spammR: an R package designed for analysis and integration of spatial multi-omic measurements

Yannick Mahlich[1], Harkirat Sohi[1], Marija Velickovic[2], Paul Piehowski[2], Jason E McDermott[1,3], Sara J Gosline[1, 4]

1 Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA;
2 Environmental and Molecular Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA;
3 Department of Molecular Microbiology and Immunology, Oregon Health & Sciences University, Portland, Oregon 97201 USA; 4. Department of Biomedical Engineering, Oregon Healthy & Sciences University, Portland, OR 97201, USA

Abstract

Summary

Spatial omics is a young and evolving field and as such shows rapid development of novel technologies and analysis methods to measure transcripts, proteins, metabolites, and post-translational modifications at high spatial resolution. These advances in technology have enabled the simultaneous generation of abundance profiles for multiple different omics types and associated microscopy imaging data, as well as their analysis in a spatial context. However, most analytical tools are designed for spatial transcriptomics platforms and are challenging to use in other contexts such as mass spectrometry-based measurements or metagenomics.

To this end we present spammR (spatial analysis of multi-omics measurements in R), an R package that enables end-to-end analysis with a specific focus on mass-spectrometry derived spatial omics datasets with the goal of integration across multiple data types (e.g. sequencing, metabolites, proteins) within the same tissue sample.

Availability and implementation

spammR is implemented in R. The package is currently installable from GitHub (https://github.com/PNNL-CompBio/spammR).

Contact

Sara Gosline (sara.gosline@pnnl.gov)

Supplementary Information

Supplementary Methods online

1. Introduction

The field of spatial biology (Crosetto et al. 2015; Moses and Pachter 2022; Vandereyken et al. 2023) and the growth of spatially-resolved technologies (Bouwman et al. 2022; Moffitt et al. 2022; Lundberg and Borner 2019; Taylor et al. 2021) has emerged from the need to study biological systems within their spatial context and native tissue environments (Kiessling and Kuppe 2024). Measuring the molecular changes of a sample, e.g. slices of tumor tissue, in its spatial context (Nichterwitz et al. 2018; Lee et al. 2014; Zhang et al. 2021), has been enabled by both novel transcriptomics (Eng et al. 2019; Rodriques et al. 2019; Codeluppi et al. 2018; Chen et al. 2015; Wang et al. 2018) as well as proteomics technologies (Radtke et al. 2022; Black et al. 2021; Angelo et al. 2014; Lin et al. 2018). Mass spectrometry (MS) based technologies to measure proteins (Zhu et al. 2018; Piehowski et al. 2020), lipids, metabolites, and other molecules (Buchberger et al. 2018; Greer et al. 2011; Veličković et al. 2024; 2025) have furthered enabled this field to study how molecular interactions can vary across a tissue.

While the pace of computational tool development has accelerated, particularly in the field of spatial transcriptomics (Walker et al. 2022), there is a gap in spatial analysis tools tailored towards the analysis and interpretation of mass spectrometry-based multi-omics data, with most existing tools focused on specific analytical frameworks (Nimo et al. 2025) that cannot be readily applied to integrative omics analyses – e.g. comparing metabolites and proteomics in the same sample. This is especially important regarding the distinct differences between antibody imaging and mass spectrometry-based measurement. For mass spectrometry specifically, those include (1) the need for normalization of mass/charge ratios, (2) the potential need to impute data to counteract missingness and, (3) MS as a technology not being restricted to proteomics. In addition to MS-based spatial proteomics, the rise of spatial glycomics, metabolomics, lipidomics and measurements of post translational modifications (PTMs), such as phosphorylation, all necessitate a more flexible framework than those that exist for other spatial technologies.

To this end, we introduce spammR (SPatial Analysis of Multiomics Measurements in R), an R package that is well-suited for an end-to-end analysis of mass-spectrometry derived spatial omics datasets with (1) smaller sample sizes and spatial sparsity of sampling, (2) considerable missingness, and (3) no a-priori knowledge about proteins or genes of interest, relying on a fully data-driven approach. Here we describe the overall features of spammR, and how it can be used in “traditional” spatial omics contexts, using proteomics data from pancreatic cancer, and lipidomics and proteomics data from rat brain as examples, as well as more broadly as a tool to perform comparative omics studies in a spatial context using metagenomic data from the 1000 soils project (Bowman et al. 2023; Song et al. 2025).

2. Implementation

spammR is a package implemented in R and will be available via Bioconductor. We make use of the ‘SpatialExperiment’ (Righelli et al. 2022) object via Bioconductor (Gentleman et al. 2004; Huber et al. 2015) to provide the ability to integrate with other tools in the Bioconductor environment, and for increased interoperability with other tools in the spatial omics field (Feng et al. 2023; Pardo et al. 2022; Windhager et al. 2023). The package consists of five main components (**Figure 1A**): (1) Data ingest, (2) data processing, (3) feature selection, (4) functional enrichment, and (5) integrated visualization.

