

Restructuring proteomics to enable personalised cancer care

The genomics community has made great strides in our understanding of the molecular basis of cancer and these advances are slowly beginning to change the way we diagnose and treat patients. But genomic studies alone cannot capture the complete view of disease processes – a more comprehensive approach is needed. For example, genetic aberrations can provide the likelihood of developing a certain disease, but proteins can diagnose what is happening in a patient in real time. As genomics begins to catalyse the field of personalised medicine, other important disciplines, including proteomics, biospecimens and bioinformatics are coming together to elevate this important field of medicine to the next level.

The addition of protein biomarker panels to the arsenal of cancer diagnostic tests will greatly advance personalised cancer care. The discovery that proteins and peptides are leaked from tumours into clinically accessible bodily fluids provides an opportunity for clinicians to diagnose cancer at a very early stage, monitor tumour recurrence as well as treatment response. Individual protein biomarkers in plasma have already proven extremely valuable for a number of therapeutic areas in the clinic (Table 1). However, as the field of clinical proteomics continues to grow, multiplexed biomarker assays will likely become the routine diagnostic test in clinical labs because a panel of biomarkers provides much greater sensitivity and specificity over individual analytes.

To date, there are well over 1,000 protein biomarker candidates that are believed to be associated with cancer¹. With so many candidates being discovered, one would expect the cancer diagnostic

market to be flooded with new tests entering the clinic, but this is not the case. The problem is that very few of these biomarker candidates have been validated, and less than a handful have turned into a medical diagnostic product². This discrepancy between discovery and clinical validation indicates that the biomarker development pipeline – from discovery to validation – needs to be fixed.

There are three major issues at hand:

- 1. Variability within biomarker discovery.** Proteomic technologies, including mass spectrometry and affinity-based detection methods, hold great promise for the discovery of novel cancer biomarkers. However, it is a well-known fact in the proteomics community that individual laboratories collect, store and study proteins in different ways, using a variety of platforms and practice standards. This contributes to the variability that exists within the biomarker development pipeline³. A complete lack of standard

By Dr Henry Rodriguez

Proteomics

Table 1: Protein biomarkers in plasma have shown substantial clinical value

Cardiac damage	Tnl, CK-MB, Mb, MPO, BNP
Cancer	PSA, CA-125, Her-2
Inflammation	CRP, SAA, cytokines, RF
Liver Damage	ALT,ALP,AST, GGT (enzyme assays)
Coagulation	AT-III, proteins C&S, fibrinogen,VWF
Allergy	IgE against various antigens
Infectious disease	HIV-1, Hepatitis BsAg

methods and reagents for protein identification and measurement is responsible for the pervasive problems with reproducibility and comparison of results across laboratories. Such variability has greatly hindered clinical assay validation for new biomarkers.

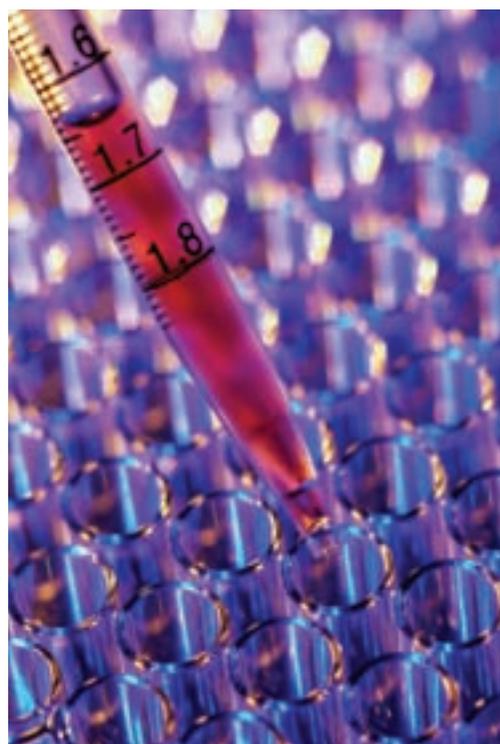
2. A gap exists in the biomarker pipeline between discovery and costly clinical validation studies. To clinically validate protein biomarkers, an enzyme-linked immunosorbent assay (ELISA) is developed for each antigen in order to test large cohorts in clinical trials. However, these tests are expensive. Each ELISA can take up to one year and millions

