Abstract. The present study aimed to evaluate a label-free tissue test for the detection of nasopharyngeal carcinoma (NPC) at early and advanced stages using Raman spectroscopy (RS). RS measurements were performed to acquire high quality Raman spectra on two groups of tissue samples: One group consists of 30 NPC patients at the early stages (I-II), and the other group is 46 NPC patients at the advanced stages (III-IV). Tentative assignment of Raman bands showed specific biomolecular changes associated with cancer development. Furthermore, effective diagnostic algorithms based on principal components analysis (PCA) and linear discriminant analysis (LDA) were applied for distinguishing Raman spectra of nasopharyngeal tissues from different stages, yielding a diagnostic sensitivity of 70% and a specificity of 78%. This exploratory work suggests that RS in conjunction with the PCA-LDA algorithms provides good diagnostic ability for the early and the advanced staged NPC tissues, and RS has enormous potential for the non-invasive detection of early and advanced stage NPC.

Introduction

Nasopharyngeal carcinoma (NPC) is a malignant neoplasm and retains a high incidence in Southeast Asia and Southern China (1). An accurate tumor-node-metastasis (TNM) staging system is important for estimating the prognosis of malignant carcinoma, constructing treatment plans, and for predicting the 5-year overall survival rate (OS) of NPC (2). The 5-year OS rates of stage I, II, III, and IV NPC are 94, 87, 77 and 65%, respectively (3). It can be difficult to make an early diagnosis of this lesion due to the anatomical position of NPC. At present, the protocols used for NPC staging include flexible fiberoptic nasopharyngoscopy pathological biopsy, blood EB-DNA test, computed tomography (CT) of the chest, magnetic resonance imaging (MRI) of the nasopharynx and neck, bone emission computed tomography (ECT) and position emission tomography. These diagnostic processes are invasive, inconvenient, and time-consuming. Therefore, an improved diagnostic method for NPC is required, that can overcome the above disadvantages.

Raman spectroscopy (RS) makes use of the inelastic scattering of light, and provides a spectrographic signature of the structure and conformation of specific molecular species, such as proteins, nucleic acids and lipids (4). At present, the application of RS is used extensively to identify cancer by analyzing excised tissue, blood plasma, saliva, urine, and seminal fluid (5-9). Tissue samples obtained timely and even continually during treatment of previously diagnosed patients, are the ideal material for confirming a precise diagnosis (10). Hitherto, RS has been widely utilized to verify diagnoses using excised samples of cancer tissues, such as esophageal cancer, lung cancer, and gastric cancer (11-13). Furthermore, Raman spectroscopy endoscopic analysis of tissues in vivo has demonstrated the potential for detecting cancers such as laryngeal cancer, esophageal cancer, lung cancer, colorectal cancer and bladder cancer (14-18). Thus, RS endoscopy may constitute an ideal technology for the diagnosis of a number of cancer types in the near future.
We have previously studied the applications of RS in nasopharyngeal cancer tissues using excised nasopharyngeal tissue samples. These have been limited to in vitro experiments, and were conducted to establish preliminary data prior to performing integrated RS combined with white-light reflectance endoscope imaging (19,20). To the best of our knowledge, no other previous studies have examined RS using NPC samples. We were able to previously differentiate noncancerous and cancerous nasopharyngeal tissues by using micro RS, which yielded a diagnostic sensitivity of 92% and a specificity of 82%. These were performed using the principal components analysis and linear discriminant analysis (PCA-LDA) diagnostic algorithms (21). This encouraging preliminary result motivated us to conduct a more systematic study, by classifying the differentiation stages of the samples based on the TNM staging system of early (I-II) and advanced (III-IV) nasopharyngeal tissues.

The aim of the present study is to discriminate the spectral differences between tissues of early stage and advanced stage NPC, and to assess the non-invasive diagnosis of NPC at different stages using RS method. The current study employs multivariate statistical methods such as (PCA-LDA), to generate multi-class diagnostic algorithms for classifying Raman spectra of different stage nasopharyngeal tissue types.

