Assessment of HR-MAS NMR Prostate Biopsy Tissue Spectra with Principal Component Analysis and Metabolite Quantification

P. A. DiCamillo[1,2], M. G. Swanson[3], J. Kurhanewicz[3], S. J. Nelson[1,2]

[1]UCSF/UCB Joint Graduate Group in Bioengineering, University of California, San Francisco, CA, United States, [2]Surbeck Laboratory of Advanced Imaging, Department of Radiology, University of California, San Francisco, CA, United States, [3]Department of Radiology, University of California, San Francisco, CA, United States

INTRODUCTION: *Ex-vivo* high resolution magic angle spinning (HR-MAS) NMR spectroscopic studies involving prostate biopsy tissues allow complete metabolic profiles to be associated with the composition and pathology of the same tissue. However, biopsy studies are particularly challenging due to the inherent tissue heterogeneity, contamination by blood and lipids, and small sample size. Automated and objective approaches to analyze this spectroscopic data and assess patient prognosis are desirable. In this study, the application of principal component analysis (PCA) to high field (11.7T) HR-MAS data of prostate biopsy tissue is described and compared to LCModel quantification.

METHODS: Transrectal ultrasound guided biopsies (N=30) were obtained from 23 patients and stored at -80°C until ready for use. HR-MAS data were acquired at 11.7 T (500 MHz for ^{1}H), 1°C, and 2,250 Hz spin rate using a Varian INOVA spectrometer, equipped with a 4mm gHX nanoprobe. 3.0 µl of D_2O containing 0.75% TSP (D_2O+TSP) was pipetted into the bottom of a 10 µl zirconium rotor and weighed, after which the tissue samples were weighed and added to the rotor. 1D "presat" spectra were acquired with 2s presaturation, 2s acquisition, 64 transients, 40,000 data points, 20,000 Hz spectral width, and a 90° flip angle.[1] The data were zerofilled, Fourier transformed, phased, referenced to TSP (0 ppm), and normalized to the peak height of the creatine resonance at 3.04 ppm. The frequency range from 2.3 to 3.7 ppm was divided equally into 300 bins, each comprising the sum of the real part of the points in its respective segment. The data were divided into three categories according to tissue composition following histological evaluation by a pathologist: healthy glandular (>40% glandular), healthy stromal (≥80% stromal), and cancer (≥20% cancer). Principal component analysis was performed on each pair of tissue groups as well as the entire dataset. These results were compared to results from Provencher's LCModel algorithm[2] applied to the same spectra, which were evaluated with Kruskall-Wallis and Mann-Whitney statistics to confirm that the key features in the PCA corresponded to the significant differences in the data.

RESULTS: Metabolic features of the data are shown in the PCA (Figure 1) and confirmed to be significant in the LCModel results (Figure 2). PCA defines and orders covarying features in a dataset. In panels D through I, features that vary directly share y-axis sign, while inversely varying features are noted on opposite sides of the y-axis. Thus panel D indicates that polyamines and citrate both increase when choline decreases – a key difference between glandular and cancer tissue. Panels D, E, and F consider two tissue groups at a time, while G, H, and I consider all three data groups together. A key component of each pairwise PCA is represented in the three-way PCA; for example the agreement between D and G indicates that the key features between glandular and cancer spectra can be seen even in the additional presence of the stromal tissue. The means of the spectra of each data class are shown as a reference (A, B, C). This method is interesting as no attempt was made to individually quantify the metabolites or the baseline, but the differences between tissue allow for discrimination between these groups. The LCModel analysis (Figure 2) depicts 8 metabolites that vary significantly between tissue groups. Importantly, citrate, cholines and polyamines agree with the principal component results. It should be noted that a low threshold criterion (20%) of cancer was used to qualify samples into the cancer group; therefore, some tissue samples from cancer infiltrating into healthy glandular tissue can account for the higher levels of citrate with respect to healthy stromal tissue.

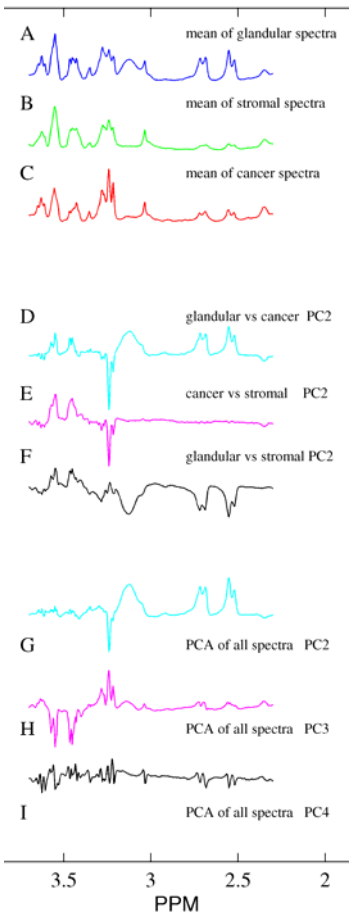

Figure 1: Prostate biopsy principal component analysis results. Mean of input spectra from glandular, stromal and cancer samples (A,B,C). Pathopneumonic differences in each pair of metabolites are seen in pairwise PCA analysis (D,E,F). The three components shown from a PC analysis on the complete dataset with all three classes together also highlight differences in Choline, Citrate and Polyamines levels (G&D, H&E, I&F). Citrate (2.5 and 2.7 ppm), Polyamines (3.1 ppm) and Choline (3.2 ppm) resonances can be clearly seen. 3.45 and 3.57 ppm resonances are due to variable levels of blood contaminants in the samples.

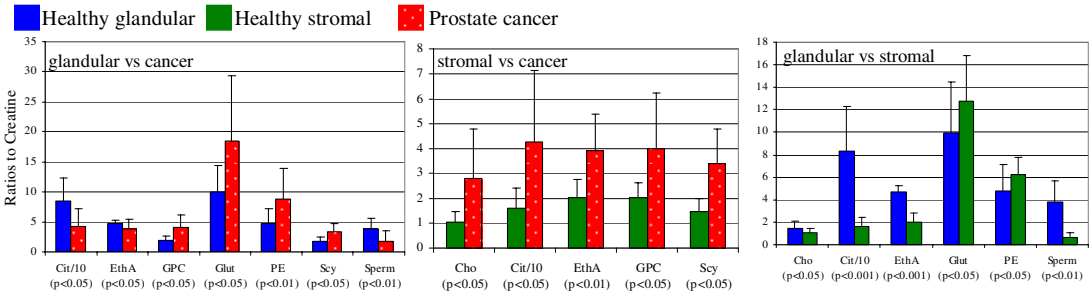

Figure 2: Metabolite differences in prostate biopsies are confirmed with LCModel analysis of same dataset. Metabolites pictured are Choline (Cho), Citrate (Cit), Ethanolamine (EthA), Glutamate (Glu), Phosphoethanolamine (PE), Spermine (Sperm), Glycerophosphocholine (GPC), and Scyllo-Inositol (Scy). The Citrate values have been reduced 10 fold in the plots due to their much larger values.

CONCLUSIONS: The results depict the major differences between the three prostate tissue classes and outline a fully automated method to analyze the data. While the metabolite quantification provides detailed information about the metabolic profile, the PCA approach provides components that can be employed as criteria for classification. An extension of this technique could provide a means to detect metabolic changes in biopsy tissues containing smaller amounts of cancer or mixed tissue.

REFERENCES: [1] Swanson MG, *et al. Proc ISMRM.* 2005; 13:263.
[2] Provencher SW. *Magn Reson Med.* 1993; 30(6):672-92.

This study was supported by LSIT01-10107.