

Agents Competent in Natural Genetic Engineering

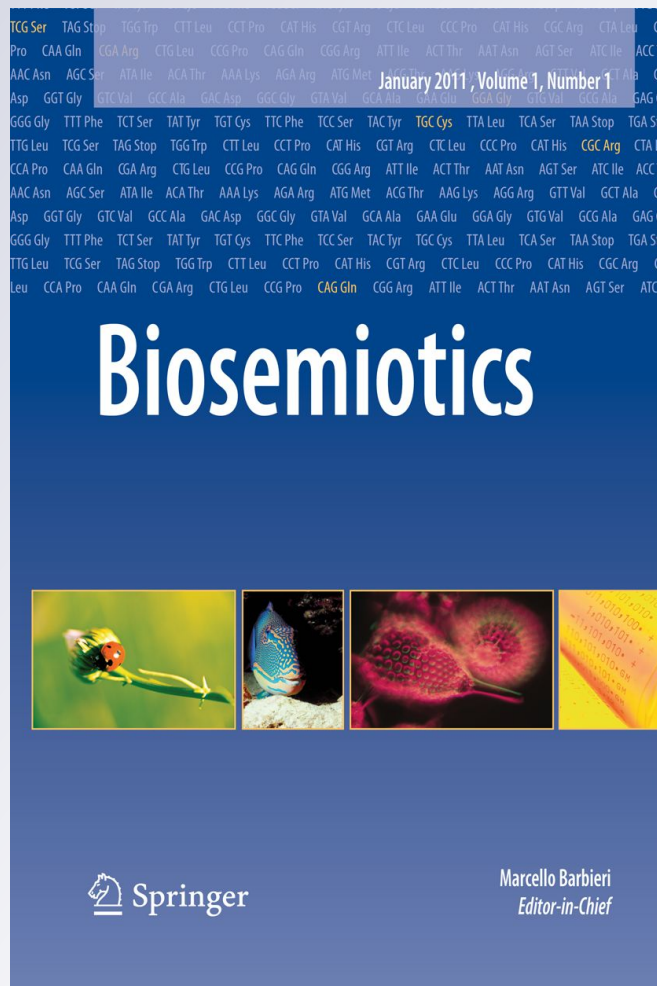
Guenther Witzany

Biosemiotics

ISSN 1875-1342

Biosemiotics

DOI 10.1007/s12304-012-9141-9



Your article is protected by copyright and all rights are held exclusively by # Springer Science+Business Media B.V.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.

Agents Competent in Natural Genetic Engineering

Review of *Evolution: A View from the 21st Century*, by James A. Shapiro (2011): New York: Financial Times Prentice Hall. 272 pages, ISBN: 978-0132780933

Guenther Witzany

Received: 12 December 2011 / Accepted: 9 January 2012
© Springer Science+Business Media B.V. 2012

Introduction

The book contains 4 main chapters (see headlines below) in which this new view of the evolutionary process is presented. Step by step all currently known empirical data are listed. The extraordinary quality of the arguments that Shapiro presents goes in parallel with the absence of any redundant information, so that this masterpiece of a reading book gives dense information in a primary text of 140 pages followed by a very helpful glossary and index, and more than 1100 of the best references in the field. All chapters contain well structured tables that help in categorising the given examples. A special service is offered by a list of far-reaching analyses and appendices that are available online.

The introduction leads to a fresh look at the basic problem of how novelty arises in evolution. Shapiro states that novelty is the critical issue in evolutionary change because without novelty selection would not have anything to act upon. In this respect, mainstream thoughts of the past 60 years were based on the assumption that inheritable novelty is the result of chance or accident or, as Darwin assumed, adaptive change resulting from natural selection applied to a number of small changes over long periods of time. The currently dominant evolutionary biology paradigm is based on the neo-darwinian paradigm that all genetic change occurs accidentally and randomly. At the end of the first half of the last century this perspective received a molecular interpretation as a number of errors in replication processes. The insistence on randomness and accident became dogmatic in order to reject all possible revivals of vitalism. But in the last 50 years there has been an abundance of empirical data that contradict this randomness and accident in variation. Cells have the capacity to change themselves adaptively and to alter their own heredity. In contrast to the predictions of the former mechanistic approaches, it is well documented that recombination has the ability to generate information and to alter the content of the genetic storage. Finally, thanks to Barbara

