Disease Markers 19 (2003,2004) 229–238
IOS Press

SELDI-TOF serum profiling for prognostic and diagnostic classification of breast cancers

Christine Laronga[b], Stephen Becker[b], Patrice Watson[c], Betsy Gregory[a], Lisa Cazares[a], Henry Lynch[c], Roger R. Perry[b], George L. Wright Jr.[a], Richard R. Drake[a] and O. John Semmes[a,*]
[a]*Departments of Microbiology and Molecular Cell Biology, Eastern Virginia Medical School, Norfolk, Virginia and the Virginia Prostate Center, Norfolk, VA, USA*
[b]*Department of Surgery, Eastern Virginia Medical School, Norfolk, Virginia and the Virginia Prostate Center, Norfolk, VA, USA*
[c]*Department of Preventive Medicine, Creighton University, Omaha, NE, USA*

Abstract. Surface enhanced laser desorption/ionization (SELDI) time-of-flight mass spectrometry has emerged as a successful tool for serum based detection and differentiation of many cancer types, including breast cancers. In this study, we have applied the SELDI technology to evaluate three potential applications that could extend the effectiveness of established procedures and biomarkers used for prognostication of breast cancers. Paired serum samples obtained from women with breast cancers prior to surgery and post-surgery (6–9 mos.) were examined. In 14/16 post-treatment patients, serum protein profiles could be used to distinguish these samples from the pre-treatment cancer samples. When compared to serum samples from normal healthy women, 11 of these post-treatment samples retained global protein profiles not found in healthy women, including five low-mass proteins that remained elevated in both pre-treatment and post-treatment serum groups. In another pilot study, serum profiles were compared for a group of 30 women who were known BRCA-1 mutation carriers, half of whom subsequently developed breast cancer within three years of the sample procurement. SELDI protein profiling accurately classified 13/15 women with BRCA-1 breast cancers from the 15 non-cancer BRCA-1 carriers. Additionally, the ability of SELDI to distinguish between the serum profiles from sentinel lymph node positive and sentinel lymph node negative patients was evaluated. In sentinel lymph node positive samples, 22/27 samples were correctly classified, in comparison to the correct classification of 55/71 sentinel lymph node negative samples. These initial results indicate the utility of protein profiling approaches for developing new diagnostic and prognostic assays for breast cancers.

Keywords: Breast cancer, protein profiling, serum

1. Introduction

Breast cancer remains the most common cancer affecting women today. However, although there are over 200,000 new cases diagnosed each year, there is still no blood test currently available for diagnostic or prognostic detection. Identification of effective biomarkers capable of serving as a screening tool for the diagnosis of breast cancer, or for prognostic indicators of disease recurrence or therapeutic response, have been the focus of intense study [3,5,8,16]. Several useful biomarkers have been identified, like Her2/neu for prognostic determinations and Ca 27.29 for assessment of disease recurrence/burden [3,8], yet mammography still remains the gold standard to which all new tests must be compared [5]. Because breast cancer presents as a spectrum of different histologies (ductal, lobular, mucinous, papillary, medullary, colloid, tubular) and stages of disease, the identification and simultaneous analysis of a panel of biomarkers may provide a greater potential for improving molecular diagnostic approaches for early

*Corresponding author: O. John Semmes, Ph.D., Department of Microbiology and Molecular Cell Biology, Eastern Virginia Medical School, 700 West Olney Road, Norfolk, VA 23507, USA. Tel.: +1 757 446 5904; Fax: +1 757 624 2255; E-mail: semmesoj@evms.edu.

ISSN 0278-0240/03,04/$17.00

detection and prognosis of breast cancers.

Among the many promising strategies in clinical proteomics, protein expression profiling using surface enhanced laser desorption/ionization (SELDI) time-of-flight mass spectrometry has emerged as a successful tool for serum based detection and differentiation of many cancer types [1,2,4,11–13,18,21–24], including breast [11,21]. This protein profiling platform has medium-high throughput capability, requires small amounts of sample material, can effectively resolve low-mass proteins (2–20 kd), and is highly reproducible when used with robotic sample handling platforms. When coupled with "learning" type classification algorithms, this approach can distinguish cancer from non-cancerous states with high accuracy [1,11–15,25].

In this study, we have applied the SELDI technology to evaluate three potential applications that could extend the effectiveness of established procedures and biomarkers used for prognostication of breast cancers. In the first set, paired serum samples from women with early stage breast cancers obtained prior to surgery and post-treatment were examined. In the second set, serum profiles were compared from a group of women who were known BRCA-1 mutation carriers who were not diagnosed with breast cancer at the time of sample procurement. Half of these women subsequently developed cancer within three years, and their serum protein profiles are compared with the cohort of BRCA-1 carriers who did not develop cancer. In the third set, the ability of SELDI to distinguish between the serum profiles from sentinel lymph node positive and negative patients was evaluated. For each group, promising results were determined, indicating multiple new approaches that can be taken to develop new diagnostic and prognostic assays for breast cancers.

2. Materials and methods

2.1. Serum samples

Female patients with an abnormal mammogram or clinical breast exam whom required an operative surgical biopsy were eligible to participate in this study for breast cancer detection. They were enrolled through the Division of Surgical Oncology, Department of Surgery at Eastern Virginia Medical School after signing an Institutional Review Board approved informed consent. Pre-treatment serum samples were obtained at the same time as pre-operative laboratory studies and therefore were separated in time by at least a week from any clinical breast examination, diagnostic imaging, or diagnostic biopsy (FNA or core). The samples were then retrospectively categorized as benign or cancerous upon pathological confirmation. Normal serum samples were randomly collected from healthy volunteers during the same time period (December 2001-May 2002) as the cancer specimens were obtained and were not drawn in concert with any clinical breast evaluation or diagnostic imaging. Blood samples were collected by venipuncture into a 10cc SST vacutainer tube and allowed to clot at 4°C for 30 minutes. Coagulated blood was spun at 3000 rpm for 10 minutes, and a portion of the serum was immediately aliquoted and frozen for storage at –80°C. All clinical information, including age, race, menopausal status, personal breast history, histological diagnosis, clinical and surgical stage, tumor type, tumor grade and receptor status were recorded in the breast study database. No attempts were made to separate the samples based on co-morbid conditions, medications taken, or timing of blood draw with respect to menstrual cycle.

