



Jul 21, 2025

CoCo-ST detects global and local biological structures in spatial transcriptomics datasets

 [Nature Cell Biology](#)

DOI

<https://dx.doi.org/10.17504/protocols.io.x54v95m6pl3e/v1>

Muhammad Aminu¹, Bo Zhu¹, Natalie Vokes¹, Jia Wu¹, Jianjun Zhang¹

¹The University of Texas MD Anderson Cancer Center

The University of Texas ...



Muhammad Aminu

The University of Texas MD Anderson Cancer Center

Create & collaborate more with a free account

Edit and publish protocols, collaborate in communities, share insights through comments, and track progress with run records.

Create free account

OPEN  ACCESS



DOI: <https://dx.doi.org/10.17504/protocols.io.x54v95m6pl3e/v1>

External link: <https://doi.org/10.1038/s41556-025-01781-z>



Protocol Citation: Muhammad Aminu, Bo Zhu, Natalie Vokes, Jia Wu, Jianjun Zhang 2025. CoCo-ST detects global and local biological structures in spatial transcriptomics datasets. **protocols.io**
<https://dx.doi.org/10.17504/protocols.io.x54v95m6pl3e/v1>

Manuscript citation:

Muhammad Aminu, Bo Zhu, Natalie Vokes, Hong Chen, Lingzhi Hong, Jianrong Li, Junya Fujimoto, Mehdi Chaib, Yuqiu Yang, Bo Wang, Alissa Poteete, Monique B. Nilsson, Xiuning Le, Tina Cascone, David Jaffray, Nicholas Navin, Tao Wang, Lauren A. Byers, Don L. Gibbons, John Heymach, Ken Chen, Chao Cheng, Jianjun Zhang, Jia Wu (2025) CoCo-ST detects global and local biological structures in spatial transcriptomics datasets. *Nature Cell Biology* doi: [10.1038/s41556-025-01781-z](https://doi.org/10.1038/s41556-025-01781-z)

License: This is an open access protocol distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Protocol status: Working

We use this protocol and it's working

Created: July 20, 2025

Last Modified: July 21, 2025

Protocol Integer ID: 222846

Keywords: local biological structures in spatial transcriptomics dataset, spatial transcriptomics dataset, st algorithm on spatial transcriptomics dataset, st algorithm for spatial domain detection, spatial domain detection, complexity of spatial domain detection, low variance spatial structure, high variance in global spatial pattern, global spatial pattern, local biological structure, relevant spatial structure, variance spatial structure, diverse tissue sample, spatial structure, main challenges with precancer analysis, precancer analysis, application of the coco, coco, contrastive learning approach, varying spatial resolution, driving tumor progression, spatial resolution, variance domain, st detect, proposed coco

Funders Acknowledgements:

National Institute of Health (NIH)

Grant ID: R01CA262425

National Institute of Health (NIH)

Grant ID: R00CA218667

National Institute of Health (NIH)

Grant ID: R01CA276178

National Institute of Health (NIH)

Grant ID: 5R50CA265307

Cancer Prevention and Research Institute of Texas (CPRI)

Grant ID: RP240117

Cancer Prevention and Research Institute of Texas (CPRI)

Grant ID: RP250399



Abstract

A protocol describing the application of the CoCo-ST algorithm on spatial transcriptomics datasets to detect biologically relevant spatial structures. We demonstrate CoCo-ST's ability to detect low variance spatial structures driving tumor progression.

Two main challenges with precancer analysis are:

- The high variance in global spatial patterns which might lead to the overlook of local, low-variance domains that hold significant biological relevance.
- The complexity of spatial domain detection across diverse tissue samples with varying spatial resolutions, necessitating a robust and scalable approach.

Here, we show CoCo-ST's contrastive learning approach for identifying both high and low-variance spatial structures, leveraging multi-sample data.

