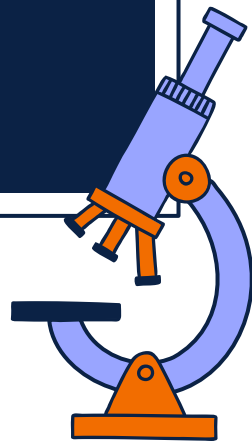


# SPATIALLY RESOLVED TRANSCRIPTOMICS

Phân tích hệ phiên mã theo cấu trúc không gian

Nguyễn Minh Huyền  
TS. Nguyễn Hồng Quang





# MỤC LỤC



01

## GIỚI THIỆU

Giới thiệu về hệ phiên mã theo cấu trúc không gian

02

## CÔNG NGHỆ VISIUM

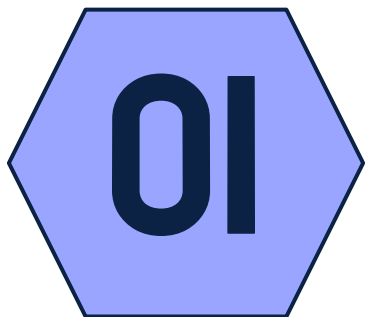
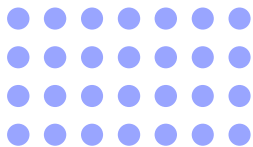
Công nghệ phân tích hệ phiên mã theo cấu trúc không gian của 10X Genomics

03

## STLEARN

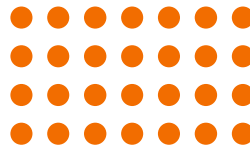
Một phương pháp ứng dụng phân tích tổng hợp dữ liệu về phiên mã không gian





# GIỚI THIỆU

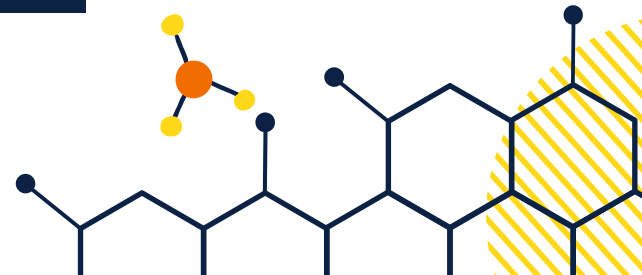
Giới thiệu về hệ phiên mã theo cấu trúc không gian





TẠI SAO LẠI THEO CẤU TRÚC KHÔNG GIAN ?

Định vị vị trí rất quan trọng



# TẠI SAO LẠI THEO CẤU TRÚC KHÔNG GIAN ?

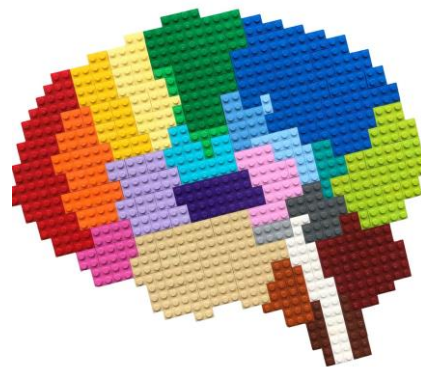
Định vị vị trí rất quan trọng

## Giải trình tự tế bào đơn

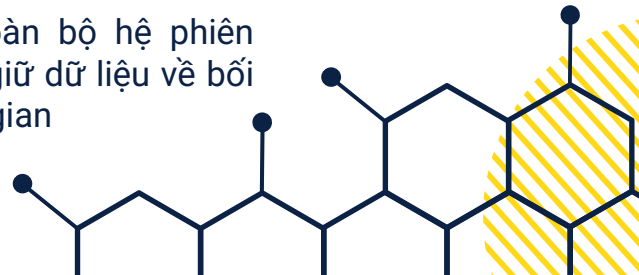


Tìm ra các loại tế bào cùng những biểu hiện gene của nó nhưng mất đi thông tin về vị trí

## Giải trình tự tế bào đơn theo cấu trúc không gian



Phân tích toàn bộ hệ phiên mã mà vẫn giữ dữ liệu về bối cảnh không gian

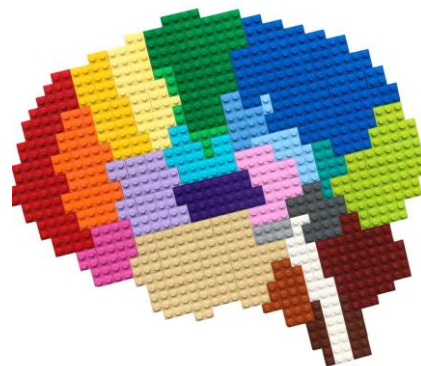


## TẠI SAO LẠI THEO CẤU TRÚC KHÔNG GIAN ?

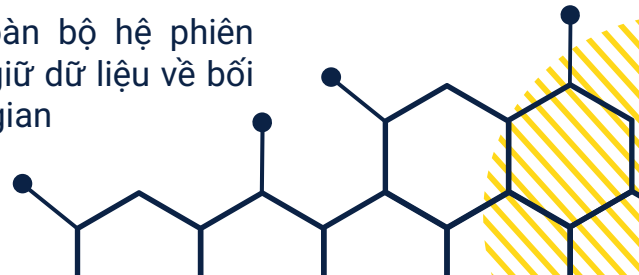
### Định vị vị trí rất quan trọng

- Bản đồ biểu hiện gene được xây dựng dựa trên việc lồng ghép hình ảnh H&E chụp lát cắt mô với bộ phiên mã của từng vị trí trên mô
- Không cần phân tách mô để phân tích hệ phiên mã
- Giữ lại tổ chức, phân bố của môi trường mô và tế bào
- Nhận diện loại tế bào trong bối cảnh mô gồm nhiều thành phần

### Giải trình tự tế bào đơn theo cấu trúc không gian

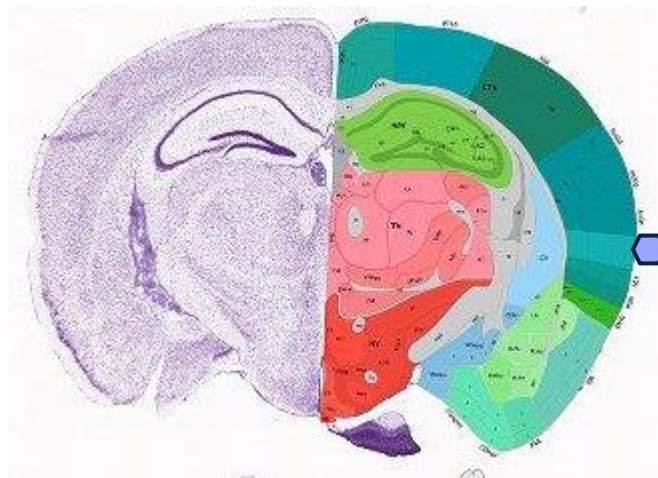
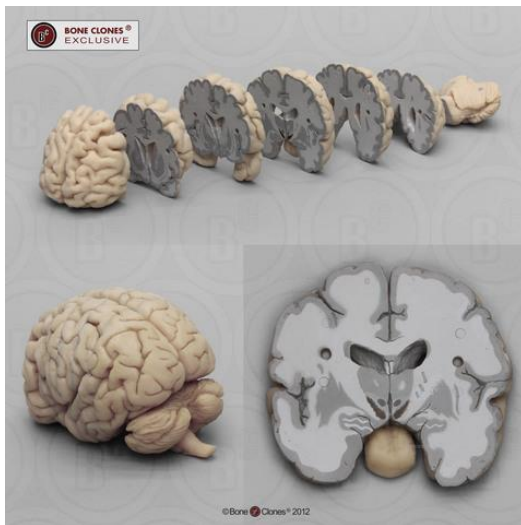


Phân tích toàn bộ hệ phiên mã mà vẫn giữ dữ liệu về bối cảnh không gian

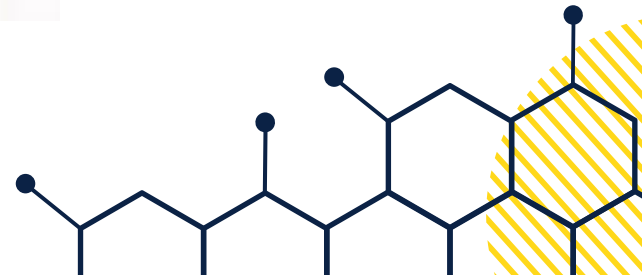


TẠI SAO LẠI THEO CẤU TRÚC KHÔNG GIAN ?

Định vị vị trí rất quan trọng

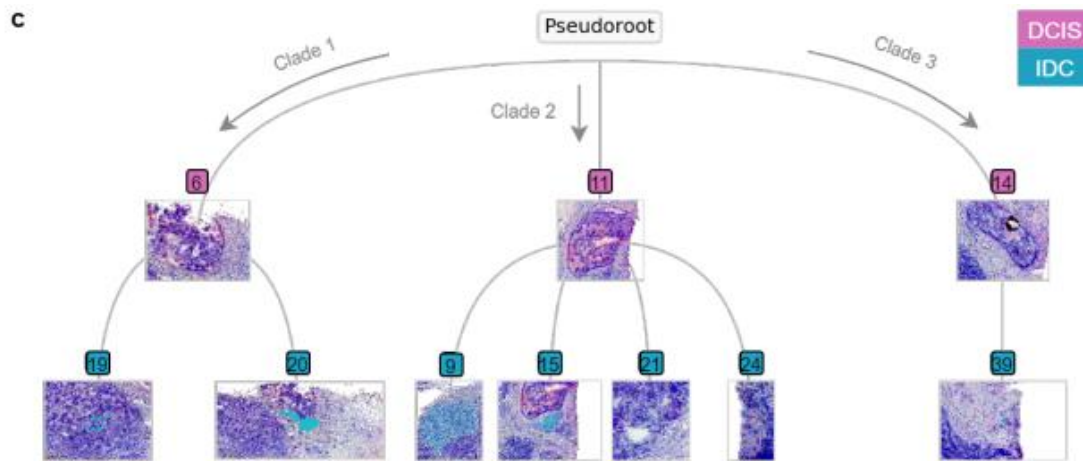
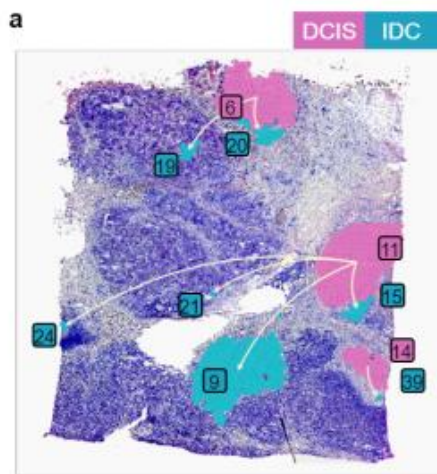


LẮT CẮT ĐỨNG  
NGANG CỦA NÃO  
CHUỘT



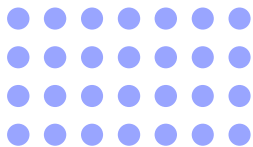
TẠI SAO LẠI THEO CẤU TRÚC KHÔNG GIAN ?

