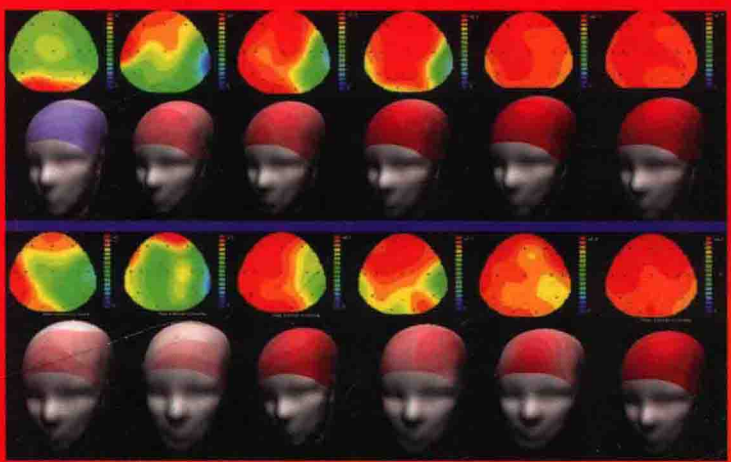
METHODS IN MOLECULAR BIOLOGY™ 448

Pharmacogenomics in Drug Discovery and Development

Edited by

Qing Yan



Humana Press

Pharmacogenomics in Drug Discovery and Development

Edited by

Qing Yan

PharmTao, Santa Clara, CA, USA

Humana Press

Editor
Qing Yan
PharmTao, Santa Clara
CA, USA
qyan@pharmtao.com

Series Editor
John M. Walker, Professor Emeritus
School of Life Sciences
University of Hertfordshire Hatfield
Hertfordshire AL10 9AB, UK

ISBN 978-1-58829-887-4 e-ISBN 978-1-59745-205-2
DOI: 10.1007/978-1-59745-205-2

Library of Congress Control Number: 2008921373

Cover illustration: Derived from Fig. 7 of Chapter 10 by Ramón Cacabelos.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Pharmacogenomics in Drug Discovery and Development

METHODS IN MOLECULAR BIOLOGY™

John M. Walker, SERIES EDITOR

460. **Essential Concepts in Toxicogenomics,** edited by Donna L. Mendrick and William B. Mattes, 2008
459. **Prion Protein Protocols,** edited by Andrew F. Hill, 2008
458. **Artificial Neural Networks:** Methods and Applications, edited by David S. Livingstone, 2008
457. **Membrane Trafficking,** edited by Ales Vancura, 2008
456. **Adipose Tissue Protocols,** Second Edition, edited by Kaiping Yang, 2008
455. **Osteoporosis,** edited by Jennifer J. Westendorf, 2008
454. **SARS- and Other Coronaviruses:** Laboratory Protocols, edited by Dave Cavanagh, 2008
453. **Bioinformatics,** Volume 2: Structure, Function, and Applications, edited by Jonathan M. Keith, 2008
452. **Bioinformatics,** Volume 1: Data, Sequence Analysis, and Evolution, edited by Jonathan M. Keith, 2008
451. **Plant Virology Protocols:** From Viral Sequence to Protein Function, edited by Gary Foster, Elisabeth Johansen, Yiguo Hong, and Peter Nagy, 2008
450. **Germline Stem Cells,** edited by Steven X. Hou and Shree Ram Singh, 2008
449. **Mesenchymal Stem Cells:** Methods and Protocols, edited by Darwin J. Prockop, Douglas G. Phinney, and Bruce A. Brunnell, 2008
448. **Pharmacogenomics in Drug Discovery and Development,** edited by Qing Yan, 2008
447. **Alcohol: Methods and Protocols,** edited by Laura E. Nagy, 2008
446. **Post-translational Modification of Proteins:** Tools for Functional Proteomics, Second Edition, edited by Christoph Kannicht, 2008
445. **Autophagosome and Phagosome,** edited by Vojo Deretic, 2008
444. **Prenatal Diagnosis,** edited by Sinhue Hahn and Laird G. Jackson, 2008
443. **Molecular Modeling of Proteins,** edited by Andreas Kukol, 2008
442. **RNAi: Design and Application,** edited by Sailen Barik, 2008
441. **Tissue Proteomics: Pathways, Biomarkers, and Drug Discovery,** edited by Brian Liu, 2008
440. **Exocytosis and Endocytosis,** edited by Andrei I. Ivanov, 2008
439. **Genomics Protocols,** Second Edition, edited by Mike Starkey and Ramnanth Elaswarapu, 2008
438. **Neural Stem Cells:** *Methods and Protocols*, Second Edition, edited by Leslie P. Weiner, 2008
437. **Drug Delivery Systems**, edited by Kewal K. Jain, 2008
436. **Avian Influenza Virus**, edited by Erica Spackman, 2008
435. **Chromosomal Mutagenesis**, edited by Greg Davis and Kevin J. Kayser, 2008
434. **Gene Therapy Protocols:** Volume 2: Design and Characterization of Gene Transfer Vectors, edited by Joseph M. LeDoux, 2008
433. **Gene Therapy Protocols:** Volume 1: Production and In Vivo Applications of Gene Transfer Vectors, edited by Joseph M. LeDoux, 2008
432. **Organelle Proteomics**, edited by Delphine Pflieger and Jean Rossier, 2008
431. **Bacterial Pathogenesis:** *Methods and Protocols*, edited by Frank DeLeo and Michael Otto, 2008

