EPFL

Genomic solutions to sustainable development

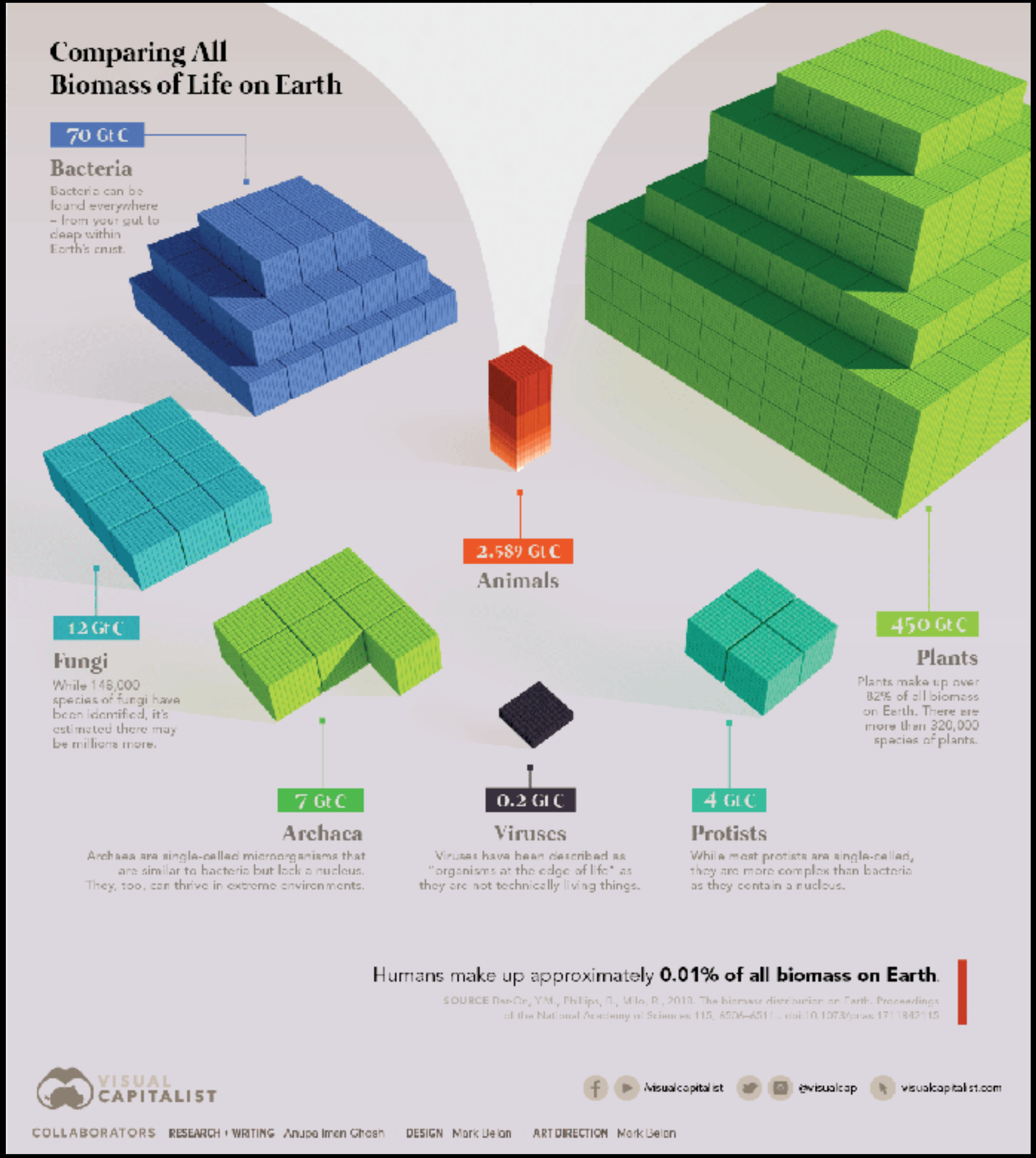
Week 6 — Environmental DNA

25 March 2025

Sebastian M. Waszak, Ph.D.
Assistant Professor, Life Sciences, EPFL
Associate Adjunct Professor, Neurology, UCSF

École
polytechnique
fédérale
de Lausanne

Global biodiversity in numbers



Comparing All Biomass of Life on Earth

70 Gt C
Bacteria
Bacteria can be found everywhere – from your gut to deep within Earth's crust.

12 Gt C
Fungi
While 148,000 species of fungi have been identified, it's estimated there may be millions more.

2.589 Gt C
Animals

450 Gt C
Plants
Plants make up over 82% of all biomass on Earth. There are more than 320,000 species of plants.

7 Gt C
Archaea
Archaea are single-celled microorganisms that are similar to bacteria but lack a nucleus. They, too, can thrive in extreme environments.

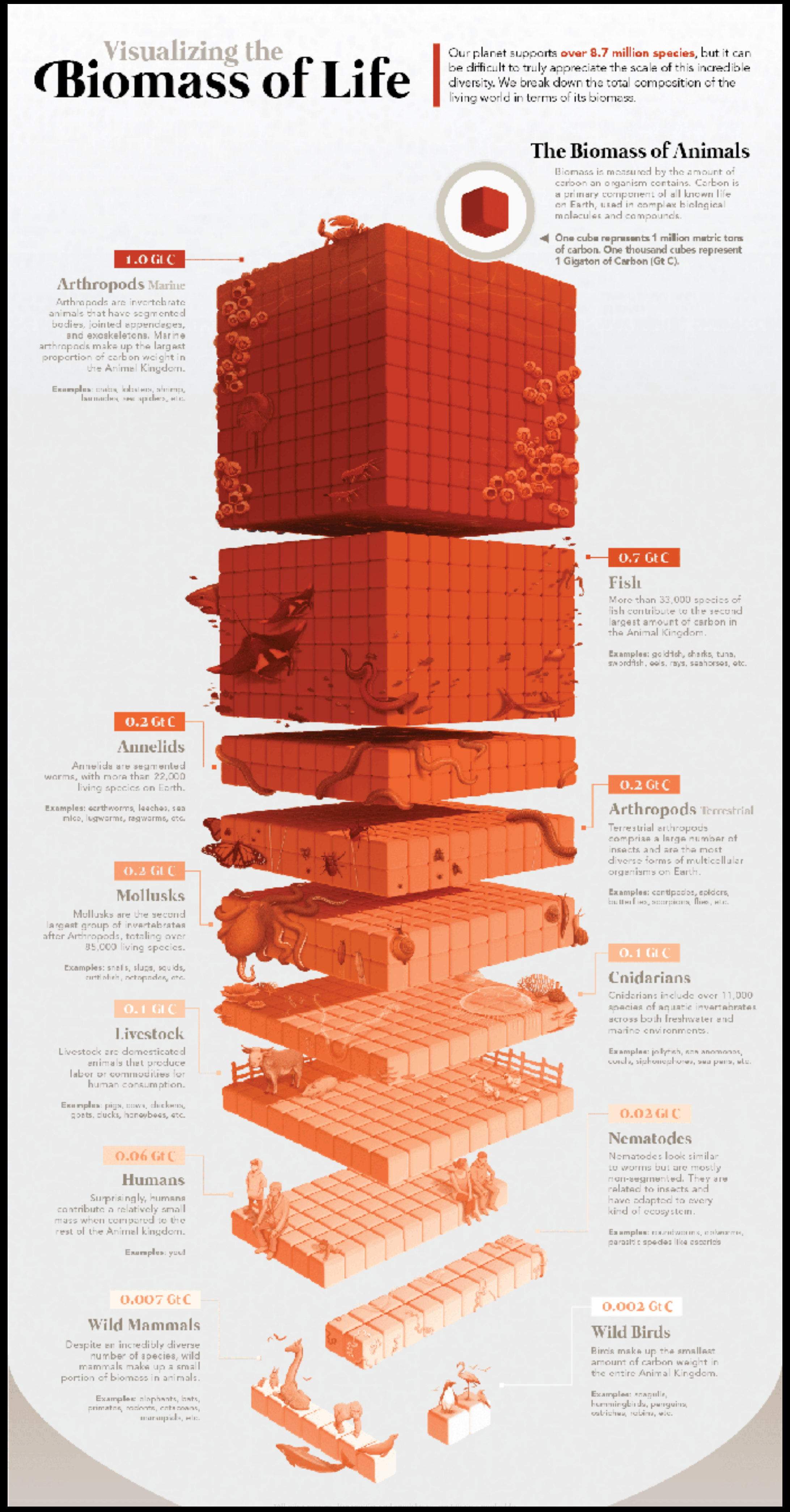
0.2 Gt C
Viruses
Viruses have been described as "organisms at the edge of life" as they are not technically living things.

4 Gt C
Protists
While most protists are single-celled, they are more complex than bacteria as they contain a nucleus.

Humans make up approximately **0.01% of all biomass on Earth.**

SOURCE Bar-On, Y.M., Phillips, R., Milo, R., 2018. The biomass distribution on Earth. Proceedings of the National Academy of Sciences 115, 6506–6511. doi:10.1073/pnas.1711842115

COLLABORATORS RESEARCH + WRITING Anupa Imen Ghosh DESIGN Mark Belan ART DIRECTION Mark Belan



Visualizing the Biomass of Life

Our planet supports **over 8.7 million species**, but it can be difficult to truly appreciate the scale of this incredible diversity. We break down the total composition of the living world in terms of its biomass.

The Biomass of Animals

Biomass is measured by the amount of carbon an organism contains. Carbon is a primary component of all known life on Earth, used in complex biological molecules and compounds.

One cube represents 1 million metric tons of carbon. One thousand cubes represent 1 Gigaton of Carbon (Gt C).

1.0 Gt C
Arthropods Marine
Arthropods are invertebrate animals that have segmented bodies, jointed appendages, and exoskeletons. Marine arthropods make up the largest proportion of carbon weight in the Animal Kingdom.
Examples: crabs, lobsters, shrimp, barnacles, sea spiders, etc.

0.7 Gt C
Fish
More than 33,000 species of fish contribute to the second largest amount of carbon in the Animal Kingdom.
Examples: goldfish, sharks, tuna, swordfish, eels, rays, seahorses, etc.

0.2 Gt C
Annelids
Annelids are segmented worms, with more than 22,000 living species on Earth.
Examples: earthworms, leeches, sea mice, lugworms, ragworms, etc.

0.2 Gt C
Arthropods Terrestrial
Terrestrial arthropods comprise a large number of insects and are the most diverse forms of multicellular organisms on Earth.
Examples: centipedes, spiders, butterflies, scorpions, flies, etc.

0.2 Gt C
Mollusks
Mollusks are the second largest group of invertebrates after Arthropods, totaling over 85,000 living species.
Examples: snails, slugs, squids, cuttlefish, octopodes, etc.

0.1 Gt C
Cnidarians
Cnidarians include over 11,000 species of aquatic invertebrates across both freshwater and marine environments.
Examples: jellyfish, sea anemones, corals, siphonophores, sea pens, etc.

0.1 Gt C
Livestock
Livestock are domesticated animals that produce labor or commodities for human consumption.
Examples: pigs, cows, chickens, goats, ducks, honeybees, etc.

0.02 Gt C
Nematodes
Nematodes look similar to worms but are mostly non-segmented. They are related to insects and have adapted to every kind of ecosystem.
Examples: roundworms, eelworms, parasitic species like ascarids

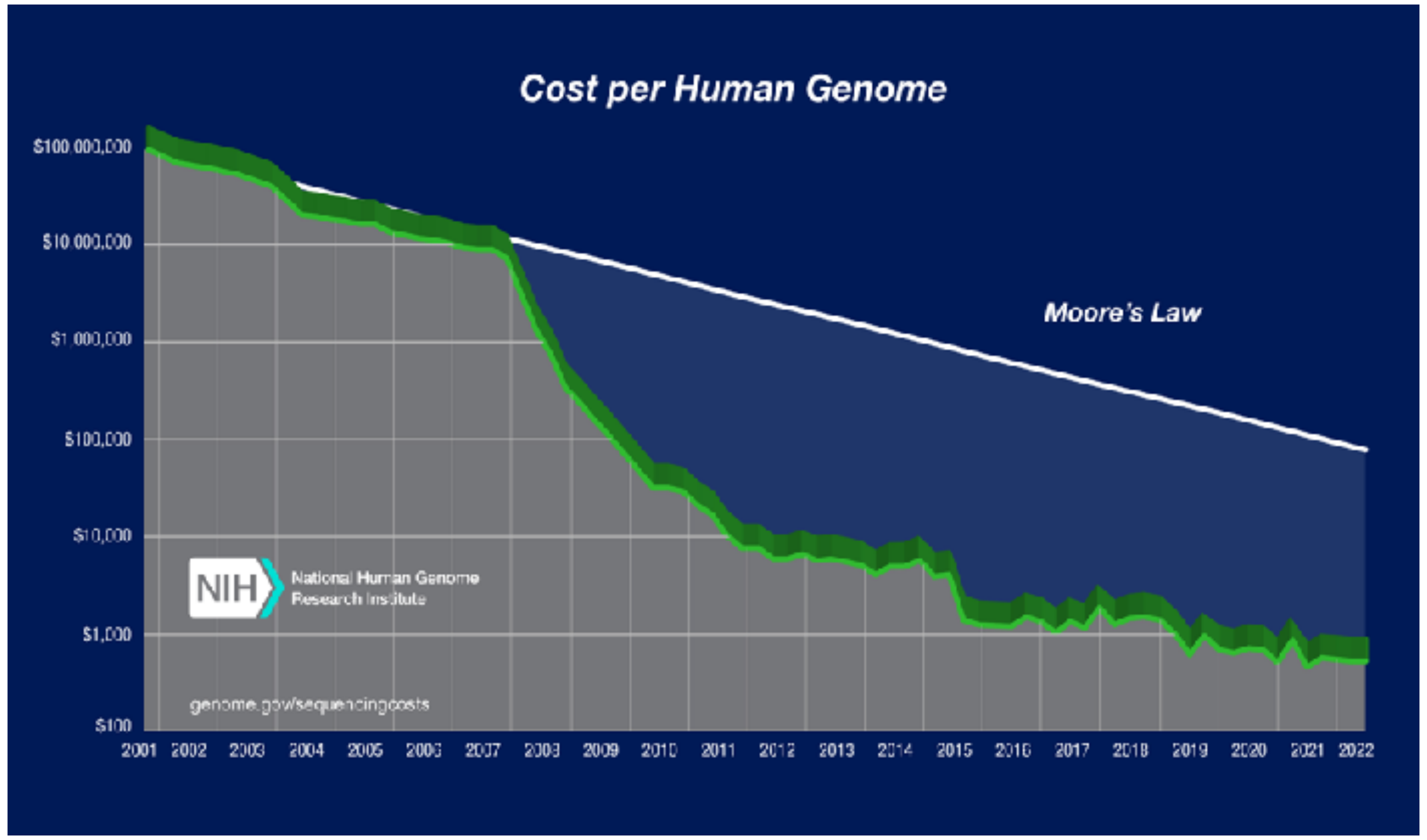
0.06 Gt C
Humans
Surprisingly, humans contribute a relatively small mass when compared to the rest of the Animal kingdom.
Examples: you!

0.007 Gt C
Wild Mammals
Despite an incredibly diverse number of species, wild mammals make up a small portion of biomass in animals.
Examples: elephants, bats, primates, rodents, cetaceans, marsupials, etc.

0.002 Gt C
Wild Birds
Birds make up the smallest amount of carbon weight in the entire Animal Kingdom.
Examples: seagulls, hummingbirds, penguins, ostriches, robins, etc.

DNA sequencing during the past 20 years

- **Moore's law**
  - Number of transistors on microchips double every 2 years
- **DNA sequencing costs**
  - Sanger DNA sequencing
  - Next generation sequencing (since 2008)
  - From $100M/genome 20 years ago to $500/genome now
- **Human genomes**
  - First human genome draft (2001)
  - Human population diversity based on 2,500 individuals (2013)
  - Human population diversity based on 1,000,000 individuals (2021)



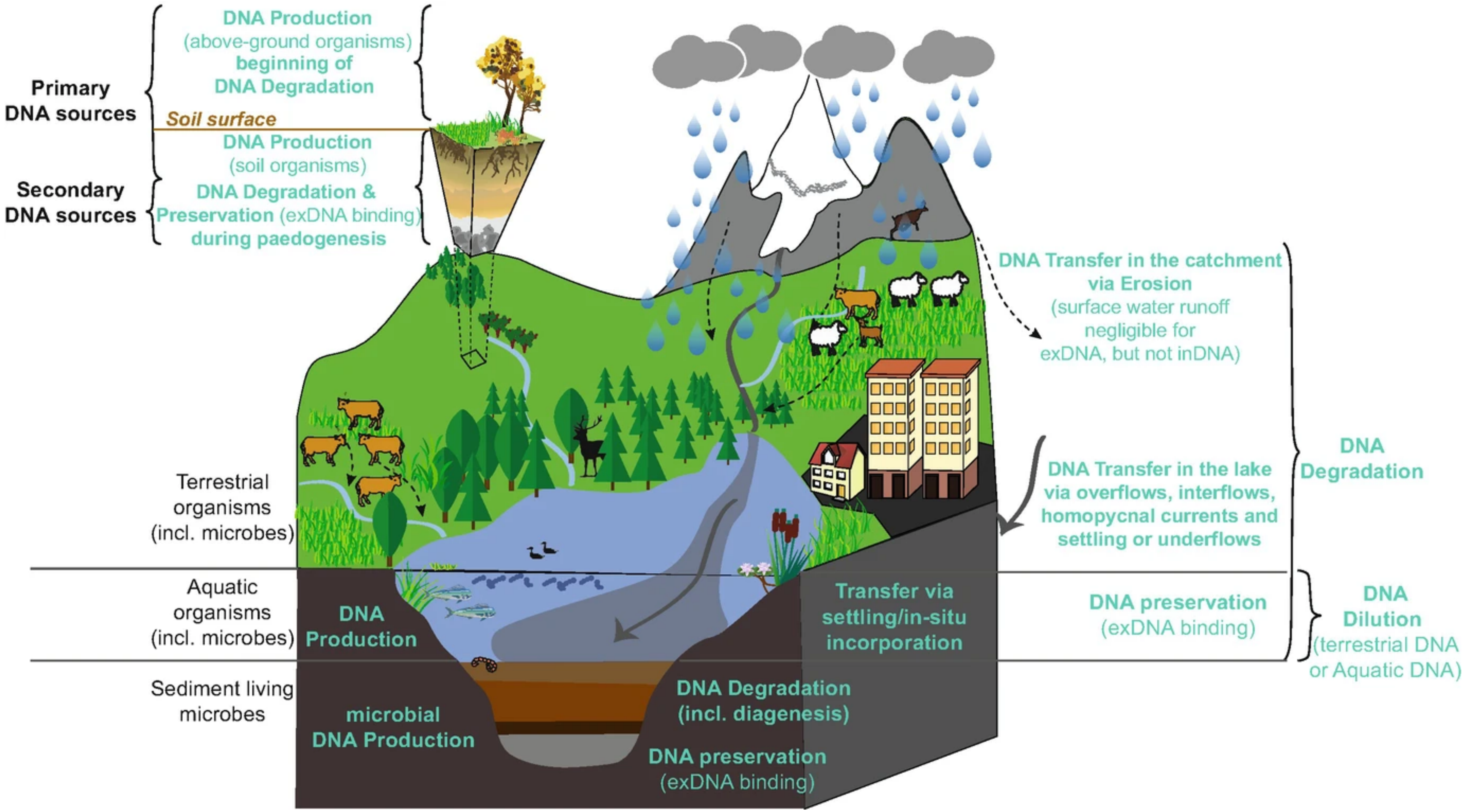
DNA sequencing is now a technology open to many scientific fields

