Mass Spectrometry in Cancer Diagnostics

Gunjan Malik, Ph.D.

Postdoctoral Fellow
The George L. Wright Jr. Center for Biomedical Proteomics,
Department of Microbiology and Molecular Cell Biology
Eastern Virginia Medical School Norfolk, VA

Eastern Virginia Medical School
Discovery Laboratory
Early Detection Research Network

Virginia Prostate Center

The George L. Wright, Jr. Center for Biomedical Proteomics
Eastern Virginia Medical School

NCI Early Detection Research Network
EVMS Biomarker Discovery Laboratory

O. John Semmes, P.I.

Proteomic Profiling of CANCERS

- Prostate
- Head and Neck
- Breast
- Bladder
- Leukemia
- Ovarian
- Colon
- Liver
- Lung

The mission of the Center is to discover a means to detect and accurately diagnose a variety of cancers long before the disease becomes life-threatening. Using mass spectrometry-based techniques to visualize and identify proteins, we hope to discover biomarkers that will detect cancer at its earliest stage, as well as markers that can predict disease and treatment outcome.

The George L. Wright, Jr.
Center for Biomedical Proteomics
Eastern Virginia Medical School

Some Facts: Why proteomics ?

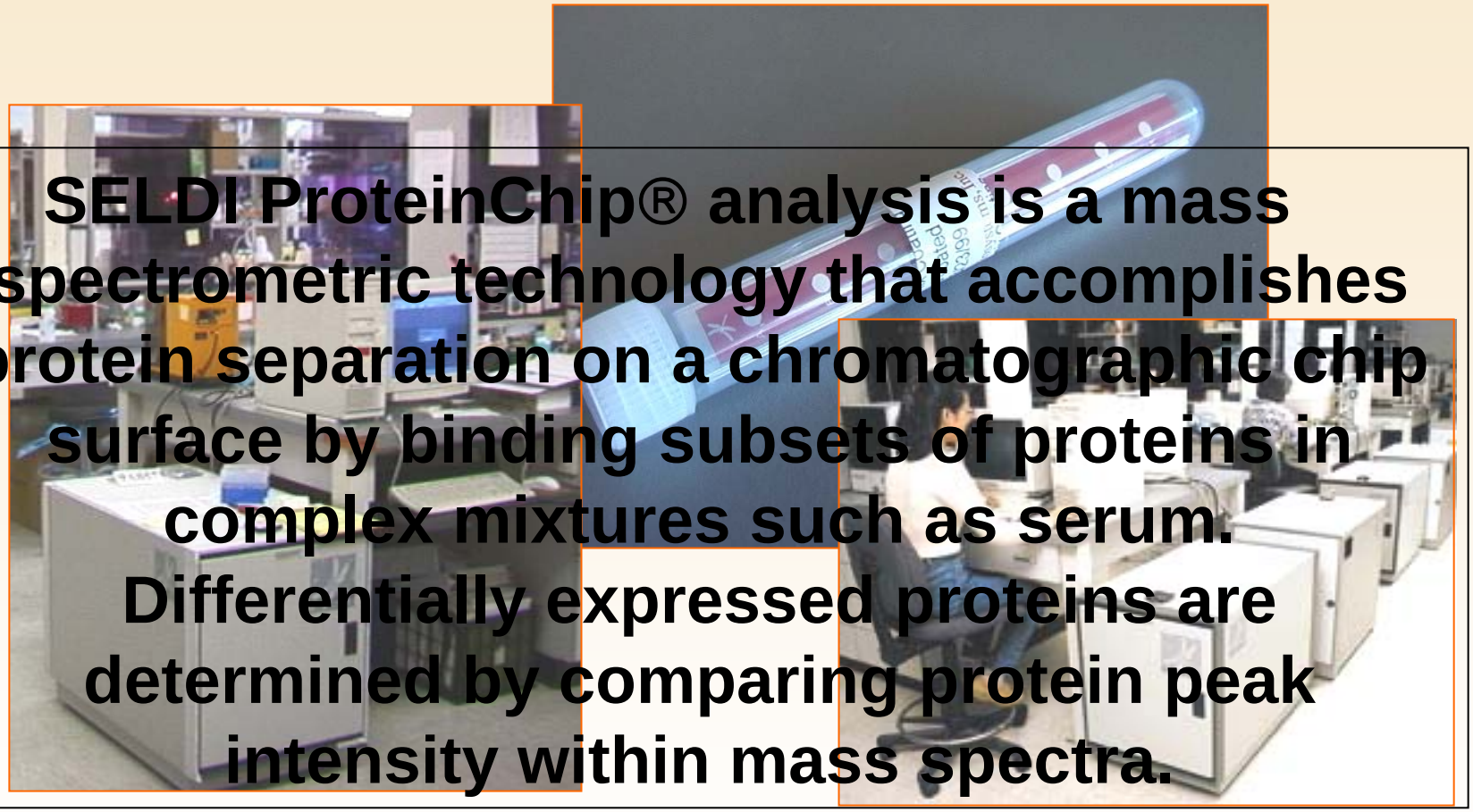
- More than 98% of all diseases are caused by multiple molecular alterations.
- There is very low correlation between mRNA abundance and protein level.
  (e.g., yeast – *S. cerevisiae*: <0.5 correlation factor)
- Gene products are modified by
  - complex gene interactions
  - cellular events
  - environmental influences

  (co- & post-translational modifications: >189 types)

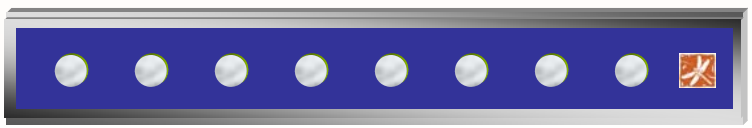
Pieces of the Puzzle

- Data acquisition, pre-processing, processing, signal versus noise
- Data interpretation phase 1, which signals are potential biomarkers
- Determining sources of variability, response to bias, analytical reproducibility
- Data interpretation phase 2, biomarker maturation

Ciphergen ProteinChip® Technology



SELDI ProteinChip® analysis is a mass spectrometric technology that accomplishes protein separation on a chromatographic chip surface by binding subsets of proteins in complex mixtures such as serum. Differentially expressed proteins are determined by comparing protein peak intensity within mass spectra.



The ProteinChip discovery platform

Expression Profiling	Purification	Characterization & I.D.	Assay Development
Rapid discovery of biomarkers: •Serum •Urine •Cell lysates •Cells •Cerebrospinal fluid •Tissue homogenates	Purification: •On-chip •Micro-chromatography spin columns •Via traditional purification strategies	Characterization and identification through: •Epitope mapping •Peptide mapping •Phosphorylation •Peptide sequencing •Glycosylation analysis •Binding domain analysis	ProteinChip array is the assay. •Antibody ProteinChip for low CV's and accurate quantity

Samples and Requirements:

Body Fluids:

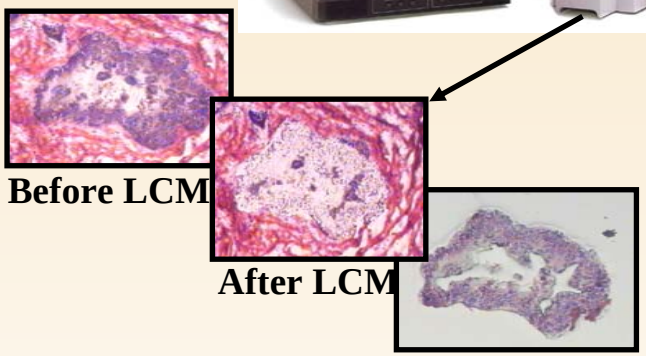
- Serum- 20µl
- Urine- 60µl
- Seminal Plasma- 20µl
- Cerebral Spinal Fluid- 60µl

Cell and Tissue Lysates:

- Microdissected cells-2000-3000
- Tissue Culture cells- whole cell lysates
- Subcellular fractionation,

(i.e. Mitochondrial or membrane lysates)

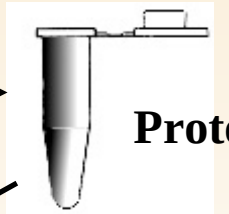
PixCell™ Laser Capture Microdissection Microscope



Before LCM

After LCM

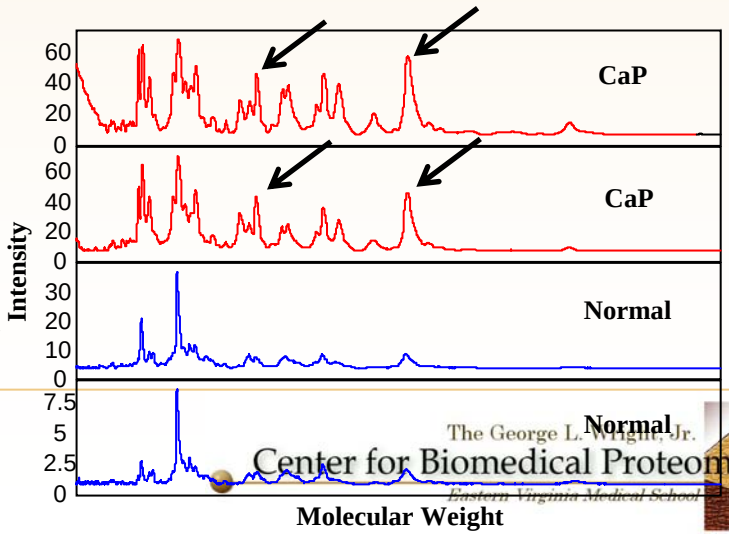
Captured Cells



Protein Extraction

ProteinChip™

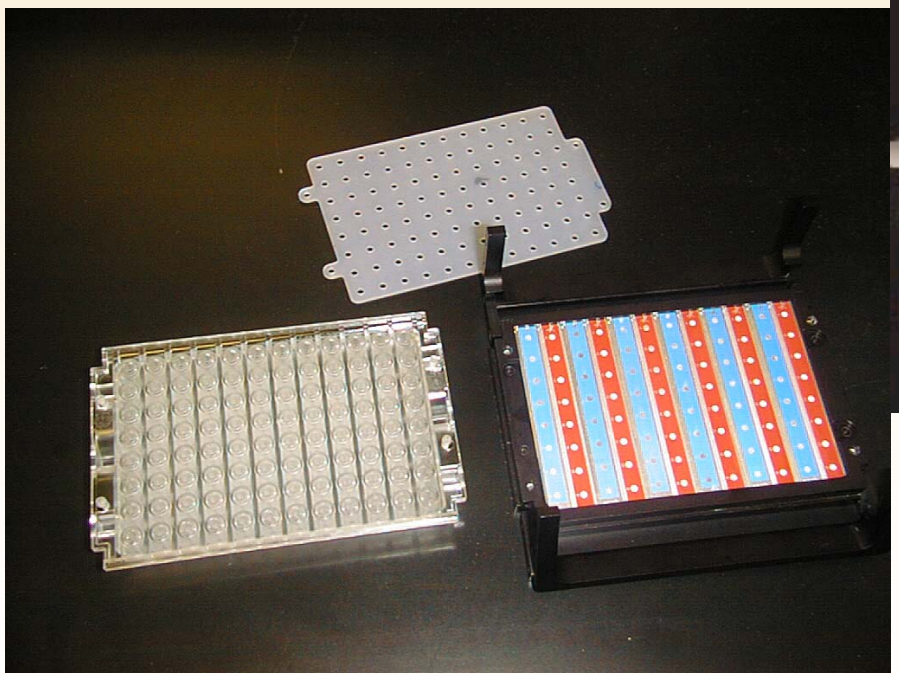
Surface Enhanced Laser Desorption/ Ionization: Time of Flight Mass Spectrometry SELDI-TOF-MS

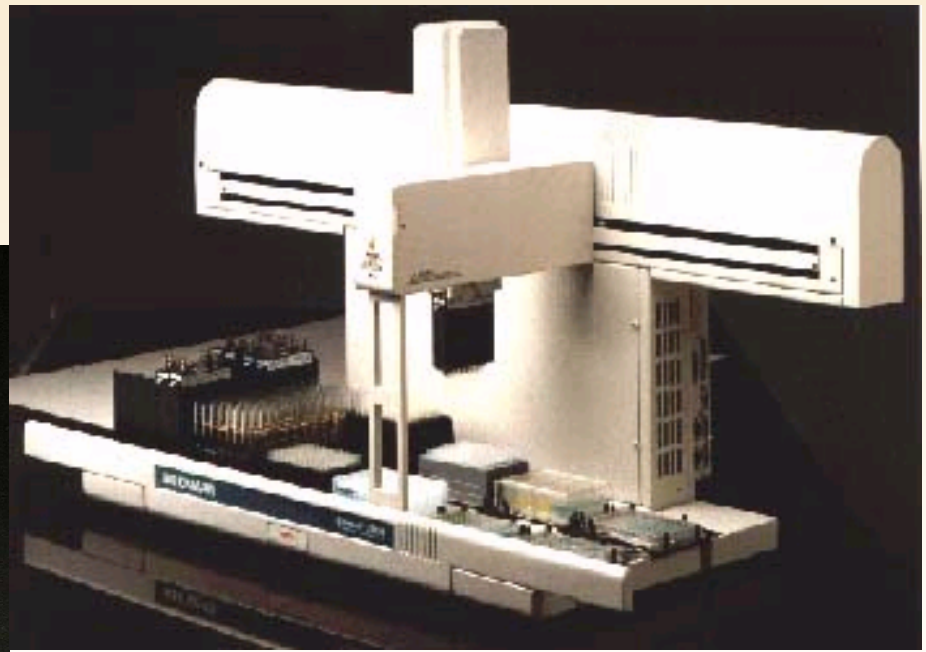


The George L. Wright, Jr. Center for Biomedical Proteomics
Eastern Virginia Medical School

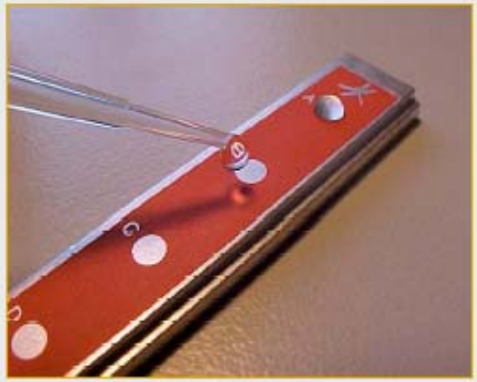
Biomek® 2000 Laboratory Automation Workstation

