Biosemiotics

Response to Denis Noble's Target Article on "The Illusions of the Modern Synthesis," Biosemiotics .

--Manuscript Draft--

Manuscript Number:	
Full Title:	Response to Denis Noble's Target Article on "The Illusions of the Modern Synthesis," Biosemiotics .
Article Type:	Commentary (1,000 - 2,000 words)
Keywords:	One gene-one protein concept; Central Dogma of Molecular Biology; Macroevolutioo and Microevolution; repetitive non-coding DNA; genomic regulatory elements; mobile DNA elements; genomic control networks; functional regulation by non-coding ncRNA molecules; complex evolutionary change through natural genetic engineering
Corresponding Author:	James A. Shapiro, PhD University of Chicago Chicago, IL UNITED STATES
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Chicago
Corresponding Author's Secondary Institution:	
First Author:	James A. Shapiro, PhD
First Author Secondary Information:	
Order of Authors:	James A. Shapiro, PhD
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	The authors of the Modern Synthesis (MS) constructed their theory of evolutionary change based on Darwin's Origin of Species and early 20th Century Mendelian and population genetics. Although early results in microbial and molecular genetics were interpreted to solidify MS ideas (via The Central Dogma of Molecular Biology), accepting the permanent truth of their MS concepts blinded adherents to significant developments in the second half of the 20th Century and 21st Century' Genomic studies of evolution based on the DNA sequence record have produced a radically different view of genome complexity and biologically-mediated evolutionary change.

Response to Denis Noble's Target Article on "The Illusions of the Modern Synthesis," *Biosemiotics* .

James A. Shapiro, University of Chicago, USA

In his Target Article "The Illusions of the Modern Synthesis," Denis Noble rightly credits the evolutionary biologists and population geneticists who developed this theory as serious scientists attempting to explicate one of the most complex questions in all of science: How do new life forms originate in the course of organic evolution? In a journal named *Biosemiotics*, it is particularly fitting that we analyse further why and how this group of highly gifted scientists came up with an explanatory scheme (the MS) that is still taught but has been rendered largely irrelevant by direct reading of the genome record in the decades since DNA was identified as the chemical core of genetic inheritance [1, 2].

I would like to suggest that the fundamental error of the MS originators was to forget that scientific explanations are inevitably temporary and provisional. In the words of their contemporary, 1983 Nobel Laureate Barbara McClintock, the authors of the MS believed they had solved the basic problems of heredity and could not see beyond what she called the "Now Explanation." By that phrase, McClintock meant our contemporary conceptual framework and the language we use to articulate it. Believing these ideas and terminology to be unchangeable facts can prevent us seeing possibilities outside the current paradigm. Now Explanations frequently make scientists unable to grasp the significance of unexpected phenomena, even when the empirical evidence is irrefutable.

The word "Synthesis" in MS stands for the merger of Darwinian ideas about evolution by random changes and gradual Natural Selection ("the Preservation of Favoured Races in the Struggle for Life") [3] with the mechanisms of Mendelian and population genetics worked out in the first half of the 20th Century. Although originally introduced by Johannsen as a neutral term for any hereditary factor subject to Mendel's rules [4], the "gene" began to take on a conceptual life of its own as the "basic unit" of inheritance. Genomes were considered to be merely collections of genes arranged on chromosomes like "beads on a string." When Beadle and Tatum published their "one gene-one enzyme" conclusion from fungal genetics experiments in 1941 [5], it was quickly generalized to "one gene-one protein," and geneticists concluded that the function of genes was to direct the synthesis of proteins involved in determining cellular and organismal characters.

The discovery of the structure of DNA provided confirmation for this viewpoint in the form of The Genetic Code, whereby a given sequence of nucleotide bases in DNA encoded the sequence of amino acids in a particular protein polypeptide chain [6]. Crick's 1958 paper was the initial statement of what he called "The Central Dogma of Molecular Biology" -- DNA directs the synthesis of more DNA as well as messenger RNAs that encode the proteins, the molecules that do the basic work to determine the properties of cells and organisms. Although (as Noble points out) Crick later had to revise his scheme to accommodate the discovery that RNA can encode DNA by the action of reverse transcriptase [7, 8], the problem of genome functioning seemed to have been solved with a very tidy molecular division of labor. Crick's use of the term "dogma" was meant quite literally to enshrine this genocentric viewpoint as a fundamental and permanent truth.

Evolutionary novelty arose randomly through unavoidable copying errors in DNA replication, leading to gradual cumulative changes in protein structure and function.

