Archaic human introgression

Population genetics course

Martin Petr
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’

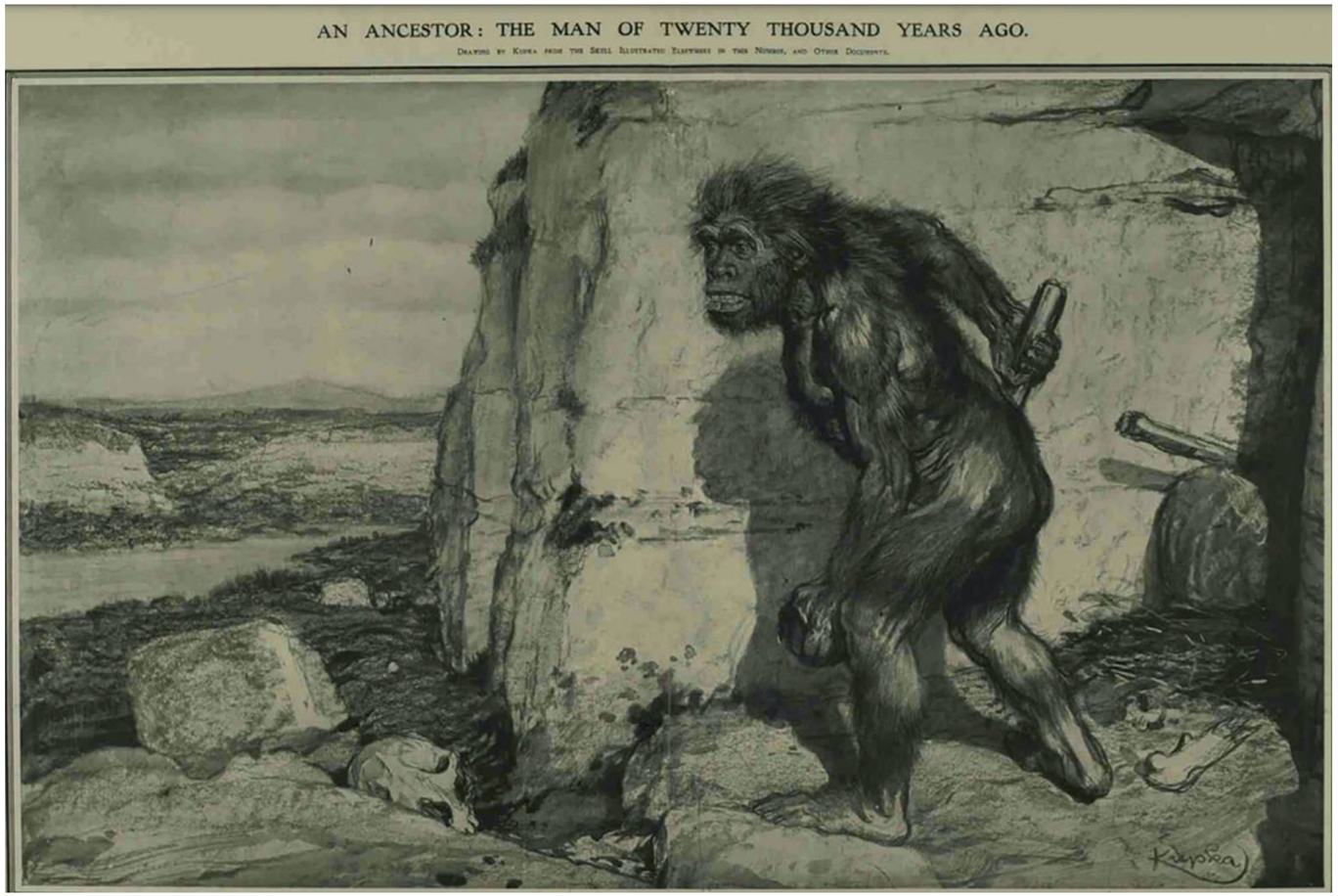
https://www.donsmaps.com/neanderthaloriginal.html

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

Neander’s valley

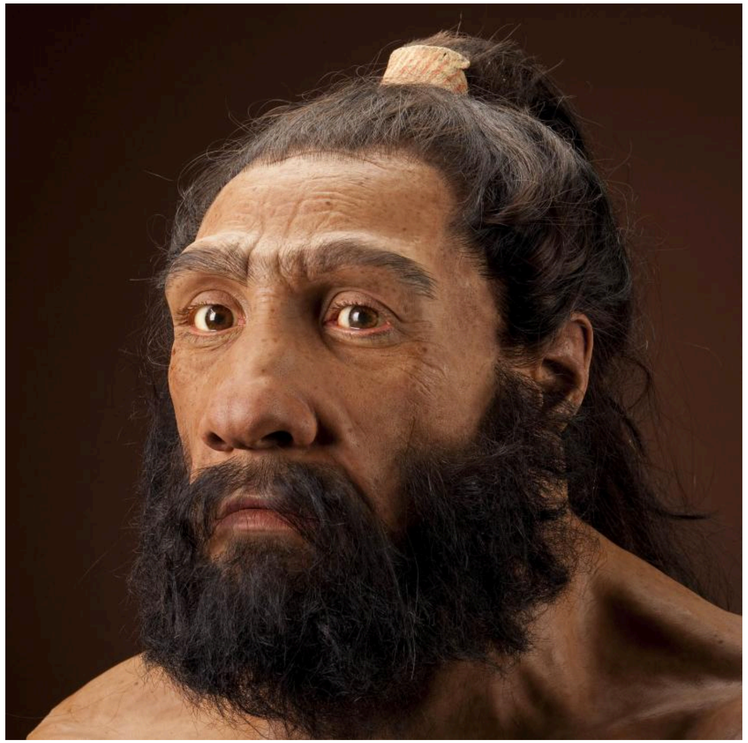
valley = Thal in German

→ **Neanderthal**



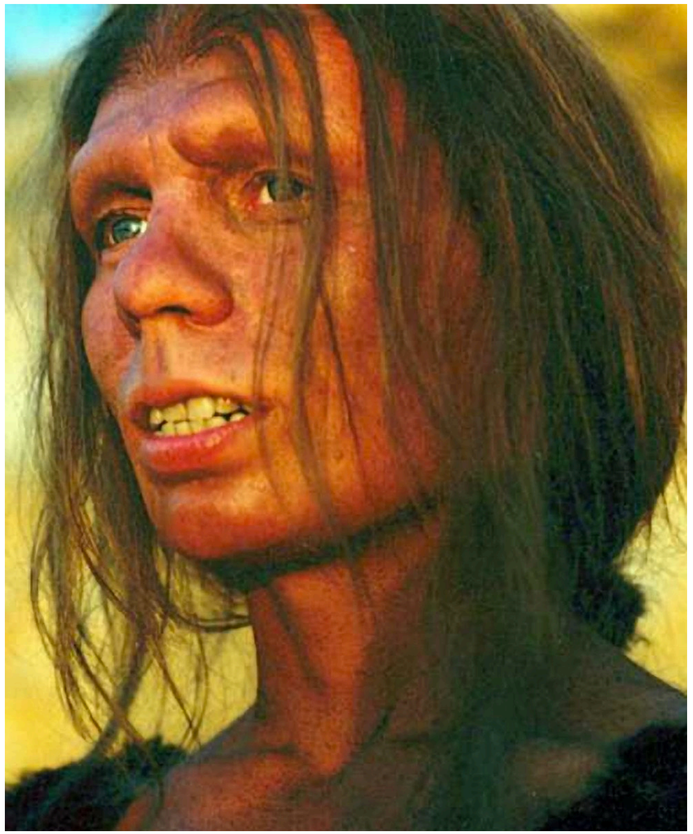
The Illustrated London News, 1909

The Neanderthal Man movie poster, 1953



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

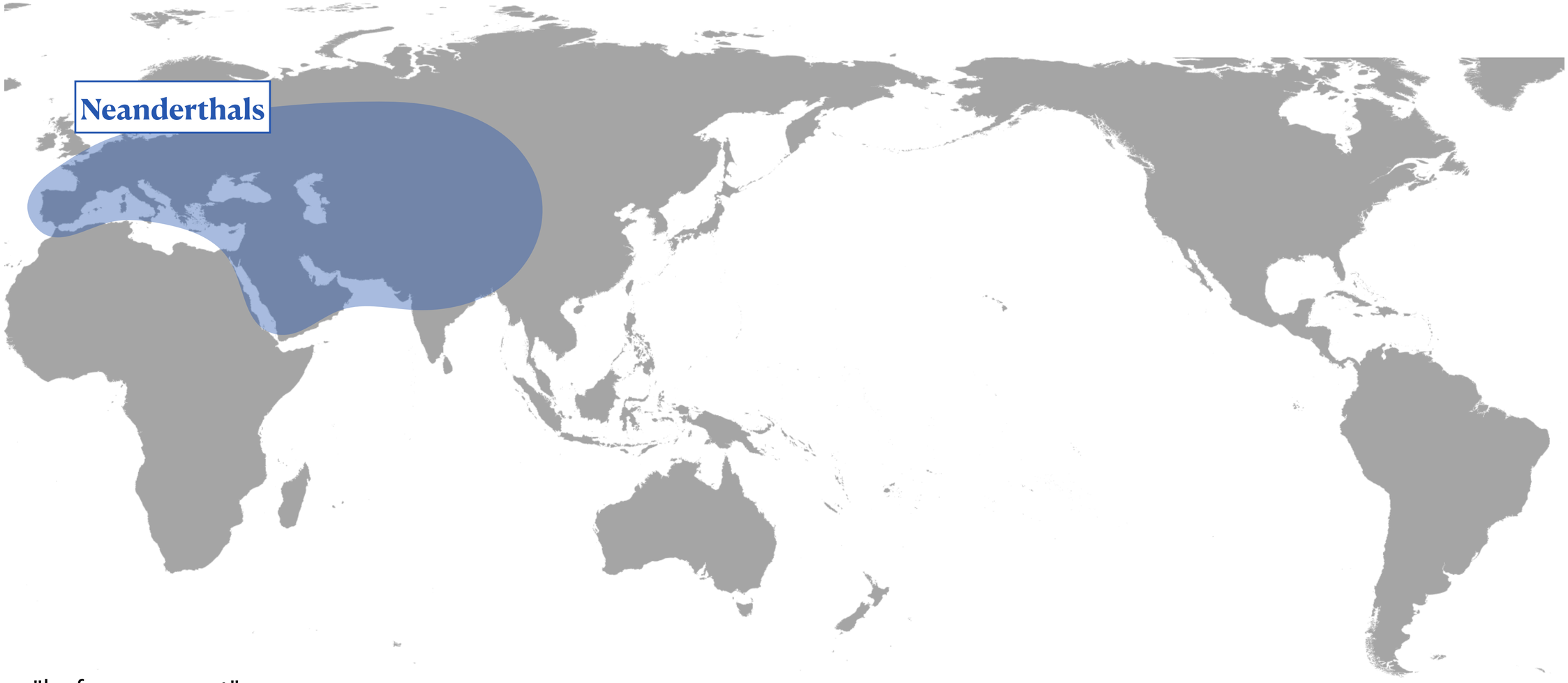
https://www.neanderthal.de



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

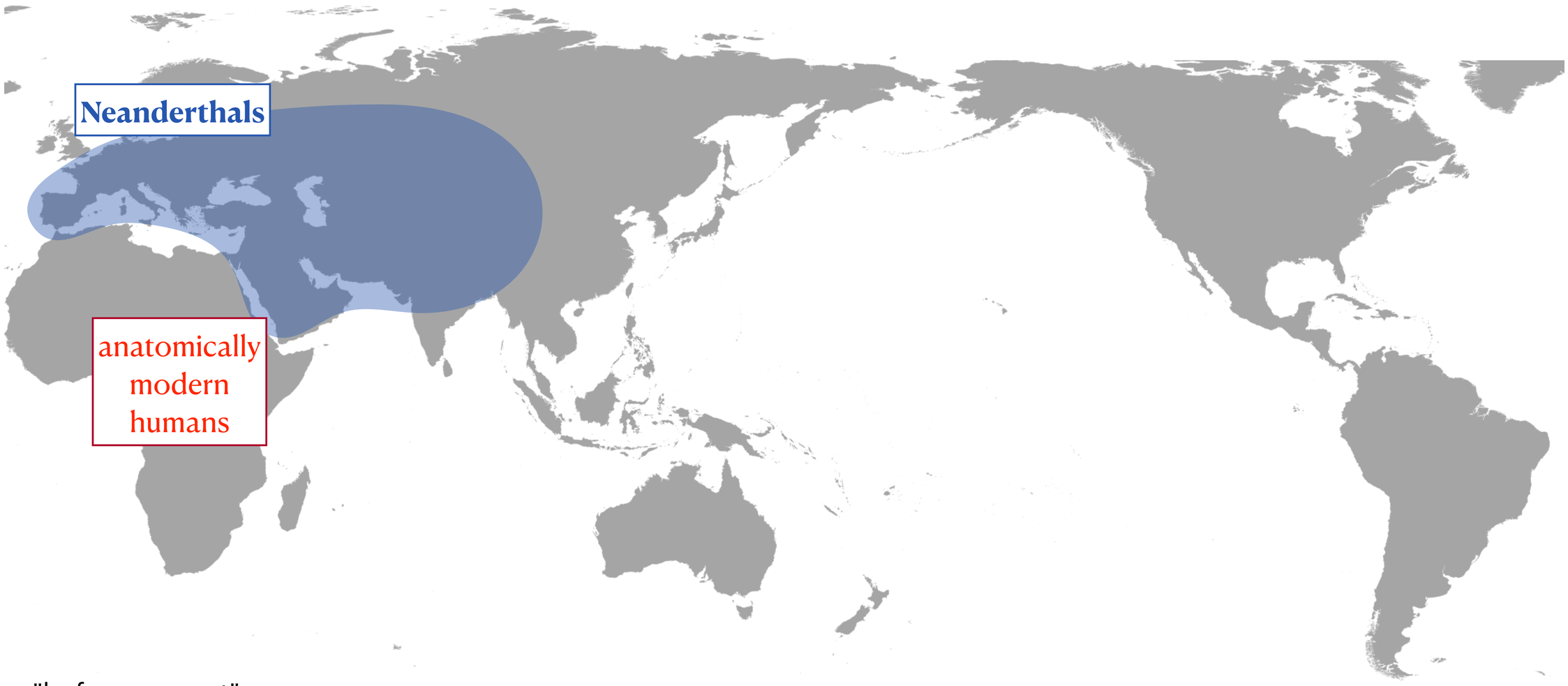
Neanderthal Museum, Mettmann, Germany

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



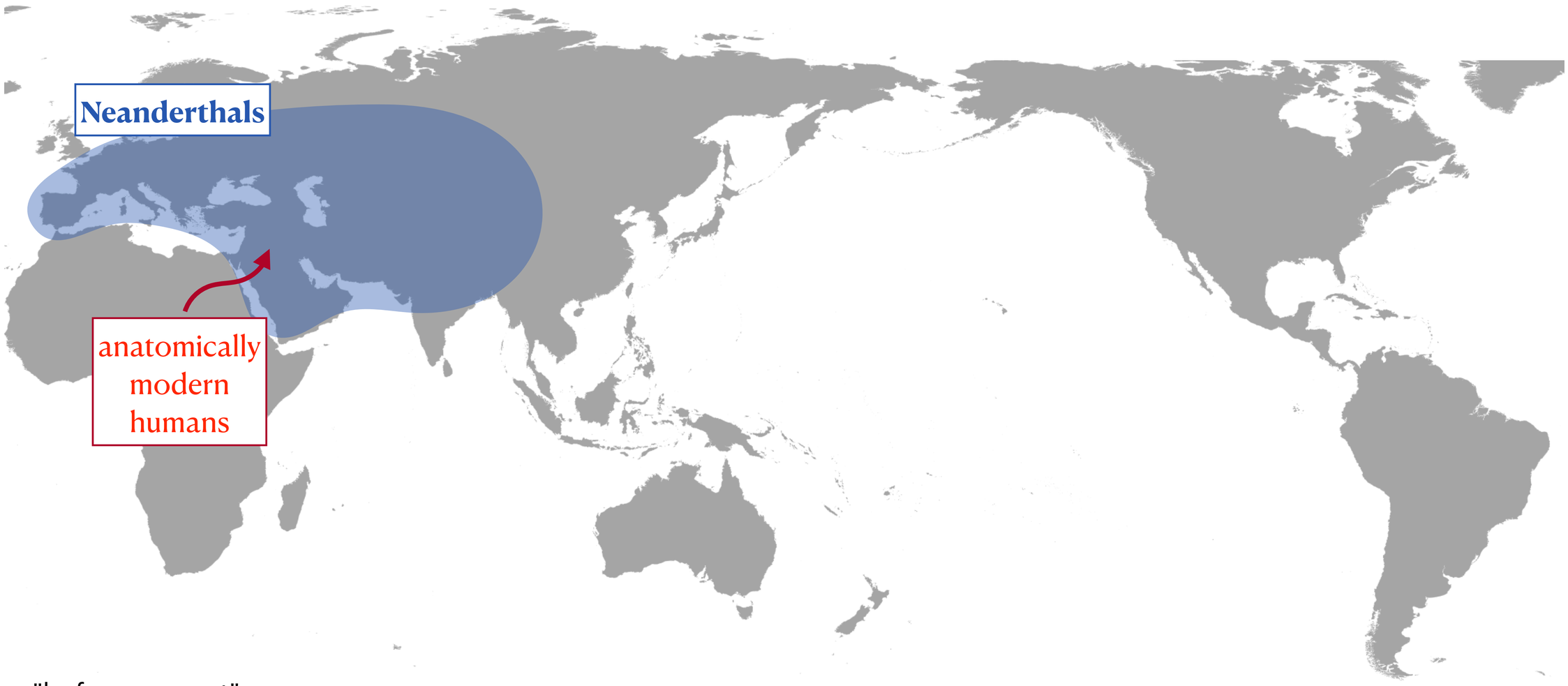
B.P. = "before present"

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



B.P. = "before present"

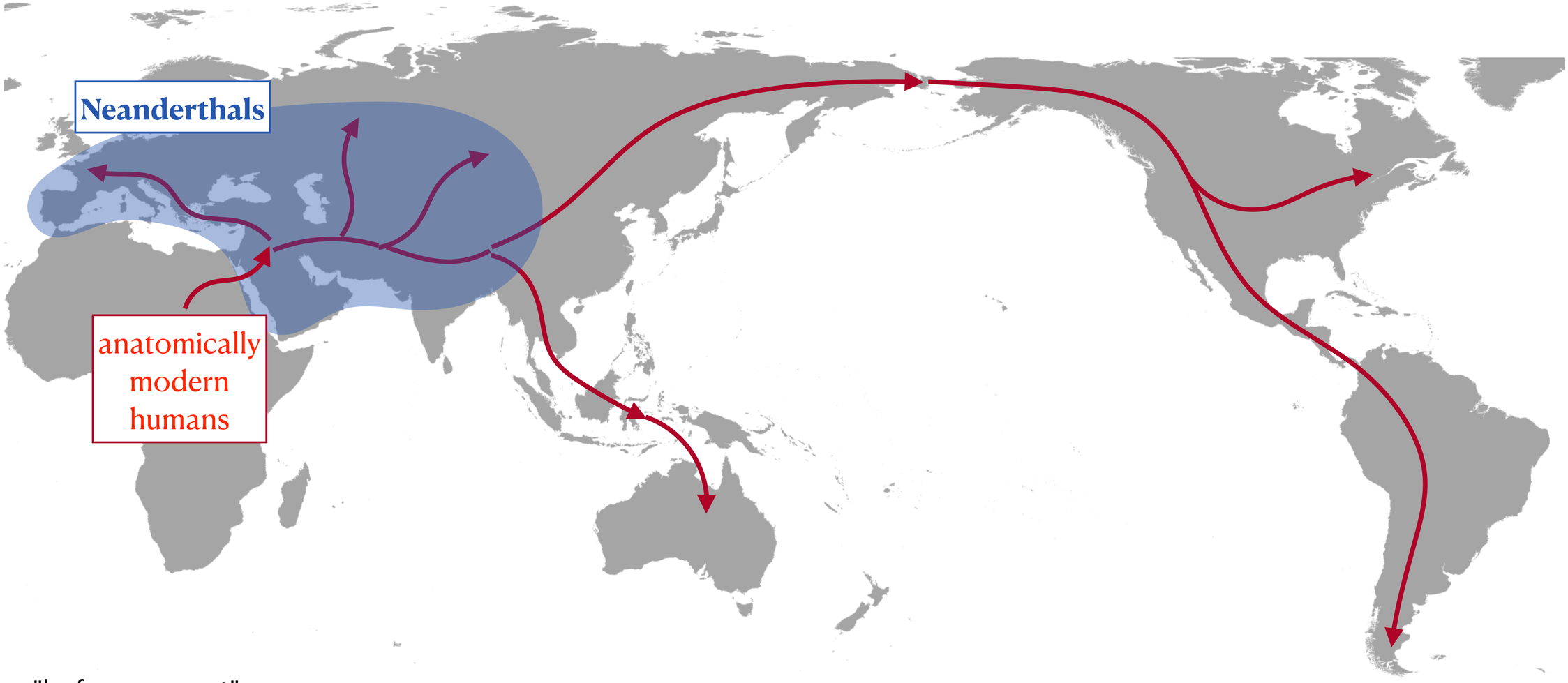
~70 kya B.P.: **modern humans migrated out of Africa**



B.P. = "before present"

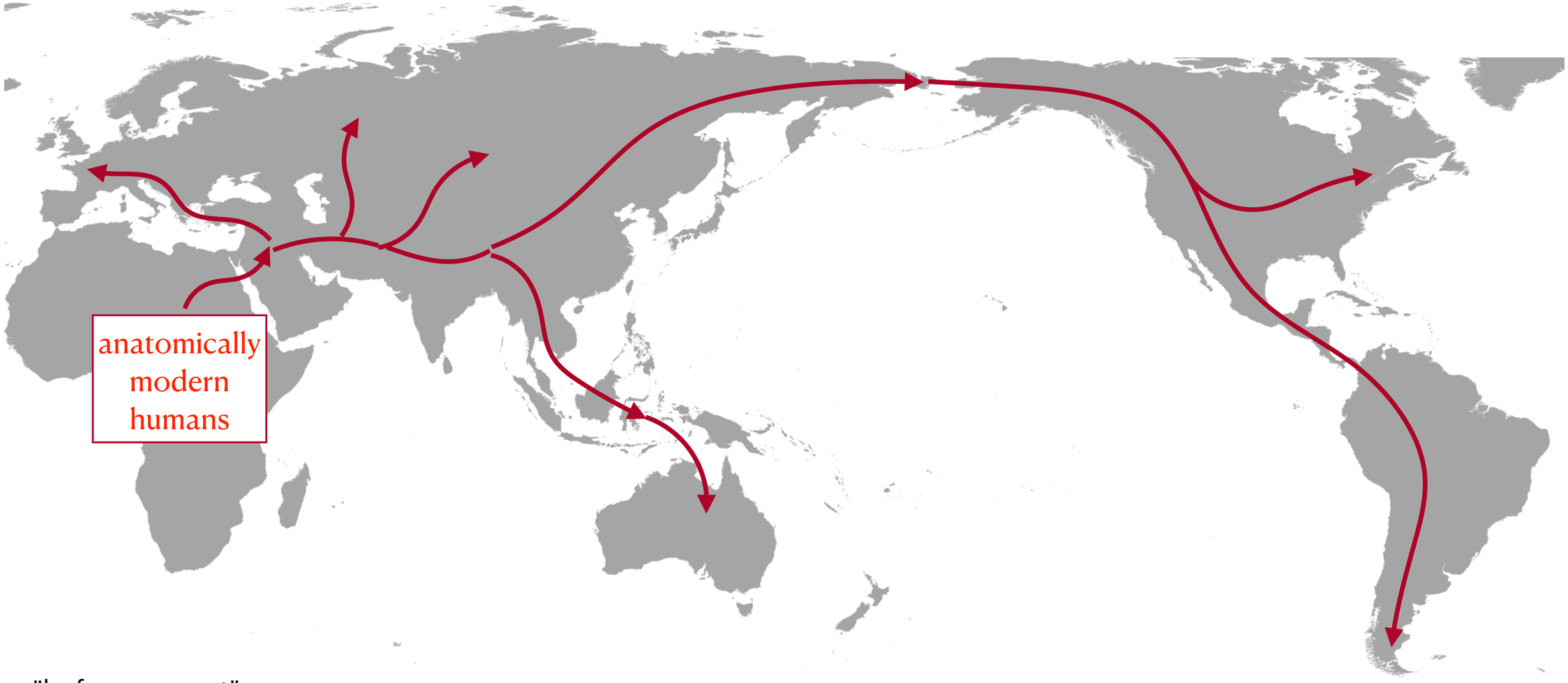
~70 kya B.P.: **modern humans migrated out of Africa**

...and collonized the entire world



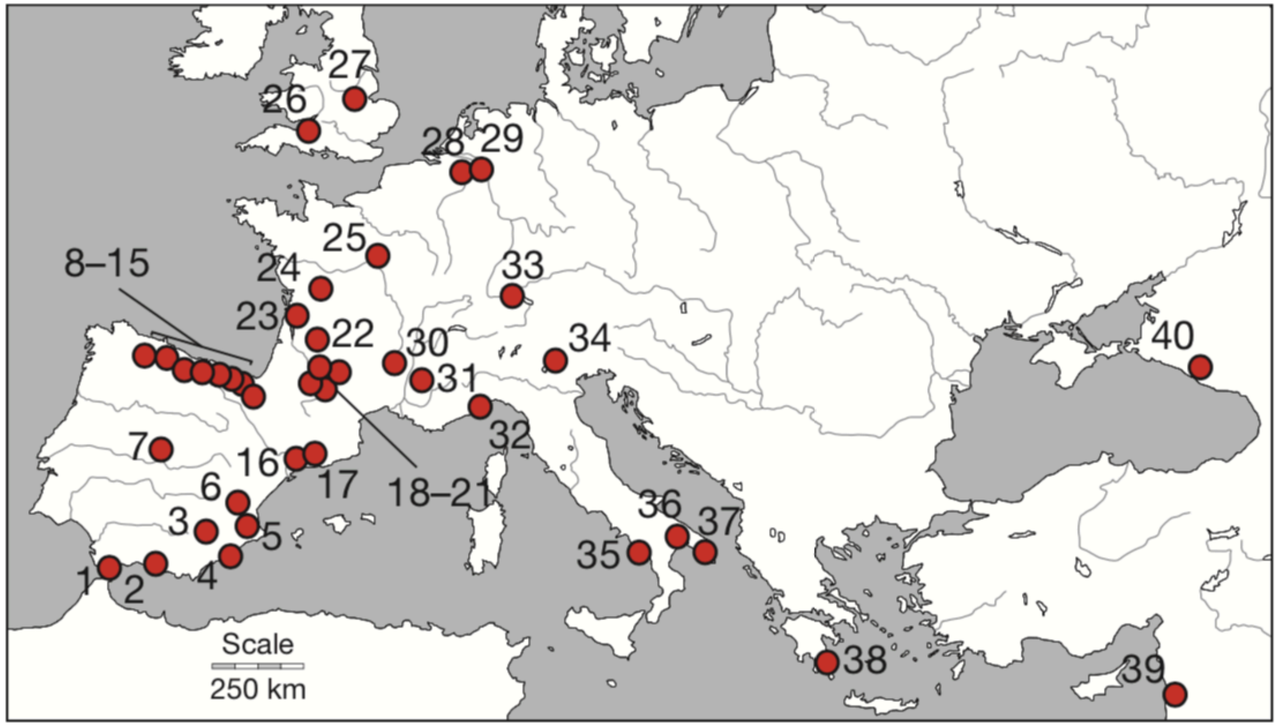
B.P. = "before present"

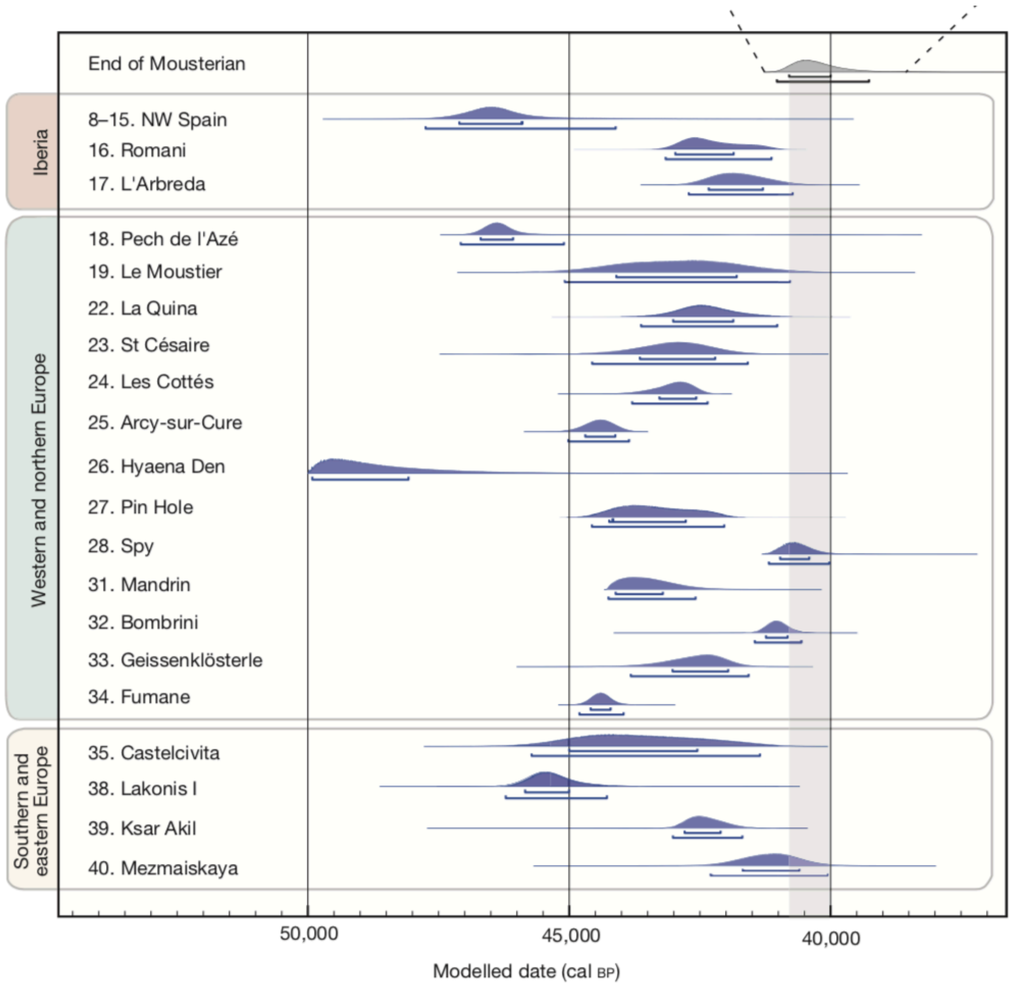
~40 kya: **Neanderthals vanished**



B.P. = "before present"

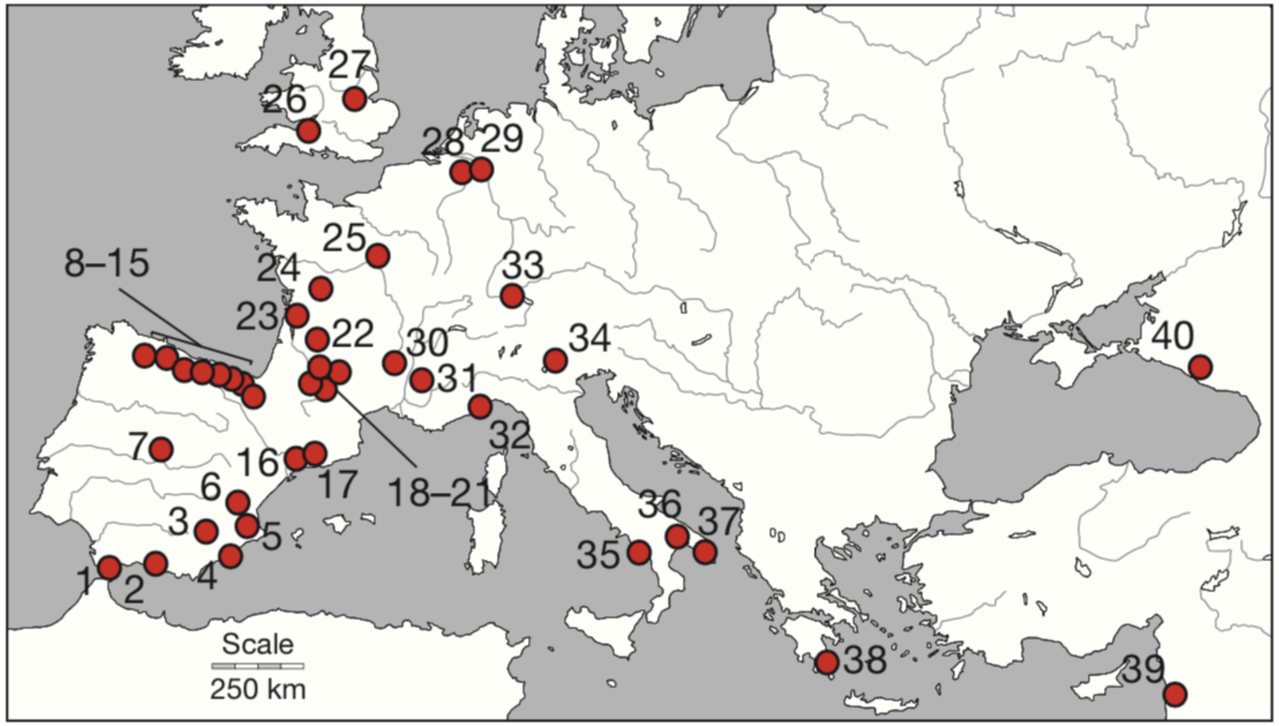
~40 kya: Neanderthals vanished

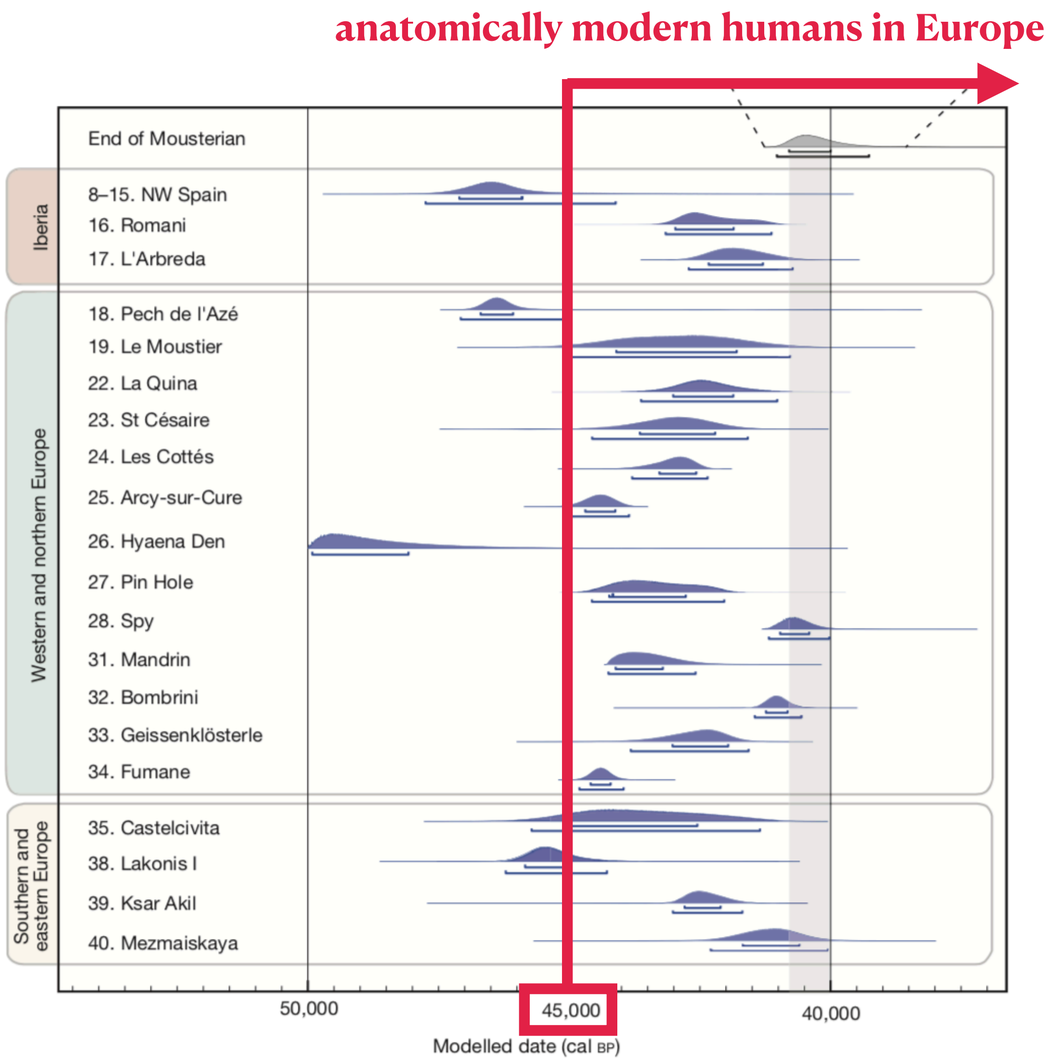




Higham *et al.* (Nature, 2014)

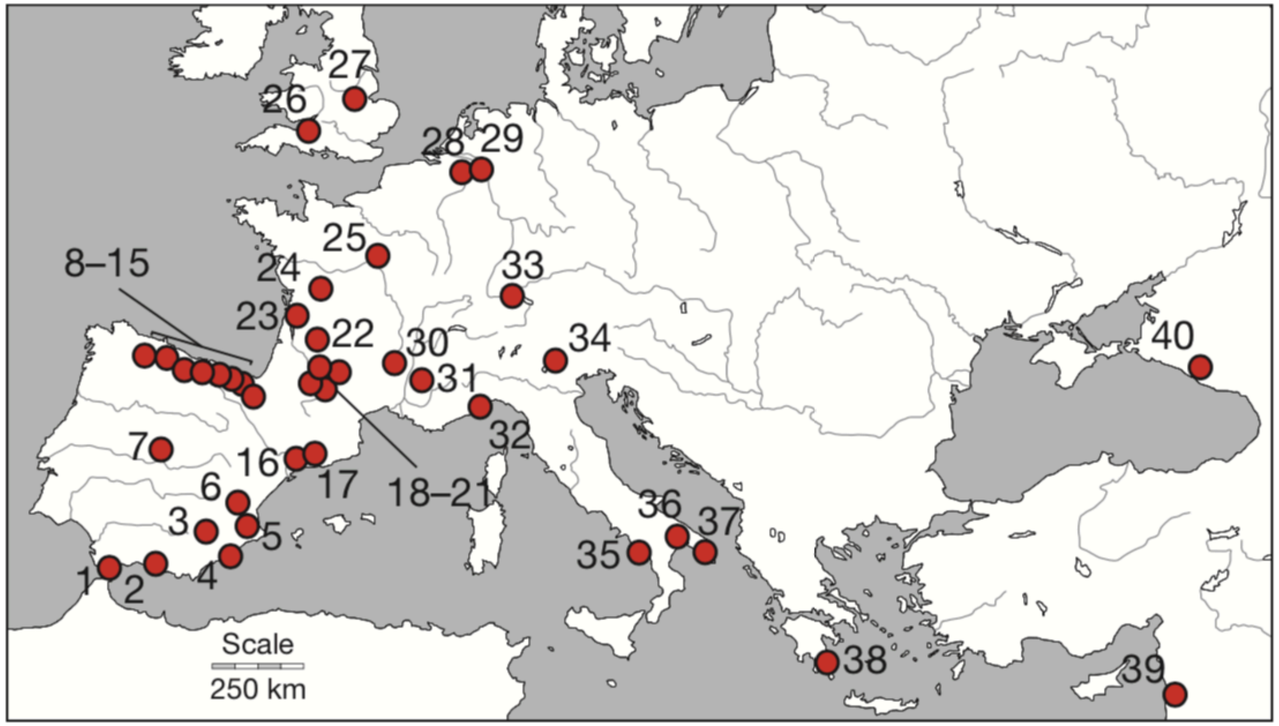
~40 kya: Neanderthals vanished

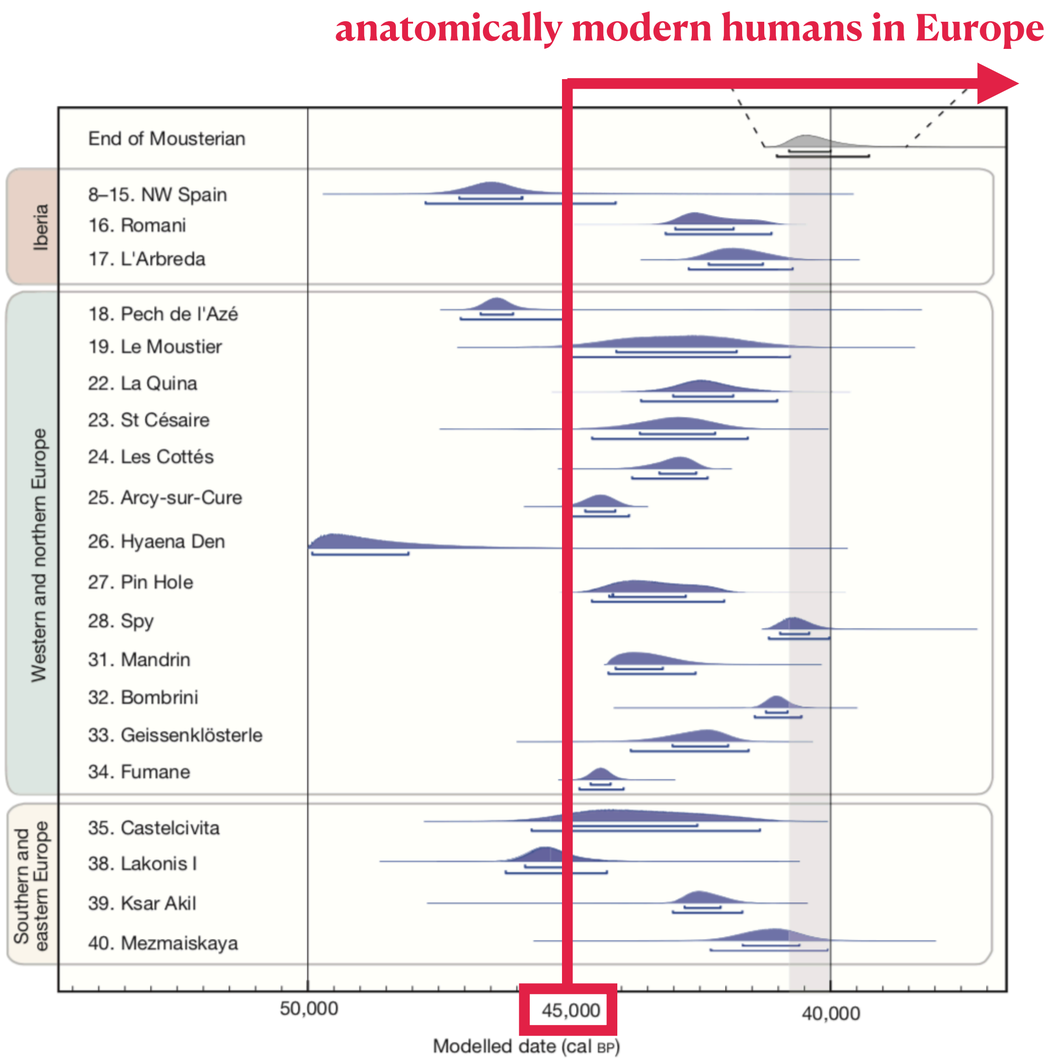




Higham *et al.* (Nature, 2014)

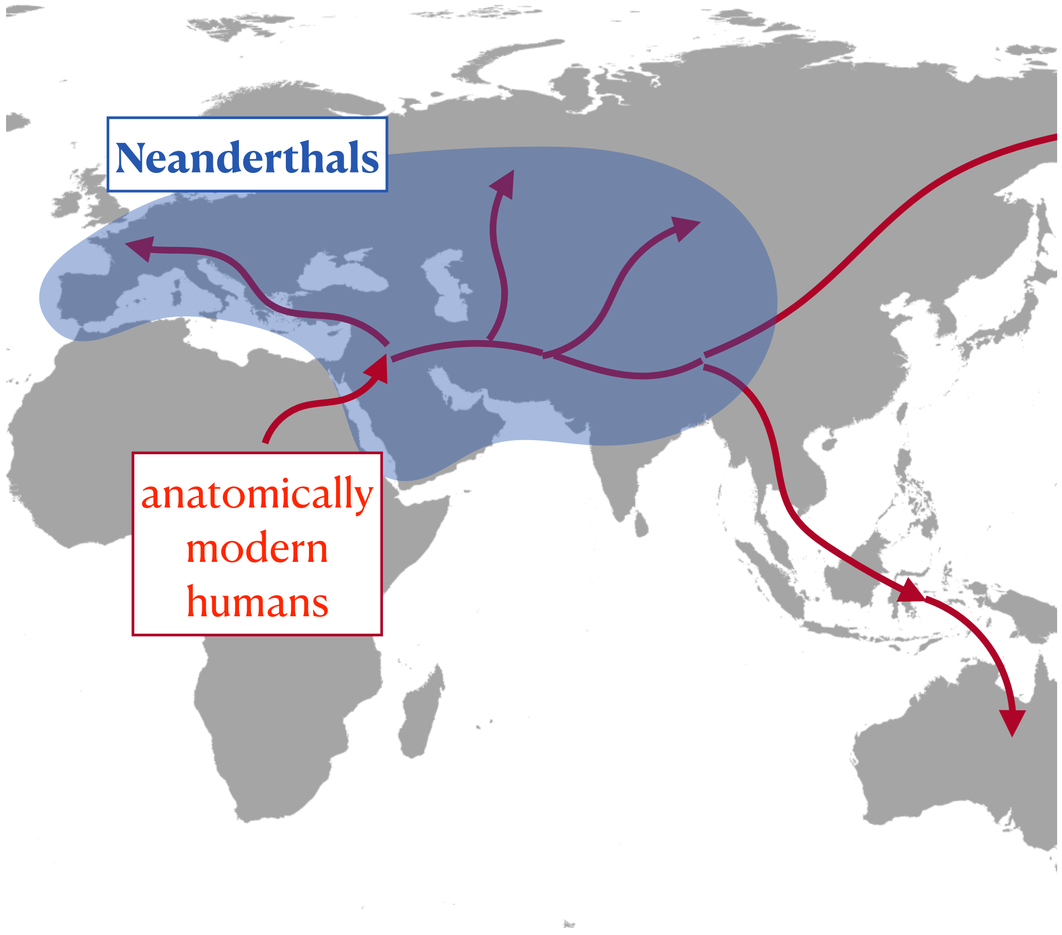
What happened in the few thousand years of overlap?



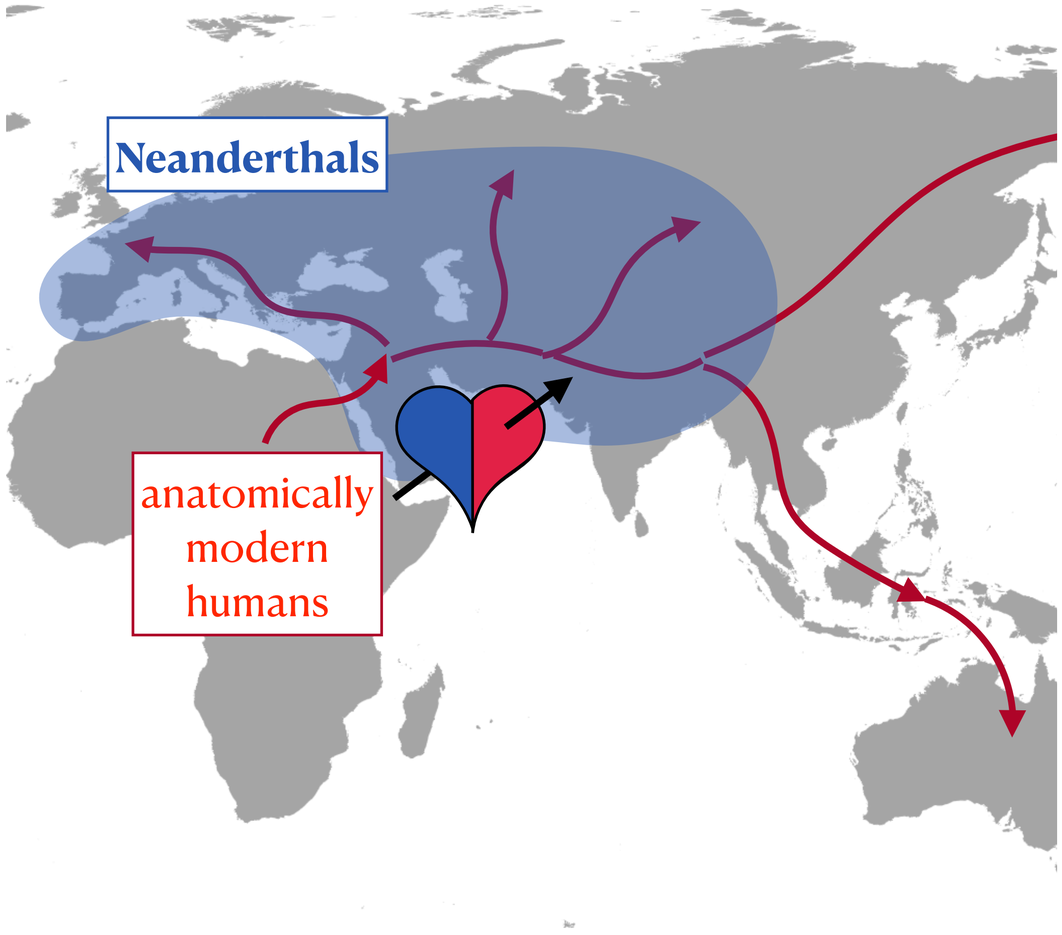


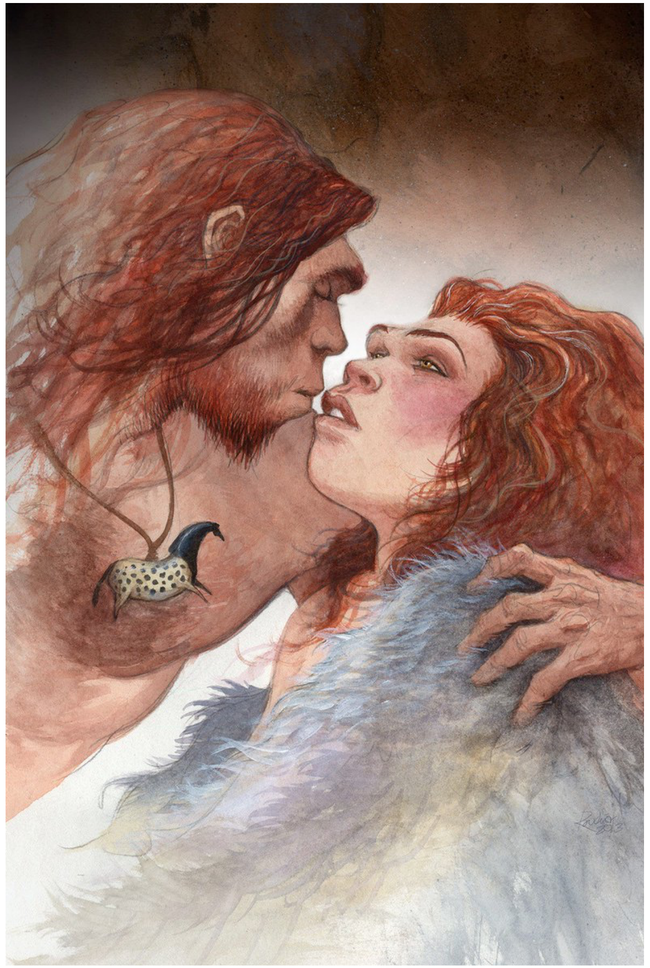
Higham *et al.* (Nature, 2014)

What happened in the few thousand years of overlap?



Introgression / gene flow / admixture?





https://twitter.com/jjhublin/status/739866080764628993

Morphological evidence?

Trinkaus *et al*. PNAS, 2003

~40 thousand years old

remains of a modern human

Peștera cu Oase, Romania

Neanderthal DNA? (ancient DNA, aDNA)

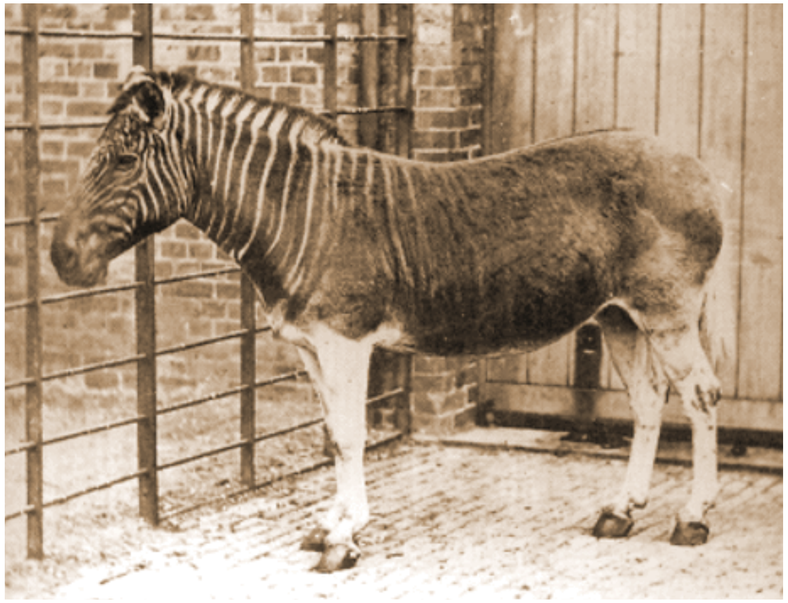
1984

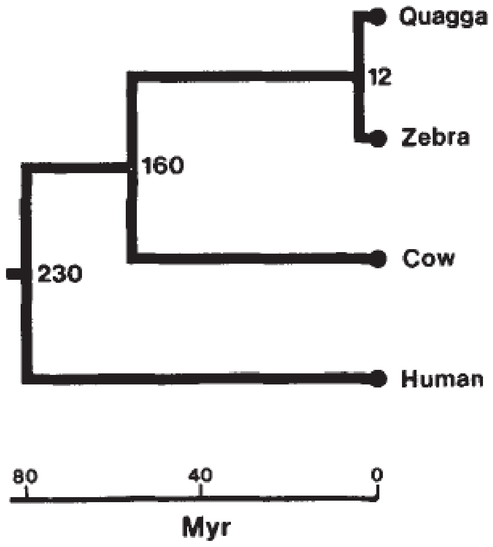
DNA sequences from the quagga, an extinct member of the horse family

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

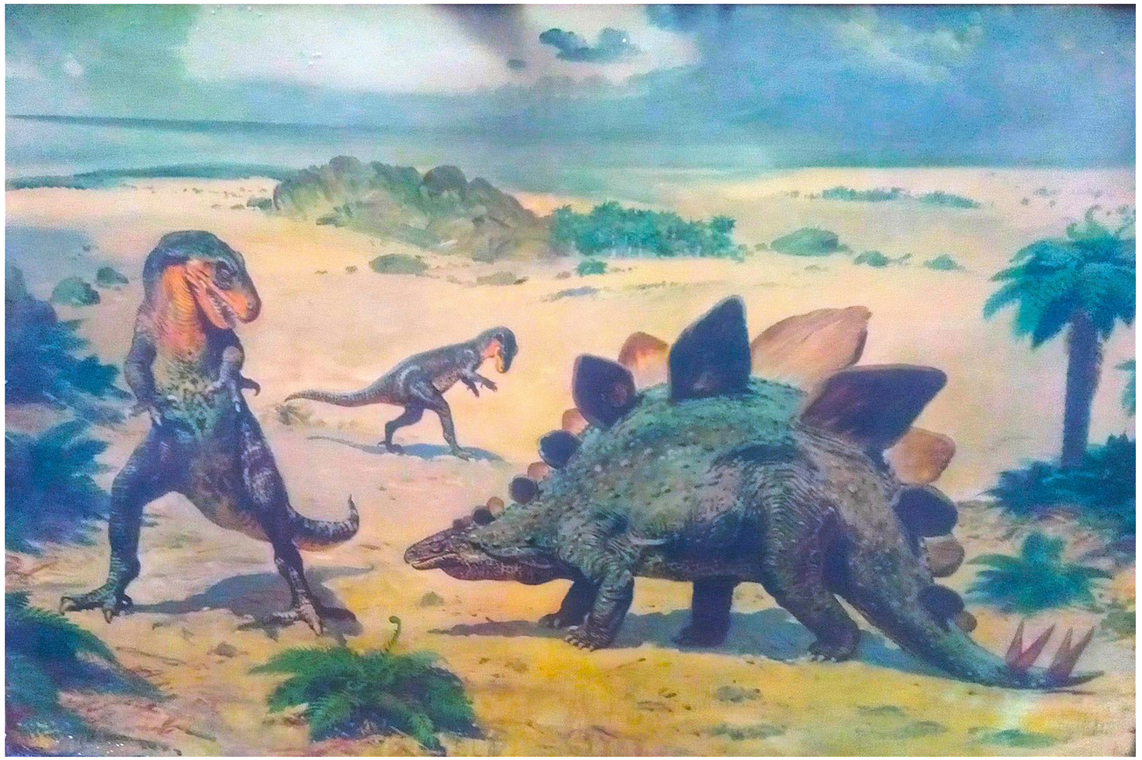
1994

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

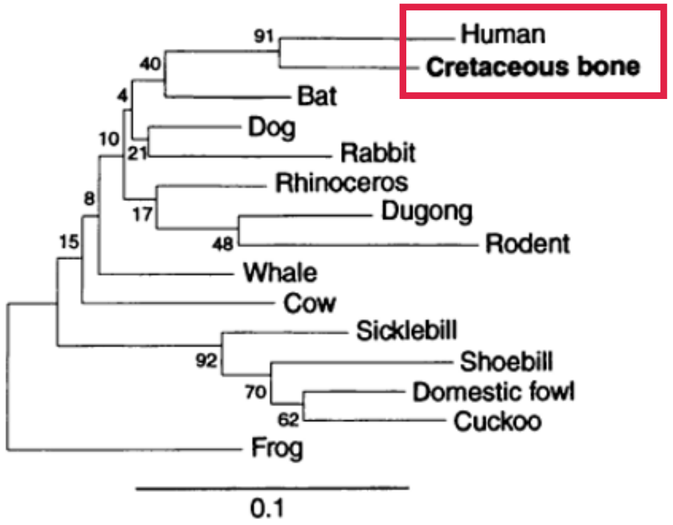
1994

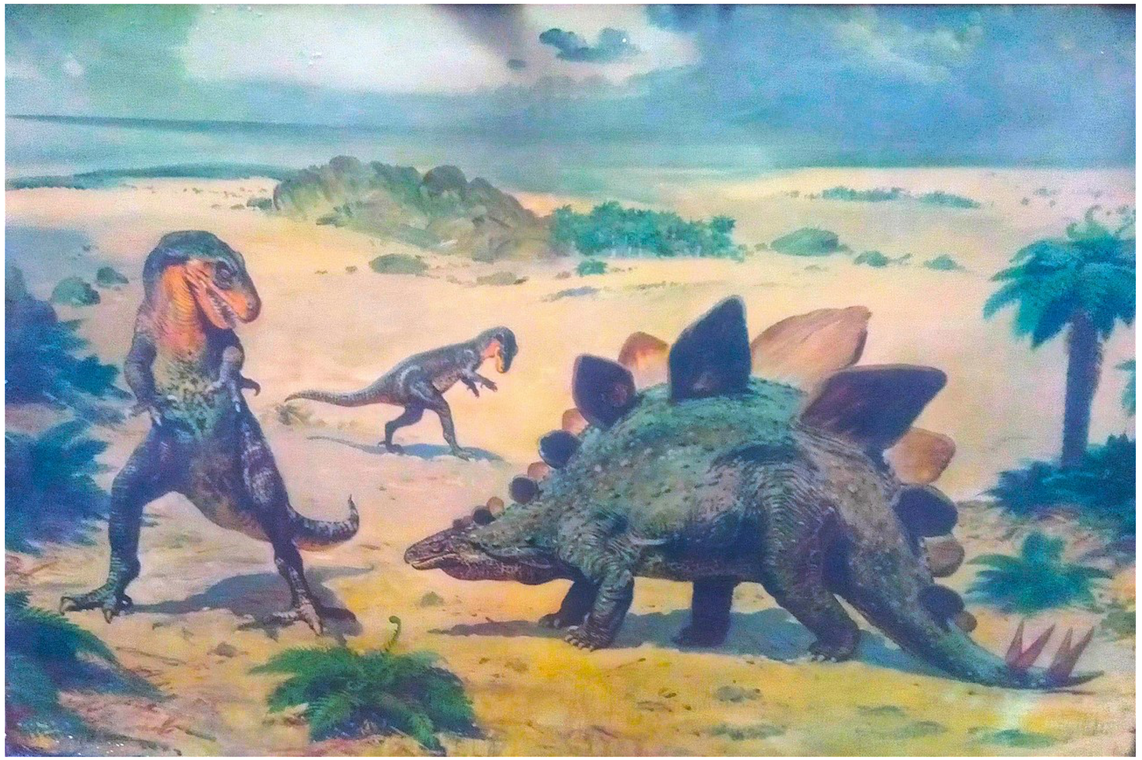
DNA Sequence from Cretaceous Period Bone Fragments

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

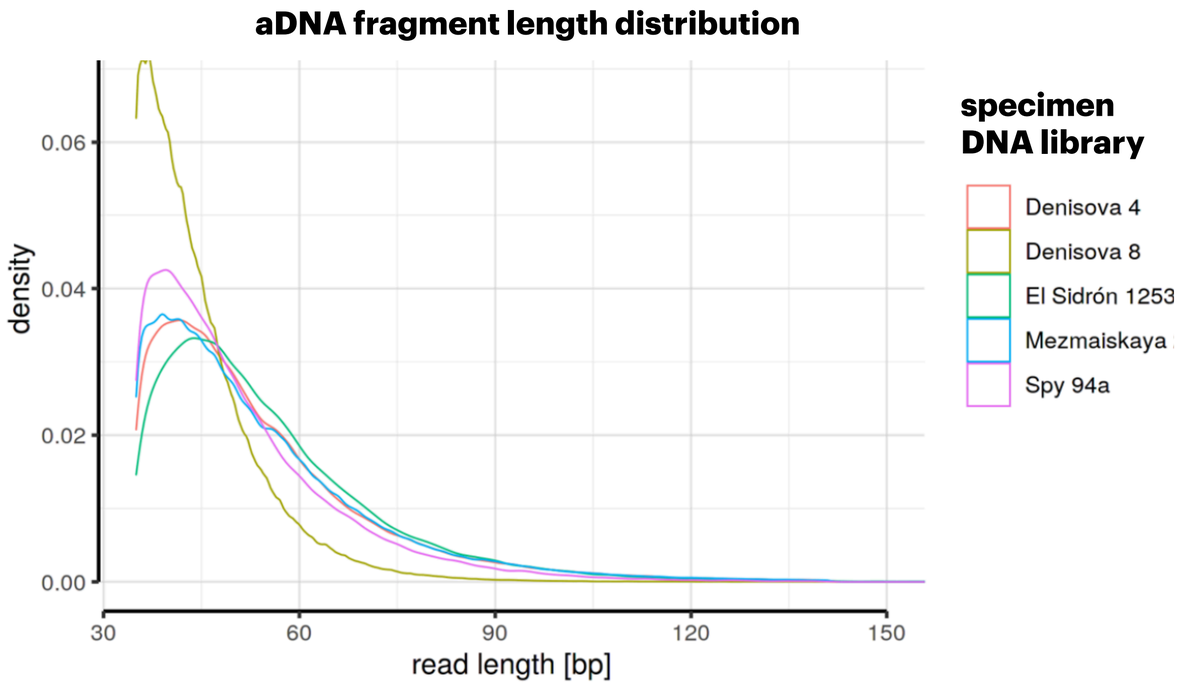




Molecular characteristics of aDNA

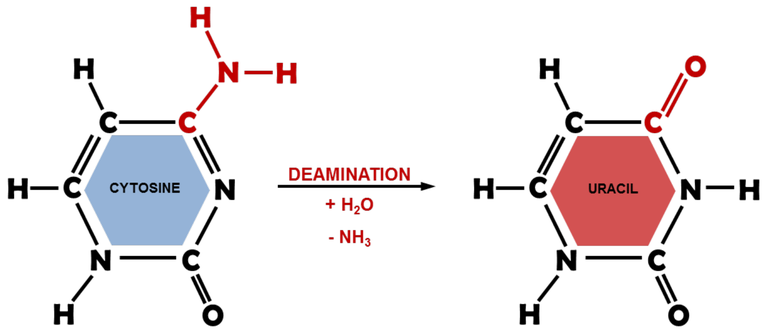
Molecular characteristics of aDNA

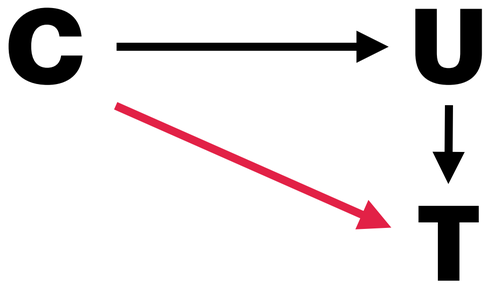
- highly fragmented

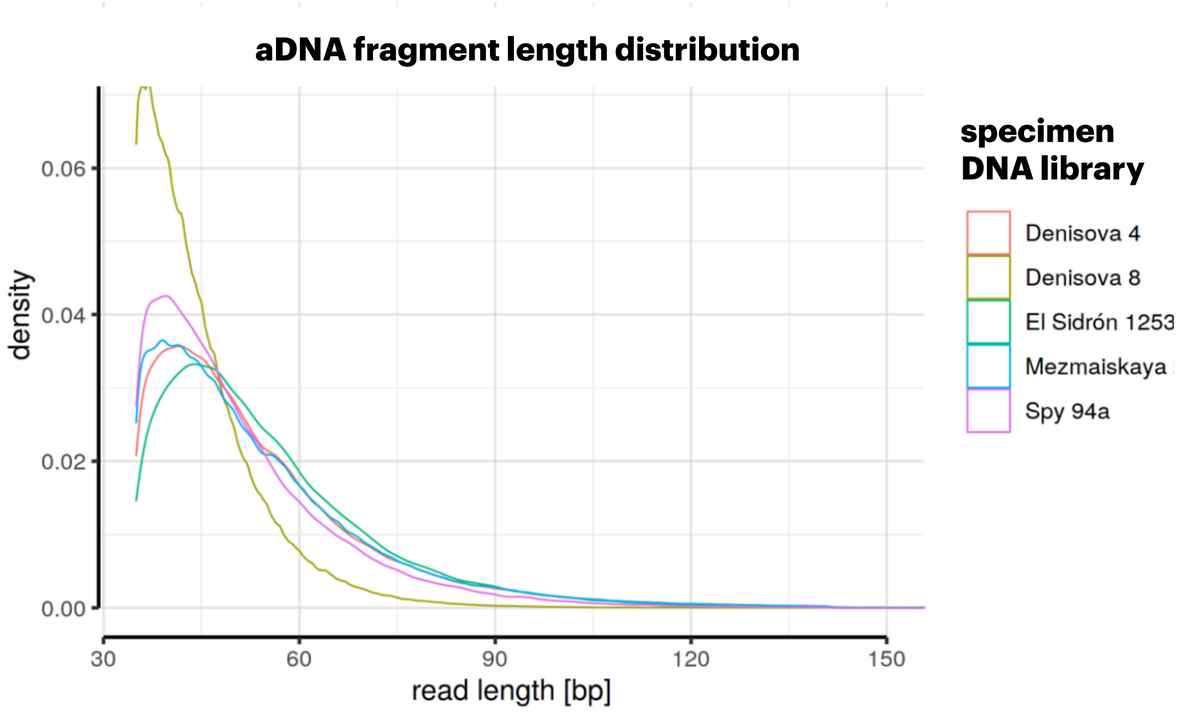


Molecular characteristics of aDNA

- highly fragmented
- post-mortem chemical modifications

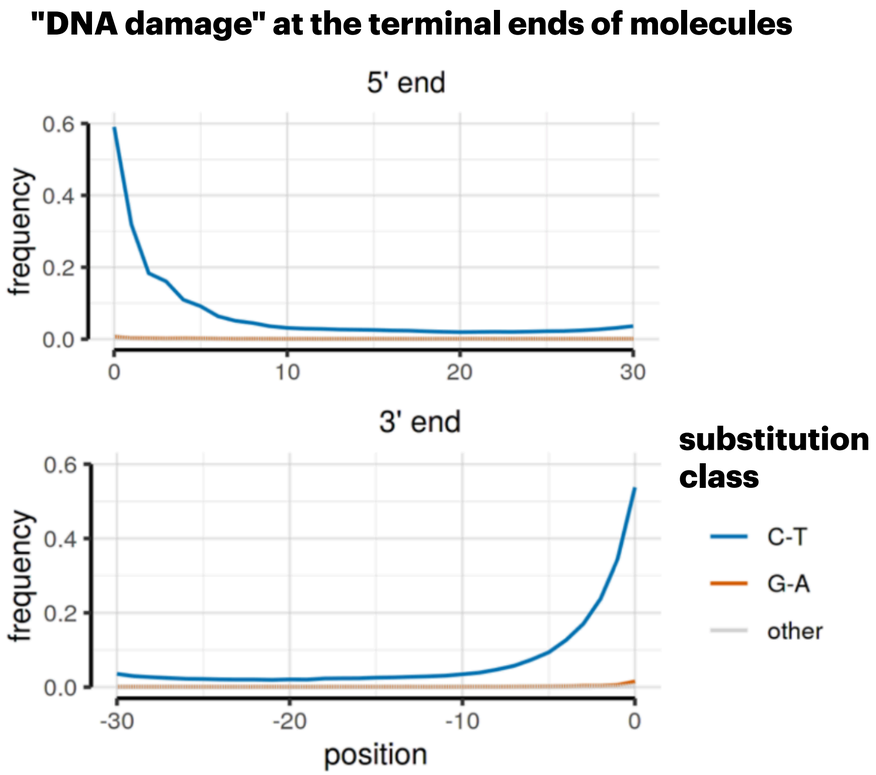


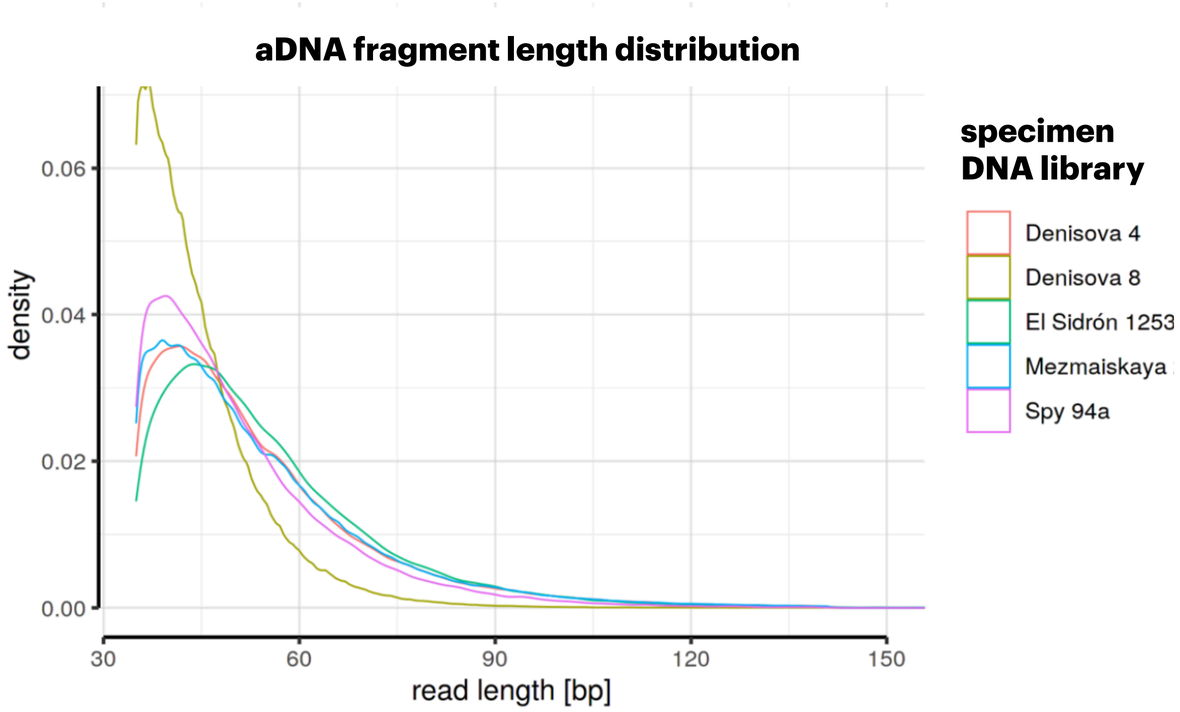




Molecular characteristics of aDNA

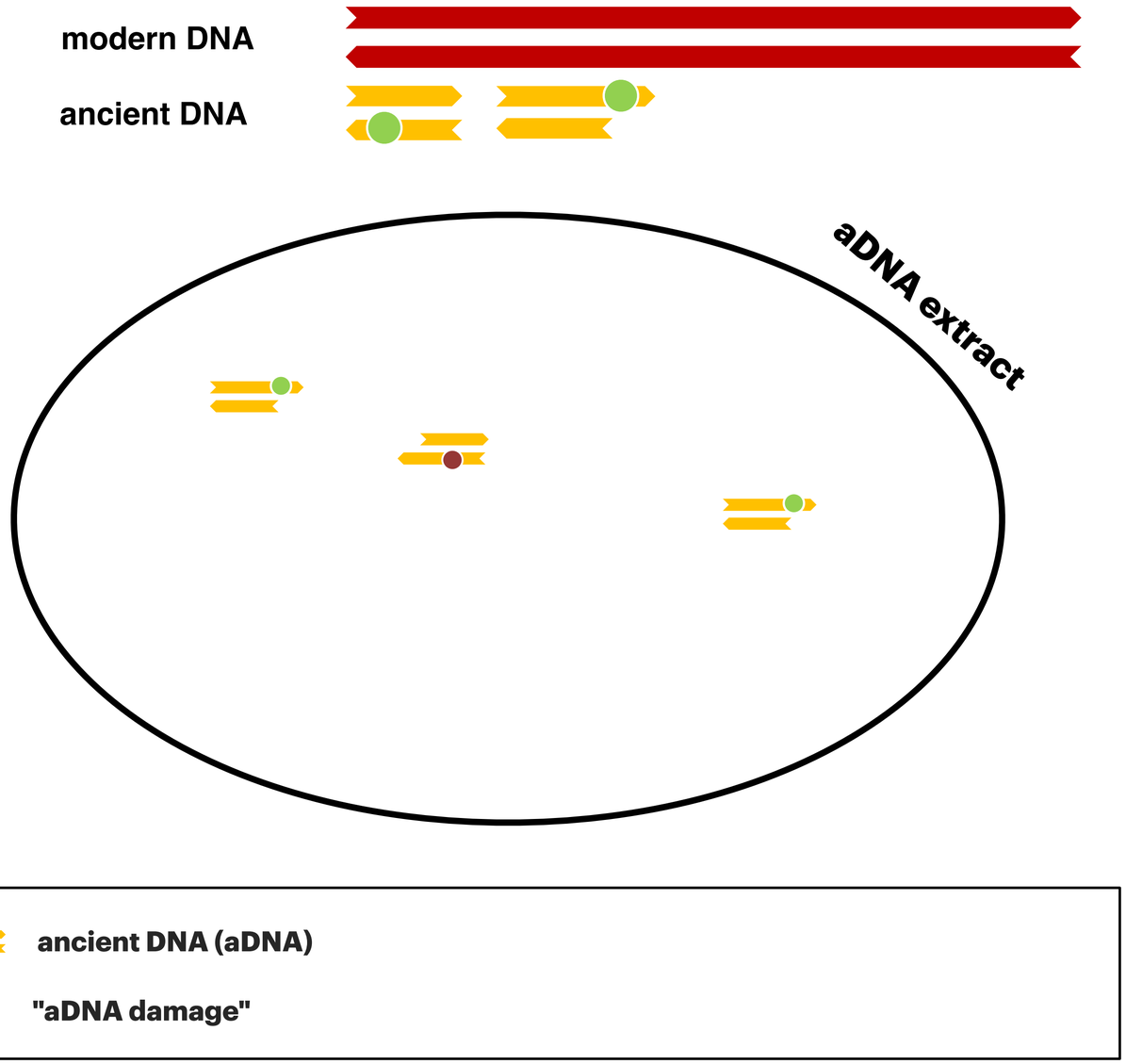
- highly fragmented
- post-mortem chemical modifications





