

Raman spectral signatures of mouse mammary tissue and associated lymph nodes: normal, tumor and mastitis

Jagdish S. Thakur,^{1*} Houbai Dai,² Gulay K. Serhatkulu,¹ Ratna Naik,² Vaman M. Naik,³ Alex Cao,¹ Abhilash Pandya,¹ Gregory W. Auner,¹ Rajah Rabah,⁴ Michael D. Klein⁴ and Carl Freeman⁵

¹ Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI 48202, USA

² Department of Physics and Astronomy, Wayne State University, Detroit, MI 48202, USA

³ Department of Natural Sciences, University of Michigan-Dearborn, Dearborn, MI 48128, USA

⁴ Departments of Surgery (Pediatric Surgery) and Pathology, Wayne State University School of Medicine and the Children's Hospital of Michigan, Detroit, MI 48201, USA

⁵ Department of Biological Sciences, Wayne State University, Detroit, MI 48202, USA

Received 10 March 2006; Accepted 7 June 2006

Raman spectroscopy involves the interaction of light with the molecular vibrations and therefore can provide information about molecular structure, tissue composition and changes in its environment. We explored whether Raman spectroscopy can reliably distinguish mammary tumors from normal mammary tissues and other pathological states in mice. We analyzed a large number of Raman spectra from the tumor and normal mammary glands of mice injected with 4T1 tumor cells, which were collected using a high-resolution (less than 4 cm^{-1}) Raman spectrometer at a fixed (785 nm) laser excitation wavelength and with 60 mW of laser power. The spectra of normal and tumor mammary glands showed consistent differences in the intensity of certain Raman bands and loss of some bands in the tumor spectra. Multivariate statistical methods – principal component analysis (PCA) and discriminant functional analysis (DFA) – were used to separate the data into different groups of mammary tumors, mastitis, lymph nodes contralateral and tumor-cell-injected sides, and normal contralateral and tumor-cell-injected sides. We demonstrate that this spectroscopic technique has the feasibility of discriminating tumor and mastitis from normal tissues and other pathological states in a short period of time and may detect tumor transformation earlier than the standard histological examination stage. Copyright © 2006 John Wiley & Sons, Ltd.

KEYWORDS: mouse mammary glands; tumor spectral signature; spectral diagnosis; PCA; DFA

INTRODUCTION

Raman spectroscopy is widely used in nondestructive biochemical analyses. Recently it has been applied for the detection of different types of cancers.^{1–4} Raman spectroscopy has also been used for classification of tissues into cancerous and precancerous stages.^{5,6} This spectroscopic technique has several advantages over Fourier-transform infrared spectroscopy (FTIR)⁷ because any wavelength can be used as a source of excitation and it provides more detailed information about the biochemical components of the sample. Another disadvantage of FTIR for tissue

examination is the strong absorption of IR radiation by the large water content of the tissue. In Raman spectroscopy,⁸ a tissue sample is irradiated with a laser which interacts with the various molecular components of the tissue. This interaction of light with the molecular vibrational modes results in scattered light with a unique shift in the frequency (Raman shift) and thereby provides information about the molecular structure and composition of the sample.⁹ Each peak position in the Raman spectrum represents a vibrational mode of a specific group (such as $-\text{CH}_2$, $-\text{CH}_3$, $-\text{NH}$, peptide, etc.) of a biomolecule present in that particular environment of the tissue. The peak height (strictly the area under the peak) is directly related to its concentration in a tissue. Transformations of normal tissue into a tumor or any other state (e.g. inflammatory) are due to changes in the

*Correspondence to: Jagdish S. Thakur, Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI 48202, USA. E-mail: jagdish@wayne.edu

concentration of the specific biomolecules. These changes are in turn reflected in the Raman vibrational spectra.

One of the main foci of cancer research is to understand the molecular changes during neoplastic transformations and translate them into diagnostic and therapeutic tools. The ability of Raman spectroscopy to uniquely map the molecular changes in transformed tissues can be exploited for rapid detection of neoplastic transformation with high specificity and sensitivity compared to other conventional methods. Moreover, because it is noninvasive¹⁰ and requires no sample preparation, and also because the Raman bands of water are quite weak, Raman spectroscopy has the potential for real-time *in vivo* cancer diagnosis applications^{11,12} using an optical-fiber probe.¹³

Earlier breast studies based on small sample sets have identified some trends in the spectra.^{3,14,15} Here we sought to determine the ability of Raman spectroscopy to differentiate normal mouse breast tissues from those with cancer on a very large data set. We further investigated whether Raman spectroscopy can be used to detect metastases before they are evident histologically and can provide some information about the transformation of normal tissues into malignant ones. We analyzed 650 Raman spectra from mouse mammary tumors and normal mammary glands along with other pathological states or tissue types. Multivariate statistical techniques separated spectra of tumors and mastitis from normal mammary gland tissues and other pathological states.

