LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-JRNL-218036

Chemical and biological differentiation of three human breast cancer cell types using time-of-flight secondary ion mass spectrometry (TOF-SIMS)

K. S. Kulp, E. S. F. Berman, M. G. Knize, D. L. Shattuck, E. J. Nelson, L. Wu, J. L. Montgomery, J. S. Felton, K. J. Wu

January 9, 2006

Analytical Chemistry

Chemical and biological differentiation of three human breast cancer cell types using time-of-flight secondary ion mass spectrometry (TOF-SIMS)

Kristen S. Kulp[1]*
Elena S.F. Berman[1]
Mark G. Knize[1]
David L. Shattuck[1]
Erik J. Nelson[2]
Ligang Wu[1]
Jennifer L. Montgomery[1]
James S. Felton[1]
Kuang Jen Wu[2]

[1]Biosciences Directorate
P.O. Box 808
Lawrence Livermore National Laboratory
Livermore CA, 94551

[2]Chemistry and Material Sciences Program
Lawrence Livermore National Laboratory
Livermore CA, 94551

*Author to whom correspondence should be addressed:

Kristen S. Kulp
[1]Biosciences Directorate
P.O. Box 808
Lawrence Livermore National Laboratory
Livermore CA, 94551
925 422-6351

kulp2@llnl.gov

Abstract

We use Time-of-Flight Secondary Ion Mass Spectrometry (TOF-SIMS) to image and classify individual cells based on their characteristic mass spectra. Using statistical data reduction on the large data sets generated during TOF-SIMS analysis, similar biological materials can be differentiated based on a combination of small changes in protein expression, metabolic activity and cell structure. We apply this powerful technique to image and differentiate three carcinoma-derived human breast cancer cell lines (MCF-7, T47D and MDA-MB-231). In homogenized cells, we show the ability to differentiate the cell types as well as cellular compartments (cytosol, nuclear and membrane). These studies illustrate the capacity of TOF-SIMS to characterize individual cells by chemical composition, which could ultimately be applied to detect and identify single aberrant cells within a normal cell population. Ultimately, we anticipate characterizing rare chemical changes that may provide clues to single cell progression within carcinogenic and metastatic pathways.

To distinguish cancerous cells from their normal counterparts, it is necessary to detect subtle chemical differences between cells. Traditionally, pre-cancerous and cancerous lesions have been identified histopathologically, based on recognizing specific, known patterns of disease progression.[1] More recently, characteristic patterns of gene expression measured by DNA microarrays as well as mass spectrometry of protein samples have been used to classify tumors into clinically relevant subgroups.[2-11] These studies have demonstrated that no single change in protein or gene can adequately identify a cell or tissue that will become cancerous. Therefore, detection methods must be able to analyze and interpret a large number of variables that together suggest progression from normal. At present, no method has been shown to both detect and image these types of changes in single cells. We are developing Time-of-Flight Secondary Ion Mass Spectrometry (TOF-SIMS) techniques to image and classify individual cells based on their characteristic mass spectra.

TOF-SIMS is a highly sensitive surface analysis technique that can be used to detect the chemical composition of a biological sample and in the case of an individual cell, produce a mass-spectral map of the spatial distribution of abundant ions at the sub-cellular level. TOF-SIMS measurements use a finely focused (~150 nm) energetic primary ion beam to desorb secondary molecular and fragment ions into a time-of-flight mass spectrometer. Because the mass spectrum from each cell captures the mass signatures of both known and unknown molecules, no prior knowledge of the target analyte(s) is needed. The technique has high spatial resolution and is highly sensitive; thousands of spectra can be measured from a 20 μm diameter mammalian cell and as little as 40 zeptomole (10^{-21} mole) of an analyte can be detected from the surface area of a typical eukaryotic cell.[12]

Several groups have been successful in identifying intracellular distributions of specific biological ions such as sodium, potassium, calcium and membrane lipid fragments using TOF-SIMS

technology.[13-22] In human breast cells exposed to the carcinogen PhIP, the target compound was found in detectable amounts within the outer leaflet membrane of the cells. However, no carcinogen was found in the cytoplasm or nucleus, despite the known DNA-binding properties of the compound.[12] Similarly, analysis of yeast cells that had been exposed to the drug clofazimine demonstrated the presence of the drug within the cells, although not at the concentrations expected.[19] Our approach in this work is not to use specific ions to characterize cells but rather the full complement of detectable ion fragments.

The strength of the ToF-SIMS technique arises from the data generated; each image is composed of 256 x 256 individual spectra, which provide highly detailed chemical maps of the sample being analyzed. However, the spectra of biological samples are extremely complex and difficult to interpret, due to the contribution of secondary ions that are generated from fragmentation of larger molecules within the sample and matrix effects that change the secondary ion yield depending on the chemical environment of the surface being sampled.[23] Further, because most of a cell's mass is comprised of proteins, which are composed of only 20 fairly homogenously distributed amino acids, there is a lack of unique peaks among different biological samples. In fact, mass spectra obtained from different types of biological materials qualitatively appear very similar. It is not possible by looking at a TOF-SIMS spectrum to select a subset of peaks that uniquely identify the sample from which the spectra are generated. Therefore, data reduction techniques must be used to analyze the large spectral data sets generated by TOF-SIMS to differentiate similar biological materials.

Principal Component Analysis (PCA) has been used to identify similarities and differences in TOF-SIMS spectra and classify spectra into groups.[23-25] PCA, a standard, unsupervised multivariate statistical technique, reduces a large data matrix to a few manageable variables that can be visualized and interpreted using a series of simple plots. PCA reduces the data by finding new axes,

which represent linear combinations of the original data, called principal components (PCs) that capture the greatest variation in the data set. The masses with the highest variance among the statistical groups are designated as responsible for the differences between samples.

