Lectin Capture Strategies Combined with Mass Spectrometry for the Discovery of Serum Glycoprotein Biomarkers*

Richard R. Drake‡, E. Ellen Schwegler‡, Gunjan Malik‡, Jose Diaz‡, Timothy Block§, Anand Mehta§, and O. John Semmes‡¶

The application of mass spectrometry to identify disease biomarkers in clinical fluids like serum using high throughput protein expression profiling continues to evolve as technology development, clinical study design, and bioinformatics improve. Previous protein expression profiling studies have offered needed insight into issues of technical reproducibility, instrument calibration, sample preparation, study design, and supervised bioinformatic data analysis. In this overview, new strategies to increase the utility of protein expression profiling for clinical biomarker assay development are discussed with an emphasis on utilizing differential lectin-based glycoprotein capture and targeted immunoassays. The carbohydrate binding specificities of different lectins offer a biological affinity approach that complements existing mass spectrometer capabilities and retains automated throughput options. Specific examples using serum samples from prostate cancer and hepatocellular carcinoma subjects are provided along with suggested experimental strategies for integration of lectin-based methods into clinical fluid expression profiling strategies. Our example workflow incorporates the necessity of early validation in biomarker discovery using an immunoaffinity-based targeted analytical approach that integrates well with upstream discovery technologies. *Molecular & Cellular Proteomics* 5:1957–1967, 2006.

Serum protein expression profiling using time-of-flight mass spectrometry approaches to identify biomarkers of disease has reached a nexus of technology development, clinical study design, and bioinformatics. Following a period of promising initial work using MALDI and SELDI mass spectrometry (1–4), new strategies to increase the utility of this approach for clinical biomarker assay development are needed (5, 6). Largely using simple chemical affinity beads or surfaces to decrease the sample complexities, these methods offer an automated, sensitive technique that consumes small amounts of clinical sample with relatively high throughput (2, 7–9). Concerns with the approach have included lack of analytical reproducibility, diminished robustness of discovered biomarkers during validation, lack of protein identification, and a fear that the dynamic range of prevalent proteins in serum or plasma prohibits identification of proteins associated with disease (10–12). Some of these concerns have subsequently been attributed to study design bias, chance, an overgeneralization of results, and sample processing issues (13, 14). On the other hand, when careful study design and sample handling is combined with carefully controlled instrument calibration, automated sample preparation, and supervised bioinformatic data analysis, serum expression profiling can be reproducible and portable across multiple laboratories (5, 8, 15–17). However, the remaining issues of protein dynamic range and complexity continue to plague these and indeed all proteomic approaches.

At this point in time, future expression profiling studies using clinical samples will require a careful balance of controlling for known problems while at the same time exploiting the rapid advances in robotic fractionation and mass spectrometry technologies. In this overview of emerging serum proteomic expression profiling strategies, we propose that use of lectin targeting of serum glycoprotein isoforms provides the desired experimental balance that minimizes known study design concerns while retaining the established strengths of high throughput approaches. The carbohydrate binding specificities of different lectins offer a biological affinity approach that complements existing chemical affinity methods and retains automated throughput options with current mass spectrometer capabilities. Specific examples using serum samples from prostate cancer and hepatocellular carcinoma subjects are provided along with suggested experimental strategies for integration of lectin-based methods into clinical fluid expression profiling strategies.

ALTERED GLYCOPROTEINS AND CANCER

It has long been known that cellular glycosylation profiles change significantly during oncogenesis (18–20); hence the search continues for tumor-secreted glycoproteins that can serve as biomarkers for diagnostics and/or tumor markers of the biological changes associated with the altered glycosyla-
tion patterns associated with development of cancer (21–24). One well characterized example is that of increased activity of N-acetylglucosaminyltransferase V, a key enzyme in the formation of branching asparagine-linked oligosaccharides that has been linked to tumor invasion and metastasis in multiple cancers (25–28). An increase of β-1,6-branched oligosaccharides within metastatic lymph nodes of breast carcinomas has been reported, and the presence of the branched oligosaccharides was associated with poor prognosis (28). The role of sialylated oligosaccharides was evaluated within primary breast tumors, and it was found that an overall reduction in the diversity of sialylated and neutral oligosaccharides occurred with disease progression (29, 30). It is also clear that fucosylated glycoproteins are elevated in individuals with liver, colorectal, and prostate cancers (18, 31). As discussed in a later section, we have also shown that a comprehensive characterization of differentially fucosylated serum glycoproteins can be used to readily distinguish hepatitis B-induced liver cancer subjects from healthy control subjects in blinded assays (32, 33).

