Paleogenetics, Part 8

CHARLES J VELLA, PHD, 2022

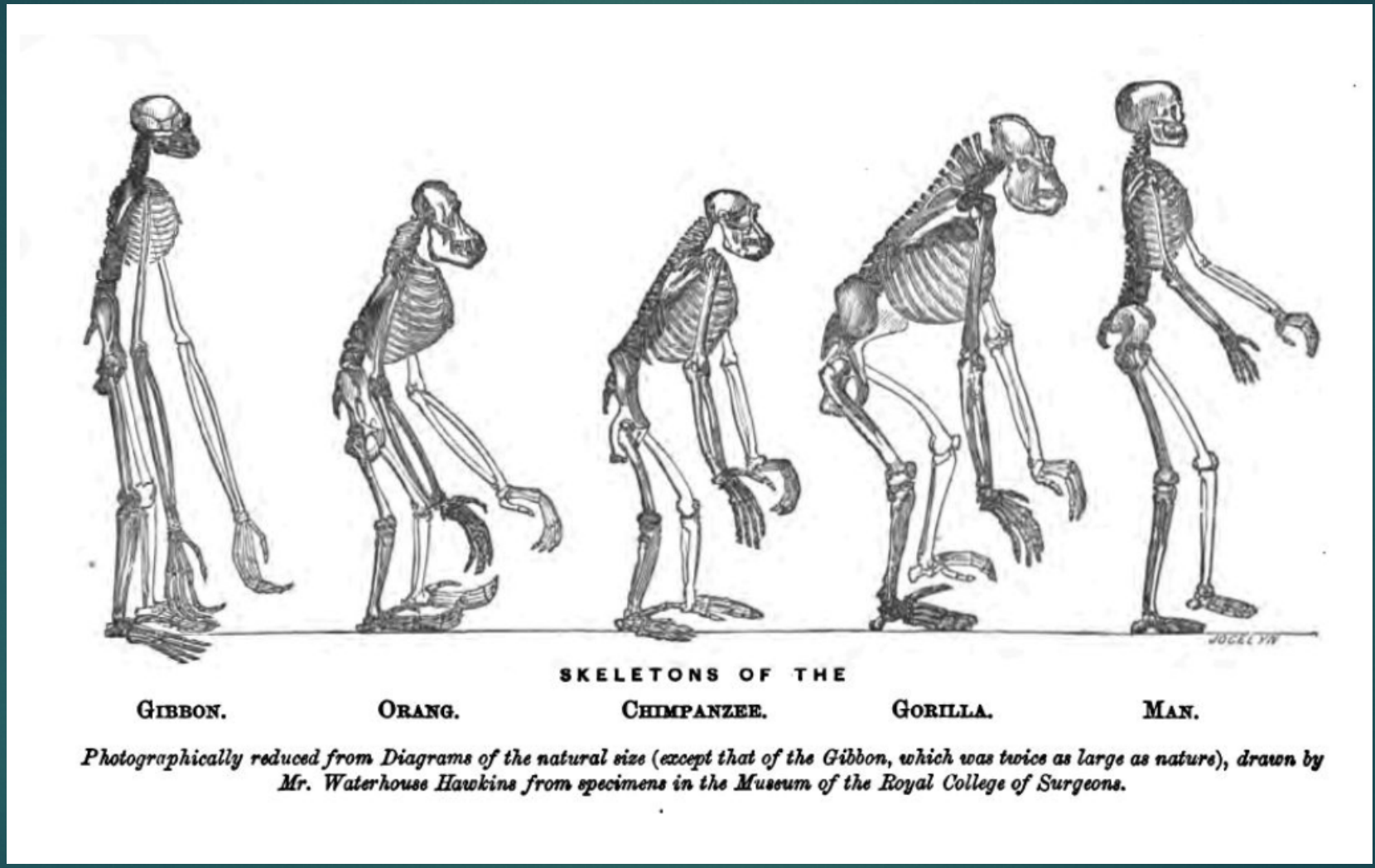
** June 2022 Updates

Don's Maps

- Resources for the study of Palaeolithic / Paleolithic European, Russian and Australian Archaeology / Archeology
- Don Hitchcock of Australia
- https://donsmaps.com/index.html

- Photos of all things archeological and paleolithic

Thomas Huxley's 1863 image of evolution in Africa



Rudolph F. Zallinger's iconic image of evolution toward human specialness

Current iconic image of human evolution: Braided river

Primate observers: Biruté Galdikas, Dian Fossey, Jane Goodall, with Mary Leakey

Some comments on a putative Denisovan molar from Tam Ngu Hao 2, Laos. Chris Stringer

- Demeter et al. reported the discovery of a human lower molar from Tam Ngu Hao 2 (Cobra Cave) in northern Laos *(Nature Communications* 2022), dated at about 150,000 years old
- The authors have done an excellent job in describing and dating the find, but at the moment I'd prefer to add the word 'putative' to its attribution as a Denisovan.
- One of the problems is that there are as yet no comparable lower molars from Denisova Cave itself, so the tooth is indirectly linked to the Denisovan fossils through a series of morphological inferences using the Xiahe jawbone (another putative Denisovan).
- The text is commendably more cautious than the title of the paper about this being an actual Denisovan fossil, acknowledging that it could even be from a Neanderthal (although the authors consider that less likely).

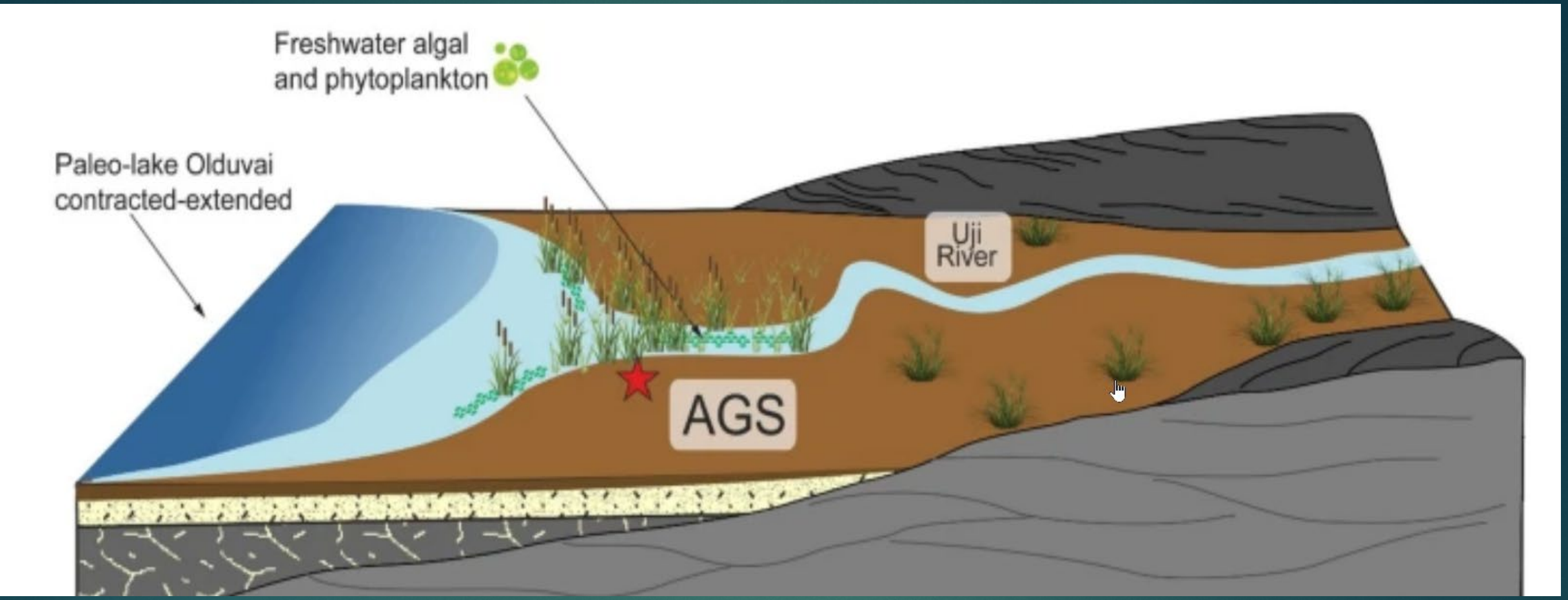
Laos: Denisovan?

- If it is a Denisovan fossil though, it would considerably extend the known physical and ecological ranges of these ancient humans into S. E. Asia, where it was previously guessed that they must have lived because of the presence of introgressed Denisovan-like DNA in extant populations in regions like the Philippines and Oceania (of course it would extend known ranges even further if it was actually a Neanderthal!).
- We will be more confident about finds such as the Tam Ngu Hao 2 molar and the Xiahe mandible being Denisovans if and when we have more complete Denisovan fossils or good quality ancient DNA or proteomes from Chinese fossils like Dali, Harbin, Xujiayao, Hualongdong and Jinniushan.
- Unfortunately, there is little chance of aDNA from the molar itself and enamel proteomic data could only identify it as from a probable female Homo individual, but there is obviously enormous potential for further exciting discoveries in this region of Laos.

A polar bear paleogenome reveals extensive ancient gene flow from polar bears into brown bears

- Polar bears (*Ursus maritimus*) and brown bears (*Ursus arctos*) are sister species possessing distinct physiological and behavioral adaptations that evolved over the last 500,000 years.
- Several extant and extinct brown bear populations have relatively recent polar bear ancestry, probably as the result of geographically localized instances of gene flow from polar bears into brown bears.
- Study: generated and analyzed an approximate 20X paleogenome from an approximately 100,000-year-old polar bear that reveals a massive prehistoric admixture event, which is evident in the genomes of all living brown bears.
- This massive admixture event mainly involved unidirectional gene flow from polar bears into brown bears and occurred as climate changes caused overlap in the ranges of the two species.

AGS site



AGS functioned as a seasonally waterlogged, low-vegetation environment characterized by dense accumulations of butchery-process debris within a wider mosaic environment that harbored both open and dense, closed ecotones

New site at Olduvai Gorge (AGS, Bed I, 1.84 Mya) widens the range of locations where hominins engaged in butchery

- Newly discovered Olduvai Gorge site (AGS) dated to 1.84 Ma, enabling an analysis of the distributions of critical local landscape resources across an explicit locus of hominin activity.
- Our results reveal that AGS was a seasonally waterlogged, largely unvegetated lakeside site situated near an ephemeral freshwater river surrounded by arid-adapted C4 grasses. The sparse vegetation at AGS contrasts with reconstructed (micro)habitats at the other anthropogenic sites at Olduvai Gorge, suggesting that central-provisioning places depended more heavily on water access than vegetation viz. woody plants as is often observed for modern hunter-gatherers.
- As hominins at AGS performed similar butchering activities as at other Bed I sites, our results suggest they did not need the shelter of trees and thus occupied a competitive position within the predatory guild.

Lavinia M. Stancampiano,, et al., 2022

New AGS Olduvai site

- Original Olduvai sites: vegetation was dominated by trees and vegetation, including palms, sedges and cattails among patchy (paleo)wetlands
- Data at AGS indicates hominins took strategic advantage of low- or unvegetated locations in addition to dense woodland thickets, low-visibility papyrus stands, and tall grassland
- Hominins at AGS would have an unencumbered view of the surrounding landscape, including precious refuge about 400 m away at FLK Zinj itself

AGS

- AGS archaeological site contains one of the highest densities of faunal remains on the Zinj Paleolandscape; suggests that the area at AGS must have been occupied for repeated instances of large carcass consumption

- Evidence of carnivore modification of carcass remains; hominins carved a competitive niche against other predators by efficiently fending off their hazard.

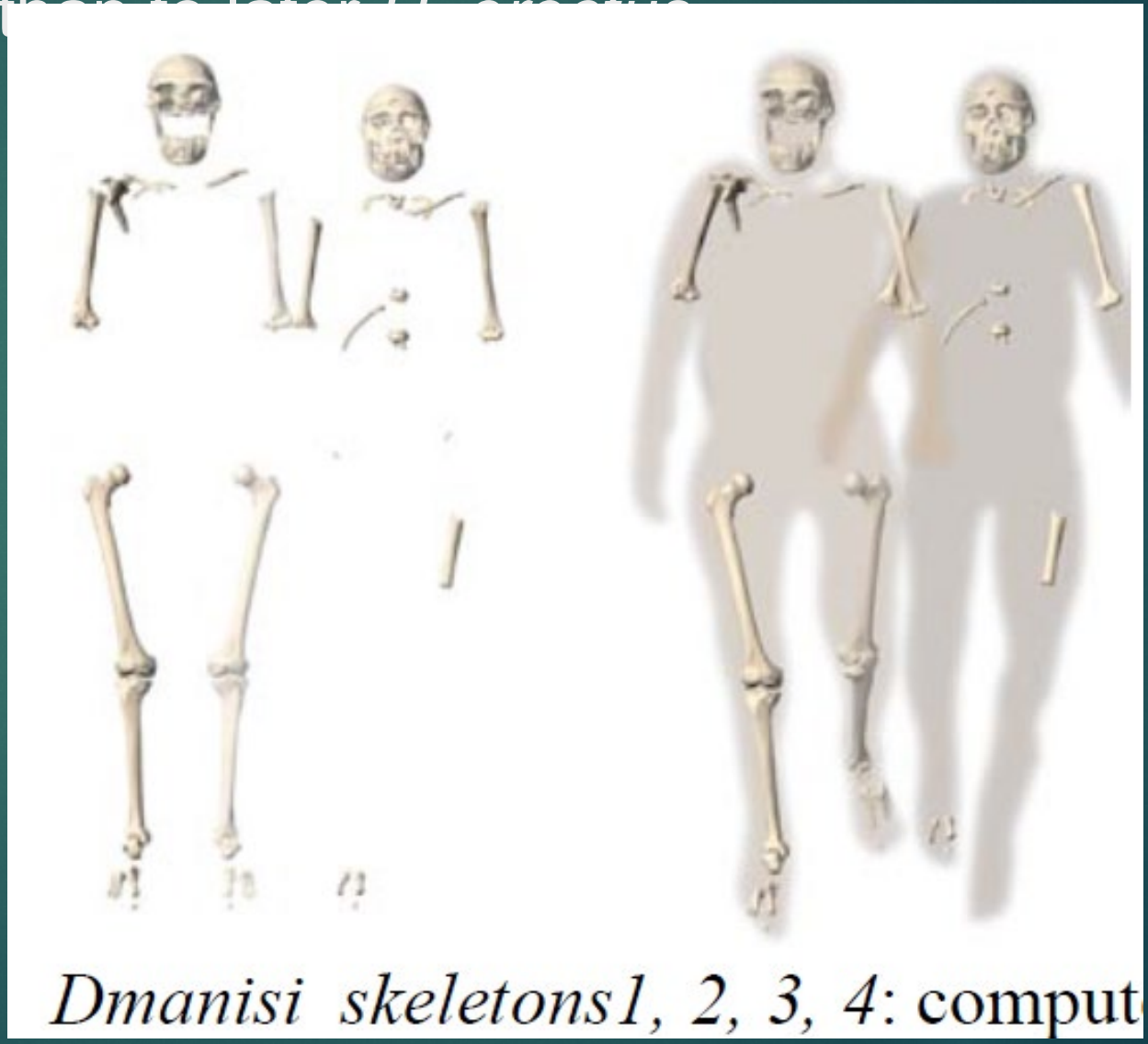
- Hominin engagement in social transport of large resources, such as bringing animal carcasses and freshwater-sourced food to AGS from the surrounding grasslands and lakeside environments.

AGS

- Conclusion: early hominins at Olduvai Gorge selected locations for cooperative resource processing—such as animal butchery—as related to water resources rather than refuge (i.e., closed thickets).
- This conclusion diversifies the environments in which anthropogenic sites occur and furthermore insinuates that hominins felt equally at ease in such environments.
- Considered together, new and old data at Olduvai reveal that hominins had reached an adaptive carnivore status by 1.84 Ma that enabled them to cope with terrestrial predation risks and fend off other carnivore competitors.

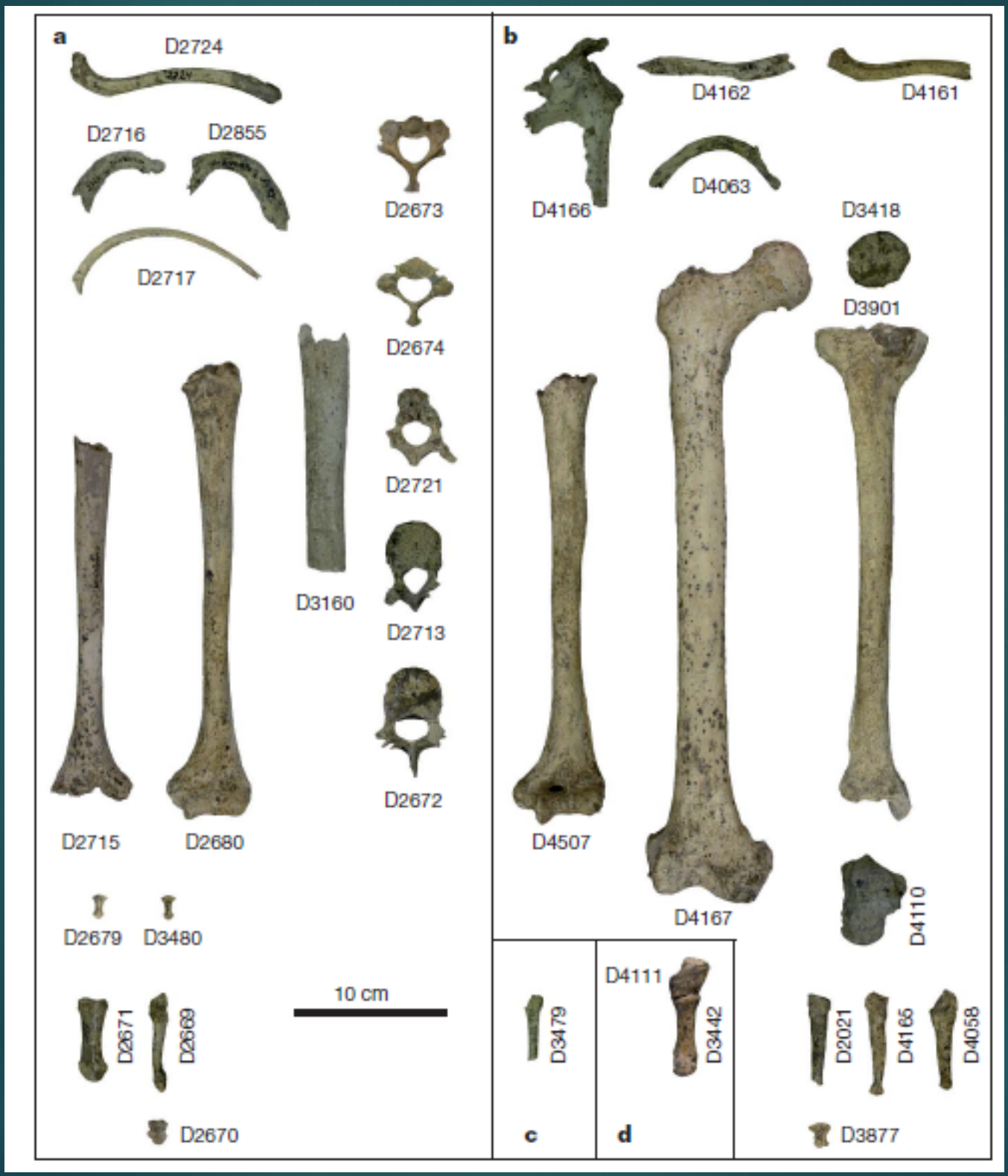
Dmanisi *H. erecti,* 1.8 Ma, **631 cc**: no fire, raw food, Oldowan tools;

Dmanisi postcranials: small body (145–166 cm; 4.8–5.4 ft) and brain size (545–775 cc), both of which are more comparable to *H. habilis* t



Dmanisi skeletons1, 2, 3, 4: comput

Dmanisi+human_femur, *Dmanisi+human_tibia*: thighbones (femora) and shinbones (tibiae) of the Dmanisi hominin (bottom of pictures) and of four modern humans
Picture credits: Tea Jashashvili, Univ. Zurich


a
D2724
D2716
D2855
D2673
D2717
D2674
D2721
D3160
D2713
D2672
D2715
D2680
D2679
D3480
10 cm
D2671
D2669
D2670
b
D4166
D4162
D4161
D4063
D3418
D3901
D4507
D4167
D4110
c
D3479
d
D4111
D3442
D2021
D4165
D4058
D3877

Mycobacterium tuberculosis

Largest case series yet of patients treated with bacteriophage therapy for antibiotic-resistant infections. 3 million antibiotic-resistant infections of all sorts occur in the United States annually, with 35,000 resulting deaths

Phage: ancestral enemy of bacteria

Phage Therapy

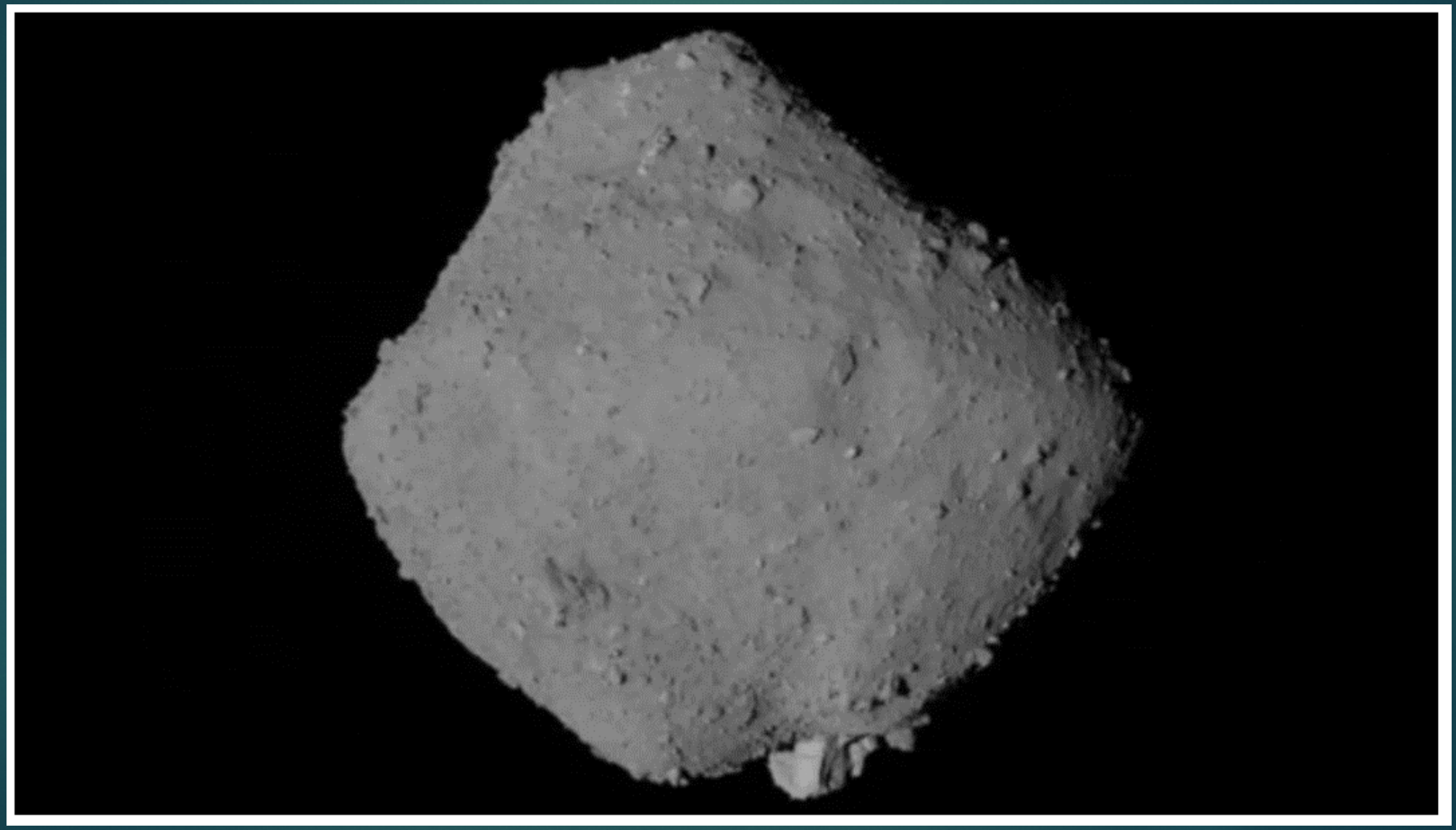
- Bacteriophages are viruses that have evolved to target and destroy specific bacterial species or strains.
- Phages are more abundant than all other life forms on Earth combined and are found wherever bacteria exist.
- 2016: experimental intravenous phage therapy to successfully treat and cure a patient who was near death from a multidrug-resistant bacterial infection.

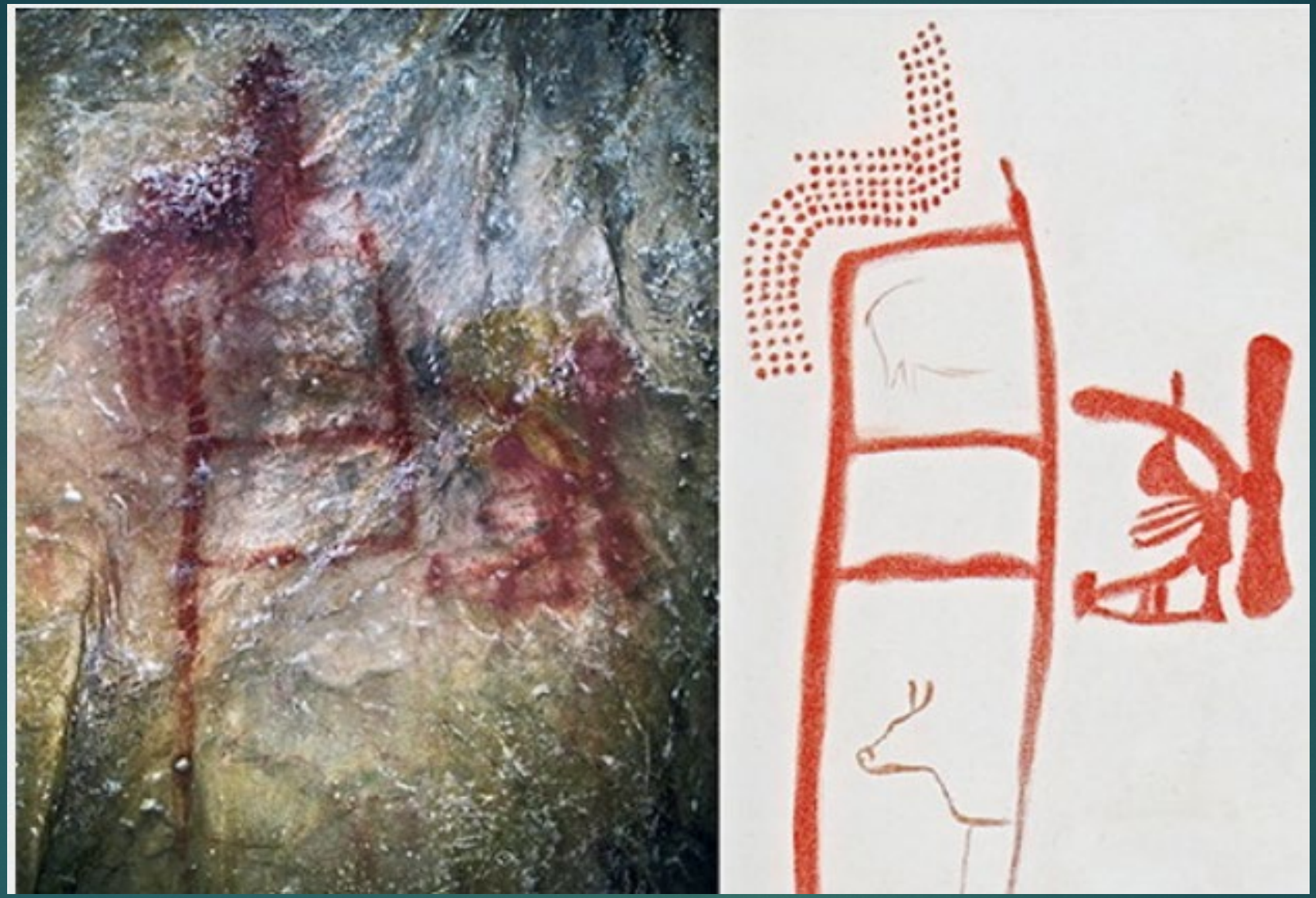
Phage therapy

- Each phage species seeks and destroys only one bacterial species and the current number of therapeutically useful phages is relatively small
- No adverse reactions to phage therapy in any of the patients, regardless of type of bacterial infection, types of phages used or method of treatment.
- Eleven of the 20 patients displayed some measure of symptom improvement or reduced bacterial presence.

Are we born with a moral compass?

Asteroid Ryugu: Rubber-ducky' asteroid 200 million miles away holds building blocks of life: 20 amino acids





Red scalariform (ladder shaped) sign, panel 78 in hall XI of La Pasiega gallery C.
This panel features the La Trampa pictorial group. The authors contended that the crust had been sampled and analyzed for a minimum age of (64 800 BP), which meant that the sign was older than this.

Critique of La Pasiega painting at 65 Ka

- **White, R., et al.**, (44 authors) 2019: Still no archaeological evidence that Neanderthals created Iberian cave art, Journal of Human Evolution:
- Urge caution in accepting our UTh dates on calcite associated with cave paintings (Hoffmann et al., 2018a) and conclusions that Neanderthals made paintings in three caves in Spain at least 65 ka ago. They argue that
- (1) uranium loss from calcite can lead to erroneously old dates and consequently UTh dates require validation by other methods;
- (2) there are problems with sampling methodology that can lead to unreliable dates; and
- (3) the existing corpus of 14C dates on cave art, including hand stencils and rectangular signs, argues against the production of any cave art before 38 ka.
- Believe results to be "especially troubling" as they "contradict more than one hundred years of research observations on the Neanderthal and modern

Hoffman et al.'s Response to White et al.'s, J. Hum. Evol. (2020)

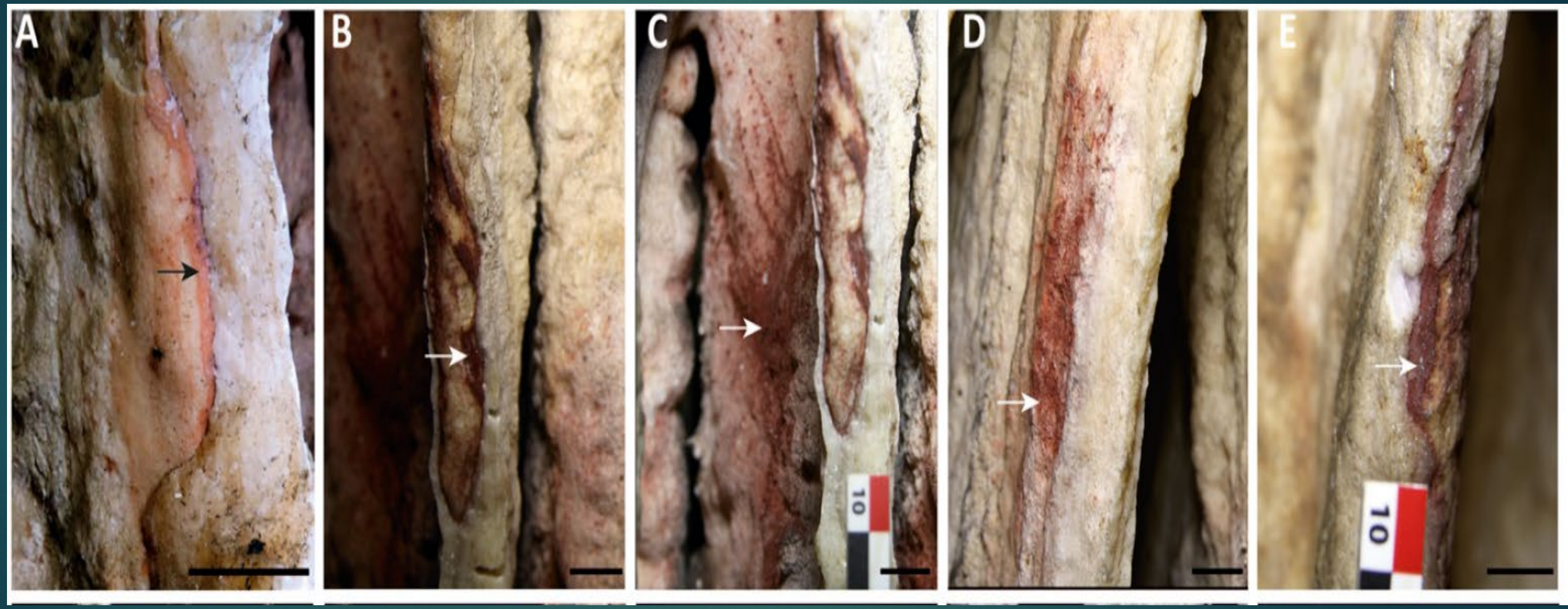
- The Iberian evidence of N art has been challenged, but all the criticisms have been exhaustively responded to.

The symbolic role of the underground world among Middle Paleolithic Neanderthals

- The dating of paintings in three caves from the Iberian Peninsula supports the view that Neanderthals developed a form of cave art more than 20,000 years before the emergence of anatomical modernity in Europe.
- This 2021 study: confirm that the paintings on a large speleothem from one of these sites, Cueva de Ardales, were N made, and we show that the pigments do not come from the outcrops of colorant material known inside the cave.
- Variations in the composition of the paint correspond to differences in the age of the paintings, supporting the hypothesis that Neanderthals used the speleothems symbolically over an extended time span.

Africa Pitarch Martía, et al., 2021

Red staining at Ardales Cave: panel II.A.3



Underground World of Ns

- At El Castillo cave in Spain, a minimum age of 40.8 ka was obtained for a red disk, consistent with Neanderthal authorship of Europe's earliest cave art as eventually corroborated by the nonfigurative paintings and hand stencils from three Iberian sites dated to >64.8 ka.
- Hand-stencil art from Borneo and a naturalistic painting from Sulawesi have yielded minimum ages of 40ka and 44 ka,.
- One of the early Iberian sites is Cueva de Ardales: microscopic and chemical analysis of panel II.A.3: >46 ka in Curtain 5, >45 ka and <49 ka in Curtain 6, and >66 ka in Curtain 8
- Analyzed natural Fe-rich coloring materials collected from the floor and walls of the cave staining is mineral in origin and cannot be interpreted as the result of microbial activity. The staining cannot be interpreted as the result of natural geological processes typically occurring in caves such as fluvial flows, infiltration from soils, percolating waters, or weathering of the walls

Underground World of Ns

- Results strongly support that the Paleolithic artist(s) used Fe-rich ochre lumps collected in geological formations from an as yet unknown source likely to be found outside the cave.
- The dating evidence implies a minimum of two incursions. Certain that our samples represent a minimum of two painting events, and we can additionally suggest that the real number is probably at least three, or maybe even four.
- Hypothesis: believe that the dome is the symbol, and the paintings are there to mark it as such, not the other way around.. Rather, it must stand instead for the renewed assertion of the symbolic value of the place or of the "canvas" itself.