EXPERIMENTAL

Animals and cell line

Female BALB/c mice of 6–8 weeks (The Taconic Farm, Germantown, NJ) were used in our experiments. All

animals were treated humanely and in full compliance with the National Institutes of Health (NIH) regulations and institutional guidelines as approved by the Animal Investigation Committee of the Wayne State University Department of Laboratory Animal Resources. The 4T1 mouse mammary tumor cell line was used to induce tumor formation. These cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum.

We used a Renishaw RM 1000 Raman microscope-spectrometer with the 785 nm near-infrared (NIR) excitation wavelength¹⁶ to record the Raman spectra. We selected this excitation wavelength to avoid the background fluorescence caused by shorter excitation wavelengths. Best spectral features¹⁴ are obtained with excitation wavelengths between 782 and 830 nm.

Induction of mammary tumors

Mammary tumors were generated in 11 BALB/c mice by orthotopic injection of 1×10^5 4T1 cells suspended in the DMEM medium into a mammary fat pad (one side only). Primary tumors were detected palpably within 10–14 days. Fourteen days after the injection of tumor cells, the mice were sacrificed, one at a time to get data from fresh tissues without any preservation from the normal mammary tissues, mammary tumors, and lymph nodes, from both tumor-cell-injected and contralateral sides. The details are given in the Table 1. Each sample was immediately divided into two halves, one was used for the Raman spectroscopy and the other was immediately fixed in 10% formalin for histological examination. Samples included 16 tissues that were tumors, 12 that appeared normal and 10 lymph nodes from the

Table 1. Tissue identification from pathology studies

Tissue name	A	B	C	D	E	F	G
Mouse1	M1	T1	LN1	M2	LN2	T2	–
Mouse2	M1	T	LN1	M2	LN2	–	–
Mouse3	M1	T1	M2	LN2	T	T2	–
Mouse4	M1	T	LN1	M2	–	–	–
Mouse5	M1-MAS	T	LN1-MAS ^a	M2-MAS	LN2-MAS ^a	–	–
Mouse6	M1	T1	LN1-MAS ^a	M2-MAS	LN2-MAS ^a	T2	–
Mouse7	M2	T	M2	M1	LN1	–	–
Mouse8	M1	T	LN1	M2 + LN2	M2	–	–
Mouse9	T1	M1	LN1	M2	LN2	T2	T3
Mouse10	M1	T	LN1	M2	LN2	–	–
Mouse11	M1	T	LN1	M2	LN2	–	–

M1 = Mammary gland from the tumor-cells-injected side (2 in data set).

M2 = Mammary gland from the contralateral side (1 in data set).

T = Tumor (T1 = first tumor, T2 = second tumor, etc.) (3 in data set).

LN1 = Lymph nodes from the tumor-cells-injected side (5 in data set).

LN2 = Lymph nodes from the contralateral side (4 in data set).

MAS = Mastitis (6 in data set).

^a Samples not used in data analysis.

side injected with tumor cells. From the contralateral side, 13 samples of mammary tissues and 8 lymph nodes were harvested. From each tissue, 12 or more Raman spectra were collected from random locations, 2 of them were collected within 30 min after the mouse tissues were harvested and 10 were collected later that day during the tissue mapping studies. After the Raman spectra were obtained the samples used to generate the Raman spectra were fixed in a 10% formalin solution and submitted for histopathologic validation to further confirm the pathological states of these tissues. All tissues were examined by a senior pathologist and here we use her categorization.

Raman spectroscopic measurements and analysis

Organs not being uniform contain several types of tissues and cells, e.g. blood, lymph, fibrous tissue, in addition to the parenchymal tissue specific to that organ. Some of these elements generate noise or fluorescence¹⁷ when excited, and therefore can interfere^{18,19} with the Raman signal. The laser beam was focused on the tissue with a spot size of 5 μm radius using a 50 \times objective, and spectra were collected in the backscattering geometry with a 10 s integration time over a spectral range 600–1800 cm^{-1} . In order to maximize the Raman spectral information of a tissue, we performed Raman mapping studies by selecting five different spots in a 100 μm \times 100 μm area on a tissue. Two mapping studies were performed for each tissue. Each spectrum represents a different point of a tissue and each point is measured 3 times before generating the final spectrum, which is a mean of 3 spectral intensities. The spectral resolution was ~ 4 cm^{-1} .

We also found that there were no changes in the spectral features of the tissues when repeated after 24 h. Data from half the mice were used to build the discriminant function model, while data from the other half were used to test the model.

