Oct 15, 2025 Version 3

Smart spatial omics (S2-omics) optimizes region-of-interest selection to capture molecular heterogeneity in diverse tissues V.3

DOI
https://dx.doi.org/10.17504/protocols.io.bp2l6zpnkgqe/v3

Musu Yuan[1], Kaitian Jin[1], Hanying Yan[1], Amelia Schroeder[1], Chunyu Luo[1], Sicong Yao[1], Bernhard Domoulin[2], Jonathan Levinsohn[2], Tianhao Luo[2], Jean R. Clemenceau[3], Inyeop Jang[3], Minji Kim[3], Yunhe Liu[4,5,6,7], Minghua Deng[8], Emma E. Furth[9], Parker Wilson[9], Anupma Nayak[9], Idania Lubo[10], Luisa Maren Solis Soto[10], Linghua Wang[4,5,6,7], Jeong Hwan Park[11], Katalin Susztak[2], Tae Hyun Hwang[3], Mingyao Li[1]

[1]Statistical Center for Single-Cell and Spatial Genomics, Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania Philadelphia, PA, USA.;
[2]Renal, Electrolyte, and Hypertension Division, Department of Medicine, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA 19104, United States.;
[3]Department of Surgery, Vanderbilt University Medical Center, Nashville, TN, USA.;
[4]Department of Genomic Medicine, University of Texas MD Anderson Cancer Center, Houston, TX, USA.;
[5]The University of Texas MD Anderson Cancer Center UTHealth Houston Graduate School of Biomedical Sciences, Houston, USA.;
[6]The James P. Allison Institute, The University of Texas MD Anderson Cancer Center, Houston, TX, USA.;
[7]Institute for Data Science in Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA.;
[8]Center for Quantitative Biology, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China.;
[9]Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, United States.;
[10]Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA.;
[11]Department of Pathology, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, Seoul National University College of Medicine, Seoul, Republic of Korea.

S2-omics

musu

OPEN ACCESS

DOI: https://dx.doi.org/10.17504/protocols.io.bp2l6zpnkgqe/v3

Protocol Citation: Musu Yuan, Kaitian Jin, Hanying Yan, Amelia Schroeder, Chunyu Luo, Sicong Yao, Bernhard Domoulin, Jonathan Levinsohn, Tianhao Luo, Jean R. Clemenceau, Inyeop Jang, Minji Kim, Yunhe Liu, Minghua Deng, Emma E. Furth, Parker Wilson, Anupma Nayak, Idania Lubo, Luisa Maren Solis Soto, Linghua Wang, Jeong Hwan Park, Katalin Susztak, Tae Hyun Hwang, Mingyao Li 2025. Smart spatial omics (S2-omics) optimizes region-of-interest selection to capture molecular heterogeneity in diverse tissues. **protocols.io** https://dx.doi.org/10.17504/protocols.io.bp2l6zpnkgqe/v3Version created by **musu**

Manuscript citation:
This protocol is associated with our paper describing S2-omics. Paper link: https://www.biorxiv.org/content/10.1101/2025.09.21.677634v1

ReadTheDocs page: https://s2omics.readthedocs.io/en/latest/

Github page: https://github.com/ddb-qiwang/S2Omics

Protocol status: Working
We use this protocol and it's working

Created: October 15, 2025

Last Modified: October 15, 2025

Protocol Integer ID: 229849

Keywords: spatial molecular profiles across entire tissue section, resulting spatial omics data, using histology image, colorectal cancer tissue section, spatial omics data, smart spatial omic, spatial molecular profile, histology image, spatial omics experiment, making spatial omics study, spatial omics study, molecular heterogeneity in diverse tissue, preserving critical spatial molecular variation, diverse tissue, critical spatial molecular variation, entire tissue section, omic, selected rois

Abstract

A protocol describing the application of S2-omics on a colorectal cancer tissue section, designing 10x VisiumHD experiment.

S2-omics is an end-to-end workflow that automatically selects regions of interest for spatial omics experiments using histology images. Additionally, S2-omics utilizes the resulting spatial omics data to virtually reconstruct spatial molecular profiles across entire tissue sections, providing valuable insights to guide subsequent experimental steps. Our histology image-guided design significantly reduces experimental costs while preserving critical spatial molecular variations, thereby making spatial omics studies more accessible and cost-effective.

We demonstrate S2-omics's utility in selecting a 6.5mm × 6.5mm ROI and predicting the cell-type labels for the whole slide image after conducting VisiumHD experiment on the selected ROI.

