

SINGLE CELL RNASEQ & SPATIAL TRANSCRIPTOMICS

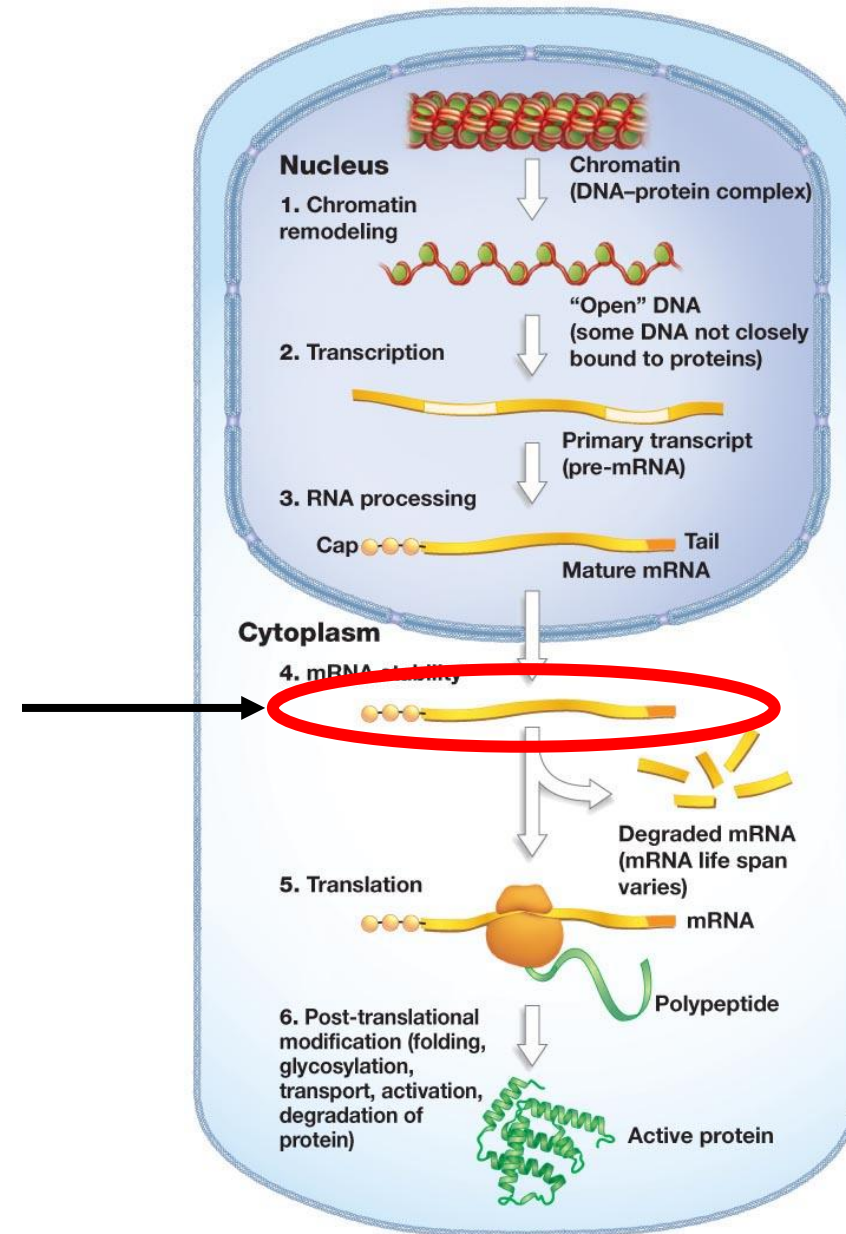
Curso Genómica 2024

Joaquin Garat

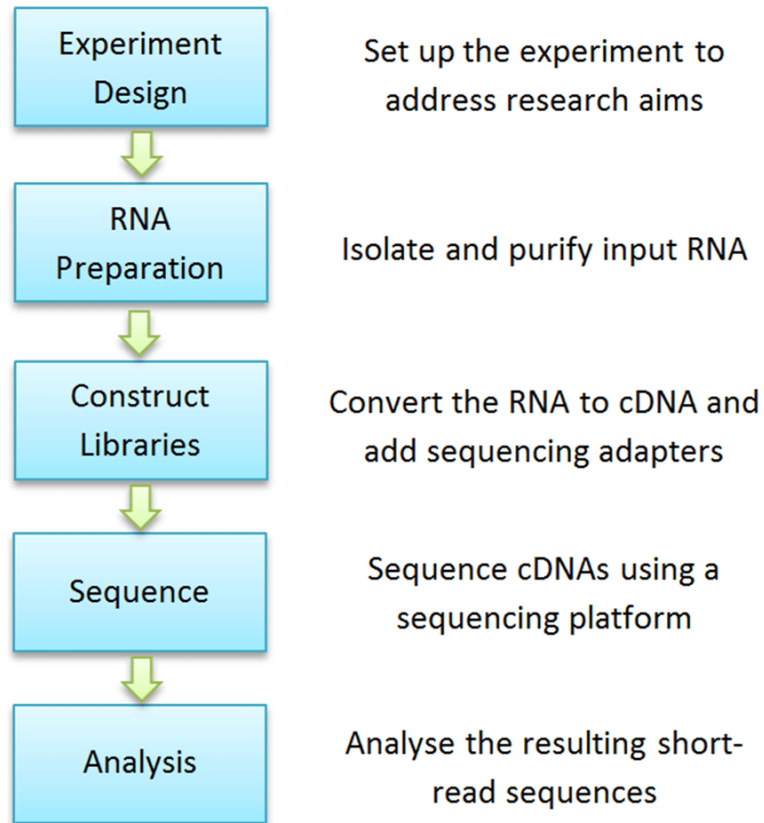


TRANSCRIPTÓMICA

Qué genes se están expresando?
Cuánto se expresan?

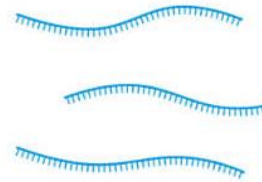


RNA-SEQ

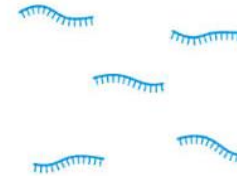


RNA Sequencing

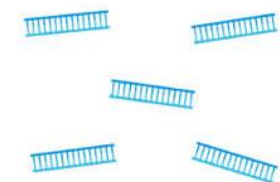
① Isolate RNA from samples



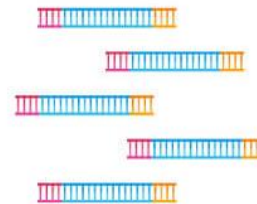
② Fragment RNA into short segments



③ Convert RNA fragments into cDNA



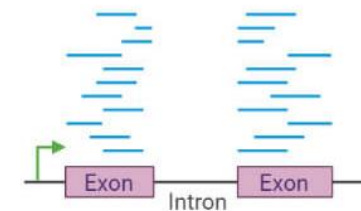
④ Ligate sequencing adapters and amplify



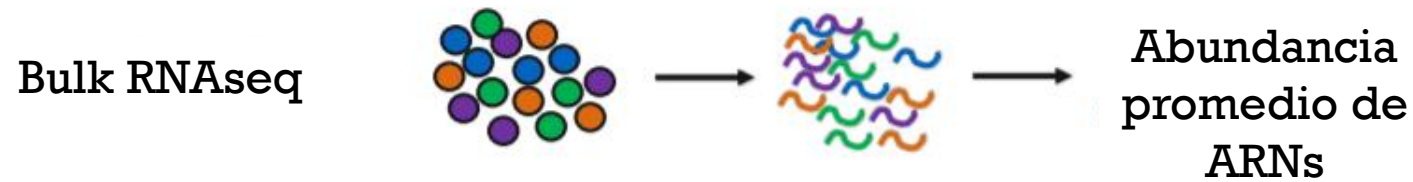
⑤ Perform NGS sequencing



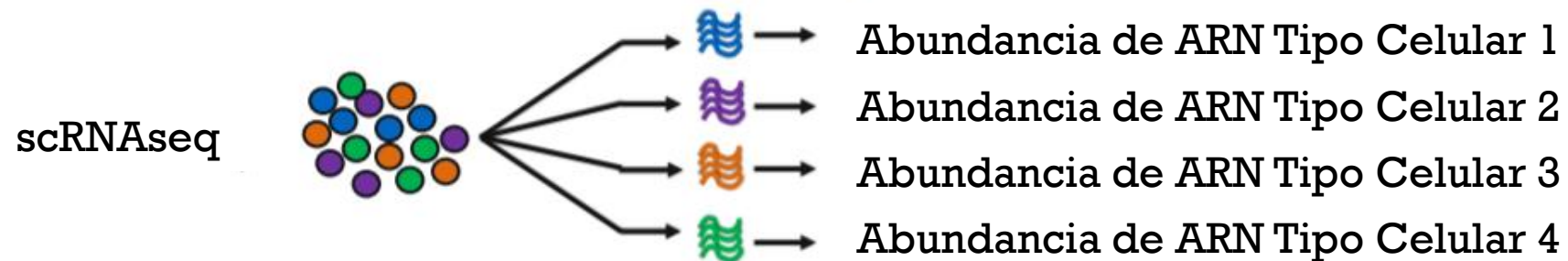
⑥ Map sequencing reads to the transcriptome/genome



TRANSCRIPTÓMICA

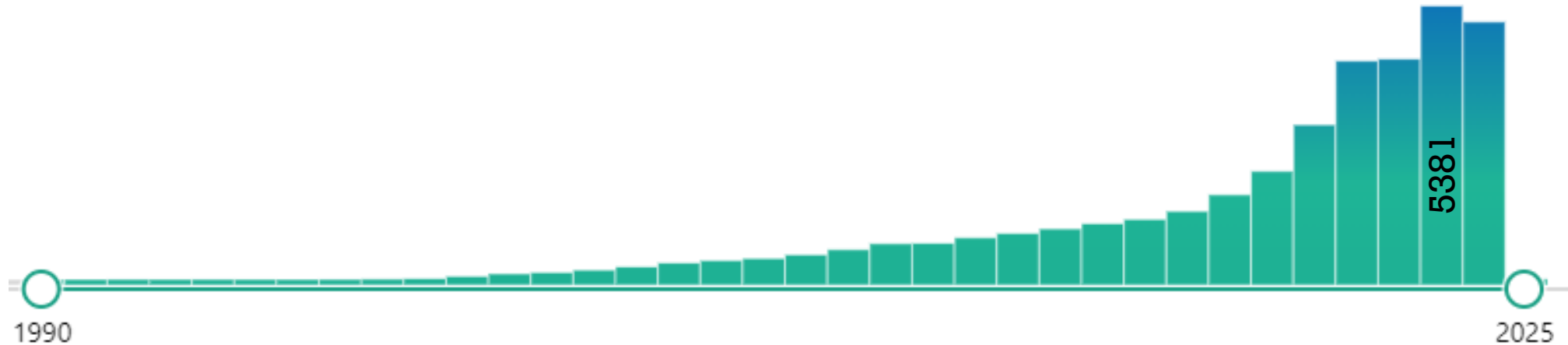


De célula única



BREVE HISTORIA...

Número de publicaciones en Pubmed que mencionan transcriptómica de célula única



Se logró amplificar ARN de células únicas

Amplificación de ARN de células únicas
+
Microarreglos

Amplificación de ARN de células únicas
+
Secuenciación Masiva



ELECCIÓN DE UN MÉTODO ADECUADO

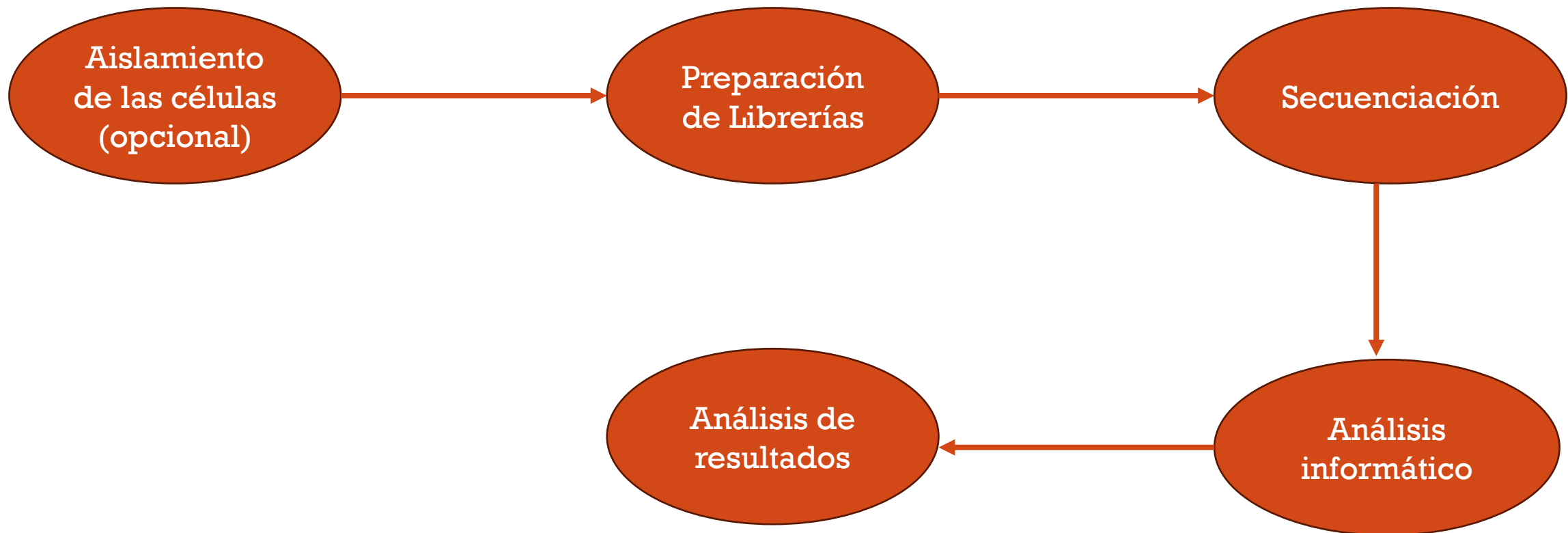
- Tipo de muestra a analizar:
 - Células eucariotas o procariotas
 - Abundancia de ARN esperada
 - Muestra fresca o post-mortem
- Objetivo planteado
 - Necesidad de replicas biológicas?
 - Descubrir poblaciones celulares poco abundantes
 - Expresión diferencial entre tratamientos o condiciones
 - Análisis de desarrollo o destino celular (Pseudotiempo)
 - Análisis multiómico (scATACseq, scTranslatomics, scProteomics)
- Posibilidades económicas



Es importante conocer los métodos disponibles!



