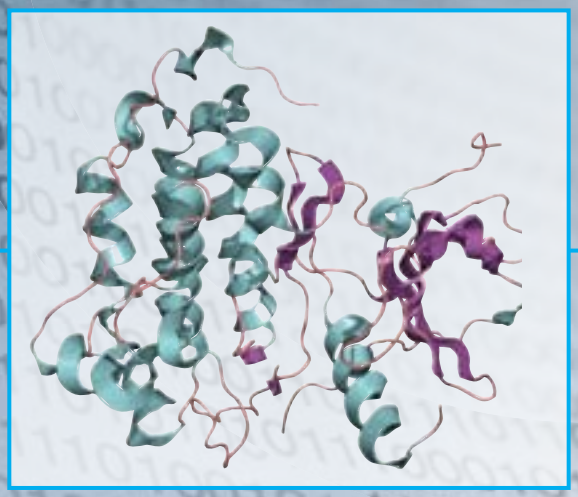
National Cancer Institute

CLINICAL PROTEOMIC TECHNOLOGIES FOR CANCER

2007 | ANNUAL REPORT



Building the Foundation for Clinical Cancer Proteomics

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

National Cancer Institute

Office of Technology & Industrial Relations

ATTN: Clinical Proteomic Technologies for Cancer

31 Center Drive, MSC 2580 | Bethesda, MD 20892-2580

Email: cancer.proteomics@mail.nih.gov

Web site: http://proteomics.cancer.gov

Table of Contents

Table of Figures

From the NCI Director

The American Cancer Society projects that almost 1.5 million cases of cancer will be diagnosed in the United States this year alone. Although we have made progress in treating many forms of cancer, it is painfully clear that we have a long way to go. Our best hope lies in being able to detect cancer early, allowing us to treat it before it can exert its devastating personal and societal effects.

The greatest promise for early detection of cancer lies in the ability to find valid molecular indicators (or biomarkers) of the disease. Progress in cancer genetics has been rapid, but it only gives us a predictive ability: We need measurements of what is happening in a patient in real time, and that means finding tell-tale protein and peptide biomarkers. Despite a great deal of work, this has proven an extremely difficult undertaking. At least part of the difficulty lies in technological and methodological variability, the extent of which we still do not understand entirely. We launched the Clinical Proteomic Technologies for Cancer initiative precisely for this reason: Our challenge is to provide the entire cancer community with the resources and tools necessary to overcome the technological and methodological barriers so that we can find those elusive markers that will lead us to our ultimate goal of reducing the burden of cancer. The progress in just the first year of this program, detailed in these pages, has been gratifying and attests to the importance of this work for the entire field of proteomics.

Everything we do at the National Cancer Institute begins and ends with the cancer patient. This singular focus gives all of our research both a sense of urgency and a sense of purpose. It is important that we remind ourselves daily that we are working for a vast number of individuals 1.5 million this year alone who we may never meet but who are counting on us to help them.

John E. Niederhuber, M.D.
Director, National Cancer Institute

From the CPTC Director

Challenges to Unleashing the Potential of Proteomics for Cancer

The tremendous advances in genetics and genomics over the past decade hold great promise for understanding, treating, and even preventing at least some forms of cancer. However, genes are only the "recipes" of the cell: The proteins encoded by the genes are ultimately the critical molecular players that drive both normal and disease physiology.

The finding that tumors "leak" proteins, peptides, and other molecules into blood, urine, and other accessible bodily fluids has led to the possibility of diagnosing cancer at an early stage simply by collecting such fluids and testing them for the presence of cancer-related proteins and peptides, or "biomarkers." Such biomarkers might also be valuable for monitoring the response to cancer during treatment or detecting the recurrence of tumors after treatment. Indeed, some blood-borne proteins are already being used as cancer biomarkers. For example, elevated levels of prostate specific antigen (PSA) suggest the presence of prostate cancer, while elevated levels of cancer antigen 125 (CA-125) may indicate cancer of the ovary or other organs. Unfortunately, both tests may result in "false negatives"—failing to detect cancer in those who have it (poor sensitivity), or "false positives"—testing positive for the presence of cancer in people who are actually cancer-free (poor specificity) (FIGURE 1). And it is clear that even true positives do not always correlate with the presence of cancer.

We do not suffer a lack of reported cancer biomarkers: The literature reports upwards of 1,200 protein biomarkers, though very few of these have been validated, and even fewer have found their way into clinical practice (FIGURE 2). It has become increasingly clear that this dichotomy can be traced —in large part—to several levels of confounding variables (FIGURE 3).

FIGURE 1. Sensitivity and Specificity of Biomarker Testing

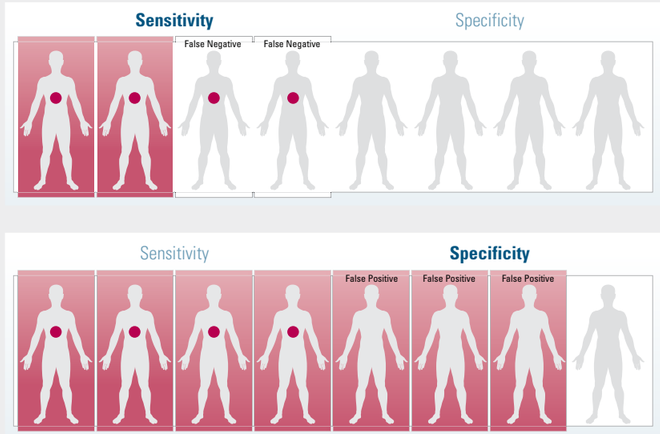
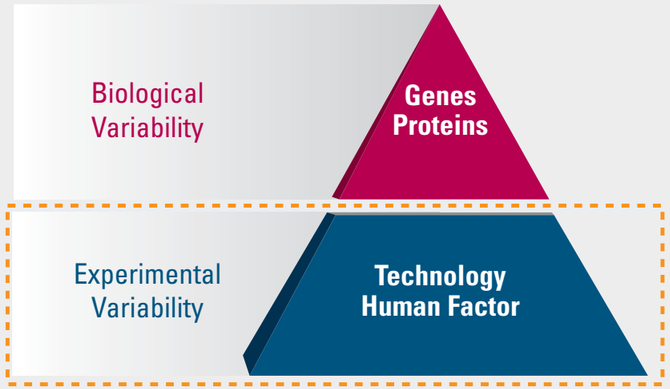


FIGURE 2. Clinical Development Pipeline

Very few cancer biomarker candidates have made their way beyond clinical trials and into patient care. The high number of biomarker discoveries reported indicates that there is a pipeline problem somewhere between discovery and the clinical translation phases.

FIGURE 3. Pyramid of Variability



The potential of proteomics for cancer detection and treatment requires that the experimental variability be reduced as much as possible—the goal of the CPTC initiative—so biological variability can be accounted for as completely as possible.

Many of these variables arise from the sheer complexity of the proteome. It is estimated that there are on the order of 100,000 to a million different proteins in the human proteome, many of which may be found in the bloodstream. These proteins are subject to a host of post-translational modifications that vary from person to person and even cell to cell, depending on constantly shifting environmental and micro-environmental conditions. Furthermore, proteins exist in a wide range of concentrations, over several orders of magnitude, making the lower abundance proteins (and the ones most likely to be the telltale signs of cancer) difficult to detect at best.

"...the proteome...is very different from cell to cell."

—Ruedi Aebersold, Ph.D.
Co-founder
Institute for Systems Biology

Emerging analytical technologies, particularly mass spectrometry and protein microarrays, carry the potential to give us the accuracy and reproducibility needed to make sense of the complex proteome. But proteomics data are being collected at a faster pace than the ability of the researchers to validate, interpret, and integrate them with other known data. The variety of platforms and standards of practice are introducing layers of variability that supplement the biological complexity. Furthermore, there is a lack of standard reagents for use by the entire proteomics community, creating further uncertainty in comparing experiments across labs or even different experiments within a single lab.

Finally, beyond the technological hurdles that researchers must overcome to maximize the use of proteomics for cancer research and diagnosis, there are also procedural and organizational hurdles. Different treatment centers are likely to collect and store samples in different ways, creating heterogeneity in the samples. In addition, the data are not always recorded, annotated, or even analyzed in the same formats or by the same methods, introducing yet additional confounding variables.

It is clear that the potential of proteomics for cancer detection and treatment requires that the non-biological sources of variability be eliminated, and that the biological variables be accounted for as completely as possible.

Recognizing the challenges facing the proteomics community, the National Cancer Institute (NCI) launched the Clinical Proteomic Technologies for Cancer (CPTC) initiative, a five-year program aimed at addressing and reducing the layers of variability that prevent progress in applying proteomic insight to clinical practice. This program was put together carefully in collaboration with the international proteomic community, through a series of meetings that outlined the key issues that needed to be addressed and offered potential solutions. As a result, CPTC is truly a reflection of what the community needs, which is defining proteomic platform performance parameters at every step of the biomarker discovery pipeline.

