“十三五”国家重点图书出版规划项目
上海高校服务国家重大战略出版工程
Translational Medicine Research
Series Editors: Zhu Chen · Xiaoming Shen
Saijuan Chen · Kerong Dai

Dongqing Wei · Yilong Ma · William C.S.Cho · Qin Xu · Fengfeng Zhou *Editor*
魏冬青　马亦龙　曹志成　徐　沁　周丰丰　著

转化医学出版工程
陈　竺　沈晓明　总　主　编
陈赛娟　戴尅戎　执行总主编

Translational Bioinformatics and Its Application

转化生物信息学及其应用

（英文版）

上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS
Springer

“十三五”国家重点图书出版规划项目
上海高校服务国家重大战略出版工程

Translational Medicine Research

Series Editors: Zhu Chen · Xiaoming Shen
Saijuan Chen · Kerong Dai

转化医学出版工程

陈　竺　沈晓明　总　主　编
陈赛娟　戴尅戎　执行总主编

Dongqing Wei · Yilong Ma · William C.S.Cho · Qin Xu · Fengfeng Zhou *Editor*
魏冬青　马亦龙　曹志成　徐　沁　周丰丰　著

Translational Bioinformatics and Its Application

转化生物信息学及其应用

（英文版）

上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

Springer

内容提要

本书系“转化医学出版工程：技术系列”之一，邀请了国内外生物、医学、信息学方面的专家学者介绍其最前沿的研究成果，从不同角度展现其研究过程中应用转化医学信息学相关方法或工具的实例，主要包括以下几个方面：(1)组学数据分析协助精准医疗；(2)通过计算生物学方法，探索成药机制，指导药物研发；(3)计算机辅助药物发现，通过定向设计与虚拟筛选更为高效快速地使众多先导化合物向临床药物转化；(4)利用大数据，通过生物统计学、生物信息学和系统生物学方法研究复杂疾病；(5)生物成像及其他信息学技术在转化医学中的应用。本书实用性强，可极大地开拓基础和临床医学工作人员的思路。

图书在版编目(CIP)数据

转化生物信息学及其应用：英文/魏冬青等著；
—上海：上海交通大学出版社，2017
转化医学出版工程
ISBN 978-7-313-18095-7

Ⅰ.①转… Ⅱ.①魏… Ⅲ.①生物信息论—英文
Ⅳ.①Q811.4

中国版本图书馆 CIP 数据核字(2017)第 219029 号

Not for sale outside the Mainland of China
(Not for sale in Hong Kong SAR, Macau SAR, and Taiwan,
and all countries except the Mainland of China)

转化生物信息学及其应用(英文版)

著　　者：魏冬青等
出版发行：上海交通大学出版社　　地　　址：上海市番禺路 951 号
邮政编码：200030　　电　　话：021-64071208
出 版 人：谈　毅
印　　制：上海锦佳印刷有限公司　　经　　销：全国新华书店
开　　本：710mm×1000mm 1/16　　印　　张：28.25
字　　数：983 千字
版　　次：2017 年 11 月第 1 版　　印　　次：2017 年 11 月第 1 次印刷
书　　号：ISBN 978-7-313-18095-7/Q
定　　价：268.00 元

Translational Medicine Research

Volume 2

Aims and Scope

In collaboration with National Infrastructures for Translational Medicine (Shanghai), the largest translational medicine research center in China, the book series "Translational Medicine Research" offers a state-of-the-art resource for physicians and researchers alike who are interested in the rapidly evolving field of translational medicine. It features original and observational investigations in the broad fields of laboratory, clinical and public health research, providing practical and up-to-date information on significant research from all subspecialties of medicine and broadening readers' horizons, from bench to bed and bed to bench.

With a focus on global interdisciplinary academic collaboration, the series aims to expedite the translation of scientific discovery into new or improved standards of management and health outcomes practice.

Series Description

Translational medicine converts promising laboratory discoveries into clinical applications and elucidates clinical questions with the use of bench work, aiming to facilitate the prediction, prevention, diagnosis and treatment of diseases. The development of translational medicine will accelerate disease control and the process of finding solutions to key health problems. It is a multidisciplinary endeavor that integrates research from the medical sciences, basic sciences and social sciences, with the aim of optimizing patient care and preventive measures that may extend beyond health care services. Therefore, close and international collaboration between all parties involved is essential to the advancement of translational medicine.

