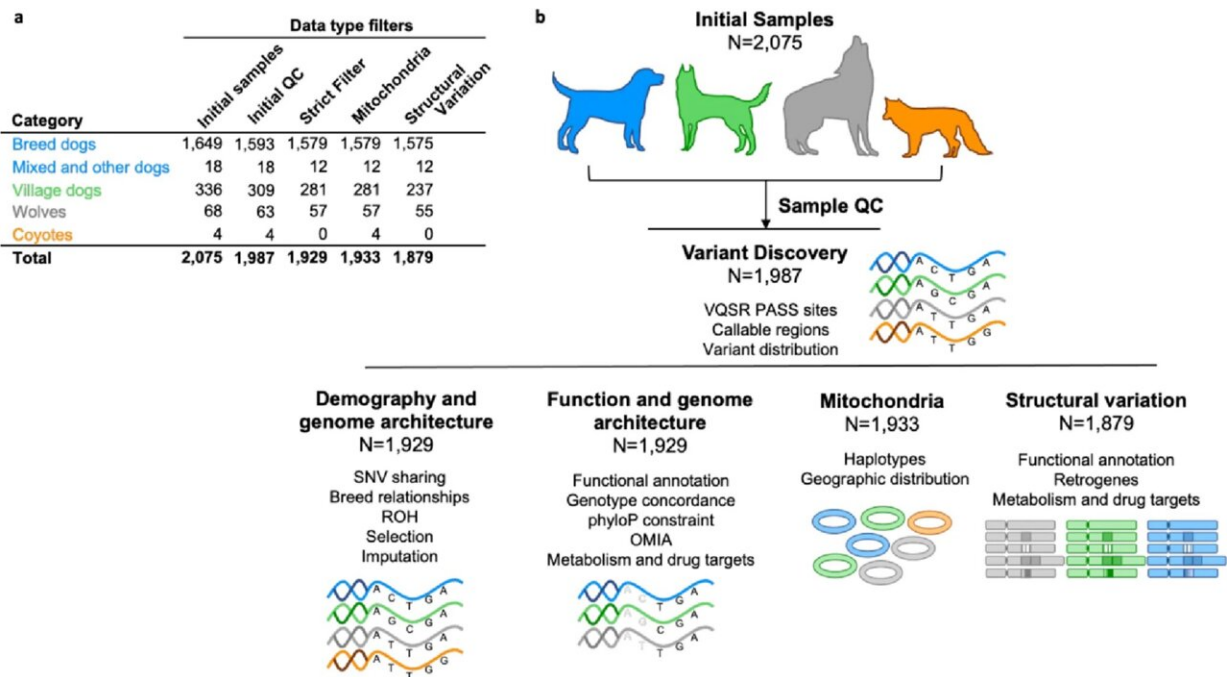


Publication of 2,000 canine genomes provides toolkit for translational research

August 18 2023



Overview of the Dog10K collection. a Sample collection and sample filtering for (b) the varied demographic, genome function, and architecture examined in the program. QC, quality control. ROH, runs of homozygosity. OMIA, Online Mendelian Inheritance in Animals. Credit: *Genome Biology* (2023). DOI: 10.1186/s13059-023-03023-7

An international research consortium has generated and analyzed 2,000 canine genomes. The resulting advanced genetics toolkit can now be

used to answer complex biological questions, spanning dog domestication, genetic differences in breed morphology, behavior and disease susceptibility, as well as the evolution and structure of the genome. The study, published in *Genome Biology*, describes the toolkit resource package and presents the first set of discoveries.

The publication is the culmination of efforts from the Dog10K consortium; 48 scientists across 25 institutions, contributing samples and resources to the immense analytical effort.

"The goal was to produce a resource the [global community](#) could access, and which they could use to speed the translation of their own research, be that in the study of the shared ancestors of dogs and wolves, or the clinical treatment of cancers. All these avenues are exciting, and all can benefit from the Dog10K catalog," says Uppsala University research scientist, Jennifer Meadows, lead co-author of the study.

The power of the Dog10K analyses lies in the depth of genetic diversity the team was able to capture. Canine samples were drawn from more than 320 of the approximately 400 recognized pedigree dog breeds, as well as niche populations of village dogs, wolves, and coyotes.

With this the team developed:

- A comprehensive catalog of single nucleotide variants, including those which can disrupt protein coding genes and may influence the extremes of dog appearance and physiology. All pedigree dogs in the study were free from disease when sampled, suggesting that some of the variants that cause a loss of protein function do not cause obvious disease in that population.
- A catalog of structural variants. This category of variants has a minimum size of 50 base pairs, and so impacts more space in the genome than a single nucleotide [variant](#). Some structural variants

were found to impact protein coding genes. This could mean that too much, or too little of the protein is made in that individual, but again did not cause obvious disease in the pedigree population.

- For both single nucleotide and structural variants, the Dog10K team illustrated how these variants are distributed within and between pedigree dog breeds, and how these variants may impact the study of potential drug therapies.
- Publicly available data processing pipelines. Combining the results from different studies can be made more difficult when different settings are used to map sequencing data, and then to call variation. The pipeline used by the Dog10K consortium is open to the community so that new or existing data can be processed and easily combined with this set.
- An imputation panel from the single nucleotide variants. Imputation is a process that allows researchers to infer variation at sites that they did not directly measure. The Dog10K team showed that by using the imputation panel, a dataset that contained 170 thousand single nucleotide sites, commonly used in genotyping array studies, can be expanded to up to 8 million sites. This allows researchers to reuse their valuable samples, and to revisit complex questions that need more data.

"We have just scratched the surface of the data's potential," continues Meadows. "There is yet more genetic diversity left to be found in dogs, wolves and coyotes, but the Dog10K team looks forward to seeing how this first effort is applied by the canine science community."

More information: Jennifer R. S. Meadows et al, Genome sequencing of 2000 canids by the Dog10K consortium advances the understanding of demography, genome function and architecture, *Genome Biology* (2023). [DOI: 10.1186/s13059-023-03023-7](https://doi.org/10.1186/s13059-023-03023-7)

Provided by Uppsala University

Citation: Publication of 2,000 canine genomes provides toolkit for translational research (2023, August 18) retrieved 2 May 2024 from <https://phys.org/news/2023-08-canine-genomes-toolkit.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.