

INTRODUCTION TO PALAEOPROTEOMICS AND PALAEOGENOMICS

MUNI
SCI

EVA CHOCHOLOVÁ

LABORATORY OF BIOLOGICAL AND MOLECULAR ANTHROPOLOGY

DEPARTMENT OF EXPERIMENTAL BIOLOGY

SYLLABUS

1. Introduction to palaeoproteomics and palaeogenomics. History, main principles and differences, molecular preservation.
2. Methods in palaeogenomics.
3. Palaeogenomics of human populations.
4. Metagenomics.
5. Palaeogenomics of pathogens.
6. Sedimentary ancient DNA.
7. Methods in palaeoproteomics I
8. Methods in palaeoproteomics II
9. Collagens, keratins and other proteins used in taxonomy.
10. Palaeoproteomics of human populations. Sex typing, evolution, skeletal and mummified remains.
11. Metaproteomics. Diet, health and culture studied by proteomic approach.
12. Dental calculus as a complex material in biomolecular archaeology.
13. Opportunities, challenges, and ethics in palaeoproteomic and palaeogenomics research.



SYLLABUS

1. Introduction to palaeoproteomics and palaeogenomics. History, main principles and differences, molecular preservation.
2. Methods in palaeogenomics.
3. Palaeogenomics of human populations.
4. Metagenomics.
5. Palaeogenomics of pathogens.
6. Sedimentary ancient DNA.
7. Methods in palaeoproteomics I
8. Methods in palaeoproteomics II
9. Collagens, keratins and other proteins used in taxonomy.
10. Palaeoproteomics of human populations. Sex typing, evolution, skeletal and mummified remains.
11. Metaproteomics. Diet, health and culture studied by proteomic approach.
12. Dental calculus as a complex material in biomolecular archaeology.
13. Opportunities, challenges, and ethics in palaeoproteomic and palaeogenomics research.

LESSON PLAN

- Quiz
- Lecture
- Collection of important concepts/terms (1-3)
- Online resources for voluntary self-directed learning.

ASSESSMENT METHODS

- Written examination with possibility to take an oral exam.
(A: 100-93 points, B: 92-85 points, C: 84-77, D: 76-69, E: 68-60, F: 60-0)
- Students can obtain up to 20 bonus points during the semester for voluntary assignments (presentation, paper analysis, involvement in discussion).

Flash talk (5 points/each talk, up to 3 talks)

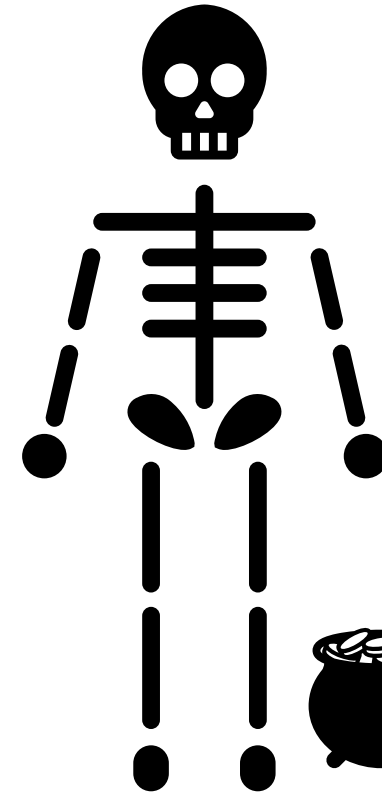
- 1) Choose a paper based on given topic
- 2) Read and analyse paper
- 3) Choose most important and interesting information
- 4) Prepare 1 slide with this information
- 5) Present in 3 minutes

Case study (5 points)

Last lecture – choose a case study (archaeological situation), prepare and share suggested solution and what analyses you would recommend.

PALAEOGENOMICS, PALAEOPROTEOMICS

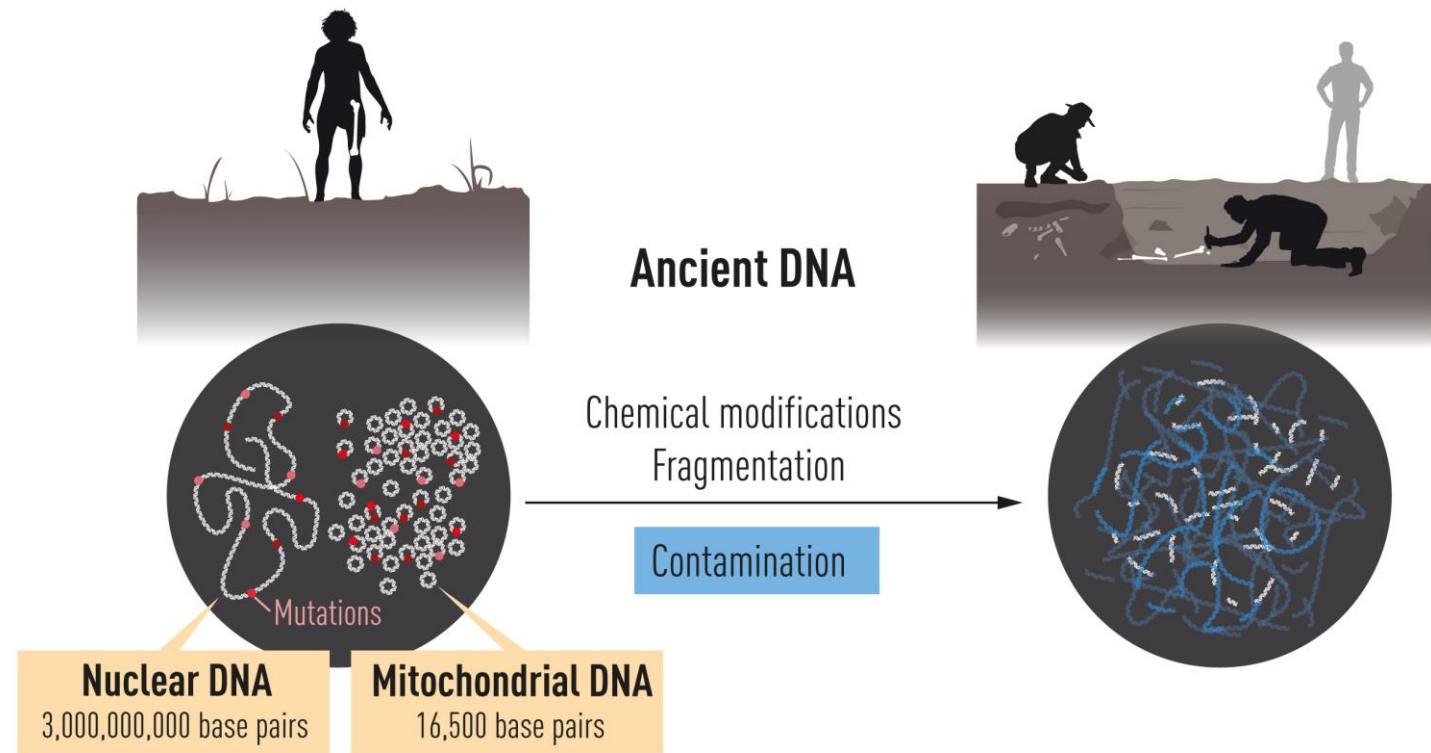
- Importance
 - Life of ancient populations
 - Migrations and contact
 - Evolutionary processes, adaptations
 - Current organisms and diseases
 - Extinct organisms
 - Epidemiology and public health implications
 - Agriculture
 - ...



NOBEL PRIZE – SVANTE PÄÄBO



The Nobel Prize in Physiology or Medicine 2022 was awarded to Svante Pääbo "for his discoveries concerning the genomes of extinct hominins and human evolution"



NOBEL PRIZE – SVANTE PÄÄBO

Thanks to Svante Pääbo's discoveries, we now understand that archaic gene sequences from our extinct relatives influence the physiology of present-day humans. One such example is the Denisovan version of the gene EPAS1, which confers an advantage for survival at high altitude and is common among present-day Tibetans. Other examples are Neanderthal genes that affect our immune response to different types of infections.

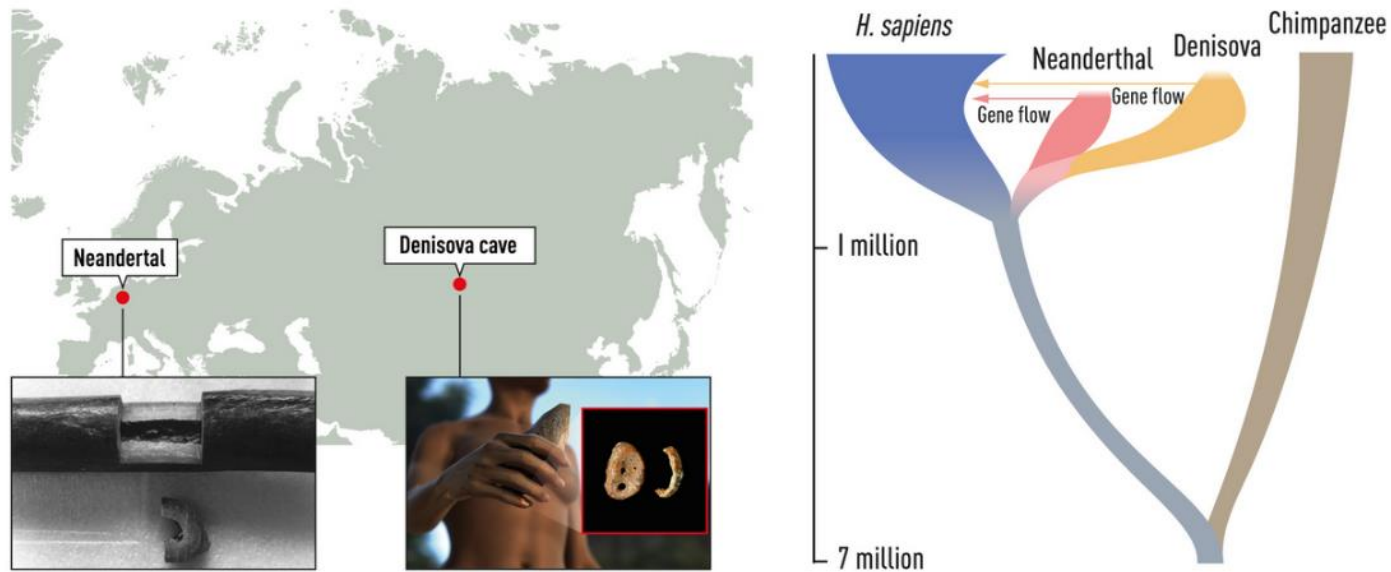


Figure 2. A. Pääbo extracted DNA from bone specimens from extinct hominins. He first obtained a bone fragment from Neanderthal in Germany, the site that gave name to the Neanderthals. Later, he used a finger bone from the Denisova Cave in southern Siberia, the site that gave name to Denisovans. B. Phylogenetic tree showing the evolution and relationship between *Homo sapiens* and the extinct hominins. The phylogenetic tree also illustrates the gene flows discovered by Pääbo.

