

# Scientists find a fingerprint of evolution across the human genome

April 8 2008

---

The Human Genome Project revealed that only a small fraction of the 3 billion “letter” DNA code actually instructs cells to manufacture proteins, the workhorses of most life processes. This has raised the question of what the remaining part of the human genome does. How much of the rest performs other biological functions, and how much is merely residue of prior genetic events?

Scientists from Cold Spring Harbor Laboratory (CSHL) and the University of Chicago now report that one of the steps in turning genetic information into proteins leaves genetic fingerprints, even on regions of the DNA that are not involved in coding for the final protein. They estimate that such fingerprints affect at least a third of the genome, suggesting that while most DNA does not code for proteins, much of it is nonetheless biologically important – important enough, that is, to persist during evolution.

## Conservation of genetic information

To gauge how critical a particular stretch of DNA is, biologists often look at the detailed sequence of “letters” it consists of, and compare it with a corresponding stretch in related creatures like mice. If the stretch serves no purpose, the thinking goes, the two sequences will differ because of numerous mutations since the two species last shared an ancestor. In contrast, it’s believed that the sequences of important genes will be similar, or “conserved,” in different species, because animals

with mutations in these genes did not survive. Biologists therefore regard conserved sequences as a sign of biological importance.

To test for conservation, researchers need to find matching stretches in the two species. This is relatively easy for stretches that “code” for proteins, where scientists long ago learned the meaning of the sequence. For “noncoding” regions, however, the comparison is often ambiguous. Even within a gene, stretches of DNA that code for pieces of the target protein are usually interspersed with much larger noncoding stretches, called introns, that are removed from the RNA working copy of the DNA before the protein is made.

## **Signs of splicing**

Previous researchers assumed that mutations in the middle of introns do not affect the final protein, so they simply accumulate. In the new work, however, the researchers found signs that evolution rejects some types of mutations even in these regions of the genome. Although the selection is weak, “introns are not neutral,” in their effect on survival, says CSHL professor Michael Zhang, a bioinformatics expert who headed the research team.

To look for selection, CSHL researcher Chaolin Zhang, a doctoral candidate at Stony Brook University, looked in the human genome for a subtle statistical imbalance in how often various “letters” appear. The researchers attribute this imbalance to special short stretches of DNA that mark regions to be removed. Unless these signal sequences are sprinkled throughout an intron, the data suggest, it may not be properly spliced out, with potentially fatal consequences. Other sequences must likewise be preserved in the regions to be retained.

The scientists found a preference for some “letters” across intron regions, and the opposite preference in coding regions. Together, these

regions make up at least a third of the genome, which is thus under selective pressure during evolution. The result supports other recent studies that suggest that, although most DNA does not code for proteins, much of it is nonetheless biologically important.

In addition to demonstrating how splicing affects genetic evolution, the statistical analysis identified possible signaling sequences, some that were already known and others that are new. According to co-author Adrian Krainer, a CSHL professor and splicing expert, “the exciting thing will be to experimentally test whether these predicted elements are really true.”

“RNA landscape of evolution for optimal exon and intron discrimination” appears in the April 15, 2008 edition of the *Proceedings of the National Academy of Sciences*. The paper is available online at [www.pnas.org/cgi/doi/10.1073/pnas.0801692105](http://www.pnas.org/cgi/doi/10.1073/pnas.0801692105).

Source: Cold Spring Harbor Laboratory

Citation: Scientists find a fingerprint of evolution across the human genome (2008, April 8) retrieved 17 April 2024 from <https://phys.org/news/2008-04-scientists-fingerprint-evolution-human-genome.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.