



Published in final edited form as:

Nat Biotechnol. ; 29(7): 625–634. doi:10.1038/nbt.1900.

A targeted proteomics–based pipeline for verification of biomarkers in plasma

Jeffrey R Whiteaker^{1,3}, Chenwei Lin^{1,3}, Jacob Kennedy^{1,3}, Liming Hou^{1,3}, Mary Trute^{1,3}, Izabela Sokal^{1,3}, Ping Yan¹, Regine M Schoenherr¹, Lei Zhao¹, Uliana J Voytovich¹, Karen S Kelly-Spratt¹, Alexei Krasnoselsky¹, Philip R Gafken¹, Jason M Hogan¹, Lisa A Jones¹, Pei Wang¹, Lynn Amon¹, Lewis A Chodosh², Peter S Nelson¹, Martin W McIntosh¹, Christopher J Kemp¹, and Amanda G Paulovich¹

¹Fred Hutchinson Cancer Research Center, Seattle, Washington, USA

²Department of Cancer Biology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA

Abstract

High-throughput technologies can now identify hundreds of candidate protein biomarkers for any disease with relative ease. However, because there are no assays for the majority of proteins and *de novo* immunoassay development is prohibitively expensive, few candidate biomarkers are tested in clinical studies. We tested whether the analytical performance of a biomarker identification pipeline based on targeted mass spectrometry would be sufficient for data-dependent prioritization of candidate biomarkers, *de novo* development of assays and multiplexed biomarker

Users may view, print, copy, download and text and data- mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

Correspondence should be addressed to A.G.P. (apaulovi@fhcrc.org).

³These authors contributed equally to this work.

AUTHOR CONTRIBUTIONS

J.R.W. oversaw all proteomic experiments, guided the data analysis and assisted in writing the manuscript. C.L. performed or oversaw all data analysis and assisted in writing the manuscript. J.K. performed all verification studies using the 57-plex SRM assay. L.H. performed all SQ-SRM studies, assisted with the AIMS studies and assisted with configuration of the multiplex SRM assays. M.T. performed plasma sample processing throughout the project. I.S. led all verification studies using the 31-plex immuno-SRM assay and performed all ELISA analyses. P.Y. performed analysis of all response curve SRM data. R.M.S. assisted with the 31-plex immuno-SRM assay characterization and verification studies, identification of candidate biomarkers, depletion of plasma samples for verification studies and editing of the manuscript. L.Z. assisted with the 31-plex immuno-SRM assay characterization and verification studies. U.J.V. performed all western blot analyses. L.A.C. provided the Her2 mouse model. K.S.K.-S. and C.J.K. helped design the mouse experiments, performed all mouse husbandry and provided the required plasma samples. A.K. assisted in selection of candidate biomarkers, assembly of the inclusion lists for the AIMS analyses and analysis of the AIMS data. P.R.G., J.M.H. and L.A.J. performed the AIMS analyses. P.S.N. managed the overall consortium, contributed data for candidate selection, made intellectual contributions to the design of the study and assisted with editing the manuscript. P.W., M.W.M. and L.A. provided statistical input for the design of the experiments and the analysis of the data and assisted with editing the manuscript. A.G.P. designed experiments, coordinated execution of the project, coordinated analysis/interpretation of the data and assumed primary responsibility for writing the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

Note: Supplementary information is available on the Nature Biotechnology website.

Accession codes. All data used by or generated during this study have been deposited in the public domain. Hash codes from the ProteomeCommons.org Tranche network for the plasma AIMS data, shotgun data and integrated SRM data are given in Table 3. Additionally, Skyline documents for implementing the 57-plex SRM (Supplementary Data 1) and the 31-plex immuno-SRM (Supplementary Data 2) assays are available.

verification. We used a data-dependent triage process to prioritize a subset of putative plasma biomarkers from >1,000 candidates previously identified using a mouse model of breast cancer. Eighty-eight novel quantitative assays based on selected reaction monitoring mass spectrometry were developed, multiplexed and evaluated in 80 plasma samples. Thirty-six proteins were verified as being elevated in the plasma of tumor-bearing animals. The analytical performance of this pipeline suggests that it should support the use of an analogous approach with human samples.

For nearly a decade, hundreds of millions of dollars per year have been spent on the discovery of putative protein biomarkers of human diseases, especially plasma biomarkers. Despite this substantial investment, the number of new US Food and Drug Administration–approved biomarkers in plasma approved annually has remained relatively static, at no more than two per year¹. This low yield is especially disappointing in the light of the revolution in technologies for discovering biomarker candidates, which has enabled the identification of hundreds of biomarker candidates for each of the most extensively studied diseases.

Several factors have contributed to this low return on investment. First, although so-called 'omics technologies have revolutionized the discovery of candidate biomarkers, the majority of these candidates do not encode clinically actionable information even if they differ in abundance in disease and control samples. Second, biomarker discovery experiments are fraught with false discoveries resulting from biological variability and the large number of hypotheses being tested in small numbers of samples. Furthermore, because there are no validated approaches for prioritizing from among the droves of candidates those likely to be of clinical use, costly clinical validation studies must be performed on large numbers of candidates for a single novel biomarker of clinical utility to be identified. Third, because there are no quantitative assays for the majority of human proteins², assays (typically enzyme-linked immunosorbent assays (ELISA)) must be developed *de novo* for clinical testing of candidate biomarkers, and *de novo* assay development is prohibitively expensive for testing large numbers of candidate biomarkers. As a result, few putative biomarkers undergo rigorous validation, and the literature is replete with lengthy lists of candidates without follow-up.

Hence, current practice in typical biomarker discovery projects consists of the following stages. First, 'omics technologies are applied to plasma, proximal fluids and/or solid tissues to identify hundreds of candidate biomarkers. Next, candidate biomarkers undergo 'verification' in which each putative biomarker is quantified in a limited number (tens to hundreds) of clinical samples to confirm differential expression of the candidate in plasma from cases versus controls. Finally, beyond verification studies, clinical validation requires a large-scale case-control or cohort study to carefully examine the impact of other covariates on the proposed marker test, to determine the positive predictive values and false referral probabilities in real practice, and to compare or combine the new test with existing clinical tests. Because the odds are extraordinarily low that any one candidate will encode clinically useful information, large numbers of candidates must be tested if there is to be any hope of identifying a clinically useful biomarker. Thus begins a desperate search of commercial sources for antibodies and immunoassays for quantifying the candidates. Unfortunately, no assays are available for the vast majority of human proteins, and 50–60% of commercially

available antibodies are so poorly validated as to be useless³⁻⁵, resulting in a considerable waste of time and money. Novel ELISA assays are extremely expensive to generate and very difficult to multiplex. Ideally they also require a recombinant protein standard, and because many proteins cannot be purified to a single component in a soluble form, the failure rate can be quite high. Thus, if one is to test more than a few dozen candidate biomarkers (for which there are good commercially available assays) one must generate novel, analytically validated assays *de novo*. Because of the high costs associated with *de novo* assay generation, a small number of candidate biomarkers must be selected from the many hundreds of available candidates, and there is no validated method to guide the prioritization of candidates. As a result, despite tremendous effort, each biomarker project in the end faces little more than a stochastic chance of success.

Until technologies exist to enable high-content, quantitative proteomic profiling on large numbers of clinical samples, a successful biomarker development pipeline must enable triaging and measuring large numbers of candidates downstream of biomarker candidate discovery, thereby accommodating the presently inevitable high per-candidate failure rate. The recent emergence of targeted mass spectrometry (MS)-based proteomic technologies, such as accurate inclusion mass screening⁶ (AIMS) and selected reaction monitoring⁷⁻¹¹ (SRM), has raised the possibility¹²⁻¹⁵ of using these technologies for data-dependent triage and testing of hundreds of candidate biomarkers in a time- and cost-effective manner. Unlike untargeted (shotgun) modes of MS frequently used in biomarker candidate discovery studies, targeted MS methods are peptide sequence-based modes of MS that focus the full analytical capacity of the instrument on tens to hundreds of selected peptides in a complex mixture^{16,17}. By restricting detection and fragmentation to only those peptides derived from proteins of interest (that is, candidate biomarkers identified in biomarker discovery experiments), sensitivity and reproducibility are improved dramatically compared to discovery-mode MS methods.

Although targeted MS-based technologies have been proposed as the basis of a viable biomarker pipeline^{12,13,15,18-20}, no such pipeline has been demonstrated. Implementing such a notional pipeline is a substantial undertaking involving protocol development and benchmarking. The purpose of this study was to test the hypothesis that the analytical performance of a staged biomarker pipeline based on targeted MS would be sufficient for (i) data-dependent prioritization of hundreds of candidate biomarkers, (ii) *de novo* development of tens to hundreds of assays of sufficient precision, specificity and sensitivity for human studies and (iii) multiplex biomarker verification studies allowing testing of tens to hundreds of candidate biomarkers in hundreds of clinical samples while consuming a minimal volume of biospecimen material.

For decades, animal models have been critical for testing hypotheses before undertaking human studies. Using a mouse cancer model for this technology benchmarking study was advantageous because (i) samples could be easily generated and collected under strict standard operating protocols in a cost-effective manner, (ii) biological and environmental variations could be minimized and strictly controlled for, allowing us to focus on the pre-analytic and analytic variations associated with the technologies and (iii) precious human clinical samples were not consumed. Although biological variation among humans will

undoubtedly be greater than that among mice, the pre-analytic and analytic variations associated with the technologies are agnostic as to what species is used. Thus, although specific biomarkers identified in a mouse model may not be relevant to human disease, the mouse is a useful model system for benchmarking technologies intended for clinical applications.

RESULTS

Biomarker candidate identification

We selected a mouse model of breast cancer that has been extensively characterized using genomic and proteomic technologies^{21–24}. Therefore, considerable data were available for identifying putative circulating biomarkers at the outset of the study. Candidate biomarkers (1,908; Supplementary Worksheet 1) were identified by integrating 13 independent genomic and proteomic data sets^{21,22}, including candidates discovered in tumor tissue (and predicted to be secreted, leaked or shed) and/or in plasma (fig. S1.1 and table S1.2 in Supplementary Results Section 1). Quantitative assays (e.g., ELISAs) were not available for the majority of the candidates, severely limiting our ability to perform follow-up verification studies of candidates.

