Biochimica et Biophysica Acta xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Review

Inconvenient truth: Cancer biomarker development by using proteomics ☆

Tadashi Kondo *

Division of Pharmacoproteomics, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan

ARTICLE INFO

Article history:
Received 4 April 2013
Received in revised form 17 July 2013
Accepted 22 July 2013
Available online xxxx

Keywords:
Biomarker
Cancer
Proteomics

ABSTRACT

A biomarker is a crucial tool for measuring the progress of disease and the effects of treatment for better clinical outcomes in cancer patients. Diagnostic, predictive, and prognostic biomarkers are required in various clinical settings. The proteome, a functional translation of the genome, is considered a rich source of biomarkers; therefore, sizable time and funding have been spent in proteomics to develop biomarkers. Although significant progress has been made in technologies toward comprehensive protein expression profiling, and many biomarker candidates published, none of the reported biomarkers have proven to be beneficial for cancer patients. The present deceleration in biomarker research can be attributed to technical limitations. Additional efforts are required to further technical progress; however, there are many examples demonstrating that problems in biomarker research are not so much with the technology but in the study design. In the study of biomarkers for early diagnosis, candidates are screened and validated by comparing cases and controls of similar sample size, and the low prevalence of disease is often ignored. Although it is reasonable to take advantage of multiple rather than single biomarkers when studying diverse disease mechanisms, the annotation of individual components of reported multiple biomarkers does not often explain the variety of molecular events underlying the clinical observations. In tissue biomarker studies, the heterogeneity of disease tissues and pathological observations are often not considered, and tissues are homogenized as a whole for protein extraction. In addition to the challenge of technical limitations, the fundamental aspects of biomarker development in a disease study need to be addressed. This article is part of a Special Issue entitled: Biomarkers: A Proteomic Challenge.

1. Introduction: cancer biomarkers and proteomics

Cancer biomarkers are used as a tool to evaluate the physical and pathological status of cancer patients. There is an urgent need for diagnostic, predictive, and prognostic biomarkers to achieve a better clinical outcome [1]. Early diagnosis and treatment generally benefit most patients with malignancies. For instance, the 5-year survival rate of localized cancer of the stomach, colon, rectum, breast, uterus, and prostate can exceed 90% [2]. Early diagnosis is usually achieved by means of invasive and/or expensive examinations, such as endoscopy, computed tomography, and magnetic resonance imaging. Non-invasive, low-cost screening using plasma biomarkers for early detection and treatment may improve the clinical outcome of some cancer patients. However, such plasma biomarkers are not available for any malignancy, and patients often only visit the hospital at the progressed stage of cancer. Standard chemotherapy and radiotherapy have improved clinical outcomes for many types of malignancies. Patients at the same clinical stage of cancer present different responses to treatments, therefore, predictive biomarkers are required to optimize therapeutic strategies. For instance, in osteosarcoma, the most common primary bone malignancy, a 5-year survival rate was 20% when surgery was the only curative treatment [3]. With the development of induction and adjuvant chemotherapy protocols, the prognosis dramatically improved, and the 5-year survival rate of non-metastatic osteosarcoma patients reaches 70% if the patient responds to the induction chemotherapy. In contrast, patients who show resistance to chemotherapy have a poor prognosis, and their 5-year survival rate remains at 20% [4]. As there are variations of chemotherapy regimen, biomarkers that predict the response to standard chemotherapy may potentially lead to optimized therapeutic strategies. In the patients with stage I non-small-cell lung cancer (NSCLC), one of the most major malignancies, the surgery with curative intention is the standard treatment. However, approximately 20% of patients with stage I NSCLC have recurrence after curative surgery [5], and the prediction of such patients and adjuvant chemotherapy, which has been proven to be effective in advanced cancers will be beneficial to improve the clinical outcome. The clinical staging based on the pre-operative imaging does not always match to pathologic staging in NSCLC, and molecular biomarkers for micrometastasis and accurate staging system have been required for adequate therapy [6]. In molecular targeting therapy, the presence and type of

☆ This article is part of a Special Issue entitled: Biomarkers: A Proteomic Challenge.

* Tel.: +81 3 3542 2511x3004.
E-mail address: takondo@ncc.go.jp.

1570-9639/$ – see front matter
http://dx.doi.org/10.1016/j.bbapap.2013.07.009

mutations in target genes have been investigated for predictive biomarkers, and employed during routine examinations. For instance, in NSCLC, the presence of somatic mutation in epidermal growth factor receptor (EGFR) is an indication of EGFR-targeting small-molecule inhibitors such as gefitinib (Iressa; AstraZeneca) and erlotinib (Tarceva; OSI Pharmaceuticals, Genetech) [7–9]. Moreover, molecular targeting therapies raise the serious need for prognostic biomarkers. For instance, gastrointestinal stromal tumor (GIST), the most common mesenchymal tumor in gastrointestinal tract [10], is characterized by the mutation and overexpression of tyrosine kinase, *c-kit*, and treatment with tyrosine kinase inhibitor, imatinib mesylate (Gleevec or Glivec; Novartis), drastically reduces the tumor progression and inhibits metastasis after surgery [11–13]. As more than half of GIST patients can be cured by surgery alone, identification of patients who will have metastasis after surgery is critical [14]. Risk classification, which is based on the pathological observations, is commonly used to evaluate the malignant potentials of GIST, and the patients in the high-risk classification group are subjected to adjuvant imatinib treatments. However, the risk classification does not always correlate with metastasis, and biomarkers for more accurate prediction of prognosis have been required to complement the risk classification for risk stratification therapy in GIST. Although biomarkers may potentially benefit patients with any type of malignancy, the clinical application of biomarkers has not been achieved in entirety.

