Introduction

A Perspective on Natural Genetic Engineering and Natural Genome Editing

In 1983, when I finished my studies of philosophies of language and science—in particular pragmatic action theory—I did not know the direction my research interests would take. In the ensuing four years, I studied a number of articles concerning different subjects in biology and was struck by the key vocabulary that was used for the description of the essential activities of cellular life, such as "genetic code," "genetic information," "cell–cell communication," "nucleotide sequences," "protein coding sequences," "self/nonself recognition"—all of which connote themes of communication and exchange of information, similar to the themes I had encountered in my studies of philosophy and action theory. Of particular influence on my thinking were the articles and books of Karl von Frisch, who received the Nobel Prize for his work on the language of bees; and a book by Manfred Eigen, in which he developed a profound argument for the idea that the genetic code functions not only as an analogue of natural human language but that both the evolution of life and the evolution of the mind crucially depend on the characteristic features of languages.

According to these and many other results of the discourse in the philosophy of science in the 20th century (especially between 1920 and 1980), it was clear that if the genetic code functions like a natural language then a variety of consequences follow because several preconditions must have been met. First, considering the genetic code as a natural language requires there to be a repertoire of signs (indices, icons, symbols)[a] that can be combinatorially arranged according to syntactic rules, similar to words composed of the characters of the alphabet, to generate information. Without syntactic rules to determine correct sequence order, the combination of signs could not carry informational content, that is, meaning (similarly, if our natural language had no syntactic rules, meaning could not be ascribed to randomly generated collections of words). In other words, a coherent syntax excludes randomly derived mixtures of characters of an alphabet; and for humans, coherent syntax is generated by humans who are competent with the syntactic rules. This book, for example, could contain the same characters but in a random order; such a book would be meaningless. Without competent authors who combine the characters according to a set of coherent syntactic and semantic rules, meaning does not exist.

Signs cannot exist or function without sign-using agents, and agents generate signs to communicate. In communicative action, agents can both exchange messages about

[a] According to the founder of semiotics, Charles Sanders Peirce, we are able to differentiate three different kinds of signs: indices, icons, and symbols. Indices are, in most cases, abiotic stimuli from the environment that are interpreted in the realm of memory, for example, a plant root identifies nutrients as being relevant, just as the plant shoot does with the angle of sunlight. Icons are biotic one-to-one signals (analogue) that need no further explanation, for example, plant cells identify auxin in a hormonal coordination process. There are also symbols, that is, signs or sequences of signs, like characters of an alphabet used to generate words, sentences, codes, which do not indicate by themselves what they mean (what their function could be) but are signs through natural or cultural conventions. Such sequences may also be sequences of behaviors, like the dance of the honey bees in colder hemispheres.

Natural Genetic Engineering and Natural Genome Editing: Ann. N.Y. Acad. Sci. 1178: 1–5 (2009).
doi: 10.1111/j.1749-6632.2009.05021.x

something and coordinate (organize) common behavior, that is, group behavior. A necessary condition for coordinated behavior among more than one agent is reciprocal communication that relies on pragmatic rules of information exchange that enables successful interaction. An additional element is that these pragmatic rules of information exchange must be embedded in the real-life contexts of the sign-using agents. As we know through modern communicative action theory, one agent alone (*solus ipse*) would not be able to generate signs with which it could establish sign-mediated interactions. In other words, sign use within communicative actions is inherently interconnected with group behavior and the cultural history of group identity.

According to the needs of communicating agents, syntactically correct combined signs can transport messages with meaning (semantics). But the meaning of signs depends on the context wherein agents are interwoven. As the context varies, agents will use the same signs to transport different messages. In plant communication, for example, the hormone auxin may transport various messages either in the context of neuronal-like cell–cell communication, as a hormone to transport messages between root and shoot, or as a morphogenic sign. Real-life languages or codes that are used in communication processes to coordinate and organize appropriate group behavior may differ slightly within the same species according to different habitats. The specific circumstances of a habitat of living populations for growing up and socialization may lead to special dialects, that is, the signs that are used to communicate are identical, but the meaning of the transported messages differs depending on regional customs. This phenomenon of dialects is definitively investigated in the communication of ants, honeybees, and bacteria, but transferred to the level of genetic content arrangements (i.e., the distribution of information in the genome) may enable us to investigate species-specific characteristics of genome organization.

To explore this possibility further, I developed the theory of "communicative nature": Living nature is organized and coordinated by communication processes, that is, sign-mediated interactions within and among cells, tissues, organs, and organisms, using a variety of single and group behaviors. If the genetic code has language-like features, it must embody syntactic (combination), pragmatic (context), and semantic (content) semiotic rules, rules that are conserved but under certain circumstances can be changed, rearranged, or even developed *de novo*. The linguistic feature of the genetic code excludes randomly derived sequences. No language-like text emerges as a random mixture of the alphabet or a randomly derived mixture of characters. If the genetic code has language-like features, there must be competent "agents" that generate syntactically correct nucleotide sequences that arrange and rearrange them, repair them if they are damaged, and integrate and delete them. Because no language is spoken by itself and no code can code itself, I tried to identify linguistically competent "agents" in natural nucleotide sequence editing that are responsible for generating meaningful nucleotide sequences that code for the constituent proteins required by all life forms. Editing of meaningful sequences needs "editors" and "agents": Where and what are such agents?