In a pre- and post-surgery pilot study, 16 paired serum samples were obtained from women newly diagnosed with invasive ductal breast cancer. The first specimen was retrieved prior to initiation of any treatment. Three women received chemotherapy prior to surgery. The second sample was drawn 6–12 months post-surgery (5 post-surgery alone; 2 post-surgery and radiation; 5 post-surgery and chemotherapy; and 4 post-surgery, chemotherapy and radiation). Eight patients were African-American and 8 were Caucasian. Only patients with stage I–III were considered eligible for the study. Seven of the 9 cases of stage I disease were in the Caucasian cohort, while the African-American cohort was composed of the higher stage disease (stage III $=$ 2, stage II $= 5$). For the entire group, the mean age at the time of diagnosis was 55 yrs (range 41–74yrs). The mean age for the AA cohort was 51 yrs, and 59 yrs in the Caucasian cohort. All treatments had been finished for at least one month prior to the post-treatment serum draw, except in one patient that had just finished her last cycle of chemotherapy 2 weeks before. The categories of breast cancer included invasive ductal ($n = 15$) and invasive lobular ($n = 1$). Ductal carcinoma in situ (DCIS), without an invasive component, inflammatory breast cancer and breast sarcoma were excluded from the study.

For the BRCA-1 pilot study, thirty age-matched (± 3 years) baseline serum specimens were identified from a collection of sera obtained during genetic testing at the Hereditary Cancer Institute at Creighton University

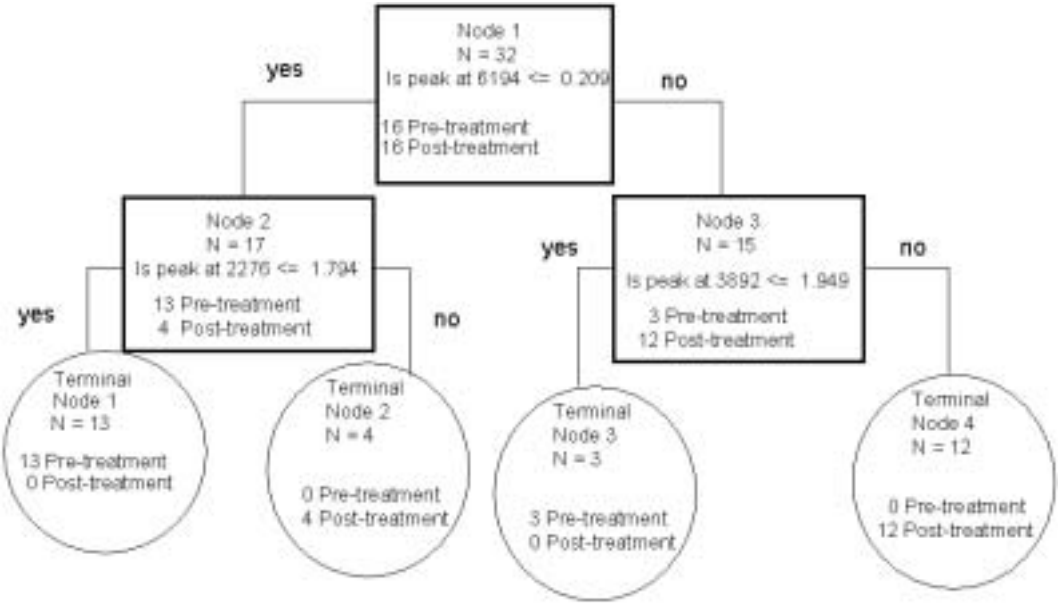

Fig. 1. Diagram of classification tree for pre-treatment and post-treatment samples. The squares are the primary nodes and the circles indicate terminal nodes. The mass value in the root nodes is followed by = the intensity value.

from 1987–2001. All were from women with germ-line BRCA-1 mutations who were not known to have cancer at baseline. The BRCA-1 women were followed until the development of breast cancer or for 7 disease-free years and divided into 2 cohorts pending development of breast cancer (BRCA-1 Ca or Carrier). All BRCA-1 Ca ($n = 15$) were diagnosed within 3 years after the baseline serum sample, while the carrier group ($n = 15$) remained cancer free for 7 years.

For the sentinel lymph node pilot study, serum samples from 98 women that underwent sentinel lymph node biopsies during surgery for invasive breast cancers samples were collected: sentinel lymph node positive (SLN+, $n = 27$) and sentinel lymph node negative (SLN−, $n = 71$). Twenty-two of the seventy-one SLN− cohort were African-American, 46 were Caucasian and 3 were "other". In the SLN+ cohort, 10 of the patients were African-American and 17 were Caucasian. The mean age at the time of diagnosis was 58 yrs in the SLN+ patients and 56 yrs in the SLN− cohort. Twenty-six of the cases in the SLN+ cohort were invasive ductal carcinoma with one case of invasive lobular carcinoma. Twenty-six of the cases in the SLN+ cohort were invasive ductal carcinoma with one case of invasive lobular carcinoma, and 8 of the 27 SLN+ patients were Her2/neu positive. Seventy of the 71 SLN− cases were invasive ductal carcinoma and there was one case of invasive lobular carcinoma.

2.2. SELDI processing of serum samples

Serum samples were processed for SELDI analysis as previously described [1,21]. Briefly, 20 μl of serum is pre-treated with 8 M urea, 1% CHAPS and vortexed for 10 minutes at 4°C. A further dilution is made in 1 M urea, 0.125% CHAPS and PBS. Diluted serum is then added to the protein chips with the aid of a bioprocessor. Protein chips are then incubated at room temperature followed by washes of PBS and water. Arrays were allowed to air dry and a saturated solution of sinapinic acid in 50% (v/v) acetonitrile, 0.5% (v/v) trifluoroacetic acid was added to each spot. The protein chip arrays were analyzed using the SELDI ProteinChip System (PBS-II, Ciphergen Biosystems, Fremont, CA). Spectra were collected by the accumulation of 192 shots at laser intensity 220 in positive ionization mode. The protein masses were calibrated externally using purified peptide standards. All samples were processed in duplicate and were randomized on the chips.

