

paradigm.

A political science

Once, biology was left to researchers—
now, everyone wants to get into the game



DO
NO
HARM



vote
YES
on stem cell
research

[ONE PERSON'S JUNK IS ANOTHER'S TREASURE]

Sorting through the human genome's vast wastelands

Nobelist Sydney Brenner had a pithy description for the huge stretches of the human genome that don't seem to code for genes: "junk DNA." Since he coined the term, scientists have debated which of these hapless-looking strands of DNA are really junk and which are key players in biological or disease processes we don't yet understand. David Haussler, director of the Center for Biomolecular Science & Engineering at the University of California, Santa Cruz, explains recent advances in sorting out our junk.

How do you separate genes from "junk"?

We haven't located every human gene precisely, but we're getting close, with somewhere between 20,000 and 25,000 protein-coding genes. Altogether, those genes account for less than 1.5 percent of the DNA in the human genome. Even the simplest species can't be pure protein-coding; there's got to be some DNA that helps to guide the process of making the protein-coding genes. But there seems to be extra bloat of material in the human genome, beyond what you would expect is needed to guide the protein-coding genes.

What do comparisons with other mammals tell us?

Most recently, we looked at the most extremely conserved things in the human genome—everything that was completely identical in the human, mouse, and rat for hundreds of DNA bases. No one knows what these segments are doing. Most of them are not making proteins. But they're definitely not junk. Those similarities have been preserved through evolution, probably because those pieces of DNA are important.

It's amazing to think about how conserved these segments are, and it makes it all the more exciting to try to figure out what they do. We have some indications that they regulate genes in the neighborhood. What's amazing is that the neighborhood can be up to a million bases away from the gene. And a very statistically significant number of these ultra-conserved elements are near genes involved in embryo development.

Overall, while it's not all ultra-conserved, another 3.5 percent of the human genome is more conserved with other species than was expected. Most of this additional conserved DNA could be regulating the protein-coding genes, but this remains to be proven. Much of it could also be making RNAs that don't code for proteins. We just don't know yet.

What else is lurking in the "junk"?

There's a lot of DNA in the human genome that seems to be evolving neutrally, just accumulating mutations through the eons. It's been called selfish DNA.



BRIAN WILLISE

Occasionally a segment of selfish DNA can make a copy of itself in a cell that's destined to become a sperm or an egg. That copy gets inserted back at a different place in the genome. If that happens, then you pass on an extra piece in your genome to your child. Through millions of years, you end up with more and more copies of these things in the genome. More than half of the human genome is stuff left over from these "transposons." They're essentially the rotting carcasses of these old selfish DNA elements. And you carry them in your genome and pass them on to your children!

Do transposons play active roles?

There was an exciting paper in *Nature* this year about one kind of transposon called long interspersed nuclear elements [LINEs]. When LINEs make copies, if they randomly insert themselves in between the protein-coding parts of a gene, they can slow down the rate at which you use that gene. Also, they occasionally can take a bit of a nearby gene with them, and when they make the copy, they put down a new bit of gene somewhere else in the genome.

Beyond genes and transposons, what makes up the rest of our DNA?

We don't know about the in-between stuff. Some will be transposons that have decayed so much from their original versions that we can't even recognize them. But some may be DNA that's doing important stuff that we just haven't recognized yet.