

Journal of Biomedical Optics

SPIEDigitalLibrary.org/jbo

Spectral analysis of human saliva for detection of lung cancer using surface- enhanced Raman spectroscopy

Xiaozhou Li
Tianyue Yang
Junxiu Lin

Spectral analysis of human saliva for detection of lung cancer using surface-enhanced Raman spectroscopy

Xiaozhou Li,^a Tianyue Yang,^a and Junxiu Lin^b

^aShenyang Ligong University, School of Science, Hunnan District, Shenyang, Liaoning, China

^bDalian University of Technology, School of Physics and Optoelectronic Engineering, Ganjingzi District, Dalian, Liaoning, China

Abstract. Surface-enhanced Raman spectroscopy (SERS) has been shown to be able to detect low-concentration biofluids. Saliva SERS readings of 21 lung cancer patients and 20 normal people were measured and differentiated. Most of the Raman peak intensities decrease for lung cancer patients compared with that of normal people. Those peaks were assigned to proteins and nucleic acids, which indicate a corresponding decrease of those substances in saliva. Principal component analysis (PCA) and linear discriminant analysis (LDA) were used to reduce and discriminate between the two groups of data, and the study resulted in accuracy, sensitivity, and specificity being 80%, 78%, and 83%, respectively. In conclusion, SERS of saliva showed the ability to predict lung cancer in our experiment. © 2012 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: 10.1117/1.JBO.17.3.037003]

Keywords: surfaced-enhanced Raman spectroscopy; saliva; lung cancer; principal component analysis-linear discriminant analysis.

Paper 11317P received Jun. 24, 2011; revised manuscript received Nov. 30, 2011; accepted for publication Jan. 9, 2012; published online Mar. 19, 2012.

1 Introduction

Lung cancer ranked first in terms of incidence and mortality, with a five-year survival rate of 15%.¹ The routine diagnosing process includes a series of examinations such as physical examination, imaging tests, and biopsies, which are time-consuming and invasive.² Raman spectroscopy provides molecular changes of tissues and biofluids in the human body, but the low sensitivity obstructs its application. Surface-enhanced Raman spectroscopy (SERS) can increase the intensity of normal Raman (NR) by an order of 4 to 15 owing to different substrates,^{3,4} and it has been widely used in the low-concentration biofluid detection that NR is hardly qualified to perform.⁵⁻⁷

Saliva has shown the ability to predict many diseases via several methods, including liquid chromatography, mass spectrometry,⁸ and enzyme-linked immunosorbent assay (ELISA).⁹ But most of them are time-consuming and require a large sample compared to SERS.¹⁰ Because of the mechanism of SERS, certain medications of nano-metal particles are usually used, such as biotinylated Ag powder¹¹ and DNA aptamers capped Ag-nanoparticle.¹² Nevertheless, some studies that apply naked SERS-active nanoparticles to the detection of body fluids have been performed. SERS of serum has exhibited differences between type II diabetes mellitus and diabetic complication groups at both the Raman shifts and the principal components.⁶ A concentration of 1 g/mL of 5-fluorouracil in saliva has been measured by the removal of thiocyanate and a pretreatment of SERS-active capillaries.¹³ A difference in saliva SERS between healthy people and oral-cancer patients at three wavenumbers has been detected using a gold nanoparticle assembly.¹⁴ In addition, SERS has been used in the differentiation of saliva for the AIDS virus by a support vector machine algorithm.¹⁵ Although it has not been studied very much, we have seen the potential of using SERS of saliva as a tool for the diagnosis of diseases.

In this paper, the saliva SERS of 21 lung cancer patients and 20 healthy people was measured. The peak variations between the healthy and lung cancer groups were investigated for the diagnosis of lung cancer. Oven-heated silver colloids were used as the SERS substrate, using capillaries as the holder of the mixture of silver colloid and liquid-state saliva to reserve the bioactivity. Then, peaks were assigned in accordance with the possible components of saliva. Finally, multivariate analysis of principal component analysis (PCA) combined with linear discriminant analysis (LDA) were used on the spectroscopy to observe the diagnostic accuracy.

