Ancient DNA

Two parts

1. Background and problems

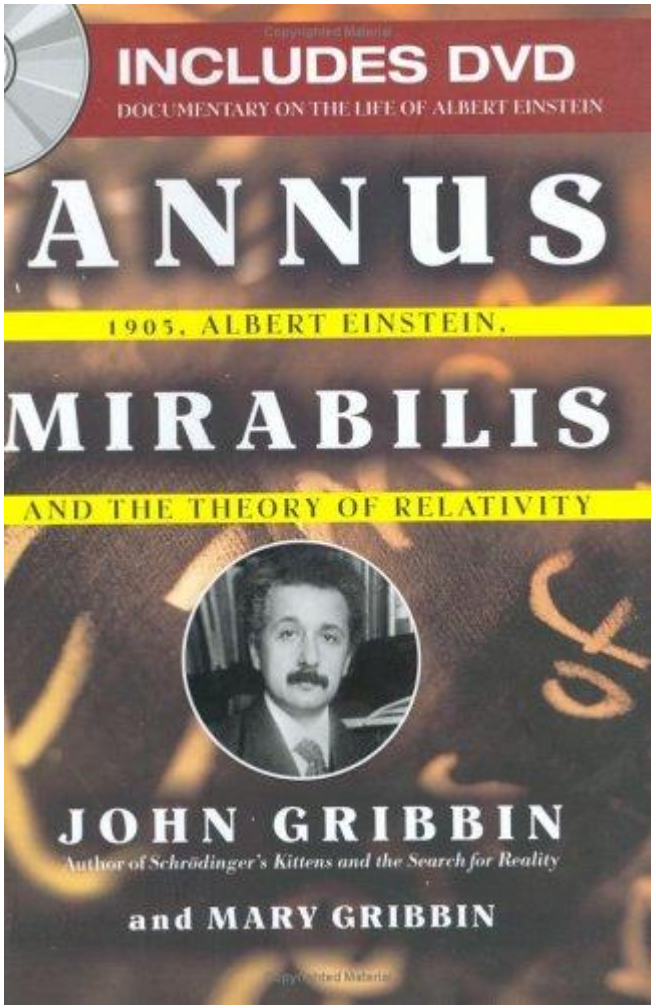
2. Examples

Part 1: What it is not about

What it is about: Genome sequencing

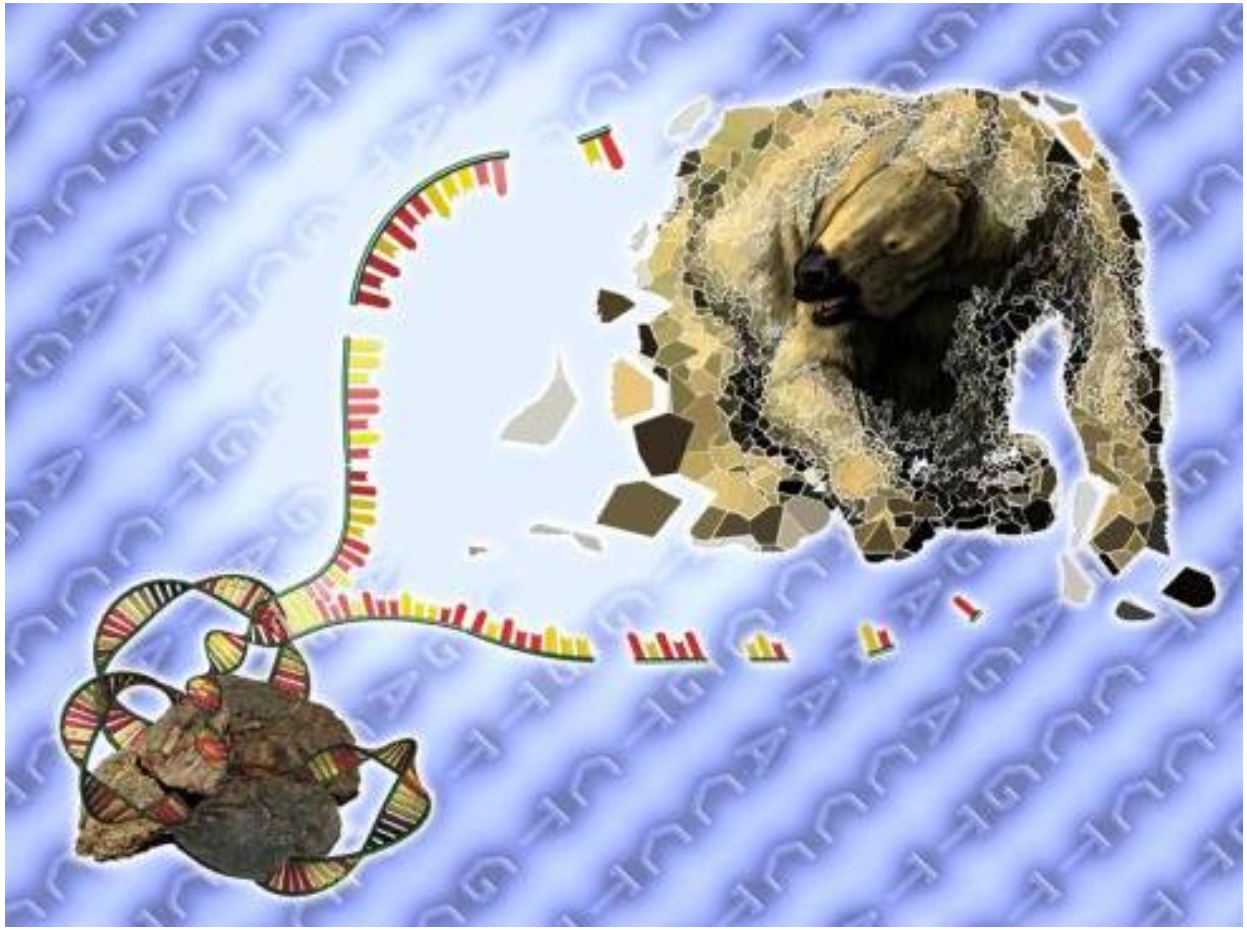
Before 2005: small data sets, limited scope

Annus mirabilis 1905



Annus mirabilis 2005: slightly less important

Birth of paleogenomics



Ancient DNA's annus mirabilis 2005

2nd June 2005

Genomic Sequencing of Pleistocene Cave Bears

James P. Noonan,[1,2] Michael Hofreiter,[3] Doug Smith,[1] James R. Priest,[2] Nadin Rohland,[3] Gernot Rabeder,[4] Johannes Krause,[3] J. Chris Detter,[1,5] Svante Pääbo,[3] Edward M. Rubin[1,2*]

Despite the greater information content of genomic DNA, ancient DNA studies have largely been limited to the amplification of mitochondrial sequences. Here we describe metagenomic libraries constructed with unamplified DNA extracted from skeletal remains of two 40,000-year-old extinct cave bears. Analysis of ~1 megabase of sequence from each library showed that despite significant microbial contamination, 5.8 and 1.1% of clones contained cave bear inserts, yielding 26,861 base pairs of cave bear genome sequence. Comparison of cave bear and modern bear sequences revealed the evolutionary relationship of these lineages. The metagenomic approach used here establishes the feasibility of ancient DNA genome sequencing programs.

Ancient DNA's annus mirabilis 2005

21st July 2005

Opinion *TRENDS in Ecology and Evolution* Vol.20 No.10 October 2005

Assessing ancient DNA studies

M. Thomas P. Gilbert[1], Hans-Jürgen Bandelt[2], Michael Hofreiter[3] and Ian Barnes[4]

[1]Ecology and Evolutionary Biology, The University of Arizona, 1041 E. Lowell St, Tucson, AZ 85721, USA
[2]Department of Mathematics, University of Hamburg, Bundesstr. 55, 20146 Hamburg, Germany
[3]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany
[4]The Center for Genetic Anthropology, Department of Biology, Darwin Building, University College London, Gower Street, London, UK, WC1E 6BT

Ancient DNA's annus mirabilis 2005

31st July 2005

nature Vol 437|15 September 2005|doi:10.1038/nature03959

ARTICLES

Genome sequencing in microfabricated high-density picolitre reactors

Marcel Margulies[1]*, Michael Egholm[1]*, William E. Altman[1], Said Attiya[1], Joel S. Bader[1], Lisa A. Bemben[1], Jan Berka[1], Michael S. Braverman[1], Yi-Ju Chen[1], Zhoutao Chen[1], Scott B. Dewell[1], Lei Du[1], Joseph M. Fierro[1], Xavier V. Gomes[1], Brian C. Godwin[1], Wen He[1], Scott Helgesen[1], Chun He Ho[1], Gerard P. Irzyk[1], Szilveszter C. Jando[1], Maria L. I. Alenquer[1], Thomas P. Jarvie[1], Kshama B. Jirage[1], Jong-Bum Kim[1], James R. Knight[1], Janna R. Lanza[1], John H. Leamon[1], Steven M. Lefkowitz[1], Ming Lei[1], Jing Li[1], Kenton L. Lohman[1], Hong Lu[1], Vinod B. Makhijani[1], Keith E. McDade[1], Michael P. McKenna[1], Eugene W. Myers[2], Elizabeth Nickerson[1], John R. Nobile[1], Ramona Plant[1], Bernard P. Puc[1], Michael T. Ronan[1], George T. Roth[1], Gary J. Sarkis[1], Jan Fredrik Simons[1], John W. Simpson[1], Maithreyan Srinivasan[1], Karrie R. Tartaro[1], Alexander Tomasz[3], Kari A. Vogt[1], Greg A. Volkmer[1], Shally H. Wang[1], Yong Wang[1], Michael P. Weiner[4], Pengguang Yu[1], Richard F. Begley[1] & Jonathan M. Rothberg[1]

The proliferation of large-scale DNA-sequencing projects in recent years has driven a search for alternative methods to reduce time and cost. Here we describe a scalable, highly parallel sequencing system with raw throughput significantly greater than that of state-of-the-art capillary electrophoresis instruments. The apparatus uses a novel fibre-optic slide of individual wells and is able to sequence 25 million bases, at 99% or better accuracy, in one four-hour run. To achieve an approximately 100-fold increase in throughput over current Sanger sequencing technology, we have developed an emulsion method for DNA amplification and an instrument for sequencing by synthesis using a pyrosequencing protocol optimized for solid support and picolitre-scale volumes. Here we show the utility, throughput, accuracy and robustness of this system by shotgun sequencing and *de novo* assembly of the *Mycoplasma genitalium* genome with 96% coverage at 99.96% accuracy in one run of the machine.

454 Life Sciences: 20 Mb sequence data

Ancient DNA's annus mirabilis 2005

11th November 2005

Ancient DNA from the First European Farmers in 7500-Year-Old Neolithic Sites

Wolfgang Haak,[1]* Peter Forster,[2] Barbara Bramanti,[1] Shuichi Matsumura,[2] Guido Brandt,[1] Marc Tänzer,[1] Richard Villems,[3] Colin Renfrew,[2] Detlef Gronenborn,[4] Kurt Werner Alt,[1] Joachim Burger[1]

The ancestry of modern Europeans is a subject of debate among geneticists, archaeologists, and anthropologists. A crucial question is the extent to which Europeans are descended from the first European farmers in the Neolithic Age 7500 years ago or from Paleolithic hunter-gatherers who were present in Europe since 40,000 years ago. Here we present an analysis of ancient DNA from early European farmers. We successfully extracted and sequenced intact stretches of maternally inherited mitochondrial DNA (mtDNA) from 24 out of 57 Neolithic skeletons from various locations in Germany, Austria, and Hungary. We found that 25% of the Neolithic farmers had one characteristic mtDNA type and that this type formerly was widespread among Neolithic farmers in Central Europe. Europeans today have a 150-times lower frequency (0.2%) of this mtDNA type, revealing that these first Neolithic farmers did not have a strong genetic influence on modern European female lineages. Our finding lends weight to a proposed Paleolithic ancestry for modern Europeans.

Ancient DNA's annus mirabilis 2005

18th December 2005

Ancient DNA's annus mirabilis 2005

20th December 2005

REPORTS

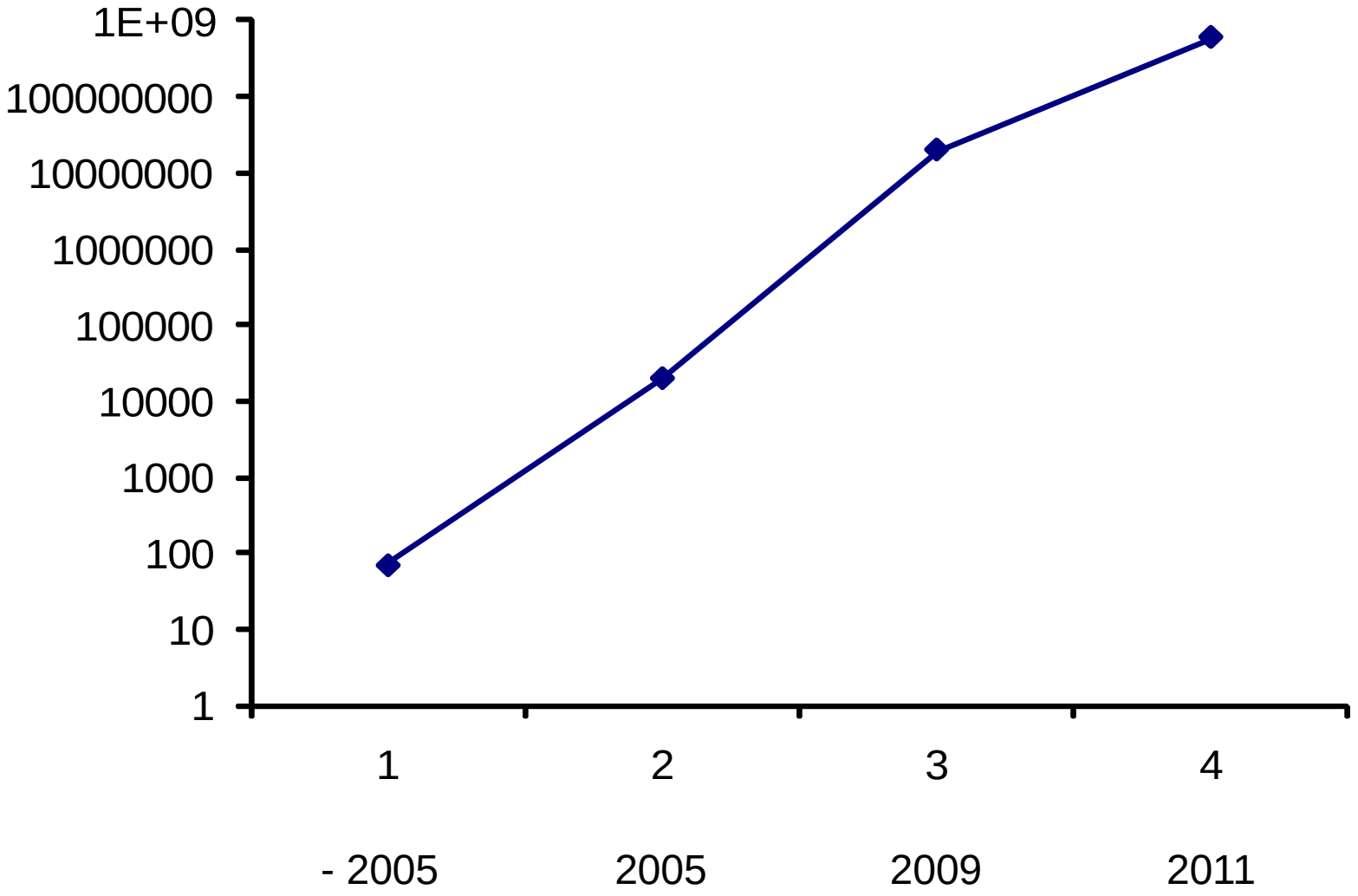
Metagenomics to Paleogenomics: Large-Scale Sequencing of Mammoth DNA

Hendrik N. Poinar,[1,2,3*] Carsten Schwarz,[1,2] Ji Qi,[4] Beth Shapiro,[5] Ross D. E. MacPhee,[6] Bernard Buigues,[7] Alexei Tikhonov,[8] Daniel H. Huson,[9] Lynn P. Tomsho,[4] Alexander Auch,[9] Markus Rampp,[10] Webb Miller,[4] Stephan C. Schuster[4*]

We sequenced 28 million base pairs of DNA in a metagenomics approach, using a woolly mammoth (*Mammuthus primigenius*) sample from Siberia. As a result of exceptional sample preservation and the use of a recently developed emulsion polymerase chain reaction and pyrosequencing technique, 13 million base pairs (45.4%) of the sequencing reads were identified as mammoth DNA. Sequence identity between our data and African elephant (*Loxodonta africana*) was 98.55%, consistent with a paleontologically based divergence date of 5 to 6 million years. The sample includes a surprisingly small diversity of environmental DNAs. The high percentage of endogenous DNA recoverable from this single mammoth would allow for completion of its genome, unleashing the field of paleogenomics.

From 454 to Illumina sequencing

Sequencing yield per machine run in kilobases



Sequencing yield per machine run in kilobases

.

Sanger sequencing	454 sequencing	Illumina sequencing
until 2005	2005	2009

Sequencing yield per machine run in kilobases

Sanger sequencing
until 2005

454 sequencing
2005

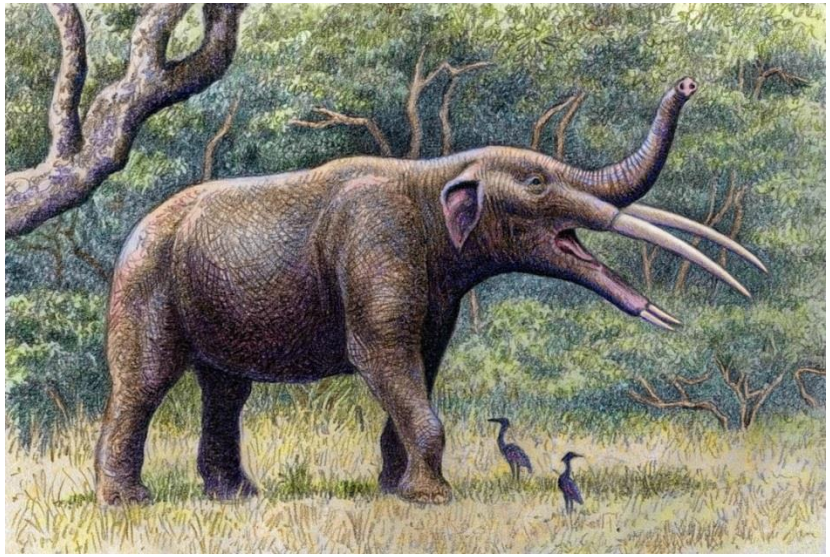
Illumina sequencing
2009

Paleogenomics: Neanderthals and human evolution

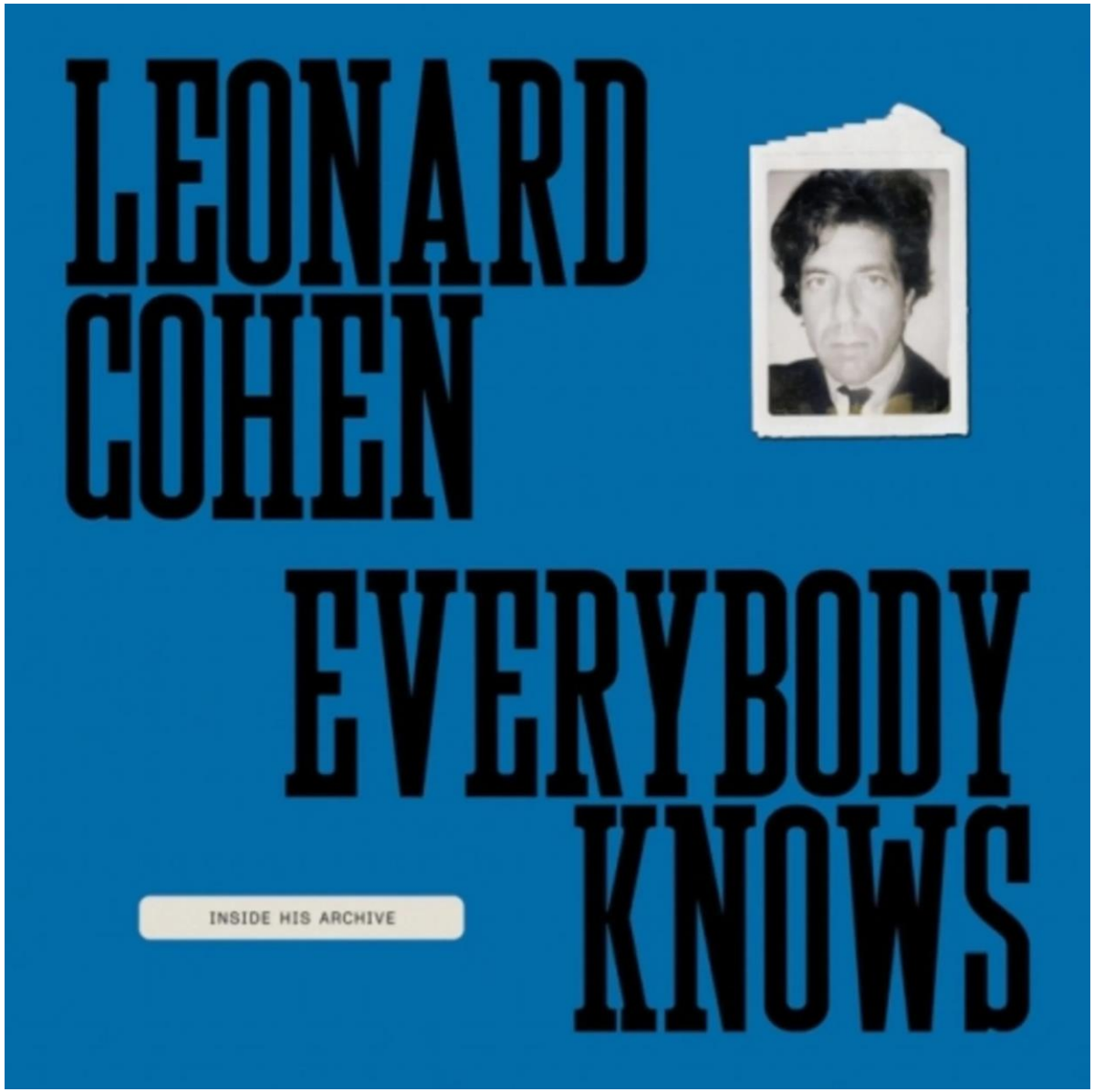
But I won’t talk about this

It’s not my job

There are much more exciting species!



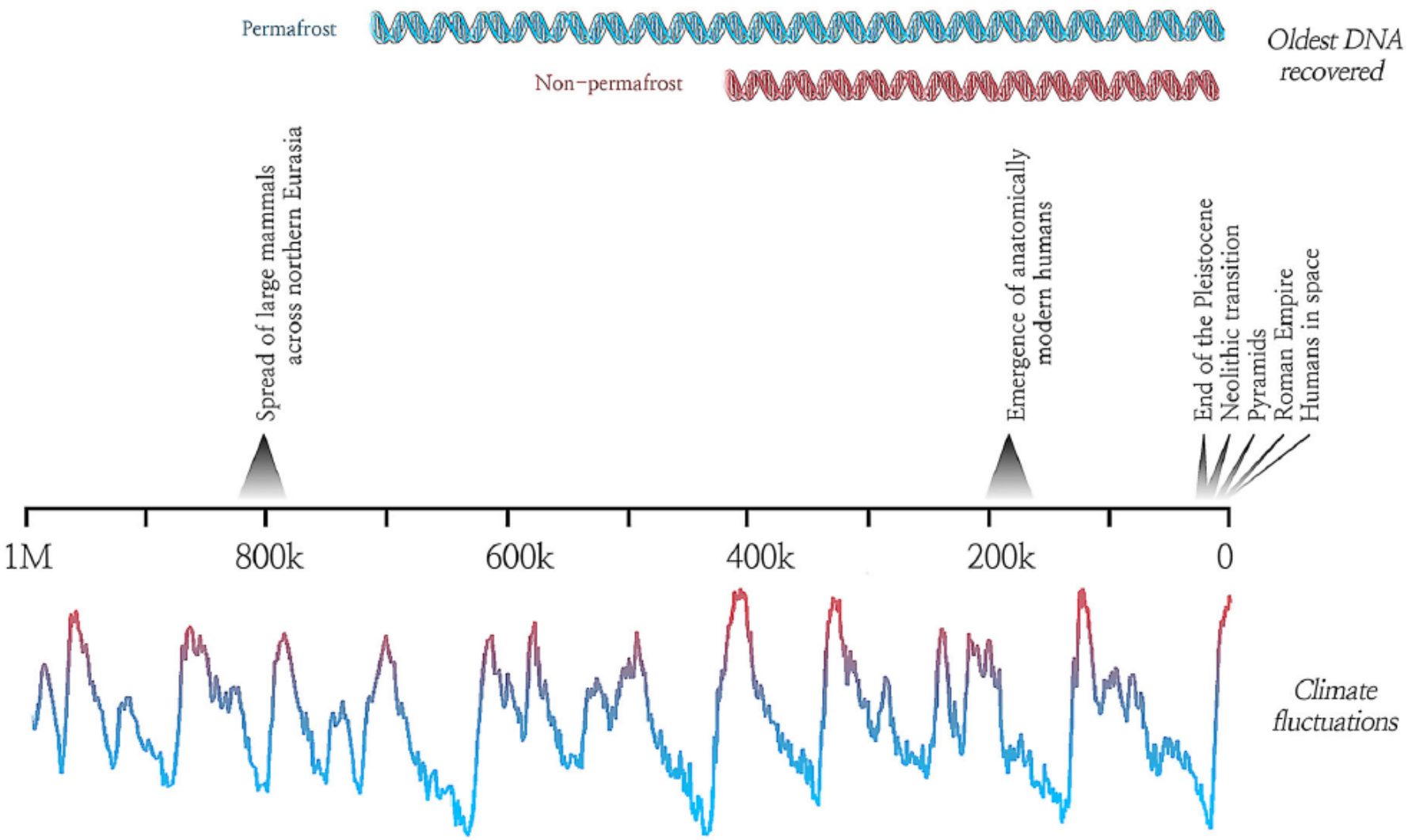
And anyway



So, what is ancient DNA?

1. No clear definition or age limit.
2. Broadly speaking, DNA from samples that have not been stored with the explicit aim of preserving DNA.
3. Sometimes distinction between museum (< 200 years), historical (< 500 years) and “true” ancient DNA (> 500 years).

And how old does it get??



Older……

Article

Million-year-old DNA sheds light on the genomic history of mammoths

https://doi.org/10.1038/s41586-021-03224-9

Received: 3 July 2020

Accepted: 11 January 2021

Published online: 17 February 2021

Check for updates

Tom van der Valk[1,2,3,17]✉, Patrícia Pečnerová[2,4,5,17], David Díez-del-Molino[1,2,4,17], Anders Bergström[6], Jonas Oppenheimer[7], Stefanie Hartmann[8], Georgios Xenikoudakis[8], Jessica A. Thomas[8], Marianne Dehasque[1,2,4], Ekin Sağlıcan[9], Fatma Rabia Fidan[9], Ian Barnes[10], Shanlin Liu[11], Mehmet Somel[9], Peter D. Heintzman[12], Pavel Nikolskiy[13], Beth Shapiro[14,15], Pontus Skoglund[6], Michael Hofreiter[8], Adrian M. Lister[10], Anders Götherström[1,16,18] & Love Dalén[1,2,4,18]✉

.....and older

Article

A 2-million-year-old ecosystem in Greenland uncovered by environmental DNA

https://doi.org/10.1038/s41586-022-05453-y

Received: 30 September 2021

Accepted: 18 October 2022

Published online: 7 December 2022

Open access

Check for updates

Kurt H. Kjær[1,27]✉, Mikkel Winther Pedersen[1,27], Bianca De Sanctis[2,3], Binia De Cahsan[4], Thorfinn S. Korneliussen[1], Christian S. Michelsen[1,5], Karina K. Sand[1], Stanislav Jelavić[1,6], Anthony H. Ruter[1], Astrid M. A. Schmidt[7,8], Kristian K. Kjeldsen[9], Alexey S. Tesakov[10], Ian Snowball[11], John C. Gosse[12], Inger G. Alsos[13], Yucheng Wang[1,2], Christoph Dockter[14], Magnus Rasmussen[14], Morten E. Jørgensen[14], Birgitte Skadhauge[14], Ana Prohaska[1,2], Jeppe Å. Kristensen[15,16], Morten Bjerager[17], Morten E. Allentoft[1,18], Eric Coissac[13,19], PhyloNorway Consortium*,**, Alexandra Rouillard[1,21], Alexandra Simakova[10], Antonio Fernandez-Guerra[1], Chris Bowler[20], Marc Macias-Fauria[22], Lasse Vinner[1], John J. Welch[3], Alan J. Hidy[23], Martin Sikora[1], Matthew J. Collins[24,25], Richard Durbin[3], Nicolaj K. Larsen[1] & Eske Willerslev[1,2,26]✉

Late Pliocene and Early Pleistocene epochs 3.6 to 0.8 million years ago[1] had climates resembling those forecasted under future warming[2]. Palaeoclimatic records show strong polar amplification with mean annual temperatures of 11–19 °C above contemporary values[3,4]. The biological communities inhabiting the Arctic during this time remain poorly known because fossils are rare[5]. Here we report an ancient environmental DNA[6] (eDNA) record describing the rich plant and animal assemblages of the Kap København Formation in North Greenland, dated to around two million years ago. The record shows an open boreal forest ecosystem with mixed vegetation

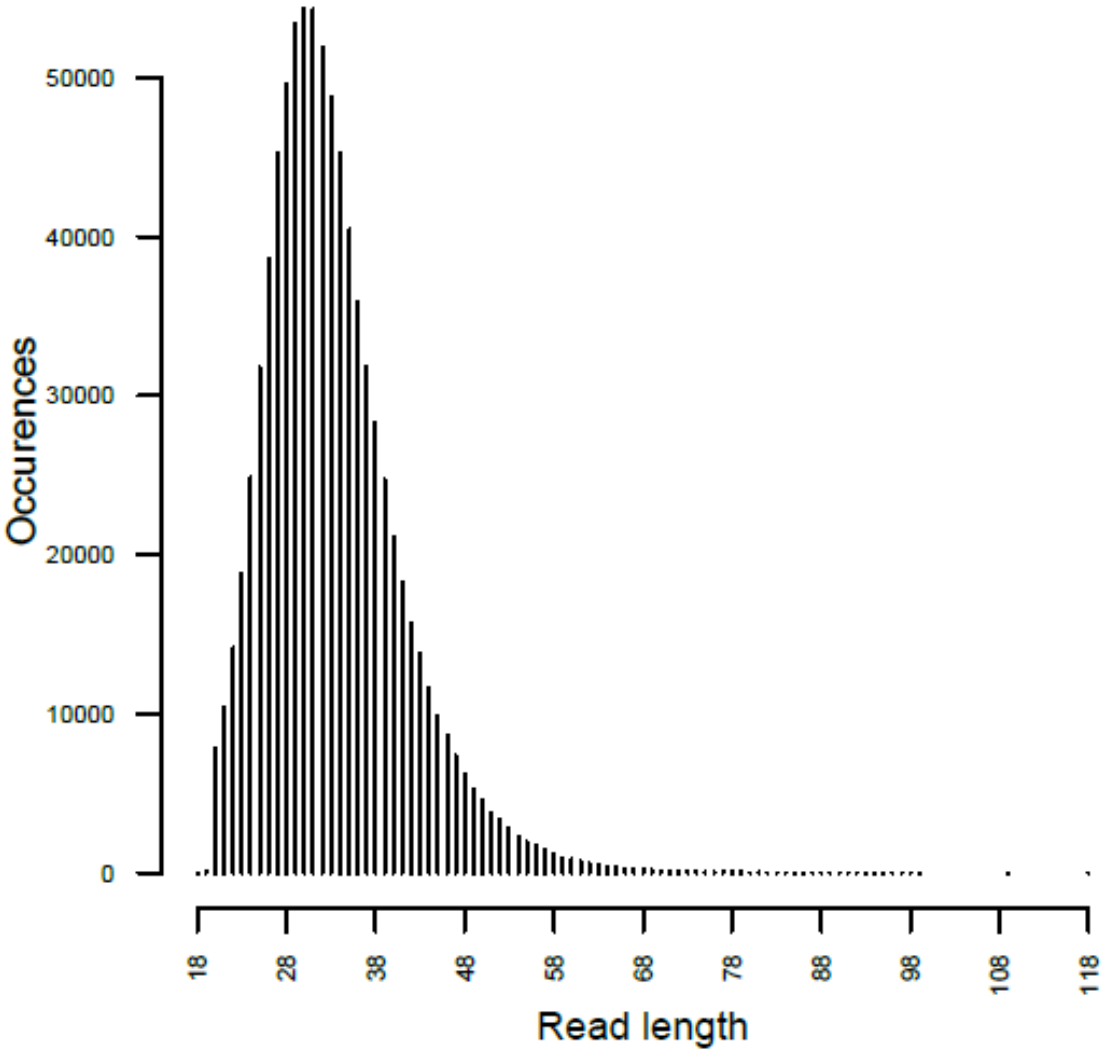
What characterizes ancient DNA?

1. Generally small amounts of DNA – kind of
2. Contamination – sometimes
3. Short fragments – largely true
4. Miscoding lesions – generally true

Properties of DNA

Properties of ancient DNA

Thus, ancient DNA is short



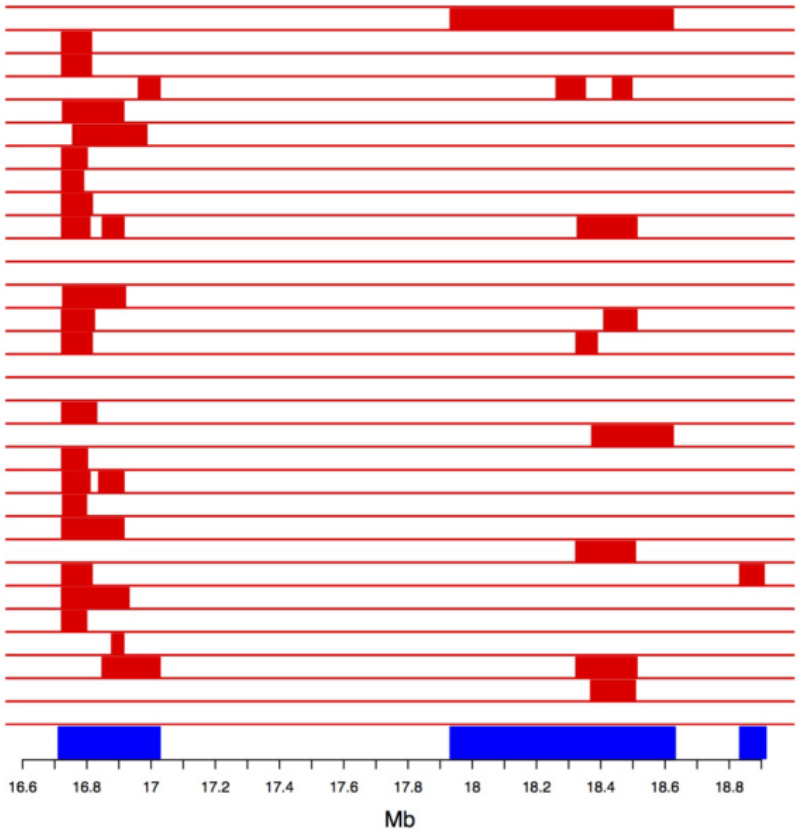
Unless it's found in modern species

LETTER

doi:10.1038/nature12961

The genomic landscape of Neanderthal ancestry in present-day humans

Sriram Sankararaman[1,2], Swapan Mallick[1,2], Michael Dannemann[3], Kay Prüfer[3], Janet Kelso[3], Svante Pääbo[3], Nick Patterson[1,2] & David Reich[1,2,4]



1. Present in low amounts?

fine powder only yielded trace amounts of DNA. In contrast, the coarse powder

Rohland & Hofreiter 2007a

Even in a well-preserved sample, aDNA is generally only present in small amounts, and the number of samples

Hofreiter et al 2013

of ancient DNA. It accounts for the fact that ancient DNA is usually highly fragmented and present only in trace amounts.

Gansauge & Meyer 2013

DNA sequences to be determined (3). Third, because ancient DNA is present in low amounts or absent in many specimens,

Briggs et al. 2007

General considerations about handling The samples used for extraction will usually contain only trace amounts of DNA. To avoid cross-contamination

Stiller et al. 2009

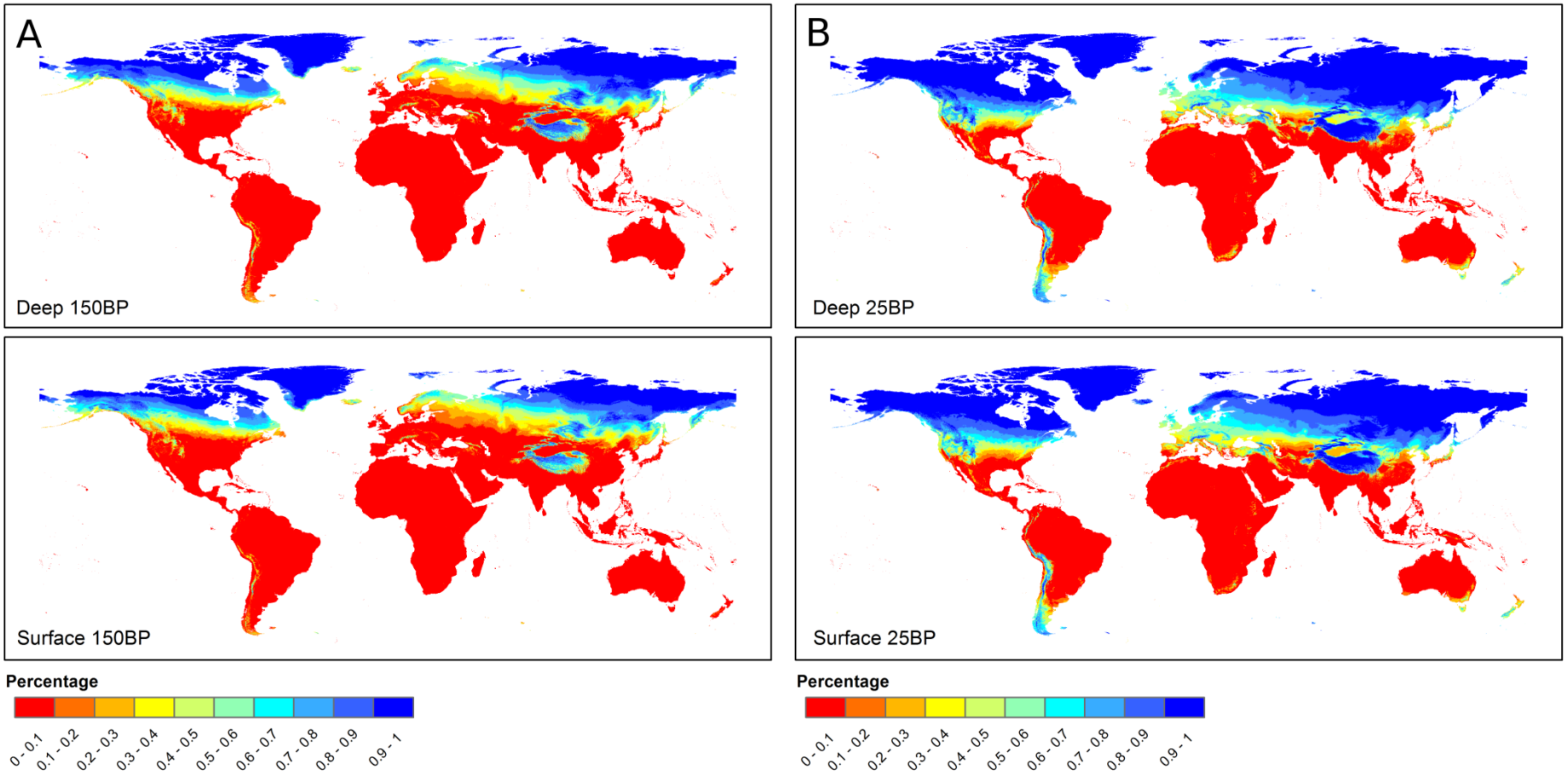
Rohland & Hofreiter 2007b

to observe evolution in "real time." However, the DNA isolated from such material, commonly referred to as "ancient DNA," is typically present in low amounts, heavily fragmented, chemically

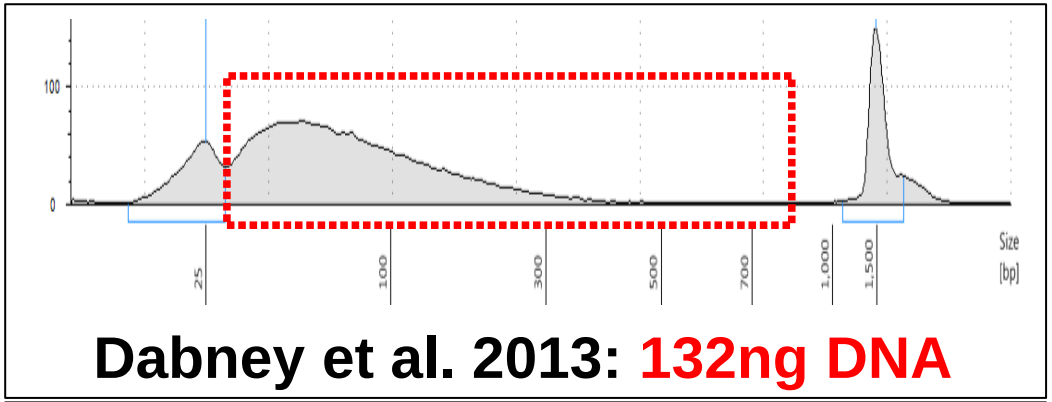
Despite these advances, sequencing ancient human DNA continues to be challenging for several reasons. First, only trace amounts of highly fragmented DNA are usually preserved in an-

Gansauge & Meyer 2014

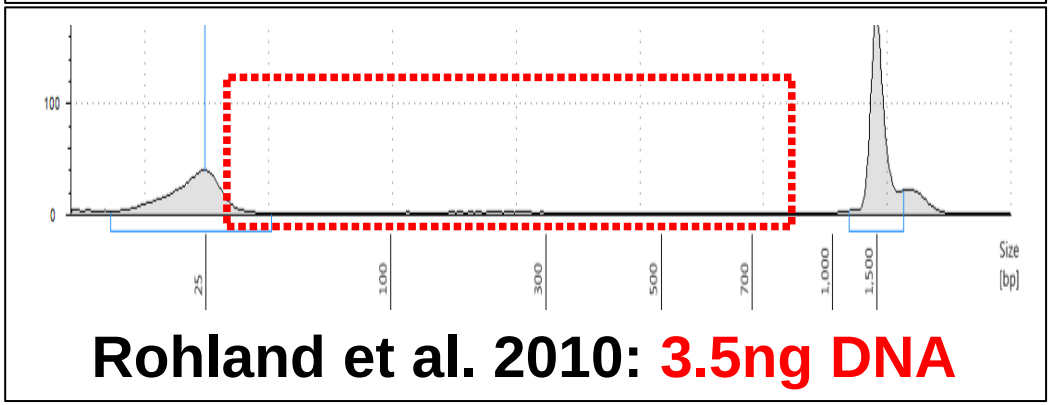
Temperature and DNA preservation



Partially a problem of DNA extraction

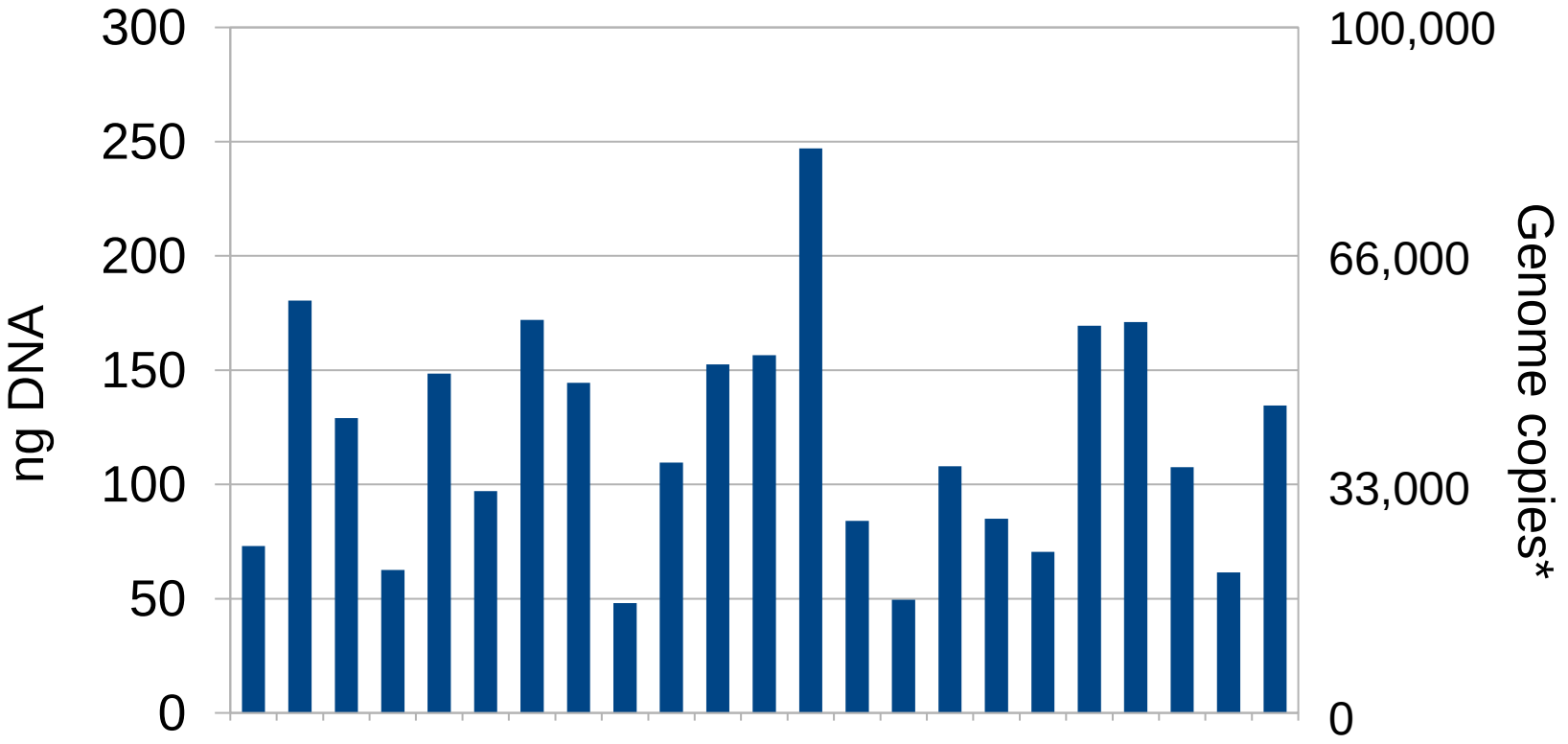


Dabney et al. 2013: 132ng DNA



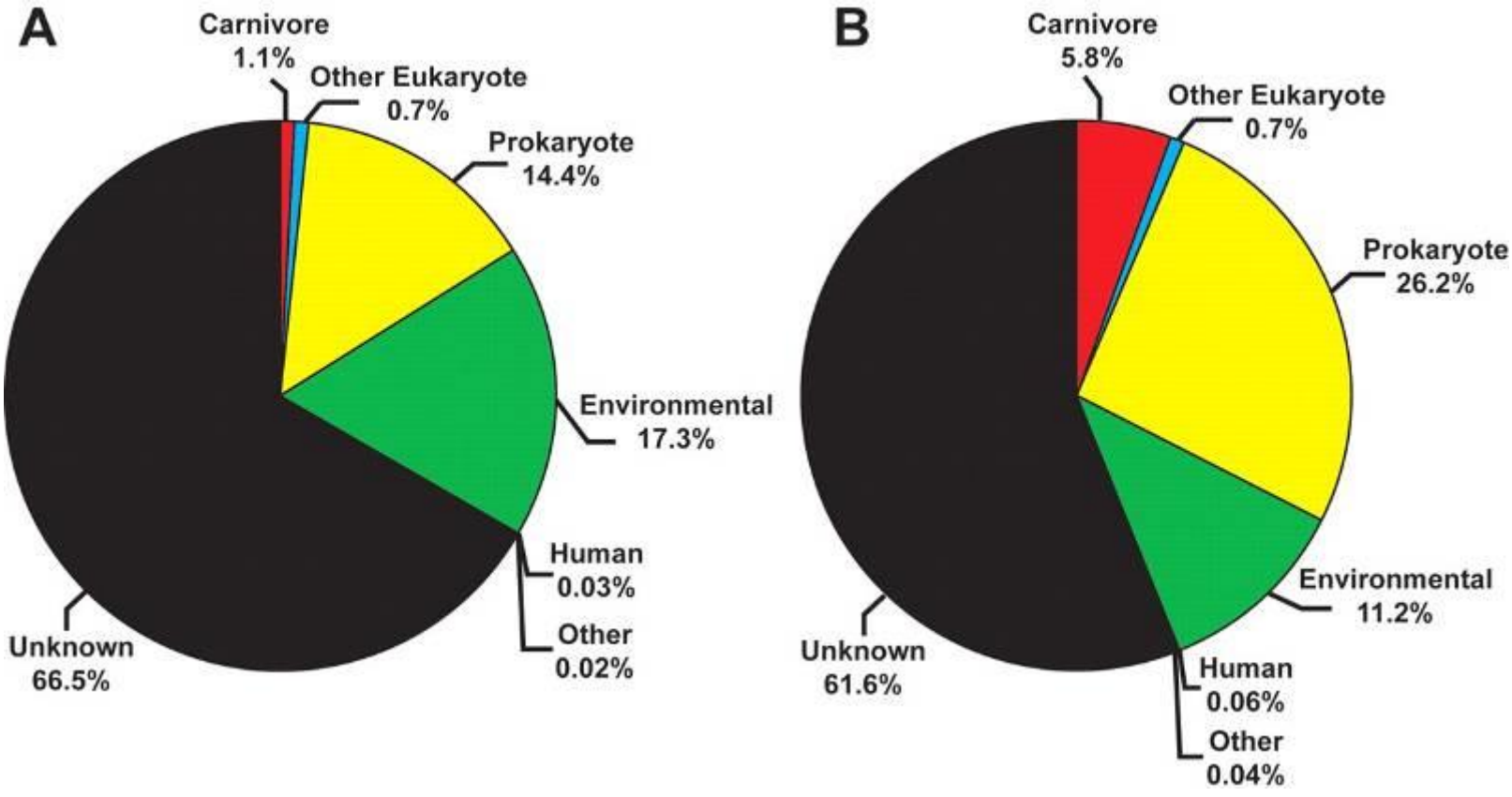
Rohland et al. 2010: 3.5ng DNA

DNA from 50 mg cave bear bone powder



*** assuming 100% cave bear DNA**

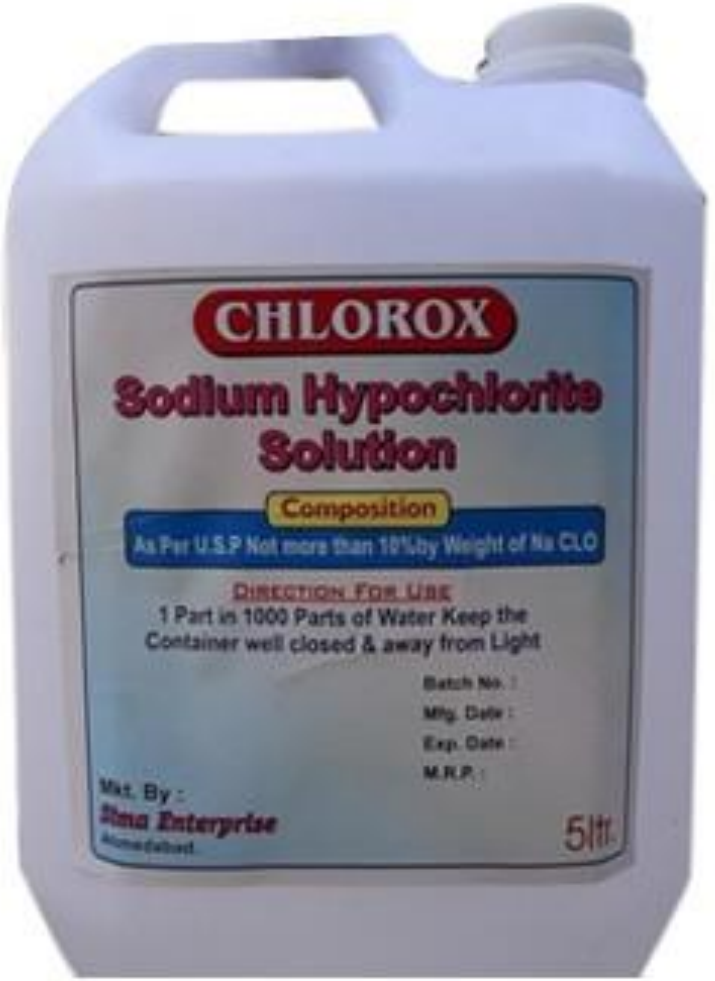
But often little endogenous DNA



What Do I Do Now

Solution 1: Get rich

Solution 2: Bleach your samples



Solution 3: Target your DNA

Solution 4: Luck

RESEARCH ARTICLES

A High-Coverage Genome Sequence from an Archaic Denisovan Individual

Matthias Meyer,[1*‡] Martin Kircher,[1*†] Marie-Theres Gansauge,[1] Heng Li,[2] Fernando Racimo,[1] Swapan Mallick,[2,3] Joshua G. Schraiber,[4] Flora Jay,[4] Kay Prüfer,[1] Cesare de Filippo,[1] Peter H. Sudmant,[6] Can Alkan,[5,6] Qiaomei Fu,[1,7] Ron Do,[2] Nadin Rohland,[2,3] Arti Tandon,[2,3] Michael Siebauer,[1] Richard E. Green,[8] Katarzyna Bryc,[3] Adrian W. Briggs,[3] Udo Stenzel,[1] Jesse Dabney,[1] Jay Shendure,[6] Jacob Kitzman,[6] Michael F. Hammer,[9] Michael V. Shunkov,[10] Anatoli P. Derevianko,[10] Nick Patterson,[2] Aida M. Andrés,[1] Evan E. Eichler,[6,11] Montgomery Slatkin,[4] David Reich,[2,3‡] Janet Kelso,[1] Svante Pääbo[1‡]

contamination. The fraction of hominin endogenous DNA is commonly smaller than 1% and rarely approaches 5% (*1*, *7*), which makes shotgun sequencing of the entire genome economically and logistically impractical. The only known exception is the Denisovan phalanx, which contains ~70% endogenous DNA. However, an

Solution 5:

Solution 5: choose wisely

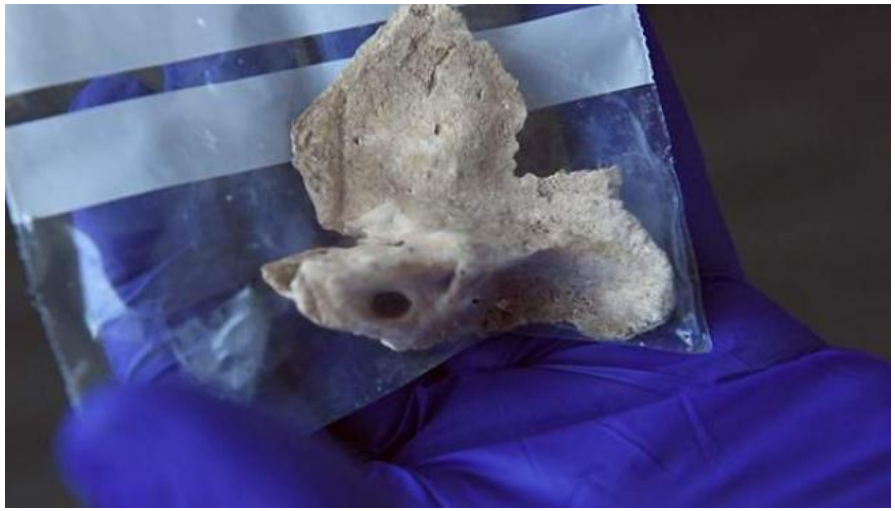
ARTICLE

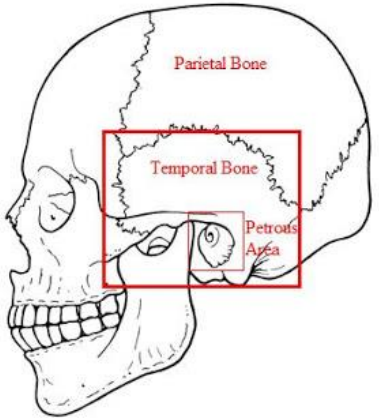
Received 4 Apr 2014 | Accepted 11 Sep 2014 | Published 21 Oct 2014

DOI: 10.1038/ncomms6257 OPEN

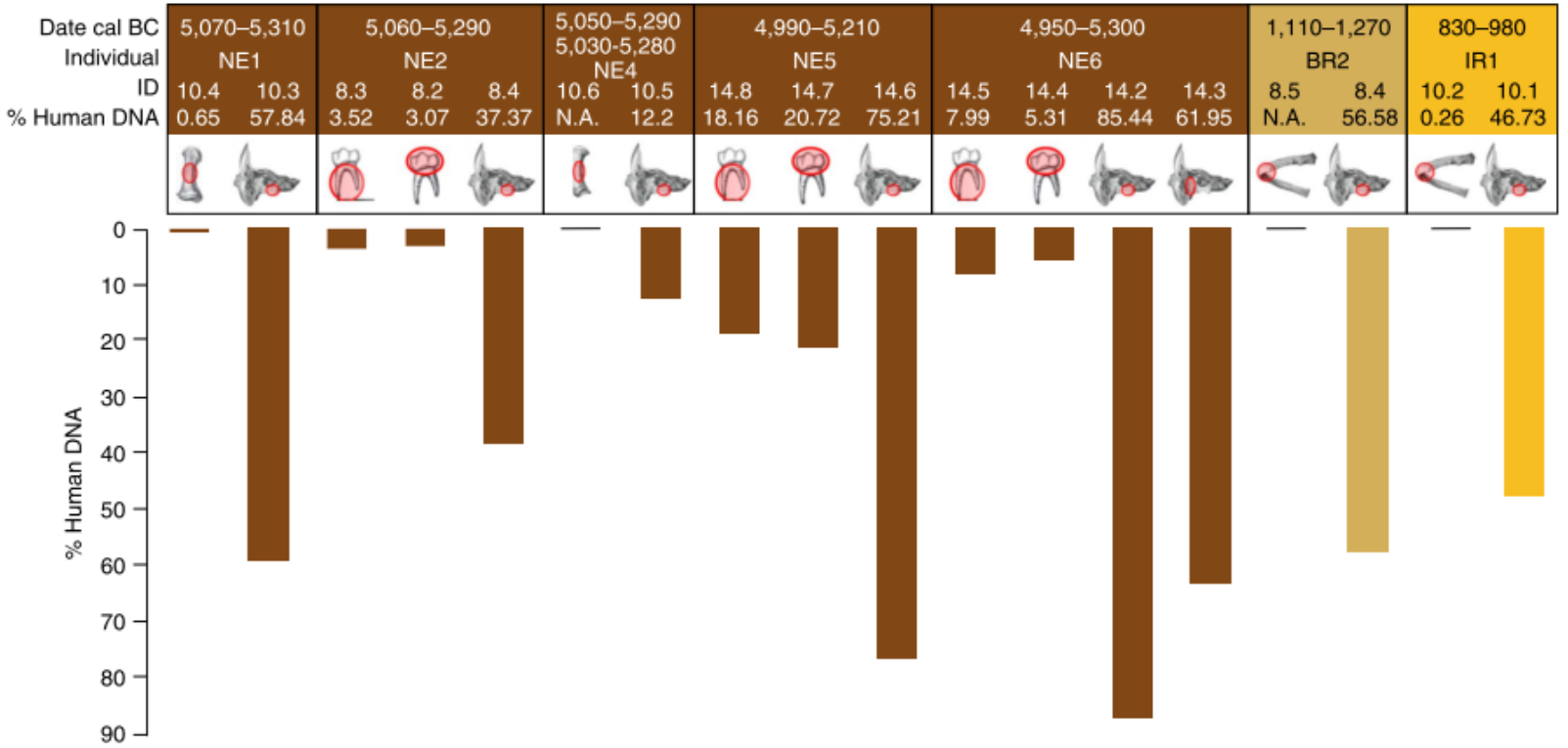
Genome flux and stasis in a five millennium transect of European prehistory

Cristina Gamba[1,2,3], Eppie R. Jones[3], Matthew D. Teasdale[3], Russell L. McLaughlin[3], Gloria Gonzalez-Fortes[4], Valeria Mattiangeli[3], László Domboróczki[5], Ivett Kővári[6], Ildikó Pap[7], Alexandra Anders[8], Alasdair Whittle[9], János Dani[10], Pál Raczky[8], Thomas F.G. Higham[11], Michael Hofreiter[4], Daniel G. Bradley[3,*] & Ron Pinhasi[1,2,3,12,*]

