

The Diagnostic Potential of the Human Plasma Proteome

Leigh Anderson, Ph.D.
Plasma Proteome Institute



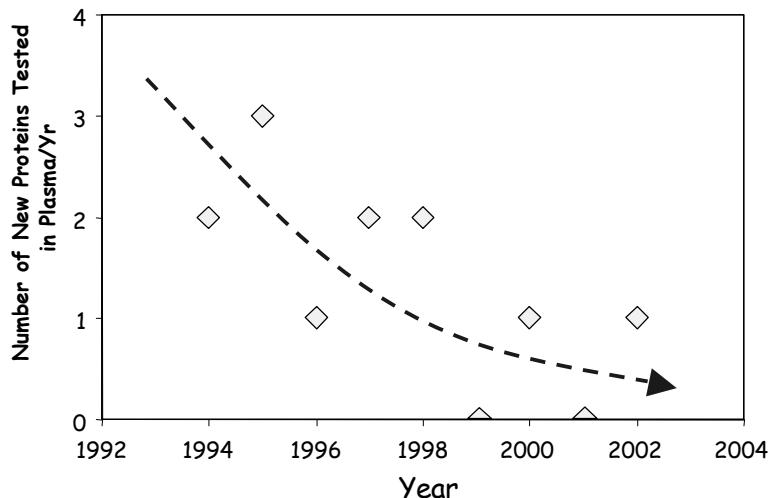
Protein Biomarkers

- Closer to function than DNA or mRNA (mRNA:protein correlation often < 0.5)
- Accessible in blood (plasma) - the primary clinical specimen
- Already widely used in clinical practice (e.g., definitive role of heart muscle troponin in diagnosing acute myocardial infarct)
- Potentially capable of defining all disease states (in a multivariate sense): the strong biomarker hypothesis
- Valuable in drug and device development to confirm efficacy (mechanism), detect toxicity, define disease state, select and stratify patient populations



Introduction of New Clinical Biomarkers is Slowing

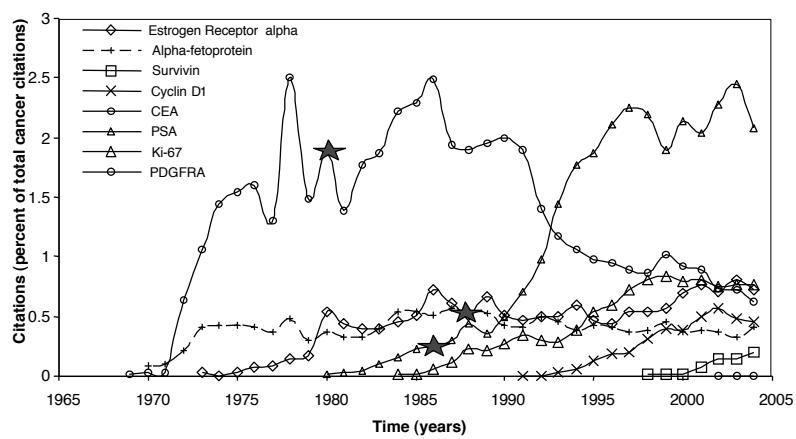
New FDA-Approved (CLIA) Diagnostic Protein Tests in Serum/Plasma Declined for the Last Decade

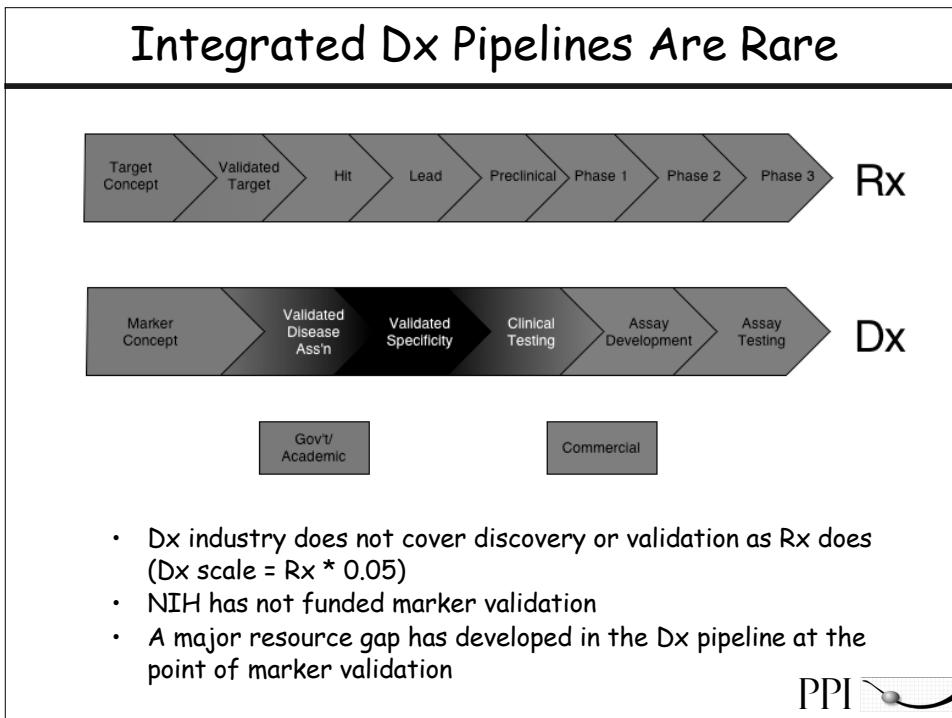


From: The human plasma proteome: history, character, and diagnostic prospects.
Anderson, N. L. Anderson, N. G., Mol Cell Proteomics (2002) 1:845-67.



Timescale for Diagnostic Marker Development Has Been Long: Cancer Markers





"The appealing notion that research advances travel from bench to bedside is laudable, but conceptually flawed. Even though the U.S. Congress fully anticipates that funding to the National Institutes of Health (NIH) will result in advances in clinical medicine and that other forces, presumably non-governmental, will translate the latest in exciting science into health technologies, under the system of healthcare we have today, this advancement is not likely to happen."

Floyd Bloom
President, AAAS
Science 300:1680-1685 (2003)

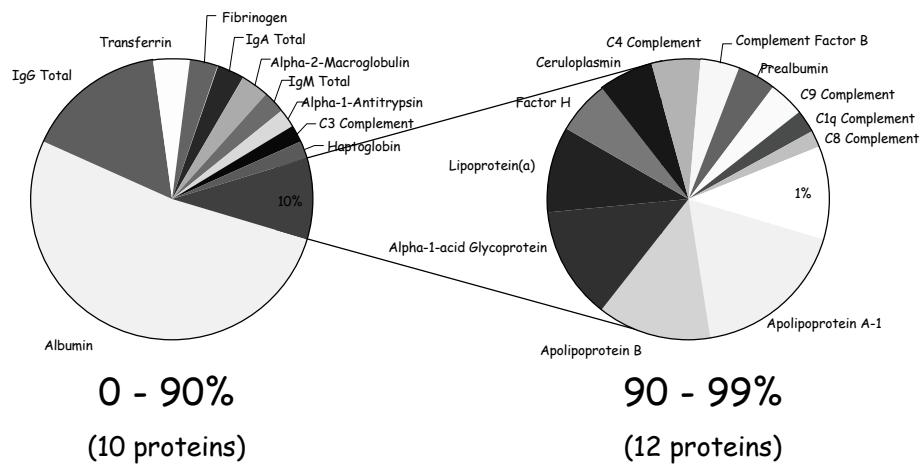


Plasma Biomarker Discovery: Major Challenge for Proteomics

- Plasma is:
 - the largest (most proteins) and deepest (widest dynamic range) representation of the human proteome in one sample
 - in communication with essentially all tissues and physiological processes (Venetian canal effect)
 - the most universally available clinical sample
 - the source of >100,000 kg of pure therapeutic protein/yr
- Thus for proteomics plasma is:
 - the most difficult sample
 - the most important sample

