The Handbook of Biomarkers

The Handbook of Biomarkers

Kewal K. Jain MD, FRACS, FFPM

Jain PharmaBiotech, Basel, Switzerland

Springer

Kewal K. Jain
Jain PharmaBiotech
Blaesiring 7
4057 Basel
Switzerland
jain@pharmabiotech.ch

ISBN 978-1-60761-684-9 e-ISBN 978-1-60761-685-6
DOI 10.1007/978-1-60761-685-6
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2010920089

Printed on acid-free paper

Humana Press is part of Springer Science+Business Media (www.springer.com)

Preface

This book is an overview of the state of the art of biomarkers. A biomarker is a characteristic that can be objectively measured and evaluated as an indicator of a physiological as well as a pathological process or response to a therapeutic intervention. Although there is nothing new about biomarkers such as glucose for diabetes and blood pressure for hypertension, the current focus is on molecular biomarkers, which have taken the center stage in the development of molecular medicine. Molecular diagnostic technologies have enabled the discovery of molecular biomarkers and are helping in the definition of their role in the pathomechanism of disease. Biomarkers form the basis of development of diagnostic assays as well as targets for drug discovery. Effect of drugs, in clinical trials as well as in practice, can be monitored by biomarker assays.

There is a tremendous amount of literature on biomarkers, but there is no comprehensive source of information on the topic. Of the thousands of biomarkers that are being discovered, relatively few are being validated for further applications, and it is difficult to evaluate the potential of a biomarker. This book describes different types of biomarkers and their discovery using various "-omics" technologies such as proteomics and metabolomics along with the background information for evaluations of biomarkers as well as the procedures for their validation and use in clinical trials. Biomarkers are first described according to technologies and then according to various diseases. An important feature is the correlation between diseases and classifications of biomarkers, which provides the reader with a guide to sort out current and future biomarkers.

This book would be an important source of information on biomarkers for scientists as well as physicians and those involved in drug discovery and development. Many of the regulatory issues concerning biomarkers are related to proteomics, molecular diagnostics, and pharmacogenomics/pharmacogenetics. By facilitating the combination of therapeutics with diagnostics, biomarkers will play an important role in the development of personalized medicine, which is an important emerging trend in health care.

Kewal K. Jain, MD
Basel, Switzerland

Contents

List of Abbreviations

2D GE	2-dimensional gel electrophoresis
AD	Alzheimer disease
BNP	B-type natriuretic peptide
CHF	congestive heart failure
CNS	central nervous system
CRADA	cooperative research and development agreement (between a US federal laboratory and one or more non-federal parties)
CRP	C-reactive protein
CSF	cerebrospinal fluid
CT	computer tomography
EGFR	epithelial growth factor receptor
ELISA	enzyme-linked immunosorbent assay
EST	expressed sequence tags
FDA	Food and Drug Administration, USA
FISH	fluorescent in situ hybridization
FMRI	functional magnetic resonance imaging
GC	gas chromatography
GFAP	glial fibrillary acidic protein
GWAS	genome-wide association study
Hs-CRP	high-sensitivity C-reactive protein
IHC	immunohistochemistry
IL	interleukin
LC	liquid chromatography
LCM	laser capture microdissection
LDH	lactic dehydrogenase
Lp-PLA2	lipoprotein-associated phospholipase A2
MALDI	matrix-assisted laser desorption/ionization
MALDI-MS	matrix-assisted laser desorption mass spectrometry
MCP-1	monocyte chemoattractant protein-1
miRNA	microRNA
MRI	magnetic resonance imaging
MS	mass spectrometry
mtDNA	mitochondrial DNA

NCI	National Cancer Institute
NIH	National Institutes of Health, USA
NMR	nuclear magnetic resonance
NO	nitric oxide
PCR	polymerase chain reaction
PET	positron emission tomography
PKC	protein kinase C
POC	point of care
PPAR	peroxisome proliferator-activator receptor
PSA	prostate-specific antigen
PSMA	prostate-specific membrane antigen
RCAT	Rolling circle amplification technology
RNAi	RNA interference
RT-PCR	real-time PCR
SELDI-TOF	surface-enhanced laser desorption and ionization-time of flight
sICAM-1	soluble intercellular adhesion molecule-1
SNP	single nucleotide polymorphisms
SPR	surface plasma resonance
USPTO	United States Patent & Trademark Office

Chapter 1
Introduction

Definitions

There are several definitions of biomarkers. A biomarker is a characteristic that can be objectively measured and evaluated as an indicator of a physiological as well as a pathological process or pharmacological response to a therapeutic intervention. Classical biomarkers are measurable alterations in blood pressure, blood lactate levels following exercise, and blood glucose in diabetes mellitus. Any specific molecular alteration of a cell on DNA, RNA, metabolite, or protein level can be referred to as a molecular biomarker. In the era of molecular biology, biomarkers usually mean molecular biomarkers and can be divided into three broad categories:

1. Those that track disease progression over time and correlate with known clinical measures.
2. Those that detect the effect of a drug.
3. Those that serve as surrogate end points in clinical trials.