Định vị vị trí rất quan trọng



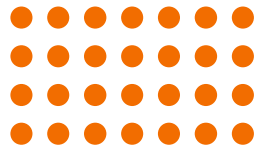
Minh họa quỹ đạo của vùng Ung thư biểu mô tuyến vú thể ống tuyến không xâm nhập (DCIS) và Ung thư biểu mô ống dẫn sữa xâm lấn (IDC)



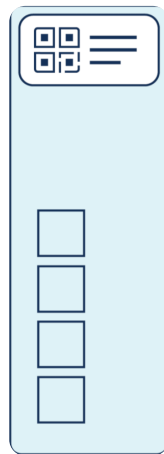


# CÔNG NGHỆ VISIUM

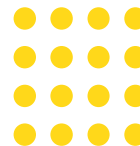
Công nghệ phân tích hệ phiên mã theo cấu trúc không gian của 10X Genomics



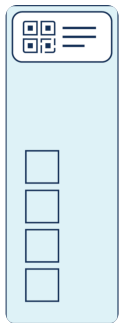
# WORK FLOW



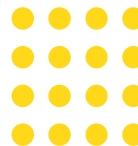
Chuẩn bị mẫu



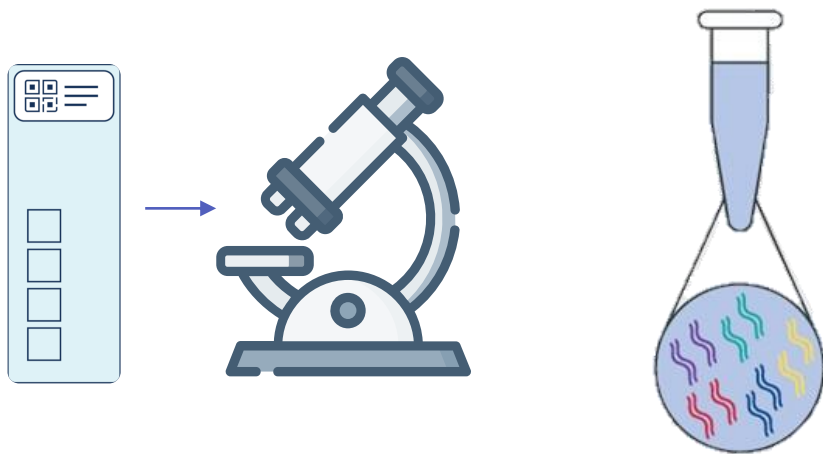
# WORK FLOW



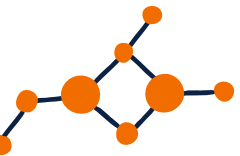
Cố định, nhuộm &  
chụp ảnh



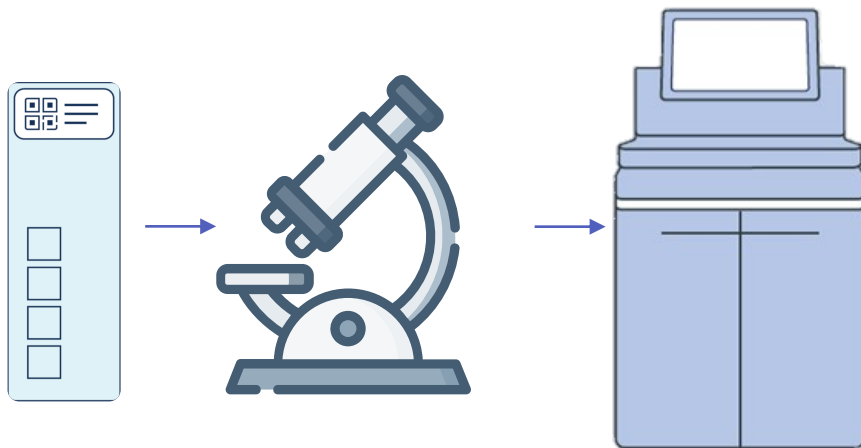
# WORK FLOW



Gắn thẻ & Chuẩn bị  
thư viện

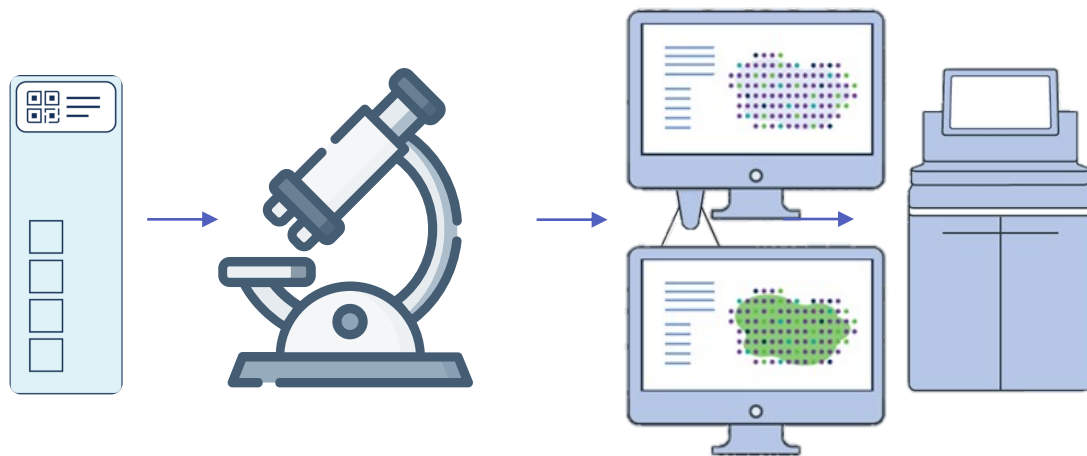


# WORK FLOW

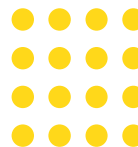
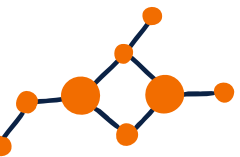


Giải trình tự

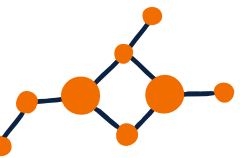
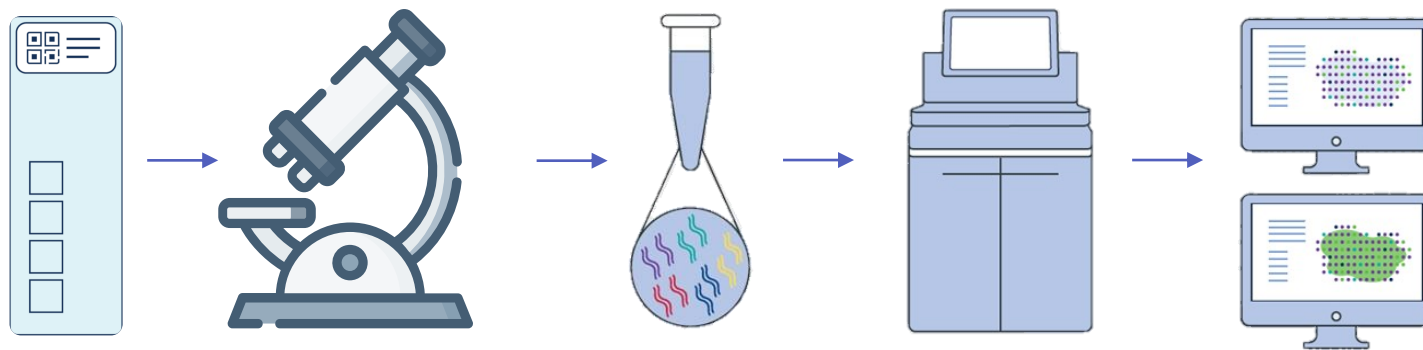
# WORK FLOW



Phân tích và hiển thị  
dữ liệu



# WORK FLOW

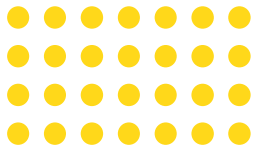
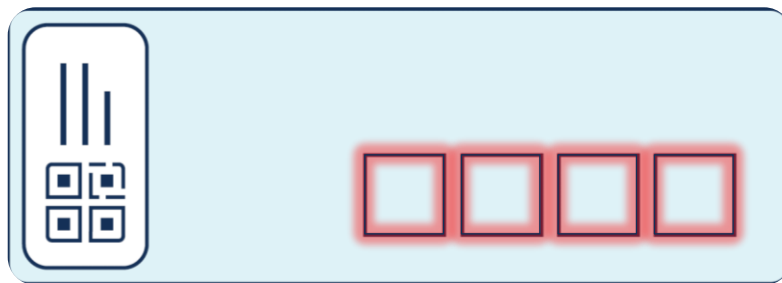


# CHUẨN BỊ MẪU

## VÙNG BẮT

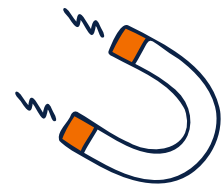
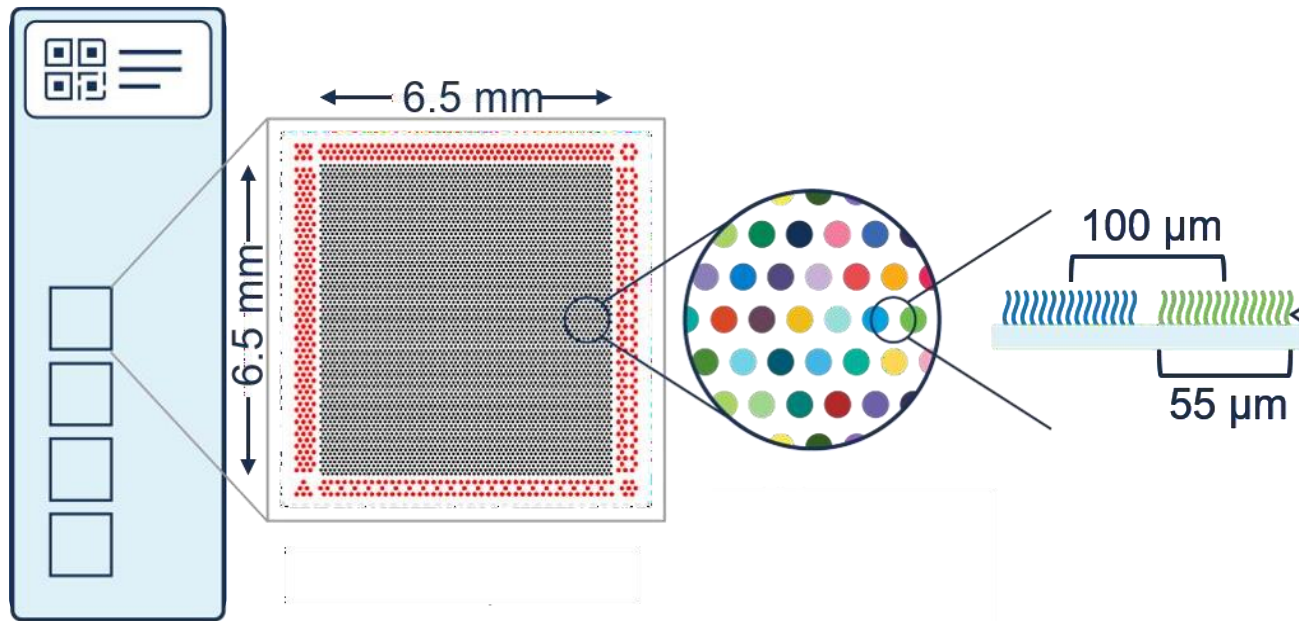


Vùng bắt (Capture Areas) trên mỗi tấm kính Visium (tấm kính biểu hiện gene theo cấu trúc không gian)

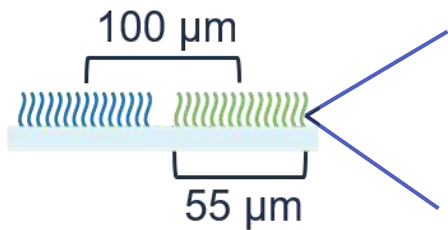




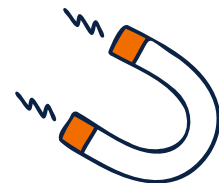
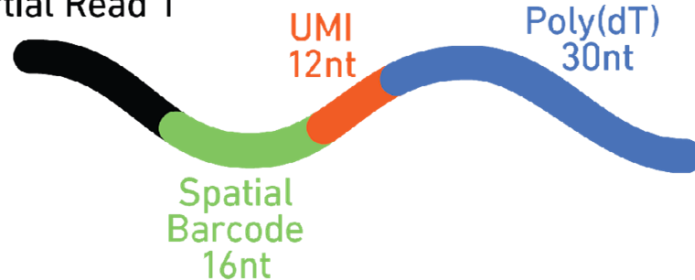
# CHUẨN BỊ MẪU