Process 48 chips/384 samples per day

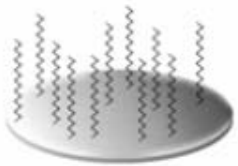




ProteinChip® Arrays: A Variety of Surfaces



Chemical Surfaces – Protein Expression Profiling:



Hydrophobic

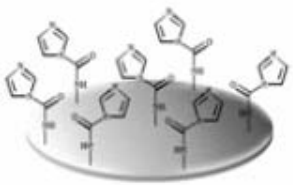
Anionic

Cationic

Metal Chelate

Normal Phase

Biological Surfaces – Protein Interaction Assays:



PS-10 or PS-20

Antibody - Antigen

Receptor - Ligand

DNA - Protein

CIPHERGEN®

ProteinChip® Array Preparation

Step 1: Complex protein sample is placed on a ProteinChip Array

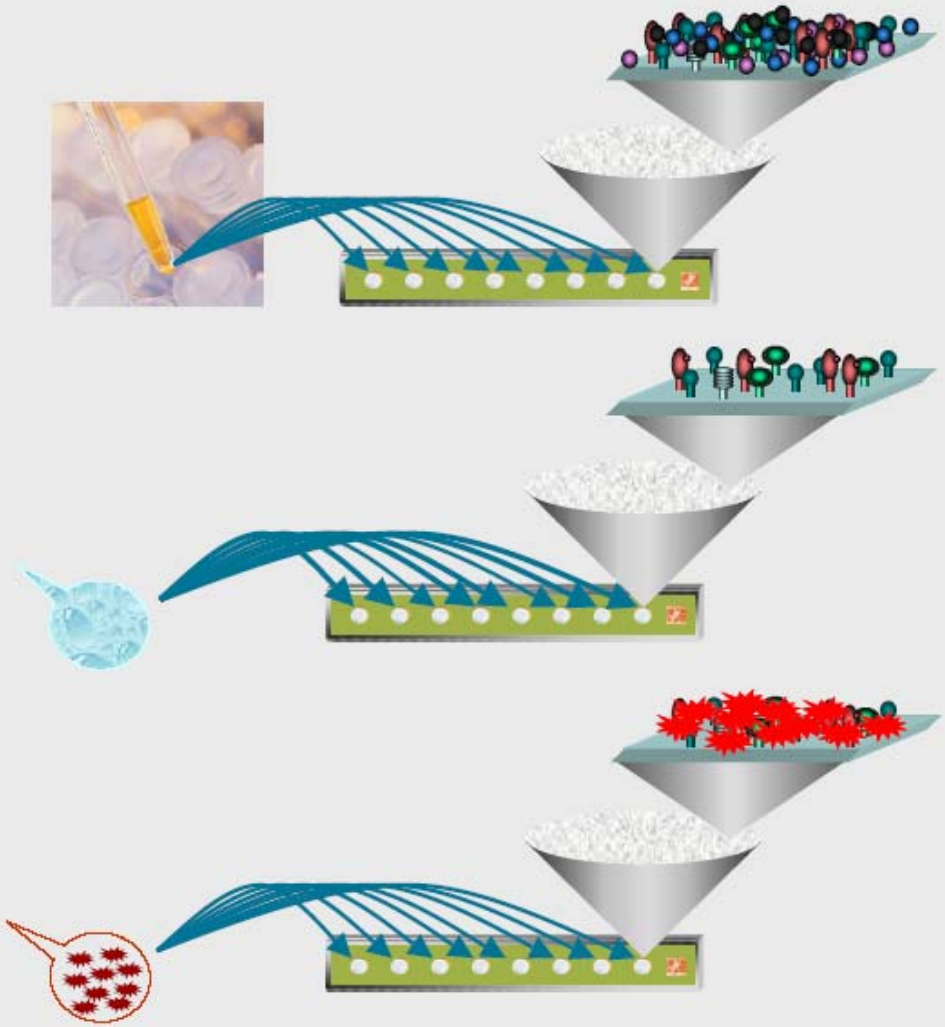
- Proteins bind to chemical or biological sites on the ProteinChip surface

Step 2: Remove unbound proteins

- Wash the ProteinChip with appropriate stringency buffer
- Bound proteins are retained

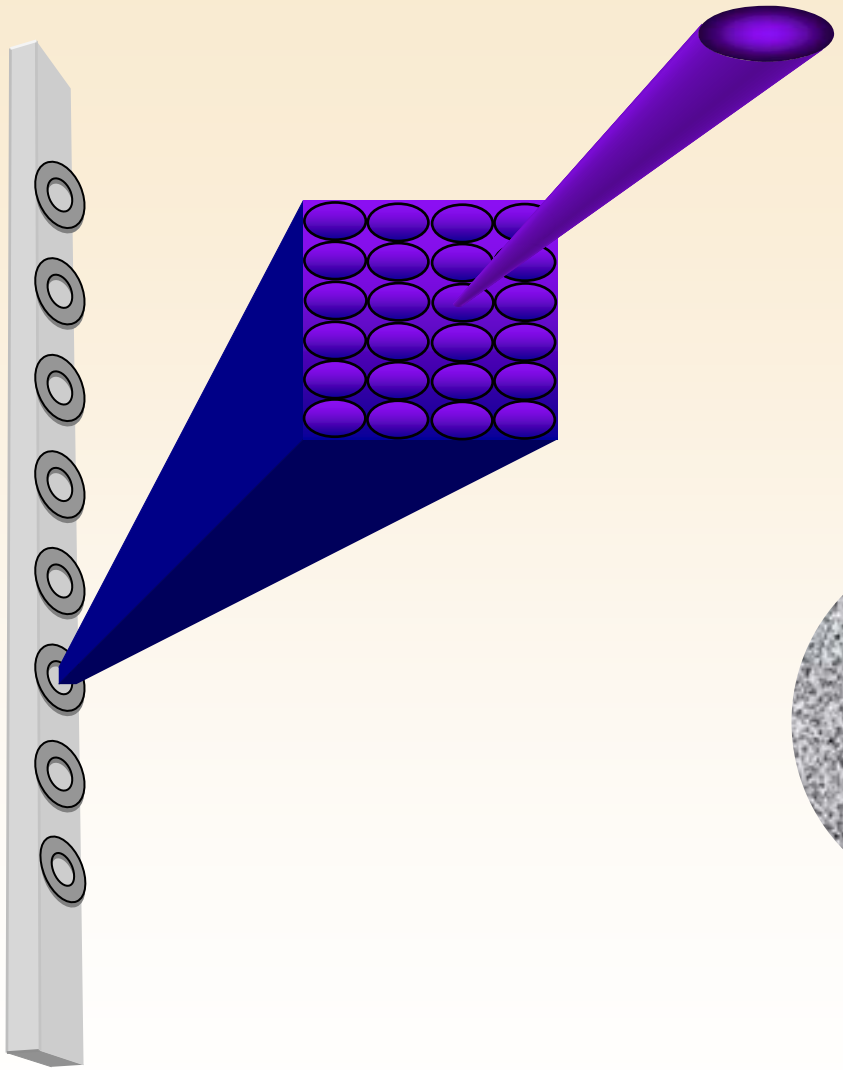
Step 3: Add Energy Absorbing Molecules or "Matrix"

- EAM is applied to each spot to facilitate desorption and ionization in the TOF-MS Chip Reader

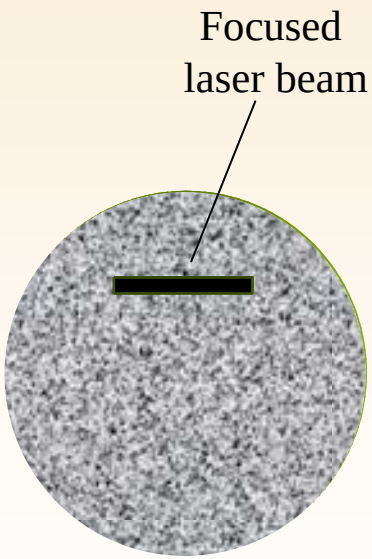


CIPHERGEN®

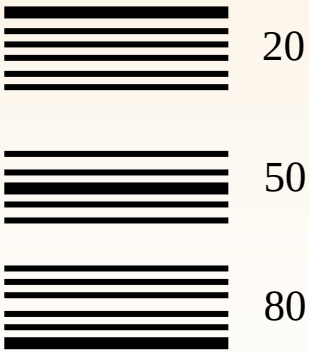
SELDI BioChip Arrays



} Addressable locations
100 µm x 100 µm units
(10,000 per cm^2)

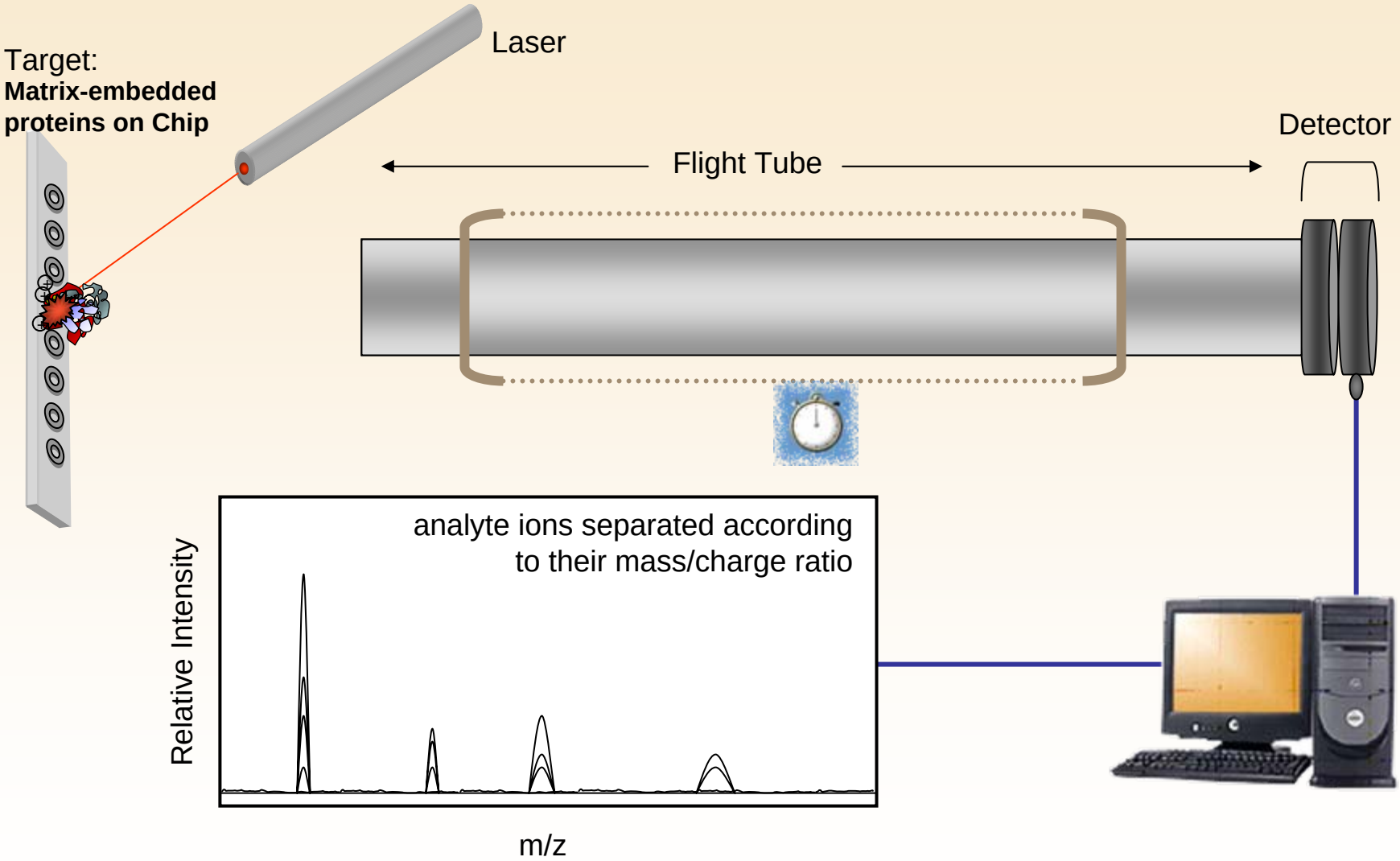


Sample spot

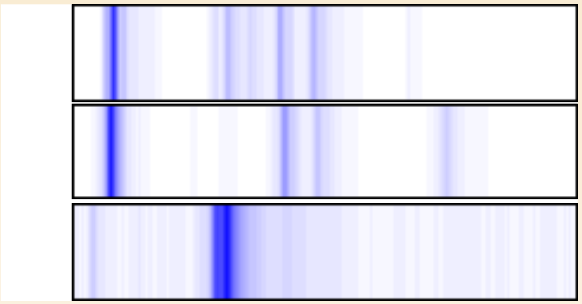


Laser queries multiple
Positions on sample spot

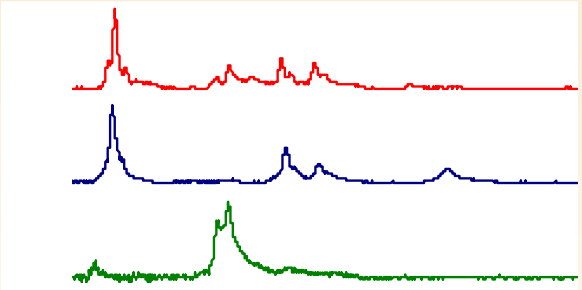
Laser Desorption/Ionization
Time-Of-Flight (TOF) Mass Spectrometry