OBTENCIÓN DE TRANSCRIPTOMAS DE CÉLULA ÚNICA

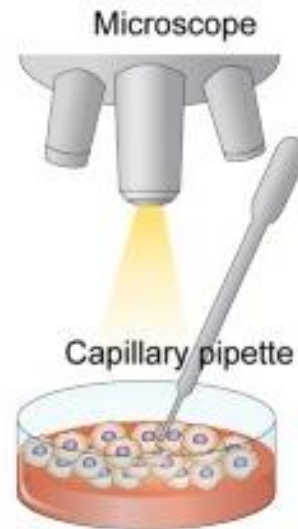


AISLAMIENTO CELULAR

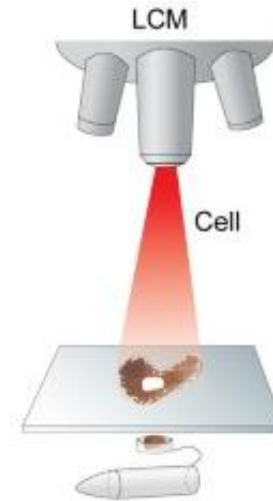
Dilución al límite



Micromanipulación



Microdissección Láser

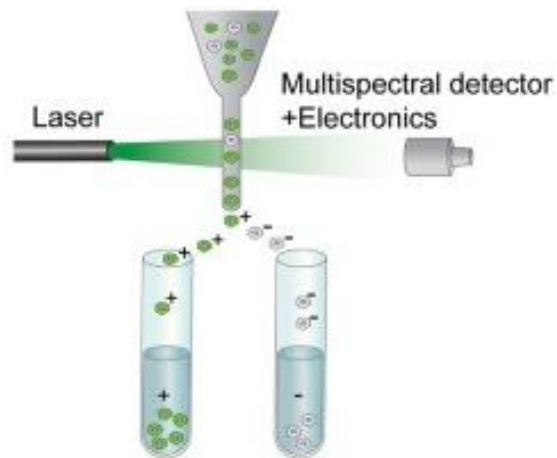


Técnicas poco automatizadas y poco eficientes
Requieren mucho tiempo para aislar grandes números de células



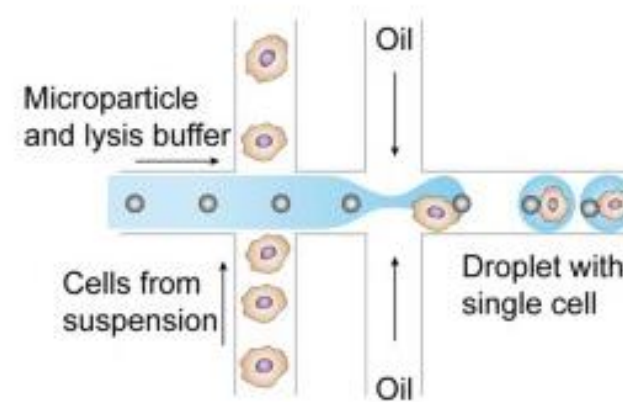
AISLAMIENTO CELULAR

FACS



Smart-seq2

Microfluidos



10x Genomics

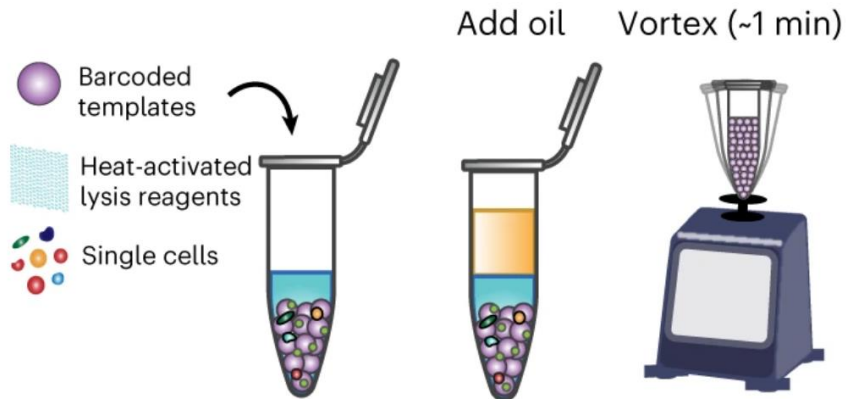
Técnicas con mayor automatización y eficientes (mayor número de células)
Los protocolos de scRNAseq más utilizados utilizan alguno de estos métodos



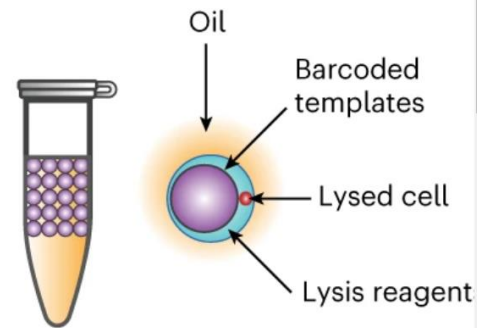
AI SLAM IENTO CELULAR

- Nuevo método low cost basado en droplets
- Similar a método de microfluidos de 10x Genomics pero no requiere equipamiento especial!

Particle-templated emulsification

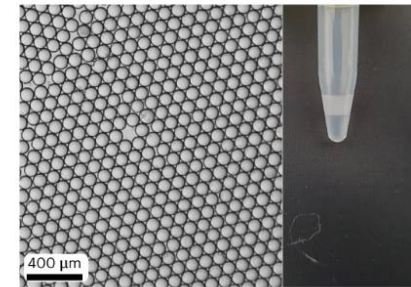


Triggered lysis, mRNA capture

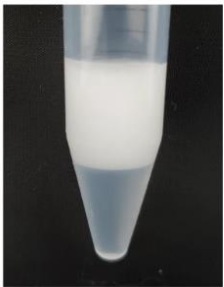
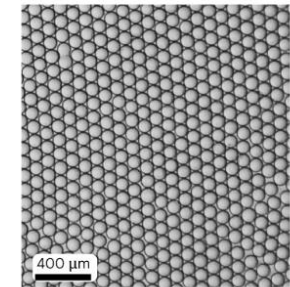


PIP-seq

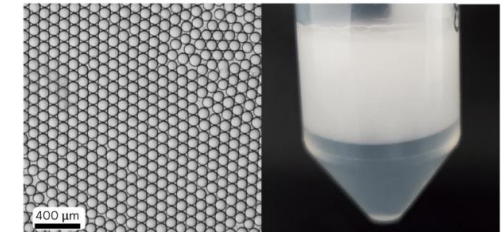
500- μ l tube (35 μ l PIPs ~4,000 cells cell input)



15-ml Falcon (2 ml PIPs ~250,000 cells cell input)



50-ml Falcon (10 ml PIPs ~1 million cells cell input)



PREPARACIÓN DE LIBRERÍAS

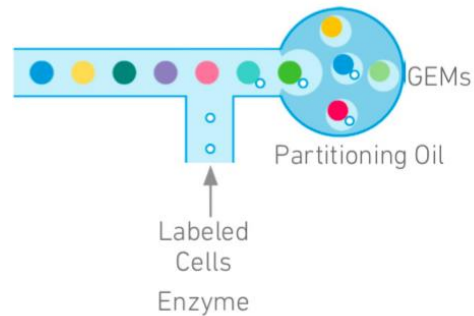
- Dependen del método de aislamiento celular elegido

En placas



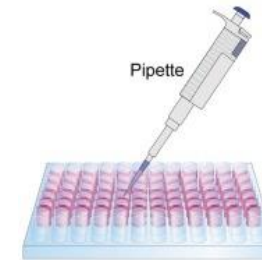
Smart-seq2
Smart-seq3

En Droplets



Chromium 10xGenomics
BacDrop
PIP-seq

En placas sin previo
aislamiento celular



Split-seq
PETRI-seq
MicroSPLIT

Todas comparten los siguientes pasos previos:

Lisis
Celular *

Transcripción
Reversa a
ADNc

Síntesis
segunda
hebra ADNc

Amplificación
del ADNc



CHROMIUM 10X GENOMICS

Equipamiento específico:
U\$\$ 40000 – U\$\$ 95000

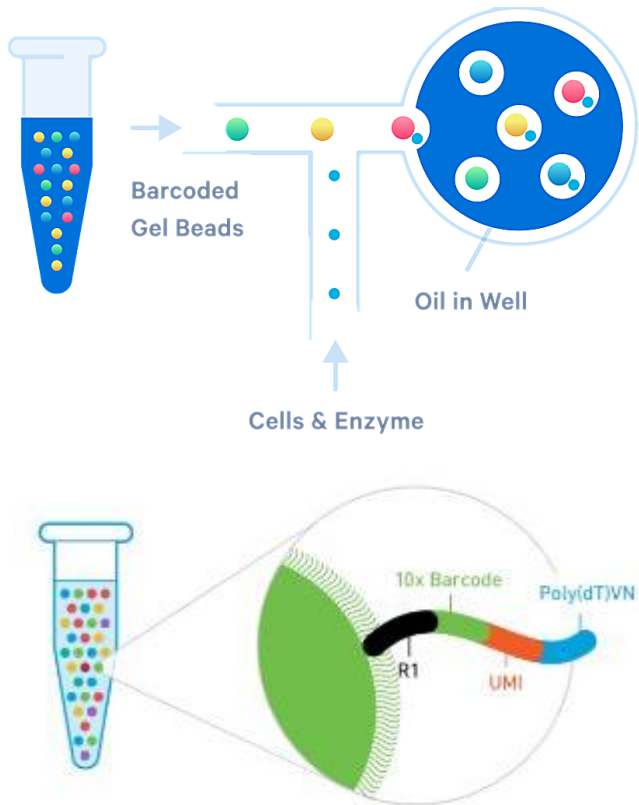


Hay un equipo en el Institut Pasteur

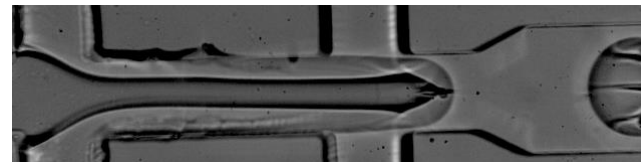
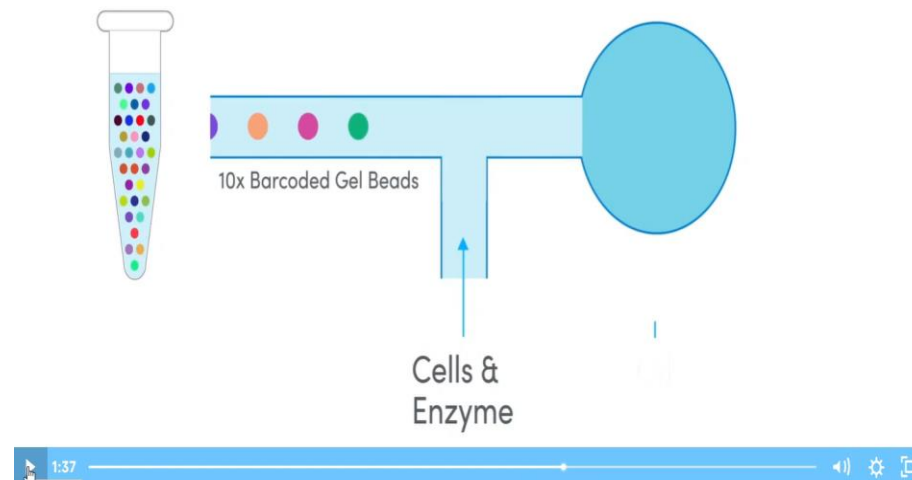


10X GENOMICS

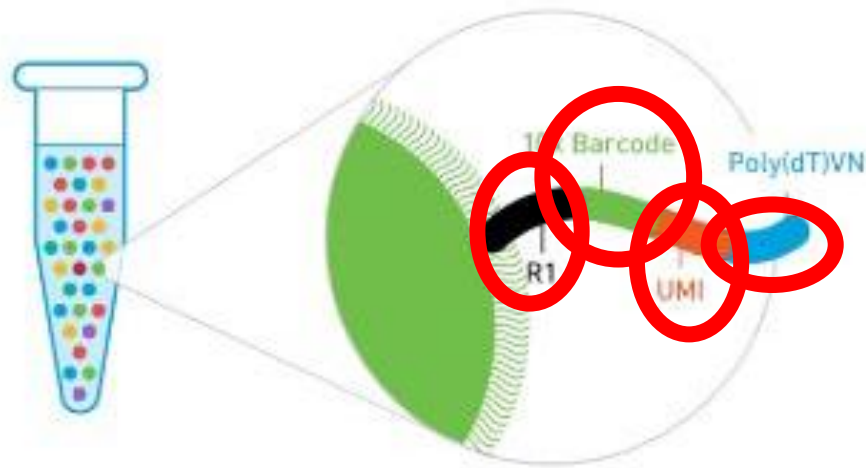
Lisis, Captura de ARNm y preparación de librerías en gotas (droplets)



10x Next GEM Technology for Single Cell Partitioning



OLIGONUCLEÓTIDO DE CAPTURA DE ARN



Captura de ARNm poliadenilado. (Lo más usual)

Secuencias de 16nt denominada barcode (única para cada droplet)

Secuencias de 12nt denominada UMI (única para cada oligo de captura)

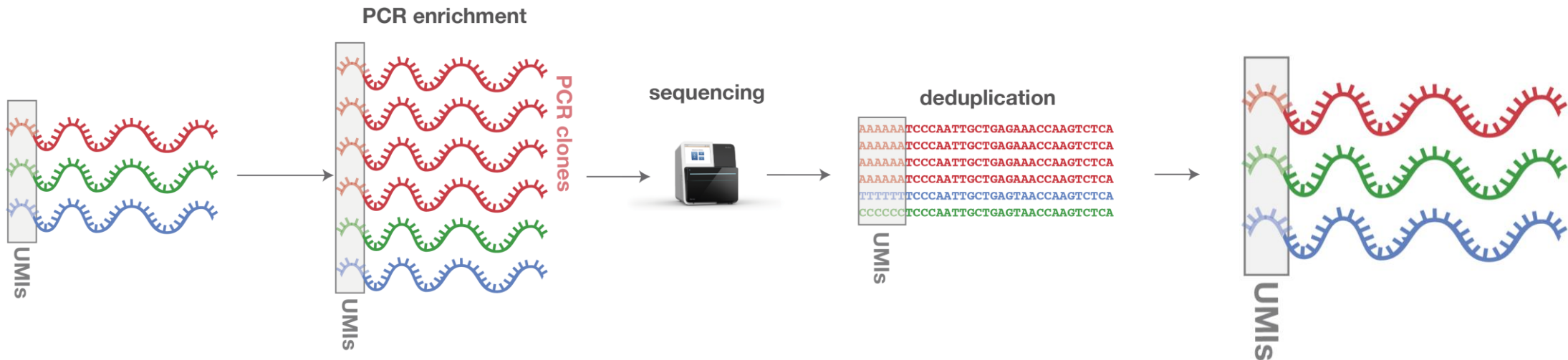
Secuencia adaptador para preparación de librerías de secuenciación masiva



PARA QUÉ SIRVEN LAS SECUENCIAS CONOCIDAS?

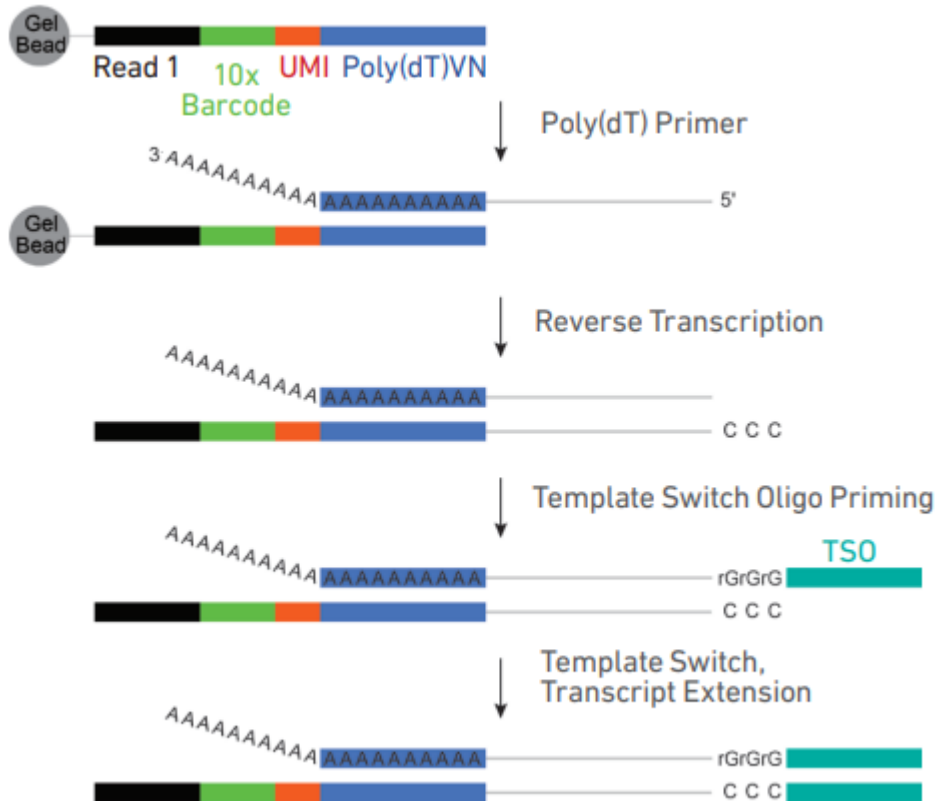
Estas secuencias se mantendrán en los fragmentos que se envían a secuenciar.