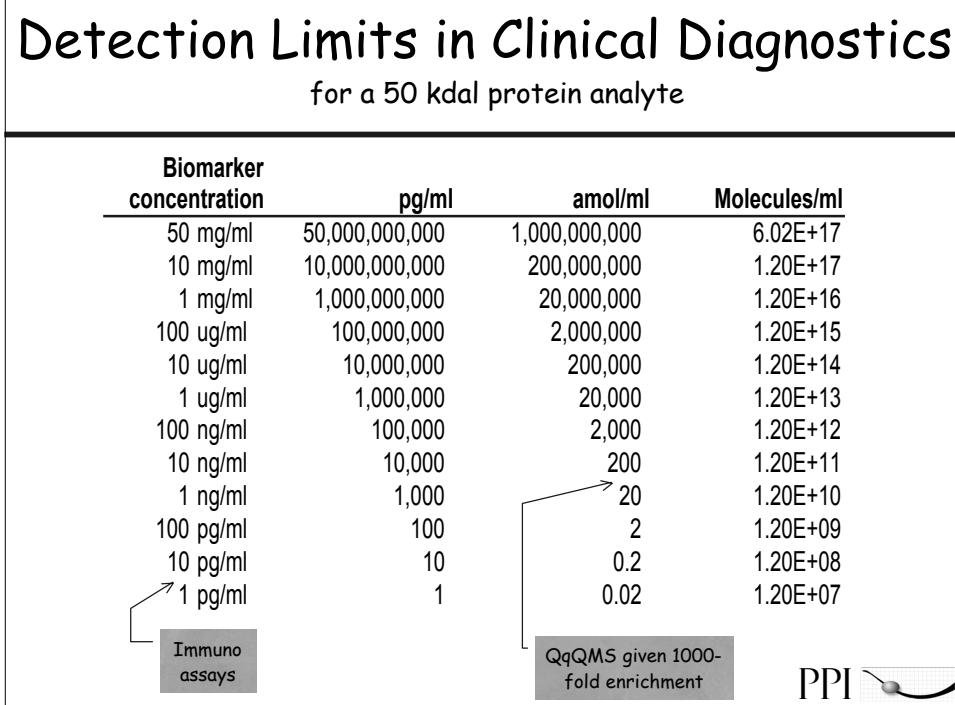
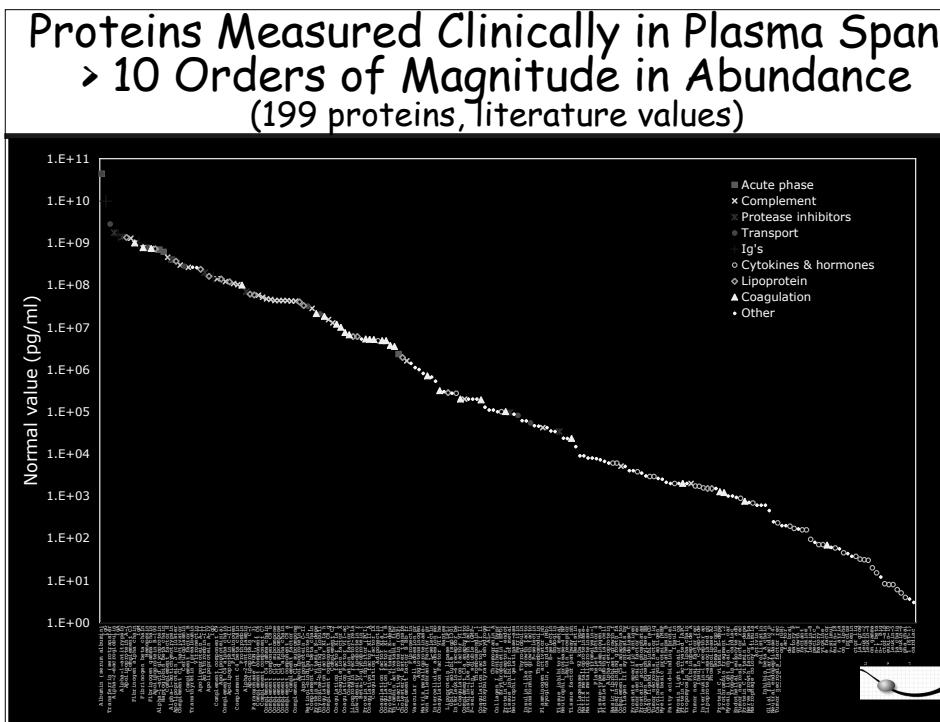
PPI

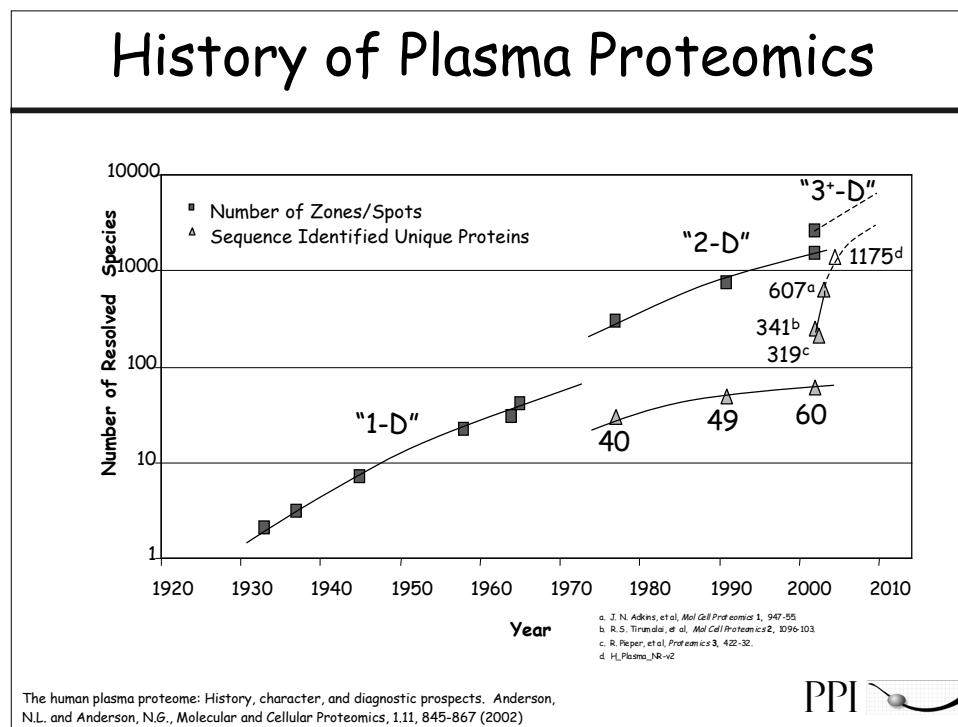
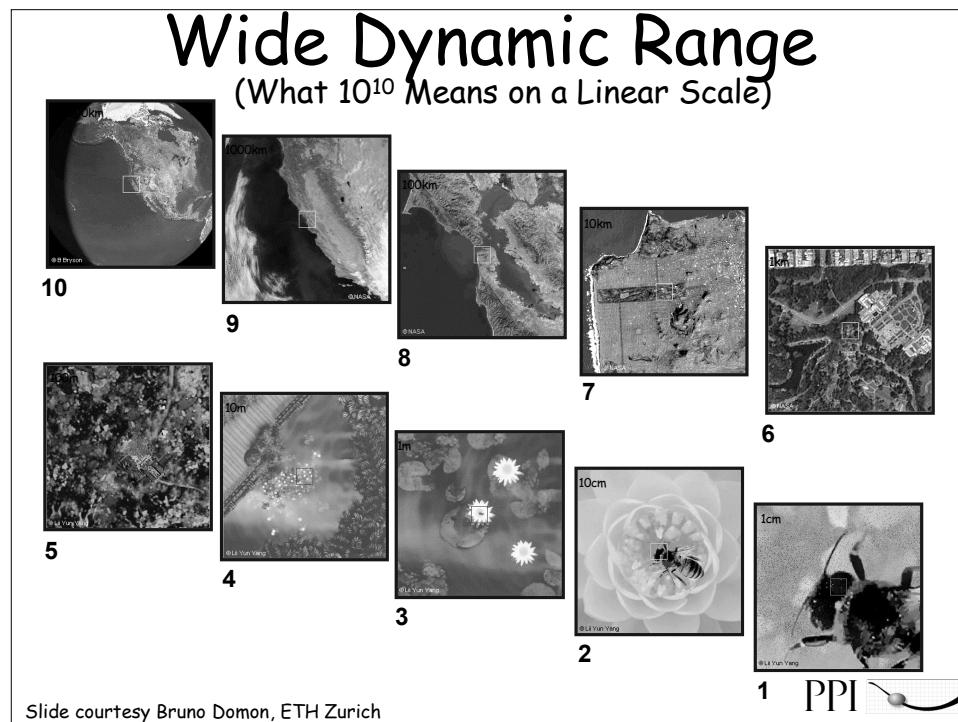
A Small Number of Proteins Make Up the Top 99% of Plasma by Mass

