Abstract # 4489

A novel technological approach for the discovery of low molecular weight serum biomarkers based on "smart" hydrogel nanoparticles applied to ovarian cancer.

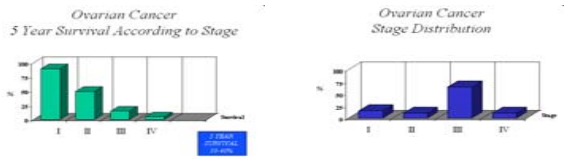
Francesco Meani[1,2], Alessandra Luchini[1], Claudia Fredolini[1], Gudo Gambara[1], Caterina Longo[1], Akrita Bhatnagar[1], Weidong Zhou[1], Mark Ross[1], Davide Tamburro[1], Sergio Pecorelli[2], Antonella Ravggi[2], Franco Odicino[2], Paul Russo[1], Virginia Espina[1], Emanuel F. Petricoin III[1], Lance Liotta[1]

[1] Center for Applied Proteomics and Molecular Medicine, George Mason University, [2] University of Brescia - Italy

Introduction

The "SILENT KILLER"

Ovarian Cancer is the fourth most frequent cause of cancer death and the most lethal of all gynaecologic tumours in women from North America. Each year, approximately 20,000 American women are diagnosed with ovarian cancer and about 15,000 women die of the disease. Despite its incidence being ten times lower than breast cancer, it is three fold more lethal and it is usually not diagnosed until its advanced stages, mainly because of the absence of early signs and symptoms and the lack of a screening protocol with adequate specificity and sensitivity.



The prognosis for survival from ovarian cancer largely depends on the extent of disease at diagnosis. **Detection of early stage disease offers an opportunity to reduce mortality.** Proteomics holds great promise as the development of panels of biomarkers, which can generate a high level of diagnostic specificity and ultimately will lead to early detection and tailored therapy.

Challenges in Biomarker Discovery

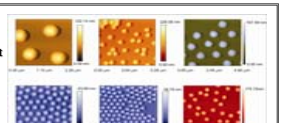
Analytical and physiological barriers hinder biomarkers discovery in an extraordinarily complex matrix biospecimen such as blood:

1) Recent interest has focused on the low molecular weight (LMW) "peptidome" as a potentially rich and relatively unexplored source of circulating markers: analytes that likely represent the most clinical important markers are too low in abundance for even the most sophisticated MS technology to identify.
2) Overwhelming abundance of resident proteins such as albumin and immunoglobulins, accounting for >95% of circulating plasma proteins, which confound and mask the isolation of rare biomarkers, often associated with the resident proteins.
3) Degradation of candidate biomarkers occurs rapidly by endogenous and exogenous proteinases immediately after the blood sample is drawn from the patient.

"Smart" Nanoporous Particles: *N*-isopropylacrylamide (NIPAm) based polymers + AAc bait

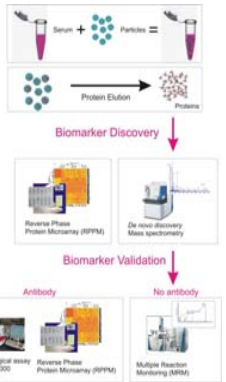
Four independent functions within minutes, in one step, in solution:

- a) molecular size sieving (<20.000D basic proteins positively charged at physiological PH)
- b) concentration of all solution phase target molecules (affinity capture)
- c) complete protection of harvested proteins from enzymatic degradation
- d) complete exclusion of albumin and high abundance proteins



Atomic Force Microscope image of NIPAm-AAc particles

Novel Workflow for Biomarker Discovery



In order to overcome and directly address the challenges associated with biomarker discovery, we have invented a novel technology: a biomarker "harvesting" nanoparticle that can nearly instantly in solution, collect, concentrate, and preserve from degradation LMW analytes will excluding high abundance HMW molecules. De novo candidate biomarker discovery with upfront nanoparticle-based sample fractionation proceeds with high resolution mass spectrometry followed by bioinformatic data mining and data processing/ database queries *(The Biomarker SIEVE: a computational framework for lable-free MS-based Proteomic Biomarker Discover. Askenazi, Zumwalt; © 2008 Sequest Technologies Inc.)*

Once the biomarker candidates have been identified, verification and validation of the sensitivity and specificity of the candidates are revealed by analysis using independent sample sets.

