Raman spectroscopy, a potential tool in diagnosis and prognosis of castration-resistant prostate cancer

Lei Wang
Dalin He
Jin Zeng
Zhenfeng Guan
Qiang Dang
Xinyang Wang
Jun Wang
Liqing Huang
Peilong Cao
Guanjun Zhang
JerTong Hsieh
Jinhai Fan
Raman spectroscopy, a potential tool in diagnosis and prognosis of castration-resistant prostate cancer

Lei Wang, a Dalin He, a Jin Zeng, a Zhenfeng Guan, a Qiang Dang, a Xinyang Wang, a Jun Wang, b Liqing Huang, b Peilong Cao, c Guanjun Zhang, c Jer Tong Hsieh, d and Jinhai Fan a

Xi’an Jiaotong University, First Affiliated Hospital of Medical College, Department of Urology, Xi’an, Shaanxi, China
Xi’an Jiaotong University, School of Science, Department of Optical Information Science and Technology, Xi’an, Shaanxi, China
Xi’an Jiaotong University, First Affiliated Hospital of Medical College, Department of Pathology, Xi’an, Shaanxi, China
University of Texas Southwestern Medical Center, Department of Urology, Dallas, Texas

Abstract. Purpose: We evaluated the feasibility of Raman spectroscopy (RS) in diagnosis and prognosis of castration-resistant prostate cancer (CRPC) in patients with prostate cancer (PC). Materials and methods: Raman spectra are detected from PC cell lines (LNCaP and C4-2) and tissues using a Labram HR 800 RS. Then, principal component analysis (PCA) and support vector machine (SVM) are applied for prediction. A leave-one-out cross-validation is used to train and test the SVM. Results: There are 50 qualified patients, including 33 with androgen-dependent prostate cancer (ADPC) and 17 with CRPC. The spectral changes at 1126, 1170, 1315 to 1338, and 1447 cm⁻¹ between CRPC and ADPC are detected in both cells and tissues models, which are assigned to specific amino acids and DNA. PCA/SVM algorithm provided a sensitivity of 88.2% and a specificity of 87.9% for diagnosing CRPC tissues. Furthermore, 14 patients with ADPC progressed to CRPC within 12 months. These patients are separated into two groups depending on whether their cancers progressed to CRPC within 12 months. PCA/SVM could differentiate these two groups with a sensitivity of 85.7% and a specificity of 88.9%. Conclusions: RS has the potential in diagnosis and prognosis of CRPC in clinical practice.

Keywords: Raman spectroscopy; prediction; androgen-dependent prostate cancer; castration-resistant prostate cancer.

1 Introduction
Prostate cancer (PC) is the most frequently diagnosed cancer in men. Androgen deprivation therapy (ADT) remains the mainstay of management of advanced PC. However, patients who receive ADT progress to castration-resistant prostate cancer (CRPC) in a median interval of 12 to 24 months. It had been proven that the median survival of patients with CRPC is only 15 to 19 months. Thus, disease progression and death invariably and rapidly follow CRPC.

Currently, there is still a lack of reliable methods for early identification and prediction of CRPC from clinical specimens. Although serum prostate specific antigen (PSA) has been considered the main indicator to assess the treatment response of the original publication, including its DOI. DOI: 10.1117/1.JBO.18.8.087001

Keywords: Raman spectroscopy; prediction; androgen-dependent prostate cancer; castration-resistant prostate cancer.

1.1 Cell Culture and Preparation
Human PC LNCaP and C4-2 cells were previously reported and cultured in RPMI-1640 (Gibco) medium with 10% fetal bovine serum (Hyclone) at 37°C with 5% CO2 in humidified incubators. Cell samples were prepared as previously described. Cells were seeded in a culture dish at a density of 2 × 10⁴ cells cm⁻² and grown to 80% confluence in growth medium as described above. The growth medium was then aspirated and remaining cells were washed with 4°
C phosphate-buffered saline (PBS) three times. The cells were dispersed from their culture dish using 0.25% Trypsin–ethylenediamine tetra-acetic acid (Sigma), transferred to a 15-mL universal tube and centrifuged at 92 g for 5 min at room temperature. After centrifugation again, the sediment was resuspended in 300 μL PBS and placed on glass slides with cytopsin. Cells were then fixed in 4% paraformaldehyde for 10 min at room temperature, rinsed, and immersed in PBS. Spectra would be measured from the cells within 6 h of removal from the culture medium.

