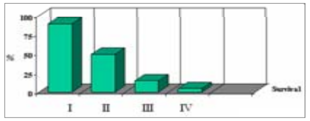
IDENTIFICATION OF A FOUR MARKER PANEL FROM THE SERUM FRAGMENTOME FOR THE CLASSIFOCATION OF STAGE I OVARIAN CANCER

S. Kuzdzal, M. Lopez, W. Patton, A. Mikulskis (PerkinElmer, MA), D. Sarracino (Harvard-Partners Center for Genetics & Genomics, MA), M. O'Gorman (Nonlinear Dynamics, UK), K. Rosenblatt, P. Gurnani, & A. Nandi (UTSW, TX), G. Whiteley (NCI, MD), D. Fishman (New York University, NY), E. Petricoin, L. Liotta (George Mason University, VA).

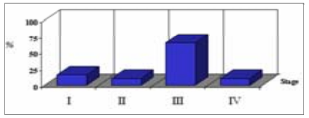
INTRODUCTION

All cancer patients benefit from early detection. The need for development of early screening methods for epithelial ovarian carcinoma is particularly urgent. Ovarian cancer is the fourth leading cause of cancer related deaths among women in the United States. Regrettably, 70-75% of new diagnoses will be Stage III or IV with a predicted 5-year survival of ca.15%. This survival rate has remained essentially unchanged over the past 40 years despite the advent of intensive radical surgical approaches and new therapeutic interventions. In contrast, if cancer is detected when confined to the ovary (Stage I), the 5-year survival approaches 90%, requires less radical operations, and may not require adjuvant chemotherapy. More sensitive and specific tests for earlier detection would greatly improve patient survival rates by facilitating early treatments such as surgical intervention.

Five Year Survival According to Stage



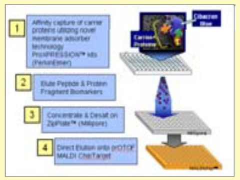
Stage Distribution of Clinical Detection



METHODS

The ovarian cancer serum biomarker study was performed using the PerkinElmer BioXPRESSION Platform (below).

Sample Preparation

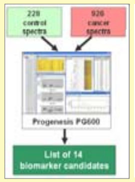


High-Throughput Biomarker Screening

MALDI O-TOF MS

Mass accuracy
<10 ppm external calibration
Mass stability
10 ppm over several hours using external calibration
Resolution
15,000 (FWHM)
Detection Limit:
Sub-femtomole
Throughput
>10,000 samples/day

Biomarker Detection



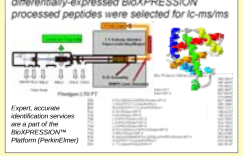
PROGENESIS

PG600 Software

Robust Biomarker Discriminant Analysis Software from Nonlinear Dynamics:
- Intelligent Peak Detection
- Spectral Display Options
- Rapid Differential Display
- UPGMA & NJ Clustering
- Grouping & Averaging According to Class

Biomarker Identification

Accurate mass, isotopically-resolved, differentially-expressed BioXPRESSION processed peptides were selected for lc-ms/ms



Expert, accurate identification services are a part of the BioXPRESSION™ Platform (PerkinElmer)

Study & Clinical Sample Information:

The ovarian study set consisted of 560 total samples. The number of serial samples collected from individual patients varied from 3 to 22 samples per patient. Samples from cancer patients were also assayed for CA125 levels. There were 213 samples with low (below 22 ng/ml) and 239 samples with high (above 22ng/ml) CA 125.

Healthy control individual serum samples were collected from unaffected women (no serial samples), ovarian cancer samples were procured in a gynecologic oncology clinic from symptomatic women who were later surgically staged and found to have epithelial ovarian carcinoma. Each sample was accompanied by a verified pathology diagnosis. Specimens were collected in red-top Vacutainer Tubes and allowed to clot for 1 h on ice, followed by centrifugation at 4 °C for 10 min at 3000 rpm. The serum supernatant was divided in aliquots and stored at -80 °C until needed. Classification model construction was constructed by Martin O'Gorman (Nonlinear Dynamics, UK). Models, created from a training set of 114 disease and control spectra (out of a total of 1154 spectra) were then applied to the testing set (remainder of samples).