2.3. SELDI data analysis

For each cohort, protein peaks were labeled and their intensities normalized for total ion current to account for variation in ionization efficiencies, using SELDI software (Version 3.1). Peak clustering was performed using the Biomarker Wizard software (Ci-

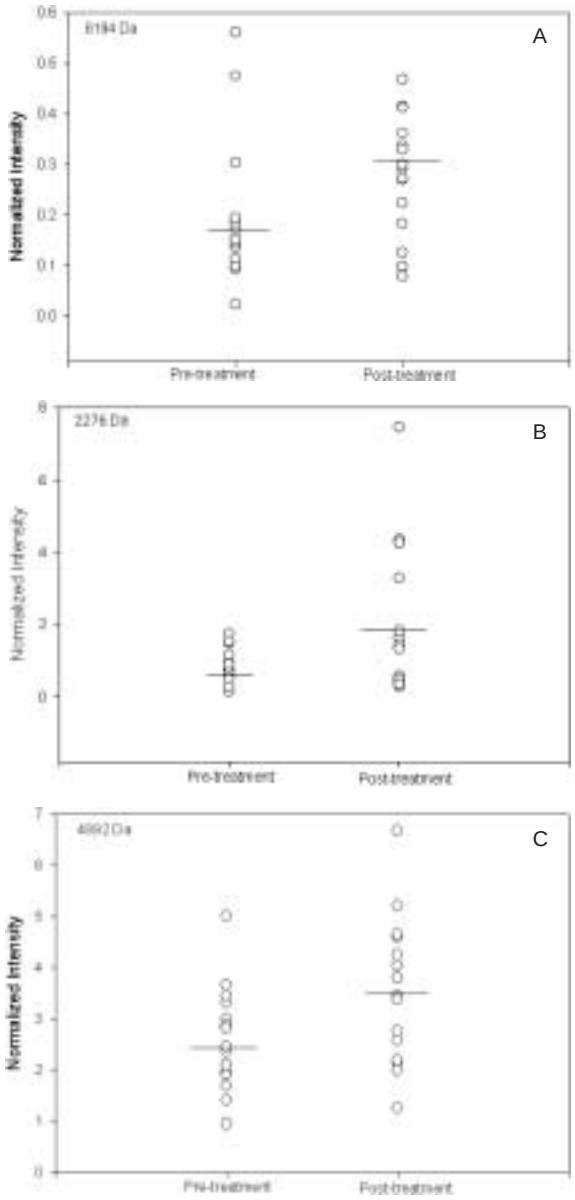

Fig. 2. Expression level of the A. 6194 Da, B. 2276 Da, and C. 4892 Da proteins used in the decision tree classification for distinguishing sera profiles of pre-treatment patients compared with sera from post-treatment patients. —-, mean normalized intensity; O, values of individual patients.

phergen Biosystems Inc.) with the following settings: for the IMAC surface: signal/noise (first pass): 3, minimum peak threshold: 5%, cluster mass window: 0.2%, signal/noise (second pass): 2. The samples were analyzed for peaks only within the range of 1.5–30 kDa. For the SAX surface: signal/noise (first pass): 3, minimum peak threshold: 5%, cluster mass window: 0.3%, signal/noise (second pass): 2. The spectra from the SAX surface was analyzed within the 30–100 kDa weight range. Peak mass and intensity were exported to an Excel file, and the peak intensities from each duplicate spectra were averaged. Pattern recognition and sample classification were performed using the Biomarker Pattern Software (Ciphergen Biosystems Inc.) as described earlier [21]. In brief, multiple decision trees were initially generated using all the peaks as variables and increasing the "cost" i.e. misclassification of a sample by 0.1. The peaks that formed the main splitters of the tree(s) with the highest prediction rates in the cross-validation analysis were then selected to examine the mean intensities of each peak in the tested samples, represented as scatter plots. Additionally, the number of samples used to test the tree during cross-validation was increased from 10 to 20. While choosing a smaller number of samples i.e. 5, to test the tree can result in over-fitting the data, increasing the number of samples used to test the tree is an acceptable means of insuring the validity of the splitters in a small sample population.

3. Results

3.1. SELDI-TOF MS profiling of pre- and post-treatment samples

Serum samples were obtained from 31 women, 15 normal, and 16 matched pre and post cancer treatment. As described in detail in the Materials and methods, there were 9 stage 1, 5 stage 2, and 2 stage 3 subjects. Each serum sample was applied robotically in duplicate to IMAC-Cu and SAX protein chips for SELDI-TOF analysis, using established procedures [21]. In brief, for the processing of each resulting protein profile, each protein peak was labeled and its intensity was normalized for total ion current to account for variation in ionization efficiencies. Peak clustering was performed using optimal settings that provide a 5% minimum peak threshold, 0.2% mass window and 0.1%–0.3% signal/noise determination. Peak mass and intensity were exported to an Excel file, and the peak intensities from each duplicate spectra were averaged. Pattern recognition and sample classification were performed using the Biomarker Pattern Software (Ciphergen Biosystems Inc), in which multiple decision trees are initially generated using all the peaks as variables. During the analysis, a pruning step occurs in which branches are removed and the cost of the removal determined to establish a minimal tree size. Second, the

Table 1
Results from cross-validation analysis of the decision classification trees for each sample set

Condition	Sensitivity	Specificity
Pre- vs. Post-treatment	87% (14/16)	87% (14/16)
Pre-treatment vs. normal	93% (14/15)	87% (13/15)
Post-treatment vs. normal	93% (13/14)	73% (11/15)
BRCA-1 vs. BRCA-1 Ca	87% (13/15)	87% (13/15)
SLN (+) vs. SLN (−)	81% (22/27)	77% (55/71)

tree is subjected to cross-validation, which separates the data into partitions that are individually evaluated against the remaining data set. The peaks that formed the main splitters of the tree(s) with the highest prediction rates in the cross-validation analysis were then selected to examine the mean intensities of each peak in the tested samples. This same process was applied for each sample set analyzed in this study.