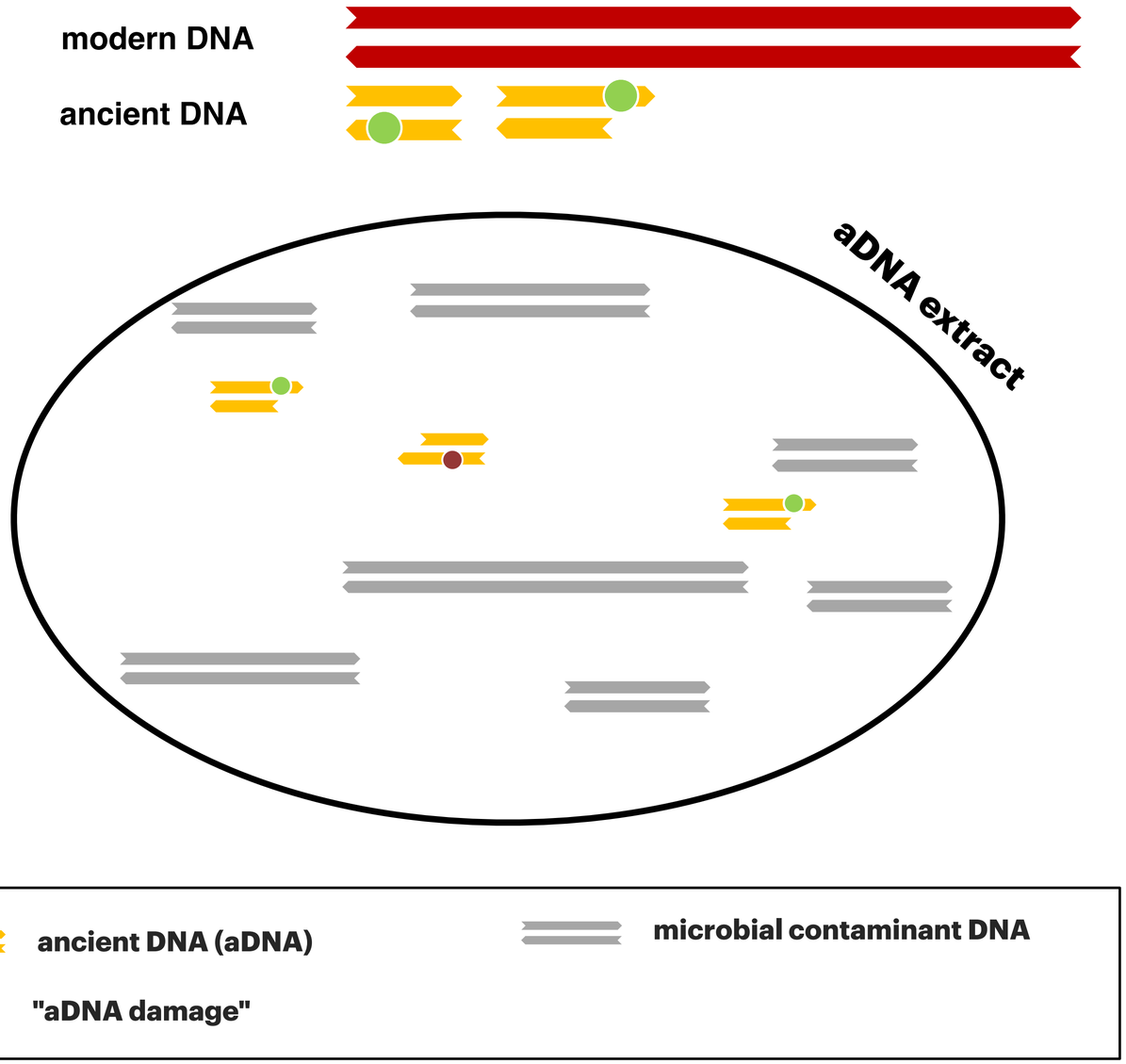
Molecular characteristics of aDNA

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



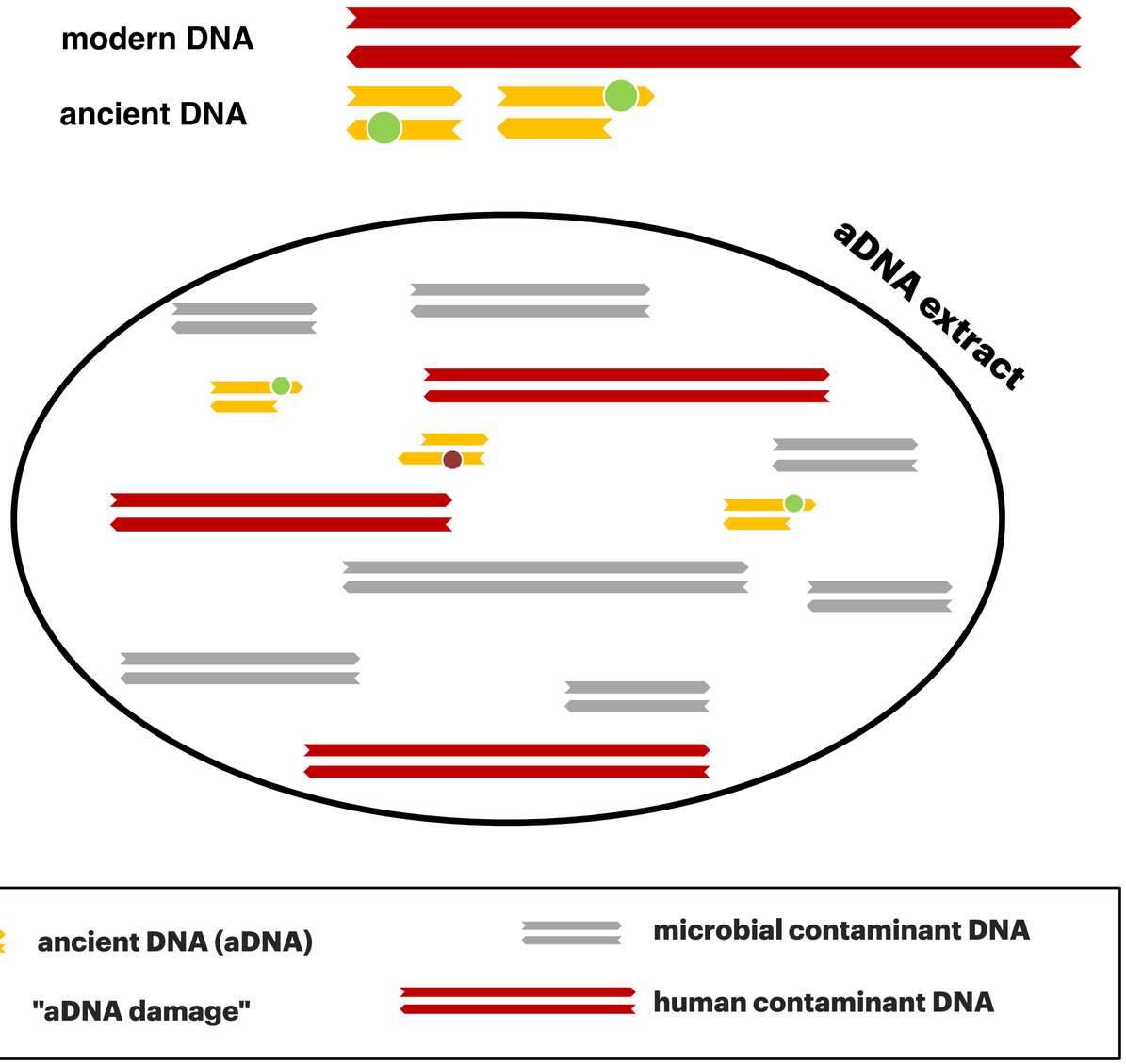
Molecular characteristics of aDNA

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



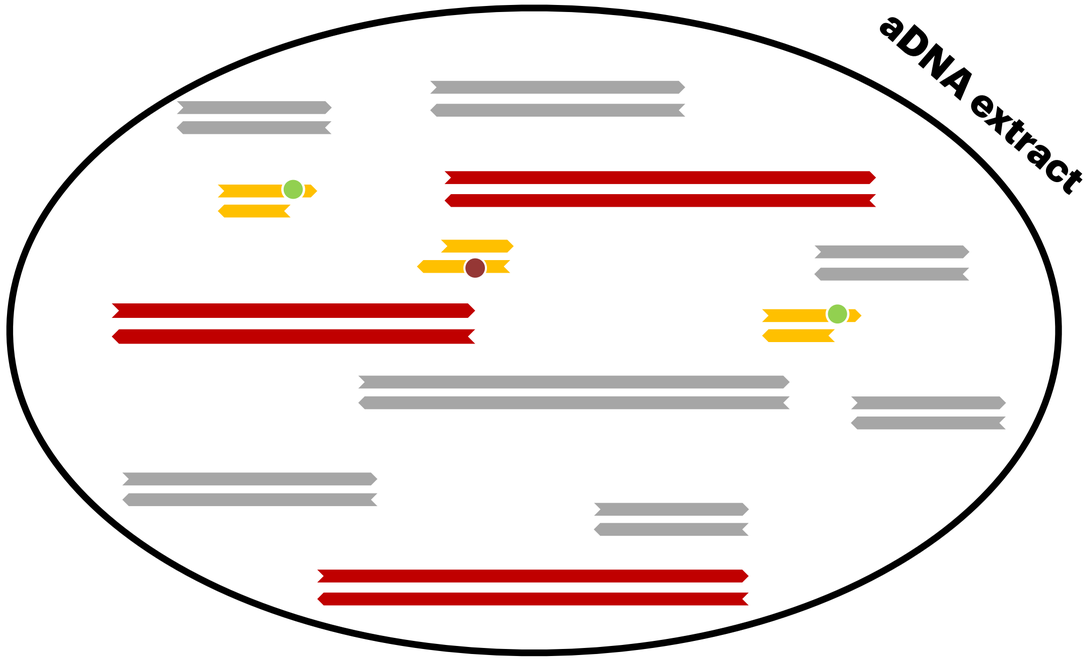
Molecular characteristics of aDNA

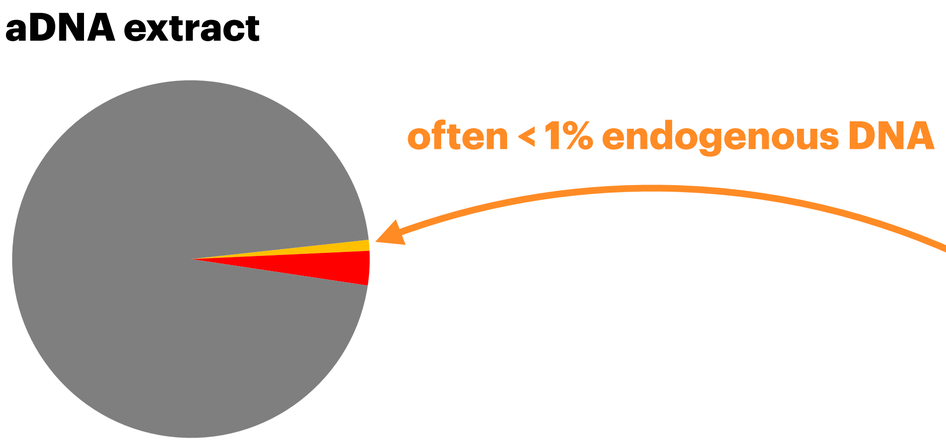
- highly fragmented
- post-mortem chemical modifications
- contamination:
  - microbialDNA
  - human DNA (excavation, museum, lab)

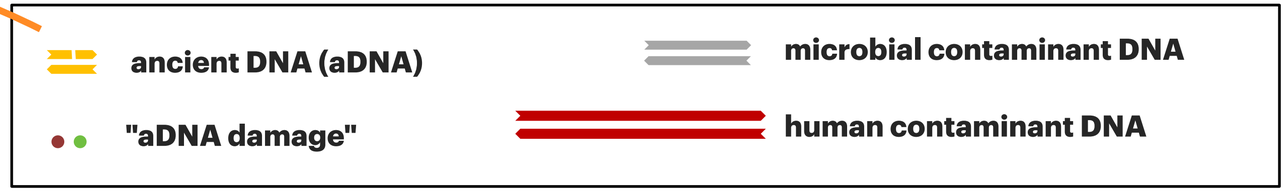


Molecular characteristics of aDNA

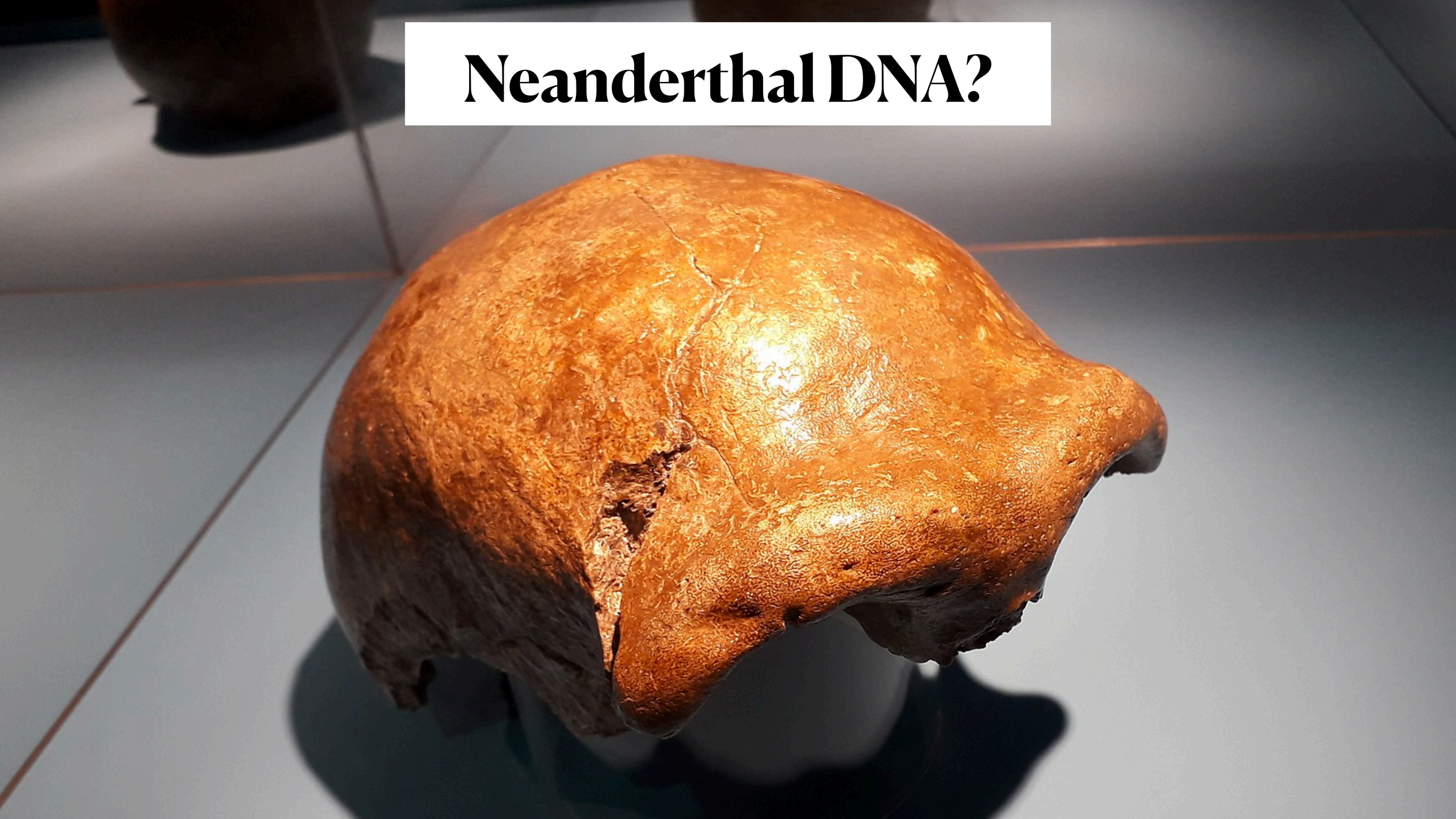
- highly fragmented
- post-mortem chemical modifications
- contamination:
  - microbialDNA
  - human DNA (excavation, museum, lab)







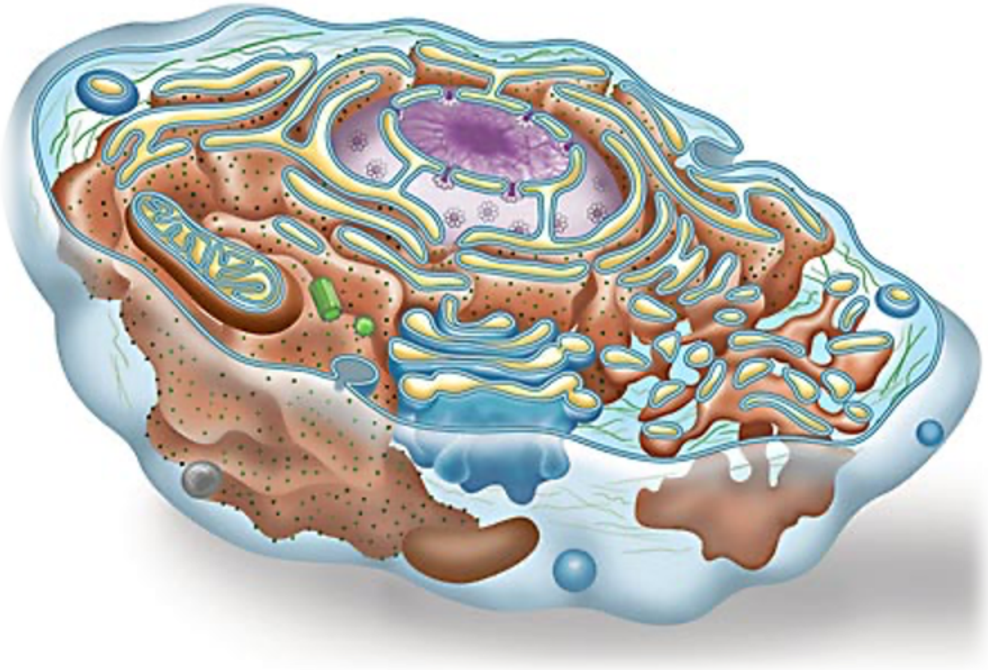
Neanderthal DNA?



Svante Pääbo

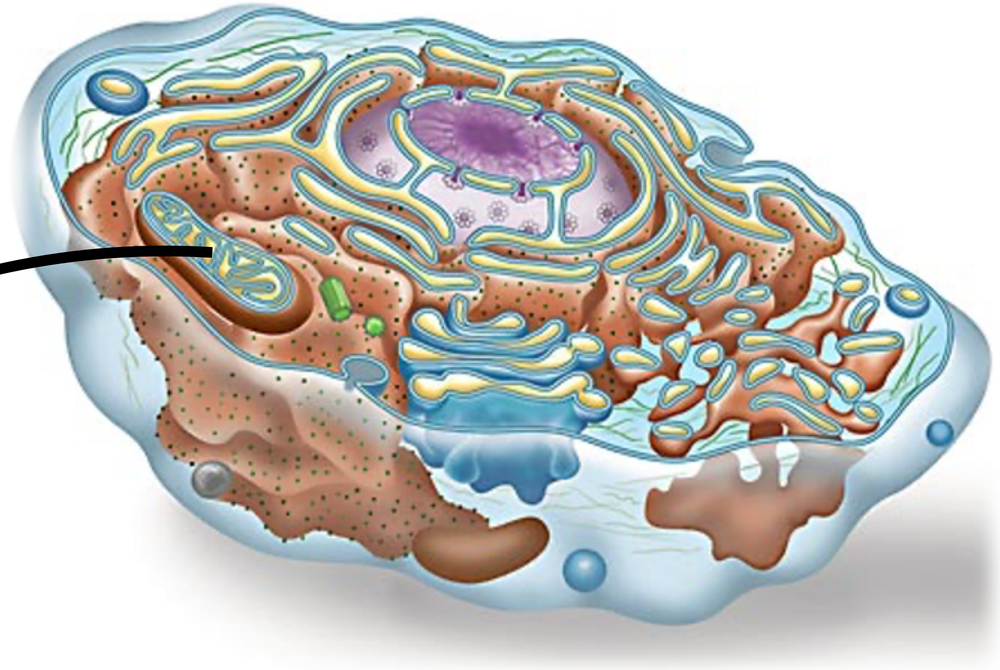
@maxplanckpress

Two sources of DNA



www.science-art.com

Two sources of DNA

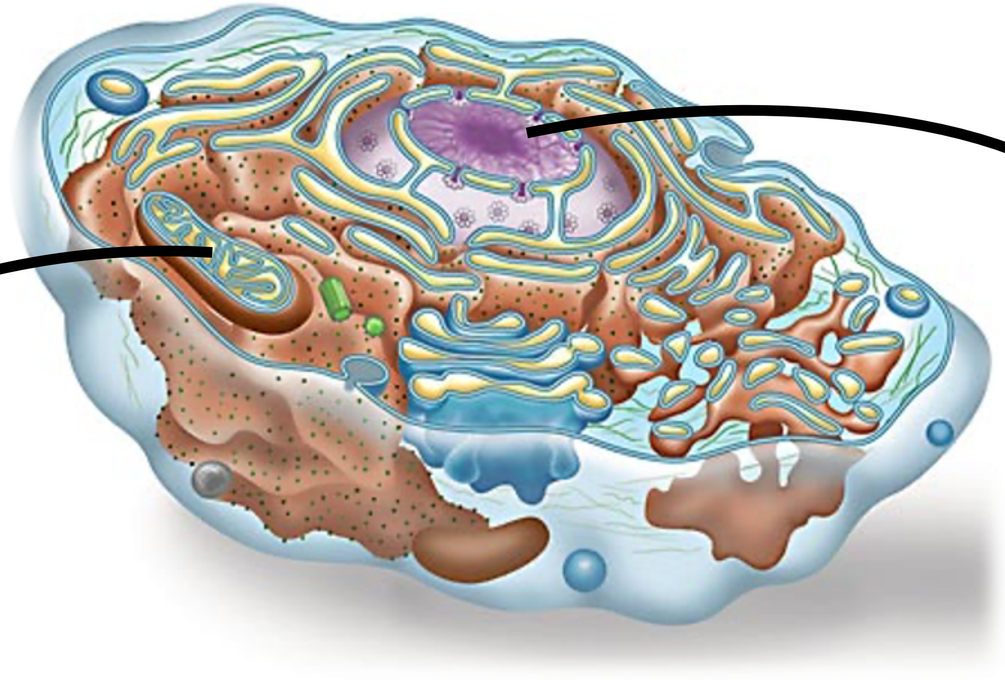


mitochondrial DNA (mtDNA)
16569 bp

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

www.science-art.com

Two sources of DNA



mitochondrial DNA (mtDNA)
16569 bp

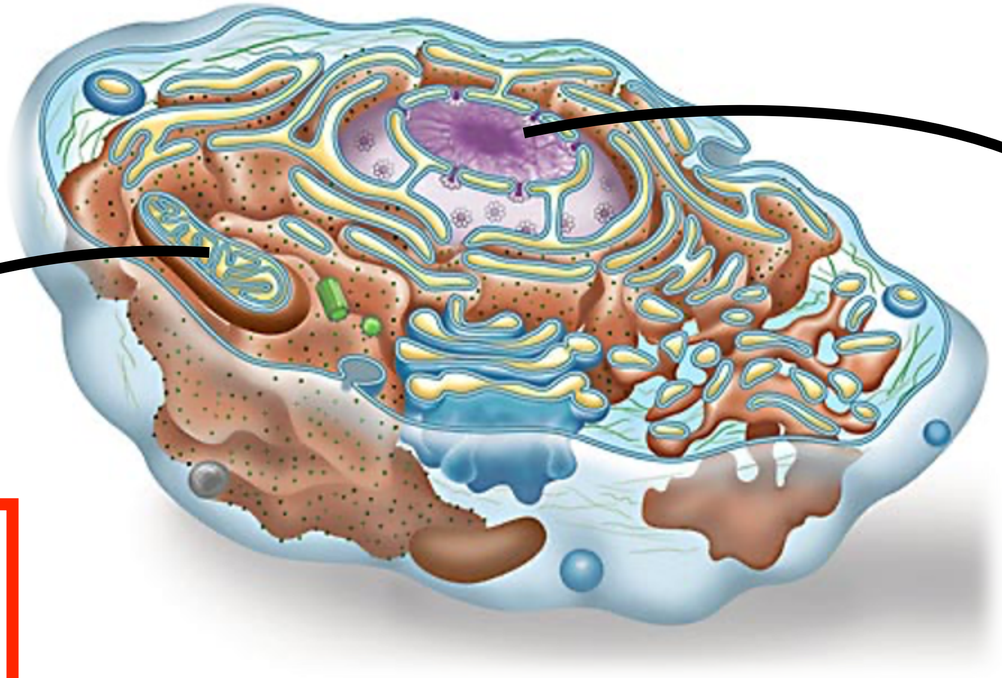
- just one locus
- transmitted along maternal line
- fast mutation rate
- **thousands of copies in a cell**

nuclear DNA
~3 Gb

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

www.science-art.com

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**

www.science-art.com

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*

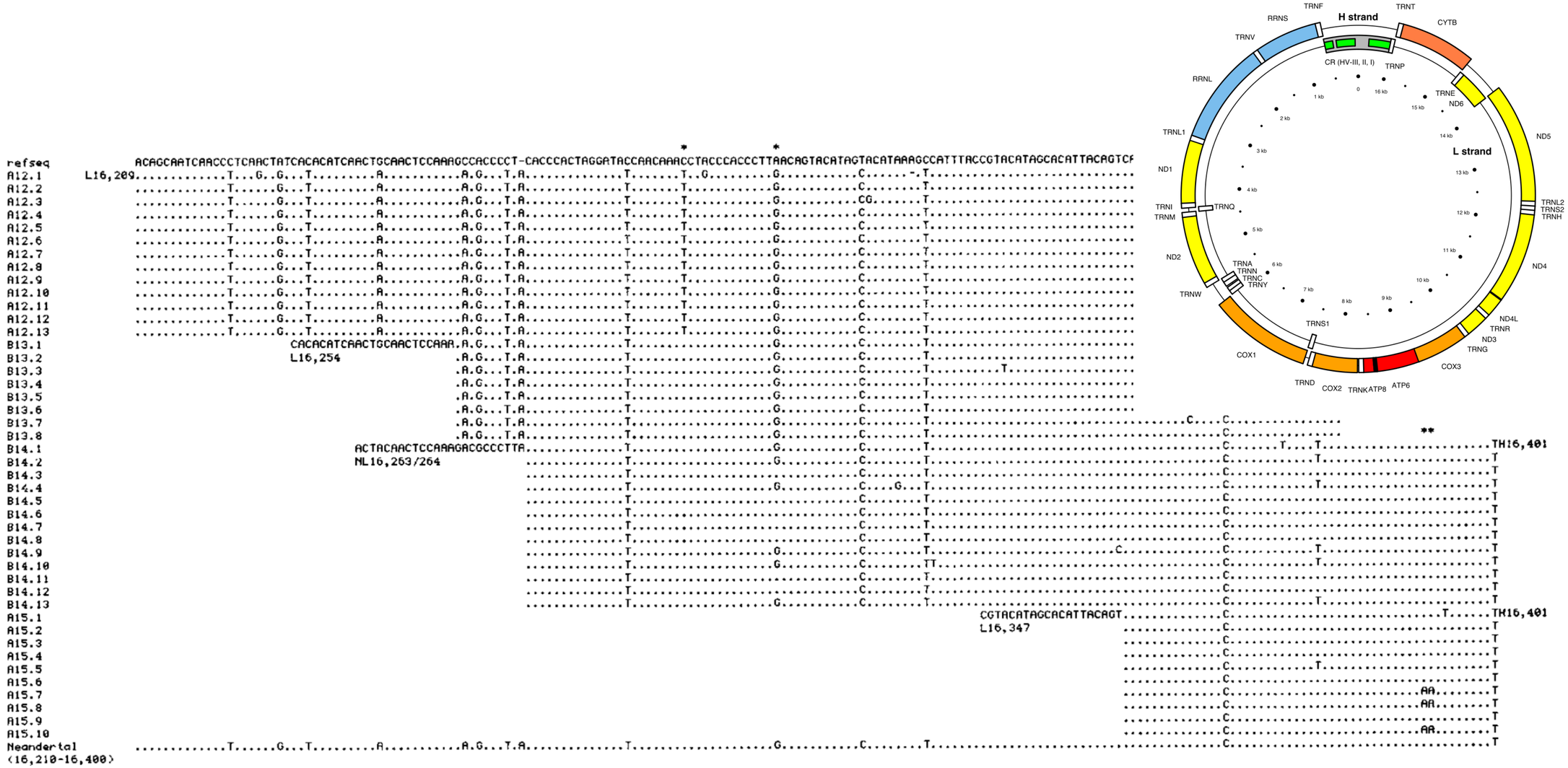
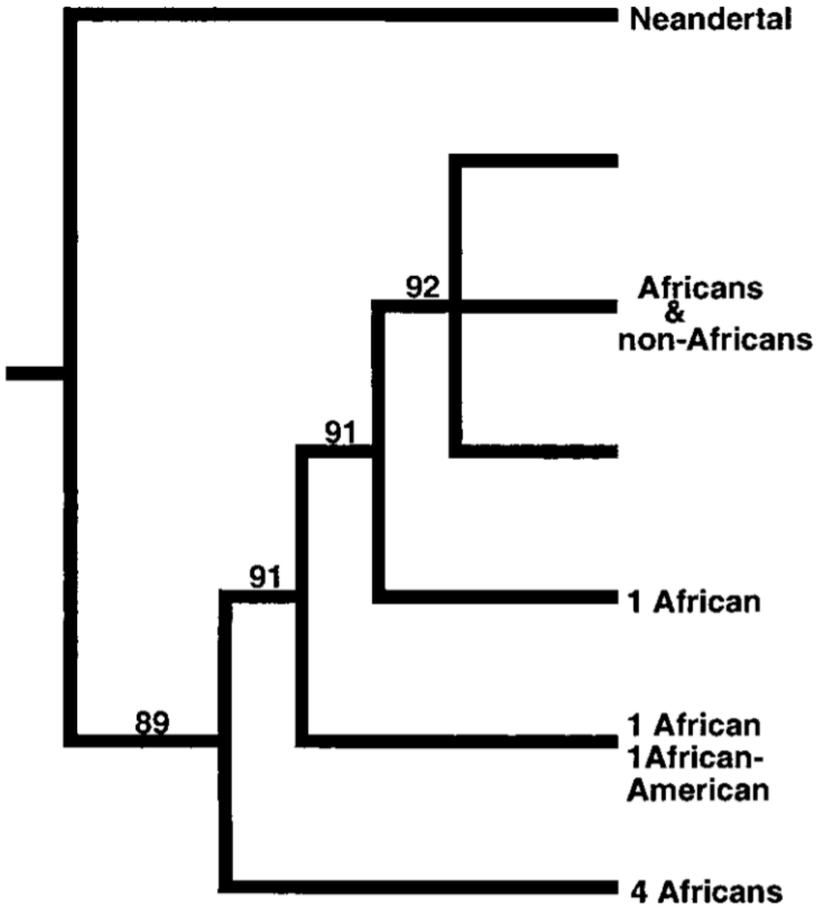
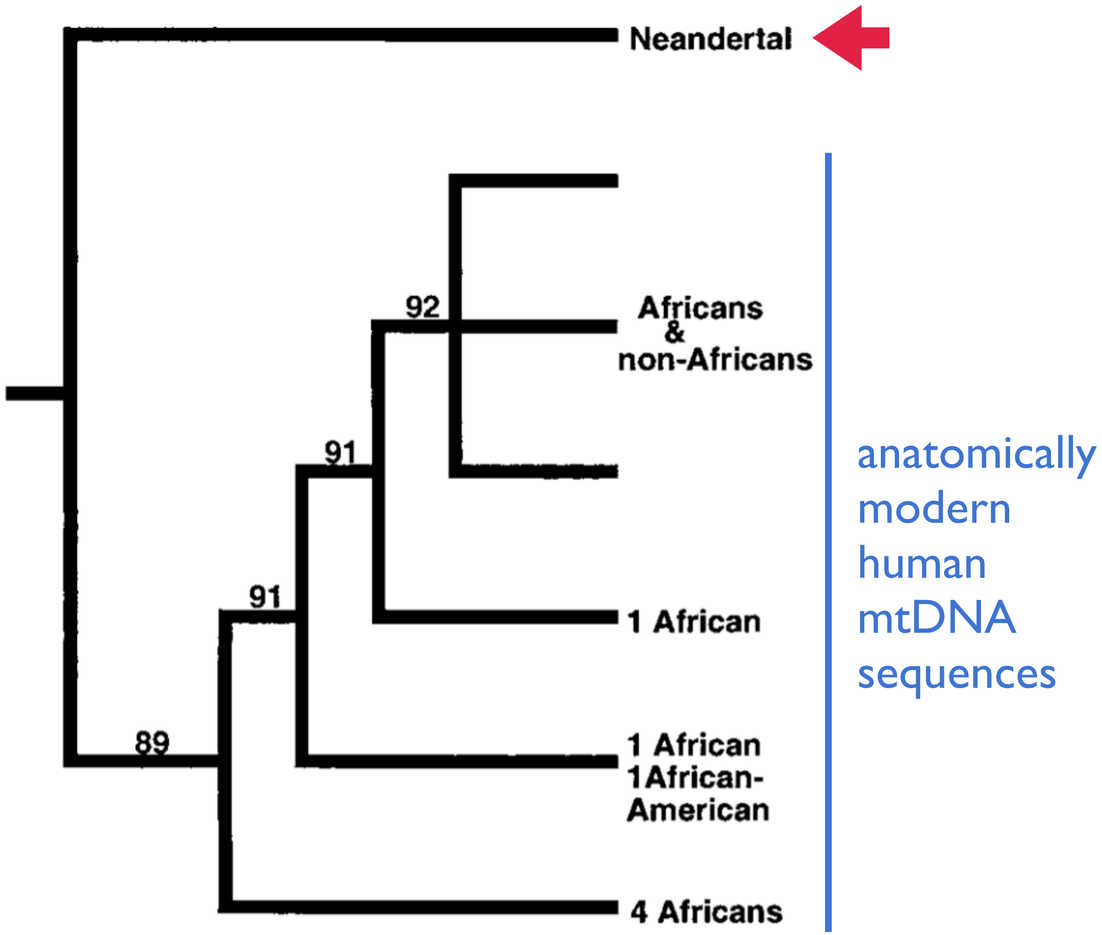


Figure 5. The DNA Sequences of Clones Used to Infer the Sequence of the Hypervariable Region I of the Neandertal Individual

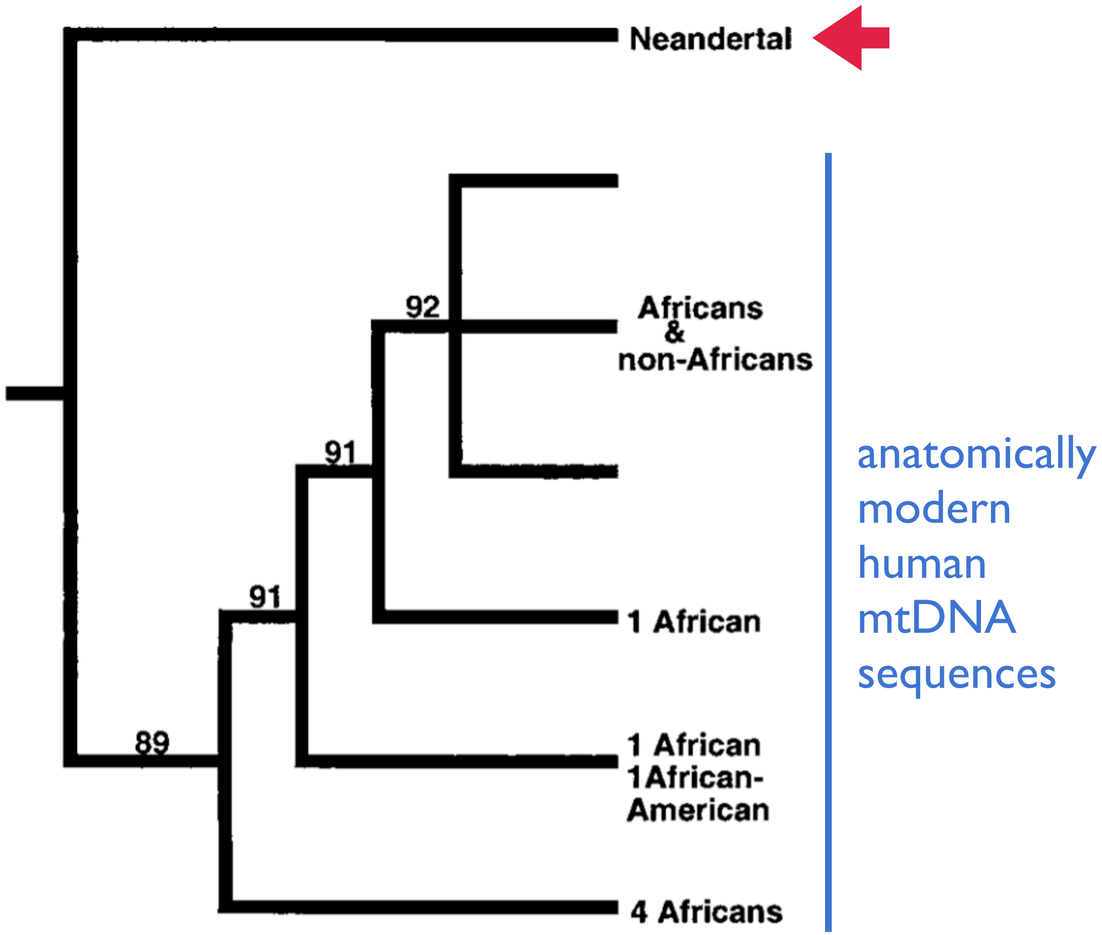


anatomically
modern
human
mtDNA
sequences



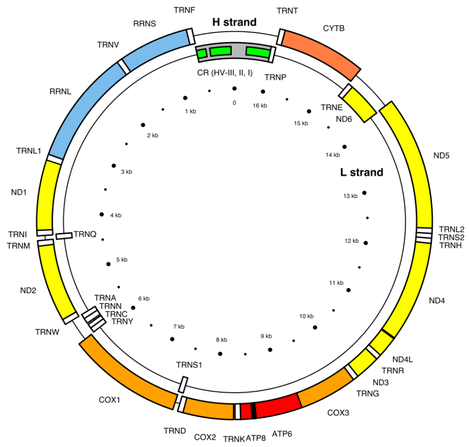
anatomically
modern
human
mtDNA
sequences

No introgression?

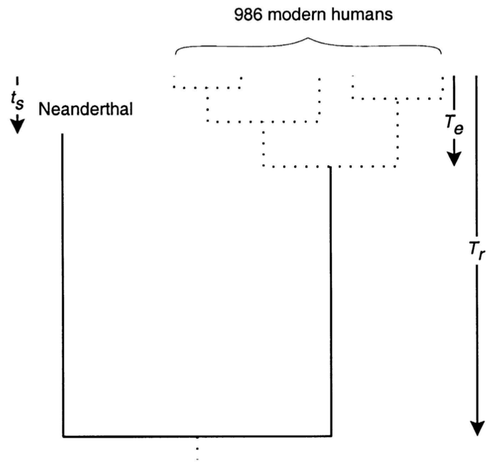


When the Neandertal mtDNA sequence is used to root a neighbor joining tree of modern human mtDNA sequences (Figure 7a), the first three branches consist exclusively of African sequences. The Neandertal mtDNA sequence thus supports a scenario in which modern humans arose recently in Africa as a distinct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

mitochondrial DNA



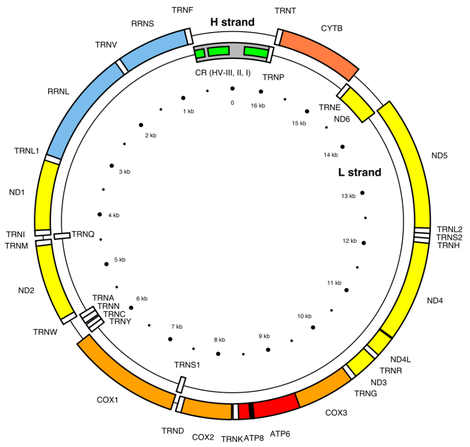
16 kb



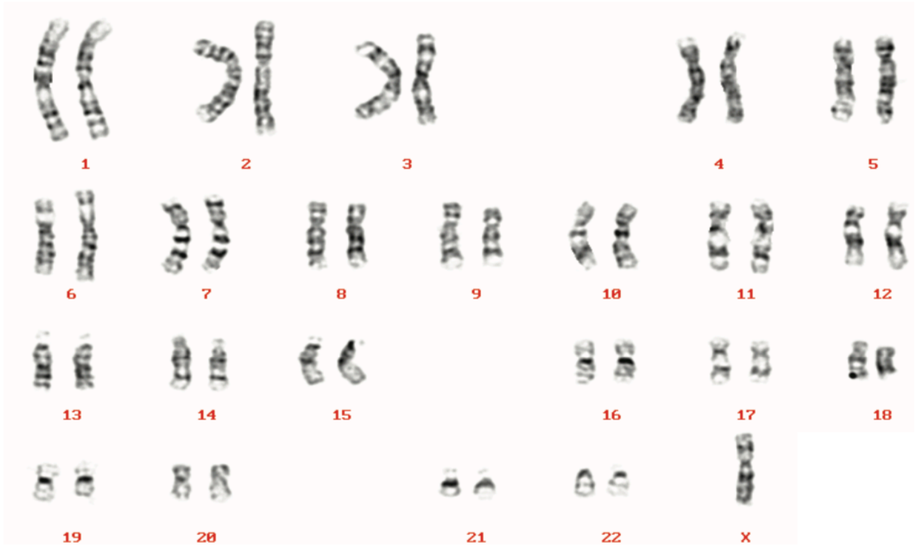
one mtDNA phylogenetic tree

(maternal history)

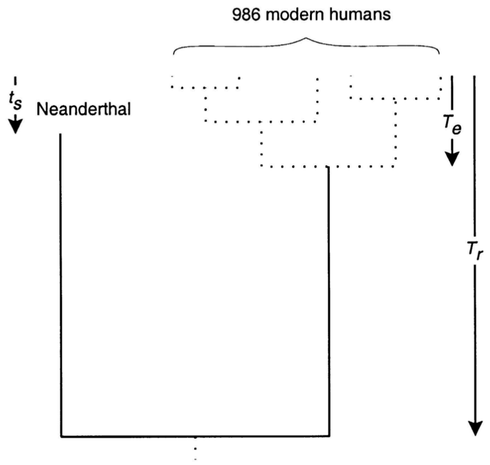
mitochondrial DNA **vs** nuclear DNA

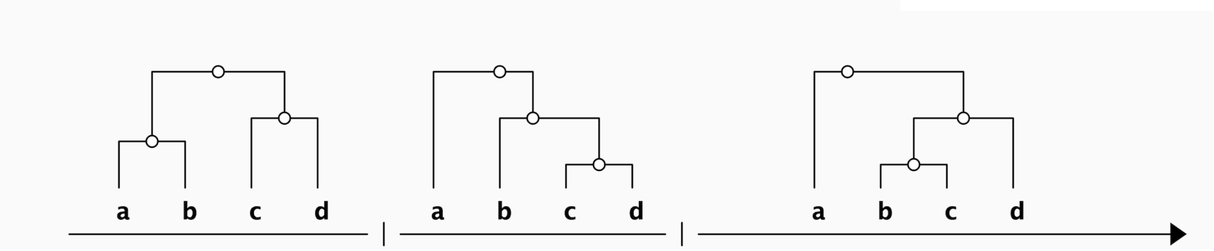


16 kb



3 Gb



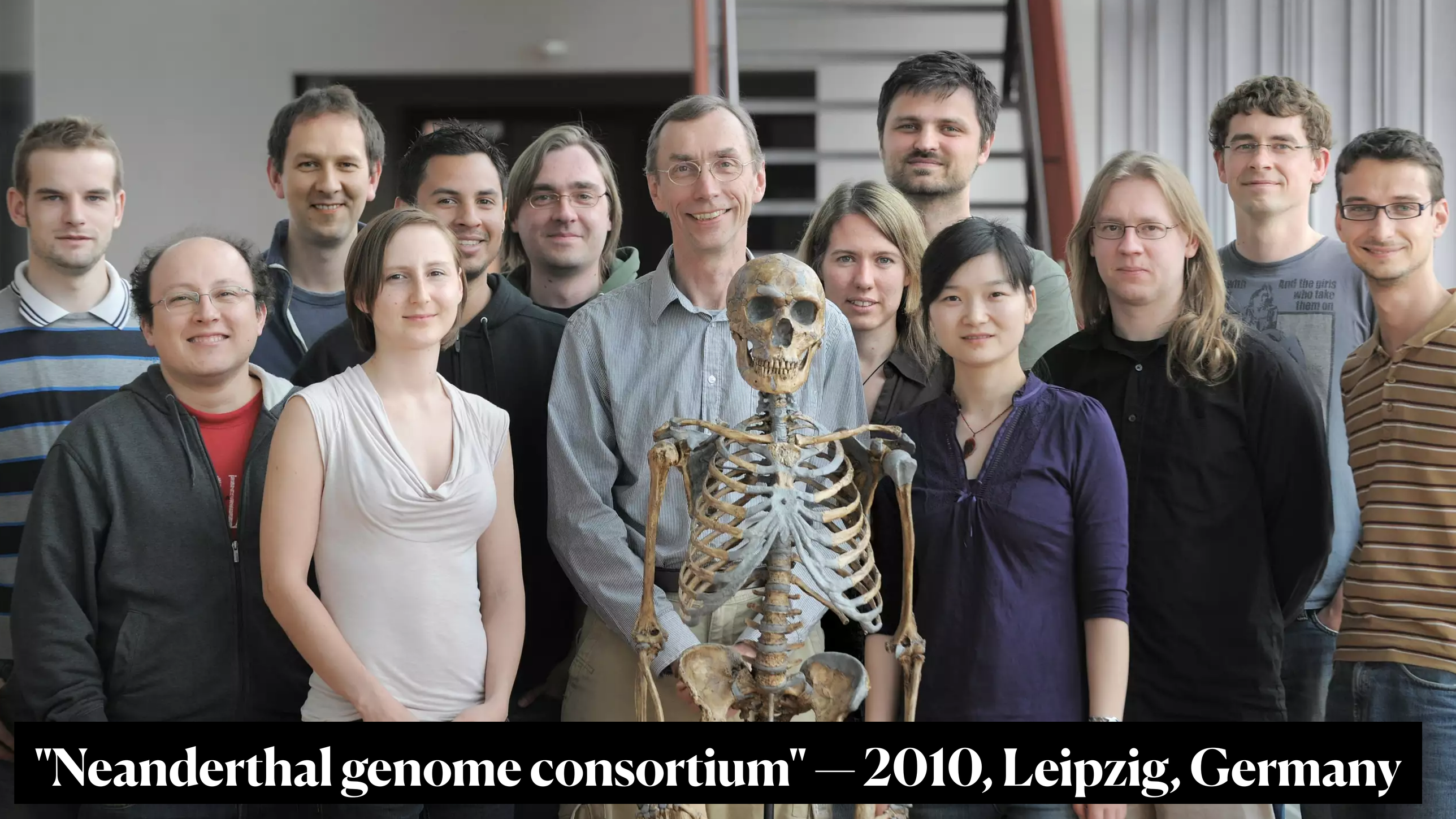


one mtDNA phylogenetic tree

(maternal history)

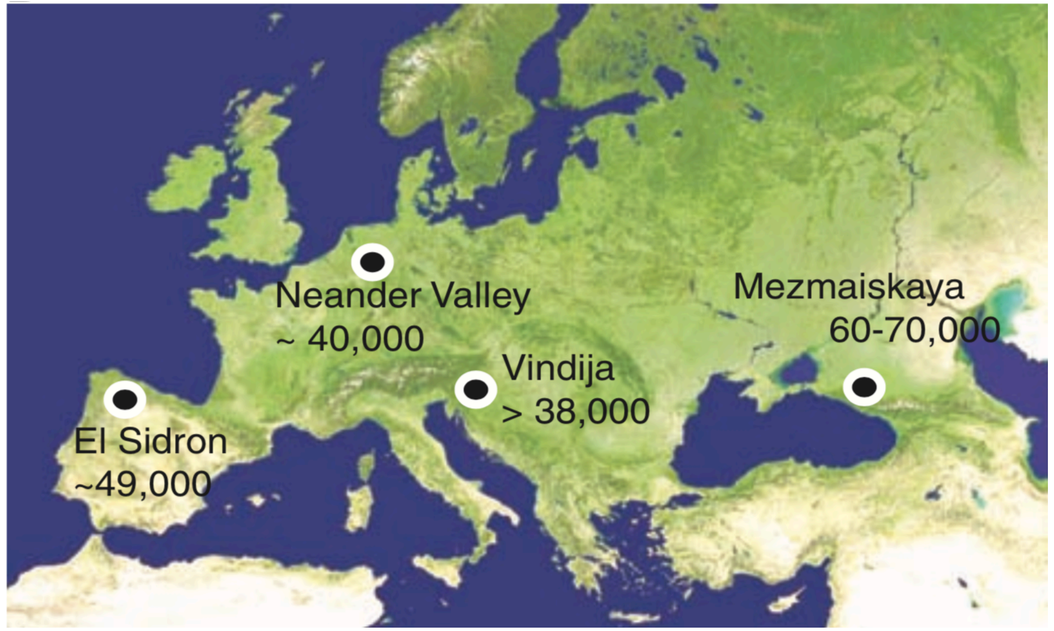
thousands of trees along the nuclear genome

(result of recombination)



"Neanderthal genome consortium" — 2010, Leipzig, Germany

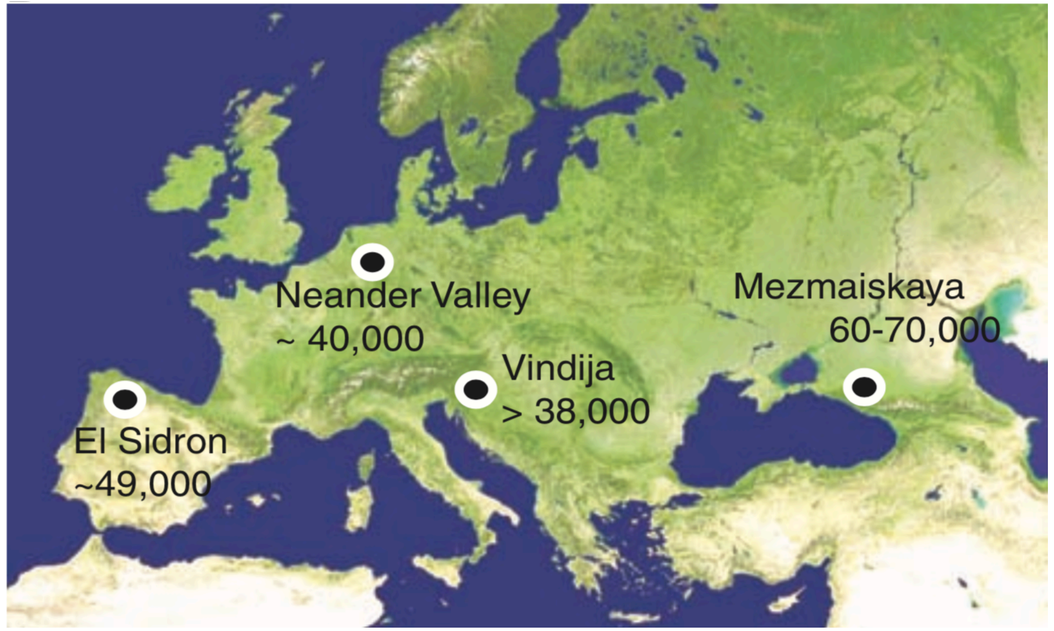
2010: "draft" Neanderthal genome



- DNA extracted from bone powder of three bones

Green *et al*. (Science 2010)

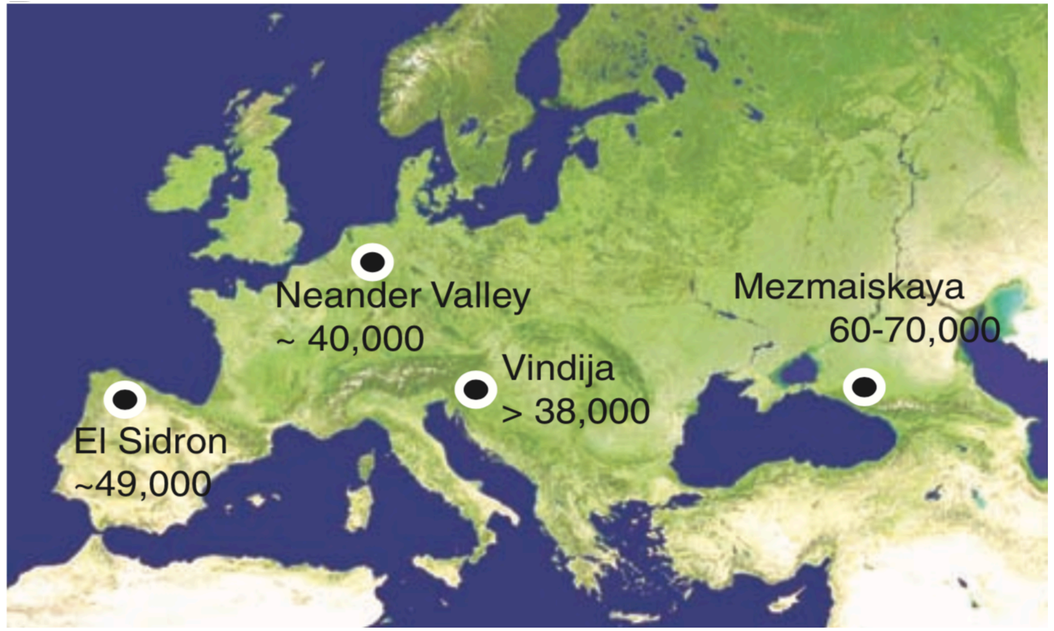
2010: "draft" Neanderthal genome



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

Green *et al*. (Science 2010)

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

Green *et al*. (Science 2010)

Typical aDNA workflow

Typical aDNA workflow

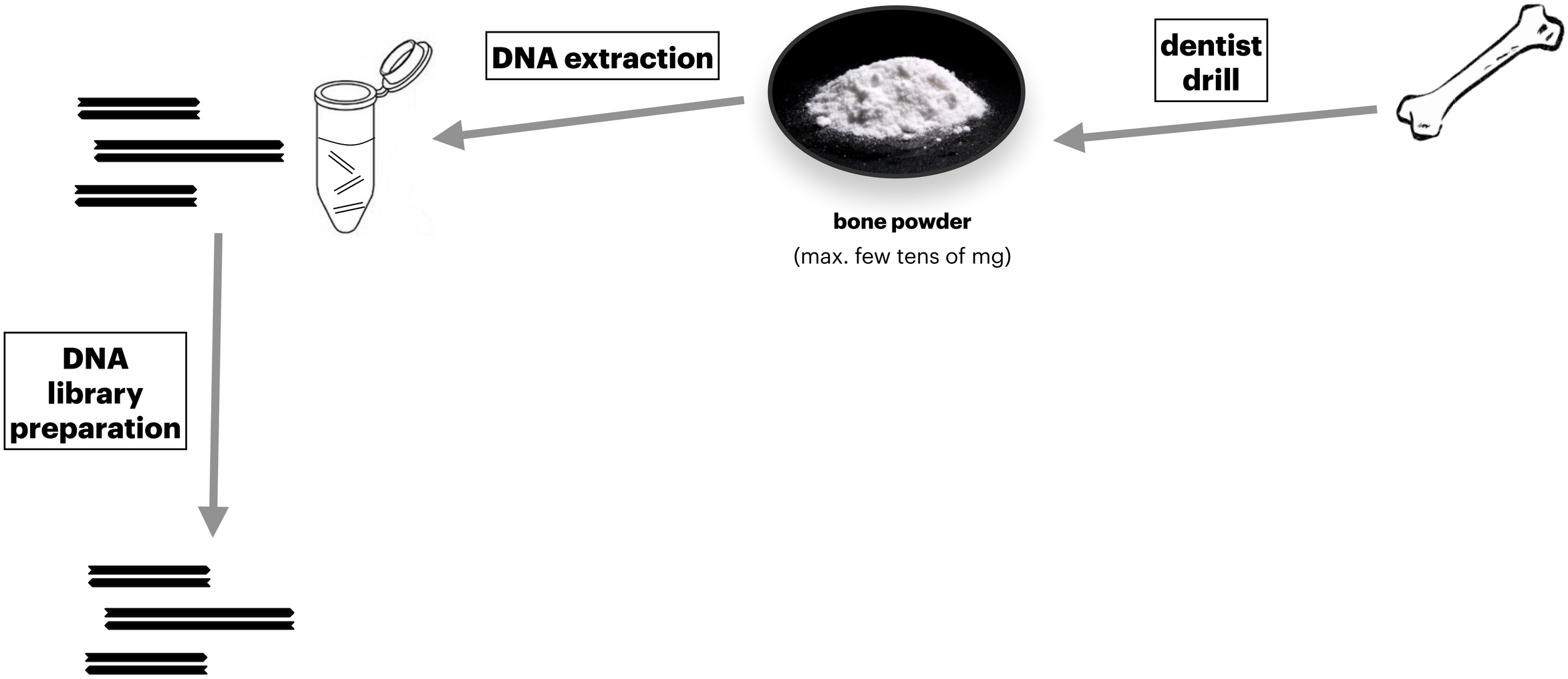
bone powder

(max. few tens of mg)

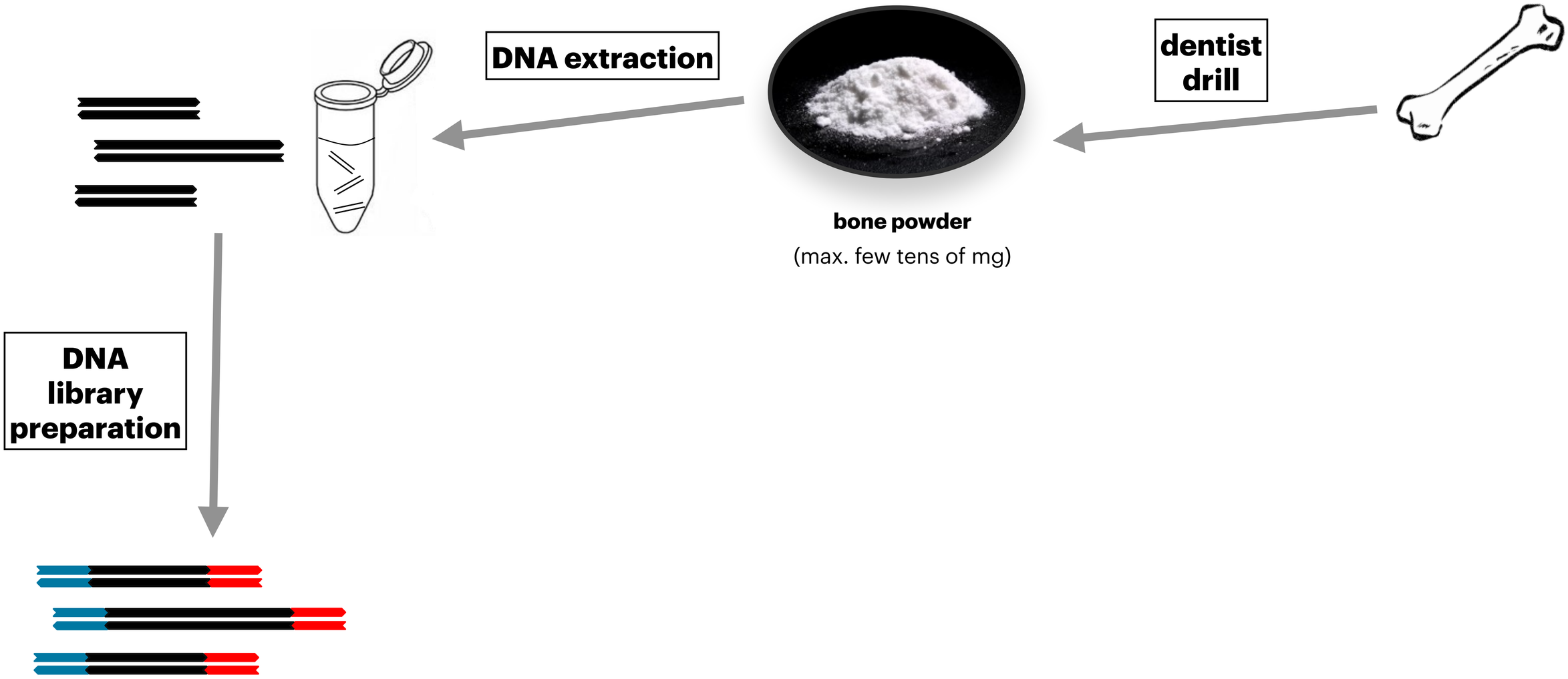
dentist drill

Typical aDNA workflow

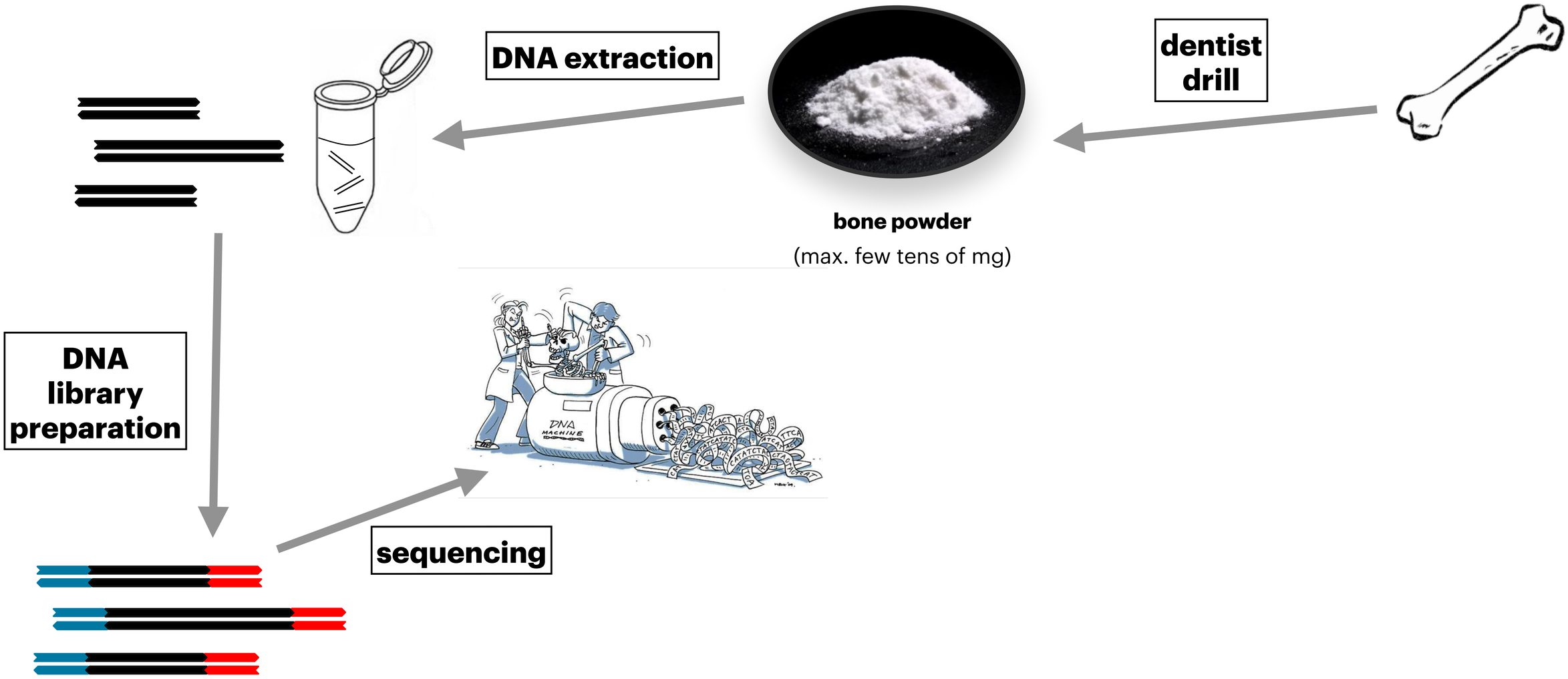
Typical aDNA workflow



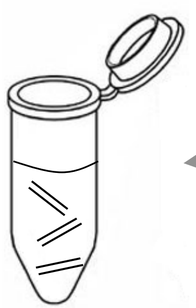
Typical aDNA workflow



Typical aDNA workflow



Typical aDNA workflow



DNA extraction

bone powder

(max. few tens of mg)

dentist drill

DNA library preparation

sequencing

bioinformatics

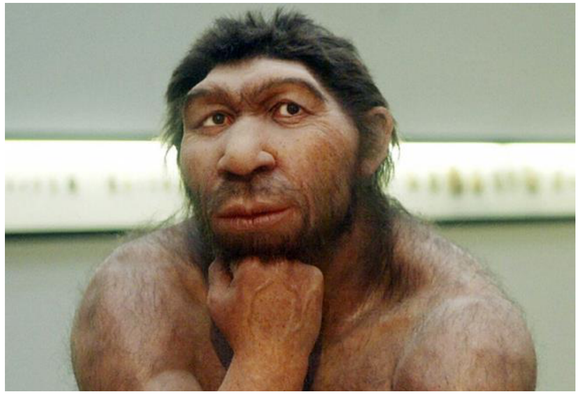
statistics

population genetics

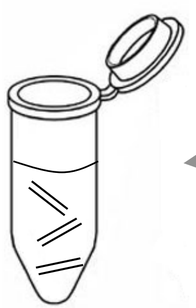
simulations

mathematical modeling

? ? ? ? ?