Prostate-specific antigen (PSA)1 is one of the best characterized examples of a secreted glycoprotein used in cancer diagnostics, and multiple glycoforms of PSA have been described (31, 34–36). PSA is a 28,400-Da glycoprotein with one defined N-linked oligosaccharide side chain at Asn-45 and is a serine protease in the kallikrein family (kallikrein 3) (37). PSA is secreted primarily by prostatic epithelial cells into the seminal plasma where it can reach concentrations of 0.5–3 mg/ml (38). The glycoforms of PSA from seminal plasma have been shown to differ from the glycosylation of PSA secreted by the prostate metastatic tumor cell line LNCaP (31, 34, 35). A thorough comparison of PSA glycoforms from seminal plasma and serum from healthy control and prostate cancer patients has been described (36); however, larger scale studies with more clinical emphasis on study design and sample numbers are still needed. Additionally an evaluation of any differences in the glycoforms of PSA bound by carrier proteins in serum (primarily α1-antichymotrypsin) versus that of the free circulating PSA remains to be performed.

α-Fetoprotein (AFP) is another well characterized serum glycoprotein used as a surrogate marker of the presence of liver cancers, and multiple glycoforms of AFP have been identified (33, 39–41). Specific targeting of these PSA and AFP glycoforms as the diagnosis for disease state or similar characterization of other known serum components like prostate-specific membrane antigen represents a targeted proteomic strategy amenable to quantitative mass spectrometry strategies that utilize isotope tagging or mass shift labeling techniques (42, 43).

GLYCPROTEOMICS AND LECTINS

The term glycoproteomics has been used to describe this emerging branch of proteomics that focuses on characterizing the protein and carbohydrate constituents of glycoproteins. Structural elucidation of mammalian glycoproteins has long relied upon the use of lectins, a class of proteins found in plants, bacteria, fungi, and animals that are known to bind specific oligosaccharide moieties (44–47). Unlike antigen-antibody binding affinities, the affinity constants (K_D) for the binding of monosaccharides and oligosaccharides to most lectins are in the low micromolar range but can be millimolar (47, 48). For affinity capture purposes, it is the multivalent nature of both the oligosaccharides and the lectins themselves that make these interactions useful for chromatography separations (46, 47). The most common approaches for use of lectins to capture serum glycoproteins has been to digest the serum with trypsin, isolate the glycopeptides with one or more lectins linked to a support resin, elute and degradately the bound peptides with protein-N-glycanase F. The sequence and protein identities of these peptides are determined by tandem mass spectrometry (42, 49–53) or FT-ICR mass spectrometry (54–56). These approaches have been proven to be useful for identifying low concentration serum glycoproteins, but there is very low sample throughput. Strategies that probe different lectins bound on multiple array platforms (57–59) are emerging as one approach to overcome the throughput issue. These assays, along with different nanotechnology improvements (60, 61), will continue to evolve toward potential clinical assays.

In this overview, we describe our approach to the application of different lectins to enrich for serum glycoforms found in sera from prostate cancer and hepatocellular carcinoma subjects. For the lectin affinity approach, we believe this offers several distinct biochemical advantages toward protein fractionation. In addition, there are many lectin types available commercially in pure, resin-bound, biotinylated, or fluorescently labeled forms, and they are generally inexpensive. These properties lend themselves to automation in particular bead-based automated fractionation strategies as front ends to mass spectrometry (62). Many major serum proteins, in particular albumin, are not glycosylated and are therefore not bound efficiently by lectins, offering a concomitant decrease in dynamic range of serum protein targets. Therefore, the biological affinity of lectins offers multiple uses and strategies to maximize the information gained from precious clinical samples. Issues related to uniformity in preparation of different lectins, their known weak binding constants, and reproducibility of binding across large sample numbers are clear hurdles that will need to be addressed as higher throughput automation strategies are pursued. In the following paragraphs, some of these properties are demonstrated for isola-

---

1 The abbreviations used are: PSA, prostate-specific antigen; AFP, α-fetoprotein; AAL, A. auranita lectin; AAA, A. anguilla agglutinin; ConA, concanavalin A; BPH, benign prostate hyperplasia; PCa, prostate cancer; HCC, hepatocellular carcinoma; LCH, L. culinaris hemagglutinin; Fc, fucosylated; HBsAg, hepatitis B surface antigen; HBV, hepatitis B virus.
tion of serum glycoproteins, and strategies to automate these processes while maximizing the amount of information gathered from the analysis are discussed.

FRACTIONATION AND IDENTIFICATION OF LECTIN-CAPTURED SERUM GLYCOPROTEINS

To capitalize on the many glycoproteins found in serum and the known changes in glycosylation associated with cancers we have devised a modular strategy for specific targeting of glycoproteins in sera for characterization as potential biomarkers. As shown in the schematic in Fig. 1, the up-front differential lectin affinity capture module integrates well with a variety of proteomic tools and resources currently available to enrich, identify, and characterize serum proteins. As described, our approach of stratifying whole glycoproteins is in contrast to the more typical paradigm of isolating glycopeptides prior to tandem mass spectrometry identification (Fig. 1, left branch). We are attempting to retain the larger multiprotein complexes in serum via lectin capture prior to trypsin digestion and tandem mass spectrometry (Fig. 1, right branch). Thus, the information gained will be complementary to a glycopeptide-based analysis. The modularity of the lectin affinity step is that it is not dependent on any one method or mass spectrometry platform and is inherently extensible and adaptable to new technology improvements at any point in the process. It is also amenable to automation with bead-based or chip-based robotics, which is important in large scale assessment of sera from well defined clinical sets.

