Understanding prostate cancer cells metabolome: a spectroscopic approach
Francisco Santos\textsuperscript{a,b}, Sandra Magalhães\textsuperscript{a}, Magda Carvalho Henriques\textsuperscript{b}, Beatriz Silva\textsuperscript{c}, Isabel Valença\textsuperscript{c}, Daniela Ribeiro\textsuperscript{c}, Margarida Fardilha\textsuperscript{b} and Alexandra Nunes\textsuperscript{a}

\textsuperscript{a} iBiMED – Institute of Biomedicine, Department of Medical Sciences, University of Aveiro, Aveiro, Portugal
\textsuperscript{b} Signal Transduction Laboratory, iBiMED – Institute of Biomedicine, Department of Medical Sciences, University of Aveiro, Aveiro, Portugal
\textsuperscript{c} Organelle Dynamics in Infection and Disease Laboratory, iBiMED – Institute of Biomedicine, Department of Medical Sciences, University of Aveiro, Aveiro, Portugal

Corresponding author: Alexandra Nunes
iBiMED – Institute of Biomedicine, Department of Medical Sciences
University of Aveiro
Agra do Crasto, 3810-193, Aveiro, Portugal
alexandranunes@ua.pt

Abstract
Prostate cancer (PCa) is the second most common neoplasia in men. Because it is often diagnosed at a late stage, mortality rates remain high. Studying cancer metabolome, which reflects early changes that occur in cells, has gained relevance and may contribute to the identification of early diagnostic biomarkers and understanding tumor biology. Fourier-transform infrared (FTIR) spectroscopy is a metabolomics technique that probes the biochemical composition of the analyzed samples and allows to discriminate samples with distinct metabolic profiles, allowing the discrimination between cancerous and non-cancerous samples. In this study, FTIR spectra were acquired from PCa and normal prostate cell lines and analyzed by principal component analysis (PCA). Our results indicate a clear discrimination between the different cell lines, meaning that they exhibit distinct metabolic profiles. This discrimination can be attributed to an altered lipid metabolism (3000-2800 cm\textsuperscript{-1}, 1800-1700 cm\textsuperscript{-1} and 1500-1400 cm\textsuperscript{-1}) and changes in protein conformation (1700-1600 cm\textsuperscript{-1}). These results suggest that studying cancer metabolome with FTIR spectroscopy not only allows the understanding of tumor metabolic behavior and may be useful to the development of new therapeutic targets.

Keywords: Prostate cancer, cell lines, metabolomics, FTIR, metabolic profile, principal component analysis
1. Introduction

Worldwide, 1.1 million new cases of PCa were registered in 2012, 307,000 of which resulted in death. These figures place PCa as the second most common oncological disease in men, ranking just behind lung cancer (1). Many risk factors have been associated with PCa, such as race, family history and increasing age, with more than 60% of new cases being diagnosed in men aged between 60 and 70 years (1,2). Additionally, sexually transmitted infections and dietary factors are also linked to the disease (2).

One of the most conventional methods for diagnosing PCa include prostate-specific antigen (PSA) testing, which is based on its concentration in blood serum (3). Despite being one of the most used, PSA testing is unspecific and false positive results are common, due to non-cancerous diseases such as benign prostatic hyperplasia (3). Once PCa is diagnosed, treatment options depend on the stage of the disease: when the tumor is clinically localized and of low-risk, patients may opt by active surveillance instead of radical interventions, such as prostatectomy (4,5). However, more aggressive and/or metastatic tumors must be treated with hormonal therapy, radiation, chemotherapy and prostatectomy, that can lead to physical impairments, such as sexual and urinary dysfunction. The 5-year survival rates are very optimistic (up to 100%) for patients diagnosed at an early stage, however it declines to approximately 28% for patients diagnosed with late-stage cancer, so there is a need for new and effective therapeutic strategies (5).