Fortunately, in the last few decades, it has become increasingly evident that cell functions consist of an abundant diversity of accompanying activities that are orchestrated by a variety of regulations. Together, these interconnected and fine-tuned networks represent a "high-definition information-processing organelle" (J. Shapiro). The systematics of interconnections among these cellular activities was deciphered by excellent researchers who followed the lead of Barbara McClintock. I found an excellent description of such information processing capabilities in the articles of James Shapiro. One of his articles was published in the proceedings of a congress held in 2001 called "Contextualizing the Genome: The Role of Epigenetics in Genetics, Development and Evolution." The other articles in the proceedings outlined the fact that through different reading patterns ("alternative splicing pathways generate different mRNAs from a single gene") of the genetic text induced by epigenetic regulation, such as methylation, histone modification, etc., the DNA information storage medium contains not just one meaning/function but many meanings/functions for different purposes, such as those needed in different developmental stages. The philosophical consequences of this fact were outlined by the editors, Getrudis Van de Vijver and colleagues in the first chapter of the 2001 congress proceedings. The role of repetitive elements in particular attracted my attention because they seemed to be one class of the natural genetic engineering elements proposed by James Shapiro.

After reading articles by Eva Jablonka and, later on, Frank Ryan, my attention was drawn to a book by the virologist Luis P. Villarreal, *Viruses and the Evolution of Life*. After reading the book it immediately became clear that viruses could be responsible for natural genome editing, that is, viruses would function as competent agents of integrating genetic content and genomic architecture. That is to say, viruses especially the great abundance and variety of persistent nonlytic settlers of host genomes or cytoplasm, are "linguistically competent" to combine, recombine, arrange, rearrange, repair, insert, delete, or even generate nucleotide sequences *de novo*. In addition they are evolutionarily older than cellular life forms and are coevolutionary obligate settlers of all cellular life.

I mentioned this to Alfred Winter, who works for the government of Salzburg County in Austria as one of the most successful managers in cultural affairs, and as he did with the "Gathering in Biosemiotics 6," held in 2006 in Salzburg, he immediately encouraged me to organize a new symposium and invite experts. Together with Erich Hamberger (Communication Science, University of Salzburg) and Hiltrud Oman (head administrator of the Gathering in Biosemiotics 6), I began to organize the symposium "Natural Genetic Engineering and Natural Genome Editing," which met in 2008 in Salzburg. Several articles I produced between 2006 and 2007 on biocommunication in plants, corals, fungi, bacteria, and viruses helped identify participants and many of the keynote speakers. In addition, James Shapiro and Luis Villarreal suggested other experts.

The official goals for this symposium made it clear that the presentations could serve as an outlook to the 21st century of life sciences. Over the past several years, the concept of natural genetic engineering has been advanced to encompass biochemical functions that make up the cellular toolbox for changing genome sequence composition and organization. Natural genetic engineering activities range from the introduction of point mutations by mutator polymerases, to large-scale chromosome rearrangements mediated by transposable elements and nonhomologous end joining, to incorporation of viral and microbial DNA into the genomes of host organisms. Recent literature on whole-genome sequences provides abundant evidence for the action of natural genetic engineering in evolution. Discoveries about natural genetic engineering have coincided with rapid progress in our understanding of epigenetic control and RNA-directed chromatin formation. Both natural genetic engineering and chromatin formatting exemplify the "read–write" potential of the genome as an information storage organelle. Special attention needs to be paid to the role of viruses and other so-called parasitic elements in the origin of genome formatting and natural genetic engineering capabilities. A critical question concerns the role of nonrandom genetic change operators in the production of complex evolutionary inventions. The purpose of the symposium was to bring together scientists working on genome organization, genome restructuring, genome formatting, and virus research to discuss how we could integrate these discoveries into our basic understanding of evolution, development, and disease.

The constructed sections that shaped the symposium are nearly identical to those presented in this volume. In the first part, "Information Processing Replaces Mechanics," James Shapiro pictures the radical discoveries between the elaboration of the central dogma of molecular biology and the current understanding of cell function that contradicts atomistic pre-DNA ideas of genome organization and violates the central dogma at multiple points. John Mattick proposes a new view on the evolution and genetic programming of complex organisms, suggesting that they have largely evolved by constructing more elaborate regulatory networks transacted by regulatory RNAs. Eugene Koonin describes a model of a precellular stage of biological evolution of inorganic compartments that harbored a diverse mix of virus-like genetic elements in which he not only recapitulates early ideas of J.B.S. Haldane but argues that key components of cellular life originated as components of virus-like entities. Patrick Forterre and David Prangishvili confirm the major roles of viruses in biological evolution: If capsid-encoding organisms (viruses) and ribosome-encoding organisms (cells) are the major types of living entities on our planet, it seems logical to conclude that their conflict has been the major engine of biological evolution. Eshel Ben-Jacob demonstrates that bacteria together have developed strategies of cooperation through intricate communication capabilities, such as quorum sensing, chemotactic signaling, and the exchange of genetic information.

