Discovering Clinical Biomarkers of Ionizing Radiation Exposure with Serum Proteomic Analysis

Cynthia Ménard,[1,5] Donald Johann,[2] Mark Lowenthal,[2] Thierry Muanza,[1,6] Mary Sproull,[1] Sally Ross,[2] James Gulley,[3] Emanuel Petricoin,[4] C. Norman Coleman,[1] Gordon Whiteley,[7] Lance Liotta,[2] and Kevin Camphausen[1]

[1]Radiation Oncology Branch, National Cancer Institute; [2]National Cancer Institute-Food and Drug Administration Clinical Proteomics Program, Laboratory of Pathology; [3]Laboratory of Tumor Immunology and Biology, Center for Cancer Research, National Cancer Institute, NIH-Department of Health and Human Services; [4]National Cancer Institute-Food and Drug Administration Clinical Proteomics Program, Office of Cellular and Gene Therapy, Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, Maryland; [5]Radiation Medicine Program, Princess Margaret Hospital, University of Toronto, Toronto, Ontario, Canada; [6]Department of Radiation Oncology, McGill University, Montreal, Quebec, Canada; and [7]Science Applications International Corporation-Frederick, Inc., Clinical Proteomics Reference Laboratory, National Cancer Institute Frederick, Gaithersburg, Maryland

Abstract

In this study, we sought to explore the merit of proteomic profiling strategies in patients with cancer before and during radiotherapy in an effort to discover clinical biomarkers of radiation exposure. Patients with a diagnosis of cancer provided informed consent for enrollment on a study permitting the collection of serum immediately before and during a course of radiation therapy. High-resolution surface-enhanced laser desorption and ionization-time of flight (SELDI-TOF) mass spectrometry (MS) was used to generate high-throughput proteomic profiles of unfractionated serum samples using an immobilized metal ion-affinity chromatography nickel-affinity chip surface. Resultant proteomic profiles were analyzed for unique biomarker signatures using supervised classification techniques. MS-based protein identification was then done on pooled sera in an effort to begin to identify specific protein fragments that are altered with radiation exposure. Sixty-eight patients with a wide range of diagnoses and radiation treatment plans provided serum samples both before and during ionizing radiation exposure. Computer-based analyses of the SELDI protein spectra could distinguish unexposed from radiation-exposed patient samples with 91% to 100% sensitivity and 97% to 100% specificity using various classifier models. The method also showed an ability to distinguish high from low dose-volume levels of exposure with a sensitivity of 83% to 100% and specificity of 91% to 100%. Using direct identity techniques of albumin-bound peptides, known to underpin the SELDI-TOF fingerprints, 23 protein fragments/peptides were uniquely detected in the radiation exposure group, including an interleukin-6 precursor protein. The composition of proteins in serum seems to change with ionizing radiation exposure. Proteomic analysis for the discovery of clinical biomarkers of radiation exposure warrants further study. (Cancer Res 2006; 66(3): 1844-50)

Introduction

Decades of research have been dedicated to the discovery of inherent biomarkers of radiation exposure. Prior studies have highlighted the importance of this research in the epidemiologic field of occupational exposure (1), exposure assessments in cases of environmental or industrial accidents (2), astronaut exposure in manned space missions (3–5), and for early markers of clinical response to radiation therapy for cancer (6–8). Recently, fears of radiological terror incidents have renewed the search for rapid and simple methods to identify exposed individuals within large populations and to predict health effects (9, 10).

The long-established gold standard biomarker of radiation exposure, cytogenetic analysis of peripheral blood lymphocytes, requires complex analysis by skilled workers (5, 11, 12). A number of alternative minimally invasive strategies have been explored, including the measurement of specific cytokines and metabolites in tissues and body fluids (4), micronucleus assays (13), and germ line radiation–induced apoptosis rates (2). More recently, a need for comprehensive profiling in the field of biomarker discovery has been recognized (10, 14). Clinical studies to date have focused on genomic profiling of circulating peripheral blood lymphocytes using microarray technology (15–17).

Although the ability to distinguish cancers from the profile of circulating protein constituents of serum was first shown in 1978 (18), only recent technological advances have led to robust techniques that permit high-throughput and comprehensive analysis of serum proteins (19). These techniques, including surface-enhanced laser desorption and ionization time of flight (SELDI-TOF) mass spectrometry (MS; ref. 20) and complementary mass spectroscopic fractionation (21), enable the study of global expression changes following pathologic events and physiologic stresses. As such, they hold promise as a comprehensive technique for biomarker discovery. Recently, it has been shown that most of the low-molecular-weight molecules that underpin the SELDI-TOF MS profiles exist bound to larger highly abundant proteins, such as albumin (22).

Thus, we sought to explore the merit of proteomic profiling strategies in patients with cancer before and during radiotherapy in an effort to discover clinical biomarkers of radiation exposure. We also sought to further characterize the information content of the spectral components using an albumin enrichment strategy of the same serum samples used in the high-throughput discovery series. There are three hypotheses tested in this study: First, proteomic profiling methods can discover a discriminating set of low-molecular-weight biomarkers that, in turn, can be used as a pattern-based diagnostic of human exposure to ionizing radiation. Second, a serum proteomic pattern diagnostic can discriminate a low dose-volume radiation exposure from a

high dose-volume exposure. Last, MS identification of the same information archive that underpins the observed MS peaks will reveal important clues to the identity and nature of the discriminatory entities.

