Genomes, Genes and Junk DNA

Timothy G. Standish

The discovery that genomes contain far fewer genes than expected has aggravated a problem recognized a generation of scientists ago: if there are so few genes, why do humans and many other organisms have so much DNA in their genomes? Only 3% of the 3 billion nucleotide letters in the human genome actually coded for proteins. Thus, most of our genome codes for nothing, or so it once seemed. This DNA that does not code for proteins was dismissively referred to as "junk DNA" and a number of prominent Darwinists and their followers jumped on the "junk DNA" bandwagon claiming that it was exactly what the messy process of Darwinian evolution predicts.[1] Susumu Ohno, who is widely credited with coining the term "junk DNA" put it this way:

> The earth is strewn with fossil remains of extinct species; is it a wonder that our genome too is filled with the remains of extinct genes? [2]

Some went beyond the idea that "junk DNA" is a legacy of our evolutionary past, a sort of molecular trash midden that can be excavated to gain information about an organism's evolutionary past. For example, Richard Dawkins, the proselytizing atheist and uber-Darwinist, co-opted junk DNA as *prima facie* evidence in support of his reductionistic "selfish gene theory" of evolution:

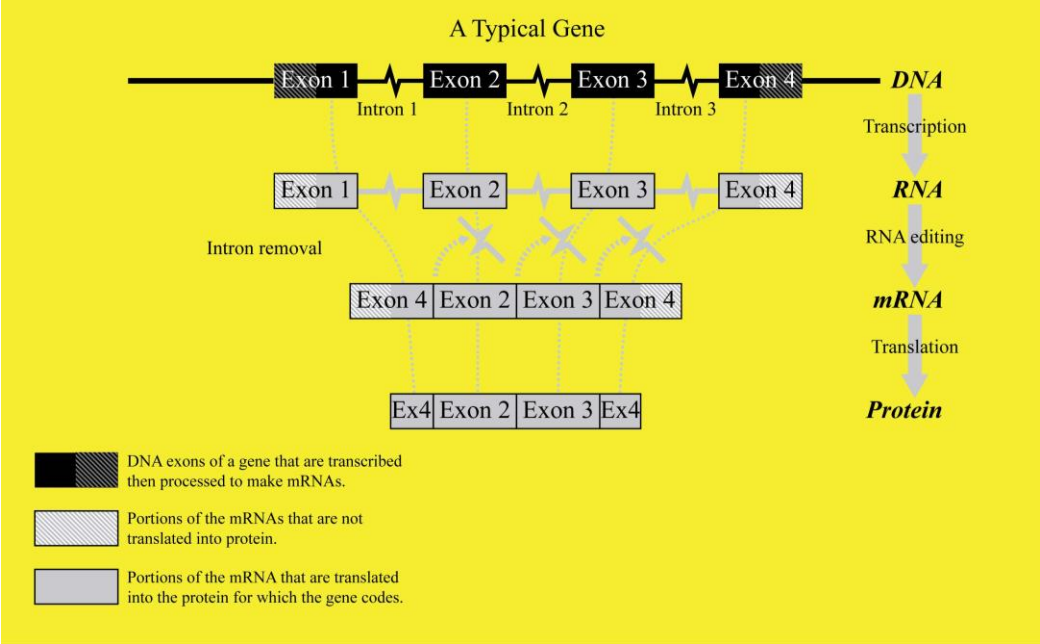
> Biologists are racking their brains trying to think what useful task this apparently surplus DNA is doing. But from the point of view of the selfish genes themselves, there is no paradox. The true "purpose" of DNA is to survive, no more and no less. The simplest way to explain the surplus DNA is to suppose that it is a parasite, or at best a harmless but useless passenger, hitching a ride in the survival machines created by the other DNA.[3]

The functionlessness of junk DNA led Dawkins to taunt creationists with the suggestion that they "might spend some earnest time speculating on why the Creator should bother to litter genomes with untranslated pseudogenes and junk tandem repeat DNA."[4] And yet, as a better understanding develops, genomes appear far more elegant than originally appreciated and apparently predicted by some Darwinists. Darwinism can accommodate design evident in genomes in the same way that other evidence of design is accommodated; by calling it "apparent" design rather than real design. Still, discovery of function in "junk DNA" calls into question Dawkins' "selfish gene" speculation along with many other ideas dependent on significant amounts of DNA actually being functionless. Further, if Darwinism explains functionality in "junk DNA" just as well as it explains functionless DNA, then it is reasonable to conclude that at least sometimes it concurrently explains everything and nothing.

In the old understanding of genes, Beadle and Tatum's Nobel Prize winning "one gene, one enzyme (protein)" concept prevailed.[5] Because humans produce more than 100,000 proteins and there appear to be less than 25,000 genes,[6] at least some genes must be capable of producing more than one protein. How might this be achieved?

