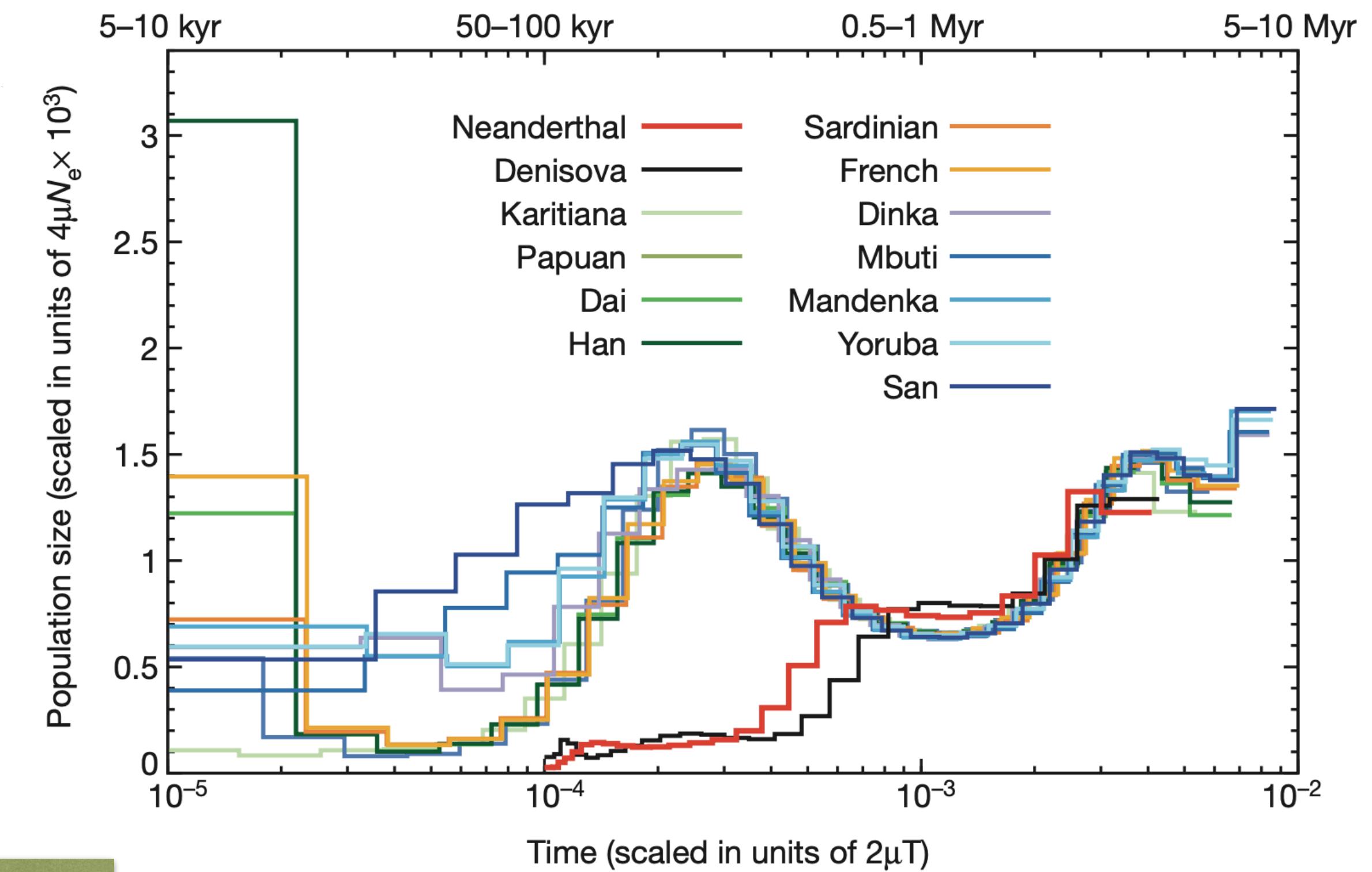


<https://womenthoushouldknow.net/population-genetics-tomoko-ohta/>



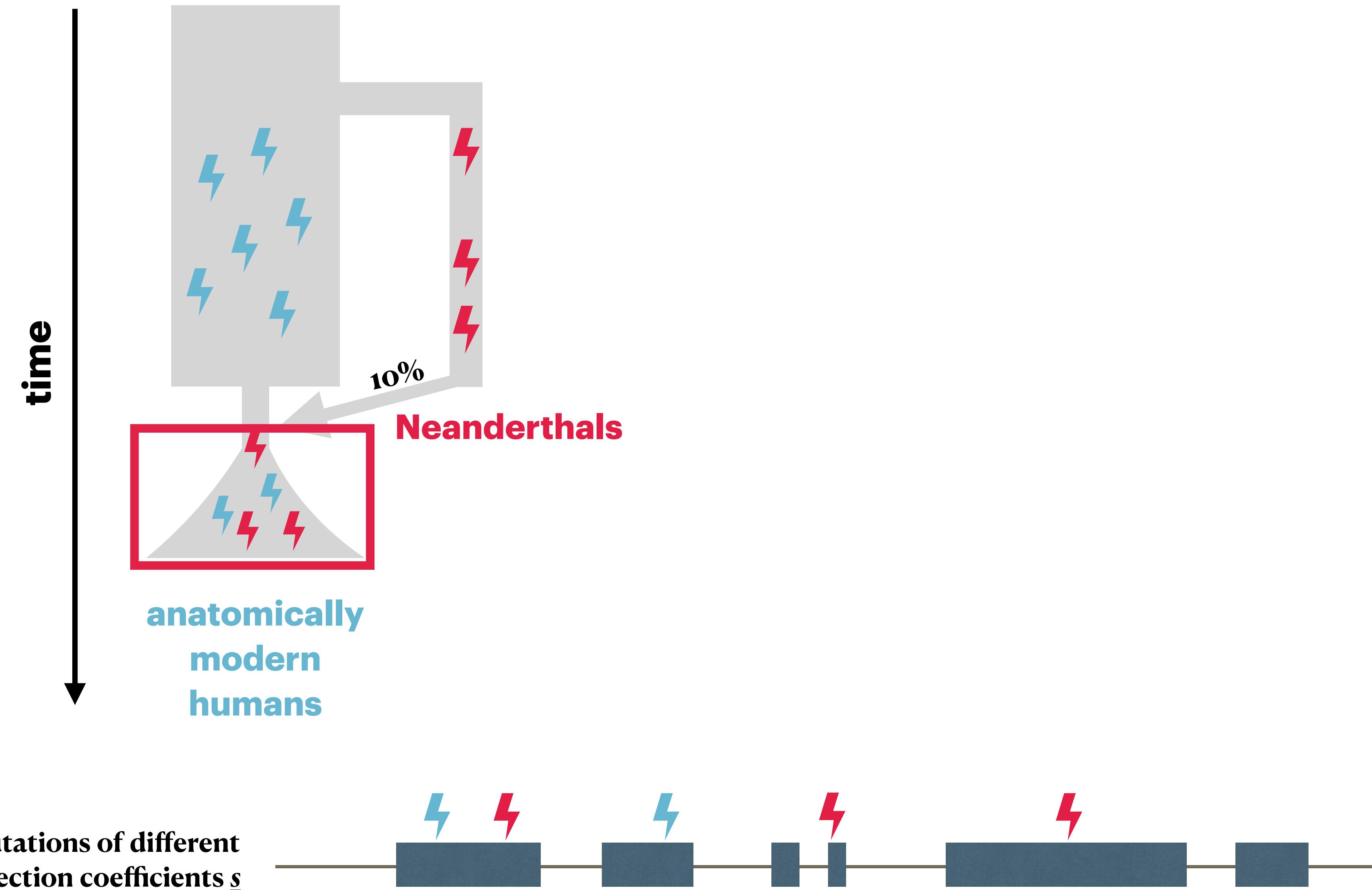
"anatomically  
modern humans"

“Neanderthals”

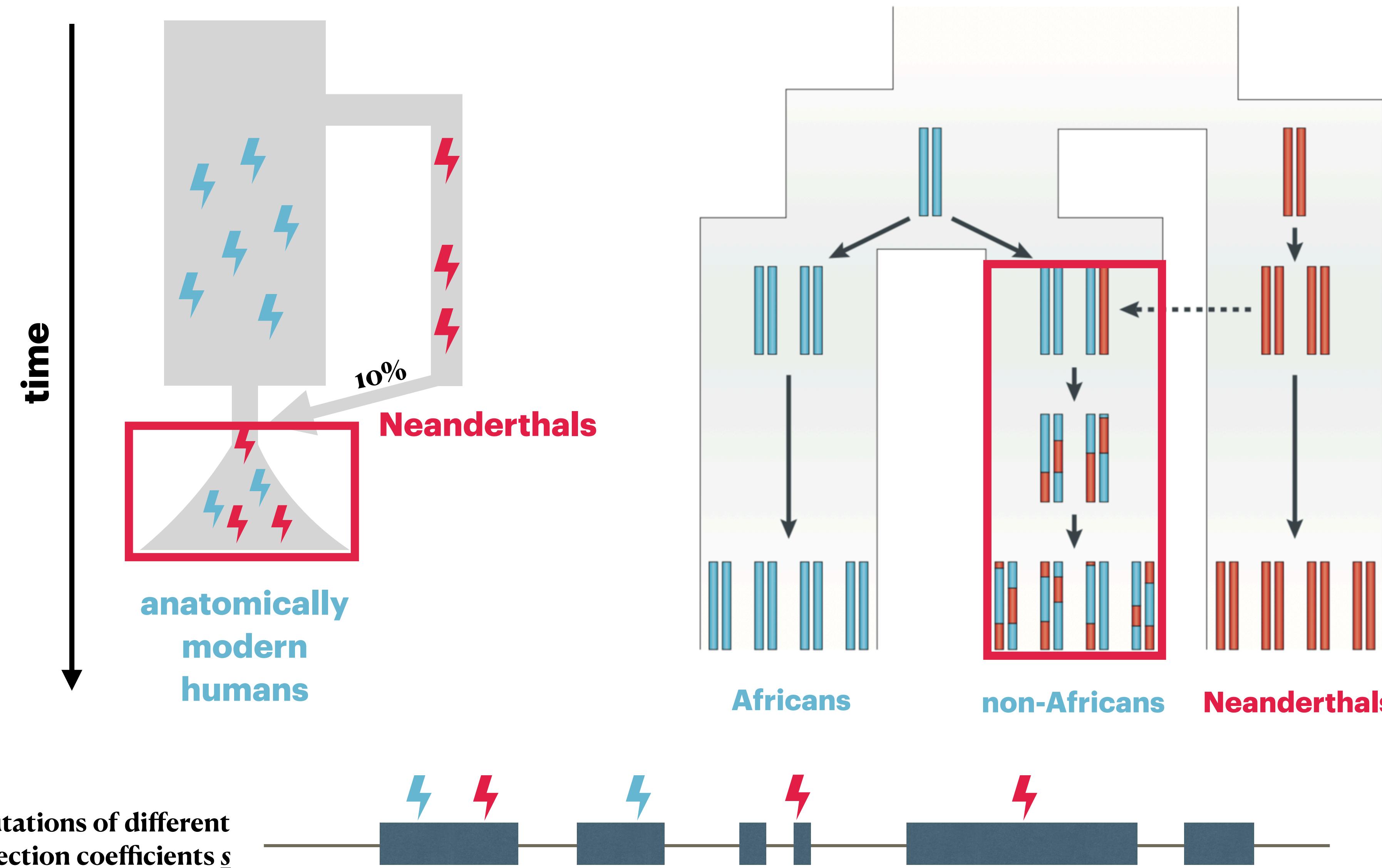


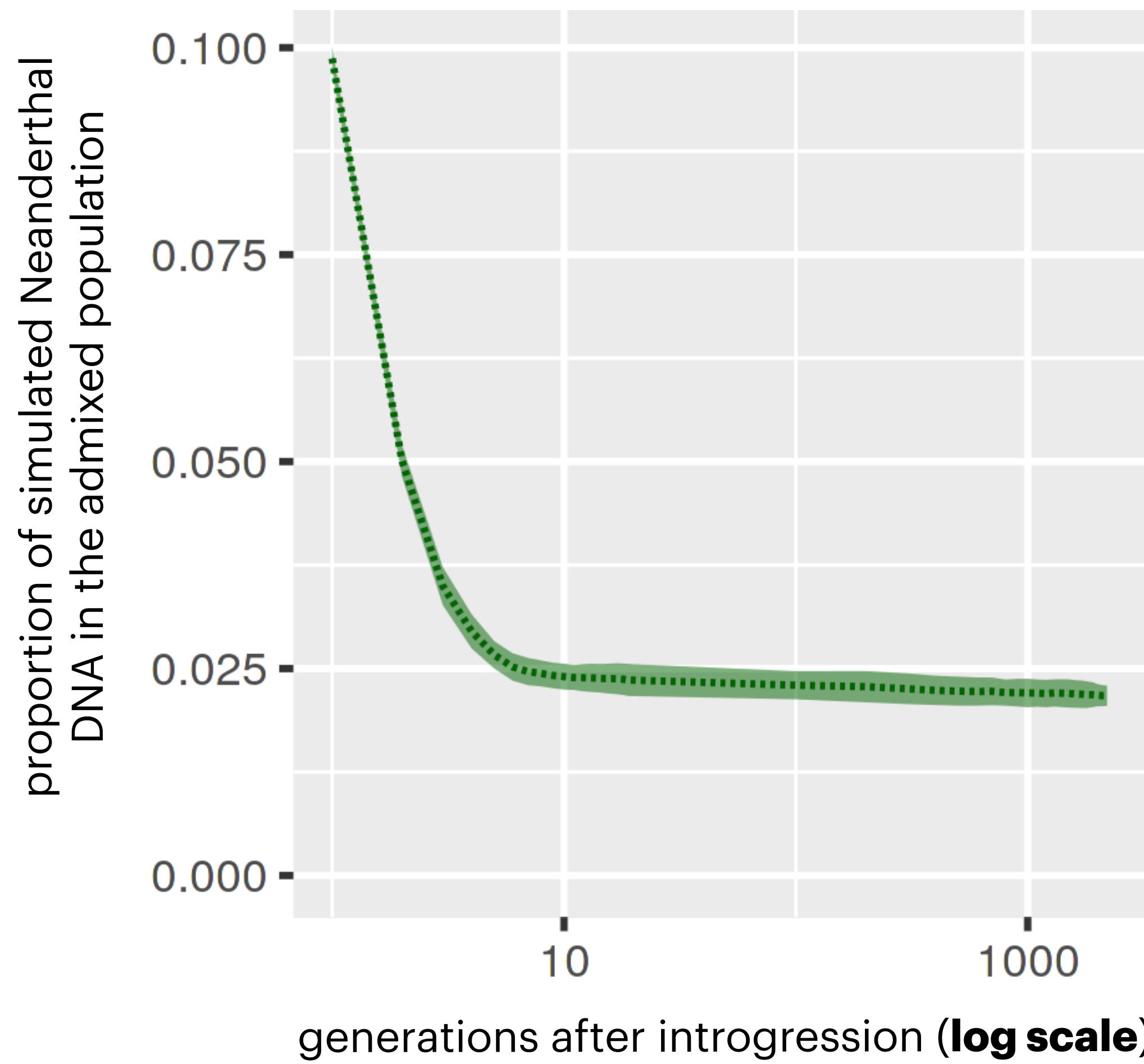
Ohta (Nature, 1973)

# Simulations: tracing the trajectory of introgressed DNA over time

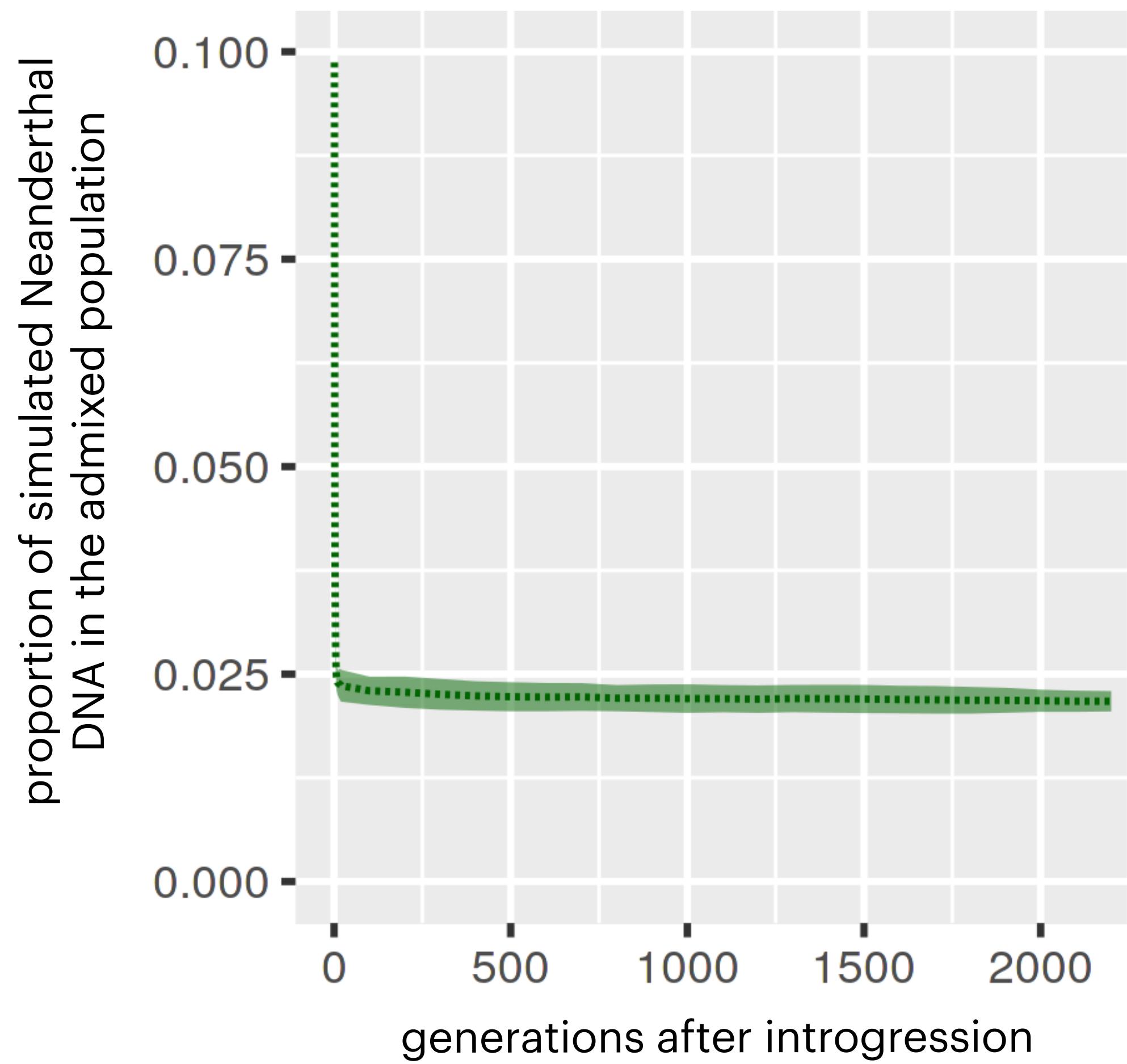


# Simulations: tracing the trajectory of introgressed DNA over time

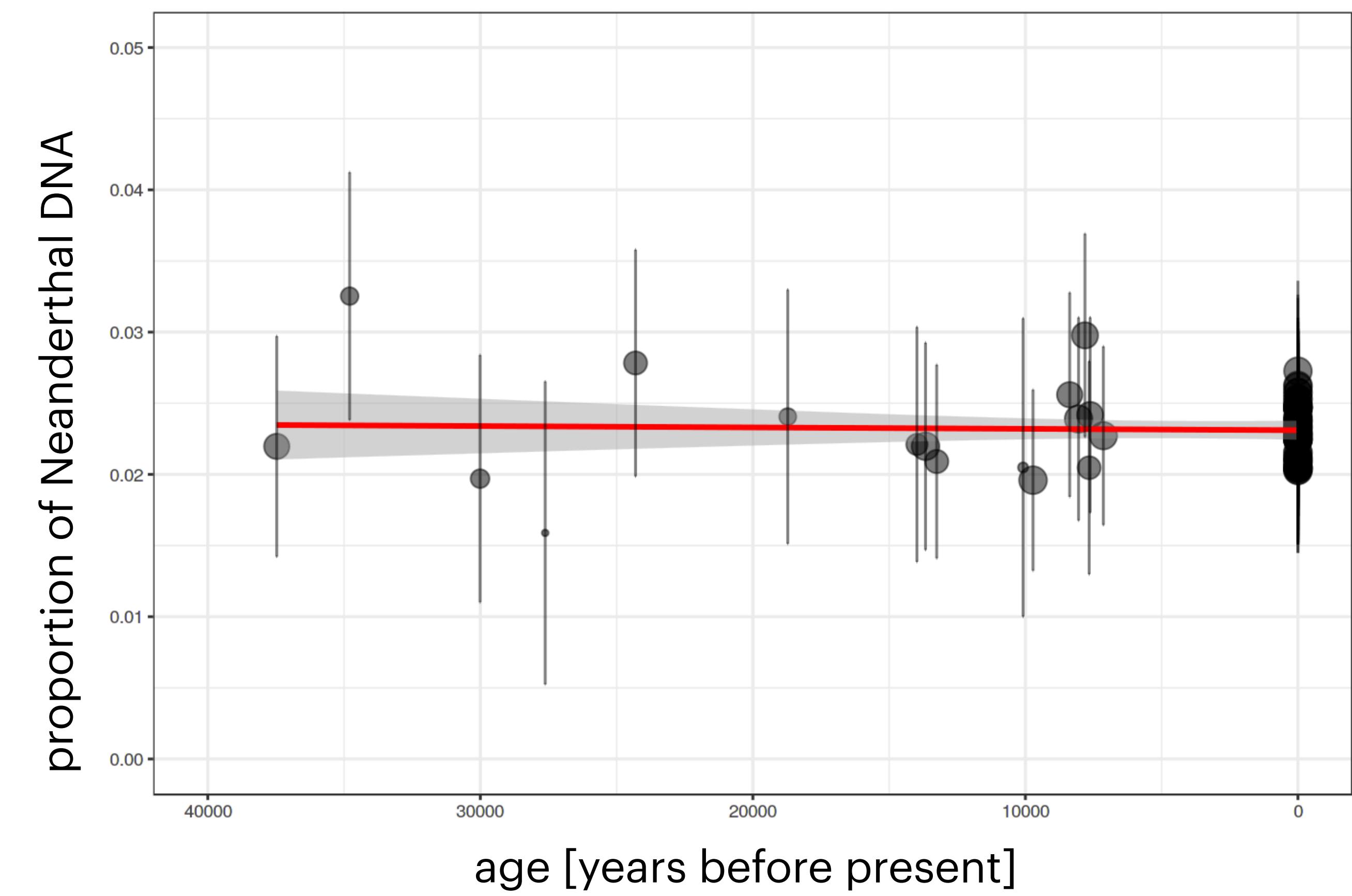
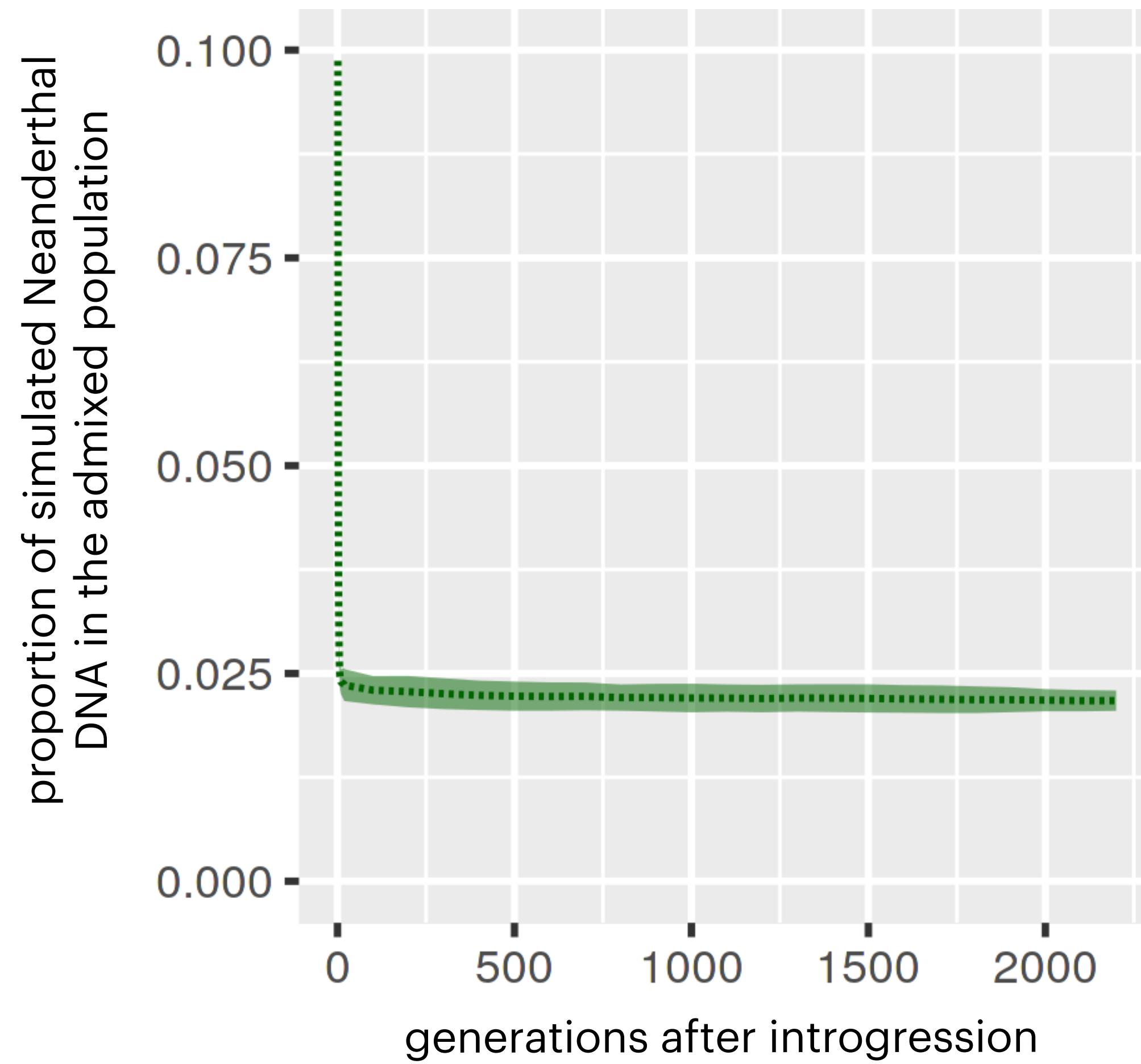




Harris et al. (2016); Juric et al. (2016); Petr et al. (2019)

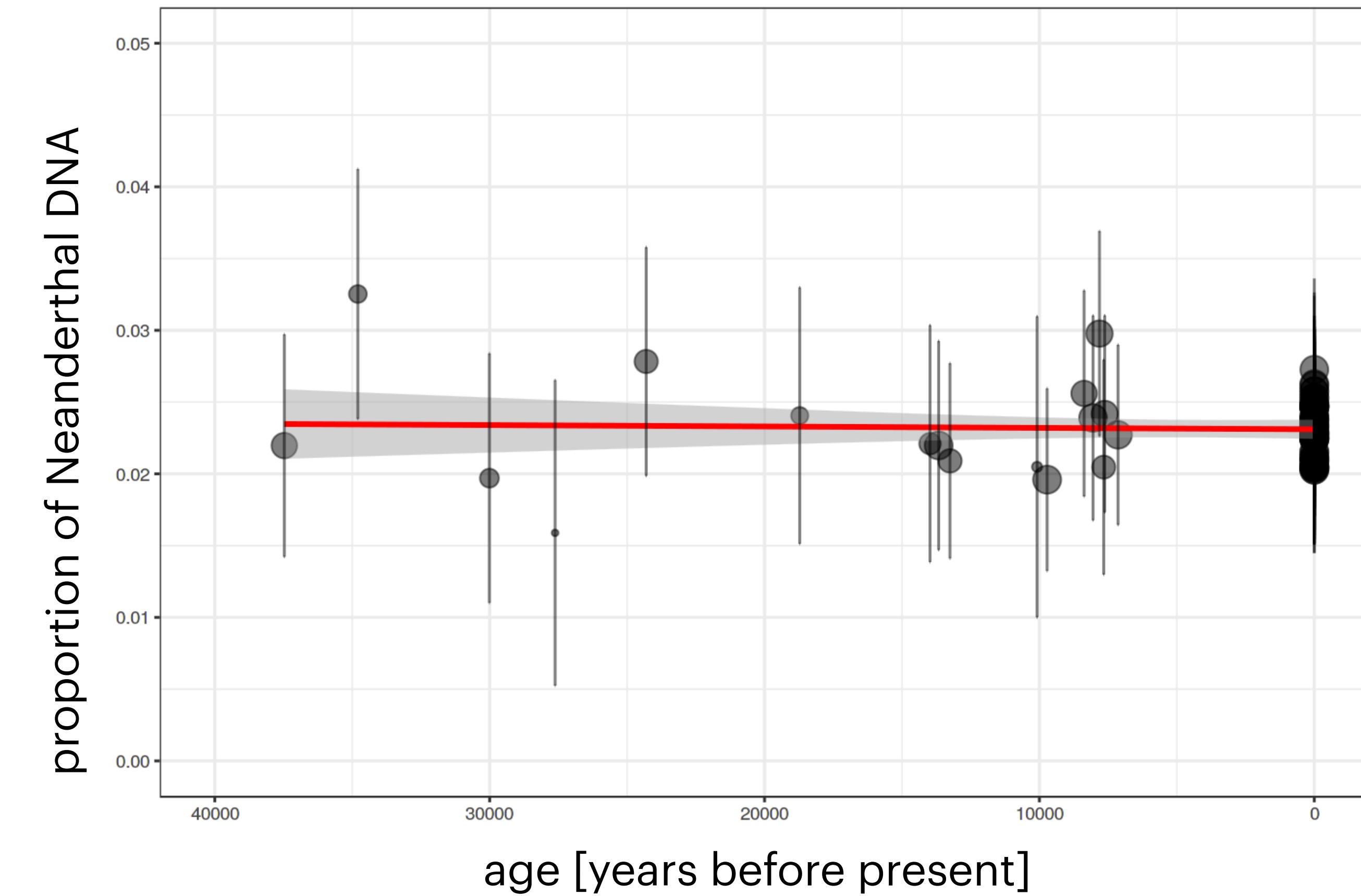
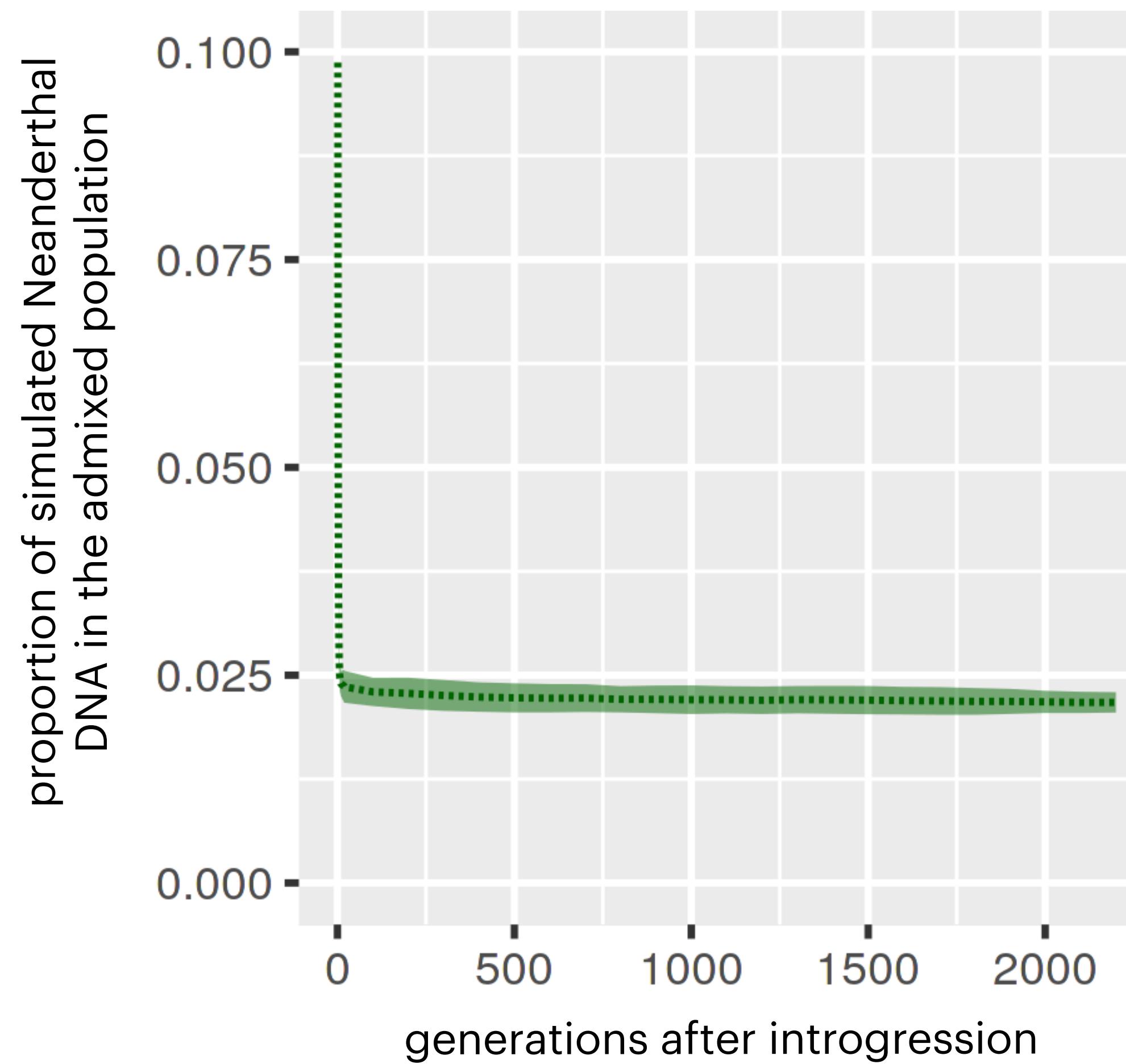