Wagner and Castner have used PCA and singular value decomposition to successfully cluster TOF-SIMS mass spectra generated from samples of single proteins and from samples of alkanethiol self-assembled monolayers, adsorbed onto gold substrates.[23, 26-28] Statistical analysis of TOF-SIMS spectra has also been employed to distinguish three species of freeze-dried yeasts based on membrane phospholipids[29] and to discriminate four yeast strains based on composite spectra from samples of yeast cultures.[30] Vegetative *Bacillis* cells were discriminated from spores based on TOF-SIMS analysis of phospholipid fragments.[31]

In a unique application of TOF-SIMS to the problem of cancer cell differentiation, we demonstrate the ability to image and identify individual cells from three human breast cancer cell lines, MCF-7, T47D, and MDA-MB-231. These three cell lines represent well-established models of different breast cancer phenotypes; MCF-7 and T47D are estrogen receptor positive (ER+) [32, 33] and MDA-MB-231 is estrogen receptor negative (ER-).[34] We also confirm the ability of TOF-SIMS and PCA to analyze and differentiate pure proteins and use the loadings plot generated by this analysis to give biological relevance to the separations. We show that PCA data reduction of TOF-SIMS spectra can differentiate cellular compartments (cytosol, nuclear and particulate) within the cell types. We are also able to differentiate homogenates from among the three breast cancer cell lines. The technique described here could ultimately be applied to detect a single aberrant cell within a normal cell population and to characterize rare chemical changes that may provide clues to single cell progression within carcinogenic and metastatic pathways.

Experimental Section

Amino Acid and Protein Analysis. Twenty amino acids and seven proteins (albumin, cytochrome C, ferritin, hemoglobin, insulin, lysosyme, and myoglobin), were obtained from Sigma (St.Louis, MO). The standards were diluted in water to an approximate concentration of 1mg/ml, spotted on silicon (Si) wafers and allowed to evaporate at room temperature. The amino acid composition of each of the proteins was based on the downloaded sequences in FASTA format from the protein's PDB (Protein Data Bank) identification code. The code used for each protein was: 1PID---despentapeptide insulin, 1A3N---Deoxy Human Hemoglobin, 1AEW---L-chain Horse apoferrin, 1AKK---Oxidized Horse Heart Cytochrome C, 1AO6---Human Serum Albumin and 1AZF---Chicken Egg White Lysozyme

Cell Culture. MCF-7, T47D, and MDA-MB-231 human breast cancer cells were obtained from American Type Culture Collection (ATCC, Manassas, VA). MCF-7 cells were grown in DMEM (Dulbecco's Modified Eagle Medium) with 5% fetal bovine serum (FBS), 1% non-essential amino acids, 10 µg/ml insulin, 2 mM L-glutamine, and 1% penicillin/streptomycin. T47D cells were grown in RPMI 1640 medium with 2 mM L-glutamine adjusted to contain 1.5 g/L sodium bicarbonate, 4.5 g/L glucose, 10 mM HEPES, and 1.0 mM sodium pyruvate and supplemented with 0.2 units/ml bovine insulin, 1% penicillin/streptomycin and 10% FBS. MDA-MB-231 cells were grown in Leibovitz's L-15 medium with 2 mM L-glutamine, 1% penicillin/streptomycin and 10% FBS. All cells were maintained at 37°C with 5% CO_2. All tissue culture supplies were obtained from Invitrogen (Carlsbad, CA).

For cell homogenization experiments, 2×10^6 cells were plated in T75 flasks and were harvested 48h later, when the cells were 75% confluent. For whole cell analysis, 8×10^5 cells were plated in a 60 mm dish containing 3 to 5 silicon chips, each about 1 cm square. The Si chips were sterilized by UV irradiation prior to seeding. Cells were grown on the polished side of the silicon chips;

no change was observed in cellular growth or morphology as compared to cells grown on the typical plastic-cell-culture ware. Cells grown on chips were freeze-fractured 48 h after plating.

Sample preparation. For homogenate experiments, cell populations were washed twice with 5 ml phosphate buffered saline (PBS) before being scraped into 3 ml PBS and transferred to a centrifuge tube. Centrifuging at 450 x *g* for five minutes pelleted the whole cells and the supernatant was discarded. The pelleted cells were resuspended in 1 ml of hypotonic lysis buffer (10mM HEPES, pH 7.9, 15mM $MgCl_2$, 10mM KCl, 1mM dithiothreitol, and Sigma protease inhibitor cocktail P 8340) and incubated for 15 minutes to allow the cells to swell. The cells were again pelleted by centrifugation, the supernatant discarded, and the pellet resuspended in 400 µl of lysis buffer. The swollen cells were lysed by passing through a 27-gauge hypodermic needle five times and the resulting solution was then centrifuged at 11,000 x *g* for 20 minutes. Trypan blue exclusion was used to monitor cell lysis in a small aliquot of cells. The supernatant, which is the cytoplasmic fraction, was removed and placed on ice. The pellet (nuclei and membranes) was then resuspended in 200 µl lysis buffer and shaken gently for 30 minutes. The nuclei were disrupted by passing through a 27-gauge hypodermic needle five times and this solution was then centrifuged at 20,000 x *g* for five minutes. The supernatant (nuclear fraction) was then removed, placed on ice and the pellet (particulate fraction) was again resuspended in lysis buffer. Less than 1 µl of each homogenate fraction was spotted on a silicon chip using a micropette with disposable tips. The samples were allowed to air dry and were stored at room temperature until TOF-SIMS analysis. Hypotonic buffer and mechanical disruption were used to lyse the cells because lysis with detergent buffers produces “sticky” spots that yield low ion counts in TOF-SIMS analysis. The experiment was repeated three times.

The samples for single cell analysis were prepared using a sandwich freeze-fracture method.[35] Approximately 10,000 latex beads (6 μm diameter, Duke Scientific Corp. (Palo Alto, CA)) in 10 μl PBS were added to the cells growing on Si chips immediately before sandwiching and fracturing. The beads act as spacers between the chips. The chips with attached cells were sandwiched with a clean silicon piece and flash-frozen in liquid nitrogen-cooled isopentane for 1 minute. The frozen sandwich was transferred to a liquid-nitrogen-cooled metal block and snapped apart with a razor blade. The substrate side of the sandwich was kept frozen on dry ice and lyophilized overnight. The samples were stored in a desiccator at room temperature until analysis by TOF-SIMS.

TOF-SIMS Analysis. The TOF-SIMS measurements were conducted on a PHI TRIFT III instrument (Physical Electronics USA, Eden Prairie, MN) equipped with a gallium ($^{69}Ga^+$) liquid metal ion gun operated at 25kV. A pulsed low-energy electron gun provided charge neutralization for all samples. Positive ion TOF-SIMS spectra were generally acquired over an area of 100 x 100 μm. The samples were held at room temperature during the course of the TOF-SIMS measurements. All TOF-SIMS spectra were calibrated to the CH_3^+, $C_2H_3^+$, and $C_4H_7^+$ peaks before PCA analysis. Five to ten spectra were recorded for each amino acid, protein or cell homogenate spot. For whole cell analysis, six or seven cells of each cell type showing the exclusion of sodium from the cellular area, which is evidence of fracturing an intact cell,[35] were imaged and the average spectrum encompassing the whole cell was recorded.

