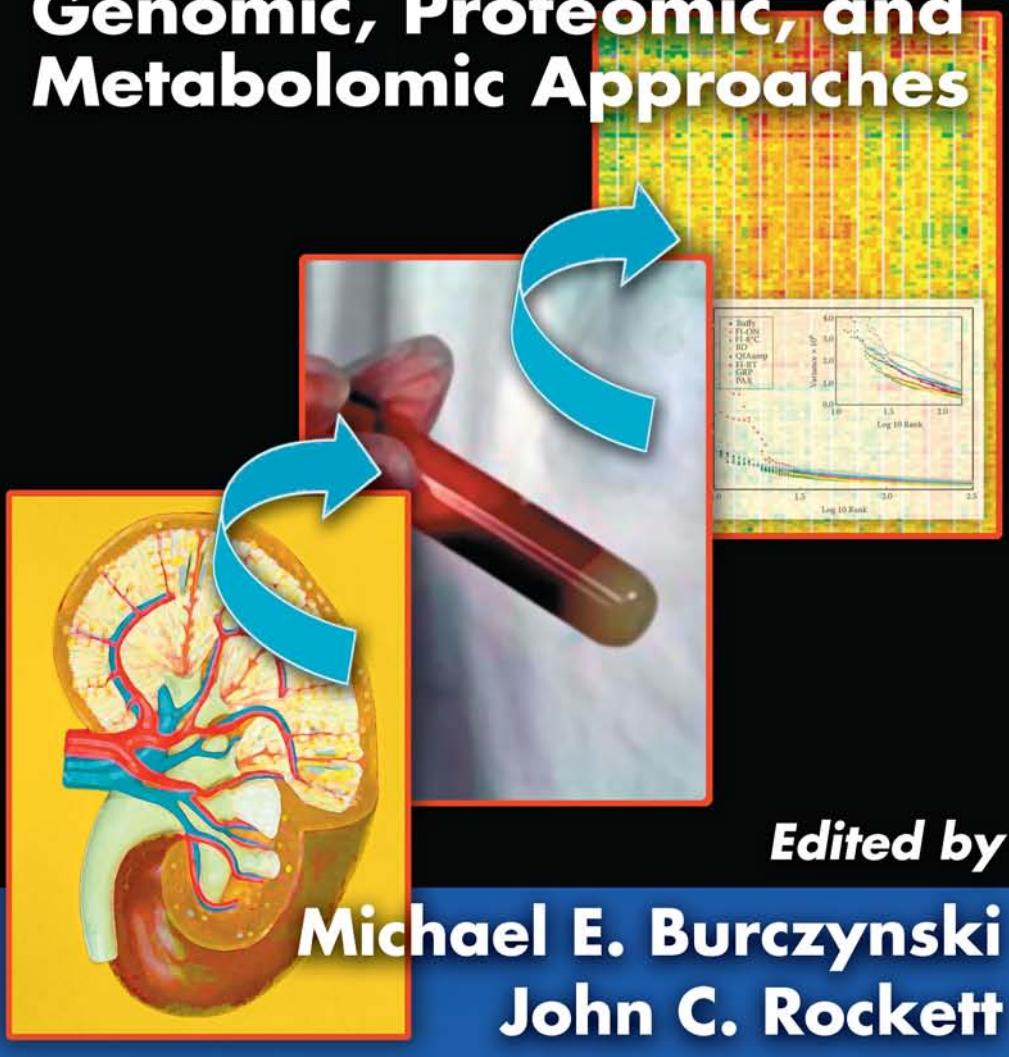
SURROGATE TISSUE ANALYSIS

Genomic, Proteomic, and Metabolomic Approaches



Edited by

Michael E. Burczynski
John C. Rockett

CRC Taylor & Francis
Taylor & Francis Group

SURROGATE TISSUE ANALYSIS

Genomic, Proteomic, and Metabolomic Approaches

SURROGATE TISSUE ANALYSIS

Genomic, Proteomic, and Metabolomic Approaches

Edited by

Michael E. Burczynski
John C. Rockett

Taylor & Francis
Taylor & Francis Group

Boca Raton London New York

A CRC title, part of the Taylor & Francis imprint, a member of the Taylor & Francis Group, the academic division of T&F Informa plc.

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20130726

International Standard Book Number-13: 978-1-4200-3809-5 (eBook - PDF)

To my lovely wife, Jennifer, and my son,
Michael William — M.E.B.

To those most dear to me — my loving wife, Gillian,
daughter Hannah Abigail,
and son Nathan David — J.C.R.

