Chapter 3
Role of Biomarkers in Personalized Medicine

Introduction

A biological marker (biomarker) is simply a molecule that indicates an alteration in physiology from the normal. For example, any specific molecular alteration of a cancer cell either on DNA, RNA, or protein level can be referred to as a molecular marker. A biomarker is defined 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. The topic of biomarkers has been discussed in a special report on this topic (Jain 2009f). The expression of a distinct gene can enable its identification in a tissue with none of the surrounding cells expressing the specific marker. Relation of biomarkers to other technologies and healthcare is shown in Fig. 3.1.

Applications of biomarkers relevant to personalized medicine are:

- The biomarker would specifically and sensitively reflect a disease state and could be used for diagnosis, for predicting response to drug, and for disease monitoring during and following therapy.
- Biomarkers can be used as drug targets in drug development.
- Biomarkers might serve to integrate diagnostics and therapeutics.

Potential usefulness of biomarkers in development of personalized medicine is illustrated by the example of the discovery of biomarkers for Huntington's disease (HD). Genome-wide gene expression profiles from blood samples of HD patients have identified changes in blood mRNAs that clearly distinguish HD patients from controls (Borovecki et al. 2005). The elevated mRNAs were significantly reduced in HD patients involved in a dose-finding study of the histone deacetylase inhibitor, sodium phenylbutyrate. These alterations in mRNA expression correlate with disease progression and response to experimental treatment. Such biomarkers may provide clues to the state of HD and may be of predictive value in clinical trials.

K.K. Jain, *Textbook of Personalized Medicine*,
DOI 10.1007/978-1-4419-0769-1_3, © Springer Science+Business Media, LLC 2009

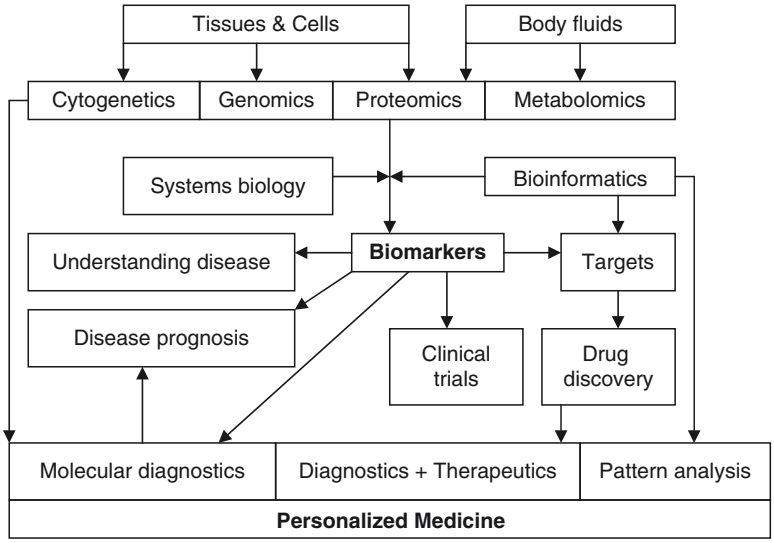

Fig. 3.1 Relation of biomarkers to other technologies and personalized medicine. ©Jain Pharma-Biotech

Technologies for Discovery of Biomarkers

Systems Biology Approach to Biomarker Identification

Ideally, a systematic approach to biomarker identification should involve multiple "-omic" technologies to investigate a disease process at all levels, including whole genome association studies to identify causative mutations or polymorphisms, as well as expression profiling, proteomics, and metabolomics to identify expression signatures and protein and small-molecule profiles that are either specific to the disease process or provide mechanistic insights into disease pathology. Uses of genomics, proteomics, and metabolomics in biomarker discovery are summarized in Table 3.1. Genomics is used to identify relevant disease genes, aberrant cellular signaling pathways and expression signatures correlated with disease. Proteomics is used to identify aberrant protein expression, post-translational modification, protein interactions, and protein profiles that are specific to a particular disorder. Finally, metabolomics is implemented to identify the presence of abnormal levels of small-molecule metabolites that are specific to and indicative of an underlying disease process.