# CHUẨN BỊ MẪU



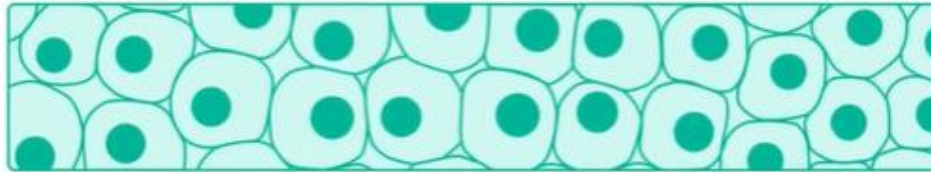
Partial Read 1



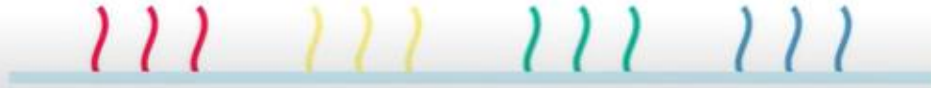
# CHUẨN BỊ MẪU

10x Visium Technology for Spatial Transcriptomics

Tissue

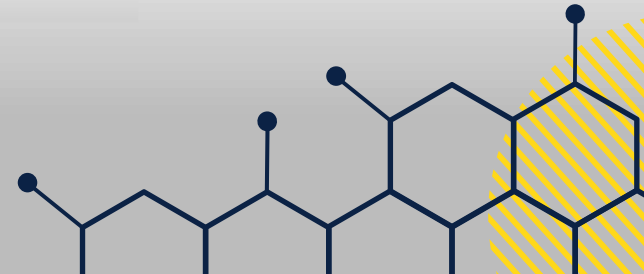
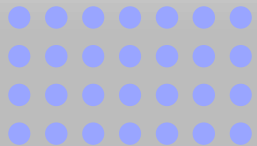


Slide



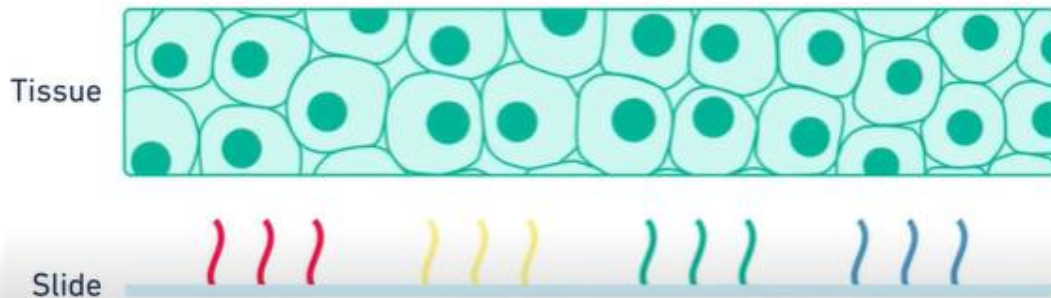
Mô

Mô được đặt trên mỗi  
vùng bắt của tấm kính  
Visium



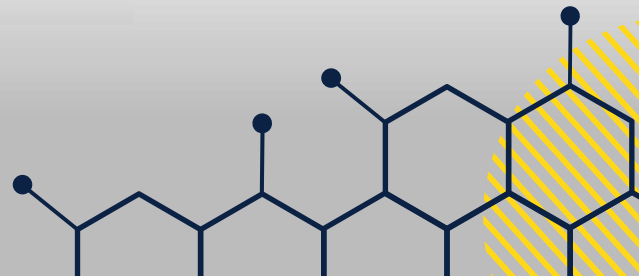
# CỐ ĐỊNH, NHUỘM VÀ CHỤP ẢNH

10x Visium Technology for Spatial Transcriptomics



Mô

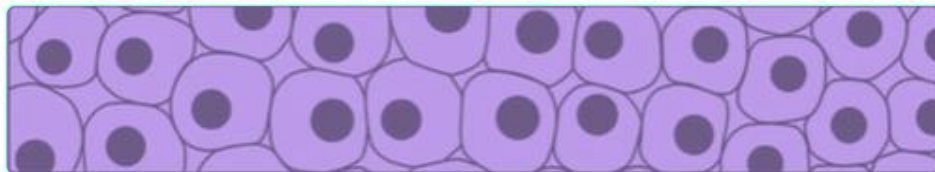
Sau đó, mô sẽ được cố định bởi methanol



# CỐ ĐỊNH, NHUỘM VÀ CHỤP ẢNH

10x Visium Technology for Spatial Transcriptomics

Tissue



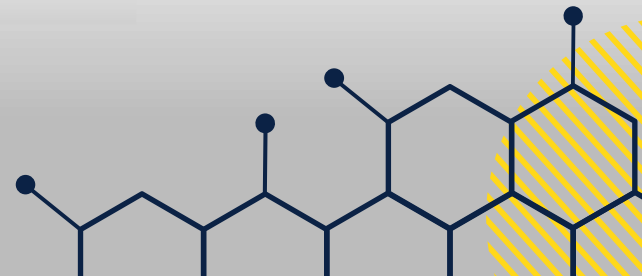
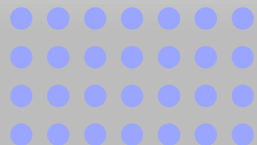
Slide



Mô đã được nhuộm

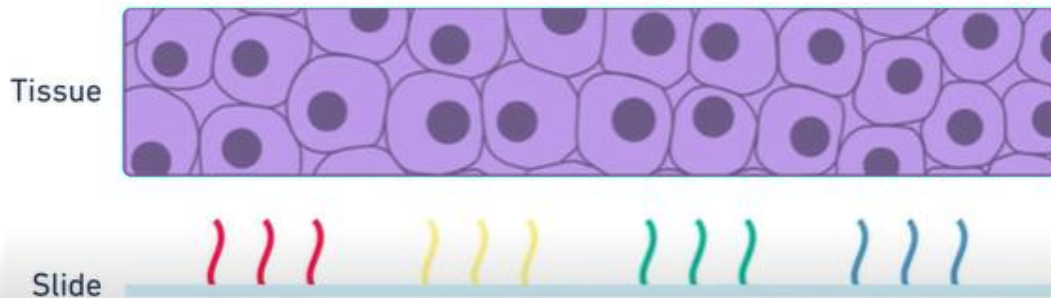
MÔ

Mô sẽ được nhuộm  
bằng hematoxylin &  
eosin



# CỐ ĐỊNH, NHUỘM VÀ CHỤP ẢNH

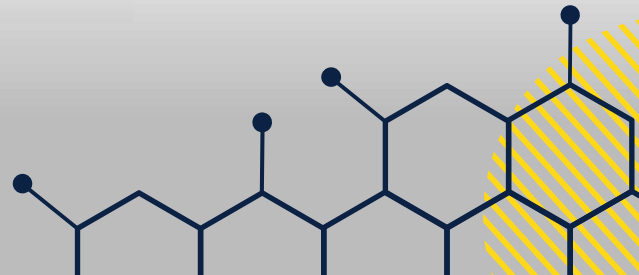
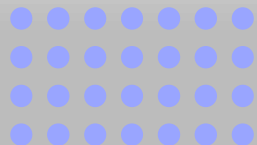
10x Visium Technology for Spatial Transcriptomics



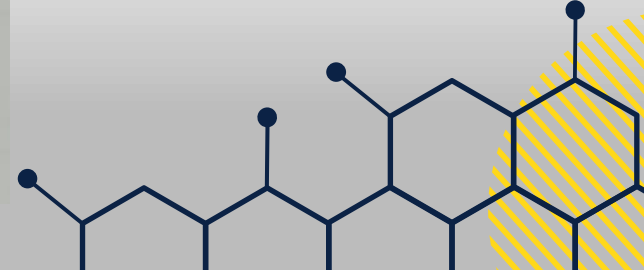
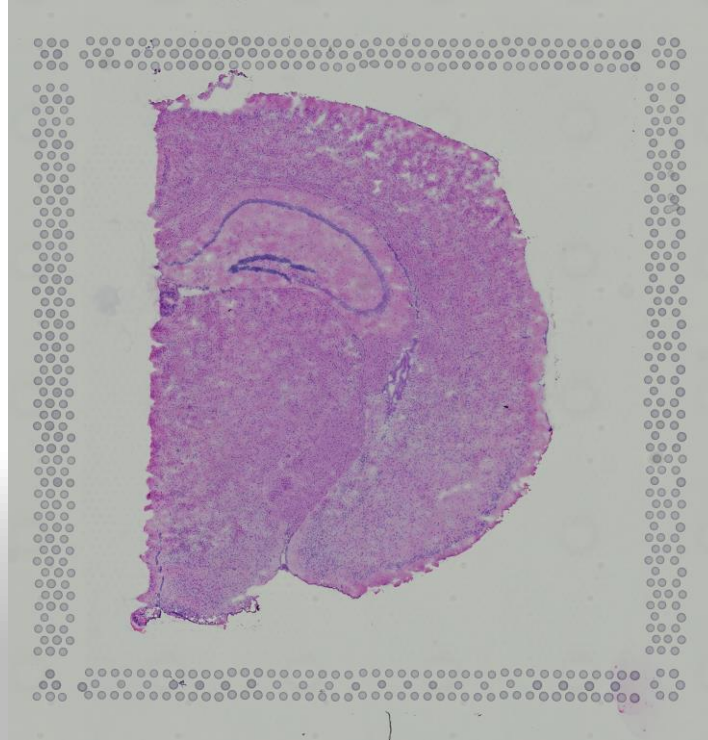
Mô đã được chụp ảnh trường sáng

Mô

Sau đó, mô sẽ được chụp ảnh với chế độ trường sáng để lấy được hình ảnh H&E của mô

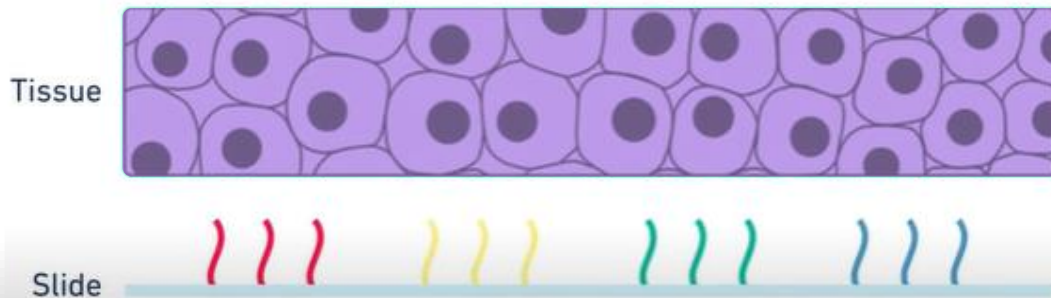


# CỐ ĐỊNH, NHUỘM VÀ CHỤP ẢNH



# CỐ ĐỊNH, NHUỘM VÀ CHỤP ẢNH

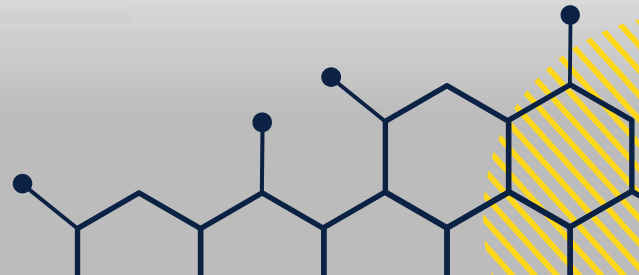
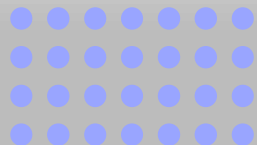
10x Visium Technology for Spatial Transcriptomics