Specific needs identified include:

1. New technologies that can quantify proteins across the entire concentration range as well as detect modified versions of proteins;
2. Optimized proteomic technologies and development of appropriate standards;
3. Common bioinformatics resources, with shared algorithms and standards for processing, analyzing, and storing proteomic data;
4. Standardized procedures for processing and storing biological samples used in proteomics research; and
5. Available high-quality reagents.

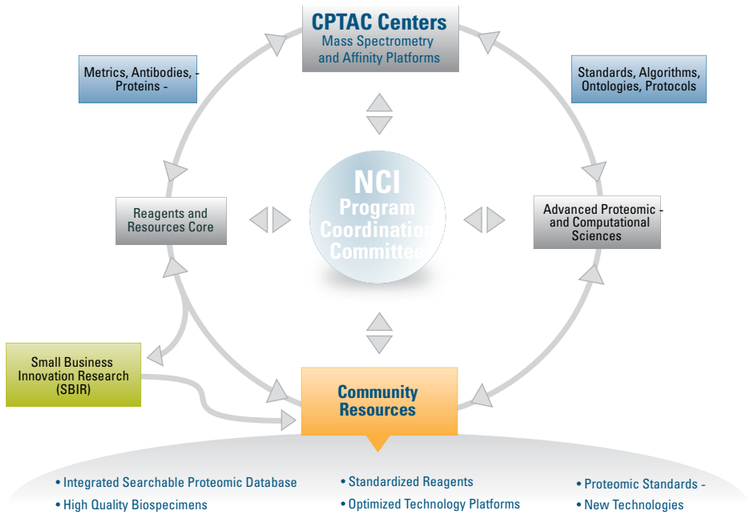
There are three major and integrated CPTC programs designed to address all of these needs (FIGURE 4): the Clinical Proteomic Technology Assessment for Cancer (CPTAC) program, the Proteomic Reagents and Resource Core, and the Advanced Platforms and Computational Sciences program. Each of these programs are described in greater detail below.

This document is a report on the first year of the five-year program. I am particularly gratified by the tremendous progress already made, which is a direct reflection of the CPTC members' dedication to the highest quality and standards and deep commitment to open and collaborative science for the sake of the entire cancer proteomics community. Their work will have mplications far beyond cancer proteomics, but their most lasting legacy will be the impact that work will have on reducing the burden of suffering and death due to cancer. And that is the ultimate reason we are all working so hard together to ensure the success of this program.

Henry Rodriguez, Ph.D., M.B.A.
Director, Clinical Proteomic Technologies for Cancer initiative
Office of the Director
National Cancer Institute

To learn more about this initiative, and to be kept up to date on our progress, please visit http://proteomics.cancer.gov.

FIGURE 4. Clinical Proteomic Technologies for Cancer–Team Science



CPTC Components:
The Clinical Proteomic Technology Assessment for Cancer (CPTAC)

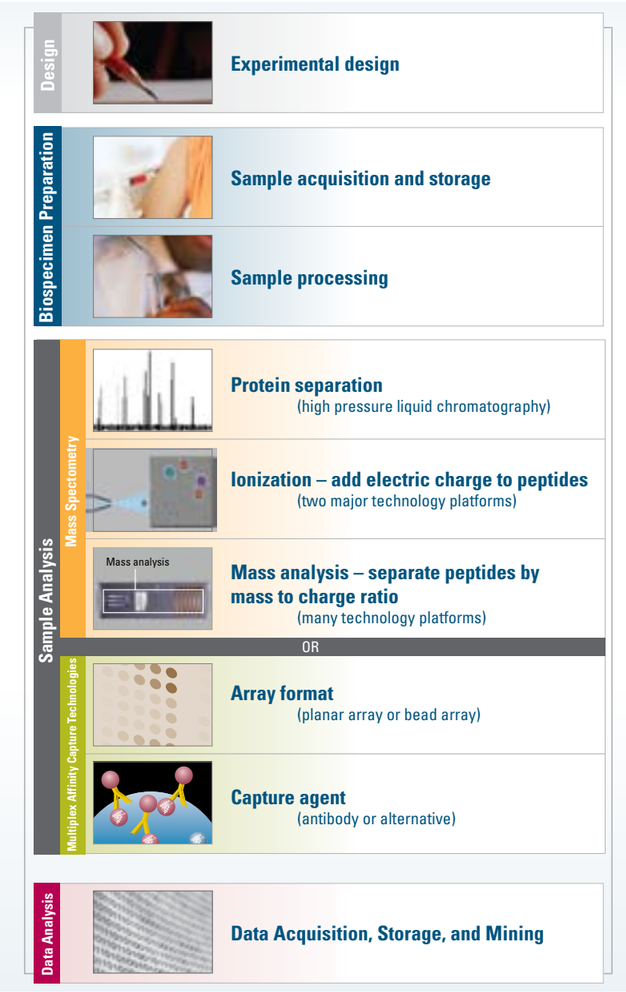
Multidisciplinary team network defining proteomic platform performance characteristics (using SOPs, reference materials, etc.) at every step of the biomarker candidate pipeline to reliably identify, quantify, and compare peptides/proteins in complex biological mixtures.

As the primary structural and functional components of the cell, proteins play vital roles in all cellular processes, including those associated with cancer. Consequently, understanding proteins and their interactions is critical to NCI's mission to reduce the burden of cancer. As a result, proteomic technologies can be used to solve mission-critical problems in cancer research, including detecting cancer processes, finding targets for novel therapeutics, and determining biological markers of treatment response.

However, in recent years, studies that have applied current protein measurement technology—including mass spectrometry (MS) and affinity-based detection methods—to clinical applications have not been as robust as had been hoped. The issue lies in the variability of the technology and its use.

Proteomic research is hampered by the variability that results from a lack of standardized technologies and methodologies, which are critically needed in order to more effectively discover and validate proteins and peptides relevant to cancer. "Shotgun" proteomics is an approach used to identify proteins in complex mixtures, and it is commonly used in candidate biomarker discovery experiments (FIGURE 5). This approach requires several complex steps, and these workflows include many areas

FIGURE 5. Candidate Biomarker Discovery Workflow



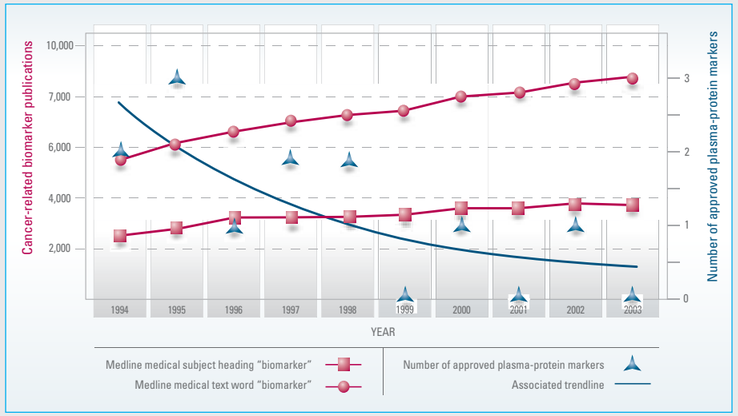
Each step within this workflow contributes to experimental variability. The goal of the CPTC initiative is to define proteomic platform performance characteristics (standard operating procedures, reference materials, etc.) at every step of the biomarker discovery pipeline, from sample collection to data analysis.

"Ultimately, we want to provide the proper tools to investigators conducting discovery proteomics so they can have the assurance that protein/peptide measurement results are due to changes in the biological sample and not to variability in the instrument, assay performance, reagents, operator, or site."

