METHODS IN PHARMACOLOGY AND TOXICOLOGY™

Biomarker Methods in Drug Discovery and Development

Edited by

Feng Wang

Humana Press

METHODS IN PHARMACOLOGY AND TOXICOLOGY™

Biomarker Methods in Drug Discovery and Development

Edited by

Feng Wang, PhD

Procter & Gamble Pharmaceutical, Inc., Mason, Ohio

Humana Press

Editor
Feng Wang
Procter & Gamble Pharmaceutical, Inc.
8700 Mason-Montgomery Road
Mason OH, USA

Series Editor
Y. James Kang
Department of Medicine, Pharmacology & Toxicology
School of Medicine
University of Louisville
Louisville, KY, USA

ISBN: 978-1-934115-23-7 e-ISBN: 978-1-59745-463-6

Library of Congress Control Number: 2007943120

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Printed in the United States of America.

Biomarker Methods in Drug Discovery and Development

METHODS IN PHARMACOLOGY AND TOXICOLOGY™

Y. James Kang, *SERIES EDITOR*

Biomarker Methods in Drug Discovery and Development
edited by **Feng Wang**, 2008

Pharmacogenomics and Personalized Medicine
edited by **Nadine Cohen**, 2008

Cytokines in Human Health: Immunotoxicology, Pathology, and Therapeutic Applications
edited by **Robert V. House and Jacques Descotes**, 2007

Drug Metabolism and Transport: Molecular Methods and Mechanisms
edited by **Lawrence Lash**, 2005

Optimization in Drug Discovery: In Vitro Methods
edited by **Zhengyin Yan and Gary W. Caldwell**, 2004

In Vitro Neurotoxicology: Principles and Challenges
edited by **Evelyn Tiffany-Castiglioni**, 2004

Cardiac Drug Development Guide
edited by **Michael K. Pugsley**, 2003

Methods in Biological Oxidative Stress
edited by **Kenneth Hensley and Robert A. Floyd**, 2003

Apoptosis Methods in Pharmacology and Toxicology: Approaches to Measurement and Quantification
edited by **Myrtle A. Davis**, 2002

Ion Channel Localization: Methods and Protocols
edited by **Anatoli N. Lopatin and Colin G. Nichols**, 2001

Preface

The pharmaceutical industry has faced many significant challenges since the early 1990s. The fundamental issue that needs to be addressed is how to improve the efficiency of drug discovery and development. The current research and development (R&D) cost for developing a new therapeutic drug is greater than $800 million. Additionally, it takes an average of 12 years to get a new drug to market with an attrition rate greater than 90%. Reviewing the overall pharmaceutical R&D process, it has become clear that many of the drug failures are due to our lack of knowledge in population diversity, which is responsible for differences in drug efficacy and toxicity. In fact, not a single approved drug is 100% safe and efficacious for all patients. For researchers, the key question is: How can we discover biomarkers that can be used to distinguish patients who will respond to the drug without adverse effects from those who will not respond and/or will have adverse effects? There is tremendous urgency to address this question. Biomarkers also fit perfectly with the vision of *personalized medicine*, the new expectation of medical practice. This is why biomarker research has been a central focus in many research labs across academia, government agencies, and the pharmaceutical industry.

There are many different ways to define biomarkers based on molecular properties, applications, and methods. The National Institutes of Health (NIH) suggested an inclusive definition for biomarker as "a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention." A biomarker can be DNA, protein, metabolite, mRNA, or lipid. This certainly increases the complexity of biomarker research. An array of technologies is needed for biomarker research to increase the success rate. Fortunately, many existing biological and analytical technologies can be, and have already been, directly applied to biomarker research. The "omics" technologies, including genomics, proteomics, and metabonomics, have been developed and can be used for identifying potential biomarkers at several different molecular levels. The unbiased nature of these "omics"

technologies is well suited for biomarker research. The goal of *Biomarker Methods in Drug Discovery and Development* is to provide a tool box for those who have a general interest in biomarker research and also for those who are currently specializing in certain technologies but want to gain an understanding of other available methodologies. Many technologies covered in this book are well validated and mature methods, whereas others are rather new but with huge promise. This book also covers some specific issues related to clinical biomarker research, such as clinical sample handling. A total of 17 chapters contributed by many experts in their research areas provide detailed descriptions of biomarker methodologies. This book is intended to be used as a guideline and a protocol reference for biomarker researchers.