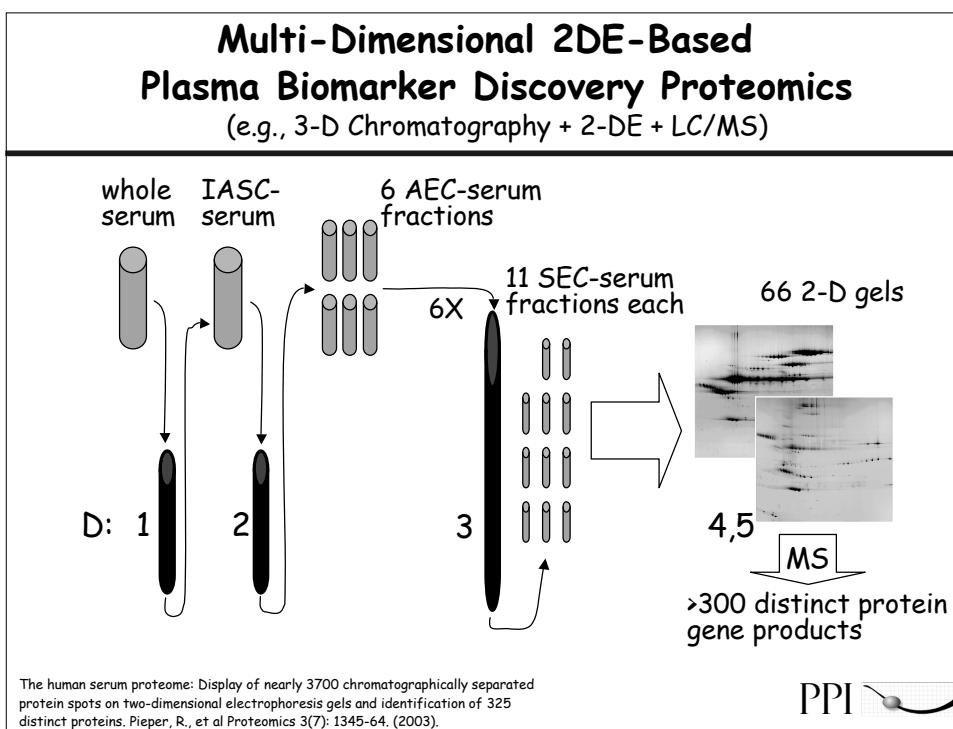
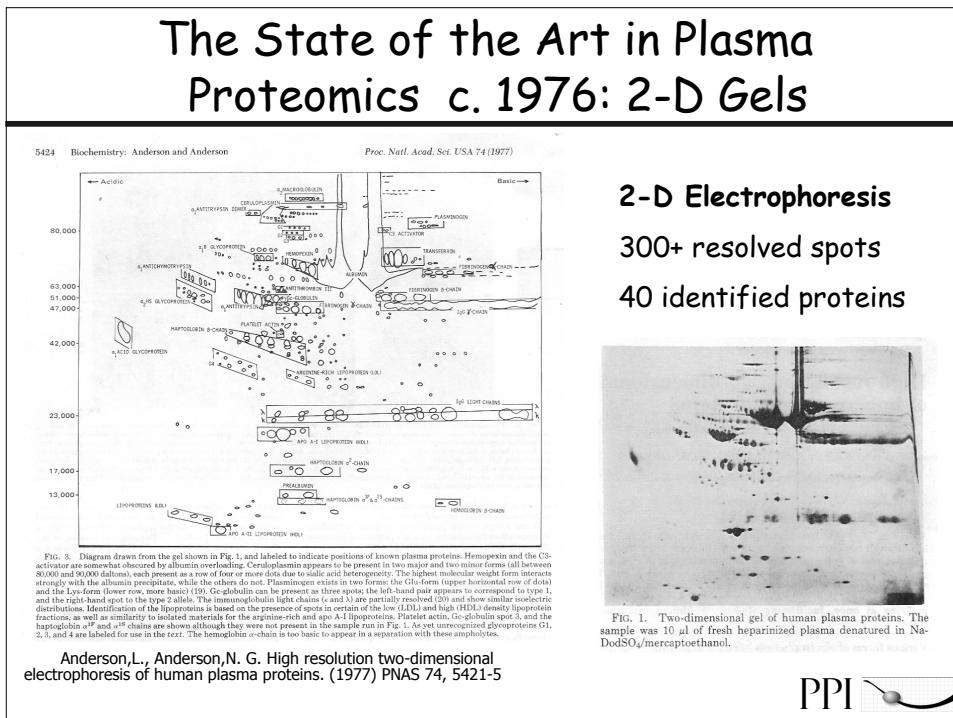


The human plasma proteome: History, character, and diagnostic prospects. Anderson, N.L. and Anderson, N.G., Molecular and Cellular Proteomics, 1:11, 845-867 (2002)

PPI







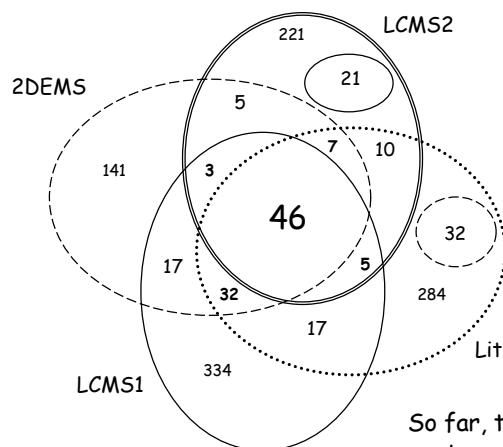
Different Technologies Provide Different Views of the Human Plasma Proteome

- Four datasets compared:
 - Base list of ~450 proteins reported in "non-proteomics" literature as measured/detected in plasma or serum
 - Three sets of 300-600 proteins each from proteomics surveys (2-D gels + MS/MS; LC/LC-MS/MS)
- Made non-redundant using methods of genomics
 - Redundancy definition: >95% homology over ≥ 15 amino acid subsequence
- Questions:
 - What is the overlap between different methods?
 - Can low-abundance proteins be detected in surveys?
 - Can a useful plasma database be collected?

From: The Human Plasma Proteome: A Non-Redundant List Developed by Combination of Four Separate Sources, N. L. Anderson et al, Molec. Cell Proteomics, 3: 311-326 (2004).