- Barcode – Único por cada droplet: Marca los transcritos que vienen de una misma célula
- UMI – Único por cada oligonucleótido de captura: Marca cada ARNm presente en una célula. Útil para corregir por sesgo de amplificación por PCR

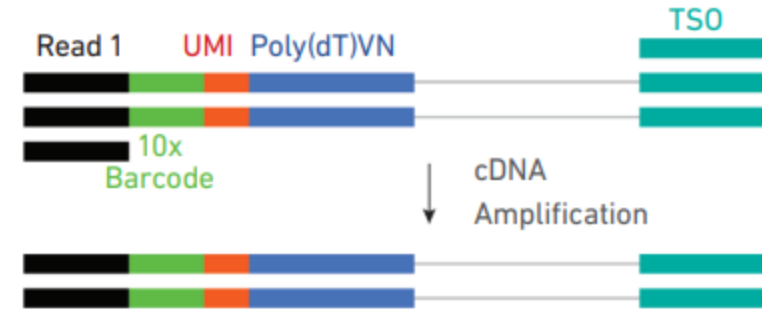


PREPARACIÓN DE LIBRERÍAS

- Síntesis de cDNA

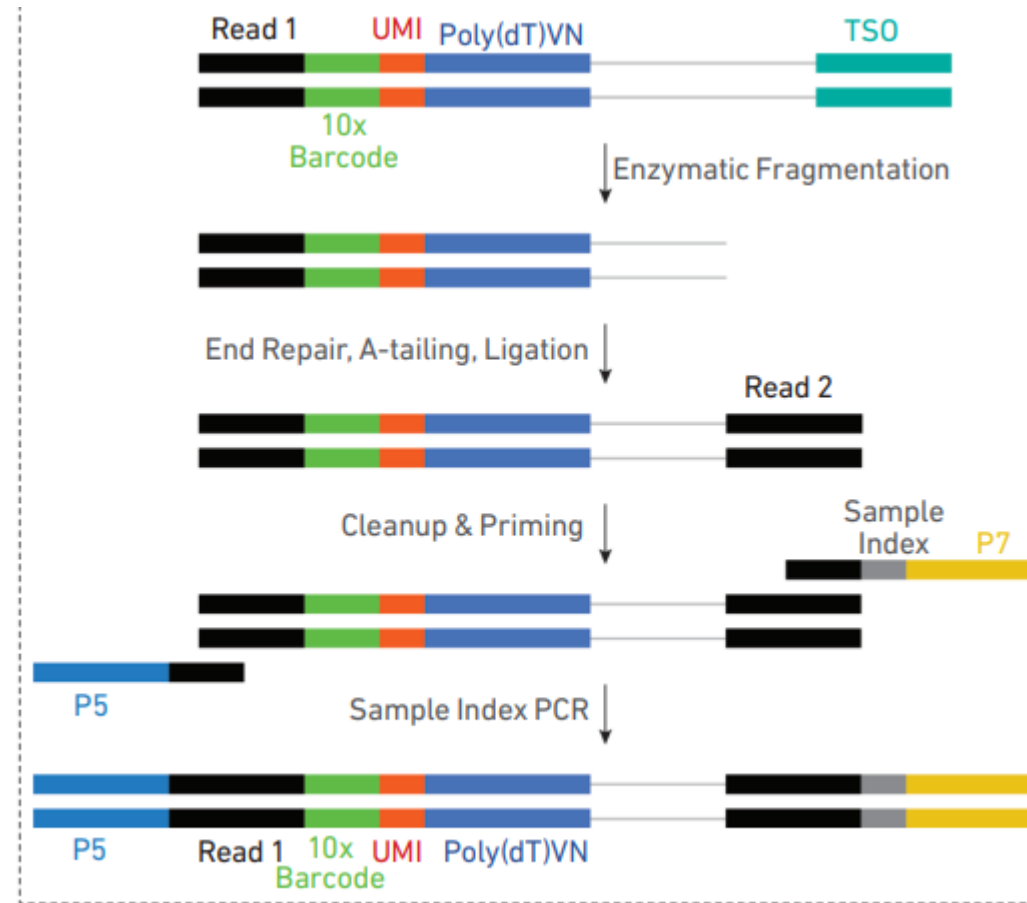


- Amplificación de cDNA



PREPARACIÓN DE LIBRERÍAS

- El producto final que se secuencia contiene el barcode, el UMI en el Read 1 y el extremo 3' del ARNm en el Read 2 (Secuenciación pareada)



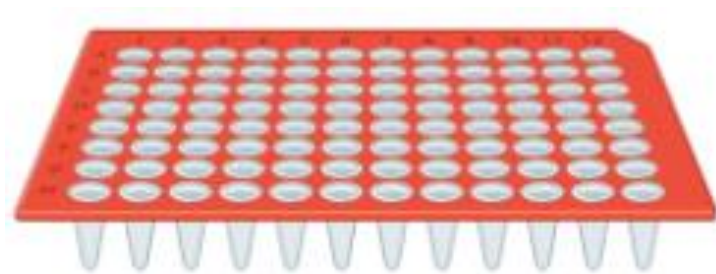
CHROMIUM 10X GENOMICS

- Necesidad de equipamiento específico
- High throughput (alta automatización)
- Baja eficiencia de captura de ARN (detección de transcritos más expresados)
- Secuencias de extremos del ARN (3' UTR)
- Bajo costo por célula (Menor a 1 U\$\$ teniendo el equipo)

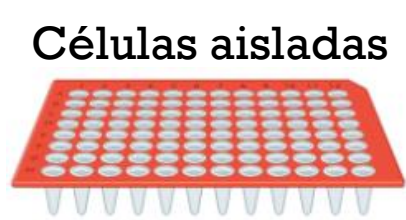


SMART-SEQ

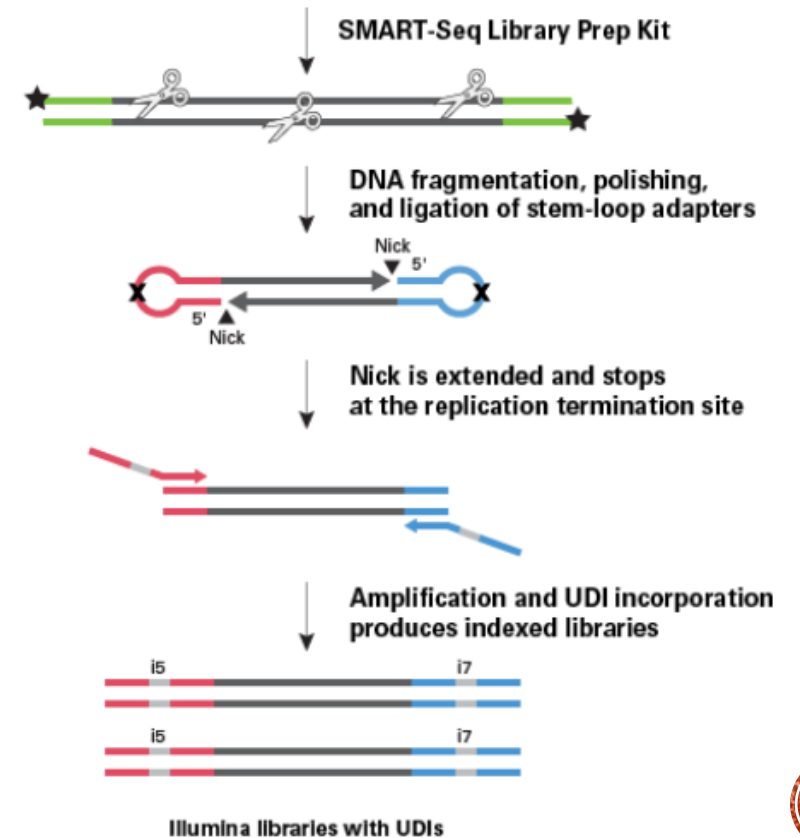
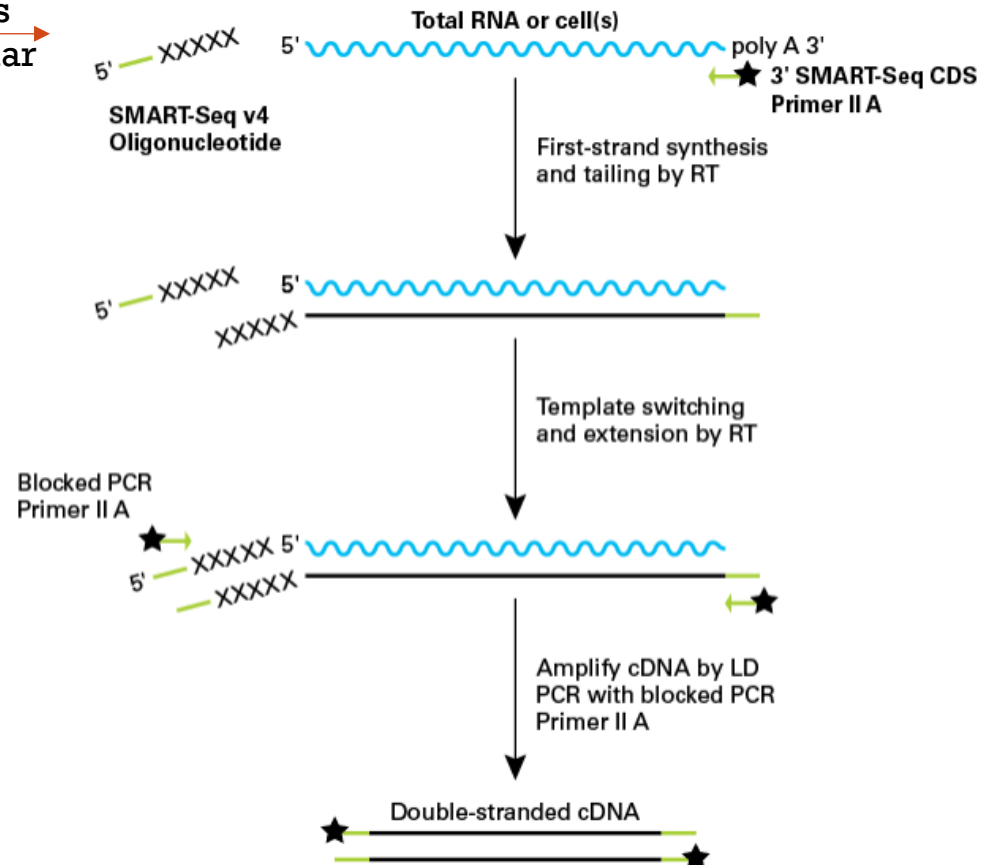
- Es el segundo método más utilizado después de 10x Genomics
- La versión más utilizada por la comunidad científica es Smart-seq2 (2014). Fácil acceso dado que se vende kit comercial (Takara).
- En 2020 se publicó Smart-seq3, y recientemente se comenzó a vender el kit comercial (Takara).
- Comienzo de preparación de librerías en multiwell (1 reacción por célula)



PREPARACIÓN DE LIBRERÍAS



Lisis
Celular

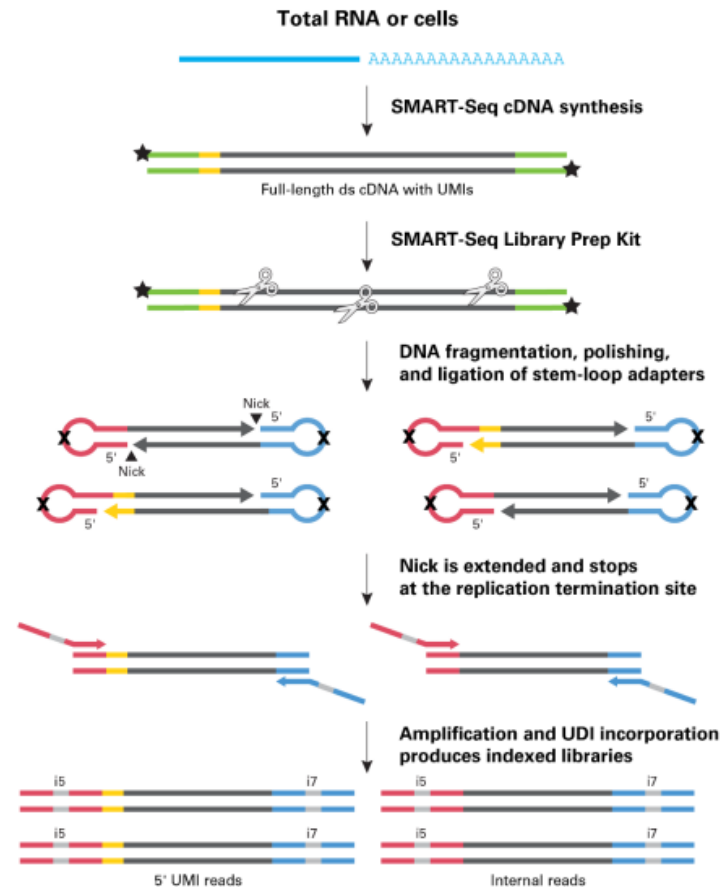
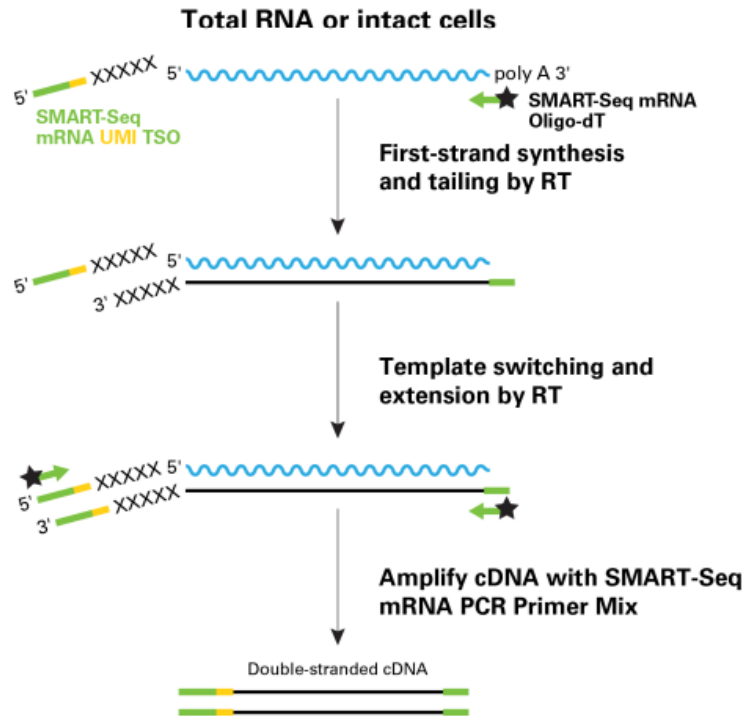


Takara Bio



SMARTSEQ3 (CON UMI)

- Extremos 5' tienen UMI y secuencias internas no



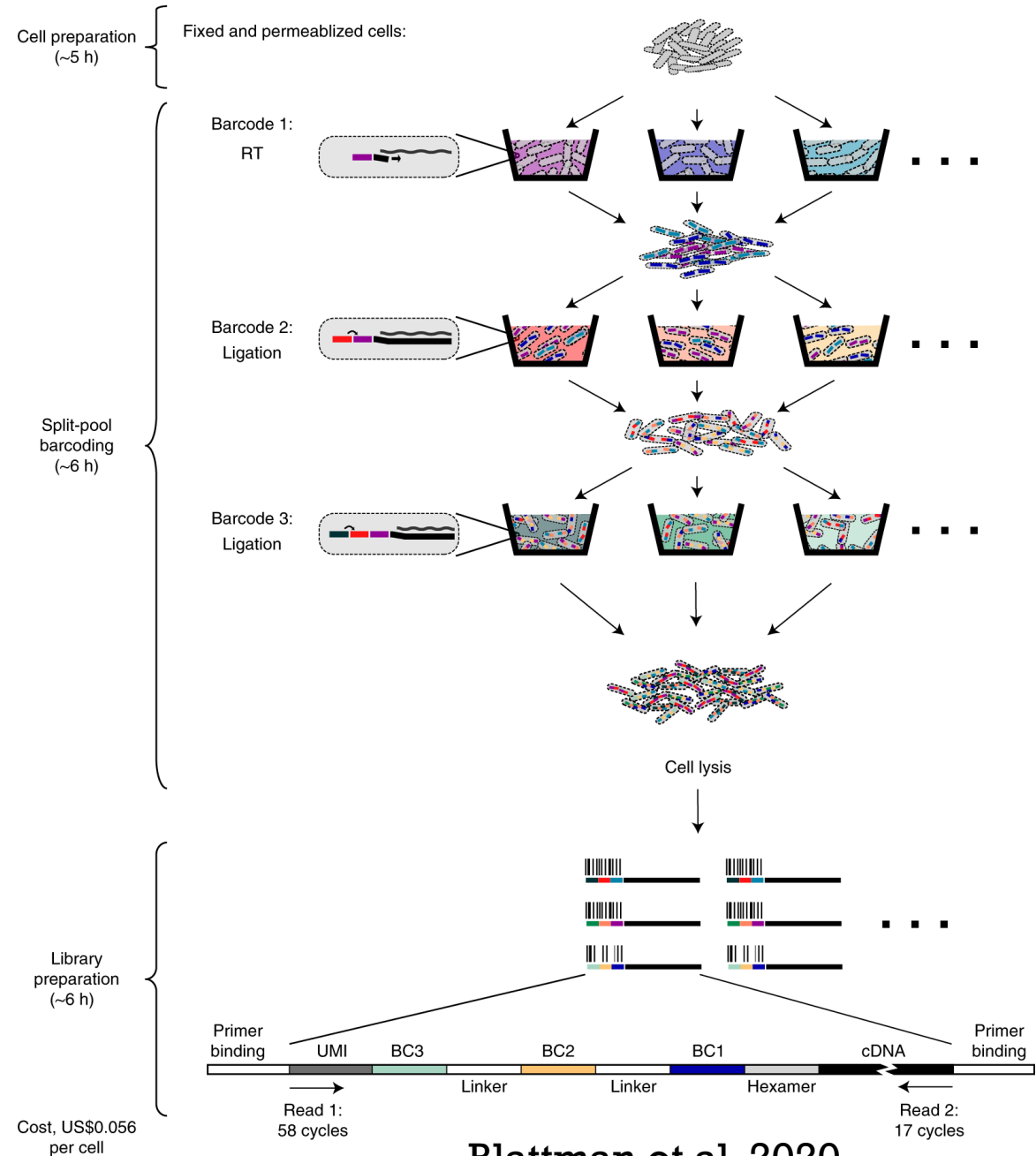
SMARTSEQ

- Necesidad de aislamiento previo de células con otro método (principalmente FACS)
- Low throughput (poca automatización)
- Alta eficiencia de captura de ARN (detección de transcritos de baja expresión)
- Secuencias de todo el transcripto (ej: Análisis de splicing)
- Alto costo por célula (Al menos U\$\$30)