2.2 Study Patients and Tumor Collection

Records were retrospectively reviewed of all patients with PC at the Department of Urology from February 2000 to June 2011, and the prostate specimens were retrieved from the Tissue Repository Facility of Department of Pathology. The study protocol has been approved by the Institutional Review Board and Research Ethical Board, Xi’an Jiaotong University. Patients with suspicion of PC gave written informed consent, and then received prostate needle biopsy guided by ultrasound and pathologic diagnosis. For treatment of advanced PC, combined androgen blockade (CAB, i.e., a combination of bicalutamide or flutamide with either bilateral orchiectomy or goserelin) was supplied to the 35 patients and single bilateral orchiectomy was offered to nine patients. Otherwise, six patients with localized PC but nodal invasion received radical retropubic prostatectomy combined with CAB. Subsequently, serum PSA was evaluated every 3 months during the first year and 6 months thereafter. The patients were selected in ADPC group if they had effective response to ADT which was recognized as decline of PSA level.

Progression to CRPC was defined as three consecutive rises in PSA, resulting in two 50% increases over the nadir, with castrate serum levels of testosterone (testosterone <50 ng/dL). Tissues of CRPC were available from 17 patients after transurethral resection of the prostate to release urinary obstruction and prostate needle biopsy guided by ultrasound. Since, the pathological stage was not available for all patients, time to CRPC progression was defined as the interval from initial of ADT to the first PSA rise.

2.3 Histopathology and Tissue Processing

Each tissue was formalin fixed and paraffin embedded and consecutively cut into 4- and 10-μm parallel sections. All sections were placed on glass slides, warmed to just above the melting point (60°C) of paraffin. The 4-μm-thick sections were dewaxed with xylene and ethanol baths, then stained by hematoxylin & eosin (H&E) for subsequent pathological examination by two experienced pathologists under Olympus BX51 upright microscope. The 10-μm-thick sections, after being dewaxed in three washes of fresh xylene and ethanol, were subjected to RS analysis. Raman spectrum was acquired using corresponding regions marked as cancerous epithelial cells on the parallel H&E-stained tissue section. Pathologic staging was performed according to the 2009 TNM classification, and grading was done according to modified Gleason score.

2.4 Instrumentation and Data Preprocessing

Raman spectra data were recorded on a Labram HR 800 (Horiba Jobin Yvon). The laser beam was set at 17 mW with a 632.8 nm He-Ne laser radiation, and accurately focused on a 1-μm² spot.
on the surface of the sample. The cells and tissues were centered and photographed using 10× objective magnifications, and measured using 100× objectives. The acquisition period was 20 s, with a 1 cm⁻¹ spectral resolution over a 700- to 1800-cm⁻¹ Raman shift range. Fifteen to twenty spectra was acquired from every tissue. The raw spectra were preprocessed by a first-order Savitsky–Golay filter for noise smoothing, and then a fifth-order polynomial was found to be optimal for fitting the autofluorescence background in the tissue-smoothed spectrum. This polynomial was then subtracted from the noise-smoothed spectrum to yield the tissue Raman spectrum.

2.5 Statistical Analysis

Principal component analysis (PCA) combined with support vector machine (SVM) in MATLAB 7.1 (The Mathworks, Inc., Natick, Massachusetts) was used to develop a diagnostic algorithm for diagnosing and predicting CRPC within the dataset. PCA is a data compression procedure that finds major trends within the spectral data set and redefines the data set using a small set of component spectra or principal components and scores. Then, SVM with a radial basis function kernel and a particle swarm optimization solver were applied to discriminate different groups in the form of their principal component scores. The spectra for each sample were averaged to give an overall spectrum for that sample. Then, leave-one-out cross-validation (LOOCV) was used to train and test PCA/SVM. In this procedure, an SVM model was initially built from the spectra of all except one spectrum. Then, the algorithm predicted the classification of omitted spectrum and stored the result. This procedure was repeated with each omitted spectrum in turn, leading to independent predictions per spectrum. For each predicted spectrum, a probability of prediction was calculated and expressed as a sensitivity and specificity for each group.

Student’s t test was performed at each Raman shift to compare the mean spectral intensities between different groups within the data set. P values of <0.01 were considered significant. Progression-free survival was calculated as the time between initial ADT and the date of progression to CRPC with the Kaplan–Meier technique and SPSS software, version 16.0 (SPSS, Chicago, Illinois). Median follow-up time was calculated by the inverse Kaplan–Meier technique and statistical significance was set at a P value of <0.05. The diagnostic accuracy of the predictions was quantified by constructing a receiver operating characteristic (ROC) curve. This represented the relationship between sensitivity and specificity at different cutoff values of the probability. The integration area under the ROC curves for PCA/SVM was calculated, illustrating the efficacy of the diagnostic algorithms.

