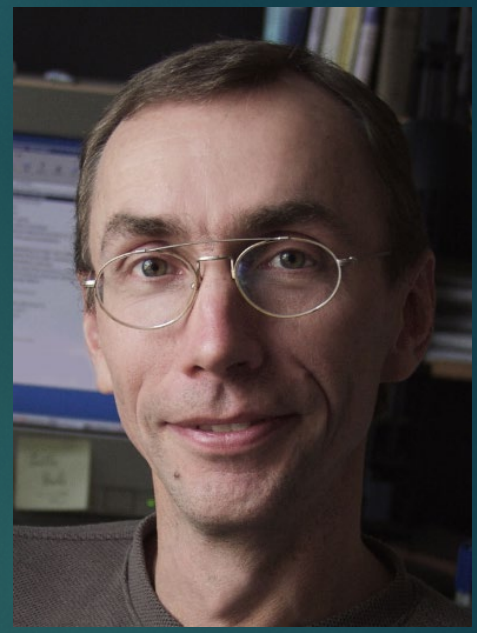
+++ Ancient DNA: A History of Human Paleogenetics

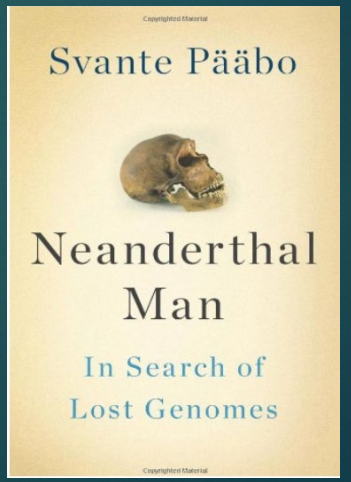
CHARLES J VELLA, PHD

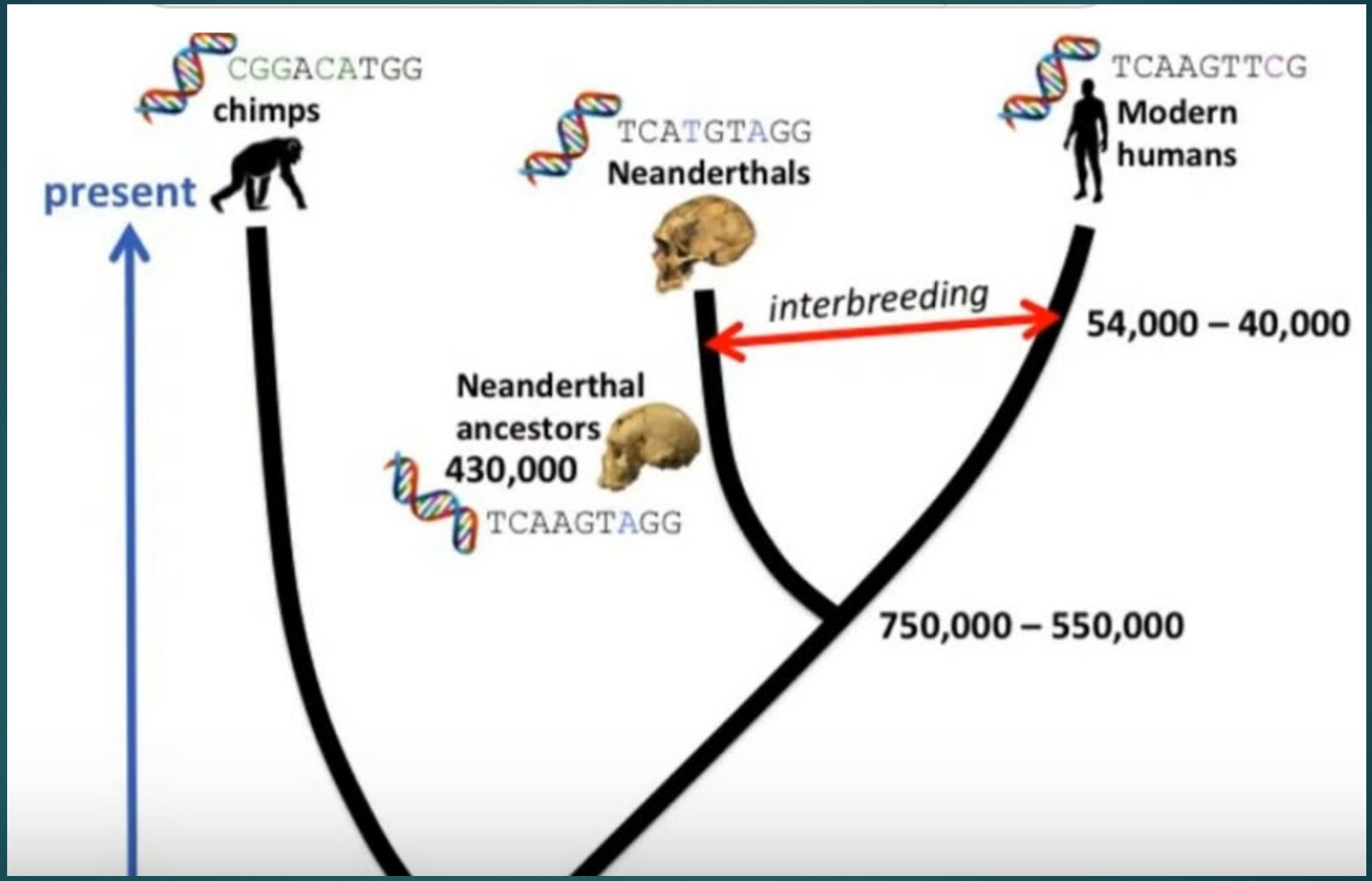
MARCH 27, 2023

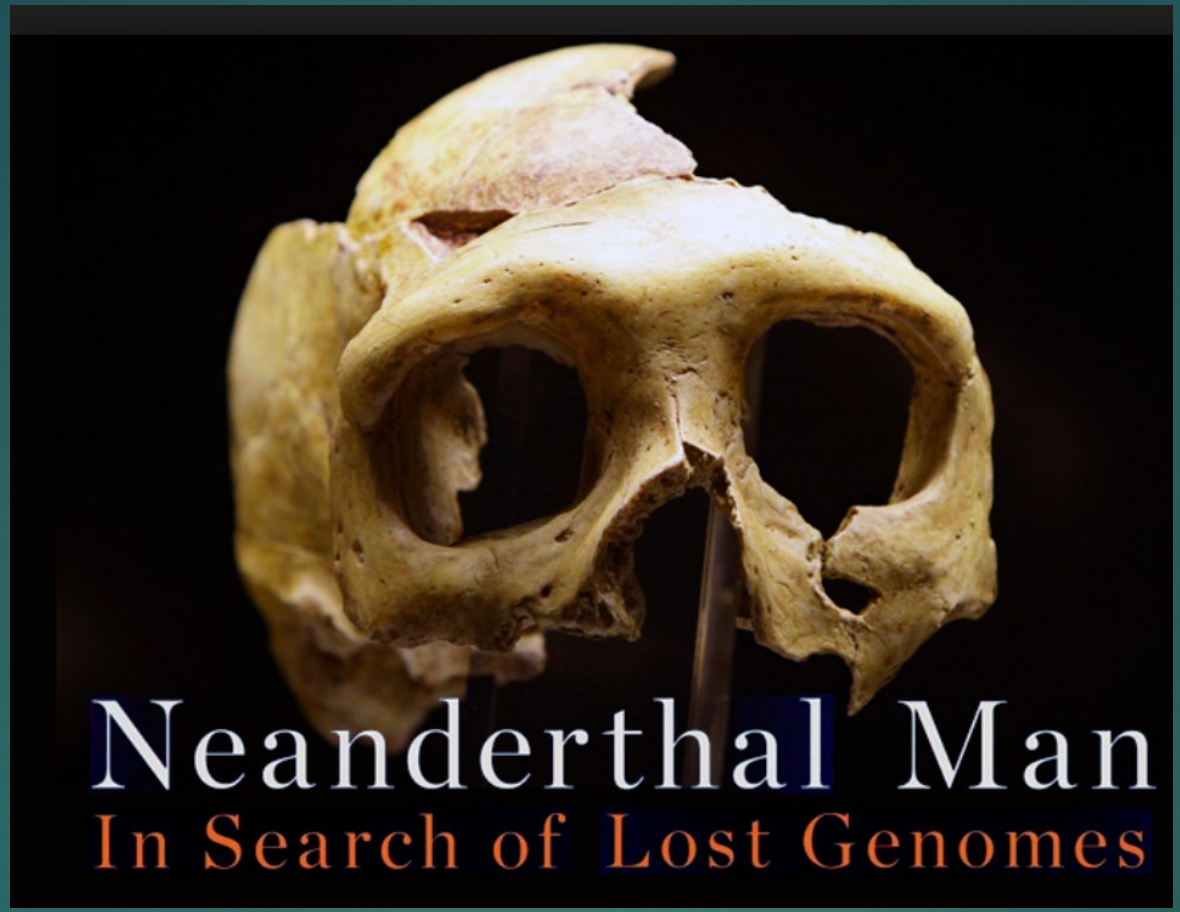
Svante Pääbo (1955-): Grandfather of Evolutionary Genetics



- Swedish biologist specializing in evolutionary genetics
- Director of genetics at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany)
- A leader in the field of molecular evolution & one of the founders of paleogenetics
- 1997: retrieve DNA from Feldhofer Cave Neanderthal; Ns were a different species

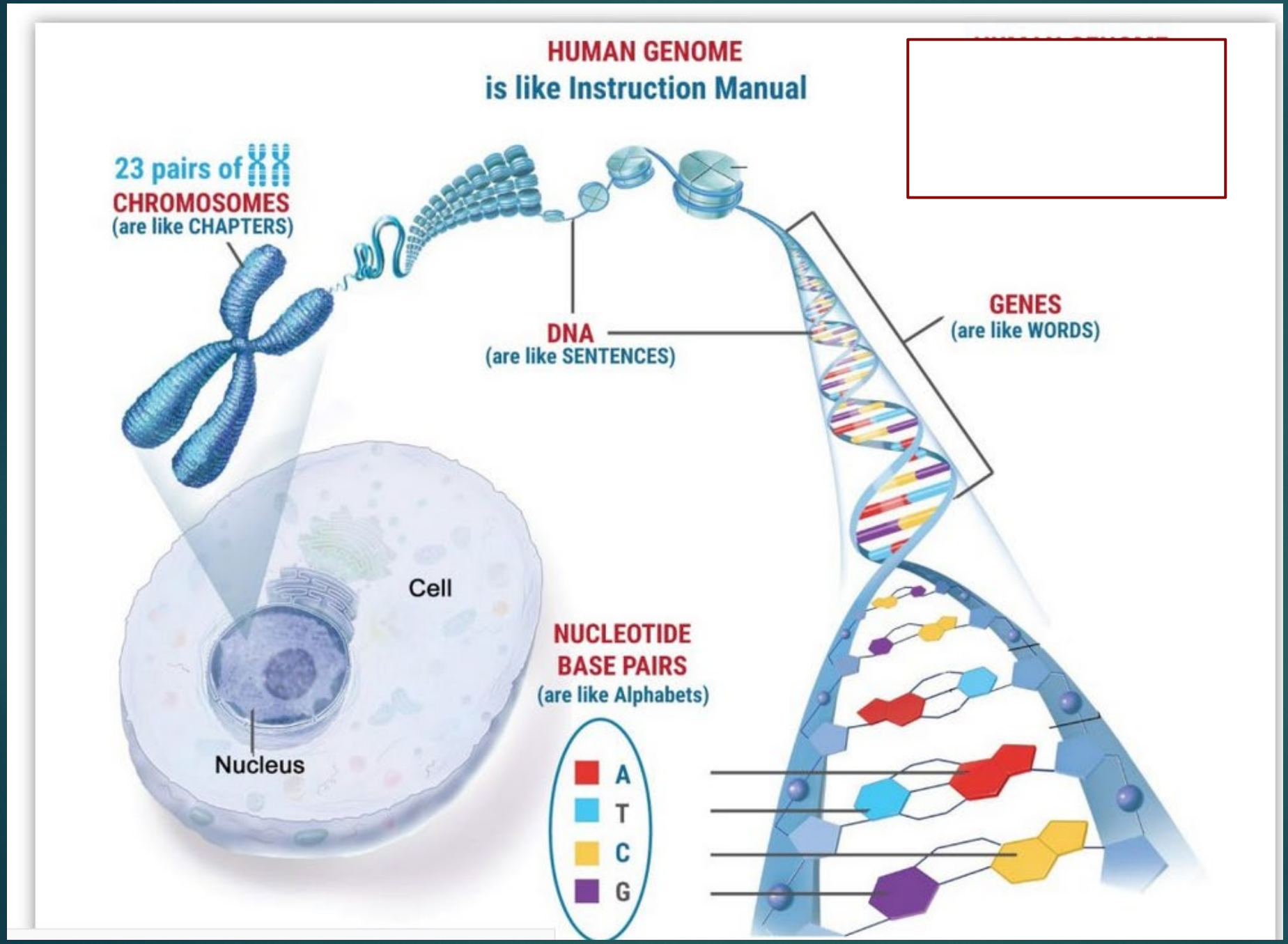



CGGACATGG
chimps
present
TCATGTAGG
Neanderthals
TCAAGTTCG
Modern
humans
interbreeding
54,000 – 40,000
Neanderthal
ancestors
430,000
TCAAGTAGG
750,000 – 550,000



By Svante Pääbo, 2014

Svante Pääbo: Nobel Laureate in Physiology & Medicine, 2022



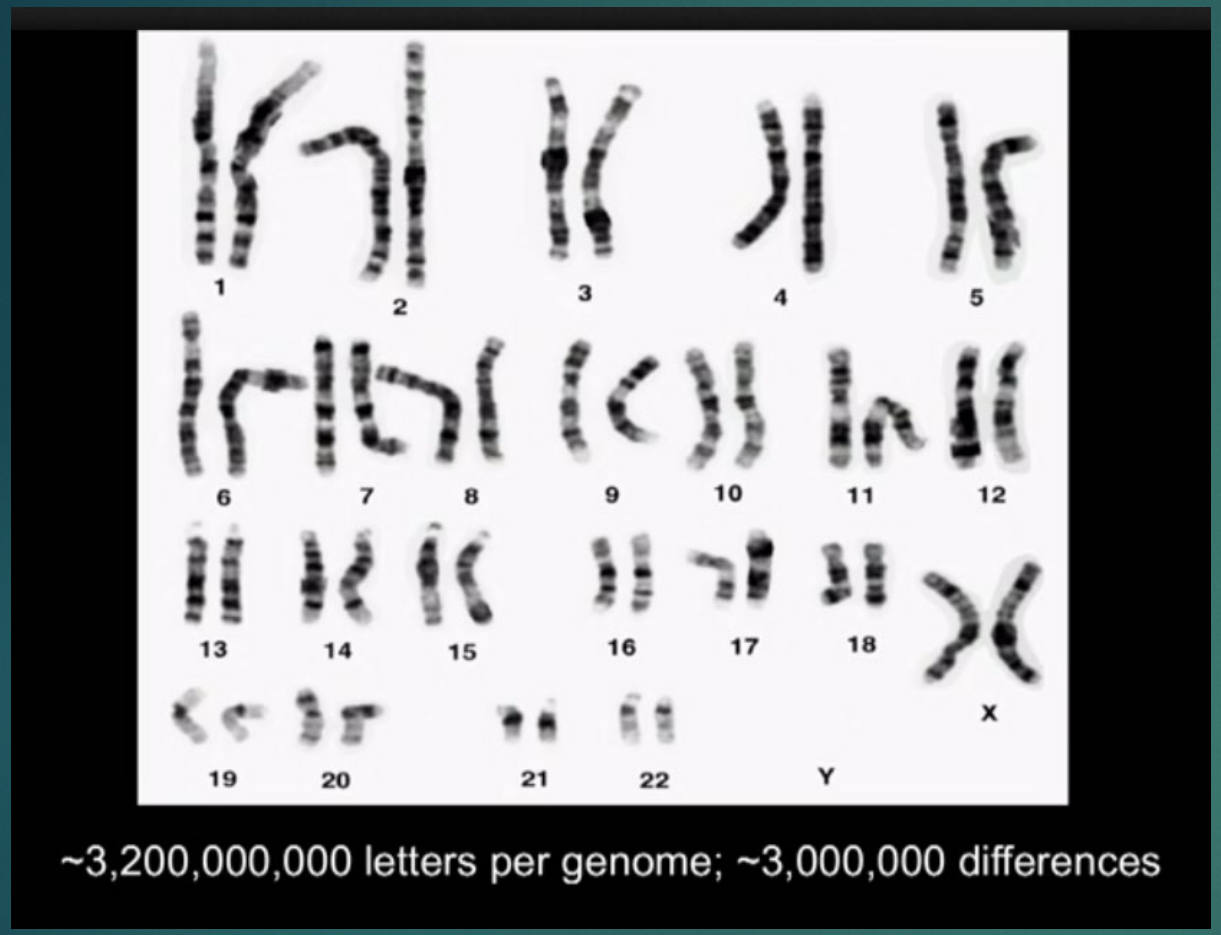
Human Genome

3 billion basepairs

20 K genes

23 chromosomes

3.2 Billion bases per genome



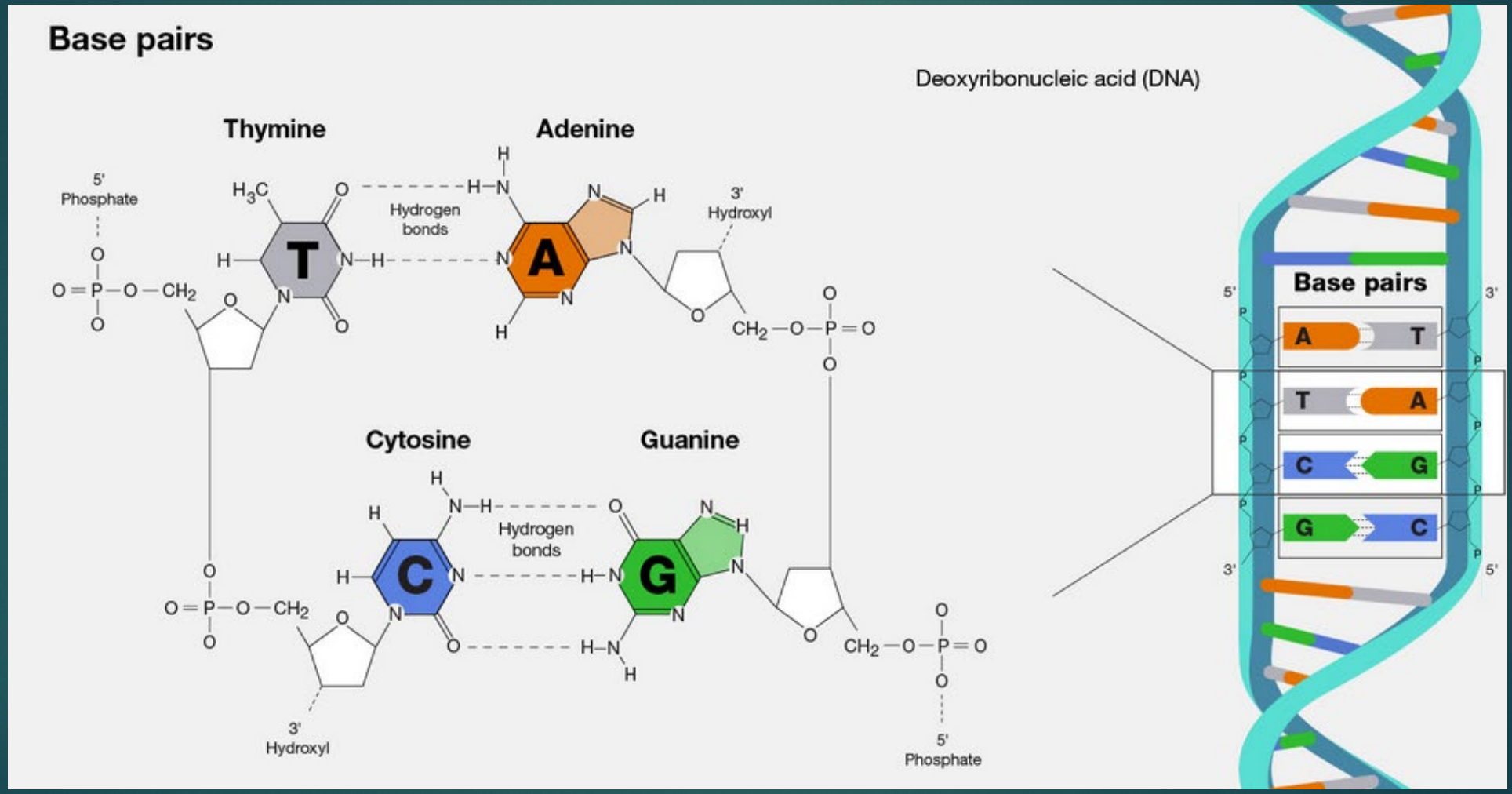
- 99.9% identical DNA between any 2 of us
- Every new baby has 100-200 new mutations
- Between any 2 humans, a mutation occurs every 1200 bases (in every 100 bases in H vs Chimp)

- Each of us has 2 genomic strands; = 6.4 billion nucleotides
- 3 million base pair differences between any 2 people (.1%)

Basic terms in Genetics

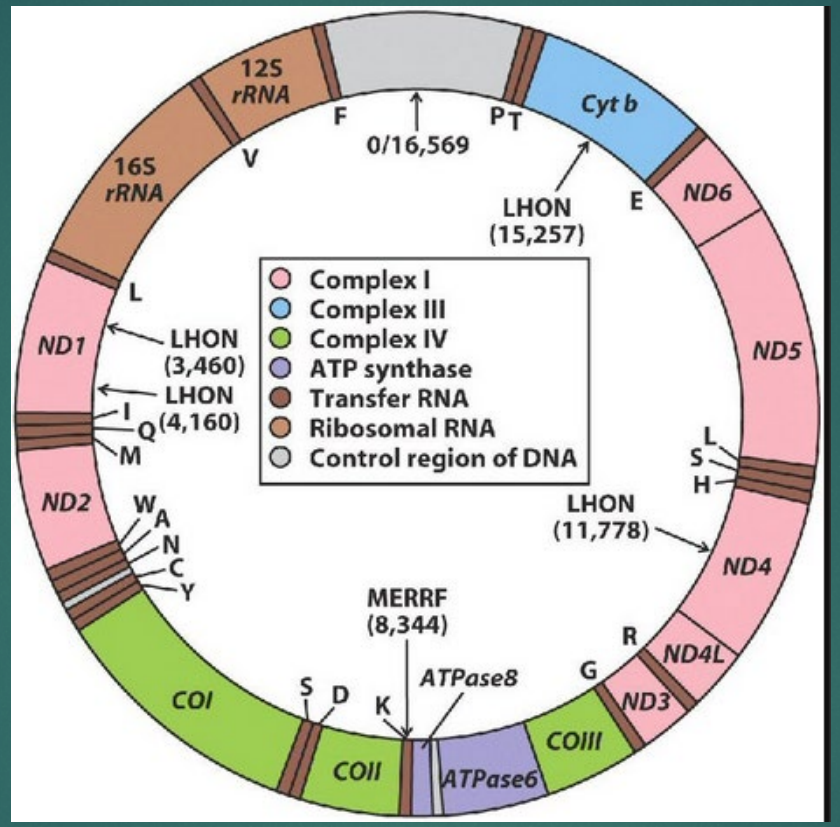
- Basepairs: A base pair consists of two complementary DNA nucleotide bases that pair together to form a “rung of the DNA ladder.”
- Genetic Loci = a specific, fixed position on a chromosome where a particular gene is located
- Allele = 1 of 2 or more alternative forms of a gene that arise by mutation; found in same place on chromosome
- Recombination is a process during meiosis by which pieces of DNA are broken and recombined to produce new combinations of alleles

Base pairs: ACTG combos on DNA strand



mtDNA

- Mitochondrial DNA (mtDNA): found in mitochondria
-
- A separate DNA genome of the mitochondria, which are maternally inherited
- 1000s of organelles found within every cell.

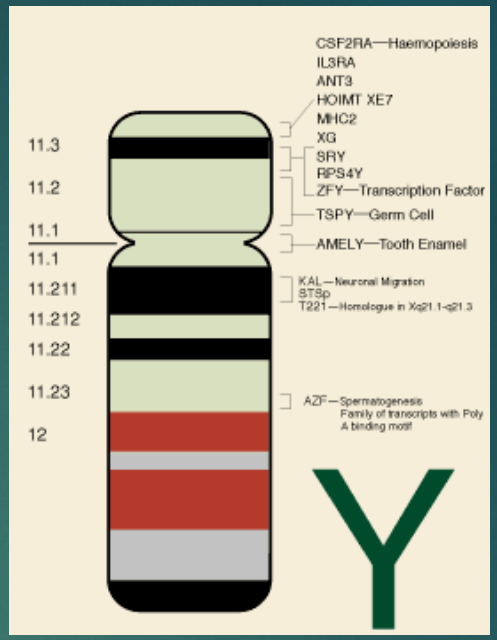


Circular;

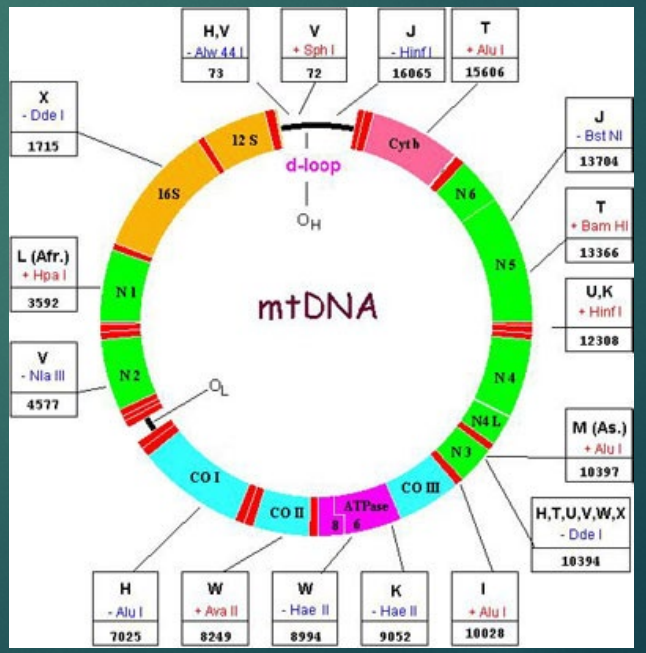
37 genes (13 protein encoding;
22 tRNA encoding, 2 rRNA)

16,569 base pairs

- Y chromosome data is used to trace paternal ancestry.



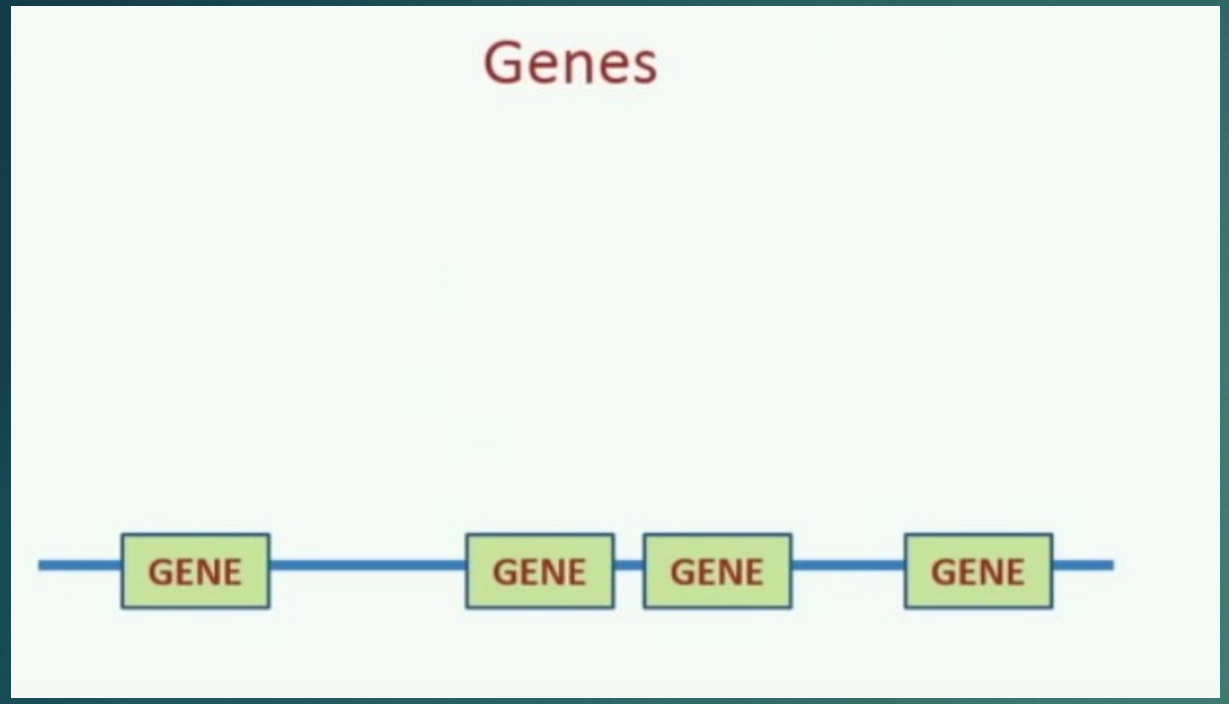
- mtDNA is used to trace maternal ancestry.