Materials and Methods

Specimen collection. Sixty-eight patients with a diagnosis of cancer provided informed consent for enrollment on one of two Institutional Review Board–approved studies permitting the collection of serum immediately before and during a course of radiation therapy between 2001 and 2003. The interval between the first and second sample collection varied according to patient preference, with the second sample typically acquired in the latter part of the course of treatment. Approximately 8 mL of blood was drawn by venipuncture and placed on ice. The samples were centrifuged within 2 hours of collection, and serum was aliquoted into Eppendorf tubes and stored at $-80°C$.

Specimen processing—SELDI-TOF. Aliquots of fresh frozen sera were sent to the National Cancer Institute (NCI)-Food and Drug Administration Clinical Proteomics laboratory. The pretreatment and posttreatment samples all had been collected and handled in an exact identical manner, not previously thawed to minimize any potential for collection and sample handling bias. Samples were thawed and 10 μL aliquots were obtained and used immediately for high-resolution SELDI-TOF analysis. Ciphergen (Fremont, CA) immobilized metal ion-affinity chromatography 3 protein arrays were used for this study. All steps were carried out on a Tecan Genesis 2000 robotic processor (Tecan, Research Triangle Park, NC) equipped with a 96-position Temo unit to reduce any potential for bias due to variation in sample preparation. The Ciphergen Bioprocessor unit was used to hold the protein arrays in place for all steps in the process. The patient-matched preexposure and postexposure samples were randomized and comingled on the same chip surface to reduce any bias that may occur due to run order and daily spectrometer fluctuations. The surfaces were activated with two applications of 100 μL per spot of 50 mmol/L nickel sulfate (Sigma, St. Louis, MO) incubated for 5 minutes per application. After aspiration, the surfaces were washed twice with 100 μL of deionized water. Dulbecco's PBS without calcium and magnesium chloride (Invitrogen, Grand Island, NY) with 0.1% Triton X-100 (Sigma) was added to each spot with a volume of 100 μL per spot and incubated for 5 minutes, aspirated, and discarded. This was repeated a second time. The chip surfaces were then dried using vacuum aspiration to remove any excess liquid. A total of 5 μL of undiluted serum was added to each spot and incubated for 20 minutes. The wash steps were carried out on the Tecan processor as follows: 100 μL of PBS was added and mixed by aspiration and dispensing. This was done for three cycles. Sterile deionized water was added and mixed by aspiration and dispensing. Remaining liquid was removed by vacuum aspiration and the chips were allowed to dry completely before the bioprocessor gasket was removed. Once removed, 1 μL matrix consisting of 12 mg/mL cinnamic acid (Fluka, St. Louis, MO) in 50% acetonitrile (Sigma) and 0.5% trifluoracetic acid (Sigma) was applied by the Tecan processor using a fixed pipette tip and a serpentine path to each spot on the arrays. The arrays were then dried at room temperature for 5 minutes and the matrix was reapplied. After the chips were completely dried, they were read on an ABI Q-Star mass spectrometer equipped with a Ciphergen PCI 1000 Protein Chip Interface as previously described (23).

Data analysis—SELDI-TOF. High-resolution MS data were derived by the unexposed and exposed blood draws from 68 cancer patients for an initial total of 136 spectra. Regions of spectra known to be dominated by matrix effect (<1,000 Da) or terminal measurement effects (>11,000 Da) were excluded *a priori*. Quality assurance/quality control (QA/QC) analyses resulted in the elimination of one spectrum from the unexposed group due to low average amplitude and total ion counts. The QA/QC tools and procedures have been previously described (23). Quartile values were computed for each spectra. Initial aggregate pattern analyses using visual data mining techniques showed interesting groupings involving the most abundant ions (third quartile or greater). Spectral data below this threshold were excluded.

High-resolution mass spectral data were then binned as a function of the nominal mass values, while generating a variety of statistical measures for the captured elements (mass and amplitude values), including, means, SDs, variances, and sums.

An amplitude-based concept hierarchy was constructed for each spectra based on percentage contributions. Normalization was achieved through transforming the summed amplitude values through the concept hierarchy mapping scheme (24). Aggregate pattern analyses were done by three-dimensional visual data mining techniques to explore for regions of interest. An example of this is illustrated in Fig. 1*A*. Given the heterogeneity present in the data due to large variety of diseases/cancers (18), as well as a wide range of both anatomic sites (9) and degree of radiation received, the visual data mining activities aided greatly in the identification of interesting ion patterns for additional pursuit (25). Custom T-SQL stored procedures were written to further aid in feature selection through the implementation of a minimax algorithmic procedure to maximize the differences between candidate ion groups while minimizing their coefficient of variation.