NOBEL PRIZE – SVANTE PÄÄBO

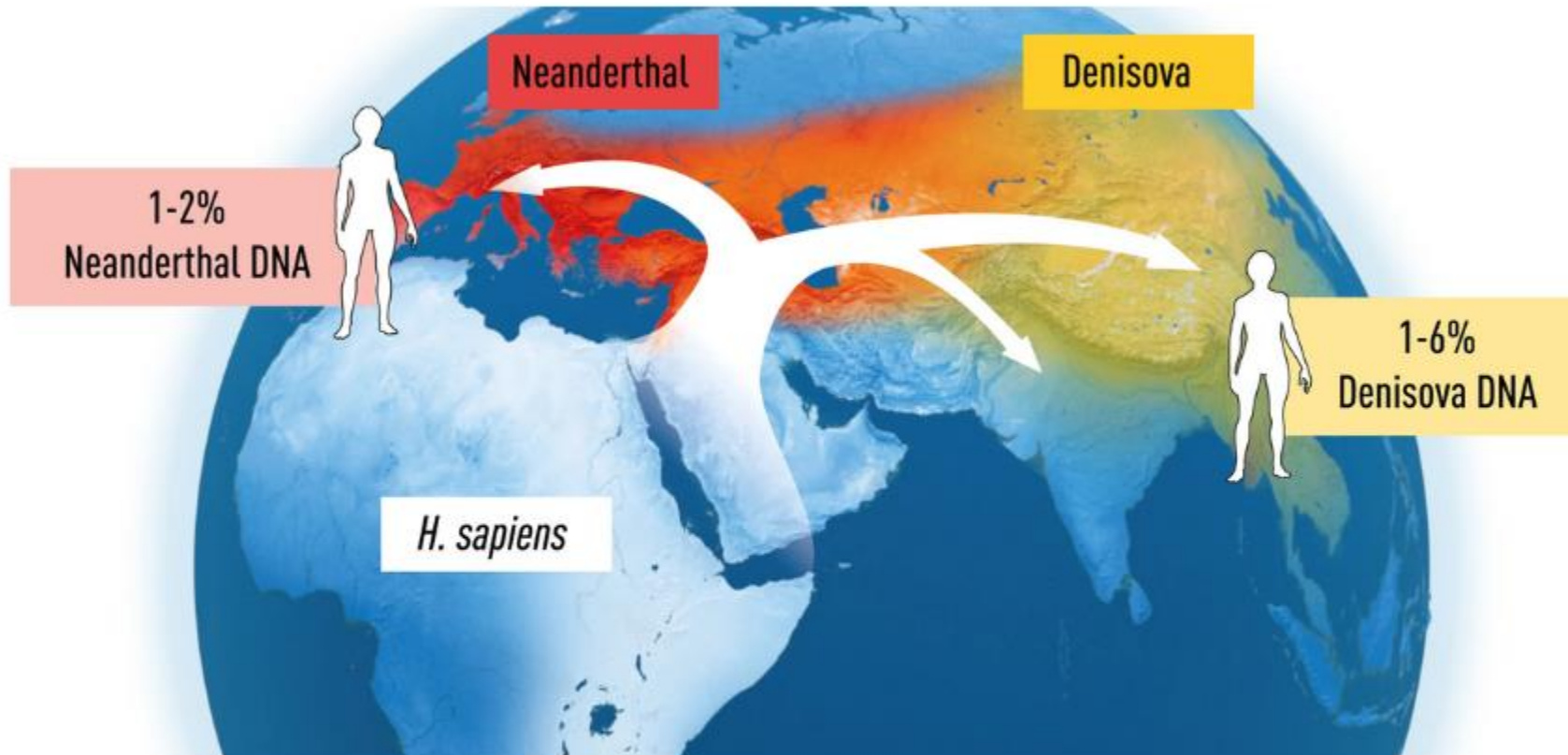
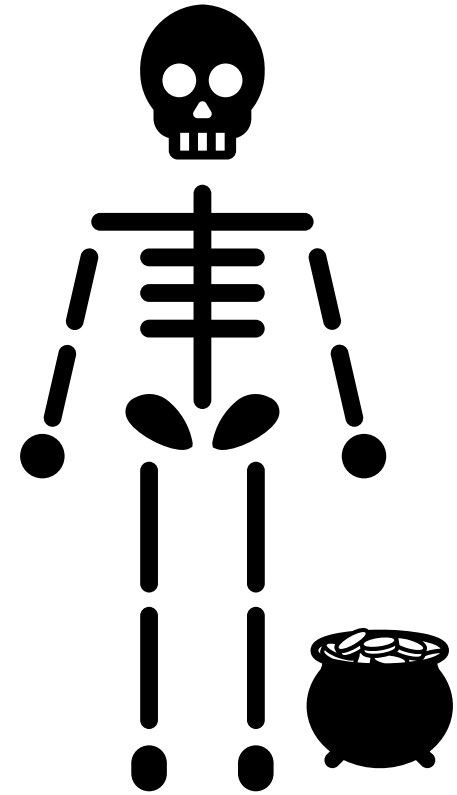
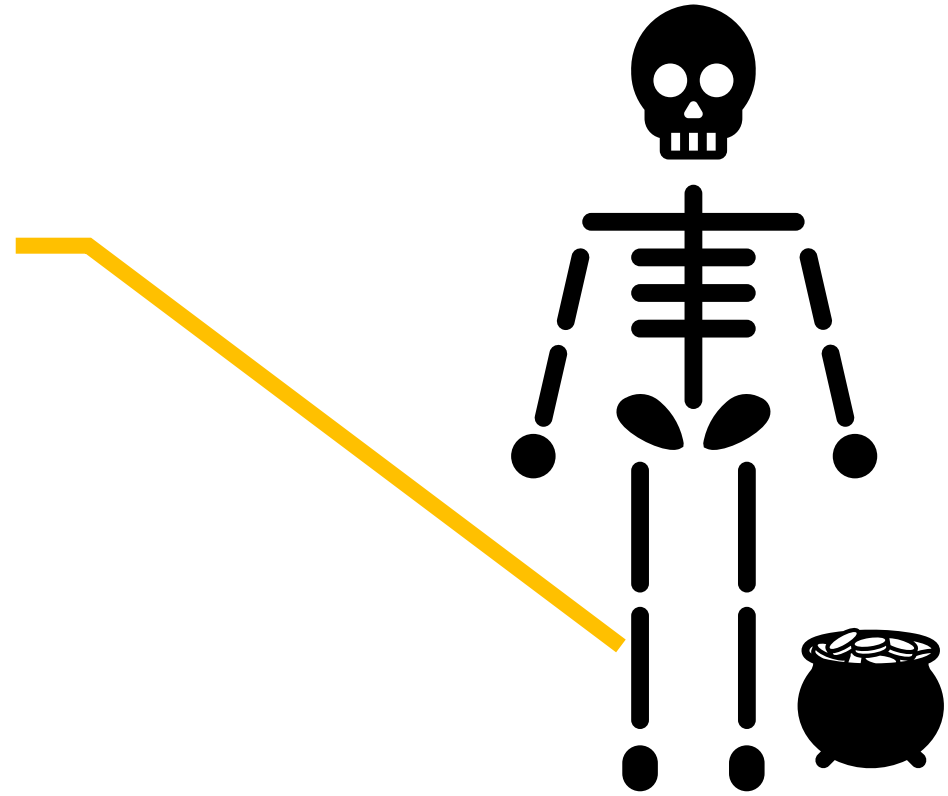
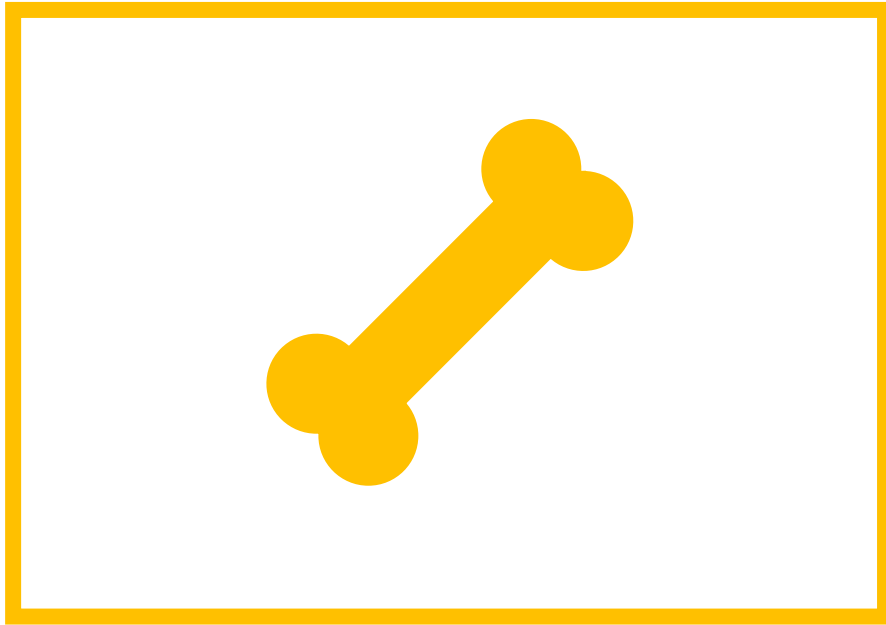


Figure 3. Pääbo's discoveries have provided important information on how the world was populated at the time when Homo sapiens migrated out of Africa and spread to the rest of the world. Neanderthals lived in the west and Denisovans in the east on the Eurasian continent. Interbreeding occurred when Homo sapiens spread across the continent, leaving traces that remain in our DNA.

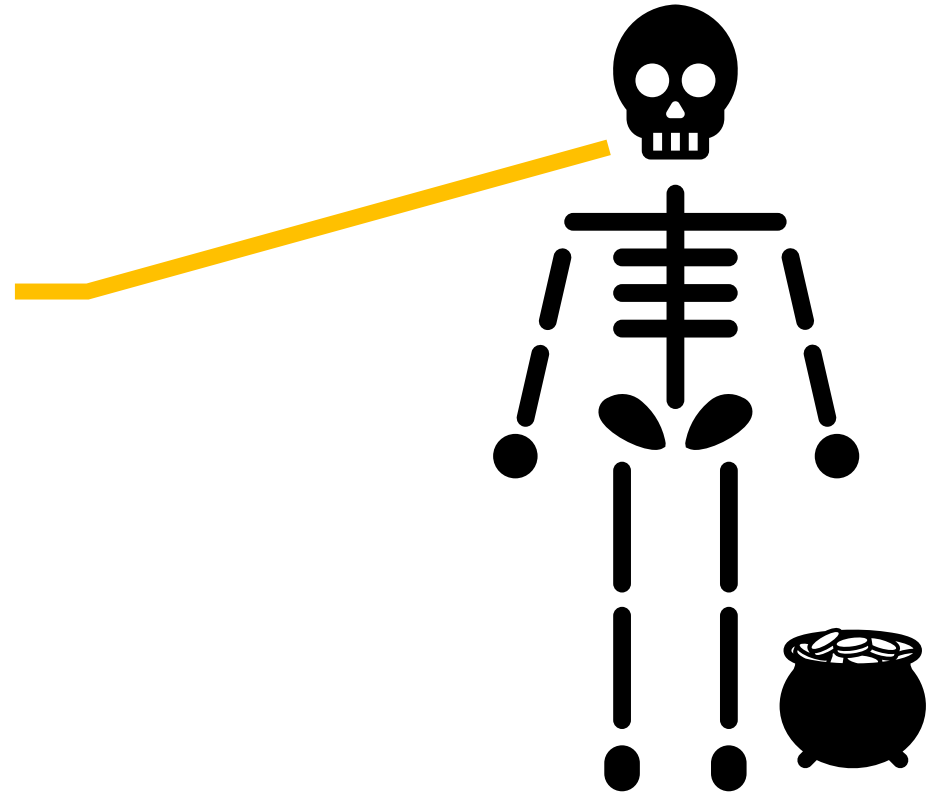
WHAT TO STUDY



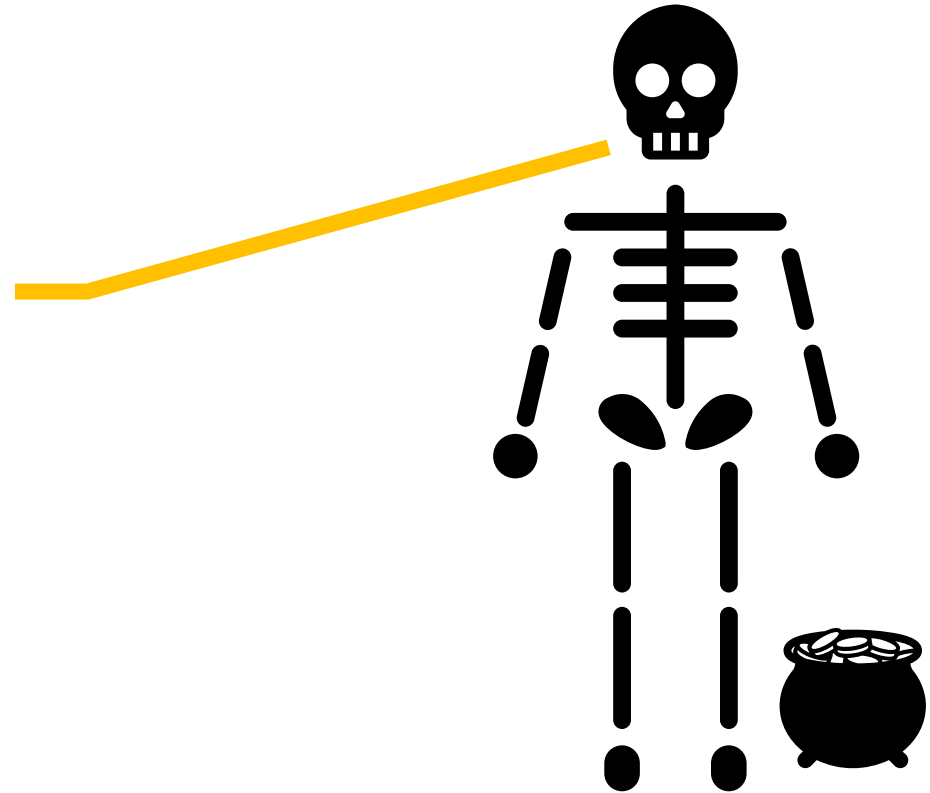
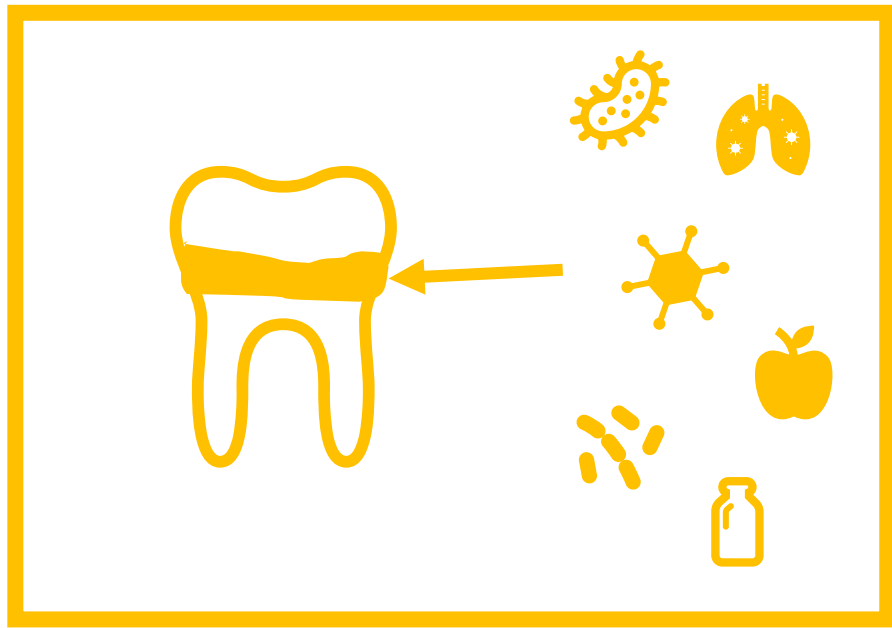
WHAT TO STUDY