The data ingest component is comprised of two functions designed to capture the diverse data modalities that spammR analyzes. The primary function is `convert_to_spe()` and consumes three types of data: a data matrix of counts or values representing the samples measured (samples as columns, features as rows), a metadata table mapping the samples to a coordinate system as well as any other phenotypic data, and the image itself. spammR assumes that the molecular data are fully normalized/batch corrected (Välikangas et al. 2018), however additional corrections can also be applied depending on the data modality consumed (e.g. RNASeq vs proteomics). The function then returns a `SpatialExperiment` object containing the information that can be used for additional processing or analysis. To account for mass spectrometry imaging data on the Metaspace2020 platform(Veličković et al. 2021), we also included a `retrieve_from_metaspace()` function that selects a specific project and downloads both the coordinate metadata as well as the molecular quantifications. This function also returns a `SpatialExperiment` object. Each function only retrieves one data modality at a time, with merging happening using utility functions described below.

The data processing component currently consists of four functions designed to aid in multiomic analysis. To account for missingness in data, we included the `calc_missingness()` function to identify missing feature values in the object, and the `impute_spe()` that can run eight basic imputation algorithms. Another key feature of the spammR is the ability to measure diverse data modalities and samples. For this we included the `spat_reduce()` function that merges two `SpatialExperiment` objects on the same coordinate space as well as the `split_spe()` function that separates out a single object into multiple smaller objects.

The feature selection component contains functions to identify features (e.g. proteins) that either show strong correlation between abundances and spatial distance based on ROIs (`distance_based_analysis()`) or display significant differential expression profiles based on sample categorization (e.g. a categorical distinction between ROIs) without incorporation of distance measures (`calc_spatial_diff_ex()`). Both functions take a `SpatialExperiment` object created by `convert_to_spe()` and return an augmented `SpatialExperiment` with the added significance rankings.

The functional enrichment component contains functions to generate over-representation statistics (ORA) for feature selection results. `enrich_gradient()` can be used to generate pathway enrichment based on results from the distance-based analysis (`distance_based_analysis()`). Analogously, `enrich_ora()` does the same for ranking results from categorical differential expression analysis. Both functions rely on leapR (Danna et al. 2021) for pathway and gene set enrichment analysis (GSEA) (Subramanian et al. 2005). Like the feature selection functions, both functions in the enrichment component will return an augmented `SpatialExperiment` object that can in turn be used for visualization of enriched pathways in ROIs.

Finally, the data contained in a `SpatialExperiment` object, either directly generated by the raw data ingestion or augmented using the feature selection components, can then be visualized using the visualization component which contains two different visualization functions. (1) `spatial_heatmap()` overlays a grid onto image data representing the different ROIs and shading the individual grid cells by the calculated differential (e.g. raw abundance measures or differential expression calculated by one of the functional enrichment components). (2) `volcano_plot()` can be used to visualize the differential expression results generated by `calc_spatial_diff_ex()`. We also enable network-style visualizations through the `spatial_network()` function, which identifies pair-wise relationships between molecules across space that can then be plotted using the tidygraph R package.

All functions are designed to be lightweight, data agnostic, and provide interoperability via the `SpatialExperiment` object to enable usage beyond the functionality outlined here.

3. Application

We included in the spammR package three examples of applications using multiomics measurements across space.

3.1. Coordinate-based spatial heatmap in spammR enables contextual interpretation of spatial differences

To demonstrate the capabilities of spammR we conducted an analysis on previously published multi-omics data of human pancreatic tissue samples (Gosline et al. 2023). Specifically, we start by using the available methods in the spammR package to generate and visualize abundance differences for insulin across regions of interest (ROIs) which predominantly contain islets or non-islet cells respectively, observing the expected behavior of higher abundance of insulin in islet grid cells (**Figure 1B**). Next, we performed a pathway enrichment analysis using spammR's feature selection and functional enrichment components by pooling images from 7 distinct pancreas samples. As expected, pathways associated with insulin regulation and secretion are enriched islet cells in comparison to pancreatic regions containing predominantly non-islet cells. Finally, using the distance-based analysis functions in spammR we were able to identify and visualize the fact that AP3S2, a subunit of the AP-3 complex that is thought to be associated with protein transport (Cowles et al. 1997), expression strength is strongly anti-correlated with the distance from islet cells, i.e. decreasing expression with increasing distance. The entire analysis is showcased in one of the vignettes (spatProt.rmd) included with the spammR package.