The unexposed ($n = 67$) and exposed ($n = 68$) case-controlled spectra were then analyzed by a variety of classification methods derived from
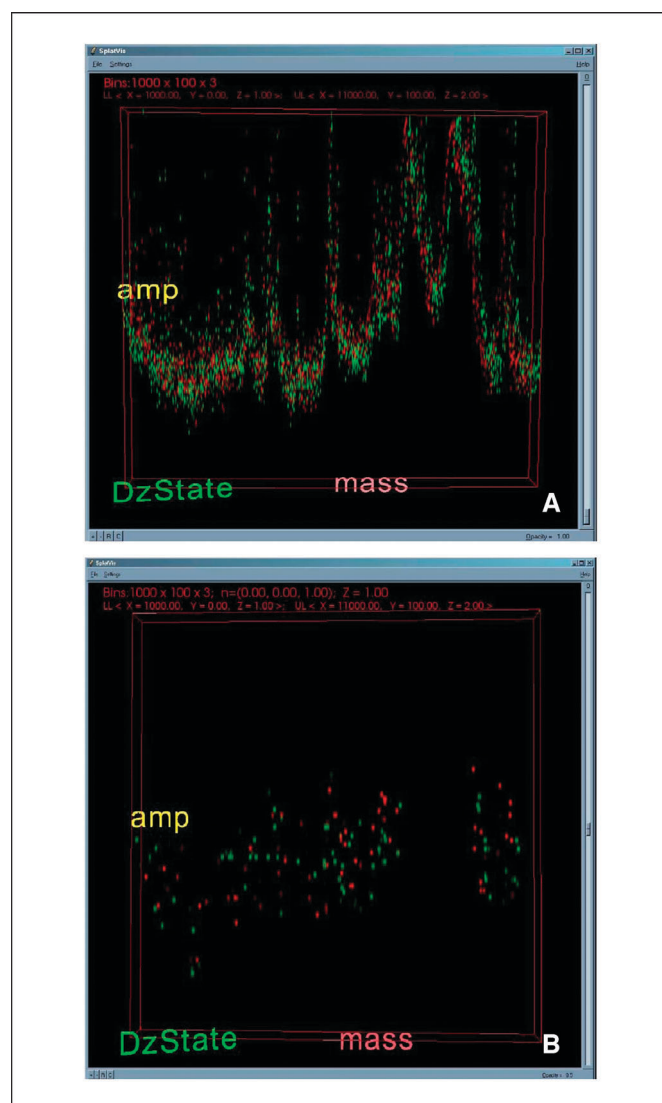

Figure 1. Interactive visual data mining (three dimensional) enables aggregate pattern analyses. *A,* visualization helps to identify areas for subsequent data mining operations. *B,* key sets of ions are revisualized following data mining. *X axis* (*mass*), *m/z* values. *Y axis* (*amp*), amplitude measurements. *Z* axis (*DzState*) and color (*red,* exposed; *green,* unexposed) are used for class grouping during dynamic data set manipulations.

Clementine, version 8.5 (SPSS, Chicago, IL). The three methods of classification tested were as follows: (*a*) C5.0 decision tree, (*b*) a hybrid classifier constructed from a multilayer perceptron neural network using an exhaustive prune strategy linked into a decision tree, and (*c*) a hybrid classifier constructed from a Classification and Regression Tree (CART) coupled to a C5.0 decision tree. Cross-validation was used to assess the classification error rates of the various models. As previously stated, due to the inherent heterogeneity in this data set and relatively small size, it was felt a validation strategy using a rotation method would better address biological relevance as well as credence in classifier accuracies. Construction of larger training data sets permitted more accurate predictive classification for this study set. The data set was partitioned into N equal-sized subgroups and models were fitted N times. N was chosen to be 10. Each model used $N - 1$ of the subgroups for training, and then applied the resulting model to the remaining subgroup. Accuracy values were averaged over the N holdout subgroups.

These computational techniques were also used to construct supervised predictive classifiers to discriminate high ($n = 32$) versus low ($n = 36$) dose-volume radiation exposure from the exposed group. The method for dividing exposure dose-volumes into high and low was done by a statistical stratification method using quartile analysis of dose-volume measurements coupled to aggregation of both cancer type and site of radiation. This revealed a natural cut in the dose-volume measurements at 190 Gy/m^2; therefore, measurements >190 Gy/m^2 were recorded as a high exposure and the remainder as low. In the low dose-volume classification, the mean dose-volume was 126 Gy/m^2 (range 44-190); in the high dose-volume classification, the mean dose-volume was 350 Gy/m^2 (range 200-1,050). Attention to disease/cancer type and anatomic site of radiation was taken to avoid bias. For instance, there are 18 prostate cancer patients in the low dose-volume group, and 15 in the high dose-volume group. In this study, the units for dose-volume measurements are Gy × fraction body volume exposed to >50% radiation dose. This definition was chosen *a priori* and assumes that both radiation dose and irradiated volume contribute equivalently to changes in the serum proteome.

Deliberate steps were taken to produce the most general predictive models and to avoid overtraining. To this end, strategies used included the selection of overtraining prevention options, the use of global pruning with a pruning severity of 75%, requiring minimum records per child branch to be greater than or equal to a threshold value (i.e., 4), and winnowing of attributes for a more parsimonious solution. To enhance predictive accuracy of the decision tree models, boosting was used.

Visual data mining techniques were also used in an iterative fashion as candidate ions were meticulously winnowed. In many cases, candidate ion sets, whether they were considered for promotion or elimination, were further analyzed by high-resolution visual data mining methods. These techniques use a high-end PC graphics card (NVIDIA Quardro FX 3000, NVIDIA Corp., Sunnyvale, CA) coupled to 9-megapixel high-resolution IBM T221 computer graphics monitor (IBM, Armonk, NY). Thus, high-resolution mining of extremely complex data sets may be addressed. Figure 1*B* illustrates aspects of a three-dimensional visual data mining exercise involving an intermediate set of candidate ions.

