A TECHNOLOGY-NEUTRAL SOFTWARE PACKAGE FOR INTEGRATED SPATIAL MULTI OMICS VISUALIZATION AND DATA ANALYSIS

Nathan Heath Patterson[1], Wanqiu Zhang[1], Maria Jose Q. Mantas[1], Phuong Thao Tran[1], Reza R Choubeh[1], Baruch Berger[1], Fang Xie[2], M. Reid Groseclose[2], Pedro Machado Almeida[3], Alice Ly[1], Peter Roelants[1], Thomas Moerman[1], Nico Verbeech[1], Marc Claesen[1]

[1] Aspect Analytics NV, C-mine 12, 3600 Genk, Belgium
[2] GSK, Collegeville, PA
[3] Lunaphore Technologies SA,, Tolochenaz, Switzerland

Introduction

Advances in spatial omics technologies now enable detailed molecular analyses directly from tissue. In parallel, combining different multi-omics approaches e.g. Mass Spectrometry Imaging (MSI) for lipids with spatial proteomics approaches such as Imaging Mass Cytometry (IMC) or multiplexed immunofluorescence (mIF), is increasingly in demand. Conducting integrated spatial-multi-omics data analysis is a daunting bioinformatics task as data is typically acquired on serial sections, at different spatial resolutions, in a variety of data formats, requiring expertise across multiple domains to obtain the best results from each individual technology. We present Weave®, a single view, collaborative spatial multi-omics bioinformatics software platform that enables efficient integration, joint visualization, and management of data from different spatial assays.

Experimental Data Collection

Lymph node dataset: Serial FFPE mouse lymph node sections (4 µm) were deparaffinized, underwent antigen retrieval (40 mins), and then incubated in a cocktail of 16 antibodies that were either conjugated to metal probes for IMC or to photocleavable Miralys probes (AmberGen) overnight at 4°C. The section for MALDI-IHC was incubated in a UV light box (10 min), coated in matrix, then measured on a rapifleX (Bruker) at 10µm pixel size in positive mode. The IMC-prepared section had the ROIs selected and imaged on a Hyperion system (Standard Biotools). Cell segmentation of IMC data was conducted using Mesmer[1].

Mouse brain dataset: Serial mouse brain cryosections (10 µm) were mounted on SuperFrost Plus Gold slides (Fisher Scientific) for mIF or ITO-coated slides (Bruker) for MSI. Multiplexed sequential immunofluorescence (seqIF™) was performed on a COMET (Lunaphore), using a 14-marker antibody panel. MALDI MSI for lipids was performed on a timsTOF fleX (Bruker) in positive mode at 20 µm spatial resolution after coating the section with norharmane using an M5 sprayer (HTX Technologies). The mIF- and MSI-measured sections were H&E stained and digitised after their respective measurements. Data-driven cell segmentation was conducted using UnMICST[2]. Morphological feature extraction of the H&E staining from the mIF section was conducted according to Zhang et al.[3].

PART 1

Leveraging cloud computing to manage spatial multi-omics data and collaborative projects

(A) The core of the Weave platform is a central hub for dataset management, co-registration and coordination of collaborative efforts in spatial biology studies. Datasets are uploaded to the platform (e.g. MSI with mIF and relevant histological images), along with metadata (e.g. which sections were used for what technique, what data analysis was conducted etc). As the platform is cloud-based, data storage and computing requirements are scalable to study needs, and accessible via browsers, and thus not dependent on specific operating systems. This allows sharing of data, analysis results and reports between project partners via a simple URL. The platform provides snapshotting and commenting functionality, facilitating collaborative analysis across multidisciplinary teams. While data analysis tools are provided in Weave, bioinformatics analysis of spatial data is a highly dynamic field; APIs and SDKs are therefore available to integrate with other workflows. The platform is certified for the ISO/IEC 27001 standard for information security.

(B) In this example of a project Dashboard, users can see their current "Projects", "Inventory", "Tools", "User Groups" and "Access".

PART 2

Spatial data registration and integration

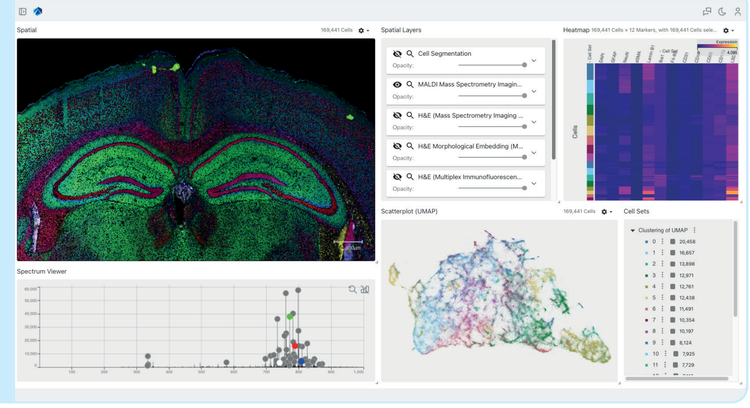
Our approach for spatial multi-omics data analysis consists of multiple steps. Prior to data analysis, metadata labeling of individual section data informs the creation of stacks and helps identify sequences of tissue section's where stack creation is feasible. Accurate, non-rigid image registration (manual or automatic, where appropriate) is first used to create a single coordinate system across serial sections and their measurements. When using the Weave co-registration tool **(A)**, a minimum of three points are required to establish the initial coordinate system and generates a view of the overlaid datasets. Following this, as many points as required can be placed to satisfactorily align the datasets, with live feedback provided to users. For data that are generated from the same section, this approach can be accurate to the single-cell level **(B-D)**.