Harris et al. (2016); Juric et al. (2016); Petr et al. (2019)



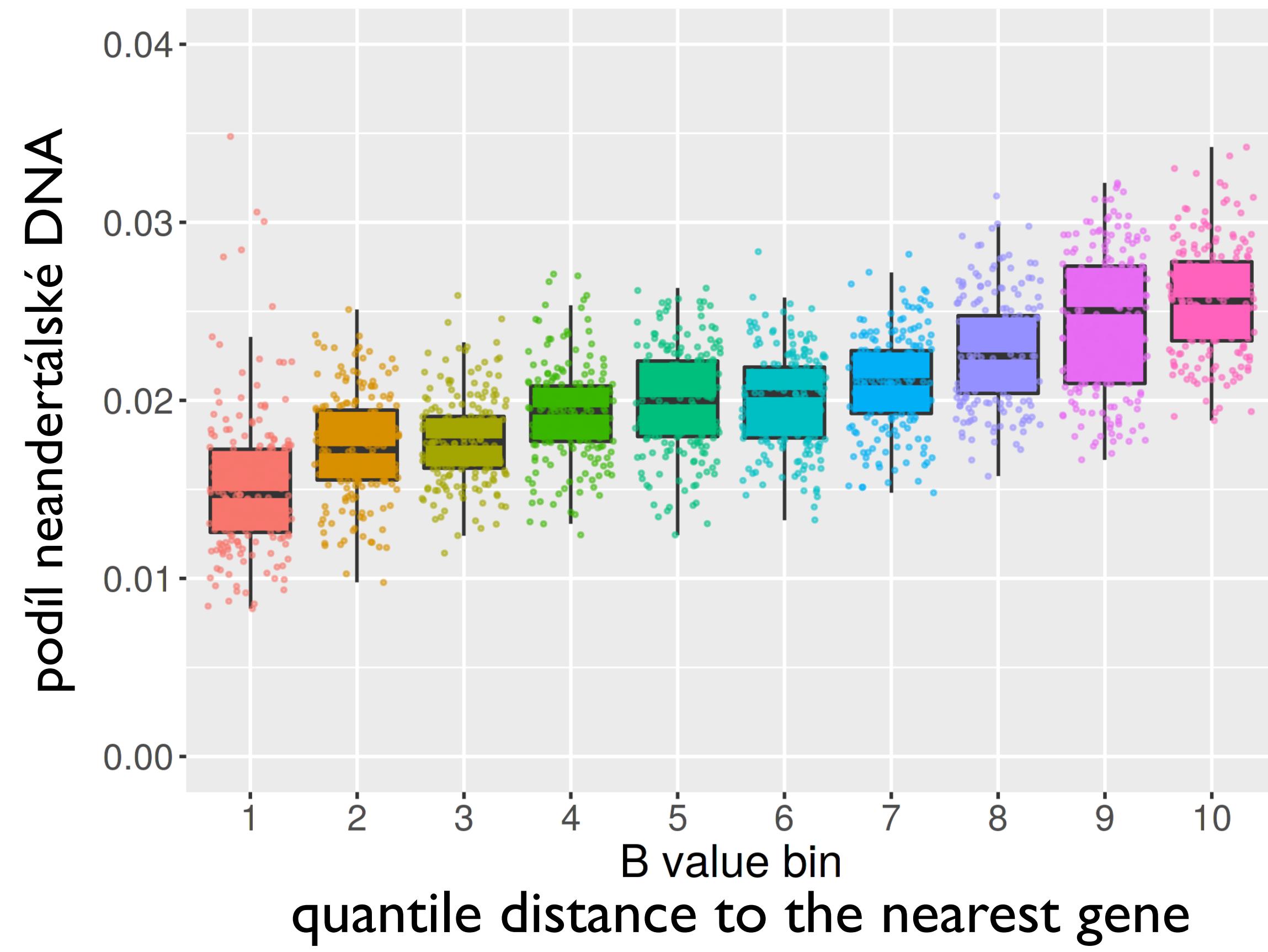
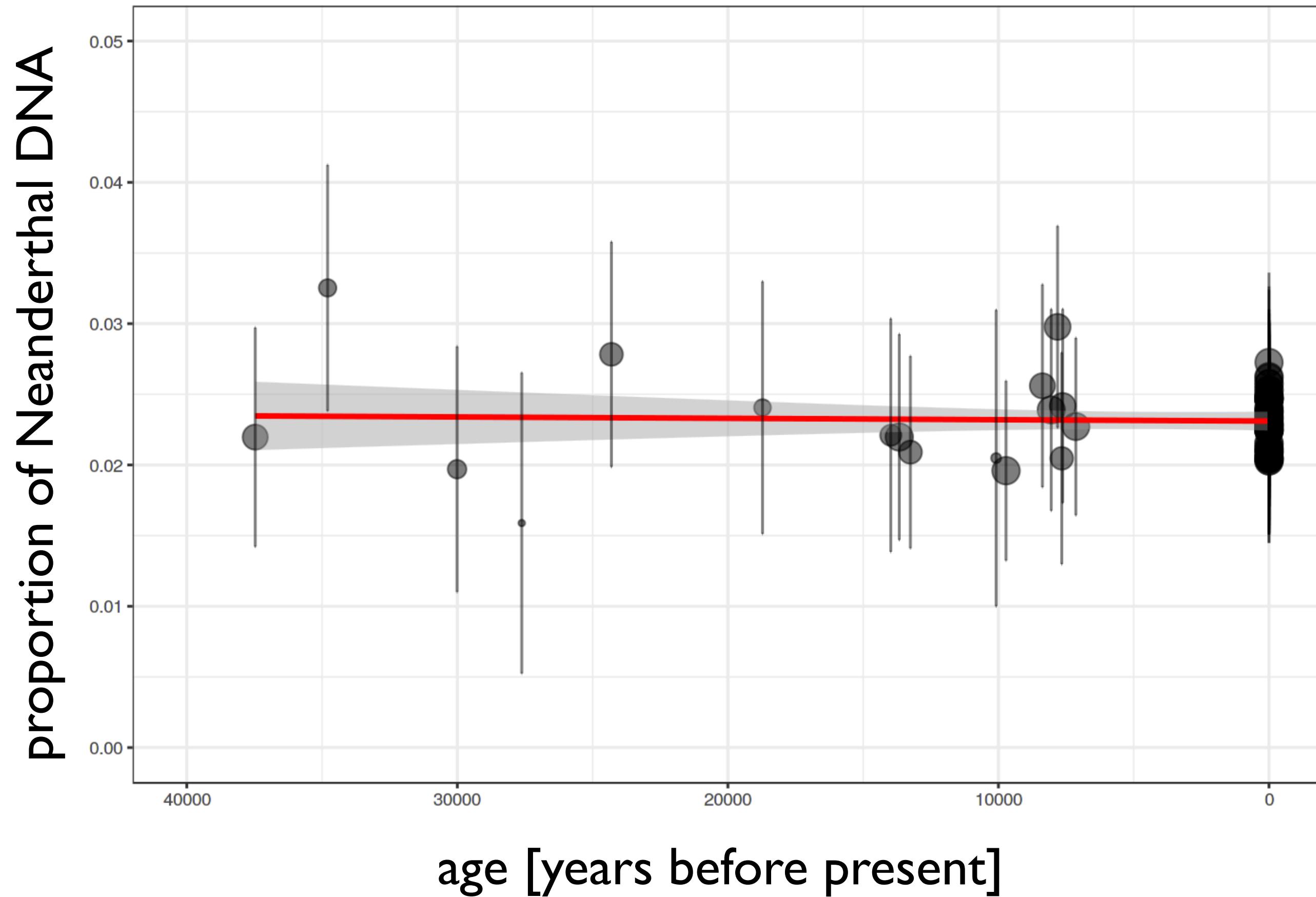
Harris et al. (2016); Juric et al. (2016); Petr et al. (2019)

# Neanderthal ancestry trajectory matches simulations...

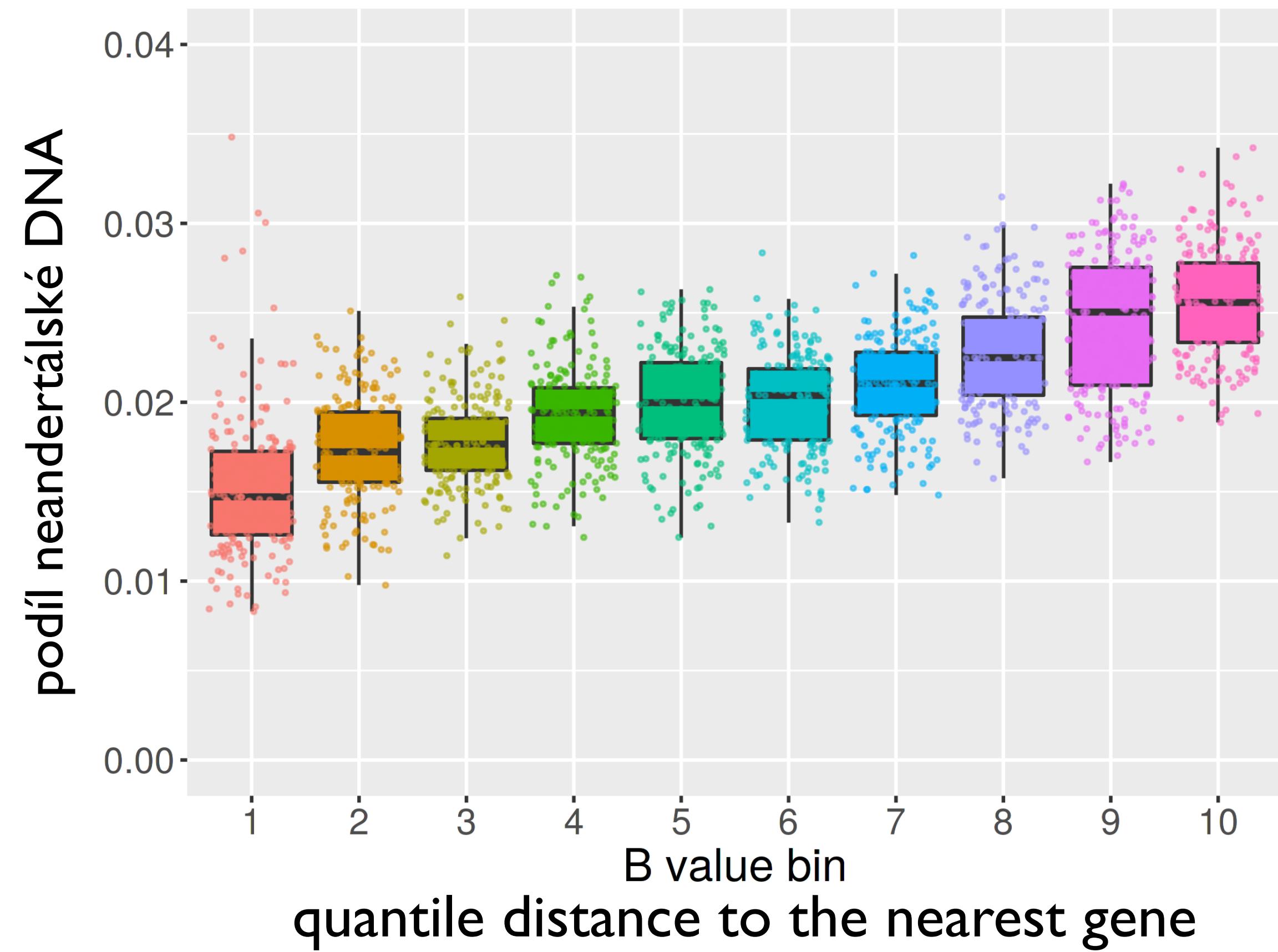
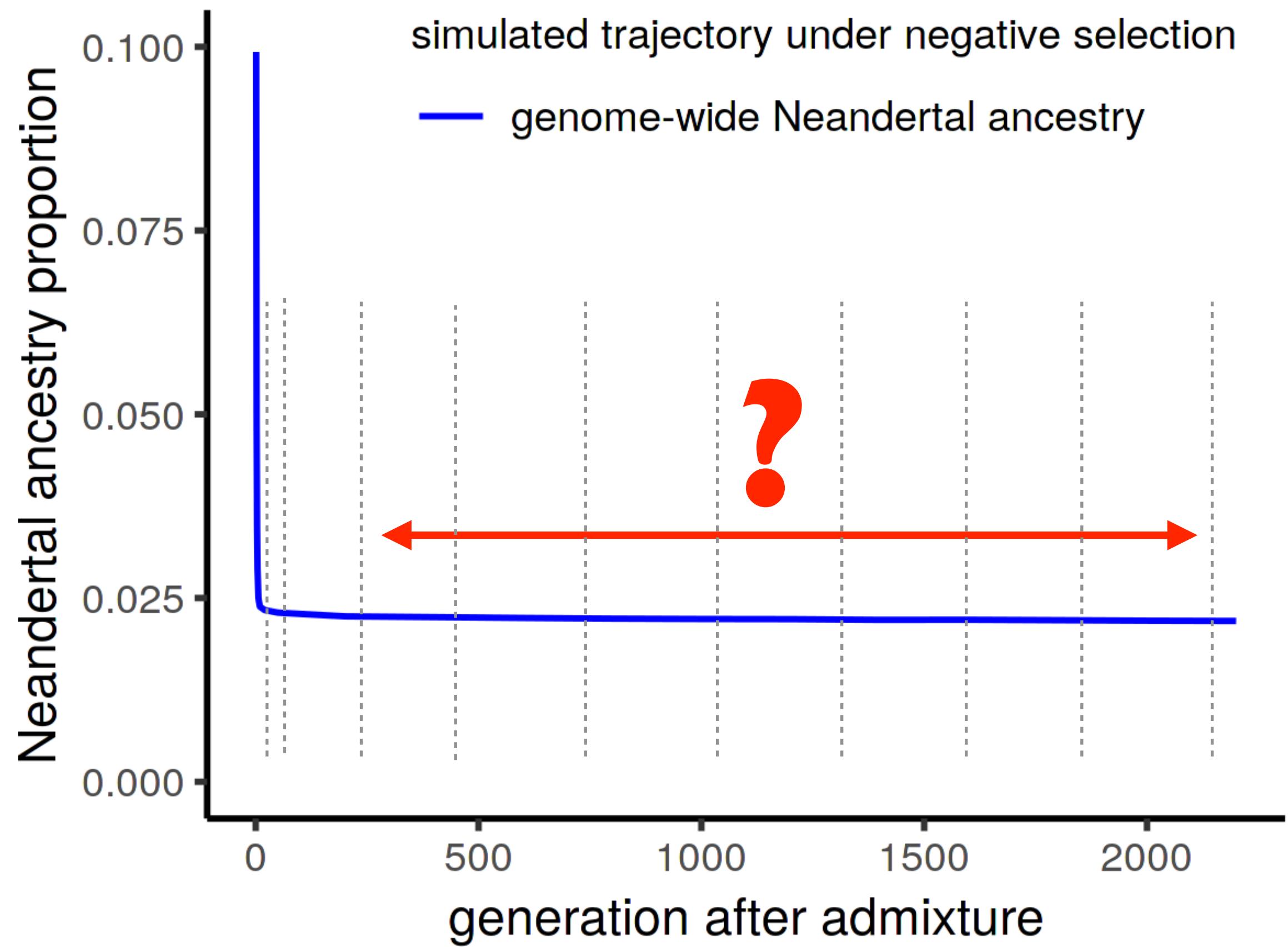


Harris et al. (2016); Juric et al. (2016); Petr et al. (2019)

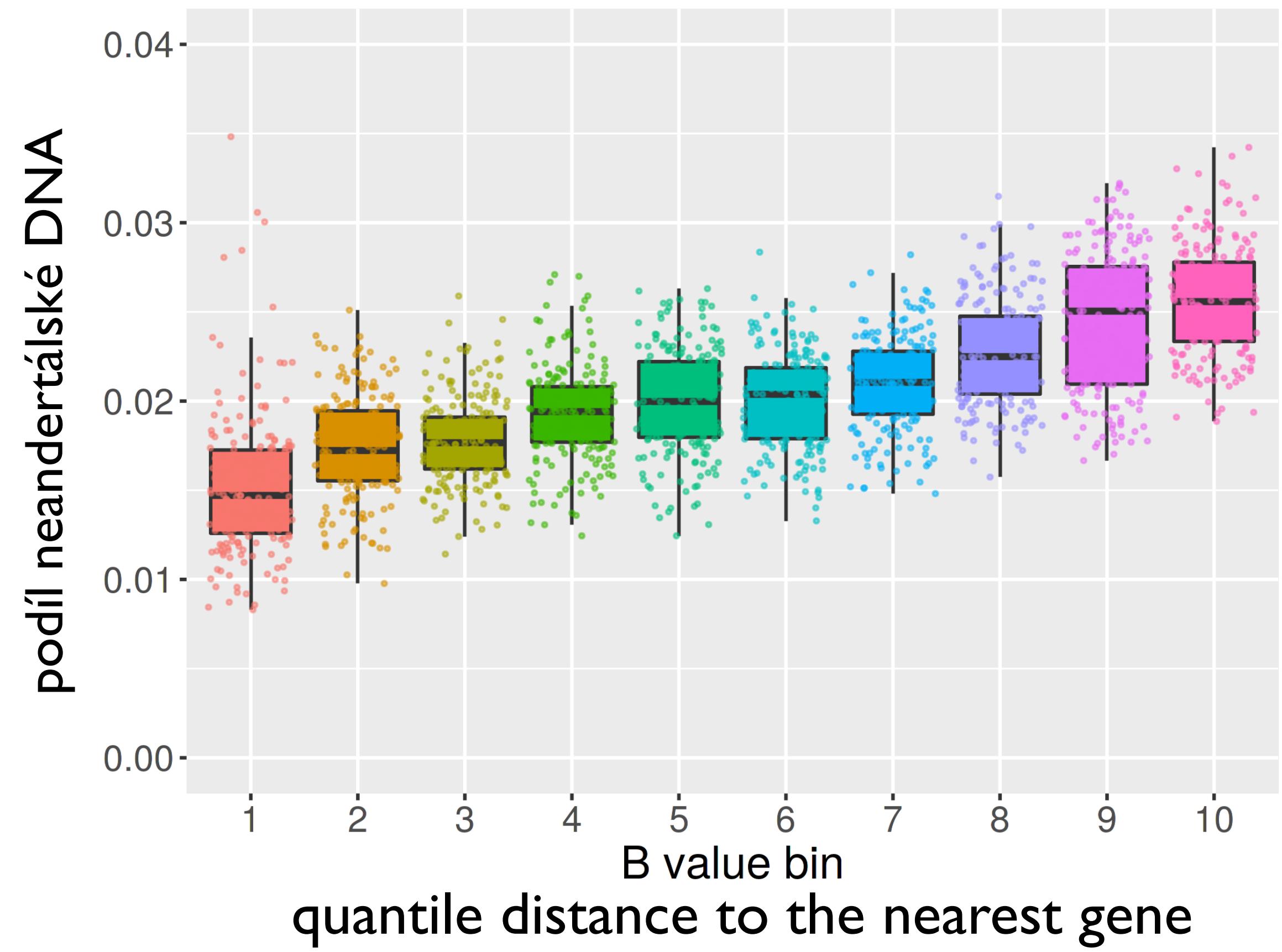
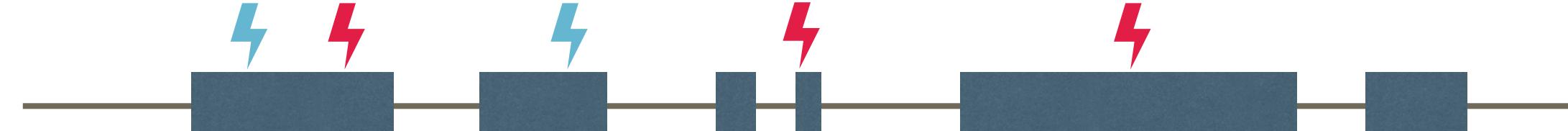
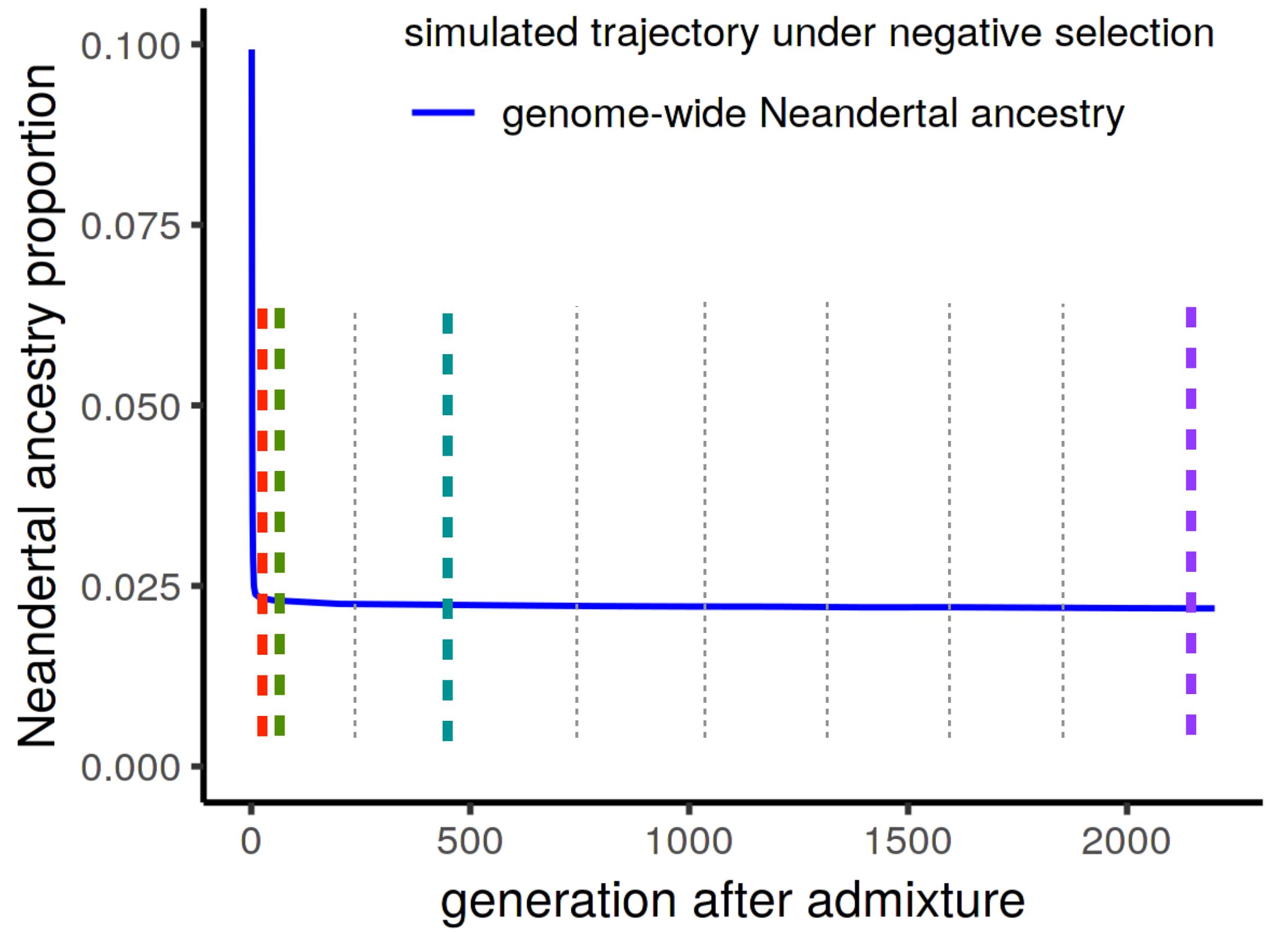
# ... what about the whole-genome negative selection signal?



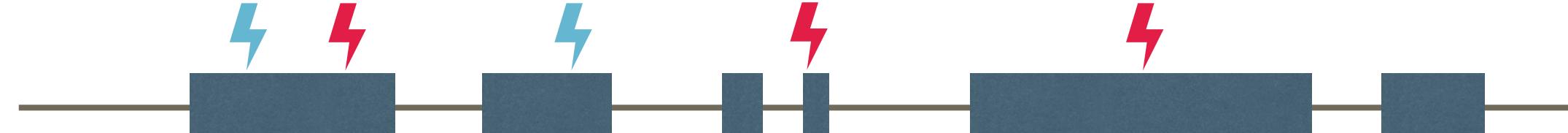
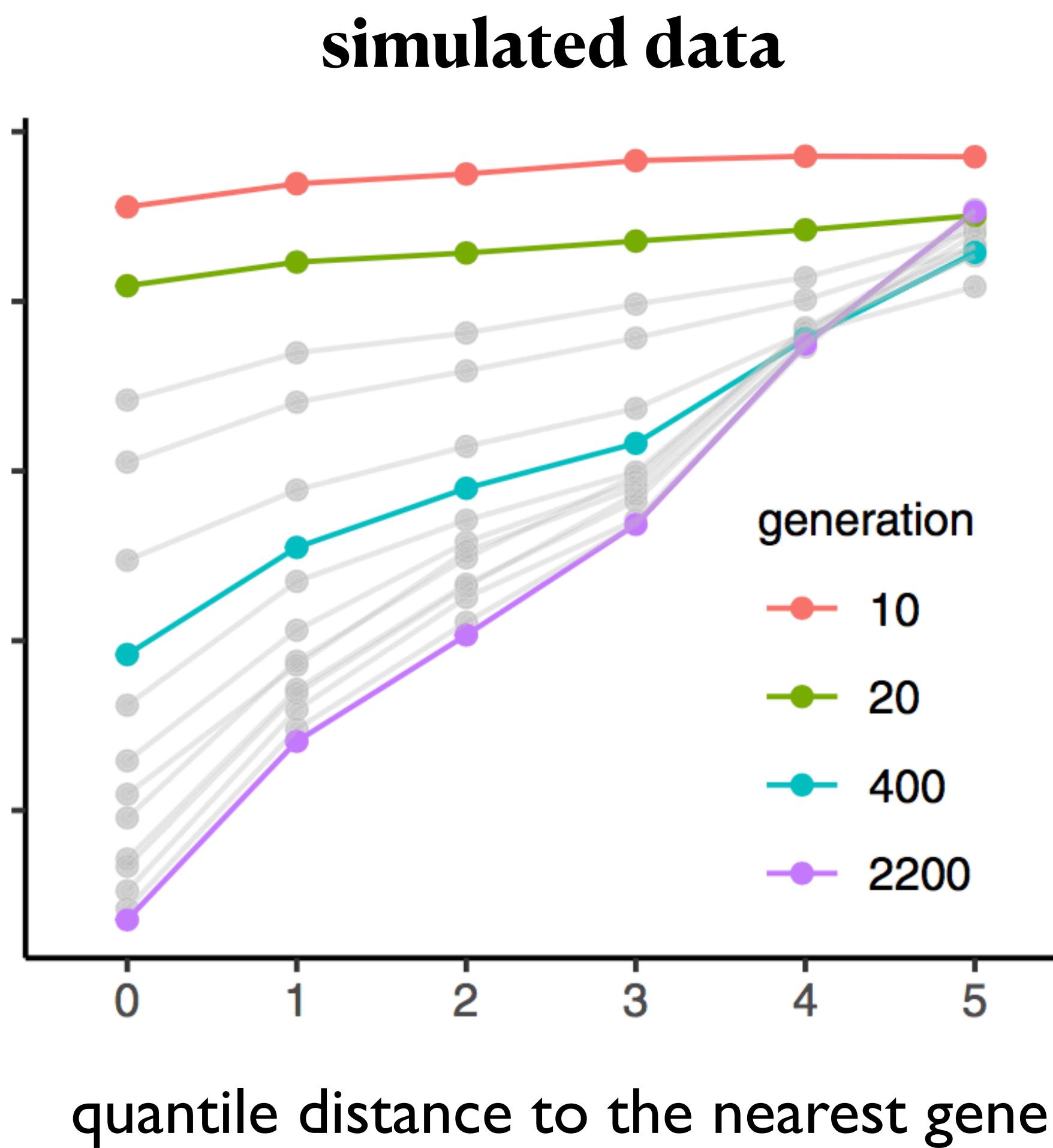
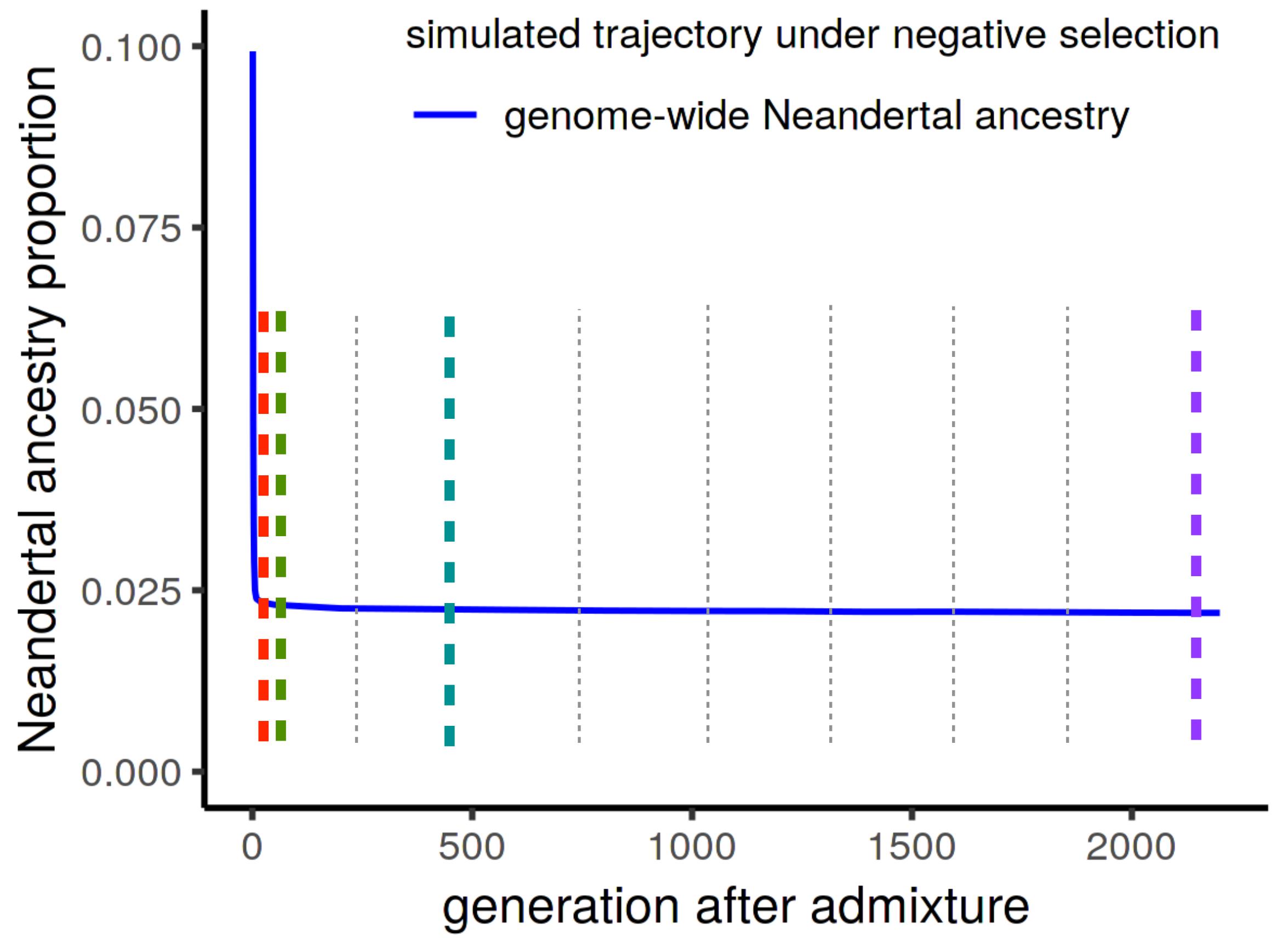
# ... what about the whole-genome negative selection signal?