Both a high-end computer workstation and server were used to accomplish the aforementioned computational tasks. These are located at the NCI, Center for Cancer Research, Clinical Proteomics Reference Laboratory (CPRL, Gaithersburg, MD). The server consists of dual processor AMD Opertron CPU module (AMD Corporation, Sunnyvale, CA) housed on a Newisys motherboard (Newisys Corp, Austin, TX) fitted with a total of 16 GB of main memory. Server disc capacity consists of an Adaptec RAID housing assembly containing twelve 146 GB IBM disc drives in a RAID 5 configuration, allowing one disc for parity checking and one disc for a hot spare. Thus, the effective storage capacity for this server at the CPRL is ~ 1.3 TB. Following CPRL specification, the server system was custom built by Colfax International (Sunnyvale, CA).

The server computer operating system is the Windows 2003 Advanced Server (Microsoft Corp, Redmond, WA). Database activities are done through the use of Microsoft SQL Server 2000 Enterprise Edition. Computer workstations at the CPRL consist of one to two Xeon processors

(Intel Corporation, Santa Clara, CA), with 2 to 4 GB of memory, running the Windows XP Professional operating system. Visual data mining activities are done on workstations outfitted with high-end graphics cards and displays. Data intensive computational tasks are off-loaded to the server.

MS-MS protein identification of pooled sera. From the original study set, a subset of 25 patients was randomly selected and their serum was pooled for MS-based sequencing and peptide identification from samples obtained before and during radiation therapy.

Albumin and associated low-molecular-weight bound peptide purification was done in the following fashion and has been further described previously (22). Forty microliters of preradiation or postradiation treatment pooled serum (~ 5 mg protein) were diluted to 200 μL with an equilibration buffer (Millipore) and run twice through a Montage (Millipore, Billerica, MA) albumin-specific affinity column. The bound protein was washed thoroughly and eluted from the column by equilibrating with 70% ACN/30% H$_2$O/0.2% trifluoroacetic acid for 30 minutes, followed by a slow spin-through of the elution mixture. The eluate was lyophilized to <10 μL in a HetoVac rotofor (CT 110) and reconstituted in a 95% H$_2$O/5% ACN/0.1% formic acid buffer (buffer A). Samples were sometimes desalted with a ZipTip cleanup or with Vivaspin 500 centrifugal membranes and always reconstituted in a 1:1 mixture of water and SDS sample buffer (20 μL total volume).

Analysis then proceeded with one-dimensional gel separation and digestion. Twenty microliters of sample in SDS sample buffer (40 μL of original serum) were boiled for 5 minutes at 95°C and run on a one-dimensional precast gel (4-20% Tris-glycine or 4-12% bis-Tris) to separate albumin from the low abundant proteins/peptides/fragments of interest. The gel was stained with Coomassie blue for 1 hour and destained overnight in 30% methanol/10% acetic acid solution. The entire lane containing stage-specific serum proteins was excised from the gel and finely sliced into very small molecular weight regions (~ 40 slices/lane). Gel bands were reduced and alkylated with 10 mmol/L DTT and 55 mmol/L iodoacetamide, incubated at 4°C for 1 hour in porcine trypsin (20 ng/μL; Promega) and allowed to digest overnight at 37°C in 25 mmol/L NH$_4$HCO$_3$. The following morning, peptides were extracted from the gel with a 70% ACN/5% formic acid solution.

Samples were lyophilized to near dryness and reconstituted in 6.5 μL of buffer A for mass spectrometric analysis. Microcapillary reverse phase liquid chromatography (LC)/MS/MS analysis was done with Dionex's LC Packings LC system (Dionex, Sunnyvale, CA) coupled online to a ThermoFinnigan LCQ Classic ion trap mass spectrometer (ThermoFinnigan, San Jose, CA) with a nanospray source. Reverse phase separations were done with an in-house, slurry-packed capillary column. The C$_{18}$ silica-bonded column is 75 μm ID, 360 μm OD, 10-cm-long fused silica packed with 5 μm beads with 300 Å pores (Vydac, Hesperia, CA). A μ-precolumn PepMap C$_{18}$ cartridge (Dionex) acts as a desalting column. Sample is injected in microliter pick-up mode and washed with buffer A for 5 minutes before a linear gradient elution with buffer B (95% ACN/5% H$_2$O/0.1% formic acid) up to 85% over 95 minutes at a flow rate of 200 nL/min. Full MS scans are followed by three MS/MS scans of the most abundant peptide ions (in a data-dependent mode) and collision-induced dissociation is done at a collision energy of 38%.

Data analysis was done by searching MS/MS spectra against the European Bioinformatics Institute of the nonredundant proteome set of Swiss-Prot, TrEMBL, and Ensembl entries through the Sequest Bioworks Browser (ThermoFinnigan). Peptides were considered legitimate hits after filtering the correlation scores and manual inspection of the MS/MS data (26–30).

Accepted peptide hits are required to have an X_{corr} ranking = 1 relative to all other peptides in the database. The albumin extraction, gel electrophoresis, protein digestion/extraction, and LC/MS/MS analysis was repeated in six distinct trials (three for preradiation treatment and three for postradiation treatment)—each time yielding diminishing returns of new identifications for low abundance peptide hits. Repetitive sequencing of identical peptides in multiple trials further validates our experimental procedure—both within and between radiation treatment stages.

