

Nanoscale

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Pursuing shell-isolated nanoparticle-enhanced Raman spectroscopy (SHINERS) for concomitant detection of breast lesions and microcalcifications

Chao Zheng^{1,2†}, Wanting Shao^{1†}, Santosh Kumar Paidi², Bing Han¹, Tong Fu¹, Di Wu¹, Lirong Bi³, Weiqing Xu⁴, Zhimin Fan^{1*} and Ishan Barman^{2,5*}

¹Department of Breast Surgery, The First Hospital of Jilin University, Changchun 130021, China

²Department of Mechanical Engineering, Johns Hopkins University, Baltimore, Maryland 21218, USA

³Department of Pathology, The First Hospital of Jilin University, Changchun 130021, China

⁴State Key Laboratory for Supramolecular Structure and Materials, Jilin University, Changchun 130012, China

⁵Department of Oncology, Johns Hopkins University, Baltimore, Maryland 21287, USA

†These authors contributed equally to this work.

*Corresponding authors: ibarman@jhu.edu & fanzhimn@163.com

Abstract

Although tissue staining followed by morphologic identification remains the gold standard for diagnosis of most cancers, such determinations relying solely on morphology are often hampered by inter- and intra-observer variability. Vibrational spectroscopic techniques, in contrast, offer objective markers for diagnoses and can afford disease detection prior to alterations in cellular and extracellular architecture by furnishing a rapid “omics”-like view of the biochemical status of the probed specimen. Here, we report a classification approach to concomitantly detect microcalcification status and local pathological state in breast tissue, featuring a combination of vibrational spectroscopy that focuses on the tumor and its microenvironment, and multivariate data analysis of spectral markers reflecting molecular expression. We employ the unprecedented sensitivity and exquisite molecular specificity offered by Au@SiO₂ shell-isolated nanoparticle-enhanced Raman spectroscopy (SHINERS) to probe the presence of calcified deposits and distinguish between normal breast tissues, fibroadenoma, atypical ductal hyperplasia, ductal carcinoma in situ (DCIS), and invasive ductal carcinoma (IDC). By correlating the spectra with the corresponding histologic assessment, we developed partial least squares-discriminant analysis derived decision algorithm that provides excellent diagnostic power in the fresh frozen sections (overall accuracy of 99.4% and 93.6% using SHINs for breast lesions with and without microcalcifications, respectively). The performance of this decision algorithm is competitive with or supersedes that of analogous algorithms employing spontaneous Raman spectroscopy while enabling facile detection due to the considerably higher intensity of SHINERS. Our results pave the way for rapid tissue spectral pathology measurements using SHINERS that can offer a novel stain-free route to accurate and economical diagnoses without human interpretation.

Introduction

Examination of stained histology slides and the corresponding recognition of morphologic features remains the gold standard for definitive diagnosis and staging of most cancerous lesions. The decision-making relies extensively on a pathologist's recognition of cell types and their morphologic patterns within the stained tissues. Although early detection and useful patient stratification impact the likelihood of cancer-free survival, manual morphological determinations involve considerable diagnostic uncertainty^{1, 2} and do not elucidate the essential biochemical pathways that lead to the specific pathology. Moreover, such examination is time-consuming and expensive. While immunohistochemical approaches offer limited molecular detection, significant hurdles still remain in employing them for quantitative, automated pathology.³

In this milieu, considerable attention has been focused on the incorporation of vibrational spectroscopic methods, as they directly offer non-perturbing molecular descriptors. Raman spectroscopy, for example, provides objective spectral markers for diagnosis of disease and may permit disease detection prior to morphologic manifestation.⁴ Studies by us and others⁵⁻¹⁴ have sought to exploit this method - due to its wealth of molecular information^{15, 16} and lack of sample preparation requirements - in differentiating breast pathology and detecting the presence of microcalcifications, an important mammographic marker of breast lesions. While spontaneous Raman spectroscopy affords sub-cellular signal localization and provides an ideal *in vivo* characterization tool, its inherently weak signals have impeded its application for extensive tissue analysis.

Surface-enhanced Raman scattering (SERS), which was first observed in the 1970s on electrochemically roughened silver electrodes,^{17, 18} has alleviated this drawback with reported biomolecular detection limits beyond immunoassay sensitivities.¹⁹ As a consequence, SERS has attracted considerable interest as an ultrasensitive and highly specific tool for non-destructive and real-time diagnosis of diseases.^{15, 20-28} However, two problems still hinder the translation of SERS for biomedical applications, namely unsatisfactory substrate generality and poor measurement reproducibility.^{29, 30} To overcome these drawbacks, shell-isolated nanoparticle-enhanced Raman spectroscopy (SHINERS), featuring nanoparticles with Au core coated by an ultrathin shell (thickness from 2-20 nm) of silica or alumina, was proposed by Tian and co-

workers at Xiamen University.³¹ While the gold core of the Au@SiO₂ shell-isolated nanoparticles (SHINs) allows obvious SERS enhancement, the shell is used to protect the gold core from degradation of the minutely fabricated nanostructures as well as preclude the interaction of the bare gold nanoparticles with the probed adsorbates. By virtue of these advantages, SHINERS has been gainfully employed in investigations of live cells³² and, recently by us, in carcinoma studies using a ratiometric approach.³³

Using breast cancer as the paradigm, we report here an approach to concomitantly diagnose microcalcification status and local pathological state featuring SHINERS and multivariate data analysis of spectral markers reflecting molecular expression. In this article, we seek to exploit the heightened sensitivity and intrinsic specificity offered by the Au@SiO₂ shell-isolated nanoparticles in elucidating the biochemical content of fresh frozen specimen and, thus, in differentiating between normal breast tissue, fibroadenoma, atypical ductal hyperplasia, ductal carcinoma in situ, and invasive ductal carcinoma. By correlating the spectra with the histologic evaluation, our partial least squares - discriminant analysis derived decision algorithms show excellent diagnostic power (overall accuracy of 99.4% and 93.6% using SHINs for breast lesions with and without microcalcifications, respectively). The performance of this decision algorithm is competitive with or supersedes that of analogous algorithms developed using spontaneous Raman spectra while offering facile detection owing to the considerably higher intensity of SHINERS spectra. Notably, to the best of our knowledge, this is the first SHINERS effort that comprehensively assays microcalcifications and tissues of all key pathophysiological conditions. Taken together with the substantially higher spectral intensities for the SHINERS data in relation to that for the spontaneous Raman spectra, our findings open the door for clinical translation of SHINERS as a label-free route to accurate pathological diagnoses without human interpretation.

Materials and Methods

Human subjects and tissue preparation

Fresh breast tissue was collected from 72 patients who underwent surgical resection or vacuum assisted (Mammotome) biopsy at the Department of Breast Surgery, the First Hospital of Jilin University. Consent to participate in this study was obtained from every patient and the project and methodology were approved by the Ethics Committee of Jilin University. After operating, the samples were immediately frozen at -20 to -25°C and two contiguous sections ($6\ \mu\text{m}$ thickness) were cut from each specimen using a freezing microtome (LEICA-CM3050S, Germany) in the Pathology Department, the First Hospital of Jilin University. For each tissue, one of the two adjacent sections was stained with haematoxylin and eosin (H&E) for routine histopathological analysis by three experienced breast pathologists. The other section was used for acquiring spectra from the diseased regions (identified by H&E staining of the other section) by preserving it in liquid nitrogen. Prior to analysis, the frozen section was thawed at 22°C for 10 minutes and multiple spectra were collected from each tissue.