Initially we assessed two parameters: which lectins provided the most differential capture and whether serum depletion prior to lectin capture improved detection of glycoproteins. A panel of six lectins was used in the capture step. These included two lectins known to bind fucose residues linked to N-acetylglucosamine, *Aleuria aurantia* lectin (AAL) and *Anguilla anguilla* agglutinin (AAA); a lectin targeting 2,6-linked sialic acid residues, SNA1 (*Sambucus nigra*); and *Helix pomatia* agglutinin, which binds N-acetylgalactosamine residues. Two broad coverage lectins, wheat germ agglutinin (binds terminal N-acetylglucosamines) and concanavalin A (ConA; binds terminal mannoses and glucoses) were also included. The lectin incubations were initially done with pooled sera from matched subjects with benign prostate hyperplasia (BPH) and prostate cancer (PCa) (*n* = 5). Sera were
incubated with each lectin bound to agarose (E-Y Laboratories or Vector Laboratories) for 16 h, eluted with the appropriate target monosaccharide, and separated by SDS-PAGE as we described previously (32, 63). Representative separations of the eluted proteins from each lectin are shown in Fig. 2. It is clear from this simple gel-based analysis that use of different lectins confers selective enrichment of serum glycoproteins and some overlap and redundancy in the captured proteins. Each of these findings is consistent with the known biology of glycoproteins.

Gel slices of multiple protein bands from SDS-polyacrylamide gels, like that shown in Fig. 2, were prepared for analysis by LC-MS/MS or MALDI-TOF/TOF. A partial list of the proteins identified by this approach that are differentially expressed between BPH and PCa sera are presented in Table I, listed per lectin. Not surprisingly, these are primarily proteins found in high concentrations in serum, so that it is specifically the variant in glycosylation that is associated with disease. Such results of course require orthogonal confirmation and further validation on non-pooled sample sets. A fucosylated version of AFP is well described in sera associated with liver cancers (39, 40). Identification of an AFP variant in prostate cancer sera is a novel finding, whereas identification of the haptoglobin and apoA-I variants are also consistent with previous expression profiling studies of cancer sera (32, 64–67). These glycoprotein isoforms may represent new potential biomarkers for detection of disease, monitoring cancer treatment, or surveillance for recurrence post-therapy. Regardless of the eventual disposition of these examples as true biomarkers, this exercise demonstrates the ability to achieve disease group-specific differential lectin capture.

We next examined the effects of two serum/plasma protein depletion strategies on the types of serum glycoproteins captured by lectins. As a comparison we show a representative display of the glycoproteins that bind SNA1 lectin versus serum proteins that do not bind SNA1 lectin (Fig. 3 lanes 1 and 2). For the six lectins tested under the conditions utilized, the lack of albumin binding to the lectins has been consistent. This lack of albumin binding to lectins is a clear advantage to their use as biological affinity reagents and for up-front fractionation strategies. Indeed the lectin capture step may simultaneously concentrate classes of glycoproteins and eliminate the major blood proteins. To evaluate depletion of major blood proteins prior to lectin incubation, two kits were used with a PCa cancer sera pool: a ProteoPrep Albumin IgG depletion kit (Sigma), an immunoaffinity capture column for albumin and IgG, and a Montage albumin deplete kit (Millipore, Billerica, MA), which uses a blue dye affinity resin. Under the conditions used, both kits were very effective at removing major serum proteins (Fig. 3, lanes 8 and 9). For comparison, an aliquot of unfractionated serum is shown (Fig. 3, lane 10). The depleted protein fraction was then incubated with SNA1 lectin, and the bound (Fig. 3, lanes 4 and 6) and unbound proteins (Fig. 3, lanes 5 and 7) were separated by SDS-PAGE. Comparison of the bound protein profiles from the depleted fractions with that of the lectin alone indicate a highly similar pattern of bound proteins that is independent of depletion. Similar results were obtained with the AAL, ConA, and wheat germ agglutinin lectins (not shown). Much additional quantitative evaluation of these fractions remains to be done, but these results demonstrate that lectins can serve as excellent initial fractionation reagents for enrichment of serum glycoproteins both with and without predepletion.
Lectin Capture for Identifying Serum Glycoprotein Biomarkers

FUCOSYLATION CHANGES IN SERUM PROTEINS ASSOCIATED WITH HEPATOCELLULAR CARCINOMA (HCC) CARCINOGENESIS

The best documented change that occurs in glycosylation during the development of HCC is an increase in the level of core α-1,6-linked fucosylation of AFP (39, 40). In HCC and in testicular cancer, the glycosylation of AFP shifts from a simple biantennary glycan to an α-1,6-linked core fucosylated biantennary glycan. These changes have been observed by both direct glycan sequencing of AFP and by increased reactivity of AFP with a variety of lectins that preferentially bind to fucose-containing glycan (41). The glycoform of AFP that reacts preferentially with the lectin Lens culinaris hemagglutinin (LCH) is referred to as AFP-L3, and it has been characterized as being a more specific marker of HCC than total AFP (32, 33). A similar methodology was utilized by us in an animal model of HBV-induced HCC, and it identified a potential biomarker termed GP73 that has been shown to be 2–3 times more sensitive than AFP (32, 72).

