Noninvasive detection of nasopharyngeal carcinoma based on saliva proteins using surface-enhanced Raman spectroscopy

Xueliang Lin
Duo Lin
Xiaosong Ge
Sufang Qiu
Shangyuan Feng
Rong Chen
Noninvasive detection of nasopharyngeal carcinoma based on saliva proteins using surface-enhanced Raman spectroscopy

Xueliang Lin,a Duo Lin,a,b,x Xiaosong Ge,a Sufang Qiu,c* Shangyuan Feng,a and Rong Chen,a

Abstract. The present study evaluated the capability of saliva analysis combining membrane protein purification with surface-enhanced Raman spectroscopy (SERS) for noninvasive detection of nasopharyngeal carcinoma (NPC). A rapid and convenient protein purification method based on cellulose acetate membrane was developed. A total of 659 high-quality SERS spectra were acquired from purified proteins extracted from the saliva samples of 170 patients with pathologically confirmed NPC and 71 healthy volunteers. Spectral analysis of those saliva protein SERS spectra revealed specific changes in some biochemical compositions, which were possibly associated with NPC transformation. Furthermore, principal component analysis combined with linear discriminant analysis (PCA-LDA) was utilized to analyze and classify the saliva protein SERS spectra from NPC and healthy subjects. Diagnostic sensitivity of 70.7%, specificity of 70.3%, and diagnostic accuracy of 70.5% could be achieved by PCA-LDA for NPC identification. These results show that this assay based on saliva protein SERS analysis holds promising potential for developing a rapid, noninvasive, and convenient clinical tool for NPC screening. © 2017 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JBO.22.10.105004]

Keywords: surface-enhanced Raman spectroscopy; nasopharyngeal carcinoma detection; saliva protein; principal component analysis combined with linear discriminant analysis.

Paper 170416RR received Jun. 27, 2017; accepted for publication Sep. 21, 2017; published online Oct. 13, 2017.

1 Introduction

According to Global Cancer Statistics, cancer is one of the most common causes of death, with about 14.1 million people being diagnosed and 8.2 million people dying of cancer in 2012 worldwide.1 Nasopharyngeal carcinoma (NPC), a head and neck malignancy, is an endemic disease with high mortality rates in certain regions of southeastern Asia, including Hong Kong, Guangdong province of China, Fujian province of China, and some other South Asian countries.2–5 The five-year overall survival rate is about 90% for stage I NPC patients; however, the stage IV group has a poor five-year overall survival rate of 30.3%.6,7 Hence, early screening and effective treatment are essential for preventing the advancement of NPC and improving the five-year overall survival rate. Unfortunately, early detection remains a great challenge due to the insidious nature of NPC and the relative anatomical inaccessibility of the nasopharynx. At present, the common detection and screening methods for NPC mainly include electronic nasopharyngoscope examination, nasopharyngeal and neck magnetic resonance imaging, body positron emission tomography, bone emission CT, and biopsy, which suffer the disadvantages of being time-consuming, high-cost, complex procedures, and based on subjective judgment of clinicians with varying levels of experience.5–10 Therefore, development of an alternative diagnostic method is of significant clinical value for NPC screening.

Raman spectroscopy (RS) is a powerful optical analytical tool with many advantages over conventional optical detection techniques. For example, RS has low water interference, high signal-to-noise ratio, and high sensitivity compared with Fourier infrared spectroscopy.11 Additionally, RS has a narrow peak width and multiplexing detection ability, leading to better analytical efficiency than fluorescence spectroscopy.12,13 RS has attracted significant attentions for its material analysis and biomedical applications, especially for cancer diagnosis, due to its ability to provide structural- and biochemical-specific information of macromolecules, such as proteins, nucleic acids, lipids, and so on.14,15 However, the conventional RS technique has some deficiencies, including its inherent weak Raman scattering efficiency and strong fluorescence background. For instance, the Raman signal of a biological molecule is 1/10,000 of its fluorescence signal. To overcome these problems, surface-enhanced Raman spectroscopy (SERS) technology exploits the interaction between the biomolecules and metal nanoparticles (NPs) surface to dramatically increase the Raman signal (enhancement factors of up to $10^{13}$ to $10^{15}$), providing a more powerful and sensitive detection approach for cancer detection.16,17 Currently, SERS technology has been widely used to detect various cancers in human samples, such as DNA, RNA, cell, blood, and tissue.18 Compared with traditional polymerase chain reactions or immunoassays, biomarker detection based on SERS avoids use of expensive reagents and complex sample preparation steps. In
addition, SERS has many advantages over traditional fluorescent methods, such as better multiplex capability and less photobleaching.19,20