Mô đã được chụp ảnh trường sáng

Mô

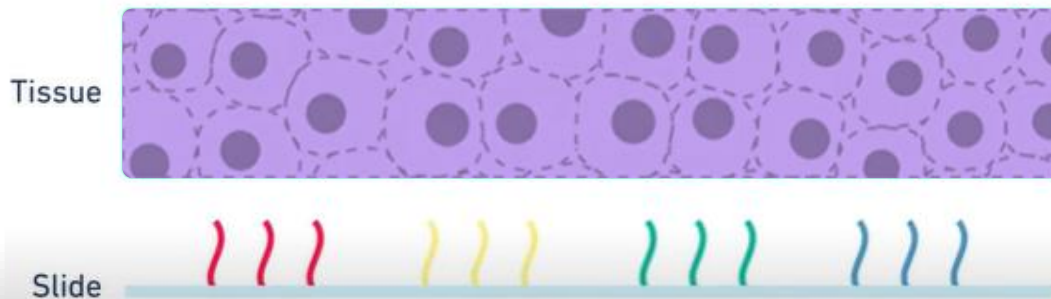
Sau đó, mô sẽ được chụp ảnh với chế độ trường sáng để lấy được hình ảnh H&E của mô





# CỐ ĐỊNH, NHUỘM VÀ CHỤP ẢNH

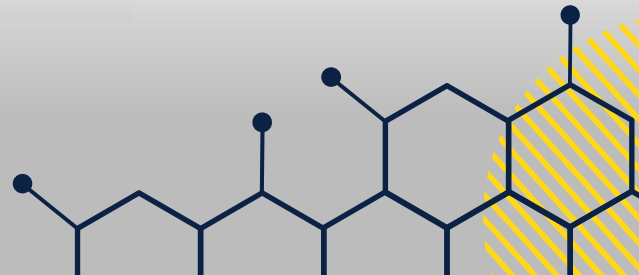
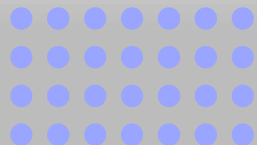
## 10x Visium Technology for Spatial Transcriptomics



Mô đã được thẩm thấu

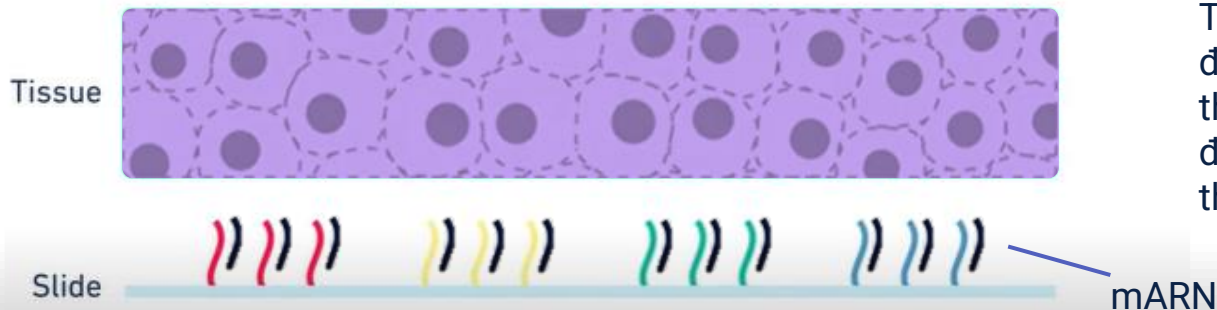
MÔ

Enzyme thẩm thấu làm cho màng tế bào có thể thẩm thấu được, từ đó mRNA có thể thoát ra được



# CỐ ĐỊNH, NHUỘM VÀ CHỤP ẢNH

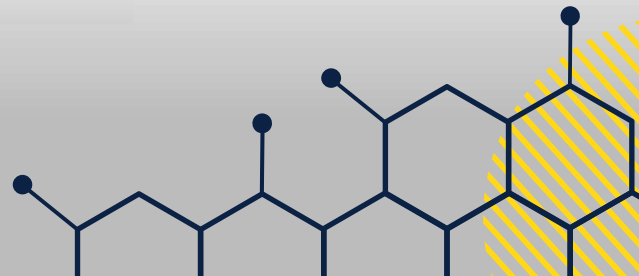
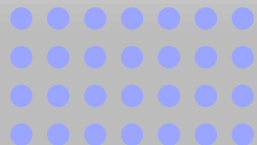
## 10x Visium Technology for Spatial Transcriptomics



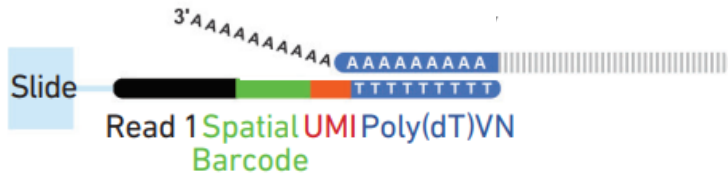
Mô đã được thẩm thấu

### TISSUE

Tissue Optimization – để xác định thời gian thẩm thấu tối ưu và đảm bảo sự tương thích của mô



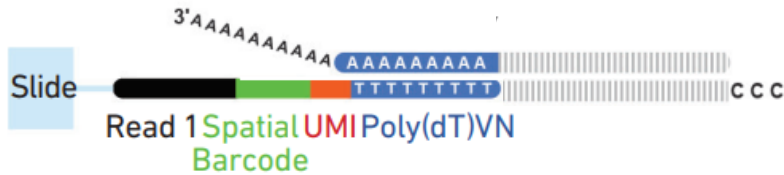
# GẮN THẺ VÀ CHUẨN BỊ THƯ VIỆN



Đuôi poly(dT) của primer  
bắt giữ đuôi poly(A) của  
mARN



# GẮN THẺ VÀ CHUẨN BỊ THƯ VIỆN



Một quá trình phiên mã ngược diễn ra với enzyme RT của 10X Genomics tạo ra mạch đơn cDNA đầu tiên và gắn trên primer. Enzyme này có đặc điểm tự động gắn thêm một số nucleotide C vào đầu 3' của cDNA sau khi đã đi hết chiều dài của mARN



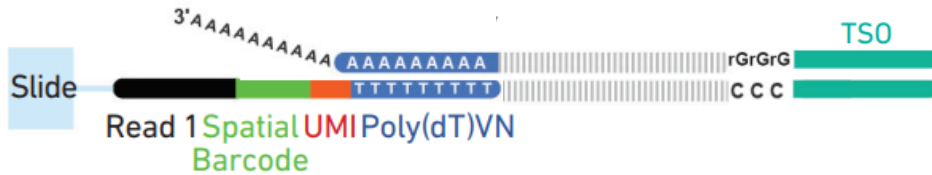
# GẮN THẺ VÀ CHUẨN BỊ THƯ VIỆN



Đuôi poly(C) của cDNA sẽ tiếp tục bắt cặp với đuôi poly(rG) của 1 trình tự gọi là TSO (Template Switch Oligo)



# GẮN THẺ VÀ CHUẨN BỊ THƯ VIỆN



Enzyme RT sẽ mở rộng phiên mã, sử dụng TSO làm khuôn mới và tiếp tục tổng hợp kéo dài cDNA



# GẮN THẺ VÀ CHUẨN BỊ THƯ VIỆN

Slide



Lúc này, mARN bị bắt giữ sẽ bị biến tính và trôi đi, để lại mạch đơn cDNA đầu tiên gắn cố định trên spot



# GẮN THẺ VÀ CHUẨN BỊ THƯ VIỆN

Một phản ứng tổng hợp mạch bổ sung diễn ra để lấy mạch đơn primer thứ hai, sử dụng TSO làm khuôn

Slide



Second Strand Primer

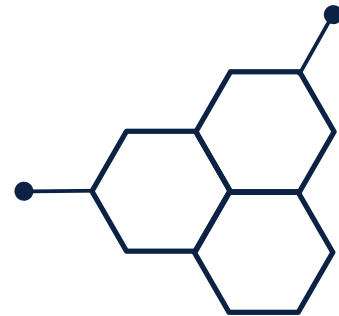




# GẮN THẺ VÀ CHUẨN BỊ THƯ VIỆN

Từ đó tổng hợp mạch đơn  
thứ hai

Slide



# GẮN THẺ VÀ CHUẨN BỊ THƯ VIỆN

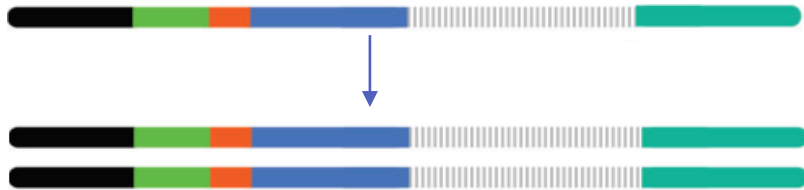
Slide



Sau khi tổng hợp, mạch đơn cDNA thứ hai mang spatial barcode bị biến tính và thu hồi, sau đó mang đi khuếch đại bằng PCR để chuẩn bị thư viện cho giải trình tự



# GẮN THẺ VÀ CHUẨN BỊ THƯ VIỆN



cDNA được định lượng số lượng cycle khuếch đại được xác định bởi qPCR. Sau đó, mạch đơn cDNA sẽ được khuếch đại bởi RT-PCR với số cycle đã được xác định.



# GẮN THẺ VÀ CHUẨN BỊ THƯ VIỆN



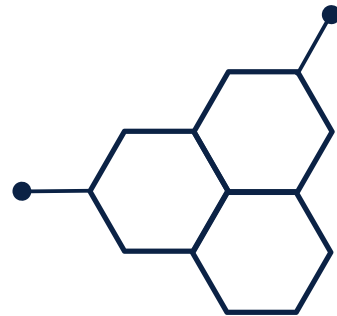
Enzyme phân mảnh cùng với SPRI cleanup cắt đi đoạn cDNA mang thông tin. Đoạn thông tin được giữ lại dài khoảng 100-400 nucleotides, tương ứng với 100-400 nucleotides ở đầu 3' của mRNA



# GẮN THẺ VÀ CHUẨN BỊ THƯ VIỆN



Gắn trình tự Read 2  
vào cDNA



# GẮN THẺ VÀ CHUẨN BỊ THƯ VIỆN



1 trình tự chỉ số duy nhất (i5) được gắn vào đầu 5' của một mạch cDNA  
1 trình tự chỉ số duy nhất (i7) được gắn vào đầu 5' của mạch cDNA còn lại



# GẮN THẺ VÀ CHUẨN BỊ THƯ VIỆN



1 trình tự chỉ số duy nhất (i5) được gắn vào đầu 5' của một mạch cDNA  
1 trình tự chỉ số duy nhất (i7) được gắn vào đầu 5' của mạch cDNA còn lại



# GẮN THẺ VÀ CHUẨN BỊ THƯ VIỆN



Thư viện cuối cùng  
bao gồm trình tự P5  
và P7 adapter để  
giải trình tự trong  
công nghệ của  
Illumina





# GẮN THẺ VÀ CHUẨN BỊ THƯ VIỆN

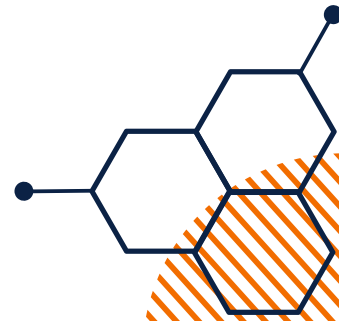


P5	i5	TruSeq Read 1	Spatial BC	UMI	Poly (dT)VN	Insert sequence	TruSeq Read 2	i7	P7
29	10	33	16	12	30	100 ~ 400	34	10	24

