

# KSBi-BIML 2024

Bioinformatics & Machine Learning(BIML)  
Workshop for Life and Medical Scientists

**생명정보학 & 머신러닝 워크숍 (온라인)**



## Cell segmentation-free in situ spatial transcriptomics data analysis using SSAM

박정빈 \_ 부산대학교



**KSBI**  
KOREAN SOCIETY FOR  
BIOINFORMATICS

| 한국생명정보학회



본 강의 자료는 한국생명정보학회가 주관하는 BIML 2024 워크샵 온라인 수업을 목적으로 제작된 것으로 해당 목적 이외의 다른 용도로 사용할 수 없음을 분명하게 알립니다.

이를 다른 사람과 공유하거나 복제, 배포, 전송할 수 없으며 만약 이러한 사항을 위반할 경우 발생하는 **모든 법적 책임은 전적으로 불법 행위자 본인에게 있음을 경고**합니다.

# KSBI-BIML 2024

## Bioinformatics & Machine Learning(BIML) Workshop for Life and Medical Scientists

안녕하십니까?

한국생명정보학회가 개최하는 동계 교육 워크숍인 BIML-2024에 여러분을 초대합니다. 생명정보학 분야의 연구자들에게 최신 동향의 데이터 분석기술을 이론과 실습을 겸비해 전달하고자 도입한 전문 교육 프로그램인 BIML 워크숍은 2015년에 시작하여 올해로 벌써 10년 차를 맞이하게 되었습니다. BIML 워크숍은 국내 생명정보학 분야의 최초이자 최고 수준의 교육프로그램으로 크게 인공지능과 생명정보분석 두 개의 분야로 구성되어 있습니다. 올해 인공지능 분야에서는 최근 생명정보 분석에서도 응용이 확대되고 있는 다양한 인공지능 기반 자료모델링 기법들에 대한 현장 강의를 진행될 예정이며, 관련하여 심층학습을 이용한 단백질구조예측, 유전체분석, 신약개발에 대한 이론과 실습 강의를 함께 제공될 예정입니다. 또한 단일세포오믹스, 공간오믹스, 메타오믹스, 그리고 롱리드염기서열 자료 분석에 대한 현장 강의는 많은 연구자의 연구 수월성 확보에 큰 도움을 줄 것으로 기대하고 있습니다.

올해 BIML의 가장 큰 변화는 최근 연구 수요가 급증하고 있는 의료정보자료 분석에 대한 현장 강의를 추가하였다는 것입니다. 특히 의료정보자료 분석을 많이 수행하시는 의과학자 및 의료정보 연구자들께서 본 강좌를 통해 많은 도움을 받으실 수 있기를 기대하고 있습니다. 또한 다양한 생명정보학 분야에 대한 온라인 강좌 프로그램도 점차 증가하고 있는 생명정보 분석기술의 다양화에 발맞추기 위해 작년과 비교해 5강좌 이상을 신규로 추가했습니다. 올해는 무료 강좌 5개를 포함하여 35개 이상의 온라인 강좌가 개설되어 제공되며, 연구 주제에 따른 연관된 강좌 추천 및 강연료 할인 프로그램도 제공되며, 온라인을 통한 Q&A 세션도 마련될 예정입니다. BIML-2024는 국내 주요 연구 중심 대학의 전임 교원이자 각 분야 최고 전문가들의 강의로 구성되었기에 해당 분야의 기초부터 최신 연구 동향까지 포함하는 수준 높은 내용의 강의를 될 것이라 확신합니다.

BIML-2024을 준비하기까지 너무나 많은 수고를 해주신 운영위원회의 정성원, 우현구, 백대현, 김태민, 김준일, 김상우, 장혜식, 박종은 교수님과 KOBIC 이병욱 박사님께 커다란 감사를 드립니다. 마지막으로 부족한 시간에도 불구하고 강의 부탁을 흔쾌히 허락하시고 훌륭한 현장 강의와 온라인 강의를 준비하시는데 노고를 아끼지 않으신 모든 강사분들께 깊은 감사를 드립니다.

2024년 2월

한국생명정보학회장 이 인 석

## Cell segmentation-free in situ spatial transcriptomics data analysis using SSAM

최근 등장한 공간전사체 (Spatial transcriptomics) 분석은 기존 단일세포 전사체 (Single-cell transcriptomics) 기반 분석법에서는 쉽게 이해하지 못했던 공간상의 세포-세포 상호작용을 이해할 수 있는 지평을 열었다. 특히 최근 들어서는 Xenium, MERSCOPE, CosMX 등 in situ 이미징 기법에 기반한 몇가지 상업화된 도구들이 출시됨으로써 그 접근성이 개선되어 더욱 주목받고 있다.

본 강의에서는 이러한 in situ 기반 공간전사체 데이터를 분석하는 데에 있어 여러가지 한계점들 중 가장 큰 문제 중 하나로 지적되는 cell segmentation 없이도 공간상에 존재하는 cell type 을 분석할 수 있는 기법인 SSAM의 동작 원리를 설명하고, 이를 활용하여 데이터를 실제로 처리할 수 있는 방법을 설명함으로써 이를 통해 공간상에서의 cell type의 분포 등을 시각화 하는 방법을 배운다.

강의는 다음의 내용을 포함한다:

- 공간전사체 개요 및 기법 소개
- SSAM 의 동작 원리 설명
- SSAM 에 기반한 cell type 분석 방법 설명
- 분석 결과 시각화 방법 소개

\* 참고강의교재: 없음

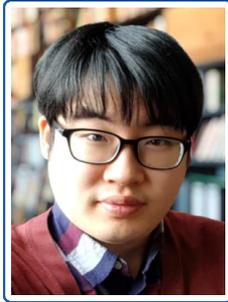
\* 교육생준비물: Google Colaboratory

\* 강의 난이도: 초급

\* 강의: 박정빈 교수 (부산대학교 의생명융합공학부)

# Curriculum Vitae

**Speaker Name: Dr. Jeongbin Park**



## ► Personal Info

Name Jeongbin Park  
Title Assistant Professor  
Affiliation Pusan National University

## ► Contact Information

Address Room# 411, Kyung-Ahm Engineering Bldg.,  
Busandaehak-ro 49, Mulgeum-eup, Yangsan-si,  
Gyeongsangnam-do, Republic of Korea  
Email jeongbin.park@pusan.ac.kr  
Phone Number 051-510-8597

---

## Research Interest

Translational bioinformatics, Machine learning and computational genomics

## Educational Experience

2012 B. Sc. in Physics, Pusan National University, Korea  
2014 M. Sc. in Physics, Seoul National University, Korea  
2020 Dr. rer. nat. in Bioscience, Heidelberg University, Germany

## Professional Experience

2018-2020 Researcher, Berlin Institute of Health (BIH) at Charite, Germany  
2020-2022 Postdoctoral Researcher, German Cancer Research Center (DKFZ), Germany  
2022- Assistant Professor, Pusan National University, Korea

## Selected Publications (5 maximum)

1. Jeongbin Park, et al., Cell segmentation-free inference of cell types from in situ transcriptomics data, *Nature Communications* 12, 3545, 2021
2. Sang-Tae Kim#, Jeongbin Park#, et al. Response to "Unexpected mutations after CRISPR-Cas9 editing in vivo", *Nature Methods* 15 (4), 239-240, 2018
3. Jeongbin Park, et al., Digenome-seq web tool for profiling CRISPR specificity, *Nature Methods* 14 (6), 548-549, 2017
4. Jeongbin Park, Sangsu Bae, Jin-Soo Kim, Cas-Designer: a web-based tool for choice of CRISPR-Cas9 target sites, *Bioinformatics* 31 (24), 4014-4016, 2015
5. Sangsu Bae#, Jeongbin Park#, Jin-Soo Kim, Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases, *Bioinformatics*, 30 (10), 1473-1475, 2014

# KSBi-BIML 2024

**Cell segmentation-free in situ spatial transcriptomics data analysis using SSAM**

부산대학교  
박정빈

## Overview

**Chapter 1.** Spatial Transcriptomics Overview & Methods

**Chapter 2.** Spatial Transcriptomics Data Analysis w/ SSAM

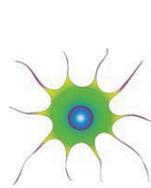
# Chapter 1.

## Spatial Transcriptomics Overview & Methods

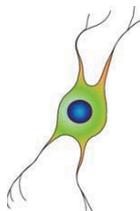
3

### Cell types

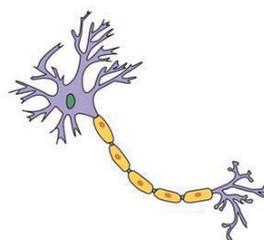
- What is cell type?



Astrocytes



Oligodendrocytes



Neurons

...

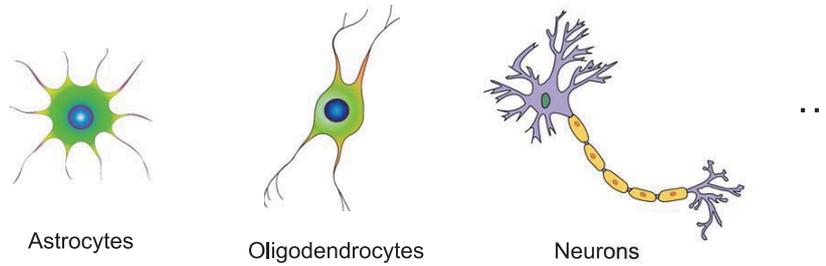
Cell type images from Wikimedia common

- Cells with different shapes and functions  
→ Cells having different gene expressions profiles

4

## Cell types

- What is cell type?

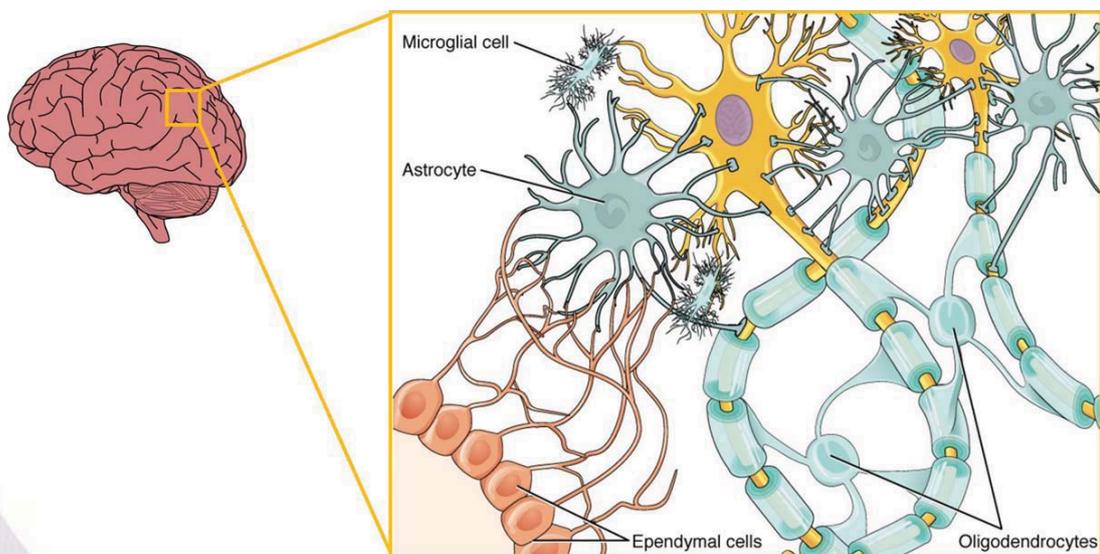


- Cells with different shapes and functions  
→ Cells having different gene expressions profiles

Really?

5

## Overview



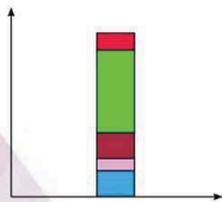
6

## Overview

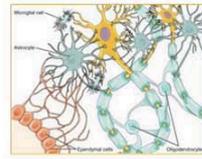
Bulk sequencing



Bulk Input



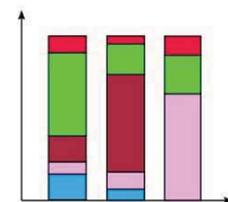
All cells are averaged



Single-cell sequencing



Microfluidics-based single cell barcoding



Difference between cells revealed

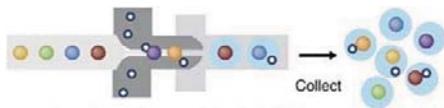
Single-cell omics (2015년경)

생명의 기본 단위인 세포를 각각 구분하여 시퀀싱이 가능해짐

7

## Single-cell omics

- 최신 미세유체학 기법 기반



- 각 세포들 간의 차이 확인 가능
- 거대 컨소시엄들이 생겨남

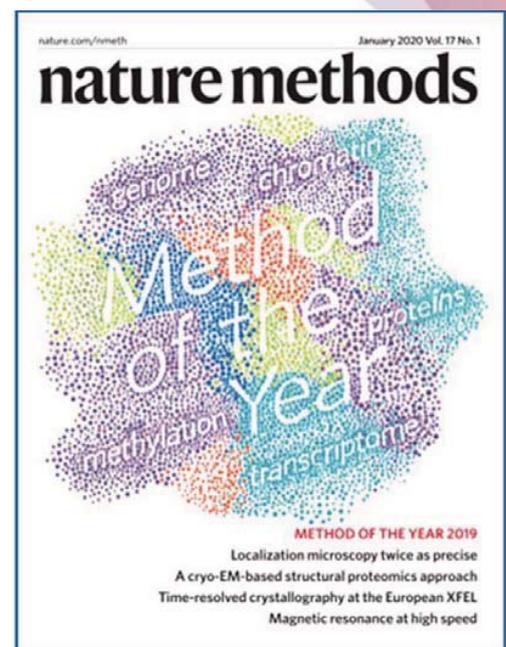


HUMAN  
CELL  
ATLAS

HCA|Organoid



인간의 모든 세포들의 지도를 만드는 프로젝트!



8

## Single-cell omics

- Scverse

- 단일세포 분석 필수 도구 모음
- <https://scverse.org/>

The image displays a grid of 10 tool cards for single-cell omics analysis. Each card includes a logo, the tool name, a brief description, and links to GitHub, documentation, and tutorials. The tools shown are: scanpy (scalable toolkit for analyzing single-cell gene expression data), anndata (Python package for handling annotated data matrices), muon (Python framework for multimodal omics analysis), mudata (format for annotated multimodal datasets), spatialdata (data framework for spatial omics), scvi-tools (library for developing and deploying machine learning models), scirpy (scalable toolkit for analyzing T-cell receptor or B-cell receptor repertoires), and squidpy (tool for analysis and visualization of spatial molecular data).

9

## Single-cell omics

- Scverse

- 단일세포 분석 필수 도구 모음
- <https://scverse.org/>

### Steering Council

The Steering Council (SC) consists of a fixed number of core team members who have additional responsibilities to ensure the smooth running of the project.

Daniila Bredikhin   Lukas Heumos   Isaac Virshup

### Management Committee

Management Committee is committed to supporting scverse and is directly involved in investing in its progress.

Francesca Finotello   Oliver Stegle   Fabian Theis   Alex Wolf   Nir Yosef

### Advisory Committee

Advisory Committee helps to shape the overall vision for scverse and to define its priorities.

Bonnie Berger   Dana Pe'er   Aviv Regev   Sarah Teichmann

10

## Human Cell Atlas

- 인간의 모든 세포들의 지도를 만드는 프로젝트
  - 현재 생물학계에서 가장 큰 프로젝트 중 하나
  - Aviv Regev, Sarah Teichmann 등 세계적인 생명정보학자들이 주도
    - <https://www.humancellatlas.org/learn-more/governance/>

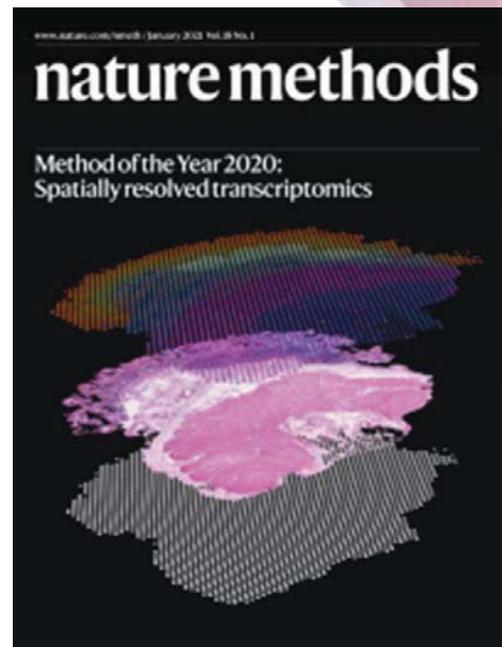


HUMAN  
CELL  
ATLAS

11

## Spatial Omics

- 단일세포체학 + 공간차원
  - 개별 세포들 간의 차이를 확인
- +
- 각 세포의 위치까지 확인가능
- 현재 가장 각광받고 있는 오믹스 기술!



