Int Microbiol (2003) 6: 75–78
DOI 10.1007/s10123-003-0106-3

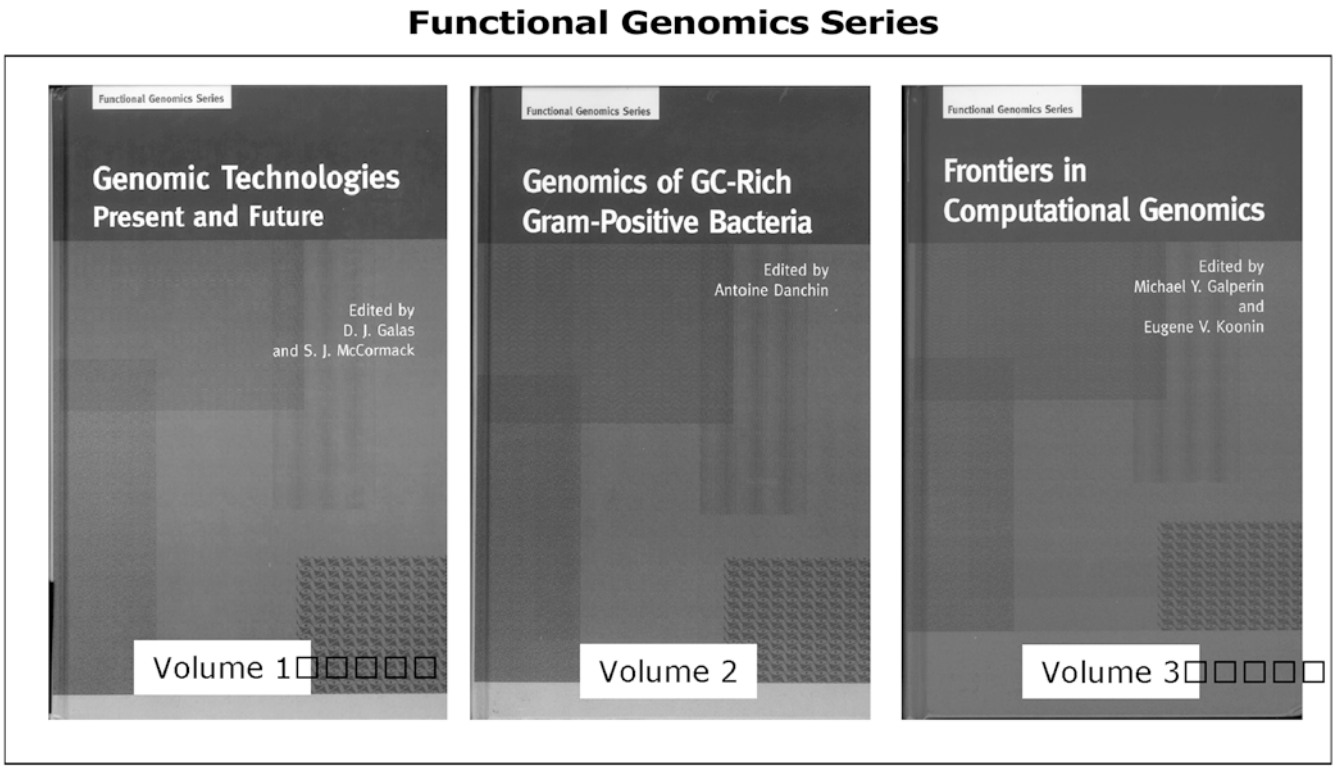
BOOK REVIEW

Mercedes Berlanga

Milton H. Saier, Jr. (series ed): Functional genomics series

Caister Academic Press, Norfolk, England

Published online: 13 March 2003



As the editors of the first volume in this series, D.J. Galas and S.J. McCormack write, "[A]dvances in understanding the 'way things work' is almost always preceded by an innovation in the technical basis for either seeing or measuring something new, or even better, doing experiments that were previously impossible." We are entering a new Golden Age, in which knowledge about bacterial genomes arrives with breathtaking speed. From the debut of whole-genome sequencing, in July 1995 (*Haemophilus influenzae*), until December 2002 80 genomes have been published, including the human genome. Genome sequencing has added a new dimension to biology; in addition to work in vivo and in vitro, we are now able to work "in silico", using computers. For the first time, the small-scale structure of living things has been revealed, portending a revolution in our thinking about biology. The *Functional Genomic Series* describes the state-of-the-art of technical means to efficiently analyze sequence information – provided by the nascent disciplines of bioinformatics and computational biology – and the need to fit the genomic pieces into the puzzle of biological function, resulting in an understanding the functioning of complex systems of genes and macromolecules.