300 ~ 600 bp

CẤU TRÚC THƯ VIỆN  
SPATIAL 3' GENE  
EXPRESSION

Sử dụng 2 chỉ số



# GIẢI TRÌNH TỰ

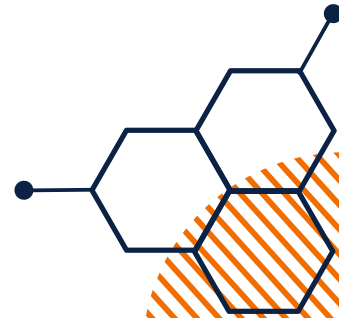


P5	i5	TruSeq Read 1	Spatial BC	UMI	Poly (dT)VN	Insert sequence	TruSeq Read 2	i7	P7
29	10	33	16	12	30	100 ~ 400	34	10	24

300 ~ 600 bp

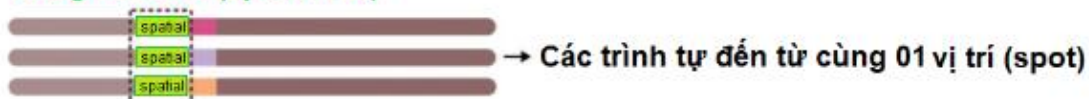
Mỗi trình tự có 4 thông tin quan trọng

- Chỉ số
- Barcode không gian đại diện cho spot
- Trình tự mang thông tin đầu 3' của mRNA
- UMI



# GIẢI TRÌNH TỰ

## Cùng barcode (spatial BC)



## Cùng barcode, transcript giống nhau, khác UMI



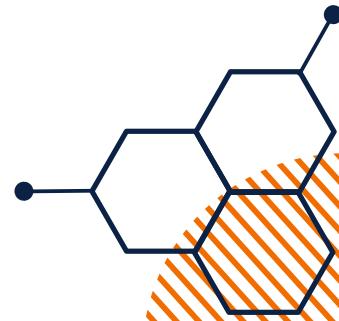
## Cùng barcode, transcript giống nhau, cùng UMI



Hình 6. Cách thức sử dụng Spatial Barcode và UMI để phân tích dữ liệu

Mỗi trình tự có 4 thông tin quan trọng

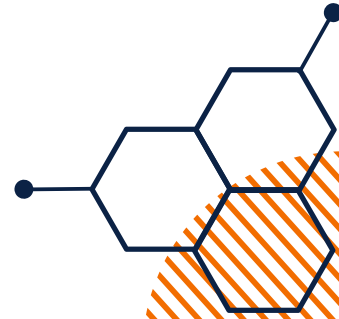
- Chỉ số
- Barcode không gian đại diện cho spot
- Trình tự mang thông tin đầu 3' của mRNA
- UMI



# GIẢI TRÌNH TỰ



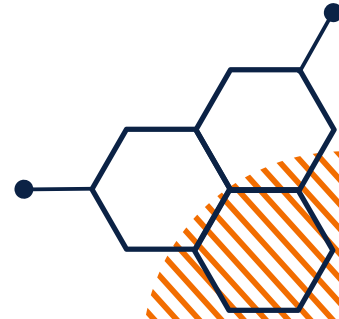
Sau khi có thư viện 10x, ta đưa vào máy Illumina để giải trình tự. Máy Illumina tạo ra file .bcl dưới định dạng binary và con người không đọc được. File .bcl là input đưa vào spaceranger mkfastq



# GIẢI TRÌNH TỰ



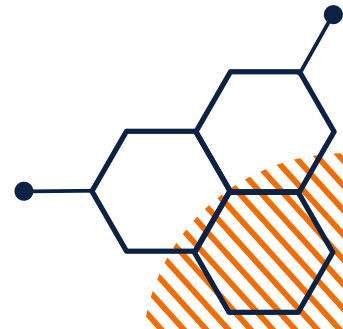
Spaceranger mkfastq sẽ giải ghép trên flow cell đó, được 2 library là library 1 và library 2, với library 2 đọc được trên 2 lanes 1 và 2 và tạo ra các file fastq.



# GIẢI TRÌNH TỰ



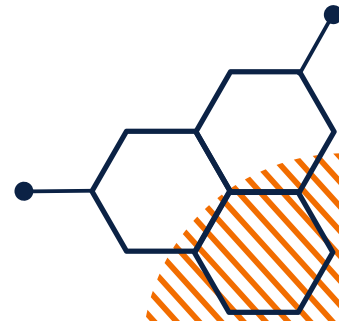
Sau khi chạy spaceranger  
mkfastq, ta chạy mỗi pipeline  
spaceranger trên một library



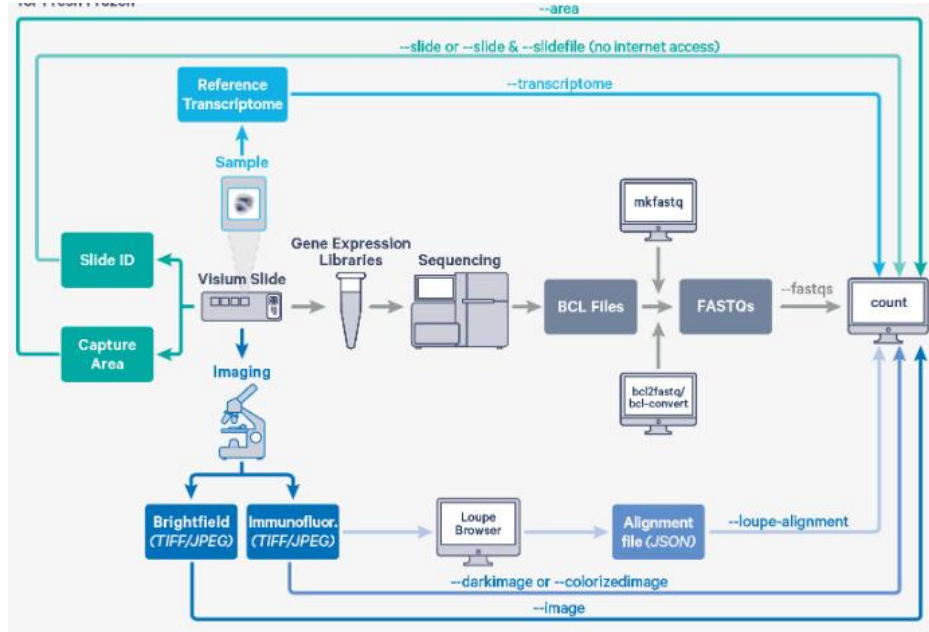
# GIẢI TRÌNH TỰ



Output của spaceranger mkfastq là các file fastq, file fastq này sẽ trở thành 1 trong những input cho spaceranger count pipeline.



## GIẢI TRÌNH TỰ



Spaceranger count được chạy trên mỗi vùng bắt trên lăng kính Visium. Input cần thiết bao gồm:

- File fastq cho một capture area
- Hình ảnh tương ứng với capture area đó ở định dạng TIFF, QTIFF hoặc JPEG
- File GPR về bố cục của slide, về cụ thể tọa độ của các spot, thể hiện bởi slide serial number và capture area arguments
- Hệ phiên mã tham chiếu cho loài mà mẫu được lấy

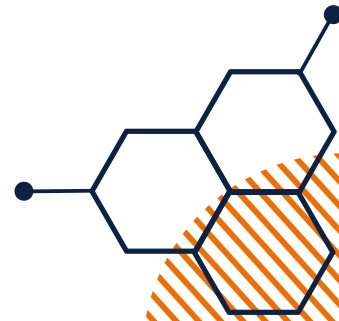




# GIẢI TRÌNH TỰ

```
V1_Adult_Mouse_Brain
├─ _cmdline
├─ _filelist
├─ _finalstate
├─ _invocation
├─ _jobmode
├─ _log
├─ _mrosource
├─ outs
├─ _perf
├─ _sitecheck
├─ SPATIAL_RNA_COUNTER_CS
├─ _tags
├─ _timestamp
├─ _uuid
├─ V1_Adult_Mouse_Brain.mri.tgz
├─ _vdrkill
└─ _versions
```

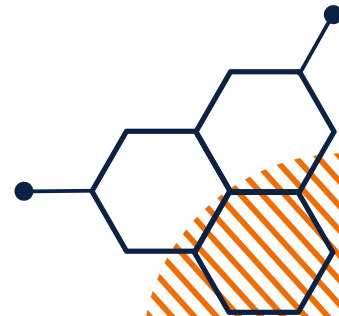
Sau khi chạy spaceranger count pipeline, tại thư mục làm việc hiện tại sẽ có thêm folder output, bao gồm metadata và các file output được tạo ra từ spaceranger count pipeline



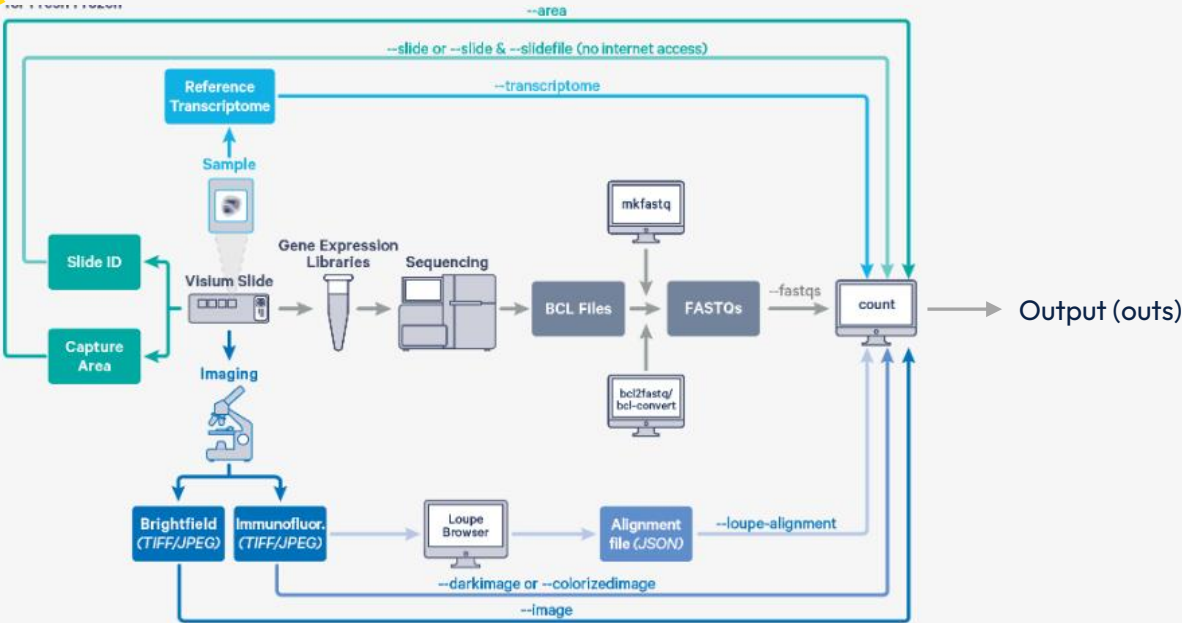
# GIẢI TRÌNH TỰ

```
V1_Adult_Mouse_Brain
├── _cmdline
├── _filelist
├── _finalstate
├── _invocation
├── _jobmode
├── _log
├── _mnosource
├── outs
├── _perr
├── _sitecheck
├── SPATIAL_RNA_COUNTER_CS
├── _tags
├── _timestamp
├── _uuid
├── V1_Adult_Mouse_Brain.mri.tgz
├── _vdrkill
└── _versions
```