Environmental DNA (eDNA)

Monitor life: any time, any place, any depth

- DNA is everywhere and derived from skin, hair, mucus, urine, feces, and decomposing organisms
- eDNA present in soil, water, ice, sediments, and air (1-1000 ng per liter; 1-10,000 ng per g)
- eDNA can capture ecosystems over time (past, present, future)
- eDNA potentially a novel approach to detect extraterrestrial life forms
- Benefits for biodiversity monitoring with eDNA sampling
  - Accessible (everyone can participate and at any site)
  - Non-invasive sampling (ie, no need to capture animals)
  - Broad sampling range (ie, many species from a single sample)
  - Cost-effective (ie, richer diversity captured per $ per bp)

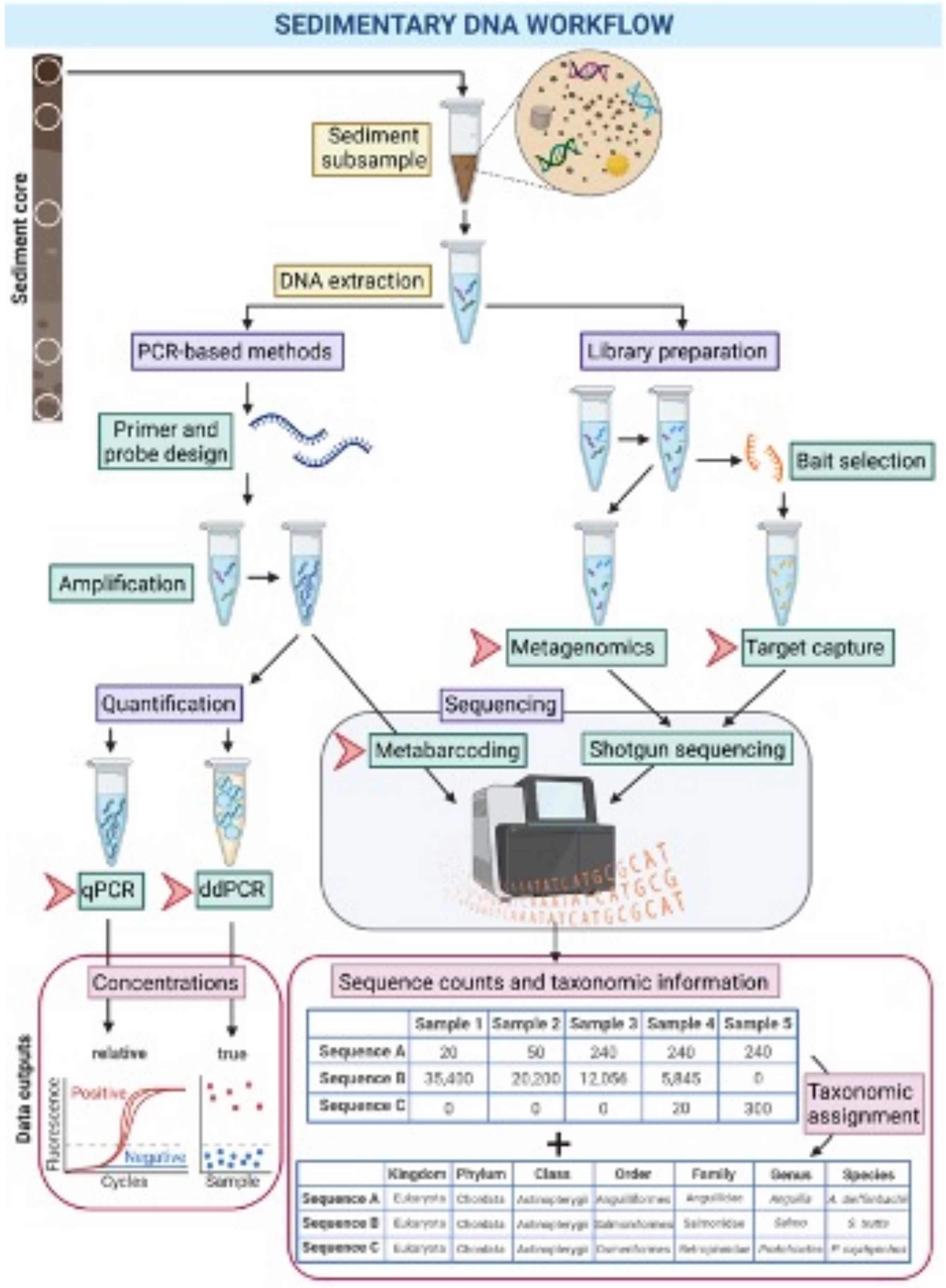
Sources and fates of eDNA in the environment



https://link.springer.com/chapter/10.1007/978-3-031-43799-1_2

inDNA = intracellular DNA
exDNA = extracellular DNA

eDNA processing workflows



Sequence counts and taxonomic information

	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
Sequence A	20	50	240	240	240
Sequence B	35,400	20,200	12,056	5,845	0
Sequence C	0	0	0	20	300

+

	Kingdom	Phylum	Class	Order	Family	Genus	Species
Sequence A	Eukaryota	Chordata	Actinopterygii	Anguilliformes	Anguillidae	Anguilla	A. dieffenbachii
Sequence B	Eukaryota	Chordata	Actinopterygii	Salmoniformes	Salmonidae	Salmo	S. trutta
Sequence C	Eukaryota	Chordata	Actinopterygii	[illegible]	[illegible]	[illegible]	[illegible]

Encyclopedia of Quaternary Science (Third Edition)
Volume 6, 2025, Pages 673-690

Complexity and preservation of eDNA (I)

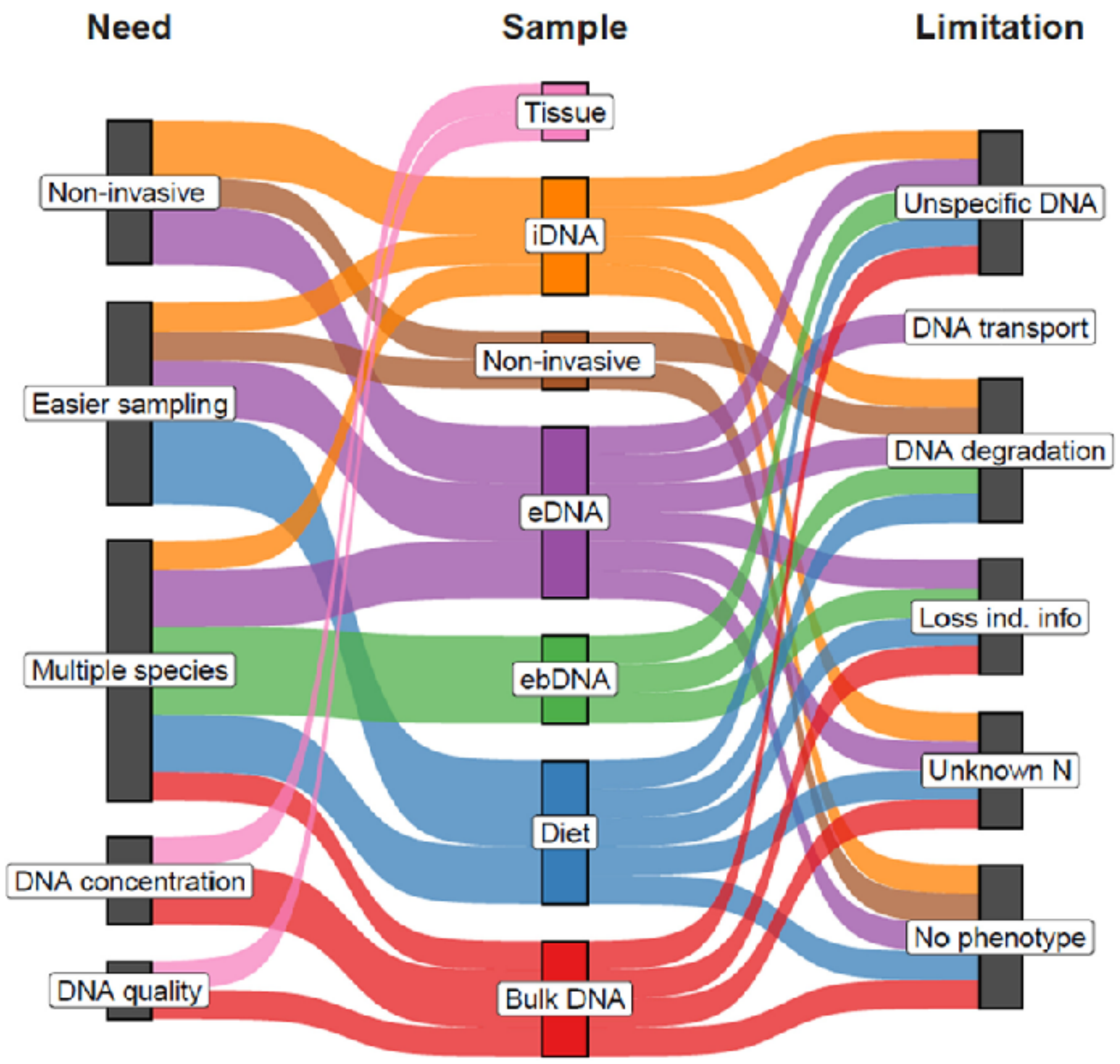
- **Complexity**: Environmental samples contain a mix of genetic material from multiple species
- **Preservation**: in certain environments, eDNA can be preserved for extended periods (freshwater and marine sediments, paleo soils, caves, permafrost)
- **Processes affecting eDNA**: degradation (via DNases), abiotic decay, absorption, and vertical migration within sediments
- **Impact of minerals**: mineral composition of sediments can impact preservation (eg, DNA absorbed by different types of clay)

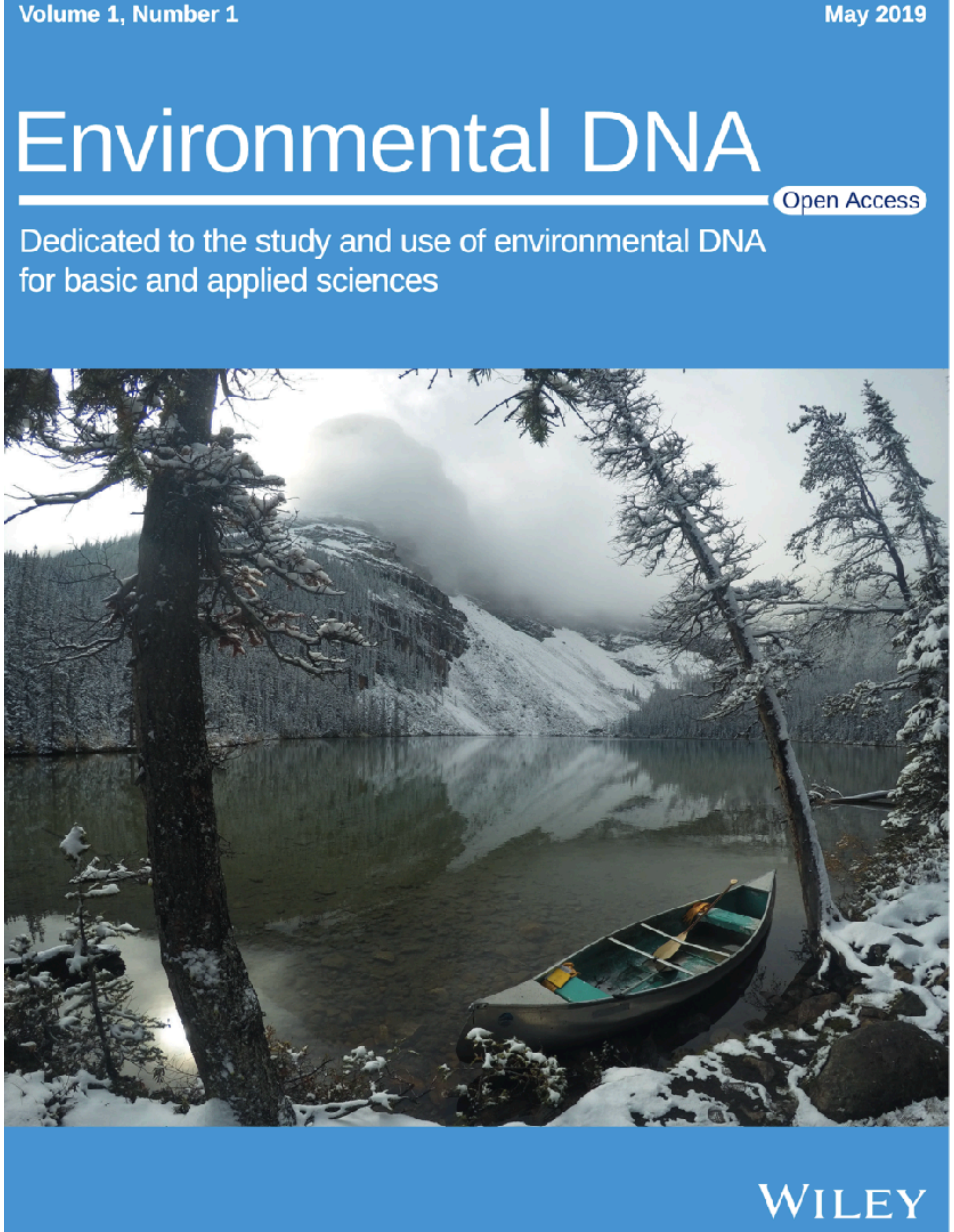
Complexity and preservation of eDNA (II)

- **Sampling:** collect environmental samples (eg, soil, water) while minimizing contamination
- **Extraction**: DNA isolation based on specialized protocols (water, soil)
- **Analysis**: targeted vs untargeted sequencing; PCR amplification and sequencing with Illumina (short-read) or Nanopore (long-read)

Limitations and challenges of eDNA in biodiversity monitoring

- Hard to separate DNA of different species (need for good reference genomes)
- DNA degradation and transport might misrepresent biodiversity
- No information about single individuals
- Unknown contributors: hard to determine number, age, or life stage of individuals





Volume 1, Number 1

May 2019

Environmental DNA

Open Access

Dedicated to the study and use of environmental DNA for basic and applied sciences

WILEY

Ancient environmental DNA

Karen Arnott/EMBL-EBI

- Unbiased profiling of ancient ecosystems
- Traditional fossil records are biased towards large and well-preserved remains
- Discovery of known species in eDNA samples based on reference genomes
- Population genomics based on ancient environmental DNA
- Relate ancient eDNA samples to current eDNA samples (long-term adaptations)
- Relate past biodiversity, climate conditions, and extinction events
- Recover past global species diversity (eg, plants)

Many plants have lost genetic diversity due to human cultivation

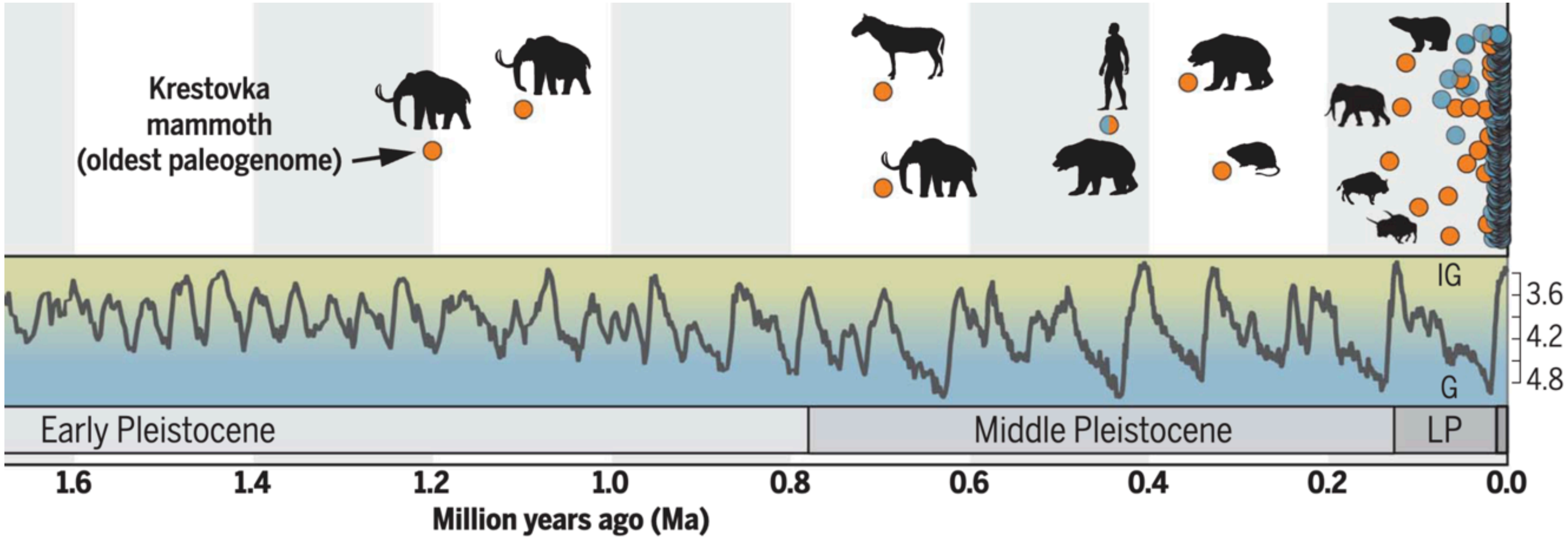
Mary Evans Picture Library/Grosvenor Prints

Ancient Environmental Genomics Initiative for Sustainability

(CHF 75M grant from the Novo Nordisk Foundation and Wellcome Trust)