We have examined the level of Fc-GP73 and Fc-hemopexin in a small patient cohort containing a total of 80 patients with varying degrees of liver disease (n = 20 each for healthy subjects, HBV carriers, HBV cirrhosis, and HCC). Analysis of these samples was performed for total GP73 level, for the level of Fc-GP73, and for the level of Fc-hemopexin. Total GP73 was analyzed by immunoblot using whole serum, and fucosylated species were analyzed by LCH lectin extraction of 5 µl of serum followed by immunoblotting of the fucosylated fraction as described previously (32, 33). A representative blot from a subset of these samples and the effects on the sensitivities to 100%. These results are only from a small sample set, and larger studies are ongoing, but it demonstrates that specific fucosylated isoforms of serum glycopro-
Proteins represent a rich pool of new biomarker targets for HCC and likely many other cancers.

AUTOMATION OF LECTIN CAPTURE FOR MASS SPECTROMETRY-BASED PROTEIN EXPRESSION PROFILING

One of our goals in developing the lectin capture strategies was to incorporate this into a relatively high throughput serum expression profiling platform. This would facilitate analysis of the large number of clinical samples necessary to accommodate the many disease variables associated with cancers as well as equally important epidemiological, study design, and biostatistical issues. One approach is summarized in Fig. 1. This involves capture of the intact glycoproteins by lectins followed by trypsin digestion to generate peptides and final affinity separation of peptides prior to MALDI-TOF and/or tandem mass spectrometry analyses. All of these steps can be fully automated using bead-based supports coupled to liquid sample handling robotics. Commercially available configurations to accomplish this are already available, typified by the Bruker Daltonics ClinProt system (9, 62), which we have used for our analyses. Shown in Fig. 5 is an example of this approach as applied to pooled sera from biopsy-proven early prostate cancer and benign disease with PSA values for both sets ranging from 2 to 4 ng/ml. In the left panel, a gel image of ConA-bound serum glycoproteins before and after trypsin digestion is shown. An aliquot of the undigested and digested proteins were incubated with IMAC-Cu²⁺ magnetic beads and processed robotically, and eluted peptides/proteins were spotted on a steel plate for MALDI-TOF analysis. In the right panel we show a comparison of the MALDI-TOF spectra of the two fractions with or without trypsin digestion. The increased detection of \( m/z \) peaks in the trypsin-digested samples is consistent with assessment of high molecular weight proteins that are not well resolved as whole proteins. In addition, we can observe specific differences in the spectra between PCa and BPH. On a larger scale, we propose that this approach can be used to generate spectra for expression profiling purposes and bioinformatic analyses and also serve as targets for further tandem MS analysis. In our specific example, this would be done using automated MS/MS using MALDI-TOF/TOF. However, we have also used this approach successfully when coupled with differential labeling (iTRAQ) and analysis using ESI-MS/MS. In point of fact, the lectin affinity capture of whole proteins should be amenable to most mass spectrometry-based analysis.

A PREVALIDATION STRATEGY FOR QUANTITATIVE ASSESSMENT OF PROTEIN BIOMARKER ISOFORMS

We have demonstrated that lectin-based fractionation strategies can be an excellent initial front-end step for serum glycoprotein isolation and a powerful approach toward biomarker discovery. However, establishment of initial discovery approaches should be complimented with compatible analytical confirmation and prevalidation strategies. For this purpose we have recommended targeted affinity methods sometimes referred to as immuno-MS (73, 74). The targeted candidate biomarker may be either specific isoforms of whole proteins or subunits and fragments of larger proteins. We recently demonstrated the power of immuno-MS in the characterization of an isoform of apoA-II, the overexpression of which was confirmed to be specific to BPH and/or PCa as compared with healthy controls (73). Interestingly although apoA-II is known to exist as an 8.7-kDa protein in serum, we found that a unique isoform with a mass of 8.9 kDa was associated with the disease state. The individual serum isoforms of apoA-II can be revealed by immunocapture and mass spectrometry analysis, whereas standard immunoassay techniques cannot generally discriminate isoforms. In Fig. 6 we show the specific capture of this unique isoform of serum apoA-II from disease specimens only. However, although we
could easily capture purified apoA-II, we were unable to capture significant amounts of wild-type apoA-II from serum. Thus, combining immunocapture with mass spectrometry allows for analysis of specific and altered forms of the target protein and is a useful intermediate step toward the maturation of a biomarker from discovery to utility.