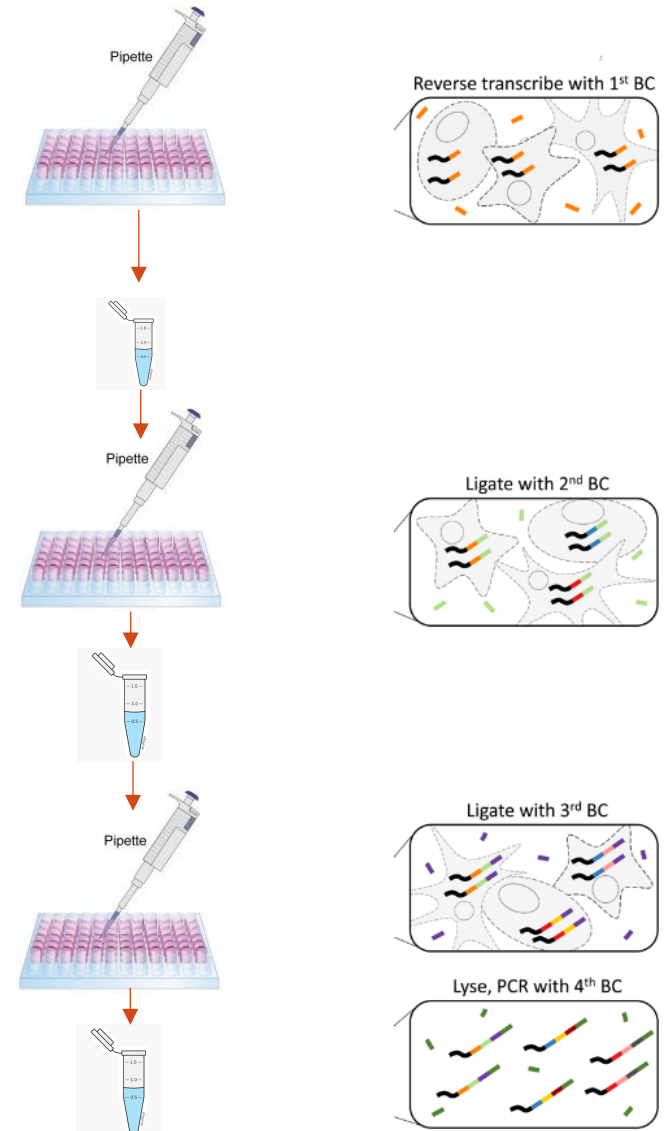
SPLIT-SEQ

- Sin aislar células
- Low Cost (no necesita equipo especial)
- High throughput
- Basado en combinación de barcodes



SPLIT-SEQ

- 1) Repartir células fijadas en multiwell 96 pocillos
- 2) Retrotranscripción incorporando barcode único en primer (random y poly(dT)) por pocillo (sin lisis)
- 3) Mezclar las células
- 4) Repartir células en multiwell
- 5) Ligación barcode
- 6) Mezclar las células
- 7) Repartir células en multiwell
- 8) Ligación barcode + UMI
- 9) Mezclar las células
- 10) Lisis y PCR



Rosenberg et al, 2018



SPLIT-SEQ

- Sin Necesidad de aislar células
- High throughput (En el ejemplo anterior, con multiwells de 96 pocillos se podría preparar librerías para 1 millón de células)
- Baja eficiencia de captura de ARN (detección de transcritos de alta expresión)
- Secuencias de extremo del transcripto (UTR)
- Bajo costo por célula y sin equipamiento específico (menos de U\$\$0.1)



ESTOS MÉTODOS SIRVEN PARA BACTERIAS?

- Particularidades:
- Pequeño tamaño (equipos de microfluidos optimizados para células eucariotas)
- Poca cantidad de ARNm por célula (femtogramos)
- Complejidades para el paso de lisis celular (Pared celular)

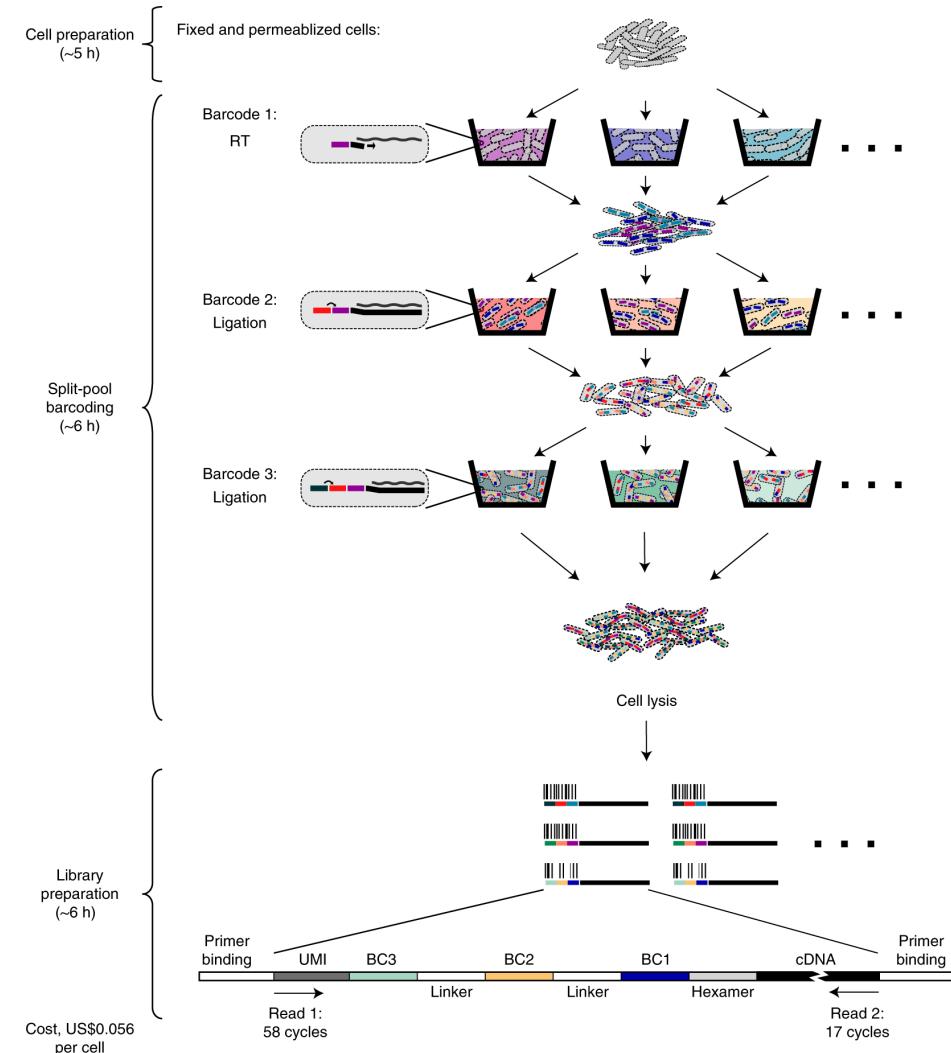
Alternativas:

PETRI-seq
MicroSPLIT
BacDrop



PETRI-SEQ Y MICROSPPLIT

- Primeros métodos de scRNAseq en bacterias con high throughput.
- Siguen la estrategia de Split-seq.
- Permeabilización depende de bacteria con la que se trabaje: Lisozima, lisostafina. Debe ser optimizado para comunidades complejas.
- Uso de polyA polimerasa de E.coli (MicroSPLIT).
- Luego se sigue el protocolo de manera similar a Split-seq. Los primers de la RT dependerán si hubo poliadenilación previa.
- Problema rRNA en ambos métodos. Necesidad de secuenciar mucho para secuenciar ARNm
- Detectan 5-10% del ARNm total en las células



BACDROP

- Basado en métodos anteriores (introduce barcode en bacterias permeabilizadas)
- Utiliza oligos anti-sentido RNAsa H para depletar ARNr. Combina 1er barcode con nuevo barcode en droplets.

1. rRNA and DNA depletion in fixed and permeabilized cells



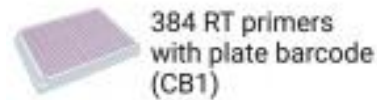
RNase H ↓ rRNA depletion



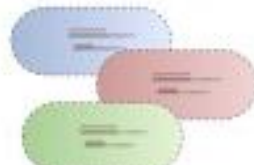
DNase I ↓ DNA removal



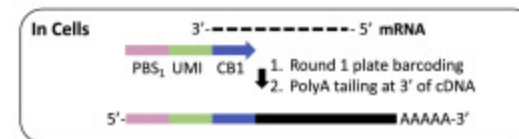
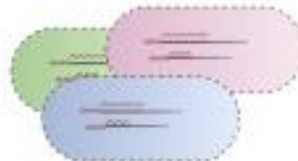
2. Reverse Transcription



cDNA synthesis and round 1 plate barcoding

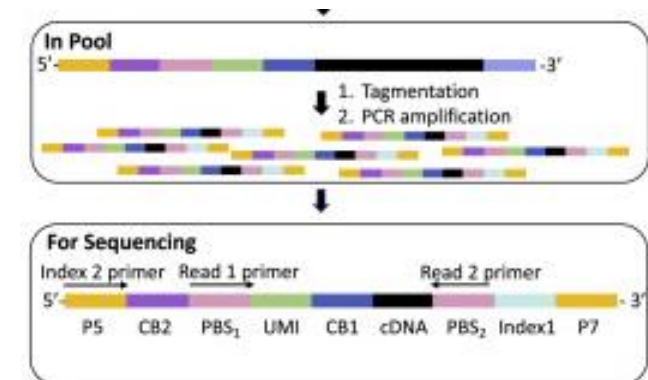
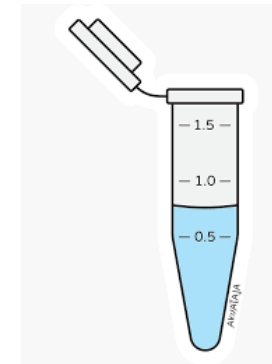
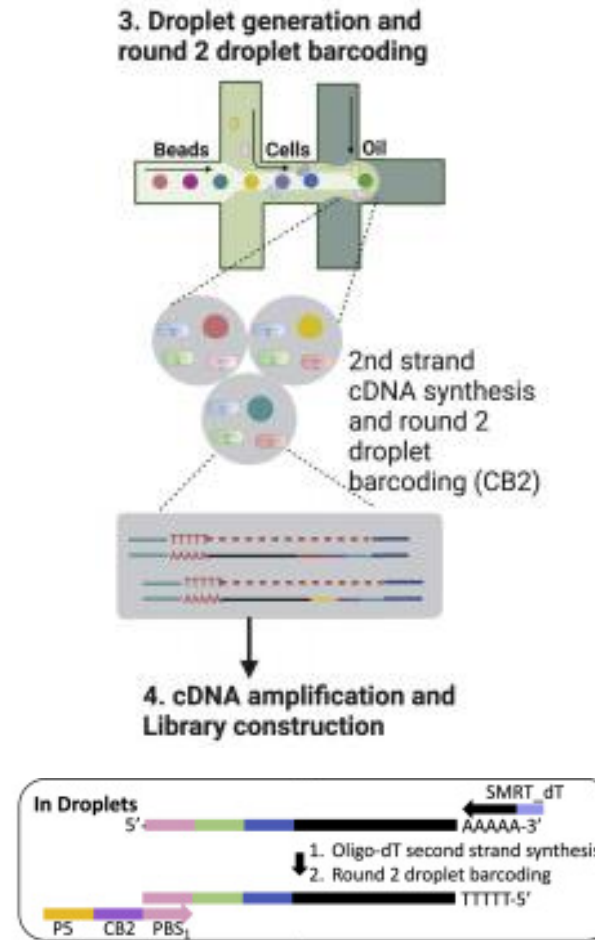


Poly A tailing at 3' of cDNA



BACDROP

- Más de una bacteria por Droplet
- Ya tienen barcode del paso anterior (combinación permite distinguir)
- Poseen polyA por paso anterior (captura de ARNm igual que en protocolo clásico de 10x Genomics)



BACDROP

- High Throughput (Automatizado)
- Permeabilización debe adaptarse para comunidades complejas
- Utiliza Plataforma más utilizada de scRNAseq (Chromium)
- Baja eficiencia de captura de ARN (detección de transcritos de alta expresión)
- Requiere equipamiento específico



SECUENCIACIÓN

- Luego de preparar las librerías se debe secuenciar.
- La profundidad de secuenciación dependerá del objetivo y del protocolo elegido previamente (saturación no es igual para todos los métodos).
- Es conveniente usar secuenciadores de alta capacidad (se requiere muchas lecturas).
- Necesidad de lecturas pareadas (barcodes en Read 1 y ARNm en Read 2)

Smartseq	10x Genomics	Split-seq	PETRIseq MicroSPLIT	BacDrop
500k	50k	50k	50k	50k
Reads por célula	Reads por célula	Reads por célula	Reads por célula	Reads por célula

En un experimento con 25000 células: 12500M reads en Smartseq, 1250M reads en los otros métodos