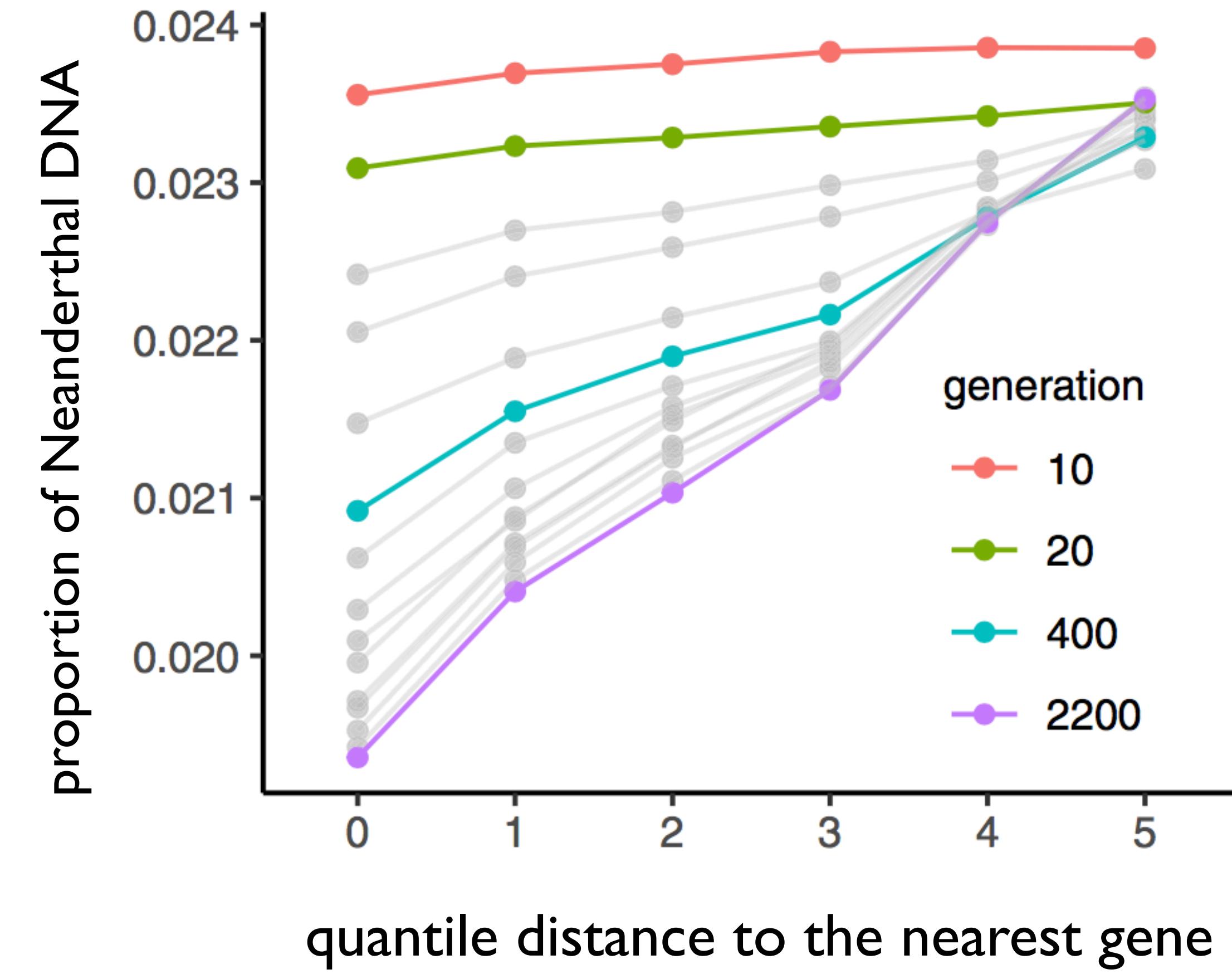
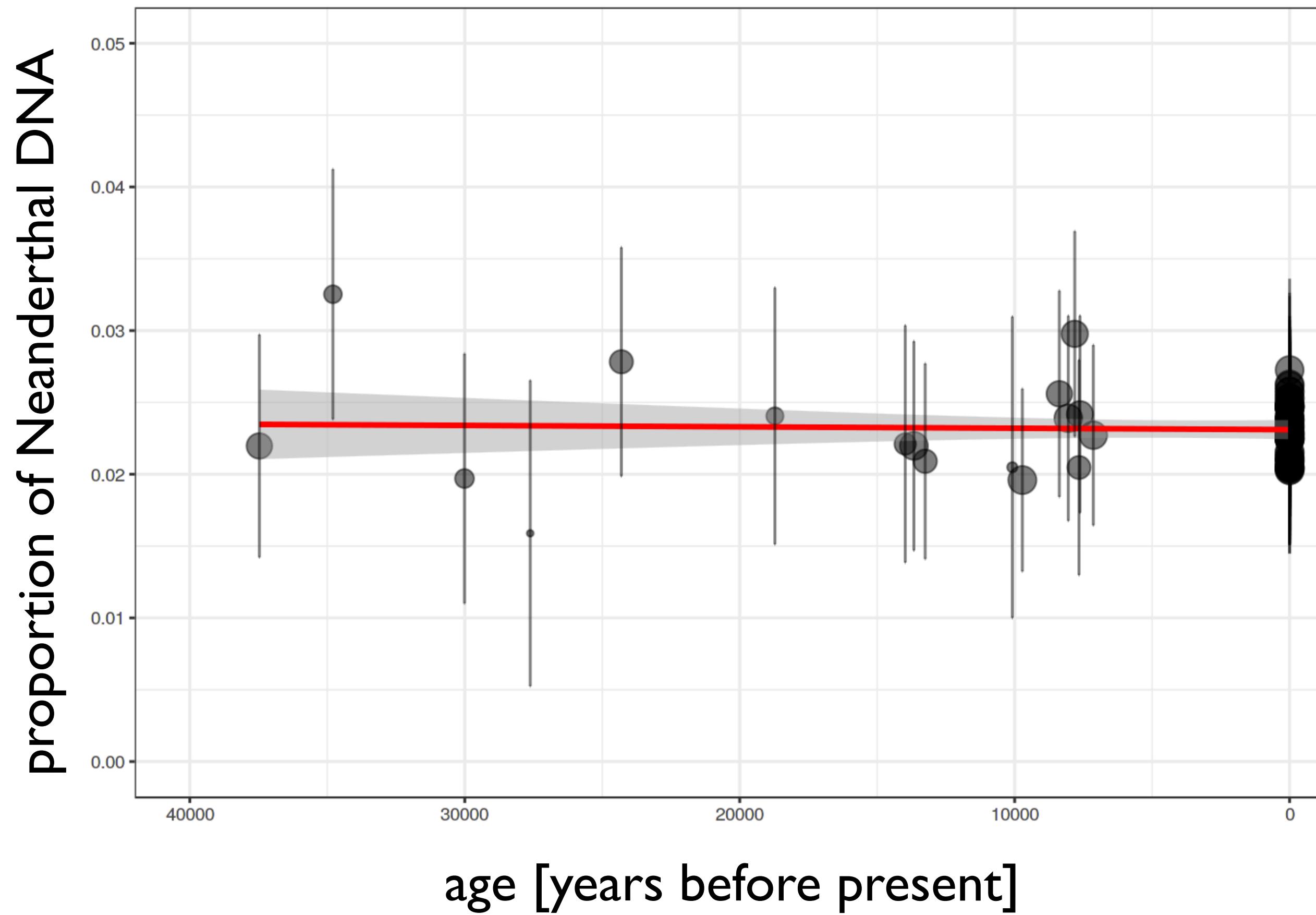
# ... what about the whole-genome negative selection signal?

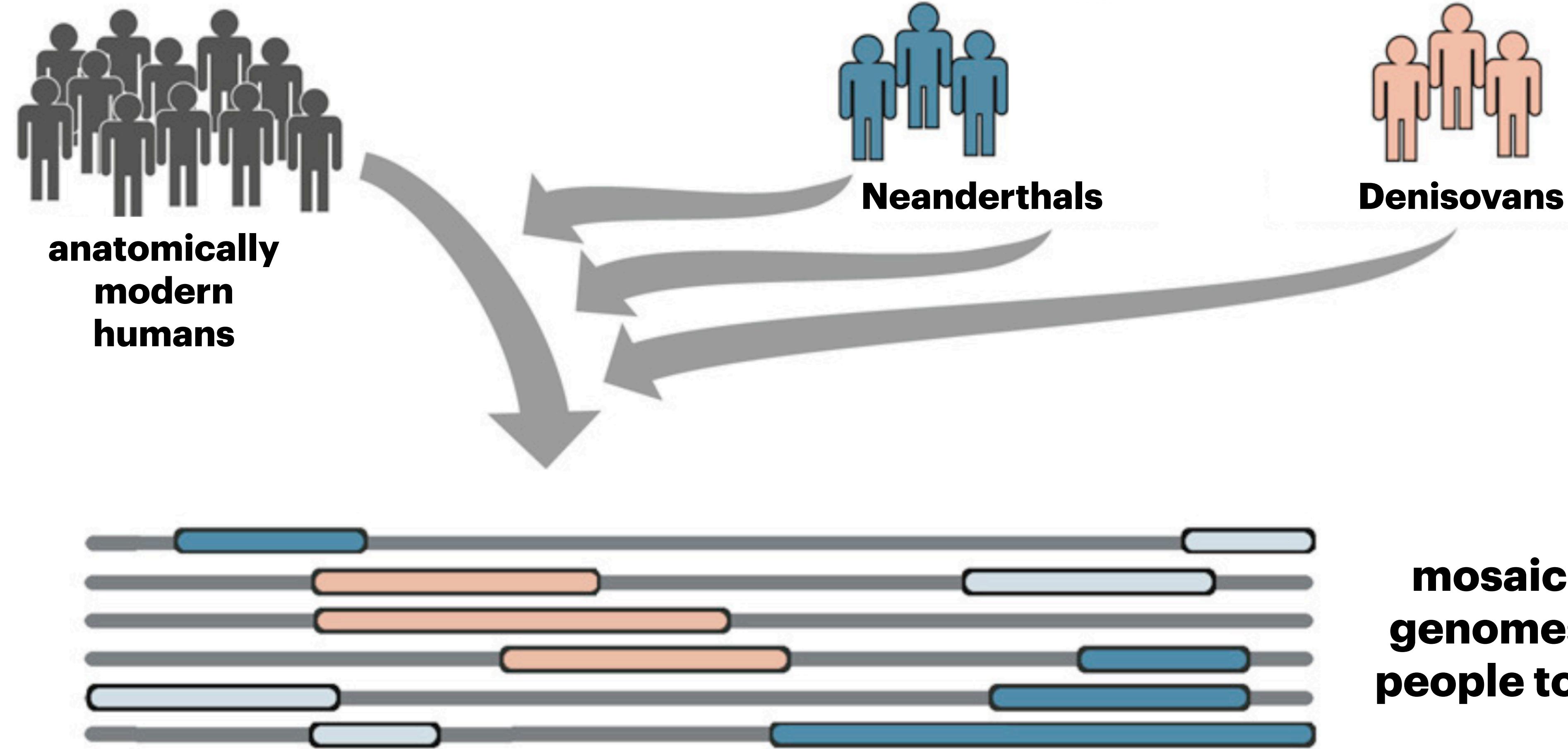


# ... what about the whole-genome negative selection signal?

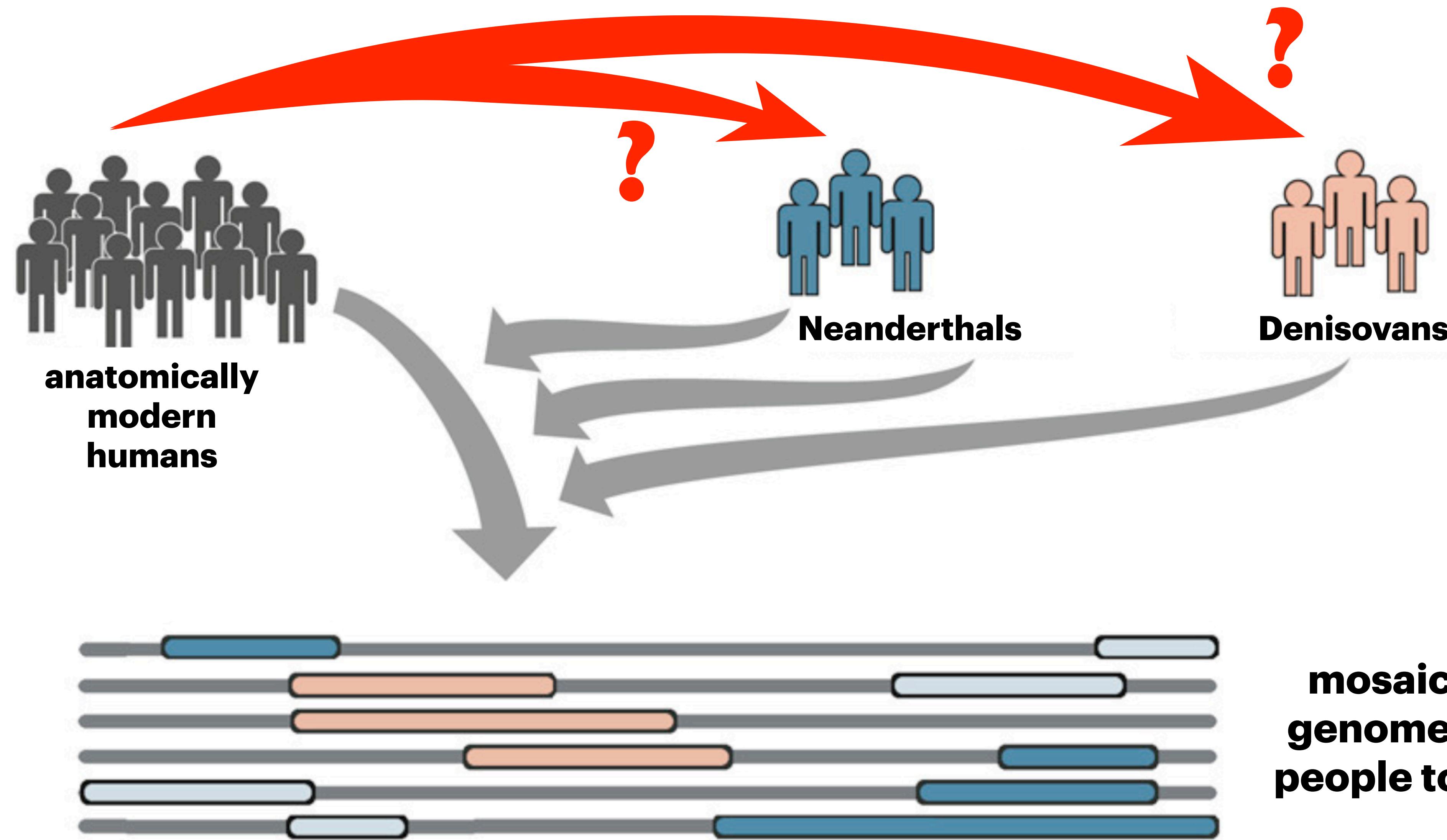


# Both empirical results match population genetic theory!

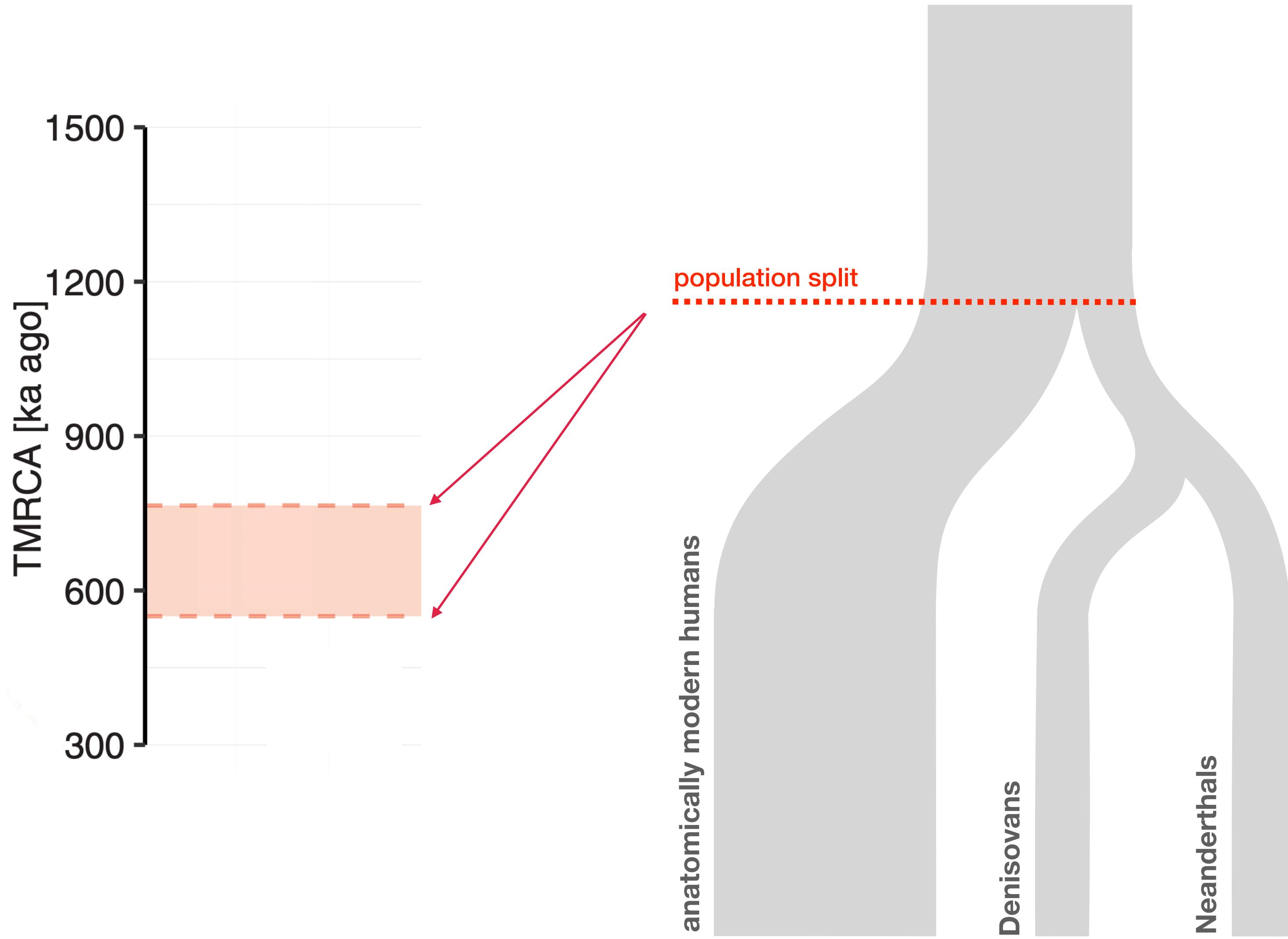




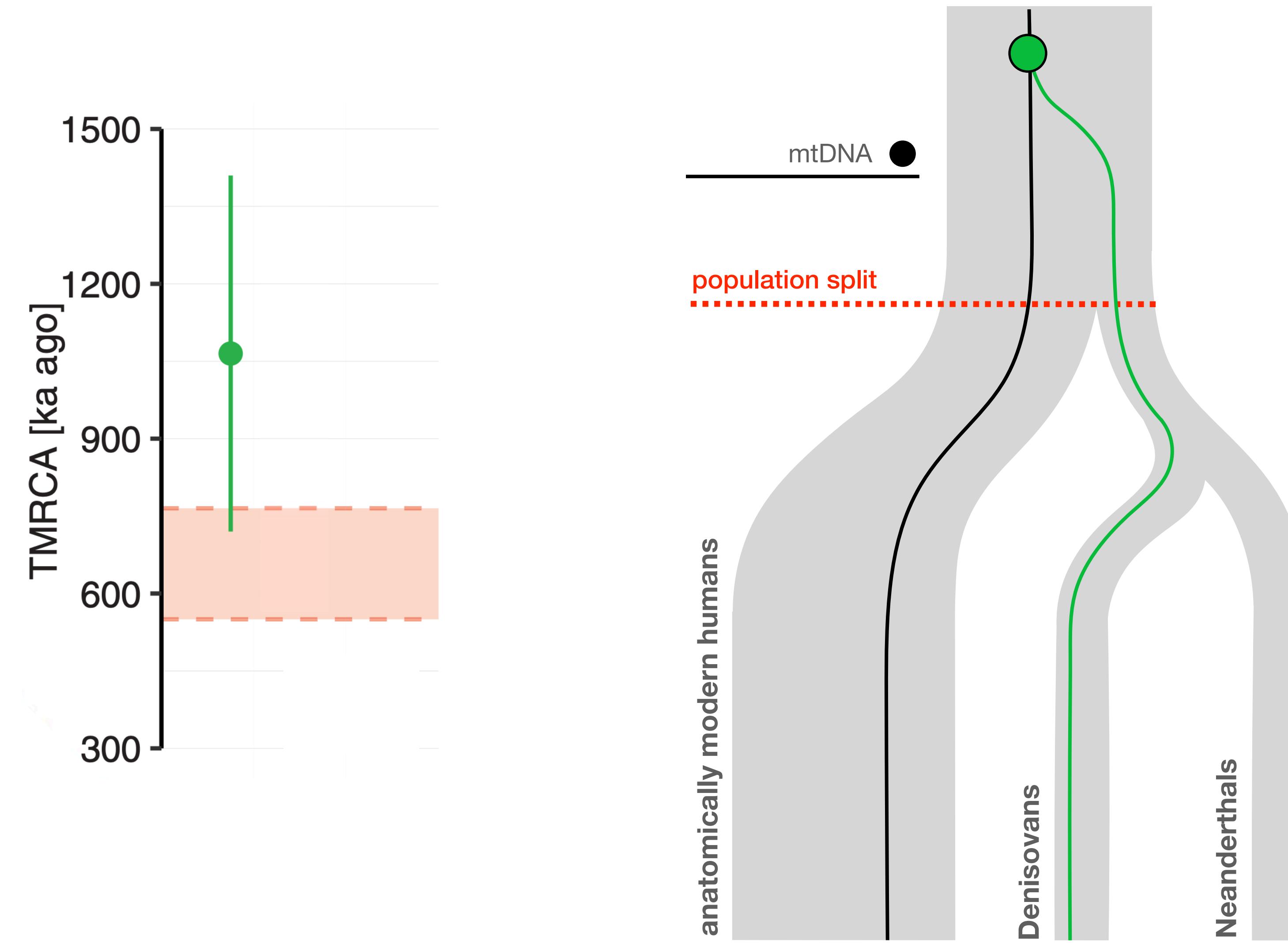
adapted from Browning et al., Cell, 2018



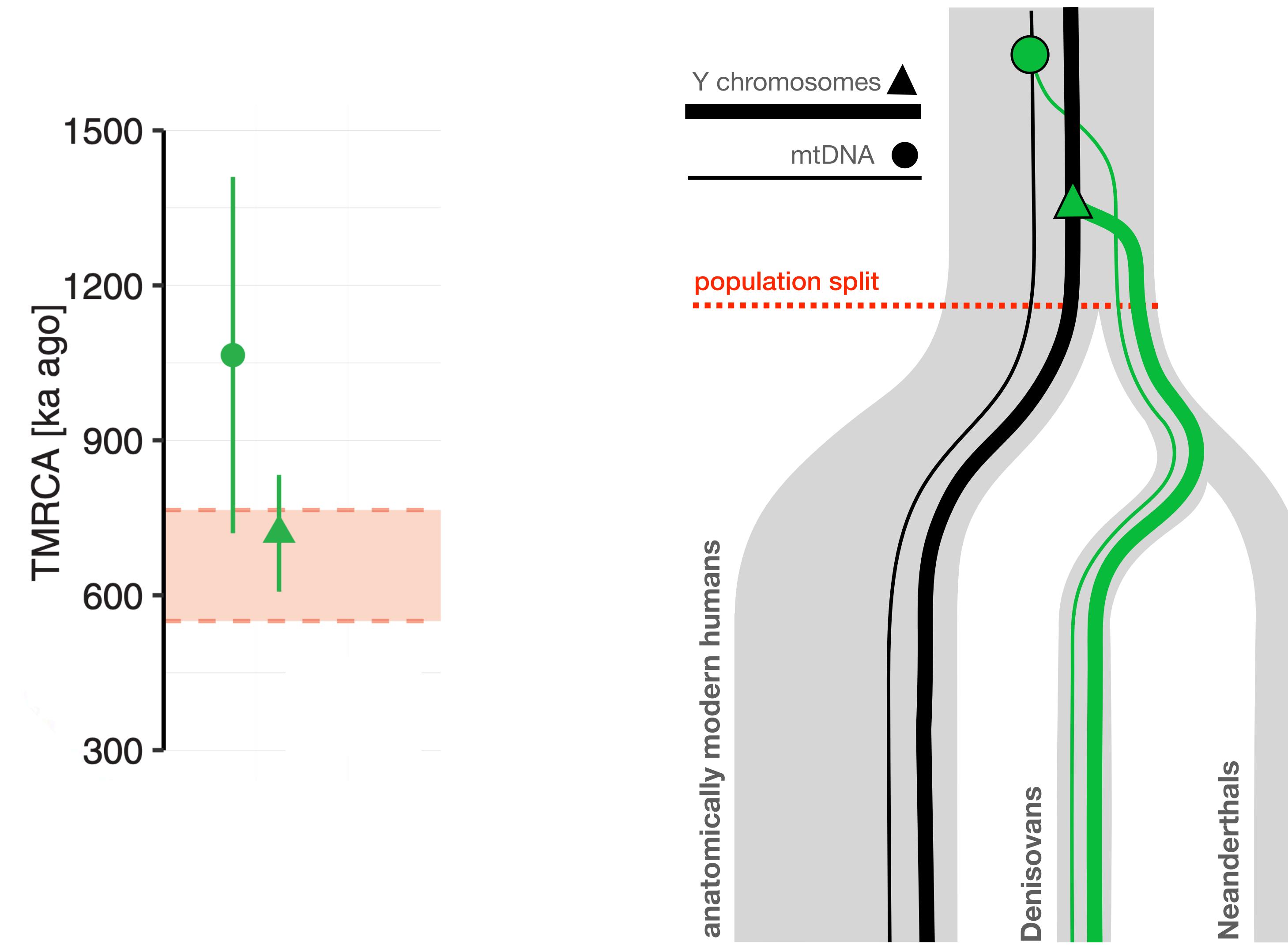
adapted from Browning et al., Cell, 2018



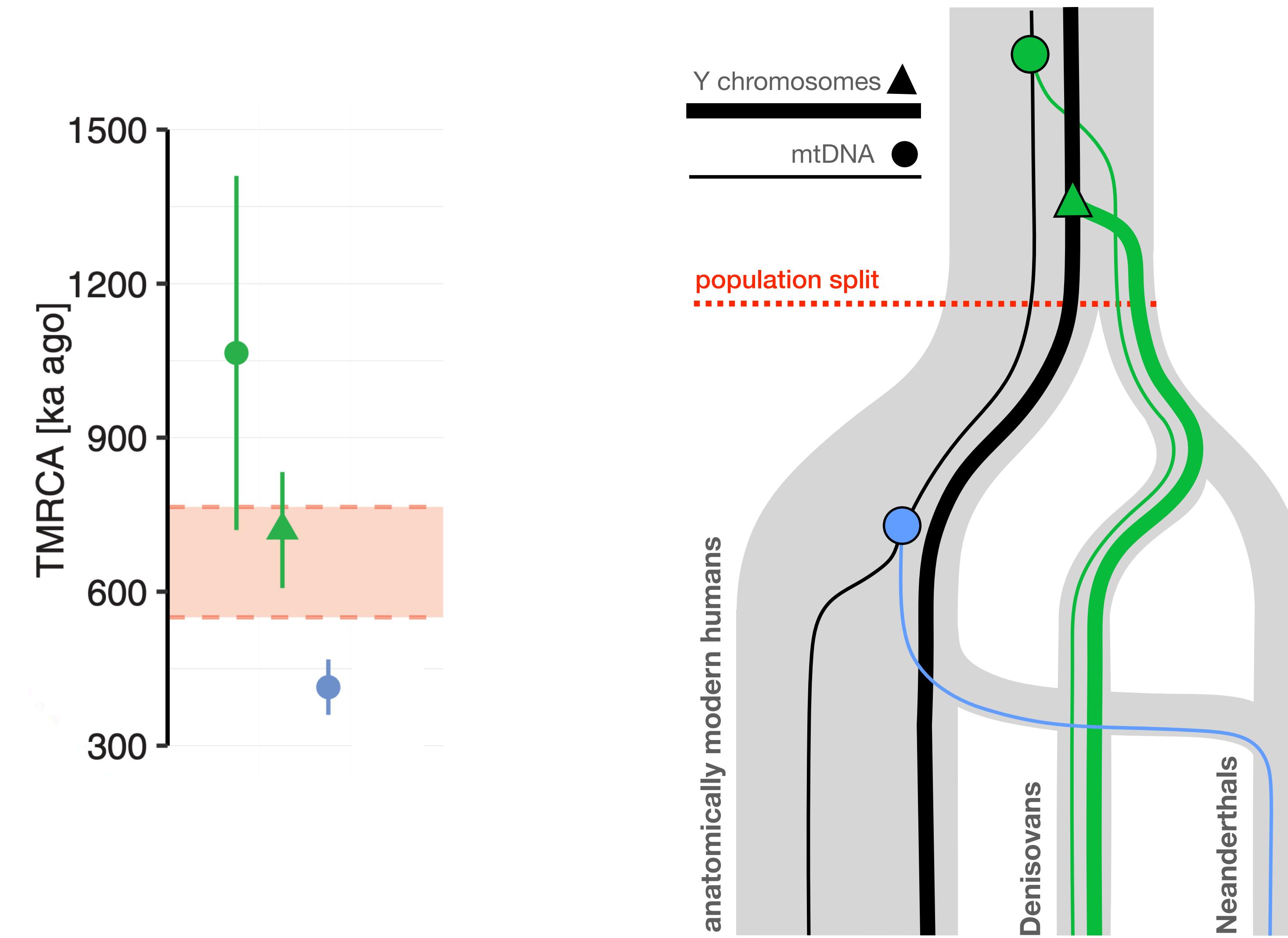
# Denisovan mtDNA follows known population history



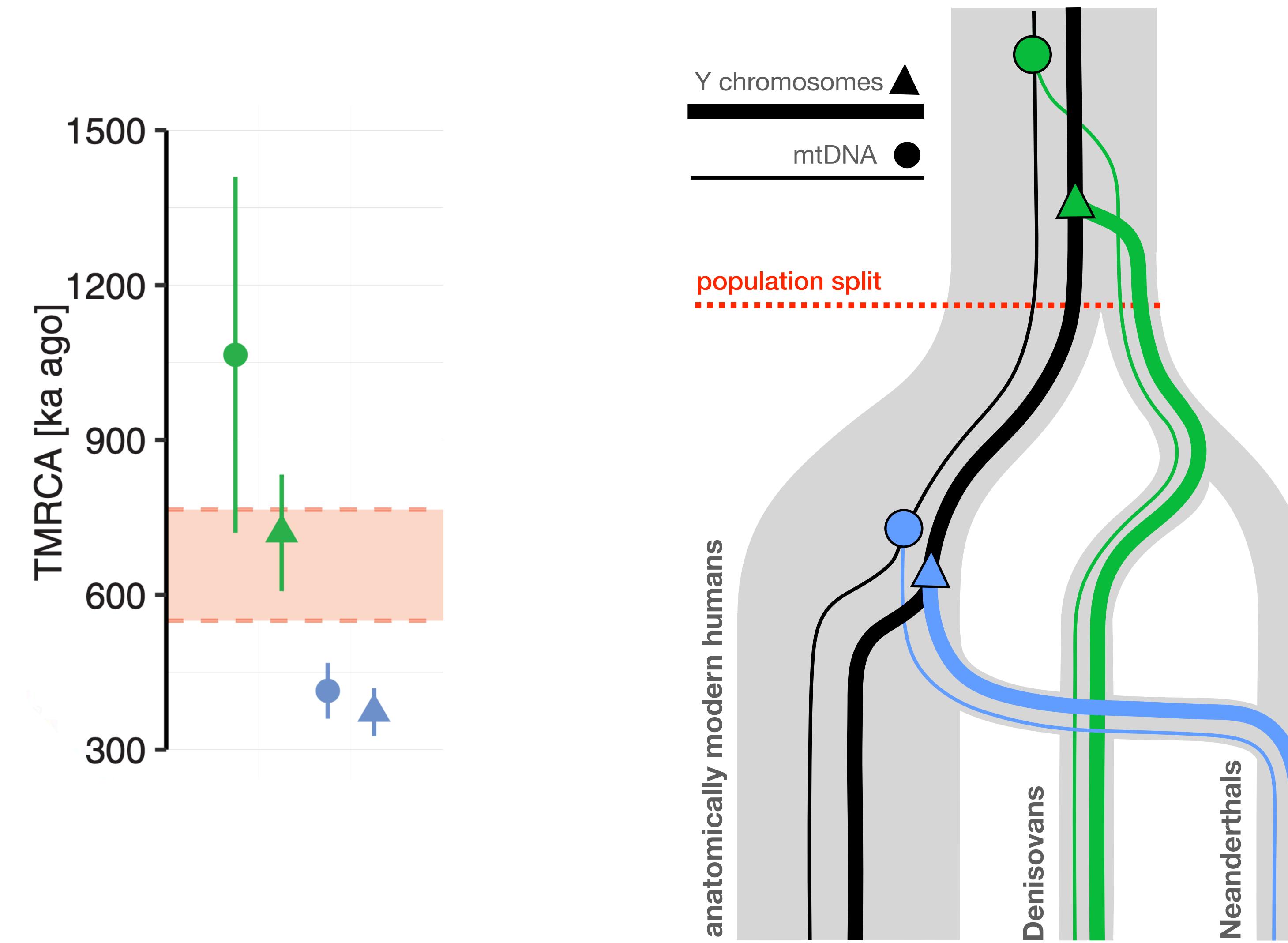
# Denisovan mtDNA follows known population history



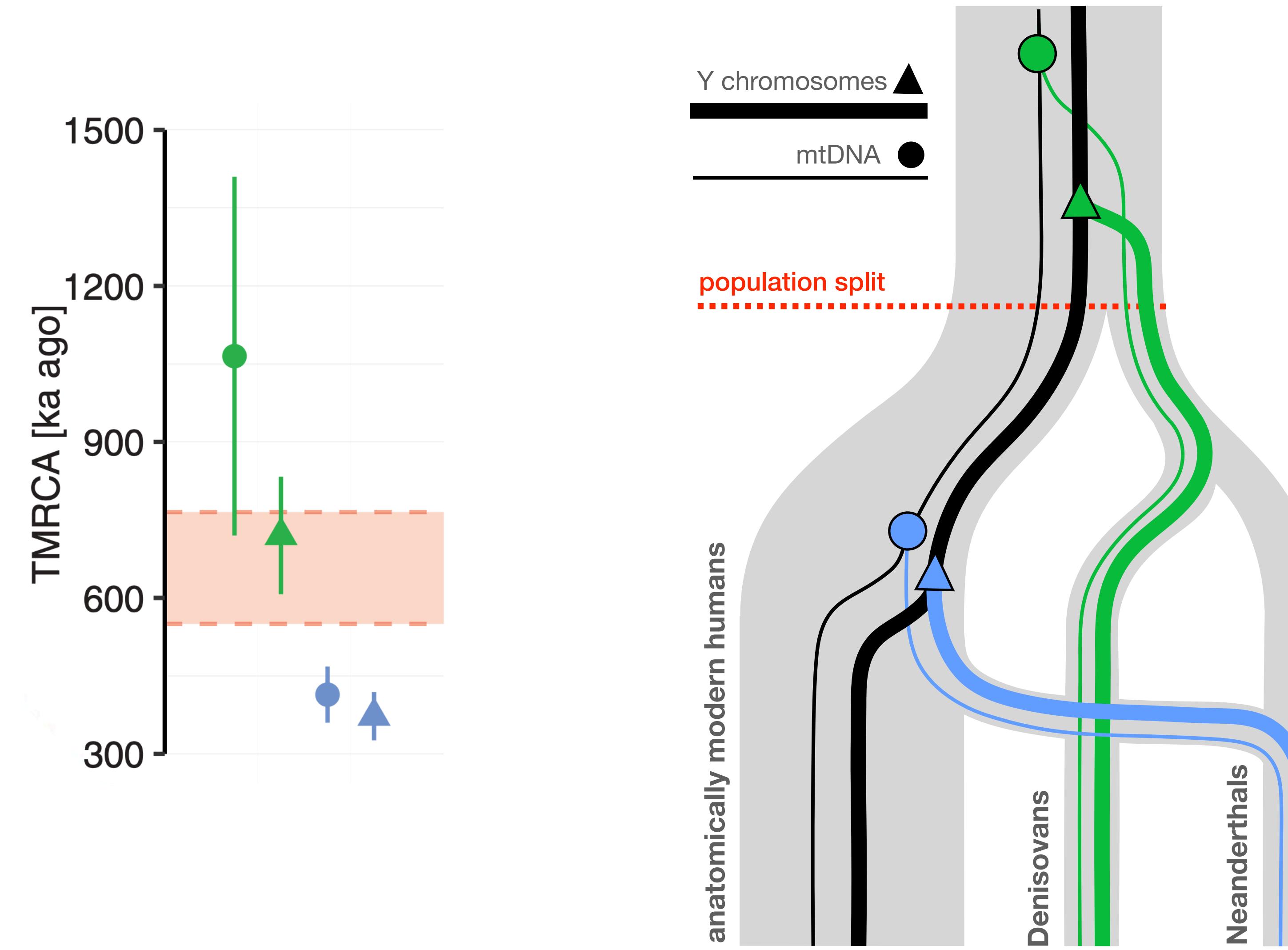
# Neanderthal mitochondrial DNAs contradict this!