Materials and methods

Subjects. Two groups of tissue samples were examined in this study: One group of 30 NPC patients at early stages (I-II), and a second group of 46 NPC patients with advanced stages (III-IV). The patients had been previously treated at the Fujian Provincial Cancer Hospital, were of similar ethnic and socioeconomic backgrounds and had been diagnosed with NPC by histopathology. The tissue samples were obtained from each patient by nasopharyngeal endoscopy and cut into pathological sections (3-10 µm thick), which were stained using hematoxylin and eosin (HE), and confirmed by pathologists. If the pathological sections were unclear using HE, immunohistochemical staining for cytokeratin 5/6 and Epstein-Barr virus-encoded RNA was used for additional verification. Each patient was diagnosed by at least two pathologists. The staging classification used was the Chinese 2008 staging system for NPC (21). Ethical approval and informed consent from all patients were obtained for the present study. The tissue samples were stored in a -80˚C freezer and were defrosted at room temperature prior to experimental RS measurements being taken.

RS measurements. Tissue samples were fixed onto a sheer aluminum plate, and all RS measurements were recorded from 800-1750 cm⁻¹, with an integration time of 30s, using an inVia micro-Raman system (Renishaw PLC, Wotton-under-Edge, UK) with a 50x objective. The tissue samples were excited using a 785 nm diode laser with a maximum power output of 30 mW.

Data Processing. All raw RS data were processed to fit the broad autofluorescence background using a fifth-order polynomial fitting method, based on that described in our previous study (22). This polynomial, essentially representing the autofluorescence, was subtracted from the raw spectrum to acquire the Raman spectrum of the tissue sample. Each RS was then normalized to the integrated area under the curve for making comparisons of spectral shapes and correcting variations in spectral intensity. Additionally, multivariate statistical analysis was performed on the pre-processed Raman data using the SPSS software package (version 15.0; SPSS Inc., Chicago, IL, USA).

Multivariate statistical analysis. Efficient diagnostic algorithms for differentiating RS spectra between the early and advanced stages of NPC patients were applied using PCA and LDA, as described in our previous study (23). Briefly, to decrease the dimensions of the spectral data, PCA is used to extract a set of orthogonal principal components (PCs) that interpret the maximum difference in the dataset, for further diagnosis and characterization. The most diagnostically valuable PCs (P<0.05) were determined by the independent-sample t-test, which were then input for the development of LDA for tissue discrimination. The performance of the diagnostic algorithms, rendered using the PCA–LDA models for precisely forecasting the tissue groups (e.g., normal vs. cancer) was evaluated in an unbiased manner using the leave-one subject-out, cross-validation method on all model spectra. In this way, one sample (i.e., one spectrum) was excluded from the dataset, and the algorithm based on PCA-LDA was redeveloped using the remaining tissue spectra. To compare the performance of the PCA-LDA model for the different NPC stage classification using the Raman spectroscopic data set, receiver operating characteristic (ROC) curves were generated by successively varying the thresholds to determine discrimination sensitivity and specificity for all samples.

Results

Raman spectra of early and advanced stage NPC. The average RS, obtained from 76 NPC patients (early stage, n=30 and advanced stage, n=46), are presented in Fig. 1. In both groups of tissue samples, characteristic bands are observed at 853, 937, 1004, 1032, 1094, 1209, 1286, 1304, 1340, 1450, 1527 and 1660 cm⁻¹.
Spectral differences between early and advanced stage NPC. The spectral differences between the early and the advanced tissue samples were examined (Fig. 2), and the results indicate that RS possesses diagnostic potential for identifying different stages of differentiation in NPC. The significantly different bands (P<0.05) are shown at 853, 937, 1004, 1032, 1094, 1209, 1268, 1304, 1340, 1450, 1527 and 1660 cm\(^{-1}\). In particular, the band intensities located at 1032 and 1209 cm\(^{-1}\) were significantly higher in the advanced tissues than those in the early tissues (P<0.05), while the band intensities located at 853, 937, 1004, 1094, 1209, 1304, 1340, 1450, 1527 and 1660 cm\(^{-1}\) were significantly lower (P<0.05) in the advanced tissues. These intensity differences suggest that in different stage NPC tissue, there are significant increases and decreases of specific biomolecules that leave an RS signature, and are correlated with the total Raman-active components.

Multivariate analysis. To identify the most significant Raman features for tissue analysis and identification, the multivariate statistical technique PCA-LDA accompanied with independent-sample t-test was carried out, and revealed that only two PCs (PC3 and PC6, P<0.05) were diagnostically significant (Fig. 3).

