Published in final edited form as:

Proteomics Clin Appl. 2008 July 30; 2(9): 1223. doi:10.1002/prca.200780073.

Prostate cancer serum biomarker discovery through proteomic analysis of alpha-2 macroglobulin protein complexes

Earle F. Burgess^{1,2}, Amy-Joan L. Ham^{3,4}, David L. Tabb^{3,4,5}, Dean Billheimer^{1,6}, Bruce J. Roth^{1,2}, Sam S. Chang^{1,7}, Michael S. Cookson^{1,7}, Timothy J. Hinton², Kristin L. Cheek^{3,4}, Salisha Hill^{3,4}, and Jennifer A. Pietenpol^{1,4}

¹Vanderbilt-Ingram Cancer Center, 777 PRB, Vanderbilt University School of Medicine, Nashville, TN, USA

²Department of Medicine, D3100 MCN, Vanderbilt University School of Medicine, Nashville, TN, USA

³Mass Spectrometry Research Center, 9144C MRBIII, Vanderbilt University School of Medicine, Nashville, TN, USA

⁴Department of Biochemistry, 652 PRB, Vanderbilt University School of Medicine, Nashville, TN, USA

⁵Department of Biomedical Informatics, 400 EBL, Vanderbilt University School of Medicine, Nashville, TN, USA

⁶Department of Biostatistics, 571 PRB, Vanderbilt University School of Medicine, Nashville, TN, USA

⁷Department of Urologic Surgery, A1302 MCN Vanderbilt University School of Medicine, Nashville, TN, USA

Abstract

Alpha-2 macroglobulin (A2M) functions as a universal protease inhibitor in serum and is capable of binding various cytokines and growth factors. In this study, we investigated if immunoaffinity enrichment and proteomic analysis of A2M protein complexes from human serum could improve detection of biologically relevant and novel candidate protein biomarkers in prostate cancer. Serum samples from six patients with androgen-independent, metastatic prostate cancer and six control patients without malignancy were analyzed by immunoaffinity enrichment of A2M protein complexes and MS identification of associated proteins. Known A2M substrates were reproducibly identified from patient serum in both cohorts, as well as proteins previously undetected in human serum. One example is heat shock protein 90 alpha (HSP90 α), which was identified only in the serum of cancer patients in this study. Using an ELISA, the presence of HSP90 α in human serum was validated on expanded test cohorts and found to exist in higher median serum concentrations in prostate cancer ($n = 18$) relative to control ($n = 13$) patients (median concentrations 50.7 *versus* 27.6 ng/mL, respectively, $p = 0.001$). Our results demonstrate the technical feasibility of this approach and support the analysis of A2M protein complexes for proteomic-based serum biomarker discovery.

Keywords

Biomarker; Heat-shock protein 90; Prostate cancer; Serum

Introduction

Great need persists for the development of improved serologic biomarkers in prostate cancer. Prostate-specific antigen (PSA) is routinely used for diagnostic and prognostic purposes but lacks adequate sensitivity and specificity in most clinical settings [1]. Use of MS for analysis of human blood has emerged as a standard methodology for protein biomarker discovery [2, 3]. Unfortunately, the convenience of sample collection is greatly overshadowed by the difficulties inherent to MS-based analysis of blood [4]. The non-cellular portion of blood is a complex mixture of proteins that varies in abundance up to ten orders of magnitude [5]. However, only 22 proteins account for 99% of serum protein mass and thus often mask detection of less abundant species [2]. To address this problem, investigators have increasingly relied upon depletion of these highly abundant proteins prior to MS-based analysis [6–10].

Although depletion strategies can improve sensitivity for detection of less abundant proteins in blood, this approach has limitations. For example, most published methods can only reduce highly abundant proteins by two orders of magnitude. Further, cytokines, growth factors and proteases are known to associate with highly abundant “carrier” proteins in blood and could be excluded from analysis by co-depletion [11–15]. These findings question whether MS-based discovery of biologically relevant proteins is sufficiently improved to justify the added expense and technical limitations associated with depletion of highly abundant blood proteins.

Alpha-2-macroglobulin (A2M) is a highly abundant plasma glycoprotein, often targeted by depletion strategies, that functions as a universal protease inhibitor [16]. Renewed interest for serologic detection of cancer-associated proteases emerged after the observation that proteolytic fragments of abundant serum proteins could have diagnostic and prognostic potential [17]. A2M covalently binds all major protease classes, thus A2M complexes in blood may contain the responsible disease-specific proteases that have thus far eluded detection. A2M can also bind various cytokines, growth factors and other nonproteolytic molecules [18,19]. Therefore, we hypothesize that targeted analysis of A2M protein complexes from blood would be a rational approach for detection of less abundant, biologically relevant candidate biomarkers in prostate cancer. To test this hypothesis, we affinity purified A2M protein complexes from the sera of patients with androgen-independent metastatic prostate cancer and similar aged controls without evidence of malignancy. Prior reports have used MS-based approaches to identify proteins associated with other highly abundant serum proteins, particularly albumin and immunoglobulins [20]. These proteins likely serve as passive carrier molecules, or a physiologic buffer in the case of albumin. A2M-associated proteins have not previously been analyzed by discovery proteomics. We chose to enrich for A2M protein complexes in this study based on the active role that A2M plays in clearance of potential biologically relevant, low-abundance proteins such as proteases. Our results demonstrate technical feasibility and inter-sample reproducibility of analyzing low-abundance A2M-associated proteins for serum biomarker discovery. We also show multiple proteins previously unreported in serum and identify HSP90 α as a novel candidate biomarker in prostate cancer.

2 Materials and methods

2.1 Patient samples

The sample collection protocol was approved by the Vanderbilt University Institutional Review Board. Serum or plasma samples were collected from patients with advanced meta-static

prostate cancer or from similarly aged control patients without evidence of malignancy. Peripheral blood was collected in a standard red top (serum) or green top (plasma) blood collection tube. For serum collection, the sample was allowed to coagulate 30 min, centrifuged at 1000×g for 10 min at 4°C. Samples were aliquoted and stored at –80°C.