3.2 spammR's agnostic approach to multi-omics integration enables the investigation of metagenomic data in a geospatial context

The implementation of spammR's spatial component is inherently agnostic to whether the distances within the investigated sample are on a micrometer, millimeter or even kilometer scale. To demonstrate this, we retrieved KEGG (Kanehisa et al. 2025) Ortholog (KO) presence data for 108 metagenome samples (54 sampling sites, samples each at two depths) from the 1000 soils project pilot study of MONet (Molecular Observation Network). Note that this is different from mass-spectrometry proteomics data in so far that in the 1000 soils data the number of predicted protein sequences from metagenome-assembled genomes (MAGs) that are associated with a given KO is recorded (i.e. presence), whereas traditional MS-proteomics data reports quantitative abundance measurements. This leads to significantly lower "abundance values" as well as lower variance between the individual values. In addition to the KO abundance, each sample contains further metadata, like soil temperature, water content, pH, metal abundances, and most importantly for spatial context, latitude, longitude, elevation and whether the sample originates from the top or the bottom of the extracted soil core. Using the metadata, we retrieved 575 KOs that show significant differential abundances across samples from desert vs. non-desert sampling locations (483 "up regulated", 92 "down regulated"). We then utilized leapR's functionality to retrieve enriched KEGG Pathways based on the KO abundance data. The pathway enrichment detected several enriched pathways, for example "sulfur cycle". Utilizing map data from the US census bureau, the terra R package as well as the latitude and longitude annotations we are then able to visualize the pathway enrichment using the `spatial_heatmap()` function (**Figure 1C**). A complete workflow can be found in the spatMicrobiome vignette that is part of spammR.

3.3 Integration of proteomic and lipidomic measurements in brain tissue

To show how spammR can be used to integrate multiple omics from the same sample, we leveraged data from Vandergrift et al. (Vandergrift et al. 2025) that measured lipidomics and proteomics from a tissue sample from rat brain. We used the `retrieve_metaspace_data()` to fetch the experimental data from the Metaspace portal (Palmer et al. 2017) and coordinates from the lipids, then downloaded the proteomics data from the original publication. The original lipid data, shown in **Figure 1D,** spans the entire image, while the protein data (carved out as squares in the image) only represented a subset of the image. We therefore used the spat_reduce() function to align the two coordinate systems to map the average lipid expression to the regions that correspond to protein measurements, shown in **Figure 1E**. Using this combined object, we could then build a correlation graph, shown in **Figure 1F**, that identifies specific proteins correlated with a particular lipid of interest, in this case a phosphatidylserine identified as highly variable across the image (PS(28:6(10Z,13Z,16Z,19Z,22Z,25Z)/12:0)). We then ran leapR analysis on the correlation results to identify a KEGG pathway significantly enriched (negatively correlated in this case) with the lipid, depicted in **Figure 1G**.

4. Conclusion

Through spammR, we address multiple limitations in the current state-of-the-art tools for computational analysis of diverse spatial omics measurements with a suite of lightweight methods that can be broadly applied across mass spectrometry and sequence-derived measurements. spammR is built upon the existing Bioconductor framework for spatial omics measurements and can therefor interoperate with the ever-growing set of tool for spatial transcriptomics, but adds specific expansions to explicitly (a) account for issues that generally arise with mass spectrometry such as missingness and (b) align omics from diverse modalities in the same image.

We worked to include a diverse set of features that could be used in the spatial omics community, but made distinct choices about when to include functionality vs. provide an interface to others. For example, in the

design of our `impute_spe()` we tried to address the importance of missing data (e.g. in data-dependent acquisition (DDA)(Wu and MacCoss 2002) proteomics data). In general, missing abundance for individual features does not equate to proteins not being expressed and can broadly be categorized into values missing (completely) at random (MCAR & MAR) and values missing not at random (MNAR). MAR features can mostly be attributed to experimental design or instrument errors. Hence, missing value intensities do not necessarily follow a pattern, making them more difficult to impute. On the other hand, MNAR values can often be traced back to technical limitations of the instrument, e.g. proteins of low expression not being detected due to intensities below the detection capabilities of the instrument. Given the different nature of MAR and MNAR values where missing values are likely to follow different intensity distributions, existing imputation strategies do not universally perform well for both types of missing values(Lazar et al. 2016; Webb-Robertson et al. 2015; Wei et al. 2018). For example, assuming MNAR for observed missing values, naïve single value imputation methods like SampMin (minimal value of the sample) can show good performance especially in downstream analyses like differential expression analysis(Harris et al. 2023; Kong et al. 2022; Liu and Dongre 2021). Especially when dealing with spatial data, where cell type-specific proteins that may only be expressed in one region of the tissue, those naïve imputation strategies are a possible approach. To account for diverse use cases, `impute_spe()` currently includes eight imputation methods including spatially-aware imputation method that adapts k-nearest neighbor imputation to assign values to missing abundance data points based on molecular abundances only from samples that are spatially close to each other (rather than close to each other in expression value). Finally, considering that spammR is built on widely used data structures with a focus on interoperability with other R packages, users also have the ability to integrate other available imputation strategies from libraries like imputeLCMD(Lazar et al. 2022) or MsCoreUtils(Rainer et al. 2022) into their workflow.