12

## Spatial Omics

- SpaceTx 컨소시엄



- 공간전사체학 실험 및 분석 방법론 비교 분석 / 표준 정립
  - <https://spacetx.github.io/>
    - Our group: Final Analysis / Paper Writing 그룹
- Allen Institute for Brain Science 주도, Chan-Zuckerberg Initiative 지원
- 인간/생쥐 전체 뇌 공간체학 지도 제작 목표

13

## Spatial Omics

- Spverse

- 공간체학 분석 필수 도구 모음?
- <https://spatialomics.github.io/>

### Team



Prof. Oliver Stegle  
Supervisor



Prof. Moritz Gerstung  
Supervisor



Elyas Heidari  
Organizing Committee



Mostafa Shahhosseini  
Organizing Committee



14

## Spatial Omics

- Spverse
  - 공간체학 분석 필수 도구 모음?
  - <https://spatialomics.github.io/>

## Hackathon Schedule

Day 1

Wednesday, July 19th, 2023

Brainstorming

13:30 - 13:45: Reception
13:45 - 14:00: Openning Talk: <b>Elyas Heidari</b>
14:00 - 14:45: Openning Talk: <b>Dr. Oliver Stegle</b> and <b>Dr. Moritz Gerstung</b>
15:00 - 15:30: <b>Luca Marconato</b> , Status Quo of SpatialData framework for standadization of spatial data types and analysis
15:30 - 16:00: <b>Jeongbin Park</b> , Status Quo of SSAM for segmentation-free spatial annotation
16:00 - 16:15: Break
16:15 - 17:00: <b>Louis Kuemmerle</b> , Status Quo of TXsim, best practices of QC and panel design
17:00 - 17:30: <b>Florian Wuennemann</b> , Status Quo of MCMICRO for preprocessing, quality control, and workflow automation
17:30 - 18:00: <b>Vitalii Kleshchevnikov</b> , Cell-cell communication
18:00 - 18:30: <b>Artem Shmatko</b> , Diffusion models for digital histopathology and ideas towards integration of histopathology and transcriptomics
18:30 - 18:45: Break
18:45 - 19:30: Conclusion, Brain Storming and Team building >> <b>Session Leaders</b>

15

## Single Cell & Spatial Omics Korea (SCSOK)

1<sup>ST</sup> SEMI ANNUAL WORKSHOP **ibS**

### Single Cell- and Spatial Genomics

Date: 12:00 (KST), 3<sup>rd</sup> - 4<sup>th</sup>, January, 2023  
 Venue: Auditorium (S236) IBS Science Culture Center, Expo-ro 55, Yuseong-gu, Daejeon, Korea  
 Zoom: id > 812.7464.0930  
 pw > 2023CGE

**Day 1 (3<sup>rd</sup>, January)**

Session I	12:00 - 13:15	Bon-Kyoung Koo (IBS) & Secret Speaker
	13:15 - 14:00	Chang Ho Sohn (YSU)
Session II	14:00 - 14:45	Jong-Eun Park (KAIST)
	14:45 - 15:30	Min Kyu Yum (KAIST)
	15:30 - 15:50	Coffee Break
	15:50 - 16:35	Jeongbin Park (PNU)
Session III	16:35 - 17:20	Sekyu Choi (POSTECH)
	17:20 - 18:05	Ik Soo Kim (GMU)

**Day 2 (4<sup>th</sup>, January)**

Session III	09:15 - 10:00	Junil Kim (SSU)
	10:00 - 10:45	Stephen Preibisch (HHMI Janelia)
Session IV	10:45 - 11:30	Jong Kyoung Kim (POSTECH)
	11:30 - 12:30	Lunch Break
	12:30 - 13:15	Jihwan Park (GIST)
	13:15 - 14:00	Jinwook Choi (GIST)
	14:00 - 14:30	Closing Remarks

Contact us: Heetak Lee, leeheetak@ibs.re.kr  
 Bon-Kyoung Koo, koo@ibs.re.kr

CENTER for GENOME ENGINEERING



16

## Single Cell & Spatial Omics Korea (SCSOK)



17

## Spatial Transcriptomics Methods

### FISH-Based

1. Sequential RNA FISH
  - osmFISH
  - RNAScope
2. Combinatorial FISH
  - MERFISH
  - (MERSCOPE)
  - seqFISH+

### ISS-based

1. In Situ Sequencing (ISS)
2. Xenium (FISSEQ + ISS)
3. STARmap
4. ExSeq

### NGS-based

1. Visium  
(formerly **Spatial Transcriptomics**)
2. STomics (Stereo-seq)
3. Pixel-seq
4. Seq-Scope

18

## Sequential FISH-based Methods

osmFISH and RNAScope

19

### osmFISH

Brief Communication | [Published: 30 October 2018](#)

## Spatial organization of the somatosensory cortex revealed by osmFISH

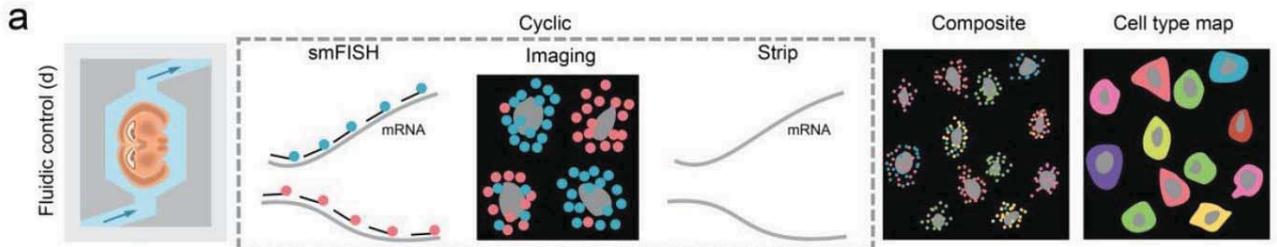
[Simone Codeluppi](#) , [Lars E. Borm](#), [Amit Zeisel](#), [Gioele La Manno](#), [Josina A. van Lunteren](#), [Camilla I. Svensson](#) & [Sten Linnarsson](#) 

[Nature Methods](#) **15**, 932–935 (2018) | [Cite this article](#)

**17k** Accesses | **227** Citations | **123** Altmetric | [Metrics](#)

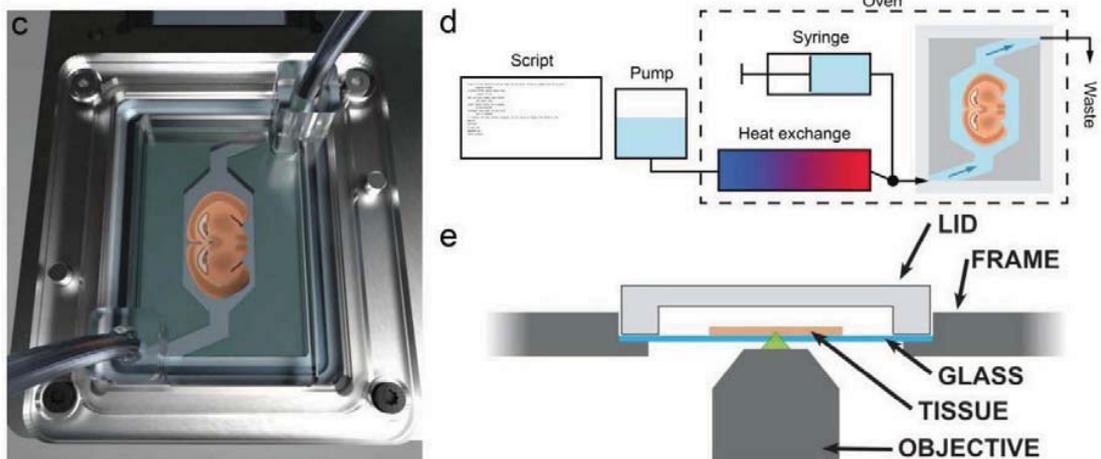
20

# osmFISH



21

# osmFISH



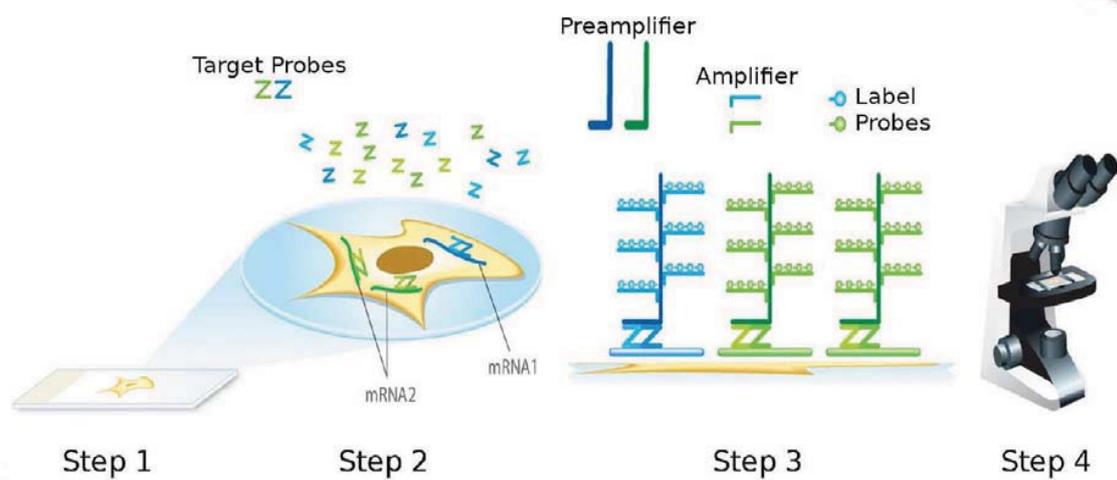
22

## RNAScope

The screenshot shows the ACD website homepage. At the top left is the ACD logo with the tagline "a bioelectron brand". To the right is a search bar with the text "Enter keywords to search" and a magnifying glass icon. Further right are links for "PLEASE SIGN IN" and "MY CART". Below the search bar is a navigation menu with the following items: Products, Services, Areas of Research, Technology, Support, Partners, Diagnostics, and Image Calendar. The main content area features a large image of a tissue section with colorful fluorescent spots. Overlaid on this image is the text: "Leader in RNA ISH Technology for Spatial Gene Expression". Below this is a paragraph: "RNAScope™ ISH is a pioneering multiplex fluorescent and chromogenic *in situ* hybridization technology. ACD's patented probe design amplifies target-specific signals, *in situ*, enabling detection and quantification of the single RNA molecule providing spatial gene expression results."

23

## RNAScope



24

## Combinatorial FISH-based Methods

MERFISH (MERSCOPE) and seqFISH+

25

## MERFISH (MERSCOPE)

 RESEARCH ARTICLE

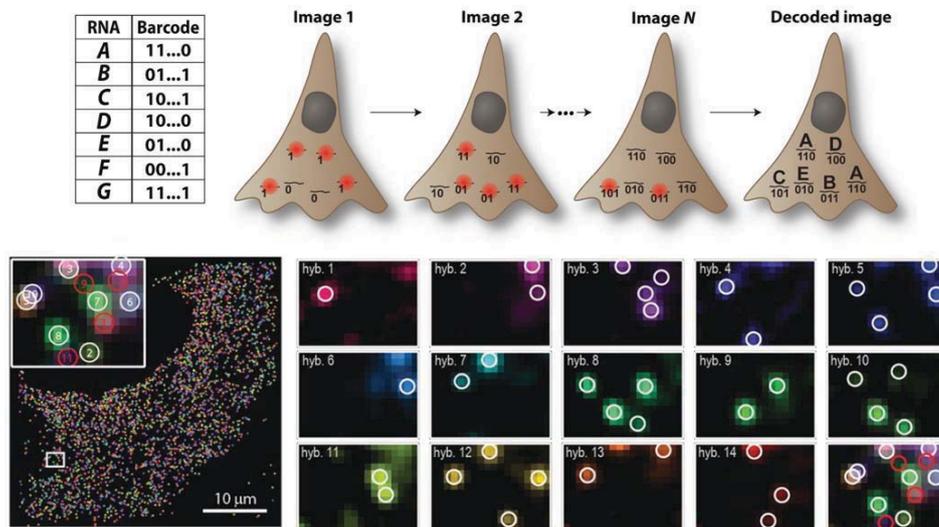
# Spatially resolved, highly multiplexed RNA profiling in single cells

KOK HAO CHEN, ALISTAIR N. BOETTIGER, JEFFREY R. MOFFITT, SIYUAN WANG, AND XIAOWEI ZHUANG [Authors Info & Affiliations](#)

SCIENCE • 9 Apr 2015 • Vol 348, Issue 6233 • DOI: 10.1126/science.aaa6090

26

## MERFISH (MERSCOPE)



27

## SeqFISH+

Letter | [Published: 25 March 2019](#)

# Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+

[Chee-Huat Linus Eng](#), [Michael Lawson](#), [Qian Zhu](#), [Ruben Dries](#), [Noushin Koulena](#), [Yodai Takei](#), [Jina Yun](#), [Christopher Cronin](#), [Christoph Karp](#), [Guo-Cheng Yuan](#) & [Long Cai](#) ✉

[Nature](#) **568**, 235–239 (2019) | [Cite this article](#)

**93k** Accesses | **645** Citations | **293** Altmetric | [Metrics](#)

28



## In Situ Sequencing (ISS)

Published: 14 July 2013

# *In situ* sequencing for RNA analysis in preserved tissue and cells

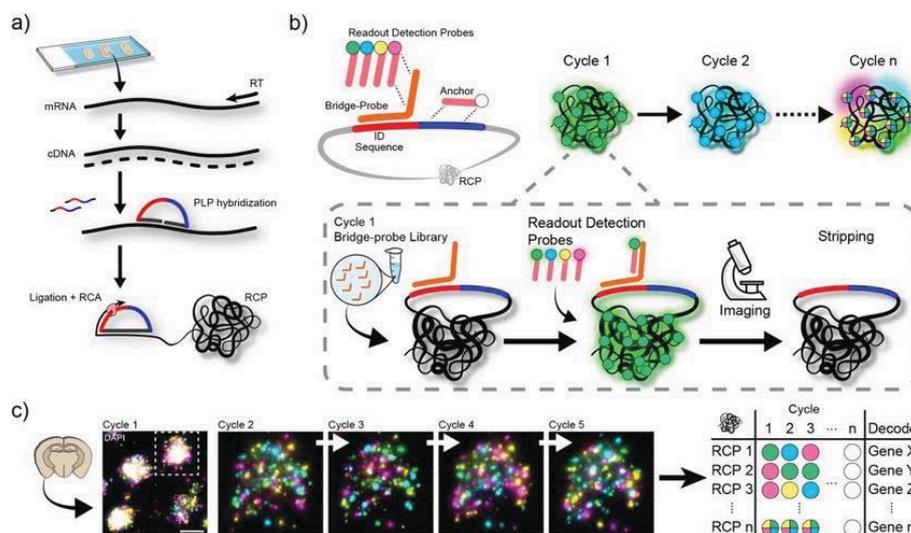
Rongqin Ke, Marco Mignardi, Alexandra Pacureanu, Jessica Svedlund, Johan Botling, Carolina Wählby & Mats Nilsson

*Nature Methods* 10, 857–860 (2013) | [Cite this article](#)

42k Accesses | 419 Citations | 84 Altmetric | [Metrics](#)

31

## In Situ Sequencing (ISS)



Gyllborg *et al*, *Nucleic Acids Research* (2020)

32

## Fluorescent in situ sequencing (FISSEQ)

REPORT



# Highly Multiplexed Subcellular RNA Sequencing in Situ

[JE HYUK LEE](#), [EVAN R. DAUGHARTHY](#), [JONATHAN SCHEIMAN](#), [REZA KALHOR](#), [JOYCE L. YANG](#), [THOMAS C. FERRANTE](#), [RICHARD TERRY](#), [SAUVEUR S. F. JEANTY](#), [CHAO LI](#), [RYOJI AMAMOTO](#), [DEREK T. PETERS](#), [BRIAN M. TURCZYK](#), [ADAM H. MARBLESTONE](#), [SAMUEL A. INVERSO](#), [AMY BERNARD](#), [PRASHANT MALI](#), [XAVIER RIOS](#), [JOHN AACH](#), AND [GEORGE M. CHURCH](#) [fewer](#) [Authors Info & Affiliations](#)

33

## 10X Xenium

10X GENOMICS

Products Area of Interest Resources Support Company Careers Search

Xenium In Situ

### High-performance in situ from the single cell leader

[Request pricing](#) [See performance](#)

ERBB2 (HER2)  
ESR1 (ER)  
PGR (PR)

In this breast cancer sample (Stage II-B, ER+/PR-/HER2+), Xenium identifies a previously unknown triple-positive region.

34

## 10X Xenium

# 10x Genomics to Acquire ReadCoor for \$350M, Cartana for \$41.2M

Oct 05, 2020 | [staff reporter](#)

*This story has been updated to include the Cartana acquisition price.*

NEW YORK – 10x Genomics announced Monday that it will acquire two firms involved in developing *in situ* analysis technologies: ReadCoor and Cartana.

10x has entered into a definitive agreement to buy Boston-based ReadCoor, a spinout from George Church's laboratory at Harvard Medical School and the Wyss Institute, for \$350 million in cash and stock. ReadCoor [launched its RC2 spatial multiomics platform](#), based on [fluorescent \*in situ\* sequencing technology](#), in February at the Advances in Genome Biology and Technology conference. 10x said it expects the deal to close by the end of the month, pending ReadCoor shareholder approval. Goldman Sachs is acting as financial advisor to ReadCoor.

10x also revealed it acquired Stockholm-based [Cartana](#), developers of *in situ* RNA analysis technology, in late August for \$41.2 million. Other details of that transaction were not disclosed.