2.2 Immunoprecipitation and immunoblot

Serum or plasma was thawed and diluted 1:2 with PBS. Endogenous immunoglobulins were depleted by incubation with Protein A Sepharose beads (Amersham Biosciences) for 1 h at 4°C prior to immunoprecipitation. Precleared samples were subjected to A2M immunoprecipitation by incubation with human A2M antibody Ab7337 (Abcam) and protein A Sepharose beads at 4°C. Beads were washed three times with wash buffer (150 mM NaCl, 1% NP40, 50 mM Tris pH 7.5, 1 mM EDTA). Immunoprecipitates were resolved by 10% SDS-PAGE under reducing conditions, and proteins were visualized in gel by colloidal CBB staining (Invitrogen). Immunoblot analysis was performed with A2M antibody Ab7337 (Abcam) as previously described [21].

2.3 In-gel tryptic digest

Following visualization of A2M immunoprecipitate with colloidal blue staining, gels were excised into equal regions (approximately 1 cm×1 cm×0.75 mm per slice); excluding immunoglobulin bands corresponding to heavy and light chains. Gel pieces were cut into 1-mm squares and subjected to tryptic digest similar to published methods [22–24]. Briefly, gel pieces were incubated with 100 mM ammonium bicarbonate for 15 min at room temperature. Ammonium bicarbonate was exchanged, and 45 mM DTT was added. Samples were heated at 50°C for 15 min. After cooling, 100 mM iodoacetamide was added, and samples were incubated at room temperature for 15 min in the dark. Supernatant was removed, and the samples were washed twice with ACN:50 mM ammonium bicarbonate (50:50). The solution was replaced with 100% ACN for 10 min or until gel slices turned white. Samples were desiccated and rehydrated with trypsin in 25 mM ammonium bicarbonate at a ratio of 1:100 w/w trypsin:protein. Tryptic digests were allowed to proceed at 37°C overnight. Following digestion, supernatants were collected, and peptides were extracted with 60% ACN, 0.1% formic acid in water and evaporated to dryness in a vacuum centrifuge. Peptides were resuspended in 0.1% formic acid for LC-MS analysis.

2.4 LC-MS/MS and data analysis

LC-MS analysis of the resulting peptides was performed using a ThermoFinnigan LTQ IT mass spectrometer equipped with a Thermo MicroAS autosampler and Thermo Surveyor HPLC pump, nanospray source, and Xcalibur 1.4 instrument control. The peptides were separated on a packed capillary tip, 100 mm×11 cm, with C₁₈ resin (Jupiter C₁₈, 5micron, 300 Å, Phenomenex, Torrance, CA) using an inline solid phase extraction column that was 100 mm×6 cm packed with the same C₁₈ resin, using a frit generated from liquid silicate Kasil 1 similar to that previously described[25], except the flow from the HPLC pump was split prior to the injection valve. The flow rate during the solid phase extraction phase of the gradient was 1 mL/min and during the separation phase was 700 nL/min. Mobile phase A was 0.1% formic acid, while mobile phase B was ACN with 0.1% formic acid. A 95-min gradient was performed with a 15-min washing period (100% A for the first 10 min followed by a gradient to 98% A at 15 min) to allow for SPE and removal of any residual salts. After the initial washing period, a 60-min gradient was performed where the first 35 min was performed as a slow, linear gradient from 98 to 75% A, followed by a faster gradient to 10% A at 65 min and an isocratic phase at 10% A to 75 min. MS/MS scans were acquired using an isolation width of 2 *m/z*, an activation time of 30 ms, and activation Q of 0.250 and 30% normalized collision energy using one microscan and maximum injection time of 100 ms for each scan. The mass spectrometer

was tuned prior to analysis using the synthetic peptide TpepK (AVAGKAGAR). Typical tune parameters were as follows: spray voltage of 1.8 KV, a capillary temperature of 150°C, a capillary voltage of 50 V and tube lens 100 V. Centroided MS/MS spectra of the peptides were acquired using data-dependent scanning in which one full MS spectra, using a full mass range of 400–2000 amu, was followed by three MS/MS spectra and the exclusion list was set to 250 with exclusion duration of 60 s.

The captured peak lists were transcoded to mzData v1.05 format by a version of the open-source ReAdW tool that had been modified to support mzData conversion (<http://www.mc.vanderbilt.edu/msrc/bioinformatics/index.php>). The software was configured to transcode only MS/MS spectra; MS scans were excluded. MS/MS spectra were identified to peptides from the IPI Human database version 3.22 (57 846 sequences) by the MyriMatch algorithm, version 1.0.321 [26]. The sequence database was doubled to contain each sequence in both normal and reversed orientations, enabling false discovery rate estimation (FDR). MyriMatch was configured to expect all cysteines to bear carboxamidomethyl modifications and to allow for the possibility of oxidation on methionines. Candidate peptides were required to feature trypsin cleavages or protein termini at both ends, though any number of missed cleavages was permitted. Precursor error was allowed to range up to 1.25 m/z in either direction, but fragment ions were required to match within 0.5 m/z . The IDPicker algorithm [27] filtered the identifications for each RPLCrun to include the largest set for which a 5% identification FDR could be maintained (as described in [28]), and these identifications were pooled for the gel corresponding to each patient. FDR rates were computed by this formula: $FDR = (2 * reverse)/(forward + reverse)$. Proteins were required to have at least two different peptide sequences observed among all gels to be included. Proteins that could not be distinguished based on observed peptides were given the same Protein Group ID (Supporting Information). A homology algorithm clustered proteins by shared peptide content and reported a minimal set of protein groups to explain all peptides, processing peptides from all gels in a single homology analysis. The algorithm reported the number of spectra and number of distinct sequences observed for each protein in each gel. The HTML files listing the peptide sequences identified for each protein are available in Supporting Information.

2.5 HSP90 α ELISA

HSP90 α protein concentration was quantified from patient sera using a commercially available sandwich ELISA (Assay Designs). Patient serum was diluted 1:25 or 1:50, and the assay was performed according to protocol instructions using purified HSP90 α as a positive control and for generation of a standard curve. Each patient sample was analyzed in duplicate. HSP90 α expression differences between cancer and control groups were evaluated using the Wilcoxon rank sum test. This procedure is the nonparametric counterpart to the two-sample t -test.

2.6 Statistical analysis

For each patient, the spectral counts for each protein were considered a semi-quantitative measurement of protein expression. Expression differences between cancer and control groups were evaluated using the Wilcoxon rank-sum test. This test is the nonparametric counterpart to the two-sample t -test. Each protein was evaluated separately at the 0.05 significance level. There was no correction for multiple hypothesis testing in this proof-of-concept study.