2 Materials and Methods

2.1 Saliva Sampling

Whole saliva was collected from 21 lung cancer patients and 20 healthy people. In each case, about 1 mL of saliva was collected between 9:00 a.m. and 11:00 a.m. to prevent the interference of food. Saliva providers were asked to rinse their mouths three times using pure water. A nonstimulated collection method was used: providers spit the saliva into a disposable cup. Then the saliva was transferred to a 1.5-mL centrifuge tube and was centrifuged at 14,000 rpm for 10 min to remove the oral mucous epithelial cells and food debris.⁸ The whole collection process was performed at a temperature of 4 °C.

The average age of the lung cancer patients was 56, and that of the healthy volunteers was 60. Most of the patients were smokers (Table 1). For better comparison between the two groups, a control group with similar age and smoking history was also selected.

2.2 SERS Measurement

Silver hydrosol was synthesized by the deoxidizing method, using nanoparticles and silver nitrate heated in a microwave

Address all correspondence to: Xiaozhou Li, Shenyang Ligong University, School of Science, Hunnan District, Shenyang, Liaoning, China. Tel: 86-13700009301; E-mail: biohy@163.com

Table 1 Demographics of the study population.

	Lung cancer (<i>n</i> = 21)	Healthy (<i>n</i> = 20)
Age (years, s.d.)	56 (12)	60 (10)
Gender (% F)	48	60
Smoker (%)	67	65

oven, that is described in Liu's article.¹⁶ This method can produce more uniform silver nanoparticles due to the heating mode of the microwave oven. A 2- μL sample of centrifuged saliva and 2 μL of silver colloid were mixed thoroughly by an oscillator for 10 min, and then the mixture was sucked into capillaries for later spectroscopy collection. Spectroscopy was collected from an inverted microscope (inVia, Renishaw plc., England) in the

range of 500 cm^{-1} to 2000 cm^{-1} , with a 632.8-nm He-Ne laser at a power maintained at 3.5 mW. The exposure time is 100 s.

2.3 Spectral Analysis

Raw spectra were processed by vector normalization, smoothing, and baseline correction to avoid errors or artificial interference during the sample preparation and spectral acquisition. Changes of Raman peaks between the two groups were compared, and the intensity change was defined as:¹⁷

$$\Delta I = \frac{I_{\text{cancer}} - I_{\text{control}}}{I_{\text{control}}} \times 100\%.$$

Raman peaks were assigned in accordance with the specifications of Raman spectroscopy or SERS of tissues, body fluids, or bio-molecules. Peaks with wavelength difference within