While researchers are studying all three categories, biotechnology and pharmaceutical companies favor using biomarkers as drug discovery tools – not only to detect biological responses to experimental drugs but also to aid in the discovery of new targets for therapeutic intervention. A biomarker can be as simple as a laboratory test or as complex as a pattern of genes or proteins. From a practical point of view, the biomarker would specifically and sensitively reflect a disease state and could be used for diagnosis as well as for disease monitoring during and following therapy. The term "negative biomarker" is used for a marker that is deficient or absent in a disease.

Surrogate end point is a biomarker that is intended to serve as a substitute for a clinically meaningful end point and is expected to predict the effect of a therapeutic intervention. A clinical end point is a clinically meaningful measure of how a patient feels, functions, or survives. Clinical end points may be further classified as intermediate end points, which are clinical end points that are not the ultimate outcome, but are nonetheless of real clinical usefulness, e.g., exacerbation rate, and ultimate clinical outcomes, which are clinical end points reflective of the accumulation of irreversible morbidity and survival. These definitions indicate

K.K. Jain, *The Handbook of Biomarkers*, DOI 10.1007/978-1-60761-685-6_1,

a clear hierarchical distinction between biomarkers and surrogate end points. While numerous laboratory biomarkers may be associated with a particular disease state, the term "surrogate" indicates the ability of a biomarker to provide information about the clinical prognosis or efficacy of a therapy. The word "surrogate" implies a strong correlation with a clinical end point, but in order to be clinically useful a surrogate must provide information about prognosis or therapeutic efficacy in a significantly shorter time than would be needed by following the clinical end point.

Historically, successful surrogates have linked effects on biomarkers for single effects in large populations but this framework needs to be expanded because it does not recognize multidimensional quality of clinical response and thus conflicts with current goals for individualized therapy. There is also the need to include possibility that multiple biomarkers may provide useful information in aggregate. A biomarker is valid if:

1. It can be measured in a test system with well-established performance characteristics.
2. Evidence for its clinical significance has been established.

Historical Aspects of Biomarkers

Historical landmarks in discovery and development of biomarkers are shown in Table 1.1.

Table 1.1 Historical landmarks in discovery and development of biomarkers

Year	Landmark
1847	The first laboratory test for a protein cancer biomarker, the Bence Jones protein in urine
1954	Test for the measurement of transaminases in myocardial infarction (Karmen et al. 1954)
1960s	The term "biomarker" started to appear in the literature in connection with metabolites and biochemical abnormalities associated with several diseases
1967	An improved test for myocardial infarction based on a biomarker – serum creatine phosphokinase (Rosalki 1967)
1971	Report of carcinoembryonic antigen (CEA) as biomarker of cancer (Moore et al. 1971)
1987	Troponin I as a biomarker for myocardial infarction (Cummins et al. 1987)
Early 1990s	Accelerator mass spectrometry used for analysis of biological samples for biomarkers
1995	Applications of proteomics for discovery of biomarkers and use in molecular diagnostics
1999	Emergence of metabolomics for study of biomarkers
2000	Sequencing of the human genome completed opening the way for discovery of gene biomarkers
2005	Discovery and application of biomarkers becomes a major activity in biotechnology and biopharmaceutical industries

Classification of Biomarkers

A classification of biomarkers is shown in Table 1.2.