Clinical biomarker research often uses patient samples and specimens. This creates significant challenges in sample collection and handling compared with tissue and animal experiments. There are many aspects of sample collection and preservation that need to be considered in order to control and reduce experimental variability. Chapter 1 provides critical insights into careful study planning to ensure that robust data can be generated from clinical samples. The examples given in this chapter cover sample collection and handling for DNA, RNA, and protein analyses from peripheral blood. This chapter also highlights many parts of the assay under development that need to be examined so that the performance characteristics can be well understood. It is important to ensure that the biomarker data generated from the assay will be solid so that accurate conclusions can be reached.

One significant outcome of the Human Genome Project was the development of numerous genomic technologies over the past decade. Many areas of research such as drug target identification, target validation, pharmacogenomics, pharmacogenetics, as well as biomarkers have benefited greatly from these new technologies. In fact, genomics has become a major tool used in biomarker research. This book highlights six genomic methodologies including gene expression, single nucleotide polymorphism (SNP), DNA methylation, and laser capture microdissection, a very useful tissue sample retrieval technique. Adverse effects and toxicity are still being predominately identified at clinical stages of drug development using lengthy and costly approaches. Good toxicity biomarkers with high predictive value are highly desired in drug R&D. Because microarray platforms offer unique advantages in identifying novel mechanism-based biomarkers, it is a powerful method to interrogate perturbations induced by experimental drugs and to pinpoint individual genes or gene sets regulated in parallel with a toxic reaction. Chapter 2 uses three examples of

gene expression–based biomarkers for hepatotoxicity, nephrotoxicity, and general toxicity signatures in blood to describe applications for microarray platform technology in toxicity biomarkers. The detailed methodology of the fluorescent microspheres (microbeads) gene expression platform is described in Chapter 3. Using a panel of the "signature" gene expression pattern, the microsphere approach offers advantages in flexibility over the traditional whole-genome microarrays. Real-time PCR is another valuable and widely used gene expression methodology. High-throughput whole-genome microarrays enable screening of large numbers of genes to identify potential biomarkers whose expression levels are correlated with disease state, clinical outcome, and treatment regimens. These candidate biomarkers need to be validated with different sets of samples and, preferably, different methods. Real-time PCR technology fits extremely well for this purpose with excellent design flexibility and fast turnaround. Chapter 4 provides an overview of real-time PCR and practical assay protocols.

Gene expression analysis platforms are also useful in SNP identification. Chapter 5 describes the serial analysis of gene expression (SAGE) methodology and the bioinformatics approach for the applications of SNP analysis. The uniqueness of SAGE is that any molecular biology lab can easily perform the protocol without relying on specialized, expensive equipment. This chapter provides detailed methods and notes so one can readily follow the experimental procedures. Another important genomic biomarker tool is Pyrosequencing, which is covered in Chapter 6. Pyrosequencing is a genotyping method based on sequencing by synthesis. This technique offers accurate and quantitative analysis of DNA sequences without the presence of a restriction enzyme site. It can also be used to identify triallelic, indel, and short repeat polymorphisms, as well as to determine allele percentages for DNA methylation. Chapter 6 provides an overview for the Pyrosequencing method and assay details for commonly analyzed and clinically relevant polymorphisms such as SNPs in the cytochrome P450, as well as assay protocols for DNA methylation measurement.

Biological tissues have high degrees of cell heterogeneity. For some studies aimed at identifying specific biological pathways, analysis is preferably done using the targeted cell type. In Chapter 7, a relatively new tissue selection and retrieval method, laser capture microdissection (LCM), is described. This chapter demonstrates the utility of LCM when it is coupled with gene expression analyses using primate endometrium tissue. The LCM-collected samples certainly can be used for analysis utilizing other biomarker platform technologies.

Six protein biomarker analysis methodologies are described in this book (Chapters 8 to 13) in order to represent a variety of commonly used

technologies. As in genomics, there are numerous protein analysis methods that have been used in biomarker research. Traditional two-dimensional gel electrophoresis (2-DE) is a powerful protein separation method that has been constantly improved in reproducibility and ease of use. Chapter 8 uses a clinical biomarker study as an example to illustrate the application of 2-DE/mass spectrometry (MS) method. A specific drug adverse effect was investigated using 2-DE/MS with patient plasma samples. Human plasma is a convenient sample source for clinical biomarker research. However, the large dynamic range of plasma protein concentration poses a significant challenge for analysis of medium- or low-abundance proteins, which are most likely more biologically relevant. Prefractionation step(s) are often necessary before any type of proteomic analysis. Chapter 8 also describes a detailed immunochemistry method that is commonly used in many labs to deplete the most abundant plasma proteins. It demonstrates that the 2-DE/MS approach can be used for complex samples to reveal potential protein biomarkers. In an attempt to simplify the 2-DE process, the difference in-gel electrophoresis (DIGE) method was developed. As an adaptation of conventional 2-DE, DIGE uses novel fluorescent labels so that two to three samples can be resolved on a single gel under identical electrophoretic conditions. It offers advantages in simplifying image analysis, increasing sample throughput, and reducing 2-DE experimental variation. Chapter 9 provides an overview and easy-to-follow protocols for the DIGE technology.

