Building a Robust and Reliable Proteomics Biomarker Pipeline: NCI's CPTAC Network

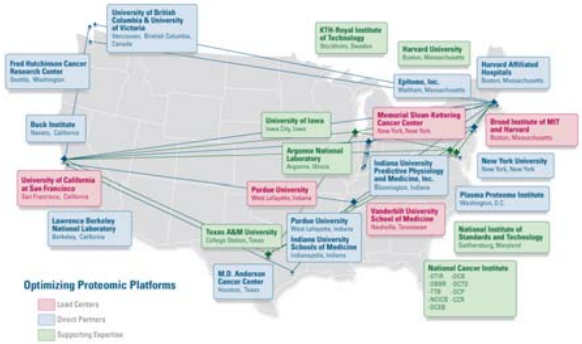
Christopher R. Kinsinger, Tara Hiltke, Mehdi Mesri, Henry Rodriguez
National Cancer Institute, Bethesda, MD

Abstract

Translation of candidate proteins as indicators of cancer status from discovery to validation biomarkers of clinical utility is a major barrier in cancer research. While the scientific literature annually reports many cancer biomarkers, few prove sufficiently resilient, reproducible, and real to translate into clinically useful assays. Consequently, the cancer research community is in need of a systematic approach to streamline the biomarker translation process from discovery to validation. We propose here a strategy to verify the biological signal of protein biomarker candidates. This biomarker verification step would provide a means to demonstrate that a biomarker's discriminative ability is due to biology as opposed to chance or bias in the samples or technology platform.

Such a pipeline requires three components: discovery, verification, and validation. NCI's Clinical Proteomics Technology Assessment for Cancer (CPTAC) is providing support for each step of this strategy through various means. Within the discovery step, CPTAC is developing metrics and reagents to assess the variability and reliability of technology platforms. CPTAC is conducting analogous studies within the verification step in order to benchmark MRM-based platforms for high-throughput, targeted studies. Additionally, CPTAC is collecting clinical samples for verification studies. These samples were collected through a novel protocol that minimizes bias. Finally, to support the validation step, CPTAC is producing reagents critical to clinical assay development. While still in process, CPTAC has made significant progress in all three pipeline components for a proteomics technology platform that targets peptides derived from protein candidates.

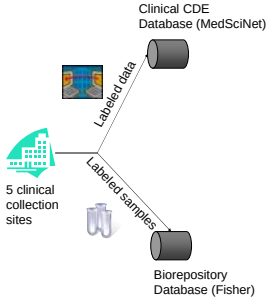
This verification strategy integrates the essential elements of biomarker development far ahead of clinical validation. Implementation of a this pipeline will streamline the biomarker development process, shorten the timeline, and improve the success rate to bring validated cancer biomarkers into the clinic.



Biospecimen plan – phase I

CPTAC center biospecimen collection:

- Plasma samples for technology validation
- 5 collection sites
- Common cohort: pre-biopsy women with a scheduled breast biopsy
  - -Minimize bias
  - -Resulting sample set with 25%/75% cancerous/benign lesions
- Common collection protocol: CPTAC plasma collection protocol
- Common processing, shipment, and storage: Fisher Bioservices
- Clinical common date elements stored in MedSciNet database
- In phase II, samples will be shipped to proteomic analysis labs to test the ability of technology platforms to reproducibly detect cancer-relevant features.



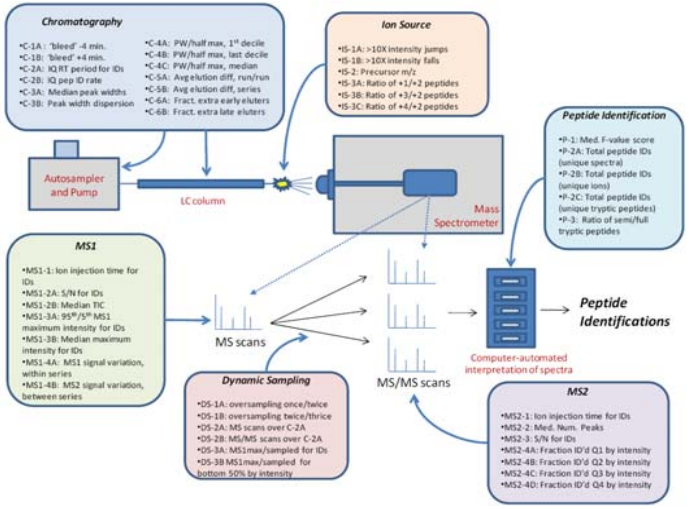
Path to Discovery

Goal
- Provide an assessment of run-to-run and lab-to-lab variability for a complex mixture of proteins
- Provide a measure of the sensitivity of a mass-spec platform with such a mixture

