
Preface

Since beginning my career in pharmaceutical research and development over 10 years ago, I have seen a remarkable acceleration in the merger of basic and applied genomic research. The pharmaceutical industry, and indeed much of the academic biomedical research community, initially viewed comparative genomics as a limited venture confined to the "holy trinity species" of medical research: mouse, rat, and human. Of course, an exception has always been infectious diseases, for which comparative genomics plays a vital role in understanding viral, bacterial, and parasitic pathogens — although the importance of looking at nonpathogenic, evolutionary immediate species was often a tough sell. However, that view is changing. Through rigorous comparative analysis, the genomes of cold-blooded vertebrate, avian, and other mammalian species are providing new understandings of the human genome. Moreover, genomic sequences are becoming available for several species that are important for drug research, such as dogs and primates, as well as more specialized applications such as bovine models for osteoarthritis and zebrafish as a model for a variety of developmental and neurological conditions.

The chapters in this book are roughly equally distributed between two sections: "Basic Research in Comparative Genomics" and "Applied Research in Comparative Genomics." My goal for organization is not to create further stratifications or subdisciplines in the field but rather to point out the commonalities and synergies across the broad field of comparative genomics. Database administrators and software engineers would ask me to select or prioritize the public genomic sequences for integration into our internal bioinformatics environment. Much to their chagrin, my stock response for selection was, "All of them!" Fortunately, that message was soon accepted and embraced. Because of the growing repertoire of species genomes, comparative genomic analysis, in particular molecular evolutionary approaches, is increasingly important in drug discovery. I hope those readers in the applied sciences see the important opportunities for mining species genomes beyond those of immediate practical utility to their field and are enlightened about technological advances in DNA sequencing and phylogenetic methods as well as understanding the impact of comparative genomics on shaping conceptual thought on the evolution of species and populations. Conversely, those with a perspective focused on more basic evolutionary issues might gain an appreciation of the utility of comparative genomics in biomedical and agricultural research.

Rather than a comprehensive volume covering every aspect of comparative genomics, this book embodies the diverse interests of prominent researchers in the field. The first section, "Basic Research in Comparative Genomics," begins with three chapters covering different challenges in the field and the methodologies used to address them. Appropriately, Michael Metzker leads with a review of the revolutionary advances in DNA sequencing technology that promise to tremendously accelerate the generation of new genomic data. Next, expanding our insight into evolution

relationships among species is one of the key benefits of comparative genomics, yet the organization and analysis of large phylogenetic data sets will require bold new approaches such as those described by Bernard Moret. The virology community has been dealing with comparative genomic data analysis longer than any other group, so the description by Chris Upton and Elliot Lefkowitz on the organization and methods applied to viruses is an example of a mature and sophisticated bioinformatics genomics resource.

The remaining four chapters in the first section cover the impact of comparative genomics on our basic understanding of the evolution and genomics of several key groups of organisms. William Martin, Tal Dagan, and Katrin Henze discuss theories derived from comparative genomics on one of the most important and controversial areas of "deep" evolution study — the evolution of the eukaryotic cell and the mitigating role of organelle biogenesis. As the most diverse and largest metazoan group, the genomics of invertebrates is now poised to provide insights into their evolution as well as the origin of vertebrates, which is discussed by Takeshi Kawashima, Eiichi Shoguchi, Yutaka Satou, and Nori Satoh. The DNA sequencing projects for many additional vertebrate species are either in progress or in the planning stage, and James Thomas provides an overview of resources and fundamental principles that are the basis for contemporary studies in comparative vertebrate genomics. Completing the basic research section is a chapter by Michael Barnes on human populations that has two messages: the application of comparative genetic analysis at the intraspecific level and insights into genetic polymorphisms linked to diseases, which is a natural segue into the second section of this book.

In the section "Applied Research in Comparative Genomics," I open with a general overview, with some examples, on the utility of comparative genomics in pharmaceutical research. The next three chapters concern the role of comparative genomics in the treatment of infectious agents. Diarmaid Hughes discusses the relevance of bacterial pathogen genomes in the renewed and urgent efforts toward novel antimicrobial drugs. Malaria and other eukaryotic parasites are the most deadly killers in the developing world, but genomic sequence data hold the promise of finding new therapies as described by Emilio Merino, Steven Sullivan, and Jane Carlton. Philippe Lemey, Koen Deforche, and Anne-Mieke Vandamme discuss the application of comparative genomics of human immunodeficiency virus (HIV) in support of acquired immunodeficiency syndrome (AIDS) research, with particular emphasis on the critical concern of drug resistance.

The next four chapters concern other human diseases and drug safety issues. Cancer cells are highly polymorphic, and understanding the patterns of mutations and chromosomal aberrations among tumor types is another application of comparative genomics as described by Timon Buys, Wan Lam, and colleagues. Another chapter by Alice Kuo, Wan Lam, and colleagues covers the emerging field of epigenomics, with an emphasis on the role of DNA methylation in cancer and the opportunities for epigenomic-based drug therapies. Understanding the universe of human drug targets and their role in disease is of critical importance to the pharmaceutical industry, and Steven Foord discusses in depth the genomics of G protein-coupled receptors with respect to neurological diseases. Evaluation of the safety of drugs and chemicals involves different model organisms, and the role of increasingly

sophisticated comparative analysis of multispecies transcriptomic data for safety assessment and toxicology studies is described by Joshua Kwekel, Lyle Burgoon, and Timothy Zacharewski.

Of course, comparative genomics has wider applications beyond biomedical and pharmaceutical research. The final two chapters examine the field of genomics in agricultural research. Michael Francki and Rudi Appels review the increasing number of plant genomics projects and their role in advancing the improvement of important crop species. Leif Andersson provides an overview of advances in domestic animal genomics that are bolstering the thousands of years of selective animal breeding for desirable traits.

Space and time did not permit comprehensive coverage of all areas of comparative genomics in this volume. In addition to environmental metagenomics, the impact of comparative genomics on bioremediation and bioprocessing is missing. Researchers for other human diseases are using genomic data from multiple species to advance their work as well. These topics are fertile grounds for some future review.

The various contributions in this book should give the sense that there is already a healthy cross-disciplinary interaction among researchers working on applied and fundamental aspects of comparative genomics. Every advance in science is built on the foundations laid earlier. If this book serves to further enlighten only a few about the excitement of comparative genomics as well as the crucial interaction and interdependency of applied and basic research, then it will have overwhelmingly achieved its objectives.

James R. Brown