Table 1.2 Classification of biomarkers

Disease biomarkers: type 0 biomarkers
Clues to pathomechanism of a disease
Diagnostic biomarkers: early detection of disease
Tracking disease progression over time
Prognostic biomarker for prognosis or outcome of disease
Diagnostic biomarkers
Molecular diagnostics, e.g., CA-125 for ovarian cancer
Biomarkers as links between diagnostics and therapeutics
Pattern diagnosis, e.g., serum protein biomarker pattern diagnosis of ovarian cancer
Biomarkers for drug discovery
Target biomarker: reports interaction of the drug with its target
Disease biomarkers as targets for drug discovery
Predictive biomarkers
Biomarker associated with a risk for disease as a candidate for a screening test
To predict disease at presymptomatic stage: autoantibodies
To predict the effect of a drug on disease
To predict the toxicity of a drug
Biomarkers to detect drug effects: type I biomarkers
Efficacy biomarker: indicator of beneficial effect of a drug
Mechanism biomarker: reports a downstream effect of a drug
Toxicity biomarker: reports toxicological effect of a drug in an in vitro or an in vivo system
Translation biomarker
A biomarker that can be applied in both a preclinical and a clinical setting
Biomarkers as surrogate end points in clinical trials: type II biomarkers
As a substitute measure for clinical outcome, e.g., cholesterol levels in statin therapy
In vivo imaging as end point: MRI of multiple sclerosis lesions in interferon therapy
Valid biomarkers: validated in clinical trials

Biological Marker as Response to Therapeutic Intervention

A biological marker can be a pharmacologic response to therapeutic intervention (Biomarkers Definitions Working Group 2001). A pharmacogenetic test is an assay intended to study interindividual variations in DNA sequence related to drug absorption and disposition (pharmacokinetics) or drug action (pharmacodynamics), including polymorphic variation in the genes that encode the functions of transporters, metabolizing enzymes, receptors, and other proteins.

A pharmacogenomic test is an assay intended to study interindividual variations in whole-genome or candidate gene, SNPs, haplotype markers, or alterations in gene expression or inactivation that may be correlated with pharmacological function and therapeutic response. In some cases, the pattern or profile of change is the relevant biomarker, rather than changes in individual markers.

Pharmacokinetic/Pharmacodynamics Biomarkers

Mechanism-based pharmacokinetic/pharmacodynamic (PK/PD) models differ from empirical descriptive models in that they contain specific expressions to characterize processes on the causal path between drug administration and effect. Mechanism-based PK/PD models have much improved properties for extrapolation and prediction. As such, they constitute a scientific basis for rational drug discovery and development. Within the context of mechanism-based PK/PD modeling, a biomarker is defined as a measure that characterizes, in a strictly quantitative manner, a process, which is on the causal path between drug administration and effect. The new classification system of biomarkers distinguishes seven types of biomarkers (Danhof et al. 2005):

1. Genotype/phenotype determining drug response.
2. Concentration of drug or drug metabolite.
3. Molecular target occupancy.
4. Molecular target activation.
5. Physiological measures.
6. Pathophysiological measures.
7. Clinical ratings.

Predictive Biomarkers

Biomarkers may be used to predict the efficacy or toxicity of a drug. Finding reliable biomarkers that are indicators of a certain response is difficult. So when looking for biomarkers that can predict a certain clinical outcome the task becomes even more challenging. Biomarkers for predicting toxicity, which is often dose related, are difficult. These effects are usually studied by increasing the dose of a compound until toxicity is observed. However, the predictive value of such an approach in patients is very limited. What is needed is a biomarker that will predict toxicity in a certain patient population.

In the chemoinformatics approach, chemistry-related toxicity can be predicted with the help of databases of known drugs that links phenotypic toxicity to a specific characteristic of a compound. However, other approaches are required for determining genomic-based toxicity.

Biomarkers are used in toxicogenomics as well. Toxicogenomics is based on the idea that if the environment inside a cell is altered by an external stimulus, some of the cell's genes will likely express themselves in an atypical way. The more toxic the external stimulus, the greater the number of genes that will be altered. Conversely, if the stimulus is benign, then very few genes will change. Predictive toxicogenomics, i.e., the acquisition of advanced knowledge of the safety profile of a compound using genomic biomarkers, is a technology that provides much optimism for improving early drug discovery decisions. Toxicogenomics creates an opportunity to shift attrition to earlier stages in drug development to a point

where course-corrective action can be taken with relatively lower financial costs, thus improving the efficiency of the drug development process. Toxicogenomics can be used for predicting toxicity, both in vivo and in vitro, by using classification algorithms and toxicogenomic databases for biomarker discovery and validation (Fielden and Kolaja 2006).