In Fig. 1, a representative decision tree is shown in which 12/16 pretreatment samples and 14/16 post-treatment samples were accurately classified, with a cross-validation of 75% sensitivity and 87% specificity (Table 1). A scatter plot of the differences in the peak intensities of three differentially expressed peaks present in the pre- and post-treatment samples is presented in Fig. 2. These peaks represent potential biomarker proteins that will be targeted in future follow-up studies with greater sample numbers. The pre- and post-treatment sample SELDI-profiles were further evaluated against the serum profiles obtained from 15 normal healthy volunteer female subjects. In comparison of the pre-treatment cancer patients relative to the normal patients, the results of the cross-validation were 87%/93% sensitivity and specificity. This is consistent with the results obtained for a previous SELDI profiling study of serum samples that distinguished breast cancer patients from healthy normals [21]. To our knowledge, a question that has not been previously evaluated is how the profiles of the post-treatment samples compare with the normal healthy subjects. Does successful treatment for breast cancer restore a "normal healthy" serum protein profile, or are there markers present indicative of treatment outcome and residual disease? In the comparison of post-treatment samples with normal healthy samples, only 1 of the post-treatment samples was misclassified, and 4 normal healthy samples misclassified in the cross-validation analysis (Table 1). This yielded a final sensitivity of 93% to separate post-treatment samples from normals (73% specificity).

While it is feasible that post-treatment samples should be similar to normal healthy profiles, it was obvious that there were clear proteomic differences between the two sample sets. After comparison of each of the common significant peak values present in the three analyses (N vs. pre; N vs. post; pre vs. post), 5 low mass peptides remained elevated in the pre- and post-treatment samples relative to the healthy subjects (*m/z* 2146, 3161, 3686, 3820 and 6679). In Fig. 3, representative SELDI peak profiles from each of the patient sub-classes is presented for the *m/z* 6679 protein. Similar representative profiles of under-expression of the indicated peaks in normal health individuals can be generated for each of the other four proteins (data not shown). These peaks represent potential prognostic markers for further evaluation of treatment outcomes and potential disease recurrence.

3.2. SELDI-TOF MS profiling of BRCA-1 mutation carriers

BRCA-1 mutations predispose women to early onset breast cancer but 20% of mutation carriers will never develop breast cancer [9]. Our objective was to determine if SELDI-TOF analysis could distinguish women with BRCA-1 mutations that will develop clinically diagnosable breast cancer in the near future (BRCA-1 Ca) from those who will not (Carrier). Analysis of SELDI-TOF spectra revealed 107 differentially expressed protein peaks ($p < 0.05$) in the 1.5 to 20 kDa range between the BRCA-1 Ca and Carriers. Upon cross-validation of a representative decision tree (Fig. 4), 13/15 BRCA-1 Ca subjects were correctly identified with a sensitivity/specificity of 87%/87% (Table 1). Several potential low-mass biomarkers were identified, including one peak at 5.9 kDa that was a primary identifier. In further studies, the spectra of BRCA-1 carriers closely resembled spectra from healthy normal serum samples, and both were distinguishable from BRCA-1 Ca and sporadic breast cancers (data not shown). Further analysis of these different groups with other chip surfaces and greater sample numbers are in progress.

3.3. SELDI-TOF MS profiling of sentinel lymph node positive and negative samples

Serum samples from 98 women that underwent sentinel lymph node biopsies during surgery for invasive breast cancers samples were evaluated by SELDI-TOF to evaluate whether sentinel lymph node positive (SLN+, $n = 27$) and sentinel lymph node negative (SLN−, $n = 71$) samples could be distinguished by their proteomic patterns. Figure 5 shows a decision

