

High-Sensitivity Blood-Based Detection of Breast Cancer by Multi-Photon Detection Diagnostic Proteomics

Andrzej K. Drukier,^{†,‡} Nathasha Ossetrova,[†] Elena Schors,[†] Galya Krasik,[†] Ivan Grigoriev,[†] Christina Koenig,[†] Miko Sulkowski,[†] Jerzy Holcman,[†] Larry R. Brown,[‡] John E. Tomaszewski,[§] Mitchell D. Schnell,^{||} Richard Sainsbury,[⊥] Anne E. Lokshin,[#] and Jasminka Godovac-Zimmermann^{*,^}

BioTraces Inc., 12455 Sunrise Valley Dr., Herndon, Virginia 20171, OncoTraces Ltd, 3 George Street, London, W1U 3QG, United Kingdom, Department of Pathology and Laboratory Medicine, University of Pennsylvania, 3400 Spruce Street, Philadelphia, Pennsylvania 19104, Department of Radiology, University of Pennsylvania, 3400 Spruce Street, Philadelphia, Pennsylvania, 19104, Department of Surgery, University College London, 67–73 Riding House Street, London W1W 7EJ, United Kingdom, Division of Hematology/Oncology, University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania 15213, and Department of Medicine, Rayne Institute, University College London, 5 University Street, London WC1E 6JF, United Kingdom

Received March 8, 2006

We have developed several new methods for blood-based cancer detection by diagnostic proteomics. Ultrasensitive methods of immunoassay using multiphoton-detection (IA/MPD) increase sensitivity by 200- to 1000-fold (1 femtogram/mL). This has allowed the measurement of cancer biomarkers with very low concentrations in blood that could not be measured for full patient cohorts with conventional immunoassays. Sensitivity and specificity in cancer detection have been found to be potentiated by use of immunoassay panels which include tissue-specific cancer biomarkers as well as cytokines and angiogenic factors. The ultrasensitive immunoassays revealed that patient to patient variations in the concentrations of individual biomarkers in blood can extend over many orders of magnitude (up to six) and that the distributions of biomarker concentrations over patient cohorts are non-Gaussian. New methods of data analysis which correlate abundances of multiple, different biomarkers have been developed to deal with such data sets. Sensitivity and specificity of about 95% have been achieved for blood-based detection of breast cancer in pilot studies on 250 patients and 95 controls. Pilot studies indicate that this methodology may also allow differentiation of malignant breast cancer from benign lesions and can provide similar sensitivity and specificity for other epithelial cancers such as prostate cancer, ovarian cancer and melanoma. The methods developed for selection, application, and evaluation of very high sensitivity biomarker panels are expected to have general relevance for diagnostic proteomics.

Keywords: breast cancer • proteomics • immunoassay • blood • biomarkers • multiphoton detection

Introduction

Breast cancer is the most common female cancer worldwide, with an incidence of 1.1 million new cases each year.¹ There are approximately 300–400 thousand deaths a year with the rate varying significantly between countries. Trials have repeatedly, and convincingly, confirmed that breast cancer is a progressive rather than a systematic disease. The stage at which

disease treatment is started has a significant impact on clinical outcome and progression of breast cancer can be prevented through early detection and effective early treatment. Early diagnosis, including mammographic screening, is a key factor in the control of breast cancer. In recent years, breast cancer mortality rates have declined as a result of earlier detection and more effective therapy.^{2,3} The five-year survival rate for breast cancer is as high as 97% if the cancer is small, of low grade, and has not spread to the lymph nodes.

Currently, diagnosis is by triple testing (clinical examination, imaging with mammography and ultrasonography, and biopsy, either by fine-needle cytology or core biopsy).⁴ There remains a need for more effective screening, especially in younger age groups where mammography is less sensitive. Better blood-based testing may aid in early diagnosis, may reduce the need for open biopsy and could provide new modalities for monitor-

* To whom correspondence should be addressed. E-mail: J.Godovac-Zimmermann@ucl.ac.uk.

[†] BioTraces Inc.

[‡] OncoTraces Ltd.

[§] Department of Pathology and Laboratory Medicine, University of Pennsylvania.

^{||} Department of Radiology, University of Pennsylvania.

[⊥] Department of Surgery, University College London.

[#] University of Pittsburgh Cancer Institute.

[^] Rayne Institute, University College London.

ing of therapy. In addition, post therapy monitoring of patients for recurrence is currently rudimentary and improved methods are required. Many serum biomarkers have been described for breast cancer, but no single biomarker has proven effective. The search for new biomarkers and the simultaneous use of multiple biomarkers is therefore an active area of diagnostics proteomics research.^{5,6}

The biology of other epithelial cancers is similar to breast cancer, even if gender, age and the speed of progression are quite different, and there are also unmet needs for better diagnosis of these cancers. For example, ovarian cancer is among the top three causes of cancer-related death.⁷ Between 1 and 2% of all women develop epithelial ovarian cancer. The paucity of symptoms and lack of reliable diagnostic modalities mean that most ovarian cancers are discovered at an advanced stage and the overall five year survival rate is about 20%. Prostate cancer is the most common form of cancer among men in the United States,⁸ and there is an increasing need for early diagnosis. The standard method for early detection of prostate cancer is screening for prostate specific antigen (PSA). This test, however, has limited specificity and sensitivity and is not specific, so that complementary biomarkers are needed.

Hundreds of proteins have been suggested as putative cancer markers. About 10% are higher abundance proteins and the rest are low abundance proteins. We define as low abundance proteins (LAPs) those proteins for which more than 20% of patients have abundance < 1 pg/mL. There are indications that the low abundance proteins are better biomarkers of cancer, but low abundance proteins could not be reliably measured with prior-art assays. We have solved the problem of assay sensitivity with new immunoassay-multiphoton-detection (IA/MPD) and Super-ELISA assays.^{9–11} The dramatically improved assay sensitivity permits use of a multiplicity of low abundance proteins as biomarkers, and permits better than 95% sensitivity (correct detection of cancer) and specificity (correct rejection of healthy) in diagnostic tests for epithelial cancers. Epithelial cancers also cause changes in blood concentrations of angiogenic and inflammatory factors,^{12–14} but these proteins (growth factors, cytokines, etc.) typically have very low abundance in blood. In the present paper, we show that immunoassay panels using low abundance proteins that include tissue-specific markers together with angiogenic and inflammatory markers can dramatically improve blood-based diagnosis of breast cancer. However, the extreme, non-Gaussian distribution of serum concentrations of the low abundance proteins over patient cohorts requires nonstandard methods of analysis in order to establish correlations between different biomarkers. Initial data for other epithelial cancers (prostate, ovarian, pancreatic) and other malignancies (melanoma) suggests that similar methods can dramatically improve blood-based diagnostic detection of these cancers. More generally, the present studies reveal new complexities in the selection and application of effective panels of biomarkers, emphasize the need for very high assay sensitivity, demonstrate correlations among different biomarkers and present new methods of data analysis for panels involving correlated biomarkers.

Materials and Methods

Healthy and Breast Cancer Samples. Two sets of serum samples from women with nontreated breast cancer (NT-breast cancer) were studied. One set of samples sets (80 woman samples) were acquired from Zeptomatrix, Buffalo, NY. The second (184 woman samples) was obtained from University of

Pennsylvania, Philadelphia, PA. Three sets of serum samples from healthy women were studied. The first set of 60 came primarily from young women (20–50 years) who were blood donors at the Hospital Charite, Berlin, Germany. The second set from 30 middle aged American women (age 40–60 years) and the third set from 30 older women (age 60–90 years) came from Zeptomatrix, Buffalo, NY. The full set of proteins in SET1 (see text) was measured for only 95 of the healthy samples.