G. Witzany (✉)
Telos—Philosophische Praxis, Vogelsangstraße 18c, 5111 Buermoos, Austria
e-mail: witzany@sbg.at

McClintock's work, it has become clear that organisms can engineer their DNA. Natural genetic engineering—a term originally introduced by Shapiro himself—represents the “ability of living cells to manipulate and restructure the DNA molecules that make up their genomes”.

Sensing, Signaling and Decision-making in Cells

In part one Shapiro shows that we can look at how cells do not act blindly, but in a well coordinated manner based on incoming information, monitoring and then act accordingly, i.e. guiding the processes that are vital for survival, growth and reproduction. In this respect cells have to (i) adjust their metabolism to available nutrients, (ii) control their progress through the cell cycle, (iii) control that progeny is complete at the time of division, (iv) repair damage and interact in coordination with other cells. This includes programmed cell death as an appropriate tool in developmental processes of multicellular organisms, i.e. suicidal actions for the benefit of the entire population. “Life requires cognition at all levels”. In this respect Shapiro formulates why he prefers the systems biology approach. It is because it seems to serve for appropriate understanding of “how groups of molecules work coordinately (as a system) to achieve some useful function dependent upon conditions. Gone is the atomistic view that molecules act independently and automatically.”

In the first example of how *E. Coli* chooses the best sugar to eat, Shapiro demonstrates how cells organise and generate information in their DNA to a changing environment. First of all there is no Cartesian dualism between information molecules and operation molecules. All classes of molecules participate in sensing, information transfer, information processing and also transport and catalysis. Second, information is then transferred from the cell surface or intracellular sensors to the genome using secondary messenger molecules. Third, protein-DNA recognition occurs at special recognition sites. Fourth, DNA binding proteins and their formatting signals operate in a coordinated manner, and fifth, proteins operate as microprocessors in regulatory circuits which enables them to behave differently depending on the variety of interactional motifs with other proteins or molecules.

As a second example, Shapiro demonstrates that there are important features of proofreading and repair in DNA replication processes. In contrast to former conventional thoughts on evolutionary change by accidents and errors in replication processes, modern research has established that DNA proofreading and repair are central features of cellular management of genome structure. If we look again at *E. Coli* we can identify less than 1 mistake in a billion new nucleotides incorporated in DNA reproduction. *E. Coli* duplicates its 4.6 MB genome in 40 minutes, which means that they reproduce at a rate of 2000 nucleotides per second. Proofreading and repair require well coordinated agents, such as polymerases that sense distortions in a sequence, plus a variety of proteins for mismatch repair and a complicated and coordinated exchange of non-correct sequences. Eukaryotic organisms have many more different DNA polymerases and proofreading exonucleases. Astonishing is the precision of proofreading in all its steps and substeps.

In a third example Shapiro presents the various ways to repair DNA damage (as a consequence of UV radiation or mutagenic chemicals) with the SOS response system that involves expression of repair, recombination and cell division inhibition proteins.

To start with, the UV-damaged DNA is removed. Then DNA polymerases produce homologues by recombination, halt cell division, alter cell metabolism, inhibit normal DNA replication and stimulate exit from the SOS process after repair is complete.

The fourth example is the cell cycle checkpoints: several checkpoints serve as feedback controls and include one for damage in the growth phase, one for replication problems in the DNA replication phase, one for damage in the second growth phase and one for abnormalities in behaviour and alignment of the duplicated chromosomes on the mitotic spindle. The result of this monitoring system is that chromosome separation, telophase and cytokinesis do not occur until all chromosome pairs are properly aligned on the spindle. This process works with a precision of 99.9%.