3 Results

3.1 RS Could Predict the CRPC Cell Accurately

A total of 49 and 50 spectra were obtained from androgen-dependent (LNCaP) and castration-resistant (C4-2) cells, respectively (Fig. 1). As compared with LNCaP, C4-2 showed higher intensities at 1126, 1171, and 1447 cm⁻¹, while lower at 783, 1299, and 1584 cm⁻¹. Black histogram in Fig. 1 presented the corresponding significance (1 – P) of Raman spectral changes (e.g., Raman peak intensities, positions, and spectral bandwidths broadening or narrowing) between C4-2 and LNCaP. PCA/SVM algorithm could predict the C4-2 cell with a sensitivity of 94.00% and a specificity of 95.92%.

3.2 RS Revealed the Raman Changes Between CRPC and ADPC Tissues, and Effectively Diagnosed CRPC

After the removal of obviously heterogeneous spectra, 826 spectra from 50 enrolled patients were analyzed, including 17 patients with CRPC and 33 patients with ADPC (Table 2). Among these 17 patients of CRPC, specimens biopsied prior to ADT were available from nine patients (Table 2), and spectra were compared both before and after developing CRPC. Figure 2(a) illustrated mean spectrum of CRPC and ADPC tissues and corresponding significance of Raman spectral changes (e.g., Raman peak intensities, Raman peak positions and

| Table 2 | The characteristics of nine paired patients in their ADPC and CRPC status. |
|----------|----------------|------------------|----------------|----------------|
| Age (years) | PSA (ng/mL) | Gleason score | TNM stage | Treatment | PSA (ng/mL) | Gleason score | Time to CRPC (months) |
| 79 | 35.62 | 6 | T3N1M1 | BO + B | 15.14 | 8 | 72 |
| 77 | 46.19 | 7 | T3N0M0 | BO | 11.35 | 9 | 27 |
| 64 | 789.80 | 7 | T3N1M1 | BO + B | 6.06 | 8 | 9 |
| 79 | 10.27 | 9 | T4N0M0 | BO + F | 18.90 | 8 | 13 |
| 67 | 51.20 | 7 | T3NXM1 | B + G | 4.44 | 8 | 30 |
| 59 | 107.60 | 7 | T3N1M1 | BO + B | 139.56 | 7 | 51 |
| 72 | 91.94 | 8 | T3N0M1 | B + G | 3.99 | 10 | 13 |
| 65 | 15.14 | 6 | T2N1M0 | RP + B + G | 10.20 | 8 | 10 |
| 74 | 76.90 | 6 | T4N1M1 | BO + B | 9.04 | 8 | 11 |

BO: bilateral orchiectomy; B: bicalutamide; F: flutamide; G: goserelin; RP: retropubic prostatectomy.
skeletal band-widths broadening or narrowing). To further confirm the spectral changes, spectra from nine paired patients were compared both before and after their CRPC status [Fig. 2(b)]. The peaks at 1126 (C\(\text{C}/\text{CN}\) stretching of protein), 1170 (tyrosine), and 1447 (CH\(_2\) deformation of protein; deoxyribose) cm\(^{-1}\) were significantly increased in CRPC, while significant peaks in ADPC were 1315 to 1338 (tryptophane; adenine) cm\(^{-1}\) (Refs. 15, 16, 22, and 25). This indicated that there was an increase or decrease in the percentage of a certain type of biomolecules relative to the total Raman-active constituents from ADPC to CRPC transformation. Figure 2(c) and 2(d) denoted the corresponding pathologic tissue with H&E stain. Figure 2(e) was scatter plots of first principal component versus third principal component of Raman spectra to demonstrate

![Figure 2](https://www.spiedigitallibrary.org/journals/Journal-of-Biomedical-Optics)

Fig. 2 The mean spectrum of castration-resistant prostate cancer (CRPC) and androgen-dependent prostate cancer (ADPC) tissues in total (a), nine paired patients (b), and the corresponding significance (1 - P value) of Student’s t test along with the spectral axis (histogram). The corresponding pathologic tissues of CRPC and ADPC with hematoxylin & eosin (H&E) stain (c and d). A scatter plot of the spectra for CRPC and ADPC (e) used to develop the model projected onto a plane in a principal component.