Several problems with the reductionist Central Dogma view of genomes as vehicles to reproduce collections of protein-coding genes as basic units of heredity were apparent early on. For the MS, the most important challenge was the distinction Richard Goldschmidt pointed out between Macroevolution, based on chromosome restructuring that originated species, and Microevolution, based on gene mutations leading to gradual improvement of adaptations within existing species. Goldschmidt documented this distinction extensively in his 1940 book, *The Material Basis of Evolution* [9]. The MS community dealt with Goldschmidt by a form of scientific excommunication, either totally ignoring him or relegating his views on evolution to what they characterized as the obviously preposterous hypothesis of dramatically novel creatures Goldschmidt called "hopeful monsters." Today, of course, the hopeful monster idea is seen as the beginning of studies on the evolution of multicellular development (Evo-Devo).

The empirical breakdown in the MS came from a variety of sources, four of which are particularly significant:

1. The discovery of mobile genetic "controlling elements" in maize plants by McClintock in the late 1940s and early 1950s [10]. McClintock unexpectedly discovered these genome elements that move from one chromosomal location to another and reprogram the expression of adjacent genetic loci, but she quickly realized the tremendous potential they had for executing major genomic and phenotypic changes. Her communication of this discovery, however, was greeted with incomprehension and anger by her colleagues. It was the rediscovery in the late 1960s and early 1970s of similar forms of mobile DNA in bacteria, yeast, *Drosophila* and ultimately all organisms that brought this major form of rapid hereditary variation into the scientific mainstream [11].

2. The study of how genome expression is regulated, ranging from bacterial metabolism of sugars to cellular differentiation in complex eukaryotes [12, 13]. These studies identified two features of genomes that were not considered by the MS: (i) the presence of DNA sequence elements recognized by regulatory proteins to control expression of every genetic locus, and (ii) the sharing of regulatory sequence elements between different loci to establish control networks across the genome. Such networks determine the expression of protein-coding genes under changing conditions and are the reason so many phenotypes have multifactorial genetic determination. The origination and coordinated evolution of these networks presents a particular explanatory challenge for MS gradualism.

3. The discovery of repetitive and largely non-coding DNA in the genomes of humans and other complex eukaryotes [14]. Repetitive DNA is clearly different from the largely unique protein coding sequences believed to be the basic functional units of heredity. To explain large amounts of repetitive DNA sequences in genomes, MS theoreticians assumed them to be non-functional freeloaders, labelling them either "junk DNA" [15] or "selfish DNA" [16]. Richard Dawkins erected an entire MS-

centered philosophy based on the existence of *The Selfish Gene* [17]. However, repetitive and non-coding DNA elements are essential to the evolution of complex organisms. At the dawn of the 21st Century, genomics revealed our own genome contains over 30 times more repetitive DNA than protein-coding sequences [18]. A study comparing protein-coding and non-coding DNA with respect to organismal complexity found that coding DNA levelled off at approximately 3 x 10^7 bp, while the non-coding DNA increased linearly with greater complexity up to 2-3 x 10^{10} bp [19]. In other words, non-coding DNA tracked organismal complexity better than the protein-coding genes. Moreover, repetitive DNA is the most labile genome component in evolution. We now understand that repetitive DNA elements act as mobile elements like McClintock's or provide shared regulatory sites connecting genomic networks [20] -- and quite often do both [21]. That last point is important because it tells us that mobile DNA elements truly serve as controlling elements to evolve the distributed but coordinated genome networks needed for the phenotypes of complex organisms.

4. Finally, molecular genomic analysis by the encyclopedia of DNA elements (ENCODE) consortium has documented that the vast majority of the supposed non-coding DNA is transcribed in a regulated manner to produce non-coding ncRNAs with a wide range of functional properties [22]. Some ncRNAs control transcription of DNA into RNA or translation of messenger RNA into protein. Many short ncRNAs play important roles in epigenetic formatting to control expression over broad regions of the genome, while ongoing research is finding that long non-coding lncRNAs execute a growing number of functions in processes as diverse as forming transcription complexes, cellular differentiation, stem cell pluripotency, sex determination, flower timing, fruit ripening, and neural connectivity in mammalian brains [21, 23].

None of the preceding four major developments in our current understanding of what a genome is and how it can evolve could possibly have been anticipated by the authors of the MS. Yet all four are solidly established by abundant molecular and genomic evidence. They tell us clearly;

- genomes contain a lot more than protein-coding genes;
- genomes change because of internal biochemical processes, such as the transposition of mobile DNA elements;
- Cellular "natural genetic engineering" processes provide the molecular basis for evolving distributed networks encoding complex phenotypes [21].