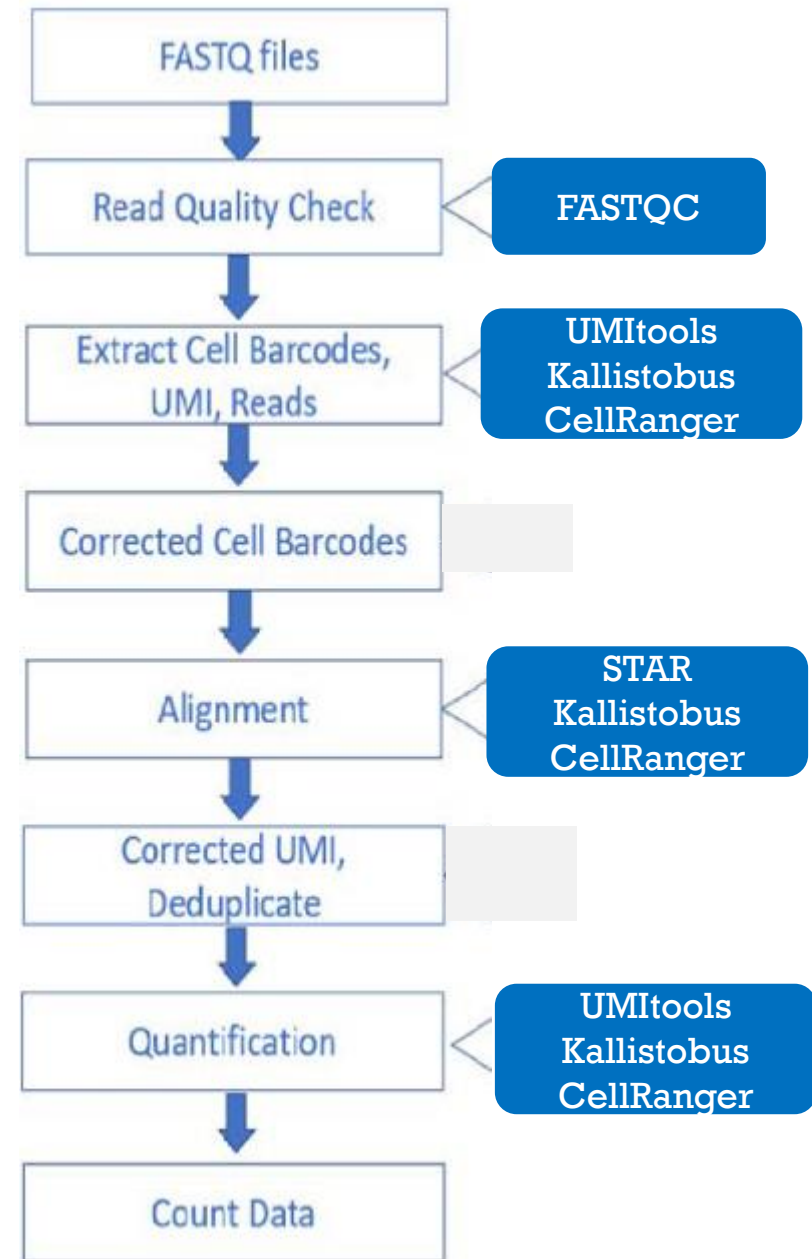
VENTAJAS DE CADA TÉCNICA

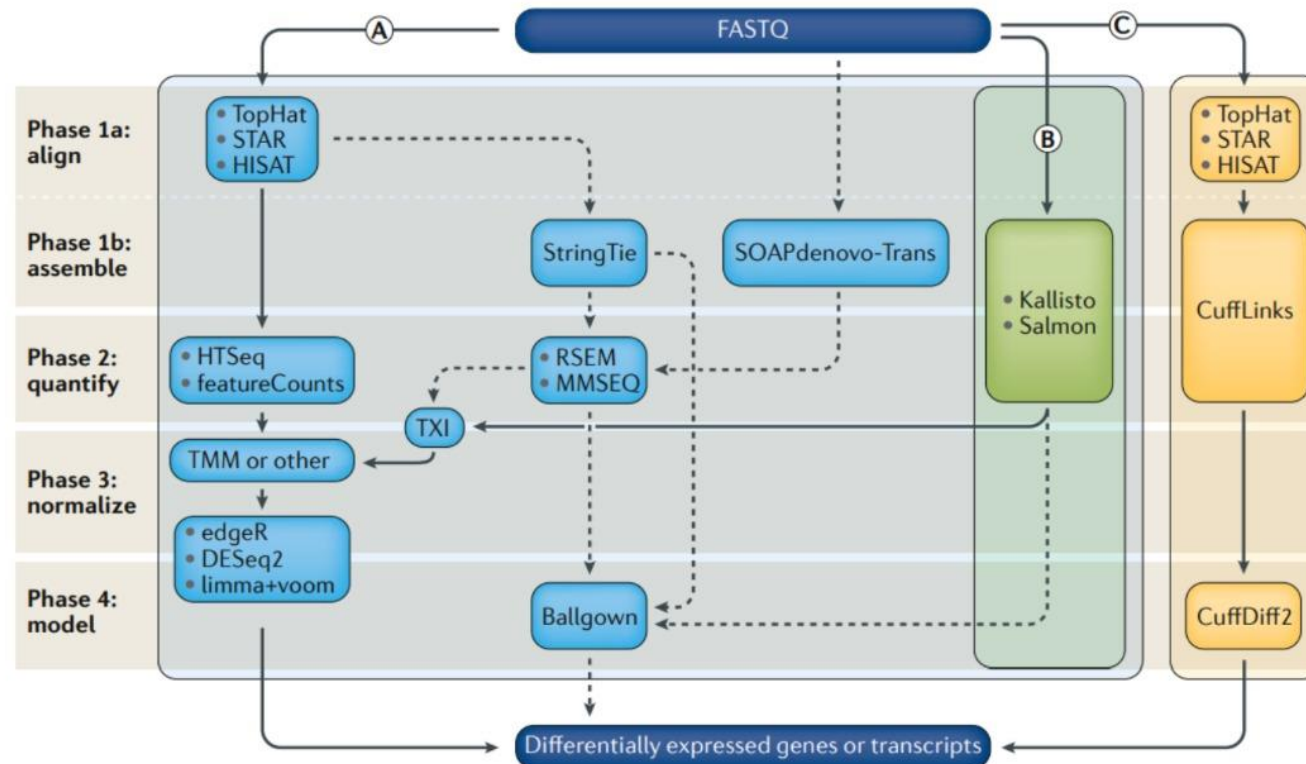
- **Smartseq**
 - Detecta transcritos de alta y baja expresión (Eficiencia de captura superior a otros métodos)
 - Detecta lecturas para todo el transcripto
- **10x Chromium**
 - Altamente automatizado
 - Gran capacidad de salida de datos
 - Costos más accesibles que Smartseq
 - Útil para analizar grandes cantidades de células
- **SplitSeq**
 - Costo más bajo
 - No necesita equipos especiales (ni para aislar células ni para preparar librerías)
- **PETRI-seq y MicroSPLIT**
 - Similar a SplitSeq pero adaptado para bacterias
- **BacDrop**
 - Similar a 10x Chromium pero adaptado para bacterias



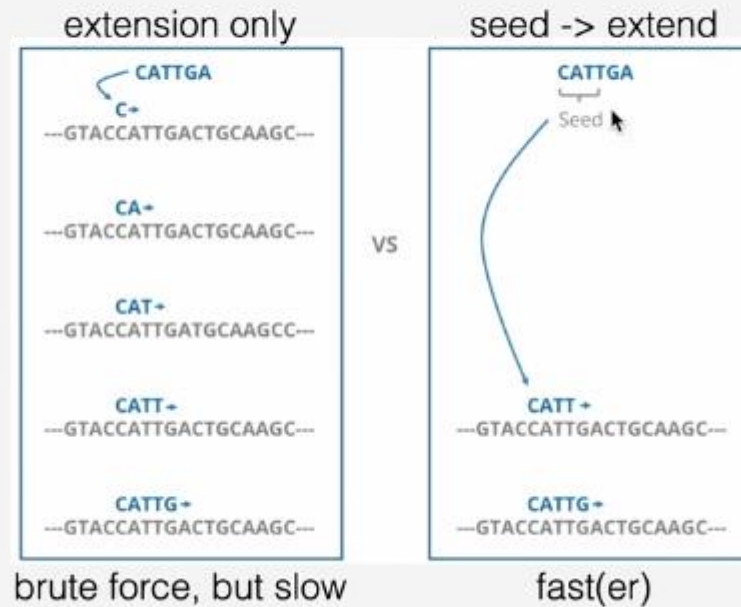
ANÁLISIS DE DATOS

- Tenemos las secuencias... y ahora?
- Los pasos iniciales son muy similares a bulk RNAseq pero...
- Pueden ser miles de millones de lecturas (capacidad computacional)
- Análisis de calidad
- Alineamiento (teniendo en cuenta barcodes y UMIs)
- Análisis de la matriz de conteos (Detección de poblaciones celulares, expresión diferencial, etc)





‘Seed and Extend’ alignments



<https://blog.sbggenomics.com/short-read-alignment-seeding/>

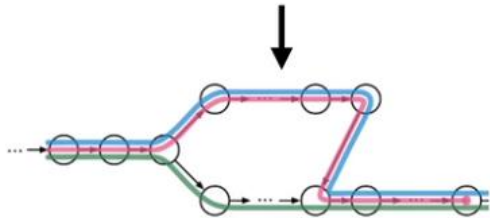


PSEUDOALIGNMENT

mapping reads with Kallisto



create transcriptome de Bruijn graph (T-DBG)
using k-mers (31 nucleotides)