Epigenomic Technologies

Epigenomics is one of the many 'omics' that have developed in the wake of the Human Genome Project. DNA methylation sites throughout the human genome

Table 3.1 Use of "-omic" technologies for discovery of biomarkers

Level of analysis	Tissue source	Technologies	Application
Genomics	Nucleated cells	Positional cloning	Mapping of disease loci
	Nucleated cells	SNP genotyping	Identification of disease gene
	Nucleated cells	Microsatellites	Mapping of disease loci
	Pathologically affected cells	Expression arrays	Identification of dysregulated genes
	Pathologically affected cells	Comparative genomic hybridization arrays	Detection of gene amplification and loss of heterozygosity
Proteomics	Affected tissues Body fluids: urine, blood, saliva	2D gel electrophoresis Liquid chromatography-mass spectrometry (MS) ICAT-MS	Identification of protein biomarkers
Metabonomics	Body fluids: urine, blood, saliva	Nuclear magnetic resonance (NMR) MS	Identification of small molecules
Glycomics	Body fluids: urine, blood, saliva	NMR Oligosaccharide arrays	Identification of carbohydrates Identification of glycoproteins

were mapped during Human Epigenome Project (HEP). The Human Genome Project provides the blueprint for life, but the epigenome tells us how this whole thing is executed, what determines when and where genes are switched on and off to produce a person. And knowing more about the human epigenome may provide clues to what goes wrong in cancer and other diseases. The latest information on this can be obtained at the HEP web site: http://www.epigenome.org/. As a prelude to the full-scale HEP, a pilot study of the methylation patterns within the Major Histocompatibility Complex (MHC) has been completed. This region of chromosome 6 is associated with more diseases than any other region in the human genome. Methylation variable positions (MVPs) were identified in the vicinity of the promoter and other relevant regions of approximately 150 loci within the MHC in tissues from a range of individuals. This provides an unprecedented insight into the complex relationship between genetics and epigenetics that underlies both normal cellular homeostasis and disease states, in particular autoimmune diseases. For the pilot project, an integrated genomics-based technology platform was developed. The pipeline involves the automated bisulphite treatment of DNA from minute tissue biopsies, gene-specific bisulphite PCR and large-scale sequencing of PCR amplicons. Analysis and quantification of methylation patterns is achieved by mass spectrometric and microarray assays.

Discovery of Methylation Biomarkers

Methylation is the only flexible genomic parameter that can change genome function under exogenous influence. Hence it constitutes the main and so far the missing link between genetics, disease, and the environment that is widely thought to play a decisive role in the etiology of virtually all human diseases. Methylation occurs naturally on cytosine bases at CpG sequences and is involved in controlling the correct expression of genes. Differentially methylated cytosines give rise to distinct patterns specific for tissue type and disease state. Such MVPs are common epigenetic markers. SNPs promise to significantly advance our ability to understand and diagnose human disease. DNA methylation is an important cellular mechanism modulating gene expression associated with aging, inflammation, and atherosclerotic processes. Global DNA hypermethylation is associated with inflammation and increased mortality in cardiovascular disease (Stenvinkel et al. 2007).

In the last few years, DNA methylation has become one of the most studied gene regulation mechanisms in carcinogenesis. Advances in the technologies that enable detection of DNA methylation in a variety of analytes have opened the possibility of developing methylation-based tests. A number of studies have provided evidence that specific methylation changes can alter the response to different therapeutic agents in cancer and, therefore, be useful biomarkers.

Proteomic Strategies for Biomarker Identification

Proteomics approach has been used to identify novel biomarkers. Although two-dimensional (2D) gel electrophoresis is used widely, ProteinChip has a greater potential for identification of biomarkers. Antibody arrays can be used for screening. Proteomic approaches for biomarker discovery have been used in many diseases. A 2D approach has been used for tumor marker identification in a number of cancers. Laser capture microdissection has been used in conjunction with ProteinChip to study protein expression profiles in cancer. The advantage of ProteinChip over 2D gel electrophoresis is that the chip platform used to identify the biomarker can also be used to develop a high-throughput assay.

Proteomics is a key technology for the discovery of biomarkers for pharmaceutical and diagnostic research. Although gene expression provides the level of proteins that is the key to the effect of the gene, it can be due to other factors in addition to the concentration of mRNA that codes for it. These factors include protein post-translational modifications, turnover, transport, and excretion. Therefore quantitative proteomics is essential for monitoring different pathways in blood samples of patients. Such biomarkers help in differential diagnosis as well as provide an understanding of pathomechanism of the disease and assessment of response to treatment. Non-invasive measurement (e.g., in serum) is the key feature of a biomarker that can be identified in diseased tissue. Multidimensional protein fractionation schemes are used to achieve appropriate sensitivity.