The rapid advancement in MS technologies has resulted in many significant developments in proteomics and protein biomarker research. Currently, there are two MS-based strategies commonly used in quantitative global proteomics. Chapter 10 covers the bottom-up strategy with the shotgun tryptic peptide liquid chromatography–mass spectrometry (LC-MS) approach, and Chapter 11 describes the top-down strategy using a set of novel protein labels to identify and quantify the differences between multiple protein samples. In the bottom-up approach, proteins are first digested by an enzyme into peptide fragments that are analyzed by LC–tandem mass spectrometry (MS/MS) and then identified by database searching. Protein quantification is achieved by measuring chromatographic peak intensity. Chapter 10 provides a comprehensive review of the technique along with the author's insights pertaining to data analysis and bioinformatics. This chapter also provides a case study as a practical example to highlight the experiment design details and data interpretation. This approach is highly sensitive and can be easily automated, offering great feasibility for large-scale protein biomarker analysis. As the top-down proteomics approach involves the analysis of intact proteins, it promises the ability to characterize posttranslational modifications and to reduce false-positive identifications

because of its multiplexing capabilities. The top-down approach is well illustrated in Chapter 11 with a newly developed isobaric mass tagging technology, ExacTag Labeling system. The concept of this technology is very similar to isotope code affinity tag (ICAT). But this method labels whole proteins rather than proteolytic peptide fragments, and the subsequent protein mixtures can be enriched under the identical conditions. Experimental variation can be significantly reduced as multiple samples are tagged and mixed before any processing steps.

Other than the traditional 2-DE/MS and extensive proteomic profiling technologies introduced previously, *Biomarker Methods in Drug Discovery and Development* also covers two unique methods that are well suited for specific protein biomarker analysis needs. The surface-enhanced laser desorption ionization mass spectrometry (SELDI-MS) described in Chapter 12 is a high-throughput screening platform especially applicable for large numbers of solution samples such as plasma, urine, and conditioned media. It is a good method to quickly identify the differences among samples based on MS patterns. Further identification of the molecular entities responsible for the differences requires extra effort and most likely other technologies. When a protein biomarker is identified through any approach, validation of the biomarker with a large number of samples is required. The enzyme-linked immunosorbent assay (ELISA) array is the method of choice for validation as well as a diagnostic platform. Chapter 13 provides excellent technical details of how to generate high-quality ELISA microarrays with easy-to-follow directions and notes.

Metabolic biomarkers are also attractive in pharmaceutical biomarker research. This is because metabolite levels can be regarded as the final process readout of biological systems combining both internal factors (genetic) and external factors (disease or drug treatment). Many minor changes at the transcript and protein levels are significant enough to be biologically meaningful. These minor changes themselves may not be detectable. However, after the biological and metabolic process, they can lead to major changes at the metabolite level, which makes analysis more feasible. As an important component of systems biology and biomarker discovery, metabonomics (or metabolomics) technologies have been continually improved. MS and nuclear magnetic resonance (NMR) spectroscopy are the two core technologies used in metabolite analysis. Coupled with liquid chromatographic separation, the MS technique offers advantages in automation and high resolution. Chapter 14 discusses LC-MS methodology and its biomarker applications. It provides detailed descriptions on each component of the LC-MS platform and the multivariate statistical data analysis methods. NMR methodology is covered in Chapter 15 where a

variety of NMR analysis approaches and applications are demonstrated. Chapter 16 addresses methodology for a specific group of metabolites, nonpolar analytes, using gas chromatography–MS (GC-MS). Traditional LC-MS and NMR metabolite analyses do not cover molecules with low polarities because of the differences in separation and ionization mechanisms. Cellular lipids play important roles in membrane biology and metabolic dysfunctions. Quantitative lipid analysis, as a biomarker tool, is of interest in the pharmaceutical industry. Chapter 17 introduces MS methodologies in analyzing glycerophospholipids, the main constituents of cellular membranes.

Biomarkers offer tremendous potential and promises for transforming pharmaceutical research and development processes. Biomarker research will positively impact not only new therapeutics but also diagnostics, as well as our overall understanding in the general life sciences. We need to apply multiple technologies at different molecular levels to work on this monumental and complex task. I hope this book will be used as a biomarker technical guideline and reference to stimulate more exciting biomarker research and more technology development.