Preface

Each human is genetically distinctive, and responds differently to disease-causing factors as well as to drugs. Mechanisms inside human bodies that control drug responses are complex and multifactorial. Pharmacogenomics arose in response to such recognition of the necessity of personalized medicine, a medicine that deals with the complexity of the human body. The development of pharmacogenomics represents the evolution of biomedicine from treating the general disease itself to treating the malfunction of an individual person, the "root" of diseases. With the change of focus from diseases to humans, pharmacogenomics brings hope for the transformation from disease treatment to disease prevention.

Pharmacogenomics is considered the future of drug therapy. For the drug development industry, pharmacogenomics is useful in identifying drug targets to obtain optimal drug efficacy for certain patient populations. Because of the diversity of patients' biological backgrounds, the same disease may be caused by genetic variations in different people, who will respond differently to the same drug. Such situations require individualized treatment that avoids adverse drug responses and ensures the best possible results.

However, many challenges need to be resolved before pharmacogenomics can be applied in the clinic. These challenges include the identification of biomarker genes and pathways, the understanding of interactions between genes and drugs, and the correlation of genotypes to disease and drug response phenotypes.

In this book, we approach these challenges from three aspects. We first introduce some important cutting-edge technologies that are useful for the development of systems-based pharmacogenomics to solve the complexity; these technologies include bioinformatics, microarray, and association studies. These technologies can help us with the identification of biomarker genes and pathways and in understanding the associations among genes, drugs, and diseases.

These systems-based approaches use bioinformatics methods for studies in pharmacogenomics and systems biology to manage, organize, and understand the overwhelming information. Integrated methodologies and procedures for applying bioinformatics analysis in pharmacogenomics are presented in this book, as bioinformatics has become indispensable for almost all biopharmaceutical studies today. Pharmacogenomics-related resources, including databases and tools, are collected and provided.

Microarrays and biochips are powerful technologies for high-throughput (HTP) analysis that may enable systematic understanding of genomics and proteomics as well as large drug response data sets. The applications of microarrays in pharmacogenomics, genotyping, and clinical diagnosis, as well as the evolution and development history of the technology, are introduced in this book. Different techniques, platforms, and tests are also discussed.

Association study is a useful method in pharmacogenomics for investigating how individuals with unique genetic variants respond to a drug treatment. Confounding caused by population structure and admixture can contribute to the lack of replication of association study results. Methods for detecting and adjusting confounding are explained, as are their advantages and disadvantages.

The second aspect of this volume includes approaches to studying gene–drug interactions, that is, how drugs act and how they are processed in the human body, including drug absorption, distribution, metabolism, and excretion. Biomarkers and molecules such as ion channels, membrane transporters, receptors, and enzymes are playing increasingly essential roles in drug design and pharmacogenomics studies. These biomarkers provide critical links between drug discovery and diagnostics efforts. Updated introductions and detailed methods about studies in these molecules are provided in this book.

For example, membrane transporters are profoundly involved in drug disposition through transporting substrate drugs between organs and tissues. Investigations of genetic variations, genotyping methods, and substrate identification of membrane transporters are helpful for drug design and development. Different methods for assessing functional significance of transporter polymorphisms in vitro and in vivo as well as the application of transporter genetics in clinical pharmacology are described. Clinical significance of pharmacogenomics studies in drug-metabolizing enzymes and drug transporters for certain treatments, such as chemotherapy, is discussed in detail.

