

# Junk DNA indicted

John Woodmorappe

Discoveries of function in erstwhile junk DNA are occurring at an ever-increasing pace. It is now realized that DNA has numerous functions beyond that of encoding peptides. A previously unsuspected world of widespread noncoding RNA, transcribed from intergenic DNA, introns and pseudogenes, has been discovered. This form of RNA appears to regulate many genes, and may even be the very foundation of human development. Antisense transcription of DNA is far more common in humans than supposed until recently.

DNA that appears to lack conservation of sequence is probably nevertheless functional in terms of general isochore composition (serving as 'background' for proper gene function), as a spacer between genes and their long-distance regulatory elements, and, in the case of introns, in terms of physical length alone. Monotonous DNA repeats (STRs) probably serve a regulatory function for certain genes. Still other potential functions for junk DNA, such as those related to 'epigenetic' control of DNA, await clarification. It is easy to see that the junk DNA concept, whose longevity owes at least partly to tacit evolutionary assumptions, has very much delayed our understanding of the genome.

One of the pillars of evolutionary thought has been the premise that living things are the products of accumulated accidents, not foreordained design. It is therefore hardly surprising that evolutionists have historically been prone to assume that living things are full of features that may have once been useful to the organism but no longer serve a function.

When the human genome was first studied decades ago, it was noted that the DNA molecule served as a template for the synthesis of proteins. Researchers soon determined that only a small fraction of it (less than about 3%<sup>1</sup>) appeared to perform this role in the human genome, leading them to suspect that the other 97% (noncoding DNA) was a useless evolutionary leftover—hence junk DNA (figure 1, top). Similar conclusions were reached for the DNA of many other organisms. However, over the years, more and more stretches of noncoding DNA have been found to be functional.<sup>2</sup> Previous reviews, from a scientific creationist perspective, have been provided by Walkup,<sup>3</sup> Standish<sup>4</sup> and Batten.<sup>5</sup> Owing to the rapidity of discoveries relative to junk DNA, an update of developments in this area is needed. The present review article is written with this in mind.

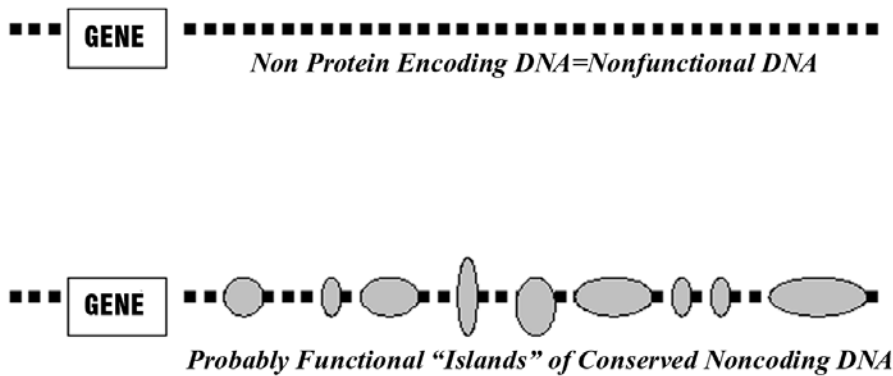
A recent article in *Scientific American*,<sup>6</sup> intended for the semi-technical reader, calls attention to some recent developments of our understanding of junk DNA. Moreover, it shows the flawed reasoning that has sustained the idea of junk DNA for so long, and exposes the hindering effects this concept has held over the biological sciences—doing so with an almost excoriating tone of writing.

Owing to the breadth of this topic, discussion herein centers upon intergenic DNA (the DNA situated between protein-encoding genes), pseudogenes (allegedly functionless remnants of protein-encoding genes), and introns (those stretches of a protein-encoding gene's DNA whose RNA transcripts are excised by the cell machinery before the mature RNA transcript is translated into a peptide; figure 2). Other forms of junk DNA are briefly mentioned.