Instrumentation and data acquisition

A confocal Raman system (LabRAM ARAMIS, Horiba Jobin Yvon, Edison, NJ, USA) with a $\sim 0.7\ \mu\text{m}$ spatial resolution, and a 5 mW, 633nm HeNe laser as excitation source was used for the collection of Raman spectra. The detection of Raman signal was carried out with a Synapse Thermoelectric cooled charge-coupled device (CCD) camera (Horiba Jobin Yvon, Edison, NJ, USA). Raman scattered light was collected with a 50x microscope objective lens (0.50 NA, LMPLFLN, Olympus, Japan) that was also used for focusing the excitation laser light. The laser beam focused on the tissue formed a spot of $1.5\ \mu\text{m}$ diameter. A 4-notch filter (Horiba Jobin Yvon, Edison, NJ, USA) was used to block strong Rayleigh-scattered light. Extended scan spectra with a spectral range of $600\text{--}1800\ \text{cm}^{-1}$ were acquired using an integration time of 60s and 3 accumulations. The Raman shift axis was calibrated using characteristic vibration bands recorded from silicon wafer. These settings were kept constant for all the spectral measurements to enable direct comparison of spectral sensitivity and specificity, especially between spontaneous Raman and SHINERS datasets.

Locations of lesions and other possible features of interest were determined from the H&E sections and the corresponding sites on its frozen contiguous sections were probed using the Raman system. About fifteen to twenty spectra were collected from different locations for each sample to ensure representative sampling and incorporate spot-to-spot variability in the recorded signals. After obtaining spontaneous Raman spectra, SHINs were added to the surface of frozen sections and the respective SERS spectra were recorded from the same spots. SHINs having an average Au core diameter of 55nm and covered by 2nm silica shells were acquired from Professor Tian's group for use in this study. On each tissue section 10 μ L of 100 nM concentrated SHINs solution was added uniformly prior to collecting the spectra. Fig. 1 shows a representative TEM image of the SHINs used for our study.

Spectral data analysis

The spectra were subjected to baseline correction by fitting and subtracting a third-order polynomial by NGS LabSpec software (Horiba Jobin Yvon, Edison, NJ, USA) and were subsequently smoothed using a 15-point averaging algorithm. These corrected spectra were used to determine the peaks characteristic of the class the spectra belong to while the intensity-normalized spectra were used in the remaining data analysis. The latter were subjected to principal component analysis (PCA), a widely used data exploration method, to capture the (subtle) spectral variances in a set of abstract orthogonal axes. For every spectrum belonging to a different pathological class, PC scores were plotted using Radviz and VizRank, radial visualization modules in Orange data mining software.³⁴ The nonlinear multidimensional radial visualization algorithm, RadViz, maps n data dimensions (PCs) onto two dimensional circular space while VizRank, offers a heuristic search technique to guide the ordering of variables and evaluating the resulting radial projections by their ability to discriminate between classes. These radial visualization plots illustrate the clustering of spectra belonging to same class (assigned based on pathological assessment) and how well the classes are separated based on the PC scores. In this study, we used only the single best-ranked projection for clustering displays. Finally, the spectra belonging to various pathologies were subjected to partial least squares-discriminant analysis (PLS-DA)³⁵ to quantify the diagnostic power of the recorded spectra. A PLS-DA derived decision model was built based on training data and the diagnostic power was independently tested by invoking an independent set of test spectra. In particular, 60% of the

spectra belonging to each class were randomly selected to train the model and the remaining 40% were tested using the model developed to check the classification capability of PLS-DA in identifying the pathology of the tissue. 1000 iterations were performed keeping the same mixture of training: test data to obtain a more robust evaluation of the classification performance.

Results

The Raman spectral dataset obtained was classified as belonging to one of the following classes based on the consensus histological assessment. 17, 9, 11 and 20 tissue sections were diagnosed as fibroadenoma (FD), atypical ductal hyperplasia (ADH), ductal carcinoma in situ (DCIS) and invasive ductal carcinoma (IDC), respectively. Additionally, 15 normal breast tissue (NB) sections were investigated for comparison with benign, premalignant and malignant pathologies. Among the diseased tissue sections, some exhibited the presence of type II microcalcifications (3 FD, 3 ADH, 5 DCIS and 4 IDC tissue sections) and were studied separately as they form an important class of pathological indicators in breast cancer diagnosis.

Spontaneous Raman spectroscopy and SHINERS of tissues without microcalcifications

Fig. 2 shows the mean spontaneous Raman and SHINERS spectra of breast tissues without calcifications belonging to the different classes, *i.e.* NB, FD, ADH, DCIS and IDC (spectra are displayed vertically offset for visual clarity). It can clearly be seen that the Raman features were significantly enhanced by using SHINs enabling us to readily characterize the differences in spectral profiles of NB, FD, ADH, DCIS, and IDC tissue sections. The characteristic band assignments of the breast tissue, performed in our recent publication,³³ indicate the presence of a similar set of chemical compounds, namely lipids, proteins, nucleic acids and β -carotenes. The main spectral features of normal breast tissue include peaks around 1300, 1442, and 1745 cm^{-1} that can be expectedly attributed to lipids^{36,37} (with the corresponding lack of prominent protein or nucleic acid peaks). Fibroadenoma shows similar characteristics as normal breast tissue except for the Raman band at 665 cm^{-1} that is characteristic of the C-S stretching mode of cystine.^{36,37} In comparison to normal breast tissue, ADH, DCIS and IDC tissues show stronger peaks located at 1004, 1033, 1610 and 1658 cm^{-1} from vibrational modes of proteins and at 970, 1090 and 1157 cm^{-1} from DNA. These are indicative of higher concentration of proteins and nucleic acids in the diseased state. Closer inspection reveals that the concentrations of proteins and nucleic acids increase gradually while lipid content reduces in the pathological transition from ADH to

IDC. Additionally, DCIS and IDC tissues feature the blue shifted and broad peak of CH₂, which suggests fracture of acyl backbone of lipids and proteins.³⁸ Also, the sensitive Raman features at 1090 cm⁻¹, characteristic of the phosphate stretching vibration in DNA, shows a perceptible shift to 1086 cm⁻¹ for DCIS and IDC hinting at the possible rupture of DNA strands with the further progression of cancer. A comprehensive list of the spectral features, and the associated peak shifts, corresponding to the different pathological classes is provided in our previous report.³³

