BIOCHIP JOURNAL, Vol. 2, No. 3, 160-174, September 2008

Mini Review

Cancer Biomarkers in 'Omics Age

Young Il Yeom[1], **Seon-Young Kim**[1], **Hee Gu Lee**[1] **& Eun Young Song**[1]

[1]Medical Genomics Research Center, Korea Research Institute of Bioscience and Biotechnology, 111 Gwahangno, Yuseong-gu, Daejeon 305-806, Korea
Correspondence and requests for materials should be addressed to Y.I. Yeom (yeomyi@kribb.re.kr)

Accepted 5 August 2008

Abstract

In spite of the remarkable scientific and technological advances in medicine achieved during the last two decades, cancer incidence is still increasing worldwide and the general ratio of deaths to new cancer cases remains as high as 49% overall, necessitating the discovery of more effective cancer therapeutics and diagnostics. Diagnosis of cancer involves areas like disease susceptibility, detection, prognosis, and monitoring. Biomarkers are the key element of modern diagnostics and their value in medicine is ever increasing. They can be used to diagnose disease risk or presence of disease in an individual, or to tailor treatments for the disease in an individual. Biomarkers are also useful for understanding pathological mechanisms as well as for the development of therapeutics. To date, the role of biomarkers in cancer has been mainly focused to disease detection and prognosis. However, with the personalized medicine coming into practice and thanks to the availability of 'omics-based enabling technologies following the completion of Human Genome Project, the spectrum of current biomarker application is being rapidly expanded such that considerable emphasis also goes to the role of biomarkers in therapeutic response prediction and pharmacodynamics of drug activity. The 'omics technologies have been generating huge amount of data for human genomic variations, which can be transformed into cancer susceptibility biomarkers when combined with epidemiological data. The increasing demand for cancer prevention is likely to leverage on the cancer susceptibility biomarkers in the future to eventually accomplish the personalized medicine in the perspectives of cancer prevention. This concept is increasingly becoming an attainable objective as the remarkable innovations in sequencing technology are, sooner or later, very likely to make the age of 'personal genome' a reality. We review the current trend of biomarker discovery and application, emphasizing their diverse roles in medical practice and drug development.

Keywords: Cancer, Biomarker, Diagnostics, 'Omics, Personal medicine

Introduction

Cancer is the number two killer disease after cardiovascular disease worldwide. According to the World Health Organization (www.who.int/en), approximately 25 million people have cancer in the economically developed countries of Japan, Europe and North America, and there will be 15 million new cases every year by 2020. In the economically developing countries of Africa, Latin America and Asia, the number of new cancer cases is expected to reach 10 million a year in 2015 as reported by the United Nations International Atomic Energy Agency (IAEA, www.iasea.org), and the total number of cancer cases will double by 2015 to account for more than a quarter of all deaths in many countries. Statistics from the WHO also indicate that 10.1 million new cases of cancer are diagnosed each year worldwide, and that the annual cancer deaths worldwide are expected to double to 12 million in 2020 from six million in 2000.

The patterns of cancer epidemiology are sharply contrasted according to economic power and scale of a country. The economically developed countries tend to have relatively high rates of cancers of the colon and rectum, and of the hormone-related cancers of the female breast, the endometrium and the prostate. The economically developing countries, on the other hand, tend to have high rates of cancers of the mouth and pharynx, larynx and esophagus, and of the stomach, liver, and cervix. Lung cancer, mainly caused by use of tobacco, is the most common cancer throughout the world.

In spite of the remarkable medical and scientific advances achieved during the last two decades to fight cancers, the general ratio of deaths to new cancer cases worldwide still remains as high as 49% overall, necessitating the discovery of more effective cancer therapeutics and diagnostics. Diagnosis of cancer involves areas like disease susceptibility, detection, prognosis, and monitoring. In current hospital settings diagnostics generally influence 60-70% of all critical decision making, including admittance, discharge and medication.

The performance of current technologies to detect cancers is rather disparate in two histologically distinct categories of cancer. For the cancers of blood such as lymphomas, myelomas and leukemias the detection technology is mature and well established in such a way that routine hematological blood count screens for blood cell abnormalities, followed by bone marrow analysis and flow cytometric studies of blood cell surface markers to arrive at a specific diagnosis. In contrast, the diagnostic modalities for the detection and management of solid tumors are performing relatively poorly and require much technological improvements and even innovations.

Biomarkers are the key element of modern diagnostics and their value in new diagnostics development is ever increasing so that biomarker R & D bears a direct relationship with the market trend of diagnostics. This review describes different types of cancer biomarkers and their application, including their role in driving recent trends in personalized cancer therapy.

History of Biomarker-based Diagnostics and New Trends in 'Omics Age

A biomarker can be any kind of molecule indicating the existence, past or present, of living organisms. In the fields of medicine, a biomarker is generally defined as a substance whose detection indicates a particular disease state, and is often represented by a change in expression or state of a protein, RNA, DNA or metabolite that correlates with the risk or progression of a disease or with the susceptibility of the disease to a given treatment. Thus biomarkers can be used to diagnose disease risk or presence of disease in an individual, or to tailor treatments for the disease in an individual (choices of drug treatment or administration regimes).

Biomarkers are also useful for understanding the pathological mechanisms and form a basis for therapeutics development. Indeed, among the many applications of biomarkers the most important ones are in drug discovery and development including clinical trials. Biomarkers can facilitate the combination of therapeutics with diagnostics and thus will play an important role in the development of personalized medicine by mediating the use of pharmacogenetics, pharmacogenomics and pharmacoproteomics. A recent biomarker market survey reported that currently, 20% of pharmaceutical R & D is improved by postgenomic biomarkers but this is expected to increase to 80% by 2010[1]. It is also predicted that the number of clinical trials using biomarkers will increase over the next five years, and that most of clinical trials at major pharmaceutical companies will have biomarkers included in the protocol by 2010. In some cases biomarkers will be used mainly to identify responders to treatment prior to enrollment, while in other cases, a biomarker strategy will be needed to gain the management's approval for compounds to advance.

The first biomarker-based laboratory test for cancer diagnosis was carried out in 1847 by detecting a protein marker called 'Bence Jones protein' from the urine of a multiple myeloma patient[2]. A Bence Jones protein is a free light chain of a monoclonal antibody immunoglobulin produced by neoplastic plasma cells. Much later, in 1954, Karmen *et al.* reported the measurement of transaminases in myocardial infarction[3]. Then, in the 1960s, the term "biomarker" started to appear in the literature in connection with metabolites and biochemical abnormalities associated with several diseases. Since the 1960s the mainstay of cancer diagnosis has been the histological analysis of biopsied tissue using microscopy. The performance and accuracy of the histology-based cancer diagnosis was significantly improved later with the arrival of immunohistochemical stains (IHC) that use fluorescent and colorimetric labeled antibodies to detect enzymes, proteins and other analytes in tissue as well as with the introduction of *in situ* hybridization techniques such as fluorescent *in-situ* hybridization (FISH) that use DNA probes to detect molecular structures in cancer cells. Biomarkers that can be detected from body fluids have the advantage of being more accessible and more likely to be of clinical use because serum or urine can be obtained by non-invasive method. Therefore, much of the efforts for biomarker-based diagnostics development was also directed to non-invasive serum-based protocols. One example is the improved diagnosis of myocardial infarction using serum creatine phosphokinase as the biomarker[4]. In 1971 Moore *et al.* reported carcinoembryonic antigen (CEA) as a biomarker of cancer[5]. Technological advances in the 1990s introduced the mass spectrometry for biomarkers analysis in biological samples. Another remarkable achievement in biomarker discovery was the sequencing of human genome in 2003, opening the way for discovery of gene biomarkers in a large scale. Various biomarker-based imaging techniques have also been developed in the 2000s and routinely used in clinical practice to diagnose cancers *in vivo* in a non-invasive manner. There is also a continuous movement of new tests and technologies coming to the fore for early cancer detection. Now in the mid-2000s the discovery and application of biomarkers has become a major activity in biotechnology and biopharmaceutical industries.

