

Dieu My Nguyen
MCDB 5520
April 23, 2018
Extra credit essay

How much of the human genome is functional? The Encyclopedia of DNA Elements (ENCODE) project estimated as high as 80% based on data of mapped regions of transcription, transcription factor association, chromatin structure, and histone modification¹². The ENCODE researchers asserted this number from their finding that the majority (80.4%) of the human genome participates in at least one biochemical RNA and/or chromatin associated event in at least one cell type. Critical scrutiny of their results include the semantic difference between biochemical activity and biological function.³

More recently, Dan Graur from the University of Houston suggested that about 10 to 15% of the genome has a function, at the upper limit of 25%, using deleterious mutation rate and the replacement fertility rate⁴. He defined the functional portion of the genome as that which has a selected-effect function, arising through natural selection. These functions include gene that code for proteins, genes that specify RNA, and DNA receptors. Only functional parts can be damaged by deleterious mutations. Because of these mutations, each couple in the world must produce slightly more than 2 children to maintain a constant population size. Based on Graur's model, if 80% of the human genome is functional, couples would each have an average of 15 children and 13 would have to die or fail to reproduce, inconsistent with historical data of human population which shows fertility rates from 2.1 to 3.9 children per couple for the past 200,000 years.

I don't have opinions that are well thought out and certainly no years of experimentation or mathematical models to present, but here, my (possibly uneducated) guess. ENCODE defined "functional" to mean the genome portions that have some biochemical activity, while critiquers uses the word to only count DNA that is acted upon by natural selection and contributes to phenotypes. I will (arbitrarily) estimate that around 10% of the human genome is "functional," using the latter, stricter definition of the term. ENCODE reported seeing 8% of places in the genome with "specific DNA:protein contact."⁵ If we also count exons, we have 9% total. These regulatory sequences may be regions subject to evolutionary action. Is the remaining 90% of DNA junk? That term seems too certain and branding for DNA that we simply haven't understood well. Whatever the estimate of "functional" portion of the genome now, we should perform knockouts on these regions of "junk" and see whether there are hidden treasures. Overall, this is my guess, and I am completely okay with being completely wrong.

¹ ENCODE Project Consortium (2004). The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* 306, 636–640.

² Kellis M (2013) Defining functional DNA elements in the human genome. *PNAS* 111(17):6131-6138.

³ Brunet TDP and Doolittle WF (2014) Getting "function" right. *PNAS* 111(33):E3365

⁴ Graur D (2017) An Upper Limit on the Functional Fraction of the Human Genome. *Genome Biol Evol* 9(7):1880-1885.

⁵ Q&A with ENCODE by Ewan Birney: <http://ewanbirney.com/2012/09/encode-my-own-thoughts.html>