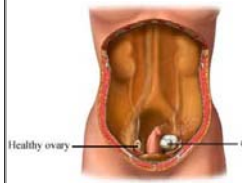
Immunological or quantitative mass spectrometry technologies are used depending on antibody sourcing: 1) If only a single validated antibody exists for the biomarker, then Reverse Phase Protein Microarrays (RPMA) can be used. 2) If an antibody sandwich pair exists then a clinical grade ELISA or bead assay can be used. 3) If no antibodies exists, or the analyte is a uniquely sized protein isoform, Multiple Reaction Ion Monitoring (MRM) using triple quadropole mass spectrometry can be used to detect a specific peptide fragment of the analyte. MRM technology does not require a specific antibody.

At George Mason University we have applied such protocol to serum samples from prostate, skin (melanoma) and ovarian cancer patients in order to single out candidate biomarkers possibly useful for early diagnosis, prediction of response to chemotherapy and follow up, of these clinically relevant diseases.

Materials & Methods

Samples

The protocol was applied to a study set of 20 early stage (Ia - IIb) epithelial ovarian cancer serum samples (including six borderline cases) and 20 benign controls (of which 13 were classified in the category of intermediate risk at US pelvic examination). Medium age was 47 in the disease group and 48 for the controls.



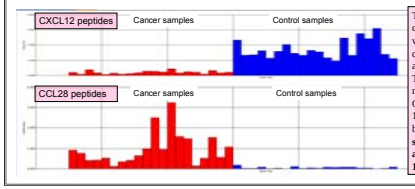
CASES

Age	Menopausal Status	Ultrasound Risk	Stage	Pathology
64	Post-	High	II	S. Bord.
63	Post-	High	IC	M. Bord.
56	Post-	Unknown	IIA	C. C. Ca.
57	Post-	High	IA	C. C. Ca.
48	Pre-	Intermediate	IA	M. AC.a.
42	Pre-	High	IA	S. Bord
24	Pre-	High	I	AM. Bord.
24	Pre-	Intermediate	IA	S. Ca.
19	Pre-	Intermediate	I	M. ACa.
54	Post-	High	IIA	E. ACa.
58	Post-	High	IA	Cl. C. Ca.
20	Pre-	Intermediate	IA	M. Bord.
39	Post-	Intermediate	I	M. Bord.
58	Post-	Unknown	IC	ACa.
68	Post-	Low	IC	ACa.
56	Post-	Low	IC	Unknown
33	Pre-	Unknown	IA	ACa.
53	Pre-	Unknown	IA	E. ACa.
59	Post-	High	IC	ACa.
29	Pre-	Low	IC	ACa.

CONTROLS

Age	Menopausal Status	Ultrasound Risk	Pathology
21	Pre-	Intermediate	M.C.T.
21	Pre-	Intermediate	M.C.T.
39	Pre-	Intermediate	C. L C.
39	Pre	Intermediate	E. C.
36	Pre	Intermediate	F. C.
68	Post-	Intermediate	S. CAd.
50	Pre-	High	F. C.
35	Post-	Intermediate	S. CAd.
41	Pre-	Intermediate	Unknown
52	Pre-	Intermediate	M.C.T.
58	Post-	Intermediate	S. CAd.
52	Pre-	Intermediate	S. CAF.
65	Post-	Intermediate	M.C.T.
65	Pre-	Unknown	F. C.
59	Unknown	Unknown	B. C.
58	Post-	Unknown	S. CAF.
74	Post-	Unknown	M. CAd.
53	Unknown	Unknown	S. CAd.
48	Pre-	Low	S-M. Cad.
36	Pre-	Low	S. CAd.

M.C.T. = Mature Cystic Teratoma
F. C. = Follicular Cyst
S. CAd. = Serous Cystoadenoma
M. C. = Mucinous Cystoadenoma
S-M. CAd. = Serous-Mucinous Cystoadenoma
S. CAF. = Serous Cystis-AdenoFibroma
B. C. = Dermoid Cyst
C. L C. = Corpus Luteum Cyst
E. C. = Endometrioxic Cyst
E. ACa. = Endometrioid AdenoCarcinoma
M. C. = Mucinous
ACa. = AdenoCarcinoma
Cl. C. Ca. = Clear Cell Carcinoma
S.=Serous



The addition of internal standards was performed as quality control in the process. The samples (400ul) have been spiked with commercially available protein standards to allow quantitative characterization and monitoring of MS analytical performance.