PCA-LDA analysis. To classify different stages of NPC tissue, the two diagnostically significant PCs were fed into the LDA model, associated with the leave-one subject-out, cross-validation technique. PCA-LDA algorithms based on the tissue Raman data calculated a diagnostic sensitivity of 70% (21/30) and specificity of 78% (36/46) for discriminating between the early and the advanced NPC tissues (Fig. 4).

ROC curve analysis. For further evaluation of the efficacy of the PCA-LDA diagnostic algorithm for different stage NPC tissue diagnosis, receiver operating characteristic (ROC) curves (Fig. 5) were constructed from the scatter plots in Fig. 4. The area under the ROC curve was 0.81, indicating that use of the PCA-LDA diagnostic algorithm based on Raman spectroscopy is viable for clinical diagnosis at the molecular level in the different stages of NPC. The predictive accuracy 75% (57/76) [sensitivity of 70% (21/30) and specificity of 78% (36/46)] could be determined for the independent validation dataset, proving the efficacy of the PCA-LDA algorithms based on Raman spectroscopy diagnosis of early and advanced stage in NPC patients.

Discussion

Raman spectroscopy is a non-invasive technique that can be used to detect NPC based on the quantitative information it provides about the specific biochemical differences between early and advanced staged tissues in terms of their unique proteins, DNA, and lipids. In the present study a diagnostic sensitivity of 70% (21/30) and specificity of 78% (36/46) was obtained for different stage cancer identification on the basis of the PCA-LDA algorithms, while a predictive accuracy of 75% (57/76) was demonstrated, indicating that RS combined with the PCA-LDA algorithms is a potential technique for distinguishing the early and the advanced stage nasopharyngeal carcinoma tissues.
RS IN THE DIFFERENTIATION OF EARLY AND ADVANCED STAGE NPC TISSUES

Table I. Peak positions and tentative assignments of the major Raman bands of early and advanced stage NPC tissue samples.

<table>
<thead>
<tr>
<th>Peak position, cm⁻¹</th>
<th>Vibrational mode</th>
<th>Major assignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>853</td>
<td>v (C-C)</td>
<td>Pro</td>
</tr>
<tr>
<td>937</td>
<td>v (C-C) in α-helix conformation</td>
<td>Pro and Val</td>
</tr>
<tr>
<td>1004</td>
<td>vs (C-C) breathing</td>
<td>Phe</td>
</tr>
<tr>
<td>1094</td>
<td>v (C-N)</td>
<td>Lipids, DNA</td>
</tr>
<tr>
<td>1209</td>
<td>v (C-C6H5)</td>
<td>Tryp and Phe</td>
</tr>
<tr>
<td>1268</td>
<td>δ (=C-H), δ (N-H)</td>
<td>PLPs, amide III</td>
</tr>
<tr>
<td>1304</td>
<td>δ (CH2)</td>
<td>Proteins and lipids</td>
</tr>
<tr>
<td>1340</td>
<td>(CH3CH2) wagging</td>
<td>Proteins and nucleic acids</td>
</tr>
<tr>
<td>1450</td>
<td>δ (CH2)</td>
<td>PLPs and collagen</td>
</tr>
<tr>
<td>1660</td>
<td>v (C≡C)</td>
<td>Lipids</td>
</tr>
</tbody>
</table>

NPC, nasopharyngeal carcinoma; v, stretching mode; vs, symmetric stretching; δ, bending mode; Pro, proline; Val, valine; Phe, phenylalanine; Tryp, tryptophan; PLPs, phospholipids.