Overlap of Four Plasma Proteome Datasets (Number of NR proteins)

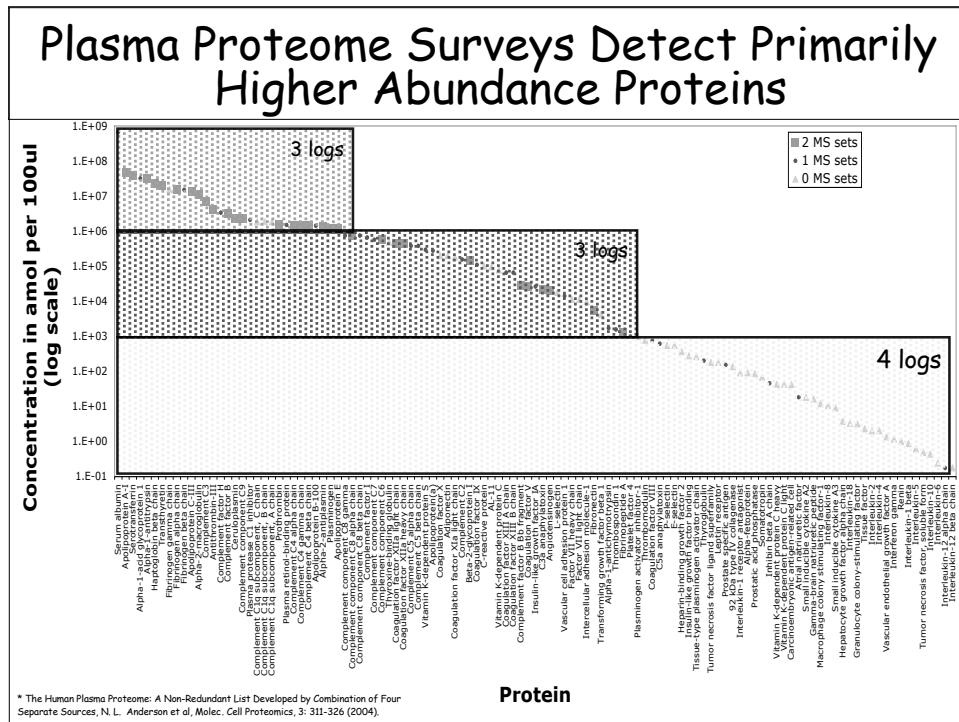


- 46 proteins in all four lists
- 195 proteins in 2 or more lists
- 1175 NR proteins total

So far, there is no comprehensive exploratory proteomics platform for plasma

From: The Human Plasma Proteome: A Non-Redundant List Developed by Combination of Four Separate Sources, N. L. Anderson et al, Molec. Cell Proteomics, 3: 311-326 (2004).

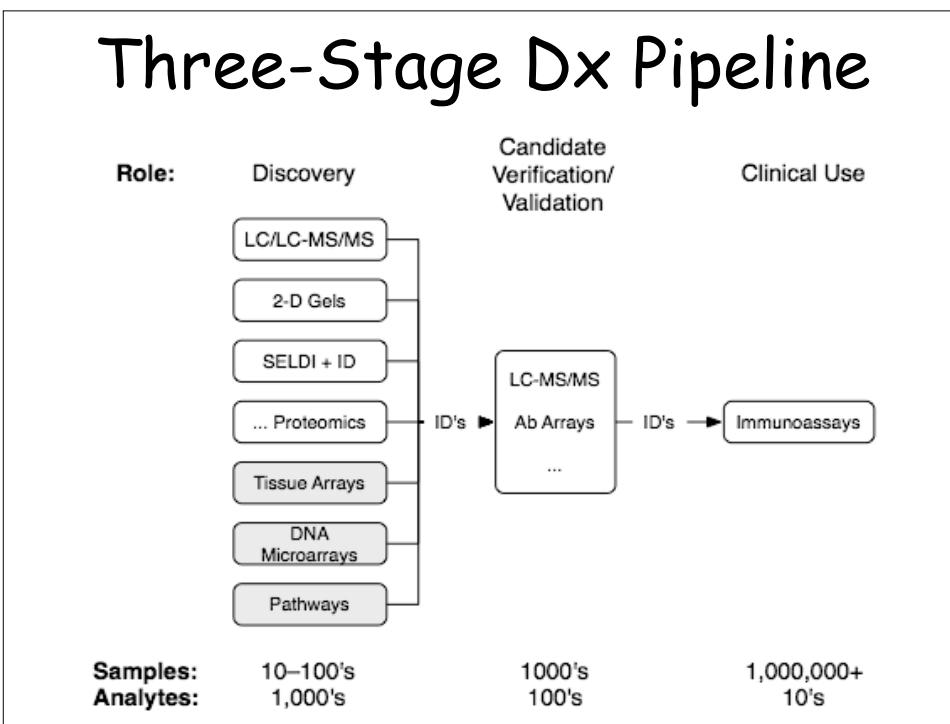
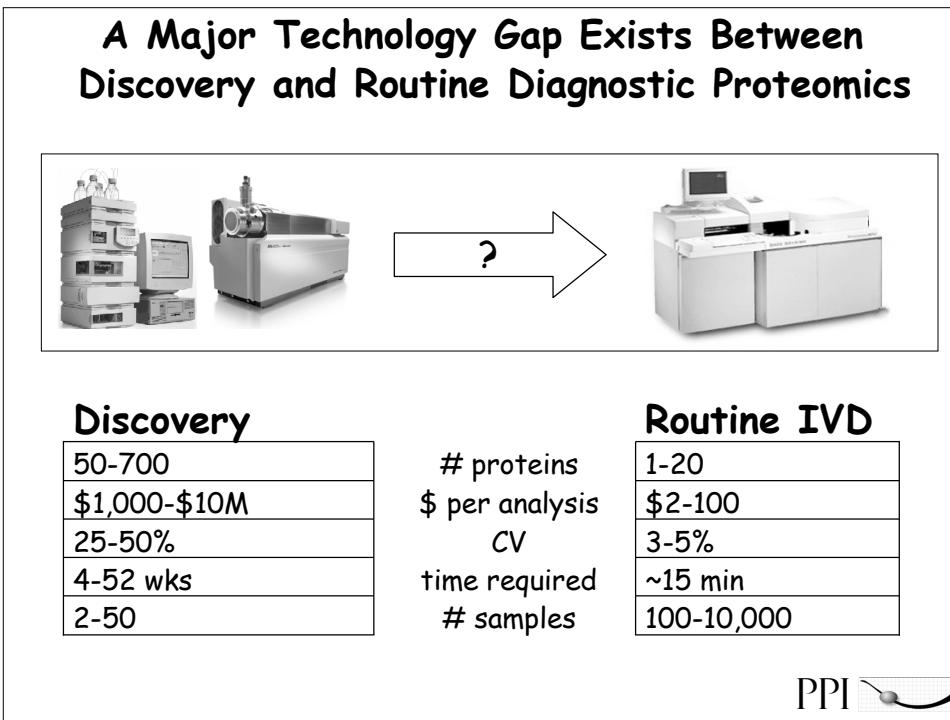




Strategic Directions for Plasma Biomarker Discovery

- Pursue biomarker discovery at the site of biology (typically within tissue) where concentration and differential are highest
- Extend fractionation technologies for deeper plasma exploration, especially for PTM's
- Use specific assays to investigate candidates from proteomics and other approaches in plasma (directed approach: validation platform)





A Candidate-Based Approach: Specific Assays for Identified Candidates

- Specific assays (e.g., immunoassays) can have much greater sensitivity than survey methods
- Rich sets of biomarker candidates are emerging:
 - 177 cardiovascular
 - Candidate-Based Proteomics in the Search for Biomarkers of Cardiovascular Disease, Leigh Anderson, *J. Physiol.*, in press 2004
 - ~1200 cancer
 - Manuscript submitted
- The challenge is to construct a specific assay for any sequence-defined protein candidate quickly, at low cost