Our next aim was to correlate the spectroscopic measurements in both cases with the histological evaluation. Fig. 3 shows a representative set of spontaneous Raman and SHINERS spectra recorded from ADH and IDC tissue sections, respectively, with the latter showing the presence of microcalcifications. Panel (D) of the figure exhibits the corresponding H&E stained section as well as the frozen contiguous sections without the addition of SHINs (Panel (B), from which the spontaneous Raman data was recorded) and post the addition of SHINs (Panel (C), from which the SHINERS data was recorded). Based on the PCA decomposition, the most significant PC scores were then employed to create a radial visualization map to evaluate the clustering of spectra belonging to the same pathological class-or the lack thereof. Fig. 4 displays the cluster maps for the four types of breast lesions (FD, ADH, DCIS, IDC) using five PC scores for spontaneous Raman (A) and SHINERS spectra (B), respectively. (The corresponding PC loadings are given in Fig. S1 of Supplementary Information.) From the figure, we observe that the sites belonging to a particular class show a tendency to cluster together, although there is considerable overlap between few of the classes particularly for (A) (spontaneous Raman spectra-based PC scores map). The Raman spectra-derived overlap is consistent with our previous report,¹² where significant segmentation between the pathologies was not observed visually-despite the absence of ADH from that dataset. This overlap is significantly reduced in Fig. 4(B), which yields reasonably satisfactory (though not perfect) class boundaries. We attribute this qualitative enhancement in spectra-based class separation to the enhancement of the subtle vibrational mode-specific spectral features by SHINs.

Quantitative assessment of the diagnostic capability can be obtained from the results of PLS-DA-derived decision algorithm summarized in Table 1. Using independent training and test sets, we computed the overall accuracy of the PLS-DA derived decision algorithms in classifying NB, FD, ADH, DCIS and IDC to be 72.8%, 91.0%, 97.2%, 87.4% and 96.9%, respectively using

spontaneous Raman spectroscopy. The analogous PLS-DA derived decision models for the SHINERS data yielded mean classification accuracy of 82.0%, 92.3%, 97.5%, 92.7% and 91.8%, respectively. Despite the extremely challenging cohort studied here (as reflected by the presence of ADH, DCIS and IDC sites), Raman spectroscopy (both spontaneous and enhanced) offer significant spectral markers of the (downstream) molecular determinants in the epithelial and stromal tissue components. Specifically, when used in conjunction with chemometric methods, these spectral profiles enable objective disease detection with high diagnostic accuracy. The overall classification performance rose for all categories from spontaneous Raman to SHINERS, barring IDC, with significant improvements for normal breast tissue and DCIS segmentation. The inferior performance of DCIS classification using spontaneous Raman spectra is reflective of the large spread observed in Fig. 4(A), while the reduced classification accuracy of IDC for SHINERS can also be predicted based on the dispersion of the IDC sites in Fig. 4(B).

Spontaneous Raman spectroscopy and SHINERS of tissues with type II microcalcifications

Raman spectroscopy has been previously investigated in formalin-fixed, paraffin-embedded sections for detection of type I and II breast microcalcifications as well as for the classification of type II microcalcifications related to benign and malignant lesions.³⁹ More recently, we have also developed Raman spectroscopy algorithms to detect microcalcifications in fresh breast needle biopsy tissue cores.¹¹ Here, Fig. 5 shows the mean spectra of tissues showing presence of microcalcifications with and without SHINs measured from fresh frozen sections of FD, ADH, DCIS, and IDC tissue. The major peak at *ca.* 960 cm^{-1} is a well-known calcium hydroxyapatite peak and can be ascribed to the totally symmetric stretching mode of the “free” tetrahedral phosphate ion.³⁹⁻⁴¹ The other noticeable peaks appear at 1002, 1072, 1156, 1446, and 1663 cm^{-1} that are attributed to amino acid residue (phenylalanine),⁴² nucleic acids, lipids, carotenoids, and amide-I, respectively.^{36, 37} FD and ADH microcalcifications spectra exhibit a peak at 1072 cm^{-1} , which corresponds to the O-P-O stretch of nucleic acids, whereas in DCIS and IDC tissues the feature is observed to be red-shifted to 1078 cm^{-1} . The spectra belonging to ADH show a stronger characteristic amide-I peak at 1657 cm^{-1} while DCIS tissue sites harbouring microcalcifications show the presence of relatively stronger peaks corresponding to amino acid residue (phenylalanine) at 1003, 1031 cm^{-1} and lipids (CH_2 and CH_3 bending) at 1301, 1441 cm^{-1} .

Also, compared to the DCIS sites, the spectra recorded from IDC sites with microcalcifications show stronger nucleic acids peaks and weaker lipids peaks.

Fig. 6 shows the corresponding radial visualization plot employing selected PC scores to depict maximum class separation of the four diseased classes. (The corresponding PC loadings are given in Fig. S1 of Supplementary Information.) Once again, it is evident that the addition of SHINs improves the separation between sites belonging to the same pathological class. A comparison of Fig. 4 and 6 reveals that the separation for both spontaneous Raman and SHINERS data is much improved for the tissue sites harboring microcalcifications. To quantitatively evaluate the diagnostic advantage provided by microcalcifications in classifying the diseased tissues, PLS-DA was employed to develop decision models, as detailed earlier. The results obtained are summarized in Table 2. In this analysis restricted to sites harboring microcalcifications alone, classification accuracies of 99.0%, 98.1%, 96.9% and 100% were achieved for the classes FD, ADH, DCIS and IDC respectively, using spontaneous Raman spectroscopy. The PLS-DA derived decision models for SHINERS, on the other hand, yielded accuracies of 99.9%, 99.9%, 97.5% and 98.2% for the same pathologies. Clearly, the multivariate analysis corroborates the visual findings that the accuracy of classification is higher for tissues harboring microcalcifications in comparison with those that show no microcalcifications. Detection of such biomineralization processes is of value not only in examining breast pathology but also in recognizing mineralized deposits in other organs, including the thyroid⁴³ and middle ear.⁴⁴

Discussion

The results of the present study demonstrate the ability of label-free Raman spectroscopy to differentiate closely related breast pathologies. Furthermore, the improvement of the classification results and the class separation in radial plots with the addition of SHINs underscores their diagnostic utility in tissue spectropathology. These two label-free approaches collectively offer a powerful tool to measure both localized cellular changes (especially in the case of SHINERS, since SERS presents a very surface selective method due to the decay of field enhancement with r^{-3})⁴⁵ and the tissue microenvironment. While a critical role for the stromal environment is acknowledged in the onset and progression of cancer,⁴⁶ existing diagnostic methods largely focus on the characterization of epithelial cells—a barrier that can be overcome by the employment of the employed photonic tools.