3 Results

3.1 Detection of A2M protein complexes in blood

Previous studies have reported characteristic gel shifts representing high molecular weight, covalently cross-linked A2M-substrate complexes [16,29]. The universal ‘protease trapping’ mechanism used by A2M involves a reactive thiol group that is exposed following protease-

mediated cleavage of a bait region [30]. To determine if sufficient steady-state levels of A2M-substrate complexes exist *in vivo* for detection and purification, human serum was immunoprecipitated with an A2M antibody, and protein complexes were analyzed by immunoblot. Based on the predicted molecular weight of cross-linked A2M species, substrate-bound complexes were detected in both plasma and serum (Fig. 1). Subsequent LC/MS-MS analysis confirmed the presence of A2M in the protein bands observed by immunoblot (Supporting Information Tables 1 and 2). Since serum is the non-cellular portion of clotted blood, it is noteworthy that coagulation involves a cascading series of activated proteases. A fraction of the cross-linked A2M species observed in serum may represent an artifact generated *ex vivo* by A2M complex formation with components of the clotting cascade. However, similar cross-linked A2M species were detected in plasma (Fig. 1, Lane 1); therefore, sufficient *in vivo* levels of steady-state A2M protein complexes exist that are amenable to detection and purification. We successfully immunoprecipitated both free A2M and substrate-cross-linked A2M complexes (Fig. 1).

After demonstrating the feasibility of A2M complex immunoprecipitation from human serum, we analyzed A2M-associated proteins from the serum of six patients with metastatic prostate cancer and six control patients without evidence of malignancy (median ages 64 and 68, respectively). Patient characteristics are shown in Table 1. Patients with advanced cancer were chosen as study subjects based on the presumption that levels of novel tumor-associated biomarkers would be highest in patients with advanced disease and thus most amenable to detection. A2M protein complexes were immunoprecipitated from the serum of all 12 patients and resolved by SDS-PAGE. Following gel staining, visualized proteins were uniformly excised and analyzed by LC-MS/MS. Spectra from each patient were identified and initially filtered to achieve a peptide identification FDR of 5%. Resulting peptides were assembled to proteins, with each protein requiring support from at least two unique peptide sequences from the set of 12 patients for inclusion in the master list (Supporting Information).

3.2 Identification of known A2M substrates

LC/MS-MS analysis of major high-molecular weight protein bands confirmed immunoprecipitation of A2M with 151 peptides identified through 16 517 spectra overall. Validation of the technical approach was established by identification of known A2M-associated proteins (Table 2), including components of the clotting cascade such as thrombin and the fibrinolytic pathway such as plasmin [16]. In addition, plasma kallikrein, a known A2M substrate and major effector of the kinin system, was identified [31]. Carboxypeptidase B2, a less abundant serum protease also reported to bind A2M, was detected in most samples of both cohorts as well [32].

A2M can non-covalently associate with proteins that lack proteolytic activity [18,19]. Since serum collection requires *ex vivo* activation of the clotting cascade, we expected the involved proteases to have sequestered the pool of non-covalently bound A2M and mask detection of known non-proteolytic A2M substrates. However, peptides of three different beta inhibin chains were detected in multiple patients (Table 2). Inhibin is a known A2M-associated protein that exists in the serum of prostate cancer patients and controls at approximately 90 and 30 ng/mL, respectively [33,34]. Transforming growth factor beta (TGF β) is also a known A2M-interacting protein that is present in serum of both prostate cancer and normal patients at approximately 30 ng/mL [35,36]. Overall, two different peptides for TGF β were observed, with spectra appearing in three patient samples (Supporting Information). These findings highlight the potential sensitivity of this approach to detect very low abundance, biologically relevant proteins.

3.3 Identification of unknown A2M-associated proteins

Following identification of known A2M substrates, we inter-rogated our dataset for potential unknown A2M-associated proteins. We were interested in the detection of serum proteases, given that A2M functions as a universal protease inhibitor. Analysis of serum was chosen in lieu of plasma based on evidence that some proteases may be more active in serum (following *ex vivo* zymogen activation) and thus more susceptible to A2M-mediated complex formation [37]. As presented above, many of the frequently identified proteins were highly abundant proteases activated by the *ex vivo* co-agulation process. However, these highly abundant species did not preclude detection of less abundant proteases such as ADAM DEC1, hepatocyte growth factor activator (HGFA) and Cathepsin D, the latter being present in serum at approximately 4 pmol/mL [38]. Data for these and other select proteases are shown in Table 3.

Of note, PSA, a serine protease, was not detected by analysis of serum using this approach. This finding is not unexpected, as PSA exists in blood primarily complexed with alpha 1-antichymotrypsin [39]. However, we also analyzed human plasma with this methodology and detected PSA in multiple patients with advanced prostate cancer (data not shown). This observation suggests that use of this approach for biomarker discovery may yield differential results depending on whether serum or plasma is selected for discovery efforts.

Extensive proteomic efforts to catalog serum or plasma constituents have yielded impressive results [40–45]. Our methodology identified multiple serum proteins previously unreported by these efforts (Supporting Information). Many of these proteins may represent non-covalent or even indirect A2M associations. Nonetheless, the presence of peptides representing these proteins in our dataset demonstrates the ability to detect a biologically relevant subproteome in serum with this approach that may improve future proteomics-based biomarker discovery. Full peptide and protein identification data from all patients are provided in detail in Supporting Information Tables 1–4.

3.4 Differential protein identification by cohort

Spectral counting provides a methodology to estimate relative quantitative differences in protein abundance using MS datasets [46]. Although this study was not powered to detect quantitative protein differences between cohorts, we searched for proteins that showed at least a twofold difference in total spectral counts per protein between cohorts. Selected proteins that meet this criterion are shown in Table 4. One example includes lactate dehydrogenase (LDH), a previously validated serum marker that is a component of a widely used clinical nomogram for prognostication of survival in prostate cancer [47]. LDH was detected with confidence from two patients in the prostate cancer cohort and none in the control cohort (Table 4). Larger cohort sizes would be required to establish statistically significant differences in protein spectral counts between cohorts, though our data suggest this methodology is capable of detecting relative quantitative differences in protein abundance between cohorts. This observation could greatly assist in prioritizing biomarker candidates that should be further pursued in validation efforts.