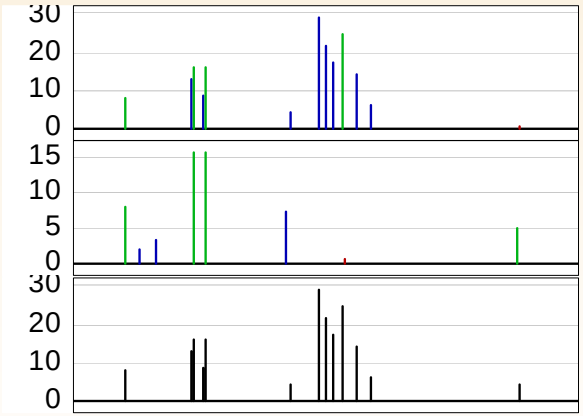
Gel view



Spectra



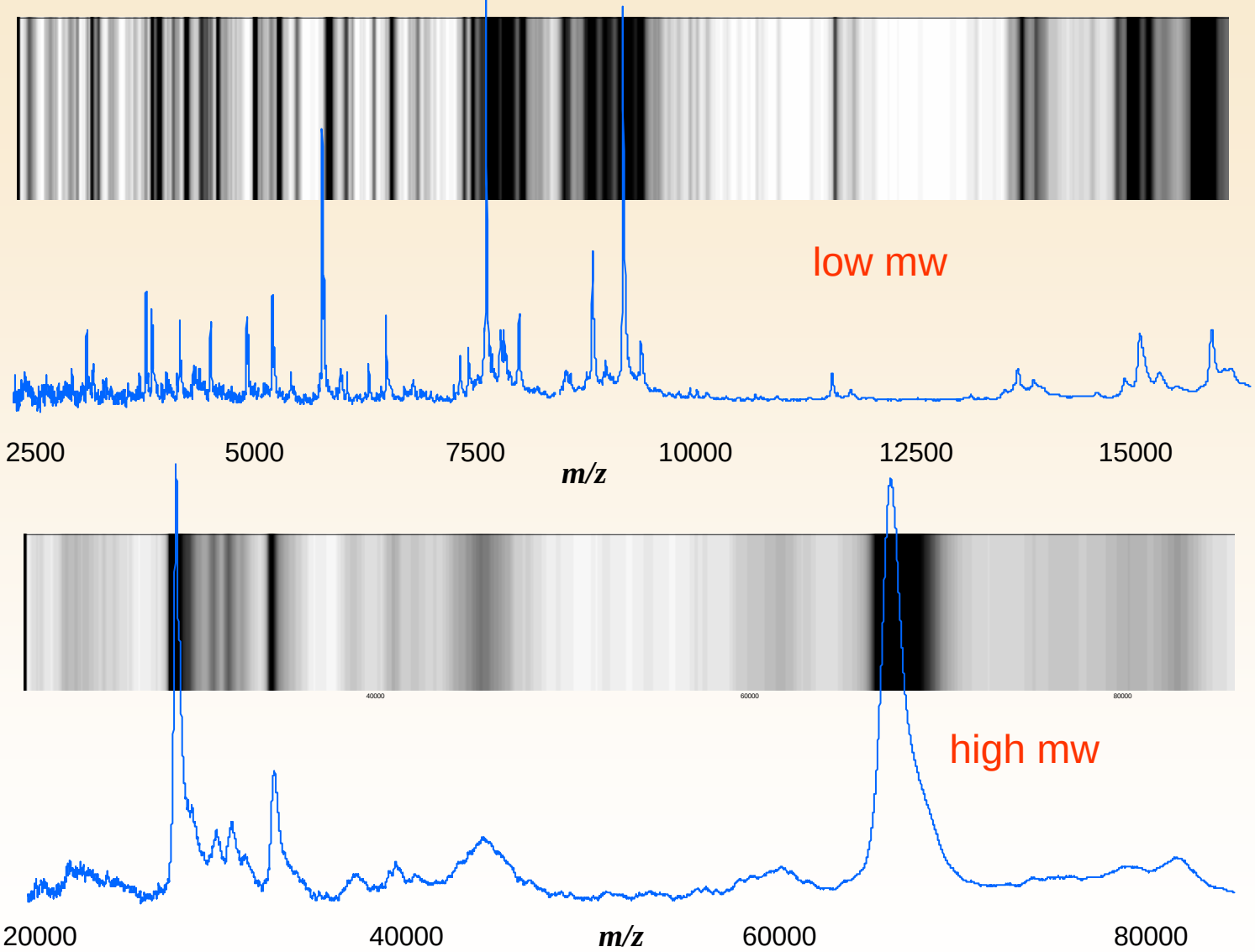
Histogram



Data presentation options

An example of SELDI output

200 peaks for serum (upto 200K Da)

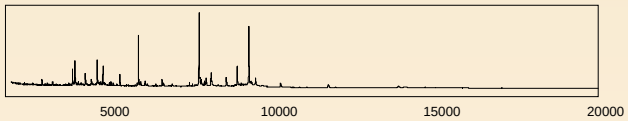


Pieces of the Puzzle

- Data acquisition, pre-processing, processing, signal versus noise
- Data interpretation phase 1, which signals are potential biomarkers
- Determining sources of variability, response to bias, analytical reproducibility
- Data interpretation phase 2, biomarker maturation

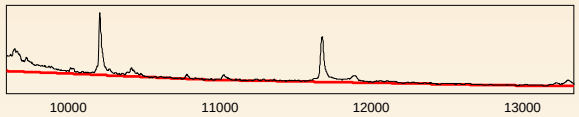
Flow diagram of spectrum analysis

SELDI-TOF MS

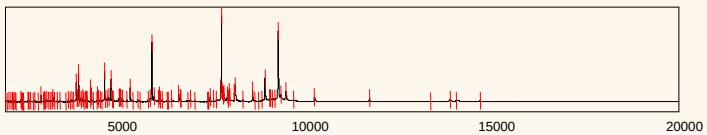


2-20 KDa

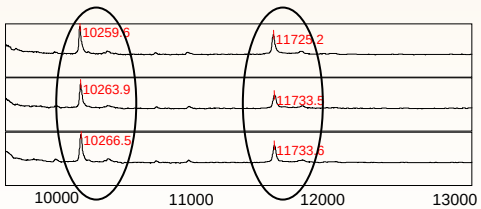
Baseline subtraction
Normalization



Peak detection



Peak alignment and clustering

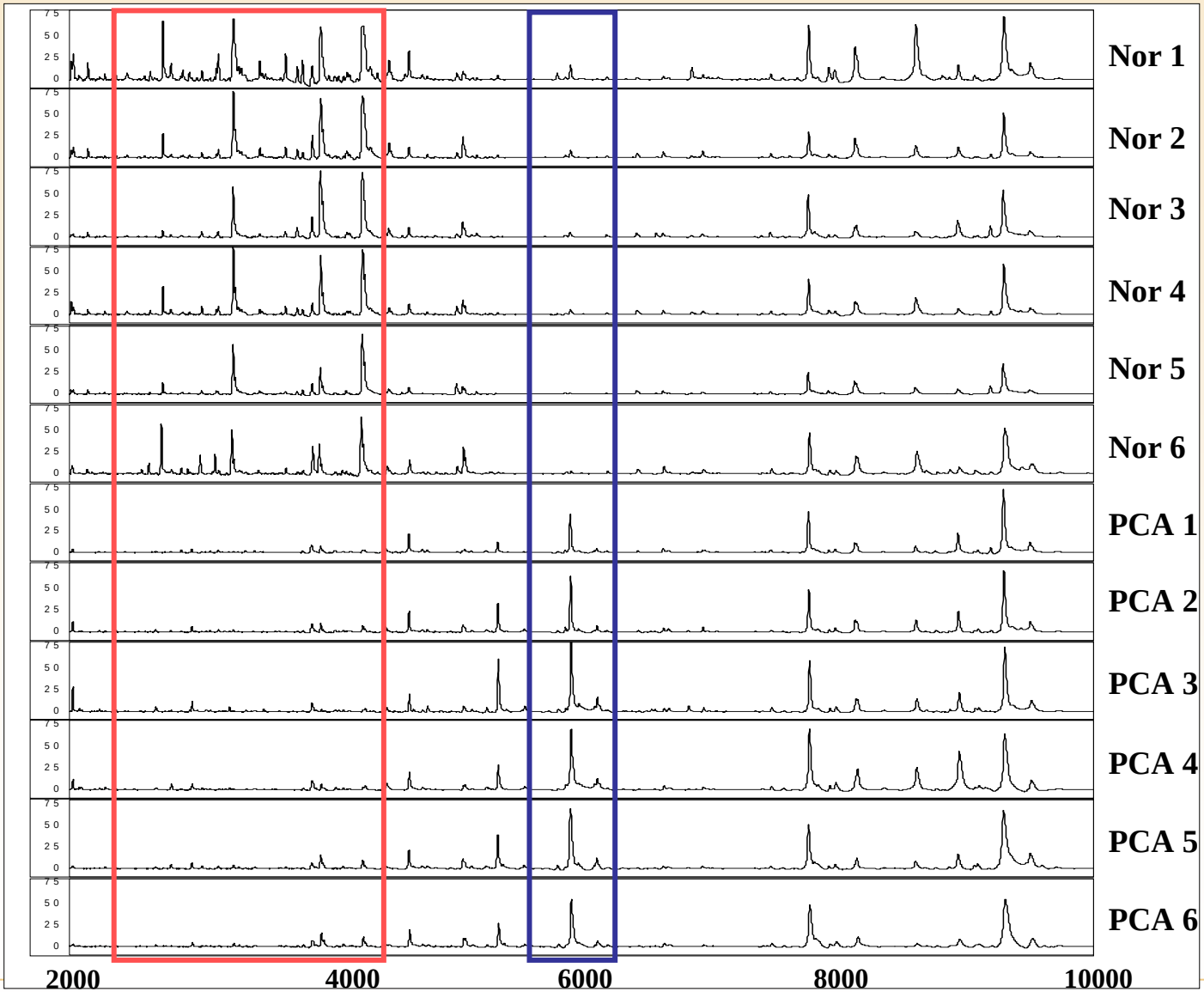


Duplicates averaged

80 peaks

Classification (Decision Tree)

SELDI Profiles of Normal vs. Cancer Serum Proteins



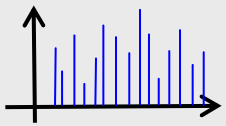
The George L. Wright, Jr.
Center for Biomedical Proteomics
Eastern Virginia Medical School

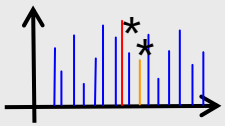
Step 1: Discovery

Training data set

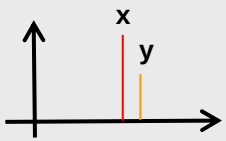
Disease → **Profile 1**

Normal → **Profile 2**





Pattern discovery



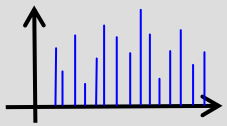
Use biomarker pattern for step 2.

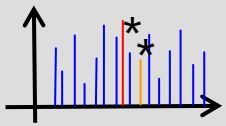
Step 2: Evaluation

Test data set

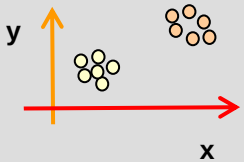
Disease → **Profile 1**

Normal → **Profile 2**





Cluster analysis



Determination of:

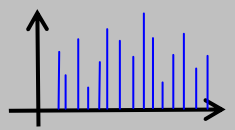
- Sensitivity
- Specificity
- Positive predictive value
- Negative predictive value

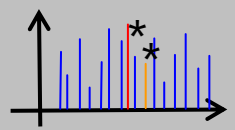
Step 3: Class prediction

Unknown data set

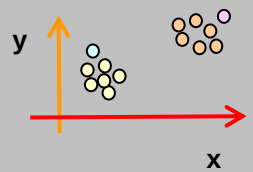
Profile 1

Profile 2





Cluster analysis



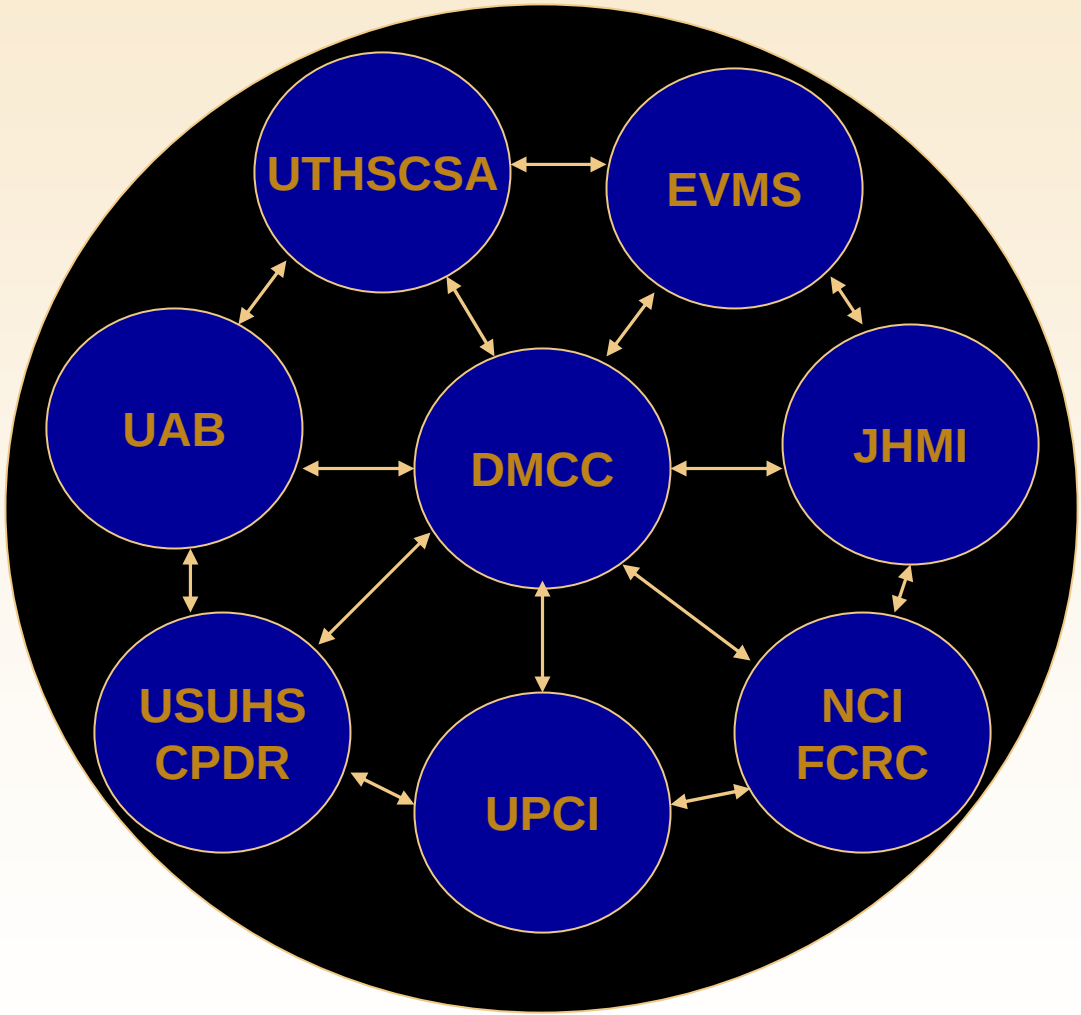
Disease **Normal**

The George L. Wright, Jr.
Center for Biomedical Proteomics
Eastern Virginia Medical School