After data registration, several data integration approaches are provided to create an integrated data structure across assays of different spatial resolutions. For example, applying Gaussian weighting to account for differences in spatial resolution when fusing ST (55 µm/px), MSI (20 µm/px) and mIF (<0.5 µm/px) datasets. Once coordinated, this data structure acts as a foundation for downstream joint data analysis and visualization.

PART 3

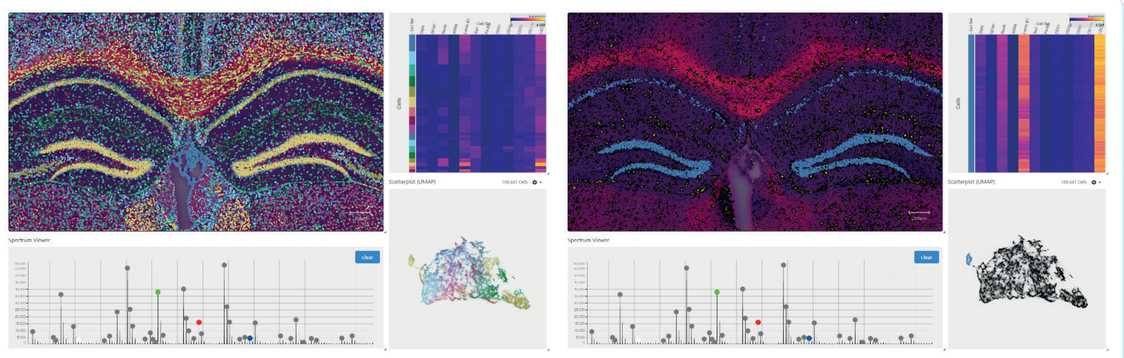
Joint visualization of integrated data

This shows the starting view in Weave of a report showing the integration of mouse brain sections measured with mIF and lipid MSI, the respective H&E images from each section, and data analysis results (cell segmentation and morphological feature extraction). In this example, the Spatial panel shows all the different datasets and analysis results overlaid in a single view. In the Spatial Layers panel, visualization and control of the different modalities is via individual dropdown menus. The Spatial panel and relevant data plots (e.g. mass spectrum for the MSI dataset, scatterplot for mIF cell segmentation) can be interactively controlled, allowing for zooming, panning, and selection of interesting features.



MSI dataset control

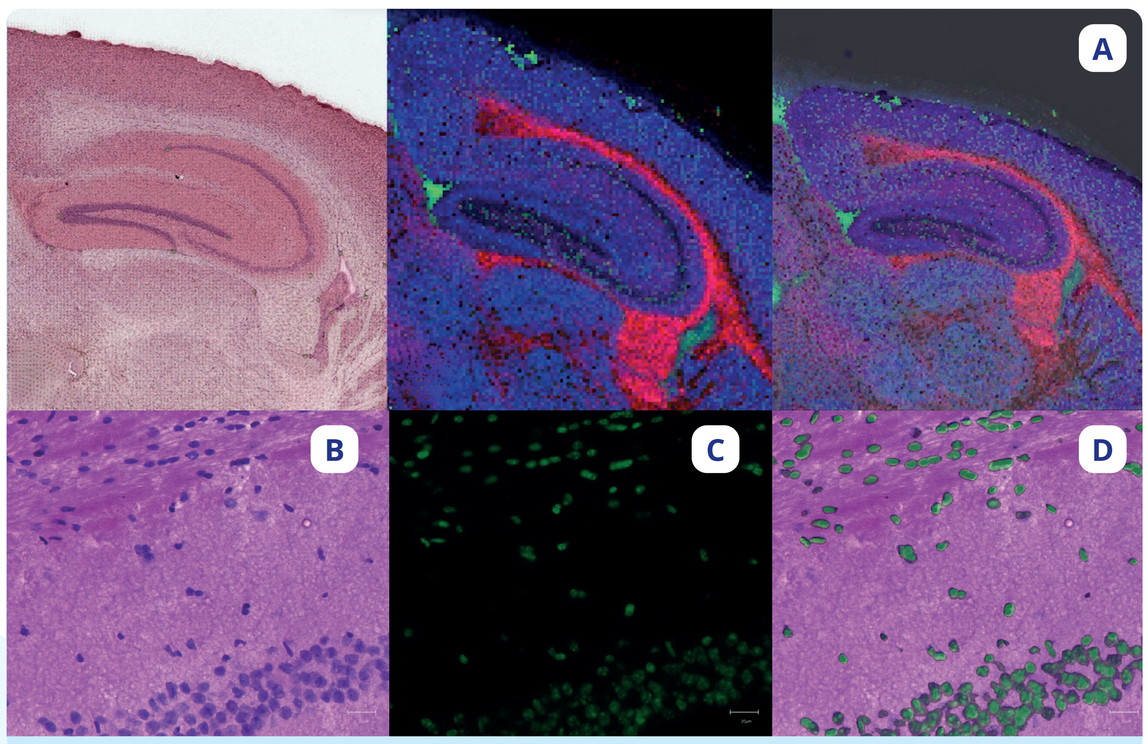
This view shows the MSI data layer and its respective dropdown menu in the Spatial Layers panel. This menu can be collapsed or expanded by clicking on the arrow ("^"/"˅") on the top right corner. Visualization of the layer in the Spatial panel is turned on/off with the eye icon; if multiple layers are turned on, selection of the magnifying glass icon turns all other layers off. m/z's for visualization can be selected via the drown down menu or from the mass spectrum panel. The software provides a variety of visualization options, such as colormap selection, control of color intensity range and layer opacity.