Journal of Biomedical Optics 087001-4 August 2013 • Vol. 18(8)
distinctively spectral clustering for CRPC and ADPC, in which each of the spectra was represented by different color and different shape. It showed a significant classification in a two-dimensional coordinate system. PCA/SVM provided an overall sensitivity of 88.2% (15/17) and a specificity of 87.9% (29/33) for diagnosing CRPC. Since a Gleason score 8 to 10 of initial diagnosis correlated with CRPC progression, stratified analysis of spectra was, respectively, performed in Gleason score 8 to 10 and Gleason score ≤7 groups to remove cross-interaction of pathological grading. Table 3 presented stratified sensitivity and specificity for different Gleason score groups.

### Table 3

<table>
<thead>
<tr>
<th></th>
<th>Total no. of cases</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>50</td>
<td>88.2</td>
<td>87.9</td>
</tr>
<tr>
<td>Gleason score ≤7</td>
<td>19</td>
<td>83.3</td>
<td>84.6</td>
</tr>
<tr>
<td>Gleason score 8 to 10</td>
<td>31</td>
<td>90.9</td>
<td>90.0</td>
</tr>
</tbody>
</table>

### 3.3 RS Prognosed the Progression Interval from ADPC to CRPC

For the 33 patients with ADPC, there was at least 12 months follow-up after initial ADT. Fourteen patients progressed to CRPC within 12 months. The patients were divided into progression and nonprogression groups depending on whether their cancer progressed to CRPC within 12 months. Figure 3(a) showed the mean spectrum of progression and nonprogression groups, however, very slight spectral changes were observed at 1207 (phenylalanine, tyrosine) and 1247 (amide III of protein; thymine) cm⁻¹ which corresponded to the changes of biomolecules in progression to CRPC (Refs. 15, 16, 22, and 25). From the PCA, the first and second principal components scores separated the two groups. A scatter plot of these principal components was shown in Fig. 3(b). The PCA/SVM diagnostic model

![Fig. 3](https://www.spiedigitallibrary.org/journals/Journal-of-Biomedical-Optics)

**Fig. 3** The mean spectrum of progression and nonprogression groups (a) and the corresponding significance (1 - P value) of Student’s t test along with the spectral axis (histogram). A scatter plot of the spectra for progression and nonprogression groups (b) used to develop the model projected onto a plane in principal component.

![Fig. 4](https://www.spiedigitallibrary.org/journals/Journal-of-Biomedical-Optics)

**Fig. 4** Kaplan-Meier curve (a) of progression-free survival rates for patients who were prognosed progression to CRPC by Raman spectroscopy (RS) (n = 15) or not (n = 18) (P = 0.027). Receiver operation characteristic (ROC) curve (b) representing the accuracy of RS discrimination with PCA/SVM. The areas under the curve for CRPC–ADPC and progression and nonprogression groups were 0.920 and 0.903.
could detect the progression to CRPC with a sensitivity of 85.7% (12/14), a specificity of 88.9% (16/19), a positive predictive value of 80% (12/15), and a negative predictive value of 78% (16/18).

Furthermore, the patients were separated into two groups depending on whether the patients were prognosed progression to CRPC by RS. Fifteen patients were predicted as progression to CRPC by RS and 18 were not. The overall median progression-free interval was 13.5 months. By using log-rank test, there was a significant difference in progression-free survival between these two groups [P = 0.027, Fig. 4]. To further evaluate the performance of the PCA/SVM algorithms together with the LOOCV method, the ROC curve [Fig. 5] was also generated. The areas under the ROC curve for CRPC and ADPC and progression and nonprogression groups were 0.920 and 0.903, respectively, confirming that the PCA/SVM diagnostic model was powerful for clinical diagnosis at the molecular level.

4 Discussion
The multistep carcinogenesis of prostate epithelia initially manifest as an androgen-dependent organ-confined cancer then progress to metastatic, castration resistant phenotype. PC cells encompass many genetic and epigenetic changes in epithelial cells as well as surrounding microenvironment. Currently, many mechanisms with a wide spectrum of molecular alterations have been identified for the onset of CRPC. However, due to the heterogeneity of PC, a single gene or protein marker was not likely to identify CRPC with a great accuracy based on clinical experience.