Human saliva, as a body fluid, contains a large number of serum components, abundant protein, and metabolites, which may change when associated with neoplastic transformation.21 Comparing the above-mentioned human samples, saliva is regarded as an ideal medium for cancer diagnostics due to its rapid, noninvasive, and convenient collection procedure.22–24 Recently, application of SERS in saliva analysis was investigated extensively to detect several types of cancer.25,26–28 These preliminary results demonstrated the potential of the saliva SERS method for cancer detection.

To further verify the efficiency of saliva SERS in NPC detection, in this work, a large number of NPC saliva samples were collected, and saliva sample analysis based on membrane protein purification was developed for saliva SERS detection. Moreover, principal component analysis combined with a linear discriminant analysis (PCA-LDA) diagnostic algorithm was employed to classify the saliva protein SERS spectra from the NPC group and the healthy group.

2 Materials and Methods

2.1 Preparation of Human Saliva Samples

In this study, the saliva samples were collected from 170 NPC patients with confirmed clinical and histopathological diagnosis of NPC lesions and 71 healthy volunteers. Detailed information on these samples is presented in Table 1. This study obtained ethical approval from Fujian Provincial Cancer Hospital (Fuzhou, Fujian, China). The collection process of the saliva samples is as follows: (1) after 12 h of overnight fasting, 1.5 ml of saliva was collected from the study subjects who washed their mouth with water three times between 7 and 8 am; (2) the collected saliva was centrifuged at 13,900 rpm for 10 min to remove oral epithelial cells and any residual food debris; and (3) the pure saliva samples were frozen at −80°C until used.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>NPC patients</th>
<th>Healthy volunteers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number</td>
<td>170</td>
<td>71</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>106</td>
<td>41</td>
</tr>
<tr>
<td>Female</td>
<td>64</td>
<td>30</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤50</td>
<td>92</td>
<td>36</td>
</tr>
<tr>
<td>&gt;50</td>
<td>78</td>
<td>35</td>
</tr>
<tr>
<td>Clinical stage of NPC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I and II</td>
<td>39</td>
<td>NA</td>
</tr>
<tr>
<td>IV and III</td>
<td>131</td>
<td>NA</td>
</tr>
</tbody>
</table>

Note: NPC, nasopharyngeal carcinoma and NA, not applicable.

2.2 Preparation of Silver Colloids

Silver (Ag) colloids were synthesized using hydroxylamine hydrochloride and Ag nitrate according to the deoxidizing method reported by Li et al.28 At first, 4.5 ml sodium hydroxide solution was mixed with 5 ml of 0.06 M hydroxylamine hydrochloride solution. After that, the mixture was added rapidly to 90 ml of 0.0011 M silver nitrate aqueous solution. The resulting mixture was stirred to obtain uniformly gray Ag colloids. Meanwhile, the maximum absorption peak of Ag colloids was 418 nm, and the Ag colloids NPs sizes were represented with a mean diameter of 35 nm by transmission electron microscopy (TEM) [Fig. 1(a)]. Finally, the final Ag colloids were obtained by centrifuging the solution at 10,000 rpm for 10 min and removing the supernatant as the SERS substrate.