3.5 Inter-sample reproducibility of protein identification

We evaluated the reproducibility of protein identification across the six patients within both cohorts (Fig. 2). Proteins for which few peptides were present might be expected to be more variable than those for which numerous peptides were observed. As a result, we scaled the threshold for protein presence from three peptides to at least nine peptides. Correspondingly, the number of proteins listed as present decreased from 427 (>2 peptides) to 216 (>8 peptides) for the cancer patients and from 400 to 203 proteins for the control patients. The proportion of proteins common to all six patients, however, did not change substantially in response to this

increased peptide requirement. On average, 49% of the proteins were found in all six cancer patients while 50% of the proteins were found in all six control patients. The second most common class of proteins was found in only one of the six patients. This class accounted for 19% of the proteins observed in cancer patients and 16% of the proteins in control patients. Approximately two-thirds of observed proteins were either universal to all patients or idiosyncratic to a single patient.

3.6 Measurement of HSP90 α serum levels

Multiple HSP90 α peptides were detected in two patients from the cancer cohort. We chose to validate this finding given the well-established role of HSP90 in tumor biology [48]. HSP90 α serum levels were measured in expanded test cohorts of advanced prostate cancer and control patients using a sandwich ELISA (Fig. 3). We validated the presence of HSP90 α in human serum and found a higher median concentration and range of the protein in the cancer cohort ($p = 0.001$). The observed median concentration in 18 cancer patients was 50.7 ng/mL (range 25.5–378.1), while the median for 13 control subjects was 27.6 ng/mL (13.9–46.5). These data provide a separation of groups with an area under the ROC curve of 0.83 (95% confidence interval of 0.68 to 0.95). Validation of the MS-based discovery of HSP90 α in human serum and differential serum levels by cohort supports the analysis of A2M protein complexes in blood for biomarker discovery.

4 Discussion

Targeted analysis of highly abundant blood protein complexes will exclude many serum proteins from detection, but identification of specific protein classes, such as proteases, may be improved with this approach. We present an approach for identifying low-abundance, potentially disease-associated proteins in human serum by analysis of A2M protein complexes from the sera of prostate cancer patients. In our dataset, proteins with suggested spectral count differences between cohorts are associated with tumor biology but have unknown prognostic significance. For example, plastin 2 was detected with two or more peptides from all patients except one control, though more than twice as many total spectra were identified to this protein in the cancer cohort than in the control cohort. Plastin 2 is frequently up-regulated in cancer cells and can induce proliferation and invasion [49,50]. Plastin 3 was also identified and exhibited an increase in spectral counts in the cancer cohort vs. control ($p = 0.048$). E-cadherin is another well-studied protein relevant to tumor biology that has been demonstrated in the serum of advanced prostate cancer patients [51]. We detected a pair of E-cadherin peptides in two patients of the cancer cohort only, thus supporting prior observations. Proteasome activator subunits (PSME) 1 and 2 were detected with higher peptide and spectral counts in the cancer cohort as well. This finding is of particular interest given the recent validation of elevated PSME3 levels in the serum of colon cancer patients [52]. Periostin is a secreted protein that has shown acquired overexpression in breast and colon cancer and supports tumor growth by suppression of apoptosis and stimulation of angiogenesis [53,54]. Serum periostin levels may serve as a marker of metastatic skeletal involvement in patients with breast cancer and may provide prognostic insight to patients with lung cancer [55,56]. Although the occurrence of elevated serum periostin levels has not been reported for patients with prostate cancer, we detected twice as many spectra in our cancer cohort than in the control one.

Another observation often overlooked in biomarker discovery efforts is the identification of proteins that may be more abundant in the control patient cohort (Table 4). For example, thrombospondin-2, a secreted anti-angiogenic factor with tumor suppressive properties, was detected in eight spectra from the control cohort and only in two from the cancer cohort [57, 58]. Confirmation of these findings would generate new hypotheses regarding tumorigenesis and raise the possibility of detecting serum markers that may signify resistance to cancer

development. Identification of such proteins could revolutionize population-based screening strategies for cancer and other diseases of interest.

Unlike MALDI or SELDI-based serum discovery approaches, gel-LC-MS/MS-based discovery methodologies are far more resource intensive and do not facilitate high-throughput analysis. Consequently, cohort size for gel-LC-MS/MS based discovery can be limited by resource availability. Establishing quantitative spectral differences between cohorts for a particular protein of interest is limited by traditional statistical evaluation of small cohort sizes. To address this issue, we propose two criteria to guide validation efforts when using gel-LC-MS/MS-based proteomics discovery methodologies, as we did. First, preclinical evidence of biological relevance to the disease cohort may best assist candidate prioritization. In our study, we chose to pursue validation of HSP90 α , given its well-characterized role in the biology of various solid tumors [48]. Our findings demonstrate the value of incorporating biologic knowledge of the underlying disease cohort into the decision about which biomarker candidates to pursue. Secondly, proteins that demonstrate multiple identified peptides in patients of the disease cohort and no identified peptides in the control cohort may indicate a higher median serum concentration in the disease cohort. In our study, multiple HSP90 α peptides were identified in two cancer patients but not in control patients. Subsequent quantification of HSP90 α levels in these two patients demonstrated actual serum concentrations at the upper range of detection, and all other patients with no identified peptides had serum concentrations well below this level (data not shown). Therefore, HSP90 α serum levels appeared to correlate with MS spectral and peptide counts. A similar spectral pattern was observed with LDH, a previously validated prostate cancer tumor marker [47]. Criteria for how best to pick candidates for validation efforts following gel-LC-MS/MS-based discovery will continue to evolve. These proposed guidelines only reflect observations from our dataset.

Confirmation of detectable HSP90 α in human serum and of differential levels in advanced prostate cancer patients is of particular interest, given that therapeutic inhibition of HSP90 function in various solid tumors is an active area of clinical research, including advanced prostate cancer [59]. Our findings represent the first validation of soluble HSP90 α in human serum. Though best characterized as a cytosolic protein, secretion of HSP90 has been demonstrated and appears to promote cancer cell invasion through enhancement of matrix metalloproteinase 2 activity [60]. This observation poses the question of whether soluble HSP90 α levels in serum may indicate which patients are best suited for pharmacologic HSP90 inhibition. We do not yet propose to use serum HSP90 α levels to differentiate metastatic prostate cancer patients from healthy control patients. Indeed, this is not a clinically useful activity. However, confirmation of differential HSP90 α serum levels by cohort serves only to validate the discovery methodology. Further evaluation to define a clinically meaningful diagnostic threshold for serum HSP90 α levels in prostate cancer is warranted. In conclusion, the biomarker discovery approach described herein provides an alternate method for proteomic analysis of human blood that may improve candidate discovery efforts and identifies soluble HSP90 as a potential novel serum biomarker in prostate cancer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Abbreviations

A1ACT	alpha-1 antichymotrypsin
A2M	alpha-2 macroglobulin
FDR	false discovery rate

HGFA	hepatocyte growth factor activator
LDH	lactate dehydrogenase
PSA	prostate-specific antigen
PSME	proteasome activator subunit
TGF β	transforming growth factor beta

Acknowledgments

This work was supported by the Ingram Charitable Trust (J. A. Pietenpol) and National Institutes of Health Grants T32 CA009592 (E. F. Burgess), R01 CA126218-01 (D. L. Tabb), ES00267 and CA68485 (Core services).

