SCiLS Lab Software

SCiLS Lab 2D: Spatial segmentation with edge-preserving image denoising

MALDI-imaging mass spectrometry is an emerging spatially-resolved mass spectrometric technique which can obtain spatial distribution of hundreds of molecules in a thin tissue section directly from the sample surface. Presently, data mining of a MALDI-imaging dataset, which consists of thousands of individual mass-to-charge (m/z) values, is mostly done manually and this represents a serious bottleneck in the data analysis workflow. In the SCiLS Lab 2D software, automatic spatial segmentation can be used as a first step of data mining, providing an overview of the dataset and allowing quick detection of prominent features.

AN INFORMATION CHALLENGE

A state-of-the-art MALDI-imaging dataset comprises a huge amount of spectra (typically 10,000 to 100,000 spectra), where an individual spectrum represents intensities measured at 10,000-25,000 m/z-bins (for ToF-MS) or at 1,000,000 or more m/z-bins for FT-ICR-MS measurements (FIG. 1). Understanding and interpreting such a tremendous dataset requires computational data mining strategies.

The SCiLS Lab 2D software implements computational methods for mining large MALDI-imaging data, in particular using spatial segmentation. In this approach applied to MALDI-imaging data, similarities of spectra is statistically determined, and similar spectra are grouped into one cluster. All spectra of a particular cluster are then assigned a selected color and displayed as a spatial segmentation map in which all pixels are color-coded according to their cluster assignment. Interactive exploration of the data can be done by browsing the hierarchical clustering dendrogram[1].

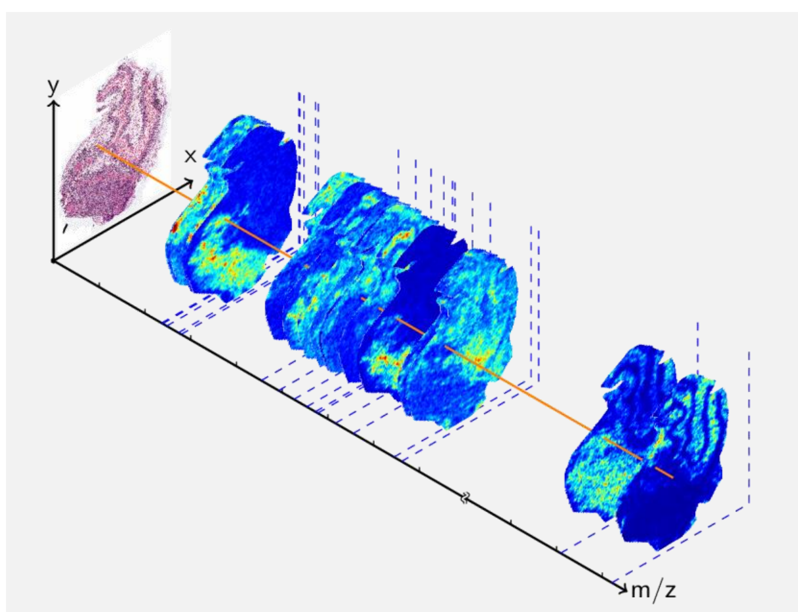

FIG. 1 MALDI-imaging dataset of a neuroendocrine tumor section shown as a hyperspectral image. A typical dataset consists of billions of measured intensities values and a manual evaluation is infeasible.

SPATIAL SEGMENTATION PIPELINE

The following pipeline for analysis and interpretation of MALDI-imaging data using spatial segmentation is implemented in SCiLS Lab 2D. First, a peak picking algorithm is applied. Considering both peak intensity and peak shape, the aim of the algorithm employed is to select only informative peaks and discard m/z-values which represent noise or baseline[2].

Second, for suppressing spectrum-to-spectrum (i.e. pixel-to-pixel) variation, edge-preserving image denoising is applied to the m/z-images of all selected peaks (FIG. 2). There is a marked increase in the quality of the resulting segmentation map if advanced image denoising is applied prior to the clustering analysis[3].

The final step of our segmentation procedure is to cluster all preprocessed spectra consisting of denoised intensities of all selected peaks. An efficient hierarchical clustering algorithm performs clustering for any arbitrary number of spectra in a reasonable time on a workstation (a handful of minutes for the dataset reported in this note).