Preface

The "omic" revolution has spurred a variety of investigative techniques in a host of model systems. One of the many fields of biomedical inquiry that has benefited from the proliferation of high-throughput molecular screening methods has been the field of surrogate tissue analysis. The combination of "omic" technologies with surrogate tissue analysis has led to a rapid increase in the amount of data concerning levels of transcripts, proteins, metabolites, and other molecules present in surrogate tissues. Concomitant with this exponential increase in knowledge has been the simultaneous need to understand the relevance of these observations, and how they may be put to beneficial use.

Surrogate tissue analysis refers in general to the assessment of nontarget or off-target tissues in the body for biochemical, molecular, or cellular correlates or indicators. At its core, surrogate tissue analysis can lead to the identification of bona fide biomarkers with applications in drug discovery and development, toxicity and risk assessment, and even clinical patient management. The main attraction of surrogate tissue analysis lies in its obvious accessibility — the sampling of cerebral spinal fluid (CSF) to determine the effectiveness of a drug inhibiting neurodegeneration is eminently more feasible than the harvesting of a brain biopsy for the same purpose. Thus, it is in this manner that understanding molecular and cellular events in surrogate tissues in the context of disease, therapeutic intervention, and toxic exposure may ultimately provide the greatest benefit.

The present textbook, *Surrogate Tissue Analysis: Genomic, Proteomic, and Metabolomic Approaches*, represents a collection of chapters describing initial applications and considerations for "omic" technologies in the field of surrogate tissue analysis. The introductory chapter sets the stage for this field of inquiry and highlights some of the important issues to consider prior to conducting profiling studies in surrogate tissues. The next three sections of the textbook review specific advances in the field of genomic, proteomic, and metabolomic approaches in surrogate tissues.

In the first of these three sections, transcriptional profiling approaches in surrogate tissues are covered, and the preponderance of chapters focused on peripheral blood profiling provides hardy evidence that this field is rapidly spawning its own subspecialty — that of hemogenomics. Chapter 2 reviews the important considerations in peripheral blood profiling in great detail and summarizes results achieved when evaluations of various blood preparation platforms are used for the purpose of transcriptional profiling. Chapters 3 and 4 cover the relatively novel application of transcriptional profiling in neurological and oncological disease settings, respectively. Chapter 5 reviews the nature of surrogate tissue profiles of toxic exposure in preclinical studies where transcriptional effects in both target and surrogate tissues can be compared. Finally, Chapter 6 focuses on transcriptional profiling in a non-blood-based tissue, semen, which is utilized as a surrogate tissue for paternal exposure.

The next section focuses on proteomic and protein-based methods for identifying markers in surrogate tissues. Chapter 7 highlights mass spectrometry approaches for assessment of proteins in serum, with a focus on the obvious implications of protein-based biomarkers for detecting and monitoring early stages of cancer. Chapter 8

assesses the ability of circulating lymphocyte integrins to indicate endometrial receptivity, and Chapter 9 demonstrates how the surrogate tissue of nipple aspirate fluid can be used to detect and monitor breast cancer in afflicted patients.

The next section explores metabolomic approaches along with other novel molecular screens that can be applied in surrogate tissues for the purpose of finding biomarkers. Metabolomics is somewhat unique in that it is particularly suited to surrogate tissue analysis, since in contrast to most DNA, RNA, and intracellular proteins in the body, only metabolites (and secreted polypeptides) are freely found in surrogate tissues. Chapters 10 and 11 therefore review the field of metabolomics and how this technology is rapidly developing into a powerful technique for biomarker identification. Chapter 12 provides an excellent overview of a subfield of metabolomics, which focuses exclusively on the measurement of lipids and is appropriately termed lipidomics, and explores how the field of lipidomics can be used in surrogate tissues to provide an understanding of dynamic inflammatory responses in hosts. Chapter 13 reviews a PCR-based approach to detect and monitor metastatic cells in the circulation, and Chapter 14 covers a methylation profiling approach that can be used to accomplish a similar end.

The final section of the textbook attempts to look toward the horizon in more general terms, with chapters that focus on regulatory, economic, and pan-omic strategies, all of which will undoubtedly influence surrogate tissue analysis in the future. Chapter 15 summarizes generally applicable regulatory issues that will undoubtedly be important considerations for those biomarkers discovered in surrogate tissue profiling studies that support drug/co-diagnostic registration and require regulatory approval. Chapter 16 provides an esoteric and interesting evaluation of the value of profiling approaches to drug development in general; these sorts of economic analyses will prove of greater and greater value as the parameters affecting the "value" of biomarkers and profiling approaches become better understood. Chapter 17 reviews current concepts in pan-omic approaches during drug development where a compendium of data generated by multiple profiling approaches is assessed and evaluated—otherwise known, at least in part, as the holy grail of systems biology.