Bruniquel Cave in France

Underground World of Ns

- The result of graphic behaviors intent on perpetuating the symbolic significance of a space.
- The evidence from Bruniquel cave, in France, shows that Middle Paleolithic Neanderthals were involved in symbolic activities taking place deep inside the karst that included the intentional modification of speleothems and their use in the construction of complex arrangements.
- The evidence from Cueva de Ardales supports the notion that speleothems played a fundamental role in the symbolic systems of some Neanderthal communities

Southern Spain art sites

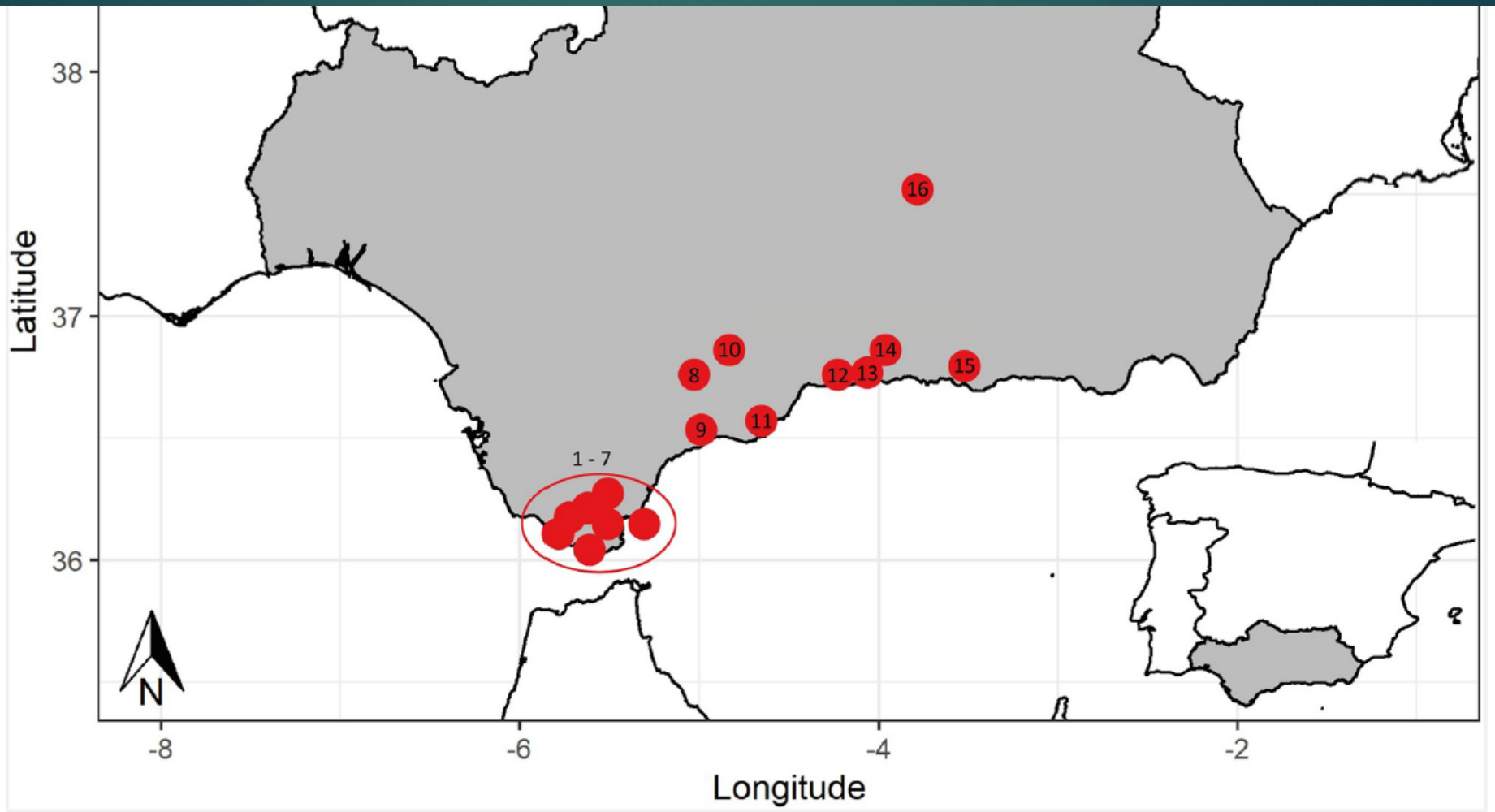


Fig 18. Distribution of rock art sites with non-figurative red paintings and handstencils in Andalusia and Gibraltar: 1: Moro, 2: Estrellas, 3: Palomas I, 4: Palomas IV, 5: Atlanterra, 6: Horadada, 7: Gorham, 8: Pileta, 9, Pecho Redondo, 10: Ardales, 11: Calamorro, 12: Navarro, 13: Victoria, 14: Higuerón, 15: Nerja, 16: Malalmuerzo.

The nature and chronology of human occupation at the Galerias Bajas, from Cueva de Ardales, Malaga, Spain

- The Cueva de Ardales is a hugely important Palaeolithic site in the south of the Iberian Peninsula owing to its rich inventory of rock art. From 2011–2018, excavations were carried out in the cave for the first time ever.
- The excavation focused on the entrance area of the cave, where the largest assemblage of non-figurative red paintings in the cave is found.
- A series of 50 AMS dates from the excavations prove a long, discontinuous, occupation history spanning from the Middle Palaeolithic to the Neolithic.
- A large assemblage of ochre lumps was discovered in the Middle Palaeolithic layers.

Ramos-Muñoz, J. et al. (2022) . *PlosOne*

Ardales Cave

- Cueva de Ardales is the most outstanding cave with Palaeolithic rock art in southern Iberia. The cave is located near the village of Ardales. in 1821 after an earthquake exposed a cave entrance; previously sealed by colluvial deposits.
- Cueva de Ardales was not a campsite, but was mainly visited for the production of rock art or the burial of the dead.
- It was not until 1918 that Henri Breuil recognized the Palaeolithic rock art .
- The site is a multi-branched karstic system that is divided into five areas. The cave contains over 1,000 paintings and engravings found on a wide variety of surfaces

Ardales Cave

- Dots and lines are the most common motifs in European Palaeolithic rock art.
- Reflect a very long tradition of marking cave walls beginning in the Middle Palaeolithic, or even earlier, and through to the Upper Palaeolithic or even younger phases.
- A series of 50 radiocarbon and 12 U/Th dates obtained within the framework of the archaeological excavation confirms a long history of human occupation in Cueva de Ardales.
- Neanderthals entered the cave in the Middle Palaeolithic, over 65,000 years ago. They left traces of symbolic practices on the cave walls and of tool maintenance in zones 3 and 5. Thereafter the cave was repeatedly visited by humans all the way to the late Neolithic period. However, the excavation has also revealed long hiatuses in human activity.

Ardales Cave

- Abstract red motifs predominate in the entrance and adjacent areas: signs of different shapes and sizes, and two black handprints painted in negative.

- On the other hand, in the interior there are many engravings and paintings of animals (horses, cattle, goats, snakes, etc.), which have been observed in styles from the Gravettian, Solutrean and Magdalenian periods.

- There are areas where black pigments overlap other older reds.

Red painting/staining of speleothems

Cueva de Ardales, Malaga, Spain

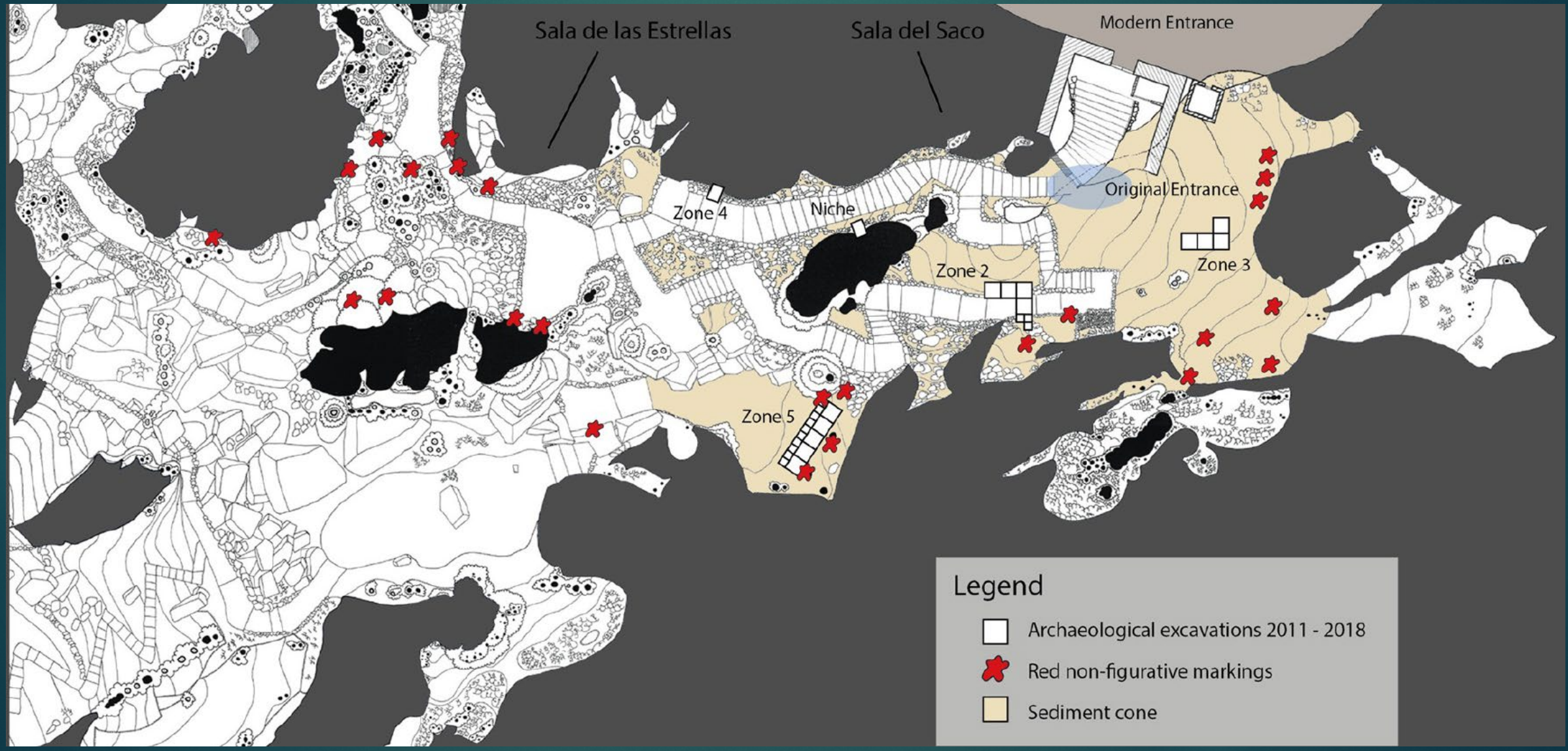
- All this spatial and pigment distribution suggests that the red abstract motifs are the earliest representations including walls, ceilings, ground rocks, speleothems and collapsed blocks. They are mainly dated to the Upper Palaeolithic, although the recent U/Th dating of carbonate crusts on abstract red depictions revealed that some of them are of Neanderthal authorship .
- Throughout the entrance area, rock art is found in the form of abstract depictions of varying size and shape and hand-stencils. Except for two black negative hand-stencils located in the Sala de las Estrellas, the other paintings were executed in red colour.
- The entrance area is different from the interior zones of the cave, where over 90% of the animal representations were documented

Ardales Cave

- In 2018, different samples of carbonate crusts deposited on the red paintings of stalagmitic curtains were dated using uranium-thorium, obtaining amazing dates of between 65 ka and 45 ka.
- This indicated a use of this site by Neanderthals during 20 ka, since it was the only human species that inhabited the area in that chronology. That paper also gave dates of about 65 ka for paintings from two other Spanish caves: a handprint in Maltravieso (Cáceres) and a ladder-shaped sign in La Pasiega (Cantabria).
- This work was contested with reservations about the dating method, especially in Maltravieso and La Pasiega, while they did recognize a minimum date of 47 ka for the Ardales pigments, although with doubts about its natural origin.
- In 2022 response, the authors returned to provide more information on the validity of the method and subsequently demonstrated that the pigment of the Ardales stalagmite had a human origin and was not natural.

DL Hoffmann et al, 2018; L. Slimack et al, 2018; DL Hoffmann et al, 2018; A. Pitarch Martí et al, 2021

Cueva de Ardales, Malaga, Spain



Ardales Cave

- This 2022 study: summarizes the work carried out between 2011 and 2018 in five excavation zones.
- In zone 5, a Middle Paleolithic level stands out, with a blade and a flake, Gravettian materials (33 Ka), with at least 25 stone artifacts (points, blades, flakes and remains of carving) as well as a perforated deer canine and other personal adornments. made with mollusk shells, and an Aurignacian level, although it only has two flakes and requires more study. On the surface, there is a calcified rope from the 16th or 17th century.
- In zone 3: where the oldest dates have been obtained, with samples between 52 and 56 ka.

Ardales Cave

- Neandertals entered Ardales for the first time more than 58 ka ago, and used the cave until about 50 ka ago and then again between 43-46 ka ago.
- Subsequently, there is a temporary hiatus of more than 7 ka, which precedes the arrival of modern humans. This begins with a possible Aurignacian occupation (36 ka), which will require more studies, since it is difficult to determine the human presence in the south of the Iberian Peninsula during that period. Subsequent occupations in the Gravettian, Solutrean and Magdalenian periods also did not leave abundant traces of activity.

Ardales Cave

- The Neandertal presence in the cave was sporadic.
- The evidence of domestic activities is very scarce (few tools and lack of fires that indicate a possible camp): they probably did not live there, but rather the
- cave was used for symbolic reasons by different human groups, beginning with the Neanderthals, who between 65 ka and 45 ka left various graphic interventions in the form of signs on the speleothems.

Ardales Cave

- In Cueva de Ardales, most abstract red motifs are located in the entrance area and the adjacent zones rather than at the back areas.
- Painted and engraved animal depictions are, on the contrary, dominating the interior areas.
- This spatial distribution, as well as four instances in which black pigments were laid over earlier red suggests the non-figurative red motifs are the earliest representations in the cave.
- The U-series dating of calcite accretions superposing some of these non-figurative paintings indicate that they were made at least 65 ka ago.

Ardales Cave

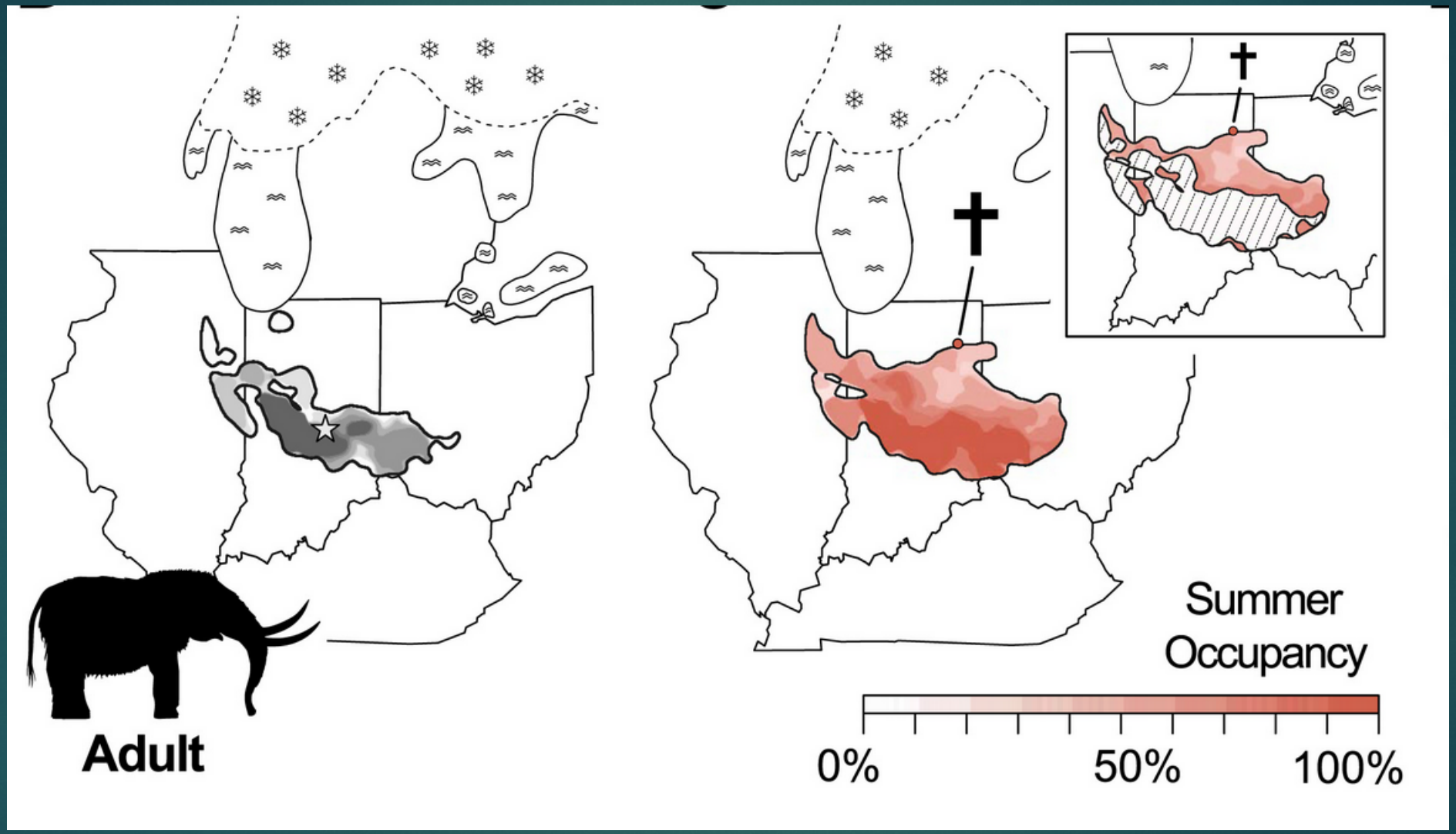
- In addition, excavations resulted in the discovery of a significant number of potential ochre lumps in all chronological phases, including Middle Palaeolithic stratigraphic units.

- All this evidences supports the hypothesis that non-figurative paintings represent, indeed, the beginning of a long rock art tradition in Cueva de Ardales.

Fred the Mastodon from Indiana: studying strontium and oxygen isotopes signatures recorded in his tusk

- Over 13,000 years ago, an American mastodon roamed what is today the American Midwest.
- Year after year, he returned to an area in northeast Indiana — believed to be a mating ground. It was there that he died in battle.

Joshua H. Miller, et al., 2022


Adult
Summer
Occupancy
0%
50%
100%

Early-onset familial Alzheimer's disease genes (Onset < age 60)

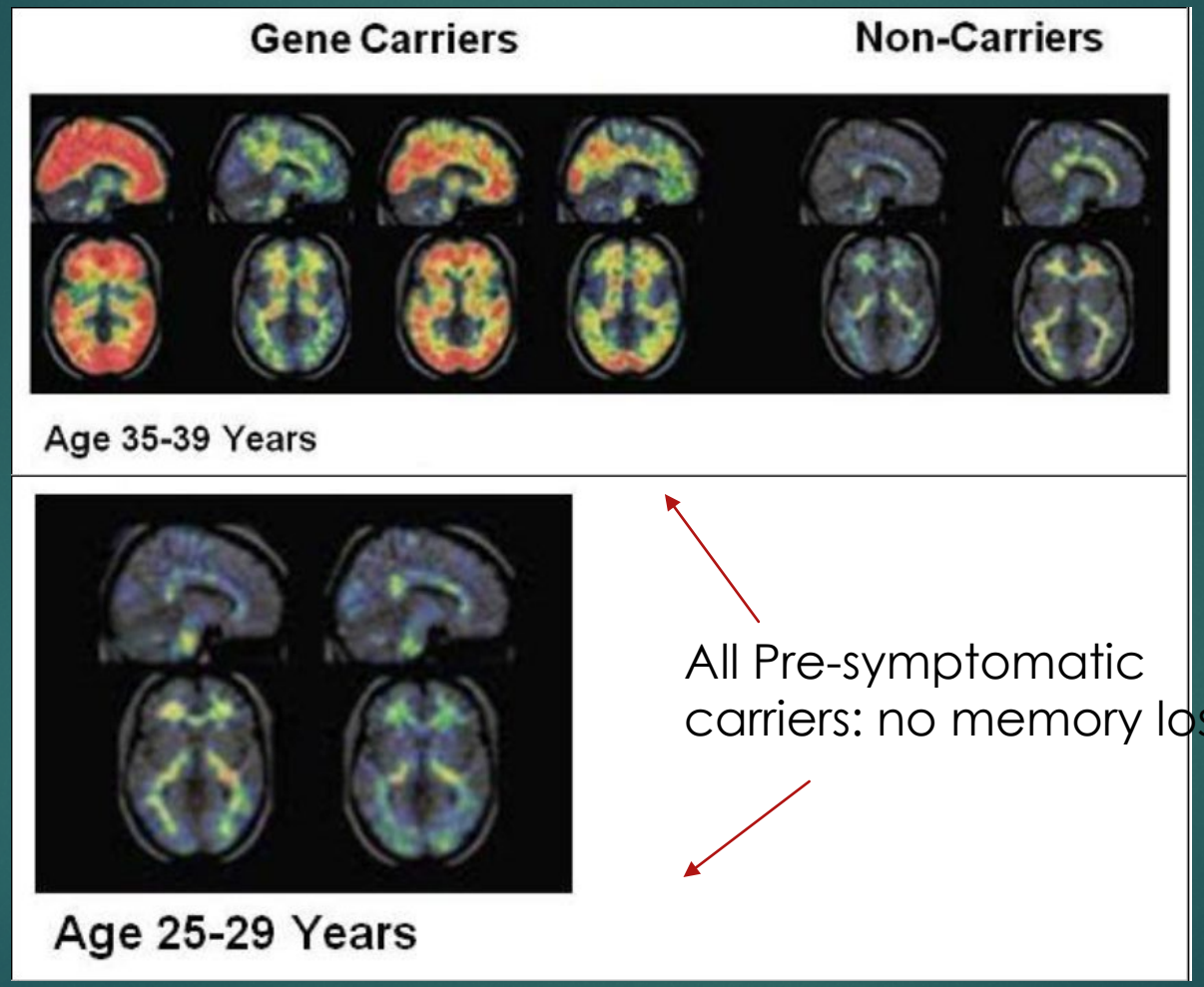
- Deterministic genes in Familial AD (FAD):
  - 3 Causative Mendelian Genes = 1% of all AD; with autosomal dominant (50%) transmission;
  - early onset, age 40-50s; less than 1% of AD is genetic; ~450 families in world:
    - Presenilin 1 (PS1) (chromosome 14) – most common
    - Presenilin 2 (PS2) (chromosome 1)
    - Amyloid Precursor Protein (APP) (chromosome 21)
  - All 3 AD genes (+ APOe4) create excessive accumulation of Aβ peptide

Ken Kosik, 2021: Colombia: Early onset Alzheimer's clust

Antioquia, Medellin region, Colombia: 6,000-member Family: Amyloid Deposition in Genetic Mendelian (Presenilin 1) AD

Mild NCD: median age 44

Major NCD: median age 49



Hope for near future:
Colombian Prevention Study

- Eventually treat AD like HTN and heart disease preclinically
- Colombian study: extended clan of 5,000 people who live in Antioquia, Colombia with early onset AD
- N = 242; 5 year trial; Genentech drug, Crenezumab (removes BA fibrils, but not plaque) injection every 2 weeks, discovered initial dose insufficient, then doubled it; discovered that 54% had no BA at beginning; then infusions in last 2 years; massive pre and post imaging, plasma, cognitive testing; 94% completion rate!
- First Data in June 15, 2022: Trial did not meet - **Crenezumab did not slow or prevent cognitive decline**
- **Final data in August.**

Surveying the landscape of rare mutations behind neurodegenerative diseases in Colombia

- Kenneth S. Kosik, UCSB Lab: Colombian family afflicted with a genetic form of early-onset Alzheimer's disease & consequences of the colonization carried out by the Spanish almost five centuries in the past.

- Family members that carried the so-called Paisa mutation (PSEN1 E280 mutation) would, like clockwork, begin to develop the hallmark forgetfulness of Alzheimer's—what they called la bobera ("the foolishness")—in their 40s.

Multiple mutation variants discovered

- New research: a concentration of unrelated families with different mutations that resulted in the same observable characteristics.
- Eleven deleterious variants were identified in the PSEN1 gene
- Among these PSEN1 variants, six were of European origin, three were Native Americans, and one African
- Common ancestor for each of three families originating at about the same time
- All the families with pathogenic PSEN1 mutations had autosomal dominant inheritance
- Five of the eight disease associated variants identified in the ALS panel were of Native American origin while only two were of European ancestry

Due to founder effects

- Demographic history plays a significant role in shaping a population's genetic risk for disease.
- In the Colombian population, founder effects led to a large number of ancestral disease-causing alleles from each of three admixed continents
- The famous Paisa mutation had been traced back to a single founder a conquistador from early Habsburg Spain, while another mutation was traced back to West Africa. Yet another was identified with Native American roots.
- How had they all emerged in the rural countryside of Colombia?

Founder effects

- Admixture among Native Americans, Spanish invaders, and enslaved Africans, all of whom passed through a population bottleneck due to widespread infectious diseases that left small isolated local settlements. As a result, the current population reflects multiple founder effects derived from diverse ancestries

- During the time of Spain's conquest of the Americas, Europeans first showed up on the northern shores of what would become Colombia in the early 16th century. Later that century, West Africans would be added into the mix as enslaved labor.

Colombia: dementia genes and infections

- People weren't the only ones mixing in Colombia at that time.
- People do not travel alone. People travel with their bugs.
- According to the researchers, the people traded their bugs with others who had never encountered them before and the population suffered massive mortality from numerous infectious diseases, including smallpox, influenza, syphilis, hepatitis, measles, encephalitis, tuberculosis, diphtheria, cholera, typhus, scarlet fever and meningitis.

Influx of Spanish diseases caused population crash (90% death rate in indigenous); after genetic diseases, populations map to isolated geographic settlements

Basis of the Population Structure

Gene flow carried by indigenous migratory streams

Caribs (Catías, Nutabes & Tahamíes) in the North. Quimbaya in the South

Spanish invasion Basque, Castillano, Sephardic

Tri-continental admixture

Genetic diversity

West African slave importation from diverse populations

Bottleneck

Smallpox
bubonic plague
chicken pox
pneumonic plague
cholera
diphtheria
influenza
measles
scarlet fever
typhus
tuberculosis
whooping cough

Geographically localized expansion

Result: Genetic disease map corresponds to the geographic map

Pre-Columbian
12.5-70,000 yrs ago → Conquistadores
1492-1550 → Colonial period
1550-1810

Slaves in Colombia

- In the 16th century, West Africans were imported into the country as enslaved people, and, like the conquistadores of the same era, mixed and mingled with the local population

- Smaller African family presenting the same dementia symptoms, unrelated to Spanish family

- Their mutations originated on two separate continents, in two different populations. Both PSEN1 mutations persisted.

Colombia: dementia genes and infections

- The collapse of the population caused a narrowing of the gene pool—a genetic bottleneck in which the few remaining survivors, who also happened to possess these rare mutations, became the founders of subsequent populations, passing their mutations to their offspring.
- A long period of colonization followed in which migration slowed and people mostly stayed put. The lack of new genes, isolated pockets of people and large families helped establish the mutations within the population.
- In Colombia is that the genetic map overlaps with the geographic map, because you can actually match people with a mutation to a certain region.

Colombia: dementia genes and infections

- When you put a selective pressure, such as pathogens, into a population, the question that arises is, do the survivors have any benefit over the ones who died?
- Could having these mutations—in particular variants of the PSEN1 (presenilin) gene—have conferred some sort of protection to their carriers against the onslaught of infectious disease in the early days of colonization, thereby ensuring their survival?
- Classic example: people who are heterozygous (one copy of a mutated gene from one parent and one normal copy from the other) for sickle cell anemia are resistant to severe malaria. The mutation that leads to the sickle-shaped cells also "interferes with the reproductive cycle of the malaria parasite.

Genetic bottlenecks, AD variants, and infections

- Identified 13 different pathogenic PSEN1 variants from multiple ancestral origins, nearly all attributed to founder effects.
- The PSEN1 mutations emerged from a small effective population in each of the early settlements that constituted a patchwork of bottlenecks dispersed throughout the country.
- Because people tended to remain geographically isolated, the rare variants represent a local genetic footprint.
- Survivors who emerged from the bottleneck had escaped the large number of infectious diseases responsible for decimating the population.

Genetic bottlenecks, AD variants, and infections

- During the historical period of colonization, populations in these settlements grew rapidly as the incidence of diseases diminished, which favored the segregation of potentially damaging variants at higher rates.

- The question arises as to whether the PSEN1 mutations could be under positive selection or are the mutations completely explained by drift.

- PSEN1 mutations cause the production of excess amyloid-beta, which may function as an anti-microbial peptide (AMP).

- In this manner, PSEN1 mutations may have been positively selected as protection against the enormous mortality of infectious diseases.

Genetic bottlenecks, AD variants, and infections

- Antimicrobial peptides function as an ancient component of the innate immune system that target bacteria, mycobacteria, enveloped viruses, fungi, and protozoans. Beta Amyloid is active against at least eight common and clinically relevant microorganisms, and several anti-amyloid-beta clinical trials have reported increased rate of infections among the participants
- This is still hypothetical.
- This work reinforces the need to include diverse populations for gene-trait association studies including populations that underwent bottlenecks as a source for gene discovery.

Native American gene

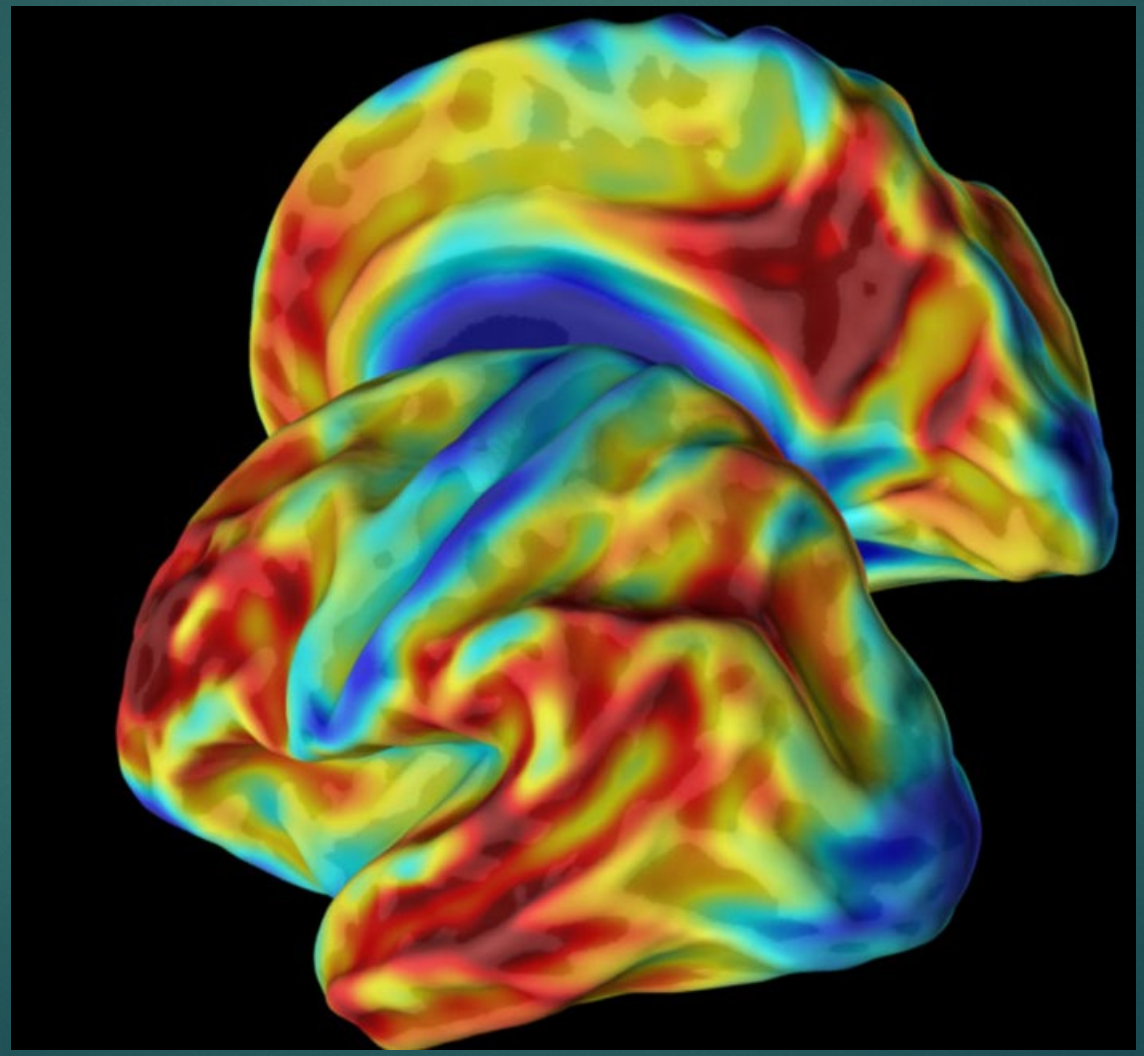
- Other findings: the same mutations and pathologies that present themselves in the European context as ALS (also known as Lou Gehrig's disease) present in the Native American lineage as frontotemporal dementia.

- Additionally, tracing the Native American dementia mutation backward through time, found that it originated with the first population of the Americas.

Sex-biased admixture: Native American women, European men

- The mitochondrial haplogroups of the probands were predominantly Native American (83.4%)
- The Y-chromosome haplogroups were mostly of European and of Mediterranean origins (92.8%),
- supporting the conclusion than multiple cohorts of Colombian origin show sex-biased admixture with Native American maternal lineages and paternal lineages from Europe

One lucky Colombian woman with PSEN1:
massive BA load,
but no dementia



Christchurch variant of Psen1

- One woman in the family with the Paisa mutation lived her entire life (into 70s) without developing Alzheimer's,
- possibly because a second copy of the rare genetic mutation decoupled the two hallmark signs of the disease:
- amyloid beta plaques (which was massively found in her brain after her death), and tau tangles (which were not found).
- BA without Tau does not produce dementia

Tool use and orangutans

- Among mammals, humans are the most-proficient tool users, with our close relatives, wild chimpanzees, placed second on that list. The pattern isn't perfect — wild bonobos and gorillas display much less tool use than do chimpanzees, and all three African great apes are closer to us on the primate family tree than orangutans are.
- Motes-Rodrigo et al. reasoned that if — when given the right materials — orangutans could perform the same basic techniques for stone-tool manufacture as early hominins, then the common ancestor of humans and orangutans might have had the same ability
- None of the orangutans followed the sequence that might be expected to correspond to early hominin behavior of striking the hammer on the core, and then using the resulting sharp flake to complete a task (making a cut to access the food box). But did exhibit percussive behavior (hitting stones on walls).

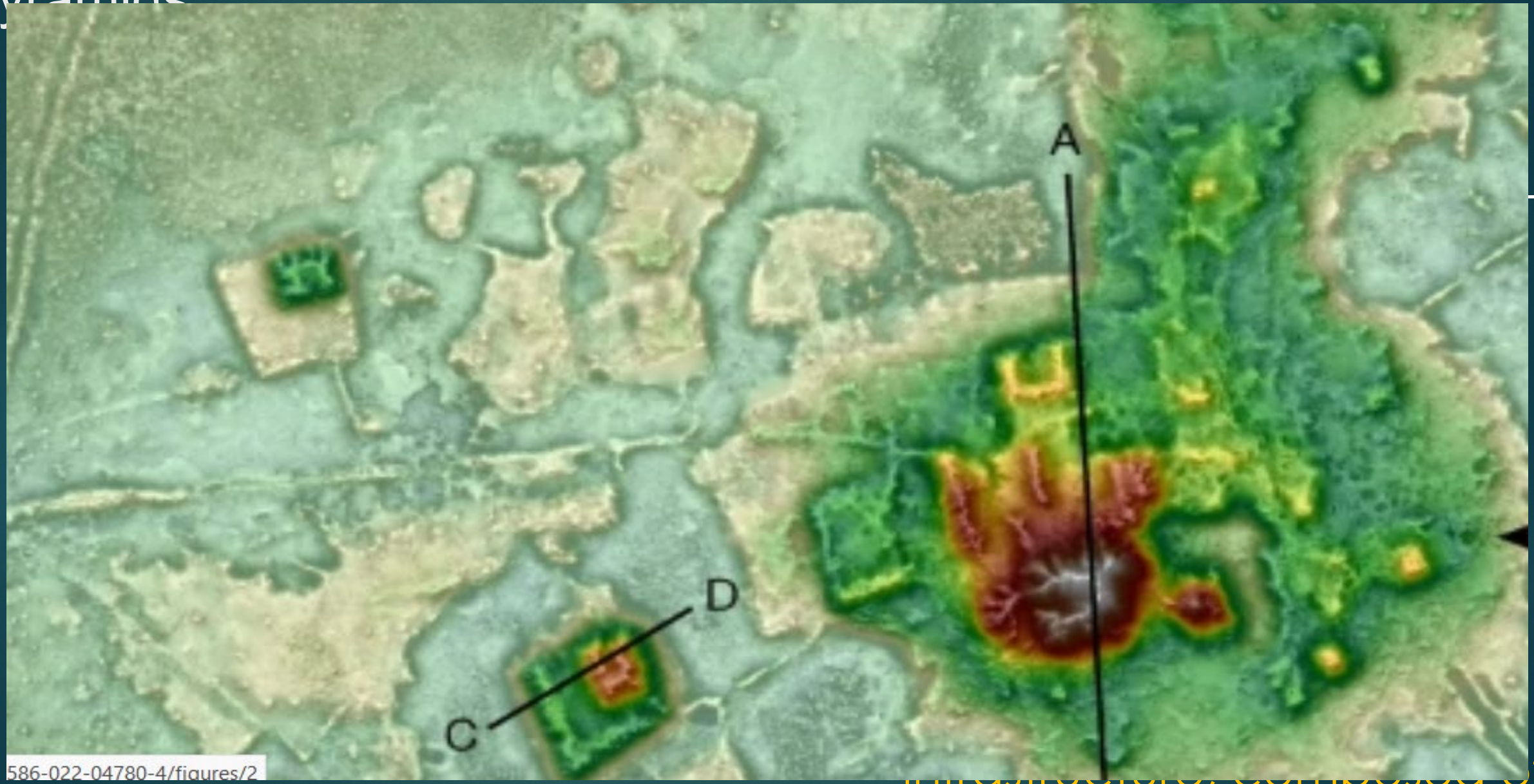
Orangutans

- Some wild orangutans make and use stick tools to extract seeds from tough shelled fruit and to prise insects from tree holes. However, stone-tool use has never been seen in wild orangutans, despite decades of observations.
- In the 1990s, when captive capuchin monkeys (Sapajus apella) were given similar equipment to the orangutans, they broke the stone cores and used the resulting flakes to cut through a barrier to get food. The capuchins (who are adept at living on the ground and in trees) even innovated a technique whereby they used one stone piece as a chisel that was hammered through the barrier using a second stone. Unlike modern orangutans, many wild capuchin groups both use and break stones.
- Captive primates have a greater aptitude for tool use than do their wild counterparts and the exposure of captive apes to humans (or enculturation) is one proposed reason for this difference (along with captive animals' greater amount of free time, potentially lower stress levels and greater access to new objects).