An altered metabolism has been recognized as one the hallmarks of cancer (6). Metabolomics has been a valuable tool in the field of oncology, allowing the discovery of biochemical profiles and, therefore, of differences between healthy and cancer metabolic phenotypes (7). Since the metabolome reflects even slight changes that occur in the cell, metabolomics may allow for an early intervention because metabolic alterations are believed to precede any morphological alterations that might occur (8–10). Also, studying cancer metabolome may improve the knowledge on tumor behavior and help to develop new targeted therapeutic strategies, which are crucial for a good prognosis (11,12).

In recent years, FTIR spectroscopy has emerged as a tool for metabolic profiling. This approach is based on the vibrations of the atoms in a molecule caused by the interaction of infrared (IR) radiation with matter (13–15). The molecular vibrations that result from this interaction are specific to the biochemical composition of the analyzed samples and lead to spectral differences between samples, producing a fingerprint of metabolism. FTIR spectroscopy is rapid, non-invasive and reagent-free, thus it is a suitable screening technique that may allow for the discrimination between healthy and cancer samples (16–19).

Regarding PCa, FTIR spectroscopy has been applied to tissue biopsies and cell lines, showing promising results and elucidating on biochemical alterations inherent to the disease (11,20–22). Cell lines present certain advantages that validate their use in
preliminary studies, such as the fact that they exhibit a defined cell state that allows the analysis of a target metabolic status. Also, they evade confounding factors characteristic of biofluids and tissues (for example, age, diet and gender) (23–25). Nonetheless, cell culture suffers for not being able to replicate cell-cell and cell-matrix interactions in the tumor microenvironment, which are vital for metabolic alterations that occur with tumor progression. Therefore, models that reproduce structure, function and metabolism of in vivo tumors should be considered for further validation of in vitro studies (23,26).

The main objective of this study was to perform a FTIR spectroscopy study of a PCa cell line (22Rv1) and two normal prostate epithelial cell lines (PNT1A and PNT2) and identify spectral differences by principal component analysis (PCA). To the best of our knowledge, this work is the first to analyze 22Rv1 and PNT1A by FTIR spectroscopy.

2. Material and Methods

2.1 Cell culture

Prostate cancer (22Rv1) and normal prostate (PNT1A and PNT2) cell lines were cultured in RPMI-1640 culture media, supplemented with 10% FBS and 1% penicillin/streptomycin mixture, and maintained in a humidified atmosphere at 37°C containing 5% CO₂. All the reagents used in this study were purchased from Thermo Fisher Scientific. PNT2 cell line was kindly given by Dr. Ricardo Perez-Tomás (University of Barcelona, Spain). 22Rv1 and PNT1A were a kind gift from Dr. Fátima Baltazar (University of Minho, Portugal). A detailed description of each cell line can be consulted in table 1.

2.2 Preparation of cells

After reaching around 90% of confluency, cells were prepared for FTIR analysis. To assess viability, cells were analyzed under the microscope. The culture medium was removed from each plate and cells were detached from the plate using a solution of trypsin 0.05% and EDTA 1%. To stop the reaction, culture medium was added to the plates and the cell suspensions were transferred to separate tubes for centrifugation (3 minutes, 1000 rpm). Following centrifugation, the supernatant was discarded, and cells were resuspended in culture medium. Cells were counted using a TC20™ Automated Cell Counter, from BIO-RAD, and a total of 10⁶ cells were used for each biological replicate. Cells were centrifuged for 3 minutes at 1000 rpm to remove the medium and resuspended in phosphate-buffered saline (PBS). Washing cells with PBS is essential to remove residual medium and trypsin (27). PBS was removed by centrifugation (3 minutes, 1000 rpm) and cell pellets were kept on ice until FTIR analysis.

