Software workflow and statistical analysis tools for evaluating multiomics MALDI MSI studies

Connor West[1], Tobias Boskamp[2], Sören Deininger[2], Mark Lim[3], Gargey Yagnik[3], Natalie Agar[4], Sylwia Stopka[4], Rick Drake[5], Stacy Malaker[6], Rachel Stubler[5], Kenneth Rothschild[3,7], Katherine Stumpo[1]

[1]Bruker Daltonics, Billerica, Massachusetts, USA
[2]Bruker Daltonics GmbH & Co. KG, Bremen, Germany
[3]AmberGen, Inc., Billerica, Massachusetts, USA
[4]Harvard Medical School. Boston Massachusetts, USA
[5]The Medical University of South Carolina, Charleston, South Carolina, USA
[6]Yale University, Dept. of Chemistry, New Haven, Connecticut, USA
[7]Boston University, Boston, Massachusetts, USA

Introduction

MALDI MSI is used for targeted and untargeted spatial profiling in lipidomics, metabolomics, proteomics, and glycomics.

Multiple acquisitions from the same tissue sample using different protocols are possible, enabling multiomics studies combining data from multiple imaging omic regimes.

Multiomics data analysis requires analyzing spatial correlations between the distributions of molecules across the omic regimes.

We present a unified data fusion workflow for multiomics MSI data analysis, allowing to apply state-of-the-art visualization tools and computational analysis algorithms.

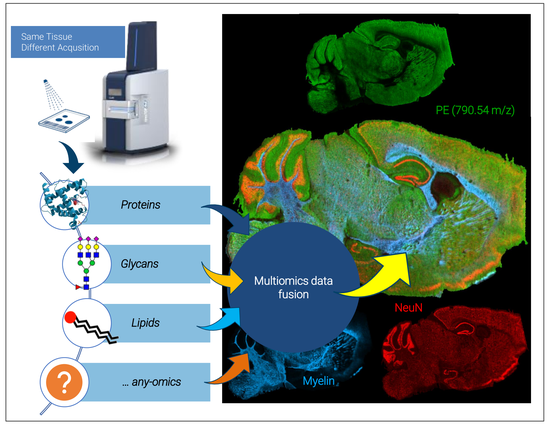

Figure 1. Multiomics data fusion concept. Starting with different datasets obtained from various MALDI-imaging acquisition methods, ion image data is co-registered together within a fused dataset, allowing direct statistical analysis.

Methods

MSI data was acquired using different protein, lipid, and glycan imaging protocols. Multiomics data analysis workflow consists of:

a) importing each dataset into the SCiLS Lab data format
b) co-registering datasets based on optical or ion images
c) extracting spectral features from one dataset and mapping them to the other dataset's spot raster
d) applying standard visualization and statistical analysis tools for joint multiomics evaluation

In addition, a scripting API makes the multiomics data available for customized analysis or reporting workflows, or for integration with third party software.

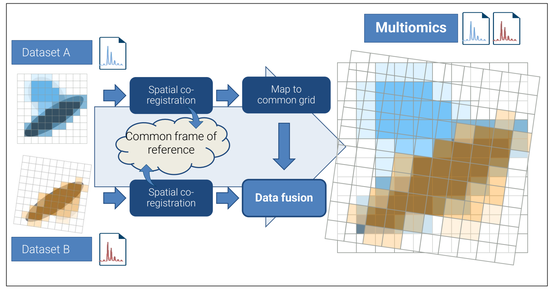

Figure 2. Data fusion workflow. Co-registering dataset A onto dataset B using a common frame of reference, such as fiduciary markings or anatomical landmarks. Data fusion is achieved using the SCiLS Ion Image Mapper tool by mapping ion images from dataset A onto dataset B as separate distinct features.

Results

Lipidomics and MALDI HiPLEX-IHC imaging in mouse brain

Acquisition A: Targeted (Myelin) protein imaging using photocleavable mass tagged antibodies (MALDI HiPLEX-IHC)

Acquisition B: Untargeted lipid imaging, reflector TOF, negative ion mode, 20 µm lateral resolution. Analysis with SCiLS Lab feature finding and principal component analysis (PCA)

Multiomics analysis: SCiLS Lab feature finding and principal component analysis (PCA) on lipid MSI and co-registration of lipid PCA with selected HiPLEX ion image.

