Discrimination of liver malignancies with 1064 nm dispersive Raman spectroscopy

Isaac J. Pence,[1] Chetan A. Patil,[2,*] Chad A. Lieber,[3] and Anita Mahadevan-Jansen[1]

[1]*Department of Biomedical Engineering, Vanderbilt University, Nashville, Tennessee 37235, USA*
[2]*Department of Bioengineering, Temple University, Philadelphia, PA 19122, USA*
[3]*Prozess Technologie, 6124 Delmar Blvd, Saint Louis, Missouri 63112, USA*
[*]*c.patil@temple.edu*

Abstract: Raman spectroscopy has been widely demonstrated for tissue characterization and disease discrimination, however current implementations with either 785 or 830 nm near-infrared (NIR) excitation have been ineffectual in tissues with intense autofluorescence such as the liver. Here we report the use of a dispersive 1064 nm Raman system using a low-noise Indium-Gallium-Arsenide (InGaAs) array to discriminate highly autofluorescent bulk tissue *ex vivo* specimens from healthy liver, adenocarcinoma, and hepatocellular carcinoma (N = 5 per group). The resulting spectra have been combined with a multivariate discrimination algorithm, sparse multinomial logistic regression (SMLR), to predict class membership of healthy and diseased tissues, and spectral bands selected for robust classification have been extracted. A quantitative metric called feature importance is defined based on classification outputs and is used to guide the association of spectral features with biological indicators of healthy and diseased liver tissue. Spectral bands with high feature importance for healthy and liver tumor specimens include retinol, heme, biliverdin, or quinones (1595 cm^{-1}); lactic acid (838 cm^{-1}); collagen (873 cm^{-1}); and nucleic acids (1485 cm^{-1}). Classification performance in both binary (normal versus tumor, 100% sensitivity and 89% specificity) and three-group cases (classification accuracy: normal 89%, adenocarcinoma 74%, hepatocellular carcinoma 64%) indicates the potential for accurately separating healthy and cancerous tissues and suggests implications for utilizing Raman techniques during surgical guidance in liver resection.

OCIS codes: (170.5660) Raman spectroscopy; (170.3890) Medical optics instrumentation; (170.6935) Tissue characterization; (170.6510) Spectroscopy, tissue diagnostics; (070.5010) Pattern recognition.

References and links

1. A. M. Enejder, T. W. Koo, J. Oh, M. Hunter, S. Sasic, M. S. Feld, and G. L. Horowitz, "Blood analysis by Raman spectroscopy," Opt. Lett. **27**(22), 2004–2006 (2002).
2. H. Q. Wang, N. Y. Huang, J. H. Zhao, H. Lui, M. Korbelik, and H. S. Zeng, "Depth-resolved in vivo micro-Raman spectroscopy of a murine skin tumor model reveals cancer-specific spectral biomarkers," J. Raman Spectrosc. **42**(2), 160–166 (2011).
3. M. S. Bergholt, W. Zheng, L. Kan, K. Y. Ho, M. Teh, K. G. Yeoh, J. B. Y. So, and Z. W. Huang, "On-line image-guided Raman endoscopy for in vivo diagnosis of gastric cancer," J. Gastroenterol. Hepatol. **26**, 192 (2011).
4. C. A. Lieber, S. K. Majumder, D. L. Ellis, D. D. Billheimer, and A. Mahadevan-Jansen, "In vivo nonmelanoma skin cancer diagnosis using Raman microspectroscopy," Lasers Surg. Med. **40**(7), 461–467 (2008).
5. H. Lui, J. Zhao, D. McLean, and H. Zeng, "Real-time Raman Spectroscopy for In Vivo Skin Cancer Diagnosis," Cancer Res. **72**(10), 2491–2500 (2012).
6. N. Stone, R. Baker, K. Rogers, A. W. Parker, and P. Matousek, "Subsurface probing of calcifications with spatially offset Raman spectroscopy (SORS): future possibilities for the diagnosis of breast cancer," Analyst (Lond.) **132**(9), 899–905 (2007).

#238354 Received 22 Apr 2015; revised 11 Jun 2015; accepted 12 Jun 2015; published 2 Jul 2015