In a further chapter on signaling from the cell surfaces to the genome, the role of intercellular (cell-cell) signals in cell death (apoptosis), Shapiro shows how the central dogma of molecular biology is violated by demonstrating a variety of examples where proteins modify sequence information (i) in DNA, (ii) in RNA (splicing and the modes of posttranscriptional modifications) and (iii) in other proteins (proteolytic cleavage, peptide excision and attachments). This goes against the central dogma of DNA-RNA-protein-phenotype and the Cartesian dualist view of Francis Crick that nucleic acid sequences contain information and proteins execute encoded instructions.

The Genome as a Read-Write Storage Medium

Part two describes the currently identified cell processes that lead to hereditary genomic changes. This new conceptual landscape changes the traditional view of the DNA as a read-only memory system to a “read-and-write system”. This part documents a variety of examples in which nucleotide and protein interactions target natural genetic engineering processes to different genomic locations, e.g. the adaptive immune system.

Genomes contain the essential formatting elements to build control circles of expression of coding sequences in complex ways. These control sections integrate sequence motifs that serve as recognition sequences for interaction with proteins or molecules and are found all over the genome in many places. Repetitive elements in most cases are not part of coding sequences but serve as regulatory tools. In this way epigenetic modifications lead to more complex regulation patterns in different ways with or without heritable changes in gene sequences. Histone modifications can mark genomes to be active for transcription or to be silenced. In this way genome formatting serves as an appropriate tool for replication, localisation and transmission to daughter cells and provides several distinct sites for initiating DNA replication, for completing DNA replication, for ensuring transmission at cell division or for governing subnuclear localisation.

From my perspective, part 2 of the book is the most striking proof of the thesis of Shapiro, because it presents a great variety of active agents that engineer genome contents. In the chapter “Distinct classes of DNA in the genome” and in the following “The molecular mechanisms of natural genetic engineering”, he describes the core elements that invent, arrange, rearrange and translocate genetic contents in a very clear and distinct way. Exactly these elements are identified by a different biological science, virology, as descendents of viral infection events that persist in host organisms as genome editors, some of them defective, some still active players for host

organisms (e.g., endogenous retroviruses in building syncytin as essential for placentation in mammals). Shapiro makes a long list of key elements of natural genetic engineering in cellular life, such as

- DNA import and export systems (horizontal gene transfer),
- homologous recombination (genetic uptake and transfer),
- non-homologous end-joining (repair DNA breaks and rejoining in adaptive immunity),
- site-specific reciprocal recombination,
- DNA transposons that can move from one (donor) site to another site,
- long terminal repeat (LTRs) retroelements (8% of the human genome),
- non-LTR retroelements (34% of the human genome),
- retrosplicing group II introns, inteins,
- diversity generating retroelements (a bacterial system involving a reverse transcriptase and a repeat element for changing the variable region of a nearby coding sequence) and
- a list of others such as the agents of the adaptive immune system in mammals and a similar system of adaptive immunity in prokaryotes (CRISPRs).

From this description it becomes undoubtedly clear that a network of mobile genetic elements is the real basis of natural genetic engineering. A list of tables enumerates them in detail. This opens an entirely new perspective for the genome theory of the 21st century by showing that cells can actively rewrite their genomes in evolutionary times with numerous natural genetic engineering processes.

Evolutionary Lessons from Molecular Biology

Part three describes the genome DNA record concerning those cell functions that possibly played a relevant role in change over evolutionary times. Many important features we can find in present day genetic sequences are the result of natural genetic engineering of cells, e.g. the 3 million examples of transposon or retrotransposon activity.

In this part of the book interesting aspects are found such as antibiotic resistance and horizontal gene transfer, symbiogenesis and the origin of the eukaryotic cells, and a very important aspect is the role of natural genetic engineering and evolutionary genomic innovation. Additionally, clarifying explanations on the use and reuse of evolutionary inventions are presented as are whole genome doubling events at critical stages of evolutionary innovation and divergence.