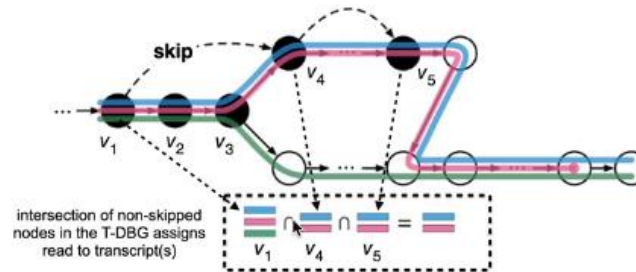


Bray et al., Nat. Biotech., 2016

mapping reads with Kallisto



Don't even need to search for all
possible K-mers

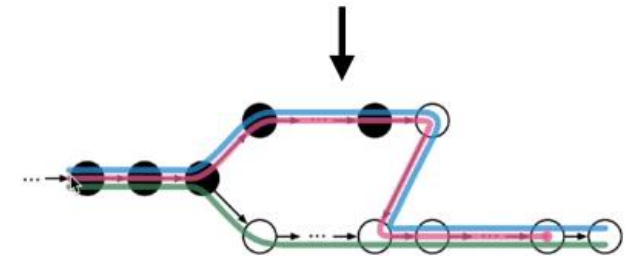


Bray et al., Nat. Biotech., 2016

mapping reads with Kallisto

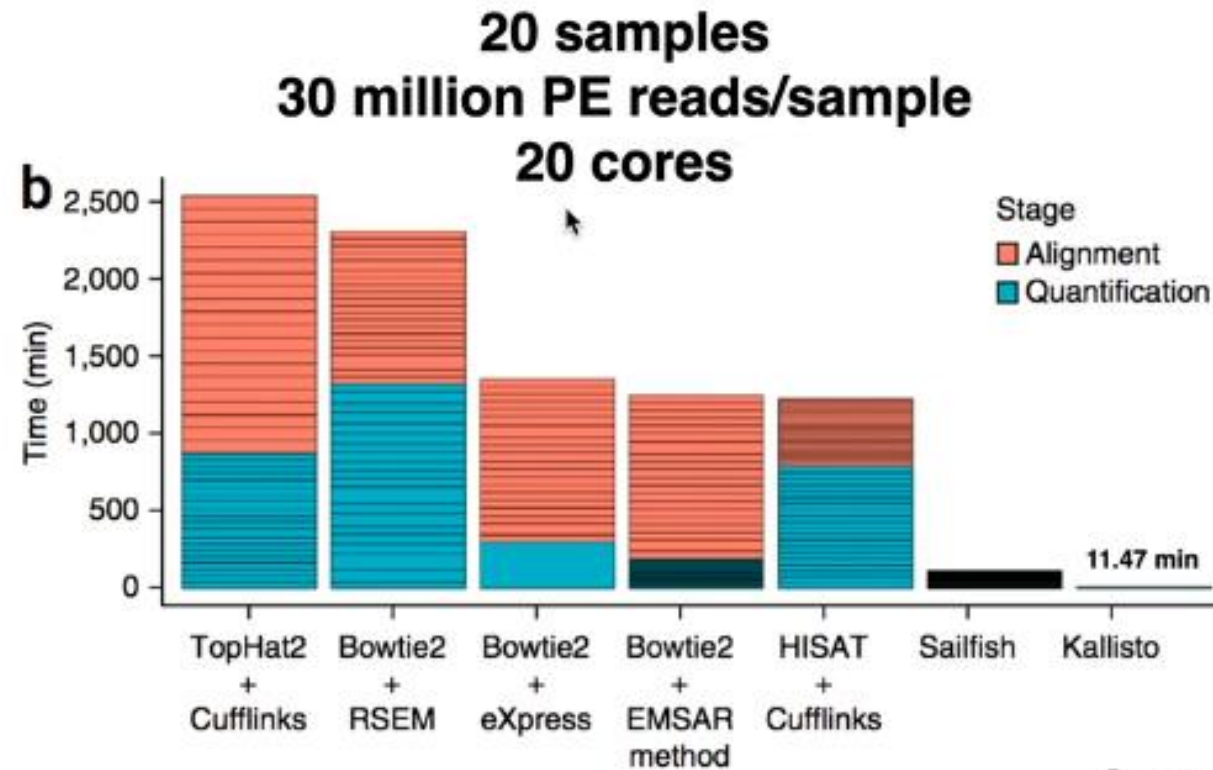


Don't align, instead find transcripts that are
compatible with the read



Bray et al., N

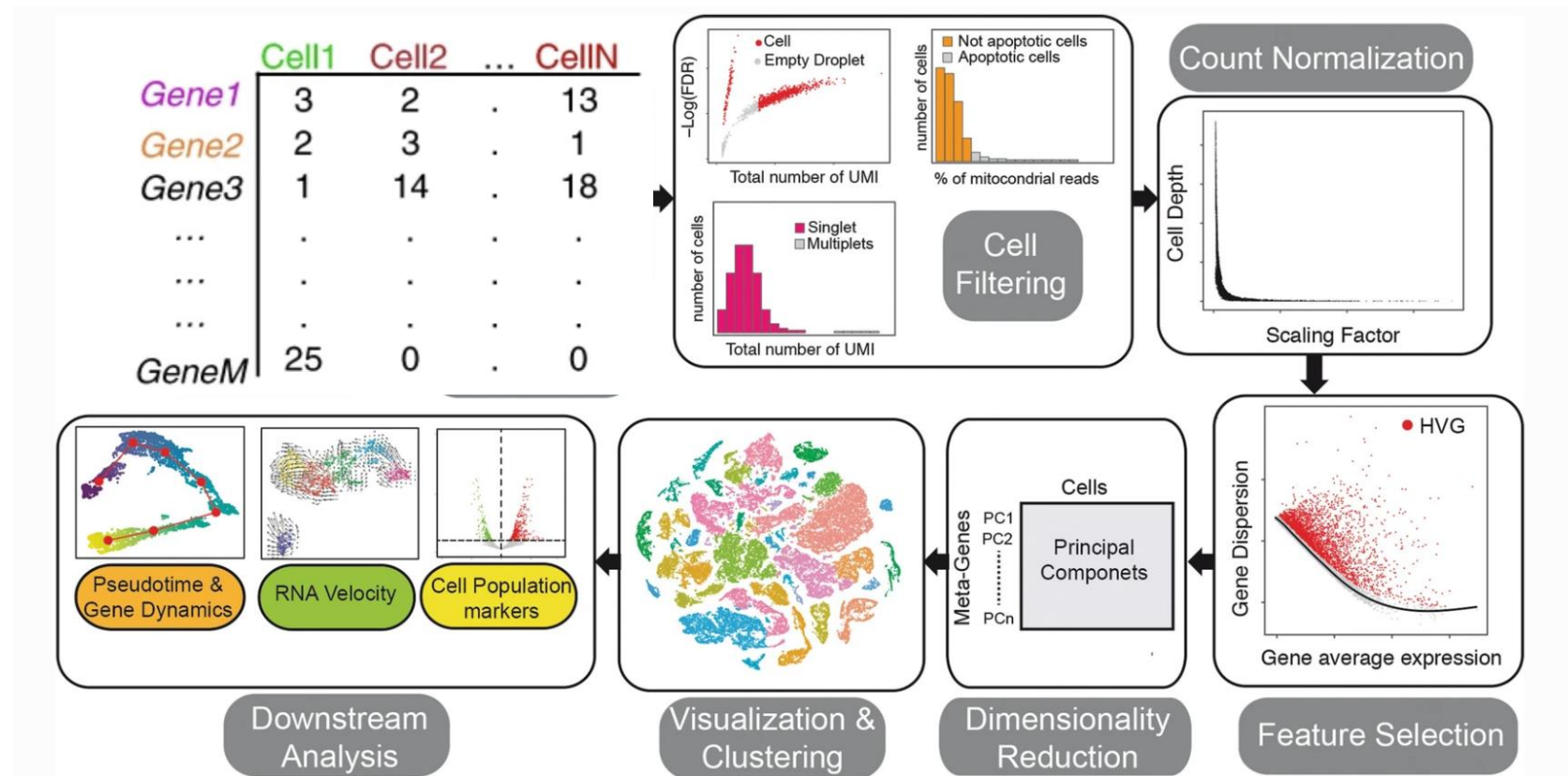




Bray et al., Nat. Biotech., 2016

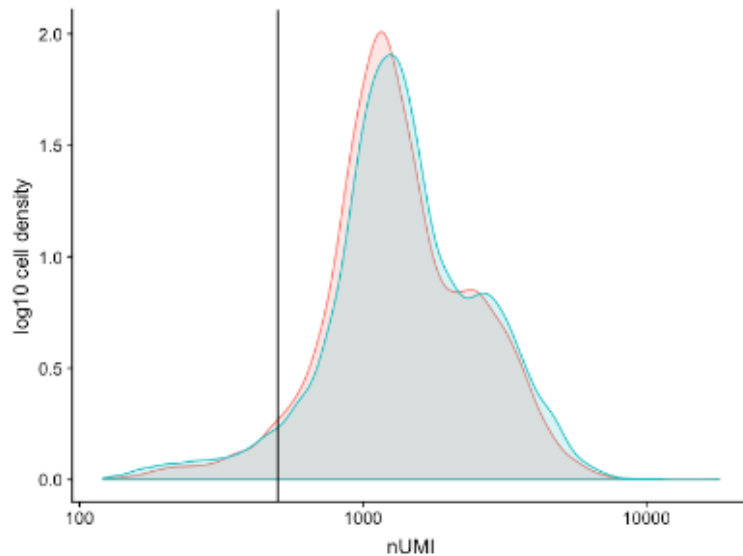


ANÁLISIS DE LA MATRIZ DE CONTEOS

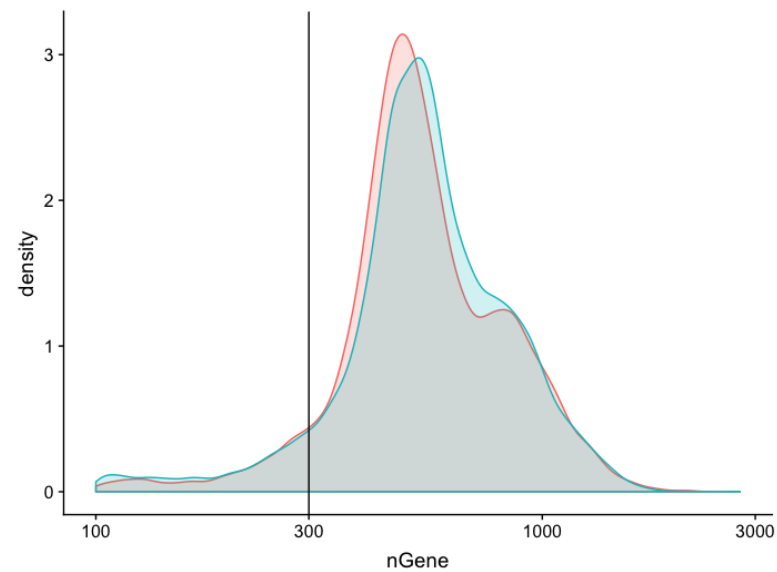


CONTROL DE CALIDAD CELULAR

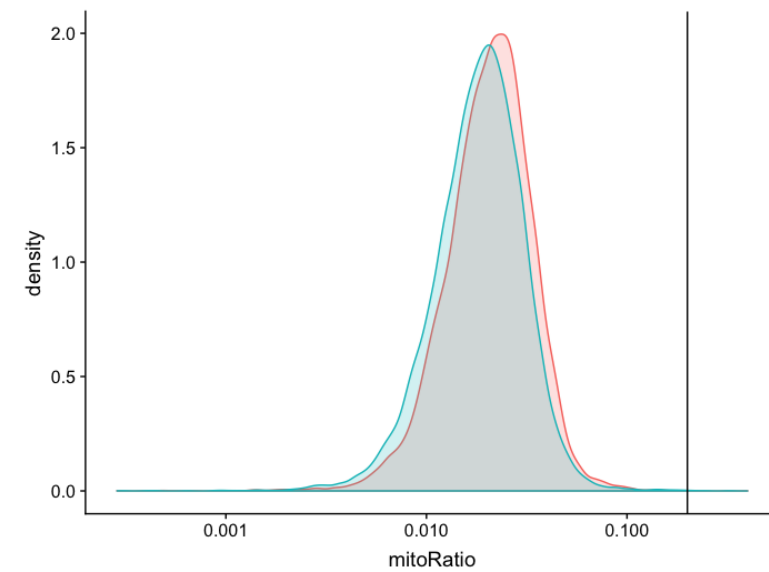
Número de UMI por célula



Número de genes por célula



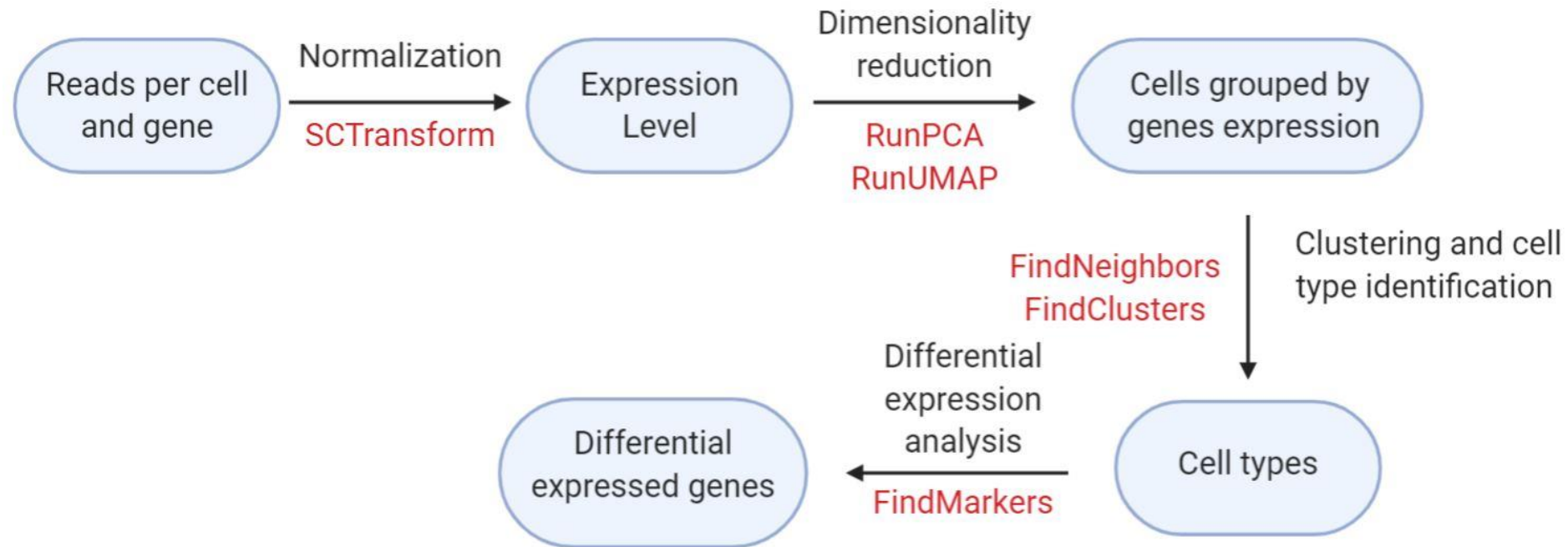
Fracción de genes mitocondriales



ANÁLISIS DE MATRÍZ DE CONTEOS



ANÁLISIS DE MATRIZ DE CONTEOS

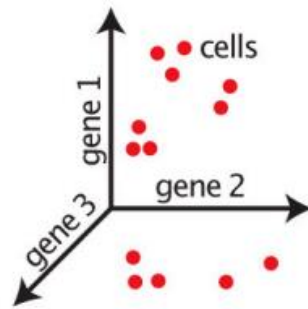


IDENTIFICACIÓN DE TIPOS O ESTADOS CELULARES

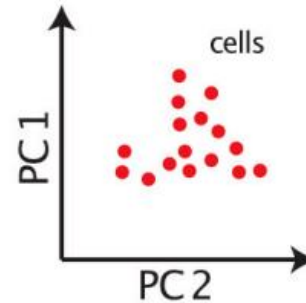
Matriz de conteos

	cell 1	cell 2		cell 66,000
gene 1	0	0	..	1
gene 2	1	5	..	2
....	...	...	...	...
gene 26,000	4	15	...	0

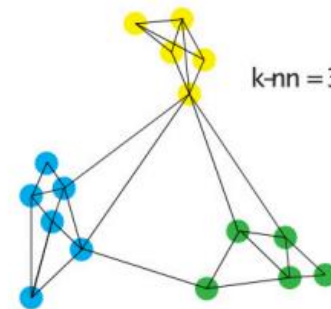
Espacio de N-dimensiones
Ejemplo con 3



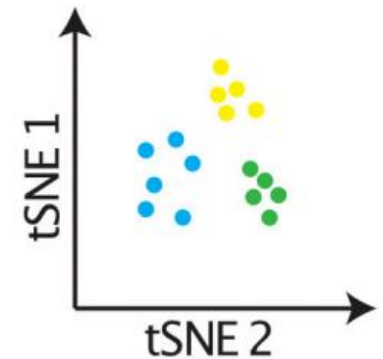
Reducción de dimensiones PCA
a 50 PC



Clustering KNN
50 PC aprox.

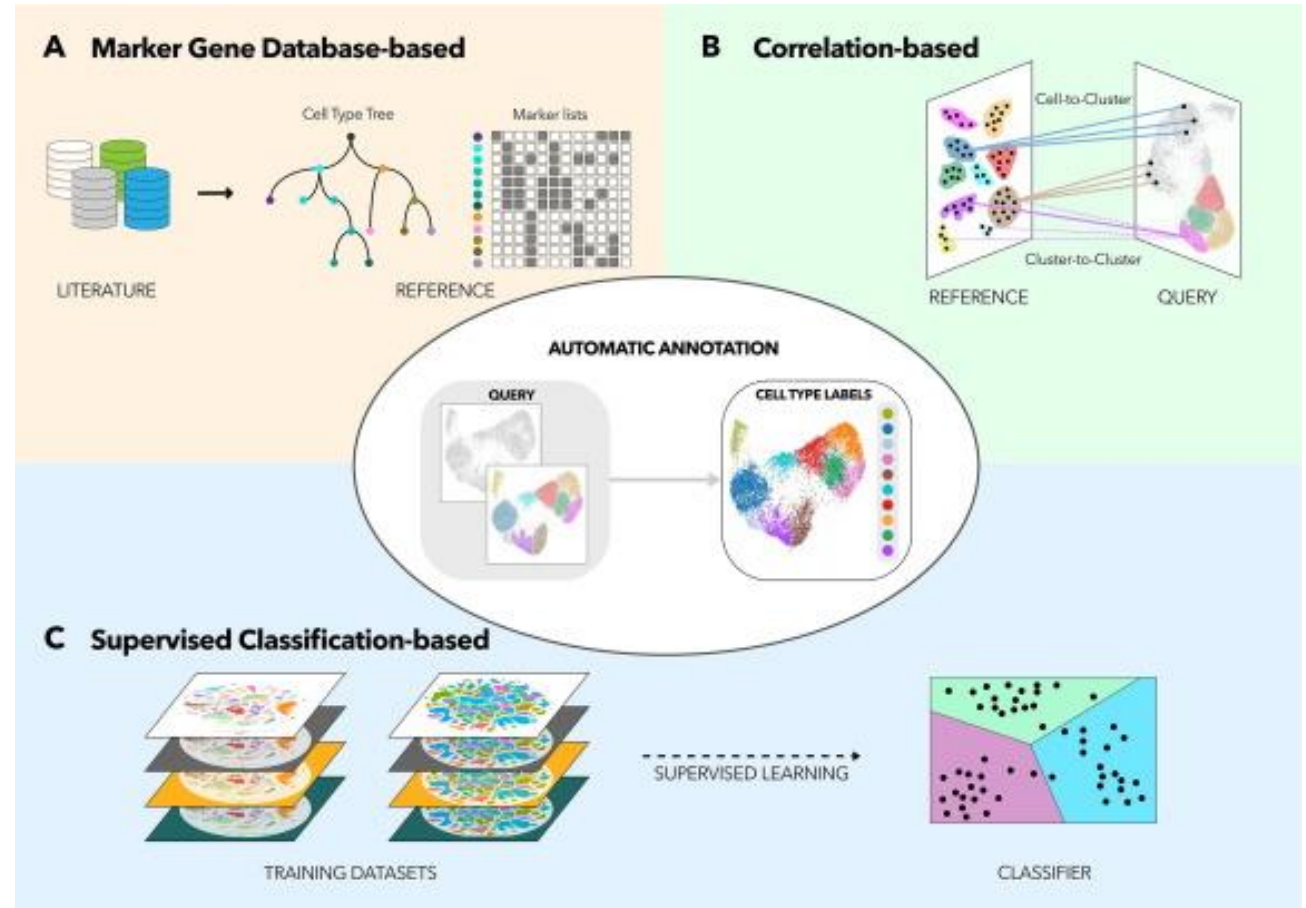


Visualización de los clusters con t-SNE o UMAP



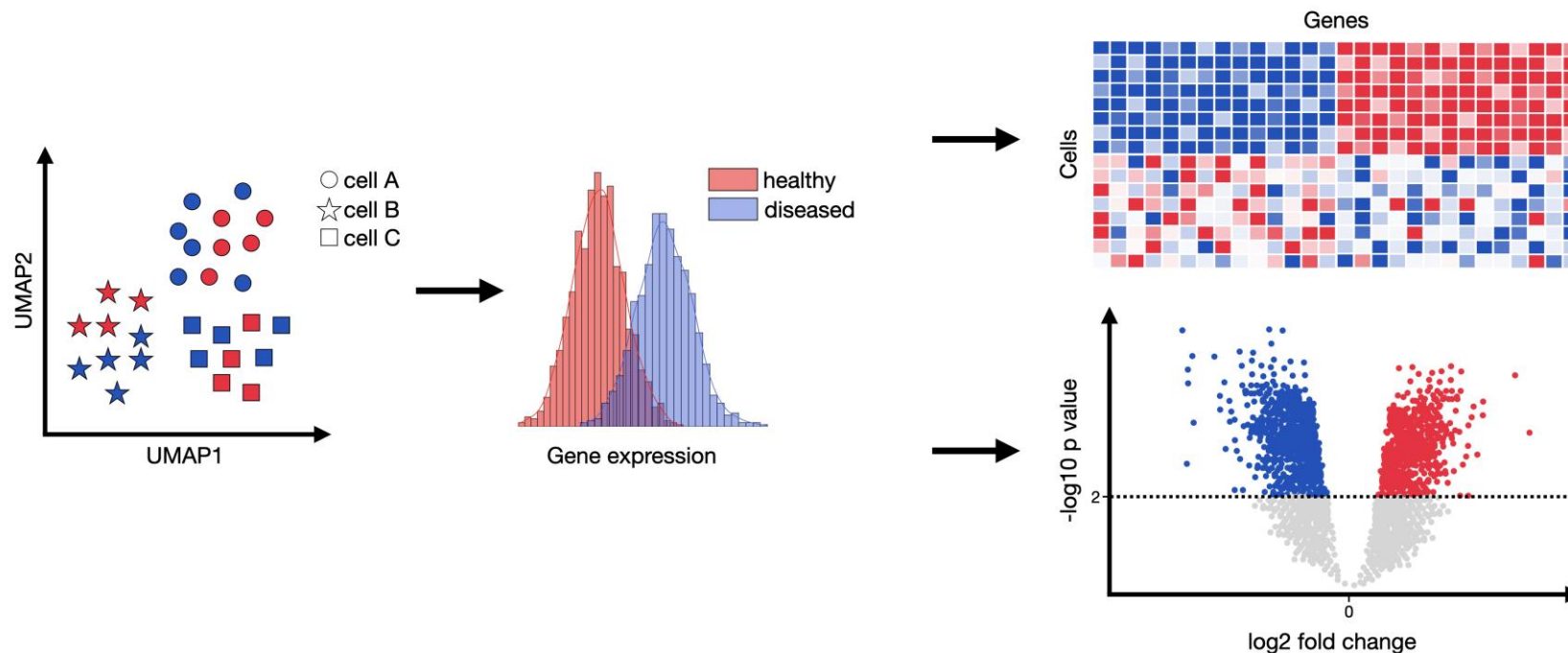
IDENTIFICACIÓN DE TIPOS CELULARES

- Estrategias Manuales
 - Expresión diferencial entre clusters
 - Análisis en base a literatura
 - Anotación de tipos celulares
- Estrategias automatizadas
 - Basadas en correlación con set de datos previo
 - Basados en machine learning con multiples set de datos previos



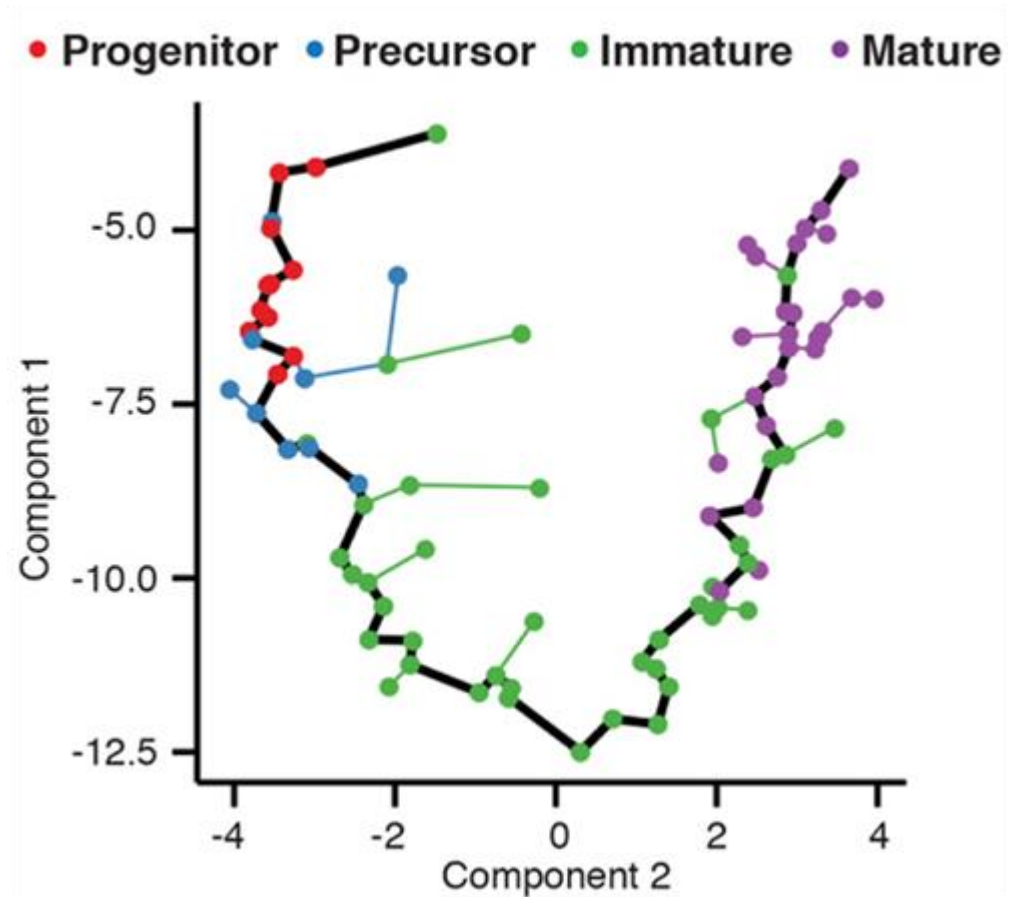
ANÁLISIS DE EXPRESIÓN DIFERENCIAL

- Entre clusters (generalmente para identificar tipos celulares) o entre condiciones (ej: sano vs enfermedad)
- Células vs Células (Wilcoxon Rank sum Test) o Pseudobulk (DESeq2, EdgeR)