I sincerely thank all the chapter authors who tirelessly took the extra time beyond their busy daily research activities to contribute to this book, and I express my genuine gratitude to all the authors for their expert knowledge and for their efforts.

Feng Wang, PhD

Contributors

CATALIN BARBACIORU, PhD • *Molecular Biology Division, Applied Biosystems, Foster City, California*

KERRY G. BEMIS, PhD • *Indiana Centers for Applied Protein Sciences (INCAPS), Indianapolis, Indiana*

ERIC A.G. BLOMME, DVM, PhD • *Abbott Laboratories, Global Pharmaceutical Research and Development, Department of Cellular, Molecular and Exploratory Toxicology, Abbott Park, Illinois*

DAVID BRAMWELL, PhD • *Nonlinear Dynamics Limited, Newcastle upon Tyne, United Kingdom*

PIUS BRZOSKA, PhD • *Molecular Biology Division, Applied Biosystems, Foster City, California*

JANETE M. CERUTTI, PhD • *Laboratory of Molecular Endocrinology, Department of Medicine, Federal University of São Paulo, São Paulo, Brazil*

CAIFU CHEN, PhD • *Molecular Biology Division, Applied Biosystems, Foster City, California*

M. ARUNI DESILVA, MS • *Department of Chemistry, Purdue University, West Lafayette, Indiana*

PRASAD DEVARAJAN, MD • *Nephrology and Hypertension, Cincinnati Children's Hospital Medical Center, University of Cincinnati, Cincinnati, Ohio*

KIM EKROOS, PhD • *Department of Molecular Pharmacology, AstraZeneca R&D, Mölndal, Sweden*

SALLY-ANN FANCY, PhD • *Pfizer Global Research and Development, Department of Exploratory Medicinal Sciences, Sandwich, Kent, United Kingdom*

DAWN P.G. FITZPATRICK, BS • *Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, Indiana*

DAVID S. GIBSON, PhD • *Queen's University Belfast, Arthritis Research Group, Musculoskeletal Education and Research Unit, Belfast, United Kingdom*

RACHEL M. GONZALEZ, PhD • *Pacific Northwest National Laboratory, Richland, Washington*
BRAD JARROLD, MS • *Procter & Gamble Pharmaceutical, Inc., Mason Business Center, Mason, Ohio*
KENTON D. JUHLIN, PhD • *The Procter & Gamble Company, Miami Valley Innovation Center, Cincinnati, Ohio*
KEITH KERKOF, BS • *Amgen Inc., Department of Molecular Sciences, Seattle, Washington*
DAVID KEYS, PhD • *Molecular Biology Division, Applied Biosystems, Foster City, California*
LISA D. KIVMAN, BS • *Amgen Inc., Department of Molecular Sciences, Thousand Oaks, California*
NATALIA LESHINSKY, MS • *Amgen Inc., Department of Molecular Sciences, Seattle, Washington*
KELLY LI, PhD • *Molecular Biology Division, Applied Biosystems, Foster City, California*
XIN LU, PhD • *National Chromatographic R&A Center, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, China*
SHARON MARSH, PhD • *Internal Medicine, Washington University School of Medicine, St. Louis, Missouri*
JORGE M. NACIFF, PhD • *The Procter & Gamble Company, Miami Valley Innovation Center, Cincinnati, Ohio*
KIMBERLY H. NOTARI, MS • *Amgen Inc., Department of Molecular Sciences, Thousand Oaks, California*
WILLIAM C. OKULICZ, MPH, PhD • *Department of Physiology, University of Massachusetts Medical School, Worcester, Massachusetts*
CRAIG E. PARMAN, PhD • *PerkinElmer Life and Analytical Sciences, Waltham, Massachusetts*
SCOTT D. PATTERSON, PhD • *Amgen Inc., Department of Molecular Sciences, Thousand Oaks, California*
WAYNE F. PATTON, PhD • *PerkinElmer Life and Analytical Sciences, Waltham, Massachusetts*
DANIEL RAFTERY, PhD • *Department of Chemistry, Purdue University, West Lafayette, Indiana*
WILLIAM A. REES, PhD • *Amgen Inc., Department of Molecular Sciences, Seattle, Washington*
GREGORY J. RIGGINS, PhD • *Department of Neurosurgery, Johns Hopkins University Medical School, Baltimore, Maryland*
KRISTINA M. ROBSON, PhD • *Amgen Inc., Department of Molecular Sciences, Thousand Oaks, California*
GARY F. ROSS, PhD • *Bio-Rad Laboratories, Hercules, California*