To triage and verify a large number of biomarker candidates, we configured a pipeline (Fig. 1) using targeted proteomic technologies, strategically staged in a manner intended to enable us to test as many candidates as possible while containing costs. Specifically, the technology with the lower expense and higher capacity to triage large numbers of candidates (AIMS) was used first, whereas the lower capacity and more expensive quantification technology (SRM) was used last (on only the most highly credentialed candidates).

Biomarker candidate prioritization (AIMS and SQ-SRM)

It is imperative to choose from among the large number of initial candidates and identify those most likely to be of use as a blood-based marker before committing resources to quantitative assay configuration. Because many candidates originate from discovery experiments based in tissue (or proximal fluids), and the number of candidates from this category may be very high (table S1.2 in Supplementary Results Section 1), the approach used should have a large capacity for determining the presence of each candidate in plasma. We chose to use AIMS⁶ for this purpose.

For AIMS analysis, peptide mass-to-charge ratios (m/z) on a user-generated software inclusion list are monitored in each scan using an Orbitrap MS system, and MS/MS spectra are acquired only when a peptide from the list is detected with both the correct accurate mass and charge state⁶. The fragmentation patterns from these acquired spectra are used to confirm the presence of the peptide. We used AIMS⁶ to confirm that a given peptide (and thus the protein from which it is derived) was detectable in the plasma (and thus might form the basis of a blood-based test), providing a bridge from discovery to MS-based targeted assay development. Only candidates detected in the plasma by AIMS were advanced to assay development.

The AIMS study is described in detail in Supplementary Results Section 2. Briefly, a pool of plasma with equal contributions from 20 tumor-bearing animals was collected, abundant plasma proteins were depleted, and plasma was proteolyzed with trypsin and subjected to strong cation exchange (SCX) chromatography to generate ten fractions. Each fraction was subjected to 21 AIMS runs and six liquid chromatography (LC)-MS/MS shotgun runs. For AIMS analysis, we targeted 16,961 proteotypic peptides (experimentally observable peptides that each uniquely identify a specific protein or protein isoform) from 1,144 candidate biomarkers. Of these, 764 candidates were not targeted owing to capacity limitations and the absence of empirical proteotypic peptides; 572 candidates (50%) were detected (Supplementary Worksheet 2). The highest success rates for detection (table S2.2 in Supplementary Results Section 2) came from those candidates identified in both tissue and plasma (79%) and those discovered in plasma but not tissue (77%). An interesting feature of this analysis is the reasonably high (17%) success rate for candidates originally discovered exclusively in the tissue data sets.

AIMS analysis added little benefit (1–5% additional candidates) over the standard shotgun analysis for candidates that had previously been discovered in plasma (table S2.3 in Supplementary Results Section 2). The added value of the inclusion list was quite apparent for candidates previously identified exclusively in tumor tissue, where a large number ($n = 48$; 26% additional candidates) of candidates discovered in tissue were confirmed by AIMS only (that is, not detected by shotgun analysis). This was expected because candidates originally discovered only in tissue are likely present at a lower concentration in plasma compared with those candidates discovered directly in plasma; thus the use of the inclusion list (and associated added sensitivity) is beneficial.

Reagent costs (e.g., synthetic stable isotope–labeled standard (SIS) peptides) prevented us from developing quantitative SRM assays for all the candidates confirmed by the AIMS analysis, necessitating a further prioritization step. For this we used a semiquantitative application of SRM²⁵ that involves normalization of signals, as described below. In the targeted SRM mode of MS, a specific tryptic peptide is selected as a stoichiometric representative of the protein from which it is cleaved. For semiquantitative SRM (SQ-SRM) evaluation, SRM signals from housekeeping proteins (that is, plasma proteins with the same abundance in tumor-bearing and control animals) were used to normalize SRM signals from candidate proteins across LC-SRM-MS runs (Supplementary Worksheet 3). The AIMS/shotgun data were used to empirically select 9,647 transitions (i.e., a pair of precursor and fragment ion m/z for a given peptide) representing 383 candidates that were measured in triplicate on depleted plasma pooled from 20 tumor-bearing mice and 20 control mice. Of these, 189 candidates were not targeted because the observed peptides did not meet our SRM peptide criteria (Supplementary Results Section 3). SRM signals were curated by hand to ensure proper peak integration and retention time alignment of multiple transitions emanating from each peptide; only transitions showing consistency across runs with a signal-to-noise ratio ≥ 8 were included in the analysis. Candidates were binned based on confidence level (figure s3.2 in Supplementary Results Section 3). High confidence was assigned to the 164 candidates for which we detected at least three transitions from at least one proteotypic peptide. Medium confidence was assigned to the 38 candidates for which

we detected two transitions from one proteotypic peptide. The remaining 171 candidates were detected by a single transition only and thus were considered to be unsubstantiated. The 164 high-confidence candidates were further prioritized by receiver operating characteristic (ROC) analysis (Supplementary Results Section 3), narrowing the list to those 118 candidates with ROC > 0.7. Thus the triage step was successful in narrowing the number of candidates to a reasonable and manageable list, and for determining the subset of candidates for which an antibody (that is, immuno-SRM assay) would be required for robust measurement.

Ninety-one proteins were selected for *de novo* quantitative assay generation (Supplementary Results Section 3). Seven candidates were selected from the housekeeping proteins and 49 candidates were selected from the high-confidence candidates, whose transition signals are likely to be sufficient for quantification by SRM-MS. To evaluate the false-negative rate of the pipeline for prioritizing viable lower abundance candidates, we also selected an additional 35 candidates for which signals in the SQ-SRM analysis were too low for assessment. For these 35 candidates, we generated affinity-purified, polyclonal antibodies for immunoaffinity enrichment of proteotypic peptides upstream of SRM-MS. This technique, stable isotope standards and capture by antipeptide antibodies²⁶ (SISCAPA), achieves limits of quantification (LOQ) of proteins in the low ng protein/ml range^{21,27-29} from 10 μ l and low pg protein/ml range²⁹ from 1 ml plasma. Development and characterization of the SRM and immuno-SRM assays are described in the following sections.

Quantitative SRM (Q-SRM) assay development

We configured 57 peptide analytes (representing 49 candidates and seven housekeeping proteins) into a multiplex SRM-MS assay, and SIS peptides were used as internal standards (Supplementary Results Section 4 and Supplementary Worksheet 4a-c). A list of peptides and SRM transitions is provided as a Skyline document; see Supplementary Data 1. The analytical performance of the multiplex assay was determined by generating response curves for candidate biomarkers in a trypsinized, depleted plasma matrix (Supplementary Appendix A). Performance characteristics of the assays for the candidate biomarkers are summarized in Table 1. The median analytical coefficient of variation (CV) across all analytes was 5.7% at the LOQ. The median limit of detection (LOD) was 68 ng protein/ml plasma, and median LOQ was 157 ng protein/ml plasma.

Quantitative immuno-SRM assay development

We developed 31 affinity-purified polyclonal antibodies to proteotypic peptides from 30 of the 35 selected candidates described above, and these antibodies were used to construct a 31-plex immuno-SRM assay, using SIS peptides as internal standards (Supplementary Results Section 5 and Supplementary Worksheet 5a-d). A list of peptides and SRM transitions is provided as a Skyline document; see Supplementary Data 2. The analytical performance of the 31-plex assay was determined by generating response curves in a trypsinized plasma matrix (Supplementary Appendix B). Performance characteristics of the assays are summarized in Table 2; median CV across all analytes was 7.4% (triplicate SISCAPA captures), median LOD was 26 ng protein/ml plasma and median LOQ was 48 ng

protein/ml plasma. Note that LOD/LOQs are not directly comparable to those for the multiplex Q-SRM assay described above, as the Q-SRM assays tested a different set of analytes and were run on depleted plasma using a Qtrap 5500 instrument, whereas immuno-SRM assays were run on nondepleted plasma using a Qtrap 4000.

Verification studies

The analytically validated, 57-plex SRM and 31-plex immuno-SRM assays were used in a quantitative biomarker verification study (Supplementary Results Section 6). No mice used for discovery, AIMS or SQ-SRM were used for the verification studies, which used a separate cohort of animals. All 88 assays were run in triplicate on 80 plasma samples derived from three cohorts of mice, as described below (Supplementary Worksheet 6a–g and Supplementary Appendix C). Median assay precision (during the verification studies) was 3.4% CV (analytical variation) for the 57-plex SRM and 12.5% CV (preanalytical and analytical variation combined) for the 31-plex SISCAPA assays.

The first cohort included ten control and ten tumor-bearing mice, where tumors were clinically apparent (~1 cm, equivalent to the tumor burden in the mice used for discovery). For the 57-plex assay, 46 of the targeted peptide analytes were present above their assay LOQs in at least 50% of the cancer-bearing animals (Supplementary Worksheet 6a). Thirty proteins (2310057J18Rik, Aldoc, Chi311, Comp, Ctsb, Ctstl, Dsg2, Ext2, Fetub, Fn1, Hsp90b1, Hspa5, Itih3, Kng2, Lcn2, Lcp1, Mug1, Nrp1, Orm2, Spp1, P4hb, Papln, Pdia3, Ppic, Ptk7, Ptpf, Spint1, Tnc, Itih1, Pzp) showed significant elevation in tumor-bearing animals (area under the curve (AUC) ≥ 0.8 , $P \leq 0.01$; Table 1). These success rates indicate that the SQ-SRM analysis of pooled plasma successfully identified candidates that (i) could be measured by SRM in depleted plasma without antibody enrichment, and (ii) had a high probability of being verified in quantitative studies examining individual (that is, not pooled) plasma samples.

For the 31-plex immuno-SRM assay, 17 of the analytes were present above their assay LOQs in at least 50% of the cancer-bearing animals (Supplementary Worksheet 6a). Of these, six proteins (B4galnt1, Ext1, Man2a1, Mfge8, Pcdh18, Prdx4) showed significant elevation in tumor-bearing animals (AUC ≥ 0.8 , $P \leq 0.01$). The results are summarized in Table 2, and representative examples are shown in Figures 2–4. Where available, commercial antibodies were used for western blot analysis or ELISA of plasma and confirmed the results of the SRM-based assays (Figs. 3, 4 and figure S6.4 in Supplementary Results Section 6). These results demonstrate that generation of an anti-peptide antibody and configuration of an immuno-SRM assay enhanced the SRM signals from target analytes, allowing quantification in plasma.