of dollars to develop, which is a harrowing thought for commercial development given that most protein biomarkers ultimately fail in the clinic. A more efficient biomarker development pipeline will require coupling discovery in tissue and proximal fluid to verification in plasma prior to clinical validation. The presence of such a bridge to rapidly and reliably triage a lengthy list of candidates prior to investing very large sums of time and money on the development of antibodies suitable for use in an ELISA would eliminate an enormous roadblock.

3. A lack of data sharing among the proteomics community. Most researchers recognise the importance of sharing data. However, the proteomics community faces challenges of technical variability, infrastructure standardisation and policy decisions which currently inhibit data sharing.

It takes a community

To bring the next generation of diagnostic tests to the clinic, the proteomics community must first invest in much needed technologies, resources and infrastructure in order to build a better biomarker development pipeline. Recognising the promise of clinical proteomics for personalised cancer care and that the challenges to be overcome are far too great an endeavour for a single institution, the United States National Cancer Institute (NCI) in



2006 launched the Clinical Proteomic Technologies for Cancer (CPTC) initiative.

The CPTC initiative has brought together the best minds in proteomics, building an entire community devoted to building a more refined, efficient and reliable biomarker development pipeline. CPTC is a highly collaborative public-private partnership effort, made up of scientists from more than 50 federal, academic and private-sector organisations (Table 2). Together, this multi-disciplinary team of scientists are laying the foundation for clinical proteomics through the following initiatives:

- Optimising current and emerging proteomic technologies and developing standard protocols and performance reagents so proteomics data will become reproducible across laboratories.
- Standardising procedures for collecting, processing and storing biological samples used in proteomics research because the output – the data – is only as good as the input.
- Developing high-quality reagents critical for proteomics research, including well-characterised monoclonal antibodies.
- Developing technologies that can quantify proteins across a large dynamic range.
- Developing common bioinformatics resources with shared algorithms and standards for processing, analysing and storing proteomic data.
- Implementing a verification step in the protein biomarker pipeline for triaging candidates prior to clinical validation.

Laying the foundation

In just a few short years, the CPTC community has made significant advances in the field that will affect the way every investigator does protein biomarker research. These can be broken down into three major milestones:

Community resources. Discussions with representatives from all parts of the cancer research community revealed a deep concern about the lack of access to affordable, well-characterised and validated affinity reagents and supporting resources. In order to drive the development of a central community core that would help accelerate biomarker discovery and validation, cancer diagnostics development and therapeutics monitoring, NCI launched the Proteomic Reagents and Resources Core. This programme within CPTC provides tools, reagents, enabling technologies and other critical resources to support protein/peptide measurement and analysis efforts.

Table 2: All organisations 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
European Bioinformatics Institute
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
Intrinsic Bioprobes, Inc
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
National Institute of Standards and Technology
New York University Medical Center
Northeastern University
Plasma Proteome Institute
Predictive Physiology and Medicine, Inc
Purdue University
Quadrascop, Inc
Rana Biosciences
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

Proteomics

Table 3: Examples of reagents and resources being made available to the scientific community

REAGENTS	
●	Reference Materials
●	Multiple Reaction Monitoring (MRM) Peptides
●	MRM Standard Kits
●	Antibodies
●	Hybridomas
●	Expression Vectors
●	Mouse Biospecimens
RESOURCES	
Tranche	
●	Standard Operating Procedures
●	Software Packages
●	FDA Documents ('Mock' 510K)

In October 2008, the Reagents and Resources Core announced the launch of the Reagents Data Portal, a Web-based service created to make all reagents and resources developed through the CPTC initiative available to the scientific community for little to no cost. The Reagents Data Portal can be accessed through the CPTC website at <http://proteomics.cancer.gov/>.

