Glycoconj J (2014) 31:403–407
DOI 10.1007/s10719-014-9544-8

Strategy for development of clinically useful glyco-biomarkers

Hisashi Narimatsu

Published online: 5 September 2014

Abstract Recent advancements in proteomics technology have stimulated the widespread research and development in the area of biomarker discovery using mass spectrometry (MS). The final goal of biomarker discovery and development is to establish clinically useful and reliable diagnostic methods for various diseases. Specific alterations in the nature and composition of glycans attached to proteins are seen during the development and progression of a number of diseases and disorders. Therefore, development of glyco-biomarkers, which detect disease-specific glycoproteins and changes in glycoforms, is gaining much attention. The combined use of multiple technologies, not solely MS, is the key to the discovery of clinically significant and reliable biomarkers. We have employed the combination of quantitative real-time polymerase chain reaction (PCR), lectin microarray, liquid chromatography/mass spectrometry-based technique with isotope-coded glycosylation site-specific tagging (IGOT-LC/MS), and bioinformatics to successfully develop a novel diagnostic kit for the quantitative evaluation of liver fibrosis. Efforts to develop highly effective glyco-biomarkers for other diseases are also currently underway.

Keywords Glyco-biomarker · IGOT-LC/MS · Lectin microarray · FastLec-Hepa · Bioinformatics · Liver fibrosis

Introduction

The definitive goal of biomarker development is to establish reliable, accurate, and clinically applicable methods for the diagnosis of diseases. Structures of glycans and glyco-modifications undergo significant changes during the development and progression of diseases, including neurodegenerative diseases, cancers, and cardiovascular diseases. Therefore, glyco-biomarkers, particularly the disease-specific glycoforms of proteins, often serve as reliable indicators of specific diseases and their progression. For realization of clinically useful diagnostic markers, the first step is choosing the disease to be targeted. This is done by a thorough discussion with physicians specialized in the diseases, clarifying the problems in the target disease and the unmet need for diagnostic technologies, followed by selecting appropriate biological samples. The experiments start with a comparison between a small number (usually approximately ten each) of diseased and unaffected control samples; here, the selection of the control is very important. For using patient samples, informed consent must be obtained with patients' understanding on the aim of the study, which will be the accurate diagnosis for their welfare in the future. During the early stages, a large number of candidate molecules (glycoproteins) are identified. These are then tested in a larger pool of samples (more than 100) and ranked based on feasibility of each candidate to be used as a marker (Fig. 1). Clinically useful diagnostic methods must be highly specific and sensitive, as well as easy to perform. Antibodies have been extensively used in diagnosis, and continue to serve as excellent probes for the detection of biomarkers. However, in the absence of appropriate recombinant antibodies, alternative probes are necessary for the detection of the selected candidate molecules present in clinical samples.

Our final goal is to develop simple diagnostic kits for the detection of disease-specific glycoproteins based on a lectin-antibody sandwich assay system, consisting of a monoclonal antibody for the detection of a carrier protein and a specific lectin for the detection of the glycan structure exhibiting disease-specific changes.

H. Narimatsu (✉)
Glycomedicine Technology Research Center (GTRC), National Institute of Advanced Industrial Science and Technology (AIST), Central 2, 1-1-1, Umezono, Tsukuba, Ibaraki 305-8568, Japan
e-mail: h.narimatsu@aist.go.jp

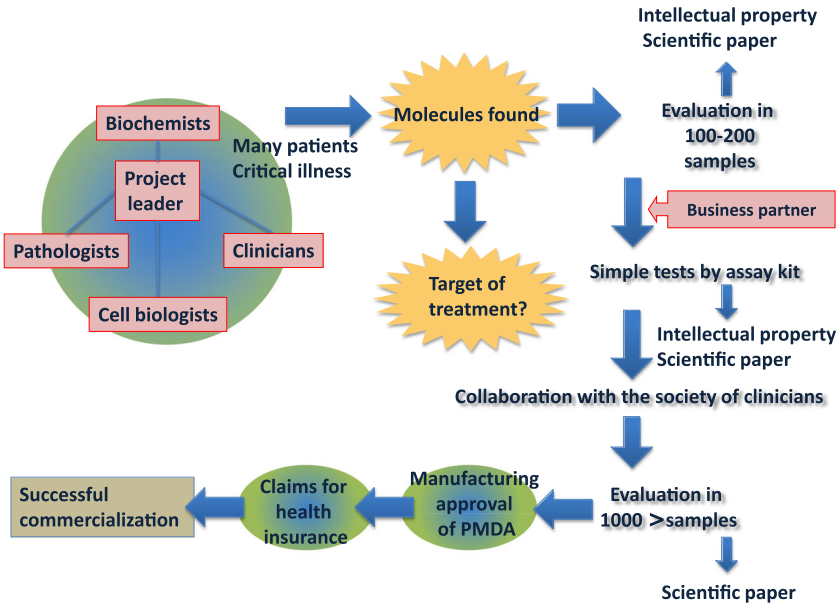

Fig. 1 A scheme for a successful biomarker development. Discovery of truly significant clinical biomarkers is supported by a strategic and rational approach and many collaborators

