KSBi-BIML 2026

Bioinformatics & Machine Learning(BIML) Workshop for Life Scientists

생명정보학 & 머신러닝 워크샵 (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 2026 워크샵을 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 행위자 본인에게 있음**을 알립니다.

KSBi-BIML 2026

Bioinformatics & Machine Learning (BIML) Workshop for Life Scientists

한국생명정보학회가 주최하는 BIML-2026 동계 Bioinformatics &Machine Learning 교육 워크숍에 여러분을 초대합니다.

BIML 워크숍은 생명정보학 연구자들이 최신 AI바이오 분야의 인공지능 기반 분석 기술과 바이오 데이터 분석 기법을 이론과 실습을 통해 체계적으로 배울 수 있는 전문 교육 프로그램입니다. 2015년에 시작된 BIML 워크숍은 올해로 12년 차를 맞이하며, 국내 생명정보학 분야의 최초이자 최고 수준의 교육 프로그램으로 자리 잡았습니다. 이번 워크숍은 크게 인공지능바이오(AI바이오) 분야와 디지털바이오 분야, 두 분야로 구성됩니다.

AI바이오 분야에서는 생명정보 분석에 폭넓게 응용되고 있는 다양한 인공지능 기반 자료 모델링 기법을 다룰 예정입니다. 특히, 인공지능 심층학습을 활용한 단백질 구조 예측, 유전체 분석, 신약 개발에 대한 이론 및 실습 강의가 진행됩니다.

또한 디지털바이오 분야에서는 단일세포오믹스, 공간오믹스, 멀티오믹스, 메타오믹스에 대한 강의도 마련되어 있어, 연구자들의 분석 역량 강화에 실질적인 도움을 줄 것으로 기대됩니다.

또한 2024년부터 추가된 의료정보 자료 분석을 다루는 강의를 올해도 지속해서 운영하고자 합니다. 이는 최근 의료정보 자료 분석에 관한 연구 수요 증가를 반영한 것으로, 관련 연구를 수행하는 의과학자 및 의료정보 연구자들에게 유용한 지침을 제공할 것입니다.

또한, 올해도 생명정보학 기술의 다양화에 발맞춰 온라인 강좌를 대폭 확대했습니다. 올해는 무료 강좌 10개를 포함한 총 40개 이상의 강좌가 개설되며, 연구 주제에 맞는 강좌 추천과 강연료 할인 혜택도 제공합니다.

BIML-2026는 국내 주요 연구 중심 대학의 전임 교수 및 각 분야 최고 전문가들의 강의로 구성되어 있으며, 기초 이론부터 최신 연구 동향까지 아우르는 심도 있는 교육의 장이 될 것으로 확신합니다.

여러분의 많은 관심과 참여를 기대합니다!

2026년 2월

한국생명정보학회장 류 성 호

강의개요

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

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 | 한국생명정보학회

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

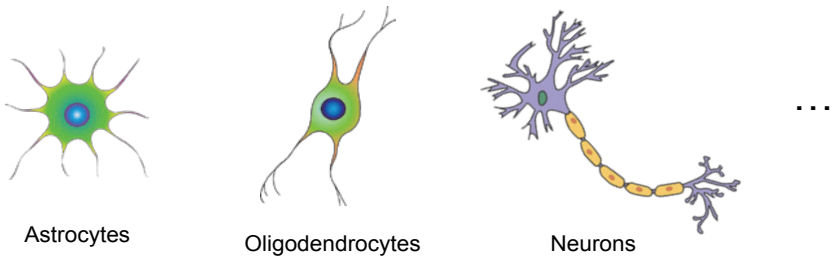
2

Chapter 1.
Spatial Transcriptomics Overview & Methods

3

Cell types

- What is cell type?



Astrocytes Oligodendrocytes Neurons …

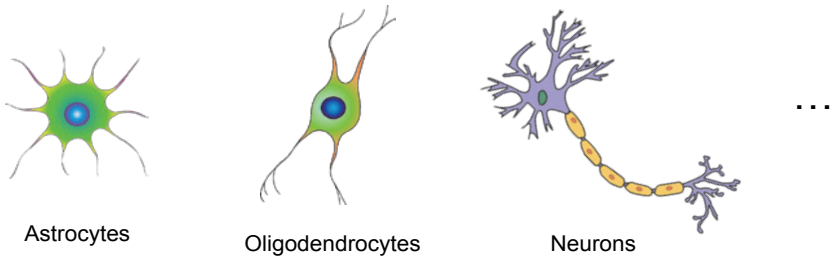
Cell type images from Wikimedia common

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

4

Cell types

- What is cell type?



Astrocytes Oligodendrocytes Neurons ...

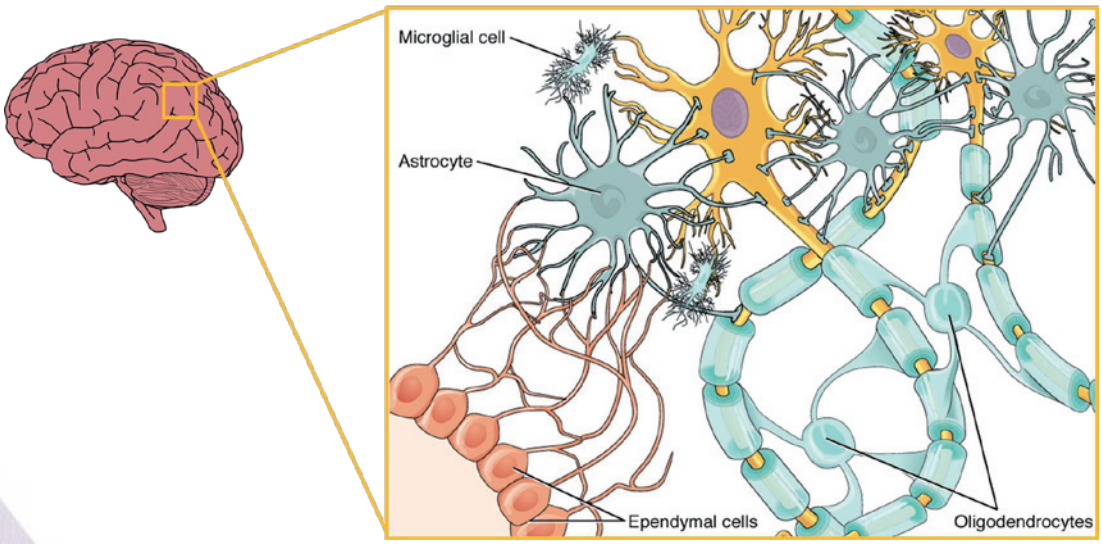
Cell type images from Wikimedia common

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

Really?

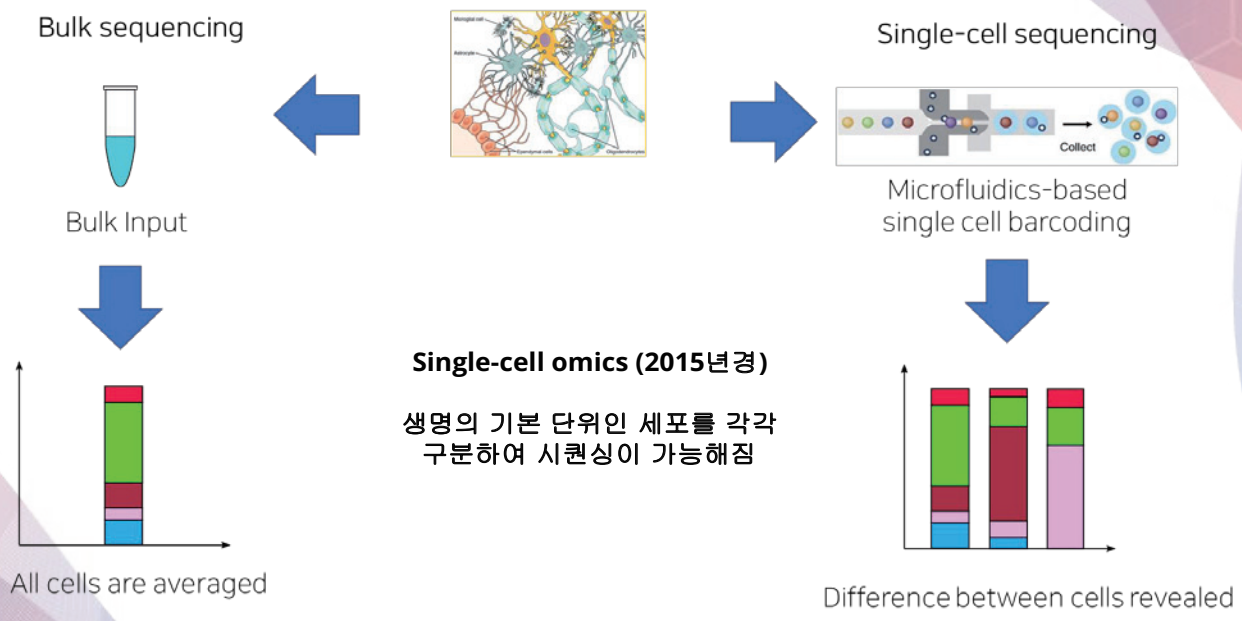
5

Overview



6

Overview

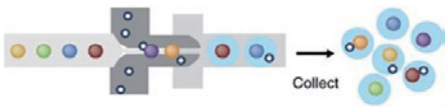


Single-cell omics (2015년경)

생명의 기본 단위인 세포를 각각 구분하여 시퀀싱이 가능해짐

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!

Single-cell omics

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

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.

Danila 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/

11

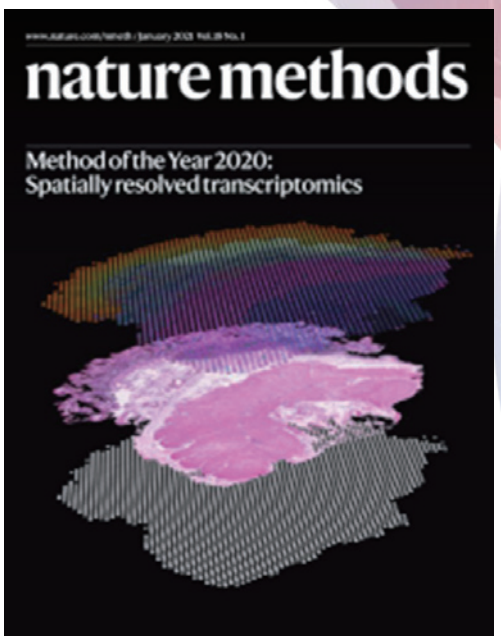
Spatial Omics

- 단일세포체학 + 공간차원
- 개별 세포들 간의 차이를 확인

+

각 세포의 위치까지 확인가능

- 현재 가장 각광받고 있는 오믹스 기술!



12

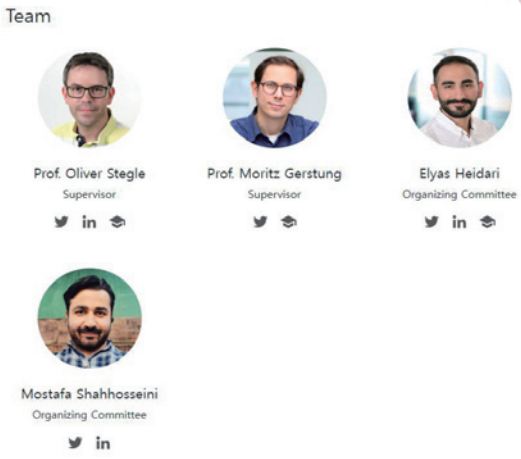
Spatial Omics

- 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/



14

Spatial Omics

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

Hackathon Schedule

Day 1

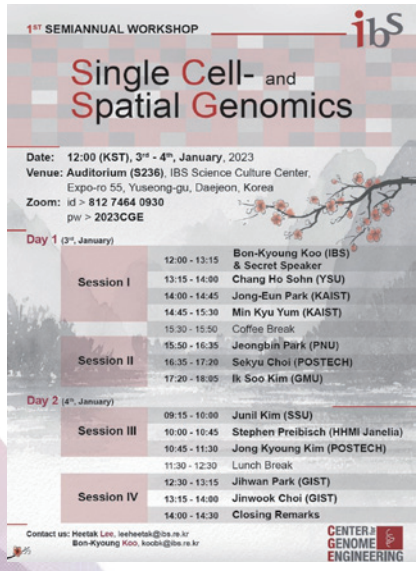
Wednesday, July 19th, 2023

Brainstorming

13:30 - 13:45: Reception
13:45 - 14:00: Openning Talk: **Elyas Heidari**
14:00 - 14:45: Openning Talk: **Dr. Oliver Stegle** and **Dr. Moritz Gerstung**
15:00 - 15:30: Luca Marconato, Status Quo of SpatialData framework for standadization of spatial data types and analysis
15:30 - 16:00: Jeongbin Park, Status Quo of SSAM for segmentation-free spatial annotation
16:00 - 16:15: Break
16:15 - 17:00: Louis Kuemmerle, Status Quo of TXsim, best practices of QC and panel design
17:00 - 17:30: Florian Wuennemann, Status Quo of MCMICRO for preprocessing, quality control, and workflow automation
17:30 - 18:00: Vitalii Kleshchevnikov, Cell-cell communication
18:00 - 18:30: Artem Shmatko, 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 >> **Session Leaders**

15

Single Cell & Spatial Omics Korea (SCSOK)



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

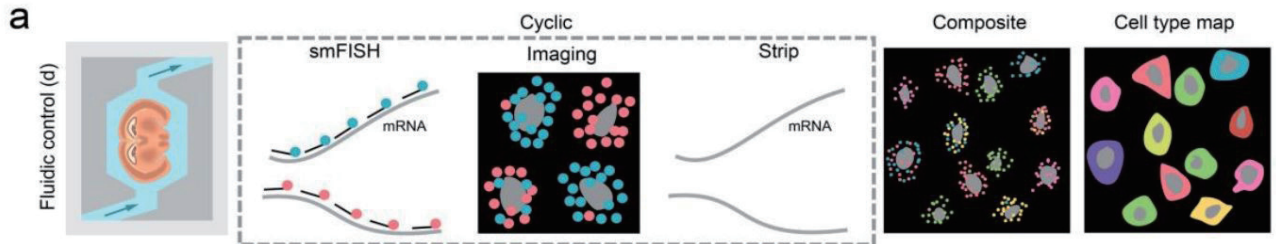
Brief Communication | Published: 30 October 2018