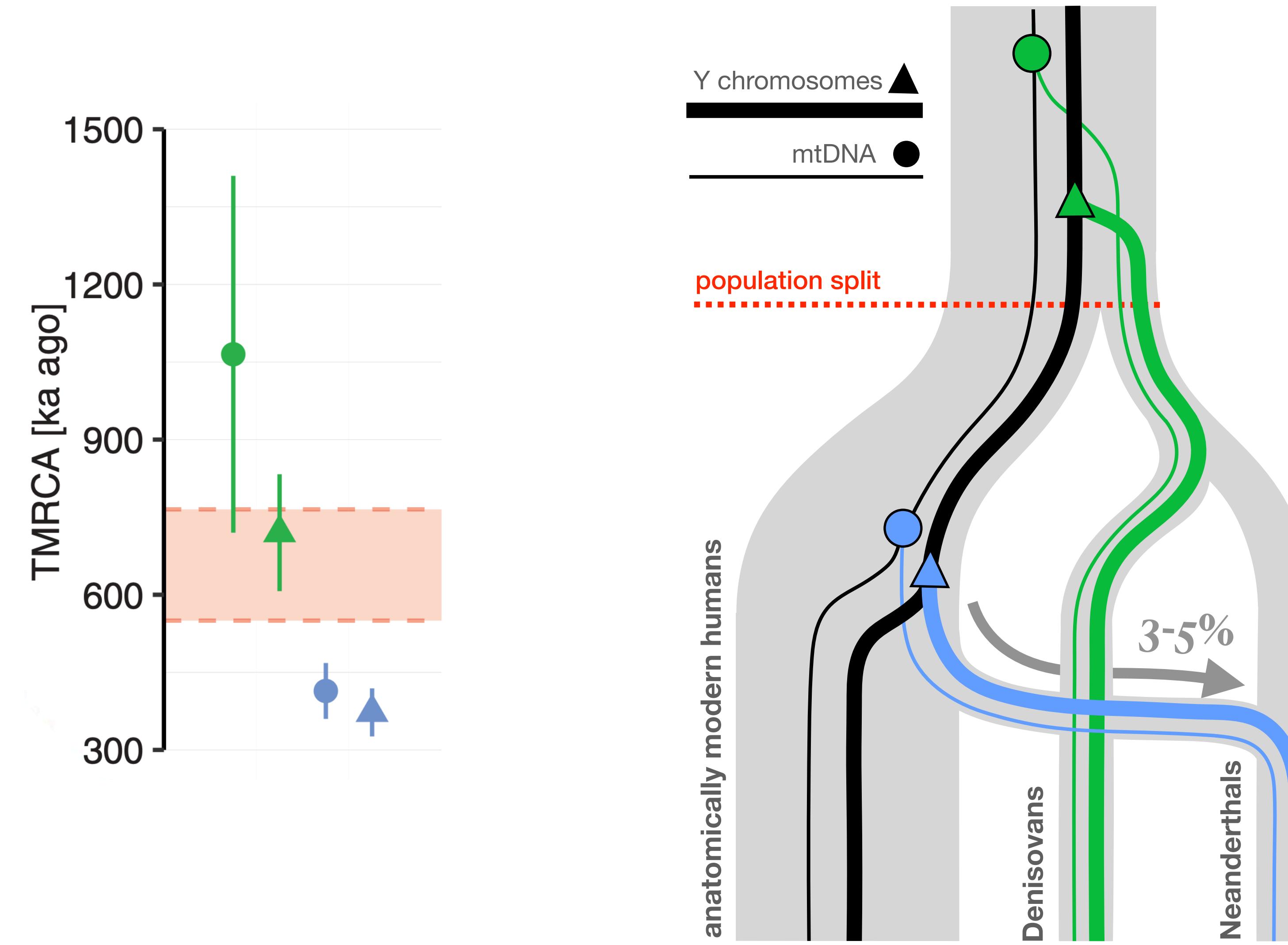
# Neanderthal Y chromosomes also contradict this!



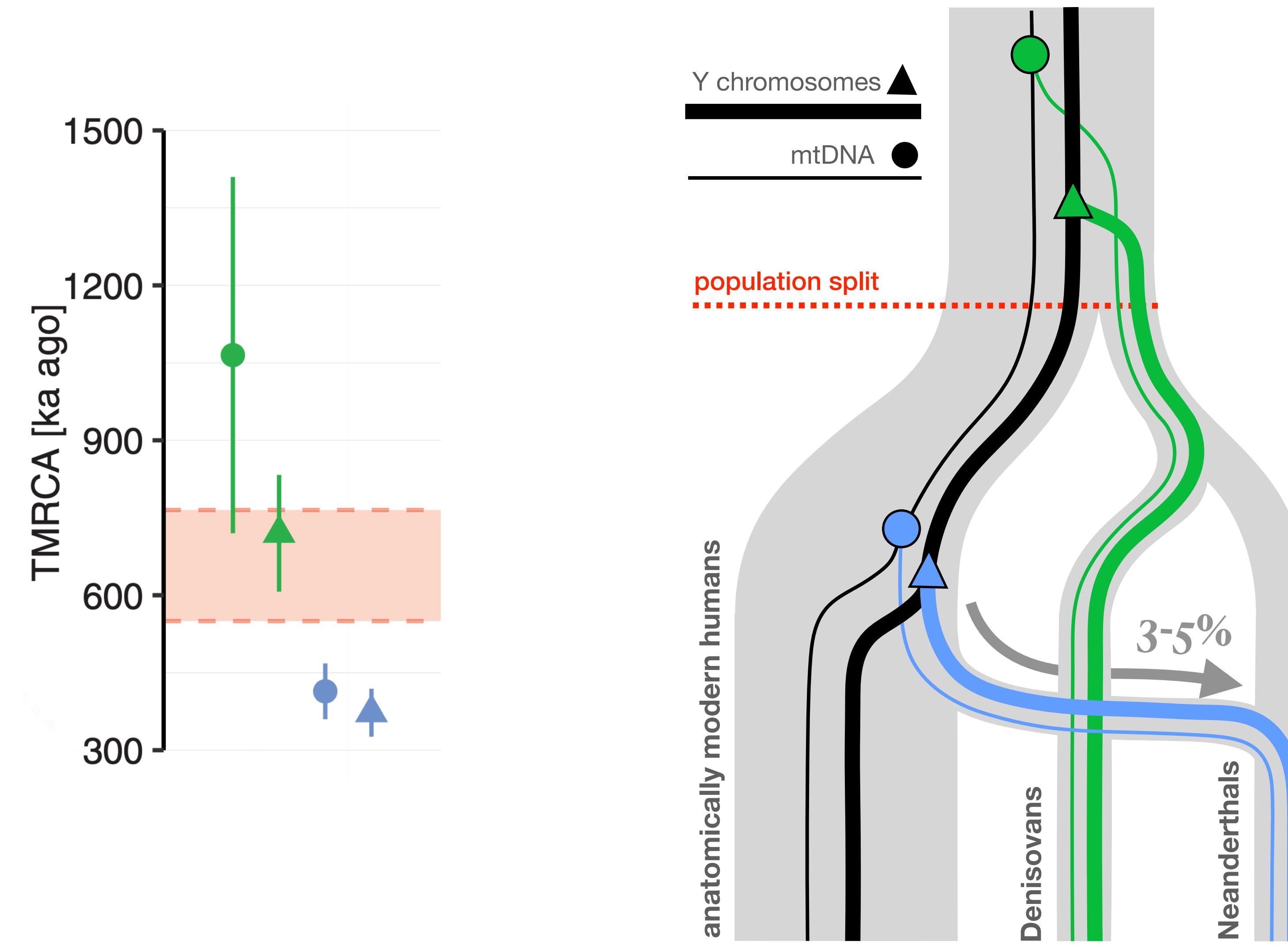
# Original Neanderthal mtDNA/Y experienced replacement



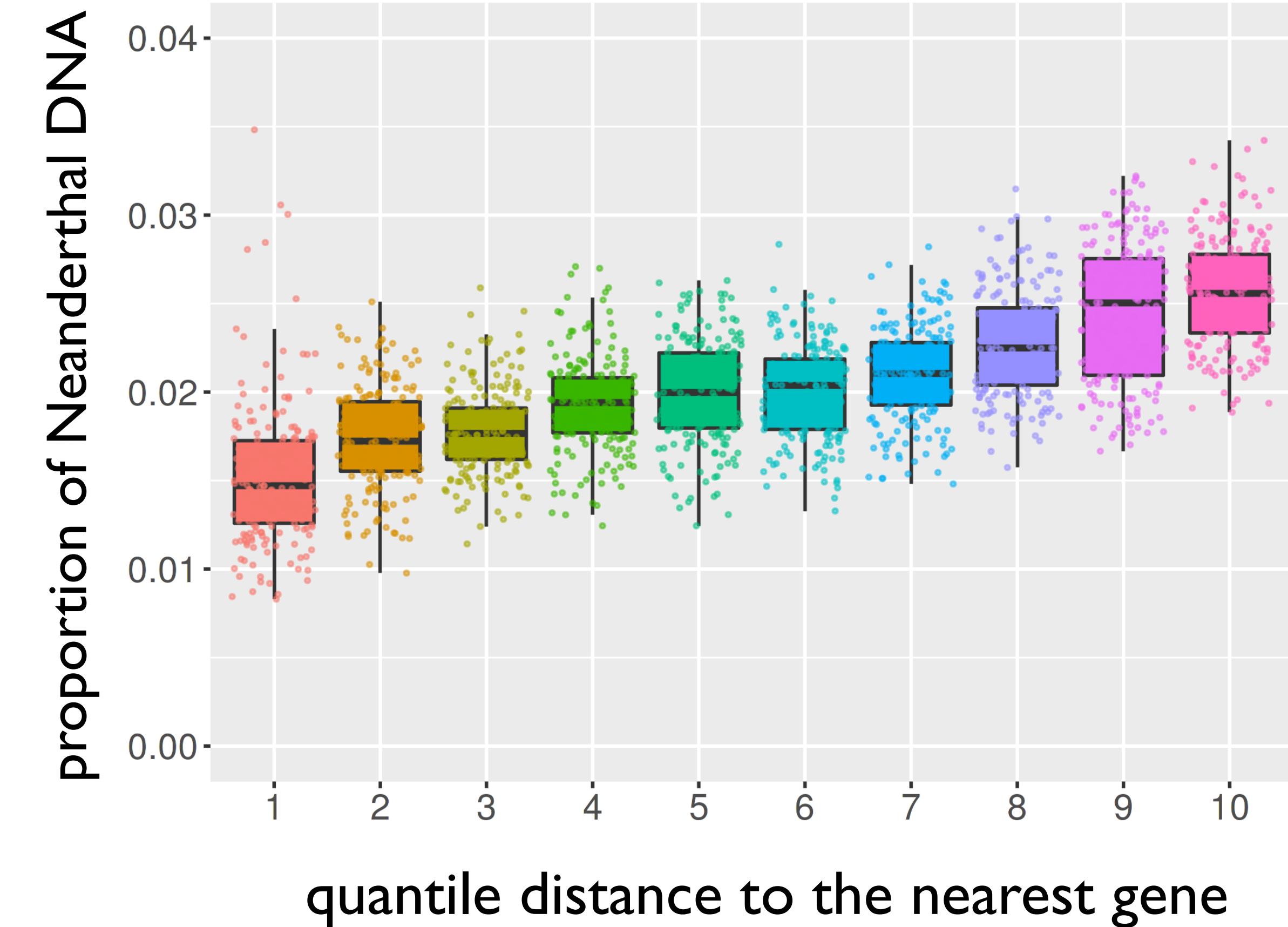
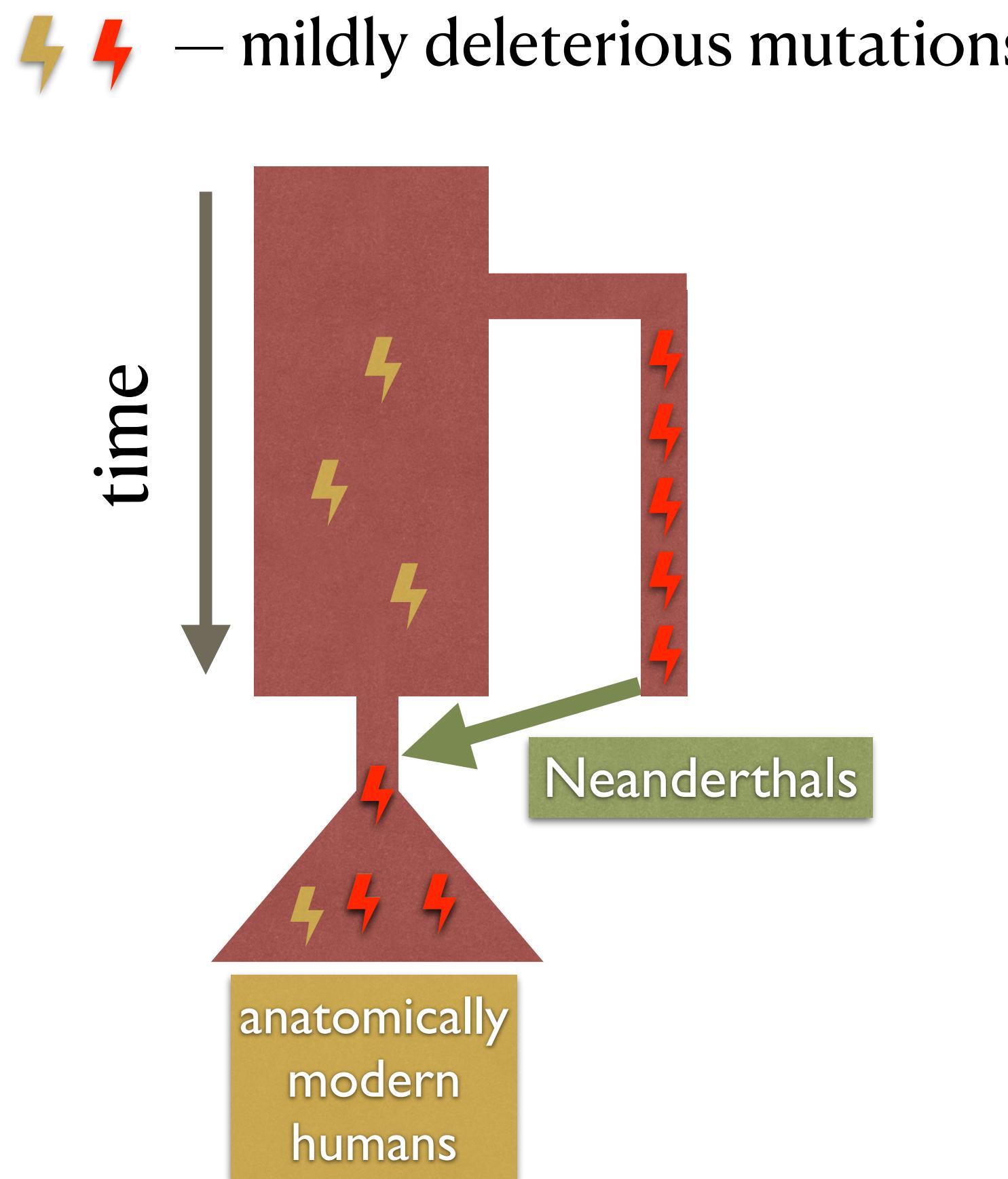
# Original Neanderthal mtDNA/Y experienced replacement



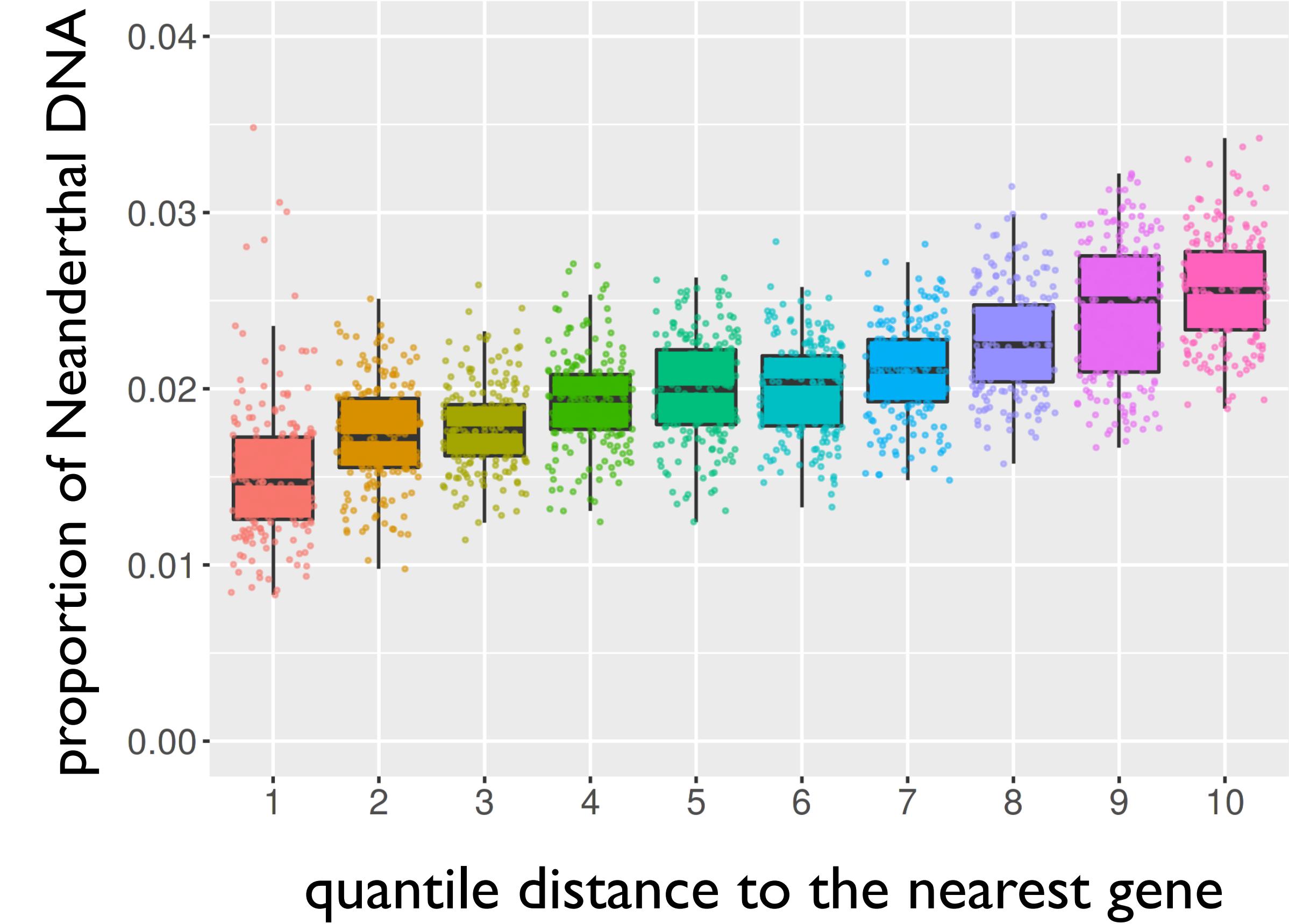
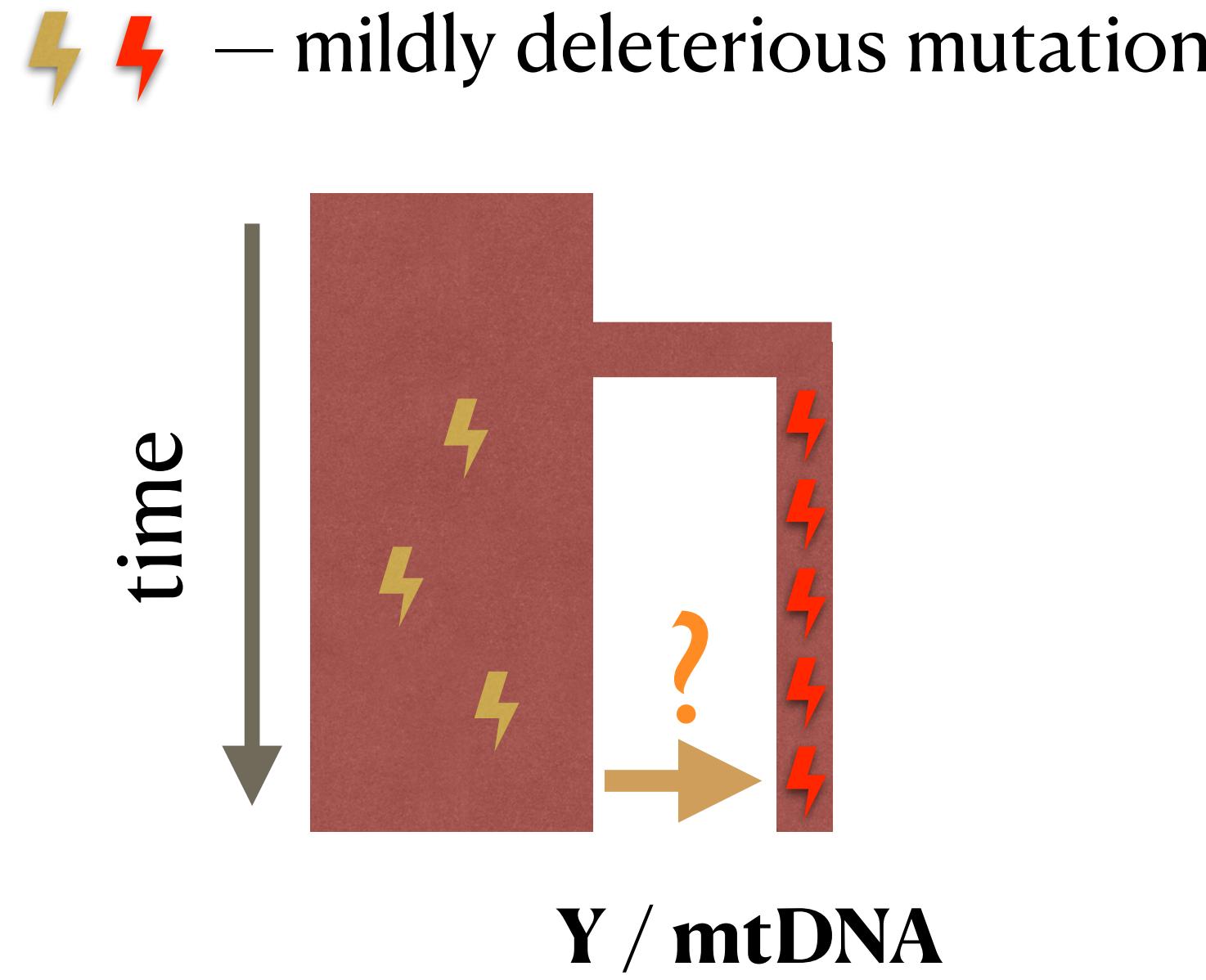
# Original Neanderthal mtDNA/Y experienced replacement (very unlikely under neutrality!)



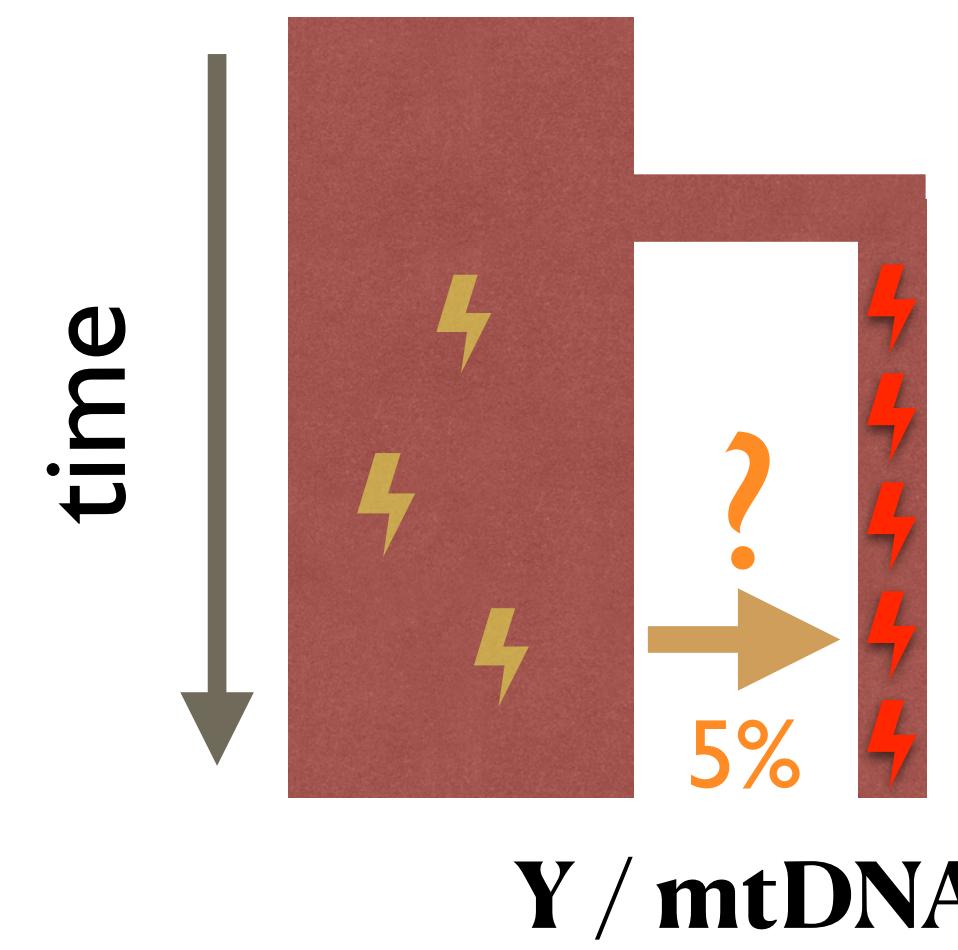
# ...but we know introgression was not neutral!



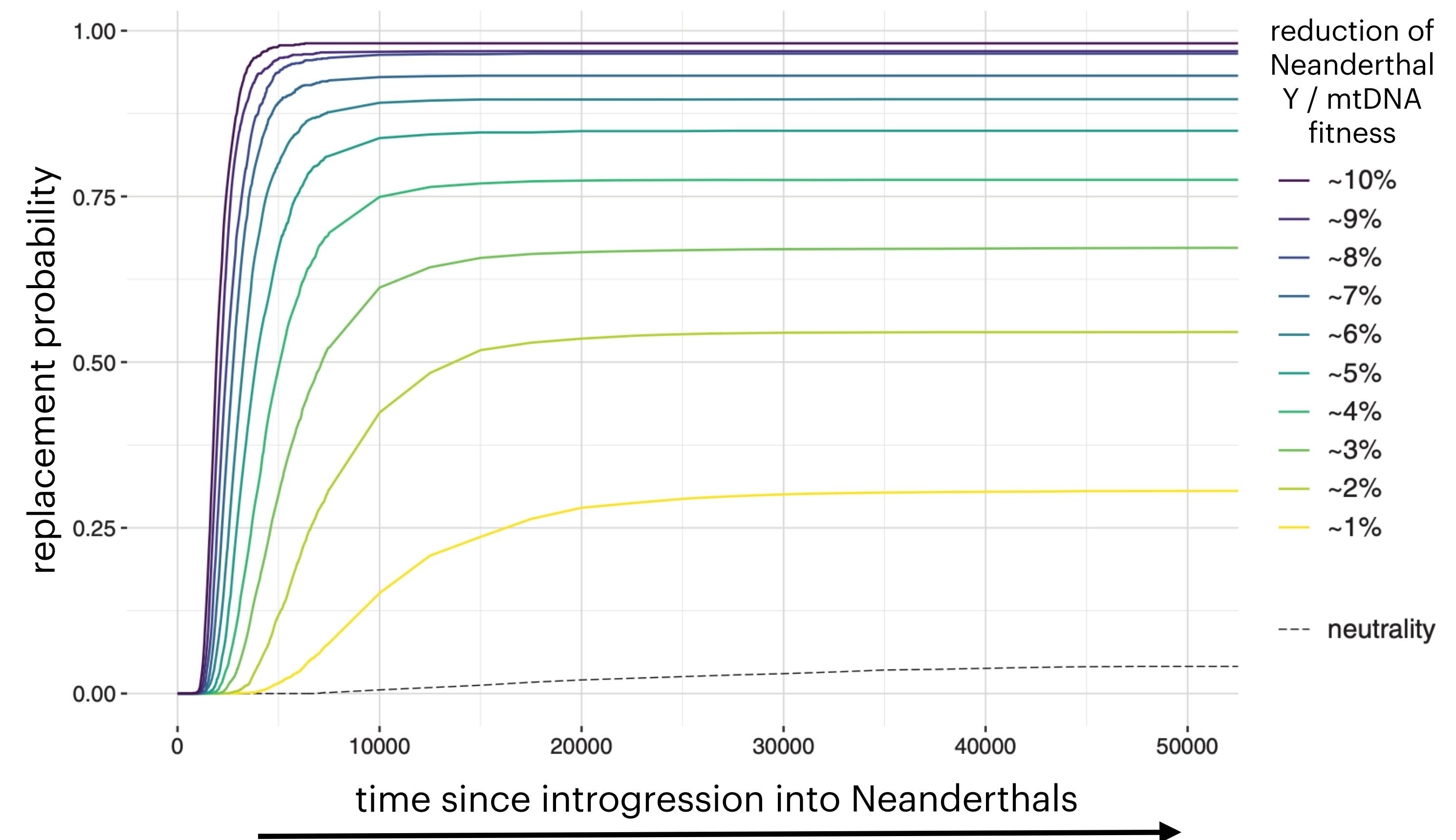
# What is the expected scenario for Y / mtDNA introgression?