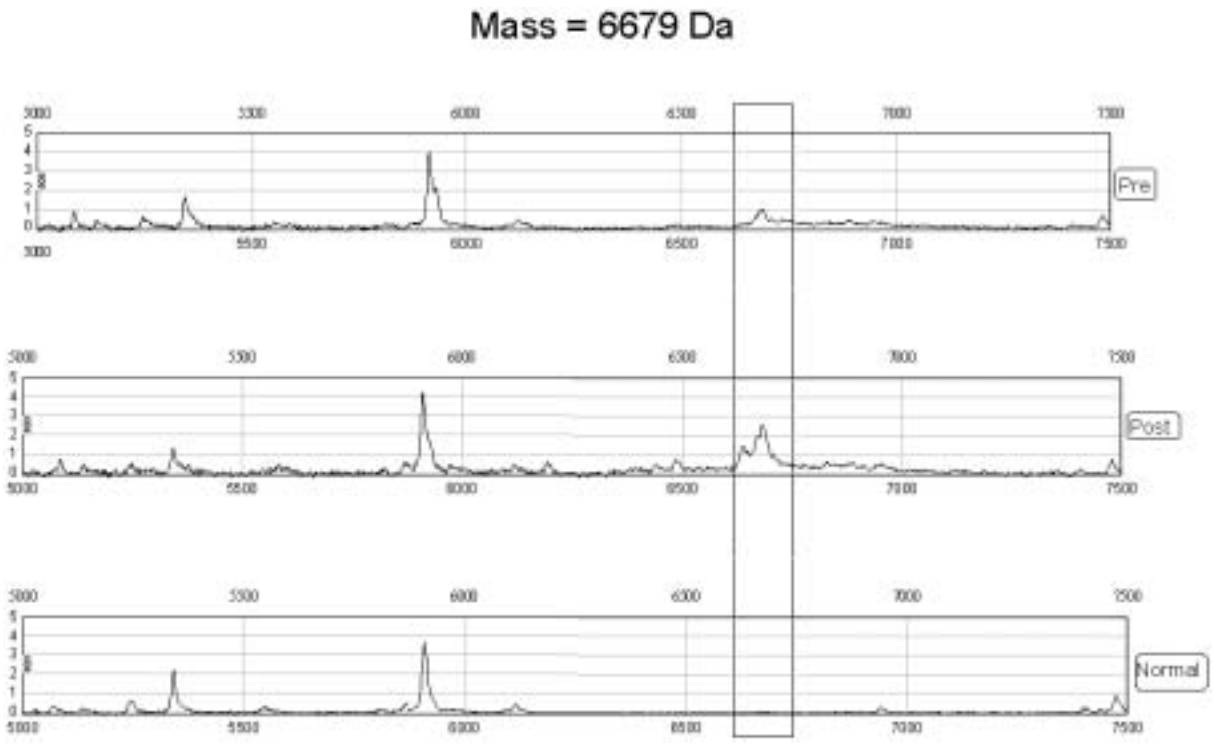

Fig. 3. Representative SELDI spectra comparison from pre-treatment, post-treatment and normal healthy sera ranging from *m/z* 4,000 to 6,000 is shown. The "box" identifies a peak with an average mass of 6679 Da that is underexpressed in the normal healthy samples compared to pre- and post-treatment samples.

tree that discriminates the SLN+ and SLN− groups with 81% sensitivity and 77% specificity in the cross-validation analysis (Table 1). A total of 6 protein peaks, 3 low mass peaks from IMAC-Cu and 3 high mass peaks from SAX chip surfaces, formed the main splitters for the final decision tree (Fig. 5).

4. Discussion

The use of SELDI-TOF and other mass spectrometry platforms for proteomic profiling of body fluids has many potential uses and advantages as a clinical assay. Using readily accessible clinical samples, the technology has proven to be reproducible, has adequate throughput, minimal risk to patients, and is relatively inexpensive [24]. Serum protein profiling with SELDI has been primarily applied to early cancer detection studies, with minimal application to prognostic assay evaluations. In a previous study, 134 serum samples from breast cancer patients, benign breast disease patients and women with no evidence of breast disease were analyzed by SELDI using the IMAC-Cu and SAX protein chip surfaces, and the same type decision tree classification algorithm described herein [21]. Protein profiles from each individual chip surface yielded sensitivities and specificities in the 78–83% range for distinguishing cancer samples from normal patients and those with benign disease. By combining differentially expressed peaks obtained on both chip surfaces, a specificity of 93.3% and sensitivity of 90% were obtained for separating a subset of 30 cancer-affected and 30 healthy patients [21]. Thus, utilizing the already established optimized assay conditions from the previous study allowed us to efficiently analyze the different sample sets for potential diagnostic and prognostic applications.

It could be predicted a priori that because there are distinct protein patterns associated with a breast cancer serum profile relative to a profile for a normal healthy state, then a woman effectively treated for breast cancer would likely revert to a normal profile and women with progression of disease or recurrence would continue to have a cancer profile, or a subset thereof. Identification of those women who do not revert to normal, or a distinct post-treatment profile indicative of successful therapy, would allow for better prognostication and therapy options. To test this hypothesis, serum samples

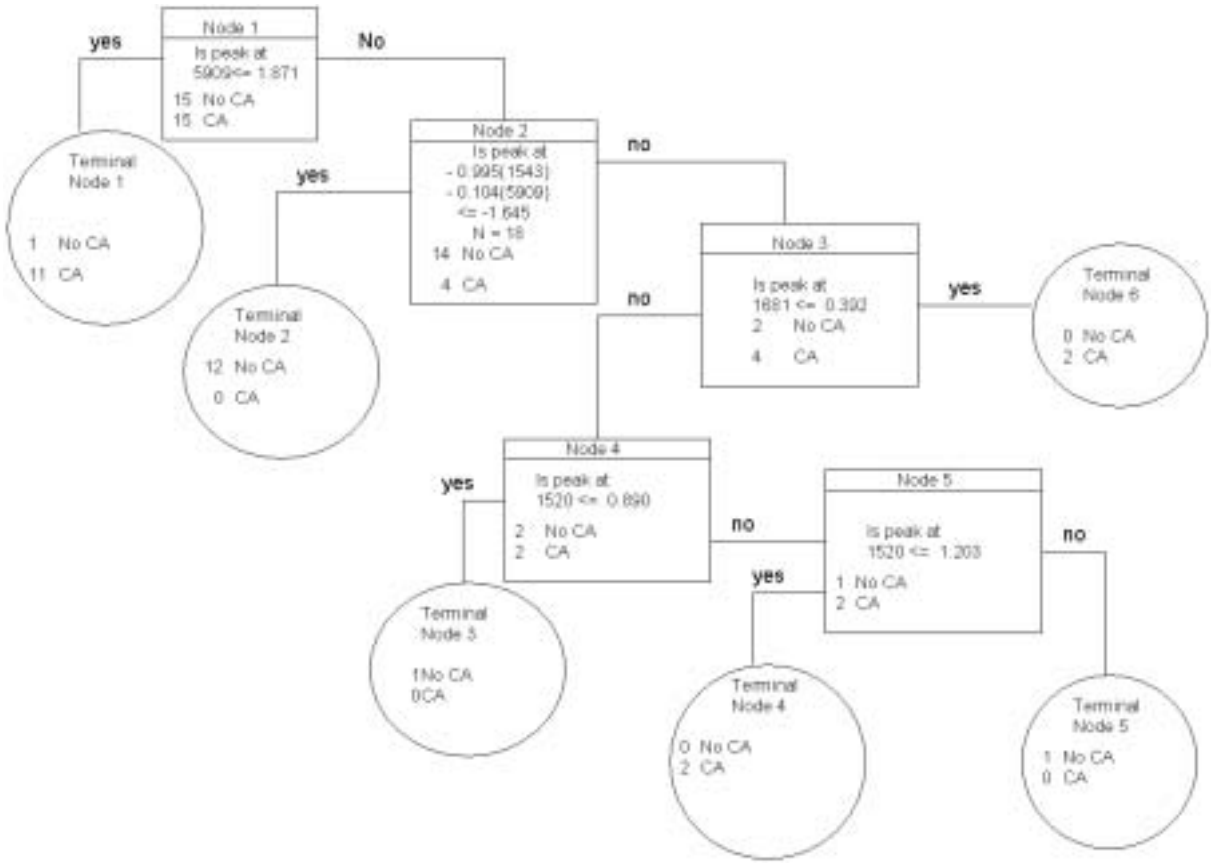

Fig. 4. Diagram of classification tree for BRCA-1 carriers with (CA) and without cancer (No CA). The squares are the primary nodes and the circles indicate terminal nodes. The mass value in the root nodes is followed by = the intensity value.

