BIOLOGICAL INFORMATION

NEW PERSPECTIVES

This page intentionally left blank

BIOLOGICAL INFORMATION

NEW PERSPECTIVES

Proceedings of a Symposium held May 31 through June 3, 2011 at Cornell University

Editors

Robert J. Marks II
Baylor University, USA

Michael J. Behe
Lehigh University, USA

William A. Dembski
Discovery Institute, USA

Bruce L. Gordon
Houston Baptist University, USA

John C. Sanford
Cornell University, USA

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data
Biological information--new perspectives : proceedings of a symposium held May 31, 2011 through June 3, 2011 at Cornell University / Robert J. Marks II, Baylor University, USA, Michael J. Behe, Lehigh University, USA, William A. Dembski, Discovery Institute, USA, Bruce L. Gordon, Houston Baptist University, USA John C. Sanford Cornell University, USA.
pages cm
Includes bibliographical references and index.
ISBN 978-9814508711 (hardcover : alk. paper)
1. Genomics--Congresses. 2. Molecular genetics--Congresses. 3. Cell interaction--Congresses. 4. Mutation (Biology)--Congresses. 5. Intelligent design (Teleology)--Congresses. I. Marks, Robert J., II (Robert Jackson), 1950–
QH426.B58 2013
572.8'629--dc23

2013016707

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Disclaimer: The Contents of this book reflect the diverse scientific perspectives of the 30 contributing authors, and these views in no way reflect the Publisher's views.

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore

Biological Information — New Perspectives

Proceedings of a Symposium held May 31, 2011 through June 3, 2011 at Cornell University

Edited by

Robert J. Marks II
Baylor University
Department of Electrical & Computer Engineering
Waco, TX, USA

Michael J. Behe
Lehigh University
Department of Biological Sciences
Bethlehem, PA, USA

William A. Dembski
Discovery Institute
Seattle, WA, USA

Bruce L. Gordon
Houston Baptist University
Department of Philosophy
Houston, TX, USA

and

John C. Sanford
Cornell University, NYSAES
Department of Horticultural Science
Geneva, NY, USA

Franzine D. Smith — Technical Editor
Chase W. Nelson — Editorial Assistant

This page intentionally left blank

Contents

Title Page v
Acknowledgements xi
General Introduction xiii

Section One **Information Theory & Biology: Introductory Comments** **1**
Robert J. Marks II

1.1.1 Biological Information — What is It? 11
Werner Gitt, Robert Compton and Jorge Fernandez

1.1.2 A General Theory of Information Cost Incurred by Successful Search 26
William A. Dembski, Winston Ewert and Robert J. Marks II

1.1.3 Pragmatic Information 64
John W. Oller, Jr.

1.2.1 Limits of Chaos and Progress in Evolutionary Dynamics 87
William F. Basener

1.2.2 Tierra: The Character of Adaptation 105
Winston Ewert, William A. Dembski and Robert J. Marks II

1.2.3 Multiple Overlapping Genetic Codes Profoundly Reduce the Probability of Beneficial Mutation 139
George Montañez, Robert J. Marks II, Jorge Fernandez and John C. Sanford

1.3.1 Entropy, Evolution and Open Systems 168
Granville Sewell

1.3.2 Information and Thermodynamics in Living Systems 179
Andy C. McIntosh

Section Two **Biological Information and Genetic Theory: Introductory Comments** **203**
John C. Sanford

2.1 Not Junk After All: Non-Protein-Coding DNA Carries Extensive Biological Information 210
Jonathan Wells

2.2 Can Purifying Natural Selection Preserve Biological Information? 232
Paul Gibson, John R. Baumgardner, Wesley H. Brewer and John C. Sanford

2.3 Selection Threshold Severely Constrains Capture of Beneficial Mutations 264
John C. Sanford, John R. Baumgardner and Wesley H. Brewer

2.4 Using Numerical Simulation to Test the "Mutation-Count" Hypothesis 298
Wesley H. Brewer, John R. Baumgardner and John C. Sanford

2.5 Can Synergistic Epistasis Halt Mutation Accumulation? Results from Numerical Simulation 312
John R. Baumgardner, Wesley H. Brewer and John C. Sanford