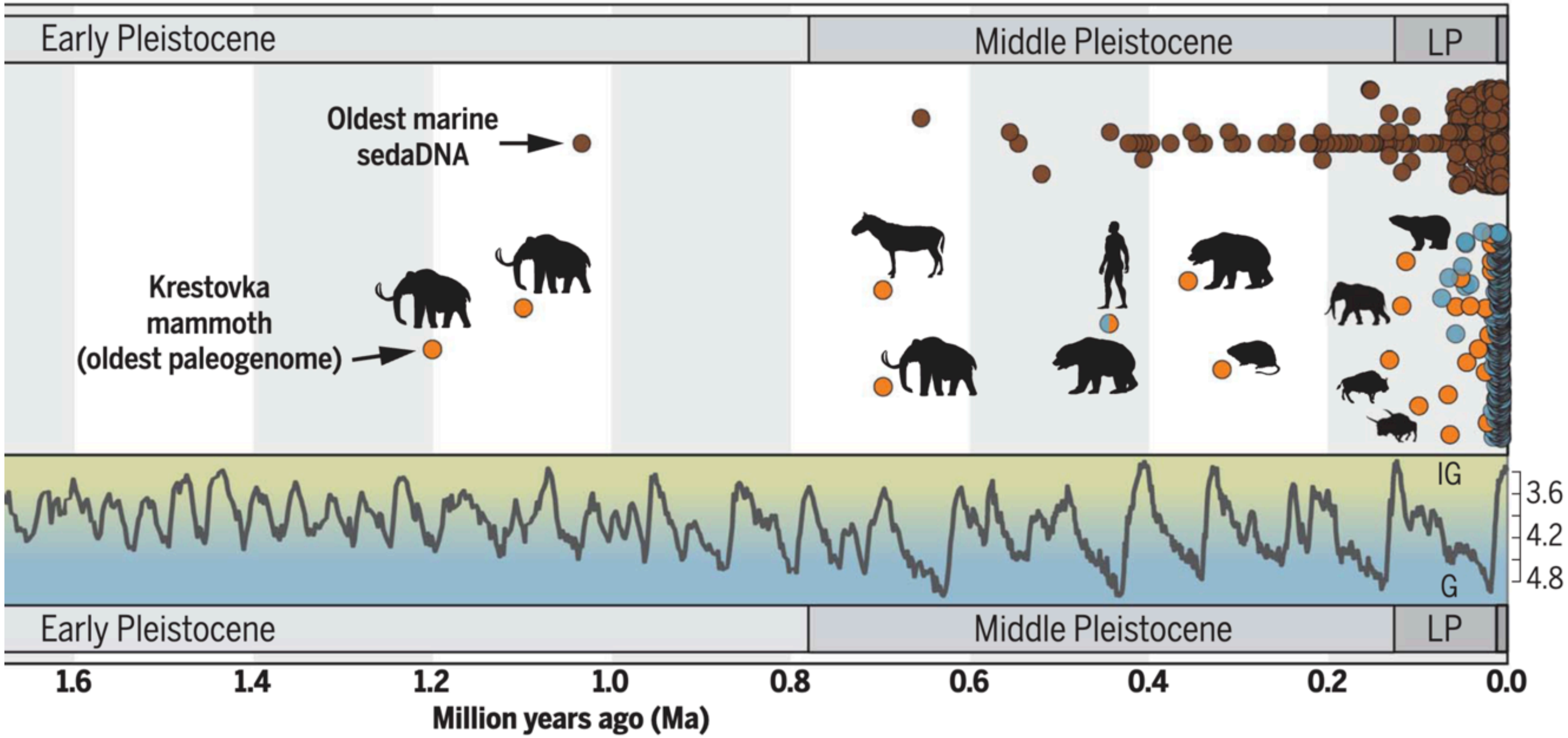
https://novonordiskfonden.dk/en/news/climate-resilient-crops-novo-nordisk-foundation-and-the-wellcome-trust-award-a-grant-of-up-to-dkk-585-million-to-the-ancient-environmental-genomics-initiative-for-sustainability/

Most fossil-based ancient DNA studies focused on the past 50,000 years



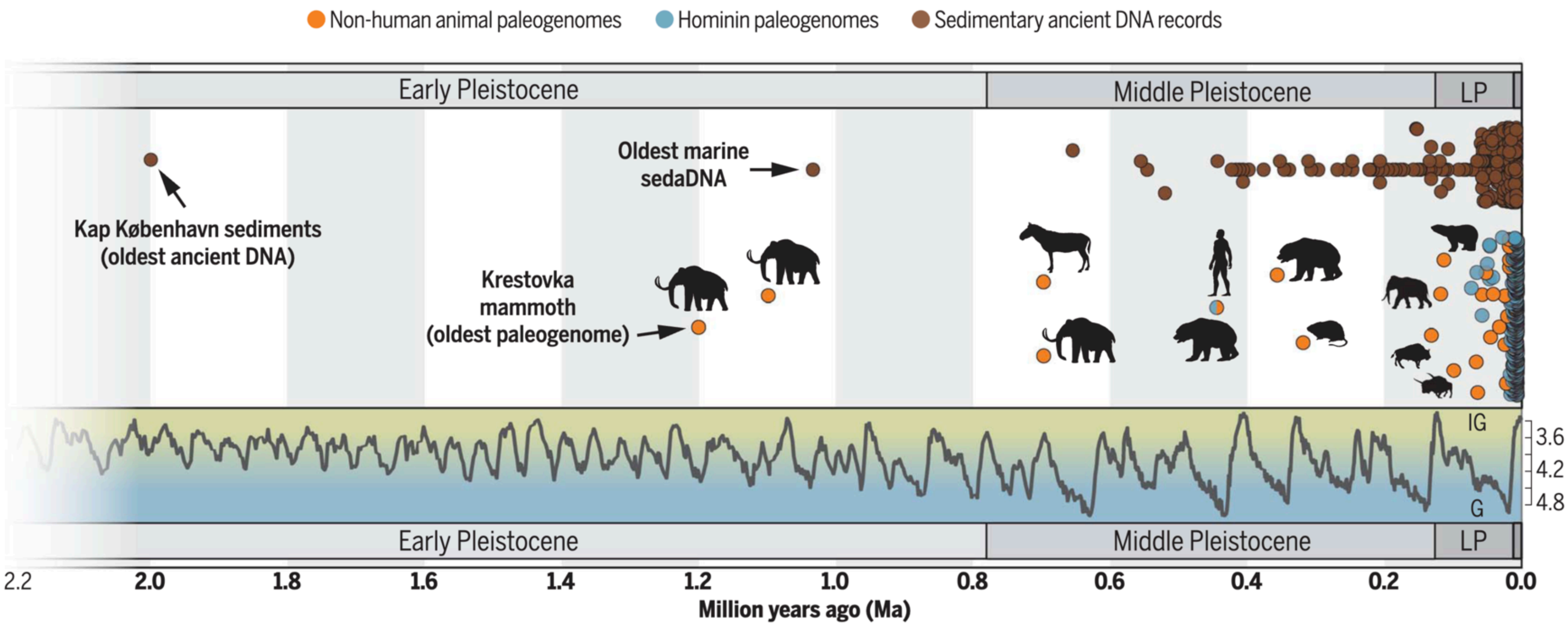
Dalén et al., Science 382, 48–53 (2023)

Ancient eDNA studies go deeper in time



Dalén et al., Science 382, 48–53 (2023)

Oldest ancient DNA uncovered with eDNA

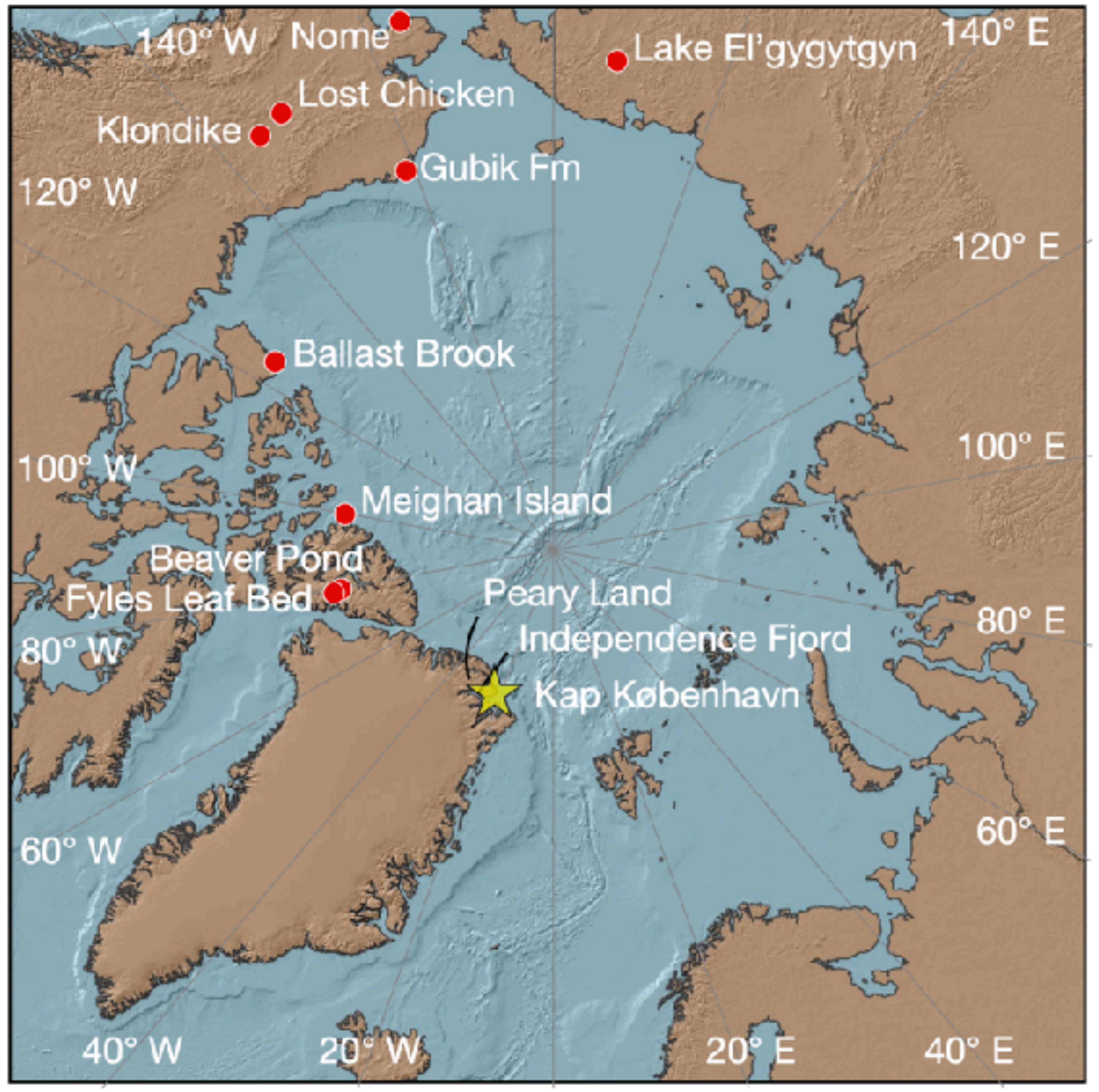


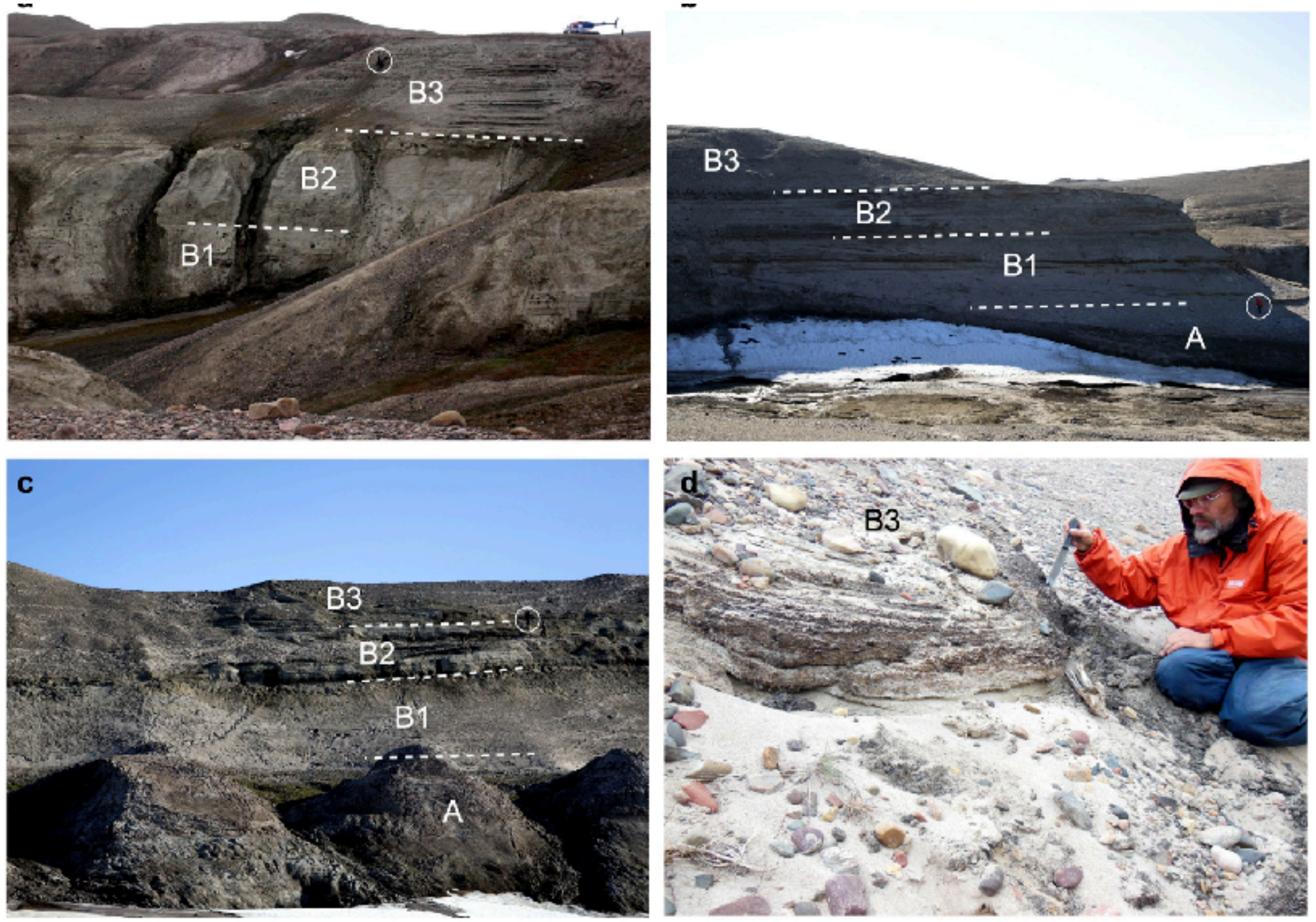
Dalén et al., Science 382, 48–53 (2023)

2-million-year-old ecosystems

eDNA sequencing informs about the paleoclimate

- Kap Kobenhavn Formation (Greenland) — dated to 2.4 Myrs ago — houses 100m thick successions of shallow marine nearshore sediments with ancient fossils



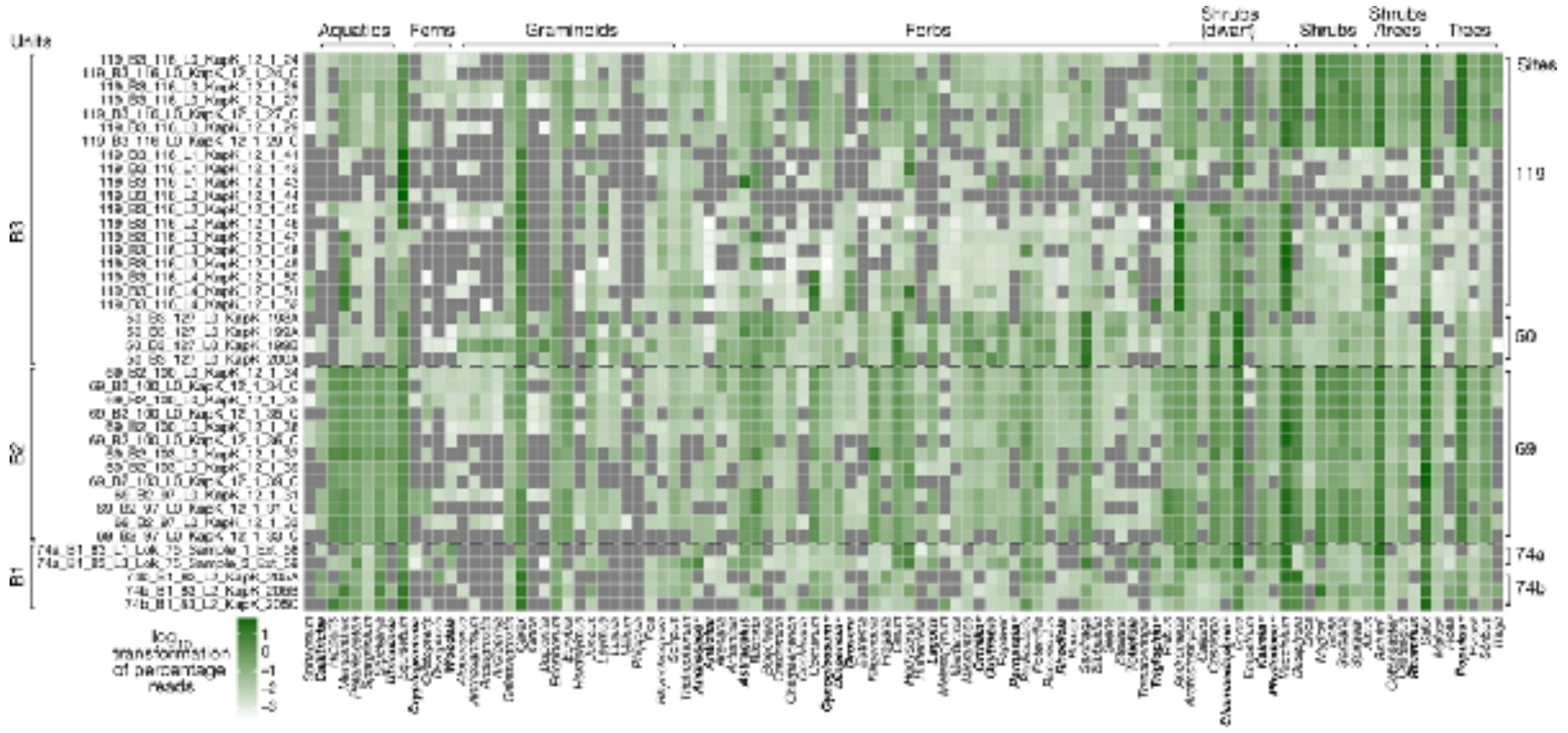


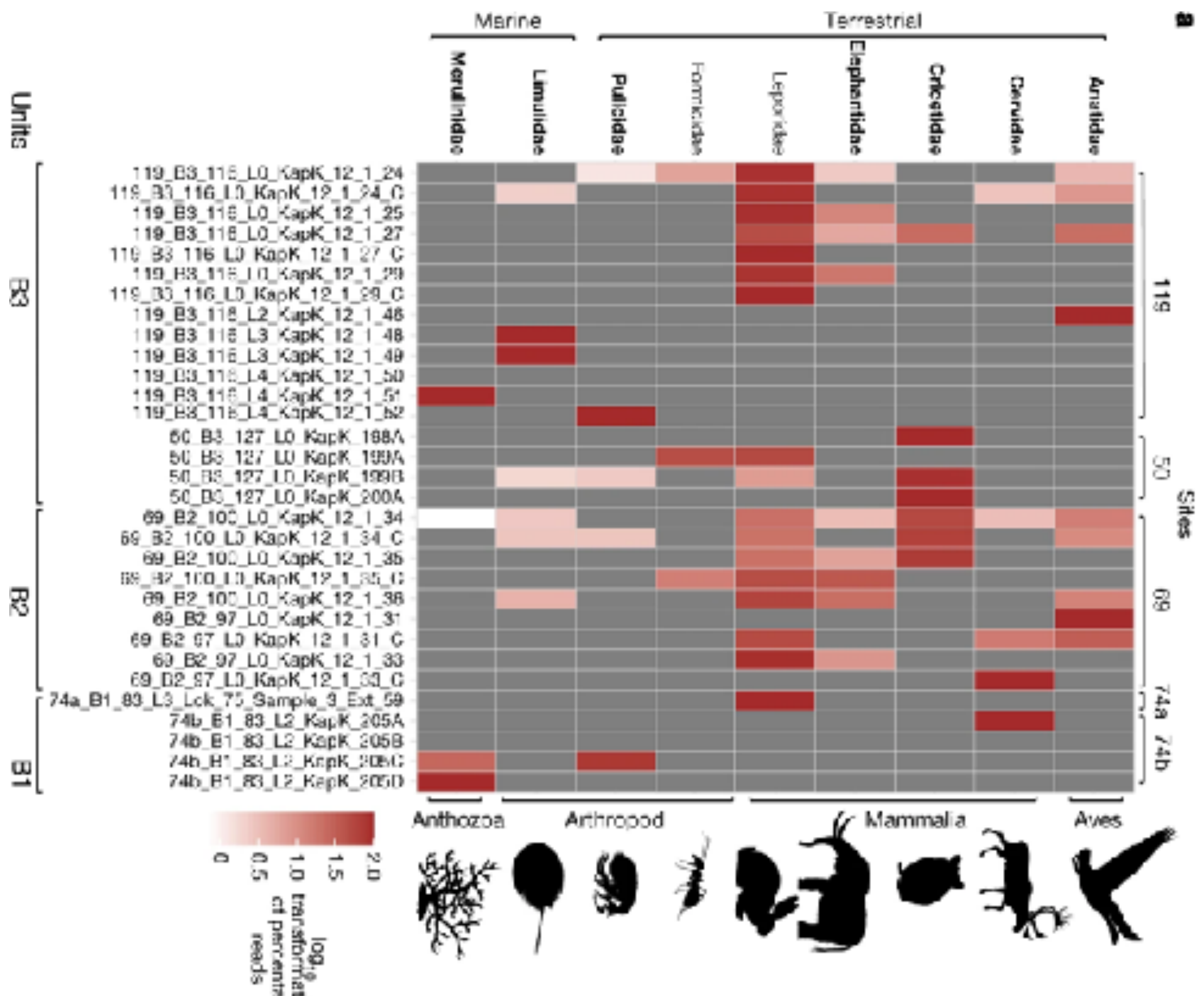
Nature **612**, 283–291 (2022)