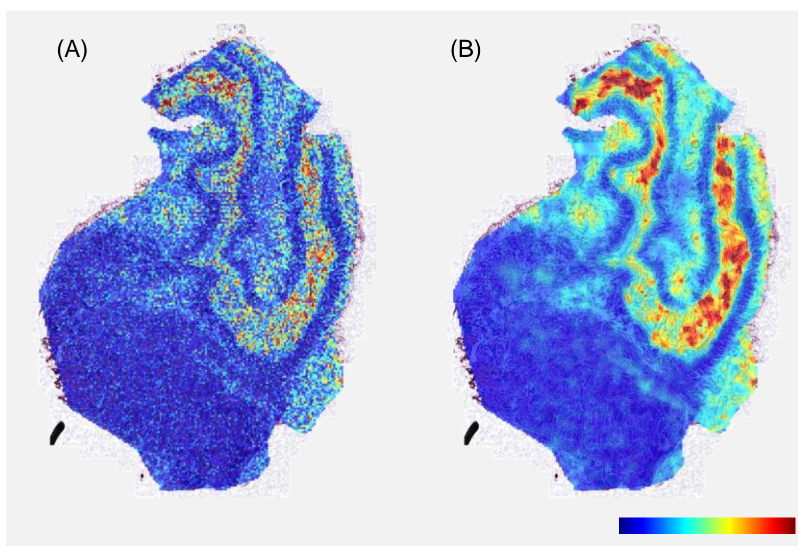

FIG. 2 Spatial distribution of m/z-value 7,159 in the neuroendocrine tumor section without (A) and with (B) edge-preserving image denoising applied. Visualized in SCiLS Lab 2D after the automatic hotspot removal display applied.

DISCLAMER: SCiLS Lab 2D is a part of Bruker Daltonik's MALDI Molecular Imager solution. Only MALDI-imaging data obtained using flexImaging (Bruker Daltonik) can be analyzed in SCiLS Lab 2D by means of multivariate analysis.

EVALUATION OF RESULTS

Once a segmentation map is generated, it can be interpreted to detect prominent spatial regions. Clustering results can be exported into the flexImaging software (Bruker Daltonik) and superimposed with high-resolution microscopy images. Similar to flexImaging, SCiLS Lab 2D provide dendrogram visualization of the clustering results, in which major clusters are displayed in a tree-like view which can be interactively explored.

In addition, SCiLS Lab 2D offers an intuitive operational concept allowing the user to split or merge a specific cluster with a single mouse click, effectively navigating along the hierarchical dendrogram in a graphical way.

NEUROENDOCRINE TUMOR DATA

We applied the procedure for spatial segmentation available in SCiLS Lab 2D to a section of a neuroendocrine tumor specimen. Tissue preparation and MALDI-imaging data acquisition have been made according to standard protocols[3]. Altogether 27,361 spectra have been acquired with a lateral resolution of 50 µm, each spectrum covering the range 3,200-18,000 m/z.

FIG. 3 (A)–(C) show segmentation maps with two, three and four clusters, the hierarchical between clusters is shown in the dendrogram in FIG. 3 (D). The segmentation maps coincide with the histological annotation visualized in FIG. 3 (E).

After segmentation, we applied correlation analysis to detect m/z-values co-localized with 2 clusters of the segmentation map. For both clusters we found highly co-localized m/z-images with a correlation higher than 0.5 and p-values of $p \leq 0.05$ (FIG. 4) .

REFERENCES

1. Deininger et al. *J Proteome Res* 2008, 7(12):5230–5236.
2. Alexandrov et al. *J Proteomics* 2011, 75(1):237–245.
3. Alexandrov et al. *J Proteome Res* 2010, 9(12):6535–6546.

SUMMARY

- Spatial segmentation provides data overview and is able to extract prominent features.
- An efficient algorithm for spatial segmentation is implemented in SCiLS Lab 2D allowing to analyze arbitrarily many spectra.
- A hierarchical clustering dendrogram allows one to explore the data interactively.
- For specific regions the co-localized m/z-values can be calculated.

