Random Forest Classification of Raman Spectra for Reliable Diagnosis of Hepatocellular Carcinoma

T. Tolstik[1#], C. Marquardt[1#], C. Matthäus[2], C. Beleites[2], C. Bielecki[1], C. Krafft[2], O. Dirsch[4], J. Popp[2,3], A. Stallmach[1]

[1]Clinic for Internal Medicine IV, Jena University Hospital, Jena, Germany
[2]Institute of Photonic Technology, Jena, Germany
[3]Institute of Physical Chemistry and Abbe Center of Photonics, Friedrich Schiller University, Jena, Germany
[4]Institute of Pathology, Friedrich Schiller University, Jena, Germany

[#] These authors contributed equally to this work

Hepatocellular carcinoma (HCC) is the fifth most common malignant tumor worldwide and nearly 600,000 HCC-related deaths per year [1]. The disease is often coexisting with hepatitis infections or cirrhosis of the liver, which makes early diagnosis and prognosis of HCC even more complicated [2]. For this reason, development of novel diagnostic methods is a crucial step in HCC management. Raman imaging allows label-free determination of the molecular composition of biological samples on cellular and subcellular level. By using classification algorithms together with information of the morphological structures, Raman spectroscopy provides a powerful tool for early HCC detection and offers the potential to prevent therapeutic mistakes due to false diagnoses of malignancy.

In this study we investigated molecular information of human liver tissue by Raman imaging in order to differentiate, classify and predict malignant and non-malignant regions. Therefore, 23 surgically resected human liver specimens were measured using confocal Raman microscopy. Raman maps were acquired, correlated with HE-stained microscopy images and classified by an experienced hepatopathologist into malignant and benign regions of HCC (n=12) and fibrosis (n=17). The average Raman spectra corresponding to the types of tissue showed distinct spectral differences with increasing intensities for HCC samples at 2855, 1656, 1435, 1298, 1083, 1063, and 890 cm^{-1} and increasing intensities at 1660, 1245, 934, and 738 cm^{-1}. In order to classify malignant and non-malignant liver tissue, random forest models were used and validated with 101×iterated 7-fold cross-validation (patient-wise), resulting in a sensitivity of 76% and a specificity of 93% for HCC.

The presented results demonstrate the high potential which Raman spectroscopy has for future diagnostics of liver cancer.

Financial support from the European Union via the “Europäischer Fonds für Regionale Entwicklung (EFRE)” and the “Thüringer Ministerium für Bildung, Wissenschaft und Kultur (TMBWK)” is highly acknowledged.

References:

[1] T. Behne, M. S. Copur, *Int J Hepatol*. 2012, 7 pages (2012), doi:10.1155/2012/859076.
[2] K. Zhu, Z. Dai, J. Zhou, *J Biomol Res*. 1(1), p. 1-10 (2013).
[3] C. Krafft, B. Dietzek, J. Popp, *Analyst* 134(6), 1046–1057 (2009).
[4] C. Matthäus, C. Krafft, B. Dietzek, B. Brehm, S. Lorkowski, J. Popp, *Anal. Chem.* 84(20), 8549–8556 (2012).
[5] S. Dochow, C. Krafft, U. Neugebauer, T. Bocklitz, T. Henkel, G. Mayer, J. Albert, J. Popp, *Lab chip* 11(8), 1484–1490 (2011).