Published in final edited form as:

Cancer J. 2011 ; 17(6): 423–428. doi:10.1097/PPO.0b013e3182383cab.

Application of proteomics to cancer early detection

Sam Hanash and Ayumu Taguchi

Fred Hutchinson Cancer Research Center

Abstract

Strategies to achieve personalized medicine and improve public health encompass assessment of an individual's risk for disease, early detection, and molecular classification of disease resulting in an informed choice of the most appropriate treatment instituted at an early stage of disease development. An unmet need in this field for which proteomics is well suited to make a major contribution is the development of blood based tests for early cancer detection. This is illustrated in proteomic studies of epithelial cancer that encompass analysis of specimens collected both at the time of diagnosis and specimens collected before onset of symptoms that are particularly suited for the identification of early detection markers. This overarching effort benefits from the availability of plasmas from subject cohorts and of engineered mouse models that are sampled at early stages of tumor development. Integration of findings from plasma with tumor tissue and cancer cell proteomic and genomic data allows elucidation of signatures in plasma for altered signaling pathways. The discovery and further development of early detection markers takes advantage of the availability of in-depth quantitative proteomics methods and bioinformatics resources for data mining.

Introduction

While the concept of blood based cancer tests is simple, the development of such tests has been challenging, to the point that very few cancer-related biomarker assays have been FDA approved in recent years. The development of biomarker panels for early cancer detection is particularly challenging given the need for sensitivity and specificity tests at a time when tumors are in early development. The difficulty in identifying markers with suitable performance for early detection is illustrated in the recent findings from an ovarian cancer protein marker validation study.¹⁻³ Numerous studies in hundreds of publications have reported the identification of protein biomarker candidates that showed promise for the early detection of ovarian cancer. To further assess the performance of candidate markers, a recent validation study encompassed 28 ovarian cancer protein candidates that were selected for assay in pre-diagnostic serum specimens from ovarian cancer cases and controls from the Prostate, Lung, Colorectal, and Ovarian Cancer (PLCO) screening trial. None of the individual markers exhibited performance characteristics that matched those of CA-125 the current best marker for ovarian cancer. Even when biomarkers were combined into panels based on predefined models, there was little improvement in performance compared with CA-125 alone. In another study of proteomic biomarker candidates together with CA 125, a sensitivity of 84% at a specificity of 98% was observed in an analysis of sera from patients who had stage I disease at the time of surgery relative to controls which was significantly better than CA 125 alone, leading to a validation study in pre-diagnostic sera from the

PLCO trial. The findings also from this study indicated that the panel did not augment the performance of CA125 alone.⁴ The negative results observed in these validation studies highlight the challenges involved in the development biomarkers with sufficient performance to be suitable for early cancer detection. They have engendered a negative outlook on biomarkers⁵ necessitating a reassessment of strategies to discover and develop cancer biomarkers. Such reassessment is relevant not only to protein biomarkers but also more broadly to various types of markers.

Here we present strategies for finding proteins biomarkers and emphasize the need for rigor in experimental design, the critical importance of the choice of biospecimens for analysis and the need to apply technologies with the pre-requisite sensitivity and quantitative accuracy to identify and quantify low abundance proteins that signal the presence of cancer at an early stage. We present findings from the application of these strategies to solid tumors, and in particular to lung cancer.

The need for blood based tests to detect cancer

The rich content of blood in thousands of proteins that have the potential to inform about the health status of an individual provides an ideal compartment to develop non-invasive diagnostics for cancer. Protein markers currently in clinical use which include CA125 for ovarian cancer, CA19-9 for pancreatic cancer, CEA for colon cancer and PSA for prostate cancer have limited utility for cancer detection.⁶ Other common cancers notably breast and lung lack established blood based biomarkers with demonstrated clinical utility in a screening setting. Imaging modalities have utility for detecting cancer early. However they have substantial limitations. In the case of colorectal cancer (CRC), while current screening methods have had an impact on mortality associated with this disease,⁷ it is estimated that ~60% of subjects over the age of 50 in the United States are not screened at recommended intervals.⁸⁻⁹ Even when subjects are referred by their physician for colonoscopy, there is only a ~50% rate of adherence.¹⁰ In the case of lung cancer, imaging based screening to detect lung cancer has shown promise as demonstrated in a recently completed NCI-initiated national lung screening.¹¹ However, the thin slices provided by multi-detector row scanners have led to the detection of large numbers of small pulmonary nodules that turn out not to be non-malignant. The frequency of detecting non-calcified nodules on a single CT varies from 5-60% in a lung cancer screening population.¹²⁻¹³ Given the high probability of false positive findings associated with CT screening for lung cancer, there is a substantial need for additional modalities to complement imaging. Blood based biomarkers provide for simple non-invasive tests for screening as well as disease classification and monitoring for cancer progression and regression.