PCA analysis. Unit mass binning was applied to each spectrum before further analysis. For data reduction of the proteins, masses 50 to 500 of each TOF-SIMS spectrum were normalized to the total ion count of these masses for that spectrum. For cell homogenates and individual cells, *m/z* = 69

(implanted gallium ions) and *m/z* = 73, 147, and 207 (PDMS contamination peaks) were removed from the data set and then masses 58 through 500 were normalized to the total ion count of these masses for each spectrum. All resulting data matrices were mean-centered and reduced by principal component analysis (PCA) using MATLAB software v. 7.0 (MathWorks Inc., Natick, MA) along with PLS Toolbox v. 3.5 (Eigenvector Research, Manson, WA). The PCA software generates a scores plot to visualize data relationships and a loadings plot to determine masses important for sample differentiation. For the homogenate samples, *m/*z = 72, a strong fragment from the protease inhibitor, was removed from the data set for PCA analysis. Ninety percent confidence ellipses were drawn using the error_ellipse.m code by J. Andrew Johnson of Binghamton University, acquired from the MATLAB Central File Exchange.

Results and Discussion

TOF-SIMS analysis of pure protein standards

To serve as a reference for interpreting the peaks important in the separation of the samples, mass spectra were taken from all 20 naturally occurring amino acids. Table 1 lists the molecular weight of each amino acid and the major or diagnostic fragments from the mass spectra. As low mass fragments tend not to be chemically specific, only amino acid fragments with masses greater than 50 are presented. The higher masses contain the unique amino acid fragments that we expect to be characteristic of a particular amino acid. This analysis utilizing a Ga^+ Liquid Metal Ion Gun (LMIG) compares well with earlier amino acid analyses by Wagner et al. using a Cs^+ LMIG.[26, 28] The advantage of the Ga^+ LMIG is that molecular ion peaks are abundant and can be employed as diagnostic peaks in addition to the fragments identified by Wagner et al.

To confirm the data reduction scheme that would eventually be used on individual cells, positive ion TOF-SIMS mass spectra from pure protein standards were reduced using PCA analysis. Seven proteins: albumin, cytochrome C, ferritin, hemoglobin, insulin, lysosyme, and myoglobin, were analyzed. The amino acid composition of each of the proteins used in the analysis, based on the sequences obtained from the PDB database is provided in Table 2. Figure 1A shows the PCA data reduction scores plot for the seven proteins. The numbers in parentheses on each axis are the percent of variance captured by that axis. In this plot, PC 1 and PC 2 capture over 74% of the variance of the data set. The TOF-SIMS mass spectra of the proteins are well separated as shown by the minimal overlap in the 90% confidence limit ellipses.

The loadings plot, figure 1B, shows the masses that are important for separating the spectra of the proteins. By identifying the ions in the loading plot, we can demonstrate that the masses that represent the largest variance in the spectra are those that correlate to known properties of the proteins. Iron (*m/z* = 56) clearly separates the proteins on PC1, from lowest to highest iron content. Ferritin, an iron storage protein that may contain up to 4500 iron atoms [36] is well separated from the remaining proteins on the positive side of PC1. Cytochrome C, myoglobin, and hemoglobin, which each contain the same iron-porphyrin prosthetic group (heme) that binds one iron atom,[37] are further separated from lysosyme, albumin, and insulin, which contain no iron. Clearly, iron content explains much of the observed differentiation of these proteins.

A comparison of the ions identified in the loadings plot as responsible for the separation of the proteins to the ions found in the fragmentation of the pure amino acid standards (Table 1), suggests the identity of several of the ions important for the protein separation along PC2. The ion at *m/z* = 72, which is predominant in the valine spectrum, is responsible for the separation of insulin, hemoglobin, and albumin from the other proteins. Table 2 confirms that these proteins have the highest

percentage of valine residues. Another ion that contributes to the separation on PC2 is *m/z* = 55, which is seen in both lysine and arginine spectra. As shown in Table 2, Cytochrome C and myoglobin have the largest percentage of lysine residues and lysosyme has the largest percentage of arginine. The peaks at *m/z* = 57 (a common minor amino acid fragment), *m/z* = 110 (an ion seen in the histidine spectrum), *m/z* = 120 (an ion seen in the threonine and phenylalanine spectra), and *m/z* = 73 and 81, are also important for the protein separation. Thus, Figure 1 demonstrates that the TOF-SIMS spectra from pure protein standards group well together. Although the fragment ions from free amino acids could be different from the fragment ions from proteins, it appears that the differentiation of these proteins along PC2 can be explained by free amino acid fragments and subtle differences in the protein amino acid composition.

Wagner et al. have also performed TOF-SIMS analysis of pure protein standards using a Cesium ion source.[26, 28, 38] These studies showed grouping of the protein spectra and correlations to relevant amino acid fragments in the loadings plots using PCA analysis. In addition, two other multivariate supervised methods, linear discriminant analysis (LDA) and discriminant principal component analysis (DPCA) were used for data reduction.[38] The PCA method that we describe is identical to that used by Wagner et al., confirming their results. We have chosen to use PCA for our analyses because it is a well-accepted, straightforward multivariate analysis technique, and is sufficient for our purposes. In addition, because PCA is an unsupervised technique, the samples are not identified in the analysis and there are no a priori assumptions as to the grouping of spectra. Therefore, spectral grouping and separation by PCA stems solely from the spectral data and is very robust. For more complicated samples, it may be necessary to use a supervised data reduction technique to effect a sufficient separation.