Antibodies and Reagents. Antibodies came from commercial suppliers. For IA/MPD and Super-ELISA, ABs for TNF α , IL-6 and IL-8 were from CLB, Amsterdam, NL. ABs for PSA were from Alpha Diagnostics, Houston, Texas. ABs for VEGF were from R & D Inc., USA. Antibodies used in the Luminex experiments are described elsewhere.^{15,16} ¹²⁵I labeled streptavidin was obtained from Amersham, UK. Streptavidin-polyHRP was from CLB, Amsterdam, NL. The enzymatic substrate and red-stop were from Neopren, CA, USA.

Read-Out of Immunoassays. Three different methods of signal detection for immunoassays have been used. Immunoassay/multiphoton-detection (IA/MPD) methods used ¹²⁵I labeled streptavidin.^{9–11} MPD-Imager/96 have been constructed by BioTraces Inc., Herndon, VA, USA with a sensitivity of a few zeptomoles per pixel. These imagers are capable of simultaneous MPD measurements in 96 or 384 well formats for microtiter plates.⁹ This method gave limits of detection of about 1 femtogram/mL. This proprietary technology of BioTraces Inc. uses ¹²⁵I to achieve ultrahigh sensitivity in immunoassays. IA/MPD was used to establish very high sensitivity Super-ELISA protocols, which were then used for production measurements of all patient samples. Super-ELISA is similar to IA/MPD except that streptavidin-polyHRP(*n*), where *n* = 20, 40, or 80, and a colorimetric plate reader are used.⁹ Super-ELISA does not use radiolabels, although it is MPD-enabled in the sense that the background-suppression methods developed for IA/MPD have been used to achieve detection limits of 10–50 femtogram/mL depending on the protein. The detection sensitivity of Super-ELISA was adequate to measure serum abundances over virtually all patients for all biomarkers used in the present study (see text).

Luminex Measurements. Using fluorescent, microbead technology from LUMINEX Corporation, the group of Dr. A. Lokshin has developed assays for a wide range of potential biomarkers. These measurements are described elsewhere.^{15,16} The Luminex technique has an estimated limit of sensitivity of about 10 pg/mL.

Data Analysis Procedures. Because the distributions of protein abundances over the cohorts were found to be non-Gaussian, standard statistical packages could not be used to analyze the immunoassay data. A series of programs (300 000 lines of code) which have been written to carry out the data analyses are outlined here and will be described in more detail elsewhere (Drukier et al., in preparation).

Results

Ultrasensitive Immunoassays. Multiphoton-detection (MPD) is an ultrasensitive method for counting of single decay events for isotopes such as ¹²⁵I that decay by the electron capture mechanism.^{9–11} These isotopes can be counted at levels well below background radiation levels and, because single decays are counted, MPD detection is inherently linear over 8–9 orders of magnitude. Immunoassay-multiphoton-detection (IA/MPD) uses ¹²⁵I labeled streptavidin in sandwich formats to achieve low zeptomole (10⁻²¹ moles) sensitivity in immunoassays and

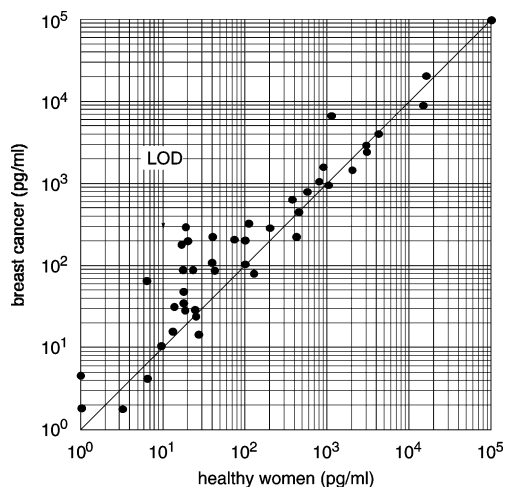


Figure 1. <Average> concentrations of proteins in serum for healthy women (HW) vs breast cancer (BC) patients as measured by Luminex. The proteins include 30 cytokines, chemokines, growth and angiogenic factors (IL-1 β , IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-12, IL-13, IL-15, IL-17, IL-18, TNF α , TNFRI, TNFRII, DR5, IFN γ , IFN β , GM-CSF, EGF, VEGF, G-CSF, bFGF, HGF, RANTES, MIP-1 α , MIP-1 β , MCP-1, MIG, and VEGF), 10 putative cancer antigens (CA-125, CA-19-9, CEA, CA-15-3, ErbB2, EGFR, KLK8, Fas, FasL, and Cyfra 21-1), and three matrix metalloproteinases (MMP-2,3, and 7). For non-Gaussian distributions of biomarker concentration with ranges of up to 5–6 logs over a cohort of samples (for both healthy and cancer cohorts, see text), arithmetic averages are highly sensitive to extreme outliers. For a cohort of 100 patients the average can be changed by a factor of 10 by a single outlier. The <average> concentrations shown in the figure are therefore calculated following exclusion of a small number of outliers at greater than 2 σ (for both healthy and cancer cohorts). Samples where concentration was below the limit of detection (LOD) were assigned a value of zero in calculating the averages. For some proteins this resulted in the <average> being below the LOD.

requires only about 1 nCi of radioactivity for immunoassays in 96 well microtiter plates.^{9–11} Highly effective background suppression is critical to obtaining extreme sensitivity with IA/MPD.

Super-ELISA is similar to IA/MPD except that streptavidin-polyHRP(n), where $n = 20, 40$ or 80 , and a colorimetric plate reader are used. Super-ELISA does not use radioactivity, but MPD was essential to optimize the measurement protocols and to reject particular sources of background. IA/MPD typically shows 100- to 1000-fold gains and Super-ELISA shows about 20- to 100-fold gains in sensitivity compared to conventional ELISA assays.⁹ This has enabled routine immunoassay of serum proteins that could not be measured by conventional ELISA assays.^{9–11} For the first time a wide selection of cytokines could be measured over complete patient cohorts in serum from healthy women or breast cancer patients (see below).

Selection of Biomarker Panels. A wide variety of proteins have been proposed as potential markers of breast cancer. Initial studies were therefore undertaken using the Luminex methodology to identify promising candidates for further studies. Because the sensitivity of the Luminex measurements is about 10 pg/mL, only proteins with relatively high abundance in serum can be analyzed. Figure 1 shows the <average> concentrations of 43 biomarkers in serum for healthy vs breast cancer patients (a small number of extreme outliers have been excluded from the <average>, see the caption to Figure 1). For the majority, the <average> concentrations are largely unchanged, suggesting low predictive power of these biomarkers for breast cancer. For some biomarkers, however, <average> concentrations over a cohort showed significant differences between healthy woman and breast cancer patients. A majority of these promising proteins were clustered near the limits of detection by Luminex (10–50 pg/mL, Figure 1) and could not be measured for all individuals. For example, PSA abundance in serum from females, which is a critical, new marker for breast cancer (see below), is too low for Luminex measurements. Experiments using the more sensitive IA/MPD and Super-ELISA methods, led to the selection of the two panels of biomarkers shown in Table 1 for more detailed investigations.