Table 1. Examples of 'omics-based cancer diagnostics (Source: Ref. 6).

Company (year founded)	Description	Marketed tests or tests in late development	Key disclosed alliances
Agendia (2003)	Uses gene expression profiling and computer algorithms to predict risk of cancer spread, recurrence, response to certain drugs or primary site of a tumor	MammaPrint (breast cancer recurrence), CupPrint (primary tumor identification)	Netherlands Cancer Institute supplies samples and expertise
Aureon Laboratories (2001)	Uses *in situ* RNA and protein imaging plus other clinical factors to predict recurrence and stage cancer	Prostate Dx (prostate cancer recurrence)	Baylor College and Memorial Sloan-Kettering supply samples and expertise
Correlogic Systems (2000)	Develops tests through a patented pattern recognition approach using pattern discovery and recognition software, mass spectrometry and proprietary lab procedures. Validating tests in breast, colon, ovarian, and prostate caners	OvaCheck (early detection of ovarian cancers; in discussion with US FDA)	Charles Stark Draper Laboratories, Univ. of Alabama at Birmingham, Johns Hopkins Univ., Univ. of the Health Sciences Walter Reed Army Medical Center
Dako (1966)	Develops diagnostics using immunohistochemistry and flow cytometry	Multiple tests, including ones that detect c-Kit, HER-2/neu, EGFR expression	No alliances disclosed
Orion Genomics (2002)	Uses proprietary methylation microarrays to detect, type and stage of cancers	None yet	Univ. of Glasgow, Johns Hopkins and Washington Univ. supply samples and expertise
Prediction Sciences (2000)	Creates panels of known markers integrating genetic variation, gene and protein expression	GeneR (integration of patient and disease traits/ medical chart data to predict patient response)	Burnham Institute and Mayo Clinic supply samples and expertise
Rubicon Genomics (1998)	Amplifies abnormally methylated DNA released into blood and urine by diseased tissue	TransPlex (whole genome amplification kit for expression profiling), Contract services	Abbott Molecular Diagnostics; analyzing samples for Genome Institute of Singapore
Epigenomics (1998)	Identifies DNA methylation patterns in tissue and blood	Colon cancer detection test; services in methylation proofing and sequencing	Roche Diagnostics, several pharmaceutical companies and universities
Genomic Health (2000)	Identifies gene expression patterns in tumor samples to predict response to treatment and risk of recurrence	Oncotype Dx	ImClone, Bristol-Myers Squibb; National Surgical Adjuvant Breast and Bowel Project provided samples and expertise
Genzyme Genetics (1986)	Licenses markers and sells tests to type cancer and other conditions	Sells tests assessing EGFR expression, Glivec resistance, irinotecan toxicity and more	Multiple academic labs, Third Wave Technologies
Monogram Bioscience (1995, formerly Virologic)	Identifies activated proteins signaling pathways in cancer cells; assesses drug sensitivity of HIV in individual patients	Sells HIV resistance tests	Collaborating on clinical trials with Pfizer and other pharmaceutical companies
OncoMethylome (2003)	Identifies gene methylation patterns to detect cancer, predict response to therapy and predict recurrence	Developing tissue- and urine-based assays for early diagnosis of multiple cancers	Johnson & Johnson, Veridex; clinical collaboration with Exact Sciences

Recently, the pattern of diagnostics development is undergoing significant changes as the advances in genomics and proteomics demonstrate the possibility of discovering abundant, informative biomarkers that

Table 2. Desirable characteristics of cancer biomarkers.

Purpose	Characteristics				
	Noninvasive	Low cost	Simple to perform	Accurate*	Informative (Discriminatory)
Screening	+++	+++	+++	+++	+++
Predisposition	+++	+++	+++	+++	+++
Early detection	++	++	++	+++	+++
Prognosis	+	+	+	++	++
Drug response	+++	++	++	+++	+++
Target for drug	NA	+	NA	+++	NA

+=low importance, ++=medium importance, +++=high importance, NA=not applicable, *low rate of false-negative results. (Source: Ref. 1)

could revolutionize the medicine and health industry in the future. Biomarkers are considered as critical functional units in transforming the 'omics information into clinically useful forms that are applicable for human health management. Thus, the 'omics paradigms (such as, genomics, proteomics, metabolomics) and enabling technologies (DNA and protein chips, microfluidics, giga-sequencing) are rapidly redefining the taxonomy of disease as biomarker-identified subgroups that can be addressed in novel ways. These advances have allowed the creation of new 'omics-based *in vitro* diagnostic (IVD) tests that quantitatively measure response to therapy, can monitor disease progression or predict disease recurrence (Table 1). In addition, introduction of targeted therapeutics has called for the development of matching biomarker-based diagnostics that can predict the responsiveness of a patient to the therapy so that responder groups can be pre-selected during clinical treatment as well as in clinical trial for drug development. This situation is promoting biomarker-driven strategic alliances between pharma and diagnostic partnerships. Accordingly, patients' view relative to disease shift towards personalized medicine such that patients are informed on the disease very specifically based on biomarkers and understand that the information subsequently can guide the patient and physician to manage the disease in a tailored way. In addition, DNA biomarkers generated from the genomics-based analysis of human samples in association with clinical information promise for early cancer assessment in healthy individuals, especially in at-risk groups of people.

Types of Cancer Biomarkers and Their Application

Biomarker Discovery and Validation

Jane indicated that desirable characteristics of molecular markers for cancer should include noninvasiveness of application, low cost and simplicity of detection, low false-negative rate (accuracy), and high informativeness (discriminatory power)[1]. Table 2 summarizes the relative importance of these characteristics in different application purposes. No one test meets all these requirements but these should be kept in mind for selection of diagnostic tests. There is an urgent need for cancer biomarkers with more accurate diagnostic capability, particularly for early stage cancer[1].

It has been suggested that the development of a useful biomarker-based diagnostic test requires three steps: discovery, development and evaluation[7]. 'Discovery' is the process by which candidate genes, proteins, antigens or imaging tools are identified. Most of the biomarker discovery so far has been achieved through a one-at-a time approach. Many of the well-known (diagnostic, blood, and serological) tests have been identified based on clear biological insight from physiology or biochemistry. This means that only a few markers at a time have been considered. One example of this way of biomarker discovery is the use of injections of inulin for measuring kidney function. From this, one discovered a naturally occurring molecule, creatinine, which enabled the same measurements to be made easily without injections.

The recent interest in biomarker discovery is because new molecular biologic techniques promise to find relevant markers rapidly, without detailed insight into mechanisms of disease. Molecular features that are the focus of current biomarker discovery in cancer include protein biomarkers showing significant changes in expression, localization or posttranslational modification in cancers, RNA biomarkers whose expression changes in cancer in association with clinical parameters or therapeutic response, DNA biomarkers such as DNA copy number variation (CNV), gene mutation, single nucleotide polymorphism (SNP), microsatellite variation, epigenetic

Table 3. Currently developed cancer biomarkers.

Cancer	Body fluid	DNA biomarker	RNA biomarker	Protein biomarker
Lung cancer	Saliva, Serum	RAS/TP53 mutations, MS alterations, methyl-p16/MGMT	Cytokeratins, MAGE genes, CEA	CEA, CA125, telomerase, CYFRA
Head & neck cancer	Saliva, Serum	*TP53*/MS alterations, HPV/EBV DNA, methyl-p16A/MGMT/DAPK	Cytokeratins	SCC, CD44, CYFRA, telomerase
Breast cancer	Serum, Nipple aspirate	MS alterations, methyl-p16/RARβ	Cytokeratins, hMAM, MAGE genes, CEA	CA15-3 (MS-1), CEA, CA125
Colorectal cancer	Stool, Serum	RAS/APC/TP53 mutations, methylMLH1/p16	Cytokeratins, CEA	CEA, CA19-9, CA15-3, telomerase
Pancreatic cancer	Stool, Serum	RAS/TP53 mutations	Cytokeratins, CEA	CA19-9
Bladder cancer	Urine, Serum	TP53 mutations, MS alterations, methylRASSF1A/RARβ	Cytokeratins, survivin, uroplakin	CEA, CA125, CA19-9, telomerase, survivin, CD44
Prostatic cancer	Urine, Serum	Methyl-GSTP1/CD44	PSA, MAGE genes, kallikrein	PSA, free PSA, telomerase, kallikrein
Liver cancer	Serum	Methyl-p16/p15	–	AFP
Stomach cancer	Serum	Methyl-E-cadherin/p16/ p15/DAPK1/GSTP1	–	Pepsinogen, gastrin17

changes (i.e., DNA methylation) and foreign DNAs (i.e., viral genome), and metabolites that are overproduced by cancer cells or cells in tumor microenvironment. Table 3 lists some of the representative cancer biomarkers developed for the diagnosis of different cancer types.