The last chapter provides a brief survey of findings in surrogate tissues that lie outside the covers of this textbook, summarizing important studies in this young field and looking to the future as well. It also discusses the burgeoning need for well-characterized and reproducible surrogate tissue analysis approaches as the requirement for biomarkers in the field of translational medicine is realized. One of the most exciting and simultaneously difficult characteristics of interpreting results from surrogate tissue profiling experiments today lies in the fact that there is often no precedent in the literature for the findings. Why do circulating peripheral blood mononuclear cells of renal cancer patients "look" different from those of healthy individuals at the transcriptional level? Are there clues to components of diseases that have been hitherto less explored — for instance, immunological responses of peripheral circulating cells to weakly immunogenic or nonimmunogenic solid tumors — and can this new knowledge be used to identify biomarkers of disease, but possibly to exploit mechanistically relevant pathways influencing disease progression by therapeutic intervention? These types of questions along with the constant efforts and the balance of innovative thinking with careful attention to details — both biological and technical — which are currently

being exhibited by investigators in the field of surrogate tissue analysis would seem to ensure that this area of biomedical research will enjoy continued success in the years to come.

Editors

Dr. Michael E. Burczynski earned his Ph.D. in pharmacology from the University of Pennsylvania and is currently the head of Pharmacogenomics in the Biomarkers Laboratory at Wyeth Research in Collegeville, Pennsylvania. He is a member of the American Association of Cancer Research and Society of Toxicology and has authored more than 50 articles and abstracts, with some of his most recent articles appearing in *Cancer Research, Clinical Cancer Research,* and *Current Molecular Medicine*. He was the editor of *An Introduction to Toxicogenomics* published by CRC Press in 2003 and is also a published fiction author. He is currently working on his latest novel, tentatively entitled *The Orchard of Perdition*.

Dr. John C. Rockett earned his Ph.D. in biological sciences from the University of Warwick, England, and is currently a research fellow in the Preclinical Molecular Profiling group at Rosetta Inpharmatics (a wholly owned subsidiary of Merck & Co., Inc.) in Seattle, Washington. He is a past research fellow at the University of Surrey, England, and a research biologist with the U.S. Environmental Protection Agency in Triangle Park, North Carolina. He is a member of the Institute of Biology and the Society of Toxicology and has published more than 70 articles and abstracts in various scientific journals, most recently in *Biology of Reproduction, Environmental Health Perspectives*, *Genomics*, *Toxicological Sciences,* and *Toxicology and Applied Pharmacology.*

Contributors

Hikmat Al-Ahmadie
Department of Pathology
Memorial Sloan-Kettering Cancer Center
New York, New York

Satyajit Bhattacharya
Department of Pathology
Memorial Sloan-Kettering Cancer Center
New York, New York

Michael E. Burczynski
Molecular Profiling and Biomarker Discovery
Wyeth Research
Collegeville, Pennsylvania

Monica J. Cahilly
Green Mountain Quality Associates
Warren, Vermont

Katherine R. Calvo
FDA-NCI Clinical Proteomics Program
Laboratory of Pathology
National Cancer Institute
Bethesda, Maryland

Clary B. Clish
Beyond Genomics
Waltham, Massachusetts

Jennifer L. Colangelo
Pfizer Global Research and Development
Safety Sciences
Groton, Connecticut

Svenja Debey
Molecular Tumor Biology and Tumor Immunology
Clinic I for Internal Medicine
University of Cologne
Cologne, Germany

Stan N. Finkelstein
Program on the Phamaceutical Industry
Massachusetts Institute of Technology
Cambridge, Massachusetts

Ronald A. Ghossein
Department of Pathology
Memorial Sloan-Kettering Cancer Center
New York, New York

Donald L. Gilbert
Division of Neurology
Cincinnati Children's Hospital Medical Center
Cincinnati, Ohio

Tracy A. Glauser
Division of Neurology
Cincinnati Children's Hospital Medical Center
Cincinnati, Ohio

Julian L. Griffin
Department of Biochemistry
University of Cambridge
Cambridge, United Kingdom

S.M. Gupta
Immunology Division
National Institute for Research in Reproductive Health
Indian Council of Medical Research
Parel, Mumbai, India

Andrew D. Hershey
Division of Neurology
Cincinnati Children's Hospital Medical Center
Cincinnati, Ohio

Stephen A. Krawetz
Department of Obstetrics and Gynecology
Center for Molecular Medicine and Genetics
Institute for Scientific Computing
Wayne State University
Detroit, Michigan

Michael P. Lawton
Pfizer Global Research and Development
Safety Sciences
Groton, Connecticut

Lance A. Liotta
FDA-NCI Clinical Proteomics Program
Office of Cell Therapy and Gene Therapy
Food and Drug Administration
Bethesda, Maryland