Before the data were subjected to statistical analysis, they were cleaned from various types of noises. The data processing steps included cosmic ray removal, subtraction of noise, removal of tissue fluorescence (when present) and normalization of the spectral intensities. A median filter was applied to the raw data to eliminate any cosmic rays or spikes in the data. Then noise was filtered using wavelets,²⁰ which has been proved useful for de-noising and filtering of Raman data. The background fluorescence was subtracted from the de-noised spectra using a modified cubic-spline algorithm that requires no *a priori* knowledge of the spectra. This algorithm divides the spectra into N (a parameter chosen before the analysis) equal windows and picks the minimum intensity point in each window. The algorithm then checks each of the N points to see if it is a local minimum and discards it if it is not. Then the algorithm adds the endpoints of the spectra to the remaining points. These points will now serve as the control points (or node locations) for the first pass of the curve fit. A piecewise cubic Hermite interpolating polynomial is used to fit the points. The fitted curve is subtracted from the de-noised spectra to produce the corrected spectra. We check to see if the corrected spectrum contains any negative values; if it does, then we choose the minimum point of the corrected spectra and add the corresponding Raman shift value to the

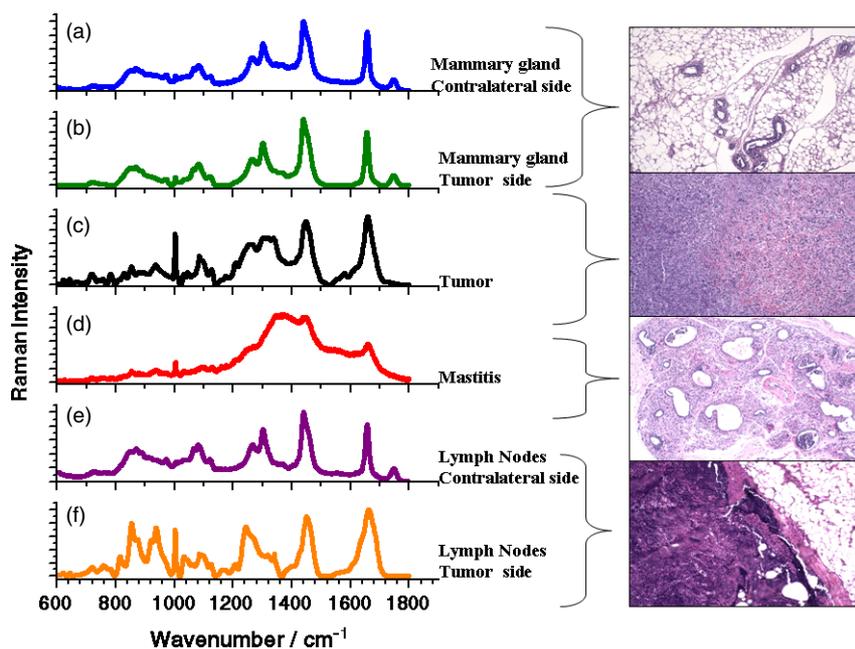


Figure 1. Representative Raman spectra for various tissue types identified in this study using 785 nm as the excitation wavelength. The various Raman intensities are from normal mammary glands contralateral side (a), tumor-cells-injected side (b), tumor (c), lymph nodes contralateral side (e), lymph-node tumor-cells-injected side (f) and mastitis (d).

set of control points. The process is re-iterated until there are only non-negative values in the subtracted spectra. This iterative method is employed to prevent any over-fitting of the de-noised spectra, and finally the corrected spectrum was normalized to unit intensity. For the data analysis we used SSPS statistical packages on desktop computers. Each Raman spectrum contains 1310 data points and we collected more than 650 spectra, but used 650 in our statistical analysis.

By definition, in the covariance matrix, the variables that fluctuate significantly are the ones that contribute significantly to the covariance matrix elements. In our spectra we identified 44 Raman peaks and out of these peaks we selected 29 peaks for statistical analysis on the basis of their standard deviation values. For this selection we determined the standard deviation of all the peaks over the whole data set and took only those peaks whose standard deviation was more than 15% of the maximum observed standard deviation value.

The diagnosis of an altered tissue state, based upon the Raman spectral variation, can involve examining the presence/absence of peaks, peak heights and energy shifts of several peaks. Accordingly, collection of large sample size of data and application of multivariate statistical procedures are required in order to make reliable predictions. To reduce the number of variables, we applied principal components analysis (PCA).^{21,22} This method mutually correlates all the n variables of the data set by calculating an $n \times n$ dimensional correlation matrix. The original variables are then combined into a smaller number of principal components that are orthogonal to one another. The orthogonal variables represent different states of the tissue. At the end of the analysis the independent trends present in the original data set can be easily observed in the reduced variables space through a series of two-dimensional graphs of the significant principal components.

Discriminant function analysis (DFA),²³ which uses a linear combination of independent variables to maximize the separation among the different groups, was used as a classification technique. Half the mice data were used to develop the algorithm and the other half were used to test the predicted classification against the pathologist's tissue assignments. Finally we used multivariate analysis of variance (MANOVA) tests to determine whether there are any significant differences among the tissues.