<https://www.genomeweb.com/sequencing/10x-genomics-acquire-readcoor-350m-cartana-412m>

## 10X Xenium = ISS + FISSEQ

35

## StarMAP

 RESEARCH ARTICLE

[f](#) [t](#) [in](#) [r](#) [s](#) [e](#)

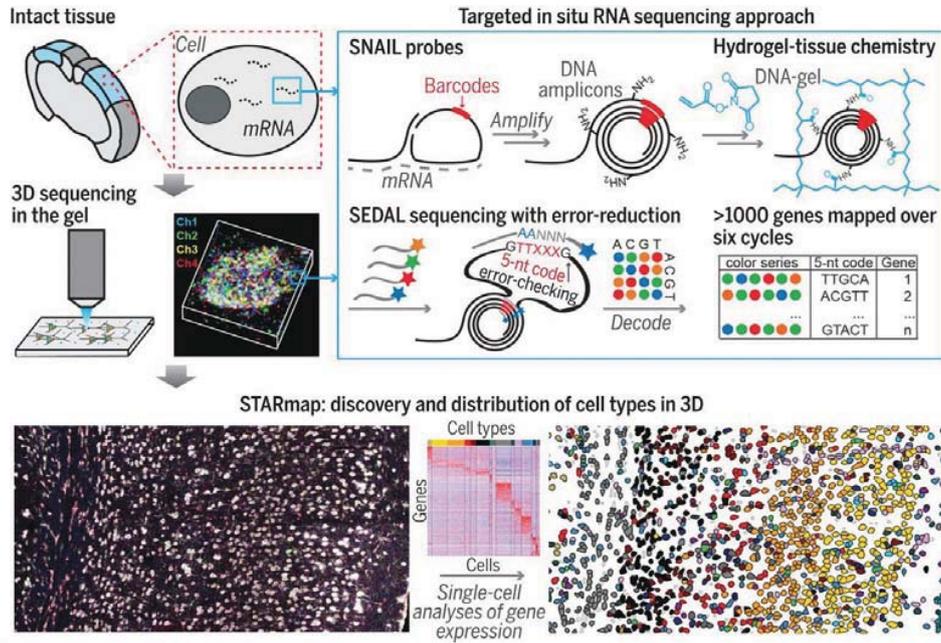
# Three-dimensional intact-tissue sequencing of single-cell transcriptional states

XIAO WANG , WILLIAM E. ALLEN , MATTHEW A. WRIGHT , EMILY L. SYLWESTRAK, NIKOLAY SAMUSIK , SAM VESUNA, KATHRYN EVANS , CINDY LIU ,

CHARU RAMAKRISHNAN , JIA LIU , GARRY P. NOLAN , FELICE-ALESSIO BAVA , AND KARL DEISSEROTH  [fewer](#) [Authors Info & Affiliations](#)

36

## StarMAP



37

## ExSeq

RESEARCH ARTICLE



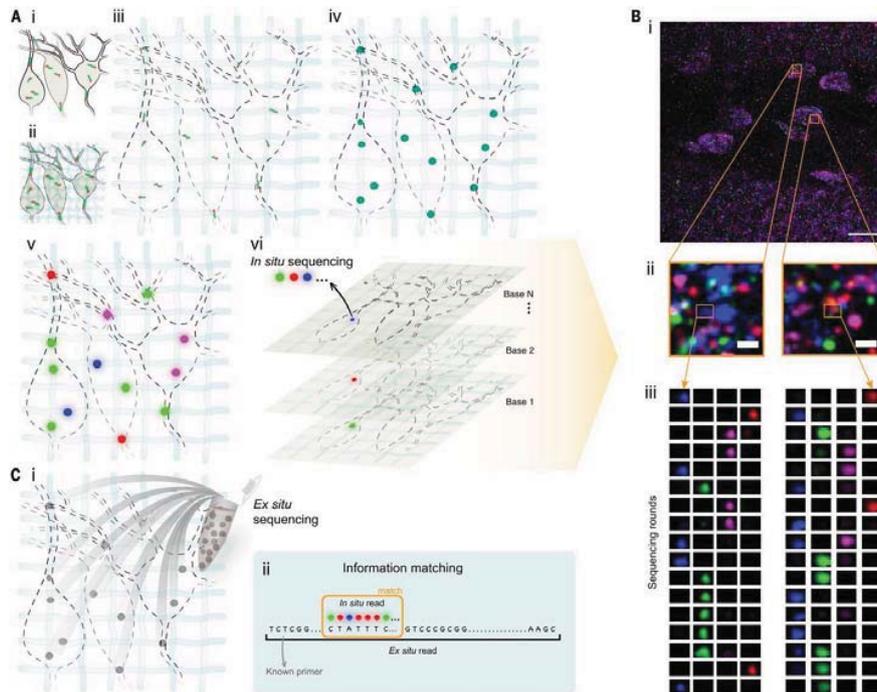
# Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems

SHAHAR ALON , DANIEL R. GOODWIN , ANUBHAV SINHA , ASMAMAW T. WASSIE, FEI CHEN, EVAN R. DAUGHARTHY , YOSUKE BANDO , ATSUSHI KAJITA, ANDREW G. XUE , KARL MARRETT, ROBERT PRIOR , YI CUI , ANDREW C. PAYNE , CHUN-CHEN YAO, HO-JUN SUK, RU WANG, CHIH-CHIEH (JAY) YU , PAUL TILLBERG , PAUL REGINATO , NIKITA PAK , SONGLEI LIU , SUKANYA PUNTHAMBAKER , ESWAR P. R. IYER , RICHIE E. KOHMAN , JEREMY A. MILLER , ED S. LEIN , ANA LAKO , NICOLE CULLEN, SCOTT RODIG, KARLA HELVIE, DANIEL L. ABRAVANEL , NIKHIL WAGLE , BRUCE E. JOHNSON, JOHANNA KLUGHAMMER , MICHAL SLYPER , JULIA WALDMAN , JUDIT JANÉ-VALBUENA , ORIT ROZENBLATT-ROSEN , AVIV REGEV , IMAXT CONSORTIUM, GEORGE M. CHURCH , ADAM H. MARBLESTONE, AND EDWARD S. BOYDEN

[fewer](#) [Authors Info & Affiliations](#)

38

## ExSeq



39

## Spatial Transcriptomics Methods

### NGS-based Methods

Visium, STOmics, Seq-Scope

40

# Visium

10x GENOMICS Products Area of Interest Resources Support Company Careers Search

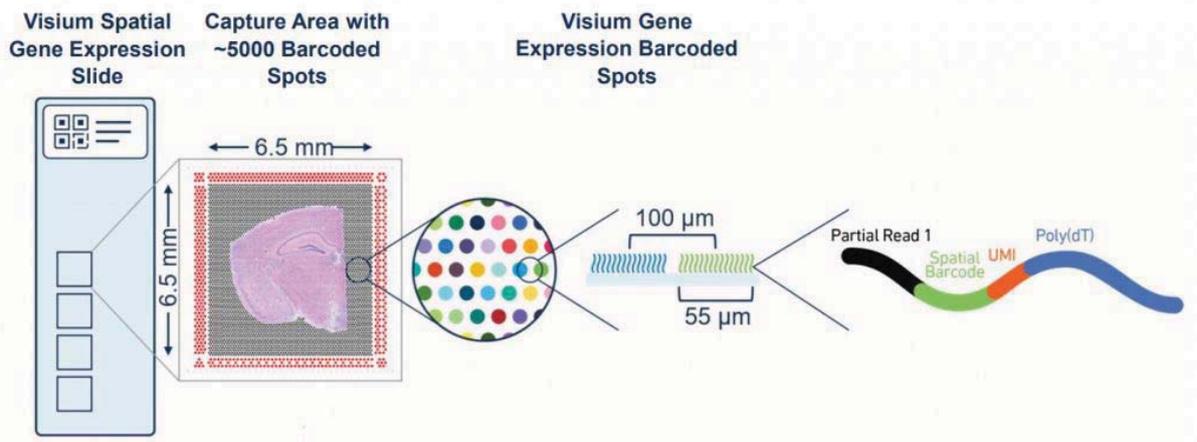
## Visium Spatial Gene Expression

### Map the whole transcriptome within the tissue context

Visium Spatial Gene Expression is a next-generation molecular profiling solution for classifying tissue based on total mRNA. Map the whole transcriptome with morphological context in FFPE or fresh frozen tissues to discover novel insights into normal development, disease pathology, and clinical translational research.

[View Pricing](#) [See how it works ▶](#)

# Visium



## STOmics (Stereo-seq)

RESOURCE | VOLUME 185, ISSUE 10, P1777-1792.E21, MAY 12, 2022

Download Full Issue

# Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays

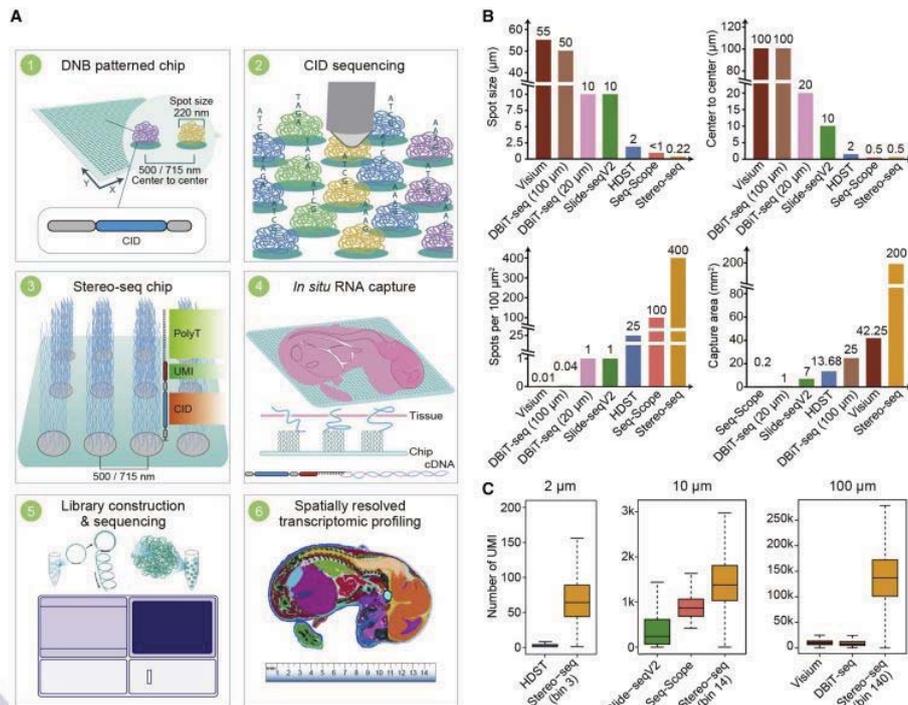
Ao Chen <sup>26</sup> • Sha Liao <sup>26</sup> • Mengnan Cheng <sup>26</sup> • Kailong Ma <sup>26</sup> • Liang Wu <sup>26</sup> • Yiwei Lai <sup>26</sup> • Xiaojie Qiu <sup>26</sup> • Jin Yang • Jiangshan Xu • Shijie Hao • Xin Wang • Huifang Lu • Xi Chen • Xing Liu • Xin Huang • Zhao Li • Yan Hong • Yuja Jiang • Jian Peng • Shuai Liu • Mengzhe Shen • Chuanyu Liu • Quanshui Li • Yue Yuan • Xiaoyu Wei • Huiwen Zheng • Weimin Feng • Zhifeng Wang • Yang Liu • Zhaohui Wang • Yunzhi Yang • Haitao Xiang • Lei Han • Baoming Qin • Pengcheng Guo • Guangyao Lai • Pura Muñoz-Cánoves • Patrick H. Maxwell • Jean Paul Thiery • Qing-Feng Wu • Fuxiang Zhao • Bichao Chen • Mei Li • Xi Dai • Shuai Wang • Haoyan Kuang • Junhou Hui • Liqun Wang • Ji-Feng Fei • Ou Wang • Xiaofeng Wei • Haorong Lu • Bo Wang • Shiping Liu • Ying Gu • Ming Ni • Wenwei Zhang • Feng Mu • Ye Yin • Huanming Yang • Michael Lisby • Richard J. Cornall • Jan Mulder • Mathias Uhlén • Miguel A. Esteban <sup>✉</sup> • Yuxiang Li <sup>✉</sup> • Longqi Liu <sup>✉</sup> • Xun Xu <sup>✉</sup> <sup>27</sup> • Jian Wang <sup>✉</sup> • Show less • Show footnotes

Open Access • Published: May 04, 2022 • DOI: <https://doi.org/10.1016/j.cell.2022.04.003>

Check for updates

43

## STOmics (Stereo-seq)



44

## Seq-Scope

RESOURCE | VOLUME 184, ISSUE 13, P3559-3572.E22, JUNE 24, 2021

[Download Full Issue](#)

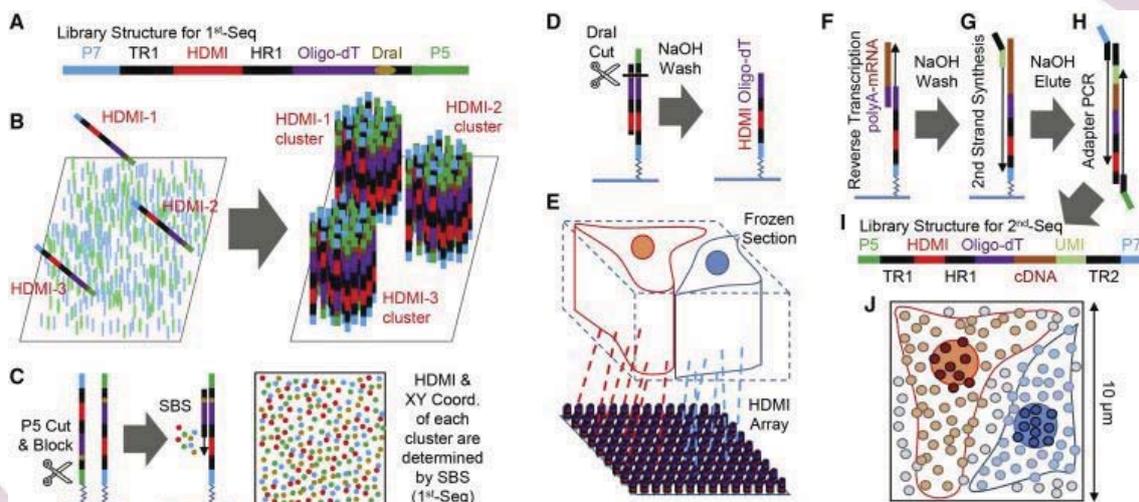
# Microscopic examination of spatial transcriptome using Seq-Scope

Chun-Seok Cho<sup>4</sup> • Jingyue Xi<sup>4</sup> • Yichen Si • Sung-Rye Park • Jer-En Hsu • Myungjin Kim • Goo Jun • Hyun Min Kang • Jun Hee Lee<sup>5</sup> [✉](#) • [Show less](#) • [Show footnotes](#)

[Open Archive](#) • Published: June 10, 2021 • DOI: <https://doi.org/10.1016/j.cell.2021.05.010> •

45

## Seq-Scope



46

## Methods Summary

### FISH-Based

- Resolution
  - 150 - 300nm
- Efficiency
  - Sequential: ~100%
  - Combinatorial: 70 - 95%
- Number of genes
  - 10 - several hundreds
- Spatial throughput
  - several mm<sup>2</sup>

### ISS-based

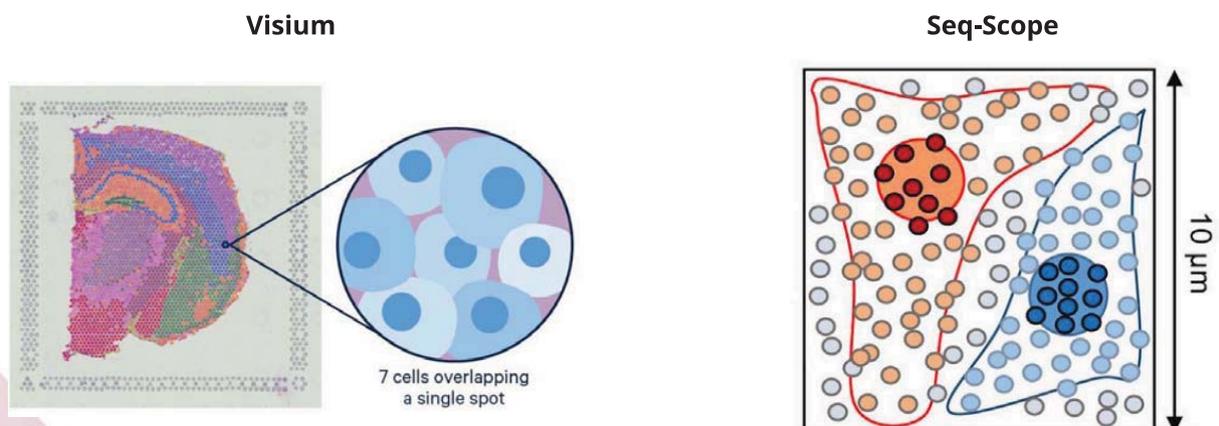
- Resolution
  - 0.5 - 1µm
- Efficiency
  - ISS: ~5%
  - FISSEQ: ~0.005%
  - ExSeq: 62% (compared to FISH)
- Number of genes
  - 100 - 1,000
- Spatial throughput
  - several mm<sup>2</sup> - cm<sup>2</sup>

### NGS-based

- Resolution
  - 0.22 - 100µm
- Efficiency
  - Visium: 6.9% or higher
  - Seq-Scope: 2.8x ~ 15.6x of Visium
  - Stereo-Seq: 0.2x of Seq-Scope
- Number of genes
  - Full transcriptome
- Spatial throughput
  - several cm<sup>2</sup>

47

## Spatial Resolution of Visium

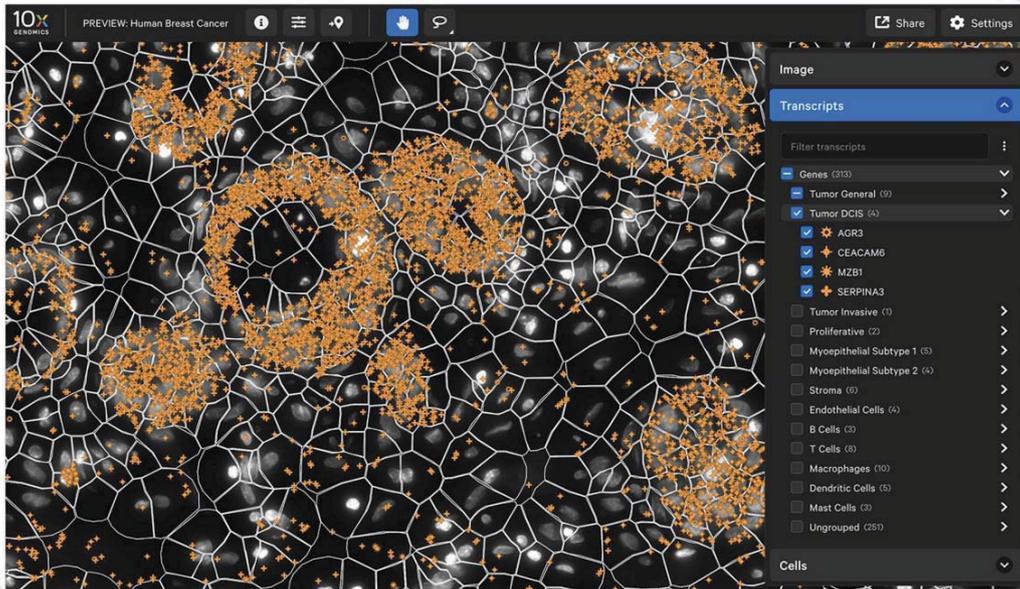


When the spatial resolution is not sufficient to distinguish single cells (e.g., Visium), you will need **deconvolution methods (e.g., Cell2location, Tangram, etc.)**

Read this benchmark result:  
<https://www.nature.com/articles/s41592-022-01481-8>

48

## Cell segmentation



**정확한 세포의 경계 찾기 (Cell Segmentation) 가 매우 중요!**

“Baysor” can improve cell segmentation, read:

<https://www.10xgenomics.com/cn/resources/analysis-guides/using-baysor-to-perform-xenium-cell-segmentation>

49

## Xenium vs. MERSCOPE vs. CosMx SMI

New Results

Follow this preprint

### Systematic benchmarking of imaging spatial transcriptomics platforms in FFPE tissues

Huan Wang, Ruixu Huang, Jack Nelson, Ce Gao, Miles Tran, Anna Yeaton, Kristen Felt, Kathleen L. Pfaff, Teri Bowman, Scott J. Rodig, Kevin Wei, Brittany A. Goods, Samouil L. Farhi  
doi: <https://doi.org/10.1101/2023.12.07.570603>

This article is a preprint and has not been certified by peer review [what does this mean?].