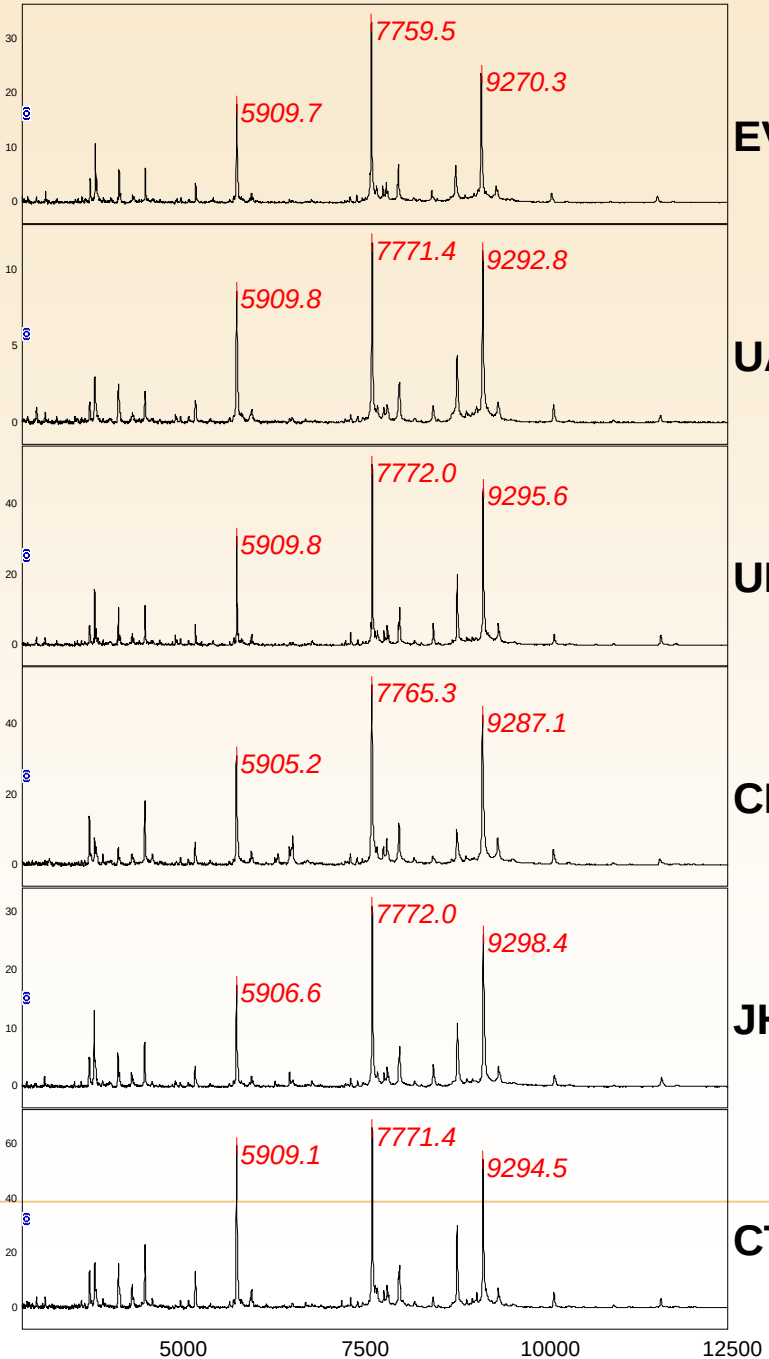
Pieces of the Puzzle

- Data acquisition, pre-processing, processing, signal versus noise
- Data interpretation phase 1, which signals are potential biomarkers
- Determining sources of variability, response to bias, analytical reproducibility
- Data interpretation phase 2, biomarker maturation

Validation Team



Semmes OJ and EPSIC members. Clin Chem. 2005 Jan;51(1):102-12.



EVMS

UAB

UPitt

CPDR

JHU

CTRC San Antonio

Table 2b Inter-Lab variability

		Mass	Intensity	S/N	Resolution
Peak 1	**average**	5906.47	26.57	163.06	460.73
	stdev	6.70	9.67		107.72
	CV	**0.0011**	**0.36**		**0.23**
Peak 2	**average**	7768.61	35.94	242.75	505.54
	stdev	8.41	6.25		82.77
	CV	**0.0010**	**0.17**		**0.16**
Peak 3	**average**	9289.18	30.96	244.03	439.28
	stdev	9.89	4.70		77.35
	CV	**0.0011**	**0.15**		**0.18**

The George L. Wright, Jr.
Center for Biomedical Proteomics
Eastern Virginia Medical School

SELDI-Prostate Cancer Validation

- **Phase I:** Synchronization of SELDI Instruments and Validation of Robotic Sample Processing: Can each site synchronize their instruments and procedures to match other sites?

  (STATUS: **Successful Completion last year**)

- **Phase II**: Population Validation: Can multiple sites get the same result in a geographically diverse cross-section study?

  (STATUS: **Initiating this year**)

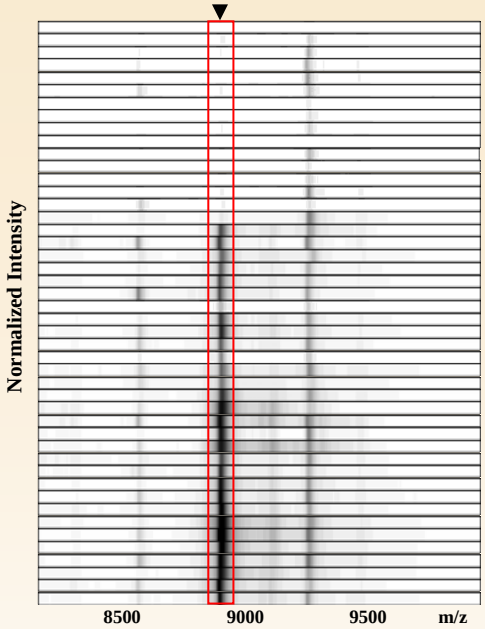
- **Phase III:** Clinical Validation: Can multiple sites get the same result in a True Early Detection Analysis?

Pieces of the Puzzle

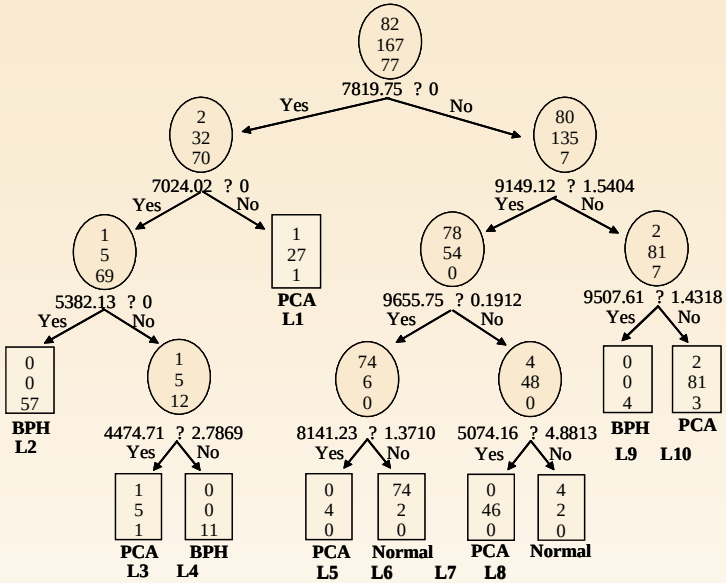
- Data acquisition, pre-processing, processing, signal versus noise
- Data interpretation phase 1, which signals are potential biomarkers
- Determining sources of variability, response to bias, analytical reproducibility
- Data interpretation phase 2, biomarker maturation

Summary of Biomarker Discovery and Identification

SELDI-TOF



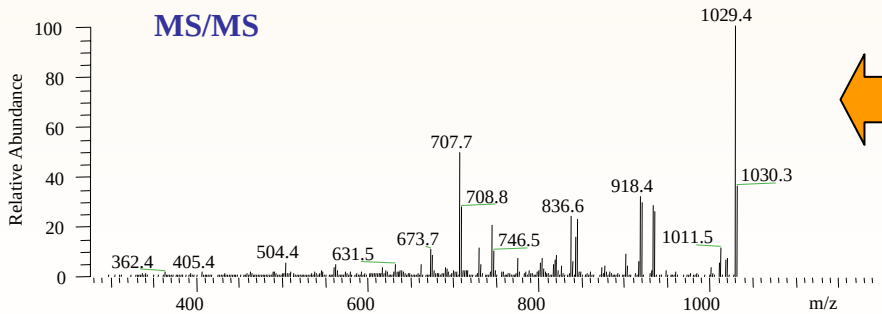
Classification and Regression Tree Analysis



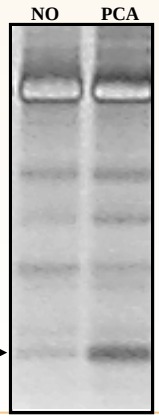
Validation

Western

Identification



Purification

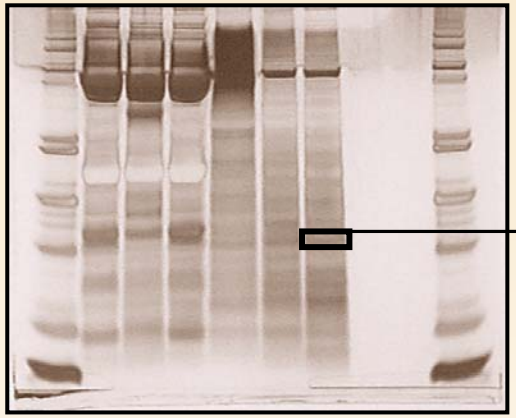


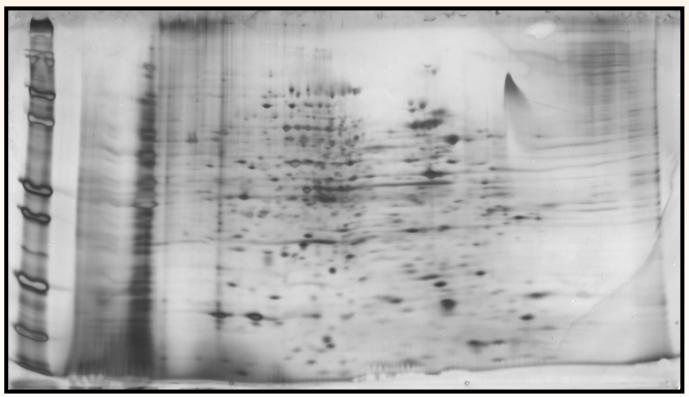
SDS-PAGE

The George L. Wright, Jr. Center for Biomedical Proteomics
Eastern Virginia Medical School

Sample Preparation

1-Dimensional PAGE





2-Dimensional PAGE

- Remove Stain
- Dehydrate with 100% Acetonitrile
- Dry in Speed Vac
- Re-hydrate slice in 12.5 ng/µL Trypsin Solution
- Incubate 45 minutes on Ice
- Add Ammonium Bicarbonate to Keep Moist
- Incubate Overnight
- Extract with Formic Acid/Acetonitrile - Dry in Speed Vac
- Resuspend in Buffer “A”

For each scan in the Chromatogram, there are 3 “parent” ion masses each with an associated list of fragment ions:

Peptide List (1046.7):

1029.4	16410.2
918.4	5211.5
836.6	6187.2
746.5	817.3
707.7	1181.3
673.7	10527.8
…..	

Parent Ion

Fragment Ions
+
Intensity

There are ~ 2500 scans/sample with a 90 min gradient

TurboSequest Search

Peptide List (1046.7):

1029.4	+2
918.4	+1
836.6	+1
746.5	+2
707.7	+2
673.7	+1
…..	

Human NR Database

Peptide List (gi | 148566):

769.2	+1
740.9	+2
1765.3	+1
1634.8	+2
879.5	+3
1028.1	+1
…..	