Valid Biomarkers

A valid biomarker is defined as a biomarker that is measured in an analytical test system with well-established performance characteristics and for which there is an established scientific framework or body of evidence that elucidates the physiologic, toxicologic, pharmacologic, or clinical significance of test results. Validation of a biomarker is context specific and the criteria for validation will vary with the intended use of the biomarker. The clinical utility (e.g., predict toxicity, effectiveness, or dosing) and use of epidemiology/population data (e.g., strength of genotype – phenotype associations) are examples of approaches that can be used to determine the necessary criteria for validation. Table 1.3 lists the terms used for disease biomarkers in clinical development, which is an expansion of type 0 biomarkers listed in Table 1.2. Regulatory aspects of biomarker validation will be discussed in Chapter 10 .

Table 1.3 Terminology of biomarkers of disease relevant to clinical development

Term	Application
Predisposition biomarker	To identify predisposition to a disease, e.g., genetic
Screening biomarkers	To identify those suffering from a disease
Staging biomarker	To determine the stage of progression of the disease
Prediction biomarker	To predict the course of the disease
Prognostic biomarker	To assess disease progression and outcome
Recurrence monitoring biomarkers	To identify recurrence of the disease

Types of Biomarkers

There are many ways of classifying biomarkers as reflected in the rest of this report. The biomarkers may be simple molecules such as metabolites, carbohydrates (e.g., glucose), steroids, and lipids. Less simple are peptides and proteins such as insulin, hemoglobin A and C, prostate-specific antigen, and C-reactive protein. More complex biomarkers are cells such as platelets or T cells and autoantibodies. Patients as clinical phenotypes are most complex but this topic will not be discussed in this report.

Genes as Biomarkers

A gene is a sequence of chromosomal DNA that is required for the production of a functional protein or a functional RNA molecule. Genes range in size from small (1.5 kb for globin gene) to large (approximately 2,000 kb for Duchenne muscular dystrophy gene). A gene includes not only the actual coding sequences but also adjacent nucleotide sequences required for the proper expression of genes, i.e., for the production of a normal mRNA molecule. Mature mRNA is about one-tenth the size of the gene from which it is transcribed. The same DNA strand of a gene is always translated into mRNA so that only one kind of mRNA is made for each gene. Transcription is gene in action. Genes are often described as blueprints of life and transmit inherited traits from one generation to another.

The activity of a gene, the so-called gene "expression" means that its DNA is used as a blueprint to produce a specific protein. Not all the genes are expressed in a typical human cell and those that are expressed vary from one cell to another. Patterns in which a gene is expressed provide clues to its biological role. Malfunctioning of genes is involved in most diseases, not only inherited ones. All functions of cells, tissues, and organs are controlled by differential gene expression. As an example, red blood cells contain large amounts of the hemoglobin protein that is responsible for carrying oxygen throughout the body. The abundance of hemoglobin in red blood cells reflects the fact that its encoding gene, the hemoglobin gene, is actively transcribed in the precursor cells that eventually produce red blood cells. In all other cells of the body, the hemoglobin gene is silent. Accordingly, hemoglobin is present only in red blood cells. It is now well established that differential gene expression results in the carefully controlled (or regulated) expression of functional proteins, such as hemoglobin and insulin.

Proteins as Biomarkers

Proteins are fairly large molecules, made up of strings of amino acids linked like a chain. There are 20 amino acids, and proteins range in length from a few to over a thousand amino acids. Different combinations of amino acids link to form tens of thousands of proteins. Proteins usually contain thousands of atoms precisely arranged in a 3D structure that is unique for each type.

As a protein is made, it "folds" itself into a complex, 3D shape, like a piece of ribbon that has been crumpled up. Each protein has one folded shape, and consistently folds into it, usually in less than a second. That complicated folded shape dictates how the protein works, and also how it interacts with other entities.

The specific sequence of amino acids that make up each protein is coded by a gene in the DNA of living cells. A protein cannot be synthesized without its mRNA being present, but a protein can persist in the cell when its mRNA is no longer present. However, mRNA may be present in abundance but the message is not translated into proteins. There is, thus, no good correlation between mRNA and protein in a cell at any given time. Protein synthesis is a very complicated process. Ribosomes

are the cell's protein factories. RNA bridges in the ribosomes are not just support structures but also a part of the protein forming machinery.

Peptides are small proteins that play a central role in almost all biological processes. They function as biochemical messengers (for example, insulin, calcitonin, and angiotensin) or occur as metabolites of proteins.