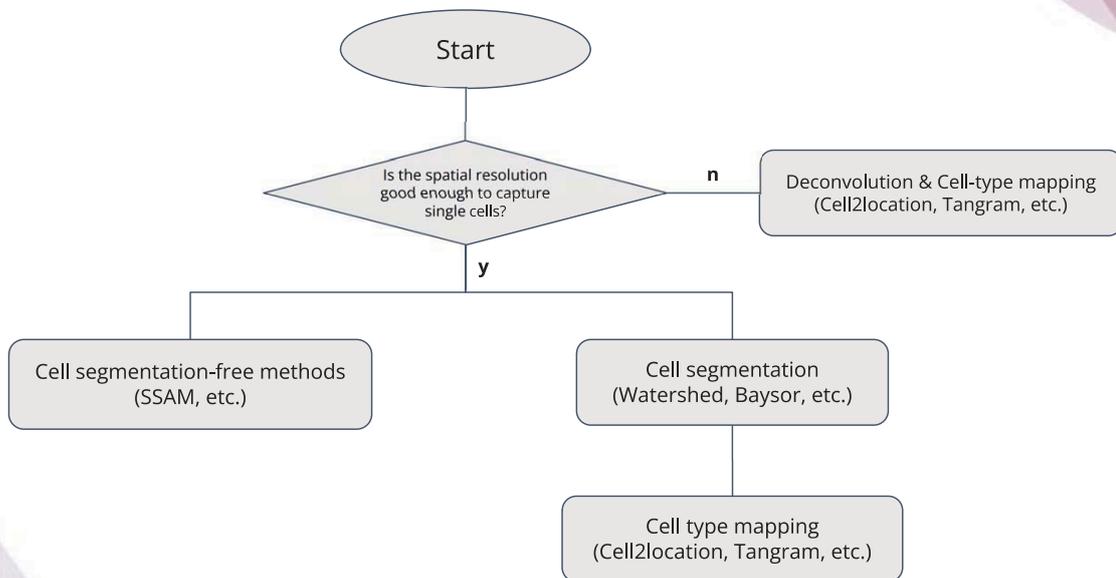
“Xenium and MERSCOPE provide more reliable true-positive signals of lowly expressed genes and that Xenium’s overall performance is less dependent on sample input quality than the other two platforms.”

“Overall, it appears that the out-of-the-box segmentation from Xenium performs poorly in terms of drawing cell boundaries specific to a single cell, while MERSCOPE and CosMx much more closely match cell boundaries.”

<https://www.biorxiv.org/content/10.1101/2023.12.07.570603v2>

50

## Spatial Omics Data Analysis Workflow

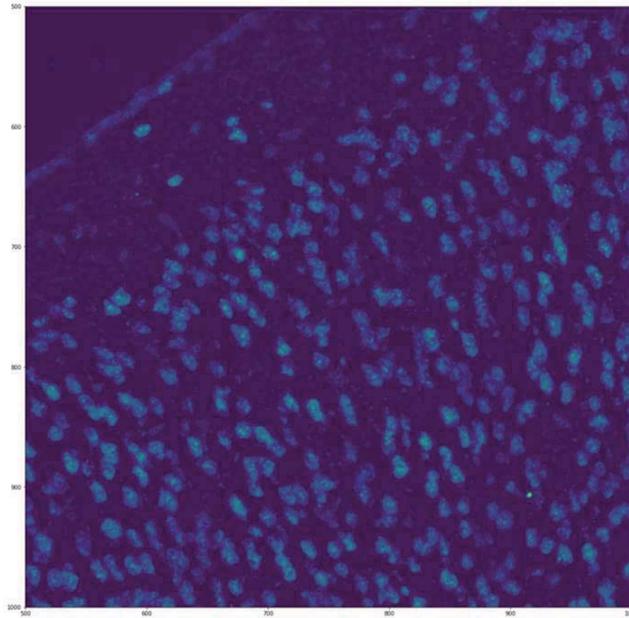


51

## Chapter 2. Spatial Transcriptomics Data Analysis w/ SSAM

52

## Spatial Transcriptomics



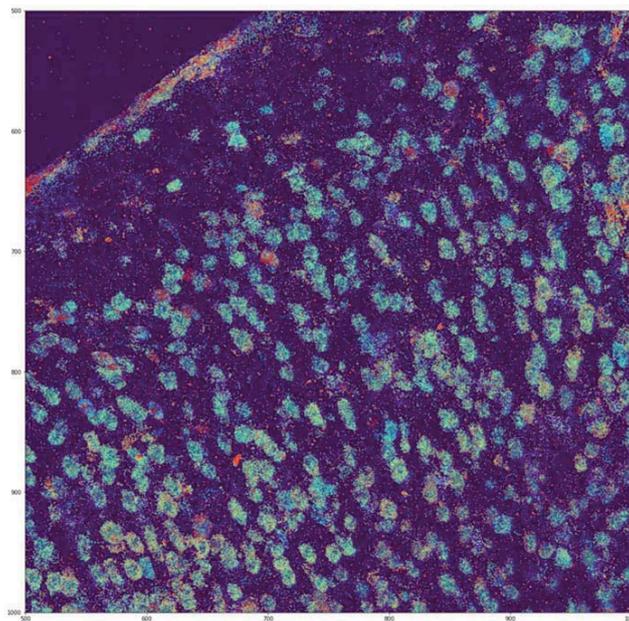
Poly-A image (all mRNAs)

53

## Spatial Transcriptomics

### Methods

- osmFISH
- MERFISH
- seqFISH+

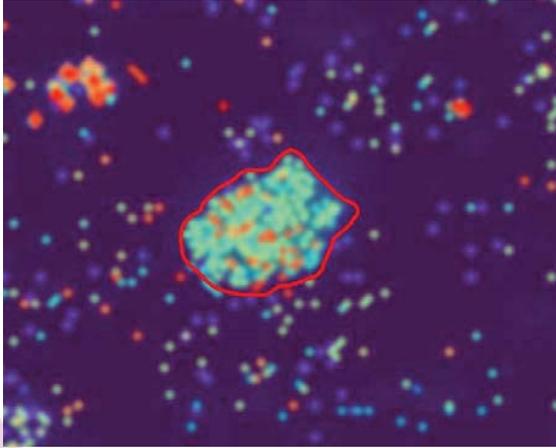


Scatter plot of mRNAs  
(33 genes, colored by genes)

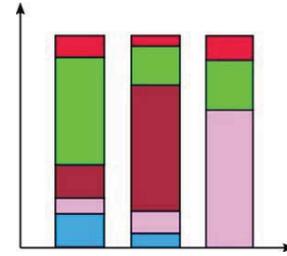
Poly-A image (all mRNAs)

54

## Spatial Transcriptomics



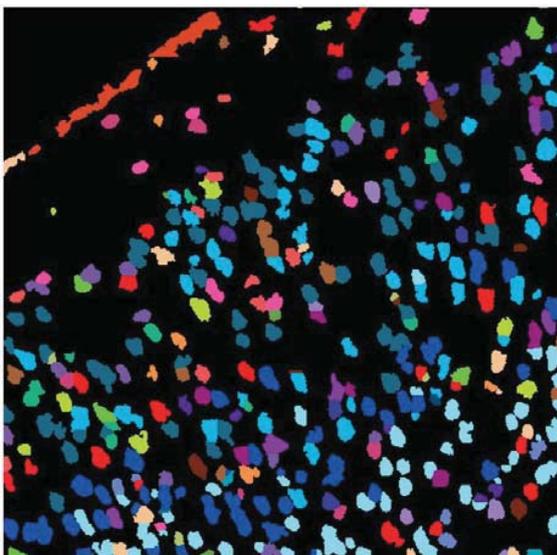
1. 세포의 경계 찾기 (Segmentation)
2. 경계 안의 mRNA 개수 세기



단일세포전사체학 데이터와 같은 방식으로 분석 가능!

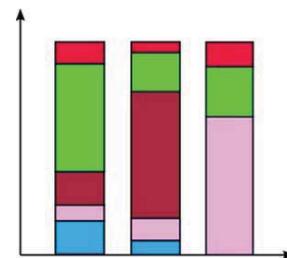
55

## Spatial Transcriptomics



Cell segments, colored by cell types

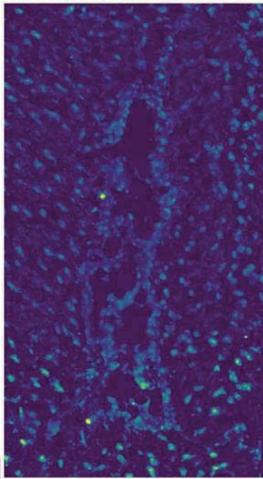
1. 세포의 경계 찾기 (Segmentation)
2. 경계 안의 mRNA 개수 세기



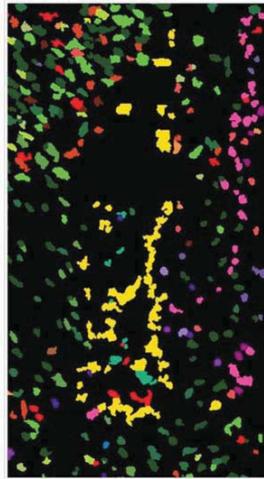
단일세포전사체학 데이터와 같은 방식으로 분석 가능!

56

## 현 공간체학 분석방법의 문제점



Poly-A Image  
(all mRNAs)



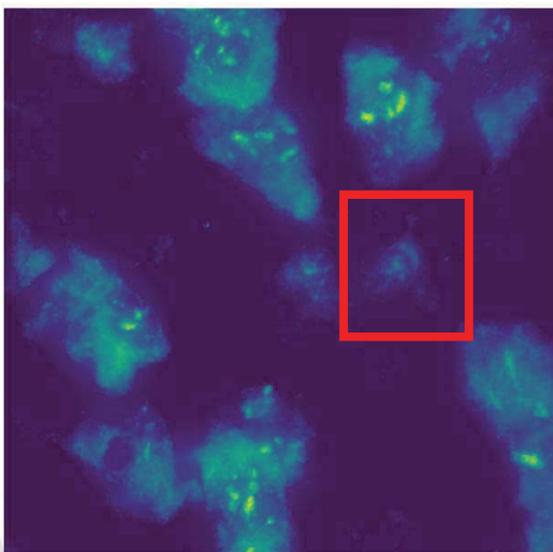
Cell segments  
colored by cell types

- Segmentation 이 쉽지 않음
  - Segmentation을 하지 않고 공간상에서의 cell-type을 분석할 수 있는 방법론 (SSAM) 개발

Park et al. *Nature Communications* **12**, 3545 (2021)

57

## 현 공간체학 분석방법의 문제점

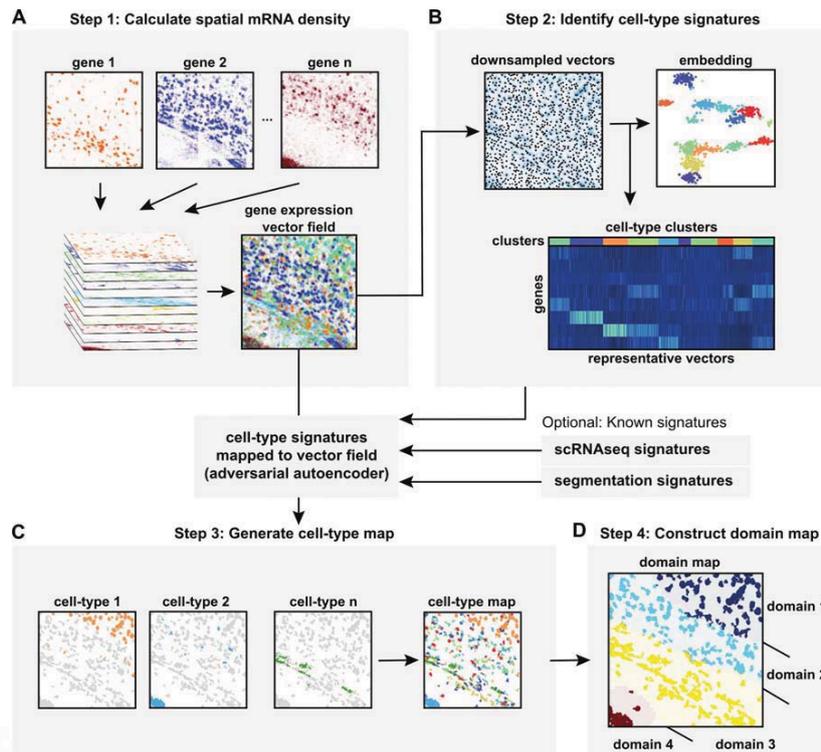


- Segmentation 이 쉽지 않음
  - Segmentation을 하지 않고 공간상에서의 cell-type을 분석할 수 있는 방법론 (SSAM) 개발

Park et al. *Nature Communications* **12**, 3545 (2021)

58

## SSAM



59

## Kernel Density Estimation

- $i$  번째 mRNA 상에 위치한 커널의 합으로 공간 밀도를 추정

$$\hat{\sigma}_h(\vec{x}) = \frac{1}{Nh} \sum_i^N K\left(\frac{\vec{x} - \vec{x}_i}{h}\right)$$

- $\sigma$  는 밀도,  $N$  은 mRNA 개수,  $K$  는 커널,  $h$  는 bandwidth 파라미터,  $x_i$  는  $i$  번째 mRNA 위치

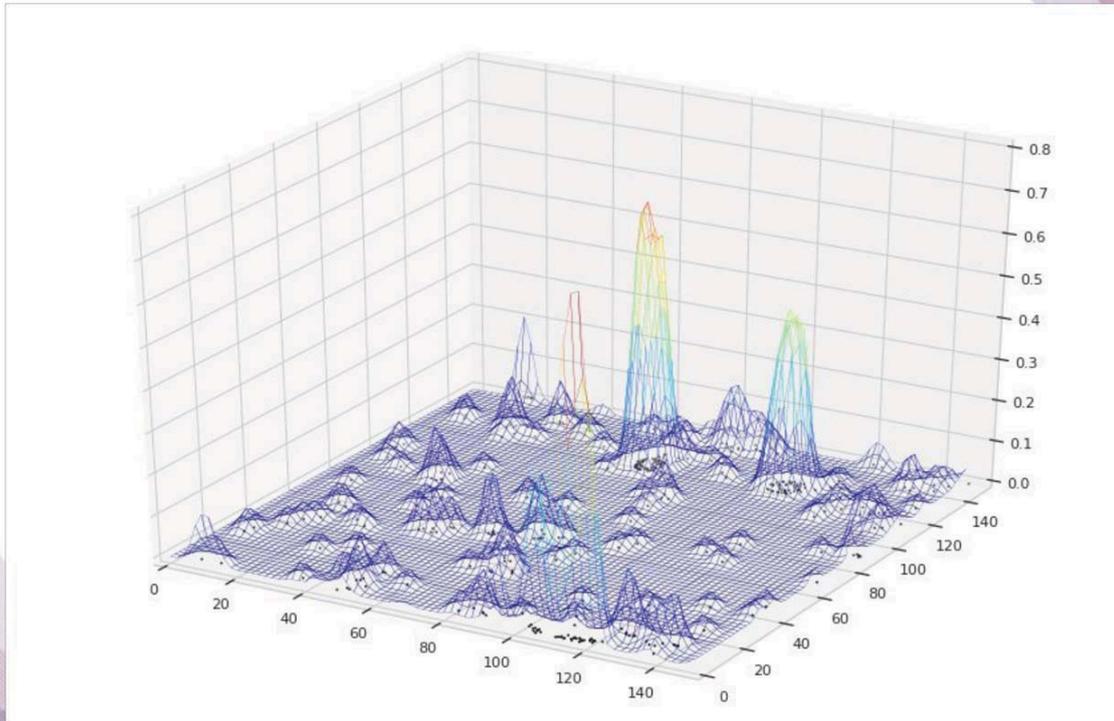
- 가우시안 커널

$$K_{gaussian}(\vec{x}) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}|\vec{x}|^2\right)$$