Typical Protein Distributions from Super-ELISA. Super-ELISA allowed all of the proteins in SET-1 to be measured for virtually all individuals. This coverage of the cohorts was confirmed by IA/MPD for the very few patients where serum concentrations were lower than the detection limits of the Super-ELISA assays (Figure 2A). The need for more sensitive assays is apparent from the limits of detection for conventional ELISA assays. For all biomarkers except VEGF, conventional ELISA assays can only measure the limited portions of the patient cohorts that show high abundances in serum (Figure 2A). Furthermore, even with the high sensitivity and accurate quantitation of Super-ELISA, no single biomarker showed a complete separation between healthy and breast cancer patients (Figure 2A).

Statistical analyses of such data often assume that biomarker abundances show Gaussian distributions around an average, but the data for the full cohorts clearly indicated that Gaussian distributions are not valid for the five markers measured by Super-ELISA (note the log scale in Figure 2A). Furthermore, there are appreciable changes in the averages as a function of age for both healthy and breast cancer patients (Figure 2B). The age-related changes might be a reflection of the menopausal status of the women, which was not available for the samples used.

Two-Dimensional Clustering. The above results indicated that the abundances of the individual proteins cannot be a reliable cancer diagnostic. If serum concentrations of these proteins are to be used, then correlations between multiple markers will be needed and due consideration of patient age

Table 1. Panels of Biomarkers Tested for Breast Cancer Diagnosis

panel	detection	cohort ^a	protein markers assayed		
			disease markers	inflammatory markers	angiogenic markers
SET 1	super-ELISA	160 BC, 95 HW	PSA	TNF α , IL-6	IL-8, VEGF
SET 2	luminex	53 BC, 60 HW	CA-125, CA-19.9, CK19 EGFR, MMP7	TNF α , IL-6, IL-15	IL-8, VEGF, EGF

^a Number of breast cancer patients (BC) and number of healthy patients (HW).

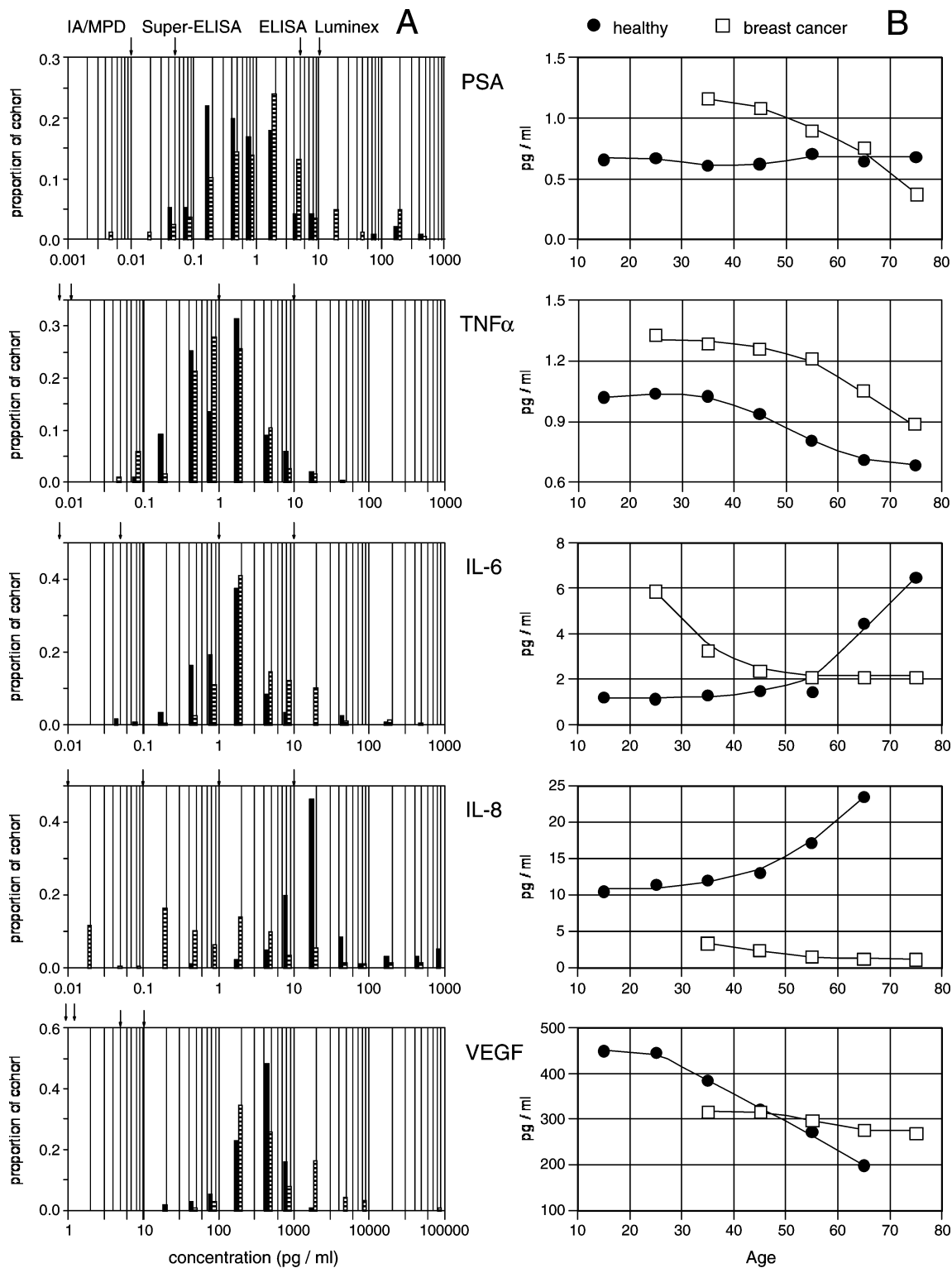


Figure 2. (A) Distributions of the serum concentrations of PSA, IL-6, IL-8, TNF α and VEGF for healthy women (solid bars) and breast cancer patients (hatched bars). The arrows at the top of each plot indicate the sensitivity of Luminex, conventional ELISA assays, Super-ELISA and IA/MPD. (B) Age dependence of the <average> concentrations in serum of PSA, IL-6, IL-8, TNF α , and VEGF for healthy women (HW) and breast cancer patients (BC).

will be required. Measurements of multiple protein concentrations inherently involve multidimensional data (5 and 11 markers imply 5 and 11 data dimensions for biomarker SETs 1 and 2 respectively). In principle such data sets can be analyzed in high-dimensional spaces to search for generalized ellipsoids

characteristic of healthy and breast cancer patients, but such analyses are difficult to visualize or to understand intuitively. We have therefore considered the protein concentrations in pairs, where correlations between different proteins can be readily observed. For N -biomarkers there are $N(N - 1)/2$

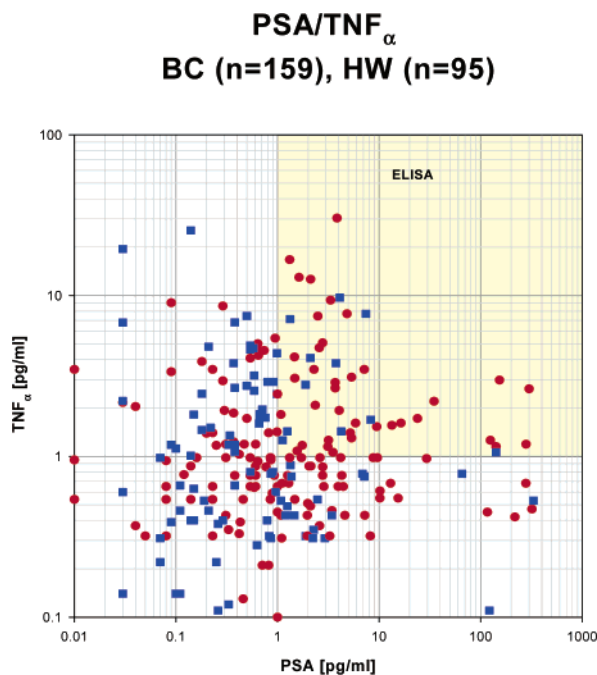


Figure 3. Correlation between the concentrations of PSA and TNF_α in serum of healthy women (blue squares) and breast cancer patients (red circles). The yellow shaded region indicates data points that are measurable with conventional ELISA assays.

different pairwise combinations of two biomarkers, which meant that 10 and 55 two-dimensional planes were examined for biomarker SETs 1 and 2.