A higher percentage of candidates assayed by Q-SRM (30/56 or 54%) were verified compared to the number verified using immuno-SRM (6/35 or 17%). This is expected because the immuno-SRM targets were unsubstantiated by SQ-SRM and likely represent lower abundance proteins. Consistent with this, 46/57 Q-SRM analytes were above their assay LOQs in at least 50% of the cancer-bearing animals, whereas only 17/31 were for the immuno-SRM analytes.

The second cohort included ten control and ten tumor-bearing mice, where the tumors were preclinical (that is, not palpable); hence, the tumor burden was considerably lower than that used in the biomarker discovery work. This second time point was examined to determine whether the SRM technology is sufficiently sensitive to quantify biomarkers that are only fractionally elevated as a consequence of an extremely low tumor burden. For the 57-plex SRM assay, 29 of the targeted analytes were present above their assay LOQs in at least 50% of the cancer-bearing animals (Supplementary Worksheet 6b). Of these, one protein (Spp1) was present at significantly ($P \leq 0.01$) elevated levels in tumor-bearing animals. For the 31-plex SISCAPA assay, 20 of the analytes were present above their assay LOQs in at least 50% of the cancer-bearing animals. Of these, one protein (Mfge8) was present at significantly ($P \leq 0.01$) elevated levels in tumor-bearing animals. Hence, as expected, detection of very low disease burden is far more challenging than detection of high disease burden.

To determine the specificity of the biomarkers elevated in cancer, we tested a third cohort of 10 control mice and 30 mice that had been injected with a compound to stimulate confounding conditions including inflammation, apoptosis, necrosis, wound healing and angiogenesis (Supplementary Results Section 6). Of the 36 proteins elevated in the plasma of animals with clinically apparent breast cancers, 17 were above their assay LOQ in at least 50% of the animals with the confounding condition. Only two proteins (Pdia3, Ppic) were elevated in a confounding condition ($AUC \geq 0.8$, $P \leq 0.01$), suggesting that the majority of biomarkers tested are not general indicators of inflammation or tissue remodeling.

The distribution of the discovery methods used to identify the verified candidates is shown in table S6.1 in Supplementary Results Section 6. The number of candidates tested in this study makes it difficult to draw conclusions regarding the general success rates of various discovery approaches. It is noteworthy that a high percentage of candidates that had been discovered exclusively in tissue-based analyses were substantiated by Q-SRM (4/4) and immuno-SRM (3/15).

DISCUSSION

The purposes of this study were to develop protocols and procedures for a large-scale biomarker pipeline based on targeted MS and to test the hypothesis that the pipeline would be sufficient to support large-scale clinical biomarker studies in humans. Previous work has characterized the performance of the individual components in isolation, such as AIMS, SRM and immuno-SRM technologies. Nonetheless, they had not been assembled in a pipeline, and were generally carried out in small numbers of reference samples and small numbers of analytes, inadequate to demonstrate the feasibility of deploying them in clinical studies.

To be useful for human biomarker studies, a pipeline must enable (i) data-dependent prioritization of hundreds of candidate biomarkers, (ii) cost-effective *de novo* development of tens to hundreds of assays of sufficient precision, specificity and sensitivity for human studies and (iii) multiplex biomarker verification studies allowing testing of tens to hundreds

of candidate biomarkers in hundreds of clinical samples while consuming a minimal volume of biospecimens.

Based on these criteria, the procedures and protocols we have described, along with our benchmarking activities, demonstrate that the analytical performance of the targeted proteomic pipeline should be sufficient to support large-scale verification studies in humans.

First, the AIMS and SQ-SRM stages of the pipeline enabled data-dependent prioritization of hundreds of candidate biomarkers. Because of the ability to schedule AIMS detection based on the known or predicted retention times of the target peptides, >1,000 parent ions can be monitored in a single LC-MS/MS run, enabling a large number of candidates to be screened for detection in plasma in a single run. This high capacity enables prefractionation of plasma, thus enabling detection of analytes in the low ng protein/ml plasma range (levels consistent with many known cancer biomarkers). Of the 1,908 candidates identified from genomic and proteomic data sets, proteotypic peptides could be identified for 1,551 candidates (81%). This already high coverage will expand further as ongoing projects designed to characterize proteotypic peptides for all human proteins mature (<http://www.mrmatlas.org/>). Of the 1,144 candidates targeted for detection in plasma by AIMS, ~50% were detected in plasma, despite the fact that many of the candidates had been discovered in tumor tissues, not plasma.

Although AIMS provided confirmation that a candidate biomarker could be detected in the plasma of tumor-bearing animals, it did not provide quantitative comparison of the levels of the candidates in cases versus controls. The gold standard for quantitative MS in the clinical laboratory is SRM-MS, which has been used for decades to quantify small molecules in human blood samples³⁰. An isotopically labeled internal standard typically is used for rigorous quantification, but label-free approaches provide a potential cost-effective method for rapid semiquantitative screening of large numbers of candidates²⁵. In this study we normalized the SRM signals of candidate biomarkers to those from a set of housekeeping proteins each of whose abundances did not, on average, differ between cases and controls (Supplementary Results Section 3). Three hundred eighty-three of the 572 candidates (67%) detected in plasma were associated with proteotypic peptides that met criteria for SRM-based quantification (Supplementary Results Section 3). Of these, 373 (97%) contained transitions that were reproducibly detectable with a signal-to-noise ratio ≥ 8 , and were thus analyzed by SQ-SRM in pooled case versus control plasma to estimate whether the corresponding biomarker candidates were differentially abundant and thus might be useful biomarkers.

By far, the most challenging stage of the pipeline was verifying the specificity of the detected transitions in SQ-SRM. The difficulty resulted from the lack of internal standards, which provide reference signals for the verification of analyte specificity. Thus, in our experiment, the choice of candidates for quantitative SRM assay development was limited to the most abundant proteins or peptides for which multiple transitions were identified. For example, of the 373 analytes tested, only 164 (44%) were associated with at least three transitions with perfectly aligned retention times, indicating high confidence for specificity of the detection for the targeted analyte. Indeed 171 of the 373 analytes (46%) could not be

substantiated with confidence, owing to difficulty in detecting transitions. The use of more affordable, albeit less pure, stable isotope standards from a process generally referred to as spot or membrane synthesis³¹ remains to be tested but may offer a viable, cost-effective, alternative strategy to circumvent this issue. Furthermore, the specificity in SQ-SRM could also be improved by using recent advances in the data-dependent inclusion of transitions, such as intelligent SRM³². It is notable that the success rate in verifying the credentialed candidates was very good (36/91 candidates or 40%). Therefore, MS-based techniques can be used to triage candidate biomarkers.

The second activity that a biomarker pipeline must support for human studies is the cost effective *de novo* development of tens to hundreds of assays of sufficient precision, specificity and sensitivity for human studies. The cost of reagents for generating Q-SRM assays are those of the SIS peptides, for which heavy-light pairs can be obtained for <\$1,000 per analyte. Reagent costs for generating a novel immuno-SRM assay with >90% success rate is <\$5,000, and the average yield of an affinity-purified polyclonal antibody allows for testing of hundreds of plasma samples to assess the utility of a marker before investing in monoclonal antibody development³³. Lead time for assay generation is ~24 weeks, including selection of proteotypic peptides, synthesis of immunogens, generation of antibodies, optimization of assay conditions and generation of response curves. In this study, 88 novel assays were developed and characterized in <1 year by one laboratory. In contrast, it would be extraordinary for a single academic laboratory to successfully configure 1–10 ELISA assays *de novo* in 1–2 years, especially at a comparable cost.

Regarding precision, as precision of the assay deteriorates, there is a nonlinear increase in the dispersion of the results, and thus the clinical signal may be drowned out by analytical noise. For example, biomarker measurements using an assay with poor precision will have a broader reference interval (due to analytical variation) and will thus be of less value for clinical classification of patients. As discussed elsewhere²⁰, statistical verification of a novel biomarker showing comparable biological variation to prostate-specific antigen would require testing plasma samples from a minimum of 500 cases and 500 well-matched controls using an assay technology associated with CV $\leq 20\%$. In this study, we observed CV < 15% for the majority of assays, demonstrating the feasibility of our approach for human biomarker verification studies.

Beyond verification studies, true clinical validation will require an even larger-scale case-control or cohort study to carefully examine the impact of other covariates on the proposed marker test, to determine the positive predictive values and false referral probabilities in real practice, and to compare or combine the new test with existing clinical tests. In the field of clinical chemistry, quality specifications for assay precision are routinely based on biological variation³⁴. It is widely accepted among clinical chemists³⁴ that the desirable level of imprecision in measurements should be <50% of the average within-subject variation (in which case the amount of variability added to true test-result variability is ~10%). Where desirable performance standards are not attainable with current methodology, the minimum level of imprecision in measurements should be <75% of the average within-subject variation (in which case the amount of variability added to true test-result variability is ~25%). Across 163 protein analytes for which human biological variation has been

determined (<http://www.westgard.com/biodatabase1.htm>), the median within-subject biologic variation is 10%; hence, the median desirable specification for assay precision is 5%. Achieving this very high bar for clinical implementation of the immuno-SRM assays in hospital laboratories will likely require optimization of the individual assays on an analyte-by-analyte basis.

Regarding specificity, SRM-based assays have a distinct advantage over conventional immunoassays, which are prone to interferences³⁵. With SRM, assay specificity is ensured by monitoring multiple transitions from each analyte and by inclusion of an internal isotopically labeled standard. Furthermore, where interferences are present, they are readily detected and can be avoided by selection of alternative transitions.

Regarding sensitivity, many known human cancer protein biomarkers are in the ng-to-pg of protein/ml plasma concentration range; early disease detection, such as in cancer, may require much higher sensitivities. In this study, assay LOQs in the low ng protein/ml plasma range were typically achieved when starting with only 10 μ l of plasma (small volumes were used in this study due to the limited plasma yield from individual mice). However, immuno-SRM assays can achieve low pg protein/ml plasma LOQs when capturing from higher plasma volumes²⁹. Immuno-SRM assays are readily amenable to the testing of hundreds of biospecimens in human studies, as sample handling is minimal and largely automated and performed in a 96-well format^{29,33}. In contrast, for Q-SRM studies (without an anti-peptide antibody), sensitivity remains the primary obstacle to use of SRM assays, without enrichment of the target analyte. Coupling SRM with abundant protein depletion and sample fractionation significantly improves LOQs³⁶⁻³⁸, but multiplex assay configuration may be difficult for hundreds of candidates, and the intensive sample processing required has a great impact on sample throughput. Hence, for large-scale biomarker verification studies on hundreds of plasma samples, it is beneficial to generate an antibody for immuno-SRM measurements, obviating the need for depletion or fractionation.