## Sequence conservation: not limited to protein-encoding DNA

Evolutionists have long been comparing the DNA sequences of different organisms. They came to believe that close similarities of given large DNA segments between different organisms implied that the DNA was conserved and therefore functional. They reasoned that natural selection would tend to remove most alterations of such sequences. Sequence conservation came to be closely associated with the DNA sequences of protein-encoding genes (figure 1, top). Conversely, the dissimilarities in the non protein-encoding DNA sequences between organisms led evolutionists to suppose that such sequences were being freely altered by mutations, over the eons, without incurring the penalty of natural selection (figure 1, top, dashes). The absence of function for noncoding<sup>7</sup> sequences seemed self-evident from its apparent large-scale selective neutrality, and the epithet 'junk DNA' became common. In addition, the ability of scientists to freely alter or remove such stretches of noncoding DNA was (incorrectly) taken as further proof that noncoding DNA is devoid of function.<sup>8</sup>

Influenced by the evolutionary premise of protein-encoding DNA being just about the only functional DNA in existence, researchers were surprised to learn that there are thousands of short, conserved blocs of junk DNA in common between the human and mouse genomes (figure 1, bottom).<sup>9</sup> Subsequent research has uncovered the astonishing fact that most of them are conserved across various mammalian orders.<sup>10</sup> These conserved blocs are probably structural and regulatory in nature, and unlikely to be, for the most part at least, previously unknown protein-encoding gene sequences. In fact, the most strongly conserved ones exhibit the signature of protein-binding sites,<sup>10</sup> consistent with a regulatory function. Some of these blocs, ironically the least conserved ones between mammalian classes,<sup>10</sup> are partly responsible for the encoding for a 'free' type of RNA that is described in the next paragraph.



**Figure 1.** The traditional evolutionary view of the genome (top): DNA gene sequences code for proteins while the intervening DNA (dashes) is nothing more than useless junk. The current view (bottom) recognizes numerous large segments of noncoding DNA as probably functional in terms of structure, regulation, and the transcription of noncoding RNA (ncRNA). Finally, the remaining noncoding DNA (dashes) can be understood as at least minimally functional for a number of reasons discussed in the text.

**Widespread transcription of ‘useless burden’ DNA**

A second major discovery involves the discovery of the existence of large amounts of RNA transcripts that, contrary to the stereotyped role of mRNA as precursor to protein synthesis, are not directly related to the function of protein-encoding genes:

‘Moreover, we found that (1) there are twice as many sequences expressed on [human] Chromosome 22 than previously thought; (2) many regions with no prior annotation are expressed and highly conserved in the mouse genome, and (3) much of the transcriptional activity exists within introns of annotated genes. Our results suggest that *a large fraction of the genome is expressed as mRNA*, and that there are many coding sequences that have not been annotated... we found a significant fraction of expression is within introns and antisense introns’ (emphasis added)?<sup>11</sup>

There have always been those who have wondered why our genome is cluttered with so much noncoding DNA if it is truly useless. The recent discoveries make it all the more unbelievable that, following evolutionary reasoning, natural selection would not have eventually removed intronic DNA, and to a lesser extent intergenic DNA, were it in fact devoid of function:

‘The fact that intron sequence contents are so pliable is the reason why introns are often considered ‘junk’. On the other hand, the enzymatic degradation of the excised introns must be a significant biochemical burden for the cell, especially

if most of the human genome is transcribed. Why would the cell go to so much trouble? Why not just get rid of the introns?’<sup>12</sup>

It has been known for over a decade that some introns function in terms of the regulation of gene expression,<sup>3-5</sup> but this function may be more widespread than previously realized, as elaborated below.

What’s more, the boundary between genic and intergenic DNA is not as clear-cut as once believed. We now know that some genes may be spread out over millions of bases, such as is the case of the well-known vertebrate dystrophin gene. Apropos to such large genes, it is now realized that introns may be at least a thousand times longer than exons, and that gene-search

programs are not particularly adept in identifying large genes. This owes to the fact that the signal to noise ratio of very large genes is quite low. Consequently, the researcher is in the proverbial position of being too close to the leaves to adequately see the forest. Or perhaps one can analogize the difficulty that pre-modern navigators experienced in recognizing and charting a large archipelago (giant gene) that was composed of tiny islands (exons) separated by large stretches of ocean (introns). Large stretches of our genome that appear to be free of genes may actually be housing very large genes, and it would only take the discovery of a relatively small number of such large-intron genes to convert most human so-called intergenic DNA to intronic DNA.<sup>13</sup>

**Some features of noncoding RNA (ncRNA)**

A variety of previously unsuspected types of RNA have been discovered in recent years. Most of these appear to have regulatory functions, and a large fraction of them are embedded in intron sequences:

‘Although many introns degrade, some contain active elements, such as microRNAs that can exploit the “RNA interference” effect to control other genes.’<sup>14</sup>

