WEAVE: A SOFTWARE PACKAGE FOR INTEGRATED SPATIAL MULTIOMICS VISUALIZATION AND DATA ANALYSIS

Wanqiu Zhang[1], Thao Tran[1], Felicia Wee[2], Craig Joseph[2], Li Yen Chong[2], Zhen Wei Neo[2], Joe Yeong[2,3], Alice Ly[1], Marc Claesen[1], Heath Patterson[1]

[1] *Aspect Analytics NV, Genk, Limburg, Belgium*
[2] *A*STAR Institute of Molecular and Cell Biology (IMCB), Singapore, Singapore*
[3] *Singapore General Hospital, Singapore, Singapore*

Introduction

There has been enormous progress in single-omics spatial technologies, which has revolutionized our understanding of the tumor micro-environment (TME). While the data generated from these methods have helped to unravel cellular intricacies within spatial context at the genomic, transcriptomic, metabolomic and protein levels, users are increasingly combining different omics readouts to obtain a holistic view of TME heterogeneity and complexity. However, spatial multi-omics data analysis presents specific bioinformatics challenges, as the data is typically acquired at different spatial resolutions, using a variety of platforms, and generates large data volumes. We present Weave, a cloud-based software for spatial omics bioinformatics, enabling efficient integration and joint visualization of different spatial-omics assays.

Methods:

Human lung cancer sections were sequentially analyzed with spatial transcriptomics using a cancer panel targeting 289 genes (Xenium, 10X Genomics), followed by multiplexed immunofluorescence using a 40-antibody panel (COMET, Lunaphore), and then H&E staining. The H&E images were digitized (Axioscan 7, Zeiss), and pathology annotation performed in QuPath. Cell segmentation was performed on the Xenium dataset using DAPI-based nuclear expansion (10X Genomics), and on the COMET dataset using CellSAM. All data were co-registered at full resolution using a non-rigid spline-based algorithm, then visualized in a web-based viewer (Weave, Aspect Analytics).

Results

Reading direction

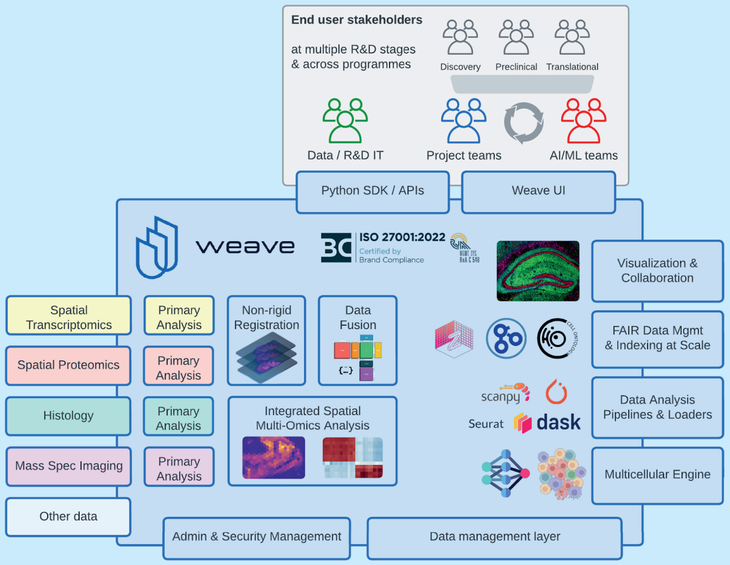

Fig. 1: Leveraging cloud computing to manage spatial multiomics data and collaborative projects. The core of the Weave platform is a central hub for dataset management, non-rigid co-registration, fusion between different spatial biology assays, and co-ordination of collaborative efforts. The platform supports a wide variety of commercially available spatial biology assays. Datasets are uploaded (e.g. multiplexed immunofluorescent spatial proteomics and relevant histological images), along with relevant metadata (e.g. which sections were used for what assay). After co-registration and data fusion, users can conduct integrated data analysis, such as differential expression of transcripts and proteins in pathologist-identified regions of interest.

As Weave is cloud-based, data storage and computing requirements are scalable to study needs and the UI 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. While data analysis tools are provided in Weave, bioinformatics analysis of spatial data is a highly dynamic field; SDKs and APIs are therefore available to integrate with other workflows. The platform is certified for the ISO/IEC 27001 standard for information security.

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 sections where stack creation is feasible. Accurate, non-rigid image registration is then used to create a single coordinate system across different measurements that can account for different sections or variations in measurement region size. For data that are generated from the same section, Weave's co- registration tool can be accurate to the single-cell level (Figure 2).

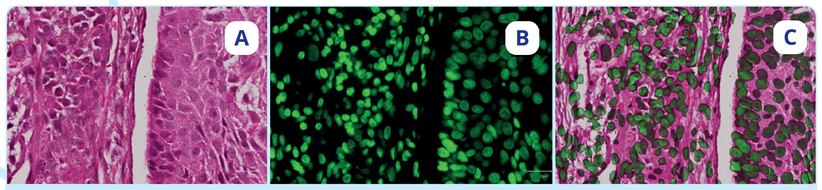

Fig. 2: Example of co-registration accuracy **A:** H&E of lung cancer biopsy **B:** DAPI channel from COMET dataset pseudo colored in green/DAPI channel from Xenium dataset pseudo colored in yellow. **C:** Overlay of the two images. Scale bar is 20 µm.