2.3 FTIR measurements

FTIR spectra were acquired in ATR mode in a FTIR spectrometer (Alpha Platinum ATR, Bruker), and processed using OPUS software (©Bruker). The spectra were obtained
over the wavenumber range 4000 – 600 cm\(^{-1}\), with a resolution of 8 cm\(^{-1}\) and 64 co-added scans. Spectra acquisition was performed in a room with controlled temperature and relative humidity (23°C and 35%, respectively). A background spectrum was acquired with the crystal empty before each different cell line and a total of 10 biological replicates for 22Rv1 and 8 for PNT1A and PNT2 were used. The cell pellet was placed on the crystal and spectra were obtained after the samples were completely air-dried. Between each measurement the crystal was cleaned with ethanol 70%, followed by distilled water.

### 2.4 FTIR data analysis
All the raw spectra were normalized to the area of amide I band and the normalized spectra were derivatized, using the second derivative and Savitzky-Golay method with 3 smoothing points. Spectra were processed using The Unscrambler X® software (v.10.4, CAMO, Oslo, Norway). The spectral regions 3000-2800 cm\(^{-1}\), 1800-1500 cm\(^{-1}\) and 1500-900 cm\(^{-1}\) were chosen for analysis and all spectral assignments were made according to widely cited literature references.

### 2.5 Multivariate analysis
PCA was applied to the normalized second-derivative spectra of all the cell lines. For multivariate analysis, we used the spectral regions between 3000-2800 cm\(^{-1}\), 1800-1500 cm\(^{-1}\) and 1500-900 cm\(^{-1}\). All the analyses were performed on The Unscrambler X® software (v.10.4, CAMO, Oslo, Norway).

### 3. Results
#### 3.1 Overview of FTIR spectra
Cell viability was confirmed by observation under the microscope, and each cell pellet consisted of mostly viable cells (unviable cells were discarded upon the removal of the culture medium before trypsinization of the cells from the plate). An area peak normalization was applied to ensure that the varying cell number between replicates would not affect the spectral variances, therefore highlighting differences in biochemical structure and not in absorbance intensity (27,28).

#### 3.2 FTIR spectral interpretation
A complete view of the averaged normalized spectra of each cell line, in the region between 4000 and 600 cm\(^{-1}\), can be observed in figure 1A. Briefly, in the 3000-2800 cm\(^{-1}\) region (figure 1B), bands at 2956 and 2871 cm\(^{-1}\) are assigned to the asymmetric and symmetric stretching of CH\(_3\) chains of lipids, respectively, whereas the bands at 2922 and 2851 cm\(^{-1}\) are attributed to the asymmetric and symmetric stretching of CH\(_2\) chains of lipids, respectively (13,18,29,30). In the region between 1800 and 900 cm\(^{-1}\) (figure 1C), the most prominent bands are associated with proteins: amide I (1640 cm\(^{-1}\)) and amide II (1540 cm\(^{-1}\)). The amide I band is attributed to C=O and C-N stretching...
of proteins, while the amide II band is assigned to N-H, C-N and C-C stretching. The amide I band is sensitive to the secondary structure of proteins and is frequently used for this type of analysis. The amide II band can also be used for studying the secondary structure of proteins, although to a lesser extent (30–34). Also, in this region, the band around 1740 cm\(^{-1}\) is assigned to the C=O stretching of phospholipid esters. In the 1500-900 cm\(^{-1}\) region, bands between 1480 and 1300 cm\(^{-1}\) are attributed to amino acid side chains and fatty acids and the major bands between 1300 and 900 cm\(^{-1}\) arise mainly due to carbohydrates (particularly glycogen) and phosphates associated with nucleic acids (13,18,29,30).

### 3.3 Principal Component Analysis

Before PCA analysis, spectra were normalized using an area peak algorithm and a second-order derivative using Savitzky–Golay algorithm was performed to resolve overlapping peaks (35,36). PCA allows to determine the most important sources of variability between prostate cancer and normal prostate cell lines. In this study, the spectral regions 3000-2800 cm\(^{-1}\), 1800-1500 cm\(^{-1}\) and 1500-900 cm\(^{-1}\) were chosen for analysis.