2.6 Computational Evolution Experiments Reveal a Net Loss of Genetic Information Despite Selection 338
Chase W. Nelson and John C. Sanford

2.7 Information Loss: Potential for Accelerating Natural Genetic Attenuation of RNA Viruses 369
Wesley H. Brewer, Franzine D. Smith and John C. Sanford

2.8 DNA.EXE: A Sequence Comparison between the Human Genome and Computer Code 385
Josiah Seaman

2.9 Biocybernetics and Biosemiosis 402
Donald Johnson

Section Three Theoretical Molecular Biology: Introductory Comments **415**
Michael J. Behe

3.1 An Ode to the Code: Evidence for Fine-Tuning in the Standard Codon Table 418
Jed C. Macosko and Amanda M. Smelser

3.2 A New Model of Intracellular Communication Based on Coherent, High-Frequency Vibrations in Biomolecules 435
L. Dent

3.3	Getting There First: An Evolutionary Rate Advantage for Adaptive Loss-of-Function Mutations *Michael J. Behe*	450
3.4	The Membrane Code: A Carrier of Essential Biological Information That Is Not Specified by DNA and Is Inherited Apart from It *Jonathan Wells*	474
3.5	Explaining Metabolic Innovation: Neo-Darwinism *versus* Design *Douglas D. Axe and Ann K. Gauger*	489
Section Four	**Biological Information and Self-Organizational Complexity Theory: Introductory Comments** *Bruce L. Gordon*	**509**
4.1	Evolution Beyond Entailing Law: The Roles of Embodied Information and Self Organization *Stuart Kauffman*	513
4.2	Towards a General Biology: Emergence of Life and Information from the Perspective of Complex Systems Dynamics *Bruce H. Weber*	533
Index		561

Acknowledgements

The Editors would like to thank Cornell University's Conference Services and Cornell University's Statler Hotel for their excellent hosting of this symposium.

We thank the various sponsors of this symposium, including: Physicians and Surgeons for Scientific Integrity; Rainbow Technologies, Inc., Aldevron, Inc.; FMS Foundation, and other anonymous donors.

We thank George Montañez and Chris Rupe for the cover artwork.

This page intentionally left blank

General Introduction

Biological information, though still in its infancy as a field of study, is widely thought to be well understood in its broad outlines. The conventional, or "old," perspective on biological information is this: biological information, in the first instance, originates through purely chemical processes. These processes produce the first replicators. Once replication is in place, the Darwinian process of natural selection acting on random variation kicks in, continually increasing the information content of these replicators. Eventually, the information generating power of chemical and Darwinian processes results in organisms as complex and sophisticated as human beings. The origin, structure, and dynamics of biological information is thus thought to reduce to a combination of stochastic chemistry and undirected evolutionary forces.

This perspective on biological information is the majority position in the scientific community. Often it fails to be fully articulated because research on chemical evolution (the chemical processes responsible for first life and thus for the first biological information) and biological evolution (the evolutionary mechanisms responsible for the subsequent history of life and thus for the increase of existing biological information) tend to be conducted by different sets of scientists with different areas of expertise. Nonetheless, one occasionally finds this perspective articulated not in pieces but fully. Nobel laureate and origin-of-life researcher Christian de Duve is a case in point. In his book *Vital Dust*, he lays out various "ages" in the history of life: The Age of Chemistry, The Age of Information, The Age of the Protocell, The Age of the Single Cell, etc. Note that chemistry starts the ball rolling and precedes information. De Duve elaborates:

> History is a continuous process that we divide, in retrospect, into ages — the Stone Age, the Bronze Age, the Iron Age — each characterized by a major innovation added to previous accomplishments. This is true also of the history of life. . . . First, there is the Age of Chemistry. It covers the formation of a number of major constituents of life, up to the first nucleic acids, and is ruled entirely by the universal principles that govern the behavior of atoms and molecules. Then comes the Age of Information, thanks to the development of special information-bearing molecules that inaugurated the new processes of Darwinian evolution and natural selection particular to the living world. [1]

