

Preface

"The hardest thing to see is what is in front of our eyes."

Johan Wolfgang von Goethe

Change, the process of becoming different, is at the heart of biology. Without it, nothing makes sense. Understanding and describing change is a fundamental endeavor in mathematics and the natural sciences. Change embraces evolutionary thought but is particularly difficult to define in complex dynamical systems, such as those encountered in life. The intricate arrangement of parts in these systems is responsible for unique relationships that must be established and emergent properties that must be explored.

What makes something in biology unique? Closely related individuals in a population may differ little in terms of their genetic makeup. Yet they can exhibit marked phenotypic differences, beginning with appearance and behavior. They can react differently when exposed to mutation or stress. Their progeny can even inherit genetic or learned features differently. As taxonomical distance widens, what is unique widens to the extreme. For example, there is only $\sim 1.2\%$ difference in the genomic makeup of chimp and human at the nucleotide level¹, yet phenotypic differences are huge. We may be tempted to explain uniqueness by citing critical differences (e.g., unique proteins, gene copy variants, or regulatory mechanisms) or general patterns (e.g., differential regulation of a large number of genes). We can claim that the repertoire of proteins or noncoding RNA, or the way the expression or functioning of these molecules is regulated are the ones responsible, or that the explanation lies in the integration of the thousands of component parts that make up biological systems. The genomic revolution now provides millions of protein sequences with which to dissect what is different, yet we are far away from understanding the complexities of life.

In face of so much diversity, we could frame the question differently. What makes something in biology common? Again, we find ourselves amazed by unabated similarities in everything from molecules, molecular structures, and cellular machinery, to body plans, behavior, and language. Why homogeneity amid so much heterogeneity? The answer I believe lies in the patterns and processes that are responsible for evolutionary change and the complexity and embedded simplicity of our evolving world. Modern phylogenetics and evolutionary bioinformatics have a lot to say about how to approach these questions, especially within the molecular realm. Similarly, the emerging field of evolutionary genomics attempts to address the complexity of entire repertoires of component parts acquired by the genomic revolution by reconciling what is common and what is unique

¹The Chimpanzee Sequencing Analysis Consortium, 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 437:69–87.

at different levels of organization and along taxonomical transects. Approaches link pragmatism with theoretical discourse, the bench with the computer, and statistical inference with the hypothetico-deductive method in a more modern framework for scientific and philosophical advance. Concepts are borrowed from unlikely disciplines, such as graph theory and networks from applied mathematics and social sciences, and biologists brush comfortably with algorithms and simulations. This is probably where frontiers in science live.

These past several months have been exciting because of what they represent. This is why this book is particularly timely. The birth of Darwin, 200 years ago, and the publication of his famous book² 50 years later are clearly of profound importance. Darwin catalyzed everlasting change in science and society, revolutionizing biology and prompting the search for deeper understanding of the fundamentals of biological change. His views took more relevance in the synthesis that ensued, and their significance was later (though slowly) embraced by molecular biology. The report of the first crystallographic structure (Protein Data Bank entry 1MBN)³ 50 years ago also marks a fundamental milestone in the exploration of the atomic makeup of the molecules of life. About 50,000 molecules later, we have better understanding of the molecular machinery of the cell as we try to reconcile models of structure with biological function. The beginning of this year also marks the acquisition of the sequences of the first 1000 organisms, which now span superkingdoms Archaea, Bacteria, and Eukarya, showing that the genomic revolution is indeed full fledged and unstoppable. Postgenomic science has changed the face of biology forever. The reductionist view of understanding the living world by atomizing systems and examining their parts is slowly caving in to more integrative approaches that impinge on the power of synthesis and the integration of knowledge. Physics enters the realm of biology and vice versa, but so do computer science, statistics, mathematics, and philosophy. A marvelous time for science!

