

Cornell part of national team to map 2,500 genomes

Cornell University

Cornell researchers are part of an international collaboration to build the most detailed map of human genetic variation. The map promises to provide a much more comprehensive understanding of the role of inherited DNA variation in human history, evolution and disease and the best methods to use to sequence DNA.

The 1,000 Genomes Project, an international public-private consortium, announced results from several pilot studies in Nature Oct. 28. The report describes the use of next-generation technology to sequence more than 1,000 human genomes from 27 populations worldwide, with ongoing data made freely available soon after it is generated. The full-scale project, to be completed by 2012, aims to sequence 2,500 individual genomes.

Cornell researchers provided statistical analysis of the vast amount of genetic data to identify errors and perform population genetics analyses. The project, headed by researchers from the Wellcome Trust Sanger Institute and the Broad Institute, takes advantage of technological advancements that have increased the amount of sequencing per dollar by about 10 times per year.

To sequence a genome, DNA is cut into fragments of a few hundreds of base pairs each, which are then "read" as the sequencing machine synthesizes new copies. Researchers must then use sophisticated computational methods to piece together where each segment belongs in the genome. To sequence 2,500 genomes, the researchers will do less sequencing for each individual, repeating the process so that each base pair is "read" about four times. Doing fewer passes over each individual's genome cuts costs but increases errors.

"The researchers at Cornell have experience working with DNA sequences in this context, where there are statistical uncertainties, so we contributed by developing robust statistical methods," said Andrew Clark, Cornell's Jacob Gould Schurman Professor of Population Genetics, a co-principal investigator for the project and a member of the project steering committee.

Cornell researchers also conducted "analysis of the population genetics of the samples, such as the distribution of frequencies of variants, quantifying population differentiation and the age of mutations," said Clark. "And of course some fraction of these variants are likely to play a role in genetic disease risk."

And Alon Keinan, assistant professor of biological statistics and computational biology, has also been involved in designing the global sampling of human populations and in contrasting variation patterns between the X chromosome and non-sex chromosomes.

Cornell part of national team to map 2,500 genomes

Published on Electronic Component News (<http://www.ecnmag.com>)

The Nature paper describes three pilot studies to test multiple strategies to catalog genetic variants.

The first pilot study, a model of the main project, sequenced the genomes of 179 individuals from European, African and East Asian populations, using low coverage, an average of four reads for each base pair. The researchers found this low coverage strategy is an effective way to discover common genetic variants between individuals.

The second study, involved sequencing six people (two families, each with two parents and a daughter) in detail -- each base pair was read an average of 20 to 60 times -- using different sequencing technologies, which uncovered the pros and cons of each technology. The results also served as a comparison group for the first pilot, which used lower coverage.

The third pilot involved high coverage sequencing of exons -- the protein-coding functional parts of genes -- of 700 individuals to augment the data for functionally important parts of the genome. This pilot also showed how worthwhile the exon data is, so the researchers plan to sequence the exons from all 2,500 individuals with high coverage in the final project.

The project is funded by many organizations, including 454 Life Sciences, a Roche company; Beijing Genomics Institute, China; the Max Planck Institute for Molecular Genetics, Berlin, Germany; and the National Human Genome Research Institute.

[SOURCE](#) [1]

Source URL (retrieved on 04/26/2015 - 5:51am):

http://www.ecnmag.com/news/2010/11/cornell-part-national-team-map-2500-genomes?qt-recent_content=0

Links:

[1] <http://www.news.cornell.edu/stories/Nov10/1000GenomesProject.html>