were obtained for pre and post treatment for all stage 1–3 disease and analyzed by SELDI. Post-treatment samples were readily distinguishable from the corresponding pre-treatment cancer samples, and consistent with our previous studies [21], normal healthy samples were distinct from the cancer samples. In the post-treatment comparison with normal healthy samples, there were 5 distinct peaks obtained from the IMAC-Cu surface that were elevated in both the cancer and post-treatment samples relative to the normal samples. Given this information, one could foresee utilizing proteomic profiles to monitor success of treatment, i.e. neoadjuvant and adjuvant chemotherapy, as well as for the early detection of recurrent breast cancer or metastatic disease, in a way complementary with current surveillance paradigms.

Mammography remains the gold standard screening test for breast cancer with a sensitivity of 75%–90% [5, 6]. The shortcomings of mammographic detection are particularly evident when examining young females who are more likely to have dense breasts [17]. Also, the power of mammography is diminished when dealing with small lesions, like recurrences. Currently, the BI-RADS (Breast Imaging Reporting and Data System) has developed a sliding scale from 1–5 (1 = negative, 2 = benign, 3 = probably benign, 4 = suspicious, 5 = highly suggestive of malignancy) in an effort to guide clinician's treatment. Women with "probably benign" or BI-RAD 3 lesions are followed with mammograms every 6 months for 2 years while those with BI-RAD 4 and 5 lesions undergo an invasive procedure to biopsy the abnormality. Proteomic profiling could be effectively utilized as an adjunct to mammography to aid the clinician in distinguishing lesions that lack specific mammographic findings, i.e. irregular borders and microcalcifications, but could still harbor a focus of malignancy. Increasing the total numbers and obtaining different longitudinal time points following surgery will be necessary to continue these efforts. The implication of the results of this pilot study are that we should expect to identify post-treatment profiles that are distinct from healthy, non-cancer samples. Determining the

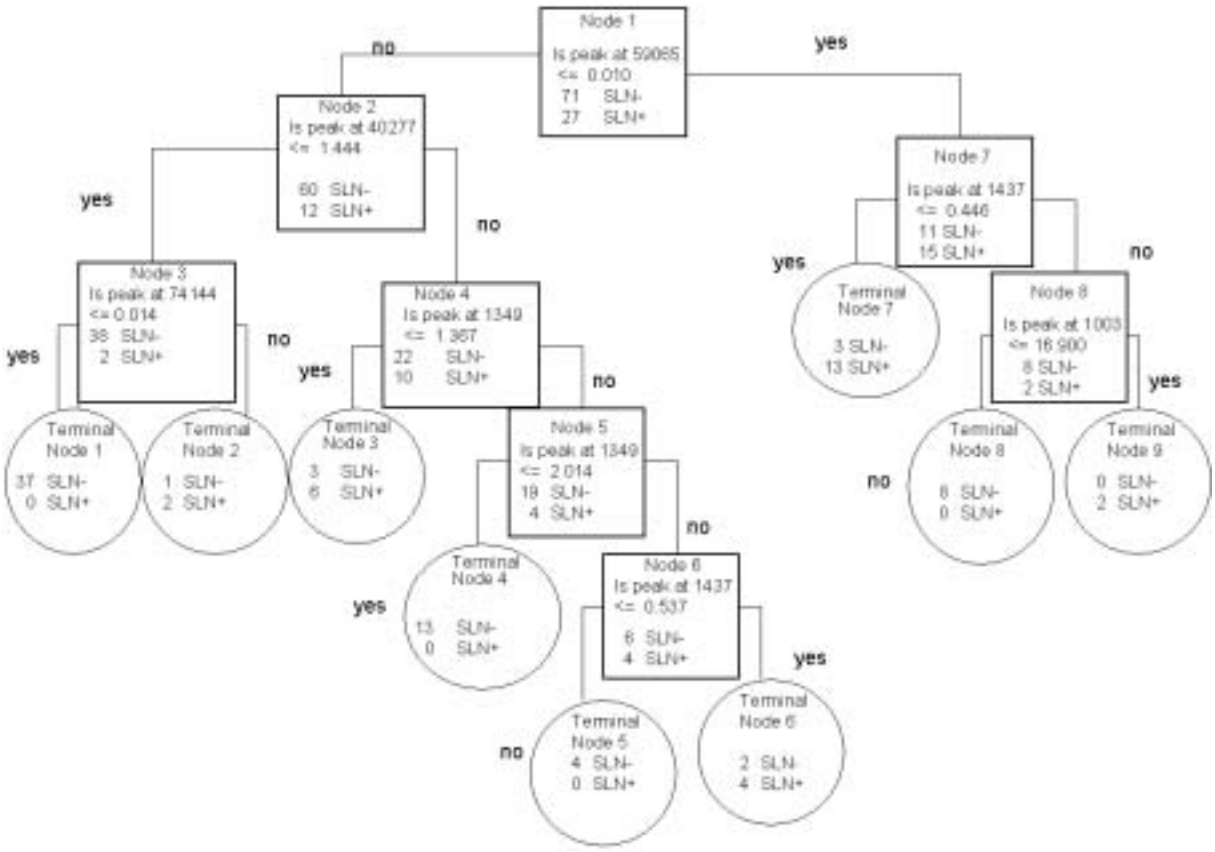

Fig. 5. Diagram of classification tree for sentinel lymph node positive (SLN+) and negative (SLN−) samples. The squares are the primary nodes and the circles indicate terminal nodes. The mass value in the root nodes is followed by = the intensity value.

prognostic significance of these profiles will have to be evaluated in the larger cohort follow-up study.

