

Preface

Purpose of This Book: This textbook on collaborative research in bioinformatics and systems biology, which are key elements of modern biology and health research, highlights and provides access to many of the methods, environments, results and resources involved, including integral laboratory data generation and experimentation and clinical activities. Collaborative projects embody a research paradigm that connects many of the top scientists, institutions, their resources and research across Europe and the world, resulting in world-class contributions to bioinformatics and systems biology.

Central Themes: A number of themes are expressed and described, which guided the selection of material and its presentation:

- This book concentrates on collaborative research projects which have a significant computational biology component. A chapter with a title such as “Cancer” therefore covers the area in this restricted context. Moreover, most large-scale collaborative projects in Europe are funded by the European Commission. These projects often unify the best laboratories in Europe, which in turn are often themselves linked to worldwide programmes. Therefore, the research tends to accurately reflect the most up-to-date and best state-of-the-art research, which usually has a computational component.
- Computational approaches are a key part of much of modern life sciences research. This book aims at making researchers aware of the central importance of bioinformatics and systems biology in modern life sciences research and the wide range of publicly available resources generated by collaborative research programmes.
- A guide is needed for researchers to access the full range of resources available. Researchers are aware of laboratories, publications, databases and tools related to their areas, such as the European Bioinformatics Institute (EBI 2007) of the European Molecular Biology Laboratory (EMBL 2007), but in the form of individual pieces of a puzzle that they need for their work. They are not aware of how these resources are being linked together, nor what has been accomplished

by doing so. This book shows how collaborative researchers are putting many of the pieces together in ways accessible to the entire biomedical community.

- Collaborative research approaches are highly productive and often essential. Extensive multilaboratory collaboration is necessary for assembling the scale of resources needed to advance in many areas depending on computational biology, especially when closely linked to experimentation. European collaborative research is highly successful owing to the autonomy and flexibility given to the researchers. Tools and resources have been assembled and developed which cover much of modern biology, ranging from gene definition and alternative splicing to protein sequence, structure, function and interaction networks, with direct application made to disease processes. Many similar projects are interlinked, leading both to broad resource development and to interrelated research programmes.
- Collaborative research results involving computational approaches represent the state of the art in many areas. This book aims to describe the most advanced research results and resources available, and to make them as accessible as possible in the form of a textbook and user manual. Often the best individual resources and results are mobilised for the best collective research.
- The research and resources described in this book are of worldwide interest and relevance. Even though the projects described are mostly funded by the European Commission with predominantly European participation, these projects are strongly interactive with worldwide resources. The resources involved include access to the databases EMBL-Bank for genome sequence, UniProt for protein sequence, Ensembl for genome browsing, MSD for protein structure, etc. Therefore, even though many tools originate from European projects, gateways are provided to worldwide research and resources.
- Science management is a key element of collaborative research. Many textbooks teach the underlying science, tools and procedures necessary to carry out research, but very few discuss how to plan and carry out a research programme, especially at the collaborative level. Each stage of a project, from planning to proposal to project organisation to project operation, requires optimal organisation and structures for optimal success. This book serves as a guide to understanding methods of modern collaborative research, and to assembling the level of resources needed for the complexity of much of modern life sciences research.
- The European Commission plays a key role in creating a collaborative research environment. Another motivation is to illustrate the role of the European Commission in health research. The Commission's health research budget is smaller than the total of that of the member states of the European Union, but it is used strategically to beneficially link resources together, and is doubling under the new Seventh Framework Programme for Research (FP7 2007) compared with the Sixth Framework Programme (FP6 2007).