– Henry Rodriguez, Ph.D., MBA
Director, CPTC

Discovery leads to candidates, not biomarkers

FIGURE 6. Reality Check: Status of Protein Biomarkers for Cancer

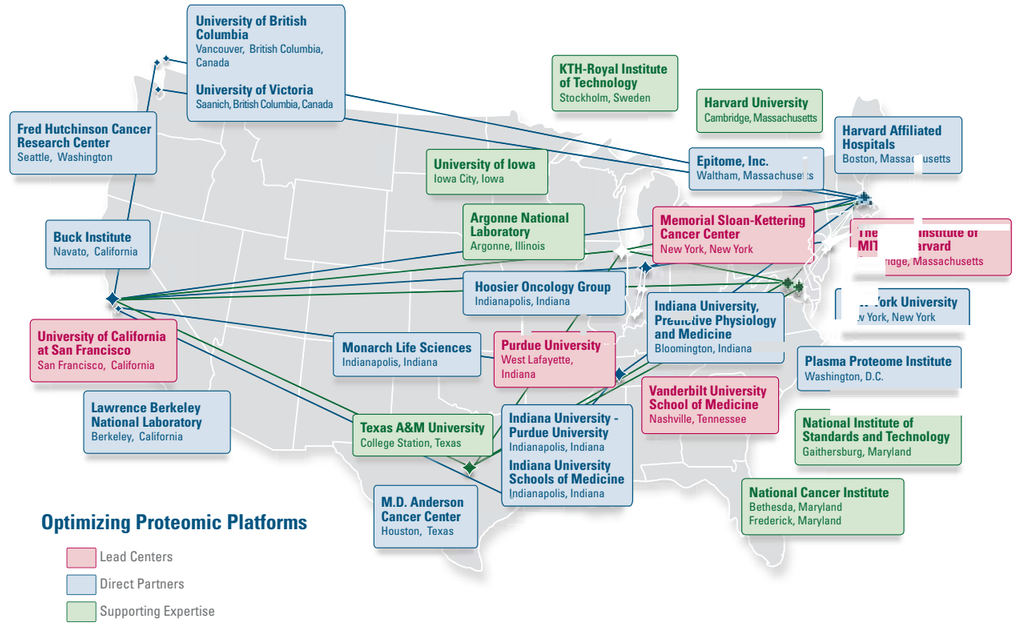


Ludwig and Weinstein. Biomarkers in Cancer Staging, Prognosis, and Treatment Selection. Nat Rev Cancer. 2005 Nov, 5(11):845-56.

"CPTC will not only standardize, but improve the quality of validation criteria that's essential to rapidly advance biomarkers into a clinical setting."

– Mark Boguski, M.D., Ph.D.
Harvard Medical School
Center for BioMedical Informatics

FIGURE 7: CPTAC Team Network



of potential variability, including the collection and storage of samples, sample preparation, chromatography, mass spectrometry, and data analysis. In addition to workflows, variability also results from the complexity of the proteomic technologies themselves.

Despite many claims for the discovery of cancer-related proteins or "biomarkers," it has proven very difficult to reproduce and validate results across either laboratories/institutions or technology platforms. If one laboratory has identified a list of biomarker candidates for the early detection of cancer, but the rest of the community cannot reproduce this list using the same patient samples, should these candidates be pursued in the clinic? Can the data be trusted? This is the reality that the cancer proteomics community is facing.

It is now necessary for protein biomarker candidates to first be verified in the laboratory before they are validated in the clinic. Biomarker verification, what the CPTC initiative would like to see implemented as a prerequisite for validation, tests the accuracy of biomarker candidates using mass spectrometry-based protocols and can help sift through this growing list of candidates. From a clinical perspective, improving the quality of candidates throughout the discovery and verification stages is a much more cost-effective way to enhance the number of successful biomarkers in the clinic. By reducing the human and experimental sources of variation, the proteomic community can improve the biomarker pipeline significantly by confidently providing greater numbers of "true," or disease-specific, candidates (FIGURE 6).

To address this critical need, the NCI established a collaborative network of five Clinical Proteomic Technology Assessment for Cancer (CPTAC) teams in September 2006. The network extends well beyond these five centers, bringing in expertise from both the public and private sectors to ensure that all of the expertise needed is brought together in a single focus (FIGURE 7).

The CPTAC collaborative studies were initiated with the goal of identifying, quantifying, and ultimately reducing sources of variability in current proteomics workflows. Longer range goals for the collaborative program are the generation of standard reference materials including samples, antibodies, data, and protocols to be made available to the community for little or no cost.

The teams are in the process of conducting rigorous assessment and optimization of two major technologies—MS and affinity capture platforms—currently used to analyze proteins and peptides during unbiased discovery (FIGURE 8).

By optimizing the technologies, developing standard protocols to ensure data reproducibility, and providing the necessary resources to the community, this program will enable all investigators conducting protein research all over the world to use proteomic technologies and methodologies effectively to directly compare and analyze their work. This should lead, in turn, to a greater number of clinically useful biomarkers for cancer.

INTER-LABORATORY STUDIES

The CPTAC network initially focused their efforts on establishing a baseline for understanding and reducing intra-laboratory and inter-laboratory variability in the discovery stage of the proteomics pipeline. The inter-laboratory studies were designed to identify and then address the sources of variability that surfaced when individual labs analyzed identical protein mixtures. All of these experiments were undertaken with an eye towards their relevance to establishing robust clinical applications.

Unbiased Discovery

The first CPTAC experiments were designed to benchmark the analytical variability of mass spectrometry platforms commonly used to discover candidate biomarkers. Since these platforms have the potential to identify any protein as a candidate, they fall under the category of "unbiased discovery" platforms.

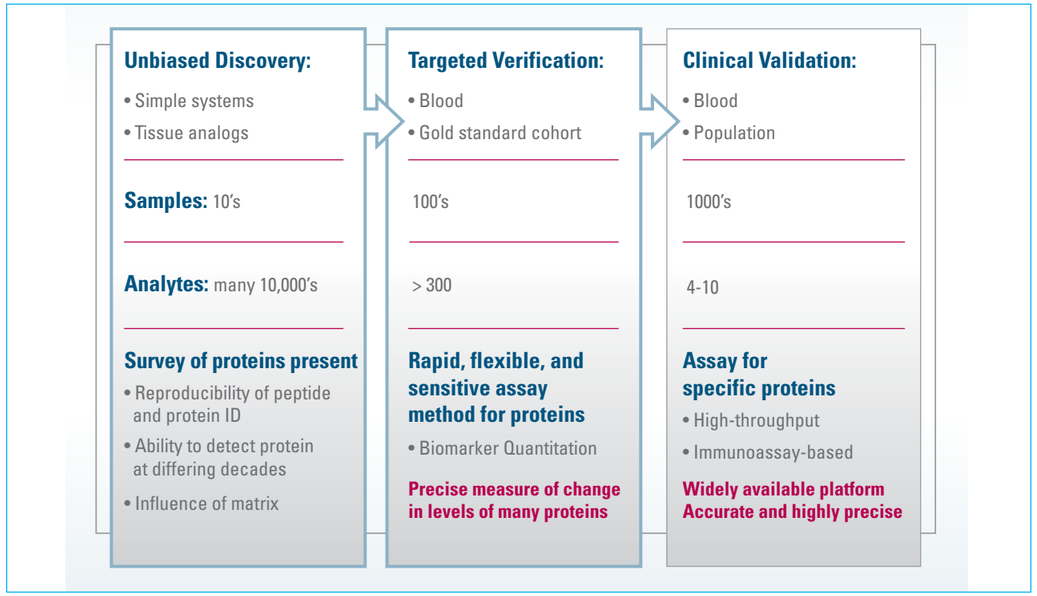
The CPTAC researchers began their analysis with a very simple mixture of proteins, drafted a standard operating procedure (SOP) based on the results, and then progressed to increasingly more complex scenarios. With each step in protein mixture complexity, the SOP was further refined to address variability issues identified in the previous step.

"When we use clinical samples, working with SOPs becomes very important."

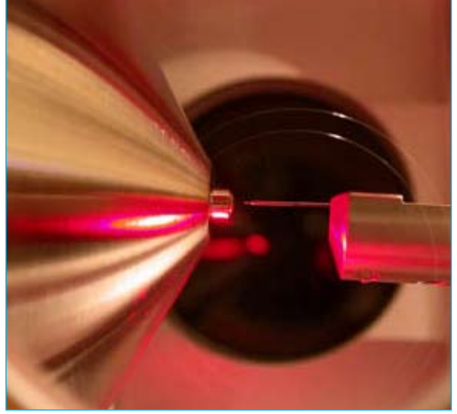
– Joshua LaBaer, M.D., Ph.D.
Director, Harvard Institute of Proteomics
Harvard Medical School

Processing of raw data, searching of candidate libraries, and comparison of analysis data submitted by separate labs was performed at the National Institute of Standards and Technology (NIST). In order to maintain consistency

FIGURE 8. Assessing Performance of Key Process Steps in the Candidate Biomarker Pipeline



The CPTAC teams are focused on reducing variability within the unbiased discovery stage of the pipeline. Targeted verification is a relatively new paradigm, which CPTAC supports. Verification is a way to rapidly triage a lengthy list of biomarker candidates prior to investing very large sums of money—and time—in clinical trials for validation.
Adapted from Rifai et al. Nat Biotechnol. 2006 Aug; 24(8): 971-83.



in the data analysis, NIST developed a common data analysis pipeline.

Measuring Baseline Variability of Current Practice

Initial CPTAC studies also focused on identifying baseline variability between laboratories. The teams analyzed a simple protein mixture developed and distributed by the NIST. This sample, the NCI-20, contained 20 human cancer-related proteins in salt buffer spanning a range of concentrations meant to resemble those found in human plasma.

An SOP was not used in this study. Each team analyzed the sample using its platform of choice in order to identify how many different platforms—and results—would be observed across the teams.