Aigang Lu
MIND Institute
University of California at Davis
Sacramento, California

P. K. Meherji
Immunology Division
National Institute for Research in Reproductive Health
Indian Council of Medical Research
Parel, Mumbai, India

Deborah P. Mounts
Bioinformatics Systems Development
Wyeth Research
Cambridge, Massachusetts

Judith L. Oestreicher
Molecular Profiling and Biomarker Discovery
Wyeth Research
Cambridge, Massachusetts

G. Charles Ostermeier
Department of Obstetrics and Gynecology
Center for Molecular Medicine and Genetics
Wayne State University
Detroit, Michigan

William D. Pennie
Pfizer Global Research and Development
Safety Sciences
Groton, Connecticut

Emanuel F. Petricoin III
FDA-NCI Clinical Proteomics Program
Office of Cell Therapy and Gene Therapy
Food and Drug Administration
Bethesda, Maryland

Raji Pillai
Genomics Collaborations
Affymetrix, Inc.
Santa Clara, California

Ruiqiong Ran
MIND Institute
University of California at Davis
Sacramento, California

K.V.R. Reddy
Immunology Division
National Institute for Research in Reproductive Health
Indian Council of Medical Research
Parel, Mumbai, India

Shawn Ritchie
Phenomenome Discoveries
Saskatoon, Saskatchewan, Canada

John C. Rockett
Preclinical Molecular Profiling
Rosetta Inpharmatics
Seattle, Washington

Edward Sauter
University of Missouri
Columbia, Missouri

Joachim L. Schultze
Molecular Tumor Biology and Tumor Immunology
Clinic I for Internal Medicine
University of Cologne
Cologne, Germany

Charles N. Serhan
Center for Experimental Therapeutics and Reperfusion Injury
Department of Anesthesiology, Perioperative and Pain Medicine
Brigham and Women's Hospital and Harvard Medical School
Boston, Massachusetts

Frank R. Sharp
MIND Institute
University of California at Davis
Sacramento, California

Anthony J. Sinskey
Department of Biology, Health Sciences and Technology and Program on the Pharmaceutical Industry
Massachusetts Institute of Technology
Cambridge, Massachusetts

Lisa A. Speicher
Translational Research
Wyeth Research
Collegeville, Pennsylvania

Sarah C. Stallings
Massachusetts Institute of Technology
Cambridge, Massachusetts

Yang Tang
MIND Institute
University of California at Davis
Sacramento, California

William L. Trepicchio
Clinical Pharmacogenomics
Millennium Pharmaceuticals
Cambridge, Massachusetts

Nigel J. Waters
DMPK Research
AstraZeneca Research and Development
Charnwood, Loughborough, United Kingdom

Maryann Z. Whitley
Expression Profiling Informatics
Wyeth Research
Cambridge, Massachusetts

Ivy H.N. Wong
Department of Obstetrics and Gynaecology
The Chinese University of Hong Kong
Hong Kong, China

Julia Wulfkuhle
FDA-NCI Clinical Proteomics Program
Laboratory of Pathology
National Cancer Institute
Bethesda, Maryland

Huichun Xu
MIND Institute
University of California at Davis
Sacramento, California

Thomas Zander
Molecular Tumor Biology and Tumor Immunology
Clinic I for Internal Medicine
University of Cologne
Cologne, Germany

Acknowledgments

The editors specifically thank the authors of the individual chapters in this textbook for their excellence in science and their dedication to the project.

I first thank Dr. Rockett for his initial suggestion of a book centered on this exciting topic to which we've both grown attached. I thank the members of my own laboratory, past and present, who have contributed to the ongoing field of surrogate tissue analysis — specifically Judy Oestreicher, Jennifer Stover, Natalie Twine, Krystyna Zuberek, and Christine Reilly — all of whom have helped make tremendous strides in understanding the power of surrogate tissue profiling in human disease. I also thank my many collaborators and colleagues at Wyeth Research in the departments of Biological Technologies, Translational Research, Translational Development, and Clinical Research and Development who are too numerous to mention but have made tremendous strides in this field possible. Specifically, I must thank Dr. Andy Dorner, Dr. Ron Salerno, and Dr. John Ryan for their unwavering commitment to pharmacogenomics and their guidance and support of Wyeth's endeavors in the field of surrogate tissue analysis. Most importantly I thank the patients, whose selfless contributions of samples to increase our understanding of disease make the clinically oriented investigations into the field of surrogate tissue analysis possible in the first place. — M.E.B.

I first of all thank Dr. Burczynski for initiating this project and showing me the ropes on my first foray into book editing. I also owe a debt of gratitude to the many dedicated, knowledgeable, and able colleagues, past and present, who have contributed practically and mentally to my scientific development and experience; in particular I thank David Dix, Sally Darney, and Bob Kavlock, whose support and encouragement were instrumental in initiating and advancing my interest and research in surrogate tissue analysis. — J.C.R.