TOF-SIMS analysis of human breast cancer cell homogenates

The initial analysis of the human cell lines was done on cell homogenates to ensure consistent access to the cell's contents and to determine chemical characteristics of the intracellular compartments of the cell lines. Low mass peaks (below *m/z* = 58) were removed from the PCA analysis to eliminate those peaks which are abundant and yet not chemically specific. Figure 2A shows the scores plot of the first two principal component axes of the positive ion spectra taken from the cytosolic, nuclear and particulate fractions of the 3 human breast cancer cell lines. These groupings, which depict both the variance within the groups and the relationship among the groups, demonstrate the chemical similarities and differences of the various fractions within and among the cell types. In this analysis, PC1 and PC2 together capture more than 81% of the variation among data groups. The ellipses drawn around each group represent the 90% confidence limit for that group. The variation seen within the groups, which results in scattered data points and larger confidence ellipses, has several sources. Some is instrument related; slight variation in TOF-SIMS spectra is found even during the analysis of homogenous solid surfaces. However, the major contribution to the variation comes from the samples themselves. Spots of cell homogenates are not uniform due to local protein aggregation and crystallization effects that occur as the liquid homogenates dry.

In Figure 2A, PC 1 separates the particulate fractions from the cytosolic and nuclear fractions of each cell line. This suggests that the small molecules and fragments measured from the spots of the particulate fractions, which are composed of the insoluble membranes and structural proteins of the cell, are chemically distinguishable from the cytosolic and nuclear fractions. The spectra measured from the T47D particulate spot are highly variable, as evidenced by the spectral scatter

and the large 90% confidence oval for that group. This large variation may be a result of lack of uniformity of this sample.

Although the particulate fractions of each of the cell types are well separated from the cytosolic and nuclear fractions along PC1, and the particulate fractions of the 2 ER+ cell lines (MCF7 and T47D) are well-separated from each other, the 90% confidence ovals for the particulate fractions of the MDA-MB-231 and T47D are overlapping. This could be due to the large variation seen in the T47D particulate spectra, or it may be that the membranes and structural proteins of these 2 cell lines are more chemically similar to each other than to the MCF-7 cells.

Because the MDA-MB-231, MCF-7 and T47D cell lines are standard models for breast cancer mechanistic studies, they have been well-characterized in the literature using traditional biological assays. The ER- MDA-MB-231 cell line has been demonstrated to be a highly metastatic and tumorigenic[39, 40] and is considered to express genes that represent a poor prognosis signature.[41] In contrast, the ER+ MCF-7 and T47D lines are classified as non-invasive.[42] Further, MCF-7 and T47D cells have been shown to express membranous E-cadherin and β-cadherin, which are powerful tumor suppressors.[42, 43] MDA-MB-231 cells do not express these proteins in their surface membranes.[42] Based on just these cited studies, we would expect that the MCF-7 and T47D particulate fractions would be more similar to each other, and different from the MDA-MB-231. However, although these studies are important descriptions of the cancer phenotype of these cell lines, they report specific differences among the cell lines that may not necessarily reflect more global differences in cell structure. In our analysis, we find that the cell types are well-separated, but the relationship of the groups does not correlate with the invasive potential of the cell line. This suggests that the differences among the cells that are important for TOF-SIMS analysis may not arise from specific, low abundant proteins discovered using traditional biological approaches.

Figure 2A also shows that the nuclear fractions of each cell type are primarily separated along PC2 and that these fractions are well separated from each other. Our cell homogenization protocol, produces nuclear fractions composed of small molecules, localized to the nucleus, and soluble nuclear proteins. The scores plot suggests that the chemical composition of these small molecules and small molecular fragments is sufficient to chemically differentiate among the 3 human breast cancer cell types.

Both multivariate axes, PC1 and PC2, are necessary to differentiate the cytosolic fractions of the cell lines. Again, as seen in the particulate analysis, the cytosolic portions of the T47D and MDA-MB-231 cells are slightly overlapping and more chemically similar to each other than to the cytosolic fractions of the MCF-7 cells. These groupings are also similar to the relationship we found among the nuclear fractions of cells. Interestingly, of the three cell compartments that we investigated, the cytosolic fractions of the cell lines were the most closely related. This suggests that there is less chemical difference in the cytosols of the cell lines than in the other cell compartments.

For all 3 cell lines, the nuclear and cytosolic fractions are closely related and for the MCF-7, the nuclear and cytosolic fractions are overlapping. A further PCA analysis of just the MCF-7 and T47D homogenate spectra (removing the MDA-MB-231 spectra which have the largest within group variation) again showed cell line dependent differentiation of the cytosolic fractions (MCF-7 and T47D cytosols were separated from each other), but no differentiation of the MCF7 cytosol and nuclear fractions (data not shown). The close relationship between the cytosolic and nuclear fractions is not unexpected. The centrifugation technique that we use to separate the cell compartments is essentially imprecise and it is likely that there are cytosolic cell components contaminating the nuclear fractions of the homogenates.

The loadings plot for the cell homogenates, Figure 2B, shows the mass peaks which are primarily responsible for the differentiation. Most of these peaks are fragments of hydrocarbons and therefore cannot be related to specific biological molecules. Some of the important peaks can be found in the spectra of pure amino acids, such as *m/z* = 70, which is found in the spectra of glutamic acid, leucine, proline, arginine and asparagine, and *m/z* = 86, which is found in leucine and iosoleucine spectra. Important peaks that are common in both the homogenate and whole cell analysis are marked with asterisks.

In contrast to the loadings plot for the proteins (Figure 1B), which provided information about the biological meaning behind the separation of the protein standards, the loadings plot for the cell homogenates is difficult to interpret. To date, little has been done to identify the small molecules or small molecular fragments that are important in biological samples. As can be seen from our results, these ions provide enough chemically specific information to differentiate specific cell types. More studies, focused on identifying the important masses in the loadings plot, need to be done before we can understand the biology that drives the separation of the cell types and cell compartments.

The results of the cell homogenate experiments were confirmed by successfully repeating the entire procedure three times over a course of 18 months. The data discussed here are representative of the results of these experiments. In addition, we performed PCA analysis of just the spectra taken from a single cell line (MCF-7) during that time and showed that although there was an expected increase in spectral variation, the molecular signature of the cell line remained essentially intact (data not shown). Given that cell culture lines are well-known for biological “drift” with changing passage number and that our data reflect cell population variations in cell cycle stage and cell density, a consistent small molecular signature for a single cell line suggests that the underlying chemical composition of a cell may be relatively constant and may be unique to that cell. Further,

once defined, this chemical signature could be used in supervised statistical techniques as a basis for identifying individual cells within a mixed cell population.