Among the CPTAC teams, eight instrument types (out of 21 instruments), as well as various liquid chromatography and data analysis tools were used to analyze the protein mixture. Not surprisingly, there was a very high degree of variability between laboratories and between instruments.

In order to minimize some of the controllable variables identified, a separate study was conducted that limited instrument type to ion trap mass spectrometers and included a limited SOP. The ion trap platform was chosen because it is the most widely used for shotgun proteomics. Although variability was reduced in this study, it was still significant, which led to a recommendation for further SOP tightening. Taking the lessons learned from the simple protein mix, the CPTAC team moved towards a more complicated protein mixture. Because of the ability to obtain sufficient material and its well characterized proteome, yeast has been chosen for analysis.

"The ultimate goal of this project is to discover cancer earlier. It's that simple."

– Leland Hartwell, Ph.D.
President and Director
Fred Hutchinson Cancer Research Center
Recipient of the 2001 Nobel Prize in Physiology or Medicine

Yeast as a Model Proteome

Although individual proteomics laboratories often have a standard mix of proteins/peptides that they use as a quality control check on their instruments, there are no globally applied complex proteome standards to benchmark analytical performance across instruments and across sites. Providing a well-characterized reference proteome as a resource to the community could help standardize proteomic technologies. One of the goals of the CPTAC network is to develop such a standard that will provide a means of comparing the performance of MS platforms:

- Over time (as a quality control);
- After the addition of new technologies, to evaluate their effectiveness compared to historic data; and
- Between laboratories, to inform optimization and troubleshooting.

Overcoming Variability through Team-based Science

Inter-laboratory Studies

The initial inter-laboratory studies were designed to answer a very basic question:

If five different laboratories analyze the exact same sample using mass spectrom etry, by how much will their results differ?

Biological samples, particularly at the protein level, are immensely complex and variable. However, an even greater level of variability is introduced by the differ ent proteomic tools and methods used to analyze those samples, and it is not easy to separate the true biological complex ity from the variables introduced by the laboratory work itself. The goal of these initial studies is to remove the layer of experimental variability so that the biology can be discovered.

To kick off these studies, each CPTAC labo ratory was asked to analyze a very basic sample, containing only 20 proteins, using their own proteomic tools and methods. When the results were compared between laboratories, the teams were amazed by how much their analyses differed.

These results underline the fundamental problem facing the proteomic community as a whole: If only five laboratories cannot reach the same conclusion using very basic protein samples, how accurate are proteomic data when patient samples are involved, which are extremely complex, with perhaps as many as 100,000 differ ent proteins? Is it any wonder that of the thousands of cancer biomarker candidates identified, only a few have made their way into clinical practice?

Significant effort is now being made by these CPTAC teams to determine why there is so much variability between laborato ries and even within laboratories in an effort to address the sources of variation.

If this initiative can successfully establish uniformity throughout the proteomics field by establishing standard practices and providing the necessary tools, then it will one day be possible for proteomics technologies to successfully make their way into clinical diagnostics and realize the vast potential they carry for detecting, treating, and even preventing cancer.

Yeast was selected as a model proteome because it has already been extensively characterized.

The yeast proteome was analyzed using a designated SOP, which was further optimized based on the previous study. Despite efforts to control for some of the variables, significant differences, although reduced, were still observed between laboratories in the number of peptides identified.

Although their work has only just begun, the CPTAC teams have made significant progress in establishing a baseline for understanding and reducing intra-laboratory and inter-laboratory variability in the unbiased discovery stage of the proteomics pipeline. In addition, valuable resources for the proteomics community have been identified: the NCI-20 protein sample and the yeast proteome sample. Re-analysis of the NCI-20 over time creates a historic record, allowing this sample to be used in subsequent studies as a performance mixture, and the yeast proteome provides an excellent performance standard and could prove to be a valuable resource for the community for benchmarking proteomic platform performance.

Manuscripts for publishing results from the above-mentioned studies are in development. Follow-up studies are currently being designed that include spiking the yeast proteome with a 48-human protein mix, using a finalized SOP. Ultimately, the CPTAC effort will move toward cell models that are relevant to human cancer.

Verification

The diagnostic biomarker pipeline begins with the discovery process, which should yield a list of candidate molecules. Before candidates can be considered true biomarkers, in a clinical sense, their presence or absence in bodily fluids and/or tissue must be quantitatively measured, or verified, in large, statistically relevant sample sets. Hence, verification analysis is mandatory in order to determine whether each candidate fulfills key requirements for use, alone or together, as a diagnostic indicator.

Multiple Reaction Monitoring (MRM)

An MS-based experimental protocol was designed and implemented to measure absolute amounts of seven proteins spiked into human plasma employing the technique of multiple reaction monitoring (MRM). MRM is currently the gold standard for identifying and quantifying drugs and metabolites in clinically relevant plasma samples due to the extremely high sensitivity and specificity of this approach.

The lower limit of detection (LOD) and protein concentrations were estimated by each team. The peptides selected for analysis varied in terms of sensitivity and percentage recovery, suggesting that two peptides from the same protein may have substantially different detection characteristics. Results across CPTAC sites were remarkably consistent in terms of percent recovery, intra-laboratory variability, and LOD.

Ongoing efforts within this initial study include, but are not limited to, extending the MRM assay dynamic range at the low end of the concentration scale in order to detect lower abundance proteins.

Optimizing Current Technologies

Intra-laboratory Studies

In addition to the enormous undertaking of establishing uniformity throughout the proteomics field, each CPTAC center is also trying to optimize current technologies in an effort to make them faster, yet more accurate and reliable for clinical use.

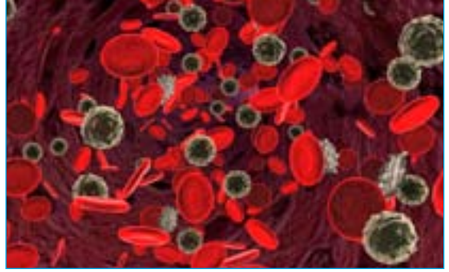
The Broad Institute of MIT and Harvard – Developing techniques to accurately and reproducibly measure the amount of protein(s) present in human plasma

Memorial Sloan-Kettering Cancer Center – Applying the use of automation in proteomics in order to eliminate human variability and enable high-throughput

Purdue University Developing simple, inexpensive technologies that will ac curately measure biomarkers in up to 10,000 patient samples per week with minimum sample workup

University of California, San Francisco Applying screening techniques to proteins that have become specifically modified during the initiation and/or progression of cancer

Vanderbilt University – Developing tech niques for detecting biomarkers in tissue, avoiding the technical challenge of detecting proteins in plasma at very low concentrations



Biospecimen Collection Protocol

The teams spent significant time enabling collaboration among CPTAC sites on studies requiring the sharing of clinical samples. This is accomplished by aligning, where possible, clinical populations, biospecimen collection, processing and storage protocols, and clinical data elements in a HIPAA-compliant manner.

"...we, too, have come to realize that you can waste an awful lot of high powered effort on unreliable samples."

– Gilbert Omenn, M.D., Ph.D.
Professor of Internal Medicine, Human Genetics, and Public Health
University of Michigan

Alignment includes identifying and reducing sources of variation and bias where possible.

Working with the NCI's Office of Biorepositories and Biospecimen Research (OBBR), the CPTAC teams are developing a set of guidelines and protocols to be used by the proteomic research community and its clinical collaborators in order to minimize the variability introduced prior to proteomic analysis.

To date, clinical protocol alignment has been accomplished at four of the five CPTAC sites for breast cancer plasma biospecimens. These biospecimens are now collected from patients with a breast lesion who are about to undergo a breast biopsy. The fifth CPTAC site is collecting an important orthogonal set of breast cancer-related plasma biospecimens in patients who have completed chemotherapy for breast cancer. A common protocol for blood collection and plasma processing has also been defined.

NCI-Frederick has been chosen as the site for a centralized CPTAC biorepository. A logistics protocol for shipping and tracking of specimens to NCI-Frederick and a distribution protocol for shipment of specimens from NCI-Frederick to CPTAC sites for group-wide CPTAC experiments are currently in development.

CPTAC CENTERS AND INTRA-LABORATORY STUDIES

In addition to the collaboration with other CPTAC members to further the optimal application of proteomics tools to cancer research, each group is applying the findings from the CPTAC collaboration to their own specific research interests.

The Broad Institute of MIT and Harvard

Team Leader: Steven A. Carr, Ph.D.
Measuring Cancer Biomarker Candidates by Targeted MS and Antibody Enrichment

The primary focus of this CPTAC research team is on the development of MRM assays for the quantitation of candidate-based protein markers in plasma. This team proposes to make MRM robust and reproducible for clinical proteomics. The use of a workflow involving strong cation exchange chromatographic fractionation of peptides and immunoaffinity enrichment on specific anti-peptide antibodies (stable isotope standards and capture by anti-peptide antibodies, SISCAPA) is being assessed. The goal is to use SISCAPA combined with MRM to obtain reliable and reproducible quantitation of signature peptides from proteins in complex digests.