7. M. D. Keller, E. Vargis, N. de Matos Granja, R. H. Wilson, M. A. Mycek, M. C. Kelley, and A. Mahadevan-Jansen, "Development of a spatially offset Raman spectroscopy probe for breast tumor surgical margin evaluation," J. Biomed. Opt. **16**(7), 077006 (2011).
8. A. Mahadevan-Jansen and R. R. Richards-Kortum, "Raman Spectroscopy for the detection of cancers and precancers," J. Biomed. Opt. **1**(1), 31–70 (1996).
9. S. Majumder, E. Kaneter, A. Robichaux, H. Jones, and A. Mahadevan-Jansen, "Near Infrared Raman Spectroscopy for in vivo diagnosis of cervical dysplasia," in *SPIE Photonics West*, 2007)
10. A. C. Society, "Cancer Facts & Figures 2014," (American Cancer Society, Atlanta, GA, 2014).
11. *World Cancer Report 2014* (International Agency for Research on Cancer, Lyon, Fance, 2014).
12. J. Bruix, M. Sherman; American Association for the Study of Liver Diseases, "Management of hepatocellular carcinoma: an update," Hepatology **53**(3), 1020–1022 (2011).
13. O. Heizmann, S. Zidowitz, H. Bourquain, S. Potthast, H. O. Peitgen, D. Oertli, and C. Kettelhack, "Assessment of intraoperative liver deformation during hepatic resection: prospective clinical study," World J. Surg. **34**(8), 1887–1893 (2010).
14. A. M. Kazaryan, I. P. Marangos, B. I. Røsok, A. R. Rosseland, O. Villanger, E. Fosse, O. Mathisen, and B. Edwin, "Laparoscopic resection of colorectal liver metastases: surgical and long-term oncologic outcome," Ann. Surg. **252**(6), 1005–1012 (2010).
15. P. Bao, T. K. Sinha, C. C. Chen, J. R. Warmath, R. L. Galloway, and A. J. Herline, "A prototype ultrasound-guided laparoscopic radiofrequency ablation system," Surg. Endosc. **21**(1), 74–79 (2007).
16. A. Rethy, T. Langø, and R. Mårvik, "Laparoscopic ultrasound for hepatocellular carcinoma and colorectal liver metastasis: an overview," Surg. Laparosc. Endosc. Percutan. Tech. **23**(2), 135–144 (2013).
17. D. Rucker, Y. Wu, L. Clements, J. Ondrake, T. Pheiffer, A. Simpson, W. Jarnagin, and M. Miga, "A mechanics-based non-rigid registration method for liver surgery using sparse intraoperative data," IEEE Trans. Med. Imaging **33**, 147–158 (2013).
18. N. Huang, M. Short, J. Zhao, H. Wang, H. Lui, M. Korbelik, and H. Zeng, "Full range characterization of the Raman spectra of organs in a murine model," Opt. Express **19**(23), 22892–22909 (2011).
19. N. Stone, R. Baker, C. H. Prieto, and P. Matousek, "Novel Raman signal recovery from deeply buried tissue components," in *Biomedical Optical Spectroscopy*, (SPIE, 2008), 68530N 68531–68515.
20. S. R. Hawi, W. B. Campbell, A. Kajdacsy-Balla, R. Murphy, F. Adar, and K. Nithipatikom, "Characterization of normal and malignant human hepatocytes by Raman microspectroscopy," Cancer Lett. **110**(1-2), 35–40 (1996).
21. K. Kochan, K. M. Marzec, K. Chruszcz-Lipska, A. Jasztal, E. Maslak, H. Musiolik, S. Chłopicki, and M. Baranska, "Pathological changes in the biochemical profile of the liver in atherosclerosis and diabetes assessed by Raman spectroscopy," Analyst (Lond.) **138**(14), 3885–3890 (2013).
22. C. Krafft, M. A. Diderhoshan, P. Recknagel, M. Miljkovic, M. Bauer, and J. Popp, "Crisp and soft multivariate methods visualize individual cell nuclei in Raman images of liver tissue sections," Vib. Spectrosc. **55**(1), 90–100 (2011).
23. A. G. Shen, B. J. Zhang, J. Ping, W. Xie, P. Donfack, S. J. Baek, X. D. Zhou, H. Wang, A. Materny, and J. M. Hu, "In vivo study on the protection of indole-3-carbinol (13C) against the mouse acute alcoholic liver injury by micro-Raman spectroscopy," J. Raman Spectrosc. **40**(5), 550–555 (2009).
24. M. C. Gaggini, R. S. Navarro, A. R. Stefanini, R. S. Sano, and L. Silveira, Jr., "Correlation between METAVIR scores and Raman spectroscopy in liver lesions induced by hepatitis C virus: a preliminary study," Lasers Med. Sci. **30**(4), 1347–1355 (2015).
25. A. Lorincz, D. Haddad, R. Naik, V. Naik, A. Fung, A. Cao, P. Manda, A. Pandya, G. Auner, R. Rabah, S. E. Langenburg, and M. D. Klein, "Raman spectroscopy for neoplastic tissue differentiation: a pilot study," J. Pediatr. Surg. **39**(6), 953–956 (2004).
26. S. Keller, B. Schrader, A. Hoffmann, W. Schrader, K. Metz, A. Rehlaender, J. Pahnke, M. Ruwe, and W. Budach, "Application of near-Infrared Fourier-Transform Raman-Spectroscopy in medical-research," J. Raman Spectrosc. **25**(7-8), 663–671 (1994).
27. B. Schrader, B. Dippel, S. Fendel, S. Keller, T. Lochte, M. Riedl, R. Schulte, and E. Tatsch, "NIR FT Raman spectroscopy - A new tool in medical diagnostics," J. Mol. Struct. **408**, 23–31 (1997).
28. C. A. Patil, I. J. Pence, C. A. Lieber, and A. Mahadevan-Jansen, "1064 nm dispersive Raman spectroscopy of tissues with strong near-infrared autofluorescence," Opt. Lett. **39**(2), 303–306 (2014).
29. C. A. Lieber and A. Mahadevan-Jansen, "Automated method for subtraction of fluorescence from biological Raman spectra," Appl. Spectrosc. **57**(11), 1363–1367 (2003).
30. B. Krishnapuram, L. Carin, M. A. T. Figueiredo, and A. J. Hartemink, "Sparse multinomial logistic regression: Fast algorithms and generalization bounds," IEEE Trans. Pattern Anal. Mach. Intell. **27**(6), 957–968 (2005).
31. G. Cassanas, M. Morssli, E. Fabregue, and L. Bardet, "Vibrational-spectra of lactic-acid and lactates," J. Raman Spectrosc. **22**(7), 409–413 (1991).
32. G. J. Puppels, "Confocal Raman microspectroscopy," in *Fluorescent and Luminescent Probes for Biological Activity*, W. Mason, ed. (Academic Press, London, 1999), pp. 377–406.
33. V. J. C. Lin and J. L. Koenig, "Raman Studies of Bovine Serum Albumin," Biopolymers **15**(1), 203–218 (1976).
34. C. J. Frank, R. L. McCreery, and D. C. Redd, "Raman spectroscopy of normal and diseased human breast tissues," Anal. Chem. **67**(5), 777–783 (1995).

#238354 Received 22 Apr 2015; revised 11 Jun 2015; accepted 12 Jun 2015; published 2 Jul 2015

35. A. Mahadevan-Jansen, M. F. Mitchell, N. Ramanujam, A. Malpica, S. Thomsen, U. Utzinger, and R. Richards-Kortum, "Near-infrared Raman spectroscopy for in vitro detection of cervical precancers," Photochem. Photobiol. **68**(1), 123–132 (1998).
36. J. Goral and V. Zichy, "Fourier-Transform Raman Studies of Materials and Compounds of Biological Importance," Spectrochim. Acta A Mol. Biomol. Spectrosc. **46**(2), 253–275 (1990).
37. R. Manoharan, Y. Wang, and M. S. Feld, "Histochemical analysis of biological tissues using Raman spectroscopy," Spectrochim. Acta A Mol. Biomol. Spectrosc. **52**(2), 215–249 (1996).
38. C. J. Frank, D. C. Redd, T. S. Gansler, and R. L. McCreery, "Characterization of human breast biopsy specimens with near-IR Raman spectroscopy," Anal. Chem. **66**(3), 319–326 (1994).
39. A. S. Haka, Z. Volynskaya, J. A. Gardecki, J. Nazemi, J. Lyons, D. Hicks, M. Fitzmaurice, R. R. Dasari, J. P. Crowe, and M. S. Feld, "In vivo margin assessment during partial mastectomy breast surgery using raman spectroscopy," Cancer Res. **66**(6), 3317–3322 (2006).
40. E. M. Kanter, E. Vargis, S. Majumder, M. D. Keller, E. Woeste, G. G. Rao, and A. Mahadevan-Jansen, "Application of Raman spectroscopy for cervical dysplasia diagnosis," J. Biophotonics **2**(1-2), 81–90 (2009).
41. P. S. Hammon, S. Makeig, H. Poizner, E. Todorov, and V. R. de Sa, "Predicting reaching targets from human EEG," IEEE Signal Proc. Mag. **25**(1), 69–77 (2008).
42. *Pathology and Genetics of Tumours of the Digestive System*, World Health Organization Classification of Tumours (IARC Press, Lyon, 2000).

