Bioimage informatics

MSIdV: a versatile tool to visualize biological indices from mass spectrometry imaging data

Eisuke Hayakawa[1]*, Yoshinori Fujimura[1] and Daisuke Miura[1]*

[1]Innovation Center for Medical Redox Navigation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan

*To whom correspondence should be addressed.

Associate Editor: Prof. Robert Murphy

Abstract

Summary: Mass spectrometry imaging (MSI) visualizes the simultaneous lateral distribution of multiple compounds on sample surface. However, it is still difficult to visualize biological indices such as energy charge index from multiple compounds because of the lack of publicly available tools. Here we present MSIdV, a visualization tool for biological indices calculated from mass spectrometry imaging data, which can effectively scan a series of mass spectra and process, calculate and visualize user-defined index measures accurately with a number of signal processing features.

Availability and implementation : MSIdV is implemented in Python 2.7 and is freely available on the web at https://sourceforge.net/projects/msidv/

Contact: eisuke.hayakawa@gmail.com

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1   Introduction

Mass spectrometry imaging (MSI) is a label-free method to acquire spatially resolved multiple molecular distribution by acquiring mass spectra across the surface of a sample (Caprioli *et al.*, 1997). Typically, matrix-assisted laser desorption/ionization (MALDI) and desorption electrospray ionization technology are applied to ionize various endogenous compounds for MSI (Takáts *et al.*, 2004). This technique has been used for a variety of samples ranging from brain tissues to bacterial colonies (Rohner *et al.*, 2005; Watrous and Dorrestein, 2011). It has enormous potential in biological studies, including pathological analysis and biomarker development.

Since MSI features huge and complex structure of mass spectrum data, it is important to develop dedicated analytical software tools. Most of the current visualization tools for MSI data were designed to visualize a single molecule of interest based on its *m/z* value. However, in many biological studies, biological indices derived from the measures of multiple molecules are frequently used to assess the biological state. For example, energy charge, which is an index of metabolic energy available from adenylate, can be defined as ([ATP] + ½ [ADP])/ ([ATP] + [ADP] + [AMP]) to indicate the energy status in cells (Atkinson and Walton, 1967). Adenylate kinase mass action ratio, which is defined as ([ATP][AMP]) / [ADP]^2 is also an important measure used to assess the kinase activity in a model. An increase of glomerular ATP/AMP and ATP/ADP ratio can be seen in the diabetic kidney (Miyamoto *et al.,* 2016). Oxidative stress can be measured by deriving index defined as the ratio of glutathione in reduced and oxidized forms (Asensi *et al.*, 1999). The visualization of such biological indices based on MSI data can be quite informative for assessing the physiological status, but the lack of publicly available dedicated tools in this area has made it difficult to use MSI data for such applications.

Here we present MSIdV, a flexible tool to visualize biological indices using MSI data. Unlike conventional MSI visualization tools, users can specify multiple *m/z* values of interest and flexibly set calculation formulae to derive index values. Additional signal processing methods can be selected to improve the accuracy of index calculations and the quality of the visualization. MSIdV runs locally on standard windows workstations with the Python 2.7 environment and related libraries.

2   Implementation

MSIdV takes imzML-formatted files (Römpp *et al.*, 2011), a common data format for MSI, as input. First, the XML part of imzML file is read to acquire the offset and length of the mass spectrum data array (*m/z* and intensity) in the corresponding binary file, as well as the spatial infor-

mation in the X-Y direction of the sample surface. Second, the binary file of the imzML is scanned for the intensity values at user-defined *m/z* locations based on the offset and length information. Users can set up to six *m/z* values for the peaks of interest, together with window size. The intensity value for the peak of interest is acquired as the mean, median, highest, or integral of the intensity values within the user-defined *m/z* range. For instance, for *m/z* = 100.0 as the peak of interest, *m/z* 0.4 as the window size, and mean chosen as the output intensity value, the mean of the intensity values in the range between *m/z* = 99.8 and 100.2 is taken as the intensity value of the peak of interest. Median and highest intensity values within a *m/z* range are particularly useful to acquire accurate images in case there is a shift in the mass calibration over time because of long acquisition times.

Index values are calculated using the acquired intensity values and a user-defined formula configured by editing a template as shown below.

$$[(C_1 I_1)O_1(C_2 I_2)O_2(C_3 I_3)] \ O_3 \ [(C_4 I_4)O_4(C_5 I_5)O_5(C_6 I_6)]$$

where I_{1-6} are intensity values of the peaks of interest and C_{1-6} are user-defined coefficient values. Operator types (O_{1-5}) can be chosen from addition, subtraction, multiplication and division. Elements of the formula can be omitted to express a simple calculation.

Visualization of the calculated index is achieved by allocating index values to a two-dimensional matrix based on the spatial information acquired from the XML metafile in the imzML file.