Typical aDNA workflow



DNA extraction

bone powder

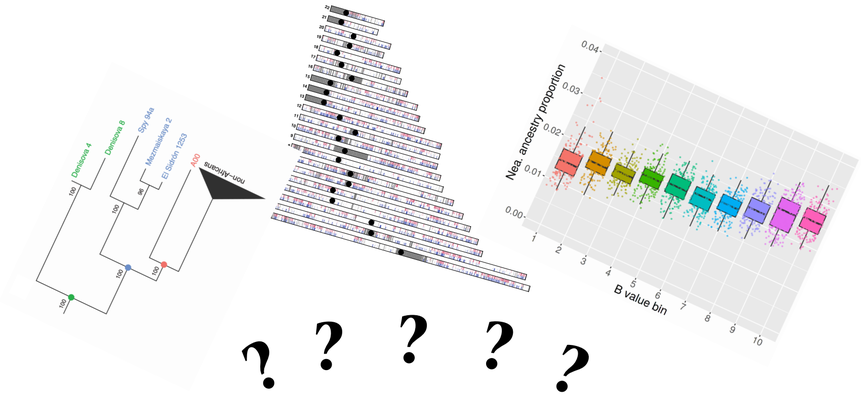
(max. few tens of mg)

dentist drill

DNA library preparation

sequencing

bioinformatics



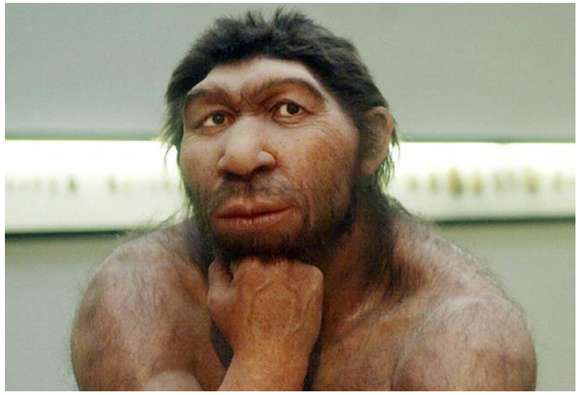
? ? ? ? ?

statistics

population genetics

simulations

mathematical modeling



So we have a Neanderthal genome… what now?

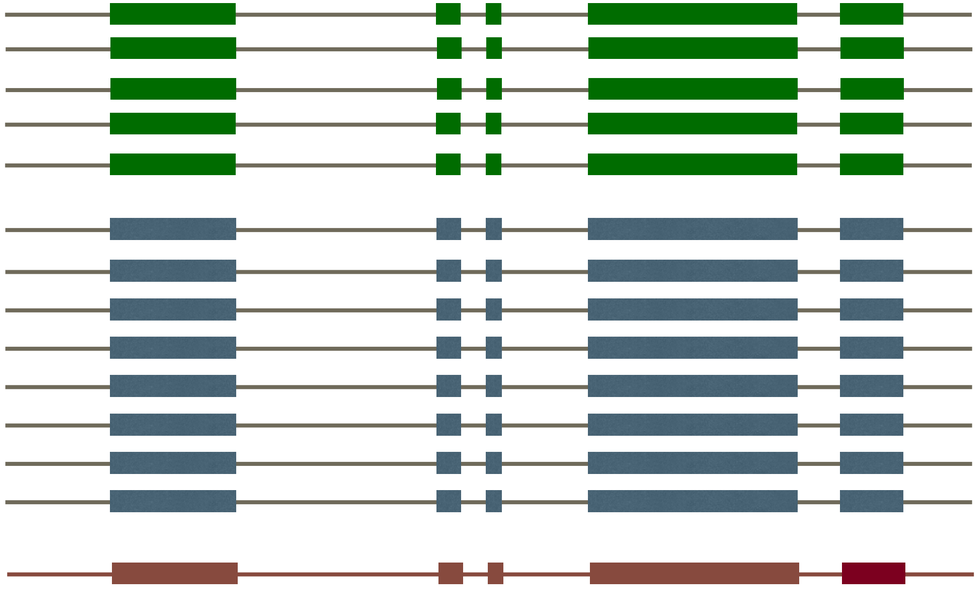
How to test the gene-flow hypothesis?

How to test the gene-flow hypothesis?

genomes of a reference population (Yorùbá)

genomes of a target population (Europeans)

Neanderthal genome



David Reich

Nick Patterson

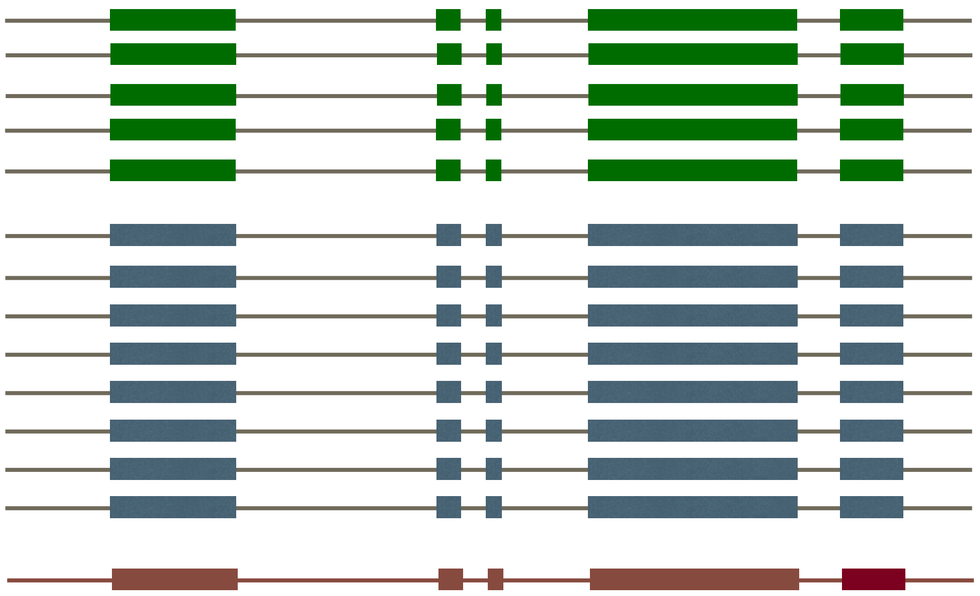
Green *et al*. (Science 2010)

How to test the gene-flow hypothesis?

genomes of a reference population (Yorùbá)

genomes of a target population (Europeans)

Neanderthal genome



G A T T A C A • • •

David Reich

Nick Patterson

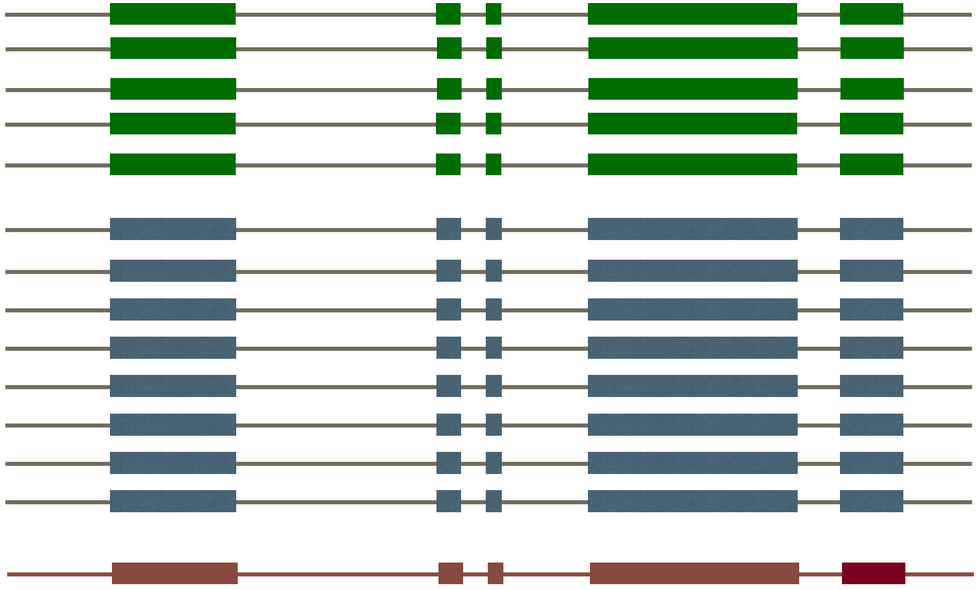
Green *et al*. (Science 2010)

How to test the gene-flow hypothesis?

genomes of a reference population (Yorùbá)

genomes of a target population (Europeans)

Neanderthal genome



G A T T A C A • • •

↑↑↑↑↑↑↑ • • •

? ? ? ? ? ? ?

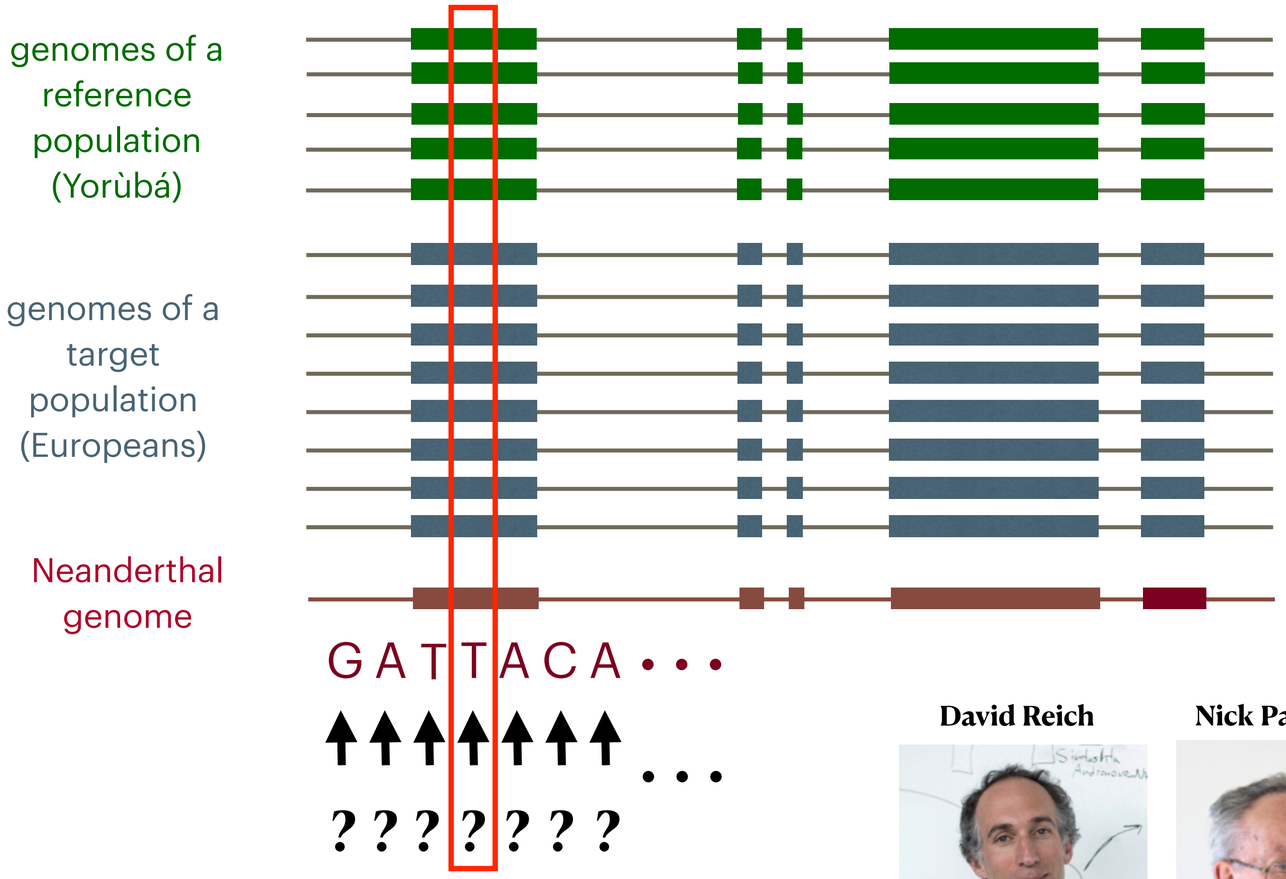
David Reich

Nick Patterson

Green *et al*. (Science 2010)

How to test the gene-flow hypothesis?

Situation at one locus...



David Reich

Nick Patterson

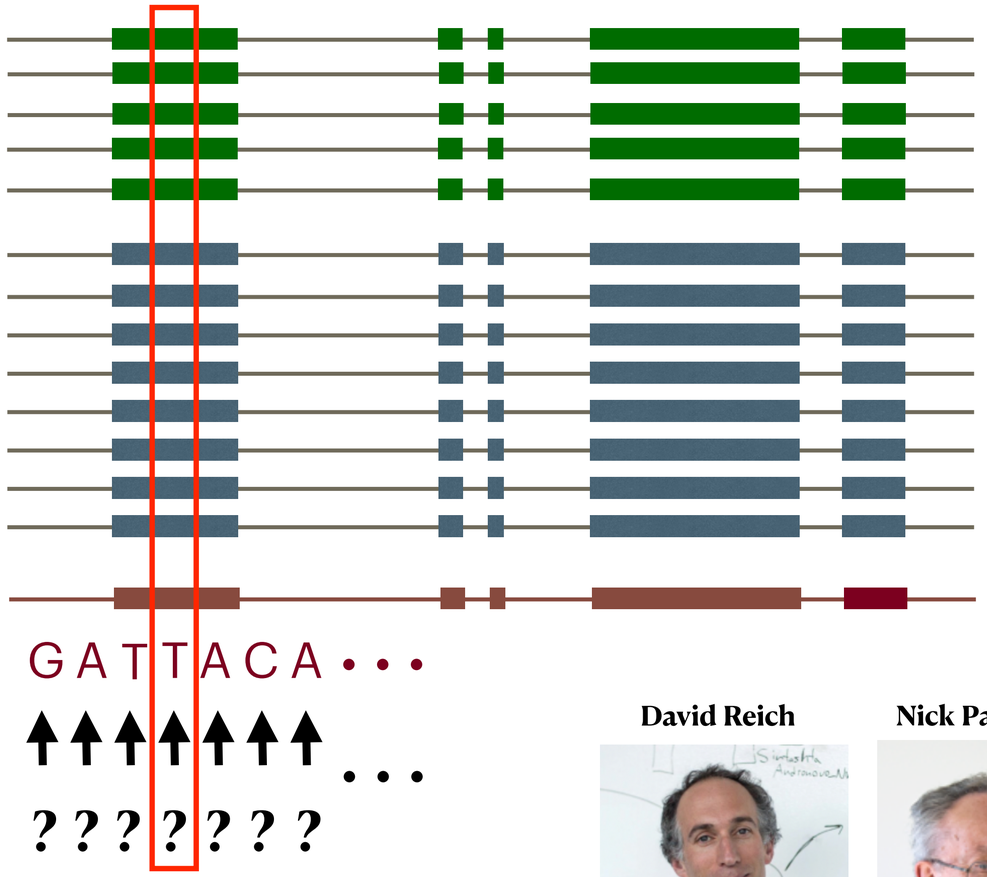
Green *et al*. (Science 2010)

How to test the gene-flow hypothesis?

genomes of a reference population (Yorùbá)

genomes of a target population (Europeans)

Neanderthal genome



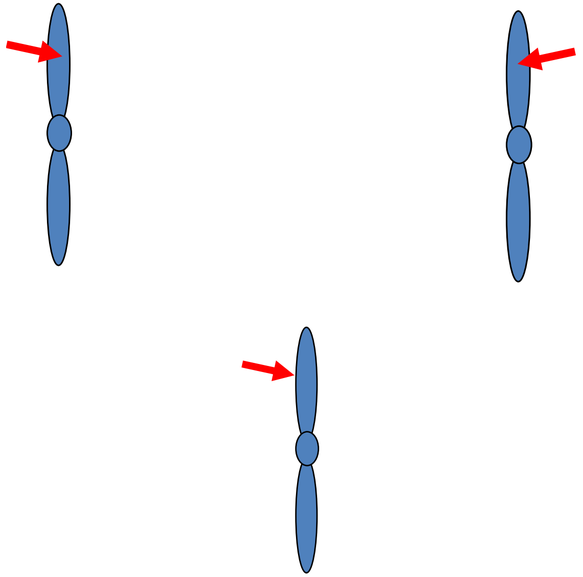
David Reich

Nick Patterson

Situation at one locus...

person of an African ancestry

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



Neanderthal

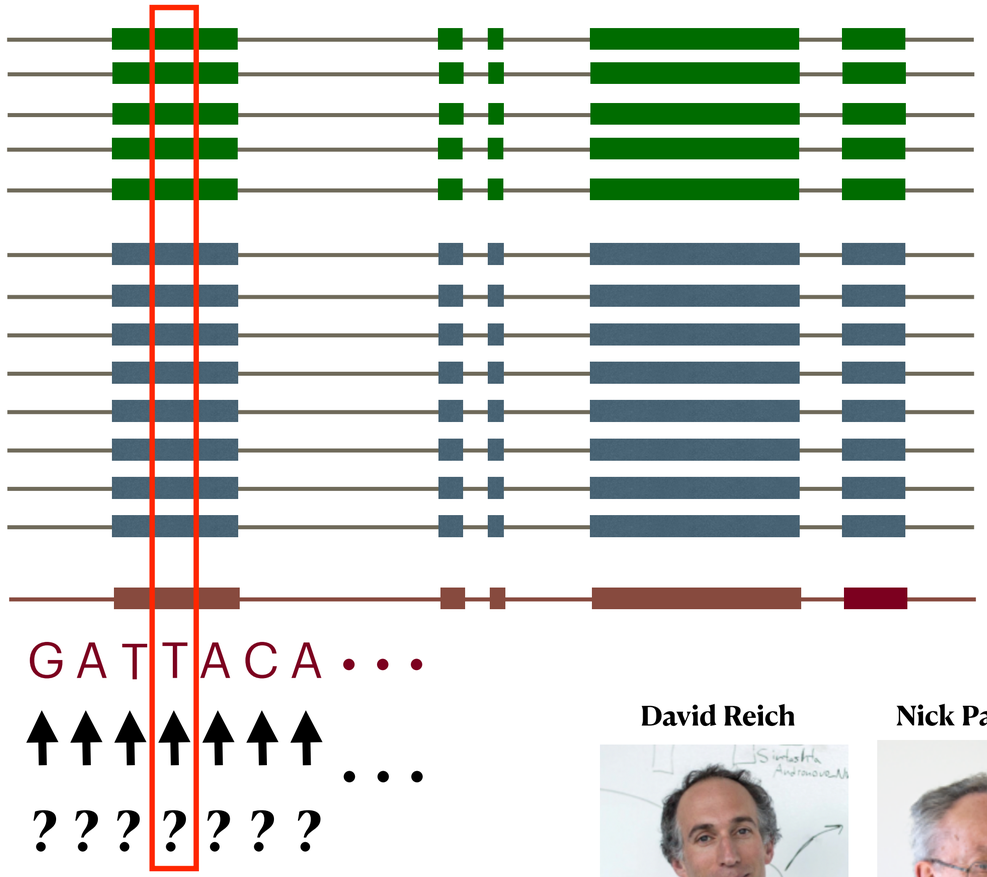
Green *et al*. (Science 2010)

How to test the gene-flow hypothesis?

genomes of a reference population (Yorùbá)

genomes of a target population (Europeans)

Neanderthal genome



David Reich

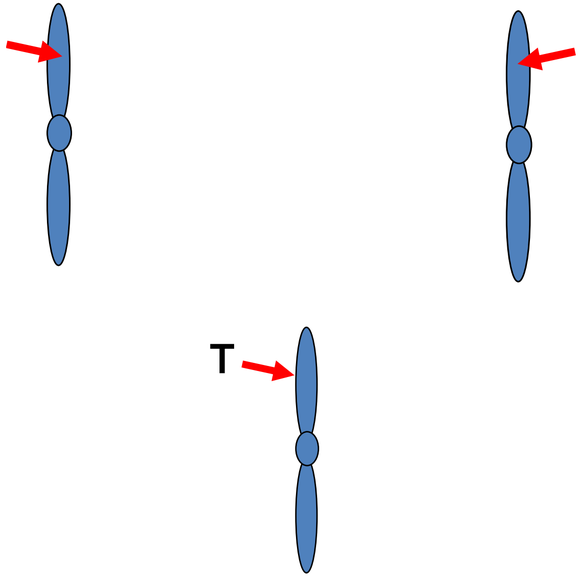
Nick Patterson

Green *et al*. (Science 2010)

Situation at one locus...

person of an African ancestry

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



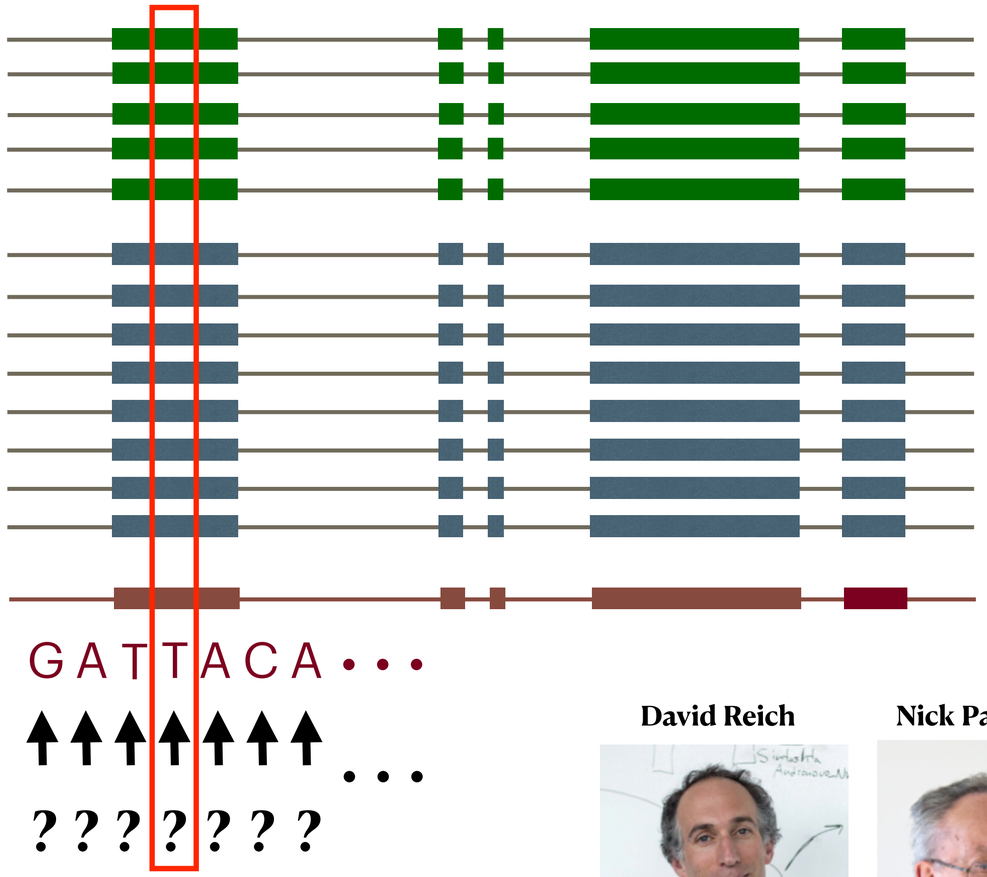
Neanderthal

How to test the gene-flow hypothesis?

genomes of a reference population (Yorùbá)

genomes of a target population (Europeans)

Neanderthal genome



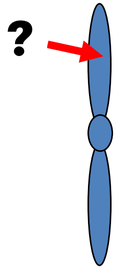
David Reich

Nick Patterson

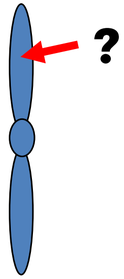
Green *et al*. (Science 2010)

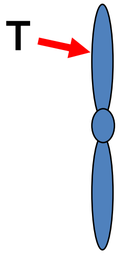
Situation at one locus...

person of an African ancestry



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





Neanderthal

Formal test of introgression

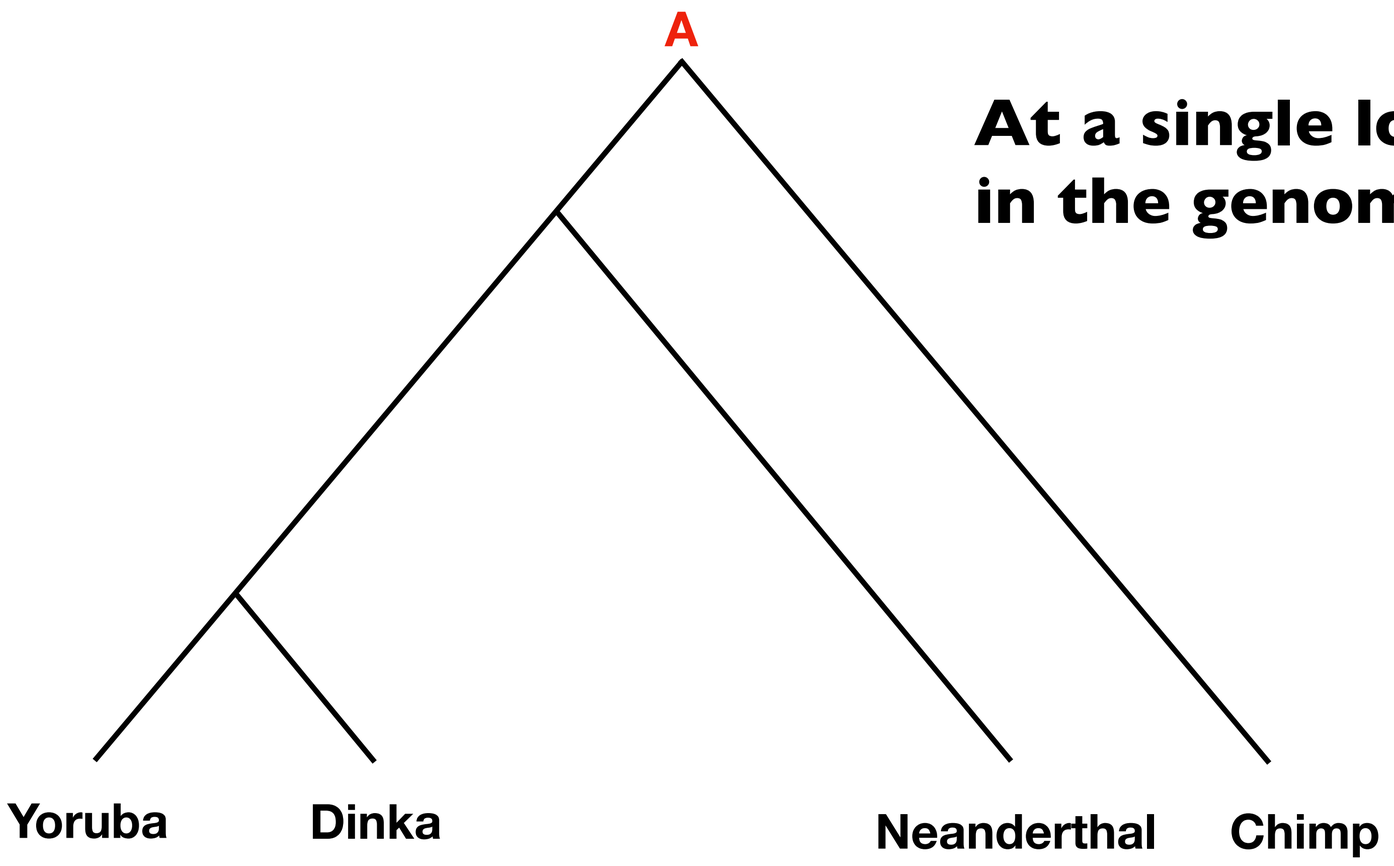
*f*4 statistic

Given a set of 4 samples, *f4* compares counts of observed **BABA vs ABBA** site patterns.

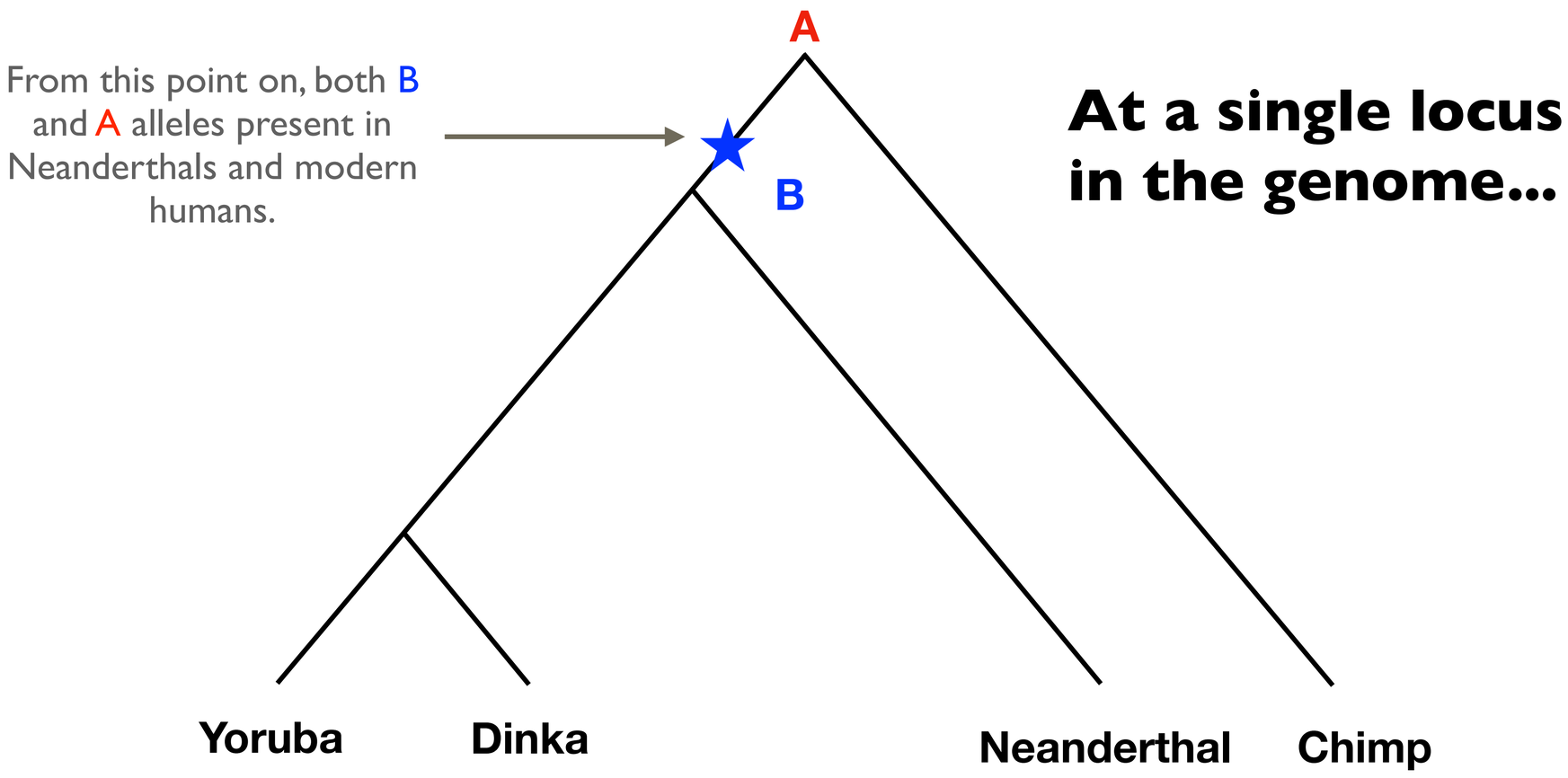
A variant of this statistic that is used equally often is known as ***D statistic***.

Nick Patterson

David Reich

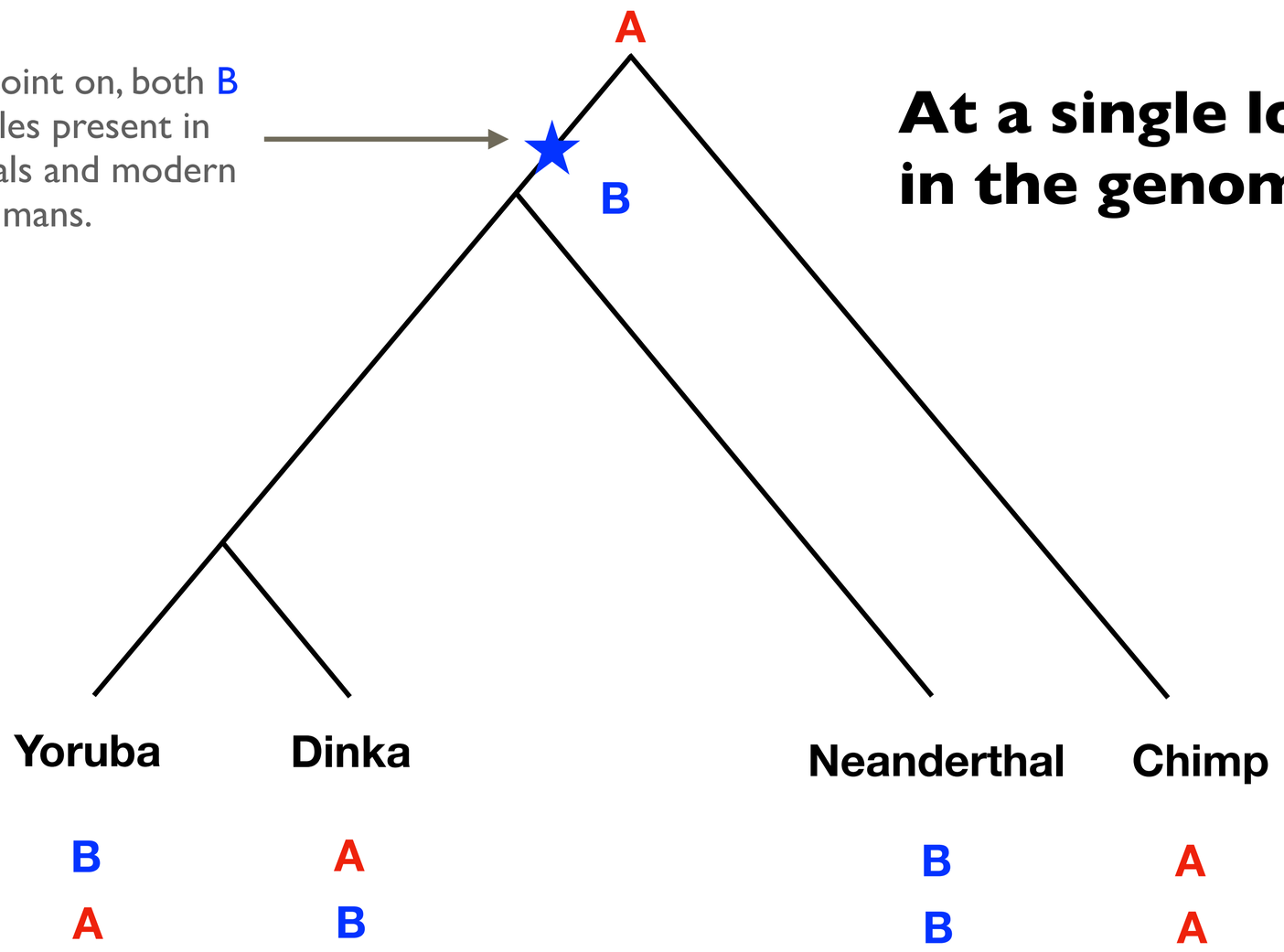


At a single locus in the genome...

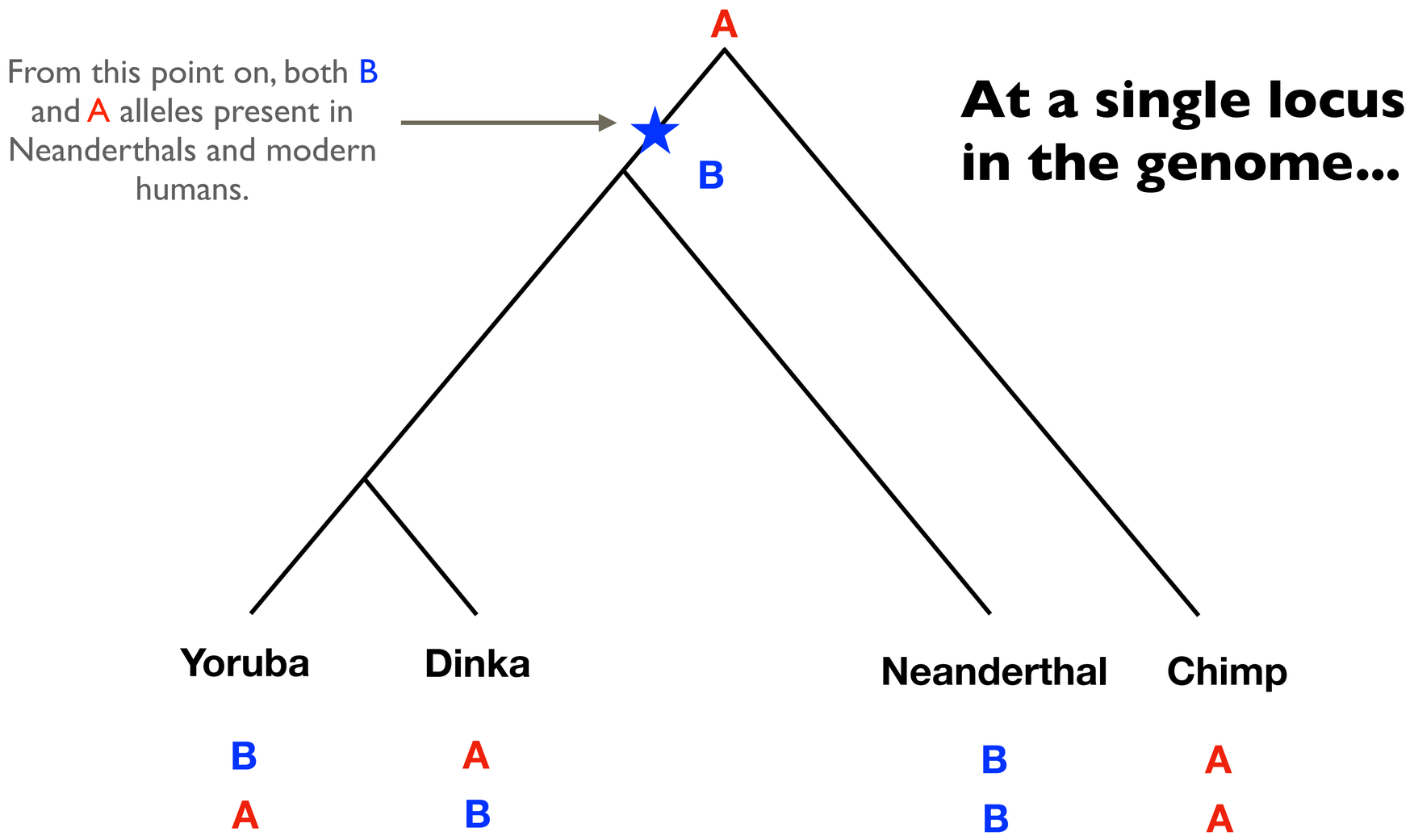


At a single locus in the genome...

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



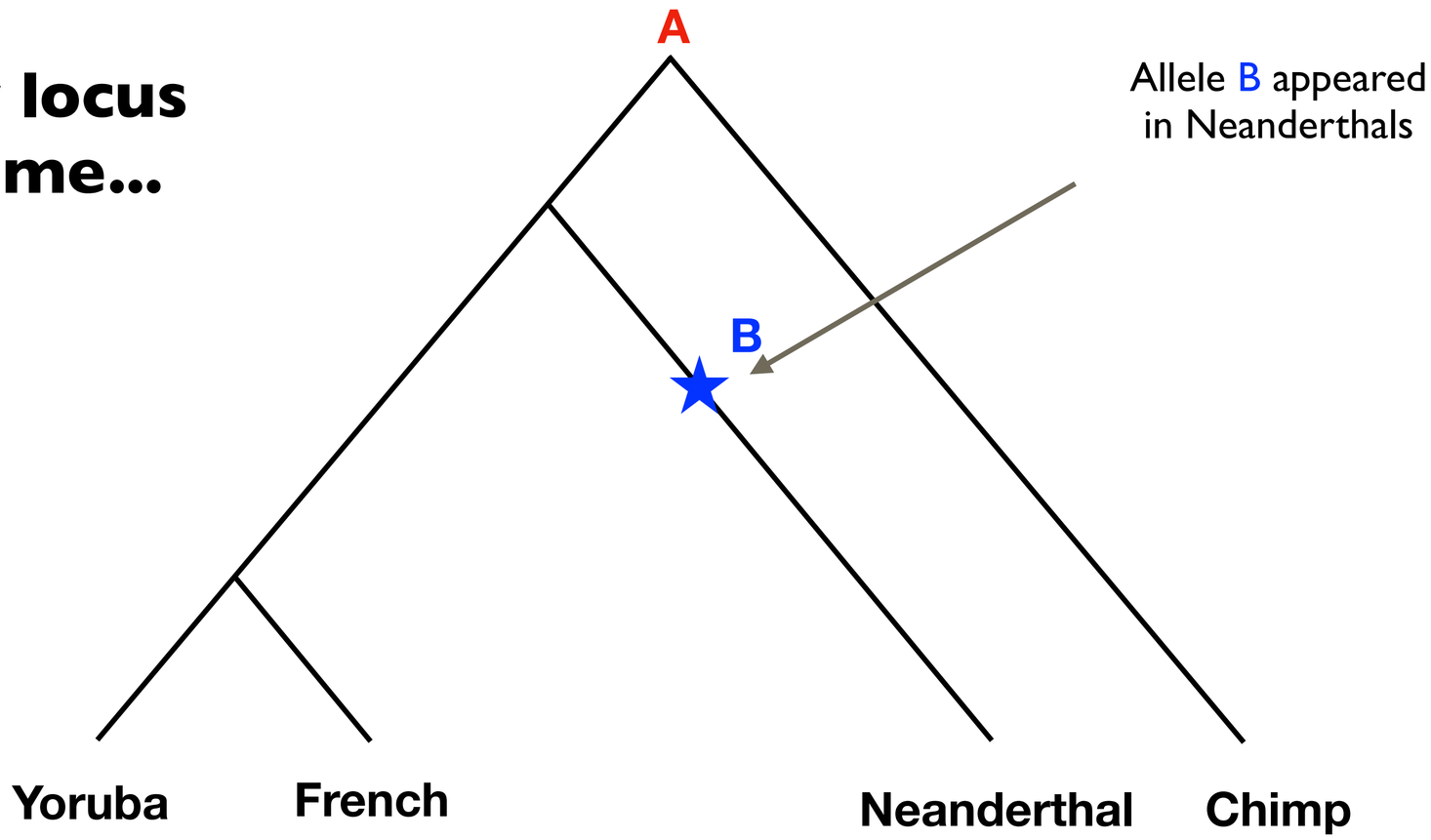
At a single locus in the genome...



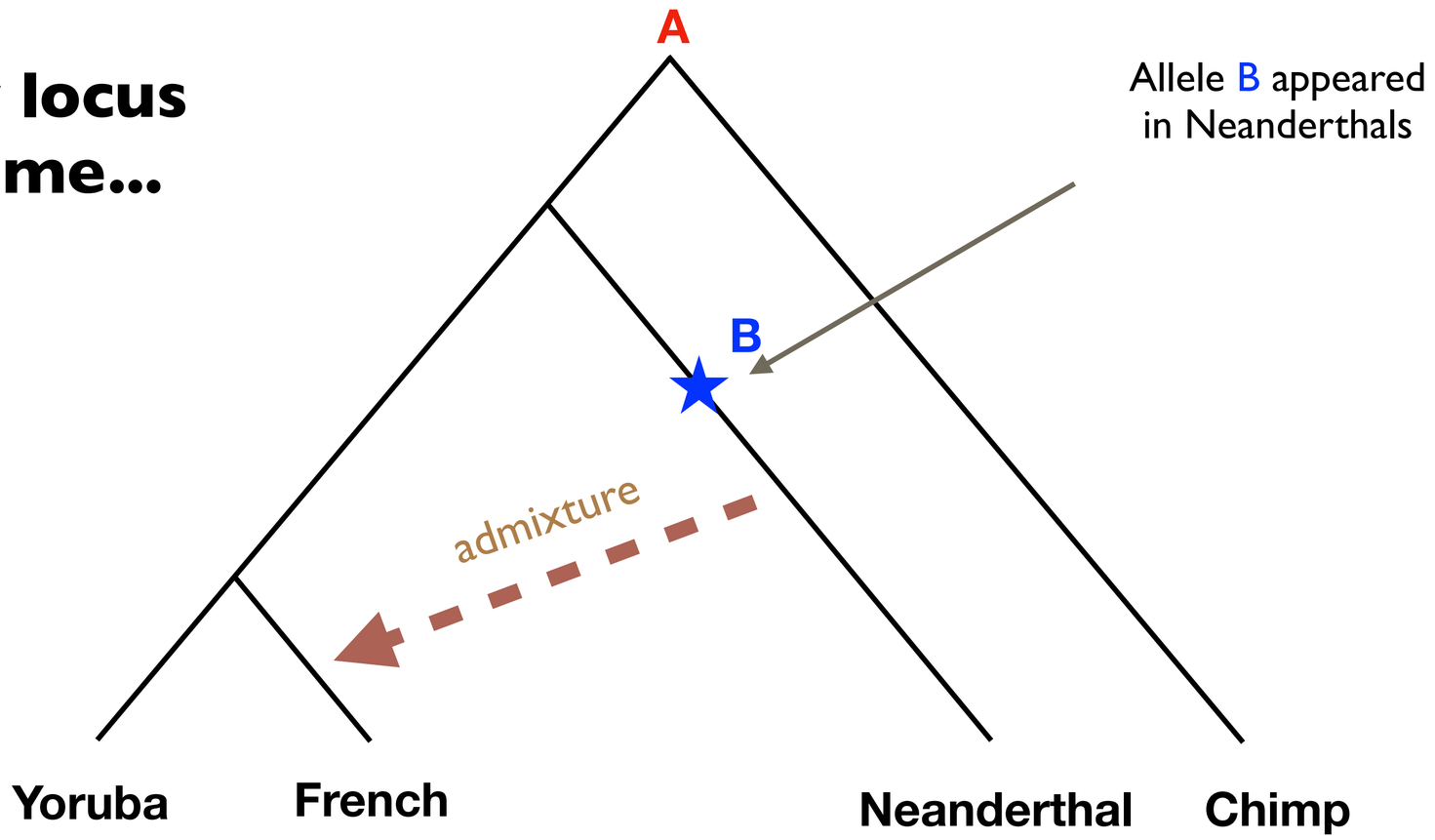
At a single locus in the genome...

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

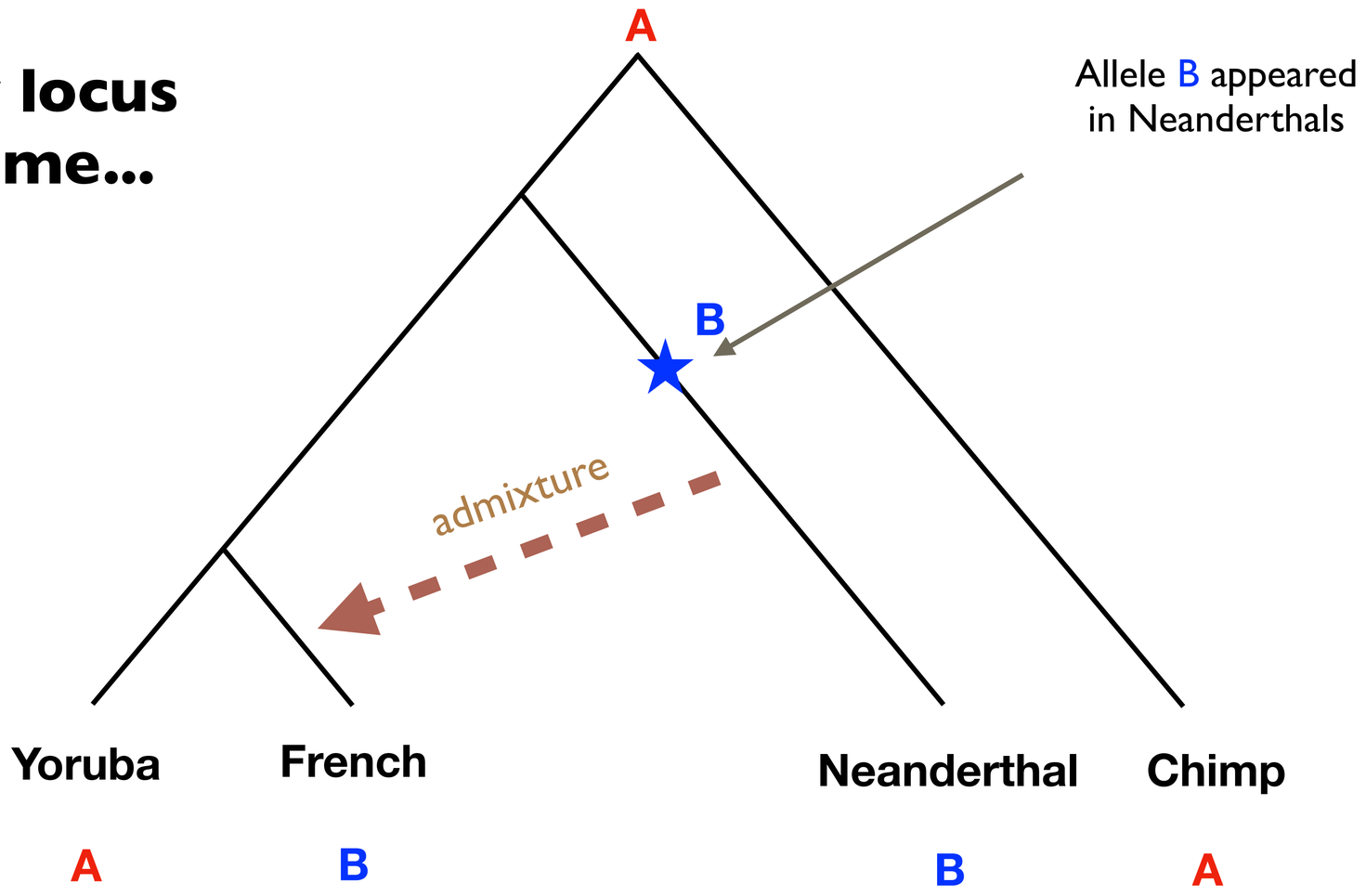
At another locus in the genome...



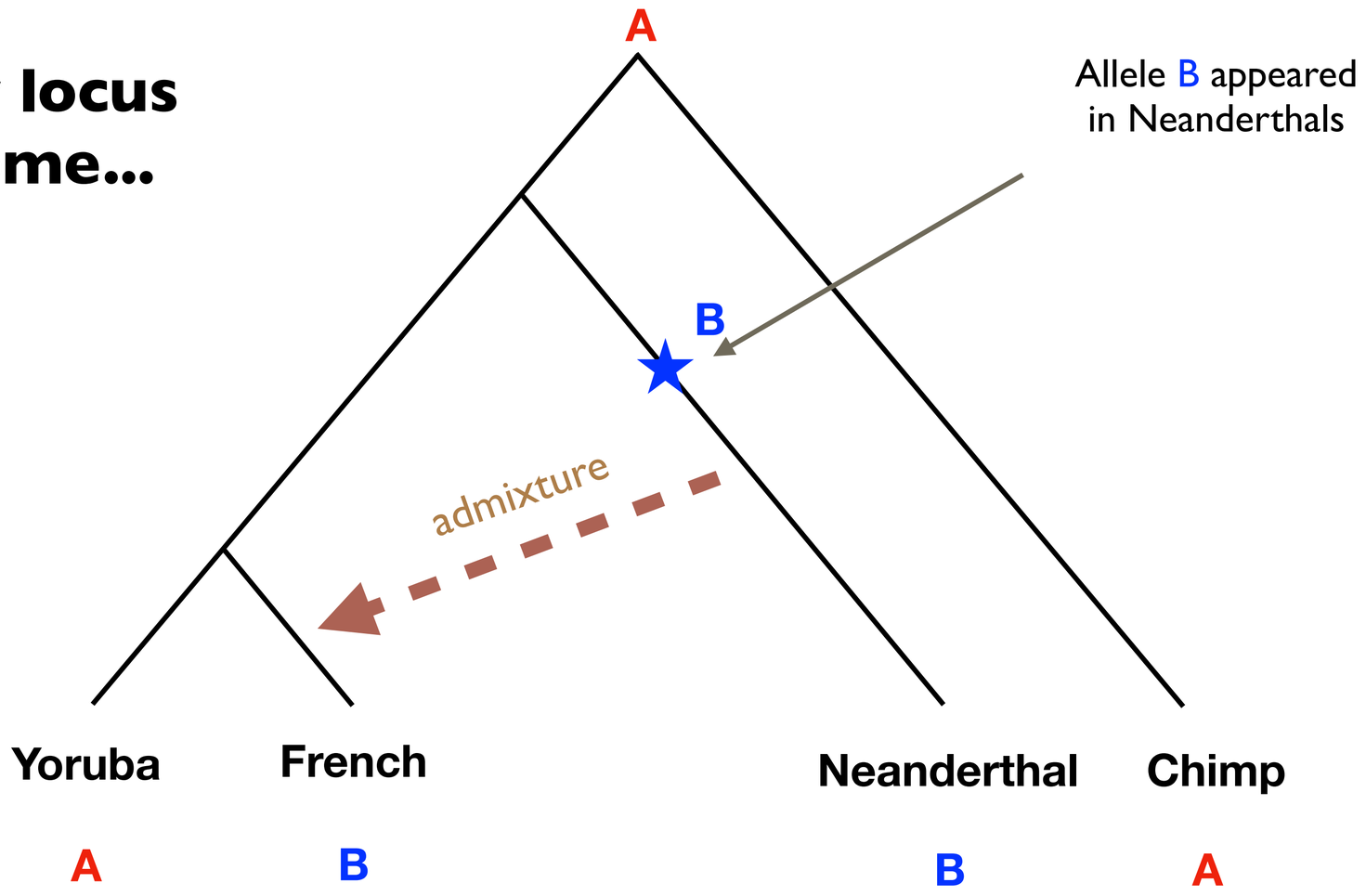
At another locus in the genome...



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(Yoruba, X; Neanderthal, Chimp) … where X = Dinka or French

Test of Neanderthal admixture **in practice**

f_4(Yoruba, X; Neanderthal, Chimp) … where X = Dinka or French

f_4 = (# BABA - # ABBA) / # SNPs

Test of Neanderthal admixture **in practice**

f_4(Yoruba, X; Neanderthal, Chimp) … where X = Dinka or French

f_4 = (# BABA - # ABBA) / # SNPs

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

Test of Neanderthal admixture **in practice**

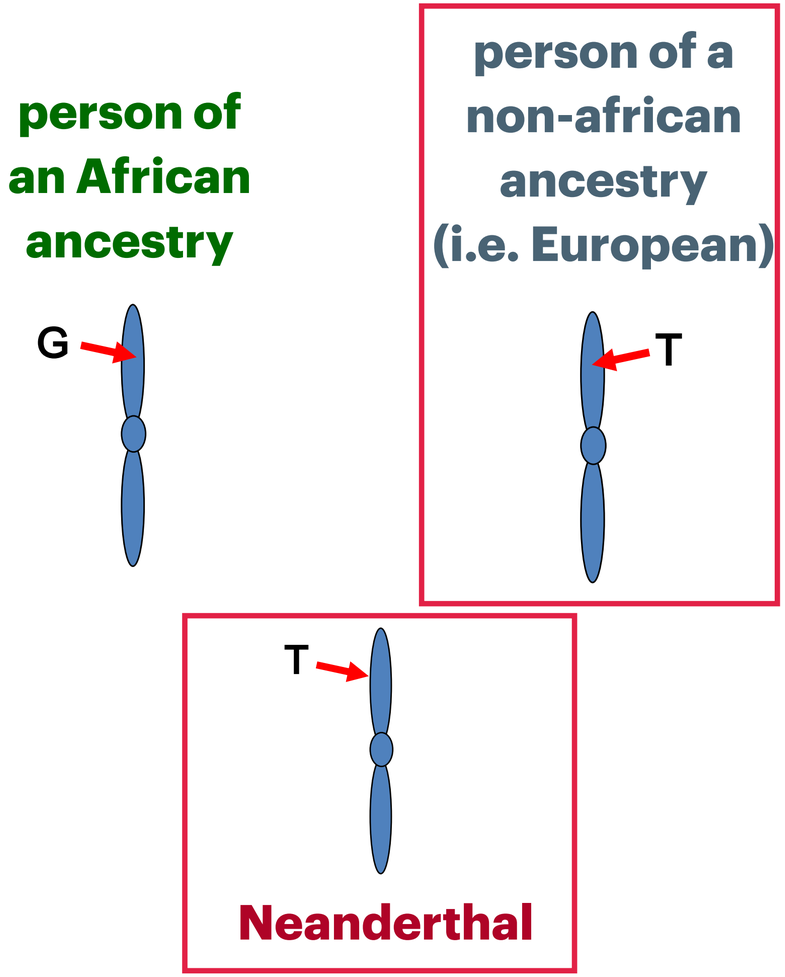
f_4(Yoruba, X; Neanderthal, Chimp) … where X = Dinka or French

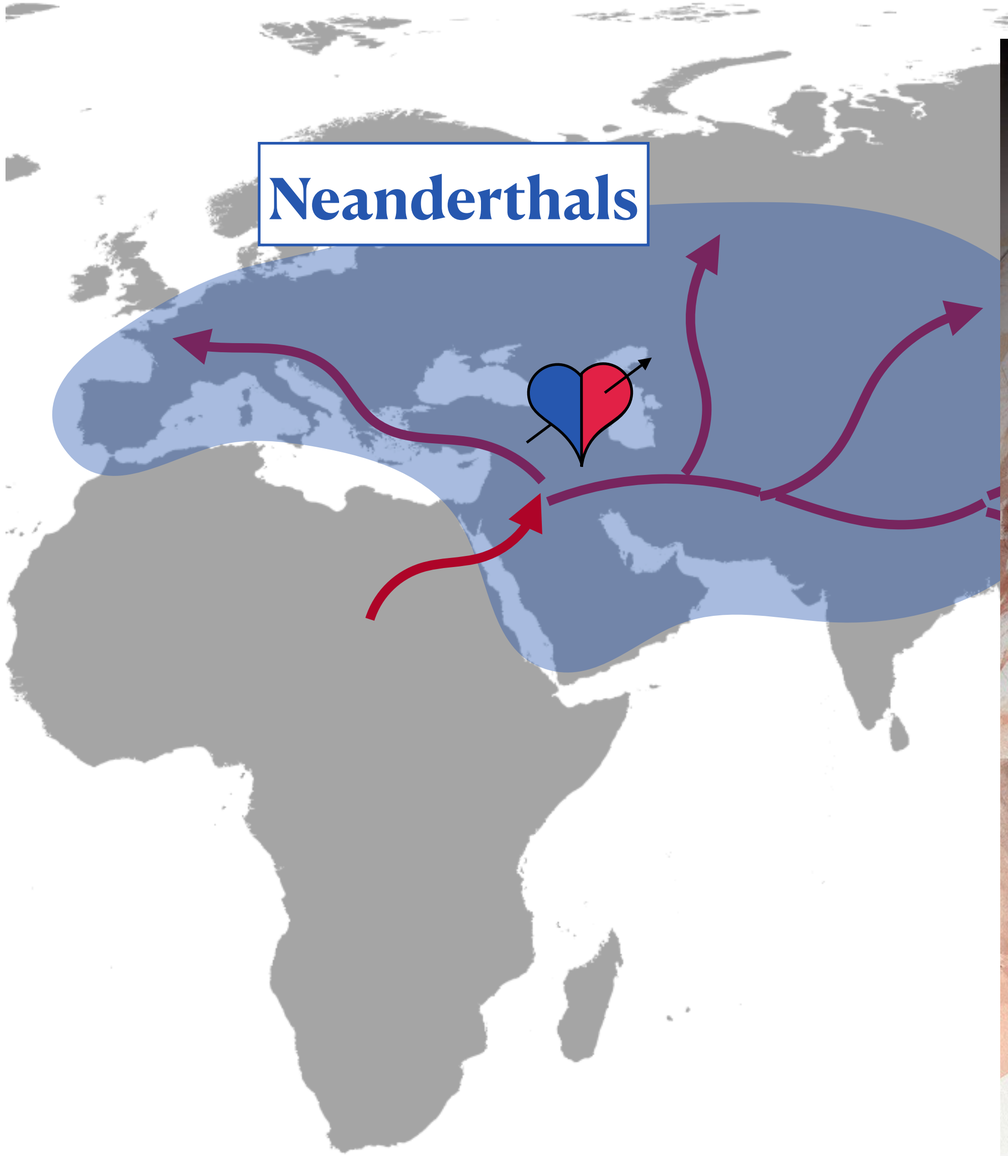
f_4 = (# BABA - # ABBA) / # SNPs

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

|Z score| > 3 considered significant

Test of Neanderthal admixture **in practice**




Neanderthals

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(Yoruba, **X**; Altai, Chimp)

X	#BABA	#ABBA	#sites	f4
French	44409	46865	1436967	-0.0017
Dinka	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(Yoruba, **X**; Altai, Chimp)

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

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(Yoruba, **X**; Altai, Chimp) / *f4*(Yoruba, **Vindija**; Altai, Chimp)

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

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(Yoruba, **X**; Altai, Chimp) / *f4*(Yoruba, **Vindija**; Altai, Chimp)

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

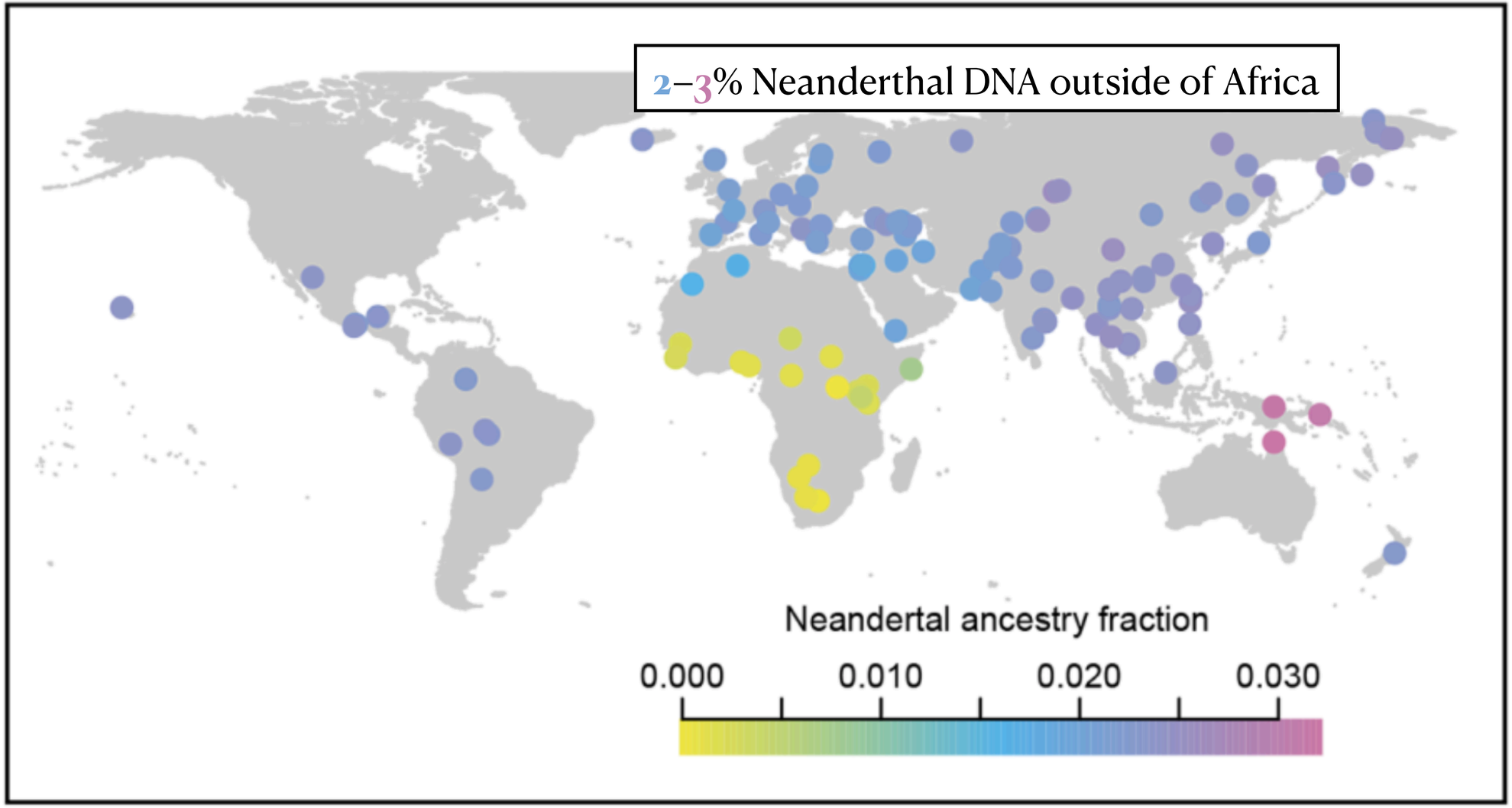
What about the proportion of Neanderthal ancestry?