To enhance the aforementioned international collaboration as well as to provide a forum for communication and cross-pollenation between basic, translational and clinical research practitioners from all relevant established and emerging disciplines, the book series "Translational Medicine Research" features original and observational investigations in the broad fields of laboratory, clinical and public health research, aiming to provide practical and up-to-date information on significant research from all subspecialties of medicine and to broaden readers' vision horizons, from bench to bed and bed to bench.

Produced in close collaboration with National Infrastructures for Translational Medicine (Shanghai), the largest translational medicine research center in China, the book series offers a state-of-the-art resource for physicians and researchers alike who are interested in the rapidly evolving field of translational medicine. Prof. Zhu Chen, the Editor-in-Chief of the series, is a hematologist at Shanghai Jiao Tong University, China's former Minister of Health, and chairman of the center's scientific advisory board.

More information about this series at http://www.springer.com/series/13024

Translational Medicine Research

Yonghao Gui
Children's Hospital Affiliated to Fudan University

Zeguang Han
Chinese National Human Genome Center at Shanghai

Lin He
Bio-X Institutes, Shanghai Jiao Tong University

Hefeng Huang
International Peace Maternity & Child Health Hospital Affiliated to Shanghai Jiao Tong University School of Medicine

Yinghao Sun
The Second Military Medical University

Huiru Tang
School of Life Sciences, Fudan University

Hongyang Wang
ShanghaiEastern Hepatobiliary Surgery Hospital Affiliated to Second Military Medical University

Shengyue Wang
Chinese National Human Genome Center at Shanghai

Yu Wang
Chinese Center for Disease Control and Prevention

Dongqing Wei
School of Life Sciences and Biotechnology, Shanghai Jiao Tong University

Fan Wu
Shanghai Municipal Center for Disease Control & Prevention

Xuemin Xu
Med-X Institutes, Shanghai Jiao Tong University

Yixin Zeng
Department of Experimental Research, Sun Yat-Sen University Cancer Center

Chunhua Zhao
Chinese Academy of Medical Sciences & Peking Union Medical College

Yupei Zhao
Chinese Academy of Medical Sciences & Peking Union Medical College

Nanshan Zhong
The First Affiliated Hospital of Guangzhou Medical University

Series Foreword

Over the years, a chasm between biomedical researchers and the patients who may benefit from their discoveries has been opened. On one hand, millions of patients with diseases such as cancer are anxiously waiting for new remedies to save their lives. On the other hand, many exciting basic science discoveries do not have opportunities to find practical applications. Recently emerging translational medicine aims to tie basic research to clinical results and optimize both patient care and preventive measures.

Translational medicine converts promising laboratory discoveries into clinical applications and elucidates clinical questions with the use of bench work, aiming to facilitate prediction, prevention, diagnosis, and treatment of diseases. With the ultimate goal to develop more effective preventive/therapeutic approaches and improve clinical outcomes and health levels, translational medicine is therefore a people (patients and the general population as a whole)-oriented medical practice.

The past three decades have witnessed tremendous advances in China in the development of living conditions, food and nutrition, and the health care system. However, while the economy grows and society rapidly transforms, the health care system faces multiple problems. China bears a complex disease spectrum: On one hand, communicable diseases frequently seen in developing countries remain a heavy burden; on the other hand, chronic diseases commonly found in developed countries are also the leading causes of death and disability in China. The situation shows that the health care system in China is facing great challenges, and a state effort is needed to meet these challenges. Therefore China is deepening its reform to improve its people's welfare. The development of translational medicine will accelerate disease control and finding solutions for health problems.

Translational medicine is a multidisciplinary program that integrates research from the medical sciences, basic sciences, and social sciences, with the aim of optimizing patient care and preventive measures that may extend beyond health care services. Therefore, close collaboration in an international scale among all the parties is essential to the development of translational medicine.

To enhance the aforementioned international collaboration as well as to provide a forum for communication and cross-fertilization among basic, translational, and clinical research practitioners, we launch the book series "Translational Medicine Research". It features original and observational investigations in the broad fields of laboratory, clinical, and public health research, aiming to provide practical up-to-date information in significant research from all subspecialties of medicine and to broaden the readers' vision and horizon from bench to bed and bed to bench.