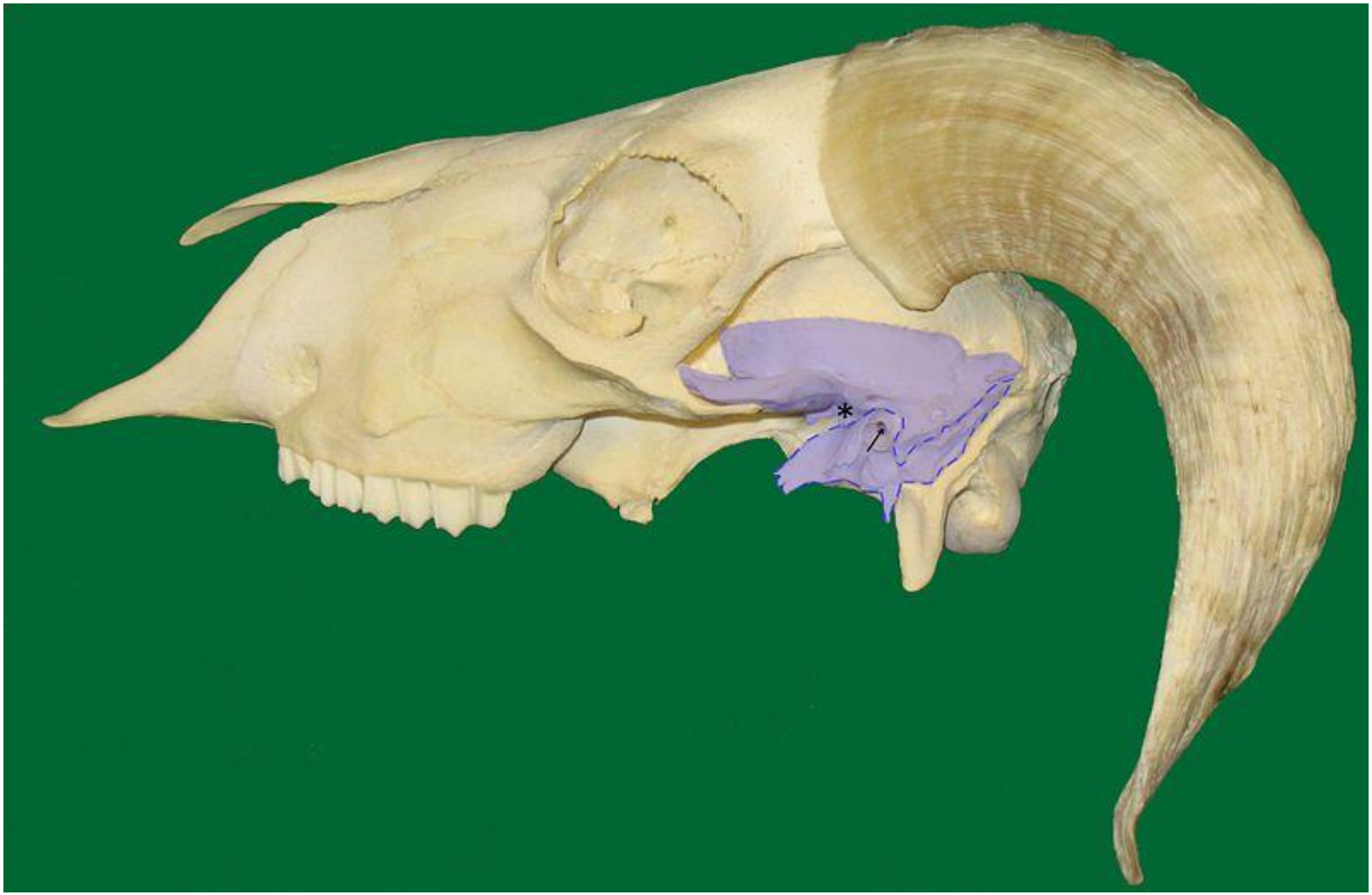




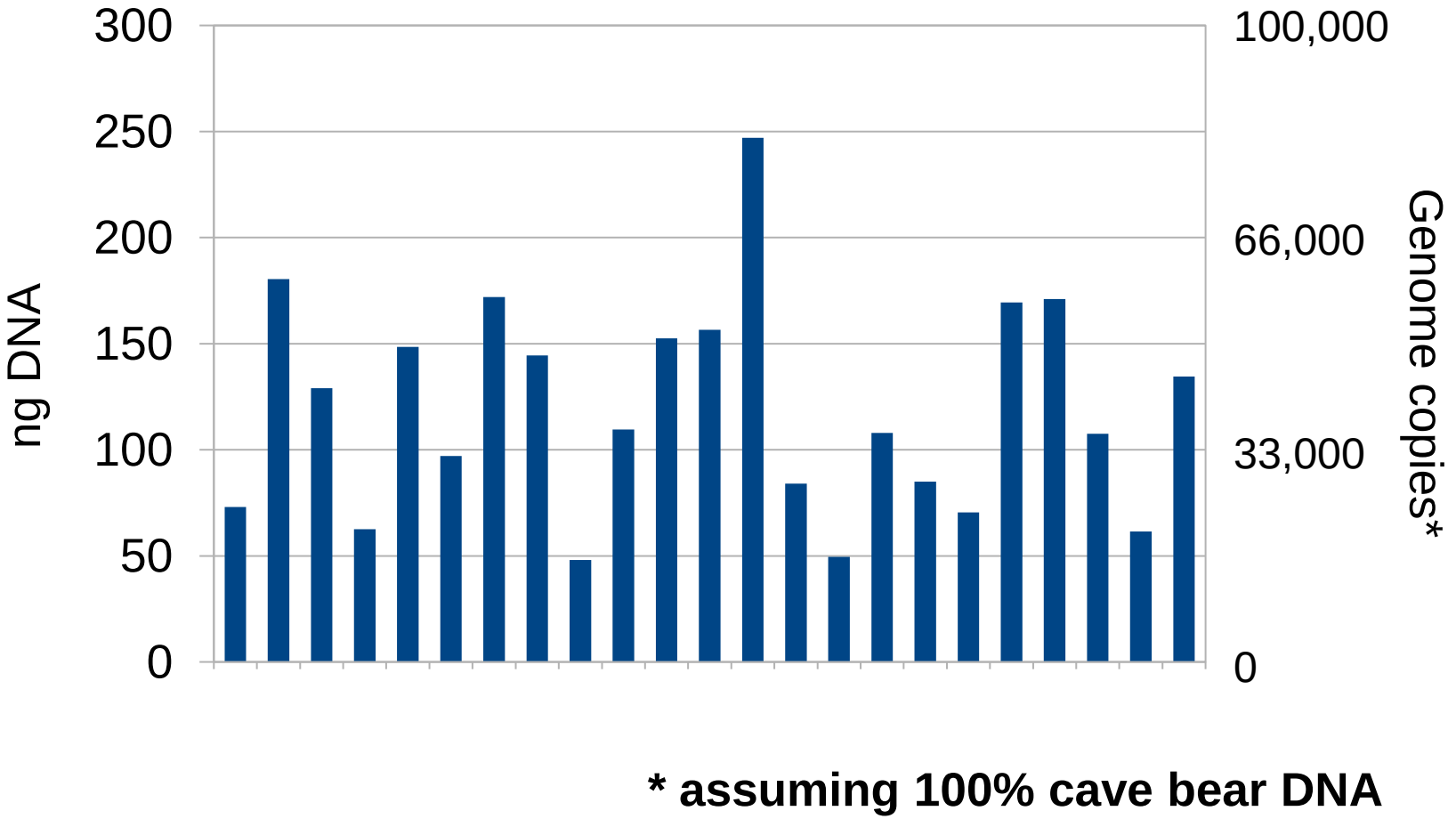
Choose the right bones



The petrous bone

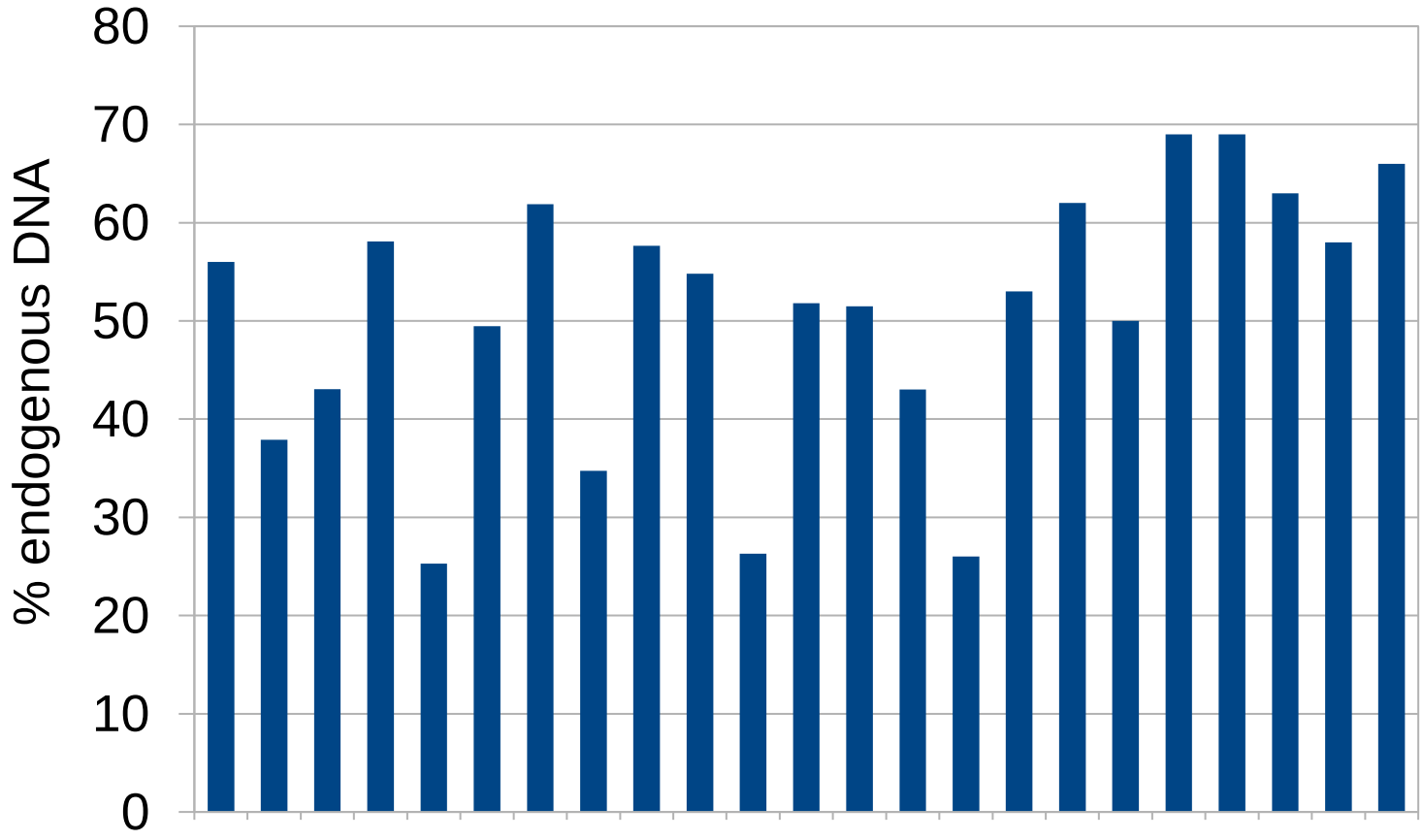


DNA from 50 mg cave bear **petrous** bone powder

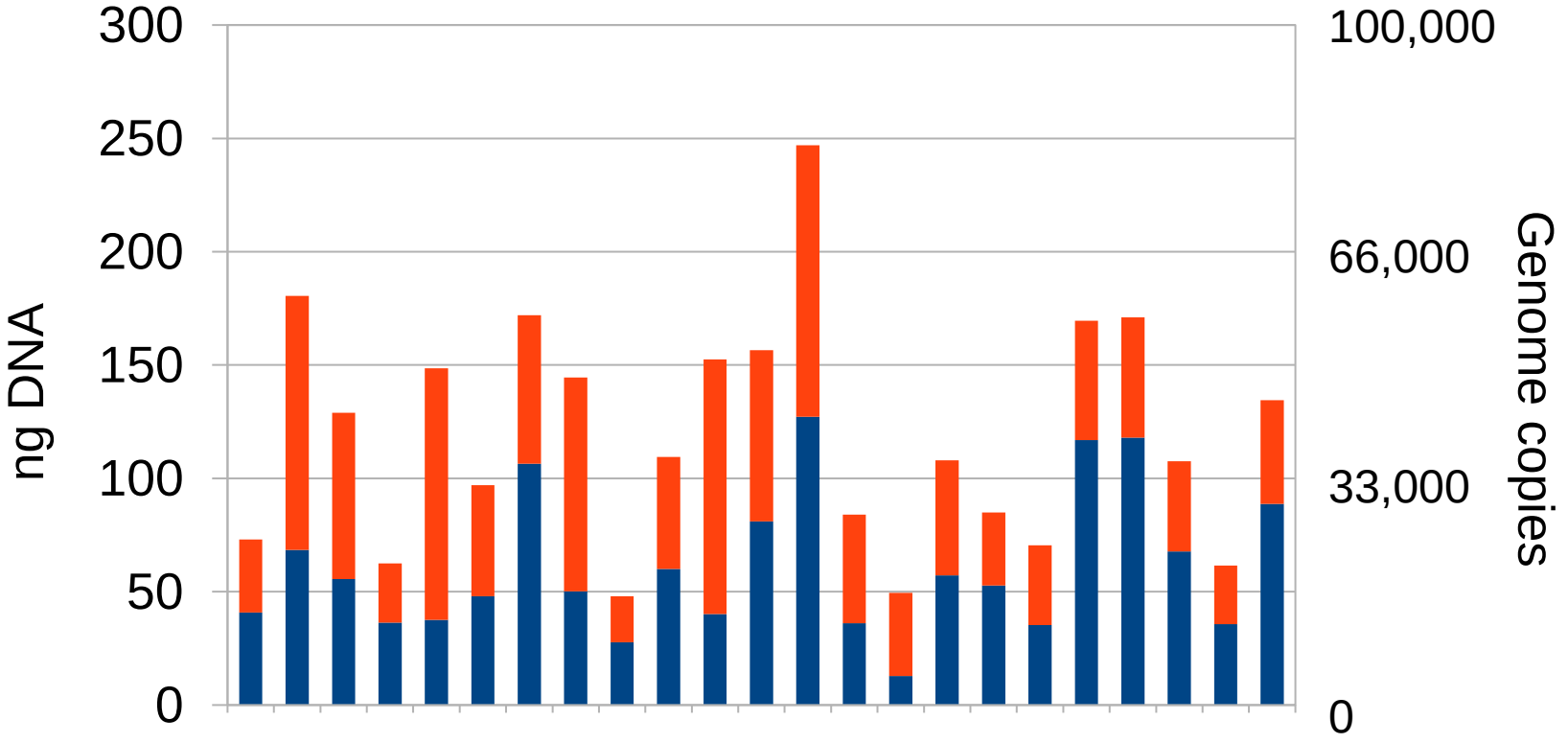


*** assuming 100% cave bear DNA**

Endogenous content



Real genome copies from 50 mg bone powder

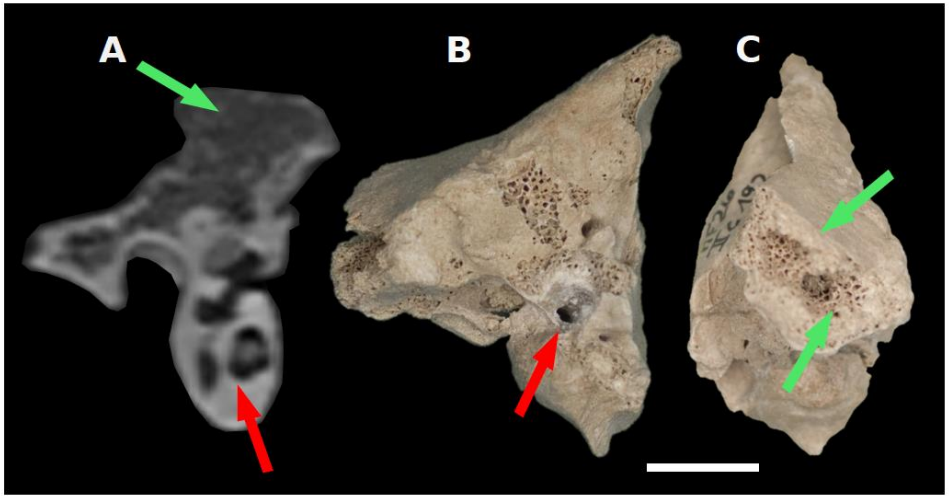


Low amounts of DNA?

Often correct.....

But you can obtain large amounts of DNA from some ancient samples.

Improve your choice



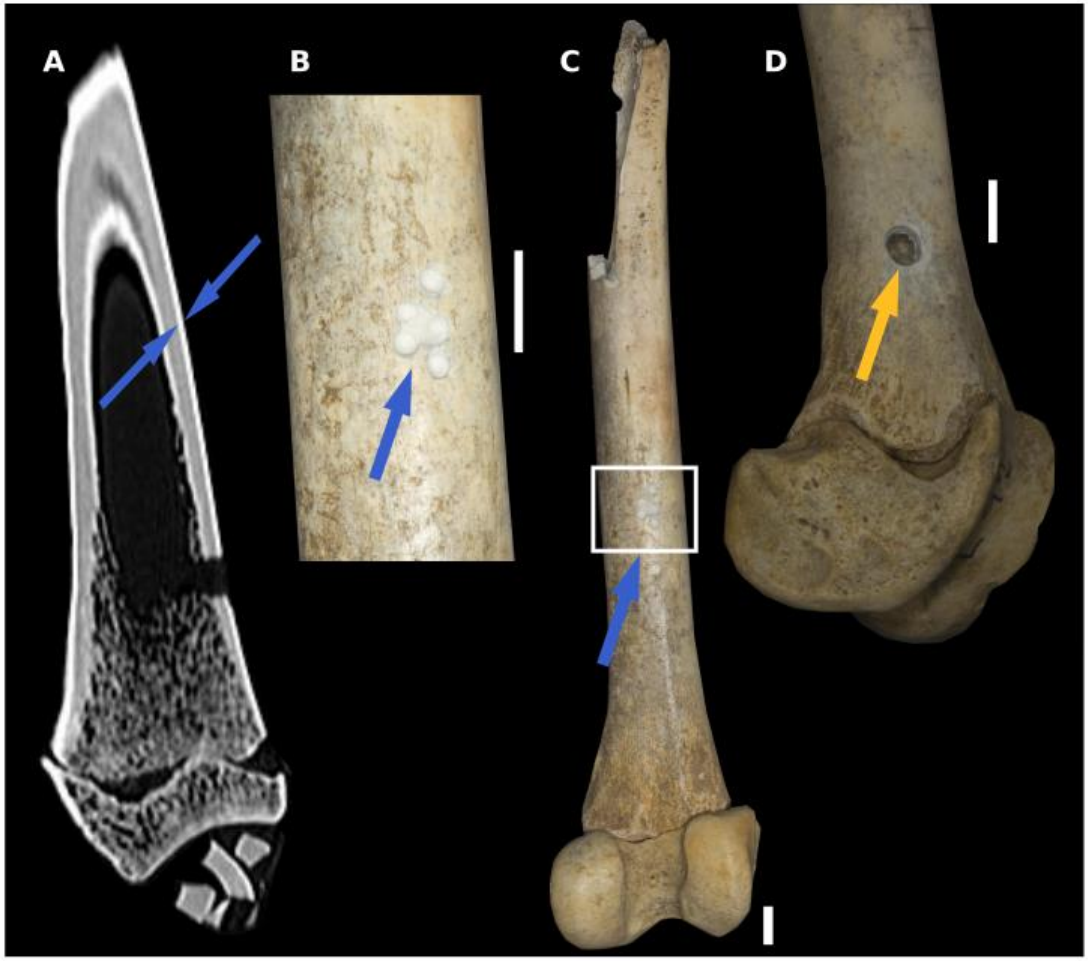
FROM THE COVER

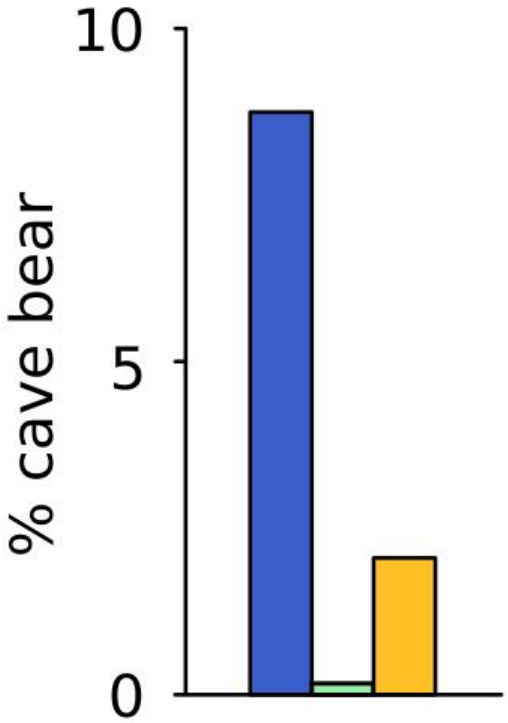
WILEY MOLECULAR ECOLOGY RESOURCES

Optimized DNA sampling of ancient bones using Computed Tomography scans

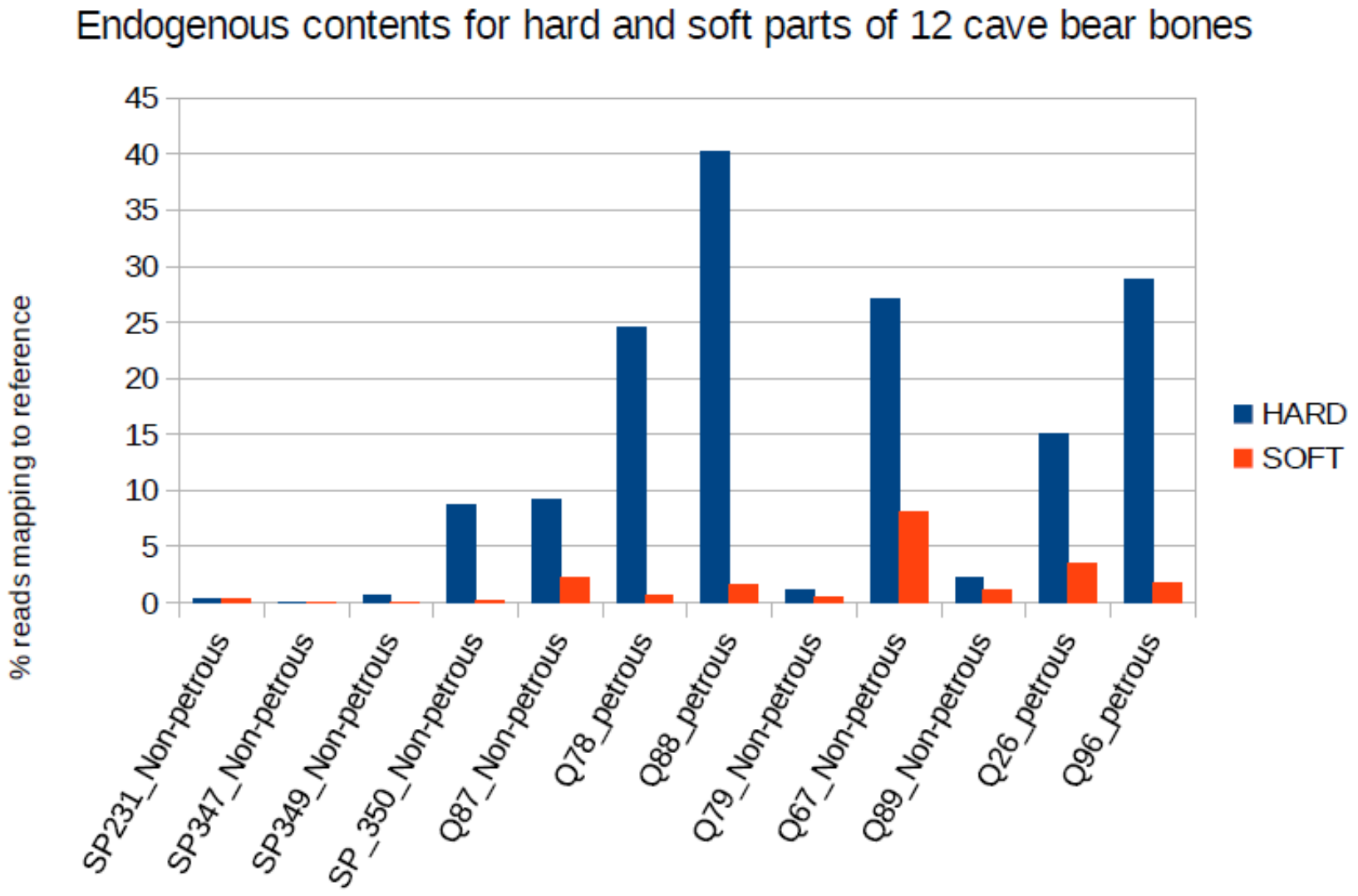
Federica Alberti[1] | Javier Gonzalez[1] | Johanna L. A. Paijmans[1] | Nikolas Basler[1] | Michaela Preick[1] | Kirstin Henneberger[1] | Alexandra Trinks[1,2] | Gernot Rabeder[3] | Nicholas J. Conard[4] | Susanne C. Münzel[4] | Ulrich Joger[5] | Guido Fritsch[6] | Thomas Hildebrandt[6] | Michael Hofreiter[1] | Axel Barlow[1]

Works also with other bones

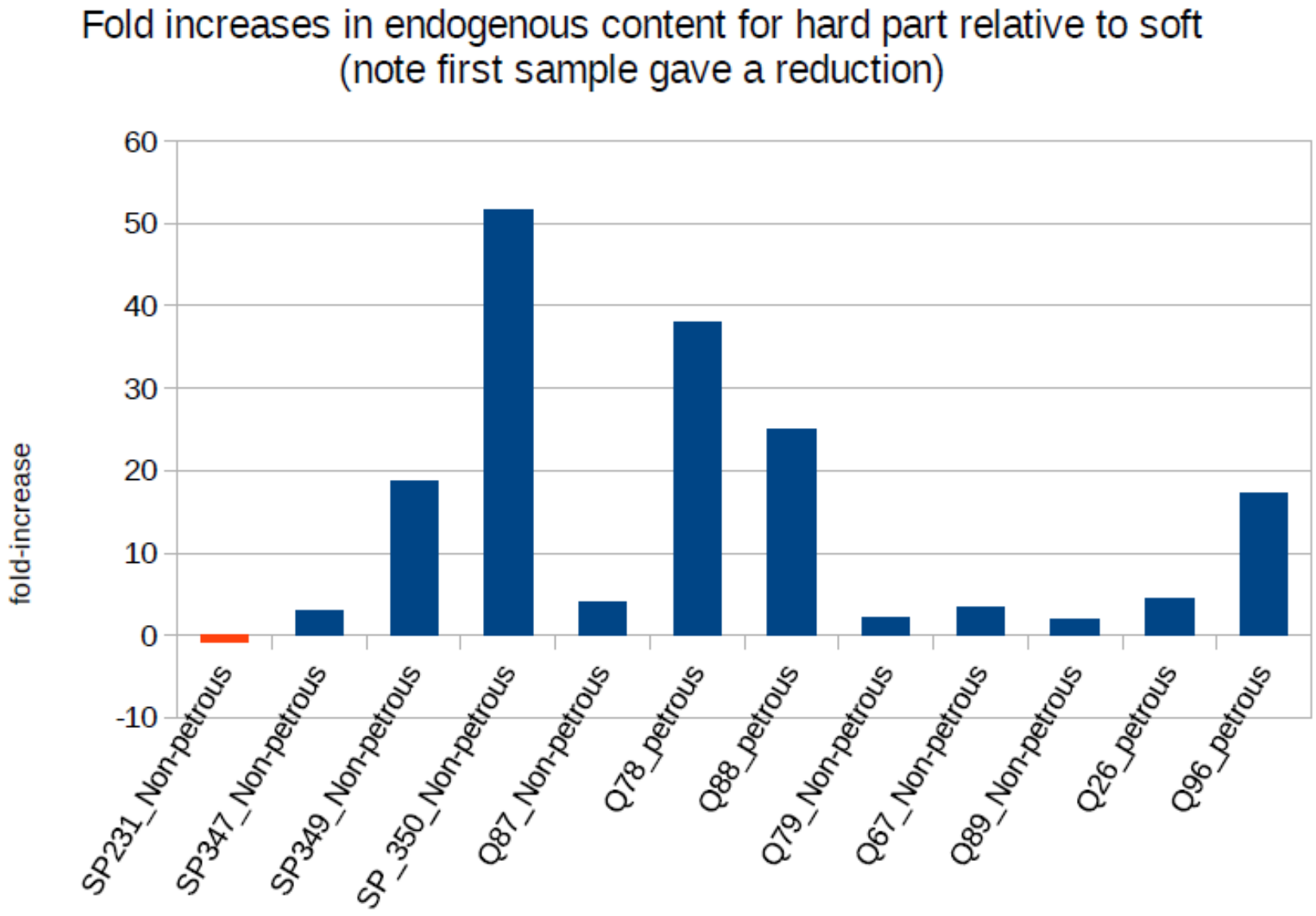




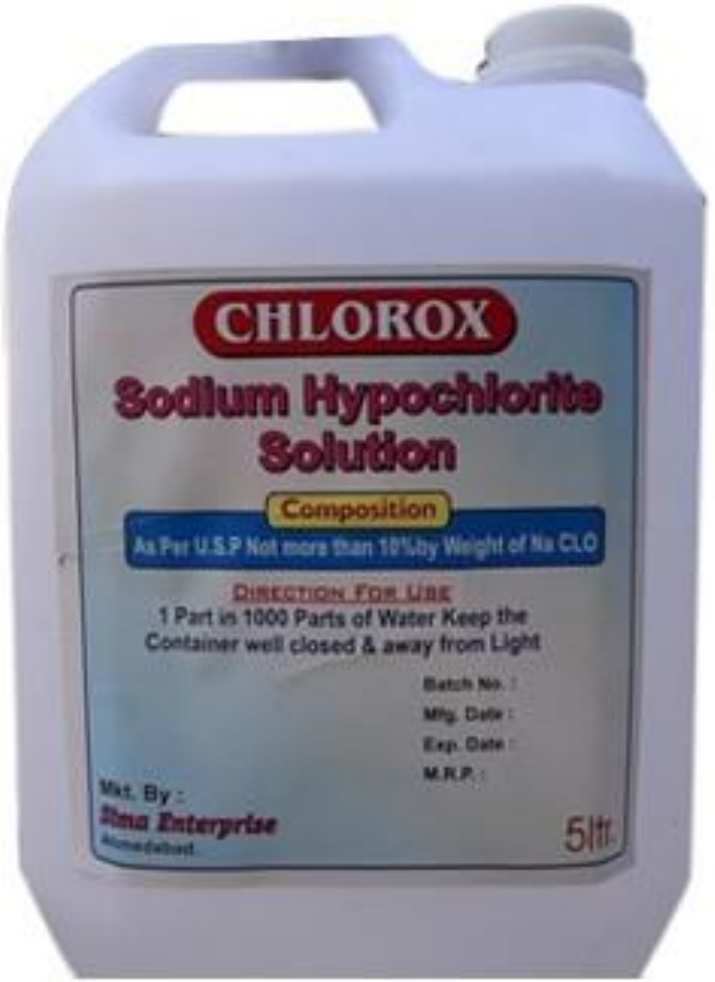
So, if you do not have a petrous – use a CT-scan



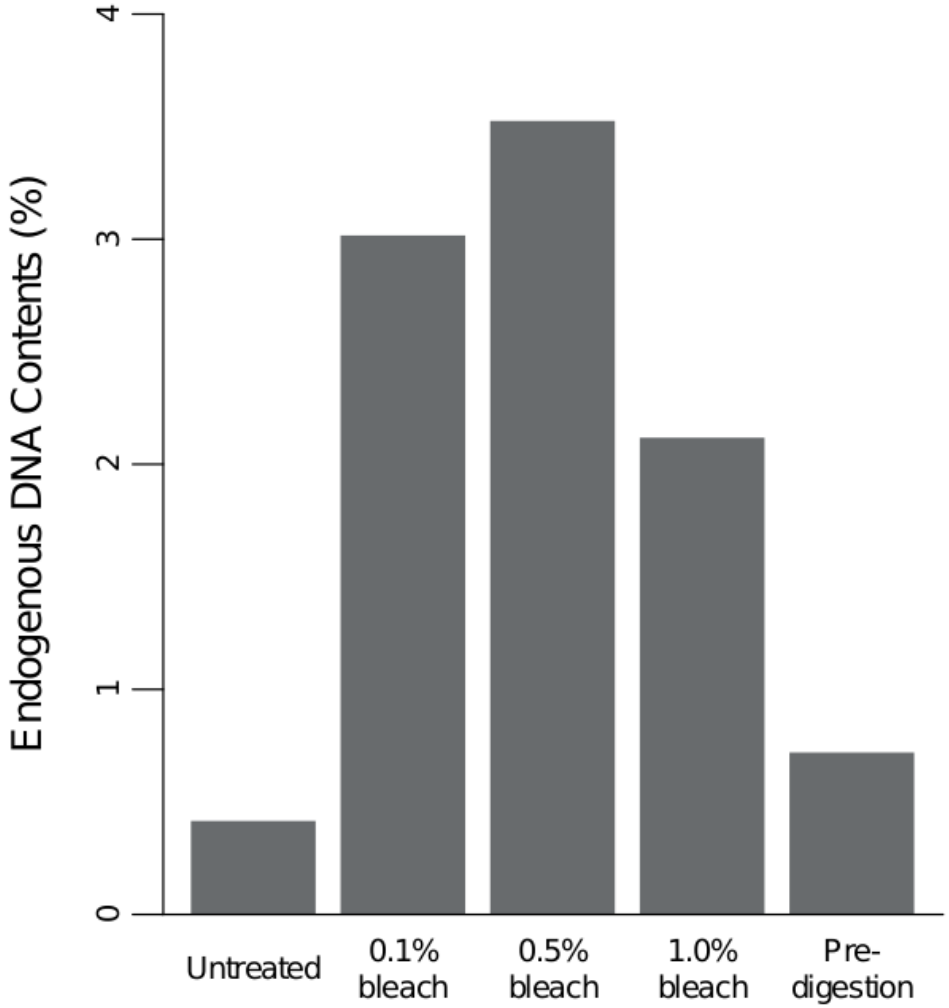
Relative increase in endogenous DNA



Solution 2: Bleach your samples



At least it helps to some extent



Basler *et al. BMC Res Notes (2017) 10:754*
https://doi.org/10.1186/s13104-017-3061-3

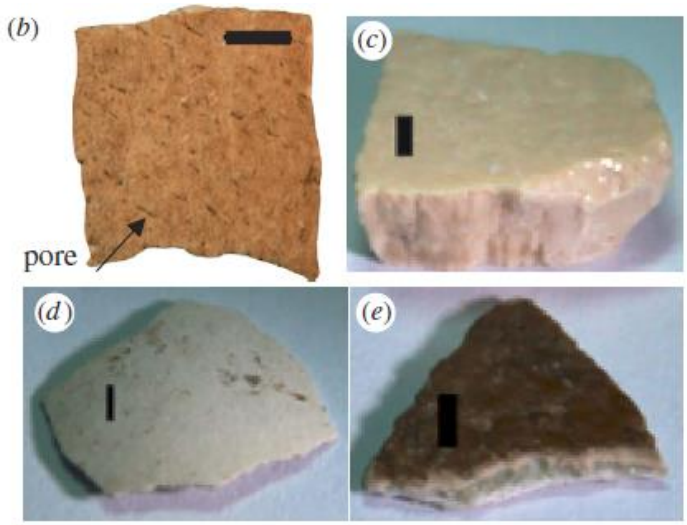
BMC Research Notes

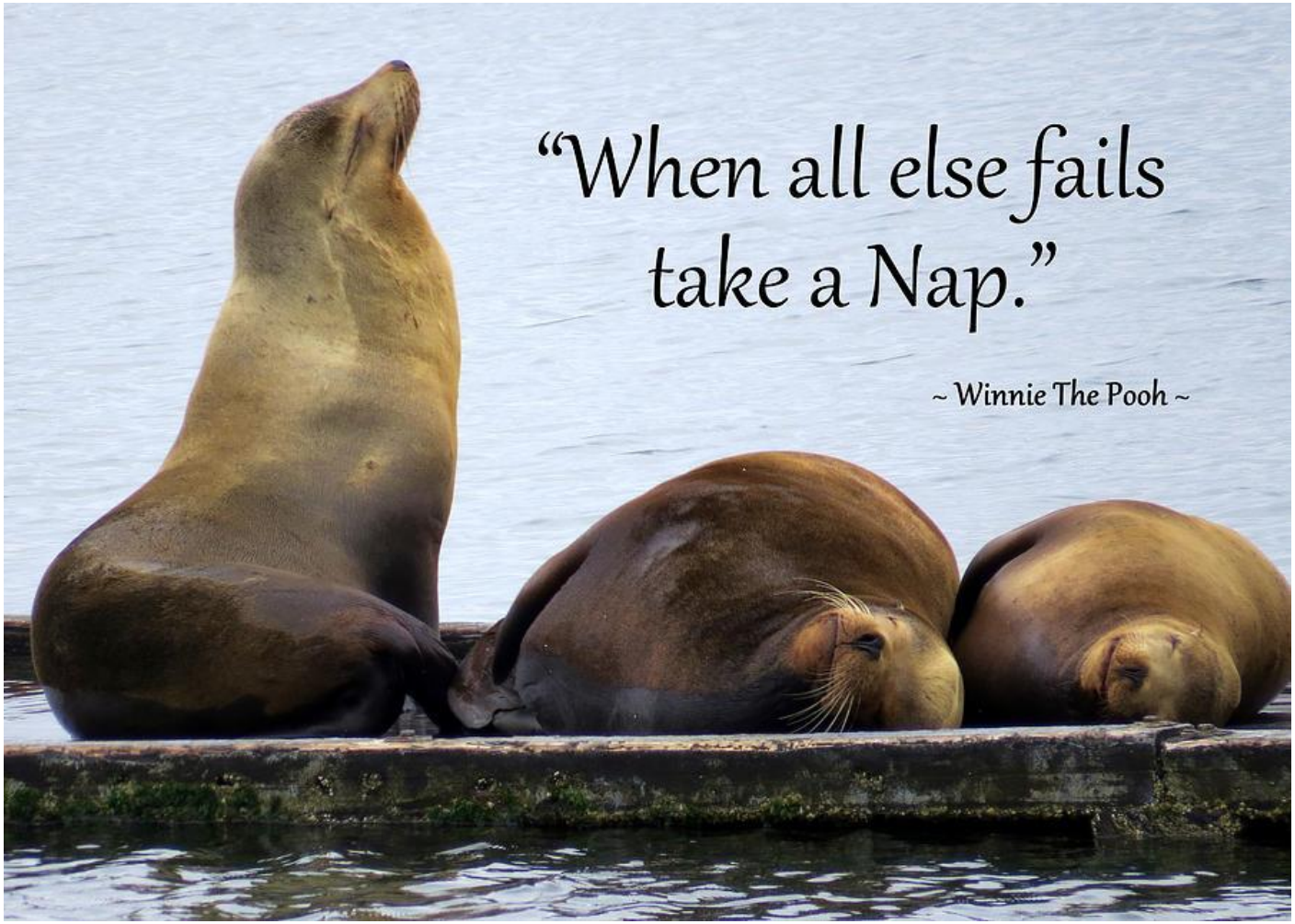
RESEARCH NOTE Open Access

Reduction of the contaminant fraction of DNA obtained from an ancient giant panda bone

Nikolas Basler[1], Georgios Xenikoudakis[1], Michael V. Westbury[1], Lingfeng Song[2], Guilian Sheng[2*] and Axel Barlow[1*]

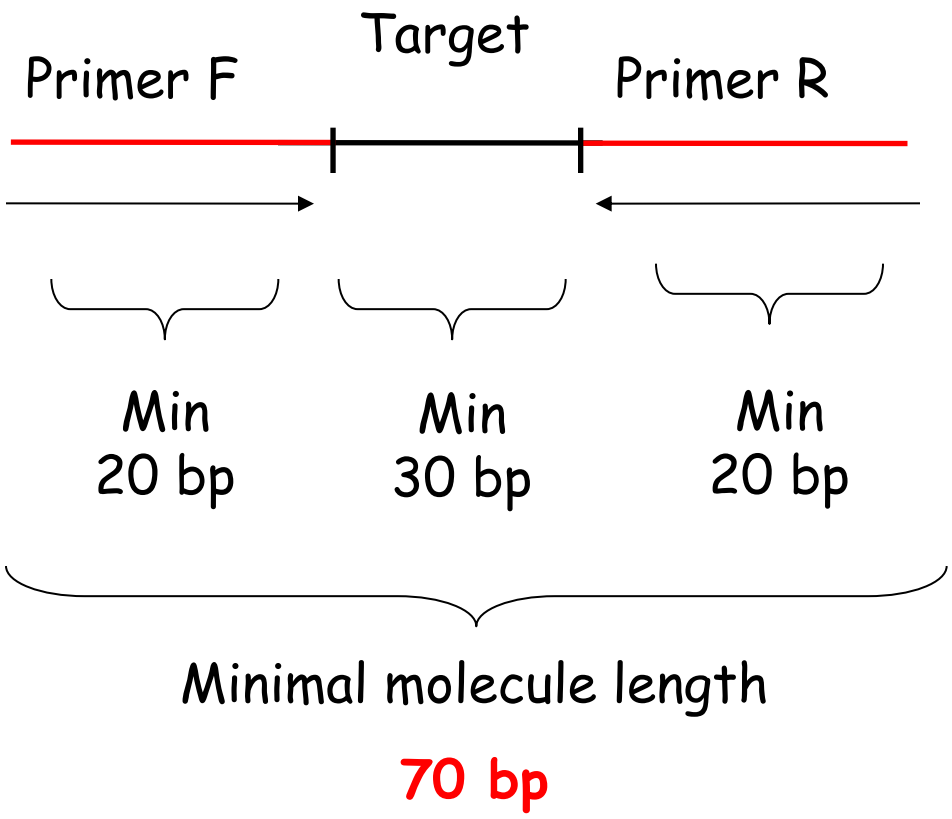
How about other aDNA substrates?



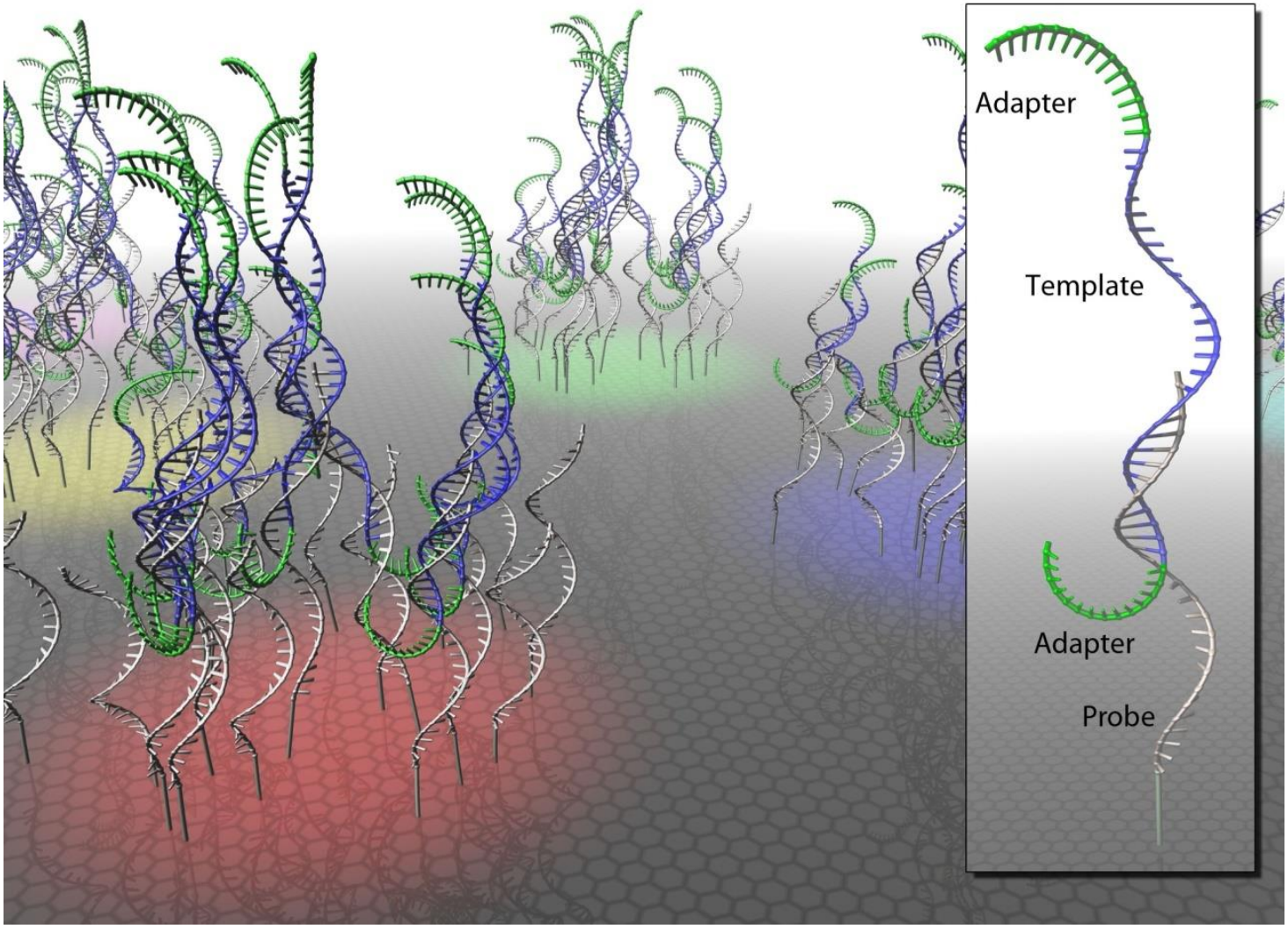

"When all else fails
take a Nap."
~ Winnie The Pooh ~

Solution 3: Target your DNA

Limitations inherent to PCR



DNA hybridization capture

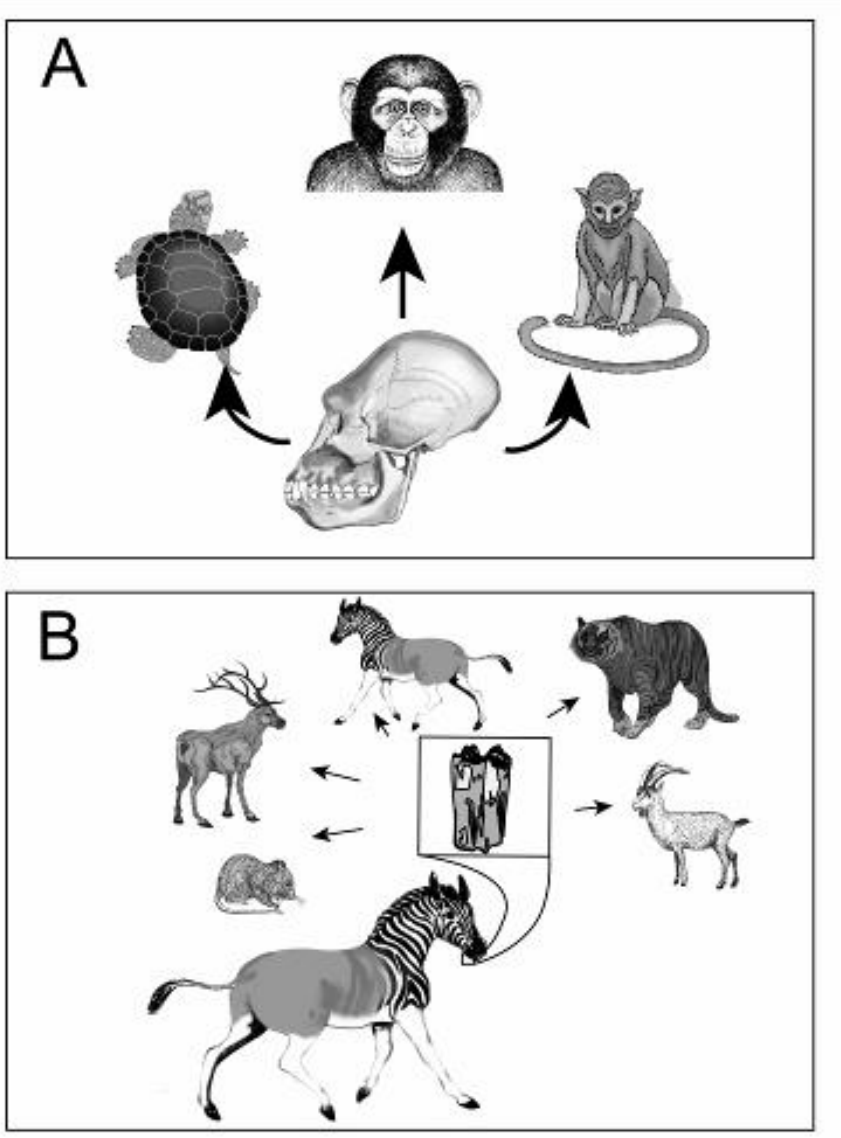


Nice but……

Sample ID	Location	Mitochondrial reference					
		Mappable reads	Mapped reads	unique mapped read	Mapped bp	Depth	Read length
510	NE China(Zhaodong)	9073046	1054870	26968	1883592	**114.21**	16492
1017	NE China(SonghuaRiver)	3066021	1413536	26647	1779650	**107.9**	16499
696	NE China(Qinggang)	2209143	1137192	24730	1587582	**96.6**	16443
564	NE China(Qinggang)	2693221	761104	24049	1454268	**88.2**	16489
508	NE China(Zhaodong)	2767554	616556	18687	1088019	**66.79**	16290
539	NE China(Daqing)	3454017	811496	19121	989439	**60.5**	16367
1016	NE China(Qinggang)	2799561	515682	13214	720338	**44.1**	16327
1011	NE China(Qinggang)	3072086	555920	12465	676867	**41.6**	16275
565	NE China(SonghuaRiver)	3160689	215725	5260	261276	**16.5**	15855
1099	S China(Fuyuan, Yunan)	2313878	29494	224	11487	**7.1**	1625
1093	S China(Fuyuan, Yunan)	2553691	1263	57	2957	**4.2**	708
15	N China(Qinhuangdao, Linxian cave)	2042719	10585	33	1937	**3.65**	531
8	N China(Liaoning Pigeon cave)	2915876	6715	56	3211	**2.53**	1269
7	N China(Liaoning Pigeon cave)	1375153	349	30	1740	**2.3**	747
1092	S China(Fuyuan, Yunan)	1827333	740	13	499	**1.4**	365
22	N China(Qinhuangdao, Linxian cave)	2375921	222	14	550	**1.11**	495
19	N China(Qinhuangdao, Linxian cave)	2118585	518	21	741	**1.08**	686

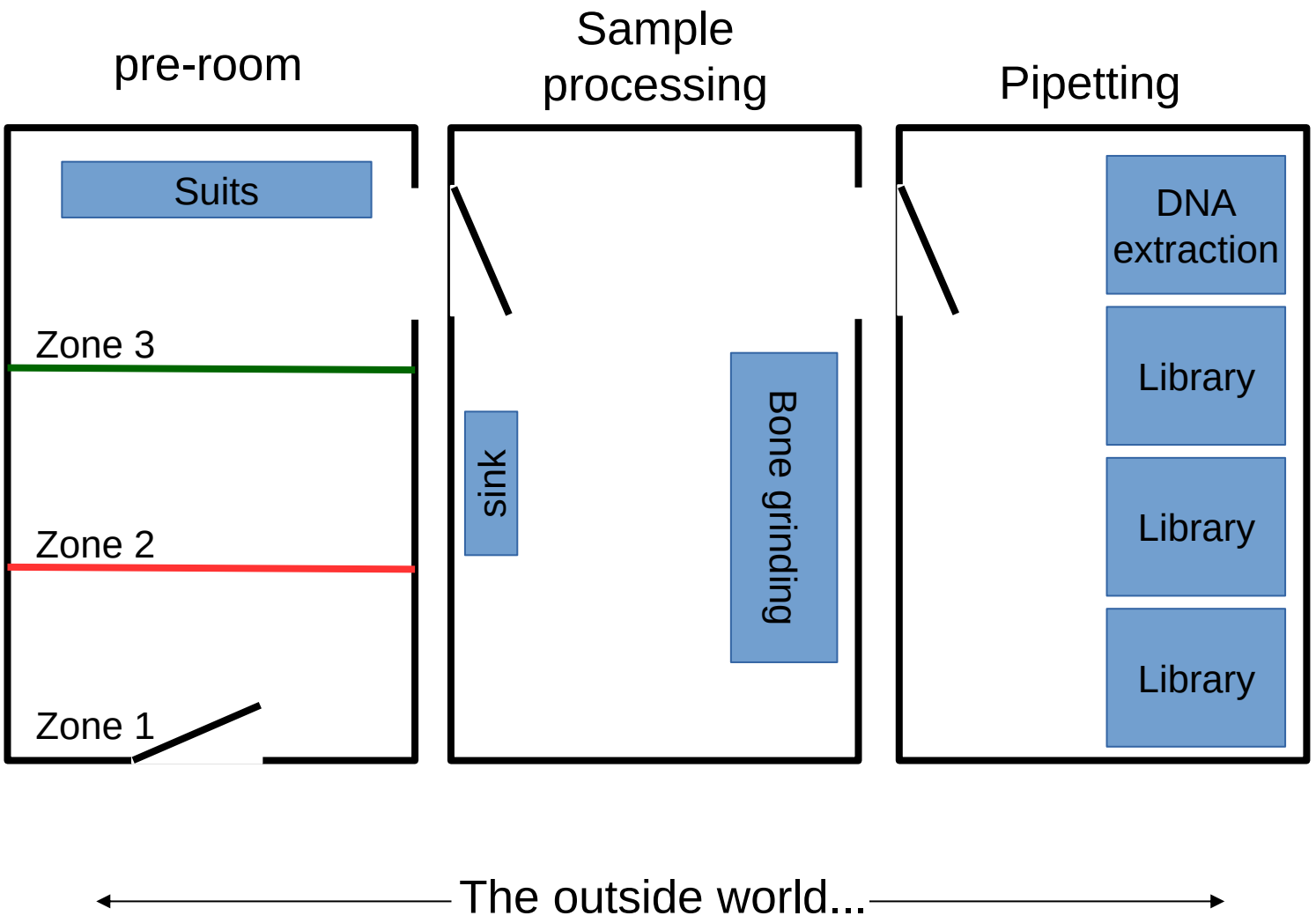
2. Contamination, after all?

Well, it does exist

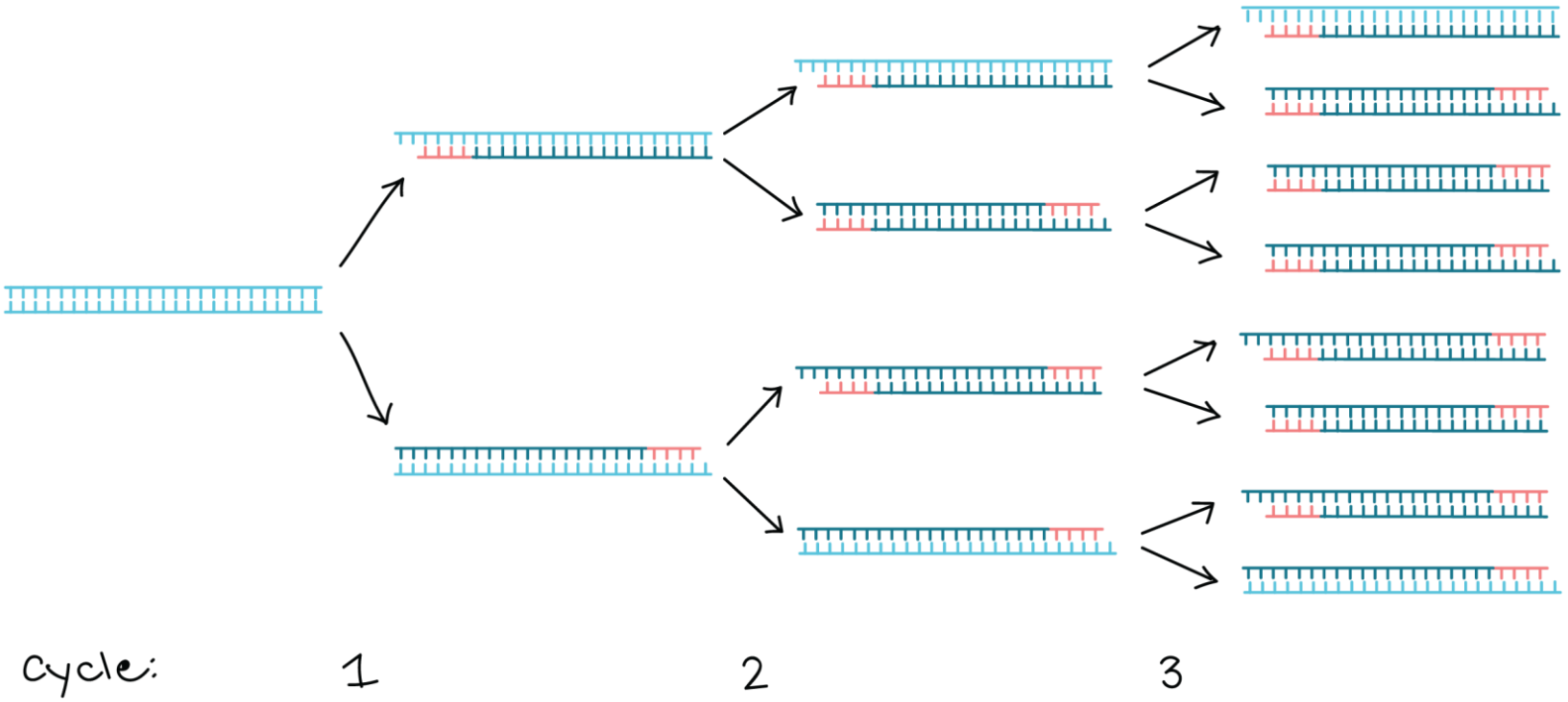


But it tends to be exaggerated

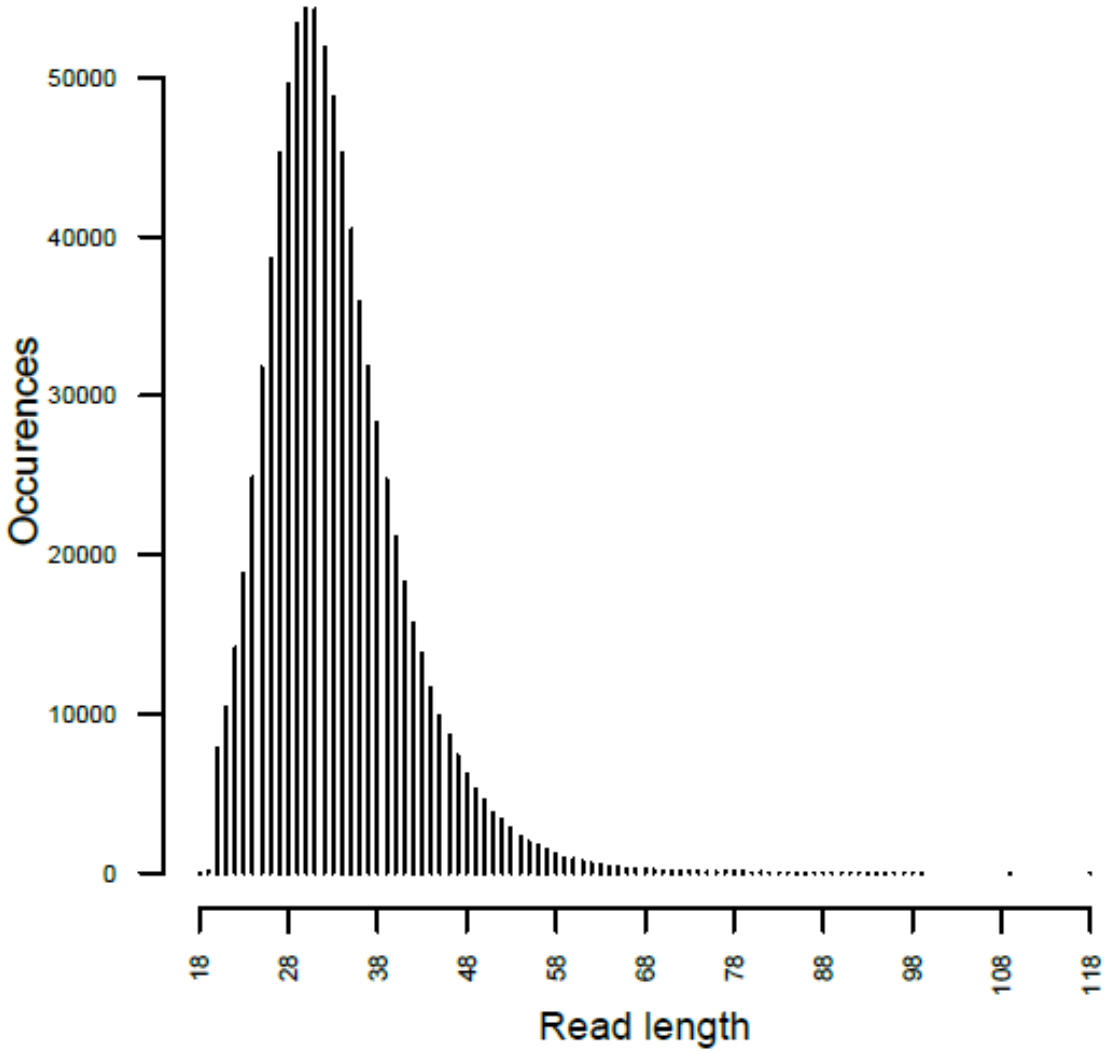
Ancient DNA lab



The real problem: amplified DNA



3. Ancient DNA is short



Ancient DNA fragment length: Rollo et al. 1988

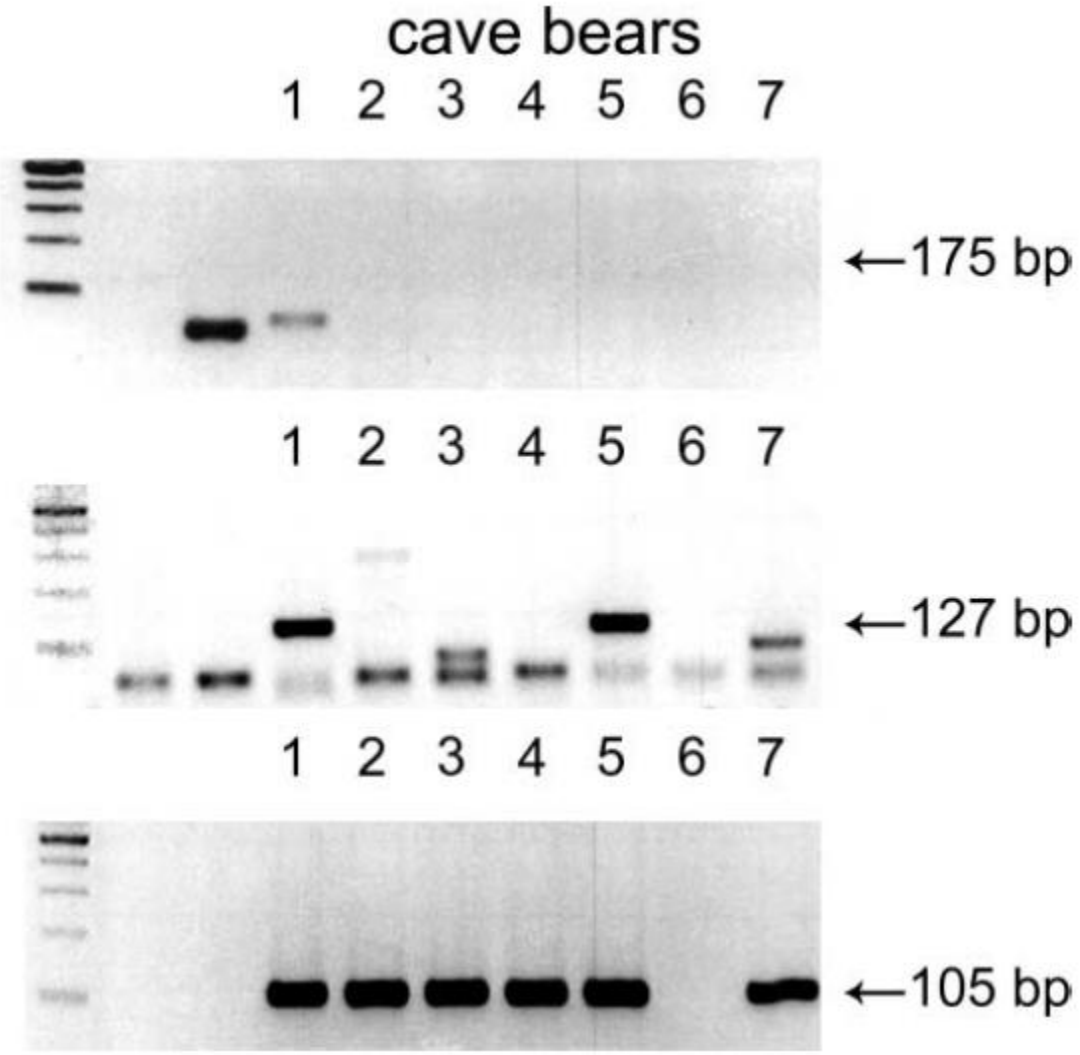
Short but faithful pieces of ancient DNA