177 Candidate Cardiovascular Disease Marker Proteins

	Name	Normal Concentration (pg/ml)	Accession#	Source for concentration	Reason
1	activin A	6.0E+02	P08476	(Eldar-Geva et al., 2001)	Released by heparin from vascular endothelium (Phillips et al., 2000)
2	adiponectin (ADPN)	4.8E+06	Q15848	(Mallamaci et al., 2002)	Higher levels in essential hypertensives (Mallamaci et al., 2002)
3	albumin	4.1E+10	P02768	(Laboratories, 2001)	Negative acute phase reactant, lower levels associated with increased risk of cardiovascular mortality (Shaper et al., 2004)
4	aldolase C	4.0E+03	P09972	(Asaka et al., 1990)	A more specific and sensitive marker of cerebrovascular disease than lactate dehydrogenase A (Asaka et al., 1990)
5	alpha 2 antiplasmin (alpha 2 API)	7.0E+07	P08697	Progen test insert	An important regulator of the fibrinolytic system
6	alpha 2 macroglobulin (alpha 2 M)	1.8E+09	P01023	(Laboratories, 2001)	Major plasma protease inhibitor
7	alpha(1)-antichymotrypsin (ACT)	4.2E+07	P01011	(Putnam, 1975)	Major plasma protease inhibitor
8	alpha1 acid-glycoprotein (AAG)	6.9E+08	P02763	(Laboratories, 2001)	Acute phase reactant
9	alpha1-antitrypsin (AAT)	1.4E+09	P01009	(Laboratories, 2001)	Major plasma protease inhibitor
10	angiotensin-converting enzyme (ACE)		P12821		Lower in stroke patients than controls (Catto et al., 1996)
11	angiotensinogen	1.5E+06	P01019	(Bloom et al., 1988)	Precursor of major blood pressure control peptide
12	anithrombin III (AT III)	2.0E+08	P01008	(Kolafatis et al., 1997)	Major inhibitor of thrombin
13	apolipoprotein A-I	1.4E+09	P02647	(Glowinska et al., 2003)	Low level associated with mortality and myocardial infarction five years after CABG(Skinner et al., 1999)
14	apolipoprotein A-II	2.4E+08	P02652	(Luo & Liu, 1994)	Lipoprotein
15	apolipoprotein A-IV	1.6E+08	P06727	(Kondo et al., 1989)	A relatively independent risk factor for CHD(Warner et al., 2001)
16	apolipoprotein B	7.3E+08	P04114	(Glowinska et al., 2003)	Major component of LDL
17	apolipoprotein C-I	6.1E+07	P02654	(Riesen & Sturzenecker, 1996)	Lipoprotein
18	apolipoprotein C-II	3.3E+07	P02655	(Buray et al., 1986)	Lipoprotein
19	apolipoprotein C-III	1.2E+08	P02656	(Onat et al., 2003)	marker of CHD independent of cholesterol(Onat et al., 2003)
20	apolipoprotein D		P05090		Lipoprotein
21	apolipoprotein E	4.0E+07	P02649		presence of epsilon4 allele a strong independent predictor of adverse events (Bacic et al., 2000)
22	apolipoprotein L1		O14791		Lipoprotein
23	aspartate aminotransferase, mitochondrial (m-type)		P00505		diagnostic for early detection of myocardial infarction (Yoneda et al., 1992)
24	basic fibroblast growth factor (bFGF)		P09038	6.0E+03	(Song et al., 2002)
25	beta(2)-glycoprotein I, nicked		P02749		sICAM-1level increases in acute cerebral infarction (Song et al., 2002)
26	B-type neurotrophic growth factor (BNNGF)	7.0E+02	P01138	(Reynolds et al., 2003)	may control extrinsic fibrinolysis via a negative feedback pathway loop (Yasuda et al., 2004)
					Candidate stroke marker (Reynolds et al., 2003)

From: Candidate-Based Proteomics in the Search for Biomarkers of Cardiovascular Disease, Leigh Anderson, *J. Physiol.*, in press 2004



Technology Alternatives for Targeted Proteomics

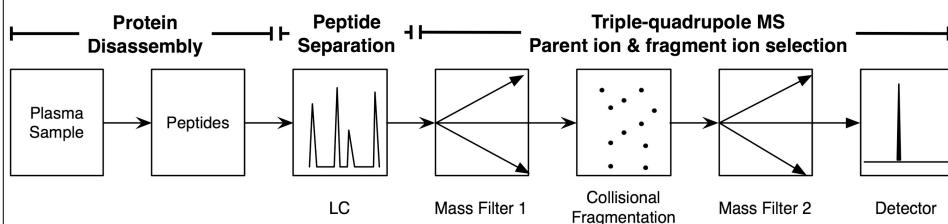
- Immunoassays (likely clinical test implementation)
 - Very sensitive
 - Expensive: IVD-quality assays cost \$2-4 million
 - Specificity issues with less well-developed assays
 - Multiplexing limits in a single assay volume
- Hybrid MS-based assays
 - Peptide MS for quantitation and identification
 - Specific enrichment for sensitivity
 - Absolute analyte specificity
 - Multiplex 25-200 assays/analysis

PPI

A General Approach to Protein Measurement via MS/MS

Objective: direct design of protein assays from sequence, using peptides as measurement surrogates:

- Variation of protein physical properties make them hard to address comprehensively
- Postulate: within every protein (good,bad) there is ≥ 1 good tryptic peptide
- Absolute quantitation using isotope dilution (stable isotope labeled peptide internal standards)

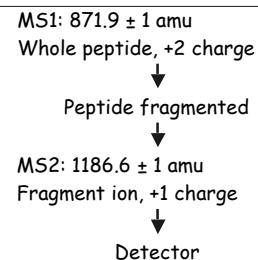


- The triple-quadrupole MS (LC-MS/MS) platform is very widely used for small molecular assays in plasma (drug metabolites, inborn errors, pesticides)
- Multiplex 100-200 assays per run

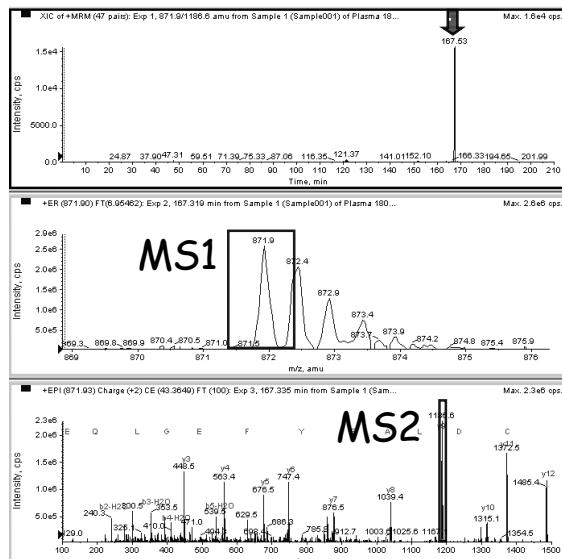
PPI

Alpha-1-acid glycoprotein 1 peptide EQLGEFYEAALDBLR 1742.8

Detected ions passing MS1 and MS2 (871.9/1186.6) over course of 180 minute LC peptide separation: single sequence-specific peak



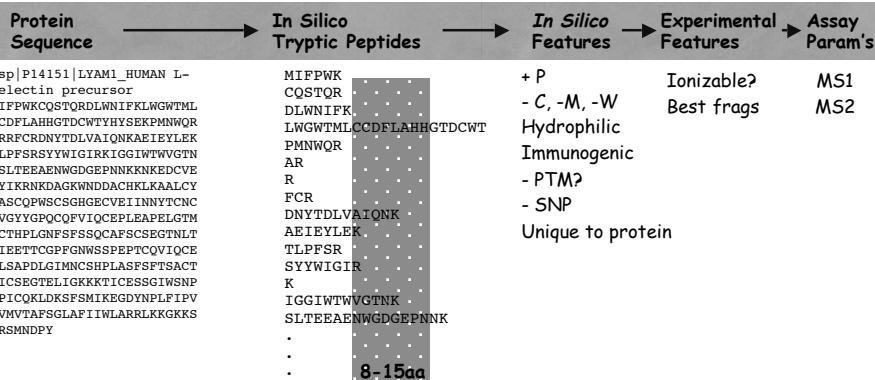
This defines an MRM assay, referring to "multiple reaction monitoring"