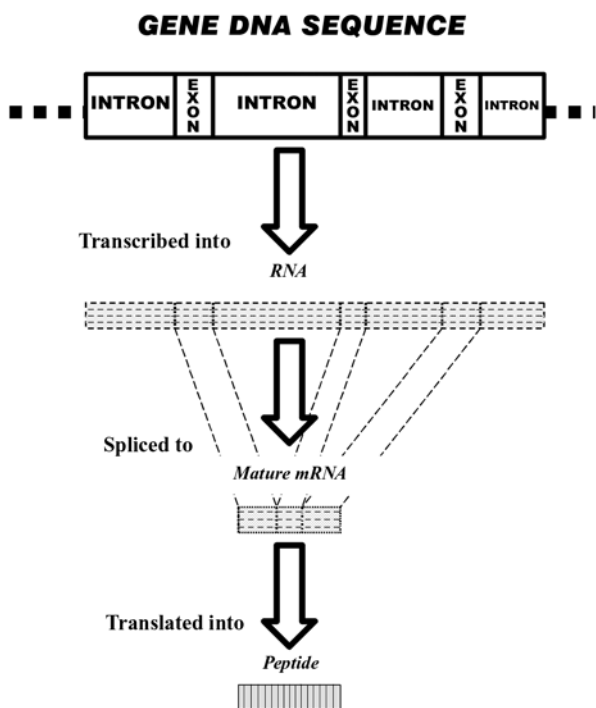
In other words, the ncRNA may inactivate the mRNA transcribed by a gene, thereby regulating the expression of the gene. One means by which this can be achieved is through the formation of dsRNA (double-stranded RNA) complexes between mRNA and its complementary antisense RNA. This process effectively binds the mRNA molecule, preventing it from being translated.

It turns out that ncRNA can be transcribed from an unexpected part of the DNA molecule. The DNA molecule

normally occurs as a union of two strands, the sense strand and the antisense strand. Normally, only one of the two strands (called the sense strand) of the DNA molecule is transcribed into RNA, and the other strand (called the antisense strand) is not used for this purpose. However, recent discoveries have demonstrated that antisense transcription, long thought to be very rare, is actually quite common in the human genome. These discoveries are elaborated in the section on pseudogenes below.

The sequences of ncRNAs tend to be short (as little as 21 nucleotides). This at least partly explains why they, and the noncoding DNA sequences responsible for their transcription, have eluded detection for so long. Moreover, they are probably common in the genome. For every protein encoding gene in the human genome, there may be at least one usually-short DNA sequence that codes for ncRNA.<sup>15</sup> Consider also the following sobering reality:

‘Recent reports indicate that these two stRNAs [small temporally regulated RNAs—a rough synonym of ncRNA, or at least a subset of nrRNAs] are indeed likely to represent only the top of an iceberg with hundreds or more of additional micro-RNAs (miRNAs) existing in metazoans. MiRNAs might thus be previously underestimated key participants in the field of gene regulation.’<sup>16</sup>



**Figure 2.** An enlargement of the gene shown in fig. 1. Introns have usually been considered useless portions of protein encoding genes. Although the entire gene sequence (top) becomes transcribed (second from top), only the exon portions of the transcripts remain in the final transcript (third from top), and become translated into a protein (bottom). Introns are now known to serve a number of functions.

Some evolutionists never tire of telling us that the DNA sequence of humans is very similar to that of chimps. To begin with, this degree of similarity has been exaggerated.<sup>17</sup> Moreover, the similarities in chimp-human DNA may be of very limited relevance. It turns out that the subtle effects of ncRNA influence on genes, rather than differences between the genes themselves, may actually be the primary cause of the biological differences between humans and chimps. It is thus ironic that the real key to human distinctiveness from other forms of life, from a genomic viewpoint, may thus actually lie in noncoding RNA rather than in protein-encoding DNA:

‘On the contrary, the massive amount of ncRNA that is expressed from the genomes of higher organisms, and the complex genetic phenomena that involve RNA, suggests that ncRNAs may constitute an endogenous control system that regulates the programmed patterns of gene expression during their development.’<sup>18</sup>

In any case, the discoveries surrounding ncRNA have major implications for our understanding of intergenic DNA:

‘Geneticists have long focused on just the small part of DNA that contains blueprints for proteins. The remainder—in humans, 98 percent of the DNA—was often dismissed as junk. But the discovery of many hidden genes that work through RNA, rather than protein, has overturned that assumption.’<sup>19</sup>