Proteomic Technologies for Detection of Biomarkers in Body Fluids

The first decision to be made in the search for a biomarker is whether to look in a body fluid or a tissue. Body fluids have the advantage of being more easily accessible and are more likely to be of clinical use because serum or urine can be obtained by non-invasive methods as a routine. Identification of rare proteins in blood is often hindered by highly abundant proteins, such as albumin and immunoglobulin, which obscure less plentiful molecules. A solution to this problem is an immunoaffinity column, Multiple Affinity Removal System (Agilent Technologies), which comprises antibodies to the six most abundant proteins found in human blood. By merely running a sample over the matrix, one can specifically remove all six proteins at once, unveiling lower-abundant species that may represent new biomarkers for disease diagnosis and therapy. The process removes about 85% of the total protein mass. The multiple affinity removal system works with blood, cerebrospinal fluid, and urine, all of which contain the same major proteins. Blood serum is the favored source for investigators interested in large-scale proteomics, because it has the most proteins. However, so far only about 500 of the 30,000 proteins in the serum have been identified. By removing albumin and the other five major proteins, scientists will be able to dig further into the proteome.

Not only has the number of proteins that can be detected in plasma expanded dramatically from hundreds to thousands, there is increased capability to detect structural variations of proteins. Recent studies also identified the presence of complex sets of small protein fragments in plasma. This set of protein fragments, the fragmentome or peptidome, is potentially a rich source of information about physiologic and disease processes. Advances in proteomics, therefore, offer great promise for the discovery of biomarkers that might serve as the basis for new clinical laboratory tests. There are many challenges, however, in the translation of newly discovered biomarkers into clinical laboratory tests. Only 10% of the proteins in human serum can be detected with currently available approaches, indicating the potential for further discovery of biomarkers. Protein variation is an untapped resource in the biomarker space, but only a selected few forms of proteomics applications are suitable for their analysis, and such variation could have a significant impact in disease diagnostics and therapeutic intervention (Kiernan 2008).

Biomarkers for Diagnostics

Currently available molecular diagnostic technologies have been used to detect biomarkers of various diseases such as cancer, metabolic disorders, infections, and diseases of the central nervous system. Some of the newly discovered biomarkers also form the basis of innovative molecular diagnostic tests. Those relevant to personalized medicine may be categorized as pharmacogenetic tests or pharmacogenomic tests.

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.

Diagnostic systems such as DNA microarrays and proteomics enable simultaneous assessment of multiple markers. Use of proteomic technologies for detection of biomarkers will be a later section in this chapter. Progress made in recent years suggests that pharmacogenomic biomarkers have the potential to provide physicians with clinically useful information that can improve patient care through increased individualization of treatment, particularly in the management of life-threatening disease.

Biomarkers for Drug Development

The advantage of applying biomarkers to early drug development is that they might aid in preclinical and early clinical decisions such as dose ranging, definition of treatment regimen, or even a preview of efficacy. Later in the clinic, biomarkers could be used to facilitate patient stratification, selection, and the description of surrogate endpoints. Information derived from biomarkers should result in a better understanding of preclinical and clinical data, which ultimately benefits patients and drug developers. If the promise of biomarkers is realized, they will become a routine component of drug development and companions to newly discovered therapies.

Use of Biomarkers for Developing MAb Therapy in Oncology

The significance of pharmacogenomics in monoclonal antibody (MAb) therapeutics is highlighted by the association between polymorphisms in Fc receptors and clinical response to anti-CD20 MAb rituximab (Rituxan) or anti-ganglioside GD2 MAb 3F8, as well as the potential link between polymorphisms in HER2 and cardiac toxicity in patients treated with the anti-HER2 MAb trastuzumab (Herceptin). The dependence on gene copy number or expression levels of HER2 and epidermal growth factor receptor (EGFR) for therapeutic efficacy of trastuzumab and cetuximab (Erbitux), respectively, supports the importance of selecting suitable patient populations based on their pharmacogenetic profile. In addition, a better understanding of target mutation status and biological consequences will benefit MAb development and may guide

clinical development and use of this innovative therapeutics. The application of pharmacogenetics and pharmacogenomics in developing MAb therapeutics will be largely dependent on the discovery of novel surrogate biomarkers and identification of disease- and therapeutics-relevant polymorphisms. There are many opportunities as well as challenges in biomarker discovery and validation, and in implementing clinical pharmacogenetics and pharmacogenomics in oncology MAb development.