With the advent of proteomic modalities, global protein expression profiling, and biomarker development, the subject of proteomics has become very popular [15,16]. Researchers have been highly motivated by the idea that proteomics, and technologies extensively developed toward the identification of biomarker candidates, may benefit patients. Indeed, without the development of diagnostic plasma biomarkers, the technologies to measure low amounts of plasma proteins would not have been so intensively developed. In addition, progress in multidimensional chromatography separation [17], varieties of immune-depletion columns [18], and multiple reaction monitoring [19] may not have been as rapid. Clinical applications of predictive or prognostic biomarkers have encouraged early efforts for the application of laser micro-dissection in proteomics [20], as well as technologies that comprehensively examine proteins in tumor tissues and extract proteins from pathological archives [21]. The rapid examination of many clinical samples for biomarker studies facilitated the development of high throughput screening methods, such as tissue and protein microarrays [22], and antibody-based multiplex assay systems [23]. Although even without the intension of biomarker studies, present proteomics technologies may eventually become available, the idea to harness proteomics for biomarker development facilitated technology development. More than anything, large amounts of money have been invested in the field of proteomics, with the expectation that biomarkers with better clinical outcomes will be identified [24]. Although the immature proteomics technology has a risk to lead false biomarkers [25], and we have to be careful to evaluate the quality of data by novel technologies, biomarker discovery has been and will continue to be the driving force in the field of proteomics.

Besides substantial technology development and investment, unfortunately, no biomarker developed by proteomics has proven to be beneficial for cancer patients. It may simply be a matter of time before the clinical utility of biomarkers is proven; time consuming randomized control trials are required to establish the clinical utility of the biomarkers. However, needless to say biomarkers for randomized control trials, there are a few biomarkers by proteomics that clinicians want to examine in a routine medical setting. The difficulties of biomarker study are widely recognized, and a lack of the collaborative and systematic approach was pointed out [26,27]. Needless to say, biomarker development requires nation-wide collaborative studies and tremendous biological resources especially at the validation phase. In this article, several important elements in cancer biomarker discovery are discussed with concrete examples.

2. Is a technical limitation a true limitation?

In most cases, the lack of practical biomarkers identified through proteomics is thought to stem from technical limitations, such as low sensitivity, reproducibility, and throughput, and narrow dynamic range of present proteomic modalities. However, the cause of failure in biomarker studies may not be due to just technical limitations.

DNA microarrays enabled global expression studies for almost all genes by the early 21st century. It is now possible to measure the mRNA levels of thousands of genes in a quantitative and reproducible manner at relatively low cost. While several challenges remain, such as observations of individual splice variants, extremely low copy number genes, and relatively low reproducibility for genes with low expression levels, most transcripts can be observed routinely. Global expression studies of mRNA have resulted in the identification of many novel genes in carcinogenesis and during the progression of various types of cancers. The most interesting outcomes of global gene expression studies are the industrialization of 2 diagnostic kits in breast cancer, namely, Oncotype DX [28] and MammaPrint [29,30]. However, out of all the published biomarkers identified by DNA microarray, only such a few have appeared in a hospital setting, and the success rate of biomarker studies by DNA microarray is not high. Hanash pointed out that the massive investment in genomics and transcriptomics, which far exceeds any investment in proteomics, yielded a very limited number of diagnostics [26]. These examples of transcriptomic biomarker studies show us that even if every single protein species become to be measured in a reproducible and quantitative way by proteomics, these results do not always translate into useful biomarkers for clinical use.

A proteome is a more functional molecular group than a transcriptome because there are many protein features that are observed only by proteomics, such as expression level, tissue and cell localization, posttranslational modifications, protein-to-protein interactions, and degradation. All these factors are physiologically and pathologically important, and can be a source of biomarker candidates. However, dramatic improvements in biomarker studies cannot be expected just by increasing the number of observable proteins and the characteristics of proteins. In addition to the challenge of technical limitations, other problems associated with proteomic biomarker studies need to be addressed. Several fundamental problems frequently observed in proteomic biomarker studies are discussed, and the possible solutions for more effective biomarker study are proposed below.

3. Plasma biomarkers for early diagnosis: unrealistic study design

Plasma biomarkers are generally useful tools that enable early diagnosis through non-invasive examination at a low cost. Because plasma contents cannot be expected by genome and transcriptome analysis, direct observation of the plasma proteome is used to identify plasma biomarker candidates. However, plasma consists of thousands of proteins with similar physiological characteristics but different expression levels. Given the large percentage of a few proteins in the plasma proteome, it has been technically challenging to measure low protein levels in complex plasma samples quantitatively with high reproducibility. Regulatory plasma proteins, such as growth factors and existing plasma biomarkers, such as alpha fetoprotein are categorized as low level plasma proteins [31]. Moreover, the level of plasma proteins released from diseased tissues is very low [32,33]. These difficult aspects of plasma proteomics emphasize the need for novel proteomic technologies to solve these problems. However, the technical limitations may not be the only reason why we do not have the long-anticipated plasma biomarkers.

The performances of biomarkers are often evaluated by their sensitivity and specificity. The results of receiver operating characteristic (ROC) analysis demonstrate whether these data are encouraging. Indeed, without a doubt, these 2 parameters are practically important in

Please cite this article as: T. Kondo, Inconvenient truth: Cancer biomarker development by using proteomics, Biochim. Biophys. Acta (2013), http://dx.doi.org/10.1016/j.bbapap.2013.07.009