CÁC FILE OUTPUT



## GIẢI TRÌNH TỰ



Output	Description
<a href="#">web_summary.html</a>	Tóm tắt dữ liệu với những hình ảnh, chỉ số và biểu đồ có thể sử dụng cho đánh giá chất lượng, lỗi và cảnh báo liên quan đến chất lượng dữ liệu
<a href="#">metrics_summary.csv</a>	Chỉ số cho việc đánh giá chất lượng
<a href="#">cloupe.cloupe</a>	File Loupe Browser cho việc trực quan hóa và phân tích dữ liệu
<a href="#">analysis/</a>	
<a href="#">spatial/spatial_enrichment.csv</a>	Kết quả phân tích thứ cấp (clustering, DGE, Moran's I)
<a href="#">spatial/</a>	Bao gồm các file hình ảnh Visium và file miêu tả vị trí spot barcode
<a href="#">molecule_info.h5</a>	Thông tin mức độ phân tử sử dụng trong các pipeline bổ sung (spaceranger aggr, spaceranger targeted-compare)
<a href="#">Barcoded BAM (optional)</a>	
<a href="#">possorted_genome_bam.bam</a>	
<a href="#">possorted_genome_bam.bai</a>	Read alignment files
<a href="#">possorted_genome_bam.csi</a>	
Filtered GEX Matrix	
<a href="#">MEX: filtered_feature_bc_matrix/</a>	Ma trận barcode-feature chỉ bao gồm các barcode liên kết với mô, sau đó được sử dụng để phân tích bằng R hoặc
<a href="#">HDF5: filtered_feature_bc_matrix.h5</a>	
Unfiltered GEX Matrix	
<a href="#">MEX: raw_feature_bc_matrix/</a>	Ma trận barcode-feature bao gồm cả barcode liên kết với mô và barcode liên kết với background, sau đó được sử dụng để phân tích bằng R hoặc Python
<a href="#">HDF5: raw_feature_bc_matrix.h5</a>	

# GIẢI TRÌNH TỰ

Output (outs)

Output	Description
<a href="#">web_summary.html</a>	Tóm tắt dữ liệu với những hình ảnh, chỉ số và biểu đồ có thể sử dụng cho đánh giá chất lượng, lỗi và cảnh báo liên quan đến chất lượng dữ liệu
<a href="#">metrics_summary.csv</a>	Chỉ số cho việc đánh giá chất lượng
<a href="#">cloupe.cloupe</a>	File Loupe Browser cho việc trực quan hóa và phân tích dữ liệu
<a href="#">analysis/</a>	Kết quả phân tích thứ cấp (clustering, DGE, Moran's I)
<a href="#">spatial/spatial_enrichment.csv</a>	
<a href="#">spatial/</a>	Bao gồm các file hình ảnh Visium và file miêu tả vị trí spot barcode
<a href="#">molecule_info.h5</a>	Thông tin mức độ phân tử sử dụng trong các pipeline bổ sung (spaceranger agg, spaceranger targeted-compare)
<a href="#">Barcoded BAM (optional)</a>	Read alignment files
<a href="#">possorted_genome_bam.bam</a>	
<a href="#">possorted_genome_bam.bai</a>	
<a href="#">possorted_genome_bam.csi</a>	
Filtered GEX Matrix	Ma trận barcode-feature chỉ bao gồm các barcode liên kết với mô, sau đó được sử dụng để phân tích bằng R hoặc
<a href="#">MEX: filtered_feature_bc_matrix/</a>	
<a href="#">HD5F: filtered_feature_bc_matrix.h5</a>	
Unfiltered GEX Matrix	Ma trận barcode-feature bao gồm cả barcode liên kết với mô và barcode liên kết với background, sau đó được sử dụng để phân tích bằng R hoặc Python
<a href="#">MEX: raw_feature_bc_matrix/</a>	
<a href="#">HD5F: raw_feature_bc_matrix.h5</a>	



Output	Description
<a href="#">web_summary.html</a>	Tóm tắt dữ liệu với những hình ảnh, chỉ số và biểu đồ có thể sử dụng cho đánh giá chất lượng, lỗi và cảnh báo liên quan đến chất lượng dữ liệu
<a href="#">metrics_summary.csv</a>	Chỉ số cho việc đánh giá chất lượng
<a href="#">cloupe.cloupe</a>	File Loupe Browser cho việc trực quan hóa và phân tích dữ liệu
<a href="#">analysis/</a>	Kết quả phân tích thứ cấp (clustering, DGE, Moran's I)
<a href="#">spatial/spatial_enrichment.csv</a>	
<a href="#">spatial/</a>	Bao gồm các file hình ảnh Visium và file miêu tả vị trí spot barcode
<a href="#">molecule_info.h5</a>	Thông tin mức độ phân tử sử dụng trong các pipeline bổ sung (spaceranger agg, spaceranger targeted-compare)

# GIẢI TRÌNH TỰ

Output (outs)

Output	Description
<a href="#">web_summary.html</a>	Tóm tắt dữ liệu với những hình ảnh, chỉ số và biểu đồ có thể sử dụng cho đánh giá chất lượng, lỗi và cảnh báo liên quan đến chất lượng dữ liệu
<a href="#">metrics_summary.csv</a>	Chỉ số cho việc đánh giá chất lượng
<a href="#">cloupe.cloupe</a>	File Loupe Browser cho việc trực quan hóa và phân tích dữ liệu
<a href="#">analysis/</a>	
<a href="#">spatial/spatial_enrichment.csv</a>	Kết quả phân tích thứ cấp (clustering, DGE, Moran's I)
<a href="#">spatial/</a>	Bao gồm các file hình ảnh Visium và file miêu tả vị trí spot barcode
<a href="#">molecule_info.h5</a>	Thông tin mức độ phân tử sử dụng trong các pipeline bổ sung (spaceranger aggr, spaceranger targeted-compare)
<a href="#">Barcoded BAM (optional)</a>	
<a href="#">possorted_genome_bam.bam</a>	Read alignment files
<a href="#">possorted_genome_bam.bai</a>	
<a href="#">possorted_genome_bam.csi</a>	
<a href="#">possorted_genome_bam.csi</a>	
Filtered GEX Matrix	
<a href="#">MEX: filtered_feature_bc_matrix/</a>	Ma trận barcode-feature chỉ bao gồm các barcode liên kết với mô, sau đó được sử dụng để phân tích bằng R hoặc
<a href="#">HDSF: filtered_feature_bc_matrix.h5</a>	
Unfiltered GEX Matrix	
<a href="#">MEX: raw_feature_bc_matrix/</a>	Ma trận barcode-feature bao gồm cả barcode liên kết với mô và barcode liên kết với background, sau đó được sử dụng để phân tích bằng R hoặc Python
<a href="#">HDSF: raw_feature_bc_matrix.h5</a>	



<u>Barcoded BAM (optional)</u>	Read alignment files
<a href="#">possorted_genome_bam.bam</a>	
<a href="#">possorted_genome_bam.bai</a>	
<a href="#">possorted_genome_bam.csi</a>	

# GIẢI TRÌNH TỰ

Output (outs)

Output	Description
<a href="#">web_summary.html</a>	Tóm tắt dữ liệu với những hình ảnh, chỉ số và biểu đồ có thể sử dụng cho đánh giá chất lượng, lỗi và cảnh báo liên quan đến chất lượng dữ liệu
<a href="#">metrics_summary.csv</a>	Chỉ số cho việc đánh giá chất lượng
<a href="#">cloupe.cloupe</a>	File Loupe Browser cho việc trực quan hóa và phân tích dữ liệu
<a href="#">analysis/</a>	Kết quả phân tích thứ cấp (clustering, DGE, Moran's I)
<a href="#">spatial/spatial_enrichment.csv</a>	
<a href="#">spatial/</a>	Bao gồm các file hình ảnh Visium và file miêu tả vị trí spot barcode
<a href="#">molecule_info.h5</a>	Thông tin mức độ phân tử sử dụng trong các pipeline bổ sung (spaceranger agg, spaceranger targeted-compare)
<a href="#">Barcoded BAM (optional)</a>	
<a href="#">possorted_genome_bam.bam</a>	Read alignment files
<a href="#">possorted_genome_bam.bai</a>	
<a href="#">possorted_genome_bam.csi</a>	
Filtered GEX Matrix	
<a href="#">MEX: filtered_feature_bc_matrix/</a>	Ma trận barcode-feature chỉ bao gồm các barcode liên kết với mô, sau đó được sử dụng để phân tích bằng R hoặc Python
<a href="#">HD5F: filtered_feature_bc_matrix.h5</a>	
Unfiltered GEX Matrix	
<a href="#">MEX: raw_feature_bc_matrix/</a>	Ma trận barcode-feature bao gồm cả barcode liên kết với mô và barcode liên kết với background, sau đó được sử dụng để phân tích bằng R hoặc Python
<a href="#">HD5F: raw_feature_bc_matrix.h5</a>	

## Filtered GEX Matrix

MEX:

filtered\_feature\_bc\_matrix/

HD5F:

filtered\_feature\_bc\_matrix.h5

Ma trận barcode-feature chỉ bao gồm các barcode liên kết với mô, sau đó được sử dụng để phân tích bằng R hoặc Python

## Unfiltered GEX Matrix

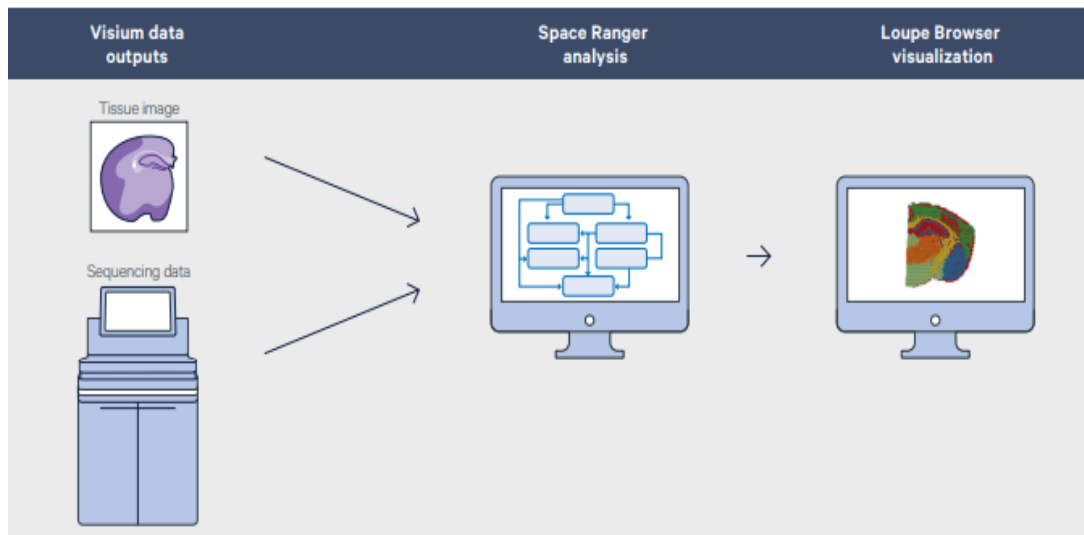
MEX: raw\_feature\_bc\_matrix/

HD5F:

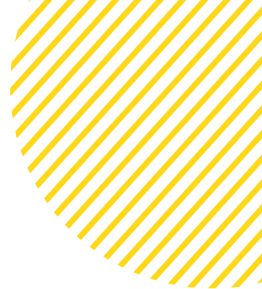
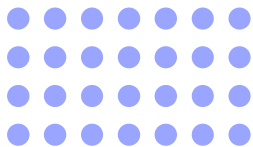
raw\_feature\_bc\_matrix.h5

Ma trận barcode-feature bao gồm cả barcode liên kết với mô và barcode liên kết với background, sau đó được sử dụng để phân tích bằng R hoặc Python

# PHÂN TÍCH VÀ HIỂN THỊ DỮ LIỆU

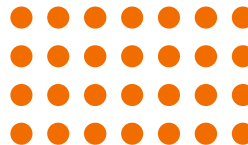


Sử dụng phần mềm Space Ranger để xử lý dữ liệu biểu hiện gene theo không gian và hiển thị kết quả với trình duyệt Loupe