Innovative methodologies for cancer protein biomarker discovery

An intense effort is ongoing to identify circulating protein markers for common cancers using a multitude of strategies (Table 1).¹⁴ This effort has engendered a large number of candidate protein markers with potential diagnostic utility for many cancer types as exemplified for lung cancer.¹⁵⁻¹⁶ In one recent study circulating levels of 47 biomarkers were evaluated against patient cohorts consisting of 90 NSCLC and 43 high-risk controls

using immunoassays.¹⁷ Multivariate statistical methods then identified a panel of six biomarkers (tumor necrosis factor- α , CYFRA 21-1, interleukin-1ra, matrix metalloproteinase-2, monocyte chemotactic protein-1 and sE-selectin) as being the most efficacious for diagnosing early stage NSCLC. These markers were tested as a panel against a second patient cohort with correct classification of 75 of 88 patients tested, pointing to the potential of novel markers to contribute to the development of a panel that has utility for early detection. While currently a large number of candidate markers have been reported in the literature that may have utility for cancer early detection, critical validation studies of utility based on prospective cohorts have been either lacking or have largely been negative as presented for ovarian cancer.

In contrast with studies that were based on a small number of pre-selected candidates for discovery of cancer markers, proteomics allows an unbiased comprehensive search for biomarkers in biospecimens. The development of mass spectrometry for protein identification, in particular electrospray ionization¹⁴ coupled with various pre-fractionation and separation schemes and protein labeling, has allowed quantitative analysis of an ever increasing number of proteins from cells, tissues and biological fluids (Figure 1). As a result, the identification of a protein product for virtually all expressed genes in a cancer cell population has become achievable. Likewise, nearly exhaustive identification of proteins associated with particular cell compartments such as proteins expressed on the cell surface or are secreted or otherwise released into the extracellular space, as well as proteins with particular post-translational modifications, is currently feasible.¹⁸⁻²¹ Also for plasma and other biological fluids, mass spectrometry has sufficiently advanced to allow quantitative proteome profiling across no less than seven logs of protein abundance.²² In parallel with developments in the field of mass spectrometry for proteomics, microarray based strategies with spotted affinity capture agents have provided a complementary approach to interrogate proteomes.²³ These technologies have been applied to tissues, biological fluids, and cell populations to identify potential cancer markers.

Illustrative of the diversity in the type of specimens used for biomarker discovery, is a strategy for lung cancer biomarker discovery based on comparative proteomic profiling of pulmonary venous effluent draining the tumor vascular bed and of systemic arterial blood obtained from the same subjects. The concept is that effluent blood contains higher concentrations of potential biomarkers compared to more distal blood. A candidate biomarker (NAP-2/ CXCL7) was identified that potentially has utility for detecting lung cancer up to 30 months prior to clinical diagnosis.²⁴

Post-translational modifications, notably glycosylation, are an important source of cancer biomarkers.²⁵ Strategies to identify glycosylated proteins associated with cancer have involved both mass spectrometry and microarray based methodology and strategies for integrating proteomics and glycomics are under development.²⁶ An antibody-lectin sandwich array method was utilized to characterize both the protein and glycan levels of specific mucins and carcinoembryonic antigen-related proteins from the sera of pancreatic cancer patients and control subjects.²⁷ MUC1 and MUC5AC proteins showed highly prevalent and distinct glycan alterations between cases and controls. The most significant elevation was the cancer antigen 19-9 on MUC1 among cases relative to controls. A

comparison of glycoproteins isolated from the serum of healthy subjects with those from patients with lung adenocarcinoma using multilectin affinity chromatography uncovered a large number of cancer-selective proteins, which included kallikrein N plasma 1 (KLKB1) and inter- α -trypsin inhibitor heavy chain 3 (ITIH3).²⁸ In another study a glycoprotein analysis approach was undertaken using high resolution Fourier transform ion cyclotron resonance mass spectrometry to analyze glycosylated proteins present in sera obtained from ovarian cancer patients and healthy controls.²⁹ Some of the glycoproteins analyzed exhibited N-linked glycan fragments in forms that were distinct from the glycans obtained from the corresponding proteins in serum from healthy controls.