RS is a well-established quantifiable optical technique with excellent reproducibility, and its spectrum might provide information in identifying structural changes of molecule. Each molecule had its own pattern of vibration that could serve as a “Raman biomarker.” RS has been used to accurately identify benign prostate hyperplasia and PC tissues, otherwise, heterogeneous cell lines (LNCaP, CaP2b, PC-3, and DU-145) with significant biochemical differences were applied in their research, but whether RS could identify congregate CRPC and ADPC cells and tissues in vivo was still confusing. We used the classic LNCaP/C4-2 cells as cell model which mimicked closely the natural history from ADPC to CRPC (Ref. 20), then selected and matched the PC samples with coincident Gleason score to reduce the intrinsic interference. Notably, spectra of the nine paired patients presented more reliable results in vivo. We compared the spectral changes between CRPC and ADPC in both cell and tissue models. It revealed the corresponding spectral changes reflected structural changes of protein or DNA, concomitant with the alteration of amino acid and basic radical. Depending on these changes of conformation or component in some biomolecules, PCA/SVM effectively identified CRPC and ADPC. However, these spectral changes only represented partial information of the biomacromolecules, such as the level of secondary protein structure. Further, molecular biological studies were needed to determine the mechanism of these alterations.

Gleason score 8 to 10, achieving nadir PSA level, and time to PSA nadir after ADT might correlate with the progression to CRPC in many retrospective analyses. However, nadir PSA level and time to PSA nadir were unpredictable in initial diagnosis, and Gleason score was nonspecific for CRPC. It was known that androgen ablation provided a selective advantage to androgen-refractory cells that grow and eventually comprise most of the tumor. Assuming that ADT effectively targeted the androgen-sensitive population of PC cells, an incomplete and sluggish response or short-term progression to CRPC was evidence of significant androgen-refractory population.

Thirty-three patients with ADPC were followed-up after initial ADT. We separated these patients into progression and nonprogression groups depending on whether their cancers progressed within 12 months. Slight spectral changes were observed, and PCA/SVM could prognose whether ADPC would progress to CRPC within 12 months accurately. We further separated the patients into two groups depending on whether the patients were prognosed progression to CRPC by RS, then compared the progression-free survival between them. A significant difference was detected by log-rank test, illustrating the efficacy of RS together with PCA/SVM algorithms for prognosing CRPC. However, the experiment must be expanded to make a more detailed stratification of time to CRPC and ensure RS was available and reproducible in clinical practice. Another prospective work is ongoing to increase the sample size and follow-up interval and also to predict survival besides CRPC through this trained algorithm.

Vibrational spectroscopy, both Raman and infrared (IR) imaging, are the prominent approaches to sense specific types of molecules or otherwise resolve the chemical species and morphologic structures. Bhargava groups coupled the IR with statistic pattern recognition of spectra to differentiate benign from malignant prostatic epithelium and the IR classified image was accurately overlaid with the H&E-stained image which had the potential of assessing the margin of cancerous tissues in surgery in real time. At the same time, the PC cells and tissues could be predicted according to different Gleason grade and tumor stage by the IR (Refs. 21 and 22). While each technology promises a specific measurement for specific situations, IR spectroscopic imaging provides a rapid and simultaneous fingerprinting of inherent biologic content, yet Raman spectroscopic imaging permits a high spatial resolution and could be applied in aqueous environments. Hence, a careful matching of vibrational spectroscopy to the clinical application could lead to useful protocols. Vibrational spectroscopic imaging directly provides the molecular descriptors, but also helps the pathologists make better diagnosis.

Nevertheless, there are two possible limitations of this study. First, overall sample number was not large enough and population with Gleason score ≤5 was not collected to demonstrate the difference between CRPC and ADPC. Because ultrastructural differences related to androgen-dependent or androgen-refractory state of cells or tissues could be detected by RS, we inferred that RS could gain similar results. Second, we used paraffin-embedded tissues for Raman analysis, therefore the influence of residual wax on RS cannot be ruled out completely. Although these tissues were dewaxed in three washes of fresh xylene and ethanol baths, slight Raman signals at 1062 and 1296 cm⁻¹ originated from residual wax have been documented. In this case, snap-freezing tissue sections could be a good alternative source for this application.

Acknowledgments
The authors thank Professor Leland W.K. Chung at Cedars-Sinai Medical Center, Los Angeles, USA, for provision of C4-2 cells, Dr. Quan-li Wang at Department of Epidemiology and Health Statistics, Xi’an Jiaotong University College of Medicine, China, for the help with statistical analysis.
Professor Victor K. Lin at Department of Urology, University of Texas Southwestern Medical Center, USA, for his excellent job critically revising this manuscript. Funding: National Natural Science Foundation of China (No. 30701009) and National Basic Research Program of China (973 Program, No. 2009CB526408).

References