RESULTS AND DISCUSSION

Raman spectral signatures

The histological examinations of the tissues identified four tissue types: normal mammary glands, mammary tumors, mastitis (fat necrosis and inflammation in two mice), and lymph nodes. Raman spectroscopy not only identified these four tissues but in many instances also distinguished mammary and lymph tissues from the injection side of the

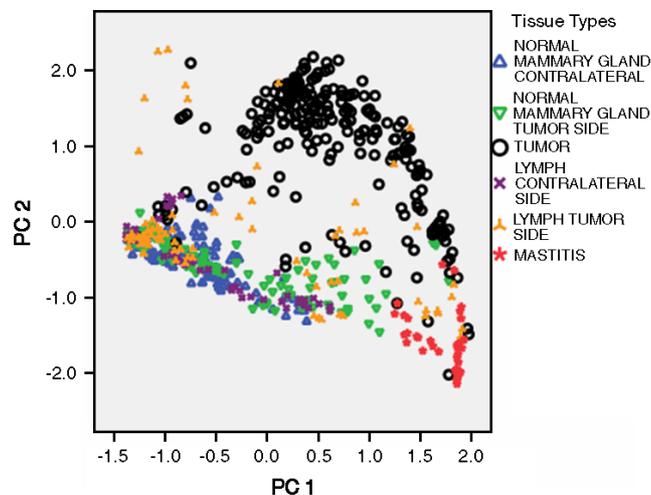


Figure 2. The scatter plots of first two important principal component scores, denoted by PC1 and PC2.

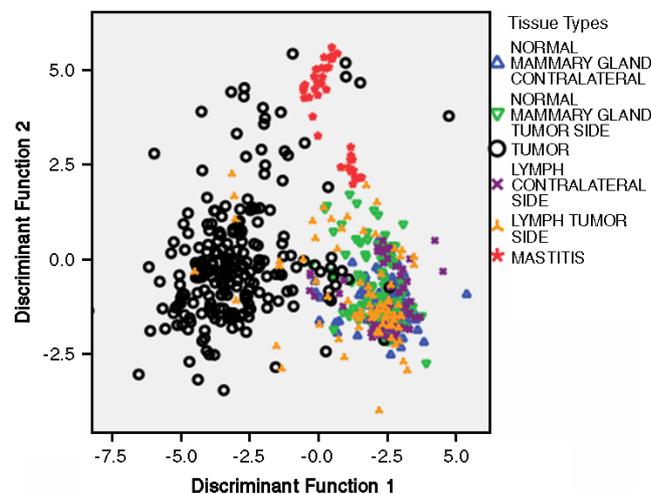


Figure 3. Scatter plots of first two important discriminant functions plotted against one another.

mouse *vs* the contralateral side. Each tissue type showed unique spectral features. As the tissue type changes (e.g. lymph to tumor) most peaks in the Raman spectrum show a variation in their intensities, though some are more dramatic than others, leading to the unique diagnostic features noted above. A typical spectrum for each pathological case as identified by the histological examinations is shown in Fig. 1. On the right-hand side of this figure are the photos of corresponding normal, tumor, mastitis and lymph-node tissues. It is interesting to note that the spectral features of the normal mammary glands are remarkably similar to those of normal human breast spectra.¹ This shows that this mice study may have direct implications to that of human breast diseases. One sees a prominent Raman band around 1748 cm^{-1} in the normal spectrum, but it is completely lost in the tumor spectrum. This band is also absent in the