XCorr = 0.02

Peptide List (gi | 23997):

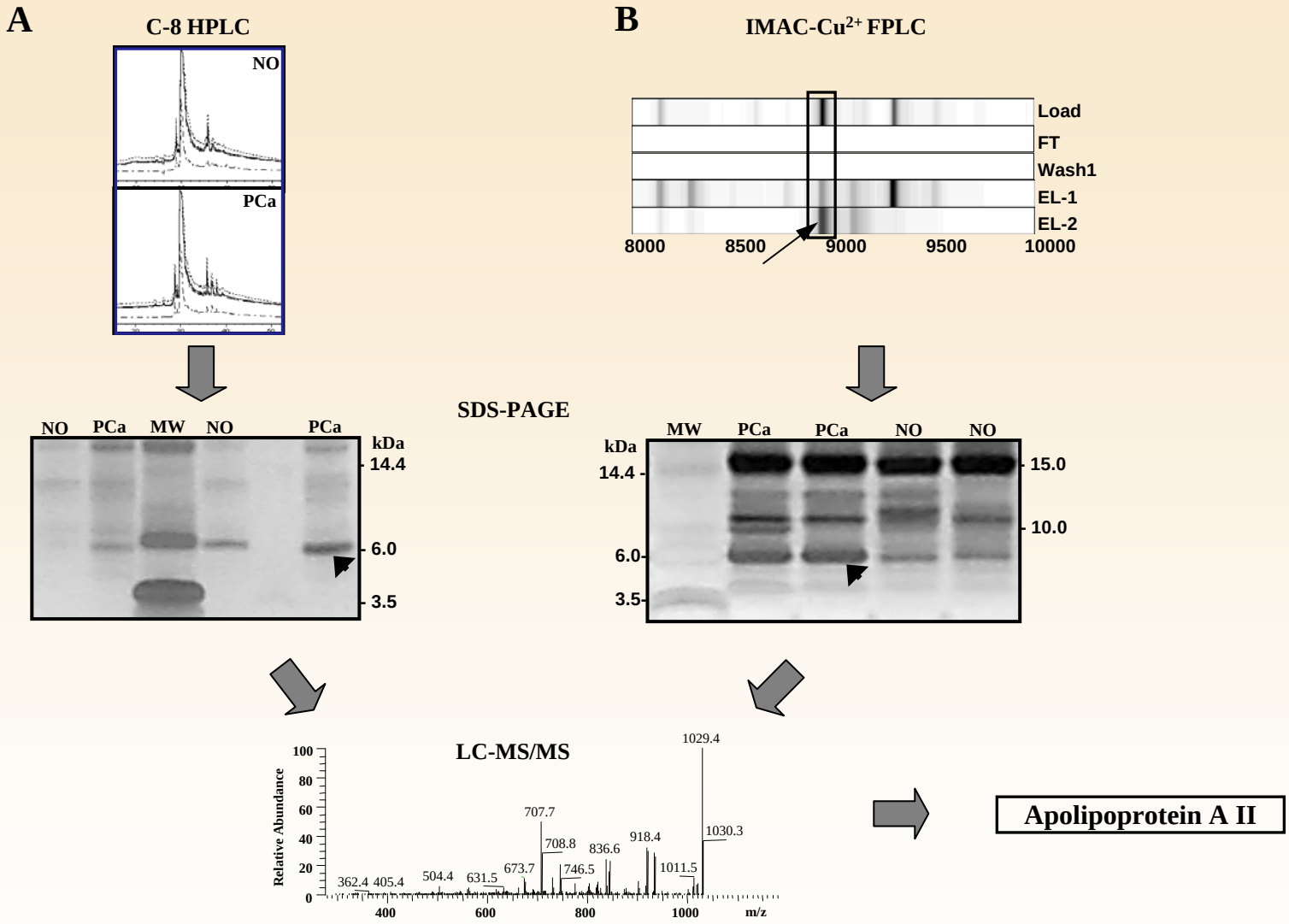
1029.1	+2
918.4	+1
836.5	+1
746.5	+2
707.2	+2
673.8	+1
…..	

XCorr = 4.73

Peptide List (gi | 428418):

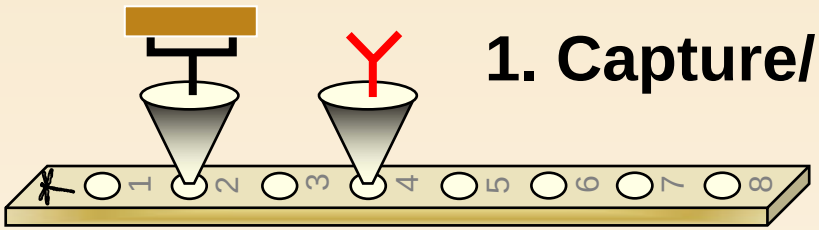
744.6	+1
1929.5	+1
1143.7	+2
1417.6	+2
862.3	+3
1596.8	+2
…..	

XCorr = 0.00

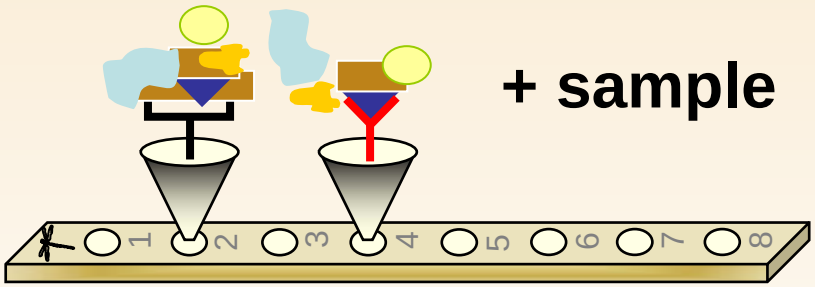


Malik, G. et al. *Clin. Cancer Res.* 2005 Feb 1;11(3):1073-85.

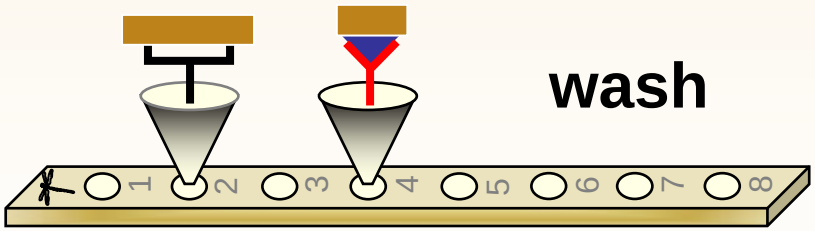
SELDI ProteinChip® Immunoassay



1. Capture/retention



+ sample



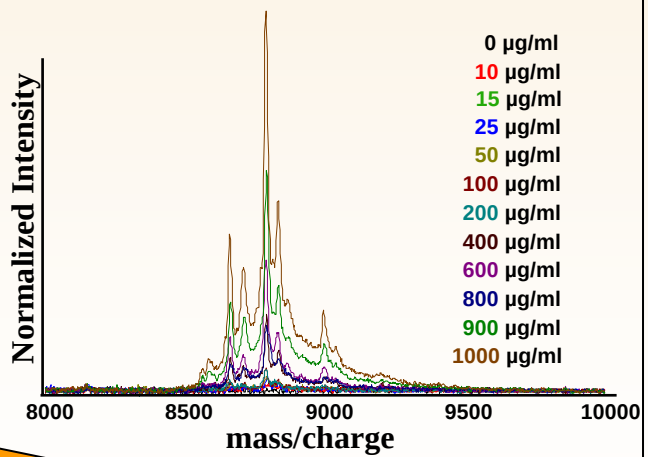
wash

Control Ab

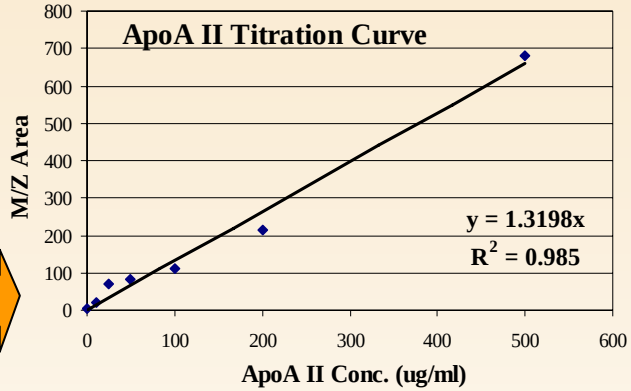
Bait Ab

+ EAM

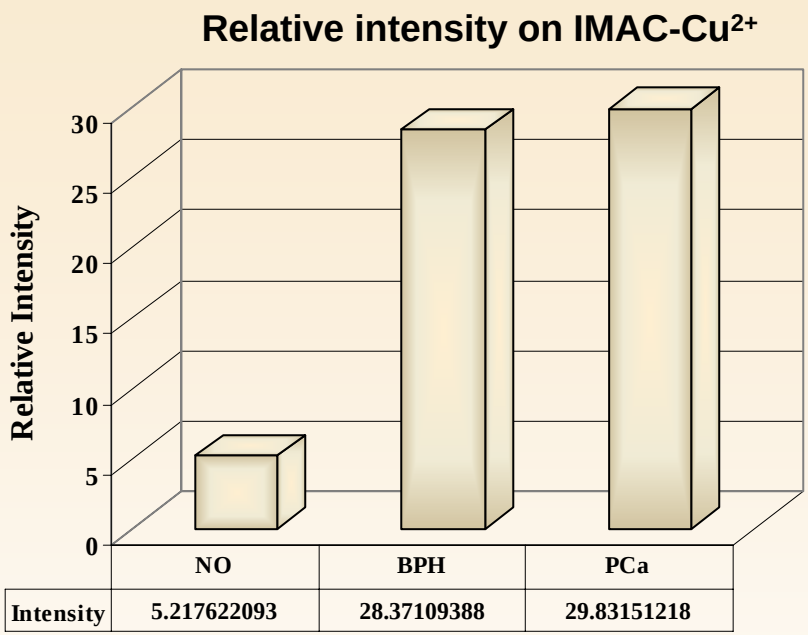
2. Detection

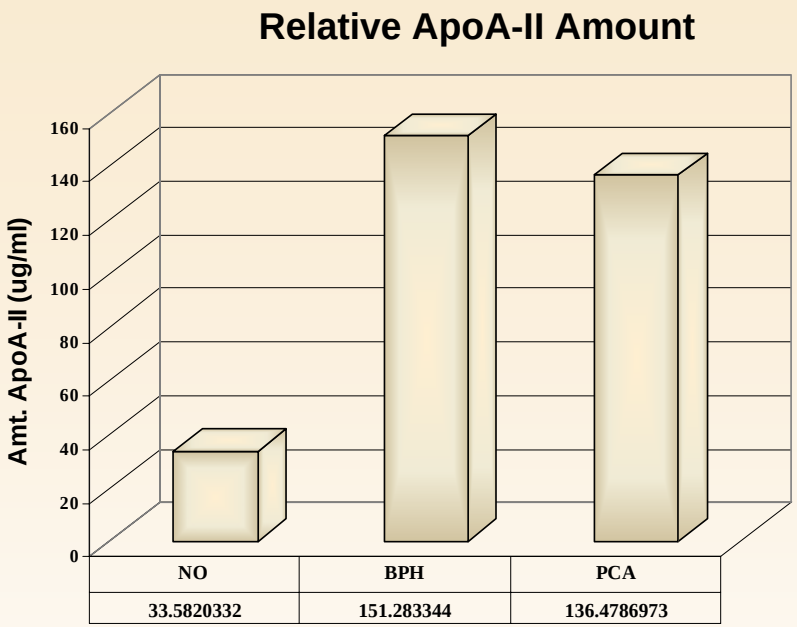


3. Data analysis



The George L. Wright, Jr.
Center for Biomedical Proteomics
Eastern Virginia Medical School

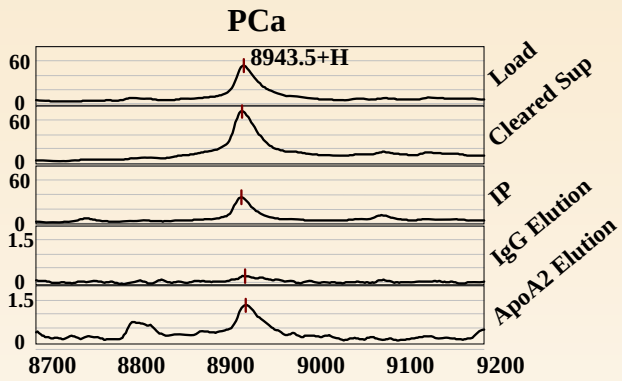




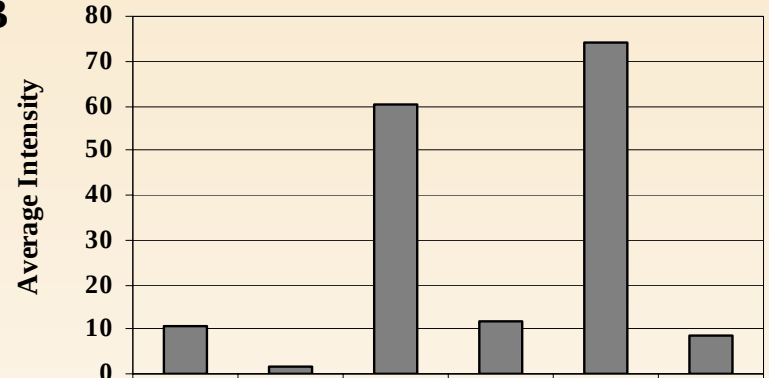
SELDI Immunoassay. A group of 50 NO (healthy), 51 BPH (benign) and 48 PCa (prostate cancer) samples were screened for ApoA-II levels using the SELDI-based immunoassay on mouse mAb Anti-ApoA-II coated PS20 ProteinChips®. The 8.9K *m/z* area observed in the serum samples was used to calculate the serum levels of ApoA-II based on the titration curve. About 4 fold increase in the protein amount in PCa samples is observed as compared to NO. (**D**) The relative normalized intensity observed in the SELDI-TOF-MS 8.9K *m/z* trace on IMAC-Cu2+ ProteinChips®.