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 ✉

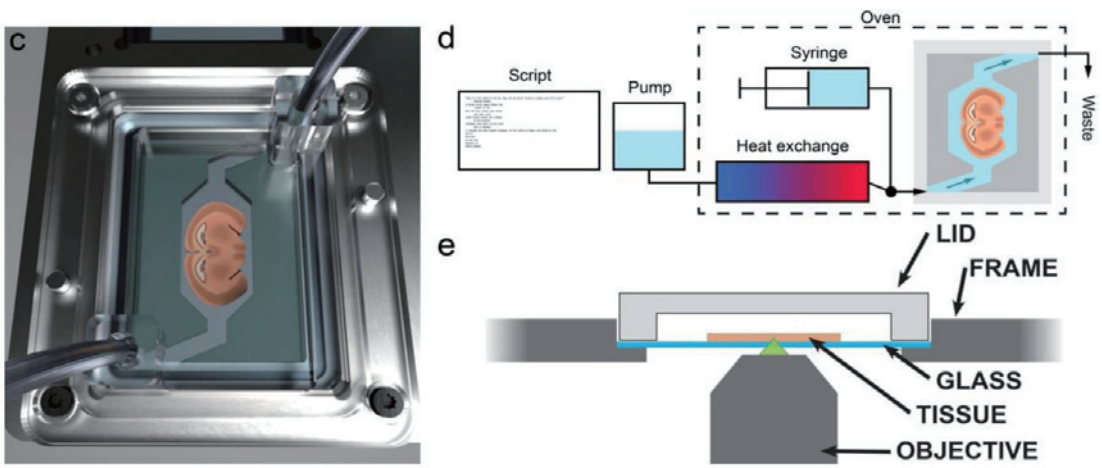
Nature Methods **15**, 932–935 (2018) | Cite this article

osmFISH



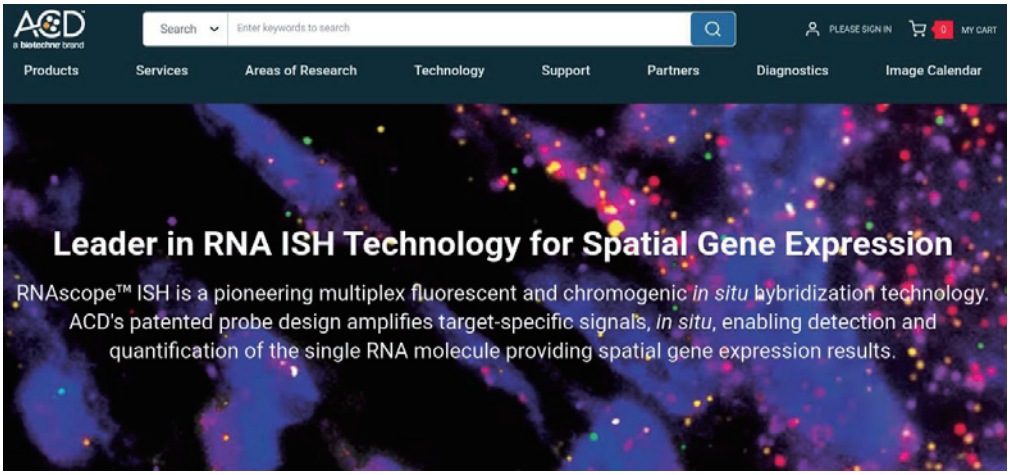
21

osmFISH

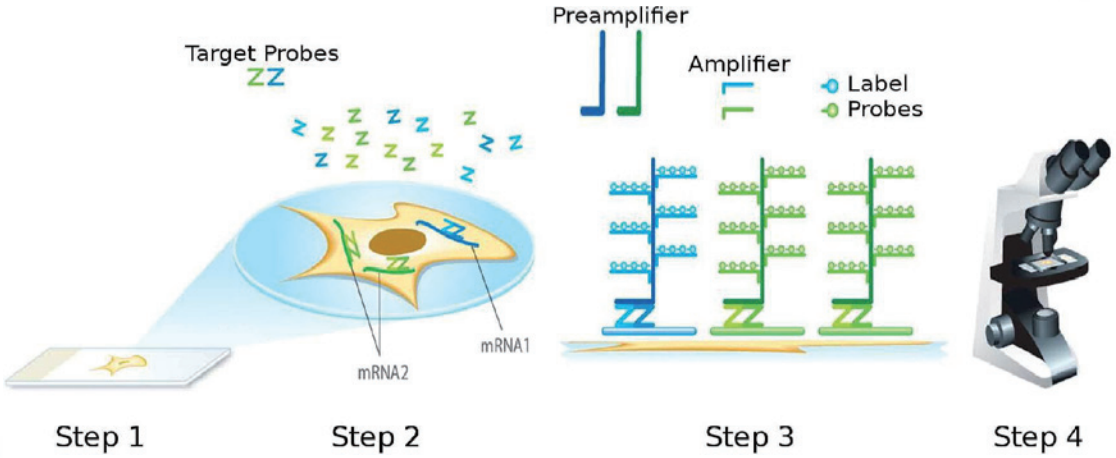


22

RNAScope



RNAScope



Combinatorial FISH-based Methods

MERFISH (MERSCOPE) and seqFISH+

RESEARCH ARTICLE

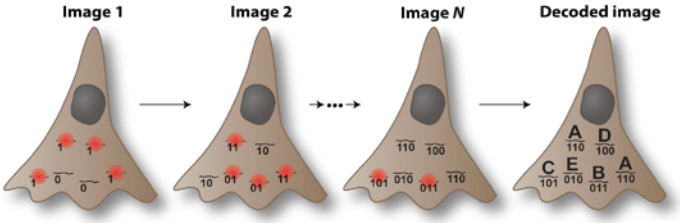
Spatially resolved, highly multiplexed RNA profiling in single cells

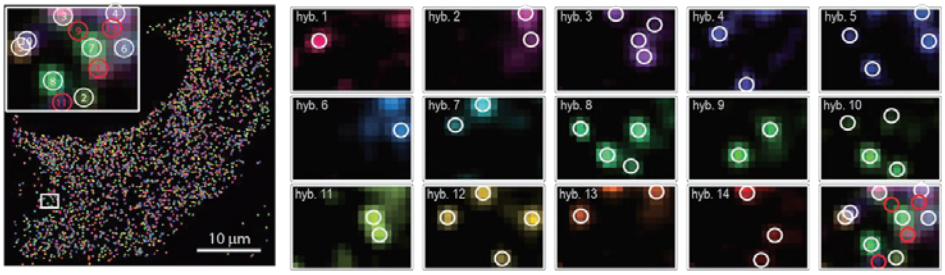
KOK HAO CHEN, ALISTAIR N. BOETTIGER, JEFFREY R. MOFFITT, SIYUAN WANG, AND XIAOWEI ZHUANG Authors Info & Affiliations

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

MERFISH (MERSCOPE)

RNA	Barcode
A	11…0
B	01…1
C	10…1
D	10…0
E	01…0
F	00…1
G	11…1





27

SeqFISH+

Letter | Published: 25 March 2019

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

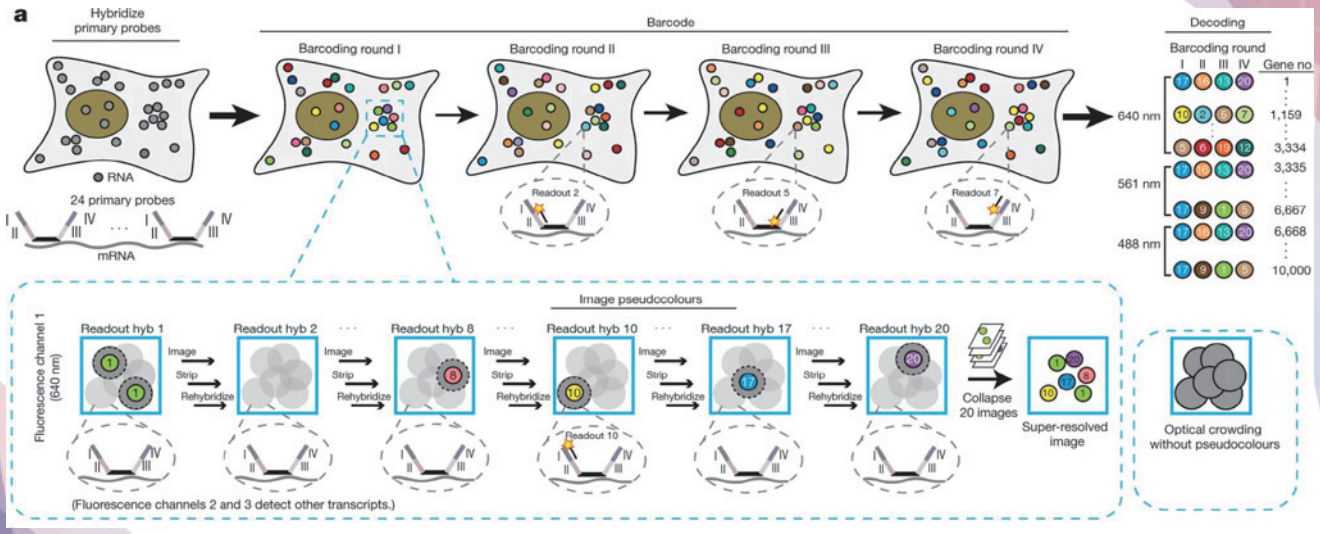
Chee-Huat Linus Eng, Michael Lawson, Qian Zhu, Ruben Dries, Noushin Koulena, Yodai Takei, Jina Yun, Christopher Cronin, Christoph Karp, Guo-Cheng Yuan & Long Cai

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

93k Accesses | **645** Citations | **293** Altmetric | Metrics

28

SeqFISH+



29

Spatial Transcriptomics Methods

ISS-based Methods

In Situ Sequencing (ISS), FISSEQ, Xenium, etc.

30

In Situ Sequencing (ISS)

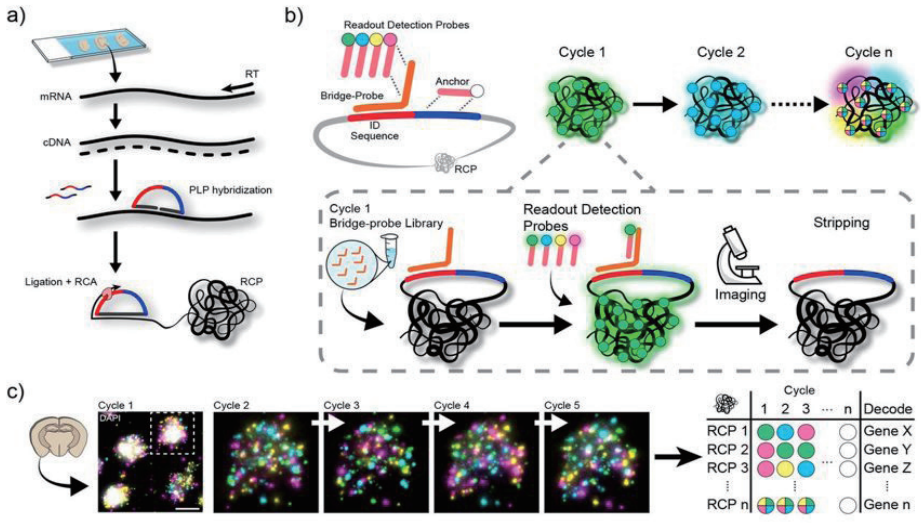
Published: 14 July 2013

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

In Situ Sequencing (ISS)



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

Fluorescent in situ sequencing (FISSEQ)