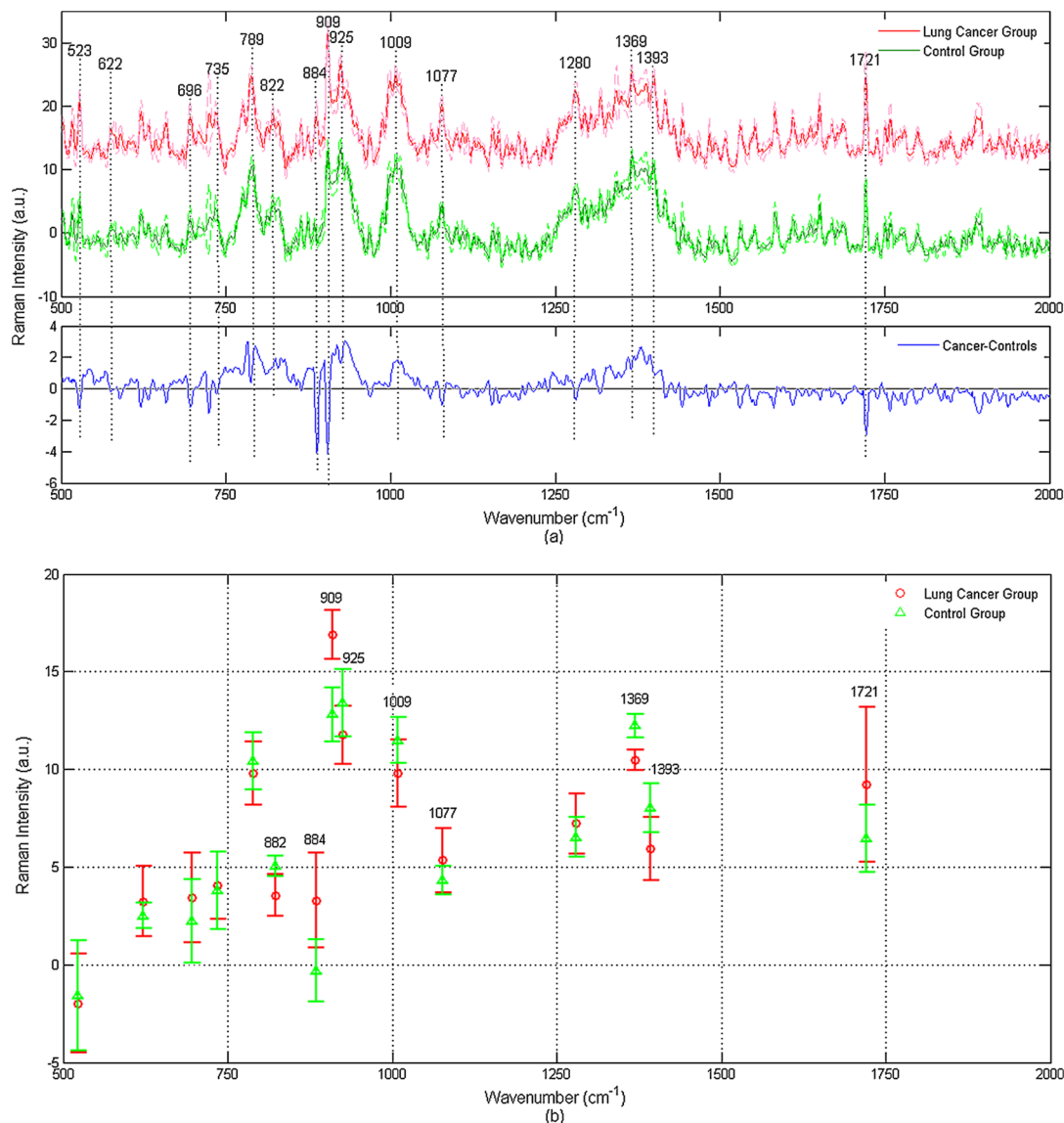


Fig. 1 (a) Comparison of the average spectra for the lung cancer patients (red line, *n* = 21) versus that of the normal group (green line, *n* = 20). The broken lines represent the standard deviations of the averages. Also shown at the bottom (blue line) is the difference spectrum. (b) Plot showing the average intensity value and the corresponding standard deviations of the selected peaks. The peaks that are marked wavenumbers are those that have *p* values lower than 0.05 of Student's *t* test.

Table 2 Assignments of SERS peaks and the changes between healthy people and lung cancer patients. (Peaks were regarded as the same Raman peaks with a wavenumber difference of less than 10 cm⁻¹.)