PCA results for the 3000-2800 cm\(^{-1}\) range are presented in figure 2A. In the score plot, it is possible to observe that these cells are separated by principal component-2 (PC2), and that 22Rv1 are in negative PC2, while PNT1A and PNT2 are in positive PC2. According to the loadings plot, 22Rv1 are mainly characterized by the spectroscopic signals located at 2917, 2874 and 2849 cm\(^{-1}\), while PNT1A and PNT2 are characterized by the peaks at 2925 and 2857 cm\(^{-1}\).

PCA of the spectral region between 1800 and 1500 cm\(^{-1}\) is illustrated in figure 2B. In this case, the discrimination between tumor and normal cells is provided by PC3: 22Rv1 are in positive PC3, while PNT1A and PNT2 are located in negative PC3. The loadings plot indicates that the main sources of variability come from the 1700-1600 cm\(^{-1}\) region. The main spectral assignments that characterize 22Rv1 are 1679, 1665, 1648, 1622 cm\(^{-1}\), and PNT1A and PNT2 are characterized by the peaks at 1653 and 1636 cm\(^{-1}\).

In the 1500-900 cm\(^{-1}\) region, 22Rv1 are present in negative PC2, while PNT1A and PNT2 are located in the positive PC2. The main spectral assignments that characterize 22Rv1 are 1427, 1384, 1152, 1127, 1073, 1036, 1022 and 982 cm\(^{-1}\). PNT1A and PNT2 are characterized by the peaks 1240, 1104, 1084, 1050 and 965 cm\(^{-1}\).

### 4. Discussion

Several studies have shown the potential of FTIR spectroscopy in distinguishing the metabolic profile of normal and cancer cells. This technique has been applied to cell lines derived from numerous types of neoplasms, such as breast (28,37), cervix (38), colon (39,40), gastric (41) and lung (42) cancers.

In this study, we have demonstrated the use of FTIR spectroscopy to differentiate between prostate cell lines derived from normal epithelium (PNT1A and PNT2) and
primary prostate tumor (22Rv1). PCA clearly discriminated spectra from 22Rv1, and PNT1A and PNT2 in the 3000-2800 cm\(^{-1}\), 1800-1500 cm\(^{-1}\) and 1500-900 cm\(^{-1}\) regions, indicating that they are biochemically different and, therefore, exhibit different metabolic profiles. In the spectral region from 3000 to 2800 cm\(^{-1}\), the main spectral assignments that allowed the separation of PNT1A and PNT2 are 2925 and 2857 cm\(^{-1}\), which arise, respectively, due to the asymmetric and symmetric stretching vibrations of CH\(_2\) chains of lipids. Furthermore, the bands at 2917 and 2849 cm\(^{-1}\), which were responsible for separating 22Rv1 from PNT1A and PNT2, can also be attributed to the asymmetric and symmetric stretching vibrations of CH\(_2\) chains of lipids. Because spectra were acquired with a resolution of 8 cm\(^{-1}\), the aforementioned peaks can be assigned to the same biochemical component. A shift to a lower frequency might suggest alterations in the lipid composition between normal and tumor cells, and an increase in disorder of the CH\(_2\) chains of membrane lipids have been associated with malignancy (17). Alterations in the membrane lipids have already been reported in prostate cancer tissue biopsies (21) and breast cancer-derived cell lines (37). Furthermore, alterations regarding lipid saturation have been detected in this spectral region. An increment in CH\(_2\)/CH\(_3\) ratio suggests an increase in the length of fatty acid chains.