Studies of G protein-coupled receptors (GPCRs) may provide insight into disease pathways, such as the involvement of the regulator of G protein signaling (RGS) protein polymorphisms in hypertension. Pharmacogenomics of GPCR studies the involvement of genetic variations in structural and functional roles, such as GPCR activation and inactivation, their relationships with diseases, and their potential uses in defining optimized novel drug targets. These investigations can be useful for refining drug discovery as GPCR disorders are associated with a wide variety of human diseases, including retinal diseases, thyroid diseases, obesity, diabetes, asthma, cardiovascular diseases, cancer, and infectious diseases.

The third aspect composes a large part of this book: a focus on how pharmacogenomics can be used in therapeutics of diseases. These diseases include cardiovascular diseases, cancer, neurological diseases, gastrointestinal disorders, autoimmune diseases, and infectious diseases. Comprehensive information for each disease system is discussed, including biomarkers involved in the disease and the associations among genes, drugs, diseases, drug response phenotypes, and the environment.

For example, epigenetics and environmental factors may play important roles in major psychiatric disorders. Detailed methods for studying these factors are given to provide a prototype model system for better diagnosis and management of

mental diseases. Asthma is another disease caused by interactions among multiple causes, including demographic, social, environmental, and genetic factors. The most common biological pathways targeted by asthma therapy and the genetic contributions to varied therapeutic responses are described.

Drug treatment in Alzheimer's disease (AD) accounts for more than 10% of direct costs, while fewer than 20% of AD patients are fair responders to conventional drugs. Pioneering pharmacogenomics studies have shown that the therapeutic response in AD is genotype specific as pharmacogenomics factors account for more than 60% of drug variability in drug disposition. This book provides a comprehensive and detailed discussion of the pharmacogenomics of AD, from functional genomics to therapeutic strategies. The integration of these pharmacogenomics protocols with AD drug discovery and clinical practice can help promote therapeutics optimization and develop cost-effective pharmaceuticals to improve both drug efficacy and safety.

For cardiovascular diseases, methods for choosing candidate genes and single-nucleotide polymorphisms (SNPs) and the association with functional studies are discussed. These mechanistic studies are particularly important when it comes to pharmacogenomics associations. These studies provide significant and clinically relevant insights into the variable drug responses in cardiovascular disease management.

In gastroenterology and hepatology, genetic variations involved in drug metabolism or disease pathophysiology have been found to have an impact on drug responses. Discussions in this book focus on clinical pharmacogenomics of inflammatory bowel disease, *Helicobacter pylori* infections, gastroesophageal reflux disease, irritable bowel syndrome, liver transplantation, and colon cancer.

For rheumatoid arthritis, the pharmacogenomics of three major disease-modifying antirheumatic drugs (methotrexate, azathioprine, and sulfasalazine) and one class of biologic antirheumatic drugs (the tumor necrosis factor antagonists) are discussed in detail.

Cancer pharmacogenomics includes studies on biomarkers such as thiopurine methyltransferase (TPMT) and epidermal growth factor receptor (EGFR). Research methods such as germline and tumor DNA studies, polymorphism selection, and biomarker screening as well as genotyping systems are described.

Using array technology in pharmacogenomics, efficacy and systemic toxicity can be evaluated for the improvement of the design and development of preclinical vaccines. Methods of applying pharmacogenomics in the evaluation of efficacy and adverse events during clinical development of vaccines are also discussed.

By covering topics from individual molecules to systemic diseases, from fundamental concepts to advanced technologies, this book intends to provide a practical, state-of-the-art, and integrative view of the application of pharmacogenomics in drug discovery and development. I would like to thank all of the authors for their contributions to this exciting new field. I also thank the series editor, Dr. John Walker, for his help with the editing.

Qing Yan
PharmTao, Santa Clara
CA, USA

Contributors

Hamid Mostafavi Abdolmaleky, MD
Laboratory of Nutrition and Metabolism at BIDMC,
Harvard Medical School, and Biomedical Engineering Department,
Boston University, Boston, Massachusetts

Tessa M. Bosch, PharmD, PhD
Clinical Pharmacy & Toxicology, Medical Center Rijnmond-Zuid, Rotterdam,
The Netherlands

W. McIntyre Burnham, PhD
Department of Pharmacology, University of Toronto, Toronto, Ontario, Canada

Ramón Cacabelos, MD, PhD, DMSci
EuroEspes Biomedical Research Center, Institute for CNS Disorders, Bergondo,
Coruña, Spain