1. Introduction

Raman spectroscopy is a real-time, non-invasive technique well suited for characterizing tissue composition without the need for exogenous dyes or contrast agents. Fundamentally, the myriad of vibrational modes that Raman scattering probes yields a spectrum of features that correspond to the molecular makeup of the sample. Biomedical applications of Raman spectroscopy include analyte [1] and biomarker investigation [2], cancer diagnosis [3–6], and surgical guidance [7]. Unlike other optical or spectroscopic techniques that are sensitive to tissue structure or presence of specific exogenous contrast agents, many biological molecules are inherently Raman active [8]. The excellent intrinsic biochemical specificity of Raman spectroscopy is beneficial for investigation of complex disease states like cancer *in situ* and can enable classification with superior sensitivity and specificity relative to competing spectroscopic approaches [9].

Liver cancer has been among the most common and deadly cancers, accounting for an estimated 33,000 new cases and 23,000 deaths in the US in 2014 [10]. The incidence has been increasing since 2006 and liver cancer mortality was the second highest amongst all cancers worldwide in 2012 [11]. Surgical intervention is often indicated as a potential treatment for liver cancers identified at early stages. Such interventions can include resection or ablative procedures depending upon patient and tumor indications [12–14]. Many techniques for guiding these surgical interventions have been explored, including mapping pre-procedure imaging to the surgical field or the integration of novel intraoperative image guidance systems [15–17]. Prior work has demonstrated the potential of Raman spectroscopy as an intraoperative tool for guiding resection of breast tumors [7] but applying a similar technique in the liver would be challenging due to the strong background autofluorescence that overwhelms the Raman scattered signal [18, 19].

Despite the strong NIR autofluorescence of bulk liver tissue, a number of groups have successfully conducted preliminary investigations of cells and tissues from the liver using specific 785 or 830 nm excitation Raman spectroscopic techniques. Much of this work for liver malignancy and disease has been shown in cell lines and thin slices of tissue, where sample preparation and the use of confocal collection geometry reduces autofluorescence [20–23] at the expense of *in vivo* translational potential. Temporal gating has also been an effective method for recovering Raman signal from liver specimens [19], but implementation requires complex and costly ultrafast laser sources that complicate *in vivo* application. Other groups that have reported Raman spectra directly from bulk liver specimens have either been able to recover only the most prominent peaks in the fingerprint region or have directly acknowledged the challenge autofluorescence presents [18, 24, 25]. For example, Huang et al reported the feasibility of recovering the spectral signature from high-wavenumber regions;

#238354 Received 22 Apr 2015; revised 11 Jun 2015; accepted 12 Jun 2015; published 2 Jul 2015

however, overwhelming autofluorescence intensity precluded the direct collection of Raman fingerprint signals with sufficient signal to noise ratios for interpretation [18]. Unfortunately, the resulting high-wavenumber spectra have not provided as rich a Raman signature for classification of tissue types in comparison to those from the fingerprint region. Alternatively, the lower photon energy of 1064 nm excitation has long been known to further reduce bulk tissue autofluorescence. FT-Raman using 1064 nm excitation has been a successful approach for producing high-quality fingerprint spectra from bulk liver specimens; however, collection times on the order of minutes per spectrum have been required [26, 27] which restricts potential *in vivo* applications.

Demonstrating Raman spectroscopy systems capable of both performing efficient collection of fingerprint spectra and reliable discrimination of diseased and healthy liver would represent valuable progress towards translation of the technique to medical applications. Recent advancements in the manufacturing and production of Indium-Gallium-Arsenide (InGaAs) arrays have resulted in detectors with high quantum efficiency (>80%), minimal or no bad pixels, and reduced dark and readout noise. The use of these detector arrays in high-throughput dispersive spectrometers has enabled a report of Raman instruments that use an excitation wavelength of 1064 nm to acquire spectra from bulk tissues with strong infrared autofluorescence [28] without the need to rely on slow FT-Raman systems. However, the previous work has neither addressed the potential for classification of the Raman spectra of the liver acquired with 1064 nm excitation and dispersive detection, nor the comparison of spectral features which may be used to discriminate liver cancer status. Here, we report the potential of 1064 nm dispersive Raman spectroscopy to differentiate malignant liver specimens from healthy liver specimens, to perform direct multivariate classification of normal liver, primary hepatocellular carcinoma, and secondary adenocarcinoma tumor specimens, and to quantify the importance of biologically relevant spectral features for future investigation.

2. Methods

Tissue measurements were acquired using two separate Raman spectroscopy systems with excitation wavelengths at 785 and 1064 nm, which are described in detail elsewhere [28]. Briefly, the illumination and collection axes of the two systems (Fig. 1) were co-aligned in order to ensure measurements were made from the same location on the sample. A flip mirror was placed in the optical path in order to toggle between 785 and 1064 nm Raman configurations. The 1064 nm system employed an InGaAs array with a multistage thermoelectric cooler locked at a temperature of −60°C (Bayspec, Inc.). The entire 1064 nm system was packaged into a single device with a common aperture for excitation and emission. Measurements were made with 30 second integration times delivering 204mW from the 1064nm diode laser at the sample for a calculated spot size of approximately 25 microns. The optical configuration of the 785 nm Raman system was designed to match the 1064 nm system as closely as possible. The detector for the 785 nm system was a back-illuminated deep-depletion CCD thermoelectrically cooled to −70°C (Princeton Instruments, Inc.). Identical absolute and relative wavenumber calibrations were performed on each system with a neon-argon lamp and common spectroscopic standards. Wavelength-dependent variations in system response were accounted for in both systems with a NIST-calibrated tungsten lamp. A three-axis micrometer-driven stage was used to position the sample. Optical alignment and recoupling efficiency was optimized by placing a silicon wafer at the sample focus and maximizing the intensity of the 520 cm^{-1} silicon peak.

#238354 Received 22 Apr 2015; revised 11 Jun 2015; accepted 12 Jun 2015; published 2 Jul 2015

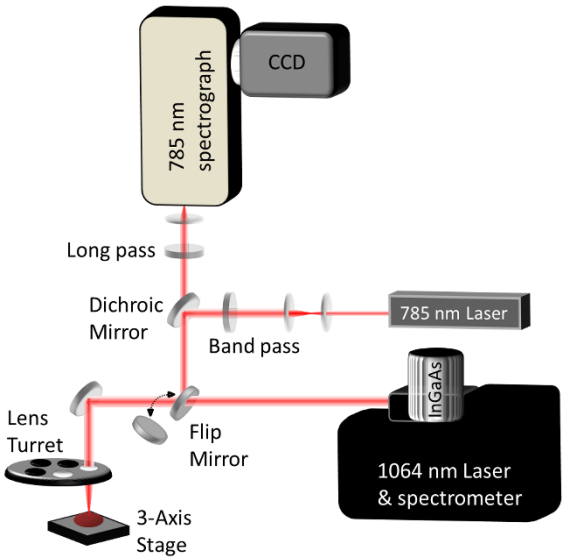

Fig. 1. Schematic of co-aligned Raman spectroscopic systems. The 785 nm system utilized beam shaping optics to best match the excitation parameters of the 1064 nm system.