M. Berlanga
University of Barcelona, Barcelona, Spain
E-mail: berlanga@ubxlab.com

Volume 1. David J. Galas, Stephen J. McCormack (eds): Genomic technologies. Present and future

Caister Academic Press, Norfolk, England, 2002.
418 pp, 16×24 cm (ISBN 0–9542464–2-X) £ 90.00

Volume 1, *Genomic technologies. Present and future*, describes in detail the new technologies available for studying the whole-genomes of organisms ranging from microbes to humans. These technologies allow us to query the structure and function of the macromolecules of life. In the first volume of the series, the basis of the methods used in genomic technologies is presented, but the protocols are not described – this is not a manual of materials and methods. The book is composed of 12 chapters. To provide an example of the different methodologies discussed, Chapter 4 (''Impact of transgenic technologies on functional genomics'') reviews the origins and current status of transgenic technologies. A combination of bacterial gene transfer technologies and molecular biology tools resulted in the birth of recombinant DNA technologies. The progress made in bacterial genetics inspired the development of similar methods to study mammalian genetics. Twentieth-century genetics was dominated by the idea of introducing exogenous genetic material into the genome of experimental organisms. Transgenic technologies have had basic, biomedical, and agricultural applications. Improvements in many economically important traits of livestock including rate of weight gain, milk yield and composition, and egg-laying frequency were achieved by hundreds of years of painstakingly slow, yet successful selective animal breeding. The potential for transgenic technology to greatly accelerate selection in livestock is great.

Three chapters focus on specific genetic technologies. Chapter 6 (''MAGIChip: properties and applications in genomic studies'') discusses the development and applications of micro-arrays of compounds immobilized on blocks of polyacrylamide-gel attached to the hydrophobic glass surface of a chip. Using this method, all steps of sequence analysis, PCR amplification, detachment of primers and products from the substrate, hybridization, and reading of the results can be carried out on the chip. This technology allows processes in the arrays to be monitored in both real time and steady-state, and have been successfully used to screen for a broad range of biologically meaningful genes, toxin and drug resistance genes. Other applications include the identification of bacteria and viruses and studies of rearrangements in human chromosomes. Chapter 7 (''Mass spectrometry of nucleic acids'') focuses on the use of mass spectrometry for the characterization of oligonucleotides and nucleic acids. The advantages of mass spectrometry for oligonucleotide analysis include high sensitivity, accurate molecular weight determinations, and the ability to derive structural information. Chapter 8 (''Detection of single nucleotide polymorphisms'') describes methods currently in use to study single-nucleotide polymorphism. Any two human genomes compared side by side are 99.9% identical. However, with a 3.2 billion base-pair genome, each person harbors some 3.2 million differences in his/her diploid genome. Most of the differences are due to single-base substitution polymorphisms, popularly known as single nucleotide polymorphisms (SNPs). While the majority of the SNPs are of no biological consequence, a fraction of the substitutions have functional significance and are the basis for the diversity found among humans. As genetic markers, SNPs can be used to follow the inheritance patterns of chromosomal regions from generation to generation and are powerful tools in the study of genetic factors associated with human diseases.

Volume 2. Antoine Danchin (ed): Genomics of GC-rich gram-positive bacteria

Caister Academic Press, Norfolk, England, 2002.
178 pp, 16×24 cm (ISBN 0–9542464–3-8) £ 75.00

Volume 2, *Genomics of GC-rich gram-positive bacteria*, focuses on the organization of bacterial genomes in a special group of Bacteria, the GC-rich gram-positive bacteria. This volume contains seven chapters. The reviews are the starting point for exploring the significance of genome sequences and all focus on biological implications. GC-rich gram-positive bacteria are particularly interesting since they include important human pathogens, such as *Mycobacterium tuberculosis* and *Mycobacterium leprae*. Other members, e.g. *Rhodococcus* spp., are able to biodegrade aromatic compounds, and others, e.g. *Streptomyces,* are producers of antibiotics. Most members of this group confused many early microbiologists, who classified actinomycetes (actinobacteria) as fungi due to their mycelial forms.

The genus *Rhodoccocus* contains diverse bacterial species inhabiting a variety of environmental niches ranging from polluted soils to plants and animals. Rhodococci play a key role in nature by recycling complex organic compounds including a wide range of aromatic hydrocarbons and most persistent xenobiotics. Reviewed here is the current knowledge of rhodococci genome organization, in particular the roles of linear and circular plasmids, mobile genetic elements and genomic rearrangements (Chap. 2, ''Genetic organization of *Rhodococcus*'').

One of the leading challenges for the coming decades is the complete eradication of tuberculosis (TB), together with AIDS and malaria. TB claims 2–3 million lives each year and remains the most ''lethal'' infectious disease due to a single microbe, i.e., *Mycobacterium tuberculosis*. The reasons for this dramatic situation include poor compliance of infected people with the intensive course of antibiotic therapy (6–9 months). Therefore, a better understanding of the host – mycobacteria interaction, leading to simpler treatment of TB,

is urgently needed. In particular, the virulence factors and genome organization of *M. tuberculosis* remain mostly unknown and need to be identified (Chap. 3, ''Functional genomics for the discovery of new *Mycobacterium tuberculosis* virulence factors''; Chap. 6, ''Genomics as a tool for identifying secreted proteins in bacteria''; and Chap. 7, ''TubercuList: a regularly updated database dedicated to the *Mycobacterium tuberculosis* genome'').