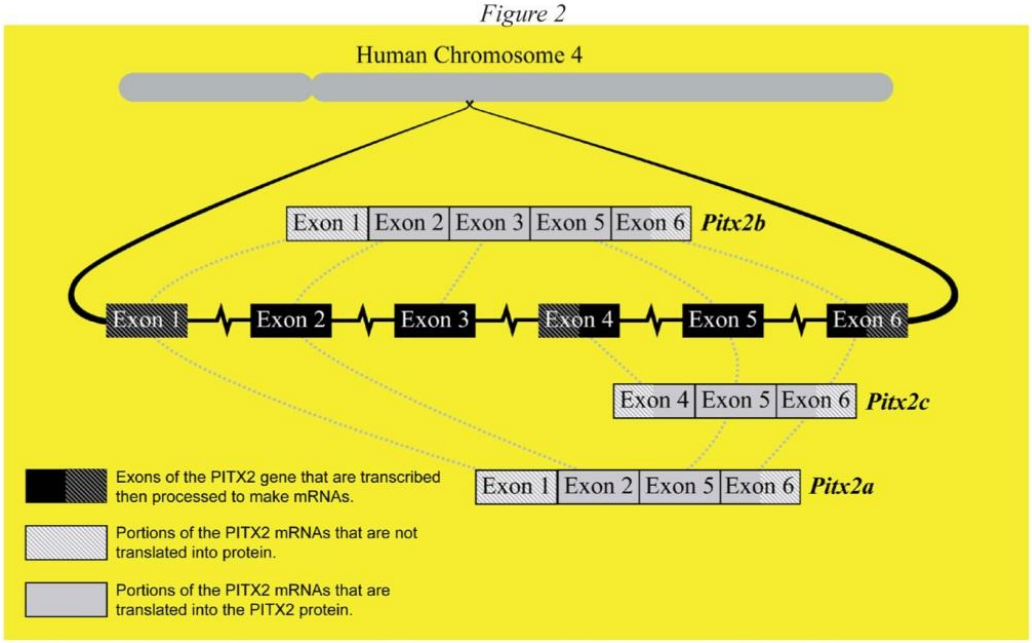
Most genes in humans and other organisms with complex cells (as well as some genes in bacteria) are made up of DNA segments called "exons" separated by intervening segments called "introns" (see Figure 1). When the protein that a gene codes for is going to be made, the first step involves making an RNA copy of the gene. This RNA transcript is then edited to remove the introns so that the protein coding exons are contiguous in mRNA. It is the mRNA that carries information out of a cell's nucleus to molecular factories called ribosomes that translate the mRNA information into specific proteins. By joining different protein coding exons from the same gene together, different mRNAs can be made and different proteins result.

Figure 1



In most complex organisms, a typical gene is composed of multiple parts only some of which actually code for a protein. The protein coding parts, called "exons" are usually separated by intervening segments called "introns." When the information coded in a gene is to be used to make a protein, the DNA is first transcribed into RNA which is then edited to remove the introns. This edited RNA, now called mRNA, then carries the information that is translated by complex cellular machinery into a protein.

Figure 2



The human *Pitx2* gene is composed of 6 exons separated by 5 introns. Joining different exons together in RNA transcripts of *Pitx2* and using different transcription and translation start sites results in several different proteins from this single gene. In this illustration, exons and introns are not drawn to scale. In the actual gene, exons vary significantly in size and introns make up the bulk of the gene. Unique mRNAs are shown for PITX2A, B and C after processing to remove introns and other RNA.

The human *Paired*-like homeo-domain transcription factor 2 (*Pitx2*) gene illustrates how RNA processing works to create several different proteins from one gene. The PITX2 protein plays a role in proper development of the head, eyes and other things.[7] *Pitx2* is composed of six exons separated by five introns (Figure 2). By joining exons 1, 2, 5 and 6 together, the mRNA for a version of PITX2 called "Isoform A", or PITX2A, is created. Joining together exons 1, 2, 3, 5 and 6 makes PITX2B mRNA and exons 4, 5 and 6 makes PITX2C. Changing parts of the protein presumably impacts how PITX2 interacts with other molecules.[8] There are many further interesting complexities involving *Pitx2*, but the point of this example is that a single gene can be used to make multiple different proteins. If that is the case, regulatory mechanisms are necessary to ensure the right proteins are constructed using the right genes.

What does this have to do with "junk DNA"? Genomes are now understood to be far more dynamic than initially imagined. While genes are less numerous than expected, they are very complex in their structure, expression and associated control systems. Information to control how genes are to be expressed has to come from somewhere. Some of that information may be imbedded in the genes themselves, but much of it clearly lies outside genes and in the DNA once thought of as leftover rubbish from the sloppy evolutionary process. To the surprise of many, much of what was once dismissed as "junk DNA" is now turning out to play a vital role in the normal function of genetic systems. As long ago as 1994, my colleague Jim Gibson, commenting on a specific class of junk DNA called "pseudogenes" noted: "[T]he argument that particular DNA sequences must not have a function because we haven't discovered any function for them is an argument from silence. To conclude that pseudogenes are junk DNA seems premature."[9]