Methods

We previously published a strategy for the development of serum glyco-biomarkers useful for clinical diagnosis [1]. This strategy is applicable to cancers and various chronic diseases.

Selection of the target disease

For reducing the patients' burden, minimally invasive diagnostic methods that use serum and body fluids (*e.g.*, urine, tears, saliva, or nasal and vaginal discharges) should be developed. Except for immunoglobulins, most of the serum proteins are secreted from the liver. With the exception of albumin, the majority of these are glycoproteins. Therefore, we chose liver disease as the target for the verification of our strategy for the development of serum glyco-biomarkers.

According to the World Health Organization [2], more than 240 million individuals are living with hepatitis B virus (HBV) worldwide, and 150 million people are infected with hepatitis C virus (HCV). These infections often cause acute hepatitis that develop into chronic hepatitis, cirrhosis, and frequently, hepatocellular carcinoma (HCC). It has been reported that approximately 90 % of HCC patients carry either HBV or HCV. The diagnosis of acute hepatitis is made by the measurement of enzymes such as aspartate aminotransferase (AST) and alanine aminotransferase (ALT) that are released into serum following hepatic injury. As the disease progresses, the tissue becomes fibrotic and the stiffness of the liver increases. There have been no definitive serum markers for quantifying the progress of chronic hepatitis. In most cases, diagnosis is made by a histopathologist who macroscopically evaluates the extent of liver fibrosis using biopsy samples. However, liver disease patients often show bleeding tendency. Therefore, the risk of performing biopsy is high, and the patients need to be admitted to hospital. Biopsy is a highly invasive procedure, but is still recognized as the most reliable method. Recently, transient elastography (FibroScan; Echosence) has been widely used for diagnostic imaging of liver fibrosis. However, there are some weaknesses pointed out; therefore a serum-based marker for the quantitative assessment of fibrosis has been long awaited. The onset of HCC can be diagnosed by computed tomography (CT) imaging. A suspicious lesion found in the liver can be HCC or intrahepatic cholangiocarcinoma. Since the therapeutic approaches to treat these cancers are different and depend on accurate diagnosis, we also searched for a specific glyco-biomarker of cholangiocarcinoma [3].

Selection of biological samples appropriate for the target disease

As described above, the selection of the biological samples appropriate for the target disease has an indispensable effect. Surgical and biopsy samples are highly useful for the development of tissue markers. Tissue markers can be used for the evaluation of malignancy or prediction of prognosis, which are essential for physicians to determine the therapeutic strategies to be employed. Tissue specimens are relatively easy to obtain at a large volume without contamination, and thus have advantages in identification of glyco-biomarker candidates (Fig. 2). In contrast, searching for an appropriate candidate molecule in serum is very difficult at an early phase of the development because serum contains various contaminants and the contents of the target glycoproteins are generally far