ApoA II Immunodepletion

A

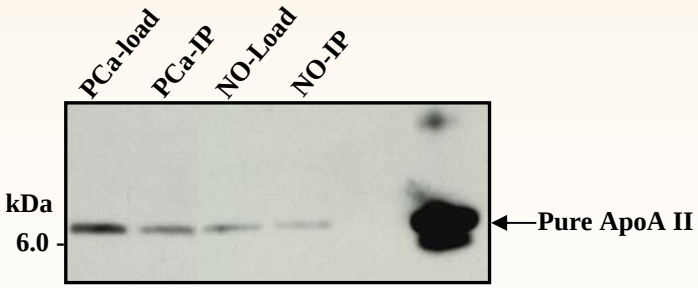


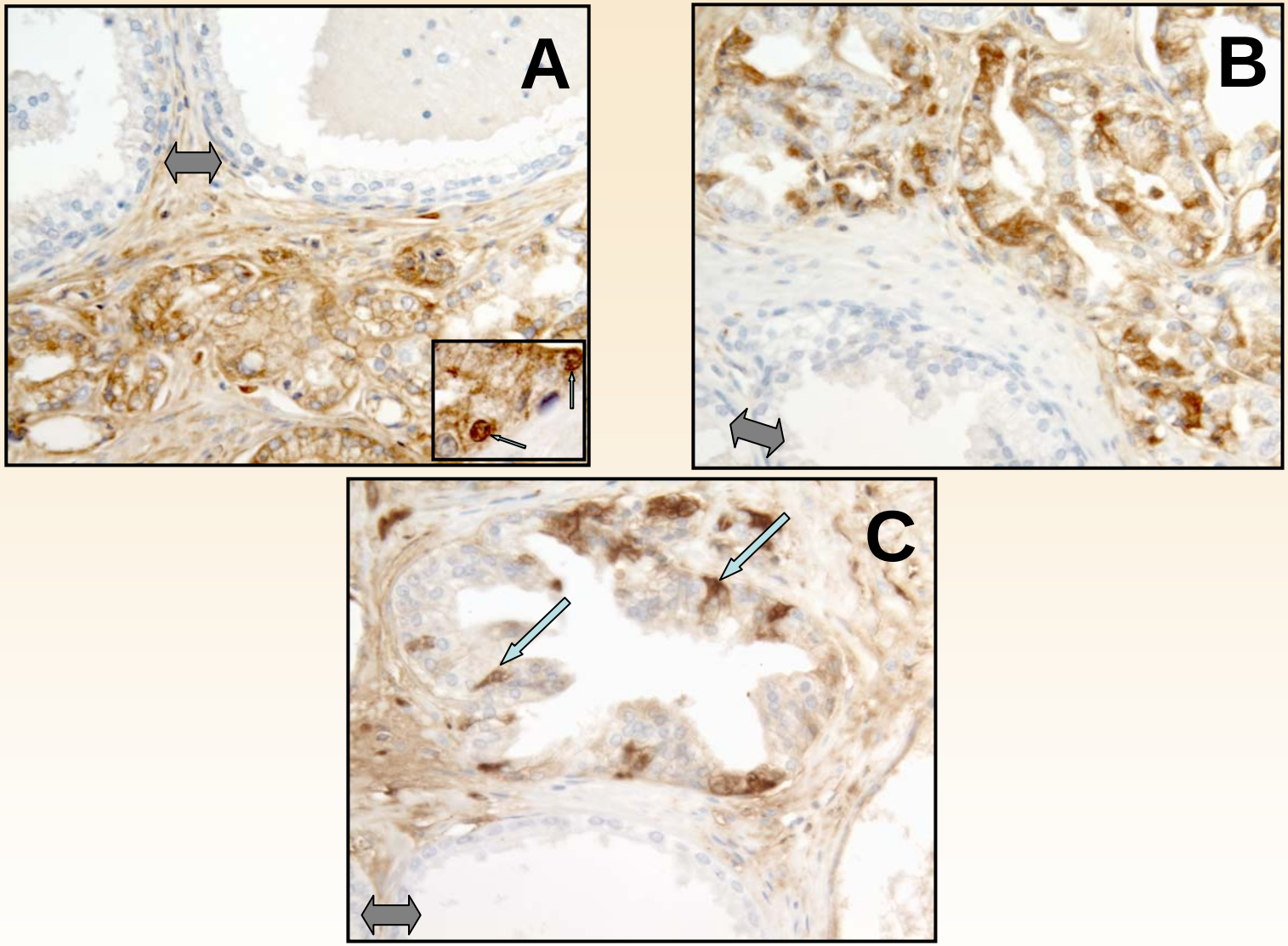
B



	NO-Load	NO-IP	PCa-Load	PCa-IP	BPH-Load	BPH-IP
Avg Intensity	10.887	1.588	60.406	11.921	74.164	8.621

C





ApoA-II Immunohistochemistry. Panel A: Prostate cancer infiltrating under uninvolved prostate glands (thick, double-headed gray arrow) which shows minimal staining even in basal cells. The pattern of staining, which is cytoplasmic, membranous and nuclear in cancer, is accentuated slightly on the advancing edge and is variable (weak to strong). Inset shows prostate cancer demonstrating scattered nuclear (thin, single-headed arrows) staining. **Panel B:** Prostate cancer infiltrating above uninvolved prostate glands (thick, double-headed gray arrow) which shows minimal focal staining. The pattern of staining in the prostate cancer is variable. **Panel C:** Weak to strong staining demonstrated in the luminal cells of PIN and prostate cancer (thin, single-headed arrows) present above uninvolved prostate glands with little staining (thick, double-headed gray arrow). Original magnification X400.

SELDI-Immunoassay for Prediction of Biopsy Outcome

Ability of ApoA-II to discriminate controls from cases with-

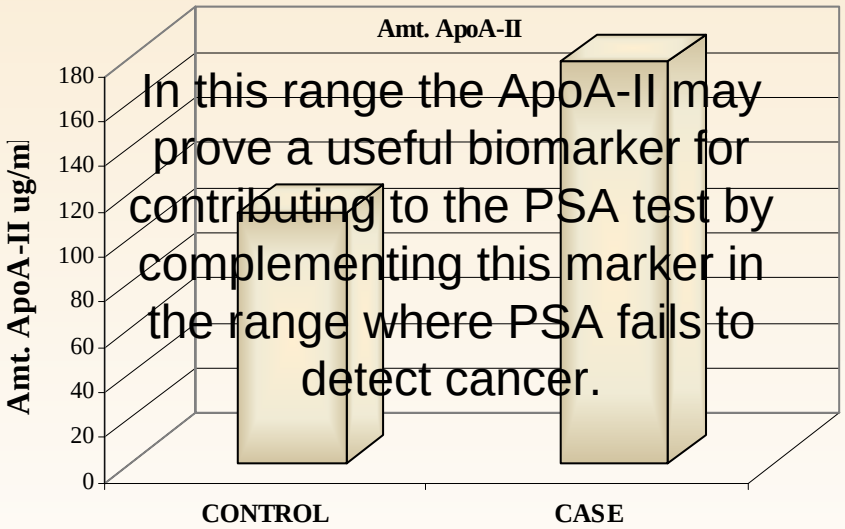
Marginal Clinical Symptoms (DRE+/PSA- or DRE-/PSA+)
PSA<4.0, Biopsy Performed

- 46 Controls in which the biopsy was negative
- 46 Cases in which the biopsy was positive

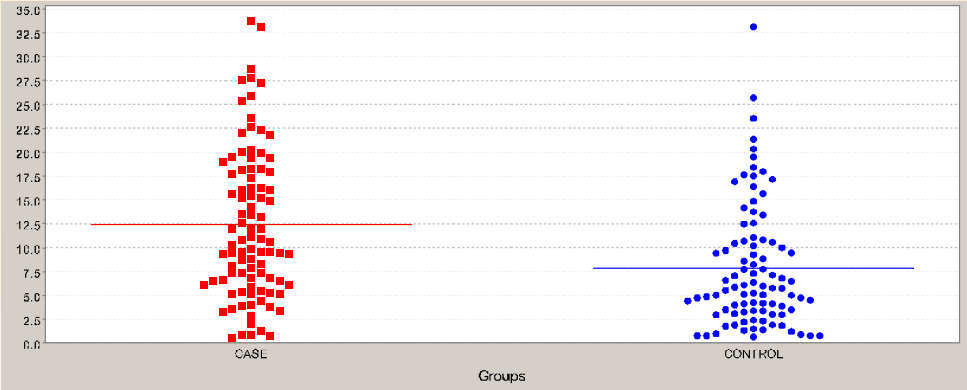
Each sample was randomized in duplicate and analyzed on antibody coated PS20 chips

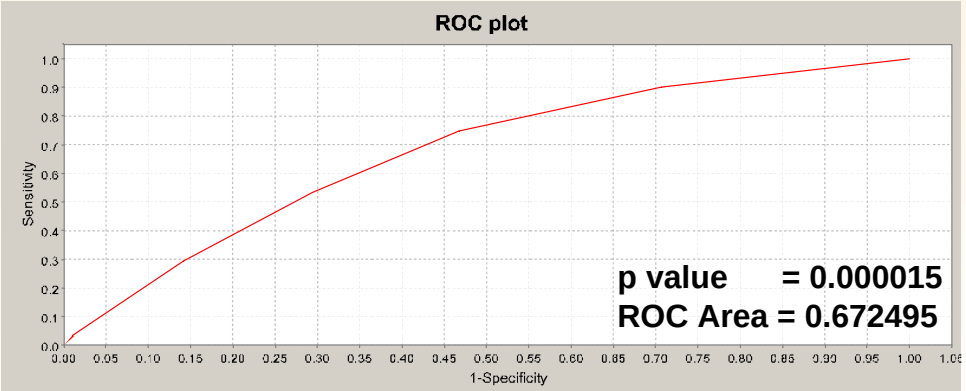
SELDI-Immunoassay for Prediction of Biopsy Outcome

Anti-ApoA-II Immunoassay



In this range the ApoA-II may prove a useful biomarker for contributing to the PSA test by complementing this marker in the range where PSA fails to detect cancer.



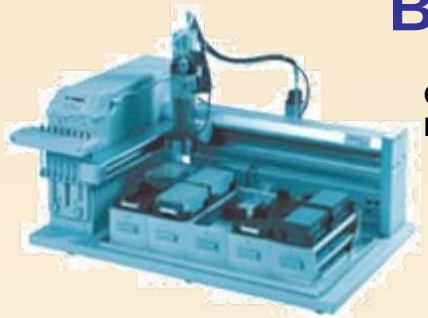


Group	Index	M/Z STD	Intensity Average	Intensity Median	Intensity STD	# of peaks	# estimated
CASE	0	5.231969	12.41333	10.6609	7.82215	92	0
CONTROL	1	6.210689	7.890178	5.8172	6.453902	92	0

WHAT NEXT?

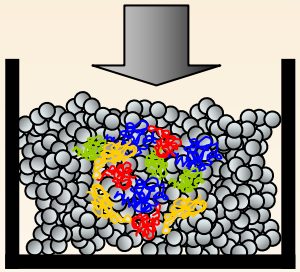
- Facilitation of protein ID
  - LC-MS/MS
  - LIFT-MALDI TOF/TOF
- Fractionation, depletion and concentration steps
  - Higher resolution (MALDI TOF/TOF)
  - Upfront Fractionation
  - Depletion of abundant protein (IgY Fractionation)
- Sensitivity vs. resolution
  - Concentration of depleted serum
  - Improved instrument sensitivity
- Mass spec-assisted Immunoassay development

Bruker ClinProt Technology Process

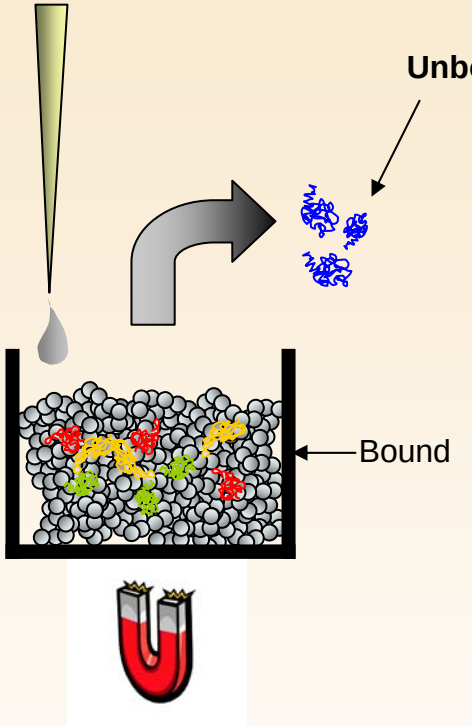


ClinProt Robotics

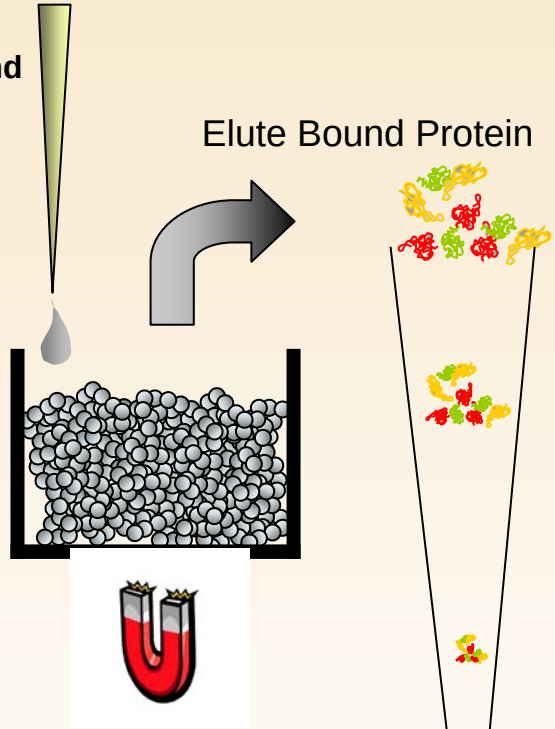
Bind Protein Mixture to Magnetic Beads



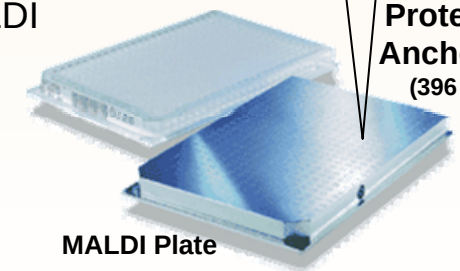
Wash



Elute Bound Protein

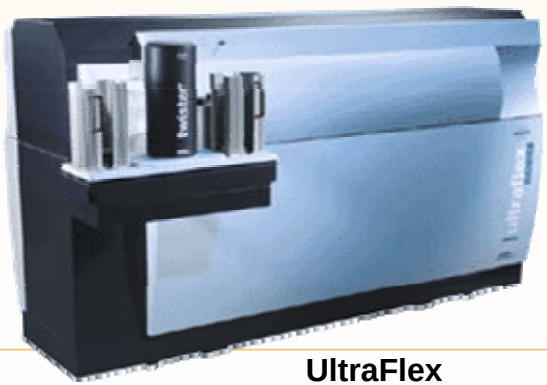


Spot eluted Proteins on AnchorPlate (396 spots)