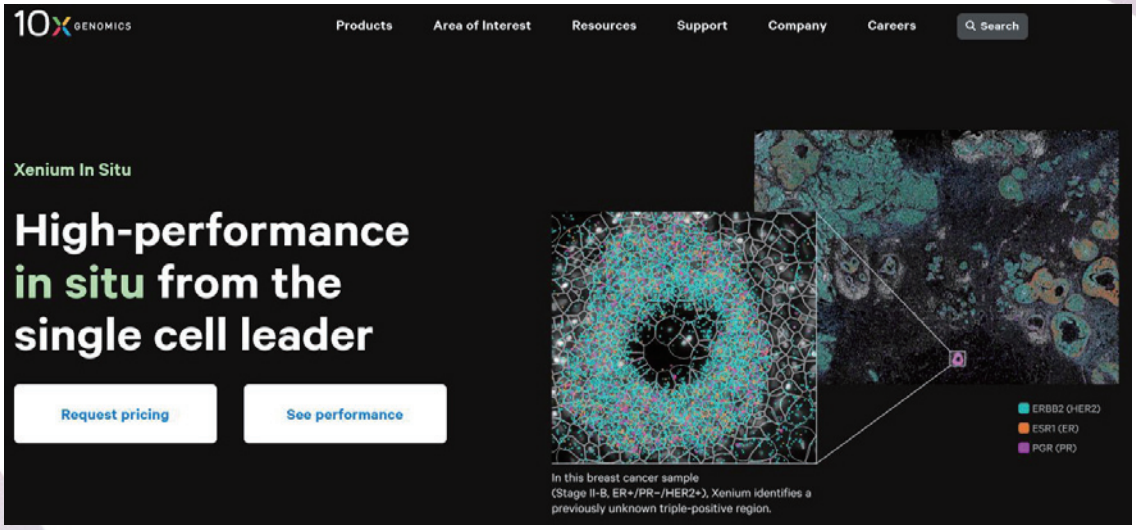
REPORT

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



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

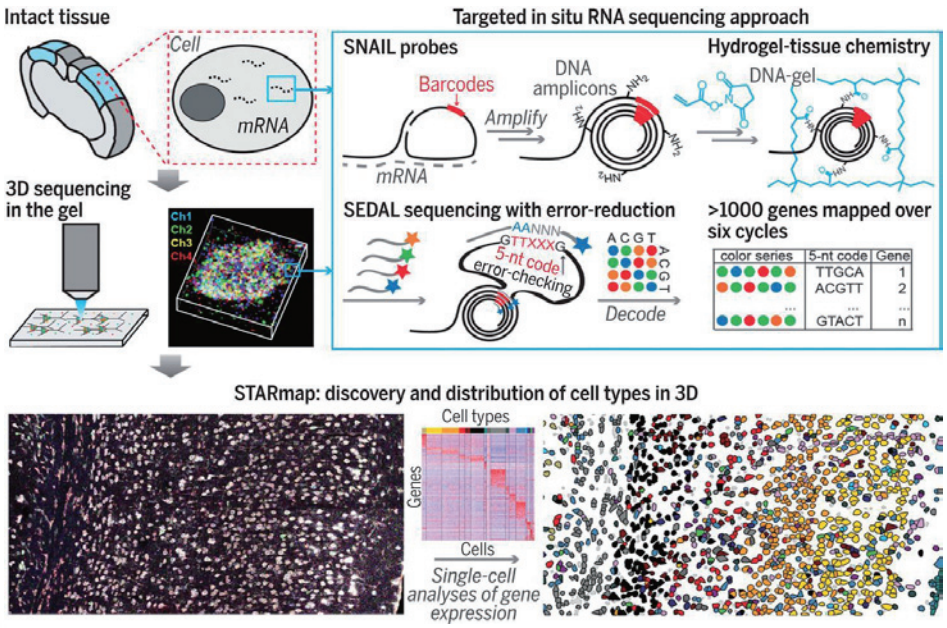
RESEARCH ARTICLE

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



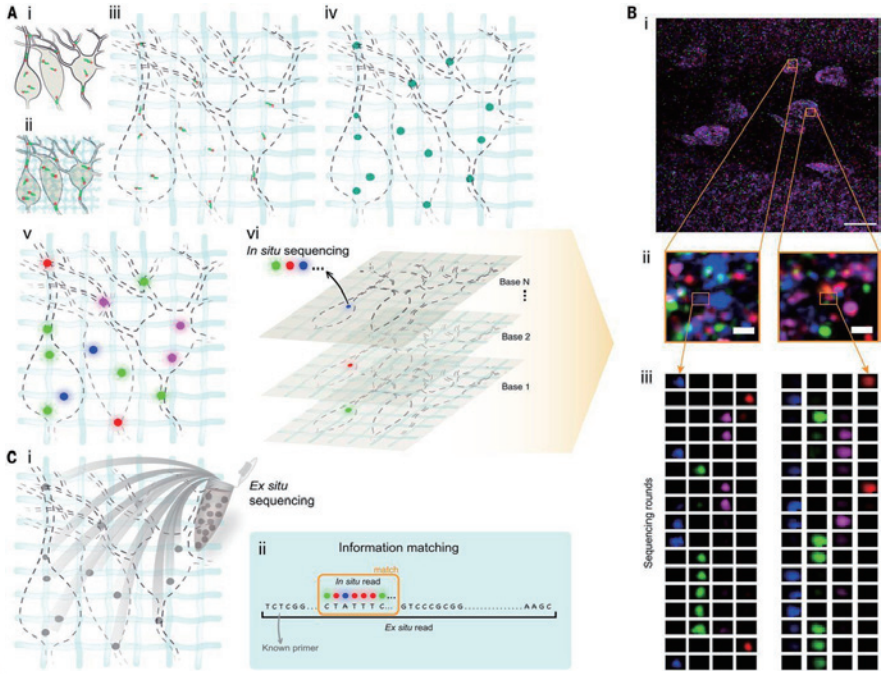
ExSeq

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

ExSeq



39

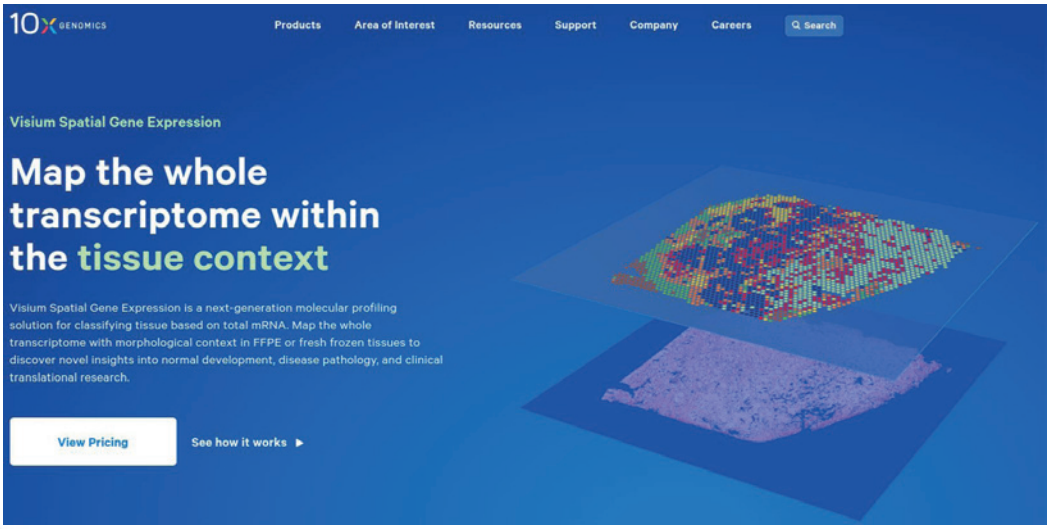
Spatial Transcriptomics Methods

NGS-based Methods

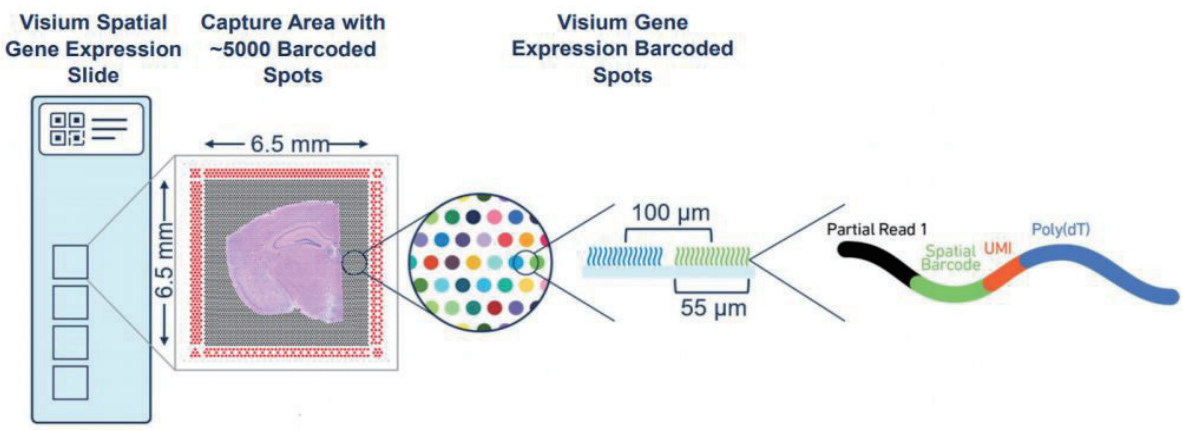
Visium, STOmics, Seq-Scope

40

Visium



Visium



STOmics (Stereo-seq)

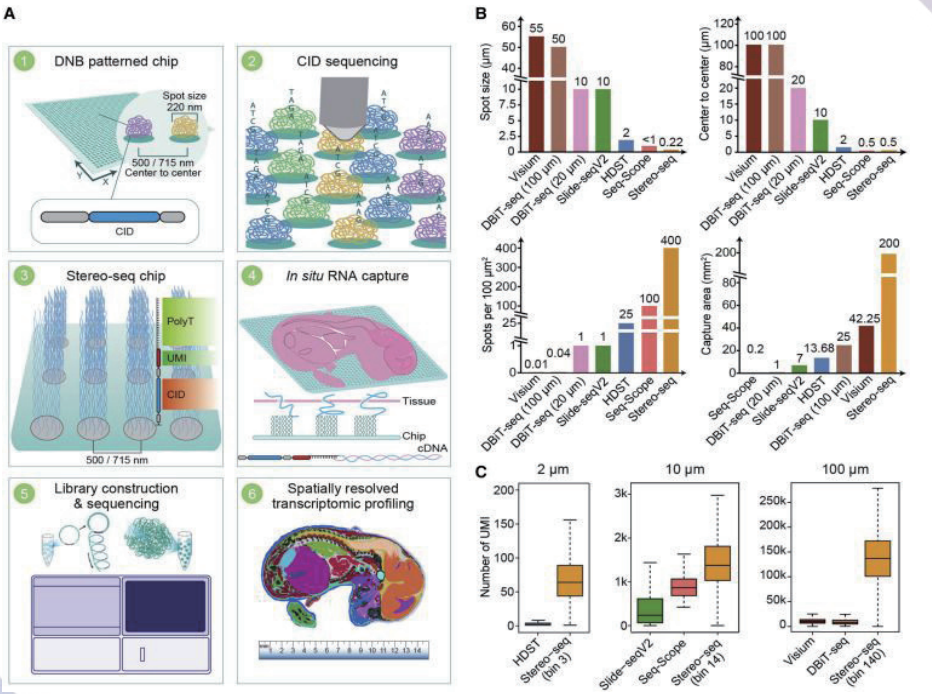
RESOURCE | VOLUME 185, ISSUE 10, P1777-1792.E21, MAY 12, 2022 Download Full Issue

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

Ao Chen [26] • Sha Liao [26] • Mengnan Cheng [26] • Kailong Ma [26] • Liang Wu [26] • Yiwei Lai [26] • Xiaojie Qiu [26] • Jin Yang • Jiangshan Xu • Shijie Hao • Xin Wang • Huifang Lu • Xi Chen • Xing Liu • Xin Huang • Zhao Li • Yan Hong • Yujia 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 • Yuxiang Li • Longqi Liu • Xun Xu [27] • Jian Wang • Show less • Show footnotes

Open Access • Published: May 04, 2022 • DOI: https://doi.org/10.1016/j.cell.2022.04.003 • Check for updates

STOmics (Stereo-seq)



Seq-Scope

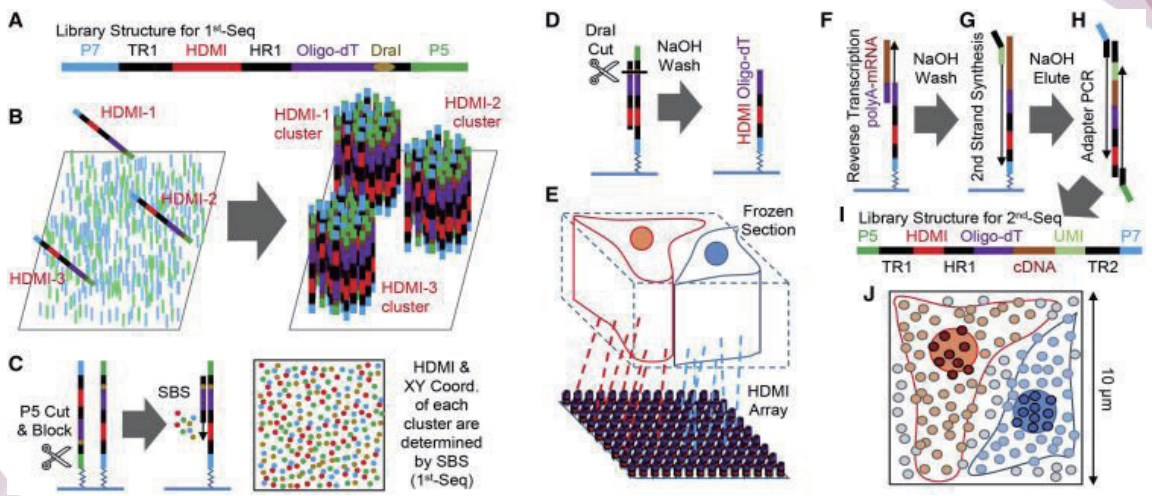
RESOURCE | VOLUME 184, ISSUE 13, P3559-3572.E22, JUNE 24, 2021 Download Full Issue

Microscopic examination of spatial transcriptome using Seq-Scope

Chun-Seok Cho [4] • Jingyue Xi [4] • Yichen Si • Sung-Rye Park • Jer-En Hsu • Myungjin Kim • Goo Jun • Hyun Min Kang • Jun Hee Lee [5] • Show less • Show footnotes

Open Archive • Published: June 10, 2021 • DOI: https://doi.org/10.1016/j.cell.2021.05.010 •

Seq-Scope



Methods Summary

FISH-Based

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

ISS-based

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

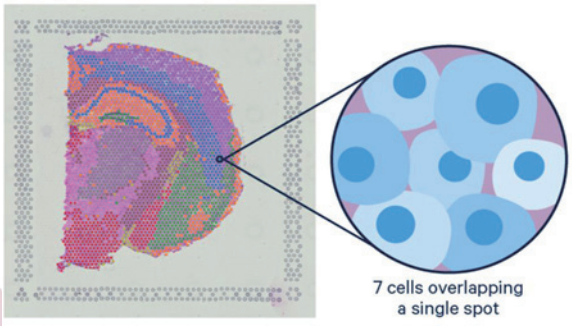
NGS-based

- Resolution
  - 0.22 - 100um
- 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^2

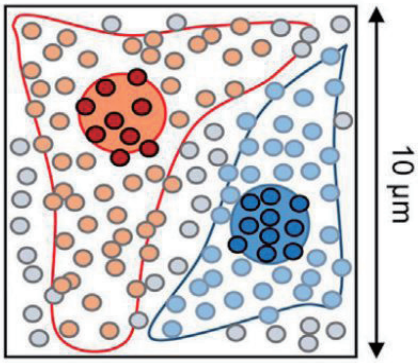
47

Spatial Resolution of Visium

Visium



Seq-Scope

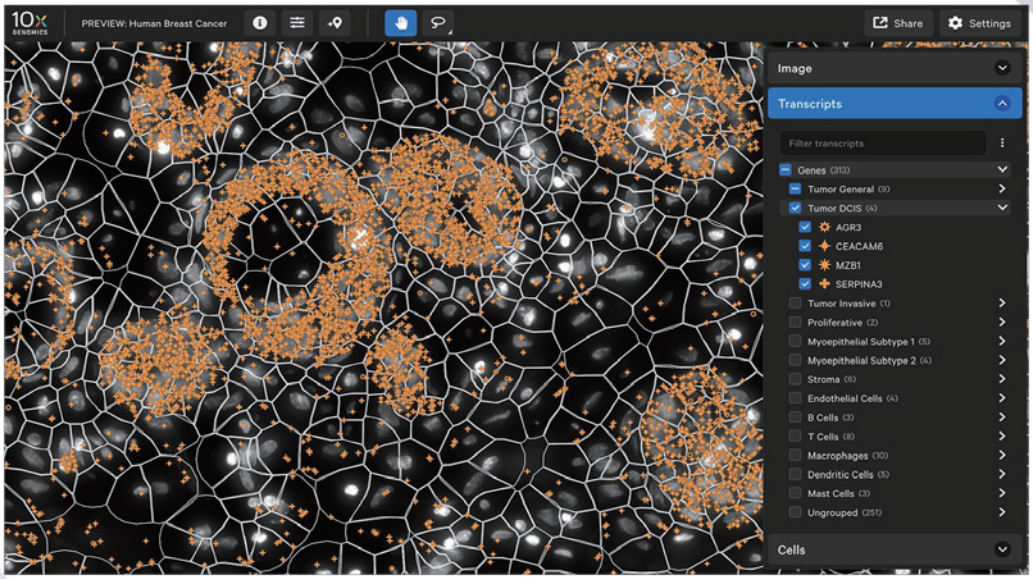


When the spatial resolution is not sufficient to distinguish single cells (e.g., Visium), you will need **deconvolution methods (e.g., Cell2location, Tangram, etc.)**