Biobanking, Biomarkers and Personalized Medicine

The Biobanking and Biomolecular Research Infrastructure (BBMRI, www.biobanks.eu), which started the preparatory phase in February 2008, will pool all of the major biobanks in Europe. Together these represent approximately 12 million blood, body fluid, and tissue samples. In the following 2 years, BBMRI will try to create the preconditions to make the biological materials and data available, as well as to standardize the analyses platforms and sample preparation. The project not only includes the organization and funding of the EU biobank, but also aims to establish a complete resource for EU life scientists, including a variety of affinity binders and molecular tools as well as a biocomputing infrastructure that will work with standardized protocols, making data generated from those materials more comparable. The BBMRI was selected for FP7 funding as one of six EU infrastructure projects that are supposed to benefit all EU researchers. It is still awaiting the grant agreement from the European Commission.

No single biobank can be large enough to generate statistically significant data of specific disease subtypes and it takes more than a few dozen or even hundreds of cases in well-defined diseases to correlate disease history or patient response to a certain therapy and to biomarkers. The 134 associated partners of the BBMRI could together provide about 2.4 million samples from population-based biobanks, and a further 10 million from disease-orientated biobanks. The project will seek to overcome the current fragmentation in biobanking, and could also become an interesting tool for the biopharmaceutical industry when validating biomarkers. The information generated from BBMRI will be useful for the development of personalized medicine.

The joint initiative, which will tie together Europe's top research groups across almost every area of molecular and cell biology, also has a political dimension. Because the protection of the data obtained from biological samples continues to be a sensitive subject, the initiative will need to conform to all the national legislations involved. For that purpose, the partners plan to establish a widely-accepted and harmonized set of practices in line with the heterogeneous landscape of European and national regulations. For instance, the protocol to be added to the Convention of Human Rights, which was approved by the EU Council in 2007 and has now been sent out to member nations for ratification, states that the confidentiality of the information obtained through diagnostic, predictive, and pharmacogenetic tests of the samples must be assured. The researchers will have to find procedures that assure a

high degree of data protection while simultaneously allowing use of the patient data to acquire deeper insights into the causes of disease. Three types of biobanks have been considered as source of biomarkers in EU (Riegman et al. 2008).

1. Population banks. Their primary goal is to obtain biomarkers of susceptibility and population identity, and their operational substrate is germinal-line DNA from a huge number of healthy donors, representative of a concrete country/region or ethnic cohort.
2. Disease-oriented banks for epidemiology. Their activity is focused on biomarkers of exposure, using a huge number of samples, usually following a healthy exposed cohort/case-control design, and studying germinal-line DNA or serum markers and a great amount of specifically designed and collected data.
3. Disease-oriented general biobanks (i.e., tumor banks). Their goals correspond to biomarkers of disease through prospective and/or retrospective collections of tumor and no-tumor samples and their derivates (DNA/RNA/proteins), usually associated to clinical data and sometimes associated to clinical trials. Those data are usually not collected for a concrete research project, except in case of clinical trials, but from the healthcare clinical records. The amount of clinical data linked to the sample determine the availability and biological value of the sample.

Expression Signatures as Diagnostic/Prognostic Tools

Gene expression signatures as determined by microarrays can be used as biomarkers for diagnosis and monitoring of therapy. The best examples are in cancer. Ipsogen SA, has used gene expression signatures to refine molecular classes of breast cancer. Utilization of these signatures together with standard clinical parameters provides a unique combination discriminating patients responding to standard anthracycline chemotherapy. The test was validated in an independent cohort with patient samples from a multicenter clinical trial.

Althea Technologies Inc.'s proprietary eXpress Profiling™ multiplexed PCR technology, which enables high throughput gene expression analysis, is being combined with Natural Selection Inc.'s bioinformatics for discovery and application of gene expression signatures for a targeted disease or drug activity. This collaboration will provide advanced methods of data mining to extract biomarkers from the large gene expression data sets.

Biomarkers for Monitoring Response to Therapy

One of the important aspects of personalized medicine is the ability to monitor response to therapy. There are some examples in various diseases mentioned in the preceding chapters. A few more examples are given here to show the value of biomarkers and their limitations in monitoring response to therapy.