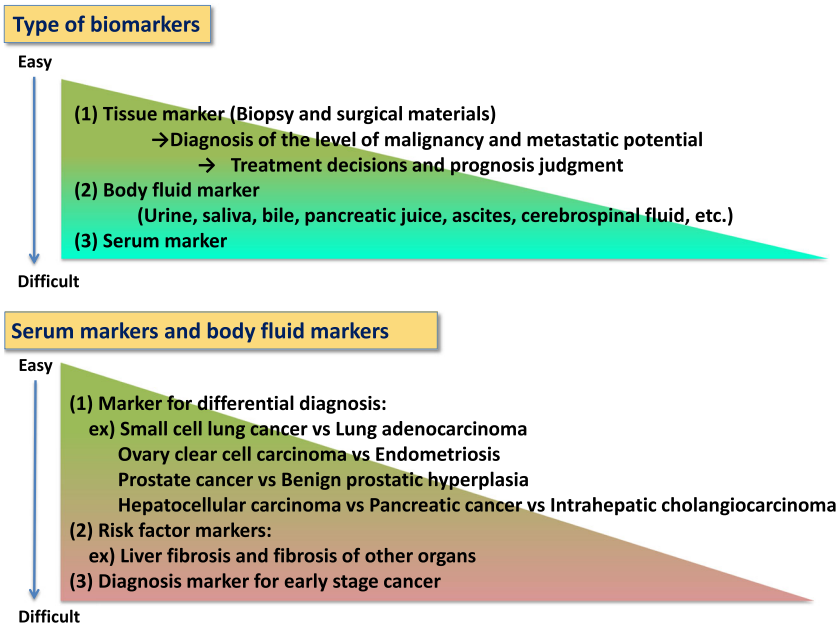

Fig. 2 Variety of cancer biomarkers of cancer. Selection of the biological samples appropriate for the target disease has an indispensable effect

lower. Therefore, we first identified the candidate glycoproteins exhibiting cancer-specific alterations in the cancerous tissues. Subsequently, we confirmed the concentrations of the candidate glycoproteins in the body fluids that contain high levels of these cancer-derived proteins. For instance, the concentrations of the candidate glycoproteins are measured in the bile and pancreatic fluids (for cholangiocarcinoma and pancreatic cancer), vaginal discharge (for endometrial carcinoma), expectoration (for lung carcinoma), and ascites fluid (for ovarian cancer). Furthermore, a simple, specific, accurate, and highly sensitive assay for the candidate protein must be established to detect the target molecule even if it presents at a trace amount diluted in the serum.

A new diagnostic system superior or complementary to the currently available methods

Some diseases are diagnosed more accurately by imaging than by serodiagnosis. Early-stage lung cancer, which currently cannot be detected using serological markers, is sensitively detected using advanced CT imaging techniques. However, CT imaging is expensive and differential diagnosis of the shadow using this technique is challenging. Whereas early diagnosis of lung cancer should be made by imaging, a feasible serum marker would be useful for the identification of not only the type of lung cancer, but also the predictors (*i.e.*, risk factors) before its onset (Fig. 3).

Ovarian cancer is latent, and the diagnostic methods are not fully advanced. Ovarian cancer is classified into four tissue types. Therefore, it is likely that there exist tissue type-specific glycoproteins and glycan structures. Among ovarian cancers, clear cell carcinoma of the ovary is one of the most commonly diagnosed forms and its prognosis is generally poor. Although CA125 is often used as an ovarian cancer marker, it is known for its poor diagnostic performance. Development of a new and accurate serum marker is therefore keenly awaited. Endometriosis has been reported as a precursor to clear cell carcinoma. Therefore, it is highly desirable to develop a new serum marker that can be used during periodical screening of patients with endometriosis to predict clear cell carcinoma.

For detecting cancers such as esophageal cancer, gastric cancer, and colon cancer, improving the sensitivity of endoscopic technologies is more realistic than developing a new serum marker. If an innovative method could reduce the medical cost, annual screening by endoscopy will contribute to the early detection, and eventually, the effective treatment of these cancers.

Although a number of specific examples can be discussed, the development goals vary largely depending on the target diseases.

Selection of clinically useful glycoproteins from the sea of marker candidates

High-throughput screening using the isotope-coded glycosylation site-specific tagging liquid chromatography/mass spectrometry (IGOT-LC/MS) method can identify several hundreds of candidate glycoproteins [1]. From this sea of glycoproteins, feasible candidates for subsequent development are selected based on bioinformatics analyses. Serum concentrations of the carrier proteins and the number of glycans attached to each molecule are the key determinants of the marker sensitivity. The expression levels of the candidate glycoproteins in a cancer lesion can be predicted from the