Contents

Section I

Introduction to Surrogate Tissue Analysis

CHAPTER 1

Introduction to Surrogate Tissue Analysis

John C. Rockett and Michael E. Burczynski

CONTENTS

1.1 INTRODUCTION

Postgenomic technologies, including those used to analyze genomic, transcriptomic, proteomic, metabonomic, and other "omic" targets, have made it possible to define molecular physiology in exquisite detail, when tissues are accessible for sampling. However, many target tissues are not accessible for human experimental or epidemiological studies, or clinical evaluations, creating the need for surrogates that afford insight into exposures and effects in such tissues. A "surrogate" can be

defined simply as "one that takes the place of another." In surrogate tissue analysis (STA), one tissue takes the place of another. More specifically, an accessible tissue takes the place of an inaccessible target tissue. For example, one might examine a patient's peripheral blood lymphocytes (PBLs) to determine whether that person has suboptimal endometrial receptivity (Chapter 8), has suffered from neurological damage (Chapter 3), has developed a nonlymphatic neoplasm (Chapter 4), or has been exposed to a toxicant (Chapter 5). An alternative STA paradigm is to measure or analyze parts or products of a target tissue that originate from the target, but are collected or measured distal to it, in the surrogate tissue. For example, it is possible to isolate and analyze sperm from semen and use the data to help understand molecular events occurring in the testis (Chapter 6). In a similar manner, peripheral blood can be a source of circulating tumor cells that have detached or have been shed from their parent neoplasm. These can be isolated and used as a source of information about the original neoplasm (Chapters 13 and 14). In other cases soluble proteins, metabolites, or lipids are secreted or excreted from target tissues, and these can be detected and measured in fluids such as blood (Chapters 7, 11, and 12), cerebrospinal fluid (Chapter 11), nipple aspirate (Chapter 9), seminal fluid, milk, saliva, and urine. Drugs, drug metabolites, and toxicants can also be detected in such fluids (Chapters 10 and 11).

Surrogate "tissue" is a convenient, though perhaps misleading term. Where "tissue" is specified, the term is in fact used broadly to refer to any biologically derived material (biospecimen) used to report on events in a specific target tissue. Indeed, the majority of samples that offer potential application in STA are usually not considered tissues according to traditional definitions. The majority of surrogate tissues (Table 1.1) consist of either body fluids (e.g., urine, milk, tears, saliva, blood, and semen), or populations of cells extracted from body fluids (e.g., epithelial cells from urine, milk, or tears; lymphocytes from blood; sperm from semen), while some (e.g., hair follicle, hair, nail) are neither tissue nor free cells.

1.2 AREAS THAT COULD BENEFIT FROM SURROGATE TISSUE ANALYSIS

STA is not a new concept. Indeed, evidence that accessible tissues can be used to monitor events in an inaccessible tissue has been around for many years. For example, Nesnow et al. (1993) showed that the DNA adduct formation, a potential method of measuring exposure to environmental genotoxicants, exhibited a similar pattern in rat PBLs, lung, and liver following exposure to polycyclic hydrocarbons, and that this was detectable at least 56 days after treatment.

The development of "omic" technologies has led many researchers to look again, or more closely, or anew at the utility and application of STA, since such technologies have broadened both the range of tissues that can be examined and the number of targets that can be analyzed in a single experiment. In particular, there is widespread interest in how STA might be developed into a new paradigm for monitoring human health. The potential benefits include:

Table 1.1 Accessible "Tissues" That Can Potentially Be Used as Surrogate Tissues

Surrogate Tissue	Targets for Analysis	Potential Source
Blood	Cells, DNA, RNA, protein, drug metabolites, heavy metals	All
Breath condensate	Proteins, metabolites	All
Bronchial lavage	Cells, DNA, RNA, protein	All
Buccal cells	Cells, DNA, RNA, protein	All
Cord blood	Cells, DNA, RNA, protein	Postpartum females
Colostrum	DNA, RNA, protein	Postpartum females
Cerebrospinal fluid	Protein	All
Cerumen (earwax)	Protein	All
Hair shaft	DNA, protein, heavy metals, drug metabolites	All
Hair follicle	Cells, DNA, RNA, protein	All
Meconium	DNA, RNA, protein	Newborn infants
Milk	Cells, DNA, RNA, protein	Postpartum females
Nail	DNA, protein, heavy metals, drug metabolites	All
Nasal lavage	DNA, RNA, protein	All
Nipple aspirate	Cells, DNA, RNA, protein	All
Placenta	Cells, DNA, RNA, protein	Postpartum females
Saliva	DNA, RNA, protein	All
Semen	Cells, DNA, RNA, protein	Adult males
Skin	Cells, DNA, RNA, protein	All
Sputum	Cells, DNA, RNA, protein	All
Stool	DNA, RNA, protein	All
Sweat	Protein	All
Tear duct secretions	DNA, RNA, protein	All
Endocervical epithelium	DNA, RNA, protein	Adult females
Vaginal epithelium	DNA, RNA, protein	Adult females
Urine	DNA, RNA, protein, drug metabolites, heavy metals	All

Source: Adapted from Rockett, 2002.