FRANCO ROLLO
AUGUSTO AMICI
ROBERTO SALVI
ANNAROSA GARBUGLIA

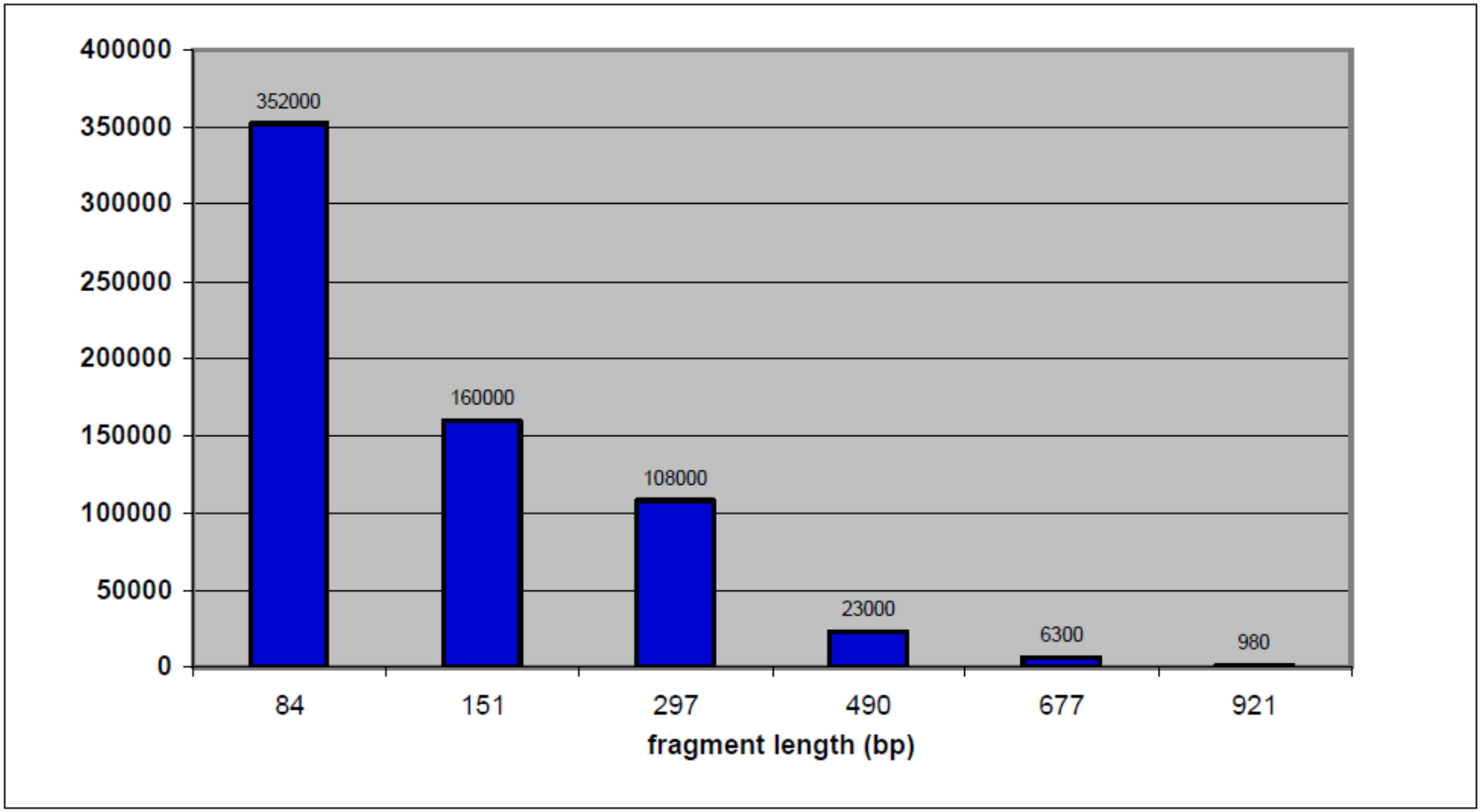
Dipartimento di Biologia Cellulare,
via F. Camerini 2,
I-62032 Camerino,
Italy

NATURE VOL. 335 27 OCTOBER 1988

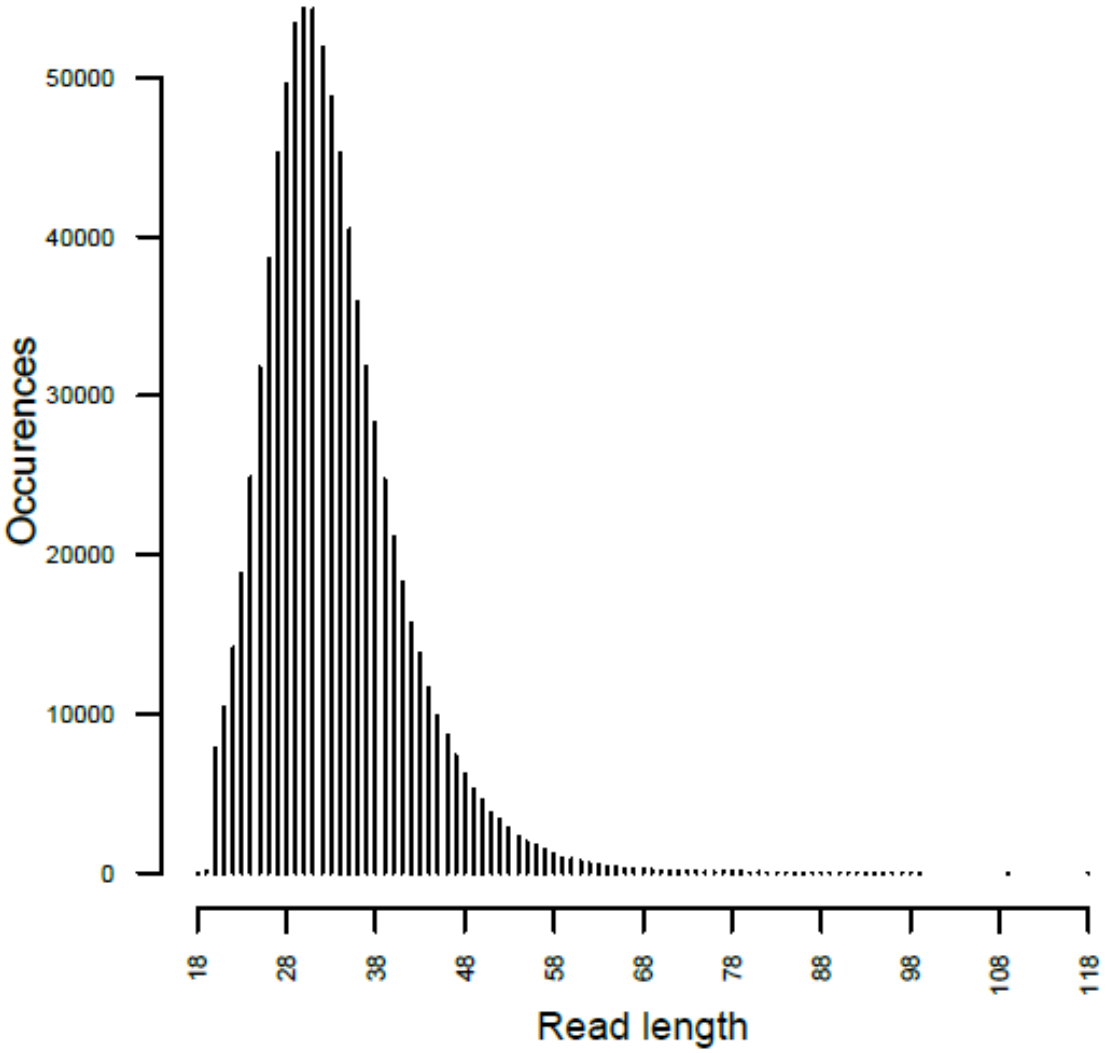
Ancient DNA fragment length: Pääbo et al. 2004



Ancient DNA fragment length: Poinar et al. 2006

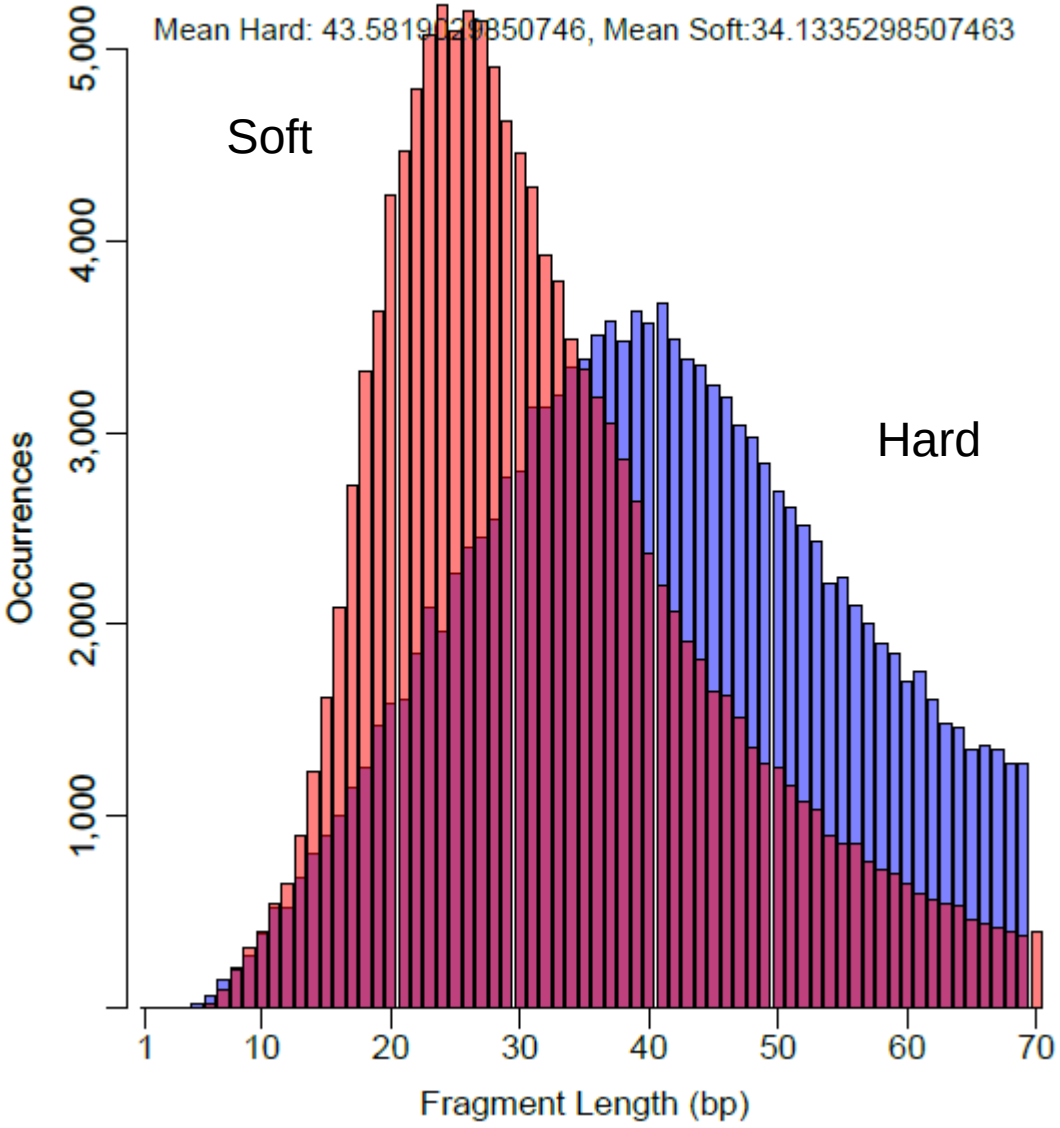


Very old DNA: ~ 120,000 years

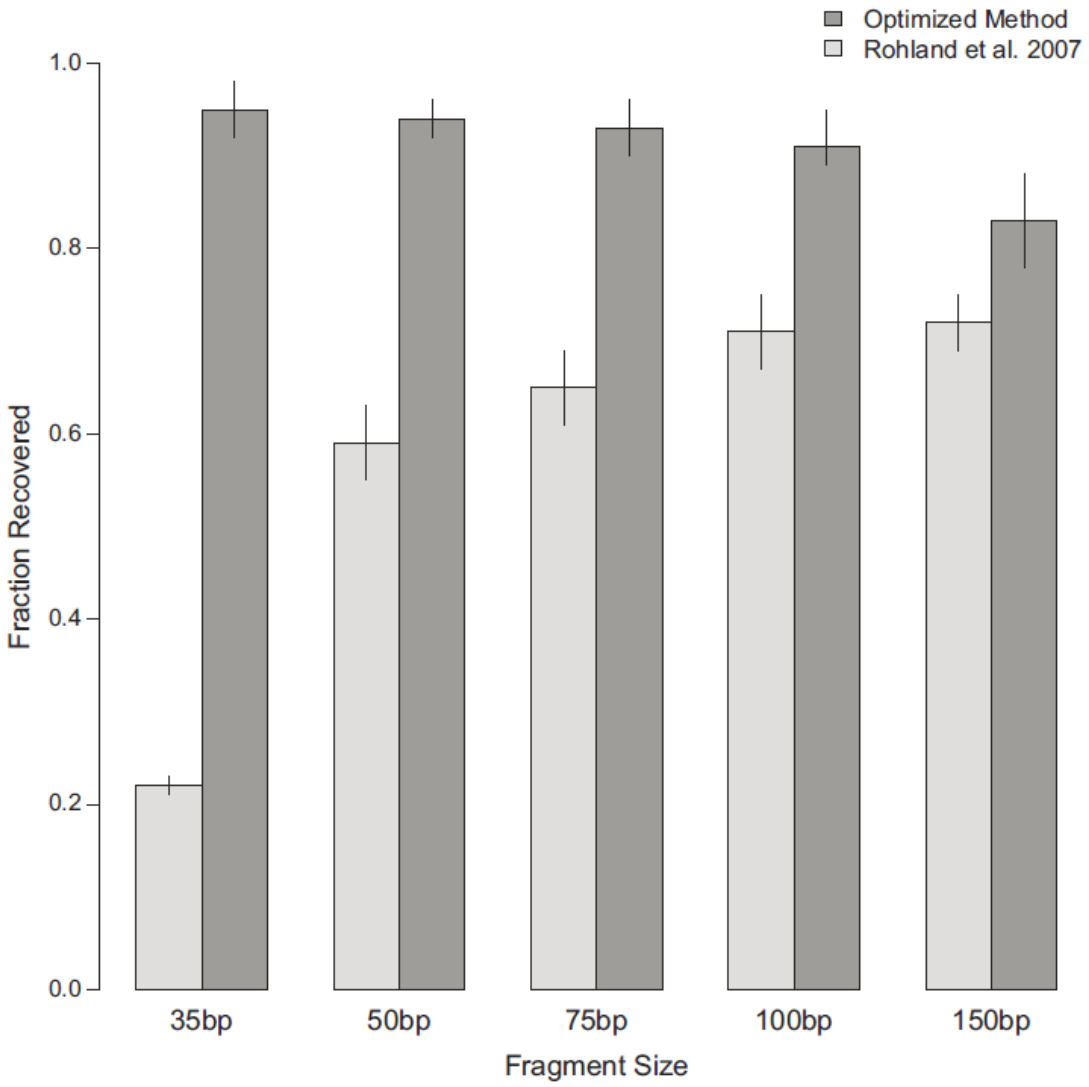


But to be fair……

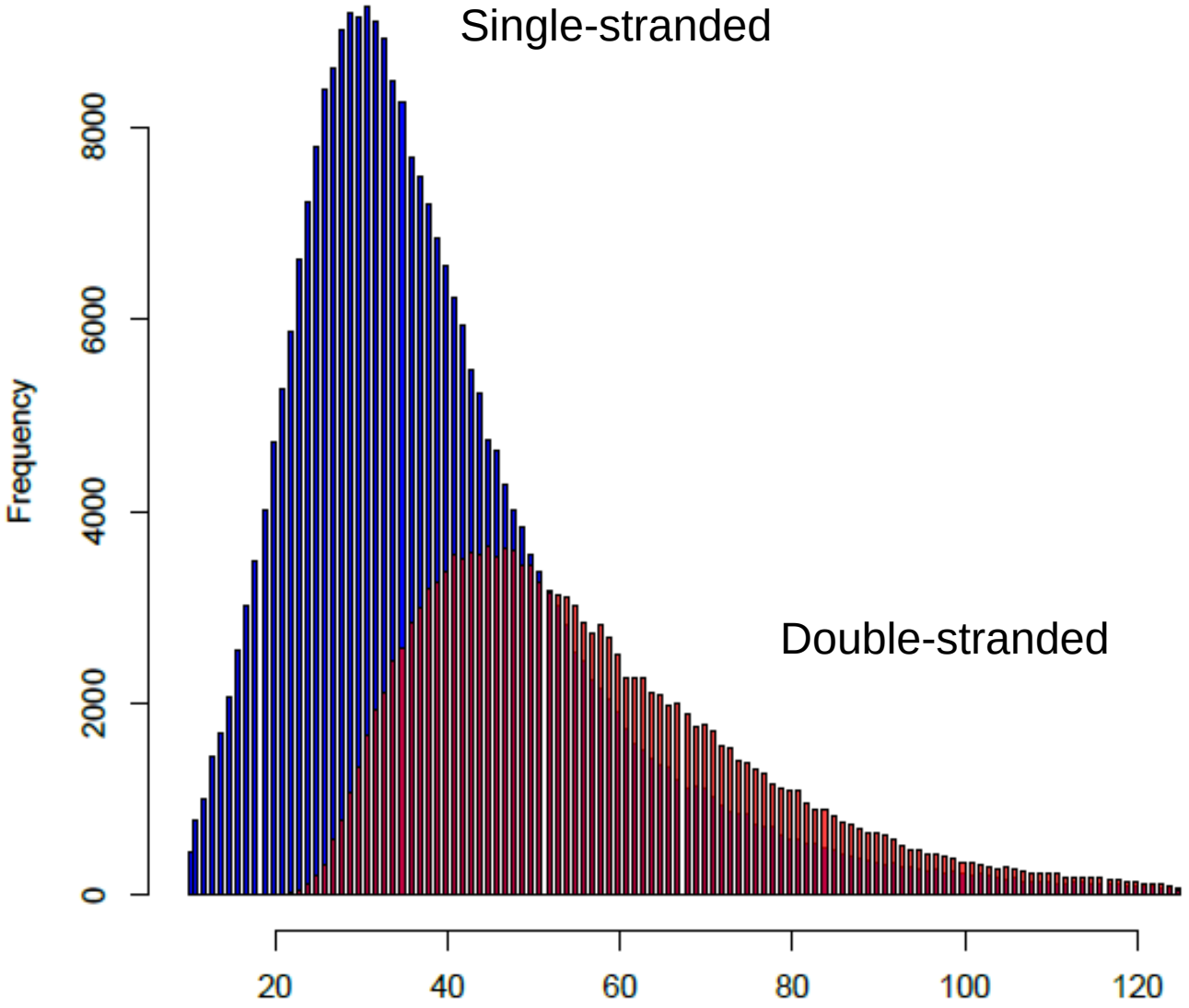
Because it depends a bit – on the bone chosen



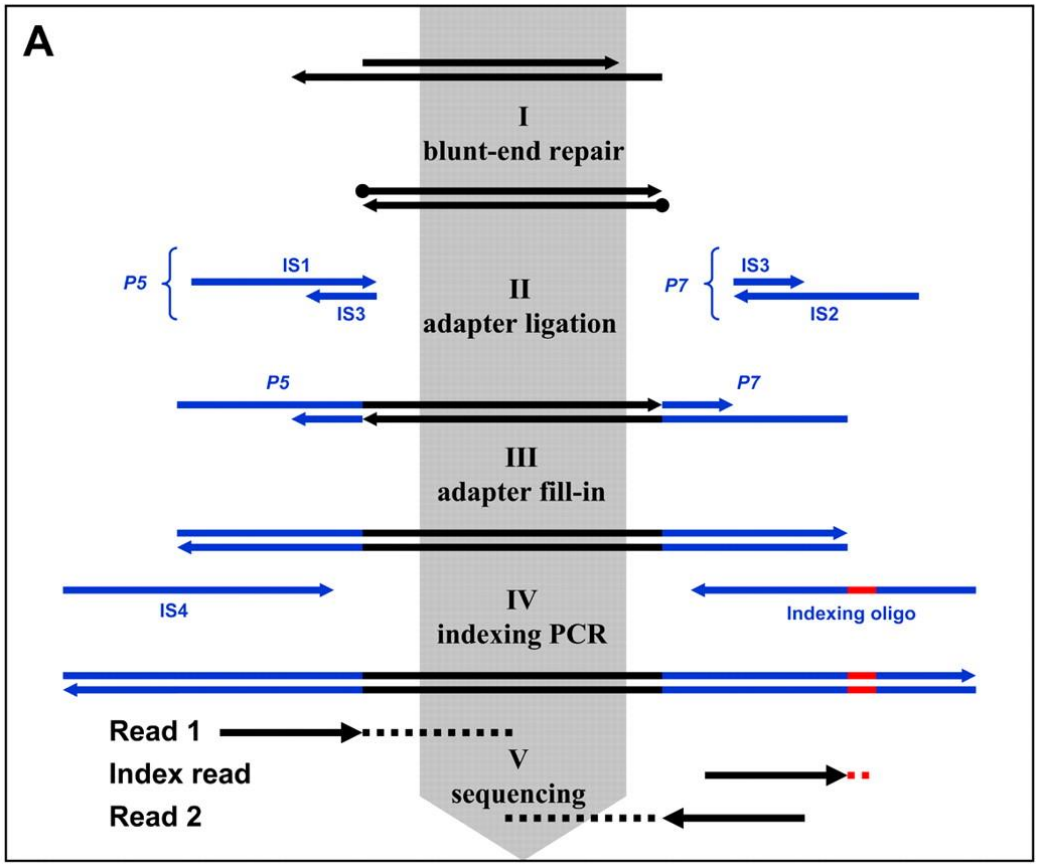
On the DNA extraction used

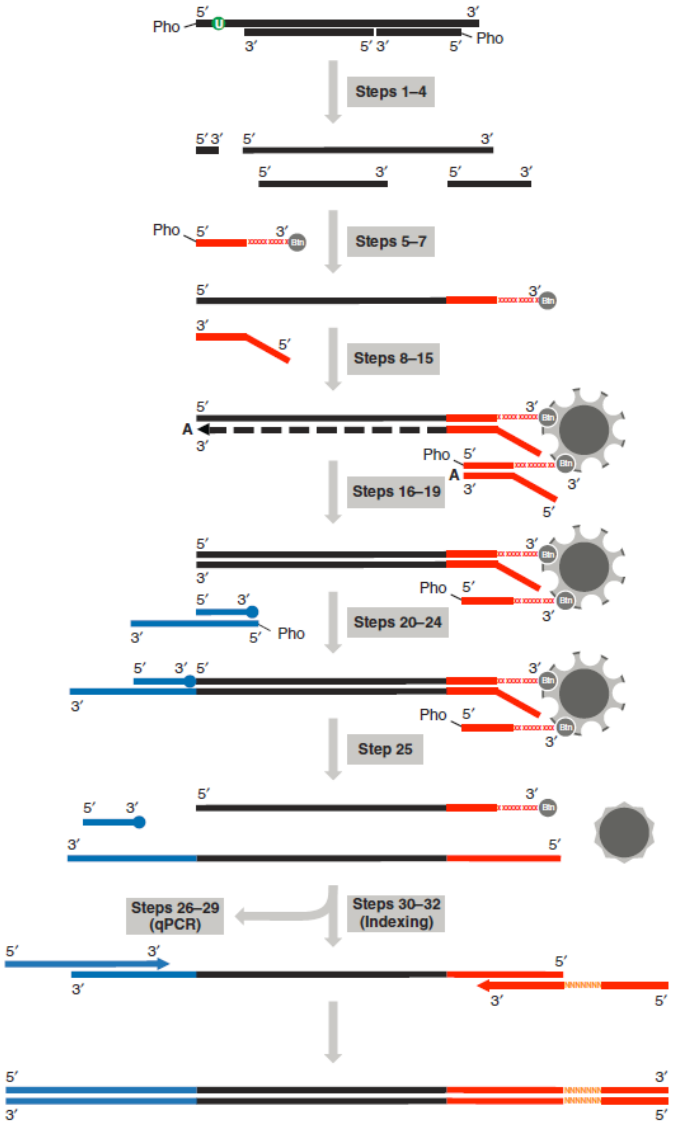


And on the library construction method used

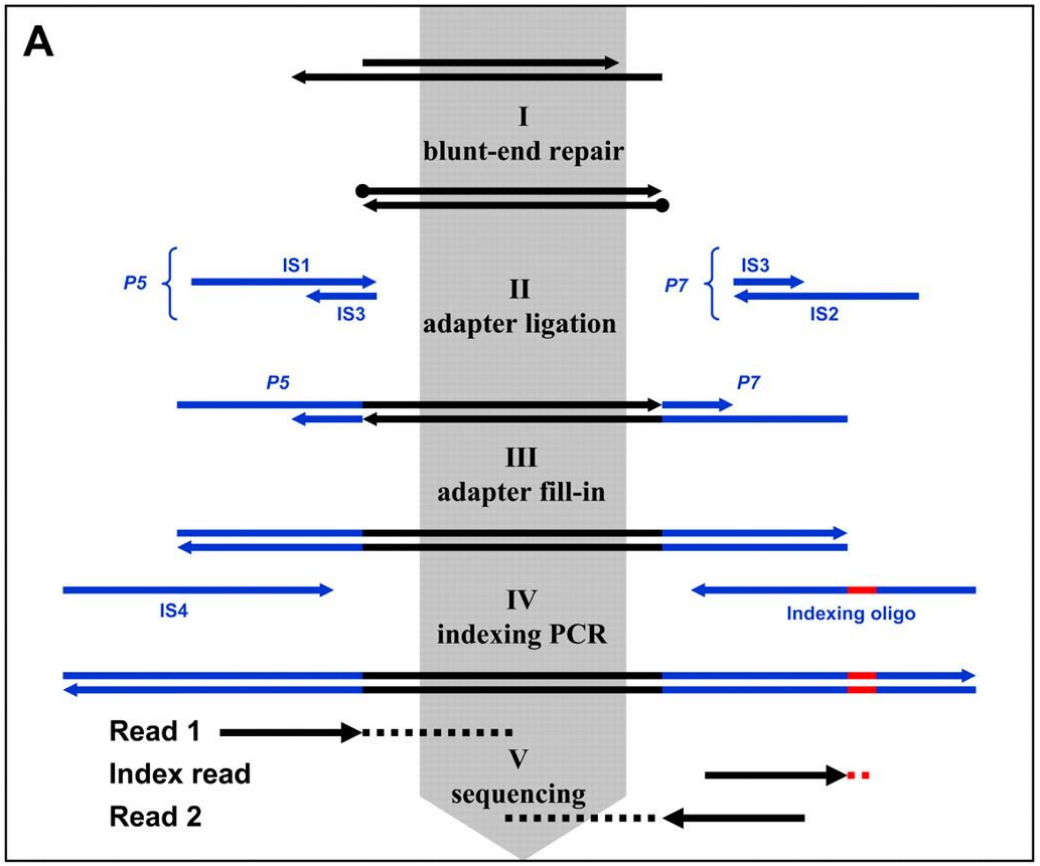


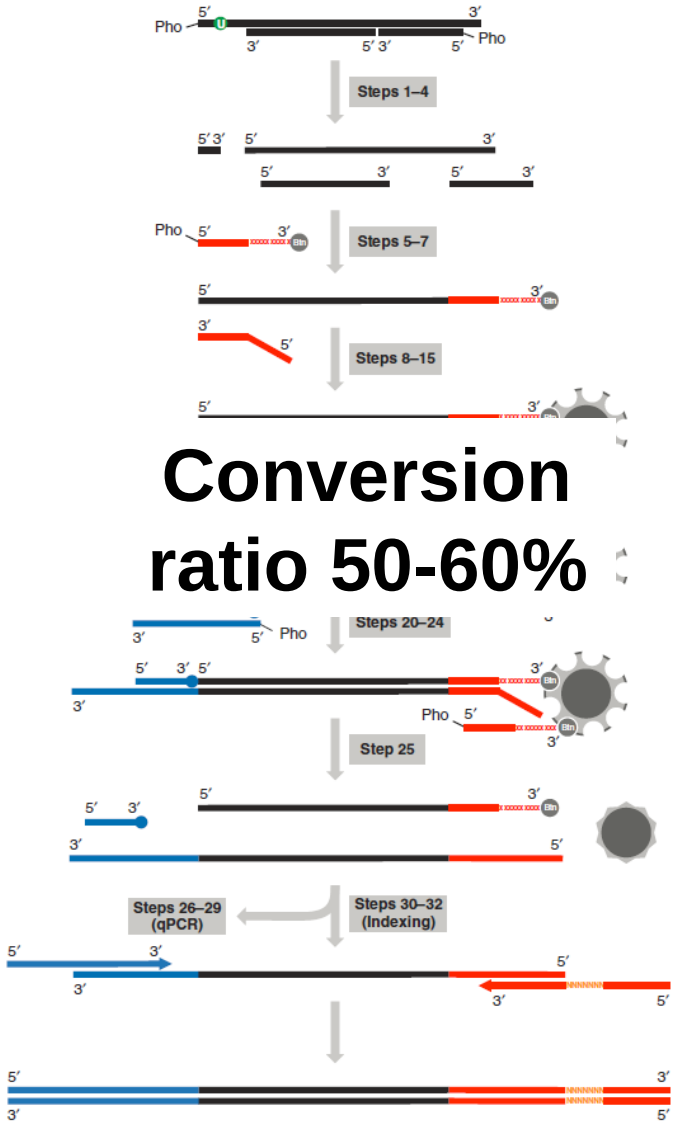
Double-stranded vs. single-stranded





Double-stranded vs. single-stranded





Conversion ratio 50-60%

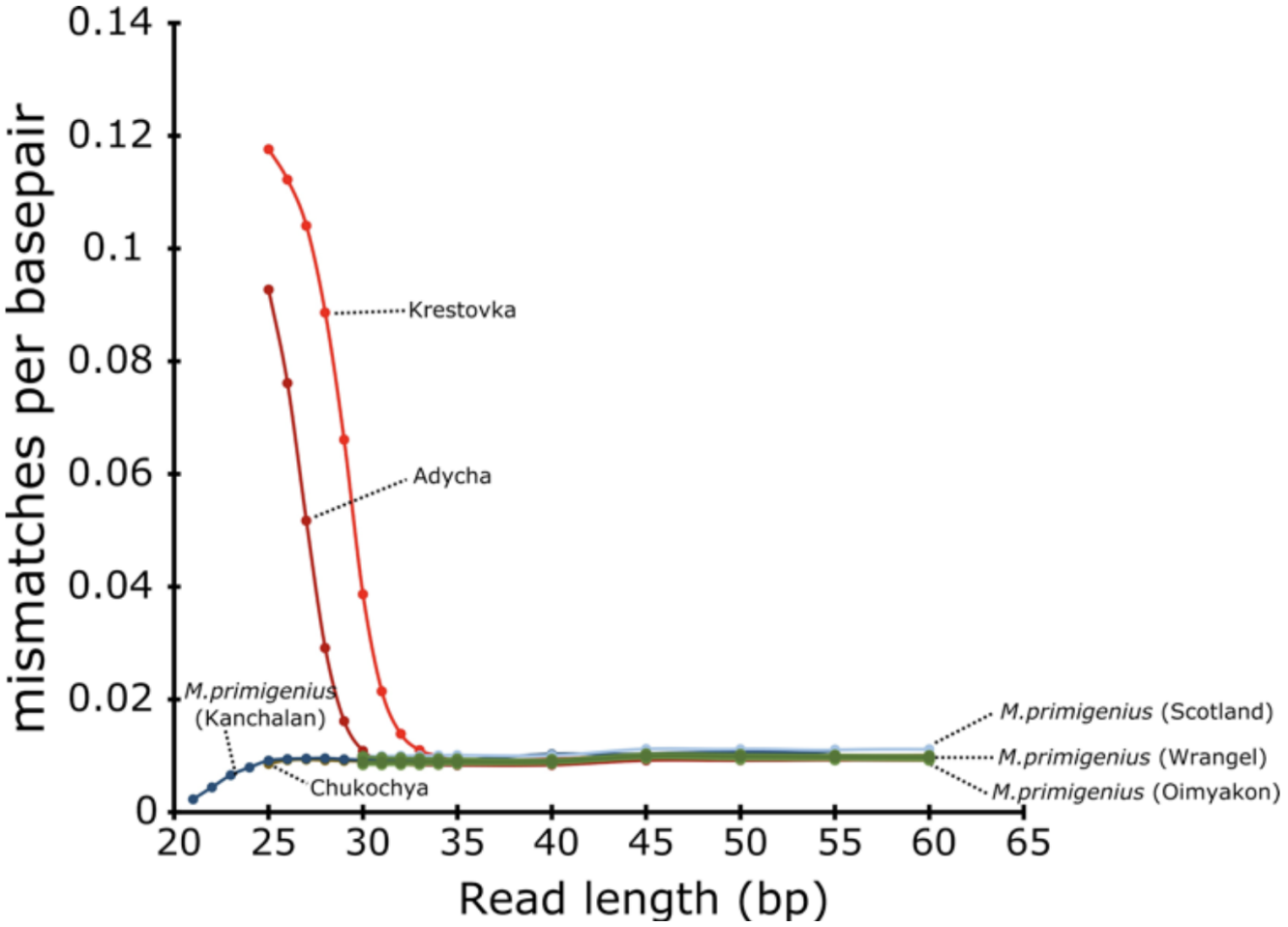
Problems of short fragments: mismapping

Sample	L=10 bp	L=15 bp	L=20 bp	L=30 bp
Ancient Samples				
KG5	22	9	9	9
874/11	5			
874/6	45	26	24	24
752/63	25			
1137/16	5			
1427/41	47	26	26	26
1427/59	1			
MAY10	7			
Modern Reference Horses				
M3160 (*LP/LP*)	375	370	370	293
M3252 (*LP/lp*)	44	44	44	44

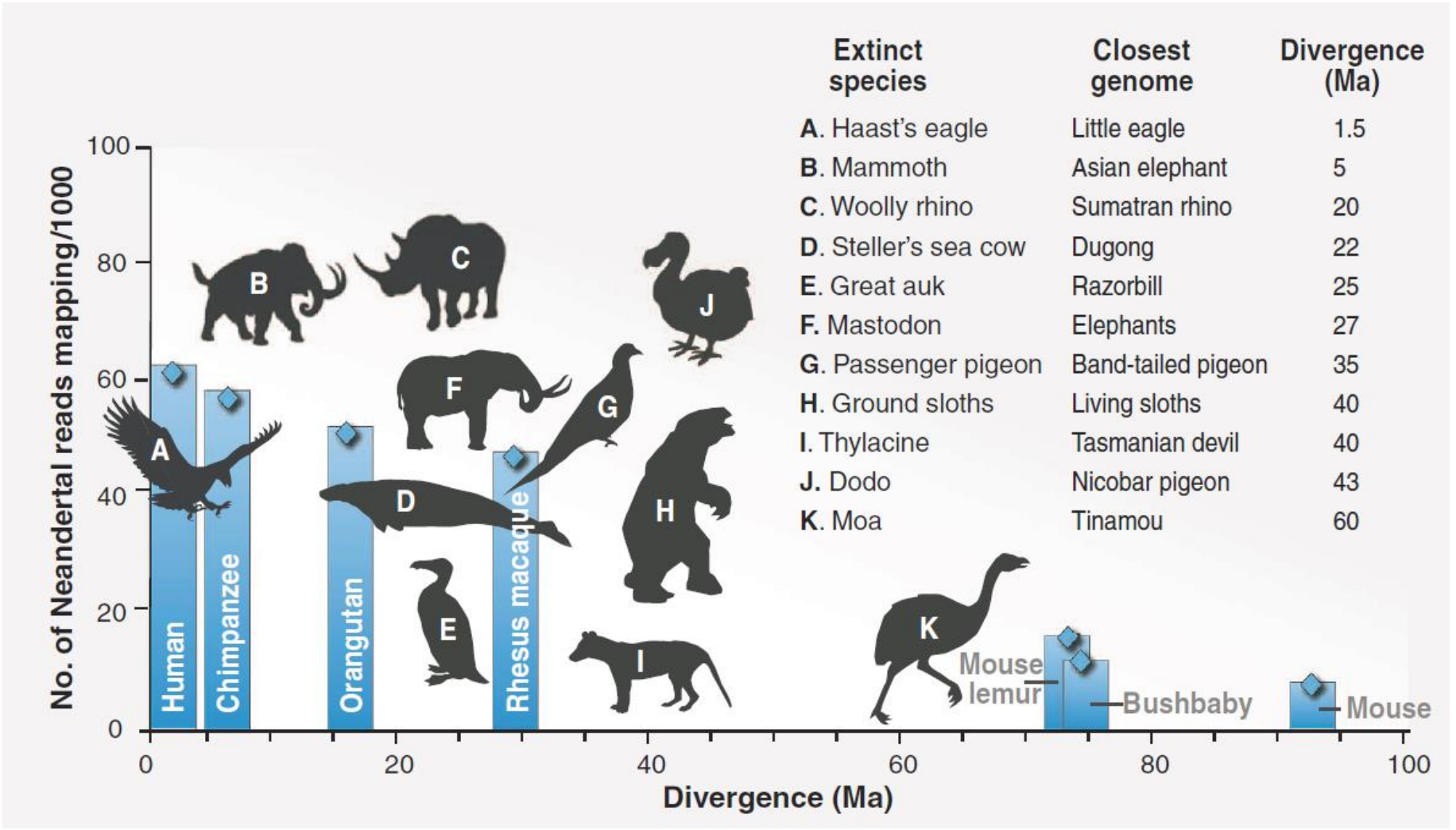
Problems of short fragments: mismapping

rl10 0 0
rl11 0 0
rl12 0 0
rl13 0 0
rl14 0 0
rl15 195 10
rl16 208 0
rl17 1156 2
rl18 2610 61
rl19 14573 1108
rl20 25000 2165
rl21 25011 2090
rl22 25014 1901
rl23 25001 1677
rl24 25008 1277
rl25 25000 758
rl26 25012 411
rl27 25002 215
rl28 25004 142
rl29 25027 124
rl30 25020 109
rl31 25017 106
rl32 25024 103
rl33 25014 115
rl34 25024 109
rl35 25025 111
rl40 25000 118
rl45 25020 102
rl50 25000 110
rl55 25025 110
rl60 25020 91

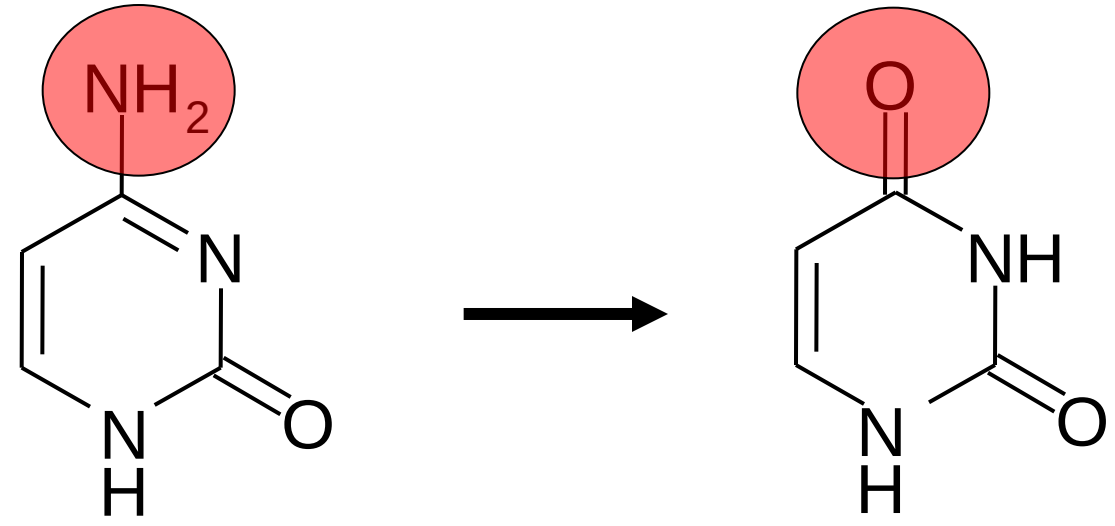
Problems of short fragments: mismapping



Problems of short fragments: no mapping at all



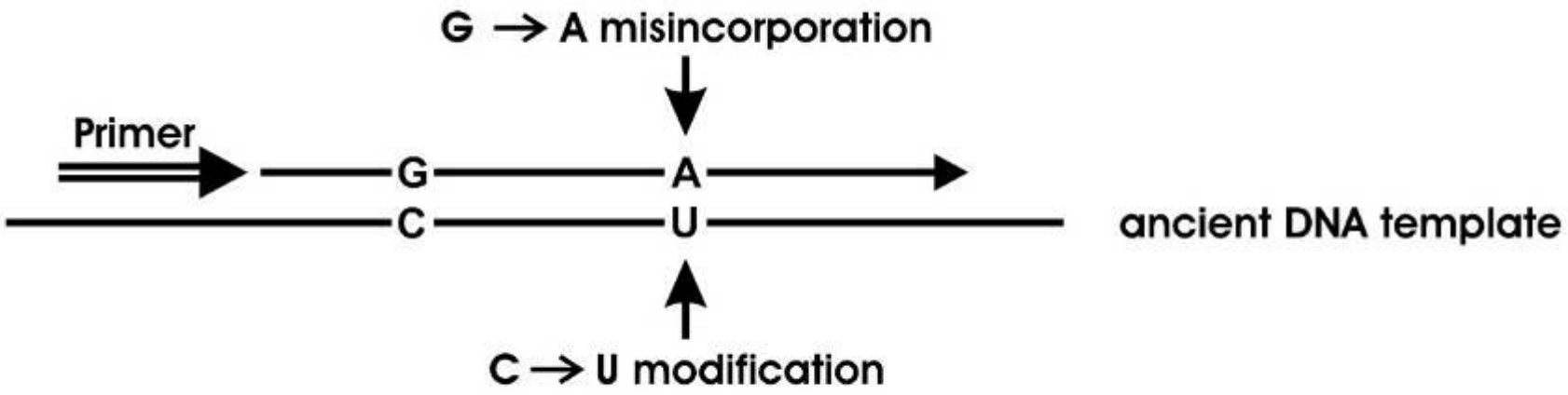
4. Miscoding lesions



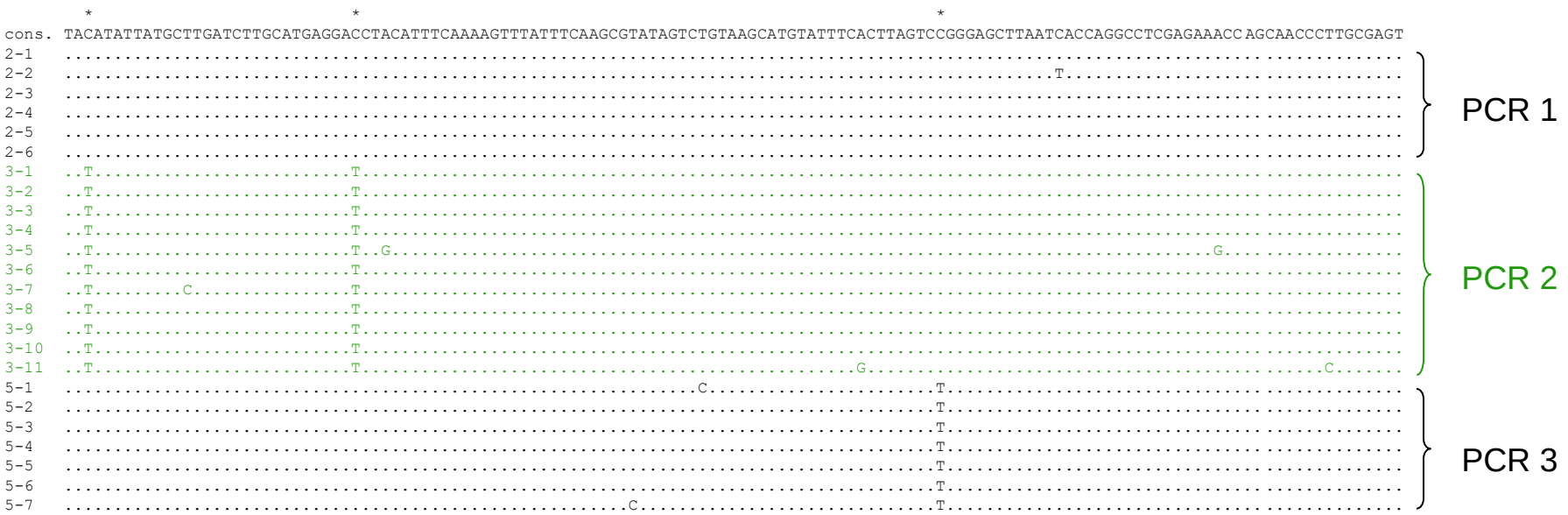
Cytosine

Uracil

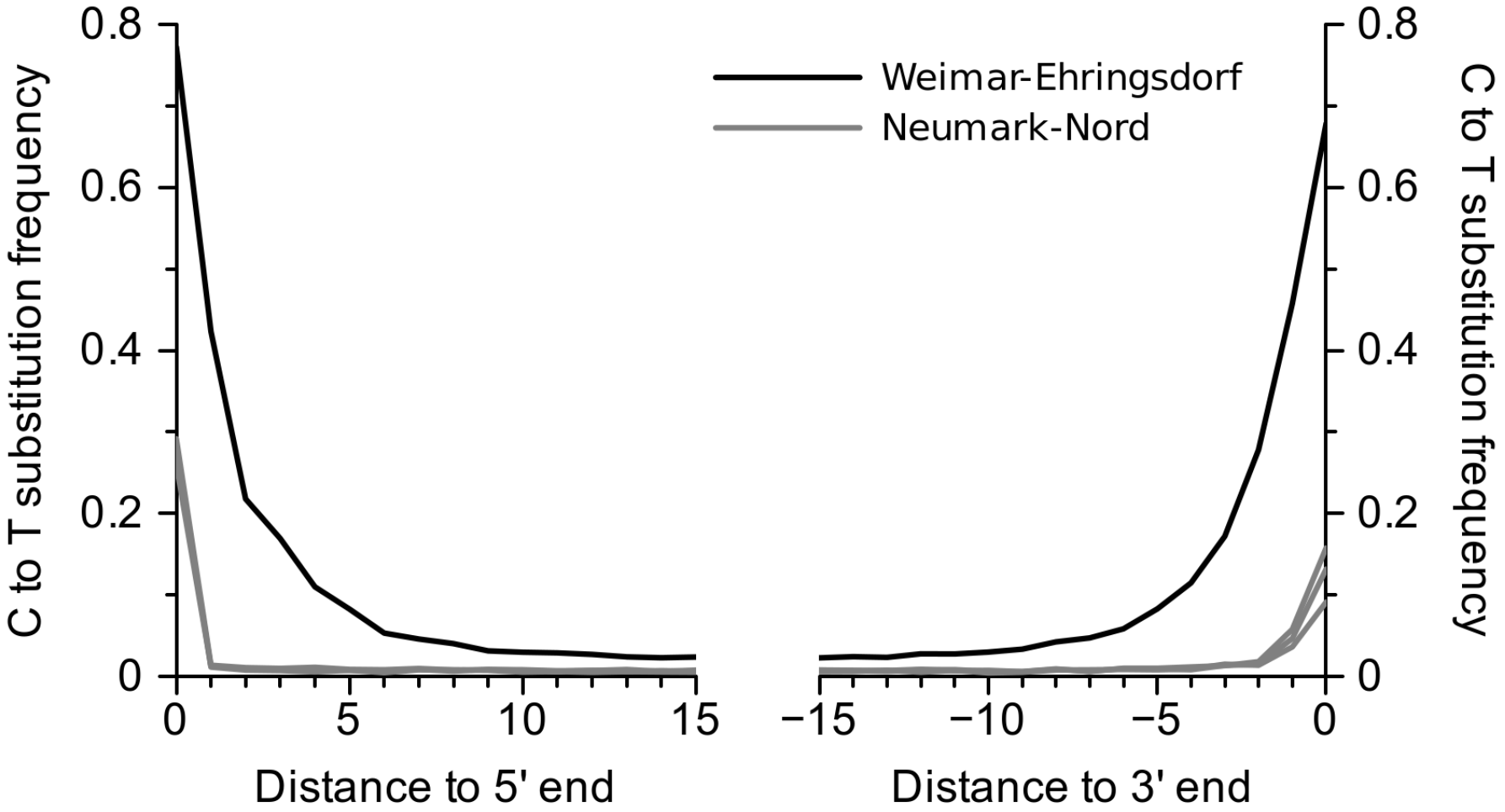
Miscoding lesions



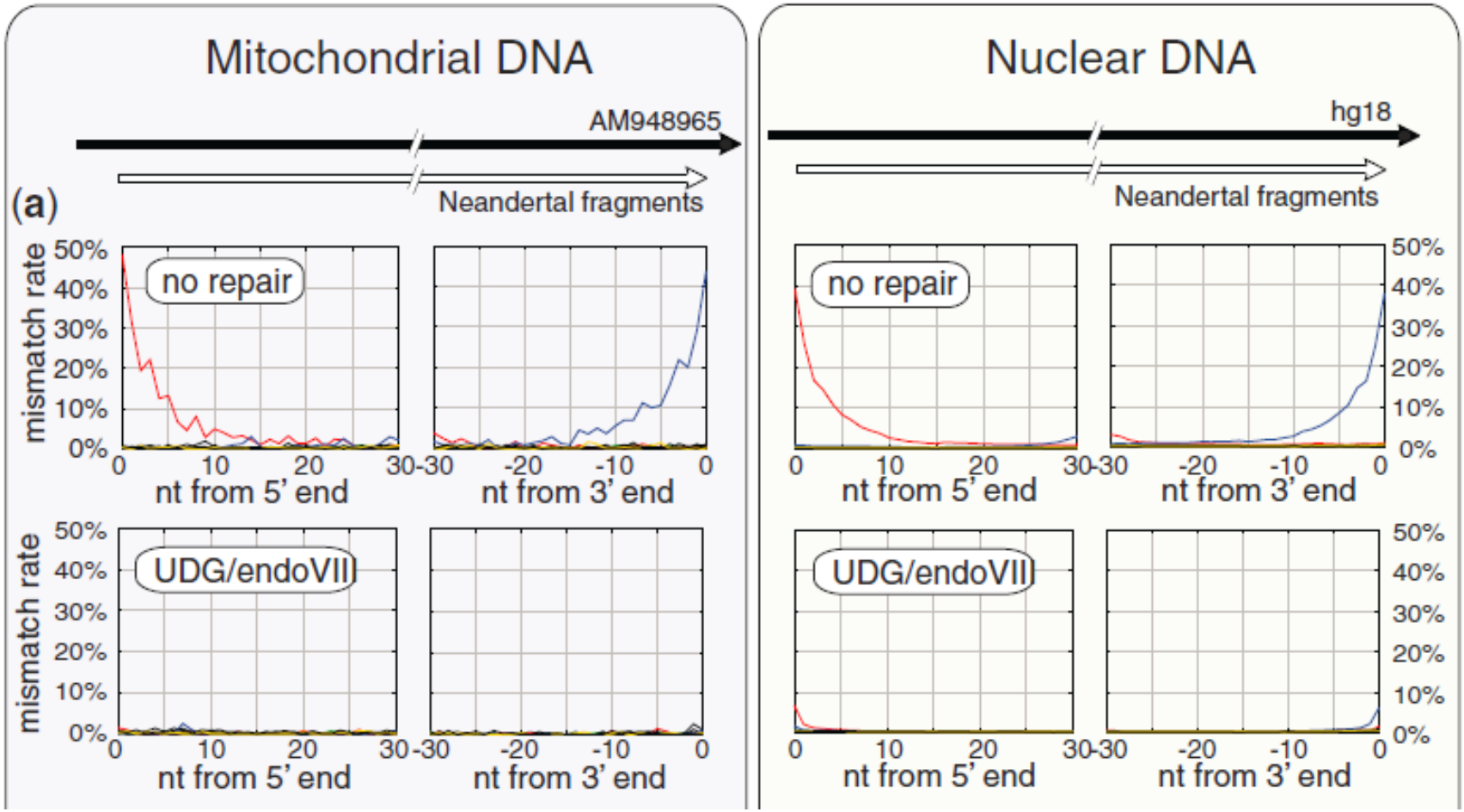
Miscoding lesions in PCR data



High frequency at fragment ends

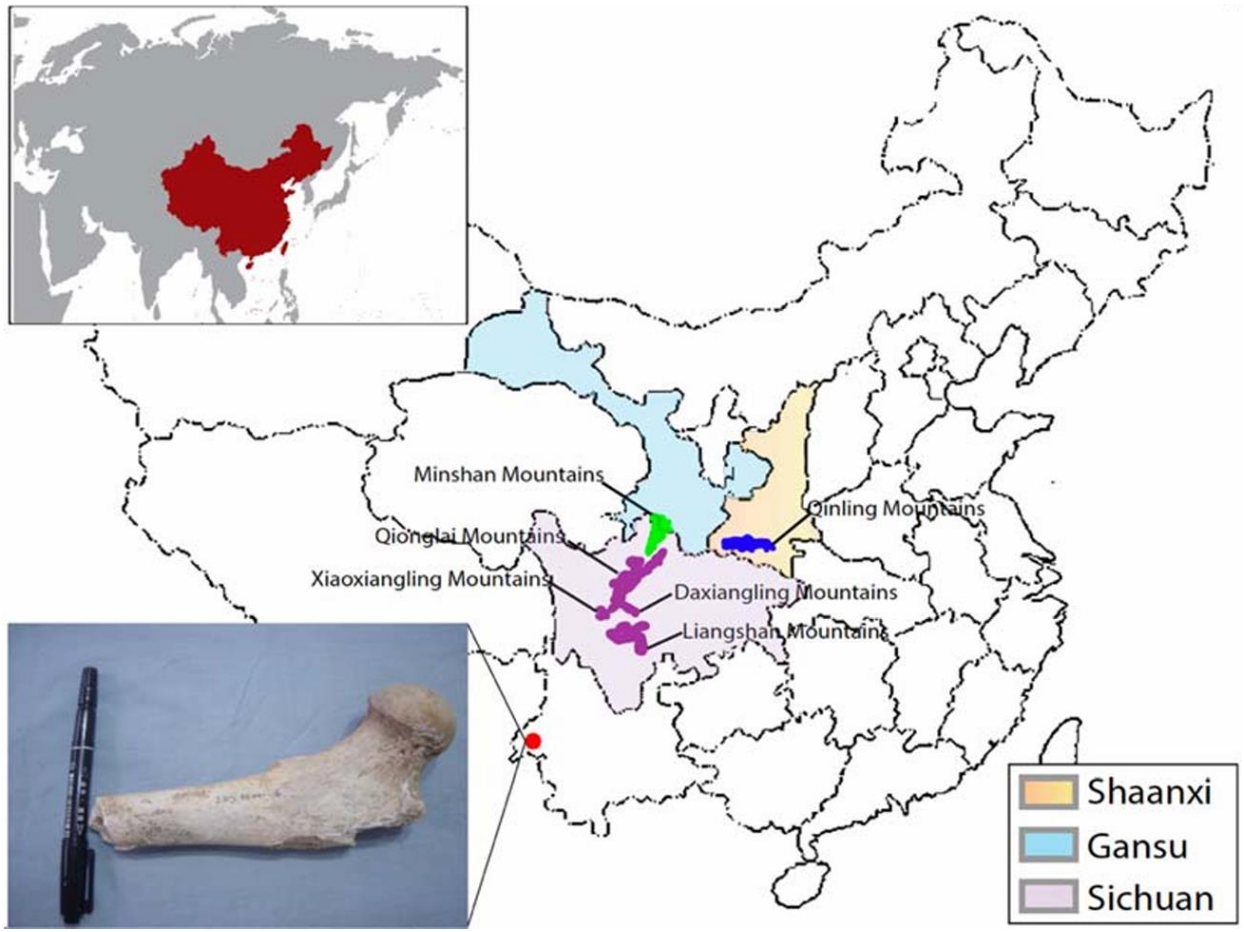


And enzymatic removal

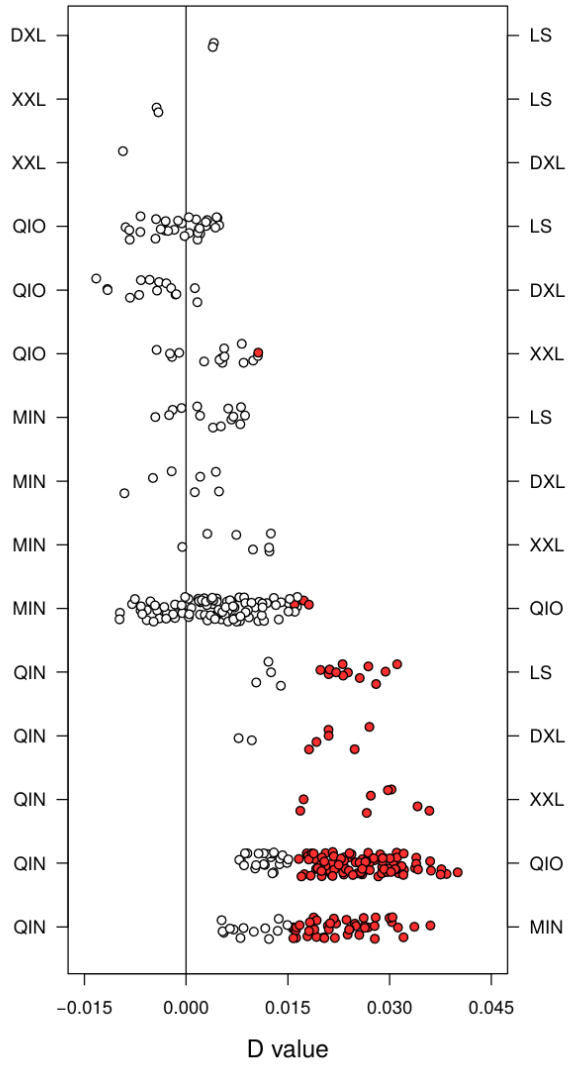


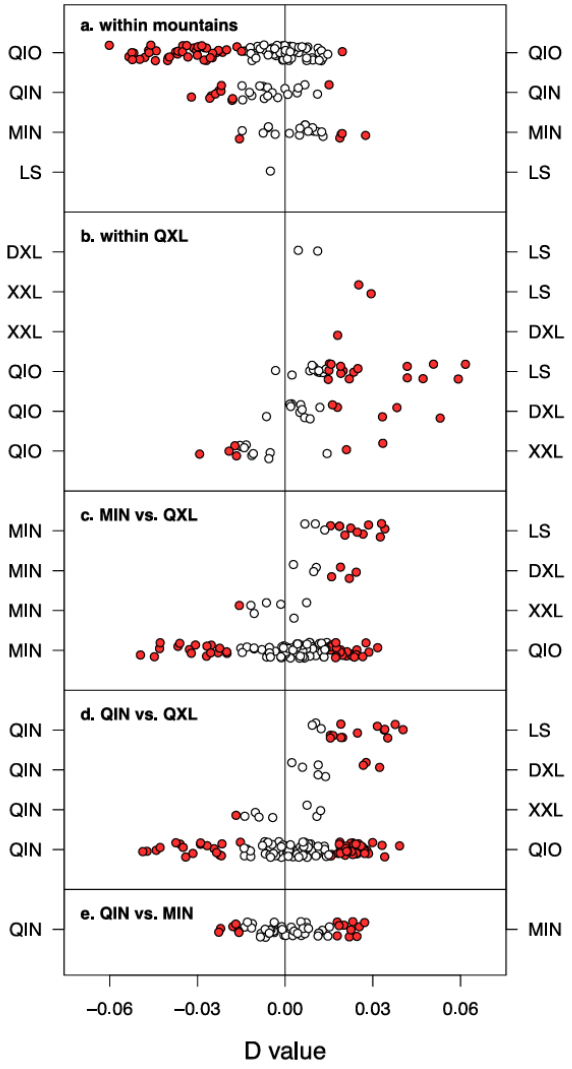
Mapping bias

Fossil giant panda

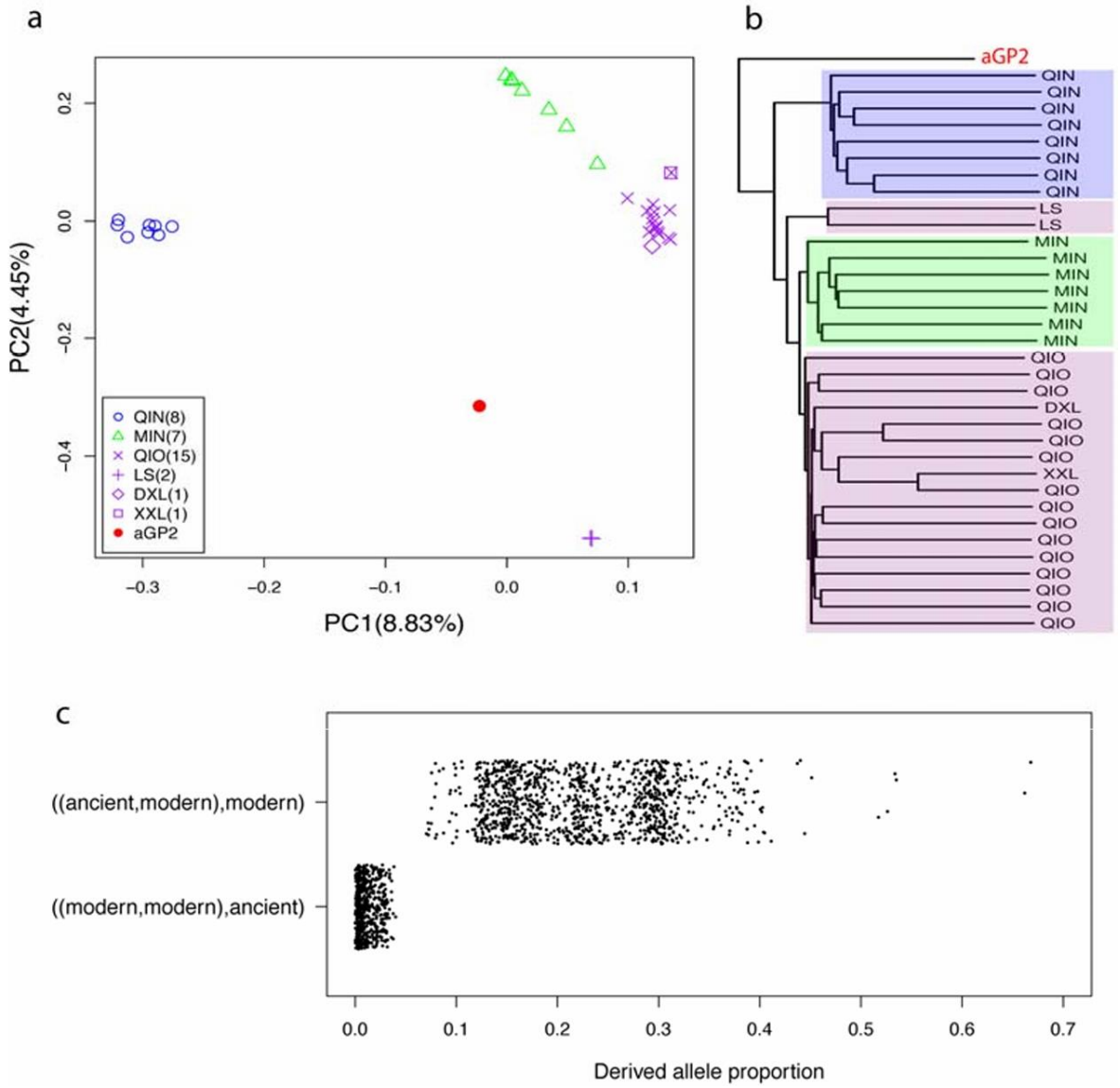


Ingroup bias

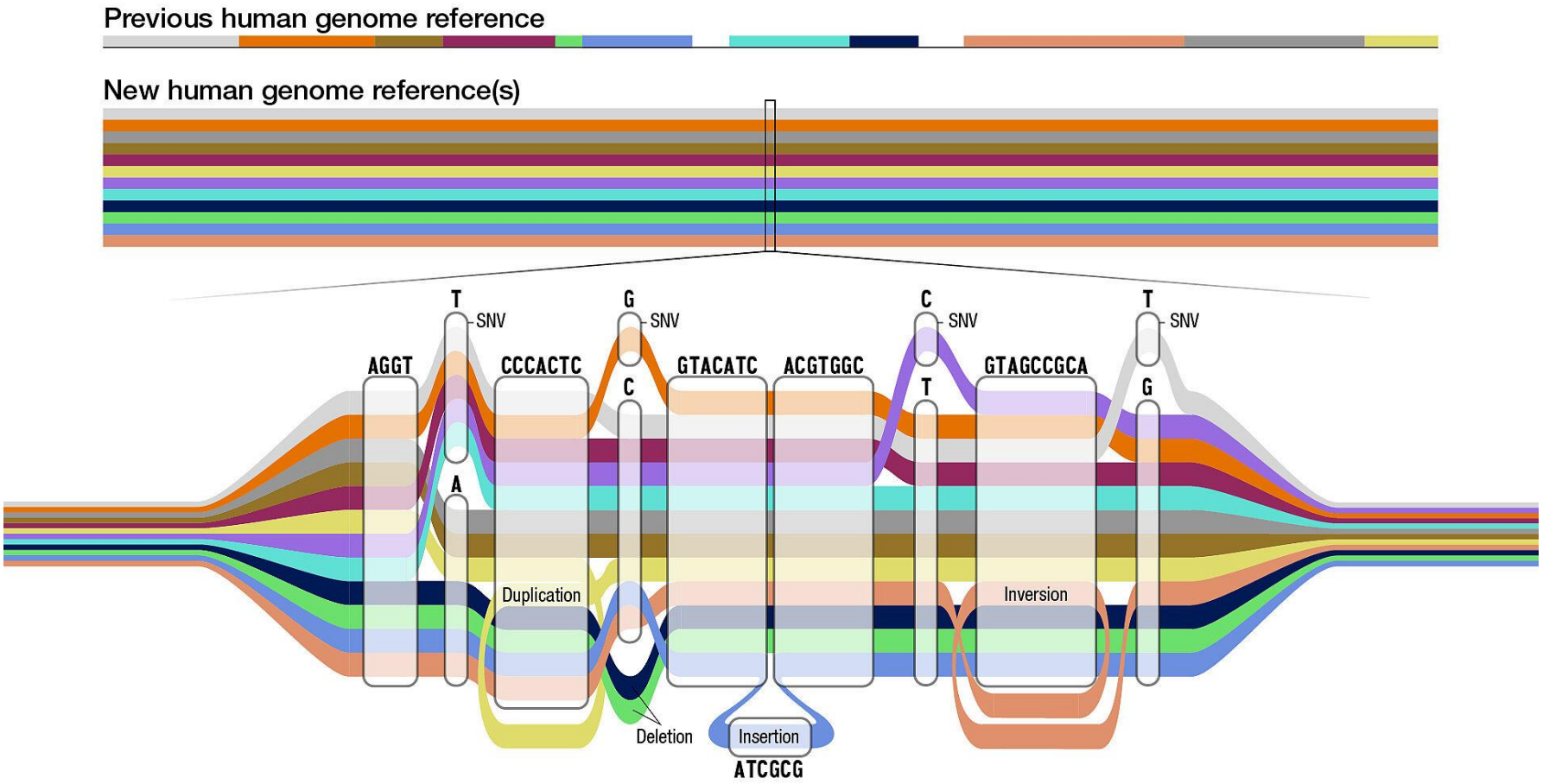




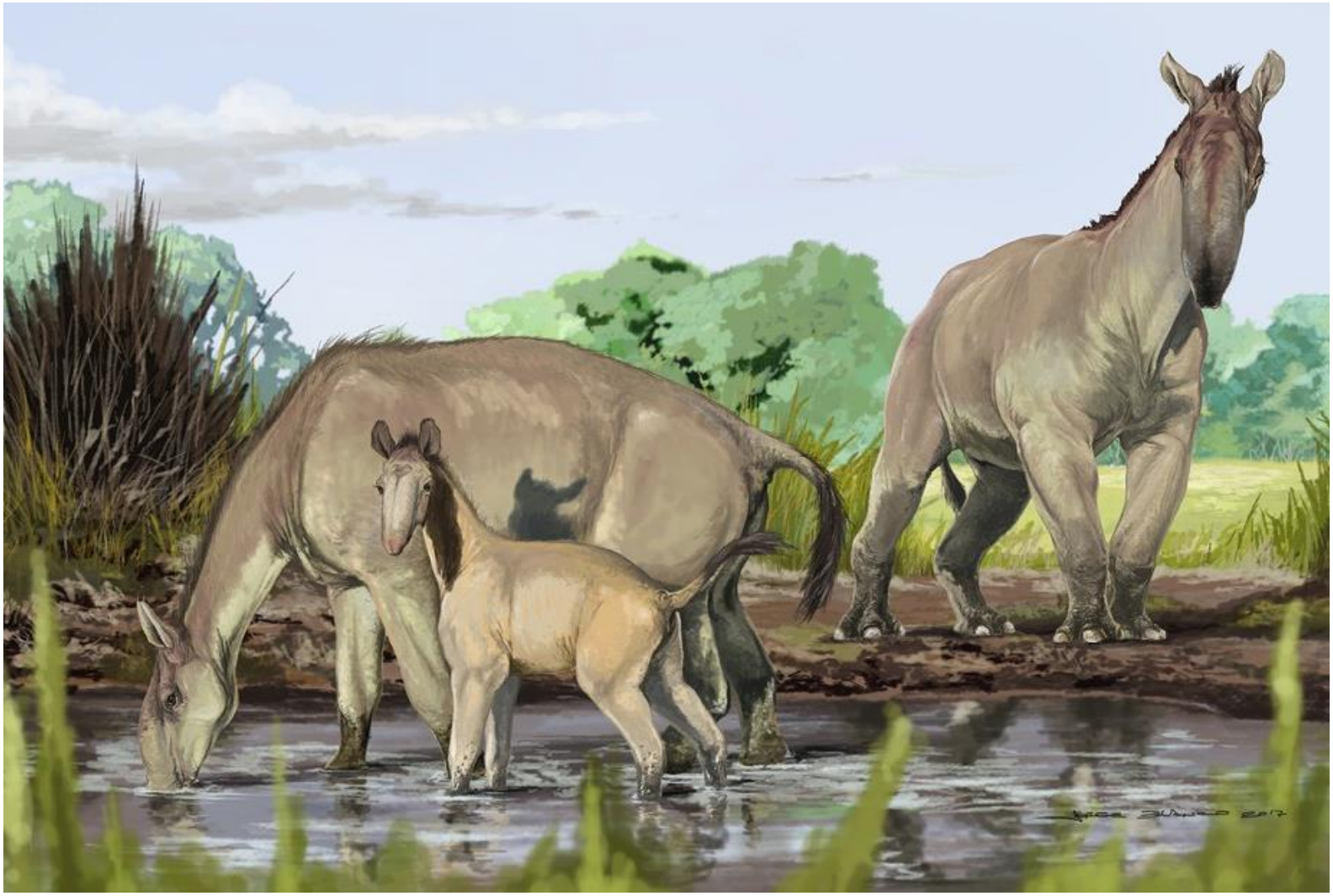
Outgroup bias?