Various methodologies exist that have been used for cancer biomarker discovery, including aCGH (array -based comparative genomic hybridization), DNA sequencing, mass spectrometry, DNA microarray, multiplex PCR, mutation-specific PCR, etc. Among these varieties high-throughput technologies based on genomics (such as DNA microarray and sequencing) and proteomics (such as high throughput differential expression analysis combined with mass spectrometry) have been highly productive in discovering a large number of biomarkers in a systematic manner. The transcriptome approach identifies genes or proteins that are either overexpressed or underexpressed in most tumours of a given type as candidates for early-detection markers (transcript or gene-expression analysis). The proteome approach searches the serum directly for protein signatures that distinguish cases from controls (proteomics). Although the studies published so far have generated enormous excitement, a great deal of further investigation is required to move from a finding of differential gene or protein expression to a clinically viable screening test (Development) and to conclusively show that the test is effective and practical for mass use (Evaluation).

Achieving this stepwise goal in practice can be greatly facilitated by implementing a standardized protocol that academic researchers, regulatory agencies and administrators of biorepositories can agree on. Pepe *et al.* have suggested, for the first time, a formal protocol to guide the process of biomarker development, where the types of studies that are required were organized into five phases[8]. These phases are ordered according to strength of evidence from weakest to strongest, and the results from earlier phases will typically be required to justify conducting later-phase studies. Although this protocol has been designed for the development of cancer diagnostics for early detection purposes, the basic principle may also apply to the diagnostics development for other purposes.

Phase 1: Preclinical exploratory studies evaluate the expression of thousands of genes or proteins in tumour and comparable healthy organ tissue to identify candidates for early detection.

Phase 2: Phase 2 studies focus on the development of clinical assays to measure markers in specimens. Assays can be developed that are more efficient than those used for discovery, including immunoassays such as ELISA to detect proteins in fluids, polymerase chain reaction tests to identify RNA or DNA, methylation tests and proteomic profiles such as those produced by SELDI-TOF. The goals are to develop clinical assays that are reproducible within and between

laboratories, to confirm the correlation between these assays and the corresponding Phase 1 studies, and to evaluate their ability to discriminate between patients with clinically established disease and population controls.

Phase 3: Phase 3 studies focus on biomarker measurements in cases before diagnosis. Samples obtained from individuals before they were diagnosed with the cancer of interest are compared with samples from healthy age-matched controls. Because the cases' samples have been obtained before their diagnosis, they allow for the evaluation of biomarker levels during the preclinical phase of the disease. Phase 3 studies are vitally important because they provide a window into the natural history of the disease and how it relates to the measurement of the biomarker under study. In the case of the prostate-specific antigen (PSA) used for prostate cancer screening, for example, Phase 3 studies provided uniquely valuable information about the amount of time by which measuring PSA could advance prostate cancer diagnosis (the lead time)[9], and the sensitivity of the test, which is typically impossible to infer from prospective screening studies. Phase 3 studies are also important because they provide information on how marker levels change over time in disease cases and in healthy individuals.

Phase 4: Phase 4 studies prospectively screen an asymptomatic population and rigorously follow up individuals who test positive to provide important information about the prevalence of detectable disease in the population and the test's specificity.

Phase 5: Phases 1-4 focus exclusively on developing tests that are feasible for widespread use and evaluating their diagnostic performance. Even if a test performs well through to Phase 4, this does not necessarily imply that the test will reduce the population burden of disease in a meaningful way. It must be shown conclusively that interventions that are used as a result of a positive test reduce mortality. Phase 5 studies directly evaluate the impact of a diagnostic test on population disease morbidity and mortality, and include randomized, controlled cancer screening trials, as well as a number of other study designs, including case-control studies, computer modelling studies and population studies.

Types and Application of Cancer Biomarkers

Two types of cancer biomarkers exist, reflecting the

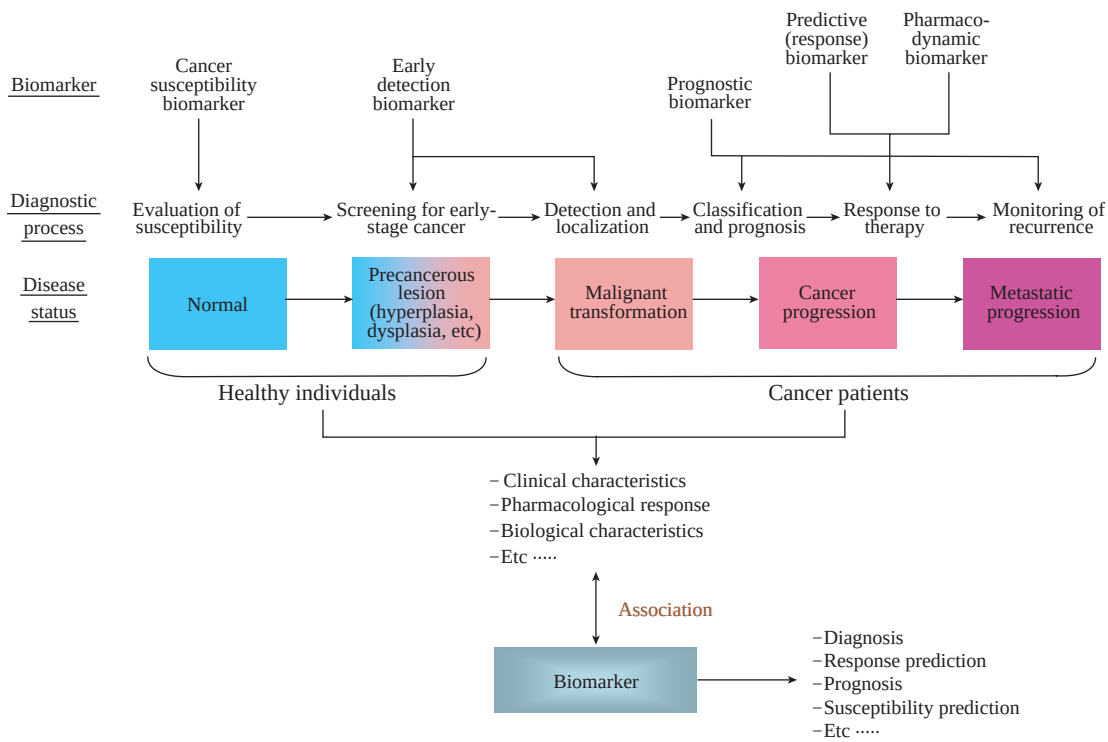

Figure 1. Functional categories of biomarkers required throughout the process of cancer development and progression or before the cancer development. Among the five types of cancer biomarkers two, cancer susceptibility biomarkers and early detection biomarkers, have major use in the risk estimation or detection of cancer as a new disease in the people that have not been diagnosed of the particular cancer type, while the other three, predictive (response) biomarkers, prognostic biomarkers and pharmacodynamic biomarkers, have major use in therapeutic purposes and drug development.

chronology of technology development for biomarker discovery. One is the classical biomarkers such as measurable alterations in blood pressure, blood lactate levels following exercise and blood glucose in diabetes mellitus. The other is referred to as molecular biomarkers, and represented by any specific molecular alterations of a cell on DNA, RNA, metabolite or protein level. In the era of molecular biology, biomarkers usually mean molecular biomarkers.