Functional connectivity identifies liberals vs conservatives

- Liberals and conservatives have noticeable and discriminative differences in functional connectivity that can be identified with high accuracy using contemporary artificial intelligence methods and that such analyses complement contemporary models relying on socio-economic and survey-based responses, esp. parental conservatism.
- Functional connectivity signatures from retrieval, empathy (disgusting image response), and monetary reward tasks are identified as important and powerful predictors of conservatism, and activations of the amygdala, inferior frontal gyrus, and hippocampus are most strongly associated with political affiliation.
- The direction of causality is unclear.

City in the Amazon via Lidar: 1500 years ago, 22 pyramids

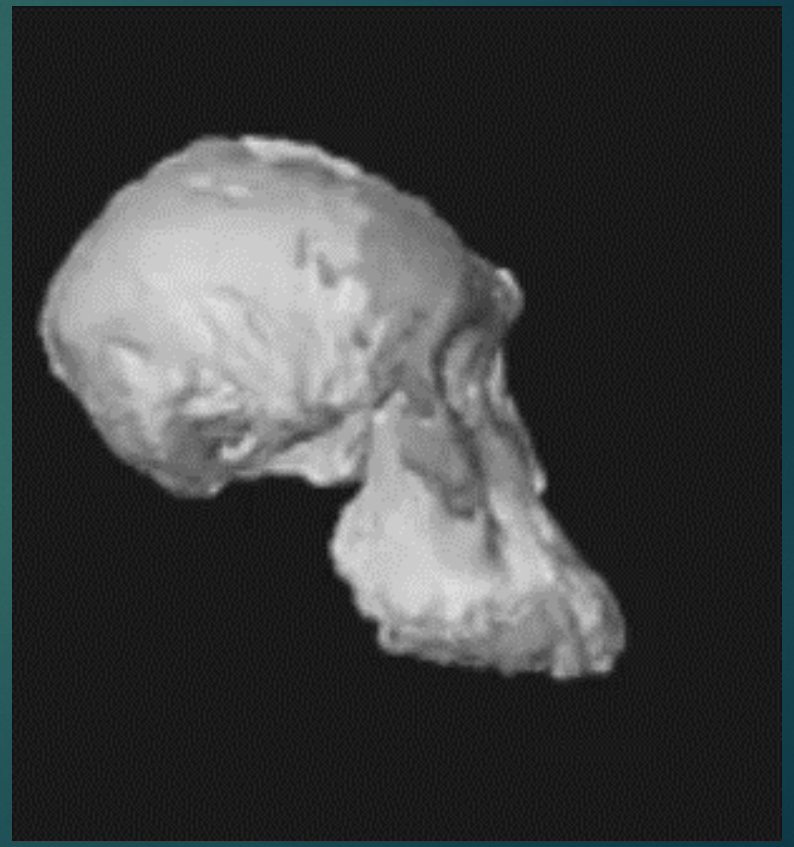


infrastructure, composed of

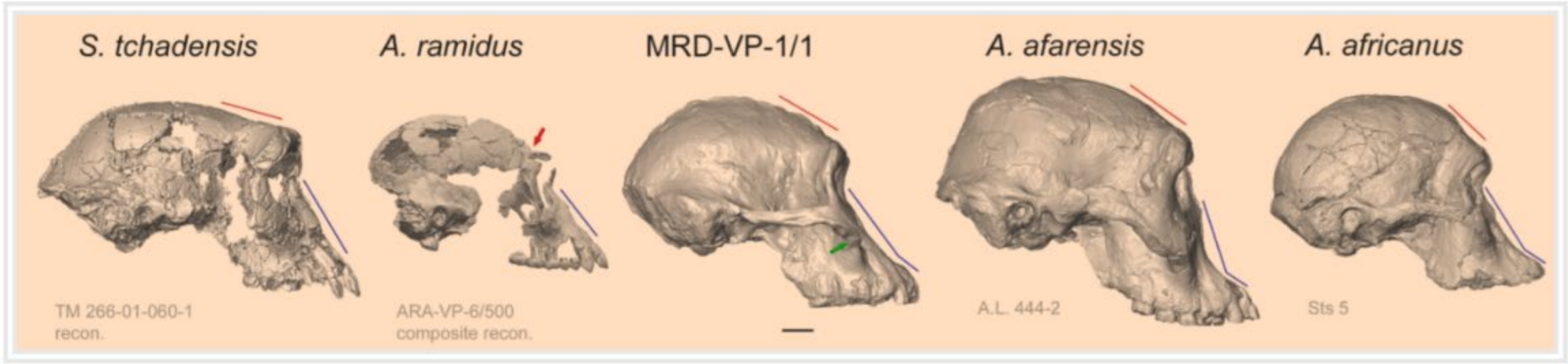
Lidar reveals pre-Hispanic low-density urbanism in the Bolivian Amazon

- Lidar, which stands for *Light Detection and Ranging,* is a remote sensing method that uses light in the form of a pulsed laser to measure ranges (variable distances) to the Earth. These light pulses—combined with other data recorded by the airborne system — generate precise, three-dimensional information about the shape of the Earth and its surface characteristics.
- A lidar instrument principally consists of a laser, a scanner, and a specialized GPS receiver.
- Airplanes and helicopters are the most commonly used platforms for acquiring lidar data over broad areas.
- Two types of lidar are topographic and bathymetric. Topographic lidar typically uses a near-infrared laser to map the land, while bathymetric lidar uses water-penetrating green light to also measure seafloor and riverbed elevations.

2019: A. anamensis MRD skull- Woranso-Mille area of Ethiopia.



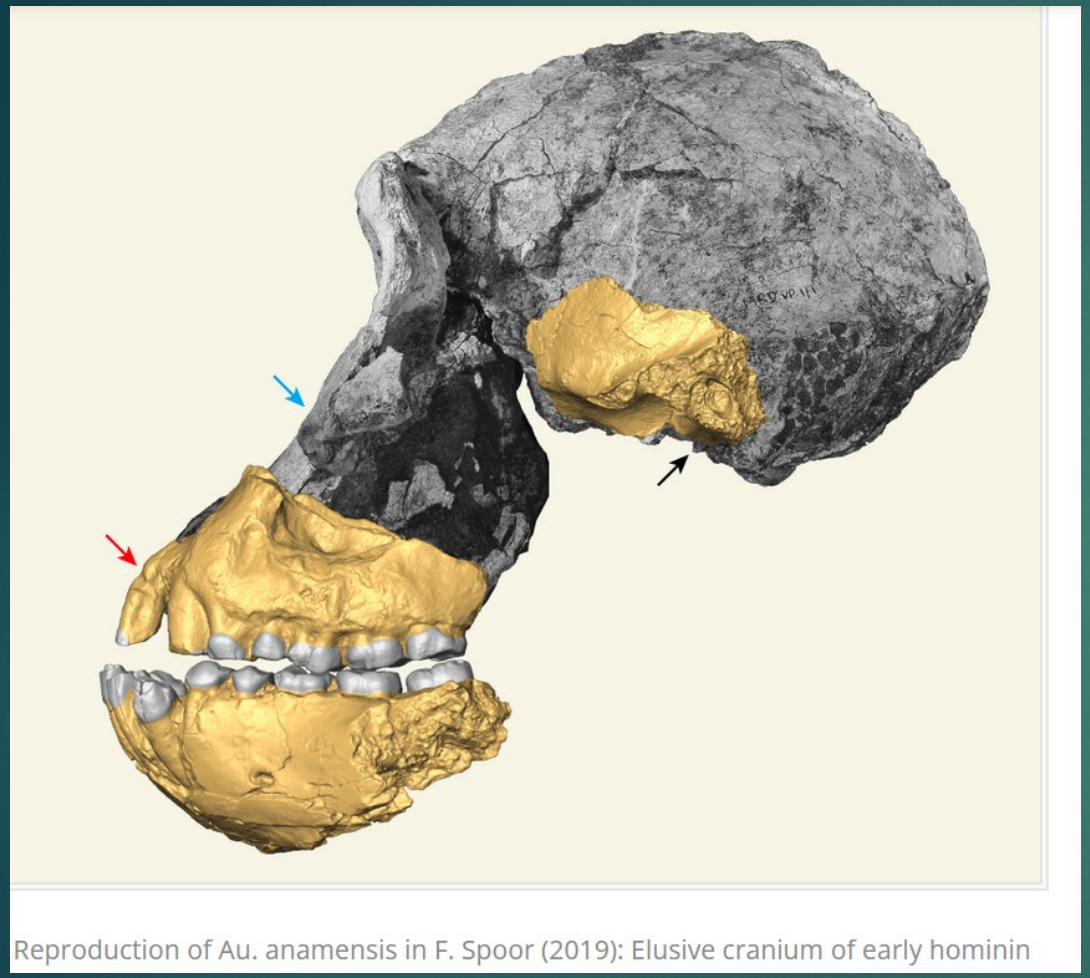
Comparison



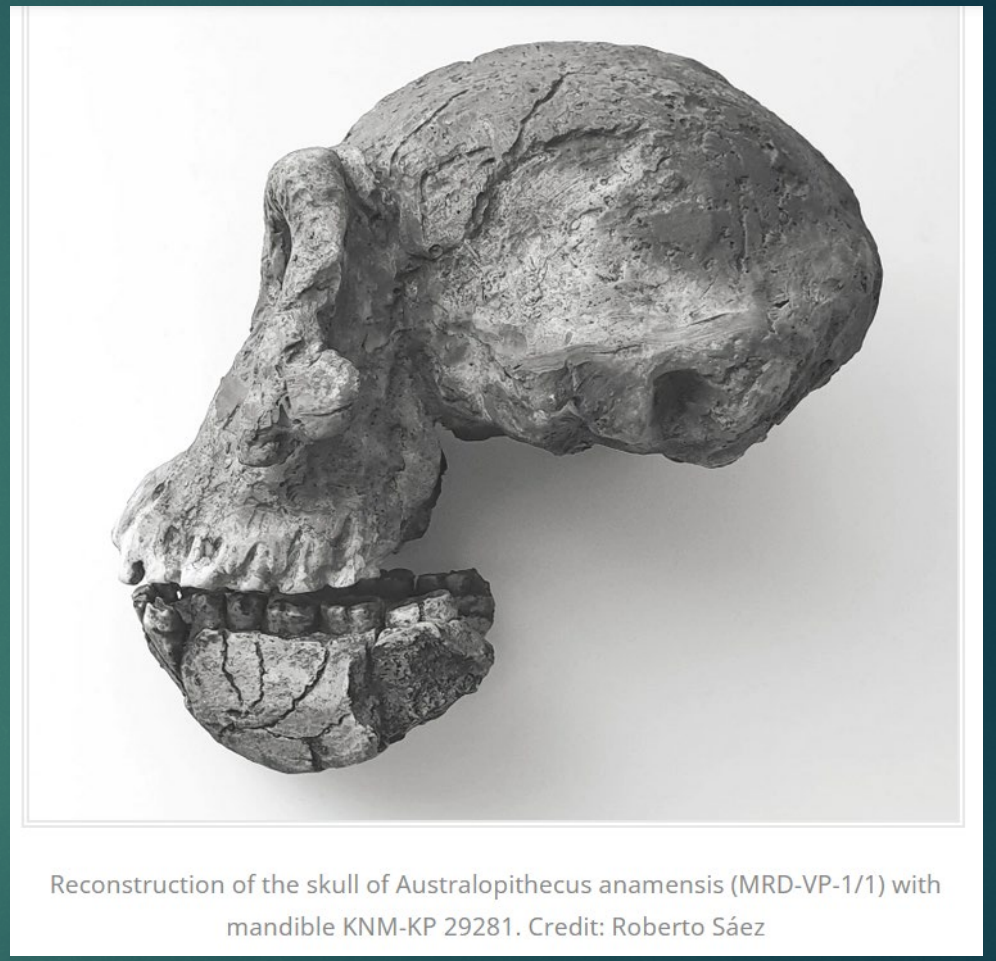
Comparison of skulls in lateral view. Credit: Haile-Selassie et al 2019

Australopithecus anamensis

Skull of Australopithecus Anamensis: 65-370 cc,



Reproduction of Au. anamensis in F. Spoor (2019): Elusive cranium of early hominin



Reconstruction of the skull of Australopithecus anamensis (MRD-VP-1/1) with mandible KNM-KP 29281. Credit: Roberto Sáez

In 2019, the first skull that could be associated with the species Australopithecus anamensis was published . It is the MRD VP-1/1 fossil, simply nicknamed MRD, found in 2016 in Woranso-Mille (Ethiopia), and dated at about 3.8 Ma .

Caffeine intake exerts dual genome-wide effects on hippocampal metabolism and learning-dependent transcription

- Caffeine is the most consumed psychoactive substance worldwide, with an inverse association between coffee/caffeine consumption and all-cause mortality.
- The impact of caffeine on human health follows an inverted bell shaped dose-response curve with benefits observable at doses of 200-400 mg per day.
- Habitual/chronic caffeine consumption normalizes synaptic plasticity and cognitive decline in ageing, Alzheimer's disease or other neuro-psychiatric conditions.
- Independently of its ability to favor arousal and attention, caffeine may exhibit cognitive-enhancing properties

Isabel Paiva, et al., 2022

Coffee helps cognition

- With caffeine reward,
  - Honeybees: able to remember a previously learned floral scent.
  - In rats: can enhance memory test performance.
  - In humans: caffeine intake immediately following learning improves discrimination performance 24 hours later.
  - Ability of caffeine to modulate hippocampal/cortical excitability in homeostatic conditions.
  - Indeed, caffeine treatment in hippocampal slices enhances basal synaptic transmission and modulates long-term potentiation (LTP) in rodents' hippocampus and sharp wave ripple complexes, that are proposed to underlie memory consolidation.
  - Caffeine also controls neuronal excitability and LTP-like effects in the human cortex

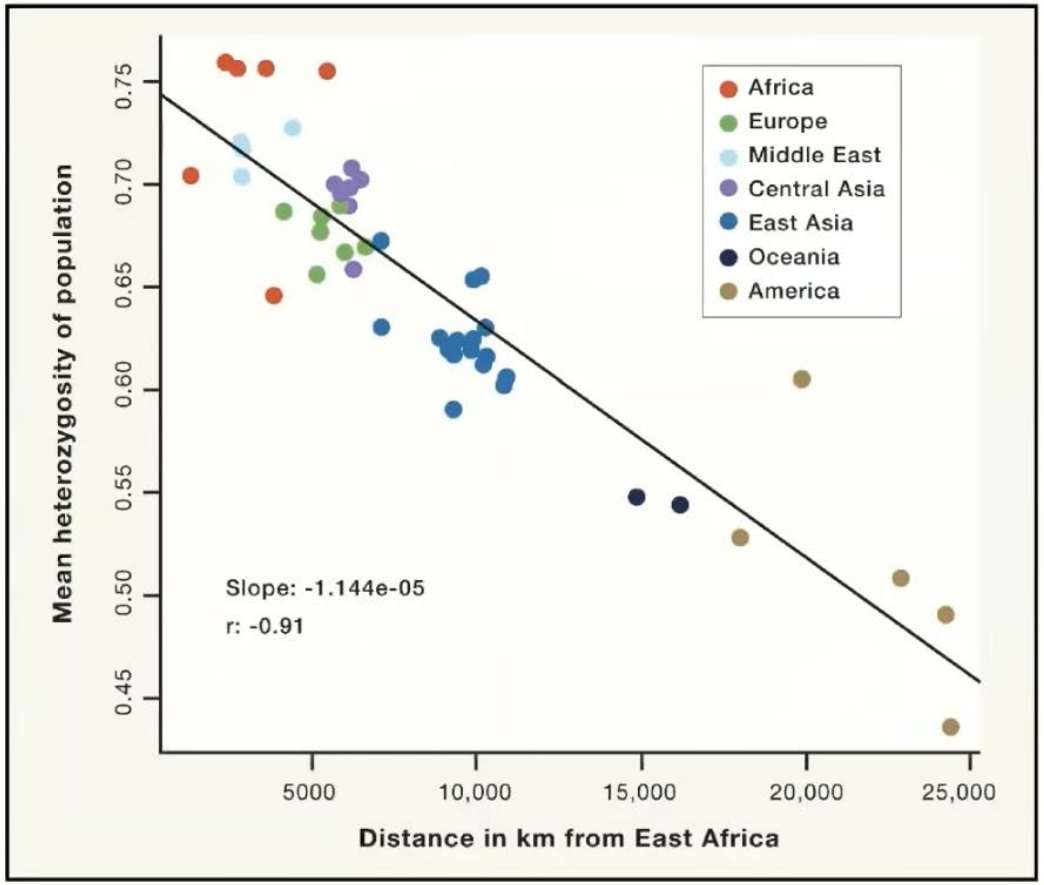
Regular caffeine intake promotes a more efficient ability of the brain to encode experience-related events.

- Mechanisms associated with habitual (chronic) caffeine consumption in the mouse hippocampus: results revealed that chronic caffeine exerts concerted effects in the hippocampus, at the epigenomic, proteomic and metabolomic levels.
- Caffeine induces neuronal-specific epigenetic changes at synaptic transmission/plasticity-related genes and increased experience-driven transcriptional activity.
- Regular caffeine intake improves the signal-to-noise ratio during information encoding, in part through a fine-tuning of metabolic genes while boosting the salience of information processing during learning in neuronal circuits. Enhances synaptic plasticity activation

Not just one wave of MHs OoA at 60 Ka

- C. Stringer:
- Multiple instances of nonviable pre-60 Ka OoA MH excursions
- Evidence in Europe of multiple MH migrations:
  - One before 50 Ka that disappeared
  - One circa 45 Ka that was supplanted
  - Final one circa 40 Ka that was permanent

We are all Africans

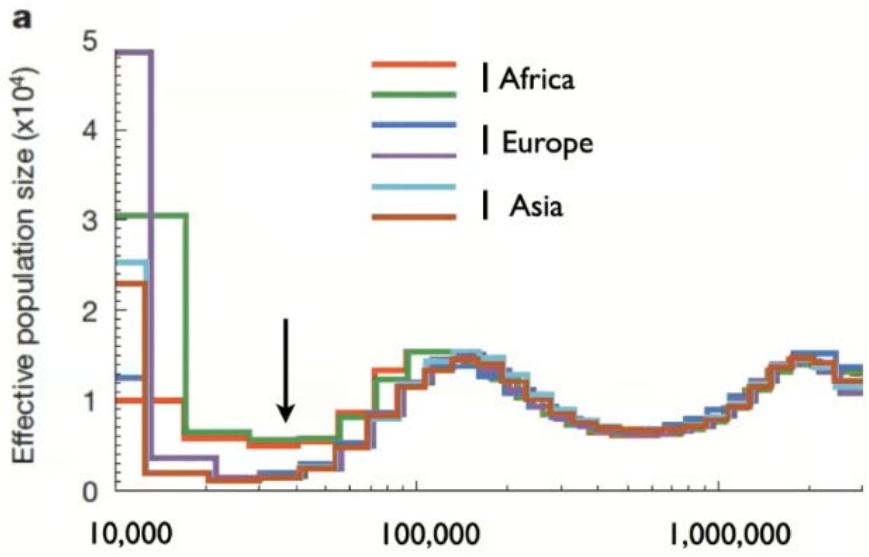


- Heterozygosity (= genetic variation) decreases with distance from East Africa (NB this is walking distance)
- Richest genetic variation found in ancestral areas

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

The bottleneck: 12 K population size at 10-60 Ka

We were very lucky



Six human genomes, measuring effective population size

Bottleneck of around 12,000 people (arrow)

Caused by climate change?

Li & Durbin (2011), Nature

Based on n =12 MH genomes

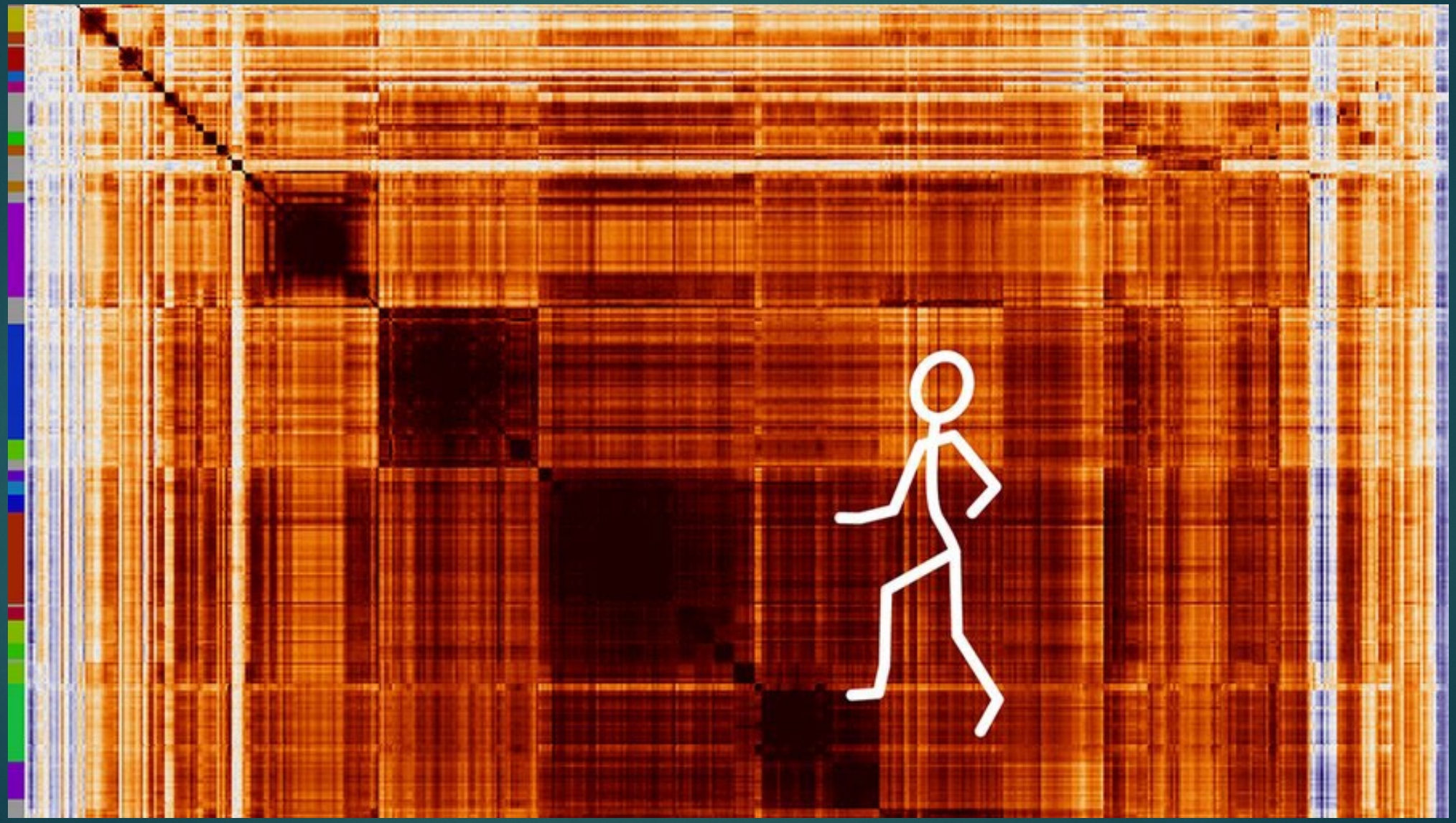
How many ancient individuals produced the variability you now see in these MH genomes

Effective population size (breeding pairs) at 10-60 Ka across Africa, reduced genetic variability

MHs could have crashed and burned

No evolutionary preferential destiny for us = we were lucky

New CRISPR-based map ties every human gene to its function



New CRISPR-based map ties every human gene to its function

- New method for id function of every human gene
- The Perturb-seq method uses CRISPR-Cas9 genome editing to introduce genetic changes into cells, and then uses single-cell RNA sequencing to capture information about the RNAs that are expressed resulting from a given genetic change. Because RNAs control all aspects of how cells behave, this method can help decode the many cellular effects of genetic changes.

Black Death plague in Kyrgyzstan

An engraved tombstone of a person who died from the Black Death plague, from the Kara-Djigach cemetery in what is now Kyrgyzstan. Credit: P.-G. Borbone, M. A. Spyrou *et al./Nature*

Origins of Black Plague in Kyrgyzstan

- People who died in a fourteenth-century outbreak in what is now Kyrgyzstan were killed by strains of the plague-causing bacterium *Yersinia pestis* that gave rise to the pathogens responsible several years later for the Black Death. The region is on the ancient Silk Road trade route.

- This was the initial wave of a nearly 500-year-long pandemic termed the second plague pandemic and is one of the largest infectious disease catastrophes in human history. Between 1346 and 1353, the Black Death laid waste to western Eurasia over an 8-year course, killing up to 60% of the populace in some places.
- Historical records suggest that the bubonic plague emerged from the east: Caffa, on the Crimean peninsula, experienced one of the earliest-recorded outbreaks of plague during a 1346 siege by the army of the Mongol Empire.

Origins of Black Plague in Kyrgyzstan

- China hosts some of the world's greatest genetic diversity of modern *Y. pestis* strains, hinting at an East Asian origin for the Black Death
- Sequenced ancient DNA from seven people whose remains were recovered, discovering *Y. pestis* DNA in three burials from Kara-Djigach.
- A pair of full *Y. pestis* genomes gleaned from the data showed that the bacteria were direct ancestors of strains linked to the Black Death, including a *Y. pestis* sample from a person who died in London. The Kara-Djigach strain was also an ancestor of the vast majority of *Y. pestis* lineages around today

Origins of Black Plague in Kyrgyzstan

- Rodents are the natural reservoir for *Y. pestis*, and humans develop bubonic plague only when a vector such as a flea passes on the infection.
- J. Krause suspects that humans' close contact with infected marmots sparked the Kyrgyzstan epidemic, whereas immunologically naive rat populations in Europe fueled the Black Death.

E. coli

Escherichia coli cells used in the long-term evolution experiment on an agar plate. Credit: Jeffrey Barrick

Legendary bacterial long-term evolution experiment

- A laboratory has been growing 12 populations of *E. coli* since 1988: long-term evolution experiment
- On 24 February 1988, evolutionary biologist Richard Lenski filled 12 flasks with sugary growth medium and seeded each with *Escherichia coli* bacteria. For the past 34 years, Lenski, at Michigan State University has nurtured the bacterial cultures, refreshing growth media daily and freezing samples for future study every couple of months.
- During their 75,000 generations of growth, the bacteria have made huge gains in their fitness — how fast they grow relative to other bacteria — and evolved some surprising traits.

E. coli: Mostly ongoing similarity

- Over the years, we've actually seen just striking amounts of similar reproducibility. So although a typical line has improved its relative fitness compared with the ancestor by maybe 70% or 80%, the variance in competitive fitness between most lines is more like just a few per cent. So they've all tremendously increased, but very similarly to one another.
- Seen some quite striking divergences between the lines. Thirty thousand generations into the experiment, one of the 12 lines evolved the ability to consume citrate, instead of just glucose.
- And after 75,000 generations , it's still the only one of the 12 lines that has evolved that ability.

Debate about source of evolution's diversity solved?

- Long debate in biology about relative importance of existing variation vs new mutations in evolution.

- Bacterial experiments normally start with identical bacterial; evolution seen is entirely due to new mutations

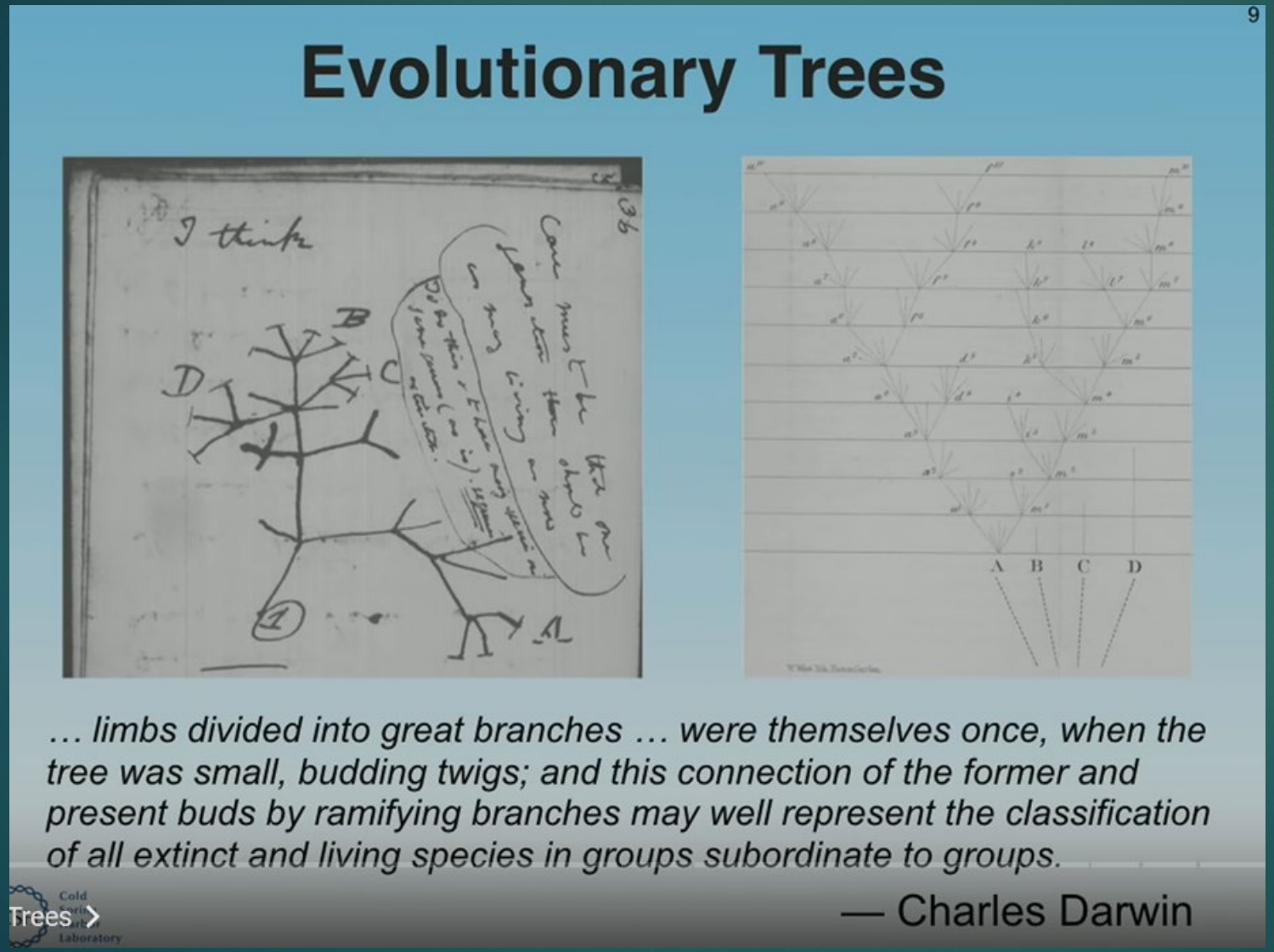
- 2022 experiment (Izutsu & Lenski): 72 populations of E. coli, ranging from genetically identical to highly diverse ones; had been feed on glucose, then fed D-serine to force to adapt

New mutations are drivers of evolution

- The genetic variation that was there to start with really mattered early in experiment, in first 50 to 100 generations; but by 500 generations (out of 2000), starting variation no longer mattered at all.
- New mutations won the debate. After 500 generations, pre-existing variation no longer mattered; all further evolution due to new mutations.
- Convincing evidence that standing variation contributed to very early stages of evolution, but that long term adaptation appears to require new mutations
- Unclear if this applies to sexually reproducing organisms

E. Coli and antimicrobial resistance evolution in large petri dish

Reconstructing ancient history from DNA – Adam Siepel



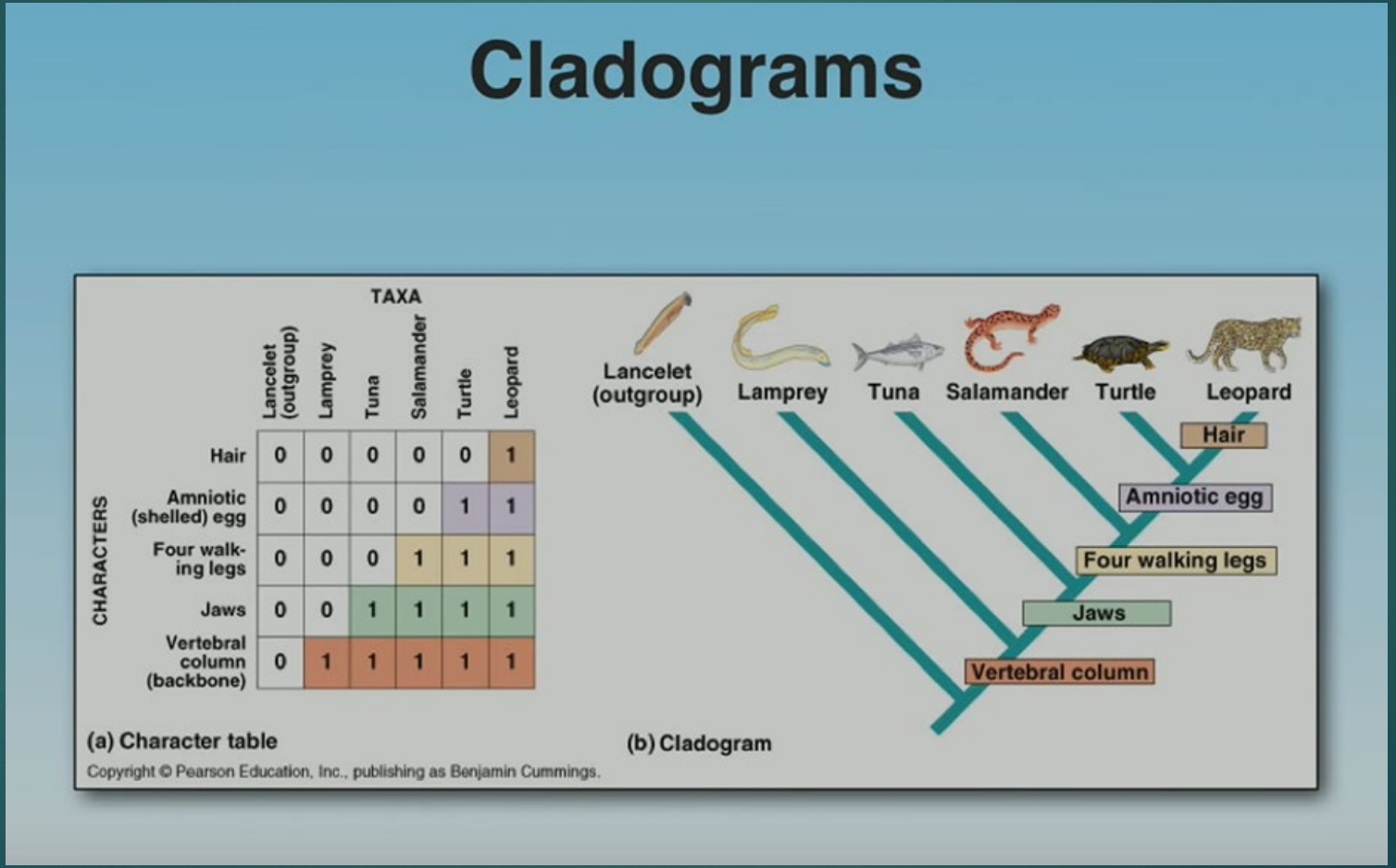
Charles Darwin:

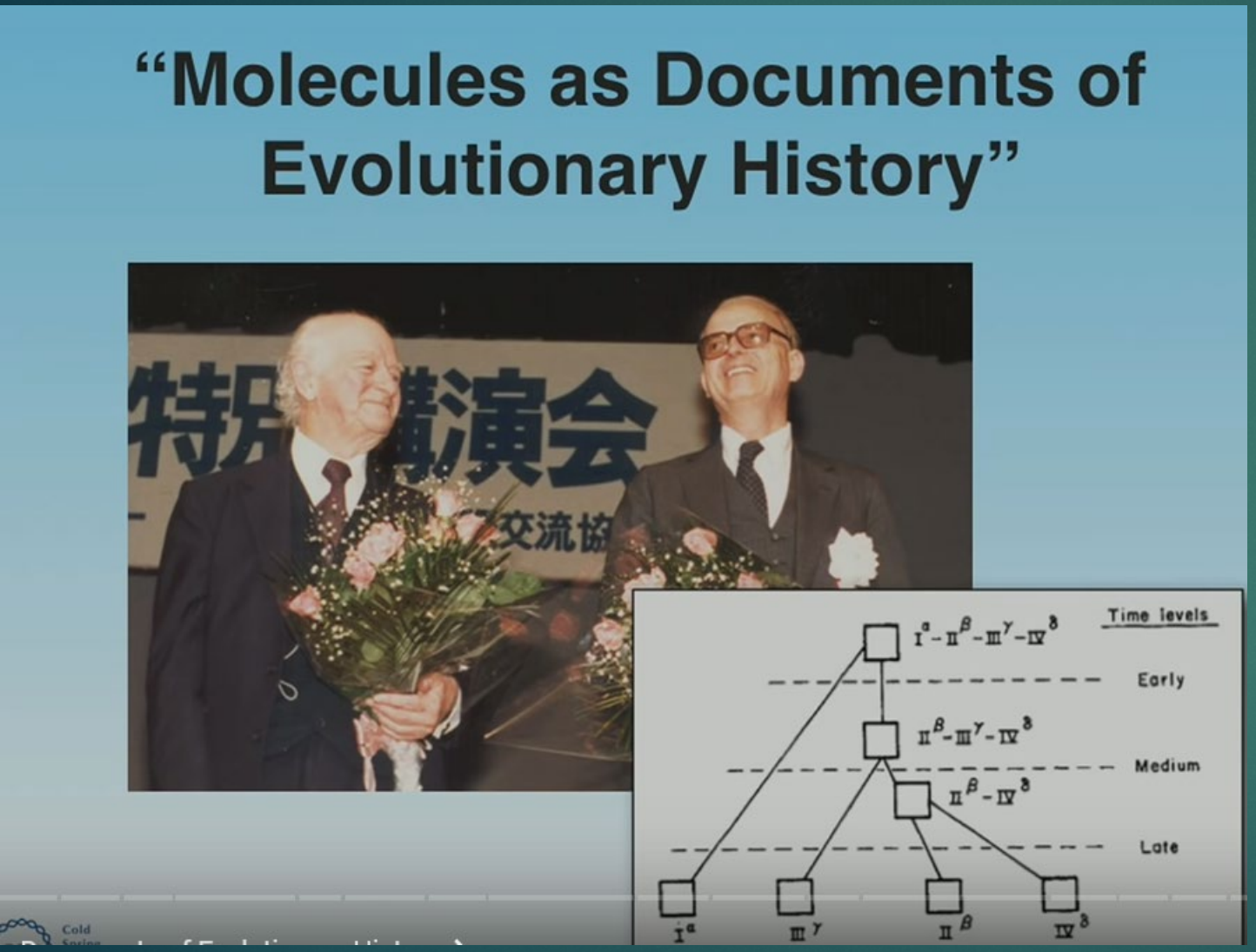
- First evolutionary tree from Notebook B, 1837:
- Branching structure with more primitive at base

- 2nd picture: only figure in 1867 book

- Linnaeus Lamarck had
- Also talked of trees

Cladistics: trees based on observable phenotypic traits: branching relationships based on morphology

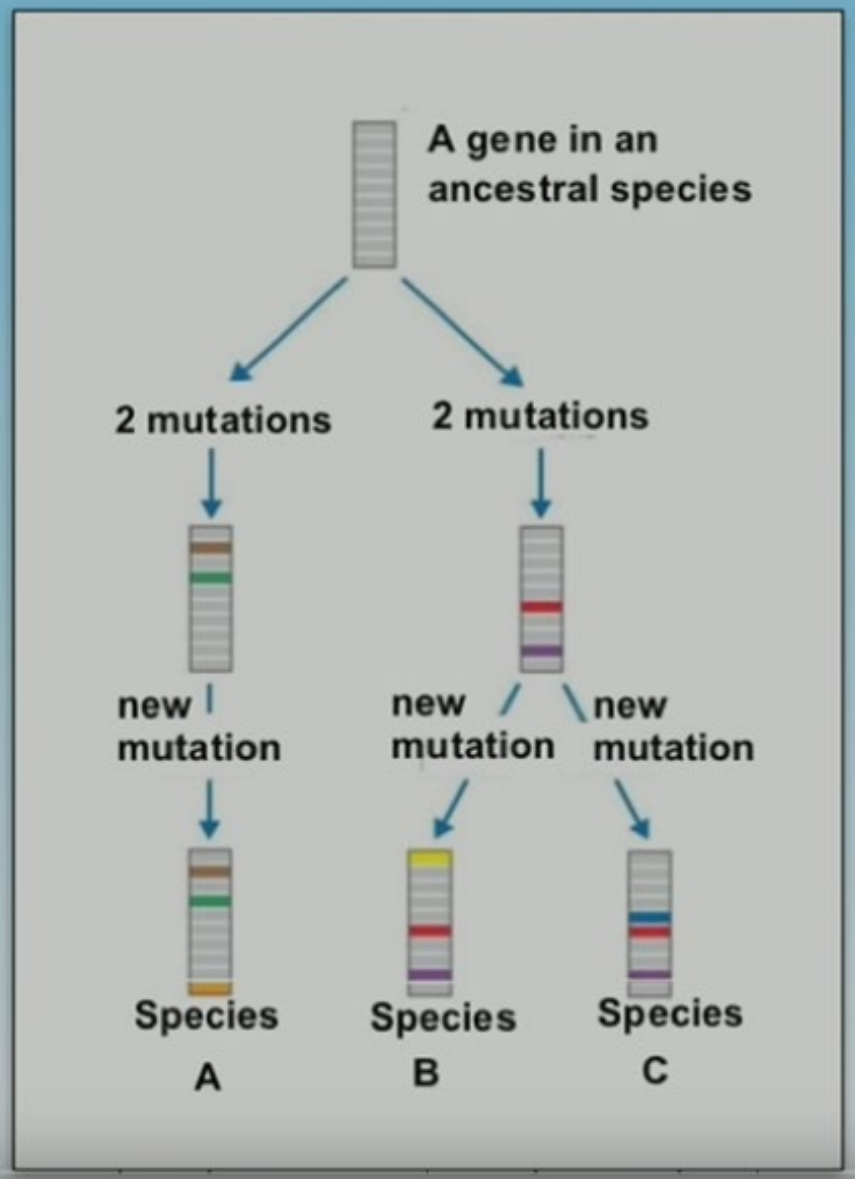




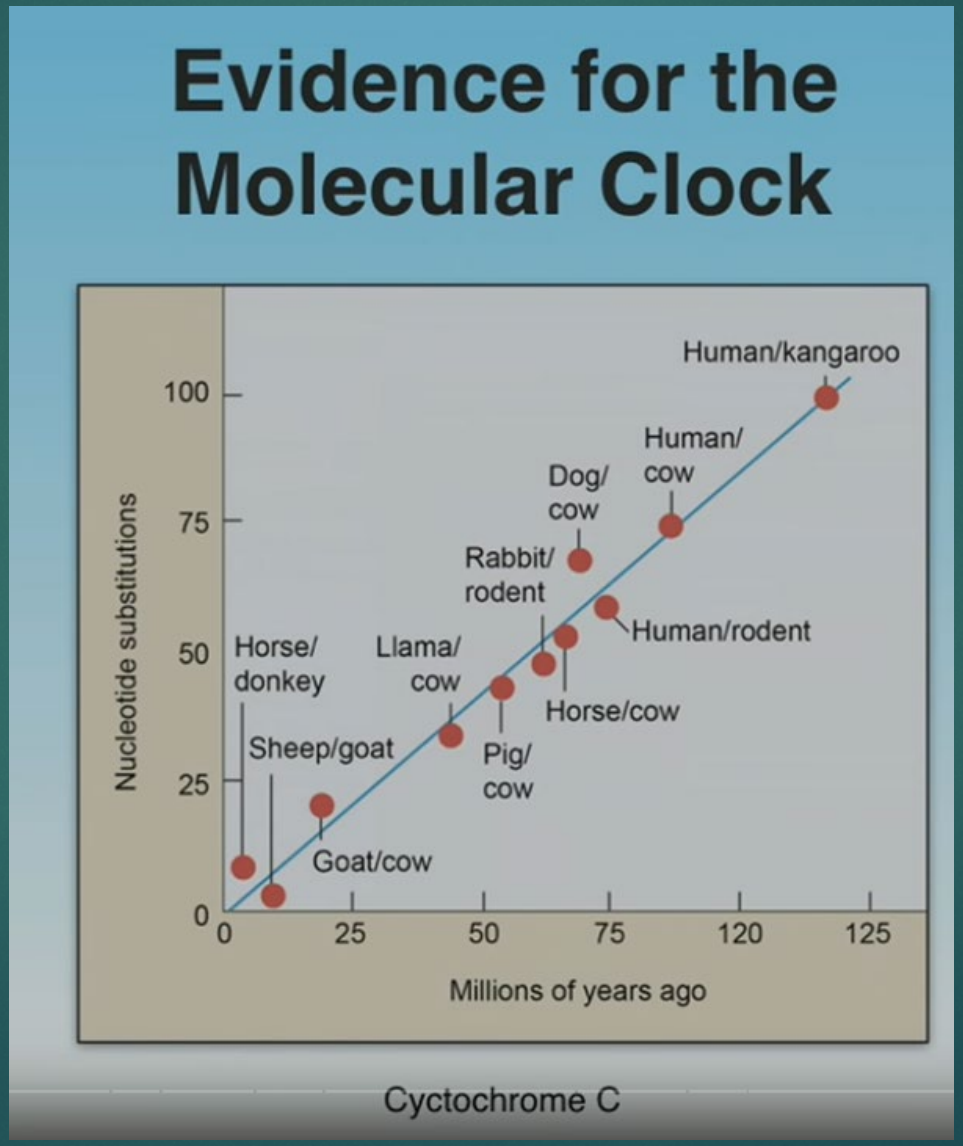
Then Molecules as basis of phylogeny: Pauling and Zuckerkandl

- Hemoglobin: number of molecules as indicator of age – molecular clock – more mutations in amino acid sequences in oldest; more closely related have fewer mutations

Molecular Clock



Date of separation: DNA nucleotide substitutions in separation of species gives you evolutionary distance time

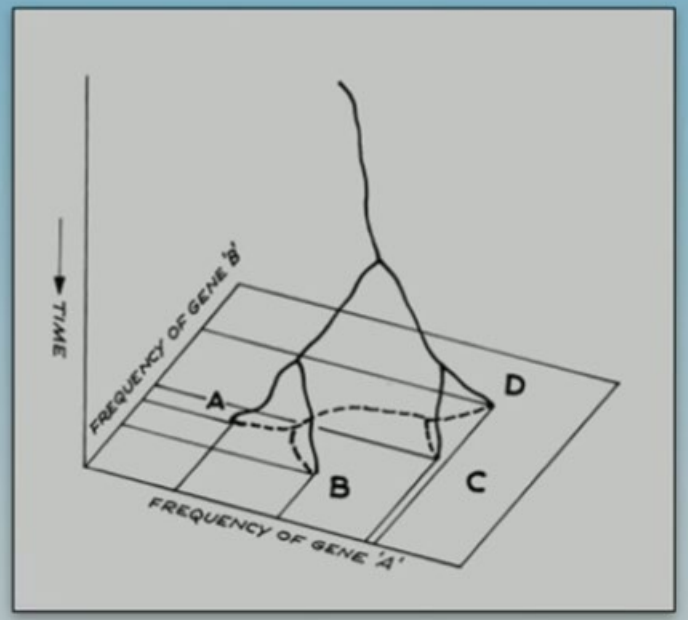


First recipe for how closely species were related and how long ago they diverged

Phylogeny Reconstruction

Luca Cavalli-Sforza

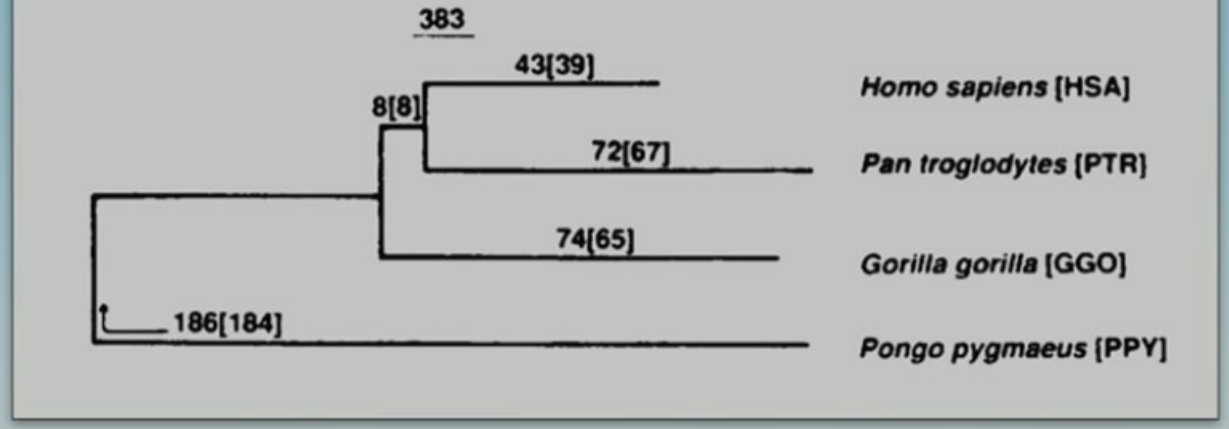
Anthony Edwards



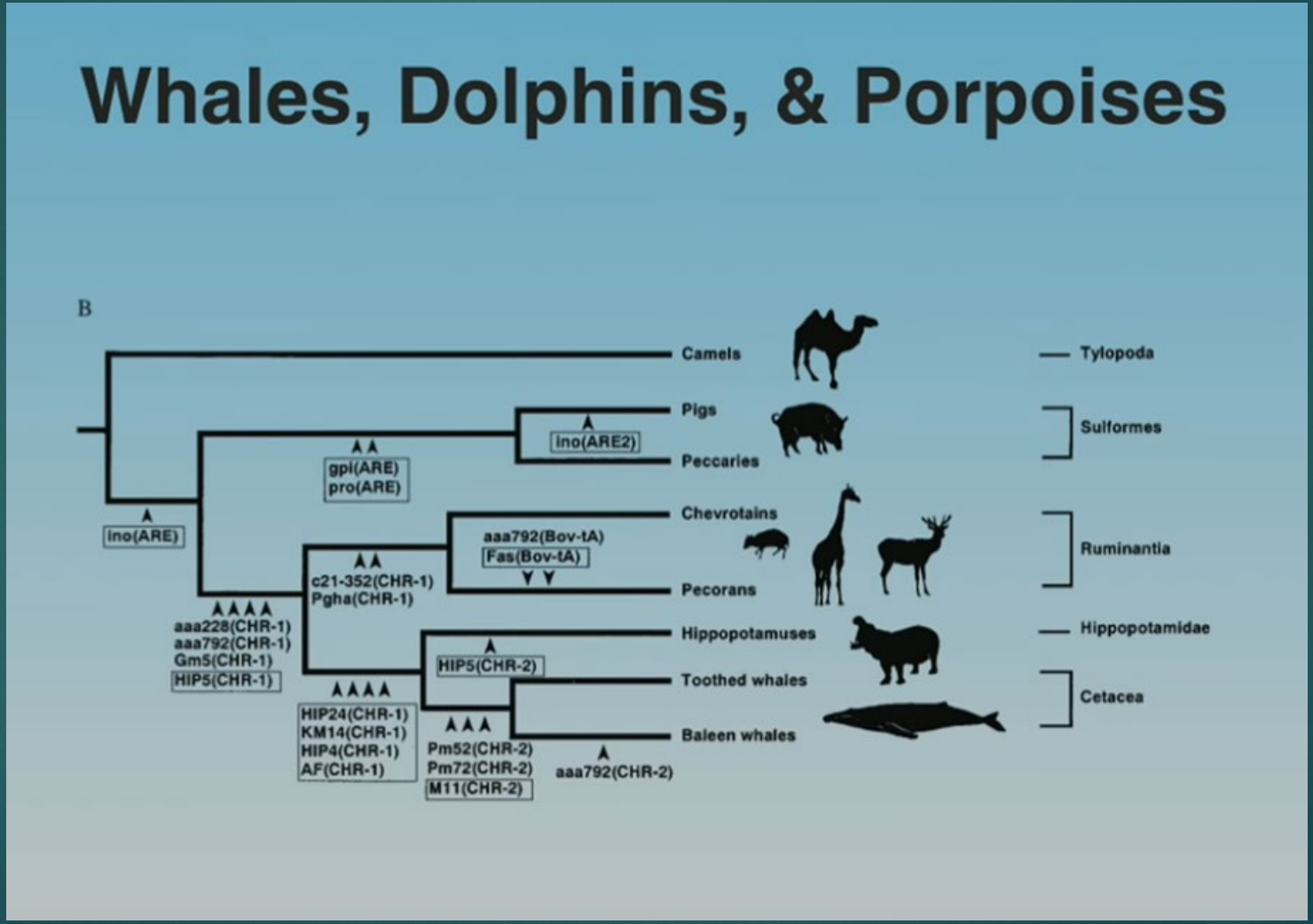
Who is closer: Famous paper showing how Humans are related to other primates, and that chimps are more closely related to us than gorillas

Phylogenetic Relations of Humans and African Apes from DNA Sequences in the ψη-Globin Region

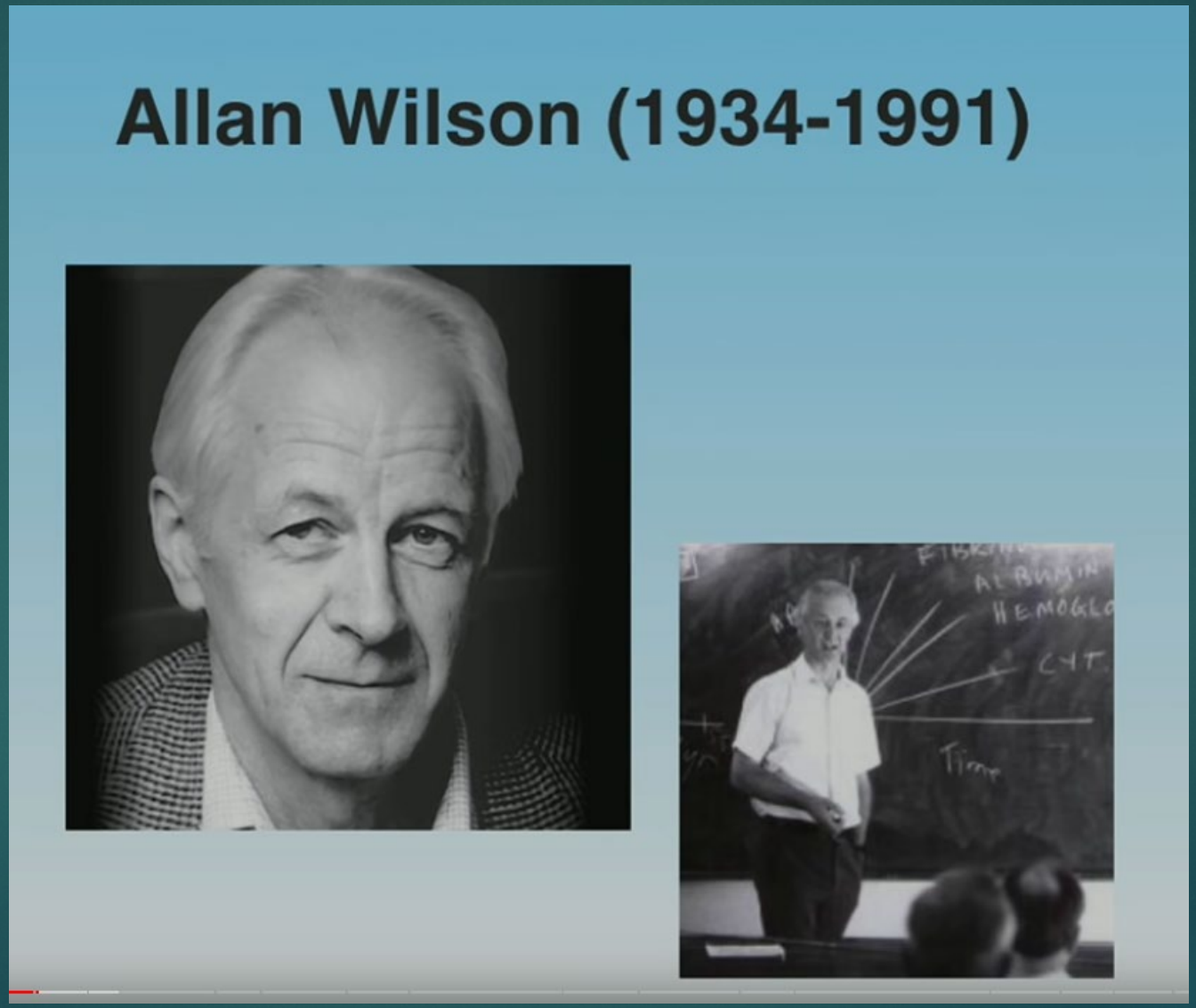
MICHAEL M. MIYAMOTO,*† JERRY L. SLIGHTOM, MORRIS GOODMAN



Cetacea: hippos were closest to whales, circa 50 Ma



Molecular phylogenetics: trained S. Paabo and Mary Claire King (BRCA gene)



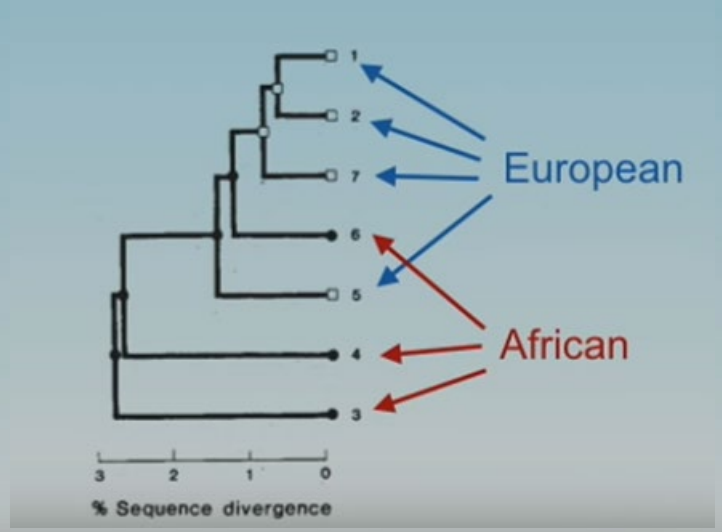
Nature 1987: 1st major study of human evolution based on mtDNA: divergence date of 200 Ka: Mitochondrial Eve

Mitochondrial DNA and human evolution

Rebecca L. Cann*, Mark Stoneking & Allan C. Wilson

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

Mitochondrial DNAs from 147 people, drawn from five geographic populations have been analysed by restriction mapping. All these mitochondrial DNAs stem from one woman who is postulated to have lived about 200,000 years ago, probably in Africa. All the populations examined except the African population have multiple origins, implying that each area was colonised repeatedly.



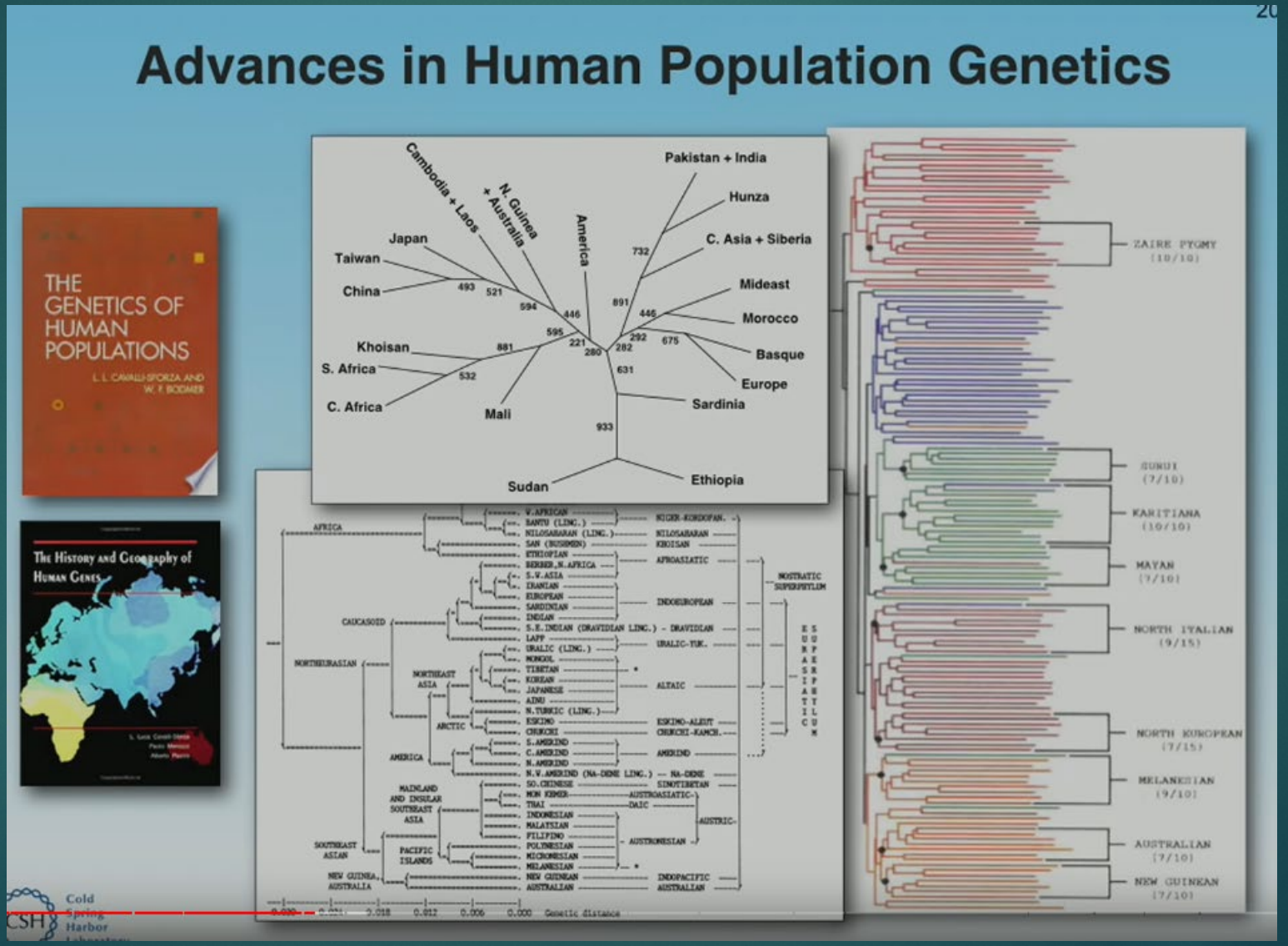
Africans mostly fell out of variation of European groups

Africa was source of human genetic variation; had greater diversity

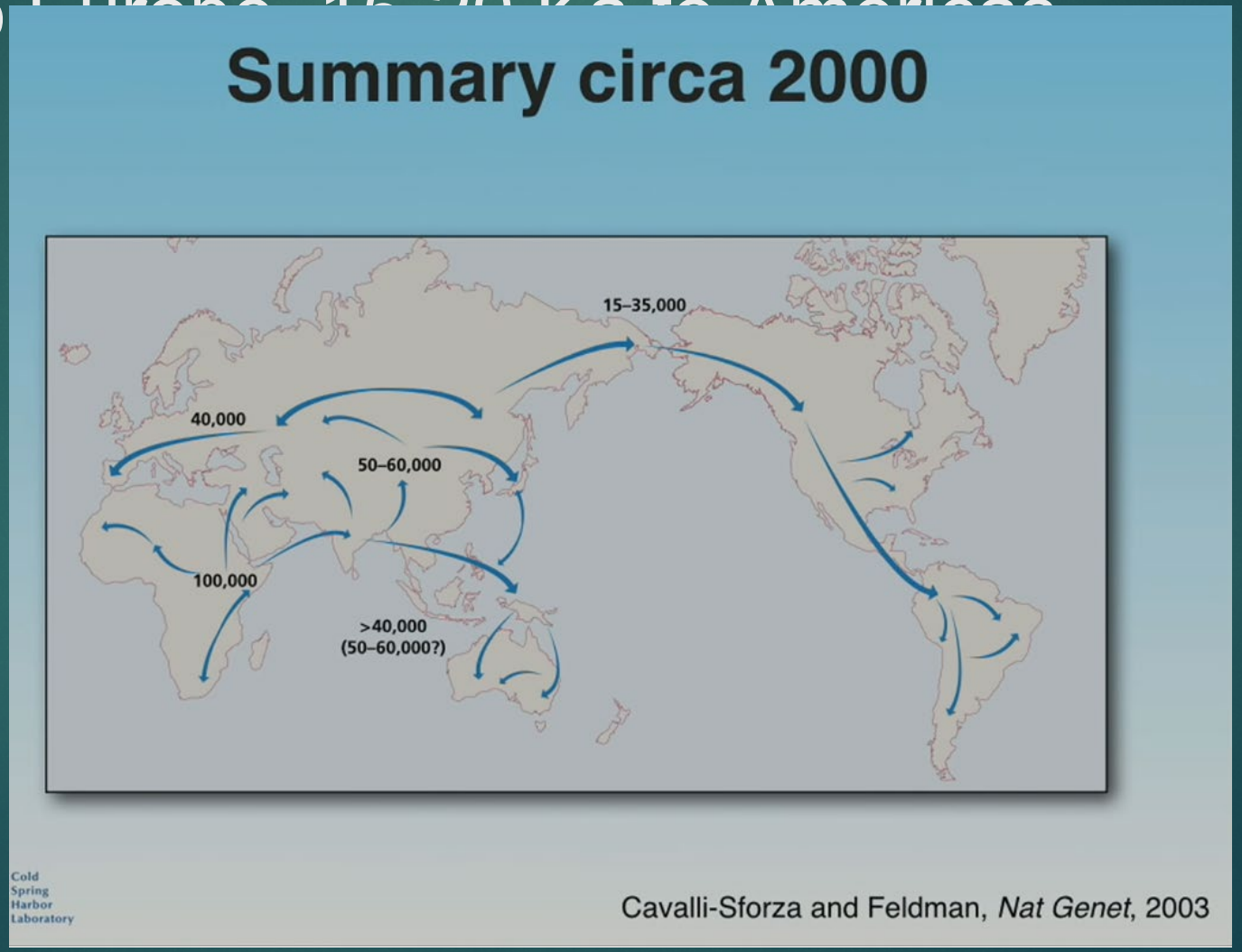
Multiple founder effects in OoA

Luca Cavalli-Sforza: genetics, linguistics, culture

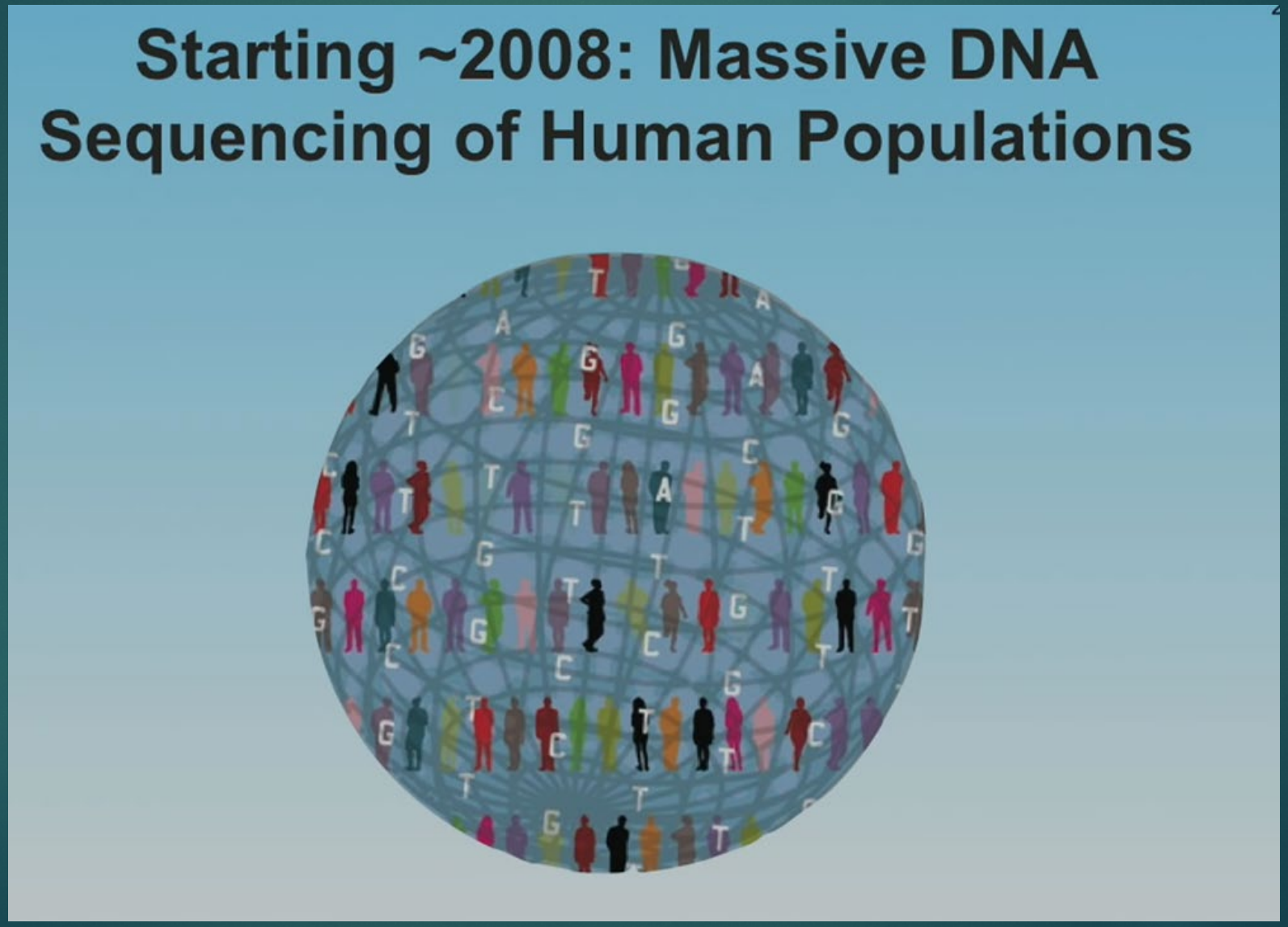
1994



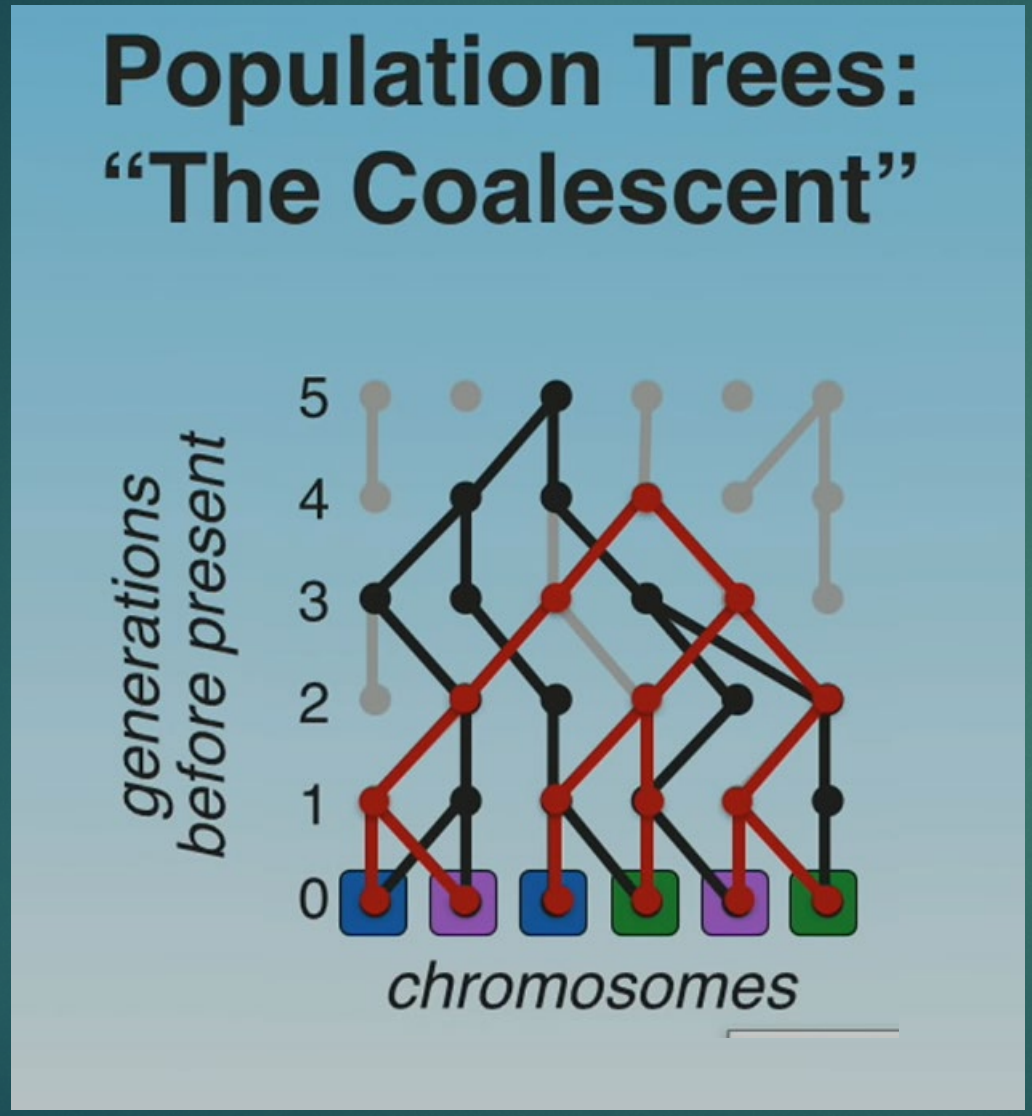
Human Evolution in 2000: 200 Ka in E Africa, 100 Ka in Arica, 60 Ka OoA, 1 Southern and 1 Northern route to East, 1 to West; 60 Ka to Australia & China, 40 Ka to