References

1. Fall K, Garmo H, Andren O, Bill-Axelsson A, et al. Prostate-specific antigen levels as a predictor of lethal prostate cancer. *J. Natl. Cancer Inst* 2007;99:526–532. [PubMed: 17405997]
2. Veenstra TD, Conrads TP, Hood BL, Avellino AM, et al. Biomarkers: mining the biofluid proteome. *Mol. Cell. Proteomics* 2005;4:409–418. [PubMed: 15684407]
3. Rai AJ, Chan DW. Cancer proteomics: Serum diagnostics for tumor marker discovery. *Ann. NY Acad. Sci* 2004;1022:286–294.
4. Hu S, Loo JA, Wong DT. Human body fluid proteome analysis. *Proteomics* 2006;6:6326–6353. [PubMed: 17083142]
5. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics* 2002;1:845–867. [PubMed: 12488461]
6. Echan LA, Tang HY, Ali-Khan N, Lee K, Speicher DW. Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics* 2005;5:3292–3303. [PubMed: 16052620]
7. Pieper R, Su Q, Gatlin CL, Huang ST, et al. Multi-component immunoaffinity subtraction chromatography: an innovative step towards a comprehensive survey of the human plasma proteome. *Proteomics* 2003;3:422–432. [PubMed: 12687610]
8. Steel LF, Trotter MG, Nakajima PB, Mattu TS, et al. Efficient and specific removal of albumin from human serum samples. *Mol. Cell. Proteomics* 2003;2:262–270. [PubMed: 12754305]
9. Ahmed N, Barker G, Oliva K, Garfin D, et al. An approach to remove albumin for the proteomic analysis of low abundance biomarkers in human serum. *Proteomics* 2003;3:1980–1987. [PubMed: 14625860]
10. Liu T, Qian WJ, Mottaz HM, Gritsenko MA, et al. Evaluation of multiprotein immunoaffinity subtraction for plasma proteomics and candidate biomarker discovery using mass spectrometry. *Mol. Cell. Proteomics* 2006;5:2167–2174. [PubMed: 16854842]
11. Liotta LA, Ferrari M, Petricoin E. Clinical proteomics: written in blood. *Nature* 2003;425:905. [PubMed: 14586448]
12. Lopez MF, Mikulskis A, Kuzdzal S, Bennett DA, et al. High-resolution serum proteomic profiling of Alzheimer disease samples reveals disease-specific, carrier-protein-bound mass signatures. *Clin. Chem* 2005;51:1946–1954. [PubMed: 16081505]
13. Lowenthal MS, Mehta AI, Frogale K, Bandle RW, et al. Analysis of albumin-associated peptides and proteins from ovarian cancer patients. *Clin. Chem* 2005;51:1933–1945. [PubMed: 16099937]
14. Mehta AI, Ross S, Lowenthal MS, Fusaro V, et al. Biomarker amplification by serum carrier protein binding. *Dis. Markers* 2003;19:1–10. [PubMed: 14757941]
15. Granger J, Siddiqui J, Copeland S, Remick D. Albumin depletion of human plasma also removes low abundance proteins including the cytokines. *Proteomics* 2005;5:4713–4718. [PubMed: 16281180]
16. Harpel PC, Brower MS. Alpha 2-macroglobulin: an introduction. *Ann. NY Acad. Sci* 1983;421:1–9.
17. Villanueva J, Shaffer DR, Philip J, Chaparro CA, et al. Differential exoprotease activities confer tumor-specific serum peptidome patterns. *J. Clin. Invest* 2006;116:271–284. [PubMed: 16395409]