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

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

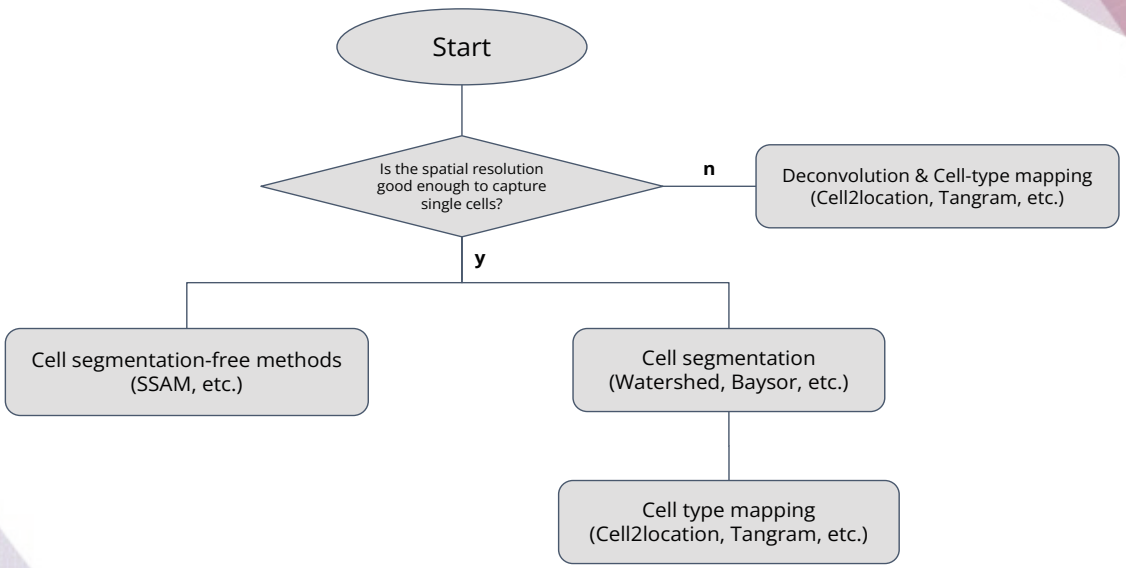
doi: https://doi.org/10.1101/2023.12.07.570603

"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

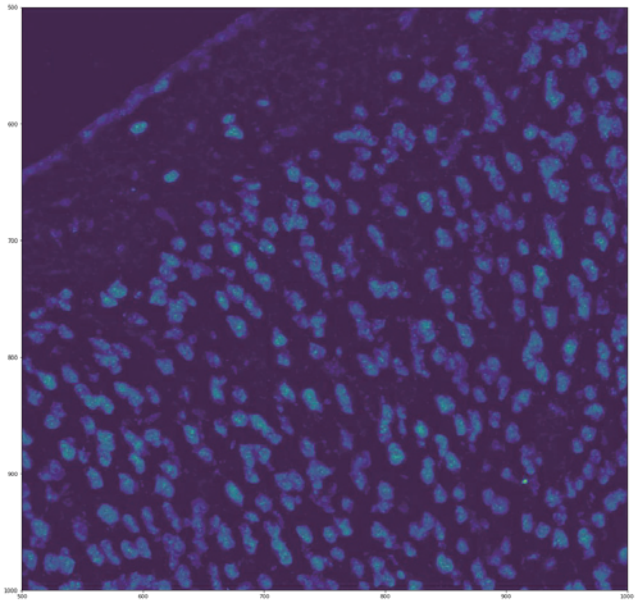
Spatial Omics Data Analysis Workflow



Chapter 2.

Spatial Transcriptomics Data Analysis w/ SSAM

Spatial Transcriptomics

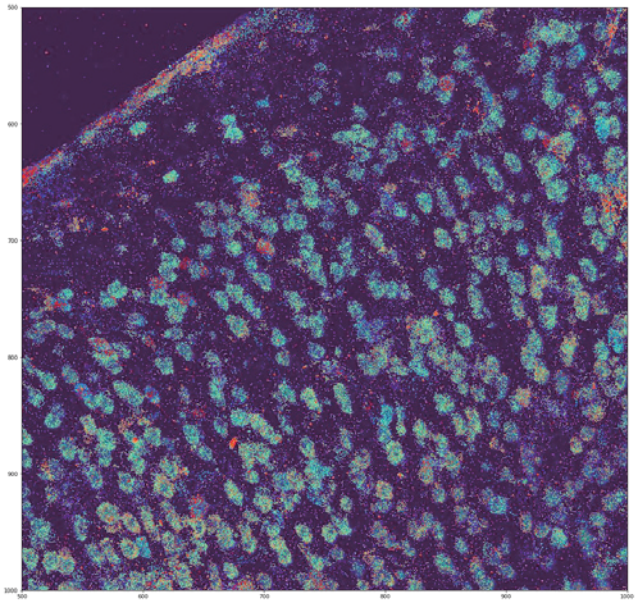


Poly-A image (all mRNAs)

Spatial Transcriptomics

Methods

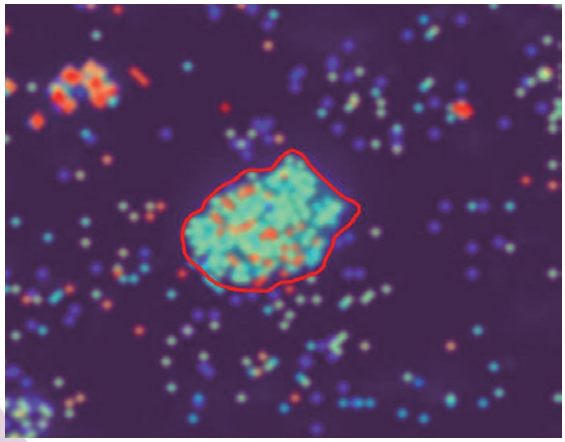
- osmFISH
- MERFISH
- seqFISH+



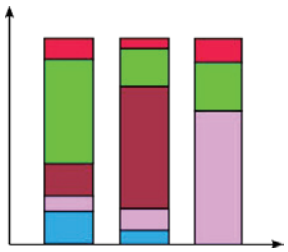
Poly-A image (all mRNAs)

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

Spatial Transcriptomics



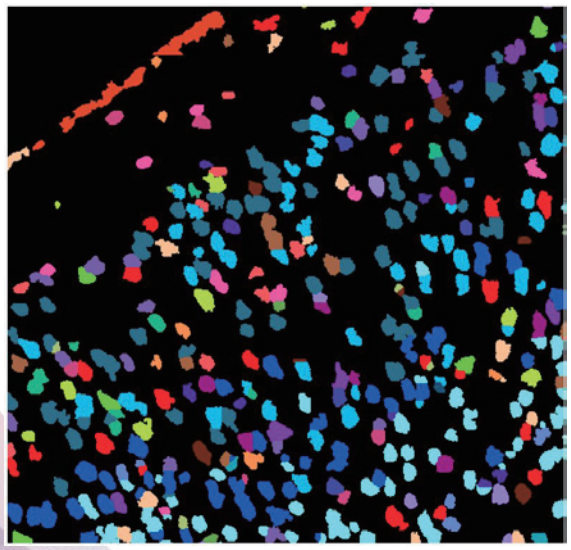
1. 세포의 경계 찾기 (Segmentation)
2. 경계 안의 mRNA 개수 세기



단일세포전사체학 데이터와 같은 방식으로 분석 가능!

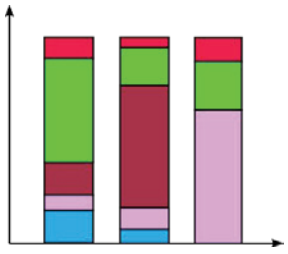
55

Spatial Transcriptomics



Cell segments, colored by cell types

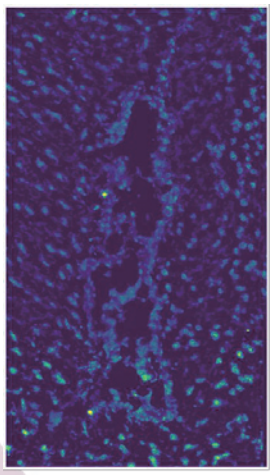
1. 세포의 경계 찾기 (Segmentation)
2. 경계 안의 mRNA 개수 세기



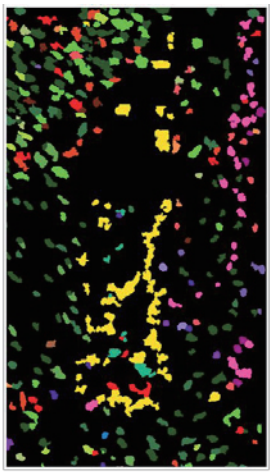
단일세포전사체학 데이터와 같은 방식으로 분석 가능!

56

현 공간체학 분석방법의 문제점



Poly-A Image
(all mRNAs)



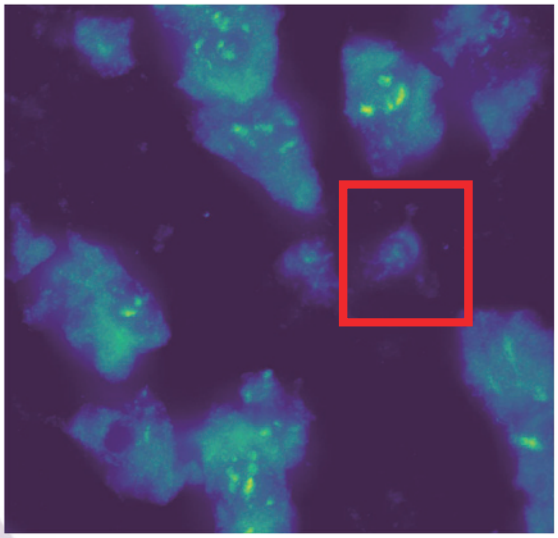
Cell segments
colored by cell types

- 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

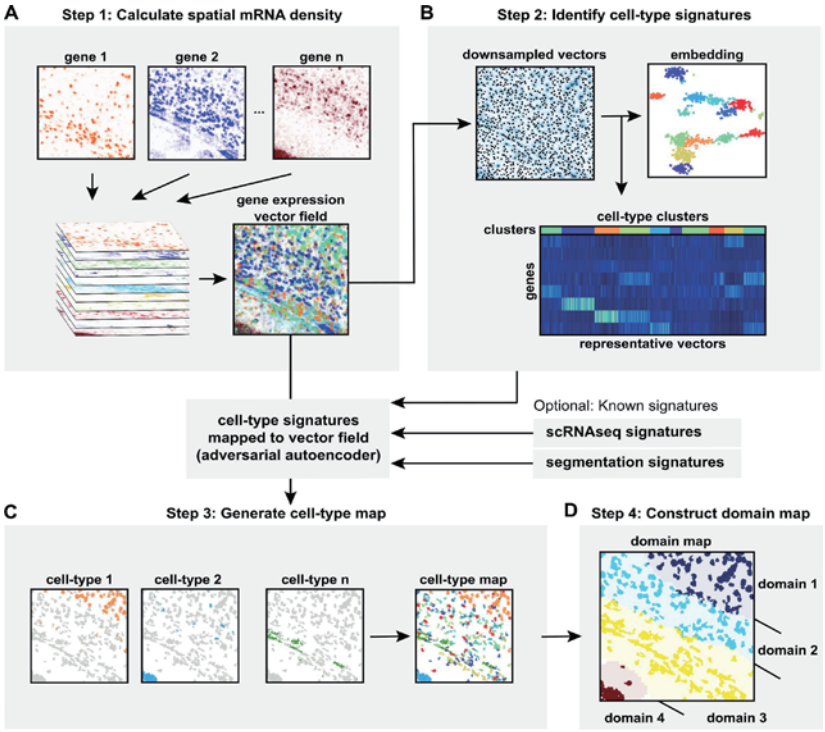


- Segmentation 이 쉽지 않음
  - Segmentation을 하지 않고 공간상에서의 cell-type을 분석할 수 있는 방법론 (SSAM) 개발

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

58

SSAM



Kernel Density Estimation

- i 번째 mRNA 상에 위치한 커널의 합으로 공간 밀도를 추정

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

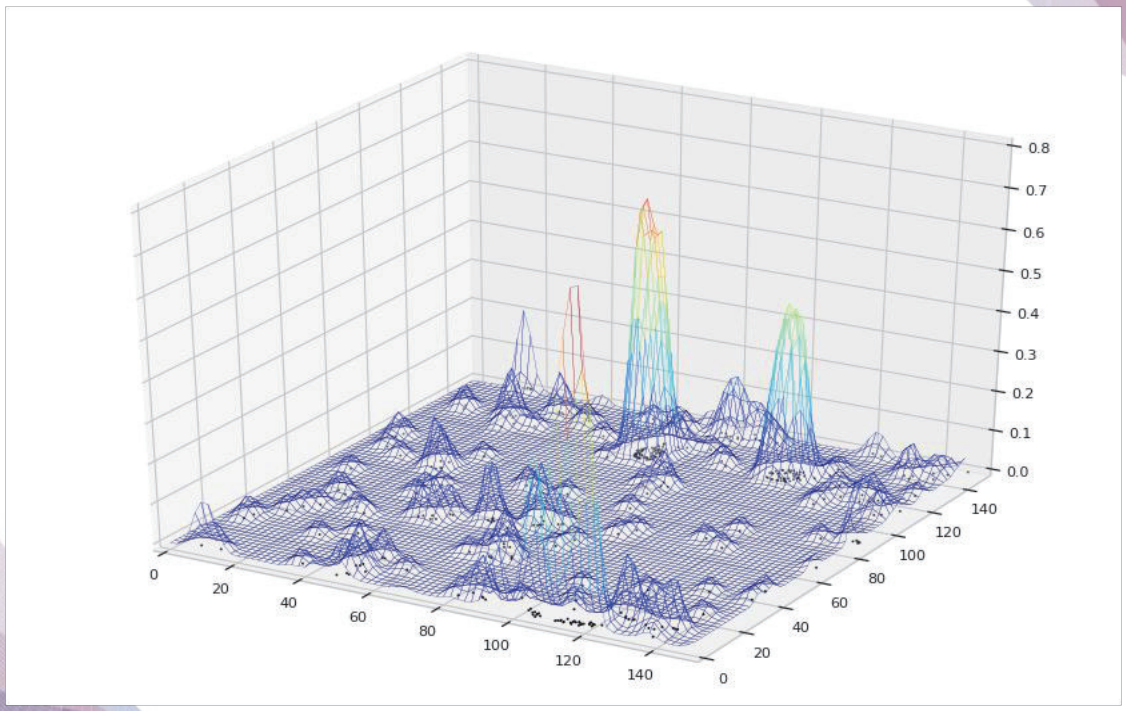
  - σ 는 밀도, N 은 mRNA 개수, K 는 커널, h 는 bandwidth 파라미터, x_i 는 i 번째 mRNA 위치

- 가우시안 커널

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

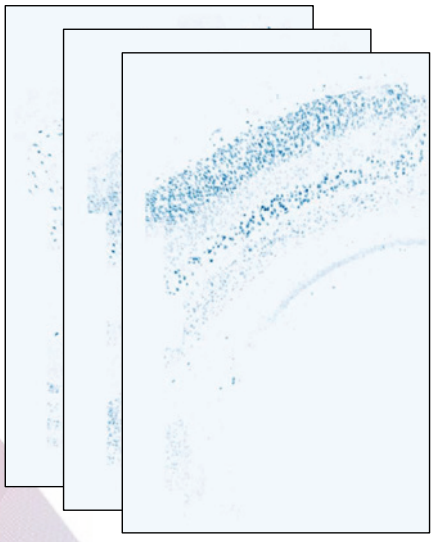
- 2*FWHM(σ) 이 세포의 지름과 비슷해지게 만드는 bandwidth 파라미터 h 선택 (세포의 지름 ~ 10um 으로 만들어주는 h=2 로 선택함)