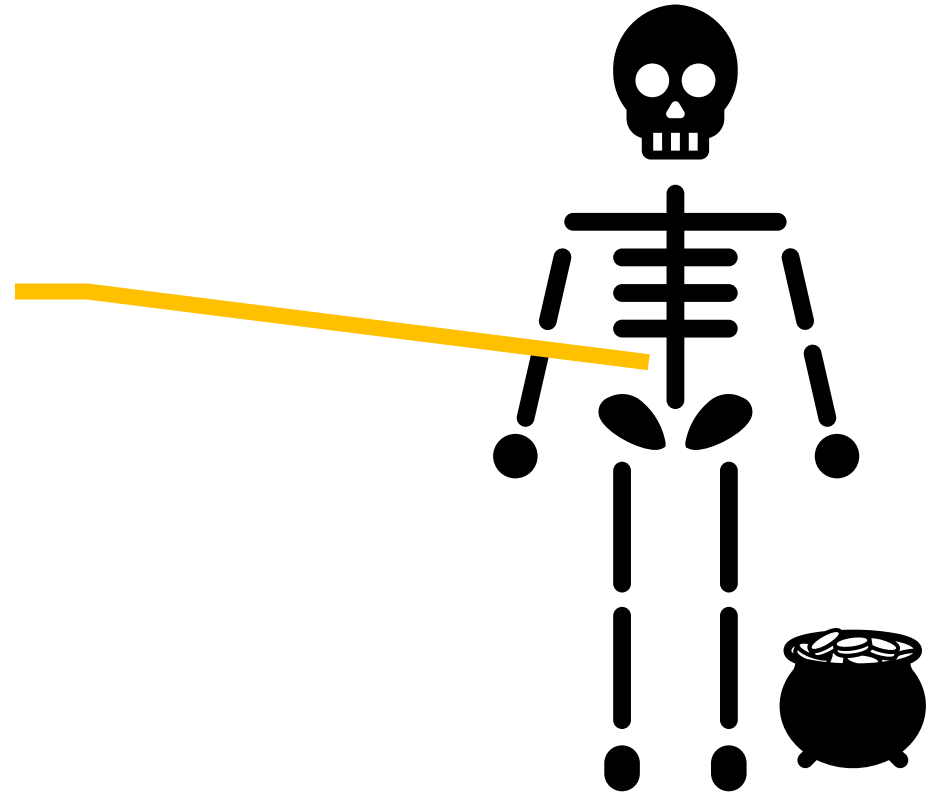
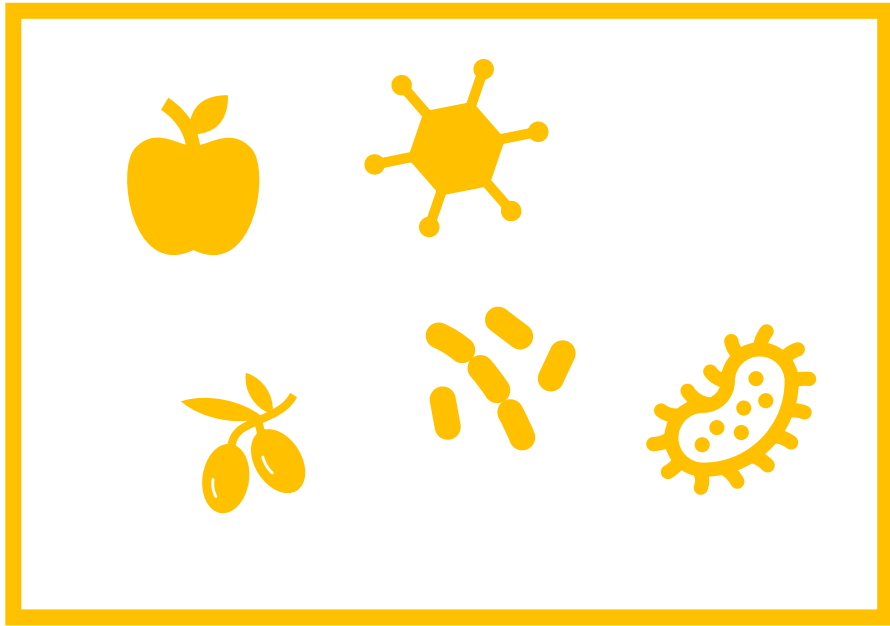
WHAT TO STUDY



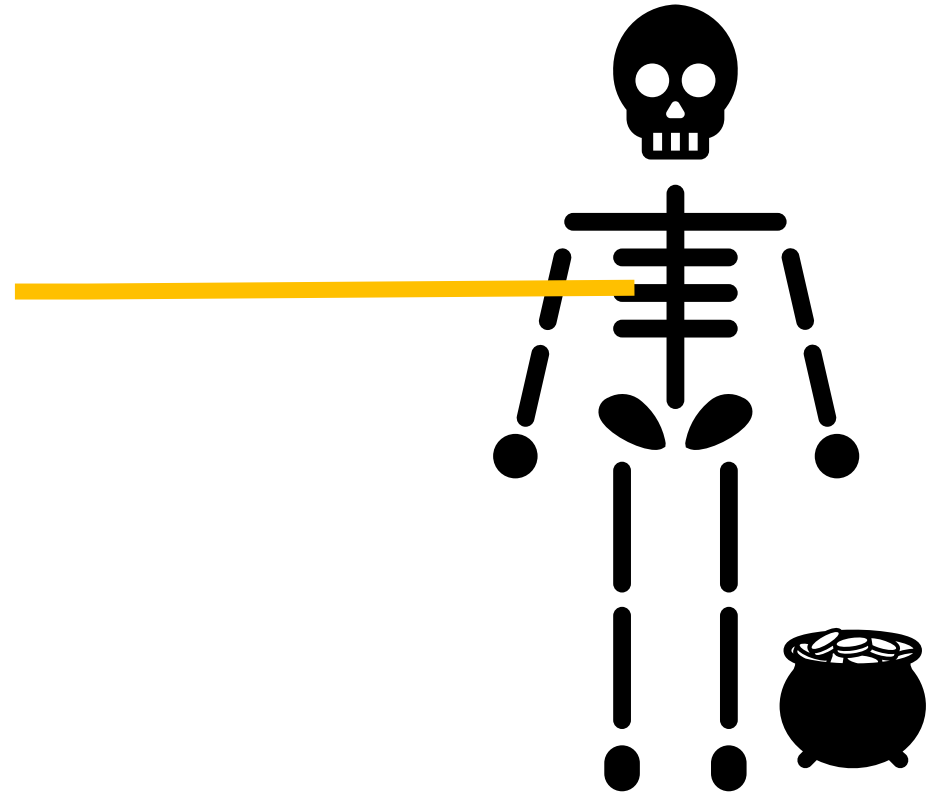
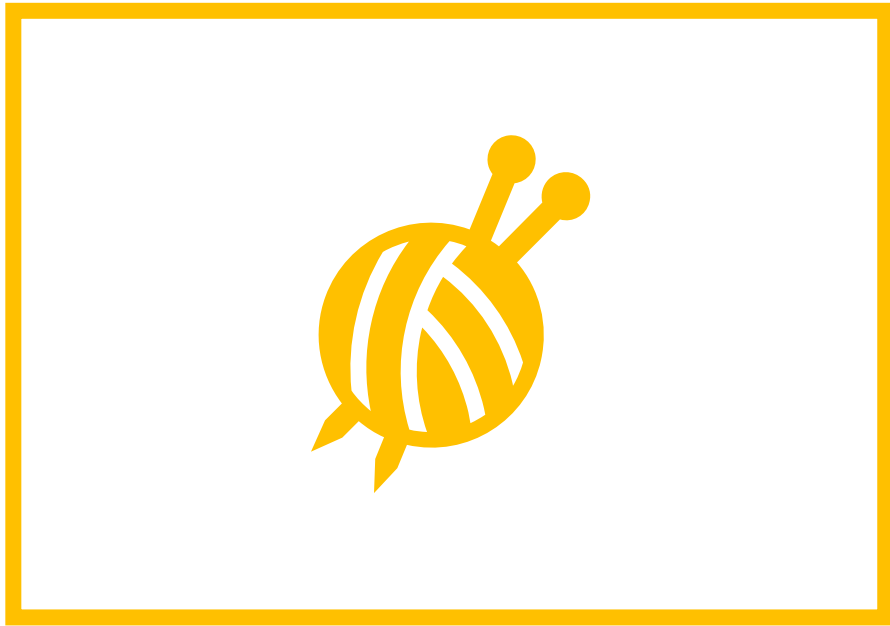
WHAT TO STUDY



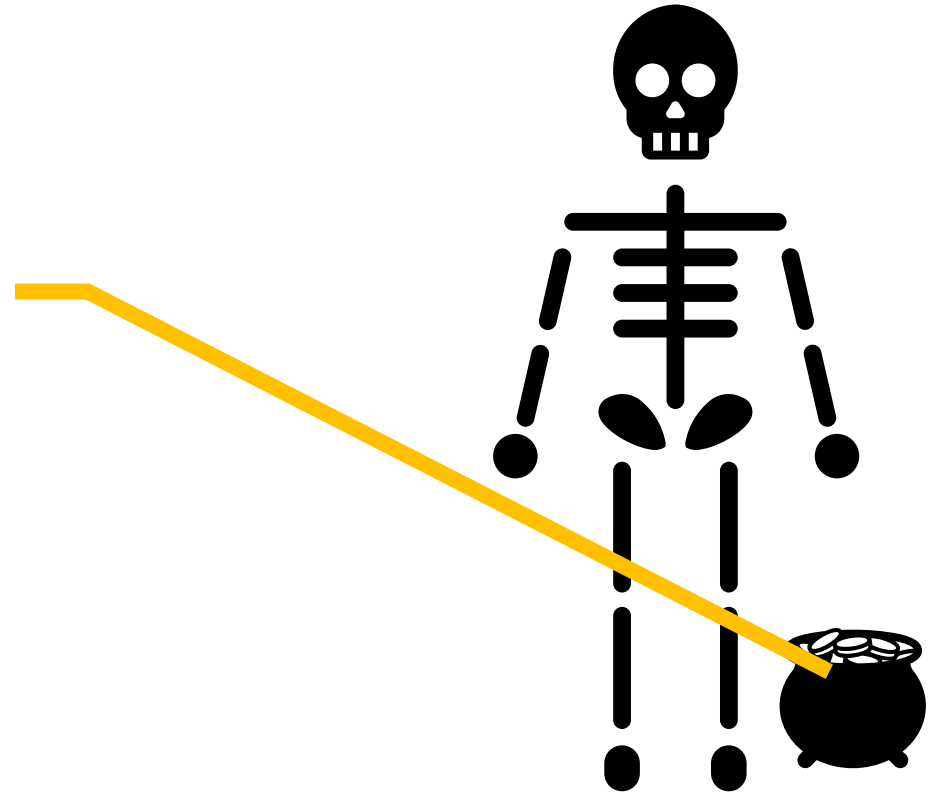
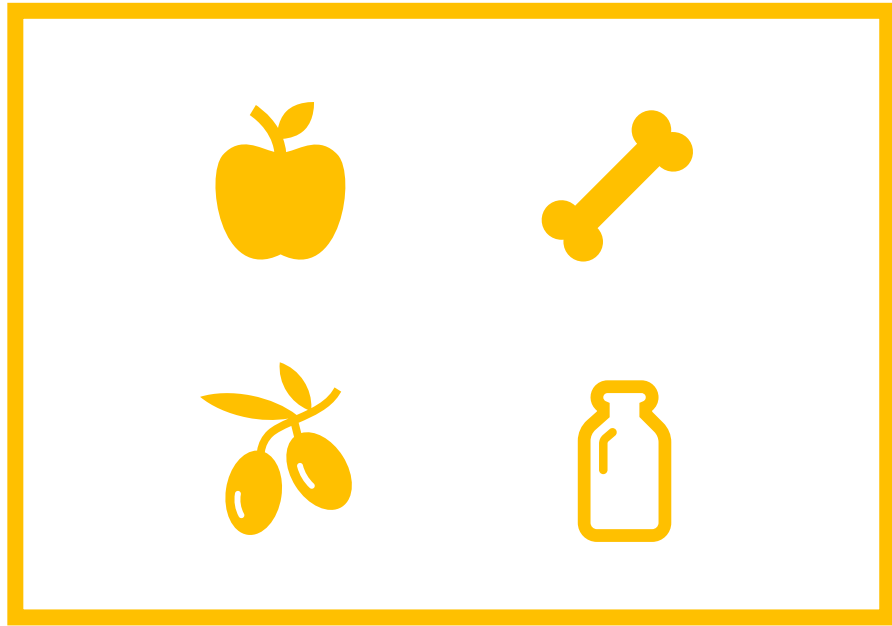
WHAT TO STUDY



