PREFACE

There have been remarkable advances in our knowledge of genomes since the previous edition of this book was published ten years ago. Back in 2007, next-generation sequencing was in its infancy and high-throughput methods for transcriptomics and proteomics were only beginning to be exploited. The application of these methods over the last ten years has resulted in an exponential increase in the number of species for which genome sequences and annotations are now available, and has enabled multiple versions of the genome of a single species to be examined. The profusion of new sequences has had a particularly dramatic impact on bacterial genomics, with introduction of the pan-genome concept and the discovery of extensive lateral transfer of genes between species. Our knowledge of eukaryotic genomes has undergone equally dramatic change, with the discovery of new types of noncoding RNA, including the vast numbers of long RNAs that are transcribed from the supposedly intergenic regions of many genomes.

Genomes 4 retains the overall structure of the previous editions, with the book divided in four parts, on genome sequencing and annotation, genome anatomies, genome expression, and genome replication and evolution. With some small changes, the order of chapters remains unchanged. However, the text throughout has been completely updated and, in many chapters, substantially revised. In particular, the development of transcriptomics and proteomics has reached the point where in *Genomes 4* it is possible to describe the processes of transcription and translation from a genomewide perspective, rather than simply through an examination of the expression of individual genes. This was my aim when I wrote the first edition of *Genomes* way back in 1999, but the information available at that time meant that these core chapters were fairly orthodox treatments of gene rather than genome expression. We are still some way from being able to describe the entire expression of a genome as a single integrated process, but we are getting there and I hope that in *Genomes 4* I have been able to convey to the reader at least some aspects of the joined-up nature of genome expression.

Genomes 4 has been a long time in the making and I would like to thank Liz Owen of Garland Science for her continued enthusiasm for the book and her gentle reminders about approaching deadlines. I also wish to thank David Borrowdale and Georgina Lucas at Garland for managing the production of the book, and Matthew McClements for his splendid artwork. As with the previous editions, *Genomes 4* would not have been finished without the support of my wife, Keri. The acknowledgment in the first edition that "if you find this book useful then you should thank Keri, not me, because she is the one who ensured that it was written" is equally true for the fourth edition.

A NOTE TO THE READER

I have tried to make the fourth edition of *Genomes* as user friendly as possible. The book therefore includes a number of devices intended to help the reader and to make the book an effective teaching and learning aid.

Organization of the Book

Genomes 4 is divided into four parts:

Part I – Studying Genomes begins with an orientation chapter that introduces the reader to genomes, transcriptomes, and proteomes, and then in Chapter 2 moves on to the methods, centered on PCR and cloning, that were used in the pre-genome era to examine individual genes. The techniques that are used for constructing genetic and physical maps, which are still important in many genome projects, are then described in Chapter 3, followed in Chapter 4 by the methodology for obtaining DNA sequences and assembling reads into draft and finished genomes sequences. Two chapters are then devoted to analysis of genome sequences: Chapter 5 on the annotation of a genome by identification of genes and other features, and Chapter 6 on functional analysis of the genes that are discovered.

Part II – Genome Anatomies surveys the anatomies of the various types of genome that are found on our planet. Chapter 7 covers eukaryotic nuclear genomes, with emphasis on the human genome, partly because of the importance of the human genome in so many areas of research, but also because our genome is the best studied of all those for which sequences are available. Chapter 8 deals with the genomes of prokaryotes and of eukaryotic organelles, the latter included here because of their prokaryotic origins, and Chapter 9 describes viral genomes and mobile genetic elements, these being grouped together because some types of mobile element are related to viral genomes.

Part III – How Genomes are Expressed describes how the biological information contained in a genome is utilized by the cell within which that genome resides. Chapter 10 addresses the important issue of how the packaging of DNA into chromatin affects expression of different parts of the genome, and Chapter 11 then describes the central role that DNA-binding proteins play in expressing those parts of the genome that are active at a particular time. Chapter 12 moves on to the transcriptome, describing how transcriptomes are studied, their compositions, and how a cell's transcriptome is synthesized and maintained. Chapter 13 gives an equivalent description of proteomics and the proteome, and Chapter 14 concludes this part of the book by exploring how the genome acts within the context of the cell and organism, responding to extracellular signals and driving the biochemical changes that underlie differentiation and development.