Figure 3 shows the distribution of the pairwise concentrations of PSA and TNF_α. Several fundamental characteristics of the full data sets are already apparent. First, conventional ELISA or Luminex-based assays can only cover a small part of the two-dimensional plane, i.e. the very high sensitivity of Super-ELISA becomes essential for obtaining high cohort coverage when considering correlations between different biomarkers. Second, there are very wide variations in the protein concentrations in serum of different patients for both PSA (5 orders of magnitude) and TNF_α (3 orders of magnitude). Finally, in two dimensions the patient data is scattered over the plane and there is no complete separation between healthy and breast cancer patients. Patients with breast cancer show pairs of (PSA TNF_α) concentrations where both proteins are either high or low compared to the average concentrations.

Composite Scores from Two-dimensional Correlations. For the five proteins assayed by Super-ELISA, there are a total of 10 possible pairwise, two-dimensional correlation planes. Each biomarker in such a correlation plane has been age corrected and the concentration scales renormalized to better handle data points spread over several orders of magnitude. Each plane is clustered into regions in which either healthy woman or breast cancer patients are the predominant, but not exclusive, population. In areas of the planes with significant overlap between healthy woman and breast cancer patients, an “uncertain” region has also been assigned. Two such planes are shown in Figure 4A. Each of these two-dimensional planes can be regarded as an independent assay for detection of breast cancer. However, they are not “cutoff” assays – both very high and very low abundances of individual markers can be an indication of breast cancer.

Following clustering, all of the 2D planes were used to calculate a composite score. For each 2D plane, the data from a single patient was assigned a value of +1 (predominantly breast cancer region), 0 (uncertain region) or -1 (predominantly healthy region). For an individual patient, the composite score is the sum over the 10 planes. If all 10 biomarker pairs fall in regions in the 2D correlation planes assigned to predominantly cancer, then the patient would be assigned a composite score of +10. Following classification of all patients, the composite scores were normalized to the range -100% to +100% (essentially division by the number of biomarkers in the panel), which provides a common scale independent of the number of biomarkers. Ideal separability would imply that all healthy women would have a strongly negative composite score and all women with breast cancer would have a strongly positive composite score. As shown in Figure 5, most breast cancer samples have a composite score (CS) > 80% (152 out of 159 samples), whereas most samples from healthy women have a composite score < -40% (92 out of 95 samples). There is only a small overlap, involving a few percent of patients, in the range between -20% and +20%.

We have also explored alternative methods for assigning composite scores to patients. In particular, it is possible to use methods such as wavelet processing so that continuous values (not just ± 1) are assigned to data points in the two-dimensional planes (Figure 4B). This has advantages for recognition of interspersed regions of healthy and breast cancer in the planes (see the plane for PSA/VEGF in Figure 4B) and is discussed in more detail elsewhere.

Overall, the panel of markers permits detection of breast cancer in almost all women (> 98%) for which cancer has been previously detected by biopsy, mammography or MRI (BC cohort). Correlations of our blood-based assay with biopsy, MRI and mammography are remarkably good, and will be discussed elsewhere. Importantly, 95% of healthy women were properly identified. In early disease detection, the FDA requires the provision of so-called ROC curves. Typically ROC curves at 80% are acceptable and curves at 85–90% are regarded as excellent. The present results correspond to ROC values of about 95%, which is about 10% better than any prior-art breast cancer detection method, including mammography. In short, ultra-sensitive immunoassays that provide data for all patients in a cohort together with the use of correlations between appropriate pairs of biomarkers substantially improves classification of breast cancer.

How Many Biomarkers are Needed? Not all biomarker pairs are equally powerful in separating healthy and breast cancer patients (Figures 3,4). To take this into account, the influence of different planes on the specificity and sensitivity of the classification of the patients has been examined. For each plane, the predictive power is defined as the average of the probability of correct classification of healthy (specificity) or of breast cancer (sensitivity) patients. Table 2 shows the predictive powers of the different planes.

Essentially, there are three classes of such 2D correlation functions, with predictive power (PP) of PP < 70%, 70% < PP < 80% and PP > 80%. For example, in the case of breast cancer diagnostics, PSA/TNF_α has a medium predictive power of less than 70%. The regions of this plane which are populated by the BC cohort are closely interspersed with regions containing members of the HW cohort (Figure 3). IL-6/IL-8 represents a 2D correlation plane with about 80% predictive power (see Figure 4A). In this case, there are several regions populated by