TOF-SIMS analysis of individual human breast cancer cells

Based on the successful separation of the homogenates of the cell lines, we applied a similar analysis technique to individual cells, utilizing the imaging capability of the TOF-SIMS instrument. Figure 3 shows total ion images of freeze-fractured MCF-7, MDA-MB-231, and T47D cells. Typically, T47D and MCF-7 cells display an epithelial cell phenotype and MDA-MB-231 cells have a more elongated, fibroblastoid shape. Although only one cell of each type is shown in Figure 3, under usual growth conditions these cell lines show tremendous morphological variation making them indistinguishable from each other using optical microscopy.

Mass spectral images were taken of six or seven fractured cells of each cell line and a composite mass spectrum of the entire cellular region was used for PCA analysis. The scores plot in Figure 4A shows a good separation of the MCF-7, MDA-MB-231, and T47D cell spectra, confirming our ability to differentiate cell types. TOF-SIMS spectra taken from the background region near each cell, and treated identically to the cell spectra, failed to separate into groups using PCA data reduction (data not shown).

In contrast to the results of the homogenate data, the spectra acquired from whole T47D cells were more similar to the MCF-7 cells, and not as closely related to the MDA-MB-231 cells. In fact, in the scores plot of the whole cells, the ER+ MCF-7 and T47D cells, which have a similar non-invasive phenotype, are closely related and well separated from the ER-, highly metastatic MDA-MB-231 cells. These results suggest that although the relationship of the cell compartments to each other did not

reflect the invasive potential of the cells, a composite analysis of all the cell's parts may provide better clues about the cancer phenotype.

The loadings plot for the individual cells (Figure 4B) shows that many of the significant masses driving the separations are important in both the cell and cell homogenate experiments. Of the ion peaks important in the differentiation, *m/z* = 184 results from the head group of phospholipids, and *m/z* = 70, 81, and 95 all appear in pure amino acid spectra suggesting they could have a protein origin. Again, as with the cell homogenate loadings plot, more work needs to be done to identify the origin of these small molecules. However, the large number of mass peaks common to the PCA separation of the cells and cell homogenates shows the expected consistency in TOF-SIMS-based cell differentiation.

Conclusion

A major advance in cancer biology in recent years has been the identification of thousands of proteins, genes, signaling molecules and their related pathways that are important for both normal and neoplastic cell growth.[44] Careful studies comparing normal cellular function to paired transformed counterparts have shown that cancer is a complex disease involving myriad molecular and cellular changes. The overarching conclusion from these previous studies has been that there is no single genetic or protein expression change or modification that can explain the development of cancer; rather it is a spectrum of subtle changes that defines the carcinogenic process.[45] Clearly, novel strategies must be developed to discover the global changes within individual cells that are mechanistically involved in cancer initiation and progression in order to develop detection, prevention and treatment strategies.

We have demonstrated the ability to differentiate 3 closely-related human breast cancer cell lines, MCF7, T47D, and MDA-MB-231, using PCA analysis of TOF-SIMS spectra acquired from individual cells. Even though the low *m/z* peaks in TOF-SIMS spectra are usually of unknown parent chemical origin, statistical characterization of an ensemble of these peaks represents an innovative method of studying and differentiating cell types at a molecular level. This study demonstrates the power of imaging mass spectrometry to discover differences among cells that may not be apparent using traditional methods. TOF-SIMS imaging is a discovery technology that can scan structural variations and functional changes in cells and provide molecular information about the genesis of cancer. This information could lead scientists to new drug development targets, detection techniques, prognosis predictors, and prevention strategies.

Acknowledgements:

The authors gratefully acknowledge the assistance of Dr. Brian Bennion in accessing the amino acid composition of the standard proteins. This work was performed under the auspices of the U.S. DOE by UC, LLNL under contract no. W-7405-Eng-48 and supported by NCI grants CA55861, CA BCRP 10IB-0077 and LDRD 04-ERD-104 (LLNL internal funding).

References

(1) Chaurand, P.; Sanders, M. E.; Jensen, R. A.; Caprioli, R. M. *Am J Pathol* **2004**, *165*, 1057-1068.
(2) Brentani, R. R.; Carraro, D. M.; Verjovski-Almeida, S.; Reis, E. M.; Neves, E. J.; de Souza, S. J.; Carvalho, A. F.; Brentani, H.; Reis, L. F. *Crit Rev Oncol Hematol* **2005**, *54*, 95-105.
(3) Goncalves, A.; Viens, P.; Sobol, H.; Maraninchi, D.; Bertucci, F. *Rev Med Interne* **2005**, *26*, 470-478.
(4) Meyerson, M.; Carbone, D. *J Clin Oncol* **2005**, *23*, 3219-3226.
(5) Ouellet, V.; Provencher, D. M.; Maugard, C. M.; Le Page, C.; Ren, F.; Lussier, C.; Novak, J.; Ge, B.; Hudson, T. J.; Tonin, P. N.; Mes-Masson, A. M. *Oncogene* **2005**, *24*, 4672-4687.

