



*[Under embargo until January 10<sup>th</sup>, 2024, at 17.00 CET]*

**PRESS RELEASE**

## **Discoveries gleaned from human ancient DNA**

*Four research articles in Nature follow the genetic traces and geographical origins of human diseases far back in time. The analyses provide detailed pictures of prehistoric human diversity and migration, while proposing an explanation for a rise in the genetic risk for multiple sclerosis (MS).*

By analysing data from the world's largest dataset to date on 5,000 ancient human genomes from Europe and Western Asia (Eurasia), new research has uncovered the prehistoric human gene pools of western Eurasia in unprecedented detail.

The results are presented in four articles published in the same issue of *Nature* by an international team of researchers led by experts from the University of Copenhagen and contributions from around 175 researchers from universities and museums in the UK, the US, Germany, Australia, Sweden, Denmark, Norway, France, Poland, Switzerland, Armenia, Ukraine, Russia, Kazakhstan, and Italy. The many researchers represent a wide range of scientific disciplines, including archaeology, evolutionary biology, medicine, ancient DNA research, infectious disease research, and epidemiology.

The research discoveries presented in the *Nature* articles are based on analyses of a subset of the 5,000 genomes and include:

- The vast genetic implications of a culturally determined barrier, which until around 4,000 years ago extended up through Europe from the Black Sea in the south to the Baltic Sea in the north.
- Mapping of how risk genes for several diseases, including type 2 diabetes and Alzheimer's disease, were dispersed in Eurasia in the wake of large migration events over 5,000 years ago.
- New scientific evidence of ancient migrations explaining why the prevalence of multiple sclerosis is twice as high in Scandinavia than in Southern Europe.
- Mapping of two almost complete population turnovers in Denmark, within a single millennium.



## THE 5000 ANCIENT HUMAN GENOMES PROJECT

The unprecedented dataset of 5,000 ancient human genomes were reconstructed by means of analysis of bones and teeth made available through a scientific partnership with museums and universities across Europe and western Asia. The sequencing effort was achieved using the power of Illumina technology.

The age of specimens ranges from the Mesolithic and Neolithic through the Bronze Age, Iron Age and Viking period into the Middle Ages. The oldest genome in the dataset is from an individual who lived approximately 34,000 years ago.

'The original aim of the ancient human genomes project was to reconstruct 1,000 ancient human genomes from Eurasia as a novel precision tool for research in brain disorders,' say the three University of Copenhagen professors, who in 2018 came up with the idea for the DNA dataset, and originally outlined the project concept:

Eske Willerslev, an expert in analysis of ancient DNA, jointly at the University of Cambridge, and the director of the project; Thomas Werge, an expert in genetic factors underlying mental disorders, and head of the Institute of Biological Psychiatry serving Mental Health Services in the Capital Region of Denmark; and Rasmus Nielsen, expert in statistical and computational analyses of ancient DNA, jointly at University of California, Berkeley, in the USA.

The objective was to produce a unique ancient genomic dataset for studying the traces and genetic evolutionary history of brain disorders as far back in time as possible to gain new medical and biological understanding of these disorders. This was to be accomplished by comparing information from the ancient DNA profiles with data from several other scientific disciplines.

Among the brain disorders the three professors originally identified as candidates for this investigation were neurological conditions such as Parkinson's disease, Alzheimer's disease, and multiple sclerosis, together with mental disorders such as ADHD and schizophrenia.

In 2018, the three professors then approached the Lundbeck Foundation – a major Danish research funding foundation – for funding to compile the special DNA dataset. They were awarded a five-year research grant totalling DKK 60 million (app. EUR 8m) for the project, which was to be coordinated at the University of Copenhagen via a newly established centre, subsequently named the Lundbeck Foundation GeoGenetics Centre.

'The rationale for awarding such a large research grant to this project, as the Lundbeck Foundation did back in 2018, was that if it all worked out, it would represent a trail-blazing means of gaining a deeper understanding of how the genetic architecture underlying brain disorders evolved over time. And brain disorders are our specific focus area,' says Jan Egebjerg, Director of Research, Lundbeck Foundation. The Lundbeck Foundation is also supporting iPYSCH consortium, one of the largest studies globally of genetic and environmental causes of mental disorders such as autism, ADHD, schizophrenia, bipolar disorder, and depression, where the focus is also on making genetic risk profiles for these disorders as precise as possible.