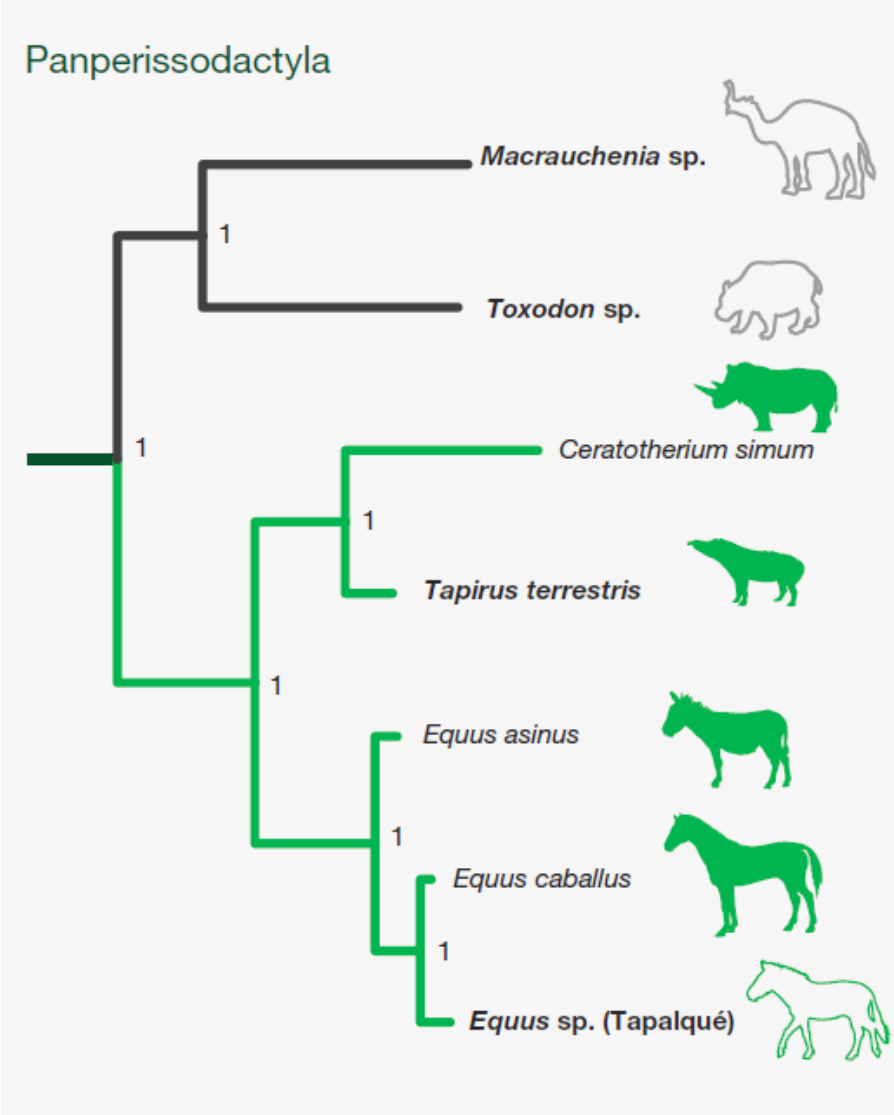
Solution: pangenome mapping?



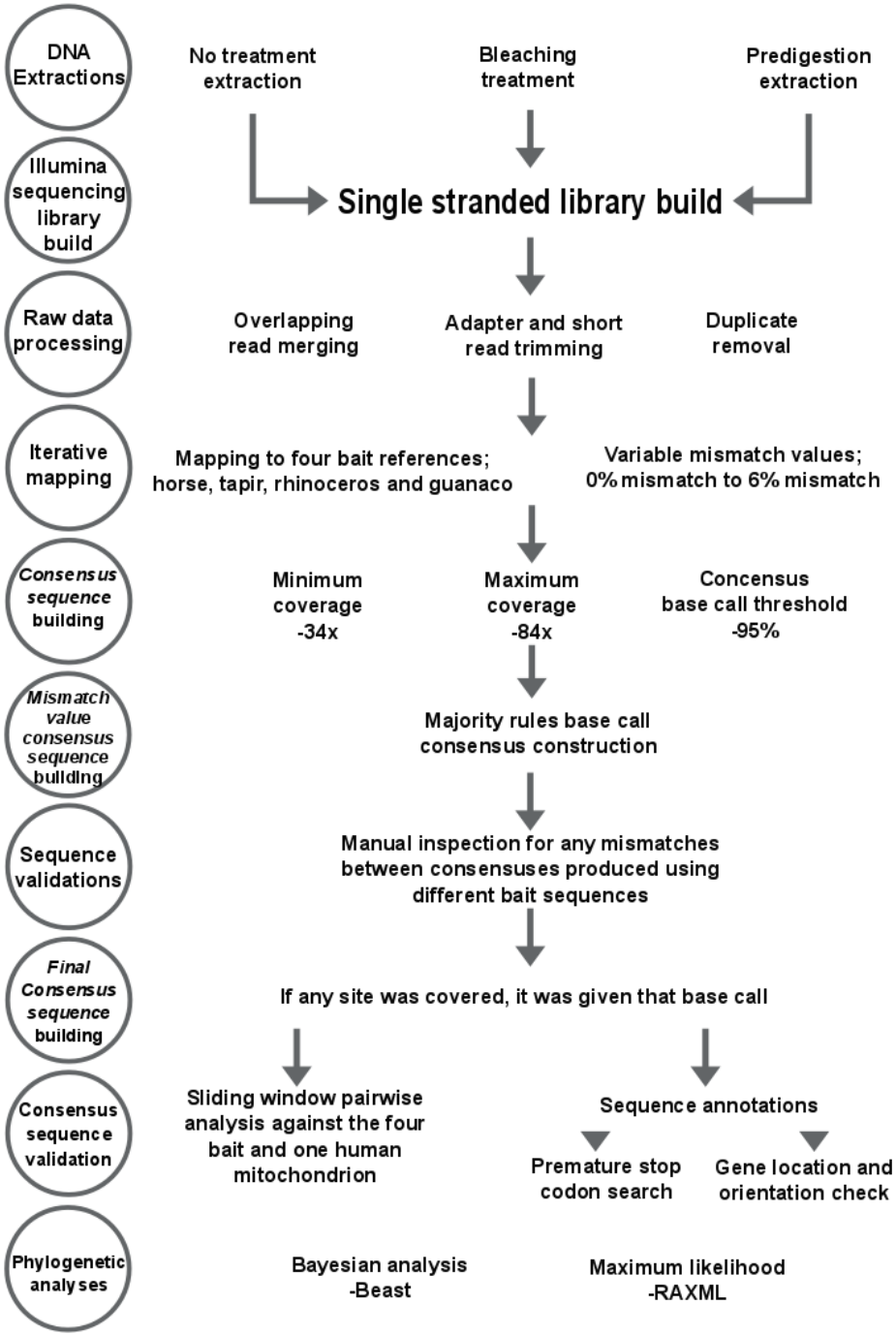
Projects on the edge of the possible



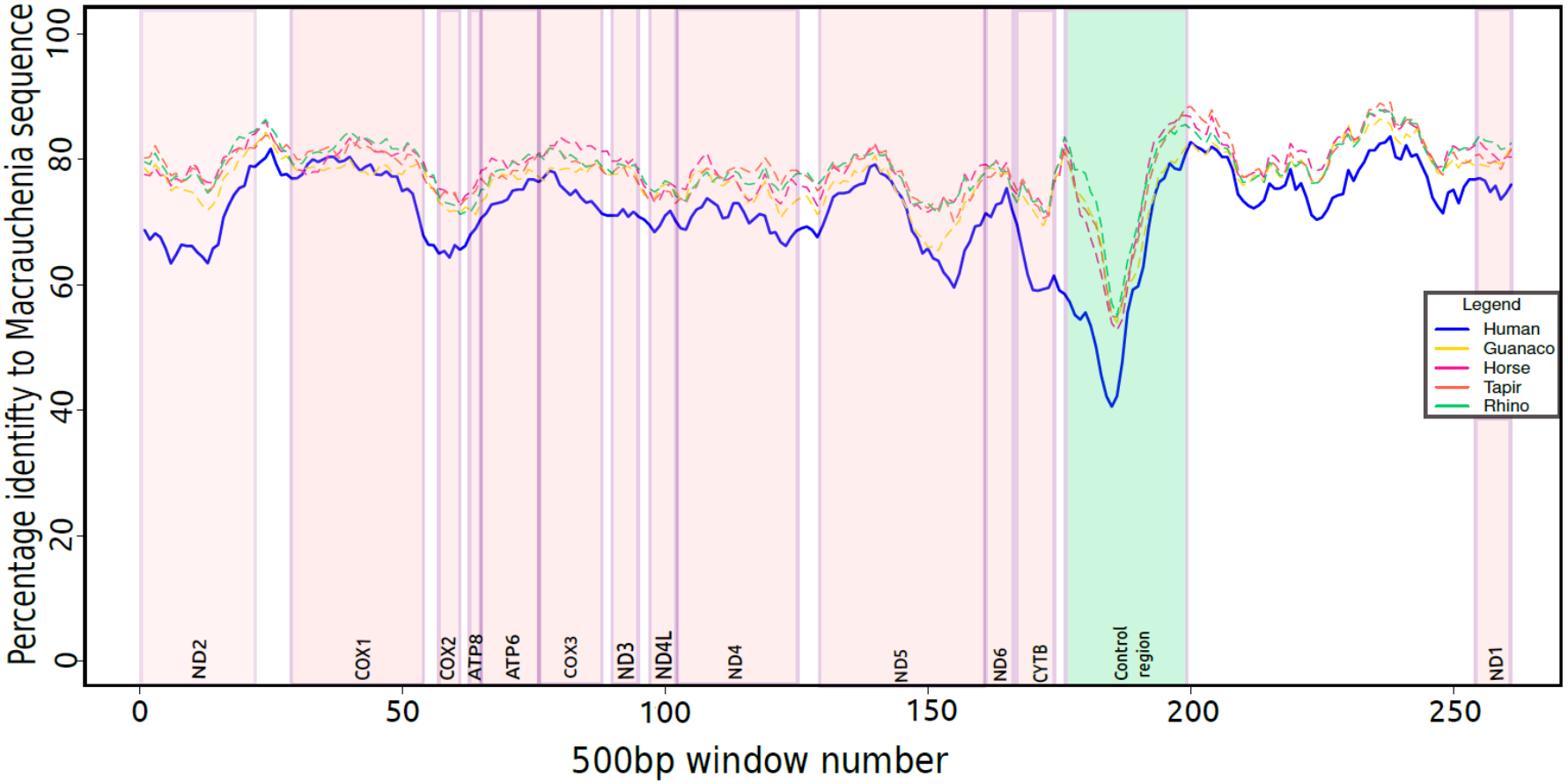
Phylogenetic distance



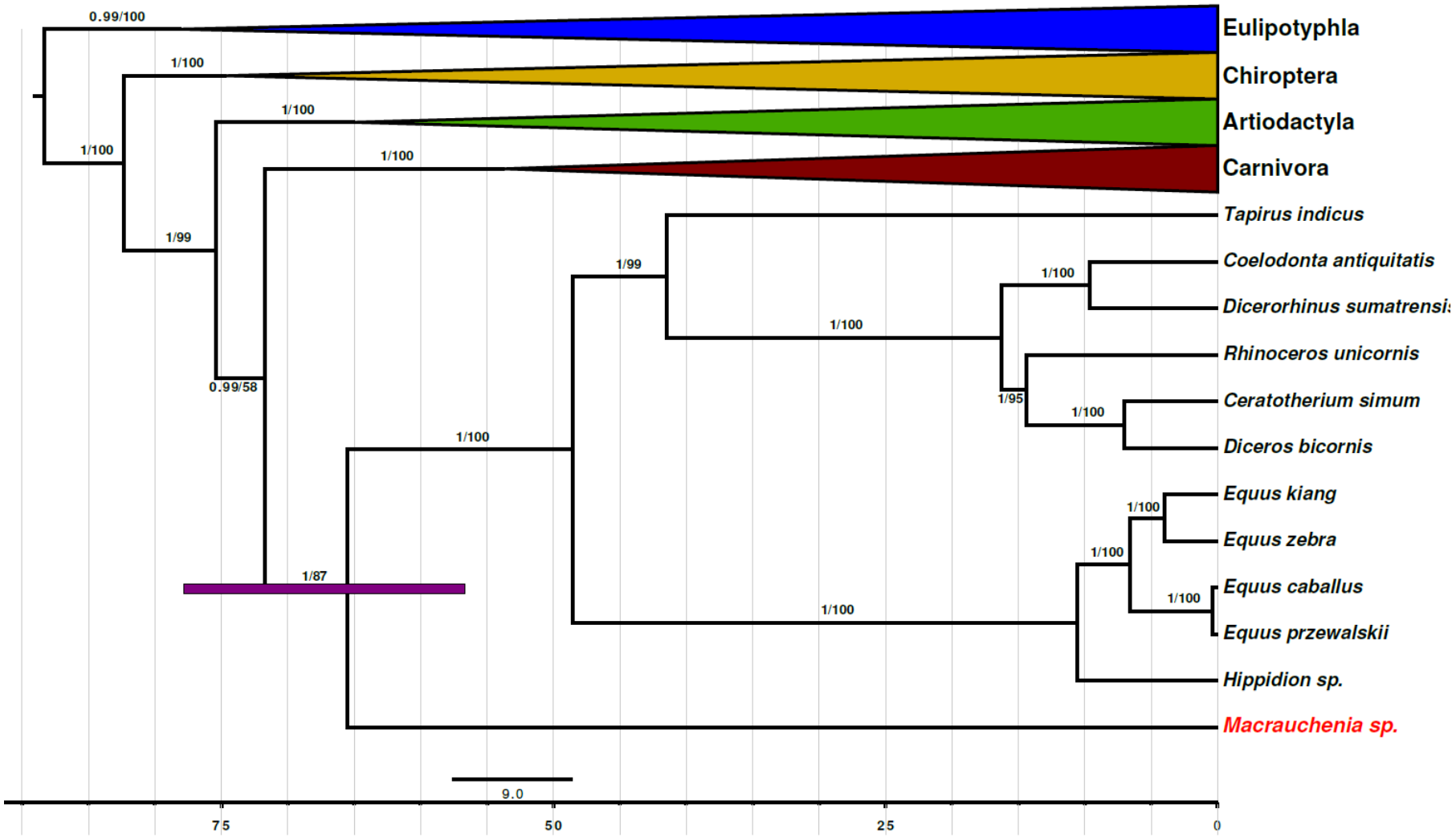
Mostly from warm to hot climate


DNA Extractions
No treatment extraction
Bleaching treatment
Predigestion extraction
Illumina sequencing library build
Single stranded library build
Raw data processing
Overlapping read merging
Adapter and short read trimming
Duplicate removal
Iterative mapping
Mapping to four bait references; horse, tapir, rhinoceros and guanaco
Variable mismatch values; 0% mismatch to 6% mismatch
Consensus sequence building
Minimum coverage -34x
Maximum coverage -84x
Concensus base call threshold -95%
Mismatch value consensus sequence building
Majority rules base call consensus construction
Sequence validations
Manual inspection for any mismatches between consensuses produced using different bait sequences
Final Consensus sequence building
If any site was covered, it was given that base call
Consensus sequence validation
Sliding window pairwise analysis against the four bait and one human mitochondrion
Sequence annotations
Premature stop codon search
Gene location and orientation check
Phylogenetic analyses
Bayesian analysis -Beast
Maximum likelihood -RAXML

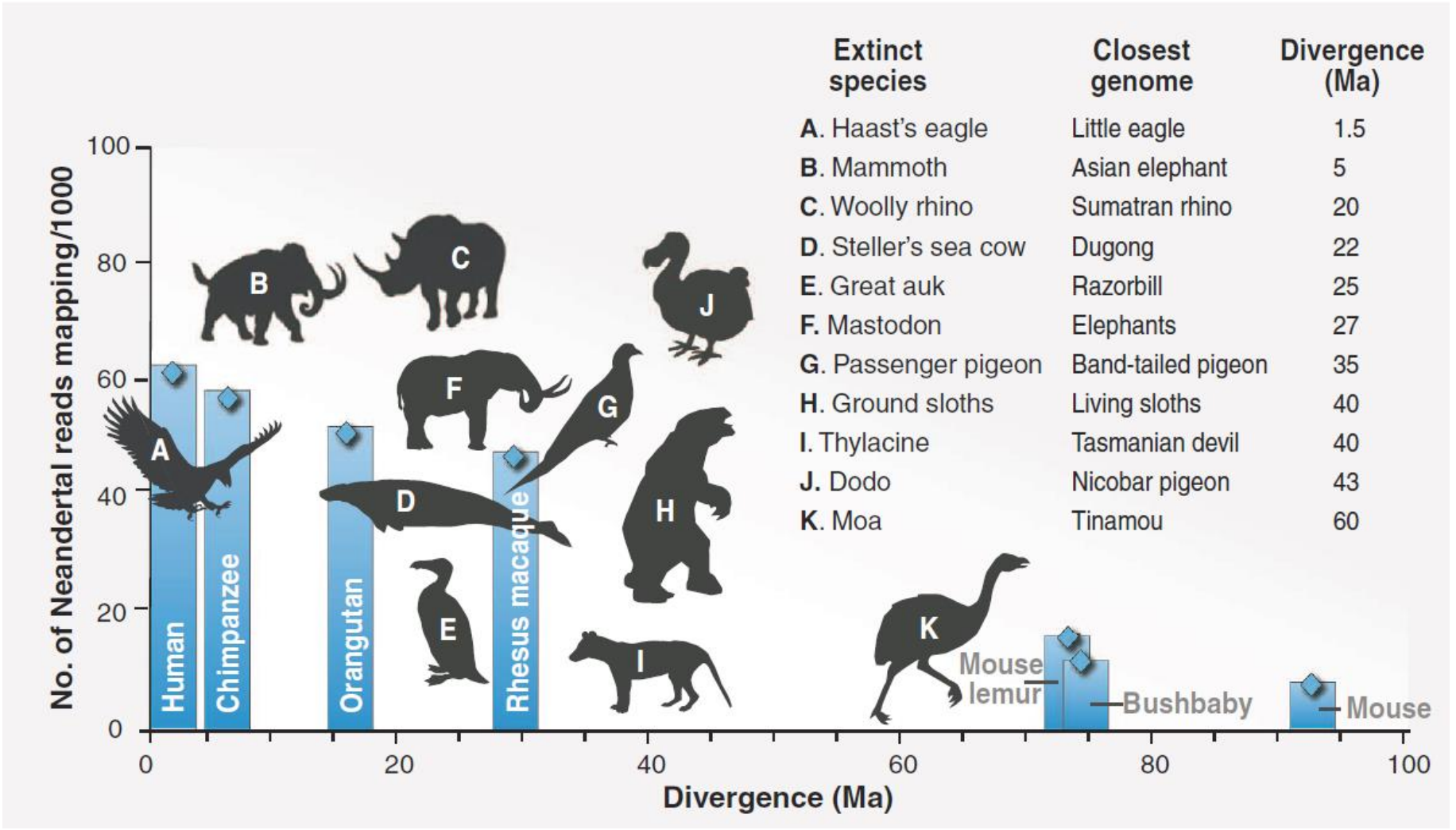
Quality check



Phylogeny makes sense (whatever that means)



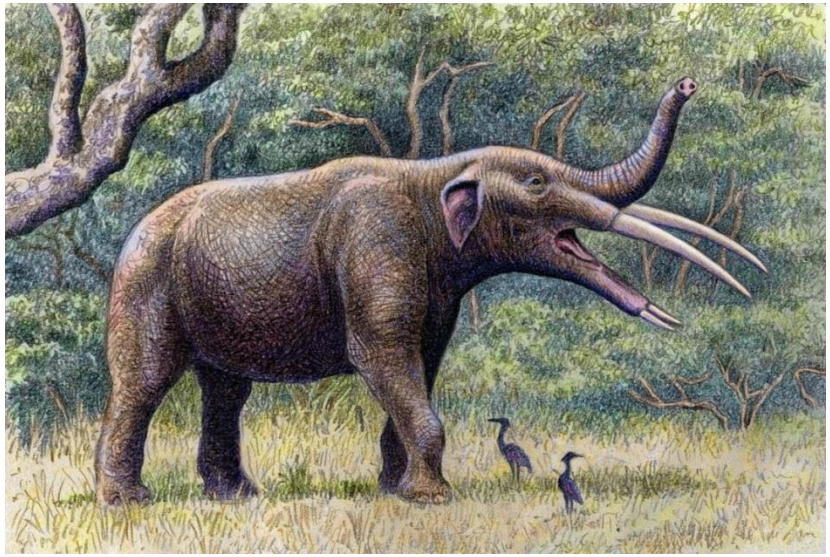
Would a nuclear genome be possible?



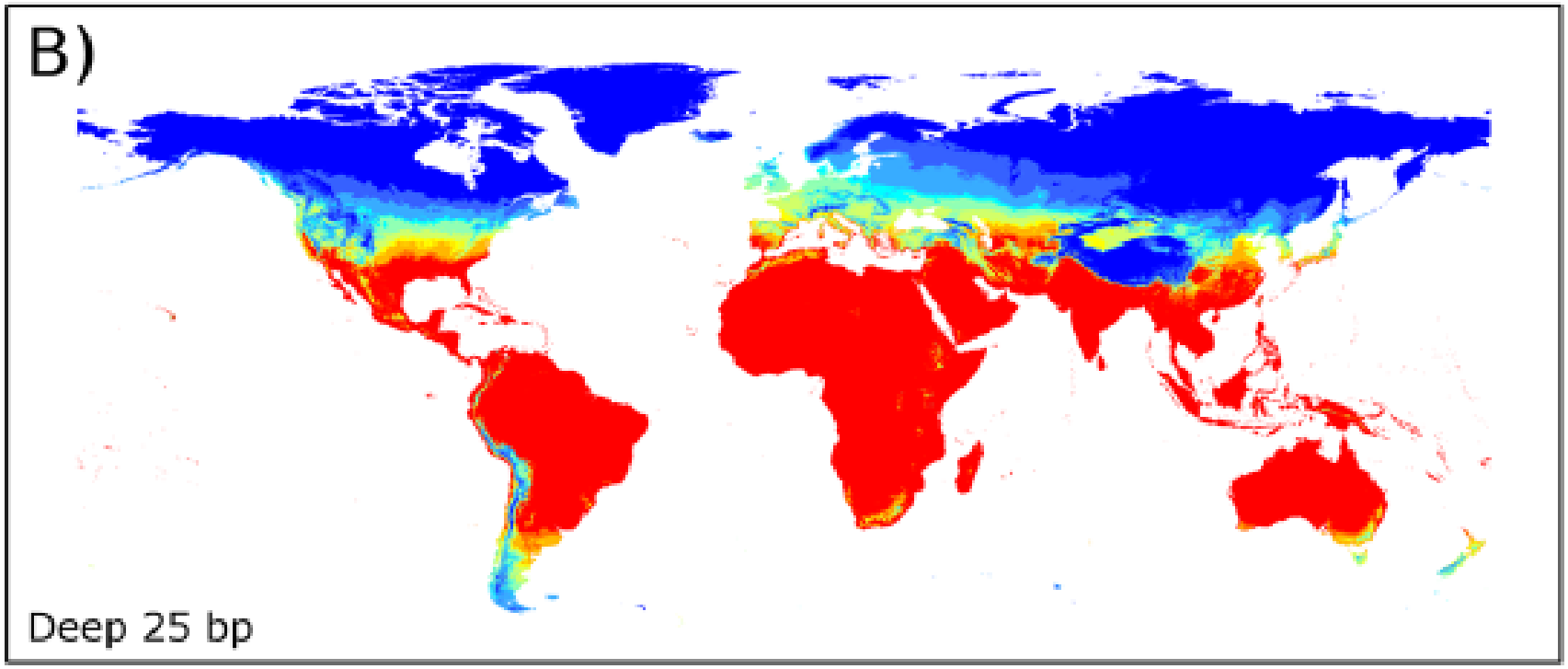
Thanks……

……..for listening so far

Part 2: What to do with ancient DNA

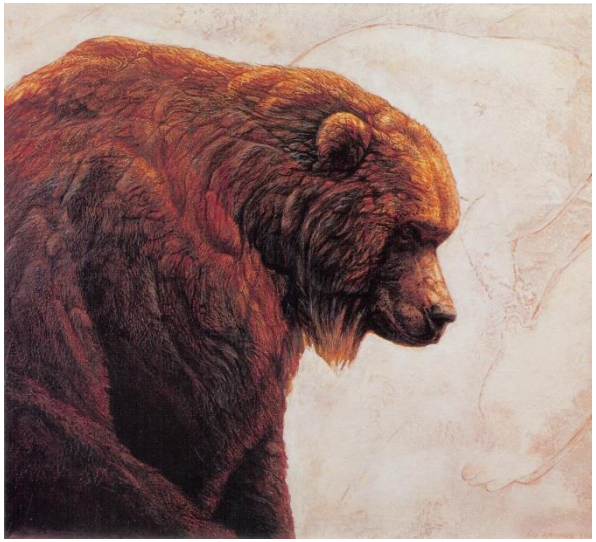


Choose samples from colder regions



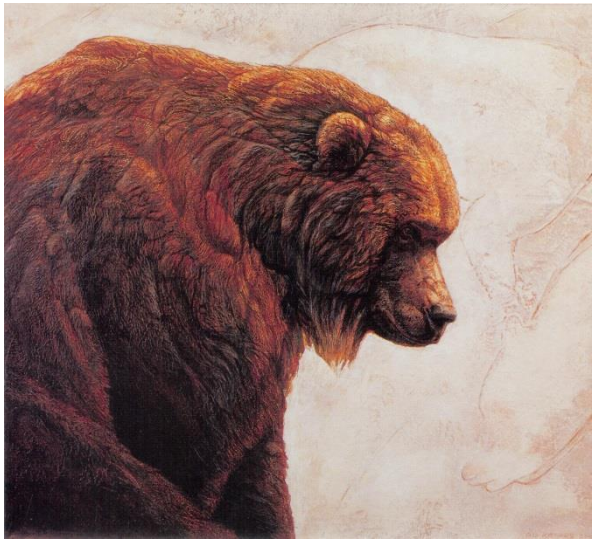
Cave bears

Bear evolution



Bear evolution and gene flow

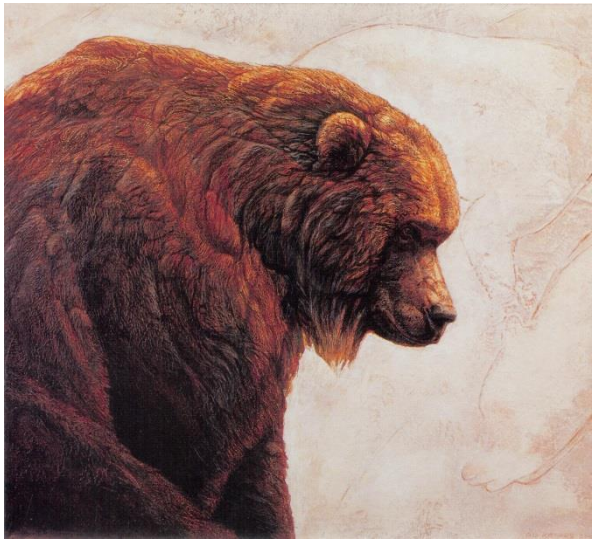
Geneflow



Gene flow among species

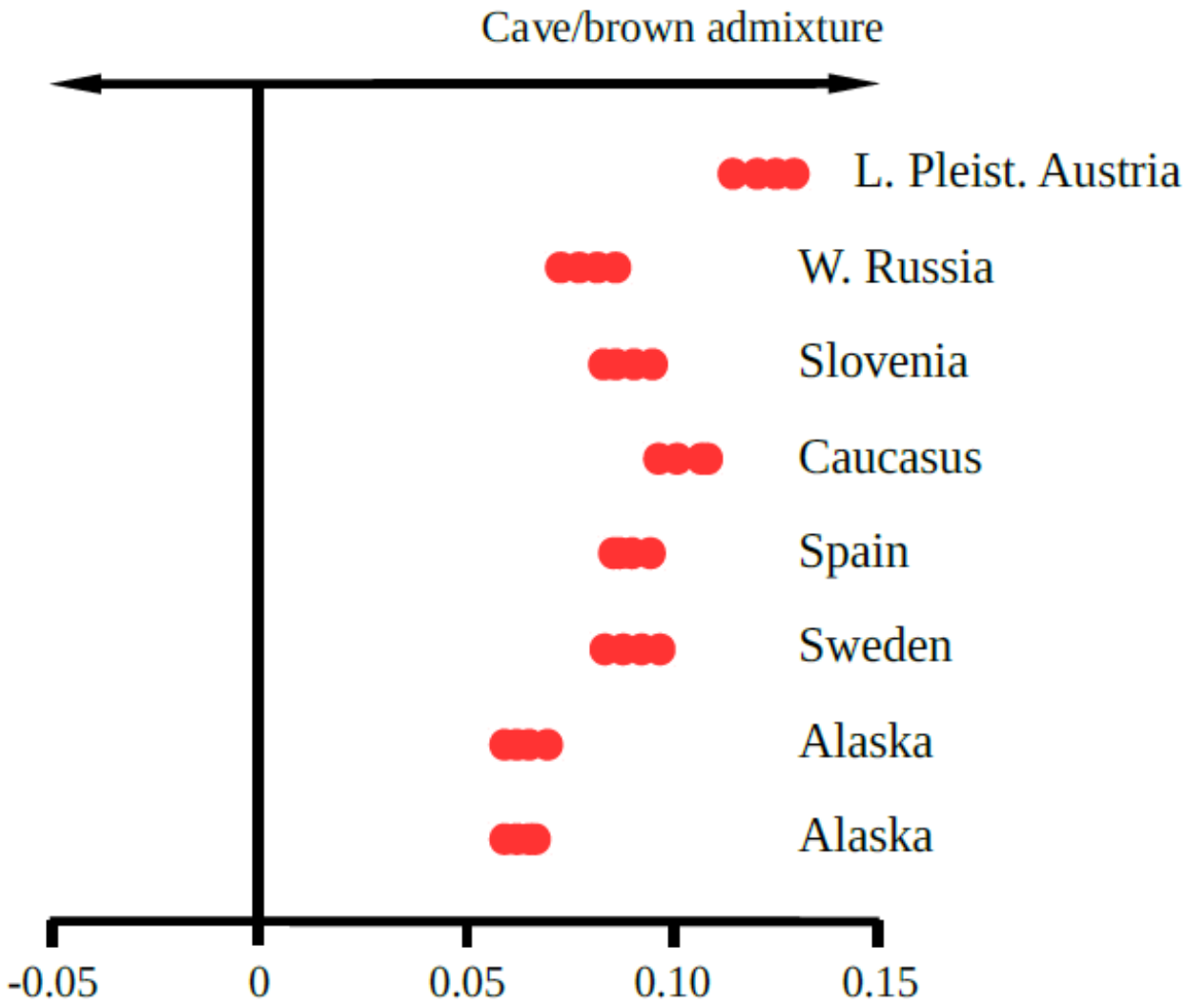
Geneflow

Geneflow?

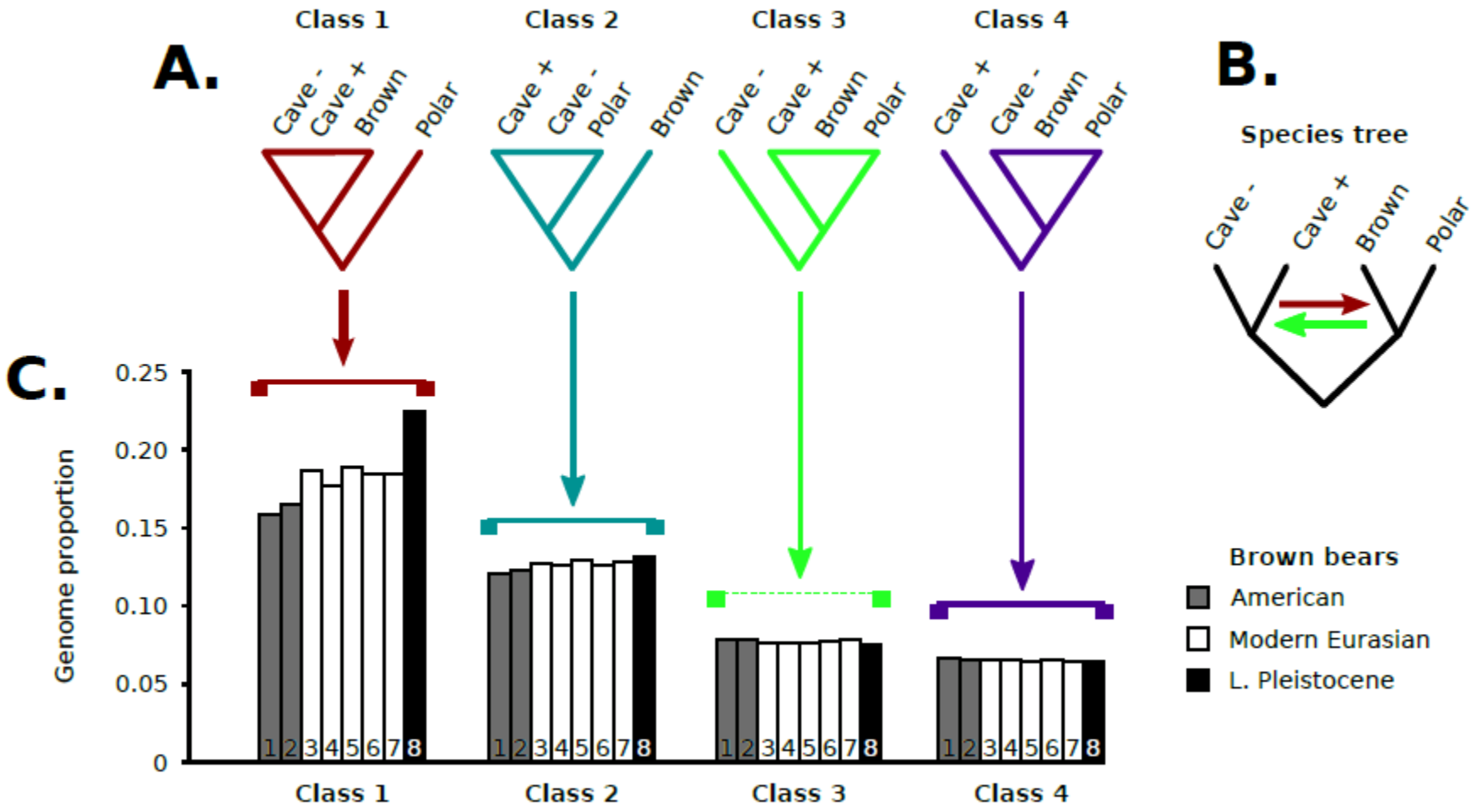


Geneflow?

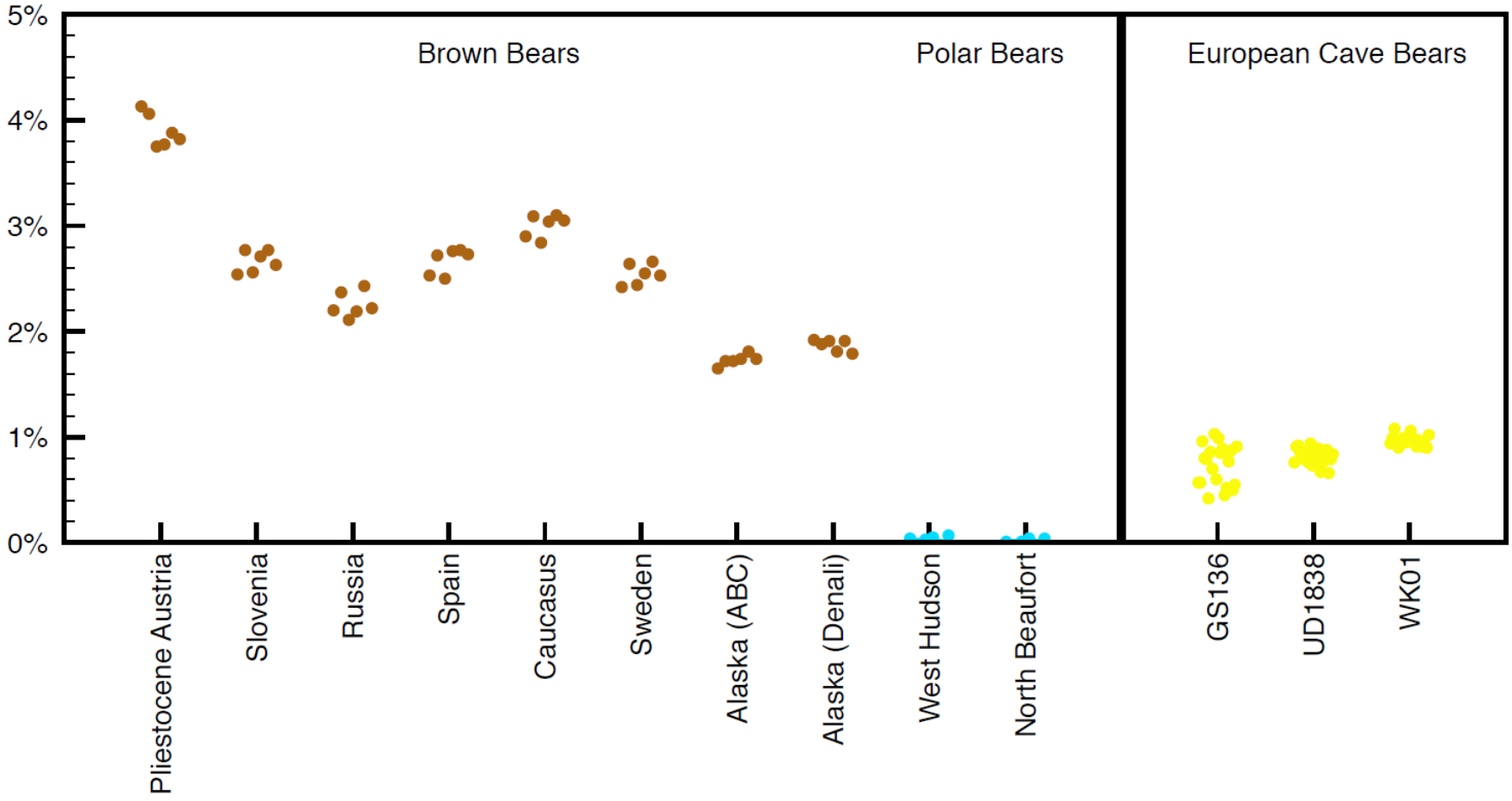
D-statistics



Phylogenetic admixture analysis

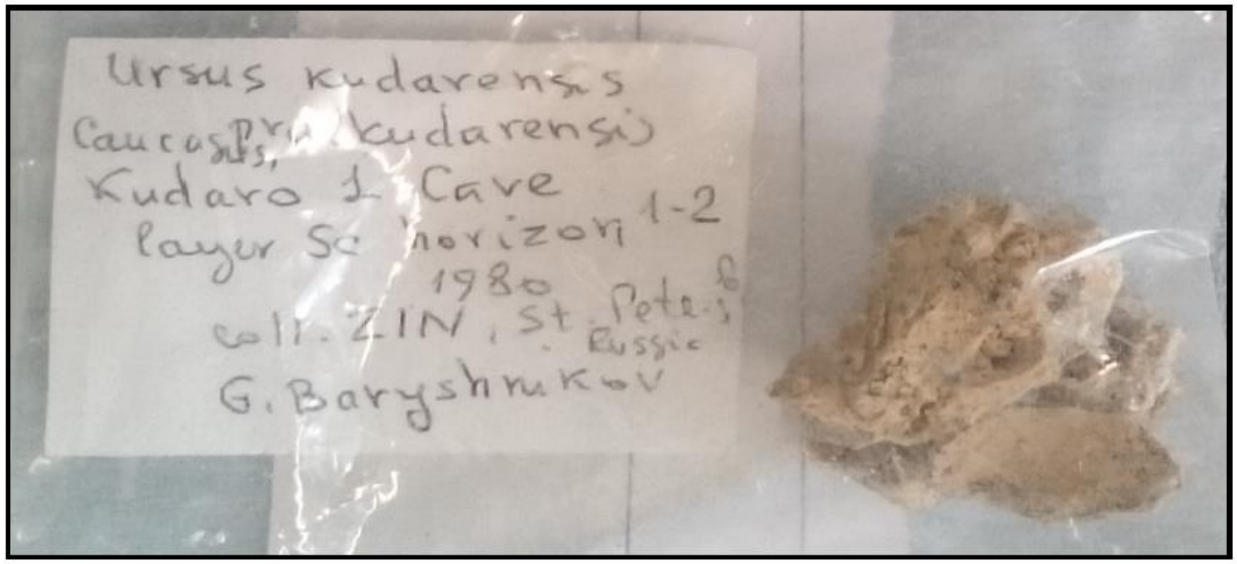


Percentage of admixture



Caucasus

Caucasus cave bears



U. kudarensis praekudarensis

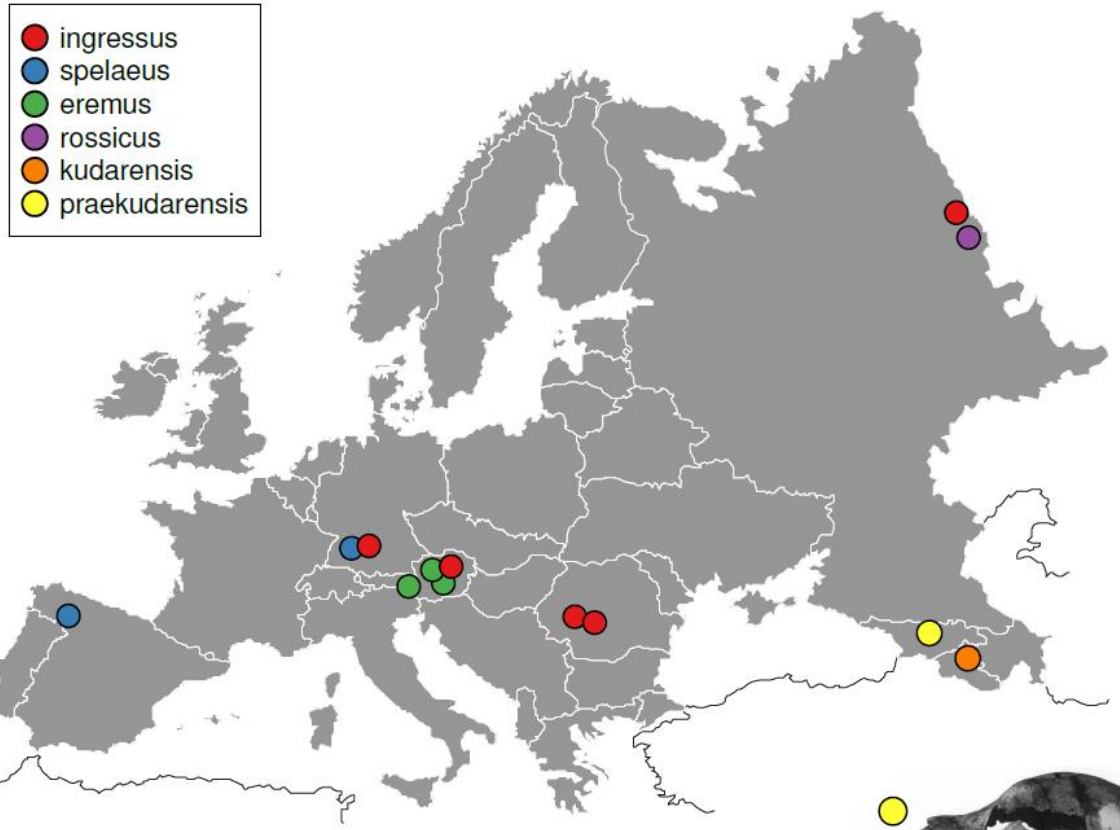
Kudaro 1 cave, Caucasus

Layer 5c, horizon 1-2

Oxygen-Isotope Stage 10

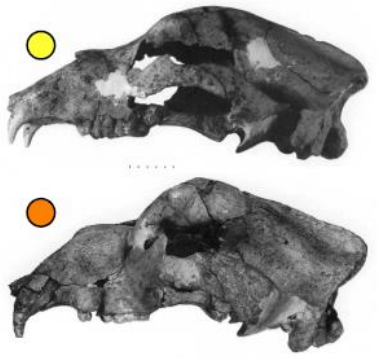
360,000 ± 90,000 years

European cave bears

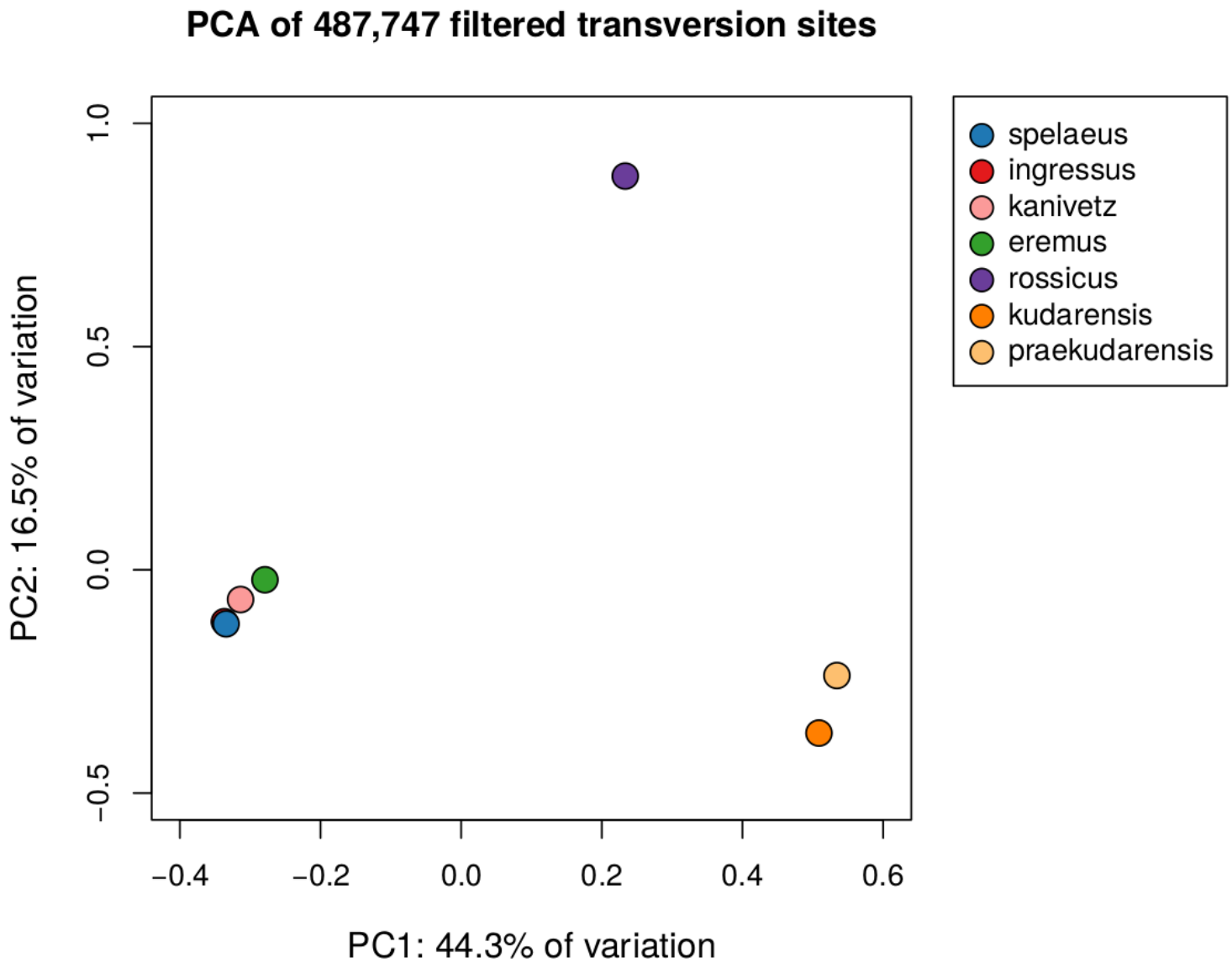


Cave Bears from the Paleolithic of the Greater Caucasus

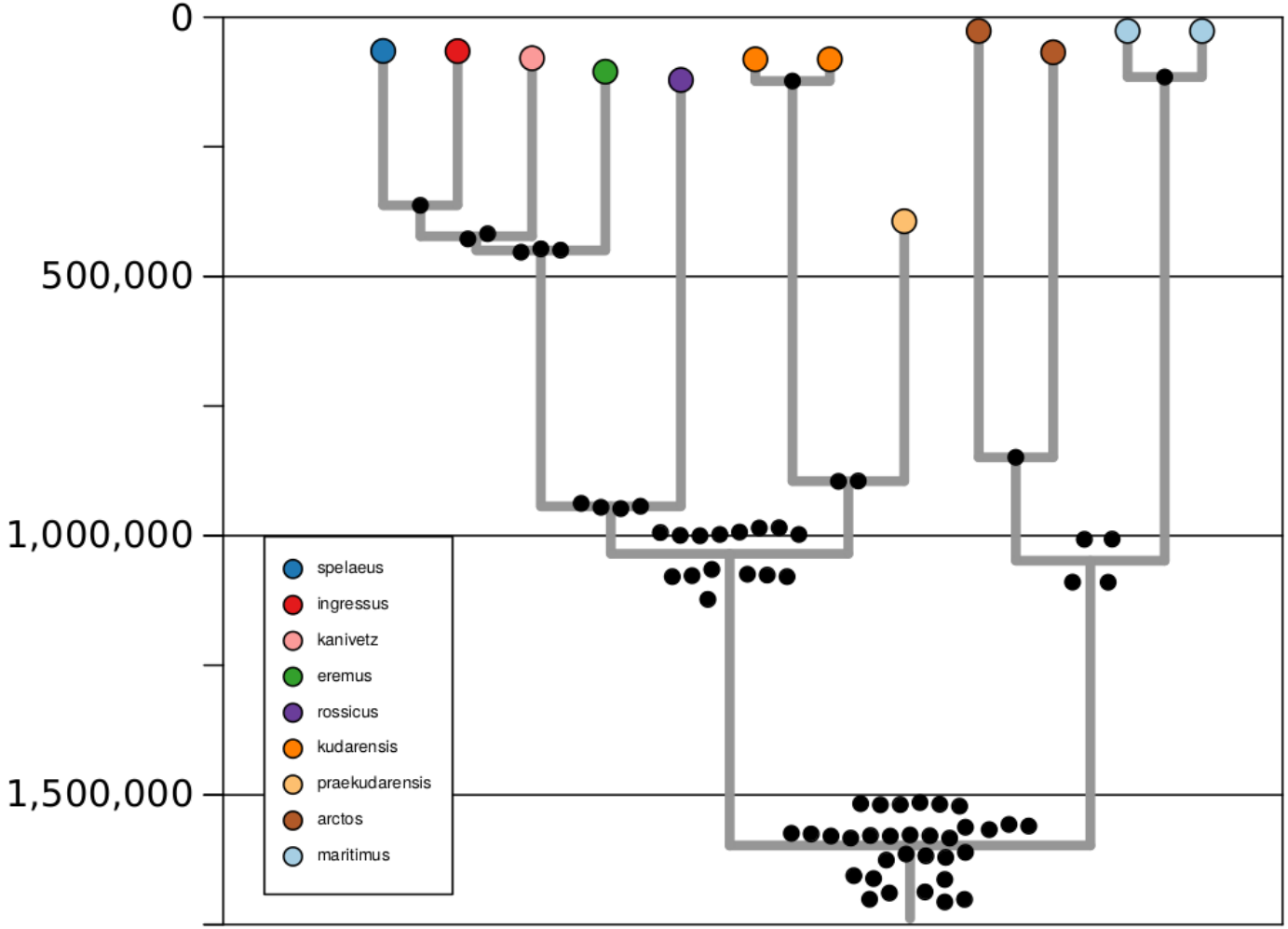
Gennady F. Baryshnikov
Zoological Institute, Russian Academy of Sciences, St. Petersburg



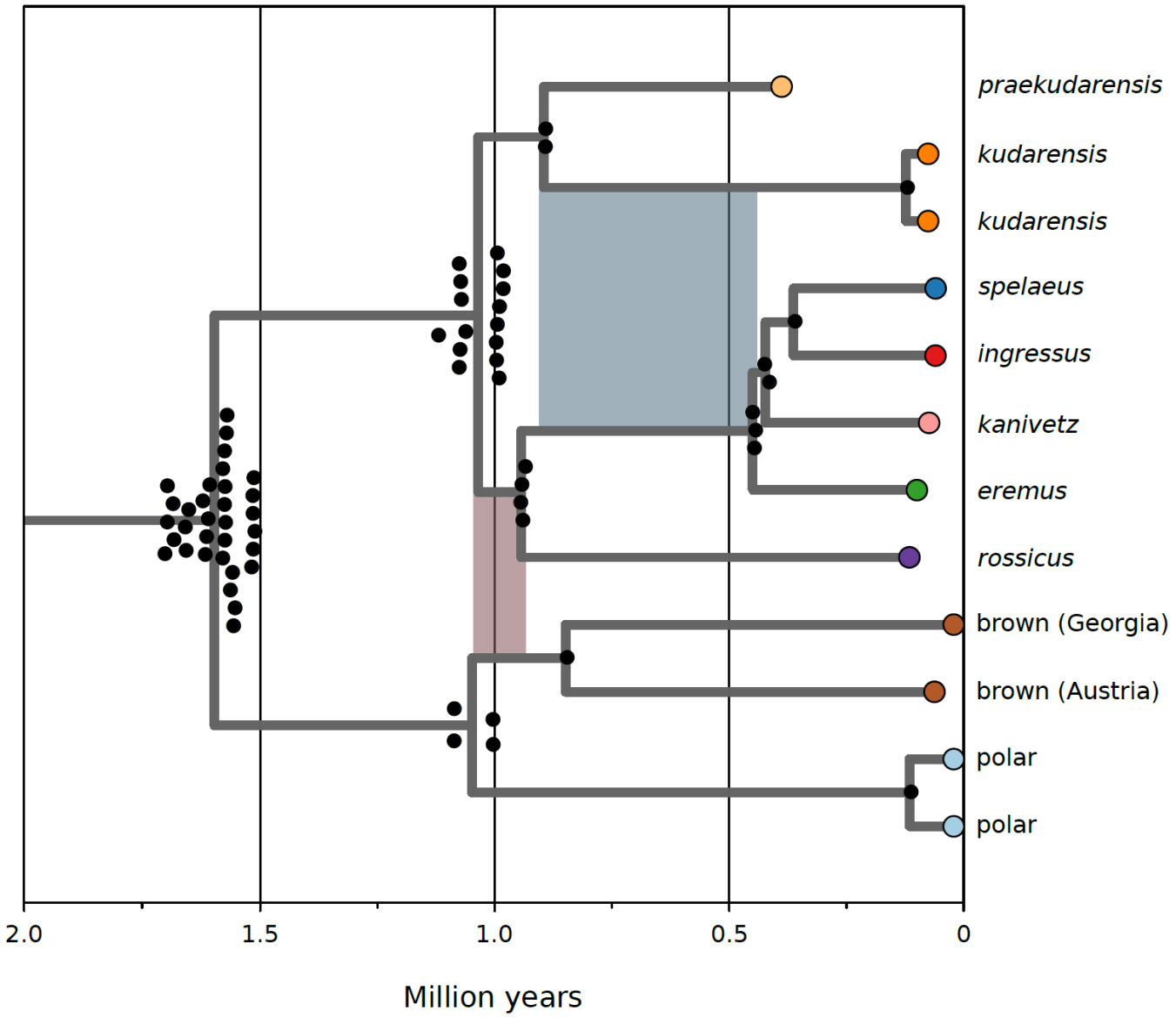
Cave bear nuclear relationships……



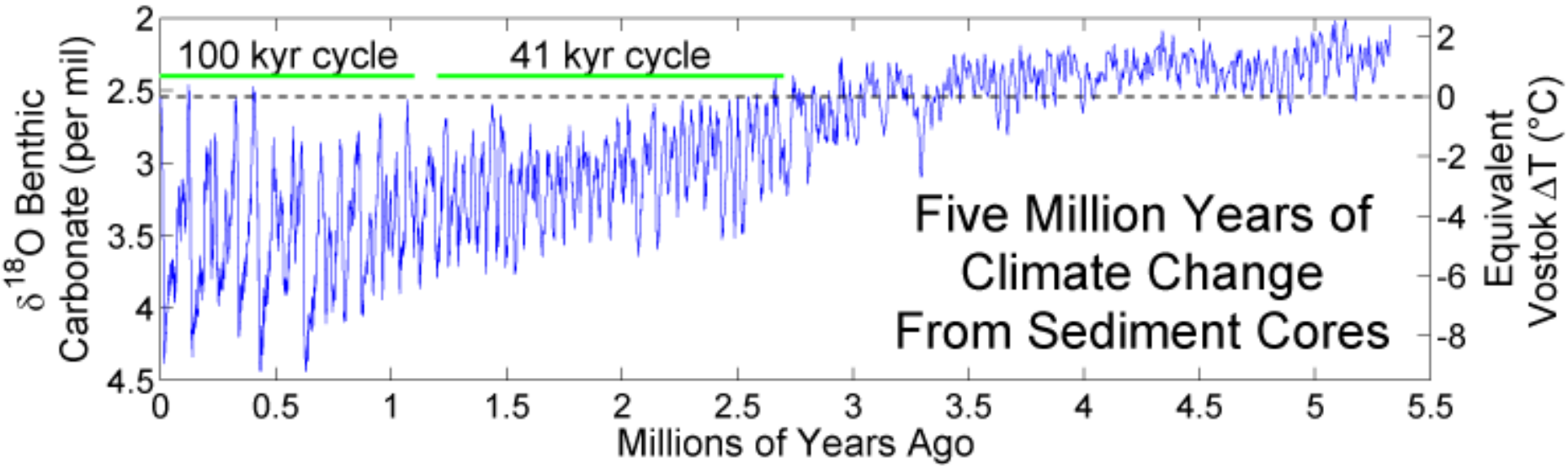
……and bear evolution from a temporal perspective



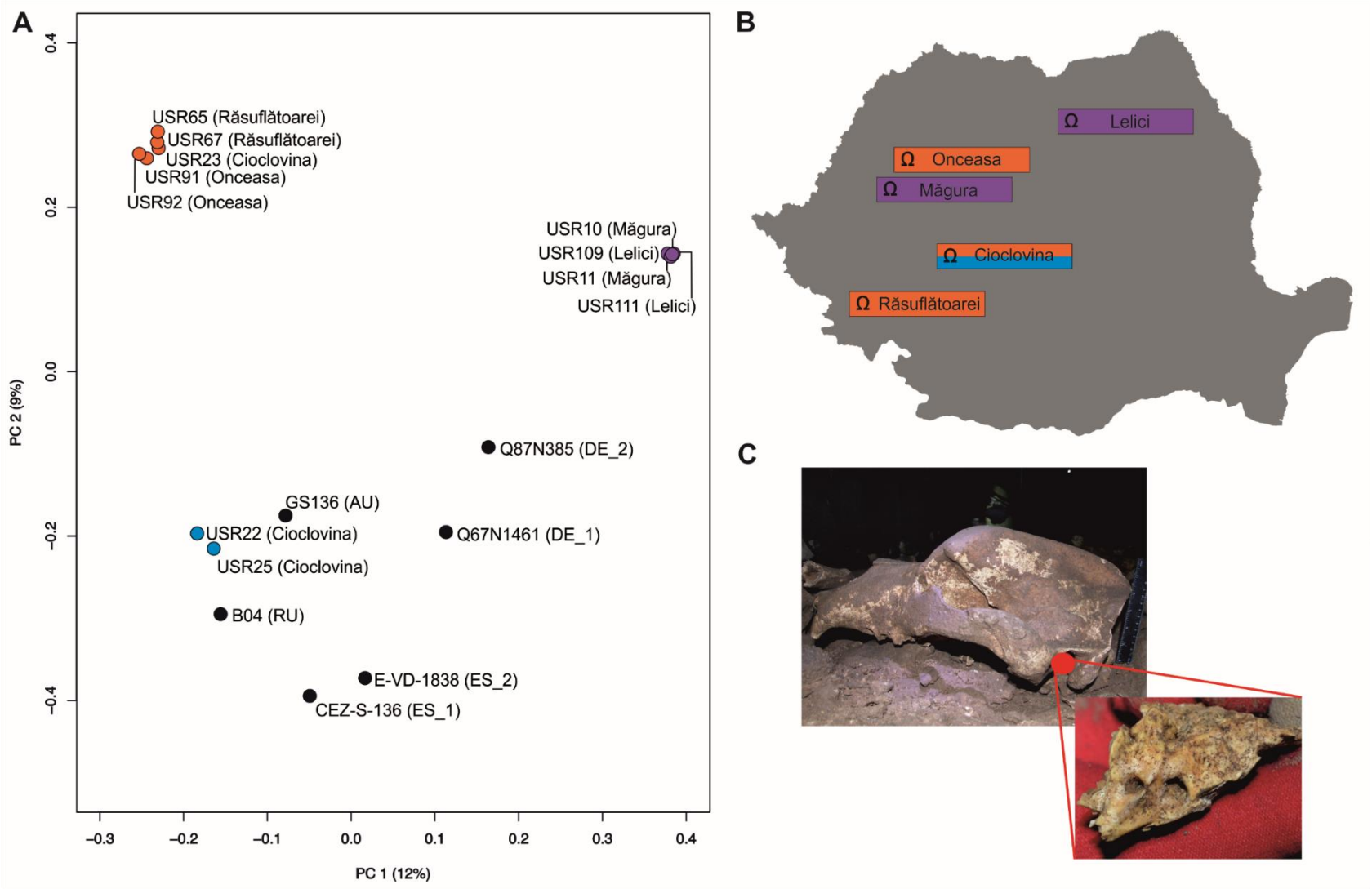
Dating admixture – 500,000 years after divergence



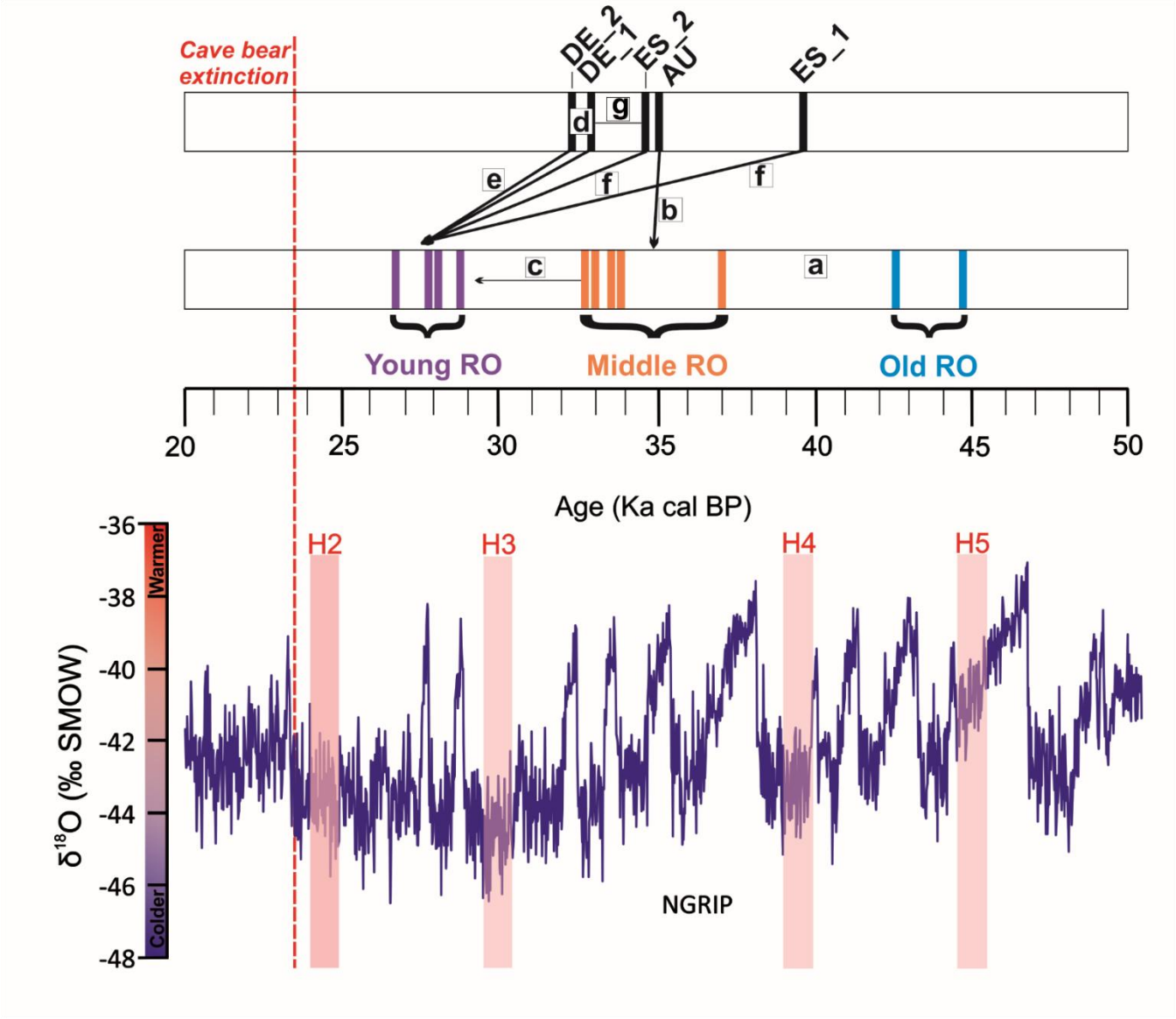
Natural climate change causing divergence?