Proteomics

The term "proteomics" indicates PROTEins expressed by a genOME and is the systematic analysis of protein profiles of tissues. The term "proteome" refers to all proteins produced by a species, much as the genome is the entire set of genes. Unlike the genome, the proteome varies with time and is defined as "the proteins present in one sample (tissue, organism, cell culture) at a certain point in time." Proteomics parallels the related field of genomics. Now that the human genome has been sequenced, we face the greater challenge of making use of this information for improving health care and discovering new drugs. There is an increasing interest in proteomics technologies now because DNA sequence information provides only a static snapshot of the various ways in which the cell might use its proteins whereas the life of the cell is a dynamic process. In addition to proteins, peptides (low molecular weight proteins) are also biomarkers of disease in body tissues and can be detected by proteomic technologies.

DNA Biomarkers

Genetic information is contained in the cells in the form of DNA. DNA consists of two strands, which resemble a ladder coiled into a spiral shape – the double helix. It is a macromolecule composed of linear array of nucleotides, each of which comprises a base plus a pentose sugar and phosphate. Only four nucleotide bases are normally found in DNA: cytosine (C), thymine (T), adenine (A), and guanine (G). The information content of the DNA is embodied in the sequential arrangement of nucleotides. The assembly of higher order structures comprising multiple proteins bound at distinct DNA sites initiates readout of information encoded in the DNA. DNA contains the instructions for making proteins. There is a need to assess DNA damage because of the impact that different insults on genetic material may have on human health.

Mitochondrial DNA

While autosomal nuclear DNA genes are confined to the nucleus, limited to two copies per cell, the mitochondrial DNA (mtDNA) genes are distributed throughout the cytoplasm and are present in numerous copies per cell. The mtDNA molecule

is relatively small containing 16,569 nucleotide pairs. Mitochondria are descendent of a "bacterium-like" organism, which had a working relationship with our ancestral cells so that they could produce energy from glucose and oxygen and store this energy in the form of high-energy phosphate bonds of adenosine triphosphate (ATP). As a remnant of its past life, each mitochondrion contains a "private" set of genes that possess the genetic blueprint for the production of proteins and other molecules that are critical to the process of cellular energy production. mtDNA encodes for proteins that are components of the mitochondrial respiratory chain and oxidative phosphorylation system. Mitochondria have a degree of autonomy within the cell by virtue of having their own genome but it is limited because replication and transcription of mtDNA is dependent on nuclear factors such as mitochondrial transcription factor a. mtDNA differs from DNA in cell nucleus in the following important respects:

- It is strictly maternally inherited, does not recombine, and therefore accumulates mutations sequentially.
- It contains few non-coding sequences.
- It has a slightly different genetic code, for example, the uridine – guanine – adenine (UGA) codon is read as "tryptophan" rather than a "stop."

Mitochondrial Mutations

There is growing evidence that defects of mtDNA causes disease. Majority of these defects are due to point mutations or rearrangements of the mitochondrial genome, while others, such as mtDNA deletions, are autosomally linked. More than 100 mutations of mtDNA been associated with a striking variety of multisystemic as well as tissue-specific human diseases. Disorders due to mutations in genes affecting mitochondrial protein synthesis may erode the bioenergetic capacity of the tissues contributing to the senescence process in aging. In contrast to the remarkable progress in our understanding of etiology, pathogenesis is only partially explained by the rules of mitochondrial genetics and remains largely unclear.

RNA Biomarkers

Ribonucleic acid (RNA) is the other major nucleic acid besides DNA but unlike DNA, it is single stranded. It contains ribose instead of deoxyribose as its sugar–phosphate backbone and that uracil (U) instead of thymine (T) in its pyrimidine bases. Like DNA, it can be assembled from nucleotides using DNA sequence as a template and RNA polymerase. The structure of an RNA molecule is also determined by its DNA-derived sequence. If proteins are the hardware, RNA is the software controlling how the genes are expressed to make proteins. RNA is unique in being able to store and transmit information as well as process that information.