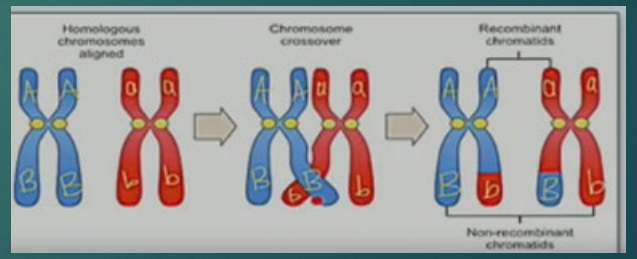
1000 genome study: complete genomes



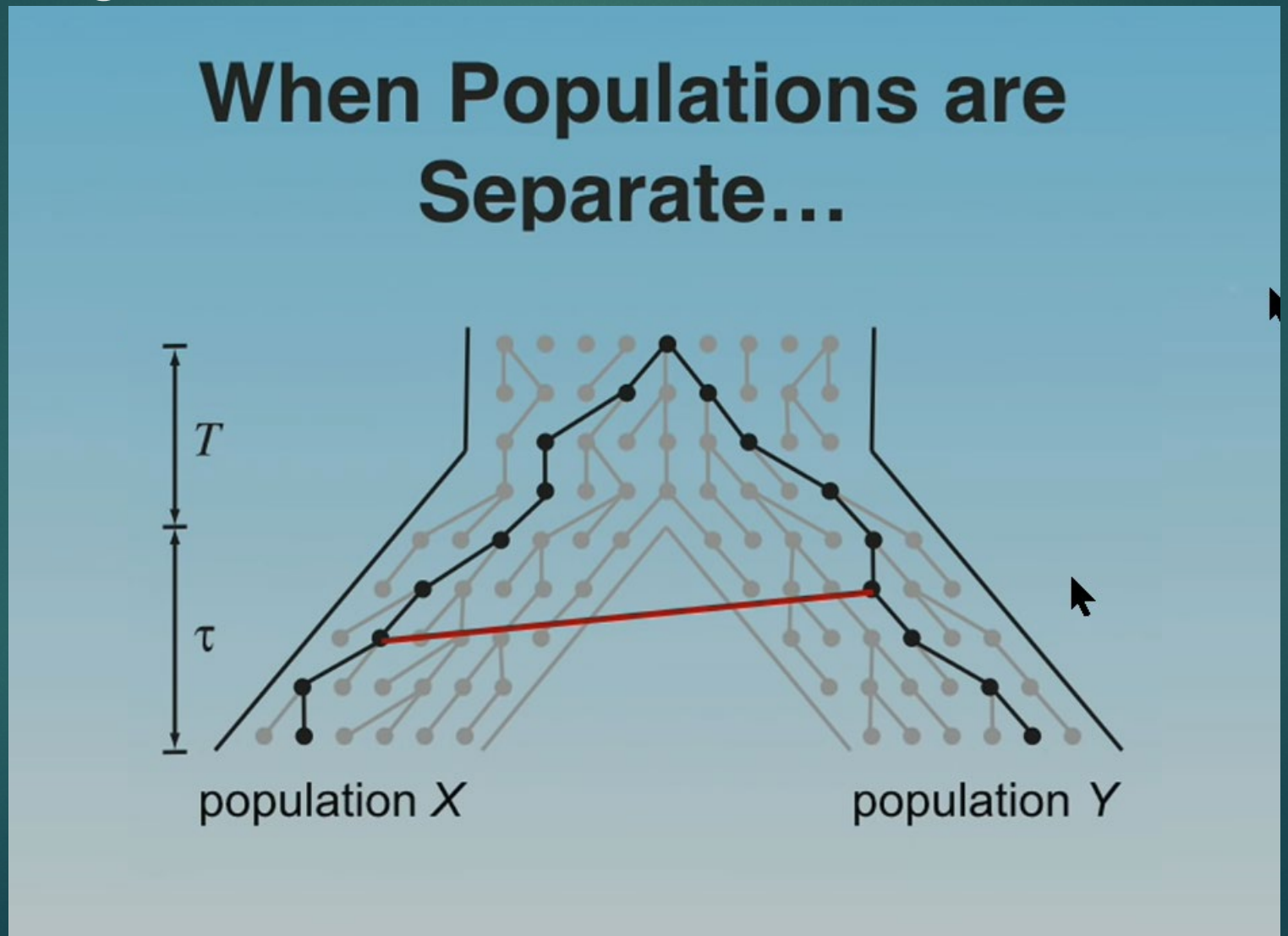
More data, more issues



- Each gene has its own evolutionary history given parental recombinatio n of 2 copies



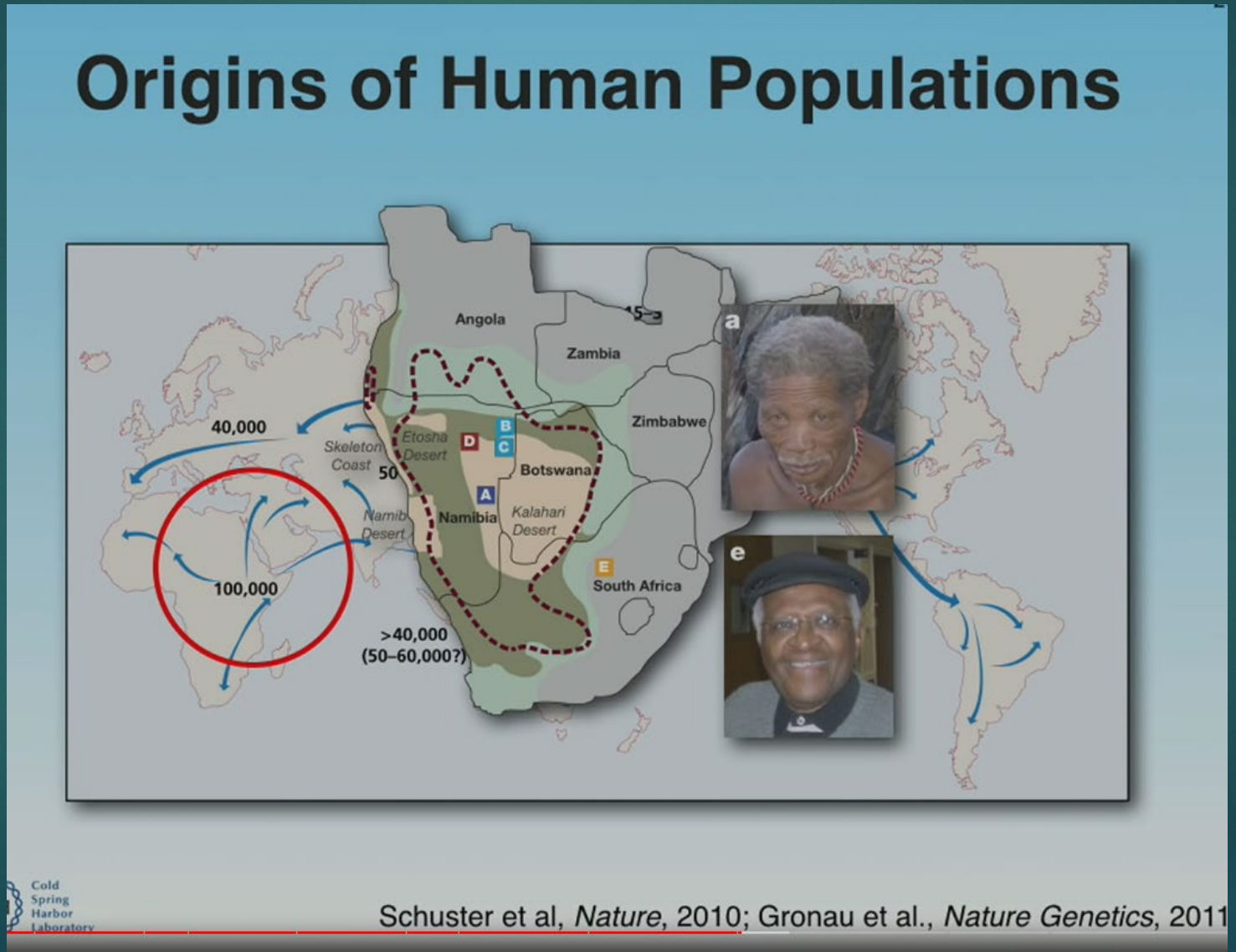
2 populations on 2 islands with common ancestor;
but later gene flow, will give you later ancestry date
= indication of gene flow



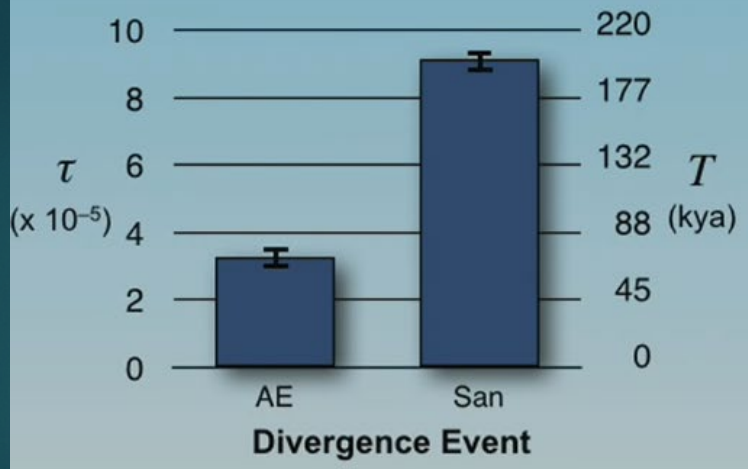
More advanced statistical techniques

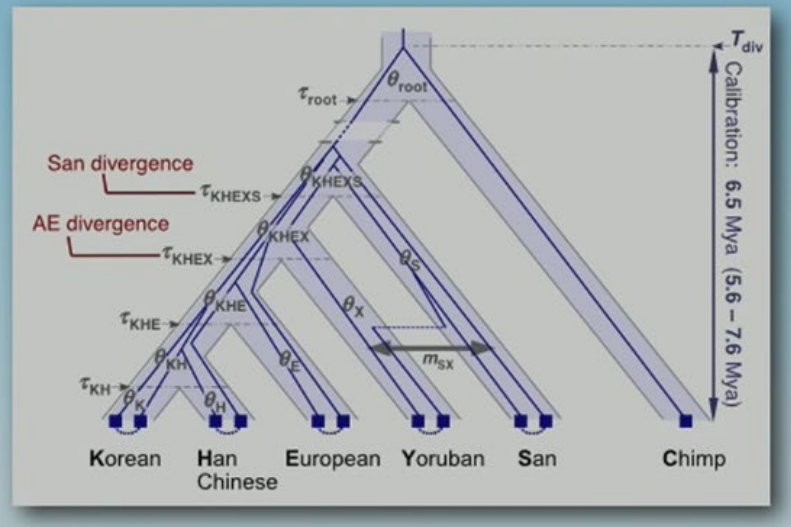
- Now use Mont Carlo statistical techniques to rule millions of comparisons of ancestral gene models to see which best fits the genomic data

Original Cavalli-Sforza data indicated San as earliest African branching group based on mt and y DNA



Divergence Time of the San





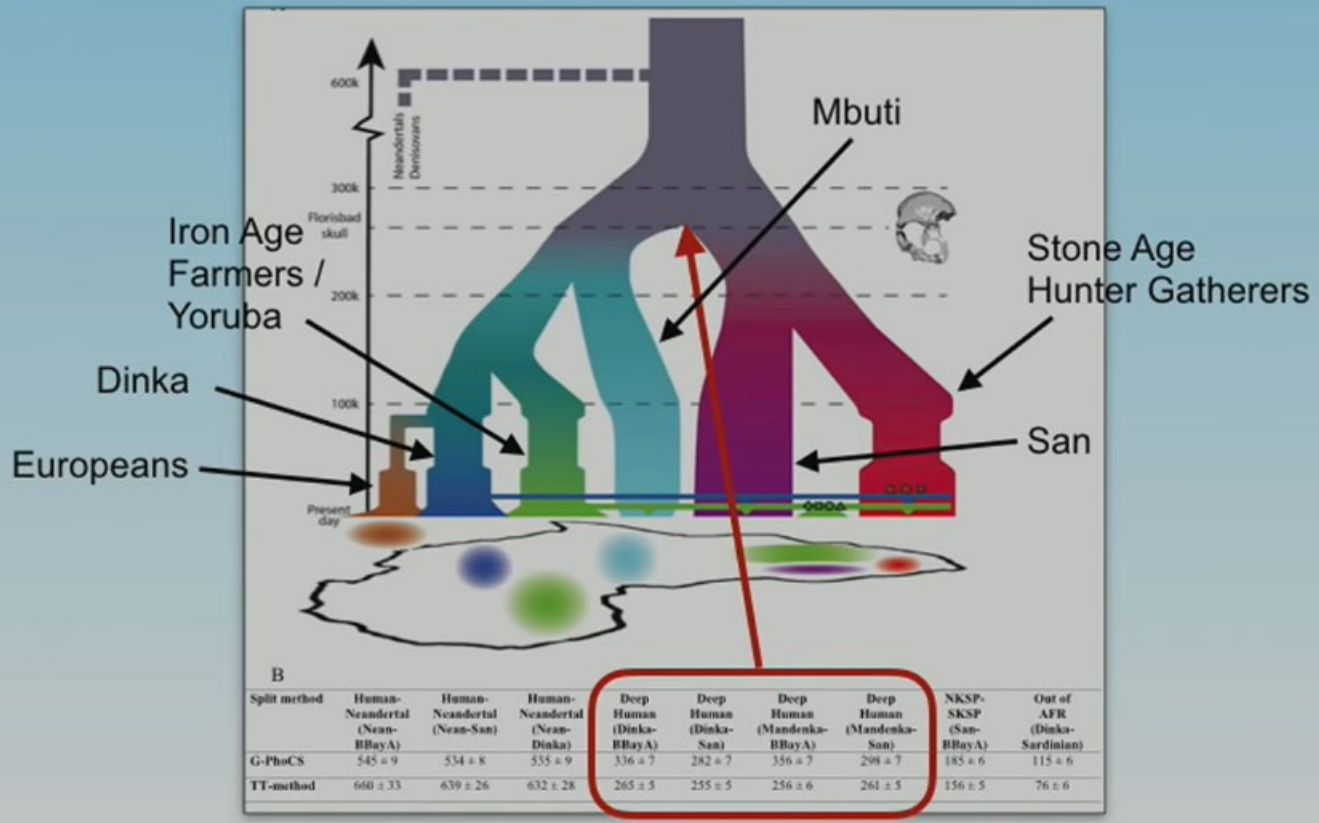
Gronau et al., *Nature Genetics*, 2011

- San diverged ~200 Ka; same time as mtEve
- AE (African-Eurasian, OofA) 70-80 Ka

Schlebusch, 2017: 260 Ka split of San

Ancient genomes from southern Africa pushes modern human divergence beyond 260,000 years ago

Carina M. Schlebusch[1,4]*, Helena Malmström[1,4]*, Torsten Günther[1], Per Sjödin[1], Alexandra Coutinho[1], Hanna Edlund[1], Arielle R. Munters[1], Maryna Steyn[2], Himla Soodyall[3], Marlize Lombard[4,5#], Mattias Jakobsson[1,4,6#]

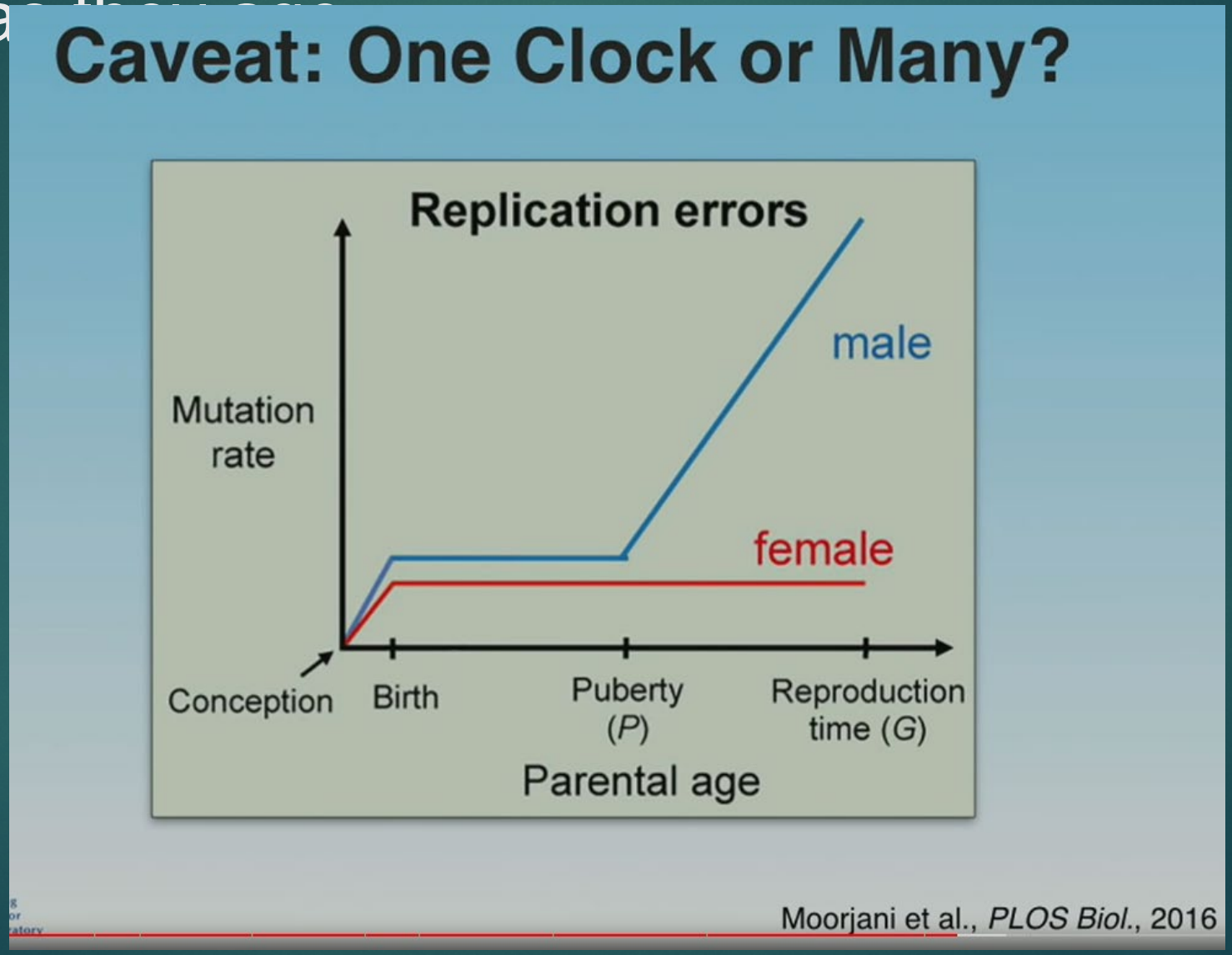


B

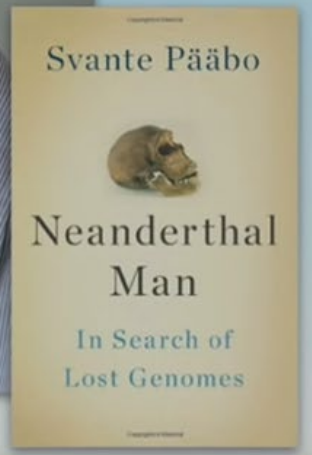
Split method	Human-Neandertal (Nean-BBayA)	Human-Neandertal (Nean-San)	Human-Neandertal (Nean-Dinka)	Deep Human (Dinka-BBayA)	Deep Human (Dinka-San)	Deep Human (Mandenka-BBayA)	Deep Human (Mandenka-San)	NKSP-SKSP (San-BBayA)	Out of AFR (Dinka-Sardinian)
G-PhoCS	545 ± 9	534 ± 8	535 ± 9	336 ± 7	282 ± 7	356 ± 7	298 ± 7	185 ± 6	115 ± 6
TT-method	660 ± 33	639 ± 26	632 ± 28	265 ± 5	255 ± 5	256 ± 6	261 ± 5	156 ± 5	76 ± 6

bioRxiv, 2017

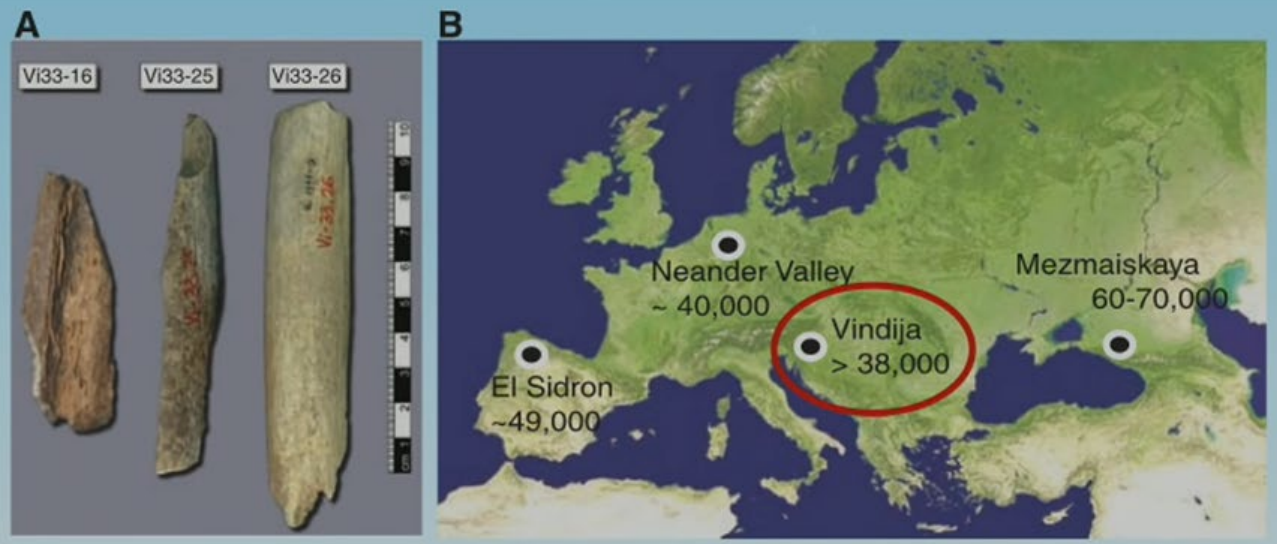
There are many molecular clocks: 1 example is differences by sex: males have more mutations in sperm cells a



Neanderthal Man: Svante Pääbo



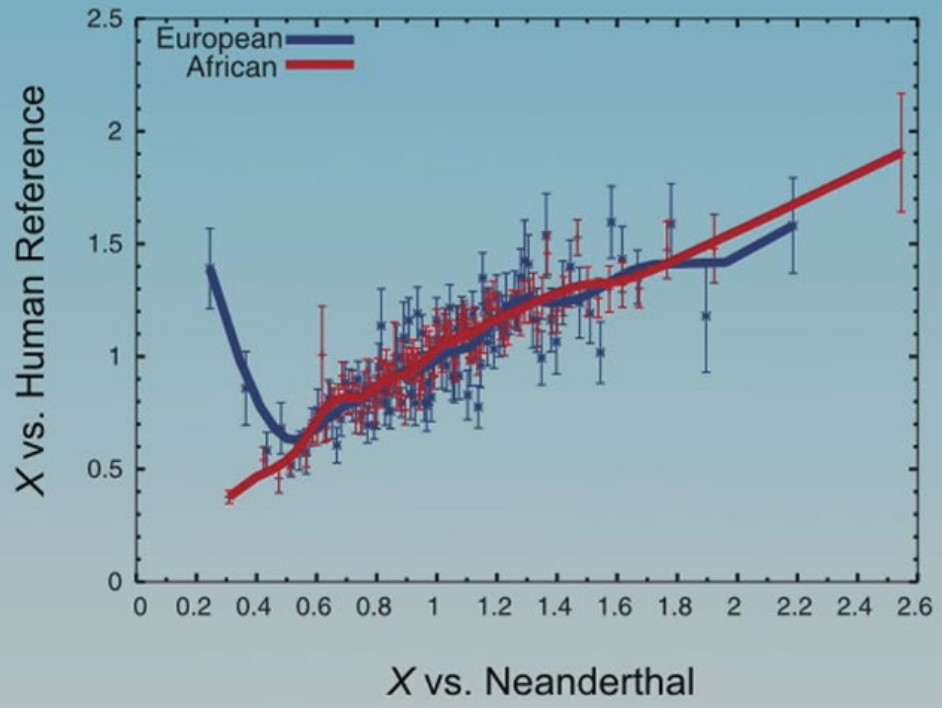
“Draft” Neanderthal Genome



Green et al, *Science*, 2010

Draft N genome by combining DNA from 3 bones from 1 cave, not from 1 individual

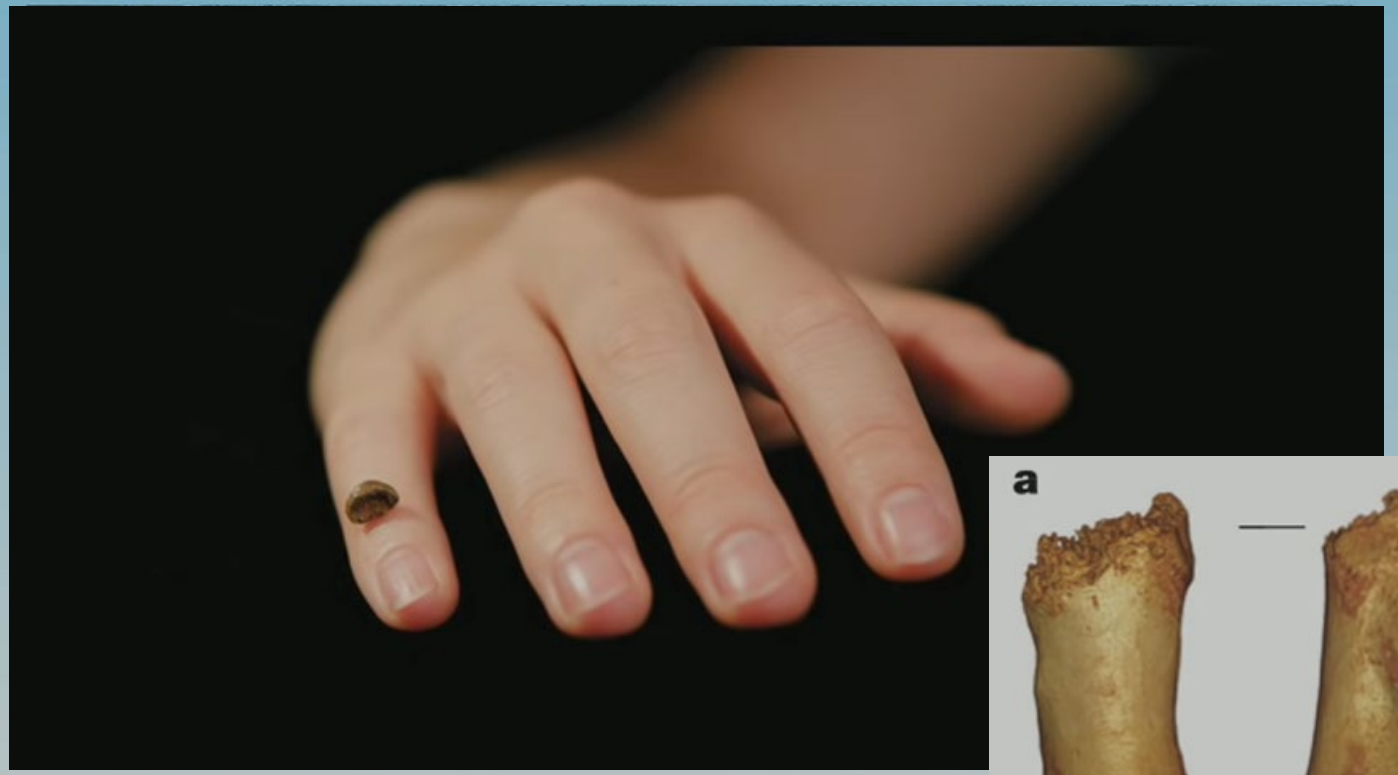
Evidence for Interbreeding

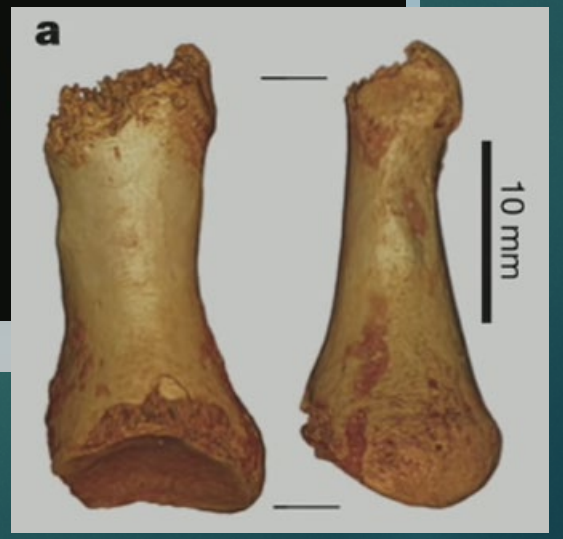


Green et al. *Science*. 2010

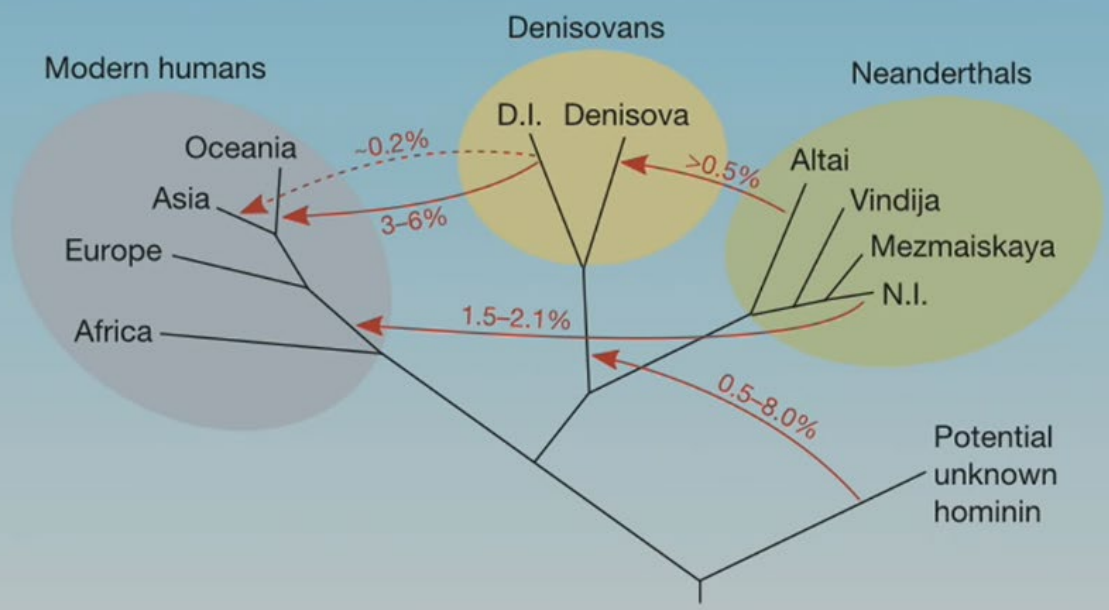
Strong evidence for interbreeding – MHs interbred with Ns after OoA

Denisova Cave



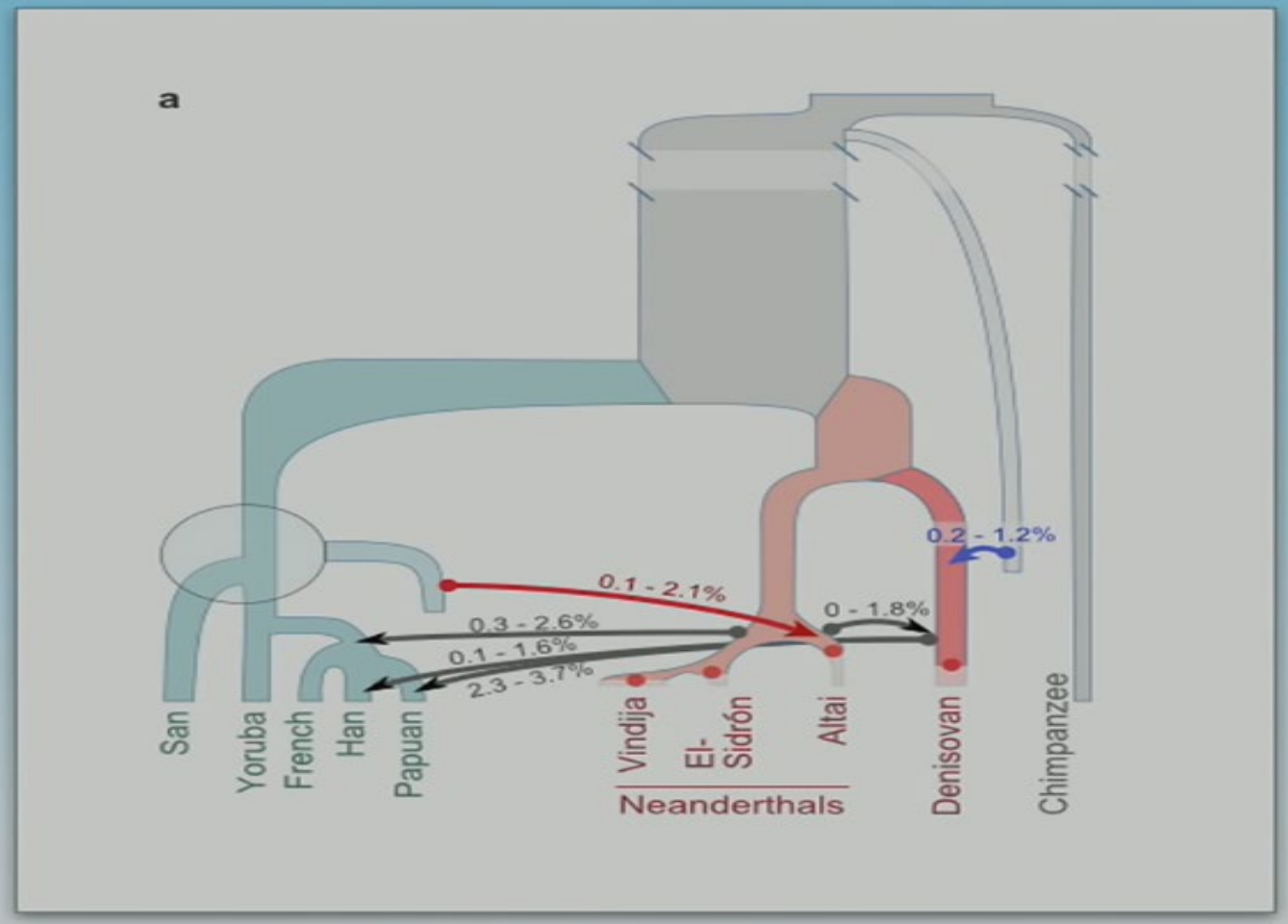


Relationships Among Archaic and Modern Hominins



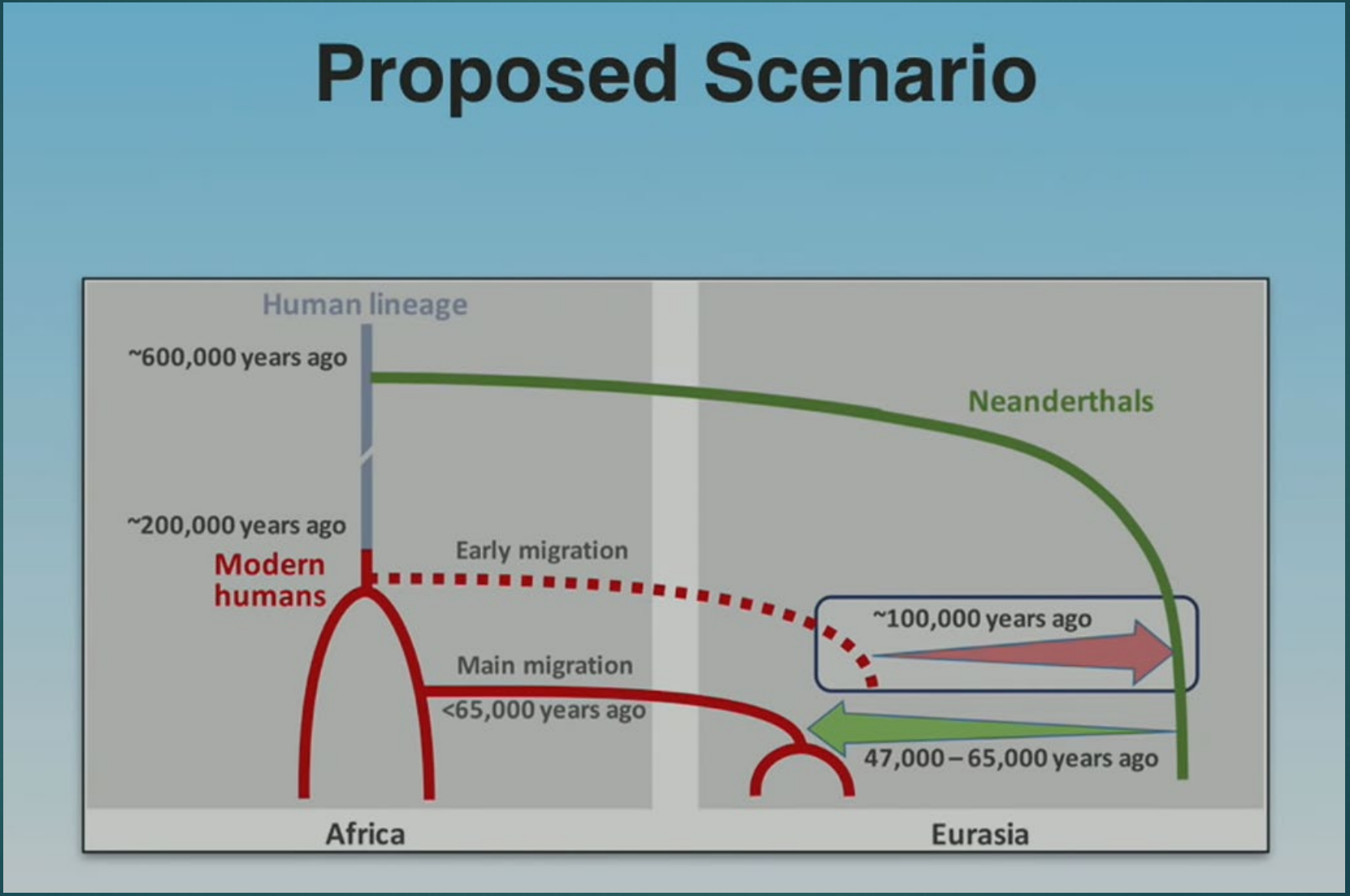
Prufer et al, *Nature*, 2013

G-PhoCS Results



Relook at prior studies: in addition saw MH introgression (red) into Ns before OoA