18. Borth W. Alpha 2-macroglobulin, a multifunctional binding protein with targeting characteristics. *FASEB J* 1992;6:3345–3353. [PubMed: 1281457]
19. Chu CT, Howard GC, Misra UK, Pizzo SV. Alpha 2-macroglobulin: a sensor for proteolysis. *Ann. NY Acad. Sci* 1994;737:291–307. [PubMed: 7524403]
20. Zhou M, Lucas DA, Chan KC, Issaq HJ, et al. An investigation into the human serum “interactome”. *Electrophoresis* 2004;25:1289–1298. [PubMed: 15174051]
21. Westfall MD, Mays DJ, Sniezek JC, Pietenpol JA. The Delta Np63 alpha phosphoprotein binds the p21 and 14-3-3 sigma promoters *in vivo* and has transcriptional repressor activity that is reduced by Hay-Wells syndrome-derived mutations. *Mol. Cell. Biol* 2003;23:2264–2276. [PubMed: 12640112]
22. Gharahdaghi F, Weinberg CR, Meagher DA, Imai BS, Mische SM. Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. *Electrophoresis* 1999;20:601–605. [PubMed: 10217175]
23. Shevchenko A, Jensen ON, Podtelejnikov AV, Sagliocco F, et al. Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc. Natl. Acad. Sci. USA* 1996;93:14440–14445. [PubMed: 8962070]
24. Wilm M, Shevchenko A, Houthaeve T, Breit S, et al. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* 1996;379:466–469. [PubMed: 8559255]
25. Licklider LJ, Thoreen CC, Peng J, Gygi SP. Automation of nanoscale microcapillary liquid chromatography-tandem mass spectrometry with a vented column. *Anal. Chem* 2002;74:3076–3083. [PubMed: 12141667]
26. Tabb DL, Fernando CG, Chambers MC. MyriMatch: Highly accurate tandem mass spectral peptide identification by multivariate hypergeometric analysis. *J. Proteome Res* 2007;6:654–661. [PubMed: 17269722]
27. Zhang B, Chambers MC, Tabb DL. Proteomic parsimony through bipartite graph analysis improves accuracy and transparency. *J. Proteome Res* 2007;6:3549–3557. [PubMed: 17676885]
28. Qian WJ, Liu T, Monroe ME, Strittmatter EF, et al. Probability-based evaluation of peptide and protein identifications from tandem mass spectrometry and SEQUEST analysis: the human proteome. *J. Proteome Res* 2005;4:53–62. [PubMed: 15707357]
29. Wang D, Yuan AI, Feinman RD. Structure of alpha 2-macroglobulin-protease complexes. *Ann. NY Acad. Sci* 1983;421:90–97. [PubMed: 6202224]
30. Feldman SR, Gonias SL, Pizzo SV. Model of alpha 2-macroglobulin structure and function. *Proc. Natl. Acad. Sci. USA* 1985;82:5700–5704. [PubMed: 2412223]
31. Harpel PC. Human plasma alpha 2-macroglobulin. An inhibitor of plasma kallikrein. *J. Exp. Med* 1970;132:329–352. [PubMed: 4101346]
32. Valnickova Z, Thogersen IB, Christensen S, Chu CT, et al. Activated human plasma carboxypeptidase B is retained in the blood by binding to alpha2-macroglobulin and pregnancy zone protein. *J. Biol. Chem* 1996;271:12937–12943. [PubMed: 8662763]
33. Sheth NA, Hurkadli KS, Sathe VS, Sheth AR. Circulating levels of inhibin in cancer. *Neoplasma* 1984;31:315–321. [PubMed: 6204214]
34. Vaughan JM, Vale WW. Alpha 2-macroglobulin is a binding protein of inhibin and activin. *Endocrinology* 1993;132:2038–2050. [PubMed: 7682939]
35. O'Connor-McCourt MD, Wakefield LM. Latent transforming growth factor-beta in serum. A specific complex with alpha 2-macroglobulin. *J. Biol. Chem* 1987;262:14090–14099. [PubMed: 2443501]
36. Wolff JM, Fandel T, Borchers H, Brehmer B Jr, Jakse G. Transforming growth factor-beta1 serum concentration in patients with prostatic cancer and benign prostatic hyperplasia. *Br. J. Urol* 1998;81:403–405. [PubMed: 9523660]
37. Hendriks D, Scharpe S, van Sande M, Lommaert MP. Characterisation of a carboxypeptidase in human serum distinct from carboxypeptidase N. *J. Clin. Chem. Clin. Biochem* 1989;27:277–285. [PubMed: 2760564]
38. Strojjan P, Budihna M, Smid L, Vrhovec I, Skrk J. Cathepsin D in tissue and serum of patients with squamous cell carcinoma of the head and neck. *Cancer Lett* 1998;130:49–56. [PubMed: 9751256]
39. Stenman UH, Leinonen J, Alfthan H, Rannikko S, et al. A complex between prostate-specific antigen and alpha 1-antichymotrypsin is the major form of prostate-specific antigen in serum of patients with

- prostatic cancer: assay of the complex improves clinical sensitivity for cancer. *Cancer Res* 1991;51:222–226. [PubMed: 1703033]
40. Adkins JN, Varnum SM, Auberry KJ, Moore RJ, et al. Toward a human blood serum proteome: analysis by multidimensional separation coupled with mass spectrometry. *Mol. Cell. Proteomics* 2002;1:947–955. [PubMed: 12543931]
 41. Anderson NL, Polanski M, Pieper R, Gatlin T, et al. The human plasma proteome: a nonredundant list developed by combination of four separate sources. *Mol. Cell. Proteomics* 2004;3:311–326. [PubMed: 14718574]
 42. Chan KC, Lucas DA, Hise D, Schaefer CF, et al. Analysis of the Human Serum Proteome. *Clin. Proteomics* 2004;1:101–226.
 43. Shen Y, Jacobs JM, Camp DG 2nd, Fang R, et al. Ultra-high-efficiency strong cation exchange LC/RPLC/MS/MS for high dynamic range characterization of the human plasma proteome. *Anal. Chem* 2004;76:1134–1144. [PubMed: 14961748]
 44. Shen Y, Kim J, Strittmatter EF, Jacobs JM, et al. Characterization of the human blood plasma proteome. *Proteomics* 2005;5:4034–4045. [PubMed: 16152657]
 45. States DJ, Omenn GS, Blackwell TW, Fermin D, et al. Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. *Nat. Biotechnol* 2006;24:333–338. [PubMed: 16525410]
 46. Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, et al. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol. Cell. Proteomics* 2005;4:1487–1502. [PubMed: 15979981]
 47. Smaletz O, Scher HI, Small EJ, Verbel DA, et al. Nomogram for overall survival of patients with progressive metastatic prostate cancer after castration. *J. Clin. Oncol* 2002;20:3972–3982. [PubMed: 12351594]
 48. Whitesell L, Lindquist SL. HSP90 and the chaperoning of cancer. *Nat. Rev. Cancer* 2005;5:761–772. [PubMed: 16175177]
 49. Foran E, McWilliam P, Kelleher D, Croke DT, Long A. The leukocyte protein L-plastin induces proliferation, invasion and loss of E-cadherin expression in colon cancer cells. *Int. J. Cancer* 2006;118:2098–2104. [PubMed: 16287074]
 50. Park T, Chen ZP, Leavitt J. Activation of the leukocyte plastin gene occurs in most human cancer cells. *Cancer Res* 1994;54:1775–1781. [PubMed: 8137292]
 51. Kuefer R, Hofer MD, Gschwend JE, Pienta KJ, et al. The role of an 80 kDa fragment of E-cadherin in the meta-static progression of prostate cancer. *Clin. Cancer Res* 2003;9:6447–6452. [PubMed: 14695147]
 52. Roessler M, Rollinger W, Mantovani-Endl L, Hagmann ML, et al. Identification of PSME3 as a novel serum tumor marker for colorectal cancer by combining two-dimensional polyacrylamide gel electrophoresis with a strictly mass spectrometry-based approach for data analysis. *Mol. Cell. Proteomics* 2006;5:2092–2101. [PubMed: 16893879]
 53. Bao S, Ouyang G, Bai X, Huang Z, et al. Periostin potentially promotes metastatic growth of colon cancer by augmenting cell survival via the Akt/PKB pathway. *Cancer Cell* 2004;5:329–339. [PubMed: 15093540]
 54. Shao R, Bao S, Bai X, Blanchette C, et al. Acquired expression of periostin by human breast cancers promotes tumor angiogenesis through up-regulation of vascular endothelial growth factor receptor 2 expression. *Mol. Cell. Biol* 2004;24:3992–4003. [PubMed: 15082792]
 55. Sasaki H, Dai M, Auclair D, Fukai I, et al. Serum level of the periostin, a homologue of an insect cell adhesion molecule, as a prognostic marker in nonsmall cell lung carcinomas. *Cancer* 2001;92:843–848. [PubMed: 11550156]
 56. Sasaki H, Yu CY, Dai M, Tam C, et al. Elevated serum periostin levels in patients with bone metastases from breast but not lung cancer. *Breast Cancer Res. Treat* 2003;77:245–252. [PubMed: 12602924]
 57. Hawighorst T, Velasco P, Streit M, Hong YK, et al. Thrombospondin-2 plays a protective role in multistep carcinogenesis: a novel host anti-tumor defense mechanism. *EMBO J* 2001;20:2631–2640. [PubMed: 11387198]