(6) Petricoin, E. F.; Ardekani, A. M.; Hitt, B. A.; Levine, P. J.; Fusaro, V. A.; Steinberg, S. M.; Mills, G. B.; Simone, C.; Fishman, D. A.; Kohn, E. C.; Liotta, L. A. *Lancet* **2002**, *359*, 572-577.
(7) Sorlie, T.; Perou, C. M.; Tibshirani, R.; Aas, T.; Geisler, S.; Johnsen, H.; Hastie, T.; Eisen, M. B.; van de Rijn, M.; Jeffrey, S. S.; Thorsen, T.; Quist, H.; Matese, J. C.; Brown, P. O.; Botstein, D.; Eystein Lonning, P.; Borresen-Dale, A. L. *Proc Natl Acad Sci U S A* **2001**, *98*, 10869-10874.
(8) Sorlie, T.; Tibshirani, R.; Parker, J.; Hastie, T.; Marron, J. S.; Nobel, A.; Deng, S.; Johnsen, H.; Pesich, R.; Geisler, S.; Demeter, J.; Perou, C. M.; Lonning, P. E.; Brown, P. O.; Borresen-Dale, A. L.; Botstein, D. *Proc Natl Acad Sci U S A* **2003**, *100*, 8418-8423.
(9) Wulfkuhle, J. D.; Liotta, L. A.; Petricoin, E. F. *Nature Reviews Cancer* **2003**, *3*, 267-275.
(10) Carr, K. M.; Rosenblatt, K.; Petricoin, E. F.; Liotta, L. A. *Hum Genomics* **2004**, *1*, 134-140.
(11) Liotta, L. A.; Petricoin, E. F.; Veenstra, T. D.; Conrads, T. P. *Endocr Relat Cancer* **2004**, *11*, 585-587.
(12) Quong, J. N.; Knize, M. G.; Kulp, K. S.; Wu, K. J. *Applied Surface Science* **2004**, *231-2*, 424-427.
(13) Chandra, S. *Journal of Microscopy-Oxford* **2001**, *204*, 150-165.
(14) Chandra, S.; Lorey, D. R.; Smith, D. R. *Radiation Research* **2002**, *157*, 700-710.
(15) Kempson, I. M.; Skinner, W. M.; Kirkbride, P. K. *Biochim Biophys Acta* **2003**, *1624*, 1-5.
(16) Roddy, T. P.; D.M. Cannon, J.; Ostrowski, S. G.; Winograd, N.; Ewing, A. G. *Anal. Chem.* **2002**, *74*, 4020-4026.
(17) Cannon, D. M.; Pacholski, M. L.; Winograd, N.; Ewing, A. G. *J. Am. Chem. Soc.* **2000**, *122*, 603-610.
(18) Cliff, B.; Lockyer, N. P.; Corlett, C.; Vickerman, J. C. *Applied Surface Science* **2003**, *203-204*, 730-733.
(19) Cliff, B.; Lockyer, N.; Jungnickel, H.; Stephens, G.; Vickerman, J. C. *Rapid Communications in Mass Spectrometry* **2003**, *17*, 2163-2167.
(20) Sjovall, P.; Lausmaa, J.; Johansson, B. *Anal Chem* **2004**, *76*, 4271-4278.
(21) Ostrowski, S. G.; Van Bell, C. T.; Winograd, N.; Ewing, A. G. *Science* **2004**, *305*, 71-73.
(22) Gazi, E.; Dwyer, J.; Lockyer, N.; Gardner, P.; Vickerman, J. C.; Miyan, J.; Hart, C. A.; Brown, M.; Shanks, J. H.; Clarke, N. *Faraday Discussions* **2004**, *126*, 41-59.
(23) Wagner, M. S.; Graham, D. J.; Ratner, B. D.; Castner, D. G. *Surface Science* **2004**, *570*, 78-97.
(24) Kargacin, M. E.; Kowalski, B. R. *Anal. Chem.* **1986**, *58*, 2300-2306.
(25) Pachuta, S. J. *Applied Surface Science* **2004**, *231-2*, 217-223.
(26) Wagner, M. S.; Castner, D. G. *Langmuir* **2001**, *17*, 4649-4660.
(27) Wagner, M. S.; Horbett, T. A.; Castner, D. G. *Langmuir* **2003**, *19*, 1708-1715.
(28) Wagner, M. S.; Castner, D. G. *Applied Surface Science* **2004**, *231-2*, 366-376.
(29) Lockyer, N. P.; Vickerman, J. C. *Applied Surface Science* **2004**, *231-2*, 377-384.
(30) Jungnickel, H.; Jones, E. A.; Lockyer, N. P.; Oliver, S. G.; Stephens, G. M.; Vickerman, J. C. *Analytical Chemistry* **2005**, *77*, 1740-1745.
(31) Thompson, C. E.; Jungnickel, H.; Lockyer, N. P.; Stephens, G. M.; Vickerman, J. C. *Applied Surface Science* **2004**, *231-2*, 420-423.
(32) Soule, H.; Vazguez, J.; Long, A.; Albert, S.; Brennan, M. *J. Natl Cancer Inst.* **1973**, *51*, 1409-1416.
(33) Keydar, I.; Chen, L.; Karby, S.; Weiss, F. R.; Delarea, J.; Radu, M.; Chaitcik, S.; Brenner, H. J. *Eur J Cancer* **1979**, *15*, 659-670.
(34) Cailleau, R.; Olive, M.; Cruciger, Q. *In Vitro* **1978**, *14*, 911-915.
(35) Chandra, S.; Morrison, G. H. *Biology of the Cell* **1992**, *74*, 31-42.
(36) Theil, E. C. *Annu. Rev. Biochem.* **1987**, *56*, 289-316.
(37) Stryer, L. *Biochemistry*; W.H. Freeman and Company: San Franciso, 1981.
(38) Wagner, M. S.; Tyler, B. J.; Castner, D. G. *Anal Chem* **2002**, *74*, 1824-1835.
(39) Price, J. E.; Polyzos, A.; Zhang, R. D.; Daniels, L. M. *Cancer Res* **1990**, *50*, 717-721.
(40) Zhang, R. D.; Fidler, I. J.; Price, J. E. *Invasion Metastasis* **1991**, *11*, 204-215.

(41) Minn, A. J.; Kang, Y.; Serganova, I.; Gupta, G. P.; Giri, D. D.; Doubrovin, M.; Ponomarev, V.; Gerald, W. L.; Blasberg, R.; Massague, J. *J Clin Invest* **2005**, *115*, 44-55.
(42) Raby, B.; Polette, M.; Gilles, C.; Clavel, C.; Strumane, K.; Matos, M.; Zahm, J.; Roy, R.; Bonnet, N.; Birembaut, P. *Int J Cancer* **2001**, *93*, 644-652.
(43) Mareel, M.; Berx, G.; Van Roy, F.; Bracke, M. *J Cell Biochem* **1996**, *61*, 524-530.
(44) Rodland, K. D. *Clin Biochem* **2004**, *37*, 579-583.
(45) Griffin, J.; Shockcor, J. *Nature Reviews Cancer* **2004**, *4*, 551-561.

Table 1. Amino acid molecular peak and six most abundant or diagnostic TOF-SIMS fragments in descending order.

