boston.com

THIS STORY HAS BEEN FORMATTED FOR EASY PRINTING

DNA study challenges basic ideas in genetics

Genome 'junk' appears essential

The Boston Globe

By Colin Nickerson, Globe Staff | June 14, 2007

A massive international study of the human genome has caused scientists to rethink some of the most basic concepts of cellular function. Genes, it turns out, may be relatively minor players in genetic processes that are far more subtle and complicated than previously imagined.

Among the critical findings: A huge amount of DNA long regarded as useless -- and dismissively labeled "junk DNA" -- now appears to be essential to the regulatory processes that control cells. Also, the regions of DNA lying between genes may be powerful triggers for diseases -- and may hold the key for potential cures.

The research, published in a set of papers in today's editions of the journals Nature and Genome Research, raised far more questions than it answered -- and in a sense was a rallying cry for more and deeper research into the functioning of the genome, often referred to as the "blueprint" for life.

"The instruction manual for life is written in a language we are only just beginning to understand," Francis Collins, director of the federal government's National Human Genome Research Institute , said at a news conference yesterday.

Collins' s institute was among the more than 80 research institutions in North America, Europe, Asia, and Australia that participated in the $42 million, four-year study, whose aim was to analyze 30 million units of human DNA -- just 1 percent of the entire human genome -- to create an inventory of biologically functional elements. The project is known as the Encyclopedia of DNA Elements, or ENCODE, and involved an exhaustive scrutiny of 44 broad "sites" in the human genome, probing not just genes, but all material in the samples.

"We're finding that a lot of the genome is as mysterious as 'dark matter' in physics; we know it is out there doing something. The challenge is to find out what and why," said Thomas D. Tullius , professor of chemistry at Boston University and one of the ENCODE researchers. "There were huge surprises; this research has upset a lot of thinking about how the genome works."

He added in an interview: "There now appear to be thousands of places in the genome that were long thought to be useless or meaningless, but which we now see to have a functional role. But we don't really understand what that role is."

Most startling, according to researchers, is that some areas of the genome looming as crucial are regions that don't contain specific instructions for making proteins. That recognition amounts to a sea change in basic biology.

There are about 20,000 genes in the human body. But they are surrounded by other DNA material whose exact purpose is unclear. Roughly 1 percent of the human genome is thought to be "protein-coding" -- that is, genes. Another 4 percent had been thought to be "non coding DNA" that serves as on-off switches for the genes, and the rest was seen as a sort of swamp with no clear purpose.

But the new work suggests that the "control regions" in the DNA are far more extensive, perhaps embracing more than half of all DNA. Functions thought to be carried out by genes alone now appear to be managed by multiple, overlapping segments of DNA. In addition, other portions of the genome are believed to be on standby, as a toolbag to be utilized as humans evolve.

"It's like clutter in the attic," said Collins. "Most of the time, the human genome is operating on the 'first and second floor,' with 5 percent of the genome doing what needs to be done on a daily basis. But over

evolutionary time, a much larger part of the genome, the stuff in the attic, becomes important. It's waiting for natural selection to call for it."

The ENCODE research builds on the historic Human Genome Project, largely completed in 2003, which cataloged the genes. Instead of the "big picture" look at the entire structure, the ENCODE project fine-combed selected sites in the genome in extraordinary detail. Half the sites were known by scientists to affect gene replication and protein coding; the other half were random samples from across the genome, including swatches of "junk."

A long standing assumption in genetics has been that cellular organisms are run by genes, which instruct cells to produce proteins thought to be the main driving mechanism in cells. But according to the study, obscure sections of the genome, the "junk DNA," may play an even more critical role in health and evolution than genes themselves.

"We're reshaping our understanding of which regions of the genome produce the critical information" that allows organisms to function and evolve, said Michael Snyder, professor of molecular biophysics at Yale University and one of the researchers.

Recent research into heart attacks and diabetes has made the startling discovery that the roots of disease may lie in noncoding portions of DNA, not in the genes themselves.

In a significant finding, researchers discovered that "gene transcription" -- essential to the process by which DNA builds proteins indispensable to life -- is occurring in regions between genes. They found that ribonucleic acid, or RNA, long seen as another type of genetic code that directs cellular machinery to make proteins, is also produced in stretches of the genome not involved in protein production. That suggests that these regions have an important purpose, though still not understood, the scientists said.

"Transcription appears to be far more interconnected across the genome than anyone had thought," said Collins, adding that the ENCODE findings are "moving us into a deeper understanding of how life works and how, sometimes, things go wrong and disease occurs."

But untangling the tantalizing implications of the new findings will be the work of years.

"It's like reading a code, text jumbled together, and you're trying to make sense of it," said Zhiping Weng, professor of biomedical engineering at Boston University and a researcher in the study. "This project provides many new insights into the complex functional landscape of the human genome." ■

© Copyright 2007 The New York Times Company