58. Streit M, Riccardi L, Velasco P, Brown LF, et al. Thrombospondin-2: a potent endogenous inhibitor of tumor growth and angiogenesis. *Proc. Natl. Acad. Sci. USA* 1999;96:14888–14893. [PubMed: 10611308]
59. Neckers L. Heat shock protein 90 inhibition by 17-allyl-amino-17-demethoxygeldanamycin: a novel therapeutic approach for treating hormone-refractory prostate cancer. *Clin. Cancer Res* 2002;8:962–966. [PubMed: 12006507]
60. Eustace BK, Sakurai T, Stewart JK, Yimlamai D, et al. Functional proteomic screens reveal an essential extra-cellular role for hsp90 alpha in cancer cell invasiveness. *Nat. Cell Biol* 2004;6:507–514. [PubMed: 15146192]

Figure 1.

A2M-substrate complexes immunoprecipitated from human serum. Immunoblot for A2M is shown. A2M exists *in vivo* as a tetramer of identical 185-kDa subunits. Under reducing SDS-PAGE conditions, non-substrate bound A2M monomers^b migrate at ~ 185 kDa. Proteolysis of the bait region leads to dissociation of the 185-kDa subunit into a 100-kDa subunit^d and an 85-kDa sub-unit not recognized by the antibody used in this experiment. Reactive cross linking of a substrate to a single 100-kDa subunit produces bands ranging from 120 to 180 kDa^c, depending on the size of the substrate protein. High molecular weight complexes (>250 kDa)^a represent multiple 100-kDa A2M subunits covalently bound to a single substrate. Immunoglobulin heavy chain (IgG HC) is shown. Lane 1: plasma; Lane 2: serum; Lane 3: A2M antibody only control. Lane 4: A2M immunoprecipitation from serum. Native and cross-linked A2M-substrate complexes are observed in both serum (Lane 2) and plasma (Lane 1) and can be immunoprecipitated (Lane 4) for analysis of associated substrate proteins.

Figure 2.

Inter-sample reproducibility of protein identification by cohort. Overlap of detected proteins amongst all six patients from the cancer cohort (A) and the control cohort (B) is shown scaled to the number of peptides required for protein identification. The proportion of proteins detected in all six patients in both cohorts did not significantly vary as the number of required peptides for identification increased.

Figure 3.

Measurement of HSP90α serum concentration. Serum was collected from an expanded cohort of cancer ($n = 18$) and control ($n = 13$) patients, and HSP90α concentration was measured by ELISA. The results are shown as a box and corresponding dot plot. Each solid dot represents an individual patient measurement. The horizontal bar transecting each box represents the median cohort value, and the boundaries of each box represent the upper and lower quartile values for each cohort. The bars extend to 1.5 times the inner quartile by convention. Outlying values are shown in the cancer cohort as hollow dots. There was an increased serum concentration of HSP90α in the cancer cohort relative to control ($p = 0.001$). The median concentration of HSP90α in the cancer patients was 50.7 ng/mL (range 25.5–378.1), while the median for control subjects was 27.6 ng/mL (13.9–46.5). These data provide a separation of groups with an area under the ROC curve of 0.83 (95% confidence interval of 0.68 to 0.95).

Patient characteristics for each cohort are shown as described. Analysis of A2M-associated proteins was performed on each patient listed below

Table 1

ID	Cohort ^{a)}	Age	PSA	Disease location ^{b)}	Treatment ^{c)}
PC1	CA	78	0.55	B, S	L/A, BI
PC2	CA	42	15.18	B, S, Li	LA
PC3	CA	70	717.9	B	OR
PC4	CA	59	159.6	B, S	L/A, BI
PC5	CA	80	500.9	B, S	OR, BI
PC6	CA	55	99.69	B	LA
Median	CA	64	248.97		
CN1	CON	67	0.81		
CN2	CON	75	7.8		
CN3	CON	75	1.44		
CN4	CON	57	N/A ^{d)}		
CN5	CON	71	1.14		
CN6	CON	65	N/A ^{d)}		
Median	CON	68	2.7975		

PSA values are reported in ng/mL (normal<4) and were obtained within 2 months of serum collection for this study. Treatment represents active therapy at time of sample collection and does not account for prior chemotherapy.

a) CA = cancer; CON = control.
b) B = bone; S= soft tissue; Li = liver.
c) LA = leutinizing hormone releasing hormone agonist; BI = bisphosphonate; OR = orchiectomy.
d) N/A = not available.

Table 2

MS identification of known A2M substrates

Protein Symbol	IPI ^{a)}	Description	Total peptides from all patients	Total spectra from all patients
<i>PLG</i>	IPI00019580.1	Plasminogen	75	872
<i>F2</i>	IPI00019568.1	Thrombin	67	1351
<i>KLKB1</i>	IPI00783921.1	Plasma Kallikrein	38	325
<i>CPB2</i>	IPI00329775.7	Carboxypeptidase B2	11	62
<i>INHBB</i>	IPI00297026.5	Inhibin beta B	2	2
<i>INHBC</i>	IPI00023314.1	Inhibin beta C	4	14
<i>INHBE</i>	IPI00011036.1	Inhibin beta E	5	11

Previously identified A2M substrates were detected in serum from both patient cohorts. Total pooled number of peptides *per* protein from all patients is shown (including patients with one unique peptide match). Total patients with multiple peptide matches are shown. Additional information regarding peptide distribution *per* sample can be found in Supporting Information.

^{a)} International Protein Index (IPI) version 3.22, released October 2006.