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

f4(Yoruba, **X**; Altai, Chimp) / *f4*(Yoruba, **Vindija**; Altai, Chimp)

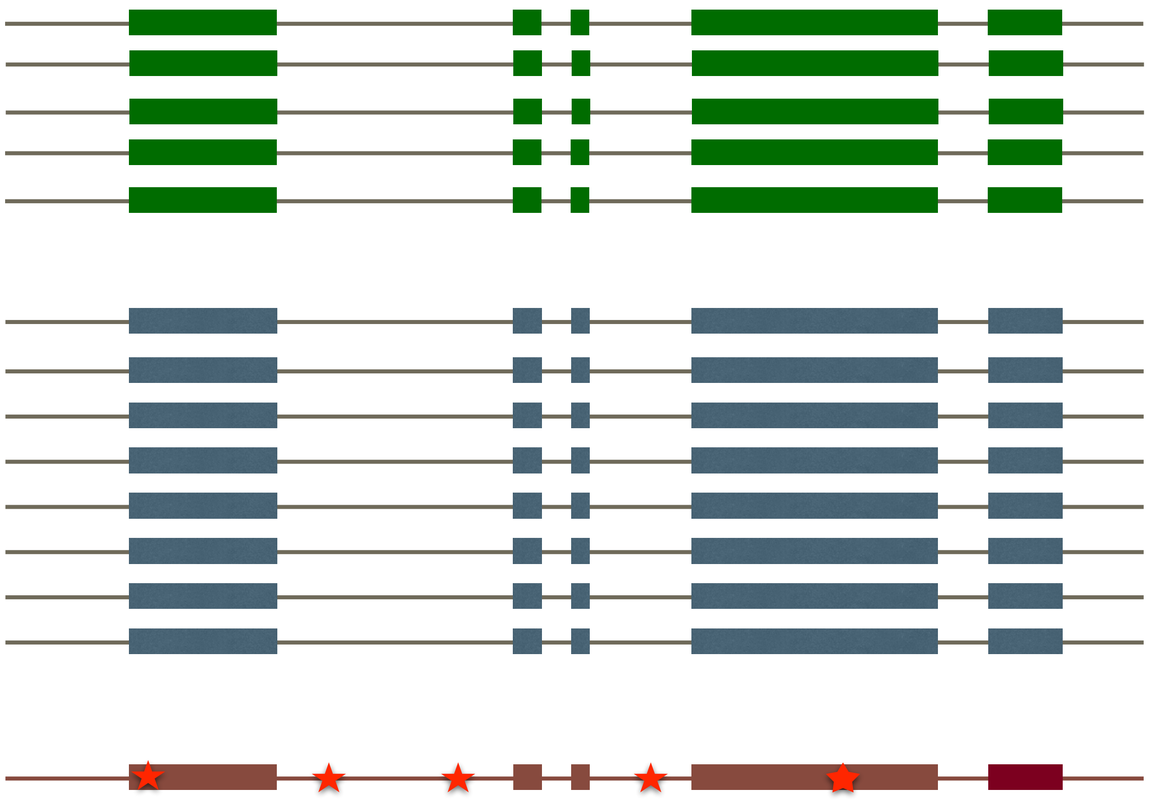
X	#BABA	#ABBA	#sites	f4	ratio
French	44409	46865	1436967	-0.0017	**2.2%**
Dinka	43025	43182	1436978	-0.0001	0.1%
Vindija	7337	118956	1436703	-0.0777	**100%**

Proportion of Neanderthal DNA in humans today



Pruefer *et al*. (Science 2017)

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



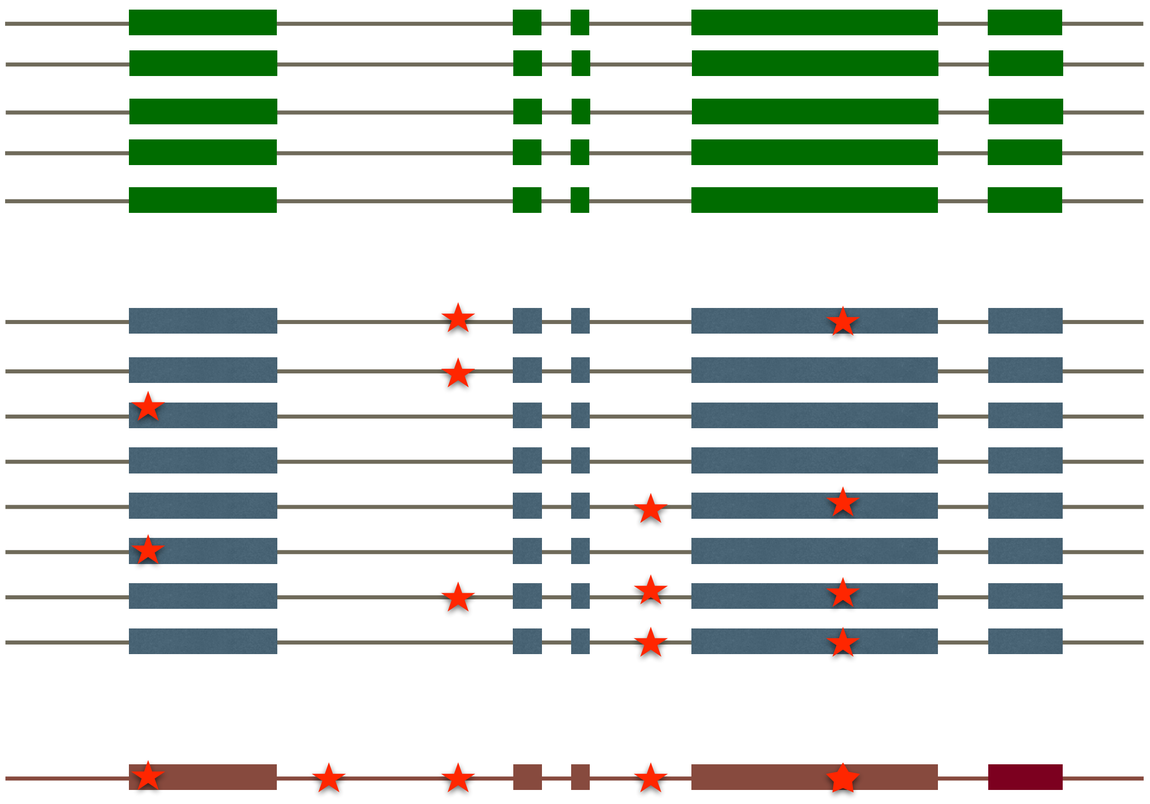
genomes of a reference population (Yorùbá)

genomes of a target population (Europeans)

Neanderthal genome

★ = Neanderthal-derived mutation

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



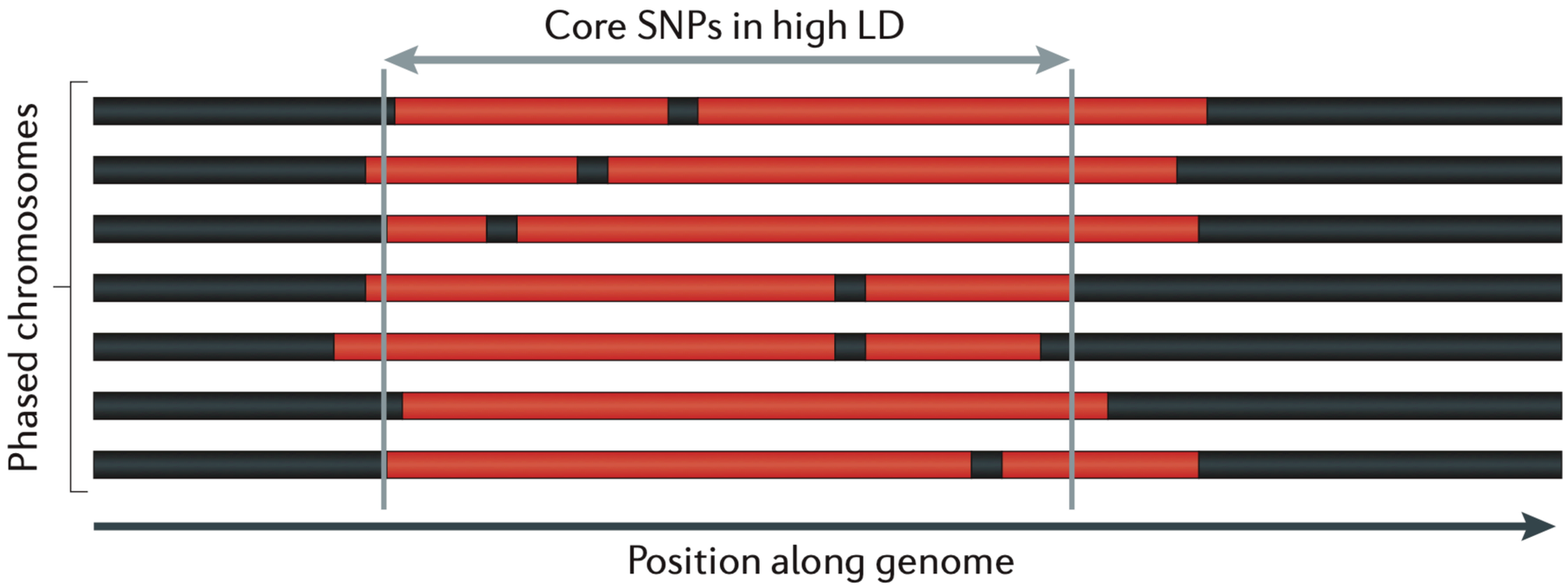
genomes of a reference population (Yorùbá)

genomes of a target population (Europeans)

Neanderthal genome

★ = Neanderthal-derived mutation

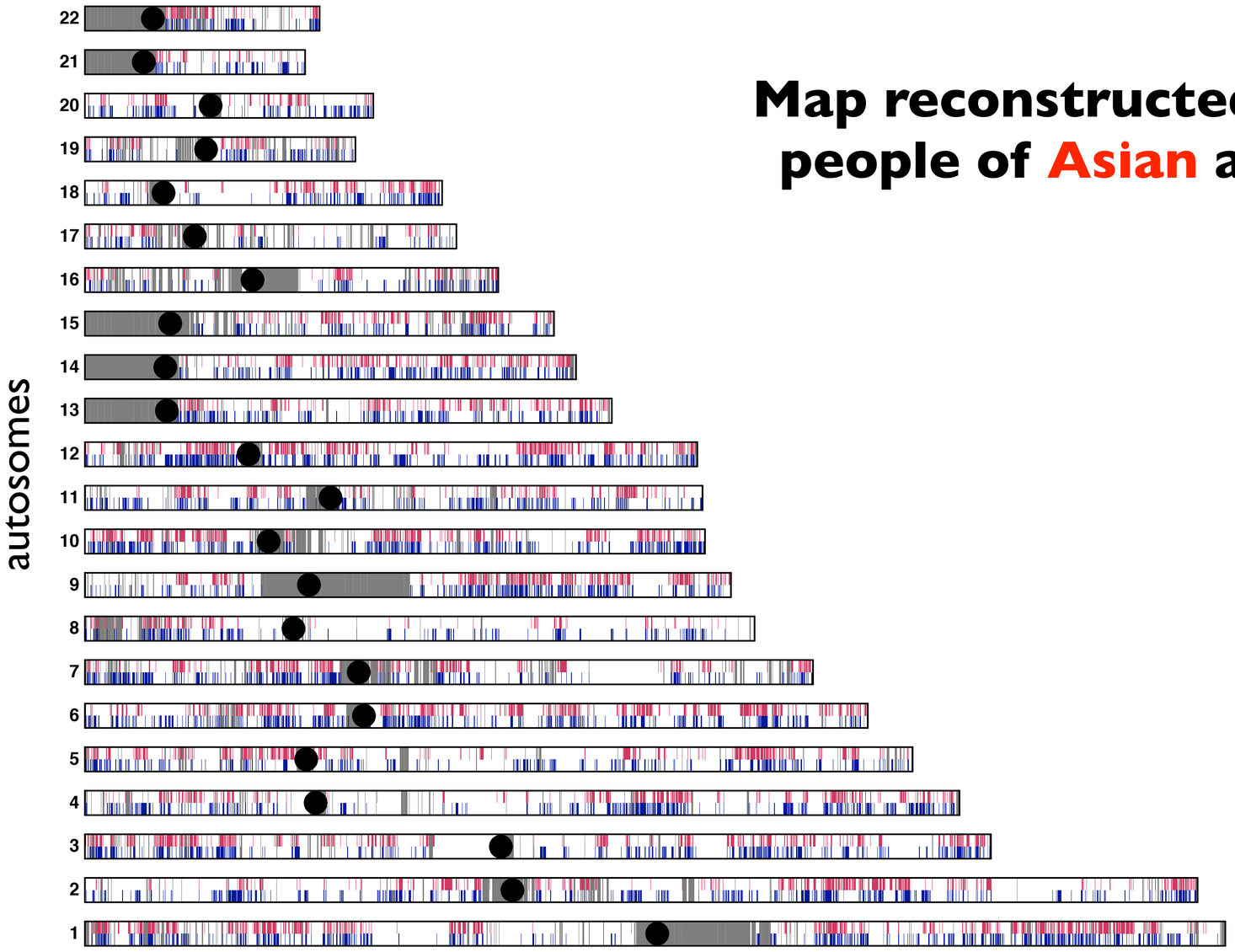
… to continuous segments of Neanderthal DNA



★ = Neanderthal-derived mutation

Racimo et al. (Nature Reviews 2015)

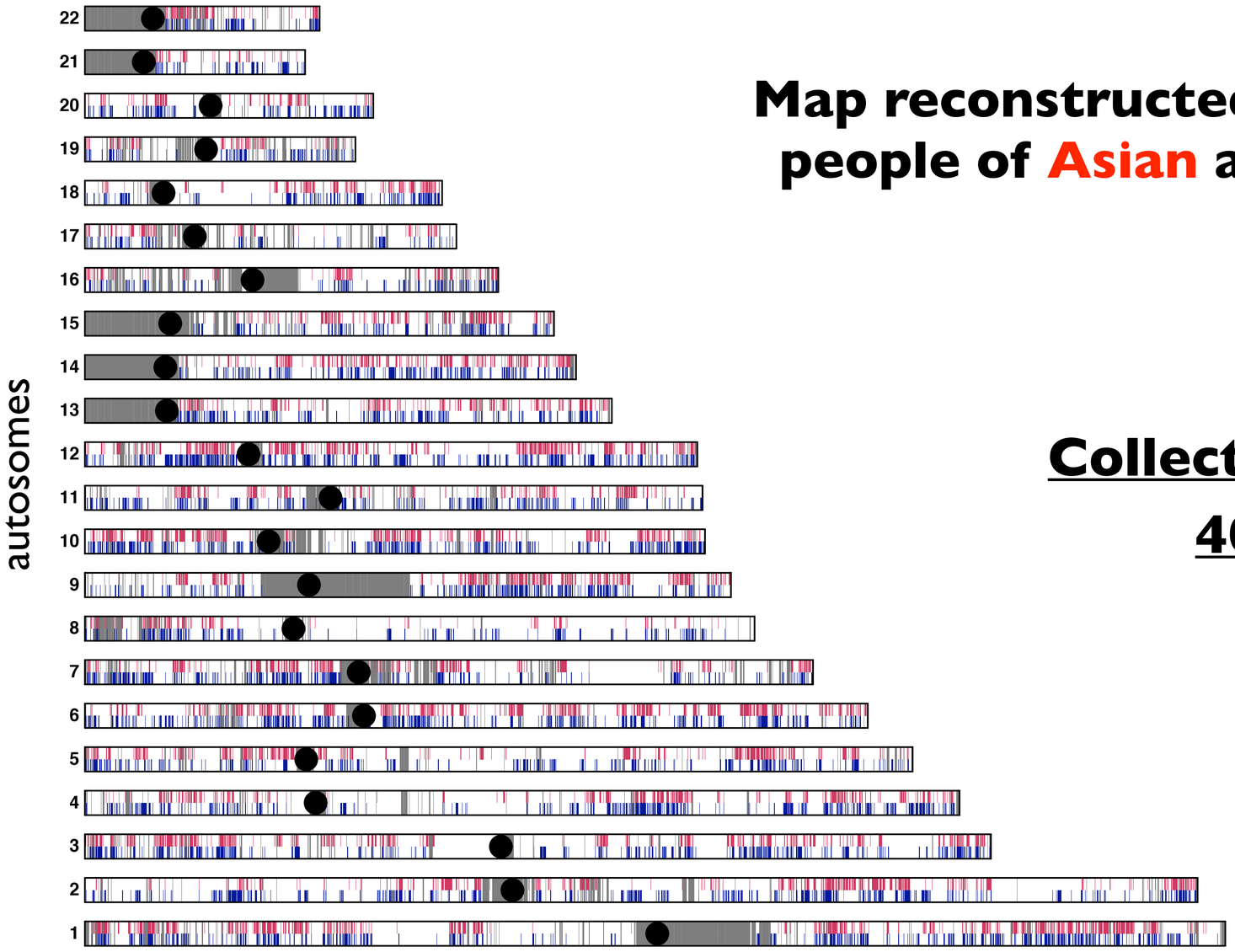
Genomic maps of Neanderthal DNA today



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

Vernot and Akey (Science 2014)

Genomic maps of Neanderthal DNA today

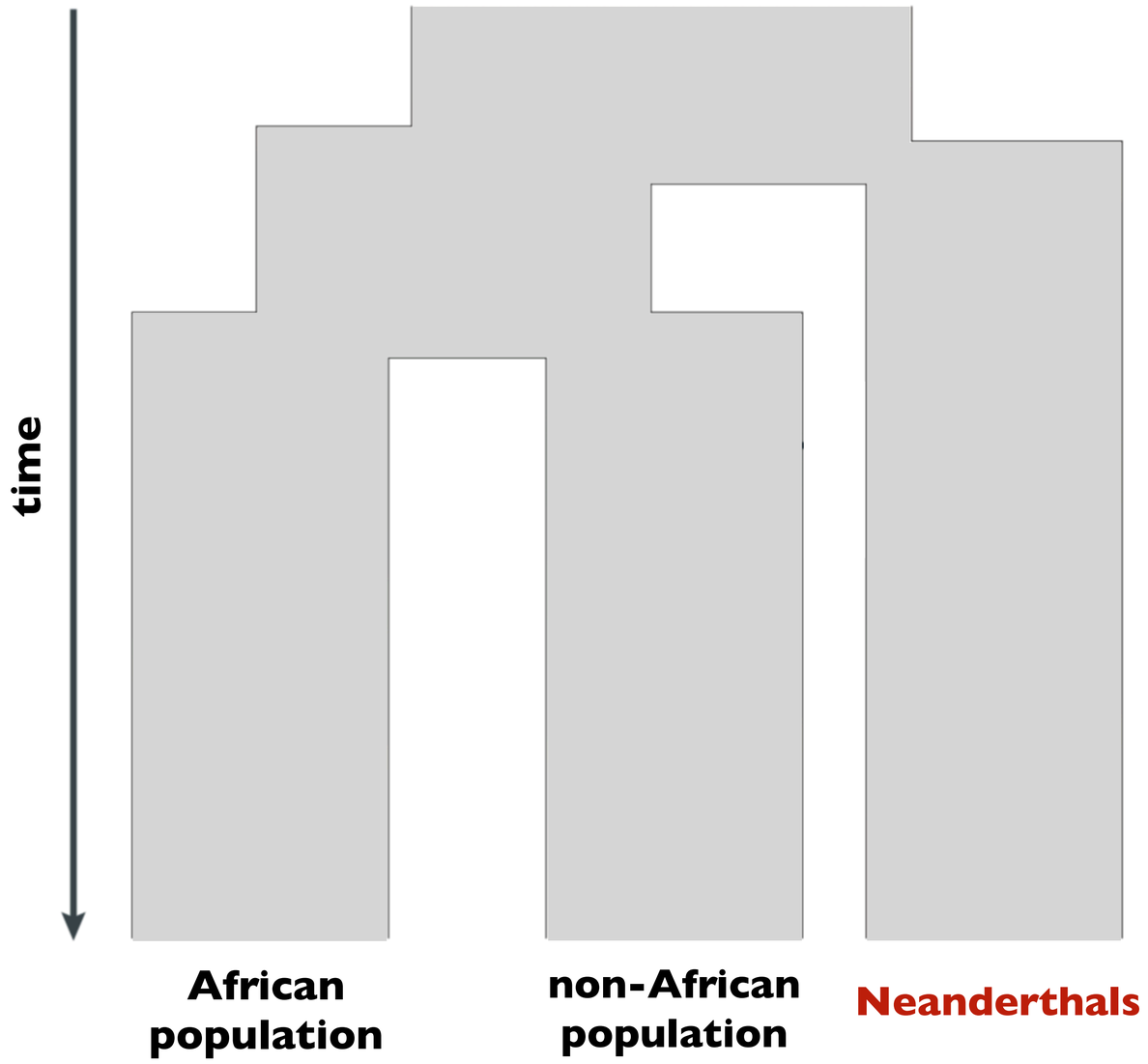


Map reconstructed from genomes of 700 people of Asian and European ancestry.

<u>Collectively we carry around 40% Neanderthal DNA!</u>

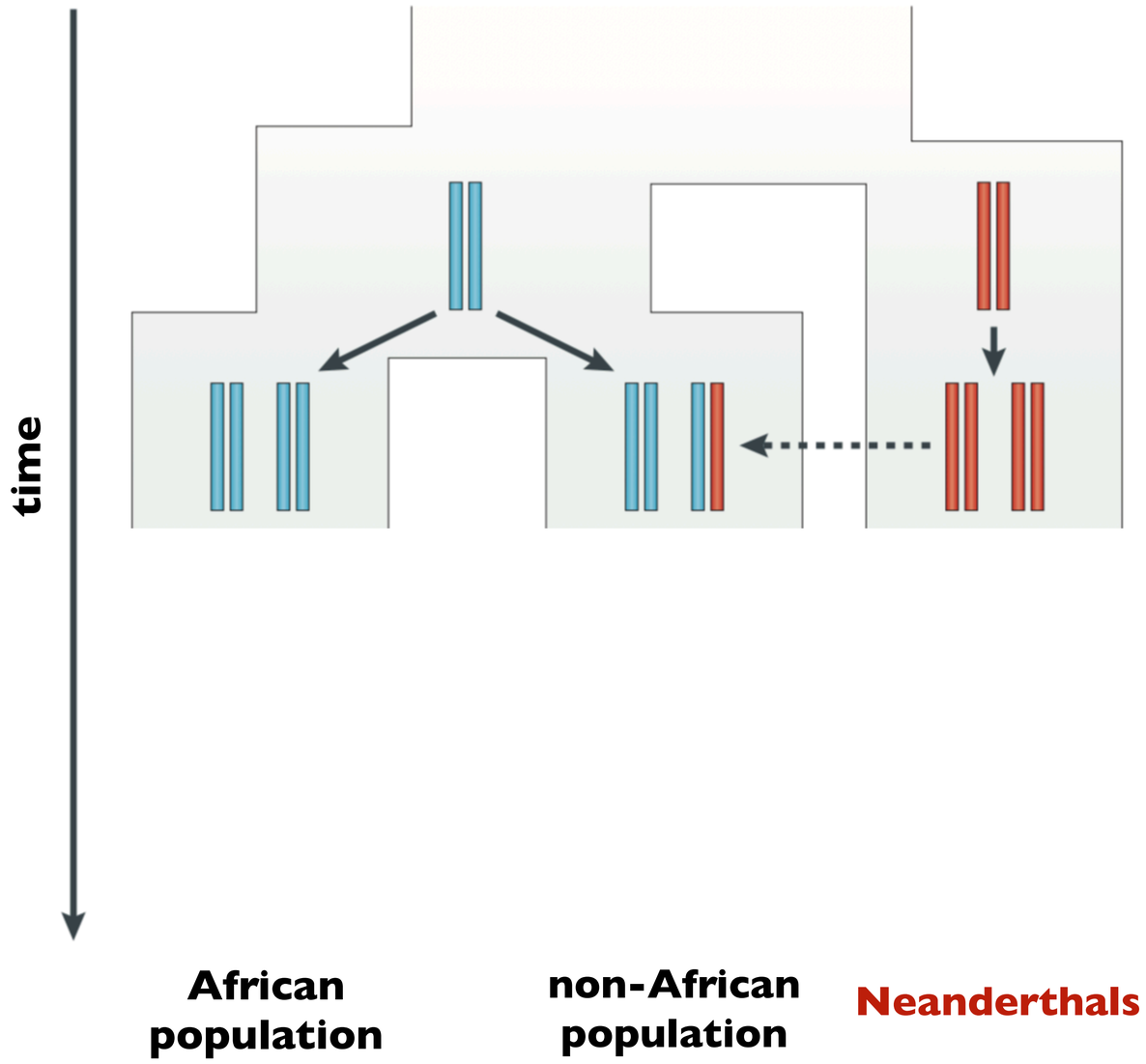
Vernot and Akey (Science 2014)

When did introgression happen?



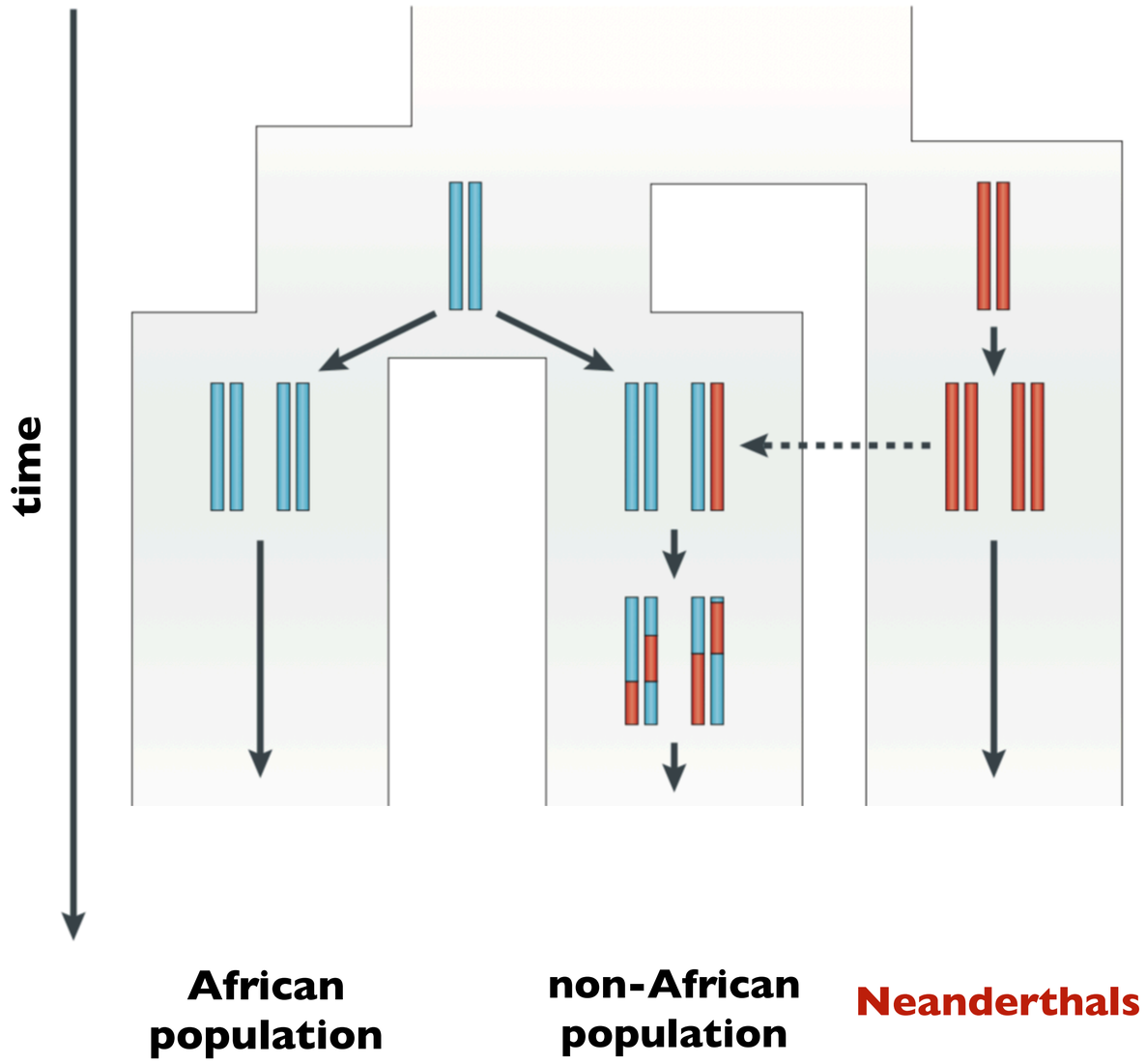
Racimo *et al*. (Nature Reviews 2015)

When did introgression happen?



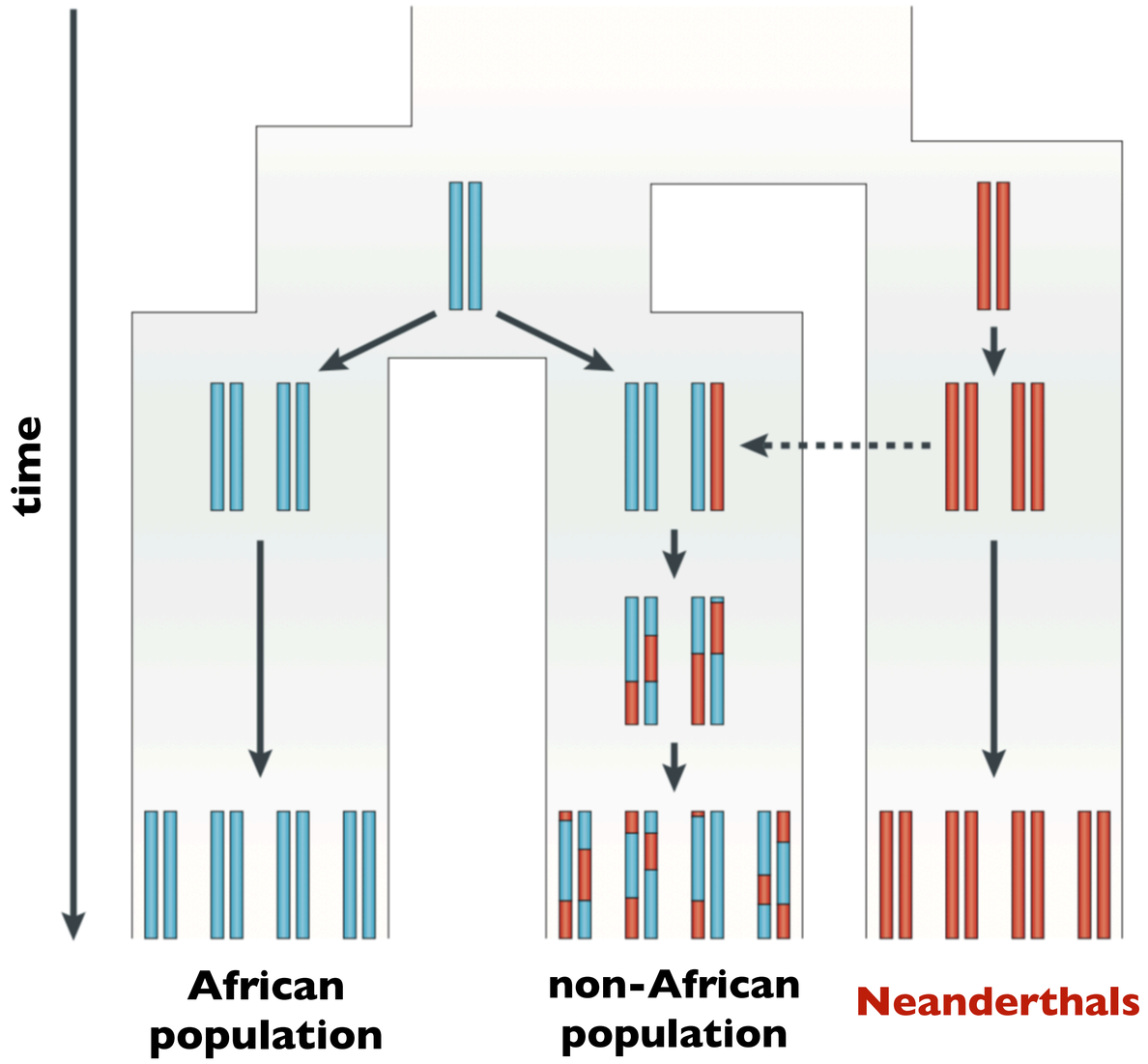
Racimo *et al*. (Nature Reviews 2015)

When did introgression happen?



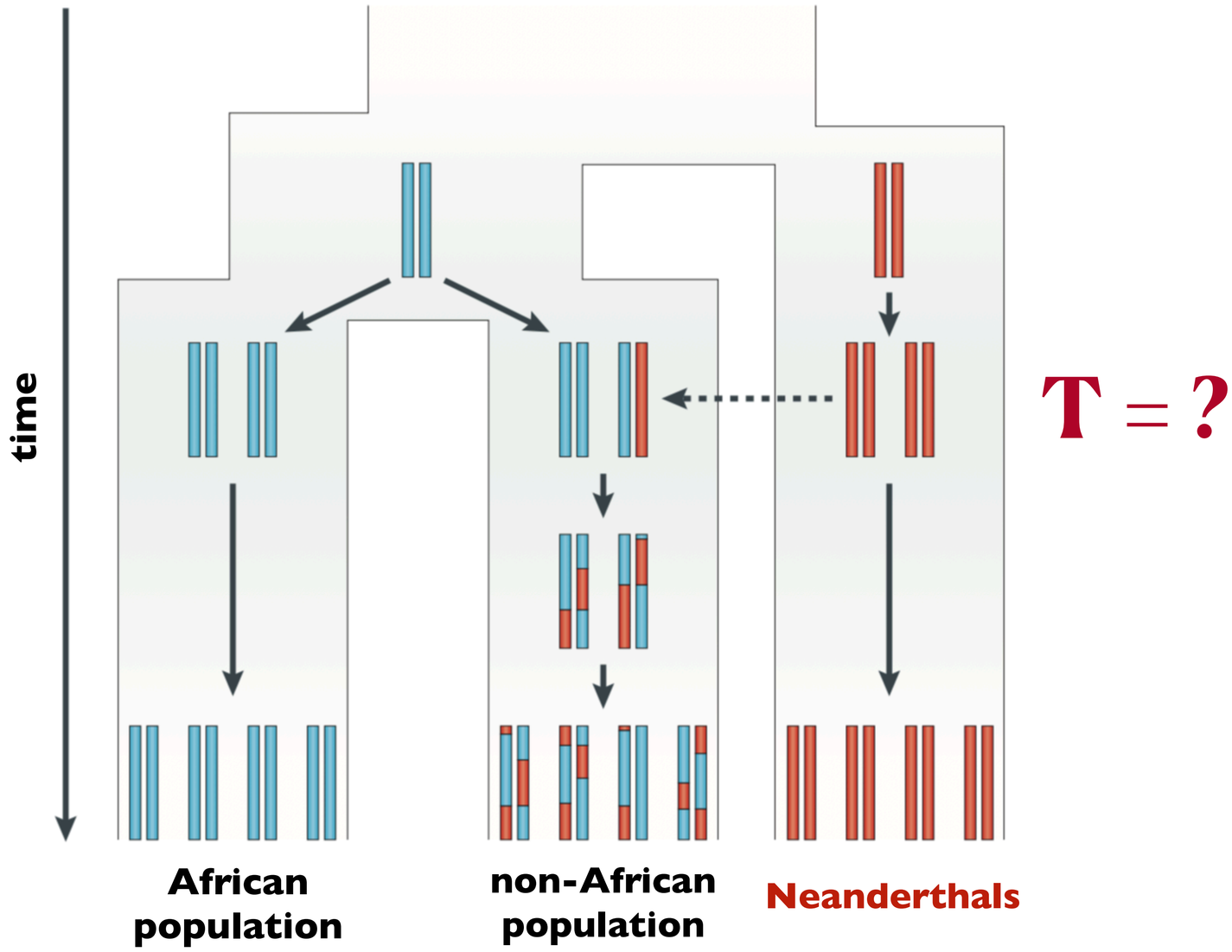
Racimo *et al*. (Nature Reviews 2015)

When did introgression happen?



Racimo *et al*. (Nature Reviews 2015)

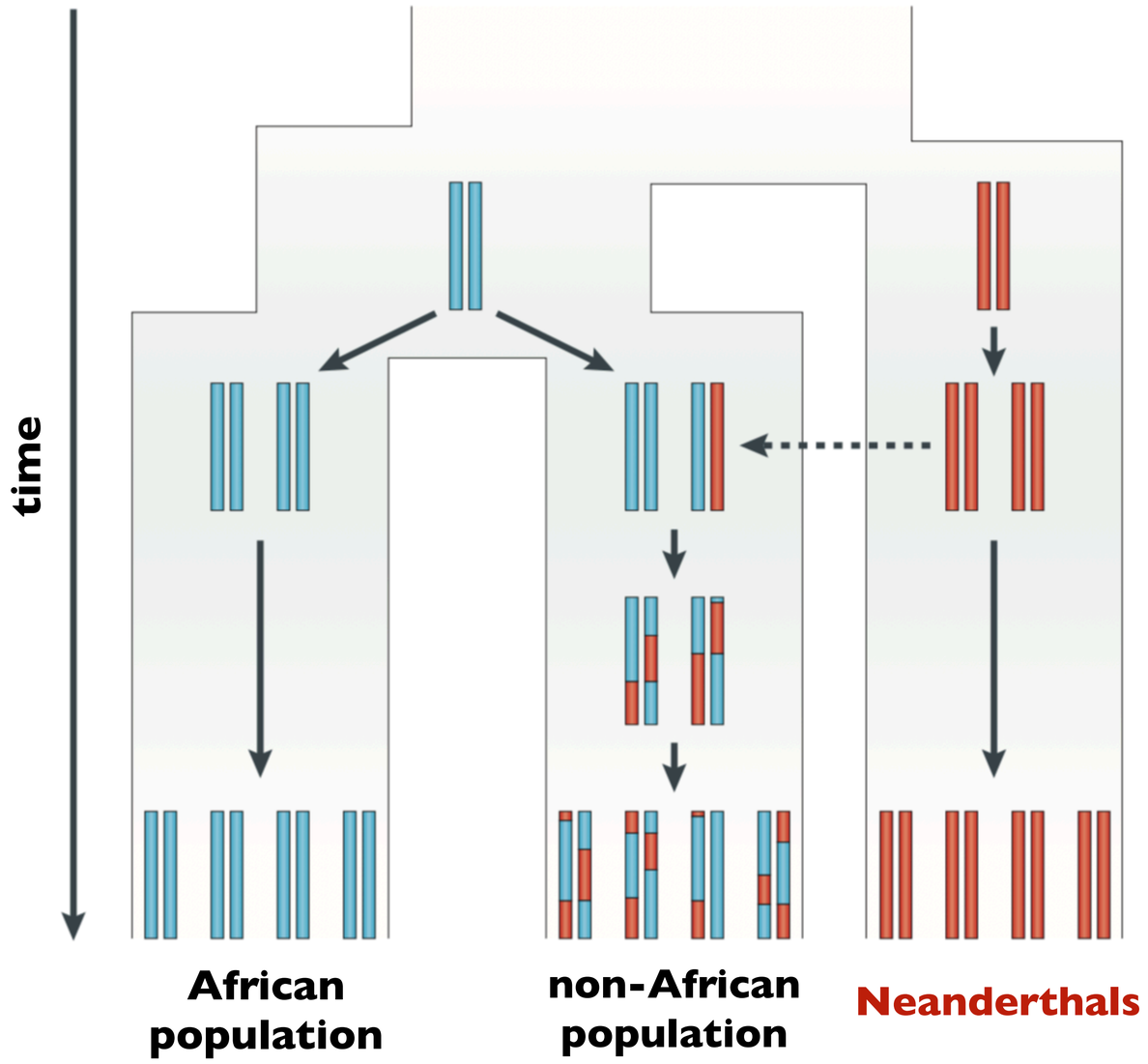
When did introgression happen?



T = ?

Racimo *et al*. (Nature Reviews 2015)

When did introgression happen?

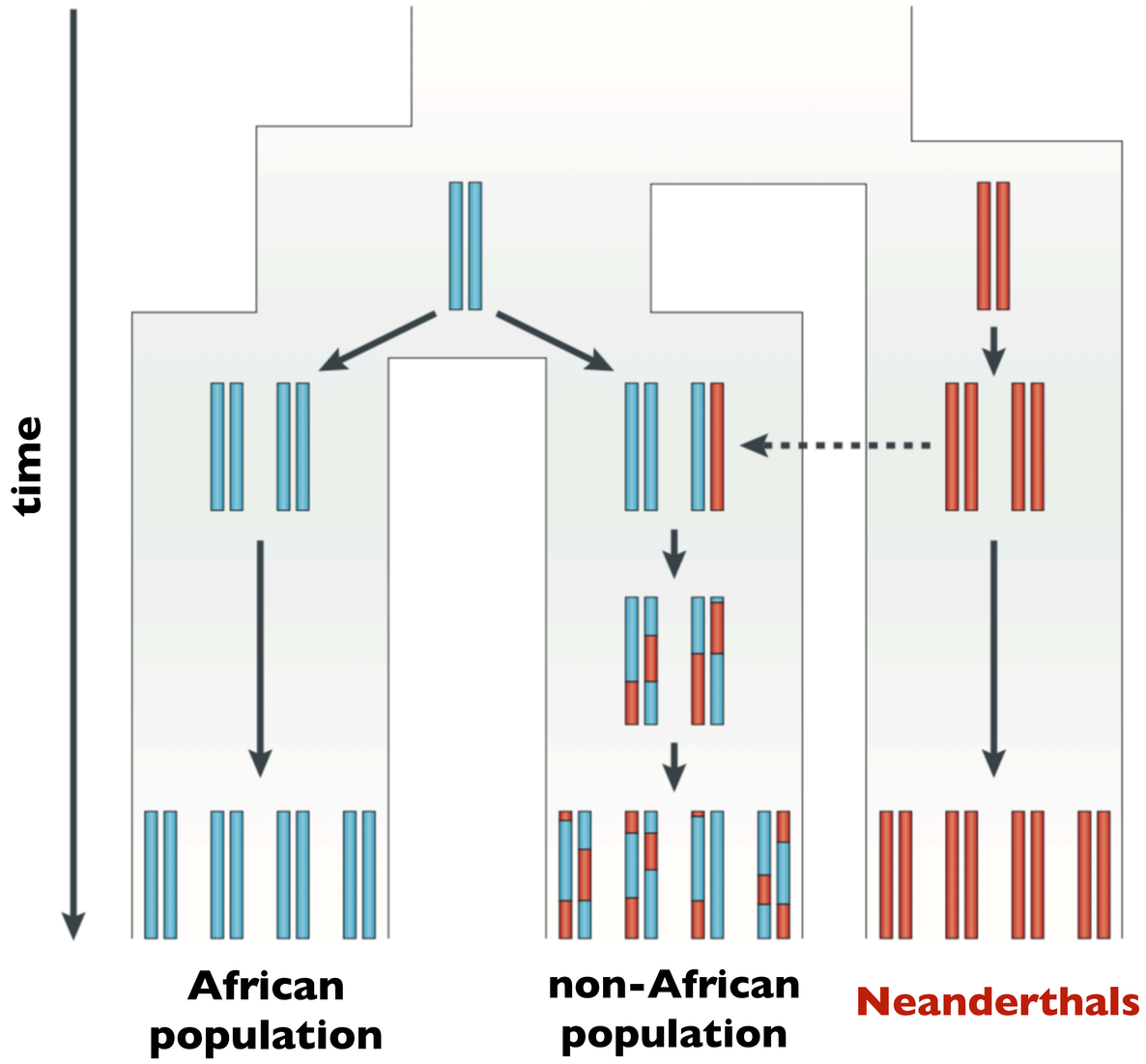


Haplotypes carrying Neanderthal alleles get shorter over time due to recombination (statistically "regularly").

$T = ?$

Racimo *et al*. (Nature Reviews 2015)

When did introgression happen?



$T = ?$

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

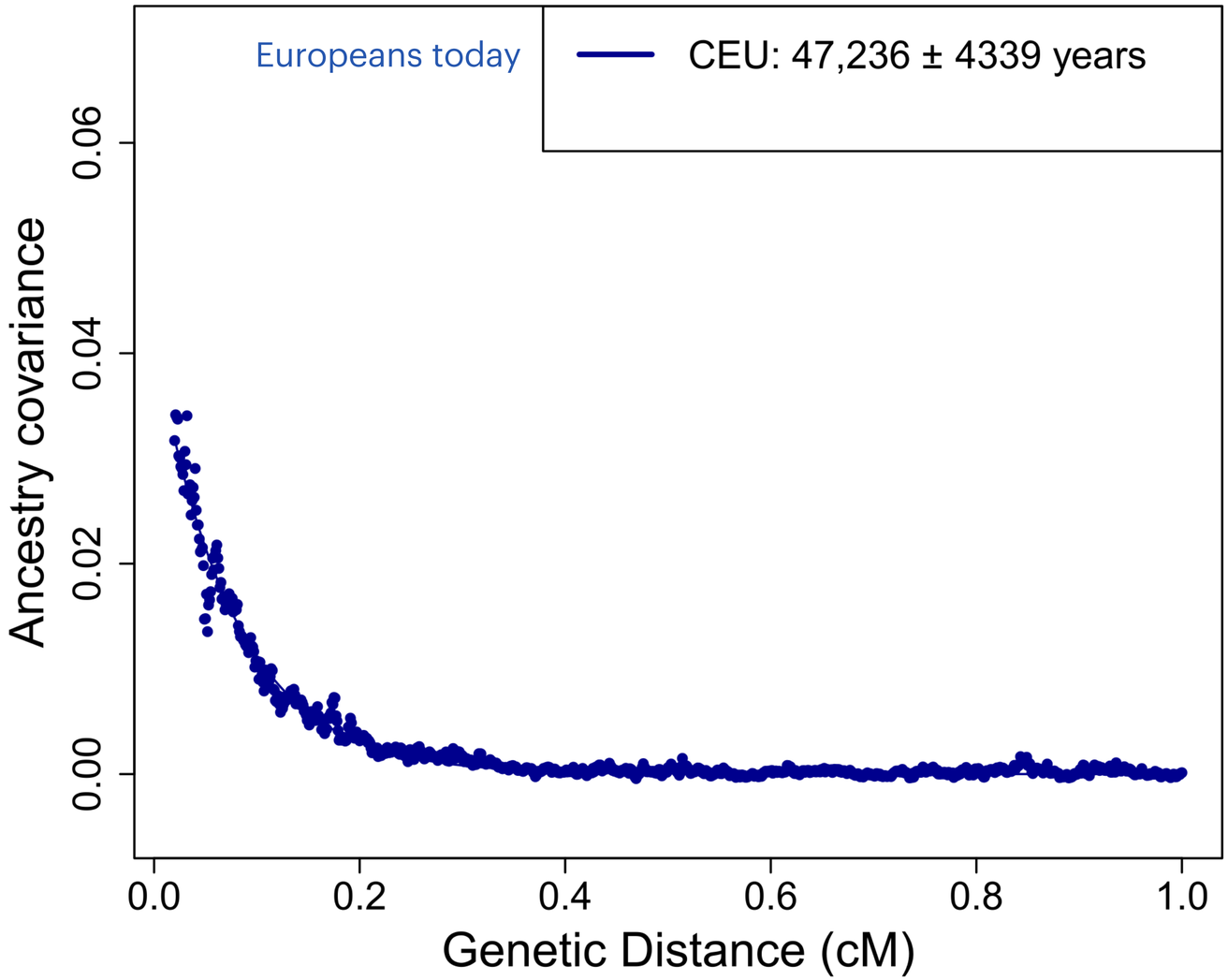
↓

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

Racimo *et al*. (Nature Reviews 2015)

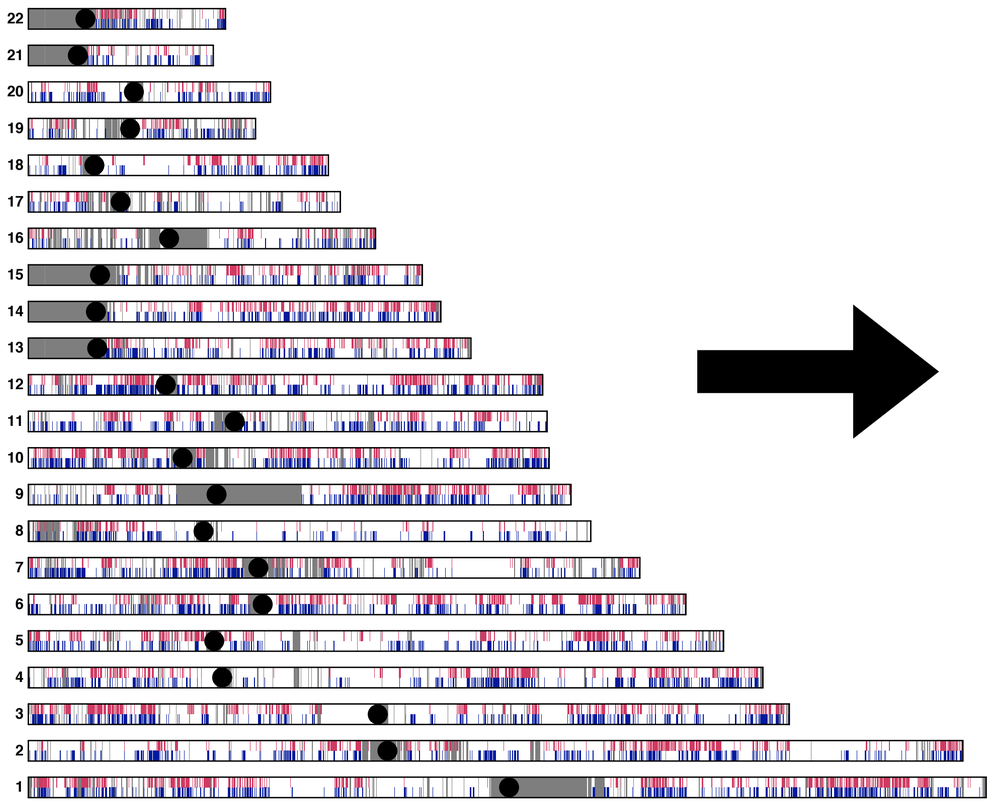
Introgression happened around 50 thousand years ago

map of Neanderthal haplotypes in people today

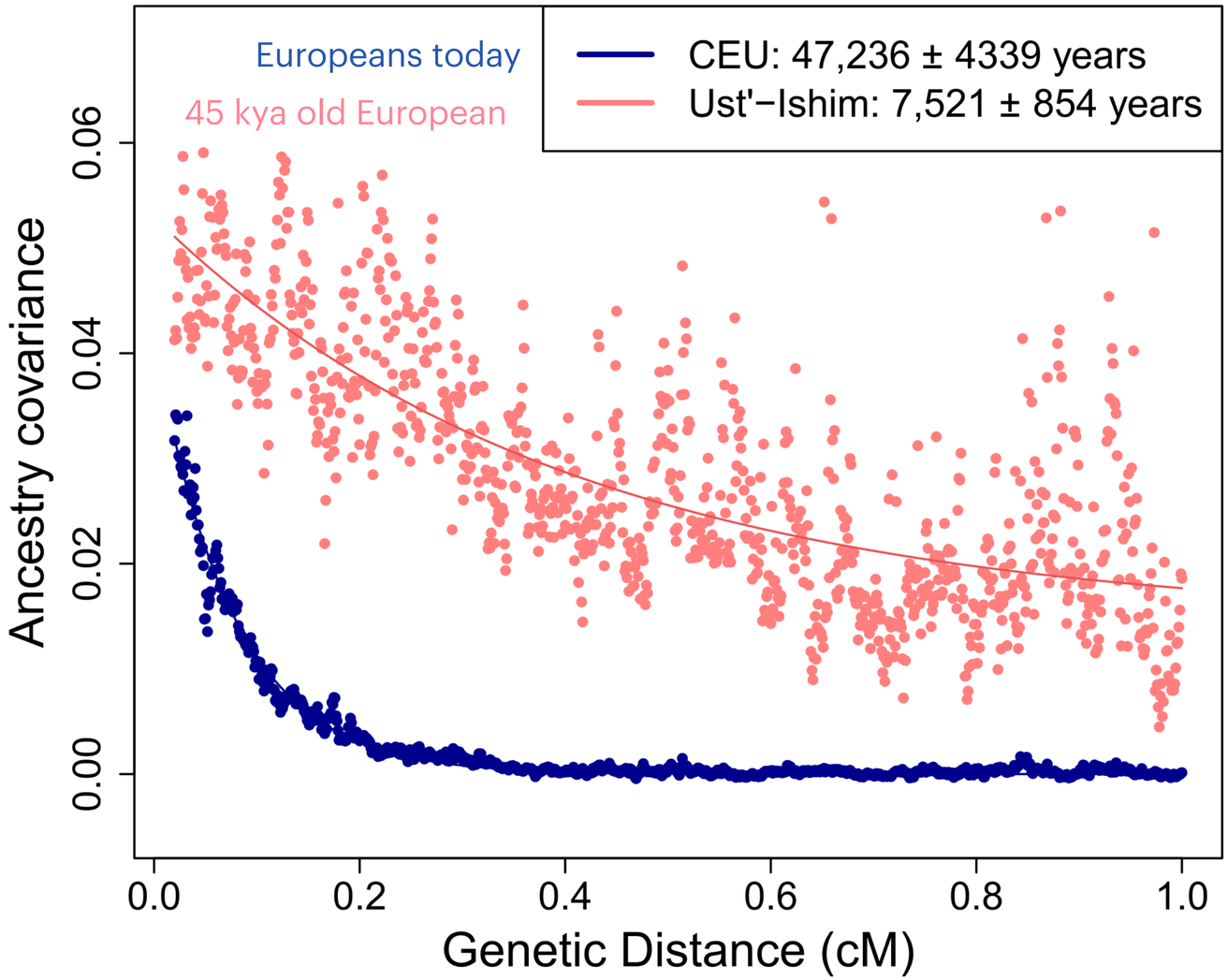


Moorjani *et al*. (PNAS 2016)

Introgression happened around 50 thousand years ago

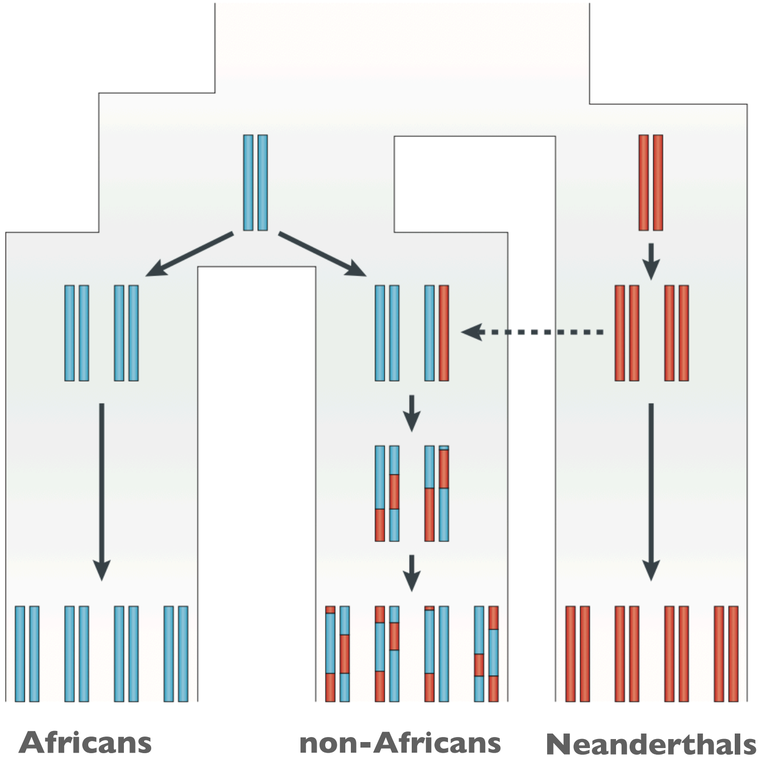


map of Neanderthal haplotypes in people today

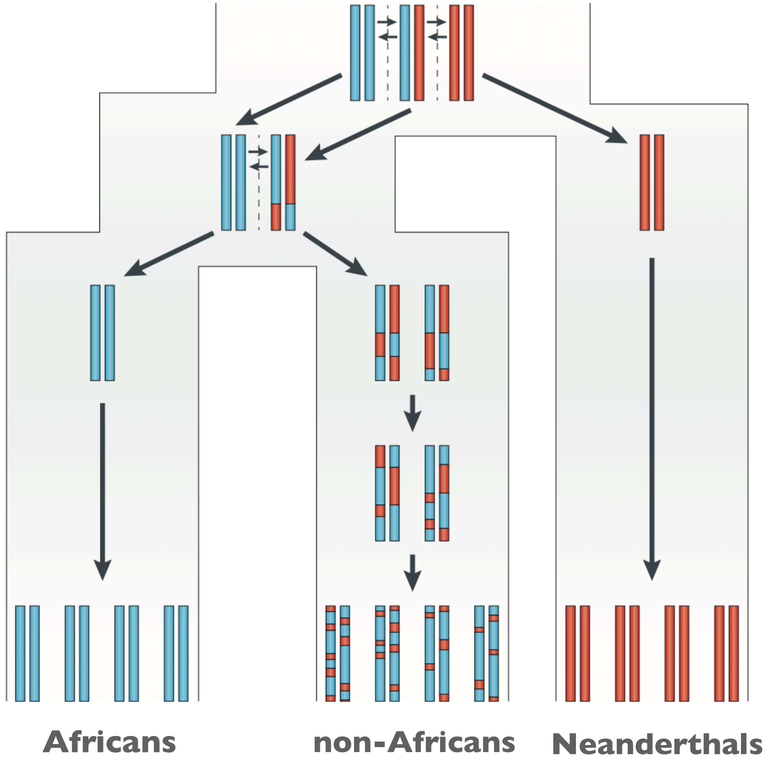


Moorjani *et al*. (PNAS 2016)

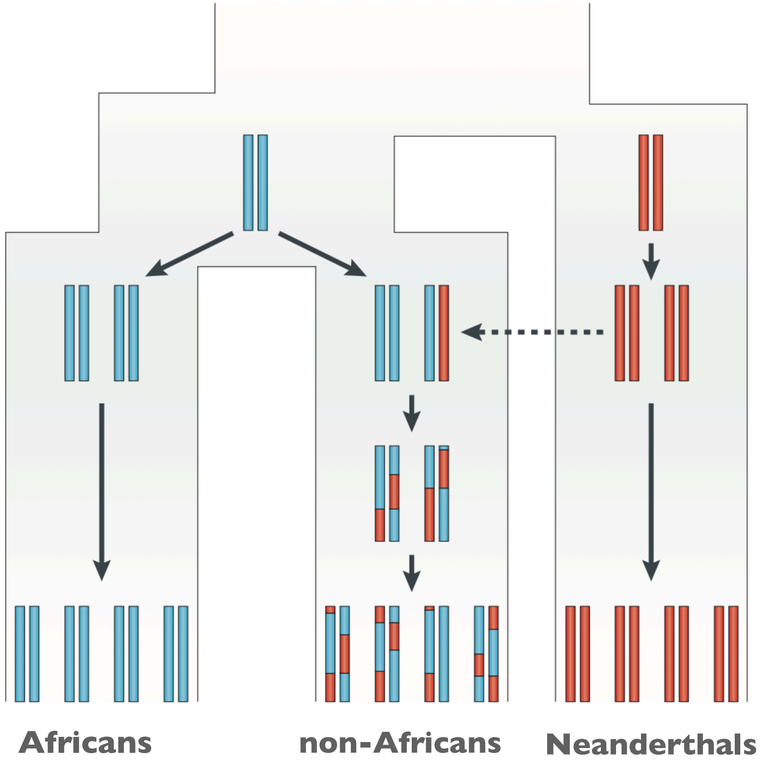
Introgression?



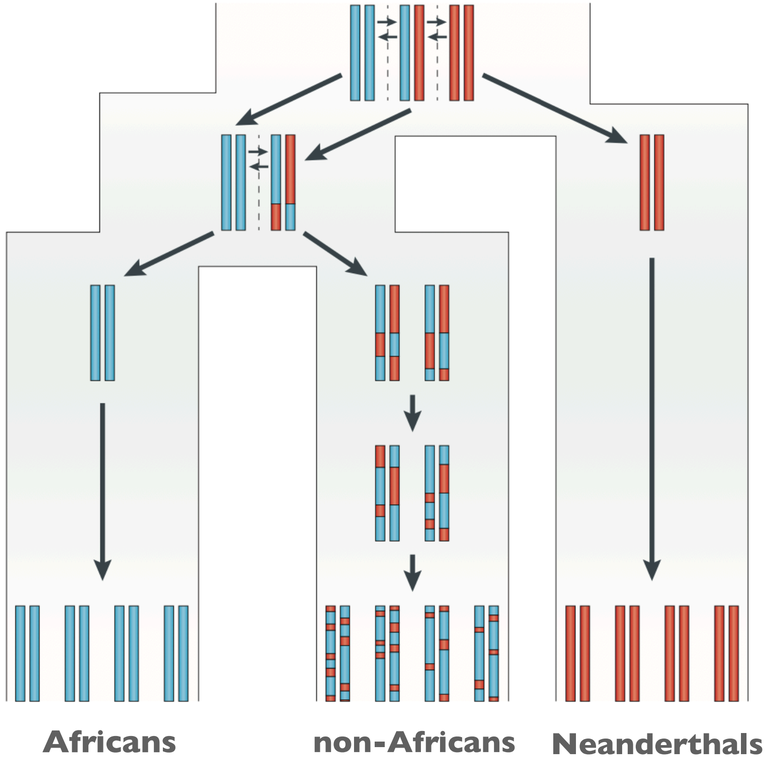
Structure?



Introgression?

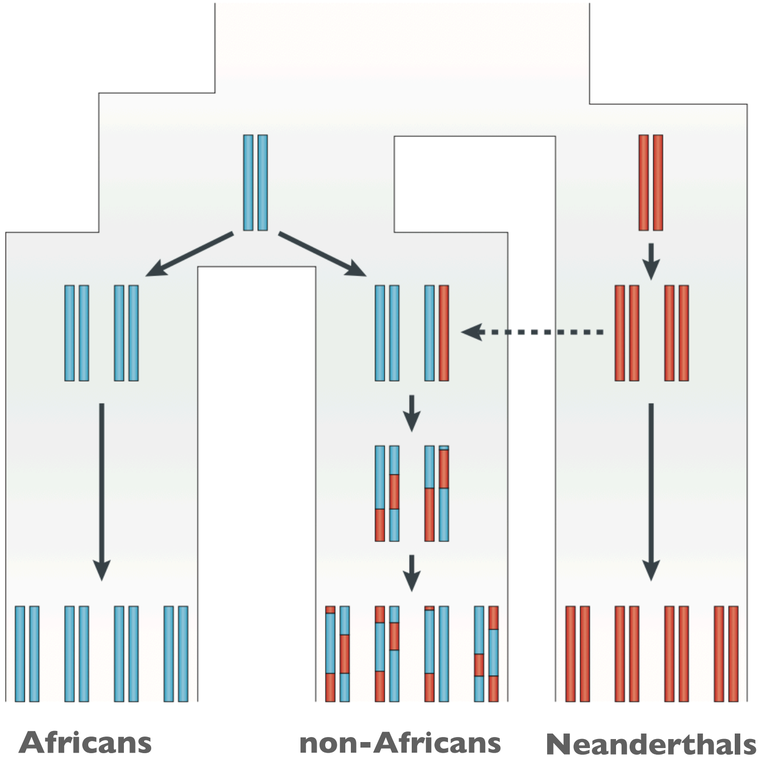


Structure?

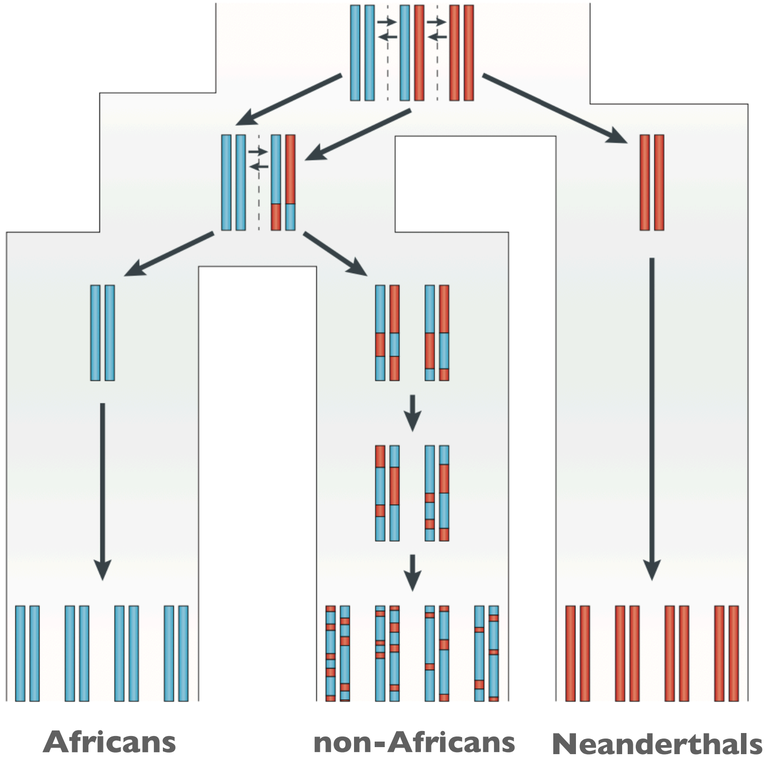


Racimo et al. (Nature Reviews 2015)

Introgression?



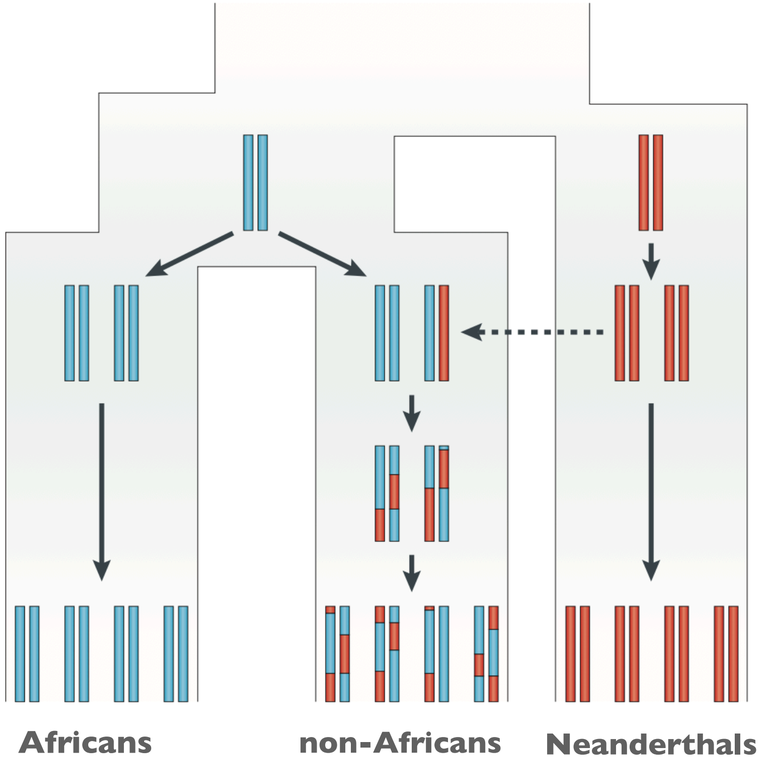
Structure?



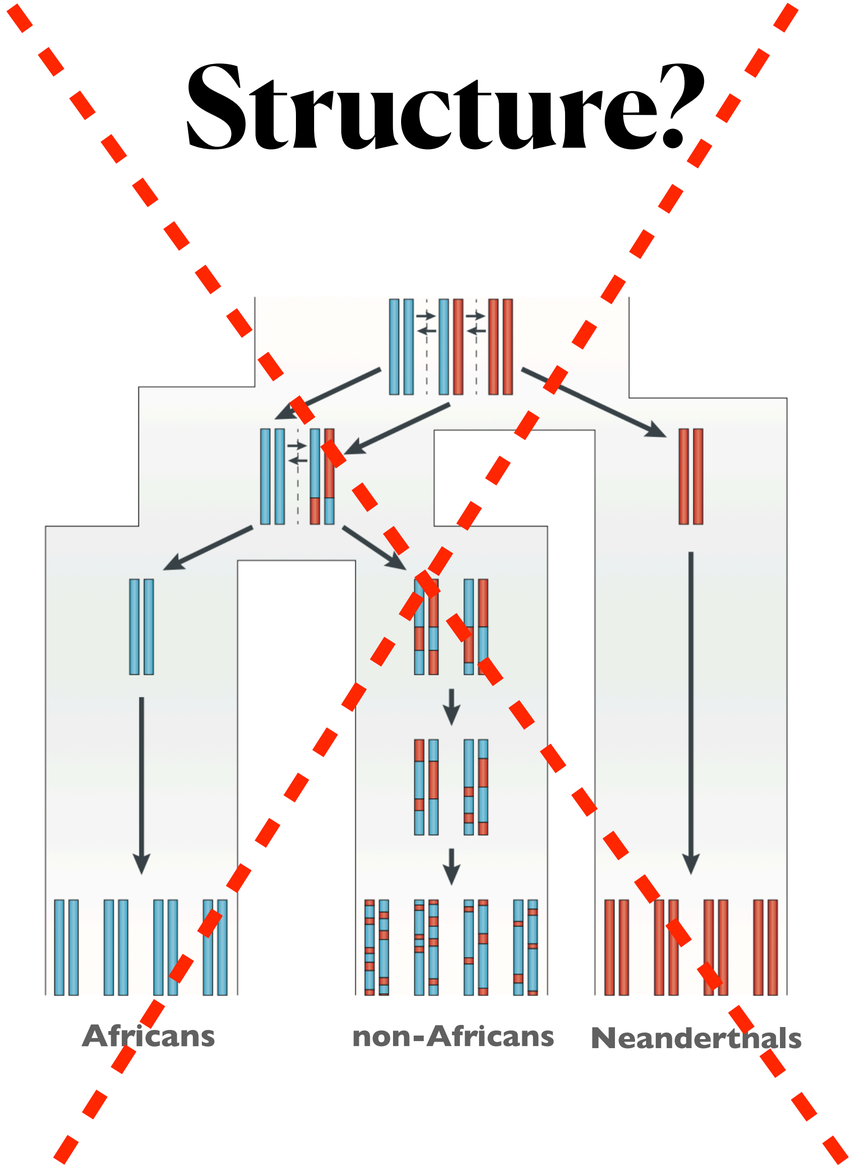
Under structure, "Neanderthal-looking" haplotypes will be much shorter!

Racimo et al. (Nature Reviews 2015)

Introgression?