WHAT TO STUDY



WHAT TO STUDY



BRIEF HISTORY OF PALAEOGENOMICS

- **Genetics**

Mitochondrial DNA,
identification



- **Genomics**

Whole genomes

- **Quagga (Higuchi et al., 1984), Egyptian mummies (Pääbo, 1985)**

Last died in Amsterdam Zoo, 12 August 1883



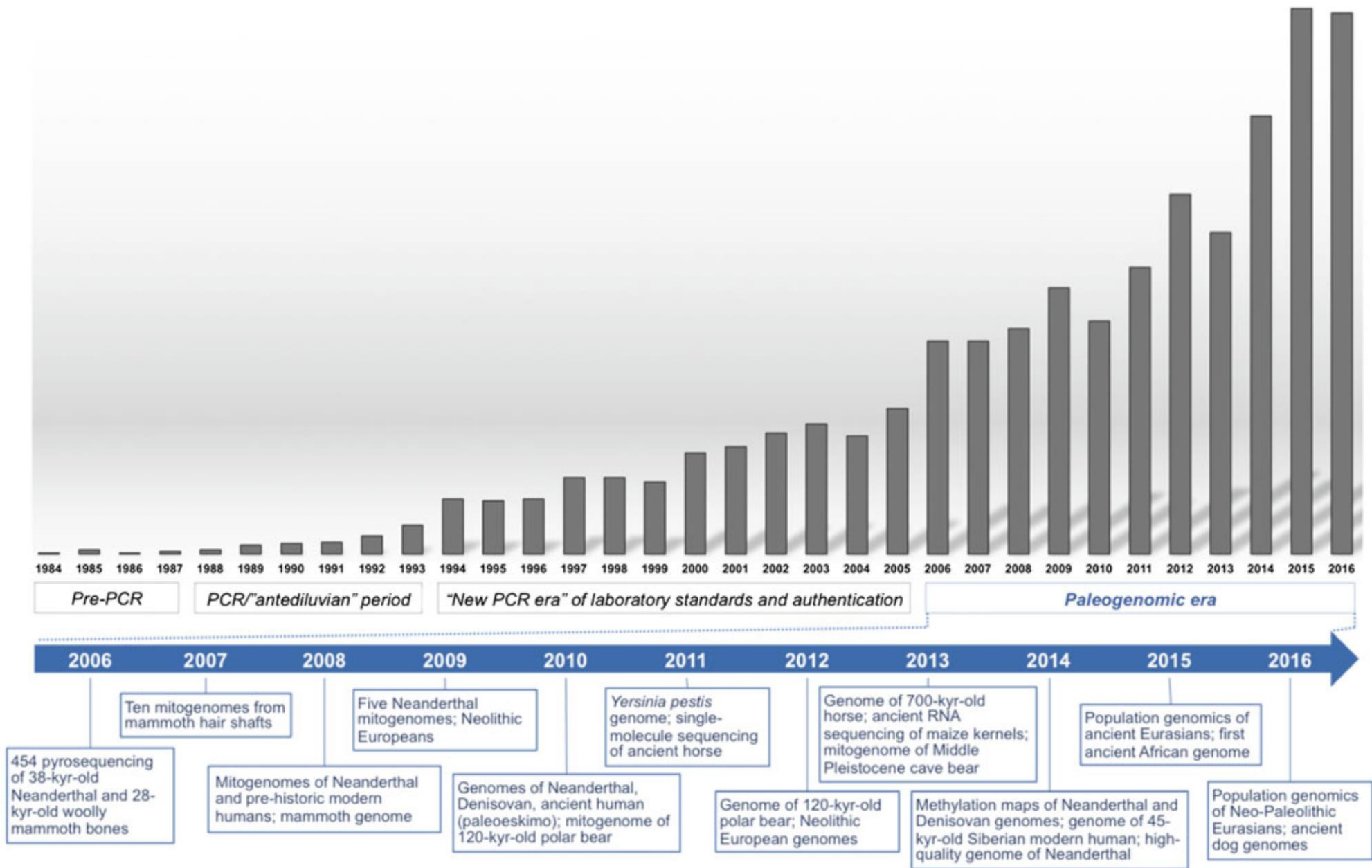
1987

Founder population, selected in Etosha in 1987 **Ricky**



2022

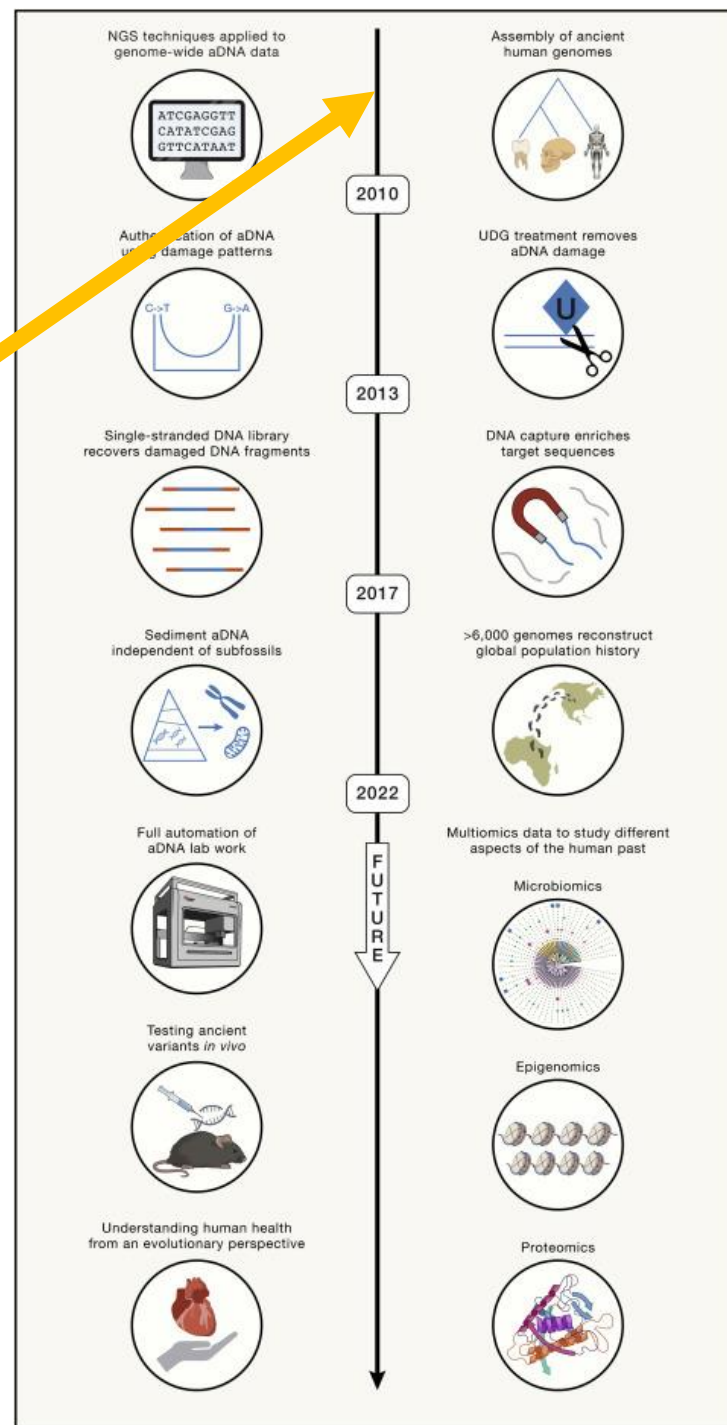
Rain Current population, Nuwejaars Wetland 2022





Neanderthal Genome Project - 2006

<https://www.mpg.de/13894984/neanderthal-genome-project>



BRIEF HISTORY OF PALAEOGENOMICS

● Non-human animal paleogenomes ● Hominin paleogenomes ● Sedimentary ancient DNA records

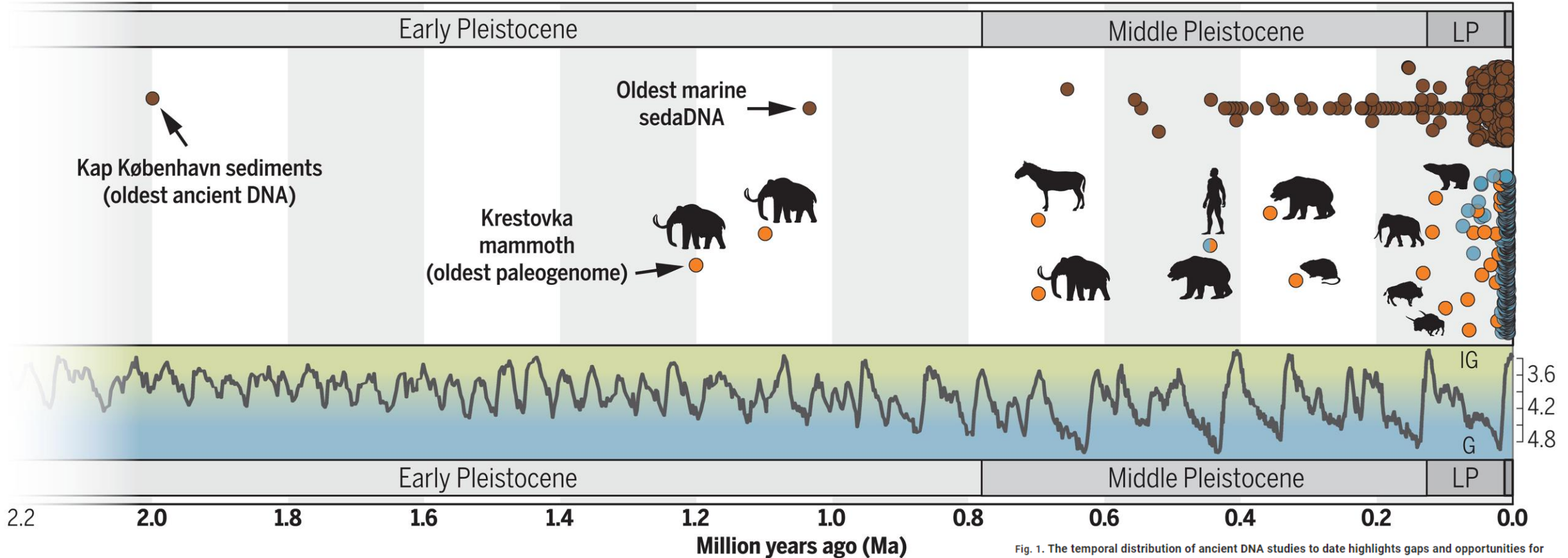
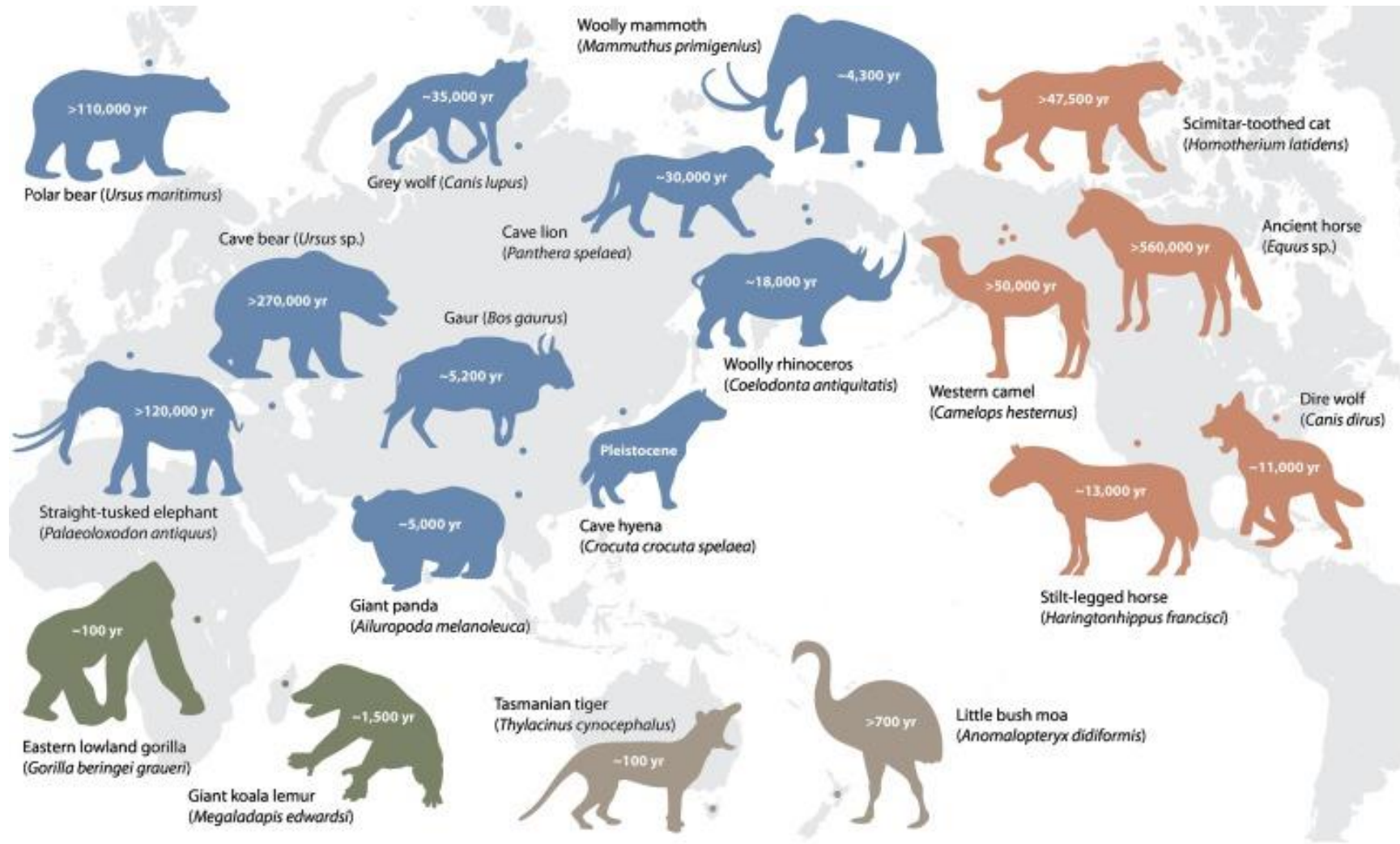


Fig. 1. The temporal distribution of ancient DNA studies to date highlights gaps and opportunities for deep-time paleogenomics and sedimentary ancient DNA.

Most ancient DNA studies fall within the last 50 ka and the most recent glacial cycle. The climate curve is based on benthic $\delta^{18}\text{O}$ -oxygen measurements [per mil; LR04 stack from (42)]. Sedimentary ancient DNA data are from the AncientMetagenomeDir (v23.06.0) (57) and (58), with metabarcoding records older than one million years excluded. Paleogenomic data are available from (59). Paleogenomes older than 100 ka are annotated with a silhouette of the study taxon, with the deep-time paleogenomes including a 130-ka steppe bison (36); 330-ka collared lemming (40); 360-ka cave bear (9); 430-ka cave bear and hominin (35, 60); 700-ka horse (8); and 700-ka, 1.1-Ma, and 1.2-Ma mammoths (10). LP, Late Pleistocene; IG, Interglacial; G, Glacial. Silhouettes are from PhyloPic <https://beta.phylopic.org/> and are in the public domain with credits to Zimices (mammoth, two bison) and Robert Bruce Horsfall (horse)