2.3 Saliva Protein Purification and Surface-Enhanced Raman Spectroscopy Detection

Figure 1(b) shows the schematic diagram of saliva protein purification and SERS detection. In brief, 10 μl of the purified saliva sample was blotted onto the cellulose acetate (CA) membrane (2 × 8 cm) by a trac pipette gun two times. In total, 20 μl of each purified saliva sample was used. After the saliva sample was completely absorbed for about 5 min, the membrane was washed for 5 min in 400 ml of a special solution made up of 180 ml of 95% ethanol, 200 ml of distilled water, and 20 ml of glacial acetic acid. This step aimed to remove any other components contained in the saliva sample and leave only the saliva proteins in the CA membrane. Then, the membrane was patted dry with a filter paper for about 10 min. After that, the position of the CA membrane that only contained saliva proteins was cut into pieces and collected in a tube. Next, 150 μl of acetic acid was added into the tube to dissolve the membrane fragments into a transparent gel. Then, 150 μl of a prepared Ag colloids solution was added into the tube and the tube was placed into 37°C water for 20 min. Finally, 2 μl of liquid supernatant (protein–Ag NP mixture) was dripped onto a clean aluminum plate (Guantai Metal, Hebei, China) for SERS measurements. All SERS spectra of the saliva protein were collected with a confocal Raman microspectrometer (Renishaw plc, Gloucestershire, UK) using a Peltier cooled charge-coupled device camera under a 785-nm diode laser (a maximum power output of 5 mW) in the range of 600 to 1750 cm−1. All spectra were recorded with 50x objective, 2 cm−1 spectral resolution, and 10 s acquisition time. The software package WIRE 2.0 was used for raw spectral acquisition.

2.4 Data Processing and Multivariate Statistical Analysis

All the raw SERS spectra of saliva proteins contain Raman scattering, fluorescence background, and noise signals. To obtain better Raman spectra, a Vancouver Raman algorithm based on a fifth-order polynomial fitting method was used to remove fluorescence background and noise signals.29 Then, all background-subtracted SERS spectra were normalized to the integrated area under the curve in the range of 600 to 1750 cm−1 for a better comparison of SERS spectral shape in the analysis. This sophisticated and robust diagnostic model based on PCA-LDA was used for Raman spectral analysis in saliva detection. In this study, PCA was used to reduce complex data sets and select the principal components (PCs) that account for the maximal...
variances in the multidimensional data sets. To further analyze saliva protein SERS data, three diagnostically significant PCs ($p < 0.05$) were selected for one independent sample $t$-test. LDA was used to generate effective diagnostics using three diagnostically significant PCs with leave-one-out cross-validation methods. Receiver operating characteristic (ROC) curves were generated by successively changing the discrimination threshold levels to additionally evaluate the classifications of the multivariate statistical method for NPC diagnosis.

3 Results and Discussions

Figures 2(a) and 2(b) show the SERS spectra of untreated saliva samples of 10 patients with NPC and saliva protein samples after membrane protein purification from the same samples, respectively. Great variations in spectral features, including spectral intensity, position, and width, could be found in the untreated saliva SERS within the cancer group, making it difficult to achieve efficient diagnostics using SERS spectral analysis. This is explainable. A saliva sample contains a variety of native proteins.
constituents (proteins, peptides, polynucleotides, and electrolytes) and exogenous substances (nonadherent oral bacteria, food remains, traces of medications, or chemical products).\textsuperscript{31,32} Although preprocessing such as washing the mouth for saliva collection was implemented, there remained some exogenous substances in the saliva samples that generated prominent SERS signals. Thus, great SERS spectral variation from untreated saliva samples was probably attributed to the different exogenous substances among subjects. Interestingly, the reproducibility of SERS spectra [Fig. 2(b)] was dramatically improved by the saliva protein purification method developed. This method not only circumvented the limitations of raw saliva SERS detection but also provided a unique opportunity to use SERS to explore the changes in saliva proteins associated with cancer transformation.