Michael Camilleri, MD
Clinical Enteric Neuroscience Translational and Epidemiological Research
(CENTER) Program, College of Medicine, Mayo Clinic, Rochester, Minnesota

David E. C. Cole, BSc, MD, PhD, FRCPC
Department of Laboratory Medicine and Pathobiology, Banting Institute,
University of Toronto, Toronto, Ontario, Canada

Sharon Cresci, MD
Department of Medicine, Washington University School of Medicine,
St. Louis, Missouri

William D. Figg, PharmD
Clinical Pharmacology Program, Medical Oncology Branch,
Center for Cancer Research, National Cancer Institute, Bethesda, Maryland

Rui Gao
Clinical Pharmacology Program, Medical Oncology Branch,
Center for Cancer Research, National Cancer Institute, Bethesda, Maryland

Erin R. Gardner, MS
Clinical Pharmacology Program, SAIC-Frederick Inc., NCI-Frederick, Frederick, Maryland

Gary Hardiman, PhD
BIOGEM, and Department of Medicine, University of California San Diego, La Jolla, California

Gregory A. Hawkins, PhD
Section on Pulmonary, Critical Care, Allergy and Immunologic Diseases, Center for Human Genomics, Wake Forest University School of Medicine, Winston-Salem, North Carolina

Donglei Hu, PhD
Institute for Human Genetics, Comprehensive Cancer Center, Department of Medicine, University of California San Francisco, San Francisco, California

Pedro A. Jose, MD, PhD
Department of Pediatrics, Georgetown University Medical Center, Washington, DC

Sharon Marsh, PhD
Division of Oncology, Washington University School of Medicine, St. Louis, Missouri

Lennart J. Nilsson
Division of Paediatrics, Faculty of Health Sciences, Linköping University, Sweden

Maire E. Percy, PhD
Department of Physiology and Obstetrics and Gynecology, University of Toronto, and Neurogenetics Laboratory, Surrey Place Centre, Toronto, Ontario, Canada

Stephen P. Peters, MD
Section on Pulmonary, Critical Care, Allergy and Immunologic Diseases, Center for Human Genomics, Wake Forest University School of Medicine, Winston-Salem, North Carolina

Prabha Ranganathan, MD
Department of Medicine, Washington University School of Medicine, St. Louis, Missouri

Karin J. Regnström
School of Pharmacy, University of Connecticut, Storrs, Connecticut

Yuri A. Saito, MD, MPH
Clinical Enteric Neuroscience Translational and Epidemiological Research (CENTER) Program, College of Medicine, Mayo Clinic, Rochester, Minnesota

Katherine A. Siminoivitch, MD
FRCPC, Department of Medicine, University of Toronto, Mount Sinai Hospital Samuel Lunenfeld and Toronto General Hospital Research Institutes, Toronto, Ontario, Canada

Cassandra L. Smith, PhD
Molecular Biotechnology Research Laboratory,
Biomedical Engineering Department, College of Engineering, Boston University, Boston, Massachusetts

Tristan M. Sissung, MS
Clinical Pharmacology Program, Medical Oncology Branch, Center for Cancer Research, National Cancer Institute, Bethesda, Maryland

Sam Thiagalingam, PhD
Departments of Medicine (Genetics Program), Genetics and Genomics, and Pathology and Laboratory Medicine, Boston University School of Medicine, Boston, Massachusetts

Miles D. Thompson, PhD
Department of Laboratory Medicine and Pathobiology, Banting Institute, University of Toronto, Toronto, Ontario, Canada

Qing Yan, MD, PhD
PharmTao, Santa Clara, California

Elad Ziv, MD
Institute for Human Genetics, Comprehensive Cancer Center,
Department of Medicine, University of California San Francisco,
San Francisco, California

Jin-Rong Zhou, PhD
Laboratory of Nutrition and Metabolism at BIDMC,
Department of Surgery, Harvard Medical School, Boston, Massachusetts