1. The ability to monitor for and measure toxicant exposure without foreknowledge of the type of exposure
2. The ability to monitor clinically healthy internal organs at the molecular level without directly sampling those organs
3. The ability to identify possible pathological events at the preclinical stage and therefore administer preventative action
4. If disease is already apparent, an ability to identify the specific type and stage without invasive biopsy
5. The ability to determine which drug regimens offer the best chance of success in treating a specific disease
6. The ability to determine if a drug is working according to its proposed mechanism of action

These benefits fall into three broad areas: monitoring toxicant exposure and effect, monitoring disease development and progression, and drug efficacy testing.

1.2.1 Monitoring Toxicant Exposure and Effect

Toxicogenomics is a postgenomic approach to toxicology that uses primarily genomic techniques to elucidate mechanisms of toxicant action by studying the genome-wide effects of xenobiotics. One of the primary tenets of toxicogenomics is that the effects of toxicants on cellular functions are mediated through gene expression changes, or at least cause gene changes to occur as secondary effects. In most cases these gene changes occur prior to clinical manifestation of toxicity, which provides a window of opportunity for preclinical diagnosis of possible toxic end points that may arise as a result of the exposure. Such a diagnosis would employ the use of gene expression profiling (GEP), either on a global or restricted scale. GEP offers the potential to classify toxicant exposures (Burczynski et al., 2000; Bartosiewicz et al., 2001; Thomas et al., 2001; Hamadeh et al., 2002a, 2002b), predict clinical outcome of such exposures (Waring et al., 2001a; Hamadeh et al., 2002c), and provide mechanistic data useful for risk assessments (Waring et al., 2001b). Recent studies have also demonstrated that early gene expression changes can predict a pathological outcome days in advance of its occurrence (Kier et al., 2004). Consequently, GEP may eventually provide a vehicle for developing exposure, diagnostic, and prognostic tests for at-risk populations or individuals.

However, using GEP to monitor for toxicant exposure and/or effect in an inaccessible tissue is a difficult prospect, since direct biopsy of such tissue is not feasible unless strong medical reason (usually indicated by clinical symptoms) dictates otherwise. A less invasive method must therefore be developed if monitoring programs are to be developed based on this toxicogenomic approach. One possible solution is the use of STA. It has been proposed that gene expression changes in accessible (surrogate) tissues (e.g., nucleated blood cells) often reflect those in inaccessible (target) tissues, thus offering a convenient biomonitoring method to provide insight into the effects of environmental toxicants on target tissues (Rockett, 2002). This subject is discussed in more detail in Chapter 5.

1.2.2 Monitoring Disease Development and Progression

One of the most intriguing concepts to have recently evolved in the field of clinical pharmacogenomics is the possibility that surrogate tissues (often the circulating cells of the peripheral blood) may contain transcriptional profiles that correlate with disease, disease status, or other clinical measures of outcome in human patients. Currently in the field of oncology it is unknown whether, in the context of solid tumor burden, such “analogous” transcriptional profiles in surrogate tissues exist. While alterations in transcriptional profiles of PBMCs of patients with cancer may not share identity with those observed in the primary tumor, such patterns would nonetheless be of tremendous physiological relevance and bear obvious diagnostic value in the assessment of this disease.

1.2.3 Drug Efficacy Testing

STA has also been used in clinical pharmacology, whereby pharmacodynamic assays are being developed for the measurement of drug action in tumor and surro-

gate tissue. The need to demonstrate that a drug is working according to its proposed mechanism is of paramount importance. Researchers at places such as the CRC in London (http://www.icr.ac.uk) are trying to determine whether such studies may be able to utilize PBLs as surrogate tissue by comparing gene expression changes in PBLs with those in cancer biopsies following administration of test drugs (http://www.icr.ac.uk/cctherap/clinical.htm). Gene expression profiling of blood has also been used to differentiate patients who respond to a drug treatment from those who do not, thus providing a mechanism for the early determination of drug efficacy. In this way, should a certain disease prove refractory to a prescribed drug, the lack of efficacy of that drug can be determined at an earlier stage than would otherwise be the case. This increases the chance of patient survival since an alternative drug regimen or treatment method can be given at an earlier stage. Examples of these and related uses of surrogate tissues in clinical pharmacology are found throughout the present text.