Kernel Density Estimation



61

Kernel Density Estimation



- 33 종류의 유전자를 쌓아올림
  → 33 차원 벡터필드

- i번째 유전자 발현량 (E_i):

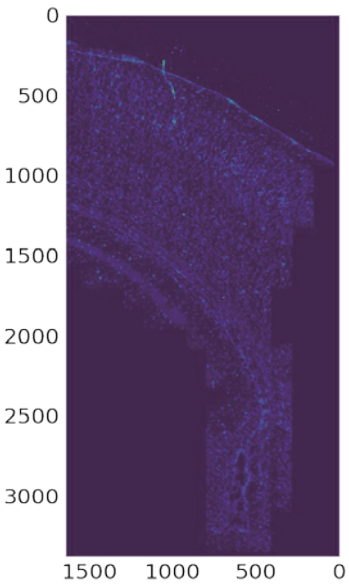
$$\hat{E}_i(\vec{x}) = \hat{\sigma}_i(\vec{x})\, N_i$$

where

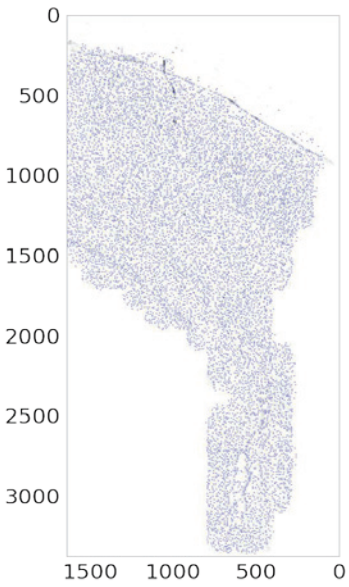
E_i : *expression of* $i-th$ *gene*
σ_i : *estimated density of* $i-th$ *gene*
N_i : *number of mRNAs of* $i-th$ *gene*

62

다운샘플링



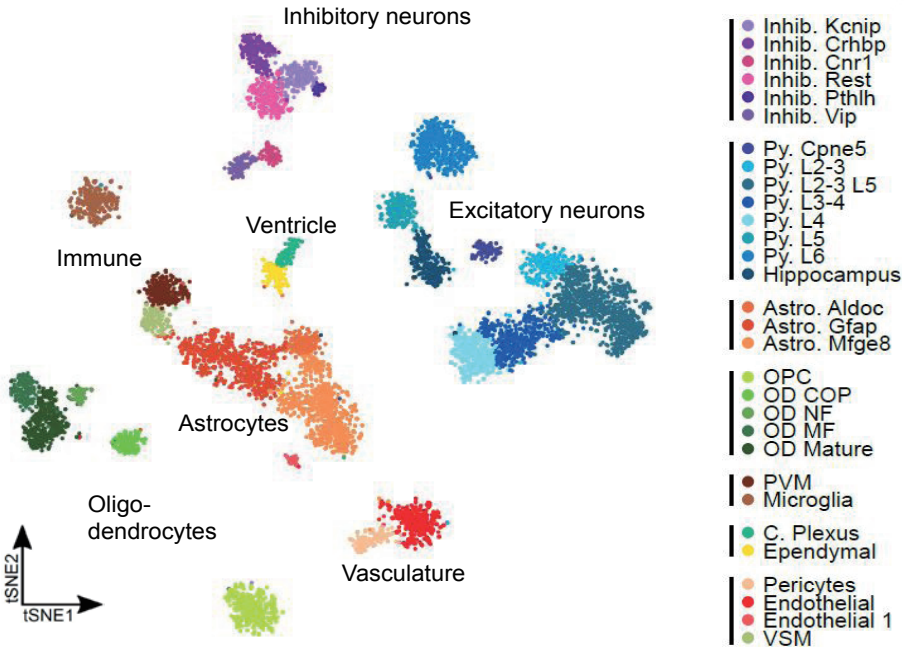
3,380 x 2,080 = 7,030,400 vectors



L1 maxima

63

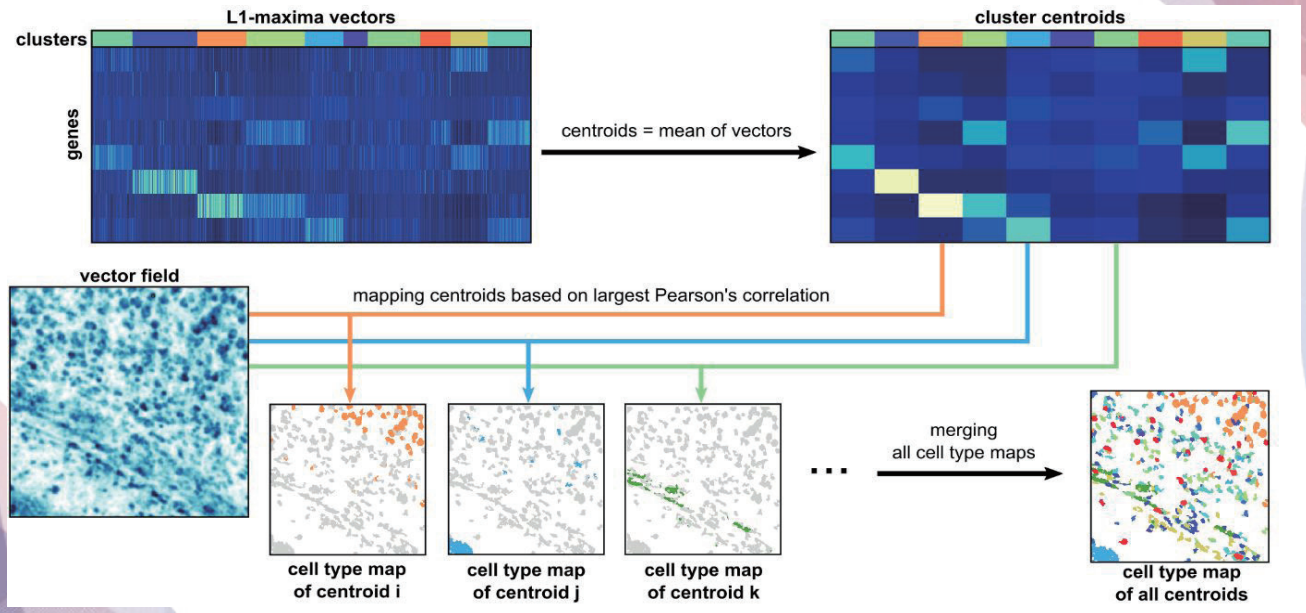
클러스터링



주의: 각 점은 "벡터" 임 (세포가 아님)

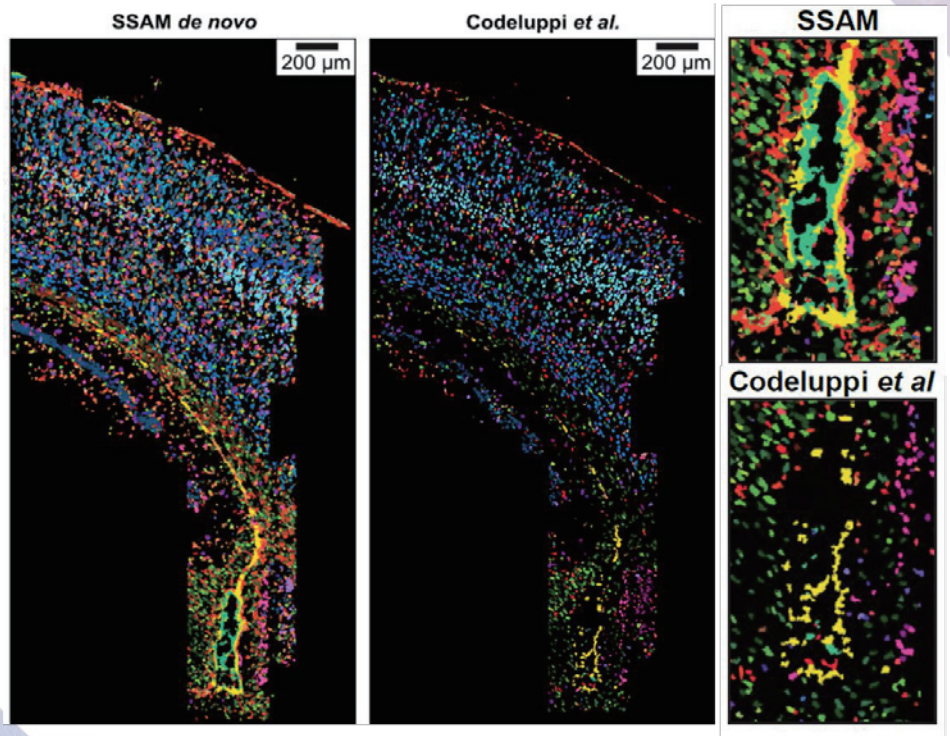
64

세포 타입 맵 (Cell-type map) 생성



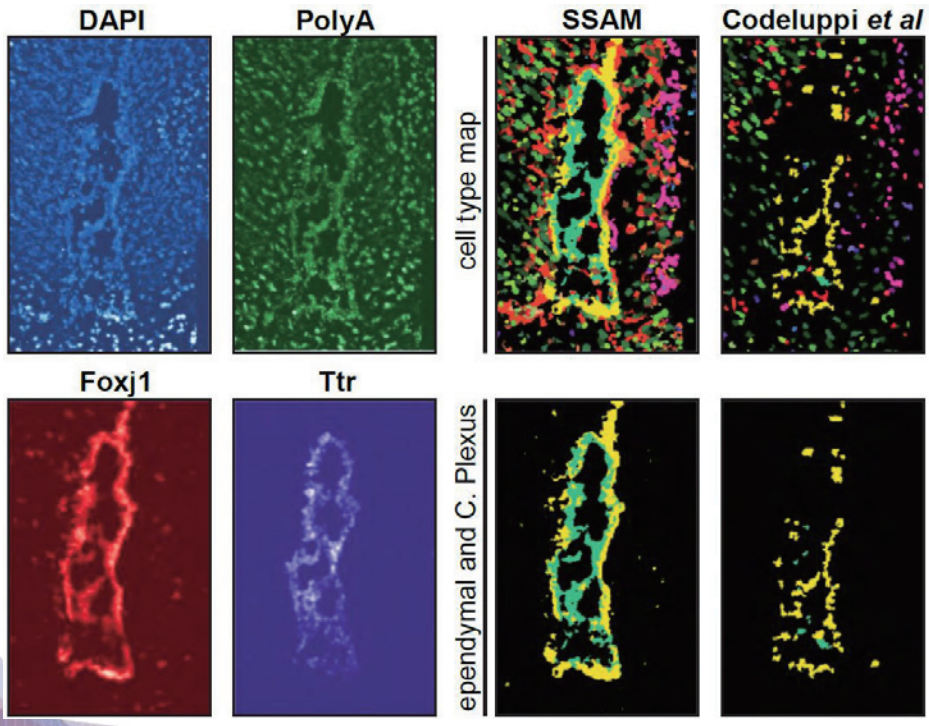
65

세포 타입 맵 (Cell-type map)