Structure of This Book: Following the Preface, Contents and introductory chapter, the book is organised in four parts which are somewhat analogous to the so-called central dogma of molecular biology: *sequence, structure, function, phenotype*. Part I (fundamental collaborative research and computational biology) shows the *sequence* of research approaches that integrates various elements of the “central

dogma” and much more besides, via bioinformatics, systems biology and developmental biology approaches. Part II (resources supporting bioinformatics and systems biology research) discusses the data and computational *structures* for research that have been created, and those *infrastructures* needed to generate the data. Part III (disease-related collaborative research and computational biology) exploits the *function* of the research and tools to study infectious and major diseases, including cancer. The chapter on genetic variation and diseases explores one of the great challenges within the “central dogma”, how to integrate all the resources on germ-line and somatic genetic variation into disease research. Finally, Part IV (science management, perspectives and conclusions) explores the overall *phenotype* of research itself, what it looks like and how it is organised, its perspectives and outstanding results.

Information Available: This book provides a snapshot of much of the current state of the art in bioinformatics and systems biology research. It is also a practical guide aimed at students, academic and industrial researchers and managers in life sciences and medical research, with information and pointers to resources. Most of the results and resources described are available worldwide through the Internet and international grid connections, and link to most of the major worldwide databases and tools.

Others besides researchers will find extensive sections of this book useful. Much of the introductory chapter and the chapter on science management is intended for the general reader, and give insights into collaborative research in general, and how it is supported in particular by the European Commission. A valuable feature of this book is that it shows how research is planned, organised and carried out in a variety of areas, in contrast to books that concentrate only on the science.

Specifically, the book discusses:

- Collaborative research paradigms
- Scientific basis and current state of the art in bioinformatics and systems biology, and their applications to disease processes
- Key scientific results and ongoing research
- Resources and infrastructures created by the projects
- Practical guidance to project and related websites and software and services
- Sources in books and the scientific literature
- Methods for accessing the knowledge and linking to existing projects
- Practical information about creating and participating in collaborative research projects
- Future perspectives

How To Use This Book: This book is intended to act as a guide for life sciences and biomedical researchers to the research and resources being developed by European Commission collaborative programmes, and to the individual laboratory resources that they link together. There are several ways of finding information:

- Table of contents: The table of contents is presented as a three-level detailed table of contents for finding individual research and resource areas.

- Summary tables and lists: Tables and lists are presented in the Chap. 1 that provide access to project websites and their participants and publications. Lists are also provided of project catalogues for the whole range of health research, and to relevant resources.
- Index: The Index provides single-phrase and word access to discussions of key scientific areas.
- References and access to websites: The main access points are over 350 websites listed in References along with over 170 key published papers by the projects discussed and supporting reference books. Reference names often correspond to project titles, and give direct access to their website home pages. These websites are often gateways and portals to many relevant tools and capabilities and data-bases. The websites themselves are vast reservoirs of information, with various forms of documentation, and extensive lists of journal publications resulting from project activities. The website documentation provides more information about the research process itself and the history and means of developing tools than may be found in the literature or instruction manuals. This book attempts to make that knowledge accessible, showing the resources available and their organisation.
- Reference forms: A text reference such as BioSapiens (2007) refers both to the project called BioSapiens and to the BioSapiens website, with the address listed in the References. The “2007” following the reference name indicates that it is a reference rather than just a project name. The date “2007” for such references means that the website was recently accessed in 2007 and is currently available, even though it may contain material from a variety of dates which may be earlier. All websites were verified on May 2008.
- Access for non-specialists: Non-specialist but scientifically oriented readers may wish to concentrate on chapters in the following order, skipping some technical sections: Chap. 1; Chap. 12, Chap. 11, Chap. 10, followed by the “Introduction” sections of the remaining chapters.
- Navigation through European Commission websites: Chap. 10 refers to various websites of those involved in the various stages of European Commission collaborative research programmes, including proposal preparation.

Brussels, Belgium
May 2008

Frederick B. Marcus



<http://www.springer.com/978-3-540-78352-7>

Bioinformatics and Systems Biology
Collaborative Research and Resources

Marcus, F.

2008, XXVI, 287 p., Hardcover

ISBN: 978-3-540-78352-7