SCiLS Ion Image Mapper to map HiPLEX-IHC ion images onto the lipid imaging dataset. Spatial correlation analysis reveals lipid species with distributions similar or contrasting to protein distribution patterns (Fig. 3).

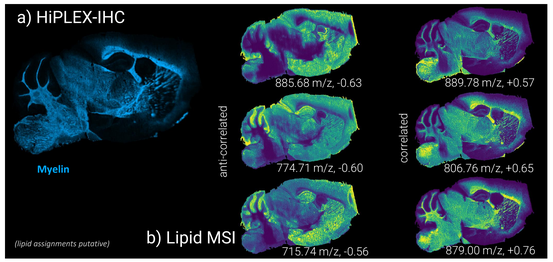

Figure 3. Co-localization analysis showing a) myelin spatial distribution from HiPLEX-IHC and b) (anti-) correlated lipid features selected based on Pearson correlation.

Glioblastoma mouse xenograft lipid and HiPLEX-IHC imaging

Acquisitions A and B as in previous example.

Multiomics analysis: Co-registration of selected ion images from lipid and HiPLEX-IHC using SCiLS API and external image registration. SCiLS Ion Image Mapper used to map HiPLEX-IHC ion images onto lipid imaging dataset. Spatial correlation analysis to detect lipid species correlated with GFAP antibody distribution (Fig. 4).

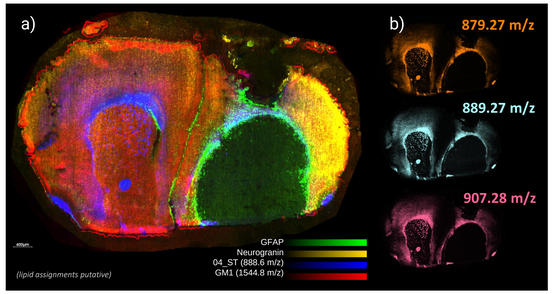

Figure 4. a) Combined visualization of glial fibrillary acidic protein (GFAP), neurogranin, 04 ST and GM1 with b) lipid features correlated to GFAP.

Human colorectal carcinoma multiglycomics analysis

Acquisition A: PNGase-F digested samples, N-linked glycans detected in reflector TOF, positive ion mode, 40 µm lateral resolution.

Acquisition B: Mucin-selective protease StcE digested samples, O-glycopeptide imaging

Multiomics analysis: Co-registration based on H&E stain optical image, data fusion of O-glycopeptides and N-linked glycan ion images using SCiLS Ion Image Mapper (Fig. 5).

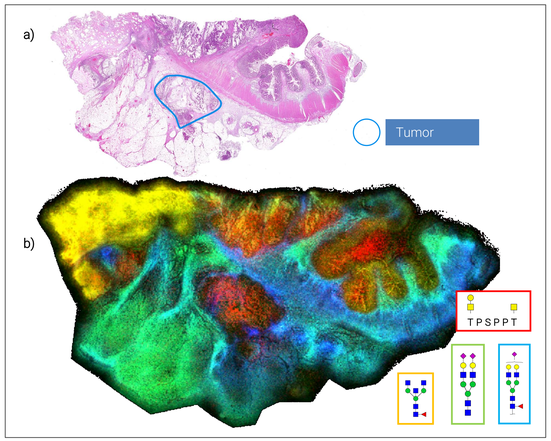

Figure 5. a) H&E stain of a human colorectal carcinoma tissue with the tumor annotated. b) Color overlay showing spatial distributions of selected N-linked glycans (yellow, green, blue) and an O-glycopeptide (red).

Conclusion

- Multiple datasets from various omics methods can be integrated into a single dataset for computational analysis and visualization
- Datasets of multiple acquisitions (targeted or untargeted) from the same tissue can be fused
- Manual co-registration based on optical or ion images. SCiLS API allows for more specialized or automated co-registration methods.
- SCiLS Ion Image Mapper performs resampling and mapping of MALDI imaging datasets to a common spot grid.