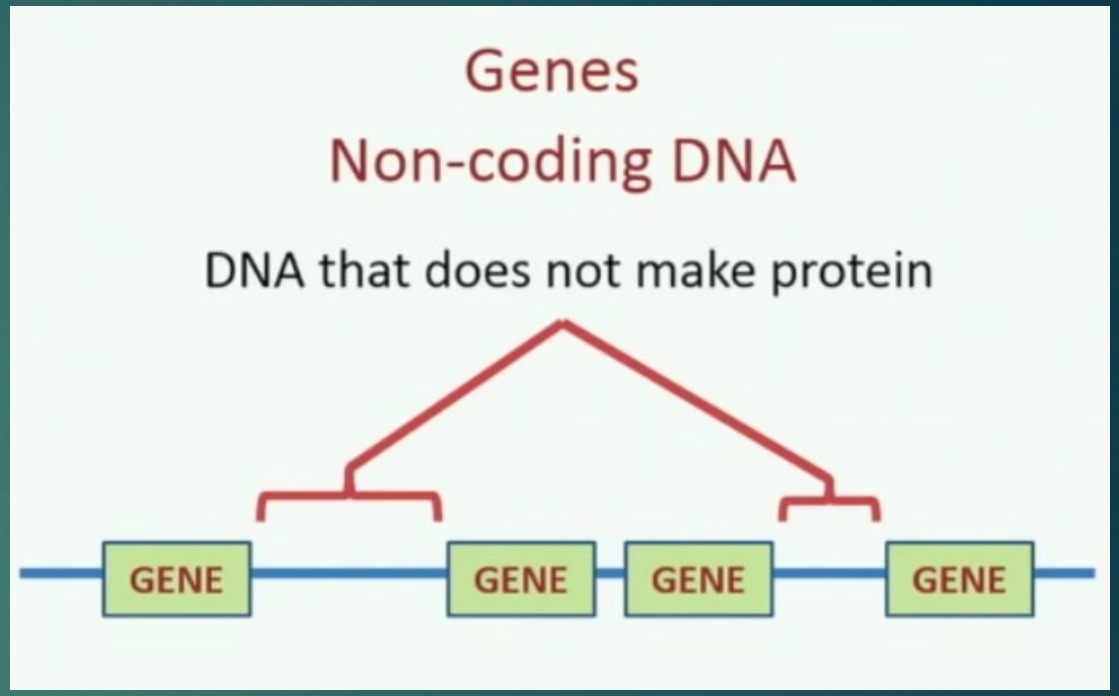


Can now compare computerized DNA sequences

2 humans: 1 difference in every 1200-1300 letters

Genes: sections of DNA that code for proteins





Old “Junk” DNA, now Regulatory DNA: carry instructions for gene regulation – Activate or inhibit gene activity

DNA Contains

- Coding Sequences (genes) (2-3%) 19,000 genes
- Non-Coding Sequences (97-98%)

Before Human Genome Project, thought we had 100 to 250 K genes based on number of proteins we make; 1 gene – 1 protein idea

In fact, we have only 19-20 K genes that make unknown number of proteins (19 K confirmed)

Non-protein coding DNA is not “Junk” DNA

- DNA outside protein producing areas are really important; not “junk
- 5-10% of the human genome is highly conserved across mammals, implying it is highly functional and required for survival
- But only ~2% code for proteins
- Regulatory DNA: Most of the non-coding areas do not code for proteins, but are regulatory, control what genes do

Functional genes: 1 %

- A little over 1% of human DNA (~19,000 genes) accounts for the proteins that carry out almost all of the critical biological processes in the body.
- The other 7% is thought to be involved in the switching on and off of genes that encode proteins
- Every mammal has approximately the same amount of functional DNA

Noncoding DNA

- Junk DNA: now appreciated that the majority of functional sequences in the human genome do not encode proteins.
- Rather, elements such as long non-coding RNAs, promoters, enhancers and countless gene-regulatory motifs work together to control gene and protein expression.
- Research of non-protein-coding elements is now 5x greater than on genes.
- There are now more than 30,000 papers per year linking SNPs and traits. A large fraction of these associations are in the once-dismissed non-coding regions

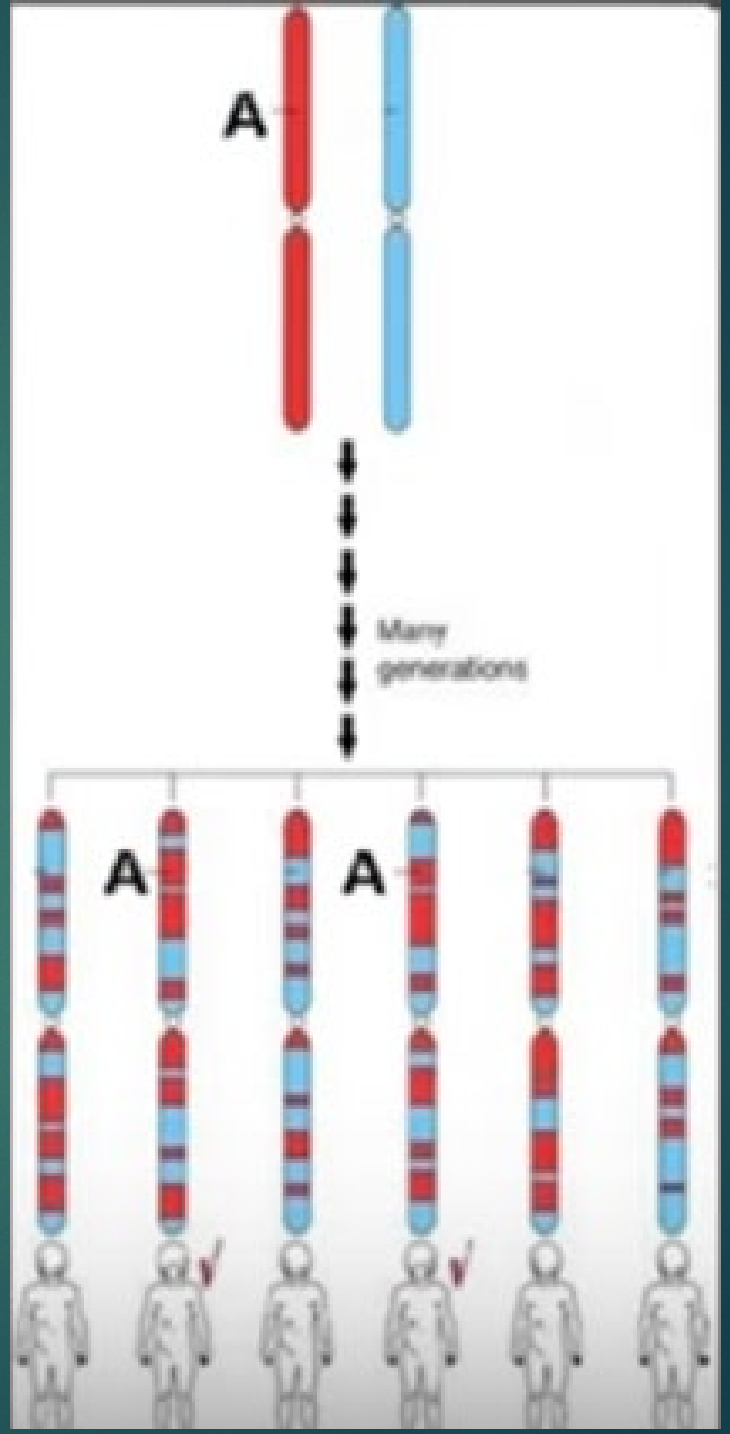
Still mostly “junk” DNA

- DNA sequence is functional only if it evolved to do something useful and if a mutation disrupting it would have harmful effect.
- DNA mutates at random due to UV radiation, or errors during cell division.
- Having too many bad mutations will kill you:
  - if most DNA was functional, most mutations would fall in good sequences & be bad for us
  - if most DNA is junk, most mutations would not affect us; which is the actual reality

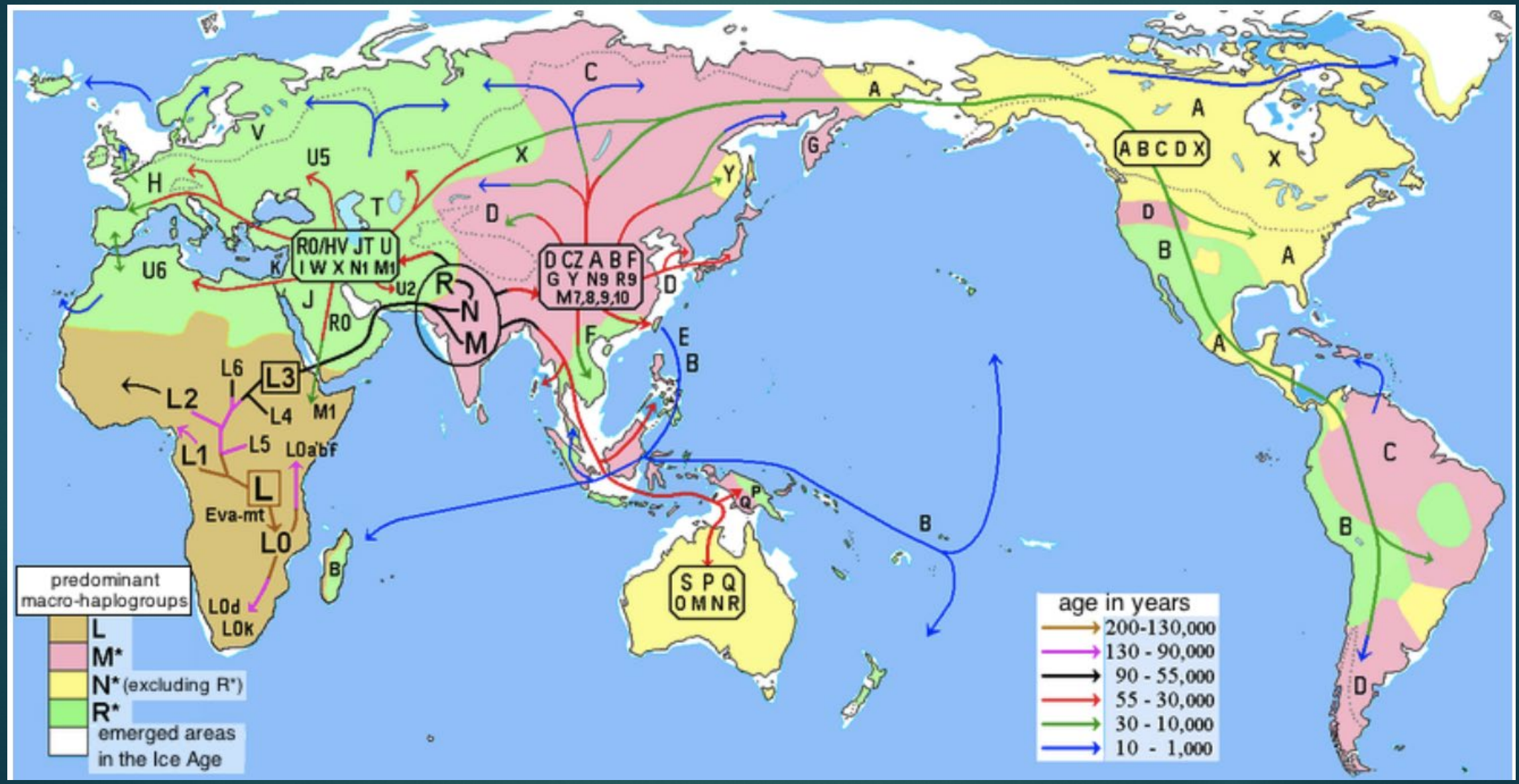
Haplotypes

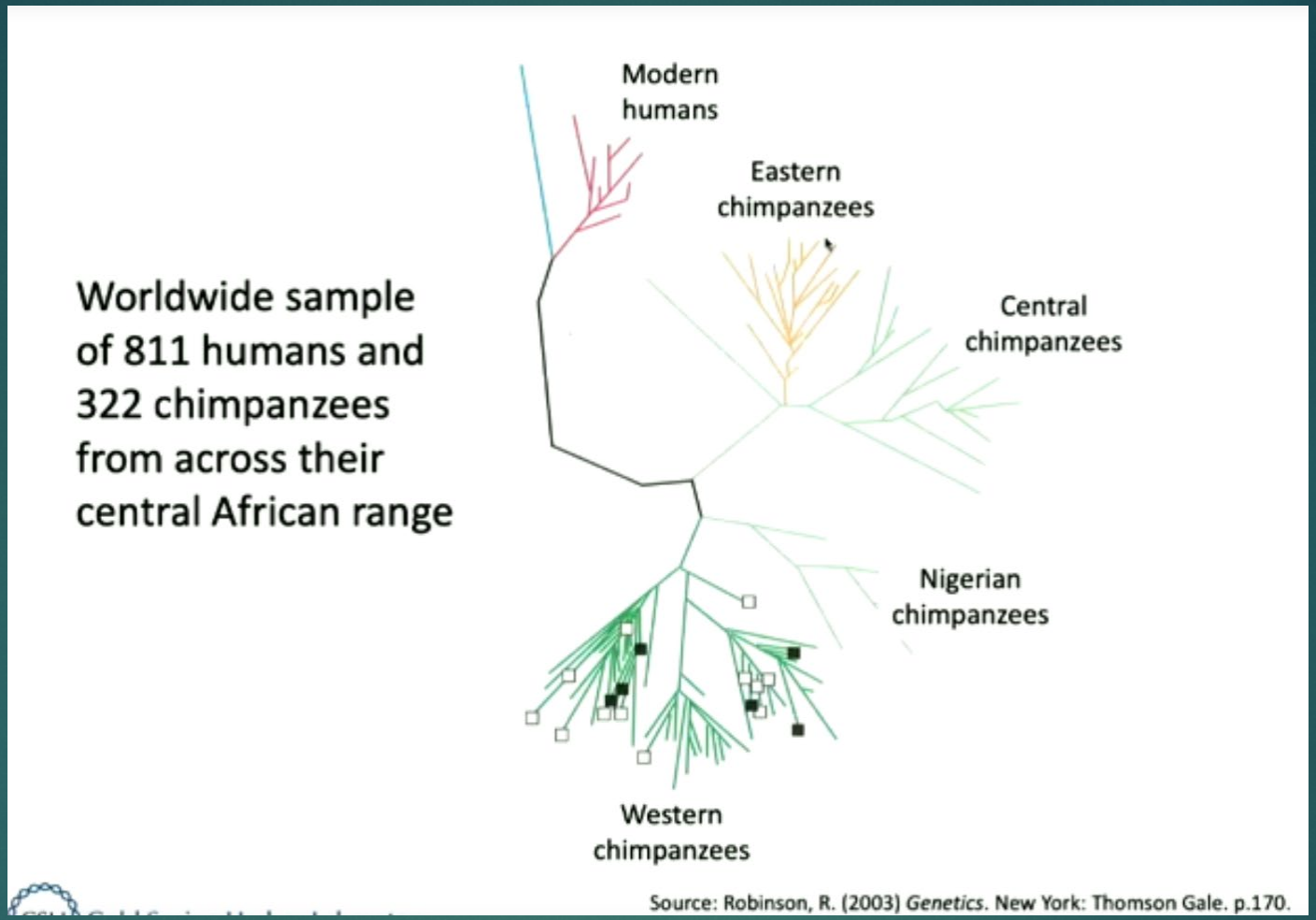
A Haplotype is a combination of alleles at different chromosome regions that are closely linked and that tend to be inherited together

In both mt and nuclear DNA



Haplogroup migrations: L is original one (mitchond Eve)





Chimps have vastly more genetic variation than modern humans.

Africans have the greatest genetic variation on earth

Ancient DNA

- Ancient DNA is the field of molecular evolutionary biology that uses DNA sequence data recovered from poorly preserved organisms, usually deceased for hundreds to millions of years.

- Involves extracting and manipulating sequence data from samples that are old and decayed in some way. But not fully fossilized.

- Current estimate of perseveration of DNA: 10 Ka to 2.0 Ma

Ancient human genomics

- Recent advancements in DNA sequencing technologies and laboratory preparation protocols have rapidly expanded the scope of ancient DNA

- Discoveries include:
  - interactions between archaic and modern humans
  - modern human population dynamics/migrations
  - including the settlement history of most world regions.

- In 2001, a draft sequence of the human genome was published. It is now a reference genome.

Yi. Liu, X. Mao, J. Krause, Qiaomei Fu, 2021

Ancient biomolecules: nucleic acids, proteins, and lipids

- The categories of ancient molecules that have arguably made the biggest contribution to elucidating evolutionary history to date are:
  - nucleic acids (aDNA, eDNA),
  - proteins,
  - lipids.

- Deoxyribonucleic acids (DNA) can show evolutionary processes with the highest resolution,
- but proteins and lipids are important on longer timescales and in geographic areas that are less favorable to DNA preservation

Ancient DNA: Ancient Biomolecules and Evolutionary Inference

- Over the last few decades, studies of ancient biomolecules have transformed our understanding of the evolutionary history of life on Earth.

- The sequencing of ancient DNA has enabled the reconstruction of speciation, migration and admixture events for extinct species.

Enrico Cappellini…Eske Willerslev, et al., 2018

Ancient Biomolecules

- Since then, the focus of aDNA studies has progressed from studying:
  - small mitochondrial and nuclear DNA fragments retrieved from a single species to
  - multiple species to
  - full genomic sequencing of one or several specimens, to
  - single-nucleotide polymorphism capture-based population genomics and
  - whole-genome shotgun sequencing, often including over hundreds of individuals.

Applications in Evolutionary Biology

- Analyses of ancient biomolecules have led to some of the biggest breakthroughs in the field of evolutionary biology.

Archaic Hominins

- Ancient genomics has been central to furthering our understanding of
- human evolution after our divergence from archaic hominins, as well as
- the evolutionary consequences of human encounters with archaic hominin groups in the Late Pleistocene.

aDNA and aProteins

- The irreversible post-death degradation of ancient DNA has so far limited its recovery—outside permafrost areas—to specimens that are not older than approximately 500 K years.

- By contrast, tandem mass spectrometry has enabled the sequencing of approximately 1.5 M-old collagen proteins, and suggested the presence of protein residues in fossils of the Cretaceous period (145 to 65 Ma)—although with limited phylogenetic use

Hybrid Neanderthal love child found in cave in Siberia

Teenage daughter of a Neanderthal mother and a Denisovan father who lived 50,000 years ago reveals how humans' ape-like cousins frequently interbred

A prehistoric 13- year- old girl who lived 50,000 years ago was the love child of two separate species of ancient human ancestor, according to a new DNA analysis of her remains.

A study of a tiny bone fragment found in a cave in Russia shows the teenager had a Neanderthal mother ...

A Neanderthal museum exhibit

Prehistoric humans did hanky-panky, shows study

Remains Of A First -Generation Child With Neanderthal And Denisovan Parents Found

Genetic analysis of bones discovered in a Siberian cave hints that the prehistoric world may have been filled with 'hybrid' humans

Berlin: Scientists say they've found the remains of a prehistoric female whose mother was a Neanderthal and whose father belonged to another extinct group of human relatives known as Denisovans.

The 90,000- year- old ...

Brown *et al.* 2016. *Nature Sci. Reports*

aDNA and protein in “fossils”

- Please note that if you have complete fossilization, it means that all the organic components have turned to minerals
- aDNA can only be found in “fossils” that have not been completely fossilized
- aDNA is organic, not fossilized
- But Ancient proteins can be discovered in completely fossilized specimens

Molecular clock to divergence time

- The molecular clock = technique that uses the mutation rate of biomolecules to deduce the time in prehistory when two or more life forms diverged.

- Data = nucleotide sequences for DNA, RNA, or amino acid sequences for proteins

- Neutral mutations (random changes) occur at a constant rate in a species

Molecular clock

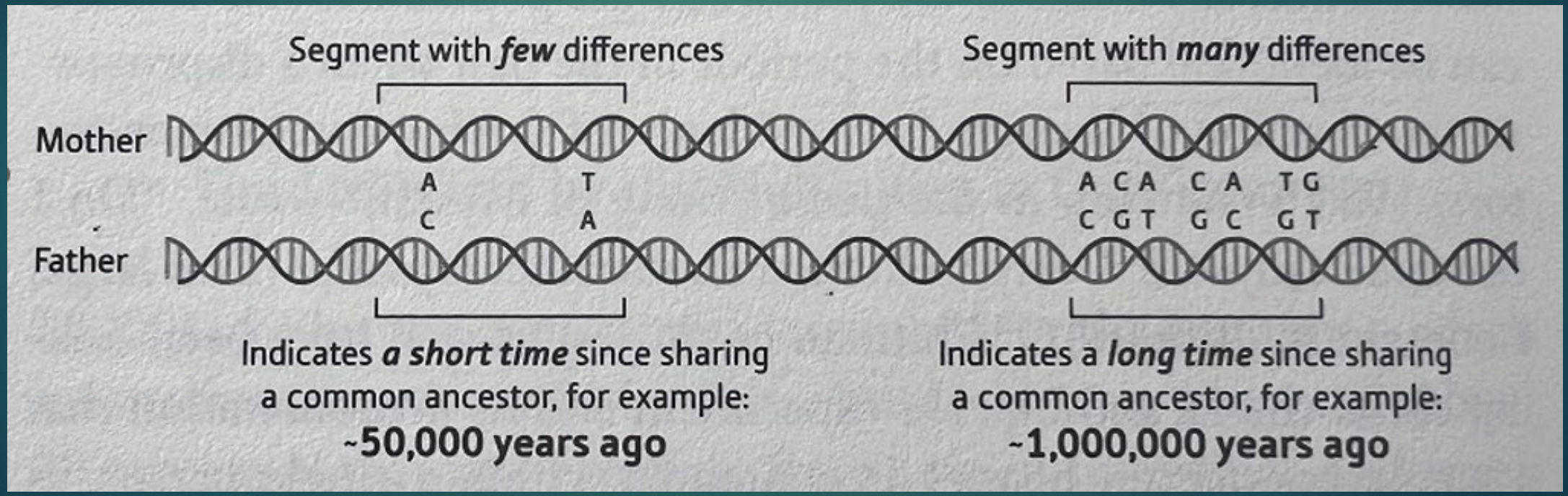
- Mutation rate = clock-like rate of molecular change; further back, more mutations
- Can compare number of mutations in 2 species to arrive at time of species divergence, a LCA
- Most phylogenies require that the molecular clock be calibrated against independent evidence about dates, such as the fossil record

DNA Molecular Clock

- DNA mutations happen one at a time
- They occur at a constant rate
- They accumulate over time
- # different mutations = time of separation between two individuals or populations
- The fewer mutation differences = more closely related
- More mutation differences (more time) = more distantly related

How we can tell how long it has been since our genes shared a Last Common Ancestor: More mutations, longer time to LCA

Each of us has two genomes: one from our mother, one from our father. Some segments are more alike than others. The more differences—or mutations—in a given segment, the longer it's been since the gene copies bequeathed to us by our parents shared a common ancestor.



Sources of Ancient biomolecules

- Bones and teeth remain the most widely used mineralized specimens for extracting aDNA

- Wealth of other suitable calcified and mineralized substrates, such as eggshells, invertebrate shells, coprolites, and dental calculus, the latter two being particularly valuable for investigating ancient microbiomes.

- Keratinous material, e.g., hair, claws, and feathers; but are scarce

- Archaeobotanical remains, such as fossilized seeds, fruit, and wood, = source of ancient plant DNA,

Ancient DNA sources

- Ancient DNA is analyzed from:
  - Mummies
  - Organisms preserved in amber
  - Plant materials found in ancient tombs
  - Bacteria
  - Bones
  - Pages in books
  - Dirt
  - Any chewed material, i.e. tar

Sources of aDNA

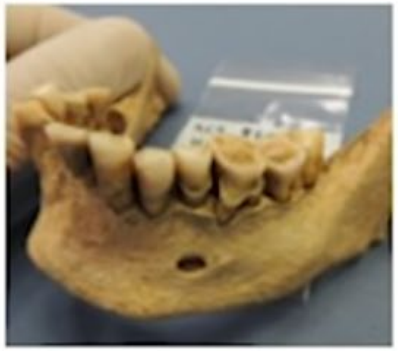
Hair
(Bonnichsen et al., 2001; Rassmussen et al., 2010)

Plants
(Goloubinoff et al., 1993; Medović et al., 2011)

Coprolites
(Sutton, 1996; Poinar et al., 1998a,b; Hofreiter et al., 2000)

Quids
(LeBlanc et al., 2007)

Tobacco lump



Dental calculus
(Ozga et al., 2016)

Ice/Soil cores
(Willerslev et al., 2003; Haile et al., 2009)

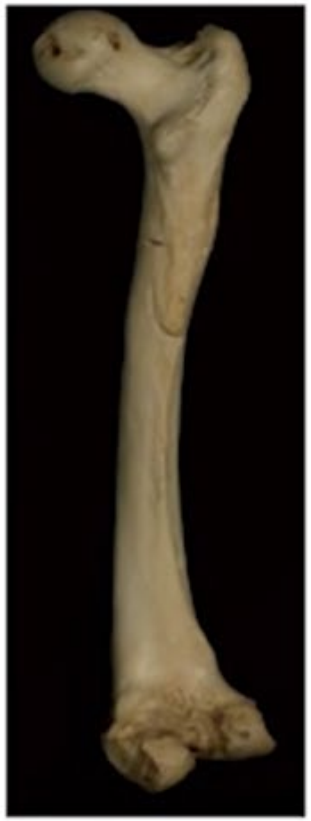
Parchments
(Parry et al., 1996; Teasdale et al., 2014)

Clothing
(LeBlanc et al., 2007; Schröder et al., 2016)

Sources of ancient DNA

First Neandertal DNA
(Krings et al., 1997)

First Denisovan DNA
(Krause et al., 2010)



Oldest DNA outside permafrost
First Middle Pleistocene hominin DNA
(Meyer et al., 2014)

Oldest DNA sequences
(Orlando et al., 2013)

3rd oldest DNA: 560–780 Ka Horse

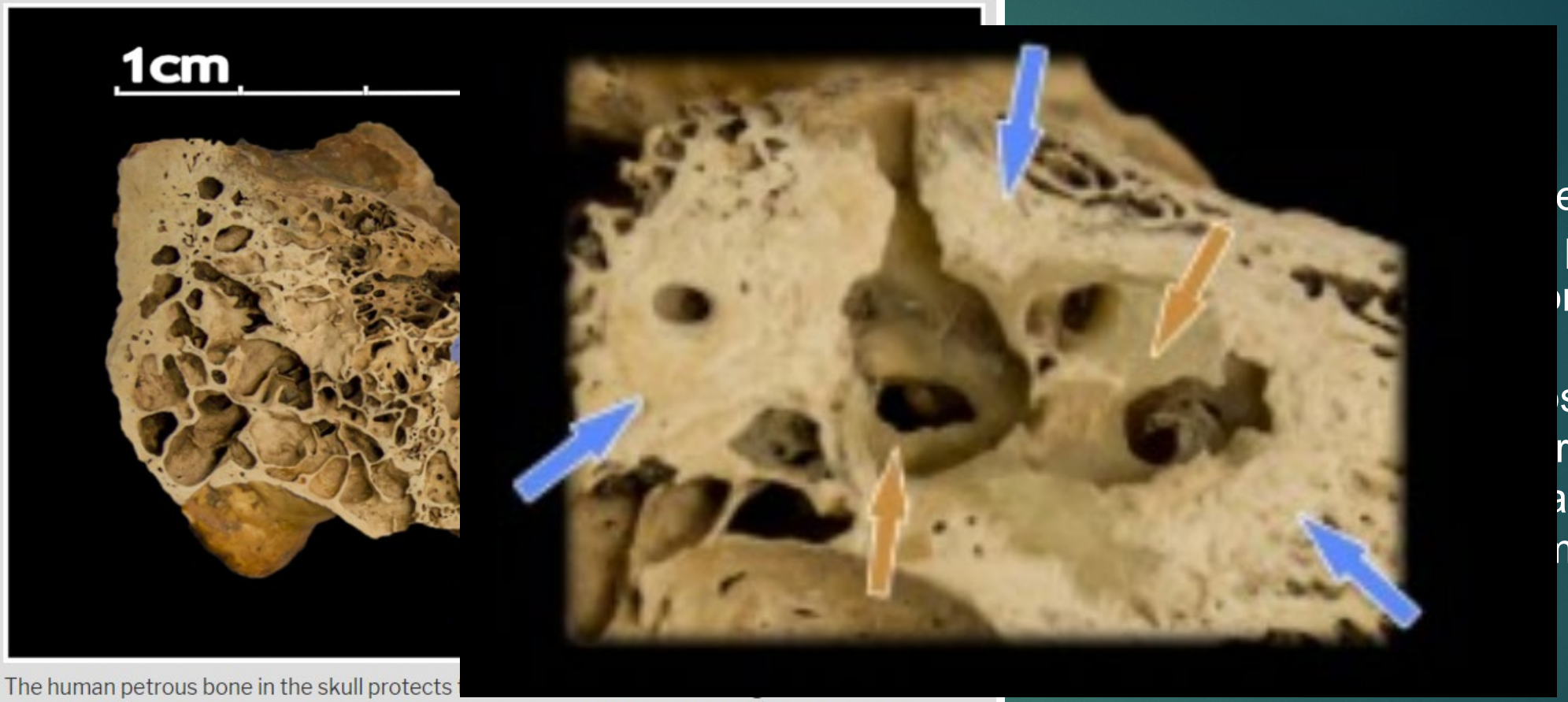
Sima de los Huesos Neandertal, 400 Ka

Oldest modern human DNA
(Fu et al., 2014)

Ust'-Ishim Siberian MH, 45 Ka

A human petrous bone being analyzed at the Max Planck Institute for the Science of Human History in Jena, Germany.

Skull's Petrous Bone; best source of paleo DNA



The human petrous bone in the skull protects
hardest, densest bones in the body, some portions (such as the area in orange, protecting the cochlea) are denser than others. Possibly because the petrous bone is so dense, DNA within the petrous bone is better preserved than in other bones. In some cases, scientists have extracted more than 100 times more DNA from the petrous bone than other bones, including teeth. Credit: Pinhasi et al., 2015, PLOS ONE.

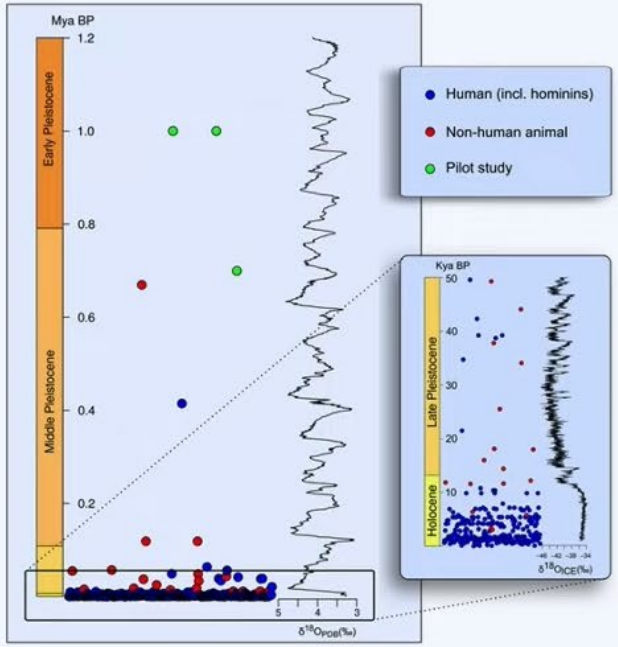
e human skull.
pyramid-shaped
ne is nicknamed

ssibly because it
ructures such as
ates sound into
nicircular canals

Ron Pinhasi, et al., 2015

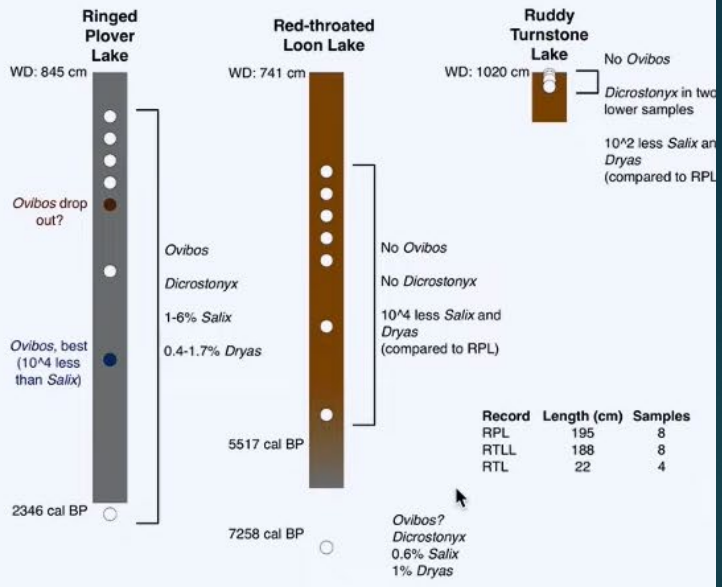
I don’t know where I am going from here, but I promise it won’t be boring – David Bowie

Very old DNA



Microbes

sedaDNA



Genetic analysis of Ancient DNA

- Main challenges to the study of ancient DNA = Two technical complications
- The first complication is molecular damage = errors in DNA sequence
- The second complication, DNA contamination = contemporary DNA contaminates almost all ancient remains and many laboratory environments.
- Molecular damage and DNA contamination give rise to erroneous computer DNA sequences used for final analysis.

1990s: Jurassic Park Hypothesis: dinosaur aDNA

Ancient DNA were awarded a special place in the public imagination by the 1993 release of Steven Spielberg's "Jurassic Park."

Would it be possible to resurrect the dinosaurs?

1993

- “Mad scientist clones dinosaur for defense purposes!”
  - New York Post

Genes from dinosaurs saved in amber?
Problem = DNA degrades

- Multiple 1990's peer reviewed journal papers claiming DNA from dinosaurs have never been retracted.

No Jurassic Park

- DNA begins as one very long strand (single continuous 3 billion bps).
- DNA Degradation: Sunshine (UV radiation) breaks down DNA in our skin, but proof-reading enzymes correct it in us when we are alive.
- Once death occurs, it begins to degrade. Breaks down into ever smaller fragments.
-
- UV radiation, oxygen, water, enzymes in gut, microorganisms in soil, etc. degrade DNA in dead cells.

Characteristics of Ancient Biomolecules

- **Ancient DNA:** Ancient DNA is normally heavily fragmented and chemically modified.

- After the death of an organism, DNA is initially degraded by normal endogenous nucleases (enzyme capable of cleaving DNA).

- This is soon followed by exogenous degradation processes, such as oxidation, hydrolysis (water damage), and background radiation, which alter the nitrogenous bases and cleave the backbone of the DNA molecules, leading to their fragmentation.

Ancient DNA & temperature: heat matters

- DNA concentration and mean fragment length declines exponentially with age, while terminal deamination (C to T and A to G at strand ends) increased with age.

- Most of DNA data cannot be taxonomically identified due to the absence of genomic references in public databases.

- Of the remaining 8%, most of the data (93%) derive from Bacteria and Archaea.

Ancient DNA

- Oldest aDNA has been sequenced from ice and permafrost ranging between 400 and 2.0 Ma in age.

- In contrast, the age of the oldest aDNA reads from the tropics is ~2 orders of magnitude lower

DNA damage

- After death, DNA strands are cut into ever smaller fragments with age.
- Greatest DNA degrader is water. DNA fragments may survive if cells dry out postmortem.
- Bones and teeth survive longest.