Materials

he-raw.jpg: Raw histology image.

pixel-size-raw.txt: Side length (in micrometers) of pixels in he-raw.jpg. This value is usually between 0.1 and 1.0. For an instance, if the resolution of raw H6E image is 0.2 microns/pixel, you can just create a txt file and write down the value '0.2'.

annotation_file.csv(optional): The annotation and spatial location of superpixels, should at least contain three columns: 'super_pixel_x', 'super_pixel_y', 'annotation'. This file is not needed for ROI selection. For an instance, the first row of this table means the cell type of 267th row (top-down) 1254th column (left-right) superpixel is Myofibroblast.

Troubleshooting

Before start

Install Python. We recommend using conda, miniconda, or mamba for this.

To run the demo, first please download the demo data and pretrained model checkpoints file from:

google drive: https://drive.google.com/drive/folders/1z1nk0sF_e25LKMHyJxJVMtROFjuWet2G?usp=sharing

Please place both 'checkpoints' and 'demo' folder under the 'S2Omics' main folder.

In this demo, we mimic the situation that we need to select a 6.5 mm*6.5 mm ROI for Visium HD experiment from a colorectal cancer tissue section. To run the ROI selection (takes about 25 minutes with GPU),

Before starting

1 Install Python. We recommend using conda, miniconda, or mamba for this.

2 To run the demo, first please download the demo data and pretrained model checkpoints file from:

3 google drive:https://drive.google.com/drive/folders/1z1nk0sF_e25LKMHyJxJVMtROFjuWet2G?usp=sharing

4 Please place both 'checkpoints' and 'demo' folder under the 'S2Omics' main folder.

5 We provide both python files and notebooks for all examples. User can download them from S2-omics's github repository https://github.com/ddb-qiwang/S2Omics

Data format

6 Here are the files needed for running S2-omics. User can further refer to the demo data to check the detailed data formats.

7 **he-raw.jpg:** Raw histology image.

8 **pixel-size-raw.txt:** Side length (in micrometers) of pixels in he-raw.jpg. This value is usually between 0.1 and 1.0. For an instance, if the resolution of raw H&E image is 0.2 microns/pixel, you can just create a txt file and write down the value '0.2'.

9 **annotation_file.csv(optional):** The annotation and spatial location of superpixels, should at least contain three columns: 'super_pixel_x', 'super_pixel_y', 'annotation'. This file is not needed for ROI selection. User can refer to the demo for more detailed input information.

ROI selection for single slide

10 For example, to select ROI on the demo colorectal cancer section:

11 python run_roi_selection_single.py --prefix './demo/Tutorial_1_VisiumHD_ROI_selection_colon/' --save_folder './demo/Tutorial_1_VisiumHD_ROI_selection_colon/S2Omics_output' --device 'cuda:0' --roi_size 6.5 6.5 --num_roi 1

ROI selection for multiple slides

12 To select ROI on the demo consecutive breast cancer sections:

13 python run_roi_selection_multiple.py --prefix_list './demo/Tutorial_3_Consecutive_ROI_selection_breast/breast_cancer_g1/' './demo/Tutorial_3_Consecutive_ROI_selection_breast/breast_cancer_g2/' './demo/Tutorial_3_Consecutive_ROI_selection_breast/breast_cancer_g3/' --save_folder_list './demo/Tutorial_3_Consecutive_ROI_selection_breast/breast_cancer_g1/S2Omics_output' './demo/Tutorial_3_Consecutive_ROI_selection_breast/breast_cancer_g2/S2Omics_output' './demo/Tutorial_3_Consecutive_ROI_selection_breast/breast_cancer_g3/S2Omics_output' --device 'cuda:0' --roi_size 1.5 1.5 --num_roi 1

Cell type broadcasting

14 To broadcast the cell type label within th selected ROI to the entire slide on the demo colorectal cancer section:

15 python run_label_broadcasting.py --WSI_datapath './demo/Tutorial_1_VisiumHD_ROI_selection_colon/' --SO_datapath './demo/Tutorial_1_VisiumHD_ROI_selection_colon/' --WSI_save_folder './demo/Tutorial_1_VisiumHD_ROI_selection_colon/S2Omics_output' --SO_save_folder './demo/Tutorial_1_VisiumHD_ROI_selection_colon/S2Omics_output' --need_preprocess True --need_feature_extraction True

License

16 For commercial use of open source software S2-omics, please contact Musu Yuan (musu990519@gmail.com) and Mingyao Li (mingyao@pennmedicine.upenn.edu).