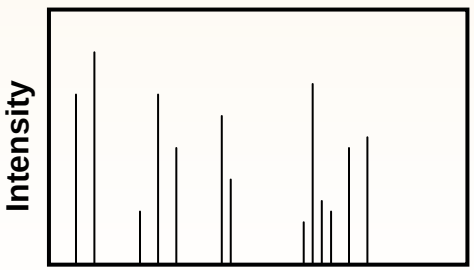


MALDI Plate

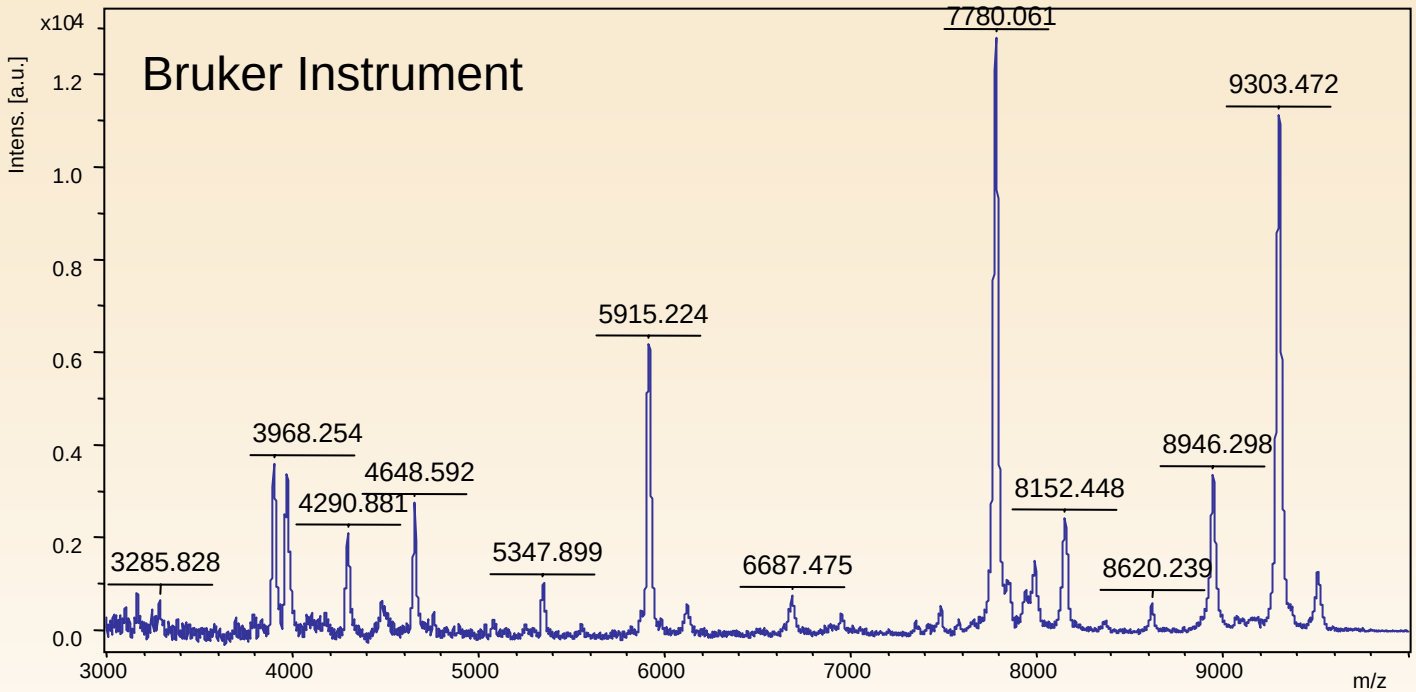
Read in MALDI TOF/TOF

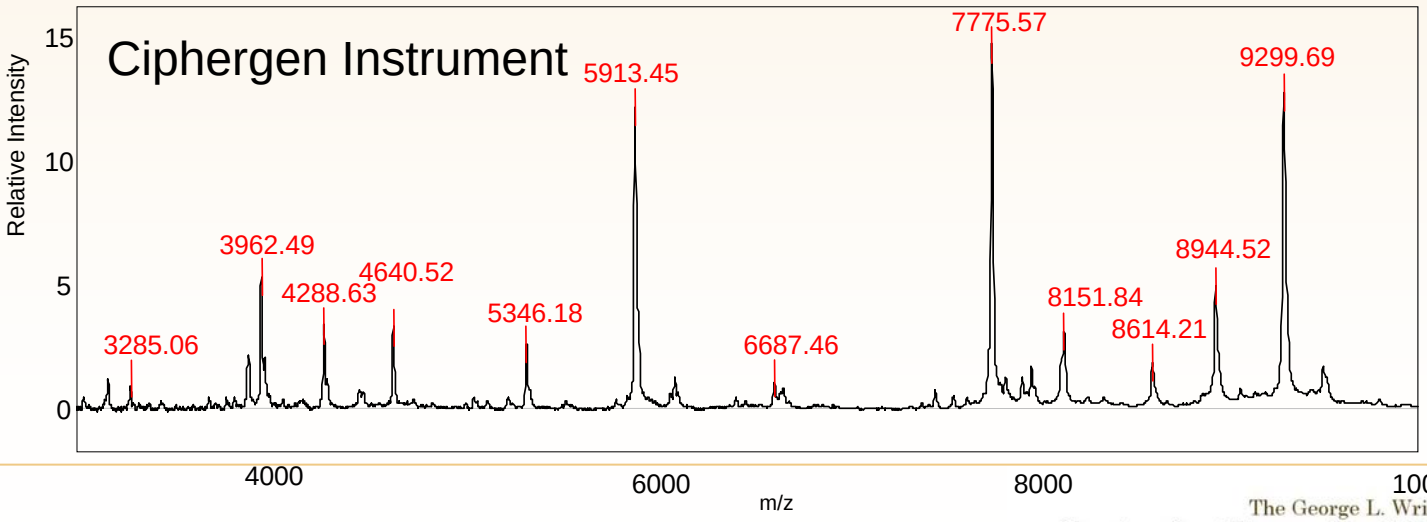


UltraFlex



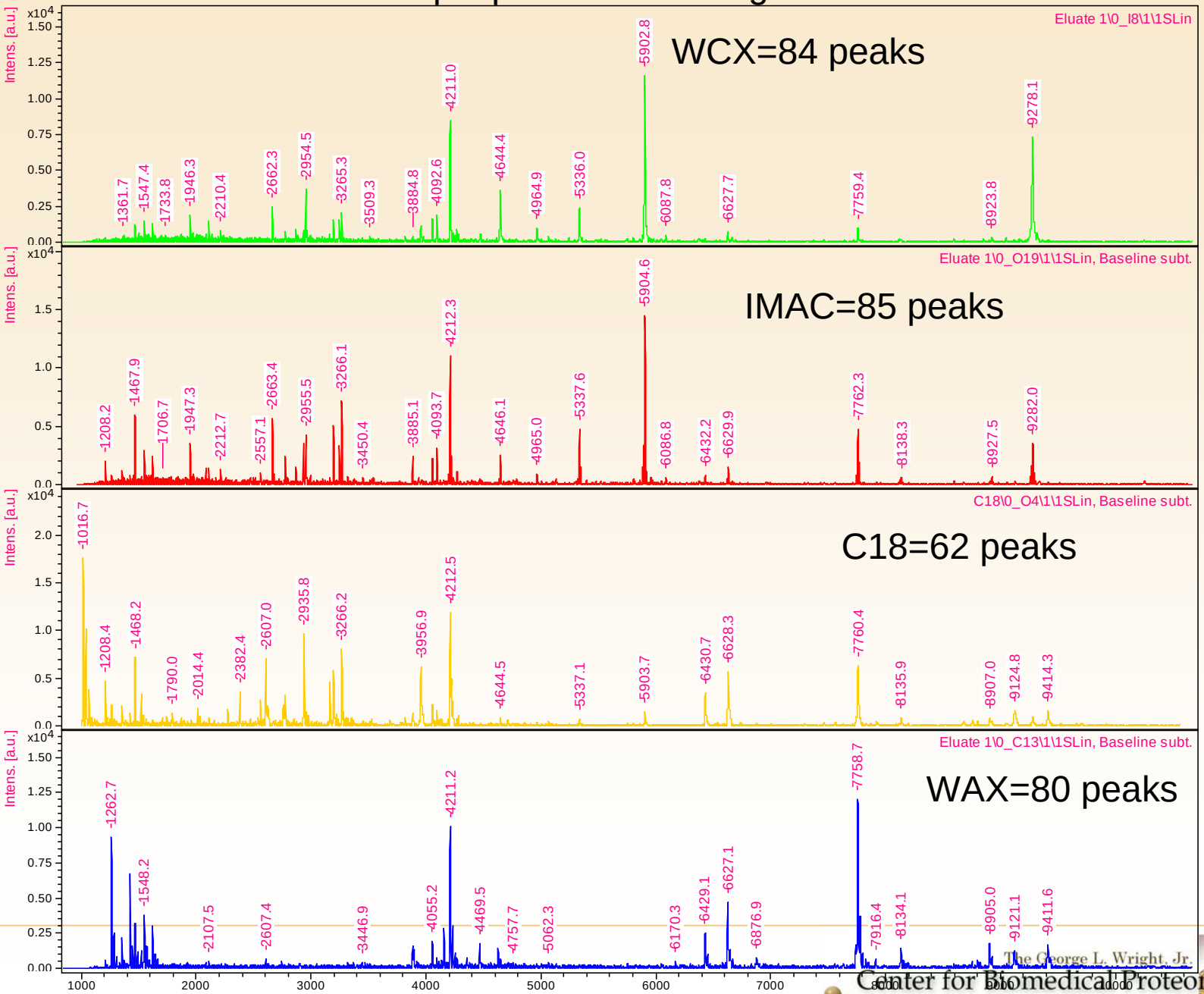
QC IMAC chip spectrum (3-10kDa)



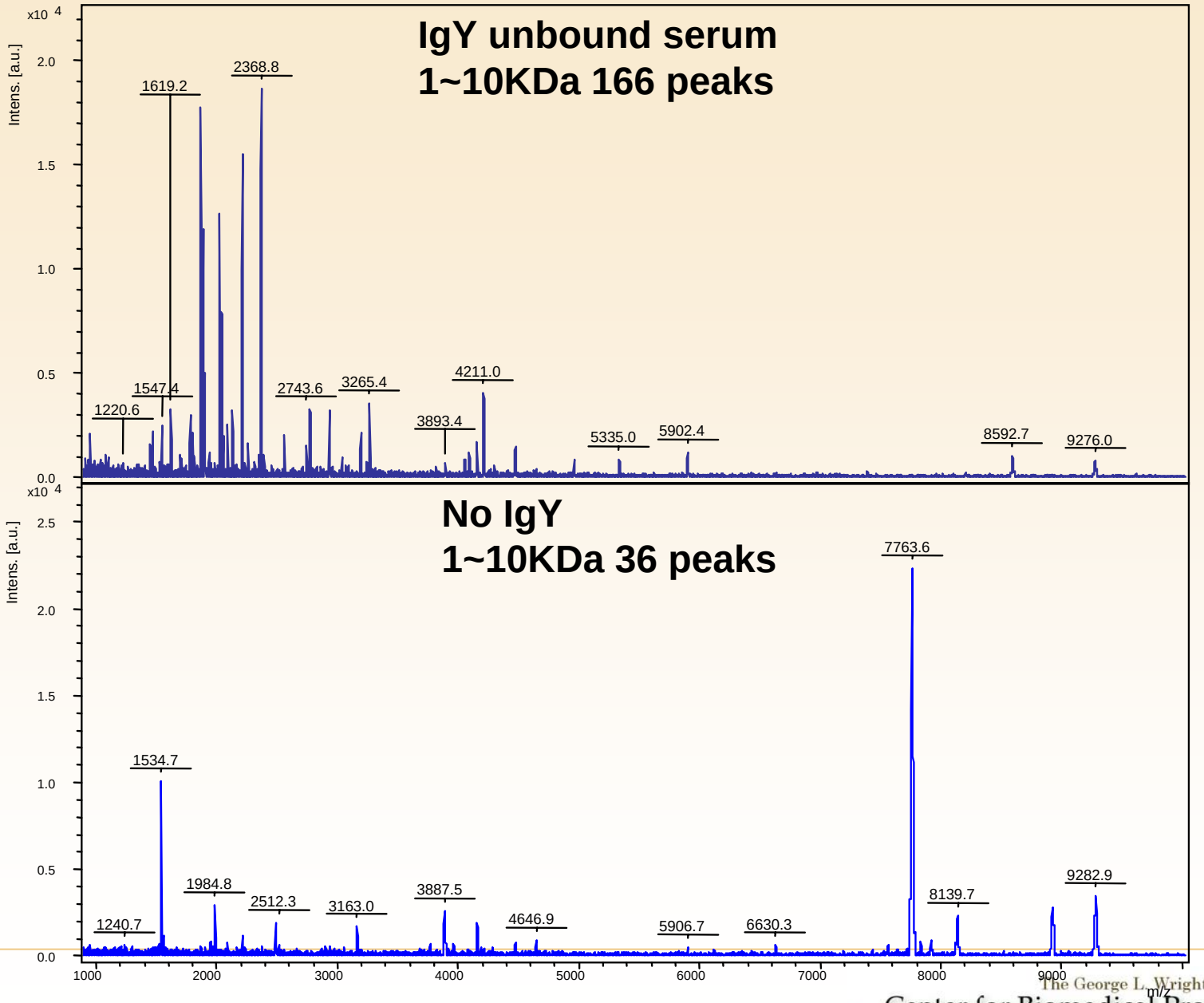


The George L. Wright, Jr.
Center for Biomedical Proteomics
Eastern Virginia Medical School