Contents

Chapter 1
The Integration of Personalized and Systems Medicine

Bioinformatics Support for Pharmacogenomics and Drug Discovery

Qing Yan

Summary Pharmacogenomics may have a deep impact on every drug treatment protocol to bring the right drug to the right patient. While pharmacogenomics can help achieve individualized medicine, the study of systems biology can help us understand the key issues in pharmacogenomics at different levels. These key issues include the associations between structure and function, the correlations between genotype and phenotype, and the interactions among gene, drug, and environment. Utilizing bioinformatics in pharmacogenomics that is conducted in a systemic way can help integrate information from different levels. At the molecular level, the detailed features of a gene and the relationship between genetic structure and function need to be explored. These detailed features include sequence analytic information such as sequence retrieval and structural modeling, sequence variation information, and sequence patterns that can correlate sequence structure to functional motifs. At the cellular level, the interactions and networks among those molecules should be examined. Higher degrees of understanding at the tissue and organism levels can help establish the correlations between genotype and phenotype. The application of bioinformatics methods in pharmacogenomics and systems biology should enable a more profound understanding of diseases at different levels and lead to both individualized and systems medicine. To facilitate up-to-date bioinformatics support, an integrated search engine and updated collections of tools are freely available at http://sysmed.pharmtao.com.

Keywords Bioinformatics; database; disease; drug; function; genotype; interaction; pathway; pharmacogenomics; phenotype; single-nucleotide polymorphism (SNP); software; structure; systems biology.

From: *Methods in Molecular Biology, vol. 448, Pharmacogenomics in Drug Discovery and Development*
Edited by Q. Yan © Humana Press, Totowa, NJ

1.1 Introduction

1.1.1 Pharmacogenomics, Systems Biology, and Drug Discovery

With the completion of the Human Genome Project and advancements in genetics and protein engineering, biology is entering the postgenomic era. Meanwhile, increasing development costs of new drugs and high-profile drug withdrawals are seen together with fewer approvals of new drugs *(1)*. In addition, with more than 30,000 genes in the human genome, the study of all the information has become a compellingly complex problem.

Biomedical science is facing the need for a new approach, a transformation from reductionism toward a holistic paradigm, from one-size-fits-all drugs toward personalized medicine. The emerging disciplines, pharmacogenomics and systems biology, provide hope for answering the questions and lighting the future of drug therapy. To encourage pharmacogenomic tests during drug development, the Food and Drug Administration (FDA) published the *Guidance for Industry: Pharmacogenomic Data Submissions* in 2003 to provide guidelines on using genetic data to make better and safer drugs *(2)*.

Pharmacogenomics studies the genetic basis of individual variation in response to therapeutic agents *(3)*. The investigation of genetic diversity in humans can make it possible to tailor optimal drug prescription and to bring the right drug to the right person. Pharmacogenomics may have a deep impact on every step of medical care, from diagnosis to drug prescription and from drug design to clinical trials. In an ideal condition, the application of pharmacogenomics based on the patient's genetic profile would enable the prediction of a patient's response to particular drugs and empower physicians to make right decisions for the treatment.

The effective approach to pharmacogenomics requires the integration of different disciplines, including structural genomics, functional genomics, proteomics, disease pathogenesis, pharmacology, and toxicology *(3)*. While pharmacogenomics may help achieve individualized medicine, the study of systems biology may help us understand the key issues in pharmacogenomics at different levels (*see* **Fig. 1.1**). These key issues include the associations between structure and function, the correlations between genotype and phenotype, and the interactions among genes, drugs, and the environment *(3)*. Systems biology investigates the roles biological molecules play in the context of complicated pathways and interactions and enables the understanding of disease and drug mechanisms at the system level *(4)*. Using computational methods, systems biology may help us simulate large networks of interacting components, organize biological principles, and create predictive models.

As shown in **Fig. 1.1**, the integration of pharmacogenomics and systems biology can help elucidate the mechanisms of diseases and drug actions at various levels and connect information between different levels. For example, altered genetic structure may cause malfunctions at the molecular level, which would influence the downstream interactions, pathways, and networks at the cellular level. Such changes may then lead to tissue or organ disorders that are disease phenotypes reflected as