In close collaboration with National Infrastructures for Translational Medicine (Shanghai), the book series "Translational Medicine Research" serves as a state-of-the-art resource for physicians and translational medical researchers alike who are interested in the rapidly evolving field of translational medicine. As the Editor-in-Chief, I welcome all the researchers in related areas to report the latest bench-to-bedside researches in this series, so that the series can promote human health by accelerating the knowledge dissemination in global community.

Shanghai, China
May 2015

Zhu Chen

Preface

It was May 2015 when I was invited to join the editorial team of the "Translational Medicine Publication Project." I proposed to edit a book entitled *Translational Bioinformatics (TBI)*. I was happy to have invited a few colleagues from China and the USA who are experts in the field to join me as coeditors, Profs. Yilong Ma, William C.S. Cho, and Fengfeng Zhou. Prof. Qin Xu from my research team and my PhD student Huiyuan Zhang spent much time in managing the project. It has been many years since I started to collaborate with Springer. Our proposal was approved quickly as a collaboration project with the Shanghai Jiao Tong University Press.

TBI is an emerging field in the study of health informatics, focused on the convergence of molecular bioinformatics, biostatistics, statistical genetics, medical imaging, and clinical or medical informatics. Its focus is on applying sound informatics methodology to the increasing amount of biomedical and genomic data to formulate knowledge, disease models, and medical tools, which can be utilized by scientists, clinicians, and patients. TBI employs data mining and analytical biomedical informatics in order to generate clinical knowledge for a wide array of applications. Furthermore, it involves cross-disciplinary biomedical research to improve human health through the use of computer-based information systems. This new field has achieved great success in the recent decade by synergic integration of the molecular and genetic footprints in tissue cultures, animal models, and patients with various diseases.

Our book tries to cover, but not limited to, the following topics:

Biomedical knowledge integration
Data-driven view of disease biology
Biological knowledge assembly and interpretation
Human microbiome analysis
Pharmacogenomics
Mining electronic health records in the genomics era
Small molecules and disease

Protein interactions and disease
Network biology approach to complex diseases
Structural variation and medical genomics
Analyses using disease ontologies
Mining genome-wide genetic markers
Genome-wide association studies
Cancer genome analysis
Medical bioinformatics: biomarkers and medical imaging
Neuroinformatics of neurological and psychiatric disorders
Neuroimaging genetics

It is a challenging task that these topics are quite diversified and involved scientists with various expertise. Finally, we tried our best to summarize these diverse topics into five Parts, as in the Contents, with the chapters 2, 6, 10, 14, 16 and 17 edited by Yilong Ma, the chapters 3, 8, 11 and 13 edited by William C.S. Cho, the chapters 5, 6 and 7 edited by Qin Xu, as well as the chapters 1, 4, 9, 11, 12, 14, 15 and 16 edited by Fengfeng Zhou. My assistants Mrs. Ruili Zhao and Ms. Qiuyuan Hu made great efforts in soliciting manuscripts. Mrs. Becky Jinan Zhao from Springer and Mrs. Min Xu and Zhufeng Zhou from the Shanghai Jiao Tong University Press give us a lot of help in formulating this book and applying for funding.

In 2015, we enter the era of "precision medicine," which integrates two major contemporary developments including various omics (e.g., genomics, proteomics and metabolomics) and Big Data. I believe the TBI would play an important role in the endeavor for precision and personalized medicine.

Shanghai, China
2016-11-13

DongQing Wei

Contents

Part I
Computer-Aided Drug Discovery

Chapter 1
Drug Discovery

Geetha Ramakrishnan

Abstract An understanding of the process of drug discovery is necessary for the development of new drugs and put into clinical practice, to alleviate the diseases prevalent in modern era. This chapter covers the basic principles of how new drugs can be discovered with emphasis on target identification, lead optimization based on computer-aided drug design methods and clinical trials. The drug design principles in the pharmaceutical industry are explained based on the target and chosen ligand using molecular docking, pharmacophore modelling and virtual screening methods. The drug design is illustrated with specific examples. The clinical trials are necessary to introduce the drugs into market after due validation.