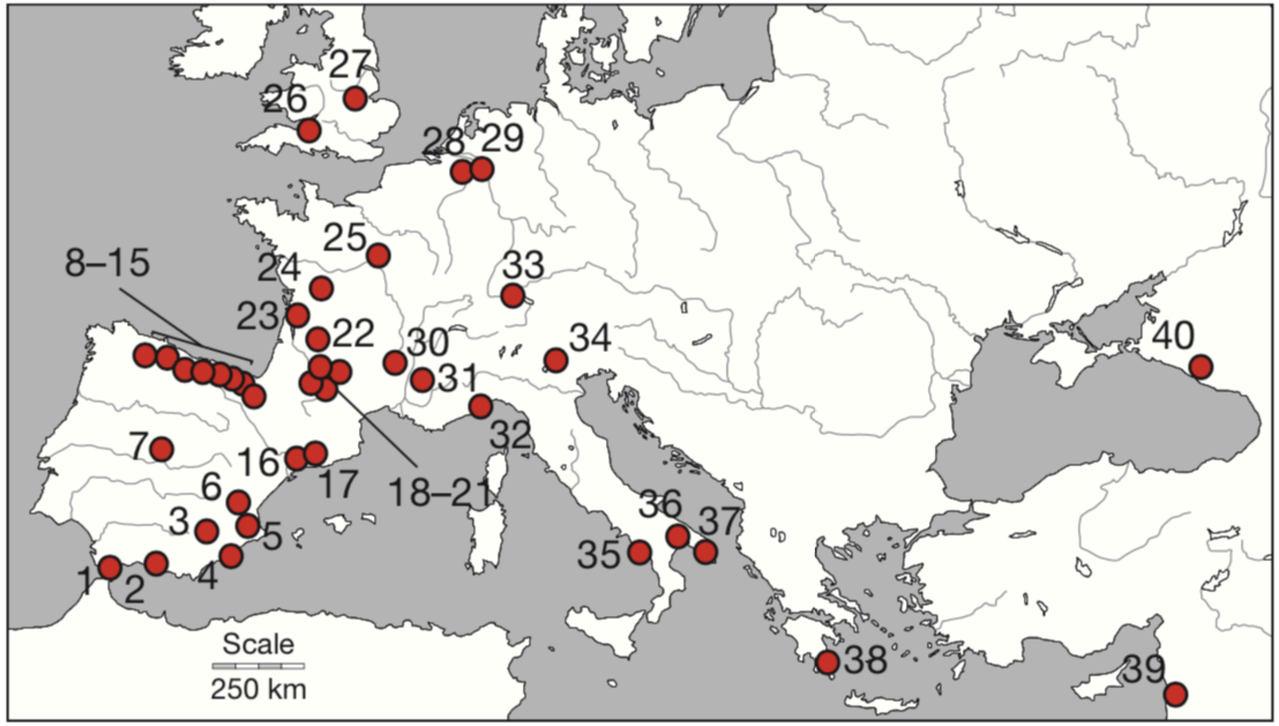
Structure?

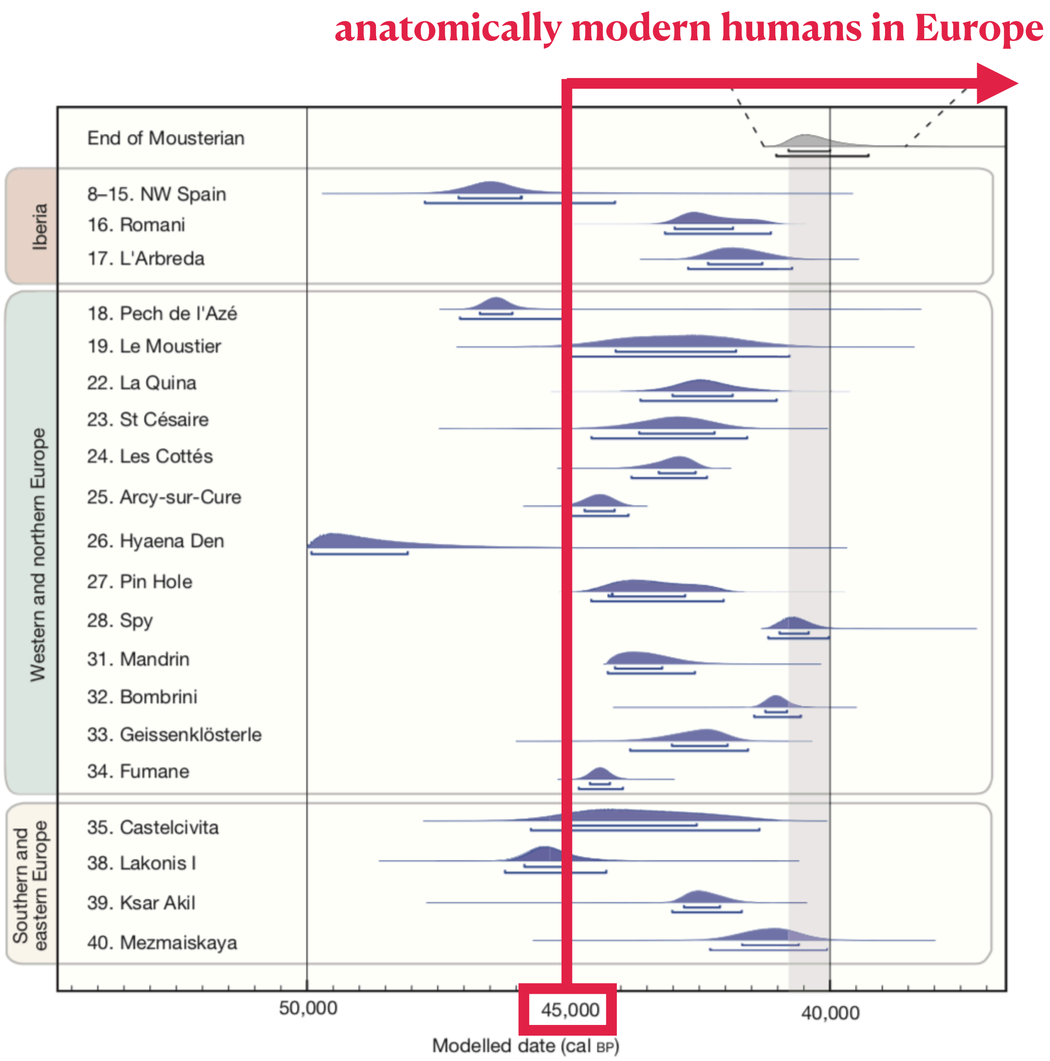


Racimo et al. (Nature Reviews 2015)

Exercises

github.com/bodkan/ku-introgression2024



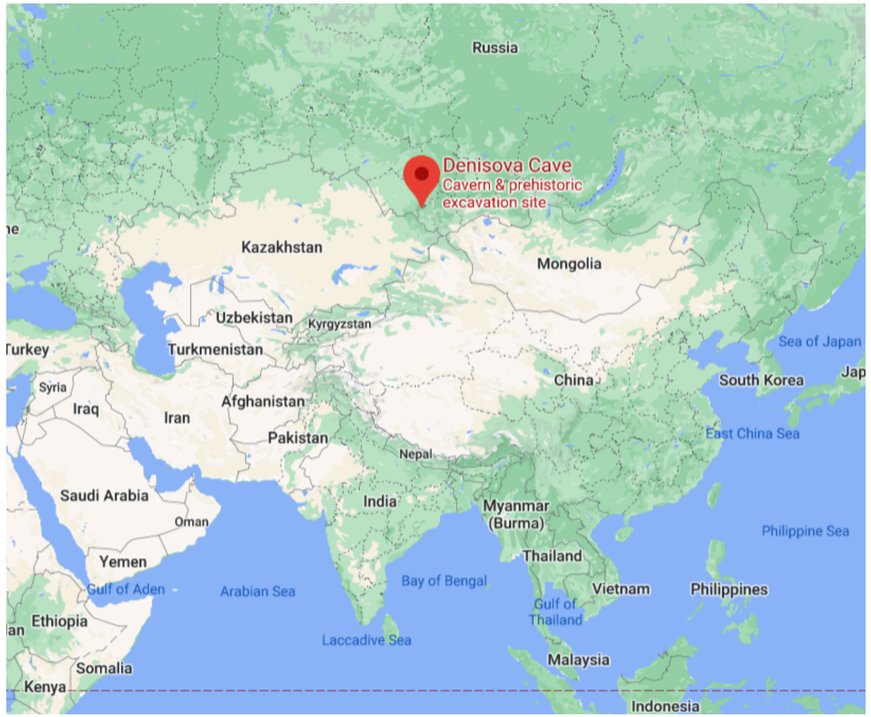


Higham *et al.* (Nature, 2014)

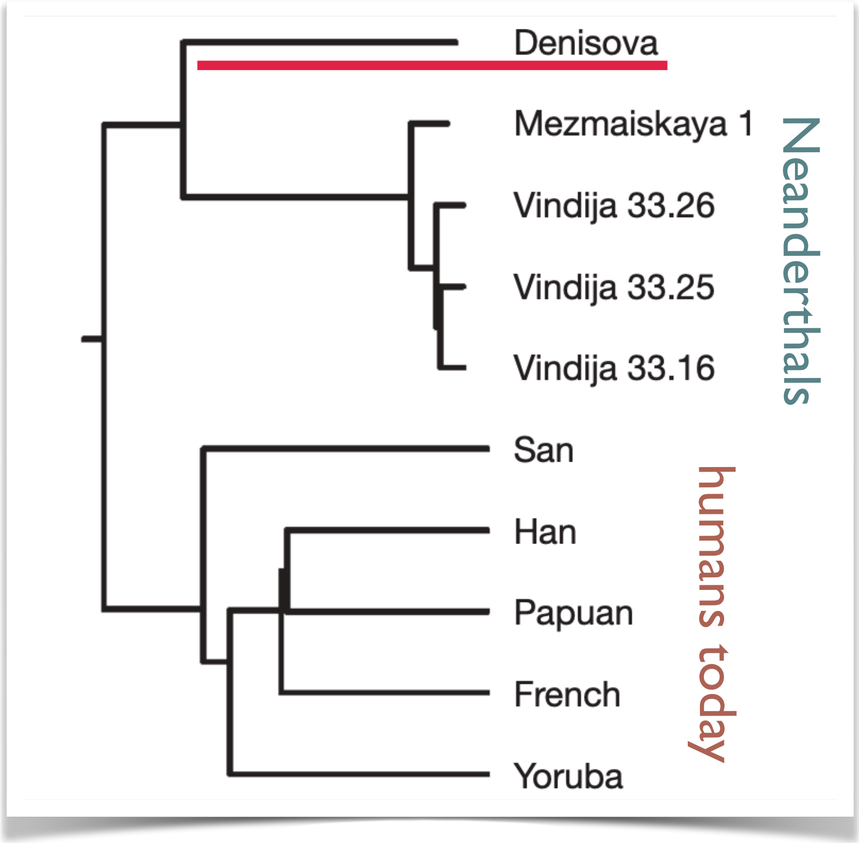
Neanderthals were not the only archaics around at the time

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



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

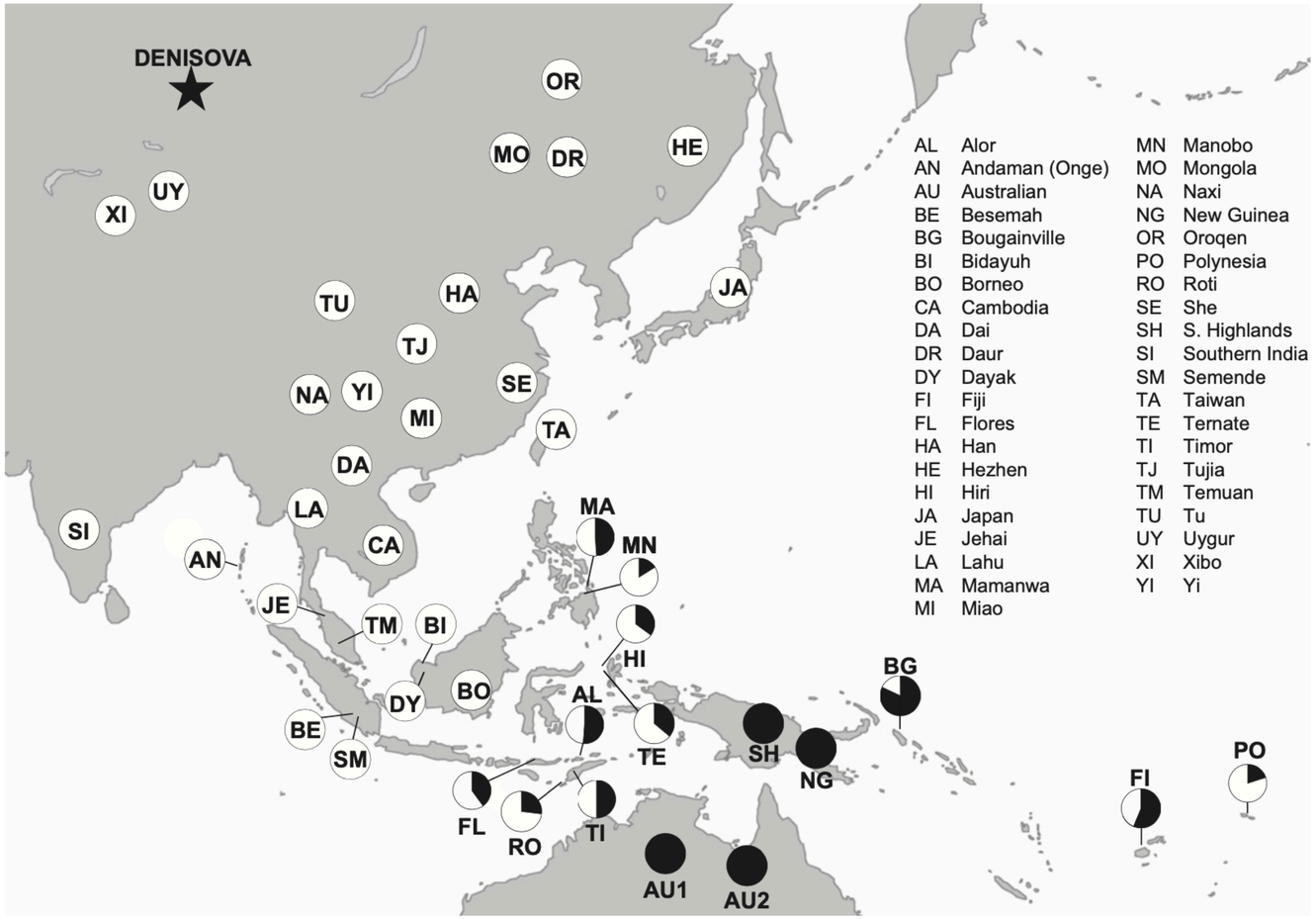
Genetic history of an archaic hominin group from Denisova Cave in Siberia

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(human 1, human 2; **Denisovan**, Chimp)



Reich *et al.*, 2011

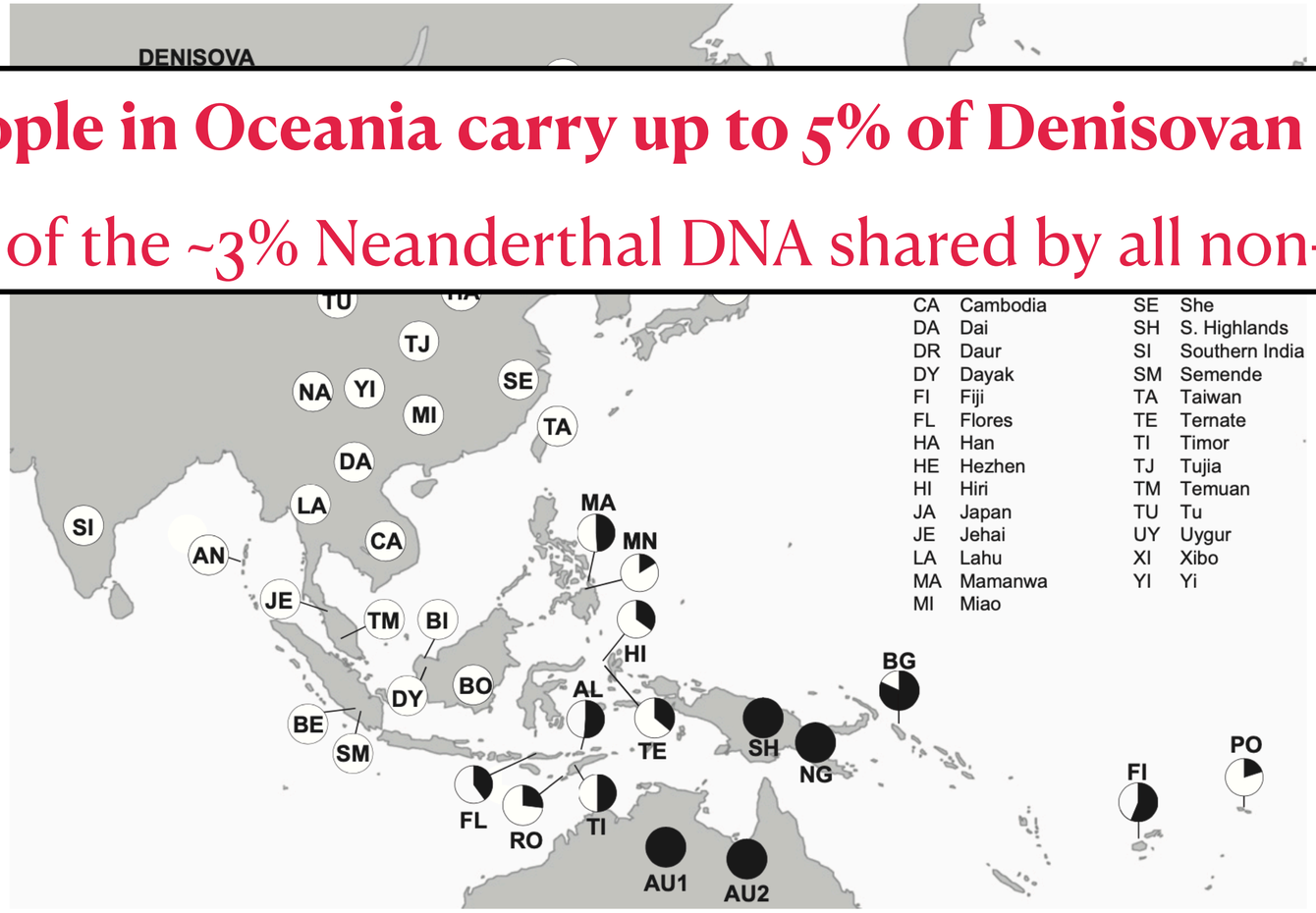
Figure 1. Denisovan Genetic Material as a Fraction of that in New Guineans

Traces of Denisovans in genomes of present-day humans

f4(human 1, human 2; **Denisovan**, Chimp)

People in Oceania carry up to 5% of Denisovan DNA!

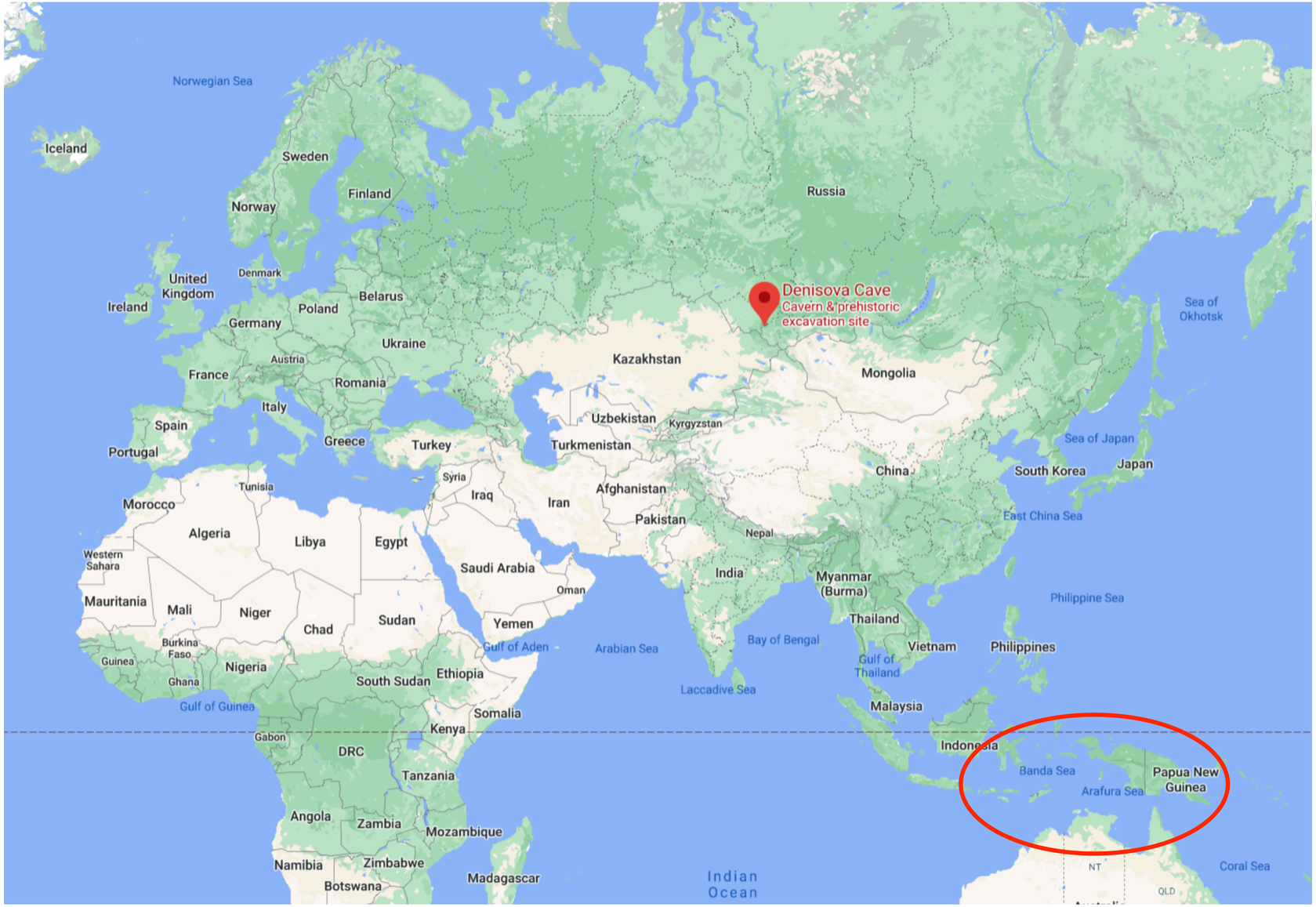
(on top of the ~3% Neanderthal DNA shared by all non-Africans!)



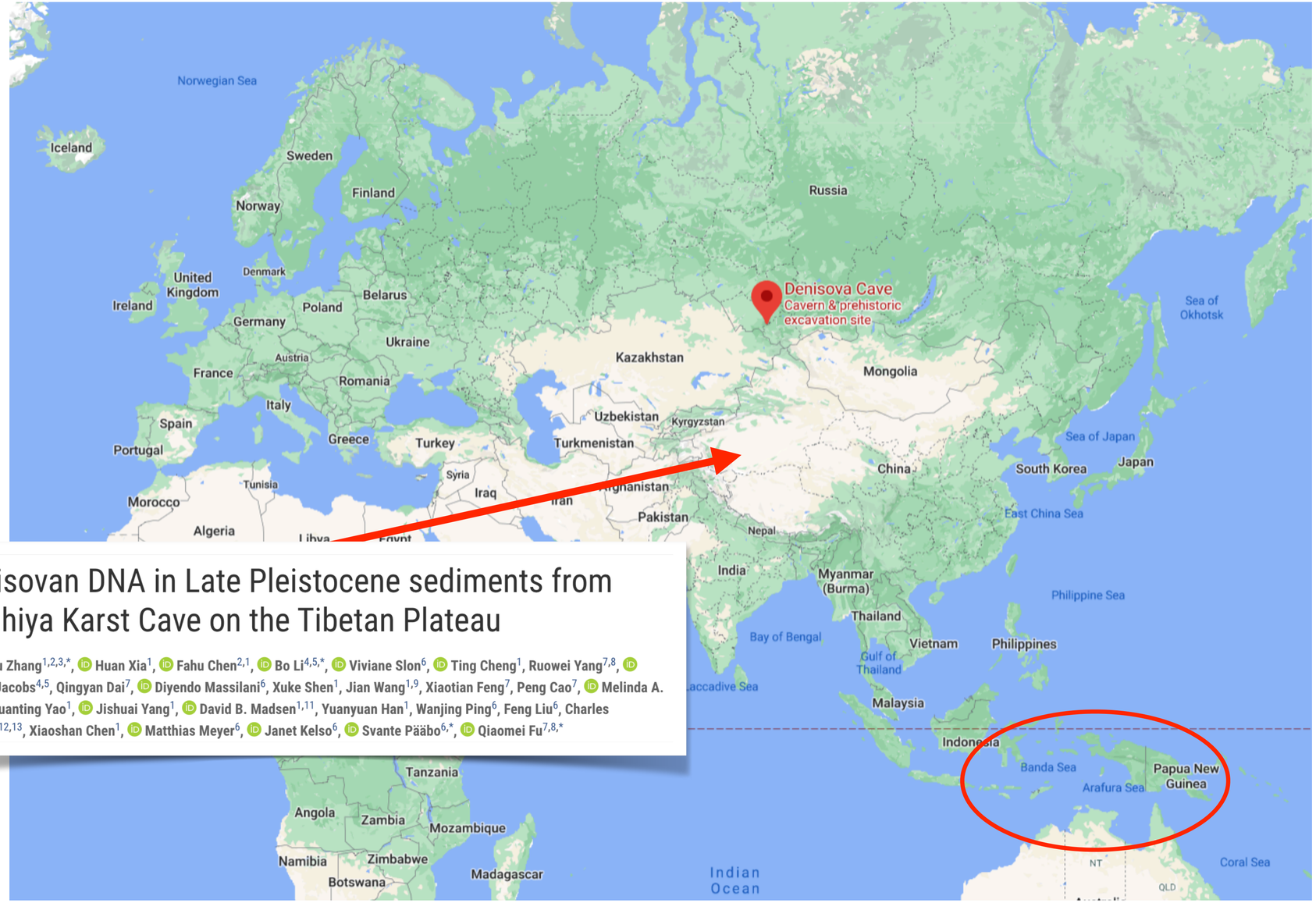
Reich *et al.*, 2011

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

Geographic range of Denisovans?

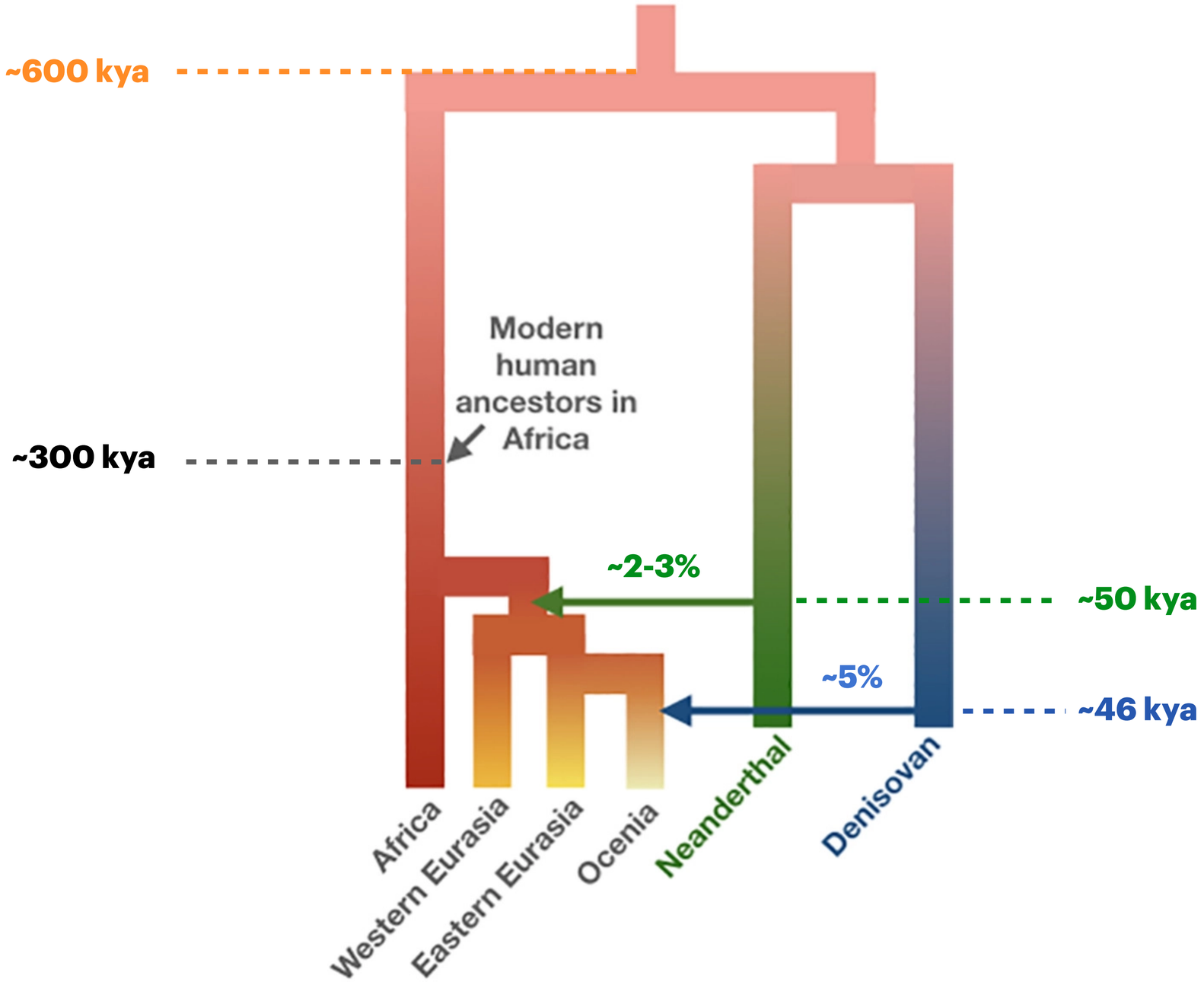


Geographic range of Denisovans?



Denisovan DNA in Late Pleistocene sediments from Baishiya Karst Cave on the Tibetan Plateau

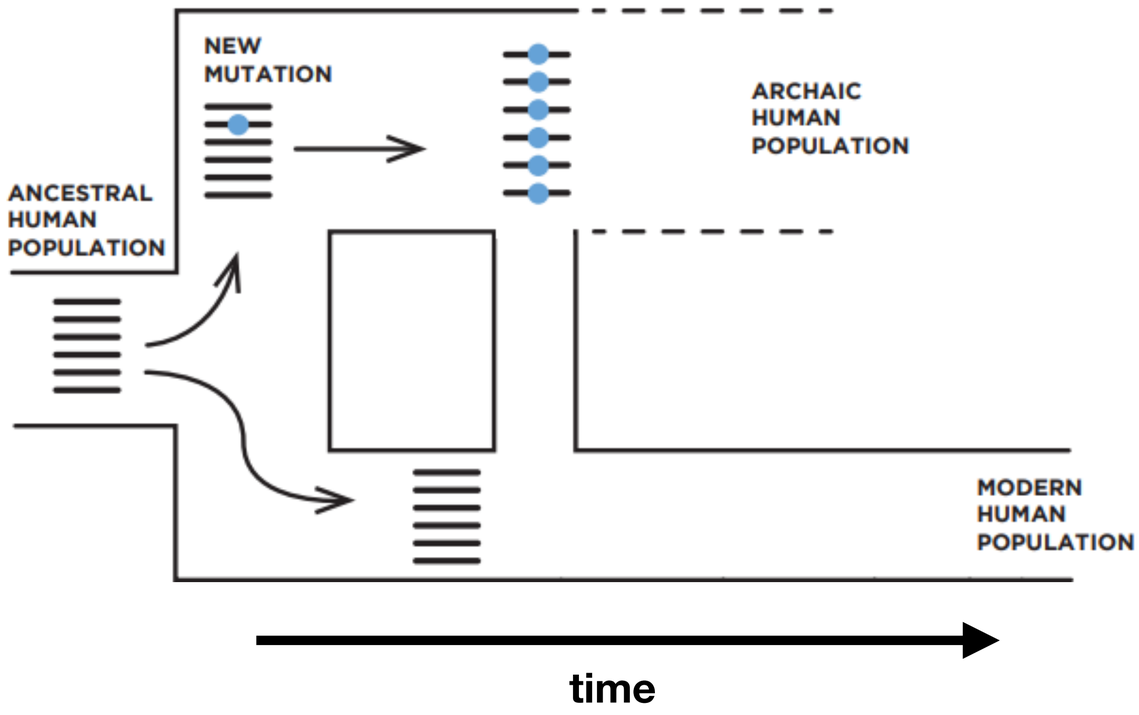
Dongju Zhang[1,2,3,*], Huan Xia[1], Fahu Chen[2,1], Bo Li[4,5,*], Viviane Slon[6], Ting Cheng[1], Ruowei Yang[7,8], Zenobia Jacobs[4,5], Qingyan Dai[7], Diyendo Massilani[6], Xuke Shen[1], Jian Wang[1,9], Xiaotian Feng[7], Peng Cao[7], Melinda A. Yang[10], Juanting Yao[1], Jishuai Yang[1], David B. Madsen[1,11], Yuanyuan Han[1], Wanjing Ping[6], Feng Liu[6], Charles Perreault[12,13], Xiaoshan Chen[1], Matthias Meyer[6], Janet Kelso[6], Svante Pääbo[6,*], Qiaomei Fu[7,8,*]



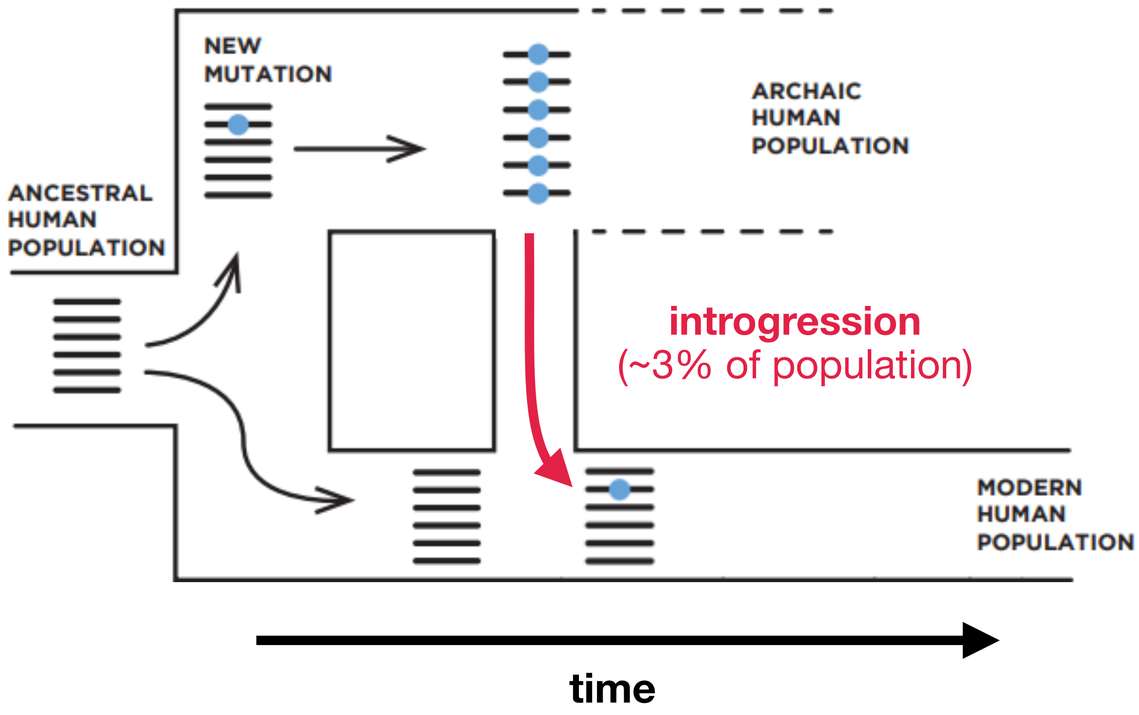
adapted from the Yearbook of Physical Anthropology (2019) by Gokcumen

Consequences of introgression on human functional biology?

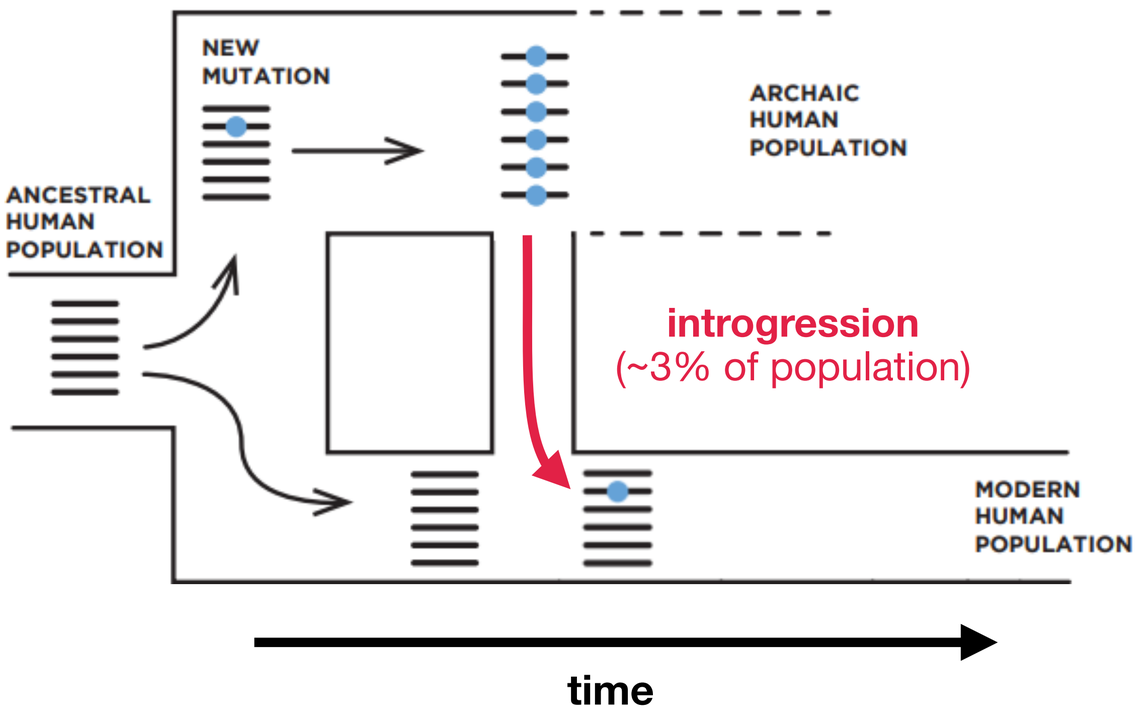
What are the possible fates of introgressed DNA?

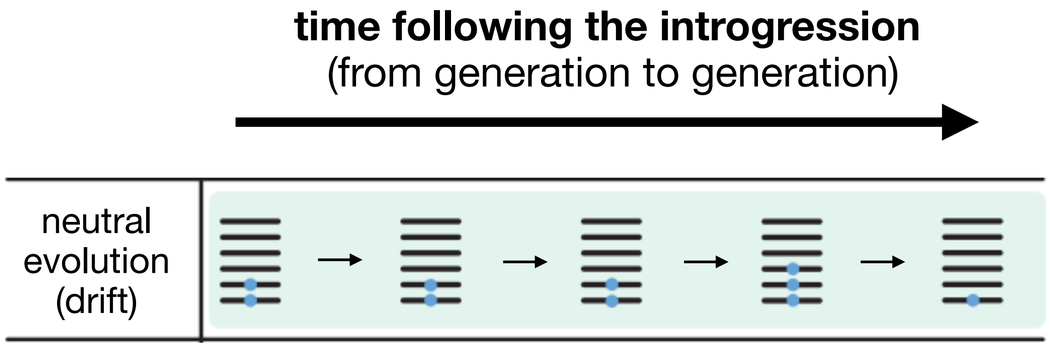


What are the possible fates of introgressed DNA?

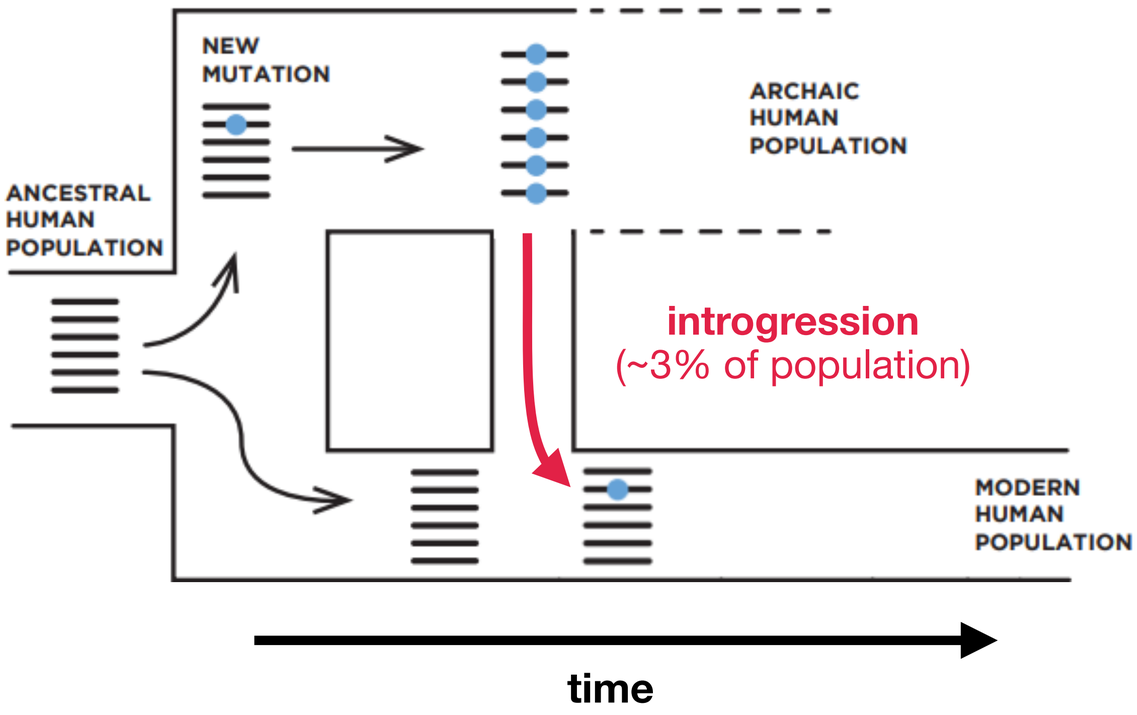


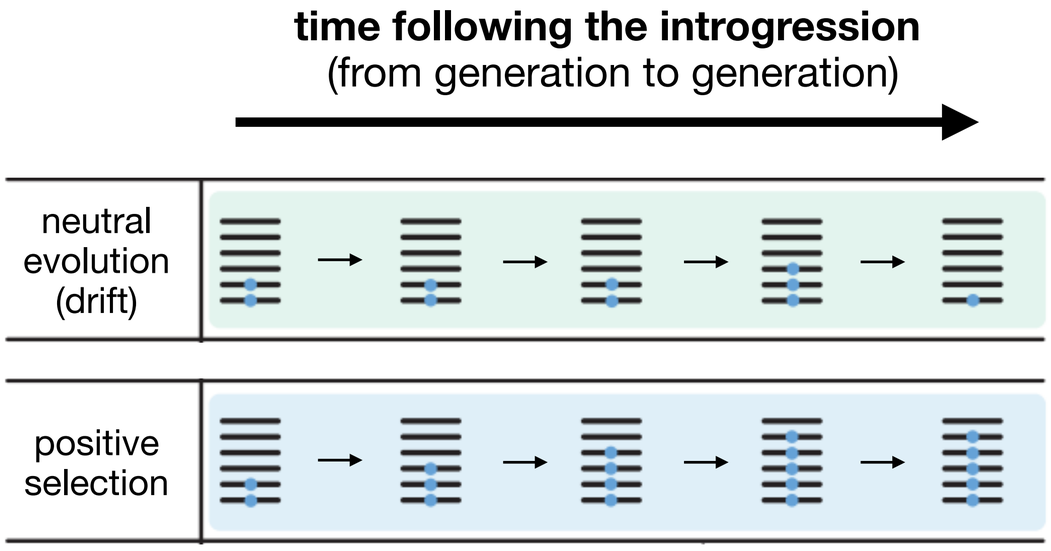
What are the possible fates of introgressed DNA?



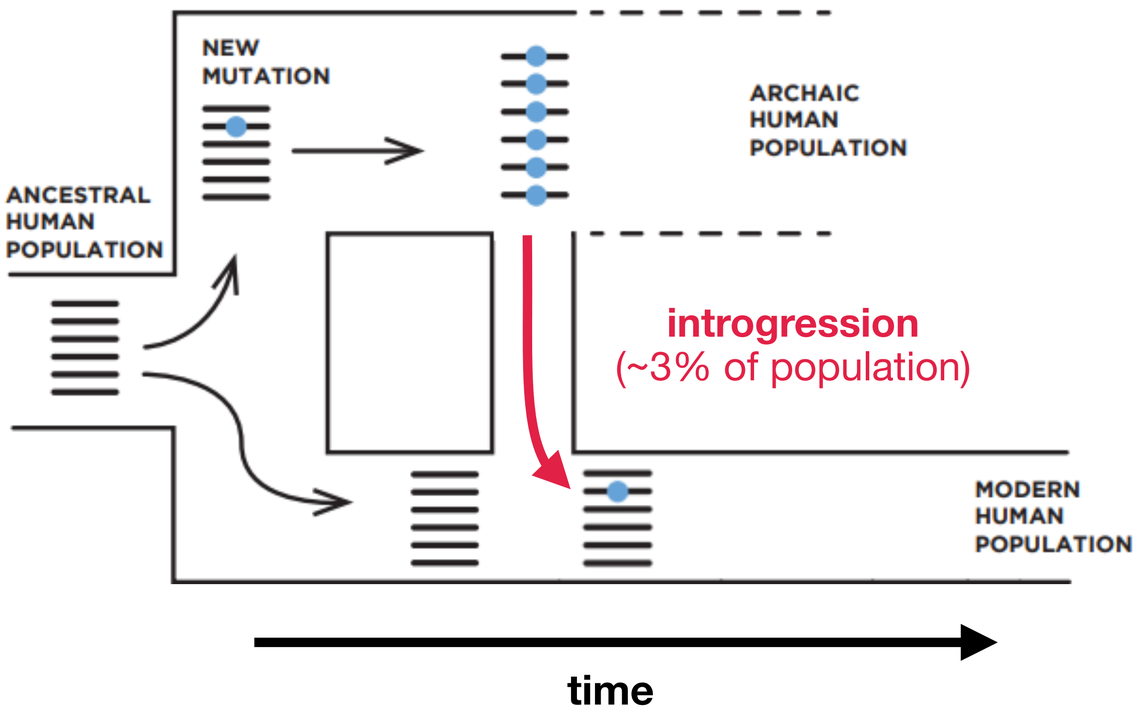


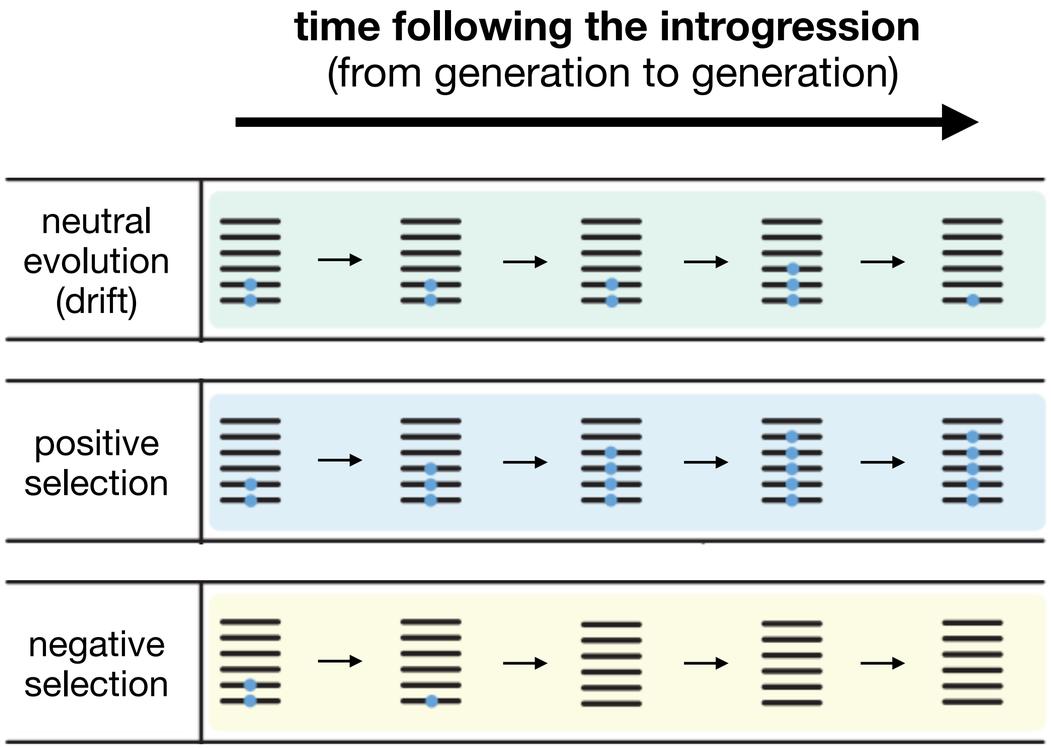
What are the possible fates of introgressed DNA?

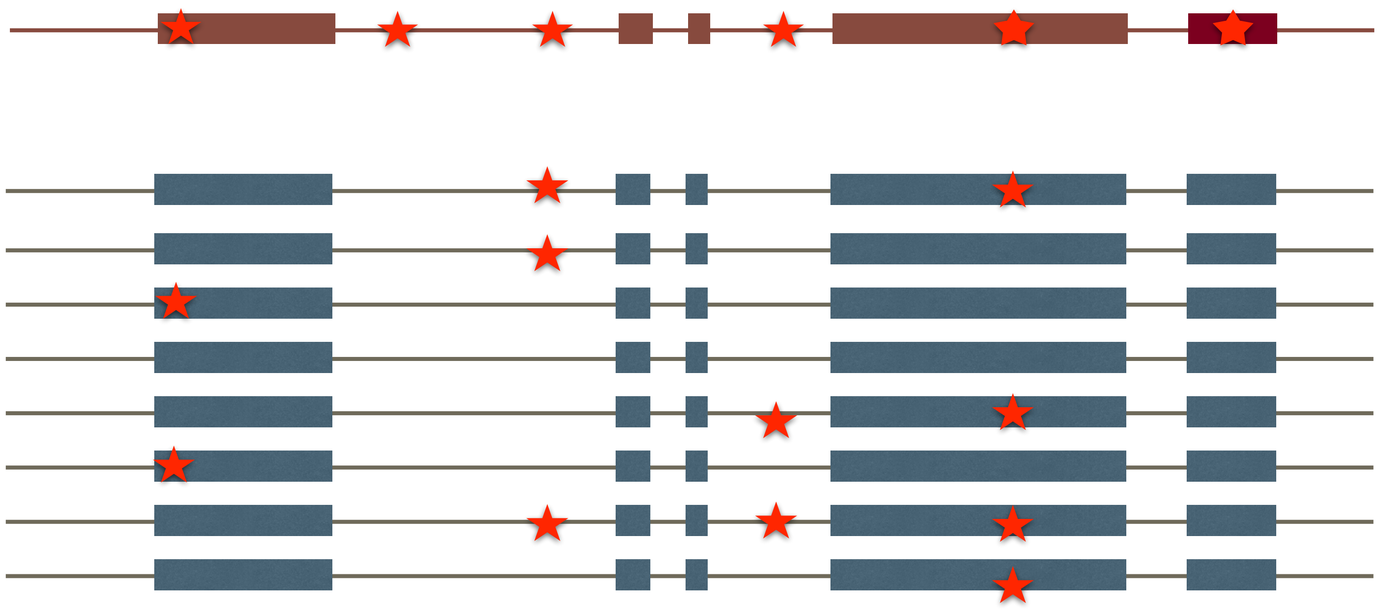




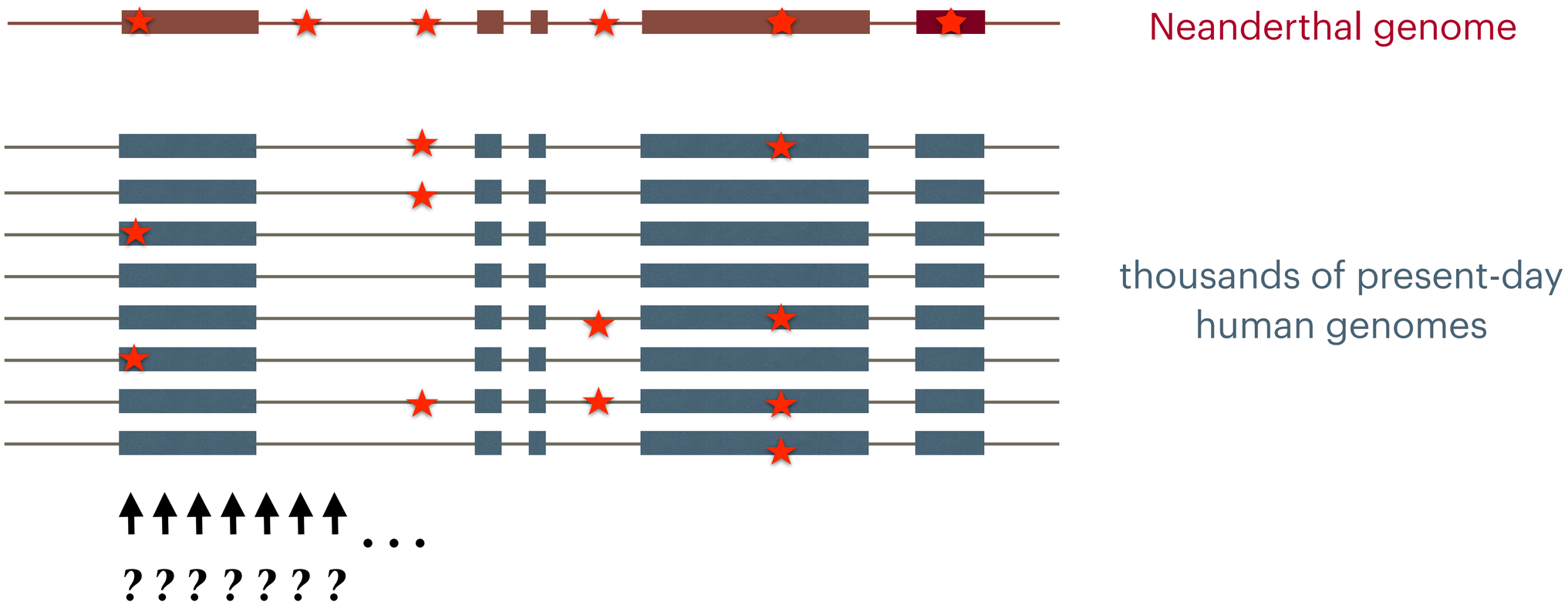
What are the possible fates of introgressed DNA?





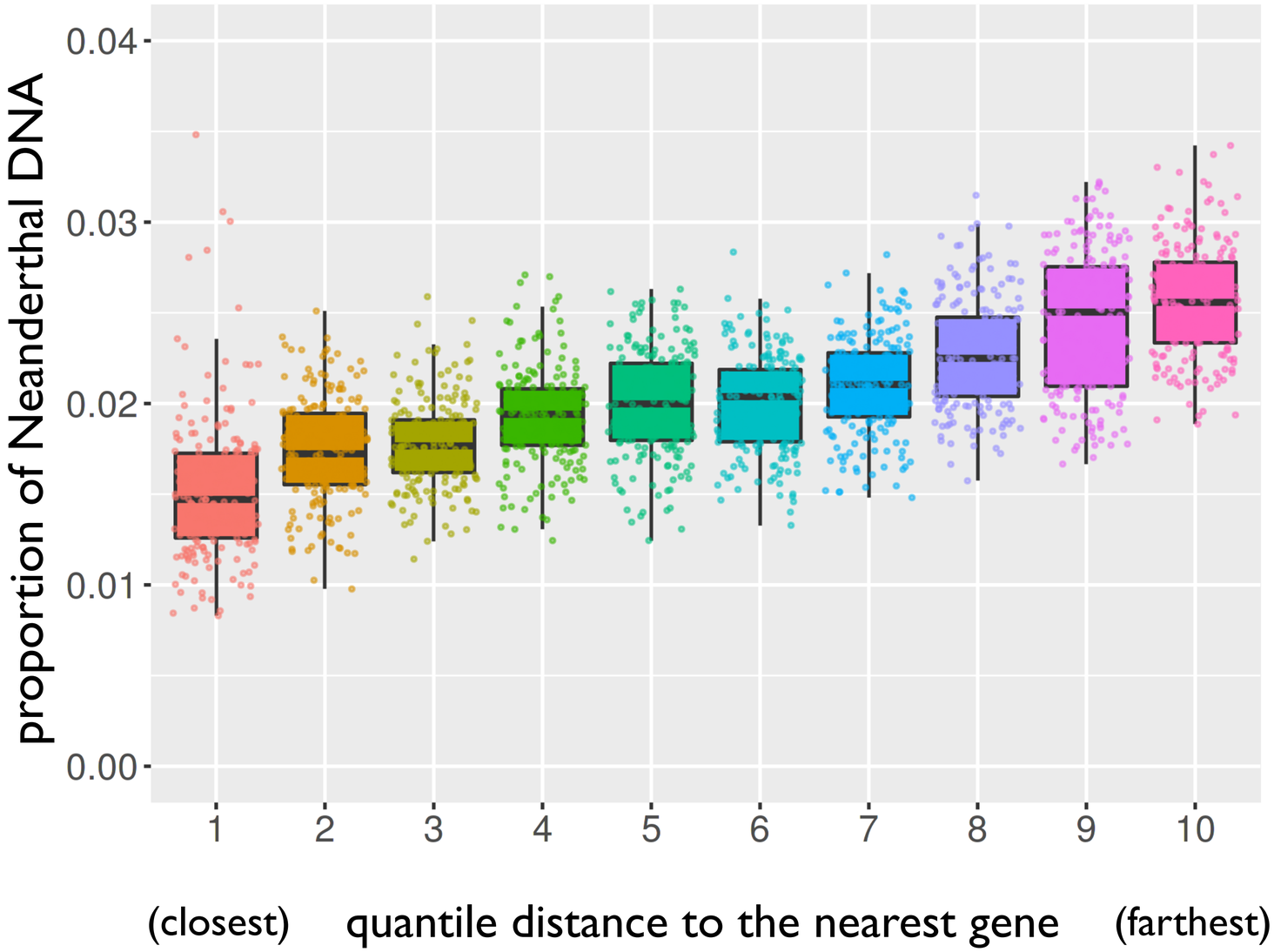

Neanderthal genome
thousands of present-day human genomes
= Neanderthal-derived mutation
= exon / promotor / enhancer / etc.

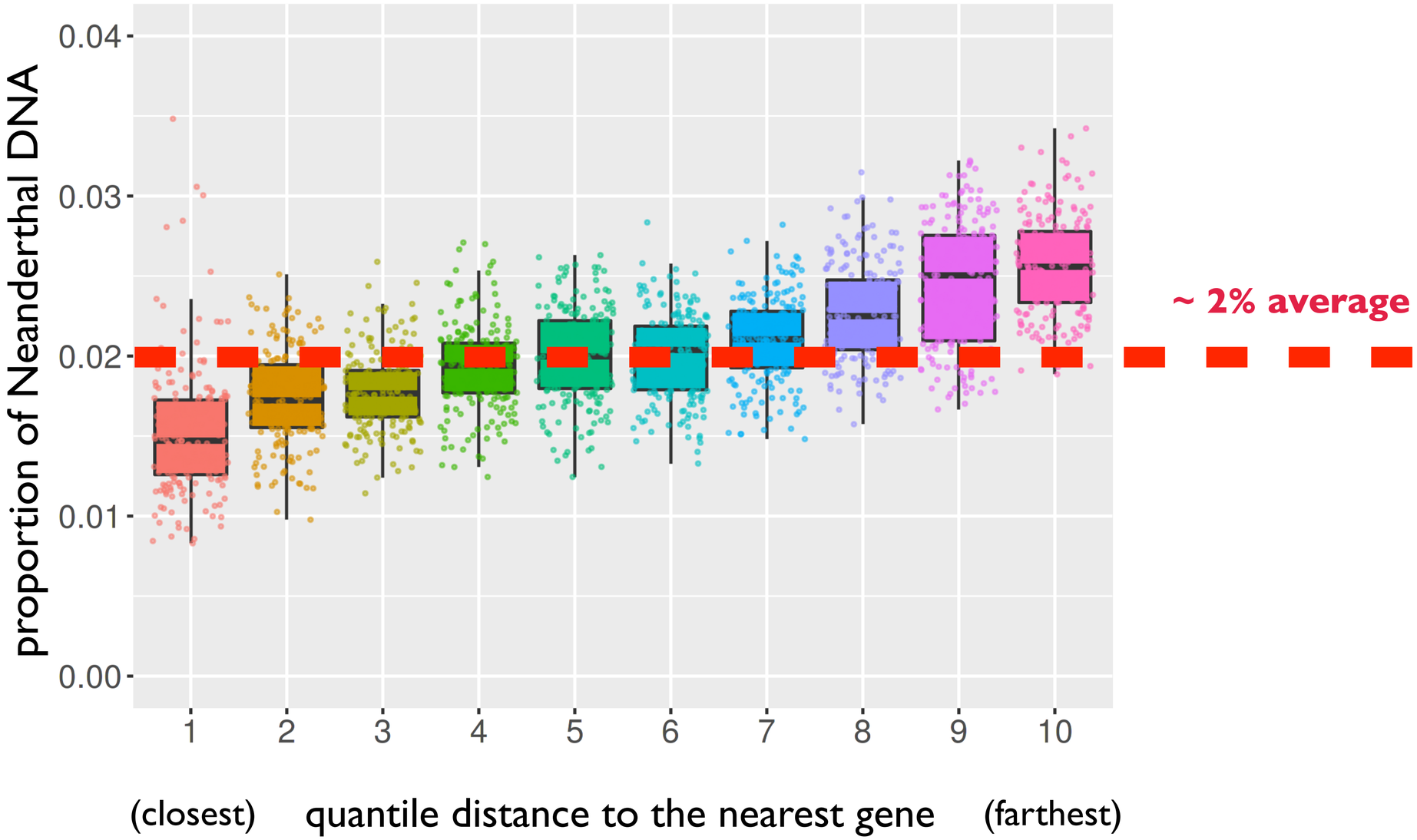
Is there a relationship between the presence of a Neanderthal allele and its distance to the nearest important gene?



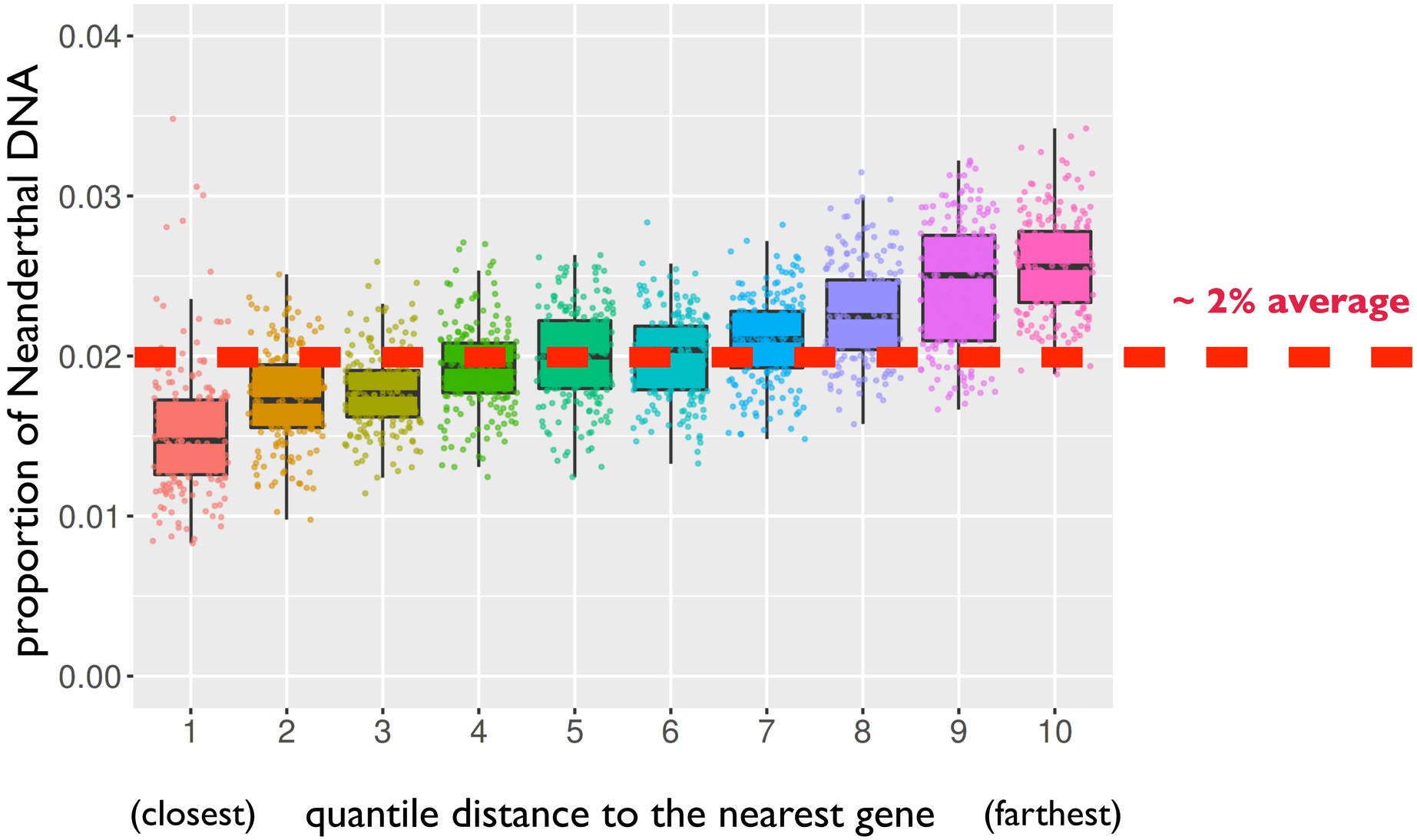
= Neanderthal-derived mutation

= exon / promotor / enhancer / etc.

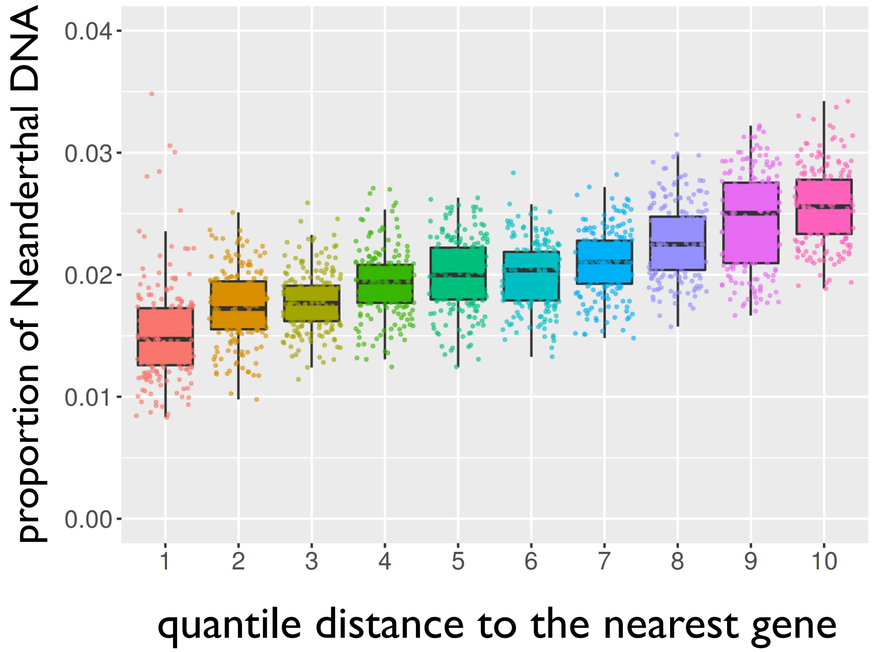

proportion of Neanderthal DNA
0.04
0.03
0.02
0.01
0.00
1
2
3
4
5
6
7
8
9
10
(closest)
quantile distance to the nearest gene
(farthest)

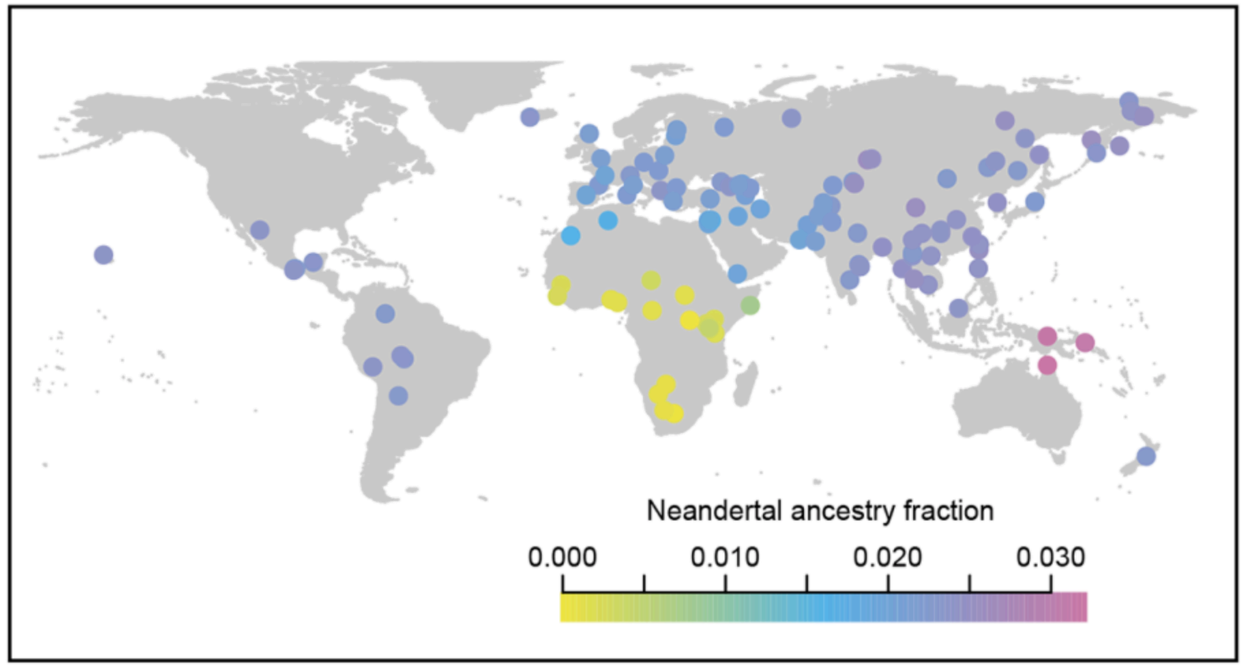

proportion of Neanderthal DNA
0.04
0.03
0.02
0.01
0.00
1
2
3
4
5
6
7
8
9
10
~ 2% average
(closest)
quantile distance to the nearest gene
(farthest)

Genome-wide selection *against* Neanderthal DNA!