BRIEF HISTORY OF PALAEOGENOMICS



Trends in Ecology & Evolution

BRIEF HISTORY OF PALAEOGENOMICS

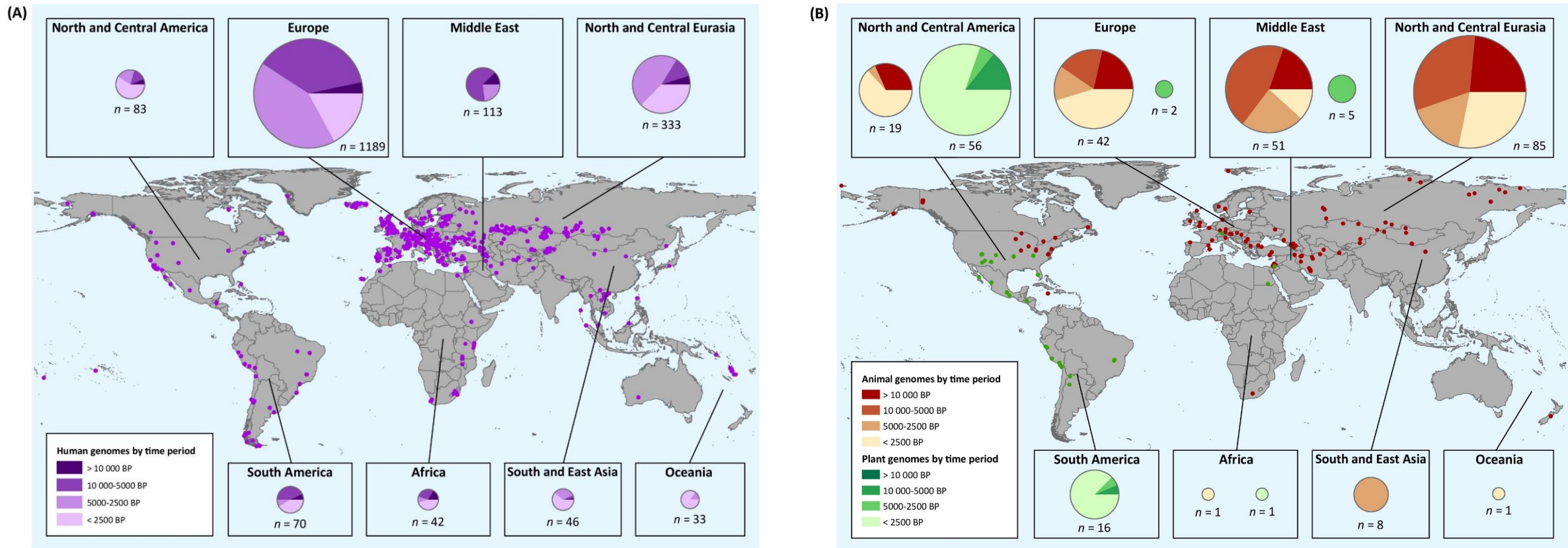
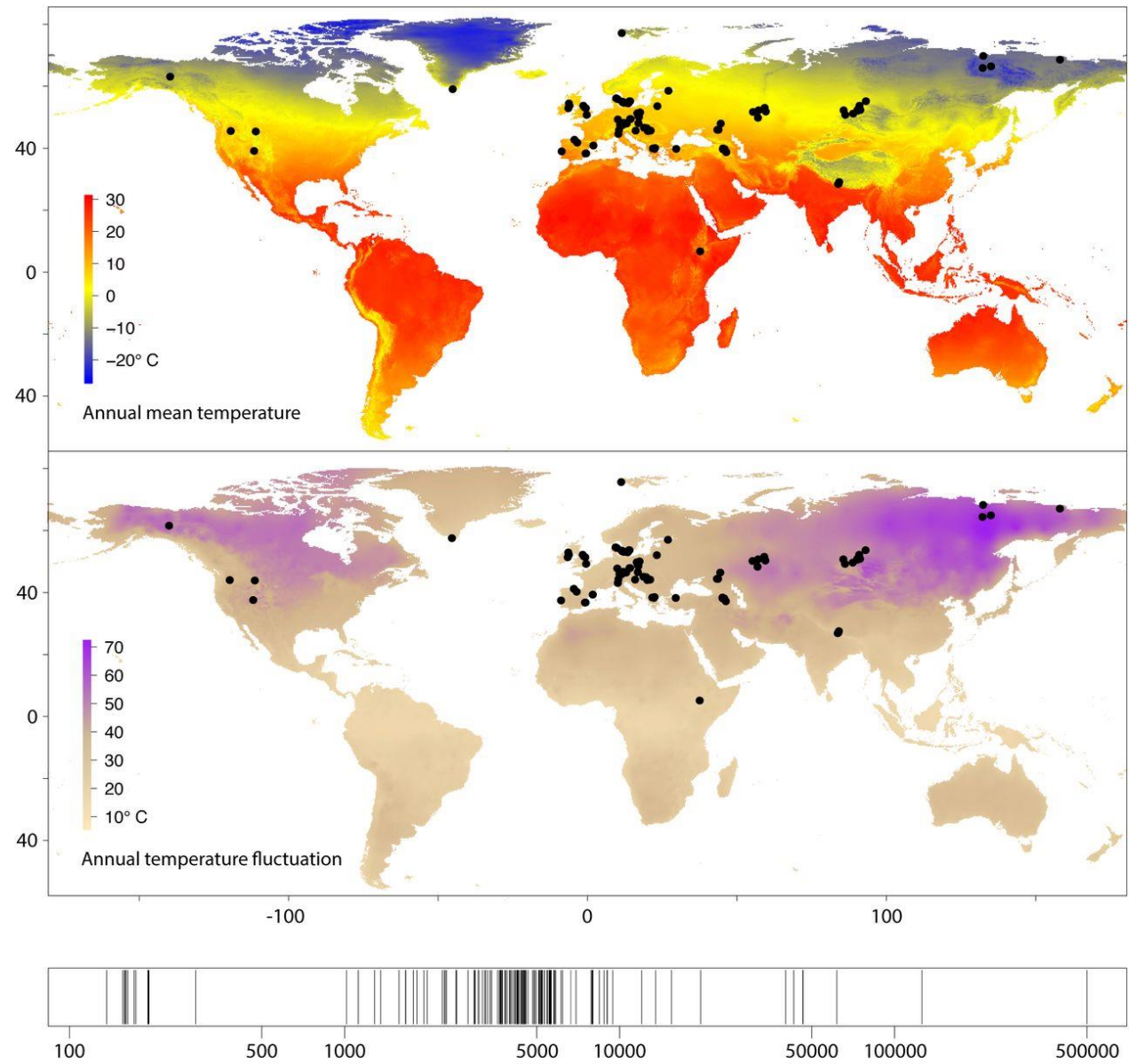


Figure 1 Geographic and Temporal Distribution of Ancient Genomes Published through December 15, 2018. Marciniak and Perry [3] previously reported a map of published human sequences. Here we generate an enhanced version of this figure, that reports both published genomic data from (A) human individuals (points in purple, updated from Marciniak and Perry, $n = 1909$), and (B) non-human animal (red, $n = 207$) and plant (green, $n = 80$) species. For the human samples, we restrict to individuals with $>0.025X$ coverage on a genome-wide set of informative single nucleotide polymorphism positions. For animals and plants, we restrict to samples in which multiple autosomal loci have been characterized. Samples with only mitochondrial or chloroplast genomes are not included. Data can be found in the online supplementary data. Maps drawn by Miriam Rothenberg in ArcGIS.

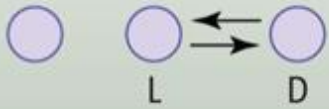


Sample age distribution (years)

Locations of 185 samples (n=94 unique sites) used in paleogenomic meta-analysis, global variation in mean temperature and temperature fluctuation, and timeline of sample ages. Note the absence of sites with annual mean temperature >20°C, reflecting known preservation bias toward cooler climates (22).

BRIEF HISTORY OF PALAEOPROTEOMICS

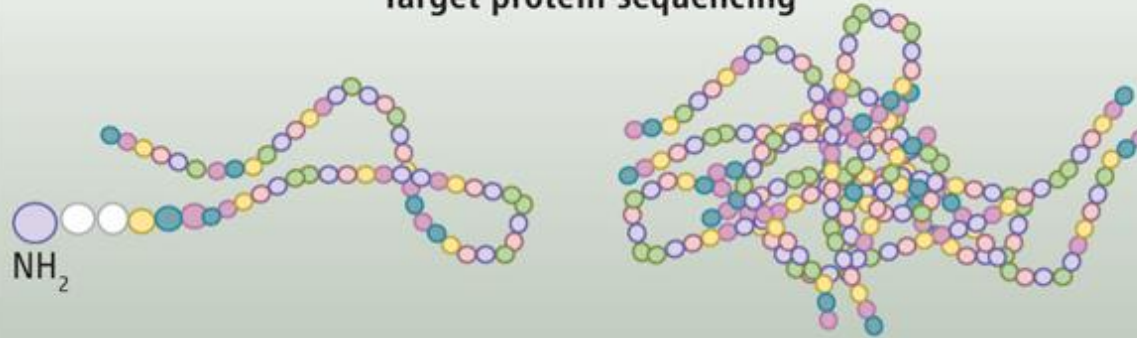
Amino acids



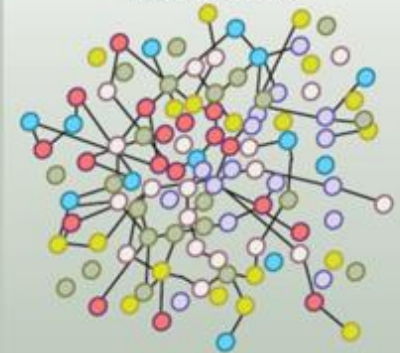
Epitopes



Target protein sequencing



Proteomics



1954

Amino acid profile



1967

Chirality



1974

Antibody detection



1980

Edman sequencing



2000

MALDI post source decay



2007

Collagen sequence?



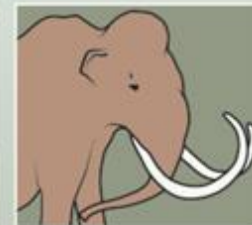
2010

Species identification (sheep vs goat)



2012

Mammoth bone



2014

Infection and immunity in dental calculus



Mass spectrometry

Identification

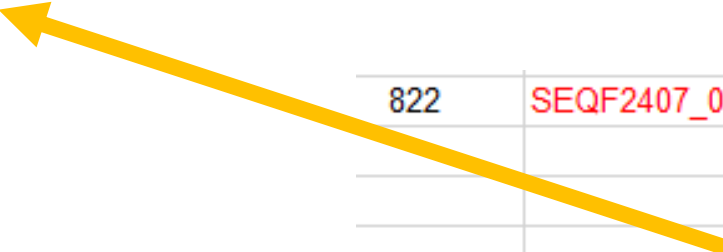
Sequencing

Expression

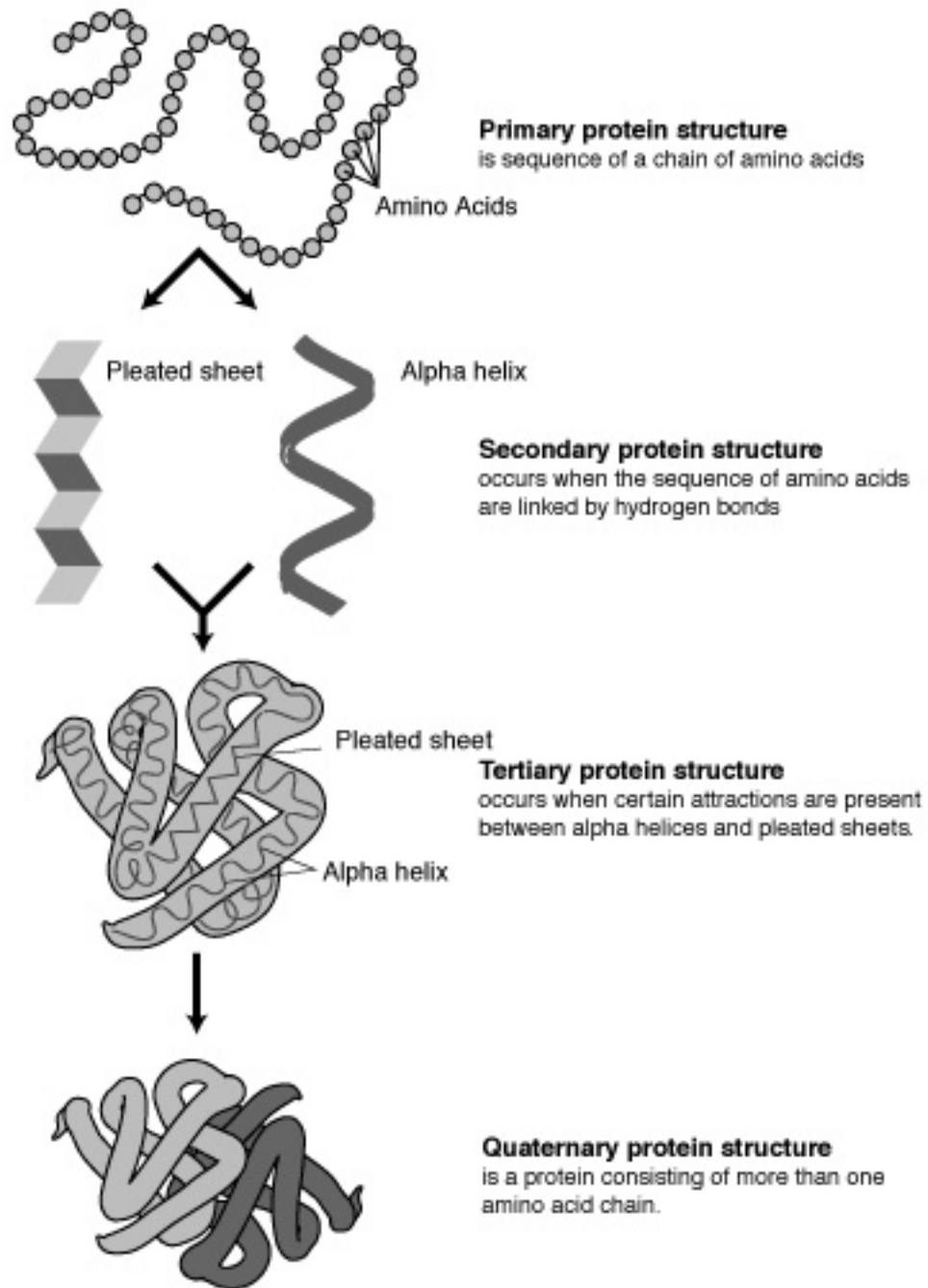
Quantification

PROTEINS

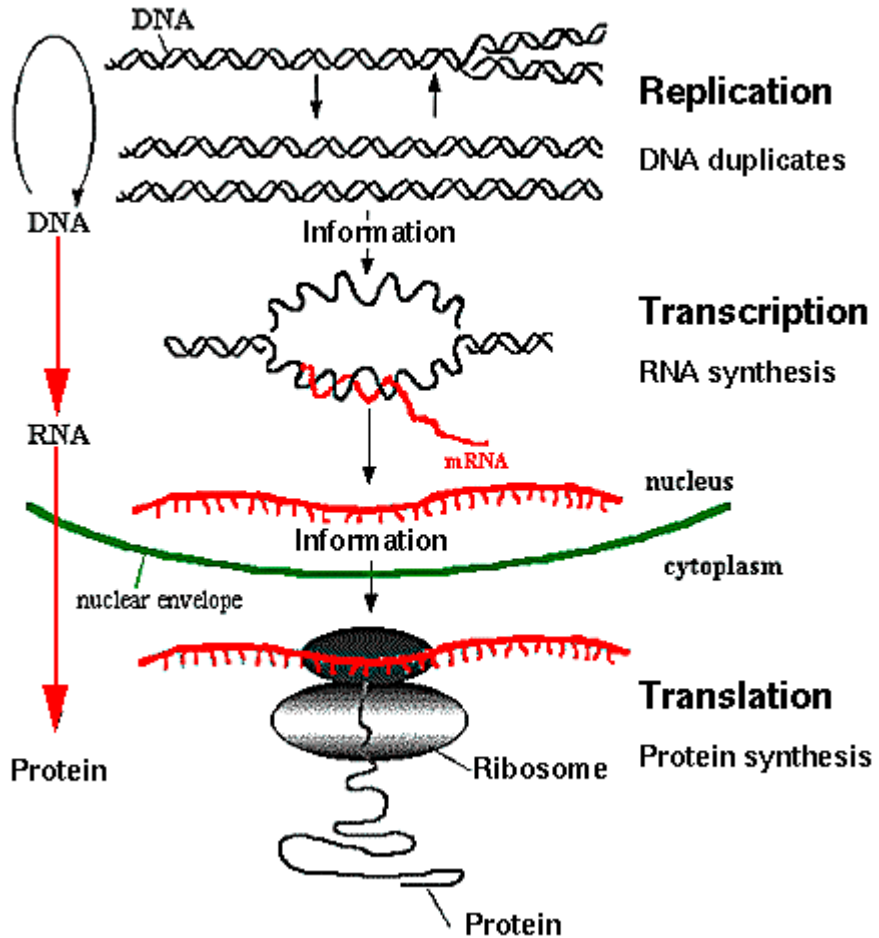
- Macromolecules, polypeptides
- Peptide – less than 30 AA
- Shape crucial for function