KLAUS RUMPEL, PhD • *Pfizer Global Research and Development, Department of Exploratory Medicinal Sciences, Sandwich, Kent, United Kingdom*

CHRIS B. RUSSELL, PhD • *Amgen Inc., Department of Molecular Sciences, Seattle, Washington*

RAYMOND R. SAMAHA, PhD • *Molecular Biology Division, Applied Biosystems, Foster City, California*

CAITRIONA SCAIFE, MSc • *University College Dublin, Conway Institute of Biomolecular and Biomedical Research, Proteome Research Centre, Belfield, Dublin, Ireland*

NARASIMHAMURTHY SHANAIAH, PhD • *Department of Chemistry, Purdue University, West Lafayette, Indiana*

I-MEI SIU, PhD • *Department of Neurosurgery, Johns Hopkins University Medical School, Baltimore, Maryland*

SANDRO J. DE SOUZA, PhD • *Ludwig Institute for Cancer Research, Sao Paulo Branch, Hospital Alemão Oswaldo Cruz, São Paulo, Brazil*

SID SUGGS, PhD • *Amgen Inc., Department of Molecular Sciences, Thousand Oaks, California*

JAY P. TIESMAN, PhD • *The Procter & Gamble Company, Miami Valley Innovation Center, Cincinnati, Ohio*

SUZANNE M. TORONTALI, BS • *The Procter & Gamble Company, Miami Valley Innovation Center, Cincinnati, Ohio*

ALEX VARBANOV, PhD • *Procter & Gamble Pharmaceutical, Inc., Mason Business Center, Mason, Ohio*

SUSAN M. VARNUM, PhD • *Pacific Northwest National Laboratory, Richland, Washington*

FENG WANG, PhD • *Procter & Gamble Pharmaceutical, Inc., Mason Business Center, Mason, Ohio*

MU WANG, PhD • *Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, Indiana*

YULEI WANG, PhD • *Molecular Biology Division, Applied Biosystems, Foster City, California*

SCOTT E. WARDER, PhD • *Abbott Laboratories, Global Pharmaceutical Research and Development, Department of Cellular, Molecular and Exploratory Toxicology, Abbott Park, Illinois*

BING XIE, PhD • *PerkinElmer Life and Analytical Sciences, Waltham, Massachusetts*

GUOWANG XU, PhD • *National Chromatographic R&A Center, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, China*

JIN-SAM YOU, PhD • *Indiana Centers for Applied Protein Sciences (INCAPS), Indianapolis, Indiana*

JINSHENG YU, MD, PhD • *Departments of Pathology & Immunology, Washington University School of Medicine, St. Louis, Missouri*

RICHARD C. ZANGAR, PhD • *Pacific Northwest National Laboratory, Richland, Washington*

SHUCHA ZHANG, MS • *Department of Chemistry, Purdue University, West Lafayette, Indiana*

Color Plates

Color plates follow p. 230.

Color Plate 1 Fig. 1, Chapter 2: Comparison of serum alanine aminotransferase (ALT activity) and liver gene expression profiles induced by a hepatotoxicant. Rats were treated for 3 days with either the vehicle or the test article at daily doses causing moderate hepatotoxicity. Serum was collected and ALT activity levels quantified. Livers were sampled, and gene expression profiles were generated with microarrays. Whereas significant interindividual variability can be seen in serum ALT levels, the expression profiles are strikingly consistent among the eight test article–treated rats. Hierarchical cluster analysis was performed using Rosetta Resolver version 6.0 (Rosetta Inpharmatics, Seattle, WA). Genes shown include genes that were upregulated or downregulated by at least twofold with a p value less than 0.01. Green indicates downregulation, and red indicates upregulation (*see* discussion on p. 31).

Color Plate 2 Fig. 6a, b, Chapter 4: Validation of microarray results using TaqMan® Gene Expression Assay data set as reference (*see* complete caption on p. 76 and discussion on p. 74).

Color Plate 3 Fig. 7, Chapter 4: Validation of potential prognostic markers by TaqMan® assay–based real-time PCR. Eighty-five marker genes including the minimal set of 54 genes identified to best distinguish the luminal A and the basal-like subtypes were validated by TaqMan® Gene Expression assays (*see* complete caption on p. 78, 79 and discussion on p. 77).

Color Plate 4 Fig. 3, Chapter 9: Schematic representation of the processes involved in DIGE image analysis (*see* discussion on p. 204).