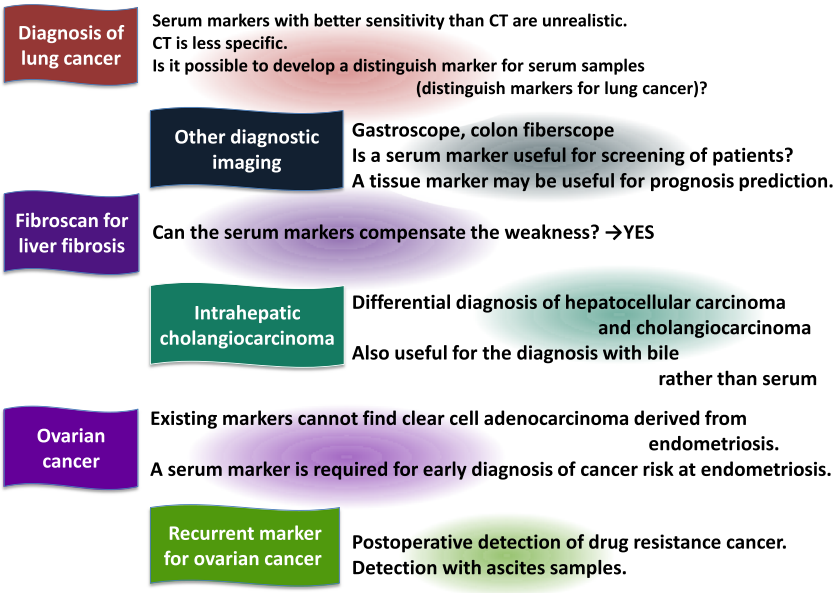

Fig. 3 Concerns in development of biomarkers useful for diagnosis. A new diagnostic system must be superior or complementary to currently available methods

reported expression levels of mRNA found in public databases. The extracellularly secreted proteins including membrane proteins are the good marker candidates. Because the production of monoclonal antibodies that recognize *N*-glycans is challenging, an alternative detection probe such as lectins must be developed to recognize cancer-specific glyco-alterations. The interaction between lectins and glycans is weaker than that between antibodies and proteins, but the greater number of binding sites compensates for such differences. For instance, Mac-2-binding protein (M2BP) has seven *N*-glycans on one molecule, and it forms a doughnut-shaped polymer consisting of 10 to 15 molecules; in total, approximately 100 *N*-glycans are clustered on the polymer [4, 5]. This allows for the formation of multiple linkages between the lectin probe and glycans, an interaction as strong as that between the protein and antibody if the binding lectin (WFA) is concentrated and immobilized in the assay system. This strong binding enables direct and highly sensitive detection of a serum marker without the need of sample pretreatment for the purification of the target glycoprotein. Therefore, we selected glycoproteins containing many glycans for further development based on the bioinformatics analyses [6, 7]. The candidates with more glycans are given higher priority.

Can the *N*-glycan structures alone be used to diagnose cancer?

It is almost impossible to detect a trace amount of a cancer-specific glycoprotein present together with abundant contaminants in serum. Attempts are being made to detect cancer-specific glyco-alterations by analyzing *N*-glycan structures using mass spectrometry (MS). However, sera collected from terminally ill cancer patients also contain various glycoproteins originating from non-cancerous tissues. Cancer progression is accompanied by various phenomena including immune responses, angiogenesis, tissue destruction, cachexia, and systemic changes. It is likely that the serum of patients with advanced stages of cancer contains abundant glycoproteins harboring cancer-related glyco-alterations. These glyco-alterations may not necessarily be originating from cancer cells. Even if an alteration originating from cancerous cells is detected, therapeutic intervention is limited at the terminal phase and the clinical usefulness of such markers is doubtful. The limitations of MS-based analyses for clinical applications are as follows: (a) glyco-alteration alone is insufficient to specifically identify cancer—the cancer-specific glyco-alterations on a specific carrier protein identify cancer; (b) a simple assay kit cannot be developed because the MS-based method requires preprocessing of serum prior to the analysis; (c) analysis requires the use of expensive equipment; and (d) a large quantity of samples are required for analysis.

Proof of the strategy: development of clinically applicable glyco-biomarkers

As mentioned above, we have developed a glyco-biomarker for assessing the progression of liver fibrosis. Because many serum glycoproteins have their origin in the liver, liver fibrosis likely affects and alters the glycan structures of many serum proteins. It is known that fucosylation is enhanced on a number of proteins with the progress of liver fibrosis [8]. However, the carrier proteins must be directly detectable without necessitating exhaustive preprocessing of serum samples. To allow for this, the probe used for detection must be a

specific lectin that binds exclusively to the target protein. Fucose-binding lectins, such as AAL and AOL, bind to abundant serum glycoproteins including immunoglobulins. Therefore, fucose-binding lectins will necessitate the preprocessing of the samples in case they were used as probes. We have once attempted to use α1-acid glycoprotein (AGP) as a fibrosis marker, but this protein also required preprocessing for purification [8, 9]. In the comparative glyco-analysis, some lectins including BPL, MAL, and WFA, had less background signals in normal serum proteins (high signal-to-noise lectins), which is a favorable characteristic of the detection probe for specific interaction [5]. We have identified a huge number of WFA-positive glycoproteins in the sera collected from HCC patients who also had liver fibrosis. Of all glycoproteins identified, those present in common in the sera of patients with normal and benign disease were eliminated. The remaining several hundreds of glycoproteins were ranked by bioinformatics analyses. The candidate proteins for which commercial antibodies were available were analyzed by western blotting and immunoprecipitation [6].