Table 3

MS identification of select proteases

Protein	Peptide count <i>per</i> patient																
	Symbol	IP1 <i>a)</i>	Protein description	Cancer						Control							
				All	1	2	3	4	5	6	All	1	2	3	4	5	6
<i>MASP1</i>	IP100290283.5	Mannan-binding lectin serine protease 1	19	10	13	6	11	9	10	18	14	14	8	5	12	8	
<i>CPV2</i>	IP100738433.1	Carboxypeptidase N subunit 2	17	11	5	4	8	12	14	18	10	3	12	15	12	9	
<i>MASP2</i>	IP100294713.3	Mannan-binding lectin serine protease 2	17	11	13	2	4	5	3	16	12	9	4	2	7	6	
<i>CPB2</i>	IP100329775.7	Carboxypeptidase B2	11	6	2	2	8	7	10	8	1	0	6	2	4	3	
<i>CNDP1</i>	IP100064667.3	Carnosine dipeptidase 1	11	0	0	2	8	3	7	12	0	0	0	3	10	3	
<i>CPV1</i>	IP100010295.1	Carboxypeptidase N catalytic chain	5	0	1	0	0	3	3	2	0	0	0	0	2	0	
<i>CTSD</i>	IP100011229.1	Cathepsin D	1	0	0	0	0	1	0	4	1	0	0	3	0	0	
<i>HGFAC</i>	IP100029193.1	Hepatocyte growth factor activator	3	0	0	0	1	2	1	0	0	0	0	0	0	0	
<i>ADAMDEC1</i>	IP100004480.1	ADAM DEC1	1	1	0	1	0	0	0	2	0	0	2	0	0	0	

Total unique peptide hits *per* patient are provided. Spectral counts per patient and peptide sequences can be found in Supporting Information

a) International Protein Index (IP1) version 3.22, released October 2006.

Table 4

Differential protein identification by cohort

Spectral count <i>per patient</i>																
IP1 <i>a)</i>	Protein description	Cancer						Control						Total		Total
		1	2	3	4	5	6	1	2	3	4	5	6	cancer	control	
IP100010471.4	Plastin-2	21	2	20	36	28	29	3	1	18	2	25	11	136	60	
IP100641231.1	Periostin	25	0	2	10	11	7	0	0	5	6	12	3	55	26	
IP100329775.7	Carboxypeptidase B2	7	2	2	9	9	14	1	0	6	2	6	4	43	19	
IP100299778.2	Serum paraoxonase/lactonase 3	4	0	4	0	5	8	0	0	2	1	1	0	21	4	
IP100216694.3	Plastin 3 <i>b)</i>	3	1	3	3	5	4	1	0	2	0	2	1	19	6	
IP100784295.1	Heat shock protein HSP 90-alpha	6	0	0	0	12	1	0	0	0	0	0	1	19	1	
IP100020996.3	IGF-binding protein complex acid labile chain	2	0	2	5	5	6	1	0	6	0	0	3	20	10	
IP100019399.1	Serumamyloid A-4 protein	0	0	6	7	4	3	0	0	3	1	0	0	20	4	
IP100604431.1	Cullin-associated NEDD8-dissociated protein 1	10	0	0	0	5	1	0	0	1	0	0	0	16	1	
IP100022391.1	Serum amyloid P-component	15	0	0	0	0	1	0	0	0	0	0	0	16	0	
IP100291410.3	Long palate, lung and nasal epitheliomcarcinoma-associated protein 1	4	0	0	3	1	4	0	0	4	0	0	0	12	4	
IP100384051.4	Proteasome activator complex subunit 2	7	0	1	2	2	0	0	0	2	1	0	0	12	3	
IP100030154.1	Proteasome activator complex subunit 1	4	0	0	1	3	4	0	0	1	0	0	0	12	1	
IP100414676.5	Heat shock protein HSP 90-beta	2	0	0	0	6	1	0	0	0	1	0	0	9	1	
IP100020986.2	Lumican	0	0	0	3	2	6	0	0	0	0	3	0	11	3	
IP100296083.2	Pulmonary surfactant-associated protein B	4	0	1	0	2	4	0	0	2	3	0	0	11	5	
IP100219217.2	L-lactate dehydrogenase B chain	4	0	0	0	4	1	0	0	1	0	0	0	9	1	

IPI a)	Protein description	Spectral count <i>per patient</i>													
		Cancer						Control						Total	
		1	2	3	4	5	6	1	2	3	4	5	6	cancer	control
IP100010295.1	Carboxypeptidase N catalytic chain	0	1	0	0	3	3	0	0	0	0	2	0	7	2
IP100217966.6	Lactate dehydrogenase A	2	0	1	0	2	0	0	0	1	0	0	0	5	1
IP100465439.4	Fructose-bisphosphate aldolase A	3	0	0	0	2	0	0	0	0	0	0	0	5	0
IP100029193.1	Hepatocyte growth factor activator	0	0	0	1	2	1	0	0	0	0	0	0	4	0
IP100744889.1	E-cadherin	0	0	0	0	2	2	0	0	0	0	0	0	4	0
IP100012269.2	Multimerin-1	0	0	2	2	2	11	0	0	8	18	7	3	17	36
IP100013933.2	Desmoplakin	0	0	0	0	0	3	19	16	0	0	0	0	3	35
IP100003351.2	Extracellular matrix protein 1	1	0	0	5	0	1	0	0	4	4	4	8	7	20
IP100554711.1	Junction plakoglobin	0	0	0	0	0	0	9	6	0	1	0	0	0	16
IP100025465.1	Mimecan	1	0	0	0	0	0	2	0	1	5	2	0	1	10
IP100025753.1	Desmoglein-1	0	0	0	1	0	1	4	4	0	0	0	0	2	8
IP100018769.3	Thrombospondin-2	0	0	1	0	0	1	0	0	0	4	3	1	2	8
IP100301288.4	SEL-OB protein	0	0	0	0	0	0	0	0	4	0	1	3	0	8
IP100373937.2	HLAR698	0	1	0	0	0	0	3	2	0	0	0	0	1	5
IP100022246.1	Azurocidin	0	0	0	0	0	0	0	0	2	2	0	0	0	4

Total spectral hits *per patient* for each protein identification are provided. Peptide sequence and distribution can be found in Supporting Information. For inclusion in this list, at least two unique peptide matches *per* protein were detected in at least one patient.

a) International Protein Index (IPI) version 3.22, released October 2006.

b) The *p* value 0.049 determined by Wilcoxon test of spectral counts between cohorts.