### Pseudogene function in the light of recent discoveries

Just as intergenic DNA was thought to be junk because it could not function as a template for protein synthesis, so also were apparently disabled copies of protein encoding genes, called pseudogenes. The long-held but outdated concept of the pseudogene rested upon the faulty premise that any inability to code for a protein necessarily implied the absence of any function. Certain pseudogene copies of protein-encoding genes had long been known to produce RNA transcripts, but such activity was thought to be little more than the last dying gasps of the pseudogene. It was supposed, after all, that mutations capable of preventing translation are more likely to occur than those that prevent transcription. Such opinions have to be completely overhauled as a result of recent findings:

‘As an example of human DNA have been found in it almost equal numbers of genes and pseudogenes—defective copies of functional genes. For decades, pseudogenes have been written off as molecular fossils, the remains of genes that were broken by mutation and abandoned by evolution. But this past May a group of Japanese geneticists reported their discovery of the first functional pseudogene.’<sup>20</sup>

This is in reference to the recently described Makorin1-p1 murine pseudogene, discussed elsewhere.<sup>21</sup> It performs an RNA-only function. Moreover, its function is completely different from that of its peptide-encoding paralogous (counterpart) gene. This is no fluke. There are two snail pseudogenes that are, respectively, functional in spite of being unable to code for a full-length protein, or for any peptide at all.<sup>22</sup> In addition, it has been shown that pseudogenes, despite being incapable of encoding peptides of appreciable length, can nevertheless encode very short peptide segments (8–11 amino acids length, with only a modest degree of sequence conservation over just this 8–11 amino acid span) that can at least potentially serve an immunobiological function.<sup>23</sup>

It is not correct to say that the Makorin1-p1 murine pseudogene described above is the first known functional pseudogene. And one must go well beyond the earlier-discovered functional snail pseudogenes. As demonstrated elsewhere,<sup>24</sup> and semantics aside, there is actually a whole set of indisputably-functional genes that qualify as functional pseudogenes in that they have major pseudogenic features, such as premature stop codons, that are circumvented by recoding processes. Partially as a consequence of this, and even if one wishes to adhere to the standard protein-encoding mentality of conventional genes, one can no longer straightforwardly assume that a given gene copy is necessarily inactivated. More recent work only underscores this fact all the more:

‘Differentiation between functional genes and disabled pseudogenes in genome annotation has proven to be a challenging and difficult task.’<sup>25</sup>

Recent discoveries relative to intergenic junk DNA itself have major implications in our understanding of pseudogenes. Consider the previously unknown world of ncRNA. As noted above, the Makorin1-p1 murine pseudogene<sup>21</sup> functions exclusively through the production of ncRNA. So does one of the two functional snail pseudogenes.<sup>22</sup> The previously documented ubiquity of overall ncRNA transcription in the genome at least suggests that ncRNA-encoding pseudogenes may be common.

Attention is now directed to the antisense transcription of intergenic and intronic DNA. Yelin et al.<sup>26</sup> have shown that antisense transcription, until recently thought of as an exotic rarity, and mostly limited to prokaryotes, occurs far more commonly in the human genome than previously supposed. This has also led to a very conservative estimate of >8% of human genes having an antisense partner.<sup>27</sup> Let us now recall that one of the two previously mentioned snail pseudogenes (antiNOS-1)<sup>22</sup> produces functional antisense RNA, albeit in a manner that differs from most of the antisense transcription in the genome. Nevertheless, the unexpected frequency of antisense transcription in our genome encourages further consideration of functional antisense-RNA transcription by pseudogenes as a common phenomenon.

Of course, having been burdened for so long with the

disabled gene assumption, studies of pseudogene function are essentially at a beginning. One factor, other than evolutionistic preconceptions, that had led to the premature rejection of pseudogene function was a form of reasoning that Gibbs calls reverse genetics, as described in his ensuing paragraph. This is a reductionist approach to the genome, and is contrasted with its remediation more broad-based analytical approach called forward genetics:

‘Reverse genetics begins with a particular gene of interest. The scientist fiddles with that gene ... , watches what happens, and then tries to deduce the gene’s function ... But the gradual realization that the genome includes hidden genes—functional sequences that were misclassified as junk—highlights a major problem of reverse genetics: *it can lead to tunnel vision* ... But forward genetics has already unearthed genetic phenomena, such as a functional pseudogene, (see main text), that no one knew were possible’ [emphasis added].<sup>28</sup>

This is indeed an intriguing time for keeping abreast of future discoveries related to pseudogene function. With forward genetics, the researcher does not merely alter the sequence of protein encoding genes and observe the consequences of this alteration. S/he also induces point mutations at random points in the genome to determine any phenotypic effects on the organism. In this way, the potential function of the noncoding DNA segment under study may be brought to light.