Figure 3(a) shows the mean normalized SERS spectrum of purified whole proteins obtained from normal (blue line, \( n = 71 \)) and NPC (red line, \( n = 170 \)) saliva samples. The prominent SERS peaks located at around 621, 642, 760, 854, 878, 935, 959, 1004, 1031, 1049, 1123, 1175, 1208, 1265, 1337, 1445, 1552, and 1684 cm\(^{-1}\) were clearly observed in both normal and NPC saliva protein samples. These prominent SERS peaks could be tentatively assigned to explain the changes in biological constituents in saliva protein samples as shown in Table 2.\textsuperscript{33–36} The strongest peaks at 760, 1004, 1265, 1445, and 1684 cm\(^{-1}\) existed in the measured saliva protein SERS spectra. Compared with the SERS spectra of NPC saliva proteins, the SERS spectra of normal saliva proteins had higher intensities at 621, 935, and 1049 cm\(^{-1}\) and lower intensities at 1004, 1031, 1208, 1123, and 1684 cm\(^{-1}\). The comparisons of saliva proteins SERS spectral intensities between the NPC and normal groups could be viewed more clearly in Figs. 3(b) and 3(c). These differences in spectral intensities demonstrated that saliva proteins SERS has a potential role for NPC detection. For instance, the SERS bands of phenylalanine (1004 and 1031 cm\(^{-1}\)) were all related to the molecular stretching and bending mode of proteins, and they showed a lower SERS signal in the normal saliva protein samples than in the NPC saliva protein samples, indicating that an increase in the percentage of these proteins content in the total SERS-active in the NPC patients. Our group also observed that the content of phenylalanine was increased when associated with malignant transformation in cervical and NPC blood plasma by the SERS technique.\textsuperscript{2,18} However, the intensity of the saliva protein SERS spectrum exhibited decreased peaks at 935 cm\(^{-1}\) (praline and valine) and 1049 cm\(^{-1}\) (proteins) in NPC subjects, indicating that cancer patient saliva may be associated with a decreased concentration of these proteins. These results revealed that specific-protein changes between the NPC and normal groups could

<table>
<thead>
<tr>
<th>Peak position (cm(^{-1}))</th>
<th>Vibrational mode</th>
<th>Major assignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>621</td>
<td>C–C twisting mode</td>
<td>Phenylalanine</td>
</tr>
<tr>
<td>642</td>
<td>ν (C–S)</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>760</td>
<td>Ring breathing mode</td>
<td>Typtophan</td>
</tr>
<tr>
<td>854</td>
<td>Ring breathing mode</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>878</td>
<td>ν (C–C)</td>
<td>Hydro-oxyproline</td>
</tr>
<tr>
<td>935</td>
<td>ν (C–C)</td>
<td>Proline</td>
</tr>
<tr>
<td>959</td>
<td>ν (C–C)</td>
<td>α-Helix, proline, valine</td>
</tr>
<tr>
<td>1004</td>
<td>ν(_s) (C–C)</td>
<td>Phenylalanine</td>
</tr>
<tr>
<td>1031</td>
<td>δ (C–H)</td>
<td>Phenylalanine</td>
</tr>
<tr>
<td>1049</td>
<td>ν (C–O), ν (C–N)</td>
<td>Proteins</td>
</tr>
<tr>
<td>1123</td>
<td>ν (C–N)</td>
<td>Proteins</td>
</tr>
<tr>
<td>1175</td>
<td>δ (C–H)</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>1208</td>
<td>ν (C–C(_6)H(_5))</td>
<td>Typtophan</td>
</tr>
<tr>
<td>1265</td>
<td>ν (CN), δ (NH)</td>
<td>Amide III, collagen</td>
</tr>
<tr>
<td>1337</td>
<td>CH(_2)CH(_2) wagging</td>
<td>Collagen</td>
</tr>
<tr>
<td>1445</td>
<td>δ (C–H)</td>
<td>Collagen, lipids</td>
</tr>
<tr>
<td>1552</td>
<td>ν (C=O)</td>
<td>Typtophan</td>
</tr>
<tr>
<td>1684</td>
<td>ν (C–C)</td>
<td>Amide I</td>
</tr>
</tbody>
</table>

Note: \( \nu \), stretching vibration; \( \delta \), bending vibration; and \( \nu_s \), symmetric stretch.
be detected by SERS, suggesting promising potential of saliva protein SERS for NPC screening.

It should be noted that the simplistic peak intensities analysis above only uses limited information of SERS peaks, and more valuable diagnostic information contained in the SERS spectra has not been employed for spectral classification. Therefore, the PCA-LDA diagnostic algorithm was used to improve the diagnostic efficiency of the SERS technique by analyzing and classifying the saliva protein SERS spectra from NPC and normal subjects.