The third activity that a biomarker pipeline must support for human studies is multiplex biomarker verification studies that permit testing of tens to hundreds of candidate biomarkers in hundreds of clinical samples²⁰, while consuming a minimal volume of biospecimens per analyte. The SRM technology is highly amenable to multiplex measurements, as demonstrated in this study and others³⁹. Current instrumentation and software allow for scheduling of the transitions being monitored (based on peptide retention times on the high-performance liquid chromatography (HPLC) column), theoretically enabling hundreds of peptide analytes to be quantified in a single LC-MS run. The high multiplex level of these technologies also allows a very large number of measurements to be made in a relatively short period of time, accelerating progress from the discovery step to verification. For example, in this study 80 plasma samples were run in triplicate for 88 analytes, totaling 21,120 assays run in one laboratory in <6 months, demonstrating the feasibility for human biomarker studies. The ability to test ten to hundreds of biomarker candidates in multiplex also makes the use of larger plasma volumes a viable option for early human biomarker studies, allowing for increased assay sensitivity (as discussed above).

In this study, we demonstrate that a staged, targeted proteomic pipeline enables triage and follow-up quantitative testing of a far larger number of candidates than would have been possible using conventional technologies, marking a substantial improvement over the current state of biomarker evaluation. Although the true impact of this or any biomarker development pipeline will not become apparent until it is successfully used to discover novel biomarkers in humans, this study uses an animal model to benchmark the performance of the proposed pipeline and thereby demonstrates the feasibility of, and thus sets the stage for, applying the pipeline to human biomarker studies. Because biological and disease subtype heterogeneity (pre-analytical variables) likely differ between mouse models and humans, biomarkers identified using mice will not necessarily be of clinical utility in humans. Nonetheless, because the analytical performance of the technologies (e.g., sensitivity, precision, specificity and multiplex characteristics) as well as pre-analytical sample processing are not species-specific, we can confidently extrapolate the analytic performance observed using a mouse model to analogous studies with humans, even if the specific biomarkers are of limited clinical potential.

METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/naturebiotechnology/>.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by a grant from The Paul G. Allen Family Foundation and the National Institutes of Health grant U24 CA126476 from the National Cancer Institute Clinical Proteomic Technology Assessment Center (CPTAC). We thank our Biomarker Project advisory board members for advice and input: R. Aebersold, L. Anderson, E. Diamandis, R. Smith, F. Appelbaum, D. Gottschling, M. Groudine, J. Roberts, V. Vasioukhin and N. Urban. We thank all members of the Biomarker Project team for input: L. Hartwell, S.M. Hanash, S.J. Pitteri, H. Wong, K.E. Gurley, D. Liggitt, D.B. Martin, T. Whitmore, A. Peterson, R. Prueitt, M. Fitzgibbon, J.K. Eng, D. May, T. Holzman, Y. Zhang, A. Stimmel, S.L. Zriny, R. Dumpit, I. Coleman and T.D. Lorentzen.

References

1. Anderson NL. The clinical plasma proteome: a survey of clinical assays for proteins in plasma and serum. *Clin Chem*. 2010; 56:177–185. [PubMed: 19884488]
2. Anderson NL, et al. A human proteome detection and quantitation project. *Mol Cell Proteomics*. 2009; 8:883–886. [PubMed: 19131327]
3. Bordeaux J, et al. Antibody validation. *Biotechniques*. 2010; 48:197–209. [PubMed: 20359301]
4. Haab BB, et al. A reagent resource to identify proteins and peptides of interest for the cancer community: a workshop report. *Mol Cell Proteomics*. 2006; 5:1996–2007. [PubMed: 16867976]
5. Stoevesandt O, Taussig MJ. Affinity reagent resources for human proteome detection: initiatives and perspectives. *Proteomics*. 2007; 7:2738–2750. [PubMed: 17639606]
6. Jaffe JD, et al. Accurate inclusion mass screening: a bridge from unbiased discovery to targeted assay development for biomarker verification. *Mol Cell Proteomics*. 2008; 7:1952–1962. [PubMed: 18534968]

7. Kuhn E, et al. Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and ¹³C-labeled peptide standards. *Proteomics*. 2004; 4:1175–1186. [PubMed: 15048997]
8. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci USA*. 2003; 100:6940–6945. [PubMed: 12771378]
9. Barr J, et al. Isotope-dilution mass spectrometric quantification of specific proteins: model application with apolipoprotein A-1. *Clin Chem*. 1996; 42:1676–1682. [PubMed: 8855153]
10. Wu S, et al. Targeted proteomics of low-level proteins in human plasma by LC/MSn: using human growth hormone as a model system. *J Proteome Res*. 2002; 1:459–465. [PubMed: 12645918]
11. Barnidge D, Goodmanson M, Klee G, Muddiman D. Absolute quantification of the model biomarker prostate-specific antigen in serum by LC-MS/MS using protein cleavage and isotope dilution MS. *J Proteome Res*. 2004; 3:644–652. [PubMed: 15253448]
12. Wang P, Whiteaker JR, Paulovich AG. The evolving role of mass spectrometry in cancer biomarker discovery. *Cancer Biol Ther*. 2009; 8:1083–1094. [PubMed: 19502776]
13. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol*. 2006; 24:971–983. [PubMed: 16900146]
14. Anderson NL. The roles of multiple proteomic platforms in a pipeline for new diagnostics. *Mol Cell Proteomics*. 2005; 4:1441–1444. [PubMed: 16020426]
15. Makawita S, Diamandis EP. The bottleneck in the cancer biomarker pipeline and protein quantification through mass spectrometry-based approaches: current strategies for candidate verification. *Clin Chem*. 2010; 56:212–222. [PubMed: 20007861]
16. Lange V, Picotti P, Domon B, Aebersold R. Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol*. 2008; 4:222. [PubMed: 18854821]
17. Pan S, et al. Mass spectrometry based targeted protein quantification: methods and applications. *J Proteome Res*. 2009; 8:787–797. [PubMed: 19105742]
18. Carr SA, Anderson L. Protein quantitation through targeted mass spectrometry: the way out of biomarker purgatory? *Clin Chem*. 2008; 54:1749–1752. [PubMed: 18957555]
19. Boja E, et al. Restructuring proteomics through verification. *Biomark Med*. 2010; 4:799–803. [PubMed: 21133699]
20. Paulovich AG, Whiteaker JR, Hoofnagle AN, Wang P. The interface between biomarker discovery and clinical validation: the tar pit of the protein biomarker pipeline. *Proteomics Clin Appl*. 2008; 2:1386–1402. [PubMed: 20976028]
21. Whiteaker JR, et al. Integrated pipeline for mass spectrometry-based discovery and confirmation of biomarkers demonstrated in a mouse model of breast cancer. *J Proteome Res*. 2007; 6:3962–3975. [PubMed: 17711321]
22. Schoenherr RM, et al. Proteome and transcriptome profiles of a Her2/Neu-driven mouse model of breast cancer. *Proteomics Clin Appl*. 2011; 5:179–188. [PubMed: 21448875]
23. Moody SE, et al. Conditional activation of Neu in the mammary epithelium of transgenic mice results in reversible pulmonary metastasis. *Cancer Cell*. 2002; 2:451–461. [PubMed: 12498714]
24. Omenn GS, Yocum AK, Menon R. Alternative splice variants, a new class of protein cancer biomarker candidates: findings in pancreatic cancer and breast cancer with systems biology implications. *Dis Markers*. 2010; 28:241–251. [PubMed: 20534909]
25. Zhang H, et al. Methods for peptide and protein quantitation by liquid chromatography-multiple reaction monitoring mass spectrometry. *Mol Cell Proteomics*. 2011; 10:M110.006593.
26. Anderson NL, et al. Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). *J Proteome Res*. 2004; 3:235–244. [PubMed: 15113099]
27. Hoofnagle AN, Becker JO, Wener MH, Heinecke JW. Quantification of thyroglobulin, a low-abundance serum protein, by immunoaffinity peptide enrichment and tandem mass spectrometry. *Clin Chem*. 2008; 54:1796–1804. [PubMed: 18801935]
28. Kuhn E, et al. Developing multiplexed assays for troponin I and interleukin-33 in plasma by peptide immunoaffinity enrichment and targeted mass spectrometry. *Clin Chem*. 2009; 55:1108–1117. [PubMed: 19372185]

29. Whiteaker JR, Zhao L, Anderson L, Paulovich AG. An automated and multiplexed method for high throughput peptide immunoaffinity enrichment and multiple reaction monitoring mass spectrometry-based quantification of protein biomarkers. *Mol Cell Proteomics*. 2010; 9:184–196. [PubMed: 19843560]
30. Want EJ, Cravatt BF, Siuzdak G. The expanding role of mass spectrometry in metabolite profiling and characterization. *ChemBioChem*. 2005; 6:1941–1951. [PubMed: 16206229]
31. Picotti P, et al. High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. *Nat Methods*. 2010; 7:43–46. [PubMed: 19966807]
32. Kiyonami R, et al. Increased selectivity, analytical precision, and throughput in targeted proteomics. *Mol Cell Proteomics*. 2011; 10:M110.002931. [PubMed: 20664071]
33. Whiteaker JR, et al. Evaluation of large scale quantitative proteomic assay development using peptide affinity-based mass spectrometry. *Mol Cell Proteomics*. 2011; 10:M110.005645. [PubMed: 21245105]
34. Fraser, CG. *Biological Variation: from Principles to Practice*. AACC Press; Washington, DC: 2001.
35. Hoofnagle AN, Wener MH. The fundamental flaws of immunoassays and potential solutions using tandem mass spectrometry. *J Immunol Methods*. 2009; 347:3–11. [PubMed: 19538965]
36. Keshishian H, et al. Quantification of cardiovascular biomarkers in patient plasma by targeted mass spectrometry and stable isotope dilution. *Mol Cell Proteomics*. 2009; 8:2339–2349. [PubMed: 19596694]
37. Stahl-Zeng J, et al. High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. *Mol Cell Proteomics*. 2007; 6:1809–1817. [PubMed: 17644760]
38. Keshishian H, Addona T, Burgess M, Kuhn E, Carr SA. Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. *Mol Cell Proteomics*. 2007; 6:2212–2229. [PubMed: 17939991]
39. Anderson L, Hunter CL. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol Cell Proteomics*. 2006; 5:573–588. [PubMed: 16332733]
40. Ekins R, Edwards P. On the meaning of “sensitivity”. *Clin Chem*. 1997; 43:1824–1831. [PubMed: 9341999]
41. Ekins R, Edwards P. On the meaning of “sensitivity”: a rejoinder. *Clin Chem*. 1998; 44:1773–1778. [PubMed: 9702979]