Temporal genomic differentiation in Romania



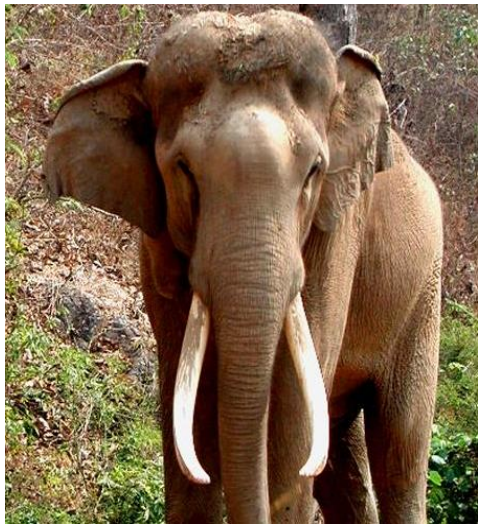
And cave bear population dynamics

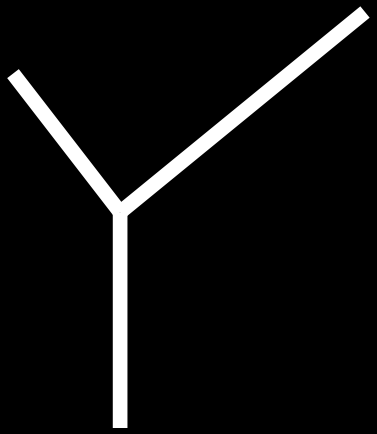


Summary cave bears

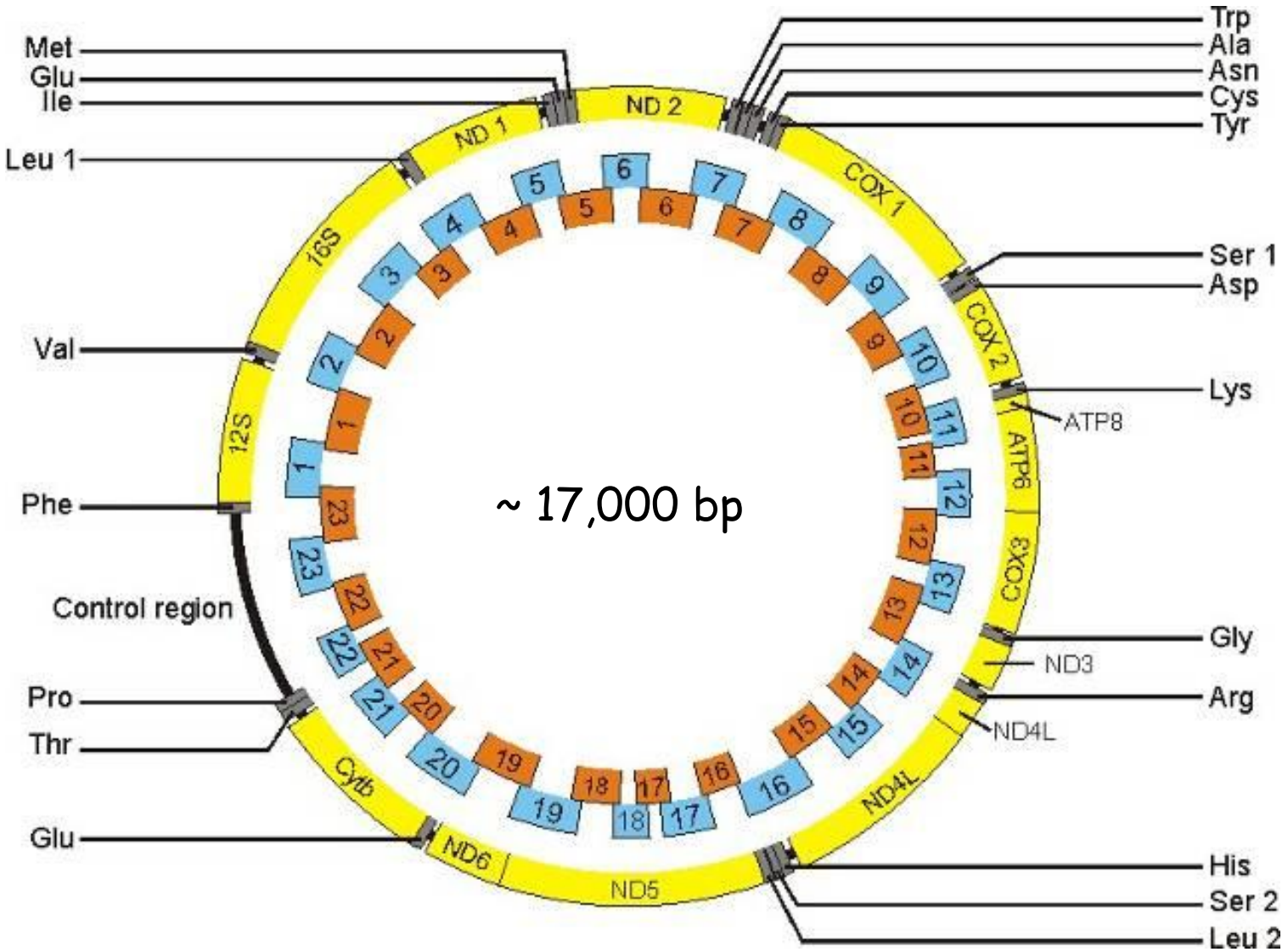
- Gene flow within and between species is an important factor in bear evolution
- The European cave bear has partially survived in the brown bear, at least o the genomic level
- Genomes from very old fossils can be used to calibrate phylogenetic trees
- Cave bears did not like cold climate phases

Elephant evolution



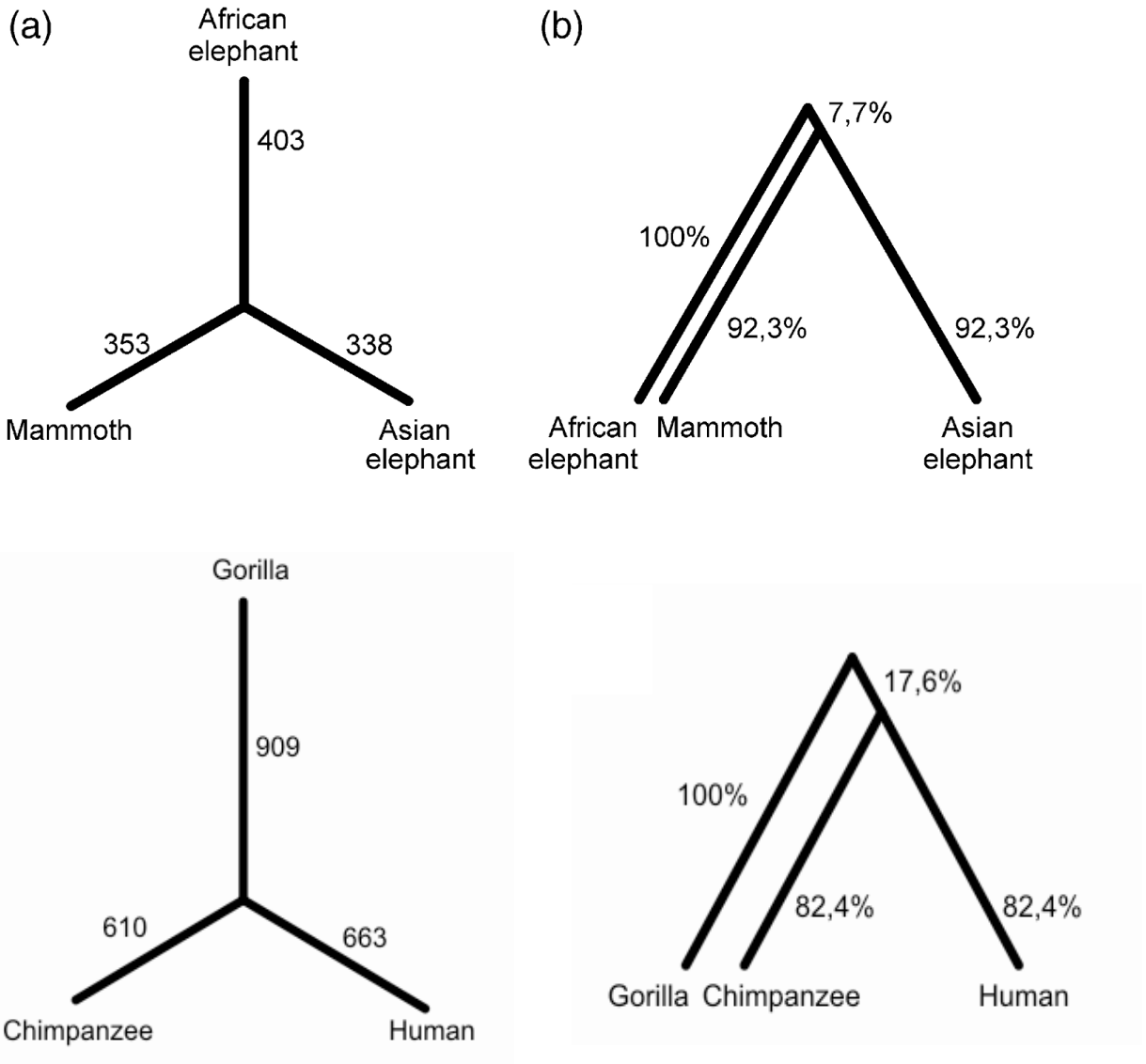


The mitochondrial genome



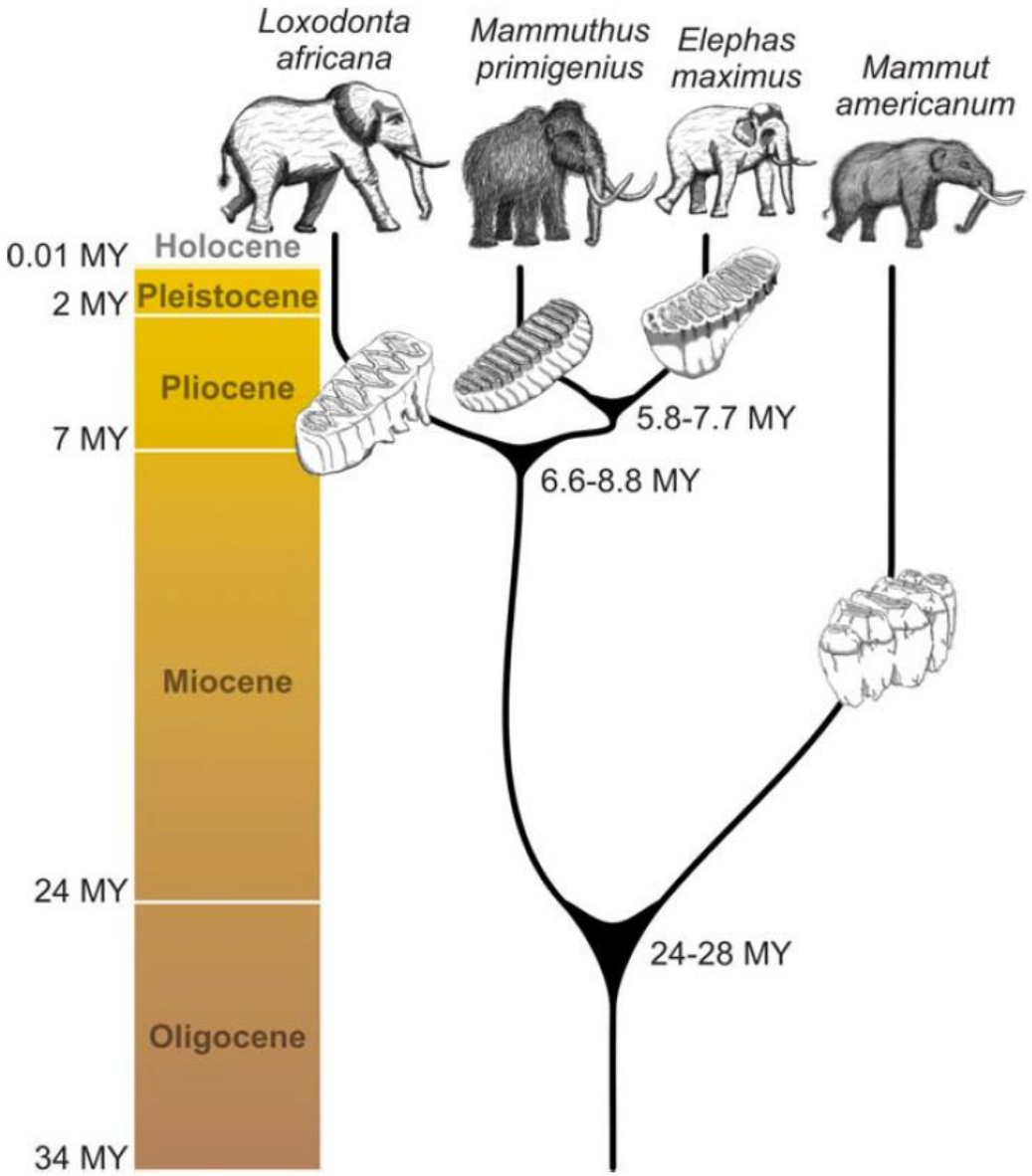
The choice of outgroup - living relatives

Phylogeny without outgroup

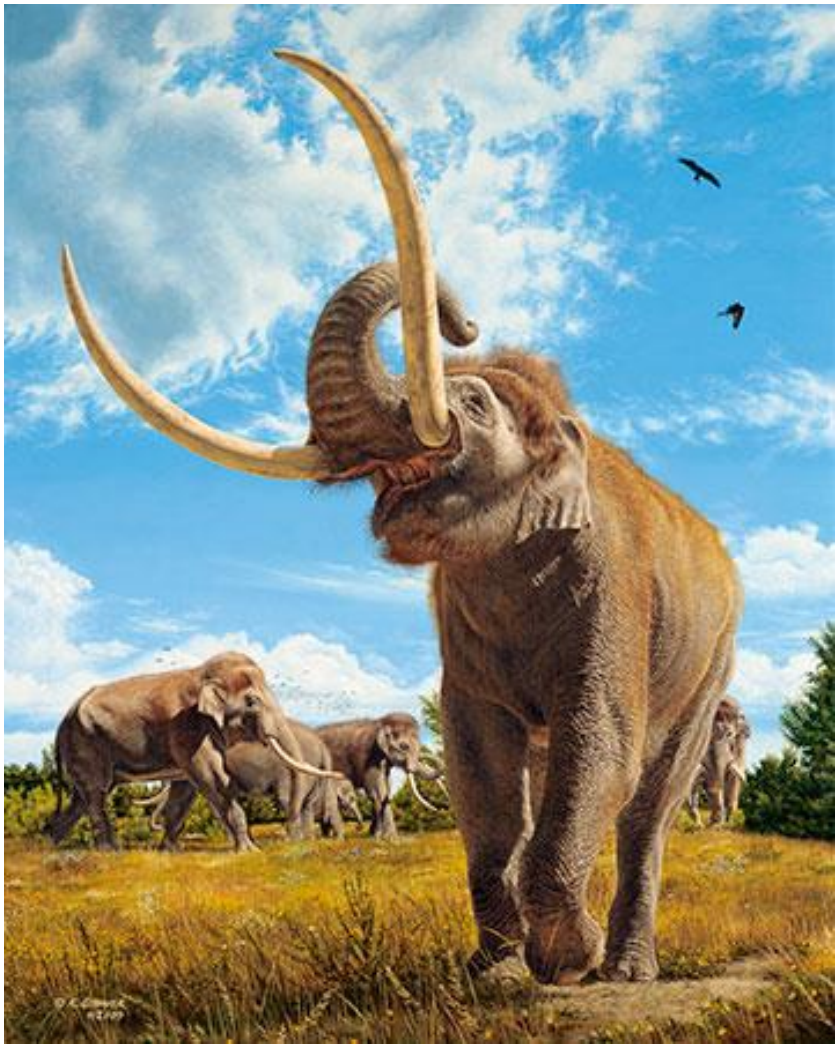


The choice of outgroup – extinct relatives

Phylogenetics

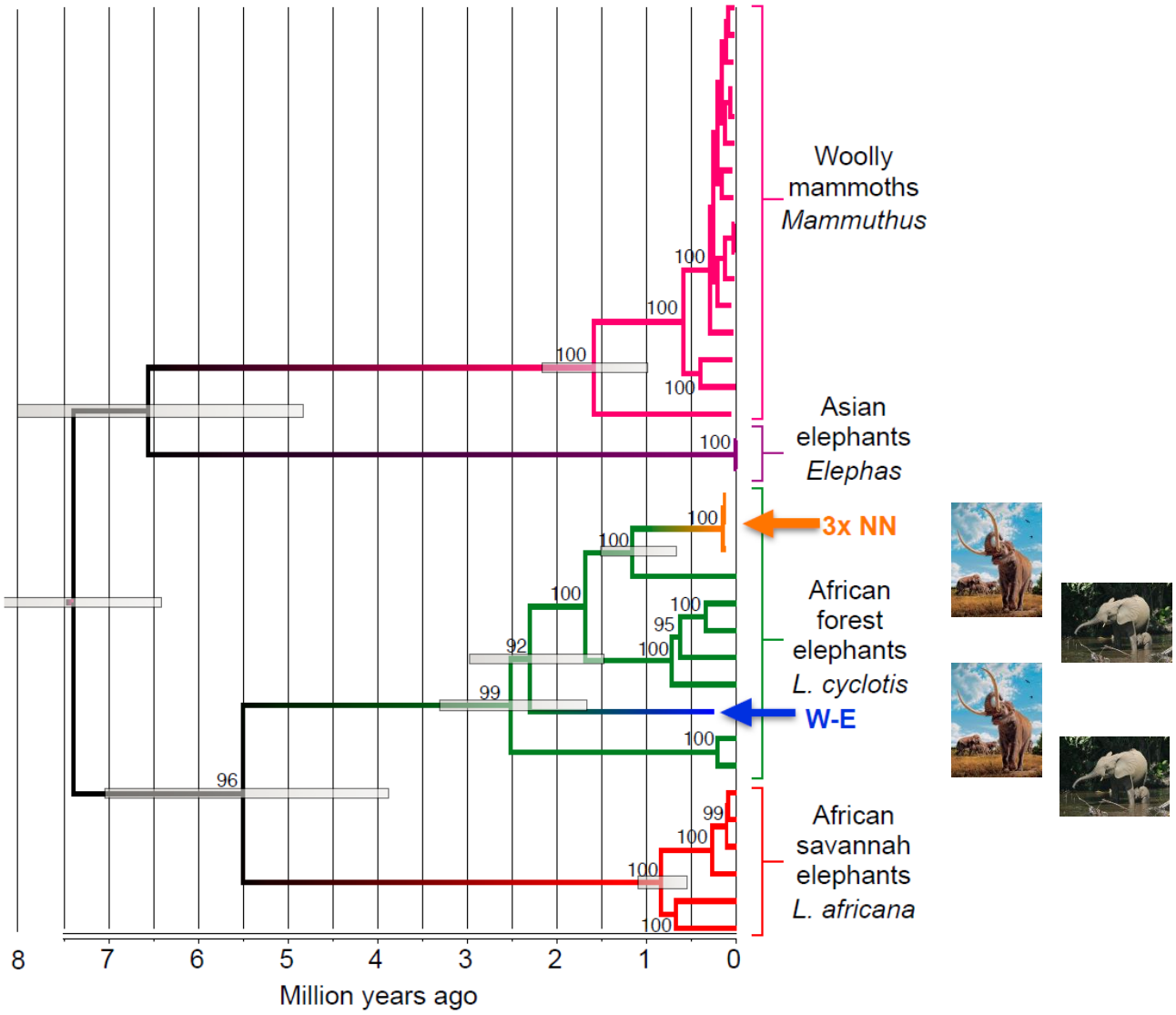


Palaeoloxodon antiquus



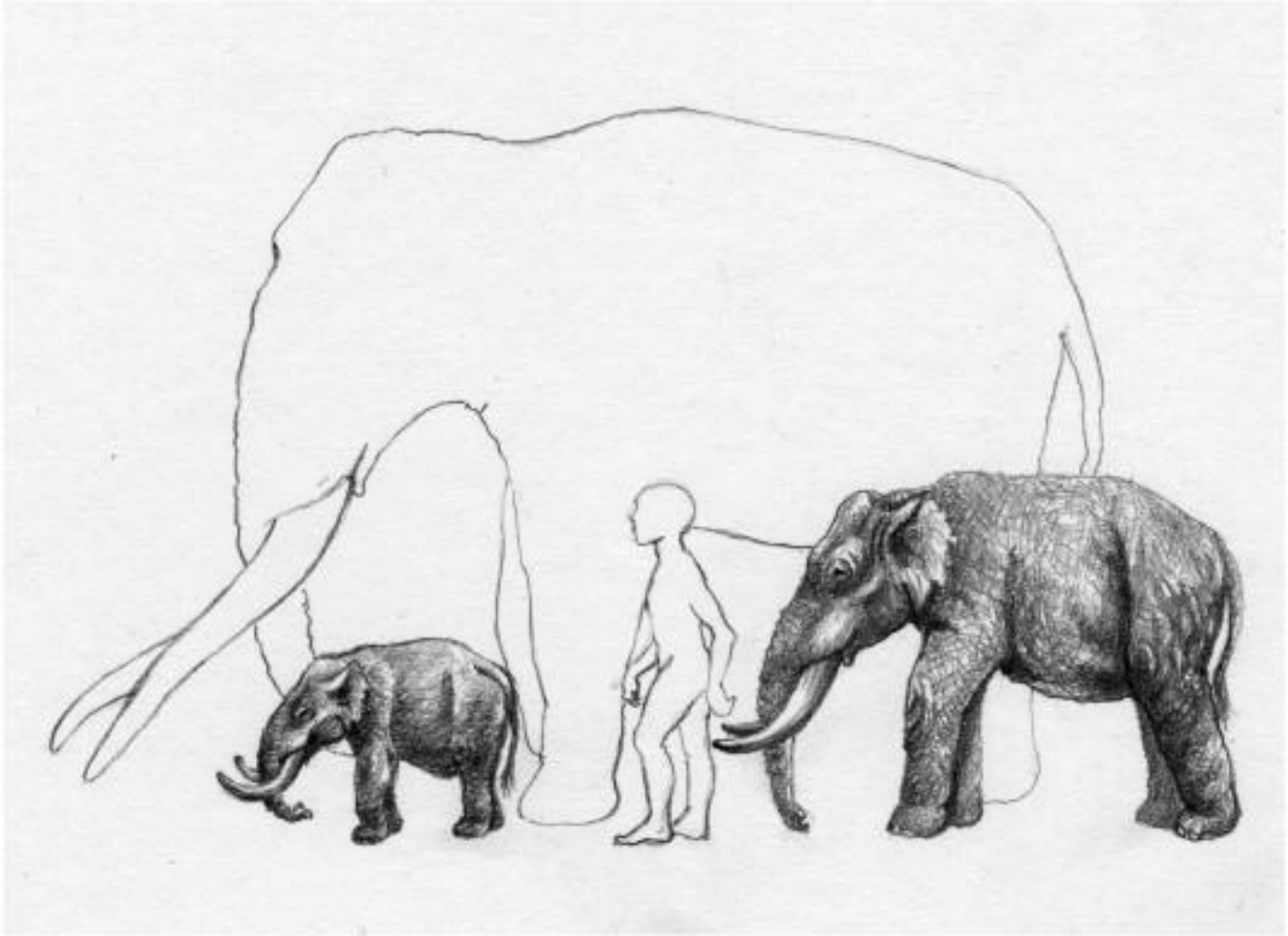
~120,000 years old

Mitogenome tree

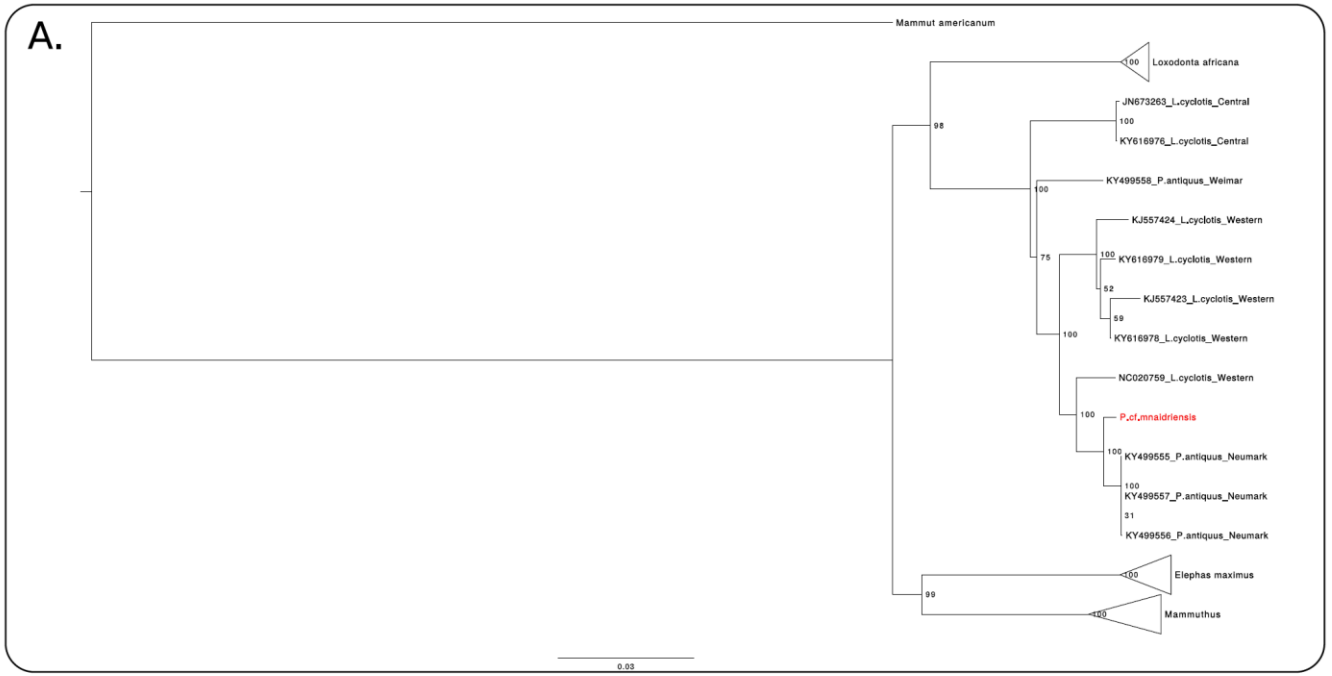


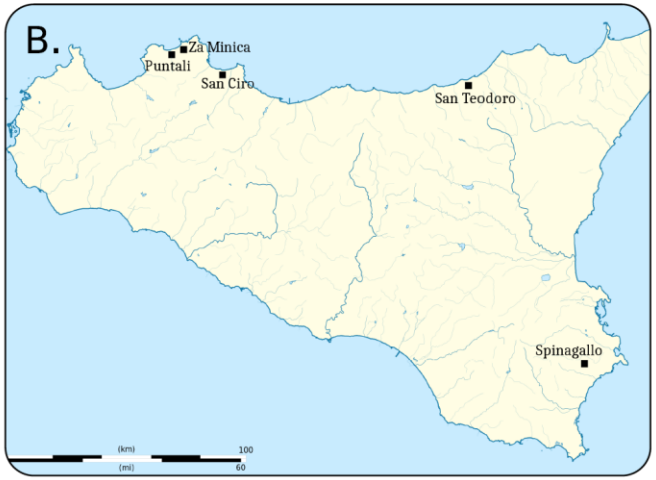
Sicily

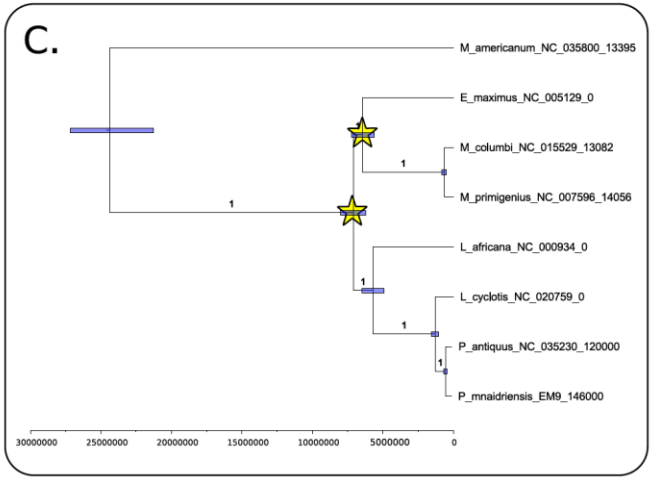
Palaeoloxodon mnaidriensis



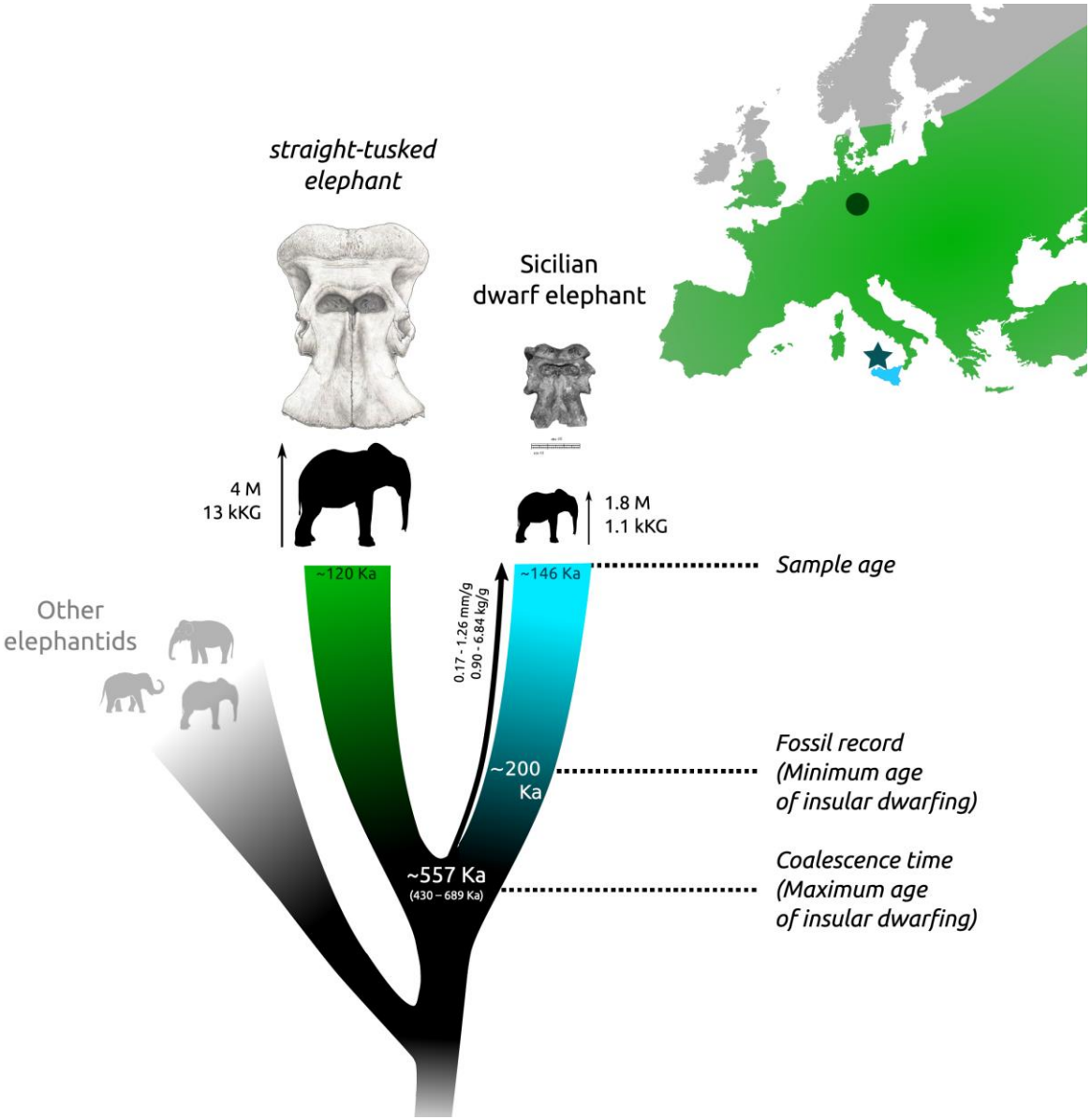
Geographical origin and mitogenome phylogeny





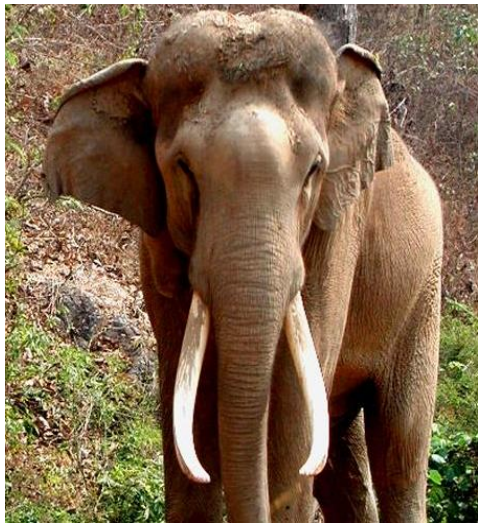


Insular dwarfing

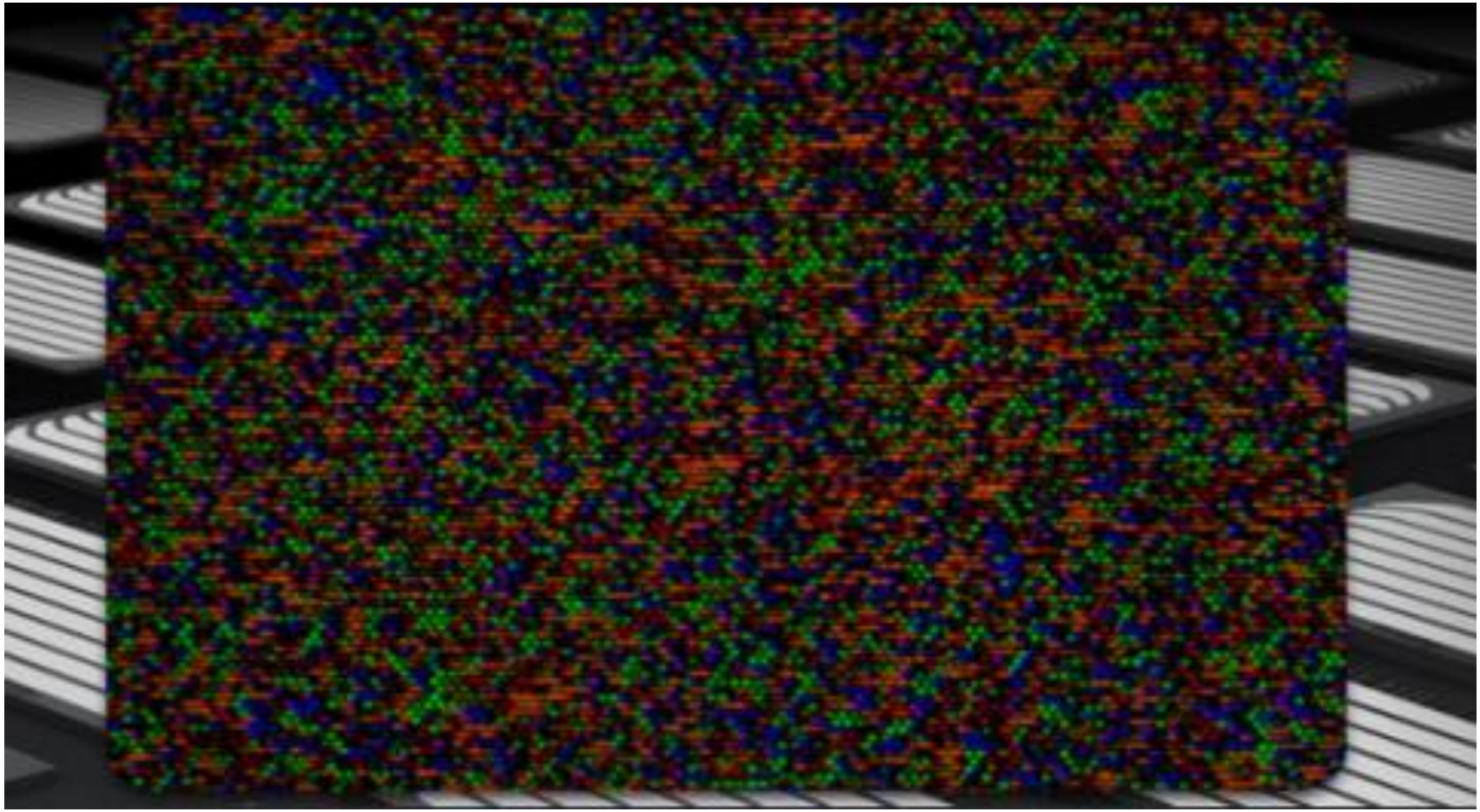


African forest elephant

How are these related?

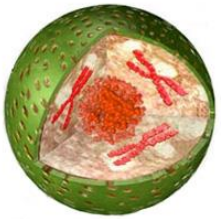


Nuclear DNA sequencing



Mitochondrial vs. nuclear phylogenies



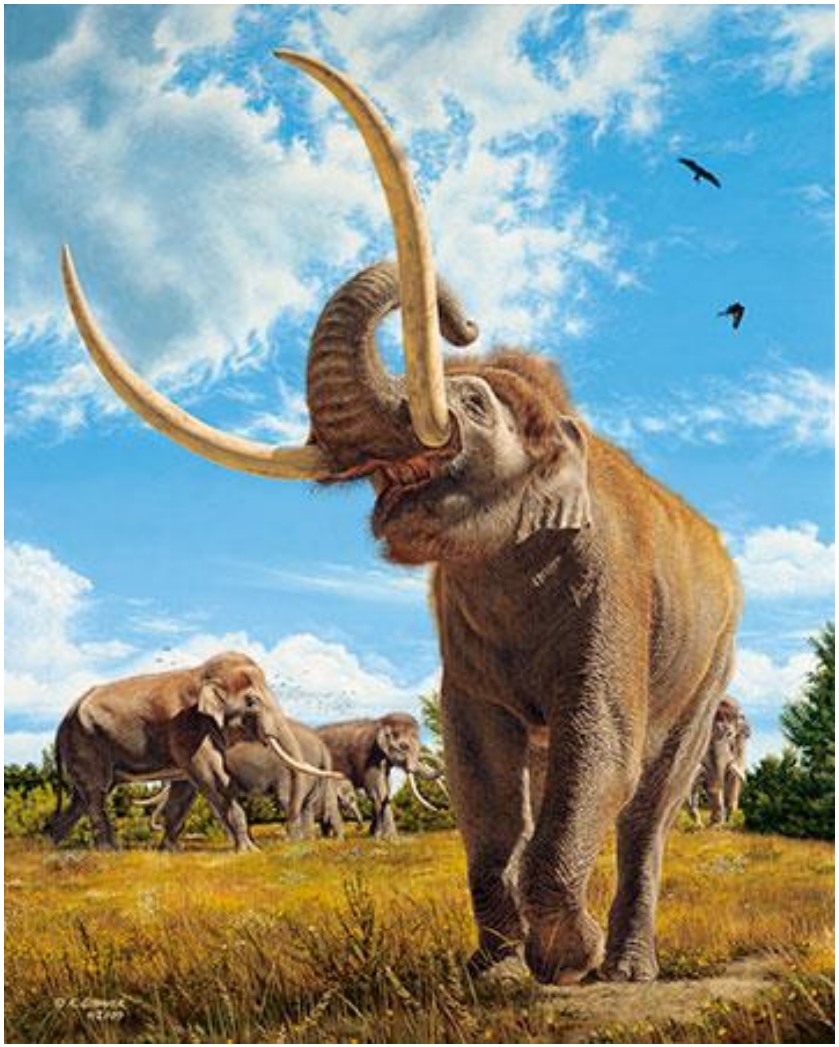



5.0
5.0
1.6
6.6



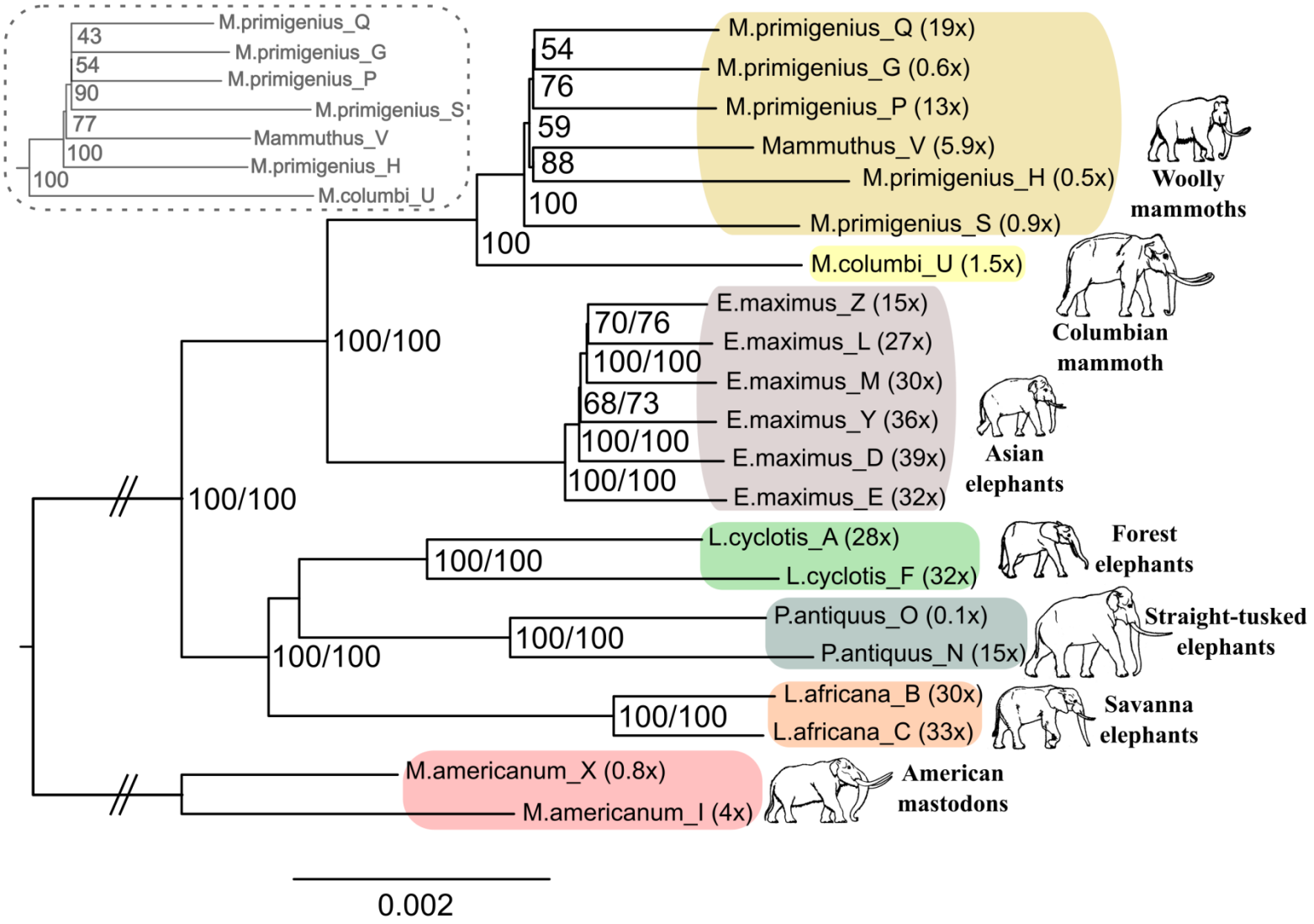
6.7
7.6
0.9

Palaeoloxodon antiquus

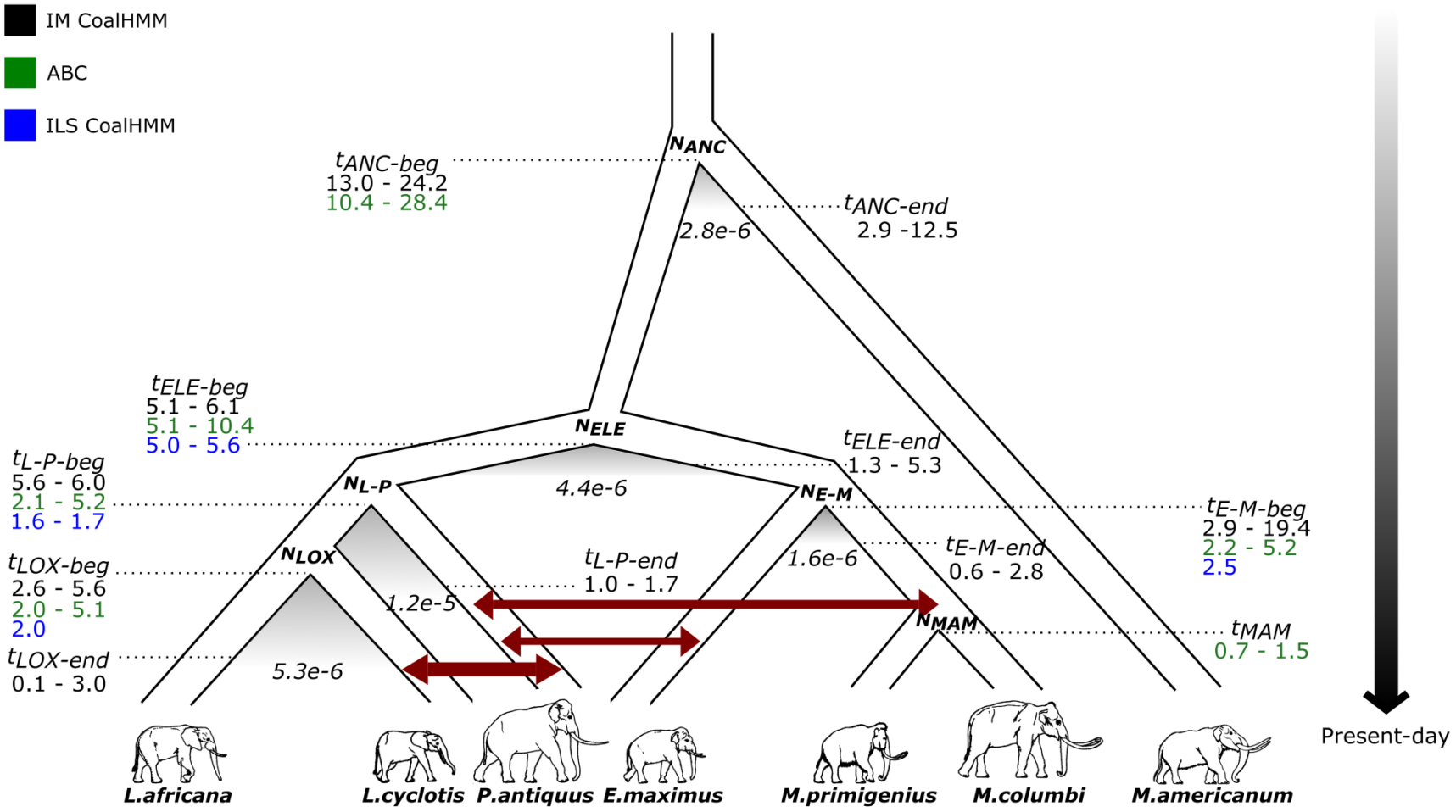


~120,000 years old

Palaeoloxodon antiquus – nuclear tree



Palaeoloxodon antiquus – an extinct hybrid species



N_{LOX}: 43,000 - 69,000
29,000 - 175,000
165,000

N_{L-P}: 46,000 -53,000
37,000 - 233,000
121,000 - 137,000

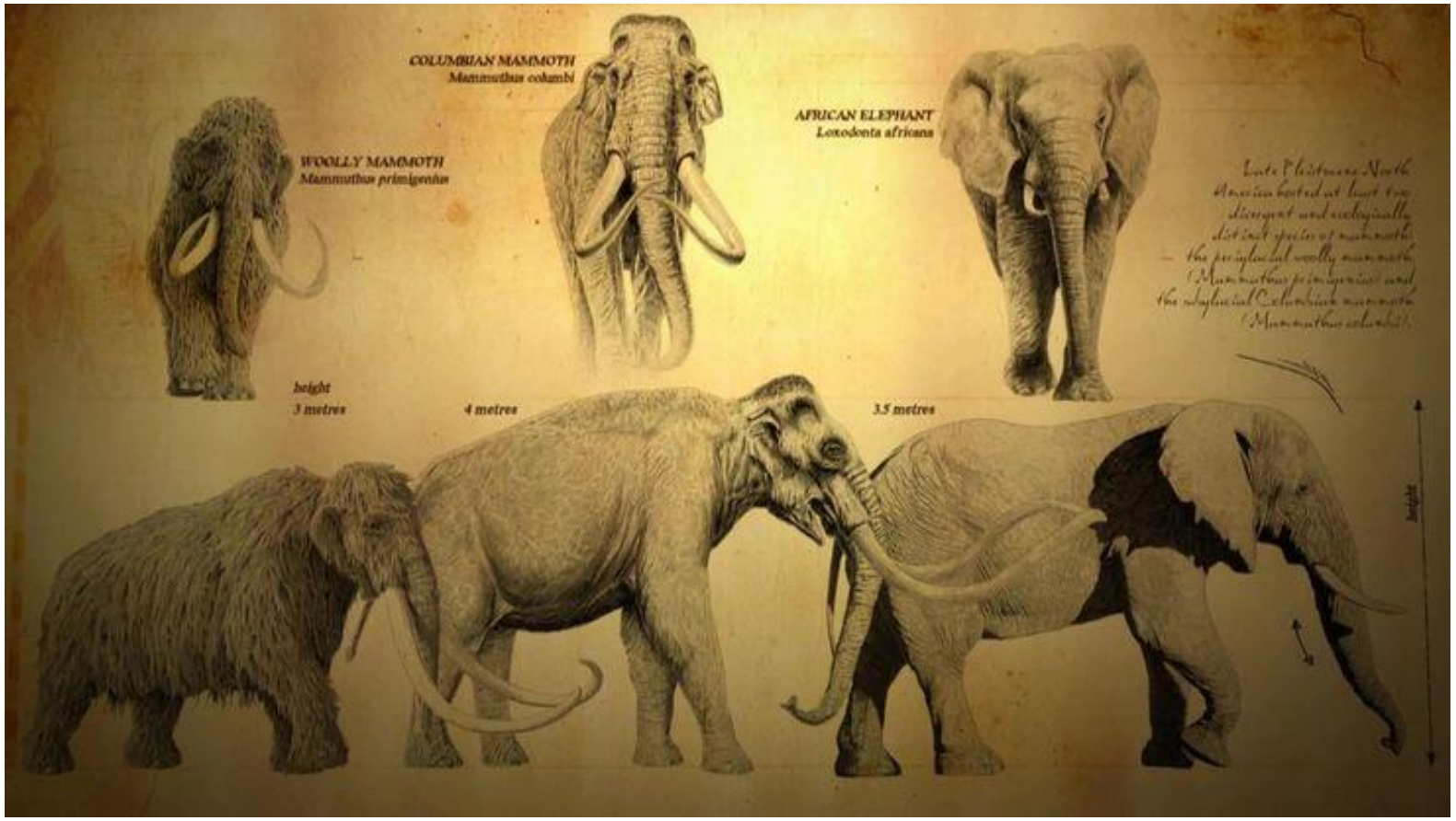
N_{ELE}: 54,000 - 68,000
7,000 - 78,000
48,000 - 53,000

N_{MAM}: 8,000 - 30,000
N_{E-M}: 27,000 - 60,000
10,000 - 130,000
72,000

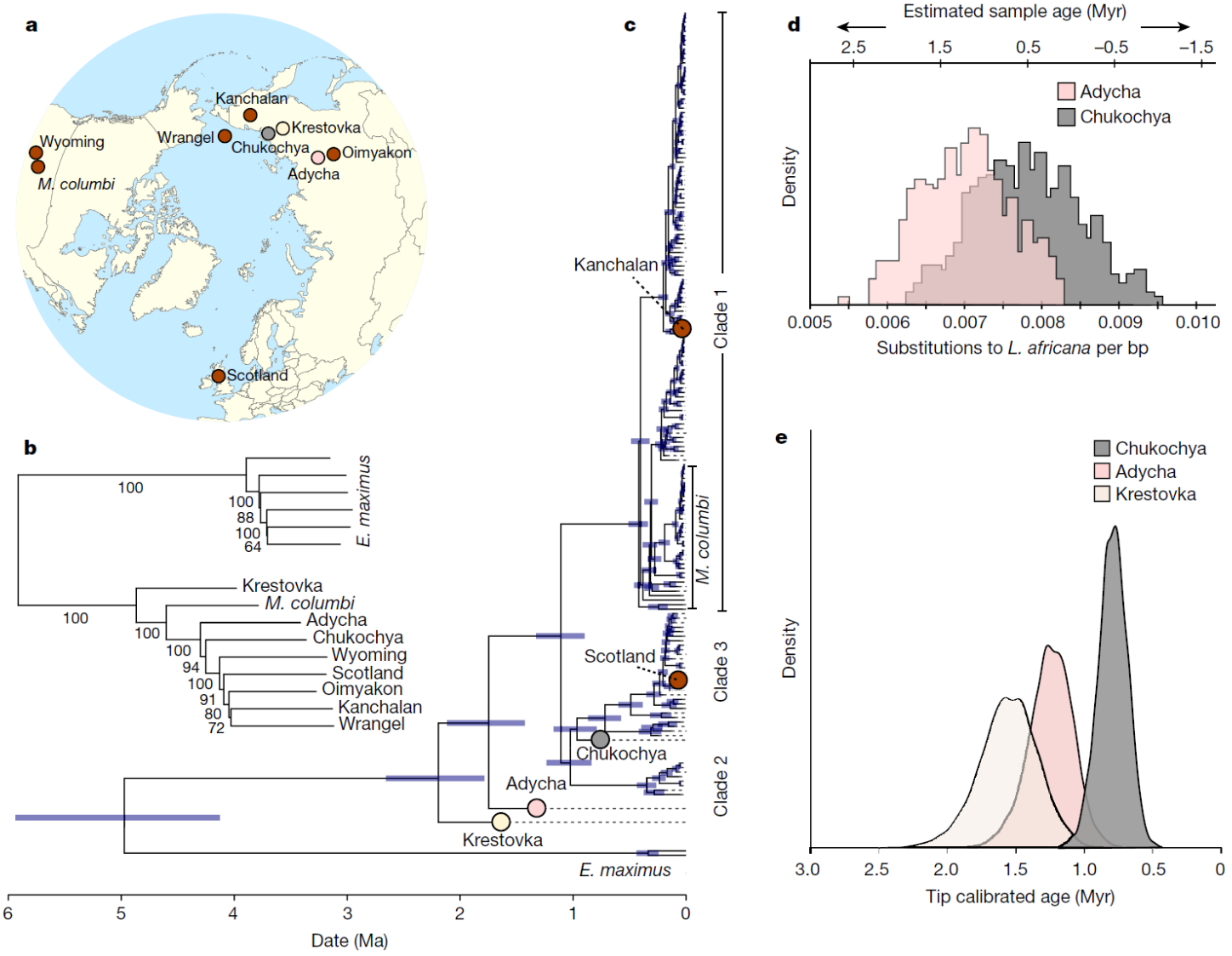
N_{ANC}: 36,000 - 77,000
2,000 - 160,000

