

Counting duplicated genome segments now possible

August 30 2009

(PhysOrg.com) -- A newly designed computational method has proven its usefulness in counting copies of duplicated genome sequences and in doing initial assessments of their contents, according to a study to be published Aug. 30 in *Nature Genetics*. The number of copies of particular DNA segments can differ from one person to the next.

The researchers named their method mrFAST, an acronym for micro-read Fast Alignment Search Tool. The study is titled, "Personalized Copy Number and Segmental Duplication Maps Using Next Generation Sequencing." The lead authors are Dr. Can Alkan, a senior fellow, and Jeffrey M. Kidd, a graduate student, both in the Department of Genome Sciences at the University of Washington (UW). Dr. Evan E. Eichler, UW professor of genome sciences, is the senior author.

Segmental duplications in the human genome have been associated with susceptibility and resistance to disease. Duplicated segments have been linked to such disorders as lupus, Crohn's disease, mental retardation, schizophrenia, color blindness, psoriasis, and age-related [macular degeneration](#). Segmental duplications often contain duplicated genes, many of which have an unknown function. Individuals have different numbers of copies of some of these duplications. Determining the number, content, and location of segmental duplications is an important step in understanding the health significance of gene copy-number variation.

"New computational methods, combined with next-generation [DNA](#)

sequencing technology, has provided for the first time an accurate census of specific genes that exist in varying number of copies," Alkan said.

"This is a way to deal with some of the most complex regions of the human genome and do what might appear to be a simple thing: Count whether a person has one, two, three or more copies of a gene," explained Kidd. "In fact, such counting is surprisingly difficult." Many standard genome analyses exclude duplication-rich or repeat-rich regions of the human genomes because their sequences are not unique.

Before this study, by using different methods scientists could analyze the entire genome of a person and say that an individual has more or fewer copies of a particular gene, but not the absolute number of copies. For example, scientists have known that some people have an increased copy-number of a gene that confer some resistance to HIV, but couldn't tell how many.

The UW researchers further examined the much-studied genomes from three healthy individuals: a European (DNA research pioneer James D. Watson), a Yoruban African individual from Nigeria, and a Han Chinese. The researchers were able to predict copy-number differences among the individuals, even when there were many copies, such as 5 in one person compared to 12 in another. The researchers conservatively validated 113 genes that were copy-number variable among the three people, but more genes are suspected to be copy-number variable. Several of the validated gene differences are known to be of biomedical relevance. They include, for example, genes related to eye and skin diseases, and many others that play a role in the immune system. The researchers noted that several human genes with the most variable copy numbers correspond to a torrent of segmental duplications that occurred within the common ancestor of apes and humans.

In talking about their study, the researchers mentioned that next-

generation technology for sequencing the human genome has far greater detection power and costs substantially less than the traditional sequencing method known as Sanger sequencing. The new technologies are beginning to distinguish subtle dissimilarities between nearly identical gene copies.

"This can provide researchers with a more accurate assessment of specific gene content and insight into functional constraints," Alkan explained.

"The newer, faster genome sequencing platforms," Alkan added, "may eventually make it feasible to detect the full-spectrum of genomic variation among many individuals, including patients suffering from diseases of genetic origin. Next-generation technology and computational methods promise low cost, rapid sequencing of different individuals and may lead to a fuller understanding of the patterns and significance of human genetic variation."

The analytical method they devised is already being tapped for the 1000 Genome Project, an international effort to catalog and compare the genomes of hundreds of people from around the world.

Alkan, Kidd, and their colleagues noted that copy number variants, including variable duplications of entire genes, are recognized as making substantial contributions to human diversity.

The ability to accurately and systematically determine the absolute copy number for any genomic segment is a notable step, the researchers added, toward a true and complete picture of individual genomes and how the genome shapes a person's characteristics.

"The next challenge," they wrote, "will be defining variation in the sequence content and the structural organization of these dynamic and

important regions of the human [genome](#)."

Source: University of Washington ([news](#) : [web](#))

Citation: Counting duplicated genome segments now possible (2009, August 30) retrieved 4 May 2024 from <https://medicalxpress.com/news/2009-08-duplicated-genome-segments.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.