We demonstrated the power of spammR for the analysis of a traditional spatial proteomics datasets using a spatial proteomics analysis of healthy human pancreatic tissue. Additionally, we showed that spammR is scale agnostic and can also be utilized on geospatial distance scale data by using a metagenome dataset highlighting the agnostic capabilities of spammR extending beyond the analysis of traditional multi-omic data. Lastly, we showed the power of integration using a small dataset that collects both lipidomics and proteomics from the same tissue.

In the future, we plan to continue expanding the functionality provided by spammR through the collection of larger, more complex datasets across tissues of interest. We aim to improve data processing through the incorporation of other spatial analysis such as the `SpatialFeatureExperiment`(Moses et al. 2023) and the direct incorporation of geoJSON format to link to tools such as quPath (Bankhead et al. 2017) or microscope vendor software. As multiomics datasets become more comprehensive, we also plan to improve our statistical methods for identification of functional pathways that account for measurements across omics.

Acknowledgements

Soil data were provided by the Molecular Observation Network (MONet) at the Environmental Molecular Sciences Laboratory (https://ror.org/04rc0xn13), a DOE Office of Science user facility sponsored by the Biological and Environmental Research program under Contract No. DE-AC05-76RL01830. The work (proposal: 10.46936/10.25585/60008970) conducted by the U.S. Department of Energy, Joint Genome Institute (https://ror.org/04xm1d337), a DOE Office of Science user facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231. A subset of soil cores was obtained by the National Ecological Observatory Network's Research Support Services which is a program sponsored by the U.S. National Science Foundation and operated under cooperative agreement by Battelle. The Molecular Observation Network (MONet) database is an open, FAIR, and publicly available compilation of

the molecular and microstructural properties of soil. Data in the MONet open science database can be found at https://sc-data.emsl.pnnl.gov/.

Author contributions

Yannick Mahlich: Analysis, Writing – original draft, **Harkirat Sohi:** Methodology, Software, **Jason McDermott:** Conceptualization, Supervision, Writing – editing, **Sara Gosline:** Conceptualization, Methodology, Supervision, Software, Writing – original draft

Funding

This work was supported by DoD CDMRP Program W81XWH-22-1-0324, NF210042; NIH grant U01-CA271412; and USDA NF-SPP 81327.

Conflict of interest

None

Figure Caption

Figure 1 Overview of spammR functionality. (A) Flow chart describing use of spammR package as well as how the methods of the package relate to one another. (B) Expression of the protein insulin in a human pancreas tissue, with labels indicating whether the image region contains islet cells or not. (C) Abundance of the sulfur cycle KEGG orthology group across the US Soil study. (D) Expression of a highly variable lipid across a rat brain sample, as well (E) the expression of that lipid only within the regions measured by proteomics. (F) Network representing the lipid of interest (pink) and highly positively or negatively correlated proteins (turquoise). (G) Expression of proteins in the KEGG Alanine, aspartate, and glutamate pathway, which was identified as significantly enriched in proteins negatively correlated with the lipid of interest shown in panel.

References

Angelo, Michael, Sean C. Bendall, Rachel Finck, et al. 2014. “Multiplexed Ion Beam Imaging of Human Breast Tumors.” *Nature Medicine* 20 (4): 436–42. https://doi.org/10.1038/nm.3488.

Bankhead, Peter, Maurice B. Loughrey, José A. Fernández, et al. 2017. “QuPath: Open Source Software for Digital Pathology Image Analysis.” *Scientific Reports* 7 (1): 16878. https://doi.org/10.1038/s41598-017-17204-5.

Black, Sarah, Darci Phillips, John W. Hickey, et al. 2021. “CODEX Multiplexed Tissue Imaging with DNA-Conjugated Antibodies.” *Nature Protocols* 16 (8): 3802–35. https://doi.org/10.1038/s41596-021-00556-8.

Bouwman, Britta A. M., Nicola Crosetto, and Magda Bienko. 2022. “The Era of 3D and Spatial Genomics.” *Trends in Genetics* 38 (10): 1062–75. https://doi.org/10.1016/j.tig.2022.05.010.

Bowman, Maggie M., Alexis E. Heath, Tamas Varga, et al. 2023. “One Thousand Soils for Molecular Understanding of Belowground Carbon Cycling.” *Frontiers in Soil Science* 3 (April). https://doi.org/10.3389/fsoil.2023.1120425.

Buchberger, Amanda Rae, Kellen DeLaney, Jillian Johnson, and Lingjun Li. 2018. “Mass Spectrometry Imaging: A Review of Emerging Advancements and Future Insights.” *Analytical Chemistry* 90 (1): 240–65. https://doi.org/10.1021/acs.analchem.7b04733.

Chen, Kok Hao, Alistair N. Boettiger, Jeffrey R. Moffitt, Siyuan Wang, and Xiaowei Zhuang. 2015. “Spatially Resolved, Highly Multiplexed RNA Profiling in Single Cells.” *Science* 348 (6233): aaa6090. https://doi.org/10.1126/science.aaa6090.