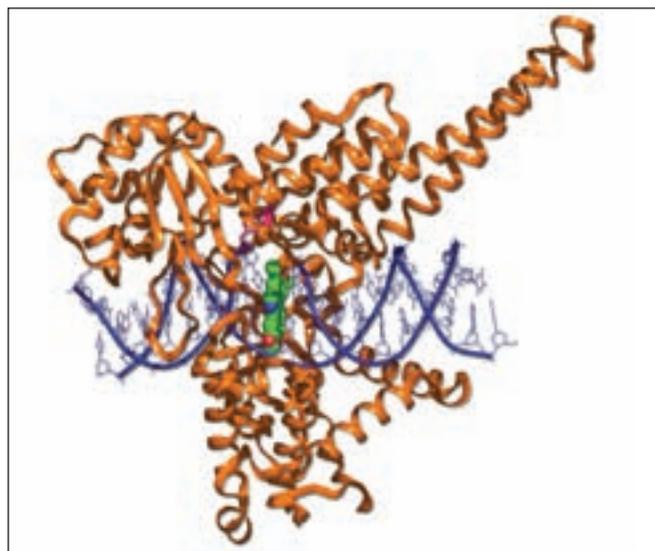
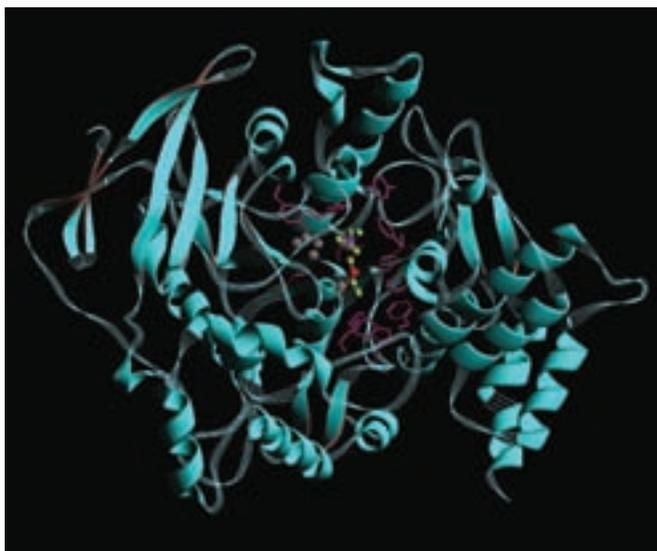
The Reagents Data Portal is continually expanding as the initiative makes way for a great number of reagents and resources in the pipeline that are needed for effective proteomic analysis (Table 3). To date, more than 25 antigens and 75 monoclonal antibodies have been generated against human cancer-associated proteins and each antibody is added to the web portal once initial characterisation (isotype, SDS-PAGE, Western Blot and ELISA) has been generated. In addition, more than 25 soft-

ware programs will soon be added to the portal to assist in biomarker discovery and verification.

Restructuring the biomarker development pipeline. Using 'shotgun' proteomics, hundreds to thousands of biomarker candidates are typically discovered at a time, most of which are false positives. There is no standard process however for triaging this lengthy list of candidates to identify the true biomarkers, those worth pursuing in the clinic. This creates a huge roadblock for clinical translation because the lack of a bridge between discovery and validation significantly raises the time and cost of clinical validation studies because an ELISA needs to be developed for each analyte so that large cohorts of patients can be tested in clinical trials.

CPTC is providing a solution to this problem by developing a new and improved biomarker development pipeline, which includes a biomarker verification, or pre-validation, step. Using targeted proteomic technologies, verification is a rapid and cheap way to assess if a given candidate is detectable in blood – critical for clinical utility – and changes in a measurable way in relation to the presence or stage of disease. This new step, once implemented, will serve as a bridge between biomarker discovery and clinical validation.