Collaborators:

- The Broad Institute of MIT and Harvard, Proteomic Platform and Cancer Program
- Fred Hutchinson Cancer Research Center and its clinical and research partners, the University of Washington and Children's Hospital and Regional Medical Center
- Plasma Proteome Institute
- University of Victoria (UVic-Genome BC Proteomics Center at Vancouver Island Technology Park and Development of Biochemistry and Microbiology)
- Harvard University and its affiliated hospitals (including Dana-Farber Cancer Institute and Massachusetts General Hospital)
- Epitome, Inc.

Memorial Sloan-Kettering Cancer Center

Team Leader: Paul Tempst, Ph.D.
Assessment of Serum Peptide Profiling to Detect Cancer-Specific Patterns

The scientists comprising this CPTAC team bring significant expertise in automated sample processing technology (robotics) to the entire CPTAC effort. This method has the potential to significantly eliminate handler variability and induced error associated with peptide measurements from clinical samples. In addition, this team has expertise in the coupling of sample fractionation using magnetic beads for the capture of peptides, prior to matrix-assisted laser-desorption/ionization time-of flight (MALDI-TOF) MS analysis. Because beads provide a larger

surface-area-to-volume ratio than flat plate protein chip designs, this enrichment process could provide a breakthrough in capturing more peptides of relevance in cancer biology. Furthermore, since peptides are metabolic products derived from proteins through the action of peptidases, an activity test has been developed for "functional" biomarker discovery. This test uses the same platform of robotics, magnetic particles, and MALDI-TOF, and the reproducibility and effectiveness will be assessed. The relevance of this technology platform could be significant for discovery researchers by enabling high-throughput and reproducibility.

Collaborators:

- Memorial Sloan-Kettering Cancer Center
- New York University Medical Center

Purdue University

Team Leader: Fred E. Regnier, Ph.D.

Translation and Clinical Proteomic Technology Assessment for Cancer: The Indiana Program

The goal of this CPTAC team is to develop simple, inexpensive analytical platform(s) that allow quantification of 10 to 50 biomarkers in 1,000 to 10,000 samples per week with minimum sample workup. The efficacy of three platforms is being compared. One is a bottom-up platform approach in which abundant protein is removed from plasma samples and then tryptic digested before fractionation with either liquid chromatography or ion mobility separators, followed by label-free quantification with ion trap or TOF mass spectrometers. A second approach exploits direct affinity selection of glycoprotein markers from plasma with lectins and antibodies prior to proteolysis and stable isotope-based comparative proteomics by mass spectrometry. The third analytical strategy utilizes large-scale immunological arrays that simultaneously select multiple analytes from several hundred plasma samples and probe protein and glycan structure in a sandwich assay format. Critical components this group brings to CPTAC are expertise in microfabrication, miniaturization, instrument development, immunological microarrays, sandwich assays that target post-translational modifications, and interferometric-based immunological assays suitable for label-free quantification in discovery and clinical proteomics.

Collaborators:

- PTM Analysis of Proteins, CPTAC Groups at University of California, San Francisco, and Buck Institute for Age Research
- NCI Cancer Centers at Purdue University and Indiana University School of Medicine
- Hoosier Oncology Group
- Indiana University
- Discovery Park at Purdue University
- Predictive Physiology and Medicine, Inc.
- Quadraspec, Inc.

University of California, San Francisco

Team Leader: Susan J. Fisher, Ph.D.

Toward Development of MS-based Screening Protocols for Early Detection of Cancer Predicated on Alterations in Alternative Splicing or Posttranslational Modifications

Unique features of this CPTAC research project include the development of novel workflows for plasma separation driven by data regarding alternative splicing and posttranslational modification (PTM) that are obtained by phenotyping the relevant tumor type. Unique isotope labeling strategies for comparing the utility of various plasma/serum isolation methods will also be employed. Other contributions include sophisticated technology platforms, a well-developed informatics infrastructure, and a proven track record of generating, implementing, and sharing novel algorithms and databases compatible with caBIG™.

Collaborators:

- Lawrence Berkeley National Laboratory
- Buck Institute for Age Research
- California Pacific Medical Center
- University of Texas M.D. Anderson Cancer Center
- University of British Columbia

Vanderbilt University School of Medicine

Team Leader: Daniel C. Liebler, Ph.D.

Standardized Proteomics Platforms for Biomarker Discovery and Verification in Cancer Research

The Vanderbilt CPTAC program is directed at the application of MS-based proteomics technologies for the discovery and verification of biomarkers to detect cancer and facilitate therapy. The team is focusing on tissue-based biomarkers for several key reasons: Many important cancer-related applications require biomarkers that can be analyzed in tissue samples; these applications meet needs related to diagnosis, prognosis, and therapeutic decision-making. Tissue-based biomarker applications are the shortest path to proteomics "success" in cancer because they avoid the technical challenge of detection in plasma at sub-ng/mL concentrations.

Collaborators:

- Vanderbilt-Ingram Cancer Center
- University of Texas M.D. Anderson Cancer Center

Providing Resources to the Entire Community

The Proteomic Reagents and Resources Core

Discussions with representatives from all parts of the cancer research community revealed a deep concern about the lack of access to affordable, well characterized antibodies and supporting resources. In order to drive the development of a central community core that would address this need, NCI launched the Proteomic Reagents and Resources Core.

This program within CPTC is a collaborative effort designed to serve the entire international proteomics community. Government agencies, academic institutions, and the private sector are working together to provide the necessary resources that are sorely needed to accelerate biomarker discovery and validation, transla tional research, molecular diagnostics, and therapeutic monitoring. Reagents and resourc es will include plasma, antibodies, databases, standard protein and/or peptide mixtures, and other standard reagents needed for effective proteomic analysis platforms.

Developing Next Generation Technologies for the Proteomic Community

Advanced Platforms and Computational Sciences

This component of the CPTC initiative was designed to aid individual investigators in the development of next generation quantitative proteomic technologies. Fifteen individual awards were provided to academic institutions across the nation.

Seven institutions and their collaborators are developing innovative technologies that are rapid, specific, reliable, and inexpensive important criteria for routine clinical use.

These include:

- University of Houston
- Northeastern University
- University of California, Los Angeles
- Institute for Systems Biology
- Emory University
- Battelle Pacific Northwest Laboratories
- Michigan State University

In parallel, eight institutions are developing extremely powerful computational tools that are necessary for accurate analysis of very large proteomic data sets.

These include:

- University of Maryland, College Park
- College of William and Mary
- Massachusetts Institute of Technology
- University of Michigan
- Fred Hutchinson Cancer Research Center
- University of Colorado at Boulder
- Vanderbilt University
- University of Virginia

"...the work of the cell is ultimately done by proteins. And so if we want to monitor what's changed in drug resistance or what's changed during successful therapy, it's most direct to look at the proteins."

– Catherine Fenselau, Ph.D.
Professor of Chemistry and Biochemistry
University of Maryland

CPTC Components:

Advanced Platforms and Computational Sciences

The Advanced Proteomic Platforms and Computational Sciences component is a comprehensive program focused on the development of innovative new tools, reagents, and the enabling of technologies for protein/peptide measurement, such as algorithm development and computational methods to interrogate emerging pre-processed data sets.

The Advanced Proteomic Platforms and Computational Sciences component supports two focus areas for protein measurement technology and application in cancer research:

- The development of innovative high-throughput technology to detect measure and characterize proteins and peptides in biological fluids that will overcome current barriers.
- The development of computational, statistical, and mathematical approaches for the analysis, processing, and facile exchange of large proteomic data sets.

Advancing the technological and analytical capabilities in proteomic research will allow the research community to better characterize and understand the differences between the normal and diseased human proteome and to develop diagnostic and treatment procedures based on these distinctions. There were 15 individual awards made, each of which is described below along with their current status.

ADVANCED PLATFORMS AWARDS

Proteomic Phosphopeptide Chip Technology for Protein Profiling

Principle Investigator:
Xiaolin Gao, Ph.D.
University of Houston

The goal of the project is to create a clinically useful microchip technology that will allow researchers to profile cancer-related proteins that have become specifically modified as a result of disease. These microchips will have wide application in clinical cancer research.

Global Production of Disease-specific Monoclonal Antibodies

Principle Investigator:
Barry L. Karger, Ph.D.
Northeastern University

The goal of the project is to generate cancer-specific reagents that will have significantly higher sensitivity and throughput than traditional approaches for detecting biomarkers that are present at very low levels. The ultimate goal at later stages is to have a library of more than 1,000 well-characterized monoclonal antibodies to low-level proteins in blood, available for screening and discovery of biomarkers. This would represent a major step in development of a true high-throughput approach suitable for large scale population studies.