Raman shift (cm ⁻¹)	Average intensity		Change (%)	p value	Biomolecules
	Control	Cancer			
523	-1.59	-1.98	25	0.639	Lysozymes, ²⁰ proteins, ²⁶ G, T ²⁷
622	2.50	3.24	29	0.736	Protein, ²¹ Phe, ²⁰ A ²⁸
696	2.23	3.41	53	0.096	Met, ²⁰ C ²⁸
735	3.81	4.05	6	0.674	Trp, ²⁷ coenzyme A, A, ⁵ C, T, ²⁹ G ³⁰
789	10.42	9.78	-6	0.196	C, U, T ²⁶
822	5.06	3.55	-30	<0.001	
884	-0.32	3.30	-1, 138	<0.001	Pro, Val, Gly, ³¹ Trp, ²⁰ Glu, ⁵ Hyd ³²
909	12.81	16.89	32	<0.001	Tyr ³³
925	13.38	11.77	-12	0.002	Pro, glucose ²⁶
1,009	11.49	9.80	-14	<0.001	Try, ²⁴ Lys, ³⁴ Phe ²⁷
1,077	4.31	5.34	24	0.013	Lipids, nucleic acids, proteins, carbohydrates ²⁶
1,280	6.53	7.23	11	0.094	Phospholipid, amide III ²¹ proteins, lipids, ³¹ T ²⁸
1,369	12.24	10.50	-14	<0.001	Trp, porphyrins, lipids, G, ²⁶ T, protein ²⁷
1,393	8.03	5.93	-26	<0.001	
1,722	6.46	9.22	43	0.007	ester group ³⁵

Nucleic acid bases: A (adenine), T (thymine), C (cytosine), U (uracil), G (guanine)

Amino acids: Trp (tryptophan), Phe (phenylalanine), Lys (lysine), Pro (proline), Glu (glutamate), Val (valine), Gly (glycine), Hyd (hydroxyprolin), Met (methionine).

10 cm⁻¹ were seen as the same peak, considering different detection conditions and experimental errors.

PCA-LDA algorithm was used to discriminate the saliva SERS of the lung cancer group from the control group. PCA is very effective in dimension reduction and can transform a large amount of data into fewer new variables. LDA is a linear method used to classify two or more groups, and it is often combined with PCA. PCA-LDA has been widely used in the classification of Raman spectroscopy for disease diagnosis, and high prediction results have been obtained.^{18,19}

We used the Curve Fitting Toolbox and Statistical Toolbox in the Matlab software (The Mathworks Inc., Natick, MA) in our data processing.

3 Results and Discussion

3.1 Raman Spectroscopy

The saliva SERS ± 1 standard deviation (SD) of both the lung cancer group ($n = 21$) and the control group ($n = 21$) were compared (Fig. 1). There were 15 major peaks at the wavelengths of about 523, 622, 696, 735, 789, 822, 884, 909, 925, 1009, 1077, 1280, 1369, 1393, and 1721 cm⁻¹.

Some differences between the two groups still existed, and the difference spectrum made them apparent. The major changes were at the nine wavenumbers of 822, 884, 909, 925, 1009, 1077, 1369, 1393, and 1721 cm⁻¹. Most of the peaks decrease

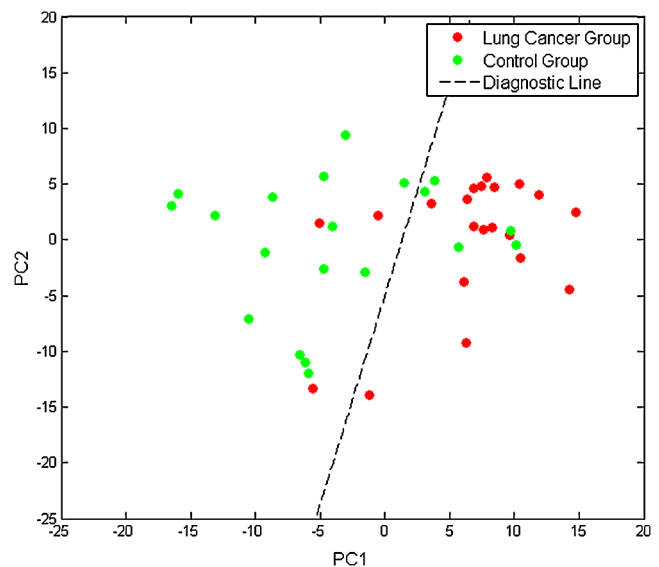


Fig. 2 The scatterplot of the first two PCs (PC1 and PC2) of PCA and the diagnostic line from LDA. The spots of the two groups distribute separately, which means that they can be discriminated well. The Fisher's discriminant function (diagnostic line) is $0.25 \times PC1 - 0.07 \times PC2 = 0.32$.