Ancient DNA degradation

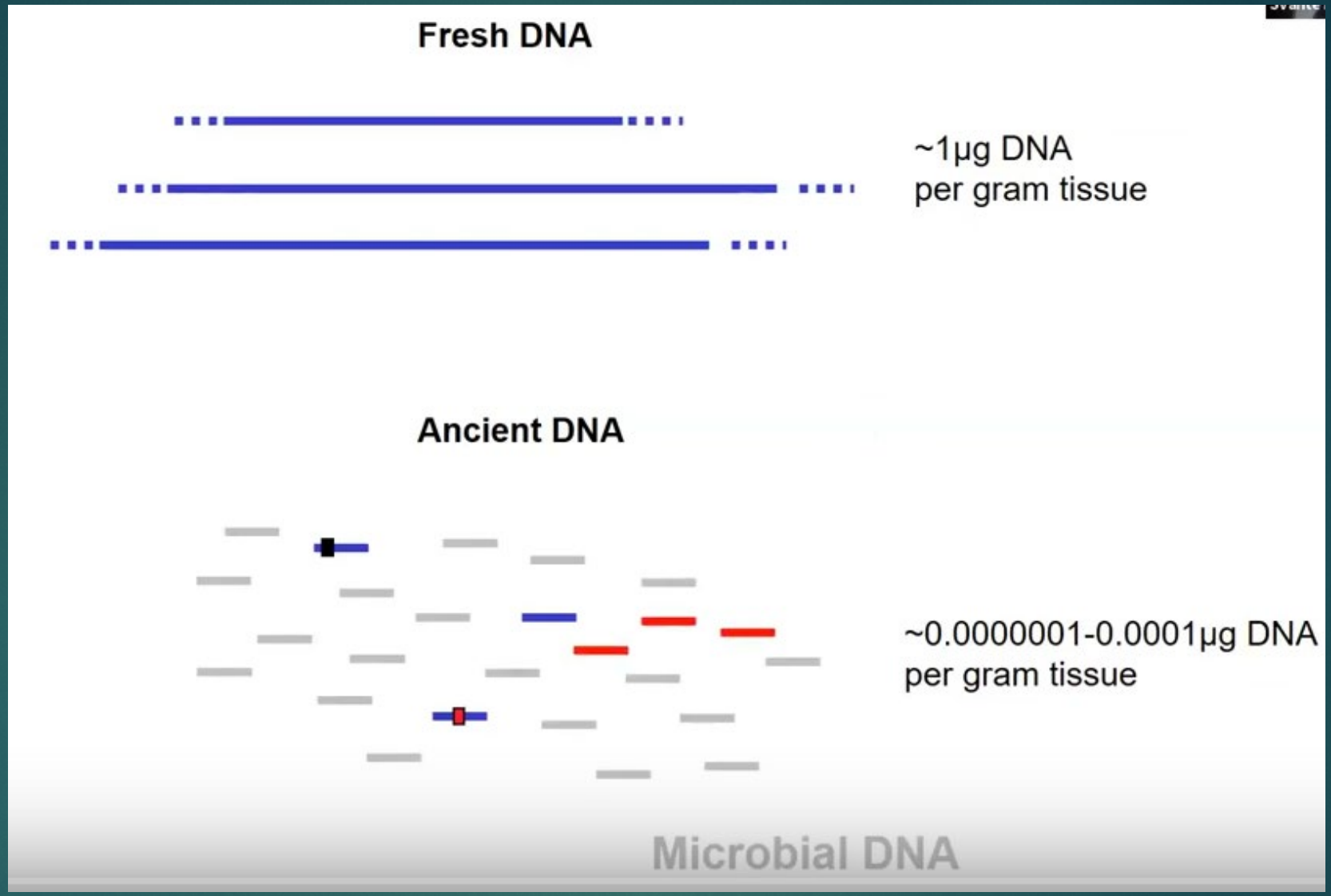
- In 1990s, Pääbo's lab: no replicable DNA from ancient amber
- His conclusion -- No dinosaur DNA: can't extract DNA from specimens that no longer have any.
- When organisms die, their DNA decomposes into minute fragments; the older the specimen, the smaller the DNA fragments.

Ancient DNA degradation

- How long this takes depends on factors like temperature, burial conditions and the number of microbes making a meal of it.

- Pääbo's initial prediction: under optimal conditions — very cold ones — DNA could survive for around 1 million years.

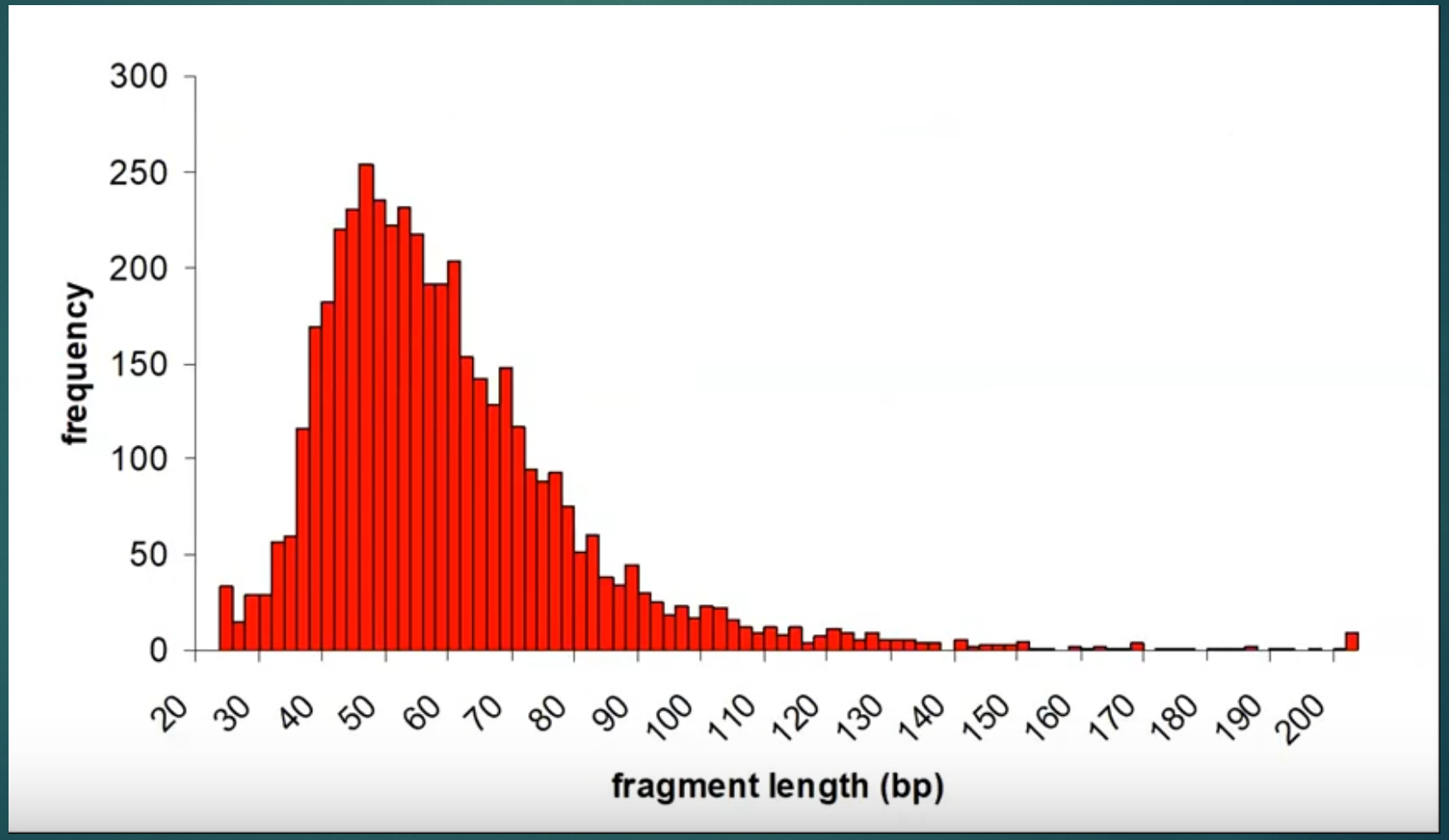
aDNA



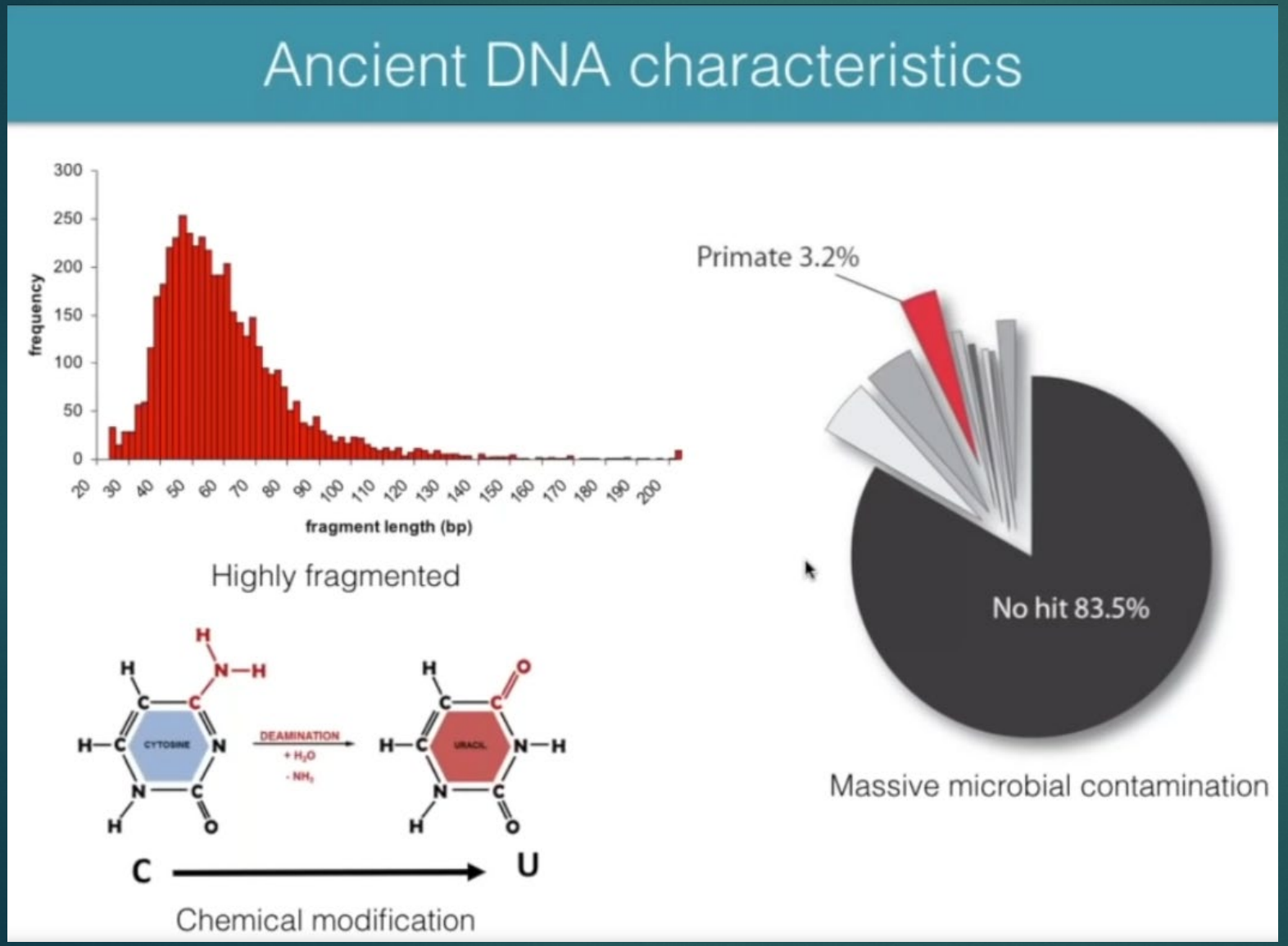
Fresh DNA:
1 millionth
of a gram

- Living DNA is in long segments; aDNA is very short, fragmented segments; 99% of aDNA is bacterial DNA and contamination from living humans.

Very short aDNA fragments even from best aDNA: 60-70 bases in length; fragments that are much longer are not aDNA

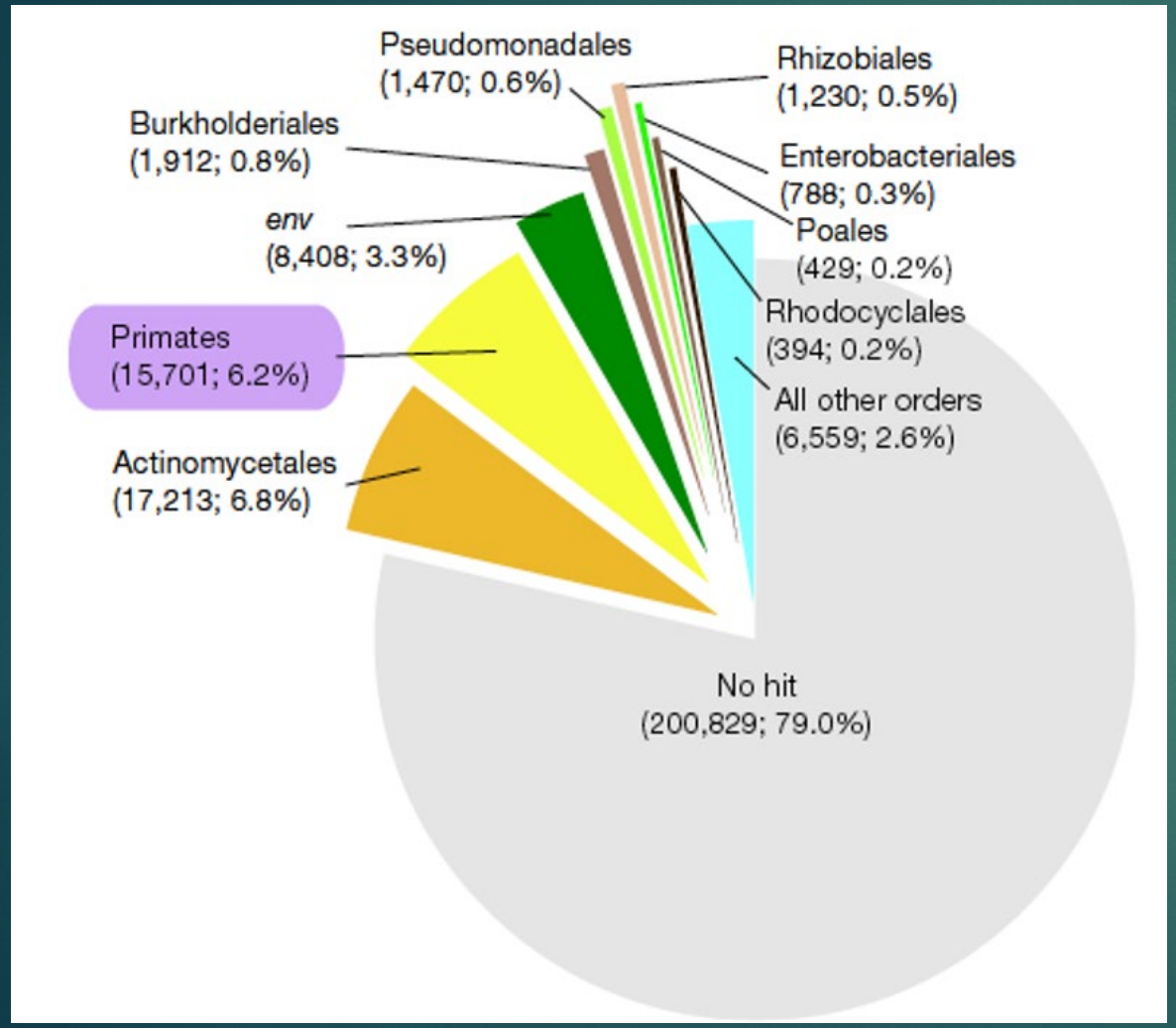


Ancient DNA characteristics required for verification:



- 1 - aDNA = is very fragmented (40-50 bps),
- 2 - It is chemically damaged (Cytosine to Uracil modification),
- 3 - It is mostly nonhuman (only 3.2% primate)

aDNA contamination = Mostly unknown soil-living microbes; Very little hominin DNA in bones: typically 3.5%



Mostly bacterial contamination

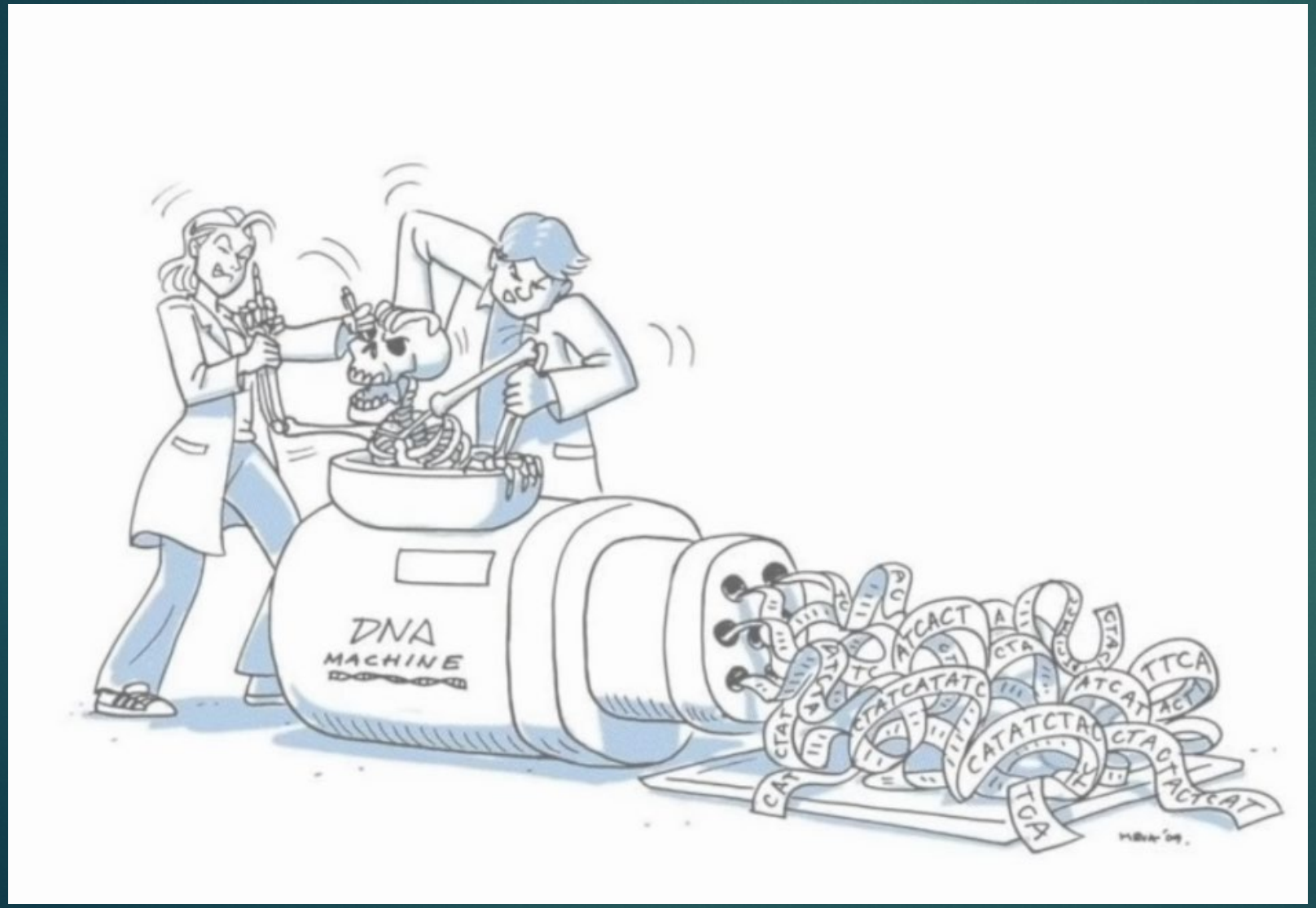
In first 12 years of work on aDNA, Pääbo found human DNA in every animal DNA sample he worked on

Chris Stringer: measured every major hominin skull in museums for his 1970s dissertation: his DNA contaminated all of them

DNA Sequencing

- DNA sequencing is the process of determining the nucleic acid sequence – the order of nucleotides in DNA.
- It includes any method or technology that is used to determine the order of the four bases: adenine, guanine, cytosine, and thymine.
- Polymerase chain reaction (PCR) is a method to rapidly make millions to billions of copies of a specific DNA sample
- Amplification refers to the production of one or more (usually millions) copies of a genetic fragment or target sequence; i.e. via PCR

aDNA Genome Sequencing: actual aDNA converted into computer digital code



Only in 2003, was sequencing ability capable of doing nuclear DNA (3 Billion letters)

aDNA research needs Ultraclean Rooms

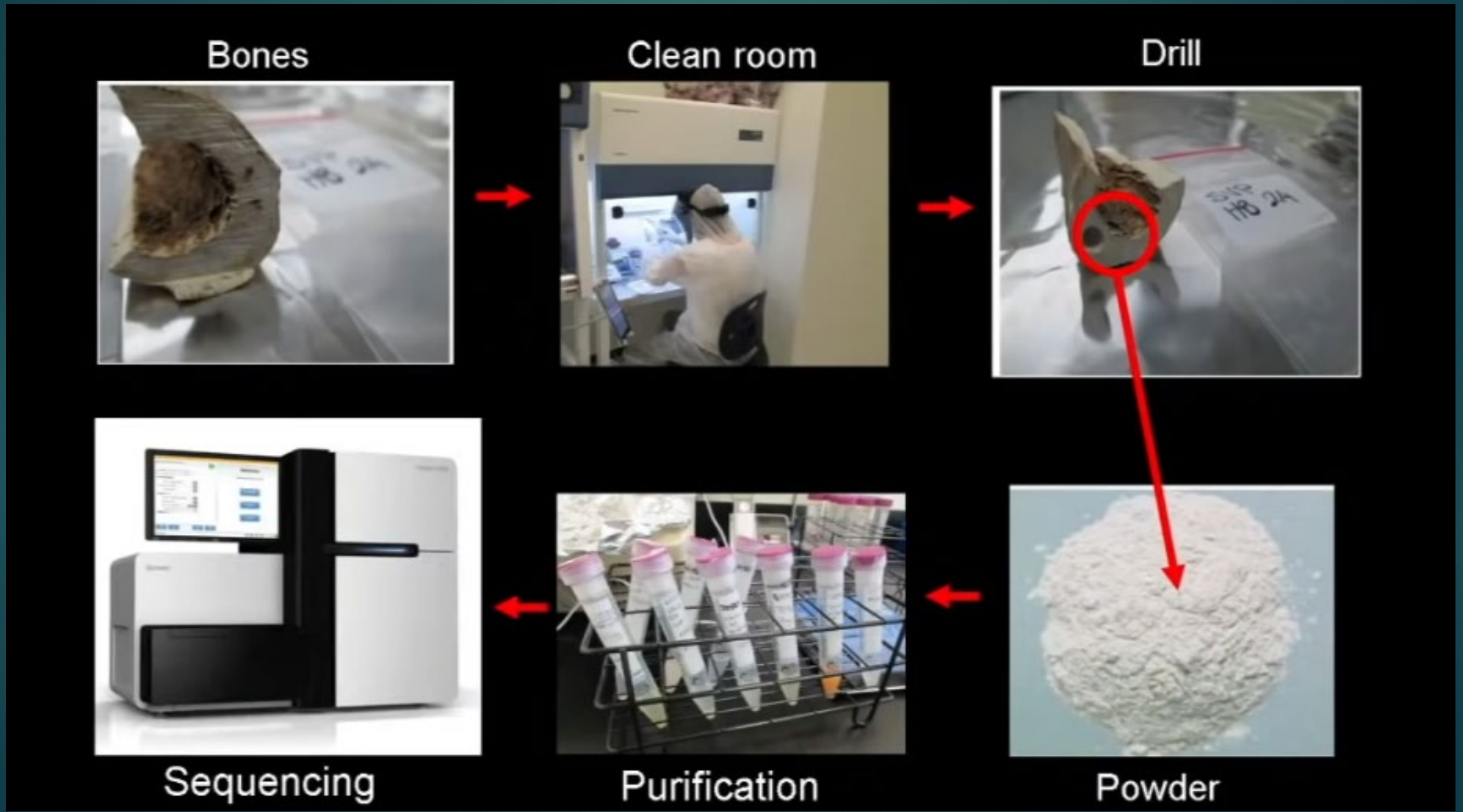
Ancient DNA Extraction

Like
silicon chip
factory
production –
no dust

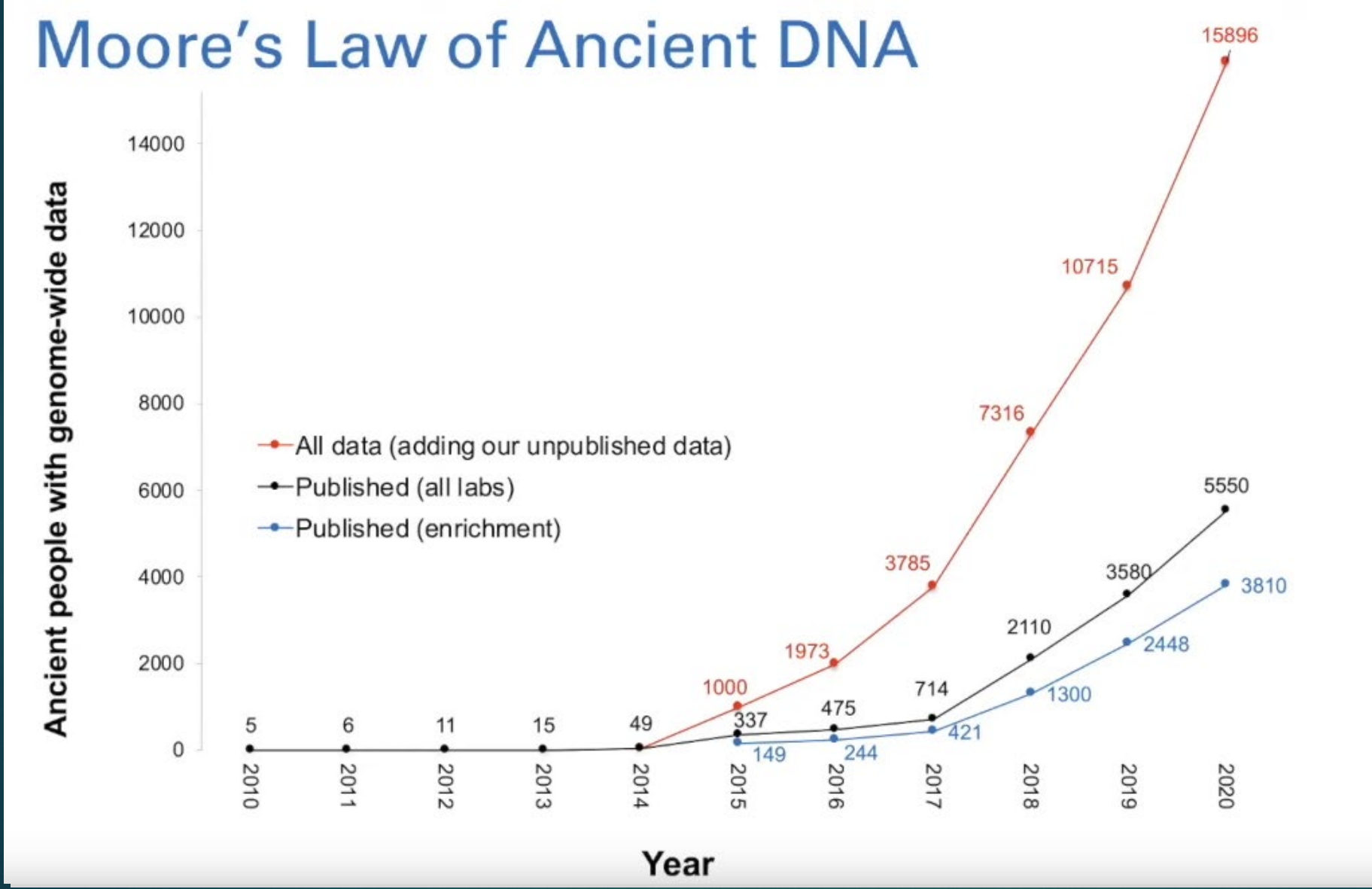
Contamination reduction: Isolation, high reverse pressure air flow, UV light

Accessing DNA in Bone – never from surface, only by drilling internally

2010 Discovery of new scientific technology (like microscope, telescope, etc.): aDNA extraction



Published ancient full genomes:



- 0 in 2009
- 50 in 2014
- 5500 in 2020
- unpublished estimate = 15,896

Reference genome: a digital summarized copy

- A reference genome is a digital nucleic acid sequence database (digital genotype), assembled by scientists as a representative example of the set of genes in one idealized individual organism of a species; usually based on multiple real genomes

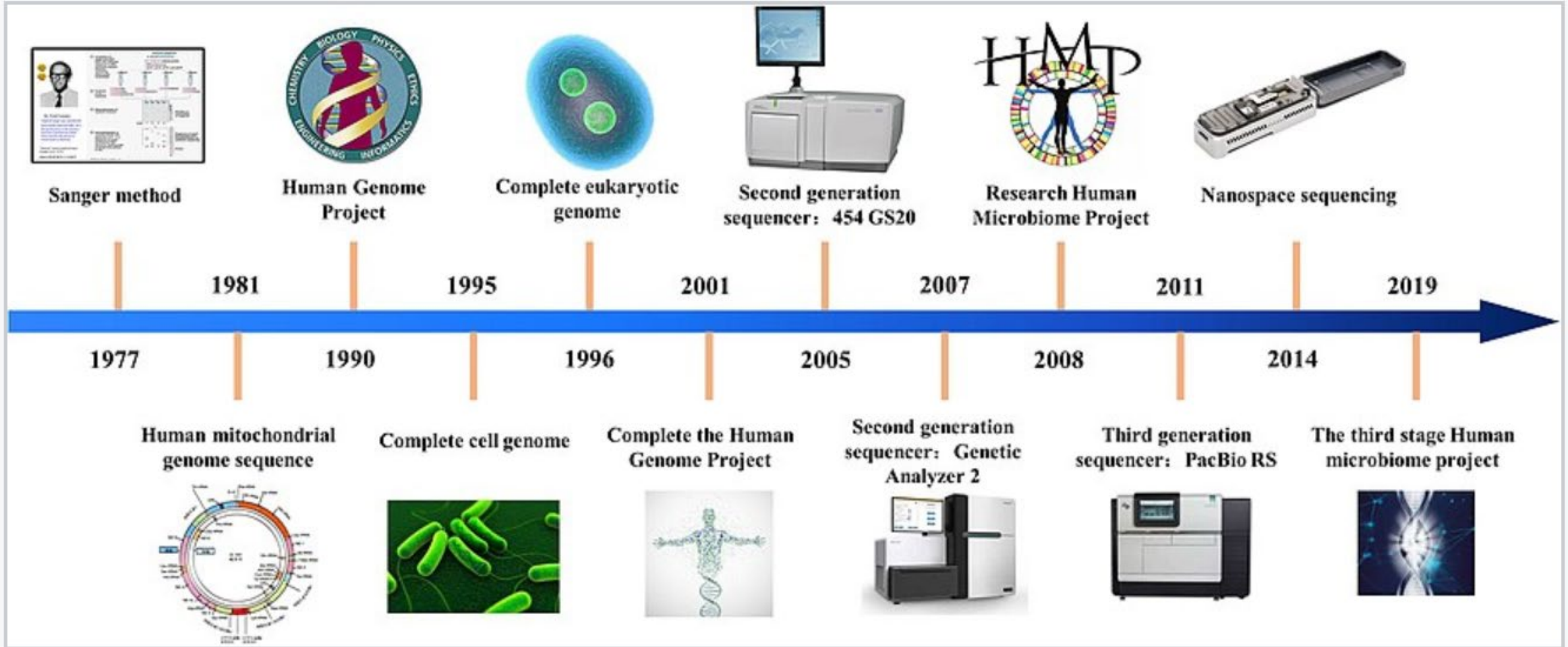
- As they are assembled from the sequencing of DNA from a number of individual donors, reference genomes do not accurately represent the set of genes of any single individual organism. A reference provides a mosaic of different DNA sequences from each donor.

- There are now reference genomes for multiple species.