Streptomycetes are remarkable in terms of their morphological and metabolic differentiation. During the later stages of development, streptomycetes generally synthesize a wide range of secondary metabolites, the best known of which are antibiotics. *Streptomyces* and closely related genera produce more than 50% of the 200 commercially available antibiotics. Another remarkable feature of *Streptomyces* spp. is the structure of their genome. They have a large chromosome, consisting of an ~8-Mb linear DNA molecule. The aim of Chap. 4 (''Global analysis of *Streptomyces*'') is to briefly describe the current data regarding genome architecture, differentiation, and secondary metabolism of *Streptomyces*.

Leprosy, an infectious disease primarily of the skin and nerves, is caused by *Mycobacterium leprae*. The reservoir of *M. leprae* is not known but it has been detected in naturally infected armadillos and sooty mangabey monkeys. The route of transmission is not clear, but it is suspected to be via respiratory route, although infection via broken skin remains a possibility. In Chap. 5 (''Impact of the *Mycobacterium leprae* genome sequence on leprosy research''), the authors discuss how knowledge of the sequence of the *M. leprae* genome may be applied to eliminate leprosy in the future.

Volume 3. Michael Y. Galperin, Eugene V. Koonin (eds): Frontiers in computational genomics

Caister Academic Press, Norfolk, England, 2002.
346 pp, 16×24 cm (ISBN 0–9542464–4-6) £ 90.00

Volume 3, *Frontiers in computional genomics*, is a compendium of computational methods the function of which is to process, store in a readily accessible form, and visualize information extracted from genomes. This volume contains eleven chapters. The book reflects the different faces of computational genomics and illustrates the difficulties of converting genomic data into valuable biological knowledge. The opening chapter describes gene prediction, emphasizing recently developed methods based on the comparison of genomic sequences. Four chapters (Chap. 2, ''Sensitive protein alignment algorithms in large DNA sequences''; Chap. 3, ''Measuring generality of knowledge-based potentials extracted from protein structure sets''; Chap. 4, ''rotein structure modeling in functional genomics''; and Chap. 5, ''Identifying domains, repeats and motifs from protein sequences'') focus on protein structure as both a key to a comprehensive understanding of protein function and the source of valuable information when trying to assign a function to every gene and gene product in a genome.

The title of Chapter 6 is ''Exploiting the variations in the genomic associations of genes to predict pathways and reconstruct their evolution''. The sequencing of complete genomes from some species has led to an explosive increase in knowledge about the phylogenetic distribution of (predicted) pathways between species. By analyzing the correlation between the genome and functionally associated proteins (proteins involved in the same pathway), an explicit model of the evolution of a pathway can be constructed. In Chap. 6, the authors map the variation and reconstruct the evolution of a single pathway in detail: iron–sulfur cluster assembly in Proteobacteria. Chapter 7 (''Using genomic context in the analysis of multi-component systems'')and Chap. 8 (''Prolegomena to the evolution of transcriptional regulation in bacterial genomes'') address related issues: (1) the analysis of domain rearrangements and the emergence of new functions in one of the most common multi-component complexes in bacteria, the ABC-type transporters; (2) a reconstruction of the evolution of prokaryotic regulatory networks. The main assumption of the comparative approach to predicting regulatory sites is based on the idea that regulons are not random sets of genes; rather, they form physiologically or metabolically logical structures.

A new discipline, ''RNomics'', is introduced in Chap. 9 (''Experimental RNomics''), which uses a combination of computational (biomathematical identification) and experimental approaches in various organisms from bacteria to humans with the goal of identifying the complete RNA of a cell. A major milestone of the genome project is the identification of all genes. This can only be achieved when genes encoding non-messenger RNAs that lack open-reading frames are not ignored. Small non-messenger RNAs do not encode proteins but have various cellular functions, such as regulation of replication, transcription, and RNA processing.

The final two chapters (Chap, 10, ''Genome-scale phylogenetic trees''; Chap. 11, ''Mathematical modeling of the evolution of domain composition of proteomes: a birth-and-death process with innovation'') take on the goal of building evolutionary genomics. Genome comparisons indicate that horizontal gene transfers are the major evolutionary phenomena, at least in prokaryotes. These events cast doubt on the very feasibility of constructing a ''tree of life'' because of the diverse histories of different genes. Phylogenies based on sequences of single molecules, rRNA, have helped to define the three primary kingdoms; however, it is the genome-scale comparisons that will enhance the resolution of the phylogenetic branches .

Functional genomics series: concluding remarks

Functional genomics series starts with methods, which allow the realization of intellectual thought. But the enormous progress continuously being witnessed in genome sequencing would not have been possible without parallel developments in computer support – bioinformatics (Volume 3). One excellent example of this is provided in Volume 2 (*Genomic of GC-rich gram positive bacteria)* in a discussion of the organization of bacterial genomes in this special group of Bacteria. Reading this volume is very rewarding for microbiologists (including myself) because we can better understand (a little more) the behavior of these curious microorganisms based on their genome. Each of the three volumes contains an extensive and updated bibliography for the reader who is interested in further exploring the topics presented in this series.