Collaboration with Christie Hunter (ABI) using 4000 Q TRAP MS



In Silico MS Assay Design via Bioinformatics



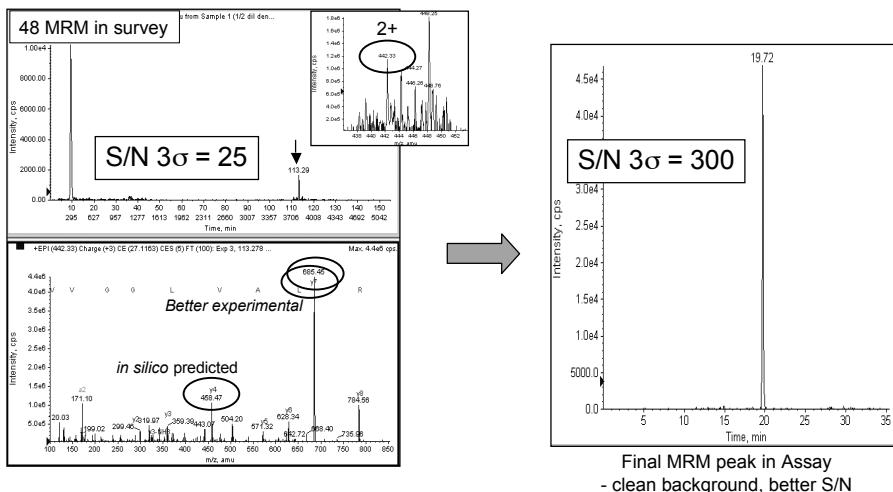
- Pure in silico yields a low proportion of good peptides
- Addition of experimental data yields >90% success

Improvements needed in prediction of ionization and fragmentation

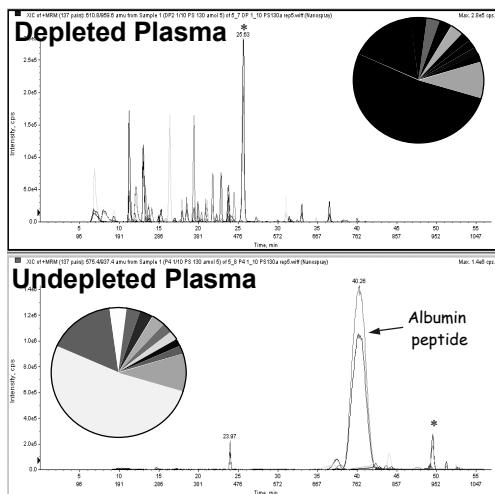


MIDAS™ Workflow for Better MRM Design

Coagulation Factor XIIa light chain - 456 fmol/uL plasma - 4 fmol on column



Depletion of Major Proteins Improves S/N for Most MRM's

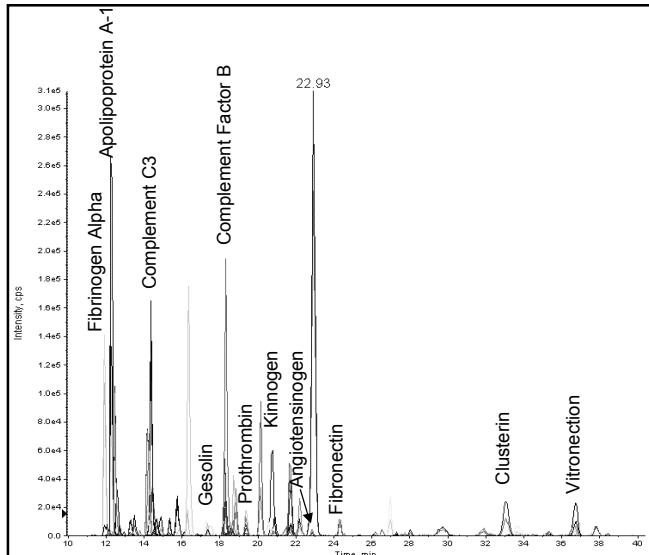


- Equivalent amounts of plasma on column (0.01 uL)
- Depletion of six of the most abundant proteins in plasma (albumin, IgG, IgA, transferrin, haptoglobin, and antitrypsin) improves chromatography and detection of lower abundance proteins.
- S/N improves ~5-10x for signals of lower abundance proteins in depleted plasma

Collaboration with Christie Hunter (ABI) using 4000 Q TRAP MS

PPI

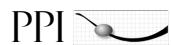
Peptide MRMs to Protein Biomarkers



Method monitors 137 MRM transitions

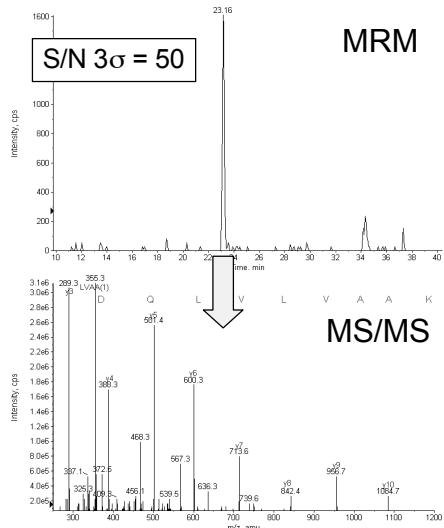
Data generated in depleted/digested plasma, 0.01 μ L loaded directly on nanocolumn

Signal / Noise good even in very complex mixture in a short (30 mins) nanoLC run



Absolute Structural Confirmation of Peptide ID

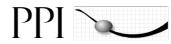
Full MS/MS scan triggered by MRM peak



MIDAS™ workflow to increase confidence in peptide MRM, confirm retention time

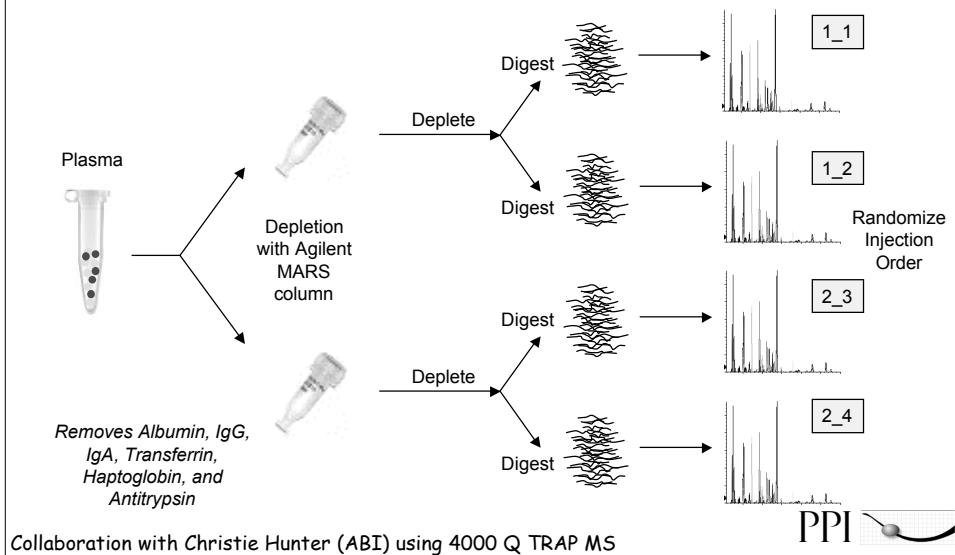
Survey experiment consists of all MRM transitions, detection of signal drives acquisition of MS/MS to confirm ID

ALQDQLVLVAAK of Angiotensin identified Present at 20116 amol/uL in plasma
200 amol on column as 0.01 μ L plasma equivalent loaded