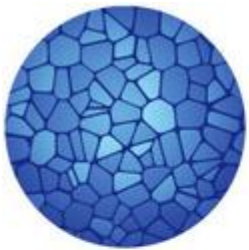
ANÁLISIS DE TRAYECTORIAS CELULARES

- Detección de tipos celulares en Desarrollo
- Análisis pseudotemporal permite ordenar las células en diferentes estadios
- Análisis de destino celular

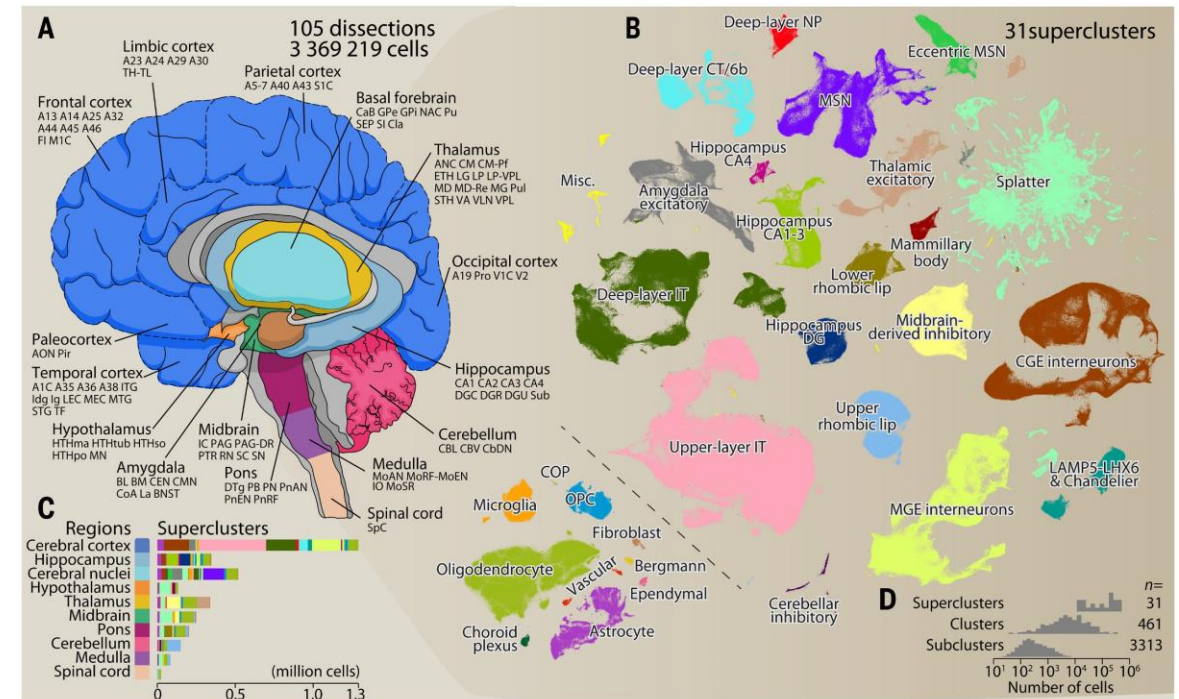
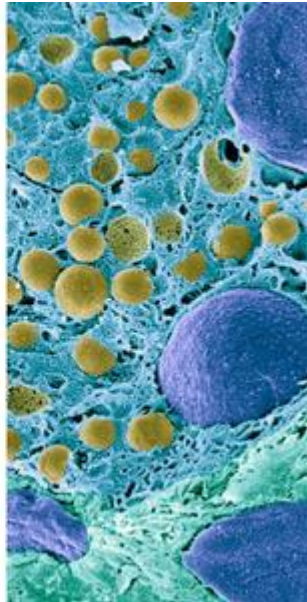
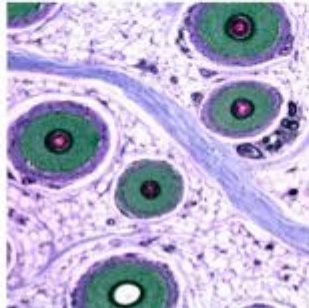


APLICACIONES

- Grandes proyectos para comprender y describir tipos y estados celulares
 - Human Cell Atlas
 - BRAIN initiative (BICCN)
 - Malaria Cell Atlas



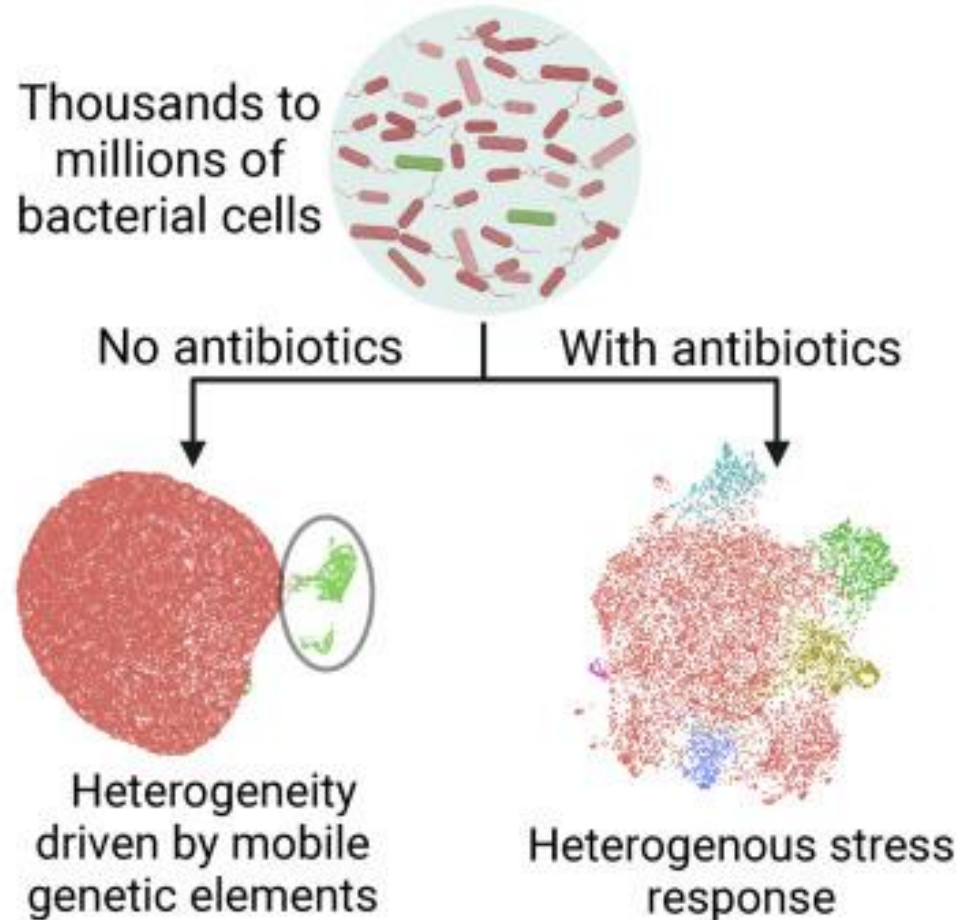
**HUMAN
CELL
ATLAS**



APLICACIONES

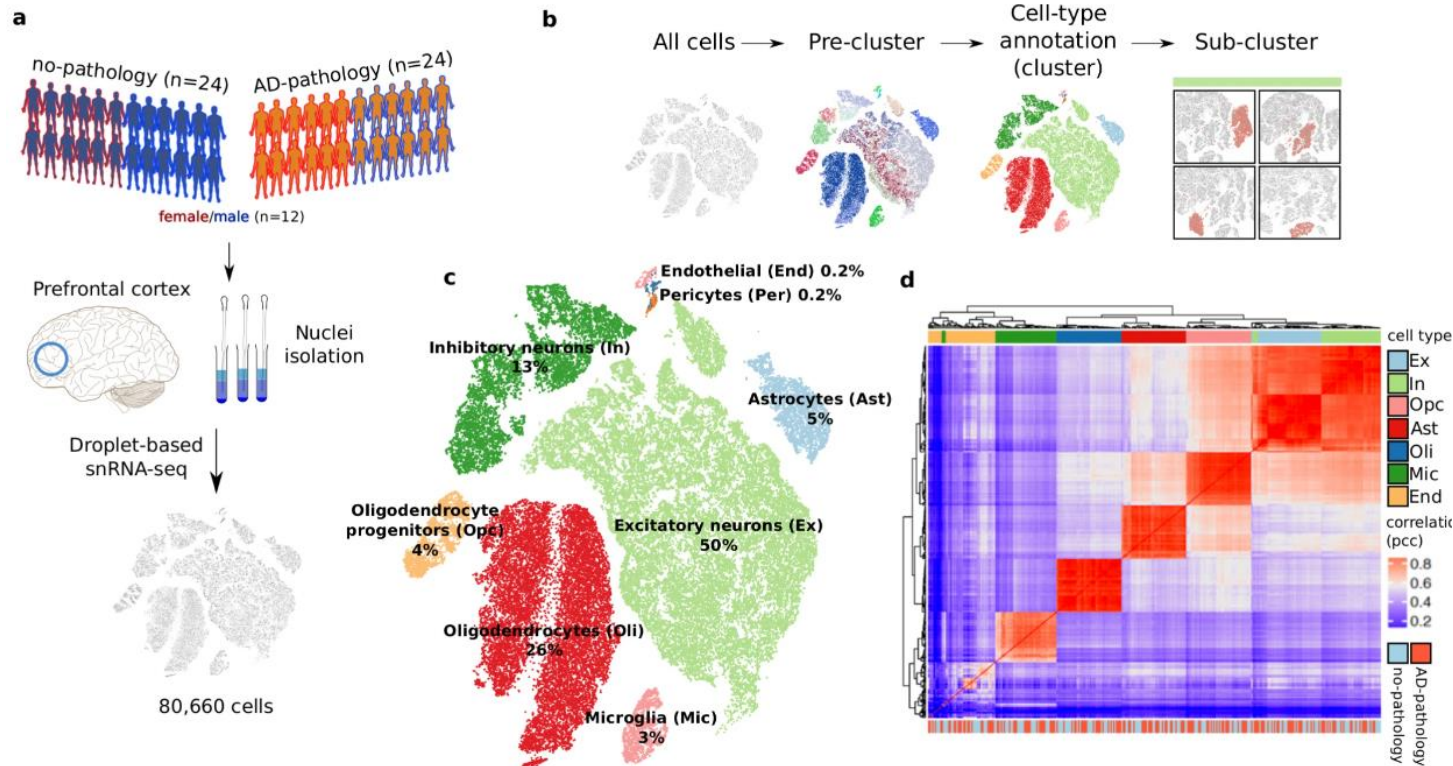
Detección de respuesta de bacterias a tratamiento con antibióticos

BacDrop: bacterial single-cell RNA-seq



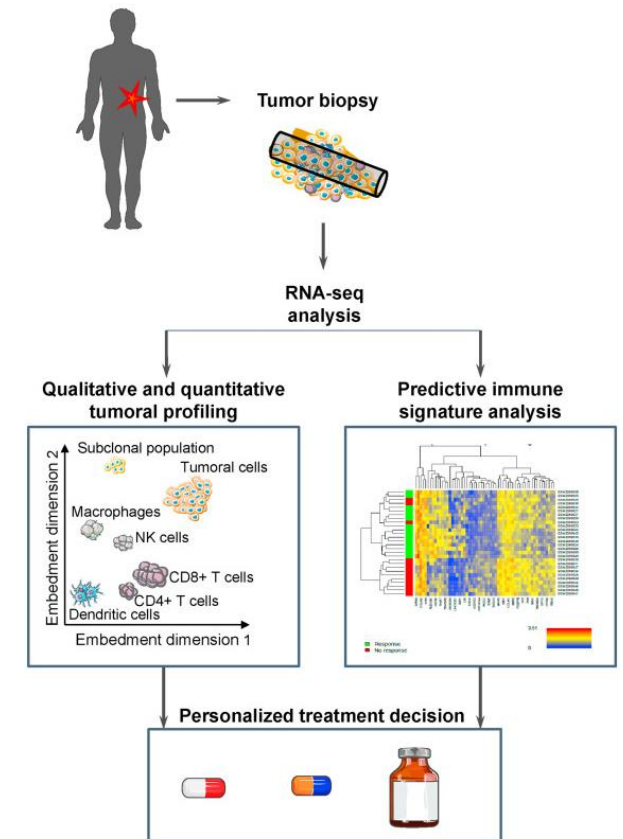
ESTUDIO DE ENFERMEDADES

Tipos celulares y genes alterados en Alzheimer's



Mathys et al, 2019

Heterogeneidad en Tumores y perspectivas de aplicaciones clínicas



Kuksin et al, 2021



OTRAS TÉCNICAS DE CÉLULA ÚNICA ESTÁN SIENDO DESARROLLADAS...

- scATACseq
- Patch-seq
- CITE-Seq (scRNAseq + proteínas específicas de superficie)
- Transcriptómica espacial
- Traductómica de célula única
- Traductómica Espacial
- Proteómica de célula única

Desafío: Integración multiómica de célula única.



TRANSCRIPTÓMICA

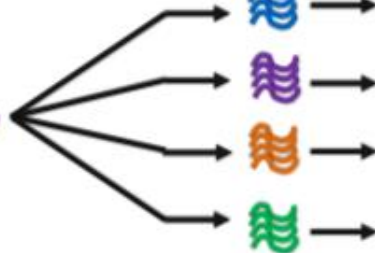
Bulk RNAseq



Abundancia
promedio de
ARNs

DE CÉLULA ÚNICA

scRNAseq



Abundancia de ARN Tipo Celular 1

Abundancia de ARN Tipo Celular 2

Abundancia de ARN Tipo Celular 3

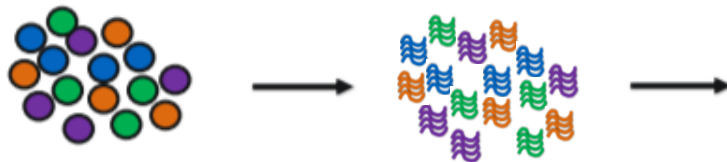
Abundancia de ARN Tipo Celular 4



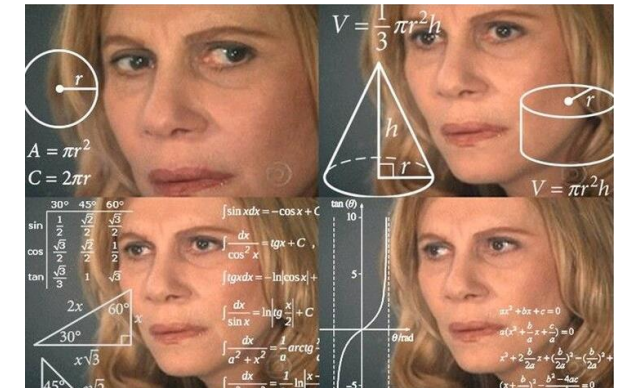
DÓNDE SE LOCALIZAN ESOS TRANSCRIPTOS?

TRANSCRIPTÓMICA ESPACIAL

Transcriptómica espacial



Abundancia de cada ARN con coordenadas de ubicación

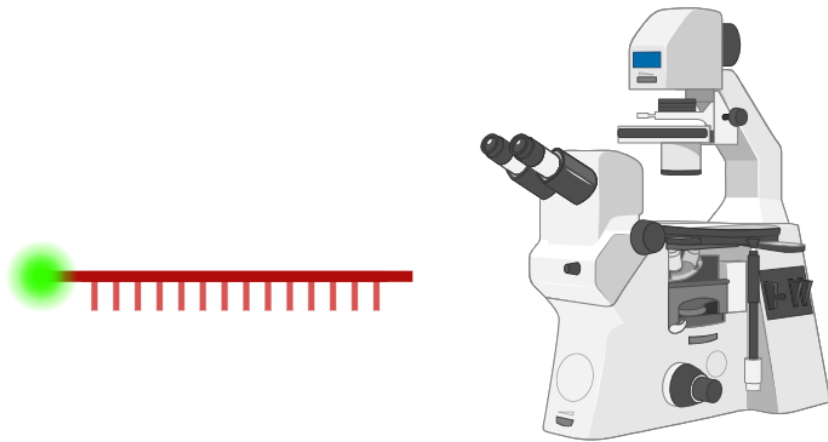


Es una FISH?