2-million-year-old ecosystems

eDNA sequencing informs about the paleoclimate

- **Preservation of eDNA in frozen sediments?**
  - Simulations suggested that 50 bp long eDNA can survive in frozen sediments (-17C) over 2 million years
- **Total eDNA sequencing and enrichment-based protocols** resulted in 2.8B short sequence (>30 bp)
- **Species quantification** based on eDNA alignment against 1,500 reference genomes
- **Reconstructed ecosystem**
  - **Plants**: mixed vegetation with poplar, birch, thuja trees
  - **Animals**: mastodons, reindeers, geese, rodents
  - **Marine species**: horseshoe crab and green algae
- **Reconstructed ecosystem has no modern analogue** (eg, 40% of detected plants no longer grow in Greenland)

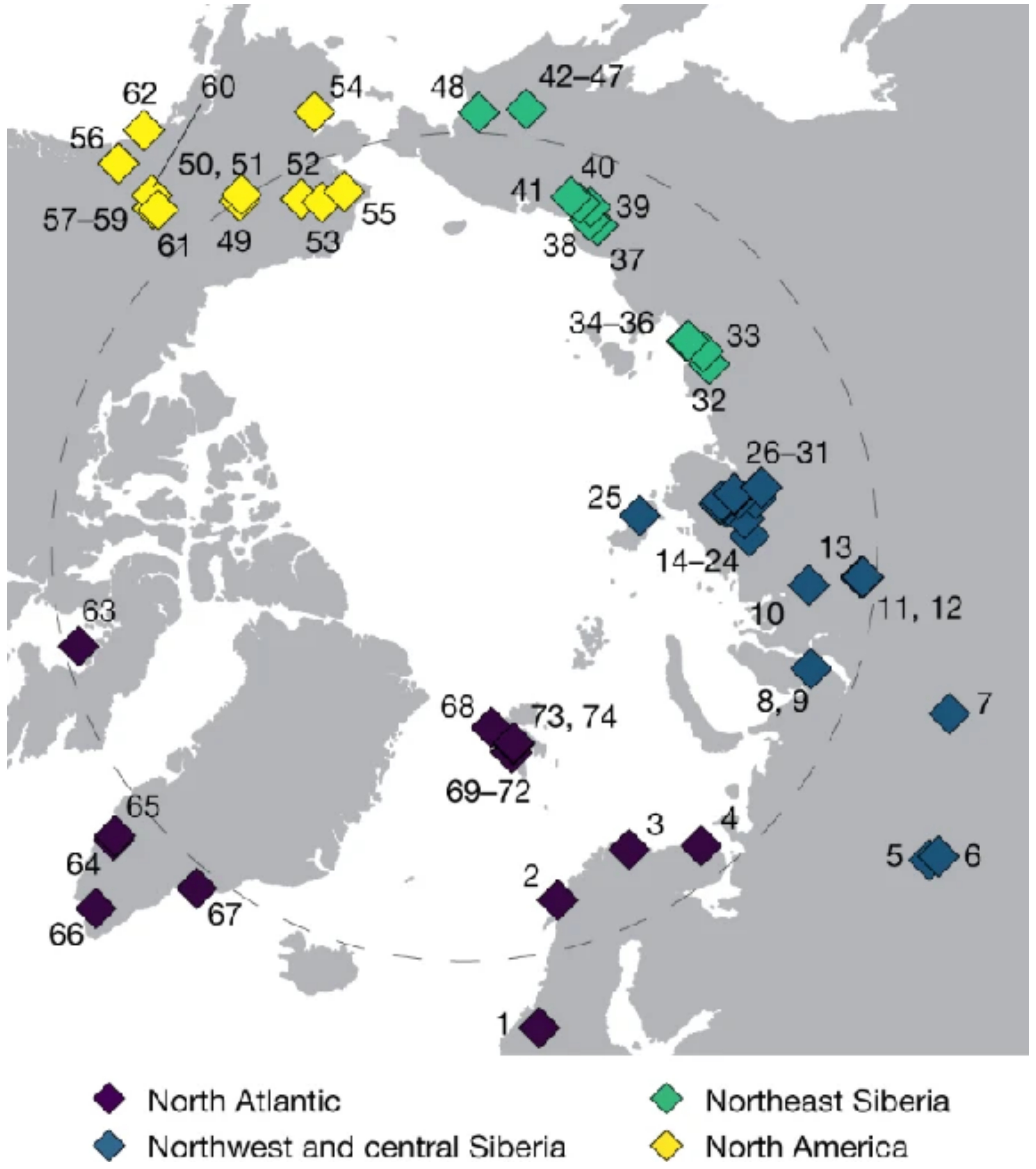




Nature **612**, 283–291 (2022).

Arctic permafrost biodiversity during the past 50,000 years

- Over **500 sediment samples** from **70 sites,** covering 50,000 year, and over 10 billion eDNA sequences
- **Steppe ecosystems** declined due to climate change (warming and rain) and less due to human activities (human overhunting likely not a main driver in the arctic circle)
- **Ancient plants** showed a shift from protein-rich dry grasslands (steppe tundra) to wetter tundras (mosses, shrubs) about 10,000 years ago. Mammoths depended likely more on nutrient-rich grasses
- **Mammoths** likely survived in isolated Arctic pockets for thousands of years after thought to be extinct (“climate refugia”)
- **Past global warming** shaped ecosystems gradually over thousands of years (vs rapid changes nowadays)



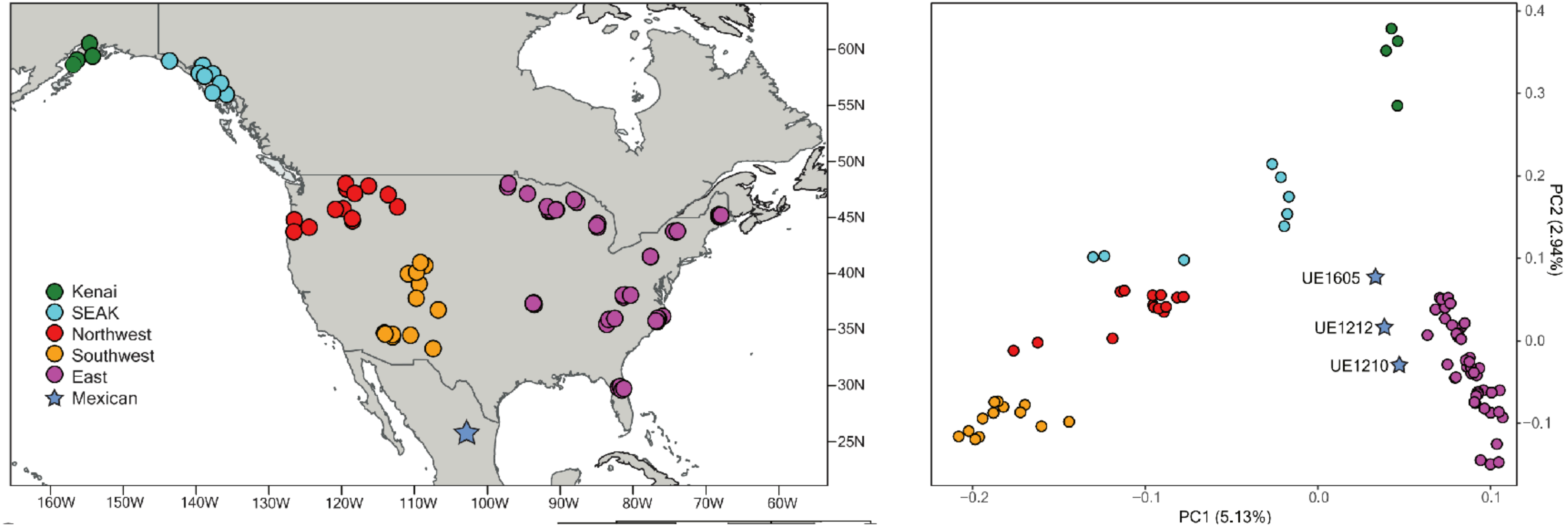
Nature **600**, 86–92 (2021).

Population genomics based on ancient eDNA

Pedersen et al., 2021, Current Biology 31, 2728–2736

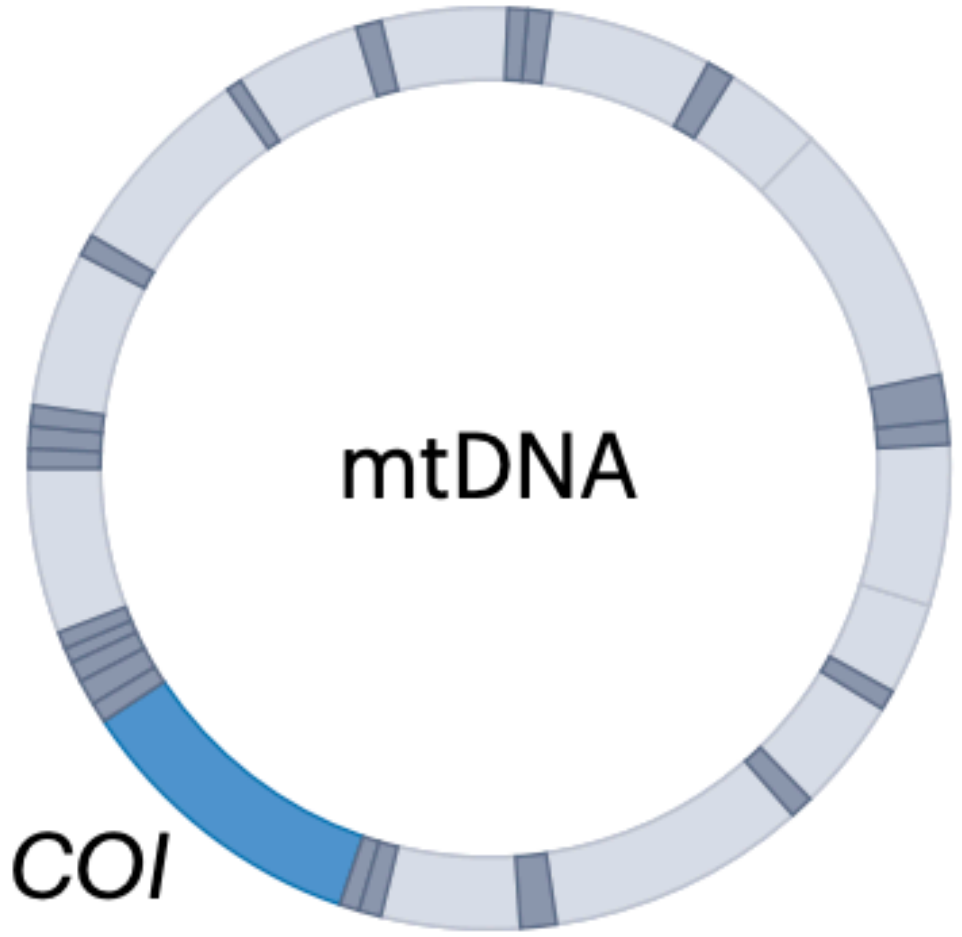
Population genomics based on ancient eDNA

- Genomes from ancient American black bears recovered from eDNA in cave sediments
- Extinct giant short-faced bear (Arctodus simus) genome recovered from eDNA in cave sediments
- Ancient Mexican black bear population more closely related to present-day Eastern American black bears



Pedersen et al., 2021, Current Biology 31, 2728–2736

Global climate change and eDNA community monitoring

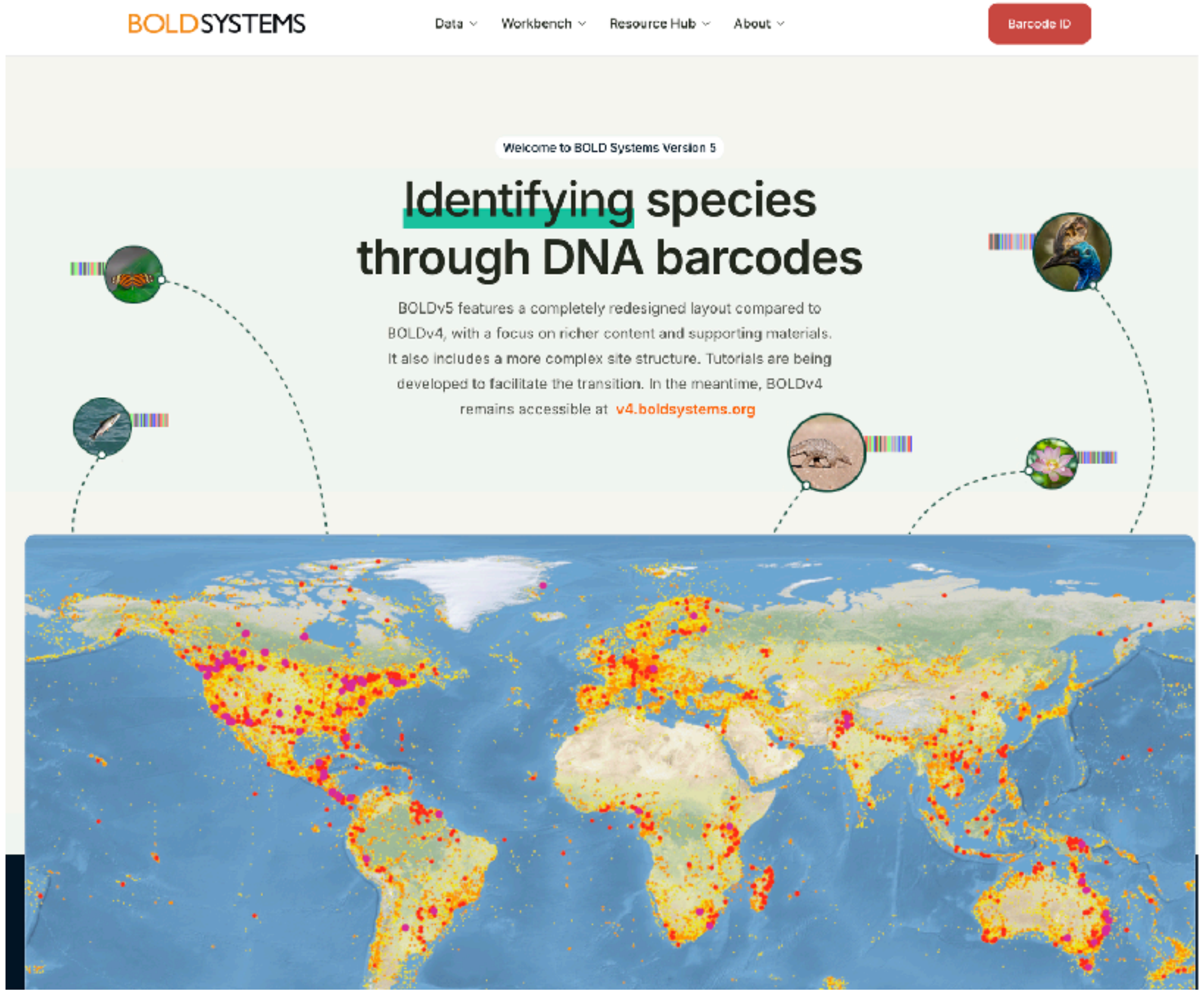


- eDNA sequencing of *barcode* regions to determine species composition of a sample
- Universal marker for species identification: a 658 bp fragment of the mitochondrial cytochrome C oxidase subunit I gene (cox1 or COI); other main barcode genes: ITS, matK, and rbcL
- Shifts in community composition over time (eg, invasive species, response to pollution, temperature changes)
- Integration of eDNA-based biodiversity, spatial information, and climate data enables predictions of habitat occupancy models under different global warming scenarios

Nature Reviews Genetics | Volume 25 | March 2024 | 165–183

Barcode of Life Data Systems (BOLD)

DNA marker for species identification (263k animals, 106k plants, 26k fungi & other)