To enhance the quality of index visualization, intensity and index values are processed further in the following ways. Index values often artificially take extreme values because the intensity value of one of the elements in a formula is unusually low or zero, because of the low quality of the mass spectrum. In such cases the pixel becomes a "hot spot" in the visualization because the index value is extremely high, thereby disrupting the dynamic range of the visualization image. To avoid this, lower and upper threshold values can be set to cut off outlier intensity or index values so that only reasonable values that fall between the thresholds are used for index visualization. Alternatively, a contrast stretching method can be applied, where the index value distribution is rescaled to include all values that fall within user-defined percentiles. Further, histogram equalization can be applied, where index values are rescaled so they are equally distributed on the histogram. The detail of histogram equalization is described in the supplemental information. These processing algorithms effectively enhance the contrast of the index image. Users can choose the signal processing method from among these options, or they can choose to apply both methods in a series. MSIdV generates the visualized index data as an image file (TIF file). It also provides raw data in the form of a two-dimensional data matrix in csv format.

3 Example usage

To demonstrate the function of MSIdV, we visualized the energy charge index of a mouse kidney. We sprayed 9-aminoacridine onto a sagittal section of mouse kidney and mass spectra were acquired using MALDI instrument in negative ionization mode. To visualize the energy charge index based on the MSI data, the mass spectral data was scanned for the mean intensity of the peaks of ATP, ADP, and AMP (*m/z* = 506.0, 426.0, and 346.1). The energy charge index was simply derived using the mean intensities of ATP, ADP, and AMP in the mass spectra. The analytical conditions and data processing are described in detail in the supplemental information. As shown in Figure 1, a higher energy charge state was observed in the renal cortex and, whereas a significantly lower index was observed in the renal medulla and pelvis.
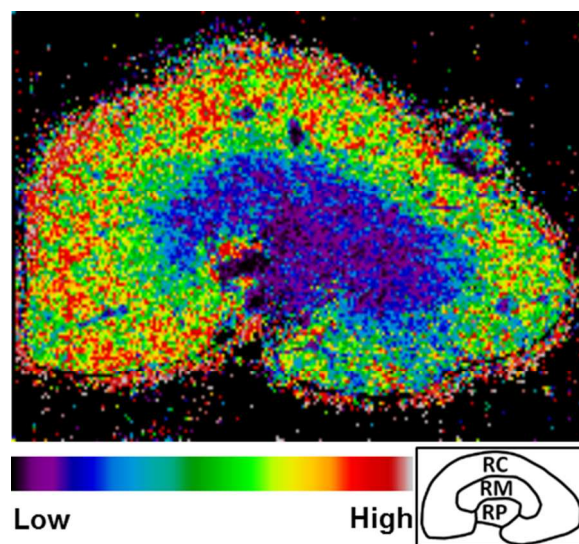
Fig. 1. Visualization of energy charge index in the mouse kidney. The colorbar shows the color coding of relative energy charge index. RC: renal cortex; RM: renal medulla; RP: renal pelvis.

4 Conclusions

MSIdV is a versatile tool for processing MSI data to flexibly and easily visualize biological indices, thereby facilitating a broad range of physiological studies based on MSI. Since it reads standard imzML-formatted data, users can build own pipelines by using external tools for signal preprocessing, such as baseline collection and alignment, to achieve even better visualization. This tool allows users to configure a formula to derive any index values of interest using a text-based easy-to-use configuration file, that requires no programming knowledge.

References

Asensi,M. et al. (1999) Ratio of reduced to oxidized glutathione as indicator of oxidative stress status and DNA damage. Methods Enzymol., 299, 267–76.

Atkinson,D.E. and Walton,G.M. (1967) Adenosine triphosphate conservation in metabolic regulation. Rat liver citrate cleavage enzyme. J. Biol. Chem., 242, 3239–41.

Caprioli,R.M. et al. (1997) Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. Anal. Chem., 69, 4751–60.

Miyamoto, S. et al. (2016) Mass Spectrometry Imaging Reveals Elevated Glomerular ATP/AMP in Diabetes/obesity and Identifies Sphingomyelin as a Possible Mediator. EBioMedicine, 7, 12-34

Rohner,T.C. et al. (2005) MALDI mass spectrometric imaging of biological tissue sections. Mech. Ageing Dev., 126, 177–85.

Römpp,A. et al. (2011) imzML: Imaging Mass Spectrometry Markup Language: A common data format for mass spectrometry imaging. Methods Mol. Biol., 696, 205–24.

Takáts,Z. et al. (2004) Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. Science, 306, 471–3.

Watrous,J.D. and Dorrestein,P.C. (2011) Imaging mass spectrometry in microbiology. Nat. Rev. Microbiol., 9, 683–94..