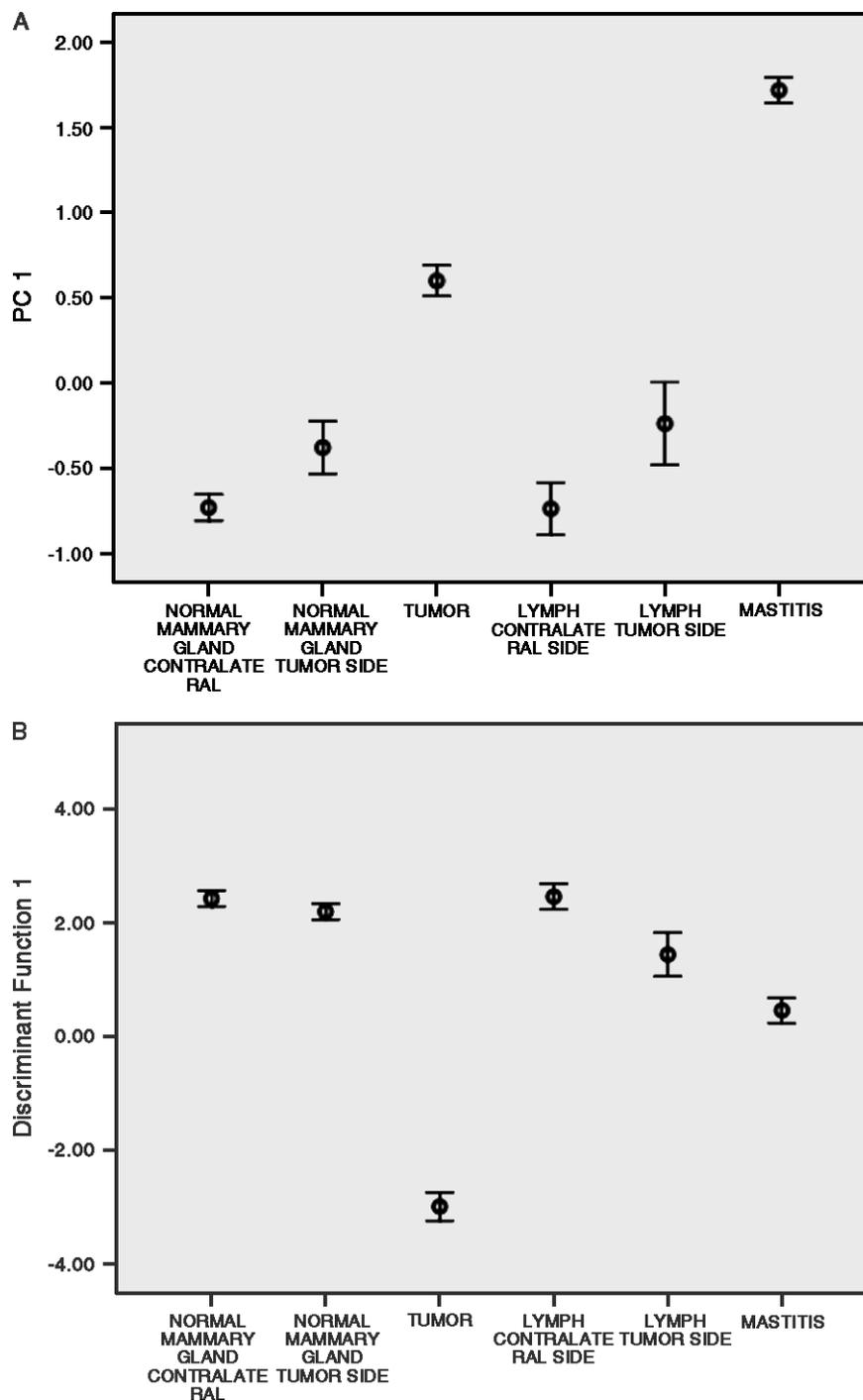


Figure 4. (A) Error bar plot for PC1 against tissue types. The Raman signals for different tissue types are clearly different. (B) Error bar plot for discriminant function 1 against tissue types.

lymph-node spectrum on the tumor-cells-injected side. Most of Raman bands of the mastitis spectrum are broadened compared to the normal spectrum.

Because of the tumor, the normal tissue exhibits several changes and develops additional features as observed in the Raman spectra: (1) The high-frequency mode at $\sim 1747\text{ cm}^{-1}$

due to C=O stretch of lipids nearly loses all its intensity indicating the loss of lipid content in the tissue. (2) The band at $\sim 1655\text{ cm}^{-1}$ becomes broad and asymmetric toward the high-frequency side. This is due to the emergence of the amide I band of proteins. We estimate the position/s of the additional band/s in the range $1660\text{--}1675\text{ cm}^{-1}$, which

are consistent with the protein α -helix, β -sheet and random coil structures. (3) Changes and increases in the intensity profile are seen at $\sim 1265\text{ cm}^{-1}$ (amide III region) along with an additional band at $\sim 1002\text{ cm}^{-1}$. (4) The low-frequency bands due to tyrosine at ~ 850 and 830 cm^{-1} are seen with increased intensity, although in the normal tissues there are some low-intensity, broad peaks near these frequencies. These changes indicate that the tissue identified as tumor by pathology shows an increased level of protein and drastically reduced levels of lipid in contrast to lipid-rich normal tissue.

PCA transformed the initial 29-dimensional data into five dimensions. First three eigenvectors capture 97% of the total variance presence in the spectral data (Fig. 2). Tumors display a very distinct pattern compared to other tissues and states. This shows that at the biomolecular level major changes have occurred in the normal mammary glands because of tumor. Similarly, the spectra of mastitis, and those of normal mammary gland on the contralateral and tumor-cells-injected sides are also different from each other. From this figure it is evident that there are strong fluctuations in the concentrations of the biomolecules of the normal tissue when it is transformed because of the cancer disease, and Raman spectroscopy is sensitive enough to detect these transformations.