Codeluppi, Simone, Lars E. Borm, Amit Zeisel, et al. 2018. “Spatial Organization of the Somatosensory Cortex Revealed by osmFISH.” *Nature Methods* 15 (11): 932–35. https://doi.org/10.1038/s41592-018-0175-z.

Cowles, Christopher R., Greg Odorizzi, Gregory S. Payne, and Scott D. Emr. 1997. “The AP-3 Adaptor Complex Is Essential for Cargo-Selective Transport to the Yeast Vacuole.” *Cell* 91 (1): 109–18. https://doi.org/10.1016/S0092-8674(01)80013-1.

Crosetto, Nicola, Magda Bienko, and Alexander van Oudenaarden. 2015. “Spatially Resolved Transcriptomics and Beyond.” *Nature Reviews Genetics* 16 (1): 57–66. https://doi.org/10.1038/nrg3832.

Danna, Vincent, Hugh Mitchell, Lindsey Anderson, et al. 2021. “leapR: An R Package for Multiomic Pathway Analysis.” *Journal of Proteome Research* 20 (4): 2116–21. https://doi.org/10.1021/acs.jproteome.0c00963.

Eng, Chee-Huat Linus, Michael Lawson, Qian Zhu, et al. 2019. “Transcriptome-Scale Super-Resolved Imaging in Tissues by RNA seqFISH+.” *Nature* 568 (7751): 235–39. https://doi.org/10.1038/s41586-019-1049-y.

Feng, Yuzhou, Tianpei Yang, John Zhu, et al. 2023. “Spatial Analysis with SPIAT and spaSim to Characterize and Simulate Tissue Microenvironments.” *Nature Communications* 14 (1): 2697. https://doi.org/10.1038/s41467-023-37822-0.

Gentleman, Robert C., Vincent J. Carey, Douglas M. Bates, et al. 2004. “Bioconductor: Open Software Development for Computational Biology and Bioinformatics.” *Genome Biology* 5 (10): R80. https://doi.org/10.1186/gb-2004-5-10-r80.

Gosline, Sara J. C., Marija Veličković, James C. Pino, et al. 2023. “Proteome Mapping of the Human Pancreatic Islet Microenvironment Reveals Endocrine–Exocrine Signaling Sphere of Influence.” *Molecular & Cellular Proteomics* 22 (8): 100592. https://doi.org/10.1016/j.mcpro.2023.100592.

Greer, Tyler, Robert Sturm, and Lingjun Li. 2011. “Mass Spectrometry Imaging for Drugs and Metabolites.” *Journal of Proteomics*, Pharmacoproteomics and Toxicoproteomics, vol. 74 (12): 2617–31. https://doi.org/10.1016/j.jprot.2011.03.032.

Harris, Lincoln, William E. Fondrie, Sewoong Oh, and William S. Noble. 2023. “Evaluating Proteomics Imputation Methods with Improved Criteria.” *Journal of Proteome Research* 22 (11): 3427–38. https://doi.org/10.1021/acs.jproteome.3c00205.

Huber, Wolfgang, Vincent J. Carey, Robert Gentleman, et al. 2015. “Orchestrating High-Throughput Genomic Analysis with Bioconductor.” *Nature Methods* 12 (2): 115–21. https://doi.org/10.1038/nmeth.3252.

Kanehisa, Minoru, Miho Furumichi, Yoko Sato, Yuriko Matsuura, and Mari Ishiguro-Watanabe. 2025. “KEGG: Biological Systems Database as a Model of the Real World.” *Nucleic Acids Research* 53 (D1): D672–77. https://doi.org/10.1093/nar/gkae909.

Kiessling, Paul, and Christoph Kuppe. 2024. “Spatial Multi-Omics: Novel Tools to Study the Complexity of Cardiovascular Diseases.” *Genome Medicine* 16 (1): 14. https://doi.org/10.1186/s13073-024-01282-y.

Kong, Weijia, Harvard Wai Hann Hui, Hui Peng, and Wilson Wen Bin Goh. 2022. “Dealing with Missing Values in Proteomics Data.” *PROTEOMICS* 22 (23–24): 2200092. https://doi.org/10.1002/pmic.202200092.

Lazar, Cosmin, Thomas Burger, and Samuel Wieczorek. 2022. *imputeLCMD: A Collection of Methods for Left-Censored Missing Data Imputation*. V. 2.1. Released June 10. https://cran.r-project.org/web/packages/imputeLCMD/index.html.

Lazar, Cosmin, Laurent Gatto, Myriam Ferro, Christophe Bruley, and Thomas Burger. 2016. “Accounting for the Multiple Natures of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare Imputation Strategies.” *Journal of Proteome Research* 15 (4): 1116–25. https://doi.org/10.1021/acs.jproteome.5b00981.

Lee, Je Hyuk, Evan R. Daugharthy, Jonathan Scheiman, et al. 2014. “Highly Multiplexed Subcellular RNA Sequencing in Situ.” *Science* 343 (6177): 1360–63. https://doi.org/10.1126/science.1250212.