Cancer is a progressive disease that has both genetic and environmental associations, and display quite distinct clinical features during the disease evolution. Therefore, distinct functional categories of biomarkers have been developed to meet different needs required throughout the process of disease development and progression or even before cancer arises in order to guide effective prevention, diagnosis and treatment of cancer. Consistent with this context, cancer biomarkers can be classified into five types according to its use in clinical practice and drug development (Figure 1). Two of them, cancer susceptibility biomarkers and early detection biomarkers, have major use in the risk estimation or detection of cancer as a new disease in the people that have not been diagnosed of the particular cancer type, while the other three, predictive (response) biomarkers, prognostic biomarkers and pharmacodynamic biomarkers, have major use in therapeutic purposes. The latter three types of biomarkers can also aid in the rational development of anticancer drugs as suggested by Sawyers[10]. The characteristics and application for each type of cancer biomarkers are described below.

Cancer Susceptibility Biomarkers

Cancer susceptibility biomarkers allow the assessment and prediction of an individual's risk to get stricken with cancer by exposure to environmental exposure and/or due to inherited genetic variations while one is still healthy. Thus, individuals can be stratified according to their cancer risk to identify those most or least likely to benefit from screening and prevention strategies. By targeting a population with an increased cancer incidence, screening tests will have a higher positive predictive value, resulting in a higher pretest likelihood of benefit. Several strategies can be used to stratify risk, including genetic testing, demographic information (including information on health habits and exposures), epigenetic markers and physiological tests.

The most common markers of susceptibility are mutations in specific genes that confer increased or decreased risk. Germline mutations are inherited from a parent and are present in each cell in the body. Up to 5% of breast cancer and up to 15% of colorectal cancer are attributable to these high penetrance hereditary germline mutations although an even larger percentage is associated with other low penetrance mutations. Genetic testing is clinically available for an increasing number of cancer susceptibility syndromes, allowing more precise risk identification than that based on family history or phenotype alone. For example, mutations in *BRCA1* and *BRCA2* are known to account for up to 50% of hereditary and familial breast cancer[11], and identification of deleterious mutations in these genes identified women with markedly increased risks of ovarian and breast cancer[12]. Rare high-penetrance mutations in *ATM*, *PTEN*, and *TP53* can also provide highly informative biomarkers for breast cancer susceptibility. Likewise, germline mutations in two mismatch repair genes, i.e., *hMSH2* and *hMLH1*, are responsible for 70% to 90% of all hereditary nonpolyposis colorectal cancer (HNPCC) cases[13], and thus can serve as the high-penetrance susceptibility biomarkers for colorectal carcinomas.

In addition to identifying genetic risk, biomarkers may predict risk based on exposures (environmental and lifestyle factors potentially related to cancer) or other epidemiological risks. Measuring exposures can be very challenging, particularly if the relevant exposures occurred in the distant past. Biomarkers can help pinpoint exposures and clarify risk by directly identifying a causative agent (e.g., the presence of high-risk human papillomavirus (HPV) DNA in cervical secretions) or by indirectly revealing effects of an exposure (e.g., the identification of DNA adducts in at-risk tissues after exposure to carcinogens). HPV selectively infects the epithelium of skin and mucous membranes. Among more than 100 genotypes of HPV that have been identified, only a subset of specific HPV types are associated with squamous cell carcinoma, adenocarcinoma, and dysplasias of the cervix, penis, anus, vagina, and vulva[14]. Using current technologies, HPV DNA can be detected in 95% to 100% of cervical cancer specimens[15], and has been called a "necessary cause" of cervical cancer[16,17]. Since most women with evidence of HPV infection do not develop cervical cancer and most infections resolve within 1 to 2 years, it is important to define high and low risk groups by finding the persistence of oncogenic HPV infection and the progression of early cervical abnormalities to invasive cancer. It has been suggested that the amount of virus present in cervical tissues, as estimated by quantitative PCR, may distinguish HPV carriers at low and high risk of development of cervical cancer[18].

Genetic polymorphisms that appear to have no relation to cancer risk may have important associations in the face of specific environmental or dietary exposures. For example, certain human leukocyte antigen (HLA)

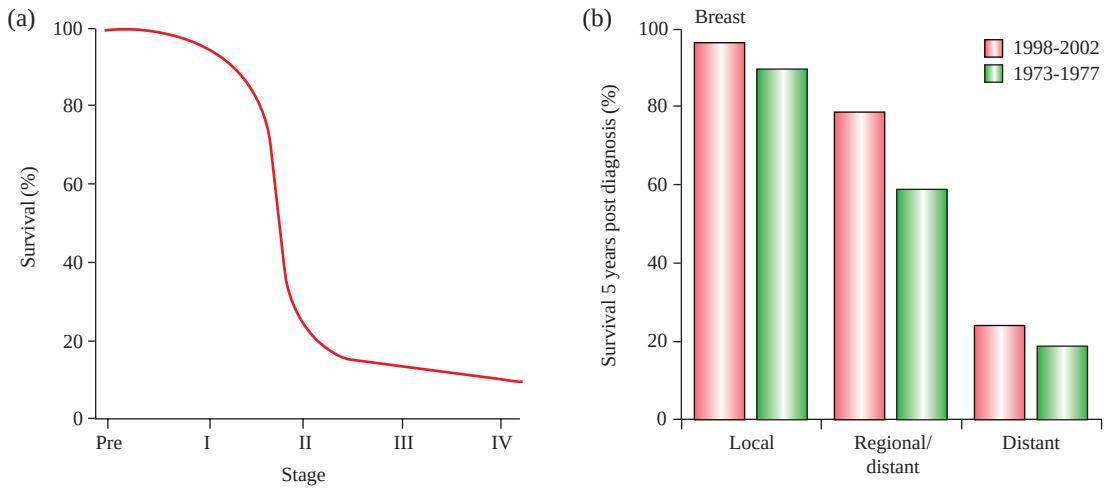

Figure 2. The value of early screening of cancer. a, If a tumour can be detected before it becomes malignant (pre-malignant stage) or while it is still small and locally confined (stage I), the chances of a cure are high. They drop sharply once the cancer spreads to local tissues and lymph nodes (II-II) or other organs (IV). b, Despite improvements in treatments since the 1970s, patients still have a greater chance of surviving if cancers are detected when they are locally confined (local) and not spread to neighbouring tissue (regional) or other organs (distant). (Cited from Ref. 23)

class II alleles in combination with exposure to specific human papillomaviruses may increase the risk of cervical carcinoma[19]. Another example is the increased colorectal cancer risk associated with the polymorphism of a metabolic enzyme, N-acetyltransferase 2 (*NAT2*), which metabolizes the carcinogen, heterocyclic amines generated from high temperature cooking of animal proteins[20]. SNPs in the coding sequence for *NAT2* give rise to different forms of the NAT2 protein and these divide Caucasian populations into slow acetylators (about 55% of the population) and rapid acetylators (about 45%)[21]. Several studies have shown that colorectal cancer cases are more likely to be rapid acetylators if they consume diets high in red meat intake[22]. Thus identification of *NAT2* genotype can lead to preventive advice for populations (i.e., changes in meat cooking practices) or individuals (i.e., those with susceptibility genotypes for common environmental exposures). Efforts are ongoing to search for additional susceptibility biomarkers representing between-person genetic variations for the genes involved in DNA repair, cell cycle control, immune response, and the inflammatory response.

Early Detection Biomarkers

Early detection biomarkers allow the detection of cancer in the earliest stage of development so that the individual can have maximal benefits from accompanying disease management. If a tumour can be detected early, i.e., before it becomes malignant (pre-malignant stage) or while it is still small and locally confined (stage I), the chances of a cure are high (Figure 2a). However, these chances drop sharply once the cancer spreads to local tissues and lymph nodes (stages II-III) or other organs (stage IV). Early detection has played a key role in the management of cervical and breast cancer, and is likely to become more important in the control of colorectal, prostate and lung cancer. Despite improvements in treatments since the 1970s, patients still have a greater chance of surviving if cancers are detected when they are locally confined (local) and not spread to neighbouring tissue (regional) or other organs (distant) (Figure 2b). Therefore, early detection represents one of the most promising approaches to reducing the growing cancer burden.