Spectral features of mammary tumors and other tissues

In the Fig. 3, discriminant function 1 is plotted against discriminant function 2, resulting in most of the tumor spectra being separated from the rest of the tissue types. The only spectra tending to slightly mix with the tumor is from the lymph nodes on the tumor-cells-injected side as marked by the circles. The second most important result in Fig. 3 is that the discriminant function 2 can clearly separate mastitis from the other tissues. Figure 3 also shows some dispersion of the lymph-node spectra. From this figure one observes that normal mammary spectra from the contralateral side are tightly clustered compared to others. These claims, that the spectra of normal mammary glands, tumor, mastitis and lymph nodes are statistically different, are further strengthened in the error bar plots of PC1 and discriminant function 1 against the tissue types as shown in Fig. 4(A) and (B), respectively. From these figures one can clearly see that the Raman signals for all the six different types of tissues are different.

Spectral features of mastitis

DFA and PCA plots showed that spectral changes due to mastitis are nearly as pronounced as those from tumor whenever mastitis is present in the tissue. In our study we found that 2 out of 11 mice showed mastitis and it was present only on the tumor-injected-side mammary glands. Given a strong correlation of mastitis and tumor, it is worthwhile to investigate its spectral characteristics and see whether

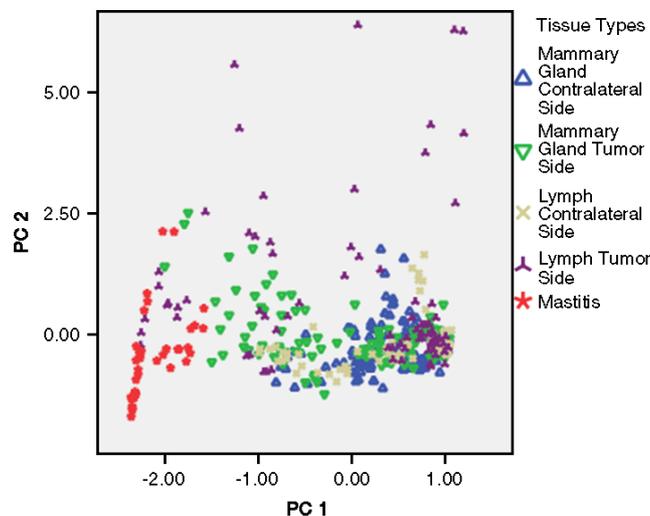


Figure 5. The scatter plots of the first two important principal components of the Raman spectra without including the Raman data from tumor-cells-injected side.

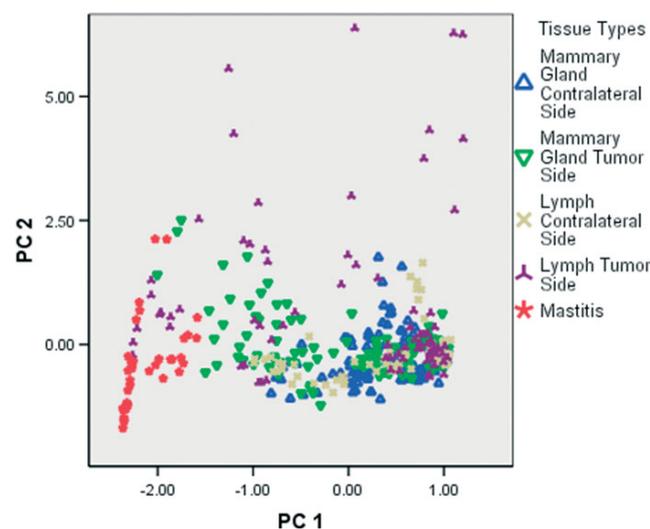


Figure 6. The scatter plots of the first two important discriminant functions against each other for the Raman data without the Raman spectra from tumor-cells-injected side.

they strongly distinguish from normal and lymph nodes. So we analyzed the Raman spectra of mastitis, normal tissue samples and lymph nodes from both injected and contralateral sides. The results of PCA are shown in Fig. 5. In this analysis, the spectra of tumor tissues (confirmed by pathological studies) were excluded from the data set. Clearly mastitis show a very different trend compared to the others and this trend is strongly projected in the DFA plots shown in Fig. 6. In Fig. 7(A) and (B), we show the error bar plot for PC1 and for discriminant function 1, respectively, against the tissue types. One clearly sees that mastitis is very different from the others. We believe that this type of Raman spectral behavior is expected from a tissue in which cancer is

not fully manifested yet but mastitis has already developed. From Fig. 6 one observes that the normal mammary gland spectra from the contralateral side are more clustered and separated from most of the data of mammary glands from the tumor-cells-injected side. The pathological or clinical significance, if any, of the subclustering of mastitis is yet to be resolved.

Surgical significance

Establishing the boundary between a tumor and the normal parenchyma tissue of an organ is critical for successful surgical removal. The separation between the normal mammary spectra from tumor-cells-injected side and contralateral sides is also seen in Fig. 7(B), which shows the discriminant function 2 against the tissue types.