Carriers of a genetic predisposition such as BRCA 1 or 2 are at high risk for the development of breast cancer at an early age (37% of the carriers of a gene mutation will develop breast cancer by age 50 [9], compared with 2% of the general population), and thus breast cancer screening is recommended to begin at an earlier age for this group. Genetic testing of high-risk women is becoming more widespread and will result in increased identification of younger women underserved by routine imaging. An alternative method of diagnosis must be devised not only for those genetically predisposed but also for improved detection in the general population. The results obtained for the SELDI profiling of the 30 BRCA1 carriers indicates that protein profiling analysis could be a contributing assay for monitoring disease in high-risk women. Whether the classification of the BRCA-1 Ca samples represents earlier detection of occult cancer or a pre-cancerous state is yet to be determined. Larger follow-up studies, including serum of BRCA-1 patients drawn at the time of cancer diagnosis, may allow more timely prophylactic strategies. Because, women identified with a BRCA-1 gene mutation are usually recommended to have prophylactic mastectomies at a young age, the ability to predict which mutation carriers will proceed to cancer would have a tremendous impact on our treatment choices.

Traditionally, assessment of the axillary lymph nodes by dissection (ALND) was performed as part of any operation for the treatment of an invasive breast cancer and allowed for accurate nodal staging followed by therapeutic decisions. Yet, ALND can result in significant permanent complications such as lymphedema and nerve injury. Therefore, in early-stage breast cancers that are less likely to present with node metastases, lymphatic mapping and sentinel lymph node (SLN) biopsy have emerged as new approaches [7]. Because the sentinel node is the first lymph node to receive lymphatic drainage from the primary breast cancer and hence metastatic tumor cells, numerous studies have shown that the findings in the sentinel node accurately

predict the status of the other axillary nodes [7,10,19, 20]. If the SLN does not contain cancer, then the other axillary lymph nodes remain in situ and thereby significantly reduce any subsequent morbidity. On the other hand, if the SLN does contain tumor, the patient must undergo a complete axillary dissection and thus, is at risk for nerve damage and lymphedema. SLN biopsy does have limitations. First and foremost, it is an operator dependent procedure. There are instances, such as in, obese patients, in which finding the SLN may be difficult and thus result in a false negative. Second, multicentric tumors that drain to several nodes may affect the accuracy of this procedure. Lastly, lymph nodes with large tumor burden may have lymphatic drainage impeded, thus limiting the accuracy. The current study was designed to determine first if SELDI profiling could predict the status of the SLN with reasonable sensitivity and specificity. The 81% correct classification of the 98 samples is promising (Fig. 5, Table 1), but much more extensive follow up remains to be done with larger sample sets. Isolation and sequencing of the discriminating proteins used in the classification tree will also be required, and this too could provide new insights into the identification of new biomarkers. Ultimately, a serum-based assay that predicts SLN status, or one that is used an adjunct to SLN determinations, could result in reduced surgical morbidity and further aid in making the decisions about whether axillary node dissection is necessary.

SELDI-TOF analysis of serum samples is an exciting technology that, within the cancer research field, has been primarily applied to studies involving the early detection of cancer. These three pilot studies presented herein indicate the potential application of SELDI profiling for more prognostic types of assays, especially for breast cancers. Because of the molecular heterogeneity of breast cancers, these pilot studies also suggest that it may be feasible to work backwards toward an early detection assay by first profiling the more homogeneous and readily available cancer samples associated with different treatment cohorts that could be evaluated for prognostic pre- and post-treatment markers. A key to any of these approaches will be the evaluation of as many samples as possible in each stratified category. Application of these samples to other proteomic profiling platforms and evaluation of the data using classification algorithms different than decision classification trees are also planned.

Acknowledgements

This study was supported by a grant from the National Cancer Institute Early Detection Research Network (CA85067).

References

[1] B.L. Adam, Y. Qu, J.W. Davis, M.D. Ward, M.A. Clements, L.H. Cazares, O.J. Semmes, P.F. Schellhammer, Y. Yasui, Z. Feng and G.L. Wright, Jr. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men, *Cancer Res* **62** (2002), 3609–3614.

[2] B.L. Adam, A. Vlahou, O.J. Semmes and G.L. Wright, Jr. Proteomic approaches to biomarker discovery in prostate and bladder cancers, *Proteomics* **1** (2001), 1264–1270.

[3] R.C. Bast P. Ravdin and D.F. Hayes, et al., 2000 update of recommendations for the use of tumor markers in breast and colorectal cancer: clinical practice guidelines of the American Society of Clinical Oncology, *J Clin Oncol* **19** (2001), 1865–1878.

[4] L.H. Cazares, B.L. Adam, M.D. Ward, S. Nasim, P.F. Schellhammer, O.J. Semmes and G.L. Wright, Jr. Normal, benign, preneoplastic, and malignant prostate cells have distinct protein expression profiles resolved by surface enhanced laser desorption/ionization mass spectrometry, *Clin Cancer Res* **8** (2002), 2541–2552.

[5] W.L. Donegan, Evaluation of a palpable breast mass, *N Engl J Med* **327** (1992), 937–942.

[6] J.C. Elmore, M.B. Barton, V.M. Moceri, S. Polk, P.J. Arena and S.W. Fletcher, Ten-year risk of false positive screening mammograms and clinical breast examinations, *N. Engl. J. Med.* **338** (1998), 1089–1096.

[7] A.E. Giuliano, D.M. Kirgan, J.M. Guenther and D.L. Morton, Lymphatic mapping and sentinel lymphadenectomy for breast cancer, *Ann. Surg.* **220** (1994), 391–401.