Innovative concepts in the application of proteomics for the discovery of early detection markers

a. Mouse to human strategies for the discovery of circulating protein biomarkers

An approach to overcome the complexity and heterogeneity of the human plasma proteome that represent important challenges toward the identification of protein changes associated with tumor development relies on the use of refined genetically engineered mouse (GEM) models of human cancer that faithfully recapitulate the molecular, biological, and clinical features of human disease. The merits of well-characterized GEM models have been exploited in several studies aimed at identifying protein biomarkers in cancer. Plasma from a GEM model of pancreatic cancer was sampled from mice at early and advanced stages of tumor development and from matched controls.³⁰ Using a proteomic approach based on extensive protein fractionation, 1,442 proteins were identified that were distributed across seven orders of magnitude of abundance in plasma. Analysis of proteins chosen on the basis of increased levels in plasma from tumor-bearing mice and corroborating protein or RNA expression in tissue documented concordance in the blood from newly diagnosed subjects with pancreatic cancer compared to controls. A panel of five proteins selected on the basis of their increased level at an early stage of tumor development in the mouse discriminated in a blinded study pre-diagnostic pancreatic cancer sera obtained before onset of symptoms from matched controls from the same cohort. In another study of lung cancer, plasma protein profiles of four mouse models of lung cancer were compared with profiles of models of pancreatic, ovarian, colon, prostate and breast cancer, and two models of inflammation (Figure 2).³¹ A protein signature for *Titf1/Nkx2-1*, a known lineage-survival oncogene in lung cancer was found in plasmas of mouse models of lung adenocarcinoma. An EGFR signature was found in plasma of an EGFR mutant model and a distinct plasma signature related to neuroendocrine development was uncovered in the small cell lung cancer model. The relevance to human lung cancer of the protein signatures identified based on mouse models was demonstrated based on assays of a subset of markers in plasmas obtained at the time of diagnosis of lung cancer as well as in pre-diagnostic plasmas. Given the availability of a large number of mouse models of cancer driven by manipulation of oncogenes and tumor suppressor genes relevant to the human counterparts, approaches to the discovery of biomarkers using such models is likely to be highly beneficial.

b. The immune response to tumor antigens as a source of biomarkers

An immune response to tumor antigens occurs early during tumor development and engenders an amplified signal detectable in the blood in the form of autoantibodies. There are numerous reports describing the identification of autoantibodies in sera of newly diagnosed subjects with solid tumors based on a multitude of technologies.³² A breast cancer study aimed at developing multi-autoantibody panels that could aid diagnosis tested a set of seven antigens (p53, c-myc, HER2, NY-ESO-1, BRCA1, BRCA2 and MUC1) using sera from 97 patients with breast cancer, 40 with ductal carcinoma in situ and 94 healthy controls. A panel consisting of six of the seven antigens achieved 64% sensitivity with 85% specificity for invasive breast cancer, while the sensitivity of individual autoantibodies ranged from 8% to 34%.³³ In another recent study, five clones from a phage cDNA expression library were subjected to blinded validation in a set of sera from patients with early colorectal cancer (CRC) and controls.³⁴ The classifier yielded high discrimination between cases and controls including among CRC subjects with low serum levels of carcinoembryonic antigen. A limitation of the classifier stems from its failure to distinguish between CRC patients and subjects with autoimmune disease.

Proteomic approaches have relied in part on interrogation of microarrays spotted with recombinant or with natural tumor derived proteins to identify novel antigens. Such a protein microarray approach was applied to discover novel lung cancer tumor antigens and to successfully validate the performance of a set of antigens previously found to be associated with autoantibodies in sera collected at the time of diagnosis of lung cancer.³⁵⁻³⁷ A blinded validation study using pre-diagnostic sera from the beta-Carotene and Retinol Efficacy Trial (CARET) cohort was done to determine whether annexin I, PGP9.5, 14-3-3 theta antigens and LAMR1 previously found to be targets of autoantibodies in newly diagnosed subjects with lung cancer are associated with autoantibodies in sera collected at the pre-symptomatic stage.³⁸ Individual sera collected from 85 subjects within a year prior to a diagnosis of lung cancer and 85 matched controls were assayed. Statistically significant reactivity was observed among cases relative to controls for autoantibodies tested with a similar frequency of reactivity in pre-diagnostic specimens as previously observed with samples from newly diagnosed subjects. An elegant proteomic approach for identifying tumor antigens consists of mass spectrometry analysis of peptides bound to soluble HLA molecules (sHLA).³⁹ Immunoaffinity purification of the sHLA molecules from plasma of multiple-myeloma and leukemia patients and controls identified thousands of sHLA peptides, including some cancer-related peptides, present among the sHLA peptidomes of the cancer patients.