Reference genome: a digital summarized copy

- All reference genomes are updateable.
- The most recent is the Human Reference Genome, *GRCh38*, from the Genome Reference Consortium is derived from thirteen anonymous volunteers. First version (1990-2003) had roughly 150,000 gaps. 13 years, $2.7 Billion
- Reference genomes are typically used as a guide on which new genomes are built. A basic comparison step in DNA sequencing

High-throughput sequencing (HTS) methods



History of sequencing technology[51]

Illumina – industry leader

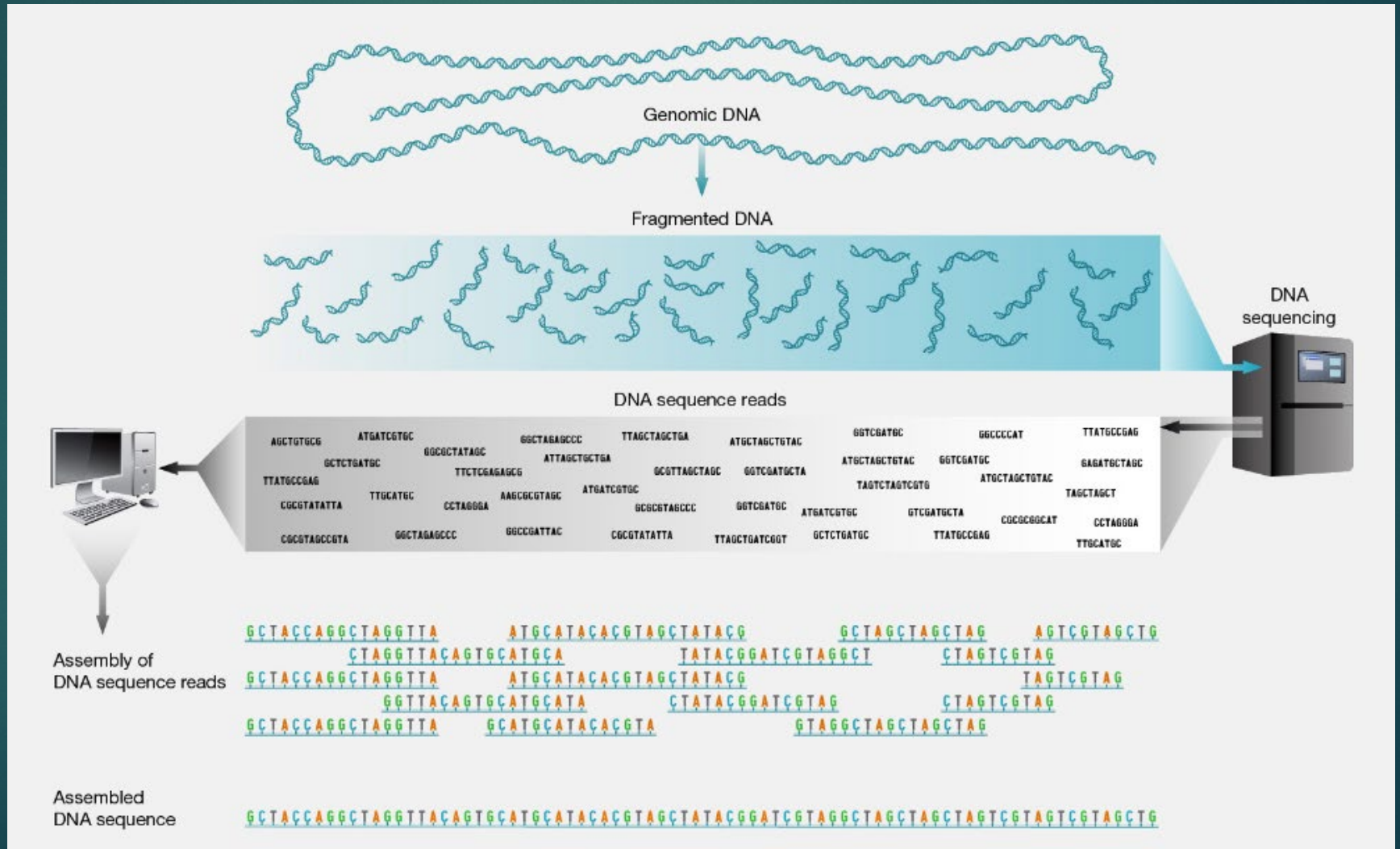
2022: Latest - NovaSeq X Series

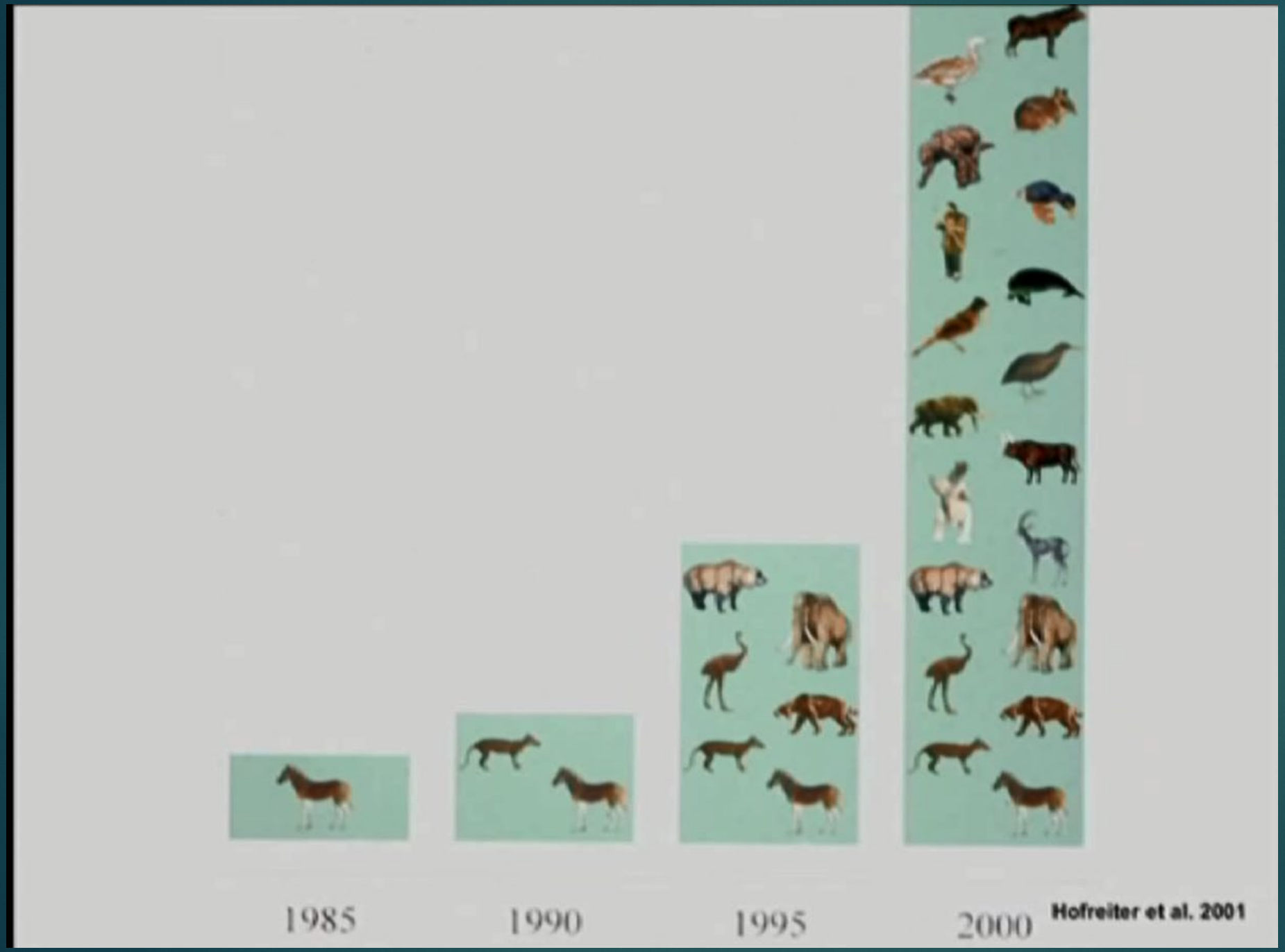
26 billion single reads per flow cell.

Cost of machine = $1.25 M

The Illumina platform has largely outcompeted the various other commercial options, primarily owing to its massive output of short DNA reads

Shotgun sequencing: randomly breaking up the genome into small DNA fragments that are sequenced individually. A computer program looks for overlaps in the DNA sequences, using them to reassemble the fragments in their correct order to reconstitute the genome. Each incorporated nucleotide is identified by its fluorescent tag.



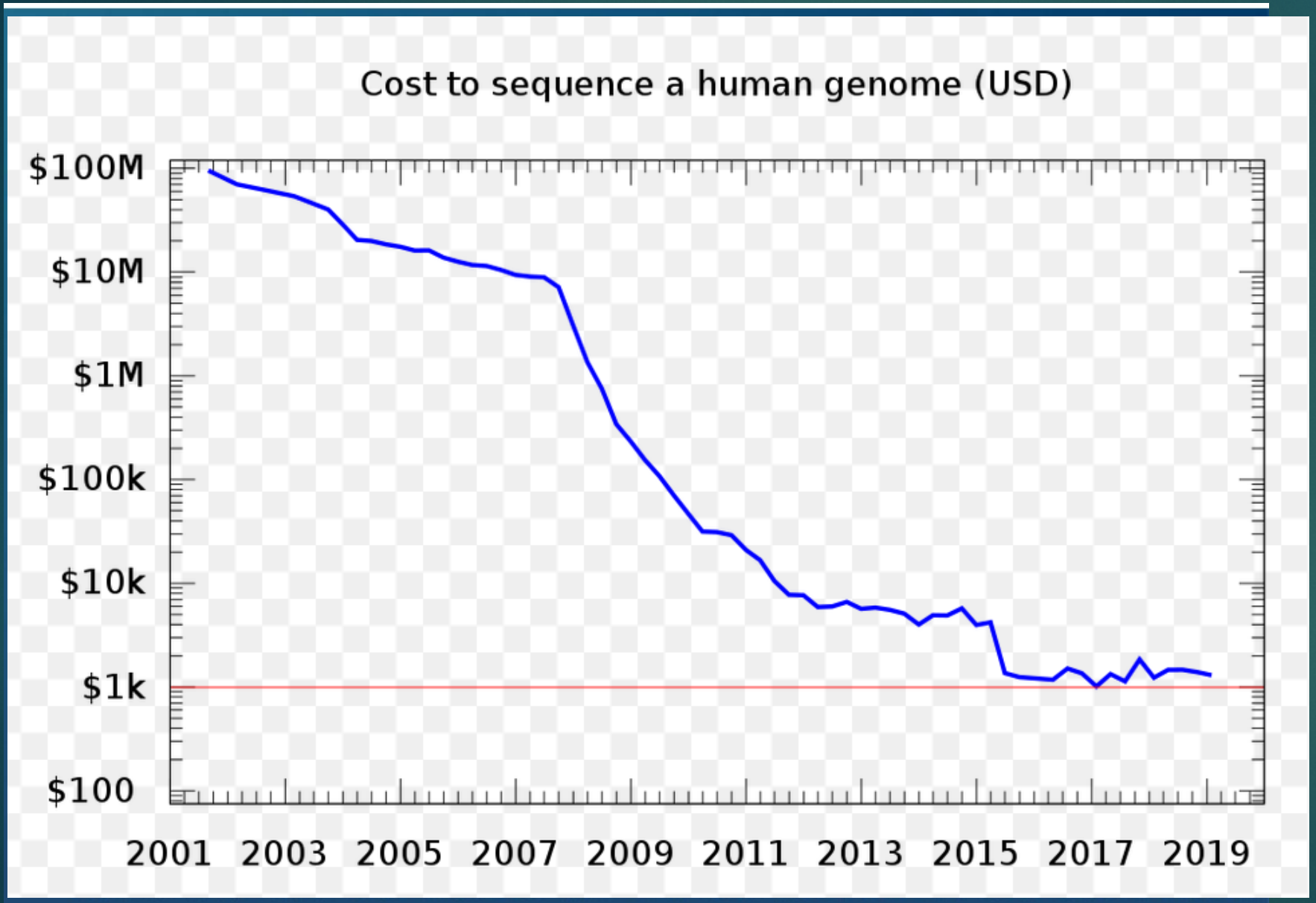


2018: 3,500 species of complex life; but only about 100 have been sequenced at “reference quality”

2018: 181 horticultural plants

2019: 1,100 plant species

2021: project to genotype all vertebrate species



1 genome = 400 GB of raw data

Amazon & Google will store your genome for $25 a year

85,000 full human genomes currently

Australian = 100,000 Genome Project

France = plan 235,000 WGS a year

GenomeAsia/100K

China = aiming for 1 Million

Whole genome for under $500

- Nuclear Genomics: $300
- Ultima Genomics of Newark, CA: $100? soon

Too much data? Need for bioinformatics

- Key difference between the PCR and NGS eras was that practitioners went from having too little to almost too much data.

- Machine sequencing technologies could produce large amounts of data that required researchers to seek or learn computational and statistical skills to interrogate it.

- From a field dominated by the laboratory scientist, aDNA research was moving into the realm of the bioinformatician/statistician.

Example of Population genetics analysis:

The allele frequency (*p*) of *EPAS1* was estimated using a maximum likelihood framework where the total number of reads across all 20 SNPs was used to calculate the allele frequency:

$$
\begin{aligned}
& l(p \mid r, t) \\
& \quad = \sum_{i=1}^{N} \log[p^2 B(r_i, t_i, 1-\varepsilon) + 2p(1-p)B(r_i, t_i, 0.5) \\
& \quad + (1-p)^2 B(r_i, t_i, \varepsilon)]
\end{aligned}
$$

The statistics courses I took at UC Berkeley did not cover this kind of analysis!!

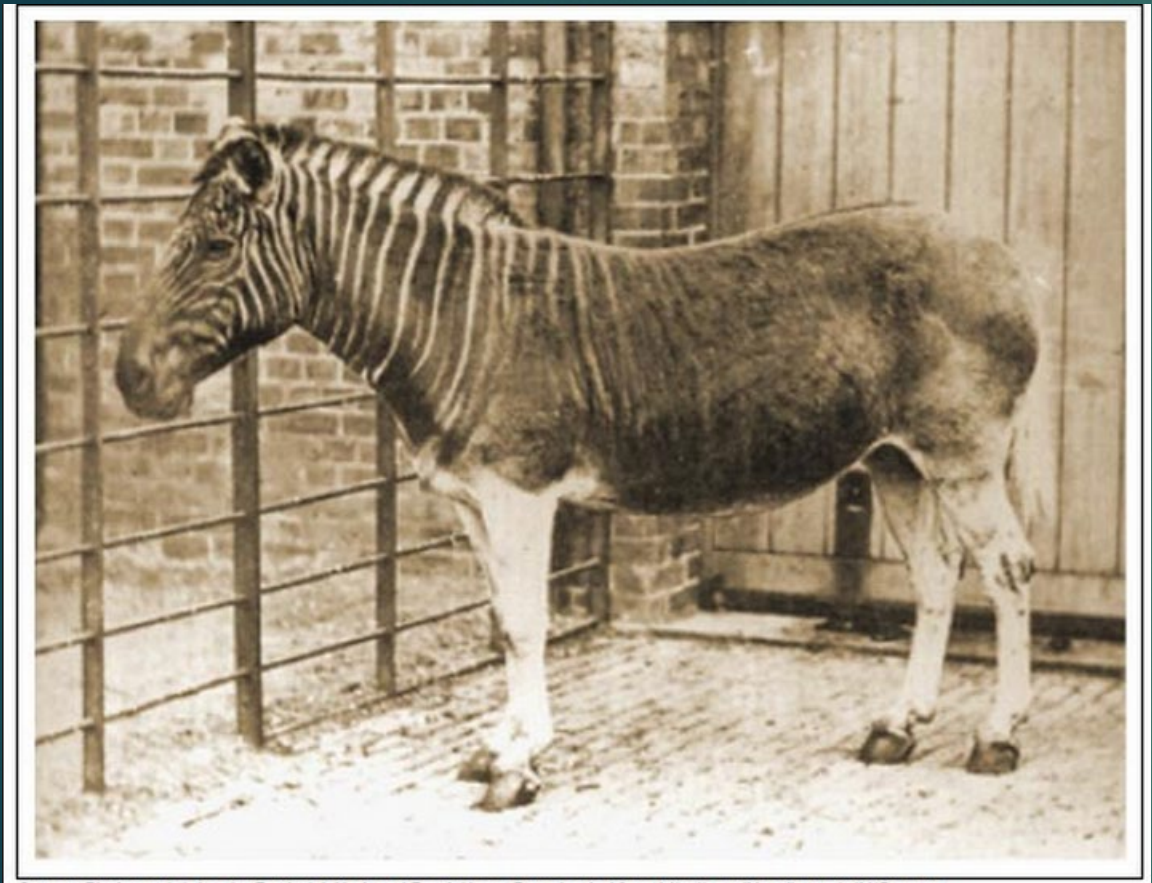
Biometric analysis

- New rule: ‘Grab as much data as you possibly can, hire a great bioinformaticist, and then start asking questions in the resulting datasets’.

- aDNA research can be seen as data mining in terms of producing data and describing its patterns without a specified hypothesis (Millar and Lambert, 2019).

- This has been criticized as a deviation from the normal scientific hypothesis-based approach.

Quagga: 1st mtDNA from extinct species, 1984



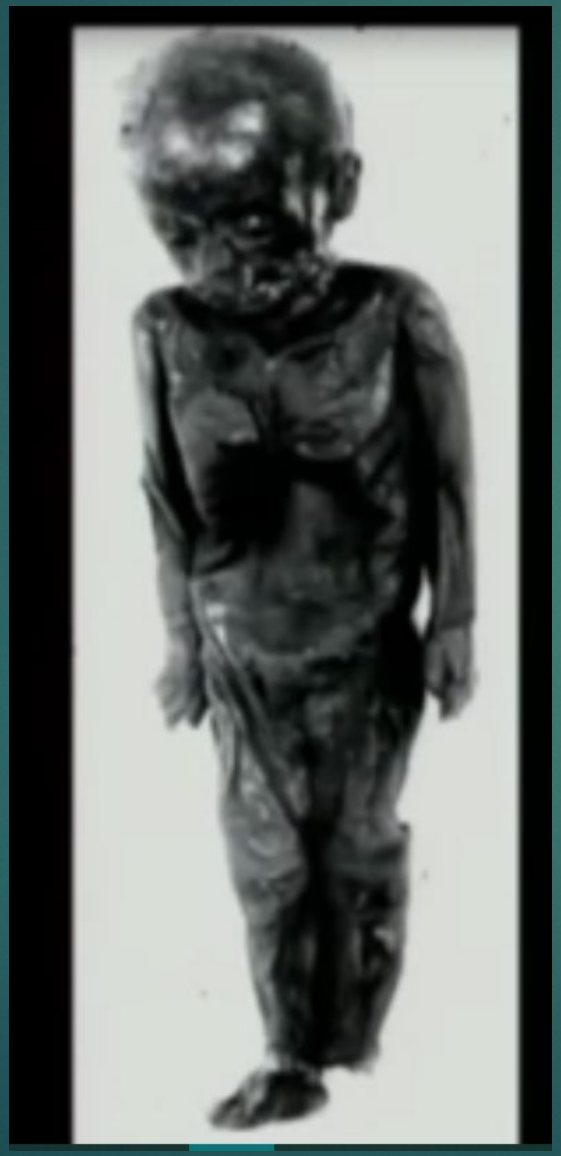
Source: Photograph taken by Frederick York and Frank Haes. Downloaded from http://en.wikipedia.org/wiki/Quagga

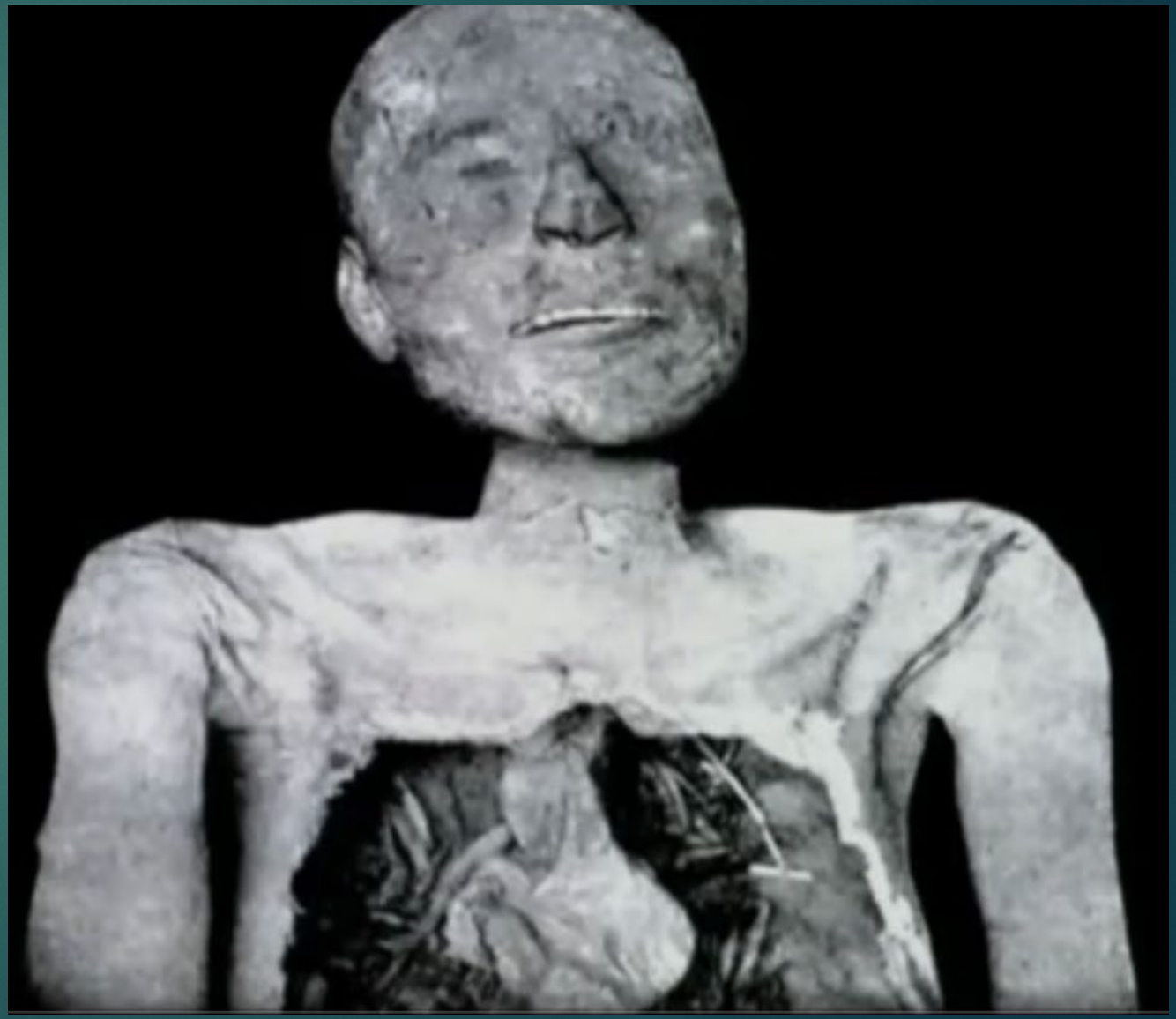
A partially striped quagga (*Equus quagga quagga*) photographed alive in 1870 in the Regent's Park Zoo in London

- Last South African zebra subspecies died at the Amsterdam Zoo in 1883.
- In 1984, Allan Wilson at UC Berkeley recovered 229 base pairs of genetic mt DNA code from a quagga.
- Achievement proved DNA could survive in dead things and spurred a new field of science: paleogenetics.

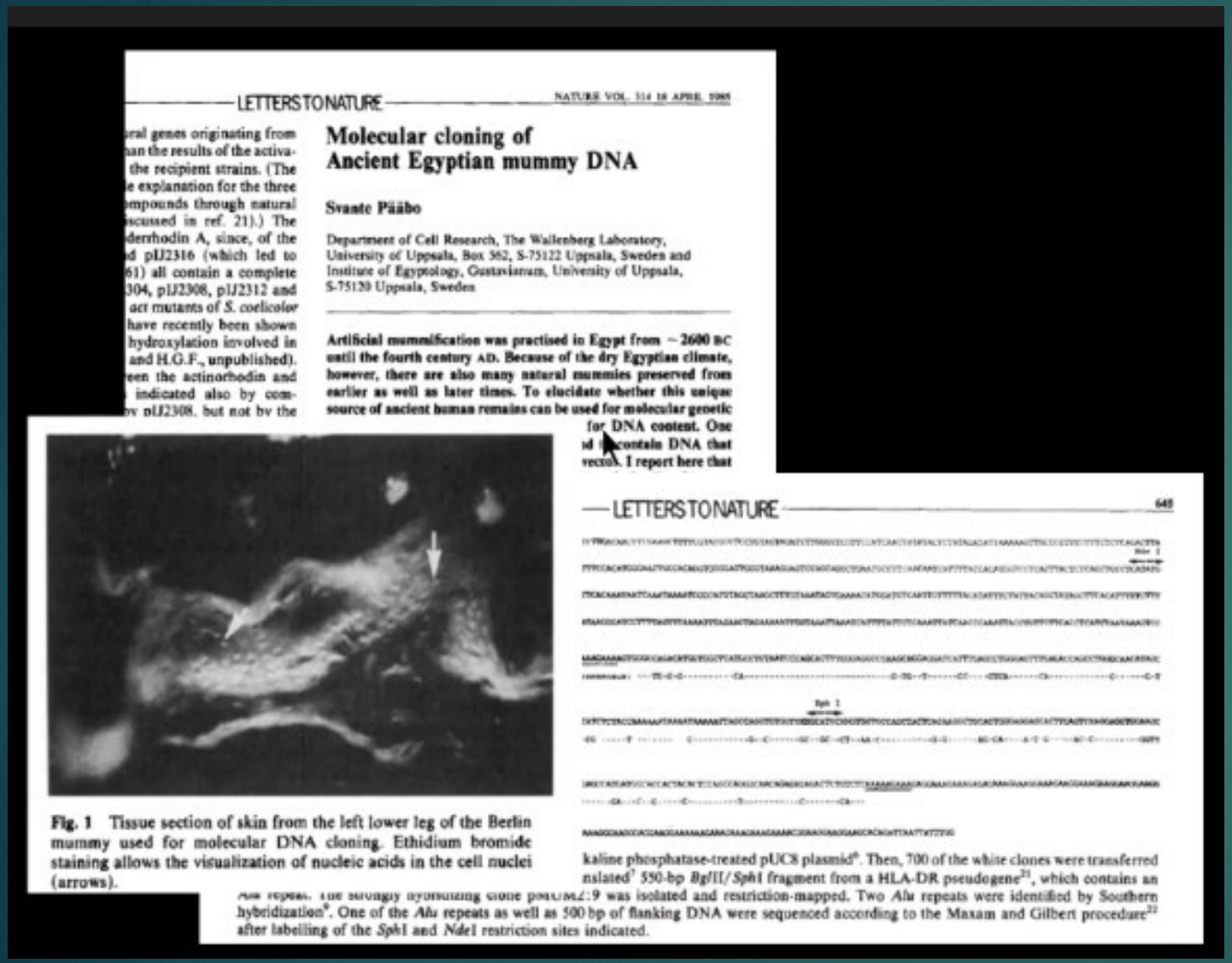
Egyptian Mummies, 1985, S. Pääbo

Very first
Human aDNA,
mtDNA, 2400 ya





Pääbo, 1985



LETTERS TO NATURE

NATURE VOL. 314 18 APRIL 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

Artificial mummification was practised in Egypt from ~2600 BC until the fourth century AD. Because of the dry Egyptian climate, however, there are also many natural mummies preserved from earlier as well as later times. To elucidate whether this unique source of ancient human remains can be used for molecular genetic

Fig. 1 Tissue section of skin from the left lower leg of the Berlin mummy used for molecular DNA cloning. Ethidium bromide staining allows the visualization of nucleic acids in the cell nuclei (arrows).

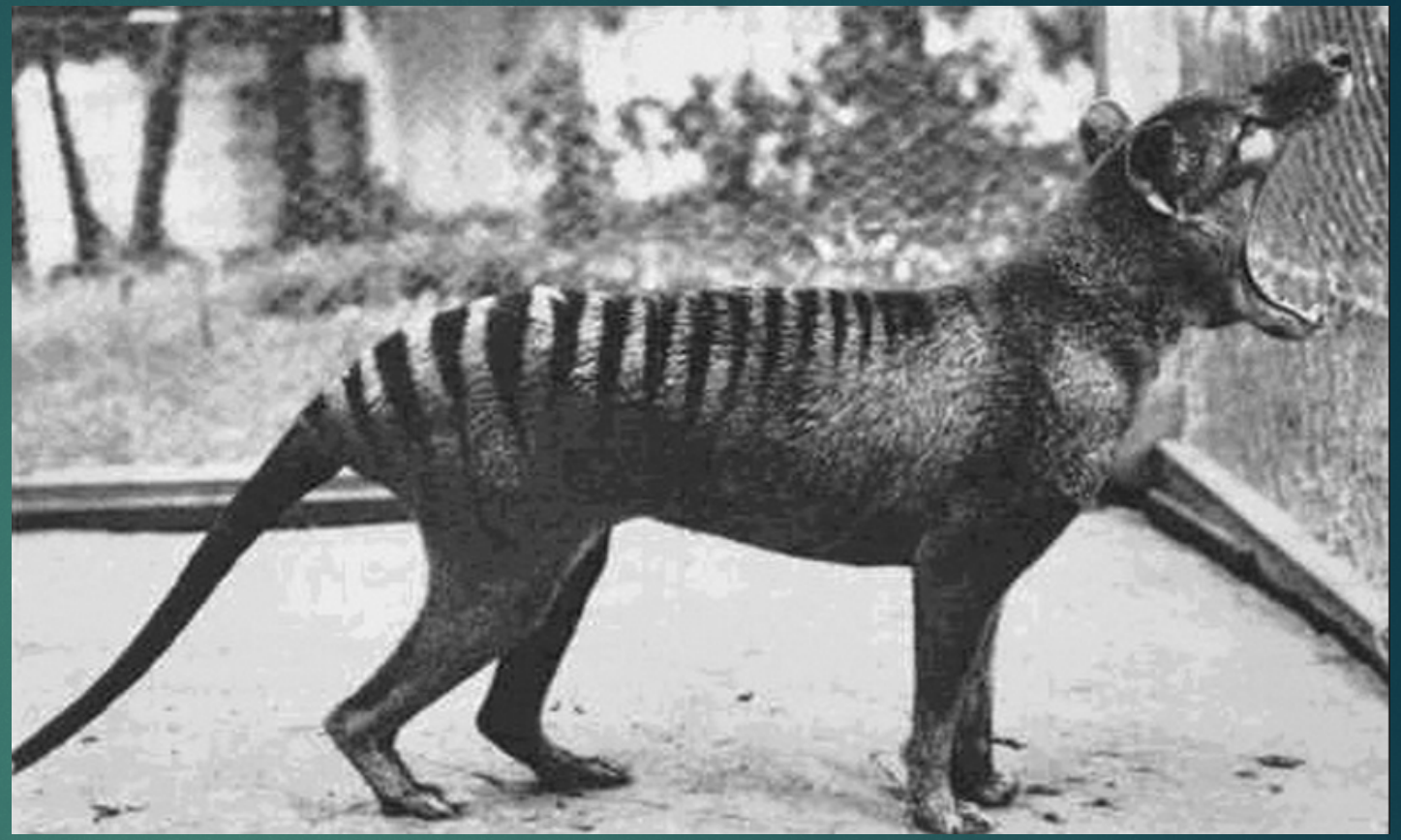
LETTERS TO NATURE 645

kaline phosphatase-treated pUC8 plasmid[6]. Then, 700 of the white clones were transferred nslated[7] 550-bp *Bgl*II/*Sph*I fragment from a HLA-DR pseudogene[21], which contains an hybridization[8]. One of the *Alu* repeats as well as 500 bp of flanking DNA were sequenced according to the Maxam and Gilbert procedure[22] after labelling of the *Sph*I and *Nde*I restriction sites indicated.