The verification technology being tested is by no means new. Multiple reaction monitoring (MRM), a mass spectrometry technology, has been around for decades, but it is only now being 're-purposed' for clinical proteomics. MRM may ultimately provide a very reliable GO/NO GO decision point in the new biomarker development pipeline, and this extra step could save the



medical diagnostic industry millions of dollars and many years of development because only the strongest candidates will move into clinical validation – and with much greater confidence.

The CPTC community is currently working to make MRM highly sensitive, on a level comparable to an ELISA, and reproducible across laboratories and platforms, so the data can be trusted.

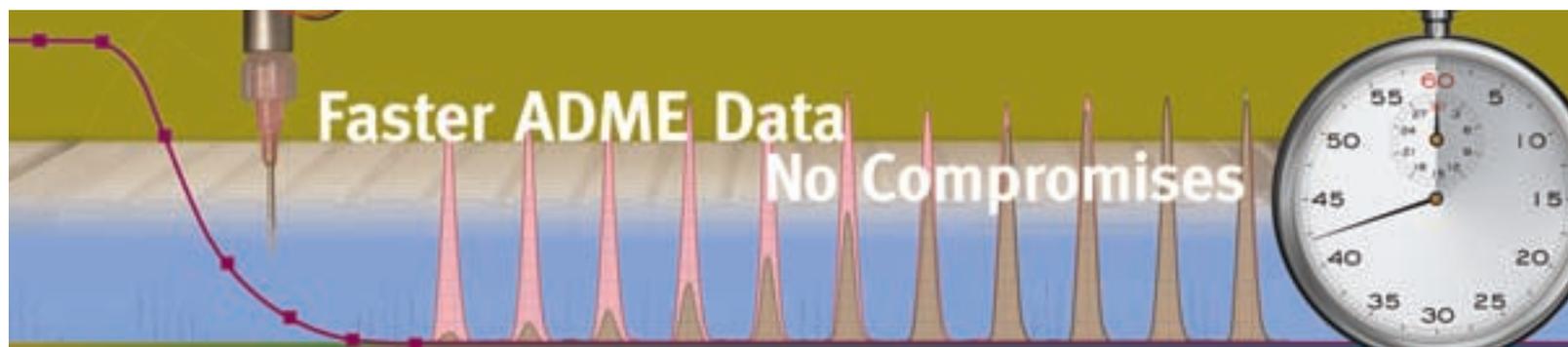
Promoting data release and sharing: the Amsterdam Principles. Advancements in science and healthcare are only made possible through widespread and barrier-free access to cutting-edge research and knowledge (data). A clear example of the benefits of data sharing comes from the explosion of available sequencing data through GenBank, which has experienced exponential growth over the past two decades. No one argues that science is better off because of public data; indeed, it is abundantly clear that the sum of the data is far greater than its individual parts.

Open data:

- Promotes and reinforces open scientific inquiry, allowing a researcher's conclusions to be validated or refuted by his or her peers,
- enables new analyses to be performed on existing data which may lead to new insights,
- provides large training and test sets for quality assessment,
- accumulation of public data provides all researchers with access to a data set that is larger than one that could ever be constructed by a single laboratory; and
- prevents unnecessary duplication of effort, although some duplication provides rigor.

On August 14, 2008, CPTC convened the 'International Summit on Proteomics Data Release and Sharing Policy' in Amsterdam, the Netherlands, to identify and address roadblocks to rapid and open access to proteomics data.

The Amsterdam Principles, an output of the



Decrease your *in vitro* ADME assay bottlenecks by streamlining your drug discovery workflow with the RapidFire® 300 Mass Spectrometry System.

The RapidFire System integrates fast, serial, on-line sample purification with mass spectrometric analysis.