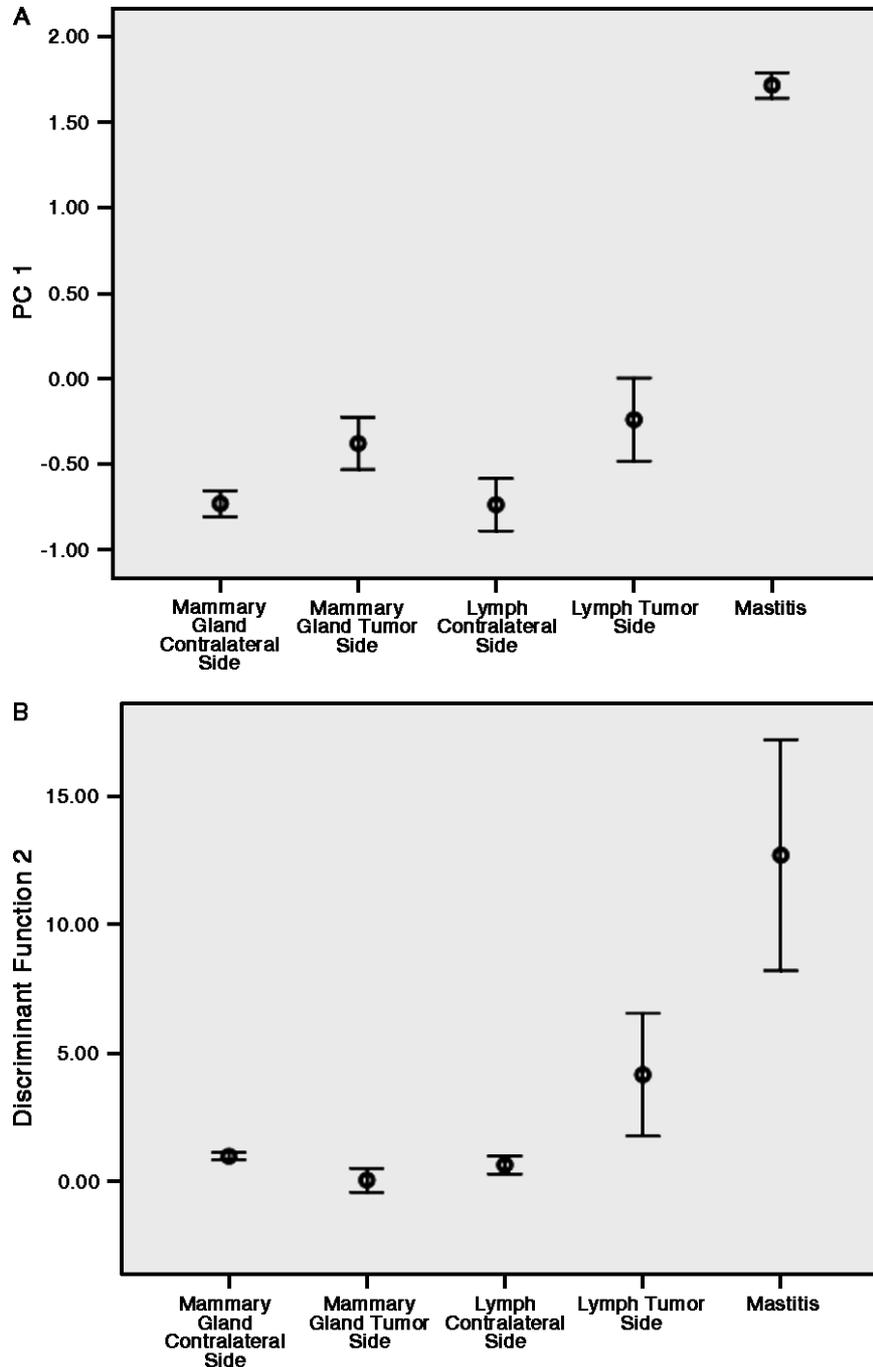


Figure 7. (A) The error bar plot for PC1 against different tissue types. This shows that the PC scores for all the five different tissue types are statistically different. (B) The error bar plot for discriminant function 1 against different tissue types. This shows that the discriminant scores for all the five different tissue types are statistically different.

This distinction could be quite useful in precisely defining surgical incision boundaries. Ultimately, one simply needs to know whether a tissue being examined is cancerous or not. Accordingly, we have collapsed our categories into two: tumor and nontumor. For the data used to construct the algorithm, we correctly classified 99.0% of the nontumor tissue and 94.5% of the tumors were correctly classified. For the data not used to construct the algorithm, we correctly classified 92.2% as nontumorous. Thus, we have, at best, 7.9% false positive results. For the data used to construct the algorithm, we correctly identified tumors 94.5% of the time. And, for the data not used to construct the algorithm, we correctly identified tumors 79.2% of the time resulting in slightly more than 7.9% false positives. When we used an expanded tissue designation to include mastitis, tumors were correctly predicted 74.3% of the time, with 16.8% being predicted to be mastitis and 6.9% predicted to be normal mammary tissue.