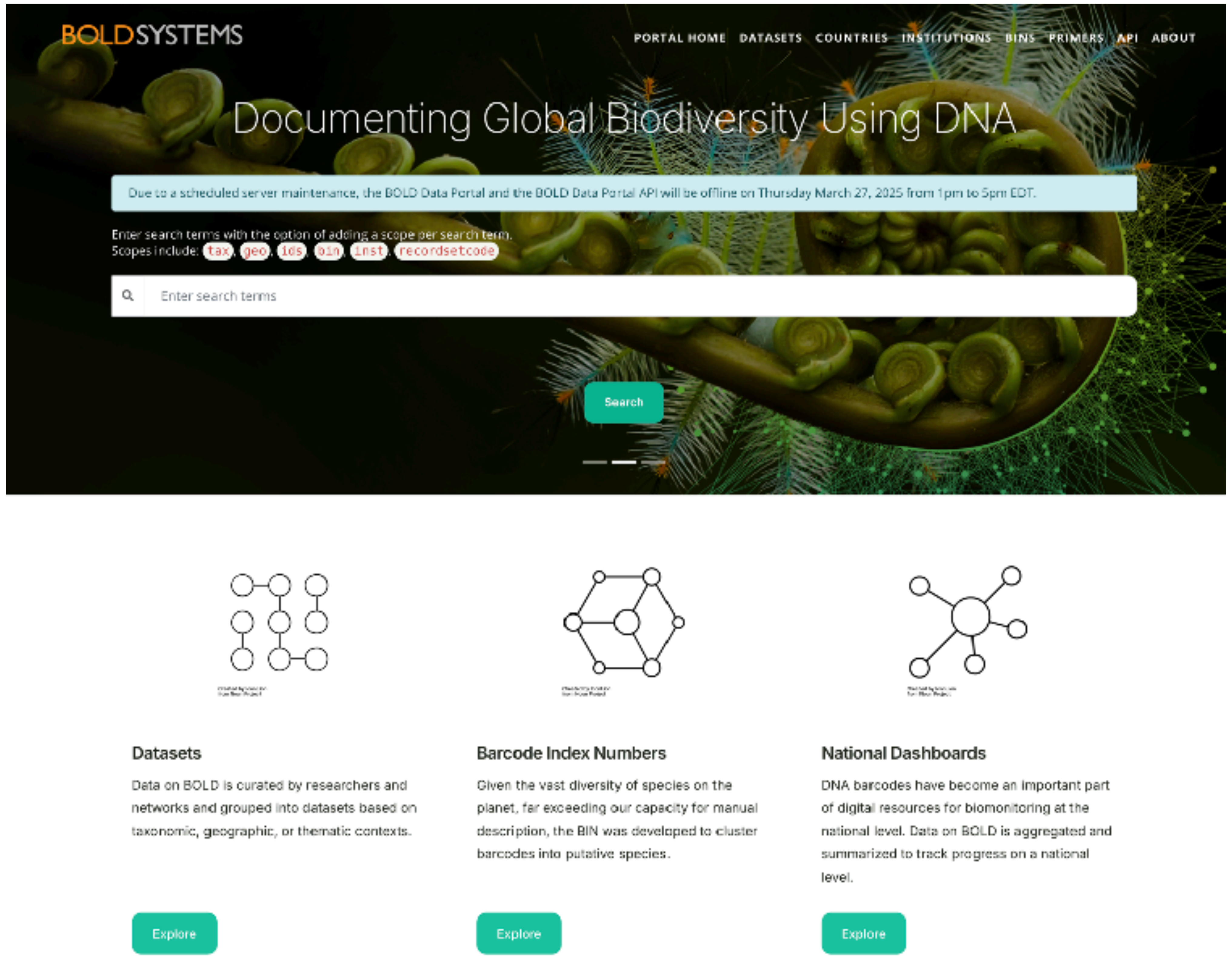


Welcome to BOLD Systems Version 5

Identifying species through DNA barcodes

BOLDv5 features a completely redesigned layout compared to BOLDv4, with a focus on richer content and supporting materials. It also includes a more complex site structure. Tutorials are being developed to facilitate the transition. In the meantime, BOLDv4 remains accessible at v4.boldsystems.org

https://boldsystems.org/



Documenting Global Biodiversity Using DNA

Due to a scheduled server maintenance, the BOLD Data Portal and the BOLD Data Portal API will be offline on Thursday March 27, 2025 from 1pm to 5pm EDT.

Enter search terms with the option of adding a scope per search term.
Scopes include: `tax`, `geo`, `ids`, `bin`, `inst`, `recordsetcode`

Datasets

Data on BOLD is curated by researchers and networks and grouped into datasets based on taxonomic, geographic, or thematic contexts.

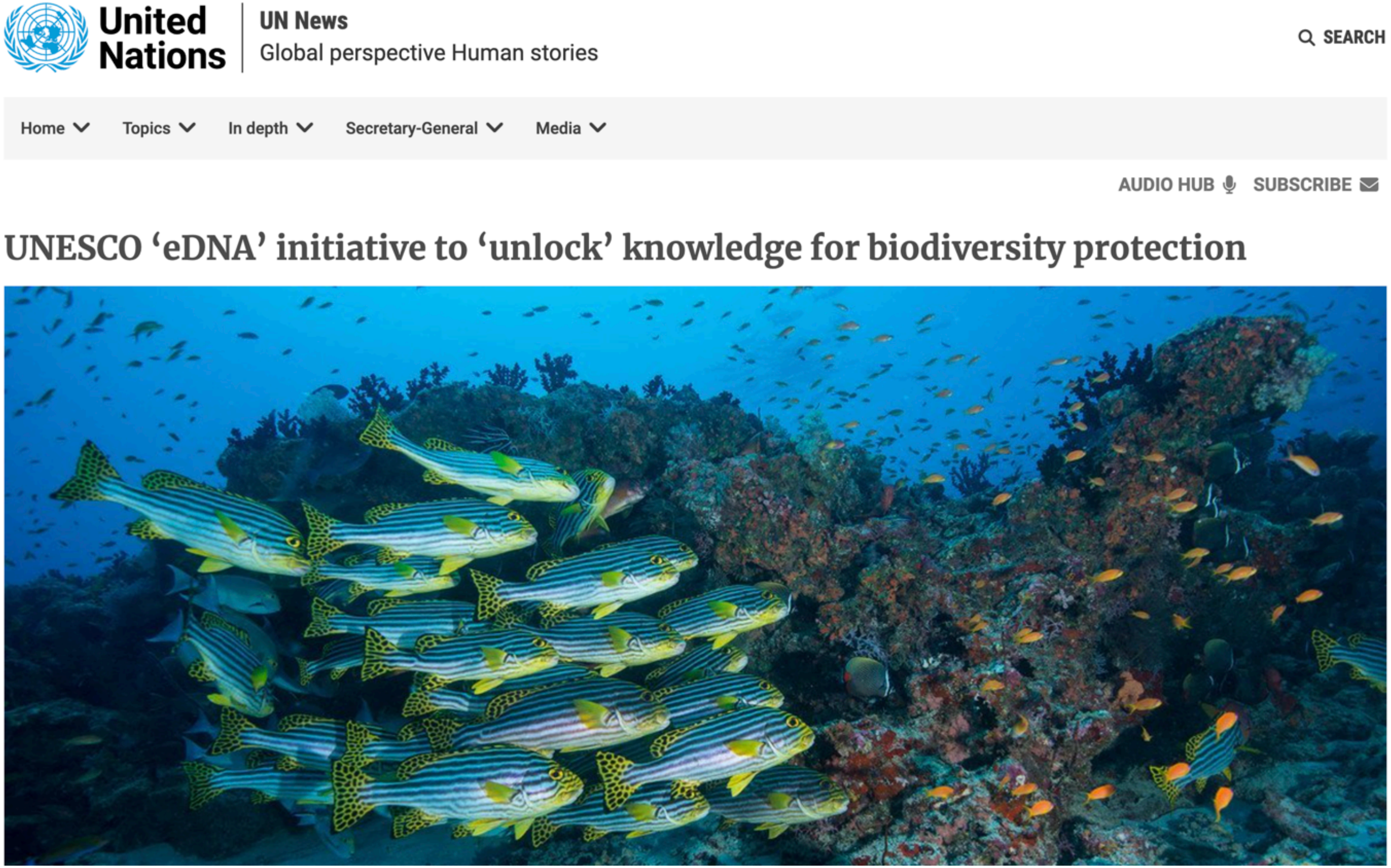
Barcode Index Numbers

Given the vast diversity of species on the planet, far exceeding our capacity for manual description, the BIN was developed to cluster barcodes into putative species.

National Dashboards

DNA barcodes have become an important part of digital resources for biomonitoring at the national level. Data on BOLD is aggregated and summarized to track progress on a national level.

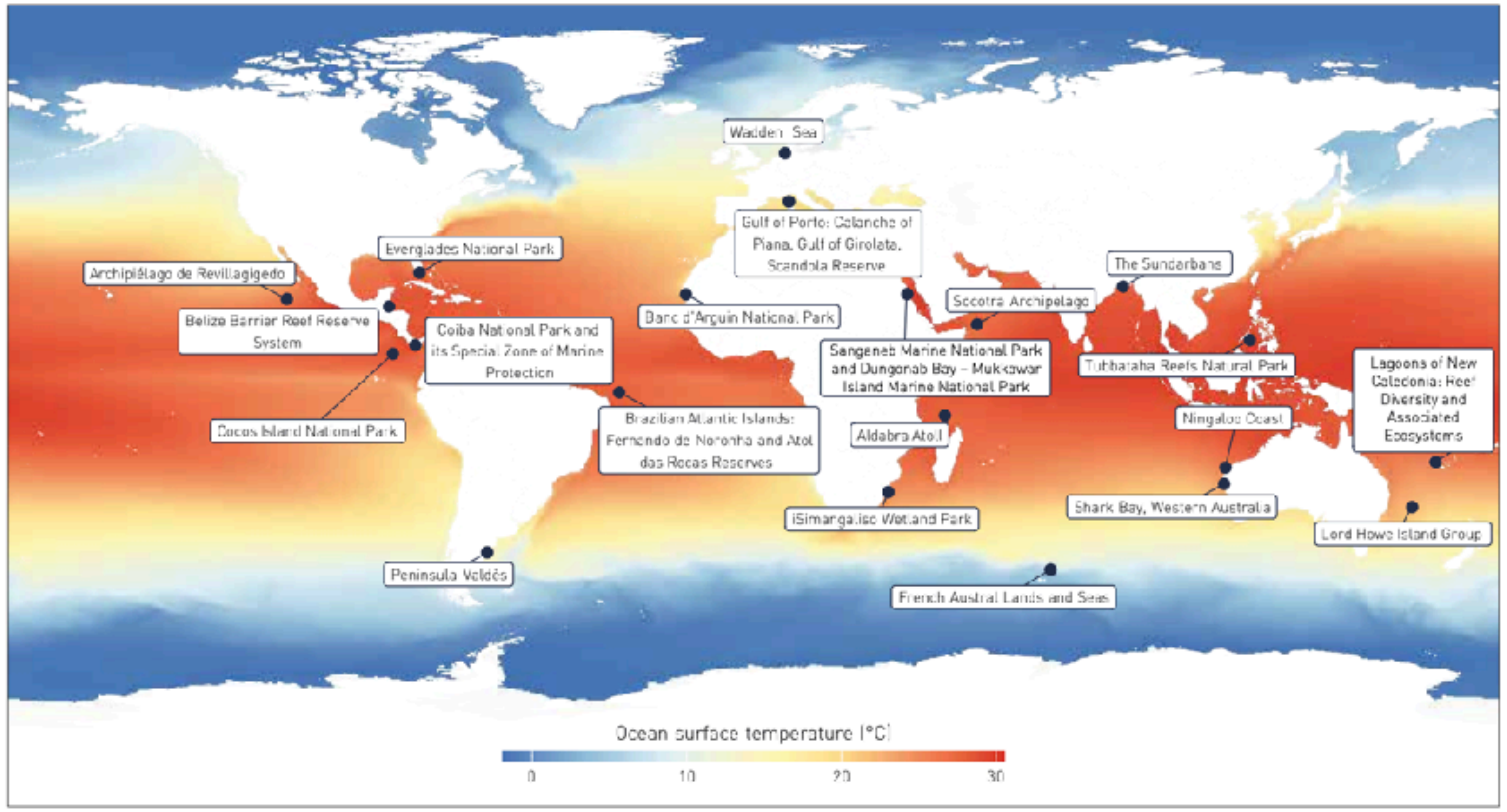
https://portal.boldsystems.org/



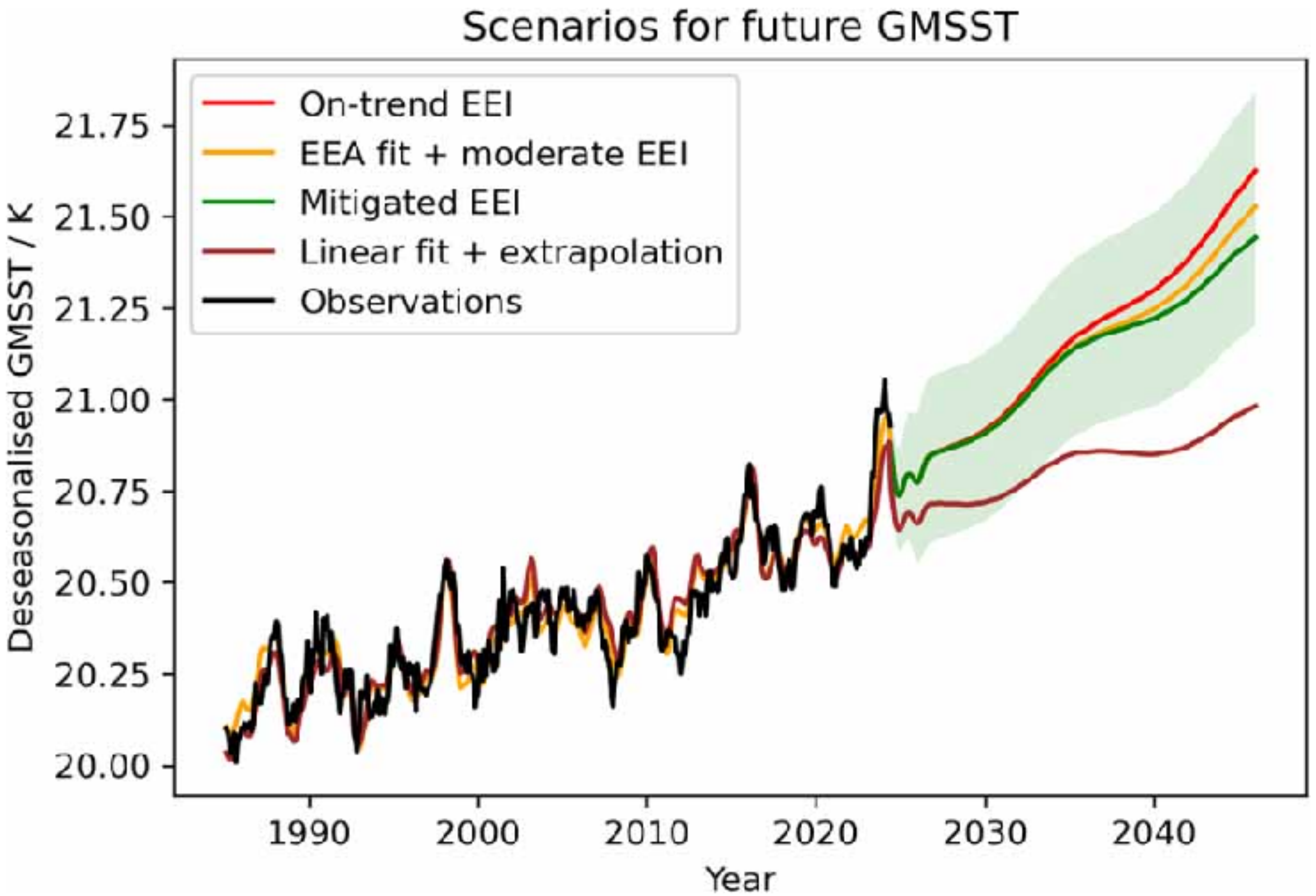
UNESCO ‘eDNA’ initiative to ‘unlock’ knowledge for biodiversity protection

21 UNESCO World Heritage marine sites participate in UNESCO eDNA Expeditions

World Heritage marine sites



Ocean temperatures are rising and hit records in 2023/2024



Merchant, C. J., Allan, R. P., & Embury, O. (2025). Environmental Research Letters, 20(2), 024037

Citizen-Science Portrait of Ocean Biodiversity

- UNESCO eDNA sampling protocols online available
- 250 volunteers identified 4,000 different marine species at World Heritage marine sites
- 120 threatened species from the IUCN RedList (75 vulnerable, 29 endangered, 16 critically endangered)
- A single campaign recorded 10-20% of local faunas (according to UNESCO-IOC OBIS and GBIF databases)
- Discovery of species with Outstanding Universal Value (OUV) status (eg, southern right whale in Argentina)
- Discovery of invasive species such as lionfish in the Belize Barrier Reef Reserve System

unesco

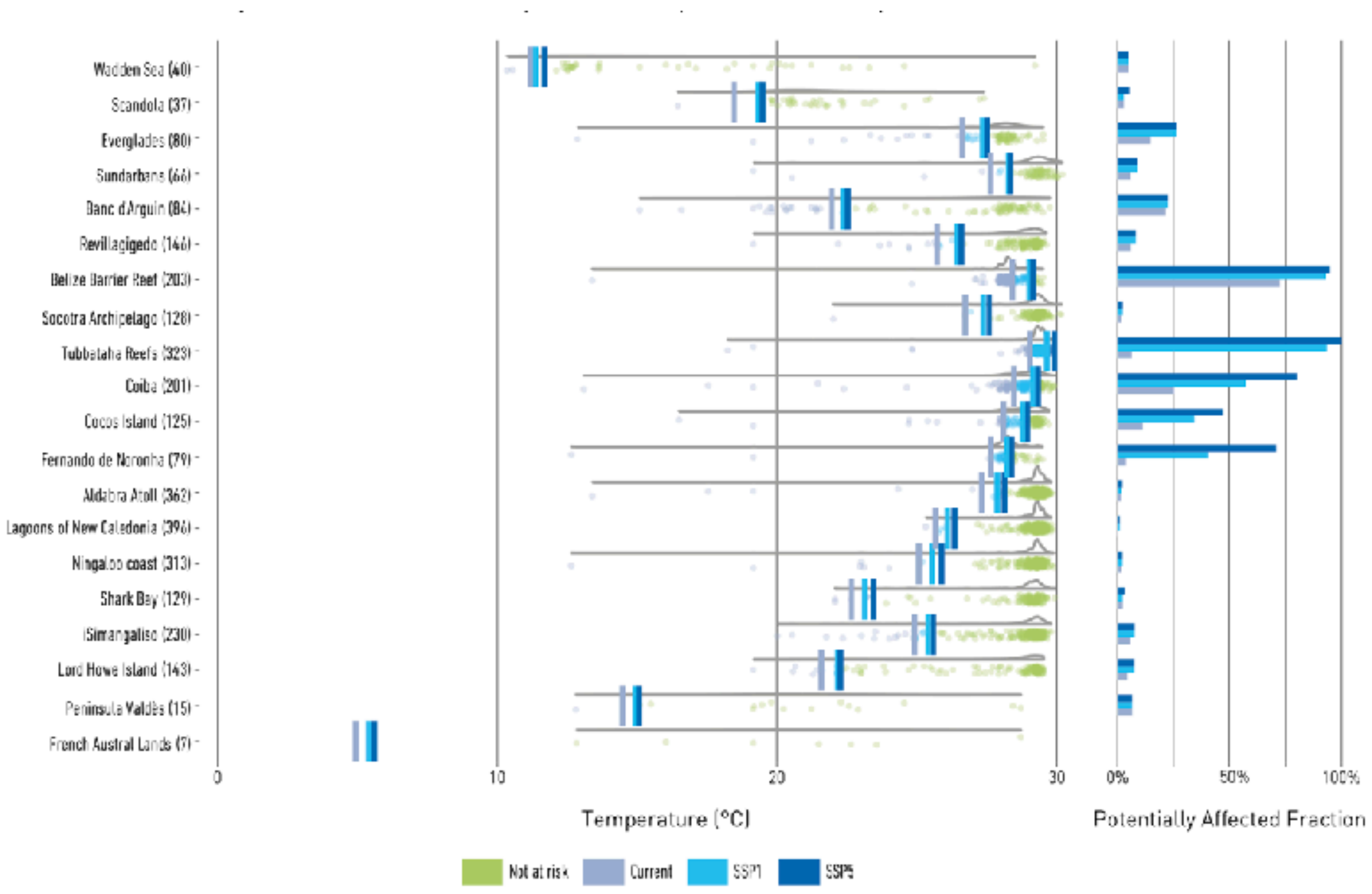
eDNA Expeditions

eDNA sampling protocol

https://www.doi.org/10.58337/CBXU3518

https://www.unesco.org/sites/default/files/medias/fichiers/2024/12/eDNA_expeditions_protocols_EN.pdf