# Natural selection in Neanderthals favors modern human Y / mtDNA



**Simulation of 5% gene flow from modern humans**



**It wasn't all bad though**

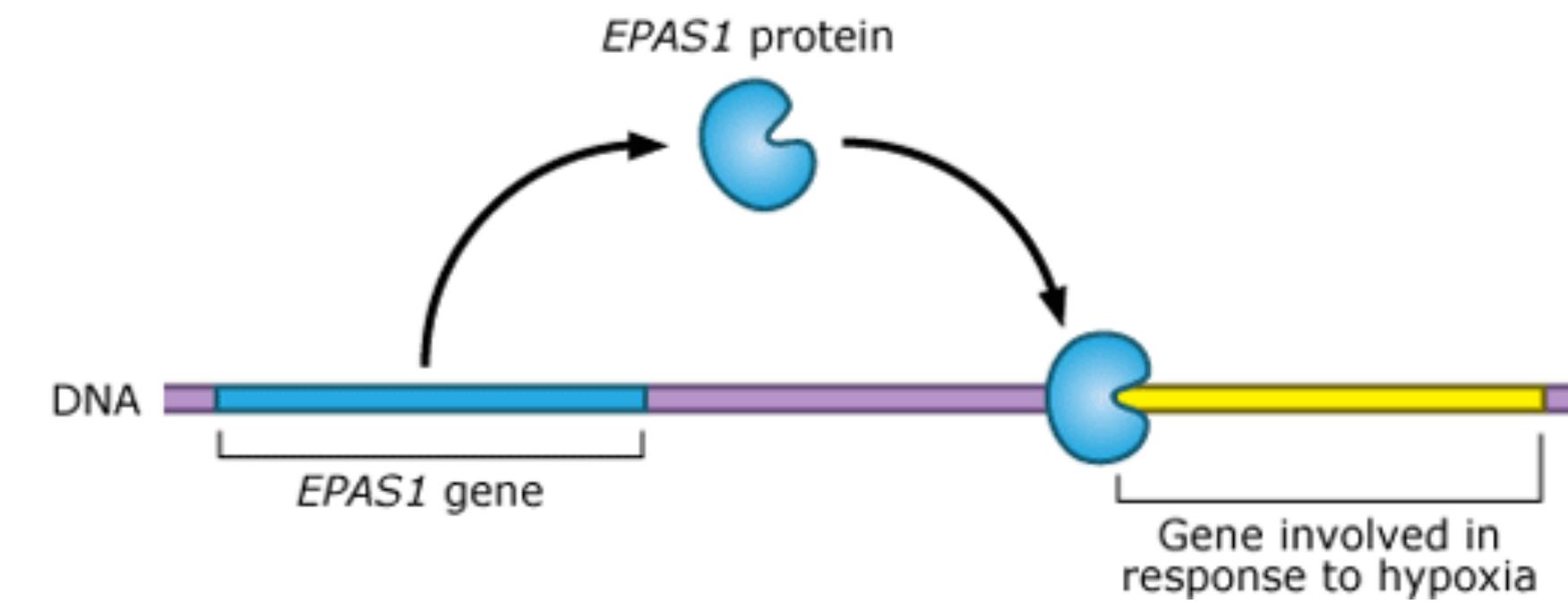
**positive selection / adaptive introgression**

# High altitude adaptation in Tibetans

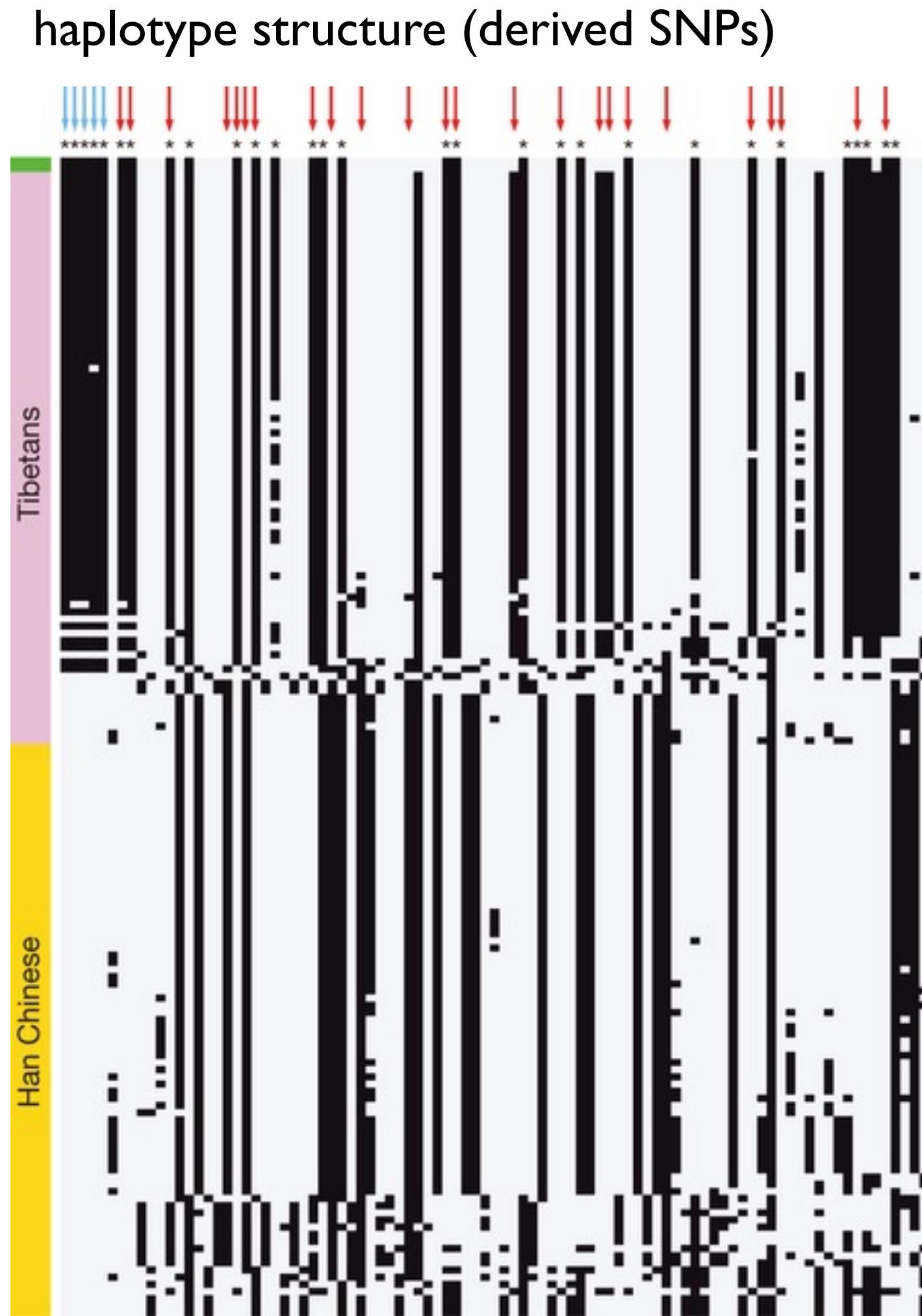
haplotype structure (derived SNPs)



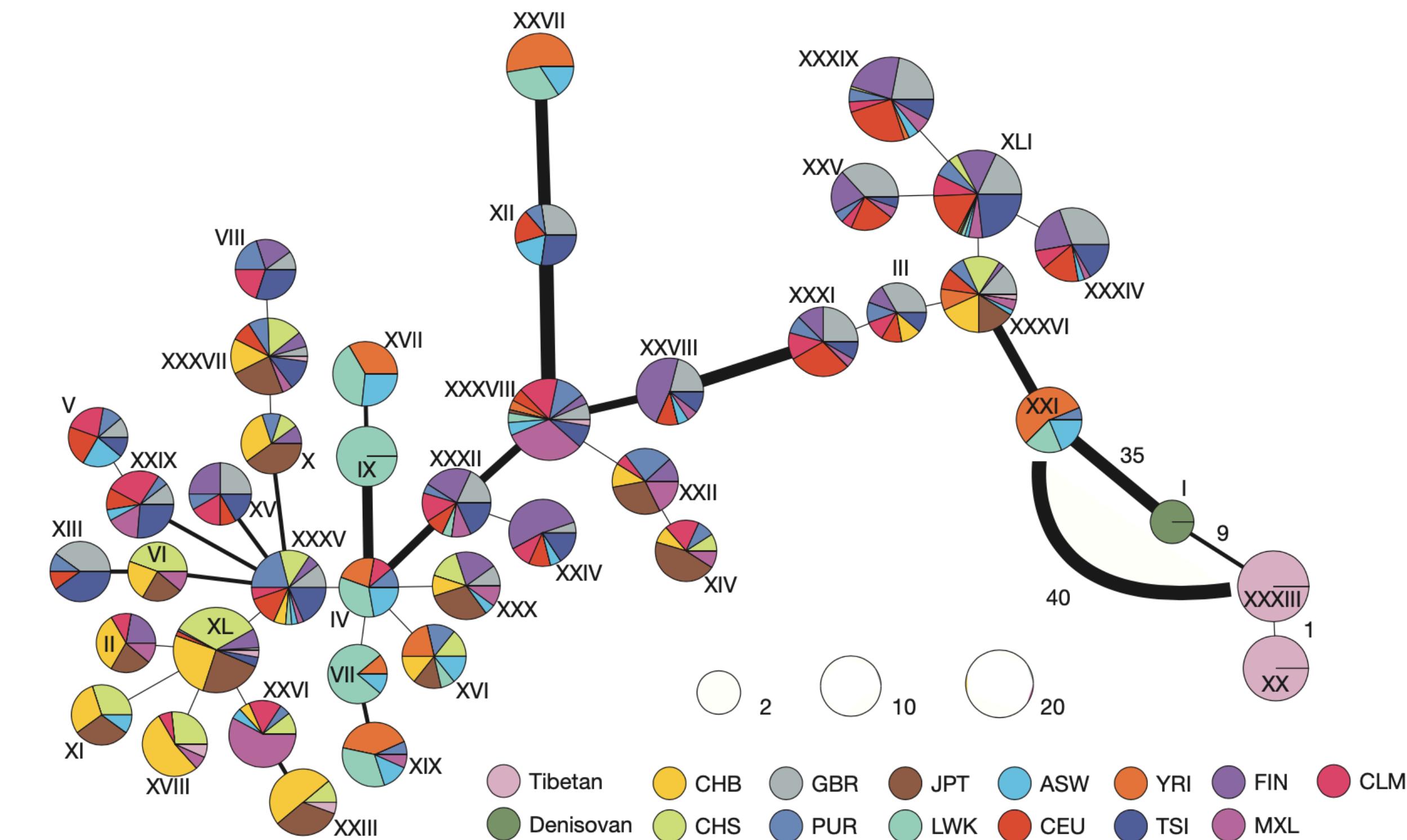
Schematic of *EPAS1* function



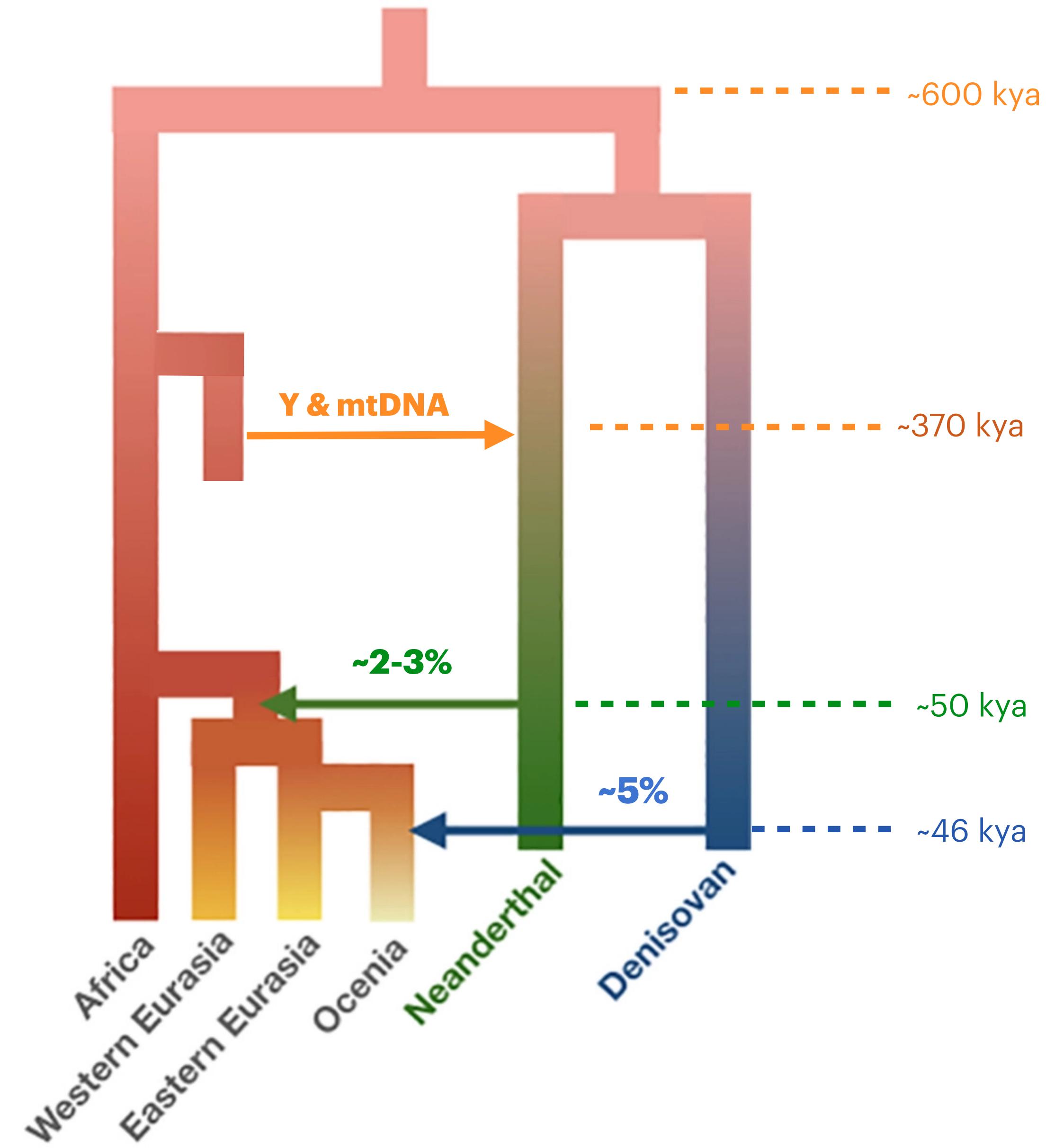
# High altitude adaptation via Denisovan introgression!



haplotype network based on pairwise-differences



**Introgression has been very frequent event in human history.**



adapted from the Yearbook of Physical Anthropology (2019) by Gokcumen

Letter | Published: 22 August 2018

## The genome of the offspring of a Neanderthal mother and a Denisovan father

Viviane Slon✉, Fabrizio Mafessoni, Benjamin Vernot, Cesare de Filippo, Steffi Grote, Bence Viola, Mateja Hajdinjak, Stéphane Peyrégne, Sarah Nagel, Samantha Brown, Katerina Douka, Tom Higham, Maxim B. Kozlikin, Michael V. Shunkov, Anatoly P. Derevianko, Janet Kelso, Matthias Meyer, Kay Prüfer & Svante Pääbo✉



1 cm

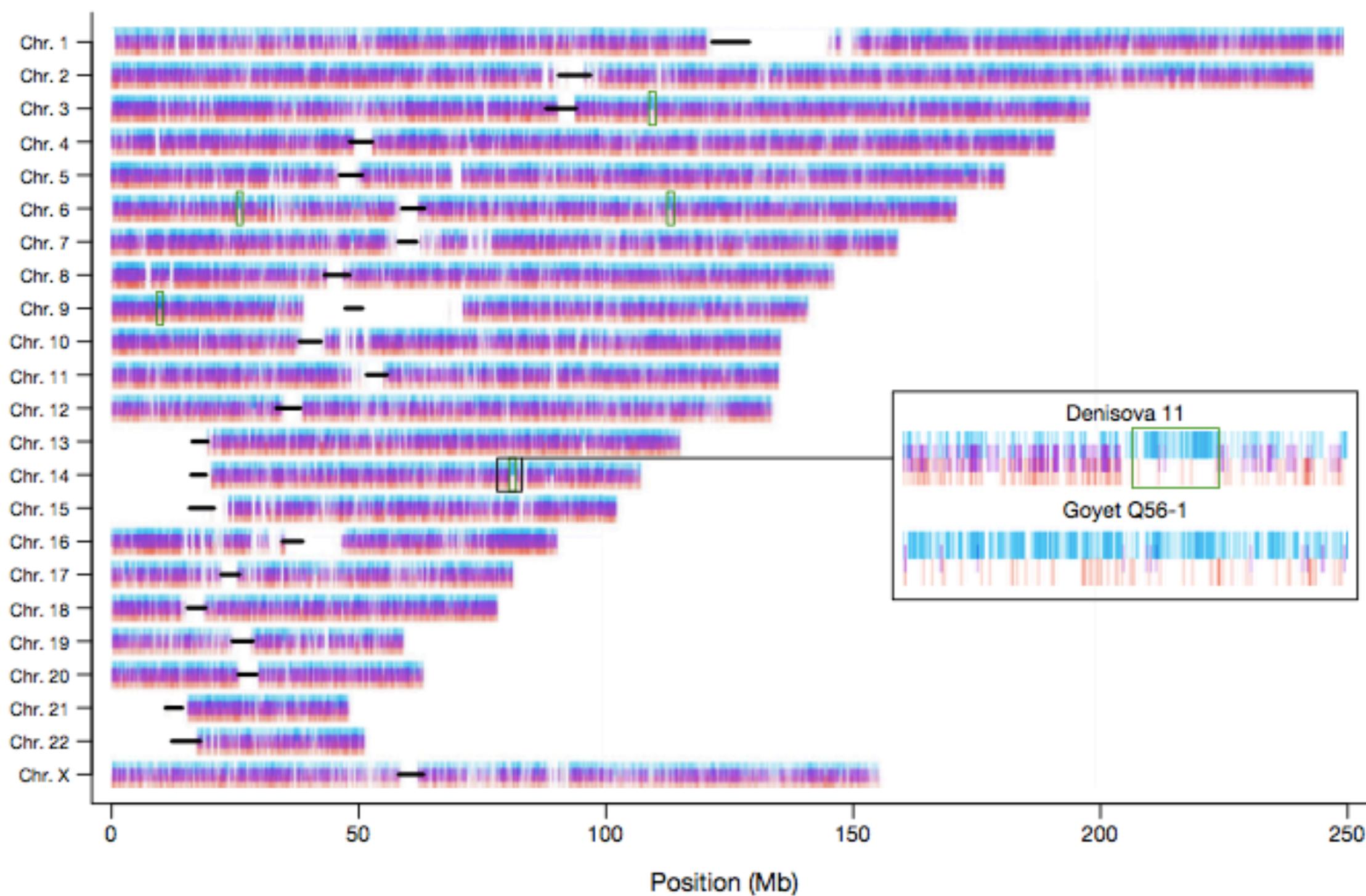


By Демин Алексей Барнаул - Own work, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=48890364>

Letter | Published: 22 August 2018

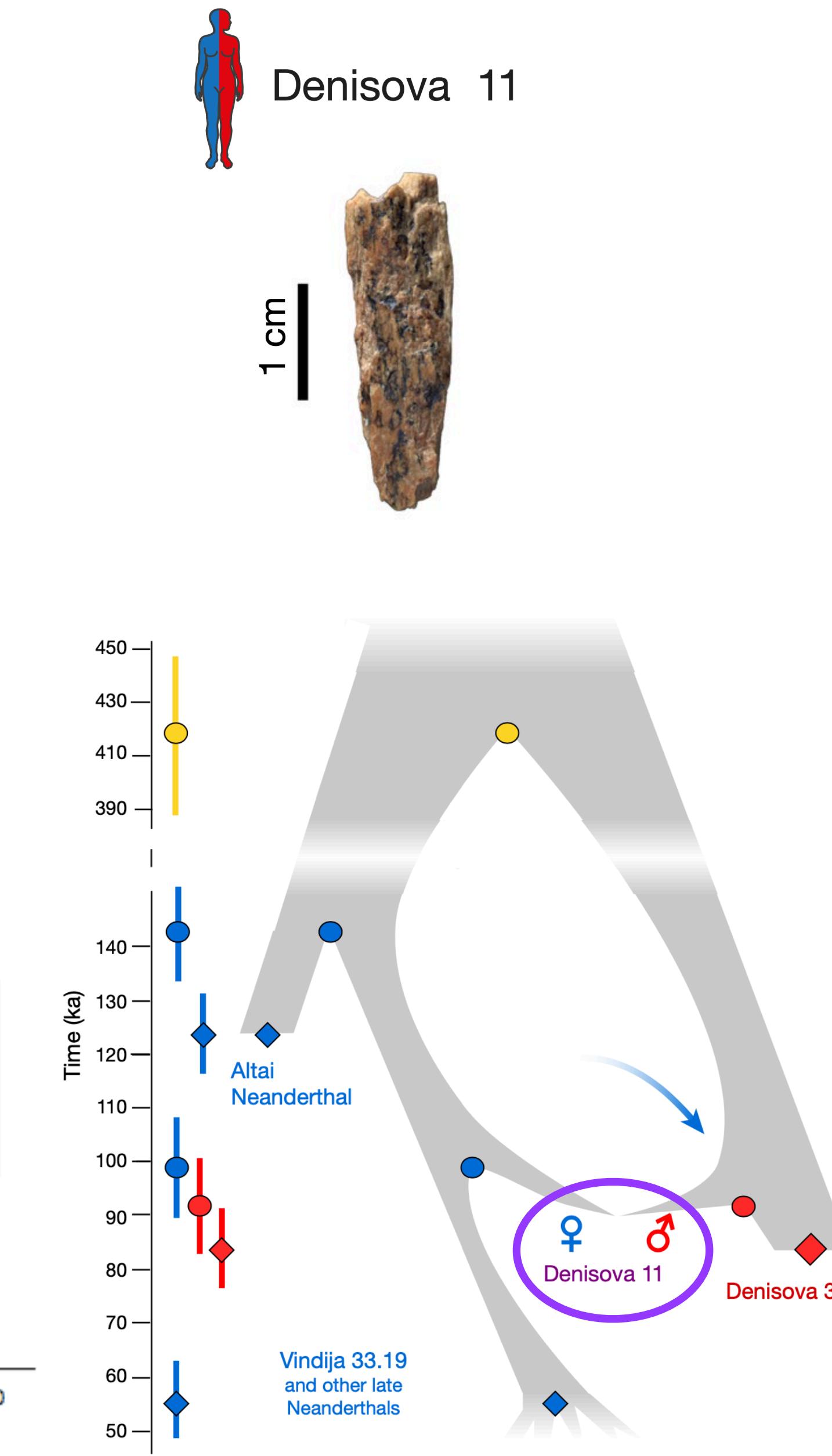
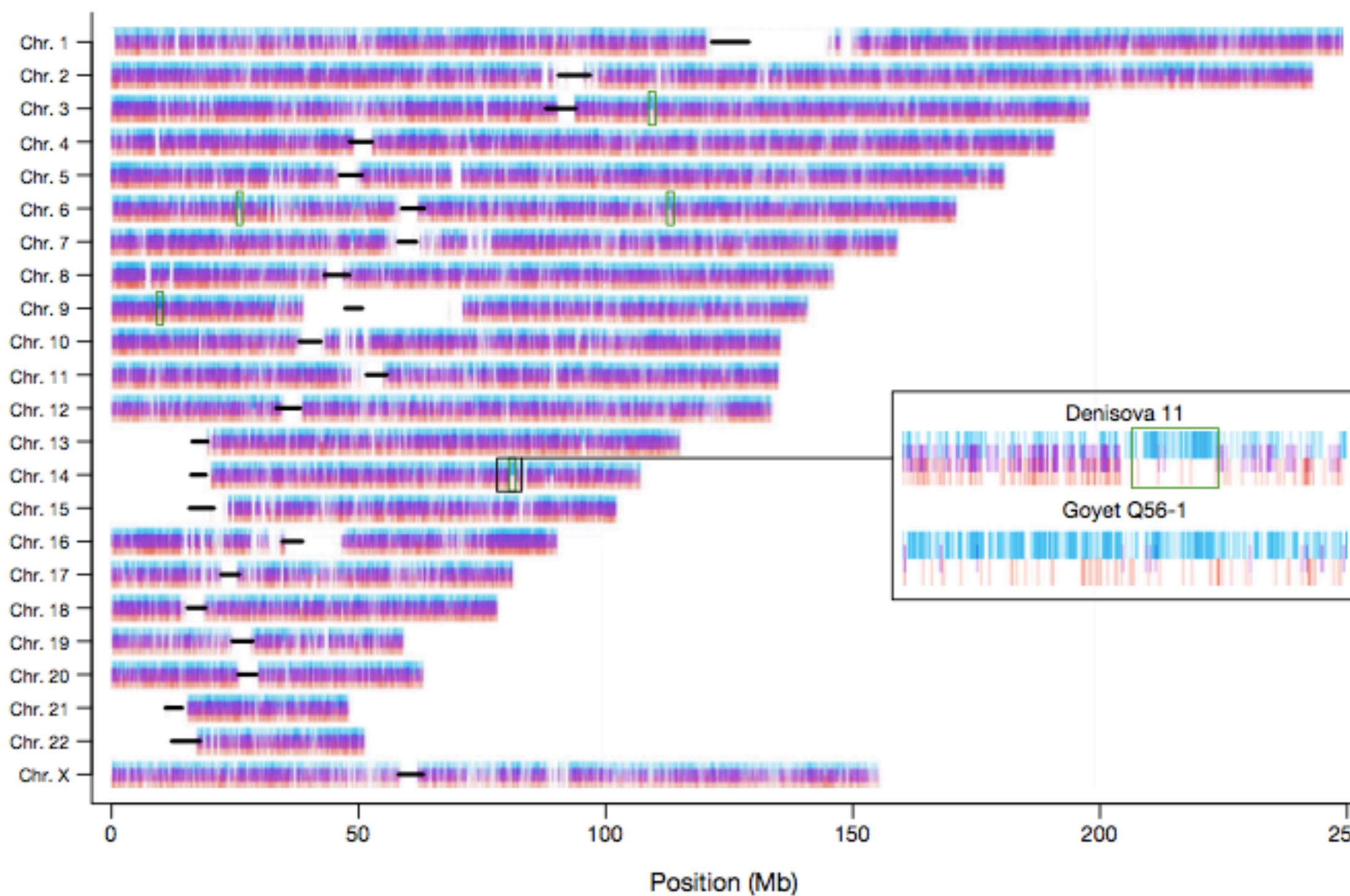
# The genome of the offspring of a Neanderthal mother and a Denisovan father

Viviane Slon✉, Fabrizio Mafessoni, Benjamin Vernot, Cesare de Filippo, Steffi Grote, Bence Viola, Mateja Hajdinjak, Stéphane Peyrégne, Sarah Nagel, Samantha Brown, Katerina Douka, Tom Higham, Maxim B. Kozlikin, Michael V. Shunkov, Anatoly P. Derevianko, Janet Kelso, Matthias Meyer, Kay Prüfer & Svante Pääbo✉



# The genome of the offspring of a Neanderthal mother and a Denisovan father

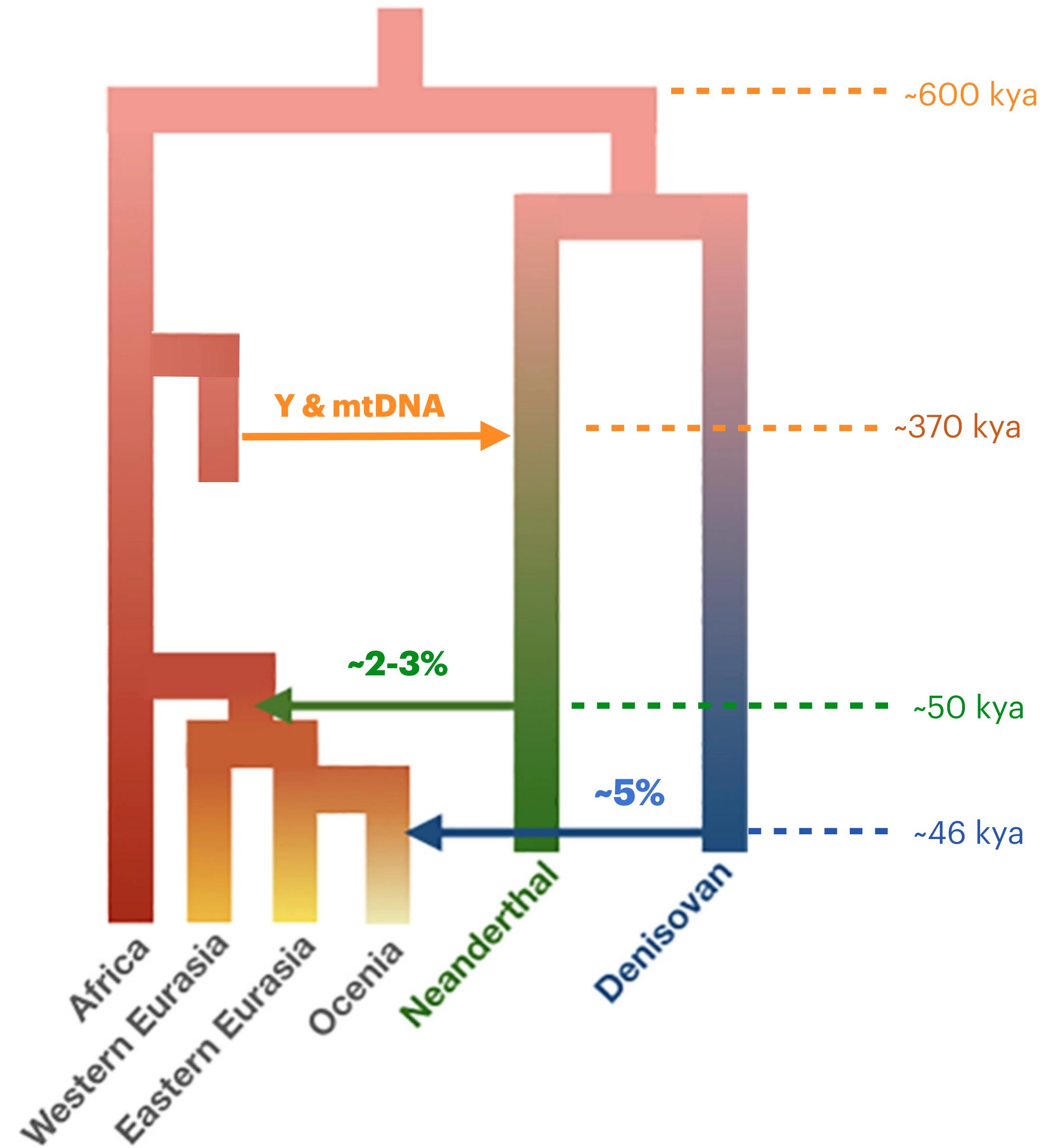
Viviane Slon✉, Fabrizio Mafessoni, Benjamin Vernot, Cesare de Filippo, Steffi Grote, Bence Viola, Mateja Hajdinjak, Stéphane Peyrégne, Sarah Nagel, Samantha Brown, Katerina Douka, Tom Higham, Maxim B. Kozlikin, Michael V. Shunkov, Anatoly P. Derevianko, Janet Kelso, Matthias Meyer, Kay Prüfer & Svante Pääbo✉



**Introgression has been very frequent event in human history.**

**Introgressed DNA has had significant impact on the biology of “recipient populations”.**

**Population genetic simulations critical for testing evolutionary hypotheses.**





?



?



?



?