The conventional perspective on biological information tends more often to be articulated in pieces. Thus Harvard chemist George Whitesides, focusing on his expertise in chemistry and setting aside the subsequent history of life, speaks to the origin of life and thus to the origin of the first biological information: "This problem [of life's origin] is one of the big ones in science. It begins to place life, and us, in the universe. Most chemists believe, as do I, that life emerged spontaneously from mixtures of molecules in the prebiotic Earth. How? I have no idea." Though short on details, Whitesides is nonetheless confident that his perspective is correct: "I believe that understanding the cell is ultimately a question of chemistry and that chemists are, in principle, best qualified to solve it. The cell is a bag — a bag containing smaller bags and helpfully organizing spaghetti — filled with a Jell-O of reacting chemicals and somehow able to replicate itself." [2]

Once life has originated and biological information is on hand, the subsequent history of life displays massive increases in information content. To explain these increases, the conventional perspective on biological information takes a thoroughly Darwinian line, elevating natural selection as the primary engine for information generation over the course of biological evolution. Richard Dawkins articulates this view as follows:

> In every generation, natural selection removes the less successful genes from the gene pool, so the remaining gene pool is a narrower subset. The narrowing is nonrandom, in the direction of improvement, where improvement is defined, in the Darwinian way, as improvement in fitness to survive and reproduce. Of course the total range of variation is topped up again in every generation by new mutation and other kinds of variation. But it still remains true that natural selection is a narrowing down from an initially wider field of possibilities, including mostly unsuccessful ones, to a narrower field of successful ones. This is analogous to the definition of information with which we began: information is what enables the narrowing down from prior uncertainty (the initial range of possibilities) to later certainty (the "successful" choice among the prior probabilities). According to this analogy, natural selection is by definition a process whereby information is fed into the gene pool of the next generation. [3]

This is the conventional, or old, perspective on the origin and evolution of biological information. All the contributors to this volume question this perspective. In its place, they propose various new perspectives — plural. Some take a clearly teleological approach, advocating intelligent agent causation as the ultimate source of biological information. Others view information as *sui generis*, as a fundamental entity not reducible to purely material factors such as chemical attraction and natural selection. And others still, while accepting a big chunk of the old perspective, think that it needs to be supplemented with self-organizational processes whose information generating powers transcend those of the old perspective. The contributors, rather than presenting a united front, attempt to explore new ground and ask insightful new questions.

But if the old perspective is so well established, why question it? Is it not a sign of recalcitrance to contradict well settled verities of the scientific community? Certainly, this can be a danger. But it is a danger only when those raising the questions are ill-informed and unqualified in the relevant sciences, and have as their main motive to derail rather than foster genuine scientific inquiry. That is the not the case with any of the contributors to this volume. Science progresses not by acceding to consensus but by breaking with it. Moreover, even with well settled scientific theories, it is healthy for science periodically to question whether those theories really hold up.

In any case, there are good reasons, readily accessible to non-experts, for thinking that the old perspective on biological information bears closer scrutiny and may well be false. Take the origin of life, where all biological information begins. Origin-of-life researchers readily admit that they don't know how life began. True, they entertain speculative ideas about life's origin, with RNA-worlds currently heading the pack. But no one in the field claims to have a precisely formulated theory with solid evidential support that explains life's origin.

Thus, Stuart Kauffman, a contributor to this volume, writes, "Anyone who tells you that he or she knows how life started on the earth some 3.45 billion years ago is a fool or a knave. Nobody knows." [4] Origin-of-life researcher Leslie Orgel similarly held that "anyone who thinks they know the solution to this problem is deluded." [5] Or consider science writer Paul Davies: "We are a very long way from comprehending the how [of life's origin]. This gulf in understanding is not merely ignorance about certain technical details, it is a major conceptual lacuna… My personal belief, for what it is worth, is that a fully satisfactory theory of the origin of life demands some radically new ideas." [6]

The origin of life is the most vexing problem facing contemporary science. It has fiercely resisted reductionist approaches to its resolution. All attempts to get life started solely through life's underlying chemistry have come up short. Could it be that although chemistry provides the medium for biological information, the information itself constitutes a message capable of riding free from the underlying medium? Could such information be a real entity — as real as the chemical constituents that embody it, and yet not reducible to them — and, dare we say, have an intelligent cause? Granted, this is itself a speculative possibility, but in a field so rife with speculation, why allow only one set of speculations (those that adhere to the old perspective) and disallow others (those that open up new possibilities)? The contributors to this volume are not offering final answers. Rather, they are raising penetrating questions precisely where the old perspective has failed to offer a promising starting point for understanding the origin of biological information.