Table 3 Diagnostic results of PCA-LDA on the SERS of saliva.

Diagnosis	Cases		Total	Accuracy	Sensitivity	Specificity
	Lung Cancer	Healthy				
Lung cancer	18	3	21	80%	78%	83%
Healthy	5	15	20			

between the control group to the lung cancer group. At 822, 884, 925, 1009, 1280, 1369, and 1393 cm^{-1} (Student's *t* test, $p < 0.05$), the peak intensities were greater for the control group than for the lung cancer group, while the bands at 909, 1077, and 1,721 cm^{-1} were more intense in the saliva of the lung cancer patients (Fig. 1).

Those Raman peaks were assigned to amino acids and nucleic acid bases (Table 2), and the decreases of peak intensities represented the corresponding decrease of those biomolecules.

3.2 PCA-LDA

PCA was used on our preprocessed SERS of saliva. The contribution rate to the total variation of spectra of the first two PCs (PC1 and PC2) was 83%; hence, the first two principal components can account for 83% of the total variability. Therefore, only the first two PCs were used for further analysis. The scatterplot of the two groups was obtained using the first two PCs, the spots representing different groups distributed separately, and they can be discriminated well. LDA was used on the first two PCs to examine the diagnostic ability of our method. The Fisher's discriminant function (diagnostic line) of LDA is as follows:

$$0.25 \times \text{PC1} - 0.07 \times \text{PC2} = 0.32.$$

Most of the spots belonged to different groups distributed at the two sides of the diagnostic line (Fig. 2), which means that our SERS can be discriminated well using PCA-LDA. An accuracy of 86%, a sensitivity of 94%, and a specificity of 81% were obtained (Table 3).

4 Discussion

Although some of the peaks were assigned to certain unit structures of proteins or nucleic acids (that is, certain amino acids or nucleic acid bases), the bonds existed in most of the units. For example, COO^- existed in all 20 common amino acids, all 5 kinds of nucleic acids have the structure of $O-P-O$, and 3 kinds of amino acids in the aromatic *R* group have aromatic rings. So the assignment may be expanded to those series of molecules. Such as the peak at 622 cm^{-1} , it was assigned to both phenylalanine²⁰ and other proteins.²¹

Our results showed an overall decrease of protein and nucleic acids in the lung cancer group than in the control group. This decrease trend may have the same cause as the Raman spectroscopy of human lung epithelial tumor cells, which is induced by the breakdown of several kinds of molecular bonds.²² As blood is one of the means for disease diagnosis, and saliva is the ultra-filtration of blood, saliva can reflect the condition of tissues to a certain degree. Hence, the decrease of intensity of certain Raman bands of saliva might be a reflection of lung tissue. Further research is needed to determine if there is a distinct corresponding relationship.

SERS is not the pure enhancement of NR. There are some wavelength shifts in SERS and new peaks will appear compared to NR.²³ As the fingerprint property of Raman spectroscopy, the same bond will have different Raman shifts in different groups. For example, $C-C$ stretch vibration has a Raman shift as 1084 cm^{-1} in lauric acid, but it moves to 1092 cm^{-1} in myristic acid, and 1099 cm^{-1} in palmitic acid.²⁴ Therefore, certain errors in the peak assignments based on the literature might exist due to totally different measuring conditions.

Some high-abundance proteins, like amylase, lysozyme, and mucin²⁵ in saliva, have a very small diagnosis significance and might immerse the Raman peaks of low-abundance biomarkers. Removal of those substances can exclude the disturbance and improve the expression of those low-abundance biomarkers. Silver colloid in SERS is more uniform than that made by traditional water-heating, but the Ag particles are smaller; thus, there is a consequent decrease in the enhancement level. A highly ordered substrate can create a higher enhancement and more reproducible results. SERS effect is influenced by many factors, like laser power, temperature, solvent of samples, SERS substrate, the state of the analyte, and exposure time. The uniformity of SERS spectrometers and measuring conditions can eliminate the above fluctuations greatly and is the future orientation of our research.