Treatment	Number of samples	Analysis replicates
Healthy (control)	108	324
Early stages of ovarian cancer (low CA125)	213	639
Late stages of ovarian cancer (high CA125)	239	717
Total samples	560	1680

Ovarian cancer serum samples were obtained courtesy of E. Petricoin, FDA, G. Whiteley, NCI and D. Fishman, NYU.

RESULTS

The BioXPRESSION Platform (PerkinElmer SCIEX) provides high-resolution, high-sensitivity analysis of the serum fragmentome. High throughput screening of carrier-protein bound protein fragments by MALDI O-TOF MS (PrOTOF 2000, PerkinElmer SCIEX) rapidly identified differentially-expressed proteins and peptides (Figure 1).

Scale-up of the BioXPRESSION Platform resulted in accurate sequence information by LC/MS/MS (Table 2). Multiple protein fragments (with very high identification probabilities) were obtained for most serum proteins (Table 3).

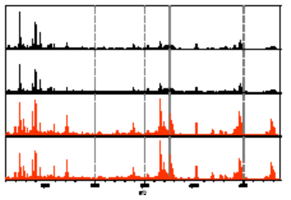


Figure 1. ProXPRESSION Biomarker HT Kit processed sera replicates analyzed by high resolution MALDI Orthogonal-TOF MS (PrOTOF 2000, PerkinElmer SCIEX). Mass range displayed = 1600 – 4300 Da, Top (black) = Control, Bottom (Red) = ovarian cancer.

Protein	Observed Fragment	MW (Da)	P (Pep)
Transthyretin gi|4507725|ref|NP_000362.1|	R.RYTIAALLSPYSYSTTAVV TNPK.E	2516.34	2.68E-04
Granzyme H gi|15529990|ref|NP_219491.	Q.EKSRKRCGGILVRKDFVL TAAHCQGSSIN.V	3173.68	2.47E-03
Yippee-like 4 gi|21450727|ref|NP_659445.1|	-.M*PSCDPGPGPACL.P	1260.51	8.73E-04
Olfactory Receptor, family 6, subfamily C, member 3 gi|50080213|ref|NP_473445.1|	F.IFM*GVTEFYILTAMSYDR YV.A	2435.17	3.85E-04
Lysosomal associated multispanning membrane protein 5 gi|5803056|ref|NP_006753.1|	D.MPHNQFIKM*M*IIFS.I	1768.86	6.40E-03
F-box and leucine-rich repeat protein 15 gi|71061449|ref|NP_077302.2|	I.PRAALARLLRDAE.G	1451.84	1.68E-03
Acetyl-Coenzyme A acetyltransferase 2 gi|5174389|ref|NP_005882.1|	D.KEIVPVLVSTRKGL.I	1538.96	2.27E-04
ELL associated factor 2 gi|21361796|ref|NP_060926.2|	A.EASLMDQM*S.S	1027.41	7.86E-03
Corneodesmosin precursor gi|67782356|ref|NP_001255.3|	R.GILNPSQPGQSSSSSQT.F	1674.79	4.80E-06
Differentially expressed in FDCP 8 isoform 1 gi|46447820|ref|NP_997397.1|	Q.YYCSHCHWNDLAVI.P	1723.74	3.72E-04
Apolipoprotein L1 isoform b precursor gi|21735616|ref|NP_663318.1|	C.SDMEGAALLRV.S	1161.59	3.22E-03
Nuclear prelamin A recognition factor isoform b gi|14165461|ref|NP_114174.1|	R.TLPMLTSACPGW.V	1276.61	2.75E-03
Transforming growth factor, beta-induced, gi|4507467|ref|NP_000349.1|	R.SVRLAPVYQKLLERM*.K	1819.03	1.95E-03
PREDICTED: proline-rich synapse-associated protein 2 gi|51476101|ref|XP_037493.5|	P.ESAADSGVEEADTRSSS DPHLETTSTISTVSSMSTLS SE.S	3988.76	3.52E-03
p21-activated kinase 7 gi|24308191	K.KKIEISGPSNFEH.R	1485.77	2.04E-05
Vacuolar protein sorting 16 isoform 3 gi|17978481|ref|NP_536338.1|	Q.GVSRILAHWACY.K	1375.69	6.55E-03
Coagulation factor XII precursor gi|4503629|ref|NP_000496.1|	K.REQPPSLTR.N	1083.59	1.25E-04
Desmoglein 1 preproprotein gi|4503401|ref|NP_001933.1|	Y.VVTGNM*GSNDKVGDF.V	1555.71	1.26E-09
Hornerin gi|57864582|ref|NP_001009931.1|	Y.GSGSGWSSSRGPY.E	1284.56	3.28E-04
Brain-specific angiogenesis inhibitor 3 gi|4502359|ref|NP_001695.1|	I.GILVFNKLVSRDGIL.D	1643.98	8.21E-03
Connector enhancer of Kinase suppressor of Ras 2 gi|41393057|ref|NP_055742.2|	A.WLDRSPF.A	920.46	4.67E-03
Zonadhesin isoform 1 gi|27881486|ref|NP_775078.1|	N.SPSCDSSLQSSMSGPGF .C	1673.68	2.28E-03
Nedd4 binding protein 2 gi|31742492|ref|NP_060647.2|	K.RDPGKEVGM*CT.Q	1208.54	1.81E-03
Sodium channel, voltage-gated, type X, alpha gi|5730033|ref|NP_006505.1|	S.EEKTKIL.G	860.51	9.20E-04
Kininogen 1 gi|4504893|ref|NP_000884.1|	K.RPPGFSPF.R	904.47	2.45E-03