The second part, "Viral Infection-driven Eukaryotic Evolution" opens with Philip Bell and his viral "eukaryogenesis" hypothesis in which he proposes a key role for viruses in the emergence of eukaryotes from a prokaryotic world environment. František Baluška reflects on cell–cell channels, viruses, and evolution, with the result that infection, parasitism, and symbiosis played major roles in the evolution of higher levels of biological complexity. This contribution is confirmed by two more detailed investigations: Manfred Heinlein reports about methods of viral infection through plant cell–cell channels via virus-encoded movement proteins with a variety of signal-mediated interactions, such as protection of viral RNA against host plant defense, and genetic as well as epigenetic changes in the progeny of infected plants. Amin Rustom investigates tunneling nanotube-related membrane connections among mammalian cells, which function in a similar way as plant cell–cell channels of plants to serve cellular transport needs and viral spread.

The third part, "Communal Evolution," begins with Lorraine Olendzenski and Peter Gogarten. In their contribution on the tree/web of life in light of horizontal gene transfer, they show that gene transfer between divergent organisms may provide an adaptive advantage that is even more pronounced within closely related organisms. The latter transfers may be neutral or nearly neutral for the recipient. Bruce Webb, Tonja Fisher, and Tyasning Nusawardani report on polydnaviruses as obligate symbionts of some parasitic wasps, with dramatic effects on both the viral genome and the delivery of viral genes into the wasp genome. Mariana Varela, Thomas Spencer, Massimo Palmarini, and Frederick Arnaud investigate friendly viruses that are observed in the special relationship between sheep and retroviruses, which clearly demonstrates the co-evolution between endogenous retroviruses and their mammalian hosts. Mahmoud El Hefnawi, Wessam

H. El Behaidy, Lamya A. El Housseiny, and Suher Zada show that via natural genetic engineering of hepatitis C virus NS5a for immune system counterattack, this protein plays an important role in destabilizing the cell environment and facilitating cancer.

In the fourth section, “Modular Interacting Agents,” Jürgen Brosius favors Ed Trifonov’s definition of a code as a more complex view of the gene as an entity composed of many subgenic modules. Luis Villarreal suggests that a consortium of persistent, nonlytic viruses constituted the adaptive immune system in jawed vertebrates for the first time. The origin of the adaptive immune system most likely occurred by massive colonization events with endogenous retroviruses. According to this unexpected and novel view on the evolution of the adaptive immune system the editor encouraged Luis Villarreal to write a more detailed contribution. Jean-Nicolas Volff demonstrates that cellular genes in animal genomes are derived from retrotransposons and retroviruses. Günther Witzany proposes some agents of natural genome editing. Noncoding RNAs could be adapted versions of persistent viral agents that now act as modular tools for cellular needs. Nika Lovšin and Matija Peterlin demonstrate that the APOBEC3 protein family protects against infections of some retroviruses, which indicates that this protein family is a remnant of a persistent retroviral infection event that now wards off competing genetic parasites.

In the fifth part, “Epigenetic Control,” I. King Jordan and Ahsan Huda report on some of the many ways that transposable elements have contributed to the epigenetic regulation of human genes and are distributed nonrandomly along chromosomes. Marcella Faria and Maria Carolina Elias investigate epigenetic controls in *Trypanosoma cruzi* and show how RNA interference is lacking in their genomes, which could be seen as a symptom of alternative epigenetic controls orchestrated by parasite–host interactions. Ehud Lamm shows that the network perspective dissolves the distinction between regulatory architecture and regulatory state, consistent with the theoretical impossibility of distinguishing *a priori* between “program” and “data” and its consequences for understanding the evolution of biological categories such as epigentic–genetic. Gertrudis Van de Vijver reflects on the meaning of current epigenetic developments in biology and the consequences for the idea of a contextual, stratified determination of living systems.

In addition to the authors that have contributed to this volume, Gil Ast, Nigel Goldenfeld, Shiv Grewal, Erich Hamberger, Kalin Vetsigian, Jerica Sabotic, and Reinhard Vlasak gave presentations. The symposium was organized in cooperation with Schatzkammer Land Salzburg, Kulturelle Sonderprojekte (Alfred Winter), and in partnership with the Leopold Kohr Academy. I would also like to thank the members of the organizing committee, Peter Eckl and Nikolaus Bresgen (Department of Cell Biology, University of Salzburg); Hiltrud Oman and Alexandra Parigiani (head of administration, organization, and reception); and Pierre Madl (Department of Biophysics, University of Salzburg) for technical support. The symposium was sponsored by the Schatzkammer Land Salzburg, Kulturelle Sonderprojekte, Tecan Sales Austria, the Austrian Federal Ministry of Science and Research, the University of Salzburg, and the Leopold Kohr Academy. We gratefully acknowledge this financial support. Finally, I would like to thank the *Annals* staff of the New York Academy of Sciences for guiding this book to press with patience and professional fine-tuning, especially Douglas Braaten and Ralph Brown.

GÜNTHER WITZANY
Telos-Philosophische Praxis
Salzburg, Austria