Clearly such targeted proteomic strategies that utilize high affinity or baiting strategies can be used to selectively enrich lower abundance proteins. Baiting strategies can also be used to identify biomarker function and to associate the biomarker with a disease pathway through identification of other proteins with which it interacts. This technology may also bridge the gap between hypothesis- and data-driven biomarker discovery by allowing functional baiting of entire classes of biomarkers that are implicated in disease and disease progression. An affinity pre-enrichment of target proteins has two advantages: low abundance proteins can be seen and quantified, and the identity of the proteins (their primary sequence) is established, allowing subsequent isoform characterization. We have begun using a recently developed technology, isotope-differentiated binding energy shift tags (IDBEST™, Target Discovery, Inc., Palo Alto, CA) (43, 75), to quantitate specific serum isoforms, including glycoforms. IDBEST uses specific tags that exploit the natural phenomenon of mass defect. The mass defect is related to the nuclear binding energy released upon formation and stabilization of the nucleus of a given element. Biomolecules have a very negligible mass shift. A maximum mass defect value of ~0.1 amu is obtained for elements with atomic numbers between 35 (bromine) and 63 (europium). These isotope-differentiated binding energy shift tags shift the peaks of all the tagged species by about 0.1 Da, allowing software to discriminate tagged from untagged species directly in the mass spectrum and thus eliminating the need for affinity cleanup of the tagged samples.

As an example of the application of this technique to biomarker assessment we labeled a target and reference samples and then followed with immunocapture of PSA. The specifically isolated protein complexes are trypsin-digested and subjected to MALDI-TOF analysis. In Fig. 7 we show an

---

**Fig. 5.** Concanavalin A-bound serum glycoproteins digested with trypsin prior to MALDI-TOF. Pooled serum from benign prostatic hyperplasia and prostate cancer sera were bound to concanavalin A lectin and then eluted. A portion of each eluate was separated on an 8–16% SDS gel (A) either intact (lanes 1 and 2) or digested overnight with trypsin (lanes 3 and 4). B, the same eluates, trypsin-digested or intact, were incubated with IMAC-Cu²⁺ magnetic beads, eluted, and applied 1:1 with α-cyano-4-hydroxycinnamic acid matrix to a steel plate for analysis on a Bruker Daltonics Ultraflex MALDI-TOF in linear mode.
example of the specific quantitation of a PSA peptide using this approach. We are currently evaluating the quantitation accuracy of this approach by parallel sampling against clinical PSA information. We anticipate using the IDBEST technology for the quantitation of specific isoforms of known biomarkers in prostate cancer, such as PSA, as well isoforms of apolipoproteins. Specific peptides that encompass structural changes (isoforms) of whole proteins can be targeted by modification of the isolation process. Thus, for glycoforms we would pretreat with glycosidase to remove the glycan and identify the peptide via the resulting mass shift. The ability to carry early discovery through confirmation and rapid prevalence, prior to the expensive and time-consuming process of developing isoform-specific or glycospecific antibodies, should prove advantageous to accelerating biomarker discovery.

**SUMMARY**

The continued improvements in proteomic mass spectrometry technologies, coupled with the human and other genome databases, has allowed unprecedented opportunities for biomarker protein discovery and analysis of complex proteomes like serum. An underlying issue will always be the quality of the starting material as this will ultimately dictate the quality of the proteomic data and utility of this for clinical purposes. In this regard, sample collection, storage and quality issues, epidemiological input, and study design biases will always influence clinical proteomic studies. Building from what was learned in the first wave of serum protein expression profiling studies, strategies striking a balance of sample throughput with improved depth of serum protein capture and protein identification need to continue to evolve. Accommodating these concerns into proteomic analysis design is facilitated by use of the lectins as front-end fractionation and enrichment tools. We have only described the use of six lectins, and there are dozens of other individual lectin types available commercially that remain to be empirically assessed and hundreds more described in the literature (21, 45). Additionally serial affinity capture strategies in which different lectins are used in tandem will increase the fractionation capabilities beyond the discussed examples. The use of a biological affinity approach targeting known glycosylation changes associated with cancer is an additional benefit. Coupling the glycoprotein characterizations to some type of simultaneous or complementary glycan analysis of the same samples will further extend the utility of this approach. Identification of altered carbohydrate content, whether it is sialic

![Fig. 6. Expression profile of apoA-II isoforms in control and prostate cancer sera. Sera were processed on IMAC-Cu²⁺ ProteinChips for SELDI-TOF MS as described previously (73). The mass region from 8000–9500 m/z is shown for purified apoA-II (top spectrum), sera from a prostate cancer patient (middle spectrum), and sera from a healthy normal patient (bottom spectrum).](image)
Lectin Capture for Identifying Serum Glycoprotein Biomarkers

Fig. 7. Quantitative immuno-MS assay for prostate-specific antigen. Purified prostate-specific antigen was labeled with a lysine-reactive IDBEST mass defect reagent (Target Discovery, Inc.), digested with trypsin, and analyzed on a Bruker Daltonics UltraFlex™ MALDI-TOF/TOF system. The raw spectrum (top panel) and mass defect spectrum (bottom panel) around the region of one of the expected PSA peptides are shown.

acid or fucose residue differences, automatically implicates the corresponding glycosyltransferases or (glycosidases) as potential participants in the oncogenesis of a particular cancer subtype. These enzyme classes can readily be assessed by other methods than proteomics, including gene or tissue microarrays, and basic biochemical enzymatic assays.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Dept. of Microbiology and Molecular Cell Biology, Lewis Hall 3144, Eastern Virginia Medical School, Norfolk, VA 23507. Tel.: 757-446-5904; Fax: 757-446-5766; E-mail: semmesoj@evms.edu.