Results

Patient characteristics for the SELDI and pooled MS study sets are described in Table 1. Due to referral patterns in our clinic, a

Table 1. Patient characteristics

Proteomic analysis method	SELDI-TOF	MS/MS*
Sample size	68	25
Diagnosis		
Prostate cancer	33	8
Breast cancer	5	2
Glioma	4	2
Ewing's sarcoma	4	3
Non-Hodgkin lymphoma	4	3
Renal cell carcinoma	2	1
Colorectal cancer	2	—
Chronic myelogenous leukemia	2	1
Acute lymphoblastic leukemia	2	—
Plasmacytoma	2	2
Rhabdomyosarcoma	1	—
Hemangioblastoma	1	—
Scleroderma	1	1
Myelodysplastic syndrome	1	—
Multiple myeloma	1	—
Malignant fibrous histiocytoma	1	1
Melanoma	1	—
Ependymoma	1	1
Site irradiated		
Pelvis	38	11
Breast/chest wall	9	4
Total body	5	1
Brain	5	2
Neck	4	3
Spine	3	1
Leg	2	1
Lung	1	1
Skin	1	1
Mean time elapsed between draws (d; range)	35 (1-55)	31(1-54)
Mean radiation dose (Gy; range)	48.6 (1.5-86.4)	42.1(6-62)
Mean dose-volume (Gy × fraction body volume exposed to >50% radiation dose; range)	2.3 (0.4-6.4)	2.4(0.4-6.4)

*MS pooled fractionation.

disproportionately high number of patients with prostate cancer and/or receiving pelvic radiotherapy are represented in this study. Overall, our study population consists of a wide range of diagnoses, body site, elapse time interval, volume, and dose of radiotherapy. All patients were exposed to ionizing radiation between the first and second blood sample acquisitions.

SELDI-TOF analysis. A summary of the classification techniques used in this study, the data sets to which they were applied, and results are contained in Tables 2 and 3. Table 2 addresses the first stated hypothesis that proteomic profiles can distinguish unexposed from radiation-exposed patient serum samples. This involves the analyses of 135 spectra, obtained by randomized and comingled placement on the SELDI chip surfaces, with the unexposed group having an $n = 67$, and the exposed group with an $n = 68$. A variety of classification methods were used for consensus and concordance purposes. As is seen by the reported sensitivity and specificity values, all classification methods did well. Different sets of key ions were discovered and used by each method. Ions common to more than one pattern-based diagnostic include ions 7,264, 2,551, 7,554, 2,698, and 10,157.

Table 3 addresses the second hypothesis that proteomic profiles distinguish high from low dose-volume radiation exposures. These analyses involve only the radiation-exposed group of samples ($n = 68$; high dose-volume = 32, low = 36). To facilitate accord, a variety of classification methods were again used to assess the overall predicative accuracy of a proteomic pattern-based diagnostic. The reported predictive results show good performance in distinguishing the degree of ionizing radiation exposure (low or high dose-volume). Common discriminating ions for these models are 7,691 and 7,889.

MS-MS protein identification of pooled sera. A total of 182 protein fragments were identified in this study, 82 of which were uniquely identified in pooled serum before patients were exposed to ionizing radiation, and 23 of which were uniquely identified in pooled serum while patients were being exposed to ionizing radiation (Table 4—ions listed in alphabetical order). The remaining 77 protein fragments were identified both before and during radiation exposure (data not shown).

Discussion

Following radiation accidents or intentional exposure, it is critically important to estimate rapidly the level of exposure in

Table 2. Sensitivity and specificity of each classifier method for distinguishing radiation exposed from unexposed SELDI-TOF serum spectra (*n* = 135)

Classifier method	Sensitivity (%)	Specificity (%)	Discriminating ions
C5.0 decision tree	100	100	10,464; 7,554; 7,379; 7,264; 6,634; 9,525; 5,782; 5,765; 1,463; 9,781; 2,551
Hybrid: CART + C5.0 decision tree	100	100	2,538; 9,748; 5,972; 2,698; 5,936; 4,711; 6,532; 6,158; 7,611; 10,157; 7,264; 2,300; 10,314; 7,214
Hybrid: neural network + C5.0 decision tree	91	97	7,554; 6,255; 10,157; 1,617; 2,551; 2,698; 5,771

persons at risk. In the moderate dose range of 1 to 10 Gy whole body exposure, gold standard lymphocyte assays have restricted value as this cell population is depleted as a function of both dose and time after exposure (31). Microarray studies of exposed cell cultures have led to the discovery of GADD45 and CDKN1a as potential biomarkers of radiation exposure. A linear gene induction was found between 2 and 50 cGy, supporting the idea that it may be possible to develop gene induction profiles that are useful markers of human radiation exposure or dose (32). However, this approach is more challenging for clinical samples given the low abundance of RNA in isolated peripheral blood lymphocytes from venipuncture (17). In the clinically critical moderate exposure dose range where lymphocytes are depleted, alternative biomarkers need to be considered and developed further.

In this study, we opted to analyze serum samples in an aggregate manner from a population of patients with known exposure to a wide range of therapeutic doses of ionization radiation, over varying time intervals, and at different body sites and volumes of irradiation. This study design was chosen with the overarching goal of identifying serum protein markers of radiation exposure that are independent of volume and body site of exposure at dose levels that may have relevance to moderate dose whole body exposures.