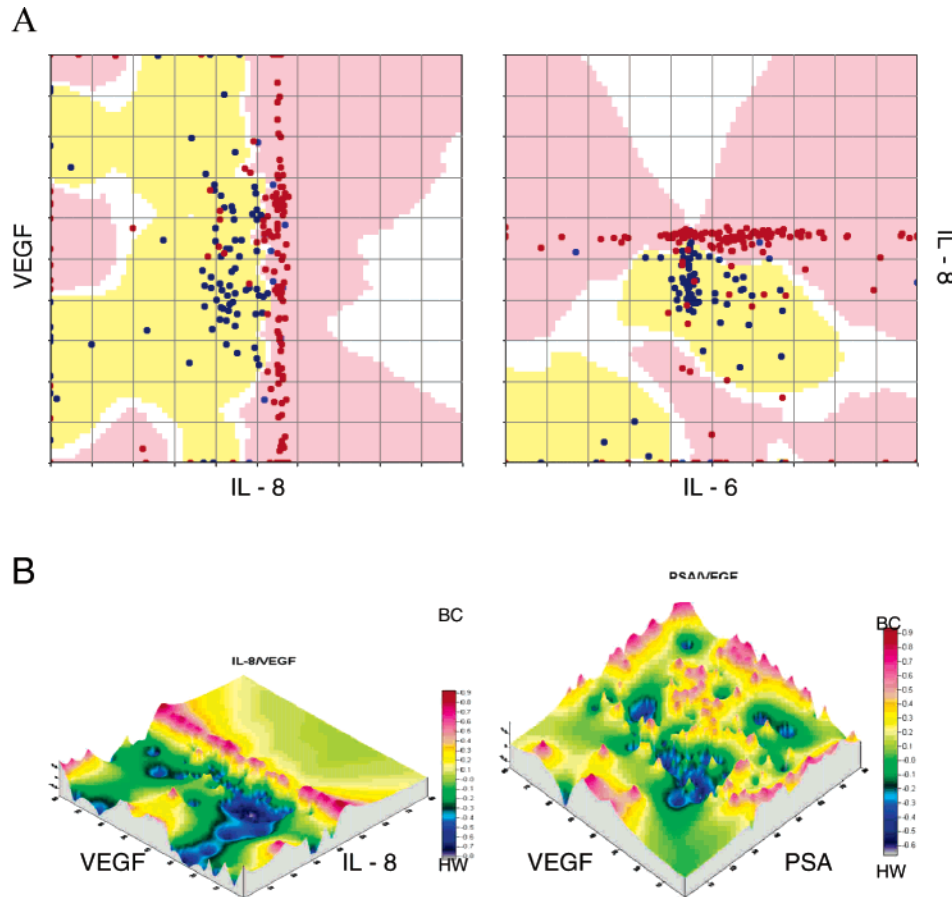


Figure 4. A. Age-adjusted and renormalized two-dimensional correlation planes for the protein pairs IL-6/IL-8 and IL-8/VEGF as measured by Super-ELISA. Healthy/cancer patients correspond to blue/red circles, respectively. The assigned regions are healthy (yellow), breast cancer (red) and uncertain (white). B. Wavelet processing of the 2D correlation data for the pairs PSA/VEGF (~85% predictive power) and (IL-8/VEGF (~95% predictive power). The behavior of each pair is typical for pairs with similar predictive power, e.g., at 85% predictive power, isolated regions for healthy patients are interspersed with cancer regions.

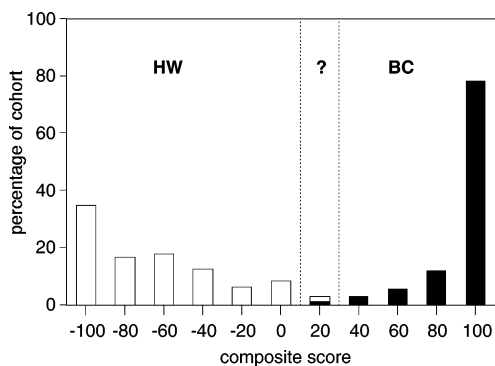


Figure 5. Percentage of healthy (HW) and breast cancer (BC) patients vs the composite score.

the HW cohort, but these regions are clustered and relatively well separated from the regions where the BC cohort dominates. Finally, the IL-8/VEGF plane has a predictive power of about 95%. This landscape shows overlapping regions for the HW cohort and two well-defined ridges for the BC cases (Figure 4).

The predictive powers and their associated landscapes suggest that some of the markers may be less important, e.g., TNF α does not appear in the very high classification for either specificity or sensitivity. For cost and efficiency of analysis it

Table 2. Predictive Power^a of Different Two-dimensional Correlation Planes

classification	2D correlations
low (<70%)	PSA/TNF α ; PSA/IL-6; TNF α /IL-6
high (70–80%)	PSA/IL-8; TNF α /IL-8; TNF α /VEGF
very high (> 80%)	PSA/VEGF; IL-6/IL-8; IL-6/VEGF; IL-8/VEGF

^a Predictive power is defined by: $PP = 100(1 - (N_f + N_p)/N)$, where N is the total number of data points, and N_f and N_p are the numbers of false negatives and false positives in the assignment of the data points to cancer or healthy.

is desirable to use the smallest number of biomarkers that is consistent with adequate levels of classification and we have therefore examined how the elimination of individual markers influences the final classifications (Table 2). PSA has a strong role in specifying the cancer as breast cancer and therefore cannot be removed (see below). Removing IL-6, IL-8 or VEGF leads to substantial decreases in specificity, i.e., the number of uncertain and/or false classifications increases considerably. The data suggests that the biomarker with the lowest predictive power is TNF α . When all measurements of TNF α are removed, there is a only few percent reduction in the specificity of the biomarker set. While this suggests that TNF α might be removed from the panel, we have observed that planes such as IL-6/TNF α help to detect differences in immune responses for different kinds of cancer (see below).

Tests of Robustness. Several tests of the robustness of the classification method have been made:¹ dependence on the precision of the data measured for individual patients,² dependence on the assay sensitivity, i.e., the proportion of the cohort for which data can be measured,³ dependence on training set, i.e., the inclusion or exclusion of data for individual patients, and⁴ dependence on the exact algorithms used for clustering of the data to regions of healthy, uncertain or breast cancer. First, the different assay methods have somewhat different precision. For Super-ELISA the measurement variance is 20% for < 0.1 pg/mL, 15% for 0.1–1 pg/mL and 5–10% for > 1 pg/mL. For ELISA it is 20% for 1–10 pg/mL and 10–15% for > 10 pg/mL. For Luminex it is 20–50% for 10–50 pg/mL and 10–30% for > 50 pg/mL. By propagating random errors of these magnitudes through the data sets, we found that the precision of the data has only a very moderate influence on the classifications. This is not surprising since the two-dimensional correlation planes cover many orders of magnitude and are very sparsely populated. Small changes in measured values do not move the points significantly in the 2D planes. Second, the specific clustering and renormalization algorithms also have only a limited influence on the classifications—here again the availability of multiple, sparsely populated 2D planes seems to be decisive. Third, the size of the data sets used in the present study was shown to be adequate in two ways. Random choice of different patients showed that for cohorts with about 40 BC and 40 HW samples, the resulting classifications were stable and gave specificity and selectivity > 95%. Furthermore, when the available data was randomly separated into a training cohort of 40 BC, HW patients and the resulting 2D planes were used to classify the remaining samples (validation cohort), specificity and selectivity was also > 95%. These analyses will be described in greater detail elsewhere (Drukier et al., in preparation).

However, the assay sensitivity (LOD) can strongly influence separability of the data into healthy and disease cohorts since insufficient sensitivity reduces the number of two-dimensional planes that can be used for classification of individual patients. The very different sensitivities of the four measurement techniques (LODs of: IA/MPD 1–5 fg/mL; Super-ELISA 10–50 fg/mL; ELISA 1–5 pg/mL; and Luminex, 10–50 pg/mL) results in marked differences in the proportion of patient samples that give correlated pairs of measurement values. Whereas both IA/MPD and SUPER-ELISA allow detection of over 95% of the 2D correlations in the patient cohorts, this is reduced to 50% for ELISA and 20% for Luminex (where only 40% of the IL-8, VEGF pairs remained). Concomitantly, the sensitivity the sensitivity and specificity were both reduced by about 5–10% (ELISA) or by 15% (Luminex). Typically this results in overlap of the composite scores for the different cohorts, producing data that is not interpretable (Figure 6).