The confluence of technological feats, such as "nextgen" sequencing, with informatics and exponential increases in our biological knowledge base prompted the consideration of advances in two seemingly distinct but emerging fields of endeavor, *Evolutionary Genomics* and *Systems Biology*. This book represents an attempt to bring timely problems and insights from a selected group of researchers to the table. The treatment of advances and challenges is therefore by no means exhaustive. In fact, the purpose was to entice thought rather than to seek exhaustive coverage. Following that spirit, the book has been organized in three sections, treating aspects of evolutionary genomics and systems biology that are crucial for the "new biology". In Part I, "Evolution of Life," six chapters cover the impact that evolutionary genomics has had on our understanding of life, from the general to the specific, exploring current challenges, controversial but central issues, and the relevance of model systems in biology. In Part II, "Evolution of Molecular Repertoires," 10 chapters discuss how levels of organization map to each other, how structure and function impinge on genomic repertoires, the centrality of evolutionary models, and the hidden world of RNA molecules that pervades biology. Finally, in Part III, "Evolution of Biological Networks," four chapters explore the complexity of the cellular makeup and wiring diagram of an organism, the dynamics and emerging properties of networks, and the role of fundamental evolutionary processes in the integration of component parts in systems. I hope the readers will find each and every chapter exciting and thought provoking.

²Darwin, C.R., 1859. *On the Origin of Species by Means of Natural Selection*. Murray, London.

³Kendrew, J.C., Bodo, G., Dintzis, H.M., Parrish, R.G., Wyckoff, H.W., and Phillips, D.C., 1958. A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature* 181:662-666

This book could have not been possible without the patience and wholehearted cooperation of its contributors, which made this project feasible. I also wish to thank those who generously took time out of their busy schedules to provide valuable comments. Their input is highly appreciated. I would also like to thank Jay E. Mittenthal for his friendship and encouragement, Derek Caetano-Anollés for artwork, and Karen Chambers and her team at Wiley for being so understanding in all issues related to the production of this book. Finally, I wish to recognize the National Science Foundation for continued and enabling support and my research team for making all things possible. From a personal point of view, I am particularly grateful to my wife, Gloria, for her patience, encouragement, and understanding, but fundamentally, for the sacrifices she bore when I embarked in the pursuit of science. Without her, none of this would have been possible.

GUSTAVO CAETANO-ANOLLÉS

Urbana, Illinois
January 2010

Daria Kuzmina, *Ph.D.*, School of Chemical Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois

Otto G. Berg, *Ph.D.*, Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois

Clara Isabel Bernier, *Ph.D.*, School of Chemical Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois
Linying Chen, *Ph.D.*, Department of Biology, National University of Colombia, Bogotá, Colombia

Erin Hartwig-Buss, *Ph.D.*, University of Marburg, Marburg, Germany

Philip E. Bourne, *Ph.D.*, School of Pharmacy and Chemical and Biomolecular Sciences, University of California, San Diego, California

Gallienne Bourque, *Ph.D.*, Department of Chemistry, National University of Singapore, Singapore

Joan M. Briggs, *Ph.D.*, School of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts

Brianne Brinkley, *Ph.D.*, School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, California

Gustavo Caetano-Anollés, *Ph.D.*, Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois

Todd A. Cooper, *Ph.D.*, Department of Biochemistry and Molecular Biology, University of Colorado School of Medicine, Aurora, Colorado

Lesley J. Collins, *Ph.D.*, School of Chemistry, University of California, San Diego, California

Antoine Danchin, *Ph.D.*, School of Chemical Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois

A. P. Jason de Koning, *Ph.D.*, Department of Biochemistry and Molecular Biology, University of Colorado School of Medicine, Aurora, Colorado

Alexander Donath, *Ph.D.*, School of Chemistry, Department of Computer Science and Information Systems, University of Leipzig, Leipzig, Germany

Bernard Dujon, *Ph.D.*, School of Chemistry, University of Cambridge, Cambridge, England