In the spectral region between 1800 and 1500 cm\(^{-1}\), the main assignments that are responsible for segregating 22Rv1 from PNT1A and PNT2 are 1758, 1679, 1665, 1648 and 1622 cm\(^{-1}\). The peak at 1758 cm\(^{-1}\) can be attributed to the C=O stretching of lipid esters (43), suggesting a dysregulated lipid metabolism, which is known to be altered in cancer. In fact, an accumulation of cholesterol has been reported in PCa. Also, de novo lipid synthesis has been described in PCa, which serves to produce fatty acids. Fatty acid synthesis plays an important role in cancer pathogenesis, and it is known that these newly synthesized fatty acids support cellular processes to promote cell proliferation and survival (44,45). Additionally, de novo lipogenesis has been shown to promote saturation of membrane lipids, which has consequences in terms of membrane dynamics and uptake and efficacy of chemotherapeutics (46).

The amide I band (1700-1600 cm\(^{-1}\)) arises due to the stretching vibrations of C=O and C-N, and provides information about the secondary structure of proteins, which has been extensively investigated in several studies (31–33). The spectral assignments specific of this region that characterize 22Rv1 are 1679, 1665, 1648 and 1622 cm\(^{-1}\). The frequency at 1679 cm\(^{-1}\) is assigned to antiparallel β-sheets, while 1665 and 1648 cm\(^{-1}\) are related to β-turns and unordered structures, respectively (47). Regarding the peak assigned to the antiparallel β-sheets, these structures are extremely present in less soluble proteins that are likely to form aggregates (32,34). An increase in the relative amount of β-sheets, relative to α-helical segments, has already been observed in colon adenocarcinoma cell lines (40). The peak at 1622 cm\(^{-1}\), however, can be specifically associated to the presence of protein aggregated strains (32,48). Cancer cells are exposed to factors that induce stress and, therefore, disrupt proteostasis, causing protein misfolding and subsequent aggregation (49). Several tumors and cancer cell lines
have been found to exhibit protein aggregation, mainly involving aggregates of tumor suppressor p53 (50–52), which can play an important role in carcinogenesis (53). Also, aggregation of p53 has been associated to platinum resistance and stem cell phenotype in a subset of ovarian cancer (54), which suggests a link between loss of proteostasis and cancer progression. On the other hand, the spectral assignments that characterize PNT1A and PNT2 are 1653 and 1636 cm\(^{-1}\), which are related to α-helices and parallel β-sheets, respectively. α-helices are usually present in more soluble proteins that are less likely to aggregate (34), and despite the presence of the assignment at 1636 cm\(^{-1}\) associated with β-sheets, this could be explained by the fact that most proteins exhibit a mixture of secondary structure (33).

A discrimination between tumor and normal cells was also observed in the spectral region between 1500 and 900 cm\(^{-1}\). The main assignments in the 1500-1300 cm\(^{-1}\) range that characterize 22Rv1 are 1427 and 1384 cm\(^{-1}\), the former being associated with the CH\(_3\) and CH\(_2\) deformation vibrations mostly due to lipid contribution, and the latter with the CH\(_3\) symmetric wagging of phospholipids, fatty acids and triglycerides (38,55). These results may confirm the lipid alterations that were detected in other regions of the spectrum. The bands between 1460 and 1400 cm\(^{-1}\) arise due to lipid contribution, and some have been explored in previous studies. For instance, the uterine cervical adenocarcinoma cell line SiSo, along with exfoliated cervical cells and cervical adenocarcinoma tissue, exhibit an increased intensity and a shift to a lower frequency of the band at 1400 cm\(^{-1}\), which might indicate structural changes of the CH\(_2\) chains of lipids (38).