Effects of Climate Change on Biodiversity in World Heritage marine sites



Belize Barrier Reef Reserve System

- Track species at marine sites that could face temperatures beyond their known tolerance limits
- Under the warmest future climate scenario (SSP5-8.5) up to 25% to 100% of fish will be beyond their thermal ranges
- Risk highest for species in tropical and subtropical oceans since already at upper thermal limits
- Currently unclear how species will cope with warming oceans (—> adaption, migration, extinction)

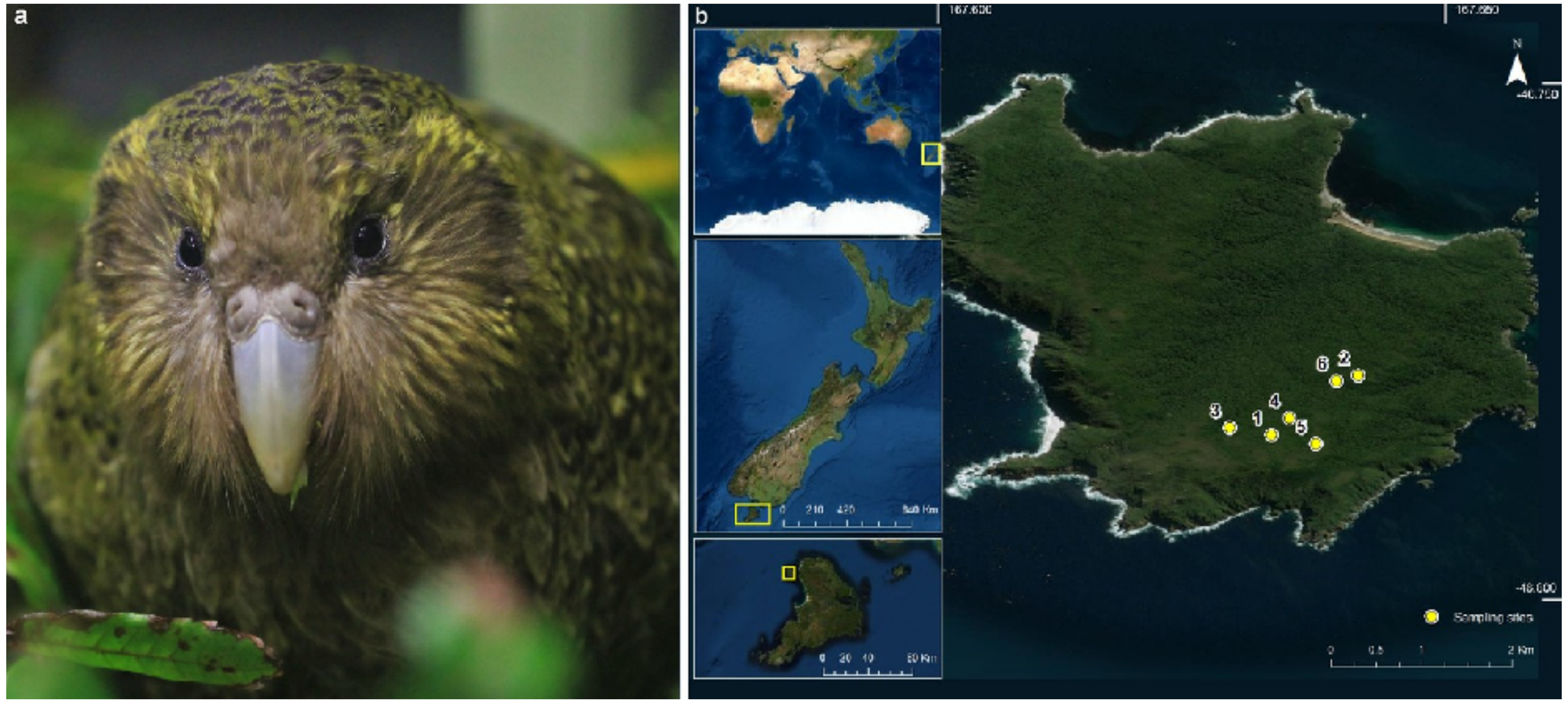
Detection and monitoring of invasive species with eDNA sequencing

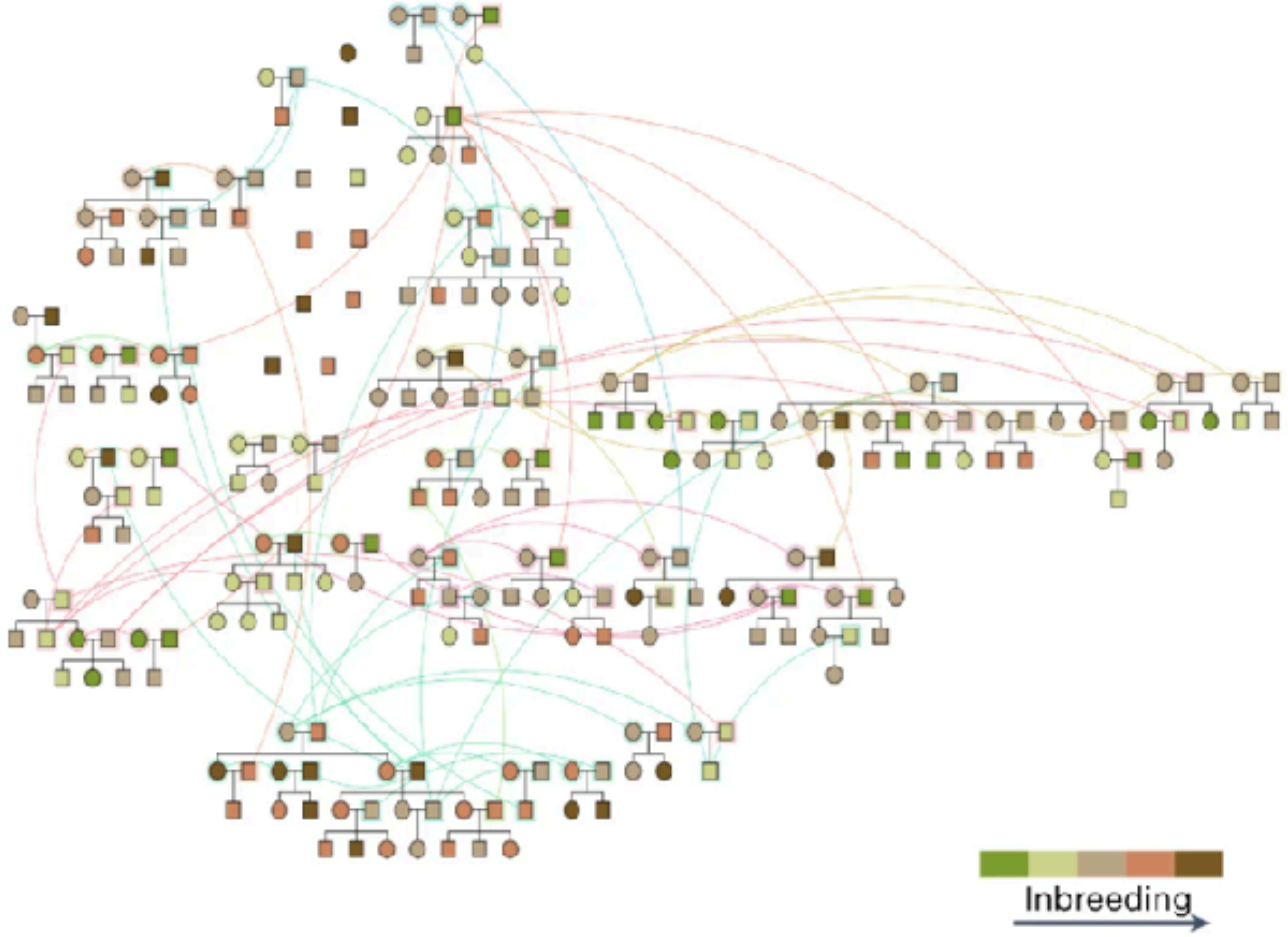
The Quagga mussel

- The Quagga mussel (*Dreissena bugensis*) originated from the Black Sea and was first discovered in Swiss waterways in 2014 based on eDNA sequencing and later confirmed in Lac Leman in 2015
- It is one of the most aggressive invasive freshwater species in the Northern hemisphere (lifespan 3-5 years)
- Up to 15,000 Quagga individuals per m^2 in Lac Leman; each mussel is able to filter 1-2 liters of water per day (incl. phytoplankton, algae)
- Rapid spread and abundance affects the lake ecosystem, impacts fishing, bathing, drinking water production, and lake water heating systems (pipe clogging)
- Increasing maintenance costs (several million CHF/year) due to damage caused by the Quagga mussel
- Likely impossible to eradicate —> monitor spread in unaffected lakes

https://www.cipel.org/en/invasion-quagga/

eDNA-based real-time genomic monitoring of endangered species

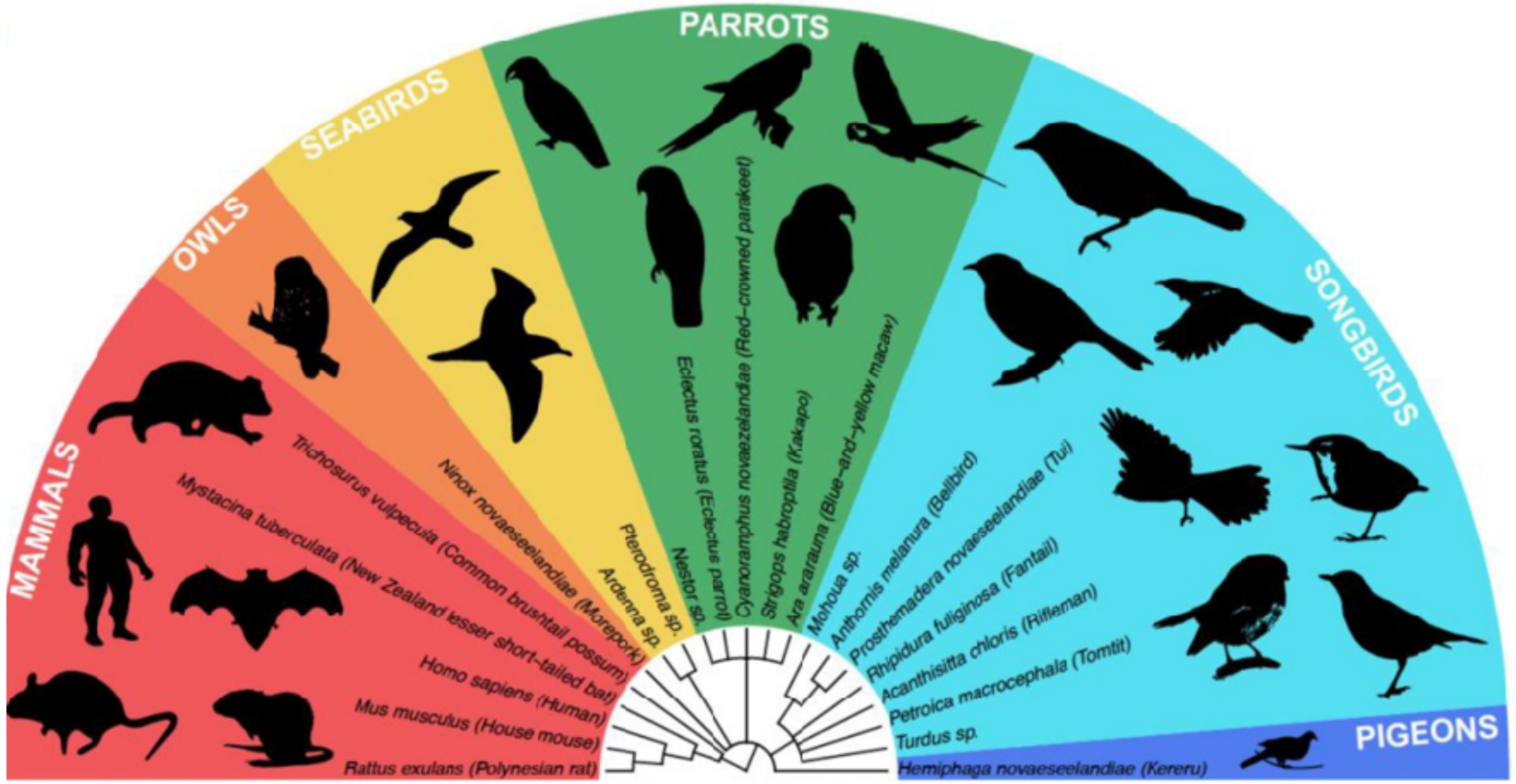




- Kākāpō, a ground-dwelling parrot that is endemic to New Zealand
- Historically important to indigenous people of New Zealand
- Only flightless parrot, heaviest parrot, and likely longest-living bird (up to 100 years)
- Critically endangered with 250 living animals
- Likely only species with genomes for the entire population

eLife 12:RP84553 (2023)
Nat Ecol Evol **7**, 1693–1705 (2023)

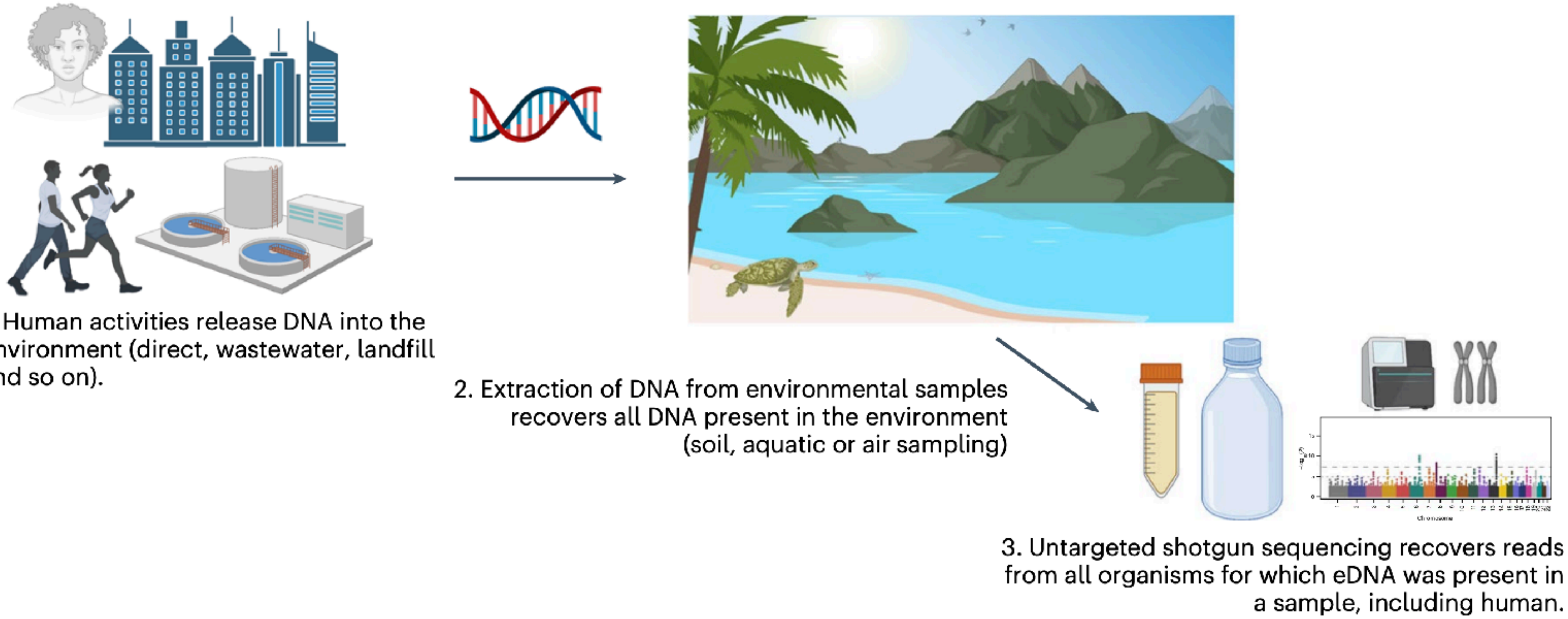
eDNA-based real-time genomic monitoring of Kākāpō parrots



Non-invasive real-time genomic monitoring of the critically endangered kākāpō *eLife* (2023) 12:RP84553.

- Detection of rare/endangered species (eg, parrots) using eDNA from soil samples
- Real-time sequencing of 12 S rRNA combined with a high-quality reference genome
- Identify individual animals from eDNA in soil samples
- Novel approach to help conservation efforts without the need to capture animals
- Monitor species distribution and genetic diversity of inbreed species

Legal and ethical considerations of eDNA studies



1. Human activities release DNA into the environment (direct, wastewater, landfill and so on).

2. Extraction of DNA from environmental samples recovers all DNA present in the environment (soil, aquatic or air sampling)

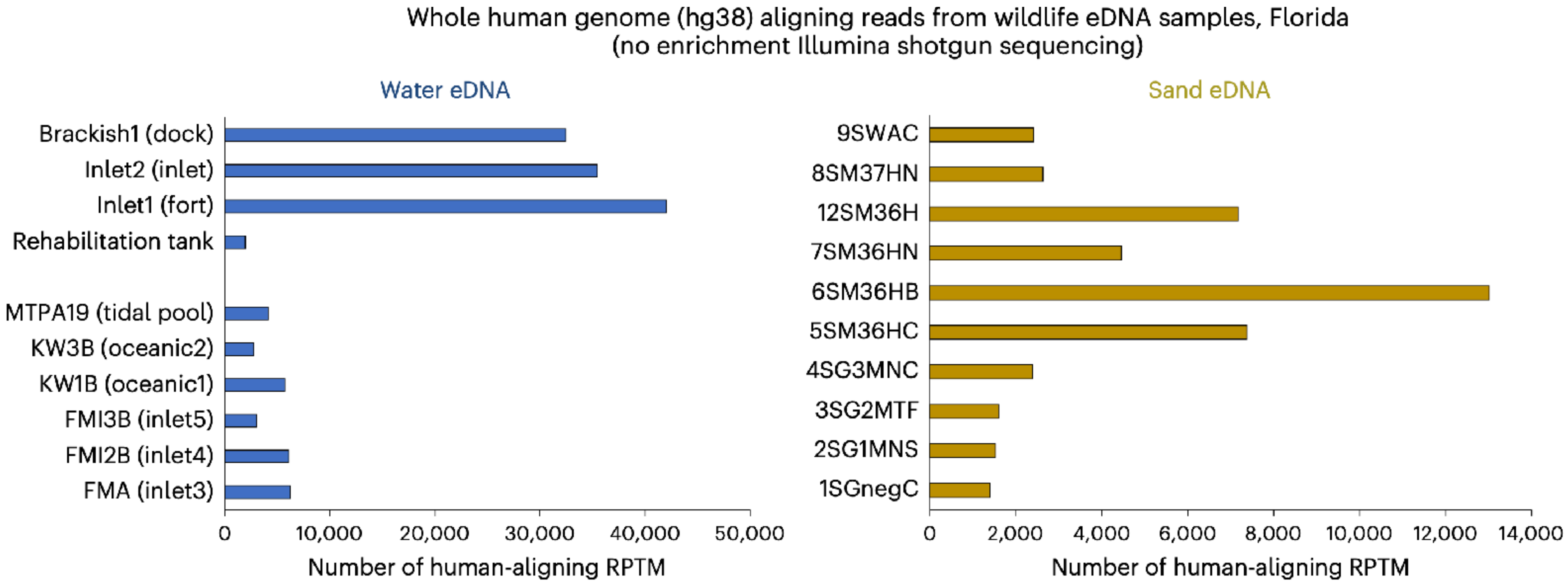
3. Untargeted shotgun sequencing recovers reads from all organisms for which eDNA was present in a sample, including human.