822	SEQF2407_00	Pyruvate kinase_Turicella otitidis
		Sequence
		SIGVLADLQGPK



CENTRAL DOGMA OF MOLECULAR BIOLOGY



Introduction of mutations
- diversity



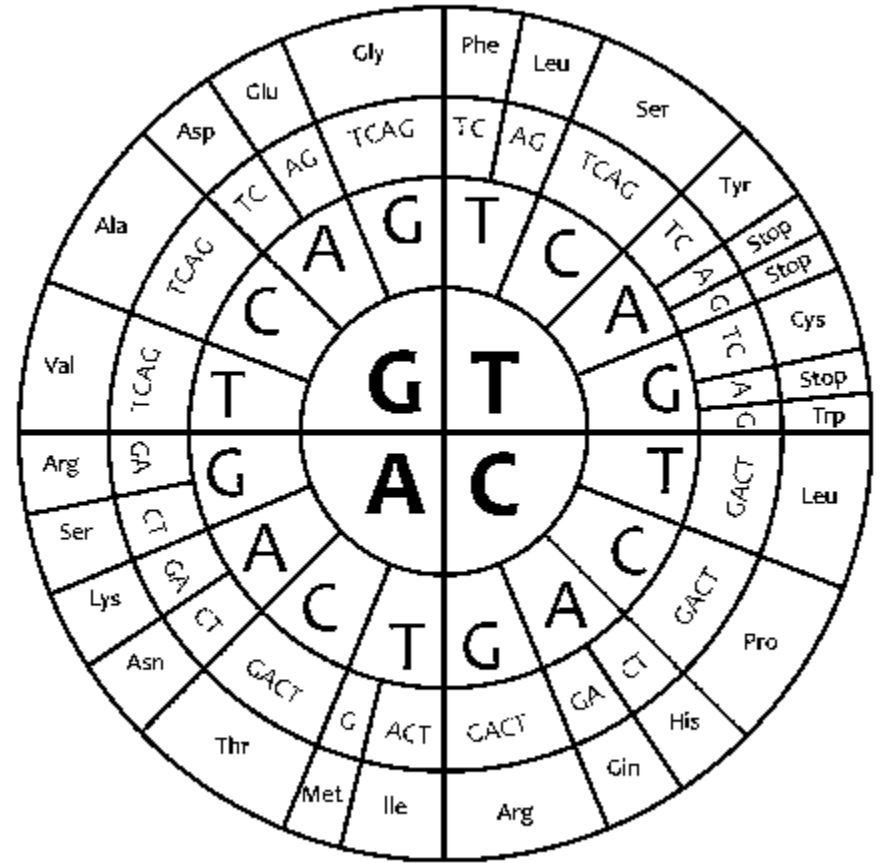
Alternation of function
usually pathological
- conservation

CONSERVED PROTEINS

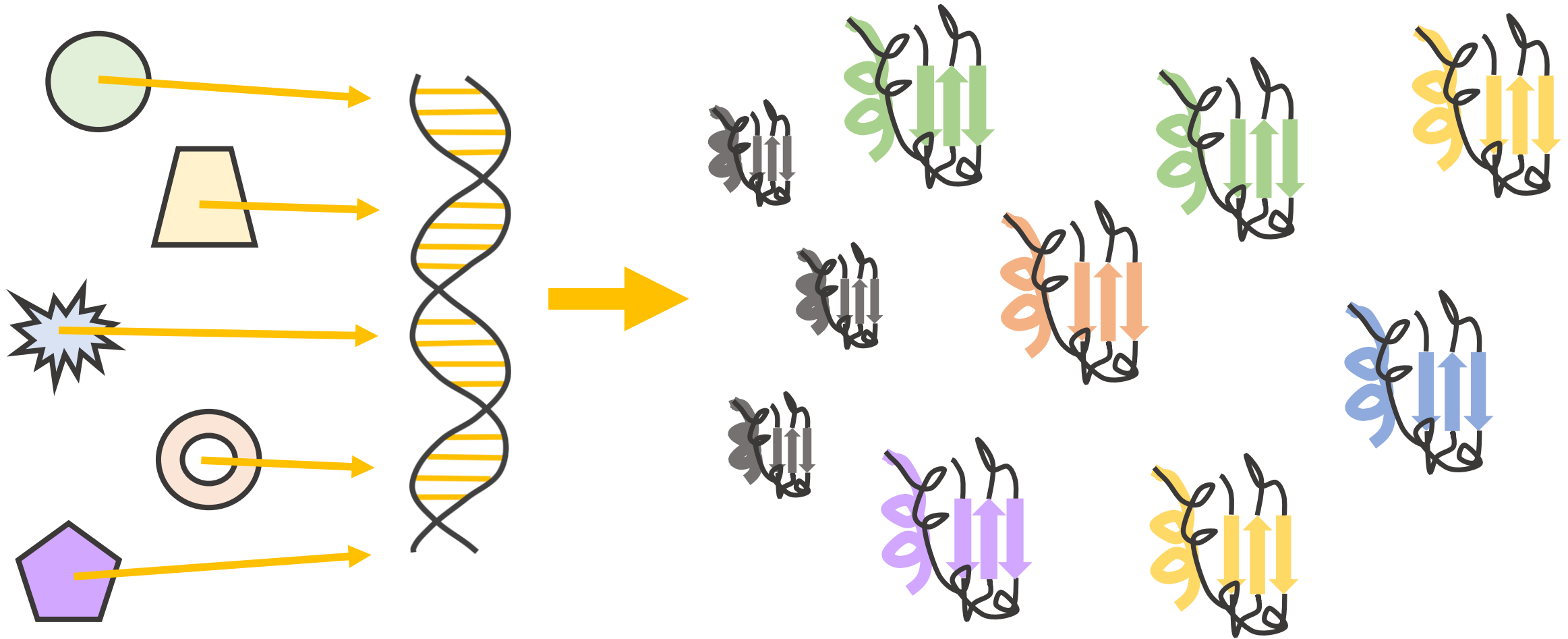
GCA GCG GCT TGC



Ala Ala Ala Trp



SAME DNA, DIFFERENT PROTEINS



DNA VS PROTEINS



AMPLIFICATION MISTAKES
AND CONTAMINATION

NO AMPLIFICATION



DNA VS PROTEINS



AMPLIFICATION MISTAKES
AND CONTAMINATION

CAN DISTINGUISH CLOSER SPECIES

NO AMPLIFICATION

OFTEN CONSERVED



DNA VS PROTEINS



AMPLIFICATION MISTAKES
AND CONTAMINATION

CAN DISTINGUISH CLOSER SPECIES

WORSE PRESERVATION

NO AMPLIFICATION

OFTEN CONSERVED

BETTER PRESERVATION



DNA VS PROTEINS



AMPLIFICATION MISTAKES
AND CONTAMINATION

CAN DISTINGUISH CLOSER SPECIES

WORSE PRESERVATION

THE SAME FOR ALL CELLS

NO AMPLIFICATION

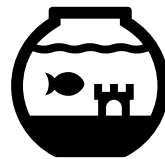
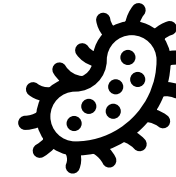
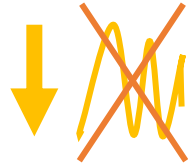
OFTEN CONSERVED

BETTER PRESERVATION

SHOWS FUNCTION AND
ACTIVE PROCESSES - TISSUE
DIFFERENCES



MOLECULAR PRESERVATION



MOLECULAR PRESERVATION

Article | [Published: 17 February 2021](#)

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

[Tom van der Valk](#) , [Patrícia Pečnerová](#), [David Díez-del-Molino](#), [Anders Bergström](#), [Jonas Oppenheimer](#), [Stefanie Hartmann](#), [Georgios Xenikoudakis](#), [Jessica A. Thomas](#), [Marianne Dehasque](#), [Ekin Saçlıcan](#), [Fatma Rabia Fidan](#), [Ian Barnes](#), [Shanlin Liu](#), [Mehmet Somel](#), [Peter D. Heintzman](#), [Pavel Nikolskiy](#), [Beth Shapiro](#), [Pontus Skoglund](#), [Michael Hofreiter](#), [Adrian M. Lister](#), [Anders Götherström](#) & [Love Dalén](#) 

Nature **591**, 265–269 (2021) | [Cite this article](#)

25k Accesses | 36 Citations | 2220 Altmetric | [Metrics](#)

DNA - 1 mil

Article | [Open access](#) | [Published: 07 December 2022](#)

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

[Kurt H. Kjær](#) , [Mikkel Winther Pedersen](#), [Bianca De Sanctis](#), [Binia De Cahsan](#), [Thorfinn S. Korneliusen](#), [Christian S. Michelsen](#), [Karina K. Sand](#), [Stanislav Jelavić](#), [Anthony H. Ruter](#), [Astrid M. A. Schmidt](#), [Kristian K. Kjeldsen](#), [Alexey S. Tesakov](#), [Ian Snowball](#), [John C. Gosse](#), [Inger G. Alsos](#), [Yucheng Wang](#), [Christoph Dockter](#), [Magnus Rasmussen](#), [Morten E. Jørgensen](#), [Birgitte Skadhauge](#), [Ana Prohaska](#), [Jeppe Å. Kristensen](#), [Morten Bjerager](#), [Morten E. Allentoft](#), [PhyloNorway Consortium](#), ... [Eske Willerslev](#)  [+ Show authors](#)

Nature **612**, 283–291 (2022) | [Cite this article](#)

136k Accesses | 50 Citations | 3492 Altmetric | [Metrics](#)

DNA - 2 mil

> [Elife](#). 2016 Sep 27;5:e17092. doi: 10.7554/eLife.17092.

Protein sequences bound to mineral surfaces persist into deep time

[Beatrice Demarchi](#)¹, [Shaun Hall](#)², [Teresa Roncal-Herrero](#)³, [Colin L Freeman](#)², [Jos Woolley](#)¹, [Molly K Crisp](#)⁴, [Julie Wilson](#)^{4 5}, [Anna Fotakis](#)⁶, [Roman Fischer](#)⁷, [Benedikt M Kessler](#)⁷, [Rosa Rakownikow Jersie-Christensen](#)⁸, [Jesper V Olsen](#)⁸, [James Haile](#)⁹, [Jessica Thomas](#)^{6 10}, [Curtis W Marean](#)^{11 12}, [John Parkington](#)¹³, [Samantha Presslee](#)¹, [Julia Lee-Thorp](#)⁹, [Peter Ditchfield](#)⁹, [Jacqueline F Hamilton](#)¹⁴, [Martyn W Ward](#)¹⁴, [Chunting Michelle Wang](#)¹⁴, [Marvin D Shaw](#)¹⁴, [Terry Harrison](#)¹⁵, [Manuel Domínguez-Rodrigo](#)¹⁶, [Ross DE MacPhee](#)¹⁷, [Amandus Kwekason](#)¹⁸, [Michaela Ecker](#)⁹, [Liora Kolska Horwitz](#)¹⁹, [Michael Chazan](#)^{20 21}, [Roland Kröger](#)³, [Jane Thomas-Oates](#)^{4 22}, [John H Harding](#)², [Enrico Cappellini](#)⁶, [Kirsty Penkman](#)⁴, [Matthew J Collins](#)¹

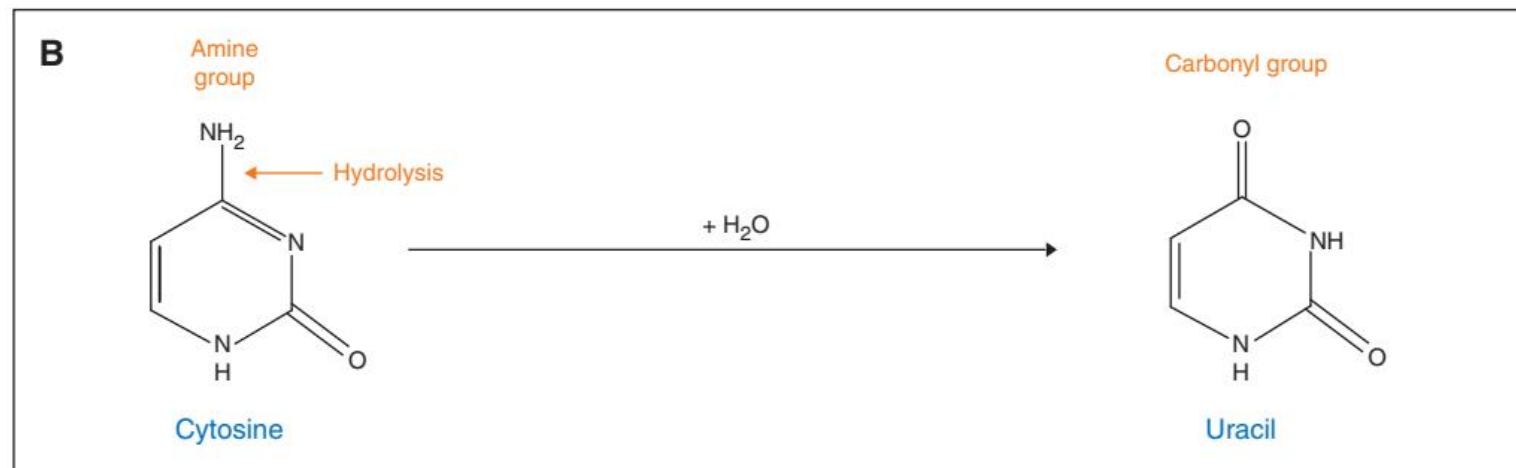
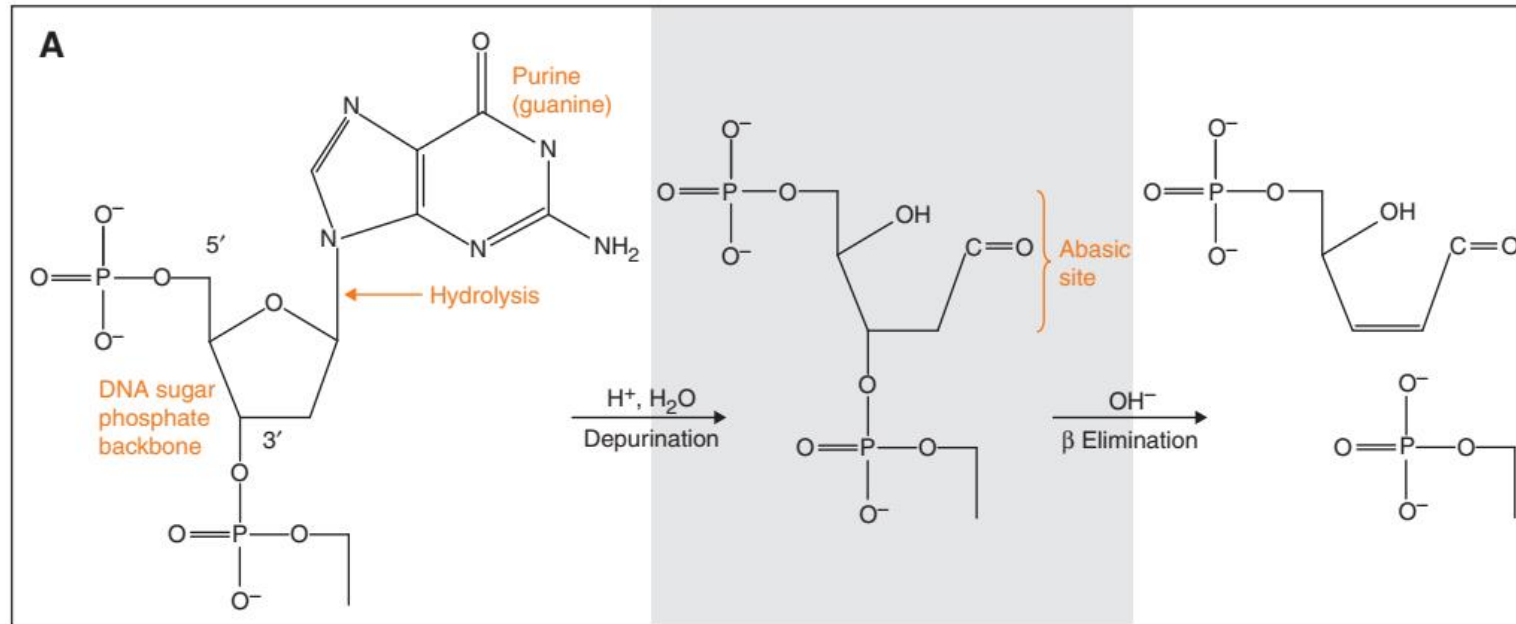
Affiliations [+ expand](#)