REFERENCES

1. Adam, B. L., Qu, Y., Davis, J. W., Ward, M. D., Clements, M. A., Cazares, L. H., Semmes, O. J., Schellhammer, P. F., Yasui, Y., Fang, Z., and Wright, G. L., Jr. (2002) Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. Cancer Res. 62, 3609–3614

2. Petricoin, E. F., Ardekani, A. M., Hitt, B. A., Levine, P. J., Fusaro, V. A., Steinberg, S. M., Mills, G. B., Simone, C., Fishman, D. A., Kohn, E. C., and Liotta, L. A. (2002) Use of proteomic patterns in serum to identify ovarian cancer. Lancet 359, 572–577

3. Campa, M. J., Wang, M. Z., Howard, B., Fitzgerald, M. C., and Patz, E. F., Jr. (2003) Protein expression profiling identifies macrophage migration inhibitory factor and cyclophilin a as potential molecular targets in non-small cell lung cancer. Cancer Res. 63, 1652–1656

4. Yanagisawa, K., Shyr, Y., Xu, B. J., Massion, P. P., Larsen, P. H., White, B. C., Roberts, J. R., Edgerton, M., Gonzalez, A., Nadaf, S., Moore, J. H., Caprioli, R. M., and Carbone, D. P. (2003) Proteomic patterns of tumour subsets in non-small-cell lung cancer. Lancet 362, 433–439

5. Semmes, O. J. (2005) The “omics” haystack: defining sources of sample bias in expression profiling. Clin. Chem. 51, 1571–1572

6. Semmes, O. J., Malik, G., and Ward, M. (2006) Application of mass spectroscopy to the discovery of biomarkers for detection of prostate cancer. J. Cell. Biochem. 98, 496–503

7. Wright G. L., Jr. (2002) SELDI proteinchip MS: a platform for biomarker discovery and cancer diagnosis. Expert Rev. Mol. Diagn. 2, 549–563

8. Villanueva, J., Philip, J., Entringer, D., Chaparro, C. A., Tanwar, M. K., Holland, E. C., and Tempst, P. (2004) Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. Anal. Chem. 76, 1560–1570

9. Pusch, W., and Kostzrwa, M. (2005) Application of MALDI-TOF mass spectrometry in screening and diagnostic research. Curr. Pharm. Des. 11, 2577–2591

10. Baggerly, K. A., Morris, J. S., and Coombes, K. R. (2004) Reproducibility of SELDI-TOF protein patterns in serum: comparing datasets from different experiments. Bioinformatics 20, 777–785

11. Ransohoff, D. F. (2005) Lessons from controversy: ovarian cancer screening and serum proteomics. J. Natl. Cancer Inst. 97, 315–319

12. Sorace, J. M., and Zhan, M. (2003) A data review and re-assessment of ovarian cancer serum proteomic profiling. BMC Bioinformatics 4, 24

13. Hu, J., Coombes, K. R., Morris, J. S., and Baggerly, K. A. (2005) The importance of experimental design in proteomic mass spectrometry experiments: some cautionary tales. Brief. Funct. Genomics Proteomics 3, 322–331

14. Ransohoff, D. F. (2005) Bias as a threat to the validity of cancer molecular-marker research. Nat. Rev. Cancer 5, 142–149

15. Semmes, O. J., Feng, Z., Adam, B. L., Banez, L. L., Bigbee, W. L., Campos, D., Cazares, L. H., Chan, D. W., Grizzle, W. E., Izbioka, E., Kagan, J., Malik, G., McLennan, D., Moul, J. W., Parlin, A., Prasanna, P., Rosenzweig, J., Sokol, L. J., Srivastava, S., Srivastava, S., Thompson, I., Welsh, M. J., White, N., Winget, M., Yasui, Y., Zhang, Z., and Zhu, L. (2005) Evaluation of serum protein profiling by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry for the detection of prostate cancer: I. Assessment of platform reproducibility. Clin. Chem. 51, 102–112

16. Caprioli, R. M. (2005) Deciphering protein molecular signatures in cancer tissues to aid in diagnosis, prognosis, and therapy. Cancer Res. 65, 10642–10645

17. Li, J., Orlandi, R., White, C. N., Rosenzweig, J., Zhao, J., Seregni, E., Morelli, D., Yu, Y., Meng, X. Y., Zhang, Z., Davidson, N. E., Fung, E. T., and Chan, D. W. (2005) Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. Clin. Chem. 48, 1296–1304

18. Hakomori, S. (1998) Tumor malignancy defined by aberrant glycosylation and sphingolipid metabolism. Cancer Res. 56, 5309–5318