Finally, we have used the Luminex data to assess whether increased numbers of biomarkers can compensate for reduced patient coverage due to lack of adequate measurement sensitivity. We focused on postmenopausal women with associated mammography tests (BC: $n = 53$; HW: $n = 60$). Only some of the biomarkers appeared likely to have useful predictive power (Figure 1). A previous analysis of the Luminex data by biostatisticians at the University of Pittsburgh achieved cancer vs healthy identification with sensitivity of about 75% and specificity of about 70%, but more than 16 biomarkers were needed (A. Lokshin, private communication). We selected a much smaller set of 11 biomarkers for detailed analysis with the more

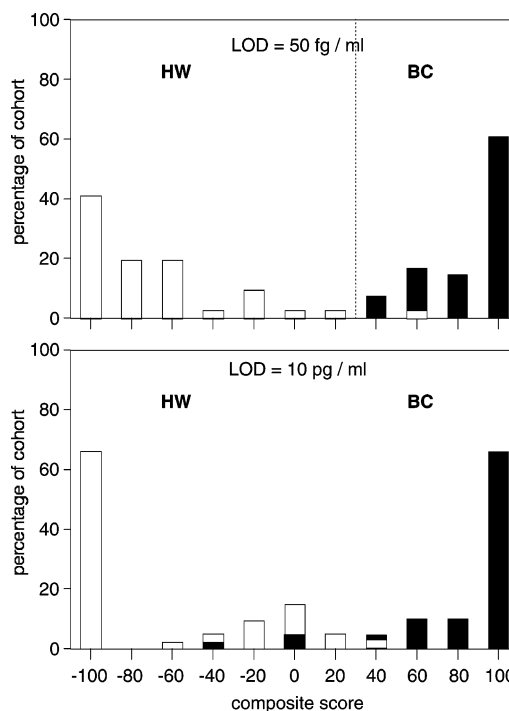


Figure 6. The effect of assay sensitivity on the correlation analysis, in particular the effect of missing data from unmeasurable serum levels. Top: Using Super-Elisa all data points can be measured. All 2-D planes can be clustered and have good quality that results in good separation between healthy and cancer cohorts. Bottom: Using a less sensitive assay, some data points are missing. The 2-D planes have lower quality or cannot be clustered. This leads to lack of separation of the cohorts.

sophisticated correlation-based algorithms. The panels of biomarkers used in the Luminex analyses include tissue markers, immune response markers and angiogenic factors (Table 1). Indeed, the Luminex markers include the cytokines and angiogenesis factors (TNF α , IL-6, IL-8, and VEGF) used in the Super-ELISA analyses. Table 3 shows the percentages of correct, uncertain and false classifications of patients using these data sets with the 2D correlation methods. The healthy and breast cancer patients could be largely separated and sensitivity and specificity better than 90% could be obtained (Table 3). Similar analyses have also been done with other types of cancer and are considered further below (see discussion). However, while these results look promising, the data sets are not large enough to be statistically certain and may not be stable with regard to random choice of patient subsets. For example, the 2D correlation of CA19–9/TNF α was found to have predictive power of 95%, but could only be measured for less than 10% of patients with Luminex (Figure 2).

Extensions of the Two-Dimensional Correlation Analysis. The ultimate goal of a measurement on an individual patient is a diagnosis, which may require decisions both about what type of cancer is present and whether the cancer is benign or malignant. Preliminary studies indicate that 2D correlation analyses may also be useful in these contexts. Figure 7 shows initial results on using an 11 biomarker analysis to distinguish benign/malignant breast cancer. Noteworthy is that no false assignments are made and that the number of patients classified as uncertain is small. Figure 8 shows very promising initial results on using an 11 biomarker analysis to distinguish

Table 3. Summary of Cancer Classification Using Two-Dimensional Correlation of Biomarkers

method/patient cohort	no. of biomarkers	correct identification	uncertain identification	false identification
Using Super-ELISA				
breast cancer (<i>n</i> = 264)	5	252/264 (95.4%)	5/264 (1.9%)	7/264 (2.7%)
healthy women (<i>n</i> = 95)		89/95 (93.6%)	4/95 (4.2%)	2/95 (2.2%)
prostate cancer (<i>n</i> = 32)	6	32/32 (100%)	0/32 (0.0%)	0/32 (0.0%)
BPH men (<i>n</i> = 60)		58/60 (96.7%)	0/60 (0.0%)	2/60 (3.3%)
prostate cancer (<i>n</i> = 32)	6	30/32 (93.8%)	0/32 (0.0%)	2/32 (6.2%)
prostatitis men (<i>n</i> = 19)		19/19 (100%)	0/19 (0.0%)	0/19 (0.0%)
Using Luminex				
breast cancer (<i>n</i> = 63)	11	63/63 (100%)	0/63 (0.0%)	0/63 (0.0%)
healthy women (<i>n</i> = 56)		56/56 (100%)	0/56 (0.0%)	0/56 (0.0%)
ovarian cancer (<i>n</i> = 51)	11	48/51 (94.1%)	2/51 (3.9%)	0/56 (0.0%)
healthy women (<i>n</i> = 56)		54/56 (96.4%)	2/56 (3.6%)	1/51 (0.0%)
pancreatic cancer (<i>n</i> = 72)	12	71/72 (98.6%)	0/72 (0.0%)	1/72 (1.4%)
healthy individuals (<i>n</i> = 53)		53/53 (100%)	0/53 (0.0%)	0/53 (0.0%)
melanoma (<i>n</i> = 172)	9	164/172 (95.3%)	7/172 (4.1%)	1/172 (0.6%)
healthy individuals (<i>n</i> = 100)		96/100 (96.0%)	3/100 (3.0%)	1/100 (1.0%)
All Cancers (<i>n</i> = 591)		567/591 (95.9%)	14/591 (2.4%)	10/591 (1.7%)
All Healthy (<i>n</i> = 383)		369/383 (96.3%)	9/383 (2.3%)	5/383 (1.4%)

breast and ovarian cancers. These data are considered further in the discussion.

Discussion

The starting concept for the present experiments was the notion that blood-based detection of cancer would be potentiated by the measurement of markers of systemic response in addition to tissue-specific disease markers. The hypothesis was that tissue/disease biomarkers permit high sensitivity and that

additional inflammatory and angiogenic biomarkers will improve the specificity. There is a substantial body of literature that motivates this idea, but it could only be tested reliably with the high sensitivity assays used here. For example, the immune response to cancer is an important defense against epithelial cancers, including breast cancer, and the present results support previous indications that measurement of the immune response helps in the detection of breast cancer. TNF α and IL-6 tend to be elevated in serum, as well as ascites, of women with breast cancer.¹⁷ It has previously been reported that serum levels of IL-8 are significantly elevated for epithelial cancers compared to normal controls¹⁸ and that levels of IL-6 are higher in epithelial cancers than in benign tumors.¹⁸ In addition to inflammation, cytokines such as IL-1, IL-6, and TNF α are also involved in some oncogenic processes.¹⁹ Considerable work on cell lines suggests the importance of pro-inflammatory cytokines in ovarian cancer,²⁰ but in-vivo studies have been much more difficult due to the very low levels of cytokines in blood. Thus, altered serum levels of cytokines are certainly associated with breast cancer, as they are with other cancers.

Similarly, angiogenesis, i.e. the development of new blood vessels, is an essential component of solid tumor growth and metastasis.²¹ Angiogenic factors are expressed by many tumors and activate host epithelium.²² The formation of the vascular stroma is crucial in the pathophysiology of malignancy and the onset of angiogenesis can mark a phase of rapid proliferation, local invasion and ultimately metastasis.²³ Vascularisation also allows greater accessibility for stray tumor cells.²⁴ The importance of angiogenesis factors in tumor growth thus suggests that they are potentially useful biomarkers for cancer. Vascular endothelial growth factor (VEGF) is a highly potent angiogenesis factor that plays a coordinated role in endothelial cell proliferation. Other important angiogenesis factors include IL-8, fibroblast growth factor (FGF) and the platelet derived endothelial growth factor (PD-EGF). Because these factors act locally at the sites of vessel formation, there is no particular barrier between them and blood. Local changes in their abundance are usually reflected in global changes in blood abundances, but these may be up or down. For example, the strong chemoattraction of a tumor can lead to local enhancement of IL-8 concentration at the tumor site, with a reduced concentration in blood (A. K. Drukier, unpublished results).