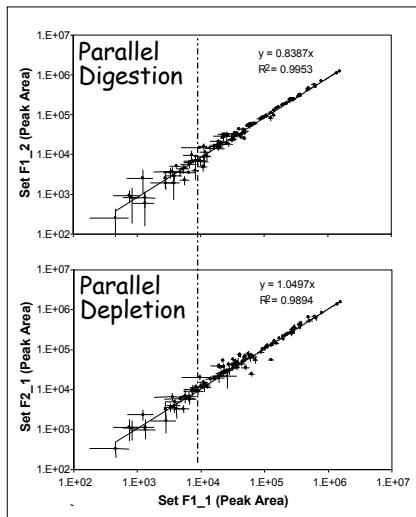


Collaboration with Christie Hunter (ABI) using 4000 Q TRAP MS

Evaluating the Reproducibility of Depleted and Digestion



Correlation Between Parallel Depletion and Digestion



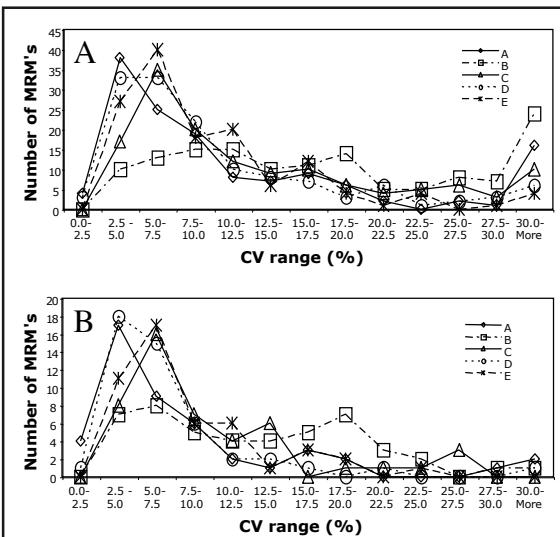
Excellent correlation between parallel digests, $R^2=0.995$ and 0.998

Parallel depletions (which include the effects of parallel digestions) also show good correlation, $R^2=0.989$ and 0.991

Collaboration with Christie Hunter (ABI) using 4000 Q TRAP MS



CV's of MRM Protein Assays



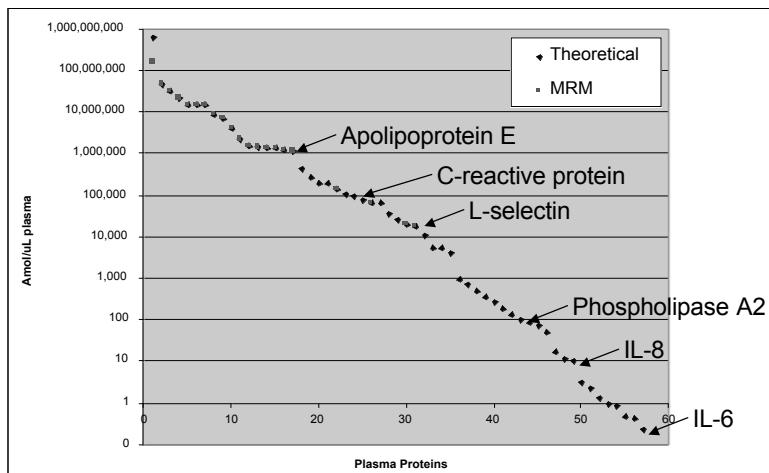
A) 137 MRM's

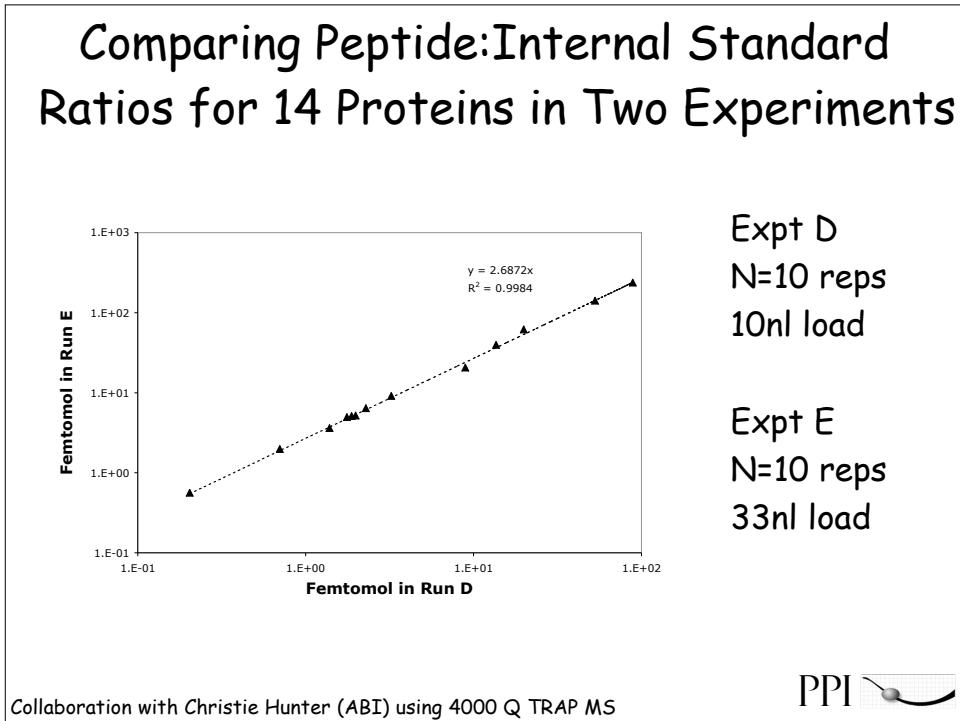
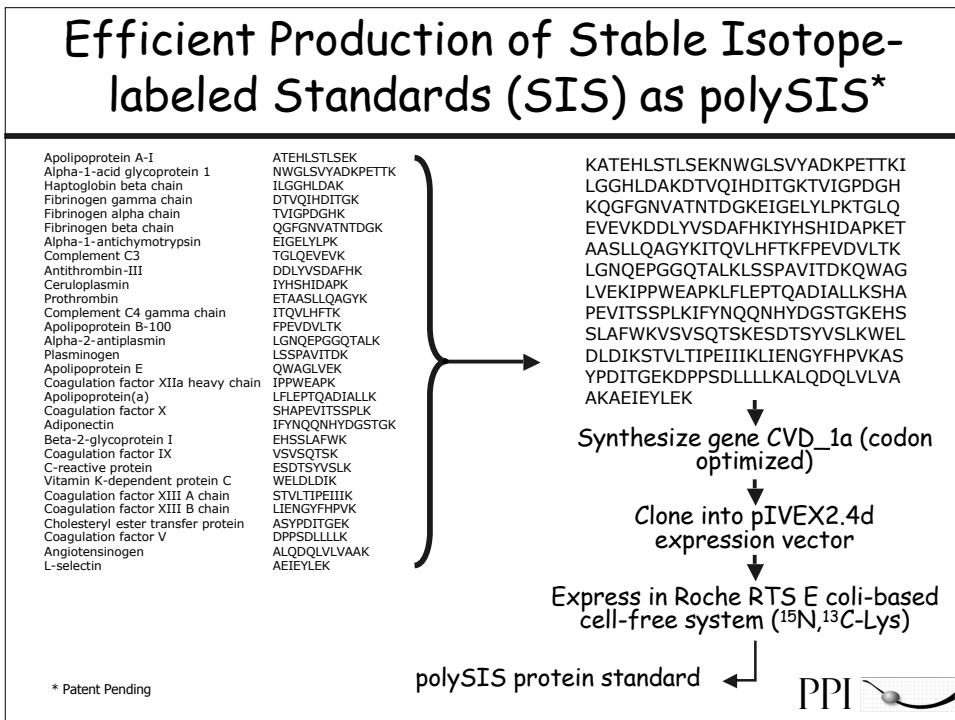
- CV's from 5 experiments, each of 10 replicate runs

B) 47 "best" MRM's



Dynamic Range of Current MRM Validation Method - Depleted Plasma





Sensitivity Enhancement via Specific Capture

Employ MS/MS assay as "2nd antibody" with absolute specificity

Add specific capture step, e.g., using anti-peptide antibody

Demonstrated 10^2 -fold sensitivity improvement with polyclonals

Up to 10^5 -fold enhancement expected with monoclonal Ab's

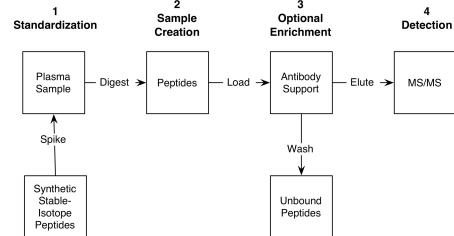
Forgiving of Ab performance

Complete specificity not required (MS does this)

Ab does not quantitate (stable isotope standards + MS do this)