Material
- Yeast lysate spike with protein mixture (Sigma UPS)

Method
- A carefully designed, multi-lab experiment using replicate analyses of the yeast lysate
- Data was centrally analyzed on a common pipeline

Results
- Over 40 performance metrics, already available in instrument output files, may now be used to determine a discovery platform's condition for a given experiment (below)
- A whole new level of quality assurance for discovery proteomics



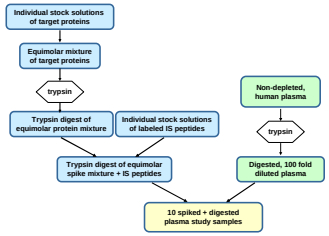
Path to Verification

Goal
- Through a team-based network, assess variability of the MRM analysis platform
- Determine the sensitivity of the MRM analysis platform

Material
- Seven target proteins spiked into stock solution

Method
- A carefully designed, multi-lab experiment using replicate analyses stock solution (flowchart below)

Results
- Over 40 performance metrics, already available in instrument output files, may now be used to determine a discovery platform's condition for a given experiment (below)
- A whole new level of quality assurance for discovery proteomics



Flowchart for sample preparation performed at NIST. An equimolar mixture of 7 target proteins and non-depleted, human plasma were reduced, alkylated and digested with trypsin separately. A calibration curve was generated by spiking the digested protein mixture (containing 100 fmol/µL of labeled internal peptide standards) into 100-fold diluted, digested plasma. All protein and peptide stock solutions were analyzed by amino acid analysis prior to sample preparation.

Protein	Sequence	% Recovery	CV	Slope	CV	T. LOD (pmol/µL)	T. LOD (ng/mL)	CV
CRP	ESDTSYVSLK	13.77	4.2	0.972	1.0	0.0027	58.1	44.8
PSA	LSEPAELTDAVK	12.50	7.9	0.988	1.7	0.0053	152.3	9.3
CRP	GYSIFSYATK	0.13	7.2	0.800	0.9	0.2289	4921.4	19.2
PSA	IVGGWECEK	5.33	8.0	0.971	2.8	0.0135	388.2	7.16
MBP	HGFLPR	17.42	8.0	0.987	2.5	0.0028	51.3	78.6
MYO	LFTGHPETLEK	21.00	7.8	0.963	1.7	0.0017	28.8	16.5
APR	AGLCQTFVYGGCR	0.84	11.0	0.919	3.1	0.0591	384.8	10.9
HRP	SSDLVALSGGHTFGK	0.13	25.7	0.845	5.1	0.2069	8033.1	24.4
LEP	INDISHTQSVSAK	5.88	22.6	0.934	3.6	0.0093	148.8	30.9
MBP	YLASASTMDHAR	0.16	38.9	0.915	9.2	0.1769	3241.3	38.7

Path to Validation

Goal
- Supply research community with robust, highly characterized monoclonal antibodies to cancer-relevant targets

Acknowledgments

This work was supported by grants from the National Cancer Institute (U24CA126476-01, U24CA126485-01, U24CA126480-01, U24CA126477-01, and U24CA126479-01), part of NCI's Clinical Proteomic Technologies for Cancer (http://proteomics.cancer.gov). A component of this initiative are the Clinical Proteomic Technology Assessment for Cancer (CPTAC) teams that include the Broad Institute of MIT and Harvard, Memorial Sloan-Kettering Cancer Center, Purdue University, University of California, San Francisco, and Vanderbilt University School of Medicine.

Dr. Paul Rudnick of NIST for analysis and visualization of CPTAC data

Data available at
-http://cptac.tranche.proteomecommons.org
-http://proteomics.cancer.gov
-http://antibodies.cancer.gov

NATIONAL CANCER INSTITUTE

CLINICAL PROTEOMIC TECHNOLOGIES FOR CANCER