

New computation method helps identify functional DNA

January 21 2015, by Krishna Ramanujan



From left to right, the structures of A-, B- and Z-DNA. Credit: Wikipedia

Striving to unravel and comprehend DNA's biological significance, Cornell scientists have created a new computational method that can identify positions in the human genome that play a role in the proper functioning of cells, according to a report published Jan. 19 in the journal *Nature Genetics*.

The <u>human genome</u> is vast, totaling some three billion <u>base pairs</u> of



nucleotides, the subunits of DNA. But only about 1.25 percent of those billions of base pairs account for genes that encode all the proteins we use. A fraction of the rest of that <u>genetic material</u> regulates genes and turns them on and off, but these have yet to be fully identified.

"This paper tackles the deep question of how to identify functional noncoding human genomic material controlling human traits and disease," said Brad Gulko, the paper's first author and a graduate student in the field of computer science. Gulko's adviser, Adam Siepel, Cornell associate professor of biological statistics and computational biology and professor of computer science at Cold Spring Harbor Laboratory, is a coauthor.

"What makes our approach unique is the straightforward combination of DNA biochemistry with recent evolutionary pressures," said Gulko. "Our method allows other scientists not only to use the results, but to readily understand them."

Insight into the human genome gained from this new computation method could be applied to personalized medicine and it may be a big step toward developing treatments for diseases like AIDS, malaria, muscular sclerosis, ALS and Alzheimer's.

Geneticists identify biologically significant DNA by looking for signals of <u>selective pressure</u> in DNA, genes and genetic material that give individuals in a population advantages and greater "fitness," or reproductive success.

The new method combines two previously used techniques to identify selective pressure. One technique looks for divergence, or differences between humans and chimpanzee genomes accumulated over millions of years; a less commonly used method looks for mutations in DNA (polymorphisms) between individual humans.



The new <u>computational method</u> clusters functionally similar markers in the genome into groups, then estimates a probability of whether a group is contributing to the fitness of the species based on associated patterns of divergence and genomic polymorphisms.

In this way, the researchers receive a "fitness consequence" (fitCons) score that predicts which genetic material might be under selective pressure and therefore biologically significant.

Compared to conventional techniques, fitCons scores demonstrate a much greater power to predict which genetic material regulates the expression of genes.

In addition, fitCons scores indicate that 4.2 to 7.5 (but probably closer to 5) percent of nucleotides in the human genome have influenced fitness since humans diverged from chimpanzees.

Provided by Cornell University

Citation: New computation method helps identify functional DNA (2015, January 21) retrieved 6 May 2024 from <u>https://phys.org/news/2015-01-method-functional-dna.html</u>

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.