

Preface

This book is intended to serve as an introduction to a new area in comparative biology known as phylogenomics. Approximately 15 years ago, concurrent with the rapid and efficient sequencing of full genomes from living organisms, Clare Fraser and Jonathon Eisen coined the term "phylogenomics," a combination of phylogeny, which refers to the process whereby evolutionary trees are generated, and genomics, which represents the endeavor of obtaining genome-level data from organisms. Phylogenomics has developed into an important and compelling discipline in its own right. We developed this book in response to the students we have encountered over the last several years who are interested in applying genomics to comparative biology, specifically to phylogenetic, evolutionary, and population genetics problems.

Phylogenomics: A Primer is for advanced undergraduate students and graduate students in molecular biology, comparative biology, evolution, genomics, biodiversity, and informatics. Depending on their educational training, students can focus on the topics in the book that are of the most interest to them. Students who do not have strong backgrounds in evolution or phylogenomics will find the chapters that discuss evolutionary principles and the manipulation of phylogenomic-level data particularly useful. Conversely, students who are adept in ecology, taxonomy, and biodiversity will have the opportunity to learn about the evolution of genes and populations at the phylogenomic level and become familiar with applying phylogenomics to their genomics research.

We believe that there is no better way to understand the information that has been obtained about genes and genomes from life on this planet than to place it into context with the grand evolutionary experiment that has unfolded over the past 3.5 billion years. To this end, we have designed this book as a journey from the basic principles on which organic life has evolved, to the role of burgeoning databases in elucidating the function of proteins and organisms, and concluding with an interpretation of linear sequence information in the framework of organismal change.

Molecules are the currency of modern genomics and have an underlying linear arrangement of their component parts; that is, proteins and DNA can be broken down into linear sequences of amino acids and nucleotides, respectively. Chapters 1 and 2 present the essential principles underlying molecular biology and describe classical techniques used to analyze molecular sequences, including several high-throughput techniques that are known as "next generation" approaches. Chapter 3 explores evolution at the population level and introduces phylogenetic tree building. As a convenience, we make a simple demarcation between the evolutionary studies that focus on populations (microevolution) and those that focus on species relationships at higher-level systematic relationships (macroevolution).

Chapters 4 through 7 discuss the storage and manipulation of genomics-level data to enable the generation of the data sets that are used in phylogenomics. These processes include accessing databases and web programs such as PubMed, GenBank, and BLAST for downloading DNA and protein sequences; aligning linear sequences and producing matrices for evolutionary analysis; and assembling and annotating genomes.

Chapters 8 through 11 focus on the construction of evolutionary trees. Various approaches to phylogenetic analysis are presented, including distance, likelihood, parsimony, resampling, and Bayesian inference. In addition, the phenomenon of incongruence in relation to tree building is described as are the methods by which this problem is addressed.

Chapters 12 through 15 focus on the application of modern phylogenomics at the gene and population level. The transformation of population genetics by the use of DNA sequence information, the detection of natural selection on genes derived from genomic data, and the application of genome-level approaches to population genetics is essential to the understanding of natural populations in an evolutionary context.

The book concludes with a discussion of the basic applications of phylogenomics in the context of modern genome research. Chapter 16 examines the use of genome content to understand evolution. The role of phylogenomics in biodiversity studies, specifically the construction of the tree of life, DNA barcoding, and metagenomics, is explored in Chapter 17. The final chapter describes how functional genomics can be applied in a phylogenomic context, specifically transcription-based approaches and protein-protein interactions.

Working through the applications described in this book does not require an extensive computer science background beyond basic skills such as using a terminal or web browser. We have developed a set of Web Features that are linked to specific methods discussed in the book and are designed to introduce students to the websites used to obtain and analyze data. These features are designed to be accessed via a laptop or desktop computer and most are Web-based. A few stand-alone programs are referenced as well, all of which can be downloaded and installed on either a Mac or PC.

Rob DeSalle
New York, New York

Jeffrey A. Rosenfeld
Newark, New Jersey

Student and Instructor Resources Websites

Accessible from www.garlandscience.com, the Student and Instructor Resources websites provide learning and teaching tools created for *Phylogenomics: A Primer*. The Student Resources Site is open to everyone and users have the option to register in order to use book-marking and note-taking tools. The Instructor Resource Site requires registration and access is available only to qualified instructors. To access the Instructor Resource Site, please contact your local sales representative or email science@garland.com.

For Students

Web Features

Web-based exercises designed to assist students in working with the programs and databases used to analyze phylogenomic data.

For Instructors

Figures

The images from the book are available in two convenient formats: PowerPoint® and JPEG, which have been optimized for display. The resources may be browsed by individual chapter or a search engine. Figures are searchable by figure number, figure name, or by keywords used in the figure legend from the book.

Resources available for other Garland Science titles can be accessed via the Garland Science website.

PowerPoint is a registered trademark of Microsoft Corporation in the United States and/or other countries.