c. Discovery studies of early detection markers based on analysis of pre-diagnostic specimens

Most studies aimed at identifying markers potentially applicable for early cancer detection, are based on specimens obtained at the time of cancer diagnosis. The molecular features and protein constituents of such specimens may not be particularly relevant to the detection of cancer at the earliest stages. The use of specimens collected at earlier time points through cohort studies has been widely implemented for risk marker studies.⁴⁰⁻⁴² An in-depth quantitative proteomic analysis of plasma comparing samples collected from estrogen

receptor-positive (ER(+)) breast cancer patients that were collected up to 17 months prior to diagnosis and from matched controls led to the identification of epidermal growth factor receptor (EGFR) as a potential predictor of breast cancer risk, leading to a validation study in an independent set of preclinical plasma samples yielding an odds ratio (OR) of 1.44; $P = 0.0008$ for women overall and an OR of 2.49; $P = 0.0001$ for current users of estrogen plus progestin (E + P) menopausal hormone therapy.⁴³ Thus the use of specimens collected at earlier time points through cohort studies for the discovery of early detection markers is feasible and may be more representative of the screening setting for early cancer detection.

Beyond technology: The need for a road map to develop protein biomarkers for early detection

Most serum and plasma biomarker studies have involved analysis of a relatively small number of samples obtained at a single institution. The challenge for the next decade will be to bring discoveries to the clinic in ways that are efficient and practical. Such an objective requires a collaborative effort that brings together scientific teams with expertise in proteomics and in various aspects of biomarker research from discovery to validation of clinical utility. Critical to this effort is the need to address some key requirements including:

1. Adequate representation of disease heterogeneity in the discovery process
2. Adequate attention to reducing sources of bias in biospecimens subjected to analysis that result in false discovery due to confounding conditions
3. Implementation of discovery and validation studies that rely on samples that are most relevant to the intended clinical application⁴⁴ and that represent independent sets.
4. Reliance on technologies that provide sufficient depth of analysis and quantitative accuracy and reproducibility
5. Integration of biomarker data from multiple independent sources
6. Implementation of a plan that envisions proceeding from discovery through various phases of validation toward the intended clinical application

References

1. Cramer DW, Bast RC Jr, Berg CD, et al. Ovarian cancer biomarker performance in prostate, lung, colorectal, and ovarian cancer screening trial specimens. *Cancer Prev Res (Phila)*. 2011; 4:365–74. [PubMed: 21372036]
2. Zhu CS, Pinsky PF, Cramer DW, et al. A framework for evaluating biomarkers for early detection: validation of biomarker panels for ovarian cancer. *Cancer Prev Res (Phila)*. 2011; 4:375–83. [PubMed: 21372037]
3. Mai PL, Wentzensen N, Greene MH. Challenges related to developing serum-based biomarkers for early ovarian cancer detection. *Cancer Prev Res (Phila)*. 2011; 4:303–6. [PubMed: 21372029]
4. Moore LE, Pfeiffer RM, Zhang Z, Lu KH, Fung ET, Bast RC Jr. Proteomic biomarkers in combination with CA 125 for detection of epithelial ovarian cancer using prediagnostic serum samples from the prostate, lung, colorectal, and ovarian (PLCO) cancer screening trial. *Cancer*. 2011
5. Buchen L. Cancer: Missing the mark. *Nature*. 2011; 471:428–32. [PubMed: 21430749]

