

Multiple genome sequencing yields detailed map of structural variants behind our genetic differences

February 2 2011

Analyzing billions of pieces of genetic data collected from people around the world, Boston College biologist Gabor Marth and his research team are playing an integral role in the global effort to sequence 1000 genomes and move closer to understanding in fine detail how genetics influence human health and development.

The most comprehensive map to date of genomic structural variants – the layer of our DNA that begins to distinguish us from one another – has been assembled by analyzing 185 human genomes, Marth and co-authors from the 1000 Genomes Project team report in the Feb. 3 edition of the journal *Nature*.

The complexity of the 1000 Genome Project draws on a range of expertise in the Marth bioinformatics lab, which receives volumes of data produced by other project teams using DNA sequencing technology, stores the data, and then analyzes it using proprietary computer software programs the Marth lab has developed.

"The tools we have developed are being used to discover a biological reality that we could not see before," said Marth, an associate professor of biology whose group is one of the lead analytics units for the 1000 Genomes Project. "There are many challenges and the work is very exciting."

The goal is to understand the genetic make up of the earth's population by analyzing genome data from as many as 2,500 individuals in order to provide new insights into the development of the human race and to understand the links between the genome and human health.

"We are working with some of the world's best research groups," said Marth, joined as a co-author on the paper with his BC colleagues Research Assistant Professor Chip Stewart and doctoral candidates Deniz Kural and Jiantao Wu.

"There are engineering, mathematical, and algorithmic challenges at every level," Marth added. "We work to make sure our computational tools are performing well, make continuous improvements and process data in a timely fashion to send to our colleagues around the world."

The researchers report in Nature the generation of a map of structural variants – those pieces of genetic code that are the base layer of instructions, also known as the genotype, that ultimately determine our outward appearances and characteristics, or phenotypes. The new map is built upon a range of structural variants, including 22,025 deletions, or missing pieces of DNA, and 6,000 insertions, pieces of DNA that have been added along the evolutionary journey, and tandem duplications.

The analysis has produced new insights into genetic selection, the introduction of large structural variants into DNA and structural variant "hotspots" formed by common biological mechanisms, the team reports in Nature. The map will play a crucial role in sequencing-based association studies, where this new understanding of human variation is applied to unlocking new ways to use the genome to understand the world's population and to inform the life and medical sciences.

"The eventual goal of studying the genotype is so we can understand how the specific genetic make-up of an individual is responsible for an

individual phenotype, such as height or weight or susceptibility to disease," said Marth. "The specific question of the 1000 [Genome Project](#) is how much divergence, or how much genetic variation, exists within different populations. That is the question we are trying to unravel."

Provided by Boston College

Citation: Multiple genome sequencing yields detailed map of structural variants behind our genetic differences (2011, February 2) retrieved 17 April 2024 from <https://medicalxpress.com/news/2011-02-multiple-genome-sequencing-yields-variants.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.