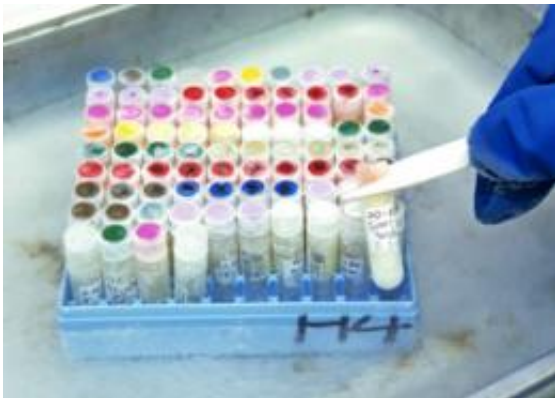


Genome 10K Project announces first 101 species for genome sequencing

November 30 2010, By Tim Stephens



Specimens available for genome sequencing include these frozen cell samples at the San Diego Zoo. Photo courtesy of San Diego Zoo.

(PhysOrg.com) -- The Genome 10K Community of Scientists and BGI (formerly the Beijing Genomics Institute) of Shenzhen, China, have announced a plan to sequence the genomes of 101 vertebrate species within the next two years, the first of an eventual 10,000 species to be sequenced by the Genome 10K Project.

The Genome 10K Project (G10K) is an international effort to gather specimens of thousands of animals from zoos, museums, and university collections throughout the world, and then sequence the genome of each species to reveal its complete genetic heritage. The project aims to assemble a genomic zoo--a collection of [DNA sequences](#) for 10,000

vertebrate species (approximately one for every vertebrate genus) by 2015.

Genome 10K cofounder David Haussler, a professor of biomolecular engineering in the Baskin School of Engineering at UC Santa Cruz, said the first 101 species to be sequenced were selected from the Genome 10K database, which catalogs specimens suitable for sequencing from more than 16,000 vertebrate species, both living and recently extinct. (A list of the selected species is [available for download](#).)

"These genomes will provide a new and exciting window into vertebrate biology," Haussler said. "The experiments of evolution represented in them carry critical information about how animals are built and how they function at the molecular level."

The planned sequencing will take place within the next 24 months at the BGI facilities in China. BGI has recently amassed the largest number of advanced DNA sequencing machines under one roof of any institute in the world and is committed to large-scale application of genome sequencing in science and medicine.

"With this joint pledge to accomplish whole-genome sequence assembly and annotation of 101 new vertebrate species--the first 1 percent of the Genome 10K target--we move much closer to providing students of biology the unabridged DNA code book for the wonder of living vertebrate forms," said project cofounder Stephen J. O'Brien, chief of the Laboratory of Genomic Diversity at the National Cancer Institute.

The sequencing project builds on a BGI project launched in January 2010, called the 1,000 Plant and Animal Reference Genome Project. According to BGI executive director Jun Wang, their project harnesses the leading sequencing platform, advanced technologies, and powerful bioinformatics analytical capability.

"The collaboration with G10K will not only significantly facilitate the progression of BGI's genomics research, but also benefit the development of global genomics researches and their relevant industrial application," Wang said.

The 101 vertebrate species chosen for this phase of the project were selected by a consortium of BGI scientists, their collaborators, sponsors, and the chairs of the Genome 10K taxonomy groups for mammals, birds, amphibians, reptiles, and fish. The species were chosen to capture phylogenetic diversity across the vertebrate radiations while avoiding overlap with public-sponsored whole-genome sequencing projects. Other considerations included the biological features that make a species interesting to humankind, the science community around each species, and the availability of vouchered specimens.

According to Genome 10K cofounder Oliver Ryder, director of genetics at the San Diego Zoo Institute for Conservation Research, the collection and banking of samples in international centers, such as San Diego's Frozen Zoo, sets the stage for rapid progress toward the sequencing goals in collaboration with major sequencing centers, such as BGI. "The data produced provide a major contribution to understanding aspects of the biology of many endangered species, thereby assisting in conservation efforts on their behalf," Ryder said.

In addition to the 101 species chosen for G10K, there are 120 vertebrate species already being sequenced in public-sector genome projects (download list [here](#)). The completion of 221 whole-genome sequences within a few years is a first step toward the G10K goal of obtaining high-quality whole-genome sequences, assembled and annotated, for 10,000 vertebrate species. To meet the quality goals of the G10K project, every species must have a reference [genome sequence](#) with chromosome-scale contiguity suitable for display on a genome browser and accurate enough for genome-based research.

Selection of future species for sequencing in the Genome 10K project will involve open-science-based input and collaboration, Haussler said. Haussler's team at UCSC's Center for Biomolecular Science and Engineering maintains the Genome 10K database and the UCSC Genome Browser.

The Genome 10K Project will hold a workshop in Santa Cruz, California, from March 14 to 18, 2011. This workshop will accomplish two goals: First, it will continue the mission to gather biological specimens of vertebrate species and describe them thoroughly on a public database. Second, it will gather [genome](#) assembly experts to develop a robust solution to the challenge of creating useful assembly with short reads and combined-scaffold assembly approaches.

More information: genome10k.soe.ucsc.edu/

Provided by University of California - Santa Cruz

Citation: Genome 10K Project announces first 101 species for genome sequencing (2010, November 30) retrieved 20 April 2024 from <https://phys.org/news/2010-11-genome-10k-species-sequencing.html>

<p>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.</p>
--