Protein	Observed Fragment	MW (Da)	P (Pep)
Thymosin-like 3 gi|34013530|ref|NP_898870.1|	E.KNPLPSKETIEQEKQAGE S.-	2113.08	3.95E-04
Neurogranin gi|5453800|ref|NP_006167.1|	R.KGPGPGGPGGAGVARG GAGGGP.S	1687.86	2.25E-12
Zyxin gi|4508047|ref|NP_003452.1|	S.LANTQPRGPPASSPAPA PKFSPVTPKFTPVAS.K	3215.72	5.72E-08
Alpha-synuclein isoform NACP112 gi|6806898|ref|NP_009292.1|	G.KTKEGVLYVGSKTKEGV VH.G	2059.15	3.83E-05
Phosducin-like gi|9943842|ref|NP_005379.2|	A.AEYPAVKFCKVKSSVIGA SSQFTRNA.L	2788.45	2.10E-04
Caldesmon 1 isoform 4 gi|15149463|ref|NP_149130.1|	K.SVQKSSGVKSTHQAAIVS KIDSRLE.Q	2655.44	2.02E-04
Myeloid-associated differentiation marker gi|19923903|ref|NP_612382.1|	M.PVTVTRTTITTTT.T	1391.77	4.03E-03
Splicing factor, arginine/serine-rich 8 isoform 1 gi|23111062|ref|NP_004583.2|	K.SGAKEEAGPGGAGGGG SRVELLVFGYACKLFRDDE RAL.A	3839.91	4.42E-05
Drebrin-like isoform a gi|21361670|ref|NP_054782.2|	E.SAVHPREIFKQKER.A	1724.96	3.67E-05
Oxysterol-binding protein-like protein 9 isoform e gi|20070331|ref|NP_078862.2|	D.FDKKLTEADAYLQILIEQL KLFD.D	2754.50	1.55E-03
p47 protein isoform a gi|20149635|ref|NP_057227.2|	S.KPRPFAGGGY.R	1049.55	1.46E-05
Developmentally regulated GTP binding protein 2 gi|4557537|ref|NP_001379.1|	H.KPNIYFKPKKGGGISFNS TV.T	2182.20	9.22E-03
Discoidin, CUB and LCCL domain containing 1 gi|27735143|ref|NP_775945.1|	G.PSWASGDSSNNHKPR.E	1639.76	1.22E-03
PiggyBac transposable element derived 4 gi|25777746|ref|NP_689808.2|	E.SDGKSSTSSDSGRSMK. W	1616.72	2.59E-03
Notch4 preproprotein gi|55770876|ref|NP_004548.3|	G.TCQNSAGSFHC.V	1154.44	3.99E-03
Calcium channel, voltage-dependent, alpha 1E subunit gi|53832005|ref|NP_000712.2|	T.EGEKDHELRGNHG.A	1477.68	1.79E-03
Wee1 tyrosine kinase gi|4507917|ref|NP_003381.1|	D.PERRPSAMALVKHSVLLS ASRKS.A	2520.42	3.72E-03
Epithelial protein lost in neoplasm beta gi|7705373|ref|NP_057441.1|	D.LKKLRRSSSLKERSRPF. T	2088.25	4.57E-03
Solute carrier family 12 (potassium/chloride transporters), member 7 gi|5730043|ref|NP_006589.1|	L.NKSQDAQLVLLNMPG.P	1627.85	6.15E-03
PREDICTED: ubiquitin specific protease 24 [gi|51458553|ref|XP_371254.3|	N.YELVGVIVHSGQ.A	1300.69	3.46E-03
MAX dimerization protein 3 gi|13775222|ref|NP_112590.1|	A.RQLKERLRSKQQSLQR.Q	2054.21	1.80E-03
Peroxisome proliferator-activated receptor (PPAR) binding protein gi|28559039|ref|NP_004765.2|	T.AVIDKLKHGVVTSGPGGE DPLDGQMGVSTNSSSHPM SSKHNM.S	4289.04	9.99E-04
Ca2+-dependent secretion activator isoform 2 gi|34452713|ref|NP_899631.1|	H.KM*TVSKNCPDQDLKIKL AVRM*DKPQNMKH.S	3428.76	2.43E-03
ATP-binding cassette, sub-family C, member 12 isoform b gi|21729888|ref|NP_660189.1|	Y.SERSPPAKGATG.P	1157.59	5.49E-03
SET binding protein 1 gi|7662122|ref|NP_056374.1|	N.KDLLLGGVAPSP.S	1166.68	2.46E-03