Significantly, the developed multivariate decision models were able to exploit the molecular structural differences - as indicated by spectral peak shifts, band shapes and relative intensity changes occurring across the Raman spectrum - to identify ADH, a benign lesion of the breast that indicates an increased risk of breast cancer. This differentiation has represented an outstanding challenge for most non-invasive photonic tools, due to its similarity to low-grade DCIS - cytologically, architecturally and on a molecular basis. In fact, early and accurate detection of ADH and DCIS prior to their progression into invasive disease represents a key goal in breast cancer diagnostics⁴⁷. We observed the spectral markers representative of vibrational modes of proteins and nucleic acids increase gradually while those of lipids reduce in the pathological progression from ADH to IDC. In addition, the CH₂ and the C=O stretching mode of phospholipids showed a consistent 2-3 cm⁻¹ blue shift in DCIS and IDC tissues, due to lipid peroxidation that occurs in the process of canceration.³⁸ Based on these findings, this study provides a significant extension to the literature by accurately differentiating these closely related pathologies (especially ADH and DCIS) in fresh frozen tissues (rather than in deparaffinised sections that may introduce spectral artifacts.⁴⁸ Perhaps surprisingly, the accuracy of classification for normal breast tissue using both spontaneous Raman spectroscopy and SHINERS was found to be lower than diseased tissues (*ca.* 73% and 82% respectively) in the recorded dataset. We attribute this to the probable presence of some uninvolved tissue sites in the diseased pathology categories that would then skew the classification model trained on spectra originating from such sites.

We have also discovered the presence of subtle, but reproducible, differences in the chemical composition of tissue sites harbouring type II microcalcifications among the four kinds of studied lesions. It was observed that the peak of O-P-O stretch of nucleic acids at 1072cm⁻¹ in the spectra of FD and ADH tissues shifted to 1078 cm⁻¹ for DCIS and IDC tissues, as a result of the changed calcification compositions of different kinds of diseases. Moreover, the ADH sites exhibit a significant amide-I peak at 1657 cm⁻¹; in contrast, the same peak loses its intensity in DCIS and IDC tissues owing to the process of canceration whereby the protein conformation is damaged to varying degrees.

A further important goal of this work was to compare the diagnostic power of SHINERS based decision models to that derived from spontaneous Raman data. We find that, on average, the

performance of the SHINERS decision models supersedes that of its spontaneous Raman counterpart. This is also reflected in the clearer clustering of the SHINERS based PC scores corresponding to the different pathologies in the radial visualization plots. A slight decrease in the classification accuracy of IDC sites (with and without microcalcifications) is observed for SHINERS in relation to spontaneous Raman measurements. While the precise rationale for this is the subject of ongoing investigations in a larger patient cohort, a plausible explanation lies in the (sharper) differences in the nature of molecular changes that appear in IDC among different patients due to its higher degree of malignancy. Following the enhancement of spectral features by SHINERS, these inter-lesion differences are amplified resulting in worse performance of the multivariate classification models.

Finally, the intensity ratios of the key lipid, protein and nucleic acid peaks underscore the improvement in signal sensitivity due to the incorporation of SHINs and indicate that a significant reduction in SHINERS acquisition time could still result in nearly equivalent spectral profiles as obtained in the spontaneous Raman measurements. This has substantive implications for Raman spectroscopic imaging, which would overcome the sampling limitations of single point recordings, and combine the molecular and spatial information content into a hyperspectral dataset. Unlike targeted SERS probes, where the SERS reporters transduce the presence of the biochemical analytes of interest into measureable signals, SHINERS allows the enhancement of the intrinsic vibrational modes of all the molecular species present in the probed specimen. This ensures the retention of the exquisite molecular specificity of spontaneous Raman spectroscopy thereby offering a unique combination of signal enhancement with facile readout and extensive multiplexing.

Acknowledgments

This work was supported by the National Natural Science Foundation of China NSFC Grant Nos. 81202078, and 21373096, and National Instrumentation Program (NIP) of the Ministry of Science and Technology of China No. 2011YQ03012408. This study was supported by the National Construction of High Quality University Projects of Graduates from the China Scholarship Council (CSC) (Grant No. 201406170141) and the JHU Whiting School of

Engineering Startup Funding to IB. The authors are grateful for the shell-isolated nanoparticles (SHINs) for SERS detection provided by Prof. Zhongqun Tian of Xiamen University.

Conflict of interest

The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval (Medical Ethics Committee, The First Hospital of Jilin University, No. 2013-168) and have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

References

1. W. C. Allsbrook, K. A. Mangold, M. H. Johnson, R. B. Lane, C. G. Lane, M. B. Amin, D. G. Bostwick, P. A. Humphrey, E. C. Jones, V. E. Reuter, W. Sakr, I. A. Sesterhenn, P. Troncoso, T. M. Wheeler and J. I. Epstein, *Human Pathology*, 2001, 32, 74-80.
2. M. R. Emmert-Buck, R. L. Strausberg, D. B. Krizman, M. F. Bonaldo, R. F. Bonner, D. G. Bostwick, M. R. Brown, K. H. Buetow, R. F. Chuaqui, D. A. Cole, P. H. Duray, C. R. Englert, J. W. Gillespie, S. Greenhut, L. Grouse, L. W. Hillier, K. S. Katz, R. D. Klausner, V. Kuznetsov, A. E. Lash, G. Lennon, W. M. Linehan, L. A. Liotta, M. A. Marra, P. J. Munson, D. K. Ornstein, V. V. Prabhu, C. Prange, G. D. Schuler, M. B. Soares, C. M. Tolstoshev, C. D. Vocke and R. H. Waterston, *Am. J. Pathol.*, 2000, 156, 1109-1115.
3. D. C. Fernandez, R. Bhargava, S. M. Hewitt and I. W. Levin, *Nat. Biotechnol.*, 2005, 23, 469-474.
4. C. A. Lieber, H. E. Nethercott and M. H. Kabeer, *Biomed. Opt. Express*, 2010, 1, 975-982.
5. A. S. Haka, K. E. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R. R. Dasari and M. S. Feld, *Proceedings of the National Academy of Sciences of the United States of America*, 2005, 102, 12371-12376.
6. P. Matousek and N. Stone, *Journal of biophotonics*, 2013, 6, 7-19.
7. D. Evers, B. Hendriks, G. Lucassen and T. Ruers, *Future oncology (London, England)*, 2012, 8, 307-320.
8. C. M. Krishna, J. Kurien, S. Mathew, L. Rao, K. Maheedhar, K. K. Kumar and M. V. P. Chowdary, *Expert Rev. Mol. Diagn.*, 2008, 8, 149-166.
9. C. Hu, J. Wang, C. Zheng, S. Xu, H. Zhang, Y. Liang, L. Bi, Z. Fan, B. Han and W. Xu, *Medical physics*, 2013, 40, 063501.
10. I. Barman, N. C. Dingari, A. Saha, S. McGee, L. H. Galindo, W. Liu, D. Plecha, N. Klein, R. R. Dasari and M. Fitzmaurice, *Cancer research*, 2013, 73, 3206-3215.
11. A. Saha, I. Barman, N. C. Dingari, S. McGee, Z. Volynskaya, L. H. Galindo, W. Liu, D. Plecha, N. Klein, R. R. Dasari and M. Fitzmaurice, *Biomed. Opt. Express*, 2011, 2, 2792-2803.
12. N. C. Dingari, I. Barman, A. Saha, S. McGee, L. H. Galindo, W. Liu, D. Plecha, N. Klein, R. R. Dasari and M. Fitzmaurice, *Journal of biophotonics*, 2013, 6, 371-381.

