ENCODE: the rough guide to the human genome

by Ed Yong

Back in 2001, the Human Genome Project gave us a near complete readout of our DNA. Somehow, those As, Gs, Cs, and Ts contained the full instructions for making one of us, but they were hardly a simple blueprint or recipe book. The genome was there, but we had little idea about how it was used, controlled or organized, much less how it led to a living, breathing human.

That gap has just got a little smaller. A massive international project called ENCODE – the Encyclopedia Of DNA Elements – has moved us from “Here’s the genome” towards “Here’s what the genome does”. Over the last 10 years, an international team of 442 scientists have assailed 147 different types of cells with 24 types of experiments. Their goal: catalogue every letter (nucleotide) within the genome that does something.

For years, we’ve known that only 1.5 percent of the genome actually contains instructions for making proteins, the molecular workhorses of our cells. But ENCODE has shown that the rest of the genome – the non-coding majority – is still rife with “functional elements”. That is, it’s doing something. According to ENCODE’s analysis, 80 percent of the genome has a “biochemical function”, but the key point is: It’s not “junk”.

And what’s in the remaining 20 percent? Possibly not junk either, according to Ewan Birney, the project’s Lead Analysis Coordinator and self-described “cat-herder-in-chief”. He explains that ENCODE only (!) looked at 147 types of cells, and the human body has a few thousand. A given part of the genome might control a gene in one cell type, but not others.

That the genome is complex will come as no surprise to scientists, but ENCODE does two fresh things: it catalogues the DNA elements for scientists to pore over; and it reveals just how many there are. “The genome is no longer an empty vastness – it is densely packed with peaks and wiggles of biochemical activity,” says Shyam Prabhakar.

Think of the human genome as a city. The basic layout, tallest buildings and most famous sights are visible from a distance. That’s where we got to in 2001. Now, we’ve zoomed in. We can see the players that make the city tick: the cleaners and security guards who maintain the buildings, the sewers and power lines connecting distant parts, the police and politicians who oversee the rest. That’s where we are now: a comprehensive 3-D portrait of a dynamic, changing entity, rather than a static, 2-D map.

And just as London is not New York, different types of cells rely on different DNA elements. For example, of the roughly 3 million locations where proteins stick to DNA, just 3,700 are commonly used in every cell examined. Liver cells, skin cells, neurons, embryonic stem cells... all of them use different suites of switches to control their lives. Again, we knew this would be so. Again, it’s the scale and the comprehensiveness that matter.

Where will it lead us? It’s easy to get carried away, and ENCODE’s scientists seem wary of the hype-and-backlash cycle that befell the Human Genome Project. Much was promised at its unveiling, by both the media and the scientists involved, including medical breakthroughs and a clearer understanding of our humanity. The ENCODE team is being more cautious. “This idea that it will lead to new treatments for cancer or provide answers that were previously unknown is at least partially true,” says Gingeras, “but the degree to which it will successfully address those issues is unknown.

“We are the most complex things we know about. It’s not surprising that the manual is huge,” says Birney.
Science in the News #10

1. Read and ‘talk’ with the text on the back of this page

2. Complete all of the sentences below

The article that I just read is about____________________________________
______________________________________________________________
______________________________________________________________
I think I am reading this article because in ___________ class we are/were studying____________________________________
______________________________________________________________
______________________________________________________________
One word from the article I did not know the meaning of was_________________________. When I looked this word up in the dictionary I find out it meant____________________________________
______________________________________________________________
The most important sentence in the article to me is “____________________________________
______________________________________________________________
This quote means____________________________________
I think this quote is the most important sentence in the article because____________________________________

After reading the article I wonder____________________________________?
Another question I still have after reading the article is____________________________________?

To ‘talk’ with the text on back:
1. underline: important facts
2. ⭐️: interesting ideas
3. 🌑: vocabulary
4. ?? : What you want to know more/ don’t understand

summarizing making connections Research Important details questioning