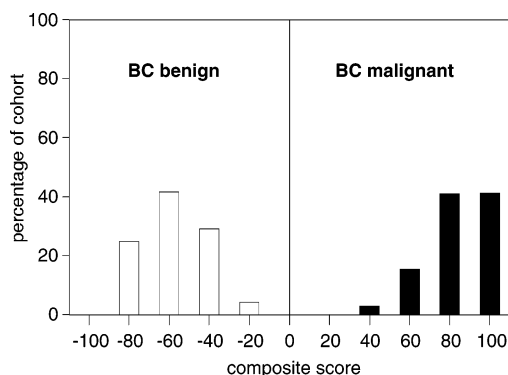


Figure 7. Percentage of patients with benign (open bars) or malignant (solid bars) breast cancer versus the composite score.

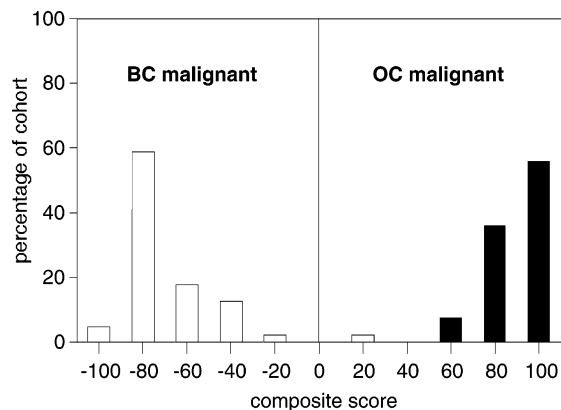


Figure 8. Percentage of patients with malignant breast cancer (open bars) or with malignant ovarian cancer (solid bars) versus the composite score.

Thus, angiogenic factors provide excellent biomarkers for tumor formation and this has been observed in the present studies—for breast cancer, the serum abundances of IL-8 and VEGF tend to decrease and increase respectively.²⁵

Prostate specific antigen (PSA) was initially believed to be specific to the prostate and has long been used as a marker of prostate cancer.^{26,27} In men, PSA is produced mainly, but not exclusively, by the prostate, and PSA blood levels are much higher in men than in women. Most PSA in the blood is bound to proteins that inhibit the proteolytic activity of PSA and high levels of the free enzyme are more suggestive of cancer than high total levels.²⁶ PSA has also been found in various female tissues and body fluids. Female breast, both normal and cancerous, produces PSA,²⁸ and this production is regulated by estrogen and progesterone. Preliminary data suggest that women with breast cancer may have a better prognosis if the level of PSA in tumors is high. A study which examined the prognostic value of PSA in a large cohort of US patients, using an assay with a detection limit of a few pg/mL, measured the level of PSA in tumor cytosolic extract of 953 women with preliminary breast cancer diagnosis.¹⁴ Detectable PSA, i.e., PSA levels > 5 pg/mL, were found to be significantly associated with smaller tumors, tumors with a small s-phase fraction, diploid tumors, younger age, and tumors with lower cellularity. Reduced risk for relapse and death remained statistically significant after accounting for other clinical and pathological variables. Because of the need for breast extracts and the limited sensitivity of the available immunoassays for PSA, these studies could not be extended to use of PSA in early diagnostics of breast cancer. We have developed assays for PSA with at least 100-fold improved sensitivity and were able to show that PSA abundances in blood provide a new tissue/disease biomarker for early breast cancer diagnosis.

Concurrent use of tissue-specific disease markers as well as markers of systemic response is strongly supported by the present results. However, the experiments presented new technical challenges, especially the requirement for more sensitive immunoassays to measure low abundance cytokines and angiogenic factors. Super-ELISA and IA/MPD assays have adequate sensitivity to monitor complete patient cohorts for PSA, IL-6, IL-8, TNF α , and VEGF, whereas the complete cohorts could be monitored only for VEGF with prior-art ELISA assays (Figure 2). The new data on PSA, IL-6, IL-8, and TNF α indicates that prior-art ELISA assays were monitoring only fractions of the complete range of abundances of these proteins in both healthy woman and cancer patients. In the case of PSA and IL-8, the sensitivity of prior ELISA assays failed to reach even the average abundance seen for the patient cohorts (Figure 2). Consequently, pairs of biomarkers such as PSA/IL-8 were only very poorly covered by the sensitivity achievable with prior-art ELISA and pairs of biomarkers such as IL-8/IL-6 that have been found to have very high predictive power for breast cancer (Table 2) were mostly inaccessible without the sensitivity of Super-ELISA.

At present, the Luminex data are mainly available for a small cohort of post-menopausal women. Much increased cohort sizes will be needed to clarify the preliminary indication that measurement of large numbers of biomarkers can at least partially offset incomplete cohort coverage. Alternatively, Luminex may be particularly valuable in a “directed-discovery” mode in which sizable groups of putative biomarkers are assessed for diagnostic potential. The greater patient coverage, but lower throughput of Super-ELISA will be useful in validating

biomarkers and essential for practical diagnosis of individual patients. The performance of the present blood-based assays for breast cancer strongly supports the use of supersensitive assays for cytokines and angiogenic factors to reduce the number of false positives in cancer diagnosis.

The importance of using correlated biomarkers is also supported by our other initial results. For example, the correlation methods show promising preliminary results for distinguishing benign/malignant breast cancers (Figure 7). This task normally requires biopsies that are costly, manpower-intensive and prone to stress/pain, i.e., a blood-based means of diagnosis would have major advantages. Similarly, the correlation methods show attractive initial results for differentiation between different types of cancers (Figure 8). Indeed, initial studies with smaller cohorts indicate that the two-dimensional correlation methods can provide high sensitivity and specificity for diagnosis of other cancers (Table 3). This suggests that the correlation methodology shows high promise for pan-cancer diagnostics. We are now exploring how many and which combinations of biomarkers would be needed for a biomarker panel capable of concurrently handling multiple types of epithelial cancers. Specific tissue/disease biomarkers are clearly needed, e.g., we have found that the combination of MIP-1a, MIP-1b, and MPA distinguishes melanoma from other cancers. On the other hand, initial indications are that there are changes in the serum abundances of cytokines and angiogenic factors that are distinctive for different cancers. For example, correlation planes for pairs of biomarkers such as IL-6/IL-8 and IL-8/VEGF are significantly different depending on the type of cancer and this could reduce the number of biomarkers that are needed for pan-cancer diagnostics. Alternatively, we have already optimized Super-ELISA assays for a wide range of other potential biomarkers that may be useful in parallel assays for different types of cancers.⁹