1.3 CHALLENGES TO THE USE OF SURROGATE TISSUES

Although there have been some promising studies in the area of STA, like all new methods and approaches there are likely to be a number of challenges to overcome before it can be determined where and when STA is both applicable and appropriate. Some challenges that have been identified so far include specimen collection, specimen availability, specimen contamination, specimen homogeneity, specimen suitability, specimen specificity, and data interpretation.

1.3.1 Specimen Collection

The biological specimens that might be used in human STA are listed in Table 1.1. With such a varied selection of samples available, one of the first challenges is to develop appropriate methods for collection, storage, and transportation of tissues at and between sites of collection and analysis. “Appropriate” means that:

1. Sufficient specimen must be collected to enable extraction of reasonable amounts of good quality target material.
2. The collection, transportation, and storage procedures must not permit degradation of the target biomolecules. For example, RNA (used for gene expression analysis) is notoriously quick to degrade in *ex vivo* samples and must be protected in such a way as to inhibit the activity of RNAses. Chapter 2 discusses this issue in depth from the perspective of blood collection for genomic analysis.
3. To obtain an accurate profile from a subject or experimental animal at the time of specimen collection, the population of RNAs (the “transcriptome”) or proteins (the “proteome”) or other “ome” under investigation in a specimen must not change between collection of the specimen and extraction of the target biomolecules (RNA, protein, etc.) from the specimen in the laboratory.

Actual measurement of the level of individual biomolecules, be they members of the transcriptome, proteome, metabonome, lipidome, or other “ome,” can be

achieved in a number of ways. However, many of the newer techniques are not yet fully validated. For example, where the use of DNA arrays is concerned, many accessible tissues provide only small amounts of sample, yielding only small amounts of RNA. To overcome this, protocols have been developed that incorporate RNA amplification steps prior to labeling and hybridization of the sample. Fink et al. (2002) used this approach successfully in carrying out microarray analysis of RNA extracted from laser capture microdissection samples. However, the reliability of array data from amplified RNA samples has yet to be fully determined. In addition, the accuracy and reliability of much of the published microarray data are still a matter of open debate, and the methods for assuring data quality are not well established (Chipping Forecast II, 2002).

1.3.2 Specimen Availability

In some cases, a potential surrogate tissue may be useful only at certain times. For example, human hair follicles exist in several different growing states, with the majority (80%) in anaphase (actively growing). These are the best for RNA extraction. In cataphase, the hair follicles are moribund, and are consequently much smaller and yield correspondingly small quantities of RNA. In other cases a potential surrogate tissue may only be available from certain populations (e.g., sperm from adult males) or at certain times (e.g., placental tissue and cord blood from postpartum females, and milk from lactating females). Another factor that might occasionally limit availability of samples is cultural, religious, or personal beliefs that prohibit the provision of certain biospecimens, most notably blood or semen.

1.3.3 Specimen Contamination

The issue of contamination must also be addressed where many surrogate tissues are concerned. This arises from the fact that since many of them are externally accessible, they may be contaminated with nonhuman biological material, including bacteria, viruses, and fungi. Stool, nail, and saliva are perhaps the best example of this.

1.3.4 Specimen Homogeneity

Many surrogate tissues are homogeneous, in that they are composed of a number of different components, including fluid (e.g., serum in blood and seminal fluid in semen) and different populations of cells (e.g., leukocytes and erythrocytes in blood and leukocytes, epithelial cells, and spermatozoa in semen). It may be necessary (depending on the cell population being sought after or the "omic" technique being used) to selectively remove or separate specific cell populations from the surrogate tissue specimen. This can be done using magnetic beads or fluorescence-activated cell sorting (FACS) if appropriate antibodies are available to cell-specific antigens, by using separation gradients, e.g., Ficoll and Percoll (Amersham Biosciences), or by using selective lysis. In the isolation of sperm from semen, for example, a wash step is included, which lyses somatic cells (epithelial and inflammatory), leaving the highly resistant sperm cells intact (see Chapter 6).

1.3.5 Specimen Suitability

Surrogate tissues vary in the types of analysis that can be carried out on them. For example, DNA can be obtained from nail and hair (Tanigawara et al., 2001), but these tissues do not yield RNA. Hair follicles, on the other hand, are a good source of RNA, and work published by Mitsui et al. (1997) indicates that as much as 900 ng of total RNA can be extracted from a single human hair follicle. Buccal cells yield both DNA and RNA. Unfortunately, since these particular cells, which are obtained by swabbing the inside cheek, are typically moribund, the RNA obtained from them is not of sufficiently good quality to use on arrays, although it has been used for reverse transcriptase-polymerase chain reaction (RT-PCR) (Smith et al., 1996).