Establishment of the simple assay kit

For validation of the assay system using a large number of samples, the system needs to be simplified. Considering all aspects mentioned above, a WFA-M2BP sandwich assay system was the most appropriate. This system was superior because: (a) very few serum glycoproteins are reactive to WFA, so pretreatment of serum samples is not necessary and direct measurement is possible; (b) because the avidity of WFA-M2BP is very strong, the washing procedure does not affect the binding.

The kit was developed to allow for rapid and automated measurement using a currently available immunoassay system (HISCL-5000; Sysmex). To provide a standard, we established HEK293T cells stably expressing WFA-reactive M2BP, and purified it. We also provided an anti-M2BP antibody that showed improved binding activity. The principle of the assay is as follows. A small volume (10 μL) of serum is applied directly to the beads on which WFA is immobilized. The beads are then washed, incubated with anti-M2BP antibody, and washed again. The standard curve gave a sufficient dynamic range. This assay kit enabled automated measurement of WFA-reactive M2BP in 17 min. Such a simple and fast assay is very important in clinical diagnostics. If the serum is collected from a patient prior to the consultation, this rapid assay ensures that the data will be available for the physician's use on the same day. We designated this system as FastLec-Hepa [5].

Conclusion

WFA-reactive M2BP is currently under validation in cooperation with a long list of hospitals and doctors. The validation is being performed using several thousands of samples and is almost completed. This is a successful proof of our strategy for the development of glyco-biomarkers, which is an achievement went through the entire development phase from the basic research to the translational research in-house.

References

1. Narimatsu, H., Sawaki, H., Kuno, A., Kaji, H., Ito, H., Ikehara, Y.: A strategy for discovery of cancer glyco-biomarkers in serum using newly developed technologies for glycoproteomics. FEBS J. **277**, 95–105 (2010)
2. World Health Organization: Health Topics-Hepatitis. http://www.who.int/topics/hepatitis/en/. Accessed December 2, 2013
3. Matsuda, A., Kuno, A., Kawamoto, T., Matsuzaki, H., Irimura, T., Ikehara, Y., Zen, Y., Nakanuma, Y., Yamamoto, M., Ohkohchi, N., *et al.*: Wisteria floribunda agglutinin-positive mucin 1 is a sensitive biliary marker for human cholangiocarcinoma. Hepatology **52**, 174–182 (2010)
4. Sasaki, T., Brakebusch, C., Engel, J., Timpl, R.: Mac-2 binding protein is a cell-adhesive protein of the extracellular matrix which self-assembles into ring-like structures and binds beta1 integrins, collagens and fibronectin. EMBO J. **17**, 1606–1613 (1998)
5. Kuno, A., Ikehara, Y., Tanaka, Y., Ito, K., Matsuda, A., Sekiya, S., Hige, S., Sakamoto, M., Kage, M., Mizokami, M., *et al.*: A serum "sweet-doughnut" protein facilitates fibrosis evaluation and therapy assessment in patients with viral hepatitis. Sci. Rep. **3**, 1065 (2013)
6. Kaji, H., Ocho, M., Togayachi, A., Kuno, A., Sogabe, M., Ohkura, T., Nozaki, H., Angata, T., Chiba, Y., Ozaki, H., *et al.*: Glycoproteomic discovery of serological biomarker candidates for HCV/HBV infection-associated liver fibrosis and hepatocellular carcinoma. J. Proteome Res. **12**, 2630–2640 (2013)
7. Research Center for Medical Glycoscience, National Institute of Advanced Industrial Science and Technology. Glyco Prot DB. http://jcggdb.jp/rcmg/gpdb/index.action. Accessed December 2 2013
8. Kuno, A., Ikehara, Y., Tanaka, Y., Saito, K., Ito, K., Tsuruno, C., Nagai, S., Takahama, Y., Mizokami, M., Hirabayashi, J., *et al.*: LecT-Hepa: A triplex lectin-antibody sandwich immunoassay for estimating the progression dynamics of liver fibrosis assisted by a bedside clinical chemistry analyzer and an automated pretreatment machine. Clin. Chim. Acta **412**, 1767–1772 (2011)
9. Kuno, A., Ikehara, Y., Tanaka, Y., Angata, T., Unno, S., Sogabe, M., Ozaki, H., Ito, K., Hirabayashi, J., Mizokami, M., *et al.*: Multilectin assay for detecting fibrosis-specific glyco-alteration by means of lectin microarray. Clin. Chem. **57**, 48–56 (2011)