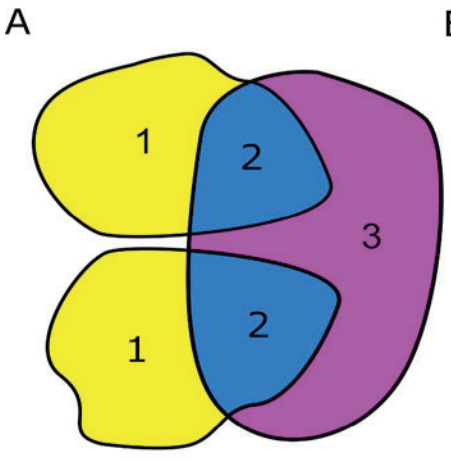
66

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



67

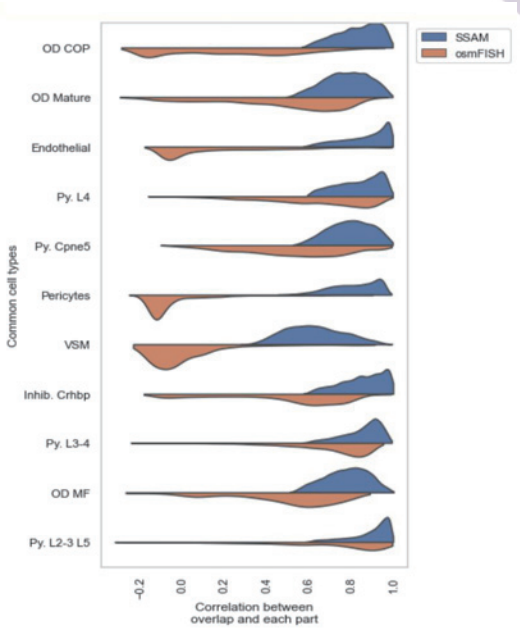
성능 비교



1: osmFISH segmentation part

2: Parts overlap

3: SSAM cell-type map part

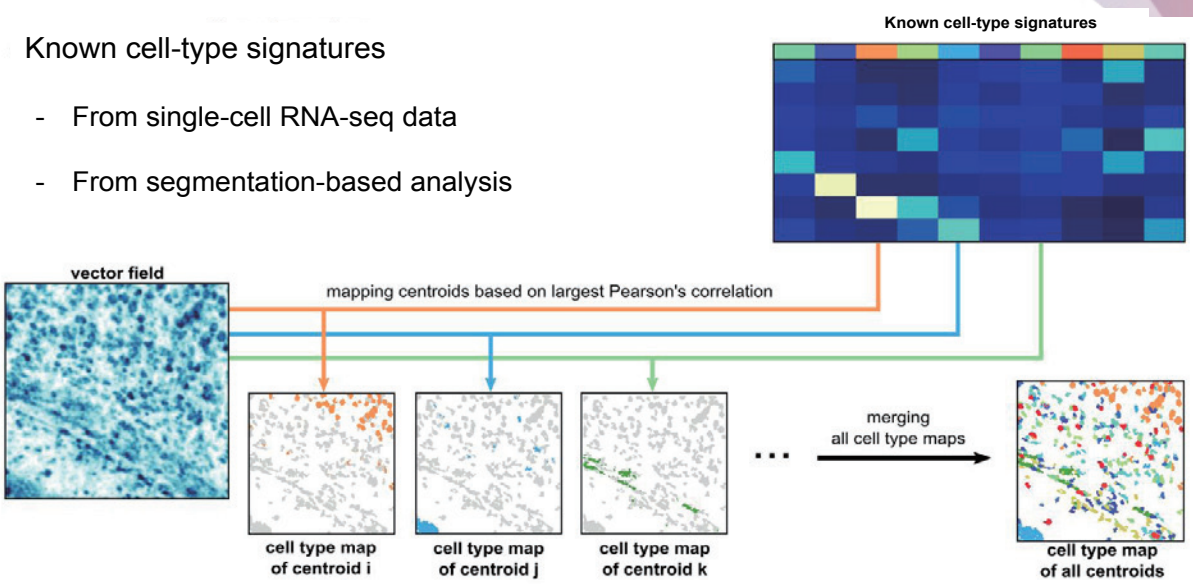


68

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

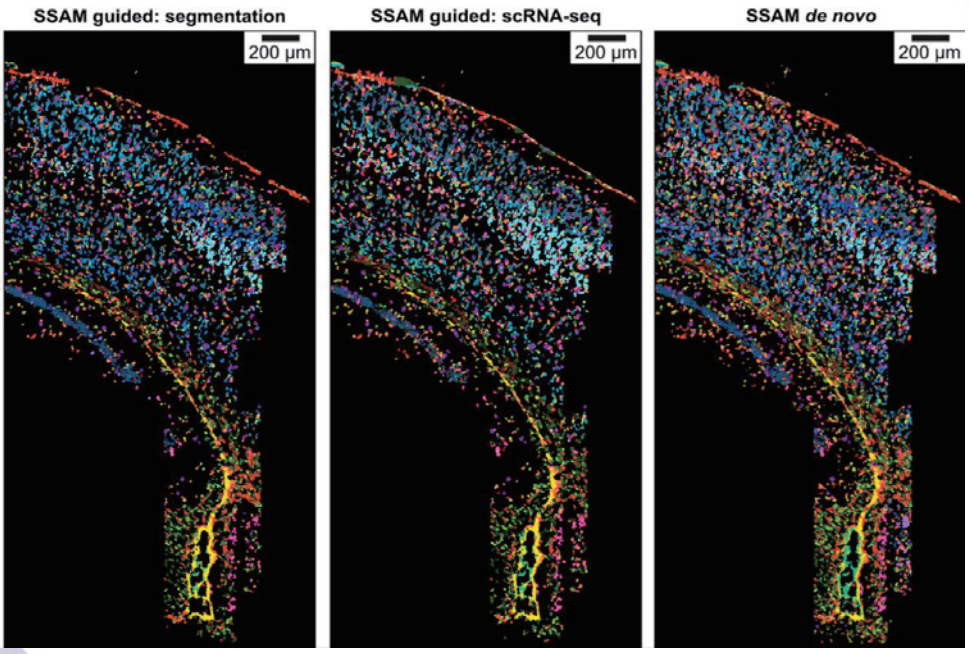
Known cell-type signatures

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



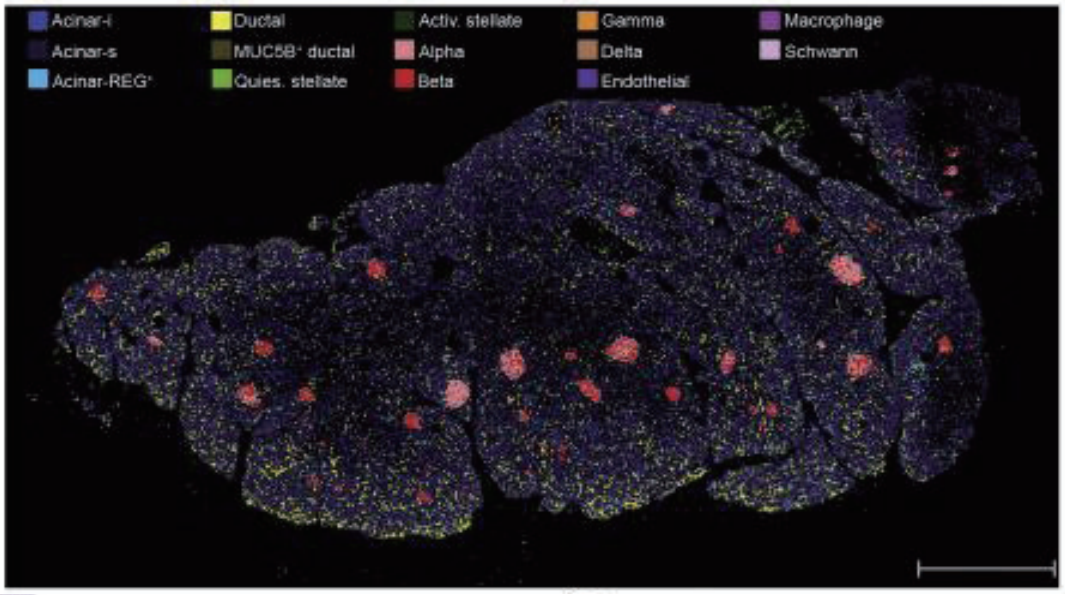
69

SSAM Guided-mode 분석



70

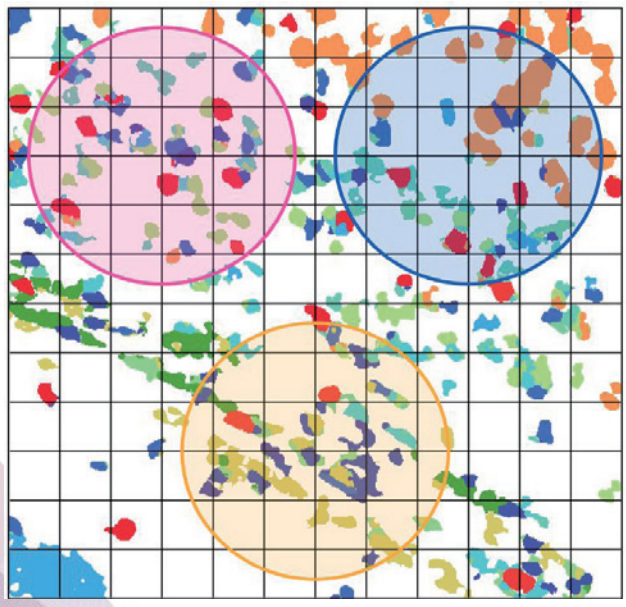
췌장 In situ sequencing 데이터 분석 (guided mode)

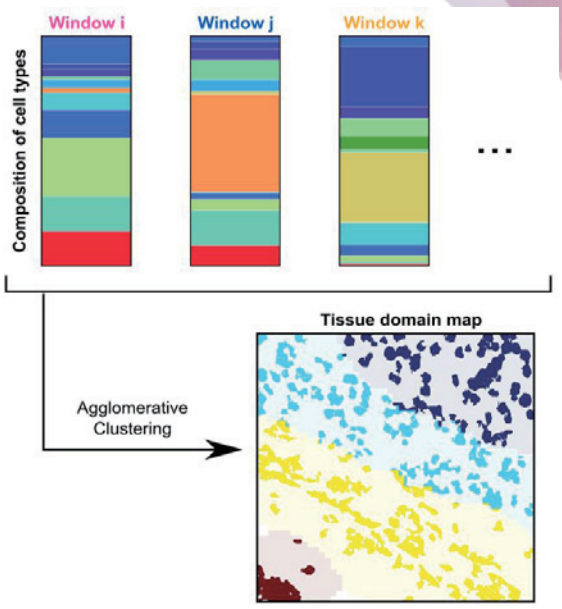


Gastroenterology 160(4) 1330-1344.E11, DOI: https://doi.org/10.1053/j.gastro.2020.11.010

71

조직 구조 (domain) 분석



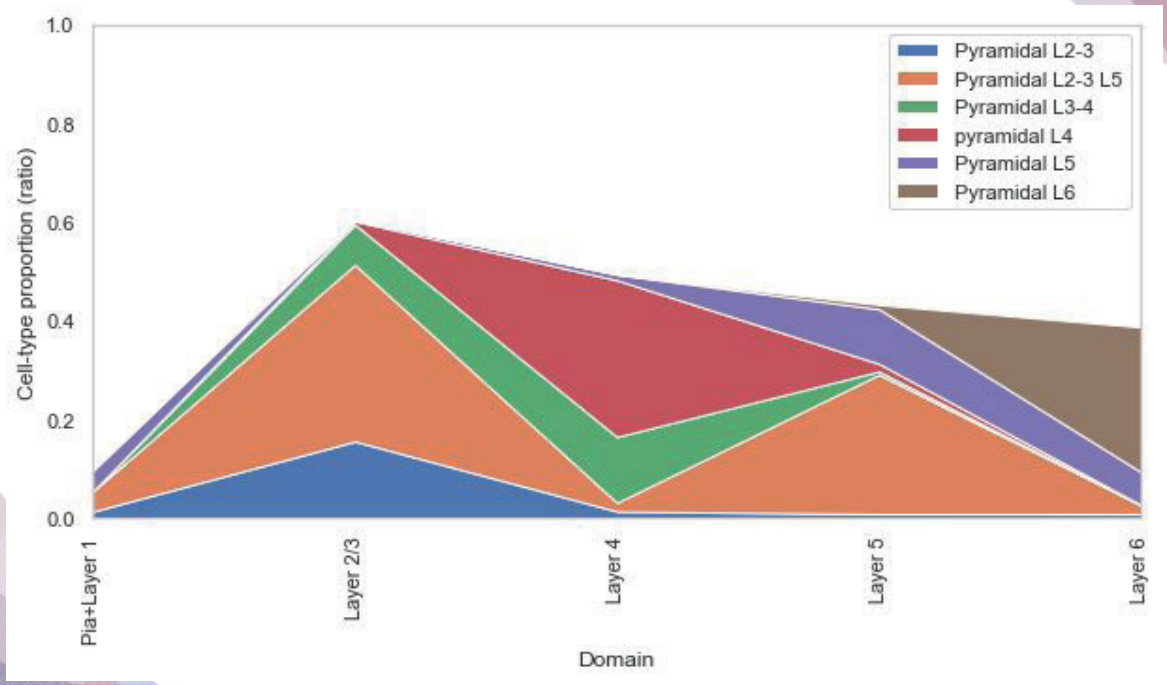


72

조직 구조 (domain) 분석

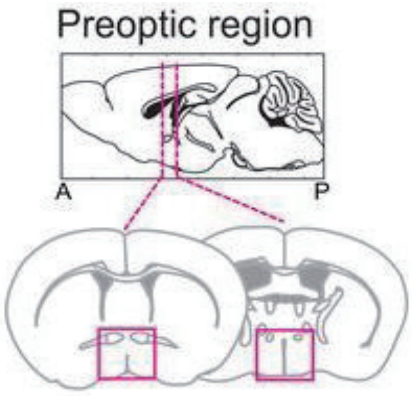
73

도메인 검증



74

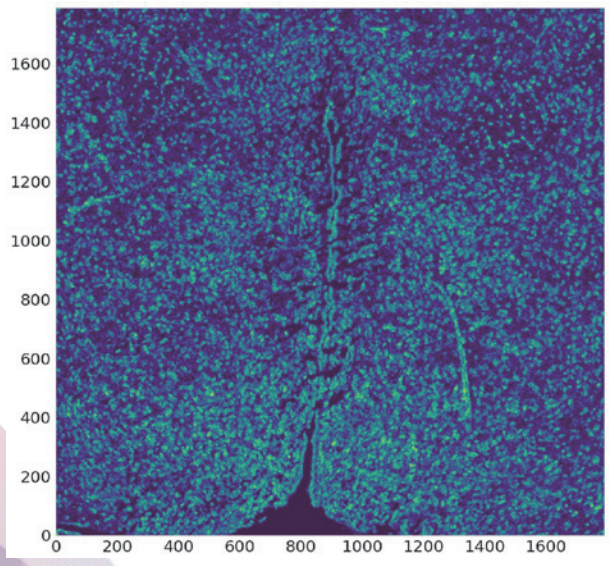
생쥐 시상하부 시각 교차 앞구역 (Preoptic area, MERFISH)



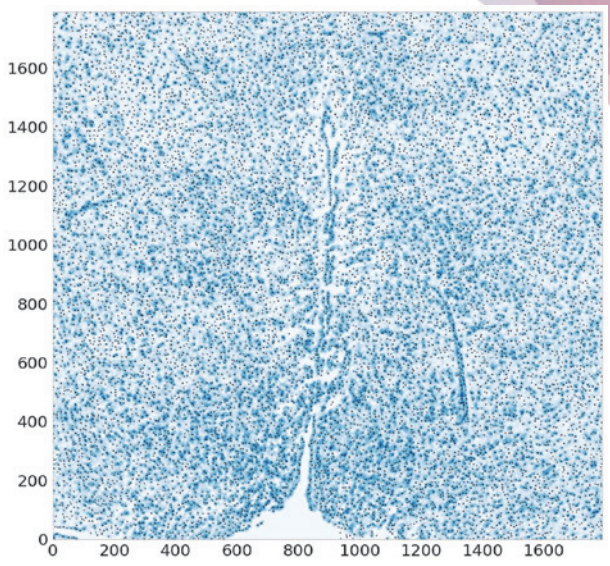
- 생쥐 시상해부 시각교차앞구역 (preoptic area, POA)
- 1.8mm x 1.8mm x 0.6mm
- MERFISH로 이미징한 mRNA 3D 위치
- 135 종 유전자

75

공간밀도추정 및 벡터 다운샘플링



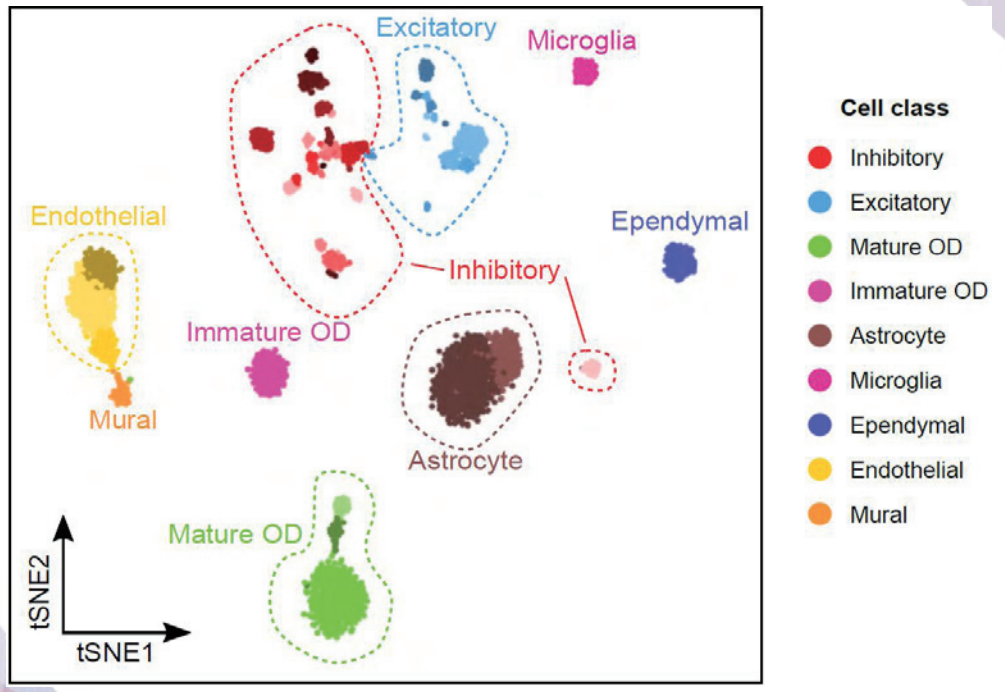
1,790 x 1,790 x 9 = 28,836,900 vectors
(z=4 um)