203 total unique peaks mass range 1000-10000



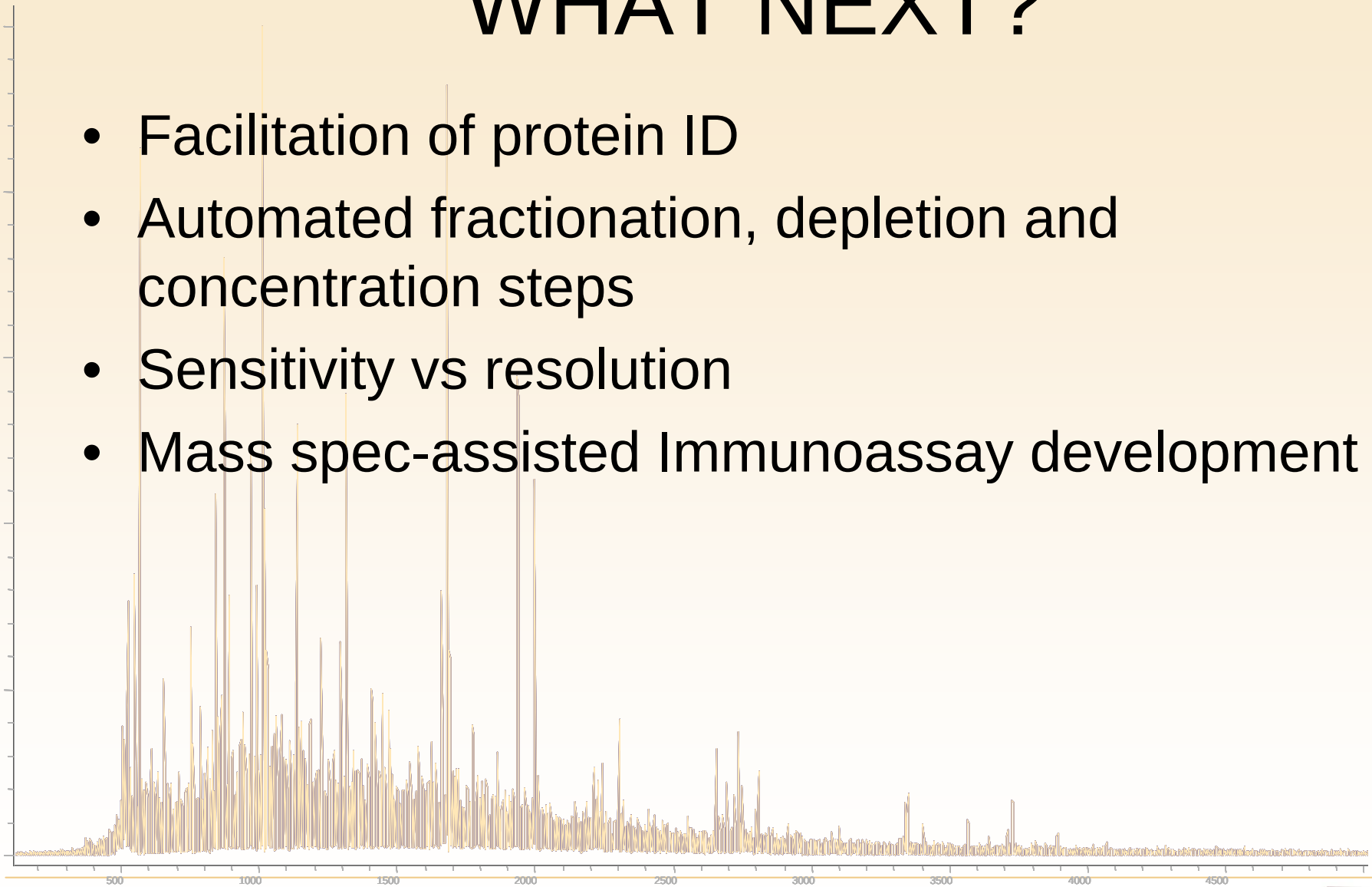
MALDI profiling of serum IgY depletion



The George L. Wright, Jr.
Center for Biomedical Proteomics
Eastern Virginia Medical School

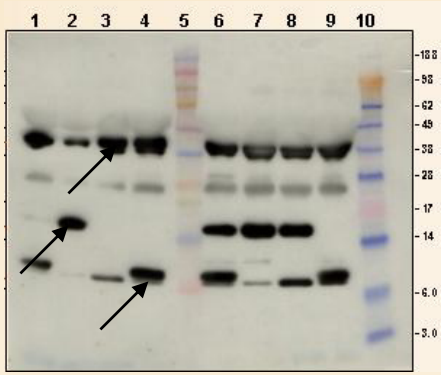
WHAT NEXT?

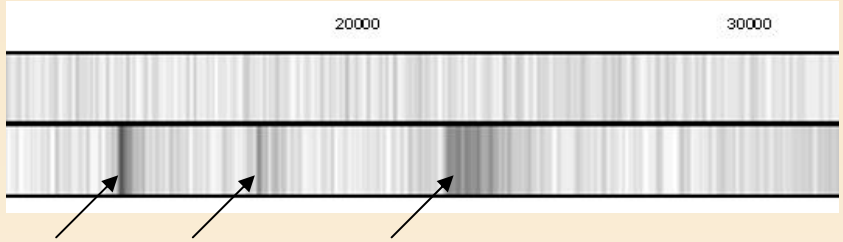
- Facilitation of protein ID
- Automated fractionation, depletion and concentration steps
- Sensitivity vs resolution
- Mass spec-assisted Immunoassay development



The George L. Wright, Jr.
Center for Biomedical Proteomics
Eastern Virginia Medical School

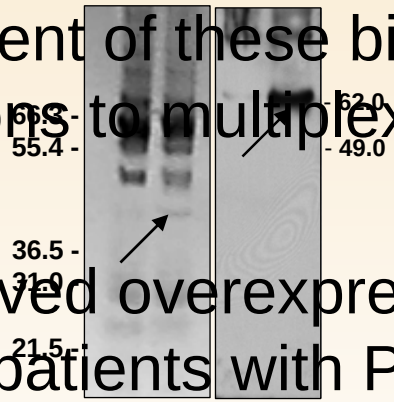
Anti- #147-148

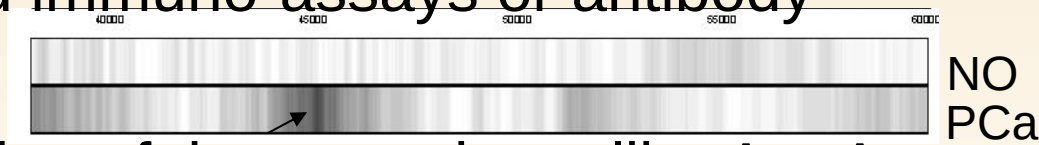




Development of these bioamarkers may provide important contributions to multiplexed immuno-assays or antibody arrays.

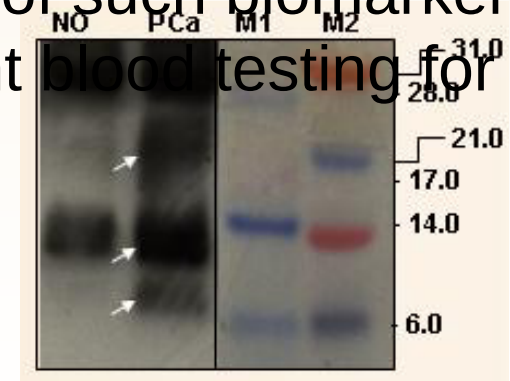
Anti-#101-102

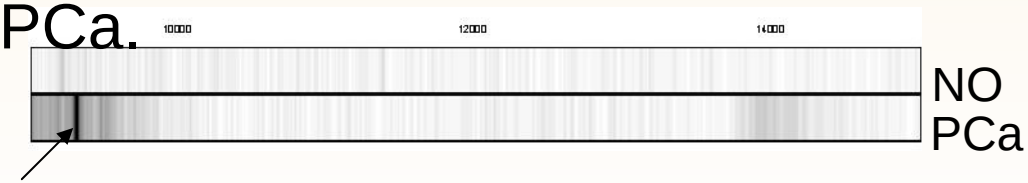




The observed overexpression of these markers, like ApoA-II, in PCa patients with PSA < 4.0 ng, suggests that analysis of such biomarkers in serum may extend the utility of current blood testing for PCa.

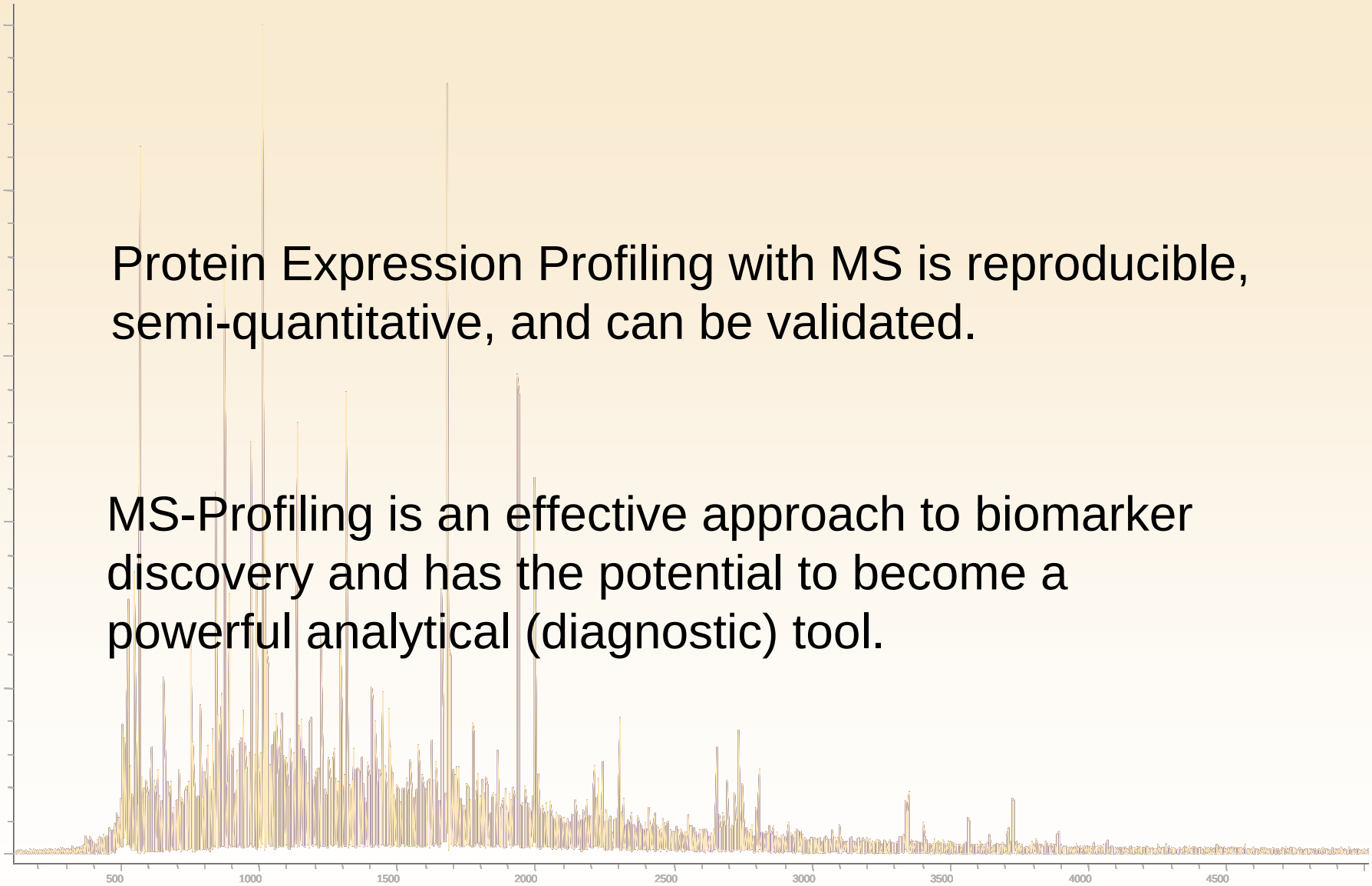
Anti-#107-109





Protein Expression Profiling with MS is reproducible, semi-quantitative, and can be validated.

MS-Profiling is an effective approach to biomarker discovery and has the potential to become a powerful analytical (diagnostic) tool.



Eastern Virginia Medical School Biomarker Discovery Laboratory

Investigators

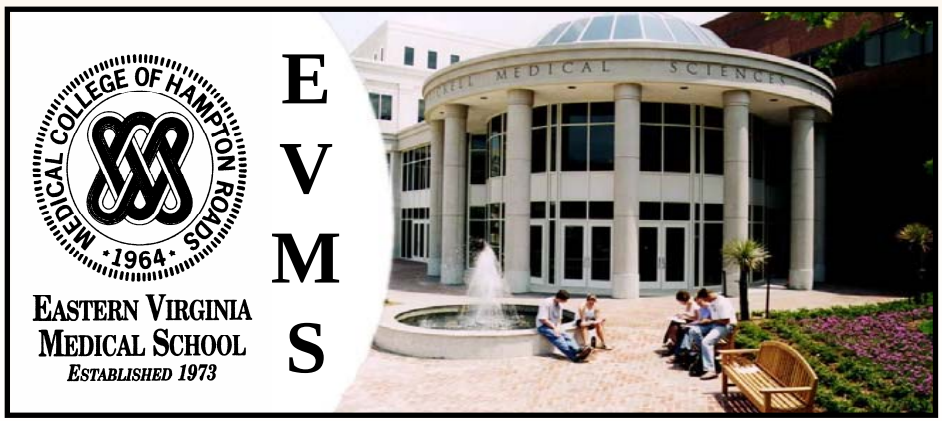
John Semmes, Ph.D.
John Davis, M.D.
Jose Diaz, M.D., Ph.D.
Rick Drake, Ph.D.
Christine Laronga, M.D.
Paul Schellhammer, M.D.
Jeffery T. Wadsworth, M.D.

Fellows

Gunjan Malik, Ph.D.
Alberto Corica, M.D.
Daniel Holterman, Ph.D.
Lining Qi, Ph.D.
Cindy, Ph.D.

Staff

Diane Brassil
Lisa Cazares
MaryAnn Clements
Tarek Kandil
Brian Main
Shamina Mitchell
Michael D. Ward



Biostatistics/Computation

WMRI
William E. Cooke, Ph.D.
Dasha I. Malyarenko, Ph.D.
Denis M. Manos, Ph.D.
Michael W. Trosset, Ph.D.
Eugene R. Tracy, Ph.D.

The George L. Wright, Jr.
Center for Biomedical Proteomics
Eastern Virginia Medical School