Lin, Jia-Ren, Benjamin Izar, Shu Wang, et al. 2018. “Highly Multiplexed Immunofluorescence Imaging of Human Tissues and Tumors Using T-CyCIF and Conventional Optical Microscopes.” *eLife* 7 (July): e31657. https://doi.org/10.7554/eLife.31657.

Liu, Mingyi, and Ashok Dongre. 2021. “Proper Imputation of Missing Values in Proteomics Datasets for Differential Expression Analysis.” *Briefings in Bioinformatics* 22 (3): bbaa112. https://doi.org/10.1093/bib/bbaa112.

Lundberg, Emma, and Georg H. H. Borner. 2019. “Spatial Proteomics: A Powerful Discovery Tool for Cell Biology.” *Nature Reviews Molecular Cell Biology* 20 (5): 285–302. https://doi.org/10.1038/s41580-018-0094-y.

Moffitt, Jeffrey R., Emma Lundberg, and Holger Heyn. 2022. “The Emerging Landscape of Spatial Profiling Technologies.” *Nature Reviews Genetics* 23 (12): 741–59. https://doi.org/10.1038/s41576-022-00515-3.

Moses, Lambda, Pétur Helgi Einarsson, Kayla Jackson, et al. 2023. “Voyager: Exploratory Single-Cell Genomics Data Analysis with Geospatial Statistics.” Preprint, bioRxiv, August 20. https://doi.org/10.1101/2023.07.20.549945.

Moses, Lambda, and Lior Pachter. 2022. “Museum of Spatial Transcriptomics.” *Nature Methods* 19 (5): 534–46. https://doi.org/10.1038/s41592-022-01409-2.

Nichterwitz, Susanne, Julio Aguila Benitez, Rein Hoogstraaten, Qiaolin Deng, and Eva Hedlund. 2018. “LCM-Seq: A Method for Spatial Transcriptomic Profiling Using Laser Capture Microdissection Coupled with PolyA-Based RNA Sequencing.” In *RNA Detection: Methods and Protocols*, edited by Imre Gaspar. Springer. https://doi.org/10.1007/978-1-4939-7213-5_6.

Nimo, Jose, Sonja Fritzsche, Daniela S. Valdes, et al. 2025. “OpenDVP: An Experimental and Computational Framework for Community-Empowered Deep Visual Proteomics.” Preprint, bioRxiv, July 16. https://doi.org/10.1101/2025.07.13.662099.

Palmer, Andrew, Prasad Phapale, Ilya Chernyavsky, et al. 2017. “FDR-Controlled Metabolite Annotation for High-Resolution Imaging Mass Spectrometry.” *Nature Methods* 14 (1): 57–60. https://doi.org/10.1038/nmeth.4072.

Pardo, Brenda, Abby Spangler, Lukas M. Weber, et al. 2022. “spatialLIBD: An R/Bioconductor Package to Visualize Spatially-Resolved Transcriptomics Data.” *BMC Genomics* 23 (1): 434. https://doi.org/10.1186/s12864-022-08601-w.

Piehowski, Paul D., Ying Zhu, Lisa M. Bramer, et al. 2020. “Automated Mass Spectrometry Imaging of over 2000 Proteins from Tissue Sections at 100-Mm Spatial Resolution.” *Nature Communications* 11 (1): 8. https://doi.org/10.1038/s41467-019-13858-z.

Radtke, Andrea J., Colin J. Chu, Ziv Yaniv, et al. 2022. “IBEX: An Iterative Immunolabeling and Chemical Bleaching Method for High-Content Imaging of Diverse Tissues.” *Nature Protocols* 17 (2): 378–401. https://doi.org/10.1038/s41596-021-00644-9.

Rainer, Johannes, Andrea Vicini, Liesa Salzer, et al. 2022. “A Modular and Expandable Ecosystem for Metabolomics Data Annotation in R.” *Metabolites* 12 (2). https://doi.org/10.3390/metabo12020173.

Righelli, Dario, Lukas M Weber, Helena L Crowell, et al. 2022. “SpatialExperiment: Infrastructure for Spatially-Resolved Transcriptomics Data in R Using Bioconductor.” *Bioinformatics* 38 (11): 3128–31. https://doi.org/10.1093/bioinformatics/btac299.

Rodriques, Samuel G., Robert R. Stickels, Aleksandrina Goeva, et al. 2019. “Slide-Seq: A Scalable Technology for Measuring Genome-Wide Expression at High Spatial Resolution.” *Science* 363 (6434): 1463–67. https://doi.org/10.1126/science.aaw1219.

Song, Young C., Cheng Shi, Kelly G. Stratton, et al. 2025. “Continental-Scale Integration of Soil Metagenomes and Metabolomes Reveals Ubiquitous Microbial Capacity for Recalcitrant Carbon Decomposition.” Preprint, bioRxiv, July 3. https://doi.org/10.1101/2025.07.03.663048.