The World Health Organization has identified a number of conditions for early detection to be an appropriate disease-control approach. First, the disease must be common and associated with serious morbidity and mortality. Second, screening tests must be able to accurately detect early-stage disease. Third, treatment after detection through screening must have been shown to improve prognosis relative to usual diagnosis. Finally, evidence must exist that the potential benefits outweigh the potential harms and costs of screening[24]. The expectation that these conditions could be satisfied for many cancers has made early detection a topic of intensive research for the past several decades.

As incidence in cancer increases, the needs of early detection of cancer are increasing. Although screen-

ing tests are in use for a range of cancers, almost none of the available tests satisfy all of these requirements, and as a result, newer and better diagnostic kit will be demanded to improve early stage diagnosis. For example, biomarkers that are currently used for ovarian cancer screening-primarily CA125-have false-positive rates that lead to an unacceptably high ratio of surgeries conducted (for confirmation of disease) to cancers detected and fail to identify many early-stage cancers[25]. In prostate cancer screening, the PSA test carries a non-trivial risk of overdiagnosis due to the test's inability to clearly differentiate indolent cases from more aggressive cancers[26].

Early-detection research has recently been revitalized by the advent of 'omics-based molecular technologies that have produced hundreds of potential biomarkers for detecting and classifying cancers. Some of these biomarkers lead directly to novel diagnostics that promise to overcome the deficiencies of existing screening tests. For example, a recent analysis of proteomic patterns in the serum of ovarian cancer patients yielded a profile that distinguished cancer cases from controls with near-perfect sensitivity and specificity[27]. Similarly, several biomarkers that have been identified through expression array analysis have been shown to be predictive of the risk of biochemical recurrence (rise in PSA levels) after initial treatment for prostate cancer[28].

Researchers try to devise straightforward, non-invasive tests, such as screening blood samples, to pick up on the earliest signs of cancerous changes. But a big challenge is that blood is a complex soup of molecules containing everything from the ubiquitous protein albumin to smaller molecules present only at trillionths of albumin's levels, while the important markers are likely to be in the low-abundance range. Nevertheless, detection of cancers at the very earliest stages is within sight due to advances in sophisticated technologies such as mass spectroscopy combined with high-throughput proteomic profiling or various sequencing- or PCR-based tools to detect mutant or modified DNAs, which make it possible to pick up very low levels of tell-tale molecules. For example, Vogelstein showed that it is possible to detect trace amounts of mutated DNA of adenomatous polyposis coli (APC) gene against a noisy background of unmutated DNA by PCR in a simple blood sample from the patients with advanced colorectal cancer[29]. Vogelstein's team also detected mutant APC molecules in more than 60% of patients with early, curable colorectal cancer where these molecules were circulating in the blood in extremely low quantities.

There is a consensus in cancer diagnostics community that using a single biomarker to diagnose a cancer is unlikely to suffice. No single biomarker can detect a given cancer with 100% sensitivity (meaning that all diseased subjects would test positive) and 100% specificity (with all healthy subjects testing negative). Panels of biomarkers with different individual sensitivities and specificities are therefore needed. For example, Sidransky's group at Johns Hopkins University tries to find ways of improving the accuracy of the blood tests used to detect PSA. PSA is produced by prostate cancers but can also result from benign prostrate enlargement or inflammation. Sidransky's team has identified an alteration to a gene called *GSTP1* that is unique to prostate cancer cells so that men with high levels of PSA in their blood could be referred for a biopsy to test for this genetic alteration, to reduce the number of false positives that result from relying on PSA levels alone[30]. Another example, Hanash's group at Fred Hutchinson has been looking for markers of early-stage lung cancer, using a two-step, proteomics-based strategy[31]. The first step is to look for proteins that have been shed by the tumour and are circulating in the bloodstream; the second step is to search for an immune response to the tumour. The host immune system may not be able to destroy the tumour cells, but it probably recognizes them as problem cells, and Hanash wants to harness that raised antibody response as a diagnostic tool. All these approaches promise new models for disease discovery and improve the future availability of multiple markers for early detection, allowing more complete coverage of the spectrum of cancers than ever before.

Prognostic Biomarkers

Prognostic biomarkers allow the natural course of an individual cancer to be predicted, distinguishing 'good outcome' tumours from 'poor outcome' tumours, and they guide the decision of whom to treat (or how aggressively to treat). The risk of cancer recurrence is high in patients who have previously had cancer, even for those who have been in remission for five or more years. Therefore, cancer survivors constitute a high risk group that could benefit from surveillance for early detection of disease recurrence. The traditional system of tumor node metastases (TNM) has been the main tool for identifying prognostic differences among patients and for guiding the treatment. But, there are many limitations of this system as a first line method for prediction and prognosis of cancer. It is difficult to distinguish related disease subtypes. Because cancer survivors and their physicians have heightened awareness of possible disease recurrence and lower thresholds for moving to costly and potentially morbid diagnostic procedures, it is

important to avoid false-positive surveillance tests. Biomarkers may be particularly helpful in these settings. As the characteristics of the initial tumor are already known, biomarkers could in some cases be tailored to detect the patient's tumor cells, thereby maximizing specificity. For example, monitoring the chronic myeloid leukaemia (CML) patients, who received bone marrow transplantion, for the persistence of the BCR-ABL translocation is an effective surveillance technique.

Biomarkers capable of early detection of recurrent cancer could result in effective treatment for a selected subset of solid-tumor patients with metastatic disease. For example, a subset of breast cancer patients diagnosed with early-stage, high-risk tumors benefits from systemic adjuvant chemotherapy or hormonal therapy in addition to surgery and local radiation[32]. Presumably, this is because micrometastatic disease is present, but not detectable, and is effectively treated by systemic therapy. In lung cancer, stage I non-small cell lung cancer (NSCLC) patients normally receive surgical treatment alone, but 35-50% of them will relapse within five years. Thus, identifying high-risk patients for adjuvant chemotherapy from stage I NSCLC is important for the appropriate treatment of the disease[33].

Combinations of biomarkers for surveillance of recurrence may be particularly advantageous. For example, serum thyroglobulin serves as an important biomarker for surveillance of previously treated thyroid cancer and can identify clinically occult disease; however, when used alone it cannot assess the risk of tumor progression or death and may lead to over-treatment of otherwise indolent disease. The combination of thyroglobulin measurements with fluorodeoxyglucose PET, however, can identify those cancers most likely to cause death and direct more aggressive treatment[34].

Accurate prognosis can also help patients avoid unnecessary treatments that are prescribed in many cancers as an adjuvant therapy following surgical resection. For example, about 70-80% of the early-stage, node-negative breast cancer patients who currently receive hormonal therapy and/or adjuvant chemotherapy would have survived with only surgical treatment[35]. In colon cancer, while the benefit of adjuvant chemotherapy to stage III colon cancer patients is well established, its benefit to stage II colon cancer patients is currently at debate. Better prognostic markers will save three-fourths of the stage II colon patients who would be cured by surgery only from unnecessary sufferings accompanied by chemotherapy[36].

The 'omics technologies have offered new opportunities for developing effective biomarkers for cancer prognosis. Gene expression profiling has been particularly useful for this purpose and allowed to develop novel prognostic biomarkers in a variety of cancer types including bladder, breast, colon, gliomas, liver, lung, prostate, and stomach cancers. Among them, breast cancer has been the most actively investigated, and two multi-gene products are already in the market for the prognosis of breast cancer patients (Table 4). In 2002, researchers at Netherland Cancer Institute (NKI) reported that gene expression profiling could better predict clinical outcome of breast cancer than conventional clinical markers[35,37]. Originally, they identified 70 prognostic genes for distant metastases within 5 years among 78 breast cancer patients who were node-negative and under the age of 55 at diagnosis. Then, the 70-gene signature was validated in a larger cohort of 295 breast cancer patients, and also in many subsequent studies. It has been approved by FDA in 2007, is now undergoing a large-scale prospective clinical trial named MINDACT (Mcroarray In Node-negative Disease may Avoid ChemoTherapy), and is commercialized and marketed by Agendia[38]. A second marketed prognostic biomarker is the Oncotype DX from Genomics Health Inc, which is based on the quantitative RT-PCR assay of 21 genes identified from the analysis of 447 samples and subsequently validated in 668 ER-positive, node-negative breast cancer patients[39]. The Oncotype DX is now widely used in the US, and a prospective clinical trial named TAILORx is going on to evaluate its utility in the guidance of treatment selection. Besides the two commercialized products, many prognostic markers for breast cancer patients are being actively developed and will be available in the market soon[38].