Keywords Lead compound • Computer-aided drug design • Molecular docking • Scoring functions • Virtual screening • Pharmacophore modelling • Quantitative structure-activity relationship (QSAR) • Clinical trials

1.1 Introduction

Drug discovery process deals with the root cause of the disease finding relevant genetic/biological components (i.e. drug targets) to discover lead compounds. Currently specialists in various fields, such as medicine, biochemistry, chemistry, computerized molecular modelling, pharmacology, microbiology, toxicology, physiology and pathology, contribute their research capability to achieve this goal. The drug discovery process (Fig. 1.1) in general is divided into three parts, namely, target identification, lead discovery and clinical trials.

The target identification will normally require a detailed assessment of the pathology of the disease and in some cases basic biochemical research such as study of the basic processes of life, body biochemistry and the use of metabolic analogues; study and exploitation of differences in molecular biology, differential

G. Ramakrishnan (✉)
Department of Chemistry, Sathyabama University, Rajiv Gandhi Salai, Chennai 600 119, Tamil Nadu, India
e-mail: icget2011@gmail.com

D.-Q. Wei et al. (eds.), *Translational Bioinformatics and Its Application*, Translational Medicine Research, DOI 10.1007/978-94-024-1045-7_1

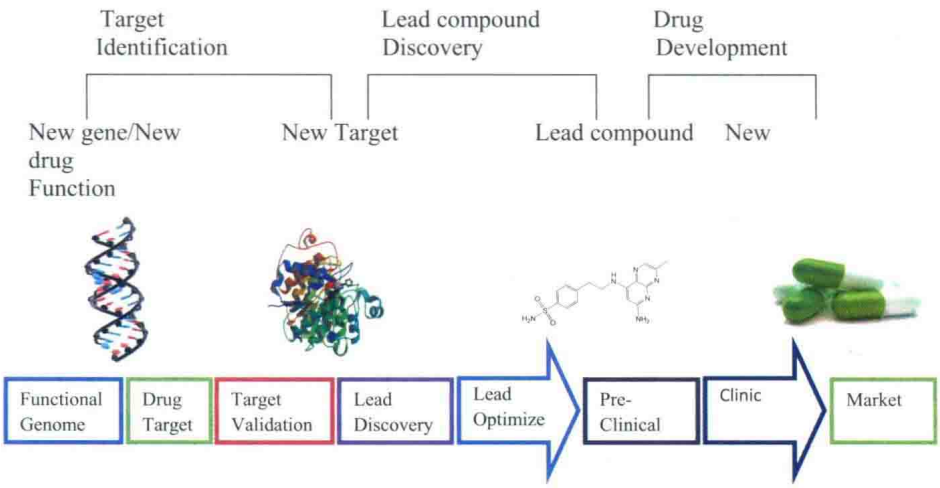

Fig. 1.1 Drug discovery process

cytology, biochemistry and endocrinology; and study of the biochemistry of diseases which will be necessary before initiating a drug design investigation.

The *lead compound* design is the most decisive step in the process of drug discovery. Methods used in lead compound design include folk/ethno-pharmacy and therapeutics, massive pharmacological screening, modification of bioactive natural products, exploitation of secondary or side effects of drugs, an approach through the molecular mechanism of drug action, drug metabolism and chemical delivery systems (Drews 1999, Bodor 1982, 1987). Numerous methods have been invented for the quantification of electronic, hydrophobic and steric effects of functional groups (Franke 1984). Statistical methods, pattern recognition/principal components analysis and cluster analysis can lead to the prediction and optimization of activity and ultimately to the design of newer drugs.

The structure of the proposed lead compound allows the medicinal/organic chemist to prepare the sample by synthetic route, and the lead compound undergoes initial pharmacological and toxicological testing. The selected lead compounds are given to animals for preclinical trials. When the lead compound has been found to be effective and safe in animal testing, it is used for human clinical trials. The lead compound is required to pass three phase clinical trials in human beings. In phase I, studies on healthy subjects are conducted to confirm safety. In phase II, studies are conducted on patients to confirm efficacy. Finally in phase III, large studies on patients are conducted to gather information about safety and efficacy at the population level.

The results of these tests enable the team to decide whether it is profitable to continue development by preparing a series of analogues, measure their activity and correlate the results to determine the drug with optimum activity.

Because of the strict prerequisites of drug authorities, which are becoming ever more demanding, the cost of drug discovery is steadily increasing. Thus, rational