- $2 * \text{FWHM}(\sigma)$  이 세포의 지름과 비슷해지게 만드는 bandwidth 파라미터  $h$  선택 (세포의 지름  $\sim 10\mu\text{m}$  으로 만들어주는  $h=2$  로 선택함)

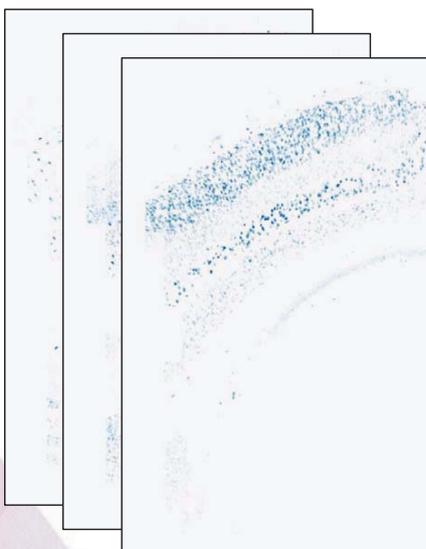
60

## Kernel Density Estimation



61

## Kernel Density Estimation



- 33 종류의 유전자를 쌓아올림  
→ 33 차원 벡터필드

- $i$ 번째 유전자 발현량 ( $E_i$ ):

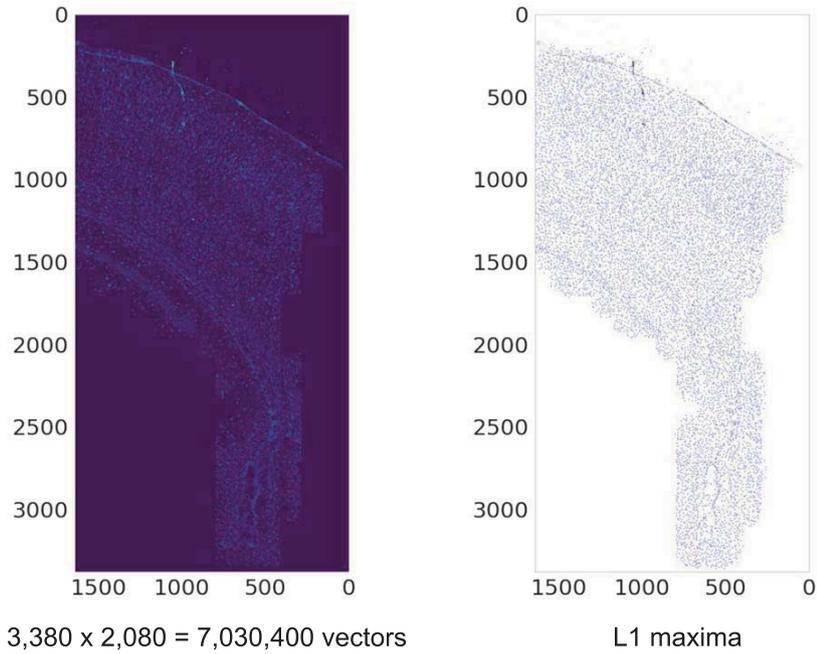
$$\hat{E}_i(\vec{x}) = \hat{\sigma}_i(\vec{x}) N_i$$

where

$E_i$ : expression of  $i$ -th gene  
 $\sigma_i$ : estimated density of  $i$ -th gene  
 $N_i$ : number of mRNAs of  $i$ -th gene

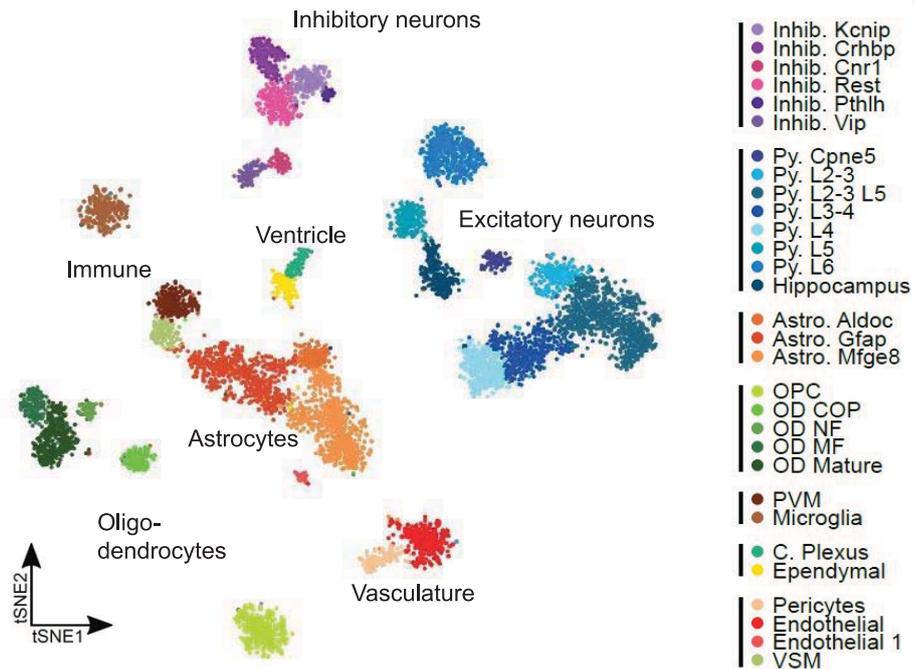
62

## 다운샘플링



63

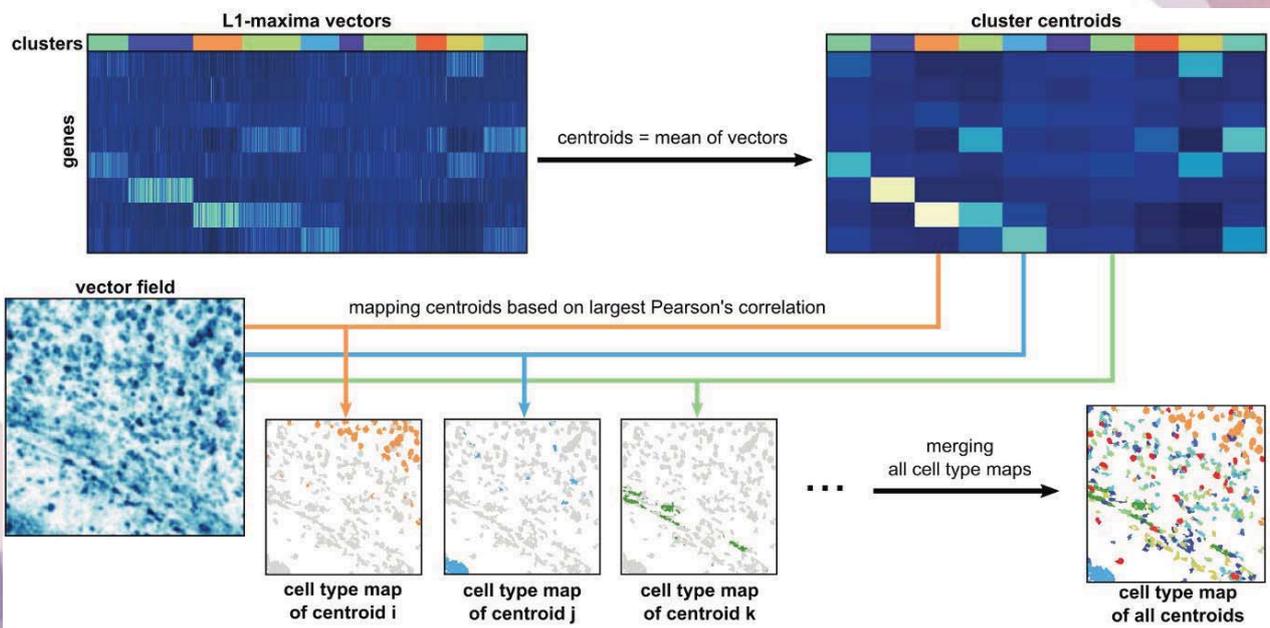
## 클러스터링



주의: 각 점은 “벡터” 임 (세포가 아님)

