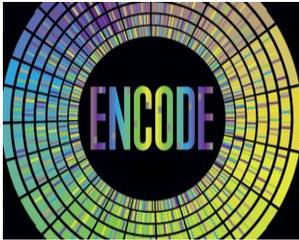


Three reasons why junk DNA makes evolutionary sense

By [Ashutosh Jogalekar](#) | September 13, 2012



ENCODE (Image: Ed Yong)

The recent dustup over the ENCODE project and its confusing finding that “80% of DNA is functional” surprises me greatly. What surprises me especially is that people are surprised by junk DNA. **Unfortunately this time the scientists are also culpable since, while the publicity surrounding ENCODE has been a media disaster, the 80% claim originated in the scientific papers themselves.** There is no doubt that the project itself – which represents a triumph of teamwork, dogged pursuit, technological mastery and first-rate science – has produced enormously useful data, and there is no doubt it will continue to do so. **What is in doubt is how long it will take for the public damage to be repaired.**

There’s a lot written about the various misleading statements about the project made by both scientists and journalists and I cannot add much to it. All I can do is to point to some excellent articles: [Larry Moran](#) has waged a longstanding effort to spread the true wisdom about junk DNA for years on his blog. [Ed Yong](#) exhaustively summarizes a long list of opinions, links and analysis. [T. Ryan Gregory](#) has some great posts dispelling the myth of the myth of junk DNA. And [John Timmer](#) has the best popular account of the matter. **The biggest mistake on the part of the scientists was to define “functional” so loosely that it could mean pretty much all of DNA. The second big mistake was not in clarifying what functional means to the public.**

But what **I found astonishing was why it’s so hard for people to accept that much of DNA must indeed be junk.** Even to someone like me who is not an expert, the existence of junk DNA appeared perfectly normal. I think that junk DNA shouldn’t shock us at all if we accept the standard evolutionary picture.

The **standard evolutionary picture tells us that evolution is messy, incomplete and inefficient. DNA consists of many kinds of sequences. Some sequences have a bonafide biological function in that they are transcribed and then translated into proteins that**

have a clear physiological role. Then there are sequences which are only transcribed into RNA which doesn't do anything. There are also sequences which are only bound by DNA-binding proteins (which was one of the definitions of "functional" the ENCODE scientists subscribed to). Finally, there are sequences which don't do anything at all. Many of these sequences consist of pseudogenes and transposons and are defective and dysfunctional genes from viruses and other genetic flotsam, inserted into our genome through our long, imperfect and promiscuous genetic history. If we can appreciate that evolution is a flawed, piecemeal, inefficient and patchwork process, we should not be surprised to find this diversity of sequences with varying degrees of function or with no function in our genome.

The reason why most of these useless pieces have not been weeded out is simply because there was no need to. We should remember that evolution does not work toward a best possible outcome, it can only do the best with what it already has. It's too much of a risk and too much work to get rid of all these defective and non-functional sequences if they aren't a burden; the work of simply duplicating these sequences is much lesser than that of getting rid of them. Thus the sequences hung around in our long evolutionary history and got passed on. The fact that they may not serve any function at all would be perfectly consistent with a haphazard natural mechanism depending on chance and the tacking on of non-functionality to useful functions simply as extra baggage.

There are two other facts in my view which should make it very easy for us to accept the existence of junk DNA. **Consider that the salamander genome is ten times the size of the human genome.** Now this implies two possibilities; either salamanders have ten times functional DNA than we do, or that the *main difference between us and salamanders is that they have much more junk DNA*. Wouldn't the complexity of salamander anatomy of physiology be vastly different if they really had so much more functional DNA? On the contrary, wouldn't the relative simplicity of salamanders compared to humans be much more consistent with just varying degrees of junk DNA? Which explanation sounds more plausible?

The third reason for accepting the reality of junk DNA is to simply think about mutational load. Our genomes, as of other organisms, have undergone lots of mutations during evolution. What would be the consequences if 90% of our genome were really functional and had undergone mutations? How would we have survived and flourished with such a high mutation rate? On the other hand, it's much simpler to understand our survival if we assume that most mutations that happen in our genome happen in junk DNA.

As a summary then, we should be surprised to find someone who says they are surprised by junk DNA. Even someone like me who is not an expert **can think of at least three simple reasons to like junk DNA:**

1. The understanding that evolution is an inherently *messy* and *inefficient* process that often produces junk. This junk may be retained if it's not causing trouble.

2. The realization that the vast differences in *genome sizes* are much better explained by junk DNA than by assuming that most DNA is truly functional.

3. The understanding that *mutational loads* would be prohibitive had most of our DNA not been junk.

Finally **as a chemist**, let me say that **I don't find the binding of DNA-binding proteins to random, non-functional stretches of DNA surprising at all**. That hardly makes these stretches physiologically important. If evolution is messy, chemistry is equally messy. **Molecules stick to many other molecules**, and not every one of these interactions has to lead to a physiological event. DNA-binding proteins that are designed to bind to specific DNA sequences would be expected to have some affinity for non-specific sequences just by chance; **a negatively charged group could interact with a positively charged one, an aromatic ring could insert between DNA base pairs and a greasy side chain might nestle into a pocket by displacing water molecules**. It was a pity the authors of ENCODE decided to define biological functionality partly in terms of chemical interactions which may or may not be biologically relevant.

The dustup from the ENCODE findings suggests that scientists continue to find order and purpose in an orderless and purposeless universe which can nonetheless produce structures of great beauty. They would like to find a purpose for everything in nature and are constantly looking for the signal hidden in the noise. Such a quest is consistent with our ingrained sense of pattern recognition and has often led to great discoveries. But the stochastic, contingent, haphazard meanderings of nature mean that sometimes noise is just that, noise. It's a truth we must accept if we want to understand nature as she really is.



About the Author: Ashutosh (Ash) Jogalekar is a chemist interested in the nature of the "central science" and its intersection with philosophy, history and culture. He is also more generally interested in the history and philosophy of science and is particularly fascinated by how science tries to mirror reality by building models. Follow on Twitter [@curiouswavefn](https://twitter.com/curiouswavefn).