Ex vivo specimens of primary hepatocellular carcinoma (N = 5), secondary liver tumor (metastatic colorectal adenocarcinoma, N = 5), and healthy human liver (N = 5) from unique donors were obtained from the Cooperative Human Tissue Network at Vanderbilt after approval of exemption from the Vanderbilt Institutional Review Board. Tissues were maintained at −80°C until the time of use, at which point they were brought to room temperature in phosphate buffered saline solution. Specimens were then positioned for measurement on the sample stage, and multiple spectra were acquired from at least five physical locations across each specimen. All spectra shown have been corrected for wavelength-dependent differences in system response. Fixed pattern noise from the InGaAs array was subtracted from the 1064 nm spectra. Additional processing steps included baseline subtraction with a modified polynomial fitting algorithm [29] and noise smoothing with a second-order Savitzky–Golay filter.

The resulting spectra were then classified with a Bayesian machine learning algorithm, sparse multinomial logistic regression (SMLR), to quantitatively determine the potential for 1064 nm dispersive Raman spectra to separate healthy and malignant liver tissues. SMLR is a versatile multiclass iterative algorithm that reduces the high dimensionality of Raman data to only those spectral basis features needed for discrimination [30]. SMLR data reduction involved creating a transformation of the original data set in which distinguishing spectral basis features were weighted based on their ability to successfully separate classes of training data. The training and classification procedure implemented here used a Laplacian prior, a direct kernel, no bias, z-scored spectral normalization, component–wise updates, and leave-one-sample-out cross validation. In this case, the leave-one-sample-out procedure trained an independent classifier for each sample (i.e. each tissue specimen). A posterior probability of class membership for each class was then calculated for each individual spectrum using a classifier trained only with spectra from other specimens. The final classification accuracy reported for each tissue type represents the percentage of all spectra correctly predicted. Two separate classification tests were performed on the data set: a binary prediction of tumor versus normal samples and a three class prediction of healthy liver, secondary colorectal metastatic adenocarcinoma, and primary hepatocellular carcinoma.

To evaluate the relative merit of spectral features identified for classification, two different outputs produced by SMLR were used. The first output was the set of averaged

#238354 Received 22 Apr 2015; revised 11 Jun 2015; accepted 12 Jun 2015; published 2 Jul 2015

weights SMLR applied to spectral basis features. In this instance, the classifier iteratively applied weights to basis features that directly corresponded to spectral bands in the original data. As a result, the average magnitude of weight assigned to a spectral feature was proportional to the band's value for classification. Another output used was the frequency with which certain spectral basis features were selected across all cross-validations. Since the human tumor specimens can exhibit substantial heterogeneity in their Raman signature within a given pathological group, cross-validation frequency can also be a valuable measure of importance. We have defined a quantitative metric called feature importance that incorporates both the SMLR feature weight and the cross-validation frequency in order to quantitatively consider both biochemical differences between distinct tissues types and spectral heterogeneity within a single tissue type. Since cross-validation frequency was inherently scaled from [0, 1] but average weights were unbounded, the magnitude of average weights for each basis feature across all classes was scaled to a range of [0, 1] in a given comparison (either binary or three class). Here we have defined feature importance as the product of feature cross-validation frequency and scaled feature weight. As frequency and scaled weights both have range [0, 1], the resulting feature importance value emphasizes features that are both heavily weighted for decisions and chosen consistently. Values for feature importance have been calculated independently for binary and three class discrimination processes.

3. Results

A comparison between calibrated spectra obtained from the liver samples with 785 nm and 1064 nm dispersive Raman spectroscopy systems prior to subtraction of background autofluorescence is displayed in Fig. 2. The spectra collected at 785 nm suffer from intense, highly variable autofluorescence that dominates the Raman scattering intensity and makes separation of Raman data and interpretation of the Raman spectrum difficult to impossible. Conversely, all spectra measured with the 1064 nm excitation system demonstrate evident Raman features prior to removal of the autofluorescent background. Visual inspection of the average spectra for normal, adenocarcinoma, and hepatocellular carcinoma samples (Fig. 3) reveals differences in the lineshapes at several regions, most notably the 1595 cm^{-1} band of the normal samples. This spectral band, and others that contain the greatest differences between classes, coincide with the features selected by SMLR as part of the multivariate signature used for discrimination. The SMLR weight, cross-validation frequency, and feature importance of the key spectral bands used for discrimination are listed in Table 1. In general, spectral bands with high average SMLR weighting correspond to those that are frequently used in cross-validation in both the binary (Fig. 4) and three class instances (data not shown), although the trend is clearly non-linear in both cases. Feature importance is encoded in Fig. 3 by the shaded vertical gray bands.

#238354 Received 22 Apr 2015; revised 11 Jun 2015; accepted 12 Jun 2015; published 2 Jul 2015

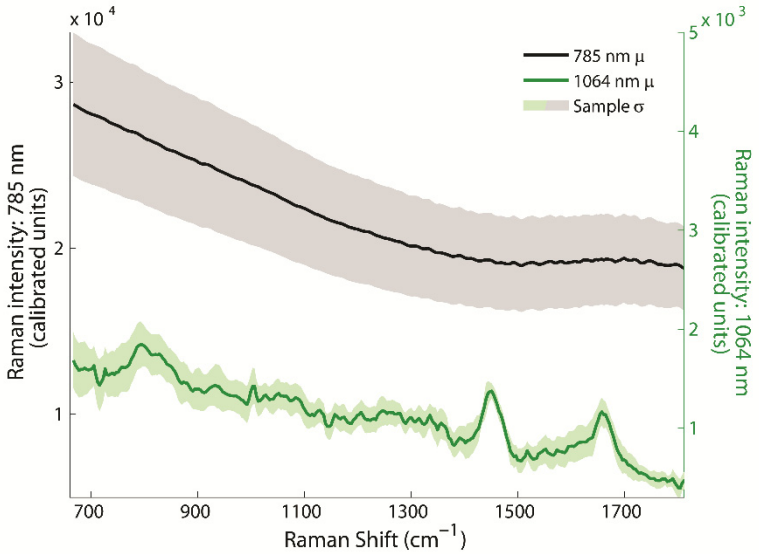

Fig. 2. Raman spectra (mean +/– standard deviation) of adenocarcinoma samples for both 785 nm and 1064 nm excitation systems. The strong and variable autofluorescent background present at 785 nm excitation overwhelms modest Raman signal from each spectrum. Spectral trends are representative of all classes measured.