This sophisticated and robust diagnostic model based on PCA-LDA has been widely used for Raman spectral analysis in tissue, cell, saliva, and blood detection. In this study, PCA was used first to reduce intensity variables within the raw Raman spectrum of saliva protein into a few PCs. Then, three PCs (PC5, PC6, and PC7) were selected to be the most diagnostically significant \( (p < 0.0001) \) as defined for discriminating normal and NPC groups by an independent sample \( t \)-test.

To further analyze saliva protein SERS data, all three diagnostically significant PCs were loaded into the LDA with the leave-one-out cross-validation method for generating an effective diagnostic model for saliva protein sample classification. Figure 4(a) shows the posterior probability plot of the healthy group (circles) and NPC group (triangles) based on the results of the PCA-LDA diagnostic model. As we can see, there are some plots overlapping in this figure. A discrimination threshold of 0.54 was employed for accurate discriminate between the healthy volunteers and NPC patients, yielding the diagnostic sensitivity, specificity, and accuracy of 70.7% (304/430), 70.3% (161/229), and 70.5% (465/659), respectively. To further evaluate the performance of saliva protein SERS for NPC diagnosis, an ROC curve was generated from the PCA-LDA data as shown in Fig. 4(b). The integration area under the ROC curve was 0.795. These results further indicated that saliva protein SERS technology combining the PCA-LDA diagnostic algorithm has the potential to be a rapid and noninvasive diagnosis method for NPC detection.

4 Conclusions

A rapid and convenient saliva SERS analysis method was developed for NPC detection. Using the membrane protein purification method, high-quality and reproducible SERS spectra of saliva proteins were obtained, making it possible to reveal specific changes in proteins associated with cancer transformation. Furthermore, a diagnostic sensitivity of 70.7% and specificity of 70.3% could be achieved by the PCA-LDA diagnostic algorithm for NPC identification, demonstrating promising potential of this saliva SERS analysis method to be a noninvasive body fluid test for clinical NPC screening.

Disclosures

The authors declare that there are no conflicts of interest related to this article.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Nos. U1605253, 61210016, 61405036, and 61575043), the Innovation Team Development Plan by the Ministry of Education of China (No. IRT15R10), the National Natural Science Foundation of Fujian, China (Nos. 2017J01499 and 2015J01436), the Key Clinical Specialty Discipline Construction Program of Fujian (China), and the National Clinical Key Specialty Construction Program.

References

Noninvasive detection of nasopharyngeal carcinoma based on saliva proteins...


Xueliang Lin received his BS degree in physics from Longyan University in 2015. He is currently working at the Institute of Laser and Optoelectronics Technology and working toward his PhD in optics at Fujian Normal University, China. His research focuses on the application of surface-enhanced Raman spectroscopy (SERS) in the diagnosis of nasopharyngeal carcinoma (NPC).

Duo Lin received his master's degree in physical electronics from Fujian Normal University in 2012. He is currently working at the Institute of Laser and Optoelectronics Technology and working toward his PhD in optics at Fujian Normal University, China. His research focuses on the application of SERS in biomedical diagnosis.

Xiaosong Ge received his BS degree communication engineering from Changzhou University in 2015. He is currently working at the Institute of Laser and Optoelectronics Technology and working toward his master's degree in physical electronics at Fujian Normal University, China. His research focuses on the application of SERS in biomedical diagnosis.

Sufang Qiu received her PhD from Fujian Medical University in 2016. Currently, she is a chief physician of Fujian Tumor Hospital, China. She is an oncologist and a master tutor. Her research interest focuses on radiation therapy and diagnosis of NPC.

Shangyuan Feng received his PhD from Fujian Normal University in 2011. He started a postdoctoral position at BC Cancer Agency, Canada, in 2013. Currently, he is an associate professor at the School of Optoelectronics and Information Engineering, Fujian Normal University, China. His research interest focuses on the application of SERS in biomedical diagnosis.

Rong Chen was the director of Key Laboratory of Optoelectronic Science and Technology for Medicine, Ministry of Education. Currently, he is a doctoral tutor at the School of Optoelectronics and Information Engineering, Fujian Normal University, China. His research interest is the application of Raman spectral to detect NPC.