5 Conclusion

Saliva has shown the potential to diagnose many diseases, and recently, SERS has been used in the detection of saliva because of the high enhancement effect. The peak changes between the lung cancer group and the control group and an 80% diagnosis accuracy in our study showed the potential of saliva SERS for diagnosing lung cancer. SERS of saliva has several advantages, such as fast, noninvasive, and low-analyte requirements, and it has prospective clinic applications. High-abundance proteins will be depleted, and better-ordered SERS substrate will be introduced in our future experiments to eliminate the disturbance of those proteins and increase the enhancement and the reproducibility of SERS measurements. Possible saliva biomarkers will be selected and detected for more exact peak assignments.

Acknowledgments

We thank The Fourth Affiliated Hospital of China Medical University for providing the saliva sample, and Dalian University of Technology for the technical support on SERS measurement.

References

1. D. M. Parkin et al., "Global Cancer Statistics, 2002," *CA Cancer J. Clin.* **55**(2), 74–108 (2005).
2. J. Langwith, "Diagnosis and Staging," Chapter 6, in *Lung Cancer*, pp. 71–77, Greenhaven Press, New York (2007).
3. X. M. Lin et al., "Surface-enhanced Raman spectroscopy: substrate-related issues," *Anal. Bioanal. Chem.* **394**(7), 1729–1745 (2009).