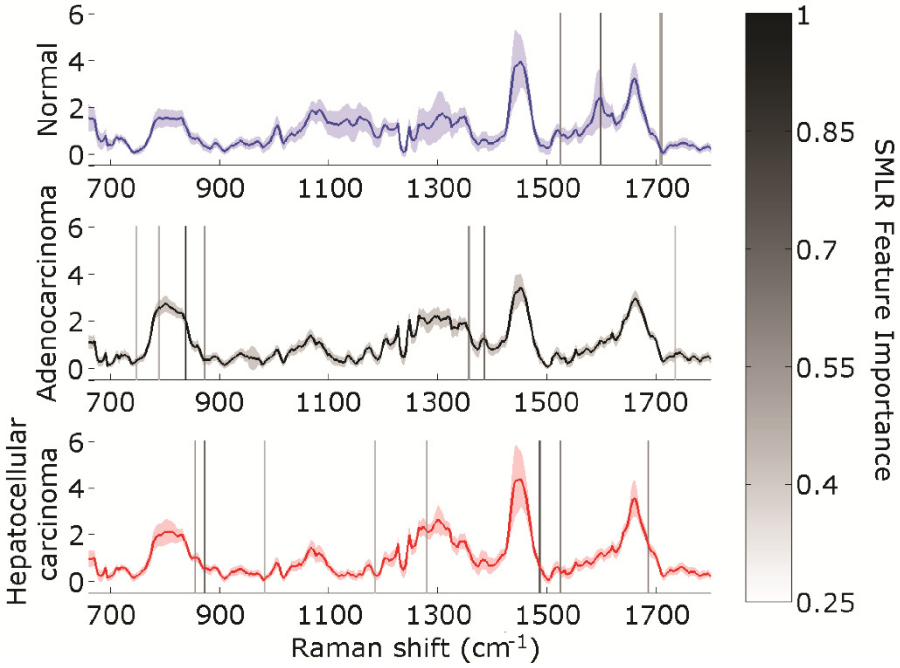

Fig. 3. Mean +/– standard deviation Raman spectra collected with 1064 nm dispersive system for healthy liver, secondary colorectal metastatic adenocarcinoma, and primary hepatocellular carcinoma samples. All spectra are corrected for wavelength-dependent spectral response and background subtracted using a modified polynomial fitting algorithm previously described. Spectral features utilized during SMLR classification are encoded with graded bands indicating the importance greater than the 0.25 threshold value.

#238354 Received 22 Apr 2015; revised 11 Jun 2015; accepted 12 Jun 2015; published 2 Jul 2015

Table 1. Key SMLR classifier spectral bands, feature importance, and assignments.

Raman shift (cm^{-1})	SMLR Weight	Cross-validation Frequency (%)	Feature Importance	Tissue Type	Assignment
747	0.46	93	0.27	Adenocarcinoma	Lactic acid [31]
789	0.53	87	0.29	Adenocarcinoma	DNA [32]
838	1.58	100	1.00	Adenocarcinoma	Lactic acid [31], Tyrosine [32]
855	0.67	100	0.42	Hepatocellular Carcinoma	Albumin [33]
873	1.06	100	0.67	Hepatocellular Carcinoma	Collagen, hydroxyproline [34]
	–0.75	100	0.48	Adenocarcinoma	
981	–0.54	87	0.30	Hepatocellular Carcinoma	–
1184	–0.49	87	0.27	Hepatocellular Carcinoma	Tyrosine, Phenylalanine [32]
1279	0.63	87	0.35	Hepatocellular Carcinoma	Amide III [35]
1356	0.70	100	0.44	Adenocarcinoma	Glucoe [35, 36]
1384	0.91	100	0.57	Adenocarcinoma	Lactic acid [31]
1485	1.00	100	0.64	Hepatocellular Carcinoma	Nucleic acid [37]
1524	0.76	93	0.45	Normal	Carotenoid [8, 38]
	–0.80	93	0.47	Hepatocellular Carcinoma	
1595	0.94	100	0.60	Normal	Retinol [21], Heme [21], Biliverdin [26], Quinones [24]
1685	0.83	93	0.49	Hepatocellular Carcinoma	Amide I [8, 34]
1706	–0.70	93	0.35	Normal	–
1734	0.46	93	0.27	Adenocarcinoma	Lipid [38]

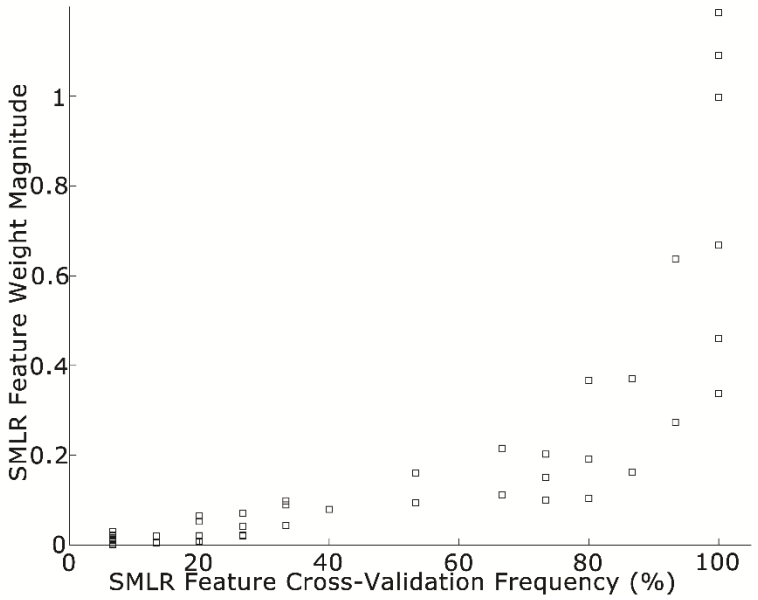

Fig. 4. The magnitude of SMLR feature weights and corresponding cross-validation frequencies for tumor samples from binary classifier. The non-linear association between these metrics holds for all tissue classes for both binary and three class discrimination tests. Frequently chosen features have a large range of weights suggesting a combination of these values will emphasize informative features with a higher feature importance value.