Top-Down Mass Spectrometry of Salivary Fluids for Cancer Assessment

Principal Investigator:
Joseph A. Loo, Ph.D.
University of California, Los Angeles

This research program focuses on the development of a new technology platform to identify relevant cancer markers in saliva. The use of saliva for disease diagnostic purposes presents an attractive potential option because of its ease of collection and its relative ease for protein profiling compared to plasma.

A New Platform to Screen Serum for Cancer Membrane Proteins

Principal Investigator:
Daniel B. Martin, M.D.
Institute for Systems Biology

In an effort to obtain better diagnostic markers of prostate cancer, a proteomic platform will be developed and implemented for the capture and analysis of prostate-specific proteins in cell culture models of the disease. The goal of this work is to define a rapid, specific, reliable, and inexpensive strategy to identify and validate prostate cancer protein markers.

A Proteomics Approach to Ubiquitination

Principal Investigator:
Junmin Peng, Ph.D.
Emory University

Proteins are labeled for degradation via a modification process known as ubiquitination. The goal of this project is to analyze the ubiquitination pattern of the proteome in mammalian tissues and human brain tumors in an effort to develop a new and powerful preparative technology for analysis of such patterns in clinical tissues.

A Proteomics Platform for Quantitative, Ultra-High-Throughput, and Ultra-Sensitive Biomarker Discovery

Principal Investigator:
Richard D. Smith, Ph.D.
Battelle Pacific Northwest Laboratories

The objective of this project is to develop a platform that will enable higher throughput and more sensitive and quantitative proteomics measurements for candidate biomarker discovery. The development of the platform will be evaluated continually in the context of low-level protein measurements in clinical biomarker discovery efforts.

Aptamer-Based Proteomic Analysis for Cancer Signatures

Principal Investigator:
Stephen P. Walton, Ph.D.
Michigan State University

The basic premise of the proposed research is the translation of protein information (difficult to analyze in parallel) to nucleic acid information (easy to analyze in parallel) via specific, high-affinity nucleic acid labels called aptamers. Aptamers provide protein-binding specificity, and the labels provide unique identification of a single aptamer in a pool. Parallel measurement of proteins becomes simply parallel measurement of aptamers (nucleic acids). The aptamers generated will allow for the design of a diagnostic technique for maximal sensitivity and accuracy.

"We need to keep our focus on technology development certainly as it's applied to health and disease, and I think this will become our door to really great discoveries in the future."

– Richard Caprioli, Ph.D.
Stanley Cohen Professor of Biochemistry
Director of Mass Spectrometry Research Center
Vanderbilt University School of Medicine

COMPUTATIONAL SCIENCE AWARDS

Proteomic Characterization of Alternate Splicing and cSNP Protein Isoforms

Principal Investigator:
Nathan J. Edwards, Ph.D.
University of Maryland, College Park

The project seeks to develop computational tools that make it possible to observe unexpected, unusual, and potentially malfunctioning versions of proteins in clinical cancer samples and cancer cell-lines.

Enhancement of MS Signal Processing Toward Improved Cancer Biomarker Discovery

Principal Investigator:
Dariya Malyarenko, Ph.D.
College of William and Mary

The goal of the project is to develop computational tools aimed at increasing the effectiveness of cancer biomarker discovery from MALDI-TOF mass spectra. The advanced signal processing algorithms and computational tools should help achieve more than an order of magnitude increase in both sensitivity and selectivity for molecular biomarker screening in a broad mass range. These improvements are toward advancing comparative mass spectrometry technology for the detection of molecular signatures of cancer in tissue and body fluids in clinical research.

A Platform for Pattern-based Proteomic Biomarker Discovery

Principal Investigator:
Denkanikota Mani, Ph.D.
Massachusetts Institute of Technology

This project is focused on developing a robust platform for analysis of liquid chromatography-mass spectrometry (LC-MS) data to enable higher throughput detection of differential changes in protein abundance across multiple clinical samples.

Analysis and Statistical Validation of Proteomic Datasets

Principal Investigator:
Alexey I. Nesvizhskii, Ph.D.
University of Michigan

One of the key problems in the field of proteomics is statistical validation, interpretation, and mining of collected datasets. The overall goal of this project is to develop statistical data analysis methods and algorithms that will enable robust, accurate, and transparent analysis of large-scale quantitative MS/MS-based proteomic datasets from human clinical specimens.

Quantitative Methods for Spectral and Image Data in Proteomics Research

Principal Investigator:
Timothy W. Randolph, Ph.D.
Fred Hutchinson Cancer Research Center

Proteomics research seeks protein-based clues about disease contained in fluids and tissue using a wide range of instruments that produce complex, high-dimensional data. This project aims to provide effective and accessible methods for data processing, interpretation, and statistical analysis of these data as they arise in cancer research studies.

Computational Tools for Cancer Proteomics

Principal Investigator:
Katheryn A. Resing, Ph.D.
University of Colorado at Boulder

The goal of this project is to develop both methods and informatics tools to understand the complexity of protein signaling, with a focus on cancer-specific signaling pathways.

New Proteomic Algorithms to Identify Mutant or Modified Proteins

Principal Investigator:
David L. Tabb, Ph.D.
Vanderbilt University

Because of genetic diversity, clinical samples may contain mutant proteins that do not match the sequences found in protein databases produced during the Human Genome Project. Ordinary protein identification techniques fail to identify these mutated sequences. This project is developing new algorithms to identify proteins even when these mutational differences are present.

PICquant-An Integrated Platform for Biomarker Discovery

Principal Investigator:
Dennis J. Templeton, Ph.D.
University of Virginia

One promising proteomic application is the potential for a complete analytic platform for urine biomarker discovery. The goals of this project are to (1) develop a new labeling reagent for peptides specifically found in urine, and (2) develop a clinical registry that links acquired urine specimens to current and prospective clinical information, including outcomes. The registry will enable multivariate clustering of disease states with quantified protein families.

NIST NATIONAL CANCER INSTITUTE

National Institute of Standards and Technology

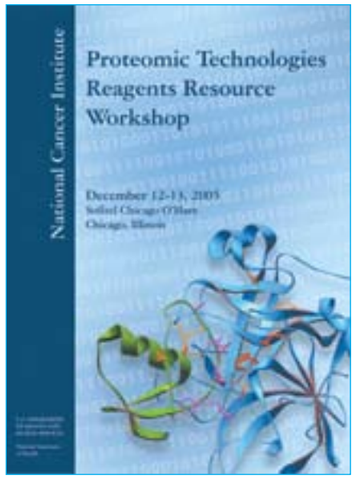
NCI has entered into an interagency agreement with NIST to develop mass spectrometry assessment materials. These materials, designed to assess the performance metrics of various instruments, will be the first of their kind developed by the NCI and will help to evaluate and compare existing proteomic technologies and compare these with emerging proteomic technologies of interest to the clinical cancer community.

CPTC Components:
The Proteomic Reagents and Resources Core

One of the most significant bottlenecks of advanced molecular diagnostic techniques like proteomics is a lack of high-quality and well-characterized reagents. This barrier was recognized by the NCI, which led to the development of the Reagents and Resources component of its Clinical Proteomic Technologies for Cancer initiative.

In order to fully realize the promise of proteomics in cancer, high-quality proteins and validated affinity capture reagents (e.g., antibodies) comprising the human proteome need to be accessible to the entire scientific community.

While tens of thousands of reagents are commercially available in today's antibody market, few are well-characterized and many are highly variable in terms of quality. Furthermore, few of these reagents are directed against novel (investigator-driven) targets, due to intellectual property issues.



The NCI held a workshop in December 2005 to outline a strategic plan to address this bottleneck and invited attendees from public, private, academic, and international institutions. Outcomes from the workshop were published in the journal *Molecular & Cellular Proteomics*. Key considerations outlined by the scientific community included:

- Making antigens freely available;
- Supporting renewable antibody production (i.e., monoclonal antibodies);
- Enabling antibody characterization using SOP-driven protocols; and
- Supporting distribution of antibodies and hybridomas for research use with no intellectual property in order to promote multiplex affinity capture platform development

It is upon these recommendations that NCI's CPTC Reagents and Resources Core is based.

This program acts as a catalyst to spur the development of pivotal resources such as antibodies that serve the entire research community. These resources are necessary for the acceleration of biomarker discovery and validation, translational research, molecular diagnostics, and therapeutic monitoring.

Reagents will be available through a central portal that will be accessible through the CPTC Web site at *http://proteomics.cancer.gov*.

OPEN-ACCESS MONOCLONAL ANTIBODIES

Among the first requests to be addressed, in addition to the establishment of protocols, was the development of a series of validated antibodies against proteins known to be related to cancer. In this past year a pipeline has been established to develop a renewable source of "open-access" (that is, available to all with minimal IP restrictions) monoclonal antibodies for use by the entire cancer proteomic community (FIGURE 9). It is expected that this resource will be up and running in 2008, with hybridomas expressing antibodies against additional proteins being added continuously.

