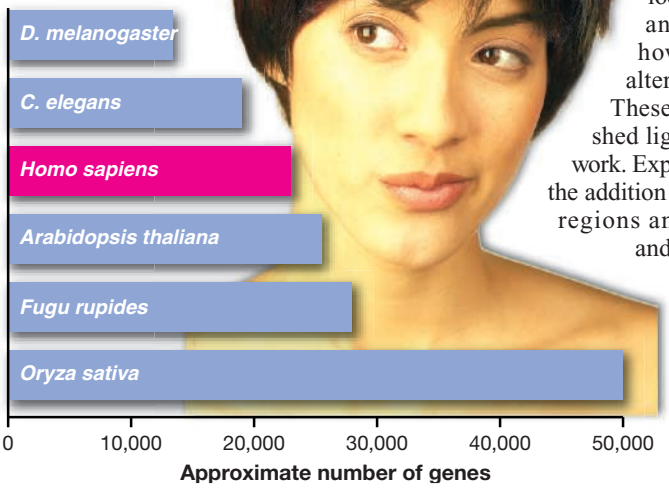
Why Do Humans Have So Few Genes

When leading biologists were unraveling the sequence of the human genome in the late 1990s, they ran a pool on the number of genes contained in the 3 billion base pairs that make up our DNA. Few bets came close. The conventional wisdom a decade or so ago was that we need about 100,000 genes to carry out the myriad cellular processes that keep us functioning. But it turns out that we have only about 25,000 genes—about the same number as a tiny flowering plant called *Arabidopsis* and barely more than the worm *Caenorhabditis elegans*.

That big surprise reinforced a growing realization among geneticists: Our genomes and those of other mammals are far more flexible and complicated than they once seemed. The old notion of one gene/one protein has gone by the board: It is now clear that many genes can make more than one protein. Regulatory proteins, RNA, noncoding bits of DNA, even chemical and structural alterations of the genome itself control how, where, and when genes are expressed. Figuring out how all these elements work together to choreograph gene expression is one of the central challenges facing biologists.

In the past few years, it has become clear that a phenomenon called alternative splicing is one reason human genomes can produce such complexity with so few genes. Human genes contain both coding DNA—exons—and noncoding DNA. In some genes, different combinations of exons can become active at different times, and each combination yields a different protein. Alternative splicing was long considered a rare hiccup during transcription, but researchers have concluded that it may occur in half—some say close to all—of our genes. That finding goes a long way toward explaining how so few genes can produce hundreds of thousands of different proteins. But how the transcription machinery decides which parts of a gene to read at any particular time is still largely a mystery.

The same could be said for the mechanisms that determine which genes or suites of genes are turned on or off at particular times and places. Researchers are discovering that each gene needs a supporting cast of hundreds to get its job done. They include proteins that shut down or activate a gene, for example by adding acetyl or methyl groups to the DNA. Other proteins, called transcription factors, interact with the genes more directly: They bind to landing sites situated near the gene under their control. As with alternative splicing, activation of different combinations of landing sites makes possible exquisite control of gene expression, but researchers have yet to figure out exactly how all these regulatory elements really work or how they fit in with alternative splicing.



In the past decade or so, researchers have also come to appreciate the key roles played by chromatin proteins and RNA in regulating gene expression. Chromatin proteins are essentially the packaging for DNA, holding chromosomes in well-defined spirals. By slightly changing shape, chromatin may expose different genes to the transcription machinery.

Genes also dance to the tune of RNA. Small RNA molecules, many less than 30 bases, now share the limelight with other gene regulators. Many researchers who once focused on messenger RNA and other relatively large RNA molecules have in the past 5 years turned their attention to these smaller cousins, including microRNA and small nuclear RNA. Surprisingly, RNAs in these various guises shut down and otherwise alter gene expression. They also are key to cell differentiation in developing organisms, but the mechanisms are not fully understood.

Researchers have made enormous strides in pinpointing these various mechanisms. By matching up genomes from organisms on different branches on the evolutionary tree, genomicists are locating regulatory regions and gaining insights into how mechanisms such as alternative splicing evolved. These studies, in turn, should shed light on how these regions work. Experiments in mice, such as the addition or deletion of regulatory regions and manipulating RNA, and computer models should also help. But the central question is likely to remain unsolved for a long time: How do all these features meld together to make us whole?

—ELIZABETH PENNISI

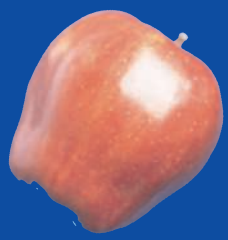
CREDIT: JUPITER IMAGES

Why is there more matter than antimatter?
To a particle physicist, matter and antimatter are almost the same. Some subtle difference must explain why matter is common and antimatter rare.

JUPITER IMAGES

Does the proton decay?
In a theory of everything, quarks (which make up protons) should somehow be convertible to leptons (such as electrons)—so catching a proton decaying into something else might reveal new laws of particle physics.

JUPITER IMAGES



What is the nature of gravity?
It clashes with quantum theory. It doesn't fit in the Standard Model. Nobody has spotted the particle that is responsible for it. Newton's apple contained a whole can of worms.

Why is time different from other dimensions?
It took millennia for scientists to realize that time is a dimension, like the three spatial dimensions, and that time and space are inextricably linked. The equations make sense, but they don't satisfy those who ask why we perceive a "now" or why time seems to flow the way it does.

Published by AAAS