The prediction results of the binary classification model, which utilized an average of 33 basis features (324 features excluded), are presented in Fig. 5 and reveal that every spectrum collected from each tumor sample is correctly classified. Most of the spectra from the healthy liver samples are correctly identified with four spectra from two normal samples errantly predicted as tumor. A binary classifier requires a minimum decision threshold of 50%. Using

#238354 Received 22 Apr 2015; revised 11 Jun 2015; accepted 12 Jun 2015; published 2 Jul 2015

this threshold, the algorithm achieved 100% sensitivity and 89.2% specificity for discriminating healthy and tumor samples. More rigorous separations of the predicted data using only spectra classified with probabilities of either >75% or >80% for class membership produces identical results for both sensitivity and specificity, which reinforces the high confidence of differentiating healthy and tumor tissues using 1064 nm dispersive Raman spectroscopy.

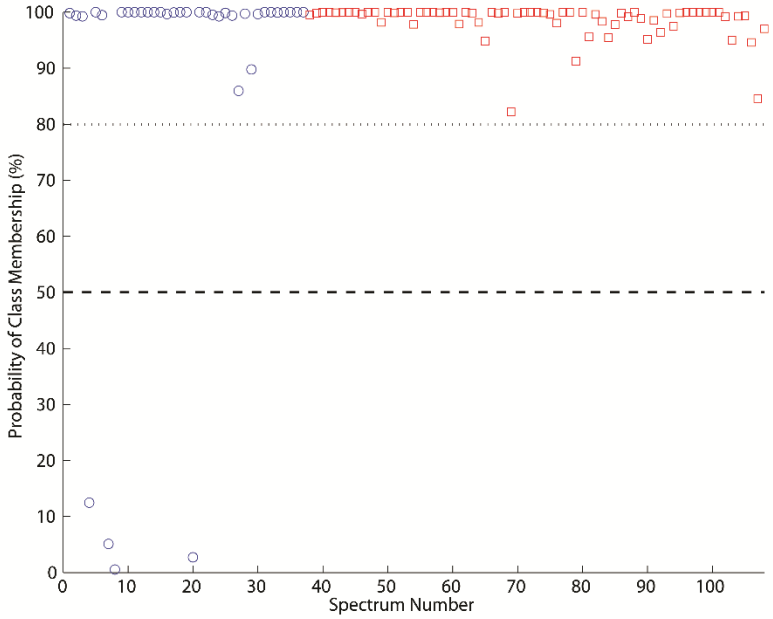

Fig. 5. Binary posterior probability of class membership as predicted by SMLR algorithm. Resulting classification yields 90.3% classification accuracy (104/108 spectra), with 100% sensitivity and 89.2% specificity for discriminating healthy (circles) and tumor samples (squares). Thresholds for >50% (dashed) and >80% (dotted) probability for class membership yield identical results.

Following the binary prediction, the same data is again used for a more challenging three class discrimination of healthy liver, primary hepatocellular carcinoma, and secondary adenocarcinoma of the liver based on a model that utilized an average of 50 basis features (314 features excluded) (Fig. 6). In a three group classifier, membership is assigned to the class with the highest probability, which can be as low as 34%. Here, four measurements from two of the healthy samples are again misclassified (three as hepatocellular carcinoma, one as adenocarcinoma) resulting in 89.2% classification accuracy. Across the tumor classes 26/35 (74.3%) adenocarcinoma and 23/36 (63.9%) hepatocellular carcinoma spectra are correctly identified. When we restrict correct classification to those spectra with greater than 50% probability to mirror the binary classifier, 33/37 (89.2%) normal, 25/35 (71.4%) adenocarcinoma, and 22/36 (61.1%) hepatocellular carcinoma spectra are correctly identified. It is important to note that while spectra from several tumor samples are misclassified by the predictor, no spectrum obtained from a tumor sample is identified as normal. When the rigor of the classification threshold is increased further, an expected falloff in the separation of the primary and secondary tumor classes is observed. A probability threshold of >75% results in correct identification of 31/37 (83.8%) normal, 17/35 (48.6%) adenocarcinoma, and 20/36 (55.6%) hepatocellular carcinoma spectra.

#238354 Received 22 Apr 2015; revised 11 Jun 2015; accepted 12 Jun 2015; published 2 Jul 2015

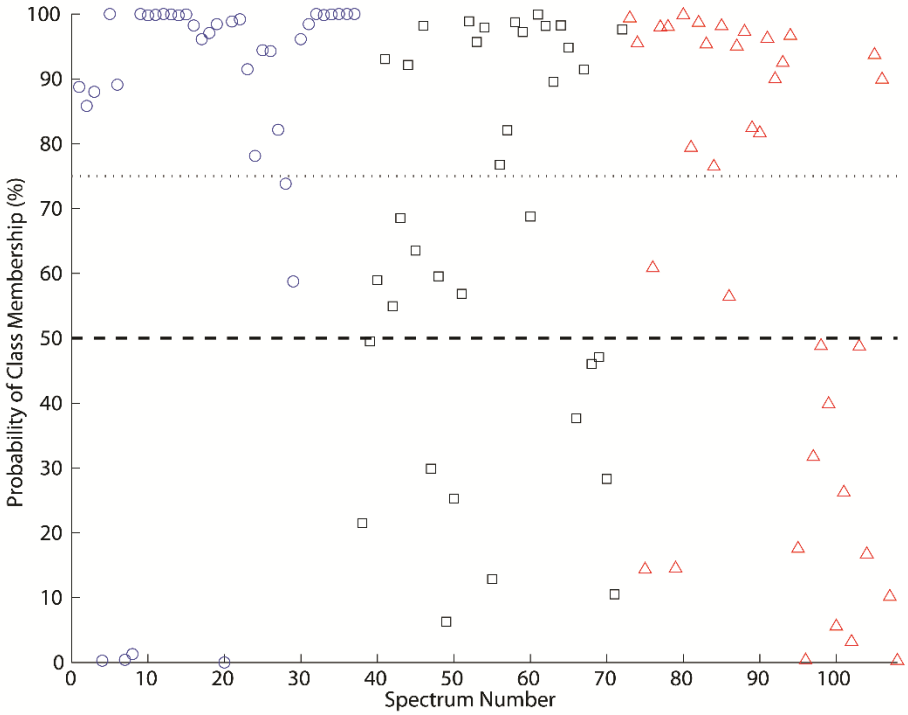

Fig. 6. Three class prediction of posterior probability of class membership from SMLR algorithm. Resulting classification yields 78.3% classification accuracy with healthy control (circles, 33/37 spectra), adenocarcinoma (squares, 26/35 spectra), and hepatocellular carcinoma (triangles, 23/36 spectra) samples correctly distinguished. Again, thresholds are indicated for >50% (dashed) and >80% (dotted) probability for class membership and demonstrate spectral heterogeneity manifested in the falloff in classification accuracy.