FIGURE 9. The Monoclonal Antibody Pipeline

Supporting Antigen Production:

Argonne National Laboratory is expressing and purifying proteins selected from a starting list of target antigens that have been identified by the scientific community.

Supporting Monoclonal Antibody Production:

A Request for Proposals (RFP) was posted for private companies to develop hybridomas against the first set of cancer-related antigens, with the goal of three monoclonal antibodies against each target. Successful RFP applicants will perform initial quality control of the antibodies. The RFP is being awarded in 2008.

Three monoclonal antibodies per antigen is the goal because different antibodies have different performance characteristics depending on the assay that is being used (e.g., ELISA, Western blotting). This will also allow researchers to target different regions of the protein.

Antibody Characterization

NCI-Frederick – *combination of ELISAs, Western blots, surface plasmon resonance, immunohistochemistry, immunoprecipitation, immunofluorescence, and immuno-mass spectrometry*

Tissue Array Research Program at NCI Center for Cancer Research

Harvard Institute of Proteomics – *Nucleic Acid Programmable Protein Arrays (NAPPA)*

Human Protein Atlas *(KTH – Royal Institute of Technology; Stockholm, Sweden) – immunohistochemistry*

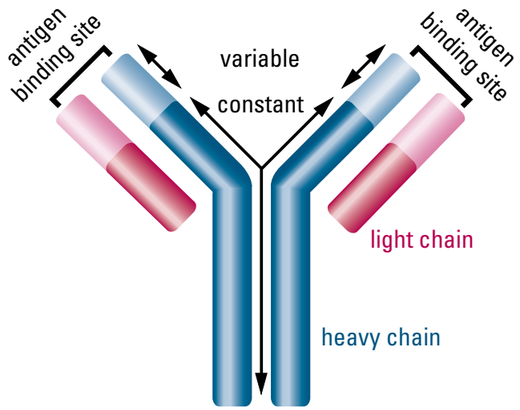
Supporting Hybridoma and Antibody Distribution:

Collection, storage, and distribution of fully-characterized antibodies and hybridomas will be carried out through the Developmental Studies Hybridoma Bank at the University of Iowa. Researchers will be charged a small fee.

"...NCI's efforts to create reagent resources will be great for science..."

– *John Yates, Ph.D.*
Professor of Cell Biology, Head of the Proteomics Mass Spectometry Lab
The Scripps Research Institute

"This program will spur the development of resources that accelerate biomarker discovery and validation, translational research, molecular diagnostics, and therapeutic monitoring," says Henry Rodriguez, Ph.D., M.B.A. "I'm convinced that when this happens, we will have built a solid foundation for proteomics in cancer."



Public - Private Partnerships

The CPTC is designed to enable infrastructure development to support clinical proteomics experiments. As such, the goals of the initiative seek to provide a broadly applicable set of tools amenable to wide use by the research community. To facilitate this mission, the NCI has formed strategic collaborations with government agencies, academic, and commercial entities.

Federal and Academic Agencies

ARGONNE NATIONAL LABORATORY (ANL)

To learn more about ANL,
visit: *http://www.anl.gov*.

DEVELOPMENTAL STUDIES HYBRIDOMA BANK (DSHB) AT THE UNIVERSITY OF IOWA

To learn more about DSHB,
visit: *http://dshb.biology.uiowa.edu*.

HARVARD INSTITUTE OF PROTEOMICS (HIP)

To learn more about HIP,
visit: *http://www.hip.harvard.edu*.

HUMAN PROTEIN ATLAS (HPA)

To learn more about HPA,
visit: *http://www.proteinatlas.org*.

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST)

To learn more about NIST,
visit: *http://www.nist.gov*.

TEXAS A&M UNIVERSITY (TAMU)

To learn more about TAMU,
visit: *http://www.tamu.edu*.

Small Business Innovation Research (SBIR) Contracts

2007 CONTRACTS

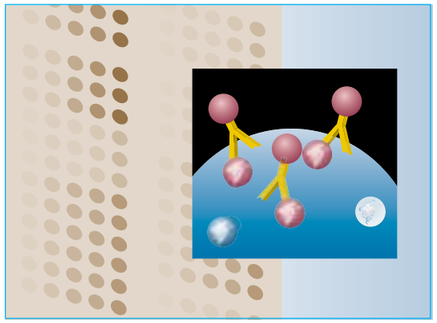
In 2006, a solicitation for Fiscal Year 2007 SBIR Grant Proposals was released to the community. Six small businesses were awarded SBIR contracts under two topics in support of the CPTC initiative.

Topic 238 – Development of Clinical Automated Multiplex Affinity Capture Technology for Detecting Low Abundance Cancer-related Proteins/Peptides

Of the hundreds of thousands of proteins believed to be found in different body fluids, it is likely that cancer-related proteins will be in relatively low abundance. The development of effective technologies to accurately measure these proteins and improve our diagnostic capabilities by discerning diseased from non-diseased states requires the development of next-generation proteomic technologies. The purpose of this project is to stimulate the development of quantitative automated affinity/protein capture multiplex technologies for measuring low abundance cancer-related proteins/peptides from bodily fluids in support of the CPTC initiative. In addition, this tool, as conceived, is to be applicable in Cancer Centers and other settings where NCI Investigators conduct clinical care.

Topic 239 – Development of Alternative Affinity Capture Reagents for Cancer Proteomics Research

Today, existing biotechnology reagent companies produce thousands of antibodies per year. Many of these are commercially available. However, many of these antibodies are known to be poorly characterized and suboptimal across multiple applications. Polyclonal antibodies lack the reproducibility of monoclonal antibodies. Likewise, the production of monoclonals is expensive and may take six to eight months to produce. Even after production, there is no guarantee that a monoclonal antibody will be specific for the target of interest, will work in the needed assay, or can be used in combination with other antibodies due to an antibody's large size and subsequent competition for overlapping binding domains. As such, the high costs associated with producing even small quantities of monoclonal antibodies represent a large barrier towards cost-effective reagents and resources for proteomic technology research and clinical adaptation. The goal of this project is to develop reproducible, highly-qualified/characterized-alternative protein capture reagents for the cancer research community. The development of these affinity capture reagents will be done in coordination with NCI's Clinical Proteomic Technologies for Cancer initiative and be targeted to a list of over 100 purified recombinant proteins being constructed and characterized through this initiative.



2008 SOLICITATION FOR CONTRACT PROPOSALS

In 2007, a solicitation for Fiscal Year 2008 SBIR Grant Proposals was released to the community. There were two topics in support of the CPTC initiative. Awards will be announced in 2008.

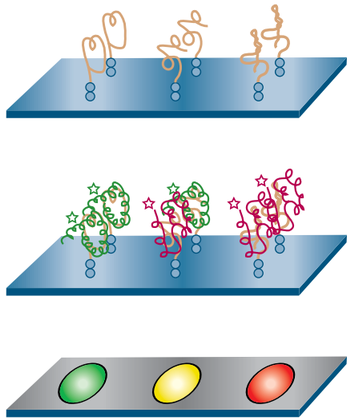
Topic 253 – Advances in Protein Expression of Post-Translationally Modified Cancer-Related Proteins

Release of a completed polypeptide chain from a ribosome is often not the last chemical step in the formation of a protein. Various covalent modifications often occur, either during or after assembly of the polypeptide chain. Most proteins undergo co- and/or post-translational modifications. Knowledge of these modifications is extremely important because they may alter physical and chemical properties, folding, conformation distribution, stability, activity, and ultimately the function of the protein. Moreover, the modification itself can act as an added functional group.

2007 SBIR Contract Recipients (Topic 238)

Meso Scale Diagnostics	Automated Multi-Array Platform for Cancer Biomarkers
Sequenom Inc.	Sensitive Protein Detection Combining Mass Spectrometry
Quadraspec Inc.	Highest Sensitivity Cancer Marker Array on Quadraspec's Bio-CD Platform
Rules Based Medicine Inc.	Automated Multiplexed Immunoassays for Rapid Quantification of Low Abundance Cancer-Related Proteins

Aptamer microarray



Examples of the biological effects of protein modifications include phosphorylation for signal transduction, ubiquitination for proteolysis, attachment of fatty acids for membrane anchoring and association, glycosylation for protein half-life, targeting, cell:cell and cell:matrix interactions. Consequently, the analysis of proteins and their post-translational modifications (PTMs) is particularly important for the study of heart disease, cancer, neurodegenerative diseases, and diabetes. Therefore, the NCI is interested in proposals that focus on the development of post-translationally modified human proteins (e.g., glycosylation, phosphorylation, acetylation, oxidation). The purpose of this project is to stimulate the development on all aspects of PTM protein expression including chemical synthesis, novel cell systems, expression vectors, and culture conditions. Proteins selected for production are to entail low abundance cancer-related proteins from bodily fluids in support of the Clinical Proteomic Technologies initiative. These proteins are to become part of CPTC's Reagents and Resource Core.