MH introgression into Ns circa 100 Ka



C. Stringer on species

- 35 current theories of what a species is
- 20% of bird and baboon species hybridize
- But N and MHs were different species based on morphology:
  - No MH has wide N pelvis
  - No MH has inner ear bones shaped like N's
  - No MH has large hollow browridges of N
  - Larger N orbits imply larger Occipital visual regions

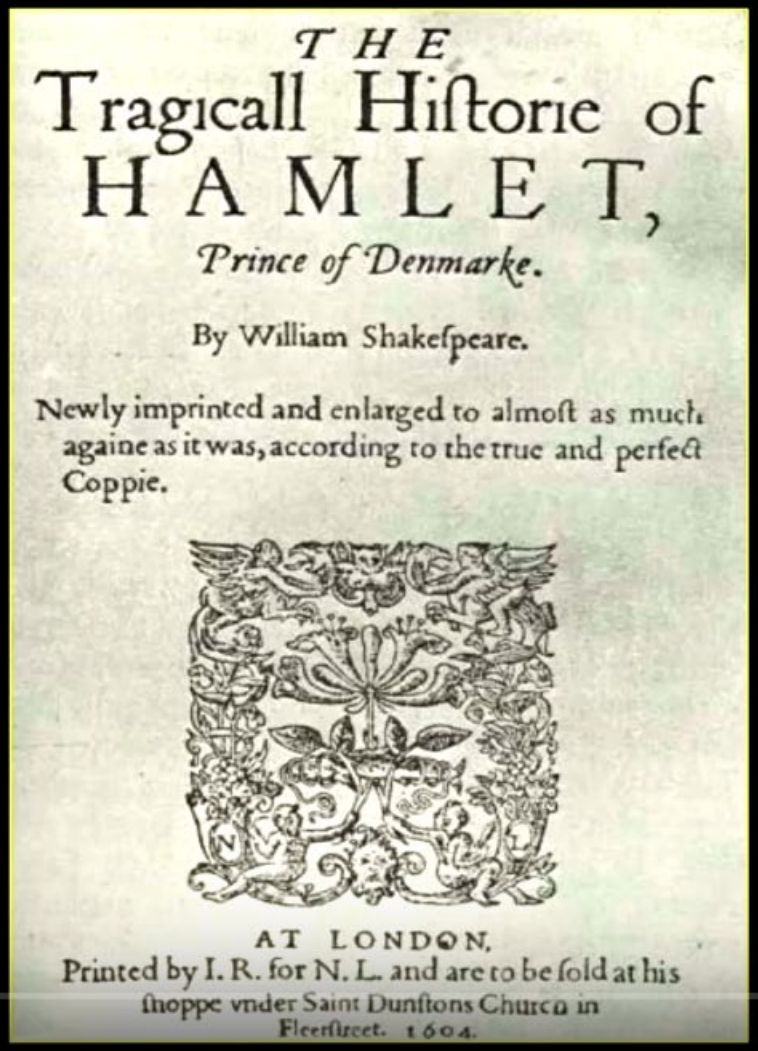
- Stringer believes that you need a certain population size for cultural traditions to take hold

Paleogenetics, Part 8:

What made Modern Humans Genetically Unique

** Novel MH genes

MH: derived and ancestral



THE
Tragicall Hiſtorie of
HAMLET,
Prince of Denmarke.

By William Shakeſpeare.

Newly imprinted and enlarged to almoſt as much againe as it was, according to the true and perfect Coppie.

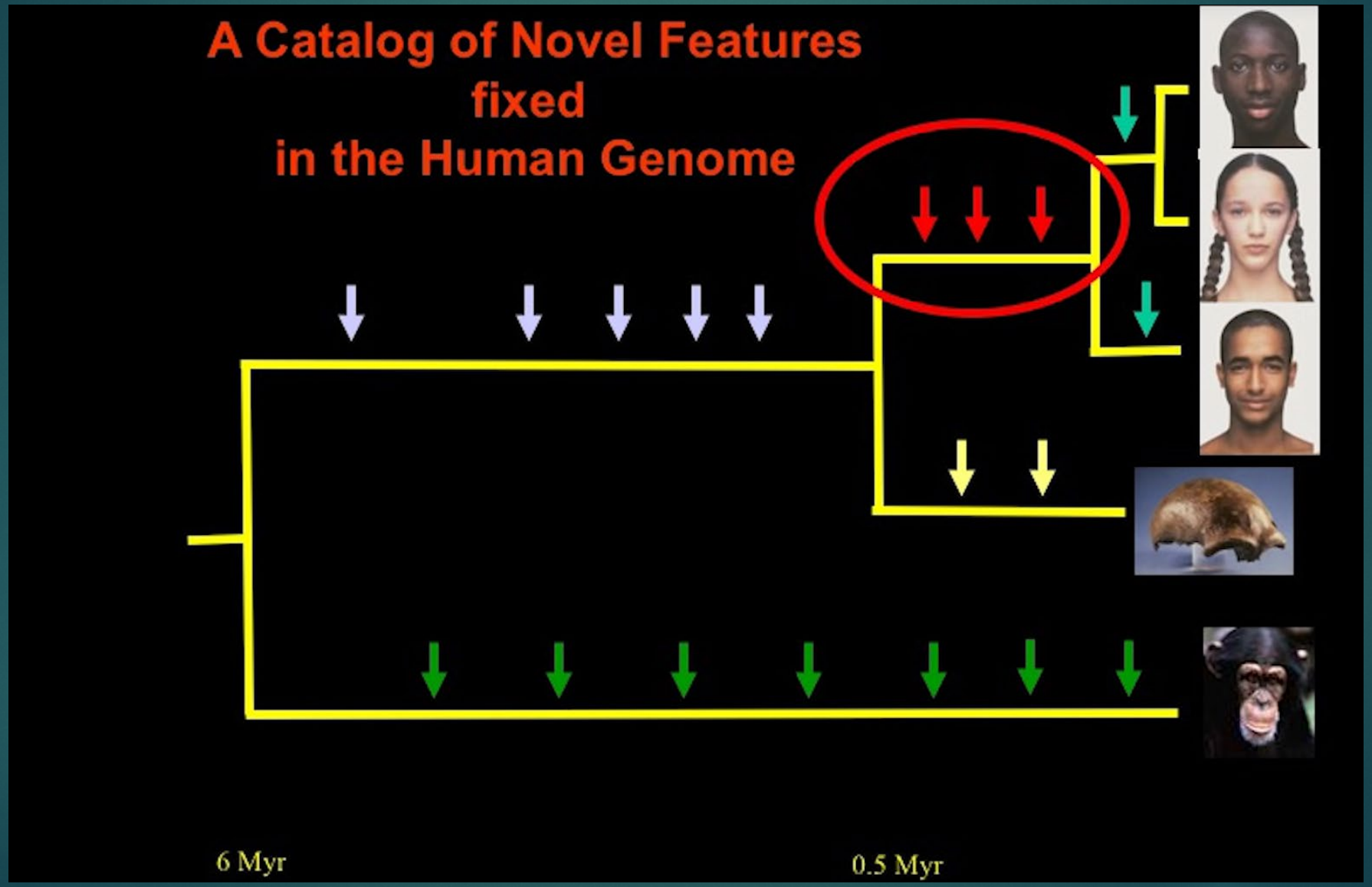
AT LONDON,
Printed by I. R. for N. L. and are to be ſold at his ſhoppe vnder Saint Dunſtons Church in Fleetſtreet. 1604.

'What a piece of work is a man!
How noble in reason!
How infinite in faculty!

In action how like an angel!
In apprehension how like a god!

The paragon of animals!'

A Catalog of Novel Features fixed in the Human Genome



Genetic Difference between humans and chimpanzee

What makes us human, genetically, compared to chimps?

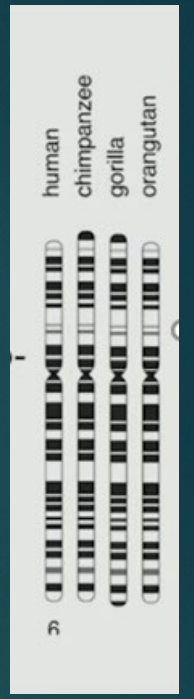
- -35 million single-nucleotide/basepair changes
- ~5 million insertion/deletions
- 9 pericentric inversions
- 1 chromosome fusion (compared to chimps): human chromosome 2 is a fusion of two great ape chromosomes.

Human/chimp genetic differences

Few large-scale chromosomal differences—one chromosomal fusion and 9 pericentric inversions

Genetically nearly identical (98.7%) or 1.27% single-base-pair substitution differences

Proteins >99% identical (30% of genes don't have a single amino- acid difference between chimpanzee and human)



MH mutations:

Genetic recipe for a modern human vs Ns & Ds

- To date, Pääbo has assembled a catalog of about 31,000 base-pair changes, or single nucleotide polymorphisms (SNPs), in which modern humans carry a different version from Neandertals and Denisovans.

- 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."

Pääbo 2014: Catalog of novel MH genetic features fixed in human genome

- 31,389 single nucleotide (basepair) changes
- 125 insertions and deletions
- 45 splice sites (at the boundary of an exon and an intron, which are part of RNA transcription)
- 3,117 regulatory regions
- 96 amino acids

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.
  - only 96 fixed amino acid substitutions in a total of 87 proteins
  - 3000 fixed changes that potentially influence gene expression in present-day humans;
  - 5 genes effect neural stem cells in the adult subventricular zone.

87 genes separate MHs from Ns and Ds

- At least 87 genes found only in modern humans that are different from the related genes in Neanderthals and Denisovans, after their ancestors branched off from Neanderthals some 600,000 years ago.

Kay Prüfer, et al., Nature, 2013

A catalog of 571 single nucleotide changes distinguishing modern humans from archaic hominins

- Narrow down the number of candidate point mutations from ~35 million differences since the split from chimpanzee when comparing only reference genomes (Consortium 2005) to 31,389 fixed human-specific changes in a previous seminal study (Pääbo 2014).
- 2018 Study: changes that are observed at high frequency in present-day humans; based on 1 Denisovan, 2 Neanderthals, & MHs
- List of 36 genes that carry missense substitutions (bp change which alters protein function) which are fixed in MHs and for which all archaic hominin individuals sequenced so far carry the ancestral state.
- In total, 647 protein-altering changes in 571 genes reached a frequency of at least 90% in the present-day human population

Martin Kuhlwilm, Cedric Boeckx, 201

Only 12 K site 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 in the genome, based on more MH genomes
- Current: 647 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

96 MH Amino Acids: from 87 genes

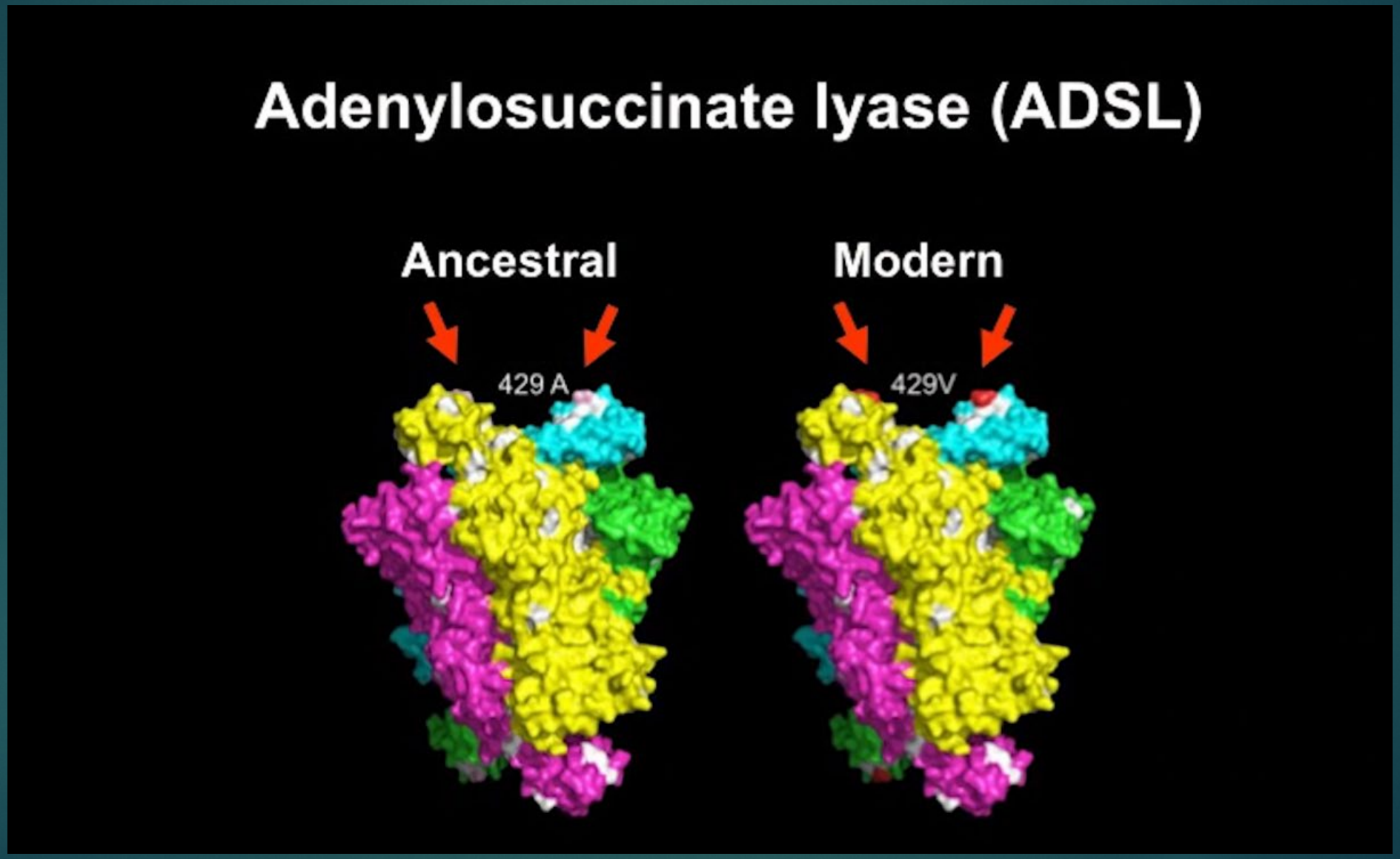
A Catalog of "all" Human-specific Amino Acids

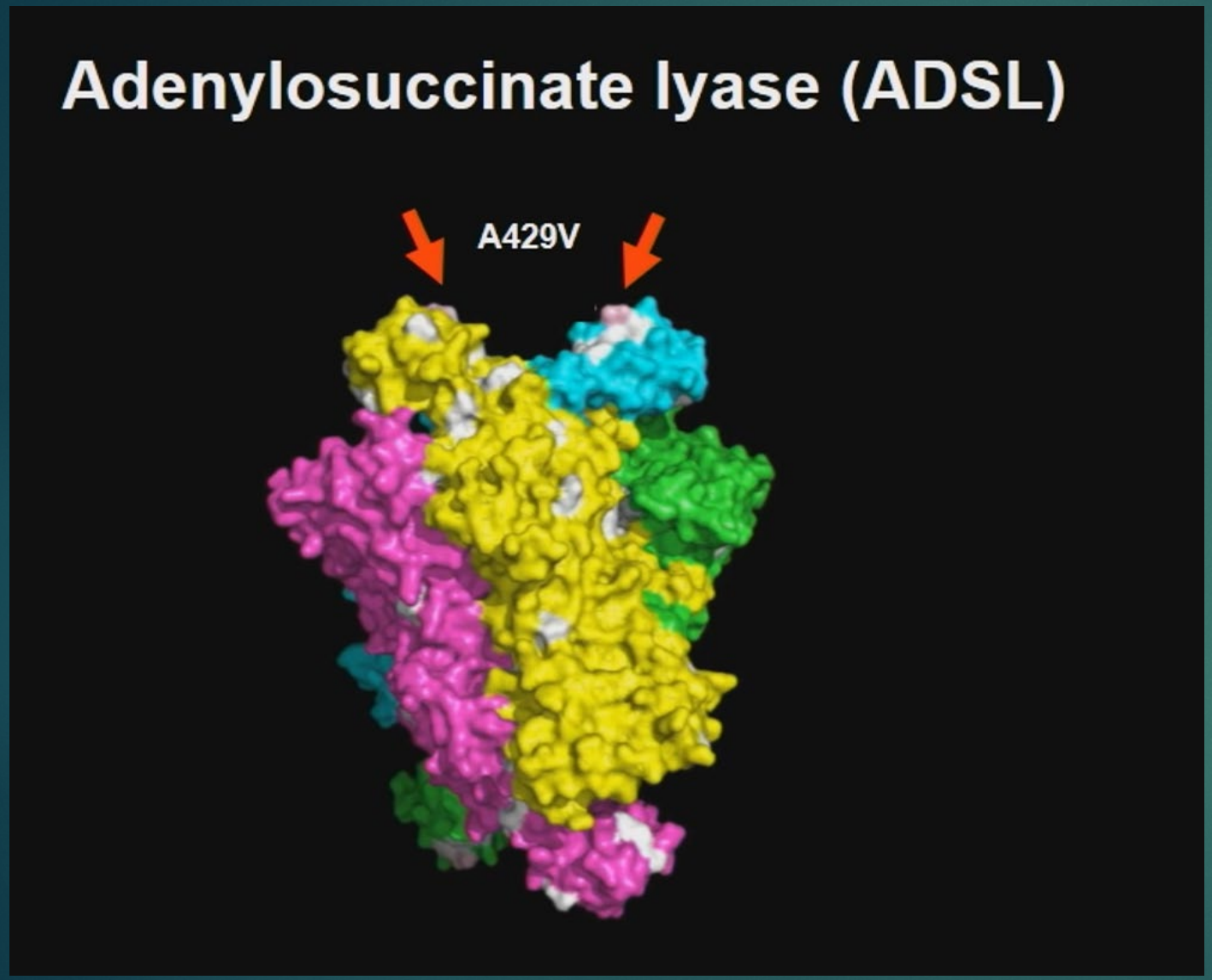
DDX53	NOP14	ADAM18*	KIF18A	SSH2
CXorf59	EVC2	RB1CC1	PLAC1L	SYNRG
Orf	HERC5*	LYPLA1	ZNHIT2	CD300LG
FRMD7	DHX29	GPT	PRDM10	TEX2
ZNF185	PTCD2	GLDC	LRTM2	ITGB4*
TKTL1	SV2C	FRRS1L	LAG3	RFNG
IFI44L	VCAN	NEK6	SCAF11	GREB1L
VCAM1	RASA1	TTF1	SLITRK1	LMNB2
SPAG17*	IRAK1BP1	FBXW5	NOVA1	MFSD12
SLC27A3	MCHR2	FAM166A	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**	

When did these new variants evolve?

- Some human traits evolved shortly after the human-chimp split, e.g. upright walking
- Others evolved later and distinguish us from ancient hominins, e.g. loss of brow ridge
- Some changes are in regions that code for microRNA molecules that regulate protein manufacture.
- What about language?
  - The FOXP2 gene is essential to speech
  - Human FOXP2 is different from chimp
  - Neandertal and Denisovan fossil DNA sequences match human FOXP2
  - But a MH FOXP2 regulator is different

MH Amino acid





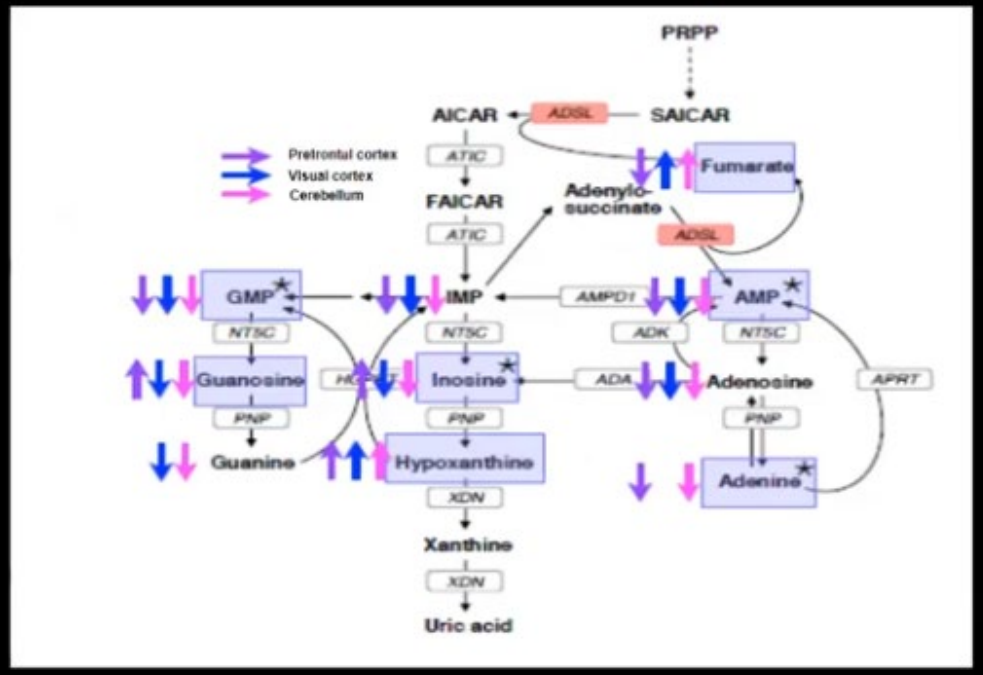
N version increases amount of Purine in cells.
MH version reduces it

In mice, humanized ADSL leads To more social dominance

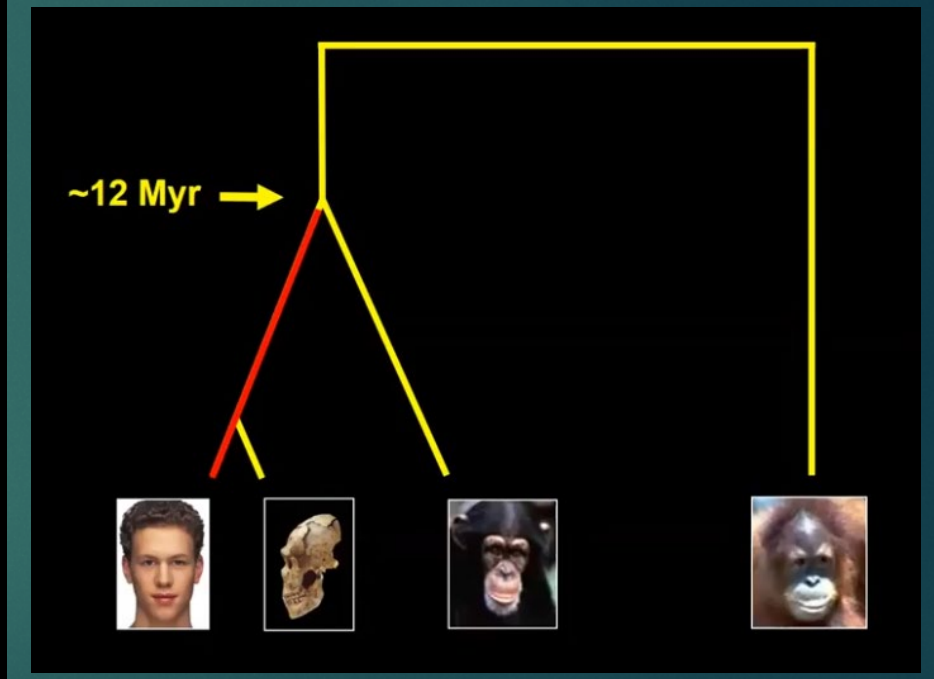
ADSL enzyme is lower in Humans

Brain metabolome
Human-specific differences

	4 x	4 x	4 x
age	7-20	10-40	21-60
# metabolites	163	185	170

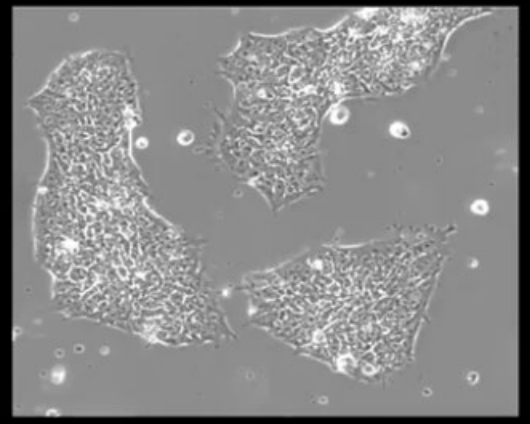


Unpublished data
Vera Stepanova *et al.*
Khaitovich Lab
Skolkovo, Moscow



CRISPR reversion to ancestorial variant: lower expression level downstream

Human cells “ancestralized” for ADSL

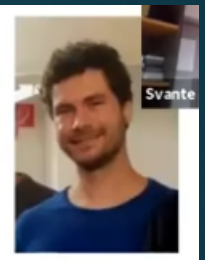


CRISPR-Cas9

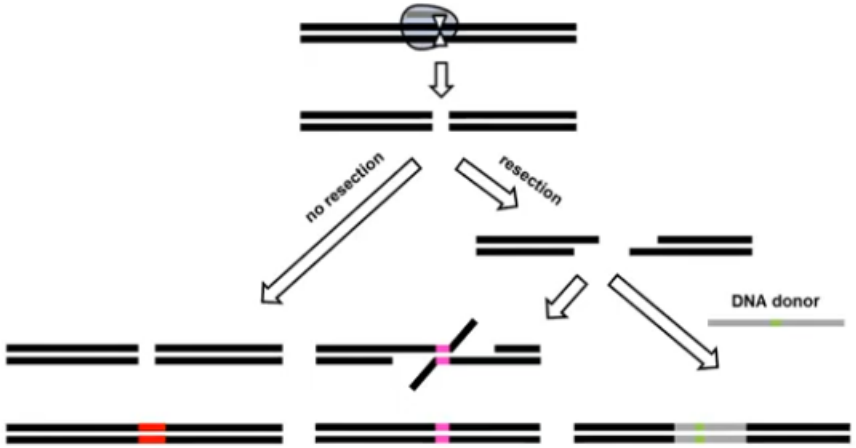
Valine → Alanine

Pos. 429

Multiplexed precise genome editing

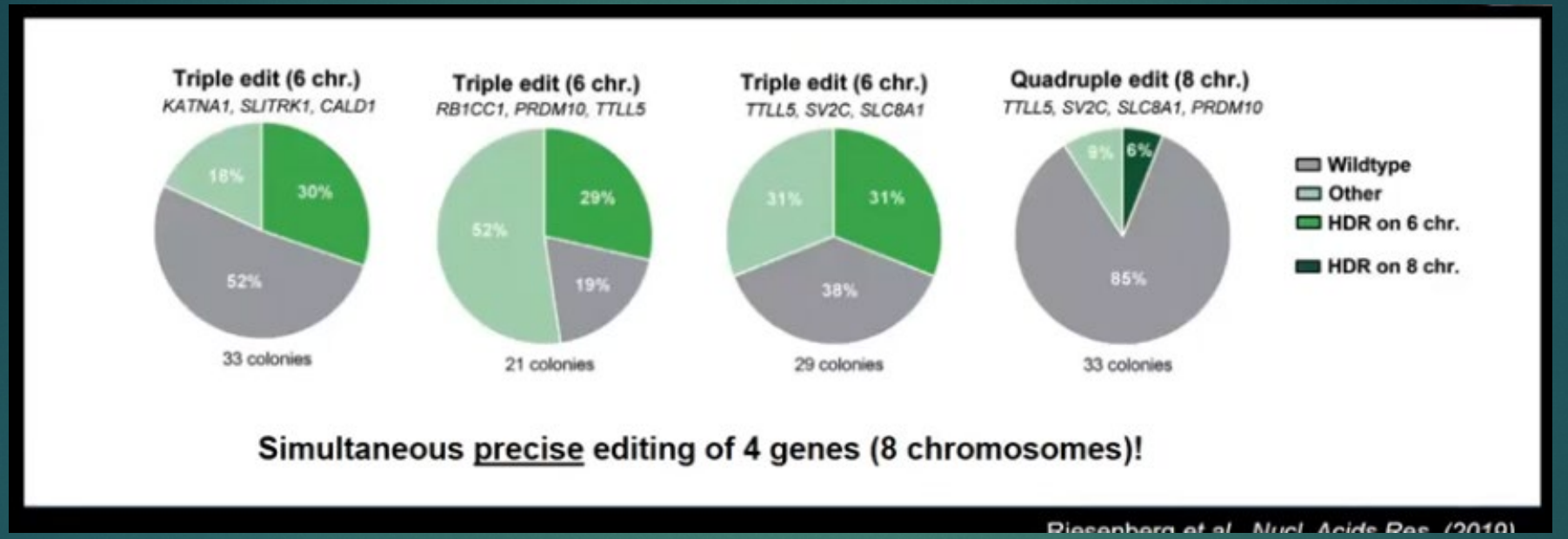


Stephan Riesenberg



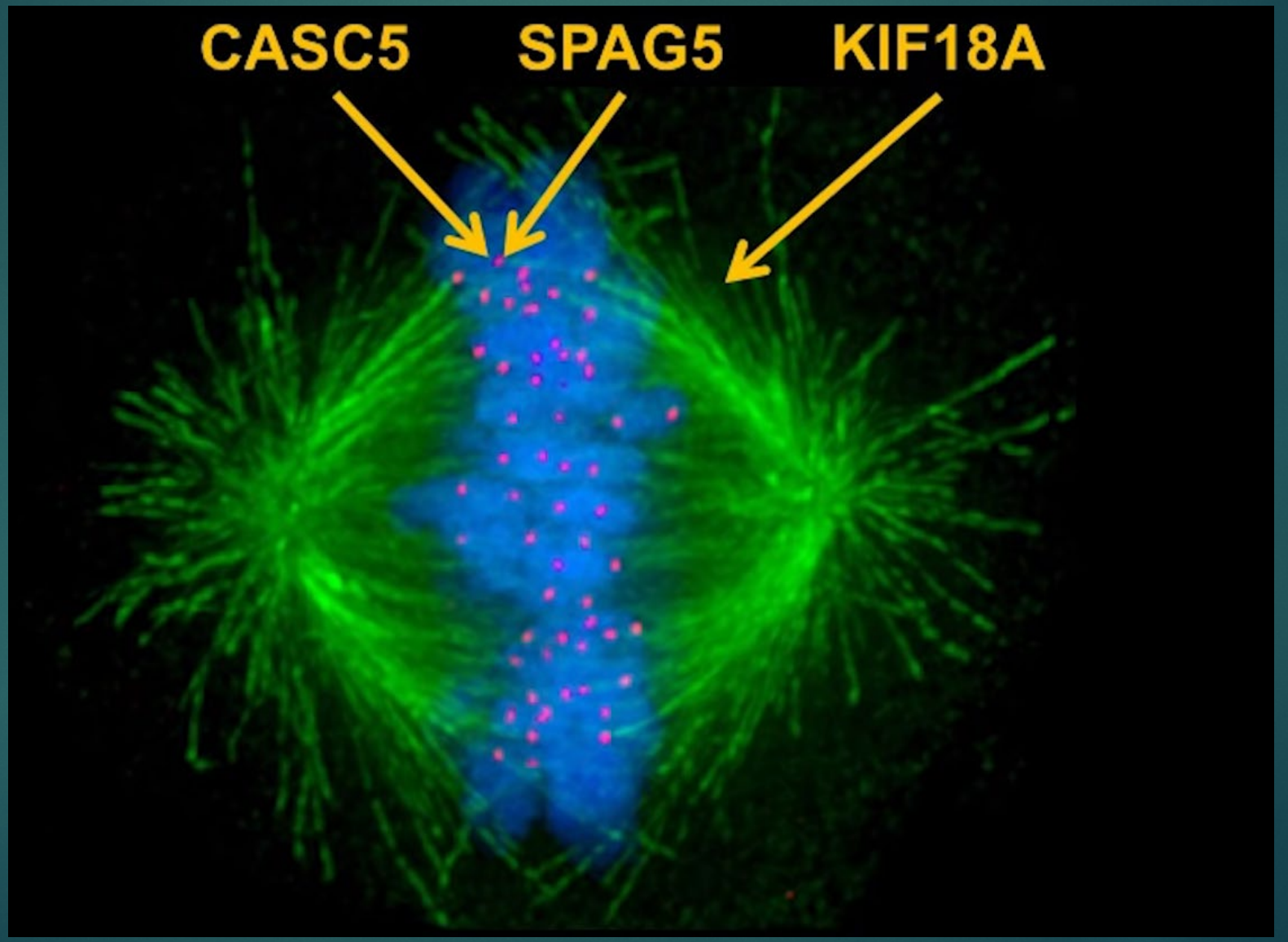
	NHEJ	MMEJ	HDR
Repair pathway:			
Editing event:	indels up to 90%	deletions	precise genome editing ~0.5 to 20%