Sensitive noninvasive strategies for monitoring treatment response in rheumatoid arthritis (RA) would be valuable for facilitating appropriate therapy and dosing, evaluating clinical outcome, and developing more effective drugs. Because different proteases are highly up-regulated in RA and contribute significantly to joint destruction, the suitability of such enzymes as in vivo imaging biomarkers for early evaluation of treatment response was investigated in a murine model of RA (Wunder et al. 2004). Using a protease-activated near-infrared fluorescence (NIRF) imaging "smart" probe, the presence and distribution of fluorescence in arthritic joints of mice with collagen-induced arthritis was examined by both noninvasive fluorescence imaging and histology. Proteases that target the Lys-Lys cleavage site, including cathepsin B, activate probe fluorescence. Treatment monitoring data, obtained following methotrexate therapy, showed that protease-activated NIRF probes are sensitive means of imaging the presence of target enzymes in arthritic joints and can be used for early monitoring of treatment response to antirheumatic drugs such as methotrexate.

Assessment of hepatic damage associated with chronic hepatitis B (CHB) currently relies on measurement of serum transaminases and assessment of hepatic histology. It was determined by serum hepatic function tests and the liver fibrosis biomarkers type IV collagen (CIV), amino-terminal propeptide of type I procollagen (PINP), amino-terminal propeptide of type III procollagen (PIIINP), and carboxy-terminal telopeptide of type I collagen (ICTP) were used for monitoring the effect of lamivudine therapy for CHB (Maxwell and Flisiak 2005). Results showed that PINP/ITCP ratio is sensitive and specific in detecting responders to treatment.

Serial measurements of biomarkers might be beneficial to assess the adequacy of medication therapy for patients with advanced heart failure. Therapy guided by N-NT-proBNP, a biomarker of heart failure, might be helpful because NT-proBNP should be lowered by therapies that decrease endogenous BNP secretion. NT-proBNP and BNP were measured in a nonconsecutive patient cohort receiving clinically indicated intravenous nesiritide (Miller et al. 2005). In this study, many patients had decreased NT-proBNP and BNP values after therapy with nesiritide, but the majority of patients did not demonstrate biochemically significant decreases in analytes despite a clinical response. Until we know more about the responses of natriuretic peptides to therapies such as nesiritide, a strategy of monitoring NT-proBNP and BNP to guide therapy cannot be universally advocated.

Drug Rescue by Biomarker-Based Personalized Medicine

Biomarkers can rescue drugs by identifying the patients that respond to them. Herceptin, approved in 1998, emerged as a $480 million-per-year winner only a decade after clinical trials showed little or no efficacy. Only when the 20–30% of women whose tumors overexpressed HER2 were singled out, was the drug's efficacy indisputable. In the pivotal clinical trial of patients with metastatic breast cancer, tumor-response rates to Herceptin plus chemotherapy were 45%, compared to 29% for chemotherapy alone.

But the response is not wholly predictable. Reported response rates for HER2-positive cancers vary from less than 20% to more than 75%. HER2-positive cells that don't respond to Herceptin may have more active forms of the kinase Akt. And HER2 belongs to a receptor family that can be activated by 11 different soluble proteins and combinations thereof. Researchers are already betting that working out the biology behind the biomarker will lead to better treatments. Another anticancer antibody based on this understanding is already in clinical trials.

Similarly, the lung-cancer drug Iressa (gefitinib) could be rescued by a diagnostic based on a biomarker. Unfavorable clinical trial results dashed high hopes for big sales, but finding the patients most likely to benefit changed prospects. Various studies found that patients who responded to Iressa had mutations in the gene for EGFR.

Future Role of Biomarkers in Personalized Medicine

Personalized medicine is being recognized by the biopharmaceutical industry, regulatory authorities, healthcare providers, and the medical profession. It should be a part of the healthcare system by the year 2013 and will mature by 2015. Genetic testing will improve predictions of disease predisposition, onset, severity, and treatments or medications that are likely to be efficacious or harmful.

Summary

This chapter introduces biomarkers and technologies for their discovery. The important points on the role of biomarkers in the development of personalized medicine are:

- Biomarkers will enable early diagnosis of disease to facilitate optimization of therapy.
- Biomarkers will play an important role in combining diagnosis with therapeutics – an important feature of personalized medicine.
- There will be an increase in the number of new drugs suitable for personalized treatment, which will be discovered by use of biomarkers.
- Validated biomarkers will play an increasing role in clinical trials for personalizing therapeutics.
- Biomarker-based monitoring of drug efficacy will guide personalized management of several diseases.Future Role of Biomarkers in Personalized Medicine