Two chemokines, CCL28 (known also as MEC, for mammary enriched chemokine. Predicted MW 14KD) and CXCL12 (SDF, stromal derived factor 1. Predicted MW 11KD), were chosen for cancer related biomarker analysis because of their small size and short half life. **Positive samples were spiked with 100ng MEC and 10ng SDF-1 as opposed to benign controls which were spiked with 10ng MEC and 100ng SDF-1.**

LC-MS/MS analysis of nanoparticle-sampled serum

Sample preparation for mass spectrometric analysis

Eluted proteins from nanoparticle were reduced by DTT, alkylated by iodoacetamide, and digested by trypsin overnight at 37 °C. The digestion mixture was desalted by ZipTip.

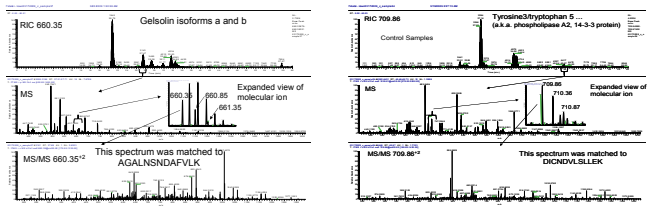
Reversed-phase liquid chromatography tandem mass spectrometry

The peptides from each sample were analyzed by reversed-phase liquid chromatography electrospray tandem mass spectrometry using LTQ-Orbitrap mass spectrometer (ThermoFisher). Reverse phase C18 column (0.5×5mm) was from Michrom BioResources. After sample injection by autosampler, the column was washed for 2 min with mobile phase A (0.1% formic acid) and peptides were eluted using a linear gradient of 0% mobile phase B (0.1% formic acid, 80% acetonitrile) to 50% mobile phase B in 50 min at 2 µl/min, then to 100% B in an additional 2 min. The LTQ mass spectrometer was operated in a data-dependent mode in which each full MS scan was followed by five MS/MS scans where the five most abundant molecular ions were dynamically selected and fragmented by collision-induced dissociation (CID) using a normalized collision energy of 30%. To decrease carryover, one blank (AngI) was run by 30 min HPLC double gradient program after analyzing one sample. To verify HPLC reproducibility, 2 standard (Enolase digestion mixture from Michrom BioResources) were run by 30 min HPLC program after analyzing 15 serum samples.

HPLC reproducibility test: standard enolase digestion was analyzed to make sure that the HPLC is reproducible.

Bioinformatic analysis

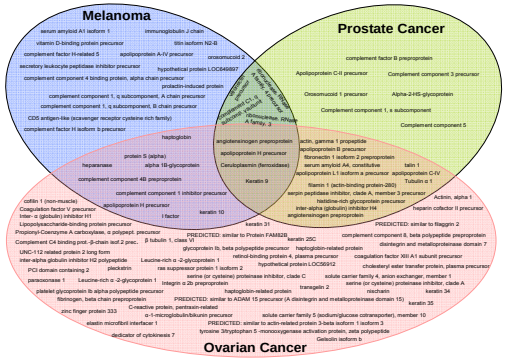
The MS data are searched against a human protein database (NCBI) with SEQUEST. Spectral counting (MS2-based) analysis is accomplished using SCAFFOLD (Proteome Software, Inc.) in order to identify likely differentially expressed candidates. The MS1-based comparative analysis is accomplished using the SIEVE software (Thermo and Vast Scientific), a probabilistically-driven tool for differential expression analysis of MS1 spectra.



Examples of MS and MS/MS chromatogram of two of the Nanoparticles yield serum proteins from ovarian cancer and control serum samples

AACR CENTENNIAL 1907-2007
100th Annual Meeting 2009
April 18-22, 2009 • Colorado Convention Center • Denver, Colorado

Results



Venn diagram of differentially expressed protein yield from OVARIAN, SKIN and PROSTATE cancers.