Simultaneous precise editing of 4 genes to analyze downstream activities

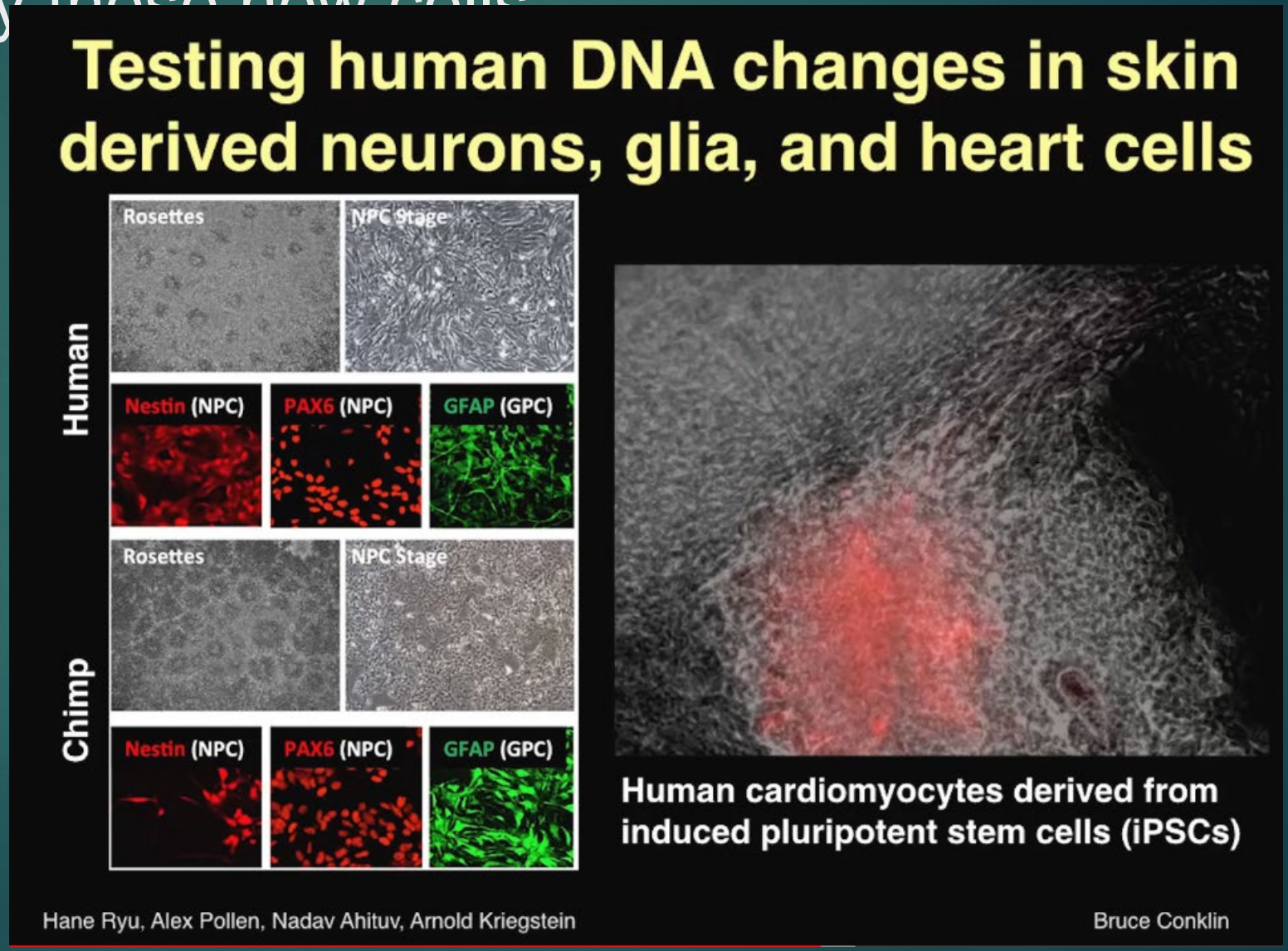


Goal: ancestralize all 100 amino acids to see unique MH effect

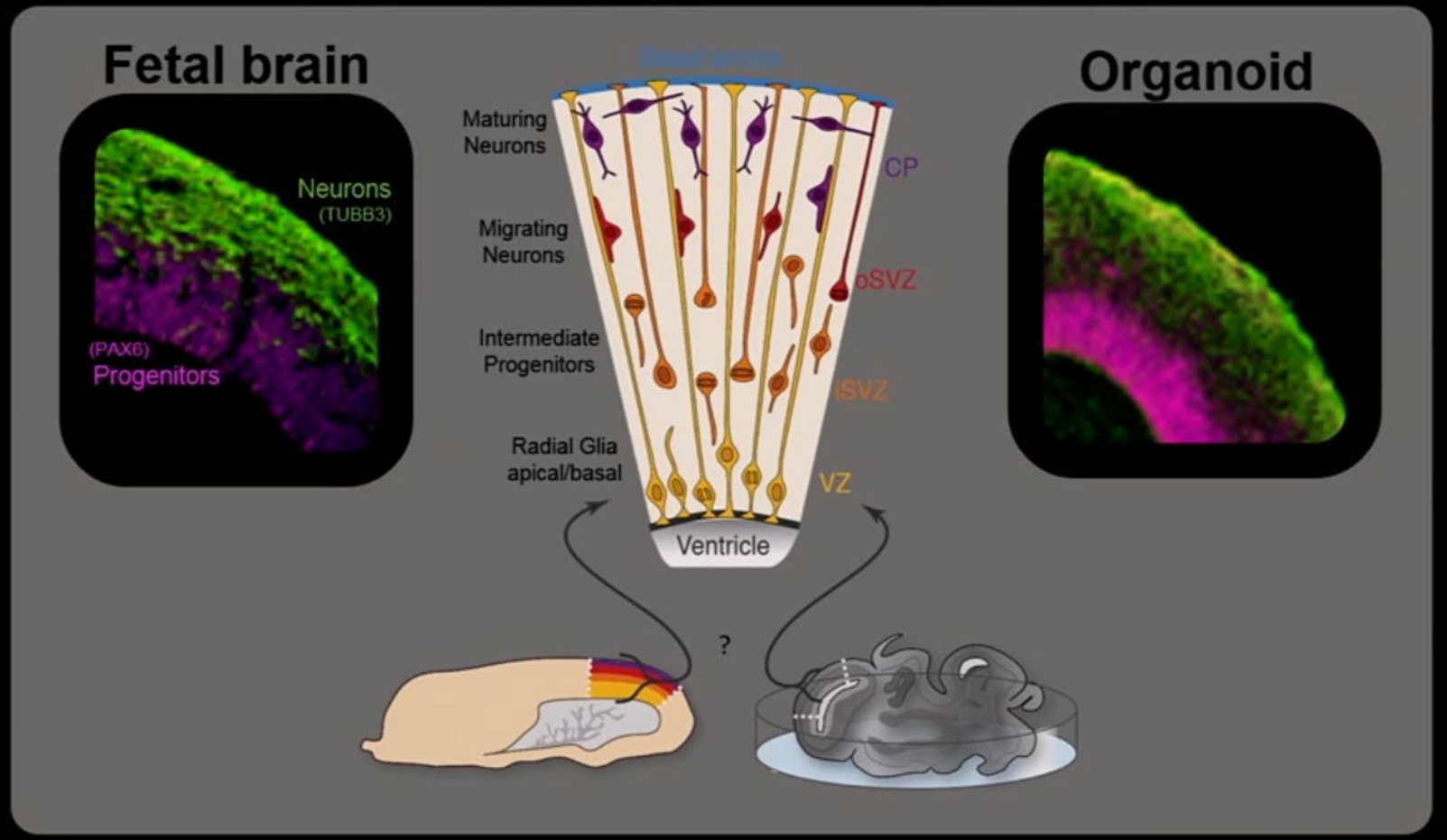
Number of unique MH genes related to spindle activity in cell division



Discovery: can turn pluripotent stem cells into other organ cells; create brain or heart cells in dish; can genetically modify those new cells



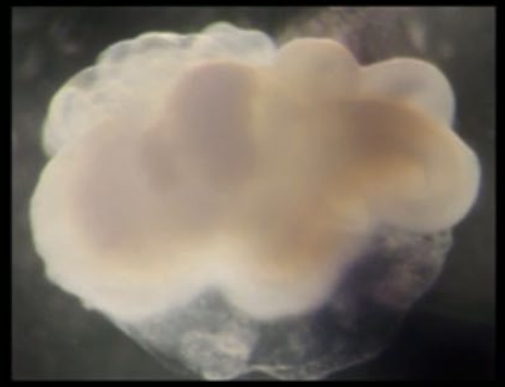
Brain Organoids



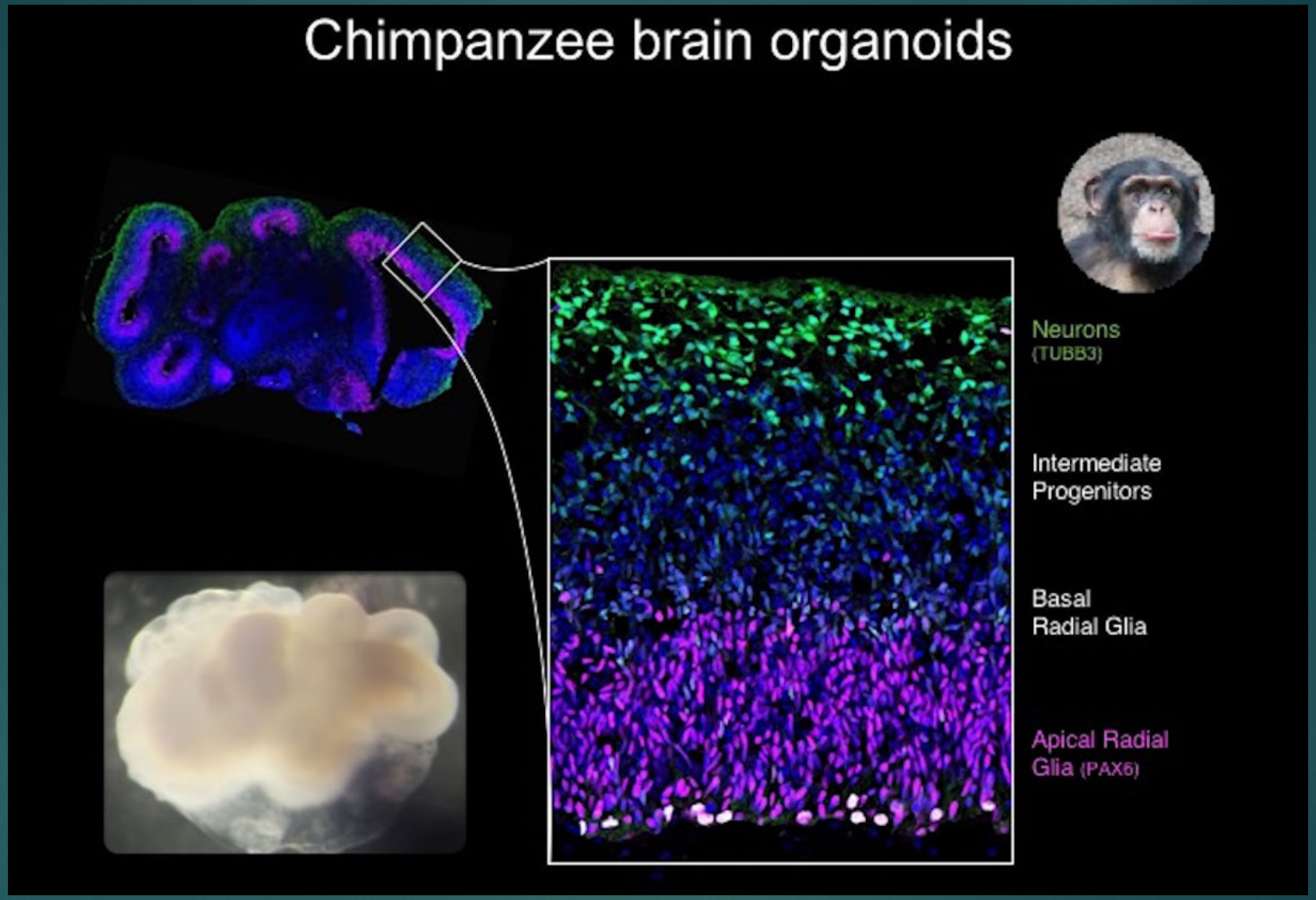
Camp *et al.*, 2016.

To clarify gene function:

Organoids

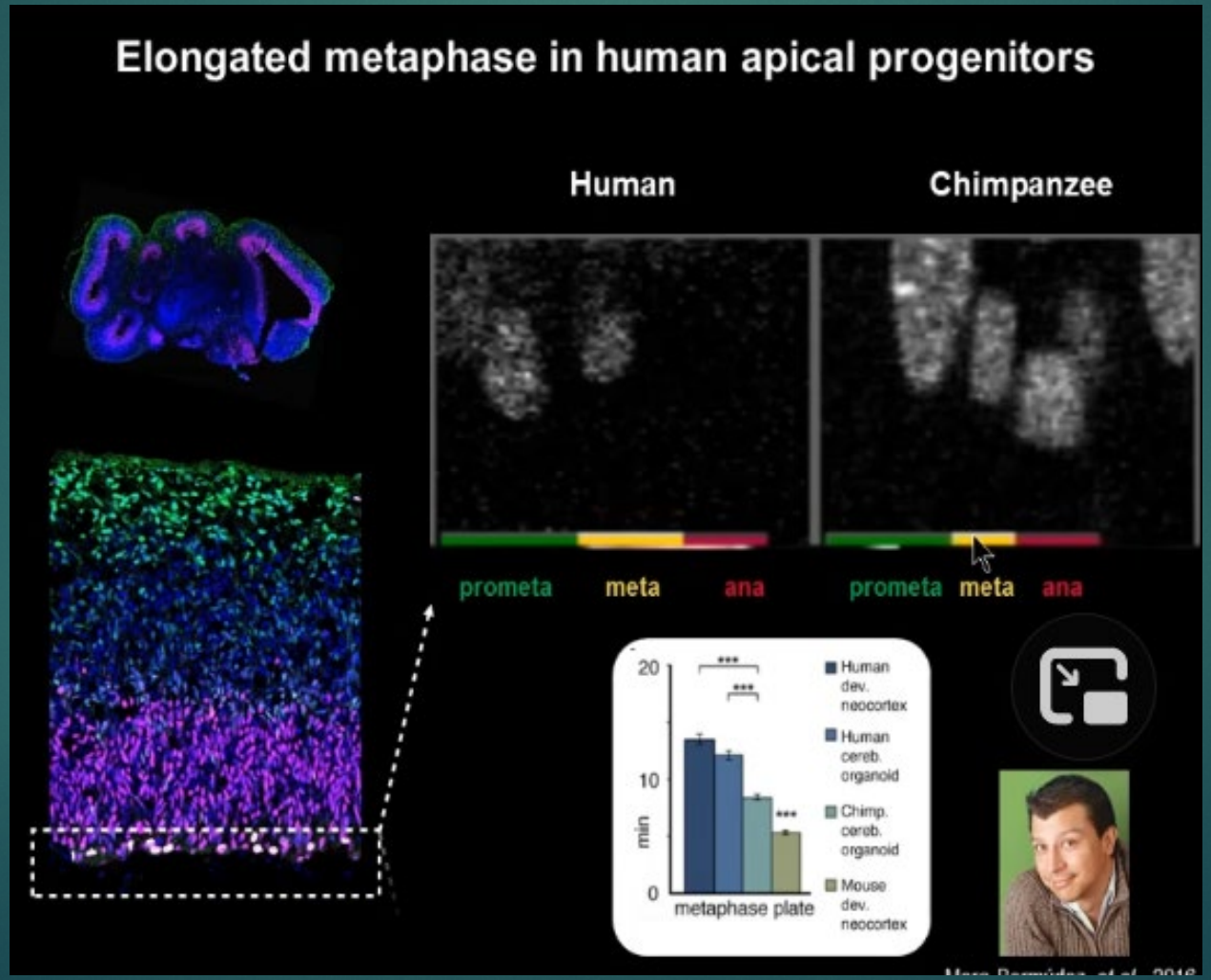


Mice



Indicate spindle genes in MHs create neuronal division that takes longer than in chimps during brain development

Comparison of chimp vs human brain organoids; longer time in human stem cell separation time in creating neurons



Genes that make humans unique

- Significant progress in the identification of genetic changes with functional evidence important in human evolution. At least 25 now.
- Copy-number differences account for more genetic difference at the basepair level among the great ape lineages than single basepair differences
- A punctuated burst of duplications in the human-great ape ancestor 8-15 million years ago
- Approximately the strongest candidates associated with evolution of the human brain and expression differences are associated with structural differences.

Atlai Neandertal

- Kay Prüfer: At least 87 genes found only in modern humans that are different from the related genes in Neanderthals and Denisovans, after their ancestors branched off from Neanderthals some 600,000 years ago.

- Pääbo: 96 functional mutations (alter proteins) that are unique to modern humans:
  - 3 involved in cell division in brain

Kay Prüfer, et al., *Nature, 2013*

How to study human-specific and Neandertal-specific genetic changes?

- Put Neandertal alleles in transgenic human
- Human alleles in transgenic chimps
- Clone a N (George Church at Harvard); Ethics committees object to this!
- Human and Neandertal changes in stem cells
- Human and Neandertal genetic changes in mice
- H and N changes in brain organoids:
  - Human substitutions in the mouse *Foxp2* gene;
  - procedural memory genes (mice with H stem cells learn faster)
- Current hypothesis: Two amino acid substitution in the human FOXP2 alters cortico-basal ganglia circuits to allow faster proceduralization of learning and aspects of speech

** HARs

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).
- It has undergone positive selection. It has 18 differences between humans and chimps; suggesting that HAR1 acquired an important new function in humans.
- 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

Human Accelerated Regions

Short
100-200 letters

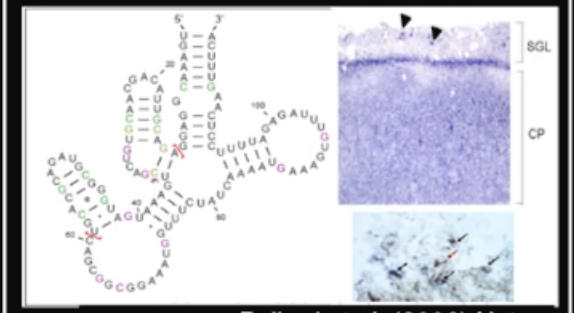
Mostly Outside Genes
Only 4% in proteins

Enriched Nearby
Transcription factors
Developmental genes
Disease genes
Duplicated genes

239 (33%) predicted developmental enhancers
24/29 (83%) of these tested positive *in vivo*
5/29 (17%) showed human-chimp differences

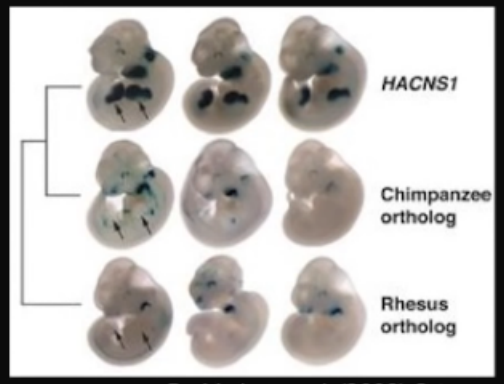
Capra et al. (2013) Phil Trans Royal Soc B

HAR1: a novel RNA gene



Pollard et al. (2006) Nature

HAR2: a novel limb enhancer



Prabhakar et al. (2008) Science

A: most very short, few hundred bps

B – mostly outside gene area; only 4% in protein areas

C – HAR1 = fastest changing area – 18 bps of 118 changed; new gene expressed in cortical development

D – some are enhancers, turn gene on or off

E – HAR2 – gene activity in thumb and wrist during development

F – regulation of gene function; turn on transcription factors, etc.

HARs

- ~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.

HAR1 & HAR2

- HAR1 is special because it does not encode a protein; It encodes RNA
- More than 50% of the genes located near HARs are involved in brain development and function.
- Products of many of these genes go on to regulate other genes.
- Human version of HAR2 relative to the version in nonhuman primates, allow this DNA sequence to drive gene activity in the wrist and thumb during fetal development; enhancing tool use?

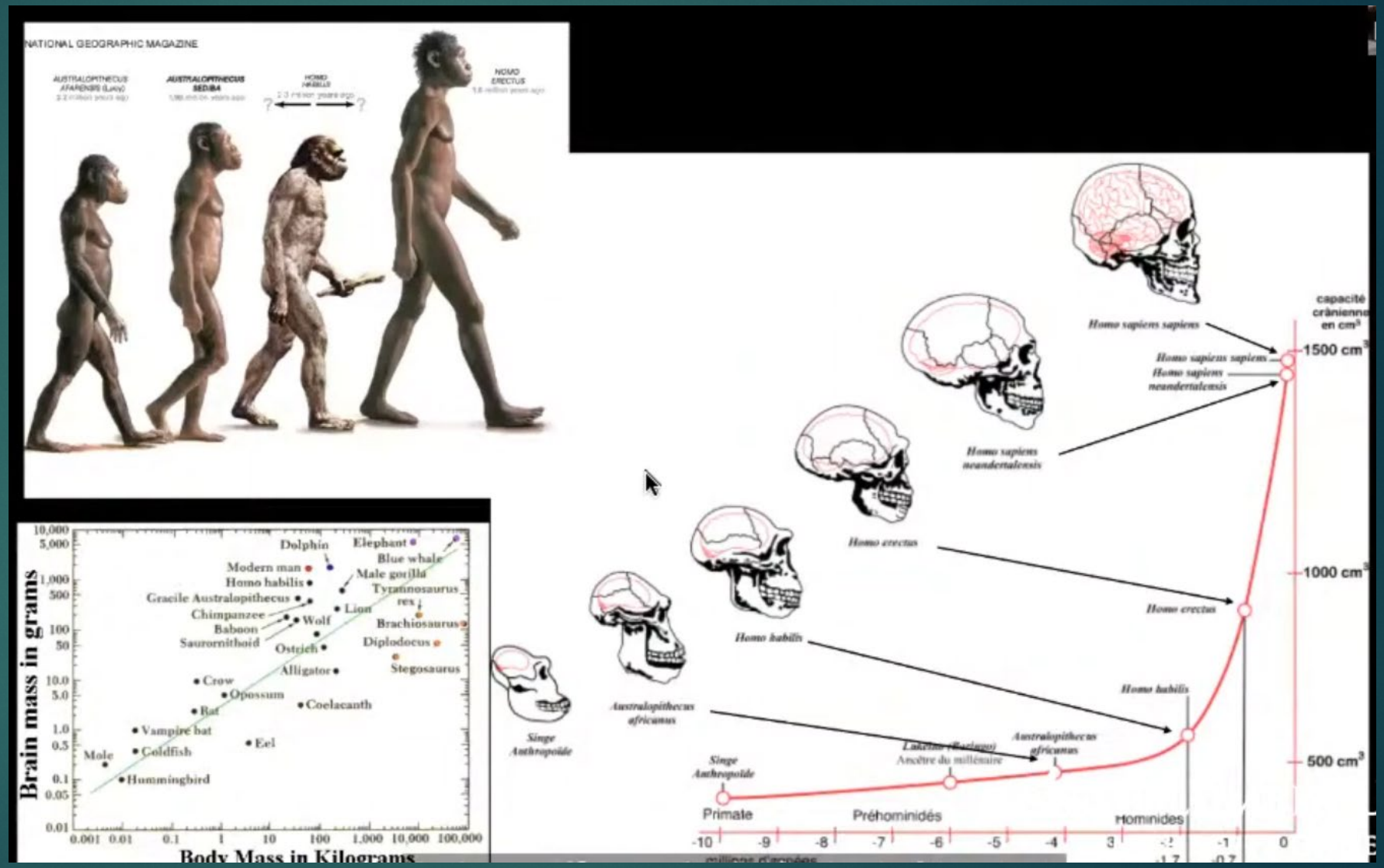
What makes us human?

Gene duplication

- Franck Polleux at Columbia: a type of genetic change called gene duplication.
- As the name implies, gene duplication occurs when a region of DNA is copied and then inserted elsewhere in the genome. So far, dozens of these human-specific gene duplications have been mapped, and his lab was among the first to ask how these gene duplications influence brain development.
- By using CRISPR and other advanced genome engineering tools, we can introduce these changes into the genome of a mouse.

Evolution of brain size in humans

Mammalian brain size varies by 100,000x;
Brain size varies by body size
So brain size is not what makes us human



What makes our brain human-specific

- Brain size?
- Neocortex/brain ratio?
- Probably:
- Neuron number
- Neuronal composition?
  Neuronal connectivity?
  Synapse types & number?

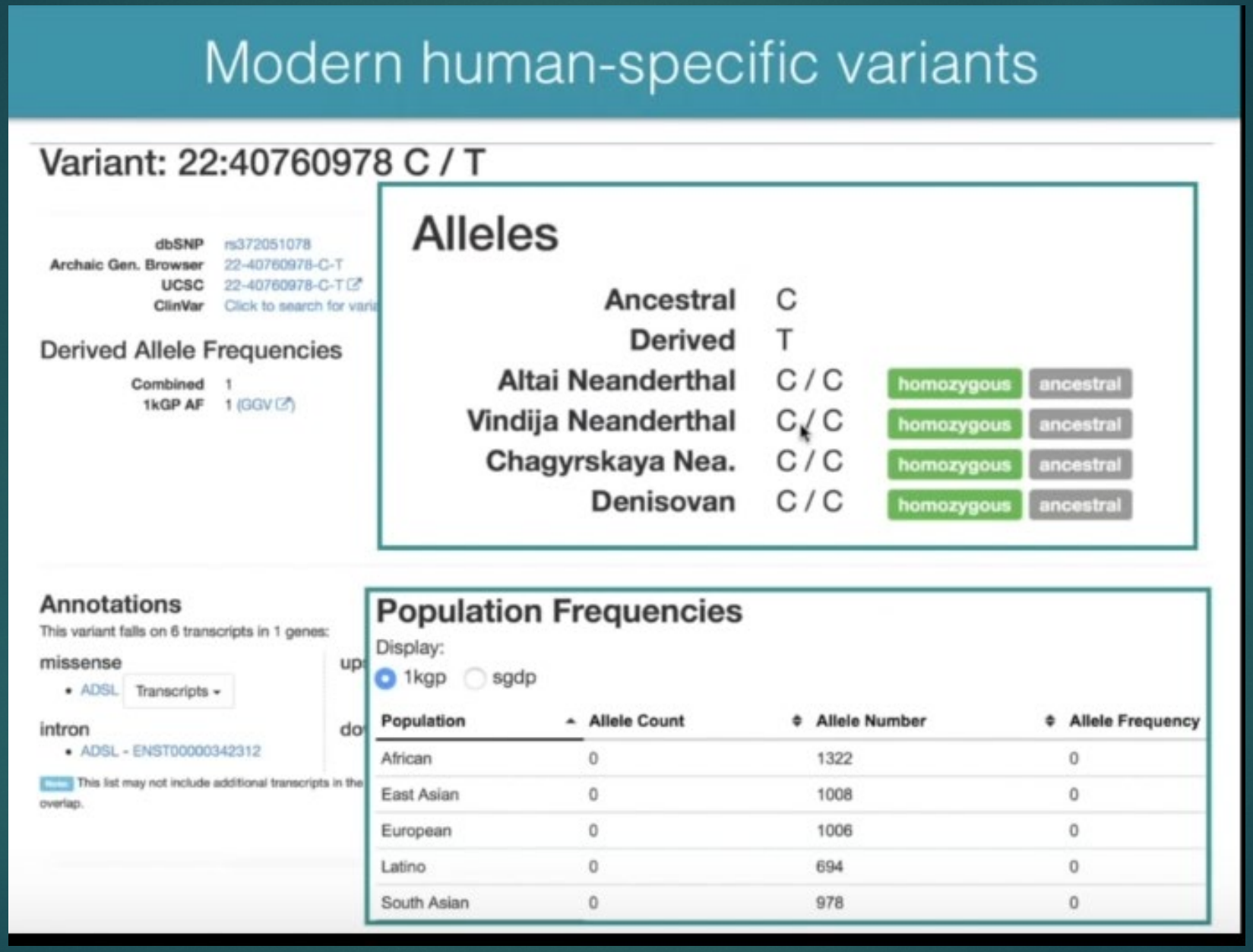
SRGAP2

- SRGAP2 is duplicated in humans.
- Have found that, in humans, SRGAP2 gene duplications drastically boost the number of synapses in the brain.
- By introducing the human-specific version of SRGAP2 into a mouse model, we can stimulate the growth of synapses.
- Causes *accelerated synaptic* maturation in mutant mice
- Third most abundant postsynaptic protein.
- Mutations frequent cause of ID/ASD

Conserved vs accelerated DNA regions

- Conserved sequence refers to regions of DNA that have been unchanged or minimally changed throughout mammalian evolution, including that of the primate lineage.
- Research: looking for human-specific changes in these conserved regions, and the body of evidence led to the definition of Human Accelerated Regions (HARs), regions of evolutionarily conserved sequence that are significantly changed in humans
- 49 such regions have been defined.
- Other research focusing specifically on genomic loss has found 500 deletions specific to humans in otherwise conserved sequences.

Modern human-specific variants



Can compare MH derived with N or D variant

HARs: non-coding regions

- HARs influence human-specific phenotypic features significantly,
- Majority of HARs and human-specific deletions are found in non-coding regulatory regions involved in gene expression.
- Many of the HAR sequences are predicted to be enhancers and other regulatory signals, while others may encode structural sites and RNA genes.
- Additionally, many changes in HARs are predicted to have created or destroyed transcription factor (DNA to RNA) genes.

HAR variants

- HAR1 (118 bps) is a novel RNA gene expressed in the neocortex during development. Does not encode a protein. If defective, lissencephaly (“smooth brain”); schizophrenia
- Also in sperm production
- HAR2 is a limb enhancer (thumb & wrist) with human-specific gene expression in the embryonic hand. Permitted the dexterity needed to manufacture and use complex tools?

HAR

- 201 other human accelerated regions have been found, most of which do not encode proteins or even RNA.
- Amazingly, more than half of the genes located near HARs are involved in brain development and function.
- Thus, even though HARs make up a minute portion of the genome, changes in these regions could have profoundly altered the human brain by influencing the activity of whole networks of genes.

2006	*HAR1A* (Pollard et al. 2006b) HARs (Pollard et al. 2006a) HACNSs (Prabhakar et al. 2006)
2007	ANCs (Bird et al. 2007)
2008	HARs (Bush and Lahn, 2008)
2010	HSDs (Sudmar
2011	hCONDELs, *GADD45G* (McLean et al. 2011) 2xHARs (Lindblad-Toh et al. 2011)
2012	HARs in archaic humans (Burbano et al. 2012) *SRGAP2* (Dennis et al. 2012; Charrier et al. 2012)

An external file that holds a picture, illustration, etc.
Object name is evx240f2.jpg

HAR1 was essentially frozen in time through hundreds of millions of years indicates that it does something very important; that it then underwent abrupt revision in humans suggests that this function was significantly modified in our lineage

21 of 49 HARs

2013	*AUTS2* (Oksenderg et al. 2013) *NPAS3* (Kamm et al. 2013a; Kamm et al. 2013b) Function of HARs as enhancers (Capra et al. 2013)
2015	SCZ-associated genes and HARs (Xu et al. 2015) *FZD8* (Boyd et al. 2015) *ARHGAP11* (Florio et al. 2015) haDHSs (Gittelman et al. 2015)
2016	Autism and HARs (Doan et al. 2016) *ARHGAP11* (Florio et al. 2016)
2017	HSDs (Dennis et al. 2017)

HARs in the context of Neandertal and Denisovan genomes

- Mutations in HARs have come to fixation faster in the human genome than other non-HAR genomic mutations due to positive selection
- HAR substitutions tend to occur episodically over time.
- 8% of these HAR substitutions were not found in either archaic hominin genome, suggesting that they had not yet become fixed in the genome of the common ancestor of modern humans and archaic hominins.
- Transcription factor *neuronal PAS domain-containing protein 3 (NPAS3)* has the largest population of noncoding-accelerated regions.
- *NPAS3* is active during mammalian brain development and the human accelerated elements within this locus predominantly appear to act as transcriptional enhancers and thus may have influence human brain evolution.

What are HAR functions?

- RNA genes = create RNA
- Many types of gene regulatory elements = turn genes on and off
- Sequences specifying DNA “activity” = can silence development genes in adults; effect epigenetics (environmental triggers for gene activity)
- Sites controlling 3D organization of DNA = DNA is 3D dynamic structure; gene activity can depend on dynamic location
- Unknown functions yet to be discovered

Table 1. Single-Gene Studies Potentially Related to the Evolution of Human Cognitive Abilities

Gene	Unique Evolutionary Feature	Reference
FOXP2	Implicated in language deficit	[14]
ASPM	Implicated in change in brain size	[64]
MCPH1	Implicated in change in brain size	[68]
PDYN	Human-specific alteration of regulatory region	[15]
GLUD2	Implicated in ape brain evolution	[69]
COX8	Potentially related to increased energy demand of brain	[70]
CMAH	A sialic acid hydroxylase activity lost in human lineage	[71]

Putting MH and N genes into Organoids

- Identified 61 unique MH genes = mutations that are unique to our species, arising some time in the last 600,000 years, and likely had a major impact on the proteins encoded by these genes.

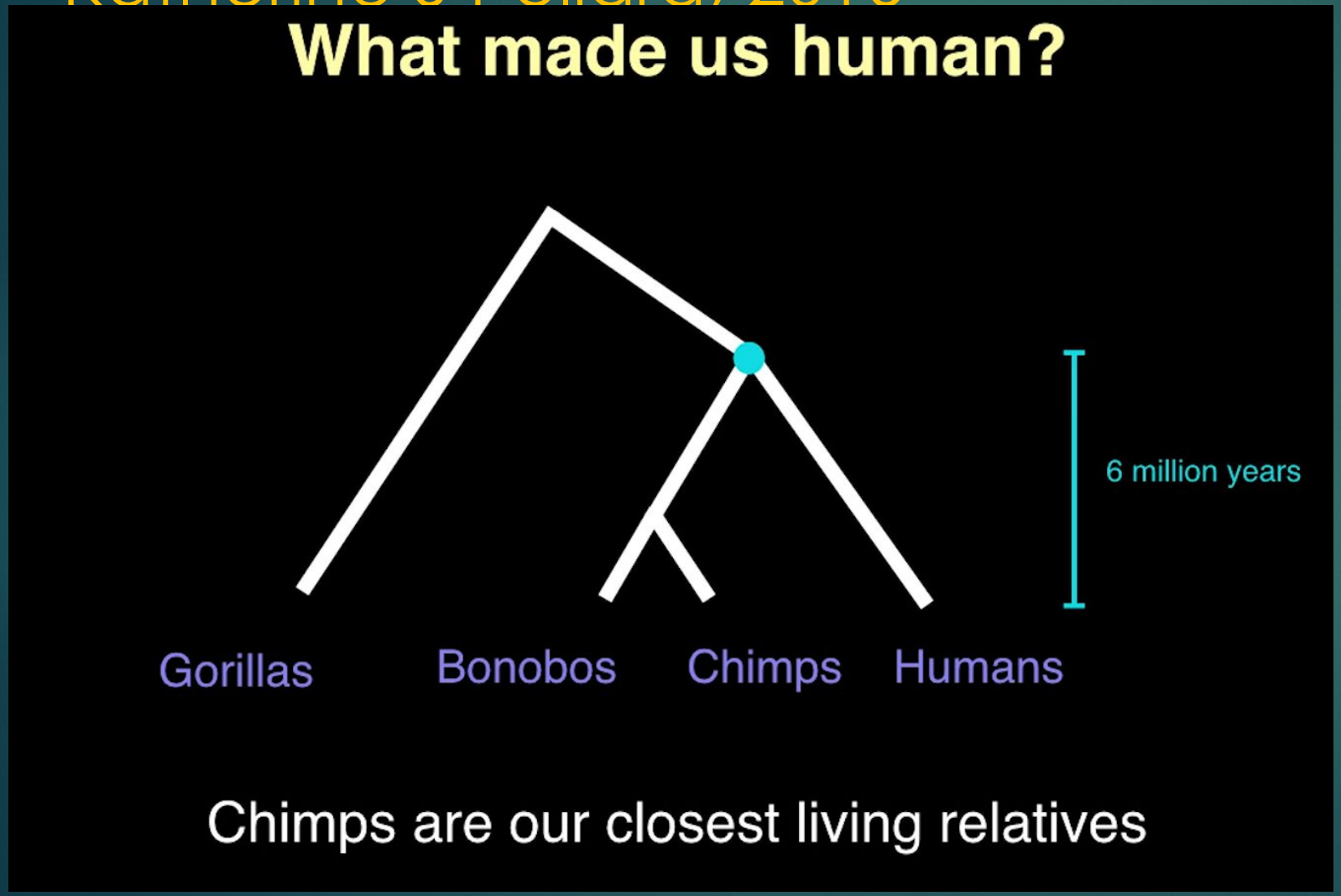
- Removal of 1 gene at a time experiments

- Reintroduction of the archaic N or D variant of *NOVA1* in cortical organoids alters neurodevelopment

- Evolutionarily conserved splicing regulator neuro-oncological ventral antigen 1 (*NOVA1*) plays a key role in neural development and function. RNA-binding protein NOVA1; in developing brain

Cleber A. Trujillo, et al., 2021

The festest evolving regions of the human genome
– Katherine S Pollard, 2016



- Efforts to identify those regions of the human genome that have changed the most since chimps and humans diverged from a common ancestor have helped pinpoint the DNA sequences that make us human.