19. Kobata, A. (1998) A retrospective and prospective view of glycopathology. Glycoconjug. J. 15, 323–331

20. Butler, M., Quehls, D., Critchley, A. J., Carchon, H., Hebestreit, H. F., Hibbert, R. G., Vilarinho, L., Teles, E., Mattheis, G., Schollen, E., Argibay, P., Harvey, D. J., Dwek, R. A., Jaeken, J., and Rudd, P. M. (2003) Detailed glycan analysis of serum glycoproteins of patients with congenital disorders of glycosylation indicates the specific defective glycan processing step and provides an insight into pathogenesis. Glycobiochemistry 13, 601–622

21. Sharon, N., and Lis, H. (2004) History of lectins: from hemagglutinins to biological recognition molecules. Glycoconjug. J. 14, 569–576

Molecular & Cellular Proteomics 5.10 1965
Lectin Capture for Identifying Serum Glycoprotein Biomarkers

23. Dennis, J. W., Granovsky, M., and Warren, C. E. (1999) Protein glycosylation in development and disease. BioEssays 21, 412–421.

24. Omtoft, T. F., and Vestergaard, E. M. (1999) Clinical aspects of altered glycosylation of glycoproteins in cancer. Electrophoresis 20, 362–371.

25. Fernandes, B., Sagman, U., Auger, M., Demetrio, M., and Dennis, J. W. (1991) β1,6 branched oligosaccharides as a marker of tumor progression in human breast and colon neoplasia. Cancer Res. 51, 718–723.

26. Yanagi, M., Aoyagi, Y., Suda, T., Mita, Y., and Asakura, H. (2001) N-Acetylgalactosaminyltransferase V as a possible aid for the evaluation of tumor invasiveness in patients with hepatocellular carcinoma. J. Gastroenterol. Hepatol. 16, 1282–1289.

27. Siddiqui, S. F., Pawelek, J., Handerson, T., Lin, C. Y., Dickson, R. B., Rimm, D. L., and Camp, R. L. (2005) Coexpression of β1,6-N-acetylgalactosamini- lyltransferase V glycoprotein substrates defines aggressive breast can- cers with poor outcome. Cancer Epidemiol. Biomark. Prev. 14, 2517–2523.

28. Handerson, T., Camp, R. Harigopal, M., Rimm, D., and Pawelek, J. (2005) β1,6-Branched oligosaccharides are increased in lymph node metastases and predict poor outcome in breast carcinoma. Clin. Cancer Res. 11, 2969–2973.

29. Dwek, M. V., Lacey, H. A., and Leatham, A. J. (1998) Breast cancer progression is associated with a reduction in the diversity of sialylated and neutral oligosaccharides. Clin. Chim. Acta 271, 191–202.

30. Dwek, M. V., Ross, H. A., and Leatham, A. J. (2001) Proteome and glyco- sylation mapping identifies post-translational modifications associated with aggressive breast cancer. Proteomics 1, 756–762.

31. Peracaula, R., Tabares, G., Royle, L., Harvey, D. J., Dwek, R. A., Rudd, P. M., and de Llorens, R. (2003) Altered glycosylation pattern allows the distinction between prostate-specific antigen (PSA) from normal and tumor origins. Glycobiology 13, 457–470.

32. Block, T. M., Comunale, M. A., Lowman, M., Steel, L. F., Romano, P. R., Fimmel, C., Tennant, B. C., London, W. T., Evans, A. A., Blumberg, B. S., Dwek, R. A., Mattu, T. S., and Mehta, A. S. (2005) Use of targeted glycoproteomics to identify serum glycoproteins that correlate with liver cancer in woodchucks and humans. Proc. Natl. Acad. Sci. U. S. A. 102, 7109–7114.

33. Comunale, M. A., Lowman, M., Long, R. E., Krauver, J., Philip, R., Seeholzer, S., Evans, A. A., Hann, H. W., Block, T. M., and Mehta, A. S. (2006) Proteomic analysis of serum associated fucosylated glycoproteins in the development of primary hepatocellular carcinoma. J. Proteome Res. 5, 308–315.

34. Prakash, S., and Robbins, P. W. (2000) Glycotyping of prostate specific antigen. Glycobiology 10, 173–175.

35. Ohyama, C., Hosono, M., Nitta, K., Koseki, K., Habuchi, T., Arai, Y., and Fukuda, M. (2004) carbohydrate structure and differential binding of prostate specific antigen to Maackia amurensis lectin between tumor invasiveness in patients with hepatocellular carcinoma and metastatic liver tumor. Trends Endocrinol. Metab. 19, 1149–1156.

36. Tabares, G., Radcliffe, C. M., Barrabes, S., Ramirez, M., Aleixandre, R. N., Horizonte, J., de Llorens, R., and de Llorens, R. (2006) Different glycan structures in prostate-specific antigen from pro- state cancer sera in relation to seminal plasma PSA. Glycobiology 16, 132–145.