Finally, we suggest that the results obtained in the present work have important implications for diagnostic proteomics in general, especially in the selection, application and evaluation of biomarker panels. There is now abundant evidence that single, golden-bullet biomarkers of diseases are very unlikely and that multiple biomarkers are essential for diagnostic proteomics. The present studies show conclusively that in breast cancer patients the same biomarker may have either high or low serum concentrations relative to the averages over patient cohorts and that predictive power is considerably improved when the abundances are used in correlation with other biomarkers. This strongly suggests that it is not the presence of a given protein, but rather it's participation in particular networks of proteins and functions that has a high predictive power for a particular disease. It is not surprising that there are correlations between different biomarkers in the physiology of breast cancer and similar behavior is highly likely for other diseases. The present 2D correlation methods provide an effective way of exploiting such correlations for diagnostic purposes and avoid the clearly inadequate assumption of averages and Gaussian statistics to describe and interpret biomarker concentrations for patient cohorts. Selection of efficient biomarker panels is likely to require judicious combination of high-throughput screening methods, very high sensitivity assay methods and, perhaps increasingly, targeted exploitation of knowledge about physiology. Present experience suggests a number of features that should be considered in designing effective panels of biomarkers.¹ Low abundance

proteins can be excellent biomarkers and may be more informative than high abundance proteins.² Combination of tissue/disease specific markers with markers of systemic response may be particularly effective.³ Validation of biomarkers is greatly facilitated if the biomarker is measurable over full patient cohorts.⁴ The panel should include combinations of biomarkers with clear correlations between different biomarkers.⁵ Biomarker combinations are especially sensitive to cohort coverage and efficient diagnosis of individual patients requires very high cohort coverage for the individual biomarkers. In addition, such panels need to be shown to be robust and disease specific. For this purpose new methods of data analysis will be needed. We believe that the methods and results obtained in the present work represent an important step toward achieving accurate, correlated measurements of the abundances of multiple, informative biomarkers that will be essential to moving beyond “diagnostic proteomics” and toward “functional proteomics” that is closely related to physiology and pathology.

Abbreviations. MPD, multiphoton-detection; IA/MPD, immunoassay with multiphoton- detection read-out; Ab, antibody; polyHRP: poly horseradish peroxidase; TNF α , tumor necrosis factor α ; TNFRI, II, TNF Receptors 1 and 2; IFN γ , interferon gamma; EGF, epithelial growth factor; VEGF, vascular endothelial growth factor; EGFR, epithelial growth factor receptor; CK19, cytokeratine 19 kDa; MMP-2, -3, -7: matrix metalloproteinase 2.3.7; PSA, prostate specific antigen; FDA, Federal Drug Administration; ROC, Receiver Operating Characteristics.

Acknowledgment. A.K.D. appreciates many discussions with Drs. S. Stass, G. Miklos, D. Coffey, Ch. Michejda, L. Gold, M. Brody, and M. Elbaum. We thank Dr. O. Kleiner, London for help during manuscript preparation. This project was made possible by NSF SBIR Phase I and II support.

References

- (1) Parkin, D. M.; Bray, F.; Ferlay, J.; Pisani, P. *Cancer J. Clin.* **2005**, *55*, 74–108.
- (2) Coleman, M. *Lancet* **2000**, *356*, 590.
- (3) Peto, R.; Boreham, J.; Clarke, M.; Davies, C.; Beral, V. *Lancet* **2000**, *355*, 1822.
- (4) Dixon, J. M. In *ABC Breast Diseases*; Dixon, J. M., Ed.; Blackwell Publishing: Cambridge, 2006, p 5.

- (5) Diamandis, E. F. *J. Natl. Cancer Inst.* **2004**, *96*, 353–356.
- (6) Robbins, R. J.; Villaneuva, J.; Tempst, P. *J. Clin. Oncol.* **2005**, *23*, 4835–4837.
- (7) Skates, S. J.; Horick, N.; Yu, Y.; Xu, F. J.; Berchuck, A.; Havrilesky, L. J.; de Bruijn, H. W.; van der Zee, A. G.; Woolas, R. P.; Jacobs, I. J.; Zhang, Z.; Bast, R. C., Jr. *J. Clin. Oncol.* **2004**, *22*, 4059–4066.
- (8) Adam, B. L.; Qu, Y.; Davis, J. W.; Ward, M. D.; Clements, M. A.; Cazares, L. H.; Semmes, O. J.; Schellhammer, P. F.; Yasui, Y.; Feng, Z.; Wright, G. L., Jr. *Cancer Res.* **2002**, *62*, 3609–3614.
- (9) Drukier, A. K.; Ossetrova, N.; Schors, E.; Brown, L. R.; Tomaszewski, J.; Sainsbury, R.; Godovac-Zimmermann, J. *J. Proteome Res.* **2005**, *4*, 2375–2378.
- (10) Kleiner, O.; Price, D. A.; Ossetrova, N.; Osetrov, S.; Volkovitsky, P.; Drukier, A. K.; Godovac-Zimmermann, J. *Proteomics* **2005**, *5*, 2322–2330.
- (11) Sainsbury, R.; Godovac-Zimmermann, J.; Drukier, A. K.; Tomaszewski, J.; Kleiner, O. *Breast Canc. Res. Tr.* **2004**, *88*, S220–221 Suppl. 1.
- (12) Kozłowski, L.; Zakrzewska, I.; Tokajuk, P.; Wojtukiewicz, M. Z. *Rocz. Akad. Med. Białymst.* **2003**, *48*, 82–84.
- (13) Rao, V. S.; Dyer, C. E.; Jameel, J. K.; Drew, P. J.; Greenman, J. *Oncol. Rep.* **2006**, *15*, 179–185.
- (14) Wang, Y.; Yang, J.; Gao, Y.; Du, Y. R.; Bao, J. Y.; Niu, W. Y.; Yao, Z. *Cell. Mol. Immunol.* **2005**, *2*, 365–372.
- (15) Lokshin, A. E.; Winans, M.; Landsittel, D.; Marrangoni, A. M.; Velikokhatnaya, L.; Modugno, F.; Nolen, B. M.; Gorelik, E. *Gynecol. Oncol.* **2006**, on line.
- (16) Gorelik, E.; Landsittel, D. P.; Marrangoni, A. M.; Modugno, F.; Velikokhatnaya, L.; Winans, M. T.; Bigbee, W. L.; Herberman, R. B.; Lokshin, A. E. *Cancer Epidemiol. Biomarkers Prev.* **2005**, *14*, 981–987.
- (17) Kozłowski, L.; Zakrzewska, I.; Tokajuk, P.; Wojtukiewicz, M. Z. *Rocz. Akad. Med. Białymstoku.* **2003**, *48*, 82–84.
- (18) Benoy, I. H.; Salgado, R.; Van Dam, P.; Geboers, K.; Van Marck, E.; Scharpe, S.; Vermeulen, P. B.; Dirix, L. Y. *Clin. Cancer Res.* **2004**, *10*, 7157–7162.
- (19) Darai, E.; Detchev, R.; Hugol, D.; Quang, N. T. *Hum. Reprod.* **2003**, *18*, 1681–1685.
- (20) Nash, M. A.; Ferrandina, G.; Gordinier, M.; Loercher, A.; Freedman, R. S. *Endocr. Relat. Cancer.* **1999**, *93*–107.
- (21) Folkman, J. *Annu. Rev. Med.* **2006**, *57*, 1–18.
- (22) Folkman, J.; Klagsbrun, M. *Science* **1987**, *235*, 442–447.
- (23) Ellis, L. M.; Fidler, I. J. *Eur. J. Cancer.* **1996**, *3*, 2451–2460.
- (24) Hanahan, D.; Folkman, J. *Cell* **1996**, *86*, 353–364.
- (25) Lewis, C. E.; Leek, R.; Harris, A.; McGee, J. O. J. *Leukoc. Biol.* **1995**, *57*, 747–751.
- (26) Duffy, M. J. *Ann. Clin. Biochem.* **1996**, *33*, 511–519.
- (27) Smith, R. A.; Cokkinides, V.; Eyre, H. *CA Cancer J. Clin.* **2005**, *55*, 31–44.
- (28) Yu, H.; Berkel, H. J. *La State Med. Soc.* **1999**, *151*, 209–213.

PR0600834