6. Kulasingam V, Diamandis EP. Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. *Nat Clin Pract Oncol*. 2008; 5:588–99. [PubMed: 18695711]
7. Hawk ET, Levin B. Colorectal cancer prevention. *J Clin Oncol*. 2005; 23:378–91. [PubMed: 15637400]
8. Seeff LC, Manninen DL, Dong FB, et al. Is there endoscopic capacity to provide colorectal cancer screening to the unscreened population in the United States? *Gastroenterology*. 2004; 127:1661–9. [PubMed: 15578502]
9. Wee CC, McCarthy EP, Phillips RS. Factors associated with colon cancer screening: the role of patient factors and physician counseling. *Prev Med*. 2005; 41:23–9. [PubMed: 15916989]
10. Denberg TD, Melhado TV, Coombes JM, et al. Predictors of nonadherence to screening colonoscopy. *J Gen Intern Med*. 2005; 20:989–95. [PubMed: 16307622]
11. Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. 2011; 365:395–409. [PubMed: 21714641]
12. Nawa T, Nakagawa T, Kusano S, Kawasaki Y, Sugawara Y, Nakata H. Lung cancer screening using low-dose spiral CT: results of baseline and 1-year follow-up studies. *Chest*. 2002; 122:15–20. [PubMed: 12114333]
13. McWilliams AM, Mayo JR, Ahn MI, MacDonald SL, Lam SC. Lung cancer screening using multi-slice thin-section computed tomography and autofluorescence bronchoscopy. *J Thorac Oncol*. 2006; 1:61–8. [PubMed: 17409828]
14. Hanash S, Taguchi A. The grand challenge to decipher the cancer proteome. *Nat Rev Cancer*. 2010; 10:652–60. [PubMed: 20733593]
15. Sung HJ, Cho JY. Biomarkers for the lung cancer diagnosis and their advances in proteomics. *BMB Rep*. 2008; 41:615–25. [PubMed: 18823584]
16. Greenberg AK, Lee MS. Biomarkers for lung cancer: clinical uses. *Curr Opin Pulm Med*. 2007; 13:249–55. [PubMed: 17534168]
17. Farlow EC, Vercillo MS, Coon JS, et al. A multi-analyte serum test for the detection of non-small cell lung cancer. *Br J Cancer*. 2010; 103:1221–8. [PubMed: 20859284]
18. Faca VM, Hanash SM. In-depth proteomics to define the cell surface and secretome of ovarian cancer cells and processes of protein shedding. *Cancer Res*. 2009; 69:728–30. [PubMed: 19155298]
19. Zeng Z, Hincapie M, Pitteri SJ, et al. A proteomics platform combining depletion, multi-lectin affinity chromatography (M-LAC), and isoelectric focusing to study the breast cancer proteome. *Anal Chem*. 2011; 83:4845–54. [PubMed: 21513341]
20. Monetti M, Nagaraj N, Sharma K, Mann M. Large-scale phosphosite quantification in tissues by a spike-in SILAC method. *Nat Methods*. 2011; 8:655–8. [PubMed: 21743459]
21. Thakur SS, Geiger T, Chatterjee B, et al. Deep and Highly Sensitive Proteome Coverage by LC-MS/MS Without Prefractionation. *Mol Cell Proteomics*. 2011; 10:M110 003699. [PubMed: 21586754]
22. Zhang Q, Faca V, Hanash S. Mining the plasma proteome for disease applications across seven logs of protein abundance. *J Proteome Res*. 2011; 10:46–50. [PubMed: 21062094]
23. Kirby R, Cho EJ, Gehrke B, et al. Aptamer-based sensor arrays for the detection and quantitation of proteins. *Anal Chem*. 2004; 76:4066–75. [PubMed: 15253644]
24. Yee J, Sadar MD, Sin DD, et al. Connective tissue-activating peptide III: a novel blood biomarker for early lung cancer detection. *J Clin Oncol*. 2009; 27:2787–92. [PubMed: 19414677]
25. Taylor AD, Hancock WS, Hincapie M, Taniguchi N, Hanash SM. Towards an integrated proteomic and glycomic approach to finding cancer biomarkers. *Genome Med*. 2009; 1:57. [PubMed: 19519948]
26. Wang H, Wong CH, Chin A, et al. Integrated mass spectrometry based analysis of plasma glycoproteins and their glycan modifications. *Nature Protocols*. 2010 In Press.
27. Yue T, Goldstein IJ, Hollingsworth MA, Kaul K, Brand RE, Haab BB. The prevalence and nature of glycan alterations on specific proteins in pancreatic cancer patients revealed using antibody-lectin sandwich arrays. *Mol Cell Proteomics*. 2009; 8:1697–707. [PubMed: 19377061]