Premature it was, both to hastily arrive at the conclusion that DNA not coding for proteins must do nothing and that this somehow supports evolutionary ideas while calling creation into question. Over the past few years "junk DNA" has yielded a treasure-trove of information about how genomes operate. The complex systems controlling exon splicing appear to be involved in sequences occupying at least one third of the human genome.[10] That is a lot more than the 3% of the human genome that was thought to be functional only a few years ago. Various small RNA transcripts appear to be involved in regulating every step in protein production. These short RNAs appear to come from all over the genome, not just areas coding for genes. In fact, it is becoming apparent that at least 70% of the genome is transcribed into RNA[11] and that both strands of the double stranded DNA molecule are commonly transcribed, not just the strand that codes a protein.[12]

Even some genes that were once thought to be broken remnants of once functional genes, the pseudogenes Gibson was referring to, have been found to have essential functions. For example, one of these "pseudogenes" appears to be essential for egg formation in mice.[13] Discovering that some pseudogenes have a function is genuinely surprising and shows how much we still have to learn about genomes. If pseudogenes have vital functions, this calls into question the logic used when invoking pseudogenes and other "junk DNA" as evidence of common ancestry.

Old arguments against a wise and good Creator God based on the false assumption that genomes are primarily leftover litter from the process of evolution are no longer tenable. That is not to say that all DNA must be perfectly functional and that we must understand that function. Some degree of degeneration in genomes since the fall is to be expected, but exactly how much remains open to question. Revisions in our understanding of how genomes work reveal an awe-inspiring level of sophistication in their design, even if some parts may be broken. It turns out that control mechanisms encoded in "junk DNA" are as important as the genes they regulate and that humans, along with all other living things, really are "fearfully and wonderfully made."[14]

Timothy G. Standish is a research scientist at the Geoscience Research Institute

[1]For a detailed discussion of the history of this argument and why, from a Darwinian perspective, it is weak, see T. G. Standish, "Rushing to Judgment: Functionality in Non-Coding or Junk DNA," *Origins* 53 (2002): 7–30.

[2]S. Ohno, "So Much 'Junk' DNA in Our Genome" in *Evolution of Genetic Systems,* ed. H. H. Smith (Brookhaven Symposia in Biology 23; New York: Gordon and Breach, 1972), 366–70.

[3]R. Dawkins, *The Selfish Gene* (rev. ed.; Oxford: Oxford University Press, 1989), 45. Note that this quote is not altered from the first edition of the Selfish Gene printed in 1976 and that the entire book can be viewed online at: http://macroevolution.narod.ru/gene/index_2en.html.

[4]R. Dawkins, "The 'Information Challenge': How Evolution Increases Information in the Genome," *Skeptic* 7 (1999): 64–69.

[5]George W. Beadle and Edward L. Tatum each shared one quarter of the 1958 Nobel Prize in Physiology or Medicine, the remaining half awarded to Joshua Lederberg for his discoveries related to bacterial genetic recombination.

[6]Despite publication of the human genome, estimating gene numbers in the genome is still an estimation, not an actual count of the genes. These estimates are based on certain assumptions that may or may not be valid. For a recent example of estimating gene numbers that includes certain evolutionary assumptions and comes up with a remarkably low estimate see M. Clamp, B. Fry, M. Kamal, X. Xie, J. Cuff , M. F. Lin, M. Kellis, K. Lindblad-Toh, and E. S. Lander, "Distinguishing Protein-coding and Noncoding Genes in the Human Genome," *Proceedings of the National Academy of Sciences USA* 104/49 (2007): 19428–19433.

[7]P. J. Gage, H. Suh, and S. A. Camper, "The Bicoid-related Pitx Gene Family in Development," *Mammalian Genome* 10 (1999): 197–200.

[8]P. Lamba, T. A. Hjalt, and D. J. Bernard. "Novel Forms of Paired-like Homeodomain Transcription Factor 2 (PITX2): Generation by Alternative Translation Initiation and mRNA Splicing," *BMC Molecular Biology* 9 (2008): no pages. Cited 3 Dec 2008. Online: http://www.biomedcentral.com/1471-2199/9/31.

[9]L. J. Gibson, "Pseudogenes and Origins," *Origins* 21/2 (1994): 91–108.

[10]C. Zhang, W.-H. Li, A. R. Krainer, and M. Q. Zhang. "RNA Landscape of Evolution for Optimal Exon and Intron Discrimination," *Proceedings of the National Academy of Sciences USA* 105/15 (2008): 5797–5802.

[11]M. Pheasant and J. S. Mattick, "Raising the Estimate of Functional Human Sequences," *Genome Research* 17 (2007): 1245–1253.

[12]RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group) and the FANTOM Consortium, "Antisense Transcription in the Mammalian Transcriptome," *Science* 309 (2005):1564–1566.

[13]O. H. Tam, A. A. Aravin, P. Stein, A. Girard, E. P. Murchison, S. Cheloufi, E. Hodges, M. Anger, R. Sachidanandam, R. M. Schultz, and G. J. Hannon. "Pseudogene-derived Small Interfering RNAs Regulate Gene Expression in Mouse Oocytes," *Nature* 453 (2008): 534-38. E-published April 10, 2008: http://dx.doi.org/10.1038/nature06904.

[14]Psalm 139:14 KJV and NIV.

1/09