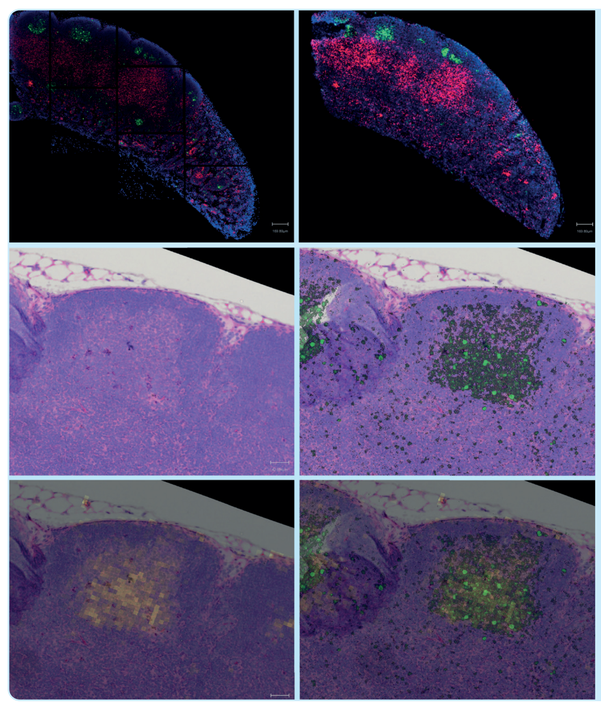


PART 4

Interactive co-visualization of multiple spatial-omics datasets

This figure shows the mIF cell segmentation result overlaid onto MSI showing m/z 788.614 and H&E. The Scatterplot (UMAP) shows a data driven clustering of the cells based on the mIF data. On the left, an outlier population in the cells can be found. A lasso was drawn over that population in the scatterplot showing the where the population corresponds to in the accompanying heatmap and its spatial localization in the sample. The cell segmentation results of the mIF can subsequently be used to extract lipid signals in the MSI through correlation and differential analysis.





PART 5

Application - Comparison of IMC with MALDI-IHC on lymph node

IMC and MALDI-IHC are two highly-multiplexed spatial proteomic techniques utilising antibody labelling and detection via mass spectrometry. In this application, researchers wanted to investigate if MALDI-IHC could be used as an initial high speed, whole section screening method before using IMC for high spatial resolution analysis. The top two panels shows a side-by-side comparison of IMC (left) and MALDI-IHC (right) results with labelling for Ki-67 (green), CD8 (red) and nuclei (blue) in a mouse lymph node. The IMC-measured section was imaged over 8 ROIs which were then stitched together in order to visualize the whole section in our software. The lower panels show the distribution of the proliferative marker Ki-67 from IMC (green) and MALDI-IHC (yellow) overlaid onto the H&E-stained MALDI-measured section, individually and simultaneouly. In both methods, the Ki-67 signal corresponds to lymph node germinal centers.

Conclusion

- Weave® software can be used for combined data analysis of datasets generated with different technologies by a large number of different vendors
- Accurate, non-rigid spline-based landmark registration is fundamental for integrating datasets, especially when conducted from serial sections
- Interactive viewer for combined visualization of datasets and data analysis elements to aid exploration and biological understanding

References: [1] Greenwald NF, et al. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. Nat Biotechnol. 2022. [2] Yapp C, et al., UnMICST: Deep learning with real augmentation for robust segmentation of highly multiplexed images of human tissues. Commun Biol. 2022. [3] Zhang W, et al. Multimodal MALDI imaging mass spectrometry for improved diagnosis of melanoma. medRxiv 2022. doi: https://doi.org/10.1101/2022.11.29.22282868

AGENTSCHAP INNOVEREN & ONDERNEMEN Vlaanderen is ondernemen · Horizon 2020 · GSK · Lunaphore a biotechne brand · FLANDERS INVESTMENT & TRADE Flanders State of the Art