37. Ward, M. A., Catto, J. W. F., and Hamdy, F. C. (2001) Prostate-specific antigen: biology, biochemistry and available commercial assays. Ann. Clin. Biochem. 38, 633–651.

38. Diamandis, E. (1998) Prostate-specific antigen: biology, biochemistry and available commercial assays. Ann. Clin. Biochem. 38, 633–651.

39. Ohyama, C., Hosono, M., Nitta, K., Oh-eda, M., Yoshikawa, K., Habuchi, T., Arai, Y., and Fukuda, M. (2004) Carbohydrate structure and differential binding of prostate specific antigen to Maackia amurensis lectin between tumor invasiveness in patients with hepatocellular carcinoma and metastatic liver tumor. Trends Endocrinol. Metab. 19, 1149–1156.

40. Tabares, G., Radcliffe, C. M., Barrabes, S., Ramirez, M., Aleixandre, R. N., Horizonte, J., de Llorens, R., and de Llorens, R. (2006) Different glycan structures in prostate-specific antigen from prostate cancer sera in relation to seminal plasma PSA. Glycobiology 16, 132–145.
glycosylated serum from hepatitis B carriers reveals polypeptides that correlate with disease status. *Proteomics* **4**, 826–838

64. Zhang, Z., Bast, R. C., Jr., Yu, Y., Li, J., Sokoll, L. J., Rai, A. J., Rosenzweig, J. M., Cameron, B., Wang, Y. Y., Meng, X. Y., Berchuck, A., Van Haften-Day, C., Hacker, N. F., de Bruijn, H. W., van der Zee, A. G., Jacobs, I. J., Fung, E. T., and Chan, D. W. (2004) Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res.* **64**, 5882–5890

65. Semmes, O. J., Cazares, L. H., Ward, M. D., Qi, L., Moody, M., Maloney, E., Morris, J., Trosset, M. W., Hisada, M., Gygi, S., and Jacobson, S. (2005) Discrete serum protein signatures discriminate between human retrovirus-associated hematologic and neurologic disease. *Leukemia* **19**, 1229–1238

66. Wilson, N. L., Schulz, B. L., Karlsson, N. G., and Packer, N. H. (2002) Sequential analysis of N- and O-linked glycosylation of 2D-PAGE separated glycoproteins. *J. Proteome Res.* **1**, 521–529

67. Jin, Y., and Manabe, T. (2005) Direct targeting of human plasma for matrix-assisted laser desorption/ionization and analysis of plasma proteins by time of flight-mass spectrometry. *Electrophoresis* **26**, 2823–2834

68. Shiraki, K., Takase, K., Tameda, Y., Hamada, M., Kosaka, Y., and Nakano, T., (1999) A clinical study of lectin-reactive α-fetoprotein as an early indicator of hepatocellular carcinoma in the follow-up of cirrhotic patients. *Hepatology* **22**, 802–807

69. Taketa, K., Sekiya, C., Namiki, M., Akamatsu, K., Ohta, Y., Endo, Y., and Kosaka, K. (1990) Lectin-reactive profiles of α-fetoprotein characterizing hepatocellular carcinoma and related conditions. *Gastroenterology* **99**, 508–518

70. Taketa, K., Ichikawa, E., Taga, H., and Hirai, H. (1985) Antibody-affinity blotting, a sensitive technique for the detection of α-fetoprotein separated by lectin affinity electrophoresis in agarose gels. *Electrophoresis* **6**, 492–497

71. Callewaert, N., Van Vlierberghe, H., Van Hecke, A., Laroy, W., Delanghe, J., and Contreras, R. (2004) Noninvasive diagnosis of liver cirrhosis using DNA sequencer-based total serum protein glycomics. *Nat. Med.* **10**, 429–434

72. Marrero, J. A., Romano, P. R., Nikolaeva, O., Steel, L., Mehta, A., Fimmel, C. J., Comunale, M. A., D’Amelio, A., Lok, A. S., and Block, T. M. (2005) GP73, a resident Golgi glycoprotein, is a novel serum marker for Hepatocellular Carcinoma. *J. Hepatol.* **43**, 1007–1012

73. Malik, G., Ward, M. D., Gupta, S. K., Trosset, M. W., Grizzle, W. E., Adam, B. L., Diaz, J. I., and Semmes, O. J. (2005) Serum levels of an isoform of apolipoprotein A-II as a potential marker for prostate cancer. *Clin. Cancer Res.* **11**, 1073–1085

74. Xiao, Z., Adam, B. L., Cazares, L. H., Clements, M. A., Davis, J. W., Schellhammer, P. F., Dalmasso, E. A., and Wright, G. L., Jr. (2001) Quantitation of serum prostate-specific membrane antigen by a novel protein biochip immunoassay discriminates benign from malignant prostate disease. *Cancer Res.* **61**, 6029–6033

75. Hall, M. P., Ashrafí, S., Obegi, I., Petersch, R., Peterson, J. N., and Schneider, L. V. (2003) “Mass defect” tags for biomolecular mass spectrometry. *J. Mass Spectrom.* **38**, 809–816