Later discovered to be contaminated DNA

For the next 15 years, Pääbo & his lab became obsessed with defeating contamination in aDNA research

1989: Marsupial “Wolf”



- Marsupial Wolf, Thylacine
- Largest carnivorous marsupial in the world
- Species extinct in 1936
- 219 bases of mitochondrial mtDNA from museum specimen

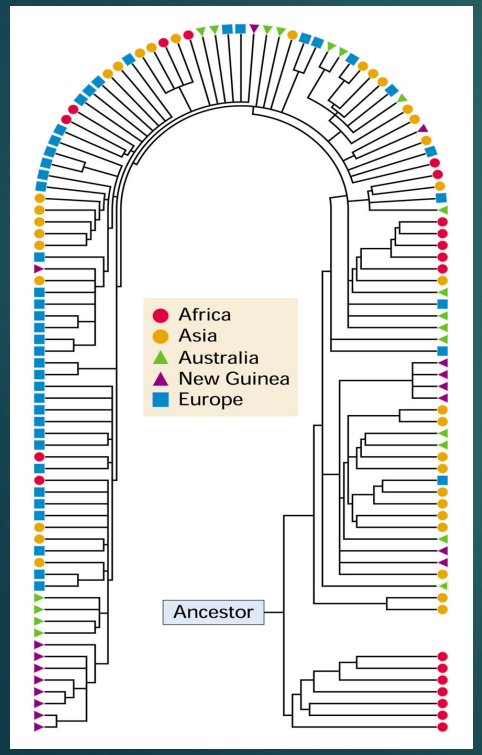
Thomas, RH, et al.,. Nature,1989

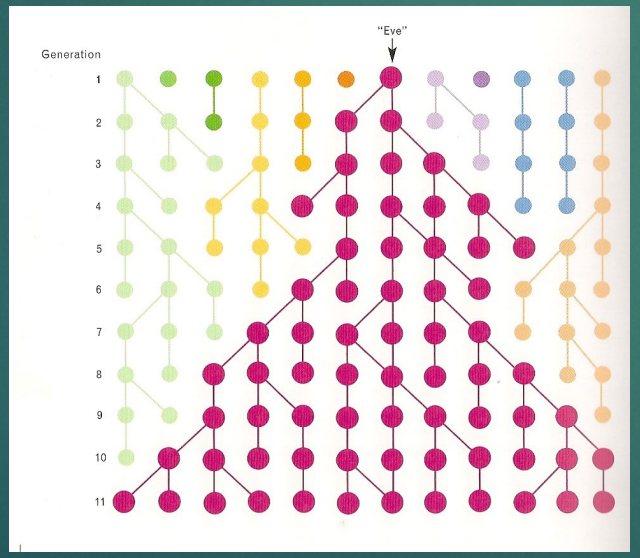
1987: Rebecca Cann: Mitochondrial Eve hypothesis

- American biochemist
- 1987: *Nature* article, elaborated the mitochondrial Eve hypothesis
- Claims a recent (ca. 200 Ka (99-148 Ka) origin for all modern humans based on a study of mtDNA haplotype links.
- We are African by DNA

Death blow for multiregionalism

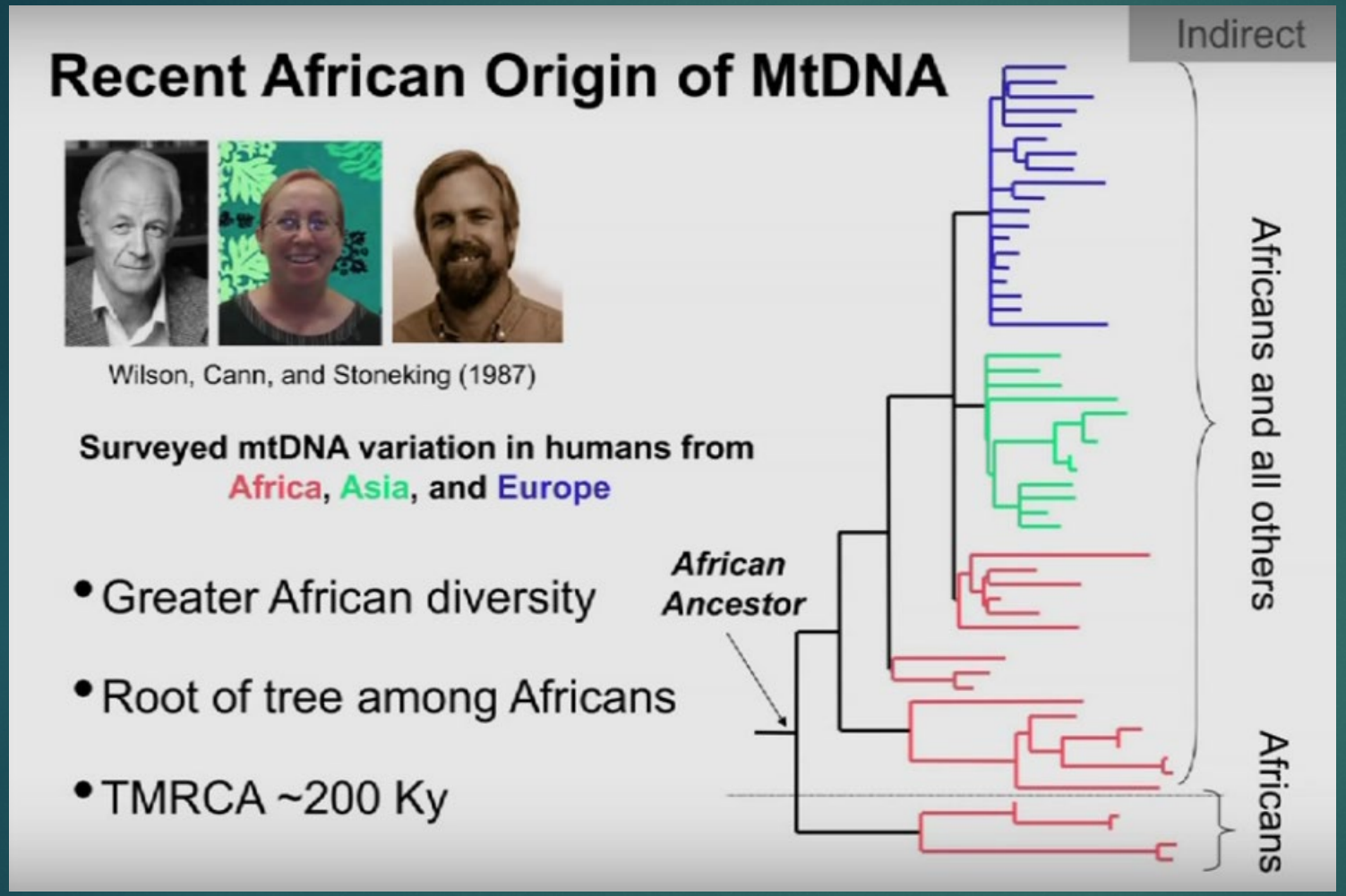
1987: Mitochondrial Eve/Most Recent Common Female Ancestor Hypothesis – not “first woman”; not LCA of *H. sapiens*; 1 woman among many who had 2 daughters; she is a ancestorial genetic phylogenetic estimate: a mathematical estimate of how far back current variants of mitochondrial DNA must go in an unbroken daughter-mother-grandmother, etc. line to converge on a single individual.
Mt haplogroup is at the root of the mtDNA phylogenetic tree





There is also a “Y-chromosomal Adam” = man from whom all living Humans are patrilineally descended (120-156 Ka)

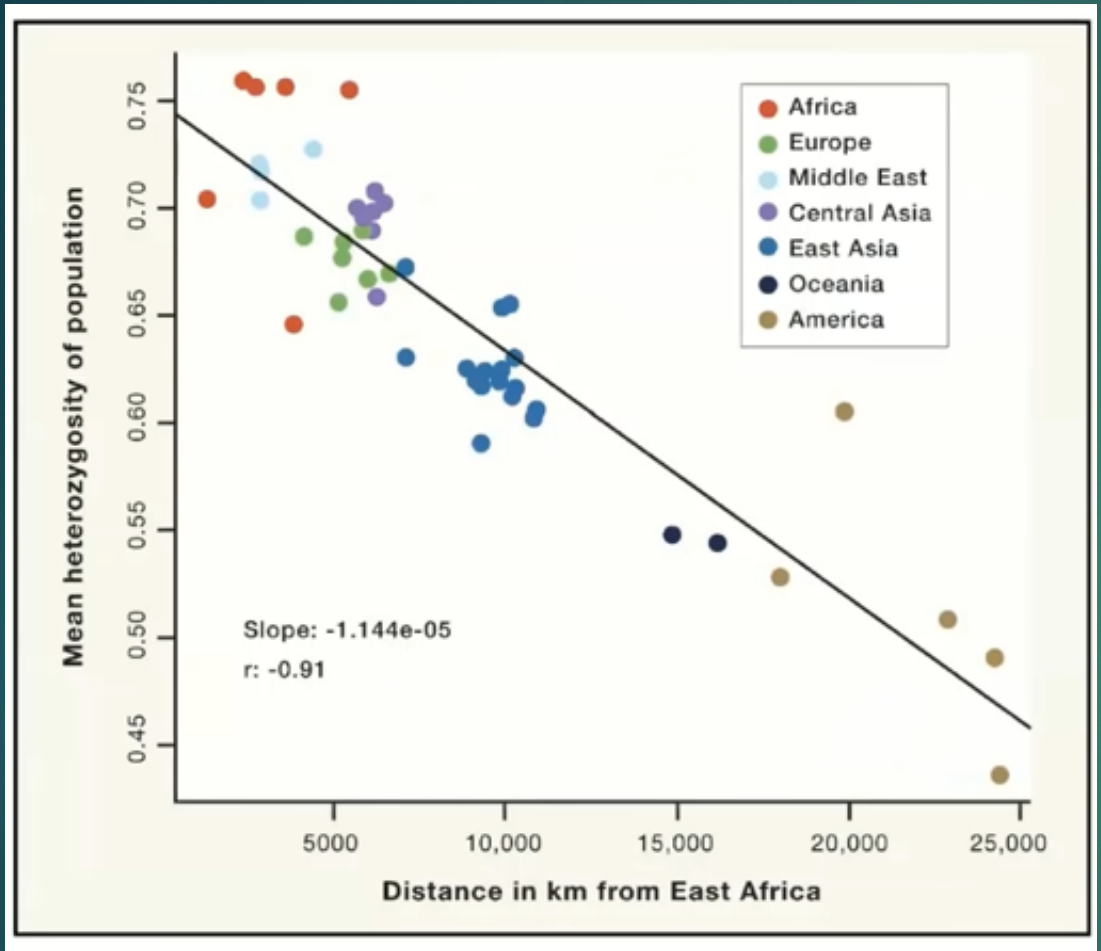
Landmark study of 1987 MRCA: mt Eve



- 1 - African lineages are longer and more mutationally diverse – more evolutionary time
- 2 – Root of tree is African; mtDNA traces to single African woman at 200 Ka

- Same results as 2000 paper that looked at whole genomes

We are all Africans



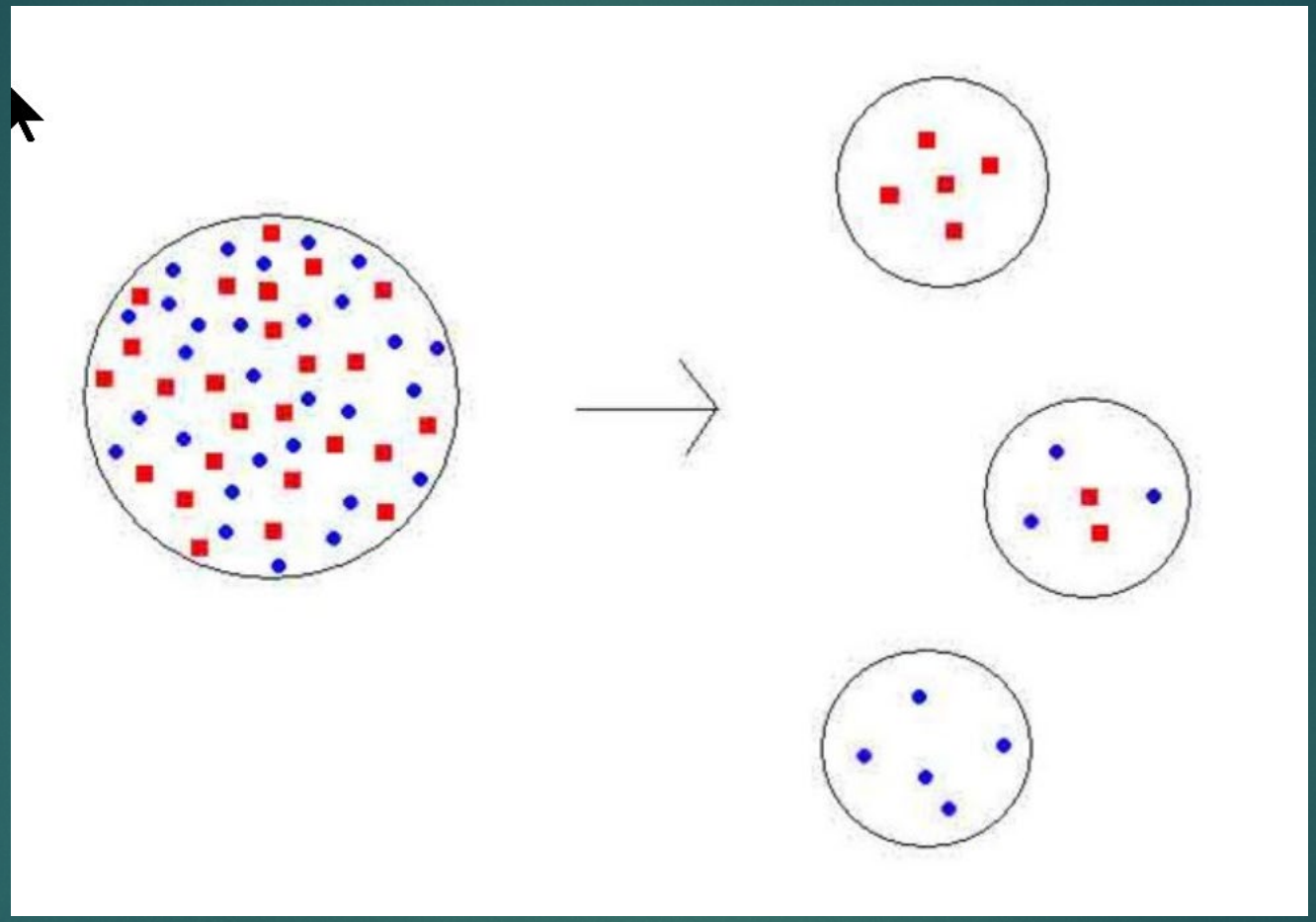
·Africa = richest genetic variation

·Heterozygosity (genetic variation) decreases with walking distance from East Africa

Founder effect/bottlenecks: group that moves away always has only a subset of total original genetic variability

Founder effect

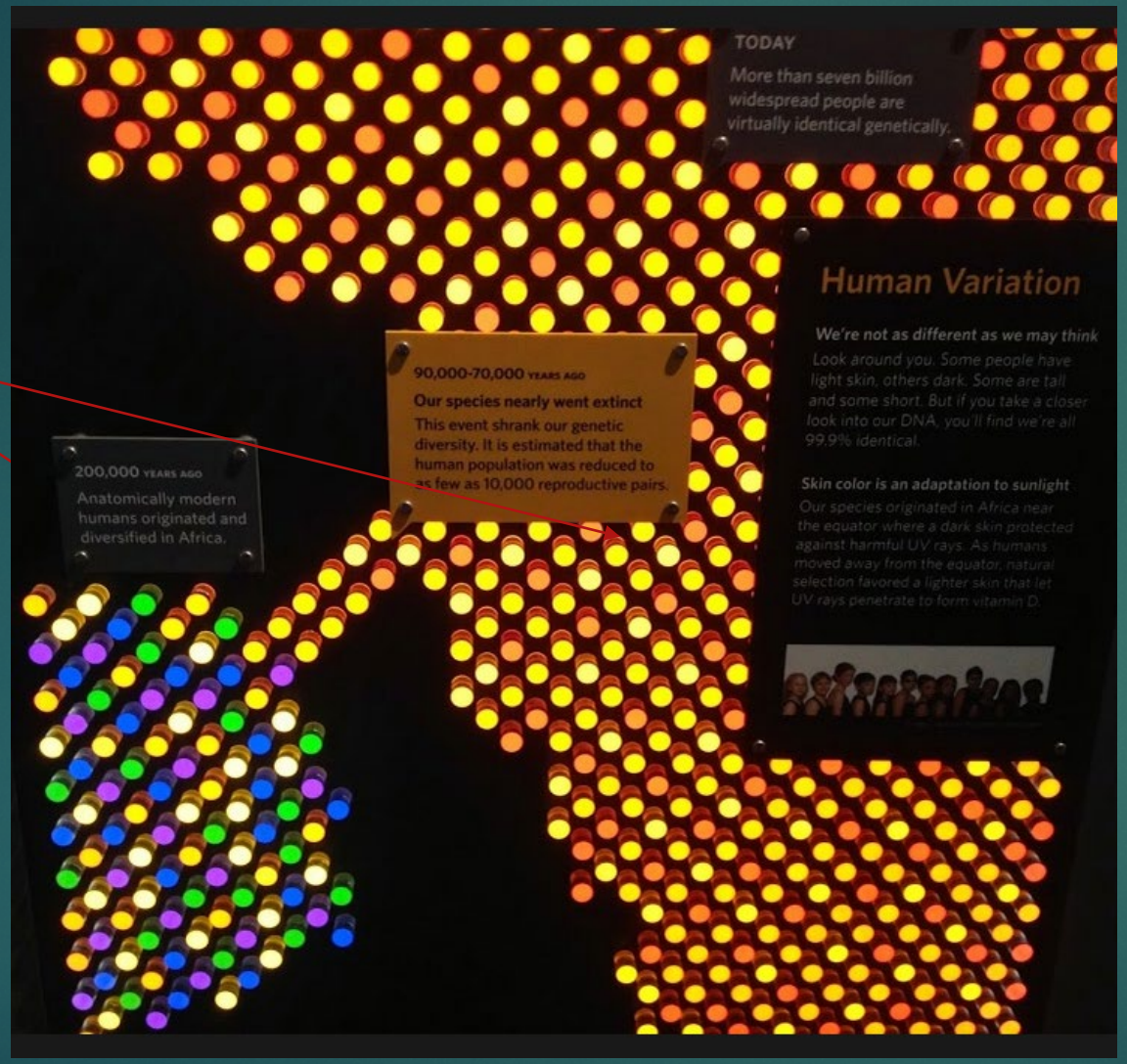
Most genetically diverse populations are in Africa



Genetic variability is reduced In each new group produced by founder effect: each new group has only the new founder's genetic mix

Human Odyssey Bottleneck exhibit: Loss of genetic variance

Africa



<u>Time course exhibit</u>: Africa at 250 Ka;

Bottleneck in Africa c 70-90K

Rest of World today with founder effects

Founder effect: reduction of genetic variation when small group starts a new population

- The Founder Effect = a type of bottleneck = a type of genetic drift describing the loss of the allelic variation that accompanies founding of a new population from a very small number of individuals (from a larger source population).

- Only a small subset of the genetic diversity of the source population is likely to be included in the new population, and the relative frequencies of these alleles may be very different from what they had been before

- Native Americans have lower diversity than Asians who have lower genetic variation than Africans

Founder effect via migrations

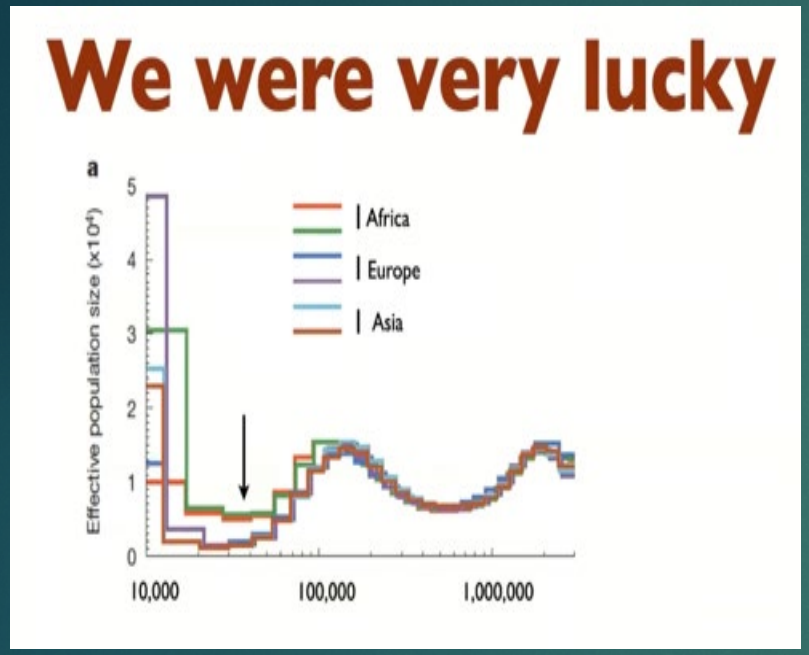
- A founder effect occurs when a new colony is started by a few members of the original population. This small population size means that the colony may have:
  - reduced genetic variation compared to the original population.
  - a non-random sample of the genes in the original population.
- For example, the Afrikaner population of Dutch settlers in South Africa is descended mainly from a few colonists.
  - Current Afrikaner population has an unusually high frequency of the gene that causes Huntington's disease

Founder Effect

- Classic African bottleneck at ~70 Ka = not due to super volcano Mt. Toba explosion in Sumatra in 74K; massive climate change

- Founder Examples:
  - polydactyly among Amish communities
  - Blue people of Kentucky (poor hemoglobin)
  - Presenilin 1 early familial Alzheimer's in Colombia
  - Mutiny on the Bounty & Pitcairn island survivors

The bottleneck: 12 K population size at 60-70 Ka



2023 study: 3000-4500 people

A founder event (bottleneck) in East Asian and European populations, associated with the human dispersal out-of-Africa event around 60 Ka

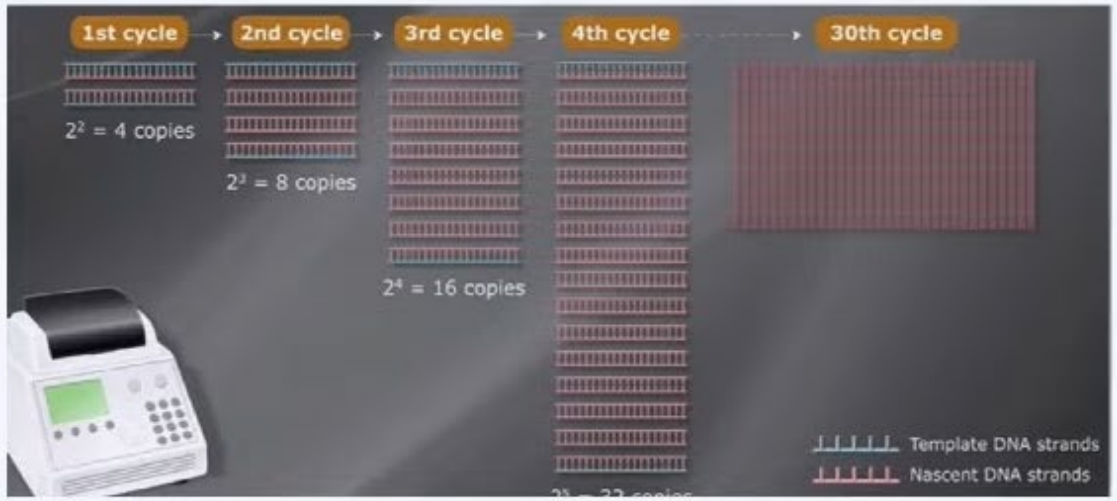
Based on n =12 MH genomes -- How many ancient individuals produced the variability you now see in these modern MH genomes

Effective population size (breeding pairs) at 60 Ka across Africa, reduced genetic variability

Just like Ns, MHs could have crashed and burned - No evolutionary preferential destiny for us = we were lucky

1983 - PCR: Mass copying of DNA: Nobel Prize in 1993

1st methodological advance; PCR



Taking advantage of the system the cell uses to copy genomes, but doing it in a tube.

Invented in 1983 and awarded the Nobel Prize in 1993.

Polymerase Chain Reaction (PCR): PCR involves using short synthetic DNA fragments called primers to select a segment of the genome to be amplified, and then multiple rounds of DNA synthesis to amplify that segment.

1989: Launch of Human Genome Project

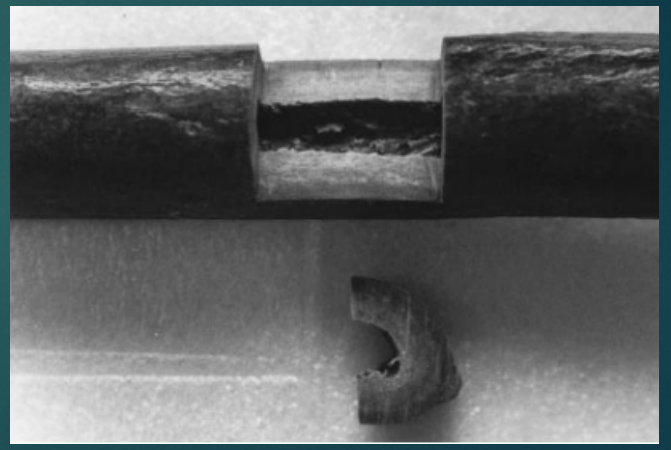
Cold Spring Harbor Lab. Library & Archives

1989: The Banbury meeting at Cold Spring Harbor Laboratory in New York before the launch of the Human Genome Project. Francis Collins and James Watson are in the top row.

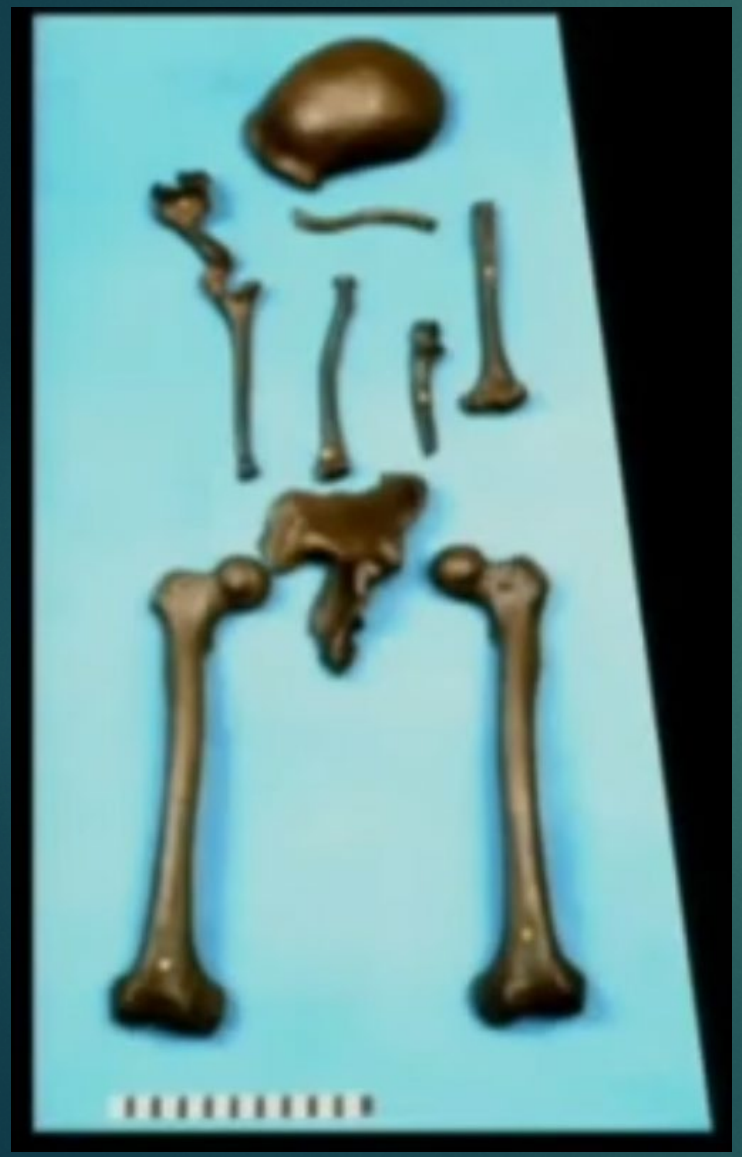
Matthias Krings:

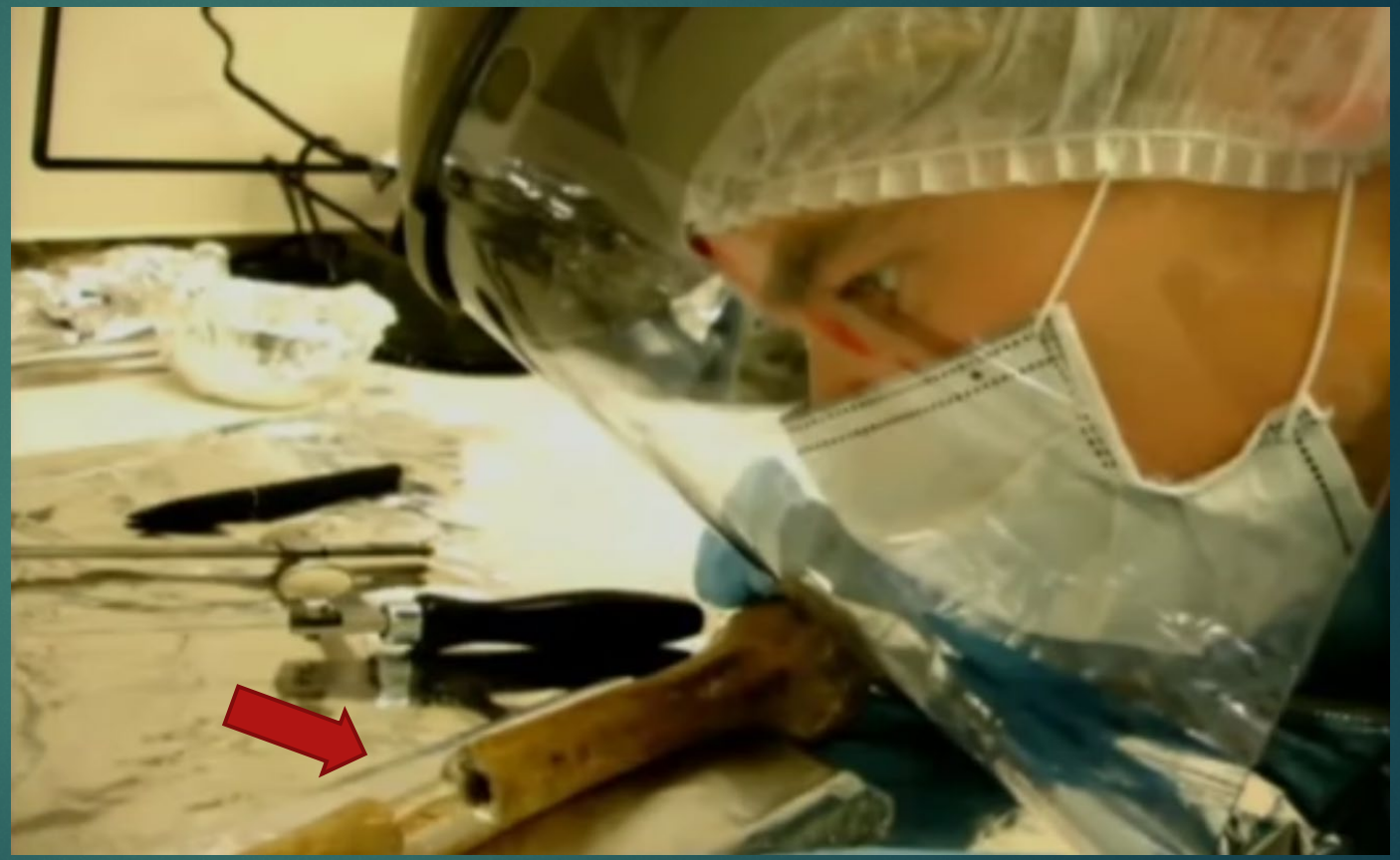
DNA Sequencing of Neanderthals

- University of Munich
- 1997: First Neandertal mitochondrial DNA sequenced (~377 bases) from Feldhofer Neanderthal, 40 Ka
- His phone message to Pääbo: “It’s not human.” = Proved modern humans and Neandertals are different “species”, which diverged from humans 690-550 Ka ago



1997: 1st N mtDNA extraction from Feldhofer Neandertal





1997: First mt DNA from a Neandertal

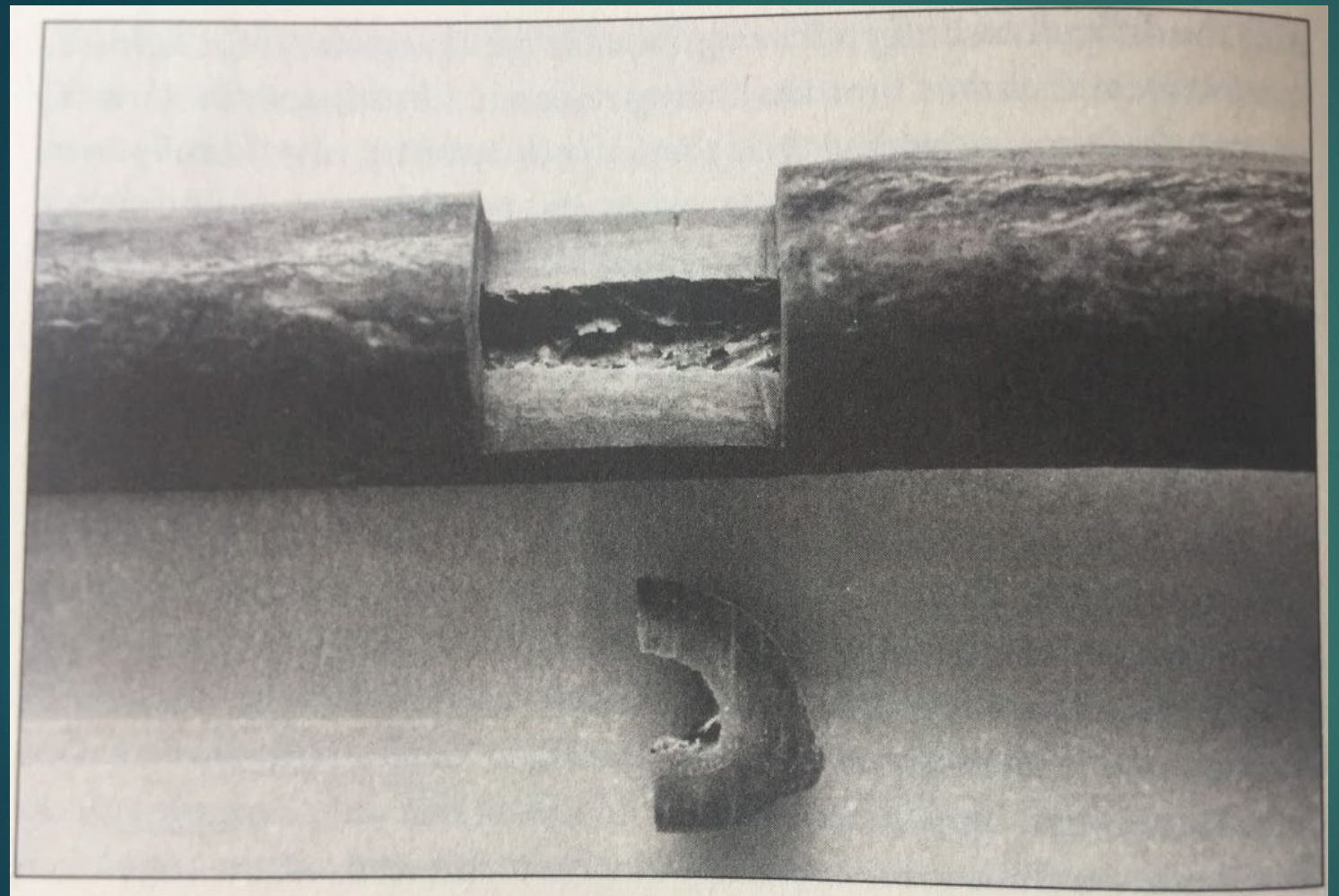


FIGURE 5.3. The right upper arm bone of the Neanderthal type specimen with the sample removed by Ralf Schmitz in 1996. Photo: R. W. Schmitz LVR-LandesMuseum Bonn

It was the original Feldhofer N individual who was first to be sampled in 1997.