Challenges unique to this particular study include a large degree of inherent heterogeneity due to two factors. First, the 68 patients exhibit 18 different medical diagnoses, which are predominately different forms of cancer. Second, there are nine different anatomic sites of radiation exposure/delivery. Nonetheless, deliberate computational strategies were used to attempt to model and understand the innate heterogeneity, while also maximizing robustness of classification methods. This was accomplished mainly through visual data mining activities that allowed for aggregate pattern analysis and discovery of key ion groupings, including an initial observation indicating a discriminatory ability in the most abundant ions (likely influenced by injury response and immune mediated events), as well as the use of *n*-fold cross-validation strategies for the testing of classifier error statistics. Novel visualization techniques facilitated aggregate spectra pattern

analyses as well as reduced data set strategies, thus enabling study of cohort heterogeneity.

As a secondary objective, we provide preliminary evidence of a dose-response relationship to the proteomic changes induced with radiation exposure. This early data must be interpreted with caution given the small sample size and lack of a patient control study design for this end point. Future studies in larger groups of patients are warranted to determine the linearity of dose-volume effects and the time course of serum proteome changes after exposure.

To further characterize the nature and identify of the SELDI-TOF profiles, we used a new method of analyzing the albumin-bound peptide archive (22). The molecules that comprise the SELDI portraits have been recently shown to emanate from carrier protein binding, which serves to amplify and enrich low-molecular-weight peptides that would otherwise be efficiently cleared by glomerular filtration (22). Two pooled sera sets, representing a randomly obtained subset from patients before and during exposure, were used for this study. The lower abundance proteins uniquely identified in pooled serum may be secreted or shed by cells as a result of signaling, necrosis, and/or apoptosis of cells in the high-dose region. In fact, one of the entities identified by this approach and that may be a component of the SELDI-TOF spectra is a fragment of interleukin (IL)-6 that was uniquely measured only in the postexposure pools. This finding is consistent with known cellular effects of ionizing radiation at the transcriptional level. A major transcription factor known to be activated by ionizing radiation is nuclear transcription factor-κB (NF-κB; refs. 33, 34). Genes induced by NF-κB following irradiation include intercellular adhesion molecule-1, galectin-3, tumor necrosis factor-α (TNF-α), IL-1β, and IL-6. In a study by Fedorocko et al. (35), mice exposed to a single 9 Gy dose of radiation were found to have elevated serum levels of IL-6 and TNF-α peaking at 5 days following exposure. This is further supported by a clinical study of patients receiving radiotherapy to their liver, where IL-6 was found to be variably elevated in serum at 24 and 48 hours postexposure (7).

The SELDI-TOF method described in this study has been optimized and adapted to a robot for the purpose of reproducibility

Table 3. Sensitivity and specificity of classifier methods for distinguishing high-dose-volume exposure from low dose-volume exposure in SELDI-TOF serum spectra (*n* = 68)

Classifier method	Sensitivity (%)	Specificity (%)	Discriminating ions
C5.0 decision tree	83	91	7,889; 7,961; 3,872
Hybrid: CART + C5.0 decision tree	100	100	7,870; 7,621; 7,195; 7,639; 7,657; 7,691; 8,152; 7,889; 7,661; 7,946; 8,037; 4,464; 7,682; 3,876; 8,295; 7,740; 8,155
Hybrid: neural network + C5.0 decision tree	100	97	7,691; 7,976; 7,993; 8,600

Table 4. Unique protein ions detected in unexposed and exposed pooled serum samples using MS

Unexposed pooled serum	Radiation exposed pooled serum
40S Ribosomal protein S4, Y isoform	Adapter-related protein complex 3δ1 subunit
60S Ribosomal protein L24 (L30)	α-2-Antiplasmin precursor
Adaptor molecule-1	Bomapin
α-2-HS-glycoprotein precursor	CRISPP peptide
α-Amylase 2B precursor	C-jun-amino-terminal kinase interacting protein 3
Apolipoprotein B-100 precursor	Collagen α1(I) chain precursor
Apolipoprotein C-I precursor	Complement factor H-related protein 1 precursor
Apolipoprotein M	Complement factor H-related protein 2 precursor
AT2 receptor-interacting protein 1	Calgranulin B
Baculoviral IAP repeat-containing protein 6	Crumbs protein homologue 1 precursor
Calgranulin B (migration inhibitory factor-related protein 14)	Hepatocellular carcinoma autoantigen
Class I MHC-restricted T cell–associated molecule	Inter-α-trypsin inhibitor heavy chain H1 precursor
Coagulation factor VIII precursor	IFN-inducible protein AIM2
Complement C1q subcomponent, B-chain precursor	IL-6 precursor
Complement component C9 precursor	MutS protein homologue 5
Corticosteroid-binding globulin precursor	N2B-titin isoform
Cytoplasmic protein NCK1	Nitric oxide synthase, inducible IID
Dedicator of cytokinesis protein 5 (fragment)	Serum amyloid A protein precursor
Fanconi anemia group G protein	Spectrin α chain, brain
Ficolin 3 precursor	Synaptotagmin-like protein 2 (exophilin 4)
Gag-Pro-Pol protein (Gag protein)	Tropomyosin 1 α chain (α-tropomyosin)
Gelsolin precursor, plasma	Vitronectin precursor
Gem-interacting protein	Zinc finger homeobox protein 2
Hemoglobin δ chain	
HLA class I histocompatibility antigen, α chain G	
Hypothetical protein CGI-96	
Hypothetical protein FLJ33718	
Immunoglobulin α-2 chain C region	
Immunoglobulin heavy chain V-III region BRO	
Immunoglobulin κ chain V-I region AG	
Immunoglobulin κ chain V-I region BAN	
Immunoglobulin κ chain V-I region DEE	
Immunoglobulin κ chain V-I region Ni	
Immunoglobulin κ chain V-I region OU	
Immunoglobulin κ chain V-II region Cum	
Immunoglobulin κ chain V-II region MIL	
Immunoglobulin κ chain V-III region NG9 precursor	