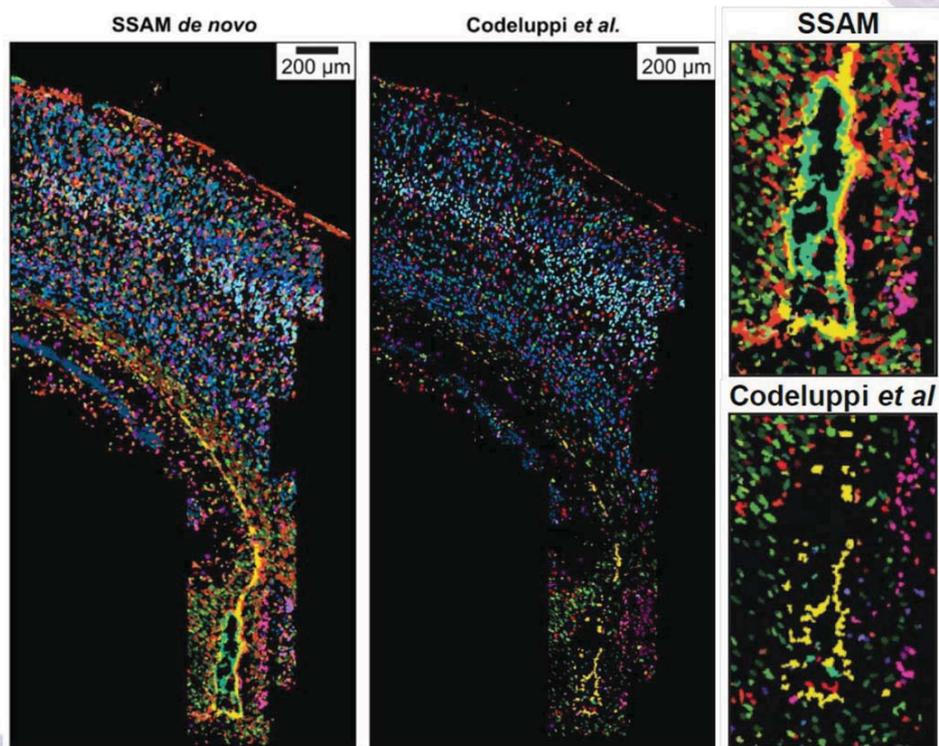
64

## 세포 타입 맵 (Cell-type map) 생성



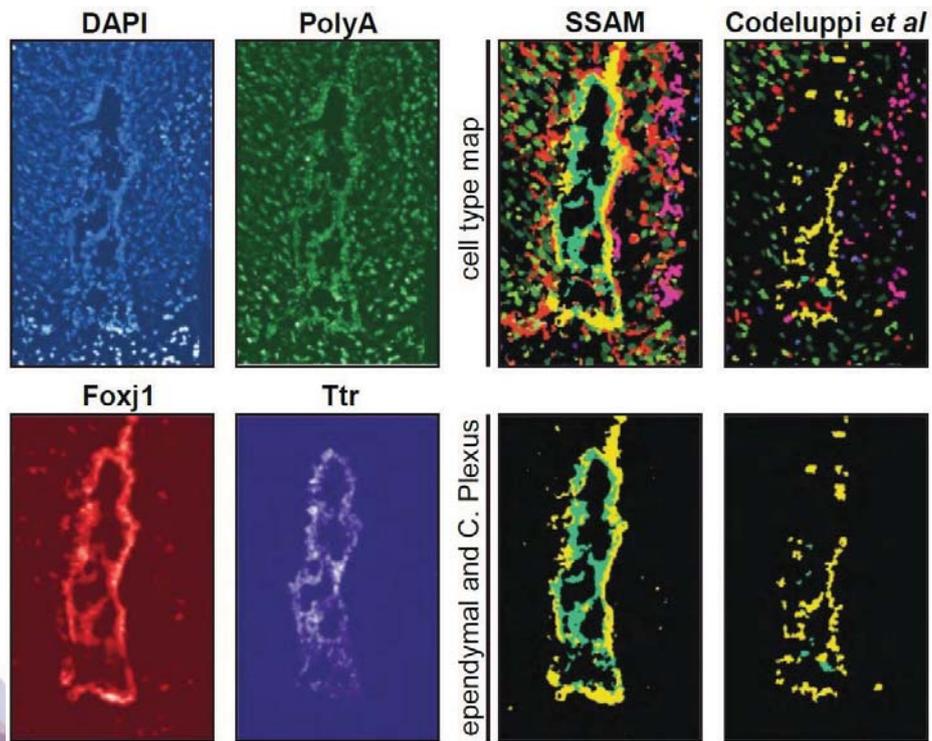
65

## 세포 타입 맵 (Cell-type map)



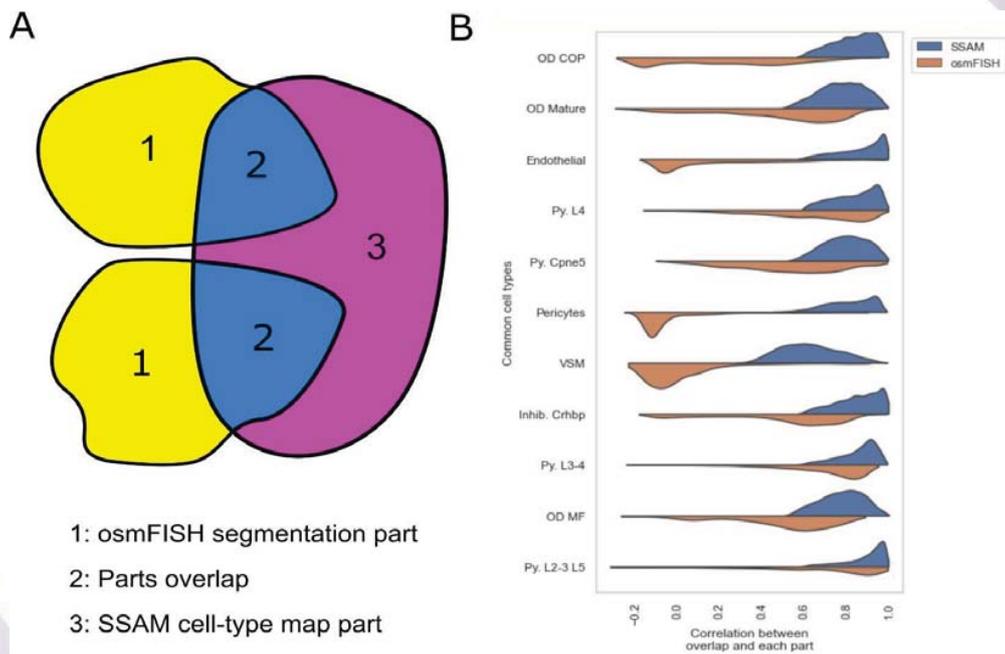
66

## 세포 타입 맵 (Cell-type map) - Ventricle 구조



67

## 성능 비교

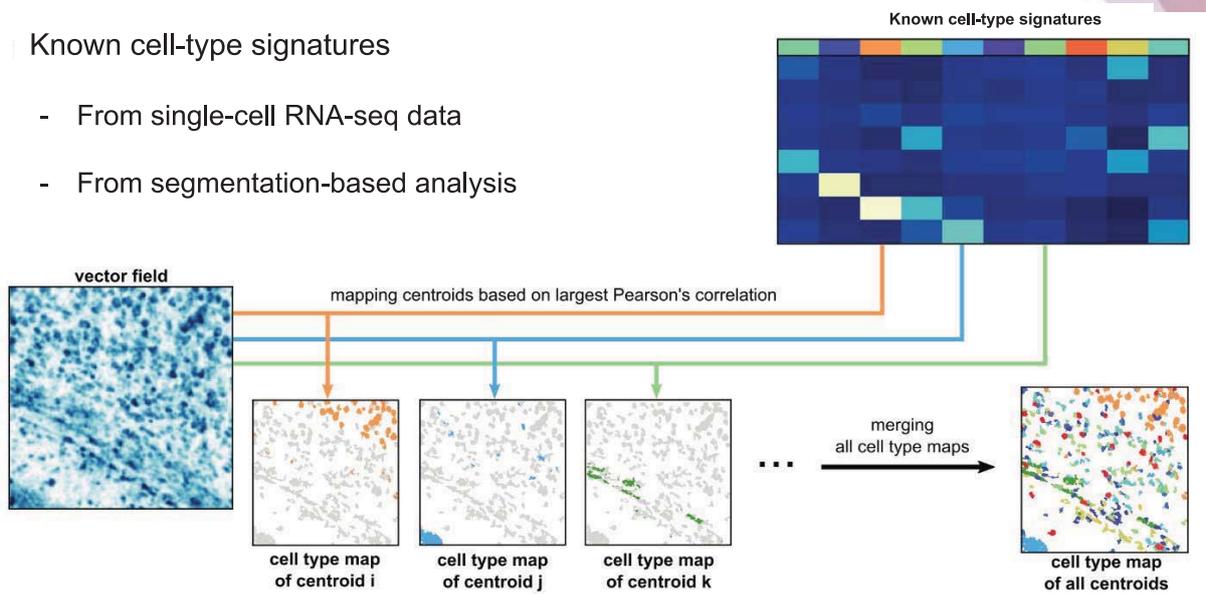


68

## 세포 타입 맵 (Cell-type map) 생성 – Guided Mode

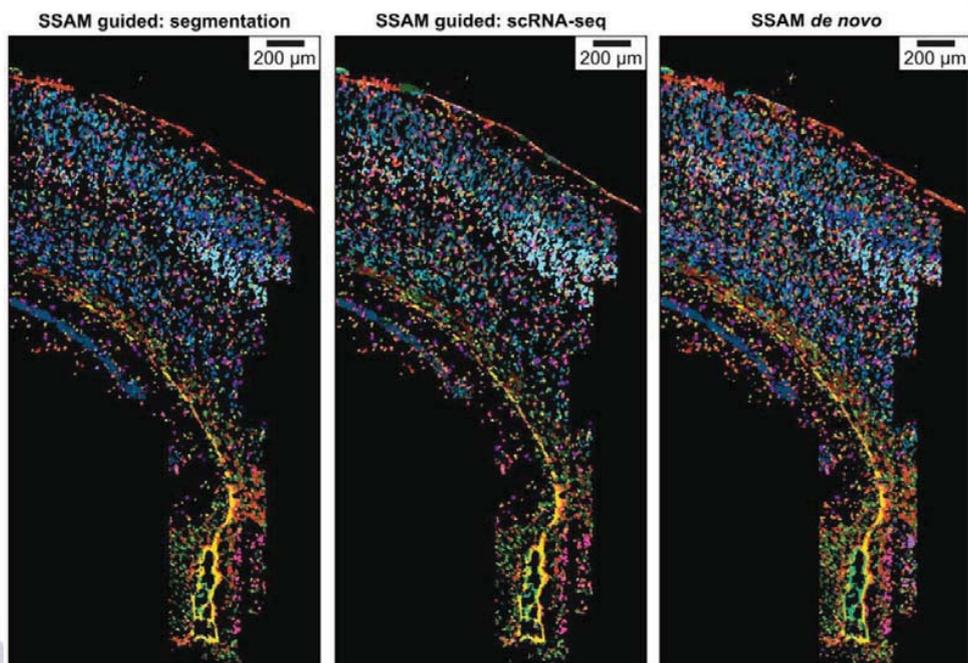
Known cell-type signatures

- From single-cell RNA-seq data
- From segmentation-based analysis



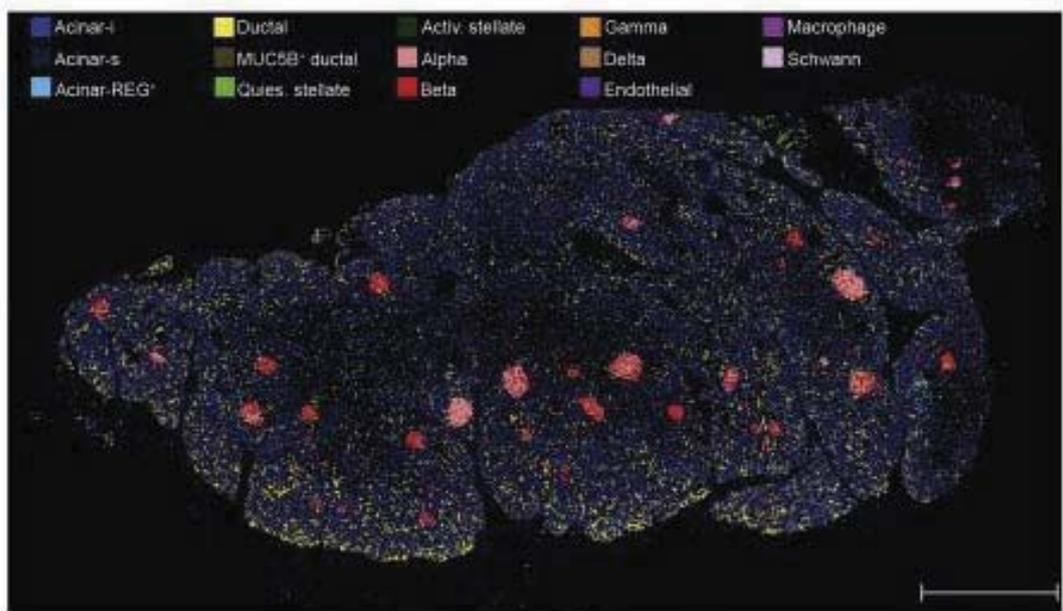
69

## SSAM Guided-mode 분석



70

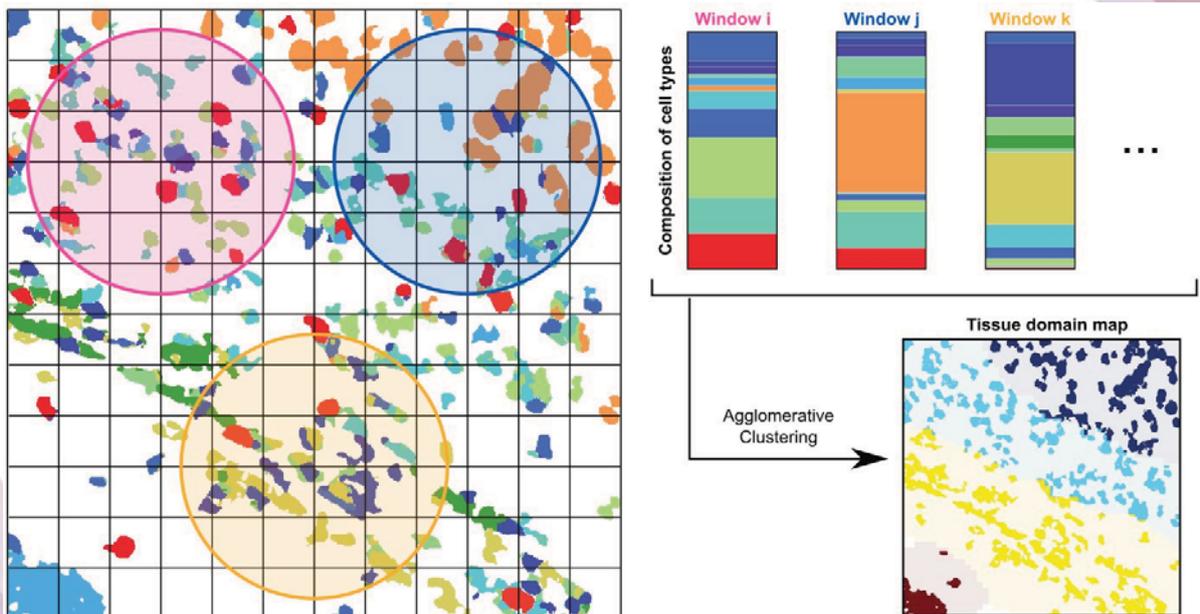
## 췌장 In situ sequencing 데이터 분석 (guided mode)



Gastroenterology 160(4) 1330-1344.E11, DOI: <https://doi.org/10.1053/j.gastro.2020.11.010>

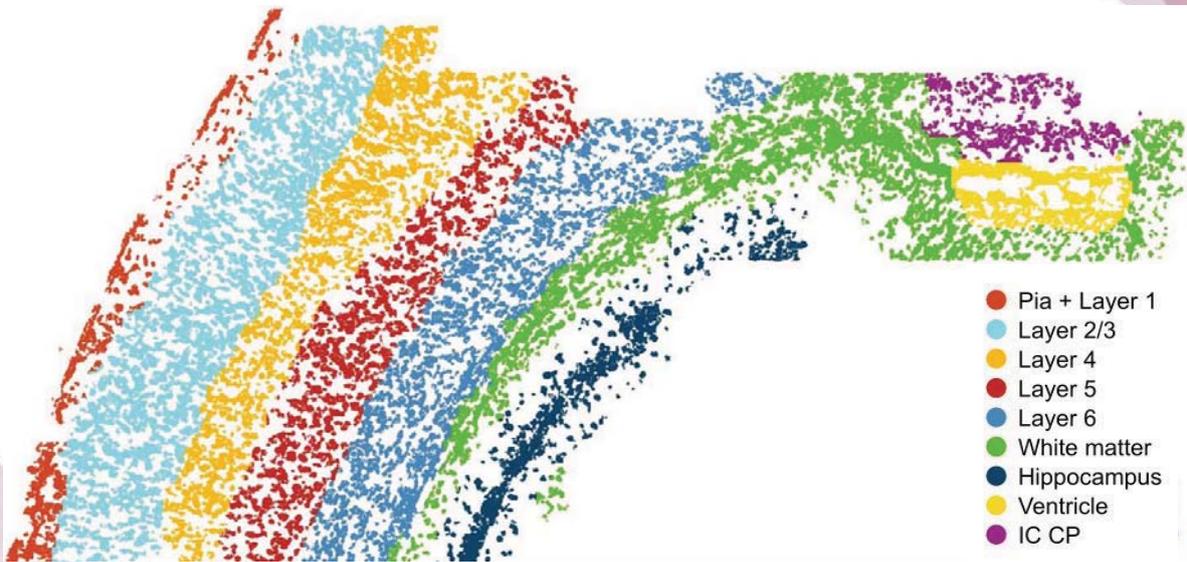
71

## 조직 구조 (domain) 분석



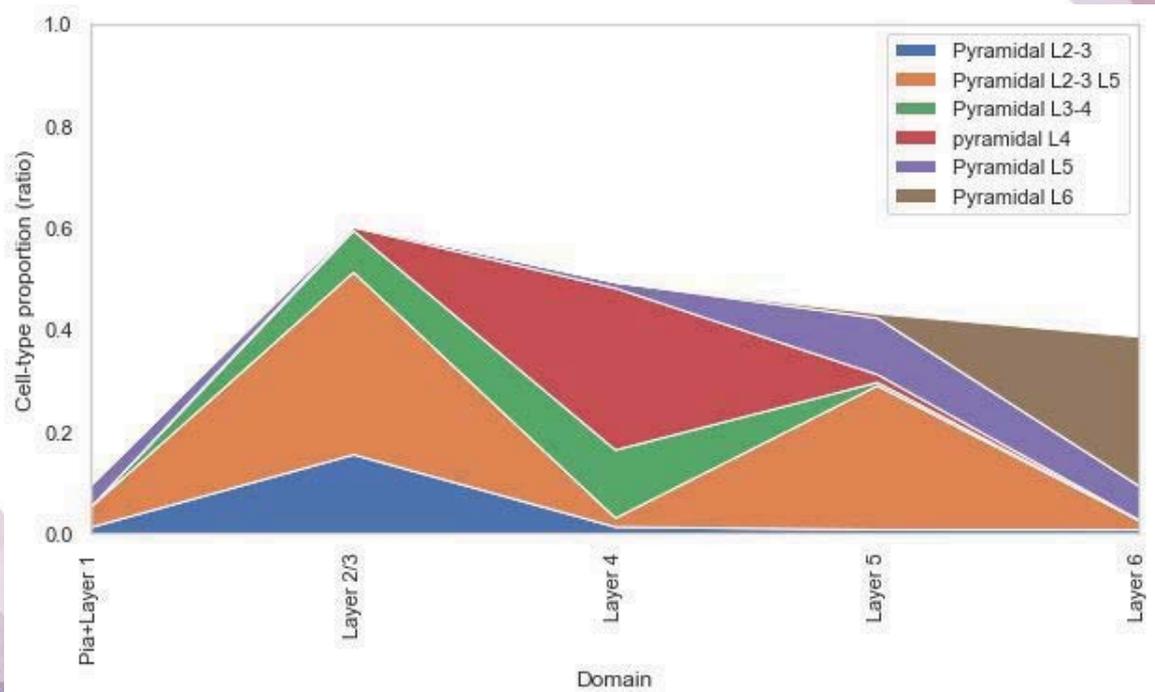
72

## 조직 구조 (domain) 분석



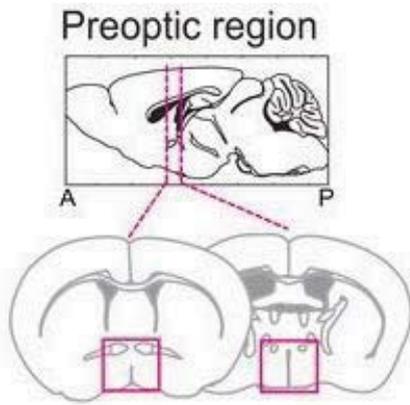
73

## 도메인 검증



74

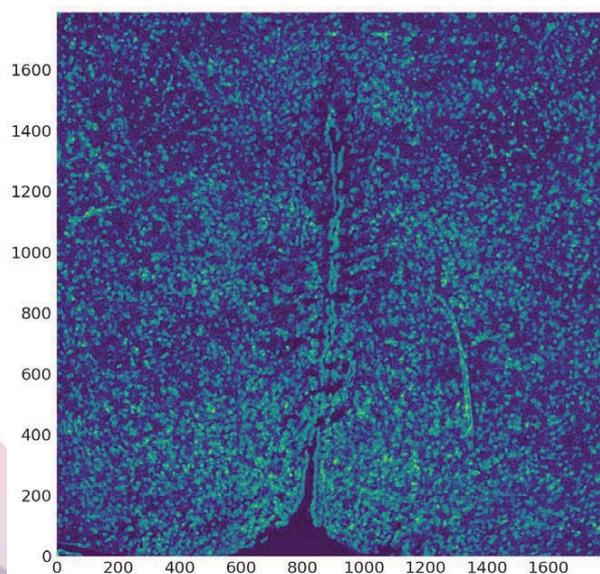
## 생쥐 시상하부 시각 교차 앞구역 (Preoptic area, MERFISH)



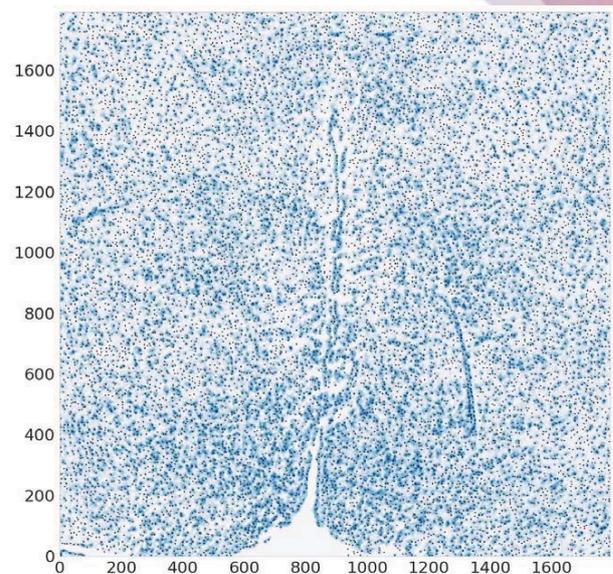
- 생쥐 시상하부 시각교차앞구역 (preoptic area, POA)
- 1.8mm x 1.8mm x 0.6mm
- MERFISH로 이미징한 mRNA 3D 위치
- 135 종 유전자

75

## 공간밀도추정 및 벡터 다운샘플링



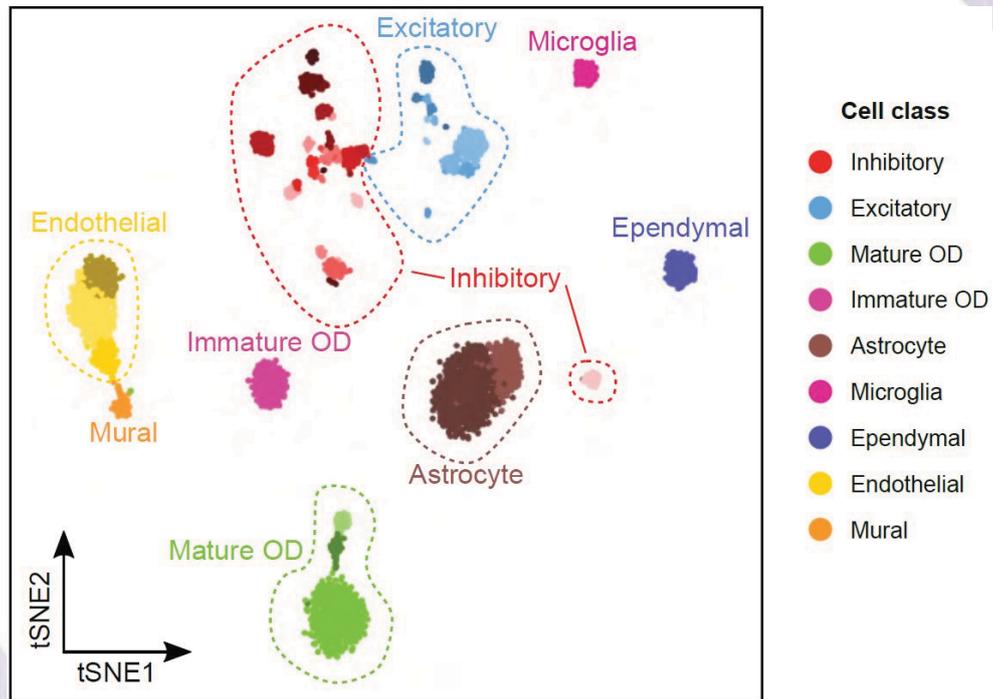
$1,790 \times 1,790 \times 9 = 28,836,900$  vectors  
( $z=4 \mu\text{m}$ )