28. Heo SH, Lee SJ, Ryoo HM, Park JY, Cho JY. Identification of putative serum glycoprotein biomarkers for human lung adenocarcinoma by multilectin affinity chromatography and LC-MS/MS. *Proteomics*. 2007; 7:4292–302. [PubMed: 17963278]
29. Li B, An HJ, Kirmiz C, Lebrilla CB, Lam KS, Miyamoto S. Glycoproteomic analyses of ovarian cancer cell lines and sera from ovarian cancer patients show distinct glycosylation changes in individual proteins. *J Proteome Res*. 2008; 7:3776–88. [PubMed: 18642944]
30. Faca VM, Song KS, Wang H, et al. A mouse to human search for plasma proteome changes associated with pancreatic tumor development. *PLoS Med*. 2008; 5:e123. [PubMed: 18547137]
31. Taguchi A, Politi K, Pitteri S, et al. Lung Cancer Signatures in Plasma Based on Proteome Profiling of Mouse Tumor Models. *Cancer Cell*. 2011
32. Tan HT, Low J, Lim SG, Chung MC. Serum autoantibodies as biomarkers for early cancer detection. *FEBS J*. 2009; 276:6880–904. [PubMed: 19860826]
33. Chapman C, Murray A, Chakrabarti J, et al. Autoantibodies in breast cancer: their use as an aid to early diagnosis. *Ann Oncol*. 2007; 18:868–73. [PubMed: 17347129]
34. Chang WJ, Wu LL, Cao F, et al. Development of autoantibody signatures as biomarkers for early detection of colorectal carcinoma. *Clin Cancer Res*. 2011
35. Brichory F, Beer D, Le Naour F, Giordano T, Hanash S. Proteomics-based identification of protein gene product 9.5 as a tumor antigen that induces a humoral immune response in lung cancer. *Cancer Research*. 2001; 61:7908–12. [PubMed: 11691811]
36. Brichory F, Misek D, Yim A, et al. An immune response manifested by the common occurrence of annexins I and II autoantibodies and high circulating levels of IL-6 in lung cancer. *Proc natl acad sci*. 2002; 98:9824–9. [PubMed: 11504947]
37. Pereira-Faca SR, Kuick R, Puravs E, et al. Identification of 14-3-3 theta as an antigen that induces a humoral response in lung cancer. *Cancer Res*. 2007; 67:12000–6. [PubMed: 18089831]
38. Qiu J, Choi G, Li L, et al. Occurrence of autoantibodies to annexin I, 14-3-3 theta and LAMR1 in prediagnostic lung cancer sera. *J Clin Oncol*. 2008; 26:5060–6. [PubMed: 18794547]
39. Bassani-Sternberg M, Barnea E, Beer I, Avivi I, Katz T, Admon A. Soluble plasma HLA peptidome as a potential source for cancer biomarkers. *Proc Natl Acad Sci U S A*. 2010; 107:18769–76. [PubMed: 20974924]
40. Wolpin BM, Wei EK, Ng K, et al. Prediagnostic plasma folate and the risk of death in patients with colorectal cancer. *J Clin Oncol*. 2008; 26:3222–8. [PubMed: 18591557]
41. Clendenen TV, Lundin E, Zeleniuch-Jacquotte A, et al. Circulating inflammation markers and risk of epithelial ovarian cancer. *Cancer Epidemiol Biomarkers Prev*. 2011; 20:799–810. [PubMed: 21467242]
42. Douglas JB, Silverman DT, Pollak MN, Tao Y, Soliman AS, Stolzenberg-Solomon RZ. Serum IGF-I, IGF-II, IGFBP-3, and IGF-I/IGFBP-3 molar ratio and risk of pancreatic cancer in the prostate, lung, colorectal, and ovarian cancer screening trial. *Cancer Epidemiol Biomarkers Prev*. 2010; 19:2298–306. [PubMed: 20699371]
43. Pitteri SJ, Amon LM, Busald Buson T, et al. Detection of elevated plasma levels of epidermal growth factor receptor before breast cancer diagnosis among hormone therapy users. *Cancer Res*. 2010; 70:8598–606. [PubMed: 20959476]
44. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. *Journal of National Cancer Institute*. 2008; 100:1432–8.
45. Rajagopal MU, Hathout Y, MacDonald TJ, et al. Proteomic profiling of cerebrospinal fluid identifies prostaglandin D2 synthase as a putative biomarker for pediatric medulloblastoma: A pediatric brain tumor consortium study. *Proteomics*. 2011; 11:935–43. [PubMed: 21271676]
46. Roy S, Josephson SA, Fridlyand J, et al. Protein biomarker identification in the CSF of patients with CNS lymphoma. *J Clin Oncol*. 2008; 26:96–105. [PubMed: 18056677]
47. Jou YJ, Lin CD, Lai CH, et al. Proteomic identification of salivary transferrin as a biomarker for early detection of oral cancer. *Anal Chim Acta*. 2010; 681:41–8. [PubMed: 21035601]
48. de Jong EP, Xie H, Onsongo G, et al. Quantitative proteomics reveals myosin and actin as promising saliva biomarkers for distinguishing pre-malignant and malignant oral lesions. *PLoS one*. 2010; 5:e11148. [PubMed: 20567502]