Mammoths

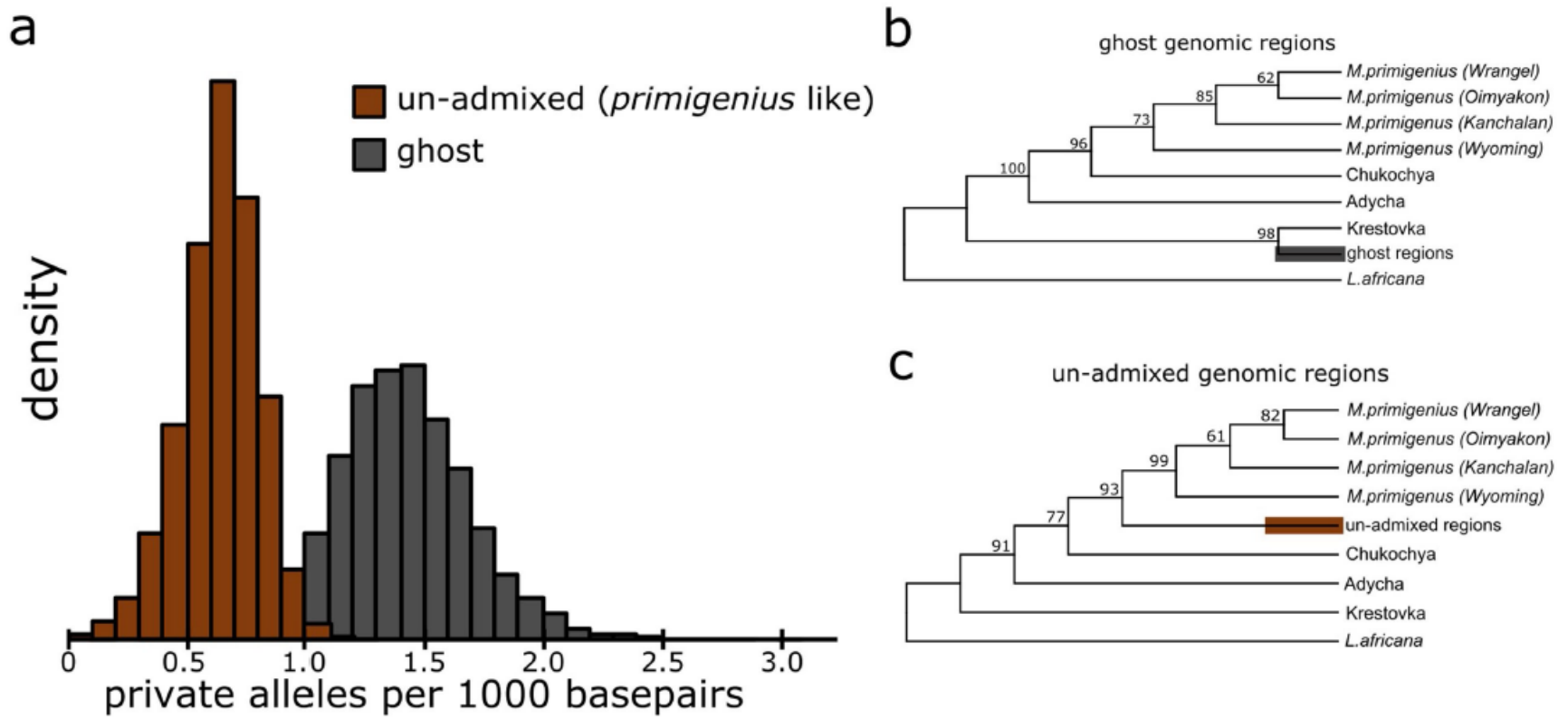
Mammoths


COLUMBIAN MAMMOTH
Mammuthus columbi
WOOLLY MAMMOTH
Mammuthus primigenius
AFRICAN ELEPHANT
Loxodonta africana
height
3 metres
4 metres
3.5 metres

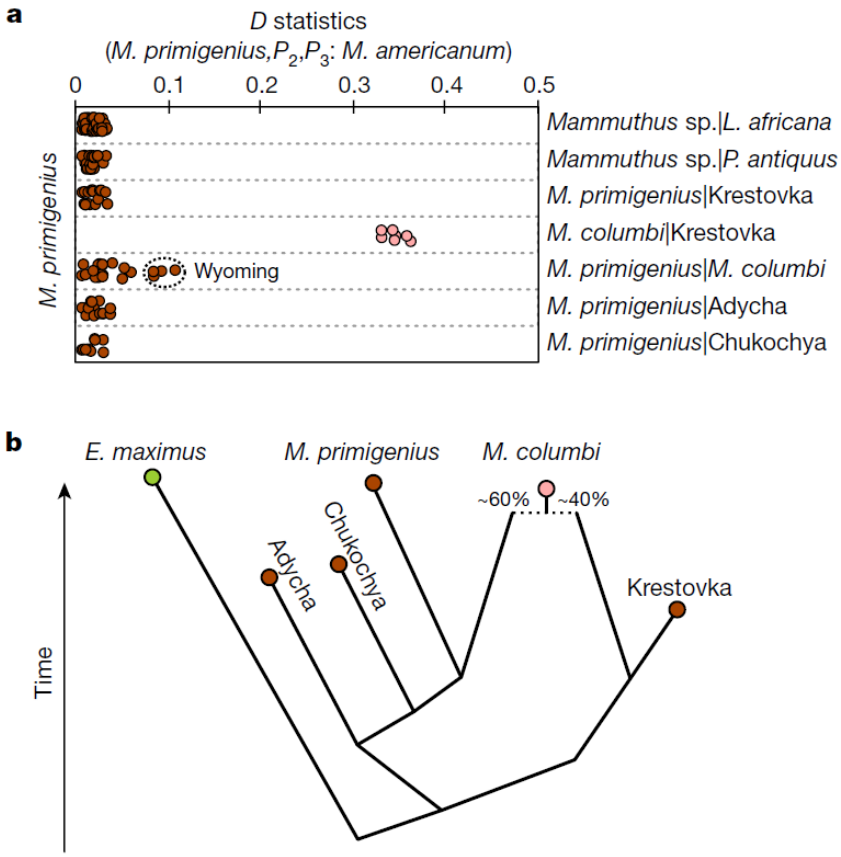
Mammoths – samples, age and phylogeny



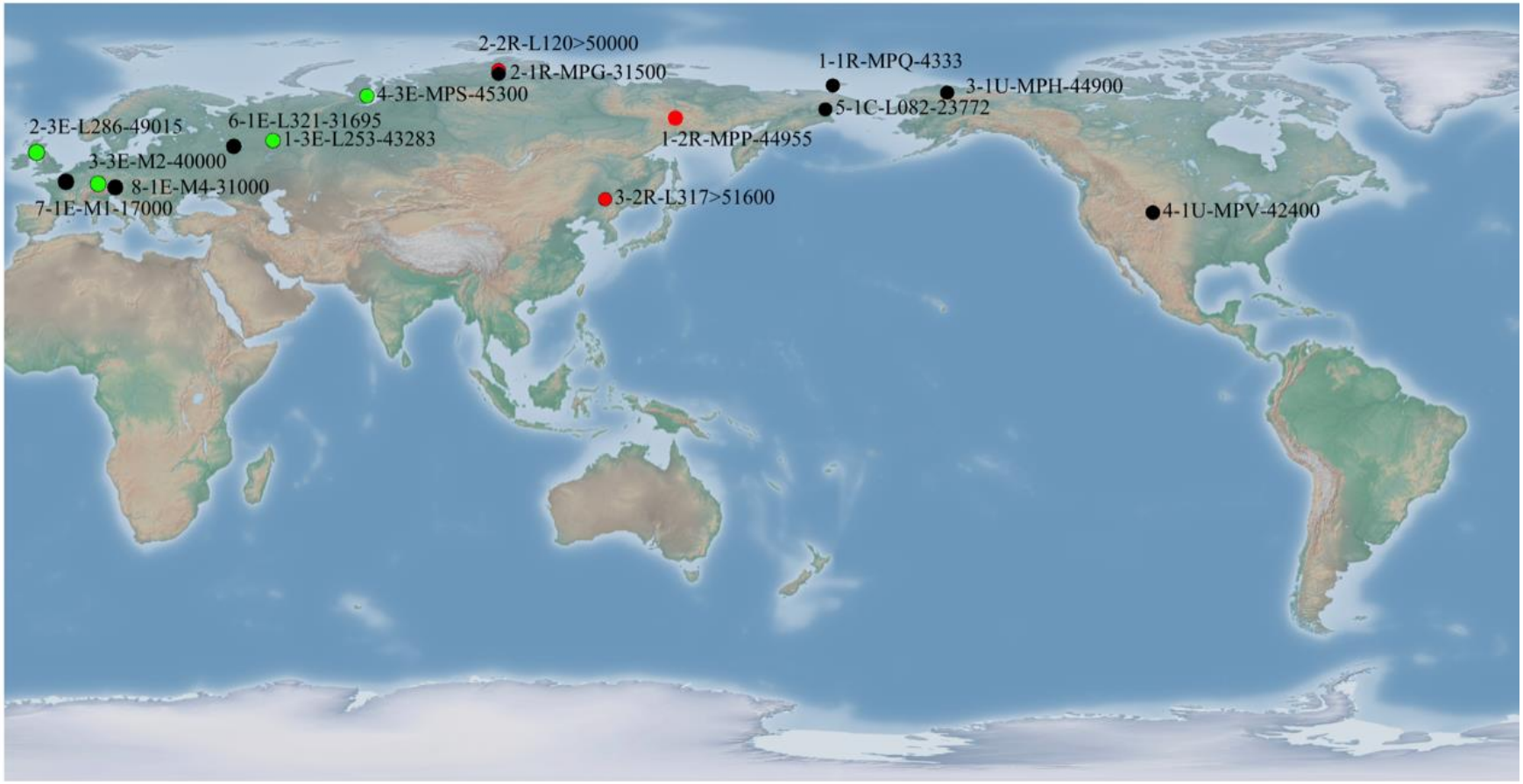
Columbian mammoth – ghost ancestry



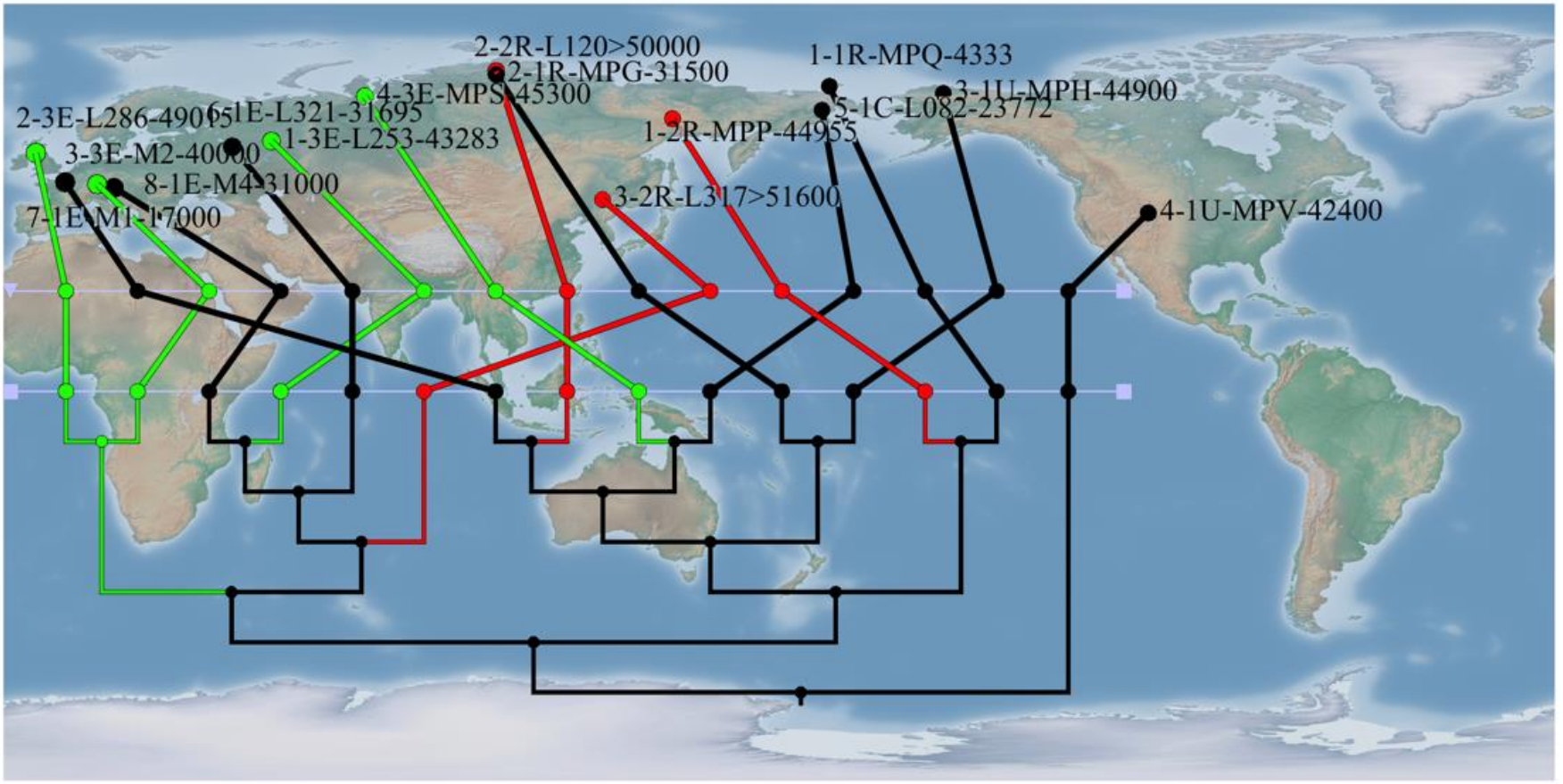
Mammoths – population history



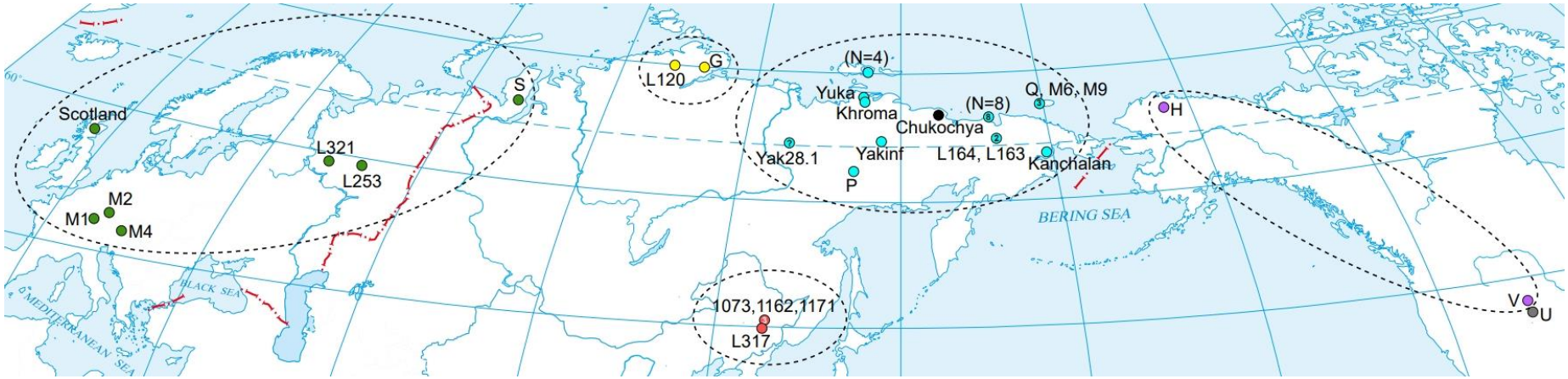
More mammoths

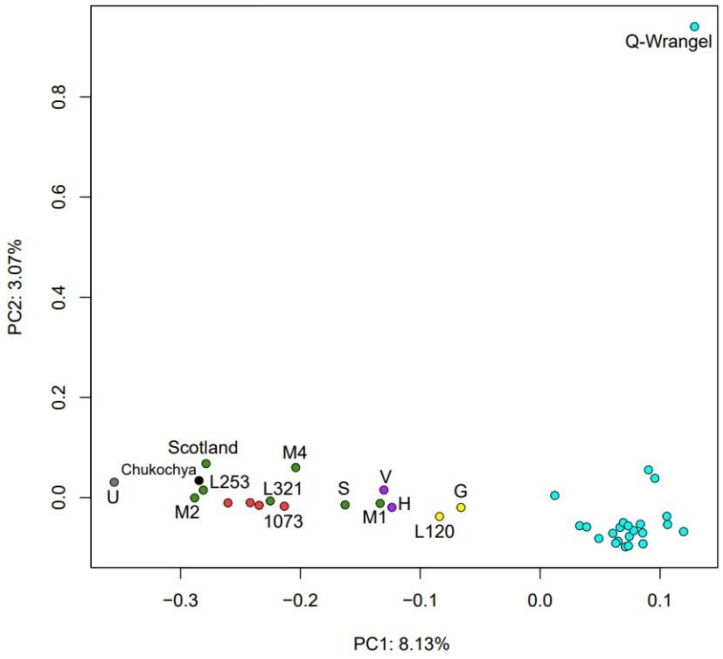


Mitochondrial haplogroups vs. nuclear tree



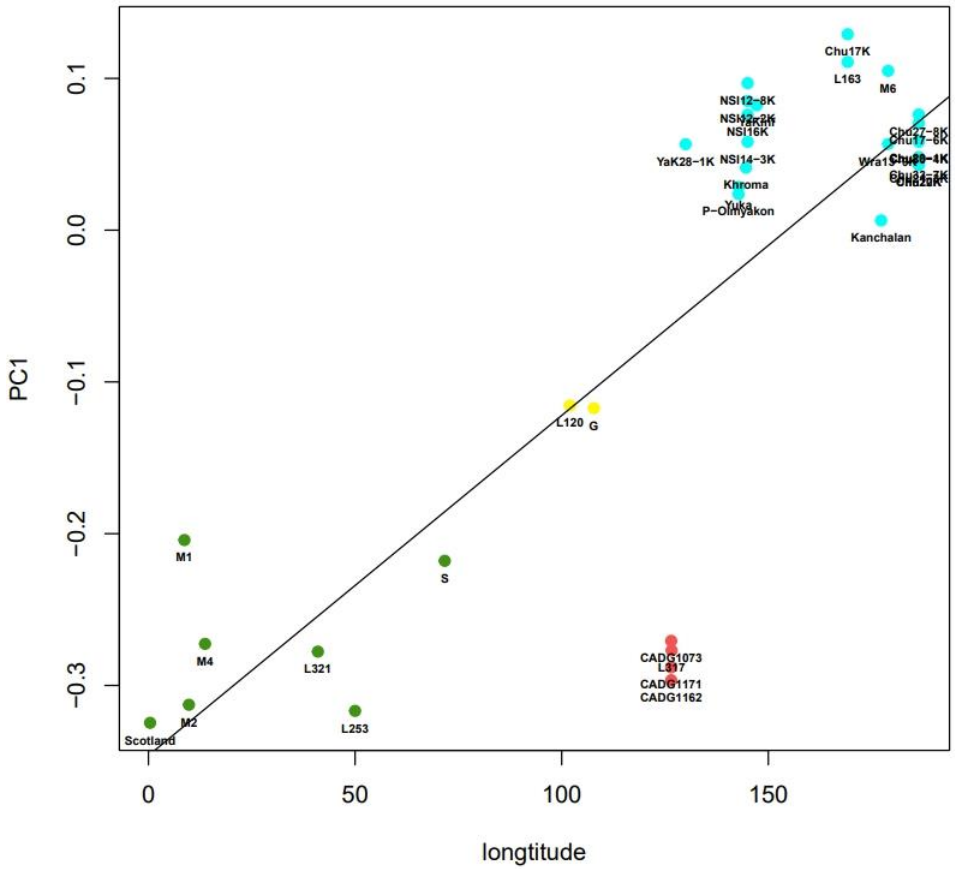
Even more mammoths





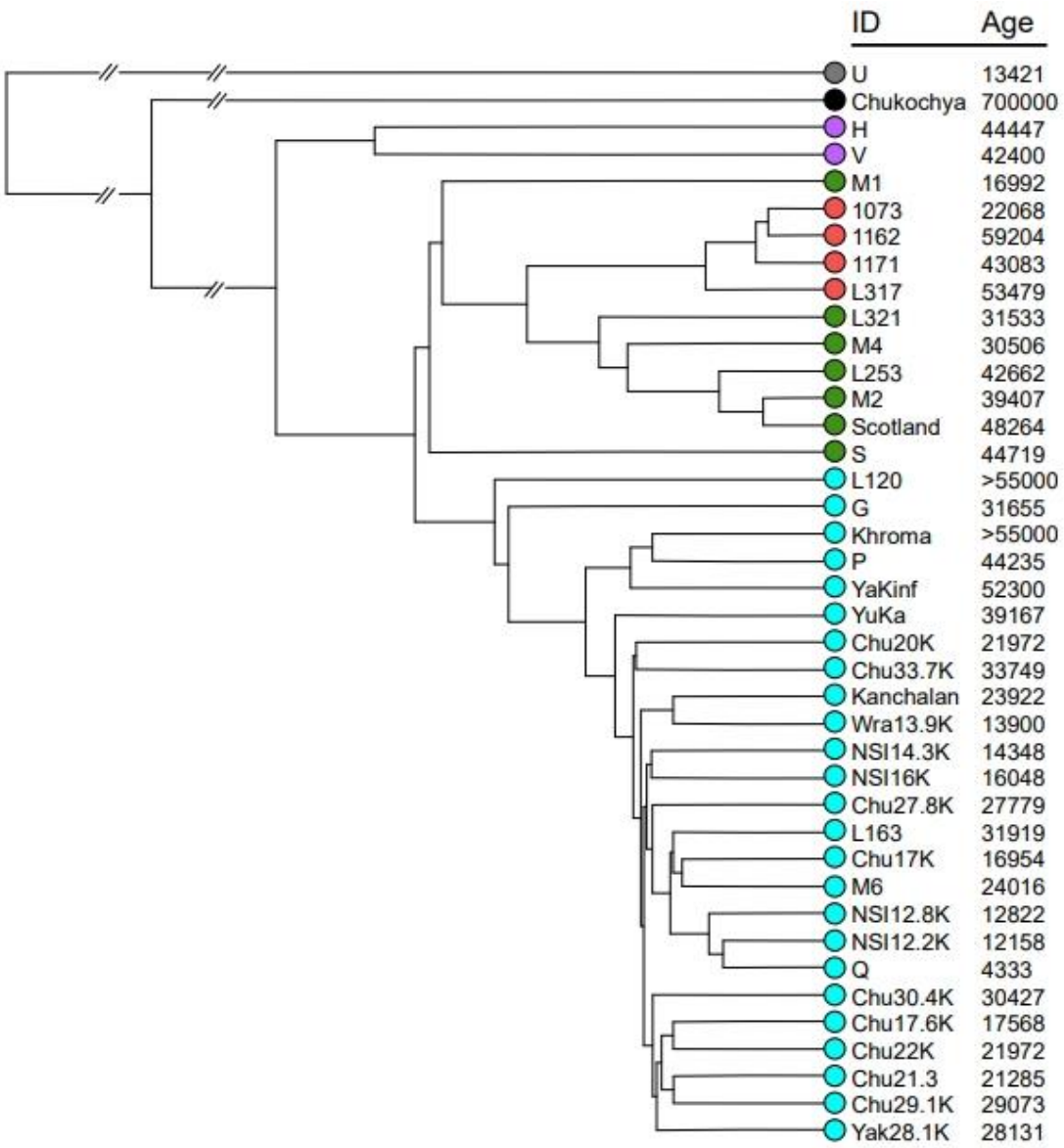
Geographical vs. genetic structure

PC1 vs longitude

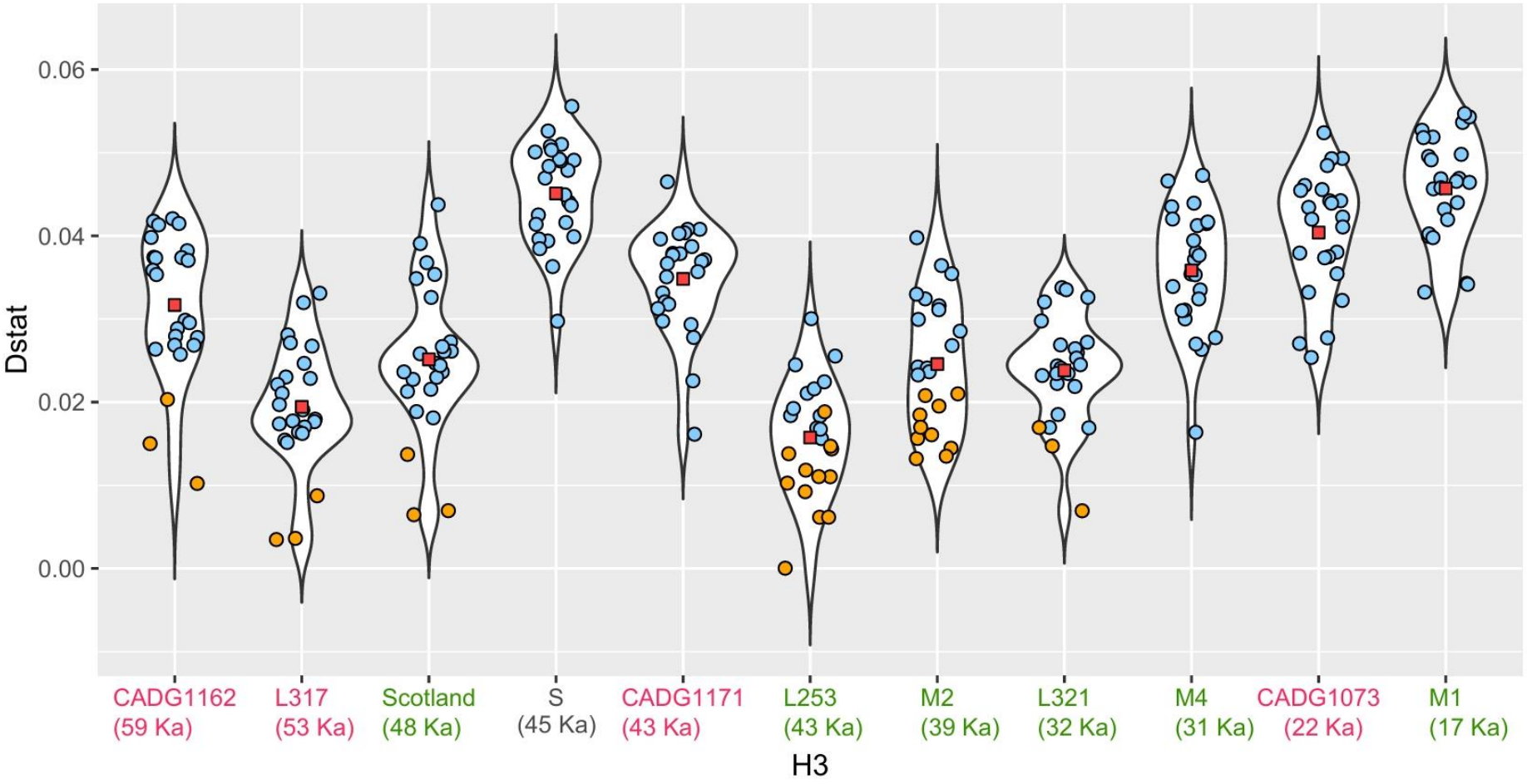


P value: 5.65e-09 Adjusted R-squared: 0.637

NJ-tree using all samples



Siberian admixture over time and space

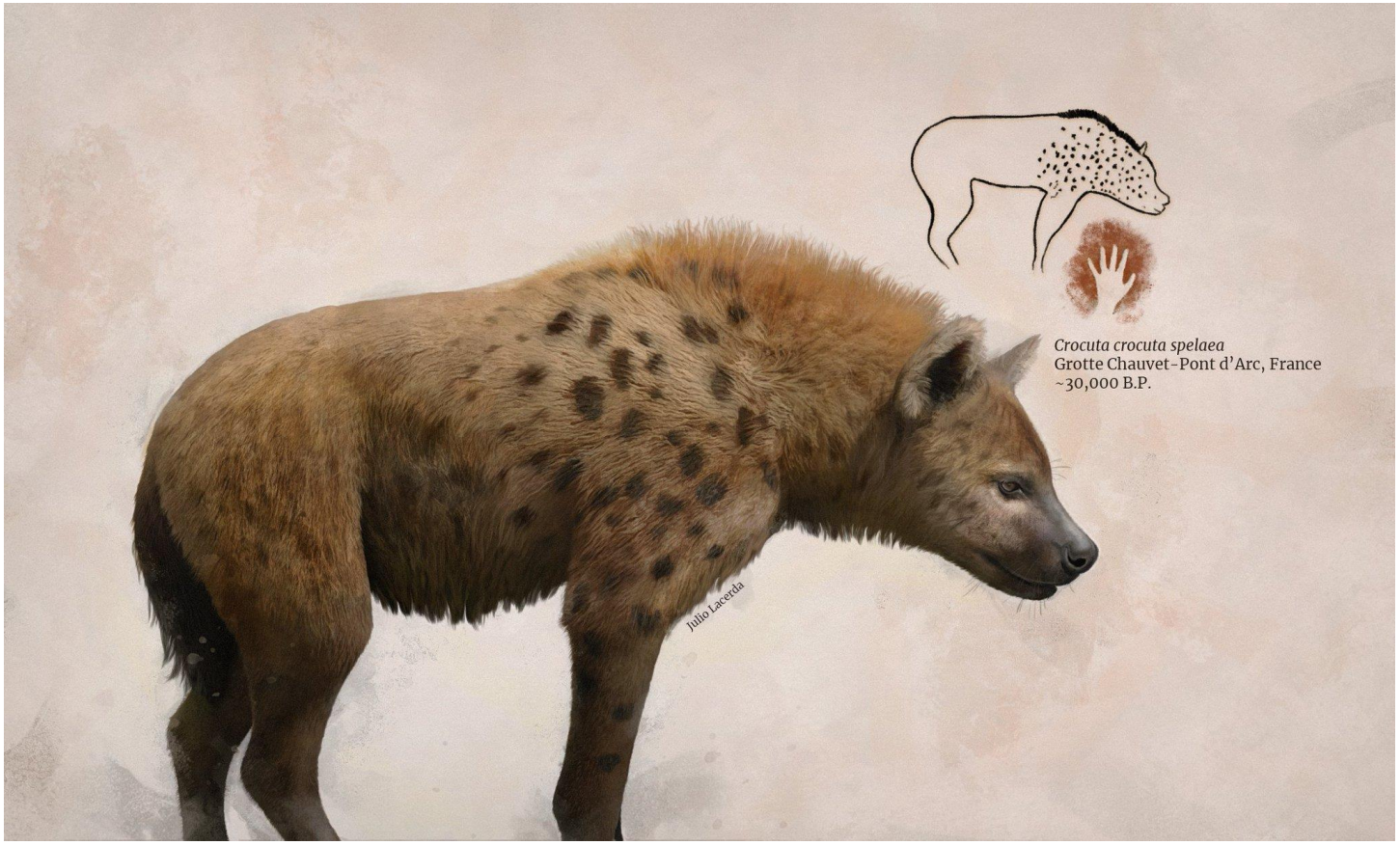


Summary elephants

- Gene flow within and between species is an important factor in elephant evolution
- The European straight-tusked elephant was a three-way hybrid
- Chinese and European mammoth are more closely related than either to Siberian ones
- Siberian influence increases over time

Cave hyenas

„Real“ cave hyena



As seen by stone age artists

Crocuta - distribution

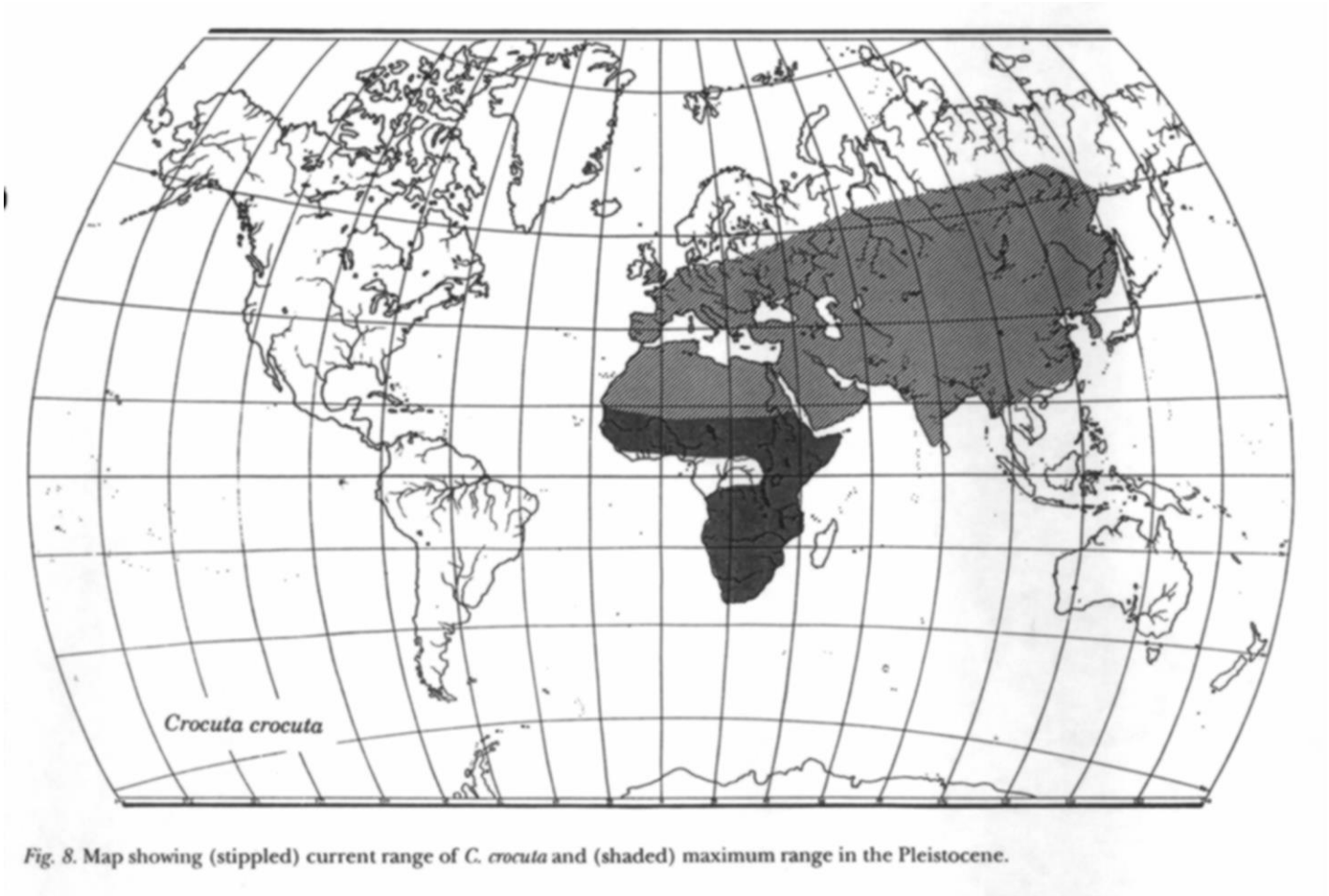
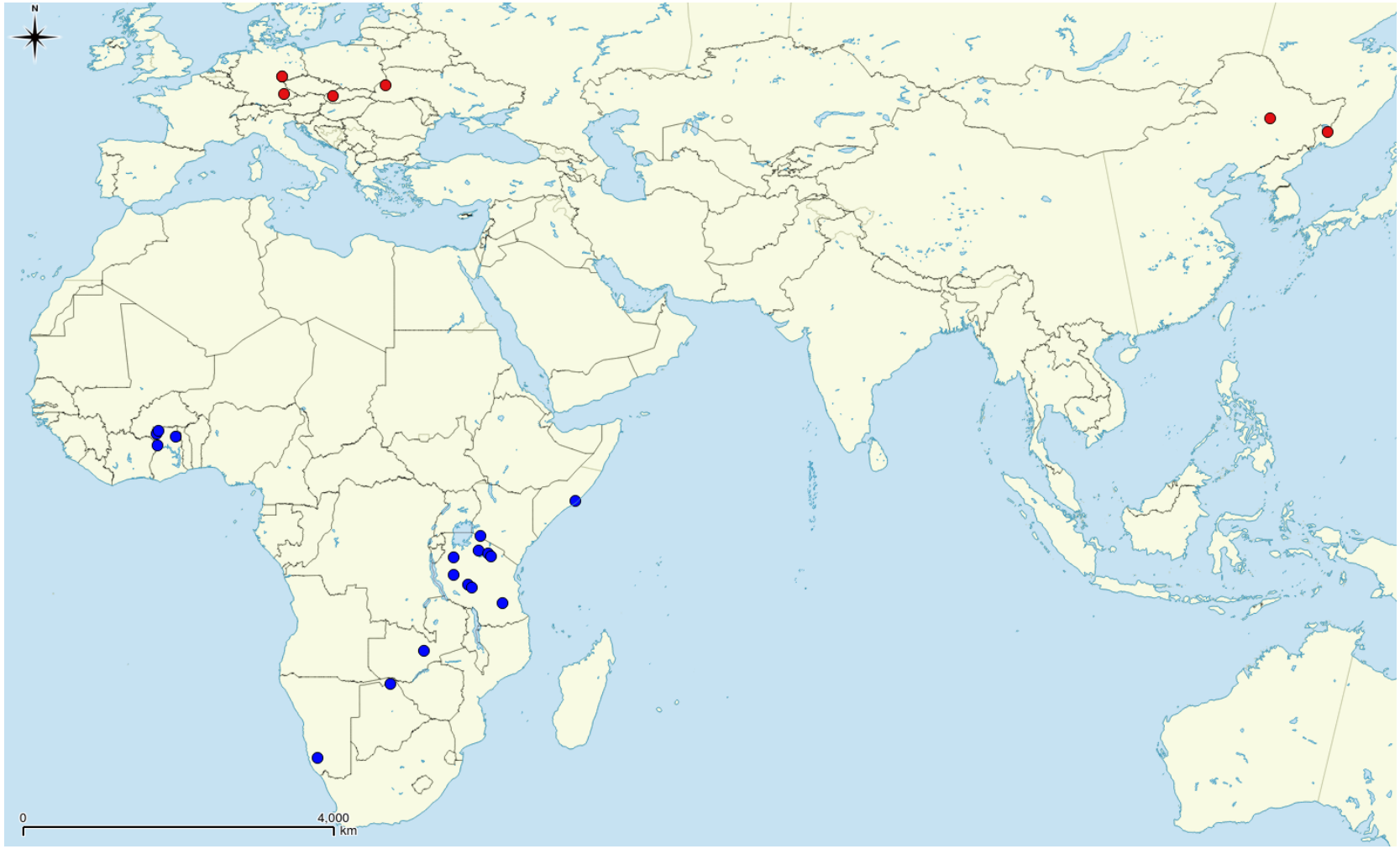
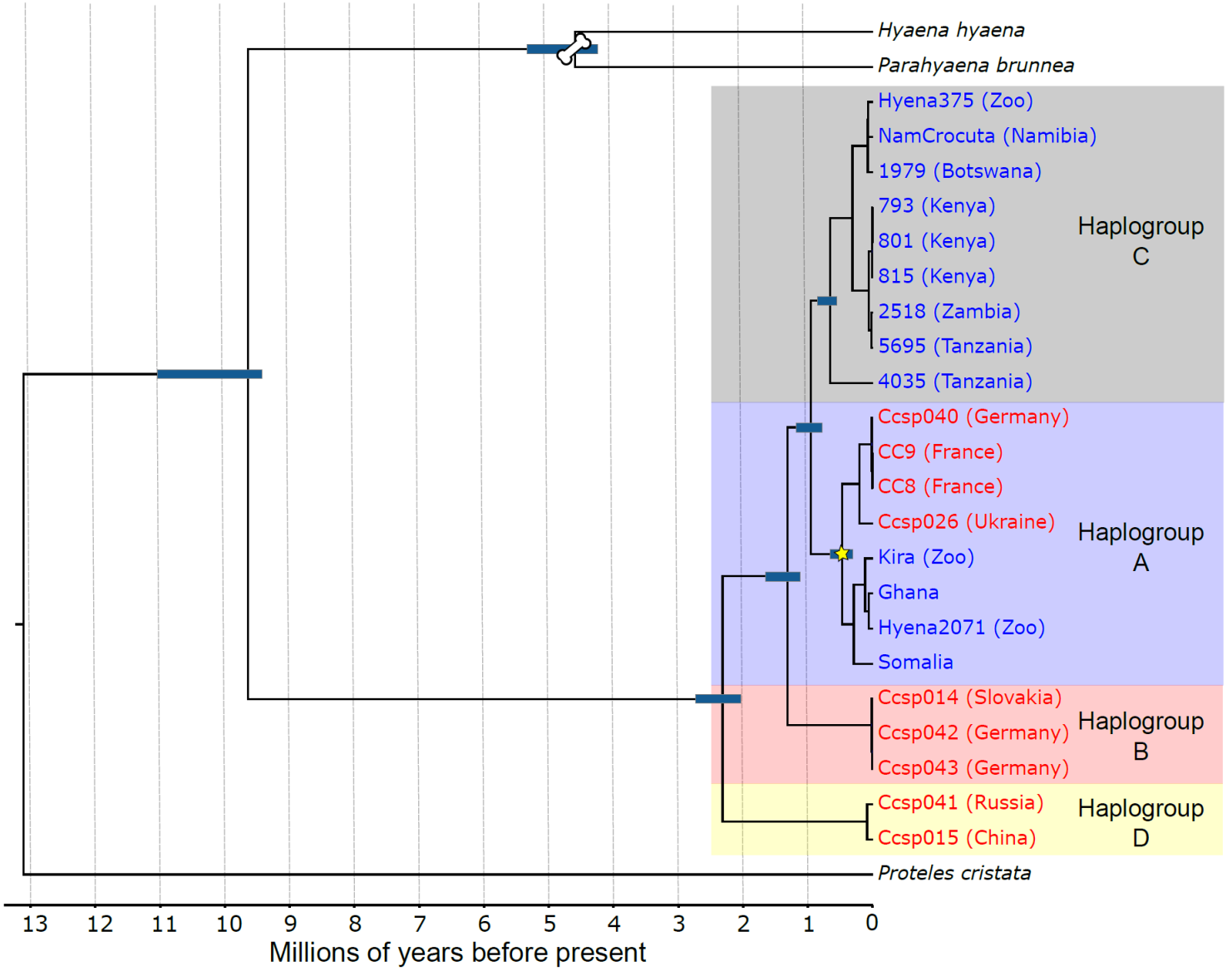


Fig. 8. Map showing (stippled) current range of *C. crocuta* and (shaded) maximum range in the Pleistocene.

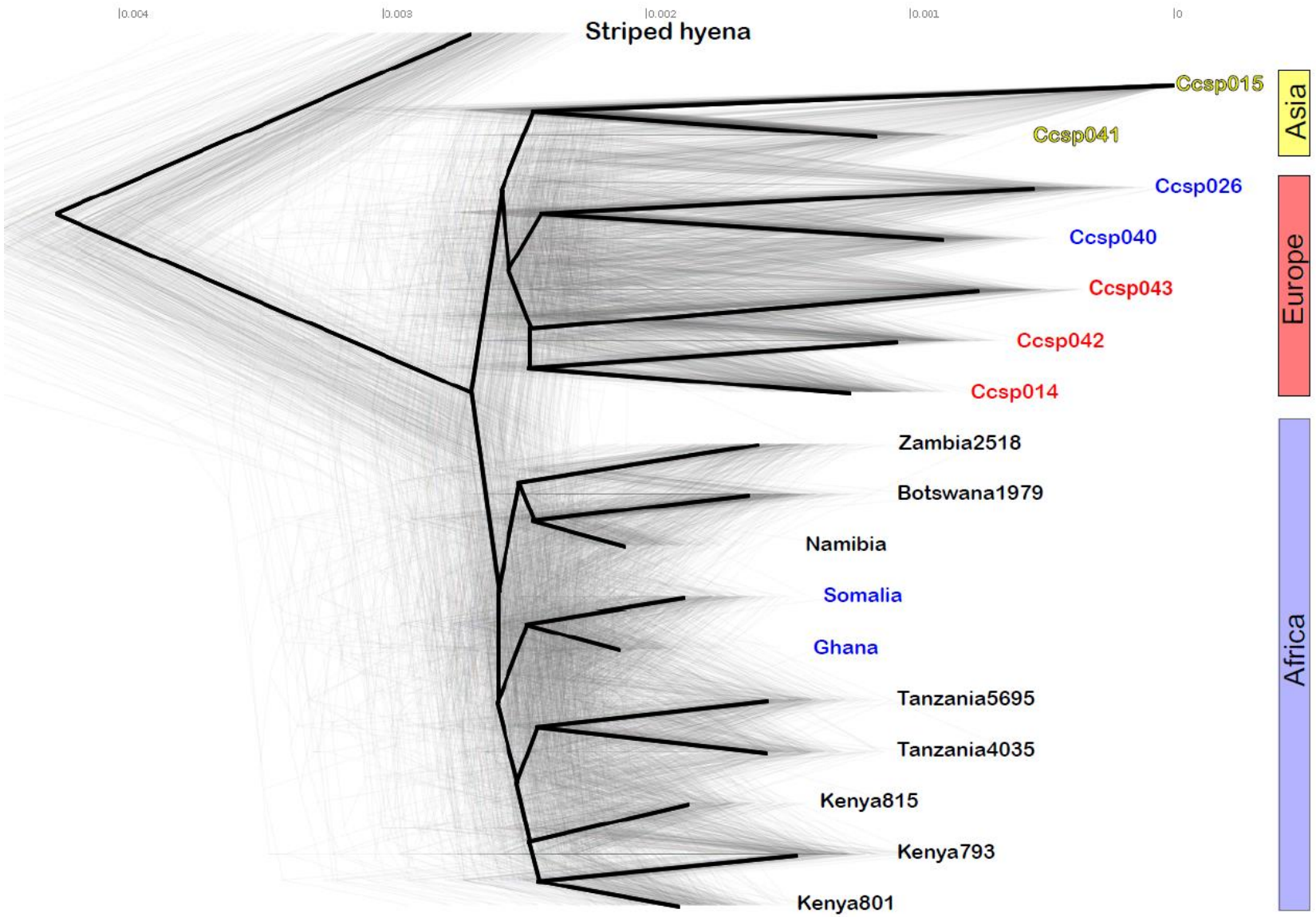
Sampling



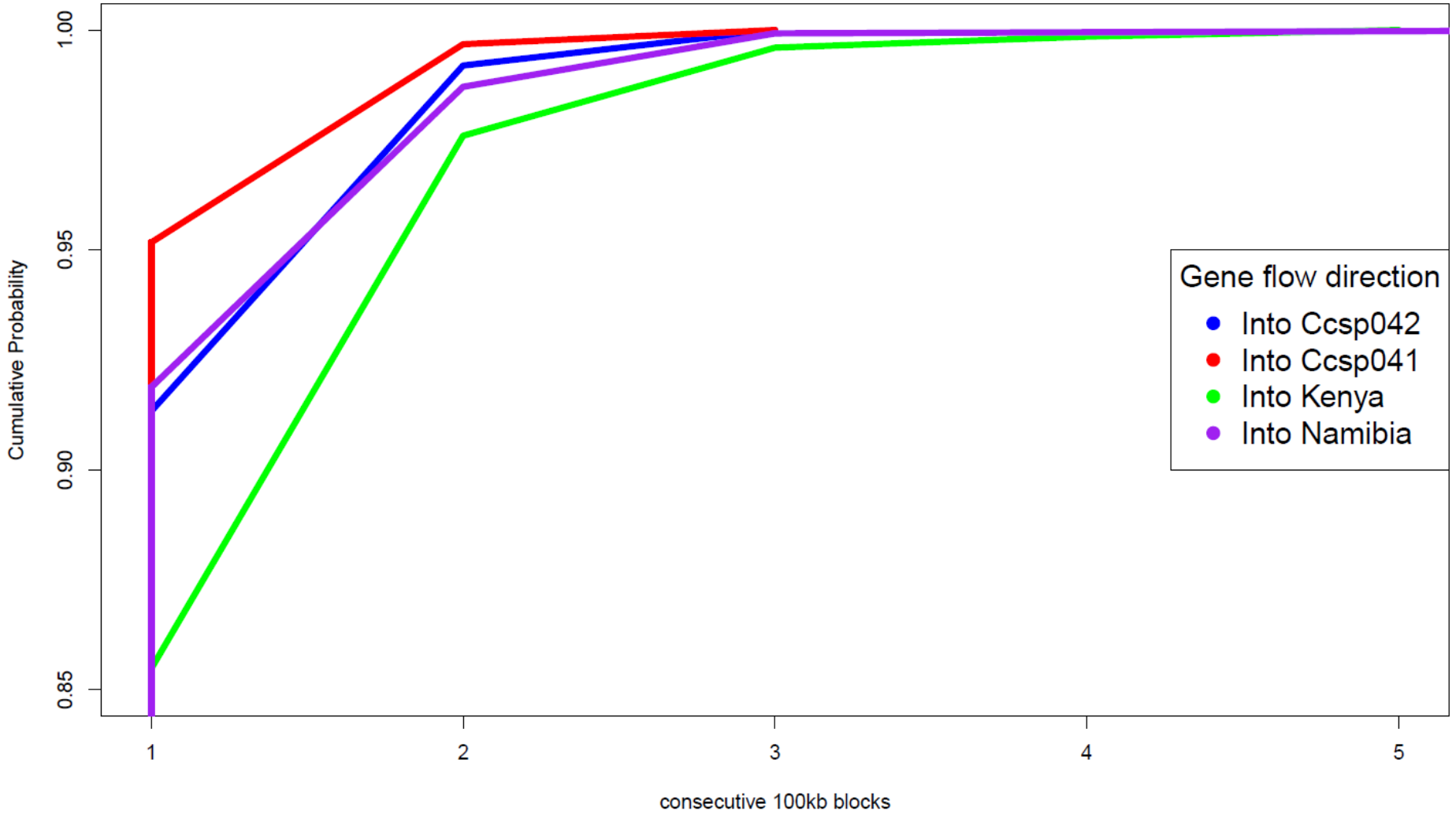
Mitogenome tree



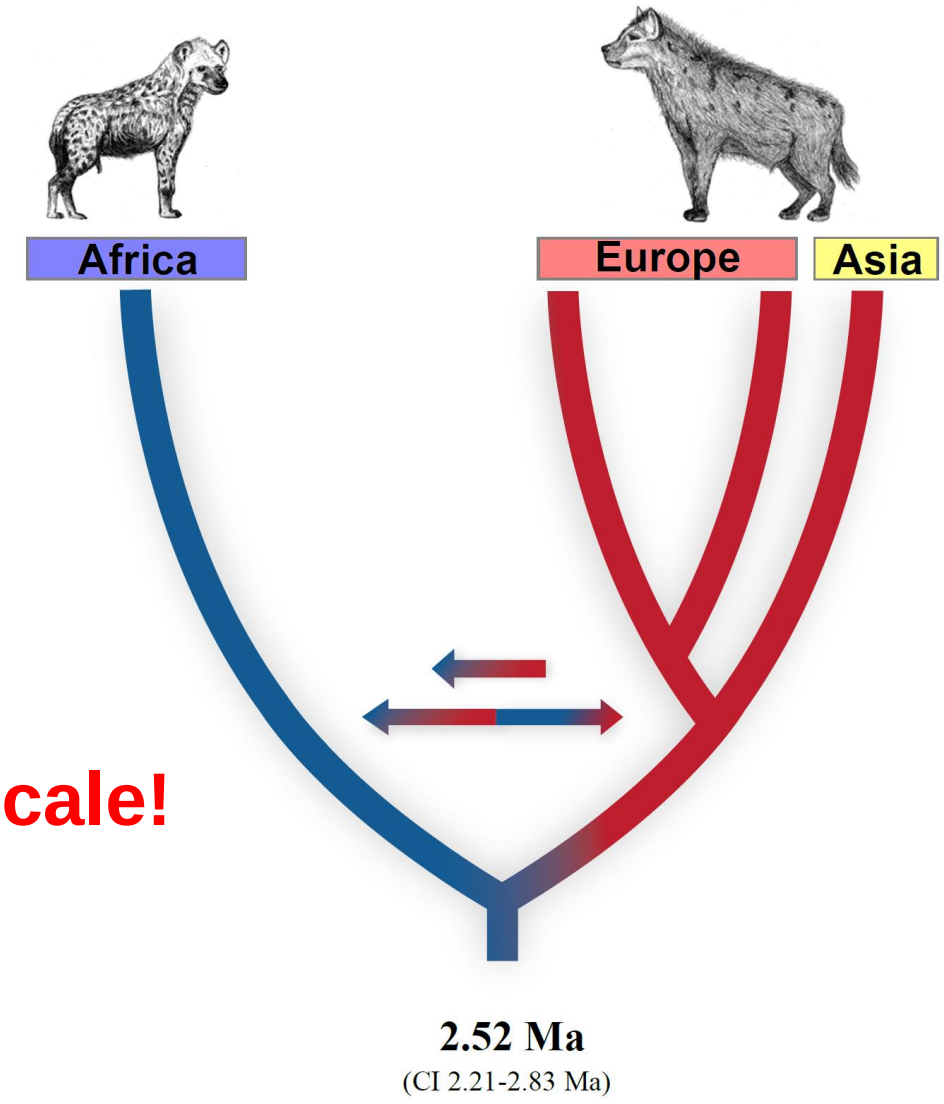
Nuclear tree



Crocuta gene flow



Hyena dated tree and admixture 2 my after divergence



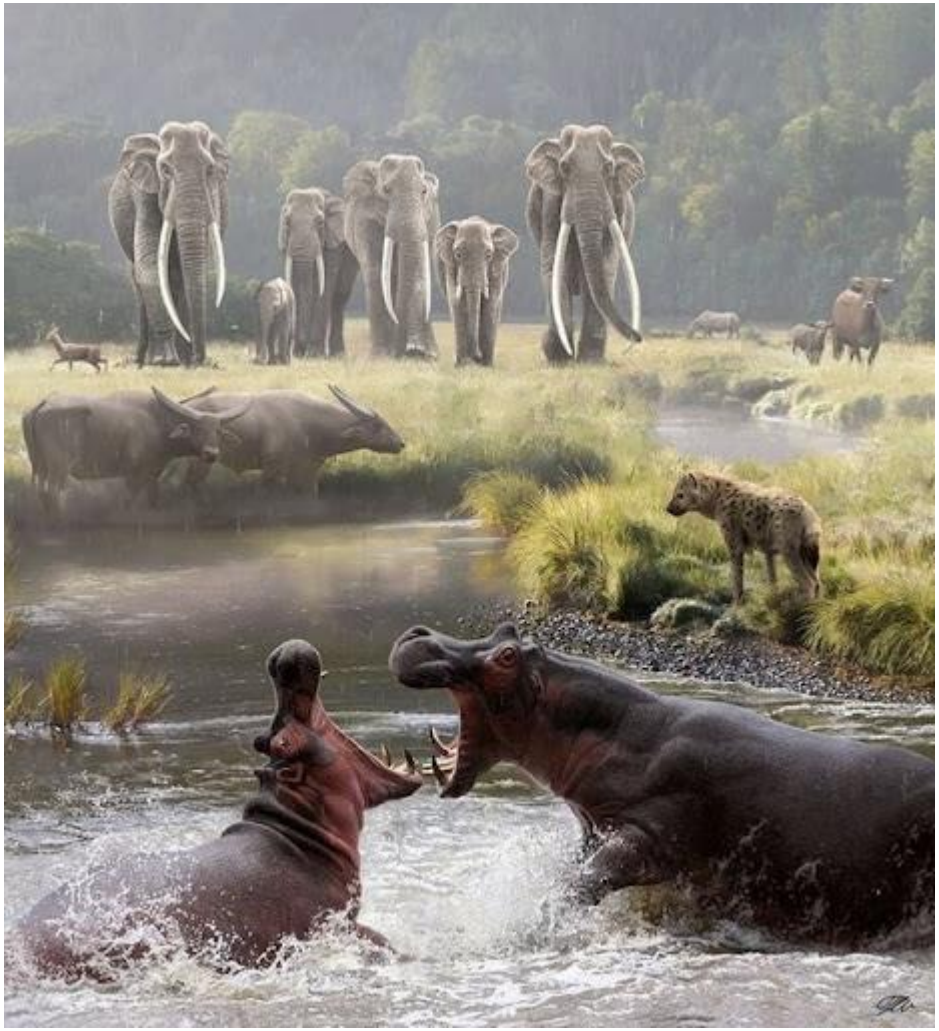
Tree not to scale!