Even so, once biological information comes on the scene at the origin of first life, don't we have a well supported theory for the increase of biological information via the Darwinian mechanism of natural selection acting on random variation? In fact, even here the old perspective on biological information comes up short. The problem, as University of Chicago molecular biologist James Shapiro notes in *Evolution: A View from the 21st Century*, is that Darwinism constitutes an oversimplification: "Molecular evidence about genome sequence changes tell us that the simplifying assumptions made in the 19th and early 20th Centuries are plainly wrong. They fail to account for the variety of cellular and genomic events we now know to have occurred." [7] Shapiro continues:

> Living cells and organisms are cognitive (sentient) entities that act and interact purposefully to ensure survival, growth, and proliferation. They possess corresponding sensory, communication, information-processing, and decision-making capabilities. Cells are built to evolve; they have the ability to alter their hereditary characteristics rapidly through well-described natural genetic engineering and epigenetic processes as well as cell mergers. [8]

The picture of life and evolution that Shapiro presents is radically at odds with the old perspective on biological information. Shapiro is not alone. Many biologists are now questioning whether conventional evolutionary theory needs to be rethought from the ground up, notably the "Altenberg 16," who started out as mainstream biologists wedded to the old perspective, but now have jumped ship because the old perspective is no longer working, at least not for them. [9]

So too, notable outsiders are beginning to question whether the old perspective is disintegrating before their very eyes. Thus Robert Laughlin, a Nobel laureate physicist who studies the properties of matter that make life possible, remarks:

> Evolution by natural selection, for instance, which Charles Darwin originally conceived as a great theory, has lately come to function more as an antitheory, called upon to cover up embarrassing experimental shortcomings and legitimize findings that are at best questionable and at worst not even wrong. Your protein defies the laws of mass action? Evolution did it! Your complicated mess of chemical reactions turns into a chicken? Evolution! The human brain works on logical principles no computer can emulate? Evolution is the cause! [10]

Note that Laughlin himself does not disavow evolution. His beef is with ill-considered conceptions of evolution and the facile use of "evolution" as a magic word to conjure away hard scientific problems, when doing so in fact merely cloaks ignorance.

Even Francisco Ayala, an otherwise staunch Neo-Darwinist (himself a protégé of Theodosius Dobzhansky, one of the key architects of the neo-Darwinian synthesis), now questions whether evolutionary theory requires fundamentally new insights: “Unfortunately, there is a lot, lot, lot to be discovered still. To reconstruct evolutionary history, we have to know how the mechanisms operate in detail, and we have only the vaguest idea of how they operate at the genetic level, how genetic change relates to development and to function. I am implying that what would be discovered would be not only details, but some major principles.” [11]

In the spring of 2011 a diverse group of scientists gathered at Cornell University with an eye on the major new principles that might be required to unravel the problem of biological information. These scientists included experts in information theory, computer science, numerical simulation, thermodynamics, evolutionary theory, whole organism biology, developmental biology, molecular biology, genetics, physics, biophysics, mathematics, and linguistics. Original scientific research was presented and discussed at this symposium, which was then written up, and constitute most of the twenty-four peer-edited papers in this volume. These papers are presented in four sections: Information Theory and Biology, Biological Information and Genetic Theory, Theoretical Molecular Biology, and Self-Organizational Complexity Theory. Each of these sections begins with an introductory chapter laying out the themes and problems to be discussed there as well providing brief summaries of the papers appearing in that section.

Many of the papers in this volume speak of biological information in the limited context of the multi-dimensional array of information encoded within a cell’s genome. Nevertheless, if we define information more broadly as “all that which is communicated,” the information within a living cell is much greater than its DNA sequence. All the components of the cell, including all the RNA and protein molecules, are continuously communicating with each other. It is recognized that there are hundreds of thousands of different types of interactions within the cell’s “interactome,” and most of these interactions in one way or another involve communication. In this sense, the amazing communication network within a cell can very reasonably be compared to the Internet.