On what time scale did this happen?

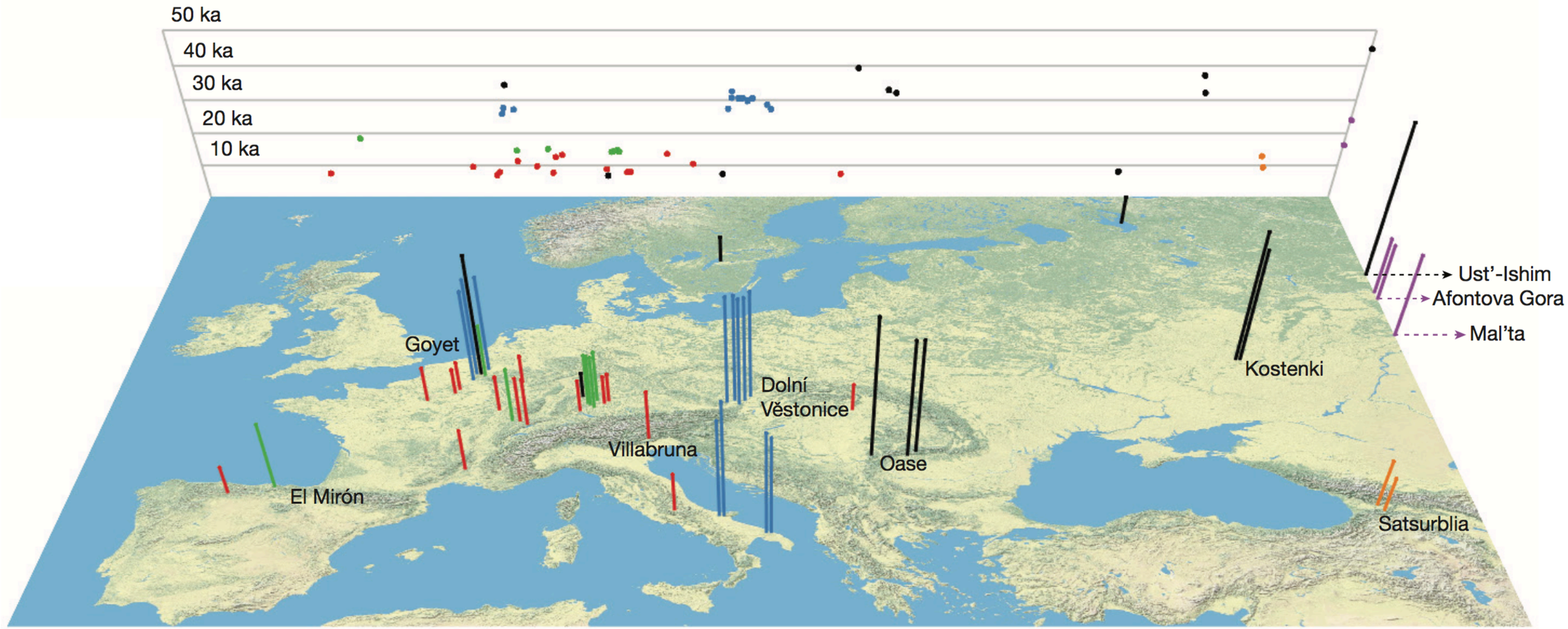




Pruefer *et al*. (Science 2017)

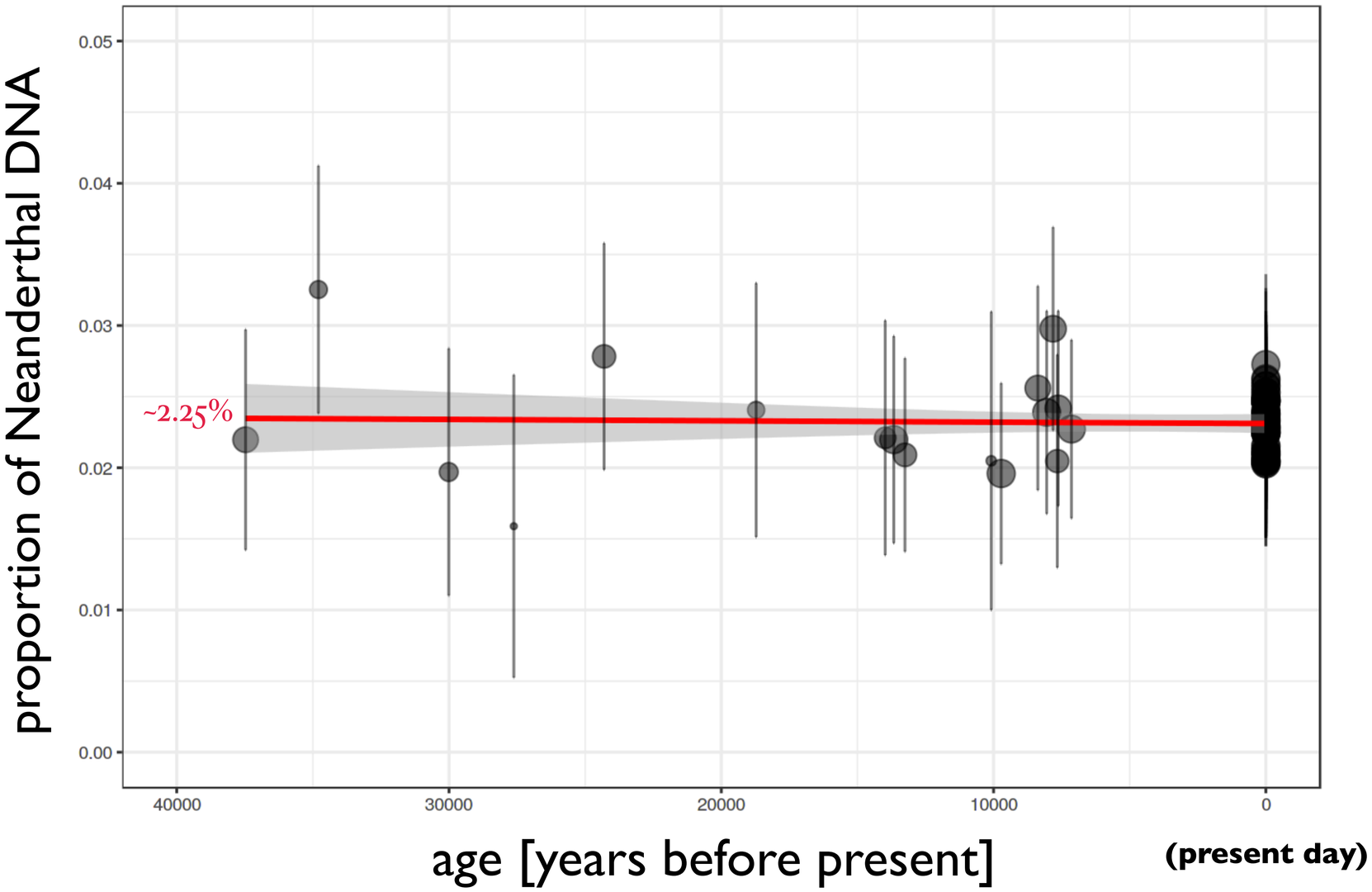
aDNA of "early-modern humans" from Europe

(45-10 kya BP)



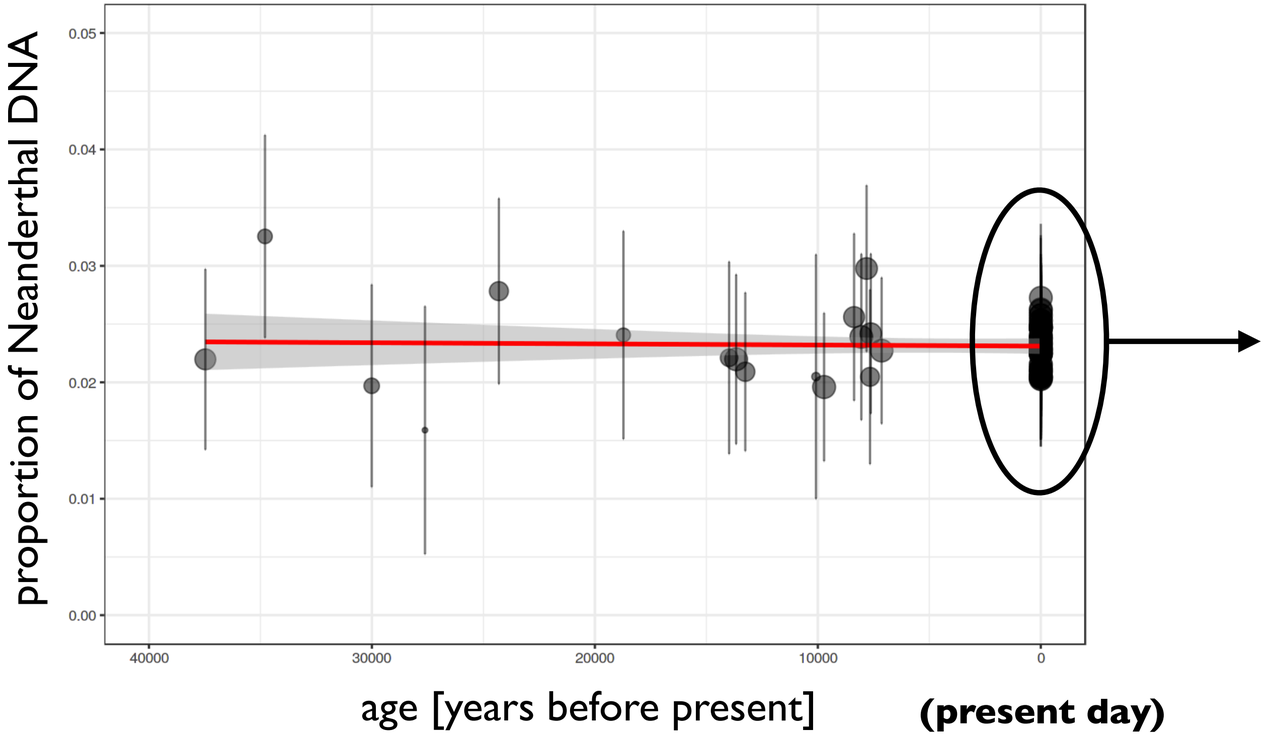
Fu *et al.* (Nature, 2016)

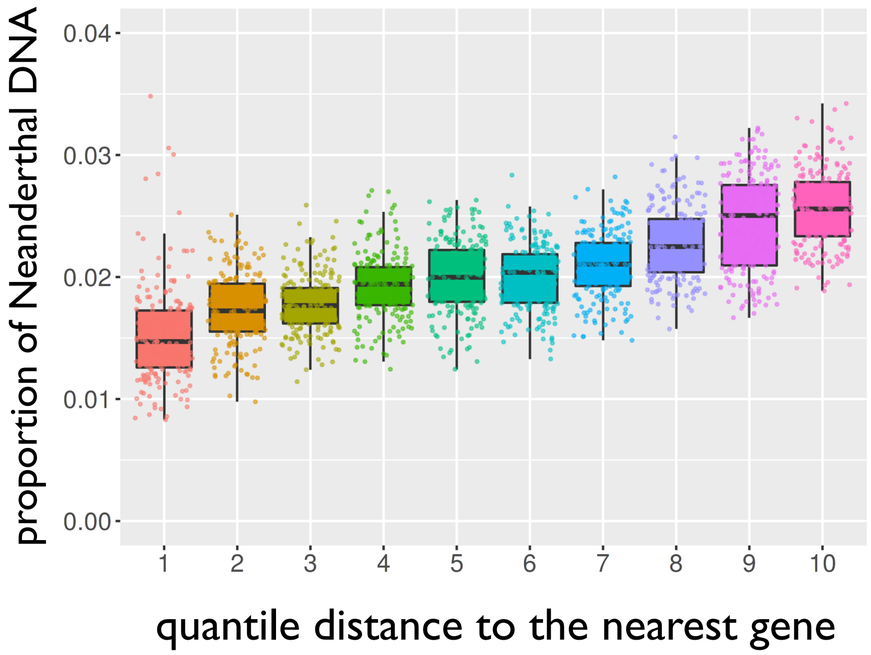
Proportion of Neanderthal ancestry in modern humans remained constant for tens of thousands of years



Petr *et al*. (PNAS 2019)

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

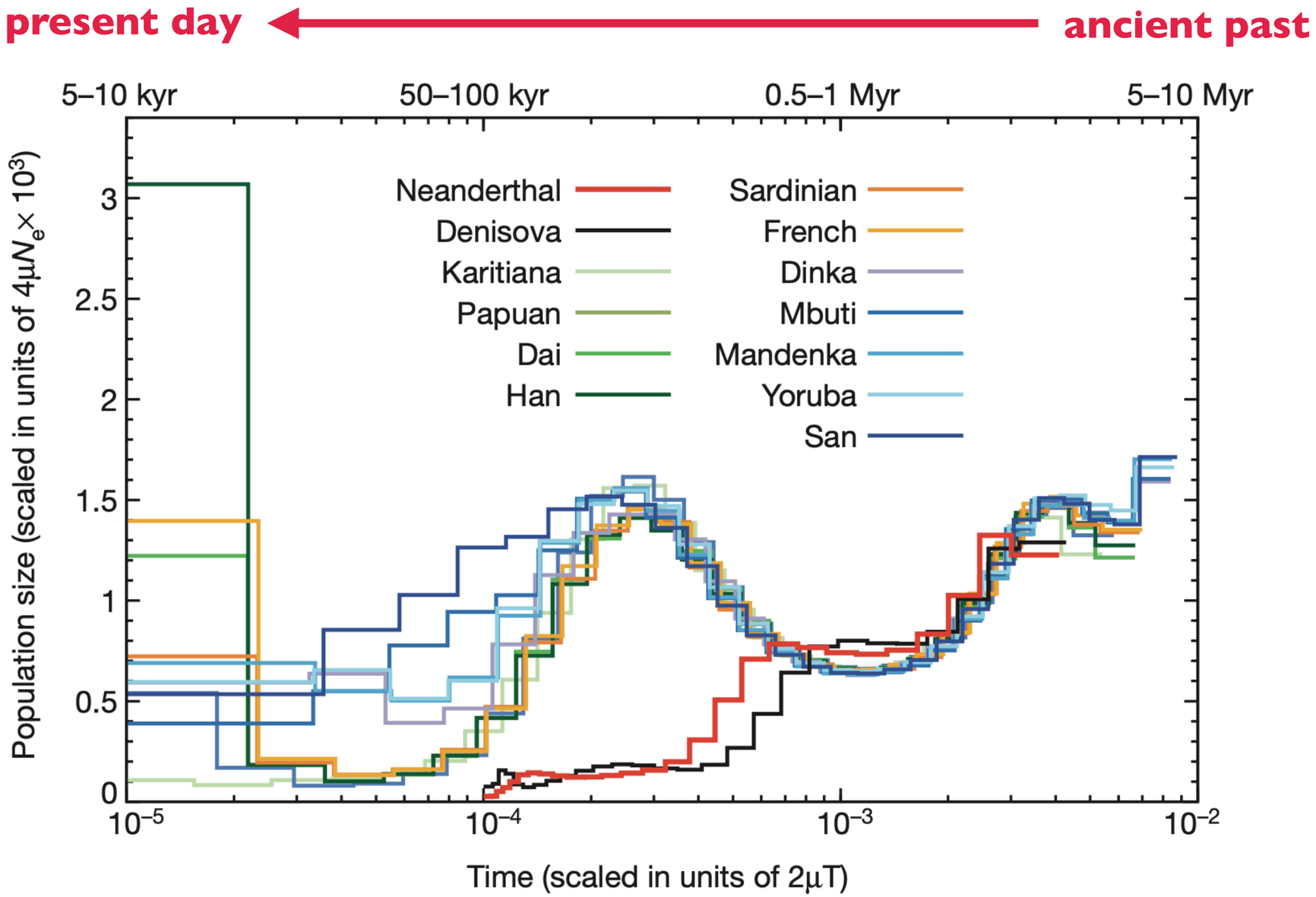




Petr *et al*. (PNAS 2019)

Demographic collapse of archaic human populations

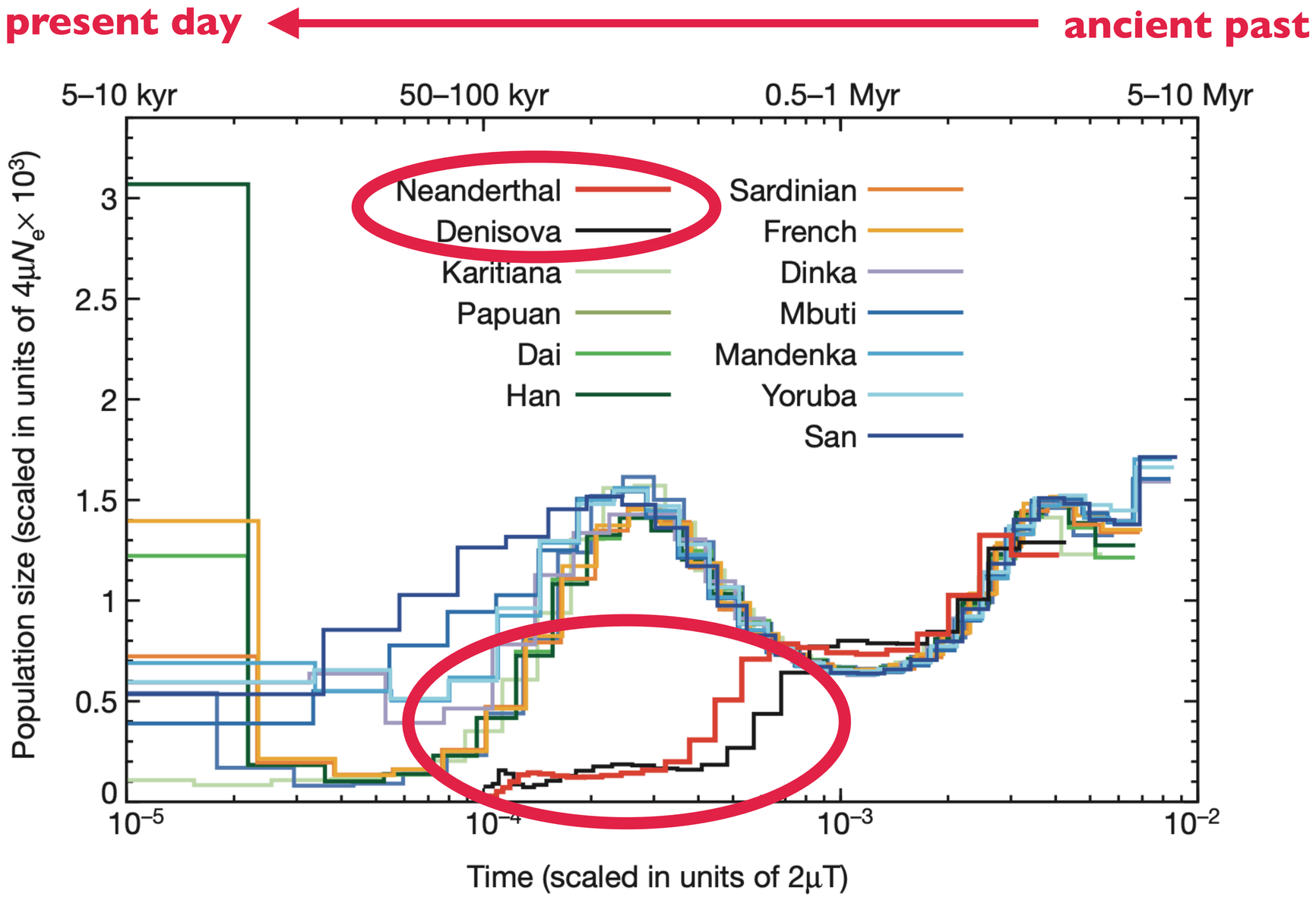
Pairwise Sequentially Markovian Coalescent (PSMC) model



Pruefer *et al*. (Nature, 2014)

Demographic collapse of archaic human populations

Pairwise Sequentially Markovian Coalescent (PSMC) model



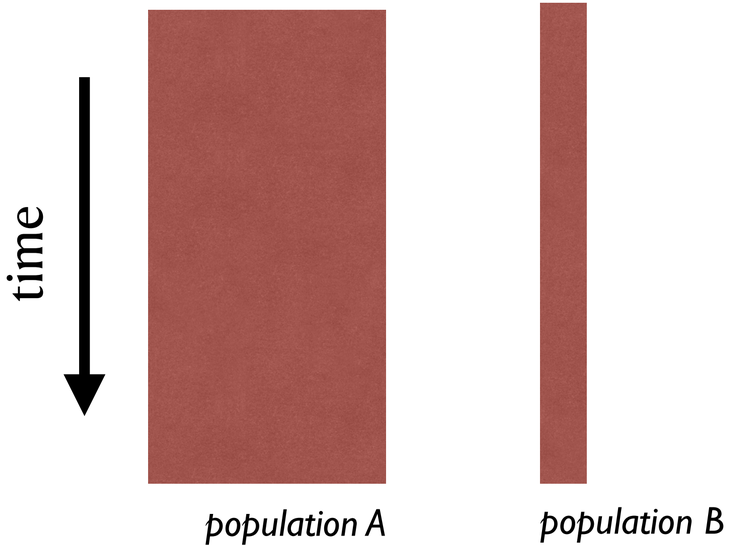
Pruefer *et al*. (Nature, 2014)

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

Nearly-neutral theory of molecular evolution

Tomoko Ohta

https://womenyoushouldknow.net/population-genetics-tomoko-ohta/

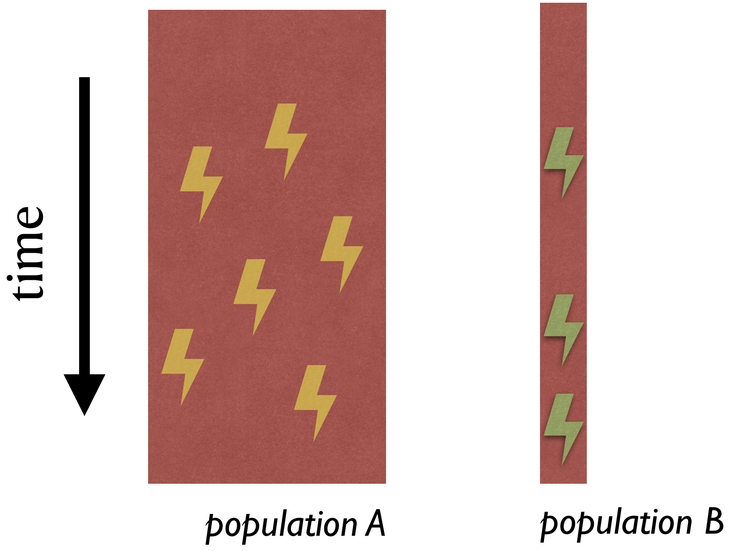


Ohta (Nature, 1973)

Nearly-neutral theory of molecular evolution

Tomoko Ohta

https://womenyoushouldknow.net/population-genetics-tomoko-ohta/



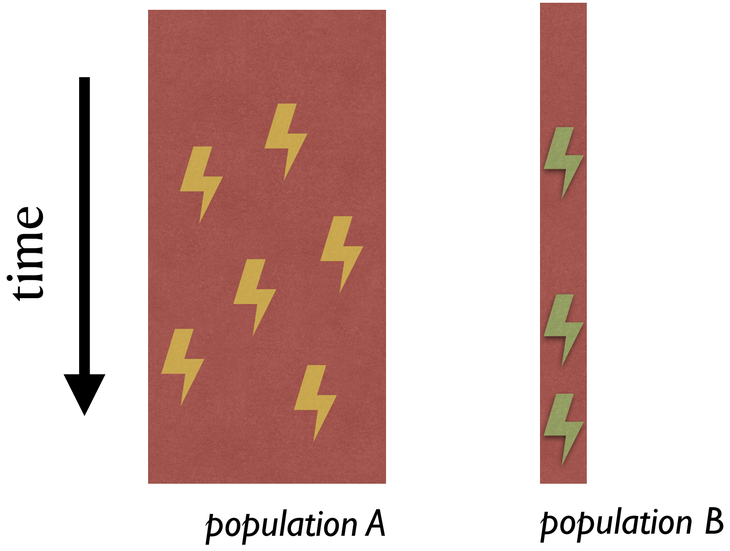
We know that genomes accumulate mutations

Ohta (Nature, 1973)

Nearly-neutral theory of molecular evolution

Tomoko Ohta

https://womenyoushouldknow.net/population-genetics-tomoko-ohta/



We know that genomes accumulate mutations

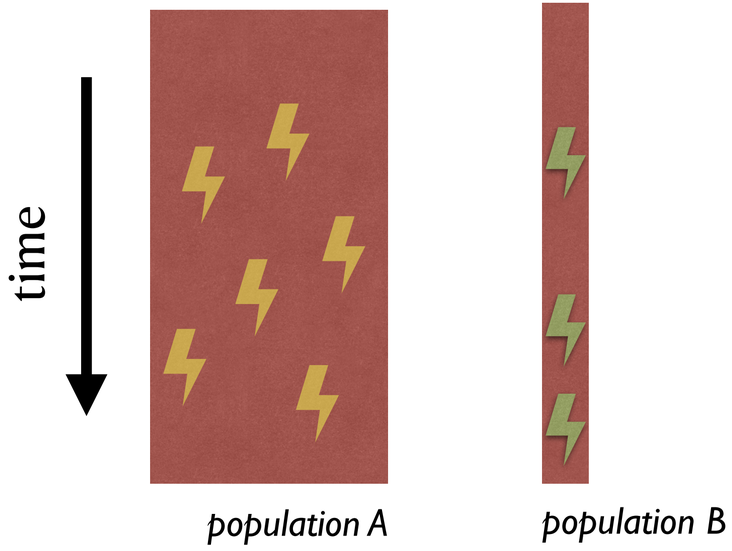
- a tiny number are extremely negative (**lethal**)
- a tiny number are positive (**adaptive**)
- a large number does nothing (**neutral**)
- many are “mildly deleterious” (**nearly neutral**)

Ohta (Nature, 1973)

Nearly-neutral theory of molecular evolution

Tomoko Ohta

https://womenyoushouldknow.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**)

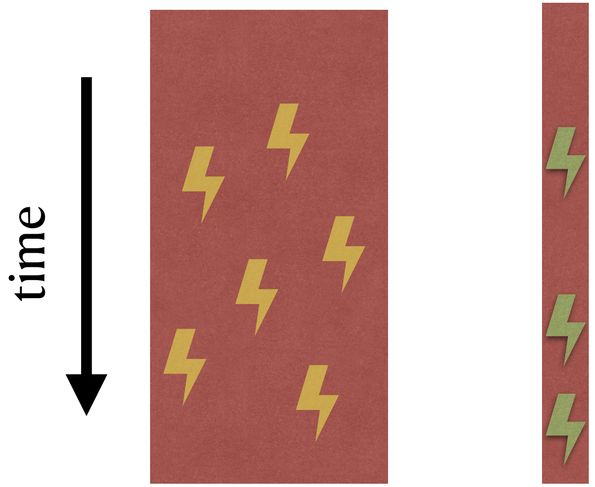
Efficacy of negative selection depends on the effectice population size (N_e).

Ohta (Nature, 1973)

Nearly-neutral theory of molecular evolution

Tomoko Ohta

https://womenyoushouldknow.net/population-genetics-tomoko-ohta/



population A *population B*

N_e(pop A) > N_e(pop B)

lower fixation of bad mutations < higher fixation of bad mutations

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**)

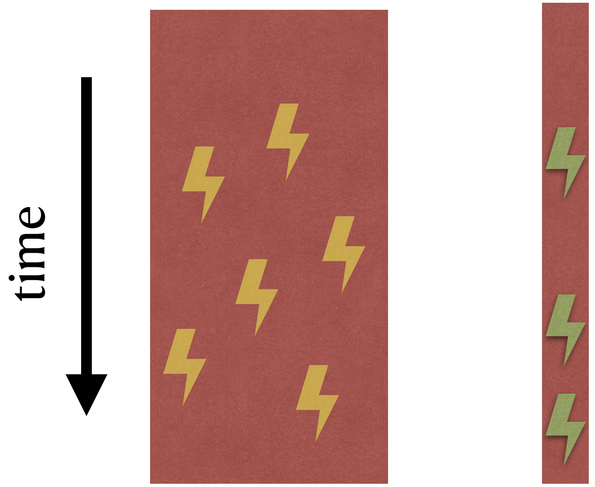
Efficacy of negative selection depends on the effectice population size (N_e).

Ohta (Nature, 1973)

Nearly-neutral theory of molecular evolution

Tomoko Ohta

https://womenyoushouldknow.net/population-genetics-tomoko-ohta/



population A *population B*

N_e(pop A) > N_e(pop B)

lower fixation of bad mutations < higher fixation of bad mutations

selection coefficient: s ~ 1 / Ne

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**)

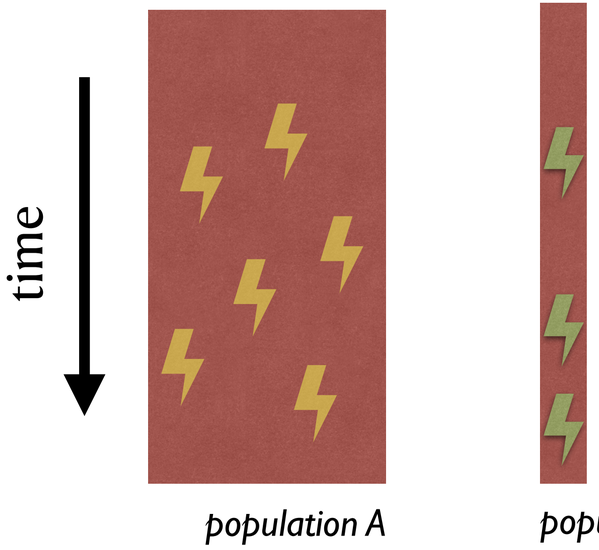
Efficacy of negative selection depends on the effectice population size (N_e).

Ohta (Nature, 1973)

Nearly-neutral theory of molecular evolution

Tomoko Ohta

https://womenyoushouldknow.net/population-genetics-tomoko-ohta/

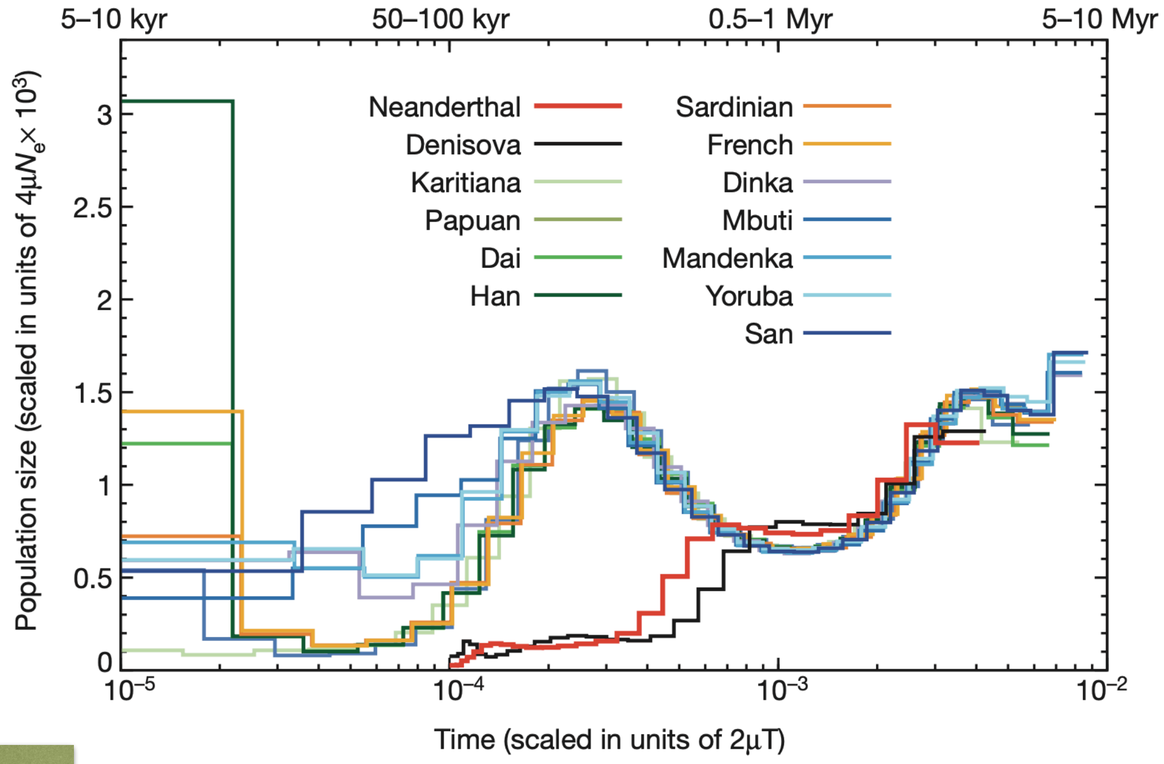


N_e(pop A) > N_e(pop B)

lower fixation of bad mutations < higher fixation of bad mutations

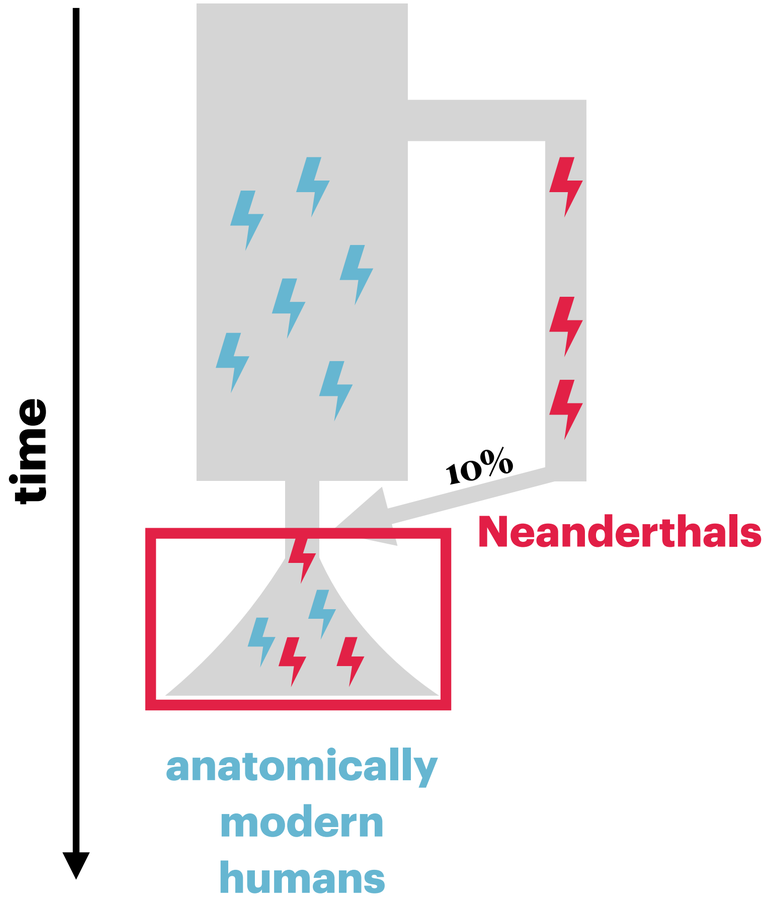
"anatomically modern humans"

"Neanderthals"

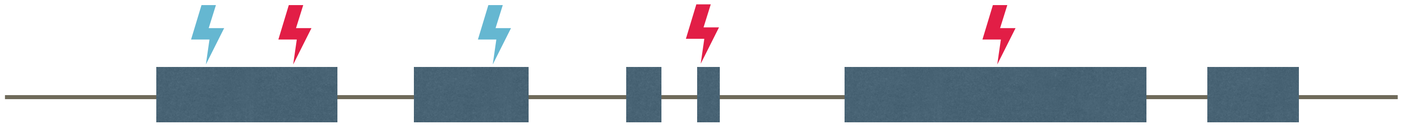


Ohta (Nature, 1973)

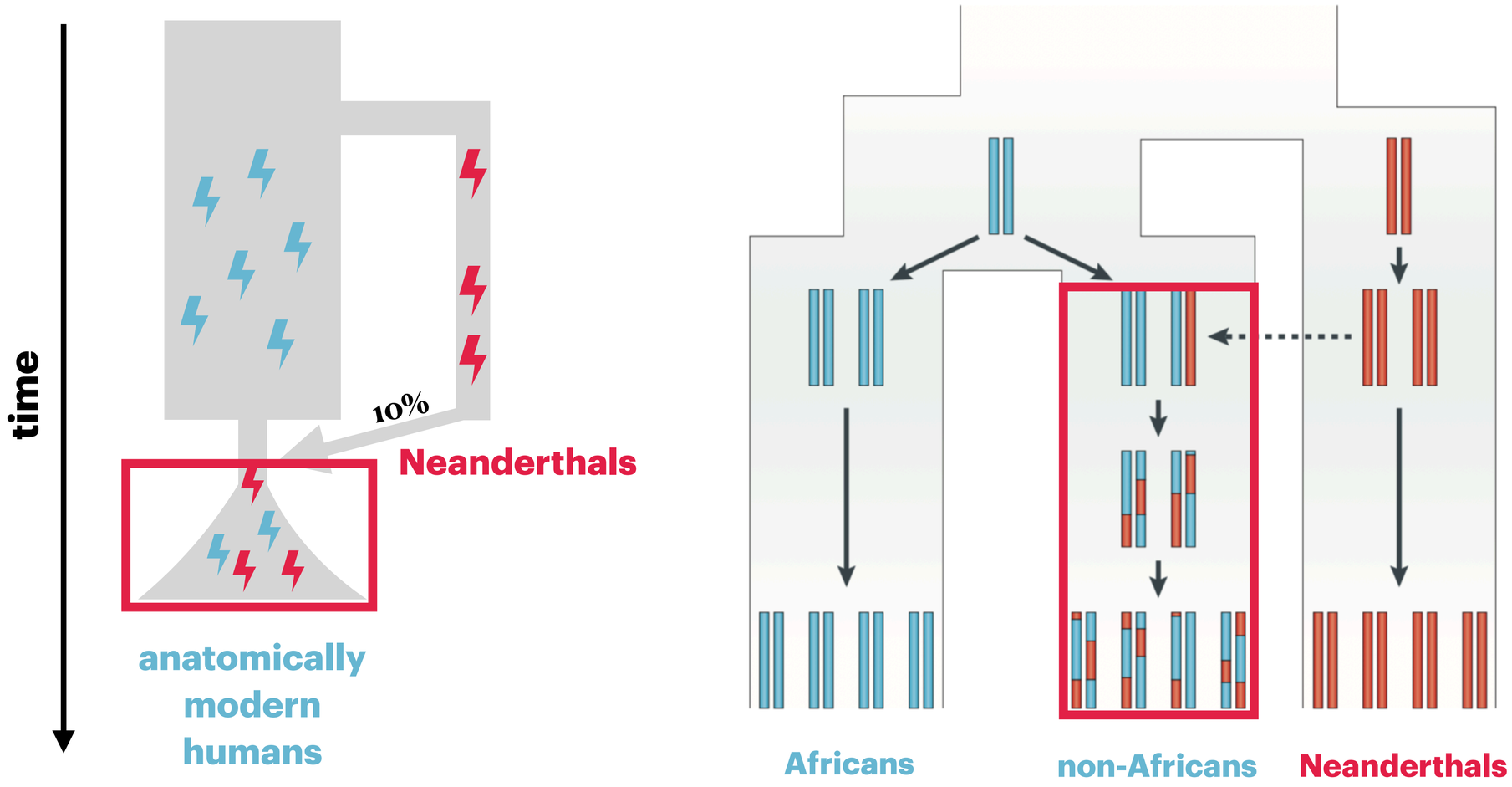
Simulations: tracing the trajectory of introgressed DNA over time



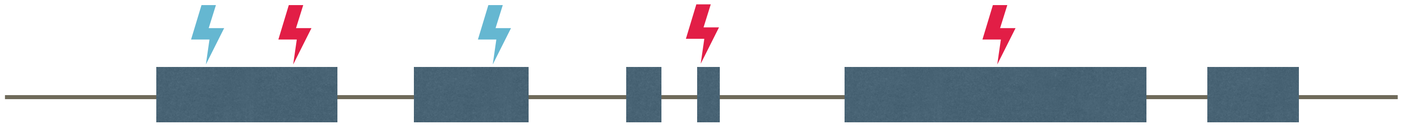
mutations of different
selection coefficients *s*

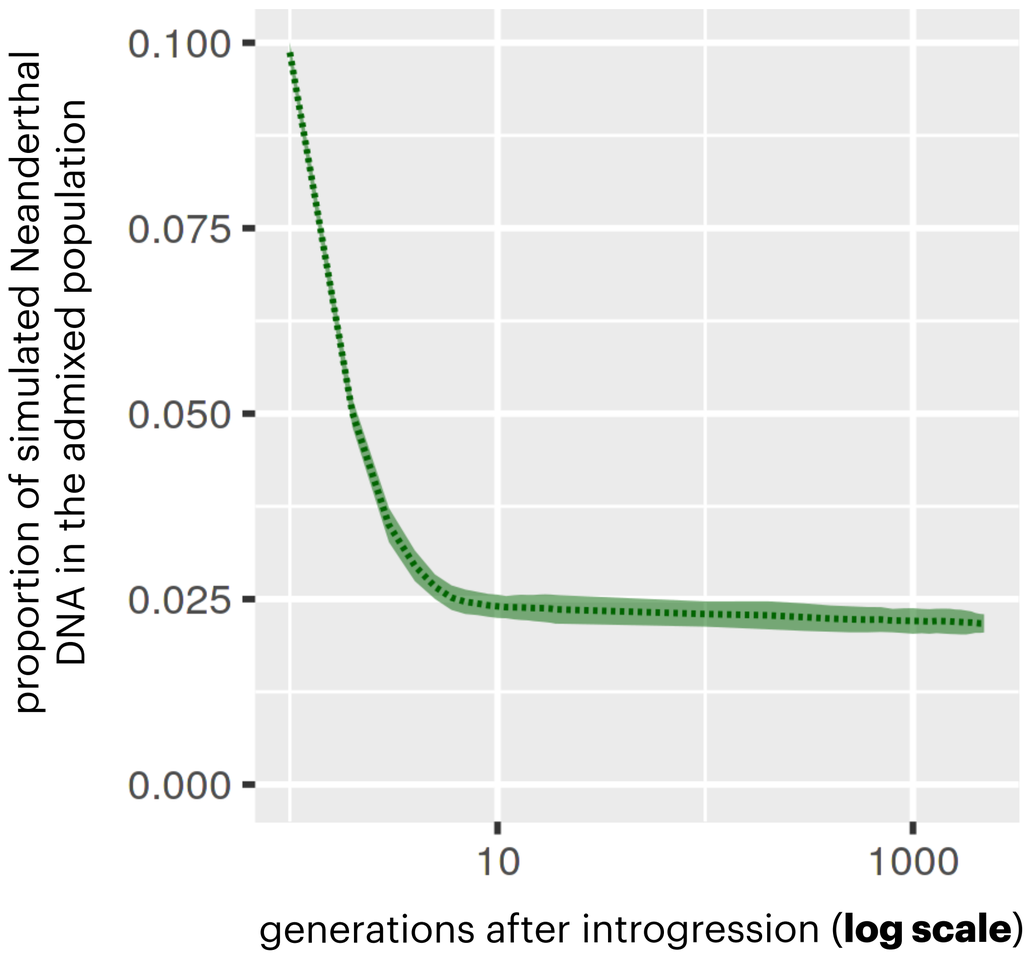


Simulations: tracing the trajectory of introgressed DNA over time

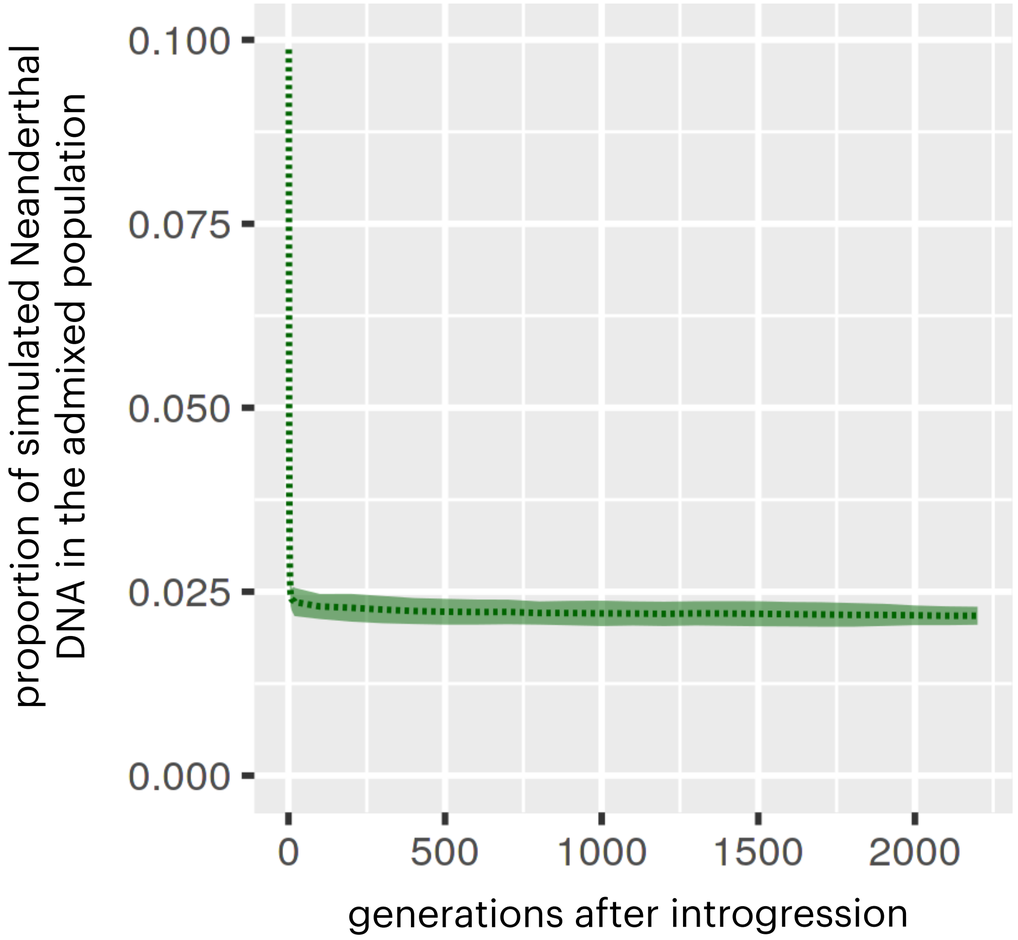


mutations of different selection coefficients *s*

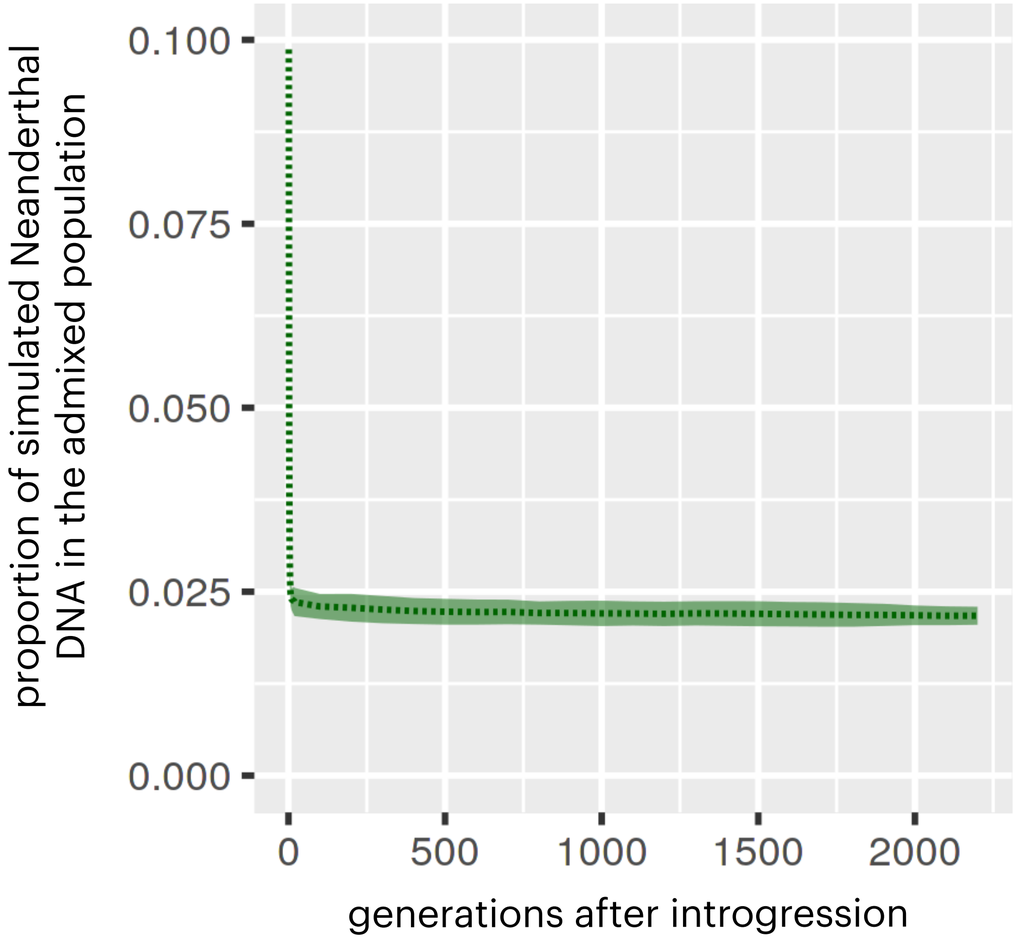


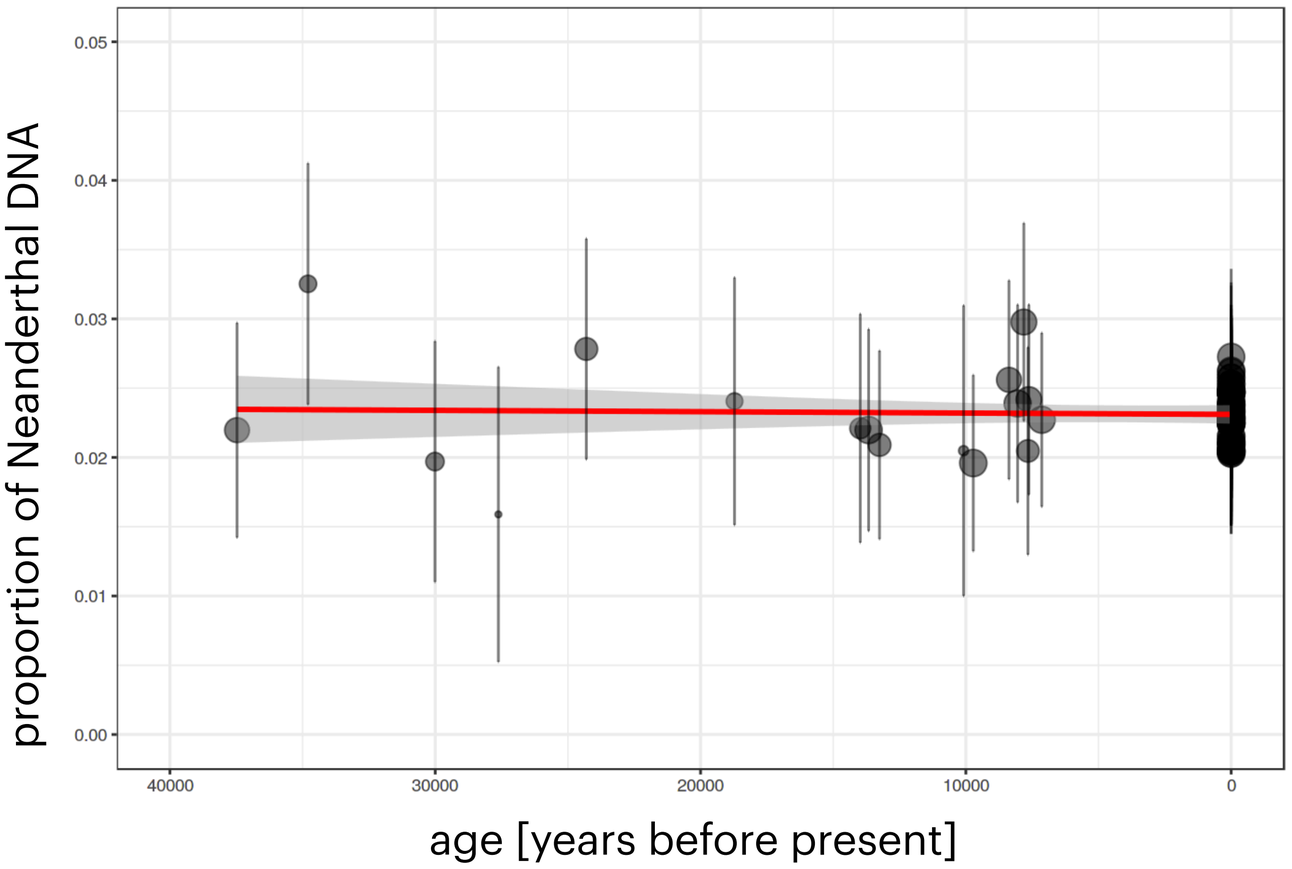


Harris et al. (2016); Juric et al. (2016); Petr *et al*. (2019)



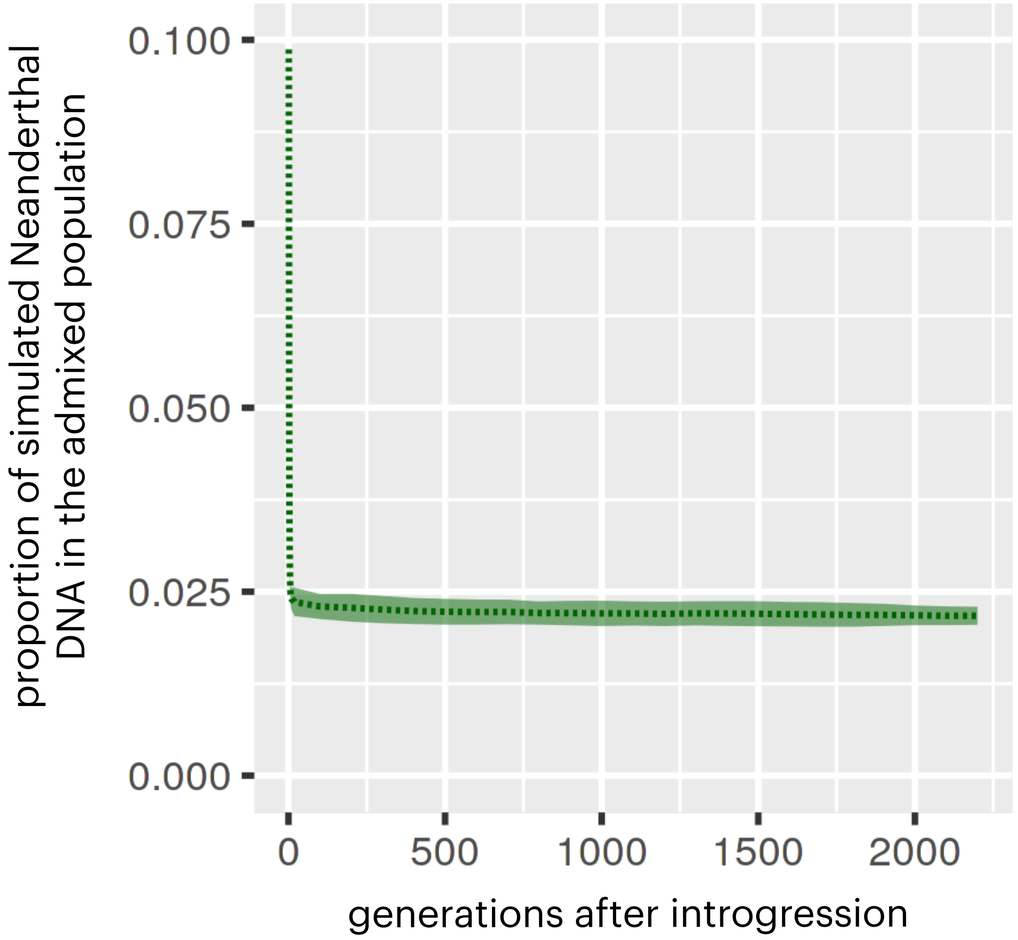
Harris et al. (2016); Juric et al. (2016); Petr *et al*. (2019)

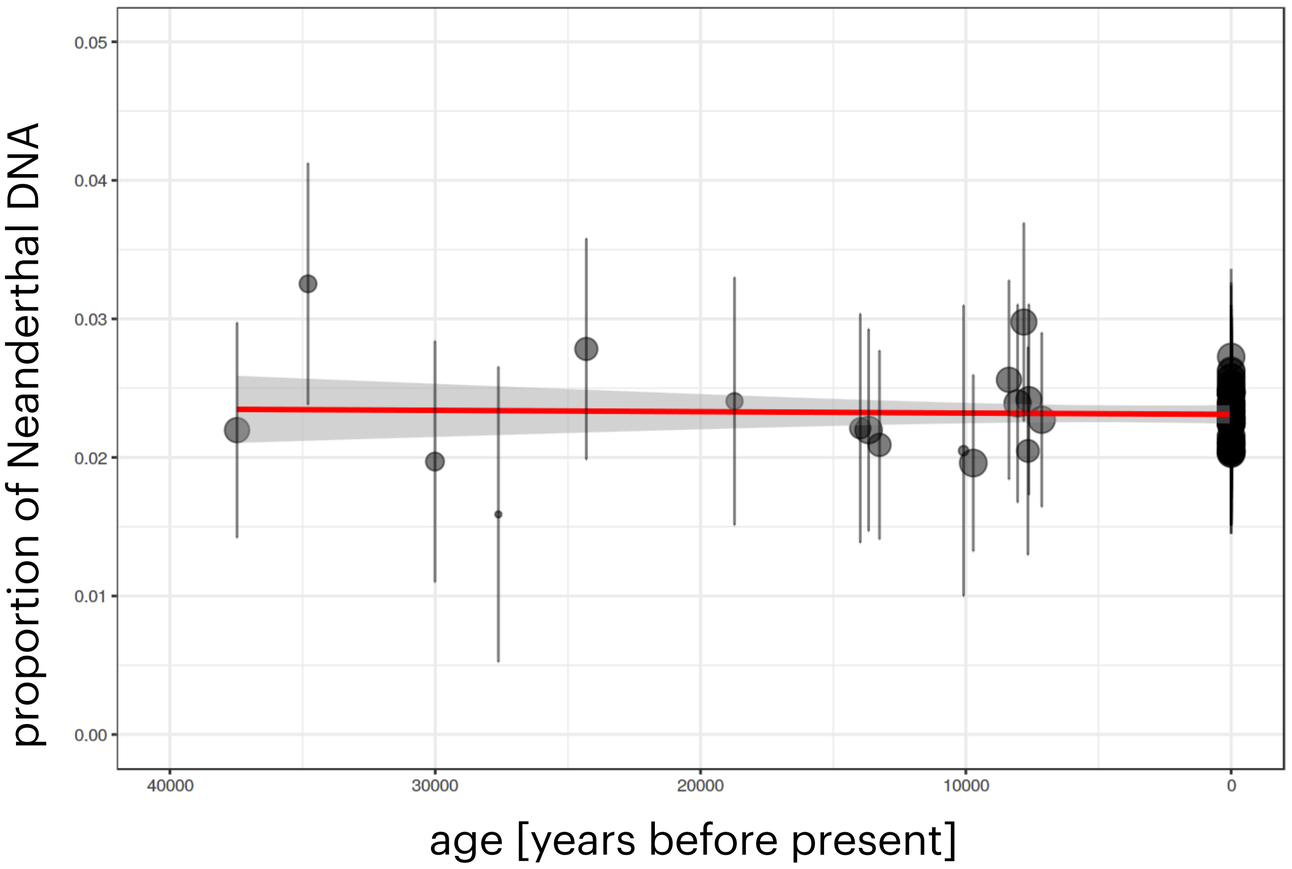




Harris et al. (2016); Juric et al. (2016); Petr *et al*. (2019)

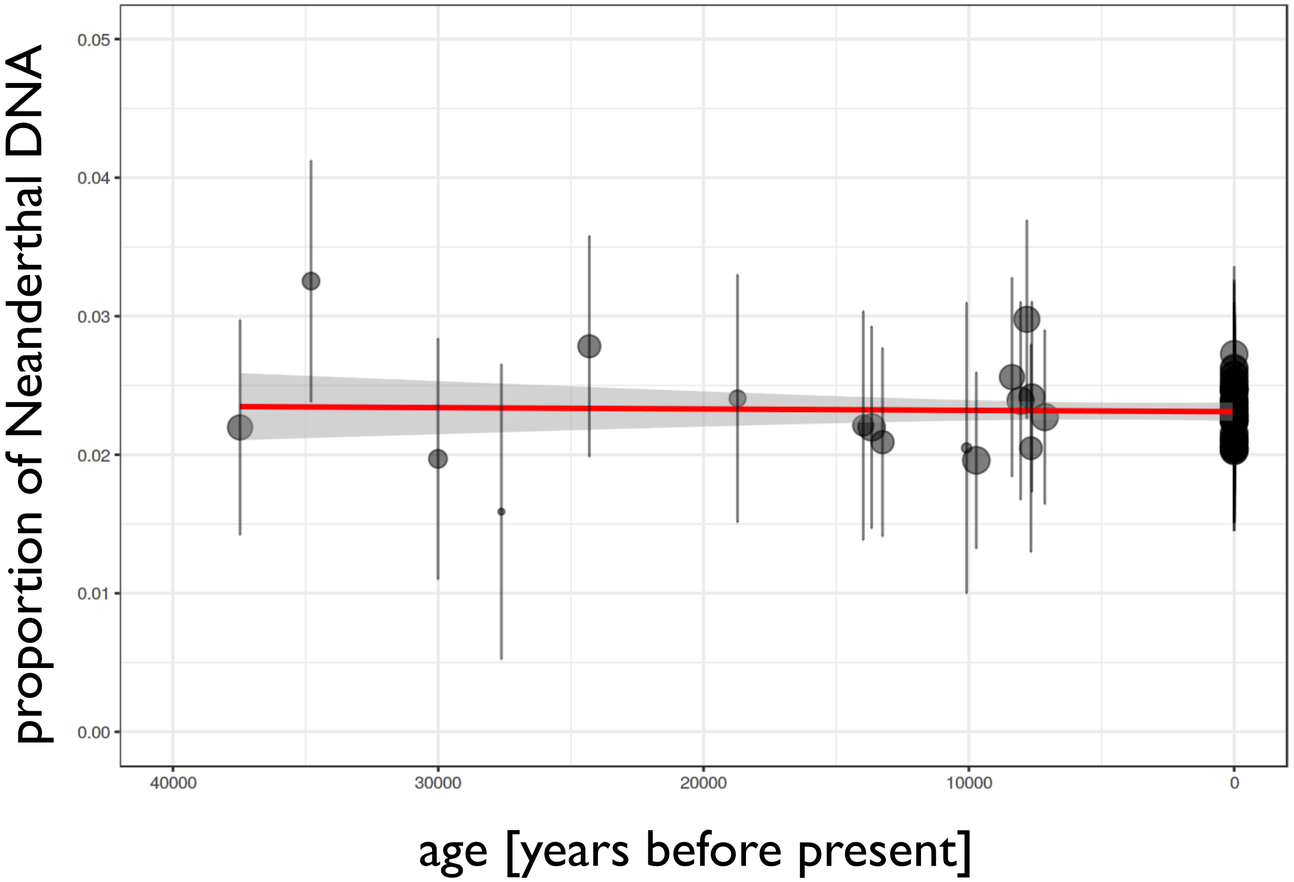
Neanderthal ancestry trajectory matches simulations...

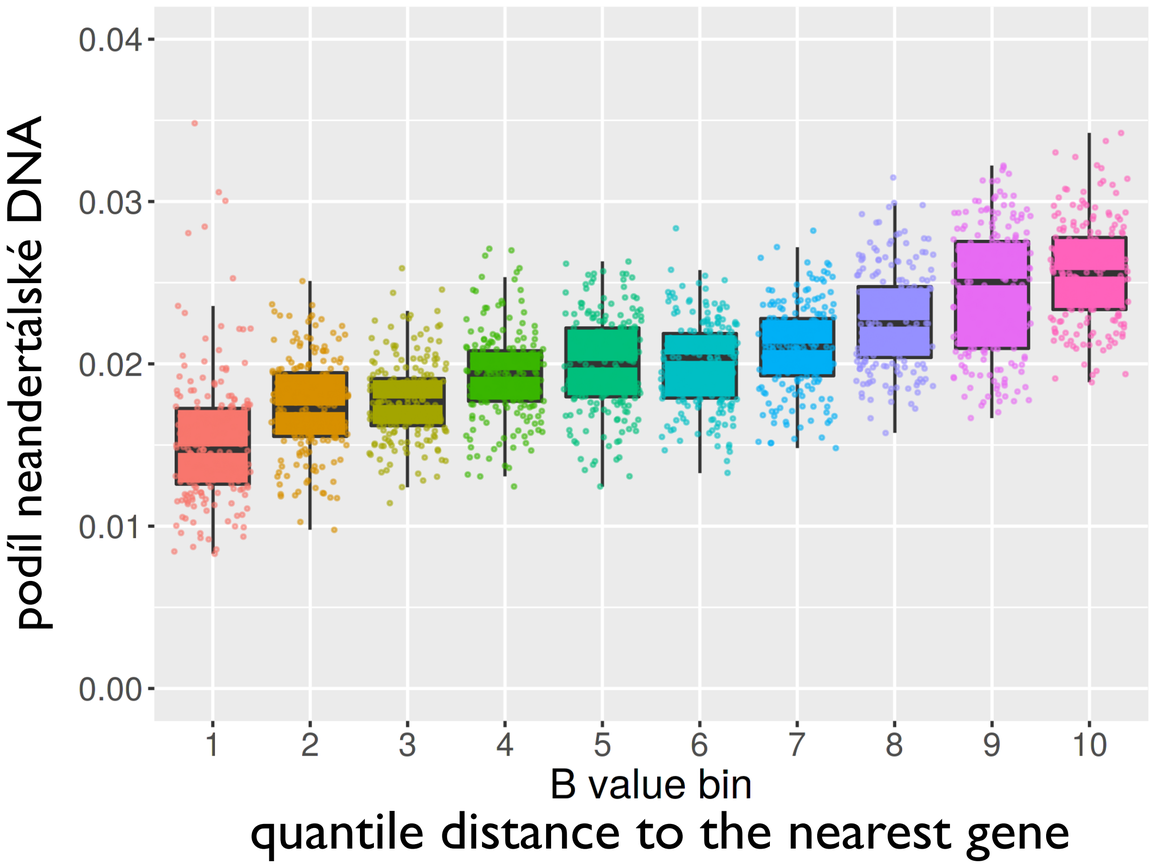




Harris et al. (2016); Juric et al. (2016); Petr *et al.* (2019)

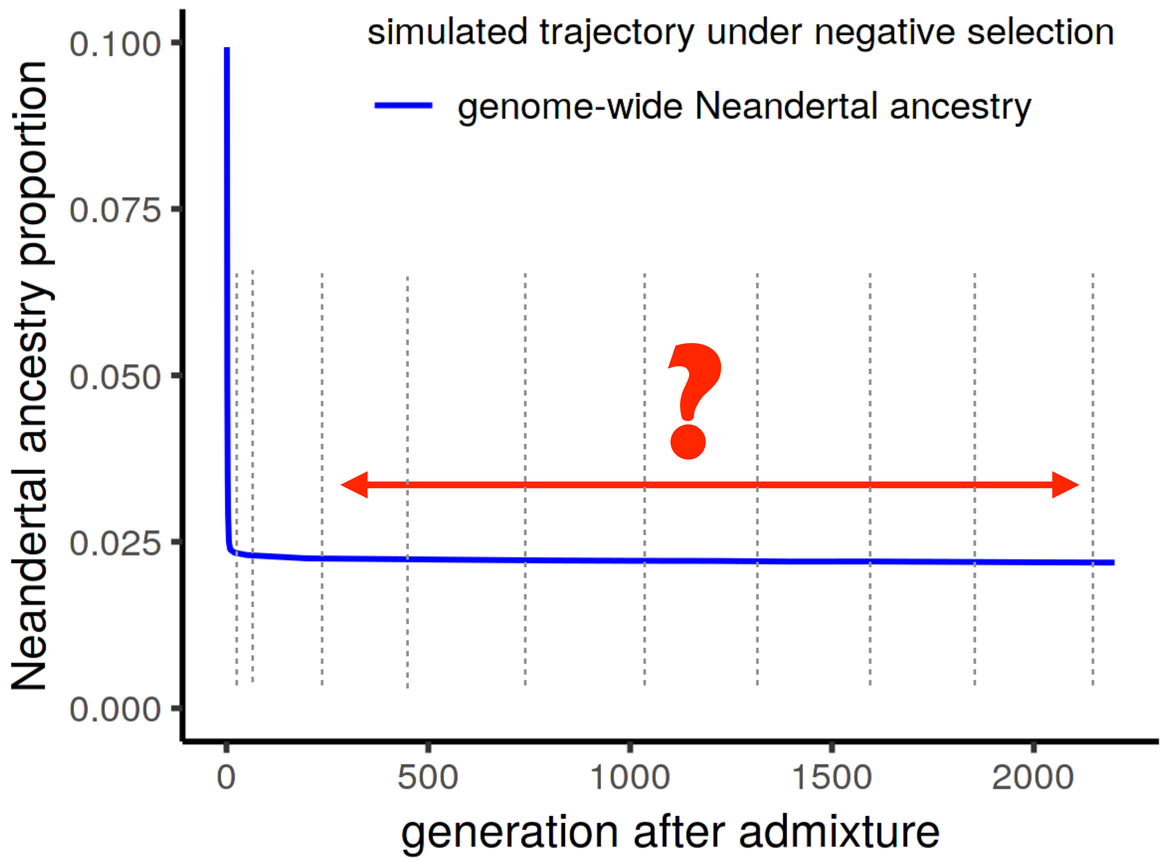
... what about the whole-genome negative selection signal?