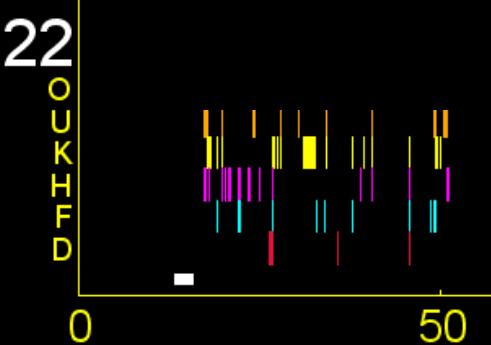
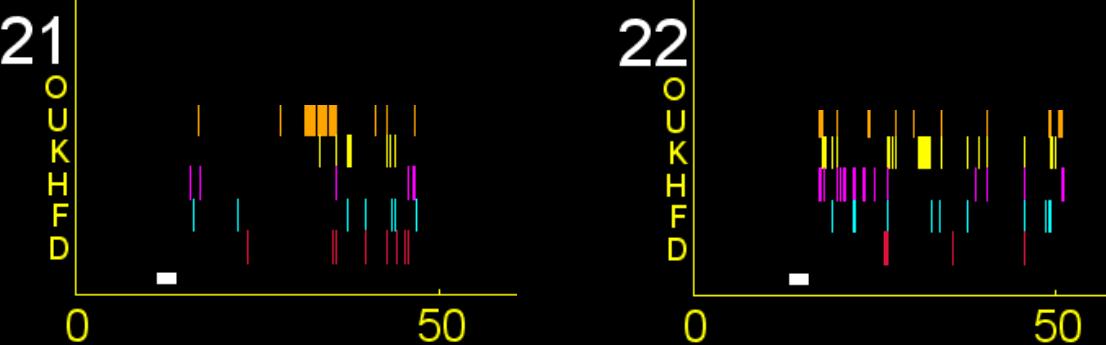
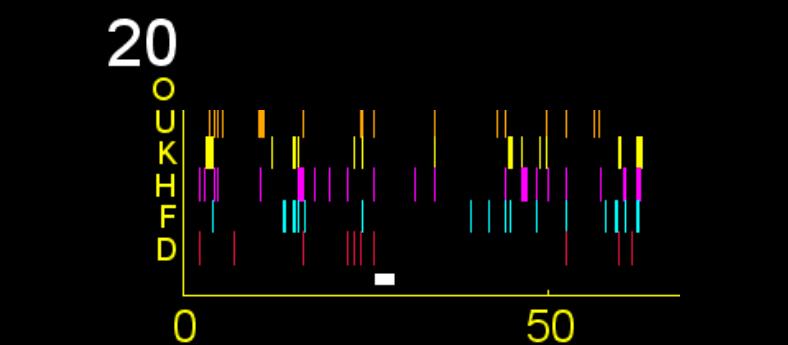
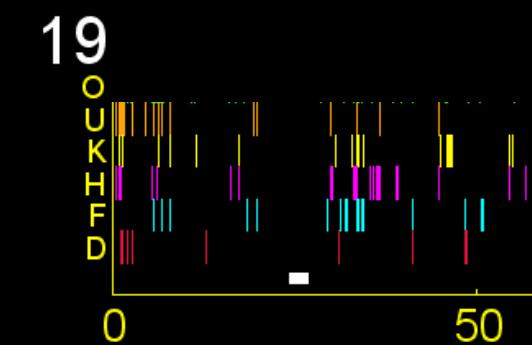
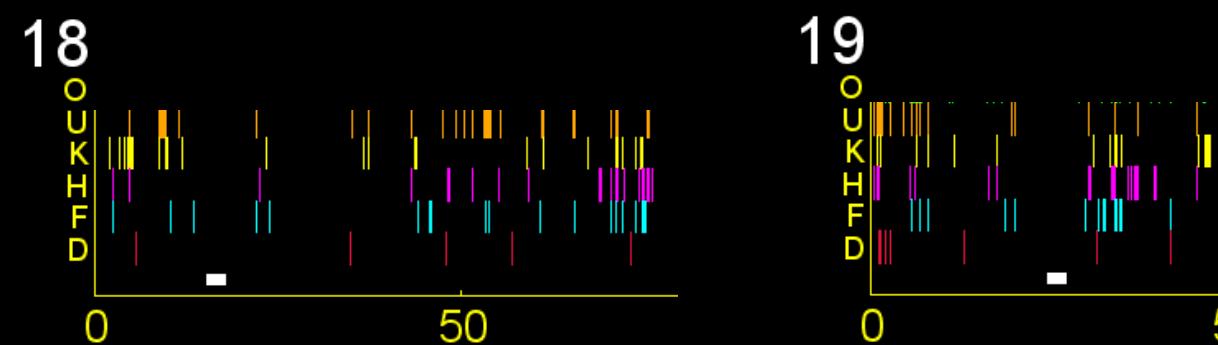
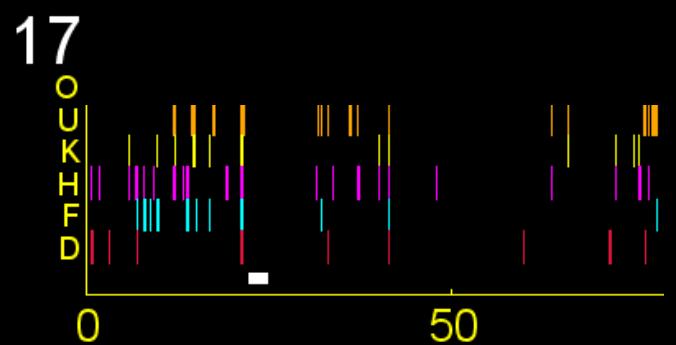
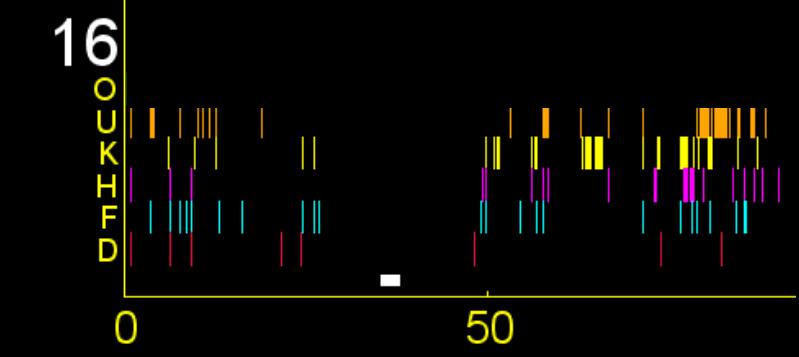
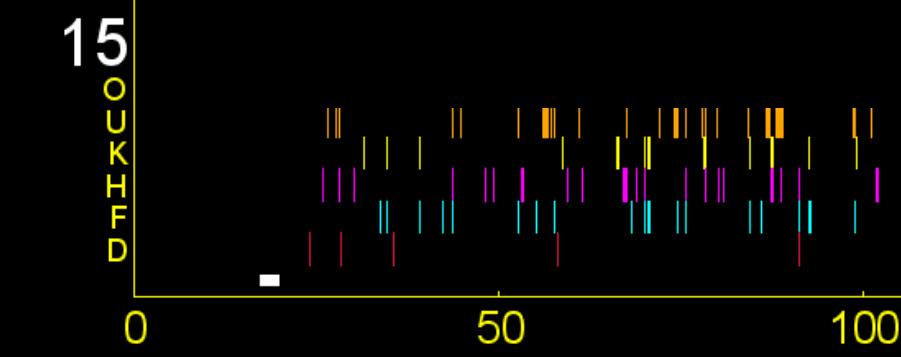
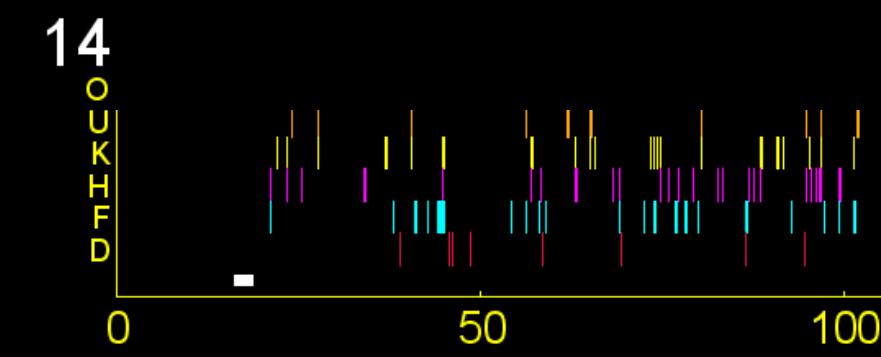
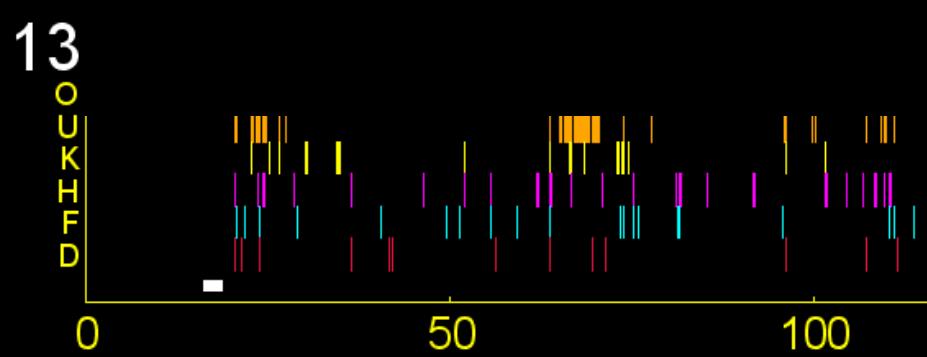
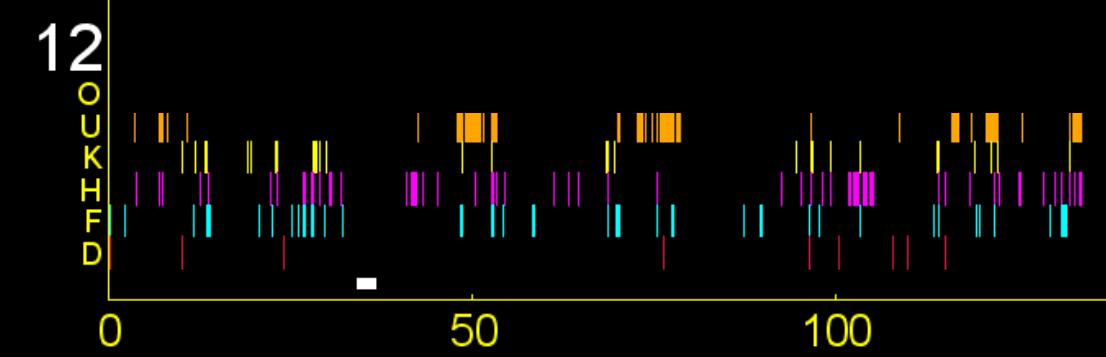
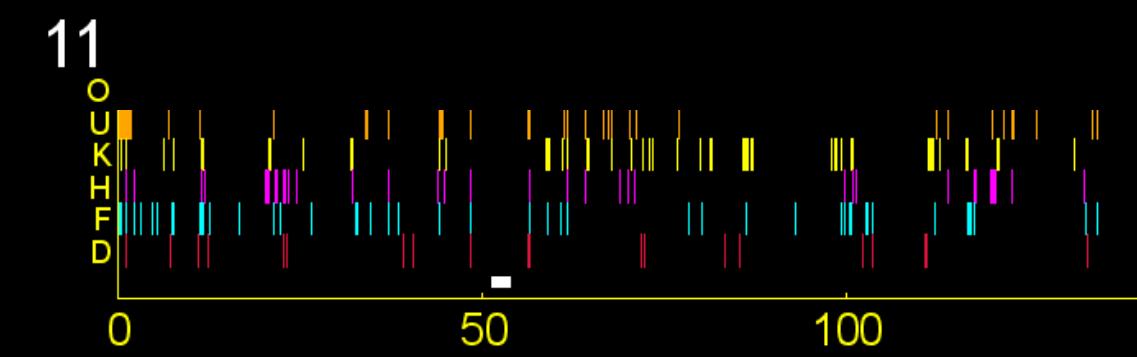
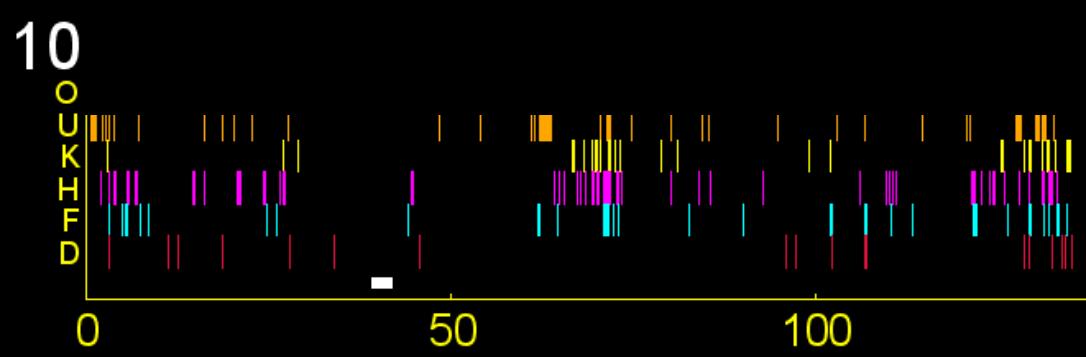
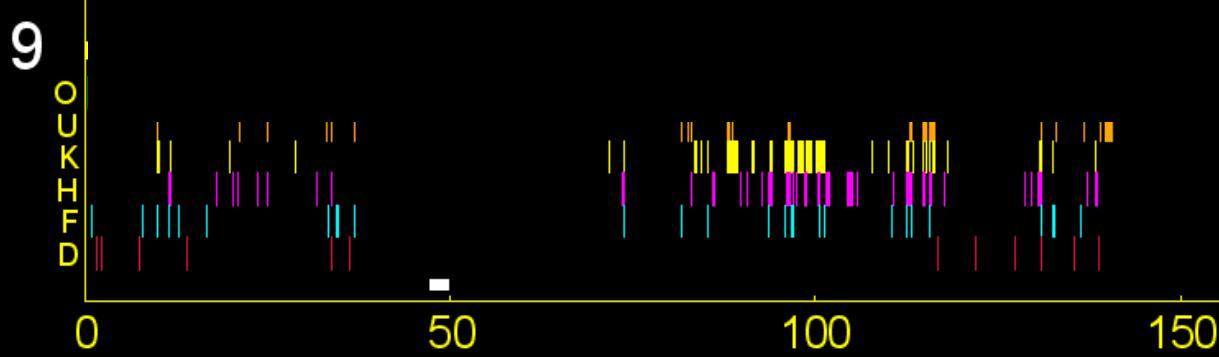
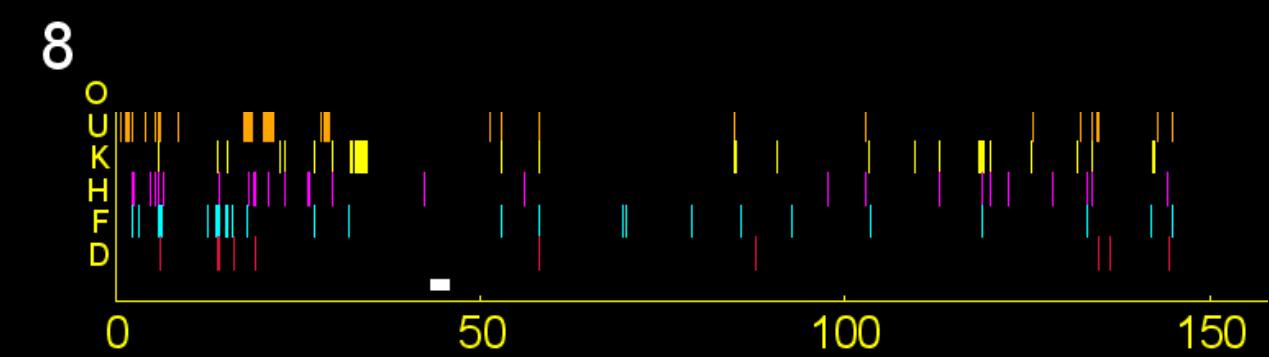
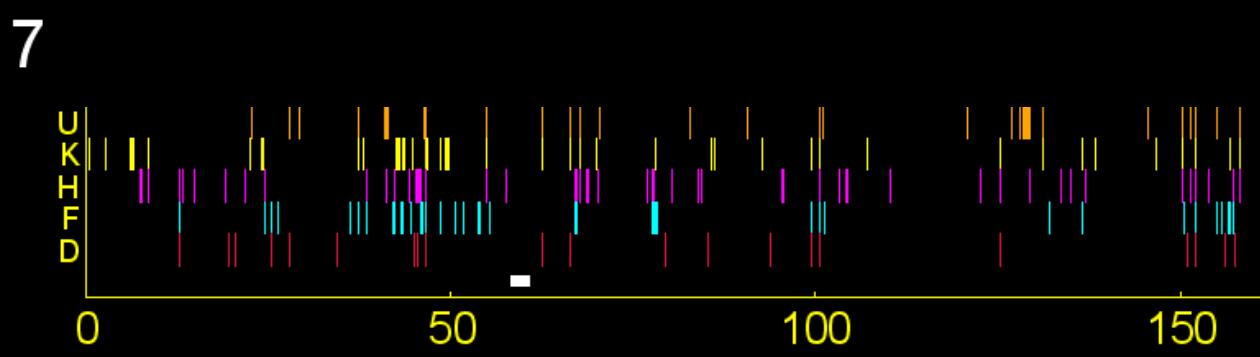
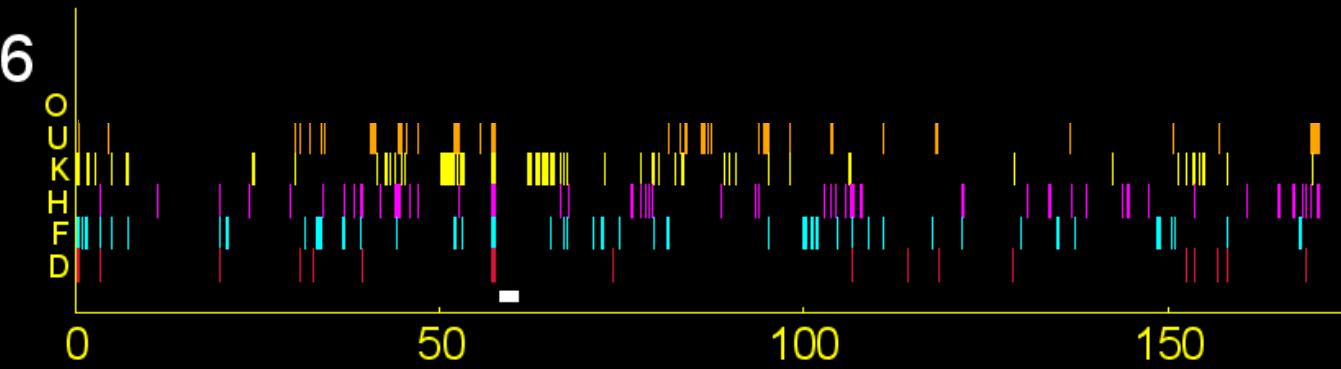
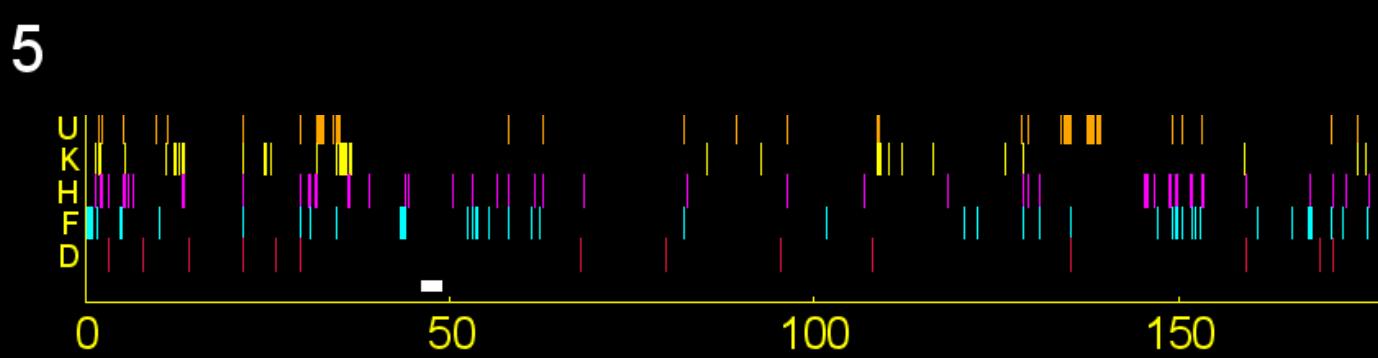
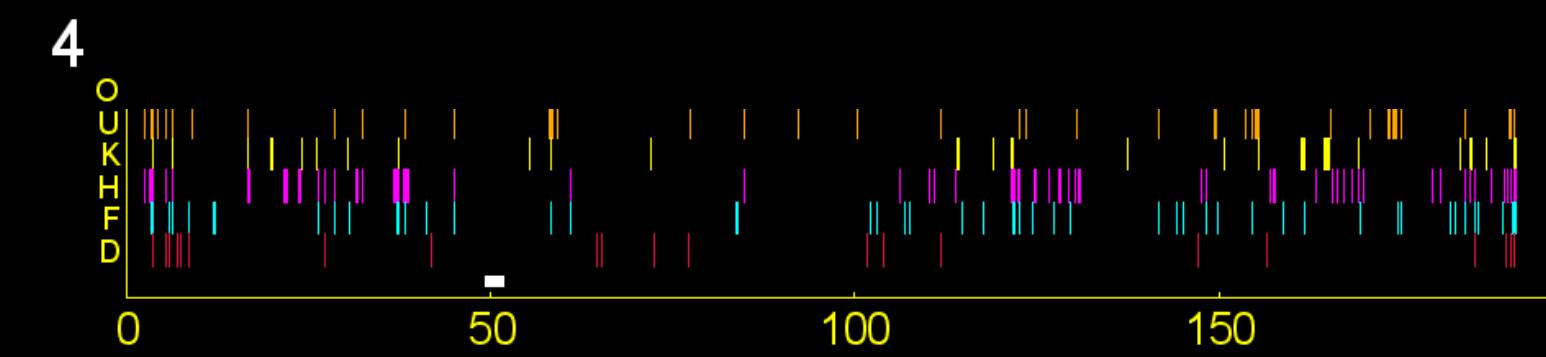
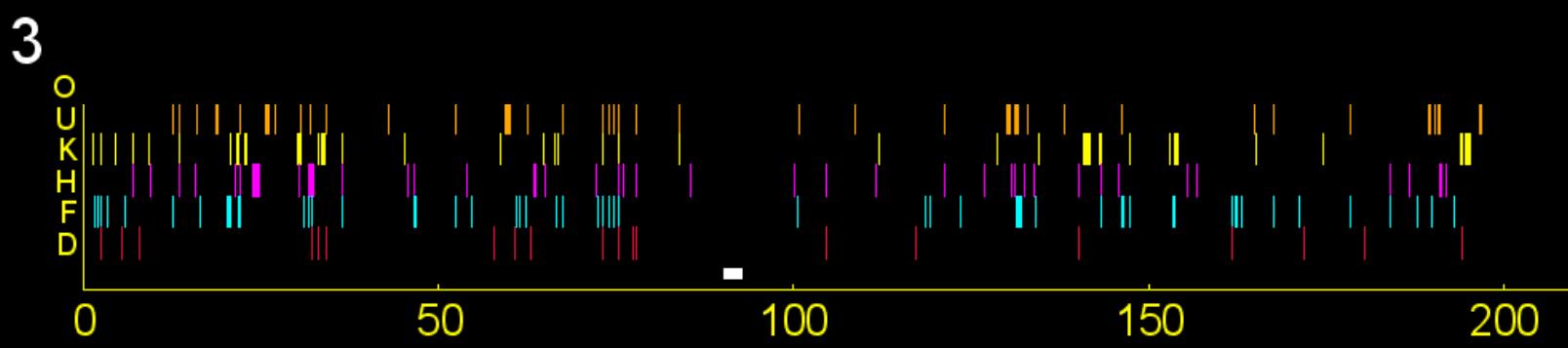
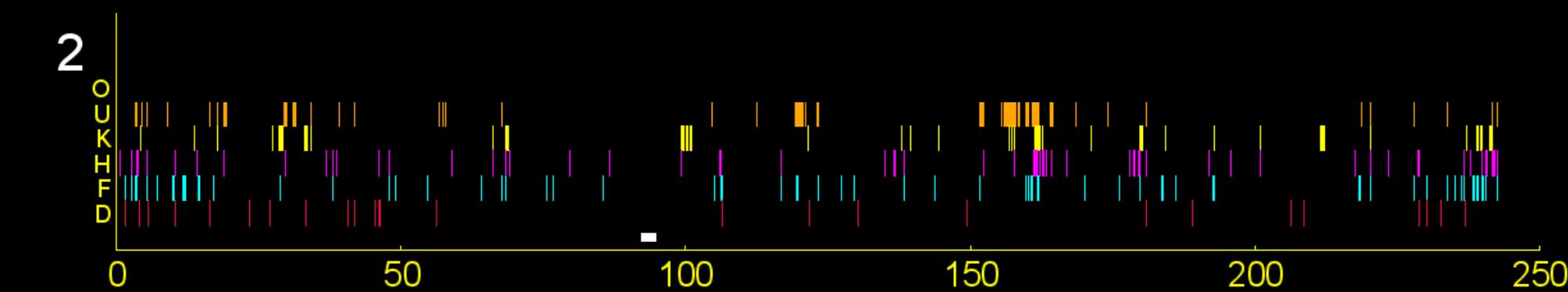
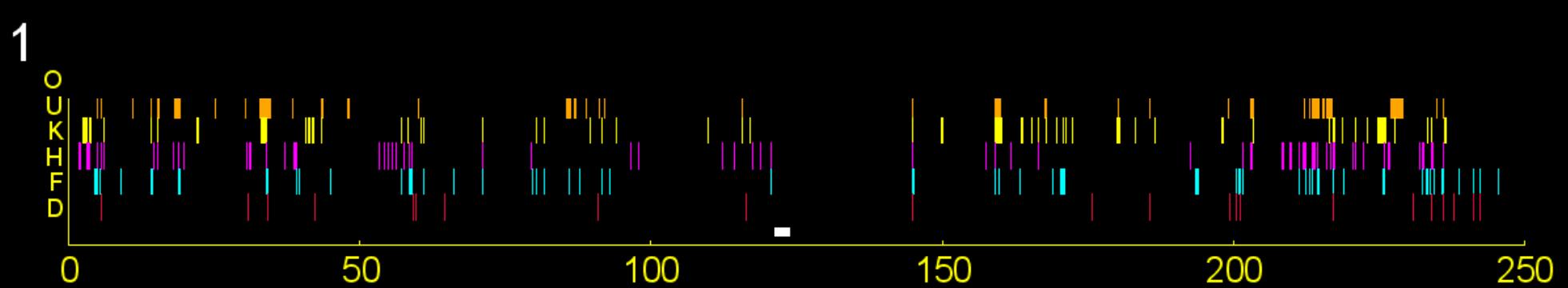
(If you like programming and would like to do a project with me, get in touch via [mp@bodkan.net](mailto:mp@bodkan.net)!)