Part IV – How Genomes Replicate and Evolve links DNA replication, mutation, and recombination with the gradual evolution of genomes over time. In Chapters 15–17 the molecular processes responsible for replication, mutation, repair, and recombination are described, and in Chapter 18 the ways in which these processes are thought to have shaped the structures and genetic contents of genomes over evolutionary time are considered. Chapter 18 then ends with a small number of case studies to illustrate how molecular phylogenomics and population genomics are being used in research and biotechnology.

LEARNING AIDS

Each chapter has a set of Short Answer Questions and In-Depth Problems, as well as an annotated Further Reading list. At the end of the book there is an extensive Glossary.

Short answer questions require 50- to 500-word answers. The questions cover the entire content of each chapter in a fairly straightforward manner, and most can be marked simply by checking each answer against the relevant part of the text. A student can use the short answer questions to work systematically through a chapter, or can select individual ones in order to evaluate their ability to answer questions on specific topics. The short answer questions could also be used in closed-book tests.

In-depth problems require a more detailed answer. They vary in nature and in difficulty, the simplest requiring little more than a literature survey, the intention of these particular problems being that the student advances his or her learning a few stages from where *Genomes 4* leaves off. Other problems require that the student evaluates a statement or a hypothesis, based on their understanding of the material in the book, possibly supplemented by reading around the subject. These problems will, hopefully, engender a certain amount of thought and critical awareness. A few problems are difficult, in some cases to the extent that there is no solid answer to the question posed. These are designed to stimulate debate and speculation, which stretches the knowledge of each student and forces them to think carefully about their statements. The in-depth problems can be tackled by students working individually, or alternatively can form the starting point for a group discussion.

Further Reading lists at the end of each chapter include those research papers, reviews, and books that I look on as the most useful sources of additional material. My intention throughout *Genomes 4* has been that students should be able to use the reading lists to obtain further information when writing extended essays or dissertations on particular topics. Research papers are therefore included, but only if their content is likely to be understandable to the average reader of the book. Emphasis is also placed on accessible reviews, one strength of these general articles being the context and relevance that they provide to a piece of work. The reading lists are divided into sections reflecting the organization of information in the chapter, and in some cases I have appended a few words summarizing the particular value of each item to help the reader decide which ones he or she wishes to seek out. In some cases, Further Reading also includes URLs for databases and other online resources relevant to the material covered in a chapter.

The **Glossary** defines every term that is highlighted in bold in the text, along with a number of additional terms that the reader might come across when referring to books or articles in the reading lists. The glossary therefore provides a quick and convenient means by which the reader can remind themselves of the technical terms relevant to the study of genomes, and also acts as a revision aid to make sure those definitions are clearly understood during the minutes of uncertainty that many students experience immediately before an exam.

INSTRUCTOR RESOURCES

The images from the book are available through www.garlandscience.com in two convenient formats: PowerPoint® and JPEG. They have been optimized for display on a computer. Figures are searchable by figure number, by figure name, or by keywords used in the figure legend from the book. Help on answering the In-Depth Problems, found at the end of each chapter, is also available.

ACKNOWLEDGMENTS

The Author and Publisher of *Genomes 4* gratefully acknowledge the contribution of the following reviewers in the development of this edition.

David Baillie, Simon Fraser University; Linda Bonen, University of Ottawa; Hugh Cam, Boston College; Yuri Dubrova, University of Leicester; Bart Eggen, University of Groningen; Robert Fowler, San José State University; Sidney Fu, George Washington University; Adrian Hall, Sheffield Hallam University; Lee Hwei Huih, Universiti Tunku Abdul Rahman; Glyn Jenkins, Aberystwyth University; Julian M. Ketley, University of Leicester; Torsten Kristensen, University of Aarhus; Gerhard May, University of Dundee; Mike McPherson, University of Leeds; Isidoro Metón, Universitat de Barcelona; Gary Ogden, St. Mary's University; Paul Overvoorde, Macalester College; John Rafferty, University of Sheffield; Andrew Read, University of Manchester; Joaquin Cañizares Sales, Universitat Politècnica de València; Michael Schweizer, Heriot-Watt University; Eric Spana, Duke University; David Studholme, Exeter University; John Taylor, University of Newcastle; Gavin Thomas, University of York; Matthew Upton, Plymouth University; Guido van den Ackerveken, Utrecht University; Vassie Ware, Lehigh University; Wei Zhang, Illinois Institute of Technology.