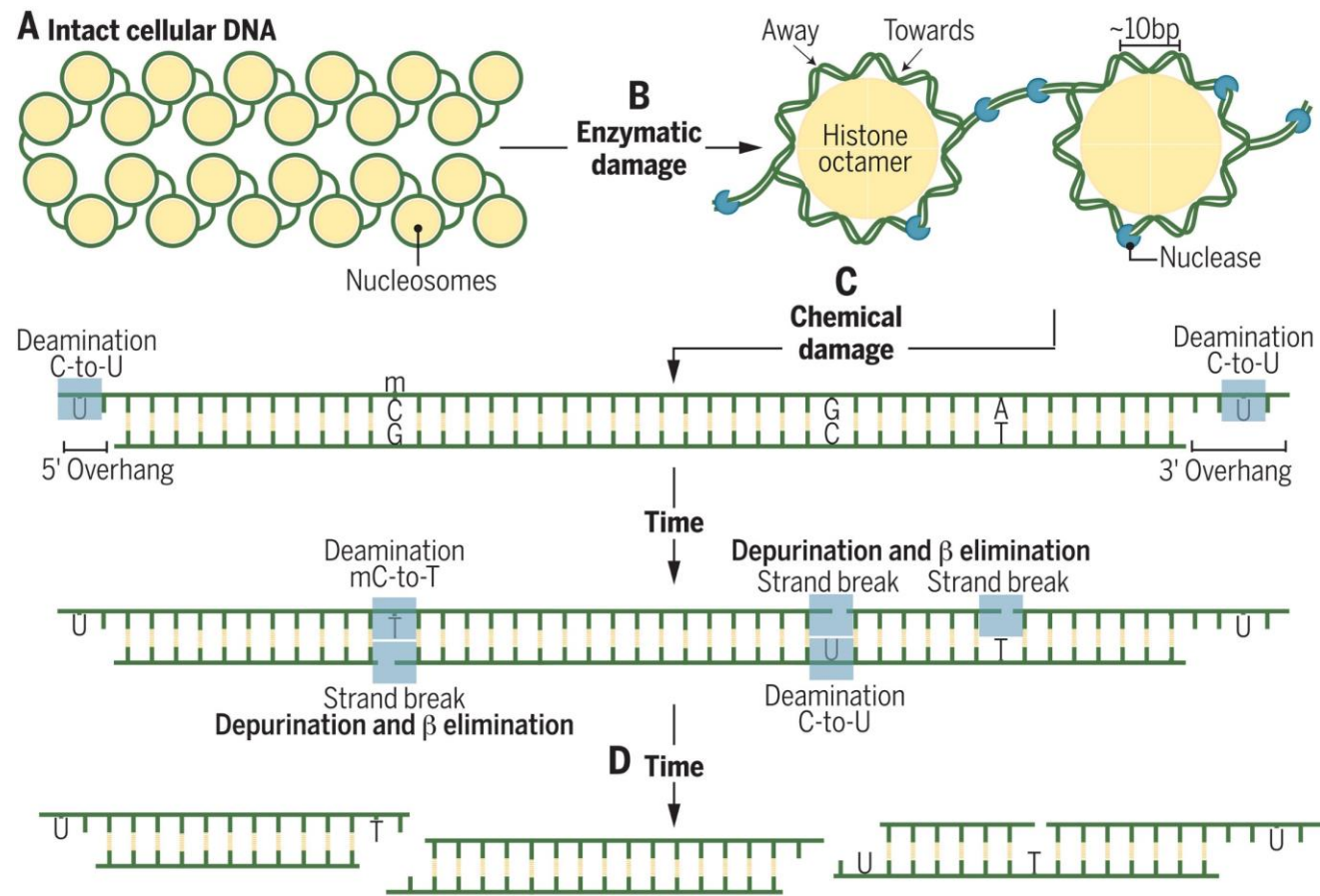
PMID: 27668515 PMCID: PMC5039028 DOI: 10.7554/eLife.17092

[Free PMC article](#)

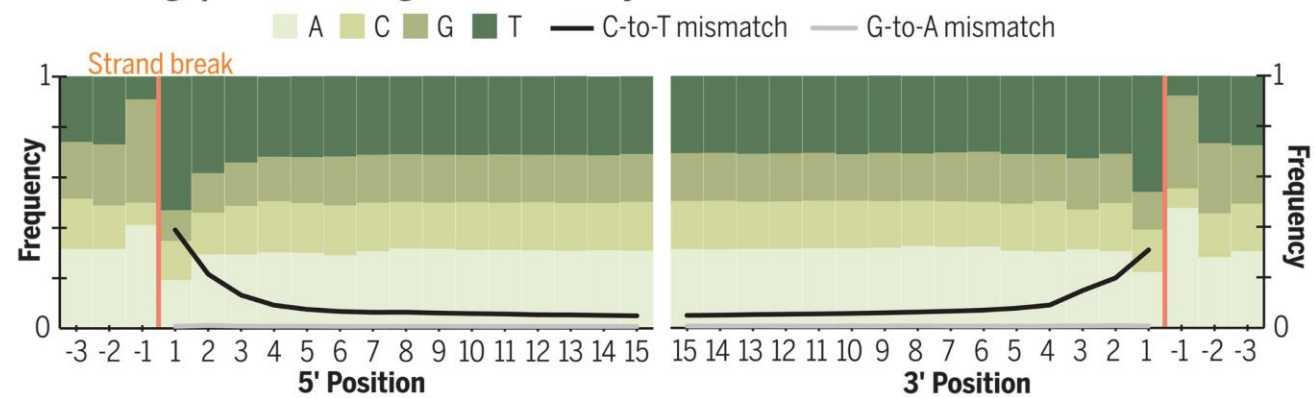
Proteins - 3,8 mil

PRESERVATION OF DNA

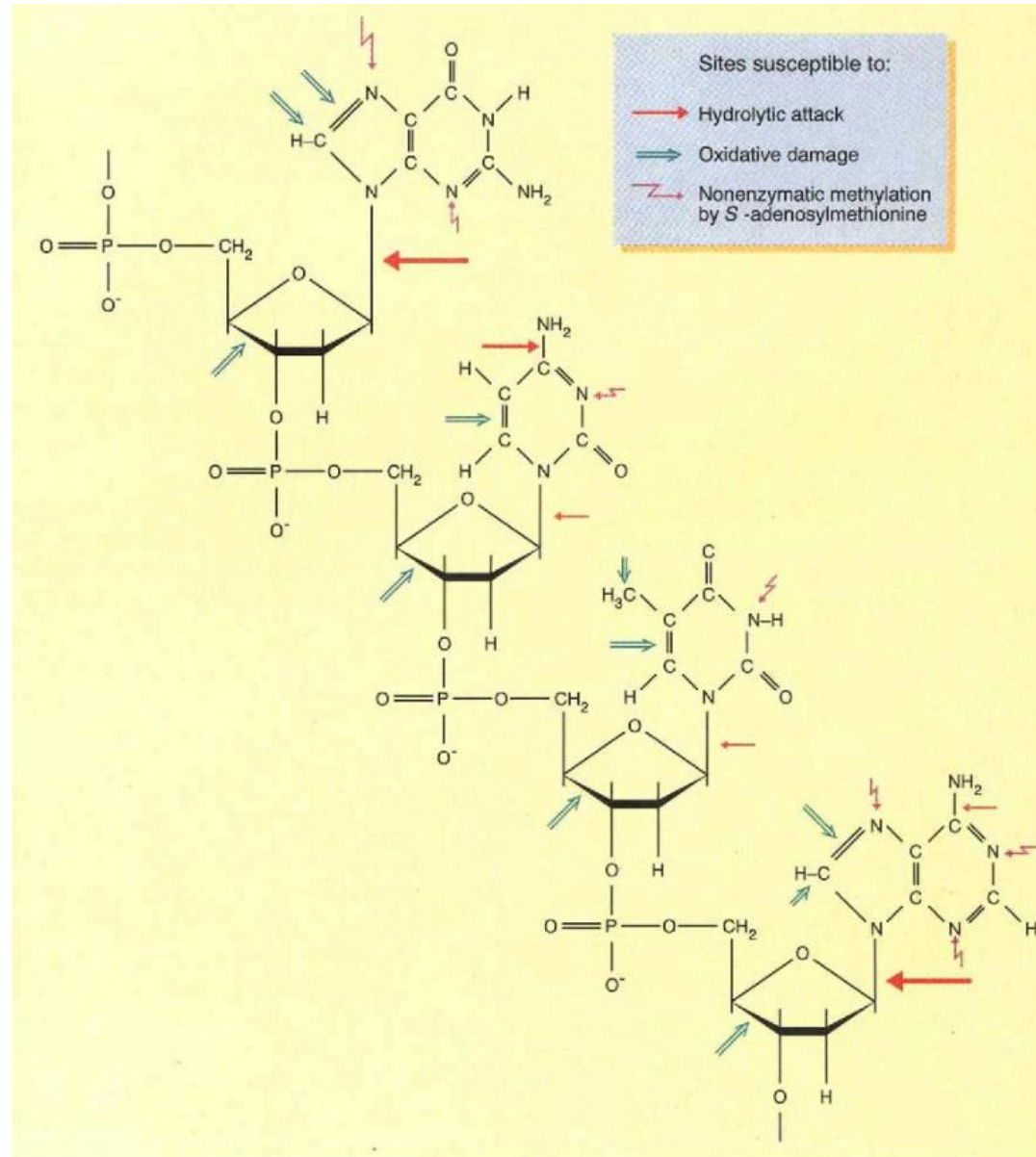




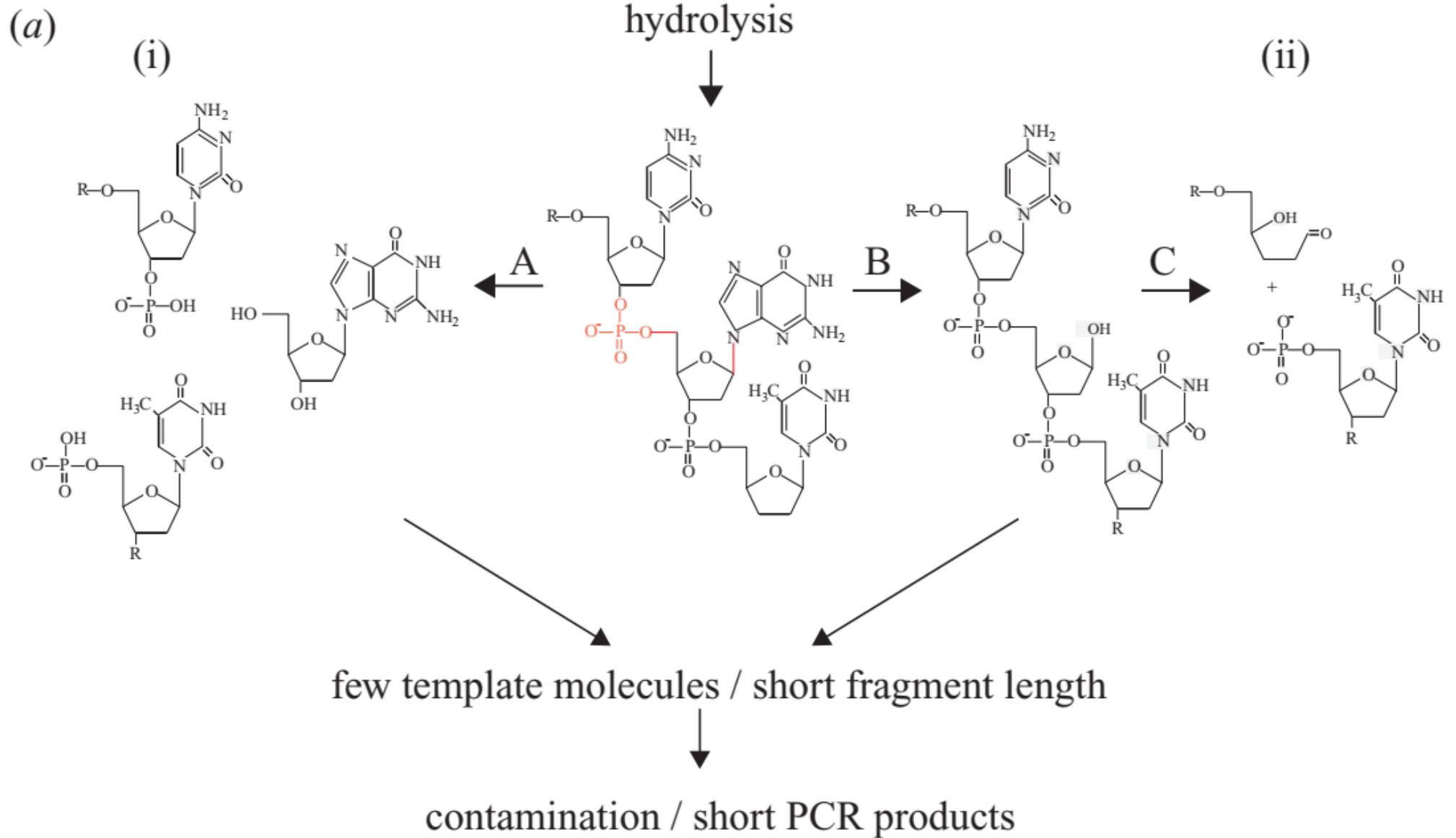
E DNA damage patterns of a single-stranded library