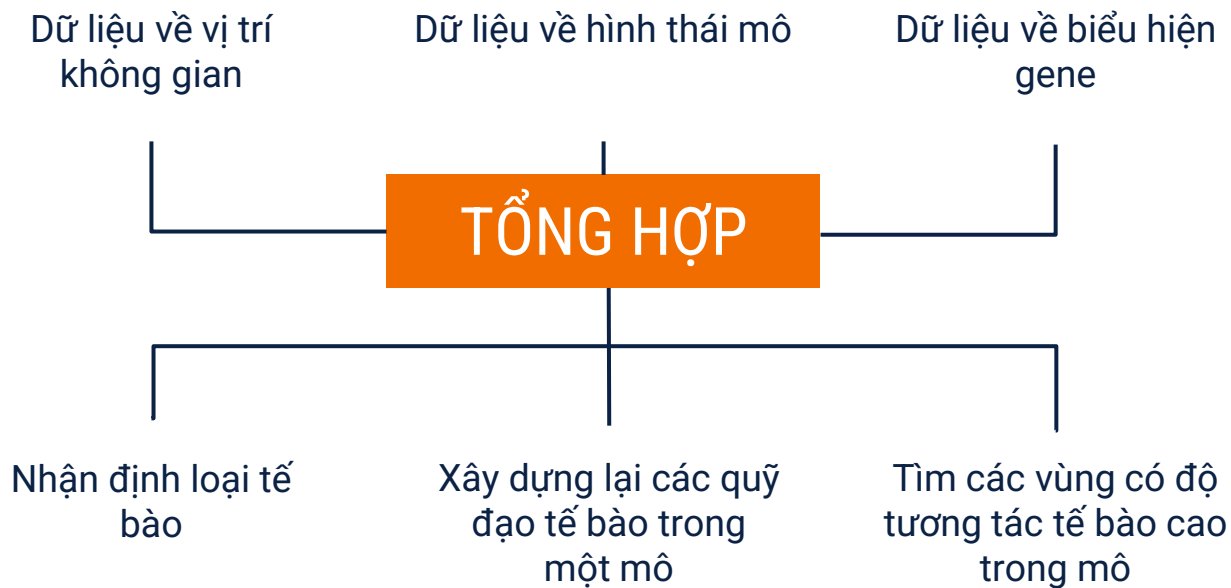
# STLEARN

Một phương pháp ứng dụng phân tích tổng hợp dữ liệu về phiên mã không gian





# STLEARN



# STLEARN WORKFLOW

## CHUẨN HÓA DỮ LIỆU

Bằng cách điều chỉnh hình thái

01

02

## QUỸ ĐẠO KHÔNG GIAN

Tìm mối quan hệ giữa các loại tế bào

03

04

## PHÂN CỤM THEO KHÔNG GIAN

Tìm các cụm được xác định theo không gian và hệ phiên mã

## TƯƠNG TÁC TẾ BÀO

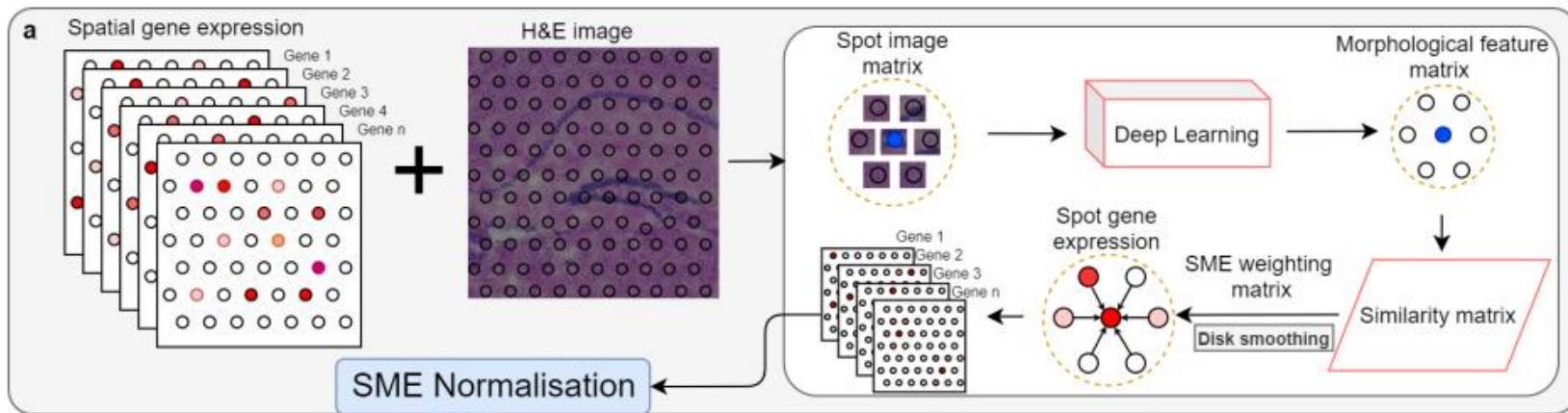
Tìm ra các điểm có tương tác nổi bật trên toàn bộ mô

# STLEARN WORKFLOW

- Chuẩn hóa SME sử dụng vị trí không gian và khoảng cách hình thái để chuẩn hóa dữ liệu biểu hiện gene
- Thay đổi giá trị biểu hiện gene trong một spot dựa trên giá trị biểu hiện của gene đó trong các spot lân cận với phạm vi bán kính d

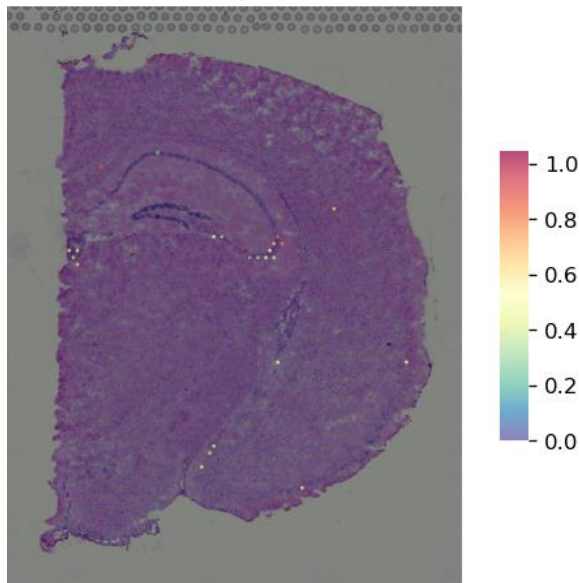
## CHUẨN HÓA DỮ LIỆU

01



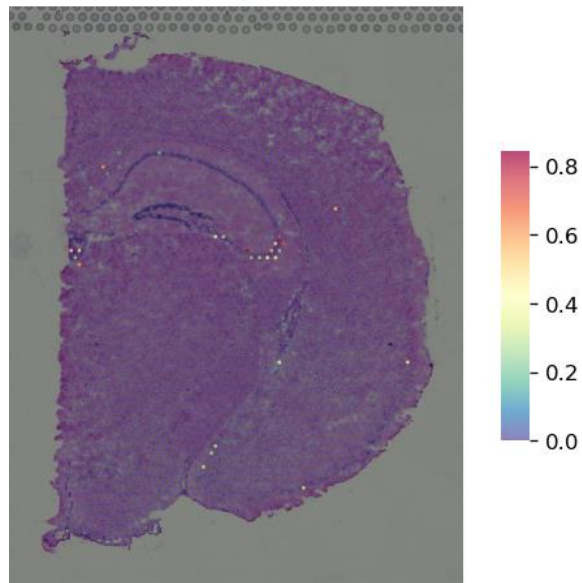
# STLEARN WORKFLOW

Lhfp11



Chuẩn hóa với thư viện scanpy

Lhfp11



Chuẩn hóa với SME



# STLEARN WORKFLOW

## CHUẨN HÓA DỮ LIỆU

Bằng cách điều chỉnh hình thái

01

02

## QUỸ ĐẠO KHÔNG GIAN

Tìm mối quan hệ giữa các loại tế bào

03

04

## PHÂN CỤM THEO KHÔNG GIAN

Tìm các cụm được xác định theo không gian và hệ phiên mã

## TƯƠNG TÁC TẾ BÀO

Tìm ra các điểm có tương tác nổi bật trên toàn bộ mô

# STLEARN WORKFLOW

02

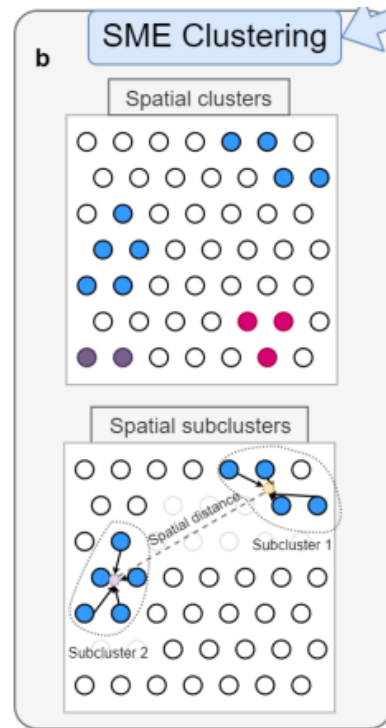
## PHÂN CỤM THEO KHÔNG GIAN

### Global Clustering

- Sử dụng ma trận đã được chuẩn hóa SME
- Từ đây có thể áp dụng K-means clustering, hoặc tạo ra sơ đồ k-Nearest-neighbor và áp dụng Louvain clustering

### Local Clustering

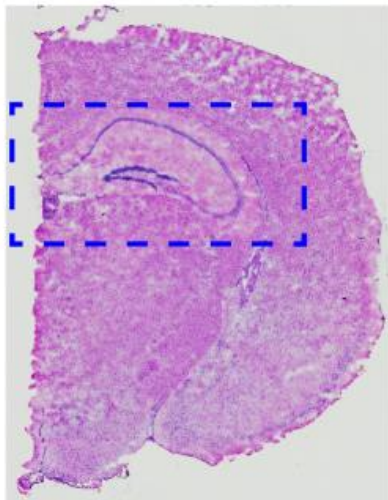
- Tại đây, thông tin về vị trí không gian được sử dụng để tìm những sub-clusters từ clusters đã được phân cụm ở bước trên mà có thể ở hai vị trí khác nhau