The results reported in *Nature*, were substantiated by comparing the ancient genomic dataset with de-identified genetic data from the large Danish iPYSCH consortium and DNA profiles from 400,000 present-day individuals registered in UK Biobank.

### **MANY CHALLENGES**

The premise for the project was experimental, recounts Professor Werge: ‘We wanted to collect ancient human specimens to see what we could get out of them, like trying to understand some of the environmental background to how diseases and disorders evolved. As I see it, the fact that the project took on such vast, complex proportions that *Nature* wanted it described in four articles is quite unique.’

Professor Willerslev comments that compiling the DNA dataset posed major logistical challenges: ‘We needed access to archaeological specimens of human teeth and bones that we knew were scattered around in museums and other institutions in the Eurasian region, and that called for many collaboration agreements. But once they were in place, things really took off – the dataset was booming, and it now exceeds 5,000 ancient human genomes. The size of the dataset has tremendously enhanced both the usability and precision of the results.

Professor Nielsen was responsible for planning the statistical and bioinformatics analyses of the information gleaned from the ancient teeth and bones in laboratories at the University of Copenhagen. And he was dealing with an overwhelming volume of data, in which the DNA was often severely degraded.

‘No one had previously analysed so many ancient genomes. Now we had to find out how to handle such vast data volumes. The problem was that the raw data is very difficult to work with because you end up with many short DNA sequences with many errors, and then those sequences have to be correctly mapped to the right position in the human genome. Plus, there is the issue of contamination from all the microorganisms present on the ancient teeth and bones. Imagine having a jigsaw puzzle consisting of millions of pieces mixed up with four other incomplete puzzle sets, and then running all that in the dishwasher for an hour. Piecing it all together afterwards is no easy task. One of the keys to our success in the end was that we teamed up with Dr. Olivier Delanau from the University of Lausanne who developed algorithms to overcome that very problem,’ says Professor Nielsen.

### **INTERNATIONAL INTEREST**

Rumours that a large ancient human genome dataset was being compiled were soon circulating in scientific circles. And since 2022 interest has been running very high, say Professors Werge, Willerslev and Nielsen:

‘We are constantly taking inquiries from researchers all over the globe – especially those investigating diseases – who typically request access to explore the ancient DNA dataset’.



The four *Nature* articles demonstrate that the large 5,000 genomes dataset serves as a precision tool capable of providing new insights into diseases when combined with analyses of present-day human DNA data and inputs from several other research fields. That in itself is immensely amazing, according to Professor Willerslev:

‘There’s no doubt that an ancient genomic dataset of this size will have applications in many different contexts within disease research. As new scientific discoveries derived from the 5,000-genome dataset become published, more data will gradually be made freely available to all researchers. Ultimately, the complete dataset will be open access for everyone’.

- END -

## CONTACT

### *Scientific contacts:*

Prof. Eske Willerslev, Lundbeck Foundation GeoGenetics Centre, Globe Institute, University of Copenhagen, Denmark, Prof. University of Cambridge, UK: ewillerslev@sund.ku.dk

Prof. Thomas Werge, University of Copenhagen, Denmark: thomas.werge@regionh.dk

Prof. Rasmus Nielsen, University of California Berkeley, Lundbeck Foundation GeoGenetics Centre, Globe Institute, University of Copenhagen, Denmark: rasmus\_nielsen@berkeley.edu

### *Media contact:*

Anna Razeto Richter, Academic Executive Assistant, Globe Institute, Lundbeck Foundation GeoGenetics Centre, University of Copenhagen, Denmark: anna.razeto@sund.ku.dk, +45 23 68 03 97

## LINKS TO ARTICLES IN NATURE

The four *Nature* articles based on information from the special DNA dataset are:

1. *Population Genomics of Postglacial Western Eurasia* (Allentoft et al.)  
<https://www.nature.com/articles/s41586-023-06865-0>
2. *The Selection Landscape and Genetic Legacy of Ancient Eurasians* (Irving-Pease et al.)  
<https://www.nature.com/articles/s41586-023-06705-1>
3. *Elevated Genetic Risk for Multiple Sclerosis Originated in Steppe Pastoralist Populations* (Barrie et al.) <https://www.nature.com/articles/s41586-023-06618-z>
4. *100 Ancient Genomes Show Repeated Population Turnovers in Neolithic Age Denmark* (Allentoft et al.) <https://www.nature.com/articles/s41586-023-06862-3>

(please note the URLs for each paper will go live on publication on the 10<sup>th</sup> of January at 17:00 CET)