- Human DNA in the environment is omni-present
- Policies on eDNA collection and use are needed

Nature Ecology & Evolution | Volume 7 | June 2023 | 873–888

BIO-312 Genomic solutions to sustainable development

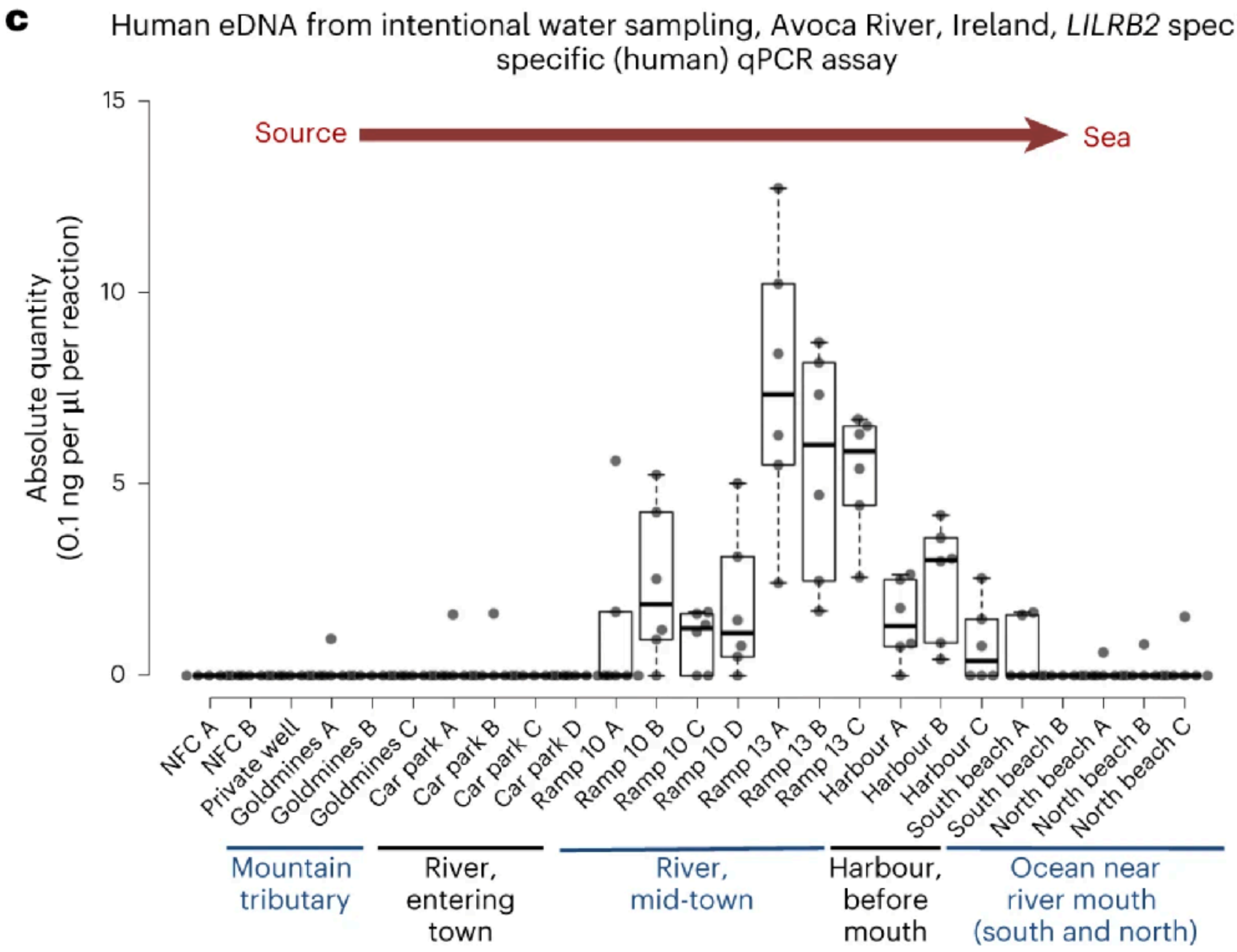
Human eDNA in water and sand samples



RPTM = reads per ten million total reads

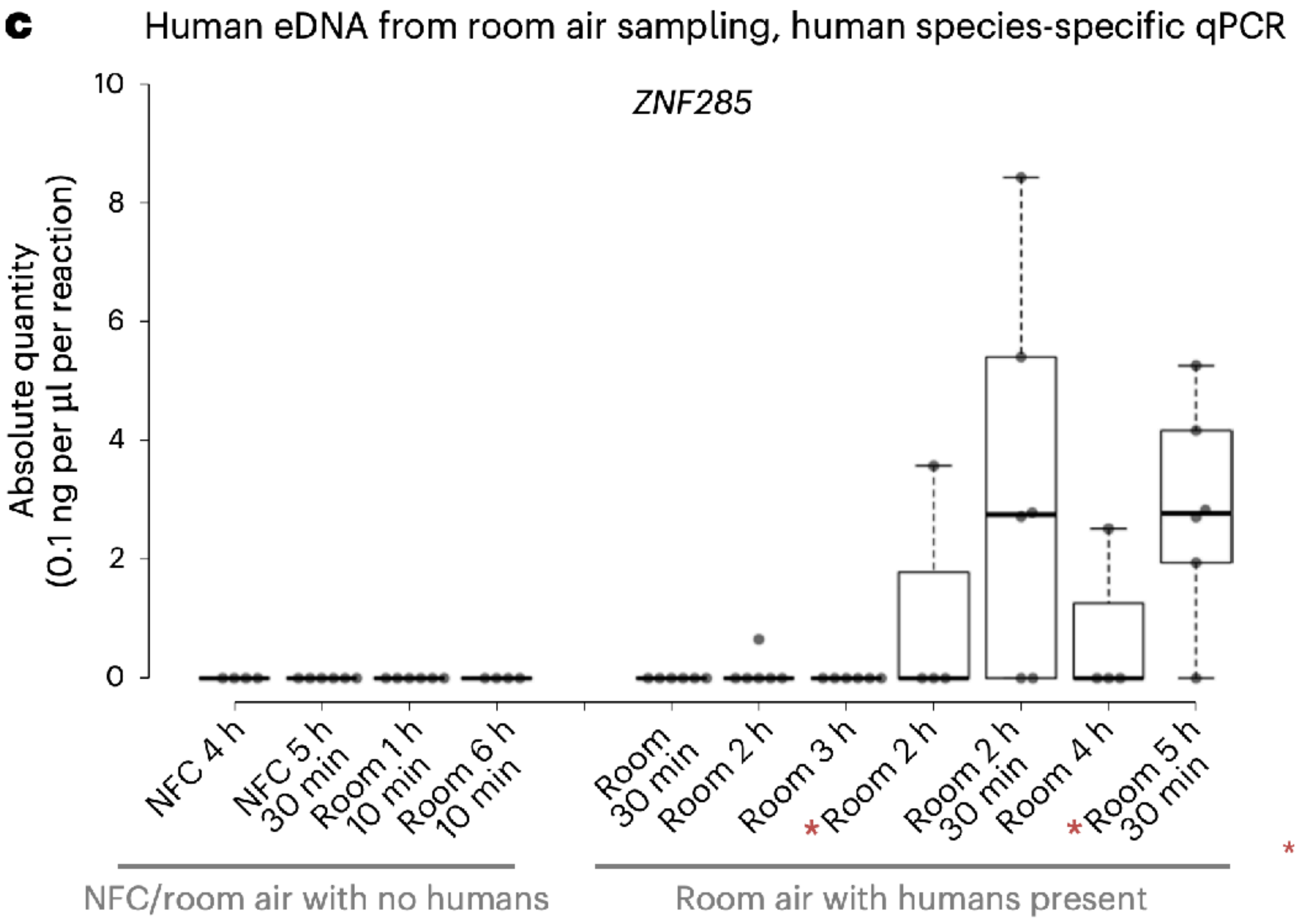
Nature Ecology & Evolution | Volume 7 | June 2023 | 873–888

Human eDNA sequencing reveals presence of human activities



Nature Ecology & Evolution | Volume 7 | June 2023 | 873–888

Human eDNA in air samples



* Nanopore sequenced

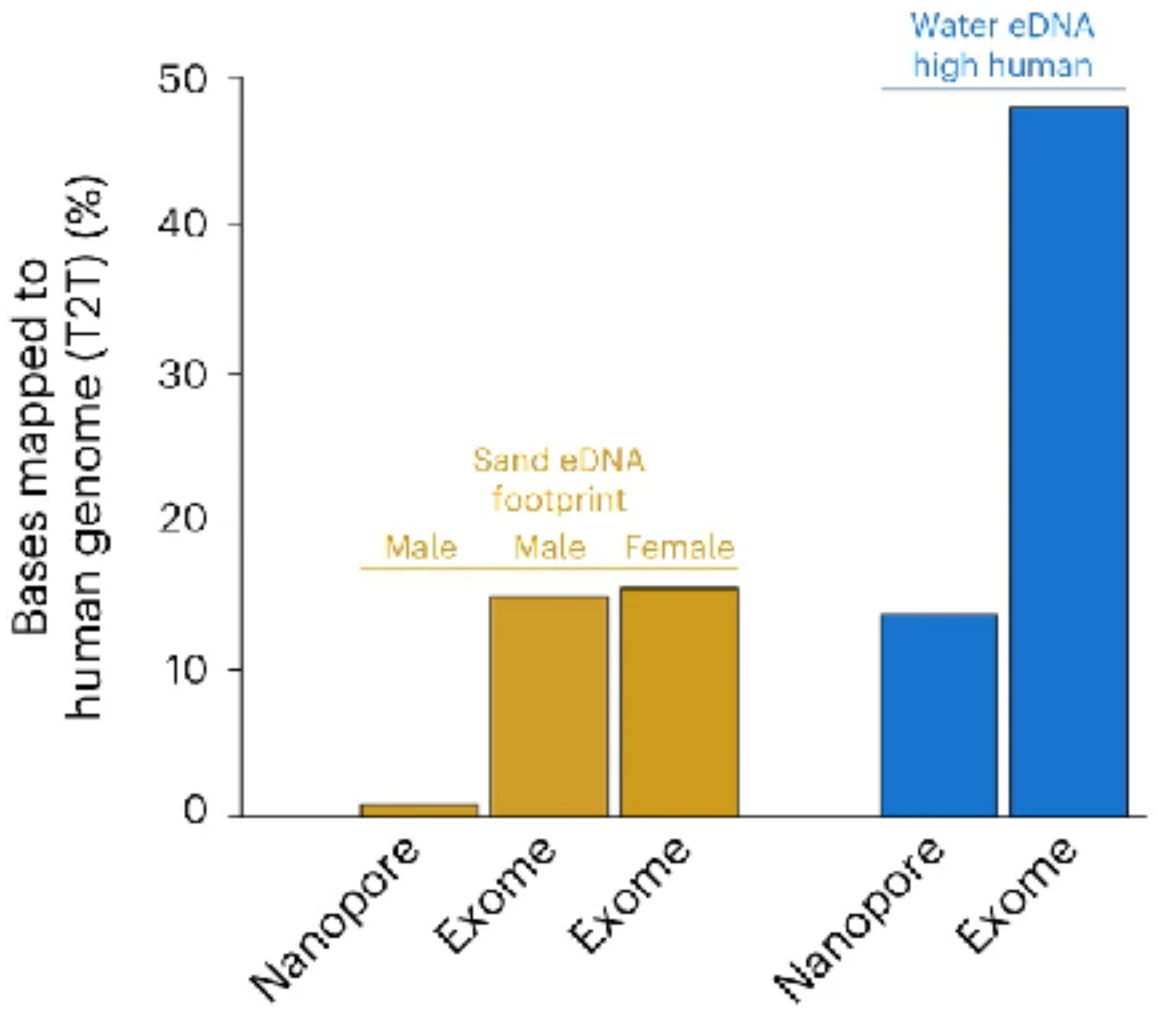
Nature Ecology & Evolution | Volume 7 | June 2023 | 873–888

Human population genomics based on eDNA

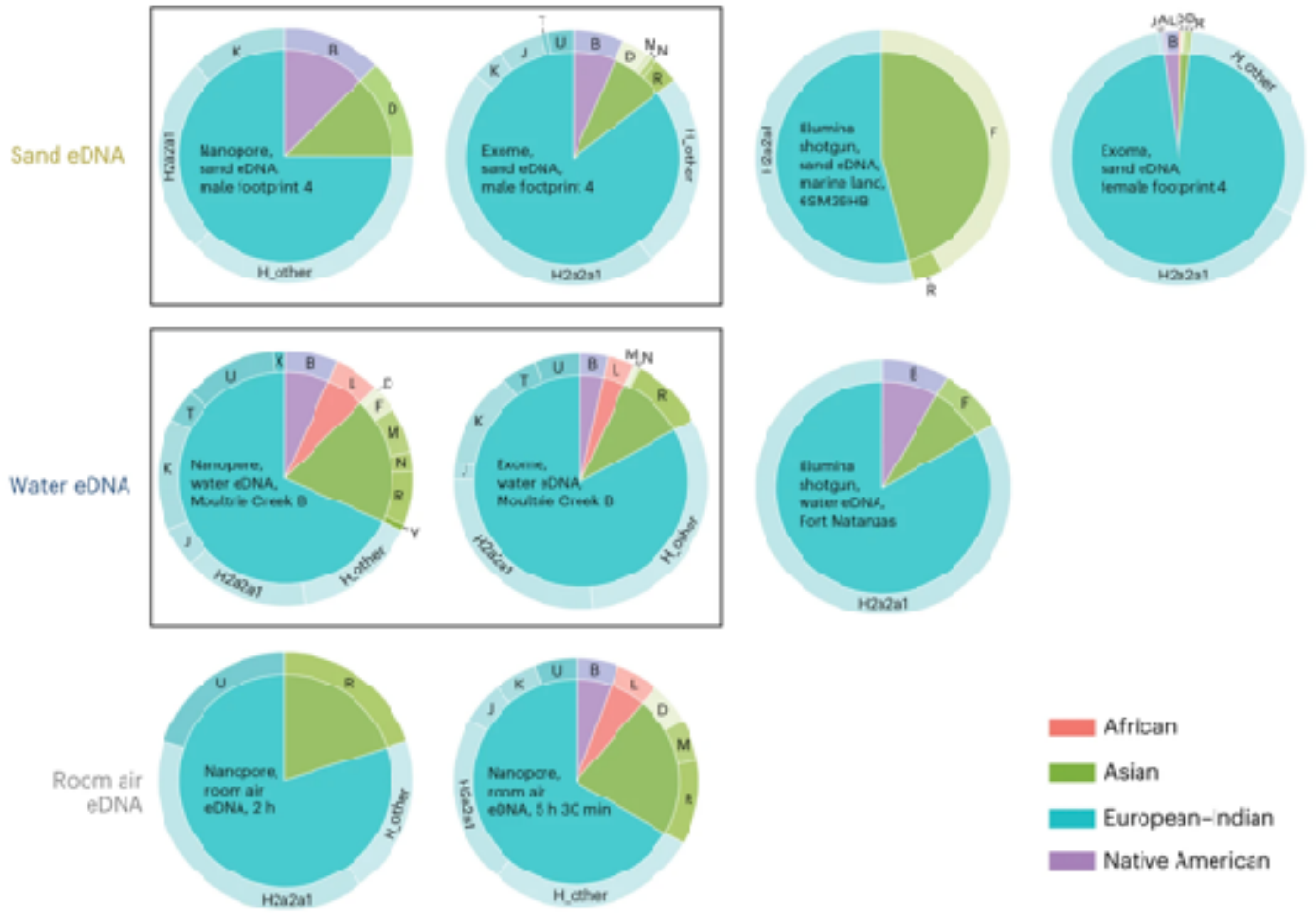
Complete human genomes from eDNA
average read length 1.5 kb
complete mtDNA genome

Longest continuous single human read per sample type
Air: 148,969 bp
Water: 120,909 bp
Sand: 39,229 bp (extraction method involves agitation of sand)

Efficient recovery of human eDNA based on human-specific DNA probes



Mitochondrial haplogroups and haplotypes in eDNA



Potential problematic implications of the capture of human genomic eDNA data

Unintended consequences

- Requirement of human-study-related ethical approvals for wildlife studies
- Lack of human subject consent/breach of privacy
- Public deposition of eDNA data including human genomic data
- Inadvertent location tracking
- Inadvertent genome harvesting

Nature Ecology & Evolution | Volume 7 | June 2023 | 873–888

Potential problematic implications of the capture of human genomic eDNA data

Potential malicious applications

- Genome harvesting—the ability to illegally/unethically harvest human genomic data from local populations/ethnic groups without their knowledge or consent
- Covert accumulation of human genetic data for malicious or commercial purposes (for example, genomic surveillance or big-data-fuelled discovery)
- Genetic surveillance—individual tracking (similar to forensics)
- Genetic surveillance—unethical tracking/locating of ethnic groups/populations
- Bio-piracy of human genetic data from populations and countries (akin to flora/fauna genetic bio-piracy)