13. A. Saha, I. Barman, N. C. Dingari, L. H. Galindo, A. Sattar, W. Liu, D. Plecha, N. Klein, R. R. Dasari and M. Fitzmaurice, *Analytical chemistry*, 2012, 84, 6715-6722.
14. R. Sathyavathi, A. Saha, J. S. Soares, N. Spegazzini, S. McGee, R. Rao Dasari, M. Fitzmaurice and I. Barman, *Scientific Reports*, 2015, 5, 9907.
15. H. Abramczyk and B. Brozek-Pluska, *Chemical Reviews*, 2013, 113, 5766-5781.
16. S. Devpura, K. N. Barton, S. L. Brown, O. Palyvoda, S. Kalkanis, V. M. Naik, F. Siddiqui, R. Naik and I. J. Chetty, *Medical Physics*, 2014, 41, 050901.
17. M. Fleischmann, P. J. Hendra and A. J. McQuillan, *Chemical Physics Letters*, 1974, 26, 163-166.
18. D. L. Jeanmaire and R. P. Van Duyne, *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry*, 1977, 84, 1-20.
19. D. S. Grubisha, R. J. Lipert, H. Y. Park, J. Driskell and M. D. Porter, *Analytical chemistry*, 2003, 75, 5936-5943.
20. M. Y. Sha, H. Xu, S. G. Penn and R. Cromer, *Nanomedicine : nanotechnology, biology, and medicine*, 2007, 2, 725-734.
21. Q. Tu and C. Chang, *Nanomedicine : nanotechnology, biology, and medicine*, 2012, 8, 545-558.
22. L. Dykman and N. Khlebtsov, *Chemical Society reviews*, 2012, 41, 2256-2282.
23. H. Chon, S. Lee, S. Y. Yoon, S. I. Chang, D. W. Lim and J. Choo, *Chemical communications (Cambridge, England)*, 2011, 47, 12515-12517.
24. J. Lin, R. Chen, S. Feng, J. Pan, B. Li, G. Chen, S. Lin, C. Li, L.-q. Sun, Z. Huang and H. Zeng, *Journal of Raman Spectroscopy*, 2012, 43, 497-502.
25. J. Yang, Z. Wang, S. Zong, C. Song, R. Zhang and Y. Cui, *Analytical and bioanalytical chemistry*, 2012, 402, 1093-1100.
26. P. Wu, Y. Gao, H. Zhang and C. Cai, *Analytical chemistry*, 2012, 84, 7692-7699.
27. Y. Yang, F. Li, L. Gao, Z. Wang, M. J. Thrall, S. S. Shen, K. K. Wong and S. T. Wong, *Biomed Opt Express*, 2011, 2, 2160-2174.
28. M. Li, J. W. Kang, S. Sukumar, R. R. Dasari and I. Barman, *Chemical Science*, 2015, 6, 3906-3914.
29. J. F. Li, X. D. Tian, S. B. Li, J. R. Anema, Z. L. Yang, Y. Ding, Y. F. Wu, Y. M. Zeng, Q. Z. Chen, B. Ren, Z. L. Wang and Z. Q. Tian, *Nature protocols*, 2013, 8, 52-65.

30. S. Park, P. X. Yang, P. Corredor and M. J. Weaver, *Journal of the American Chemical Society*, 2002, 124, 2428-2429.
31. J. F. Li, Y. F. Huang, Y. Ding, Z. L. Yang, S. B. Li, X. S. Zhou, F. R. Fan, W. Zhang, Z. Y. Zhou, Y. Wu de, B. Ren, Z. L. Wang and Z. Q. Tian, *Nature*, 2010, 464, 392-395.
32. J. F. Li, X. D. Tian, S. B. Li, J. R. Anema, Z. L. Yang, Y. Ding, Y. F. Wu, Y. M. Zeng, Q. Z. Chen, B. Ren, Z. L. Wang and Z. Q. Tian, *Nature protocols*, 2013, 8, 52-65.
33. C. Zheng, L. Liang, S. Xu, H. Zhang, C. Hu, L. Bi, Z. Fan, B. Han and W. Xu, *Analytical and bioanalytical chemistry*, 2014, 406, 5425-5432.
34. J. Demsar, T. Curk, A. Erjavec, C. Gorup, T. Hocevar, M. Milutinovic, M. Mozina, M. Polajnar, M. Toplak, A. Staric, M. Stajdohar, L. Umek, L. Zagar, J. Zbontar, M. Zitnik and B. Zupan, *J. Mach. Learn. Res.*, 2013, 14, 2349-2353.
35. R. G. Brereton and G. R. Lloyd, *J. Chemometr.*, 2014, 28, 213-225.
36. R. E. Kast, G. K. Serhatkulu, A. Cao, A. K. Pandya, H. Dai, J. S. Thakur, V. M. Naik, R. Naik, M. D. Klein, G. W. Auner and R. Rabah, *Biopolymers*, 2008, 89, 235-241.
37. N. Stone, C. Kendall, J. Smith, P. Crow and H. Barr, *Faraday discussions*, 2004, 126, 141-157; discussion 169-183.
38. H. Abramczyk, I. Placek, B. Brozek-Pluska, K. Kurczewski, Z. Morawiecc and M. Tazbir, *Spectr.-Int. J.*, 2008, 22, 113-121.
39. A. S. Haka, K. E. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R. R. Dasari and M. S. Feld, *Cancer research*, 2002, 62, 5375-5380.
40. L. Liang, C. Zheng, H. Zhang, S. Xu, Z. Zhang, C. Hu, L. Bi, Z. Fan, B. Han and W. Xu, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2014, 132, 397-402.
41. R. Baker, P. Matousek, K. L. Ronayne, A. W. Parker, K. Rogers and N. Stone, *The Analyst*, 2007, 132, 48-53.
42. M. Marcelo, R. Leandro, A. Emília Ângelo Loschiavo, S. Ana Maria do Espírito, S. Edson Aparecido Pereira dos, B. Renata Andrade and M. Airton Abrahão, *Theoretical Chemistry Accounts*, 2009, 125, 329-334.
43. M. L. C. Khoo, S. L. Asa, I. J. Witterick and J. L. Freeman, *Head & Neck*, 2002, 24, 651-655.

44. R. Pandey, S. K. Paidi, J. W. Kang, N. Spegazzini, R. R. Dasari, T. A. Valdez and I. Barman, *Scientific Reports*, 2015, 5, 13305.
45. P. L. Stiles, J. A. Dieringer, N. C. Shah and R. P. Van Duyne, *Annual Review of Analytical Chemistry*, 2008, 1, 601-626.
46. A. H. Beck, A. R. Sangoi, S. Leung, R. J. Marinelli, T. O. Nielsen, M. J. van de Vijver, R. B. West, M. van de Rijn and D. Koller, *Science Translational Medicine*, 2011, 3, 108ra113-108ra113.
47. A. Bombonati and D. C. Sgroi, *The Journal of pathology*, 2011, 223, 307-317.
48. R. Wiens, M. Rak, N. Cox, S. Abraham, B. H. J. Juurlink, W. M. Kulyk and K. M. Gough, *Analytical and bioanalytical chemistry*, 2007, 387, 1679-1689.