At that point only mtDNA could be reliably extracted, and the result bolstered the evolutionary theory dominant at the time, which proposed that Neanderthals had arisen and remained genetically isolated in Europe.

1997: mtDNA of Feldhofer Neanderthal:
First Hominin DNA

- 1997: Pääbo retrieves DNA from original Feldhofer Cave Neanderthal; Matthias Krings isolates mtDNA; 377 bp Neandertal sequence was aligned with Cambridge MH reference sequence. The alignment shows 27 differences (24 transitions, 2 transversions, 1 deletion)

- Conclusion: N mitochondrial DNA falls outside of variation of present-day MH

- Ns were totally replaced; no N mitochondrial DNA today; no Neandertal Y chromosome today

Krings et al,. *Cell*, 1997)

History of DNA sequencing

- 2002: first mouse *Mus musculus* genome

- 2002: online Genome browsers become available: such as Ensembl and the UCSC Genome Browser

- 2002: Discovery of Oase 1: Neanderthal Great-GGGGGGrandson (Oase 1, *Romania, 40 Ka,*), the jawbone of a modern human found in 2002, contained over 99% contaminant DNA. But in 2015 researchers sequenced enough authentic DNA to show that the man had a Neanderthal ancestor a mere four to six generations back.

- 2004: field of metagenomics — the reconstruction of microbial communities DNA directly from environmental samples

https://www.nature.com/articles/d42859-020-00103-7

History of DNA sequencing

- 2005: first draft DNA sequence of a non-human primate, the chimpanzee
- 2005*: rice Oryza sativa genome;* one of the last genomes to be Sanger-sequenced, clone by clone
- 2005: Next Generation Sequencing: introduction of high-throughput, massively parallel sequencing technologies able to sequence a bacterial genome at a fraction of the cost and time of traditional Sanger sequencing techniques

Comparison of the human and chimpanzee genomes

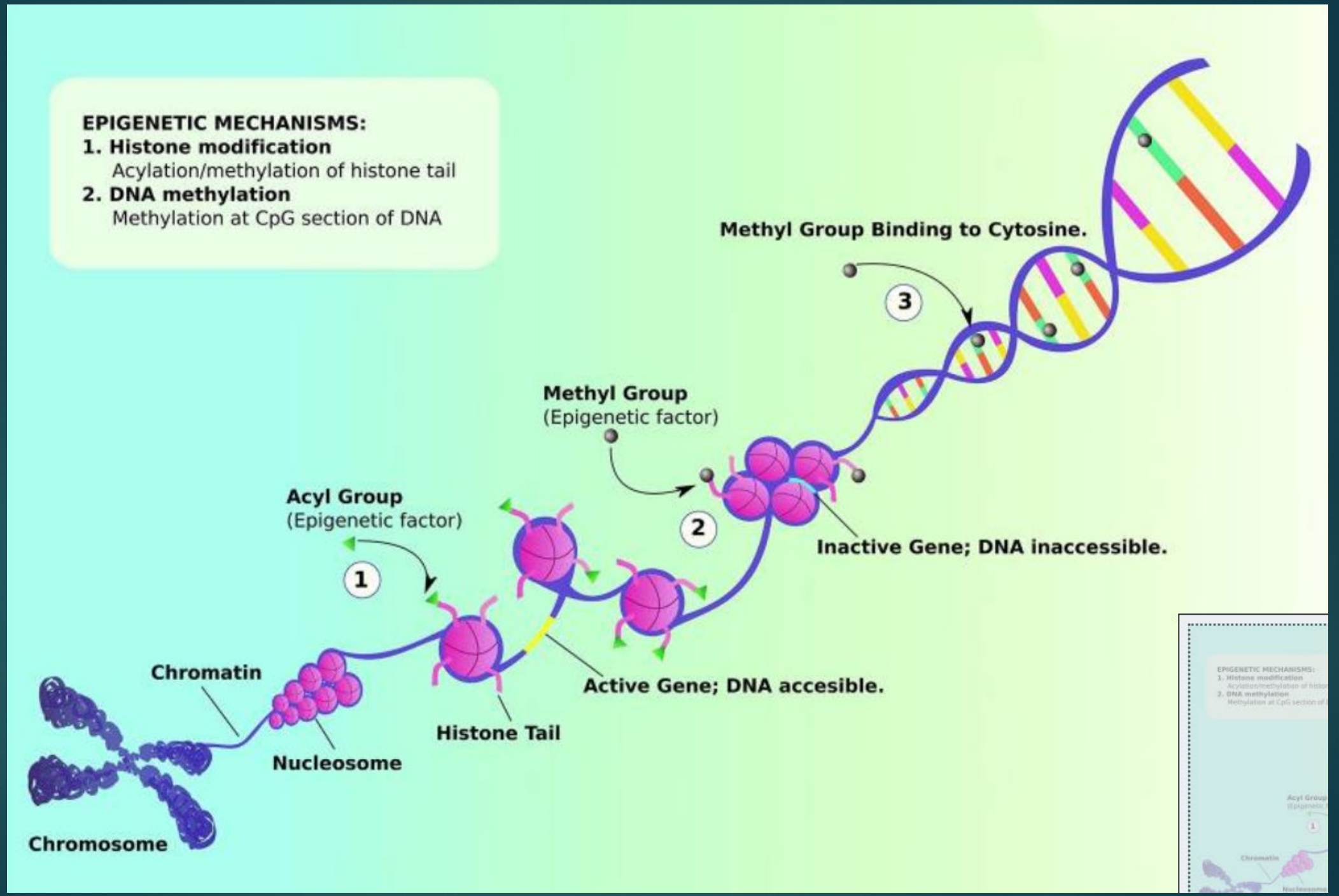
- The genome of “Clint”, the chimpanzee, was published September 1, 2005.
- 2400 million bases (of ~3160 million bases) were sequenced
- Mean nucleotide divergence between humans and chimps was 1.06%.
- Differ by 1 chromosomal fusion (human chrom. 2) and at least 9 pericentric inversions.
- 29% of all proteins compared were identical!

Paleogenetic Studies

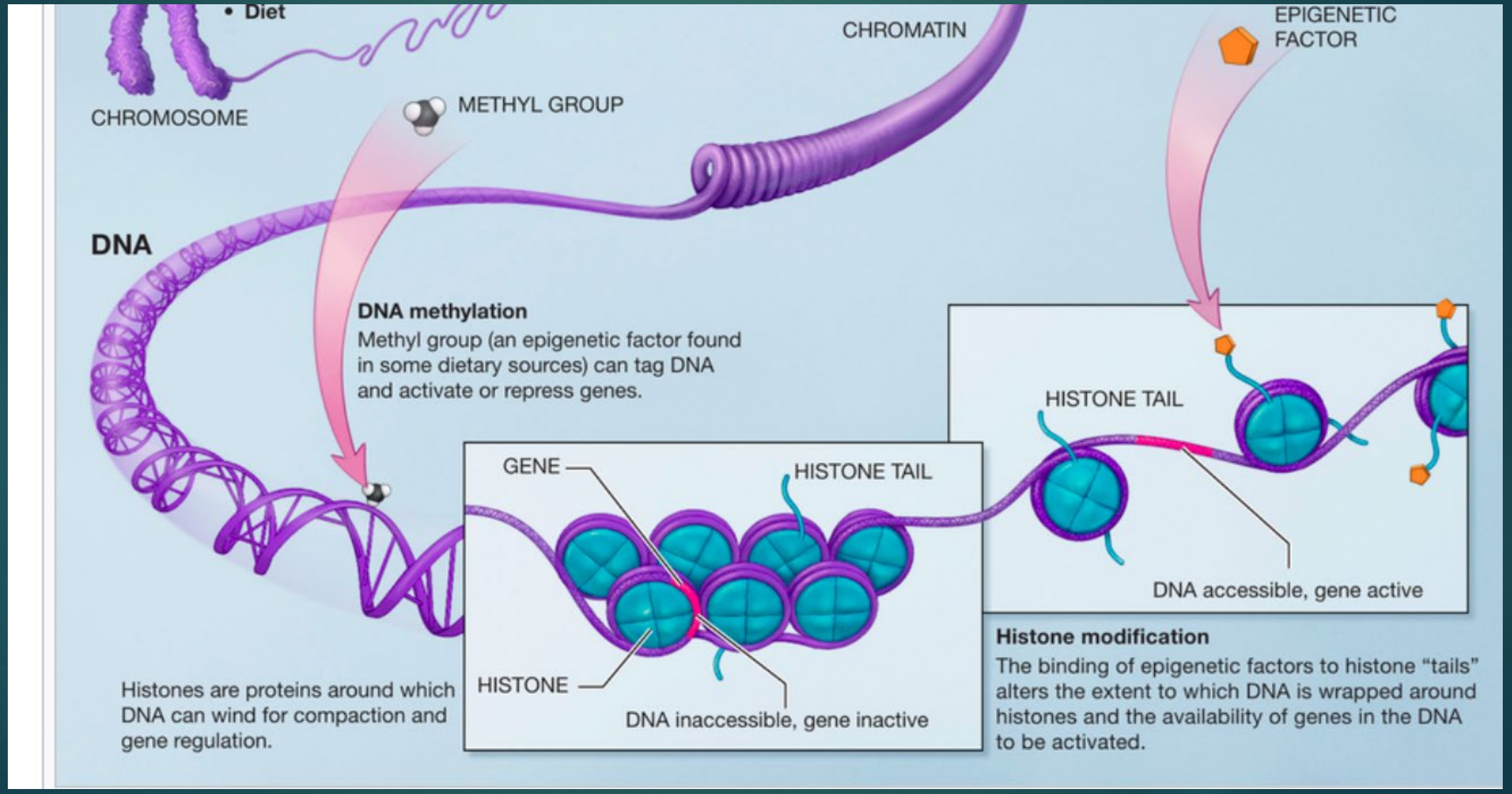
- 2006: Partial sequencing of Neandertal genomic DNA (Noonan *et al.*, *Science* **314**, 1113 (2006). Green *et al.*, *Nature* **444**, 330 (2006))
- 2007: Neandertals were in Siberia (Krause *et al.*, *Nature* **449**, 902 (2007))
- 2007: Neandertals = a red hair gene and a fair skin gene (Lalueza-Fox *et al.*, *Science* **318**, 1453 (2007))
- 2007: Neandertals and modern humans share the same variant of the language gene FOXP2 (Krause *et al.*, *Curr. Biology* **17**, 1908 (2007)

Epigenetics

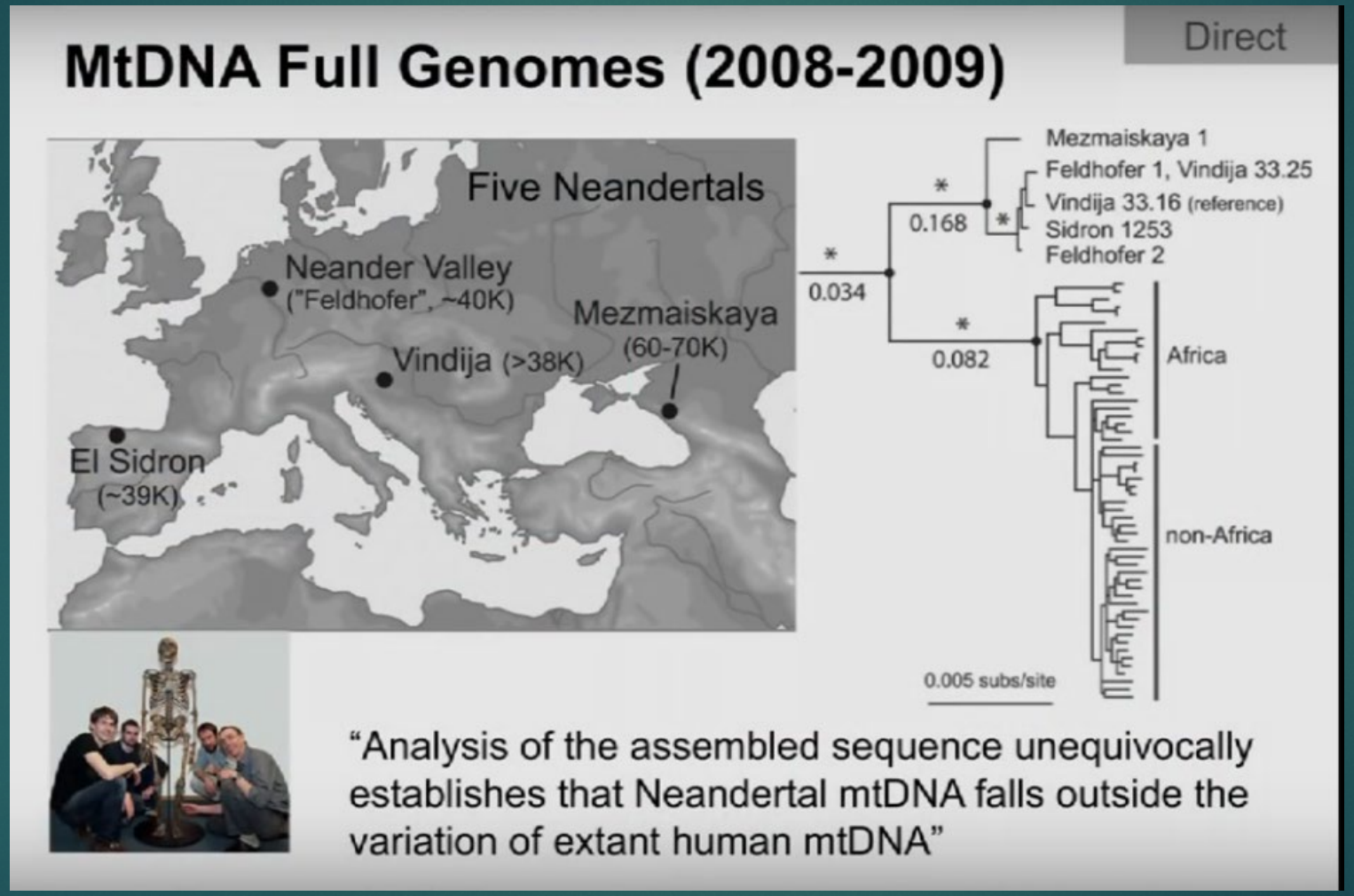
- 2007: ChIP–sequencing: determining how proteins interact with DNA to regulate gene expression -- chromatin binding patterns of different proteins – start of epigenetics
- Epigenetics refers to gene regulation, control of gene expression from noncoding areas; Tags gene via methylation and silences gene expression
- Transgenerational, therefore non-Darwinian, but “Lamarckian”
- Evidence: Holland 1944: winter starvation – starvation effects in children and grandchildren –i.e., psychiatric, obesity
- Holocaust survivors who were starved and had PTSD had epigenetic changes


EPIGENETIC MECHANISMS:
1. Histone modification
Acylation/methylation of histone tail
2. DNA methylation
Methylation at CpG section of DNA
Methyl Group Binding to Cytosine.
3
Methyl Group
(Epigenetic factor)
Acyl Group
(Epigenetic factor)
2
1
Inactive Gene; DNA inaccessible.
Active Gene; DNA accesible.
Chromatin
Histone Tail
Nucleosome
Chromosome

Loose wrap: gene activation; tight wrap: gene inactivation



2008, Pääbo's group sequenced complete MtDNA in 5 Ns: outside MH variation



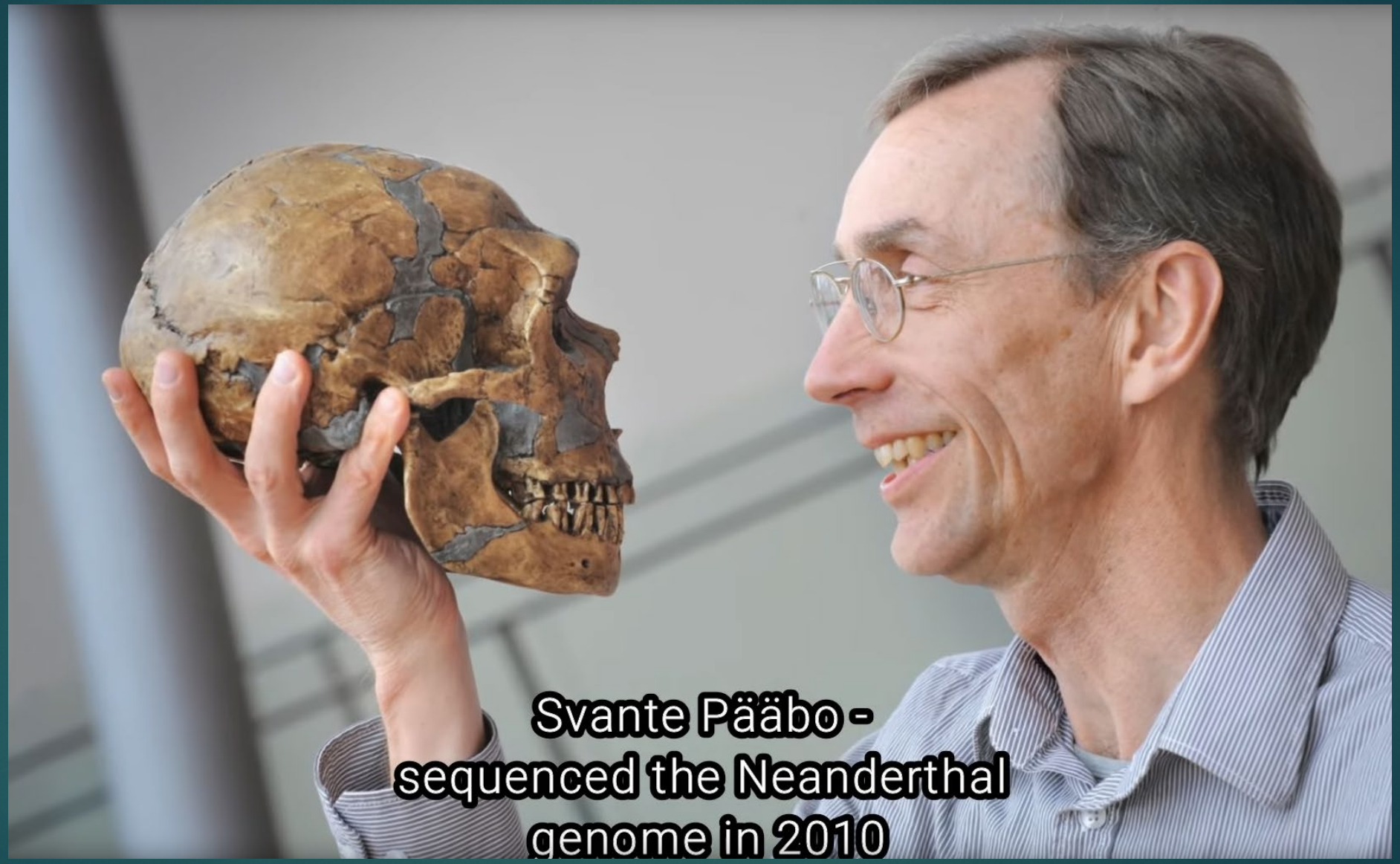
(Green *et al.*, *Cell* **134**, 416 (2008))

2009: Neandertal

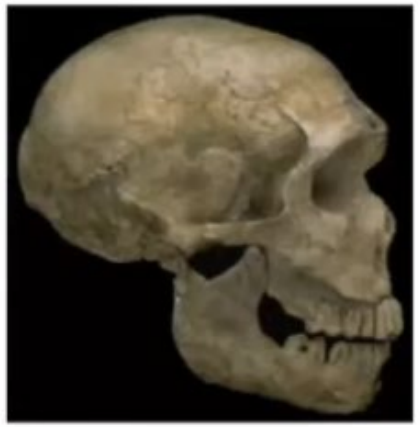
- Other studies show the existence of eastern, western and southern groups of Neanderthals (Fabre et al. 2009).
- On average, Neanderthal mtDNA genomes differ from each other by 20.4 bases and are only 1/3 as diverse as modern humans (Briggs et al. 2009).
- The low diversity signal a small population size in Ns (Briggs et al. 2009).

2010 Discoveries: Start of the aDNA revolution & paleogenetics

- 2010: First draft Neanderthal genome
- 2010: First draft Denisovan genome - first hominin species discovered solely by DNA
- Both from S. Pääbo's Leipzig Lab
- *** By 2020: 1-2% N DNA in MHs; 5-6% D DNA in Melanesians & .2% of both D & N DNA in Asians & Native Americans; .3% N DNA in Africans


Svante Pääbo -
sequenced the Neanderthal
genome in 2010

Two publications that created the field of aDNA



Svante Pääbo

A Draft Sequence of the Neandertal Genome

Richard E. Green,[1]*†‡ Johannes Krause,[1]†§ Adrian W. Briggs,[1]†§ Tomislav Maricic,[1]†§ Udo Stenzel,[1]†§ Martin Kircher,[1]†§ Nick Patterson,[2]†§ Heng Li,[2]† Weiwei Zhai,[3]†‖ Markus Hsi-Yang Fritz,[4]† Nancy F. Hansen,[5]† Eric Y. Durand,[3]† Anna-Sapfo Malaspinas,[3]† Jeffrey D. Jensen,[6]† Tomas Marques-Bonet,[7,13]† Can Alkan,[7]† Kay Prüfer,[1]† Matthias Meyer,[1]† Hernán A. Burbano,[1]† Jeffrey M. Good,[1,8]† Rigo Schultz,[1] Ayinuer Aximu-Petri,[1] Anne Butthof,[1] Barbara Höber,[1] Barbara Höffner,[1] Madlen Siegemund,[1] Antje Weihmann,[1] Chad Nusbaum,[2] Eric S. Lander,[2] Carsten Russ,[2] Nathaniel Novod,[2] Jason Affourtit,[9] Michael Egholm,[9] Christine Verna,[21] Pavao Rudan,[10] Dejana Brajkovic,[11] Željko Kucan,[10] Ivan Gušic,[10] Vladimir B. Doronichev,[12] Liubov V. Golovanova,[12] Carles Lalueza-Fox,[13] Marco de la Rasilla,[14] Javier Fortea,[14]¶ Antonio Rosas,[15] Ralf W. Schmitz,[16,17] Philip L. F. Johnson,[18]† Evan E. Eichler,[7]† Daniel Falush,[19]† Ewan Birney,[4]† James C. Mullikin,[5]† Montgomery Slatkin,[3]† Rasmus Nielsen,[3]† Janet Kelso,[1]† Michael Lachmann,[1]† David Reich,[2,20]*† Svante Pääbo[1]*†

2010: Discovery of interbreeding between Neandertals and modern humans

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

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

2010: Discovery of the "Denisovans," a previous unknown archaic population who also interbed with modern humans

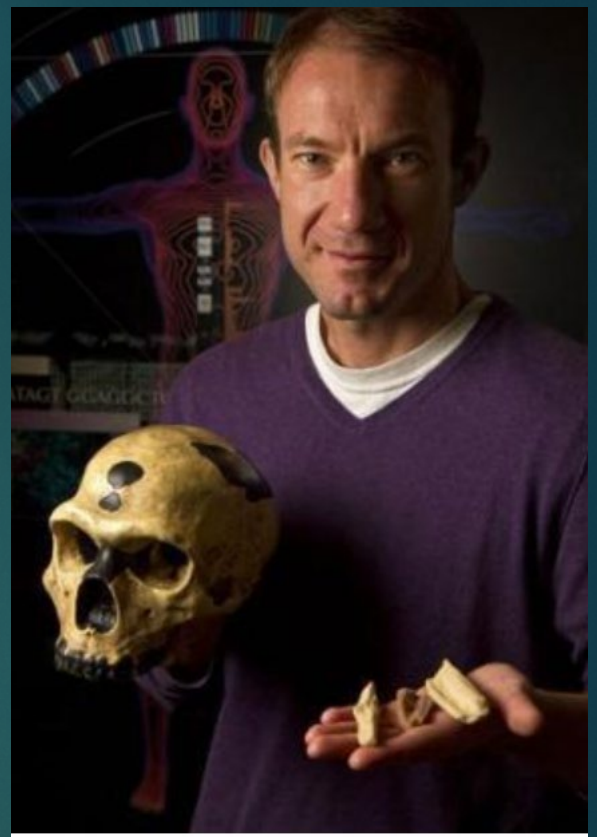
2010: Draft sequence of the Neandertal genome (Green *et al.*, *Science* **328**, 710 (2010))

Three Neandertal Bones from Vindija, Croatia: combined the DNA from all 3 bones to get first draft genome of N = ~55% of N genome

Originally thought to be a faunal bone and tossed in box with animal bones

Richard Edward Green:
1-4% Neandertal DNA in modern humans

- Computational biologist; UC Santa Cruz
- Lab of Svante Pääbo
- 2010: proved gene flow from Neanderthals to modern humans between 50-60 Ka ago
- 2010: Found 1 to 4 % of the genomes of non-Africans is derived from Neanderthals, meaning that the admixture occurred early on, probably in the Middle East;



Richard E. (Ed) Green, a computational biologist in the Baskin School of Engineering at UC Santa

Green *et al.*, *Science* **328**, 710 (2010)

MH & Ns share 99.7% of 3 billion SNPs

- Neanderthal DNA is 99.7 percent identical to present-day human DNA
- N = 98.8 percent identical to chimpanzee DNA.
- 9 million SNPs (single nucleotide polymorphisms) difference between N & MHs
- MHs have 1-2% N nuclear DNA. You have about as much Neandertal DNA as people inherit from a 4th G-grandparent.
- Split of the N and MH lineages, based on mtDNA, is dated to 760 to 550 Ka

2010 Denisovan mtDNA differences

- MH differences in mtDNA from:
  - Neanderthals: an average of 202 nucleotide positions, out of approximately 16,500
  - Denisovans: 385 positions,
  - Chimpanzees: 1,462 positions

Binomial species name battle and “Denisovans”

- Svante Pääbo and his Leipzig gang have refused to give a classic Latin binomial name (i.e. Homo denisova) to the new genetic findings from the finger bone discovered at Denisova Cave in the Altai Mountains

- Russian collaborators use “*Homo altaiensis*”

Not Us and Them

- Pääbo now recommends against imagining separate species of human evolution: not an Us and a Them, but one enormous "metapopulation" composed of shifting clusters of essentially human-ish things that periodically coincided in time and space and, when they happened to bump into one another, occasionally had sex.

- Finlayson: "Each valley could have told a different story. In one, they may have hit each other over the head. In another, they may have made love. In another, they ignored each other."

- Jon Mooallem: "a superlong elevator ride with strangers."

2013: 3rd Oldest DNA: Dawson, in Canada's Yukon Territory

700 Ka frozen sediment; bone near ash layer at 680-700 Ka; DNA of complete horse

2013: 735 K year old horse genome

- Samples from a horse leg bone from 735 KA have yielded the 3rd oldest full genome known to date.
- Cold is good. Frozen is even better, because liquid water isn't present to degrade DNA molecules.
- The six-inch (15-centimeter) horse leg bone originated in the Yukon Territory of western Canada in permafrost in 2003.
- Sequenced 12 billion DNA (mostly bacterial) molecules, of which 40 million were of horse origin

Orlando, L., *et al., Nature,* 2013

Paleogenetic Studies:

- 2012: Full sequence of the Denisovan genome
  - Matthias Meyer, et al., A High-Coverage Genome Sequence from an Archaic Denisovan Individual *Science* (30 August 2012)

- 2013: A mitochondrial genome sequence of a hominin from Sima de los Huesos = Denisovan (Matthias Meyer, et al., , *Nature*, 2013)

- 2013: The complete genome sequence of a Neanderthal from the Altai Mountains, (Kay Prüfer, et al., *Nature, 2013*); Denisova is 4900 miles from Spain – Neandertals had huge range

Sima de los Huesos: Denisovan Mitochondrial DNA

- 2013: hominin femur from Sima de los Huesos (Pit of Bones) in Atapuerca, Spain: mitochondrial DNA closer to that of Denisovans than to Neanderthals or modern humans.
- 2 of many possible explanations: Pääbo: from a prior ancestor of N & D; Stringer: *Homo antecessor* interbred with unknown species who was ancestor to both Denisovan and Sima group

Figure 2 | Femur XIII reassembled from three parts after sampling. The natural fractures are visible in the proximal third of the femur.

K. Prüfer, 2013: What makes us MH: Our genetic recipe

- MH DNA sequence changes that distinguish MHs from our nearest extinct relatives is small. In 3 billion base pairs, only:
  - 31,389 such single nucleotide substitutions
  - 4,113 short insertions and deletions (indels)
  - 105,757 substitutions and 3,900 indels shared by 90% of present-day humans.
  - 87 genes
  - only 96 fixed amino acid substitutions in a total of 87 proteins
  - 5 genes effect neural stem cells in the adult subventricular zone.