[8] D.F. Hayes B. Trock and A.L. Harris, Assessing the clinical impact of prognostic factors: when is statistically significant clinically useful? *Breast Cancer Res Treat* **52** (1998), 305–319.

[9] M.C. King, J.H. Marks and J.B. Mandell, and the New York Breast Cancer Study Group. Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2, *Science* **302** (2003), 643–646.

[10] D. Krag, D. Weaver, T. Ashikaga, F. Moffat, V.S. Klimberg, C. Shriver, S. Feldman, R. Kusminsky, M. Gadd, J. Kuhn, S. Harlow and P. Beitsch, The sentinel node in breast cancer–a multicenter validation study, *N. Eng. J. Med.* **339** (1998), 941–946.

[11] J. Li, Z. Zhang, J. Rosenzweig, Y.Y. Wang and D.W. Chan, Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer, *Clin Chem* **48** (2002), 1296–1304.

[12] E.F. Petricoin, A.M. Ardekani, B.A. Hitt, P.J. Levine, V.A. Fusaro, S.M. Steinberg, G.B. Mills, C. Simone, D.A. Fishman, E.C. Kohn and L.A. Liotta, Use of proteomic patterns in serum to identify ovarian cancer, *Lancet* **359** (2002), 572–577.

[13] E.F. Petricoin, D.K. Ornstein, C.P. Paweletz, A. Ardekani, P.S. Hackett, B.A. Hitt, A. Velassco, C. Truco, L. Wiegand, K. Wood, C.B. Simone, P.J. Levine, W.M. Linehan, M.R.

Emmert-Buck, S.M. Steinberg, E.C. Kohn and L.A. Liotta, Serum proteomic patterns for detection of prostate cancer, *J. Natl. Cancer Inst.* **94** (2002), 1576–1578.

[14] Y. Qu, B.L. Adam, M. Thornquist, J.D. Potter, M.L. Thompson, Y. Yasui, J. Davis, P.F. Schellhammer, L. Cazares, M. Clements, G.L. Wright, Jr. and Z. Feng, Data reduction using a discrete wavelet transform in discriminant analysis of very high dimensionality data, *Biometrics* **59** (2003), 143–151.

[15] Y. Qu, B.L. Adam, Y. Yasui, M.D. Ward, L.H. Cazares, P.F. Schellhammer, Z. Feng, O.J. Semmes and G.L. Wright, Jr. Boosted decision tree analysis of surface-enhanced laser desorption/ionization mass spectral serum profiles discriminates prostate cancer from noncancer patients, *Clin. Chem.* **48** (2002), 1835–1843.

[16] J.S. Ross, G.P. Linette, J. Stec, E. Clark, M. Ayers, N. Leschly, W.F. Symmans, G.N. Hortobagyi and L. Pusztai, Breast cancer biomarkers and molecular medicine, *Expert Rev Mol Diagn* **3** (2003), 573–585.

[17] I. Saarenmaa, T. Salminen, U. Geiger, P. Heikkinen, S. Hyvarinen, J. Isola, V. Kataja, M.L. Kokko, R. Kokko, E. Kumpulainen, A. Karkkainen, J. Pakkanen, P. Peltonen, A. Piironen, A. Salo, M.L. Talviala and M. Haka, The effect of age and density of the breast on the sensitivity of breast cancer diagnostic by mammography and ultrasonography, *Breast Cancer Res. Treat.* **67** (2001), 117–123.

[18] P.R. Srinivas, S. Srivastava, S. Hanash and G.L. Wright, Jr. Proteomics in early detection of cancer, *Clin Chem* **47** (2001), 1901–1911.

[19] U. Veronesi, G. Paganelli, G. Viale, A. Luini, S. Zurrida, V. Galimberti, M. Intra, P. Veronesi, C. Robertson, P. Maisonneuve, G. Renne, C. De Cicco, F. De Lucia and R. Gennari, A randomized comparison of sentinel-node biopsy with routine axillary dissection in breast cancer, *N. Eng. J. Med.* **349** (2003), 546–553.

[20] U. Veronesi, G. Paganelli, V. Galimberti, G. Viale, S. Zurrida, M. Bedoni, A. Costa, C. de Cicco, J.G. Geraghty, A. Luini, V. Sacchini and P. Veronesi, Sentinel-node biopsy to avoid axillary dissection in breast cancer with clinically negative lymph-nodes, *Lancet* **349** (1997), 1864–1867.

[21] A. Vlahou, C. Laronga, L. Wilson, B. Gregory, K. Fournier, D. McGaughey, R.R. Perry, G.L. Wright, Jr and O.J. Semmes, A novel approach toward development of a rapid blood test for breast cancer, *Clin Breast Cancer* **4** (2003), 203–209.

[22] A. Vlahou, P.F. Schellhammer, S. Mendrinos, K. Patel, F.I. Kondylis, L. Gong, S. Nasim and G.L. Wright, Jr. Development of a novel proteomic approach for the detection of transitional cell carcinoma of the bladder in urine, *Amer. J. Pathology* **158** (2001), 1491–1502.

[23] J.T. Wadsworth, K.D. Somers, B.C. Stack, Jr., L. Cazares, G. Malik, B.L. Adam, G.L. Wright, Jr. and O.J. Semmes, Identification of head and neck cancer patients using serum protein profiles, *Arch Otolaryngol Head Neck Surg* (2004, in press).

[24] G.L. Wright, Jr. SELDI proteinchip MS: a platform for biomarker discovery and cancer diagnosis, *Expert Rev Mol Diagn* **2** (2002), 549–563.

[25] Y. Yasui, M. Pepe, M.L. Thompson, B.L. Adam, G.L. Wright, Jr., Y. Qu, J.D. Potter, M. Winget, M. Thornquist and Z. Feng, A data-analytic strategy for protein biomarker discovery: profiling of high-dimensional proteomic data for cancer detection, *Biostatistics* **4** (2003), 449–463.