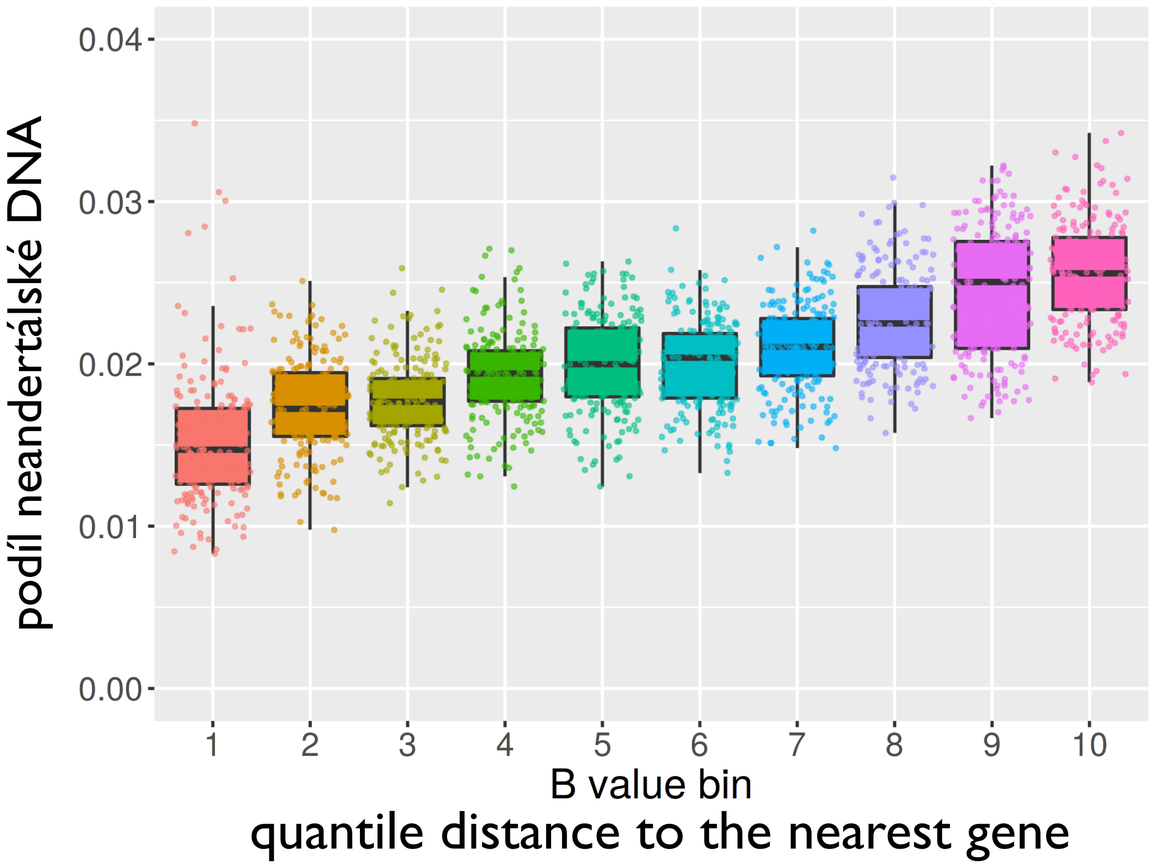




Petr *et al.* (PNAS, 2019)

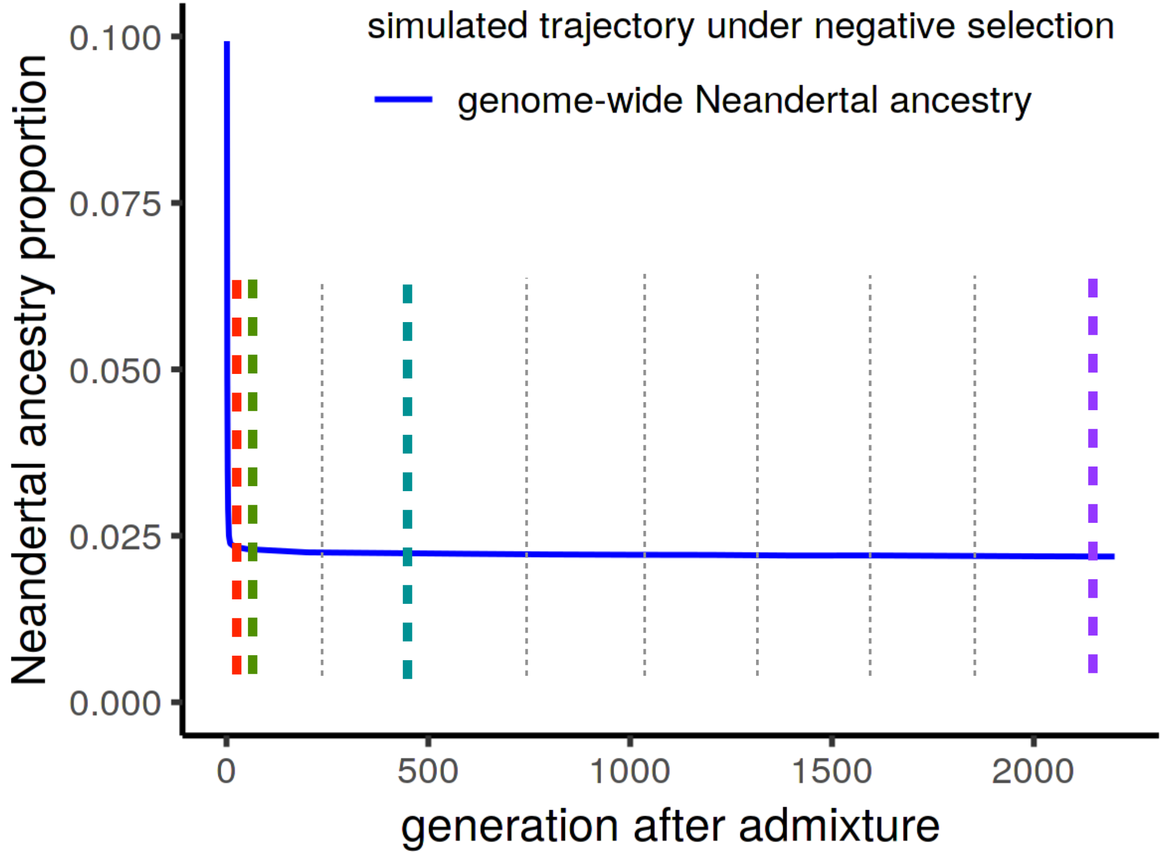
... what about the whole-genome negative selection signal?

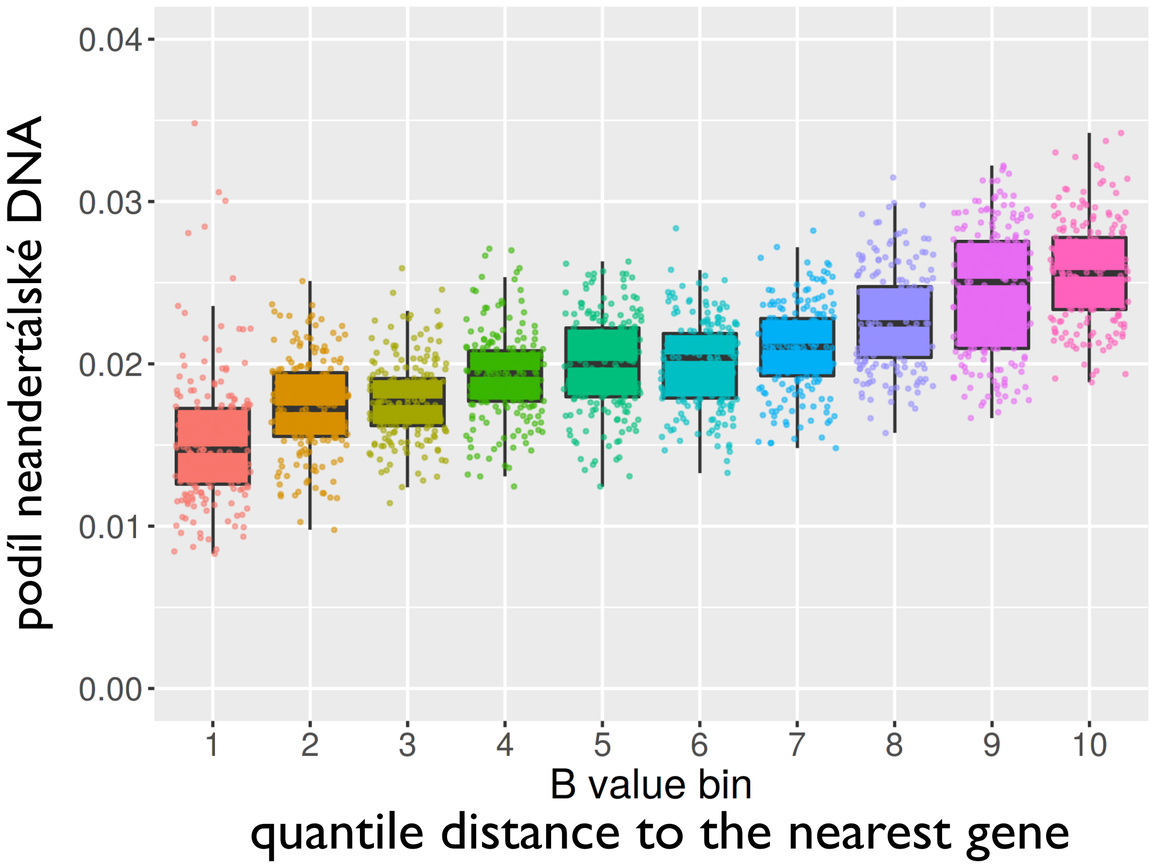




Petr *et al.* (PNAS, 2019)

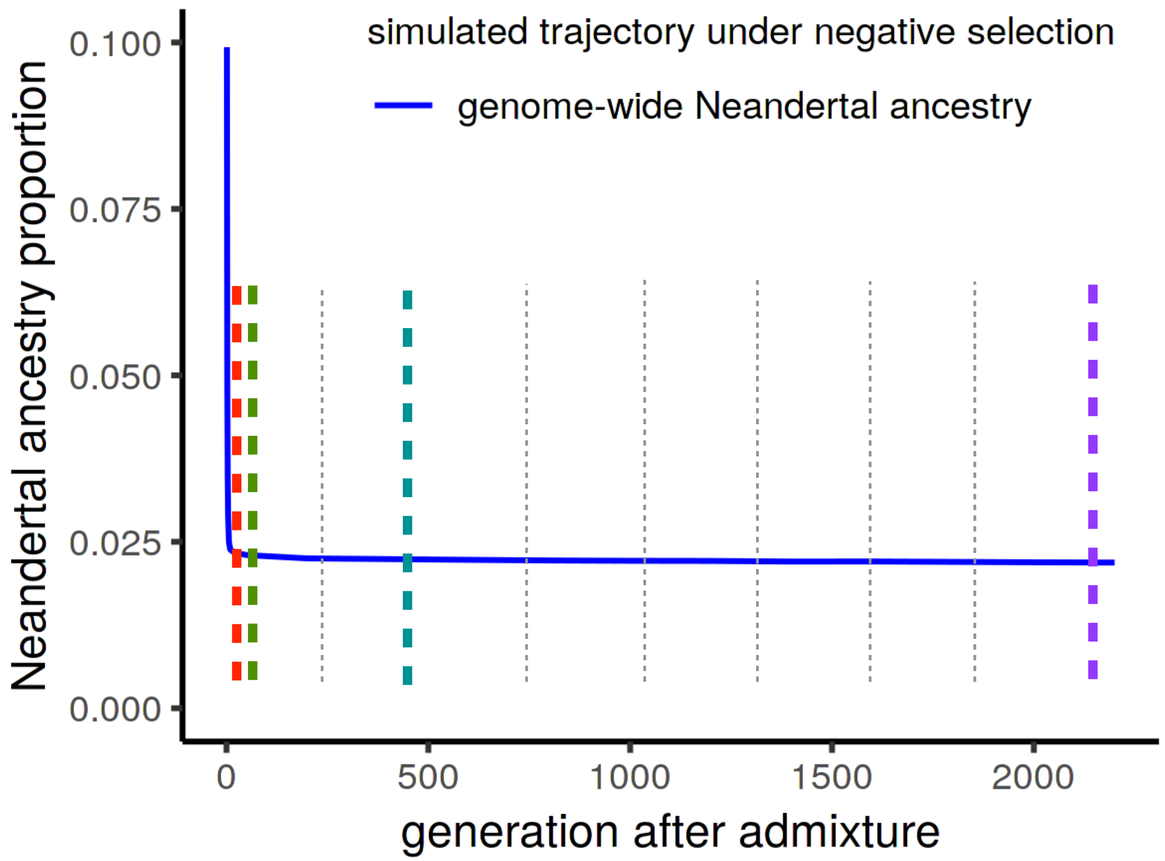
... what about the whole-genome negative selection signal?

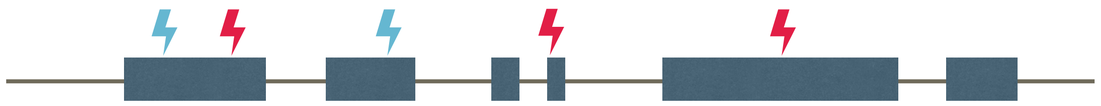




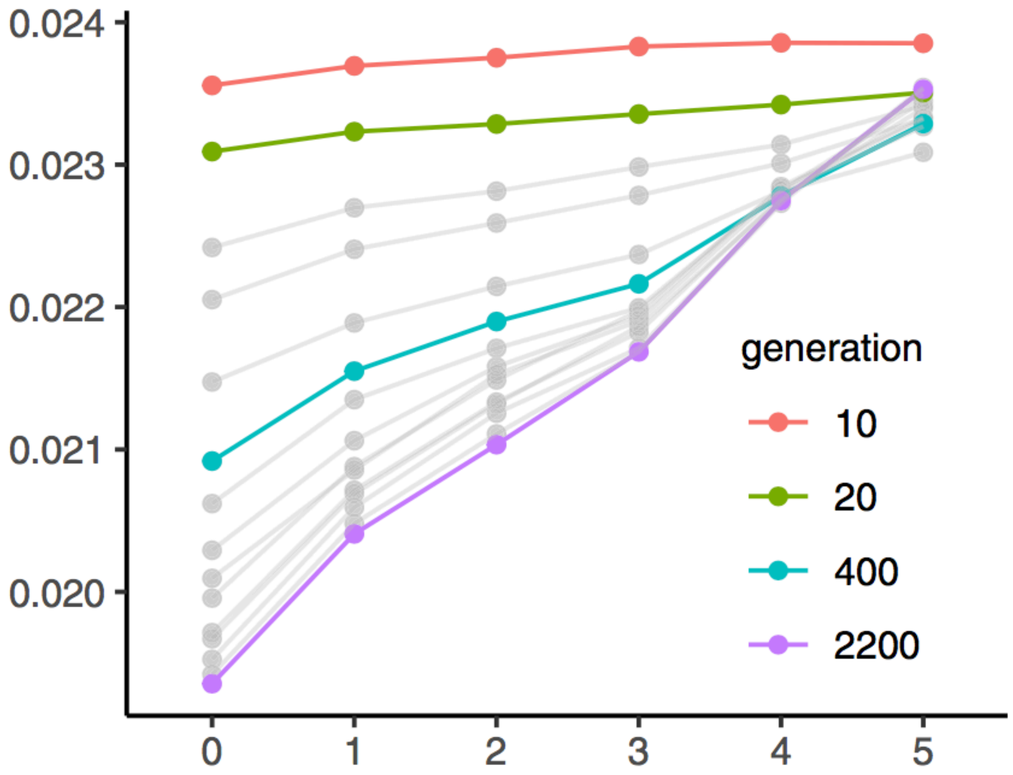
Petr *et al.* (PNAS, 2019)

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





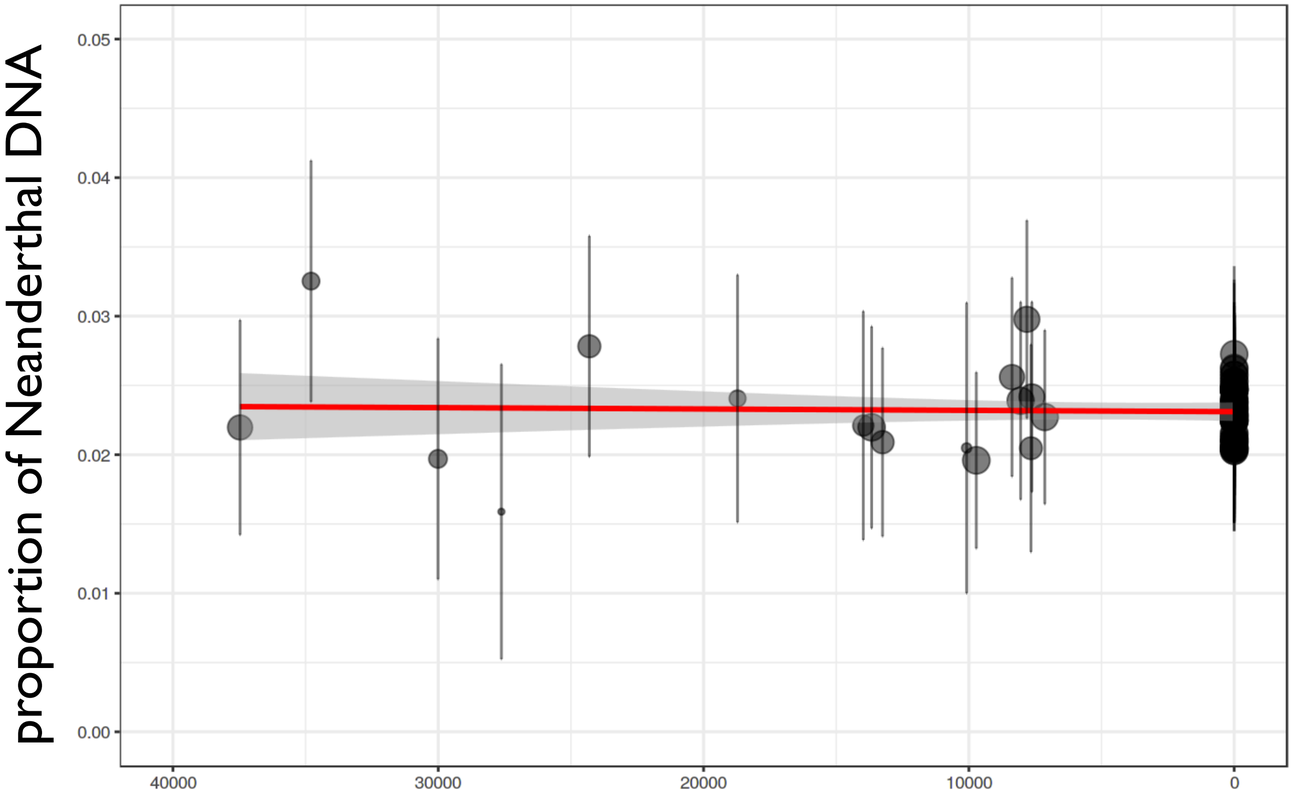
simulated data

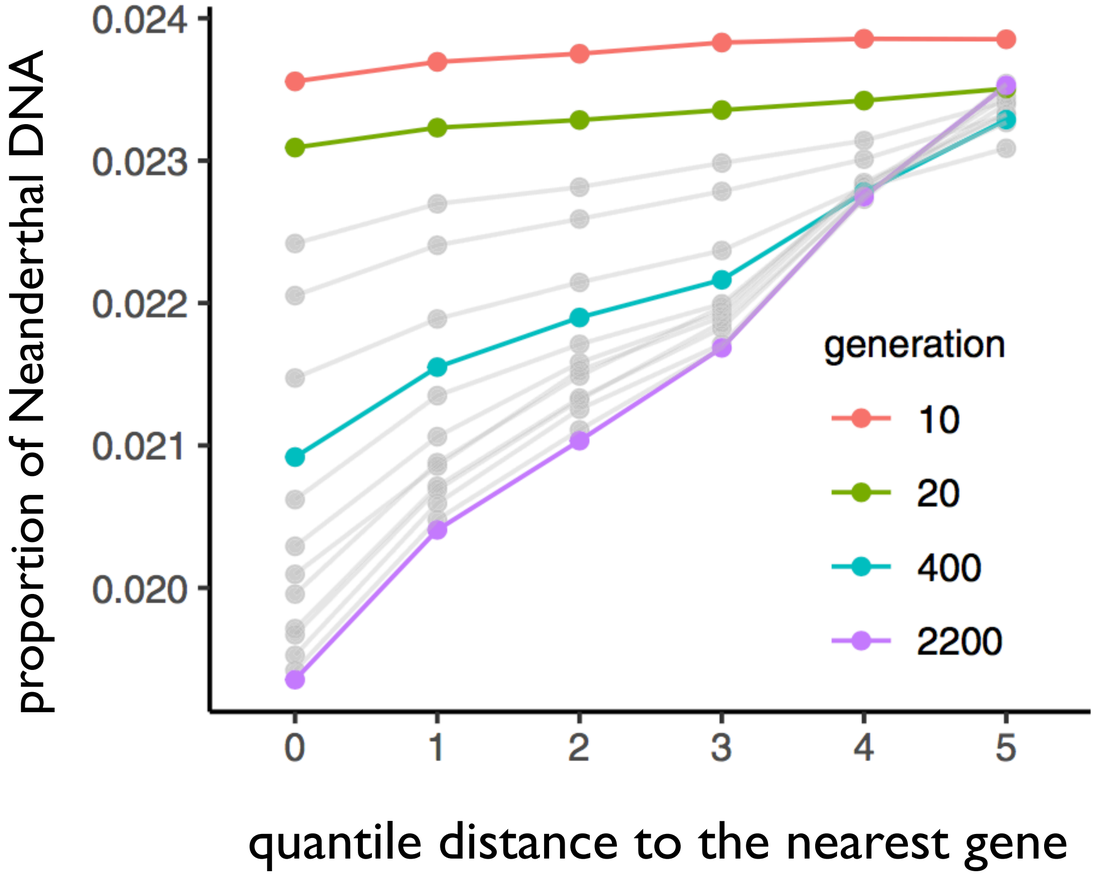


quantile distance to the nearest gene

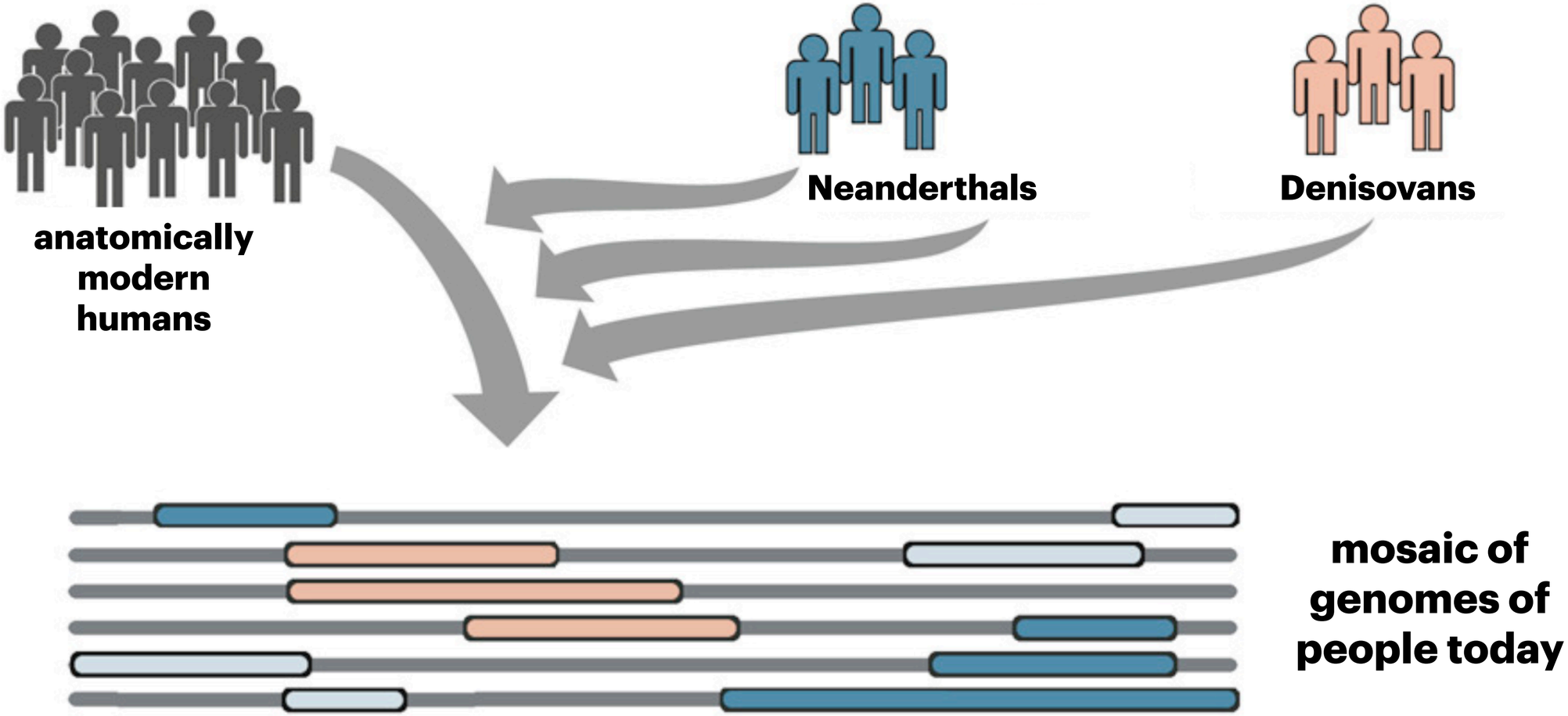
Petr *et al.* (PNAS, 2019)

Both empirical results match population genetic theory!

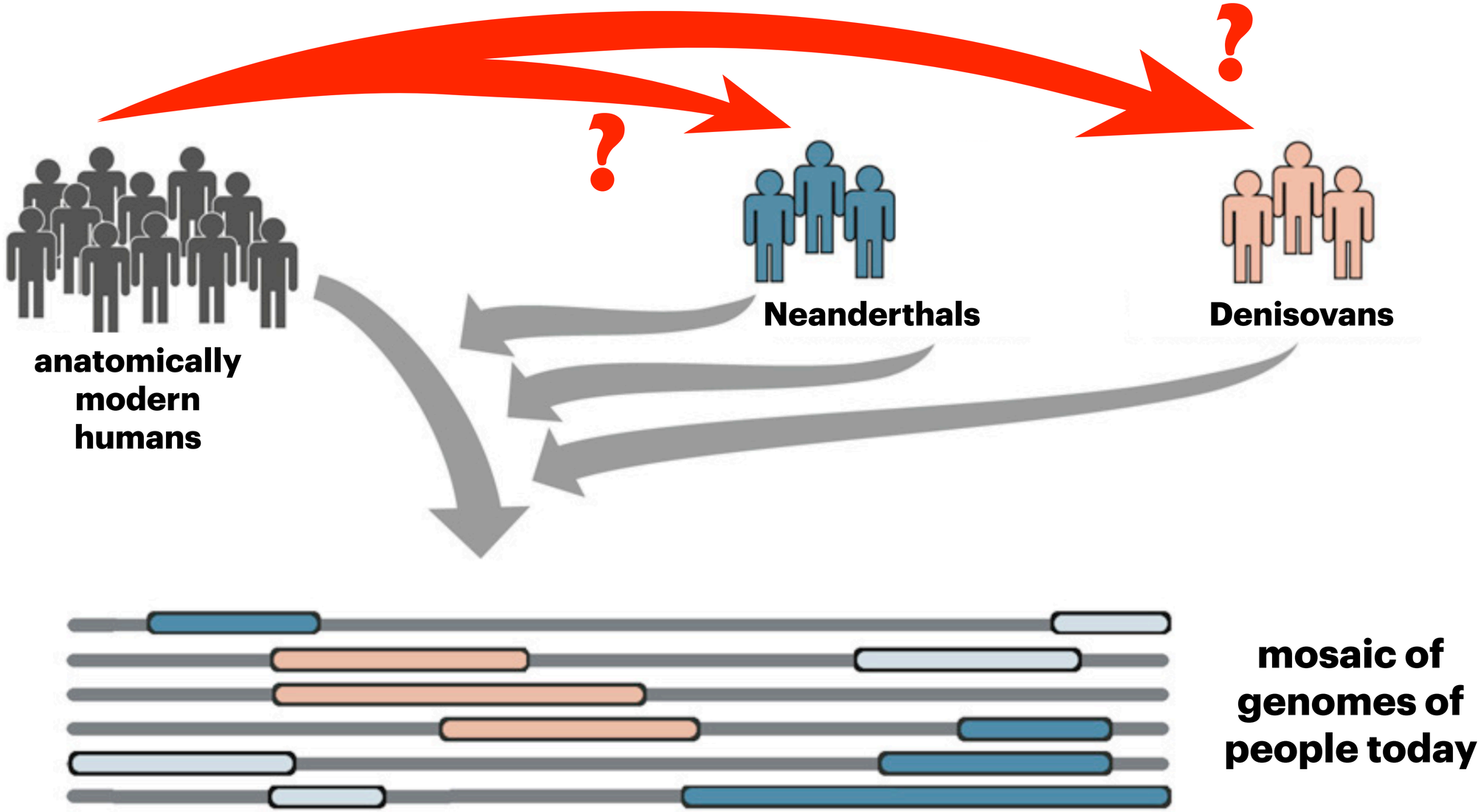




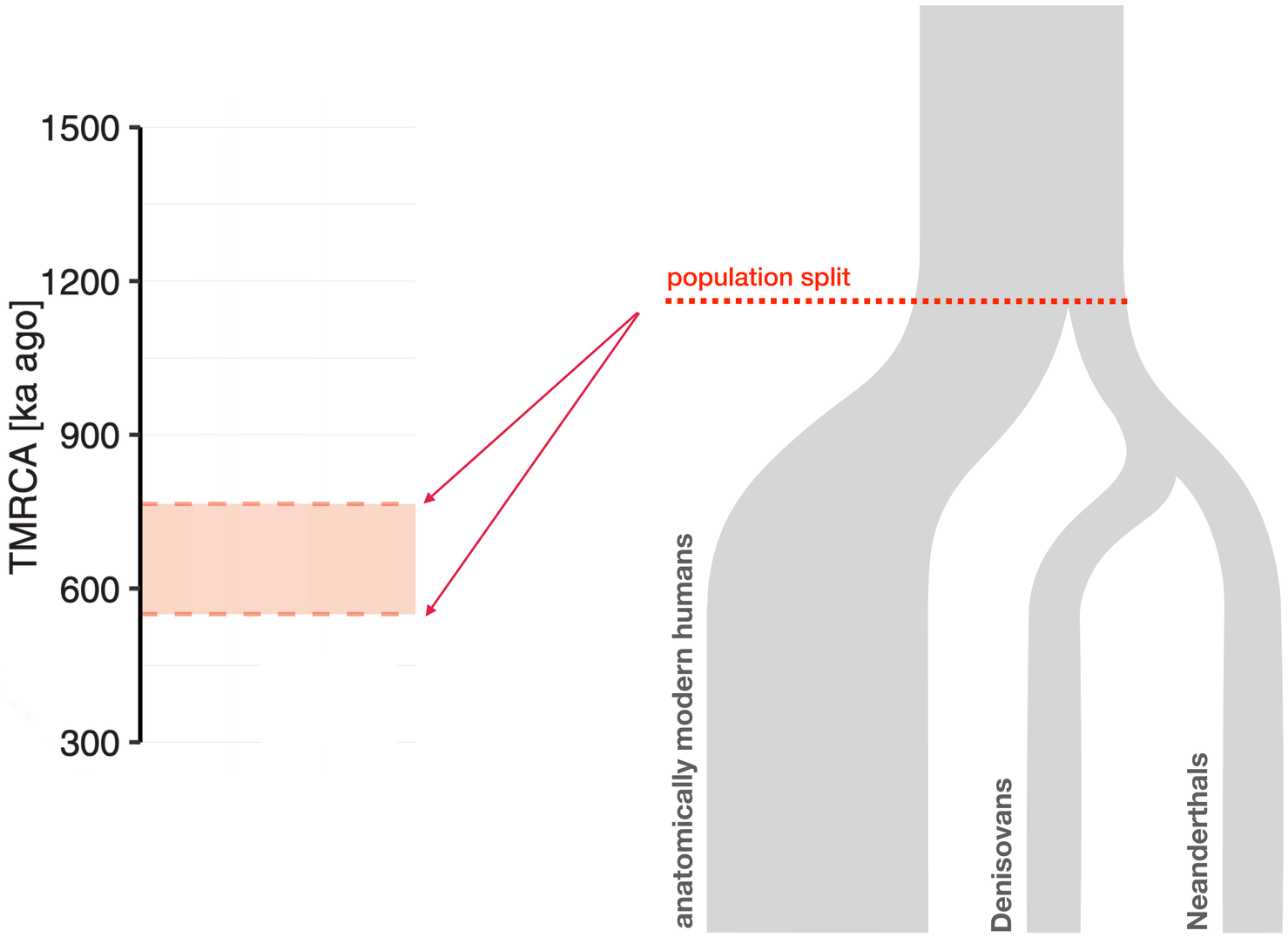
Petr *et al.* (PNAS, 2019)



adapted from Browning *et al.*, Cell, 2018

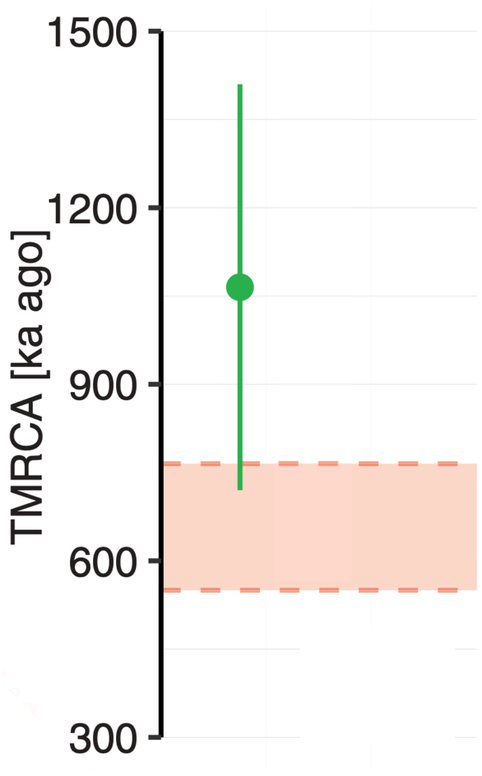


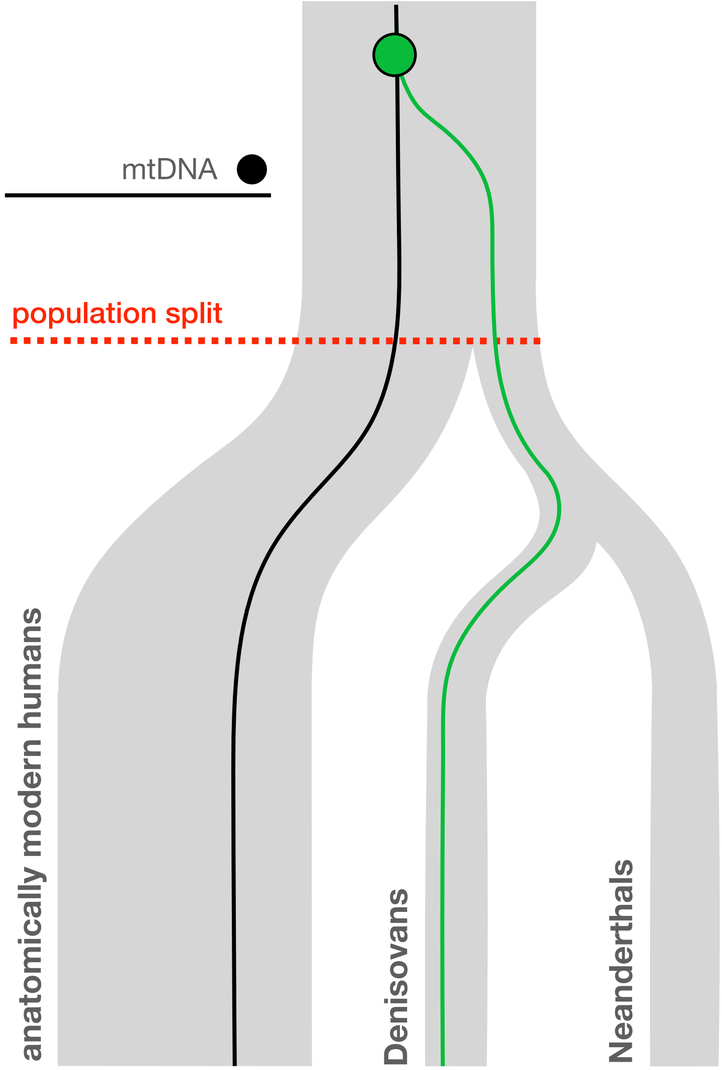
adapted from Browning *et al.*, Cell, 2018



Petr *et al*. (Science, 2020)

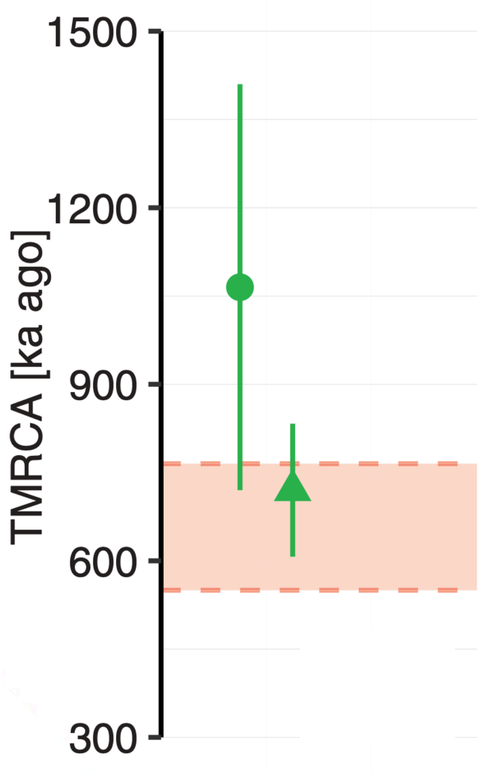
Denisovan mtDNA follows known population history

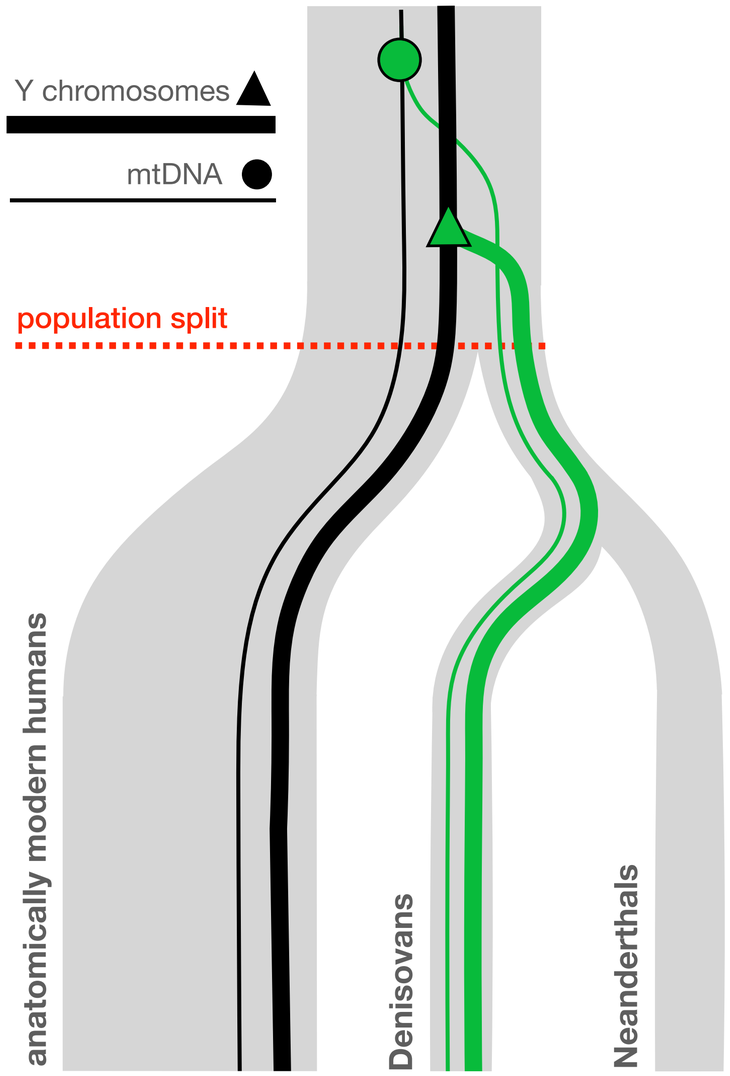




Petr *et al*. (Science, 2020)

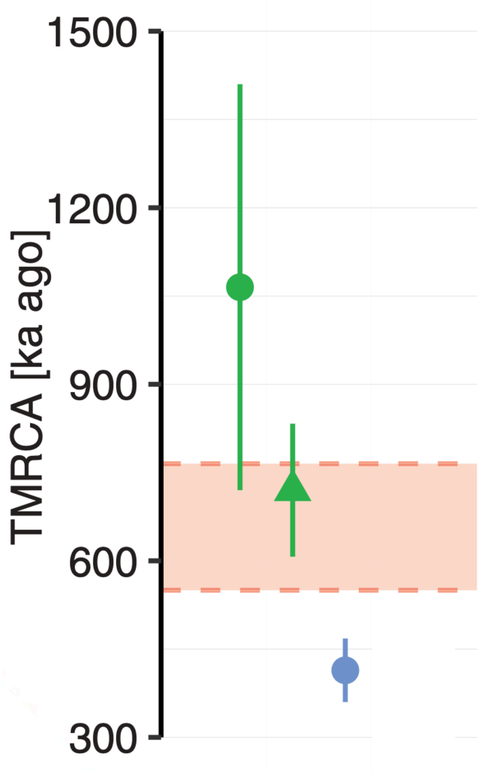
Denisovan mtDNA follows known population history

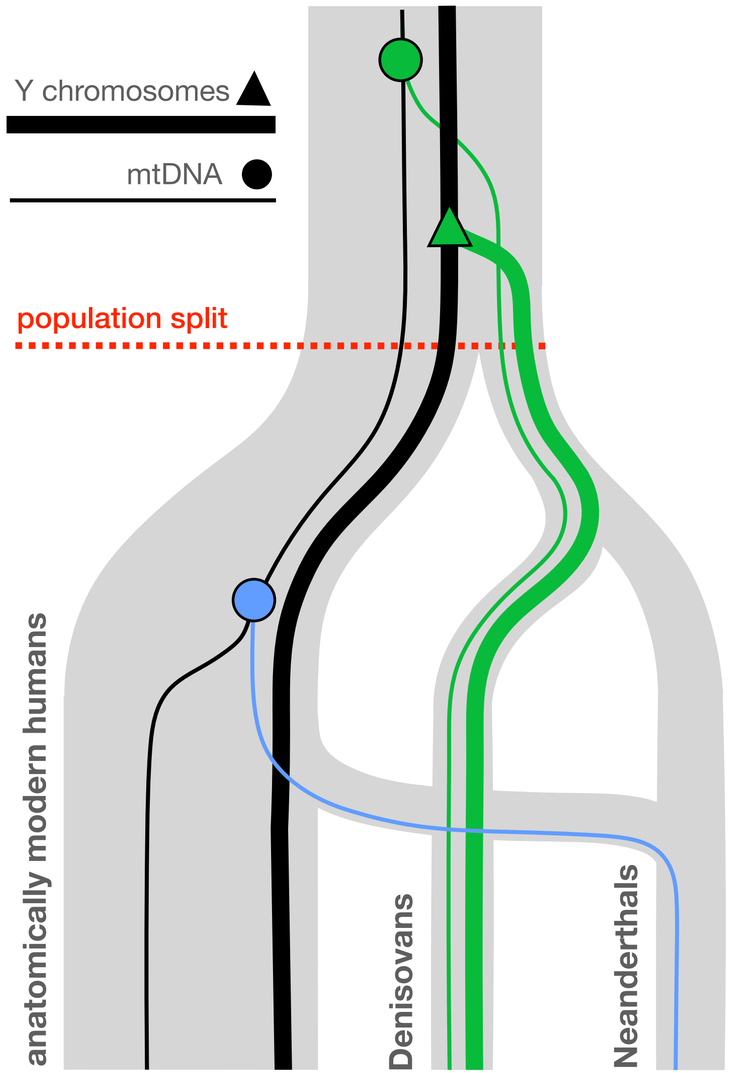




Petr *et al*. (Science, 2020)

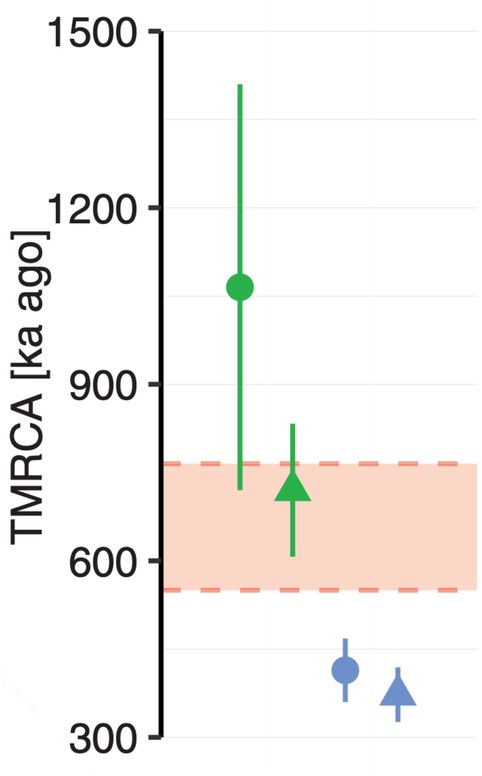
Neanderthal mitochondrial DNAs contradict this!

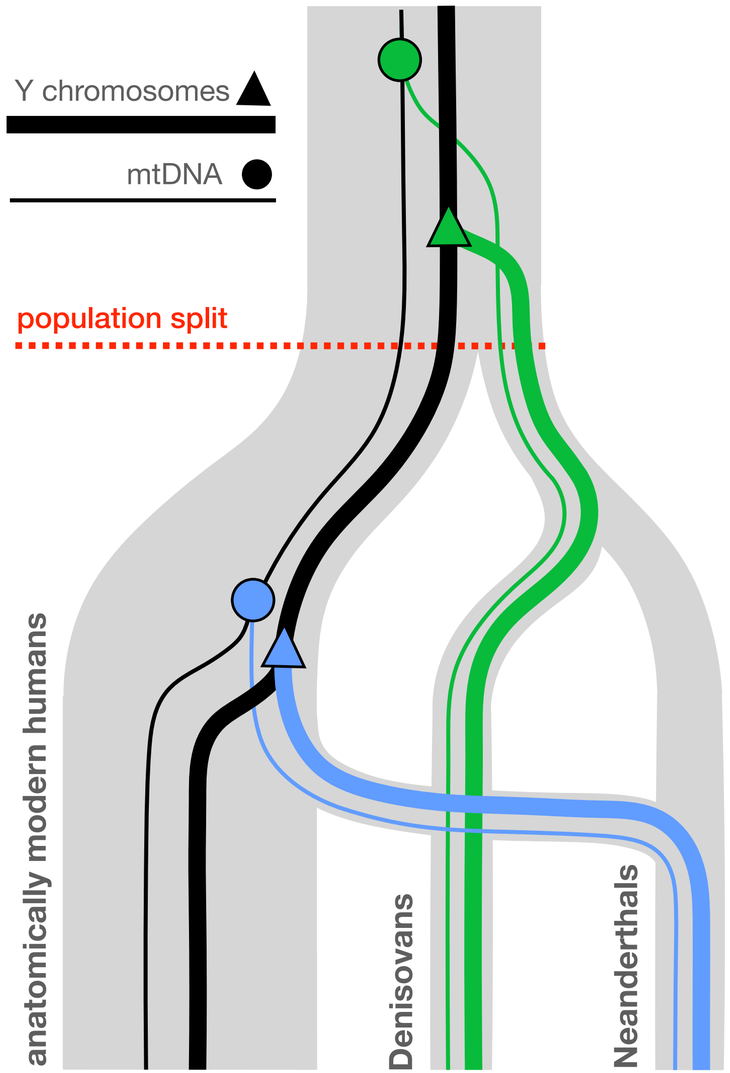




Petr *et al*. (Science, 2020)

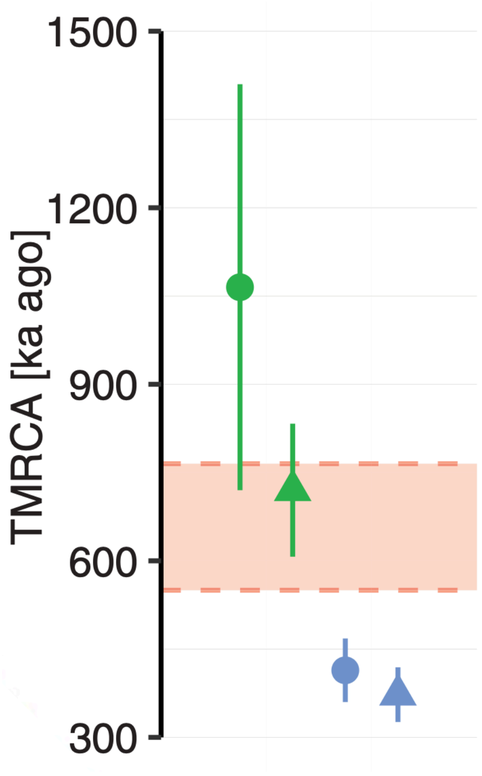
Neanderthal Y chromosomes *also* contradict this!

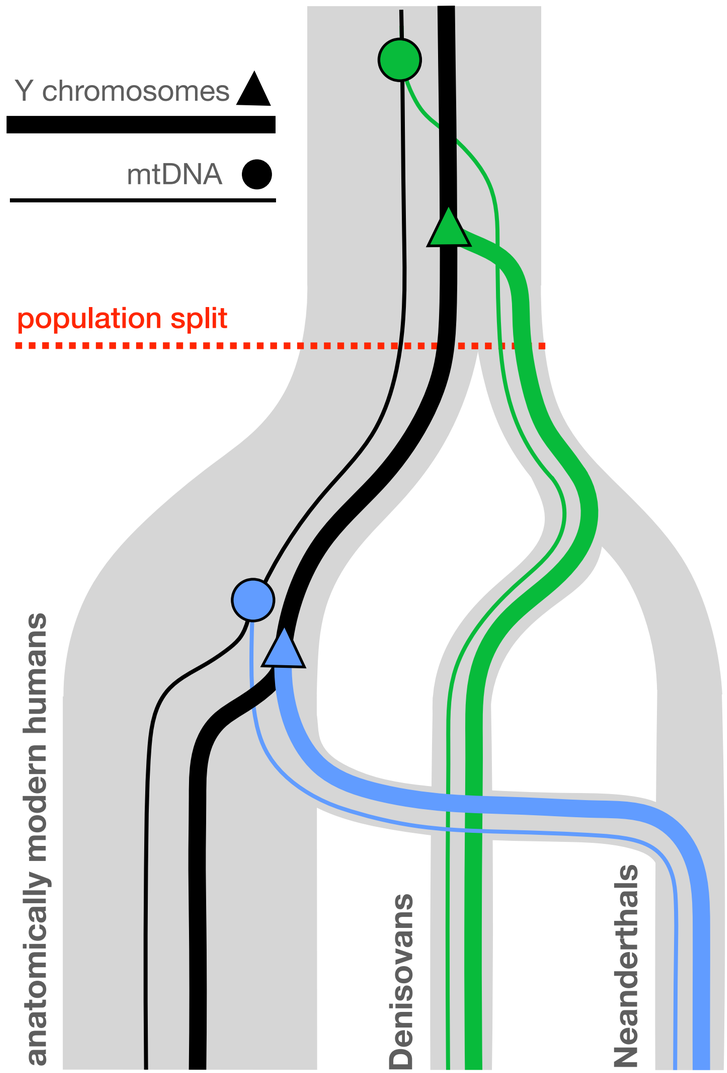




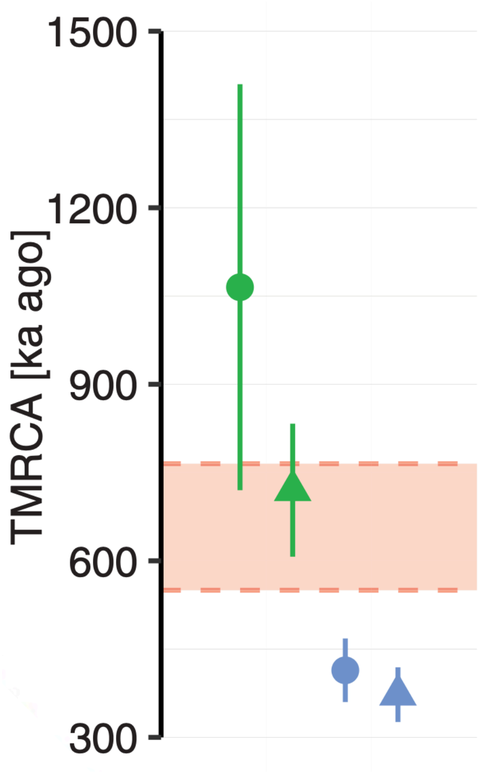
Petr *et al*. (Science, 2020)

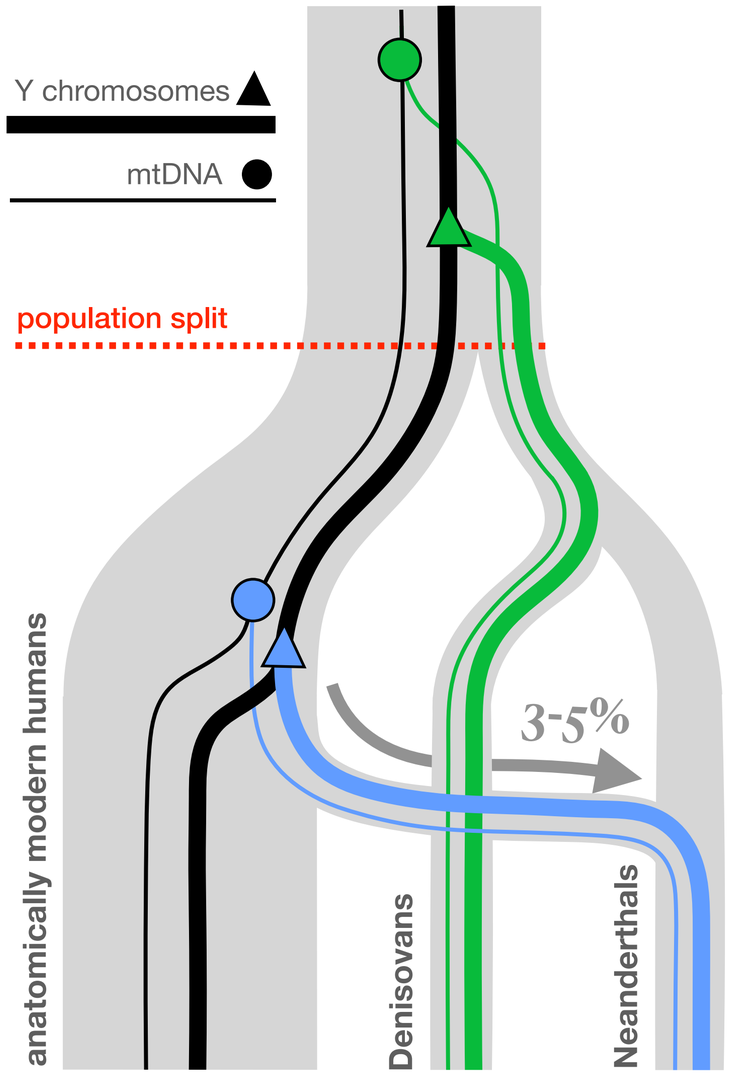
Original Neanderthal mtDNA/Y experienced replacement





Original Neanderthal mtDNA/Y experienced replacement

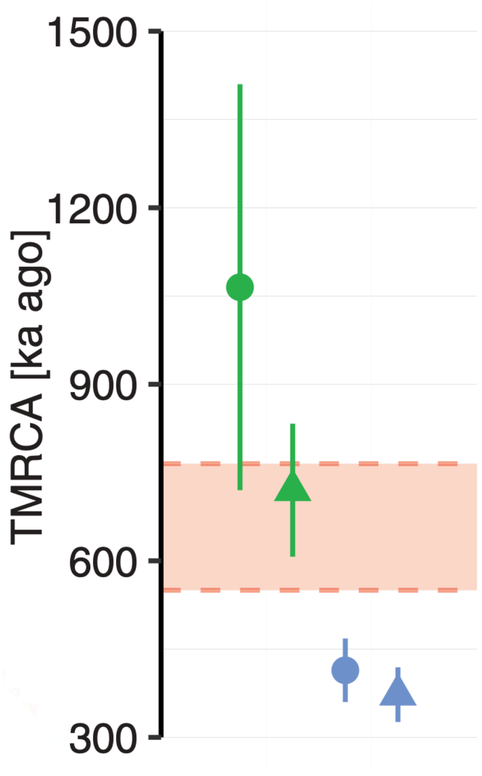


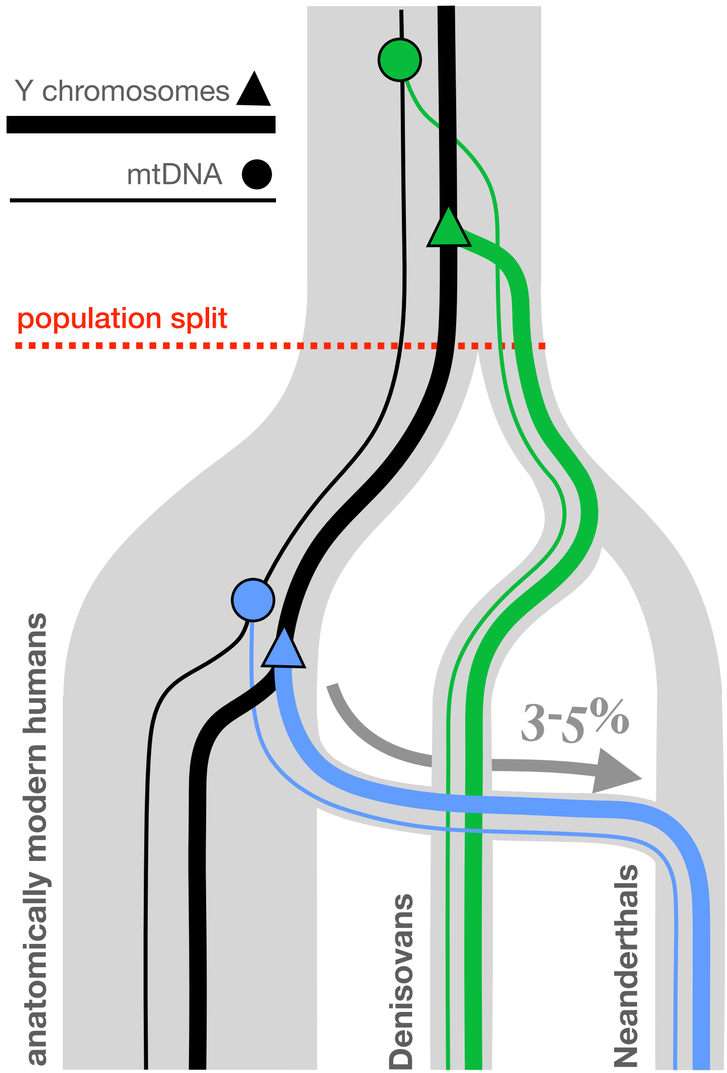


Hubisz *et al.* (PLOS Genetics, 2020)

Original Neanderthal mtDNA/Y experienced replacement

(very unlikely under neutrality!)

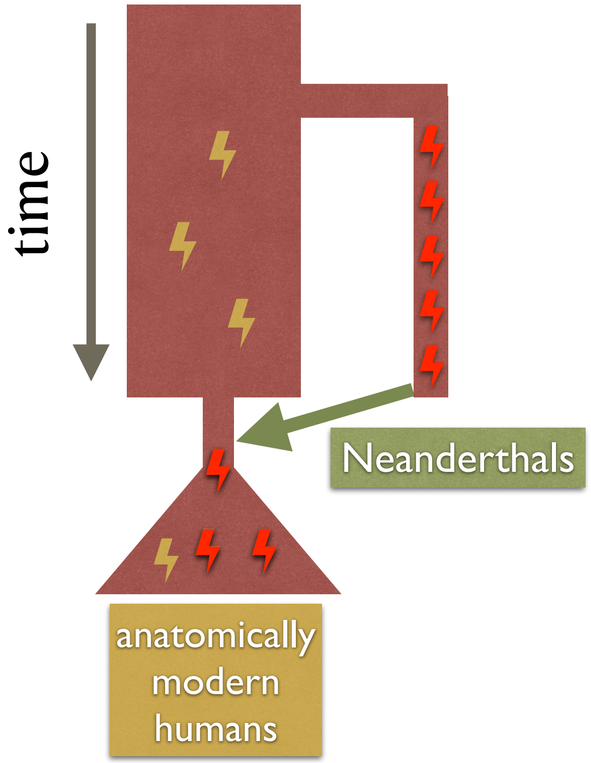


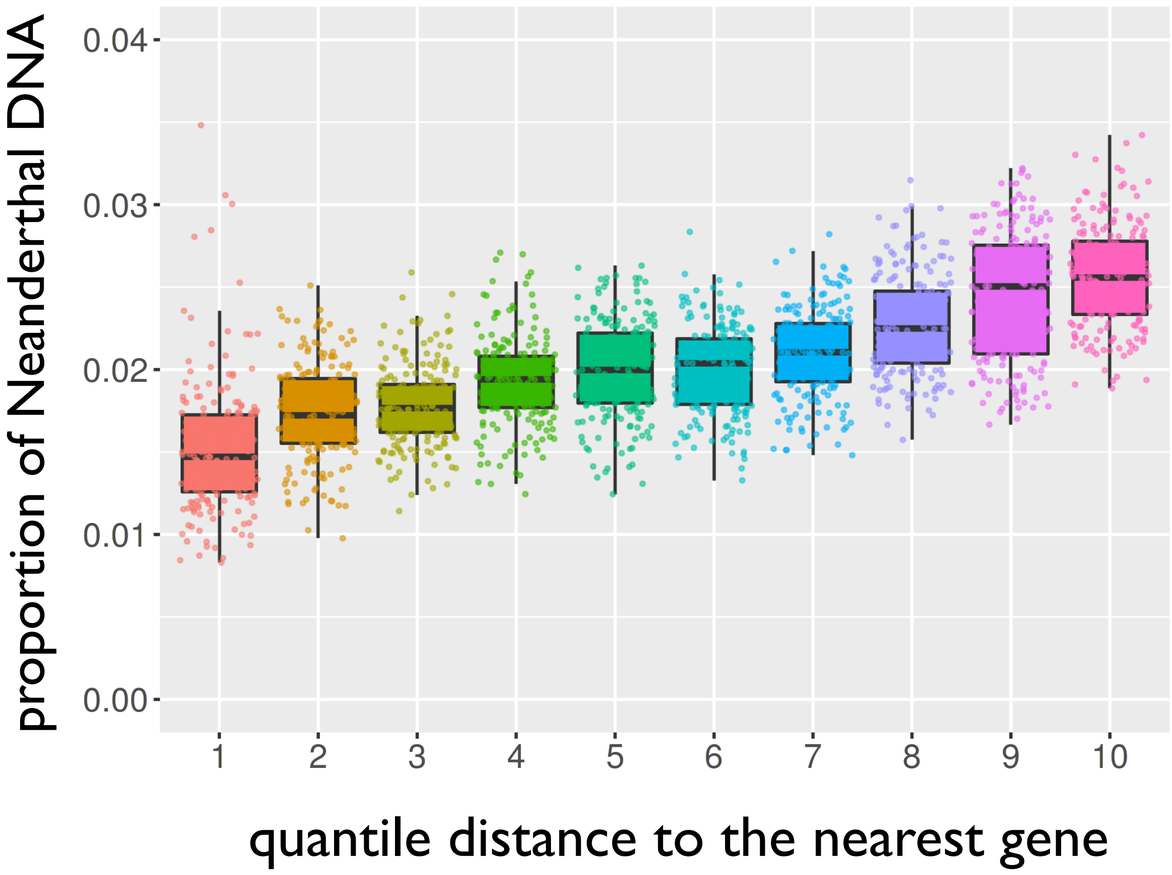


Hubisz *et al.* (PLOS Genetics, 2020)

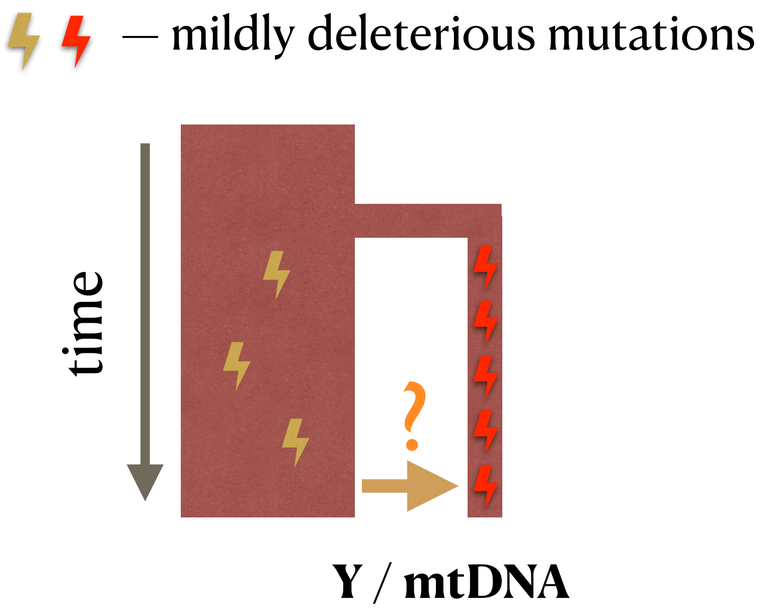
... but we know introgression was not neutral!