Joint visualization of integrated spatial multi-omics datasets in a single software

Communication and visualization of data is via web-based Weave reports, which can be shared via URLs. Figure 3 shows an example of Weave reports displaying the integrated Xenium, COMET and H&E datasets for lung cancer samples, and data analysis results (cell segmentation, clustering results). The Xenium and COMET datasets with their respective H&E images and data analysis results are overlaid in a single view in the Spatial panel. Control of the visualization of the different modalities is via individual dropdown menus in the Spatial Layers. Xenium results can be either be visualized as cells or cell contours that expressed a gene via the Gene List panel or as transcript spots or density maps via the Transcript spot panel. The Spatial panel and relevant data plots (e.g. UMAP scatterplot) can be interactively controlled, allowing for zooming, panning, and selection of interesting features.

Fig. 3: Example of a Weave report, allowing joint visualization of the Xenium and COMET datasets, H&E image, pathologist annotations (grey mask), and associated data analysis results in a single view. Users are able to interactively explore, examine and share results.

Correlation analysis of gene transcript and protein expression patterns

From the 289 genes targeted with the Xenium, 16 corresponding proteins were targeted using COMET. We conducted correlation analysis to identify which transcript-protein pairs have similar spatial expression, and if this correlation was affected by cell segmentation approach. Table 1 shows the results for the transcript- protein pairs. Some pairs had high correlation, regardless of cell segmentation (e.g. CDH1/E-cadherin), while correlation of a protein markers which refers protein complexes that derive from different genes yield variable result (e.g. CD3).

xenium	comet	correlation (comet segmentation)	correlation (xenium segmentation)
MS4A1	CD20	0.53	0.50
FOXP3	FoxP3	0.52	0.45
CD3D + CD3E	CD3	0.51	0.44
CDH1	E-Cad	0.50	0.62
CD34	CD34	0.42	0.44
CD3E	CD3	0.41	0.35
CD3D	CD3	0.41	0.34
CD14	CD14	0.40	0.44
CD8A	CD8	0.39	0.34
CD8A + CD8B	CD8	0.38	0.33
FCGR3A	CD16	0.38	0.47
CD68	CD68	0.37	0.42
KRT15 + KRT7	CK	0.34	0.38
KRT15	CK	0.34	0.44
MKI67	Ki67	0.33	0.33
CD38	CD38	0.31	0.30
CD163	CD163	0.31	0.37

Table 1: Correlation of gene transcript-protein expression when using COMET-based cell segmentation and Xenium-based cell segmentation.

Fig. 4: Joint visualization of transcript-protein pairs with different cell segmentation.
A: CDH1 transcripts (red spots) and corresponding protein E-cadherin (yellow) had high spatial correlation, regardless of whether cell segmentation approach was via Xenium (green contours) or COMET (blue contours).
B: Protein complexes, e.g. CD3 (white) have more variable correlations with respective genes CD3D (fuschia spots) and CDE (yellow spots), regardless of cell segmentation approach. Scale bar is 20 µm.

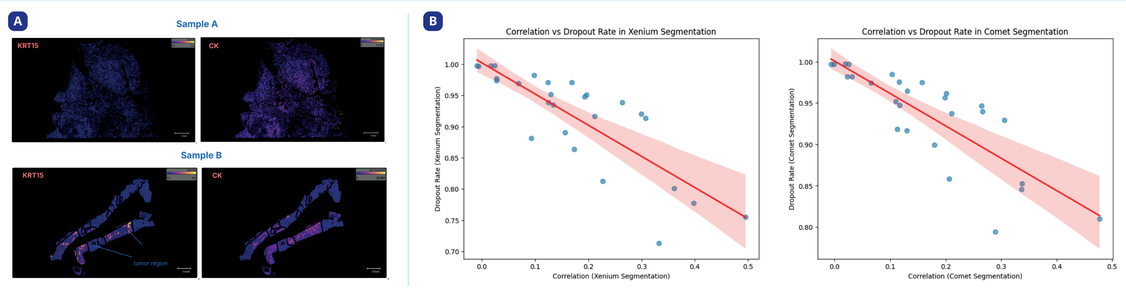

Fig. 5: A: Example of a transcript-protein pair with different spatial correlation across the two samples. KRT15-Cytokeratin had low correlation in sample A but high correlation in sample B. This may be due to the pan-cytokeratin antibody detecting more proteins than the products of KRT7 and KRT15 from the Xenium panel.

B: For both cell segmentations, gene transcripts with high dropout rates—the proportion of cells in which transcripts were undetected—tended to have lower transcript-protein correlations.

References
[1] https://www.10xgenomics.com/support/software/xenium-onboard-analysis/latest/algorithms- overview/segmentation#seg-nucleus-expansion
[2] Israel, Uriah, et al. "A foundation model for cell segmentation." bioRxiv (2023).
[3] https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.pearsonr.html

AGENTSCHAP INNOVEREN & ONDERNEMEN Vlaanderen is ondernemen
Horizon
FLANDERS INVESTMENT & TRADE Flanders State of the Art

Conclusion & Outlook

- As spatial omics is increasingly used to investigate TIME biology, Weave® software addresses the need for spatial omics and spatial multi-omics bioinformatics solutions.
- Overlays of multiple spatial omics datasets from different technologies and vendors can be integrated, enabling direct visual
- Comparison with interactive browsing of full resolution images and data analysis results.