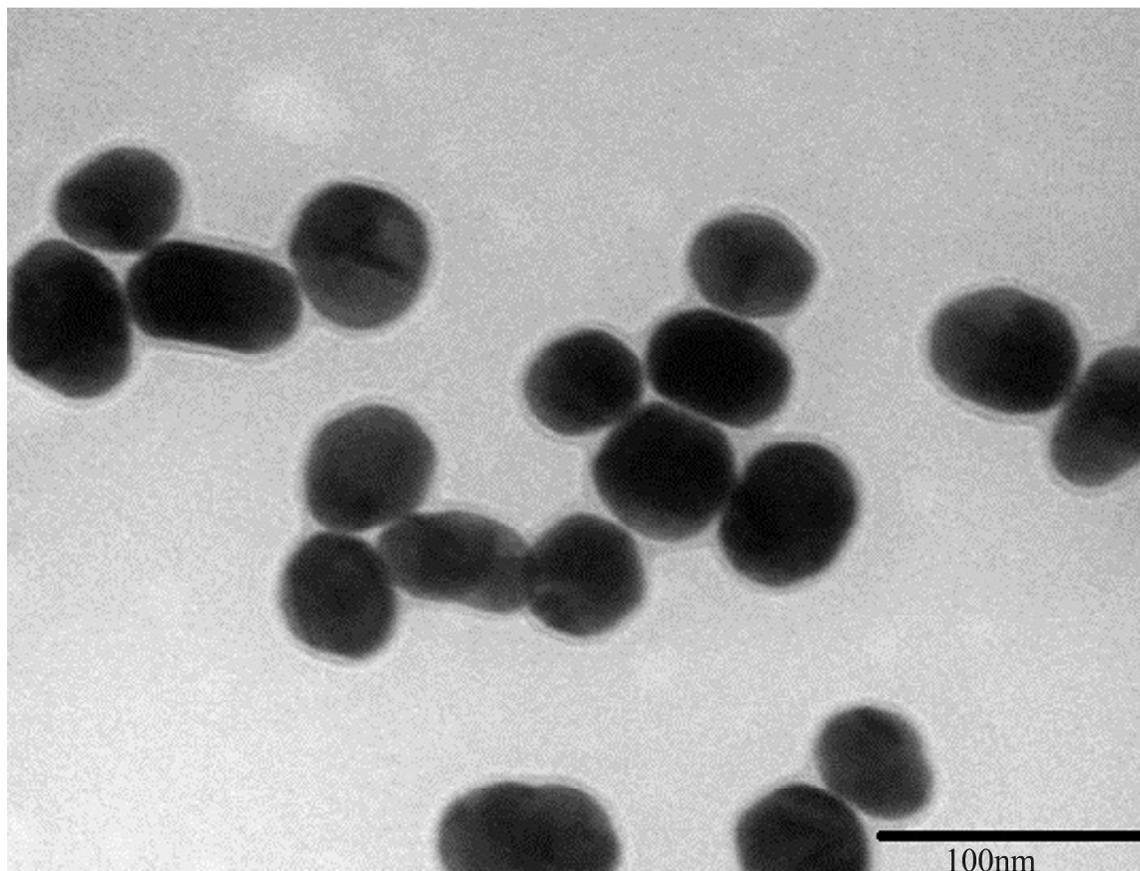


Fig. 1 Representative TEM image of Au@SiO₂ shell-isolated nanoparticles (SHINs).

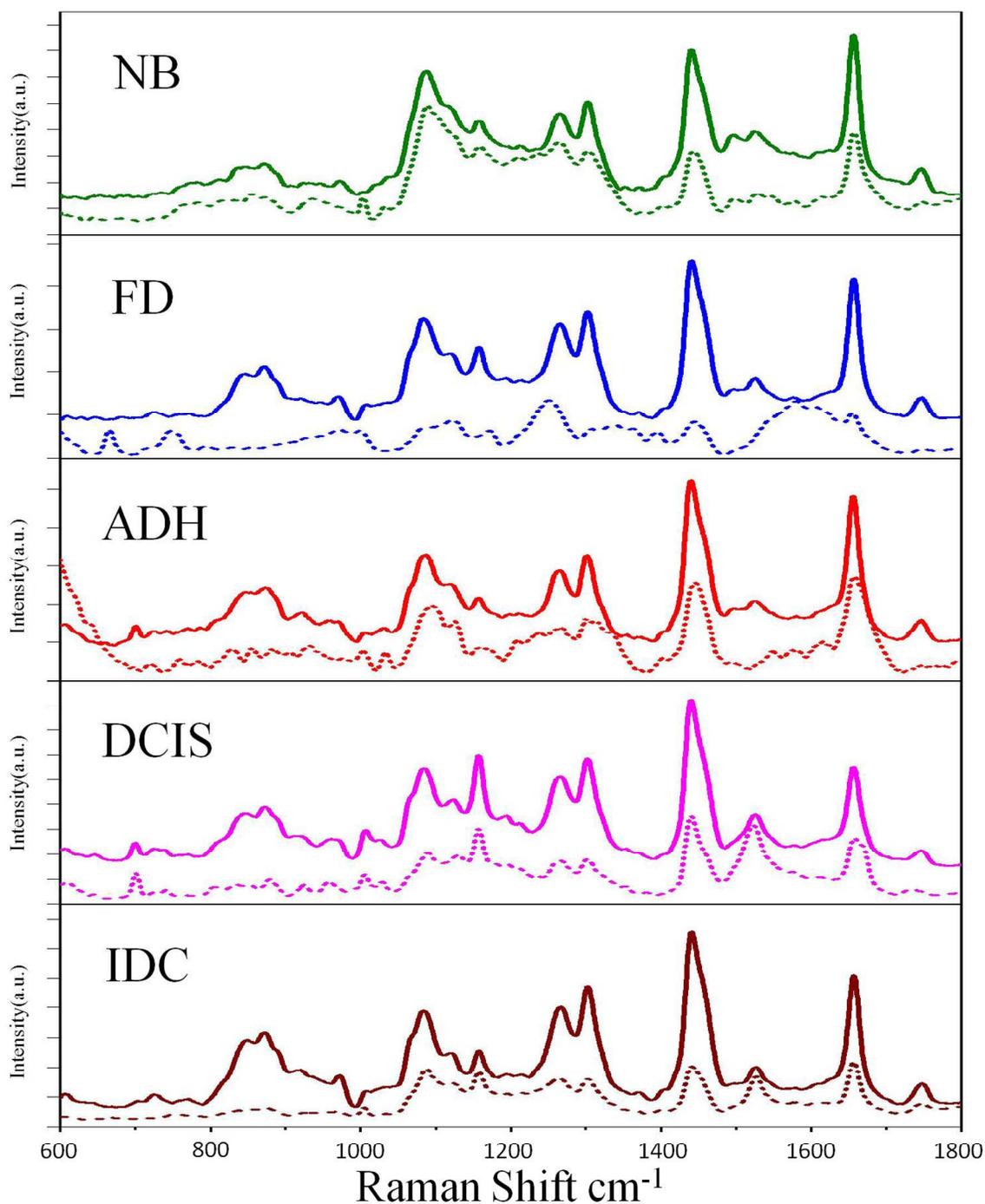


Fig. 2 Average SHINERS spectra (solid line) and spontaneous Raman spectra (dotted line) acquired from fresh frozen sections of normal breast tissue (NB) as well as those displaying the following breast lesions: fibroadenoma (FD), atypical ductal hyperplasia (ADH), ductal carcinoma in situ (DCIS), and invasive ductal carcinoma (IDC).