Contemporary genomics has turned evolutionary theory upside-down. The MS treats the evolving organism as the passive object of random mutations and Natural Selection. In the genomics view, the organism is an active agent of its own evolution, responding to ecological challenge by restructuring its genome. All the more's the pity that students do not learn this 21st Century evolution.

REFERENCES

1. Avery, O.T., C.M. MacLeod, M. McCarty, Studies on the chemical nature of the substance inducing transformation of Pneumococcal types: Induction of transformation by a desoxyribonucleic acid fraction isolated prom Pneumococcus Type III*. J. Exp. Med.*, 1944. **79**: p. 137–158. https://pubmed.ncbi.nlm.nih.gov/19871359/
2. Watson, J.D. and F.H. Crick, Genetical implications of the structure of deoxyribonucleic acid*. Nature*, 1953. **171**: p. 964-967.
3. Darwin, C., *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*1859, London: John Russel.
4. Johannsen, W., The Genotype Conception of Heredity*. The American Naturalist*, 1911. **45**(531): p. 129-159.
5. Beadle, G.W. and E.L. Tatum, Genetic Control of Biochemical Reactions in Neurospora*. Proc Natl Acad Sci U S A*, 1941. **27**(11): p. 499-506. https://pubmed.ncbi.nlm.nih.gov/16588492/
6. Crick, F., On protein synthesis*. Symp Soc Exp Biol*, 1958. **12**: p. 138-163.
7. Temin, H., S Mizutani, RNA-dependent DNA polymerase in virions of Rous sarcoma virus*. Nature*, 1970. **226**: p. 1211-1213. https://pubmed.ncbi.nlm.nih.gov/4316301/
8. Crick, F., Central dogma of molecular biology*. Nature*, 1970. **227**: p. 561-563. https://pubmed.ncbi.nlm.nih.gov/4913914/
9. Goldschmidt, R., *The Material Basis of Evolution, Reissued (The Silliman Memorial Lectures Series), 1982*1940, New Haven CT: Yale Univ.Press.
10. McClintock, B., *Discovery And Characterization of Transposable Elements: The Collected Papers of Barbara McClintock* 1987, New York: Garland.
11. Bukhari, A.I., J.A. Shapiro, and S. L. Adhya (Eds.), *DNA insertion elements, plasmids and episomes* 1977, Cold Spring Harbor, New York: Cold Spring Harbor Press.
12. Jacob, F., Monod, J, Genetic regulatory mechanisms in the synthesis of proteins*. J Mol Biol*, 1961. **3**: p. 318– 356. https://pubmed.ncbi.nlm.nih.gov/13718526/
13. Goode, D.K., et al., Dynamic Gene Regulatory Networks Drive Hematopoietic Specification and Differentiation*. Dev Cell*, 2016. **36**(5): p. 572-87. https://pubmed.ncbi.nlm.nih.gov/26923725/
14. Britten, R., Kohne, DE, Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms*. Science*, 1968. **161**: p. 529-540. https://pubmed.ncbi.nlm.nih.gov/4874239/
15. Ohno, S., So much "junk" DNA in our genome*. Brookhaven Symp Biol*, 1972. **23**: p. 366-70. https://pubmed.ncbi.nlm.nih.gov/5065367/
16. Orgel, L.E. and F.H. Crick, Selfish DNA: the ultimate parasite*. Nature*, 1980. **284**(5757): p. 604-7. https://pubmed.ncbi.nlm.nih.gov/7366731/
17. Dawkins, R., *The Selfish Gene*1976, Oxford: Oxford University Press.
18. Lander, E.S., et al., Initial sequencing and analysis of the human genome*. Nature*, 2001. **409**(6822): p. 860-921. https://pubmed.ncbi.nlm.nih.gov/11237011/
19. Liu, G., J.S. Mattick, and R.J. Taft, A meta-analysis of the genomic and transcriptomic composition of complex life*. Cell Cycle*, 2013. **12**(13): p. 2061-72. https://pubmed.ncbi.nlm.nih.gov/23759593/
20. Britten, R.J. and E.H. Davidson, Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty*. Q Rev Biol*, 1971. **46**(2): p. 111-38. https://pubmed.ncbi.nlm.nih.gov/5160087/

21. Shapiro, J.A., Living Organisms Author Their Read-Write Genomes in Evolution. *Biology (Basel)*, 2017. **6**(4). https://pubmed.ncbi.nlm.nih.gov/29211049/
22. Consortium, E.P., An integrated encyclopedia of DNA elements in the human genome. *Nature*, 2012. **489**(7414): p. 57-74. https://pubmed.ncbi.nlm.nih.gov/22955616/
23. Statello, L., et al., Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol*, 2021. **22**(2): p. 96-118. https://pubmed.ncbi.nlm.nih.gov/33353982/