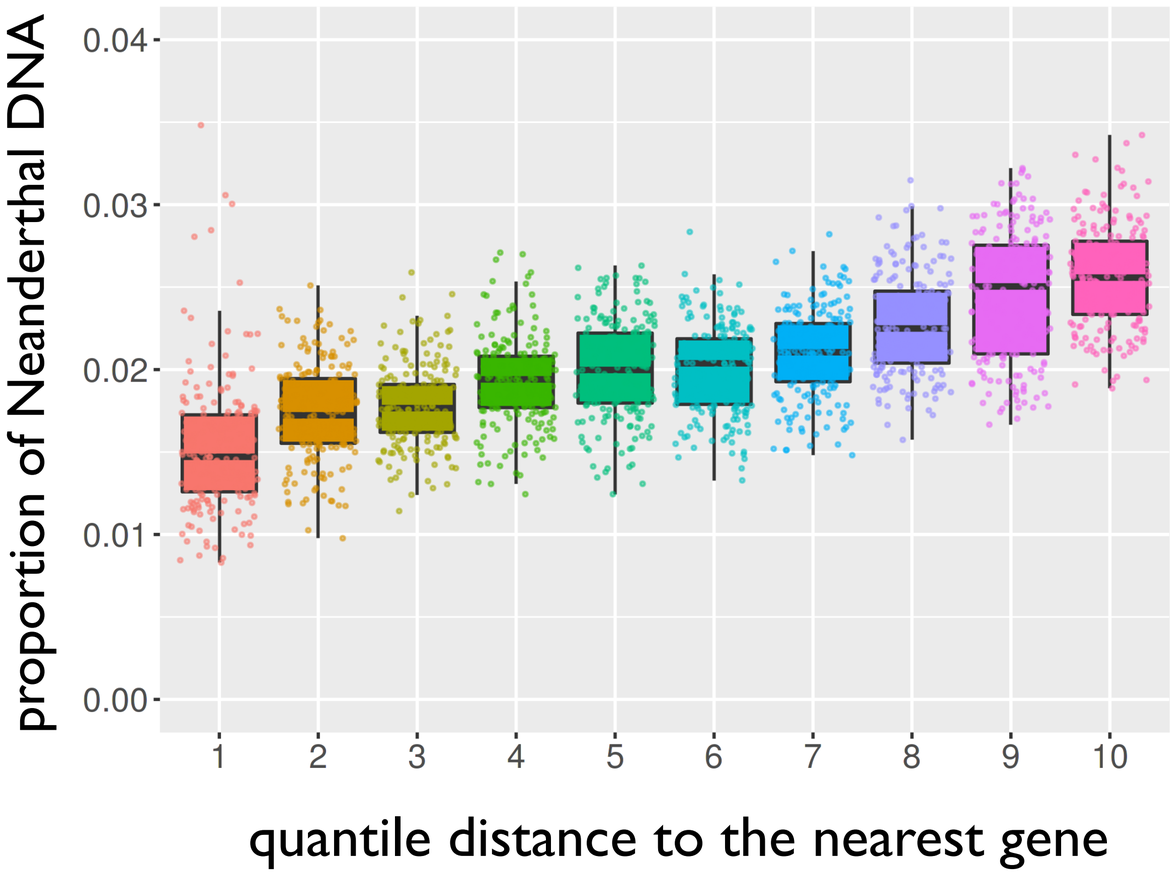
— mildly deleterious mutations



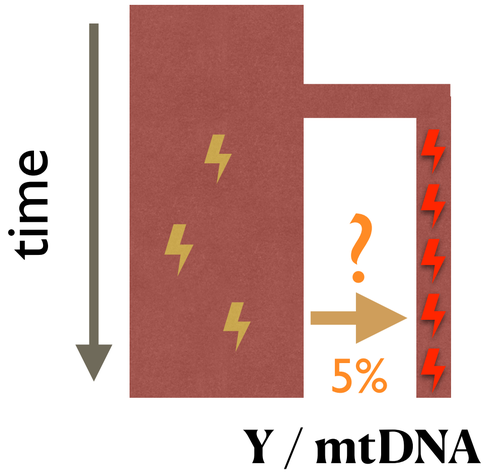


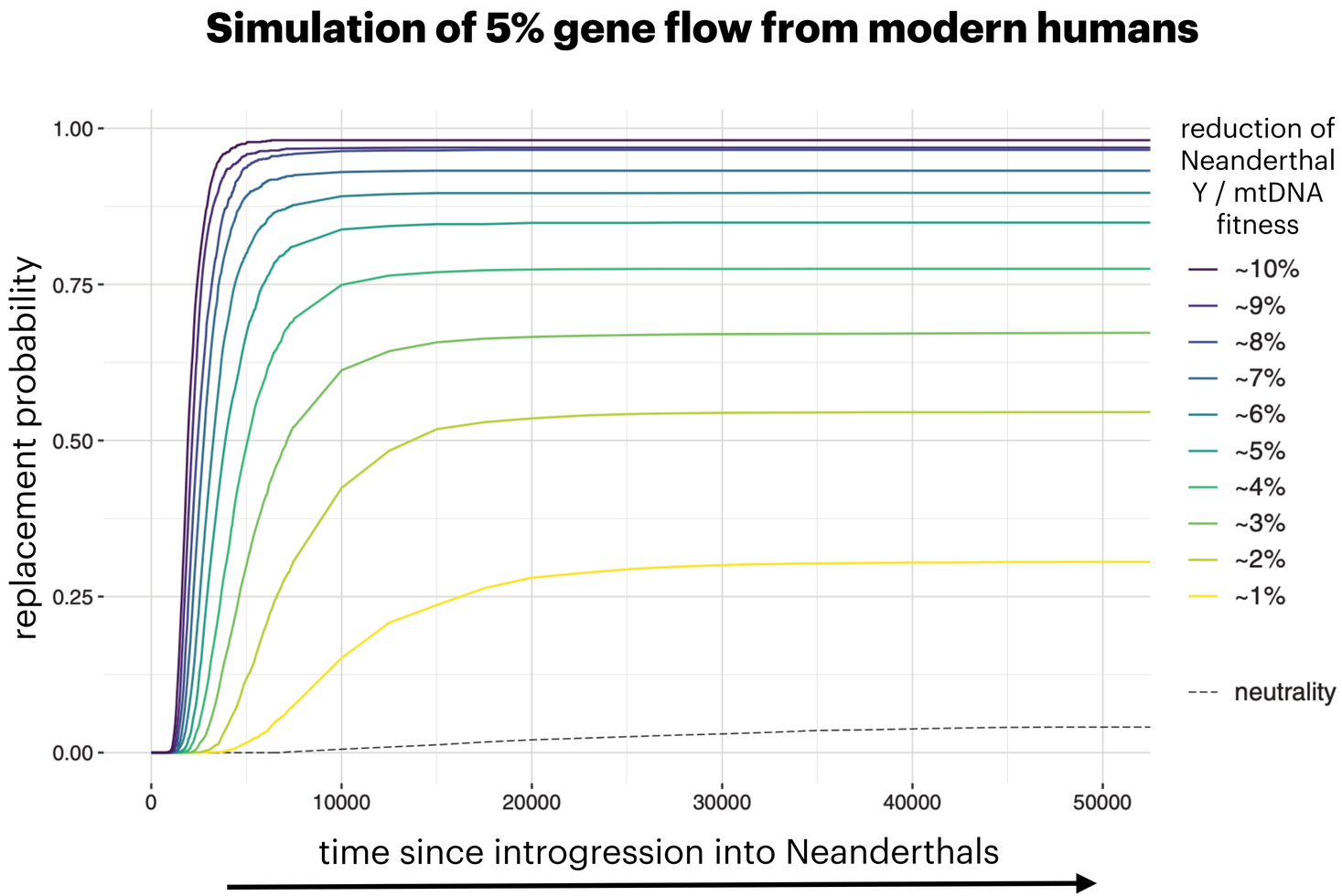
What is the expected scenario for Y / mtDNA introgression?





Natural selection in Neanderthals favors modern human Y / mtDNA





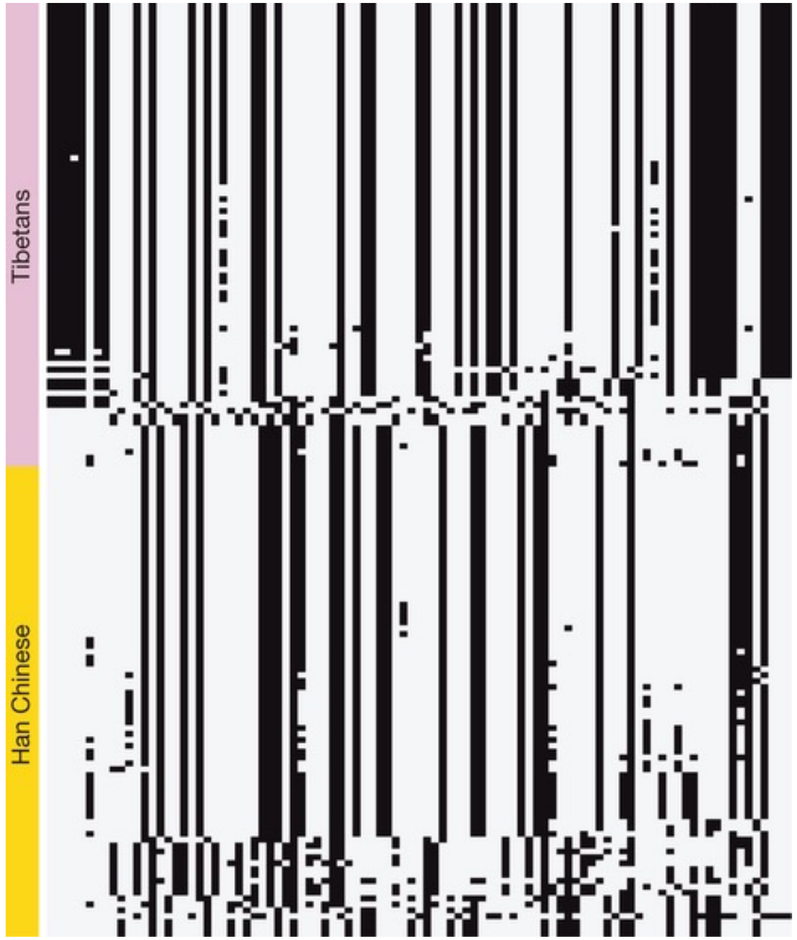
Petr *et al*. (Science, 2020)

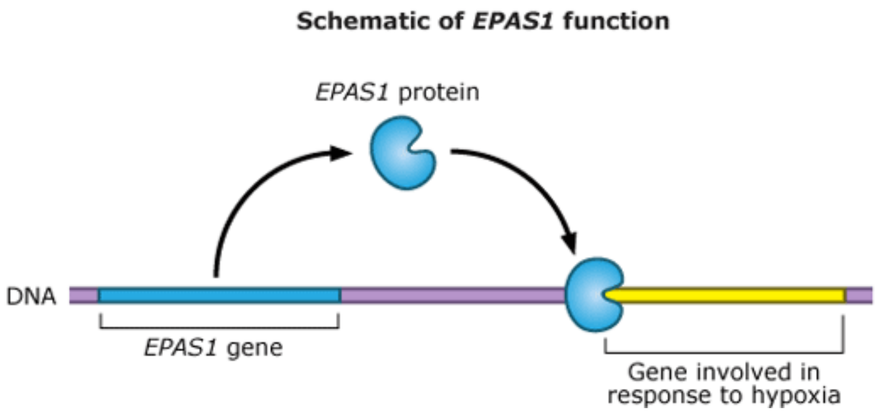
It wasn’t all bad though

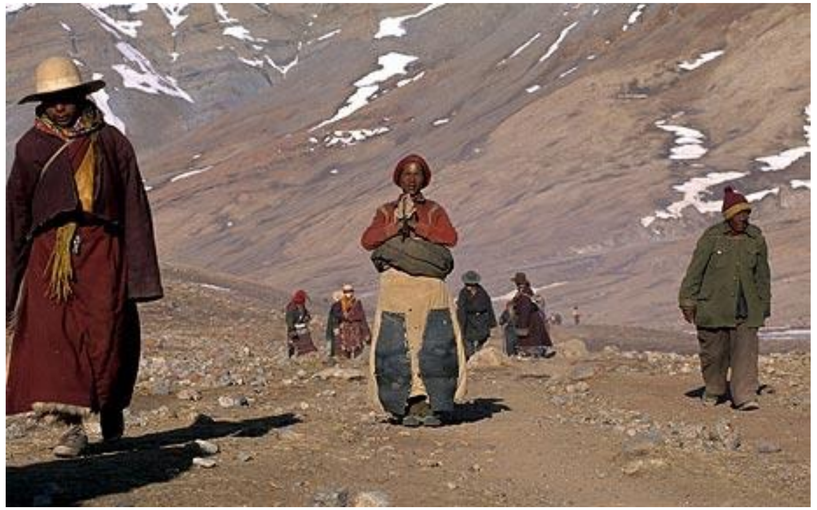
positive selection / adaptive introgression

High altitude adaptation in Tibetans

haplotype structure (derived SNPs)



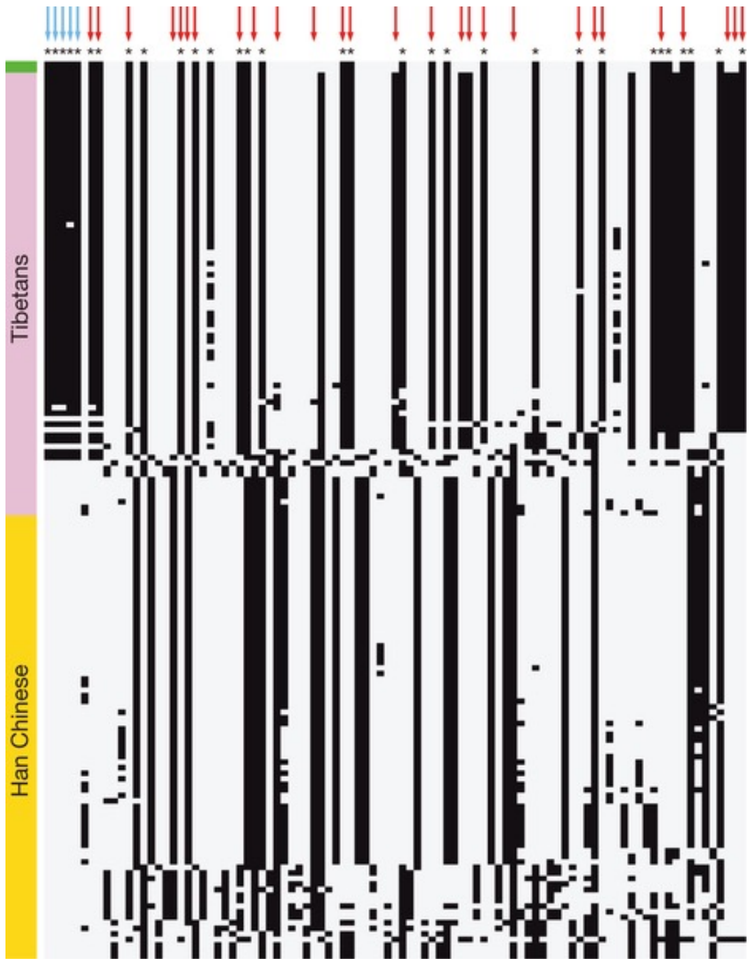




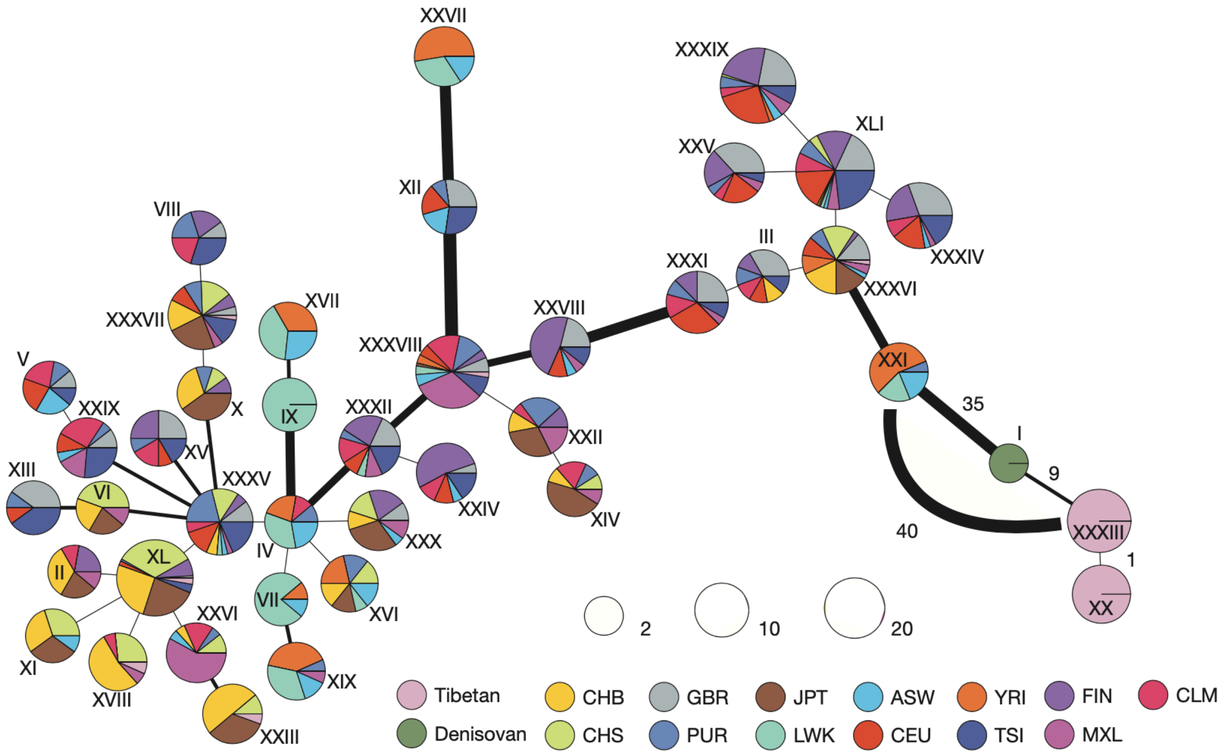
Huerta-Sánchez *et al.* (Nature, 2014)

High altitude adaptation via Denisovan introgression!

haplotype structure (derived SNPs)

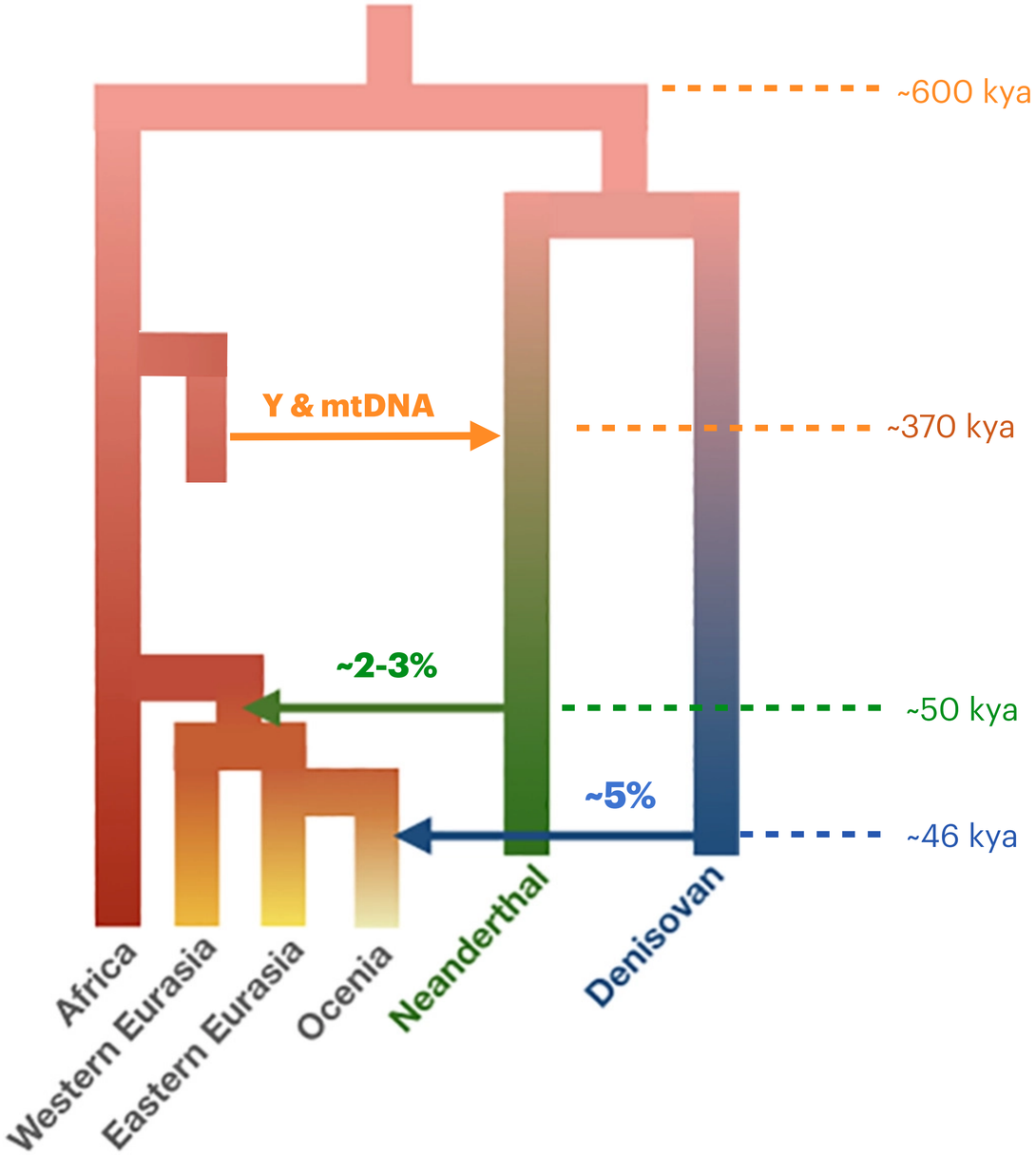


haplotype network based on pairwise-differences



Huerta-Sánchez *et al.* (Nature, 2014)

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

Denisova 11

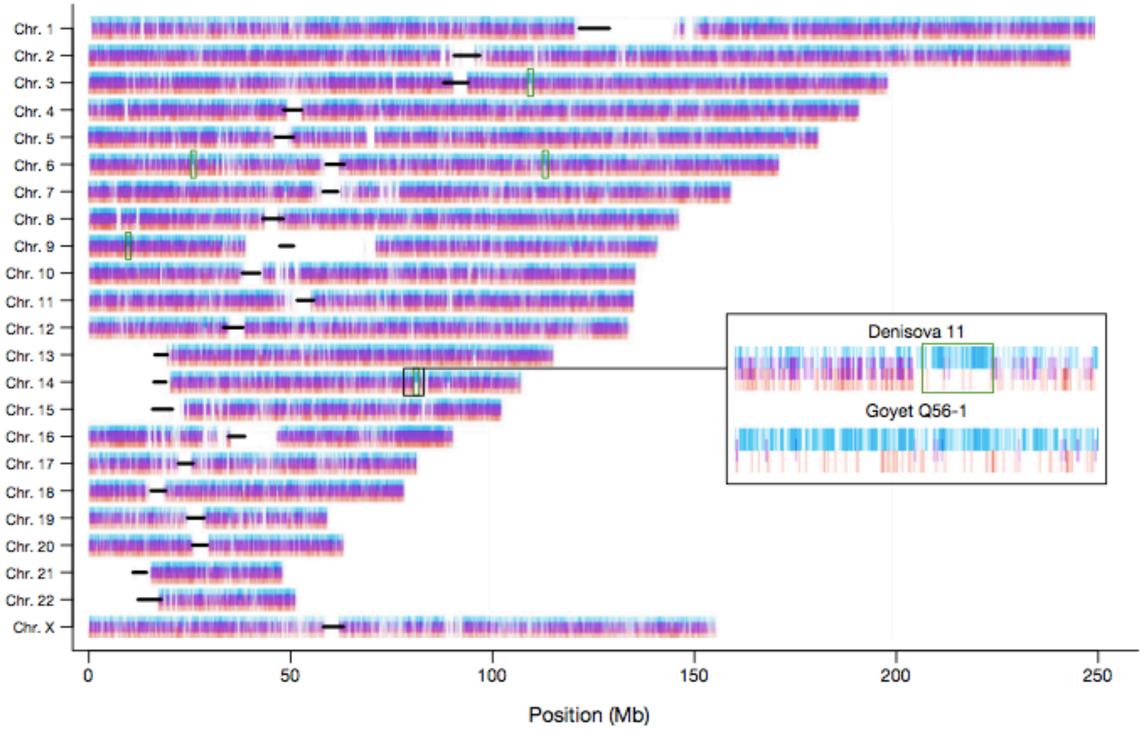
By Демин Алексей Барнаул - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=48890364

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

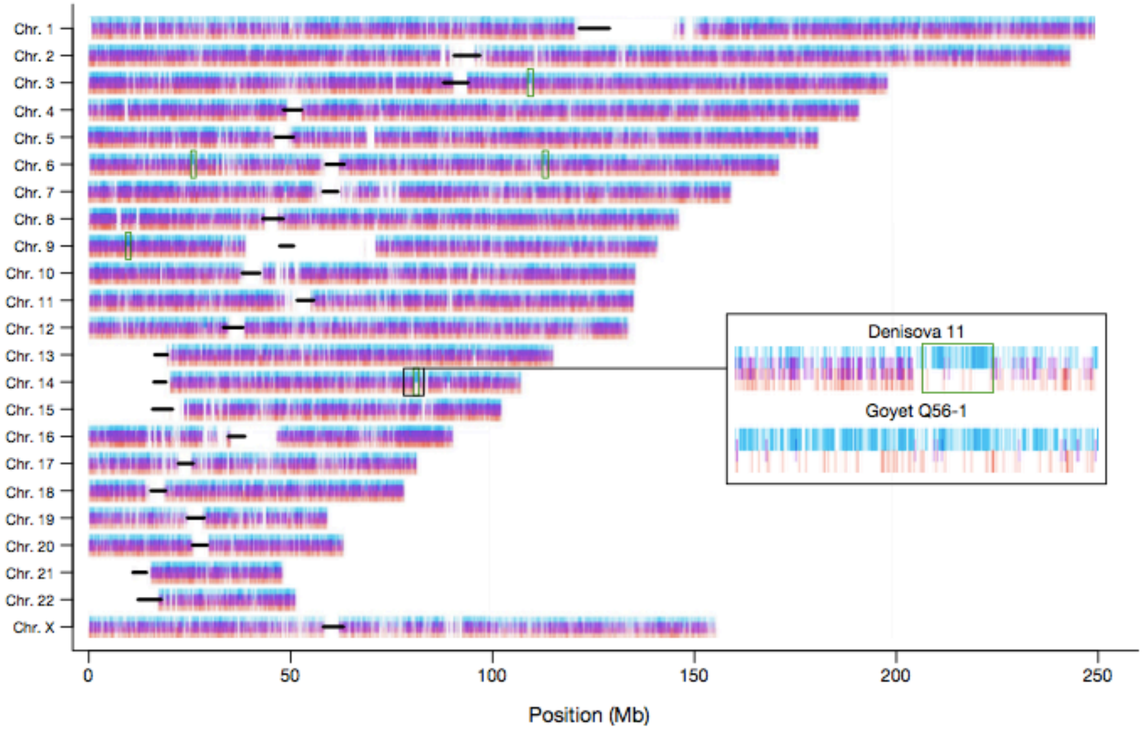
Denisova 11

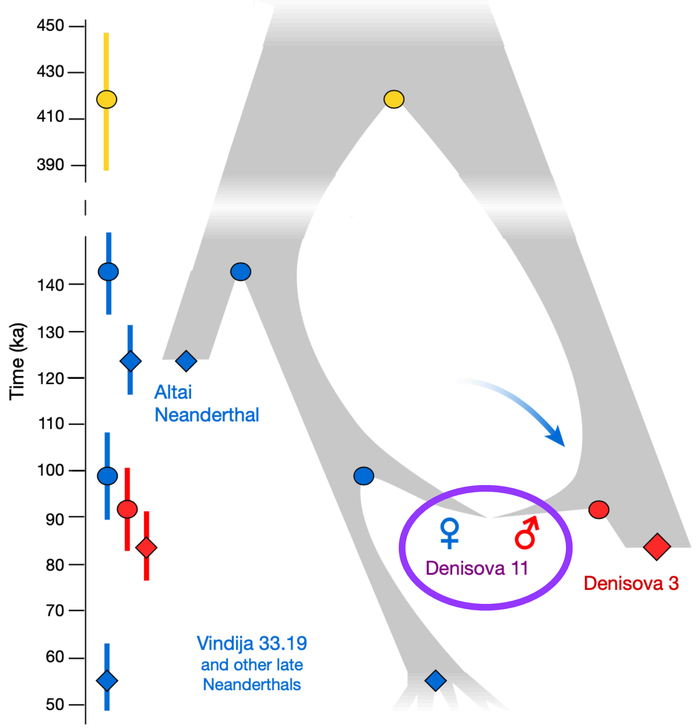


Letter | Published: 22 August 2018

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

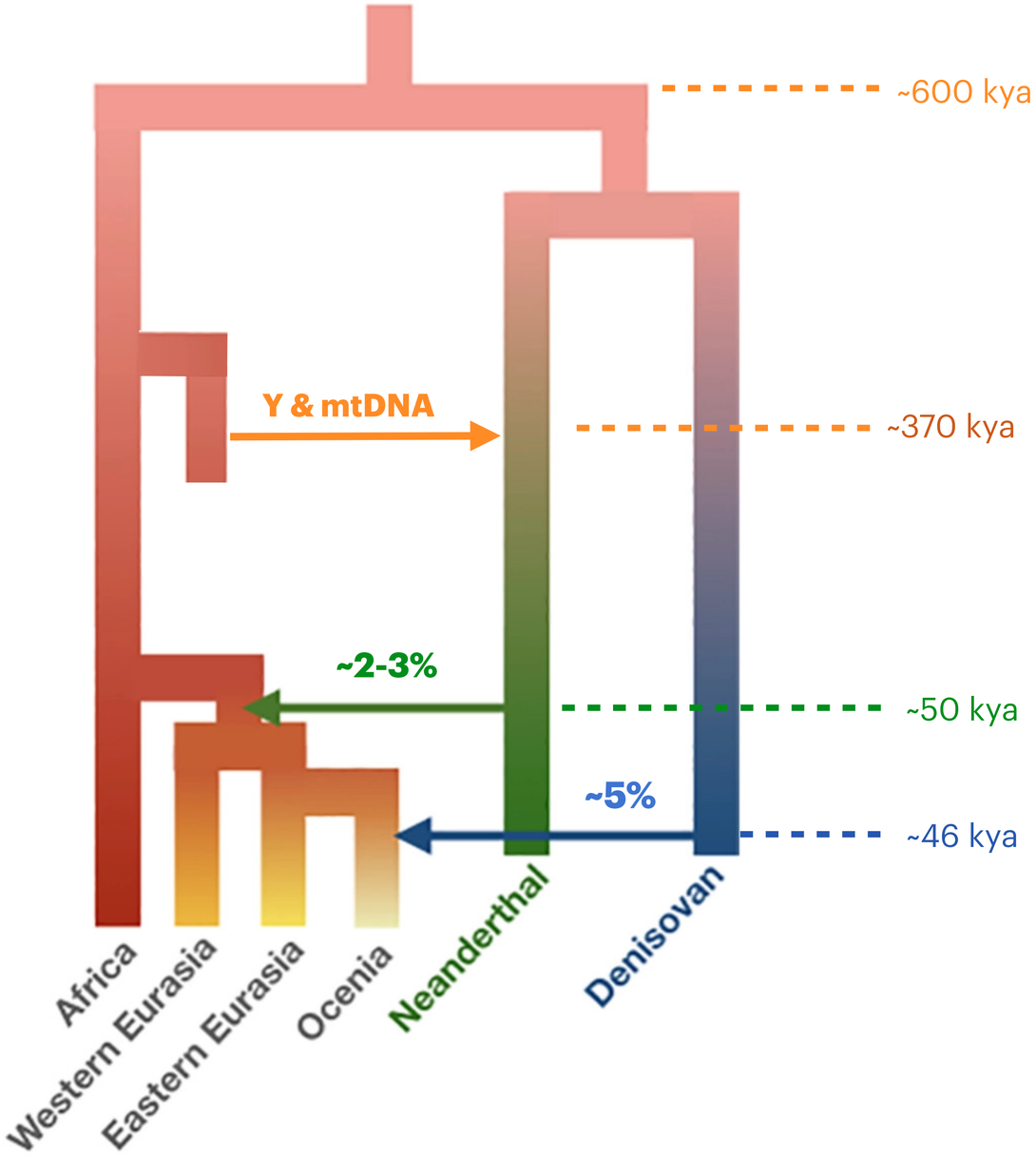




Introgression has been very frequent event in human history.

Introgressed DNA has had significant impact on the biology of "recipient populations".

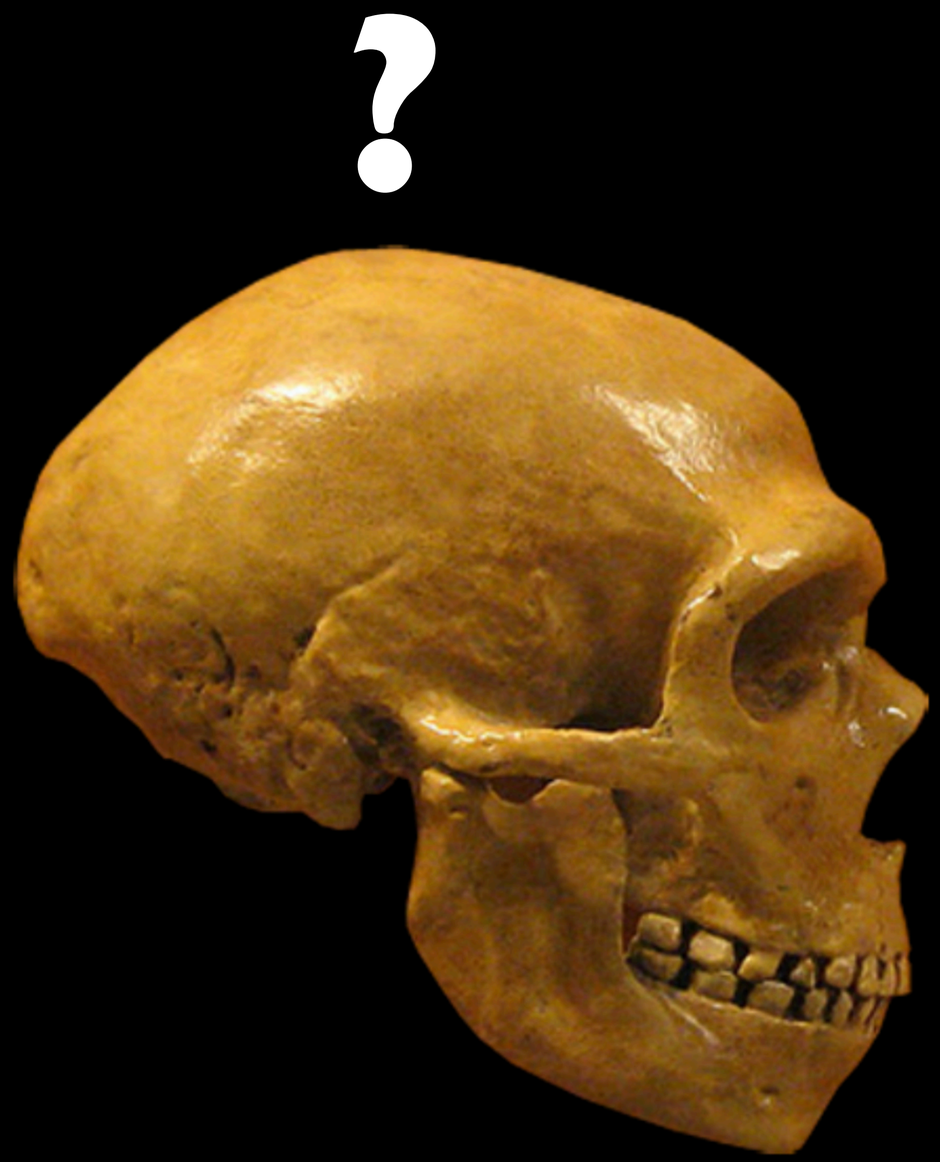
Population genetic simulations critical for testing evolutionary hypotheses.



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

?

?



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

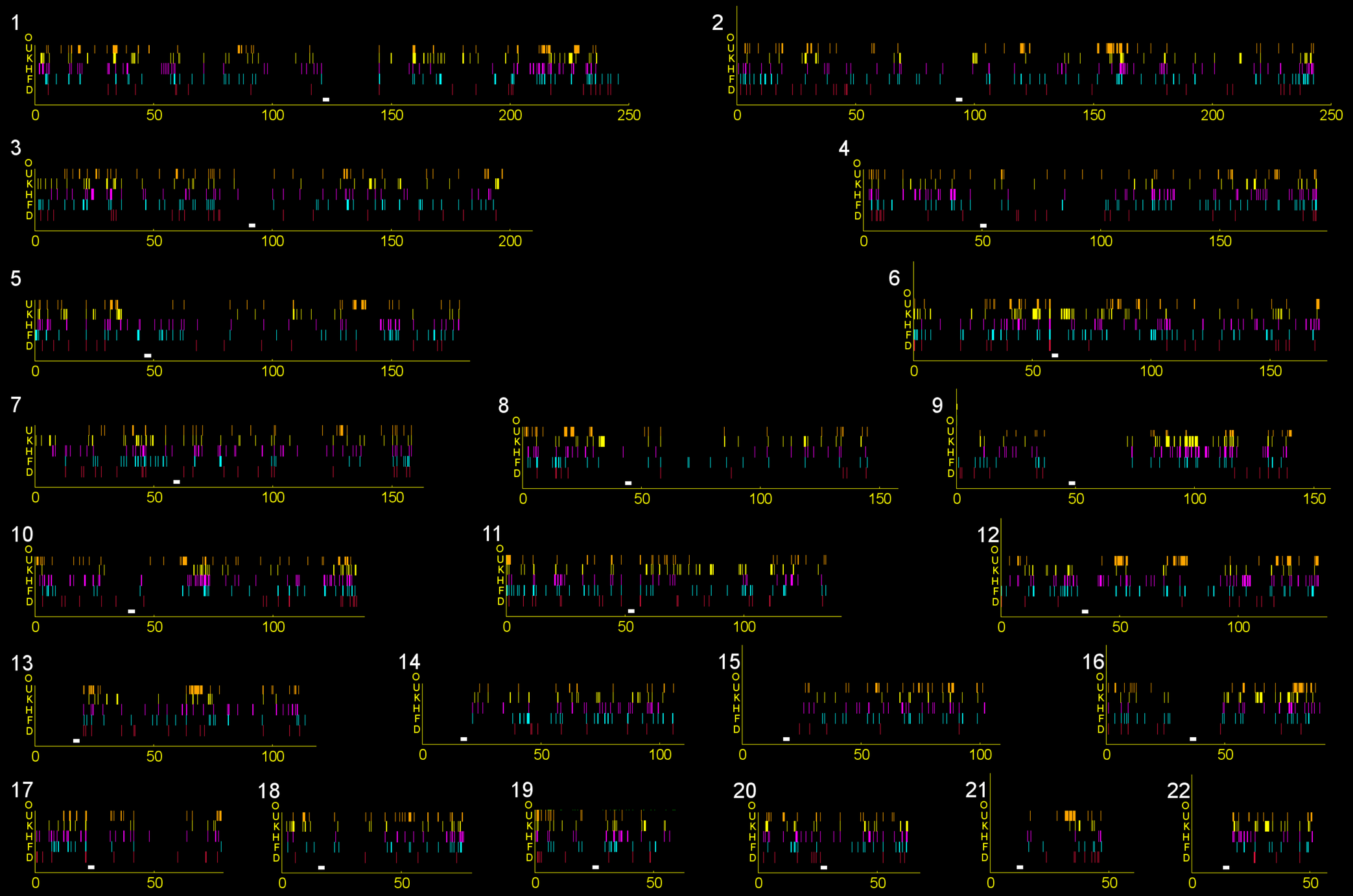
Morphological evidence?

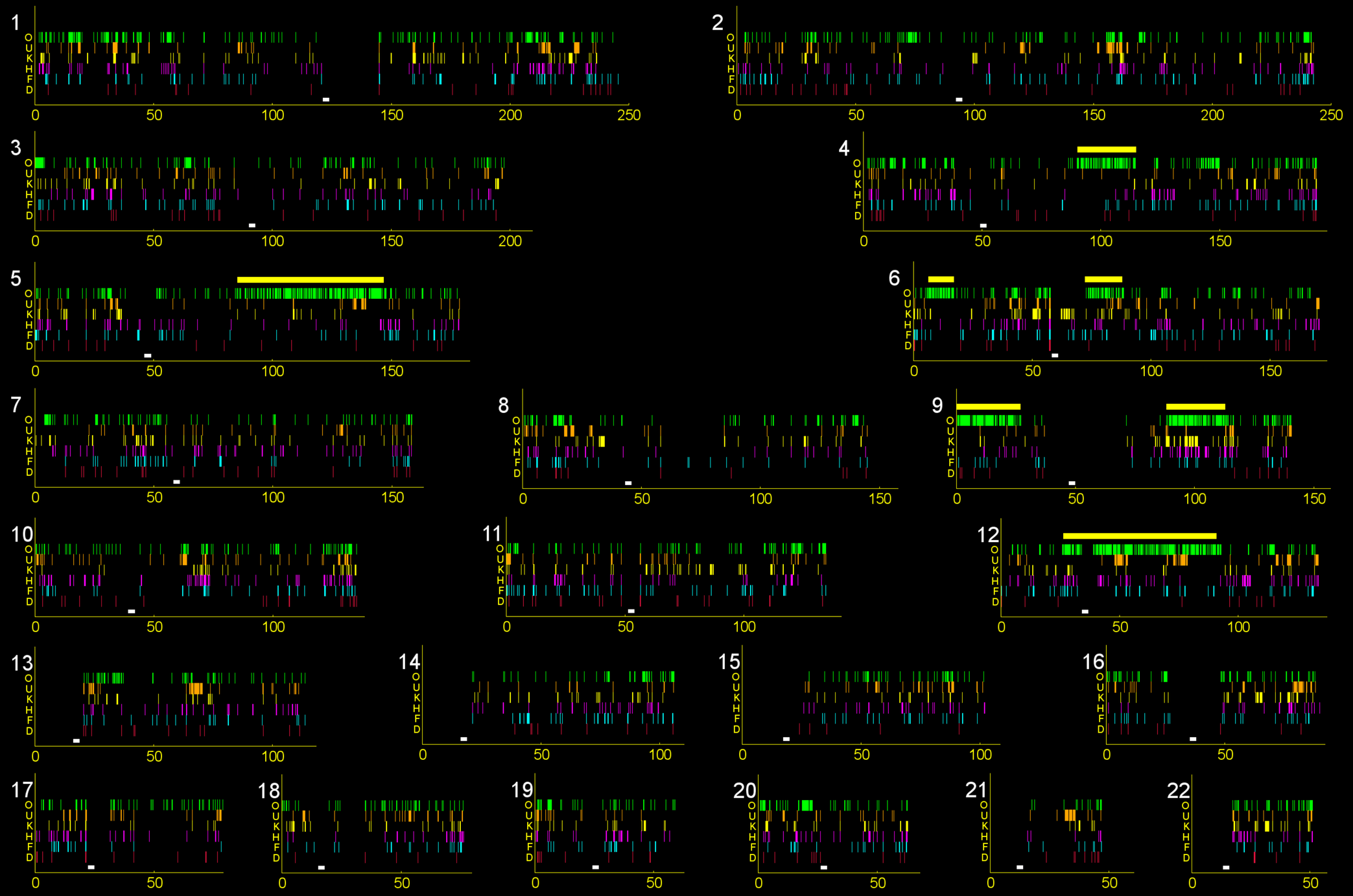
Trinkaus *et al*. PNAS, 2003

~40 thousand years old

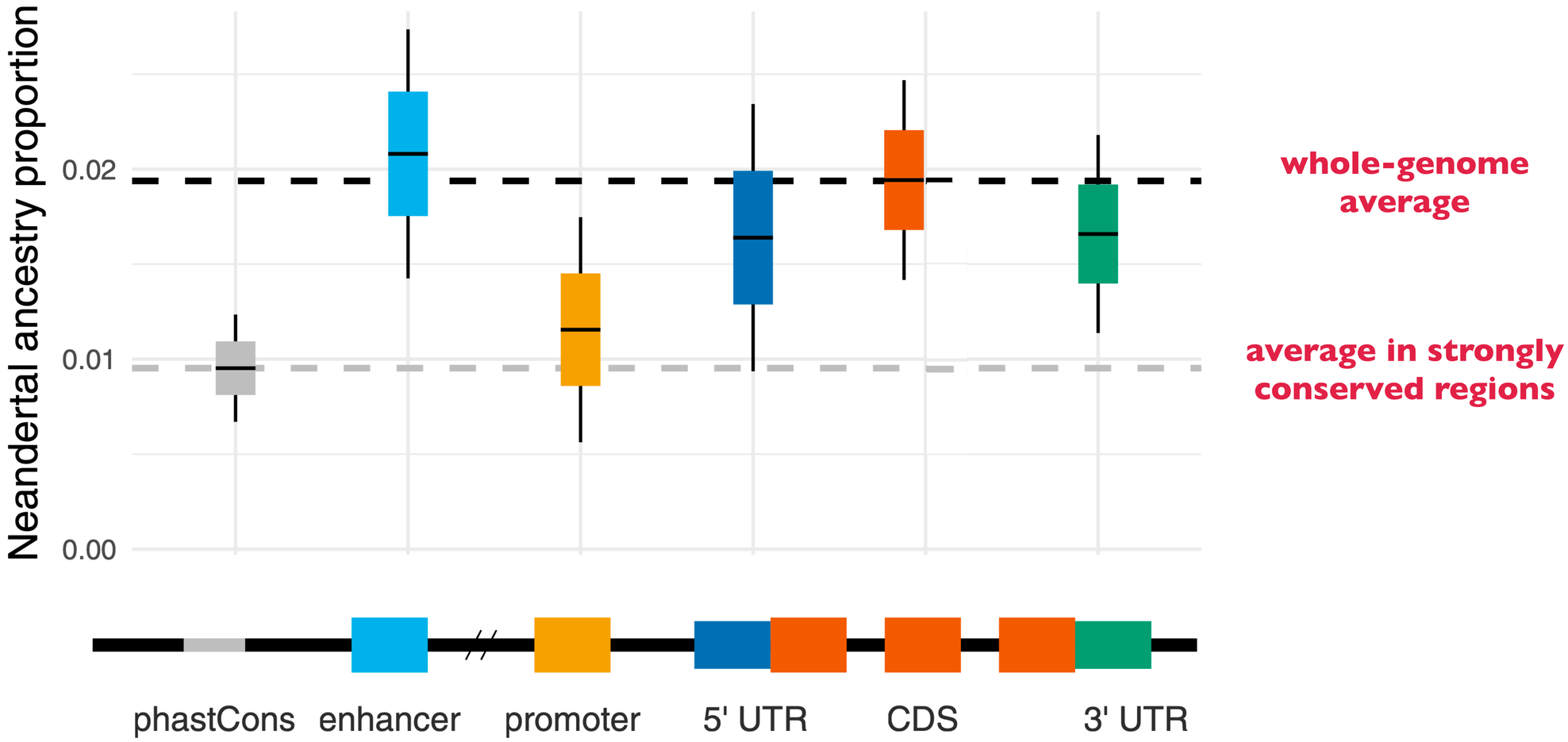
remains of a modern human

Peștera cu Oase, Romania


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
O
U
K
H
F
D
0
50
100
150
200
250

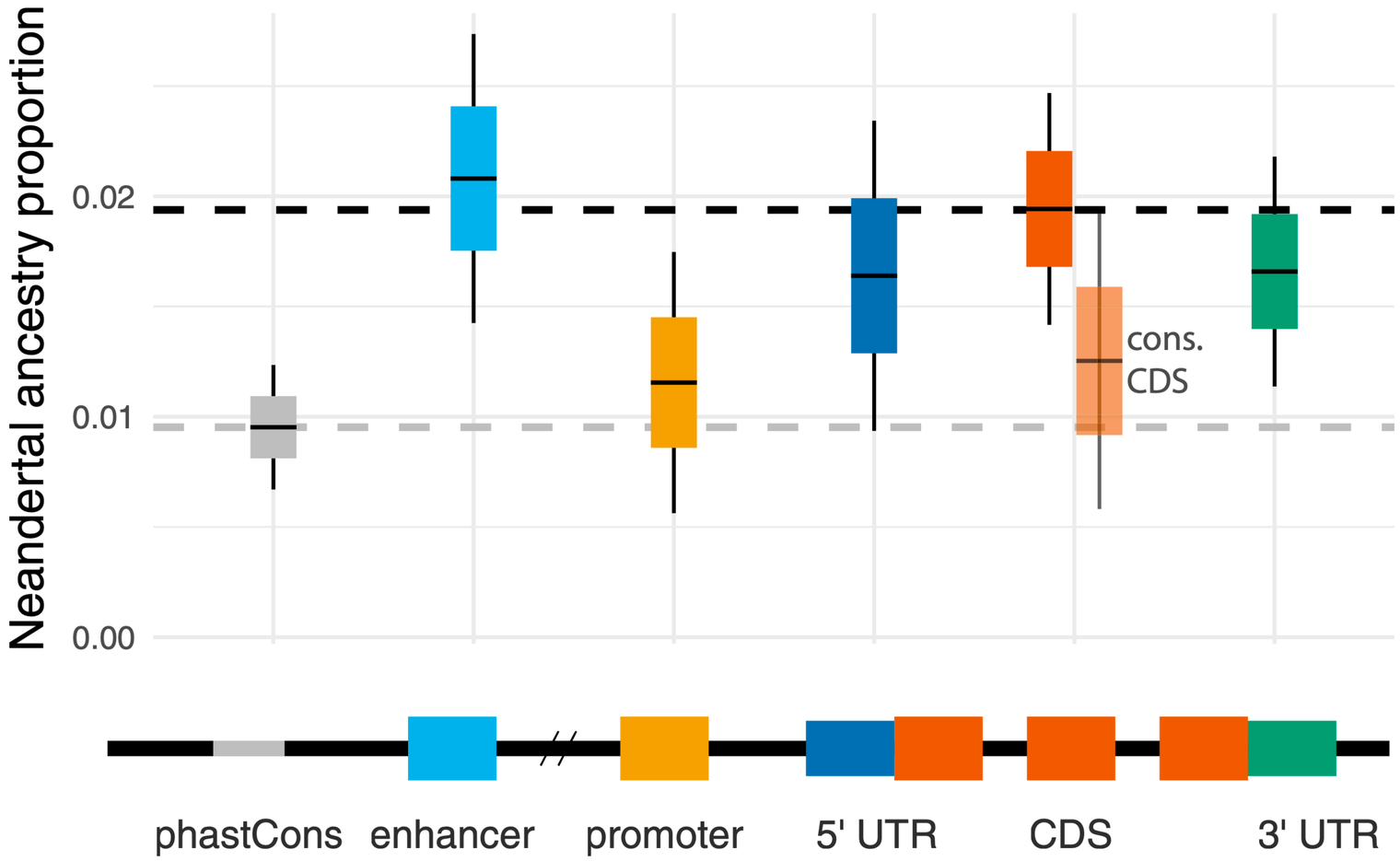

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
O
U
K
H
F
D
0
50
100
150
200
250

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



Petr *et al*. (PNAS, 2019)

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



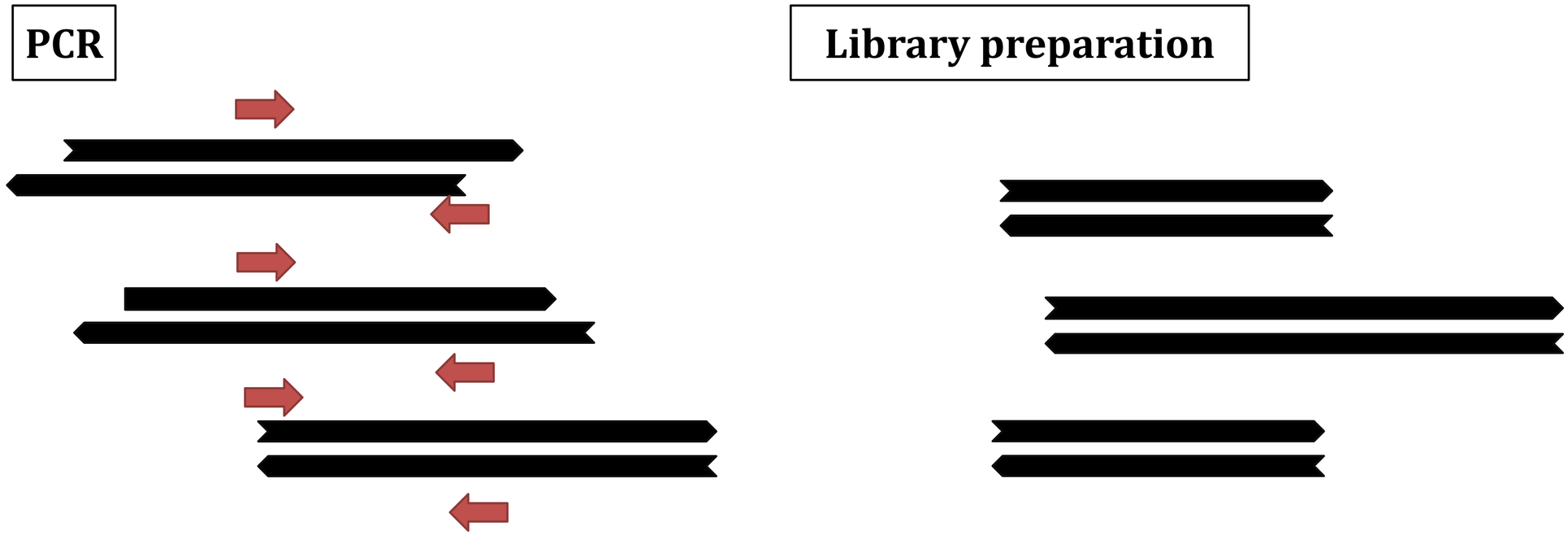
whole-genome average

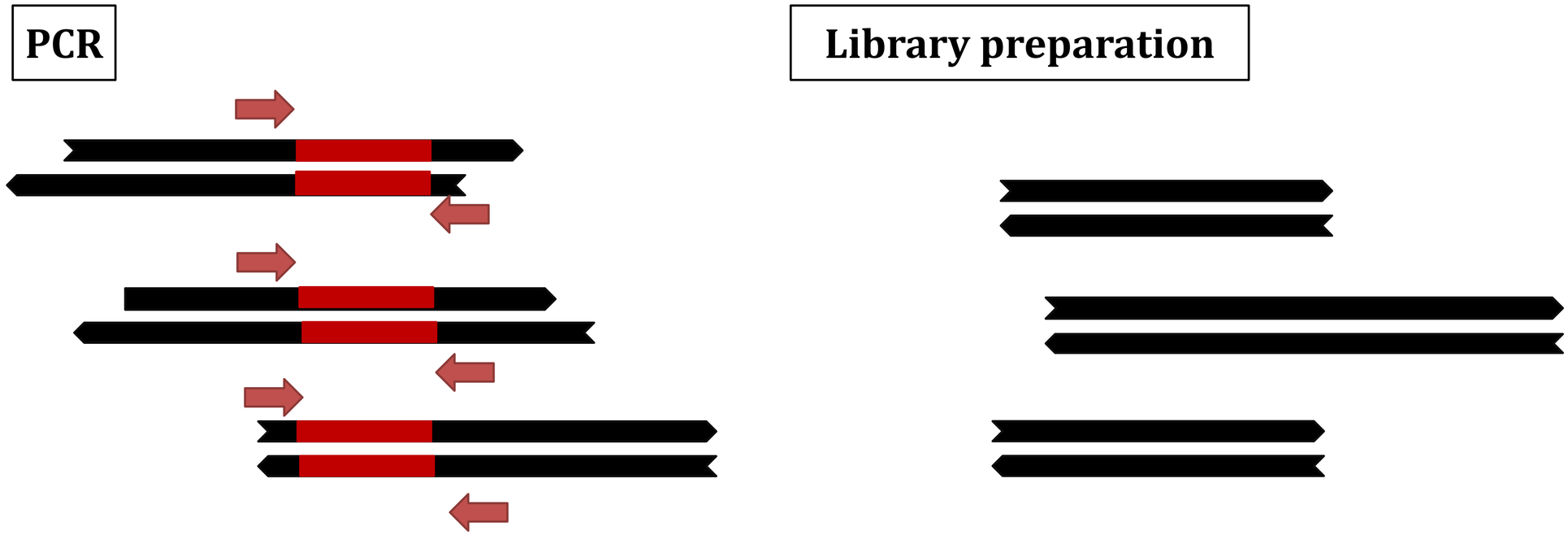
average in strongly conserved regions

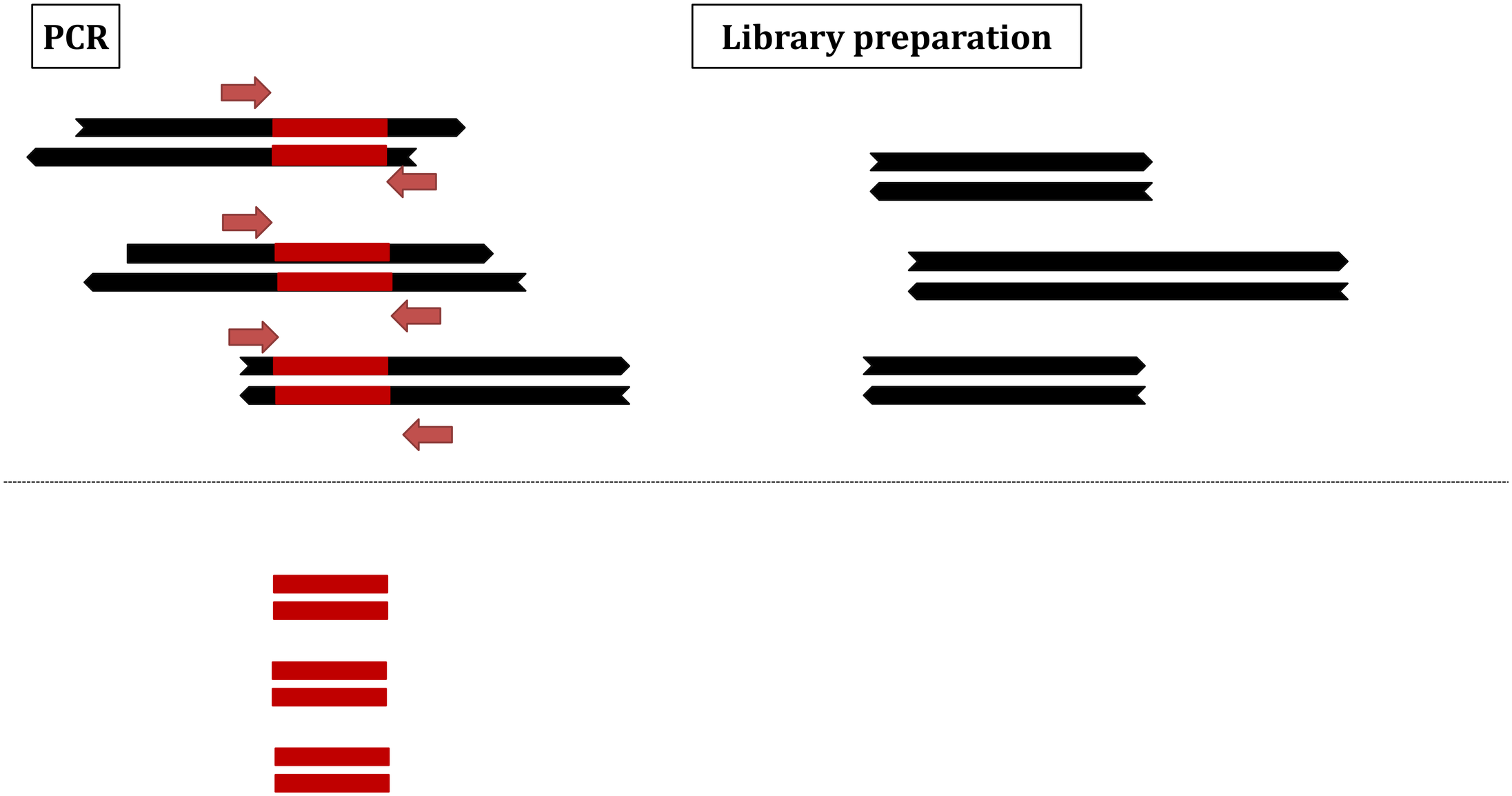
Petr *et al*. (PNAS, 2019)

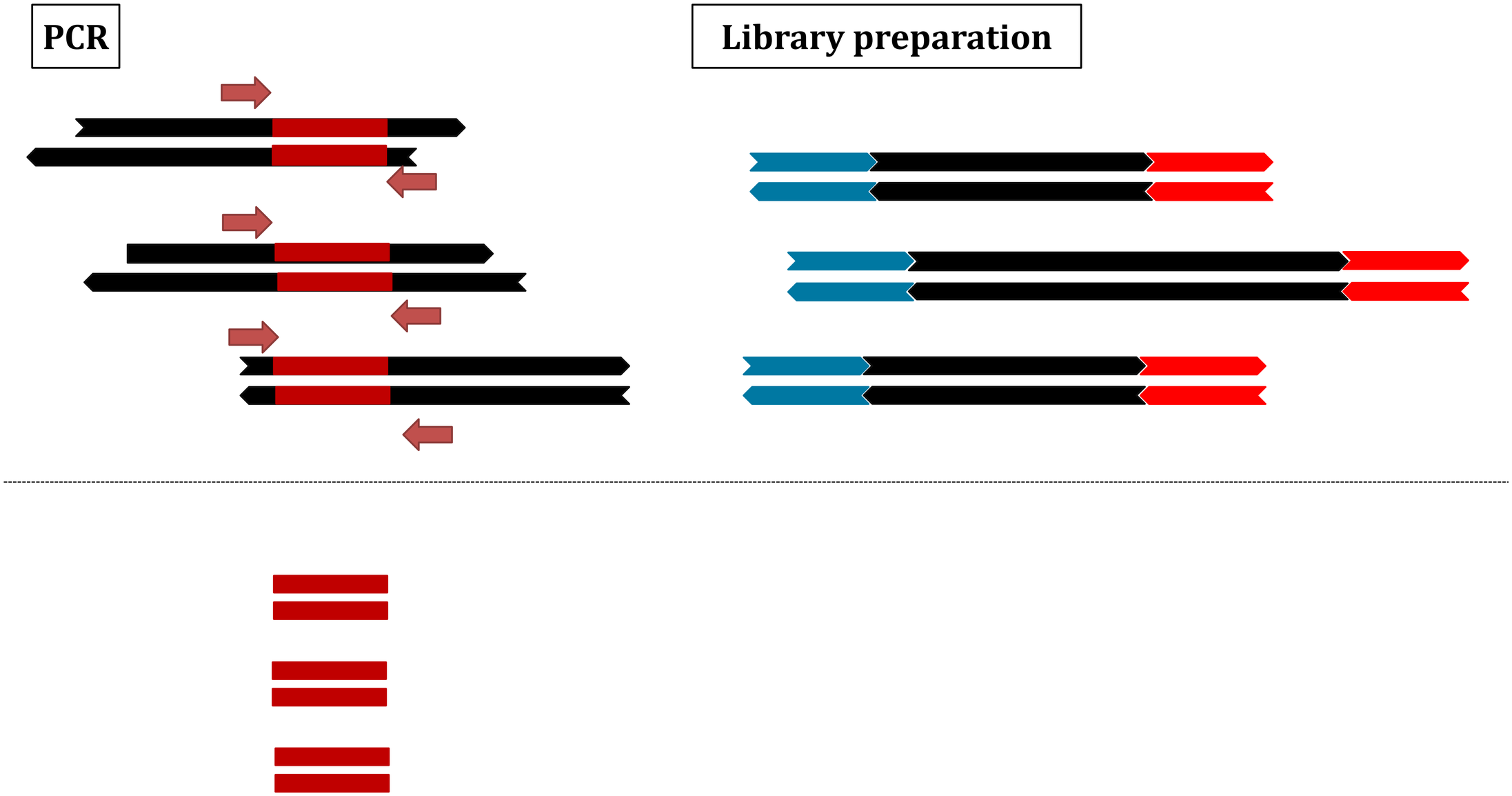
PCR

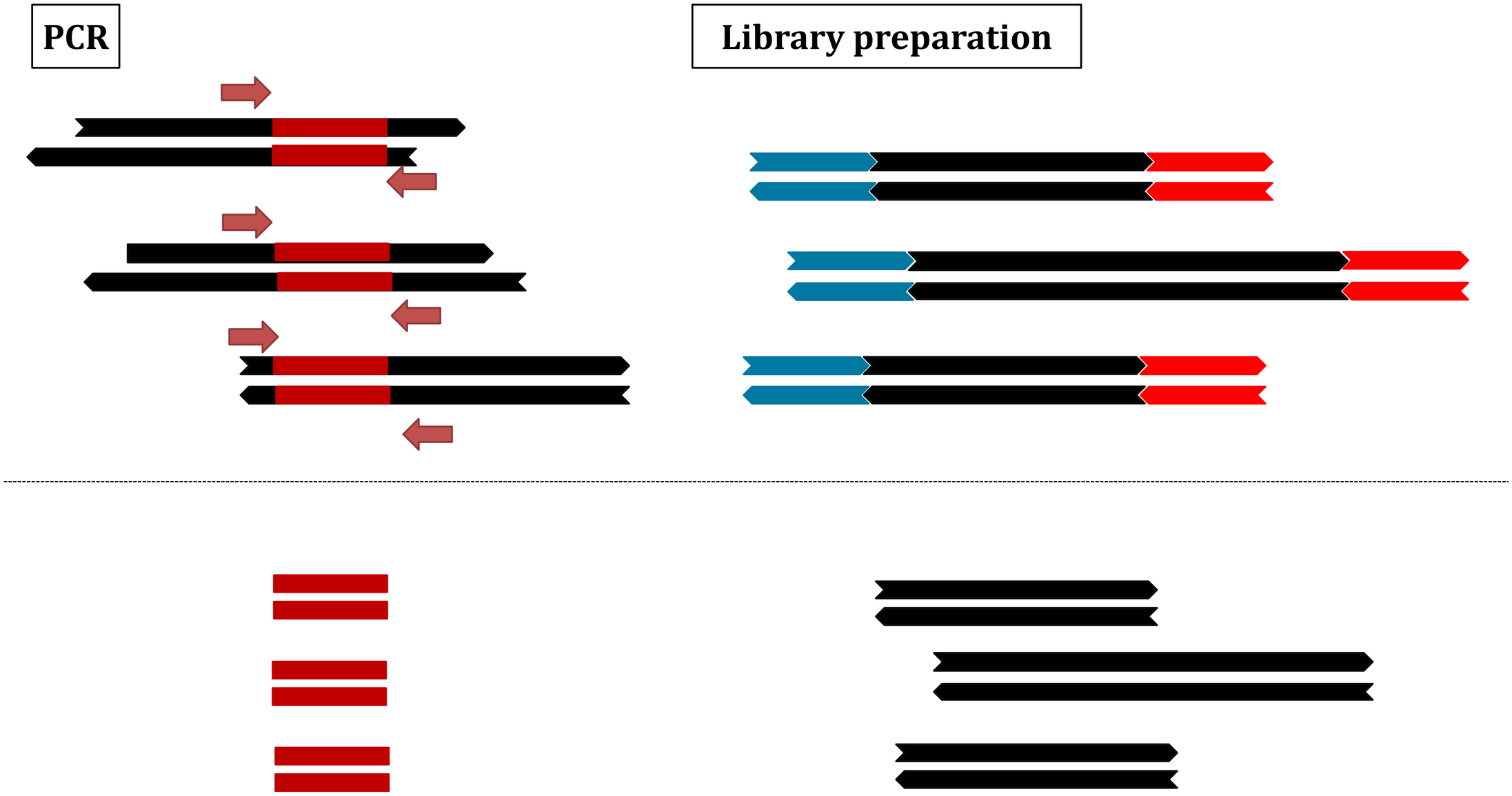
Library preparation

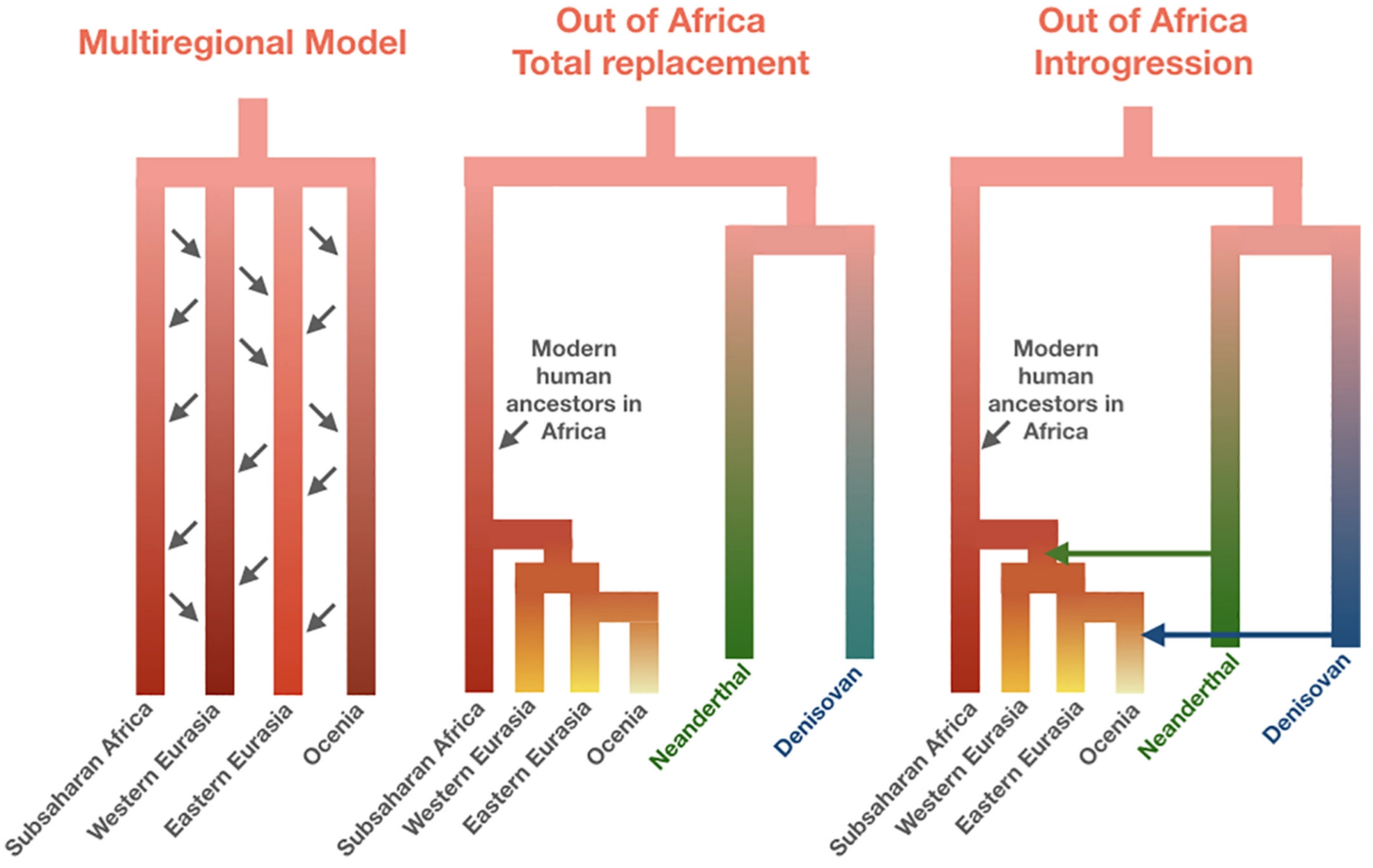

PCR
Library preparation


PCR
Library preparation


PCR
Library preparation


PCR
Library preparation


PCR
Library preparation



Gokcumen (Yearbook of Physical Anthropology, 2019)

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(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...

```
##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 genotype ind. 1 genotype ind. 2 genotype ind. 3 ...