49. Hu S, Arellano M, Boontheung P, et al. Salivary proteomics for oral cancer biomarker discovery. *Clin Cancer Res*. 2008; 14:6246–52. [PubMed: 18829504]
50. Dowling P, Wormald R, Meleady P, Henry M, Curran A, Clynes M. Analysis of the saliva proteome from patients with head and neck squamous cell carcinoma reveals differences in abundance levels of proteins associated with tumour progression and metastasis. *J Proteomics*. 2008; 71:168–75. [PubMed: 18617144]
51. Pawlik TM, Hawke DH, Liu Y, et al. Proteomic analysis of nipple aspirate fluid from women with early-stage breast cancer using isotope-coded affinity tags and tandem mass spectrometry reveals differential expression of vitamin D binding protein. *BMC cancer*. 2006; 6:68. [PubMed: 16542425]
52. Wang CL, Wang CI, Liao PC, et al. Discovery of retinoblastoma-associated binding protein 46 as a novel prognostic marker for distant metastasis in nonsmall cell lung cancer by combined analysis of cancer cell secretome and pleural effusion proteome. *Journal of proteome research*. 2009; 8:4428–40. [PubMed: 19655816]
53. Pernemalm M, De Petris L, Eriksson H, et al. Use of narrow-range peptide IEF to improve detection of lung adenocarcinoma markers in plasma and pleural effusion. *Proteomics*. 2009; 9:3414–24. [PubMed: 19609957]
54. Rodriguez-Pineiro AM, Blanco-Prieto S, Sanchez-Otero N, Rodriguez-Berrocal FJ, de la Cadena MP. On the identification of biomarkers for non-small cell lung cancer in serum and pleural effusion. *J Proteomics*. 2010; 73:1511–22. [PubMed: 20230924]
55. Oumeraci T, Schmidt B, Wolf T, et al. Bronchoalveolar lavage fluid of lung cancer patients: mapping the uncharted waters using proteomics technology. *Lung cancer (Amsterdam, Netherlands)*. 2011; 72:136–8.
56. Chen R, Pan S, Duan X, et al. Elevated level of anterior gradient-2 in pancreatic juice from patients with pre-malignant pancreatic neoplasia. *Mol Cancer*. 2010; 9:149. [PubMed: 20550709]
57. Elschenbroich S, Ignatchenko V, Clarke B, et al. In-depth proteomics of ovarian cancer ascites: combining shotgun proteomics and selected reaction monitoring mass spectrometry. *Journal of proteome research*. 2011; 10:2286–99. [PubMed: 21491939]
58. Amon LM, Law W, Fitzgibbon MP, et al. Integrative proteomic analysis of serum and peritoneal fluids helps identify proteins that are up-regulated in serum of women with ovarian cancer. *PloS one*. 2010; 5:e11137. [PubMed: 20559444]
59. Tian M, Cui YZ, Song GH, et al. Proteomic analysis identifies MMP-9, DJ-1 and A1BG as overexpressed proteins in pancreatic juice from pancreatic ductal adenocarcinoma patients. *BMC cancer*. 2008; 8:241. [PubMed: 18706098]
60. Chen YT, Chen CL, Chen HW, et al. Discovery of novel bladder cancer biomarkers by comparative urine proteomics using iTRAQ technology. *Journal of proteome research*. 2010; 9:5803–15. [PubMed: 20806971]
61. Feldman AS, Banyard J, Wu CL, McDougal WS, Zetter BR. Cystatin B as a tissue and urinary biomarker of bladder cancer recurrence and disease progression. *Clin Cancer Res*. 2009; 15:1024–31. [PubMed: 19188175]
62. Kreunin P, Zhao J, Rosser C, Urquidí V, Lubman DM, Goodison S. Bladder cancer associated glycoprotein signatures revealed by urinary proteomic profiling. *Journal of proteome research*. 2007; 6:2631–9. [PubMed: 17518487]
63. Theodorescu D, Schiffer E, Bauer HW, et al. Discovery and validation of urinary biomarkers for prostate cancer. *Proteomics Clin Appl*. 2008; 2:556–70. [PubMed: 19759844]
64. Ye B, Skates S, Mok SC, et al. Proteomic-based discovery and characterization of glycosylated eosinophil-derived neurotoxin and COOH-terminal osteopontin fragments for ovarian cancer in urine. *Clin Cancer Res*. 2006; 12:432–41. [PubMed: 16428483]