### Junk DNA: various functions unrelated to sequence conservation

As illuminating as recent rebuttals of the junk DNA have been, they still appear to focus excessively upon conserved DNA *sequence* as a necessary prerequisite for function. This does not square with the facts. Consider introns. It now turns out that they may be functional not only as repositories of regulatory sequences as well as ncRNA-encoding ‘islands’ of DNA, but also solely in terms of their physical *length*:

‘We argue that minimal introns [that is, introns that cluster around a species-specific peak at the lower end of the intron-size distribution] affect function by enhancing the rate at which mRNA is exported from the cell nucleus ... From an analysis of the yeast expression data, we will show that minimal introns can enhance mRNA synthesis rates. In essence, we present an example of selection based on conservation of intron size, as opposed to conservation of sequence content ... Perhaps the perception that introns are junk is an artifact of an overly narrow focus on conservation of sequence content as the only signature of selection.’<sup>12</sup>

Indeed! This iconoclastic reasoning should be extended further. It turns out that essentially no sequence conservation at all is compatible with function in noncoding

DNA. As elaborated elsewhere,<sup>29</sup> some forms of ncRNA may perform an antiviral immunobiological function that is not dependent upon the conservation of even short blocs of intergenic DNA for their transcription.

In addition, lengthy stretches of noncoding DNA can exert a subtle regulatory influence on distant genes based not on sequence but upon general base composition. It has been found that intergenic DNA tends to form segments longer than 300,000 bases that are relatively constant in composition. These segments, called isochores, may be either enriched in G+C (guanine and cytosine) at the expense of A+T (adenine and thymine), or vice-versa. Traditionally, it has been tacitly supposed that genes are ‘just there’, performing their functions as islands in a sea of junk DNA (Fig. 3, top). We now realize that genes do not function in isolation, but are in fact influenced by the composition of the isochores of the noncoding DNA that surround them (Fig. 3, bottom). There are a number of known manifestations of such influence on human gene behavior:

‘Considering that mRNA with long 5’ UTRs, uAUG and initiator codons with suboptimal context are translated less efficiently, our results suggest that genes requiring highly efficient translation should be mostly located in GC-rich isochores, whereas genes requiring fine modulation of expression should be predominantly located in GC-poor isochores. These indications are in agreement with independent data indicating a preferential location of housekeeping and tissue-specific genes in GC-rich and GC-poor isochores, respectively.’<sup>30</sup>

Since at least some genes depend upon their genomic *context* for their usual degree of expression, a considerable amount of intergenic DNA can be understood as serving, at very least, the minimal function of providing this contextual ‘background’.

It has been known for some time that enhancer sequences can be located hundreds of thousands of bases upstream of the genes whose expression they regulate.<sup>31</sup> Now comes evidence that such enhancers can influence genes from much greater distances and, moreover, can occur in those giant segments of noncoding DNA that had long been thought to be free of functional elements:

‘Approximately 25% of the genome consists of gene-poor regions greater than 500 kb, termed gene deserts (1) [sic]. These segments have been minimally explored, and their functional significance remains elusive . . . The demonstration that several of the enhancers characterized in this study reside in gene deserts highlights that these regions can indeed serve as reservoirs for sequence elements containing important functions. Moreover, our observations have implications for studies aiming to decipher the regulatory architecture of the human genome, as well as those exploring the functional impact of sequence variation. The size of the genome regions believed to be functionally linked to

a particular gene may need to be expanded to take into account the possibility of essential regulatory sequences acting over megabase distances.’<sup>32</sup>

Other implications of such long-distant gene regulation remain to be clarified. A few years ago, Zuckerkandl,<sup>33</sup> an eminent lifetime critic of the junk DNA concept, suggested, based on cited evidence, that it is the *distance* between gene and its regulatory element, not the sequence of the noncoding DNA situated between the two entities, that serves as a function. If so, this may be analogous to the function of an antenna.<sup>34</sup>

Let us now consider repetitive sequences of noncoding DNA. Even these may be functional after all. Consider, for example, STRs (short tandem repeats). These [e.g. GATAGATAGATA...abbreviated (GATA)<sub>n</sub>, and TATC-TATCTATC...abbreviated (TATC)<sub>n</sub>] have long been thought to be biological nonsense. However, Riley and Krieger<sup>35</sup> have presented evidence that such repeats may function as post-transcriptional signals for the mRNAs of genes that encode a variety of membrane-interacting proteins. The nature of this signaling has not been characterized, but it is suspected that (GATA)<sub>n</sub> sequences exist in order to introduce multiple stop codons in series. The latter have already been shown to prolong the stability of untranslated mRNA transcripts.