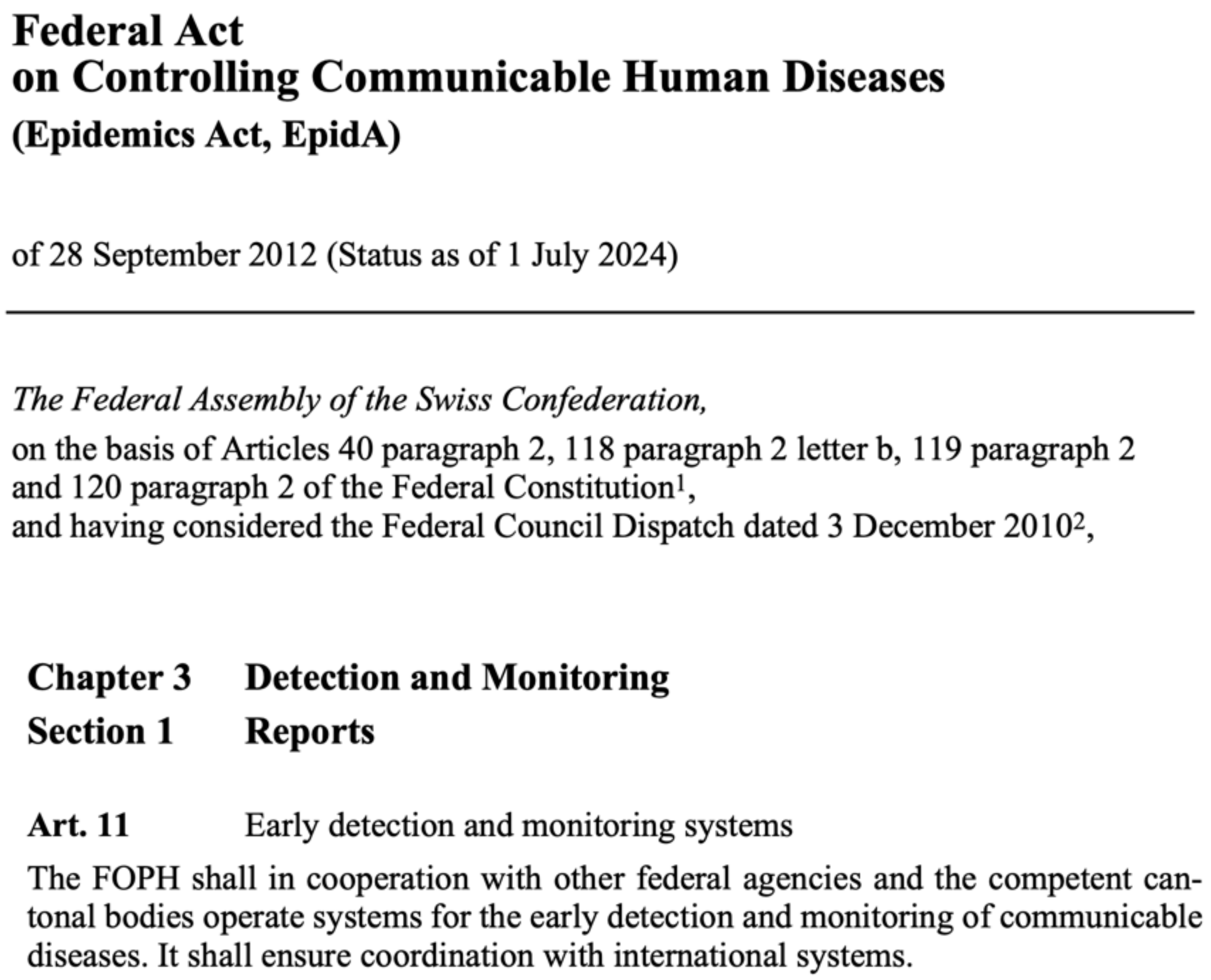
Nature Ecology & Evolution | Volume 7 | June 2023 | 873–888

Wastewater surveillance

Monitor the circulation of pathogens in populations

Photo: Eawag, Esther Miche

Wastewater surveillance in Switzerland



Federal Act
on Controlling Communicable Human Diseases
(Epidemics Act, EpidA)

of 28 September 2012 (Status as of 1 July 2024)

The Federal Assembly of the Swiss Confederation,

on the basis of Articles 40 paragraph 2, 118 paragraph 2 letter b, 119 paragraph 2 and 120 paragraph 2 of the Federal Constitution[1],
and having considered the Federal Council Dispatch dated 3 December 2010[2],

Chapter 3 **Detection and Monitoring**
Section 1 **Reports**

Art. 11 Early detection and monitoring systems

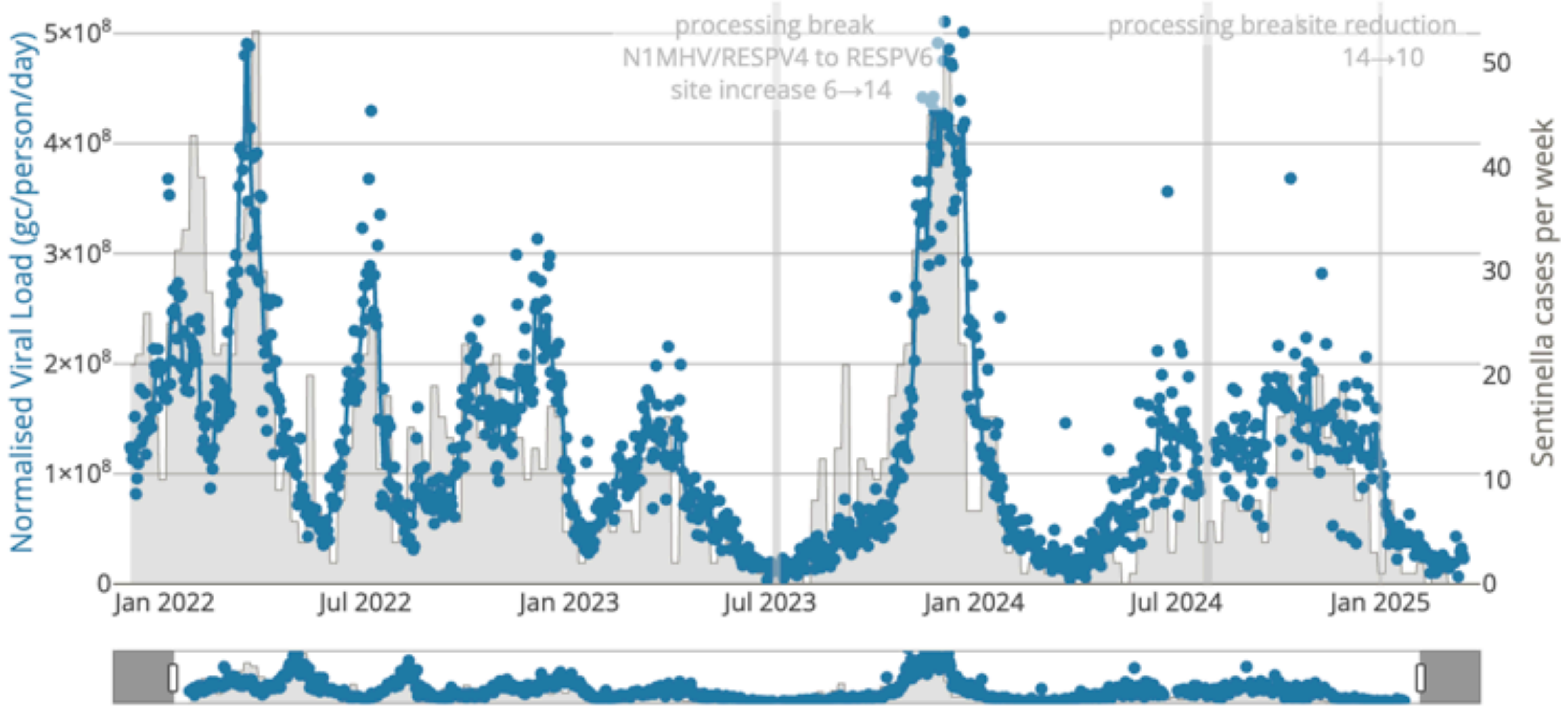
The FOPH shall in cooperation with other federal agencies and the competent cantonal bodies operate systems for the early detection and monitoring of communicable diseases. It shall ensure coordination with international systems.

Wastewater surveillance in Switzerland

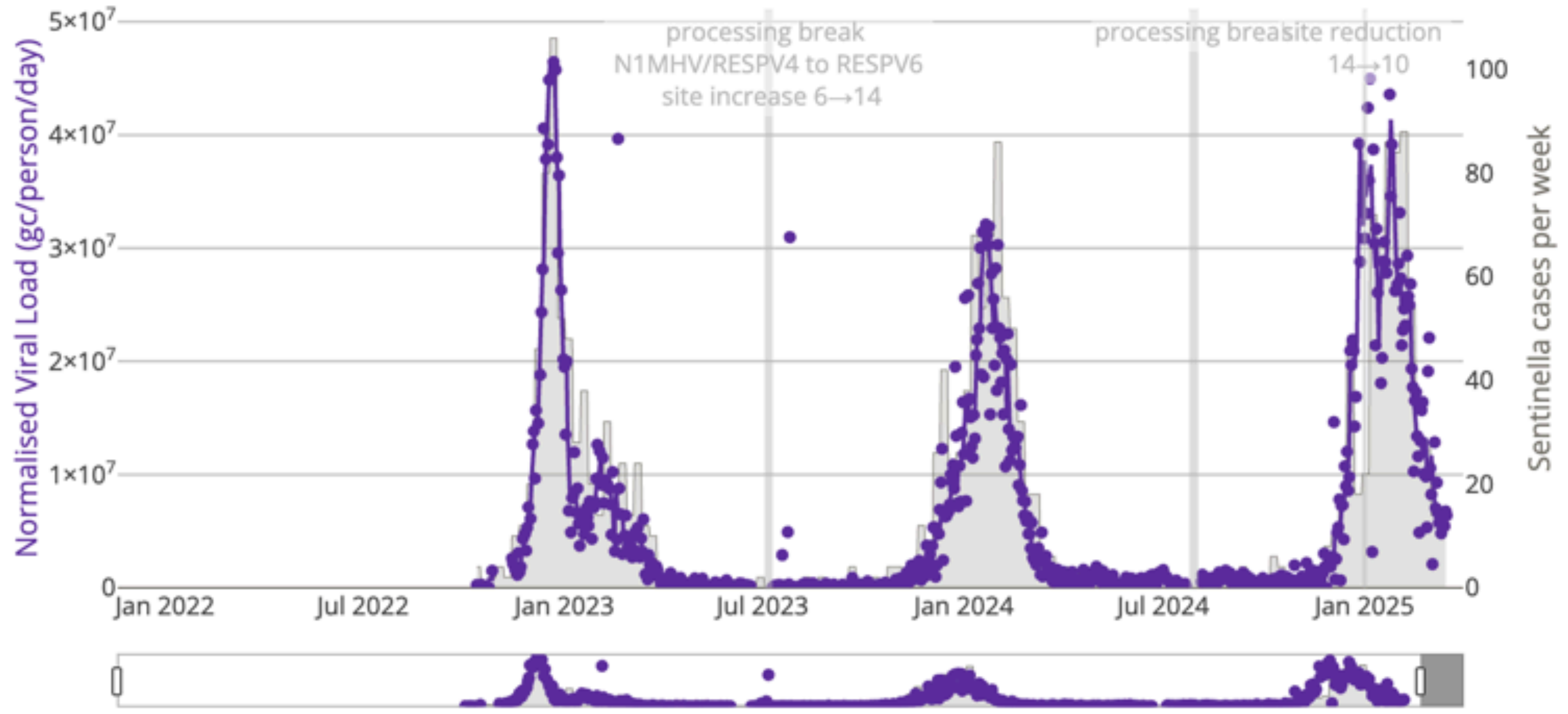
- Wastewater samples regularly taken from 14 wastewater treatment plants in Switzerland (since 2023)
- Multiple viruses are simultaneously quantified (SARS-CoV-2, Influenza A, Influence B, and RSV)
- Estimation of viral load (gene copies/day per 100,000 inhabitants) and R values (reproduction number)
- Live data available on https://wise.ethz.ch/

Wastewater surveillance of viruses in Switzerland

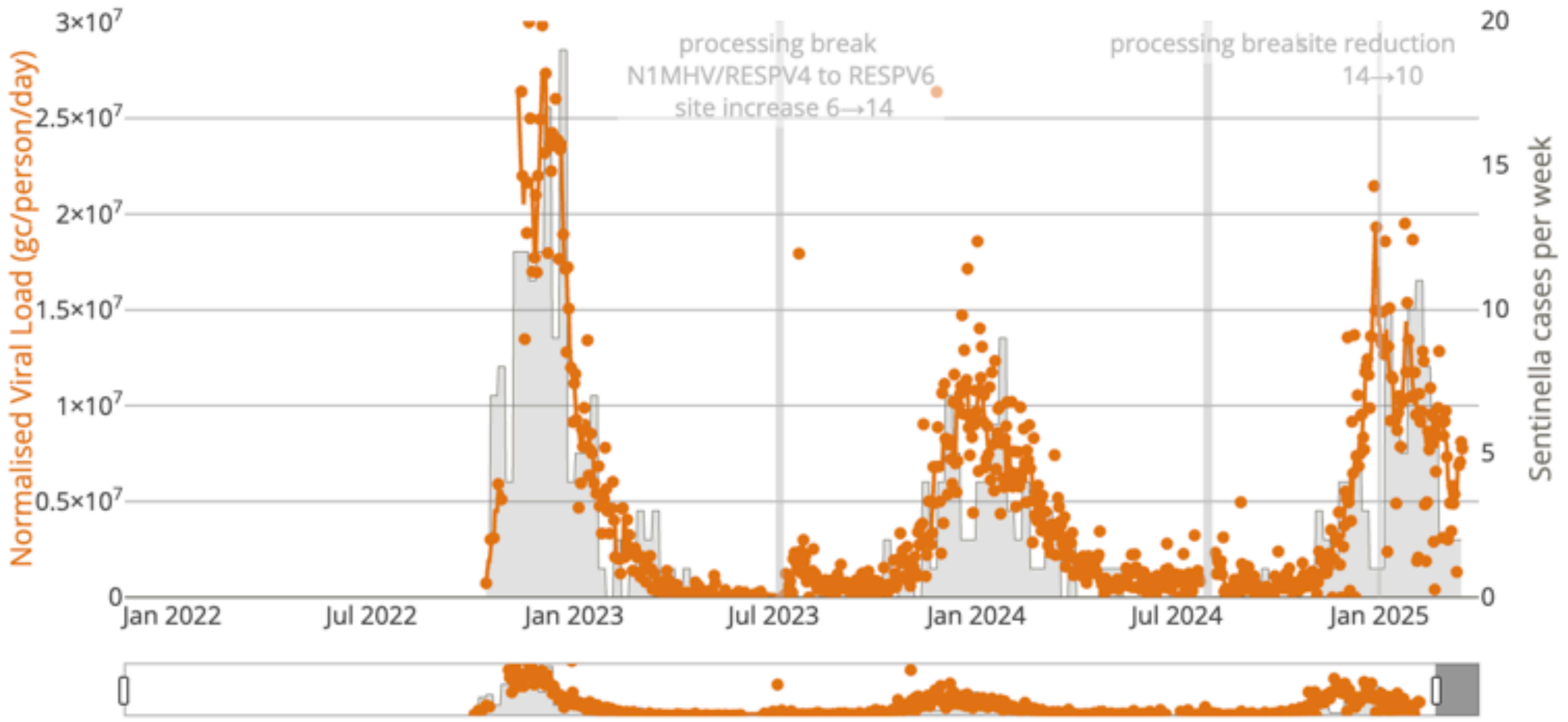
SARS-CoV-2 Viral Load



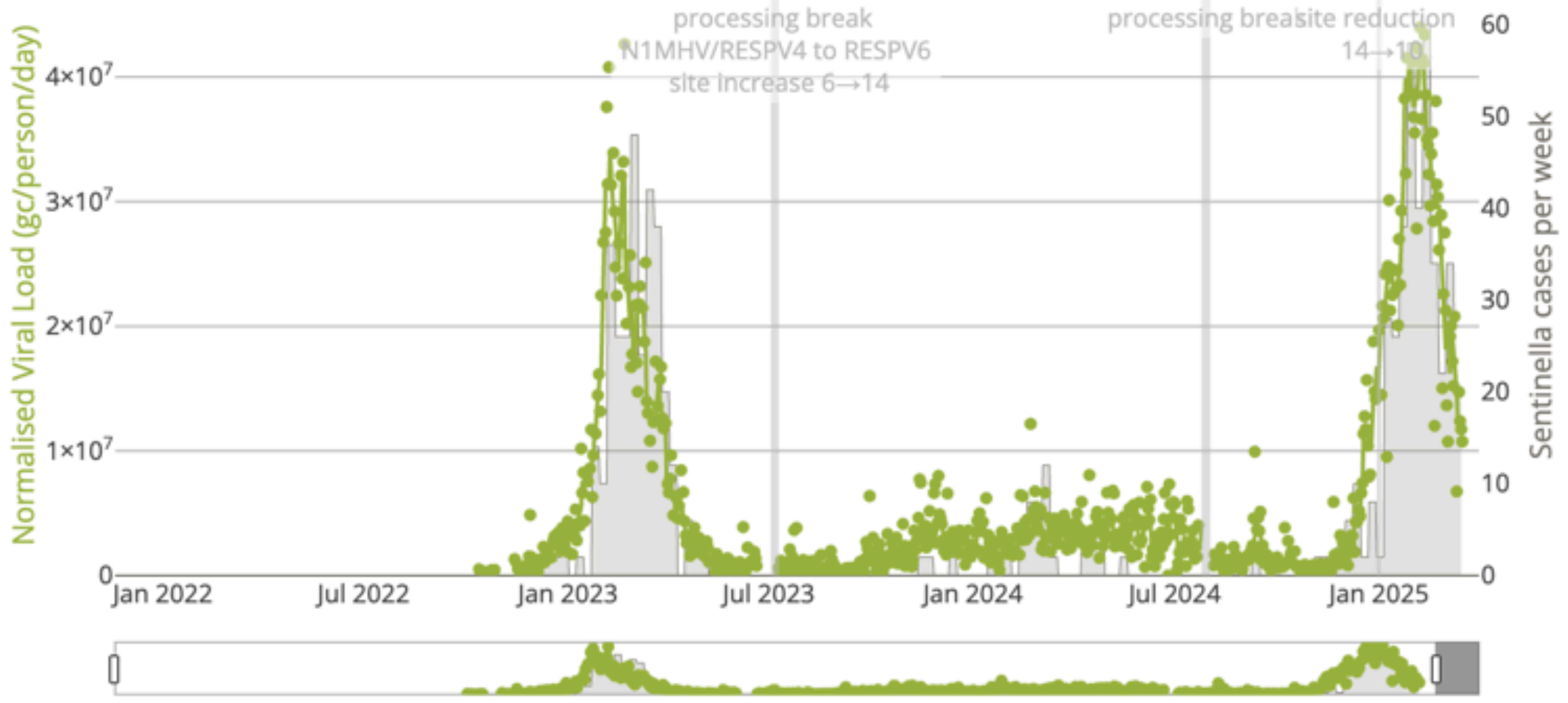
Influenza A Viral Load



Respiratory Syncytial Virus (RSV) Viral Load



Influenza B Viral Load



Case study:
eDNA sequencing to address local or global challenges

- What issue do you want to study with eDNA sequencing?
- What is your study site, target organism(s), community and where will you collect eDNA samples?
- What are potential challenges in your project?
- How does your study support sustainability (eg, conservation strategy, policy making, ecosystem management)?

10 min group discussion + 5 min presentation