Table 4. Marketed multigene biomarkers for breast cancer prognosis.

	Oncotype DX	MammaPrint
Manufacturer	Genomic Health, Inc.	Agendia BV
Website	http://www.genomichealth.com	http://www.agendia.com
Method	RT-PCR	Microarary
Sample condition	FFPE	Fresh frozen
Gene number	21	70
Indication	ER+; LN−	ER+; ER−; LN−
Clinical Trial	TAILORx	MINDACT
Reference	#49	#35

Predictive (or Response) Biomarkers

Predictive (or response) biomarkers are used to assess the probability that a patient will benefit from a particular treatment. Patients with breast cancer in which the gene *ERBB2* (also known as *HER2* or *NEU*)

is amplified (that is, extra copies are present) benefit from treatment with trastuzumab (Herceptin), whereas when the gene encoding the estrogen receptor is expressed by the tumour, the patients respond to treatment with tamoxifen instead[10]. Similarly, patients who have leukaemia with the *PML-RARA* translocation respond to all-*trans* retinoic acid, and those with the Philadelphia chromosome (which contains the *BCR-ABL* fusion gene) respond to imatinib mesylate (Gleevec or Glivec)[10].

Biomarkers for leukaemia have traditionally been assessed by using routine cytogenetic analysis, but additional predictive information can be gained by using genotype-based analysis. For example, in CML patients who develop resistance to imatinib mesylate, distinct mutations in the genetic region encoding the kinase domain of BCR-ABL predict differential sensitivity to the newer ABL inhibitors dasatinib and nilotinib[40]. In addition, mutations in the genetic region encoding the kinase domain of the epidermal growth-factor receptor (EGFR) predict the sensitivity of lung tumours to erlotinib or gefitinib[41,42]. Conversely, distinct mutations in *KRAS* predict that patients with lung cancer will fail to respond to these inhibitors and that patients with colon cancer will fail to respond to therapy with EGFR-specific antibody[43,44]. And, in glioblastoma multiforme, distinct mutations in the genetic region encoding the extracellular domain of EGFR predict sensitivity to EGFR inhibitors but only in cases in which the tumour-suppressor protein PTEN is also intact[45].

Functional imaging also offers the ability to detect early response by measuring molecular changes, rather than waiting for a change in tumor size. Therapeutic approaches can be tested quickly and abandoned if they do not work[46]. Using imaging to identify a subset of patients who respond to therapy can turn what would have been a failed clinical trial into a successful one for a defined cohort of patients. For example, the remarkable response of some patients with gastrointestinal stromal tumors overexpressing the *c*-KIT kinase to Gleevec can be seen within days of treatment through positron emission tomography (PET) imaging of glucose metabolism[47].

In addition to the single marker-based response prediction to therapy, multiple, genomics-based markers have been developed. For example, Nevins group has developed gene expression signatures that predict sensitivity to individual chemotherapeutic drugs using *in-vitro* drug sensitivity data coupled with microarray data[48]. Many of these signatures could accurately predict clinical response in individuals treated with the commonly used cytotoxic drugs. The anticancer agents traditionally used in cancer chemotherapy often lack tumor specificity, and thus their use has been seriously hampered by the adverse side effects associated with the nonspecific action mode. The development of gene expression profiles that can predict response to these agents provides new opportunities to better use them, including using in combination with existing targeted therapies.

With the targeted therapy forming a strong trend in drug development, more and more predictive biomarkers are expected to be developed in the coming years. Good biomarkers could greatly accelerate new drug development by shortening clinical trials, identifying responsive patients and revealing toxic side effects.

Pharmacodynamic Biomarkers

Pharmacodynamic biomarkers measure the near-term treatment effects of a drug on the tumour (or on the host) and can, in theory, be used to guide dose selection in the early stages of clinical development of a new anticancer drug. There is a rapidly growing body of evidence linking genetic polymorphisms with functional changes in proteins that are responsible for the metabolism and disposition of many medications. Likewise, polymorphisms in genes encoding the targets of medications (e.g. receptors) can alter the pharmacodynamics of the drug response by changing receptor sensitivity[49,50]. These genetic determinants of drug effects remain stable over a patient's lifetime, and thus only need to be measured once.

The clinical and commercial values of pharmacodynamic biomarkers can be typically exemplified by the case of warfarin[51]. Warfarin is a drug modifying the blood clotting process that is prescribed to prevent and treat thromboembolism following myocardial infarction, atrial fibrillation, stroke, venous thrombosis and various surgeries. It is one of the most widely prescribed drugs in the world but it is difficult to determine an effective warfarin dose for a patient since there is a 20-fold variation in dose requirements for therapeutic clotting times. The consequences of improper dosing are serious. Hemorrhage during warfarin therapy is a leading cause of death in Western countries and related adverse events account for 1 in 10 hospital admissions[52]. Getting initial dosing right is usually a labor-intensive and costly process but knowledge of a patient's genotype can significantly improve warfarin dosing and reduce warfarin-related adverse events. Variations in the cytochrome P450 genes, which are involved in the metabolism of warfarin, explain why certain people require a lower or higher dosage of warfarin to get its full benefits. Molecular diagnostic tests for cytochrome P450 genotype variation are available from a variety of clinical laboratories. The FDA is actively reviewing a relabeling

of warfarin to require pretherapy genetic diagnostic testing[53]. The warfarin example illustrates the opportunities for applying pharmacogenomic knowledge to better target an existing therapy, thereby improving patient outcomes and lowering the cost of care.

In cancers, one of the best-studied examples in pharmacogenetics is the genetic polymorphism of thiopurine methyltransferase (*TPMT*), which catalyses the S-methylation of the thiopurines azathioprine, mercaptopurine and thioguanine[54,55]. These agents are commonly used in the treatment of leukemia as well as rheumatic diseases, inflammatory bowel diseases, and solid organ transplantation. Thiopurines are inactive prodrugs that require metabolic conversion to thioguanine nucleotides (TGN) to exert their effects, or they are inactivated via either xanthine oxidase or TPMT. In hematopoietic tissues, TGN is mainly inactivated by TPMT pathway, and therefore, patients who inherit *TPMT* deficiency accumulate excessive TGN concentrations with standard doses of these medications. TPMT activity exhibits genetic polymorphism in populations, ~90% of individuals inheriting high enzyme activity whereas 10% have immediate activity because of heterozygosity and 0.3% have low or no detectable enzyme activity because they inherit two non-functional *TPMT* alleles[56,57]. It is well known that TPMT-deficient patients are at a high risk for severe, and sometimes fatal, hematological toxicity, and patients who are *TPMT* heterozygotes have an intermediate risk of hematological toxicity[58,59]. Patients who inherit two mutant alleles should be started on 6-10% of the standard dose of thiopurines. Heterozygous patients starting with full doses are very likely to require a dose reduction to avoid toxicity. The molecular basis for polymorphic TPMT activity has now been defined, with three mutant alleles (*TPMT*2*, *TPMT*3A*, or *TPMT*3C*) accounting for TPMT deficiency in $>95\%$ of patients[60-63], and *TPMT* genotyping is now available as a Clinical Laboratory Improvement Act (CLIA)-certified molecular diagnostic from reference laboratories (e.g. Prometheus, San Diego, CA).