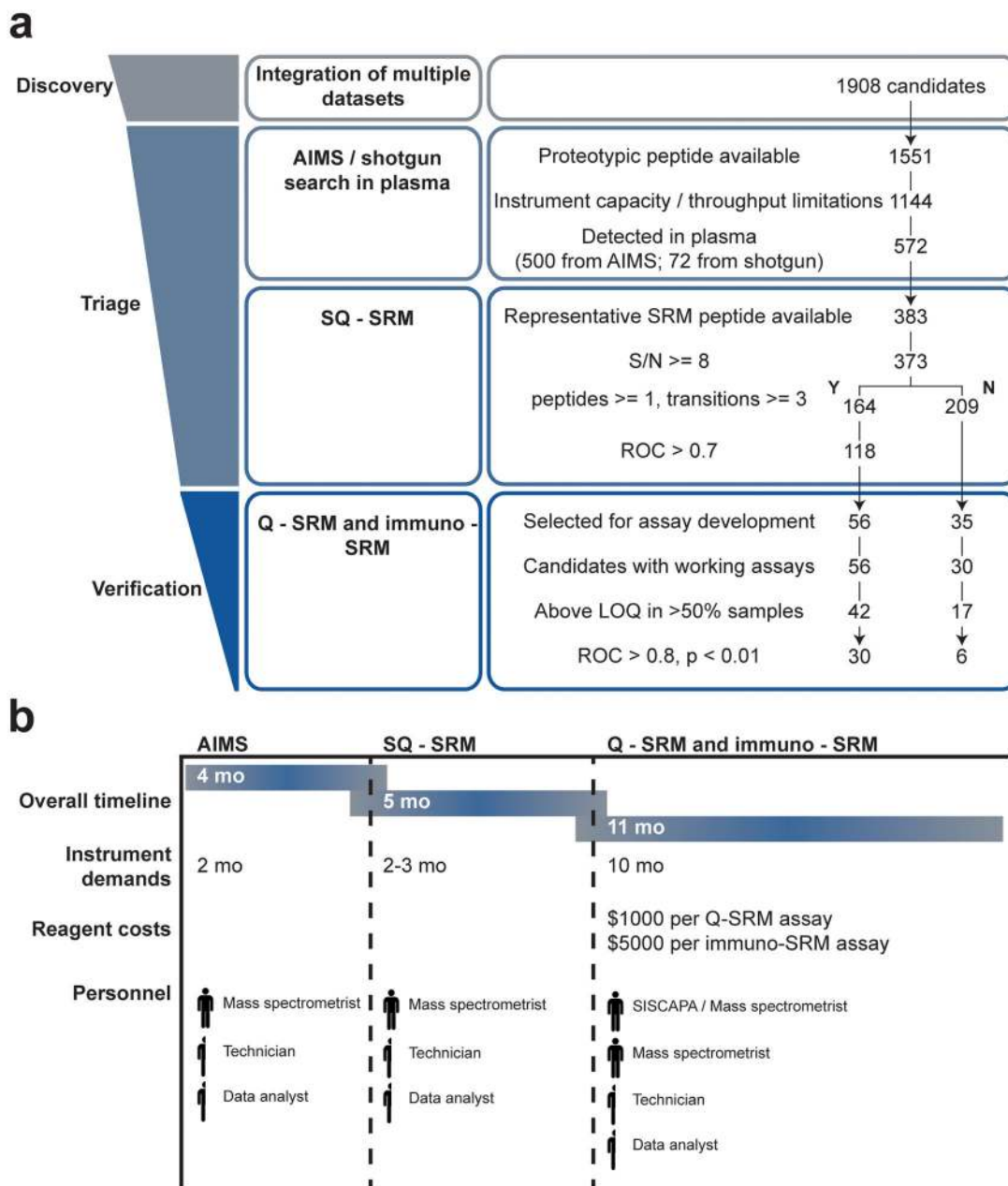


Figure 1. Multistage, targeted proteomic pipeline for triage and verification of biomarker candidates. (a) Overview of the workflow used to triage and verify candidate biomarkers, showing the flux of candidates at each stage of the pipeline. (b) Required resources for implementing the proteomic pipeline. The overall timeline includes time for data collection and analysis. For Q-SRM and immuno-SRM measurements, the overall timeline includes synthetic peptide quality control, development of SRM methods, acquisition of response curves and data analysis (but not the time required to generate antibodies, which can be interspersed with other activities). Instrument demands are summarized independently to provide an estimate of the required laboratory resources to carry out the study. Additional reagent costs (e.g.,

peptide standards and antibodies) are required for Q-SRM and immuno-SRM assays. Finally, the required personnel used in each phase of the study are denoted as full-time equivalents.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

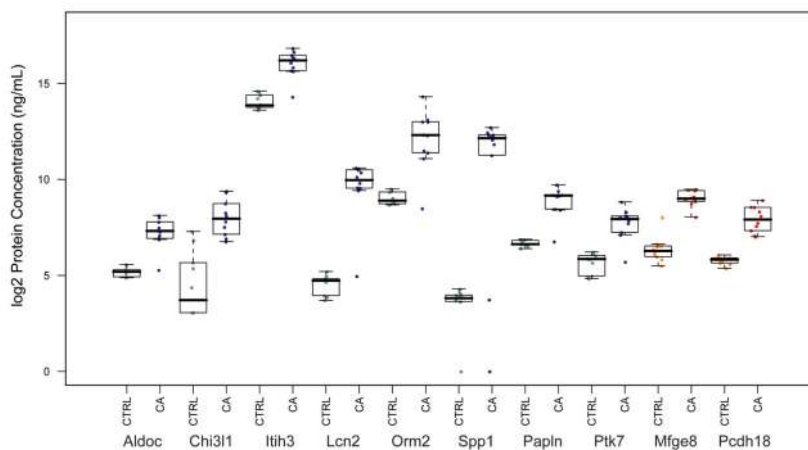


Figure 2.

Concentrations of biomarker candidates in plasma of tumor-bearing animals compared to controls. The median value is plotted as a line with each box displaying the distribution of the inner quartiles with whiskers showing 95% of the data. Results shown are derived from the SRM-MS–based measurements in Supplementary Worksheet 6a (worksheet labeled “clinically apparent”). Gray and blue data points represent Q-SRM data, whereas orange and red data points represent immuno-SRM data. Cases and controls were matched with respect to age and gender, housed in the same cage, euthanized on the same day, and derived from different animals than those used in the discovery and prioritization experiments. (Although Aldoc and Chi311 meet significance based on *P*-values, their levels are near their assay LODs.) CA, clinically apparent tumors; CTRL, matched controls.

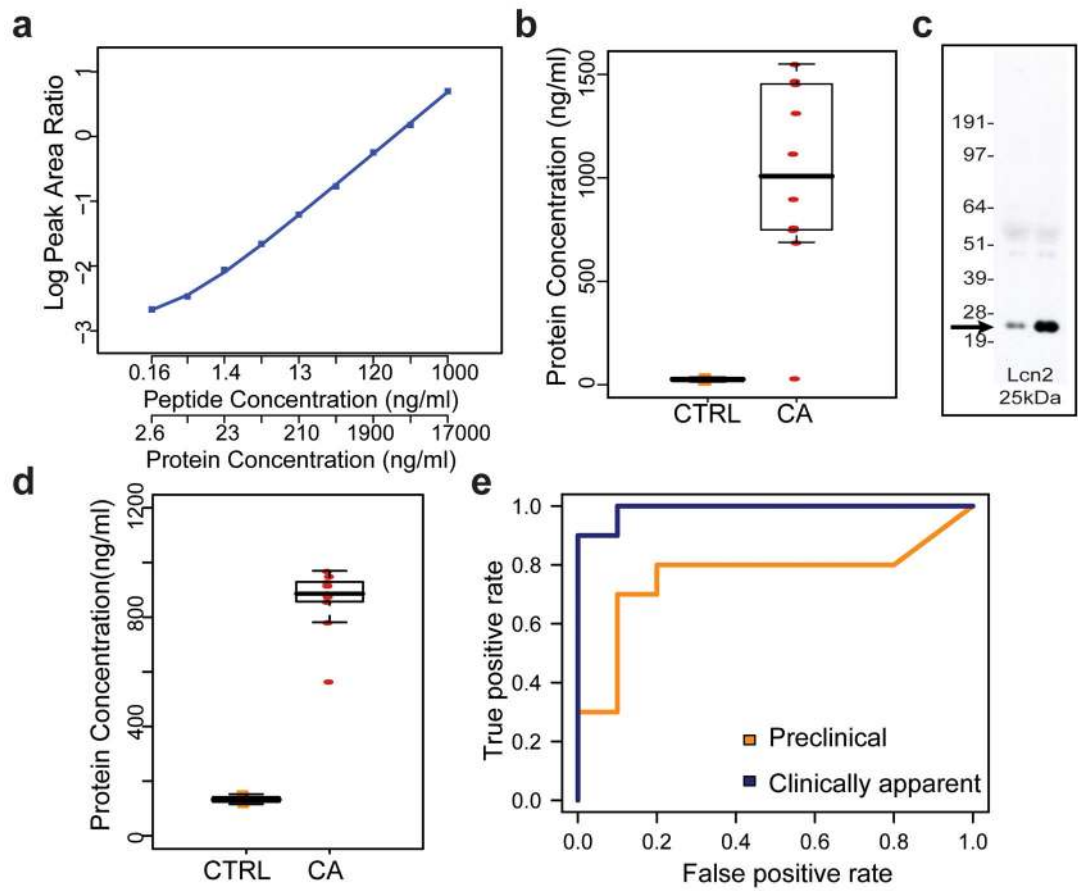


Figure 3.