### Junk DNA: how puny our understanding

For the longest time, noncoding DNA has been deemed unworthy of serious study. The prevalence and intensity of this attitude is obvious from the following recent statements of James D. Watson, writing on the 50<sup>th</sup> anniversary of his (and Francis Crick’s) discovery of the helical structure of the DNA molecule:

‘As the HGP [Human Genome Project] lurched into high gear, the debate persisted about the best way to proceed. Some pointed out that a large portion of the human genome is what we in the trade call ‘junk’, stretches of DNA that apparently don’t code for anything. Indeed, those stretches that encode proteins—genes—constitute only a small fraction of the total. Why therefore, these critics asked, should we sequence the entire genome—why bother with the junk?’<sup>36</sup>

All of the attention paid in this article to conserved blocs of DNA, and related phenomena, should not lead to the impression that researchers have gotten a reasonable handle on even this aspect of junk DNA. At present, only a single-digit aggregate percentage of human noncoding DNA is suspected of being conserved, but this is surely an underestimate of unknown magnitude, (leaving aside the fact that noncoding DNA does not have to be conserved to be functional). It is recognized that alternating multi-species comparisons of noncoding DNA, the standard method of identifying short conserved blocs of the same, probably do not saturate our ability of detecting such sequences,<sup>37</sup> at

least among mammals. Moreover, such comparisons, by their very nature, are apt to miss those conserved blocs that vary from species to species.<sup>38</sup> Noncoding RNAs (ncRNAs) can also escape detection for similar reasons, in addition to an unknown fraction of them being expressed only under very specific conditions within an organism.<sup>39</sup>

Many avenues of research into the nature of junk DNA have scarcely been even touched. Almost nothing is known about possible functions of noncoding DNA related to the fractal nature of DNA itself.<sup>40</sup> (Note that the previously discussed intron function related to intron length<sup>12</sup> is an example of DNA function caused by the fractal distribution of intron lengths). Finally, we now realize that inherited characteristics are governed not only by the DNA of the cell but also by an ‘epigenetic’ layer of information that is stored in the proteins and chemicals that surround and adhere to the DNA molecule itself.<sup>41</sup> We understand very little about this ‘epigenetic’ layer, and even less about how it may impact our understanding of noncoding DNA:

‘Geneticists have yet to decipher the complex code by which epigenetic marks interact with the other components of the genome.’<sup>42</sup>

### Conclusion

There is little doubt that, at least in terms of genomics, recent discoveries relative to noncoding DNA have been revolutionary in nature:

‘There are probably tens or even hundreds of thousands of small RNAs produced by processing of expressed noncoding RNA sequences, including

introns ... It is now clear that the assumption that most genetic information is expressed as proteins is incorrect, at least in the higher organisms ... The central dogma has therefore not only been taken to mean that most genes encode proteins, but also that proteins are sufficient in themselves to specify and organize the autopoietic programming of complex biological entities, an assumption that has pervaded molecular biology for decades. This assumption must now be reassessed.’<sup>43</sup>