The dose of cytotoxic chemotherapy that is used to determine antitumour activity in phase II clinical trials is usually the maximum tolerated dose discovered in a phase I dose escalation study. But for the drugs that have been optimized to bind to a specific molecular target, this might be a less relevant end point. For example, imatinib mesylate has been shown to block the protein-kinase activity of BCR-ABL in the tumour cells of CML patients at the same doses that induce clinical remission, which are well below those associated with toxicity. Therefore, an alternative way to determine an appropriate dose is to measure the impact of the drug on its target across a range of doses (known as a target engagement study) and then to select a dose for phase II clinical trials on the basis of the magnitude of target modulation. The utility of pharmacodynamic biomarkers might also extend beyond the clinical trial phase of drug development. Recently, the magnitude of BCR-ABL kinase activity inhibition was found to correlate with clinical outcome, possibly justifying the personalized selection of drug dose based on the results of target engagement assays[64].

It has been proposed that during Phase II clinical trials patients would be genotyped, and the genetic polymorphisms associated with a favorable response to therapy and/or toxicity risk identified. Phase III studies might then involve only those who are likely to achieve the desired response with a low risk of toxicity based on their genetic profile[65]. This would probably result in smaller and shorter Phase III clinical trials, thus reducing the costs of drug discovery and resulting in faster drug approval. Over the past decade, there has been an explosion of new knowledge about genetic polymorphisms that are responsible for inherited differences in drug efficacy and toxicity. Thus, it has been predicted that someday pharmacogenomics will inevitably be an important part of the drug development process[66]. Thus it is expected that in the future, molecular diagnostics will be approved side-by-side with the drug for which they are targeted.

Future Perspectives of Biomarker-based Cancer Diagnostics

Diagnostics have traditionally been associated with low margins, tricky reimbursement issues and difficulty of market penetration. However, with changes to healthcare economics and regulation and the adoption of companion diagnostic tests that are predictive for drug response, this trend is expected to change. Jane's market report predicts that the use of biomarkers would be widespread for clinical purposes in oncology in 5-10 years[1]. Frequently, there would be a companion diagnostic for therapeutics. Also, there will be an increase in the use of panels consisting of multiple biomarkers, e.g., DNA plus protein, RNA plus protein, and it will be facilitated by new instrumentation that allows two types of molecules to be detected simultaneously. The report also anticipates that biotechnology companies will independently develop biomarker-based diagnostic tests to predict response to drugs. Such a development may be sponsored by organizations other than the pharmaceutical

industry, e.g., payers who wish to control/rationalize expenditures. All these developments will facilitate the implementation of personalized medicine that would require collaboration of patients, physicians, diagnostic and pharmaceutical companies, academic institutions, payers, and regulatory authorities.

Organized consortia such as the Biomarker Alliance[67], consisting of expertise specializing in biomarker discovery and validation, pharmacogenomic services, *in-vivo* imaging and clinical implementation of biomarkers, may play an important role in this development. In addition to the development-oriented, closed local activities such as the Biomarker Alliance, open and global movement such as Human Variome Project[68], organized by scientists specializing in molecular genetics, genomics, molecular pathology, clinical genetics, etc, is also expected to make huge contributions to the expansion and completion of biomarker-based personalized medicine, in the long run, by providing well-catalogued information for human genetic variations in association with clinical phenotypes.

Analysis of diagnostics market by Batchelder and Miller in 2006 indicated that innovative technology is already driving down the cost of biomarker discovery, and companies that succeed both in enhancing the robustness and accuracy of tests and in marrying them with specific treatments to identify appropriate subpopulations of responders (or non-responders) will be in a good position to capitalize on the increasing demand of payers to move away from costly treatments as well as on the greater use of IVDs by physicians to enable more precise clinical decision making[51].

Although diagnostics are clearly less attractive than therapeutics in terms of the potential returns, this sector does have several advantages. First, there is no requirement in the development of diagnostics for the lengthy clinical trials process that a drug company must pursue to obtain marketing approval. Second, several different paths to commercialization are available for diagnostic manufacturers and their backers. Third, because of the lower regulatory hurdles, the time to market is also shorter. Fourth and most important of all, diagnostics have been central, and are poised to become increasingly critical components of healthcare provision. This is because as increasing number of targeted therapeutics are introduced into the market, the demand for diagnostic tests is likely to boom, especially in oncology field, which can identify the condition and facilitate prescription of the right drug for a particular disease and which can divide a particular clinical population into patients who are likely to respond or not respond to a particular therapy. Thus, the market growth for diagnostics as companions to treatments is predicted to only increase[51]. In summary, it is believed that ongoing changes in technology, regulation, adoption of companion diagnostics for drugs, reimbursement and marketing are all combining to make the diagnostics field much more interesting and attractive to all the players involved.

Acknowledgements

This work was supported by the grant for the 21C Frontier Human Genome Functional Analysis Project from Korea Ministry of Science & Technology and by a grant from Korea Food and Drug Administration.

References

1. Jain, K.K. Biomarkers. In: *A JAIN PHARMABIOTECH REPORT*. Basel. 2008.
2. Jones, H.B. On a new substance occurring in the urine of a patient with mollities ossium. *Phil. Trans. R. Soc. Lond.* **138**, 55-62 (1848).
3. Laude, J.S., Wroblewski, F. & Karmen, A. Serum glutamic oxaloacetic transaminase activity in human acute transmural myocardial infarction. *Science* **120**, 497-499 (1954).
4. Rosalki, S.B. An improved procedure for serum creatine phosphokinase determination. *J. Lab. Clin. Med.* **69**, 696-705 (1967).
5. Moore, T.L., Kantrowitz, P.A. & Zamcheck, N. Demonstration of an antigen common to several varieties of neoplasia. *N. Engl. J. Med.* **285**, 138-141 (1971).
6. Baker, M. New-wave diagnostics. *Nat. Biotech.* **24**, 931-938 (2006).
7. Etzioni, R. *et al*. The case for early detection. *Nature Rev. Cancer* **3**, 1-10 (2003).
8. Pepe, M.S. *et al*. Phases of biomarker development for early detection of cancer. *J. Natl. Cancer Inst.* **93**, 1054-1061 (2001).
9. Gann, P.H., Hennekens, C.H. & Stampfer, M. A prospective evaluation of plasma prostate-specific antigen for detection of prostatic cancer. *JAMA* **273**, 289-294 (1995).
10. Sawyers, C.L. The cancer biomarker problem. *Nature* **452**, 548-552 (2008).
11. Bennett, I.C., Gattas, M. & The, B.T. The management of familial breast cancer. *Breast* **9**, 247-263 (2000).
12. Ford, D. *et al*. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am. J. Hum. Genet.* **62**, 676-689 (1998).
13. Nicolaides, N.C. *et al*. Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. *Nature*