SRM-MS, ELISA and western blot analysis data confirm elevation of Lcn2 in plasma of tumor-bearing animals. (a) Response curve for the SRM-MS assay targeting a proteotypic peptide from Lcn2. The analytic %CV is 1.9% (at the LOQ), LOD is 10.5 ng protein/ml plasma and LOQ is 22.5 ng protein/ml plasma. (b) Distributions of concentrations of Lcn2 (inferred from the peptide concentration and assuming 100% recovery) in the mouse cohort with clinically apparent tumors (CA) and controls (CTRL). The median value is plotted as a line with each box displaying the distribution of the inner quartiles and whiskers showing 95% of the data. Results derived from the SRM-MS-based measurements in Supplementary Worksheet 6a. (c) Western blot analysis showing elevation of Lcn2 in plasma pools of tumor-bearing animals (lane 2) compared with healthy controls (lane 1). (d) Distributions of the concentrations of the Lcn2 in the mouse cohort with clinically apparent tumors (CA) and matched controls (CTRL), determined using a commercially available ELISA assay. The median value is plotted as a line with each box displaying the distribution of the inner quartiles and whiskers showing 95% of the data. Note that these are not the same plasma samples analyzed by SRM-MS in b. (e) ROC curve generated from the SRM-MS-based measurements of Lcn2 in the clinically apparent tumor cohort (blue line) and the preclinical mouse cohort (gold line).

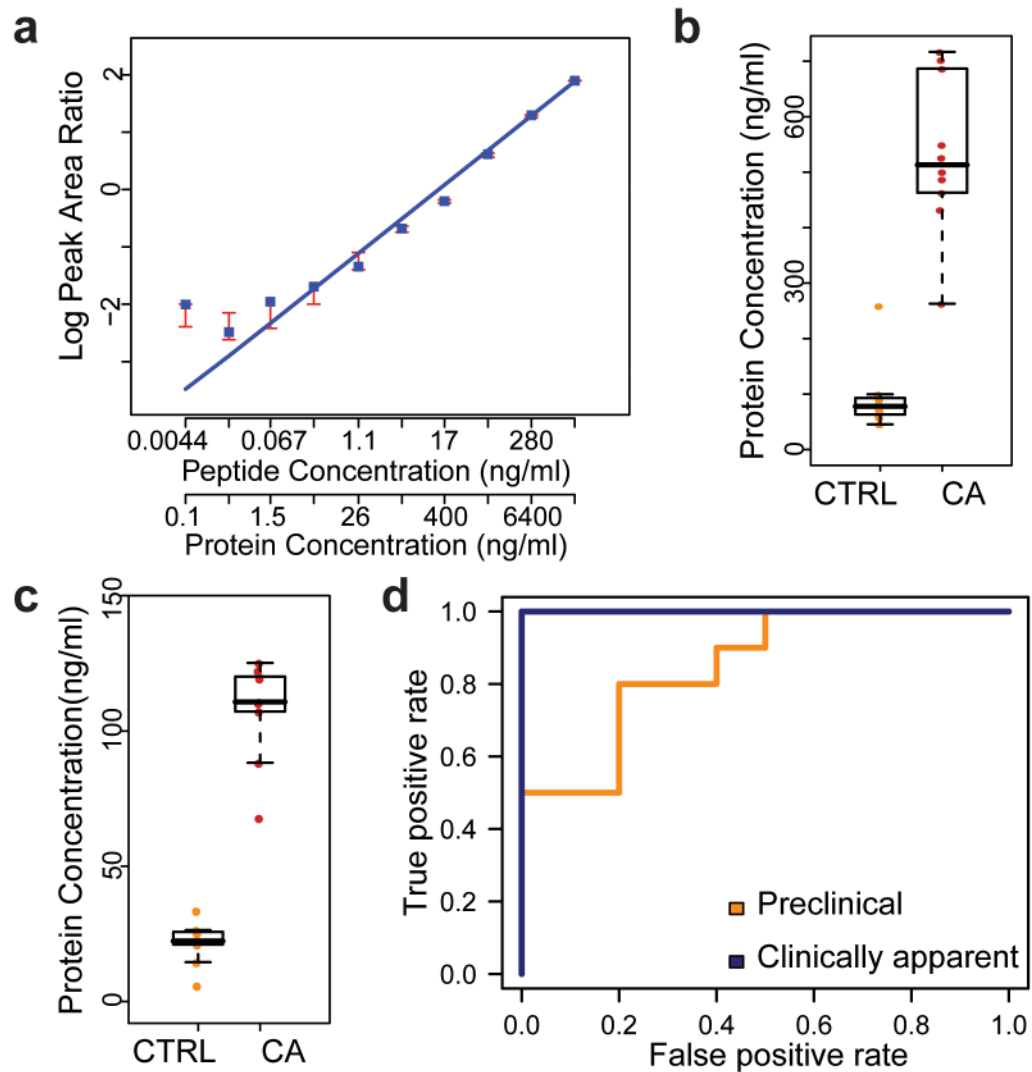


Figure 4.

Immuno-SRM and ELISA data confirm elevation of Mfge8 in plasma of tumor-bearing animals. (a) Response curve for the immuno-SRM-MS assay targeting a proteotypic peptide from Mfge8. The %CV is 10% (at the LOQ), LOD is 20.9 ng protein/ml plasma and LOQ is 48.4 ng protein/ml plasma. (b) Distributions of the concentration of Mfge8 (inferred from the peptide concentration and assuming 100% recovery) in the mouse cohort with clinically apparent tumors (CA) and matched controls (CTRL). The median value is plotted as a line with each box displaying the distribution of the inner quartiles and whiskers showing 95% of the data. Results shown are derived from the immuno-SRM-MS-based measurements in Supplementary Worksheet 6a. (c) Box plots show the distributions of the concentration of the Mfge8 protein in the mouse cohort with clinically apparent tumors (CA) and matched controls (CTRL), determined using a commercially available ELISA assay. These are the identical plasma samples analyzed by immuno-SRM-MS in b. The median value is plotted as a line with each box displaying the distribution of the inner quartiles and whiskers showing 95% of the data. (d) ROC curve generated from the immuno-SRM-MS-based

measurements of Mfge8 in the clinically apparent tumor cohort (blue line) and the preclinical mouse cohort (gold line). The plot was derived from the immuno-SRM-MS data in Supplementary Worksheet 6a,b.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1
Metrics of performance and verification results for quantitative SRM-based assays

Gene/Peptide	LOD			LOQ			Preclinical			Clin. Apparent		
	fmol	Peptide (ng/ml)	Protein (ng/ml)	fmol	Peptide (ng/ml)	Protein (ng/ml)	Median %CV	LOQ %CV	AUC	P-value	AUC	P-value
<i>2310057J18_VLYD...</i>	103.3	198.5	2636.7	261.7	502.9	6679.3	10.3	10.6	0.64	0.169	0.82	0.005
<i>Aldoc_TPSA...</i>	8.1	9.0	214.0	22.7	25.0	595.9	5.5	3.1	0.59	0.488	0.97	0.000
<i>Arhg4ib_TLLG...</i>	0.6	0.9	9.2	1.4	2.0	21.2	8.3	11.2	0.70	0.092	0.51	0.764
<i>Cabr_EQFL...</i>	0.5	0.5	17.3	1.4	1.3	43.4	5.8	9.7	0.59	0.508	0.63	0.326
<i>Camp_AVDD...*</i>	0.8	1.0	10.7	1.8	2.3	23.6	9.5	5.6	0.54	0.746	0.70	0.144
<i>Cdh1_DTAN...</i>	9.0	12.2	586.7	24.3	33.2	1592.3	4.6	7.2	0.56	0.654	0.65	0.281
<i>Chi3l1_SYGF...</i>	25.5	30.2	731.4	66.7	78.8	1911.2	10.8	10.4	0.58	0.547	0.96	0.000
<i>Comp_ELQE...</i>	0.4	0.4	21.0	0.8	0.7	41.4	2.9	4.2	0.73	0.080	0.89	0.000
<i>Csb_VAFG...</i>	12.5	14.9	310.6	27.8	33.3	691.8	6.6	7.7	0.51	0.920	0.86	0.008
<i>Csl_ENGG...</i>	1.2	1.5	30.9	2.7	3.2	68.2	3.1	3.9	0.55	0.669	0.95	0.000
<i>Dsg2_SIGN...</i>	1.1	1.3	86.5	2.5	3.1	205.3	3.5	6.7	0.56	0.689	0.97	0.000
<i>Dsm_YALY...*</i>	0.1	0.1	1.2	0.2	0.2	2.7	8.1	6.6	0.53	0.813	0.80	0.035
<i>Egfr_EISD...</i>	2.8	2.5	247.3	5.8	5.3	518.8	2.9	3.1	0.76	0.050	0.72	0.095
<i>Ext2_YVDD...</i>	2.8	2.9	154.7	7.8	7.9	424.3	3.0	2.4	0.53	0.818	0.96	0.000
<i>Fcna_YSSF...</i>	0.4	0.4	9.8	0.8	0.8	20.0	3.3	3.3	0.71	0.090	0.69	0.139
<i>Fetub_GSIQ...</i>	8.2	11.2	233.9	20.4	27.9	581.6	5.3	4.0	0.58	0.577	0.86	0.000
<i>Frl_SYTL...</i>	0.6	0.6	105.4	1.1	1.2	203.1	3.7	12.4	0.54	0.744	0.96	0.000
<i>Gdi2_FVSI...</i>	0.3	0.2	8.5	0.5	0.4	17.7	3.9	7.2	0.62	0.296	0.50	0.980
<i>Glod4_LELQ...</i>	1.8	2.2	39.0	3.9	4.8	86.7	5.1	8.6	0.55	0.694	0.74	0.087
<i>Grr1_DVDF...</i>	2.1	2.5	89.0	3.5	4.1	147.8	5.4	4.9	0.55	0.694	0.59	0.445
<i>Hsp90b1_GVVD...</i>	3.3	3.2	201.9	8.8	8.7	543.6	4.2	1.7	0.68	0.184	0.97	0.000
<i>Hspa5_NELE...</i>	3.1	2.8	151.7	8.5	7.4	410.1	2.3	0.9	0.52	0.906	0.87	0.005
<i>Hspa8_DAGT...</i>	2.7	2.2	124.3	7.1	5.7	324.6	5.1	10.1	0.52	0.873	0.55	0.721
<i>Ith3_EVSF...*</i>	5.0	3.9	330.3	10.0	7.7	660.6	13.2	11.7	0.68	0.169	0.97	0.000
<i>Kng2_LQSG...*</i>	0.4	0.3	16.6	0.8	0.7	35.9	5.2	3.1	0.65	0.271	0.94	0.000
<i>Lbp_HIQL...</i>	60.7	62.0	2156.3	166.9	170.7	5931.9	20.0	20.4	0.54	0.759	0.92	0.000
<i>Lcn2_WYVV...</i>	0.7	0.6	10.5	1.5	1.4	22.5	2.9	1.3	0.76	0.050	0.99	0.000

