

R/Python Application Programming Interface for MSI Statistical Analysis: Tumor Micro Environment Case Study

Rima Ait-Belkacem¹, Fabien Pamelard¹, Gaël Picard de Muller¹, Lauranne Poncelet¹, David Bonnel¹ and Jonathan Stauber²

¹*ImaBiotech SAS, Parc Eurasanté, Loos, France*

²*ImaBiotech Corp, Billerica, Boston, USA*

Introduction

When using high resolution imaging instruments, mass resolving power and accuracy are provided for a better confidence in results interpretation. Although, MSI data management could be difficult for memory and space requirements. Since instrument providers do not focus onto the development of specific libraries, some standard formats like imzML were born to facilitate the exchange and processing of MSI datas. Despite, it is still difficult to deal with large binary datasets in the same time. For various data science areas, R and Python language offer many libraries for computational research. The objective of this study was to explore tumor microenvironment immuno-metabolic landscape using a new in house R and Python interface developed library to interact with Big Data MSI architecture.

Methods

Melanoma D10 cell line was transplanted in immunodeficient (NSG) mice transduced CD8 T cells. Tumors were then harvested, measured and snap frozen at day 0 (control without T cells), 2, 7 and 25 post-T cell injection). Data acquisition was performed using 7T MALDI-FTICR (Bruker Daltonics, Germany) at 50 μm spatial resolution. All the data were stored into a HDFS database engine. Acquired data were then treated with a new data processing workflow based on a new home based R and Python libraries. Through Web Services, this interface allowed to easily request the Hadoop data source to get the mass filters information. In addition, statistical analysis, treatments and plots were made using NumPY, SkiImage and MatPlotLib, which allowed more biological interpretation.

Preliminary data

From a technical point of view, R and Python interface allowed to offer to the scientific community, an access to open source libraries which helps to perform Machine learning, statistical classification and images analysis. Also, it enabled getting easy methods to BigData engine and more key information input, such as: mass filters matrices or spectrum at a position. This combined to an open source libraries; we were able to retrieve previous results obtained by commercial software, with more flexibility in terms of parameters or using new statistical libraries, according to the user's needs. This interface allowed reaching higher molecular information degree regarding the biological case study. Molecular imaging analysis highlighted the molecular distribution of different amino acid, nucleotide, energetic, hexosamine and TCA cycle metabolites. In parallel, CD3-staining showed lymphocytic infiltration starting from day 7 with extensive areas at day 25. After histological and molecular images overlay, metabolites level was measured from different regions. Potential correlation between metabolites/nutrients abundance and tumor/immune cells presence was studied. A decrease of Malate, N-acetylglucosamine, GSH and GSSG was noticed in tumor cells when their level was stable or increasing in LT immune cells. Ornithine and Histidine were increasing in tumor when decreasing in LT immune cells. Finally, ADP, AMP, Inosine, UMP, UDP-N-acetylglucosamine, Phenylalanine and Tyrosine level was stable in tumor when changing in LT cells.

Novel aspect – Limit 20 words

R/Python interface combining existing libraries and BigData for a better data access and interpretation in MSI data science.