One key question of neo-Darwinism was whether the sequences of contemporary genomes fit the Darwinian predictions of change by “numerous successive slight variations”. Shapiro presents the other version, overwhelmingly rich empirical data of major genomic changes at key moments in evolution. This functions according to the integration and engineering of horizontally transferred DNA in which no inviolable taxonomic barriers exist (Table III.1 within the reviewed book). A most interesting phenomenon includes the modular and duplicative nature of protein evolution, i.e. that proteins and related domains are functioning as multiple modules for a variety of genome compositions.

Without a doubt, the symbiogenesis hypothesis has also been proved in a variety of cases which means that eukaryotic cells are a consortial integrated

system of formerly free living prokaryotes, which then transferred key translational competences to the eukaryotic nucleus. The natural genetic engineering competences that have been explained in part two now are shown as key agents of genomic innovation in creating novel exons (coding sequences) as well as novel introns (non-coding sequences), which opens the possibility of multiple protein meanings by alternative splicing. These agents, such as LINES, regulate the speed of genome expression. Others control epigenetic imprinting sites or serve as novel regulatory RNA molecules or in chromosome rearrangements. All of these data demonstrate that proteins evolve by accumulating and rearranging polypeptide domains and not by a series of individual amino acid changes. This means that evolutionary genomic changes are not stochastic, localised point mutations, but exchanges of DNA segments encoding the polypeptides that comprise these domains, typical features of transposable elements. The DNA record does not support the slow accumulation of random gradual changes transmitted by restricted patterns of vertical descent, as postulated for 50 years by neo-darwinian proponents.

A New Conceptual Framework for Evolutionary Research in the 21st Century

Part four looks at how we can integrate the former three parts into a coherent contemporary view of biological evolution. Shapiro presents a new conceptual basis for evolutionary research in the 21st century. This conceptual basis fits better into the empirical data available than neo-Darwinism and its contemporary controversial counterpart, intelligent design. In this respect it represents a third way.

In the first chapter Shapiro demonstrates that a systems theoretical approach can explain how functional novelties evolve. They evolve as features of systems that cooperate. Some novelties are established by reorganising available functions in new combinations. As shown in part 2, the various natural genetic engineering operators are the molecular agents of active genome change such as transposons and retrotransposons and their related agents. They do not reinvent every change in a new way, but reassemble functional systems of protein domains on genome formatting, as well as regulatory elements that change the content order of the genome into new functional domains. In this respect the genome is the result of systemic changes of the cooperative networks of these natural genetic engineering agents and can explain evolutionary novelties in a more coherent way than “numerous successive slight modification” of the genome by random changes (mutations). The abundance of natural genetic engineering agents that are able to change the genetic content order for the adaptational purposes of cells and multicellular organisms presented and described in parts 1–3 does not leave any magical or implausible narratives in how cells can be capable of introducing coordinated changes into different but functionally related regions of their genomes. In a certain way Shapiro resolves the mechanism-vitalism debate of the early 20th century also in demonstrating that the real driving force of evolutionary novelties is neither mechanistic nor vitalistic, but a competent agent-driven engineering process of unbelievable complexity.

Conclusion

This book is helpful to researchers in many different fields such as cell biology, molecular biology, genetics, epigenetics, virology and biosemiotics. It describes an abundance of coordinated actions on genome formatting, restructuring, transcription processes and repair mechanisms, and shows that they are essential features of cellular life. Virus-first proponents can read this book under the assumption that all these features derived from virus infections and were subsequently transferred to cells. Shapiro does not endorse this virological perspective but, being an open minded scientist, does not exclude it either and states that “the virosphere provides a major reservoir of genetic novelty.”

The book represents a state of the art scientific description of a new picture on how organisms evolve(d). It integrates the most relevant empirical data into a unifying perspective on evolutionary processes and answers nearly all the unresolved questions left by neo-Darwinism in the last 50 years. In this respect Shapiro's book represents an excellent example of both Karl Popper's falsification criterion on how a better concept integrates more of the available empirical data than the previous one, and of Thomas Kuhn's paradigm shift, in that a rather new look on evolution arises that was clearly unimaginable for the proponents of the former one.