Gene/Peptide	LOD				LOQ				Preclinical				Clin. Apparent		
	fmol	Peptide (ng/ml)	Protein (ng/ml)	fmol	Peptide (ng/ml)	Protein (ng/ml)	Median (ng/ml)	Median %CV	LOQ %CV	AUC	P-value	AUC	P-value	AUC	P-value
<i>Lcp1</i> . VYAL...	0.7	0.7	32.9	1.5	1.6	71.9	71.9	5.8	6.1	0.60	0.485	0.87	0.005	0.87	0.005
<i>Lgal3bp</i> . LASA...	9.8	11.6	420.2	24.3	28.8	1045.9	1045.9	4.2	3.6	0.64	0.346	0.50	0.975	0.50	0.975
<i>Mrc2</i> . TLGD...	0.7	0.6	74.8	1.7	1.5	189.4	189.4	8.3	9.1	0.77	0.027	0.64	0.318	0.64	0.318
<i>Mug1</i> . DPSS...	0.3	0.3	38.0	0.8	0.8	90.3	90.3	5.5	5.2	0.66	0.226	0.93	0.000	0.93	0.000
<i>Nrp1</i> . SFEG...	1.2	1.3	81.0	2.9	3.2	199.4	199.4	3.2	8.3	0.63	0.363	0.88	0.000	0.88	0.000
<i>Orn2</i> . IFAD...*	16.4	11.3	261.3	39.7	27.3	631.5	631.5	15.3	11.6	0.77	0.032	0.90	0.000	0.90	0.000
<i>Spp1</i> . GDSL...	2.5	1.6	53.2	5.6	3.5	120.5	120.5	3.0	3.5	0.94	0.000	0.845	0.010	0.845	0.010
<i>P4hb</i> . ILEF...	0.8	0.5	29.2	1.6	1.0	61.6	61.6	3.7	4.6	0.52	0.828	0.94	0.000	0.94	0.000
<i>Papln</i> . GYNQ...	0.6	0.7	55.5	1.4	1.6	131.8	131.8	5.1	5.7	0.73	0.065	0.96	0.000	0.96	0.000
<i>Pdia3</i> . ELND...	0.8	0.8	31.1	1.3	1.2	48.3	48.3	6.5	7.9	0.60	0.480	0.97	0.000	0.97	0.000
<i>Pjhl</i> . DSPS...*	0.1	0.1	1.1	0.3	0.2	2.5	2.5	8.0	8.8	0.52	0.833	0.73	0.105	0.73	0.105
<i>Postn</i> . DQGP...	4.7	5.5	294.5	12.2	14.1	754.8	754.8	14.9	5.5	0.69	0.174	0.67	0.224	0.67	0.224
<i>Ppic</i> . TVEN...*	0.5	0.4	7.1	1.2	1.1	18.9	18.9	7.7	2.7	0.54	0.786	0.91	0.000	0.91	0.000
<i>Proc</i> . TFIL...	0.8	0.5	27.7	1.9	1.3	64.3	64.3	5.6	4.1	0.81	0.015	0.68	0.172	0.68	0.172
<i>Psm2</i> . YNED...	4.8	7.0	82.6	10.5	15.4	182.0	182.0	4.3	4.1	0.73	0.060	0.6	0.448	0.6	0.448
<i>Psm4</i> . LLDL...	1.7	1.4	33.7	3.6	2.9	70.3	70.3	3.7	5.5	0.72	0.102	0.74	0.057	0.74	0.057
<i>Plk7</i> . INSV...	0.9	0.9	67.5	2.0	2.0	156.8	156.8	4.0	6.4	0.64	0.306	0.94	0.000	0.94	0.000
<i>Pprf</i> . YSIG...	0.5	0.5	66.6	1.2	1.3	168.5	168.5	11.0	11.7	0.53	0.786	0.99	0.000	0.99	0.000
<i>Sprint</i> . EGTY...	1.1	0.9	43.4	2.7	2.2	100.9	100.9	3.6	3.8	0.54	0.744	0.93	0.000	0.93	0.000
<i>Th1</i> . ILAQ...	0.9	0.9	170.9	2.1	2.1	386.7	386.7	6.5	6.5	0.70	0.132	0.67	0.189	0.67	0.189
<i>Tnc</i> . TTLT...	0.4	0.5	57.9	0.7	0.9	104.9	104.9	2.6	8.2	0.59	0.520	0.93	0.000	0.93	0.000
<i>Ywhaz</i> . TAFD...	15.3	21.8	283.5	40.2	57.1	744.6	744.6	2.7	2.8	0.59	0.530	0.65	0.256	0.65	0.256

Three to four transitions were measured for each target in the equivalent of 1.5 µl of depleted plasma, and three repeats (three LC-SRM-MS analyses) were performed at ten concentration points (and a blank sample). The best transition for each target was selected as the basis for the quantitative assay. A linear regression algorithm was used for fitting the ten serial dilution data points for each curve. The LOD of each target was obtained from the average of the blank measurements plus three times the s.d. The LOQ was obtained from the average of the blank measurements plus ten times the s.d. with additional criteria that the CV of measurements of the concentration points near the calculated LOQ were <25%^{40,41}. Protein LOD/LOQ estimates assume complete trypsin digestion.

Assays followed by an asterisk (*) were characterized by reverse curves, as described in Supplementary Results Section 4. ROC curves and corresponding AUC values were determined for each candidate, and P values were determined by permutation testing, as described in Supplementary Results Section 6. Candidates determined to be significantly elevated (that is, verified; AUC ≥0.8 and P ≤0.01 in the plasma of animals with clinically apparent tumors) are in boldface. (Although Aldoc, Chi311 and Lbp meet significance based on P-values, their levels are near their assay LODs.) Peptide sequences are abbreviated to the first four amino acids (full sequences are shown in Supplementary Worksheet 4a).

Table 2

Metrics of performance and verification results for immuno-SRM assays

Gene/Peptide	LOD				LOQ				Preclinical			Clinical Apparent		
	fmol	Peptide (ng/ml)	Protein (ng/ml)	fmol	Peptide (ng/ml)	Protein (ng/ml)	fmol	Peptide (ng/ml)	Protein (ng/ml)	LOQ %CV	AUC	P-value	AUC	P-value
<i>Agm</i> . SELF...	24.0	2.4	498.6	61.7	6.2	1280.3	18.2	23.3	0.56	0.682	0.54	0.672		
<i>Anxa5</i> . ETSG...	12.2	2.0	43.4	25.8	4.2	92.2	13.2	3.8	0.54	0.779	0.50	0.983		
<i>Aric</i> . TLHP...	3.4	0.5	22.1	7.5	1.0	48.1	15.7	13.1	0.50	0.498	0.82	0.025		
<i>B4galnt1</i> . SLAD...	1.6	0.5	9.6	2.8	0.8	16.4	19.2	14.3	0.50	0.498	1.00	0.000		
<i>Casp3</i> . SVDS...	2.2	0.3	6.8	4.1	0.6	12.9	12.1	9.8	0.65	0.256	0.80	0.015		
<i>Cc2</i> . VQDD...	8.0	1.8	45.9	16.6	3.8	95.3	11.2	4.9	0.64	0.279	0.72	0.107		
<i>Cdc37</i> . LGPG...	13.7	3.1	61.1	34.9	7.9	155.6	13.3	7.3	0.59	0.567	0.73	0.057		
<i>Coro1c</i> . VTWD...*	48.6	7.5	258.0	98.3	15.1	522.1	14.3	6.2	0.51	0.920	0.53	0.818		
<i>Cot1l</i> . FALL...	1.2	0.2	1.9	1.8	0.3	2.9	8.4	8.2	0.61	0.398	0.61	0.418		
<i>Crk</i> . IGDQ...	2.4	0.5	8.1	4.1	0.8	13.8	9.3	4.5	0.66	0.254	0.76	0.072		
<i>Dpysl2</i> . DIGA...	7.9	1.9	49.2	14.0	3.3	87.2	15.9	10.6	0.70	0.134	0.67	0.226		
<i>Egfla1</i> . YYVT...	0.8	0.1	3.9	1.0	0.1	5.1	11.5	11.5	0.64	0.341	0.61	0.470		
<i>Exf1</i> . FTAV...	2.1	0.5	18.2	3.5	0.7	30.1	8.8	6.8	0.67	0.194	0.92	0.000		
<i>Gnb2l1</i> . FSPN...	9.3	1.8	32.7	22.3	4.3	78.2	11.9	8.9	0.62	0.415	0.58	0.540		
<i>Gnb2l1</i> . HLYT...	3.3	0.8	11.7	7.2	1.6	25.4	16.3	2.8	0.65	0.254	0.61	0.353		
<i>Iqgap1</i> . LEGV...	8.6	1.8	162.2	15.5	3.2	293.4	18.8	9.8	0.65	0.249	0.51	0.963		
<i>Isoc1</i> . YFGD...	5.3	0.7	16.8	11.4	1.4	36.6	7.7	4.9	0.51	0.925	0.73	0.082		
<i>Mgprbp1</i> . SQVN...	5.5	1.7	26.1	9.6	2.9	45.3	8.5	18.0	0.56	0.610	0.51	0.945		
<i>Man2a1</i> . DWVV...	5.8	0.7	75.9	13.4	1.6	176.1	9.4	4.7	0.57	0.572	1.00	0.000		
<i>Mjge8</i> . VAHS...	4.1	0.9	20.9	9.4	2.1	48.4	21.9	7.4	0.85	0.010	1.00	0.000		
<i>Paics</i> . ITSC...*	10.1	1.7	47.6	26.4	4.5	124.2	16.3	16.9	0.69	0.142	0.60	0.483		
<i>Pal2</i> . SVID...	5.6	1.1	32.2	12.1	2.3	69.9	11.2	5.8	0.63	0.328	0.52	0.888		
<i>Pcdh18</i> . SYEL...	4.4	1.0	55.0	10.1	2.4	126.8	7.1	6.0	0.56	0.644	1.00	0.000		
<i>Pcd6</i> . AGVN...	10.2	1.5	22.3	18.4	2.7	40.2	6.0	5.4	0.50	0.498	0.60	0.391		
<i>Pcd6ip</i> . FIQQ...	1.3	0.3	12.8	1.8	0.4	17.8	13.8	7.4	0.66	0.296	0.66	0.204		
<i>Pfas</i> . LGGT...	3.4	1.0	49.5	8.2	2.5	118.5	16.7	2.2	0.56	0.622	0.63	0.378		
<i>Ppp1r7</i> . CIEN...	54.0	8.1	223.0	128.8	19.4	531.9	42.1	7.8	0.51	0.920	0.67	0.219		

