Mapping the genetic diversity of Antarctic life

ASU virologist Arvind Varsani brings specialized expertise to $18.4M NSF project sequencing Antarctic organisms, ecosystems

By Gabriela Harrod, ASU News
October 12, 2026

Antarctica is home to organisms that have evolved to survive months of darkness, frigid temperatures and some of the harshest conditions on Earth. Hidden within their genetic material are clues to how life adapts and persists in extreme environments.

A new $18.4 million project supported by the National Science Foundation will help make those clues available to scientists around the world. The initiative, called Genome Infrastructure catalyzing next Generation Antarctic Data, Access, Tools, and Analytics, or GIGA DATA, will build an extensive genomic resource representing life across Antarctica.

Arvind Varsani, a professor in the School of Life Sciences and core faculty member in the Biodesign Center for Fundamental and Applied Microbiomics, Center for Evolution and Medicine and Center for Mechanisms of Evolution, is a co-principal investigator on the multi-institutional project.

A molecular virologist, Varsani studies viruses across ecosystems and has conducted extensive research on viruses associated with Antarctic penguins, seals, fish and other wildlife.

Led by Alison Murray of the Desert Research Institute, GIGA DATA aims to produce reference-quality genomes for more than 250 Antarctic species, including algae, invertebrates, fish, birds and marine mammals. The project will also analyze 600 environmental samples from ecosystems such as lakes and soils, producing tens of thousands of microbial and viral genomic datasets.

“The main objective is to sequence at least 250 eukaryote genomes from a variety of Antarctic vertebrates and invertebrates,” Varsani said. “In addition to that, sequence microbial communities (including viruses) in metagenomes of approximately 600 samples from different environments.”

Rachel O’Neill and Jill Wegrzyn at the University of Connecticut will conduct the sequencing, while other partner institutions will contribute expertise in areas including bioinformatics (Patrick Chain and Bin Hu at New Mexico Consortium), avian and pinniped genomics (Rauri Bowie at the University of California, Berkeley and Allyson Hindle at the University of Nevada, Las Vegas).

Researchers throughout the Antarctic science community will be able to propose samples for sequencing, including specimens collected during previous expeditions and field seasons.

At ASU, Varsani will focus on ensuring that the computational workflows used to identify and analyze viral sequences are robust. He will also help make the resulting data available to the broader research community and participate in reviewing proposed samples.

Viruses require specialized expertise because their genetic material differs significantly from that of cellular organisms. Viral genomes can consist of either DNA or RNA, and the tools designed to analyze the genomes of plants or animals cannot always be applied to them.

“The genomics of viruses are quite complex and different from the genomics of organisms,” Varsani said. “Their genomes are relatively small and complex, and they can be RNA or DNA. The sequencing platforms and pipelines need to be adapted to improve the signal to noise ratio to address that.”

Creating a comprehensive genomic resource could help researchers identify genes that allow Antarctic organisms to tolerate cold, extended sunlight or darkness and other environmental stresses. Those discoveries could improve understanding of evolution and ecosystem resilience while potentially leading to applications in biotechnology or medicine.

Varsani pointed to pigments produced by bacteria exposed to long periods of sunlight as one example of how biological adaptations could inspire useful products. Antarctic fish produce antifreeze glycoproteins, which researchers have shown to have evolved from a proteinase gene to an ice-binding protein encoding gene some 5–14 million years ago.

Such discoveries may lead to the identification of closely related genes in non-antarctic organisms whose gene products could lead to new biotechnology and therapeutic innovation. Other organisms may contain genes or molecules with currently unknown function but with new protein structure prediction tools may help decipher their function.

The project also addresses a significant shortage of Antarctic genomic data. Outside of heavily studied subjects such as avian influenza, information about Antarctic viruses remains particularly limited.

“The output from the project will dramatically improve the landscape of sequence information from Antarctica,” Varsani said. “It will help us understand some of the really complex evolution associated with the animals and organisms that live in these systems.”

Producing the data will not be simple. Some Antarctic organisms, including icefish, have genomes containing highly repetitive regions that can be difficult to assemble. The team will use newer long-read sequencing technologies to capture larger stretches of genetic material and combine them with shorter sequences to construct more complete genomes.

For Varsani, those scientific and technical challenges are part of what makes GIGA DATA compelling.

“It’s a huge-scale project and a big undertaking,” he said. “It’s pretty bold.”

This story originally appeared on ASU News.

Main image



An Adélie penguin photographed in Antarctica. Photo courtesy of Arvind Varsani/ASU

Text image(s)



Arvind Varsani