K. Prüfer, 2013

- Introgressed Neanderthal DNA sequences suggest a population split from the Altai Neanderthal between 114-77 Ka ago, well after 400 Ka ago when Neanderthal features appear in the fossil record

- Allele sharing between Neanderthals and non-African populations is owing to recent admixture.

2018 data: Only 12,000 bp changes

- Previously, a number of 31,389 sites has been reported as recently fixed derived in present-day humans, while being ancestral in archaics (Pääbo 2014; Prufer et al. 2014).
- We find a smaller number of only 12,027 positions are different in the genome, based on more MH genomes
- Current: 647 protein altering changes in 571 genes; genes that were under positive selection in humans traits related to brain functions are prominently represented
  - Cell division and the brain growth trajectory: brain growth, ventricular region neuron multiplication, size of cerebellum, globular braincase shape
  - Cellular features of neurons: genes with axon-guidance-related functions, related to language
  - Craniofacial phenotype

Martin Kuhlwilm, Cedric Boeckx, 2018

MH mutations:

Genetic recipe for a modern human vs Ns & Ds

- Pääbo: "The dirty little secret of genomics is that we know next to nothing about how a genome translates into the particularities of a living and breathing individual."

96 Human-specific Amino Acids: from 87 genes

A Catalog of "all" Human-specific Amino Acids

DDX53	NOP14	ADAM18*	KIF18A	SSH2
CXorf59	EVC2	RB1CC1	PLAC1L	SYNRG
Orf	HERC5*	LYPLA1	ZNHIT2	CD300LG
FRMD7	DHX29	GPT	PRDM10	TEX2
ZNF185	PTCD2	GLDC	LRTM2	ITGB4*
TKTL1	SV2C	FRRS1L	LAG3	RFNG
IFI44L	VCAN	NEK6	SCAF11	GREB1L
VCAM1	RASA1	TTF1	SLITRK1	LMNB2
SPAG17*	IRAK1BP1	FBXW5	NOVA1	MFSD12
SLC27A3	MCHR2	FAM166A	TTLL5	NCOA6
SPTA1	ZBTB24	ARRDC1	GPR132	TP53TG5*
NFASC	KATNA1	ANKRD30A	CASC5*	C21orf62
KIF26B	LRRD1	FAM149B1	STARD9	RSPH1
SLC8A1	KLF14	FAM178A	SLC12A1	ENTHD1
NOTO	CALD1	Orf	KIAA1199	ADSL
ANKMY1	ERI1	PNLIP	CDH16	
SCAP	CSGALNACT1	UBQLN3	PIEZ01	Σ: 87 genes
OR5K4*	GSR	DCHS1	SPAG5**	

Human Accelerated Regions (HAR1)

- Our DNA blueprints are 98.4% identical to chimps. Only 15 million bps—less than 1 percent—are different in humans
- Research: Find pieces of DNA that have changed the most since humans and chimps split from a common ancestor.
- A stretch of 118 bases that together became known as human accelerated region 1 (HAR1).
- Involved in cortical development; When things go wrong in these neurons, the result may be a severe, often deadly, congenital disorder known as lissencephaly; a markedly reduced cortical surface area

HARs: mainly noncoding areas

- ~30 gene families show human-specific gene duplications
- Besides these 30 genes, there are a 1000 noncoding area sequences that appear human specific
- HARs - human accelerated regions; basepair mutations that are fixed in all mammals but changed in humans
- HARs largely serve to regulate the activity of other genes, including those that guide brain development.

Hominins like sex: 3 to 5 cases of interbreeding among four distinct hominin populations

N into MH = 1.5-2.1%

N to D = 0.5%

D to MH = 3-6%

Unknown hominin to D = 0.5-8.0%

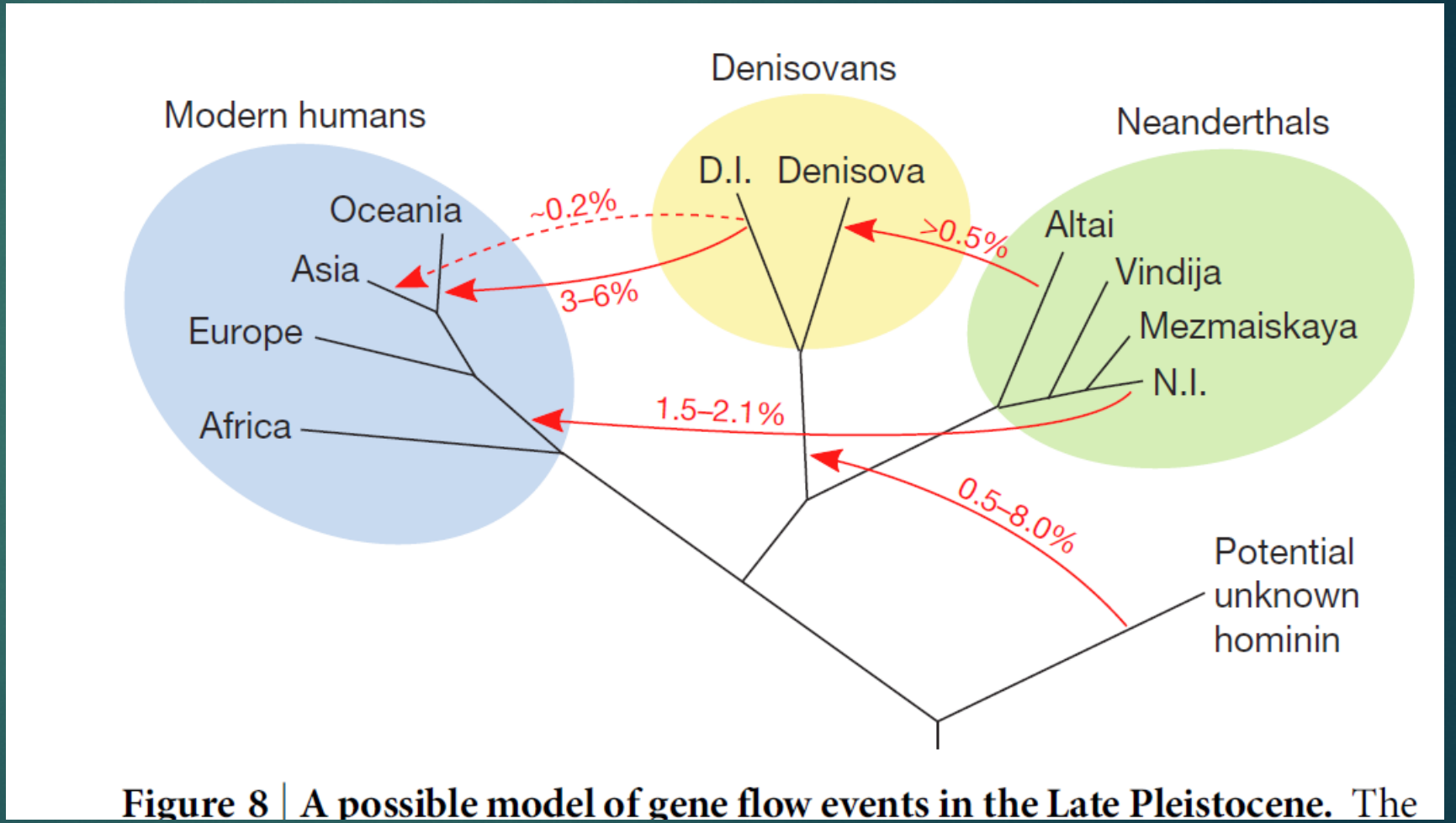


Figure 8 | A possible model of gene flow events in the Late Pleistocene. The

K. Prüfer, 2013:

Percentage of admixture: differing estimates

- The proportion of Neanderthal-derived ancestry was estimated by Green, et al., 2010 to be 1–4% of the Eurasian genome.

- Prüfer et al. (2013) estimated the proportion of N DNA in MHs:
  - 1.5–2.1% for non-Africans [now the accepted %]
  - revised in 2017 to a higher 1.8–2.6% for non-Africans outside Oceania.

- Lohse and Frantz (2014) infer a higher rate of 3.4–7.3% in Eurasia.

How much N DNA in MHs

- N & D mix in the Native Americas = 0.2%
- N mix in Modern Africans = 0.3%
- Prüfer et al. (2017) noted that
  - East Asians carry more Neandertal DNA (2.3–2.6%) than Western Eurasians (1.8–2.4%).
    - Chen et al. (2020): East Asians have 8% (not prior 20%) more Neanderthal ancestry than Europeans

Archaic human ancestry

- All modern non-African genomes are estimated to carry approximately 2%-8% archaic human sequence:
  - approximately 2% ancestry from Neanderthals
  - additional 2%-6% ancestry from Denisovans in Melanesian populations.
  - Present-day levels of archaic ancestry need not reflect initial admixture levels, which were higher

N DNA revelations:

- It is important to remember that Neanderthals:
  - were a heavily inbred population
    - Altai Neandertal's parents were related as half-siblings
  - with very low effective population sizes (possibly by an order of magnitude lower than the early AMH's)
  - which may have lowered their genetic fitness (capacity to survive and reproduce) by as much as 40%

Why the replacement?

- Why was N DNA replaced in MHs?

- Because
  - Ns had such small populations and accumulation of deleterious variants,
  - MH variants fixated because they were more functional

Neandertal effective population sizes

- Better evidence: K Prüfer, et al. 2014: estimate that the effective population size (Ne) of Neandertals was Ne = 1,000–5,000; a shorter common branch for the Neandertal–Denisovan ancestor (300 generations).

- Rogers et al., 2017: Ne of Neandertals was n = ~15,000; a longer common branch for the Neandertal–Denisovan ancestor (5,000–10,000 generations).

- Fabrizio Mafessoni & Kay Prüfer, 2017: Rogers et al. model predicts a large Ne for Neandertals that is at odds with the low heterozygosity in the Altai Neandertal. A relatively small effective population size for Neandertals remains better supported = 1,000-5000 Ne

CAS visitor question: If MH and N are 99.7% genetically identical, how can MHs have 2% Neandertal DNA per Ancestry.com

- The 99.7% refers to all 3 Billion base pairs of your genome, thereby leading to a 0.3% bp difference between MHs and Ns.

- You have ~2% N DNA.

- The 2% refers to 2% of that 0.3% bp difference -- roughly 0.0006% of 3 billion, or 180,000 N base pairs.

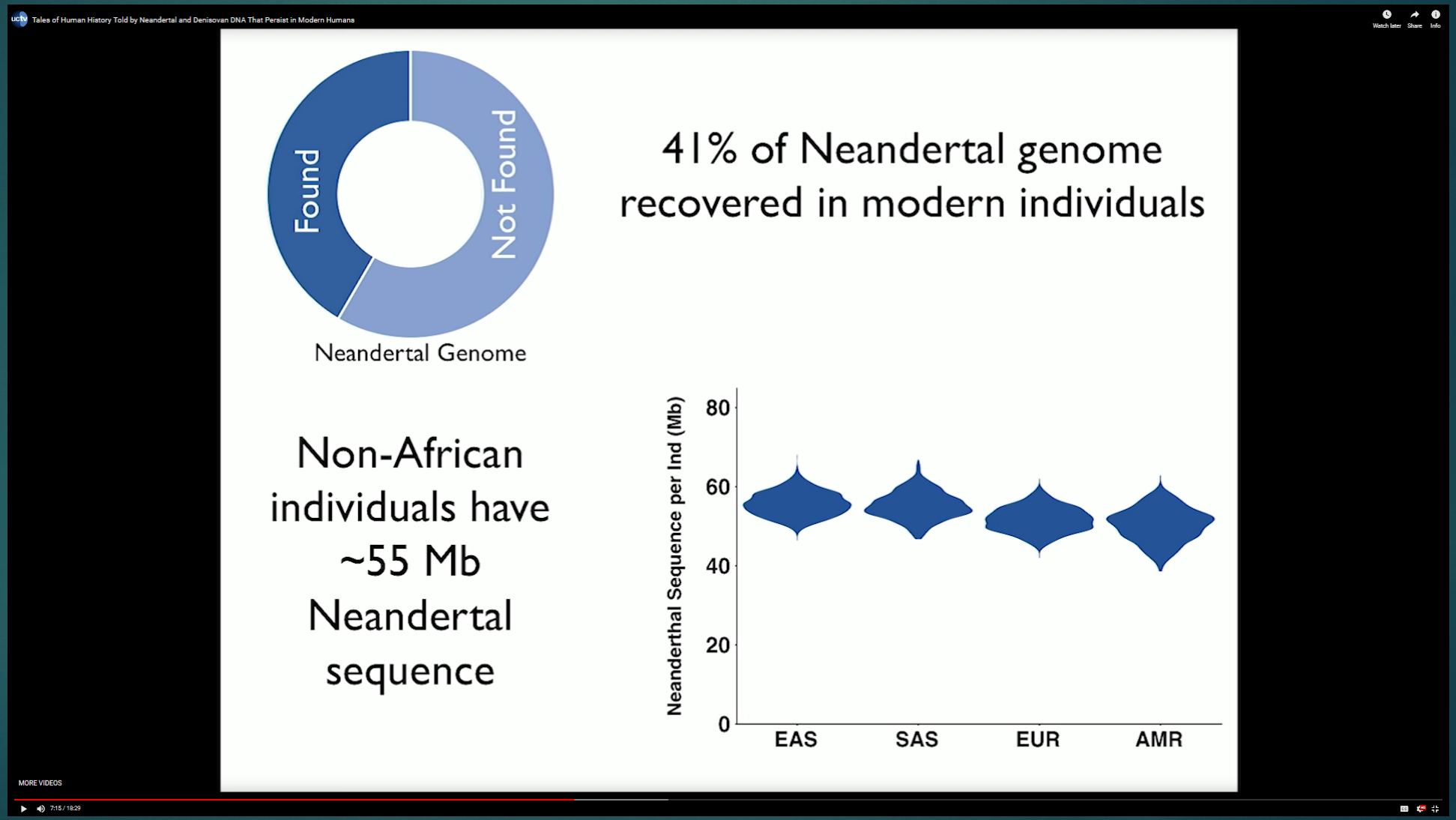
N DNA at Ancestry.com

- 135,171 N SNPs (Single Nucleotide Polymorphisms = a ACTG switch) are “Neanderthal variants“; believed to have originated in Neanderthals and later entered the modern human population via interbreeding.

- Of these, 3,731 SNPs are assayed via Ancestry Illumina’s v5 genotyping, and 1,436 SNPs are further assayed.

- For each SNP, an individual carries 0, 1 or 2 Neanderthal variant copies.

41% of total N genome present today in MHs

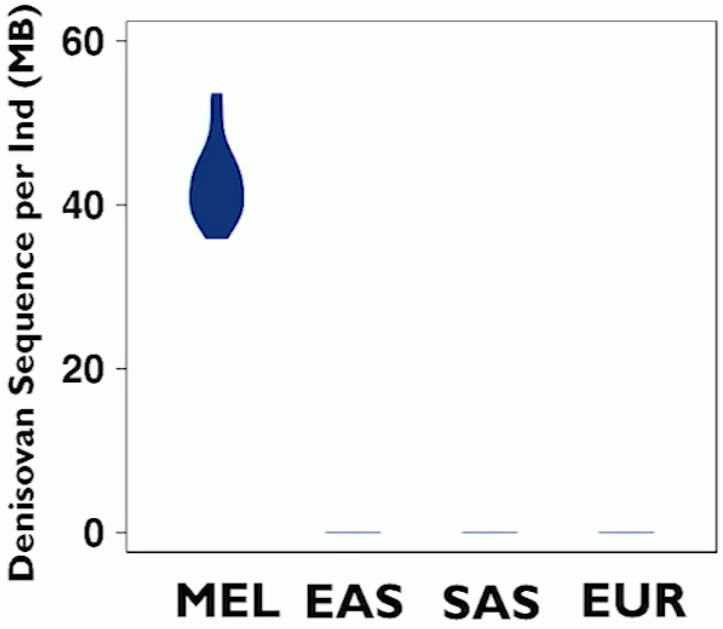


10% of total Denisovan genome is present in MHs

Denisovan Genome

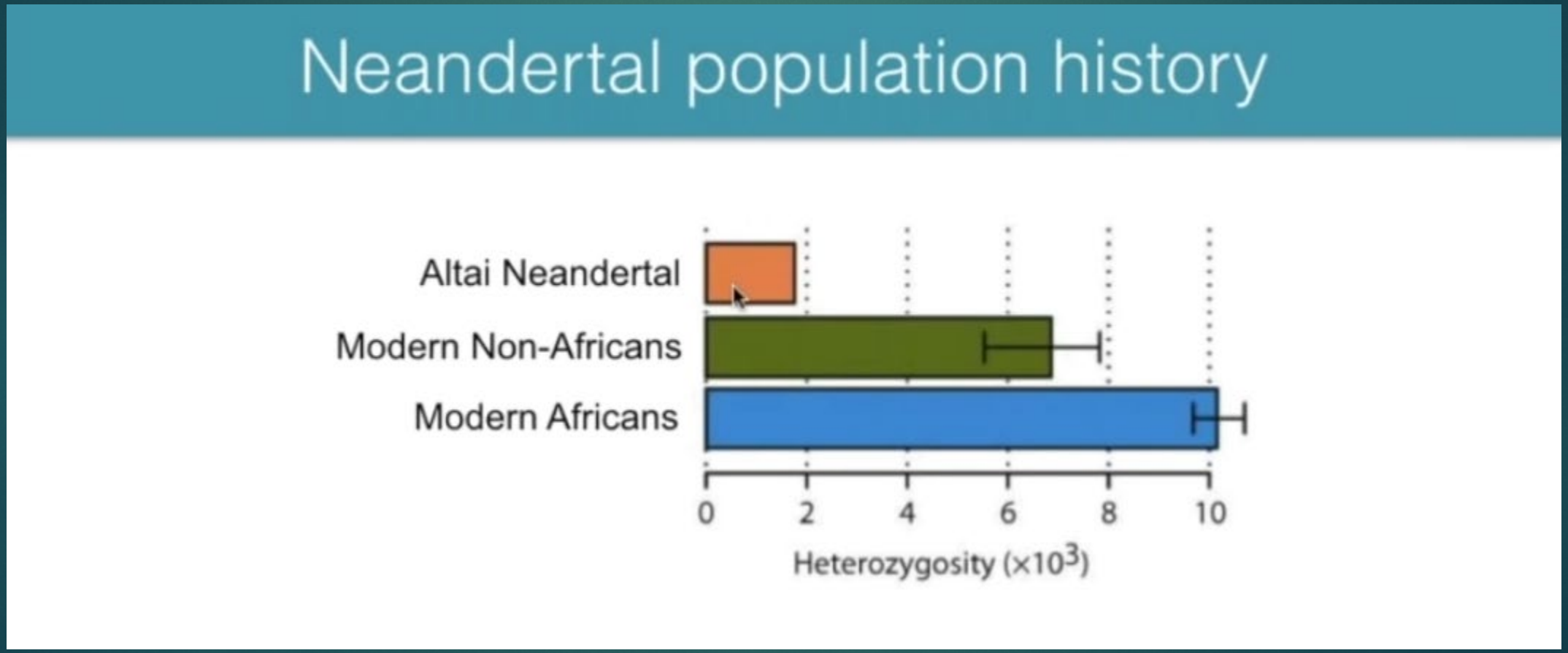
10% of Denisovan genome recovered in modern individuals

Melanesians have ~40 Mb Denisovan sequence



Based only on n=35 who are Melanesians; more to be found

Very Low N Genetic diversity: long stretches of homozygosity; lots of interbreeding; implies closely related parents



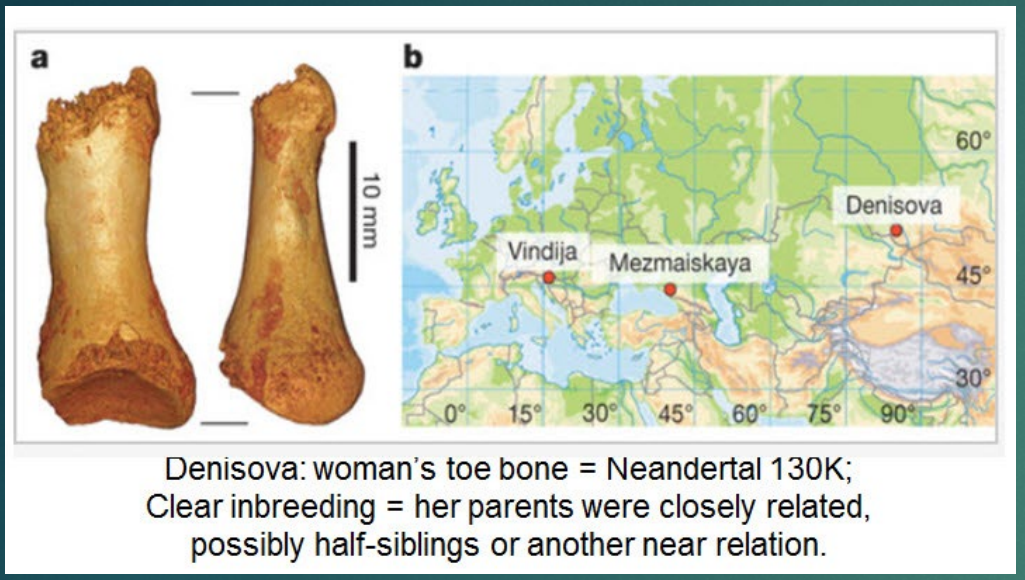
Neandertals from Genome

- Low genetic variability:
  - heterozygosity in Neanderthals as well as Denisovans appears to have been
    - lower than in present-day humans and is
    - among the lowest measured for any organism.

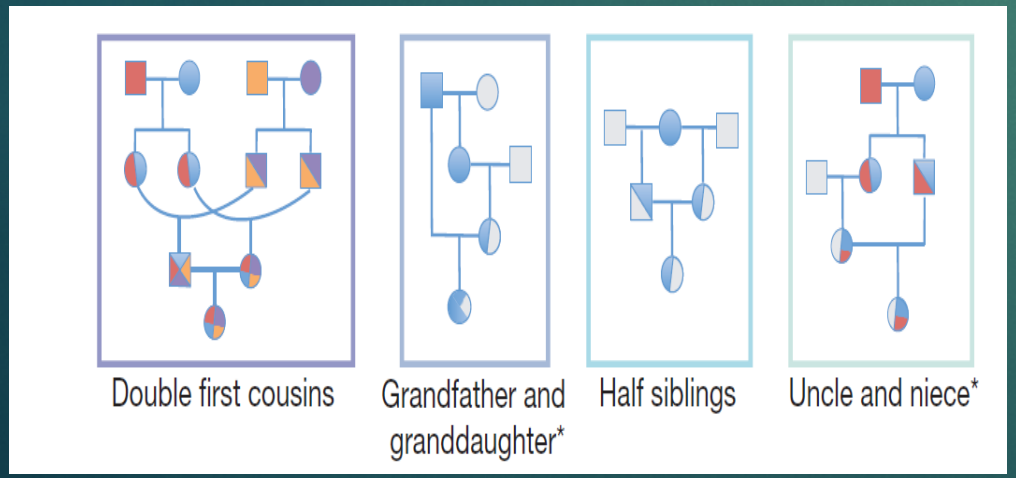
- All N genomes analysed show evidence of a reduction in population size that occurred sometime before 1.0 million years ago.

- Subsequently, the population ancestral to present-day humans increased in size, whereas the Altai and Denisovan ancestral populations decreased further in size.

2014: One Reason for Neandertal Demise: Low population number with interbreeding



Denisova: woman's toe bone = Neandertal 130K;
Clear inbreeding = her parents were closely related, possibly half-siblings or another near relation.



Denisova Neandertal woman toe bones:

Chromosome 21: Mom & Dad genetically related (19 Mb base pairs with no difference)

Highly interbreed:
Half siblings
Grandfather-granddaughter
Aunt-nephew
Double first cousins

Pruefer et al., , Nature, 2014

Neandertal DNA in Different Modern Humans Not Same

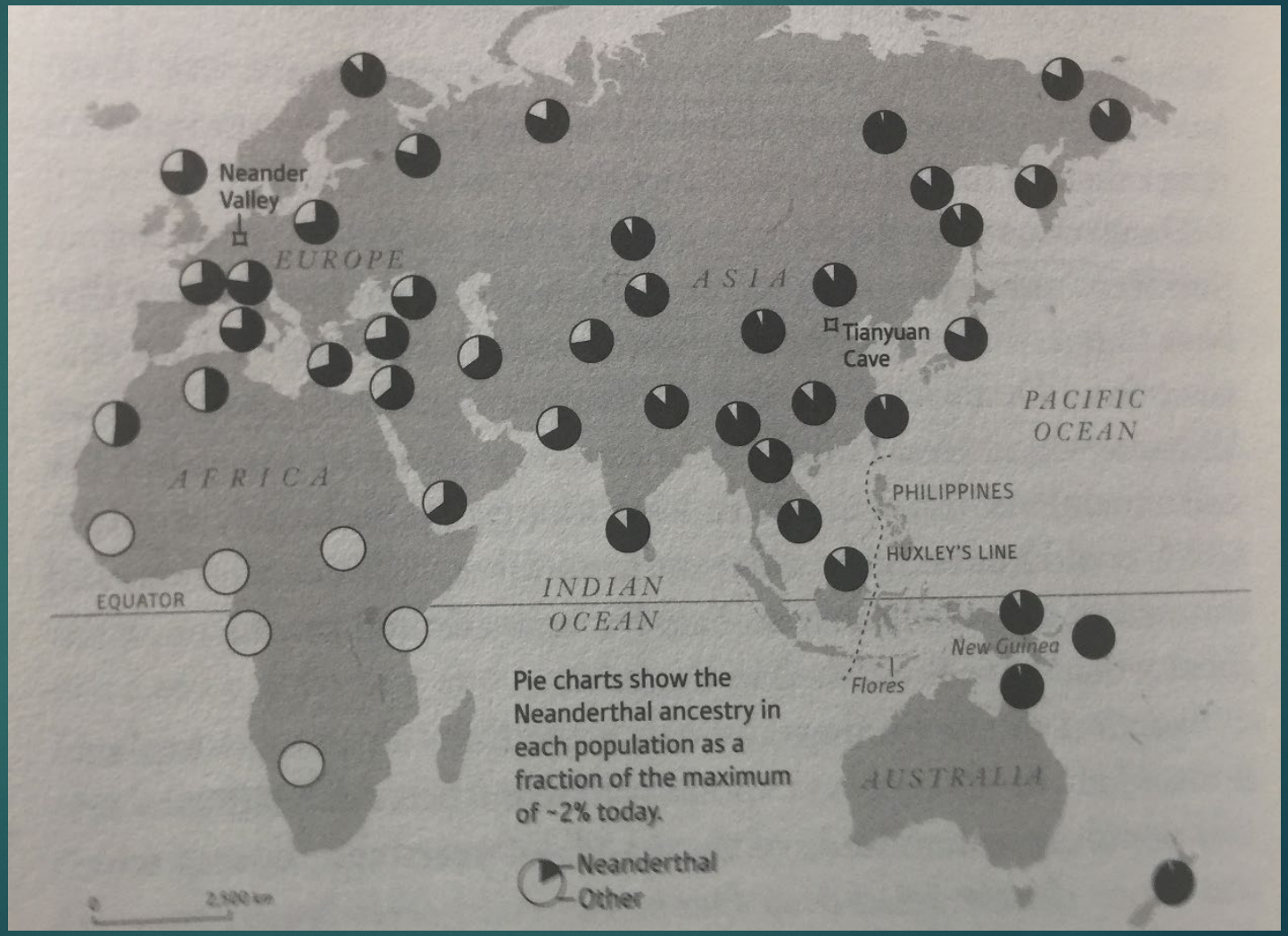
- While only 1-2% of the total genome of moderns is Neandertal, this represents 40% (20-70% range) of total Neandertal genome.

  - Living Europeans have inherited around 1.2% and
  - East Asians about 1.4% of their DNA
  - from our Neandertal cousins.

- Akey - 2,504 genomes:
  - Europeans on average had 51 M N bps
  - East Asians had 55 M N bps;
  - Africans = 17 M N bps

S. Sankararaman:

N ancestry (fraction of maximum 2%) in various Eurasian populations



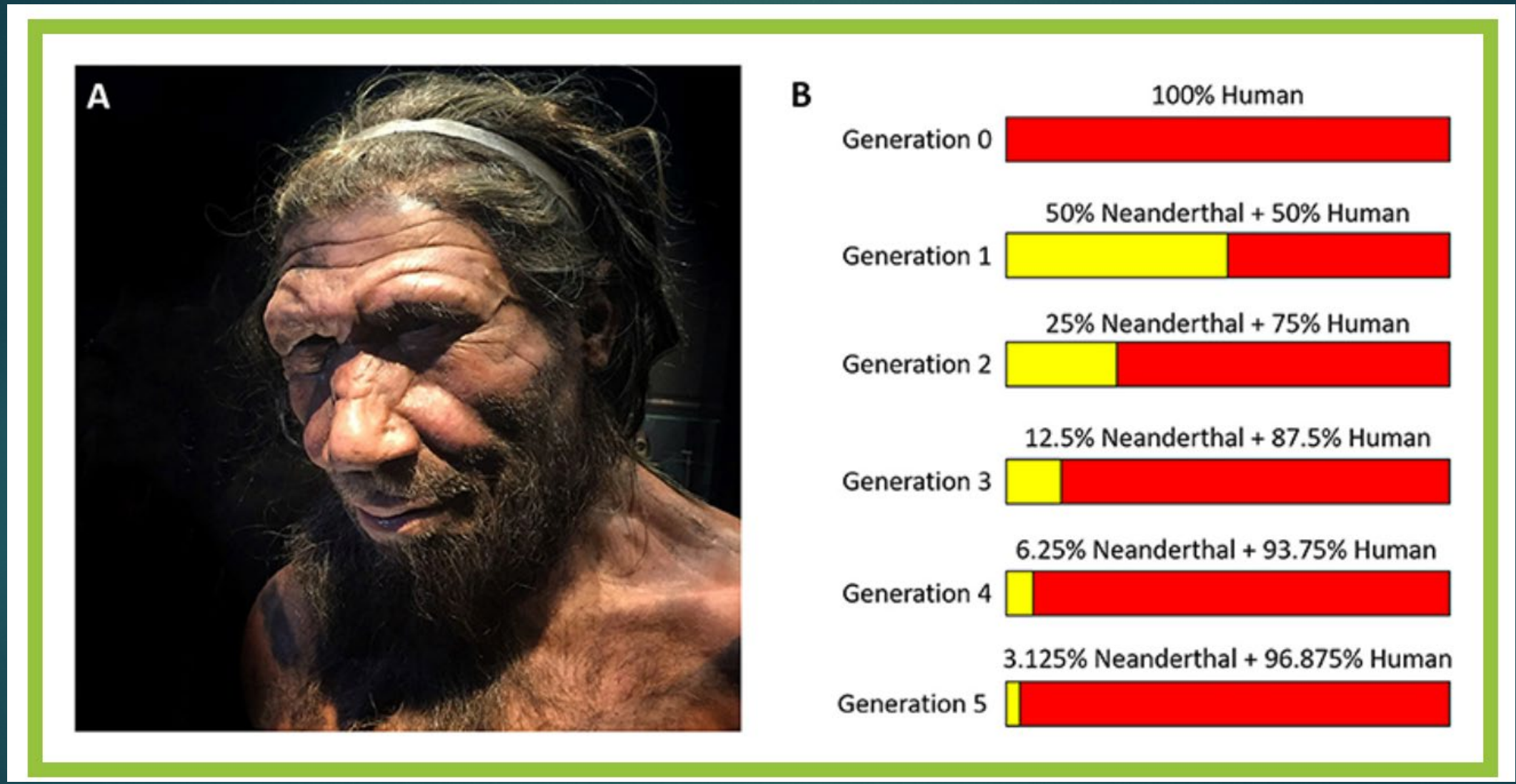
Wolf & Akey, 2018: early MHs had more N DNA

- Early MHs had more N DNA: discovery of an ancient East Asian individual, dated to 40 Ka, who was an ancestor of modern Asians and who carried 4 to 5% Neanderthal ancestry.

- Data from a 42 Ka AMH from Pęstera cu Oase, Romania, reveal this individual shared 6% to 9% of his genome with Neanderthals, more than 3× any contemporary modern humans.

- Pęstera cu Oase individual had a very recent Neanderthal ancestor (within 4-6 generations) and likely did not contribute any ancestry to modern populations.

N and MH Hybridization: why 2% N DNA today?