Gene/Peptide	LOD			LOQ				Preclinical		Clinical Apparent		
	fmol	Peptide (ng/ml)	Protein (ng/ml)	fmol	Peptide (ng/ml)	Protein (ng/ml)	Median %CV	LOQ %CV	AUC	P-value	AUC	P-value
<i>Prdx4</i> . ISKP...	6.1	1.3	18.8	14.9	3.1	46.3	4.1	8.8	0.54	0.741	0.93	0.005
<i>Pymb5</i> . DAYS...	4.6	0.7	13.2	7.4	1.1	21.2	8.6	2.2	0.75	0.057	0.69	0.147
<i>Pk3</i> . LESF...	35.1	4.8	146.6	86.4	11.8	361.1	13.6	4.5	0.67	0.177	0.55	0.498
<i>Sec14l1</i> . EGES...	0.2	0.0	1.8	0.5	0.1	4.1	12.1	8.3	0.53	0.826	0.50	0.498

Three to four transitions were measured for each target in the equivalent of 10 µl of digested plasma, and three repeats (three SISCAPA process replicates) were performed at ten concentration points (and a blank sample). The best transition for each target was selected as the basis for the quantitative assay. A linear regression algorithm was used for fitting the ten serial dilution data points for each curve. The LOD of each target was obtained from the average of the blank measurements plus three times the s.d. The LOQ was obtained from the average of the blank measurements plus ten times the s.d. with additional criteria that the CV of measurements of the concentration points near the calculated LOQ were <25%^{40,41}. Protein LOD/LOQ estimates assume complete trypsin digestion. Assays for candidates followed by asterisk (*) were characterized by reverse curves, as described in Supplementary Results Section 5. ROC curves and corresponding AUC values were determined for each candidate, and P-values were determined by permutation testing (Supplementary Results Section 6). Candidates determined to be significantly elevated (that is, verified; AUC ≥0.8 and P ≤0.01) in the plasma of animals with clinically apparent tumors are in boldface. Peptide sequences are abbreviated to the first four amino acids (full sequences are shown in Supplementary Worksheet 5a).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Accession codes for ProteomeCommons.org

Table 3

Data set	Description	Tranche hash string
AIMS and LC-MS/MS of SCX fraction 1	AIMS and LC-MS/MS RAW data	K3lKd96CyN8D6desXTBSNly9qnLgEOu C3uH0CRl+HDS1vkEodIjPoXqv7k2q8P 3iXBMT7T36wzmmTqGRMue54PEuAAA AAAAAAAeIQ==
AIMS and LC-MS/MS of SCX fraction 2	AIMS and LC-MS/MS RAW data	ehRyE2rw4mrfkVvP+AEzgfWCOQOfgpmI bCousVQbbcpit8rAR3qu8MsEKwg6x X+OelIAS9r652Wc6Hw7ehWatyhcAAA AAAAAQmA==
AIMS and LC-MS/MS of SCX fraction 3	AIMS and LC-MS/MS RAW data	9/rjdiG2NIVFBw0qQ6FL+V4vJMCuUoU 6jdlk3FEbLmXlpEjvGtBU/2h4DG0eY1XNBj CtrMzxpjd86U96n+D2VpCdf1aYcAAAAA g7g==
AIMS and LC-MS/MS of SCX fraction 4	AIMS and LC-MS/MS RAW data	Etrqd2AkdAcUMmmPnmMCxh7Ak0vHe6/1 6S7/sml8uz5ncpxmRL TTSf5GoUu/re02EC Qvzv4Zn9z3ZSR+E7xNkssqN7j8moAAA AAAAAh5Q==
AIMS and LC-MS/MS of SCX fraction 5	AIMS and LC-MS/MS RAW data	hg6it116AddlZ4hG3Poiup3n9aj1NqKg d04TpWlfgUqjBSAPeU4FOGQl0V1YC 3cA3LaYwZOWrBT441c6eygAAAAA A8Q==
AIMS and LC-MS/MS of SCX fraction 6	AIMS and LC-MS/MS RAW data	mK1cV+hKdZWRkOvty20dR4LOvJ2CE OOS5SRmu/Ry auMzIE2ay23TH+0E6p94G fLjij9H65+ASL1PUu lU0cs5Z6+MUgAA AAAAAA8Q==
AIMS and LC-MS/MS of SCX fraction 7	AIMS and LC-MS/MS RAW data	g3AUjwyHrtcdHmF5puFSf0CdHSrHkb7H 01uRbj0oTQYStTYRGzh9sCZCQe8E3IPX 4rgieHQkg2pMjhzIcRIZjsAAA AAAAAA8Q==
AIMS and LC-MS/MS of SCX fraction 8	AIMS and LC-MS/MS RAW data	cxY08d9hxVqGN6GIA0ZR/+U0khhGL8SV RAMdq46x6ovoMyKoo0NMVY9UUV/TO+ k6L6Gjii:7TEvxl/wx9uu7Ormm+qIAAAAAA AAF8Q==
AIMS and LC-MS/MS of SCX fraction 9	AIMS and LC-MS/MS RAW data	cjitrLqD3wTBYAIItoQA+dzITVWWYJrwgUzGzT4Y 9vnyZY48m6C6uqr 316FqjRwyZP4QNjhrV15hNub65jxfx+ jFhs+8AAAAA AAF8Q==
AIMS and LC-MS/MS of SCX fraction 10	AIMS and LC-MS/MS RAW data	AhG0zSjIacyC4j+pww6dbbBf8GDxQHvRz+ Nwcldl11wWPHqwhiHI/QDvNBfFRS7pdK 5ZLh8Zwz30bHhY vrbha60erA3cAAA AAgDA==
AIMS and LC-MS/MS of SCX fraction 1	AIMS and LC-MS/MS mzXML data	T61O2Omx.A97AvDQ2099GblLUNi0w14 oNgKZuFfIK0CM9/iBv2pPutYlmb+i4G TqCuGkHeXEI0MDEXS34EPo9irZ7EAAA AAAAAAL8w==
AIMS and LC-MS/MS of SCX fraction 2	AIMS and LC-MS/MS mzXML data	LoHl/2J49bDTJOOmMk4UooWsoFL61A0 eexVz7KVHIoKvz911WvIC6vPohnEtwjVPT QQ50ccPWV2s4y3oFXIEFHU0Ylc0AAAA AAAAAsvw==
AIMS and LC-MS/MS of SCX fraction 3	AIMS and LC-MS/MS mzXML data	hQ8L5AbeWvRHy9Lunkx6n/cbfcmyraDZB nMwcOWAqoDHNtUEBYOOGu10w1698E YXu63CW17PRZBD75m5VmlBiUMH4A AAAAAAIXQ==
AIMS and LC-MS/MS of SCX fraction 4	AIMS and LC-MS/MS mzXML data	5bDXI5/18jpk3j9gmI4hlL6SBPEDdi8KG d7V.sP15+OToODIQ+qdxKXJw1VLphlpY4 YCzSAZLr2amRZjZICZsLxgGOV4AAAAA AAXIXQ==
AIMS and LC-MS/MS of SCX fraction 5	AIMS and LC-MS/MS mzXML data	dge2su595zuURoADids8A/PE6REgxsCs N3n3fRPDm+w9luMhqFqsRPOAp5MJetly OMQcwDcRk+TatQ3206sLh1qD+/YAAAA AAAAAgYQ==
AIMS and LC-MS/MS of SCX fraction 6	AIMS and LC-MS/MS mzXML data	O5YapJumqtXlpeLbiRPvkMdmK5MHOlq 7UYzzDjF08vOu5PZCmrJT2Pq2MjVjFz76uHSFit9IXg6RswCdePwH3MAAAA AAAAAgYQ==
AIMS and LC-MS/MS of SCX fraction 7	AIMS and LC-MS/MS mzXML data	v4Tnih9rX+QugLO41s/h16puHrrn0XyZOiu Y9dPm2Kj5mS8NRD4hrTriKbv28eCIXJX6 2Nj1E+x26wZQ399j0Y vIAAAAAA AAAAAgYQ==
AIMS and LC-MS/MS of SCX fraction 8	AIMS and LC-MS/MS mzXML data	eqCLA6T8Tooptu5b1pFnV2A6h+WN1pspYkaj b8ia/v9RnKxvcVNoDexbYwrmfWJn5wmwZM 2Zct6MxIqIG0zuXclvMAAAAAA AAAAAgYQ==
AIMS and LC-MS/MS of SCX fraction 9	AIMS and LC-MS/MS mzXML data	atBRwAIMFiao9s1W3FmWB8p4/Z52pbP EQLMjjsKHIXVfsUnkNOB9cEIXV4B+qrU HPZ8w8TJrT8DntWzMKF6+f5h0gxYAAA AAAAAgYQ==
AIMS and LC-MS/MS of SCX fraction 10	AIMS and LC-MS/MS mzXML data	0RV8VHh/OYNqsnjSUKsmwOXB4lyzCza GWjeGnEYWwIPEXkEK2m5wMN6W8jntQ AAAAAA AAgfA==
mouse_plasma_qmmm_siscapa_data	Integration results r1h4keM2uJwCRrPe9dS5IP91eDvVFyuV from response curves	JGBayQr7Y8H8r9+cugheWgCtpr8HAHC xbCPK6rBFa/4JlhC9EBv0lulnFZA AAAAA AABawQ==