RS has been used recently for cancer detection, based on the quantitative information provided about the specific biochemical differences between the noncancerous and cancerous blood plasma or tissue samples, particularly in terms of unique proteins, DNA, and lipids (10,20,22). Typical and tentative assignments for the observed Raman bands may provide a better understanding of the molecular basis of the diagnosis, as shown in Table I. Characteristic spectral features and relative intensity differences, which can be considered as molecular and cellular changes accompanied by malignant transformation in different stages of NPC were observed between the early and advanced NPC tissue groups. For example, the Raman peak intensity at 1004 cm⁻¹, due to the vs(C-C) breathing of phenylalanine exhibited lower signal levels for the advanced stage tissues when compared to the spectrum of the early stage tissues. A decrease in RS signal at 1004 cm⁻¹ was also observed in our previous study, when comparing cancerous and noncancerous samples, indicating a decrease in the percentage of phenylalanine relative to the total RS-active constituents in cancer tissues (10). This indicates that changes in protein-related RS peaks at 1004 cm⁻¹ that are generally associated with differential stage development likely reflect special changes of protein content and conformation. Furthermore, the Raman peak at 1004 cm⁻¹, which can reflect the changes of phenylalanine in tissue, cell, blood and other tissue samples, has been demonstrated previously to be a prominent and stable RS signal (9,10,23). Recently, we have conducted SERS studies of blood plasma samples from healthy volunteers and NPC patients with different T stages. Lin et al observed a decrease of phenylalanine in T2-4 stage cancer blood samples in NPC compared to that of T1 stage by SERS (20). Therefore, the peak at 1004 cm⁻¹ has the potential to serve as a biomarker for the differentiation between the early and advanced stage NPC.

The RS bands at 1209 cm⁻¹ are characteristic of the v(C=C6H5) of tryptophan and phenylalanine, and their percentage signals are markedly reduced in early staged tissues, indicating an increase in the percentage of tryptophan and phenylalanine contents relative to the total Raman-active components in advanced stage tissue. Our previous study demonstrated that the band intensity located at 1209 cm⁻¹ for healthy tissues was significantly lower than for cancerous tissues (10). An additional study has also observed an increase in phenylalanine in gastric cancer blood plasma using SERS (24). Thus, we hypothesize the carcinoma is undergoing a transformation process with a gradual increase in tryptophan and phenylalanine content. However, further studies are required in order to prove this.

Raman spectroscopy is a convenient, quick, and repeatable technique. The differences in RS spectra between early and advanced stage NPC tissue samples may validate biomolecular changes correlated with the development of the cancer. RS demonstrates a good ability to distinguish between the early and advanced staged tissues in NPC for the first time. Furthermore, the feasibility of RS in differentially staged NPC tissue, classified by the TNM staging system, indicates a meaningful tool in the early diagnosis NPC. It may also improve the efficacy of treatment and the prognosis of NPC.
There are several limitations of the present study that should be addressed. First, our stratified study had a small sample, and therefore, further trials with larger sample sizes should be conducted to test to improve the reproducibility and efficacy of this method so that it may translate to clinical use. Secondly, as the spectroscopic instrument is of research grade, while this is useful for the development of methodology and a pilot application for NPC staging detection, it presents a challenge for future clinical applications. In the future, RS analysis may be complementary to pathological biopsy for the detection of NPC at an early stage and non-invasively. SERS is more sensitive than normal RS, and may therefore be more appropriate for the acquisition of cancer-related results. We are currently exploring the practical feasibility of using the SERS method for NPC stage determination in a preliminary study. The ability of RS has been demonstrated in the utilization of Raman imaging (RI) as a routine optical biopsy tool, to accurately characterize cancer tissues and distinguish these from healthy tissues (25). Thus, the integrated endoscopy system based on RS for NPC detection may be developed for the diagnosis of NPC in the near future.

In conclusion, Raman spectroscopy (RS) combined with the PCA-LDA algorithms is able to distinguish early and the advanced stage NPC tissues in a non-invasive and rapid manner.

Acknowledgements

The authors of the present study thank Professor Su Ying, Department of Radiation Biological Laboratory, Fujian Provincial Cancer Hospital for the NPC tissues restore and her scientific advice. The current study was funded by the National Clinical Key Specialty Construction Program (grant no. 61210016), the Key Clinical Specialty Discipline Construction Program of Fujian (grant no. 61178090), People’s Republic of China and Fujian Provincial Key Laboratory of Translational Cancer Medicine (grant no. 61178083), the National Natural Science Foundation of China (grant no. 61405036), the Science and Technology Project of Fujian Province (grant no. WKJ-FJ-01), the Fujian Province Health Commission Young and Middle-aged Talent Training Project (grant no. 2014-ZQN-JC-6), the Program for Chang Jiang Scholars and Innovative Research Team in University (grant no. IRT1115) and the Provincial Natural Science Foundation (grant nos. 2015J01436 and 2015J0105).

References