4. Discussion

A growing body of work supports the potential for disease discrimination with Raman spectroscopy. Such systems have typically used dispersive detection and excitation sources at either 785 or 830 nm; however, the naturally strong autofluorescence of tissues such as the liver has limited the feasibility of diagnostic applications. Previous work identified the strengths and limitations of using 1064 nm excitation and dispersive spectrographs with InGaAs detector arrays for collection of Raman spectra from bulk tissue [28]. In this work, high-throughput, dispersive detection facilitated acquisition times of 30 seconds or less, which represents a substantial improvement in speed in comparison with established 1064nm FT-Raman spectroscopy systems. One tradeoff is that the reduced scattering cross-sections at 1064 nm in comparison to 785 nm produce weaker Raman signals in tissues with low to moderate autofluorescence. However, in tissues with strong autofluorescence such as the liver (Fig. 2), the decreased autofluorescence seen using 1064 nm excitation is critical for obtaining spectra with sufficient quality for real-time, non-destructive biochemical analysis. Qualitative inspection of the spectrum shown in Fig. 3 reveals the presence of clear Raman signatures in the region from 750 – 1750 cm^{-1}, a region rich with features attributed to nucleic acids, proteins, metabolic factors, and lipids. In order to develop the application of Raman spectroscopy for analysis of bulk liver tissue, we chose to investigate whether robust statistical differences between healthy and cancerous tissues could be found and attempt to provide insight into the biochemical basis of the spectral features that have been used for classification.

#238354 Received 22 Apr 2015; revised 11 Jun 2015; accepted 12 Jun 2015; published 2 Jul 2015

Based on visual inspection alone, there appear to be qualitative differences in the spectral lineshapes of healthy liver, primary hepatocellular carcinoma, and secondary adenocarcinoma specimens. For example, normal tissue appears to be characterized by the presence of a peak at 1595 cm^{-1} that is absent in specimens from both tumor classes (Fig. 3). Statistical analysis of the data using SMLR strongly supports the ability to utilize both clearly visible spectral differences along with those that are visually more subtle to achieve excellent classification performance (Fig. 5). As a Bayesian classifier, one of SMLR's valuable characteristics is that it produces probabilities of class membership for all spectra. Thus, the user can define the probability threshold for classification and is free to choose more (>80% posterior probability) or less (>50% posterior probability) stringent criteria. We have shown that the spectral differences between normal and cancerous liver tissue are distinct enough that a decision threshold can be set at an 80% probability of class membership without any deterioration in performance from the minimum threshold of 50%. Even at this high level of rigor, the performance of the classifier (100% sensitivity, 89.2% specificity, 90.3% classification accuracy) compares very well with the sensitivities (86-100%) and specificities (68-100%) that have been reported for established applications of Raman techniques for diagnostics [5, 7, 39, 40] and suggests successful translation from basic science to medical applications is possible.

Partitioning the data set further for the three class SMLR discrimination of liver samples also presents promising results (Fig. 6). The same four measurements from two of the healthy samples are again misclassified, maintaining 89.2% classification accuracy. When attempting to separate the two tumor classes, the performance degraded, yielding 74.3% and 63.9% classification accuracies for adenocarcinoma and hepatocellular carcinoma, respectively. Of the 3 group misclassifications, it is notable that 25 of the 40 spectra did not receive greater than 75% posterior probability of membership for any class and are, therefore, unclassified based on this decision threshold. The reported classification performance is likely due to a combination of effects. The limited sample size of this preliminary study may not have provided the needed power to discriminate between tumor classes. By using leave-one-sample-out cross-validation, the trained models will have low bias, but may have higher variability as individual samples are omitted. However, the resulting model should generalize appropriately as sample sizes are expanded because this cross-validation technique balances complexity, bias, and variance components. Another factor that has impacted the classification performance is the similarity of the spectral lineshapes for the tumor classes. While these signals enabled confident separation of normal from tumor, the spectral similarities required the classifier to isolate more subtle differences between tumor classes which may have been subject to noise and inter-sample variations. These similarities in Raman signature for tumor classes partially account for posterior probabilities below 75% for many of the misclassified samples.

While the accuracy of statistical classification algorithms provides guidance on the potential for discriminating different tissue types, the strength of Raman spectroscopy is its inherent biochemical specificity. The ability to determine which spectral bands contribute to discrimination and the extent of their contribution would provide valuable insight into the biological and chemical basis of the differences observed between tissue types. Prior work has utilized the SMLR weights in order to evaluate the relative importance that individual features contribute to discrimination [41]. However, since SMLR promotes sparsity and weights are calculated during individual cross-validation operations, it is possible for a spectral feature to be substantially weighted but not used consistently across the entire data set. This situation could arise when conducting feasibility studies on novel applications, where it is unknown to what extent the sample spectra are representative of the overall population, as is often the case in initial investigations of Raman spectroscopy for identification of complex diseases like cancer. It is possible that there may be some bands within a subset of spectra that greatly influence the separation of pathologically identified

#238354 Received 22 Apr 2015; revised 11 Jun 2015; accepted 12 Jun 2015; published 2 Jul 2015

disease classes (high weighting), but that are not universally critical for separation of pathological classes (low frequency). This phenomena is observed in the data shown in Fig. 4, where 9 features have weights above 0.3, however the frequency with which these features are used in cross-validation varies from 80 to 100%. The extracted feature weights and cross-validation frequencies are clearly related measures, but each captures unique information about the spectral bands that inform the classification algorithm (Fig. 4). Here, we have combined spectral feature weight, which captures information that better explains differences between tissue types (between group variability), with cross-validation frequency, which indicates the heterogeneity of the spectral feature within a tissue type (within-group variability), in order to account for the complementary information provided by each. To our knowledge, this work provides the first demonstration of extracting the importance of features used for SMLR classification for Raman spectral data. While this analysis has utilized SMLR as the multivariate tool for feature reduction and cross-validated classification, all statistical pattern recognition algorithms have parameters that can be tuned to meet analytical needs and allow extraction of complementary outputs that enable direct quantitative evaluation of the influence and robustness of selected features. This technique is broadly applicable to other analytic approaches and provides a valuable basis for utilizing classification outputs to assist evaluation of spectra collected during application development. By extracting and combining both feature weight and frequency of use, it is possible to associate the strongest sources of discrimination with vibrational bonds related to cancer status (Table 1). Furthermore, by identifying a robust set of spectral features, we can begin to explore the differences in the heterogeneous subclasses of liver tumors using dispersive 1064nm Raman techniques.