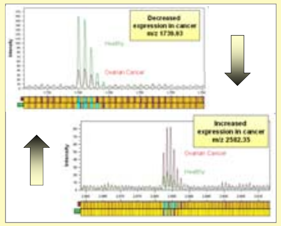
Table 2. Partial listing of some of the protein fragments identified using the BioXPRESSION Platfrom (PerkinElmer SCIEX). Thousands of fragments representing over 200 human serum fragments have been identified.

Observed Fragment	Fragment MW (Da)	P (Pep)
K.PEASFSPR.R	890.44	7.29E-04
I.PKPEASFSPR.R	1115.58	5.26E-04
P.PDVPDHAAYHPF.R	1365.62	1.11E-03
G.PPDVPDHAAYHPF.R	1462.67	1.45E-03
P.GPPDVPDHAAYHPF.R	1519.7	1.20E-03
L.PGPPDVPDHAAYHPF.R	1616.75	2.59E-08
F.FKYYLQGAKIPKPEA.S	1752.97	1.09E-04
Y.YLQGAKIPKPEASFSPR.R	1889.03	3.41E-07
T.FFKYYLQGAKIPKPEA.S	1900.04	9.07E-08
T.FFKYYLQGAKIPKPEAS.F	1987.07	6.05E-09
K.YYLQGAKIPKPEASFSPR.R	2052.09	9.15E-10
H.SGSTFFKYYLQGAKIPKPEAS.F	2319.2	5.56E-07
V.HSGSTFFKYYLQGAKIPKPEA.S	2369.23	8.10E-05
P.GVLSSRQLGLPGPPDVPDHAAYHPF.R	2627.34	8.04E-13
R.NVHSGSTFFKYYLQGAKIPKPEAS.F	2669.37	7.81E-05
R.PGVLSSRQLGLPGPPDVPDHAAYHPF.R	2724.39	1.62E-11
H.SGSTFFKYYLQGAKIPKPEASFSPR.R	2806.46	2.71E-13
N.FRPGVLSSRQLGLPGPPDVPDHAAYHP.F	2880.49	7.05E-04
N.FRPGVLSSRQLGLPGPPDVPDHAAYHPF.R	3027.56	2.97E-06
R.NVHSGSTFFKYYLQGAKIPKPEASFSPR.R	3156.63	1.91E-09
R.MNFRPGVLSSRQLGLPGPPDVPDHAAYHPF.R	3272.64	2.01E-06
R.M*NFRPGVLSSRQLGLPGPPDVPDHAAYHPF.R	3288.64	2.37E-05
R.QAGAAGSRMNFRPGVLSSRQLGLPGPPDVPDHAAYHPF.R	3970.99	1.58E-05
R.QAGAAGSRM*NFRPGVLSSRQLGLPGPPDVPDHAAYHPF.R	3986.98	2.66E-04
W.NRQAGAAGSRM*NFRPGVLSSRQLGLPGPPDVPDHAAYHPF.R	4257.13	1.08E-03