The bands present in the region between 1300 and 900 cm\(^{-1}\) arise mainly due to carbohydrates and phosphates associated with nucleic acids. The most important spectral assignments characteristic of 22Rv1 are 1152, 1127, 1073, 1036, 1022 cm\(^{-1}\). The bands at 1152 and 1022 cm\(^{-1}\) are related to carbohydrates, particularly glycogen (which is specifically assigned to the frequency 1022 cm\(^{-1}\)). A reduced glycogen content has been observed in PCa tissue biopsies (21), and cervical adenocarcinoma tissue and cervical adenocarcinoma cells (38). The decrease in glycogen content in malignant cells can be attributed to their higher metabolic activity and, thus, an increase in glycolysis (11,17,38,56). The remaining bands (1127, 1073 and 1036 cm\(^{-1}\)) account for differences associated with nucleic acids. This could be linked to the fact that nucleic acids exhibit increased hydrogen-bonding of the phosphodiester groups and a higher level of packing in cancer cells, as confirmed by several studies (40).

In this study FTIR spectroscopy and PCA were successfully applied to a PCa and two normal prostate (PNT1A and PNT2) cell lines. Our results clearly indicate that this technique, coupled with multivariate analysis, can be used to distinguish the different types of prostate cells in several spectral regions. However, in order to validate our results, other biological samples should be considered for analysis, such as human urine, serum/plasma and prostatic fluid (57). Also, because FTIR spectroscopy is uncapable of detecting specific metabolites that are responsible for causing changes, other
metabolomic techniques should be considered, such as mass spectrometry and nuclear magnetic resonance. Prostate cancer metabolome has been studied by these techniques, allowing the identification of metabolites implicated in certain metabolic pathways (58–61). Nevertheless, our results indicate that primary tumor cells can be distinguished from normal cells, and that lipid metabolism obviously plays an important role in tumorigenesis. Other alterations, such as protein aggregation, can also be an important event in cancer progression. Given the fact that the biochemical composition of the prostate cell lines can be detected by FTIR, it could be useful in providing insights on new targets for more accurate and sensitive therapeutic strategies.

Funding
This work was supported by Fundação para a Ciência e a Tecnologia, FEDER and COMPETE (grants number UID/BIM/04501/2013, POCI-01-0145-FEDER-007628, POCI-01-0145-FEDER-016728 and PTDC/DTP-DES/6077/2014). SM, MCH and IV are also funded with individuals grant of Fundação para a Ciência e Tecnologia I.P. (SFRH/BD/131820/2017, SFRH/BD/131846/2017 and SFRH/BD/101942/2014, respectively).

References
Figure 1: Normalized mean spectra of prostate cancer (22Rv1) and normal prostate (PNT1A and PNT2) cells in the 4000-600 cm$^{-1}$ region (A); Amplification of the regions between 3000 and 2800 cm$^{-1}$, and 1800 and 900 cm$^{-1}$. X-axis: wavenumber (cm$^{-1}$); Y-axis: arbitrary units (AU).
Figure 2: PCA score and loadings plots from the normalized second derivative spectra of prostate cancer (22Rv1) and normal prostate (PNT1A and PNT2) cell lines on the 3000-2800 cm\(^{-1}\) (A), 1800-1500 cm\(^{-1}\) (B) and 1500-900 cm\(^{-1}\) (C) regions.
<table>
<thead>
<tr>
<th>CELL LINE</th>
<th>DERIVATION</th>
<th>MORPHOLOGY</th>
<th>AR PROTEIN</th>
<th>DISEASE</th>
<th>TUMORIGINIC</th>
<th>REFERENCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>PNT1A</td>
<td>Normal prostate epithelial cells immortalized with SV40 genome</td>
<td>Epithelial</td>
<td>Yes</td>
<td>Healthy</td>
<td>No</td>
<td>(62)</td>
</tr>
<tr>
<td>PNT2</td>
<td>Normal prostate epithelial cells immortalized with SV40 genome</td>
<td>Epithelial</td>
<td>Yes</td>
<td>Healthy</td>
<td>No</td>
<td>(62)</td>
</tr>
<tr>
<td>22Rv1</td>
<td>Primary tumor (established from xenograft CWR22R)</td>
<td>Epithelial</td>
<td>Yes</td>
<td>Carcinoma</td>
<td>Yes</td>
<td>(63,64)</td>
</tr>
</tbody>
</table>