Down to 3% within 120 years; why is there not less N DNA in us

Answer: Adaptive introgression

- A gene variant will alter a phenotype if gene frequency increases because it proves beneficial for individual. If more babies survive with that trait.

- Some introgressed Neandertal DNA turned out to be helpful for humans and did not disappear from the human DNA.

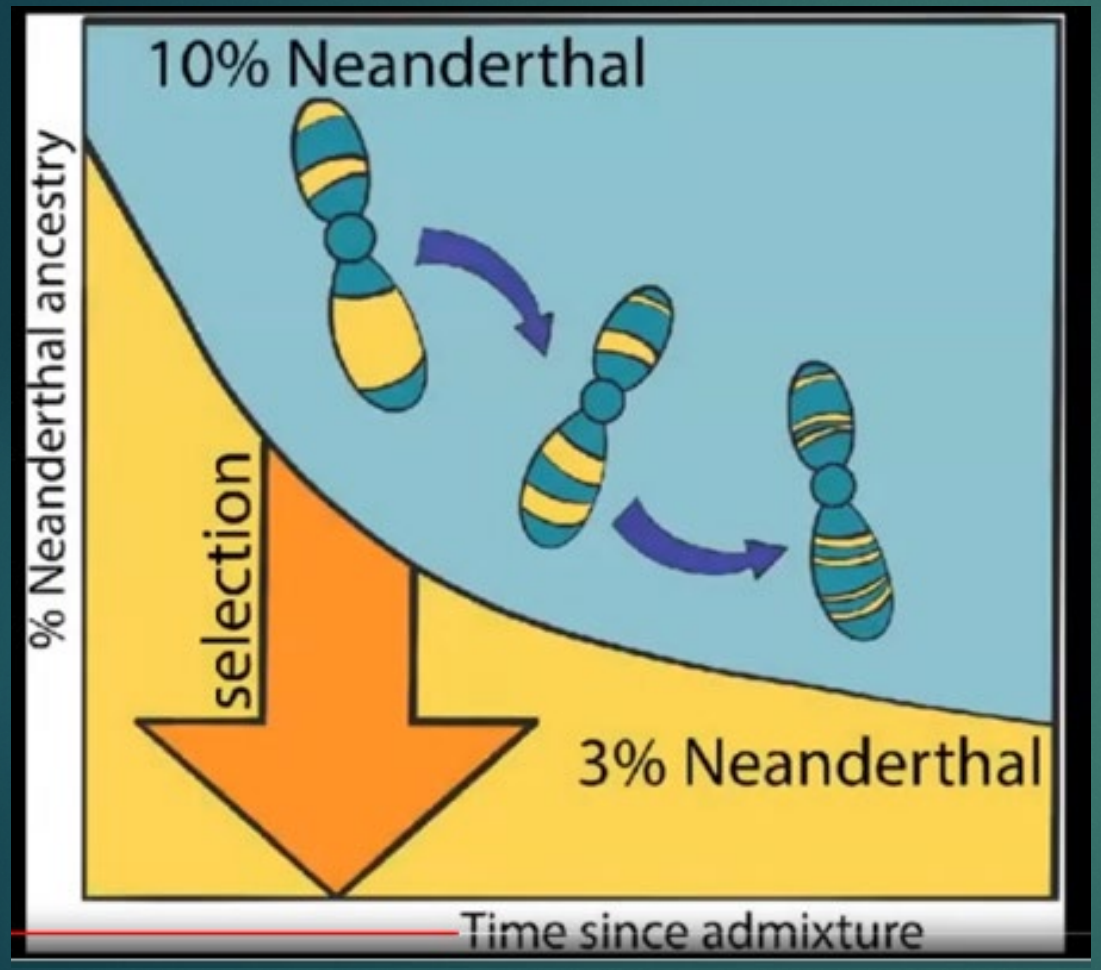
Pääbo: N-MH Assimilation

- The first modern humans in Europe may have mixed extensively with resident Neandertals
- MHs, circa 45-40 Ka, actually mixed quite happily, extensively and frequently with Neanderthals.
- Neandertals and Denisovans (at least partially) were assimilated into larger modern human populations?
- Perhaps different scenarios in different locales
- If MHs had 50x larger population, then we get 2% N DNA in MHs today

All the first Upper Palaeolithic humans in Europe had recent Neanderthal ancestry: Bacho Kiro Cave, Bulgaria, 46-43 Ka, 3.8% N DNA; at 35 Ka, in same cave, only 1.9% N

Hajdinjak *et al.*. Nature, 2021

N DNA selected away from original 10% in MHs to 2% now

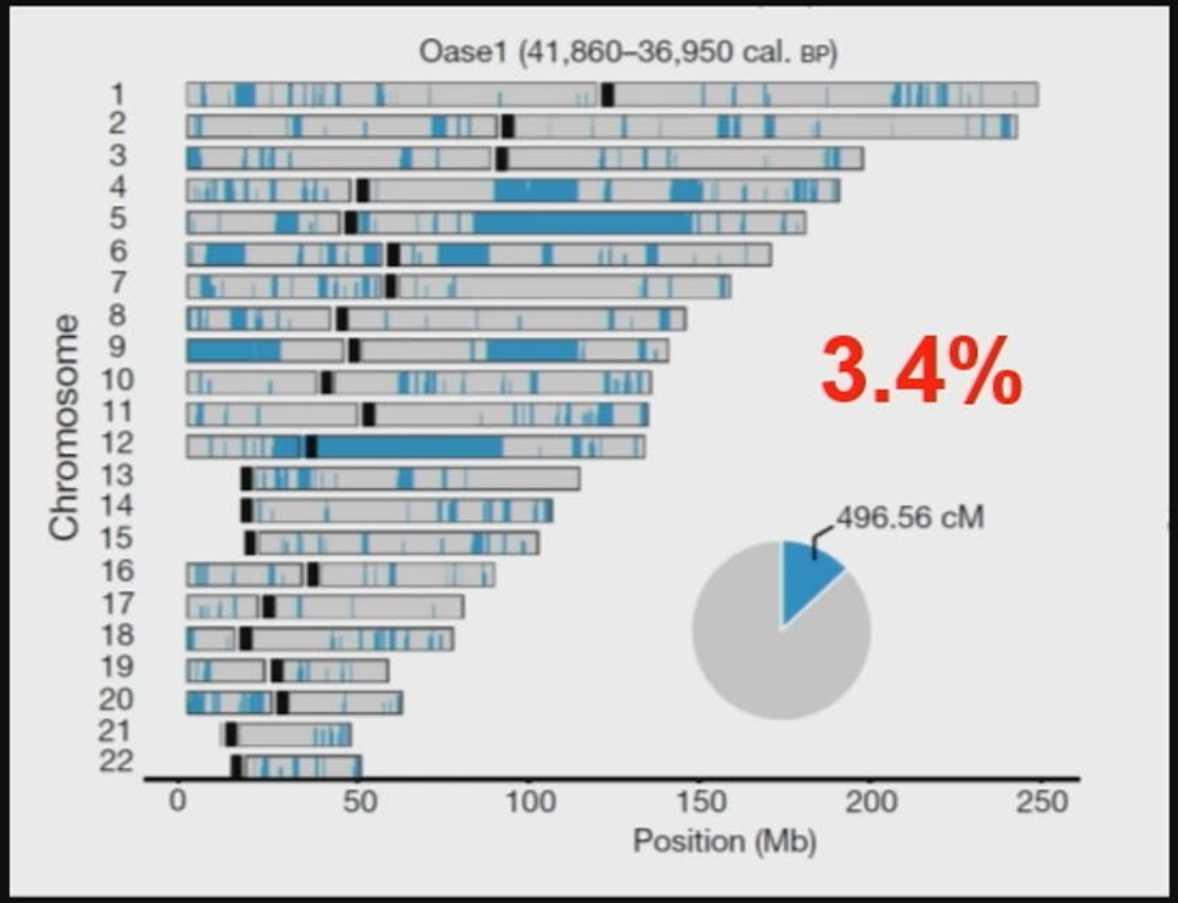


The curve shows the expected decline n the proportion of Neandertal DNA in modern humans due to natural selection.

Not only is the Neanderthal DNA proportion decreasing through time, it is also distributed in smaller and smaller segments due to the effect of recombination

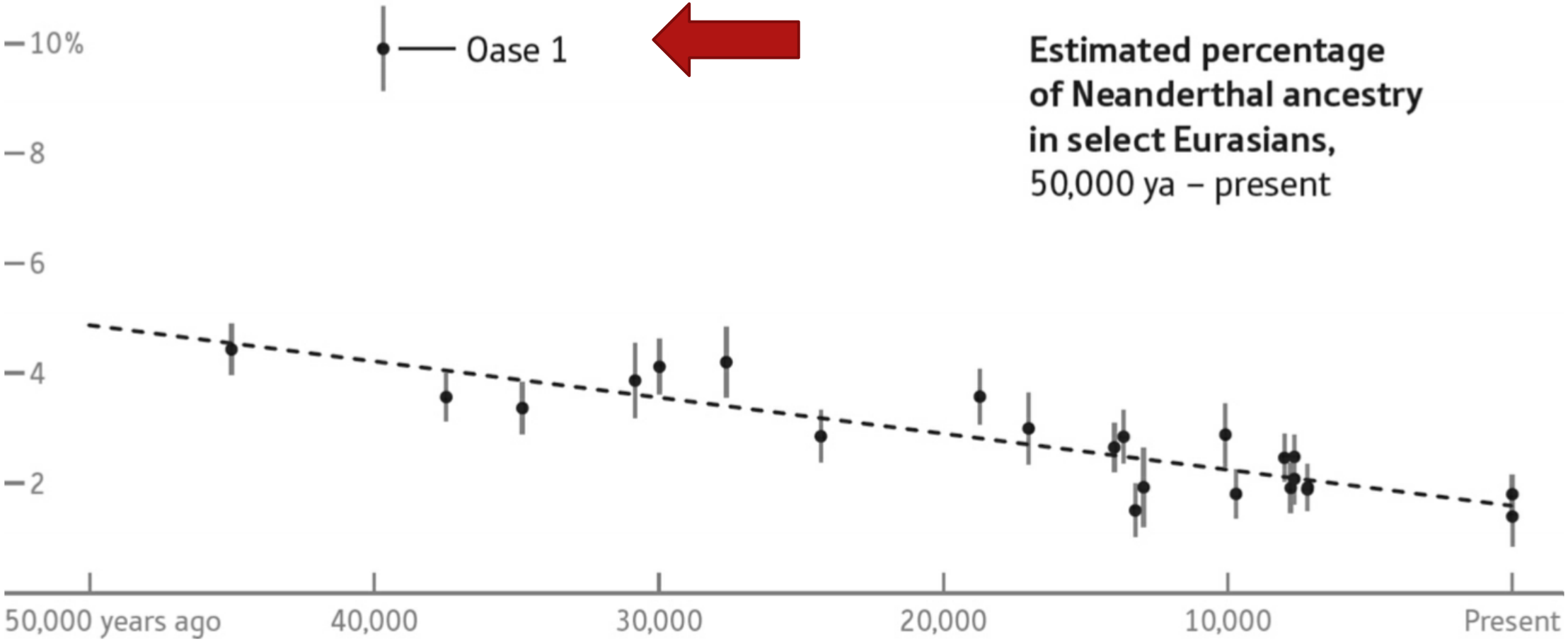
Oase 1, Romania

~40,000 yrs



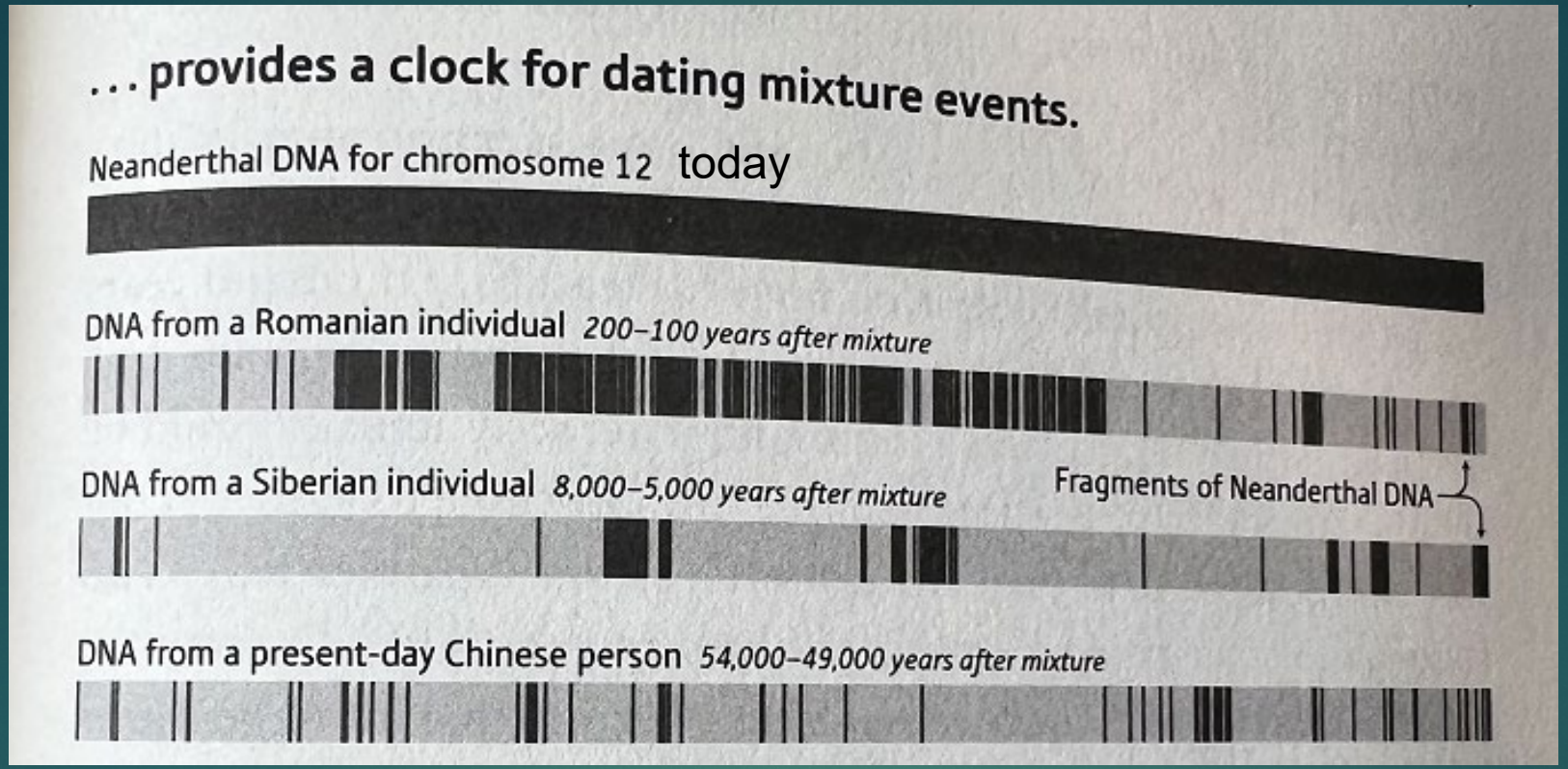
6 generations back, a full N ancestor

Neanderthal ancestry has been removed over time by natural selection.



Oldest currently known Modern Humans in Eurasia: all had N DNA

- *Oase 1*: 37-42 Ka, Romania, Modern Human-Neandertal hybrid; with 6-10% Neandertal DNA, GGGgrandparent = full Neandertal
- Bacho Kiro cave, Bulgaria, 43-46 Ka; Earliest Modern Human in Europe; 3.5% N DNA; but at 35 Ka, same cave, only 1.9% N
- Zlatý kůň in Czechia; Modern human skull, 45 Ka+; long N fragments
- Ust'-Ishim individual from Siberia, ~45 Ka – had a N ancestor 7 Ka before; 2% but much longer fragments
- Surprisingly, however, none of those pre-40 ka individuals left substantial genetic traces in present-day Eurasian populations



- N DNA dilution in MH:
  - Amount of Neandertal DNA lessens with amount of time from original admixture.
  - There is no N DNA in more than 50% of MH genome.
  - Evidence of systematic removal of N DNA by natural selection, esp. protein coding gene regions, & X & Y chromosomes

Sex between N and MHs

- Either Quest for Fire or Dancing with Wolves; the difference is the music
- Kidnapping or Romance?
- No way to tell currently
- Less mixture in X chromosome than in other chromosomes

N DNA in MHs

- Neanderthal-derived DNA has been found in the genomes of most or possibly all contemporary populations, varying noticeably by region.

- N DNA for 1.8–2.6% of modern genomes for people outside Sub-Saharan Africa, and up to 0.3% for those in Africa. Those are averages for whole genome.

- Specific regions of the genome may have degrees of Neanderthal ancestry as high as 64% in Europeans

N, D DNA in MHs

- N DNA is highest in East Asians, intermediate in Europeans, and lower in Southeast Asians.
- Denisovan-derived ancestry is largely absent from modern populations in Africa and Western Eurasia.
- The highest rates of Denisovan admixture have been found in Oceanian and some Southeast Asian populations.
- It is present in 4–6% of the genome of modern Melanesians; the highest amounts found in the Negrito populations of the Philippines. The date of Denisovan admixture was ~ 44 to 54 Ka

N, D DNA in Asians

- In addition, low traces of Denisovan-derived ancestry have been found in mainland Asia, with an elevated Denisovan ancestry in South Asian populations compared to other mainland populations.

- Mainland Asian and Native American populations may have a 0.2% Denisovan contribution

- In Africa, ghost lineage: archaic alleles consistent with several independent archaic admixture events in the subcontinent have been found. It is currently unknown who these archaic African hominins were.

Effects of Negative Selection

- No evidence of Neanderthal mitochondrial DNA has been found in modern humans = from Neanderthal male and modern human female pairings

- There is a presence of large genomic regions in MHs with strongly reduced Neanderthal DNA due to negative selection, partly caused by hybrid male infertility.

- These large regions of low Neanderthal DNA were most-pronounced on the X chromosome and testes

Evidence of Functional Archaic Admixture

- There are three lines of evidence for this selection against archaic ancestry:
  - archaic haplotypes have decreased in frequency over time,
  - archaic haplotypes are depleted in more conserved parts of the genome (those that are survival related)
  - archaic variants are less likely to have functional consequences.

DNA discoveries

- *Kennewick Man:* 2013-2015, an 8,000-year-old skeleton found in Washington state in 1996, was genetically closest to local Native Americans. The revelation ended a 20-year legal battle and allowed tribes to rebury the bones.

- *Anzick-1:* 2014, an infant buried with Clovis tools. confirmed that Native Americans mostly descend from Siberians

- 23-kyr-old Mal'ta individual sequenced: more closely related to Europeans and Native Americans than to local Siberians

Genomic Studies:

- 2015: 37-42-kyr-old European Oase 1 individual with recent Neanderthal introgression sequenced
- Ancient and modern Native Americans, Paleo-Eskimo people and the Inuit
- 2016: Oldest *Homo* nuclear DNA (*Spain 430,000 years ago*) - Sima de los Huesos humans are Neandertals
- 2019: 3 types of Denisovans; 1 needs new species name

Only 1.5 percent to 7 percent of the human genome contains uniquely human DNA

- Study examined every spot of DNA in the genomes of 279 people: For each basepair, determined whether it was MH, N, D, etc.

- Humans-only DNA tends to contain genes involved in brain development and function, hinting that brain evolution was important in making humans human. But don't yet know exactly what the genes do

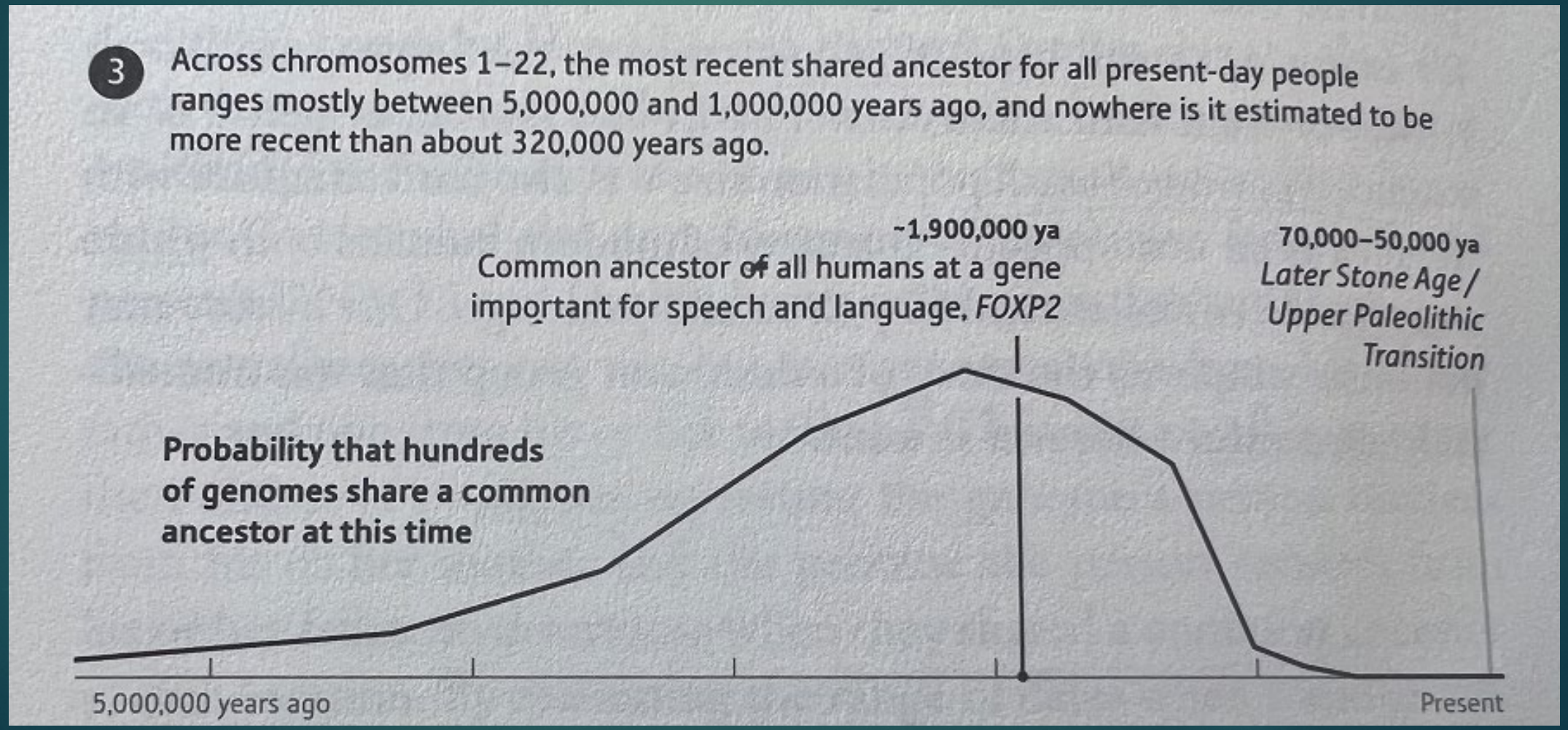
Human Migrations: Profoundly interrelated Species

- Out of Africa: Latest 2016 Nature: all non-Africans today trace their ancestry to a single population emerging from Africa between 50,000 and 80,000 years ago. All MH DNA is African by origin

- 2 humans on separate continents are closer genetically, than 2 chimps on opposite side of an African river in same jungle (100 chimps are more diverse than all 7 B MHs)

- Genetic diversity in non African MHs is incredibly low; of 14 "ancestral clusters" for all of humanity, 9 of those clusters are in Africa (due to longest time to accrue mutations)

How We Can Tell How Long It Has Been Since Our Genes Shared Common Ancestor: Most recent shared ancestor is ~320 ya -- MRCA for all present-day people: between 1-5 Ma to 320 Ka; for FOXP2, 1.9 MA



Most recent shared ancestor of MHs is ~320 ya

- MH genome: there is no gene location where all people living today share a common ancestor earlier than ~320 K ya; in effect, the approximate origin date of MHs

- This is far older time than required by Richard Klein's theory of genetic switch that made us MH ~50 Ka; disproves his theory; if he was right, would find genetic variants that were shared within last 100 Ka; but there are none

Generations	Relationship	# of ancestors	Percent of DNA
7	GGGGG-Grandparents	128	0.78
6	GGGG-Grandparents	64	1.56
5	GGG-Grandparents	32	3.12
4	GG-Grandparents	16	6.25
3	Great-Grandparents	8	12.5
2	Grandparents	4	25
1	Parents	2	50
	You	1	100

The generational relationship chart above represents the average amount of DNA that you will inherit from each of those ancestors. You may inherit more or less.

How much DNA is inherited

- Everyone inherits 50% of their DNA from their parents, but not everyone inherits half of each of their ancestors' DNA from a parent.
- Sybs can inherit different amounts.
- Sometimes, the child will inherit all of a segment of DNA from an ancestor, and in other cases, the child will inherit none.
- You have a 1 in 8.4 million chance of being unrelated to one of your grandparents.
- 2% N = equivalent of gggg-grandparent amount

Origin of MHs Debate as of 2011: Stringer vs Wolpoff

- From 1984, for 27 years, Chris Stringer and Milford Wolpoff fought about where and how MHs originated.

- Stringer, a paleoanthropologist at the Natural History Museum in London, held that modern humans came out of Africa,
  - spread around the world,
  - and replaced, rather than mated with, the archaic humans they met.

- Wolpoff, of the University of Michigan, Ann Arbor, argued that a single, worldwide species of human, including archaic forms outside of Africa,
  - met, mingled and had offspring locally,
  - and so produced *Homo sapiens*.
  - Their battle was long and bitter.

Ann Gibbons 2011

Origins

- Then in 2010, nuclear genomes of Ns and Ds came out.
- Allowed test of above models.
- Genomes appeared to refute the complete replacement concept of the Out of Africa model
- Winner: Out of Africa, but with low levels of admixture

Origin of MHs Debate

- Genomic data did not prove the classic multiregionalism model correct either.
- They suggest only a small amount of interbreeding, presumably at the margins where invading moderns met archaic groups that were the worldwide descendants of *H. erectus*
- Svante Pääbo: best model = replacement with hybridization, or 'leaky replacement.
- New picture most resembles so-called assimilation models, which got relatively little attention over the years, a la Fred Smith

Origins - 1984

- In 1984, Mitochondrial Eve: mother of all in Africa, circa 200 Ka
- Studies of living people—from Y DNA in nucleus & mtDNA—consistently found that Africans were the most diverse genetically.
- This suggests that modern humans arose in Africa, where they had more time to accumulate mutations

Origin of MHs Debate

- Meanwhile, ancient DNA technology also took off.
- Pääbo's group sequenced first a few bits of Neandertal mitochondrial DNA in 1997, then the entire mitochondrial genomes of several Neandertals—and found them to be distinct from those of living people.
- So this ancient DNA, too, argued against the idea of mixing between Neandertals and moderns.
- Over the years the replacement model became the leading theory, with only a stubborn few, including Wolpoff, holding to multiregionalism.

Origin of MHs Debate

- A few paleoanthropologists proposed middle-of-the-road models, i.e. Fred Smith:
  - most of our ancestors arose in Africa but interbred with local populations as they spread out around the globe, with archaic people contributing to about 10% of living people's genomes.

- At the University of Hamburg in Germany, Gunter Brauer similarly proposed replacement with hybridization, but with a trivial amount of interbreeding.

- But neither model got much traction. Over time, the two more extreme models moved toward the middle, with most multiregionalists recognizing that the chief ancestors of modern humans arose in Africa

Origin of MHs Debate

- Then in May 2010 came the Neandertals' complete nuclear genome,
- Pääbo's team found that a small amount—1% to 2%—of the nuclear DNA of Europeans and Asians, but not of Africans, can be traced to Neandertals.
- The most likely model to explain this, Pääbo says, was that early modern humans arose in Africa but interbred with Neandertals in the Middle East or Arabia before spreading into Asia and Europe, about 50,000 to 80,000 years ago

Origins: Denisovans

- In December 2010, the team published in *Nature* the complete nuclear genome of a girl's pinky finger from Denisova Cave, Siberia.
- It was neither a Neandertal's nor a modern human's DNA: a new group, Denisovans, more closely related to Ns.
- Denisovans/Neanderthals split from modern humans about 760 to 550 Ka

Origin of MHs Debate

- Modern humans interbred with Neandertals as they left Africa in the past 100,000 years.
- Neandertals left their mark in the genomes of living Asians and Europeans
- Later, a subset of this group of moderns—who carried some Neandertal DNA—headed east toward Melanesia and interbred with the Denisovans in Asia on the way.
- As a result, Melanesians inherited DNA from both Neandertals and Denisovans, with as much as 8% of their DNA coming from archaic people,

Origin of MHs Debate

- This means *H. sapiens* mixed it up with at least two different archaic peoples, in at least two distinct times and places.
- To some, that's starting to sound a lot like a newer version multiregionalism.
- “It's hard to explain how good I feel about this,” said Wolpoff, who says that seeing complete replacement falsified twice in 1 year was beyond his wildest expectations.”

Origin of MHs Debate: Assimilation

- Yet the interbreeding with archaic humans seemed limited—from 1% to 6% of some living people's genomes. Stringer and many others did not consider it full-scale multiregional continuity.
- Low levels of interbreeding suggest that either archaic people mated with moderns only rarely—or their hybrid offspring had low fitness and so produced few viable offspring.
- David Reich notes that at least 90% of our genomes are inherited from African ancestors who replaced the archaic people on other continents but hybridized with them around the margins.
- And that scenario most closely backs the assimilation models proposed by Smith and Brauer.

Oldest DNA genome: 1.6 Ma mammoth

- 2021: Permafrost-preserved mammoth teeth, 1.6 million years old, identify a new kind of mammoth in E. Siberia.

- Genomic DNA extracted from a trio of tooth specimens excavated in the 1970s: The samples sequenced,
  - one from an early woolly mammoth (*Mammuthus primigenius*) and
  - two assigned to a precursor known as steppe mammoths (*Mammuthus trogontherii*), had been excavated by the Russian palaeontologist Andrei Sher.

Tom van der Valk, et al., 2021

2021: Neandertal blood types

- Study of high-quality sequences of three Neanderthals and one Denisovan individuals for 7 blood group systems that are used today in transfusion (ABO including H/Se, Rh (Rhesus), Kell, Duffy, Kidd, MNS, Diego).
- These hominins already possessed the full range of blood variability found in modern humans.
- In addition, it confirms that they had:
  - an African origin,
  - low genetic variability,
  - weak fertility and
  - susceptibility to viral infections that lead to a high infant mortality rate.
  - In a Neanderthal cross with *sapiens*, there would be an 18% chance that the child would develop a hemolytic disease and die.

Condemi S, et al., 2021

Arctic Desert: Kap Kobenhavn Formation in northern Greenland today

This region today is barren and home to moss, lichen, and muskox.

Today preserves sediments from both land and a shallow ocean-side estuary.

2023: Greenland at 2 Ma: New study “A tour de force. Simply astounding”

An illustration of the Kap Kobenhavn Formation in northern Greenland two million years ago, when it was covered with poplar and birch forests and populated with mastodons.

Greenland at 2 Ma: a treasury of species

- Extracted DNA from more than 135 different species:
  - 102 different plant genera
  - 9 different animal taxa: mastodons, caribou, Arctic hares, lemmings, rodents, geese, fleas and ants. Also snippets of horseshoe crab and coral DNA, which generally live today in warmer waters.
  - Now working on 4 Ma sediment

Mineral surfaces adsorbs and preserve aDNA

- Extracted DNA from 41 organic-rich sediment samples at five different sites within the Kap København Formation, Greenland. Screened nearly 3 billion of these “reads” against libraries of living species.

- The marine depositional environment favored adsorption of DNA on the mineral surfaces (clay minerals, the mineral smectite, and quartz).

- Chemical bond with minerals reduced the rate of spontaneous chemical degradation of aDNA

- This presentation contains some copyrighted material from journals the use of which has not always been authorized by the copyright owner. Such material is made available in an effort to advance understanding of the topics discussed in this presentation. This constitutes 'fair use' of any such copyrighted material as provided for in section 107 of the US Copyright Law. In accordance with Title 17 U.S.C. Section 107, the material on this site is distributed without profit, and is used for nonprofit educational purposes. If you wish to use copyrighted material from this site for purposes of your own that go beyond 'fair use', you must obtain permission from the copyright owner. If you are the copyright owner and would like this content removed from this site, please contact me.

Contact Info

- Charles J. Vella, PhD
- www.charlesjvellaphd.com
- charlesvella@comcast.net
- 415-939-6175