Figure 1. In-depth quantitative analysis by means of liquid chromatography-mass spectrometry of proteins in a pair of plasmas from a cancer case and a control

Plasmas from case and from control are subjected to isotopic labeling of proteins followed by pooling and fractionation of the pool prior to mass spectrometry analysis of digested fraction. Columns represent individual fractions; rows represent individual proteins identified in particular fractions. The color scheme is indicative of case to control concentrations ratios (red= increased, yellow= no change and green= decreased) for individual proteins based on differential isotope labeling. The ratio for an individual protein is determined based on the isotopic envelopes of case and control for a given peptide as shown for an EGFR peptide.

Figure 2. A mouse to human search for lung cancer plasma protein markers

Unsupervised hierarchical clustering of plasma proteome profile of mouse models of cancer resulted in clustering of proteomes from lung adenocarcinomas and small cell lung cancer. Validation studies of a panel of candidate markers (EGFR, SFTP8, WFDC2, and ANGPTL3) identified in mouse models of lung adenocarcinoma yielded a significant area under the curve (AUC) for the individual markers and the combined panel in plasmas from newly diagnosed subjects with lung cancer and in pre-diagnostic plasmas relative to controls. Likewise a validation study of ROBO1, a candidate marker identified in plasma from the small cell lung cancer mouse model, resulted in significantly increase levels in plasmas from subjects with small cell lung cancer relative to controls.

Table 1

Diversity of specimen sources for proteomic profiling and biomarker discovery

Specimen	Cancer Type	Biomarker
Cerebrospinal fluid	Pediatric medulloblastoma	prostaglandin D2 synthase ⁴⁵
	CNS lymphoma	antithrombin III ⁴⁶
Saliva	Oral squamous cell carcinoma	transferrin, actin, myosin, M2BP, MRP14, CD59, catalase, profilin ⁴⁷⁻⁴⁹
	Head and neck squamous cell carcinoma	S100A9 ⁵⁰
Nipple aspirate fluid	Breast cancer	vitamin D-binding protein ⁵¹
Pleural effusion	Non small cell lung cancer	RbAp46, NPC2, PEDF ⁵²⁻⁵⁴
Bronchoalveolar lavage fluid	Non small cell lung cancer	Kallikrein-8, Histatin 3, S100A12 ⁵⁵
Pulmonary venous effluent	Non small cell lung cancer	CTAP III/NAP-2 ²⁴
Pancreatic juice	Pancreatic cancer	anterior gradient-2 ⁵⁶
Peritoneal fluids/ascites	Ovarian cancer	PKM1/2, GAPDH, MSLN, IGFBP1 ⁵⁷⁻⁵⁸
	Pancreatic cancer	MMP-9, DJ-1, A1BG ⁵⁹
Urine	Bladder cancer	apolipoprotein A1, Cystatin B, α -1B-Glycoprotein ⁶⁰⁻⁶²
	Prostate cancer	Collagen alpha-1 (III) chain, Collagen alpha-1 (I) chain, Psoriasis susceptibility candidate gene 2, Sodium/potassium-transporting ATPase γ chain ⁶³
	ovarian cancer	eosinophil-derived neurotoxin ⁶⁴