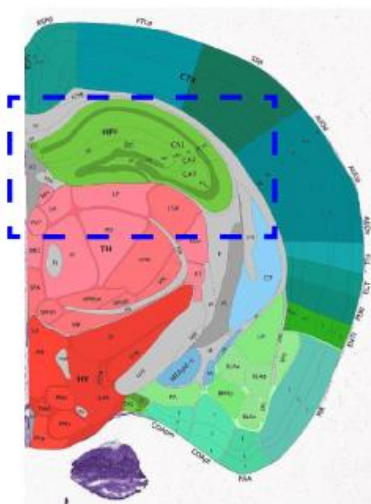
# STLEARN WORKFLOW

**a**

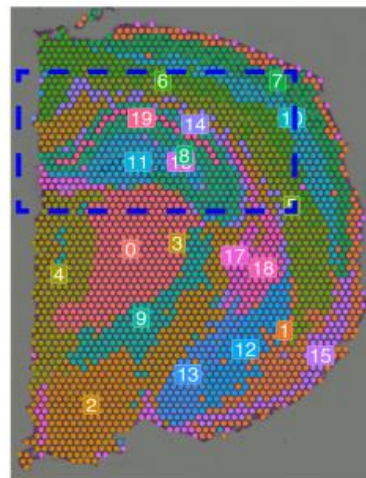
Original H&E image



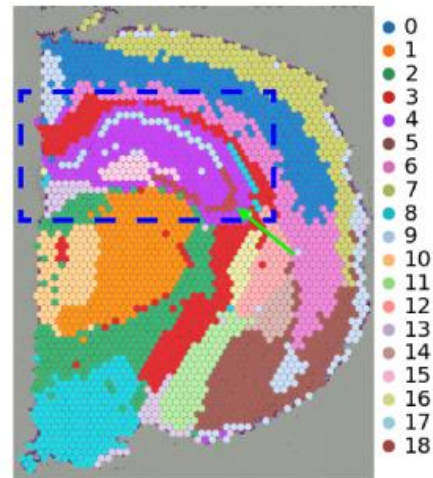
Allen mouse brain atlas



Seurat



SMEclust



# STLEARN WORKFLOW

## CHUẨN HÓA DỮ LIỆU

Bằng cách điều chỉnh hình thái

01

02

## PHÂN CỤM THEO KHÔNG GIAN

Tìm các cụm được xác định theo không gian và hệ phiên mã

03

## QUỸ ĐẠO KHÔNG GIAN

Tìm mối quan hệ giữa các loại tế bào

04

## TƯƠNG TÁC TẾ BÀO

Tìm ra các điểm có tương tác nổi bật trên toàn bộ mô



# STLEARN WORKFLOW

## Thuật toán phân tích quỹ đạo Pseudo-Space-Time (PST)

Cho hai tập spots  $U$  và  $V$  với trung tâm lần lượt là  $u$  và  $v$ , tính khoảng cách pseudo-space-time kết hợp bởi khoảng cách vật lí với khoảng cách biểu hiện gene, với  $\omega$  phản ánh sự cân bằng giữa hai thành phần trên

- Pseudo-time-distance:

$$d_{DPT(u,v)} = |DPT_u - DPT_v|$$

- Spatial distance:

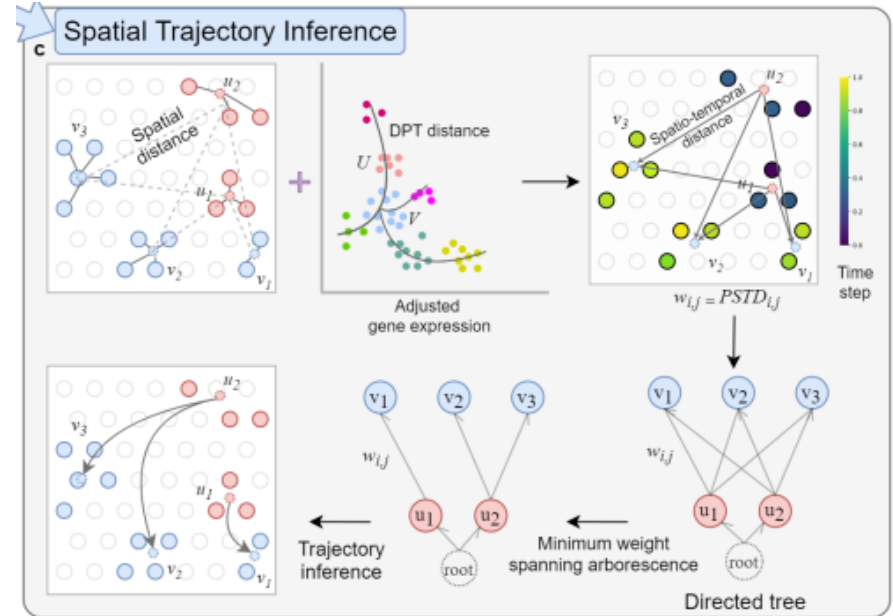
$$d_{S(u,v)} = \sqrt{(v_1 - u_1)^2 + (u_2 - v_2)^2}$$

- Pseudo-space-time distance:

$$PSTD_{u,v} = d_{DPT} \cdot \omega + d_S \cdot (1 - \omega), \omega \in [0; 1]$$

## QUỸ ĐẠO KHÔNG GIAN

03



# STLEARN WORKFLOW

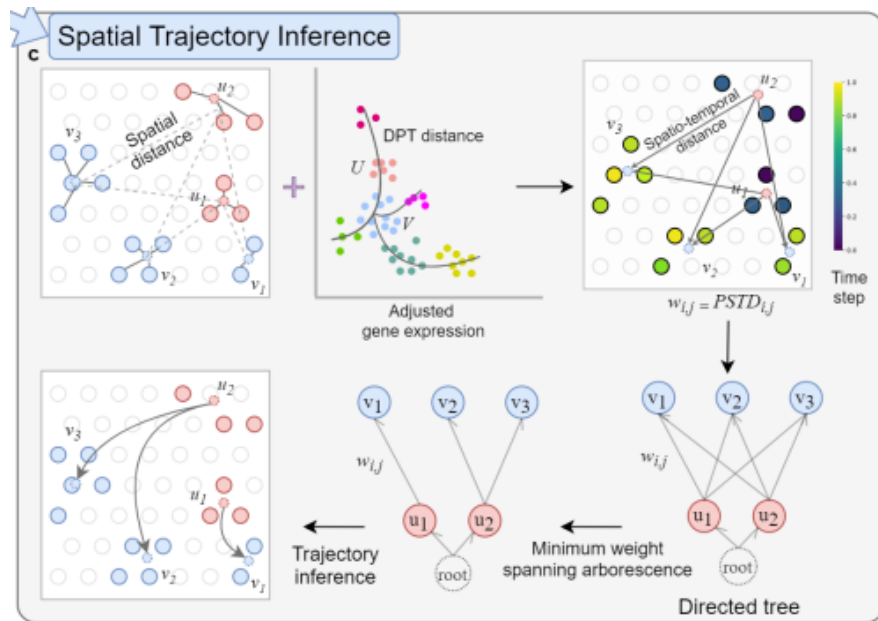
Thuật toán phân tích quỹ đạo  
Pseudo-Space-Time (PST)

QUỸ ĐẠO KHÔNG GIAN

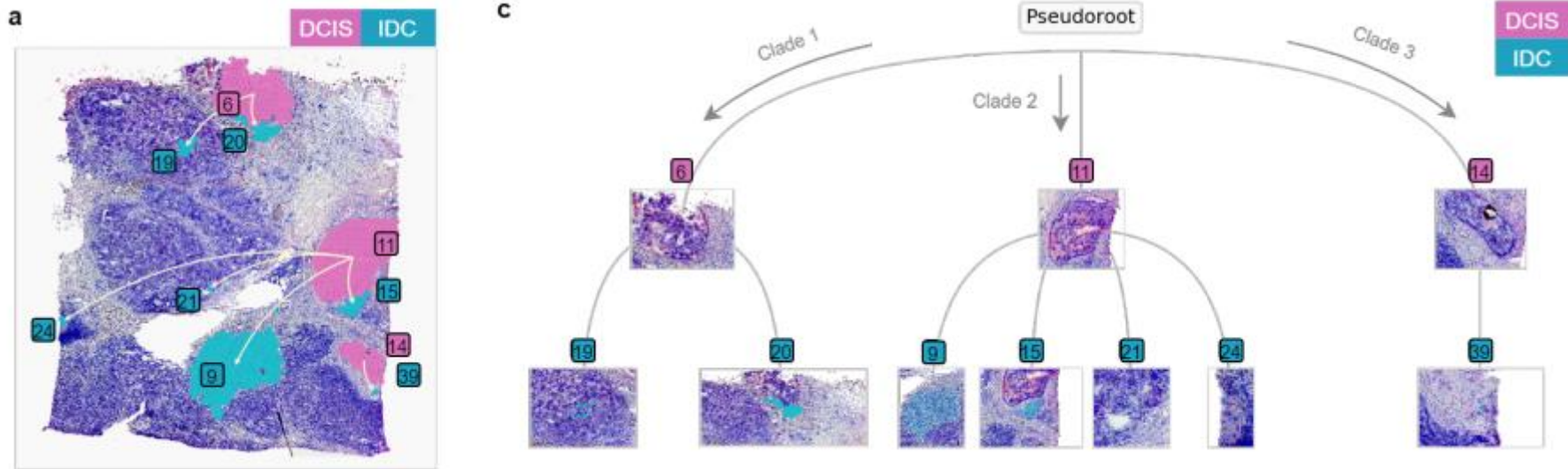
03

Với ma trận PSTD, ta xây dựng một đồ thị có hướng, sau đó tối ưu đồ thị để tìm ra đường ngắn nhất, gốc và nhánh (quỹ đạo cần tìm).

Các đường thẳng minh họa mối quan hệ về mặt phiên mã không gian giữa 1 cặp cụm. Hướng của quỹ đạo được xác định bởi khoảng cách DPT. Nếu khoảng cách DPT  $> 0$  thì là hướng thuận, nếu  $< 0$  thì là hướng nghịch.



# STLEARN WORKFLOW



Minh họa quỹ đạo của vùng Ung thư biểu mô tuyến vú thể ống tuyến không xâm nhập (DCIS) và Ung thư biểu mô ống dẫn sữa xâm lấn (IDC)

# STLEARN WORKFLOW

## CHUẨN HÓA DỮ LIỆU

Bằng cách điều chỉnh hình thái

01

02

## PHÂN CỤM THEO KHÔNG GIAN

Tìm các cụm được xác định theo không gian và hệ phiên mã

03

## QUỸ ĐẠO KHÔNG GIAN

Tìm mối quan hệ giữa các loại tế bào

04

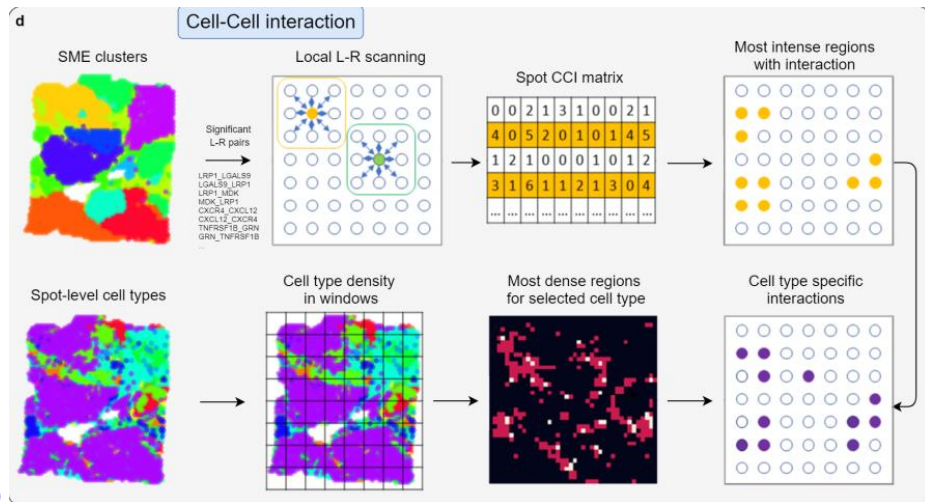
## TƯƠNG TÁC TẾ BÀO

Tìm ra các điểm có tương tác nổi bật trên toàn bộ mô

# STLEARN WORKFLOW

04

## TƯƠNG TÁC TẾ BÀO



- Đầu tiên, ta áp dụng CellPhoneDB vào kết quả phân cụm SME để tìm các cặp phối tử-thụ thể (L-R) nổi bật
- Ta tính điểm về mức độ cùng biểu hiện cặp L-R tại tất cả các điểm, sau đó phân cụm dựa trên tương tác L-R
- Tiếp theo, sự đa dạng của loại tế bào được tính dựa trên số loại tế bào trên một vùng đơn vị của mô
- Cuối cùng, kết hợp dữ liệu về sự đa dạng và dữ liệu về biểu hiện L-R, ta tìm ra được những vùng có khả năng tương tác cao nhất (những vùng có tương tác L-R cao và đa dạng loại tế bào)



**THANKS FOR LISTENING**