CONCLUSIONS

We have demonstrated that Raman spectroscopy measurements of mammary gland tissues from mice injected with 4T1 tumor cells can distinguish mammary tumors from other physiological or pathological states of the mammary glands. The diagnostic algorithms based on the Raman spectrum peak heights categorize tissues into six classes. Out of these, tumor and mastitis are separated from the rest. These separations are of vital clinical importance. However, the spectra of normal mammary glands and lymph nodes from the contralateral side overlap, but lymph-node spectra from the tumor-cells-injected side show slight mixing with tumor spectra. This study suggests that Raman spectroscopy can possibly perform a real-time analysis of the human mammary tissues for the detection of cancer.

Acknowledgements

This work was supported by the Center for Smart Sensors and Integrated Microsystems at Wayne State University, Detroit, MI, and Computer-Assisted Robotic Enhanced Surgery Program at the Children's Hospital of Michigan, Detroit, MI. We are thankful to Dr Fred Miller, Karmanos Cancer Institute, Detroit, MI, for providing 4T1 mouse mammary tumor cell line. The authors also like to thank

Wayne State University's President Reid Research Enhancement Program for funding this work.

REFERENCES

- Haka AS, Shafer-Peltier KE, Fitzmaurice M, Crowe J, Dasari RR, Feld MS. *Proc. Natl. Acad. Sci. U.S.A.* 2005; **102**: 12371.
- Malini R, Venkatakrishna K, Kurien J, Keerthilatha MP, Rao L, Kartha VB, Krishna CM. *Biopolymers* 2006; **81**: 179.
- Frank CJ, McCreery RL, Redd DC. *Anal. Chem.* 1995; **67**: 777.
- Barr H, Dix T, Stone N. *Lasers Med. Sci.* 1998; **13**: 3.
- Mahadevan-Jansen A, Rebecca RK. *J. Biomed. Opt.* 1996; **1**: 31.
- Stone N, Kendall C, Shepherd N, Crow P, Barr H. *J. Raman Spectrosc.* 2002; **33**: 564.
- Wagnieres GA, Star WM, Wilson BC. *Photochem. Photobiol.* 1998; **68**: 603.
- Raman CV, Krishnan KS. *Nature* 1928; **121**: 501.
- Hanlon EB, Manoharan R, Koo TW, Shafer KE, Motz JT, Fitzmaurice M, Kramer JR, Itzkan I, Dasari RR, Feld MS. *Phys. Med. Biol.* 2000; **45**: R1.
- Tissa R, Hata Scholz TA, Ermakov IV, McClane RW, Khachik F, Gellermann W, Pershing LK. *J. Invest. Dermatol.* 2000; **115**: 441.
- Manoharan R, Wang Y, Feld MS. *Spectrochim. Acta, Part A* 1996; **52**: 215.
- Shim MG, Song LM, Marcon NE, Wilson BC. *Photochem. Photobiol.* 2000; **72**: 146.
- Motz JT, Jason TM, Hunter M, Galindo LH, Gardecki JA, Kramer JR, Dasari RR, Feld MS. *Appl. Opt.* 2004; **43**: 542.
- Frank CJ, Redd DC, Gansler TS, McCreery RL. *Anal. Chem.* 1994; **66**: 319.
- Redd DCFZ, Yue KT, Gansler TS. *Appl. Spectrosc.* 1993; **47**: 787.
- Mahadevan-Jansen A, Mitchell MF, Ramanujam N, Malpica A, Thomsen S, Utzinger U, Richards-Kortum R. *Photochem. Photobiol.* 1998; **68**: 123.
- Lieber CA, Mahadevan-Jansen A. *Appl. Spectrosc.* 2003; **57**: 1363.
- Avrillier S, Tinet E, Ettore D, Tualle JM, Gelebart B. *Appl. Opt.* 1998; **37**: 2781.
- Panjehpour M, Julius CE, Phan MN, Vo-Dinh T, Overholt S. *Lasers Surg. Med.* 2002; **31**: 367.
- Wolthuis R, Bakker Schut TC, Caspers PJ, Bushman HPJ, Romer TJ, Bruining HA, Puppels GJ. Raman spectroscopic methods for *in vitro* and *in vivo* tissue characterization. In *Fluorescent and Luminescent Probes for Biological Activity*, Mason WT (ed.). Academic Press: London, 1999; 433.
- Geladi P, Kowalski BR. *Anal. Chem. Acta* 1986; **185**: 1.
- Jolliffe IT. *Principal Component Analysis*. Springer-Verlag: Heidelberg, 1986.
- Lachenbruch PA. *Discriminant Analysis*. Hafner Press: New York, 1975.