Classically RNAs can be classified into messenger RNAs (mRNAs), which are translated into proteins, and non-protein-coding RNAs (ncRNAs). mRNA is the short-lived intermediary in the transfer of genetic information from DNA to protein. mRNA is transported out of the nucleus and is translated into protein on the cytoplasmic ribosomes. Transcriptome is the complete set of mRNA molecules of a cell, tissue, or an organism. Transcription preserves the whole information content of the DNA sequence that it has been transcribed from, since the RNA has the same base-pairing characteristics.

ncRNA genes produce functional RNA molecules rather than encoding proteins and include transfer RNAs (tRNAs) and ribosomal RNAs (rRNA). rRNAs are highly structured and conserved molecules found in all living organisms and are well established as phylogenetic markers. During the last two decades several ncRNAs have emerged, having a diverse range of functions, from structural through regulatory to catalytic. A dominating category is that of small nucleolar (sno) RNAs, which act as guides to direct pseudouridylation and 2′-*O*-ribose methylation in rRNA. Other categories are microRNAs (miRNAs), antisense transcripts, and transcriptional units containing a high density of stop codons and lacking any extensive open reading frame.

Profiling of human mRNA in serum has been found to be useful for detection of oral squamous cell carcinoma (Li et al. 2006a). Human mRNAs are present in saliva and can be used as biomarkers of oral cancer. Saliva harbors both full-length and partially degraded forms of mRNA. RNA enters the oral cavity from different sources, and association with macromolecules may protect salivary RNA from degradation (Park et al. 2006). However, RNA is unstable and the degradation process is likely to start before the cells are shed from the tissue, limiting its value as a biomarker. The results of measurements of transcript levels in biopsies of oral tissue need to be interpreted with caution. To address the problem of RNA instability, RNA is immediately stabilized after the blood draw by PAXgene (PreAnalytiX). Total RNA is then extracted from PAXgene-stabilized blood and subjected to microarray analysis (Debey-Pascher et al. 2009). Combining RNA stabilization of peripheral blood with bead-based oligonucleotide microarray technology is not only applicable to small single-center studies with optimized infrastructure but also applicable to large-scale multicenter trials that are mandatory for the development of predictive biomarkers for disease and treatment outcome.

Transcriptomics

The focus of decoding genomic information for drug discovery has been mostly on proteomics and mRNA (cDNA) analysis. A limitation of this approach is that the information contained within the genome is first expressed in the form of “primary transcripts” before it is processed into mRNA and proteins. The primary transcripts may not lead to the formation of mRNA and proteins but perform crucial cellular functions directly. Transcriptomics is the study of the entire set of RNA transcripts of an organism.

A shared goal in transcript and proteomic profiling is the development of biomarkers and signatures of chemical toxicity. Toxicity profiling with DNA microarrays to measure all mRNA transcripts, or by global separation and identification of proteins, has led to the discovery of better descriptors of toxicity, toxicant classification, and exposure monitoring than current indicators. Biomarkers and signature profiles are described for specific chemical toxicants that affect target organs such as liver, kidney, neural tissues, gastrointestinal tract, and skeletal muscle, for specific disease models such as cancer and inflammation, and for unique chemical-protein adducts underlying cell injury (Merrick and Bruno 2004). The recent introduction of toxicogenomics databases support researchers in sharing, analyzing, visualizing, and mining expression data, assist the integration of transcriptomics, proteomics, and toxicology data sets, and eventually will permit *in silico* biomarker and signature pattern discovery.

MicroRNAs

MicroRNAs (miRNAs), small mostly non-coding RNA gene products, are molecules derived from larger segments of "precursor" RNA that are found in all diverse multicellular organisms. miRNAs are 21–25-nucleotide transcripts that repress gene function through interactions with target mRNAs (Moss 2002). miRNAs target the control of gene activity at multiple levels, specifically transcription, translation, and protein degradation, i.e., miRNAs act as meta-regulators of expression control. miRNA-mediated gene regulation is guided by the base-pairing rules of Watson and Crick (Chen 2005).

Each miRNA is thought to regulate multiple genes, and since approximately 1,000 miRNA genes have been identified in humans, the potential regulatory circuitry afforded by miRNA is enormous. Recent studies of miRNA expression implicate miRNAs in viral disease, neurodevelopment, and cancer. In higher eukaryotes, the role of miRNAs in regulating gene expression could be as important as that of transcription factors.

Metabolomics

The human metabolome is best understood by analogy to the human genome, i.e., where the human genome is the set of all genes in a human, the human metabolome is the set of all metabolites in a human. In a systems biology approach, metabolomics provides a functional readout of changes determined by genetic blueprint, regulation, protein abundance and modification, and environmental influence. Metabolomics is the study of the small molecules, or metabolites, contained in a human cell, tissue, or organ (including fluids) and involved in primary and intermediary metabolism. By definition, the metabolome should exclude enzymes, genetic material, and structural molecules such as glycosaminoglycans, and other polymeric