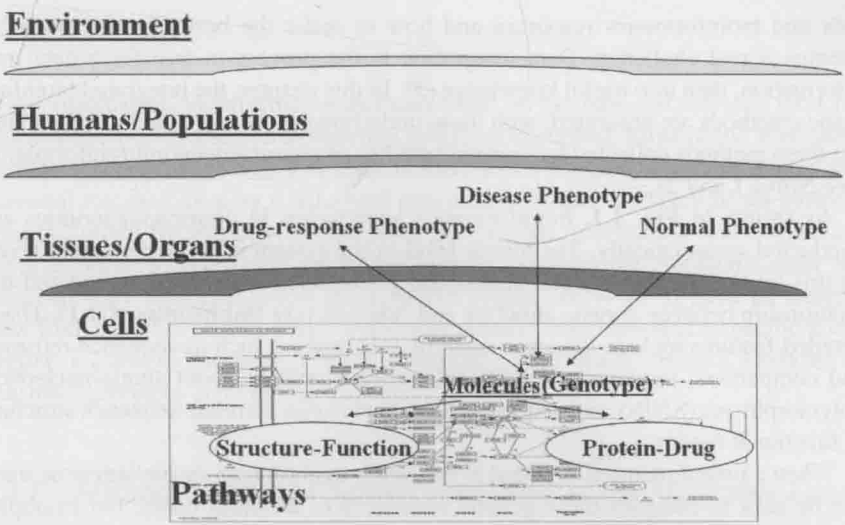


Fig. 1.1 Systematic understanding of pharmacogenomics at different levels

symptoms of the whole body. In addition, varied genetic structure and altered functions may influence the interactions between genes and drugs, which ultimately affect drug–response phenotypes. On the other hand, interactions among genes, drugs, and the environment at higher levels may also affect the structure and function of genes at the molecular level, which would in turn change downstream reactions and phenotypes, forming a feedback loop. The understanding of such an interwoven network may be the ultimate key to accurately identifying drug targets and to avoiding adverse reactions.

Handling such interwoven networks and complex feedback loops is beyond the capability of common laboratory methods, not to mention that just the complexity of scientific literature itself is already beyond measure. Help from computers and bioinformatics has become a must in today's biomedical research. In fact, bioinformatics methods have become indispensable for each step in biomedical research, from high-throughput data collection to clinical decision support. This chapter focuses on the application of bioinformatics methods in the study of pharmacogenomics, drug discovery, and systems biology.

1.1.2 Application of Bioinformatics in Pharmacogenomics and Systems Biology

Bioinformatics is using computational approaches to solve biomedical problems and to improve the communication, understanding, and management of biomedical information *(5)*. Because of the information overflow, how to integrate biomedical

data and bioinformatics resources and how to make the best use of them have become a real challenge. Data integration is the process to transform data into information, then into useful knowledge *(5)*. In this chapter, the integrated bioinformatics methods are presented, with those underlying tools and sources for carrying out these methods collected for systemic studies of pharmacogenomics information (*see* **Notes 1** and **2**).

As shown in **Fig. 1.1**, bioinformatics approaches in pharmacogenomics are conducted systematically. The lowest level in the system is at the molecular level. At this level, it is necessary to understand the detailed features of a gene and the relationship between genetic structure and function (*see* **Subheading 1.2.1**). These detailed features include sequence analytic information such as sequence retrieval and comparison, sequence variation information such as about single-nucleotide polymorphisms (SNPs), and sequence patterns that can correlate sequence structure to functional motifs.

When a time dimension is added at this level, evolutionary or phylogenetic trees can be built to compare these genetic sequences of different times. For example, tools such as the Basic Local Alignment Search Tool (BLAST) (*see* **Subheading 1.2.1.1**) and ClustalW (*see* **Subheading 1.2.1.1**) are commonly used in comparing genetic sequences and evolutionary relationships. SNP databases such as dbSNP (*see* **Subheading 1.2.1.4**) can provide information for individual genotype data.

Tools for sequence pattern analysis, including PROSITE and Pfam, are useful for correlating sequence structure to functional motifs (*see* **Subheading 1.2.1.2**). Three-dimensional (3-D) modeling of the sequence structure, such as using the database Protein Data Bank (PDB) and the program SWISS-MODEL (*see* **Subheading 1.2.1.3**), will also provide us better understanding of the structure–function relationship at this level.

In addition to sequence structural information on molecules themselves, these molecules have been categorized and classified according to their functions and interactions with drugs or other molecules. For example, some Web sites and databases can be used for studying such specific genetic molecules, including receptors and transporters.

With comprehensive examination at the molecular level, we can then scale up to the higher level to gain a more complete view of how the system works. At the cellular level, the interactions and networks among those molecules are examined. Protein–protein interaction databases and gene network and pathway databases such as Kyoto Encyclopedia of Genes and Genomes (KEGG) are usually used for this level of study (*see* **Subheading 1.2.2.1**). Resources are also available for studying even higher levels, including the tissue level and the organism level, and for making the connection between different levels. For example, some databases supply linkages between sequence variation genotypes and disease phenotypes, such as the Online Mendelian Inheritance in Man (OMIM) database (*see* **Subheading 1.2.3**).

In the following, the application of bioinformatics methods in pharmacogenomics studies is described at different levels of study to enable a relatively holistic understanding.