Cornell Finds Natural Selection in Humans

October 26 2005

The most detailed analysis to date of how humans differ from one another at the DNA level shows strong evidence that natural selection has shaped the recent evolution of our species, according to researchers from Cornell University, Celera Genomics and Celera Diagnostics.

In a study published in the Oct. 20 issue of the journal *Nature*, Cornell scientists analyzed 11,624 genes, comparing how genes vary not only among 39 humans but also between the humans and a chimpanzee, whose DNA is 99 percent identical to humans.

The comparisons within and between species suggest that about 9 percent of genes that show some variability within humans or differences between humans and chimpanzees have evolved too rapidly to be explained simply by chance.

The study suggests that positive Darwinian natural selection -- in which some forms of a gene are favored because they increase the probability of survival or reproduction -- is responsible for the increased rate of evolution. Since genes are blueprints for proteins, positive selection causes changes in the amino acid sequence of the protein for which the gene codes.

"Our study suggests that natural selection has played an important role in patterning the human genome," said the paper's lead author, Carlos Bustamante, assistant professor of biological statistics and computational biology at Cornell.

The Cornell/Celera team found that genes involved in immune function, sperm and egg production, sensory perception and transcription factors (proteins that control which genes are turned on or off) have been particularly affected by positive selection and show rapid evolution in the last 5 million years, when humans shared a common ancestor with chimps.

Likewise, the researchers found that approximately 13 percent of the genes that may vary show evidence of slightly deleterious or harmful mutations in human populations; these include genes involved in determining the basic structure of cells and muscles as well as genes that control traffic in and out of the cell. These mutations are subject to weak negative selection, according to the study.

In general, negative selection eliminates from the population very harmful changes to proteins that kill or stop reproduction. But mutations that have led to slightly deleterious versions of the gene -- mutations that may cause disease or only slightly reduce the average number of children left by those that carried the mutation -- can by chance become quite common in the population.

The authors also found a correlation between genes predicted to be under negative selection and genes implicated in certain hereditary diseases. For example, among the genes the researchers predicted to be under negative selection are those involved in muscular dystrophy and in Usher syndrome, the most common cause of congenital blindness and deafness in developed countries.

"We have a long way to go before we can predict from looking at sequences, which mutations in which genes and under which environmental conditions can ultimately lead to disease. This is a first step in identifying the classes of genes that appear to be particularly vulnerable to these types of changes," said Bustamante.

A team from Celera initiated the project and sequenced more than 20,000 genes in 39 humans and a chimpanzee. By comparing the DNA sequences of the 39 human subjects across the 20,000 genes, the Celera researchers identified DNA sites in the genome where individuals in the sample differed from one another. The chimpanzee sequence was then used to identify which form of the gene was the original ancestral form and which was the derived or new type.

The original goal of the project was to identify novel amino acid variants that could then be tested for association with human disease in subsequent studies. The Cornell researchers became involved at the analysis stage in order to make predictions about what types of changes are most likely to be functionally important.