Amino acid	MW	Most abundant	2nd	3rd	4th	5th	6th
Alanine	89	90	112	134	179	57	81
Arginine	174	70	55	175	57	59	73
Asparagine	132	133	155	177	87	56	70
Aspartic acid	133	134	155	219	279	57	88
Cysteine	121	55	122	219	241	57	76
Glutamic acid	147	175	148	70	102	84	130
Glutamine	146	147	84	130	101	56	279
Glycine	75	76	98	120	178	200	57
Histidine	155	110	156	82	95	69	55
Isoleucine	131	132	86	69	263	56	58
Leucine	131	132	86	70	55	57	263
Lysine	146	55	56	147	84	69	104
Methionine	149	61	166	120	150	56	91
Phenylalanine	165	166	120	91	51	103	77
Proline	115	116	70	231	68	117	138
Serine	105	106	128	60	57	91	116
Threonine	119	120	56	74	116	239	57
Tryptophan	204	130	159	188	117	205	143
Tyrosine	181	182	136	107	116	165	123
Valine	117	72	118	55	57	235	59

Table 2. Amino acid percent composition of selected proteins.

Amino acid	Cytochrome C	Insulin	Hemoglobin all chains	Lysozyme	Myoglobin	Albumin	Ferritin
Alanine	5.8	4.3	12.5	9.3	9.8	10.6	8.6
Arginine	1.9	2.2	2.1	8.5	1.3	4.1	6.3
Asparagine	4.8	6.5	3.5	10.9	1.3	2.9	3.4
Aspartic Acid	2.9	0.0	5.2	5.4	5.2	6.2	6.9
Cysteine	1.9	13.0	1.0	6.2	0.0	6.0	1.1
Glutamic Acid	8.7	8.7	4.2	1.6	8.5	10.6	8.6
Glutamine	2.9	6.5	1.4	2.3	3.9	3.4	6.3
Glycine	11.5	8.7	7.0	9.3	9.8	2.1	6.3
Histidine	2.9	4.3	6.6	0.8	7.2	2.7	3.4
Isoleucine	5.8	2.2	0.0	4.7	5.9	1.4	2.3
Leucine	5.8	13.0	12.5	6.2	11.1	10.4	15.5
Lysine	18.3	0.0	7.7	4.7	12.4	10.1	5.2
Methionine	1.9	0.0	1.0	1.6	1.3	1.0	1.7
Phenylalanine	3.8	6.5	5.2	2.3	4.6	5.3	4.6
Proline	3.8	0.0	4.9	1.6	2.6	4.1	1.7
Serine	0.0	6.5	5.6	7.8	3.3	4.1	5.7
Threonine	9.6	0.0	5.6	5.4	4.6	4.8	3.4
Tryptophan	1.0	0.0	1.0	4.7	1.3	0.2	0.6
Tyr (Y)	3.8	6.5	2.1	2.3	1.3	3.1	3.4
Valine	2.9	10.9	10.8	4.7	4.6	7.0	4.6

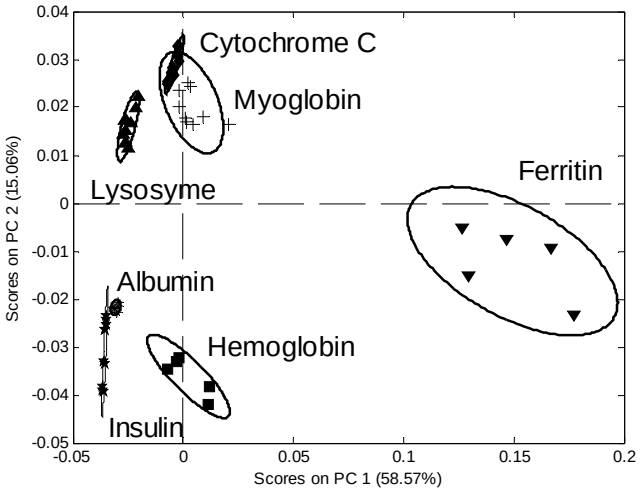
Figure Legends

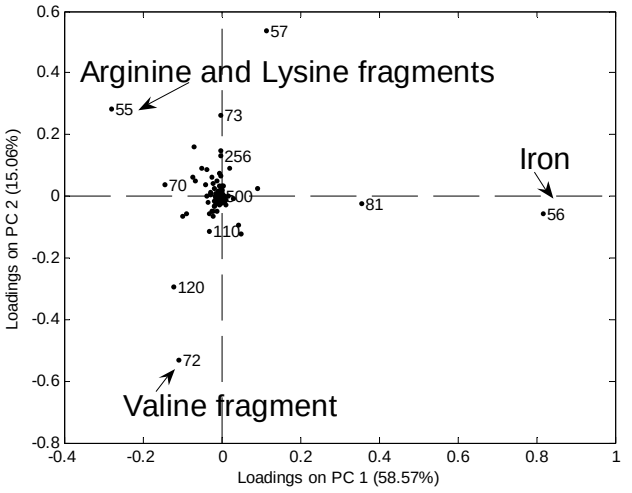
Figure 1. Scores plot (a) from PCA data reduction of TOF-SIMS positive ion spectra from proteins spotted on a silicon chip. Data points are multiple regions of interest from a single spot. Ellipses are 90% confidence intervals. Loadings plot (b) showing masses responsible for separation of protein spots.

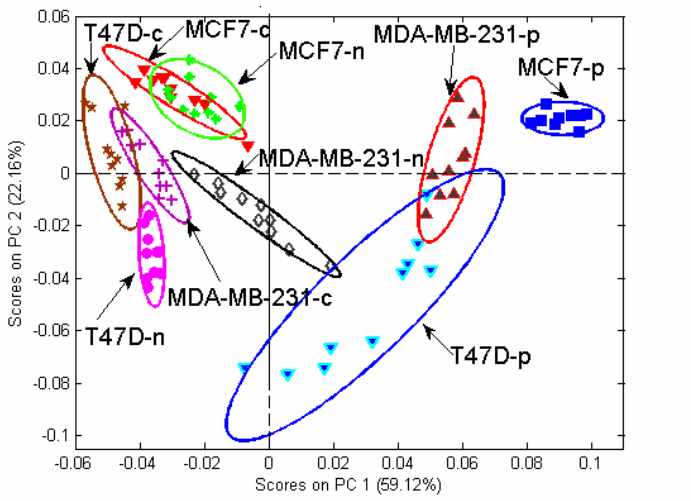
Figure 2. Scores plot (a) from PCA data reduction of TOF-SIMS positive ion spectra from cultured human cell homogenates separated by centrifugation into cytosol, nuclear, and insoluble particulate fractions. Ellipses are 90% confidence intervals. Loadings plot (b) showing masses responsible for separation of homogenate spots. Asterisks designate loading peaks that are common to the homogenate and the whole cell analysis.

