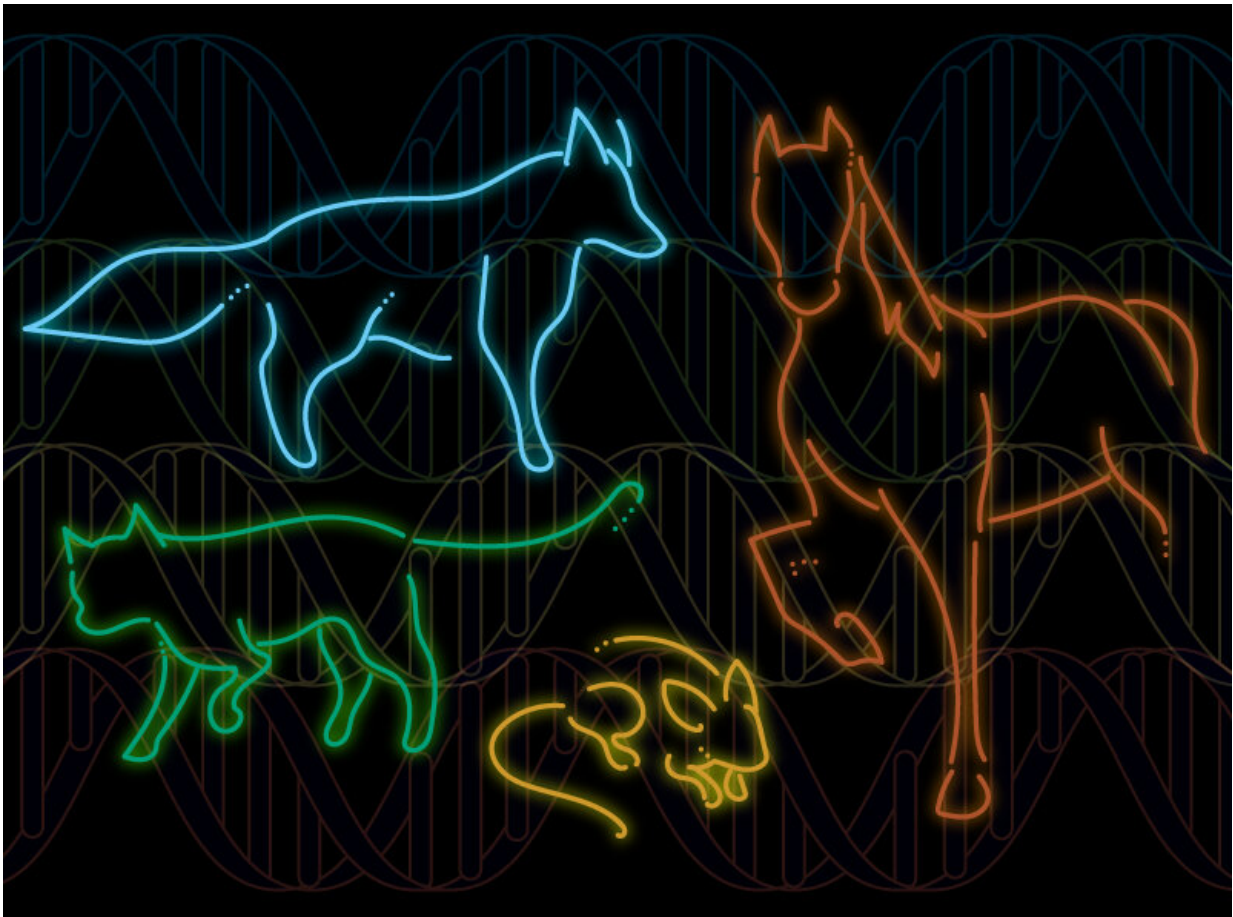# Largest set of mammalian genomes reveals species at risk of extinction

November 11 2020

A large international consortium led by scientists at Uppsala University and the Broad Institute of MIT and Harvard has sequenced the genome of 130 mammals and analysed the data together with 110 existing genomes to allow scientist to identify which are the important positions in the DNA. This new information can help both research on disease mutations in humans and how best to preserve endangered species. The study is published in *Nature*. Credit: Susanna Hamilton

An international team of researchers with an effort called the Zoonomia Project has analyzed and compared the whole genomes of more than 80 percent of all mammalian families, spanning almost 110 million years of evolution. The genomic dataset, published today in *Nature*, includes genomes from more than 120 species that were not previously sequenced, and captures mammalian diversity at an unprecedented scale.

The dataset is aimed at advancing human health research. Researchers can use the data to compare the genomes of humans and other mammals, which could help identify genomic regions that might be involved in human disease. The authors are also making the dataset available to the scientific community via the Zoonomia Project website, without any restrictions on use.

"The core idea for the project was to develop and use this data to help human geneticists figure out which mutations cause disease," said co-senior author Kerstin Lindblad-Toh, scientific director of vertebrate genomics at the Broad and professor in comparative genomics at Uppsala University.

However, in analyzing the new genomes, the authors also found that mammalian species with high extinction rates have less genetic diversity. The findings suggest that sequencing even just a single individual could provide crucial information, in a cost efficient way, on which populations may be at higher risk for extinction and should be prioritized for in-depth assessment of conservation needs.

"We wrote the paper to talk about this large, unique dataset and explain why it is interesting. Once you make the data widely available and explain its utility to the broader research community, you can really change the way science is done," said co-senior author Elinor Karlsson,

director of the Vertebrate Genomics Group at the Broad Institute of MIT and Harvard and professor at the University of Massachusetts Medical School.

Zoonomia data have already helped researchers in a recent study to assess the risk of infection with SARS-CoV-2 across many species. The researchers identified 47 mammals that have a high likelihood of being reservoirs or intermediate hosts for the SARS-CoV-2 virus.

Mapping mammals

The Zoonomia Project, formerly called the 200 Mammals Project, builds on a previous project, the 29 Mammals Project, which began sequencing mammalian genomes in 2006. The latest project extends the work by exploring the genomes of species that can perform physiological feats that humans can't, from hibernating squirrels to exceptionally long-lived bats. The project also included genomes of endangered species.

In the new study, the researchers collaborated with 28 different institutions worldwide to collect samples for genomic analysis, with the Frozen Zoo at the San Diego Global Zoo providing almost half of the samples. The team focused on species of medical, biological, and biodiversity conservation interest and increased the percentage of mammalian families with a representative genome from 49 to 82.

The project also developed and is sharing tools that will enable researchers to look at every "letter" or base in a mammalian genome sequence and compare it to sequences in equivalent locations in the human genome, including regions likely to be involved in disease. This could help researchers identify genetic sites that have remained the same and functional over evolutionary time and those that have randomly mutated. If a site has remained stable across mammals over millions of

years, it probably has an important function, so any change in that site could potentially be linked to disease.

In releasing the data, the authors call upon the scientific community to support field researchers in collecting samples, increase access to computational resources that enable the analysis of massive genomic datasets, and share genomic data rapidly and openly.

"One of the most exciting things about the Zoonomia Project is that many of our core questions are accessible to people both within and outside of science," said first author Diane Genereux, a research scientist in the Vertebrate Genomics Group at the Broad. "By designing scientific projects that are accessible to all, we can ensure benefits for public, human, and environmental health."

Provided by Broad Institute of MIT and Harvard