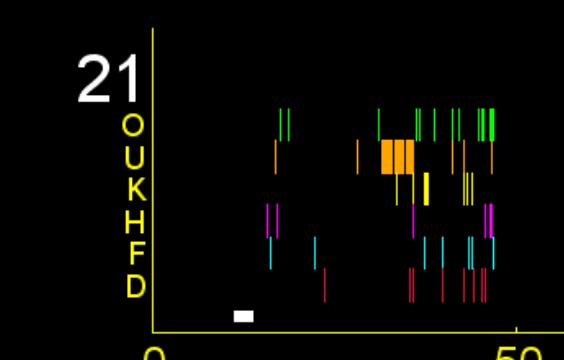
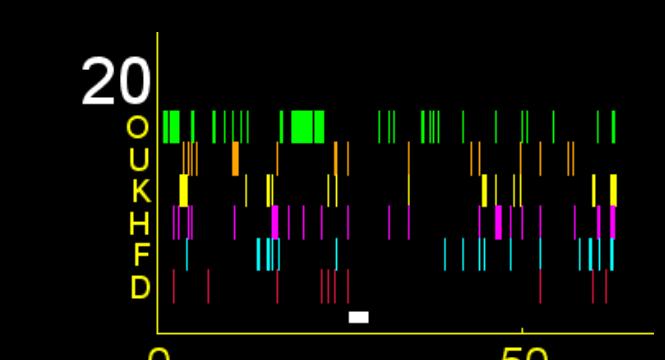
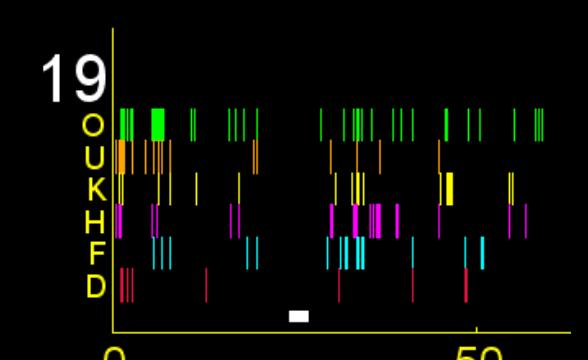
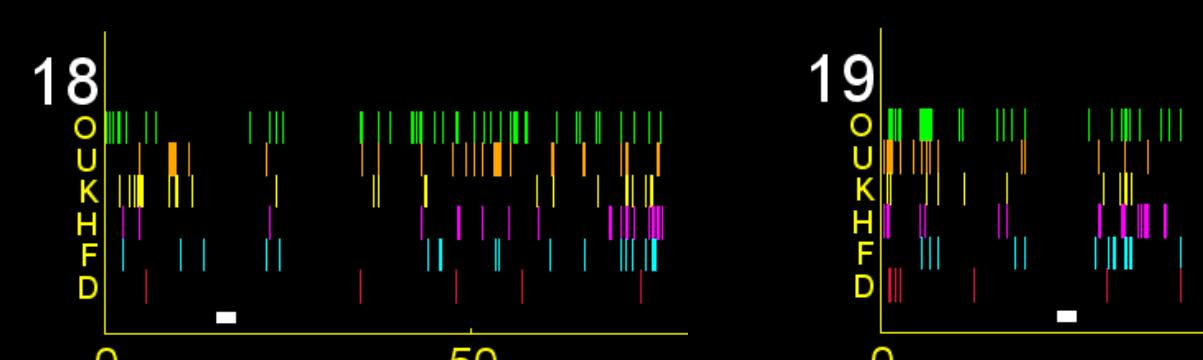
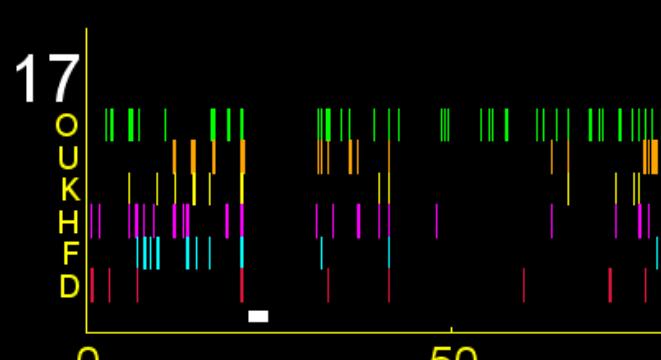
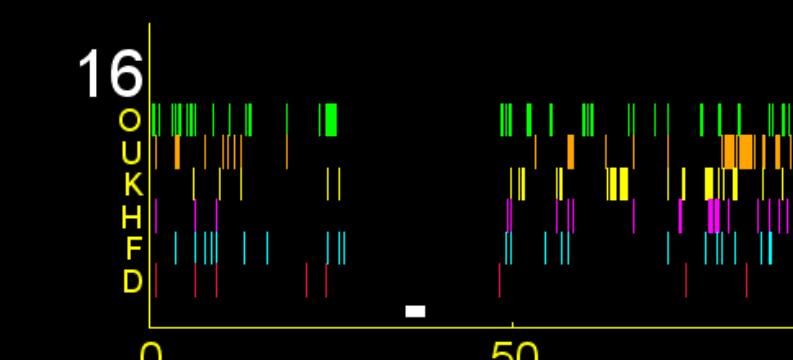
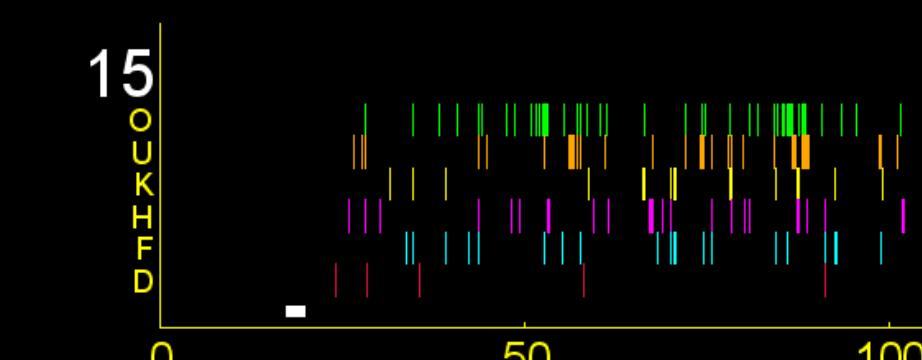
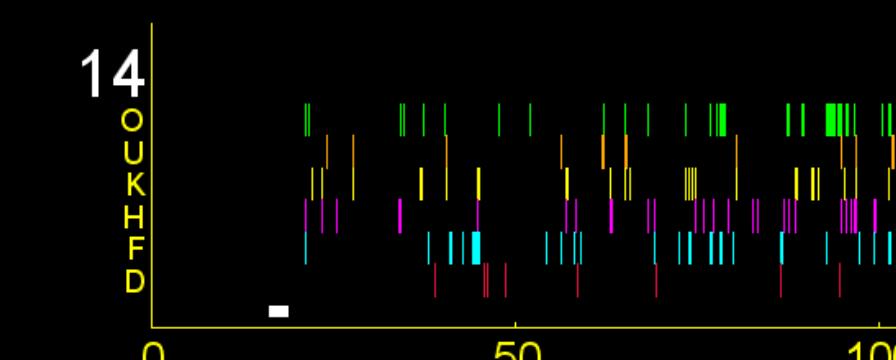
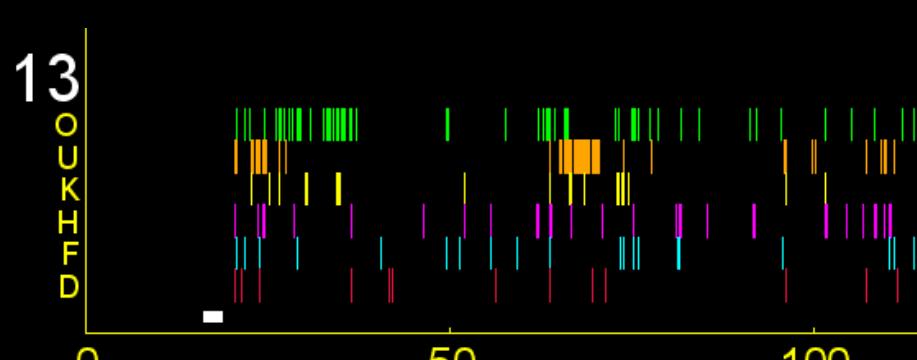
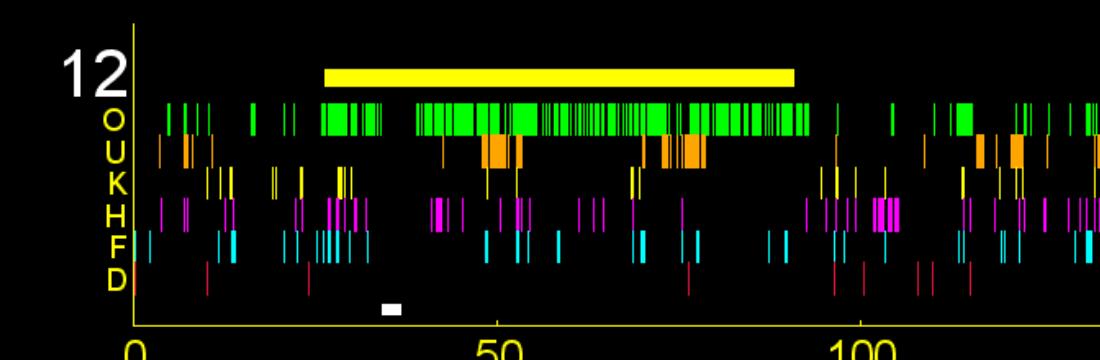
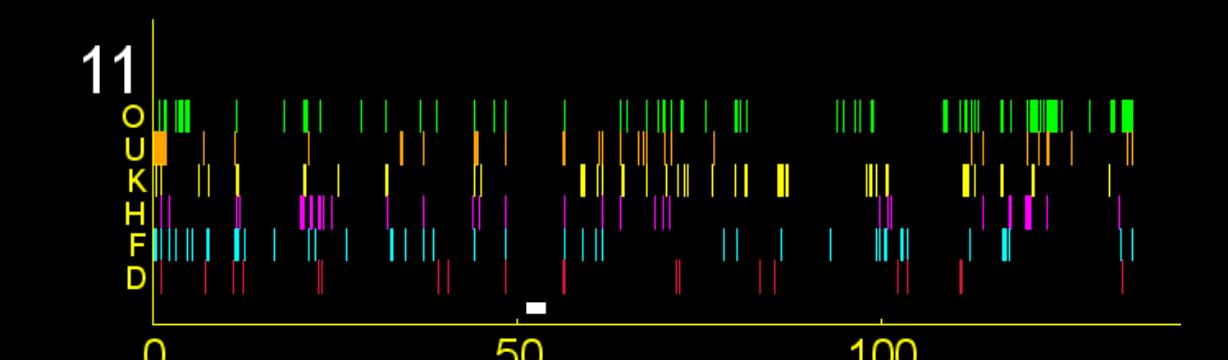
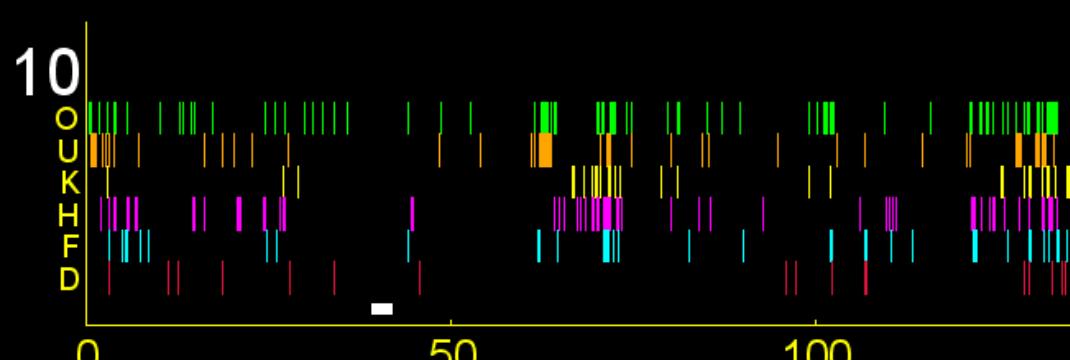
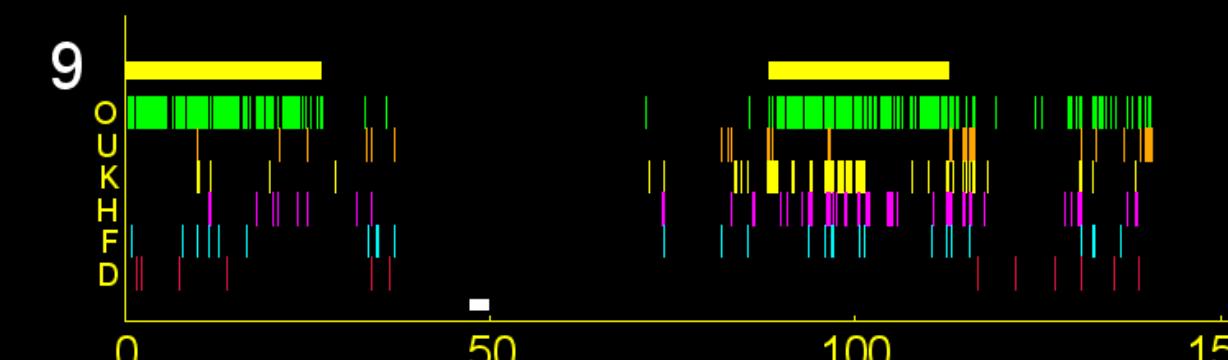
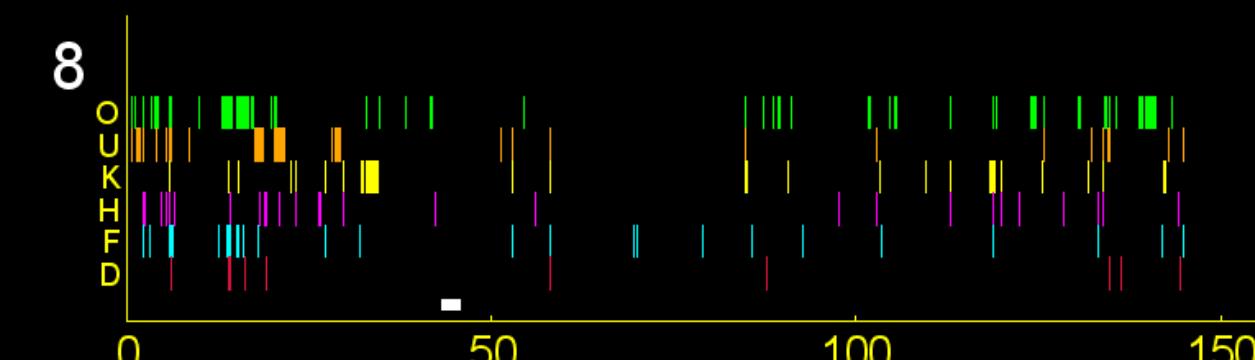
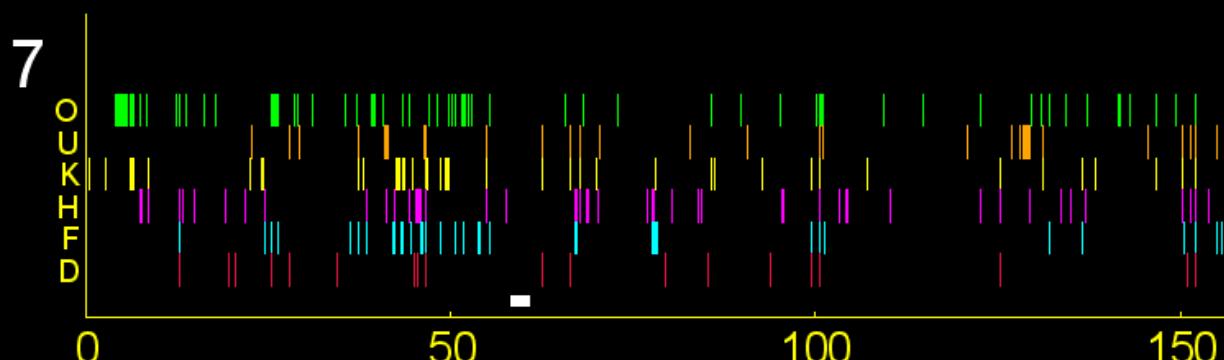
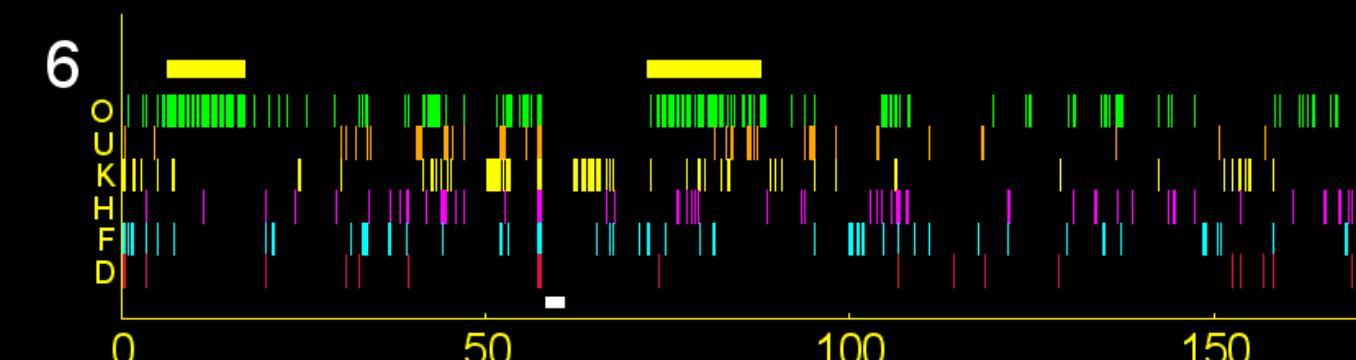
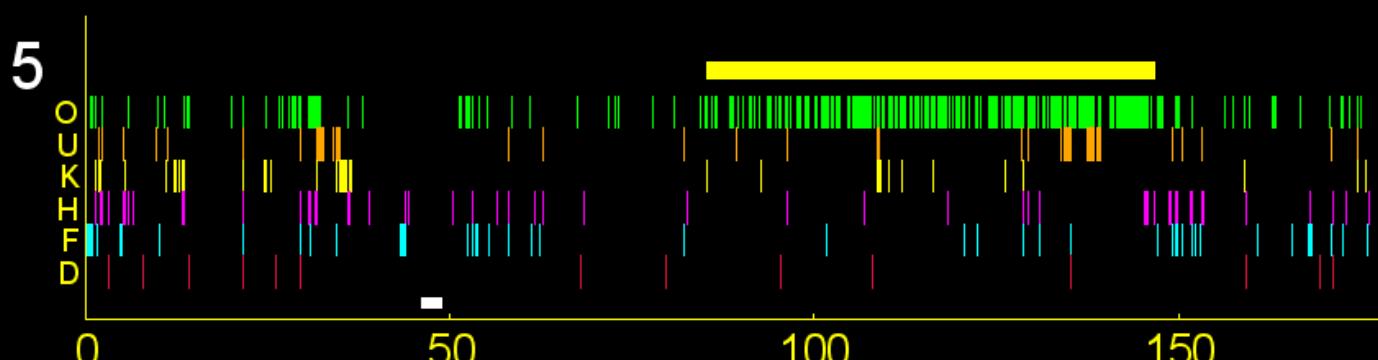
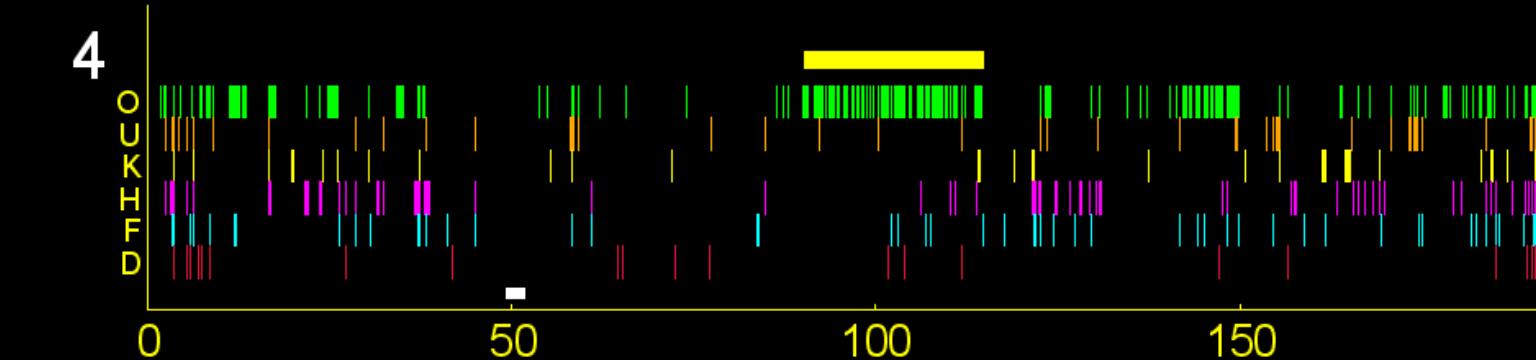
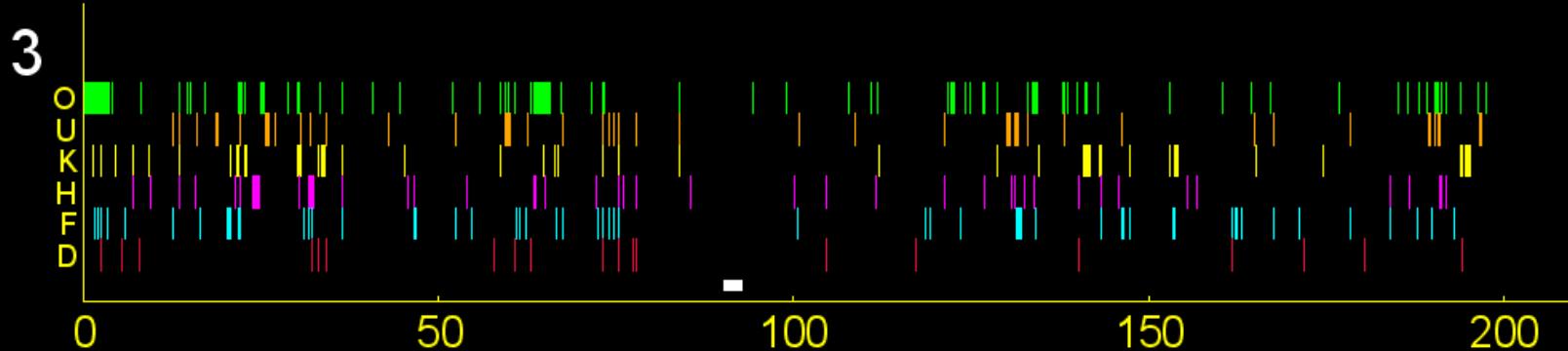
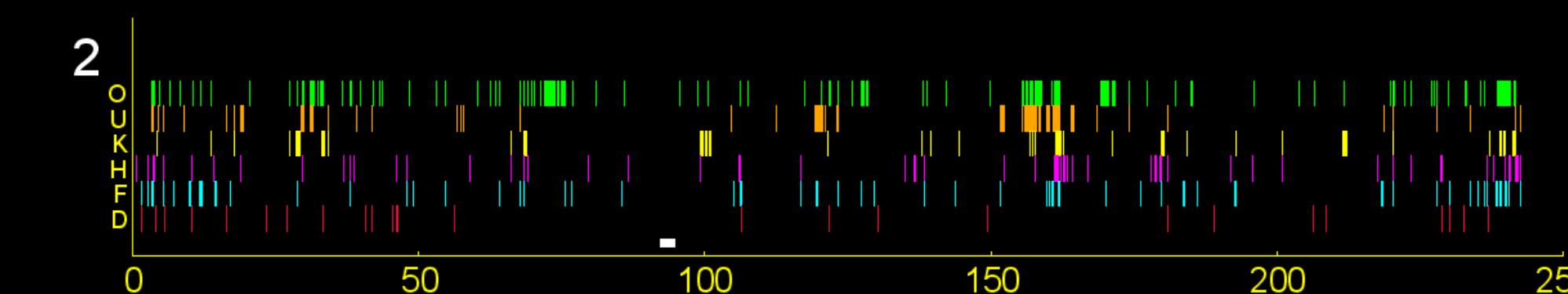
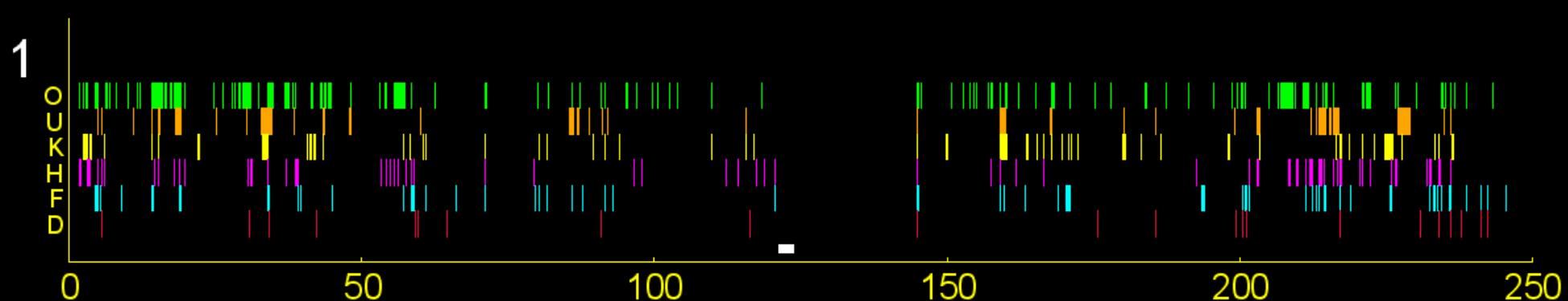


# Morphological evidence?

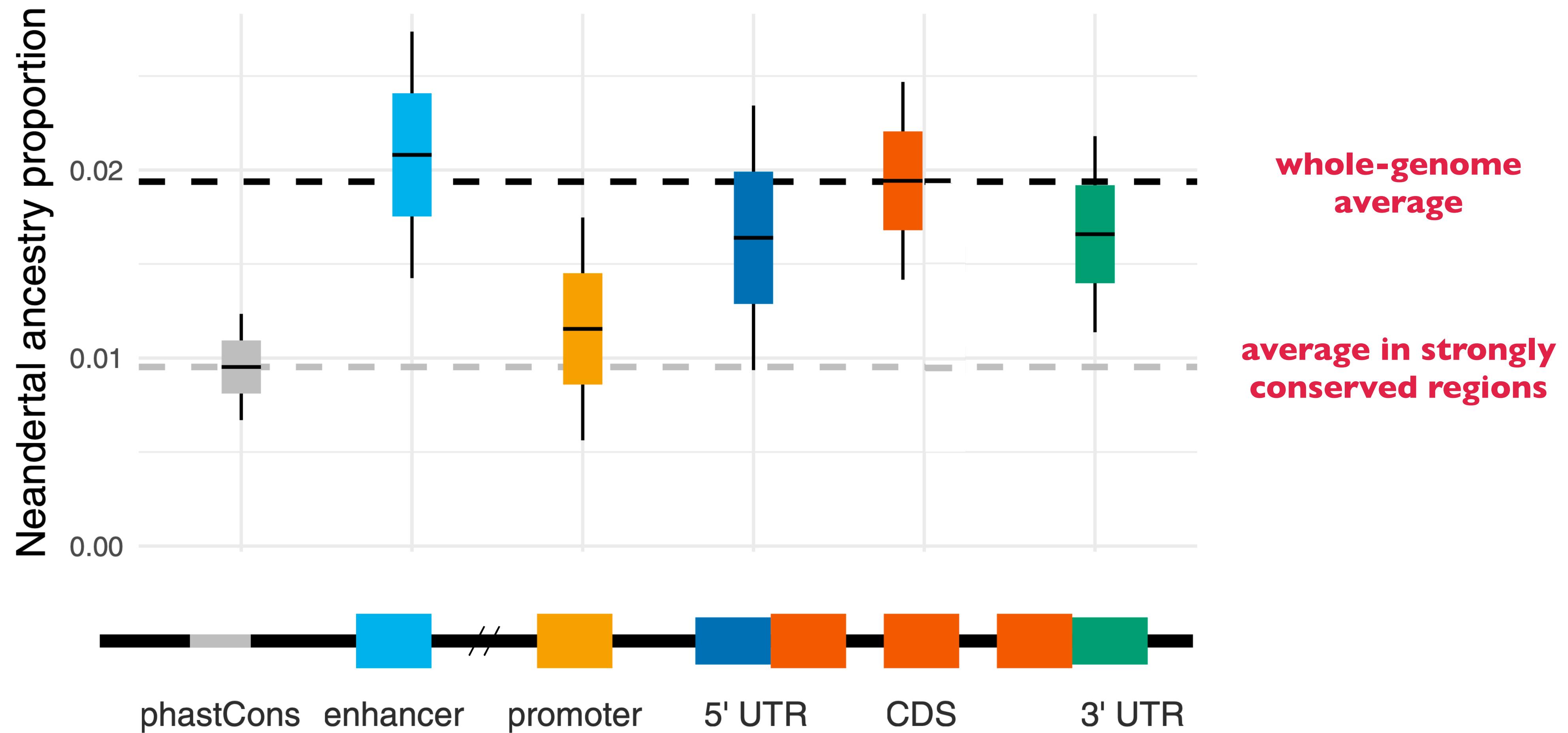


**~40 thousand years old  
remains of a modern human  
Peștera cu Oase, Romania**

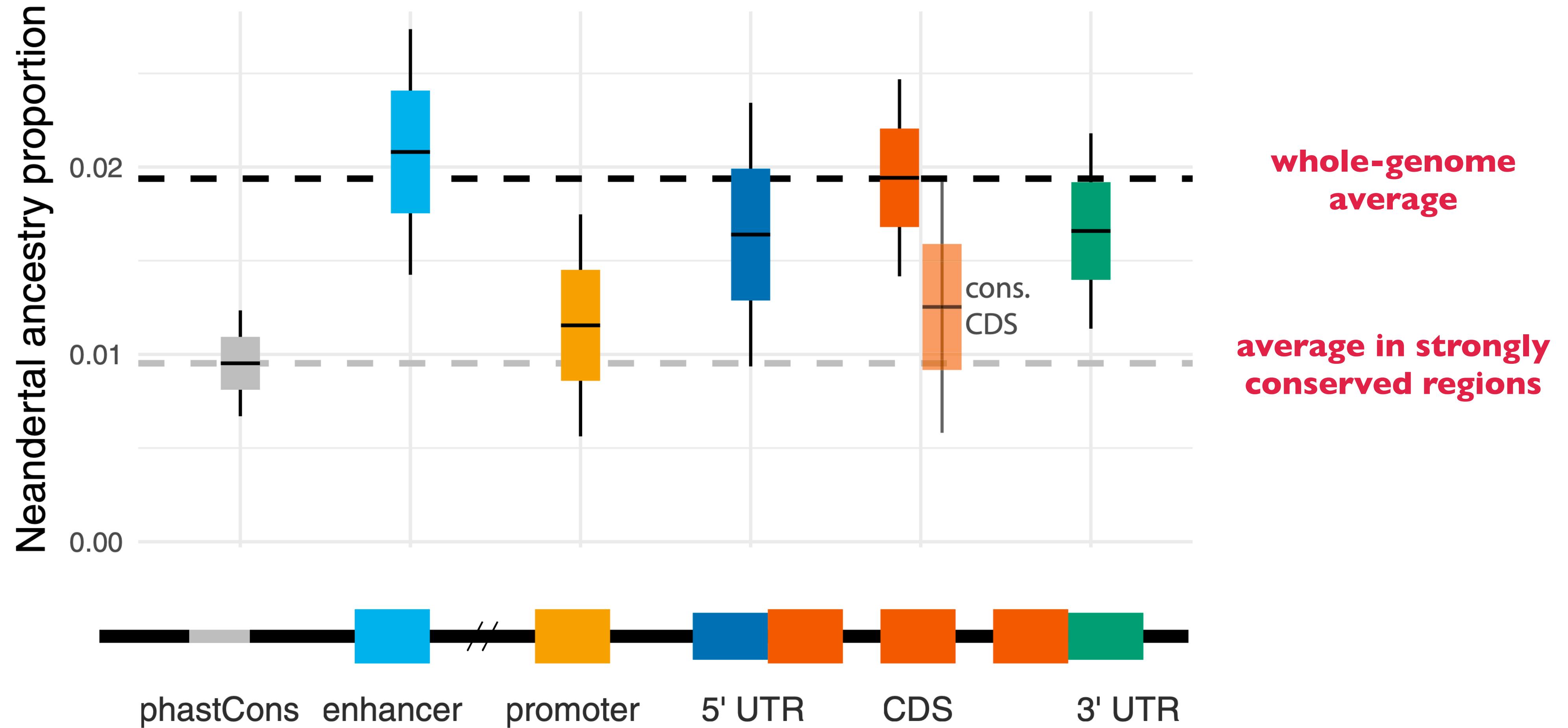




# The story is much more complicated if we zoom in...



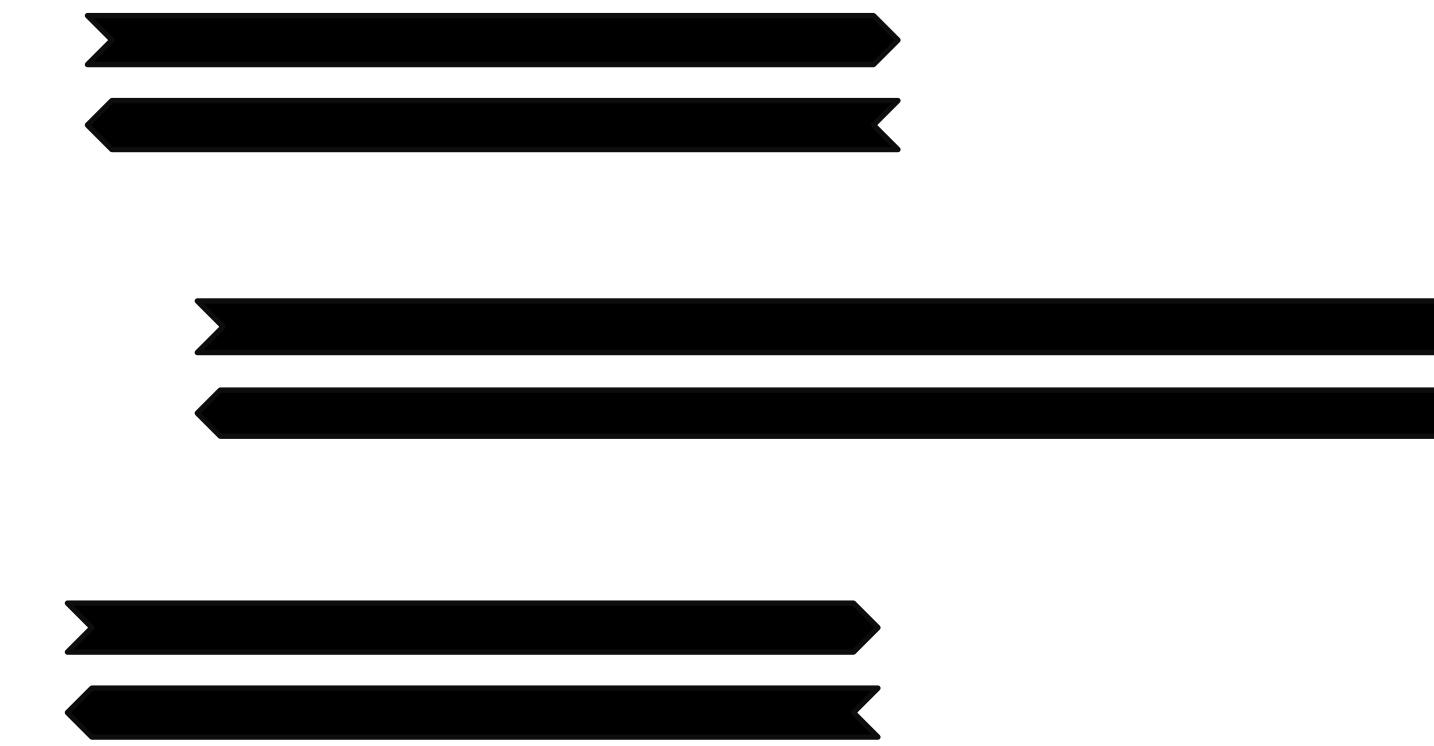
# The story is much more complicated if we zoom in...