Spotted hyenas are different

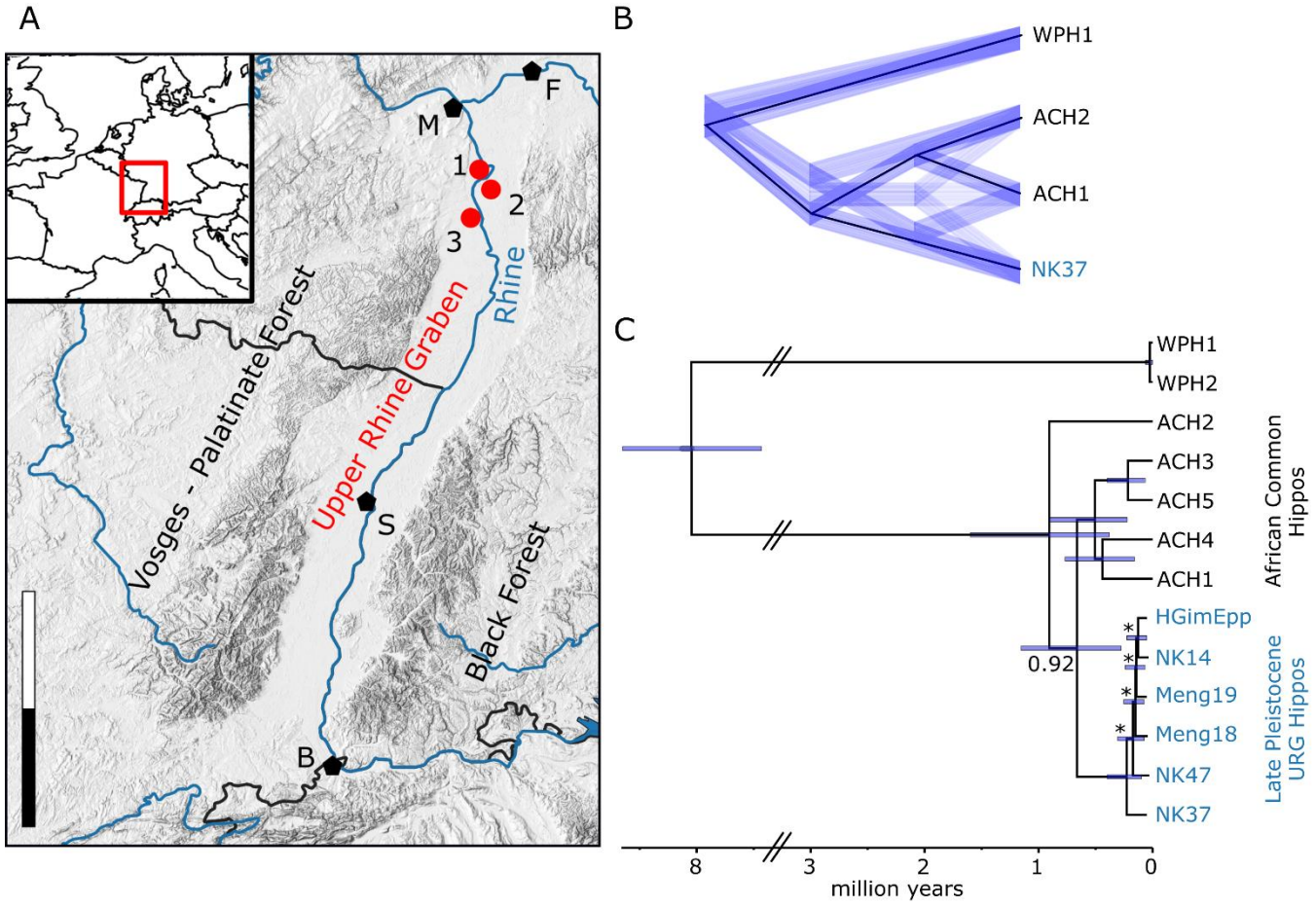
Conclusions

- Spotted and cave hyenas diverged a long time (2.5 my) ago
- There is a discordance between mitochondrial and nuclear tree similar to humans, Neanderthals and Denisovans
- We find evidence for multiple gene flow events between Africa and Eurasia, until almost 2 my after the initial divergence

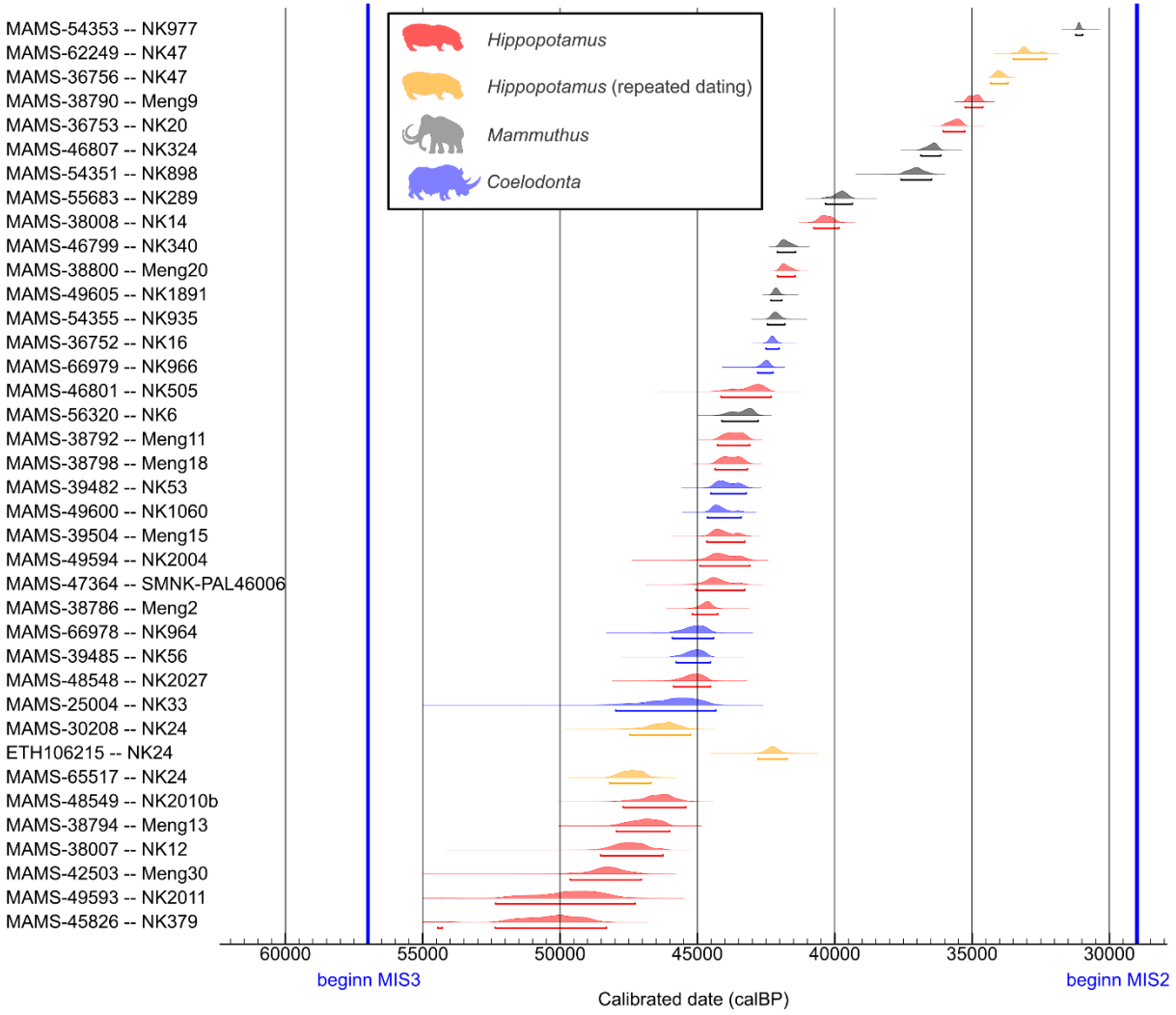
Pleistocene hippos



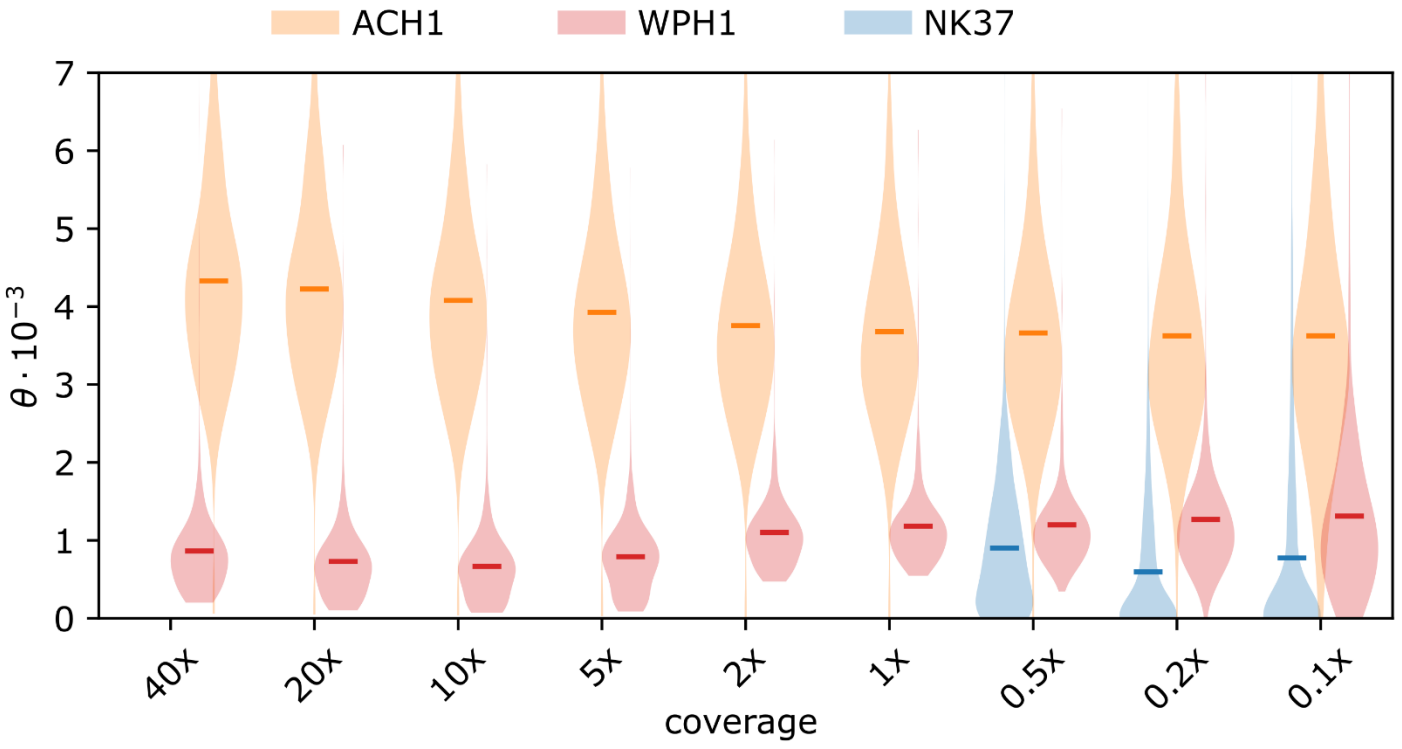
Sampling and phylogenetic trees



Sample ages


MAMS-54353 -- NK977
MAMS-62249 -- NK47
MAMS-36756 -- NK47
MAMS-38790 -- Meng9
MAMS-36753 -- NK20
MAMS-46807 -- NK324
MAMS-54351 -- NK898
MAMS-55683 -- NK289
MAMS-38008 -- NK14
MAMS-46799 -- NK340
MAMS-38800 -- Meng20
MAMS-49605 -- NK1891
MAMS-54355 -- NK935
MAMS-36752 -- NK16
MAMS-66979 -- NK966
MAMS-46801 -- NK505
MAMS-56320 -- NK6
MAMS-38792 -- Meng11
MAMS-38798 -- Meng18
MAMS-39482 -- NK53
MAMS-49600 -- NK1060
MAMS-39504 -- Meng15
MAMS-49594 -- NK2004
MAMS-47364 -- SMNK-PAL46006
MAMS-38786 -- Meng2
MAMS-66978 -- NK964
MAMS-39485 -- NK56
MAMS-48548 -- NK2027
MAMS-25004 -- NK33
MAMS-30208 -- NK24
ETH106215 -- NK24
MAMS-65517 -- NK24
MAMS-48549 -- NK2010b
MAMS-38794 -- Meng13
MAMS-38007 -- NK12
MAMS-42503 -- Meng30
MAMS-49593 -- NK2011
MAMS-45826 -- NK379
Hippopotamus
Hippopotamus (repeated dating)
Mammuthus
Coelodonta
60000
55000
50000
45000
40000
35000
30000
beginn MIS3
beginn MIS2
Calibrated date (calBP)

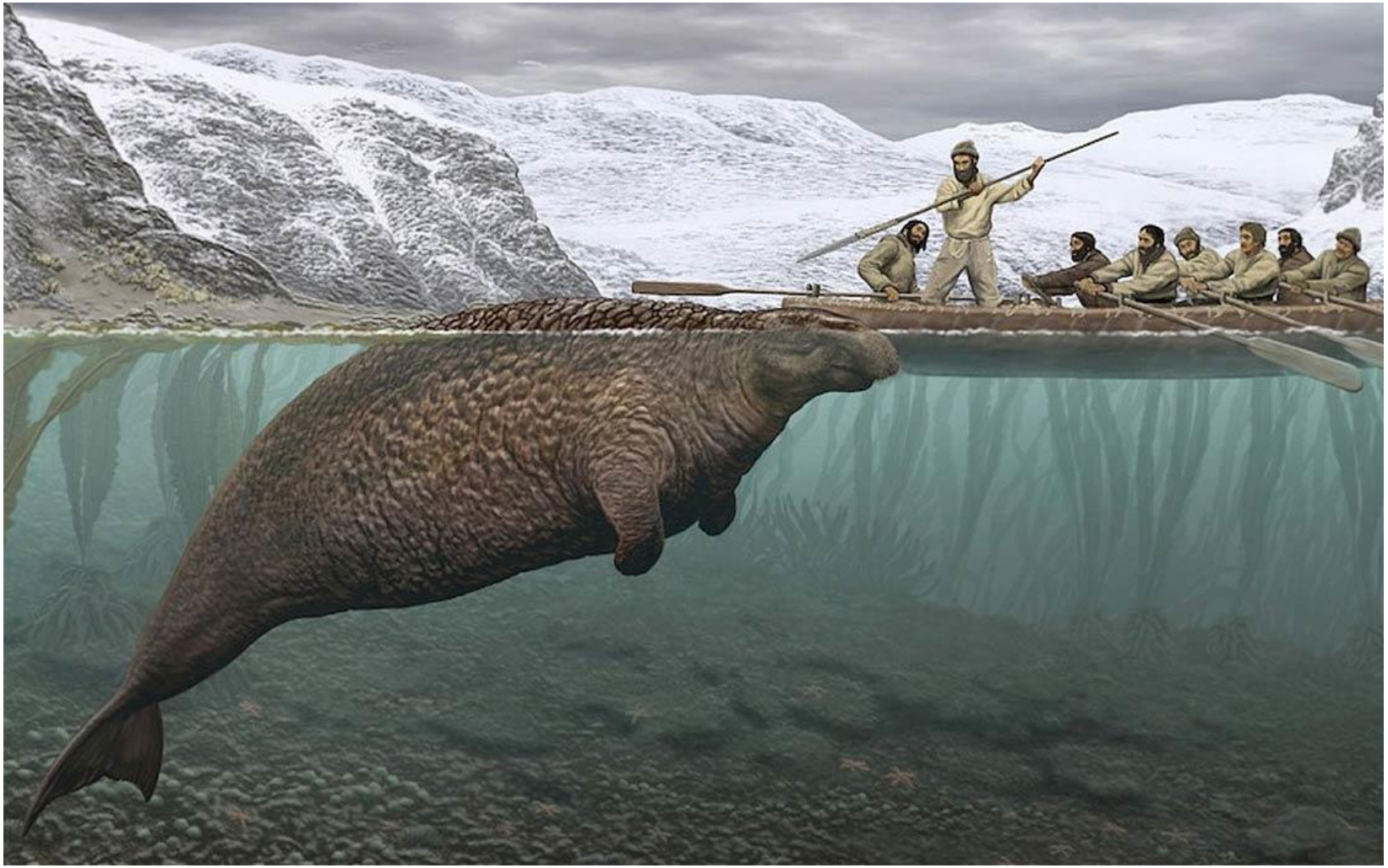
Low genetic diversity



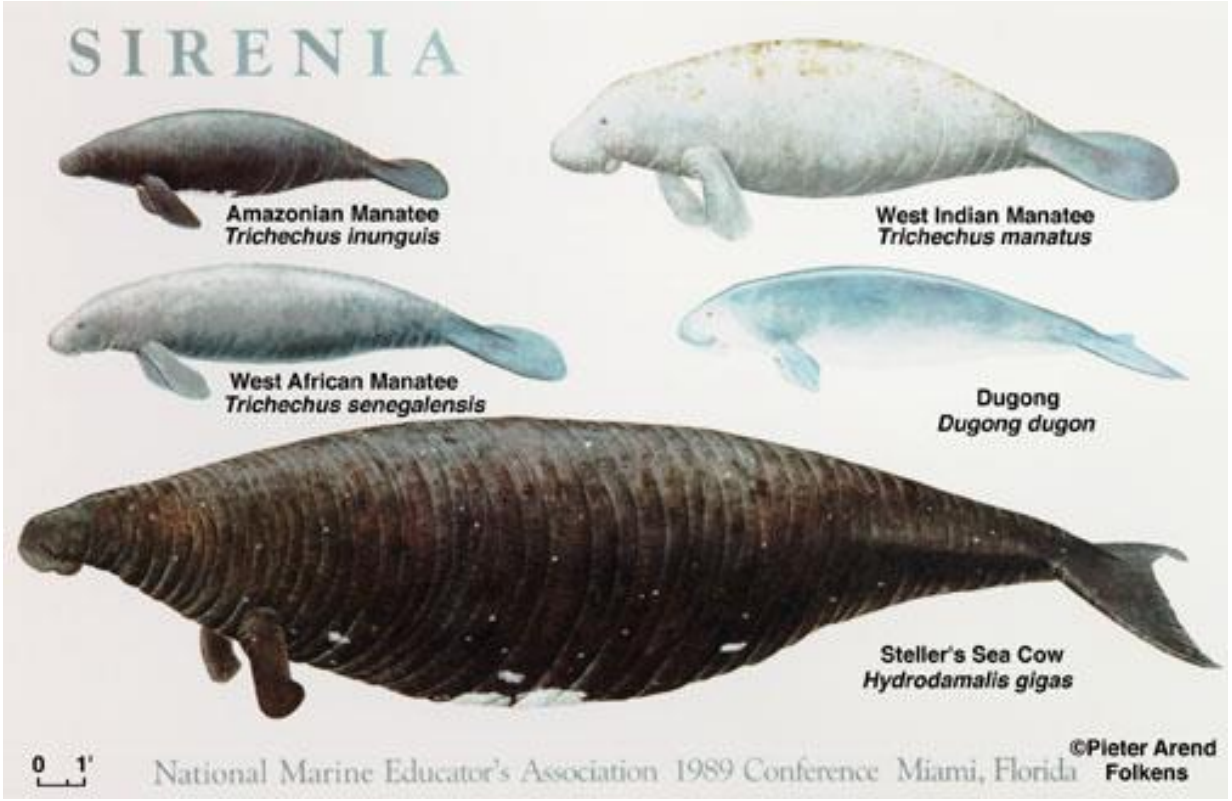
Summary hippos

- Late Pleistocene European hippos were closely related to extant African ones
- The Rhine valley hippos are not Eemian in age
- Pleistocene ecosystems were not analogous to modern ones
- The Rhine valley hippos were a relict population

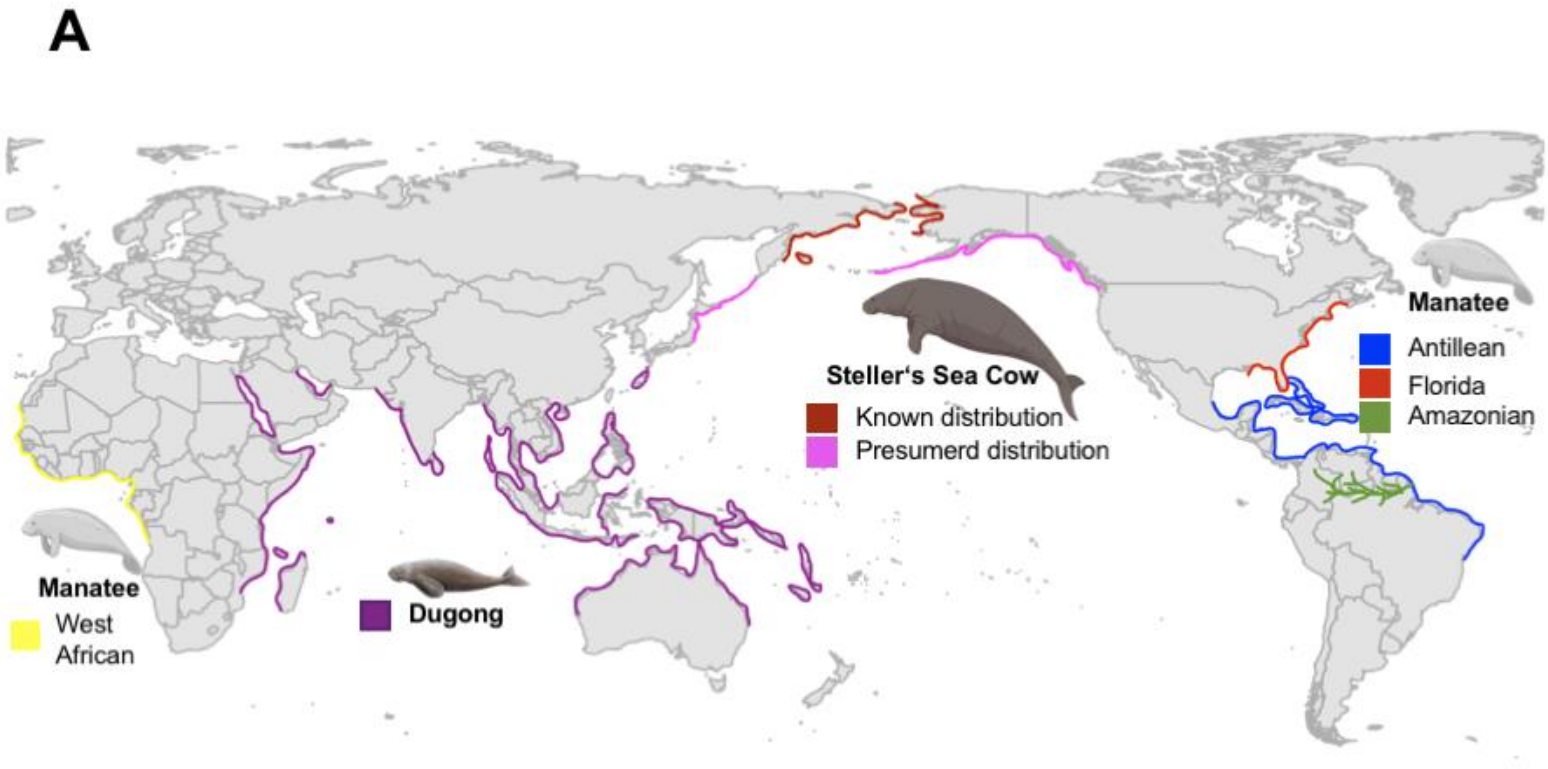
Steller‘s sea cow: † 1768



Size comparison

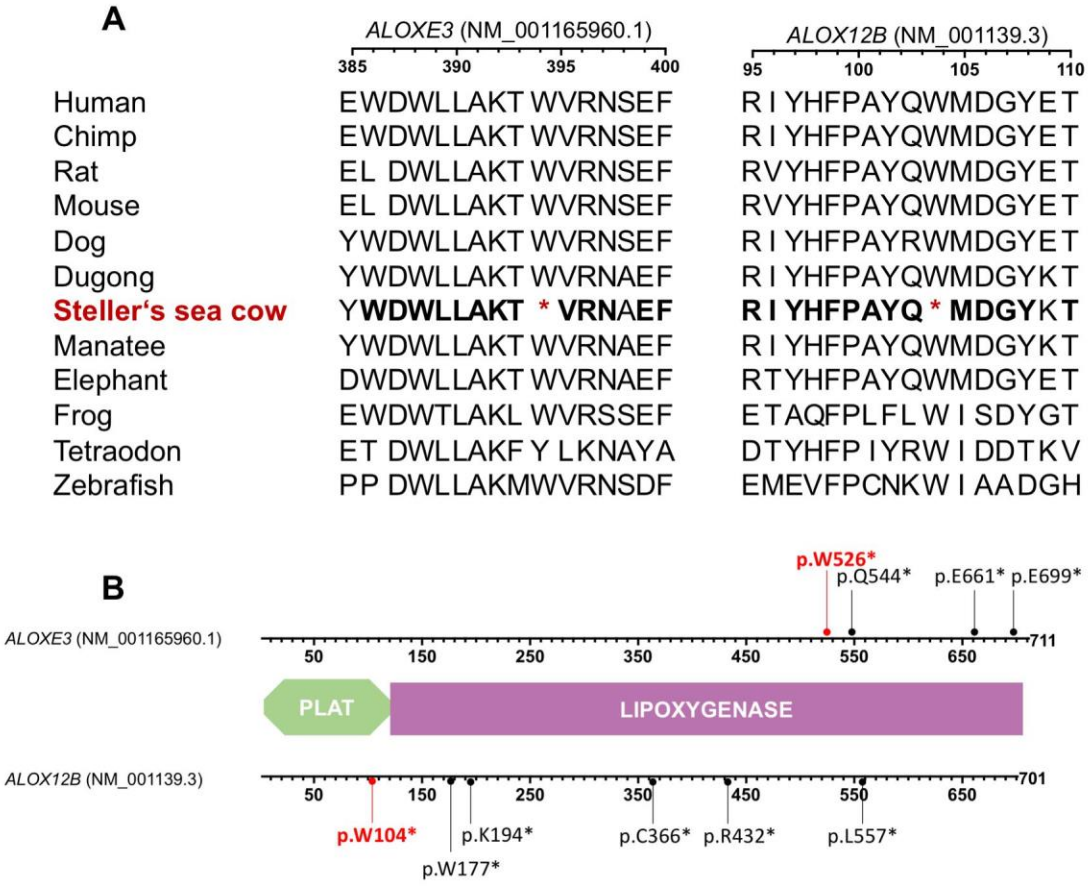


Distribution

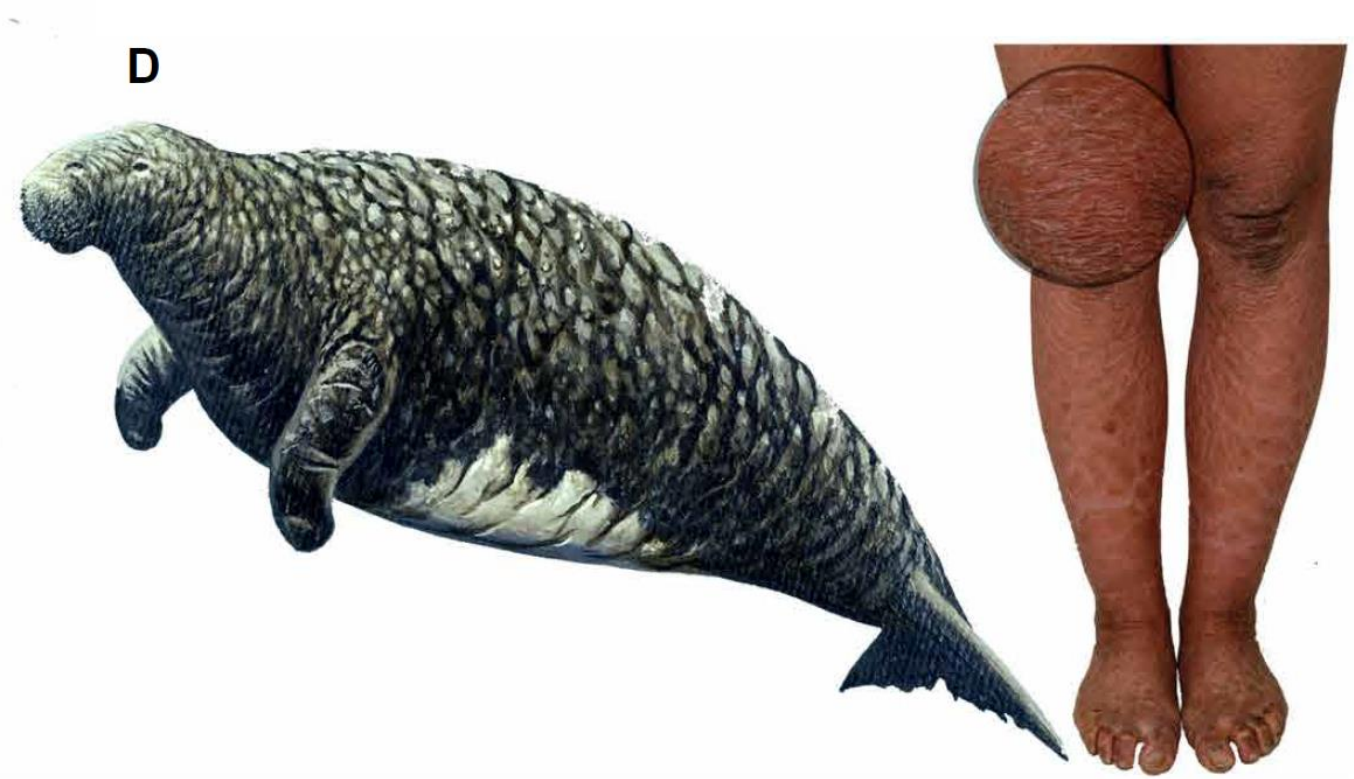


Steller‘s sea cow skin

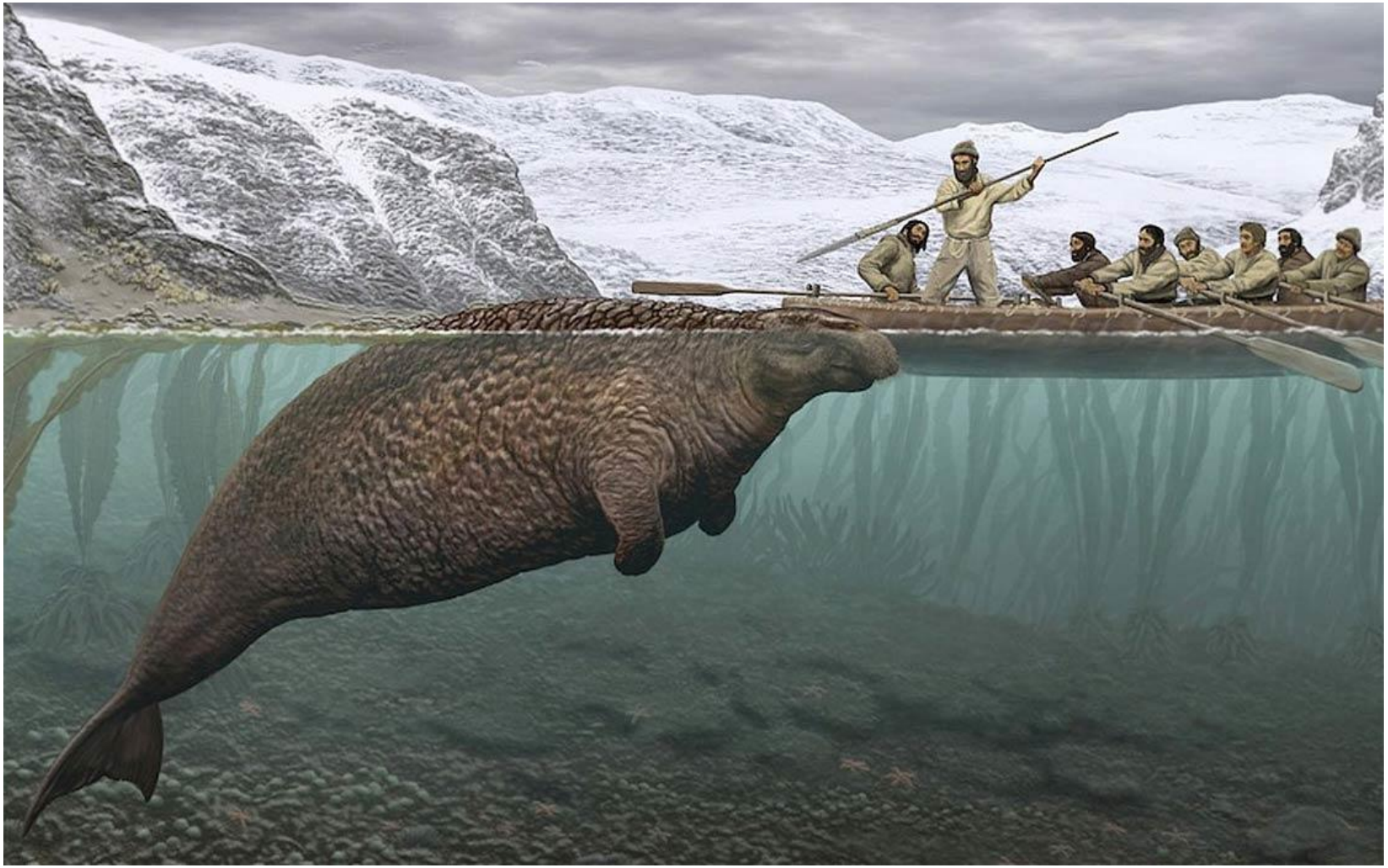
Functional analysis



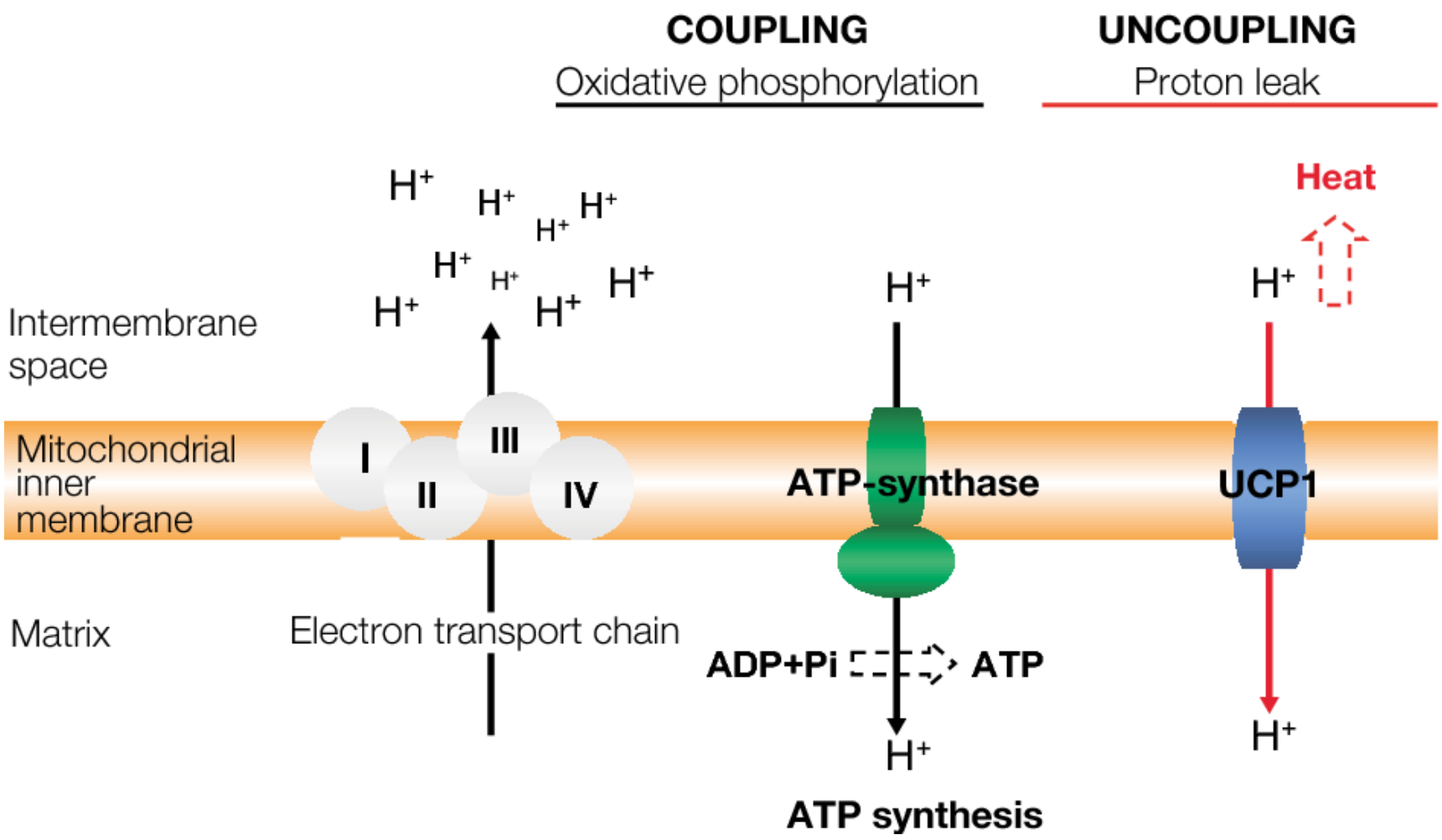
Explaining the phenotype



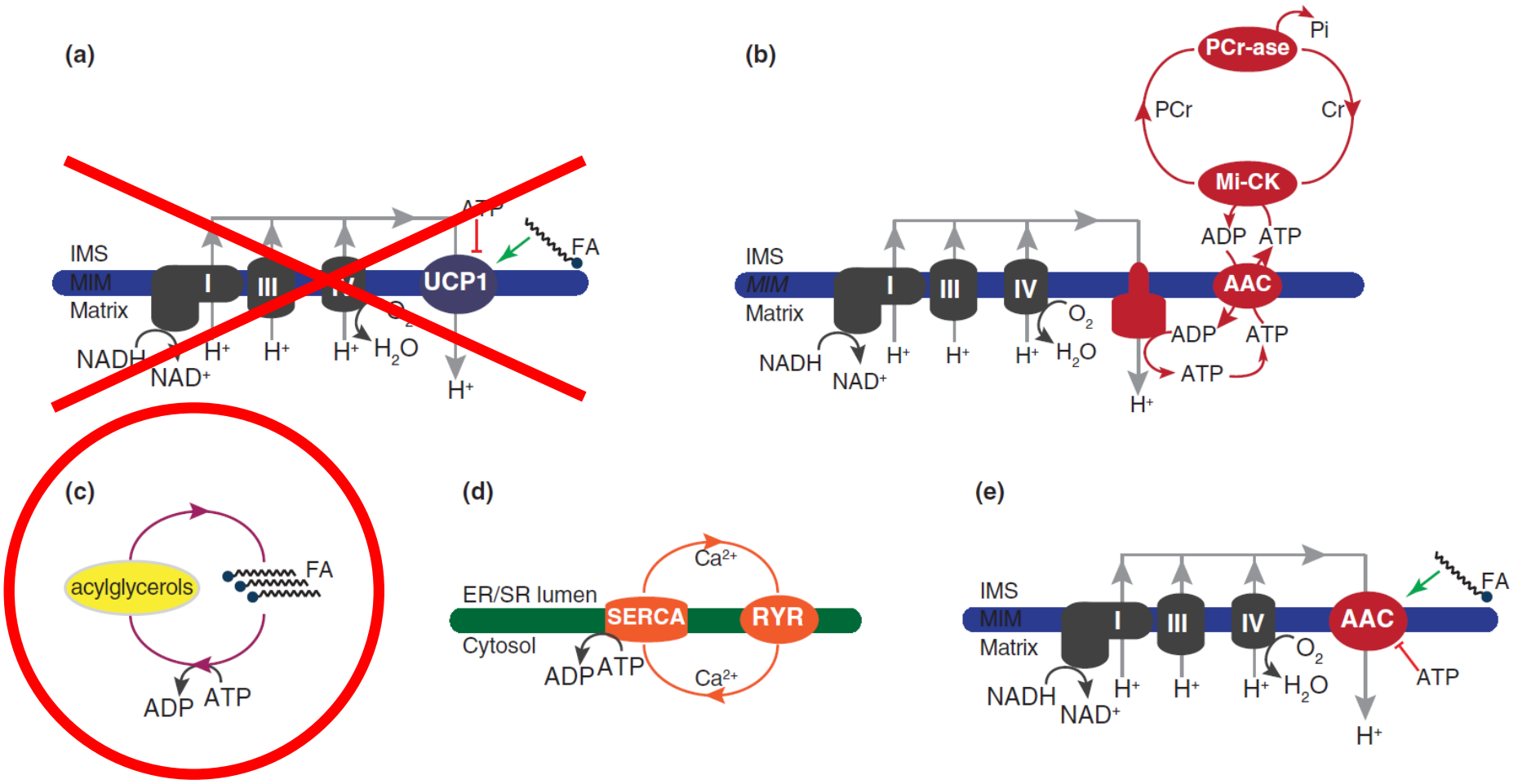
Steller's sea cow: habitat



Non-shivering thermogenesis: UCP1

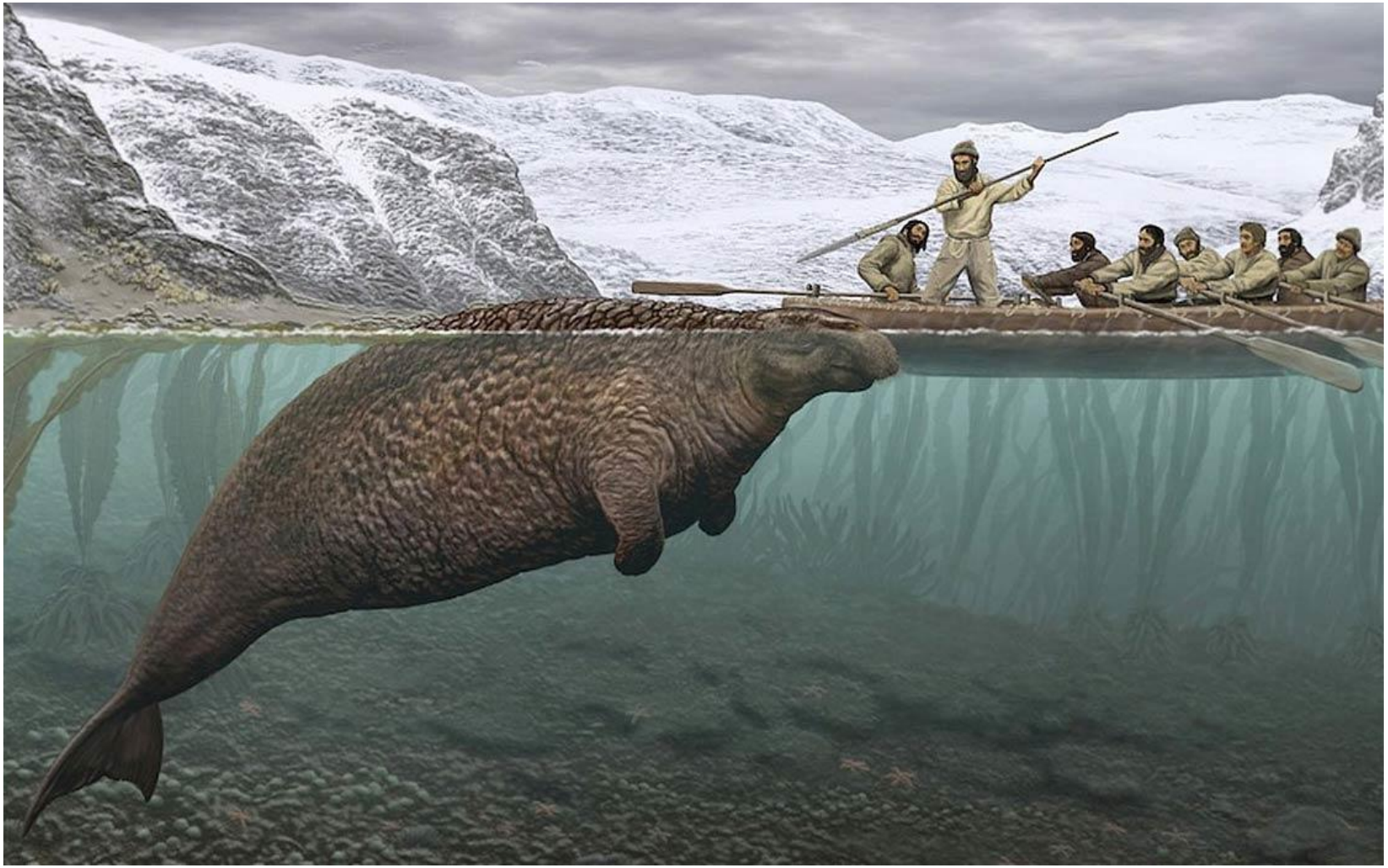


Thermogenic futile cycles

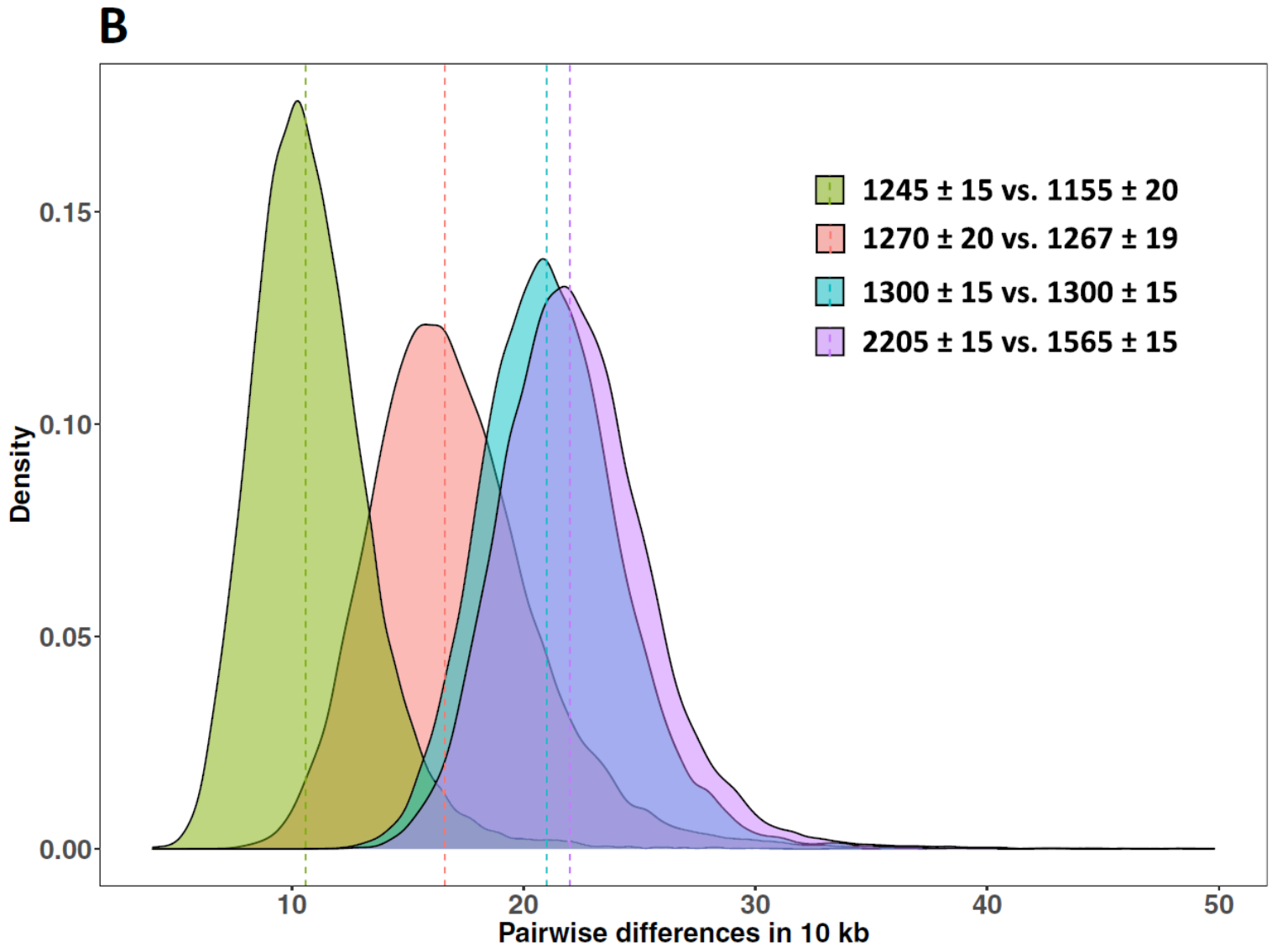


Steller‘s sea cow fat deposits

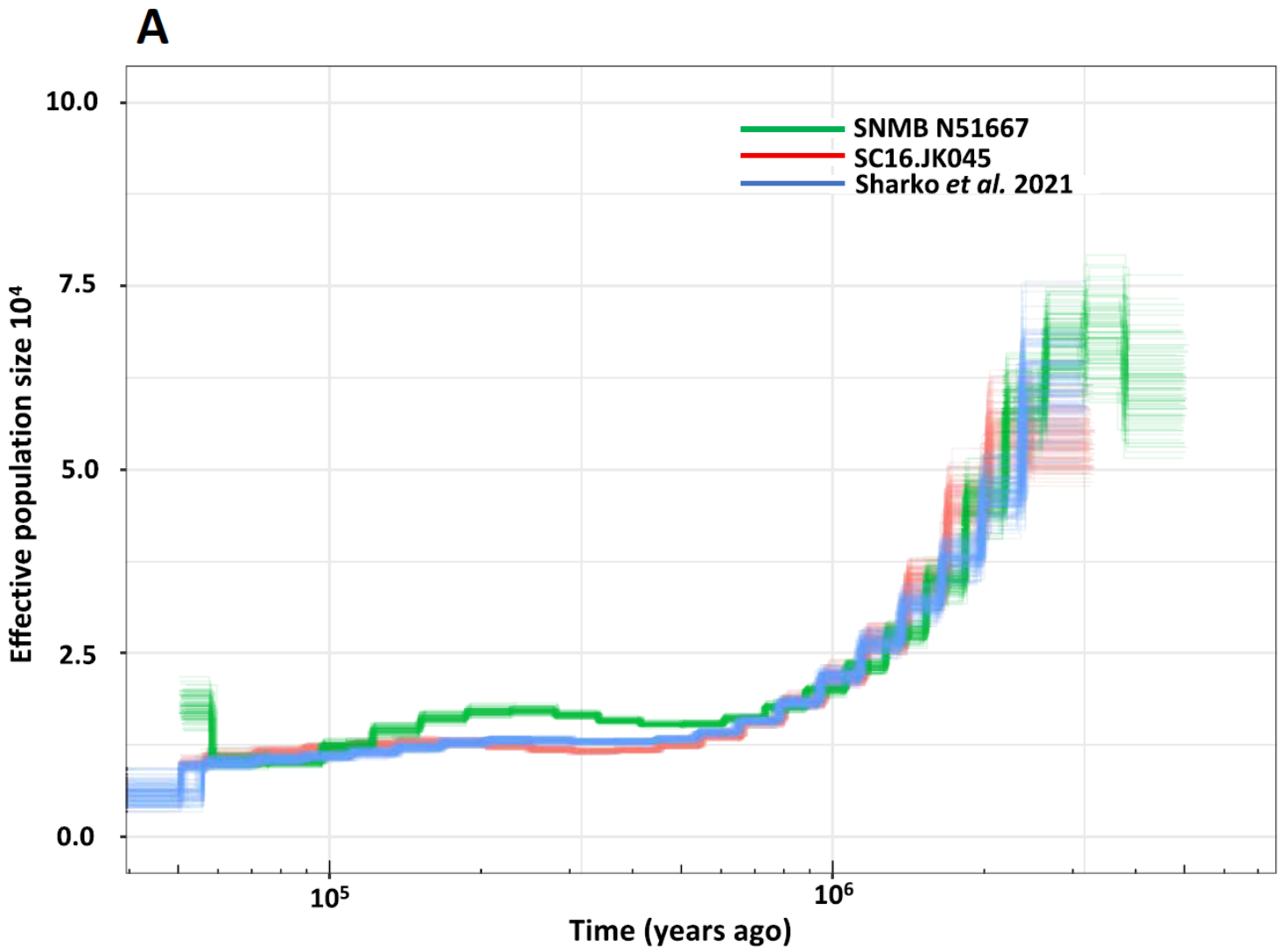
Steller‘s sea cow: † 1768



Genetic diversity over time



Population size over time

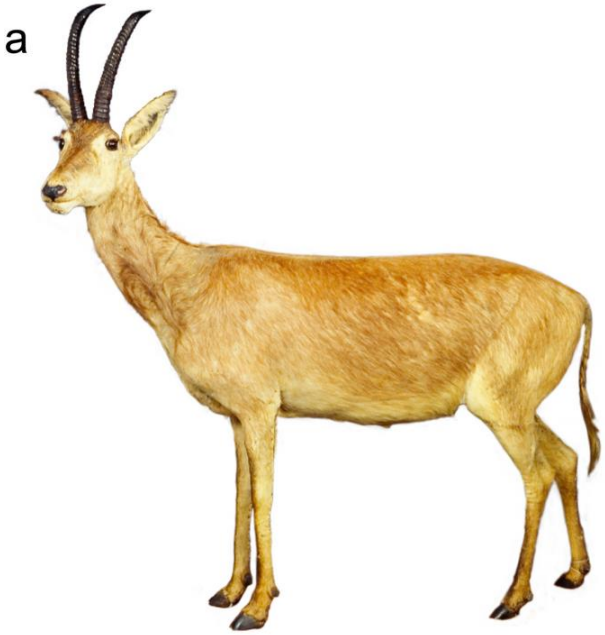


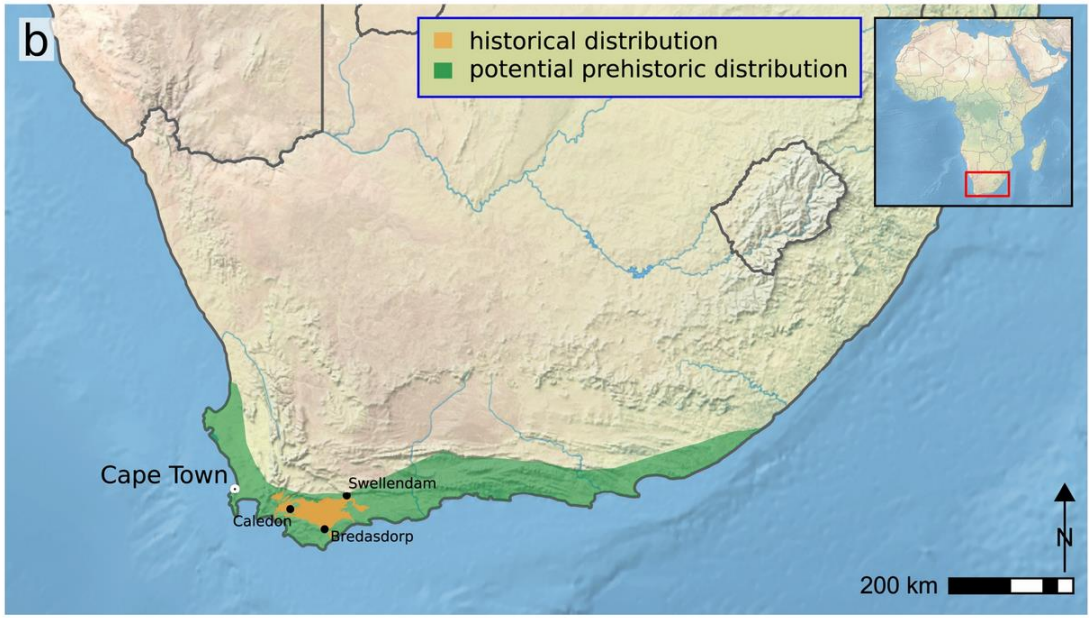
Summary Steller's sea cow

- Low population size for the last 1,000,000 years
- Further loss of genetic diversity between 2,000 and 1,000 years ago
- Genetic cause for skin phenotype and probably cold-adaptation revealed