371, 75-80 (1994).
14. Reichman, R.C. Human Papillomaviruses. In: Braunwald, E. *et al*. (eds.) *HARRISON'S PRINCIPLES of INTERNAL MEDICINE*, 15th ed. New York: McGraw -Hill. 1118-1120 (2001).
15. Bosch, F.X. & de Sanjose, S. Human papillomavirus and cervical cancer-burden and assessment of causality. *J. Natl. Cancer Inst. Monographs* **31**, 3-13 (2003).
16. Walboomers, J.M. *et al*. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J. Pathol.* **189**, 12-19 (1999).
17. Schiffman, M.H. & Castle, P. Epidemiologic studies of a necessary causal risk factor: human papillomavirus infection and cervical neoplasia. *J. Natl. Cancer Inst.* **95**, E2-E2 (2003).
18. Josefsson, A.M. *et al*. Viral load of human papilloma virus 16 as a determinant for development of cervical carcinoma *in situ*: a nested case-control study. *Lancet* **355**, 2189-2193 (2000).
19. Apple, R.J. *et al*. HLA DR-DQ associations with cervical carcinoma show papillomavirus-type specificity. *Nat. Genet.* **6**, 157-62 (1994).
20. Kize, M.G., Salmon, C.P., Pais, P. & Felton, J.S. Food heating and the formation of heterocyclic aromatic amine and polycyclic aromatic hydrocarbon mutagens/carcinogens. *Adv. Exp. Med. Biol.* **459**, 179-193 (1999).
21. Chen, Y.-C. & Hunter, D.J. Molecular epidemiology of cancer. *CA Cancer J. Clin.* **55**, 45-54 (2005).
22. Chen, J. *et al*. A prospective study of N-acetyltransferase genotype, red meat intake, and risk of colorectal cancer. *Cancer Res.* **58**, 3307-3311 (1998).
23. Spinney, L. Caught in time. *Nature* **442**, 736-738 (2006).
24. Winawer, S.J. *et al*. Prevention of colorectal cancer: guidelines based on new data. WHO Collaborating Center for the Prevention of Colorectal Cancer. *Bull World Health Organ* **73**, 7-10 (1995).
25. Bell, R., Petticrew, M. & Sheldon, T. The performance of screening tests for ovarian cancer: results of a systematic review. *Br. J. Obst. Gynaecol.* **105**, 1136-1147 (1998).
26. Etzioni, R. *et al*. Overdiagnosis due to prostate-specific antigen screening: lessons from US prostate cancer incidence trends. *J. Natl. Cancer Inst.* **94**, 981-990 (2002).
27. Petricoin, E.F. *et al*. Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* **359**, 572-577 (2002).
28. Dhanasekaran, S.M. *et al*. Delineation of prognostic biomarkers in prostate cancer. *Nature* **412**, 822-826 (2001).
29. Diehl, F. *et al*. Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proc. Natl. Acad. Sci. USA* **102**, 16368-16373 (2005).
30. Jeronimo, C. *et al*. Quantitation of GSTP1 methylation in non-neoplastic prostatic tissue and organ-confined prostate adenocarcinoma. *J. Natl. Cancer Inst.* **93**, 1747-1752 (2001).
31. Gao, W.M. *et al*. Distinctive serum protein profiles involving abundant proteins in lung cancer patients based upon antibody microarray analysis. *BMC Cancer* **5**, 110 (2005).
32. Bonadonna, G. *et al*. Combination chemotherapy as an adjuvant treatment in operable breast cancer. *N. Engl. J. Med.* **294**, 405-410 (1976).
33. Beer, D.G. *et al*. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat. Med.* **8**, 816-824 (2002).
34. Robbins, R.J. *et al*. Real-time prognosis for metastatic thyroid carcinoma based on 2-[18F]fluoro-2-deoxy-D-glucose-positron emission tomography scanning. *J. Clin. Endocrinol. Metab.* **91**, 498-505 (2006).
35. van't Veer, L.J. *et al*. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530-536 (2002).
36. Johnston, P.G. Stage II colorectal cancer: to treat or not to treat. *The Oncologist* **10**, 332-334 (2005).
37. van de Vijver, M.J. *et al*. A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* **347**, 1999-2009 (2002).
38. Ross, J., Hatzis, C., Symmans, R., Pusztai, L. & Hortobagyi, G.N. Commercialized multigene predictors of clinical outcome for breast cancer, *Oncologist* **13**, 477-493 (2008).
39. Paik, S. *et al*. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med.* **351**, 2817-2826 (2004).
40. Shah, N.P. *et al*. Sequential ABL kinase inhibitor therapy selects for compound drug-resistant BCR-ABL mutations with altered oncogenic potency. *J. Clin. Invest.* **117**, 2562-2569 (2007).
41. Lynch, T.J. *et al*. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **350**, 2129-2139 (2004).
42. Sharma, S.V., Bell, D.W., Settleman, J. & Haber, D.A. Epidermal growth factor receptor mutations in lung cancer. *Nature Rev. Cancer* **7**, 169-181 (2007).
43. Pao, W. *et al*. KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLoS. Med.* **2**, e17, doi:10.1371/journal.pmed.0020017 (2005).
44. Khambata-Ford, S. *et al*. Expression of epiregulin and amphiregulin and *K-ras* mutation status predict disease control in metastatic colorectal cancer patients treated with cetuximab. *J. Clin. Oncol.* **25**, 3230-3237 (2007).
45. Mellinghoff, I.K. *et al*. Molecular determinants of the response of glioblastomas to EGFR kinase inhibitors. *N. Engl. J. Med.* **353**, 2012-2024 (2005).
46. Weber, W.A. Use of PET for monitoring cancer therapy and for predicting outcome. *J. Nucl. Med.* **46**, 983-995 (2005).
47. Stroobants, S. *et al*. 18FDG-Positron emission tomography for the early prediction of response in advanced soft tissue sarcoma treated with imatinib mesylate

(Glivec). *Eur. J. Cancer.* **39**, 2012-2020 (2003).
48. Potti, A. *et al*. Genomic signatures to guide the use of chemotherapeutics. *Nat. Med.* **12**, 1294-300 (2006).
49. Evans, W.E. & Johnson, J.A. Pharmacogenomics: the inherited basis for interindividual differences in drug response. *Annu. Rev. Genomics Hum. Genet.* **2**, 9-39 (2001).
50. Evans, W.E. & Relling, M.V. Pharmacogenomics: translating functional genomics into rational therapeutics. *Science* **286**, 487-491 (1999).
51. Batchelder, K. & Miller, P. A change in the market-investing in diagnostics. *Nat. Biotech.* **24**, 922-926 (2006).
52. Pirmohamed, M. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *Br. Med. J.* **329**, 15-19 (2004).
53. http://www.fda.gov/ohrms/dockets/ac/05/minutes/2005-4194M1.pdf
54. Krynetski, E.Y. & Evans, W.E. Pharmacogenetics of cancer therapy: getting personal. *Am. J. Hum. Genet.* **63**, 11-16 (1998).
55. McLeod, H.L. *et al*. Genetic polymorphism of thiopurine methyltransferase and its clinical relevance for childhood acute lymphoblastic leukemia. *Leukemia* **14**, 567-572 (2000).
56. Otterness, D. *et al*. Human thiopurine methyltransferase pharmacogenetics: gene sequence polymorphisms. *Clin. Pharmacol. Ther.* **62**, 60-73 (1997).
57. Yates, C.R. *et al*. Molecular diagnosis of thiopurine S-methyltransferase deficiency: genetic basis for azathioprine and mercaptopurine intolerance. *Ann. Intern. Med.* **126**, 608-614 (1997).
58. Black, A.J. *et al*. Thiopurine methyltransferase genotype predicts therapy-limiting severe toxicity from azathioprine. *Ann. Intern. Med.* **129**, 716-718 (1998).
59. Relling, M.V. *et al*. Mercaptopurine therapy intolerance and heterozygosity at the thiopurine S-methyltransferase gene locus. *J. Natl. Cancer Inst.* **91**, 2001-2008 (1999).
60. Otterness, D. *et al*. Human thiopurine methyltransferase pharmacogenetics: gene sequence polymorphisms. *Clin. Pharmacol. Ther.* **62**, 60-73 (1997).
61. Yates, C.R. *et al*. Molecular diagnosis of thiopurine S-methyltransferase deficiency: genetic basis for azathioprine and mercaptopurine intolerance. *Ann Intern. Med.* **126**, 608-614 (1997).
62. Krynetski, E.Y. *et al*. A single point mutation leading to loss of catalytic activity in human thiopurine S-methyltransferase. *Proc. Natl. Acad. Sci. USA* **92**, 949-953 (1995).
63. Tai, H.L. *et al*. Thiopurine S-methyltransferase deficiency: two nucleotide transitions define the most prevalent mutant allele associated with loss of catalytic activity in Caucasians. *Am. J. Hum. Genet.* **58**, 694-702 (1996).
64. Shah, N.P. *et al*. Potent transient inhibition of BCR-ABL by dasatinib leads to complete cytogenetic remissions in patients with chronic myeloid leukemia: implications for patient management and drug development. *Blood* **108**, abstr. 2166 (2006).
65. Roses, A.D. Pharmacogenetics and the practice of medicine. *Nature* **405**, 857-865 (2000).
66. Johnson, J.A. & Evans, W.E. Molecular diagnostics as a predictive tool: genetics of drug efficacy and toxicity. *Trends Mol. Med.* **8**, 300-305 (2002).
67. http://www.biomarkeralliance.com.
68. http://www.humanvariomeproject.org.