Table 4. Unique protein ions detected in unexposed and exposed pooled serum samples using MS (Cont'd)

Unexposed pooled serum	Radiation exposed pooled serum
Immunoglobulin κ chain V-IV region precursor	
Immunoglobulin λ chain V-III region SH	
Immunoglobulin λ chain V-VI region WLT	
Immunoglobulin μ chain C region	
Immunoglobulin μ heavy chain disease protein (BOT)	
Inter-α-trypsin inhibitor heavy chain H2 precursor	
Kinesin-like protein KIF17	
Lactotransferrin precursor	
Laminin α-5 chain precursor	
Lipin 1	
Melastatin 1	
Mitochondrial trifunctional protein β subunit (fragment)	
Myosin-reactive immunoglobulin heavy chain variable region	
Myosin-reactive immunoglobulin light chain	
Nuclear receptor corepressor 1	
Organic anion transporter 4	
Paraoxonase 3	
Phosphatidylcholine-sterol acyltransferase precursor	
Phospholipid transfer protein precursor	
Plasma retinol-binding protein precursor	
Platelet factor 4 variant precursor	
PMS1 protein homologue 2	
Potassium voltage-gated channel subfamily H member 2	
Probable G protein-coupled receptor GPR32	
Protein kinase C, ϵ type	
Protein Plunc precursor	
Putative breast epithelial stromal interaction protein	
R26660_2, partial CDS (fragment)	
RRP5 protein homologue (programmed cell death protein 11)	
Secretogranin III precursor	
Selenoprotein P precursor	
Sex hormone-binding globulin precursor	
Sialic acid synthase	
Smad ubiquitination regulatory factor 2	
SmcY protein (histocompatibility Y antigen)	
Stonin 2 (stoned B)	
Striatin 4 (zinedin)	
T-cell activation WD repeat protein	
Tetranectin precursor	
Thyroid receptor interacting protein 13	
Titin Novex-1 isoform (fragment)	
T-lymphoma invasion and metastasis inducing protein 1	
Transforming growth factor-β–induced protein IG-H3	
TNF precursor (TNF-α)	
Zinc finger protein 10 (KOX 1)	

and enabling of a high-throughput test. This gives precision of pipetting, reproducibility of pipette dispensing position, and ensures accurate timing. In addition, the use of barcodes on both samples and Ciphergen arrays allows for traceability of sample location. The time for processing a serum sample is ~2.5 hours. Finally, once a classifier is formulated and validated, the computational or wall time required to classify an unknown spectrum is a few milliseconds.

The low-molecular-weight serum proteome is a newly discovered information archive reflective of metabolism and pathologic events. In the recent past, it has been successfully exploited for early diagnosis of ovarian cancer as well as diagnosis of other cancers. This study adapted these established principles for the successful discrimination of both radiation exposure as well as the degree of radiation exposure. Applications of such a high-throughput blood test may include personalized proteomic measures of clinical response to radiation therapy and rapid response measures for populations exposed to moderate doses of radiation. Further refinement of the methods reported in this study may lead to personalized molecular medicine approaches directly applicable to adjuvant cancer care (36). Instead of categorical radiation dosing techniques, a strategy using a serum proteomic test, with the aim to better quantify the biological effect of serial exposures, may better avoid toxicity while maximizing therapeutic efficacy.

Conclusion

The composition of proteins in serum changes with ionizing radiation exposure. Proteomic analysis for the discovery of clinical biomarkers of radiation exposure warrants further study.