- Label-free, native detection technology requires no fluorescence, luminescence or radioactivity
- High-throughput – 6 seconds per sample – obtain 7 point IC₅₀ curves in just 42 seconds
- Pre-validated ADME assays include:
 - CYP Inhibition
 - Metabolic Stability

- Complements existing assays & workflows
- Compatible with leading manufacturers' mass spectrometers
- Multiple analysis methods possible with a single setup
- Enables hassle-free unattended 24 hour operation
- Use 96 or 384 well plates

BioTrove offers ADME screening services with fast turn-around times or instrumentation for your laboratory. Contact us today and obtain mass spec quality data at the speed of plate reader-based methods!



12 Gill Street, Suite 4000, Woburn, MA 01801-1728 | 781-721-3695 | www.biotrove.com

Proteomics

References

- 1 Polanski and Anderson. Biomarker Insights. 2006; 2:1–48.
- 2 Ludwig and Weinstein. Nature Reviews Cancer. 2005 Nov; 5: 845-856.
- 3 Clinical Proteomic Technologies for Cancer. 2007 Annual Report: Building the Foundation for Clinical Cancer Proteomics. Available at: <http://proteomics.cancer.gov/library/reports.asp>. Accessed November 18, 2008.
- 4 Rodriguez, Henry, Snyder, Mike, Uhlen, Mathias, Andrews, Phil, Beavis, Ronald, Borchers, Christoph, Chalkley, Robert J, Cho, Sang Yun, Cottingham, Katie, Dunn, Michael, Dylag, Tomasz, Edgar, Ron, Hare, Peter, Heck, Albert J.R, Hirsch, Roland F, Kennedy, Karen, Kolar, Patrik, Kraus, Hans-Joachim, Mallick, Parag, Nesvizhskii, Alexey, Ping, Peipei, Ponten, Fredrik, Yang, Liming, Yates, John R, Stein, Stephen E, Hermjakob, Henning, Kinsinger, Christopher R, Apweiler, Rolf. Recommendations from the 2008 International Summit on Proteomics Data Release and Sharing Policy: The Amsterdam Principles. J. Proteome Res. [ePub 2009 Apr 3].



summit, was an effort by the global proteomics community to establish guidelines that will encourage and enable proteomics data sharing. The Amsterdam Principles address issues surrounding 1) timing, 2) comprehensiveness, 3) format, 4) deposition to repositories, 5) quality metrics, and 6) responsibility for proteomics data release. A summit report has been published that explores a framework for data release and sharing principles that will most effectively fulfill the needs of the community⁴. The release of high quality data following standardised approaches will put the pace of proteomic research on a trajectory similar to that seen in large-scale genomics research.

Conclusion

As the human genome has now been sequenced, the elucidation of protein function is the next challenge toward the understanding of biological processes in health and disease – specifically cancer. In the post genome sequencing era, the major challenge in understanding a complex disease such as cancer is to unravel the key biological functions and interactions of altered genes and their products. To meet these scientific challenges, resources and strategies need to be developed that are suitable to analyse large numbers of altered cancer genes and associated proteins in parallel, and to achieve high throughput analysis that provides meaningful and signifi-

cant information. NCI's CPTC is leading the effort in providing a foundation for pursuing the development of human cancer proteomes. But this work will have implications far beyond cancer, ultimately impacting the diagnosis and treatment of all human disease.

DDW

Dr Henry Rodriguez is the Director of the National Cancer Institute's Clinical Proteomic Technologies for Cancer programmes. With training in molecular diagnostics, regulatory affairs, policy and business finance, he currently manages the efforts of an interdisciplinary network of clinical and analytical centres dedicated to the search of proteins and other molecules known as biomarkers that may indicate cancer long before a tumour is detected. Dr Rodriguez's initiatives are paradigm shifting in that they aim to develop new, more refined, efficient and reliable biomarker discovery and verification pipelines. These pipelines are anticipated to produce better credentialled candidate leads, ultimately accelerating the discovery of new cancer biomarkers. Dr Rodriguez earned his BS in biological sciences at Florida International University, his MS in toxicology/biological sciences at Florida International University, his PhD in molecular and cellular biology at Boston University, and his MBA at Johns Hopkins University School of Business.