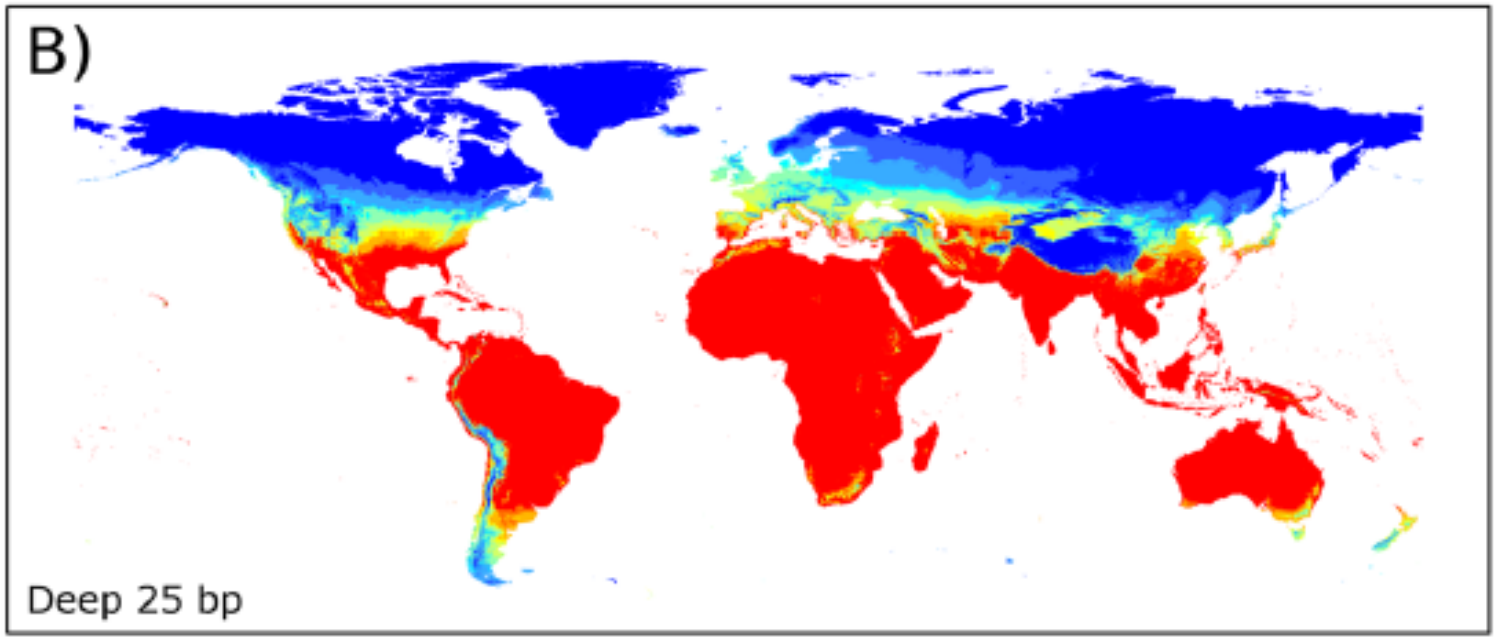
South-African bluebuck † ~1800

Distribution

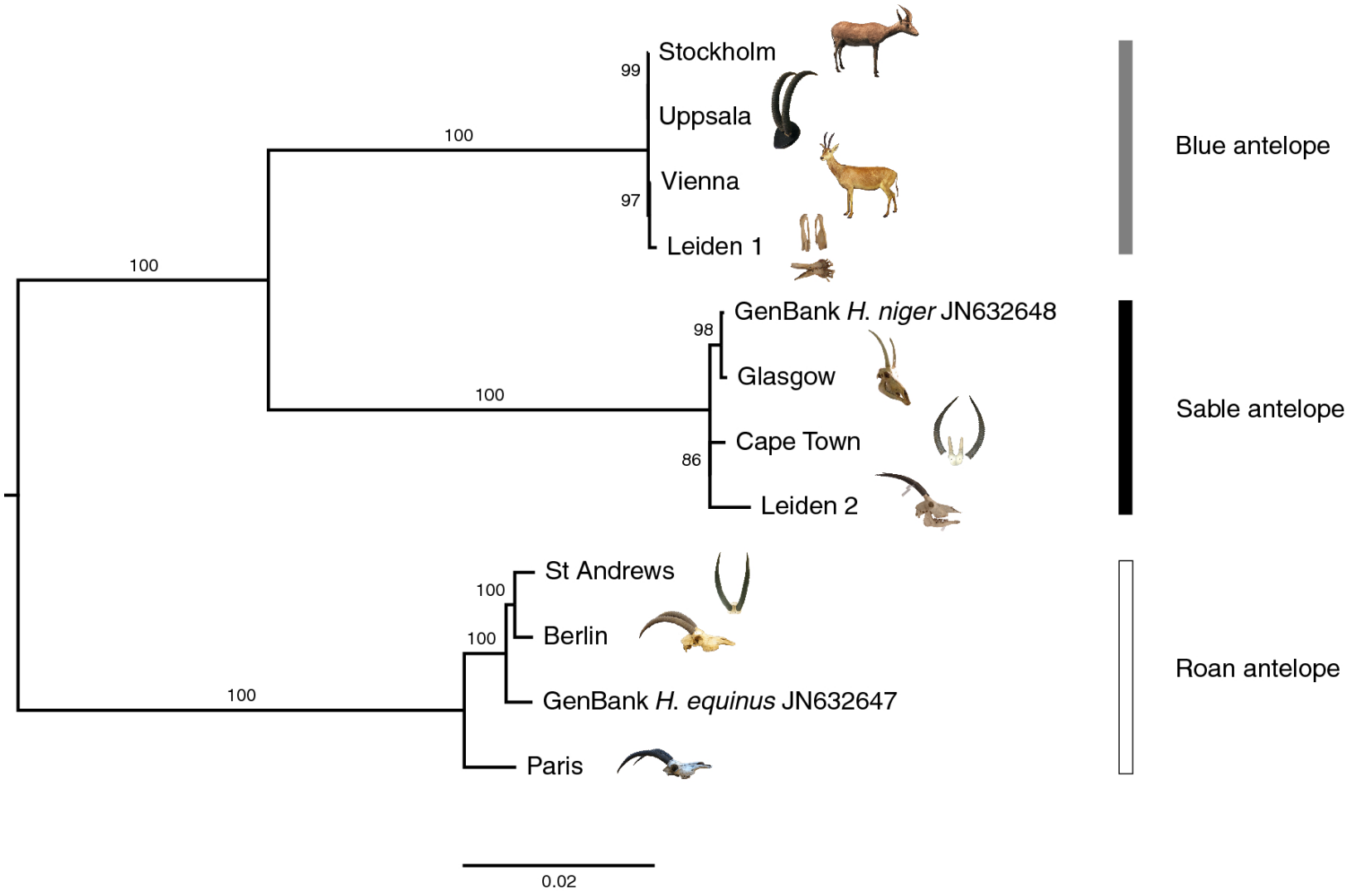




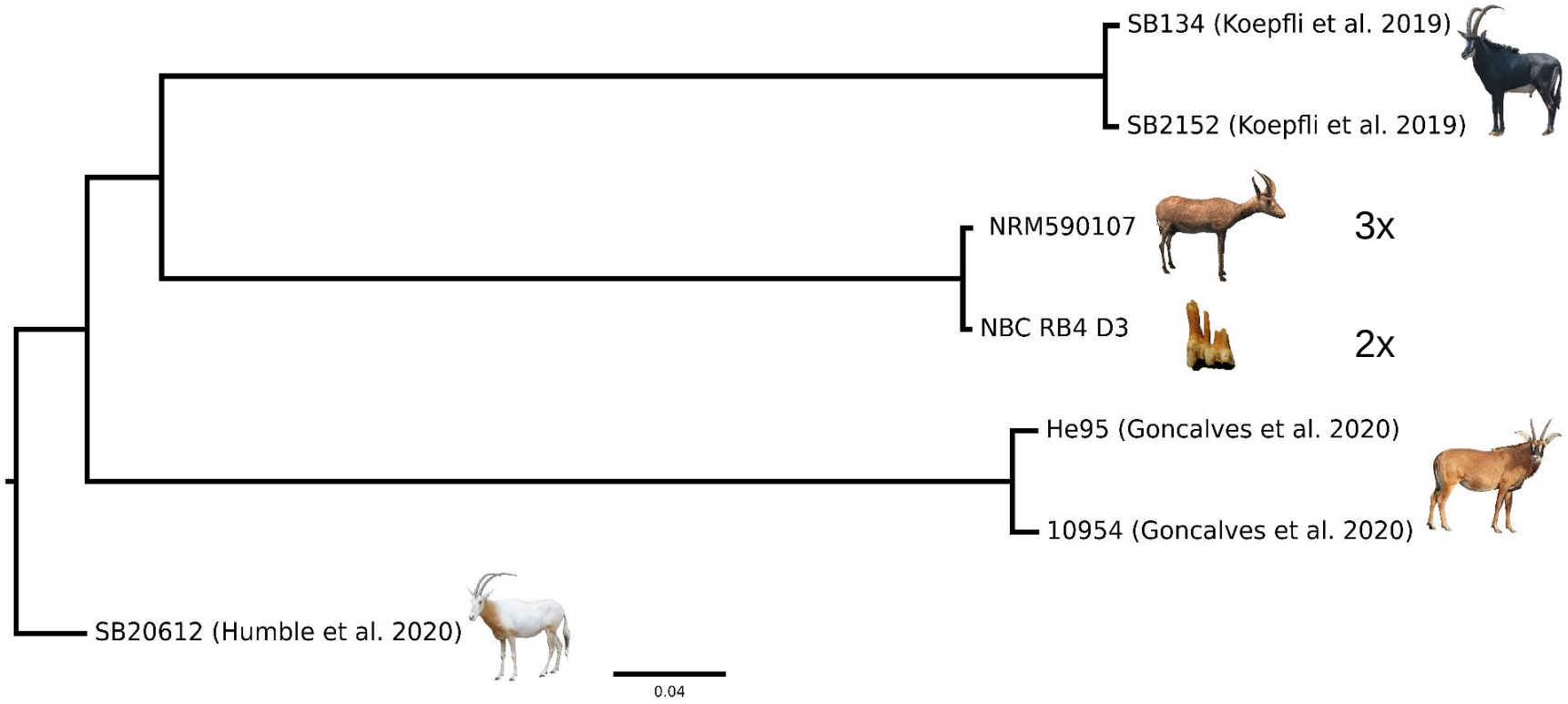
Problem 1: Poor DNA preservation



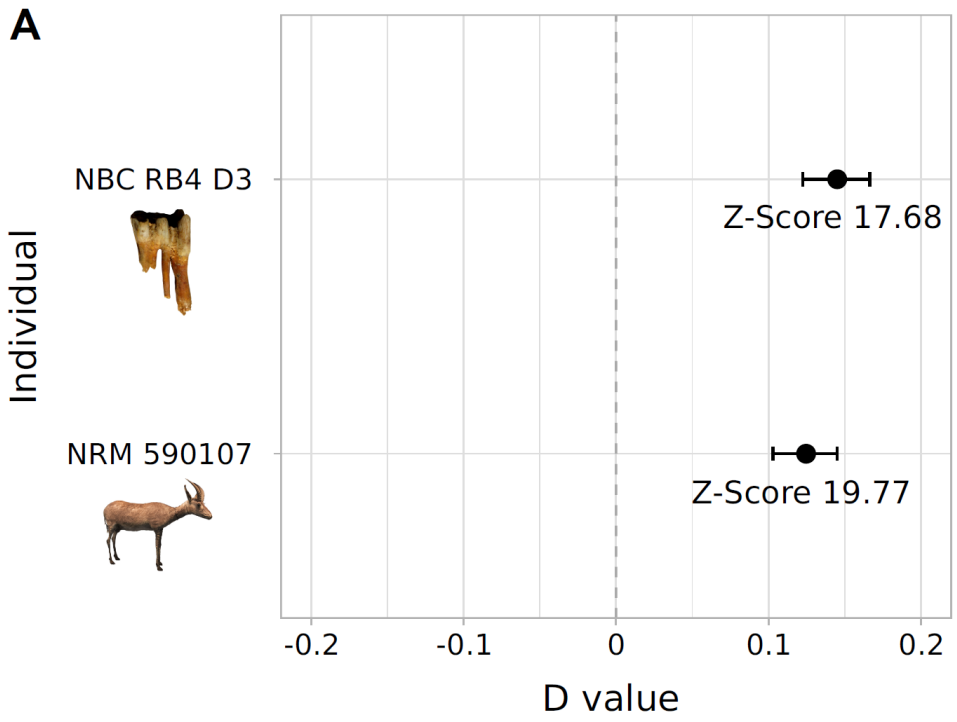
Problem 2: Few museum samples

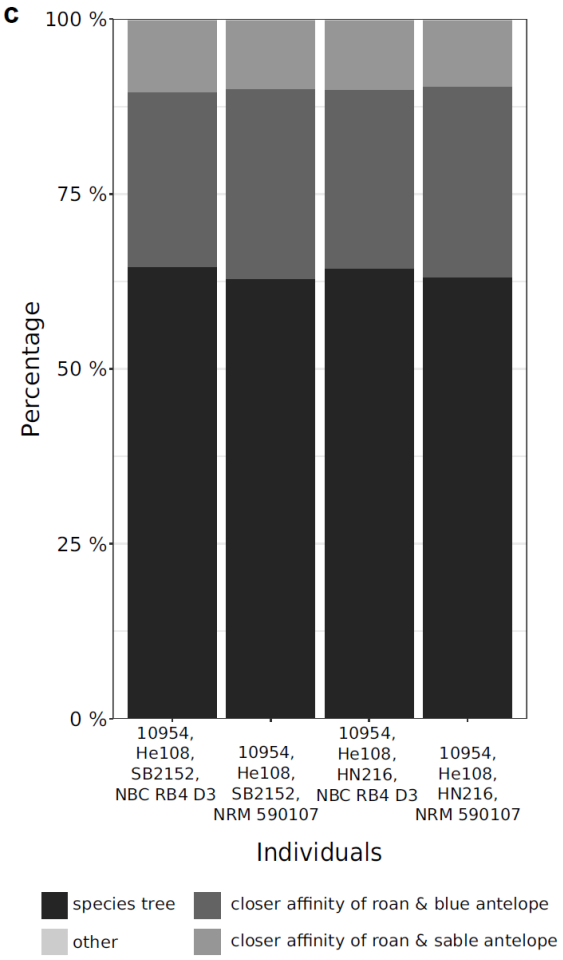


Initial nuclear sequencing at low coverage

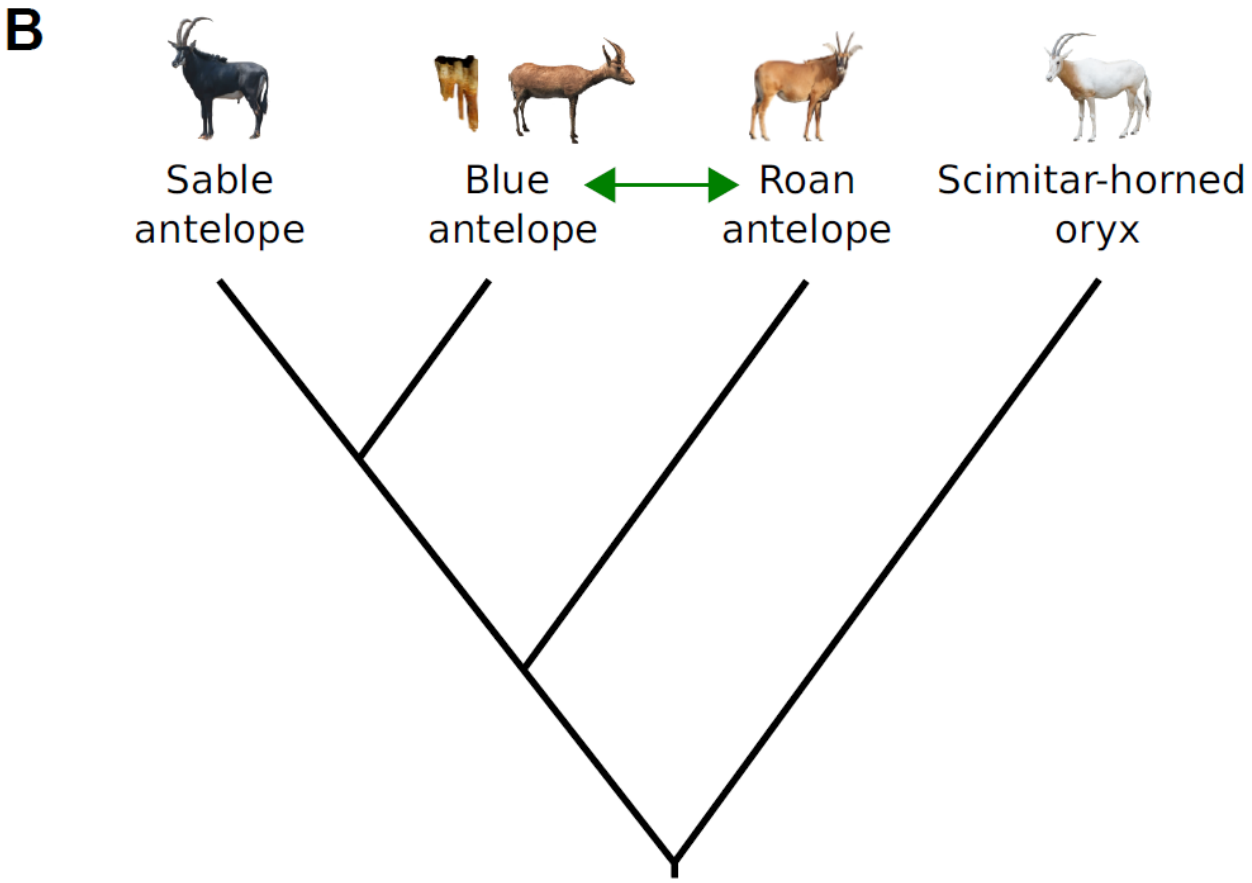


Evidence for post-divergence gene flow

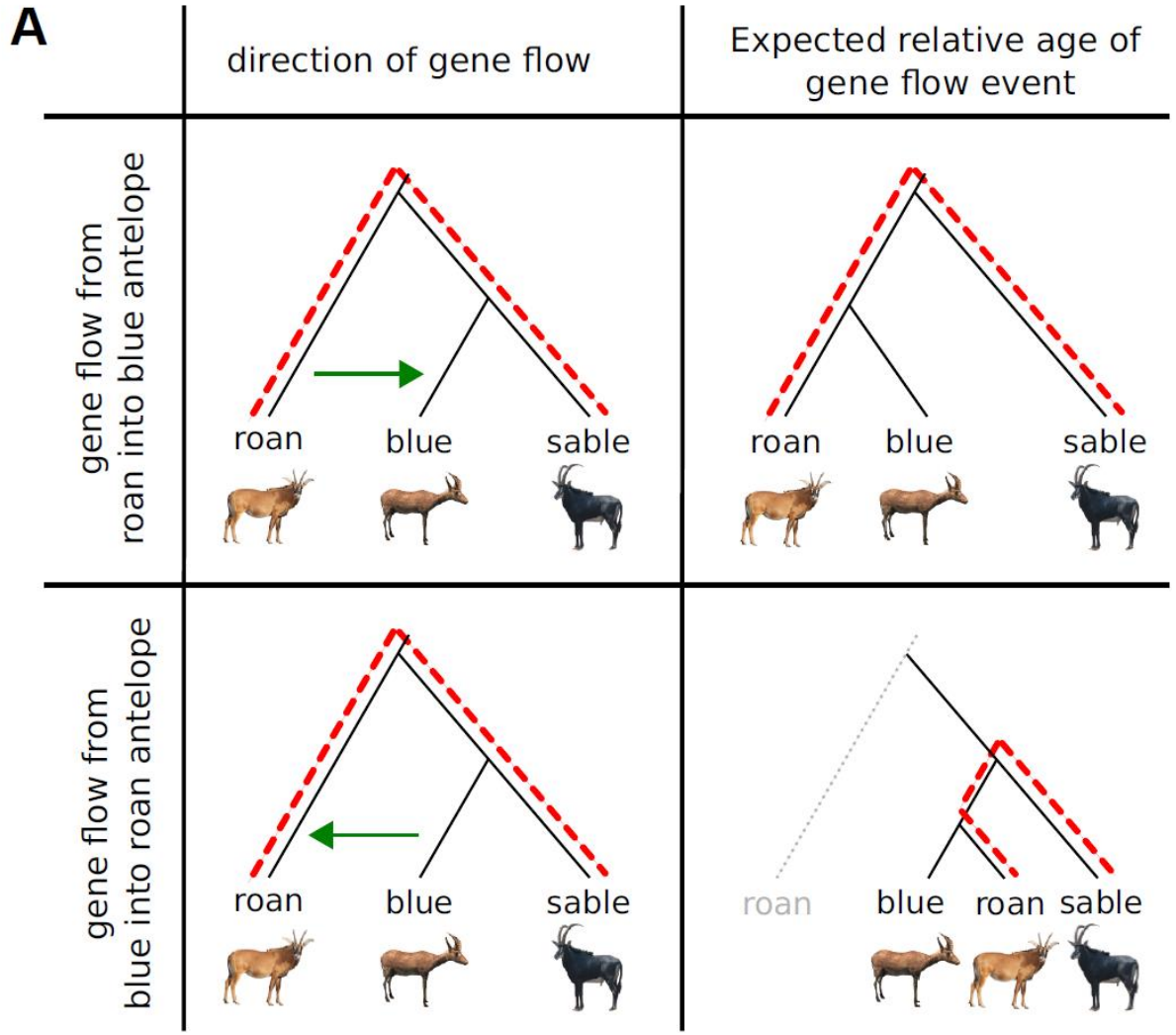




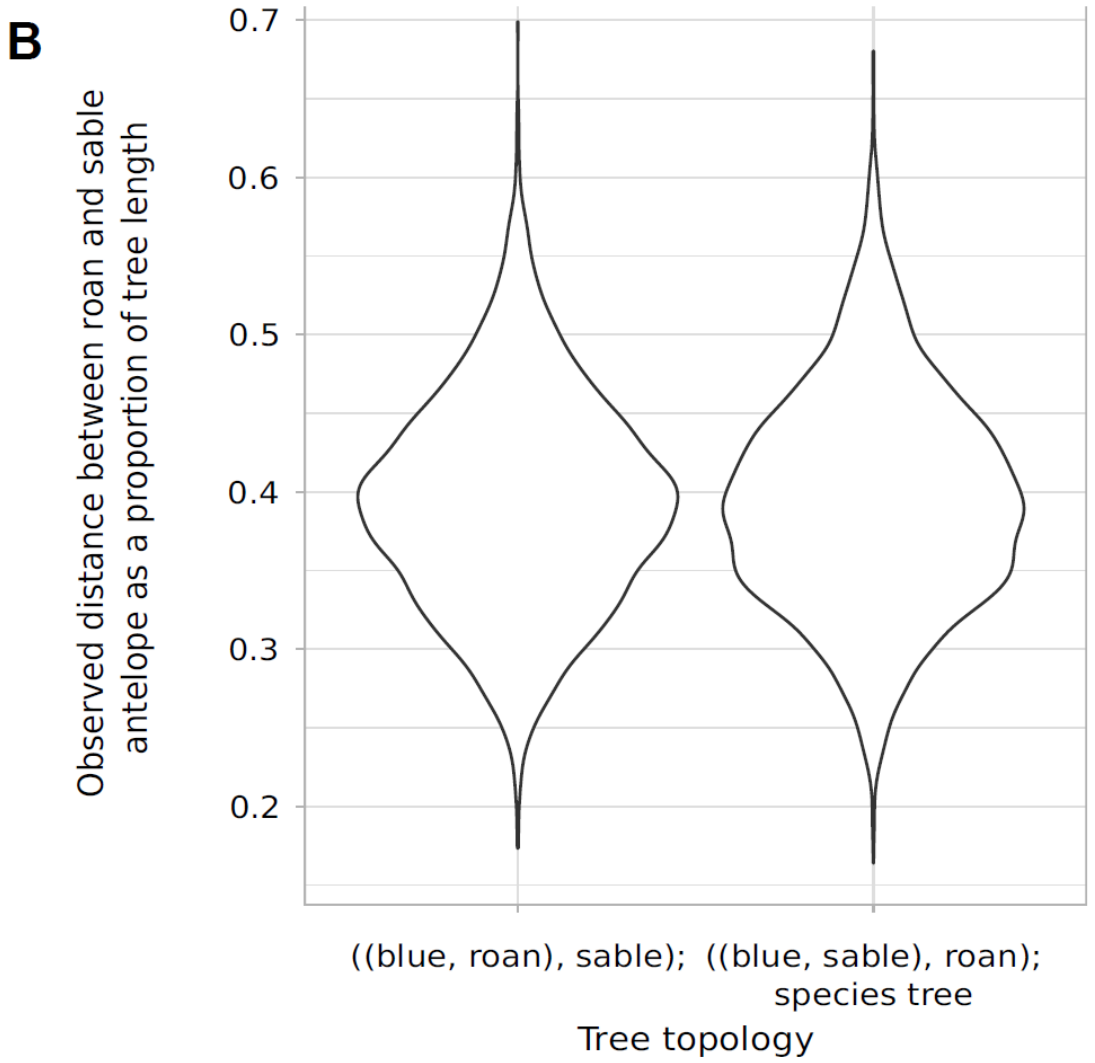
How to determine its direction?



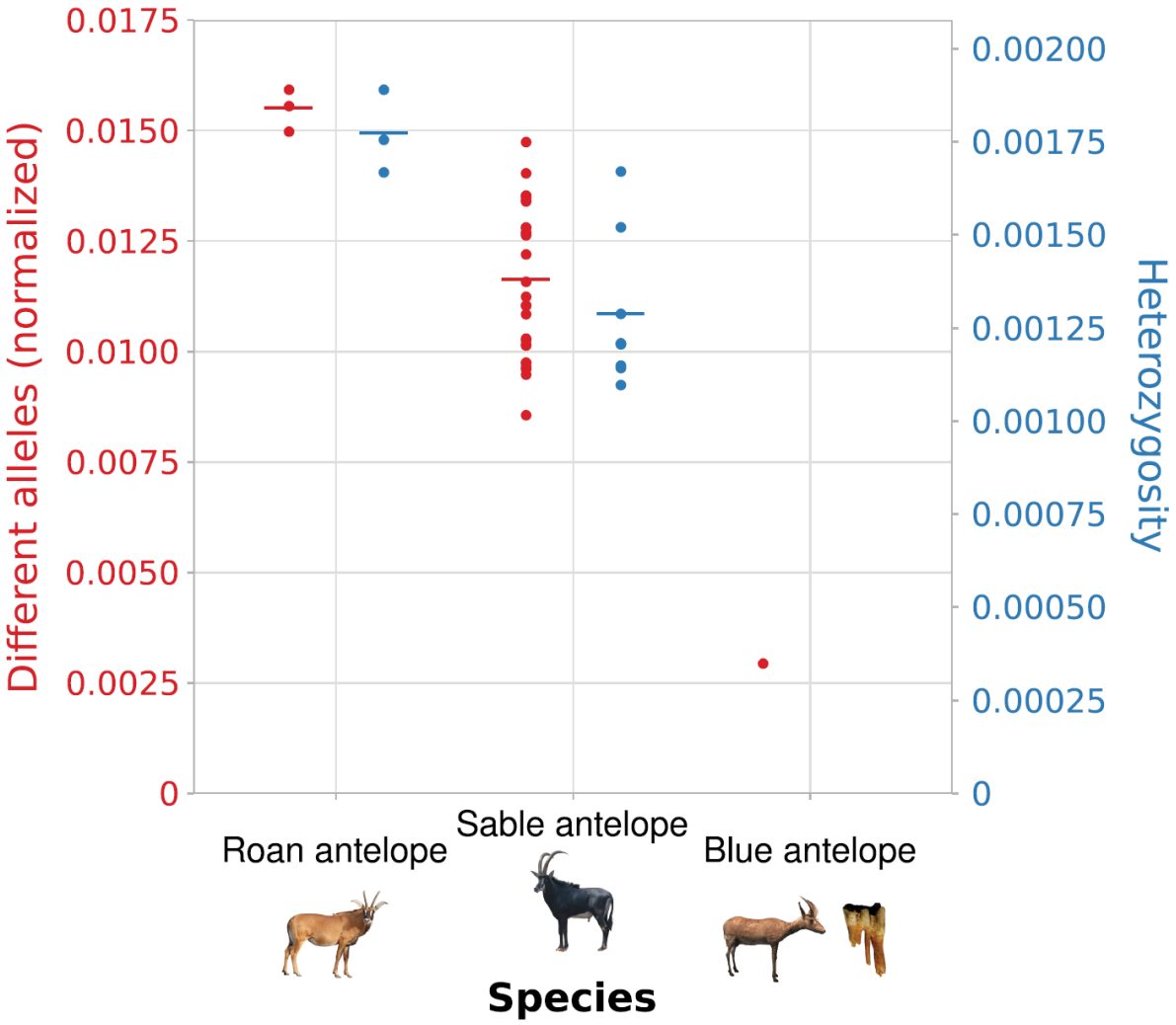
How to determine its direction?



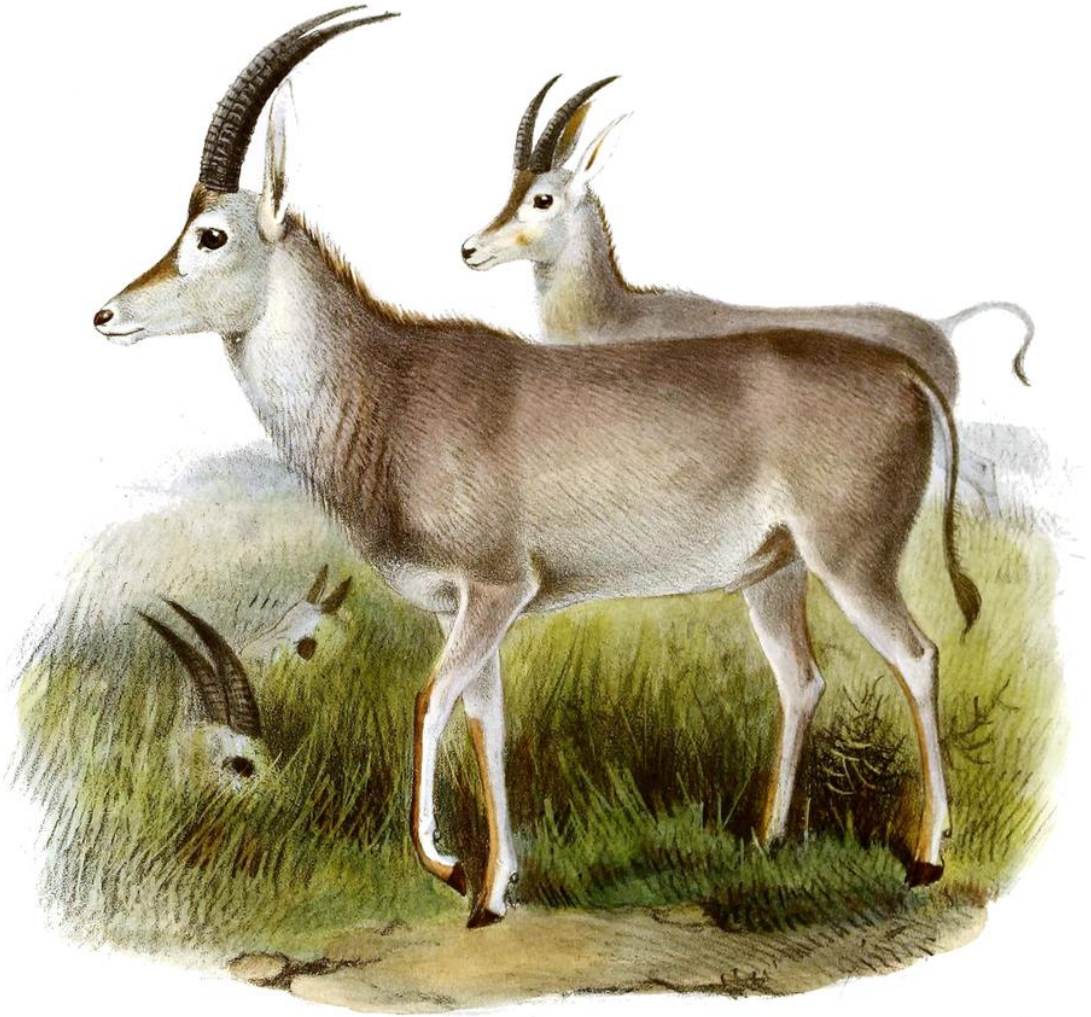
How to determine its direction?



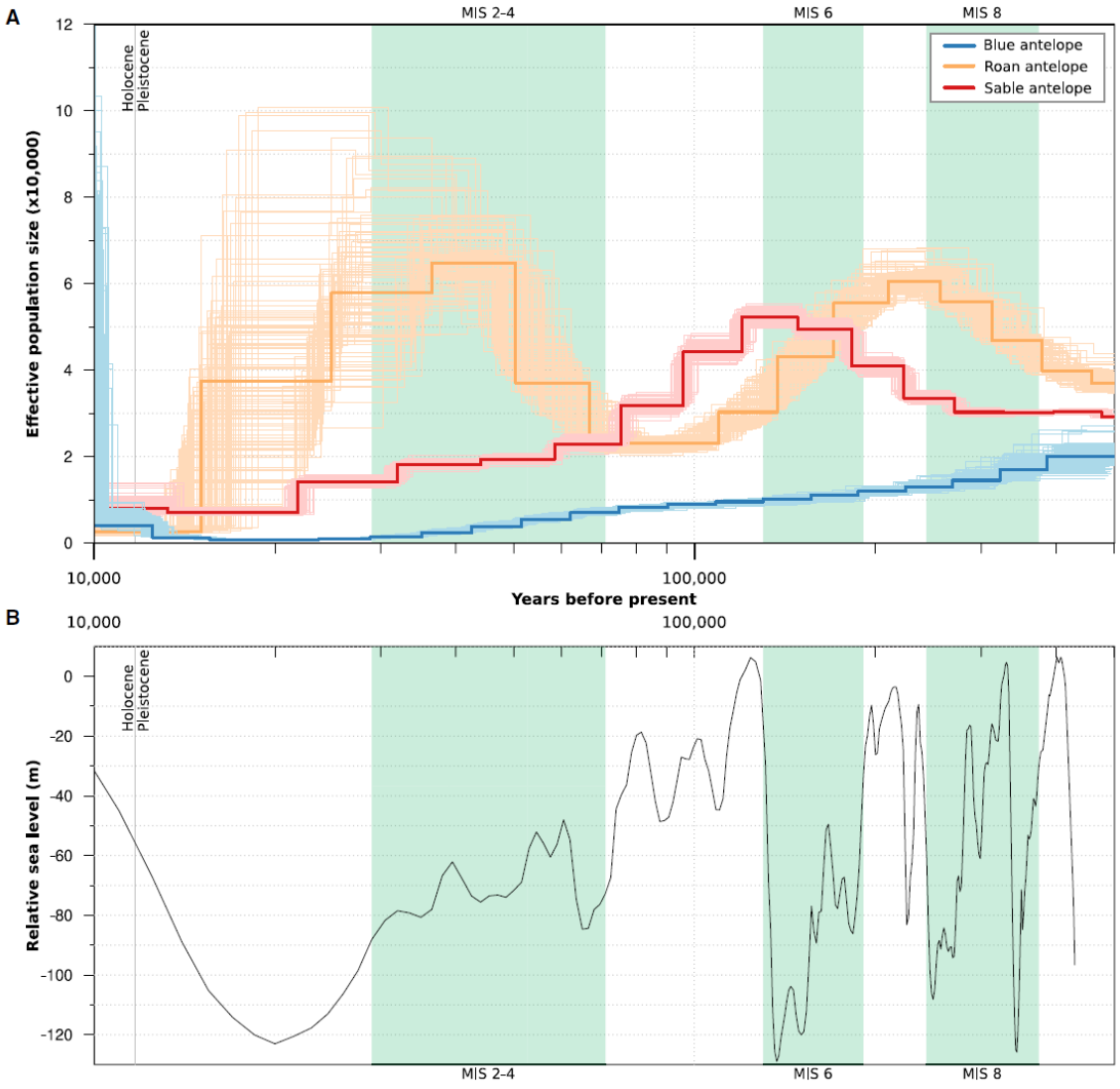
Low genetic diversity



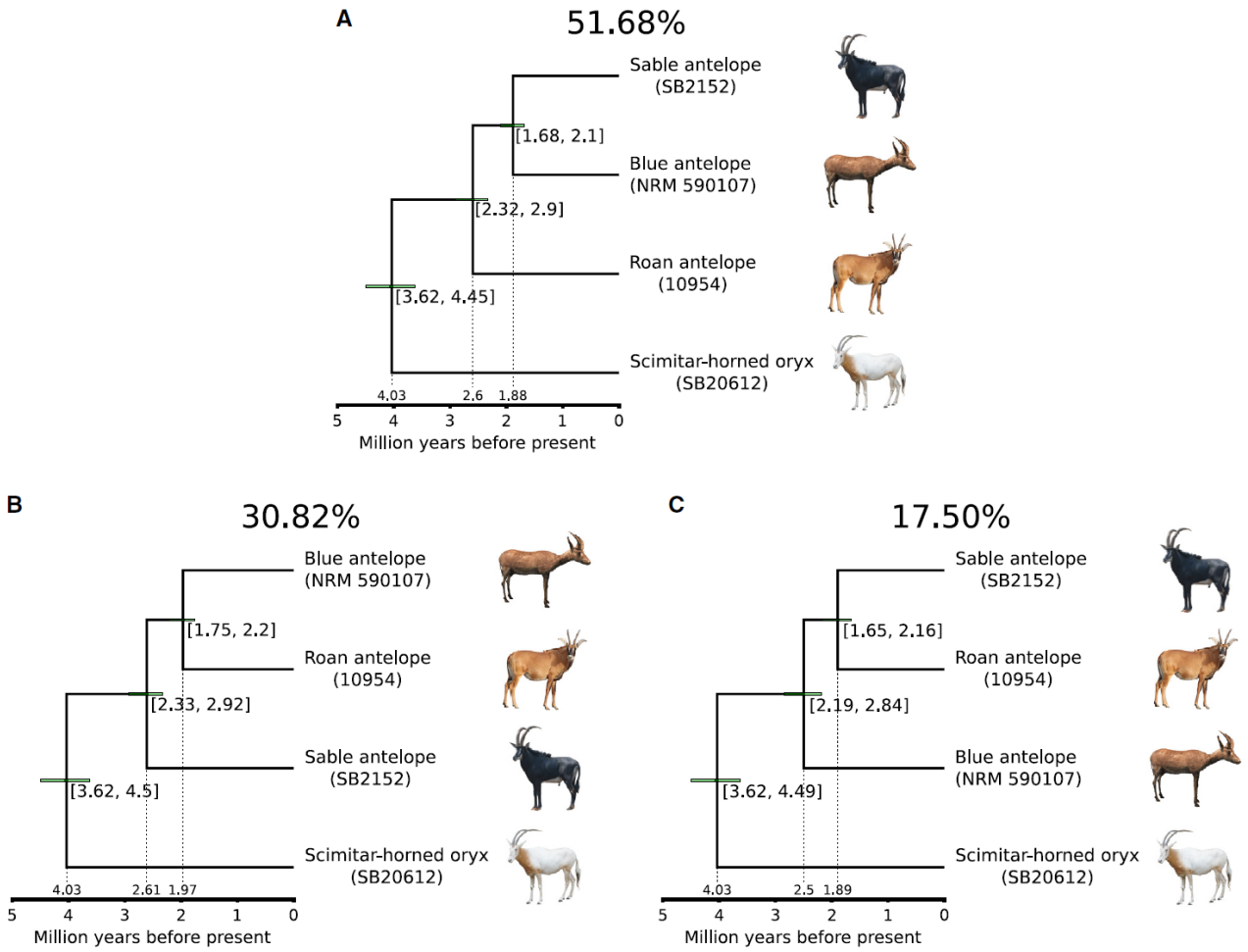
Coverage improved to 40x



Long-term low population size



Dating alternative topologies



Summary bluebuck

- Low population size for the last 500,000 years
- No connection between population size and climatic fluctuations
- Asymmetric gene flow happed during second speciation

=> not really secondary gene flow


STAR TREK VI
THE UNDISCOVERED COUNTRY

Can de-extinction become real?

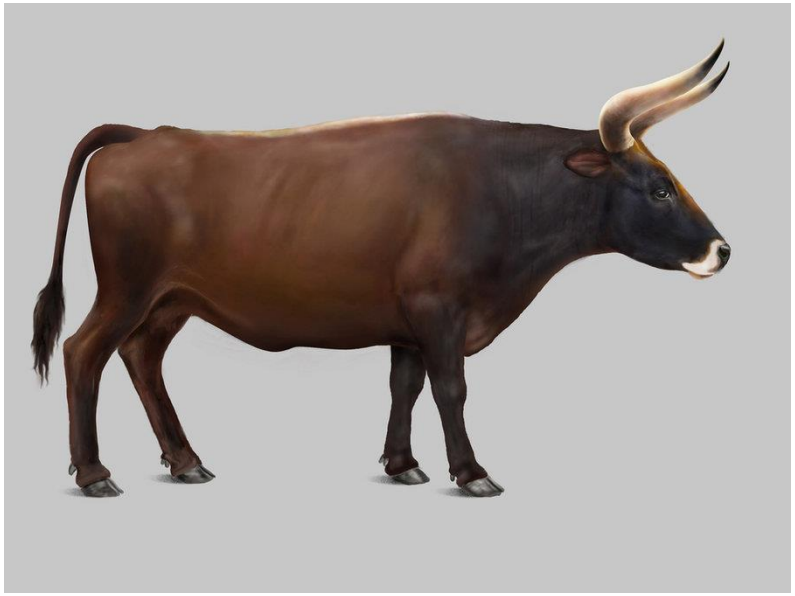
Phenotype reconstruction based on mammoth nuDNA (possibly contaminated)

Jurassic Park? No

Mammoth safari? Maybe....

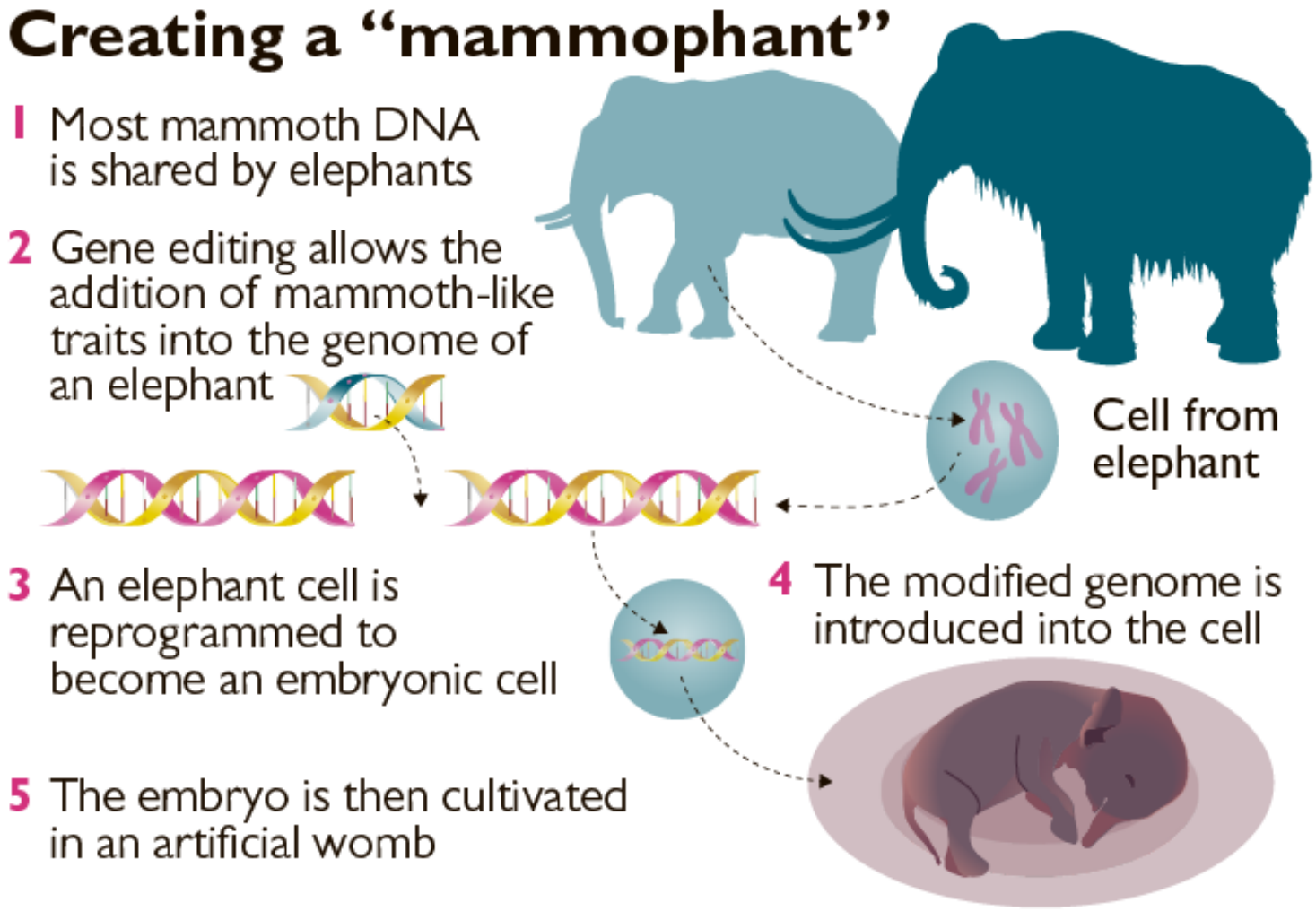
De-extinction: 3 possibilities

1. Backbreeding

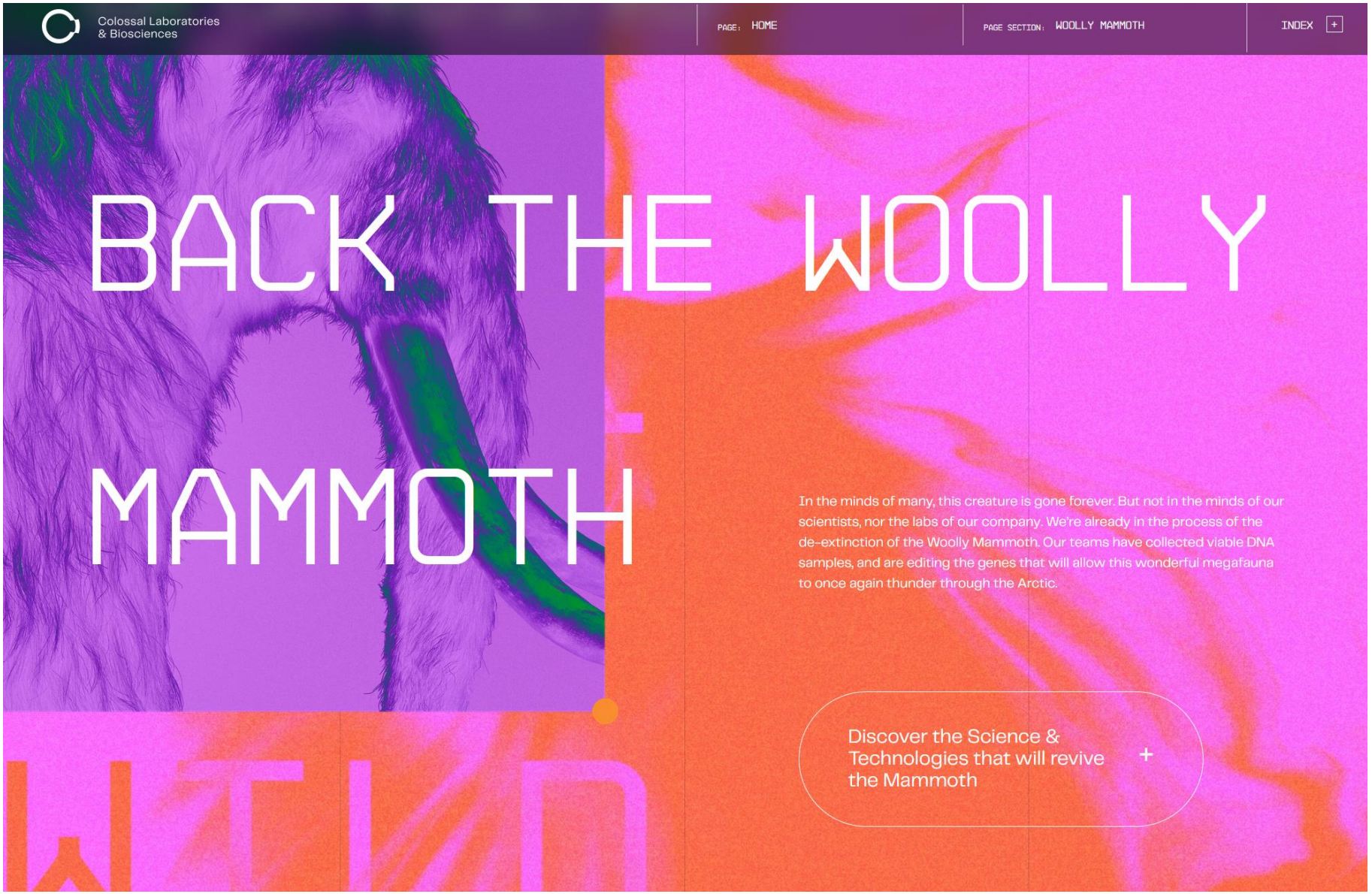


2. Cloning

3. Genome-modification

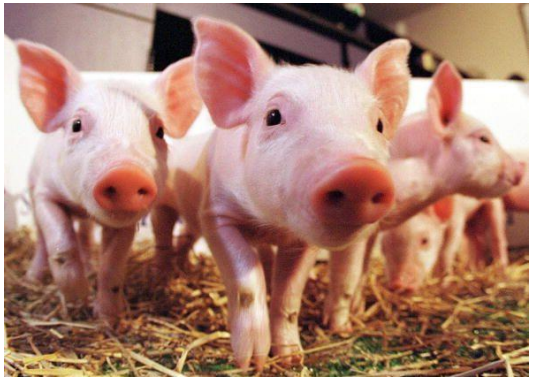


Colossal Biosciences



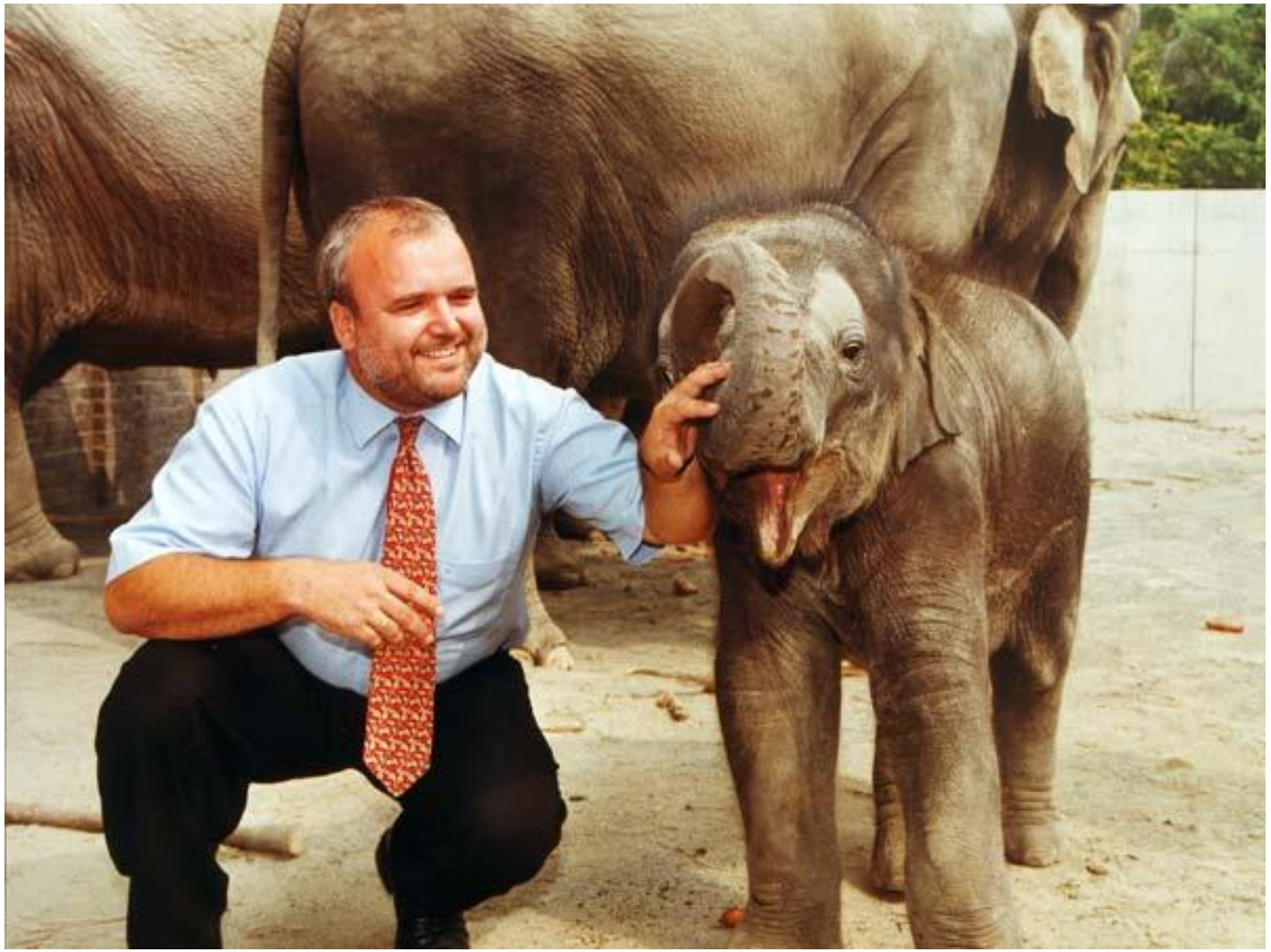
Mammoths are iconic

Cloned animals

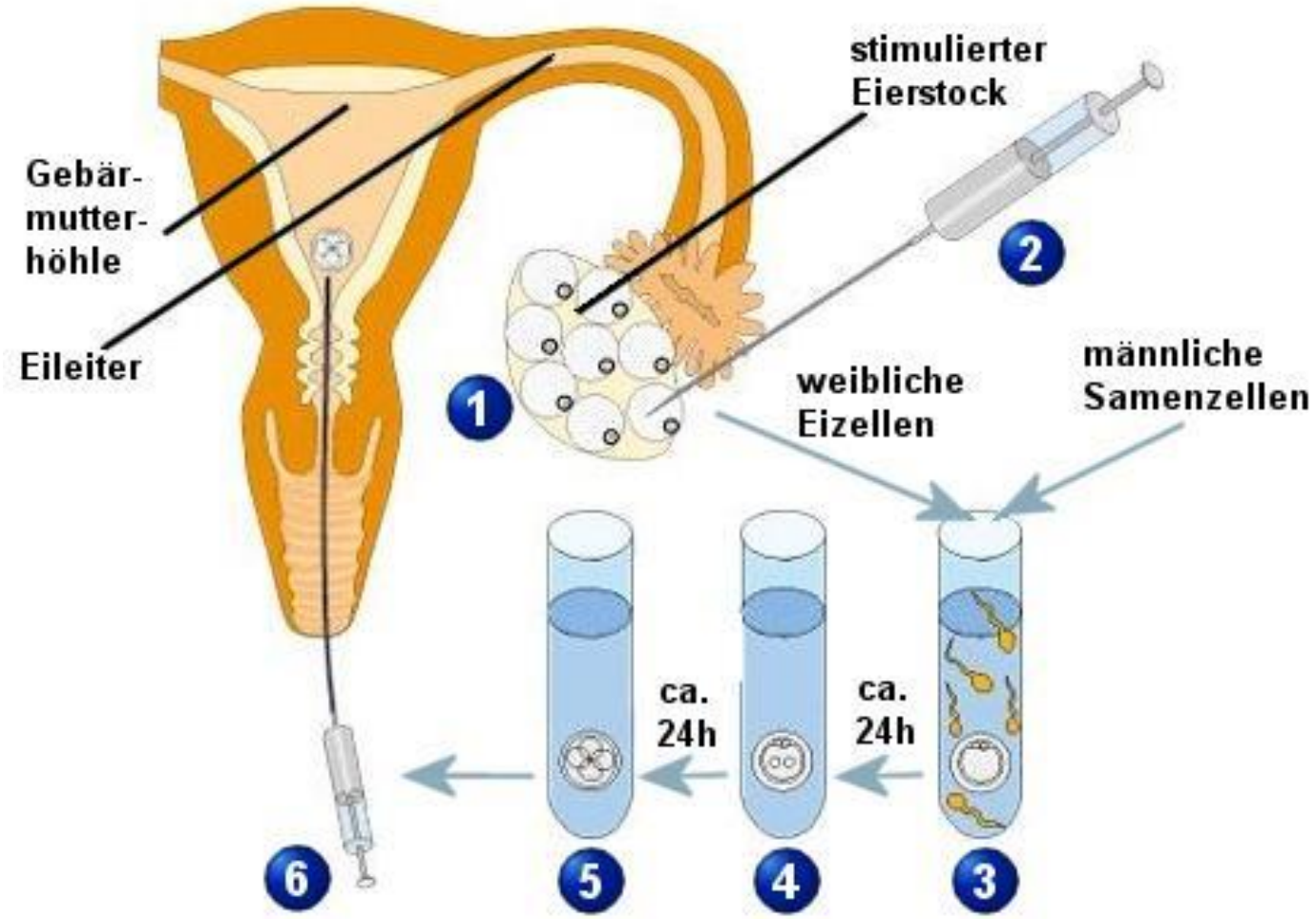


No elephants……

Problem 1: breeding elephants

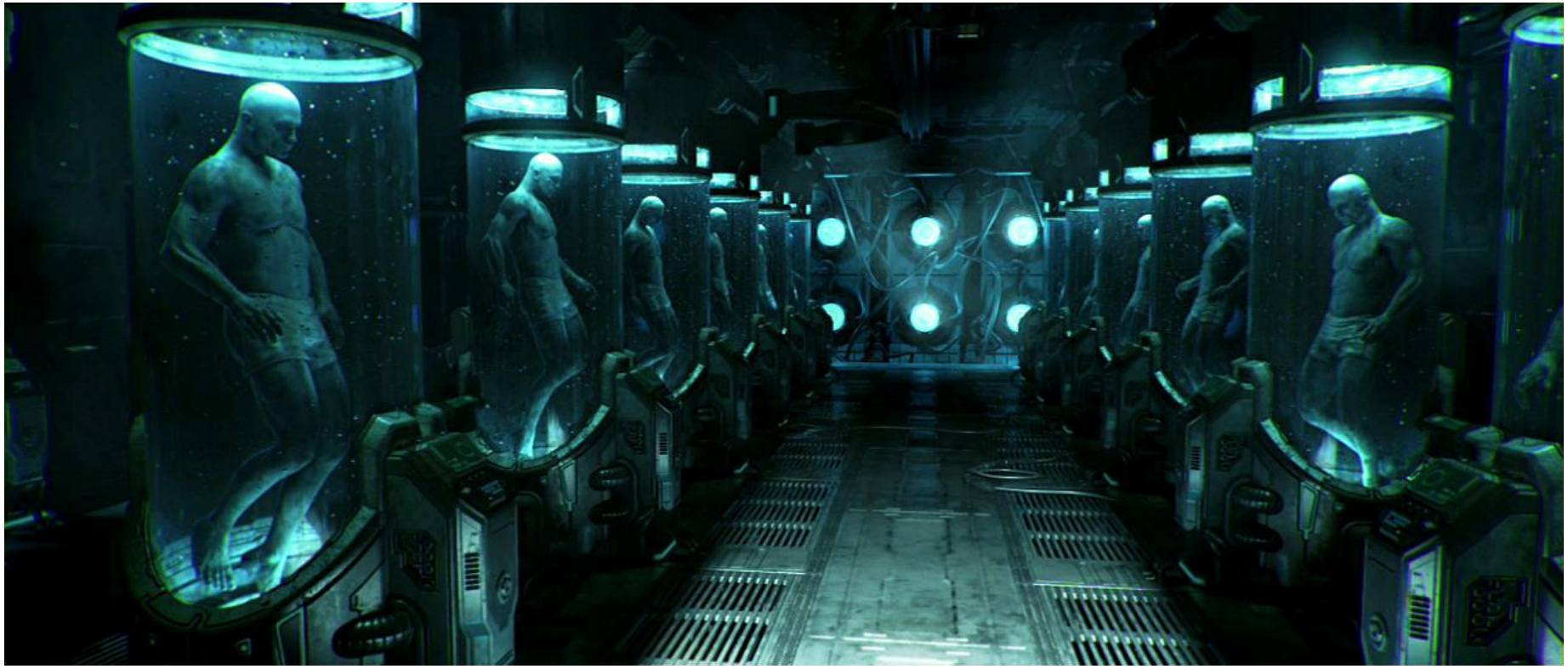


Problem 2: reproductive biology



Problem 3: generation time

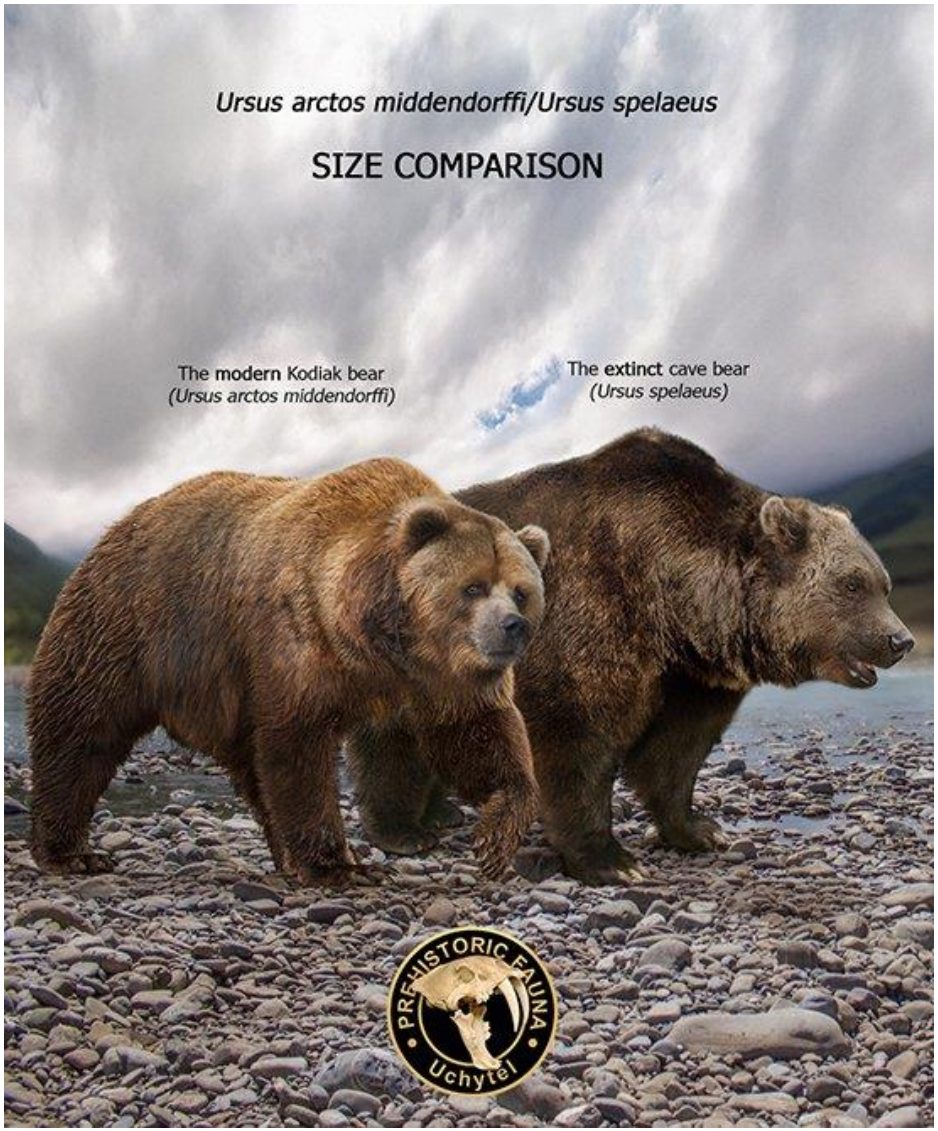
Artificial womb?



Better candidates: lava mouse

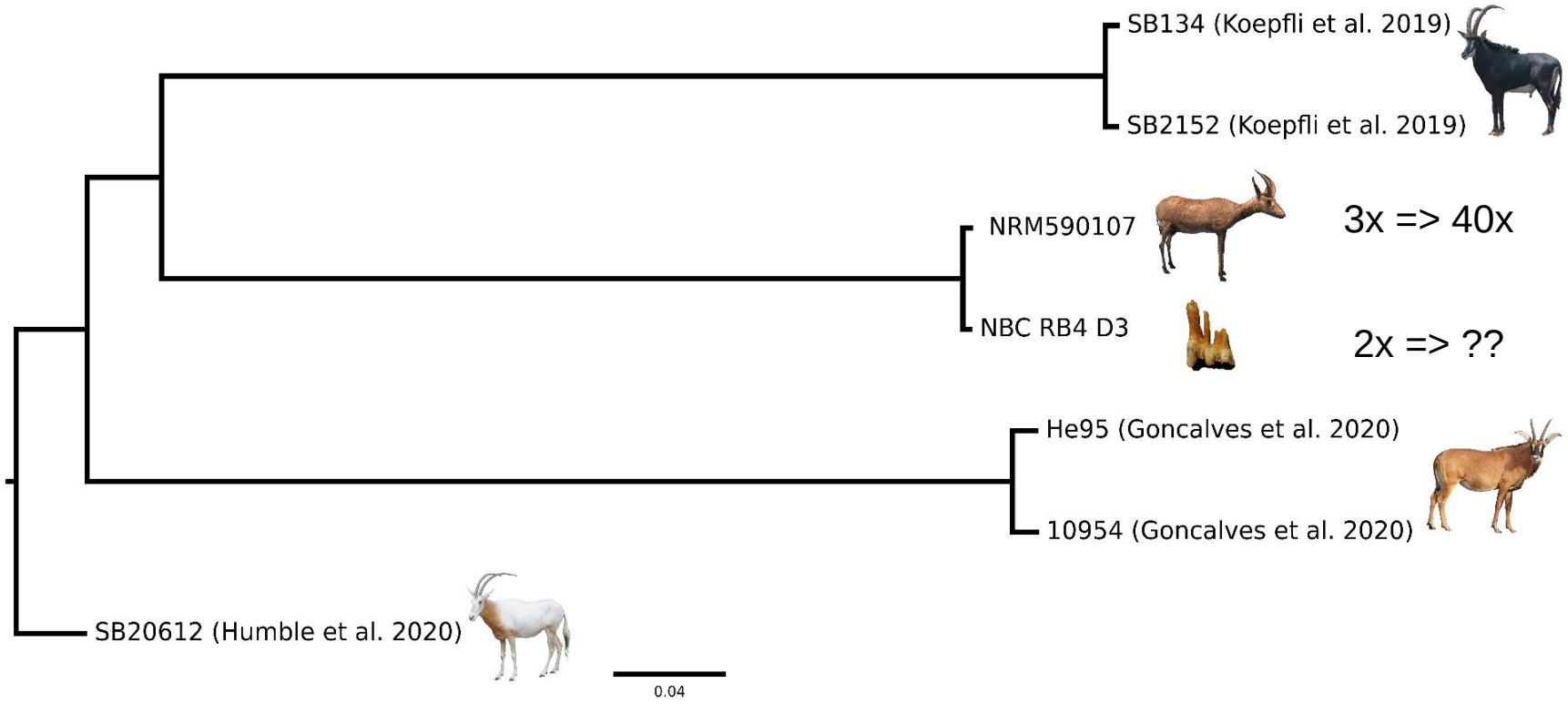
Better candidates: cave bear

In theory……

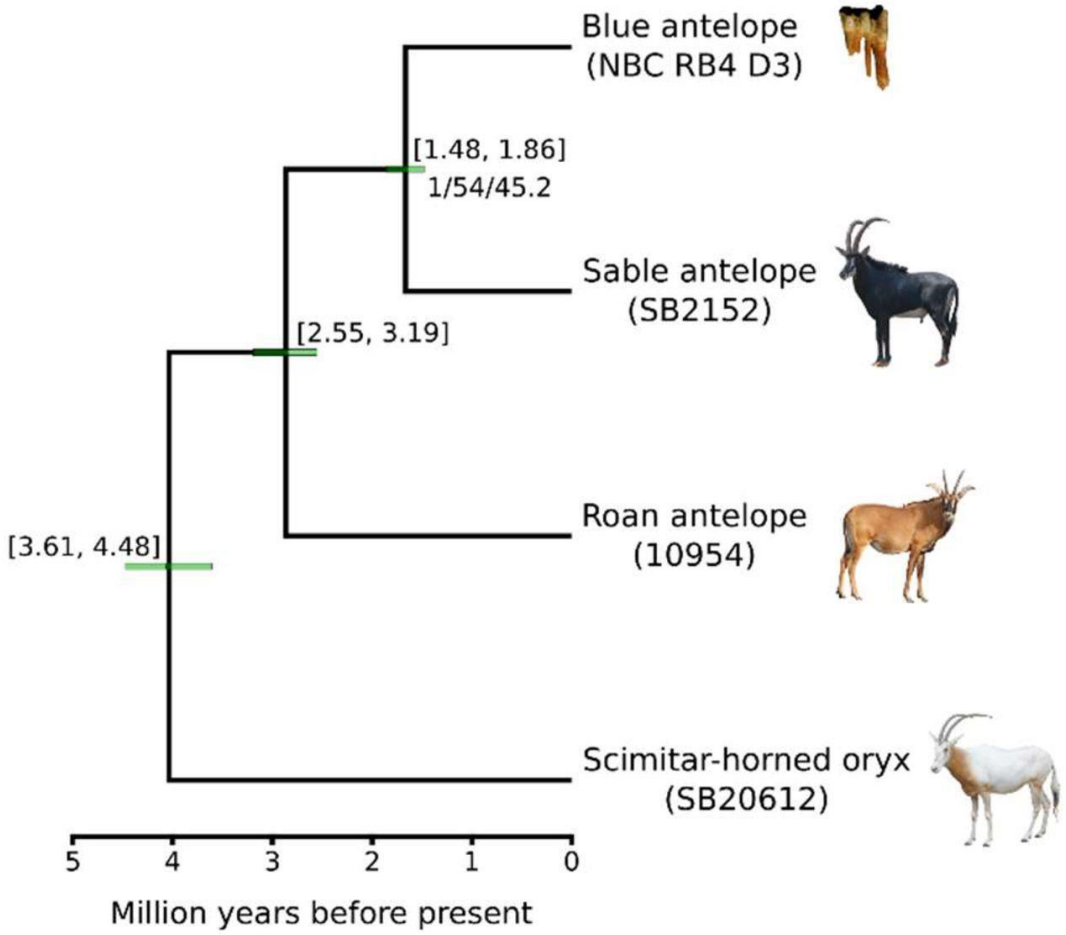

Ursus arctos middendorffi/Ursus spelaeus
SIZE COMPARISON
The modern Kodiak bear
(Ursus arctos middendorffi)
The extinct cave bear
(Ursus spelaeus)
PREHISTORIC FAUNA
Uchytel

In practice: South-African bluebuck ~1800 †

Advantages: 1) first genomes available

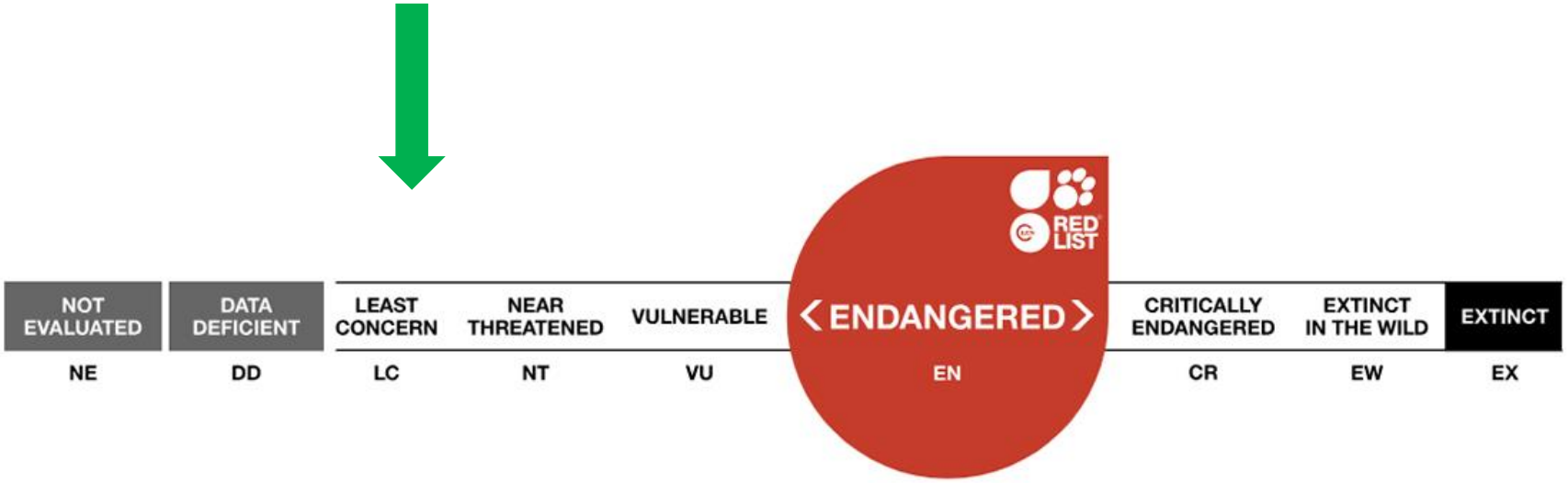


Advantages: 2) Very close extant relative



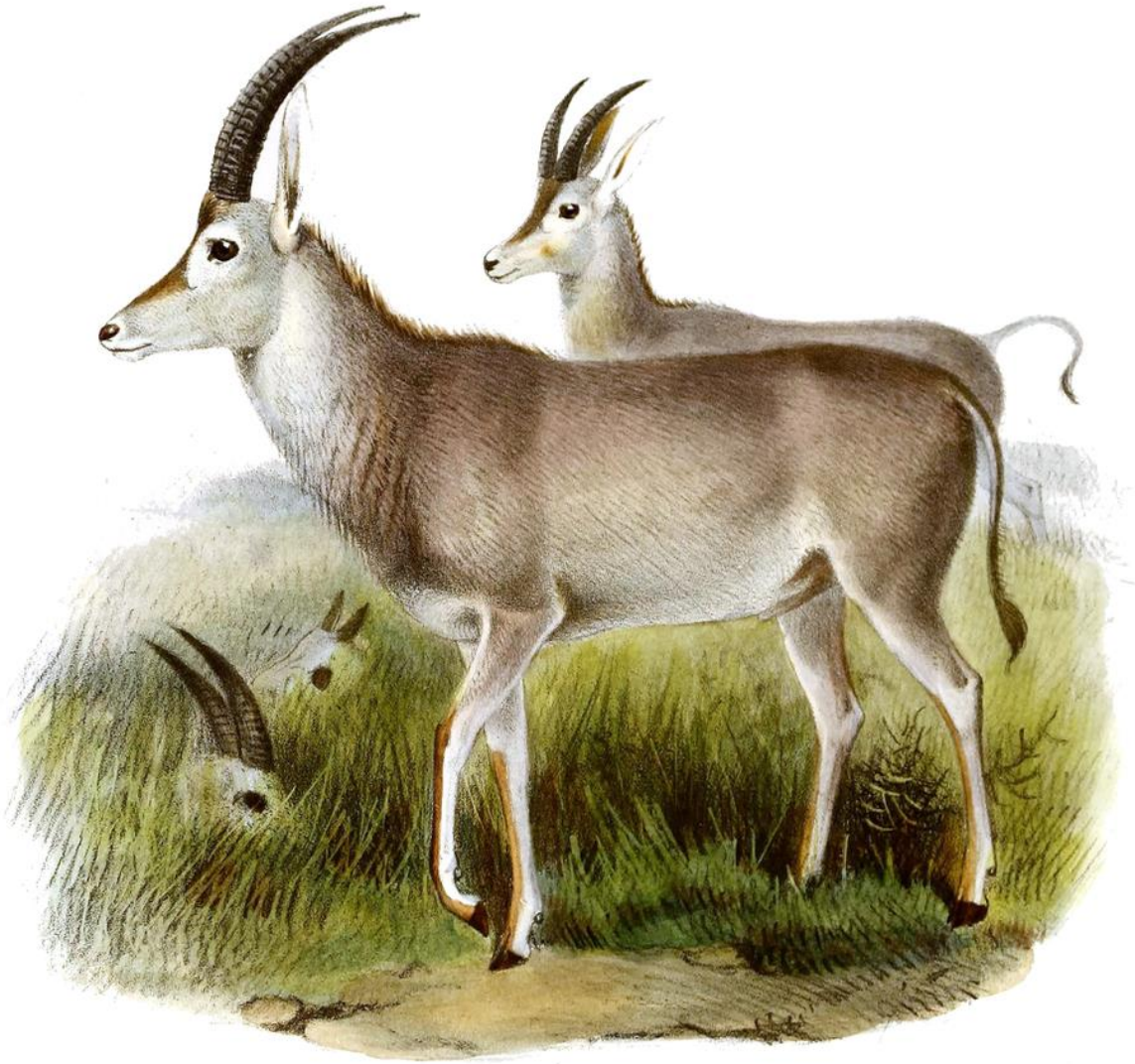
Advantages: 3) Also easy to breed

Advantages: 4) No conservation concern



Advantages: 5) It’s almost a cow (or a goat)

De-extinction? Maybe……




Ende