L1 maxima

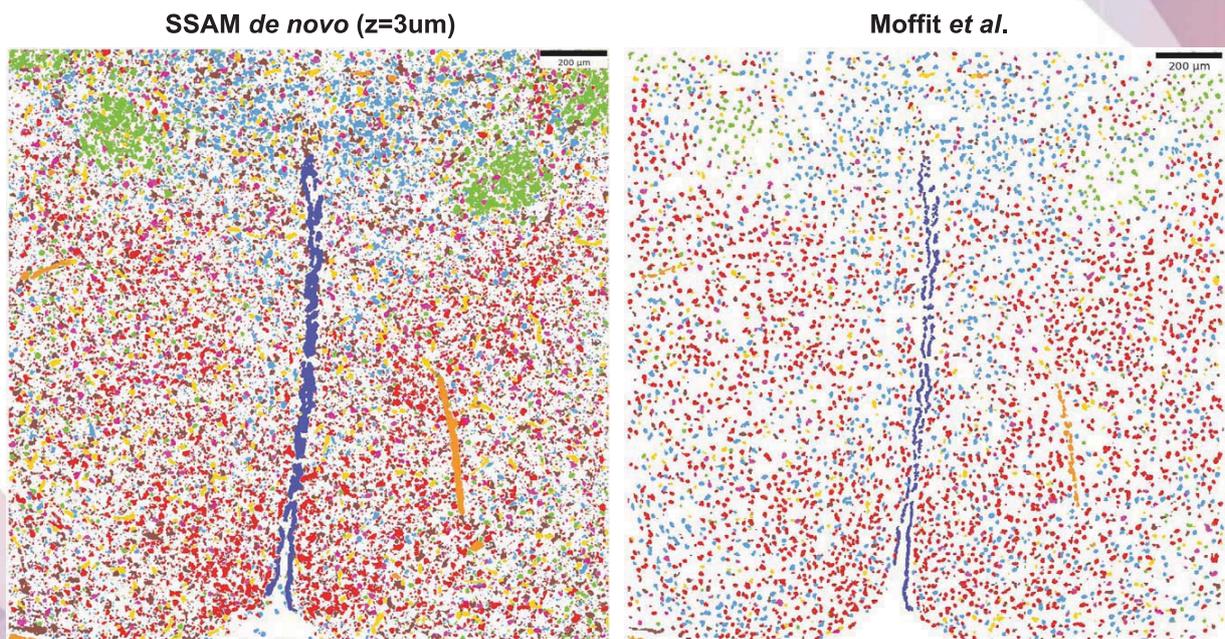
76

## 클러스터링



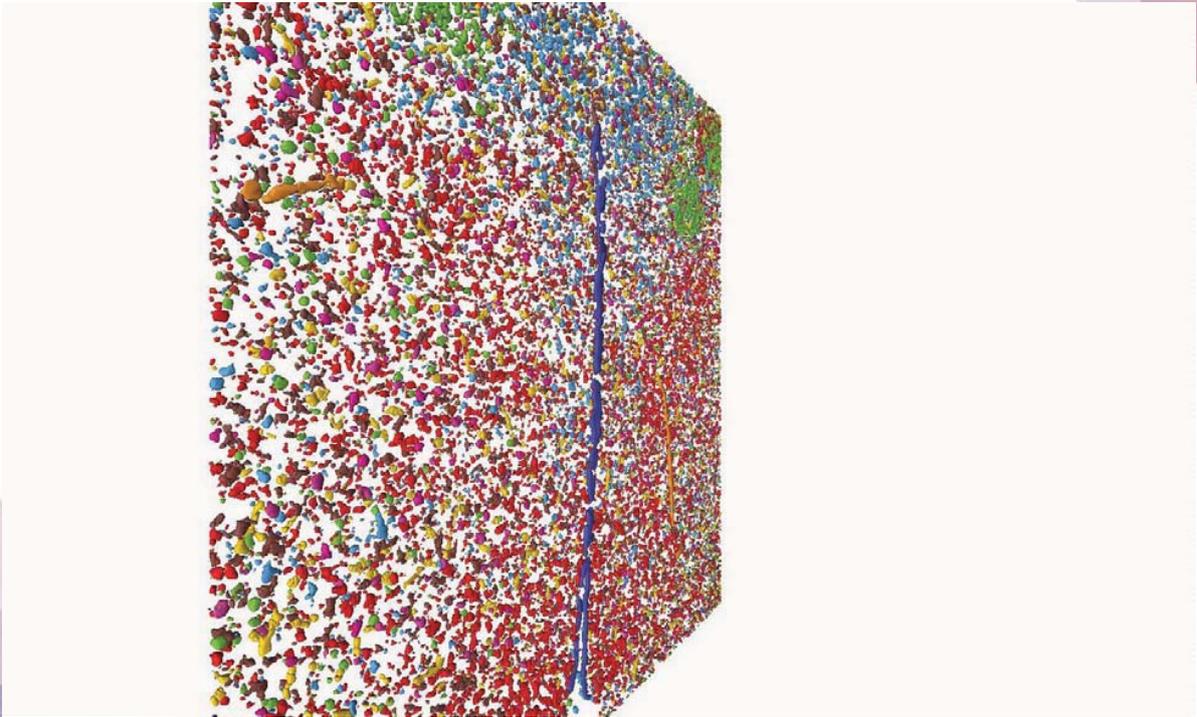
77

## 결과 비교



78

## 3D 세포 타입 맵



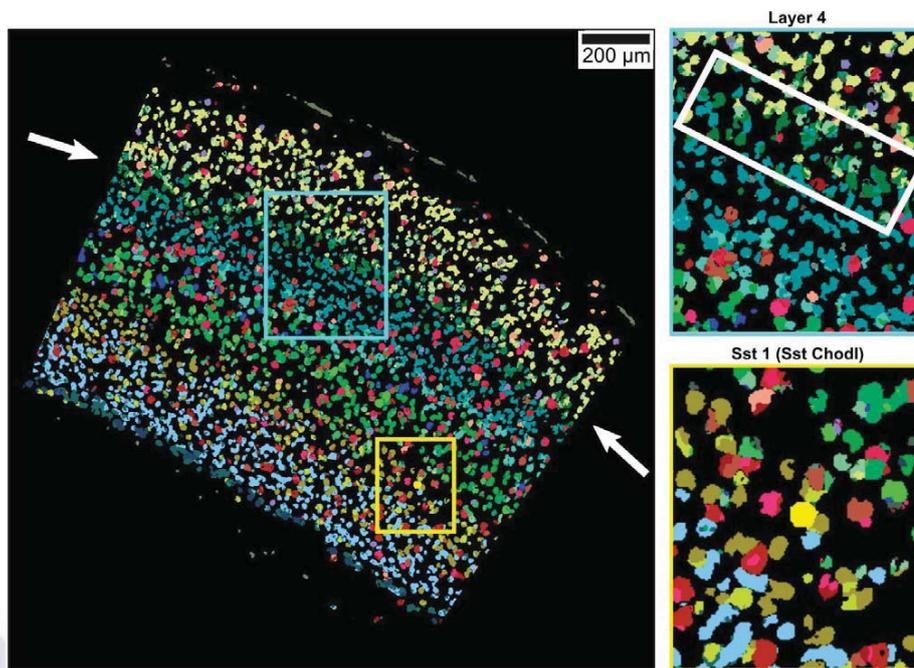
79

## Allen multiplexed smFISH 데이터

- SpaceTx 컨소시엄을 통해 생산된 데이터
- 생쥐 일차시각피질 (VISp)
- multiplexed smFISH (osmFISH와 비슷) 기법으로 얻은 2D mRNA 위치
- 22 종의 유전자

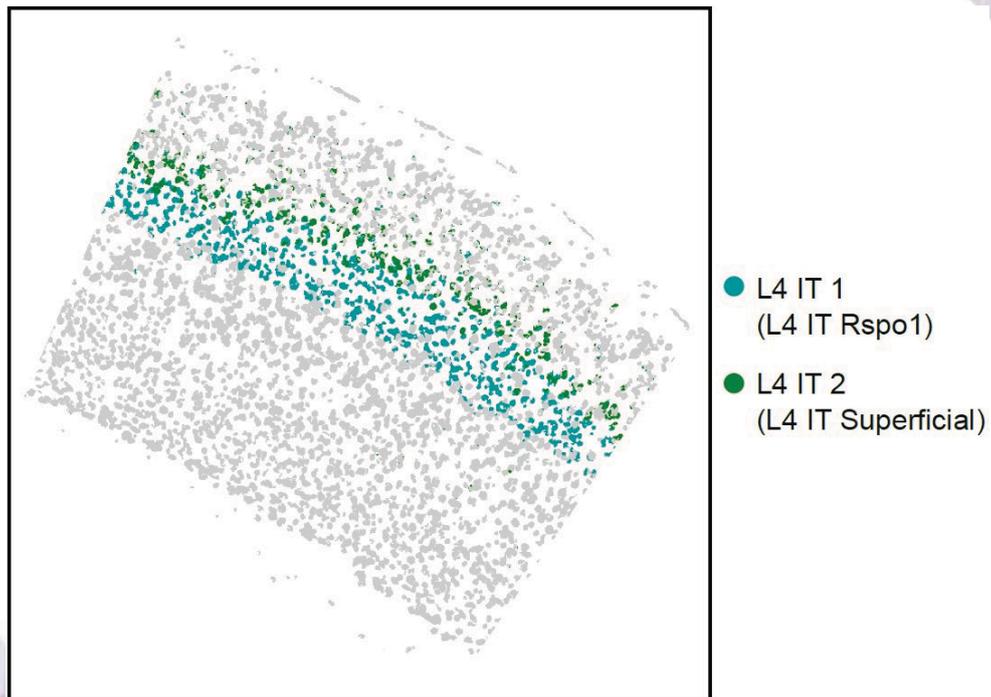
80

## 세포 타입 맵 (cell-type map)



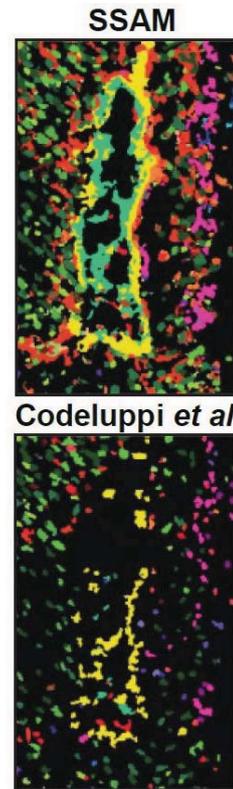
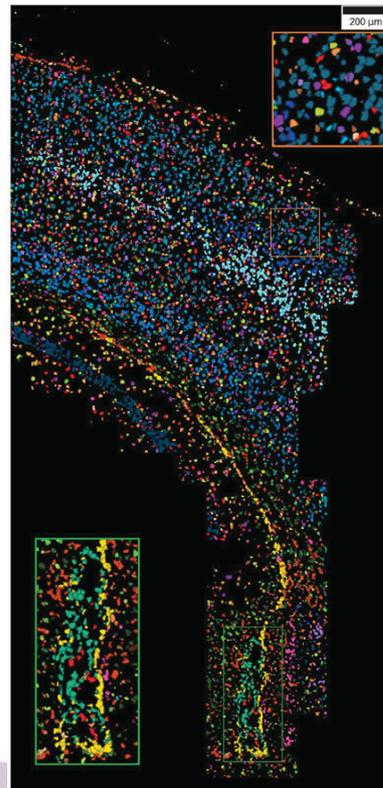
81

## L4 IT 세포 타입의 분리



82

## 세포 타입 맵을 활용한 세포 경계 검출 (segmentation)



83

## Acknowledgments

J. Park et al. *Nature Communications* 12, 3545 (2021)

• Wonyl Choi (Boston U.)

### BIH, Charite, Berlin

- Naveed Ishaque
- Roland Eils
- Luca Tosti
- Christian Conrad
- Bianca Hennig

### Linnarsson Lab, Karolinska

- Lars Borm
- Simone Codeluppi

### DKFZ, Heidelberg

- Oliver Stegle
- Daniel Hübschmann
- Matthias Schlesner
- Zuguang Gu
- Nagarajan Paramasivam
- Stephen Krämer
- Yue Zhuo

Harvard Medical School / Boston Children's Hospital

- Jeffery Moffit

Chan  
Zuckerberg  
Initiative

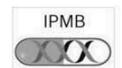


HUMAN  
CELL  
ATLAS



CHARITÉ

berlin



### Allen Brain Institute, USA

- Brian Long
- Emma Garren
- Thuc Nghi Nguyen
- Bosiljka Tasic
- Jeffery Miller
- Ambrose Carr
- Ed Lein

84

## 튜토리얼

- SSAM 설치:

```
pip install git+https://github.com/pnucolab/ssam
```

- Google Colab에서 SSAM 설치:

```
!pip install git+https://github.com/pnucolab/ssam
Collecting git+https://github.com/pnucolab/ssam
  Cloning https://github.com/pnucolab/ssam to /tmp/pip-req-build-p_qn9mq2
  Running command git clone --filter=blob:none --quiet https://github.com/pnucolab/ssam /tmp/pip-req-build-p_qn9mq2
  Resolved https://github.com/pnucolab/ssam to commit 7e037d8d2156a222caa914976b2ed810b3f2b876
  Preparing metadata (setup.py) ... done
Collecting numpy=1.24.4 (from ssam==1.1.1)
  Downloading numpy-1.24.4-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (17.3 MB)
  17.3/17.3 MB 18.7 MB/s eta 0:00:00
Collecting numba=0.57.1 (from ssam==1.1.1)
  Downloading numba-0.57.1-cp310-cp310-manylinux2014_x86_64.manylinux_2_17_x86_64.whl (3.6 MB)
  3.6/3.6 MB 38.3 MB/s eta 0:00:00
Collecting networkx=2.8.8 (from ssam==1.1.1)
  Downloading networkx-2.8.8-py3-none-any.whl (2.0 MB)
  2.0/2.0 MB 25.1 MB/s eta 0:00:00
Requirement already satisfied: scipy in /usr/local/lib/python3.10/dist-packages (from ssam==1.1.1) (1.11.4)
Requirement already satisfied: pandas in /usr/local/lib/python3.10/dist-packages (from ssam==1.1.1) (1.5.3)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.10/dist-packages (from ssam==1.1.1) (3.7.1)
```

85

## 튜토리얼

- 예제 데이터 내려받기

- [https://s3.amazonaws.com/starfish.data.spacetx/spacetx-website/data/smFISH\\_Allen/s3\\_spot\\_table.csv](https://s3.amazonaws.com/starfish.data.spacetx/spacetx-website/data/smFISH_Allen/s3_spot_table.csv)

```
!wget https://s3.amazonaws.com/starfish.data.spacetx/spacetx-website/data/smFISH_Allen/s3_spot_table.csv
--2024-01-31 06:27:14-- https://s3.amazonaws.com/starfish.data.spacetx/spacetx-website/data/smFISH_Allen/s3_spot_table.csv
Resolving s3.amazonaws.com (s3.amazonaws.com)... 54.231.199.0, 52.216.110.93, 52.216.144.125, ...
Connecting to s3.amazonaws.com (s3.amazonaws.com)|54.231.199.0|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 55124358 (53M) [text/csv]
Saving to: 's3_spot_table.csv'

s3_spot_table.csv 100%[=====] 52.57M 33.1MB/s in 1.6s

2024-01-31 06:27:16 (33.1 MB/s) - 's3_spot_table.csv' saved [55124358/55124358]
```

<https://spacetx.github.io/>

86

## 튜토리얼

- 예제 데이터 불러오기

```
[3] import pandas as pd

[5] spots = pd.read_csv("s3_spot_table.csv")
```

spots

Unnamed: 0	gene	molecule_id	confidence	cell	assignment_confidence	rotated_x	rotated_y	xc	yc
0	Fezf2	4	1.00000	1605	1.0	1296.718027	813.958480	1296.718027	813.958480
1	Fezf2	17	1.00000	2496	1.0	1302.424423	844.131209	1302.424423	844.131209
2	Fezf2	36	1.00000	1605	1.0	1293.752097	822.910508	1293.752097	822.910508
3	Fezf2	77	0.99999	1605	1.0	1292.981468	825.102131	1292.981468	825.102131
4	Fezf2	86	1.00000	3929	1.0	1254.284721	915.218969	1254.284721	915.218969
...	...	...	...	...	...	...	...	...	...
530778	Parm1	1071304	0.99995	407	1.0	115.997457	502.193148	115.997457	502.193148
530779	Parm1	1071313	1.00000	4354	1.0	134.008161	429.477827	134.008161	429.477827
530780	Parm1	1071319	0.99972	3478	1.0	107.807802	398.245527	107.807802	398.245527
530781	Parm1	1071321	0.99999	999	1.0	120.079417	451.938345	120.079417	451.938345
530782	Parm1	1072021	0.99984	2231	1.0	141.805571	65.985526	141.805571	65.985526

530783 rows × 10 columns

87

## 튜토리얼

- 예제 데이터를 형식에 맞게 변환하기

```
spots_xy = spots.rename(columns={'rotated_x': 'x', 'rotated_y': 'y'})[['gene', 'x', 'y']]
spots_xy
```

	gene	x	y
0	Fezf2	1296.718027	813.958480
1	Fezf2	1302.424423	844.131209
2	Fezf2	1293.752097	822.910508
3	Fezf2	1292.981468	825.102131
4	Fezf2	1254.284721	915.218969
...	...	...	...
530778	Parm1	115.997457	502.193148
530779	Parm1	134.008161	429.477827
530780	Parm1	107.807802	398.245527
530781	Parm1	120.079417	451.938345
530782	Parm1	141.805571	65.985526

530783 rows × 3 columns

88

## 튜토리얼

- SSAMDataset, SSAMAnalysis 객체 생성하기

```
[1] import ssam
[2] ds = ssam.SSAMDataset("exampledata")
    analysis = ssam.SSAMAnalysis(ds, verbose=True)
```

- KDE 실행하기

```
analysis.run_kde(spots_xy, width=1330, height=1220)
... Running KDE for gene Alcam...
    Saving KDE for gene Alcam...
    Running KDE for gene Chodl...
    Saving KDE for gene Chodl...
    Running KDE for gene Cux2...
    Saving KDE for gene Cux2...
    Running KDE for gene Fezf2...
    Saving KDE for gene Fezf2...
```

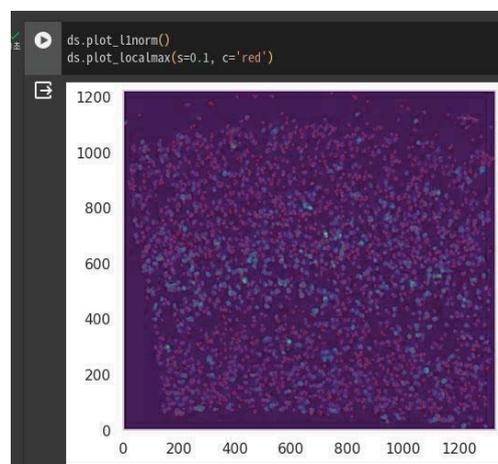
89

## 튜토리얼

- Local Maxima 찾기

```
[18] analysis.find_localmax()
Found 2787 local max vectors.
```