The spectral features ascribed the highest feature importance within the normal samples are the presence of the peak shoulders near the 1524 cm^{-1} and 1706 cm^{-1} bands and the peak at 1595 cm^{-1}. The 1524 cm^{-1} band (FI = 0.45) selected is most likely a shoulder of the 1517 cm^{-1} carotenoid peak. The 1706 cm^{-1} band (FI = 0.35) appears to be a shoulder of an unattributed 1690 cm^{-1} peak. The most important feature, the 1595cm^{-1} peak (FI = 0.6) has previously been observed in Raman spectra of healthy liver and potentially attributed to vitamin A (retinol), heme, biliverdin, or quinones. Retinol is primarily stored in the healthy liver and is a byproduct of carotenoid breakdown, whose spectral features could correspond to the other bands identified as characteristic of normal spectra. A potential explanation for heme as the signal source in non-diseased samples is that healthy liver tissue plays a central role in heme synthesis. Similarly, the liver is responsible for hemoglobin breakdown, resulting in biliverdin, which could account for the presence of this molecule. Some of the differences identified for primary hepatocellular carcinoma samples include significant bands that may be associated with collagen content (FI = 0.67), nucleic acids (FI = 0.64), carotenoids (FI = 0.47), and albumin content (FI = 0.42, Table 1). Cirrhosis, or the replacement of healthy liver tissue with scar tissue, is the primary risk factor for hepatocellular carcinoma [10]. This increased scar tissue content in primary tumor samples may account for the importance of collagen features for discrimination. Increased nucleic acid content is widely associated with cancer status. The presence of liver export proteins, including albumin, is a key histopathological feature for distinguishing primary and secondary liver tumors [42]. The secondary metastatic colorectal tumor samples have significant features associated with lactic acid (FI = 1.0 & 0.57), collagen (FI = 0.48), and glucose (FI = 0.44, Table 1). Lactic acid is an important metabolic component in tumors which rely more heavily on anaerobic glycolysis than healthy tissue. Anaerobic metabolic processes are supported by both glucose and lactic acid signals, and are consistent with the binary classifier. Furthermore, the Cori cycle occurs in healthy liver tissue, converting lactate into glucose, and may be disrupted in tumor tissues.

The spectra measured from the tumor classes are similar and, therefore, require the classifier to select features with less prominence to distinguish between primary and secondary tumor samples. Subtle features that may have informed discrimination are given

#238354 Received 22 Apr 2015; revised 11 Jun 2015; accepted 12 Jun 2015; published 2 Jul 2015

importance levels near 0.25, including DNA (FI = 0.29), lactic acid (FI = 0.27), and lipids (FI = 0.27) for adenocarcinoma samples, and collagen (0.35) and proteins (0.27) for hepatocellular carcinoma samples. Metabolic and proliferative markers are common non-specific cancer indicators and support the multivariate signature for discriminating samples with high sensitivity and specificity. These relatively weak features contribute to the complex spectral signal that has enabled discrimination of tumor samples. Although the three group classification performance decreased relative to the binary discrimination, no spectrum obtained from a tumor sample was identified as normal. Further evaluation of a larger sample set of primary and secondary tumor specimens may clarify the sources and prominence of these important spectral features.

The demonstrated sensitivity of Raman spectral signatures to disease status yields promising preliminary evidence for further investigation for the liver. In addition to simply classifying normal and tumor, rapidly and accurately identifying primary and secondary liver tumors may have significant impact to guide medical intervention. The potential for discriminating these tumors *in vivo* may facilitate improved patient outcomes for resection and ablative procedures in the future, and inform other treatment options *in situ*. The identification of a robust set of spectral features that contribute to the separation of disease classes and are associated with biologically relevant processes in the liver can also help direct future investigations. These advancements, along with the demonstrated discrimination of healthy, primary, and secondary tumor tissue samples, may have implications for improved surgical guidance to conserve liver tissue during resection or to direct ablative procedures *in vivo*. While the work performed in this manuscript relied on bulk, free-beam optics, the approach can be readily translated into a clinical setting by integration with fiber-optic probes. In addition, Raman spectroscopy performed with 1064 nm excitation and dispersive detection opens the door to potential investigations of biomarkers and analytes in the liver, as well as other tissues with strong autofluorescence that have not yet been widely investigated *in situ*, including the kidney and deeply pigmented skin. As these new lines of investigation develop, it will be important to have analytical methods in place that establish the feasibility of classification while also providing insight into the biological and chemical basis of discrimination. Feature importance is a valuable metric in this regard, and is well suited to quantifying biochemical features critical for discrimination.

4. Conclusion

The use of Raman spectroscopy for bulk tissue characterization has seen widespread investigation for identification, diagnostics, and guidance applications for several conditions including cancer. However, the high levels of background signal from tissues with strong NIR autofluorescence present a challenge for practical implementation. Previously, only limited investigation of liver tissues has been feasible using Raman instruments based on 785 or 830 nm excitation or by FT-Raman instruments that could not be easily adopted in a medical setting. Now, by using InGaAs detector arrays and high-throughput dispersive instruments with NIR excitation above 1000 nm, Raman measurements have been rapidly acquired from regions of the spectrum that are less susceptible to the strong liver autofluorescence that usually overwhelms subtle Raman signals. High quality Raman spectra obtained from liver specimens have been measured and classified using a powerful multivariate technique that has successfully separated samples based on tissue group in both normal versus tumor and normal versus primary versus secondary metastatic liver tumor regimes with high accuracy. This classification scheme has also enabled extraction of important spectral features for discrimination with Raman spectroscopic data for the first time. The spectral bands with highest feature importance for healthy and liver tumor specimens, including lactic acid (838 cm^{-1}); collagen (873 cm^{-1}); nucleic acids (1485 cm^{-1}); and retinol, heme, biliverdin, or quinones (1595 cm^{-1}), are combined with less prominent features to enable discrimination based on biologically relevant features with high accuracy. Evaluating the importance of the

#238354 Received 22 Apr 2015; revised 11 Jun 2015; accepted 12 Jun 2015; published 2 Jul 2015

spectral bands selected by the classification algorithm has demonstrated a strong association with biological features of healthy and cancerous liver and provides a foundation for further study. The combination of these advances has enabled a direct investigation of liver tissues without the impact of NIR autofluorescence and demonstrates promise for characterizing and discriminating normal and liver tumor specimens. The suppression of autofluorescence from bulk tissue may also enable Raman spectroscopic investigations of diseases and biomarkers that were not permitted by previous technology and may have potential for future applications including surgical guidance.

Acknowledgments

The authors would like to acknowledge BaySpec, Inc, particularly Dr. William Yang, for the loan of the 1064 nm Raman instrument. Note: at the time of these studies, author CA Patil was at Vanderbilt University and author CA Lieber was an employee of BaySpec. Tissue samples were provided by the Cooperative Human Tissue Network which is funded by the National Cancer Institute. Other investigators may have received specimens from the same subjects.

#238354 Received 22 Apr 2015; revised 11 Jun 2015; accepted 12 Jun 2015; published 2 Jul 2015