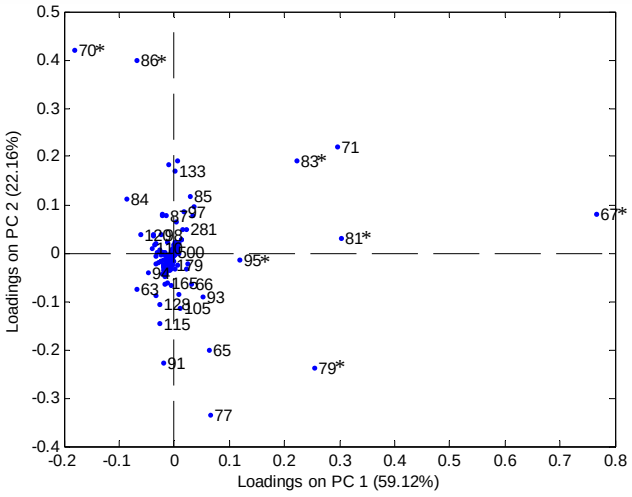
Figure 3. Ion images of cultured human MCF-7, MDA-MB-231 and T47D cells.

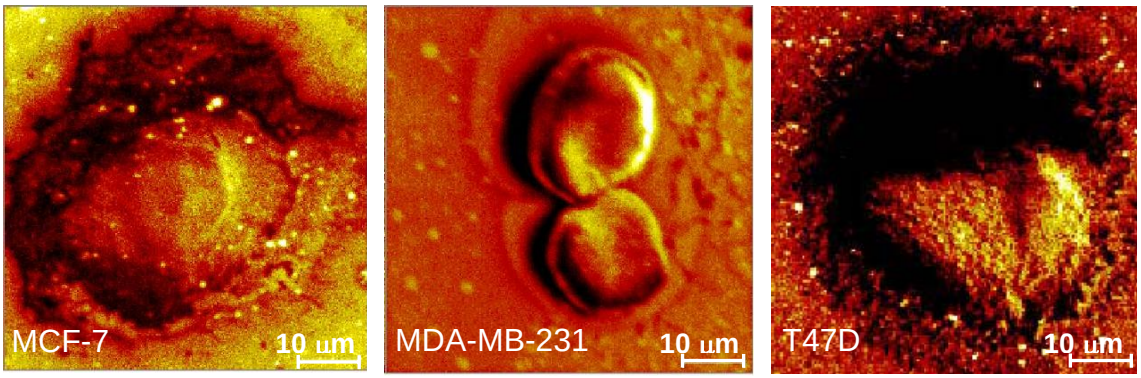
Figure 4. Scores plot (a) from PCA data reduction of TOF-SIMS positive ion spectra from freeze-fractured MCF-7, MDA-MB-231 and T47D cells. Ellipses are 90% confidence intervals. Loadings plot (b) showing masses responsible for separation of cells. Asterisks designate loading peaks that are common to the homogenate and the whole cell analysis.

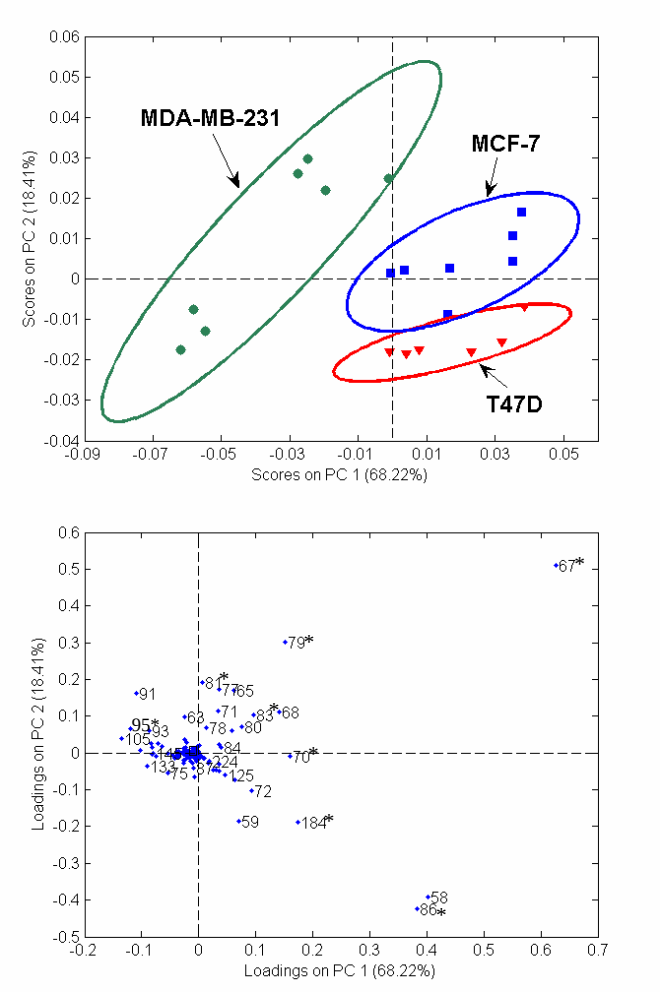

Cytochrome C
Myoglobin
Ferritin
Lysosyme
Albumin
Hemoglobin
Insulin
Scores on PC 2 (15.06%)
Scores on PC 1 (58.57%)


Arginine and Lysine fragments
Iron
Valine fragment
57
55
73
256
70
500
81
56
110
120
72
Loadings on PC 2 (15.06%)
Loadings on PC 1 (58.57%)


T47D-c
MCF7-c
MCF7-n
MDA-MB-231-p
MCF7-p
MDA-MB-231-n
MDA-MB-231-c
T47D-n
T47D-p
Scores on PC 2 (22.16%)
Scores on PC 1 (59.12%)


70*
86*
71
83*
133
84
85
67*
281
81*
95*
63
93
105
115
65
91
79*
77
Loadings on PC 2 (22.16%)
Loadings on PC 1 (59.12%)


MCF-7
10 um
MDA-MB-231
10 um
T47D
10 um


MDA-MB-231
MCF-7
T47D
Scores on PC 2 (18.41%)
Scores on PC 1 (68.22%)
Loadings on PC 2 (18.41%)
Loadings on PC 1 (68.22%)