If we extend the computer science analogy further, we can consider the genome as stored information (the “hard drive” of the cell), while the RNA, protein, and other structures can be considered the “active information” (the RAM of the cell). While many of the papers given at this symposium deal with the information within the genome, it is very important we do not forget that most biological information in the cell is above and beyond the genome. On a level entirely above and beyond all this communicated information within the cell, information is also being communicated between cells, and between organisms. On a still higher level, we have the little-understood biological information that underlies the human mind, our own intelligence, and human consciousness. All of this is biological information! There exists an unknown number of symbolic languages (the genetic code being just one of many biological codes) underlying this astounding communication labyrinth integrating all levels of biological information.

All this talk about information as a real object of study within the field of biology, however, raises the question, What exactly is information in the first place? Is it a precisely defined measurable entity? Can the study of biological information be turned into an exact science? Does biological information connect meaningfully with information theory as understood in the mathematical and engineering sciences? As University of Texas philosopher of biology Sahotra Sarkar rightly notes, "It is incumbent upon those who think that informational concepts have theoretical value in biology (that is, they explain things rather than being merely metaphors) to produce an appropriate technical concept of information for biological contexts." [12] The first section of this volume is devoted to precisely this concern. Keying off of research on evolutionary search, No Free Lunch theorems, and Conservation of Information, this section attempts to provide the theoretical underpinnings for a full-fledged theory of biological information.

In the last decades, it has become clear that biological information is crucial to our understanding of life. On completion of the Human Genome Project, former Caltech president and Nobel Prize-winning biologist David Baltimore remarked, "Modern biology is a science of information. The sequencing of the genome is a landmark of progress in specifying the information, decoding it into its many coded meanings and learning how it goes wrong in disease. While it is a moment worthy of the attention of every human, we should not mistake progress for a solution. There is yet much hard work to be done…" [13] The contributors to this volume agree and desire that their efforts here will inspire much hard work on the greater project of providing a full-fledged theory of biological information, one that is free of ideological bias and gets at the truth of the matter.

— *The Editors*

References and Notes

1. de Duve C (1995) Vital Dust: Life as a Cosmic Imperative, p. 10. Basic Books, New York.
2. Whitesides GM (2007) Revolutions in Chemistry (Priestly Medalist address). Chem & Eng News 85(13):12–17, available online at http://pubs.acs.org/cen/coverstory/85/8513cover1.html (last accessed November 10, 2011).

3. Dawkins R, “The Information Challenge,” published through the Australian Skeptics and available online at http://www.skeptics.com.au/publications/articles/the-information-challenge (last accessed November 10, 2011).
4. Kauffman S (1995) At Home in the Universe: The Search for the Laws of Self-Organization and Complexity, p. 31. Oxford University Press, New York.
5. Quoted in Jason Socrates Bardi, “Life — What We Know, and What We Don’t,” *TSRI News & Views* (the online weekly of The Scripps Research Institute) 3(30) (October 11, 2004), available online at http://www.scripps.edu/newsandviews/e_20041011/ghadiri.html (last accessed November 11, 2011).
6. Davies P (1999) The Fifth Miracle: The Search for the Origin and Meaning of Life, p. 17. Simon & Schuster, New York.
7. Shapiro JA (2011) Evolution: A View from the 21st Century, p. 128. FT Press Science, Upper Saddle River, NJ.
8. Ibid., 143.
9. Mazur S (2010) The Altenberg 16: An Exposé of the Evolution Industry. North Atlantic Books, Berkeley, CA.
10. Laughlin RB (2005) A Different Universe: Reinventing Physics from the Bottom Down, pp. 168–169. Basic Books, New York.
11. In an interview, reported in Larry Witham, *Where Darwin Meets the Bible: Creationists and Evolutionists in America* (Oxford: Oxford University Press, 2002), 90.
12. Sarkar S (2007) Doubting Darwin? Blackwell, Oxford, p. 119.
13. Baltimore D, “DNA is a Reality beyond Metaphor,” *Caltech and the Human Genome Project* (2000): available online at http://marcomm.caltech.edu/events/dna/dnabalt2.html (last accessed November 11, 2011).