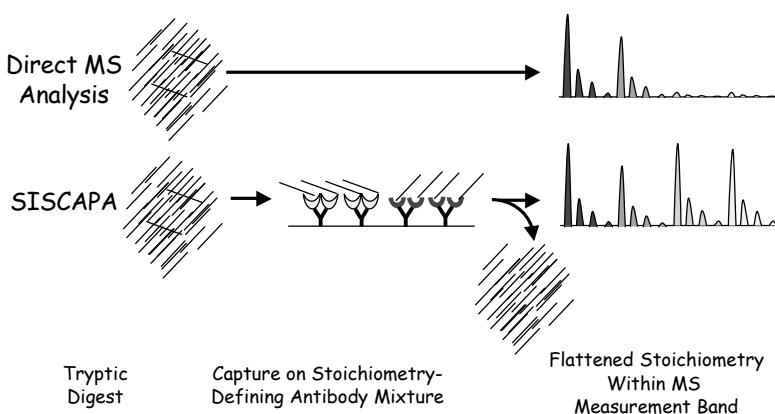
SISCAPA is an immunoassay in which the second (detection) antibody is replaced by a mass spectrometer. This provides:

- Increased assay specificity (absolute structural specificity of MS/MS)
- High sensitivity (low fmol- high amol at peptide level)
- Lower cost (one Ab instead of two)
- Lower analyte-analyte interference (and greater sample stability) by digesting proteins to peptides



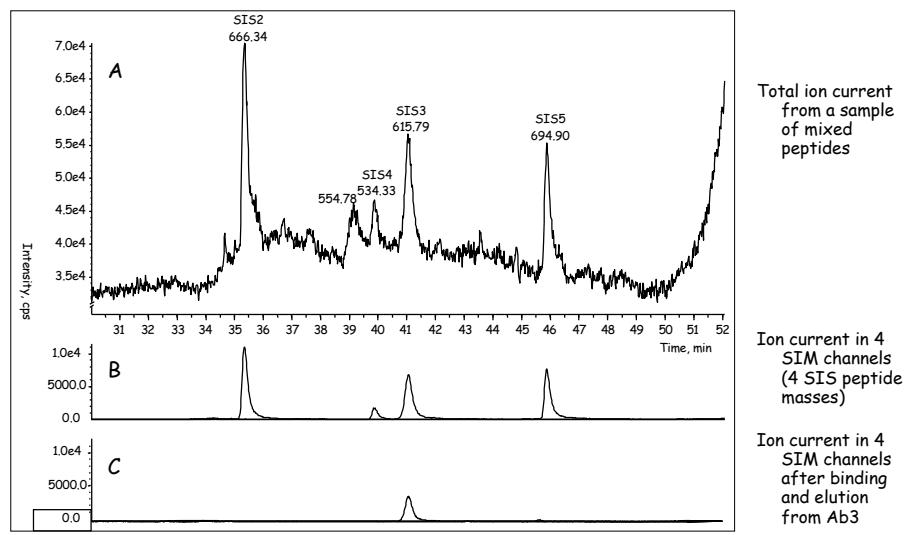
PPI

SISCAPA Captures Similar Amounts of Peptides, Flattening the Abundance Distribution to Accommodate MS Dynamic Range



PPI

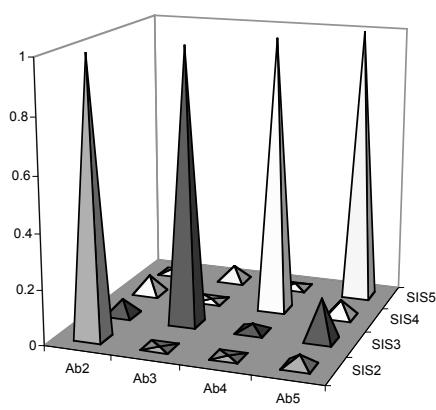
Ab Capture from SIS Peptide Mixture with Selected Ion Monitoring (SIM)



Mass Spectrometric Quantitation of Peptides and Proteins Using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA).
Anderson, N.L., et al, Journal of Proteome Research, 3: 235-44 (2004).



Relative Quantities of Four SIS Peptides Bound by Four Anti-Peptide Antibodies, Using Two-stage MS Selection (SRM) Average Peptide Enrichment by Ab > 100-fold



The signals (vertical axis) for each antibody are normalized to the largest signal for that antibody

Mass Spectrometric Quantitation of Peptides and Proteins Using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA).
Anderson, N.L., et al, Journal of Proteome Research, 3: 235-44 (2004).



Critical Objectives for Biomarker Proteomics

- Progress candidate biomarkers into large-scale validation testing
 - Bring reliable, quantitative proteomics to the clinical and epidemiology communities
- Generate unequivocal successes: biomarkers that transform clinical treatment of patients
- Answer basic biological questions:
 - Are there plasma biomarkers defining any given disease state? (i.e., resolve the biomarker hypothesis)
 - If panels are required in the general Dx case, how many protein components are usually required?
- Transform healthcare economics by providing fast, cheap, accurate diagnosis



Acknowledgements

SISCAPA Experiments

Bob Olafson, Darryl Hardy, UVIC-
Genome B.C. Proteomics Centre
Terry Pearson, Lee Haines, Department
of Biochemistry and Microbiology,
University of Victoria, B.C., Canada
John Rush, Cell Signaling

Plasma Proteome Database

Malu Polanski (PPI)
Richard Fagan, Anna Lobley,
Inpharmatica Ltd., London
Rembert Pieper, Tina Gatlin, present
address: The Institute for Genomic
Research

Radhakrishna S. Tirumalai, Timothy D.
Veenstra, Mass Spectrometry Center,
U. S. National Cancer Institute

Joshua N. Adkins, Joel G. Pounds,
Biological Sciences Department,
Pacific Northwest National
Laboratory

polySIS labeled peptide standards

Jerry Becker, Andrew Breite,
Roche Protein Expression Group,
Indianapolis

MRM Assay Development

Christie Hunter, Applied
Biosystems, Foster City
MALDI TOF/TOF
Glen Hortic, Alan Remaley, Steve
Drake, Gerry Hoehn, Clinical
Chemistry Laboratory, NIH
Clinical Center

Graphics
Arkitek Studios, Seattle

Grant Funding

National Cancer Institute (contract
23XS144A)

www.plasmaproteome.org



Recent Relevant Papers

The human plasma proteome: History, character, and diagnostic prospects. Anderson, N.L. and Anderson, N.G., Molecular and Cellular Proteomics, 1:11, 845-867 (2002)

The human serum proteome: Display of nearly 3700 chromatographically separated protein spots on two-dimensional electrophoresis gels and identification of 325 distinct proteins. Pieper, R., et al Proteomics 3(7): 1345-64. (2003).

Multi-component immunoaffinity subtraction chromatography: An innovative step towards a comprehensive survey of the human plasma proteome. Pieper, R., Su, Q., Gatlin, C. L., Huang, S. T., Anderson, N. L., Steiner, S. Proteomics 3(4): 422-32 (2003).

Therapeutic potential of the plasma proteome. Lathrop, J.T., Anderson, N.L., Anderson, N.G., and Hammond, D.J. Current Opinion in Mol. Therapeutics 5:250-257 (2003).

Mass Spectrometric Quantitation of Peptides and Proteins Using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). Anderson, N.L., Anderson, N.G., Haines, L.R., Hardie, D.B., Olafson, R.W., and Pearson, T.W. Journal of Proteome Research, 3: 235-44 (2004).

NHLBI Clinical Proteomics Working Group Report. Granger, C.B., Van Eyk, J.E., Mockrin, S.C., and Anderson, N.L., on behalf of the Working Group Members. Circulation 109: 1697-703 (2004).

Candidate-Based Proteomics in the Search for Biomarkers of Cardiovascular Disease, Leigh Anderson, J. Physiol., 563:1:23-60 (2005).

The Roles of Multiple Proteomics Platforms in a Pipeline for New Diagnostics, N. Leigh Anderson, Mol Cell Proteomics, in press 2005