Subramanian, Aravind, Pablo Tamayo, Vamsi K. Mootha, et al. 2005. “Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles.” *Proceedings of the National Academy of Sciences* 102 (43): 15545–50. https://doi.org/10.1073/pnas.0506580102.

Taylor, Michael J., Jessica K. Lukowski, and Christopher R. Anderton. 2021. “Spatially Resolved Mass Spectrometry at the Single Cell: Recent Innovations in Proteomics and Metabolomics.” *Journal of the American Society for Mass Spectrometry* 32 (4): 872–94. https://doi.org/10.1021/jasms.0c00439.

Välikangas, Tommi, Tomi Suomi, and Laura L Elo. 2018. “A Systematic Evaluation of Normalization Methods in Quantitative Label-Free Proteomics.” *Briefings in Bioinformatics* 19 (1): 1–11. https://doi.org/10.1093/bib/bbw095.

Vandereyken, Katy, Alejandro Sifrim, Bernard Thienpont, and Thierry Voet. 2023. “Methods and Applications for Single-Cell and Spatial Multi-Omics.” *Nature Reviews Genetics* 24 (8): 494–515. https://doi.org/10.1038/s41576-023-00580-2.

Vandergrift, Gregory W., Marija Veličković, Le Z. Day, et al. 2025. “Untargeted Spatial Metabolomics and Spatial Proteomics on the Same Tissue Section.” *Analytical Chemistry* 97 (1): 392–400. https://doi.org/10.1021/acs.analchem.4c04462.

Veličković, Dušan, Tamara Bečejac, Sergii Mamedov, et al. 2021. “Rapid Automated Annotation and Analysis of N-Glycan Mass Spectrometry Imaging Data Sets Using NGlycDB in METASPACE.” *Analytical Chemistry* 93 (40): 13421–25. https://doi.org/10.1021/acs.analchem.1c02347.

Veličković, Marija, Leena Kadam, Joonhoon Kim, et al. 2025. “Advanced Multi-Modal Mass Spectrometry Imaging Reveals Functional Differences of Placental Villous Compartments at Microscale Resolution.” *Nature Communications* 16 (1): 2061. https://doi.org/10.1038/s41467-025-57107-y.

Veličković, Marija, Ruonan Wu, Yuqian Gao, et al. 2024. “Mapping Microhabitats of Lignocellulose Decomposition by a Microbial Consortium.” *Nature Chemical Biology* 20 (8): 1033–43. https://doi.org/10.1038/s41589-023-01536-7.

Walker, Benjamin L., Zixuan Cang, Honglei Ren, Eric Bourgain-Chang, and Qing Nie. 2022. “Deciphering Tissue Structure and Function Using Spatial Transcriptomics.” *Communications Biology* 5 (1): 1–10. https://doi.org/10.1038/s42003-022-03175-5.

Wang, Xiao, William E. Allen, Matthew A. Wright, et al. 2018. “Three-Dimensional Intact-Tissue Sequencing of Single-Cell Transcriptional States.” *Science* 361 (6400): eaat5691. https://doi.org/10.1126/science.aat5691.

Webb-Robertson, Bobbie-Jo M., Holli K. Wiberg, Melissa M. Matzke, et al. 2015. “Review, Evaluation, and Discussion of the Challenges of Missing Value Imputation for Mass Spectrometry-Based Label-Free Global Proteomics.” *Journal of Proteome Research* 14 (5): 1993–2001. https://doi.org/10.1021/pr501138h.