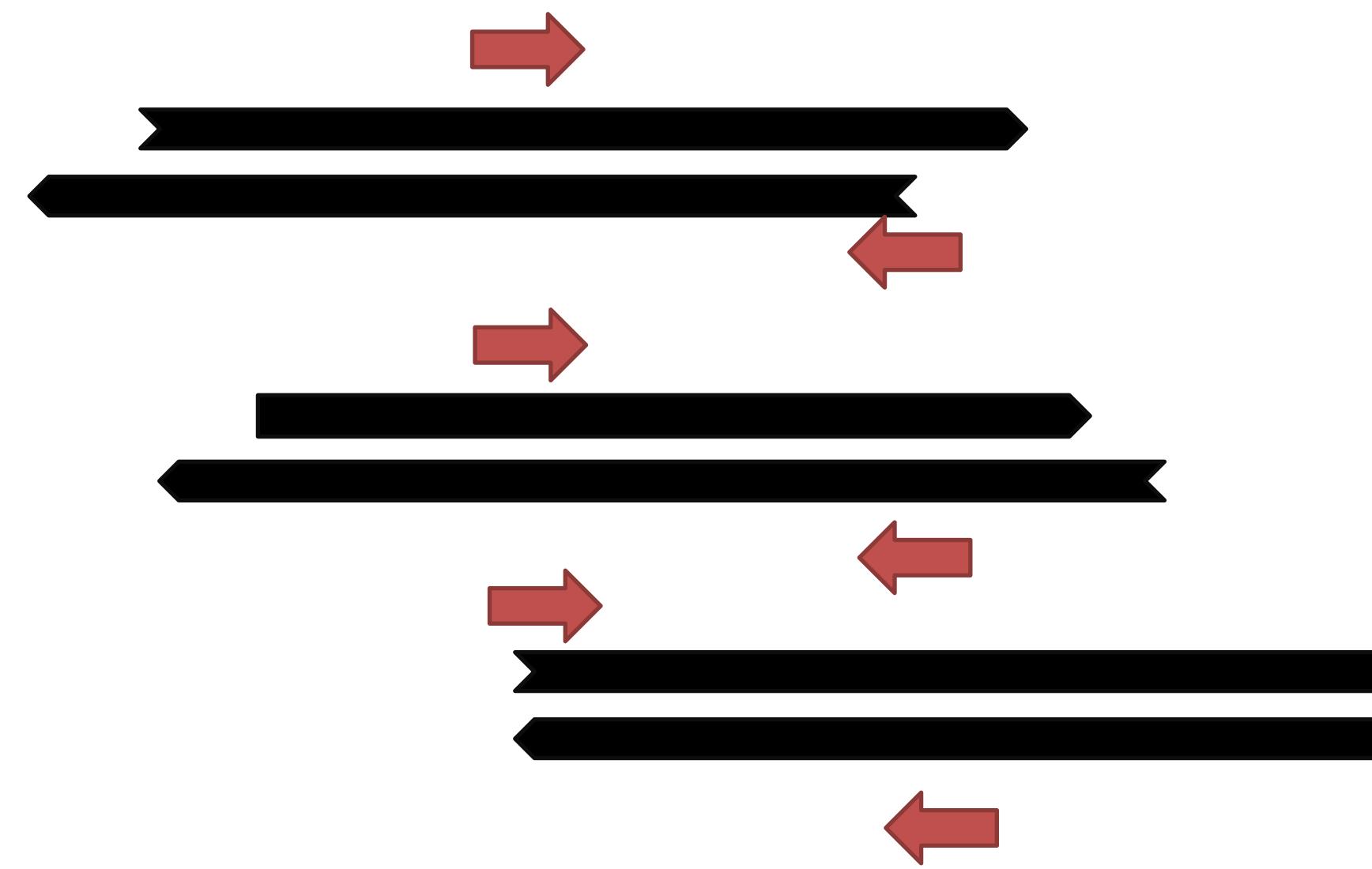
**PCR**



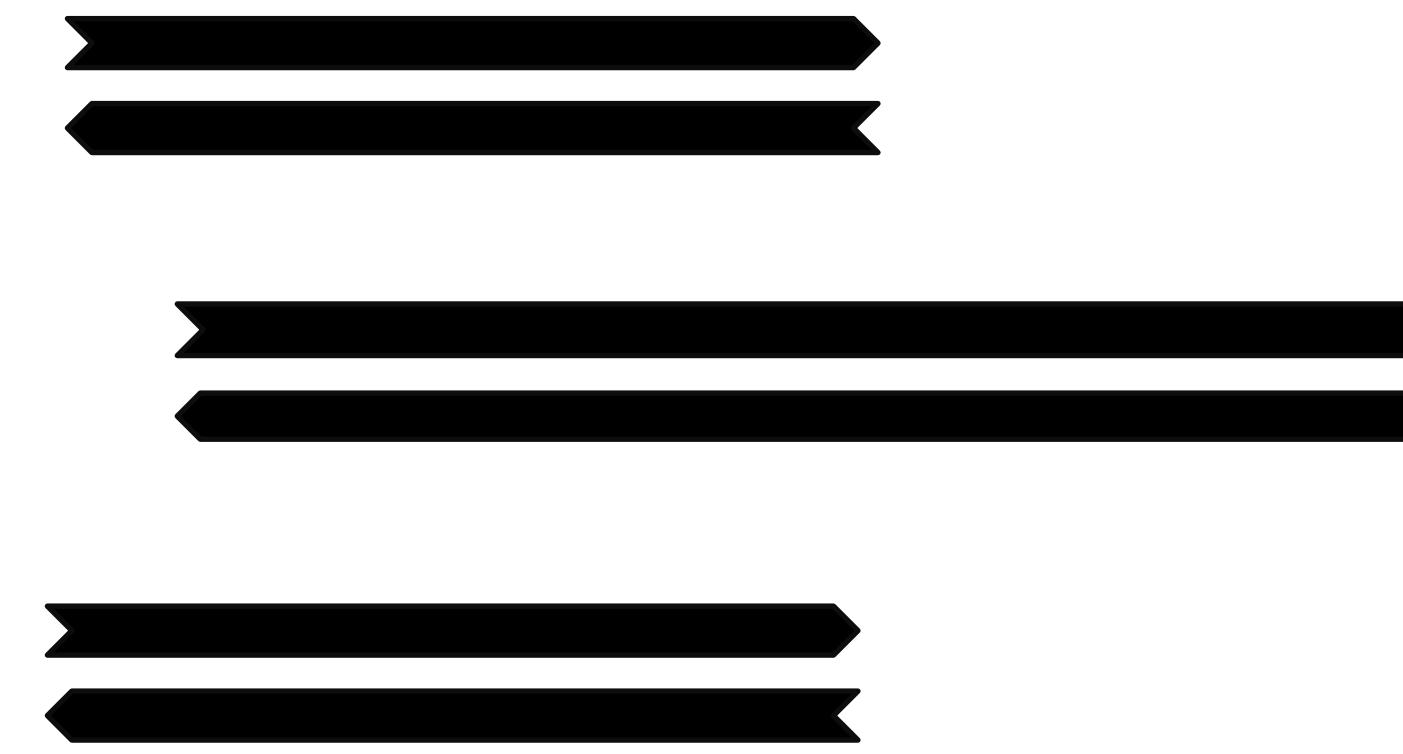
**Library preparation**



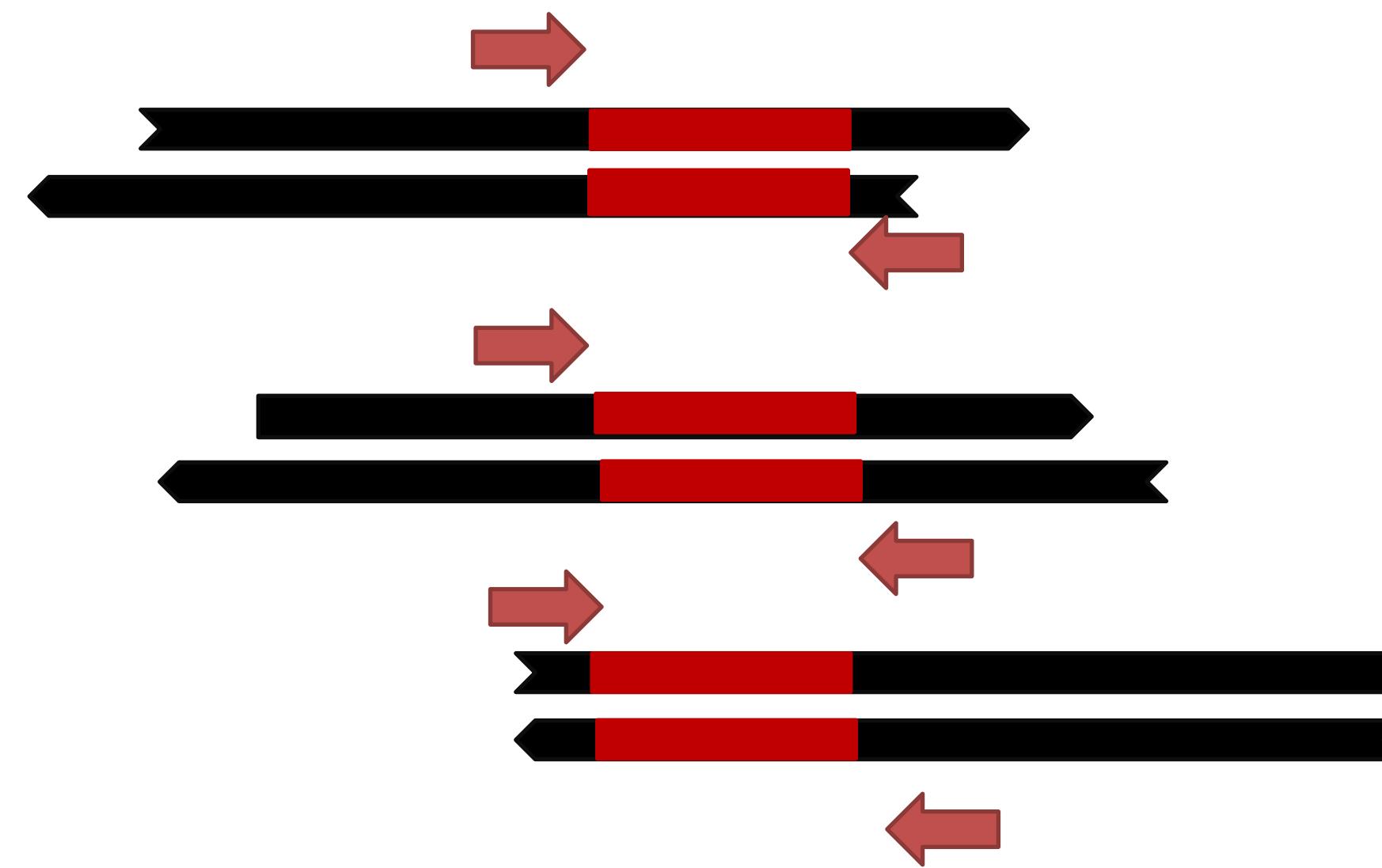
**PCR**



**Library preparation**



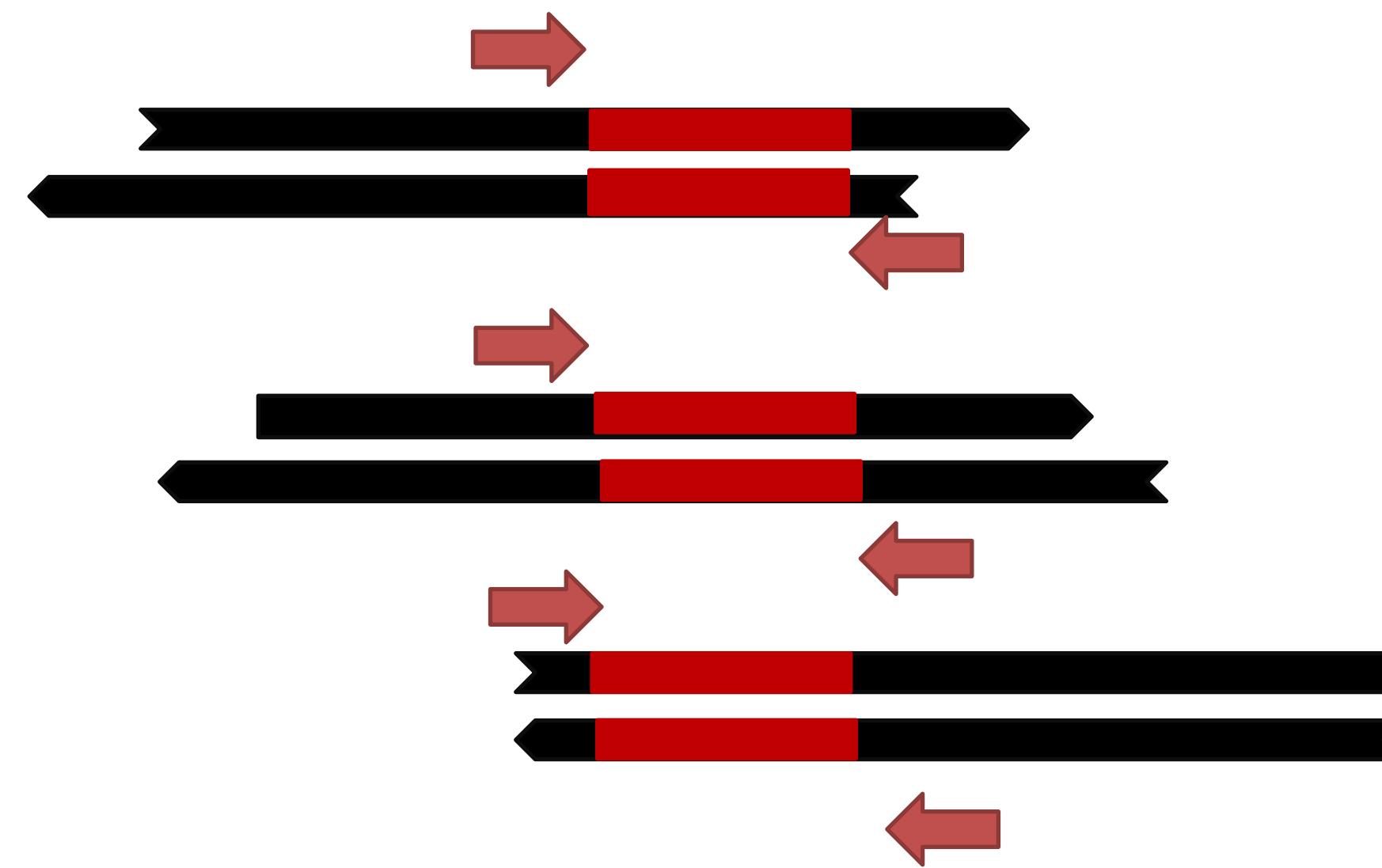
**PCR**



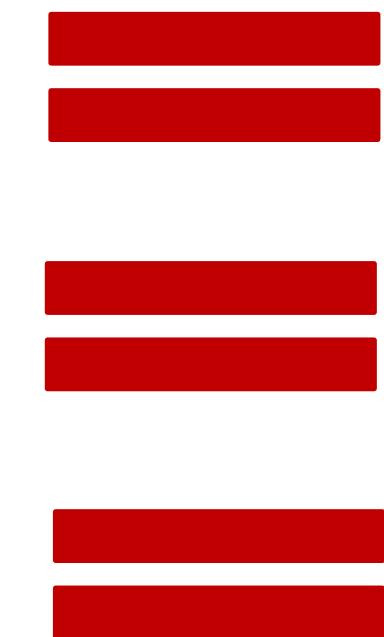
**Library preparation**



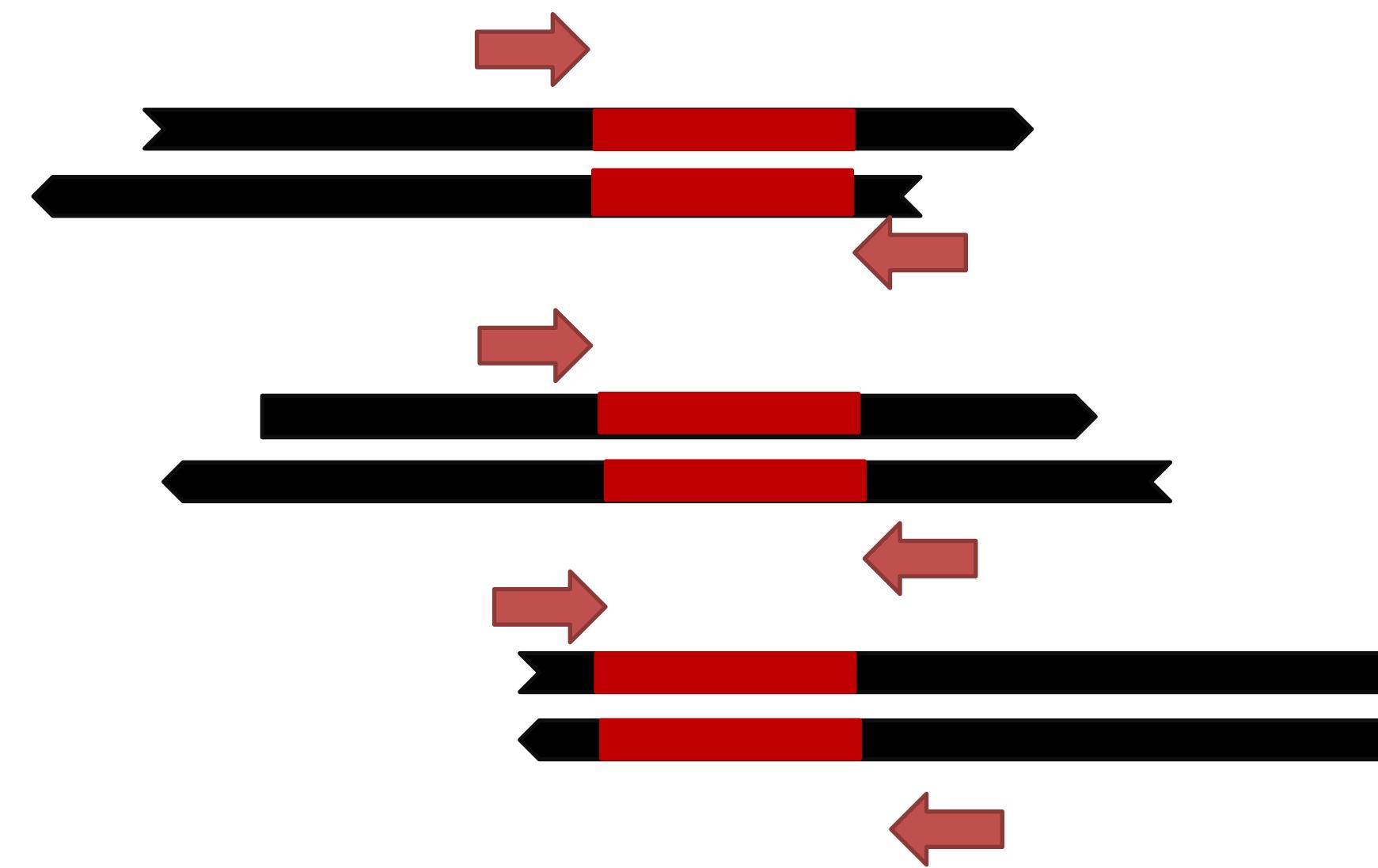
## PCR



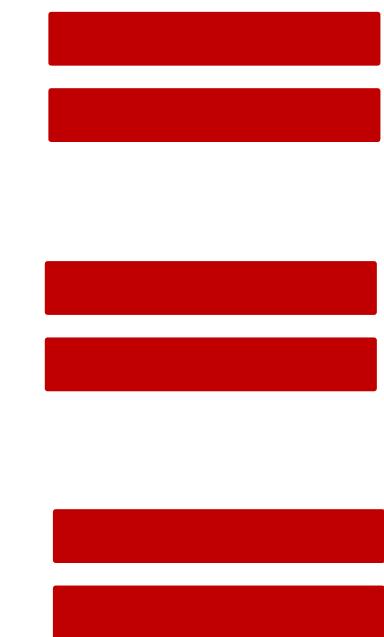
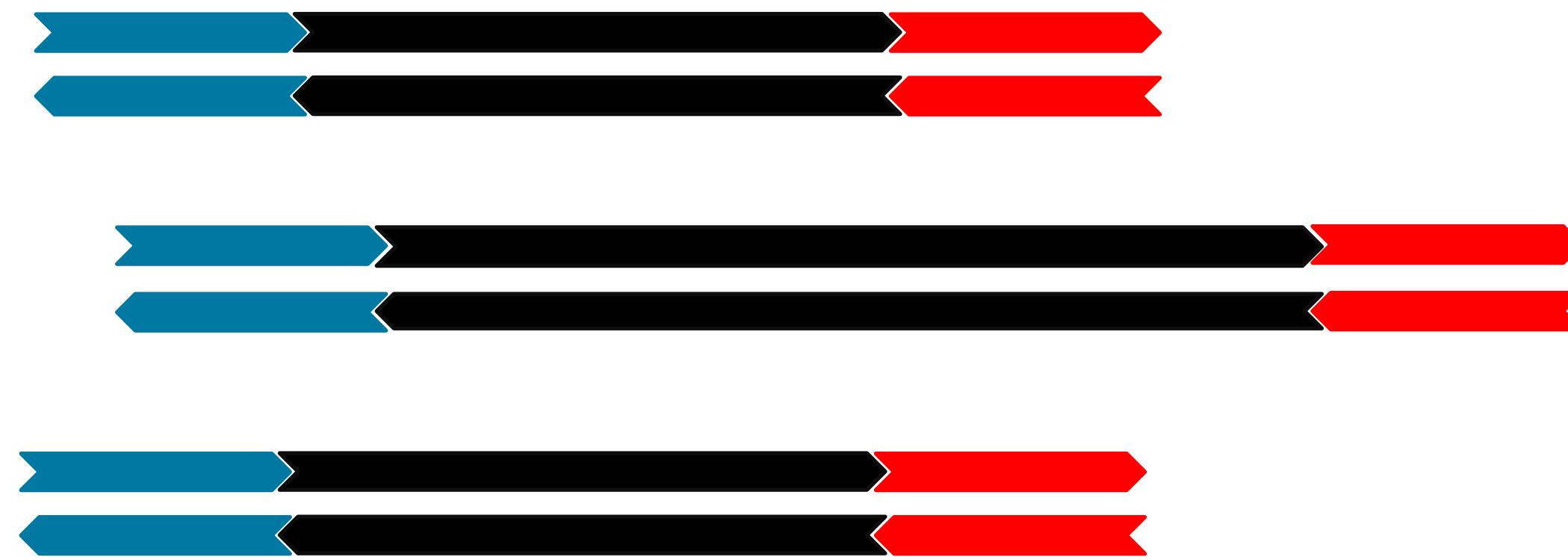
## Library preparation



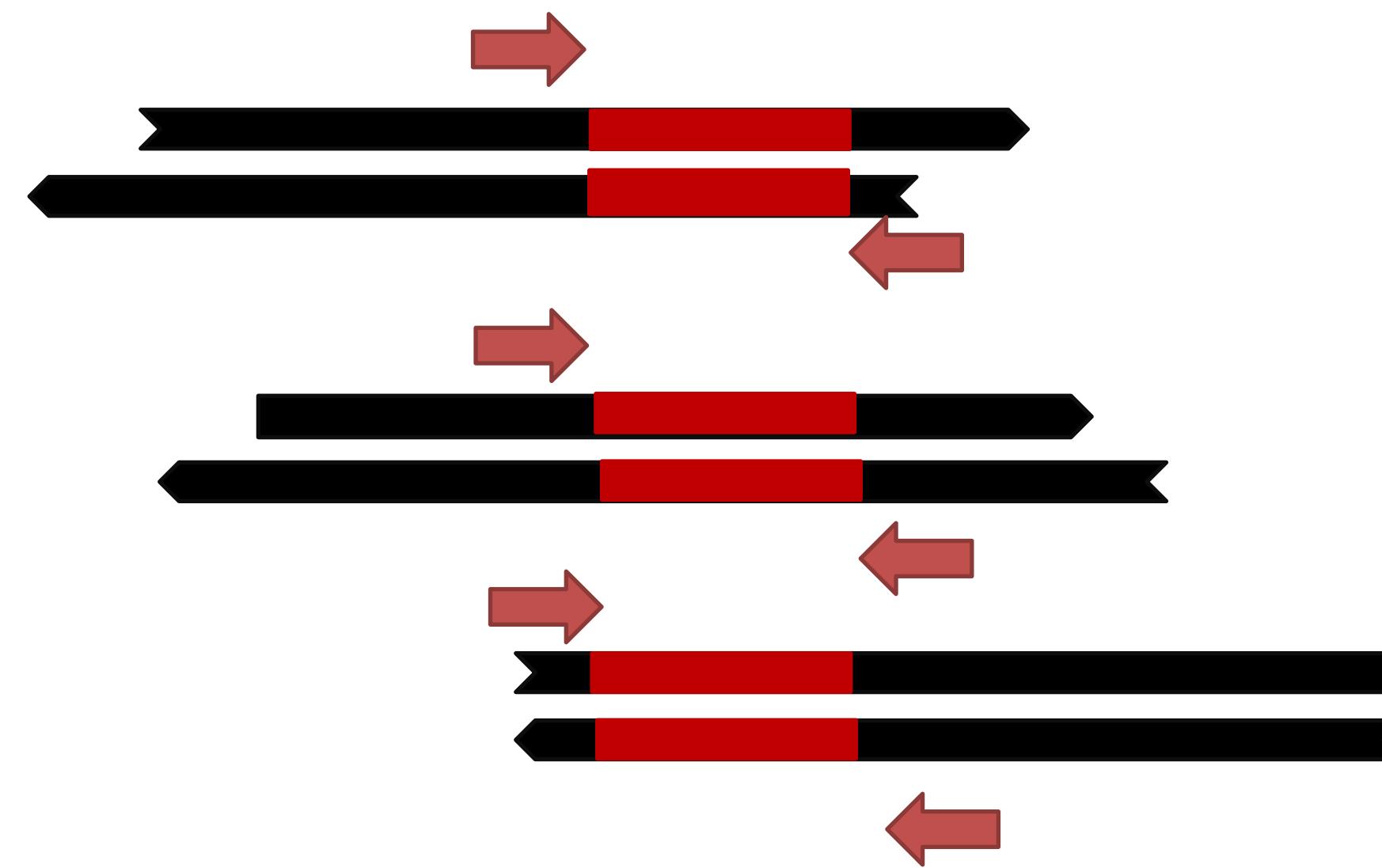
## PCR



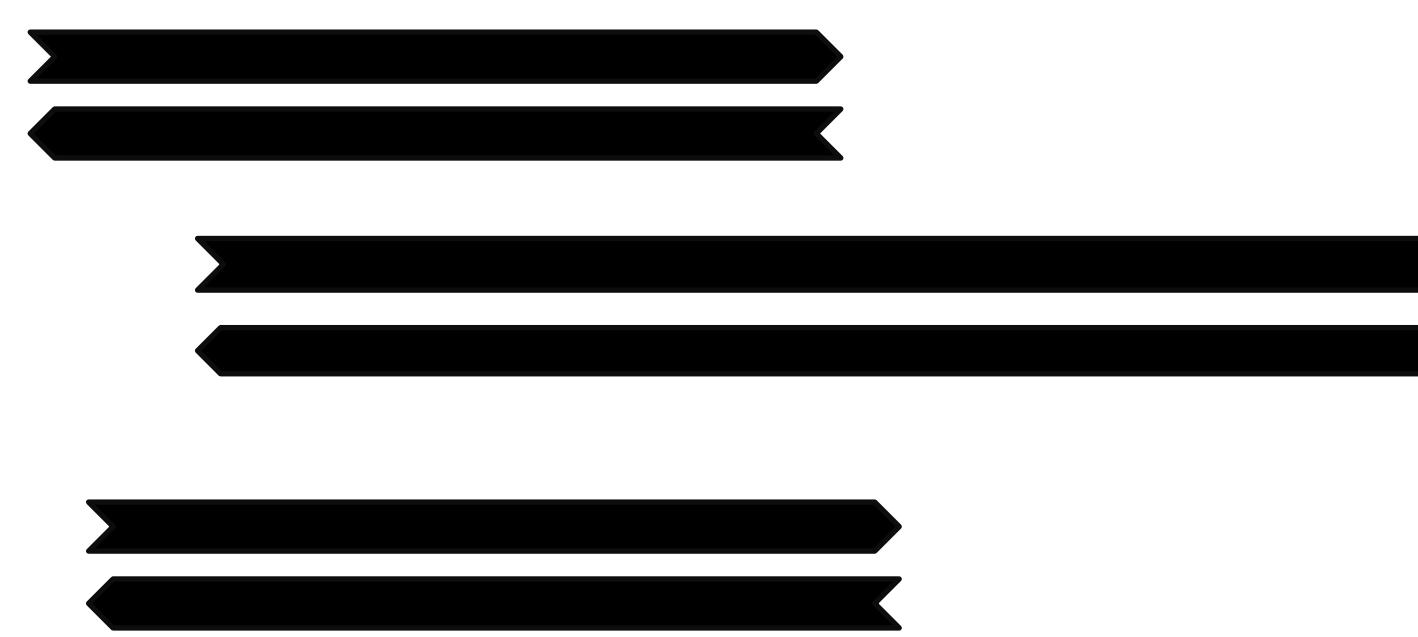
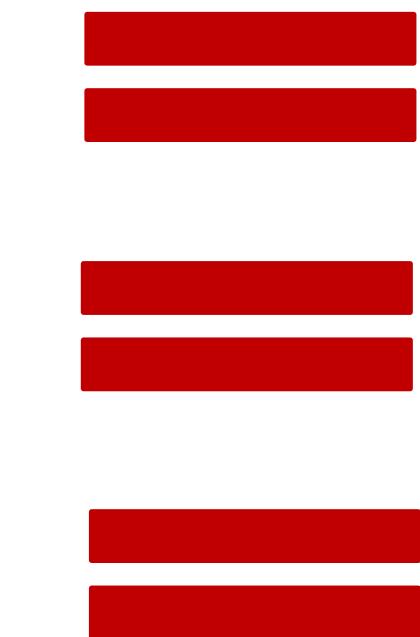
## Library preparation

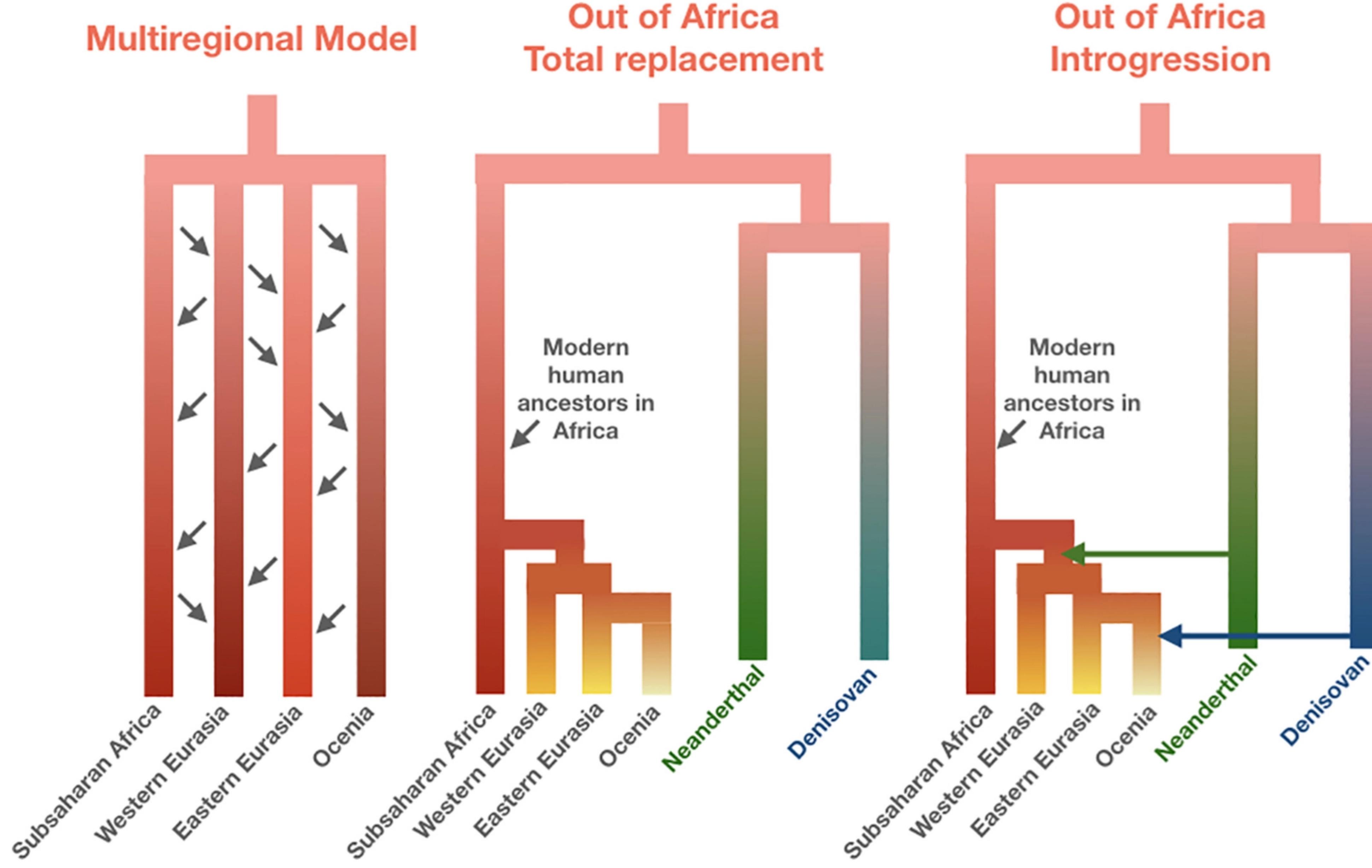


## PCR



## Library preparation





# Genomics in practice...

```
##fileformat=VCFv4.0
##source=BCM:SNPTools:hapfuse
##reference=1000Genomes-NCBI37
##FORMAT=<ID=GT,Number=1>Type=String,Description="Genotype">
##FORMAT=<ID=AP,Number=2>Type=Float,Description="Allelic Probability, P(Allele=1|Haplotype)">
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT HG00096 HG00097 HG00099 HG00100 HG00101 HG00102
12 60076 . A C 100 PASS . GT:AP 1|0:1.000,0.000 0|0:0.000,0.005 0|0:0.000,0.005
12 60252 . A G 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.300 0|1:0.010,0.590
12 60317 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 60344 . C A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.020,0.025 0|0:0.005,0.000
12 60383 . G A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 60405 . T C 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 60474 . G A 100 PASS . GT:AP 0|0:0.000,0.000 0|1:0.015,0.705 0|1:0.020,0.775
12 60614 . C A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.005 0|0:0.000,0.015
12 60628 . T C 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.070 0|0:0.000,0.000
12 60654 . G A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61021 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61107 . G T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.015 0|0:0.000,0.000
12 61172 . G A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61220 . G A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.005,0.265 0|1:0.020,0.665
12 61258 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.005,0.465 0|1:0.020,0.895
12 61272 . T C 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61329 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61341 . G A 100 PASS . GT:AP 0|0:0.000,0.000 0|1:0.010,0.560 0|1:0.020,0.855
12 61368 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|1:0.020,0.630 0|1:0.020,0.955
12 61392 . T A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61405 . G C 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61411 . C A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.015
12 61416 . G A 100 PASS . GT:AP 0|0:0.000,0.025 0|0:0.000,0.010 0|0:0.015,0.075
12 61422 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.105 0|0:0.005,0.010
12 61476 . C G 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.005,0.000
12 61510 . G A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61516 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.005 0|0:0.000,0.000
12 61552 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61604 . T G 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61687 . G A 100 PASS . GT:AP 1|0:1.000,0.000 0|1:0.015,0.625 0|1:0.020,0.960
12 61700 . C T 100 PASS . GT:AP 0|0:0.005,0.000 0|0:0.000,0.035 0|0:0.025,0.060
```

# Genomics in practice...

```
##fileformat=VCFv4.0
##source=BCM:SNPTools:hapfuse
##reference=1000Genomes-NCBI37
##FORMAT=<ID=GT,Number=1>Type=String,Description="Genotype">
##FORMAT=<ID=AP,Number=2>Type=Float,Description="Allelic Probability, P(A Allele=1|Haplotype)">
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT HG00096 HG00097 HG00099 HG00100 HG00101 HG00102
12 60076 . A C 100 PASS . GT:AP 1|0:1.000,0.000 0|0:0.000,0.005 0|0:0.000,0.005
12 60252 . A G 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.300 0|1:0.010,0.590
12 60317 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 60344 . C A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.020,0.025 0|0:0.005,0.000
12 60383 . G A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 60405 . T C 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 60474 . G A 100 PASS . GT:AP 0|0:0.000,0.000 0|1:0.015,0.705 0|1:0.020,0.775
12 60614 . C A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.005 0|0:0.000,0.015
12 60628 . T C 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.070 0|0:0.000,0.000
12 60654 . G A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61021 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61107 . G T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.015 0|0:0.000,0.000
12 61172 . G A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61220 . G A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.005,0.265 0|1:0.020,0.665
12 61258 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.005,0.465 0|1:0.020,0.895
12 61272 . T C 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61329 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61341 . G A 100 PASS . GT:AP 0|0:0.000,0.000 0|1:0.010,0.560 0|1:0.020,0.855
12 61368 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|1:0.020,0.630 0|1:0.020,0.955
12 61392 . T A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61405 . G C 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61411 . C A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.015
12 61416 . G A 100 PASS . GT:AP 0|0:0.000,0.025 0|0:0.000,0.010 0|0:0.015,0.075
12 61422 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.105 0|0:0.005,0.010
12 61476 . C G 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.005,0.000
12 61510 . G A 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61516 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.005 0|0:0.000,0.000
12 61552 . C T 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61604 . T G 100 PASS . GT:AP 0|0:0.000,0.000 0|0:0.000,0.000 0|0:0.000,0.000
12 61687 . G A 100 PASS . GT:AP 1|0:1.000,0.000 0|1:0.015,0.625 0|1:0.020,0.960
12 61700 . C T 100 PASS . GT:AP 0|0:0.005,0.000 0|0:0.000,0.035 0|0:0.025,0.060
```

chromosome position

# Genomics in practice...

#fileformat=VCFv4.0 ##source=BCM:SNPTools:hapfuse ##reference=1000Genomes-NCBI37 ##FORMAT=<ID=GT,Number=1>Type=String,Description="Genotype"> ##FORMAT=<ID=AP,Number=2>Type=Float,Description="Allelic Probability, P(Allele=1 Haplotype)">														
#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	HG00096	HG00097	HG00099	HG00100	HG00101	HG00102
12	60076	.	A	C	100	PASS	.	GT:AP	1 0:1.000,0.000	0 0:0.000,0.005	0 0:0.000,0.005			
12	60252	.	A	G	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.300	0 1:0.010,0.590			
12	60317	.	C	T	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.000	0 0:0.000,0.000			
12	60344	.	C	A	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.020,0.025	0 0:0.005,0.000			
12	60383	.	G	A	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.000	0 0:0.000,0.000			
12	60405	.	T	C	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.000	0 0:0.000,0.000			
12	60474	.	G	A	100	PASS	.	GT:AP	0 0:0.000,0.000	0 1:0.015,0.705	0 1:0.020,0.775			
12	60614	.	C	A	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.005	0 0:0.000,0.015			
12	60628	.	T	C	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.070	0 0:0.000,0.000			
12	60654	.	G	A	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.000	0 0:0.000,0.000			
12	61021	.	C	T	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.000	0 0:0.000,0.000			
12	61107	.	G	T	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.015	0 0:0.000,0.000			
12	61172	.	G	A	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.000	0 0:0.000,0.000			
12	61220	.	G	A	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.005,0.265	0 1:0.020,0.665			
12	61258	.	C	T	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.005,0.465	0 1:0.020,0.895			
12	61272	.	T	C	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.000	0 0:0.000,0.000			
12	61329	.	C	T	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.000	0 0:0.000,0.000			
12	61341	.	G	A	100	PASS	.	GT:AP	0 0:0.000,0.000	0 1:0.010,0.560	0 1:0.020,0.855			
12	61368	.	C	T	100	PASS	.	GT:AP	0 0:0.000,0.000	0 1:0.020,0.630	0 1:0.020,0.955			
12	61392	.	T	A	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.000	0 0:0.000,0.000			
12	61405	.	G	C	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.000	0 0:0.000,0.000			
12	61411	.	C	A	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.000	0 0:0.000,0.015			
12	61416	.	G	A	100	PASS	.	GT:AP	0 0:0.000,0.025	0 0:0.000,0.010	0 0:0.015,0.075			
12	61422	.	C	T	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.105	0 0:0.005,0.010			
12	61476	.	C	G	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.000	0 0:0.005,0.000			
12	61510	.	G	A	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.000	0 0:0.000,0.000			
12	61516	.	C	T	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.005	0 0:0.000,0.000			
12	61552	.	C	T	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.000	0 0:0.000,0.000			
12	61604	.	T	G	100	PASS	.	GT:AP	0 0:0.000,0.000	0 0:0.000,0.000	0 0:0.000,0.000			
12	61687	.	G	A	100	PASS	.	GT:AP	1 0:1.000,0.000	0 1:0.015,0.625	0 1:0.020,0.960			
12	61700	.	C	T	100	PASS	.	GT:AP	0 0:0.005,0.000	0 0:0.000,0.035	0 0:0.025,0.060			

chromosome position

SNP

# Genomics in practice...

VCF File Content															
#CHROM		POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	HG00096	HG00097	HG00099	HG00100	HG00101	HG00102
12	60076	.		A	C	100	PASS	.	GT:AP	1 0	1.000, 0.000	0 0	0.000, 0.005	0 0	0.000, 0.005
12	60252	.		A	G	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.300	0 1	0.010, 0.590
12	60317	.		C	T	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.000	0 0	0.000, 0.000
12	60344	.		C	A	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.020, 0.025	0 0	0.005, 0.000
12	60383	.		G	A	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.000	0 0	0.000, 0.000
12	60405	.		T	C	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.000	0 0	0.000, 0.000
12	60474	.		G	A	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 1	0.015, 0.705	0 1	0.020, 0.775
12	60614	.		C	A	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.005	0 0	0.000, 0.015
12	60628	.		T	C	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.070	0 0	0.000, 0.000
12	60654	.		G	A	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.000	0 0	0.000, 0.000
12	61021	.		C	T	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.000	0 0	0.000, 0.000
12	61107	.		G	T	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.015	0 0	0.000, 0.000
12	61172	.		G	A	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.000	0 0	0.000, 0.000
12	61220	.		G	A	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.005, 0.265	0 1	0.020, 0.665
12	61258	.		C	T	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.005, 0.465	0 1	0.020, 0.895
12	61272	.		T	C	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.000	0 0	0.000, 0.000
12	61329	.		C	T	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.000	0 0	0.000, 0.000
12	61341	.		G	A	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 1	0.010, 0.560	0 1	0.020, 0.855
12	61368	.		C	T	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 1	0.020, 0.630	0 1	0.020, 0.955
12	61392	.		T	A	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.000	0 0	0.000, 0.000
12	61405	.		G	C	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.000	0 0	0.000, 0.000
12	61411	.		C	A	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.000	0 0	0.000, 0.015
12	61416	.		G	A	100	PASS	.	GT:AP	0 0	0.000, 0.025	0 0	0.000, 0.010	0 0	0.015, 0.075
12	61422	.		C	T	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.105	0 0	0.005, 0.010
12	61476	.		C	G	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.000	0 0	0.005, 0.000
12	61510	.		G	A	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.000	0 0	0.000, 0.000
12	61516	.		C	T	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.005	0 0	0.000, 0.000
12	61552	.		C	T	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.000	0 0	0.000, 0.000
12	61604	.		T	G	100	PASS	.	GT:AP	0 0	0.000, 0.000	0 0	0.000, 0.000	0 0	0.000, 0.000
12	61687	.		G	A	100	PASS	.	GT:AP	1 0	1.000, 0.000	0 1	0.015, 0.625	0 1	0.020, 0.960
12	61700	.		C	T	100	PASS	.	GT:AP	0 0	0.005, 0.000	0 0	0.000, 0.035	0 0	0.025, 0.060

chromosome position

SNP

genotype ind. 1 genotype ind. 2 genotype ind. 3 ...