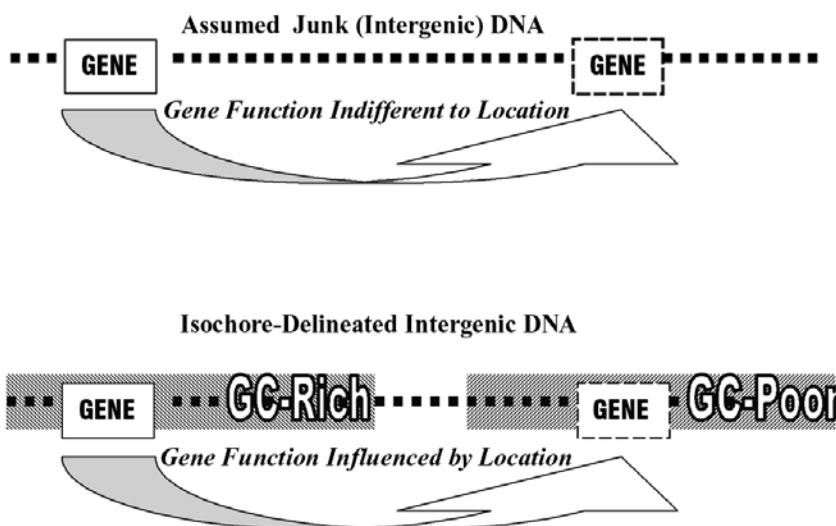
Let us consider some implications of this ‘scientific earthquake’. Evolutionist propagandists often claim that, not only is molecules-to-man evolution an indisputable fact, but that the biological sciences (or even *all* sciences) cannot even exist without it. The contrary is easily shown by recalling the history of biological concepts, many of which (e.g. taxonomy, cell theory, germ theory) were developed either before Darwin or independent of his ideas. In fact, this evolutionary propaganda can easily be turned around. One can show the ways that evolutionary concepts have actually *hindered* the progress of science in general and the biological sciences in particular. Although Gibbs does not elaborate on the evolutionistic origins of the junk DNA concept, it is difficult to escape the conclusion, from his following quotations of some apparently verbal comments of Mattick, that the evolutionary junk DNA concept has long exerted a baleful effect on the biological sciences:

‘I think that this will come to be a classic story of orthodoxy derailing objective analysis of the facts, in this case for a quarter of a century, Mattick says, “The failure to recognize the full implication of this—particularly the possibility that the intervening noncoding sequences may be transmitting parallel information in the form of RNA molecules—may well go down as one of the biggest mistakes in the history of molecular biology”.’<sup>44</sup>

Strong language indeed! It takes little reflection to realize in what direction the evolutionary concept of junk DNA is headed:

‘No one knows just what the big picture of genetics will look like once this hidden layer of information is made visible. “Indeed, what was damned as junk because it was not understood may, in fact, turn out to be the very basis of human complexity”, Mattick suggests.’<sup>45</sup>

One can, at very least, only wonder how much sooner the functions of noncoding DNA would have been discovered had the genome been recognized as the pre-planned product of an Intelligent Designer instead of the long-term outcome



**Figure 3.** Isochore composition as a gene regulatory function of noncoding DNA. Conventionally, genes were thought to function as self-contained units, independent of the surrounding DNA that, after all, was presumably redundant and useless (top). A change in their location (curved arrow) would not have altered their function. With isochore composition functioning as a contextual ‘background’ for gene expression (bottom), a change in surrounding noncoding DNA is now known to alter the expression of many genes.

of purposeless evolutionary processes.