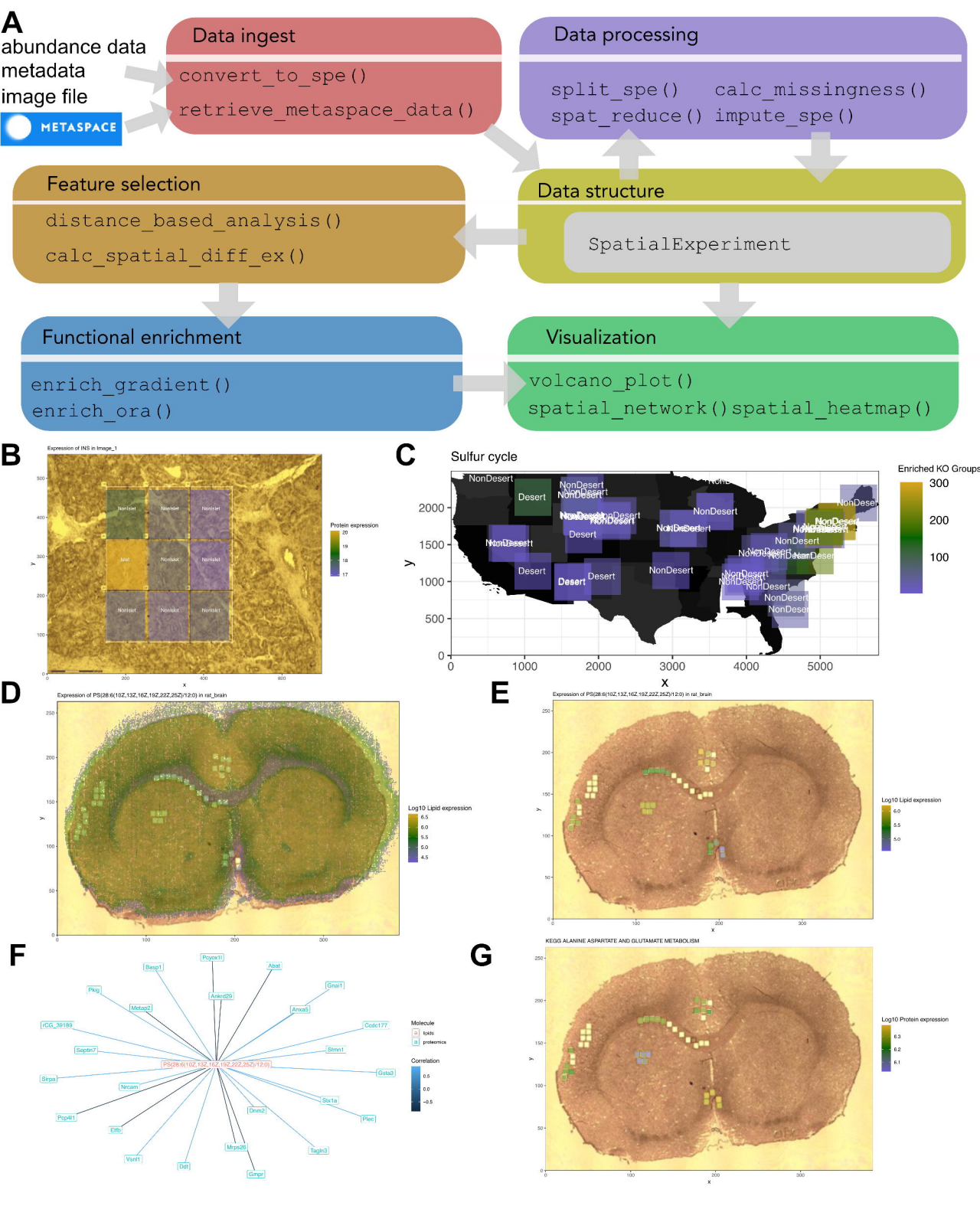
Wei, Runmin, Jingye Wang, Mingming Su, et al. 2018. “Missing Value Imputation Approach for Mass Spectrometry-Based Metabolomics Data.” *Scientific Reports* 8 (1): 663. https://doi.org/10.1038/s41598-017-19120-0.

Windhager, Jonas, Vito Riccardo Tomaso Zanotelli, Daniel Schulz, et al. 2023. “An End-to-End Workflow for Multiplexed Image Processing and Analysis.” *Nature Protocols* 18 (11): 3565–613. https://doi.org/10.1038/s41596-023-00881-0.

Wu, Christine C., and Michael J. MacCoss. 2002. “Shotgun Proteomics: Tools for the Analysis of Complex Biological Systems.” *Current Opinion in Molecular Therapeutics* 4 (3): 242–50.

Zhang, Meng, Stephen W. Eichhorn, Brian Zingg, et al. 2021. “Spatially Resolved Cell Atlas of the Mouse Primary Motor Cortex by MERFISH.” *Nature* 598 (7879): 137–43. https://doi.org/10.1038/s41586-021-03705-x.

Zhu, Ying, Paul D. Piehowski, Rui Zhao, et al. 2018. “Nanodroplet Processing Platform for Deep and Quantitative Proteome Profiling of 10–100 Mammalian Cells.” *Nature Communications* 9 (1): 882. https://doi.org/10.1038/s41467-018-03367-w.


A
abundance data
metadata
image file
METASPACE
Data ingest
convert_to_spe()
retrieve_metaspace_data()
Data processing
split_spe()
calc_missingness()
spat_reduce()
impute_spe()
Feature selection
distance_based_analysis()
calc_spatial_diff_ex()
Data structure
SpatialExperiment
Functional enrichment
enrich_gradient()
enrich_ora()
Visualization
volcano_plot()
spatial_network()
spatial_heatmap()
B
Expression of INS in Image_1
Protein expression
C
Sulfur cycle
Enriched KO Groups
NonDesert
Desert
x
y
D
Expression of PS(28:6(10Z,13Z,16Z,19Z,22Z,25Z)/12:0) in rat_brain
Log10 Lipid expression
E
Expression of PS(28:6(10Z,13Z,16Z,19Z,22Z,25Z)/12:0) in rat_brain
Log10 Lipid expression
F
Molecule
lipids
proteomics
Correlation
G
KEGG ALANINE ASPARTATE AND GLUTAMATE METABOLISM
Log10 Protein expression