1.3.6 Specimen Specificity

Another issue is that in some cases toxicant action can be very specific, and there may be no appropriate surrogate tissue. In other cases, certain surrogate tissues may be more useful than others depending on the target tissue being studied. For example, sperm is likely to be the best surrogate tissue for monitoring events occurring in the testis (Ostermeier et al., 2002), whereas intuitively one could reasonably hypothesize that PBLs are probably most useful as surrogates for thymus, spleen, tonsils, bone marrow, or glandular tissues. Indeed, when Ember et al. (2000) compared Ha-ras and p53 expression in PBLs with several target tissues (lung, liver, lymph nodes, kidneys, spleen) following exposure to a carcinogenic agent, similar expression patterns were found only in PBLs and spleen. Thus, some appropriate matching of targets and surrogate tissues is called for. Of course, there may have been many other genes that did correlate in these studies but were not analyzed. Therefore, the ability to monitor expression of many thousands of genes or proteins in one experiment, as permitted by DNA or protein arrays, makes the application of such technology to STA highly desirable.

1.3.7 Data Interpretation

Perhaps the greatest challenge of all will be the interpretation and appropriate utilization of all the “omic” and other data obtained from target and surrogate tissues. Validating the relationship between gene expression or protein profiles and toxicant exposure or disease state has already begun. If and when these relationships have been fully verified in target tissues, then the relationship between gene or protein expression in target and surrogate tissues must be established. In doing so, it will be necessary to determine whether genetic or proteomic biomarkers of toxicity or disease in target tissues are reflected in the surrogate tissue across a range of doses, time points, and disease states. Alternatively, omic biomarkers in surrogate tissues may be of high clinical value but fail to share identity with markers in the primary tissue. For example, one such scenario might involve the transcriptional response of circulating peripheral blood leukocytes due to tumor regression induced by successful chemotherapy, in which the transcriptional responses of PBMCs accurately “predict” beneficial tumor response.

One of the best hopes for successful utilization of STA lies in identifying unique biomarkers (e.g., changes in expression of a single gene/protein or a small number of such genes/proteins) of exposure and effect that show concordant modulation in surrogate and target tissues following toxicant exposure. What is ideally needed to utilize such biomarkers is one or more large relational databases through which newly generated data can be compared against previously documented toxicant exposures and effects. This would facilitate diagnosis of the type and likely outcome of any particular exposure. Of course, gene and/or protein expression levels alone may be insufficient to make an accurate diagnosis or prognosis. Other factors, such as the presence of polymorphisms in drug metabolizing and detoxifying enzymes, may need to be incorporated to improve the reliability and accuracy of this approach. Until such a time, perhaps a decade or more away, when such databases are available, it will in most cases be an enormous challenge to interpret the biological meaning and significance of the data.

1.4 SUMMARY

Surrogate tissue analysis is currently a relatively small but rapidly growing area of research. The ability to investigate biological mechanisms and obtain diagnostic and prognostic information about an inaccessible target tissue by using accessible surrogate tissues and fluids has significant and far reaching implications for health care as well as basic and clinical research. Initial proof-of-principal experiments in humans and animal models, many of which are described in this text, have provided encouraging results that suggest that STA can be applied in a large number of different scenarios. Whether STA becomes an integral component of future human health monitoring programs, a tool of limited situation-specific use, or a dead end idea, will be determined only after further studies have been conducted. However, the future of STA appears to be linked quite closely with the advancement of omic technologies, and given the large and widespread investment in these, further advances in the utility and application of STA seem quite likely.

REFERENCES

Bartosiewicz, M., Penn, S., and Buckpitt, A., 2001. Applications of gene arrays in environmental toxicology: fingerprints of gene regulation associated with cadmium chloride, benzo(*a*)pyrene, and trichloroethylene. *Environ. Health Perspect.* 109, 71–74.

Burczynski, M.E., McMillian, M., Ciervo, J., Li, L., Parker, J.B., Dunn, R.T., II, Hicken, S., Farr, S., and Johnson, M.D., 2000. Toxicogenomics-based discrimination of toxic mechanism in HepG2 human hepatoma cells. *Toxicol. Sci.* 58(2), 399–415.

Chipping Forecast II, 2002. *Nat. Gen. Suppl.,* Vol. 32, December.

Ember, I., Kiss, I., Gyongyi, Z., and Varga, C.S., 2000. Comparison of early onco/suppressor gene expressions in peripheral leukocytes and potential target organs of rats exposed to the carcinogen 1-nitropyrene. *Eur. J. Cancer Prev.* 9, 439–442.