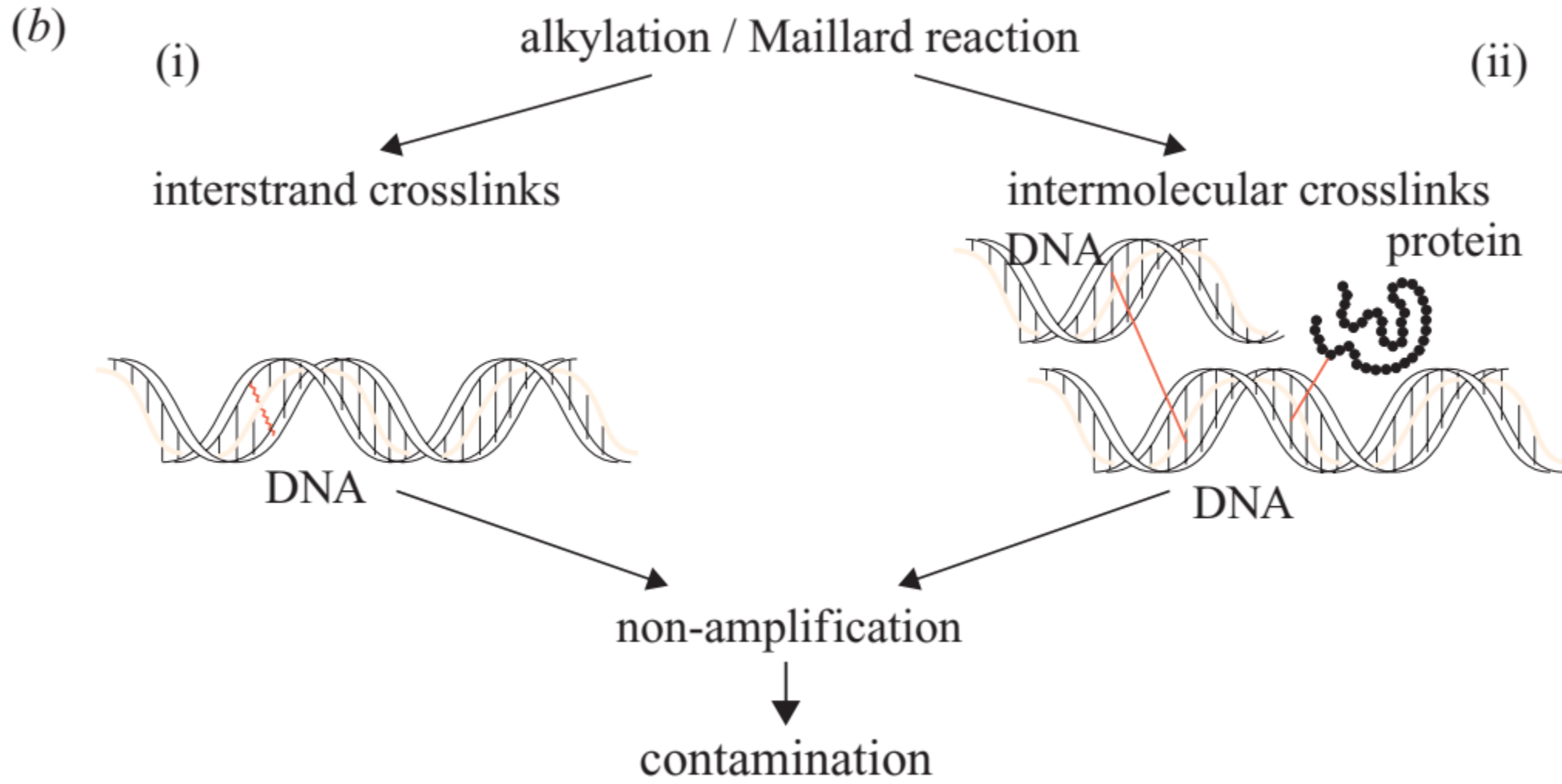
MOLECULAR PRESERVATION - DNA



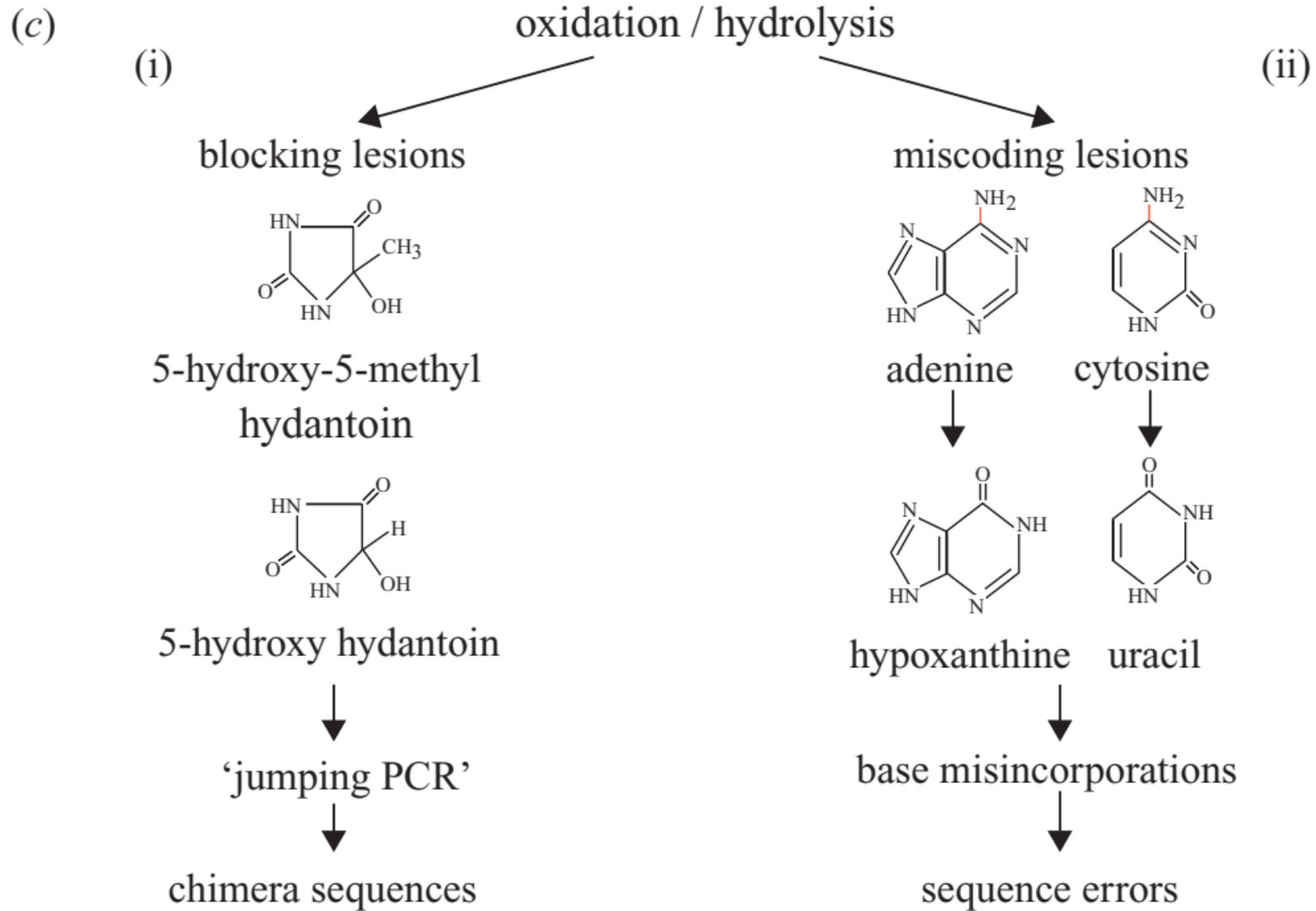
PRESERVATION OF DNA



PRESERVATION OF DNA

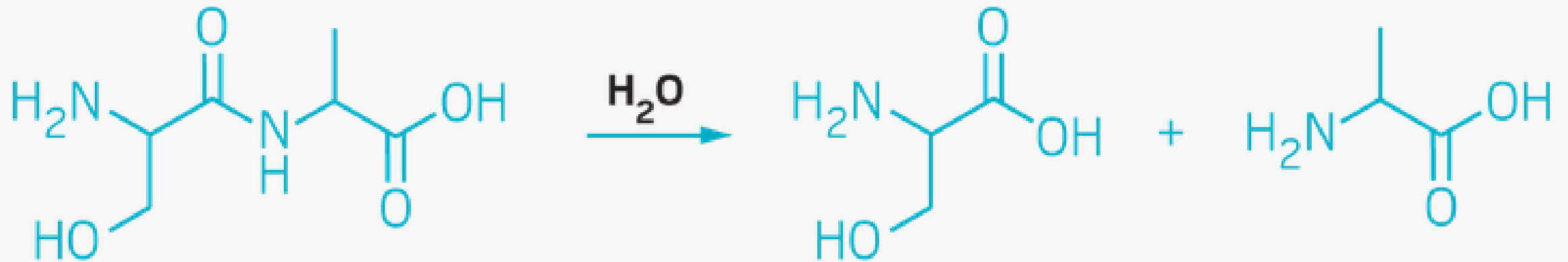


PRESERVATION OF DNA



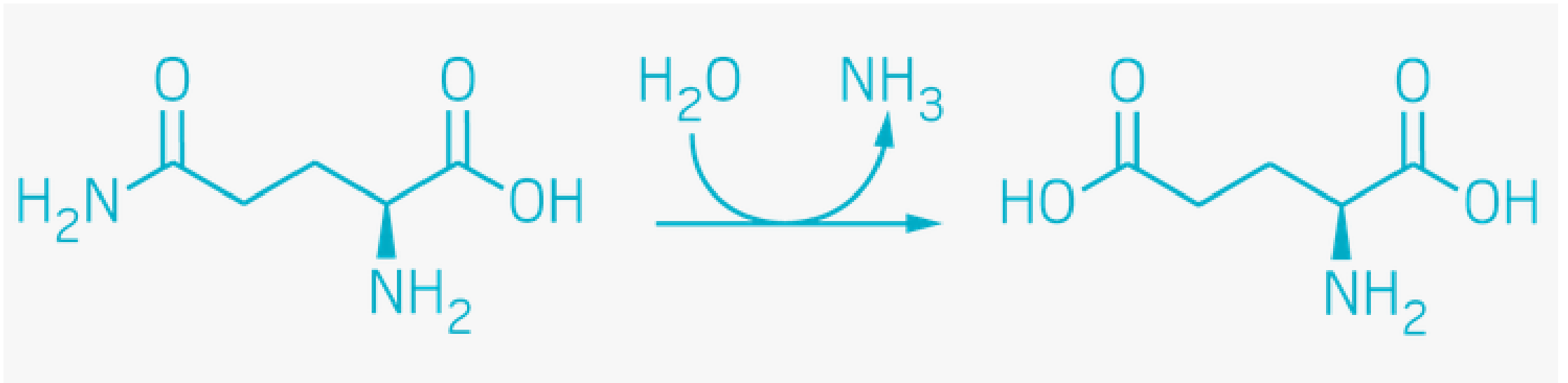
PRESERVATION OF PROTEINS

Hydrolysis, deamidation, racemination



PRESERVATION OF PROTEINS

Hydrolysis, **deamidation**, racemination



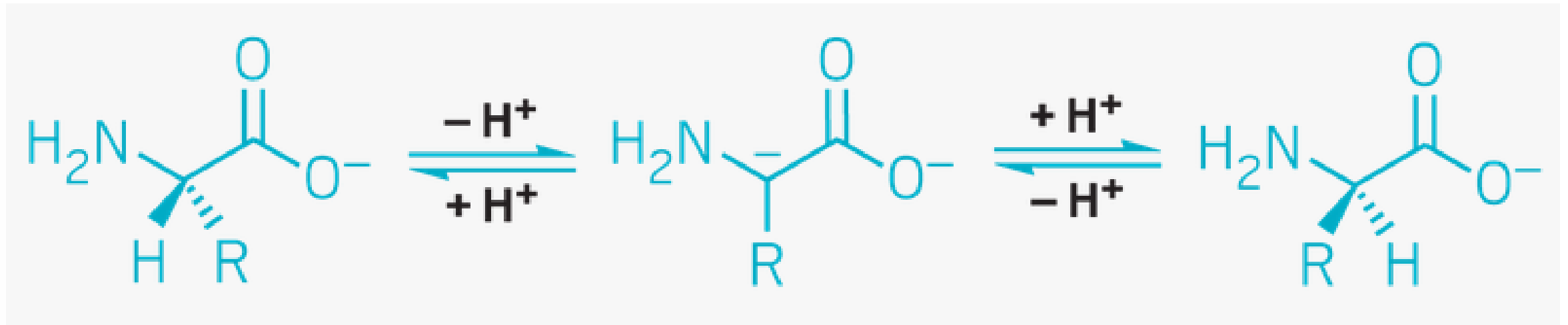
asparagine and glutamine




aspartic and glutamic acid

PRESERVATION OF PROTEINS

Hydrolysis, deamidation, **racemination**



L to D form, can break from the chain

- 
- Quagga phenotype deextinction
 - Protein vs DNA preservation – oldest
 - Research distribution in the world



2024:

- UDG treatment
- Relevance of older methods
- Extrapolation from ancient genomes is biased by sampling
- Preservation influences
- Quagga
- Genes from Denisovans
- Cultural objects
- Oldest proteins (3.8 M), oldest aDNA (1 M, 2 M)
- Deamination in aDNA, deamidation in proteins