Number of anticipated awards: 4

Topic 254 – Development of Clinical Quantitative Multiplex High-Throughput Mass Spectrometric Immunoassay for Detecting Low Abundance Cancer-Related Proteins/ Peptides in Bodily Fluids

The application of proteomics tools in the clinical setting lags far behind their use in basic science and drug discovery. In the past, protein/peptide biomarkers were tested individually to determine their value using common techniques such as ELISA, 2-D gels, and mass spectrometry. Each of these technologies has its advantages, but they still suffer from an inability to quantitatively evaluate multiple markers in a single reaction. However, recent applications of affinity mass spectrometry into clinical laboratories brought a renewed interest in mass spectrometric immunoassays as a more specific affinity method capable of selectively targeting and studying protein biomarkers. In mass spectrometry-based immunoassays, proteins are affinity retrieved from biological samples via surface-immobilized antibodies, and are then detected via mass spectrometric analysis. The assays benefit from dual specificity, which is brought about by the affinity of the antibody and the protein mass readout. The mass spectrometric aspect of the assays enables single-step detection of protein isoforms and their individual quantification.

Therefore, the NCI is interested in proposals that focus on developing a multiplexed mass spectrometric immunoassay for the detection of low abundance cancer-related proteins/ peptides from bodily fluids (examples of "bodily fluids" include plasma or serum, serous fluids collected from ductal lavage, but not cell lysates or tissue culture media). Surface enhanced laser desorption ionization (SELDI) MS will not be considered for this SBIR due to its limited ability to comprehensively measure and identify low abundance proteins in serum or plasma considered to be within the dynamic range of proteins released from cancer cells.

Number of anticipated awards: 4

2007 SBIR Contract Recipients (Topic 239)

Allele Biotechnology & Pharmaceuticals	Yeast Single Chain Antibodies as Capture Reagents
Accacia International Inc.	High-Throughput of Aptamers against Cancer Biomarkers

Community Outreach

CPTC places a premium on communicating with research stakeholders to ensure that opportunities to incorporate unique perspectives are explored fully and optimally. This commitment to communication and outreach does not stop with investigators, clinicians, and private sector representatives who translate discoveries from the bench to the bedside.

Diagnostics and therapeutics developed using support from CPTC research programs ultimately benefit patients and those disease-free individuals who will benefit from preventive approaches. CPTC engages with these beneficiaries through an organized outreach program that includes direct involvement and input from representatives of the advocacy community, including NCI's Consumer Advocates in Research and Related Activities (CARRA). CARRA members participate in a wide range of NCI activities and represent the collective viewpoint of people affected by cancer. CARRA members participate in a variety of NCI activities involving scientific research and communication of scientific results, including but not limited to, sitting on committees and boards, and attending meetings, workshops, and site visits.

Ann McNeil, R.N., Miami Children's Hospital, is one CARRA member involved in CPTC activities. At the 1st Annual CPTC Meeting, Ms. McNeil emphasized in her welcoming speech that the work being carried out by the CPTC investigators and teams ultimately will impact people's lives and treatments in positive ways. The members of CPTC would like to thank CARRA and Ms. McNeil specifically, for their dedication, hard work, and constant reminder that everything starts with a patient and ends with a patient.

Appendix A - *Participating Organizations*

ORGANIZATIONS PARTICIPATING IN THE CPTC INITIATIVE

Accacia International, Inc.

Allele Biotechnology & Pharmaceuticals

Argonne National Laboratory

Battelle Pacific Northwest Laboratories

The Broad Institute of MIT and Harvard, Proteomic Platform and Cancer Program

Buck Institute for Age Research

California Pacific Medical Center

College of William and Mary

Developmental Studies Hybridoma Bank at the University of Iowa

Discovery Park at Purdue University

Emory University

Epitome, Inc.

Fred Hutchinson Cancer Research Center and its clinical and research partners, the University of Washington and Children's Hospital and Regional Medical Center

Harvard Institute of Proteomics

Harvard University and its affiliated hospitals (including Dana-Farber Cancer Institute and Massachusetts General Hospital)

Hoosier Oncology Group

Human Protein Atlas (KTH – Royal Institute of Technology; Stockholm, Sweden)

Indiana University

Indiana University School of Medicine

Indiana University – Purdue University Indianapolis

Institute for Systems Biology

Lawrence Berkeley National Laboratory

Massachusetts Institute of Technology

Memorial Sloan-Kettering Cancer Center

Meso Scale Diagnostics

Michigan State University

Monarch Life Sciences

National Cancer Institute – Center for Cancer Research Tissue Array Program

National Cancer Institute – Frederick Advanced Technology Program

New York University Medical Center

Northeastern University

Plasma Proteome Institute

Predictive Physiology and Medicine, Inc.

Purdue University

Quadraspec, Inc.

Rules-Based Medicine, Inc.

Sequenom, Inc.

University of British Columbia

University of California, Los Angeles

University of California, San Francisco

University of Colorado at Boulder

University of Houston

University of Maryland, College Park

University of Michigan

University of Texas M.D. Anderson Cancer Center

University of Victoria (UVic-Genome BC Proteomics Center at Vancouver Island Technology Park and Development of Biochemistry and Microbiology)

University of Virginia

Vanderbilt University Medical Center

Vanderbilt-Ingram Cancer Center

Appendix B - *Publications*

INTRA-LABORATORY STUDY PUBLICATIONS

Hu J, He X, Baggerly K, Coombes K, Hennessy B, Mills G. Non-parametric quantification of protein lysate arrays. *Bioinformatics*. 2007;23(15):1986-94.

Fenyo D, Phinney B, Beavis R. Determining the overall merit of protein identification data sets: rho-diagrams and rho-scores. *J Proteome Res*. 2007;6(5): 1997-2004.

Zhang B, Chambers M, Tabb D. Proteomic parsimony through bipartite graph analysis improves accuracy and transparency. *J Proteome Res*. 2007;6(9):3549-57.

Tabb D, Fernando C, Chambers M. MyriMatch: highly accurate tandem mass spectral peptide identification by multivariate hypergeometric analysis. *J Proteome Res*. 2007; 6(2):654-61.

ADVANCED PROTEOMIC PLATFORMS AND COMPUTATIONAL SCIENCES PUBLICATIONS

Wu X, Tseng C, Edwards N. HMMatch: peptide identification by spectral matching of tandem mass spectra using hidden Mardov models. *J Comput Biol*. 2007;14(8):1025-43.

Edwards N. Novel peptide identification from tandem mass spectra using ESTs and sequence database compression. *Mol Sys Biol*. 2007;3:102.

Peng I, Shiea J, Ogorzalek Loo R, Loo J. Electrospray-assisted laser desorption/ionization and tandem mass spectrometry of peptides and proteins. *Rapid Commun Mass Spectrom*. 2007;21(16):2541-6.

Page J, Kelly R, Tang K, Smith R. Ionization and transmission efficiency in an electrospray ionization-mass spectrometry interface. *J Am Soc Mass Spectrom*. 2007;18(9):1582-90.

Kelly R, Page J, Tang K, Smith R. Array of chemically etched fused-silica emitters for improving the sensitivity and quantitation of electrospray ionization mass spectrometry. *Anal Chem*. 2007;79(11):4192-8.

Liu J, Kang S, Tang C, Ellis L, Li T. Meta-prediction of protein subcellular localization with reduced voting. *Nucleic Acids Res*. 2007;35(15):e96.

Zhu Q, Hong A, Sheng N, Zhang X, Matejko A, Jun K-Y, Srivannavit O, Gulari E, Gao X, Zhou X. Paraflo™ Biochip for Nucleic Acid and Protein Analysis. *Methods Mol Biol*. 2007;382:287-312 (book chapter)

Nesvizhskii A, Vitek O, Aebersold R. Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nat Methods*. 2007;4(10):787-97.

Witze E, Old W, Resing K, Ahn N. Mapping protein post-translational modifications with mass spectrometry. *Nat Methods*. 2007;4(10):798-806.

Ibrahim Y, Belov M, Tolmachev A, Prior D, Smith R. Ion funnel trap interface for orthogonal time-of-flight mass spectrometry. *Anal Chem*. 2007;79(20):7845-52.

National Cancer Institute

Office of Technology & Industrial Relations

ATTN: Clinical Proteomic Technologies for Cancer
31 Center Drive, MSC 2580
Bethesda, MD 20892-2580

Email: cancer.proteomics@mail.nih.gov
Web site: http://proteomics.cancer.gov

NIH Publication No. 08-6469
Printed July 2008