Case v. Control Ovarian Cancer Candidate Differentially Abundant Proteins Selected from Scaffold Results

	Spectral Count Results		Manual Analysis						
	# Assigned Spectra		# peptides	Sequest	ID Likely	Differential Abundance			
Protein	Controls	Cases	evaluated	scores	Correct			Notes	SIEVE
a disintegrin and metalloproteinase d	1	4	1 of 1	2.4, 0.1	no	possibly	Ca>Ctrl		cannot find
dedicator of cytokinesis	0	5	1 of 1	2.7, 0.3	no	possibly	Ca>Ctrl		p>0.01
nischarin	1	2	1 of 1	2.6, 0.2	no	unlikely			cannot find
PCI domain containing 2	0	3	1 of 1	2.5, 0.2	no	unlikely			p>0.01
ras suppressor protein 1 isoform	0	3	1 of 1	3.6, 0.5	yes	uncertain		ion in 1 of 4 samples	cannot find
zinc finger protein 333	0	6	1 of 1	2.3, 0.2	no	possibly	Ca>Ctrl	good MS2, no good match	cannot find
tyrosine3/tryptophan 5-...	3	5	1 of 1	3.3, 0.2	yes	possibly	Ca>Ctrl		cannot find
gelsolin isoform b	30	47	1 of 3-7	3.7, 0.4	yes	possibly	Ca>Ctrl	high variability	p>0.01
integrin alpha 2b	0	28	1 of 2-10	2.4, 0.3	yes	possibly	Ca>Ctrl	high variability	p>0.01
leucine-rich-alpha-2-glycoprotein 1	3	9	1 of 2-3	2.3, 0.2	yes	possibly	Ca>Ctrl	high variability	cannot find
lipopolysaccharide-binding protein	2	28	1 of 2-3	4.1, 0.5	yes	likely	Ca>Ctrl	high variability	p>0.01
propionyl-coenzyme A carboxylase	7	11	1 of 1	2.3, 0.2	no	unlikely		good MS2, no good match	incorrect framing
unc-112 related protein 2 long form	0	14	1 of 3-4	3.1, 0.6	yes	possibly	Ca>Ctrl	unexplained ions in MS2	p>0.01

Candidate differentially abundant peptides, chosen on the basis of their biologic function, have been evaluated by manual inspection of the raw data to confirm peptide identification and determine intensities providing us with statistics conveying the differential expression of the protein biomarkers.

Conclusions

• With this work we have begun to test a novel workflow for body fluid-based biomarker discovery. We have developed hydrogel nanoparticles to amplify, fractionate and enrich LMW analytes in a single simple step that in just a few minutes can harvest candidate biomarker analytes from raw unprocessed serum, providing a powerful pre-analytical method for mass spectrometry-based biomarker discovery.

• The list of identified protein analytes confirms that LMW proteins, protein fragments and peptides are bound by the chemical bait inside the particles, while the shell acts as a sieve, excluding high molecular weight, high abundance proteins such as albumin.

•The addition of internal standards, intended as quality control in the process, reveals that the final MS measurements collected are consistent with the expected concentration differences of the spike-in standards.

•The software driven MS analysis performed to date revealed a list of potentially differentially expressed early stage (I-IIb) ovarian cancer biomarker candidates that were elevated in the serum from ovarian cancer patients compared to serum taken from age-matched benign controls from women with adenomas and benign ovarian cysts.

• Comparison of these ovarian cancer candidates to biomarker candidates discovered using the identical workflow applied to serum sets from patients with other cancers and appropriate control groups, namely prostate and melanoma revealed only marginal overlap indicating potential for specificity in the candidates uncovered to date.

• Based on the MS analysis and subsequent manual verification of the spectra, we are focusing our validation work on the following biomarker candidates: Ras suppressor protein isoform 1, tyrosine3/tryptophan5-monoxygenase activation protein, zeta polypeptide, gelsolin isoform b, integrin alpha 2b, leucine-rich alpha-2-glycoprotein 1, LPS-binding protein, and UNC-112 related protein 2 long form.

• Future work will encompass antibody-based (e.g. ELISA, RPMA) and antibody-independent (MRM-MS) verification and validation of differential expression in independent blinded serum study sets. These study sets will include benign and inflammatory diseases as well as other cancers to understand the specificity and sensitivity of each candidate alone and in combination with each other.