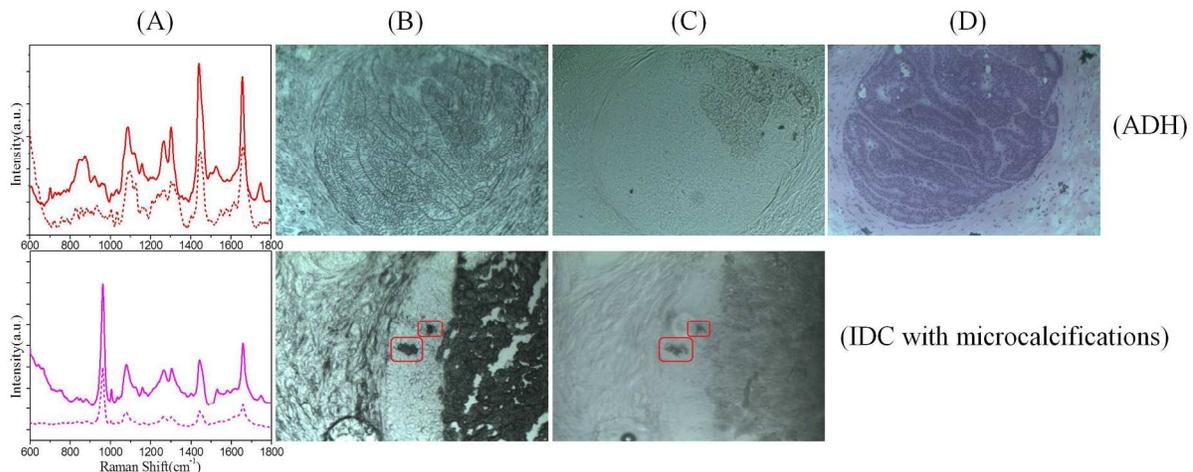


Fig. 3 Spectroscopy histopathology comparisons for tissue section exhibiting ADH (top) and IDC with microcalcifications (bottom). (A) Mean SHINERS (solid line) and spontaneous Raman (dotted line) spectra acquired from the lesions; (B) Images of the fresh frozen sections (without SHINs) that were employed in spontaneous Raman spectra acquisition. The locations of microcalcifications were highlighted by squares; (C) Images of the fresh frozen sections after addition of SHINs. The locations of microcalcifications were highlighted by squares; (D) Corresponding sections stained with H&E. The images represent an approximate area of $0.8 \times 0.6 \text{ mm}^2$.

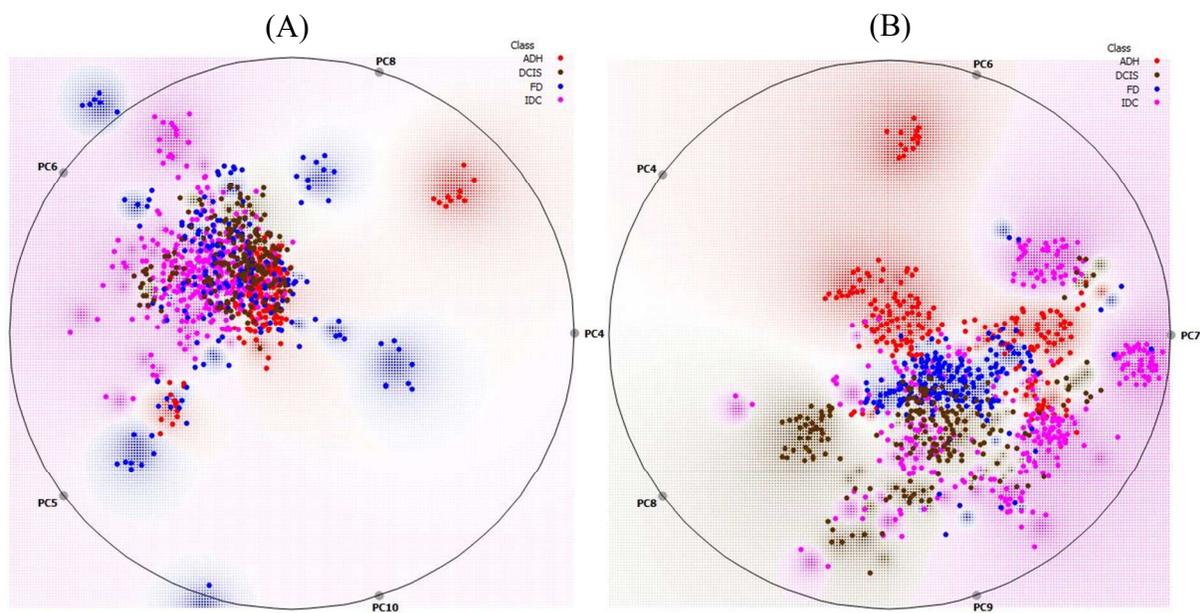


Fig. 4 Multi-dimensional radial visualization plots for principal component (PC) scores corresponding to the five most significant PCs. The specific PCs were selected using a heuristic search algorithm that enables maximal separation of the four diseased groups. (A) Spontaneous Raman spectra; (B) SHINERS spectra. The plots highlight the degree of clustering of sites belonging to a specific pathological class in the absence of microcalcifications for the two spectral datasets.