Table 3. Partial listing of observed protein fragments of Inter-Alpha (globulin) Inhibitor H4 (Plasma Kallikrein-Sensitive Glycoprotein) [P Protein = 2.71E-13]. Multiple fragments were identified for most proteins in this study.

Figure 2. Processing of 96 samples using ProXPRESSION Biomarker HT Kits. These kits fractionate/enrich carrier-protein bound protein fragments (processing may be automated on liquid handlers).

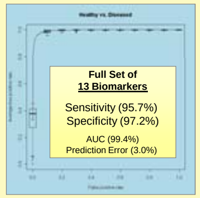
Example differentially expressed fragments from screening

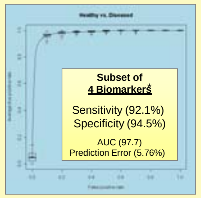


Sequence information for some of the key peptide fragments

m/z	Fold expression cancer vs healthy	p Value	Sequence ID
1690.94	-1.95	0.0000000	S.KITHRIHWESASLL.R gi|4557385|ref|NP_000055.1| complement component 3 precursor [Homo sapiens] [MA
1739.93	-2.63	0.0000000	R.NGFKSHALQLNNRQI.R gi|67190748|ref|NP_009224.2| complement component 4A preproprotein [Homo sapiens]
1777.97	-1.93	0.0000000	S.SKITHRIHWESASLL.R gi|4557385|ref|NP_000055.1| complement component 3 precursor [Homo sapiens] [MA
1865.01	-1.95	0.0000000	R.SSKITHRIHWESASLL.R gi|4557385|ref|NP_000055.1| complement component 3 precursor [Homo sapiens] [MA
2021.11	-1.47	0.0000000	R.SSKITHRIHWESASLLR.S gi|4557385|ref|NP_000055.1| complement component 3 precursor [Homo sapiens] [MA
2898.54	2.09	0.0000000	G.PRRYTIAALLSPYSYSTTAVVTNPKE.- gi|4507725|ref|NP_000362.1| transthyretin [Homo sapiens] [MASS=15887]
3027.57	2.70	0.0000000	N.FRPGVLSSRQLGLPGPPDVPDHAAYHPF.R gi|31542984|ref|NP_002209.2| inter-alpha (globulin) inhibitor H4 (plasma Kallikrein-sensitive glyco

A 4 marker subset of the initial 13 marker set delivered phenomenal sensitivity and specificity





*manuscript in preparation

CONCLUSIONS

The BioXPRESSION Platform is a powerful biomarker screening platform. Many differentially-expressed protein fragments were identified from human serum. A panel of four protein fragment biomarkers classified early stage disease from non-disease with very high sensitivity and specificity. While many of the more differentially-expressed biomarker candidates represent fragments of coagulation proteins, many interesting fragments (PPARs, Notch4, angiogenesis inhibitors, etc.) were also identified. Biomarker subsets have been submitted for publication.