Chimp and human proteins are very similar in size and sequence

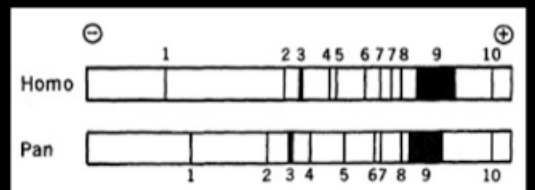


Fig. 1. Separation of human and chimpanzee plasma proteins by acrylamide electrophoresis at *p*H 8.9. The proteins are: 1, α_2-macroglobulin; 2, third component of complement; 3, transferrin; 4, haptoglobin; 5, ceruloplasmin; 6, α_{2HS}-glycoprotein; 7, Gc-globulins; 8, α_1-antitrypsin; 9, albumin; and 10, α_1-acid glycoprotein. The chimpanzee plasma has transferrin genotype *Pan* CC; the human plasma has transferrin genotype *Homo* CC and haptoglobin genotype 1-1. The direction of migration is from left to right.

Table 1. Differences in amino acid sequences of human and chimpanzee polypeptides. Lysozyme, carbonic anhydrase, albumin, and transferrin have been compared immunologically by the microcomplement fixation technique. Amino acid sequences have been determined for the other proteins. Numbers in parentheses indicate references for each protein.

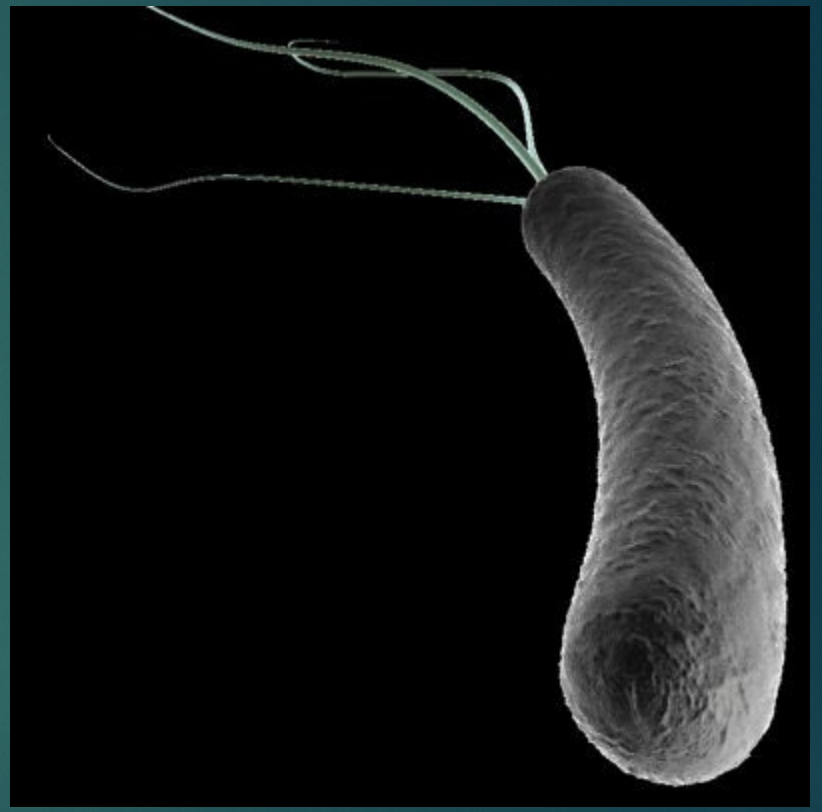
Protein	Amino acid differences	Amino acid sites
Fibrinopeptides A and B (*3*)	0	30
Cytochrome c (*4*)	0	104
Lysozyme (*13*)	~0	130
Hemoglobin α (*4*)	0	141
Hemoglobin β (*4*)	0	146
Hemoglobin $^{A}\gamma$ (*5, 6*)	0	146
Hemoglobin $^{G}\gamma$ (*5, 6*)	0	146
Hemoglobin δ (*5, 8*)	1	146
Myoglobin (*7*)	1	153
Carbonic anhydrase (*4, 12*)	~3	264
Serum albumin (*10*)	~6	580
Transferrin (*11*)	~8	647
Total	~19	2633

Evolution at two levels in humans and chimpanzees
King & Wilson, *Science* (1975)

1975 study: Only 12 genes differed; study concluded regulatory regions, not genes, were important in the differences in these 2 groups

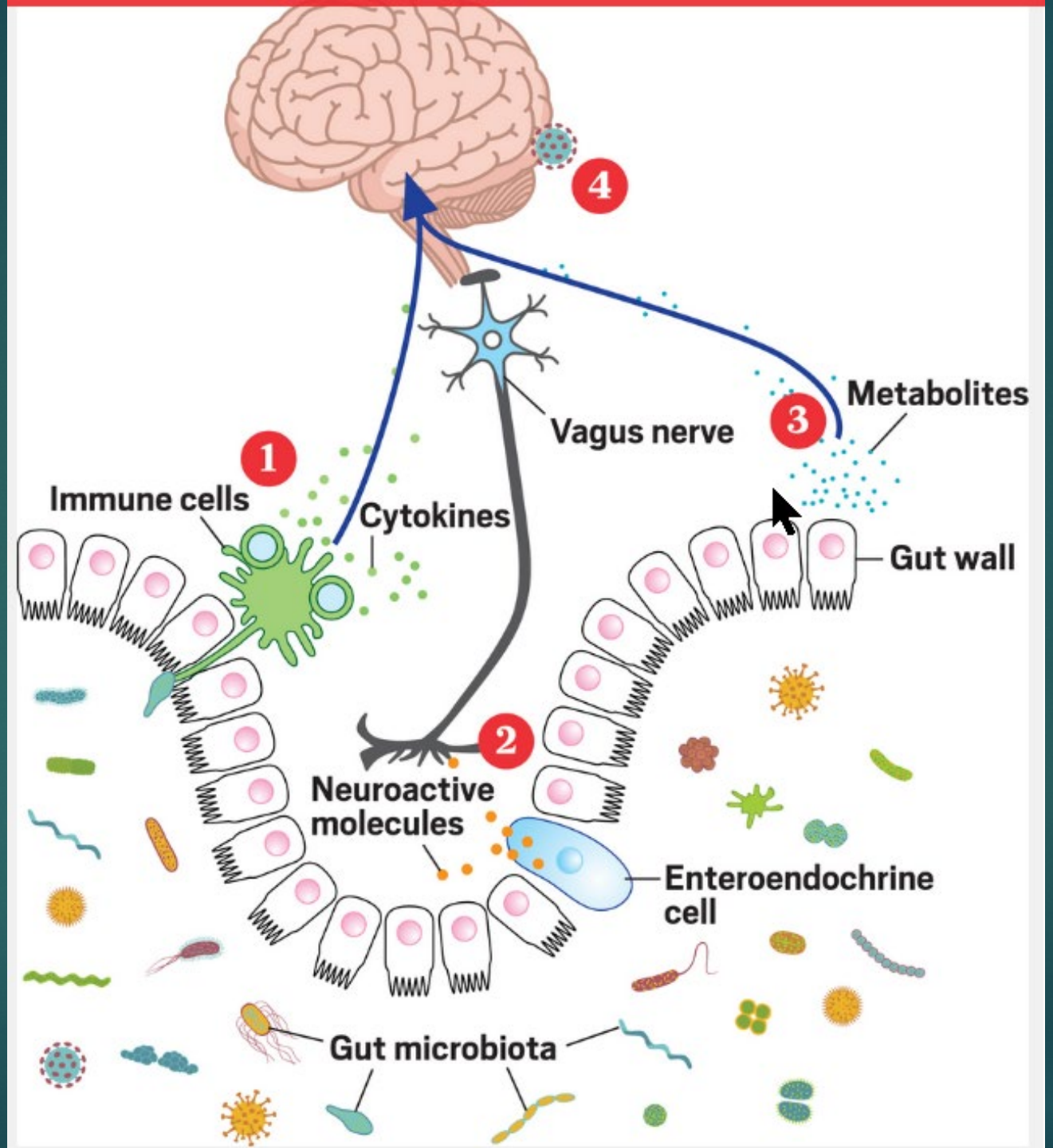
Neural System: Not much has changed in the last billion years…

- Find food; avoid predation; in a changing environment
- Basic input-output organization
- Sensory apparatus
- Movement (effector) apparatus
- More or less complex links between perception and action
- Closeness of mouth & brain: brain arose as the gut's way of controlling intake; still have genes that control both formation of gut & forebrain
- Slightly elaborated in humans…

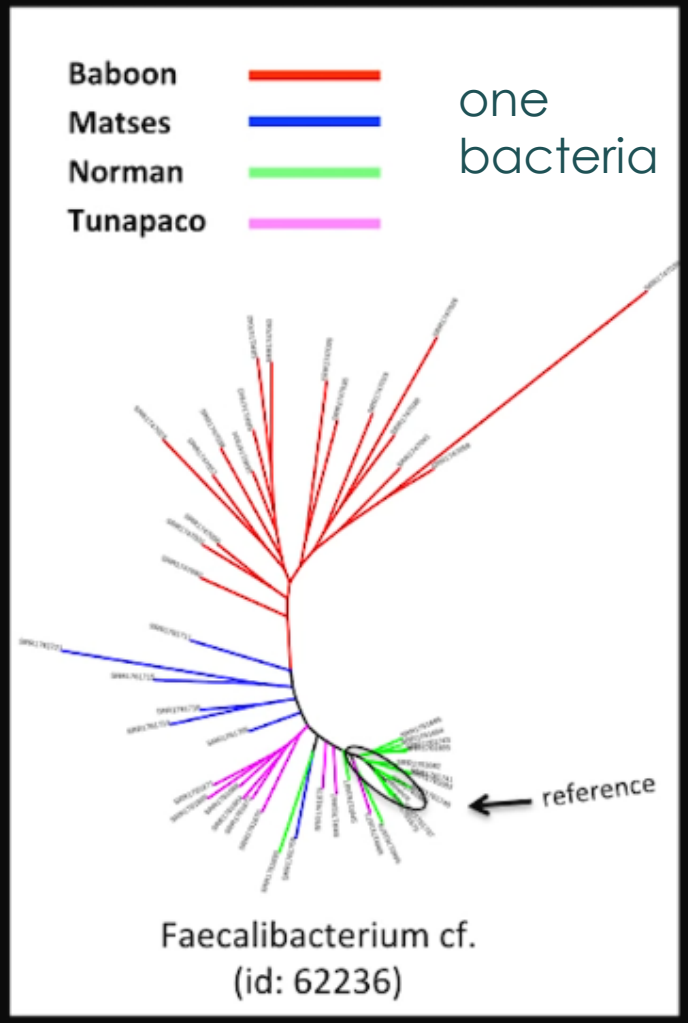


Microbiome Genome: Our Other Genome

- Microbes in our bodies
- Curled up inside us, our intestines have a surface area of around 32 m^2, or 344 ft^2. Put another way, the intestines of an average adult occupy the same area as a small studio apartment in New York City
- Makes up ~5 lbs of body weight
- Trillions of microbes; Outnumber human cells by 1-10x
- Contribute 100x more genes in these bacteria and viruses
- Produces 95% of serotonin in body
- Are essential player in normal human biology and disease
- 40-50% of our genome are viruses that have integrated with us; some are junk, others have been functionally coopted by us


4
Vagus nerve
3
Metabolites
1
Immune cells
Cytokines
Gut wall
2
Neuroactive molecules
Enteroendochrine cell
Gut microbiota

Microbiome Evolution



- Humans from different countries harbor distinct microbial communities
- Microbiomes of non-human primates are more distinct
- This is true even when two hosts contain the same species. Strains can differ significantly in their genes!

Microbiome connected to brain

- Mice chemically coaxed to produce high levels of an autism-linked gut molecule have anxiety-like behavior and unusual patterns of brain connectivity; and a decrease in myelin
- unusual connectivity patterns in brain regions related to emotion processing,
- Ketogenic diet: neurologists have known that putting children with epilepsy on a specific diet can reduce their seizures; two particular types of gut bacteria thrive in mice feasting on a ketogenic diet; may produce GABA

Evolution of genetic networks for human creativity

- I. Zwir, et al., 2021 study: Creativity could be one of the main reasons *Homo sapiens* survived and dominated over related species

- MHs: demonstrate remarkable creativity compared to other hominins: that is, they show signs of innovation, flexibility, depth of planning, and related cognitive abilities for symbolism and self-awareness that enable spontaneous generation of narrative art and language

- MHs: also more prosocial in their behaviors than archaic hominins: they maintained larger social groups, established reciprocal social networks for remote trade, and regularly cooperated with one another in groups

I. Zwir, et al., 2021

Evolution of genetic networks for human creativity

- Creativity = divergent thinking = inferior prefrontal; convergent thinking= lateral prefrontal/parietal (executive self-control network)
- Used best measures of domains of the creative personality = the Temperament and Character Inventory (TCI); Did genome-wide association studies of the TCI in three different samples with different environments and cultures (Finns, Germans, and Koreans).
- Found 42 SNP sets that were significantly associated with the character profiles and identified 727 gene loci.; 50 SNP sets/736 gene loci mapped with temperament.
- Each of the three phenotypic networks was strongly correlated with a different multi-locus genotypic network

3 genetic networks for human creativity

- The functions of the genes that mapped to the genotypic networks were found to regulate distinct systems of learning and memory underlying personality:

- (i) "emotional reactivity" network: a multi-locus network of 249 genes for regulation of emotional reactivity, associative conditioning, and social attachments

- (ii) "self control" network: a multi-locus network of 438 genes for regulation of intentional goal-seeking, such as purposeful acquisition of food, manufacture of tools, cooperative team-work, logical analysis, and symbolization

- (iii) the "self-awareness" network: a genotypic network of 574 genes for episodic learning and autobiographic memory of a person's life as a narrative with past, present, and future within which the person can explore alternative perspectives with intuitive insight and creative imagination

Creativity

- It is remarkable that 73% of the 972 genes in these three networks are unique to a single network.

- The genes identified for temperament and character accounted for nearly all the heritability of personality expected from twin studies.

- The strong relations of the three temperament-character phenotypic networks to three major genotypic networks for human adaptability provided us with valuable tools for evaluating the evolution of human creativity and other aspects of behavioral modernity

Evolution of genetic networks for human creativity

- Hypothesized that the three genotypic networks for human adaptability evolved in successive steps during the evolution of modern human personality.
- Hypothesized that (i) chimpanzees would have genes only in the emotional reactivity network, (ii) both Neanderthals and Sapiens would share many genes for intentional self-control, which was already evident in their common human lineage, and (iii) genes found only in Sapiens would be most frequent in the network for creative self-awareness
- We confirmed the hypothesis that the genes related to the character of Sapiens were over-expressed in brain regions that have been involved in human self awareness and autobiographical memory.
- Specifically, they were significantly over-expressed in late-myelinating regions of neocortex in frontal, temporal, and parietal regions

Evolution of genetic networks for human creativity

- 972 modern genes that regulate three distinct systems of learning and memory in *Homo sapiens*: emotional reactivity, self-control and self-awareness.
- New study analyzed DNA previously taken from Neanderthal (fossils, modern humans, and chimpanzees). They found that the genes related to the oldest network — emotional reactivity — were identical among *Homo sapiens*, Neanderthals and chimpanzees.
- But the chimpanzees completely lacked the genes that led to self-awareness and self-control in humans.
- Neanderthals had nearly the same genes for emotional reactivity as chimpanzees, and they were intermediate between modern humans and chimpanzees in their numbers of genes for both self-control and self-awareness

Creativity genes

- 267 of those 972 genes were unique to *Homo sapiens*.
- 95% of the 267 genes only in modern humans; were not protein-coding, but were regulatory genes.
- The genes that cluster in association with those found only in modern humans are over-expressed in brain regions involved in human self-awareness and creativity, including late-myelinating and phylogenetically recent regions of neocortex for autobiographical memory in frontal, parietal, and temporal regions.

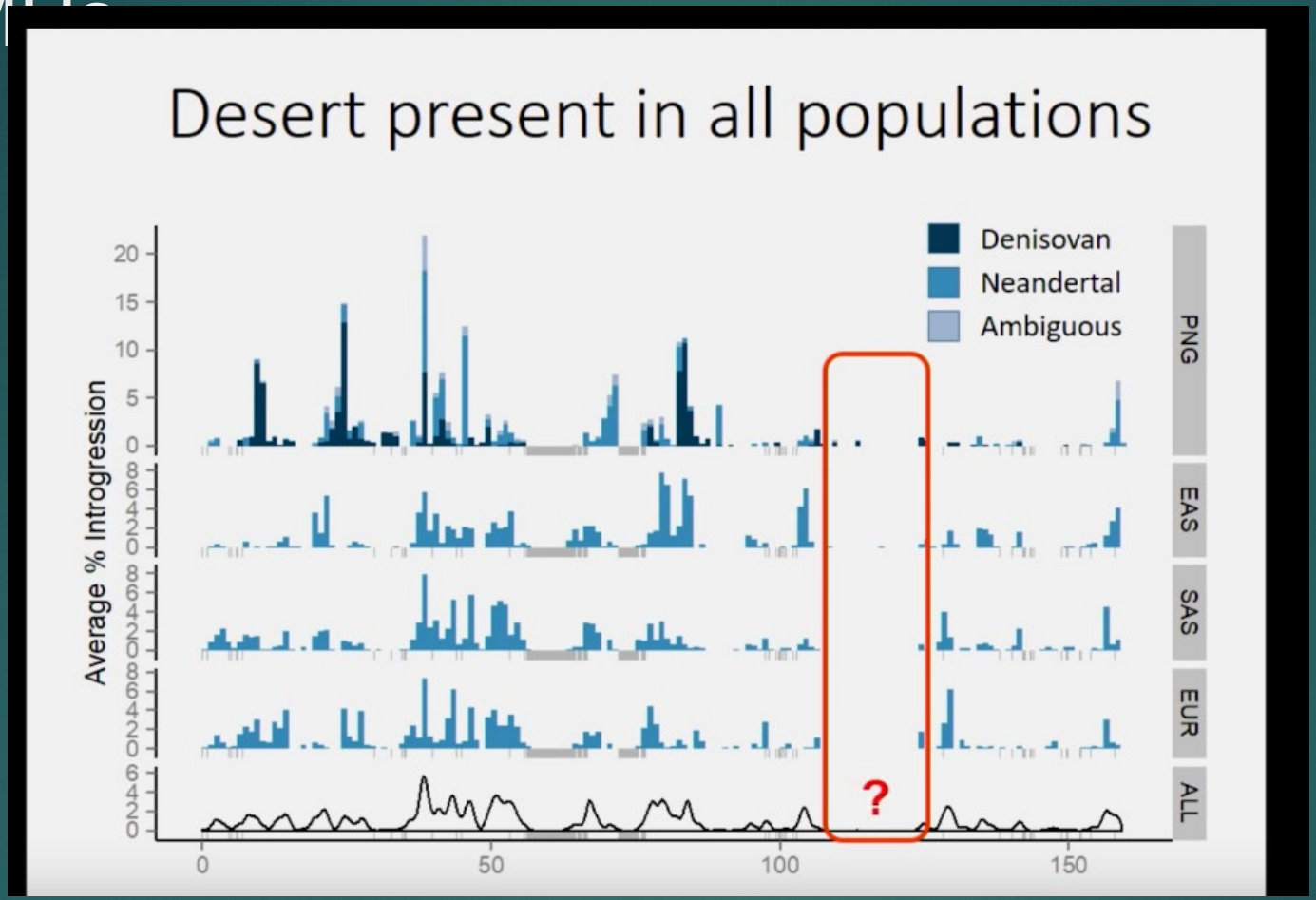
Evolution of genetic networks

- The emotional reactivity network evolved in monkeys and apes about 40 million years ago,
- self-control network evolved a little less than 2 million years ago,
- self-awareness and creativity network emerged just 100,000 years ago

Creativity genes: critique

- We do not know the causal link between genetics and these higher traits.
- Neanderthals may not have had the same genes for creativity and self-awareness, but rather their own set of genes that we don't understand.

Genetic desert areas of no N or D effect; functions unique to M



Genes in testicles: On the X chromosome, less N genes; and other areas

Complicated

- 30% of the human genome, or around 6,000 genes, govern neurological systems
- Each of the estimated 19,000 **genes** in the **human genome** makes an average of three proteins.
- Brain synapsis have 2000 proteins connected to them

New genomics: GWAS

- Genome-wide association studies (GWAS)
- which allow scientists to search thousands of individual genomes to discover genetic variants that are linked to specific traits or diseases,
- including height, obesity and susceptibility to complex diseases such as schizophrenia.

GWAS

- GWAS have successfully identified genomic regions that increase the risks of common conditions such as diabetes, coronary artery disease, schizophrenia, and Crohn's disease.
- There have been more than 10,700 GWAS done since 2005.
- But 78% of GWAS are of Europeans.

GWAS

- Have used 1,000,000 whole genomes
- 150,000 genomes that show an incredible amount of human genetic diversity.
- There are 241 million differences in people's genomes, with an average of one variant for every eight base pairs

Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression

- Major depressive disorder (MDD) is a common illness accompanied by considerable morbidity, mortality, costs, and heightened risk of suicide.
- GWAS based in 135,458 cases and 344,901 controls and identified 44 independent and significant loci.
- We found important relationships of genetic risk for major depression with educational attainment, body mass, and schizophrenia: lower educational attainment and higher body mass were putatively causal, whereas major depression and schizophrenia reflected a partly shared biological etiology.

Naomi R Wray,, et al., 2018

GWAS of Depression: Conclusions

- First, major depression is a brain disorder.
- Genetic results best match gene expression patterns in prefrontal and anterior cingulate cortex
- Genetic findings implicated neurons (not microglia or astrocytes),
- Genetic associations for major depression (as with schizophrenia) tend to occur in genomic regions conserved across a range of placental mammals. Conservation suggests important functional roles. Notably, our analyses did not implicate exons or coding regions.

Depression GWAS

- implicated developmental gene regulatory processes.
- Gene set analyses implicated genes containing binding sites to the protein product of *RBFOX1*, and this gene set is also significantly enriched for rare exonic variation in autism and schizophrenia
- found significant positive genetic correlations with measures of body mass and negative genetic correlations with years of education, while showing no evidence of genetic correlation with IQ.
- MR analysis results are consistent with both BMI and years of education being causal, or correlated with causal, risk factors for major depression. May be something correlated with MDD that drives the association.

Depression GWAS

- Significant positive correlations of major depression with all psychiatric disorders that we evaluated, including disorders prominent in childhood. This pattern of results indicates that the current classification scheme for major psychiatric disorders does not align well with the underlying genetic basis of these disorders.
- A shared biological basis for major depression and schizophrenia.
- Major depression is not a discrete entity at any level of analysis. Rather, our data strongly suggest the existence of biological processes common to major depression and schizophrenia (and likely, other psychiatric disorders).

GWAS study of positive selection on 870 polygenic traits

- Study of n = 512; moderns and ancients; a total of 870 traits
- 88% of these traits underwent polygenic change in the past 2,000–3,000 years
- Traits related to pigmentation, body measurement and nutritional intake exhibited strong selection signals across different time scale.
- The traits that seemed to be under selection ranged from skin traits such as "ease of tanning" to various body measurements.

Weichen Song,, et al., 2021

870 polygenic traits

- Genes associated with some undesirable traits increased in prevalence over time, including genes associated with conditions like skin cancer, inflammatory bowel disease and anorexia nervosa.
- Some of these disorders arise as side effects of genes that are beneficial for other reasons.
- If one variant elevates the risk of one disease but decreases the risk of another, natural selection would have little power to eliminate this variant,
- This is why disorders with complex genetics, such as schizophrenia or ADHD, persist despite natural selection.

GWAS

- Differences between two populations can appear genetic, when they are actually environmental.

- GWAS can't show that a gene causes a trait, only that they're associated.

- The results can get weird, fast. 1994 paper, chopstick skills are clearly not a gift of DNA: They're a matter of practice from a young age. But a GWAS study in a diverse population like San Francisco might very easily turn up evidence of genes associated with chopstick skills simply by revealing genes that are more common in East Asian populations than in European populations.

GWAS: Association can be environmental

- This mistake has actually happened.
- In the last decade, a number of papers came out claiming that height-conferring gene variants are more prevalent in Northern Europe than in Southern Europe and that natural selection was pushing Northern Europeans to become taller, on average, according to research published in 2012.
- But it turned out the impact of these genetic variants was overestimated.
- When looking at those same genetic variants in less diverse populations, the evidence for natural selection vanished. The study had been picking up on so-far-unknown environmental differences between northern and southern Europeans and mistaking them for something purely genetic.

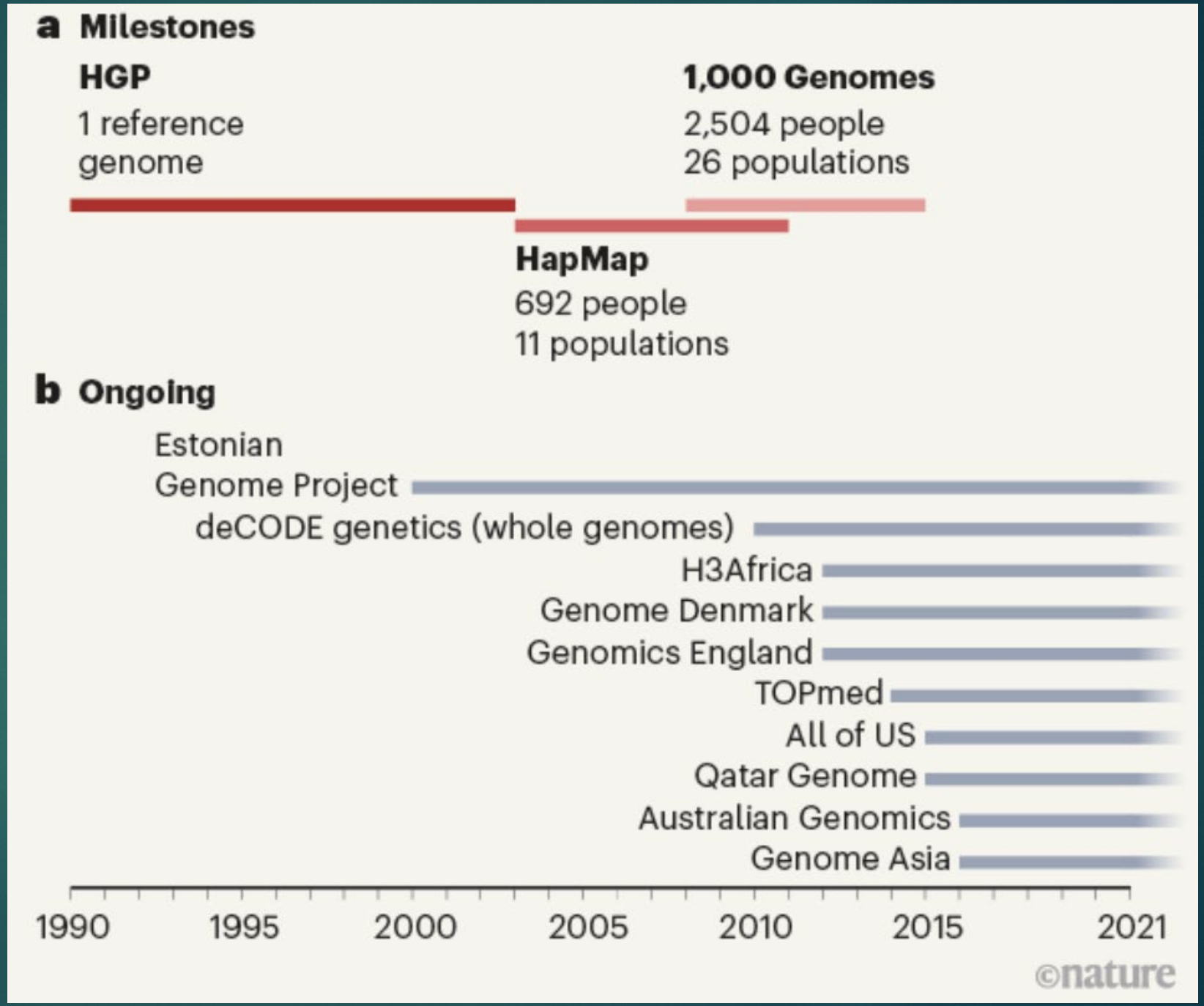
GWAS

- Higher IQ was associated with having more sexual partners but fewer children.
- ADHD and schizophrenia were both associated with having more sexual partners. These two conditions are examples of traits that might be a challenge in daily life, yet improve mating success
- When looking back over more than 100,000 years of human history, the researchers found that traits having to do with skin tone and body measurements were the most common to show selection pressure. These included things like facial measurements, height and torso length.
- For example, genes associated with face shape and size were apparently under natural selection over the past 100,000 years, which might have to do with changes to the jaw and skull associated with diet and brain growth.
- In last 3,000 years ago, inflammatory bowel disease seemed to be favored by natural selection. This could be an example of a trait that is helpful in one context and harmful in another

GWAS

- If genes tend to vary together — and many do — natural selection could be acting on a totally different trait than the one that seems most intuitive.
- For example, the variants that made skin tanning easier, which showed up as highly selected, are likely related to a lot of other traits, like rates of skin cancer, freckling and hair color

New Genomics



New Genomics

- The 1000 Genomes Project was created in 2008 to generate a more comprehensive catalogue of HGV (human genetic variants) by systematically sequencing the genomes of thousands of individuals from diverse geographical locations; now 2,504 individuals from 26 population groups on 5 continents

- Global distribution of HGV: It emerged that most common variants are shared globally, but rarer variants are shared by closely related populations, with 86% of rare variants restricted to a single continental group.

- The project also confirmed that there is greater genetic diversity in African populations than in other groups.

HGV

- Group of humans that left Africa about 100,000 years ago to populate the rest of the world carried only a subset of the variations that existed at the time in Africa; this means that the subset of HGV left behind can be studied only in Africans.

- Africa is historically under-represented in genomic studies

- Less than 2% of human genomes analyzed so far have been those of African people, despite the fact that Africa, where humans originated, contains more genetic diversity than any other continent.

New Genomics: Africa

- A 2019 study estimated that a genome representing the DNA of the African population would have about 10% more DNA than the current reference.
- Human Heredity and Health in Africa (H3Africa) consortium, reported whole-genome sequences of 426 individuals from 50 ethnolinguistic groups in Africa. H3Africa discovered more than three million new variants
- Identified 62 regions of the genome that have been evolutionarily maintained at high frequency, perhaps because of protective roles in viral immunity, DNA repair and metabolism.
- Emphasizes need to increase diversity in genome science

New genomics: Social labels

- Inferred genetic clusters might not overlap with social descriptors such as ‘Black’, ‘Latino’, ‘Asian’ and ‘European’ — an assumption that has been used by some to justify racial categorization.

- The best evidence so far suggests that social categories and genetic clusters are inconsistent.

- Identified 21 global ancestries reported that its 6,000 individuals had, on average, DNA from 4 different ancestries.

- This indicates the need for caution when using labels such as African/Black, Hispanic/Latino, Asian or European/white in genome science

New Genomics

- deCode (Iceland): certain European populations from a small pool of founders, such as Iceland's, are useful for different reasons: genetic homogeneity can help to reveal environmental factors and single-gene variations that have a strong effect.

- In 1996, the Bermuda Principles, with all parties agreeing to make the human genome sequences available in public databases, ideally within 24 hours — no delays, no exceptions.

- Now a Tower of Babel. NIH requires deposition in Database for Genotypes and Phenotypes, or dbGaP. But very difficult to use.

Genomic databases

- Most of these individual-level genomic data now live in ‘controlled-access’ databases.
- These were set up to deal with the sticky legal and ethical concerns that come with genomic data that have been linked to personal information — ‘phenotype data’ that can include health-care records, disease status or lifestyle choices.
- Even in anonymized data sets, it’s technically possible that individuals can be reidentified. So, controlled-access databases vet the researchers seeking access and ensure that the data are used only for the purposes that participants consented to.
- Then there are personal genomics companies: 23andMe, Ancestry, Family Tree DNA, Genomics, etc.

Human Genome Project

- Human Genome Project (HGP): There is no consensus on where a gene starts and ends or, surprisingly, even what sequence exactly encodes some genes

- Example: explosion of new studies after HGP:
  - gene TP53, involved in 50% of cancers; studies before HGP = 373; studies after HGP = 9824

- Intense focus on a small number of 'superstar' protein-coding genes;

- But a pivot towards non-protein-coding sections of the genome, and to understanding interactions between genetic material and proteins

New genomics: Drugs

- Before 1980 molecular and protein targets of drugs were unknown.
- The HGP changed this. Now, 100% of the targets are known for drugs licensed in the United States each year.
- Of the roughly 20,000 proteins revealed by the HGP as potential drug targets, only about 10% — 2,149 — have so far been targeted by approved drugs. That leaves 90% of the proteome untouched by pharmacology.
- The majority of successful drugs do not directly target individual disease genes. Instead, they target proteins one or two interactions away, modulating the consequences of faulty components.

New Genomics: Mendelian diseases

- The number of Mendelian diseases that have a known genetic cause went from 1,257 in 2001 to 4,377

- American College of Medical Genetics and Genomics recommends that people who have their genomes sequenced for any diagnostic purpose be informed if they carry disease-causing variants in any of 59 genes that are linked to potentially life-threatening Mendelian conditions for which pre-emptive management is available

Recommendations for improving statistical inference in population genomics

- The field of evolutionary biology is subject to misguided theoretical approaches
- A new study examines mathematical models designed to draw inferences about how evolution operates at the level of populations of organisms.
- The study concludes that such population models must be constructed with the greatest care, avoiding unwarranted initial assumptions, weighing the quality of existing knowledge and remaining open to alternate explanations.

Parul Johri, et al., 2022

Population genetics

- One of our key messages is the importance of considering the contributions of evolutionary processes certain to be in constant operation
  - such as purifying selection and genetic drift,
  - before simply relying on hypothesized or rare evolutionary processes as the primary drivers of observed population variation (such as positive selection)

Population genetics

- Population genomics arose as early efforts in the field attempted to reconcile Charles Darwin's notion of evolution by means of natural selection with the first inklings of the mechanisms of inheritance, uncovered by the Augustinian monk, Gregor Mendel.
- The synthesis culminated in the 1920s and early 30s, largely thanks to the mathematical work of Fisher, Haldane and Wright, who were the first to explore how natural selection together with other evolutionary forces would modify the genetic composition of Mendelian populations over time.
- Today, studies in population genomics involve the large-scale application of various genomic technologies to explore the genetic composition of biological populations, and how various factors, including natural selection and genetic drift, produce changes in genetic composition over time.

Population genetics

- To accomplish this, population geneticists develop mathematical models quantifying the contributions of these evolutionary processes in shaping gene frequencies, use this theory to design statistical inference approaches for estimating the forces producing observed patterns of genetic variation in actual populations, and test their conclusions against accumulated data.
- Such variation in the human genome can take several forms. One common source of variation is known as single nucleotide polymorphisms, or SNPs, where a single DNA letter in the genome is altered.
- But larger-scale variation in the genome, involving the simultaneous alteration of hundreds or even thousands of base pairs is also possible.

Population genetics

- Natural selection may occur when different variants segregating in a population have a fitness differential relative to one another.
- By designing and studying mathematical models governing the corresponding gene frequency change and applying those models to empirical data, population geneticists seek to understand the contributing evolutionary processes in a rigorous, quantitative way.
- Thus, population genetics is often regarded as the theoretical cornerstone of modern Darwinian evolution.

Population genetics: purifying selection vs genetic drift

- Although the importance of natural selection to the evolutionary process is undeniable, the role of positive selection in increasing the frequency of beneficial variants—the potential driver of adaptation—is certain to be comparatively rare relative even to other forms of natural selection.
- For example, purifying selection—the removal of deleterious variants from the population—is a constantly acting and far more pervasive form of selection.
- In addition, there are multiple non-selective evolutionary processes of great importance.

Population Genetics

- For example, genetic drift describes the many chance genetic fluctuations/mutations inherent to evolution.
- In large populations, natural selection may act more efficiently in purging deleterious variation and potentially fixing beneficial variation
- In small populations, genetic drift will be increasingly dominant.

Population genetics: think genetic drift

- According to the Neutral Theory of Molecular Evolution, most evolutionary changes at the molecular level in real populations are governed not by natural selection, but by genetic drift, (random mutational fluctuations in the frequency of a particular version of a gene in a population).
- The study emphasizes that this critical point is too often missed by evolutionary biologists.
- Natural selection is just one of several evolutionary mechanisms, and the failure to realize this is probably the most significant impediment to a fruitful integration of evolutionary theory with molecular, cellular, and developmental biology.

*** Possible lecture

- Next time: David Reich's
- Pigmentation
- High altitude/hypoxic adaptation
- Lactase tolerance
- FOXP2 Gene
- CRISPR