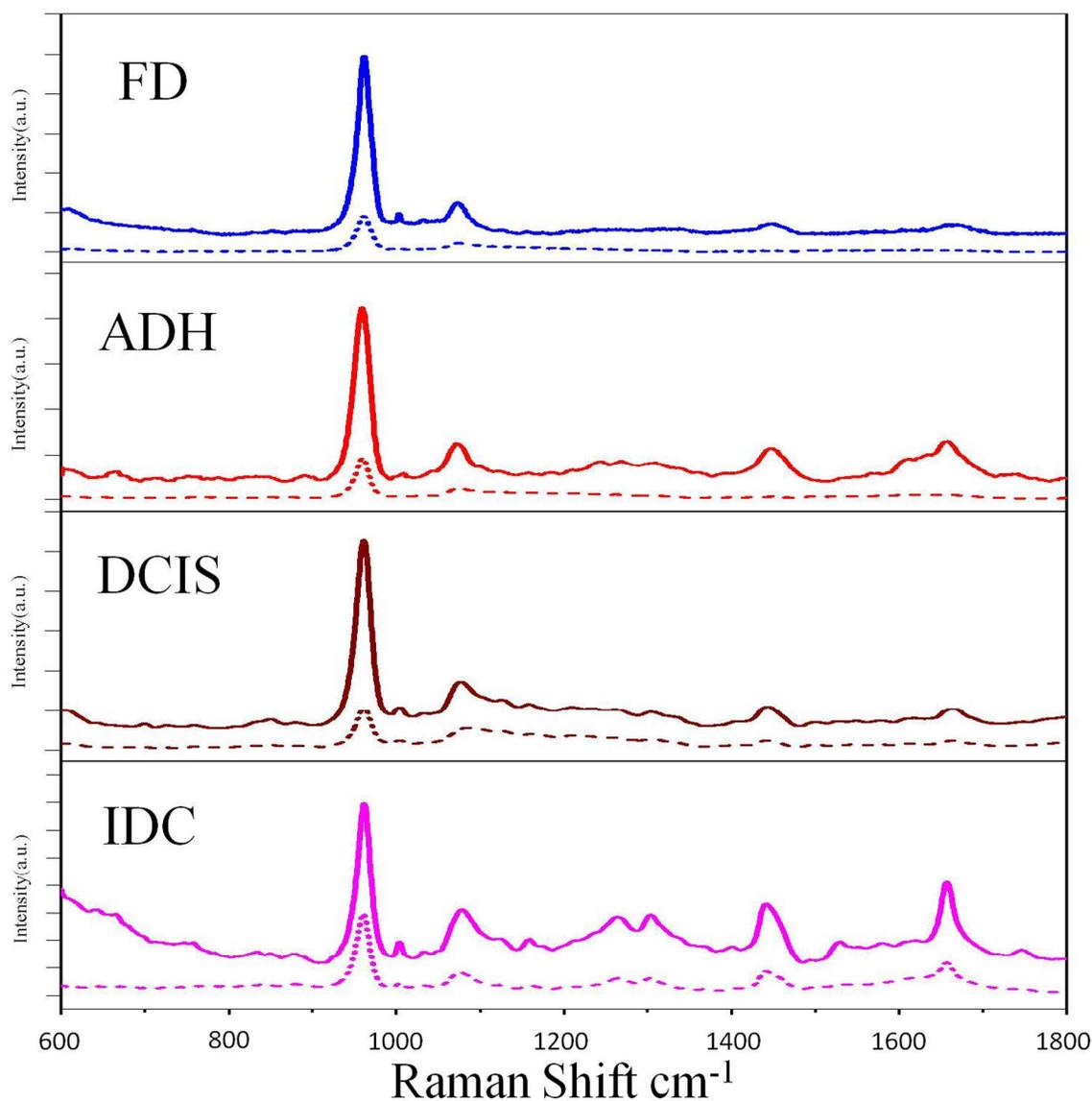


Fig. 5 Average SHINERS spectra (solid line) and spontaneous Raman spectra (dotted line) acquired from tissue sites displaying type II microcalcifications with the following breast pathologies: fibroadenoma (FD), atypical ductal hyperplasia (ADH), ductal carcinoma in situ (DCIS), and invasive ductal carcinoma (IDC).

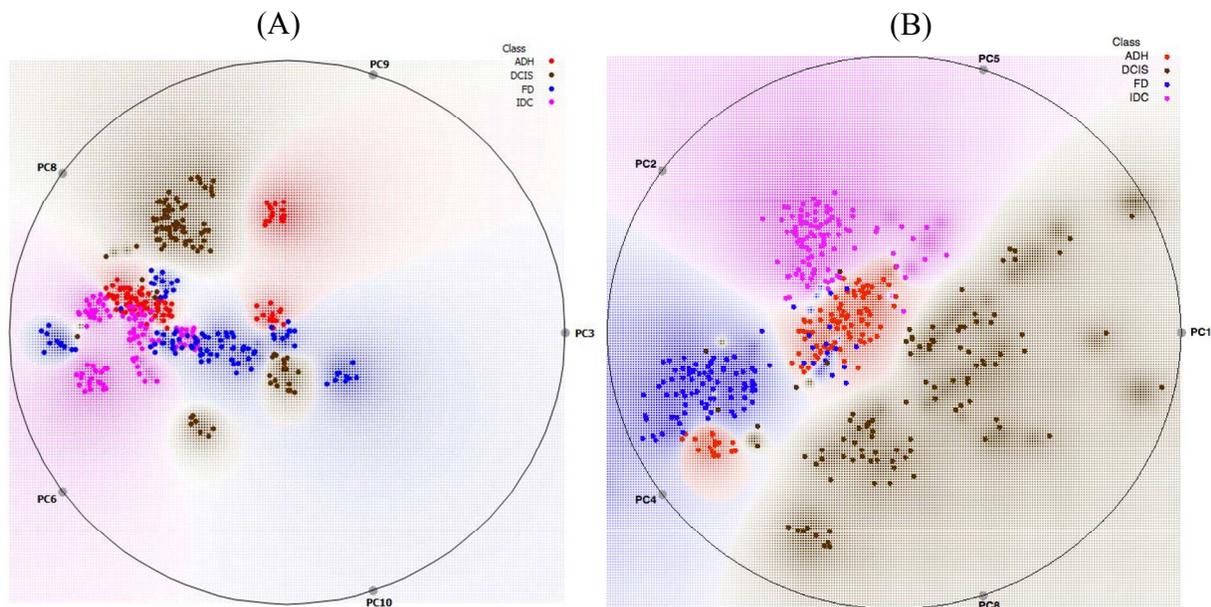


Fig. 6 Radial visualization map using principal component (PC) scores that provide maximal separation between tissues sites belonging to different pathologies and harboring microcalcifications. (A) Spontaneous Raman spectra; (B) SHINERS spectra.

Table 1. Summary of classification accuracies for PLS-DA derived decision algorithms featuring spontaneous Raman and SHINERS spectra, respectively. Standard deviations are noted in the parentheses. In this analysis, spectra and identity labels corresponding to normal breast (NB), fibroadenoma (FD), acute ductal hyperplasia (ADH), ductal carcinoma in situ (DCIS) and invasive ductal carcinoma (IDC) were included.

	Accuracy of spontaneous Raman spectroscopy	Accuracy of SHINERS
NB	72.8% (5.7%)	82.0% (5.1%)
FD	91.0% (4.4%)	92.3% (2.5%)
ADH	97.2% (1.7%)	97.5% (1.6%)
DCIS	87.4% (6.0%)	92.7% (3.9%)
IDC	96.9% (2.0%)	91.8% (4.0%)

Table 2. Summary of classification accuracies for PLS-DA derived decision algorithms featuring spontaneous Raman and SHINERS spectra, respectively. Standard deviations are noted in the parentheses. In this analysis, spectra and identity labels corresponding to tissue sites harboring microcalcifications and displaying the following pathologies were included: fibroadenoma (FD), acute ductal hyperplasia (ADH), ductal carcinoma in situ (DCIS) and invasive ductal carcinoma (IDC).

	Accuracy of spontaneous Raman spectroscopy	Accuracy of SHINERS
FD	99.0% (1.5%)	99.9% (0.2%)
ADH	98.1% (6.3%)	99.9% (1.7%)
DCIS	96.9% (2.6%)	97.5% (2.8%)
IDC	100.0% (0%)	98.2% (1.8%)