### References

1. Wong, G.K.-S., Is 'junk' DNA mostly intron DNA? *Genome Research* **10**:1672, 2000.
2. Function can be defined as the performance of an action that benefits the host organism.
3. Walkup, L.K., Junk DNA: evolutionary discards or God's tools? *TJ* **14**(2):18–30, 2000.
4. Standish, T.G., Rushing to judgment: functionality in noncoding or 'junk' DNA, *Origins (Loma Linda)* **53**:7–30, 2002.
5. Batten, D., More junk reclaimed, *TJ* **16**(2):8, 2002.
6. Gibbs, W.W., The unseen genome: gems among the junk, *Scientific American* **289**(5):48–53, 2003. The present article goes beyond an exposition of Gibbs and his citations. It also contains a fair amount of more recent material.
7. In the present context, 'noncoding DNA' relates to actual or perceived noninvolvement in the long-known process of protein synthesis. As discussed further in this article, 'noncoding DNA' can code for various short but functional biomolecules.
8. It must be reiterated that this would, at most, indicate that such DNA does not perform an *indispensable* function. It would not prove that junk DNA performs *no* function. Consider an analogy with the human organism. Parts of the human body can be surgically removed without apparent harm to the patient, but this hardly proves that these body parts lack function.
9. Dermitzakis, E.T. *et al.*, Numerous potentially functional but non-genic conserved sequences on human chromosome 21, *Nature* **420**(6915):578–582, 2002.
10. Dermitzakis, E.T. *et al.*, Evolutionary discrimination of mammalian conserved non-genic sequences (CNGs), *Science* **302**:1033–1035, 2003.
11. Rinn, J.L. *et al.*, The transcriptional activity of human Chromosome 22, *Genes and Development* **17**(4):530, 536, 2003.
12. Yu, J. *et al.*, Minimal introns are not 'junk', *Genome Research* **12**:1185, 2002.
13. Wong, ref. 1, pp. 1672–1678.
14. Gibbs, ref. 6, p. 51.
15. Mattick, J.S., Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms, *BioEssays* **25**:931, 2003.
16. Grosshans, H. and Slack, F.J., Micro-RNAs: small is plentiful, *J. Cell Biology* **156**(1):17, 2002.
17. DeWitt, D. A., >98% chimp/human DNA similarity? Not any more, *TJ* **17**(1)8–10, 2003.
18. Mattick, ref. 15, p. 936.
19. Gibbs, ref. 6, p. 48.
20. Gibbs, ref. 6, p. 50.
21. Woodmorappe, J., Pseudogene function: more evidence, *TJ* **17**(2):15–18, 2003.
22. Woodmorappe, J., Pseudogene function: regulation of gene expression, *TJ* **17**(1):47–52, 2003.
23. Woodmorappe, J., The potential immunological functions of pseudogenes and other junk DNA, *TJ* **17**(3):102–108, 2003.
24. Woodmorappe, J., Unconventional gene behavior and its relationship to pseudogenes, *Proceedings of the Fifth International Conference on Creationism, Technical Volume*, pp. 505–514, 2003.
25. Zhang, Z. and Gerstein, M., The human genome has 49 cytochrome *c* pseudogenes, including a relic of a primordial gene that still functions in mouse, *Gene* **312**:71, 2003.
26. Yelin, R. *et al.*, Widespread occurrence of antisense transcription in the human genome. *Nature Biotechnology* **21**:379–386, 2003.
27. Yelin *et al.*, ref. 26, p. 384.
28. Gibbs, ref. 6, p. 52.
29. Woodmorappe, ref. 23, pp. 106–107.
30. Pesole, G., Bernardi, G. and Saccone, C., Isochore specificity of AUG initiator context of human genes, *Federation of European Biochemical Societies Letters* **464**:61, 1999.
31. Woodmorappe, ref. 24, p. 507.
32. Nobrega, M.A., Ovcharenko, I., Afzal, V. and Rubin, E.M., Scanning human gene deserts for long-range enhancers, *Science* **302**:413, 2003.
33. Zuckerkandl, E., Why so many noncoding nucleotides? *Genetica* **115**:119–121, 2002.
34. Consider the analogy of the antenna and the radio component. Much as the antenna relies upon its cumulative distance from the radio (independent of any internal sequence, unlike the radio components themselves) for function, so also noncoding DNA can function as a spacer, or tether, with little or no regard for its sequence.
35. Riley, D.E. and Krieger, J.N., Diverse eukaryotic transcripts suggest short tandem repeats have cellular functions, *Biochemical and Biophysical Research Communications* **298**:581–586, 2002.
36. Watson, J.D., *DNA: The Secret of Life*, Alfred A. Knopf, New York, p. 179, 2003.
37. Thomas, J.W. *et al.*, Comparative analyses of multi-species sequences from targeted genomic regions, *Nature* **424**:791, 2003. In addition, 98% of non-exonic multi-species conserved sequences do not correspond to currently known regulatory elements (p. 791). This underscores the fact that our understanding of even the rudiments of the regulatory functions of noncoding DNA are only in their infancy.
38. Margulies, E.H. *et al.*, Identification and characterization of multi-species conserved sequences, *Genome Research* **13**:2513, 2003.
39. Lai, E.C., microRNAs: runts of the genome assert themselves, *Current Biology* **13**:R930, 2003.
40. Spinelli, G., Heterochromatin and complexity, *Nonlinear Dynamics, Psychology, and Life Sciences* **7**(4):338, 2003.
41. Gibbs, W.W., The unseen genome: beyond DNA, *Scientific American* **289**(6):106–113, 2003.
42. Gibbs, ref. 41, p. 108.
43. Mattick, ref. 15, pp. 933, 937.
44. Gibbs, ref. 6, pp. 49–50.
45. Gibbs, ref. 6, p. 53.

---

**John Woodmorappe** has an M.A. in geology and a B.A. in biology, from a midwestern US state university. He is a science educator by profession. He is the author of *Studies in Flood Geology*, *Noah's Ark: A Feasibility Study* and *The Mythology of Modern Dating Methods*.