This protocol is associated with our proposed CoCo-ST algorithm for spatial domain detection in Nature Cell Biology

- 1 CoCoST Code: The CoCoST algorithm can be downloaded from the following GitHub repository:

Download CoCoST from GitHub

```
source("path/to/CoCoST.R")
```

- 2 Load Background Data (Normal Tissue)

```
normalTissue <- Load10X_Spatial(  
  data.dir = "X:/maminu/Processed 129S4 Urethane  
model/53430/outs",  
  filename = "filtered_feature_bc_matrix.h5",  
  assay = "Spatial",  
  filter.matrix = TRUE,  
  to.upper = FALSE)
```

- 3 Preprocess the data

```
normalTissue <- SCTransform(normalTissue, assay = "Spatial",  
verbose = FALSE)
```

- 4 Load Foreground Data (Abnormal Tissue)

```
abnormalTissue <- Load10X_Spatial(  
  data.dir = "X:/maminu/Processed 129S4 Urethane  
model/53433/outs",  
  filename = "filtered_feature_bc_matrix.h5",  
  assay = "Spatial",  
  filter.matrix = TRUE,  
  to.upper = FALSE)
```

- 5 Preprocess the data

```
abnormalTissue <- SCTransform(abnormalTissue, assay = "Spatial",  
verbose = FALSE)
```

6 Graph Construction

```
fdata <- abnormalTissue@assays[["SCT"]@scale.data
flocation <- GetTissueCoordinates(abnormalTissue)

bdata <- normalTissue@assays[["SCT"]@scale.data
blocation <- GetTissueCoordinates(normalTissue)
```

7 Construct affinity matrices

```
rbf <- laplacedot(sigma = 0.50)

fKernel <- kernelMatrix(rbf, t(fdata))
Wf <- fKernel@.Data

bKernel <- kernelMatrix(rbf, t(bdata))
Wb <- bKernel@.Data
```

8 Run CoCoST

```
para <- 0.10
Dim <- 20
gCPCA <- CoCoST(t(fdata), Wf, t(bdata), Wb, para, Dim)
```

9 Create Dimensionality Reduction Object in Seurat.

```
CoCo <- CreateDimReducObject(
  embeddings = gCPCA[["fgComponents"]],
  loadings = gCPCA[["projMatrix"]],
  stdev = numeric(),
  key = "CoCoST_",
  assay = "SCT")
rownames(CoCo@feature.loadings) <- rownames(fdata)
abnormalTissue@reductions[["CoCoST"]] <- CoCo
```

10 Run UMAP on CoCoST representations and perform clustering.

```
abnormalTissue <- RunUMAP(abnormalTissue, reduction = "CoCoST",
  dims = 1:10,
                                n.components = 5, reduction.name =
  "CoCo_UMAP")
abnormalTissue <- FindNeighbors(abnormalTissue, reduction =
  "CoCo_UMAP", dims = 1:5)
abnormalTissue <- FindClusters(abnormalTissue, verbose = TRUE,
  resolution = 0.05)
abnormalTissue@meta.data[["coco_clusters"]] <-
abnormalTissue@meta.data[["SCT_snn_res.0.05"]]
```

11 Visualize lower dimensional representations

```
cols <- c("#E1BD6D", "deepskyblue1", "#7A3A9A", "#ED0000FF",
  "#0B775E", "#ff00ff", "#7B556C", "#44B05B")
DimPlot(abnormalTissue, reduction = "CoCo_UMAP", label = FALSE,
  cols = cols)
```

12 Generate the spatial domain cluster plot:

```
SpatialDimPlot(abnormalTissue, label = FALSE, label.size = 3,
  group.by = "coco_clusters", pt.size.factor = 4) +
  scale_fill_manual(values = c("#E1BD6D", "deepskyblue1",
  "#7A3A9A", "#ED0000FF", "#0B775E", "#ff00ff", "#7B556C",
  "#44B05B"),
                    labels = c("Normal lung", "Fibrotic/Scar
  tissue", "Adjacent normal", "Bronchus/Alveoli",
                              "Adenoma", "Membrane"))
```