Acknowledgments

References

1. Schneider J, Presek P, Braun A, Woitowitz HJ. Serum levels of pantropic p53 protein and EGF-receptor, and detection of anti-p53 antibodies in former uranium miners (SDAG Wismut). Am J Ind Med 1999;36:602–9.
2. Menz R, Andres R, Larsson B, Ozsahin M, Trott K, Crompton NE. Biological dosimetry: the potential use of radiation-induced apoptosis in human T-lymphocytes. Radiat Environ Biophys 1997;36:175–81.
3. Horneck G. Biological monitoring of radiation exposure. Adv Space Res 1998;22:1631–41.
4. Ahlers I. Changes in whole-body metabolic parameters associated with radiation. Adv Space Res 1994;14:531–9.
5. Obe G, Johannes I, Johannes C, Hallman K, Reitz G, Facius R. Chromosomal aberrations in blood lymphocytes of astronauts after long-term space flights. Int J Radiat Biol 1997;72:727–34.
6. Jack CI, Cottier B, Jackson MJ, Cassapi L, Fraser WD, Hind CR. Indicators of free radical activity in patients developing radiation pneumonitis. Int J Radiat Oncol Biol Phys 1996;34:149–54.
7. Wickremesekera JK, Chen W, Cannan RJ, Stubbs RS. Serum proinflammatory cytokine response in patients with advanced liver tumors following selective internal radiation therapy (SIRT) with (90)yttrium microspheres. Int J Radiat Oncol Biol Phys 2001;49:1015–21.
8. Hara R, Itami J, Komiyama T, Katoh D, Kondo T. Serum levels of KL-6 for predicting the occurrence of radiation pneumonitis after stereotactic radiotherapy for lung tumors. Chest 2004;125:340–4.
9. Amundson SA, Bittner M, Fornace AJ, Jr. Functional genomics as a window on radiation stress signaling. Oncogene 2003;22:5828–33.
10. Coleman CN, Blakely WF, Fike JR, et al. Molecular and cellular biology of moderate-dose (1-10 Gy) radiation and potential mechanisms of radiation protection: report of a workshop at Bethesda, Maryland, December 17-18, 2001. Radiat Res 2003;159:812–34.
11. Evans HJ, Buckton KE, Hamilton GE, Carothers A. Radiation-induced chromosome aberrations in nuclear-dockyard workers. Nature 1979;277:531–4.
12. Bianchi M, Bianchi NO, Brewen JG, et al. Evaluation of radiation-induced chromosomal aberrations in human peripheral blood lymphocytes *in vitro*: result of an IAEA-coordinated programme. Mutat Res 1982;96:233–42.
13. Dertinger SD, Chen Y, Miller RK, et al. Micronucleated CD71-positive reticulocytes: a blood-based endpoint of cytogenetic damage in humans. Mutat Res 2003;542:77–87.
14. Johann DJ, Jr., McGuigan MD, Patel AR, et al. Clinical proteomics and biomarker discovery. Ann N Y Acad Sci 2004;1022:295–305.
15. Amundson SA, Lee RA, Koch-Paiz CA, et al. Differential responses of stress genes to low dose-rate γ irradiation. Mol Cancer Res 2003;1:445–52.
16. Amundson SA, Fornace AJ, Jr. Monitoring human radiation exposure by gene expression profiling: possibilities and pitfalls. Health Phys 2003;85:36–42.
17. Amundson SA, Grace MB, McLeland CB, et al. Human *in vivo* radiation-induced biomarkers: gene expression changes in radiotherapy patients. Cancer Res 2004;64:6368–71.
18. Phillips RK, Dewhirst MW, Gillette EL, Clow DJ. Electrophoretic evaluation of sera from dogs with cancer. Am J Vet Res 1978;39:1482–6.
19. Conrads TP, Zhou M, Petricoin EF III, Liotta L, Veenstra TD. Cancer diagnosis using proteomic patterns. Expert Rev Mol Diagn 2003;3:411–20.
20. Petricoin EF, Liotta LA. SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. Curr Opin Biotechnol 2004;15:24–30.
21. Mehta AI, Ross S, Lowenthal MS, et al. Biomarker amplification by serum carrier protein binding. Dis Markers 2003;19:1–10.
22. Lowenthal MS, Mehta AI, Frogale K, et al. Analysis of albumin-associated peptides and proteins from ovarian cancer patients. Clin Chem 2005;51:1933–45.
23. Conrads TP, Fusaro VA, Ross S, et al. High-resolution serum proteomic features for ovarian cancer detection. Endocr Relat Cancer 2004;11:163–78.
24. Han J, Cai Y, Cercone N. Data-drive discovery of quantitative rules in relational databases. IEEE Trans Know Data Eng 1993;5:29–40.
25. Johann D, McGuigan M, Tomov S, Petricoin E, Liotta L. Towards a systems biology toolkit. In: Proceedings of the 17th IEEE symposium on computer-based medical systems (CBMS '04); 2004. p. 500–5.
26. Xiang R, Shi Y, Dillon DA, Negin B, Horvath C, Wilkins JA. 2D LC/MS analysis of membrane proteins from breast cancer cell lines MCF7 and BT474. J Proteome Res 2004;3:1278–83.
27. Wilmarth PA, Riviere MA, Rustvold DL, Lauten JD, Madden TE, David LL. Two-dimensional liquid chromatography study of the human whole saliva proteome. J Proteome Res 2004;3:1017–23.
28. Qian WJ, Liu T, Monroe ME, et al. Probability-based evaluation of peptide and protein identifications from tandem mass spectrometry and SEQUEST analysis: the human proteome. J Proteome Res 2005;4:53–62.
29. Peng J, Elias JE, Thoreen CC, Licklider LJ, Gygi SP. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. J Proteome Res 2003;2:43–50.
30. Zhou M, Lucas DA, Chan KC, et al. An investigation into the human serum "interactome." Electrophoresis 2004;25:1289–98.
31. Goans RE, Holloway EC, Berger ME, Ricks RC. Early dose assessment in criticality accidents. Health Phys 2001;81:446–9.
32. Amundson SA, Bittner M, Meltzer P, Trent J, Fornace AJ, Jr. Biological indicators for the identification of ionizing radiation exposure in humans. Expert Rev Mol Diagn 2001;1:211–9.
33. Raju U, Gumin GJ, Tofilon PJ. Radiation-induced transcription factor activation in the rat cerebral cortex. Int J Radiat Biol 2000;76:1045–53.
34. Brach MA, Hass R, Sherman ML, Gunji H, Weichselbaum R, Kufe D. Ionizing radiation induces expression and binding activity of the nuclear factor κB. J Clin Invest 1991;88:691–5.
35. Fedorocko P, Egyed A, Vacek A. Irradiation induces increased production of haemopoietic and proinflammatory cytokines in the mouse lung. Int J Radiat Biol 2002;78:305–13.
36. Liotta LA, Kohn EC, Petricoin EF. Clinical proteomics: personalized molecular medicine. JAMA 2001;286:2211–4.