- Local Maxima 확인하기



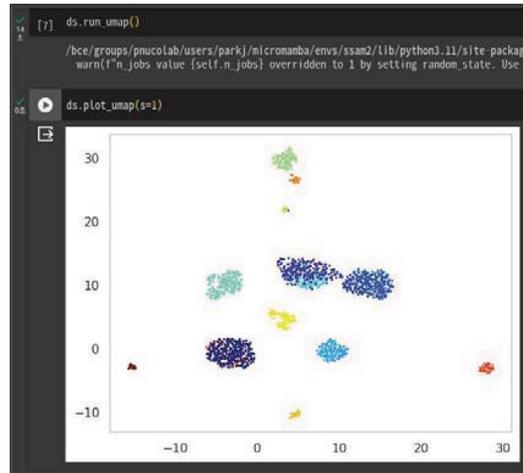
90

## 튜토리얼

- 클러스터링하기

```
[6] analysis.cluster_vectors()  
Found 15 clusters
```

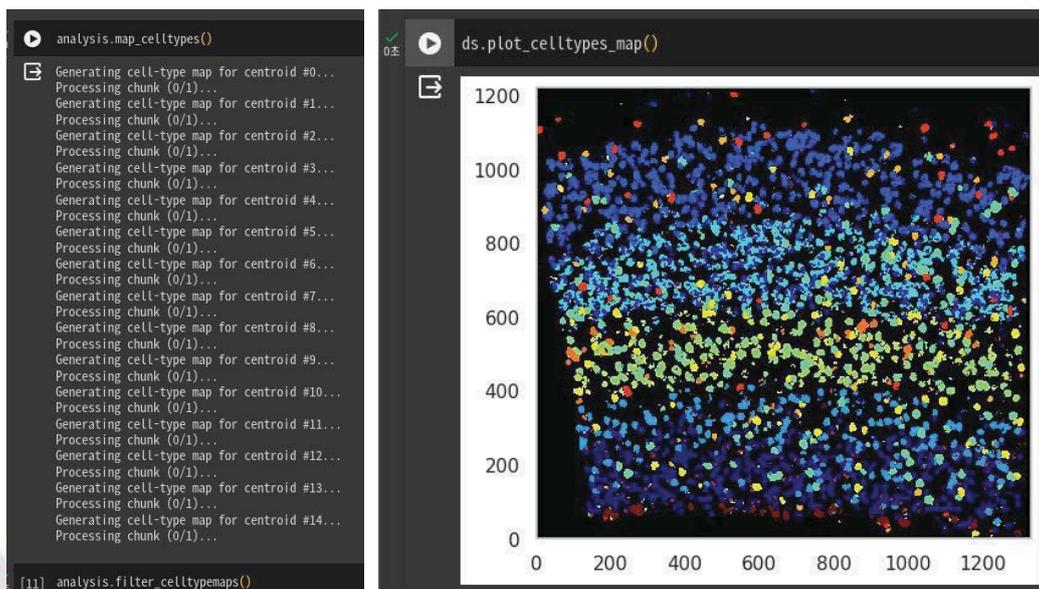
- UMAP 그리고 확인하기



91

## 튜토리얼

- Cell-type map 그리기



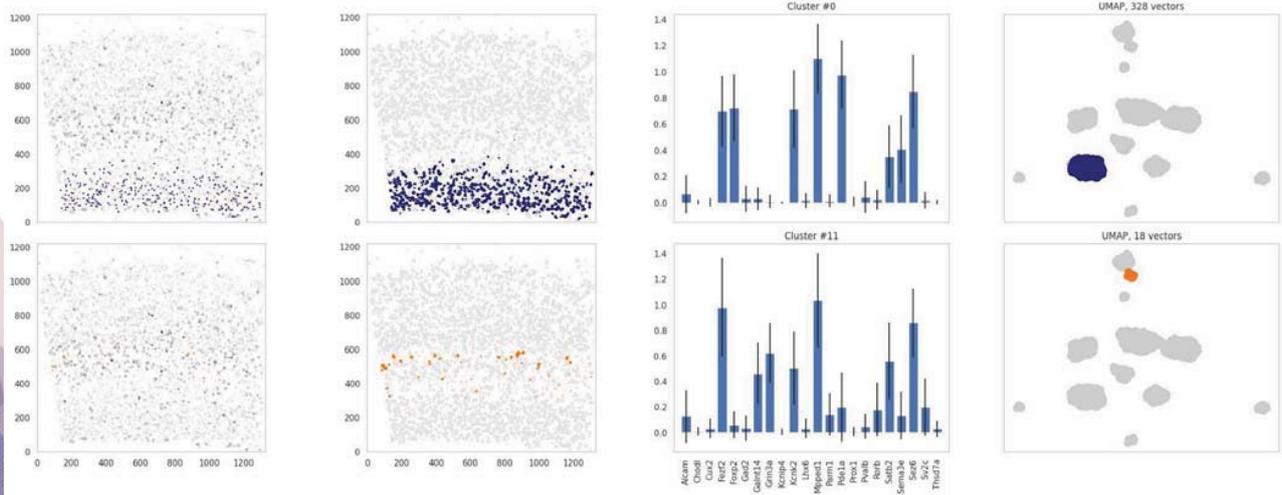
92

## 튜토리얼

- Cell type 별 gene expression 구하기 (diagnostic plots)

```
import matplotlib.pyplot as plt

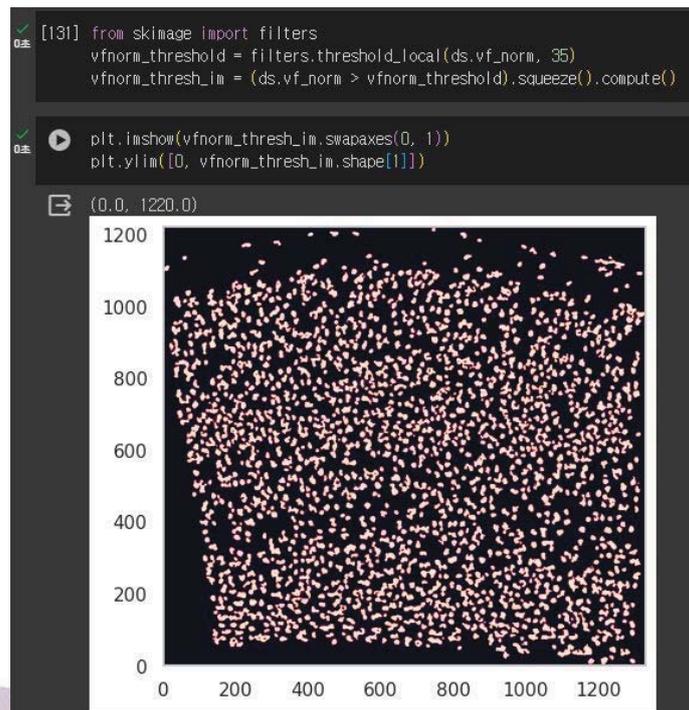
for i in range(15): # cell type 개수
    plt.figure(figsize=[30,5]) # 가로로 길게
    ds.plot_diagnostic_plot(i, use_embedding='umap')
```



93

## 튜토리얼

- Cell Segmentation 구하기 - Marker 찾기



94

## 튜토리얼

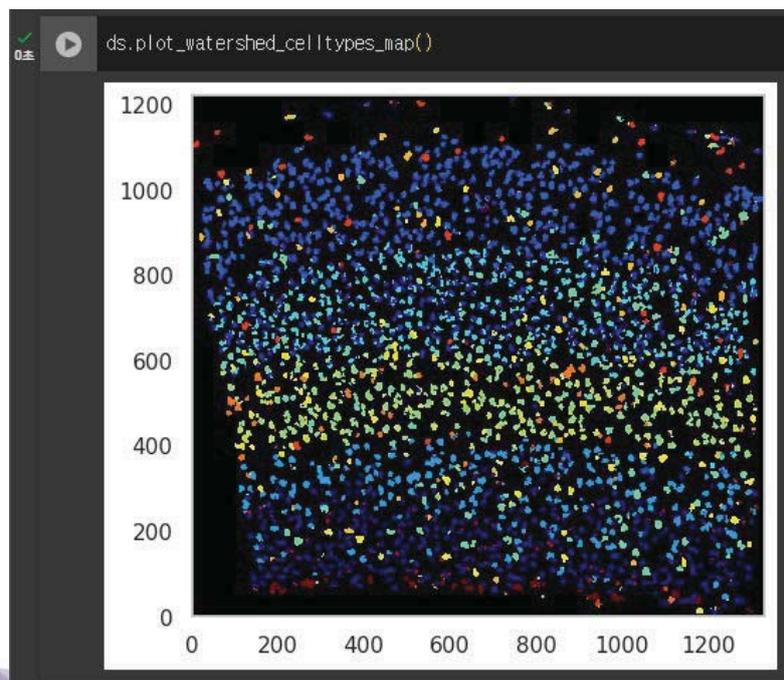
- Cell Segmentation 수행하기

```
✓ [133] analysis.run_watershed(vfnorm_thresh_im)
1초
Segmenting cell type #0...
Segmenting cell type #1...
Segmenting cell type #2...
Segmenting cell type #3...
Segmenting cell type #4...
Segmenting cell type #5...
Segmenting cell type #6...
Segmenting cell type #7...
Segmenting cell type #8...
Segmenting cell type #9...
Segmenting cell type #10...
Segmenting cell type #11...
Segmenting cell type #12...
Segmenting cell type #13...
Segmenting cell type #14...
```

95

## 튜토리얼

- Cell Segmentation 기반 Cell-type Map 그리기



96

## 튜토리얼

- Guided mode를 위한 데이터 내려받기
  - [https://s3.amazonaws.com/starfish.data.spacetx/spacetx-website/data/smFISH\\_Allen/s3\\_cell\\_by\\_gene.csv](https://s3.amazonaws.com/starfish.data.spacetx/spacetx-website/data/smFISH_Allen/s3_cell_by_gene.csv)

```
!wget https://s3.amazonaws.com/starfish.data.spacetx/spacetx-website/data/smFISH_Allen/s3_cell_by_gene.csv
--2024-01-31 15:45:54-- https://s3.amazonaws.com/starfish.data.spacetx/spacetx-website/data/smFISH_Allen/s3_cell_by_gene.csv
s3.amazonaws.com (s3.amazonaws.com) 해석 중... 52.217.65.206, 54.231.130.248, 54.231.164.112, ...
다음으로 연결 중: s3.amazonaws.com (s3.amazonaws.com)|52.217.65.206|:443... 연결했습니다.
HTTP 요청을 보냈습니다. 응답 기다리는 중... 200 OK
길이: 129436 (126K) [text/csv]
저장 위치: `s3_cell_by_gene.csv`

s3_cell_by_gene.csv 100%[=====] 126,40K 220KB/s / 0.6s
2024-01-31 15:45:56 (220 KB/s) - `s3_cell_by_gene.csv' 저장함 [129436/129436]
```

<https://spacetx.github.io/>

97

## 튜토리얼

- Guided mode를 위한 데이터 불러오기

```
import pandas as pd
cxg = pd.read_csv("s3_cell_by_gene.csv", header=1, index_col='gene_name')[ds.genes]
cxg
```

	Alcam	Chodl	Cux2	Fezf2	Foxp2	Gad2	Galnt14	Grin3a	Kcnp4
gene_name									
1	0	0	0	15	20	0	0	1	0
5	4	1	20	3	1	3	2	0	1
6	36	1	39	1	1	0	1	0	2
7	32	1	37	2	1	0	0	1	1
9	66	0	19	2	0	75	0	0	0
...	...	...	...	...	...	...	...	...	...
4664	19	1	42	2	2	1	0	0	1
4665	0	0	0	15	21	1	0	0	0
4669	57	10	6	9	5	65	23	50	4
4670	7	0	6	0	0	0	0	0	0
4690	0	0	0	2	6	0	0	0	0

2360 rows x 22 columns

98

## 튜토리얼

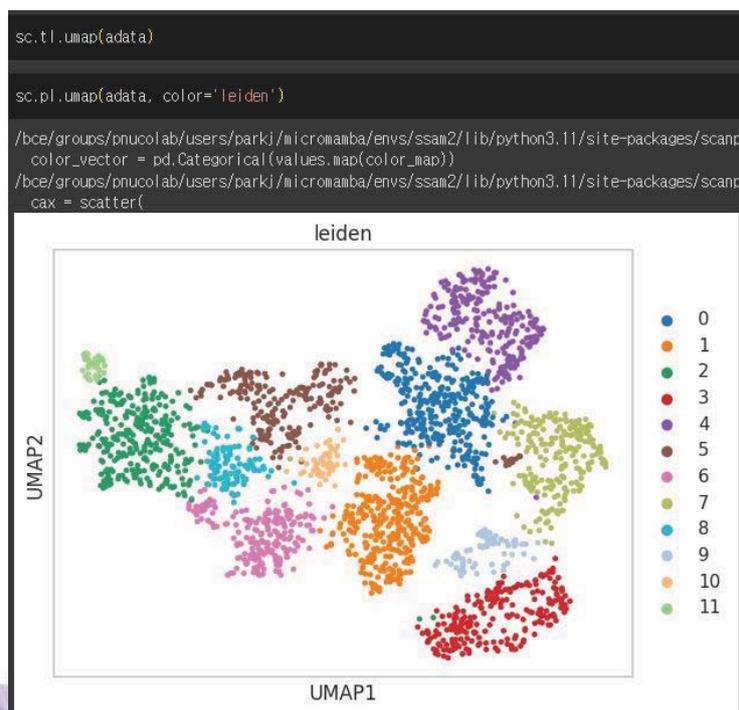
- ScanPy로 데이터 분석하기 (Single-cell 분석 튜토리얼 참고)

```
[70] import scanpy as sc # This requires `pip install scanpy`
[71] adata = sc.AnnData(cxg)
/bce/groups/pnucolab/users/parkj/micromamba/envs/ssam2/lib/python3.11/
warnings.warn("Transforming to str index.", ImplicitModificationWarn
[72] # A standard scRNA-seq processing pipeline
# e.g., https://scanpy-tutorials.readthedocs.io/en/latest/pbmc3k.html
sc.pp.normalize_total(adata, target_sum=1e4)
sc.pp.log1p(adata)
sc.tl.pca(adata, svd_solver='arpack')
sc.pp.neighbors(adata, n_neighbors=10)
sc.tl.leiden(adata, resolution=0.6)
```

99

## 튜토리얼

- ScanPy로 데이터 분석하기 (Single-cell 분석 튜토리얼 참고)



100

## 튜토리얼

- Centroid 기반 Cell-type map 생성하기

```
[75] centroids = []
      for cl in adata.obs.leiden.unique():
          centroids.append(adata.X[adata.obs.leiden == cl].mean(axis=0))

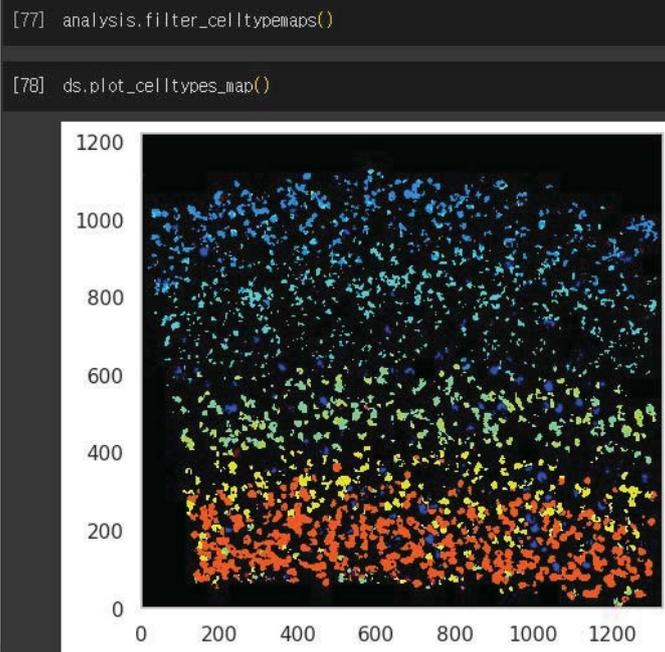
[76] analysis.map_celltypes(centroids)

Generating cell-type map for centroid #0...
Processing chunk (0/1)...
Generating cell-type map for centroid #1...
Processing chunk (0/1)
```

101

## 튜토리얼

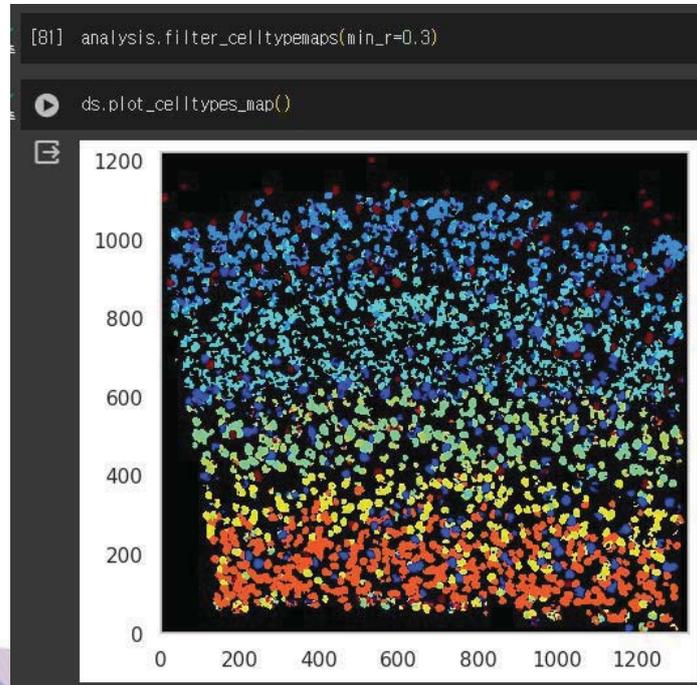
- Centroid 기반 Cell-type map 생성하기



102

## 튜토리얼

- Centroid 기반 Cell-type map 생성하기



103