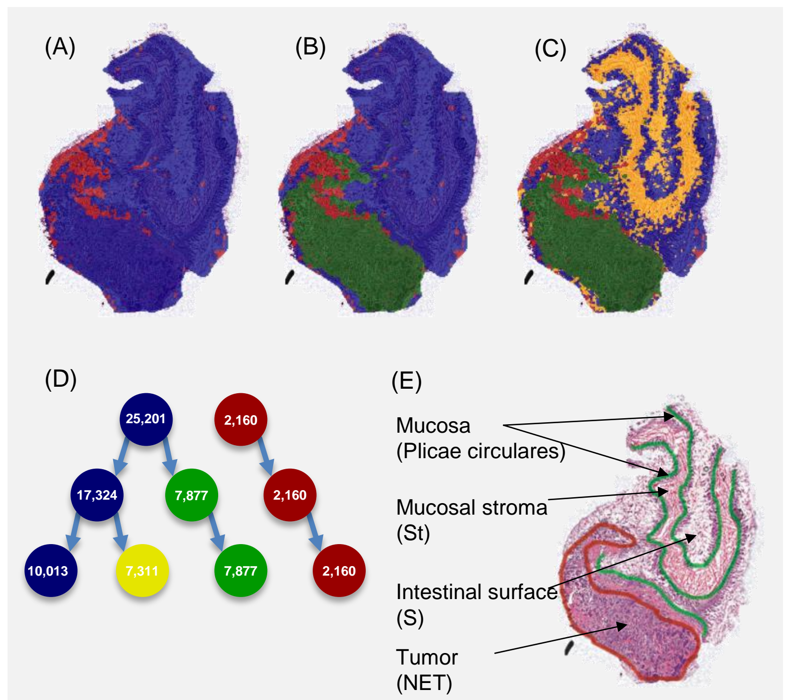

FIG. 3 The segmentation map for two (A), three (B), and four (C) clusters of the neuroendocrine tumor dataset, visualized in SCiLS Lab 2D. The hierarchical relations between clusters are illustrated with the dendrogram (D). The histological annotation is shown in (E).

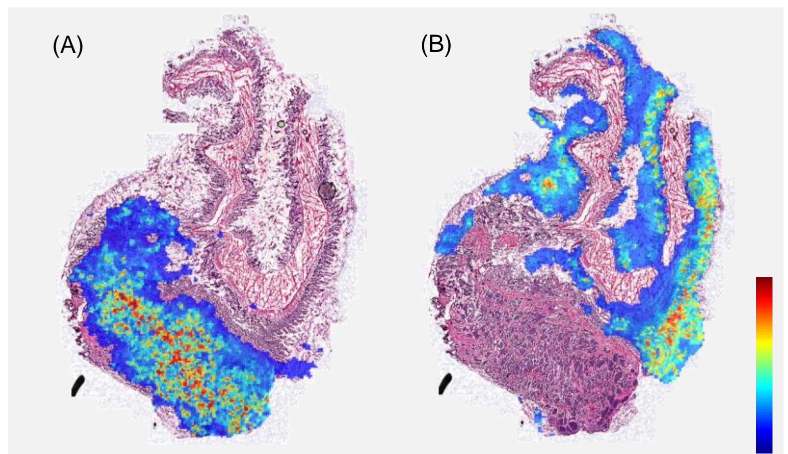

FIG. 4 Individual molecules highly co-localized with green (m/z 7,547; A) and blue (m/z 6,724; B) cluster in FIG. 3 (C). Correlation values are 0.61 and 0.56, respectively. Low intensities of the m/z-image were set transparent. Visualization was done in SCiLS Lab 2D with edge-preserving image denoising and automatic hotspot removal applied.

KEYWORDS

MALDI-imaging, data analysis, spatial segmentation, edge-preserving denoising, neuroendocrine tumor invading the small intestine

AUTHORS

D. Trede[1], S. Schiffler[1], M. Becker[2], G. Ernst[3], K. Steinhorst[1], J.H. Kobarg[1], A. Dyatlov[1], J. Oetjen[1], A. Fuetterer[2], H. Thiele[1], F. von Eggeling[3], P. Maass[1], T. Alexandrov[1]; [1]SCiLS, Bremen; [2]Bruker Daltonik, Bremen; [3]University Hospital Jena, Jena.

SCiLS BRUKER

SCiLS GmbH
Dr. Dennis Trede
Fahrenheitstr. 1
28359 Bremen
Germany

T: +49 (421) 1611 0863
E: contact@scils.de
W: www.scils.de