L1 maxima

76

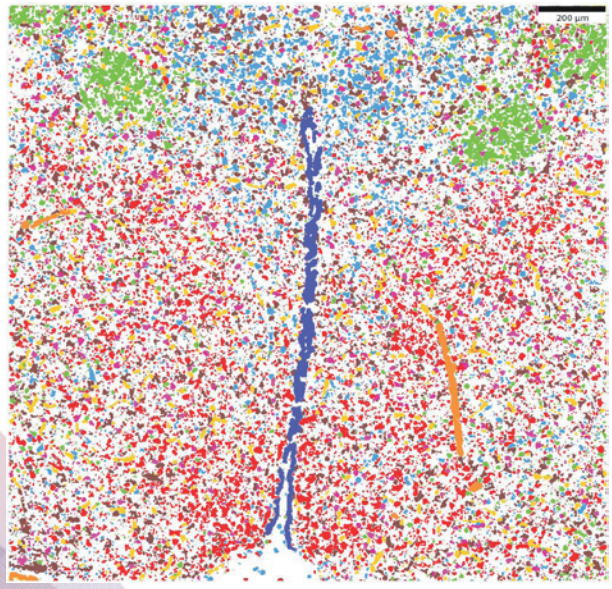
클러스터링



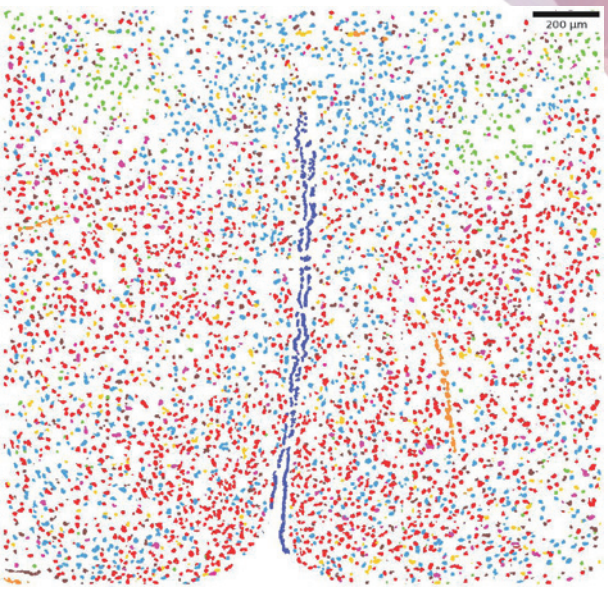
77

결과 비교

SSAM *de novo* (z=3um)

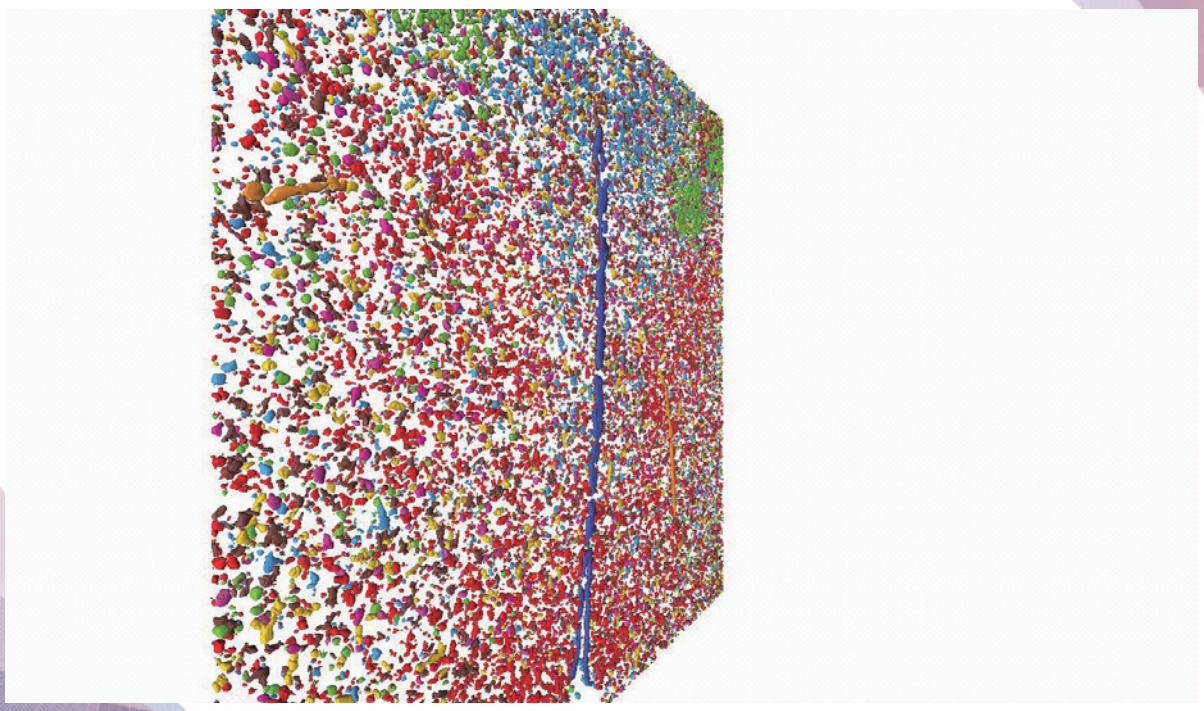


Moffit *et al*.



78

3D 세포 타입 맵



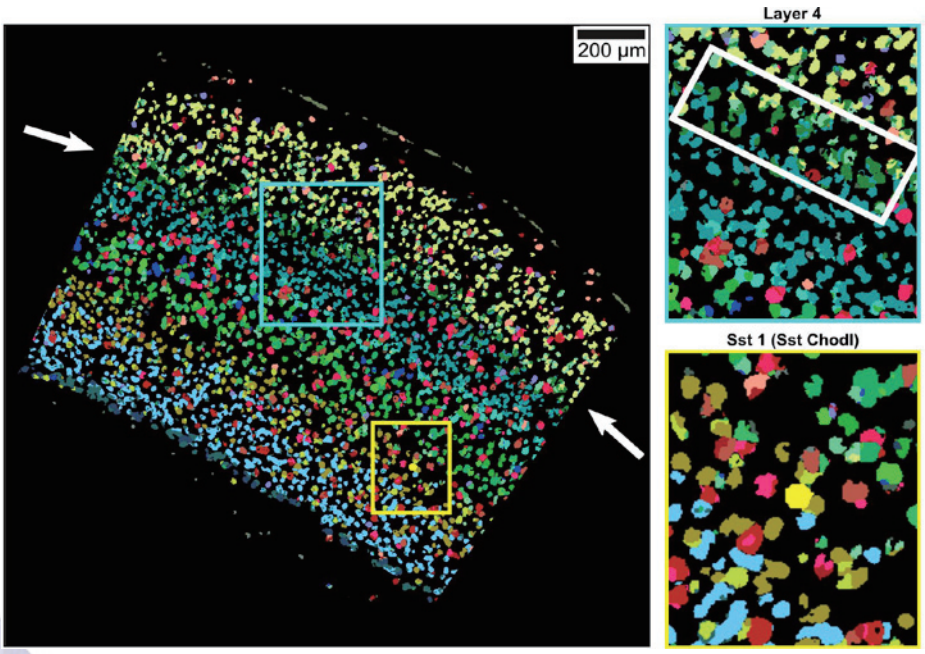
79

Allen multiplexed smFISH 데이터

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

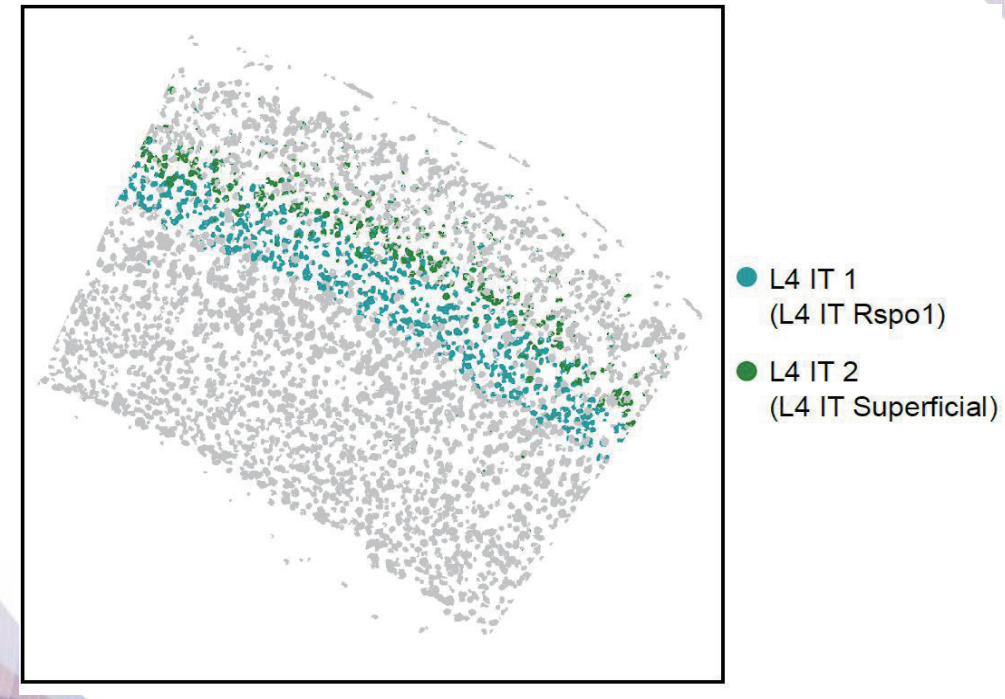
80

세포 타입 맵 (cell-type map)



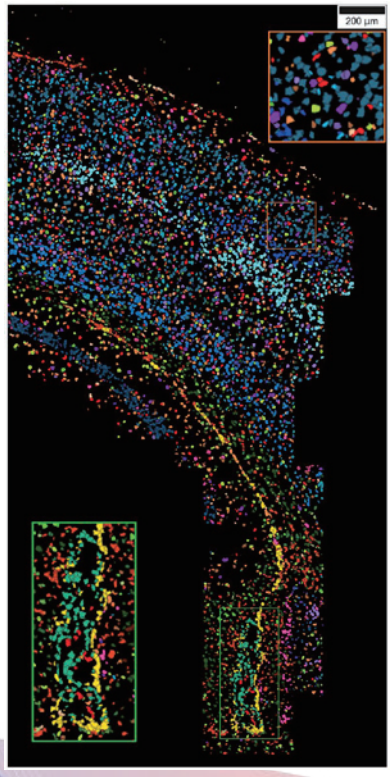
81

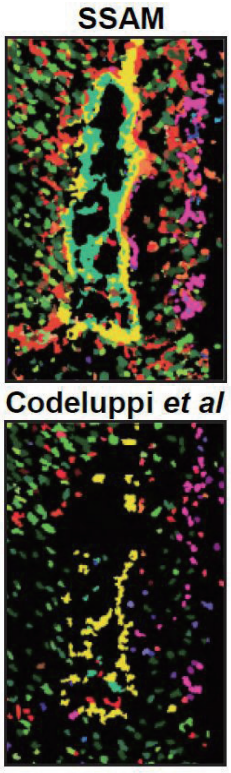
L4 IT 세포 타입의 분리



82

세포 타입 맵을 활용한 세포 경계 검출 (segmentation)





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

Allen Brain Institute, USA

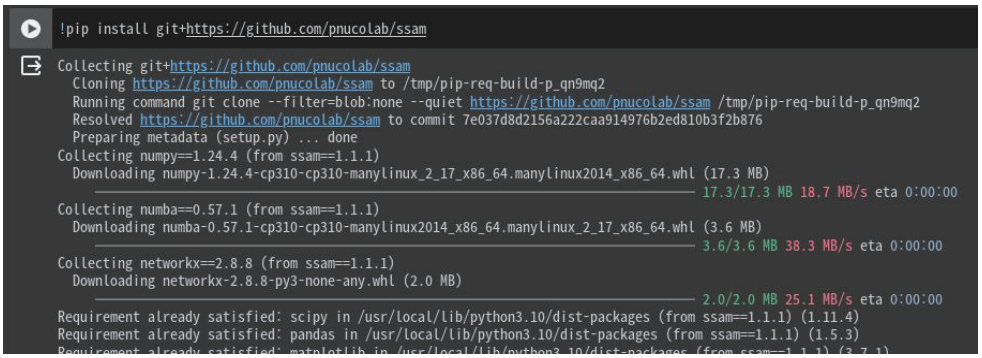
• Brian Long
• Emma Garren
• Thuc Nghi Nguyen
• Bosiljka Tasic
• Jeffery Miller
• Ambrose Carr
• Ed Lein

튜토리얼

- SSAM 설치:

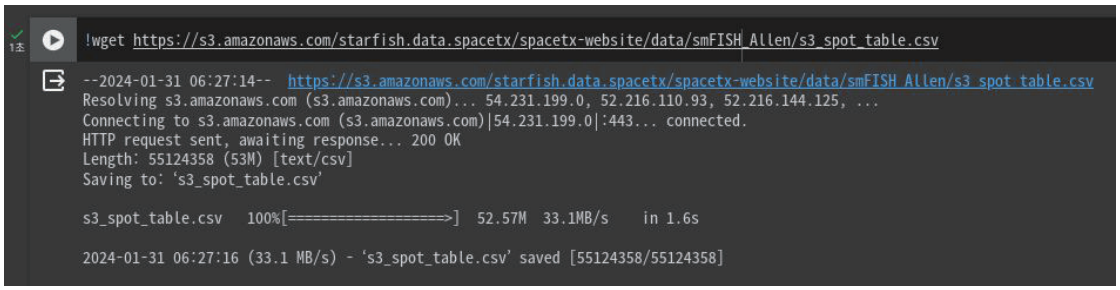
```
pip install git+https://github.com/pnucolab/ssam
```

- Google Colab에서 SSAM 설치:



튜토리얼

- 예제 데이터 내려받기
  - https://s3.amazonaws.com/starfish.data.spacetx/spacetx-website/data/smFISH_Allen/s3_spot_table.csv



https://spacetx.github.io/

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	0	Fezf2	4	1.00000	1605	1.0	1296.718027	813.958480	1296.718027	813.958480
1	1	Fezf2	17	1.00000	2496	1.0	1302.424423	844.131209	1302.424423	844.131209
2	2	Fezf2	36	1.00000	1605	1.0	1293.752097	822.910508	1293.752097	822.910508
3	3	Fezf2	77	0.99999	1605	1.0	1292.981468	825.102131	1292.981468	825.102131
4	4	Fezf2	86	1.00000	3929	1.0	1254.284721	915.218969	1254.284721	915.218969
...	...	...	...	...	...	...	...	...	...	...
530778	530778	Parm1	1071304	0.99995	407	1.0	115.997457	502.193148	115.997457	502.193148
530779	530779	Parm1	1071313	1.00000	4354	1.0	134.008161	429.477827	134.008161	429.477827
530780	530780	Parm1	1071319	0.99972	3478	1.0	107.807802	398.245527	107.807802	398.245527
530781	530781	Parm1	1071321	0.99999	999	1.0	120.079417	451.938345	120.079417	451.938345
530782	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...
```

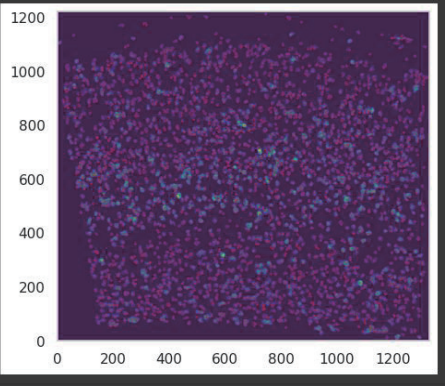

튜토리얼

- Local Maxima 찾기

```
[18]  analysis.find_localmax()

      Found 2787 local max vectors.
```

- Local Maxima 확인하기

```
ds.plot_linorm()
ds.plot_localmax(s=0.1, c='red')
```



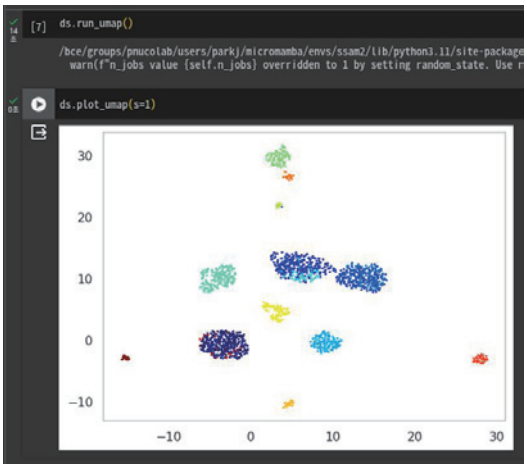
튜토리얼

- 클러스터링하기

```
[6]  analysis.cluster_vectors()

     Found 15 clusters
```

- UMAP 그리고 확인하기

```
[7]  ds.run_umap()

     /bce/groups/pnucolab/users/parkj/micromamba/envs/ssam2/lib/python3.11/site-package
       warn(f"n_jobs value {self.n_jobs} overridden to 1 by setting random_state. Use r

     ds.plot_umap(s=1)
```

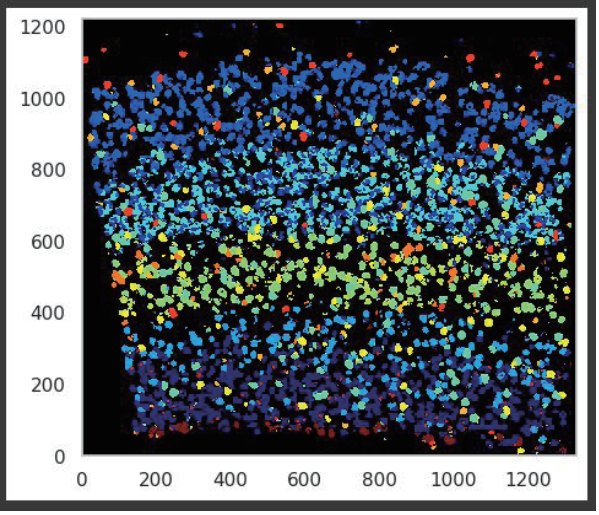


튜토리얼

- Cell-type map 그리기

```
analysis.map_celltypes()

Generating cell-type map for centroid #0...
Processing chunk (0/1)...
Generating cell-type map for centroid #1...
Processing chunk (0/1)...
Generating cell-type map for centroid #2...
Processing chunk (0/1)...
Generating cell-type map for centroid #3...
Processing chunk (0/1)...
Generating cell-type map for centroid #4...
Processing chunk (0/1)...
Generating cell-type map for centroid #5...
Processing chunk (0/1)...
Generating cell-type map for centroid #6...
Processing chunk (0/1)...
Generating cell-type map for centroid #7...
Processing chunk (0/1)...
Generating cell-type map for centroid #8...
Processing chunk (0/1)...
Generating cell-type map for centroid #9...
Processing chunk (0/1)...
Generating cell-type map for centroid #10...
Processing chunk (0/1)...
Generating cell-type map for centroid #11...
Processing chunk (0/1)...
Generating cell-type map for centroid #12...
Processing chunk (0/1)...
Generating cell-type map for centroid #13...
Processing chunk (0/1)...
Generating cell-type map for centroid #14...
Processing chunk (0/1)...

[11] analysis.filter_celltypemaps()
```

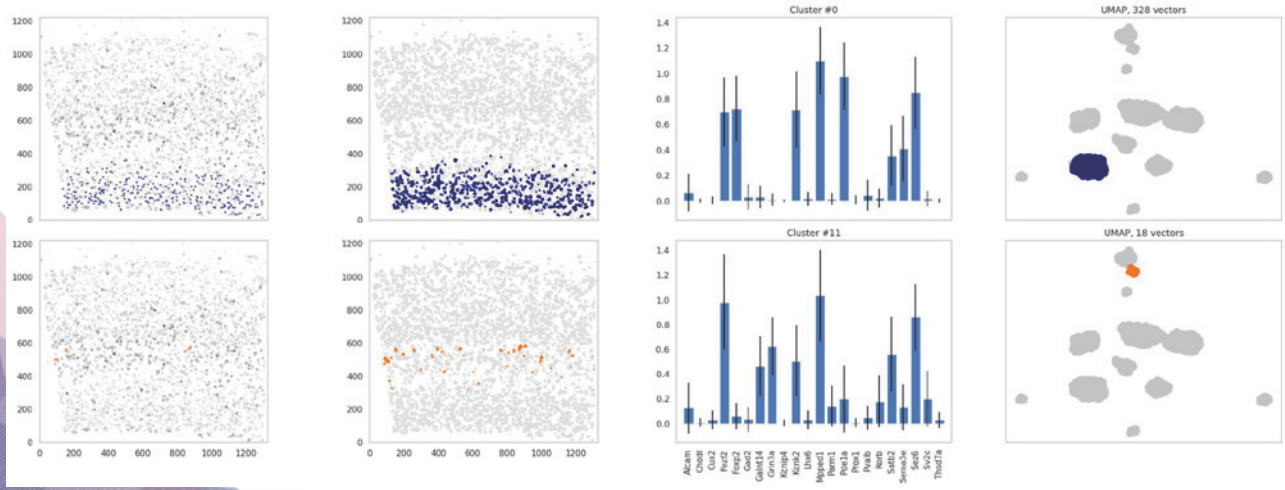
```
ds.plot_celltypes_map()
```



튜토리얼

- Cell type 별 gene expression 구하기 (diagnostic plots)

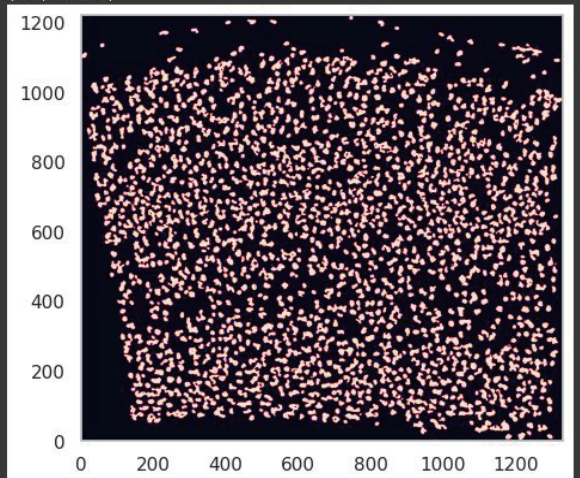
```
import matplotlib.pyplot as plt

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



튜토리얼

- Cell Segmentation 구하기 - Marker 찾기

```
[131] from skimage import filters
      vfnorm_threshold = filters.threshold_local(ds.vf_norm, 35)
      vfnorm_thresh_im = (ds.vf_norm > vfnorm_threshold).squeeze().compute()

      plt.imshow(vfnorm_thresh_im.swapaxes(0, 1))
      plt.ylim([0, vfnorm_thresh_im.shape[1]])

      (0.0, 1220.0)
```



튜토리얼

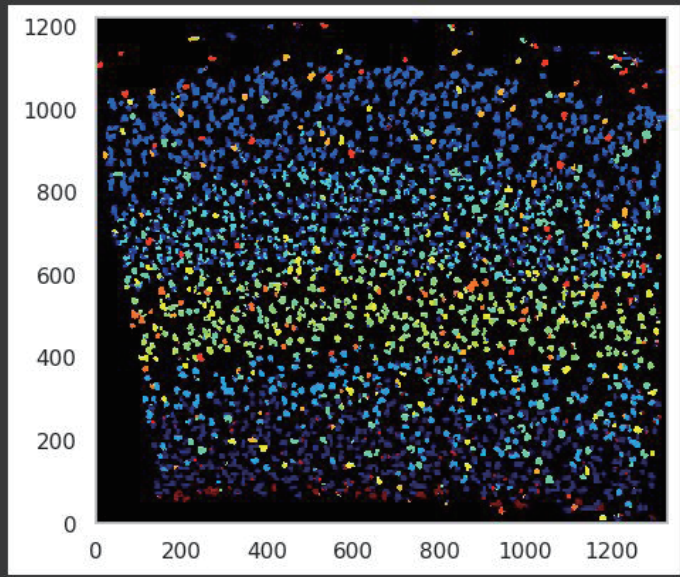
- Cell Segmentation 수행하기

```
[133] analysis.run_watershed(vfnorm_thresh_im)

      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...
```


튜토리얼

- Cell Segmentation 기반 Cell-type Map 그리기

```
ds.plot_watershed_celltypes_map()
```



튜토리얼

- Guided mode를 위한 데이터 내려받기
  - 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/s
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
```

gene_name	Alcam	Chodl	Cux2	Fezf2	Foxp2	Gad2	Galnt14	Grin3a	Kcnip4
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 × 22 columns

98

튜토리얼

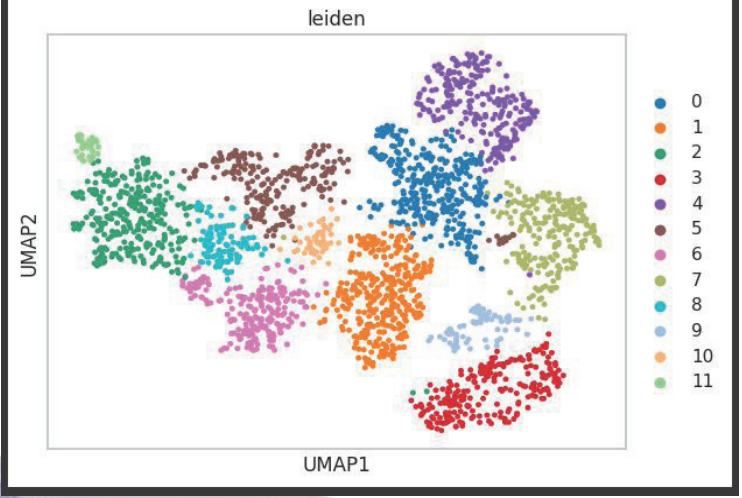
- ScanPy로 데이터 분석하기 (Single-cell 분석 튜토리얼 참고)

```
[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)
```


튜토리얼

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

```
sc.tl.umap(adata)

sc.pl.umap(adata, color='leiden')

/bce/groups/pnucolab/users/parkj/micromamba/envs/ssam2/lib/python3.11/site-packages/scanpy
  color_vector = pd.Categorical(values.map(color_map))
/bce/groups/pnucolab/users/parkj/micromamba/envs/ssam2/lib/python3.11/site-packages/scanpy
  cax = scatter(
```



튜토리얼

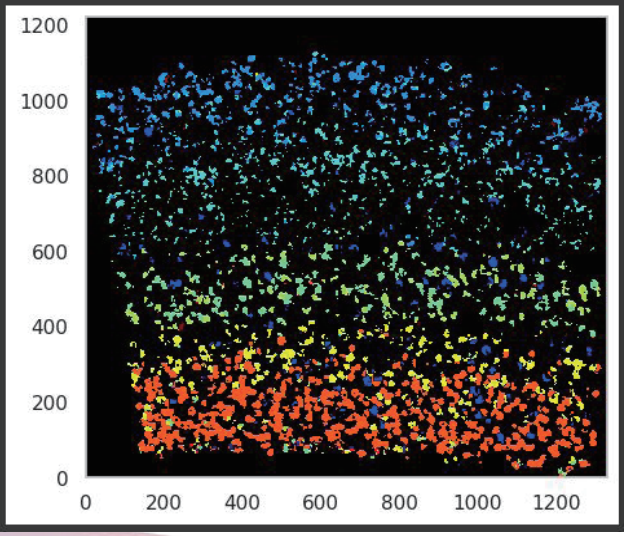
- Centroid 기반 Cell-type map 생성하기

```
[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...
```


튜토리얼

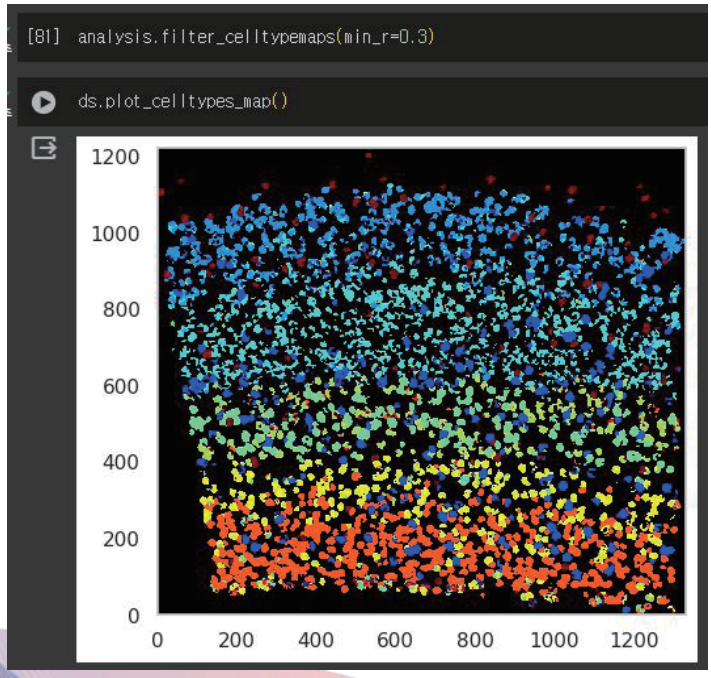
- Centroid 기반 Cell-type map 생성하기

```
[77]  analysis.filter_celltypemaps()

[78]  ds.plot_celltypes_map()
```



튜토리얼

- Centroid 기반 Cell-type map 생성하기



103