4. S. Nie and S. R. Emory, "Probing single molecules and single nanoparticles by surface-enhanced Raman scattering," *Science* **275**(5303), 1102–1106 (1997).
5. S. Feng et al., "Nasopharyngeal cancer detection based on blood plasma surface-enhanced Raman spectroscopy and multivariate analysis," *Biosens. Bioelectron.* **25**(11), 2414–2419 (2010).
6. H. Han et al., "Analysis of serum from type II diabetes mellitus and diabetic complication using surface-enhanced Raman spectra (SERS)," *Appl. Phys. B: Lasers Optics* **94**(4), 667–672 (2009).
7. K. Virkler and I. K. Lednev, "Forensic body fluid identification: the Raman spectroscopic signature of saliva," *Analyst* **135**(3), 512–517 (2010).
8. S. Hu et al. "Salivary proteomics for oral cancer biomarker discovery," *Clin. Cancer Res.* **14**(19), 6246–6252 (2008).
9. M. N. Brooks et al., "Salivary protein factors are elevated in breast cancer patients," *Mol. Med. Rep.* **1**(3), 375–378 (2008).
10. X. X. Han, B. Zhao, and Y. Ozaki, "Surface-enhanced Raman scattering for protein detection," *Anal. Bioanal. Chem.* **394**(7), 1719–1727 (2009).
11. K. Kim, H. K. Park, and N. H. Kim, "Silver-particle-based surface-enhanced Raman scattering spectroscopy for biomolecular sensing and recognition," *Langmuir* **22**(7), 3421–3427 (2006).
12. L. Fabris et al., "Aptatag-based multiplexed assay for protein detection by surface-enhanced Raman spectroscopy," *Small* **6**(14), 1550–1557 (2010).
13. S. Farquharson et al., "Detection of 5-fluorouracil in saliva using surface-enhanced Raman spectroscopy," *Vibrational Spectrosc.* **38**(1–2), 79–84 (2005).
14. K. W. Kho et al., "Surface-enhanced Raman spectroscopic (SERS) study of saliva in the early detection of oral cancer," *Proc. SPIE* **5702**, 84–91 (2005).
15. W. Yan et al., "Preliminary study on the quick detection of Acquired Immune Deficiency Syndrome by saliva analysis using surface-enhanced Raman spectroscopic technique," in *31st Annual Int Conf IEEE EMBS, IEEE, Minneapolis, MN*, pp. 885–887 (2009).
16. K. Liu et al., "Surface-enhanced Raman spectrum of rat serum in the novel silver colloid," *Spectroscopy Spectral Anal.* **28**(2), 339–342 (2008).
17. Z. Zhang and Y. Xu, "The molecular mechanism of the photosensitization of photofrin (YHPD)," *Sci. China Chem.* **21**(6), 595–601 (1991).
18. A. T. Harris et al., "Potential for Raman spectroscopy to provide cancer screening using a peripheral blood sample," *Head Neck Oncol.* **1**(1), 34 (2009).
19. S. K. Teh et al., "Diagnostic potential of near-infrared Raman spectroscopy in the stomach: differentiating dysplasia from normal tissue," *Brit. J. Cancer* **98**(2), 457–465 (2008).
20. M. C. Chen, R. C. Lord, and R. Mendelsohn, "Laser-excited Raman spectroscopy of biomolecules, V. conformational changes associated with the chemical denaturation of lysozyme," *J. Am. Chem. Soc.* **96**(10), 3038–3042 (1974).
21. J. L. Pichardo-Molina et al., "Raman spectroscopy and multivariate analysis of serum samples from breast cancer patients," *Lasers Med. Sci.* **22**(4), 229–236 (2007).
22. I. Notingher et al., "Spectroscopic study of human lung epithelial cells (A549) in culture: living cells versus dead cells," *Biopolymers* **72**(4), 230–240 (2003).
23. J. Guicheteau et al., "Raman and surface-enhanced Raman spectroscopy of amino acids and nucleotide bases for target bacterial vibrational mode identification," *Proc. SPIE* **6218**, 621800 (2006).
24. J. De Gelder et al., "Reference database of Raman spectra of biological molecules," *J. Raman Spec.* **38**(9), 1133–1147 (2007).
25. K. Virkler and I. K. Lednev, "Analysis of body fluids for forensic purposes: from laboratory testing to non-destructive rapid confirmatory identification at a crime scene," *Forensic Sci. Int.* **188**(1–3), 1–17 (2009).
26. N. Stone et al., "Raman spectroscopy for identification of epithelial cancers," *Faraday Discuss.* **126**, 141–157 (2004).
27. U. Neugebauer et al., "Identification and differentiation of single cells from peripheral blood by Raman spectroscopic imaging," *J. Biophoton.* **3**(8–9), 579–587 (2010).
28. C. Otto et al., "Surface-enhanced Raman spectroscopy of DNA bases," *J. Raman Spectrosc.* **17**(3), 289–298 (1986).
29. K. Virkler and I. K. Lednev, "Raman spectroscopic signature of blood and its potential application to forensic body fluid identification," *Anal. Bioanal. Chem.* **396**(1), 525–534 (2010).
30. A. Barhoumi et al., "Surface-enhanced Raman spectroscopy of DNA," *J. Am. Chem. Soc.* **130**(16), 5523–5529 (2008).
31. P. O. Andrade et al., "Study of normal colorectal tissue by FT-Raman spectroscopy," *Anal. Bioanal. Chem.* **387**(5), 1643–1648 (2007).
32. B. G. Frushour and J. L. Koenig, "Raman scattering of collagen, gelatin, and elastin," *Biopolymers* **14**(2), 379–391 (1975).
33. A. Synytsya et al., "Raman spectroscopy of tissue samples irradiated by protons," *Int. J. Radiat. Biol.* **80**(8), 581–591 (2004).
34. W. L. Peticolas, "Application of Raman spectroscopy to biological macromolecules," *Biochimie* **57**(4), 417–428 (1975).
35. C. Krafft et al., "Near infrared Raman spectra of human brain lipids," *Spectrochim. Acta Part A: Mol. Biomol. Spectroscopy* **61**(7), 1529–1535 (2005).