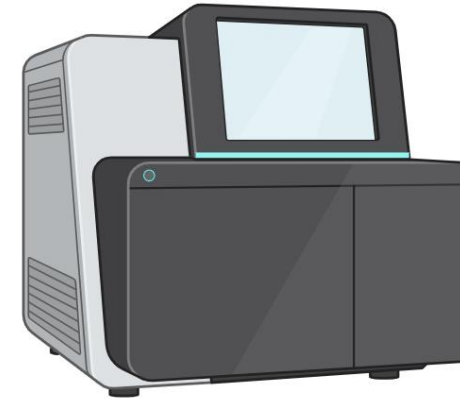
ESTRATEGIAS PARA ANALIZAR TRANSCRIPTOS...

- Métodos basados en hibridación de sondas.



Problema métodos clásicos:
No son ómicos (pocos ARNm en simultáneo)

- Métodos basados en secuenciación masiva.



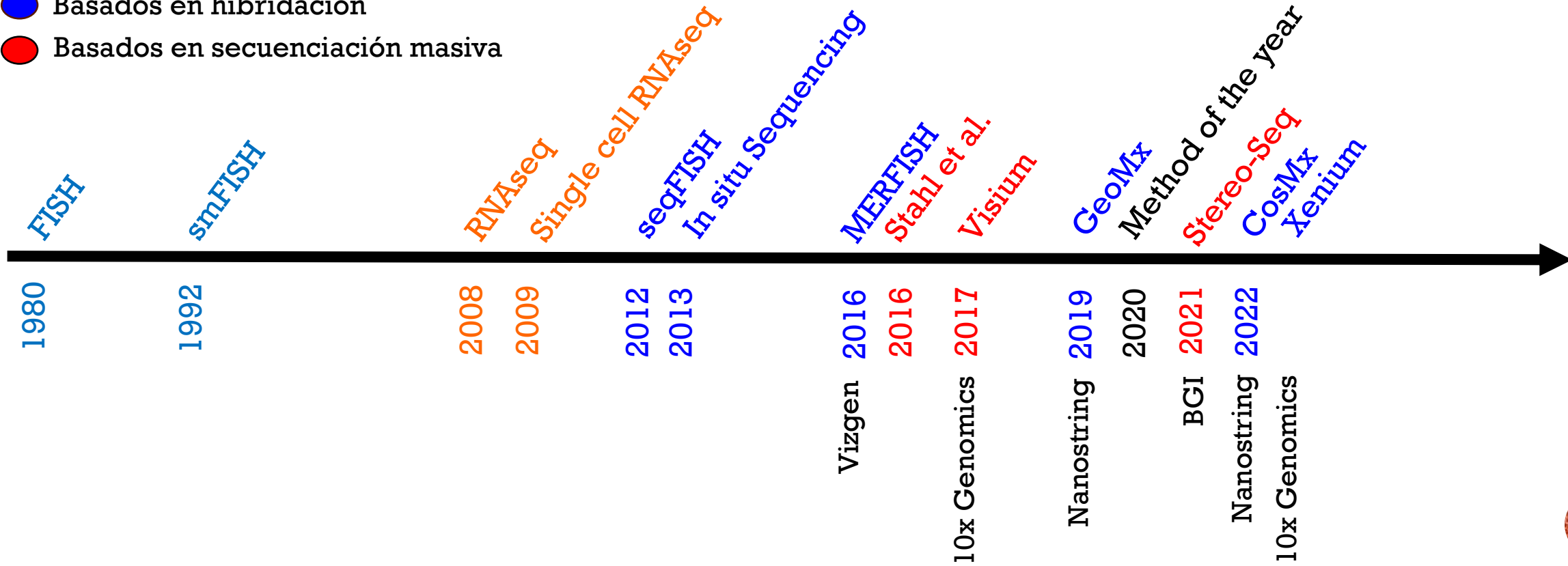
Problema métodos clásicos:
No dan información posicional de los ARNm



BREVE HISTORIA...

● Basados en hibridación

● Basados en secuenciación masiva



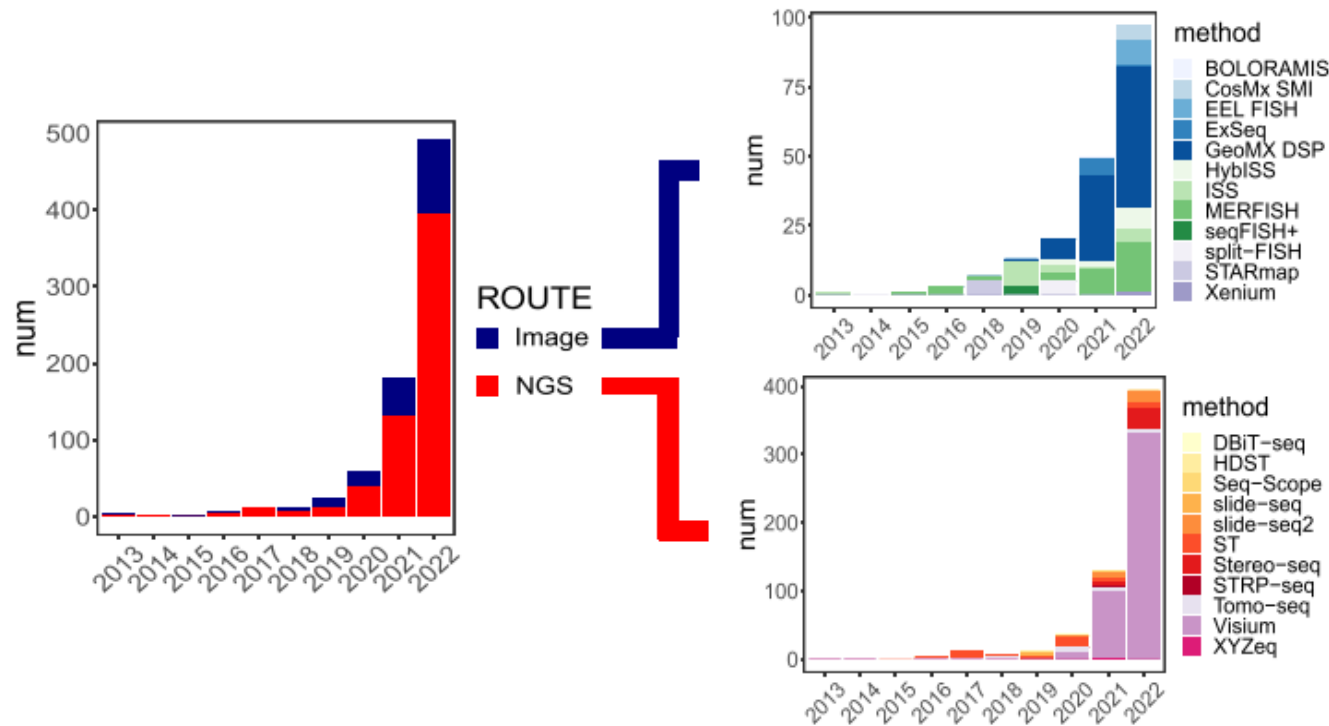
ELECCIÓN DE UN MÉTODO ADECUADO

- Tipo de muestra a analizar:
 - Células eucariotas o procariotas
 - Abundancia de ARN esperada
- Objetivo planteado
 - Necesidad de resolución? (región, célula única, subcelular)
 - Estudio de ARNm target o global? (paneles vs secuenciación)
 - Análisis multiómico (co-detección de proteínas, traductómica espacial)
- Posibilidades económicas
- Disponibilidad de equipos comerciales

Es importante conocer los métodos disponibles!



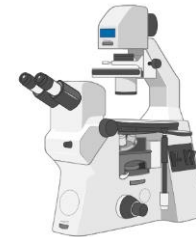
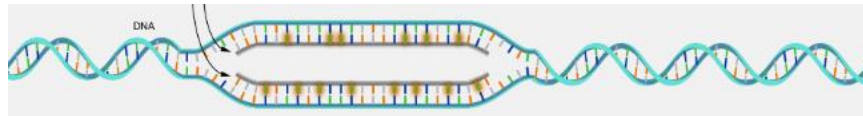
TÉCNICAS MÁS ADOPTADAS POR LA COMUNIDAD CIENTÍFICA



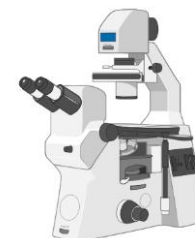
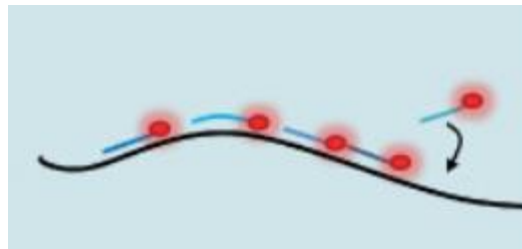
TÉCNICAS BASADAS EN HIBRIDACIÓN

Técnicas precursoras:

- Hibridización in situ fluorescente (FISH)
 - Técnica inicialmente aplicada para estudiar localización de ADN y posteriormente adaptada para detectar ARNm
 - Inicialmente era poco sensible y poco precisa para cuantificar.

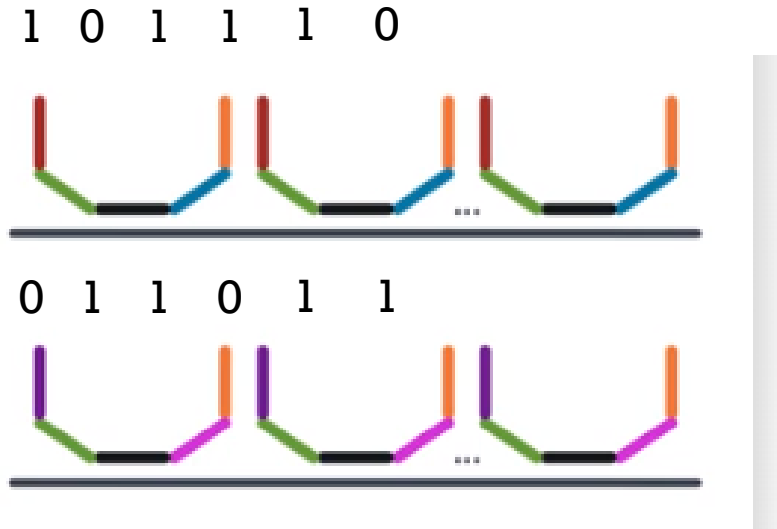


- smFISH
 - Serie de técnicas que mejoran la sensibilidad de la FISH a molécula única y aumenta considerablemente la precisión para cuantificar ARN. La clave es amplificar la señal.



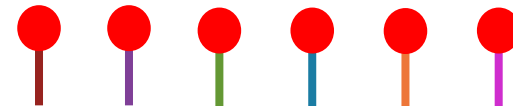
MERFISH

- Segunda técnica más utilizada actualmente dentro de los métodos basados en hibridación
- Multiplexed Error Robust Fluorescence In Situ Hybridization



Hibridación de sonda que contiene barcode.

Múltiples rondas de hibridación con secuencias que decodifican los barcode



ERROR ROBUST

- La clave de este método es que es muy robusta frente a errores por su diseño.
- En 16 rondas de hibridación se podrían detectar 65000 ARNm
- Solo una pequeña porción de las combinaciones posibles tiene significado.
- El diseño clásico incluye 4 señales por barcode. 12 rondas sin señal y 4 con señal



En el artículo original en 16 rondas de hibridación se detecta y cuantifica 140 ARNm con 80% de eficiencia de detección.



M

Multiplexed

Paneles de oligos que hibridan contra ARNm de interés y contiene cada uno un barcode diferente.

Normalmente se diseñan paneles para detectar hasta 500 genes.

Se ha utilizado para detectar 10000 genes

ER

Error-Robust

Probabilidad baja de error de identificación.

Alta eficiencia de detección.

FISH

FISH

Consta de multiples rondas de hibridación tipo FISH.

A priori el principal requerimiento de equipos sería un microscopio de epifluorescencia.

Requiere mucha manipulación para todas las rondas de hibridación.



vizgen MERSCOPE

- La empresa Vizgen ofrece un equipo especialmente diseñado para estos experimentos.
- Precio: U\$\$285.000
- Completamente automatizado (se evita la manipulación de las múltiples rondas de hibridación). Microscopio con sistema de microfluidos incorporado.
- Ofrecen paneles pre-diseñados o a elección del investigador.



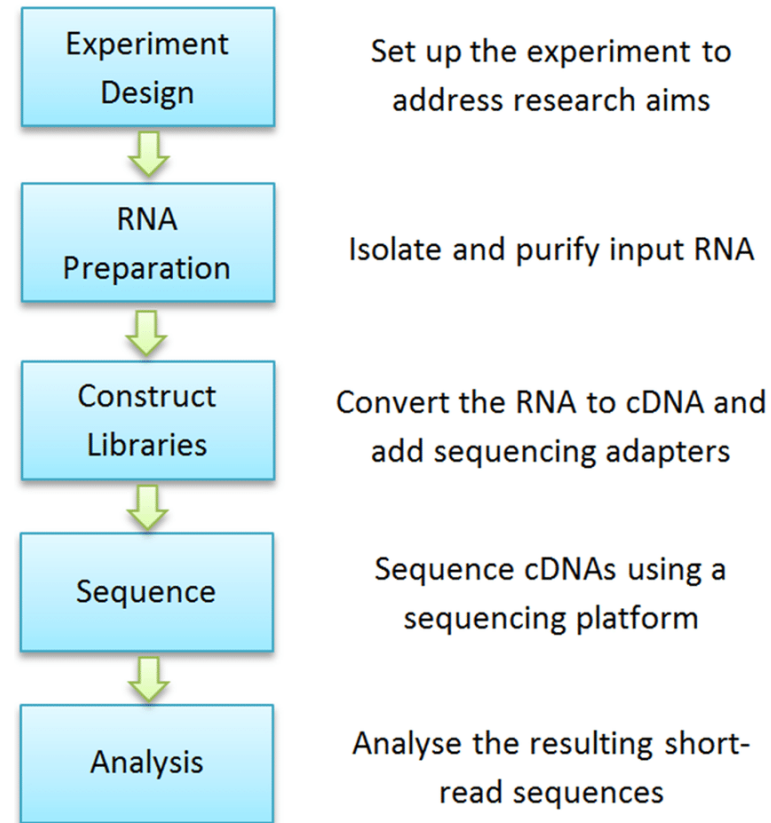
MERFISH

- Automatización: MERSCOPE (U\$\$285.000)
- Resolución: Subcelular - 200nm
- Co-detección: Proteínas, marcadores fluorescentes
- Sensibilidad: Alta
- Necesidad de paneles (Hasta 500 ARNm)
- FOV: 100 mm²



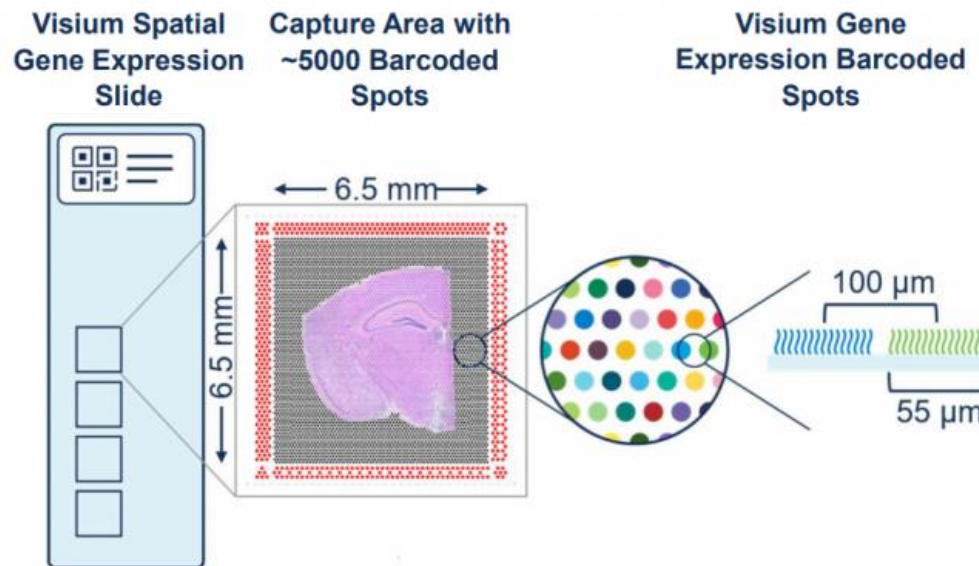
MÉTODOS BASADOS EN SECUENCIACIÓN

- Cómo agregarle información espacial al RNA-seq?



VISIUM (10X GENOMICS)

- Método más utilizado de transcriptómica espacial
- Lisis y transferencia de ARNs a sitios de captura
- La resolución viene dada por el tamaño del sitio de captura.



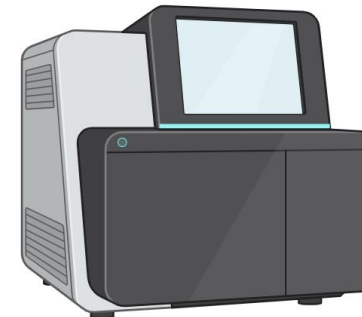
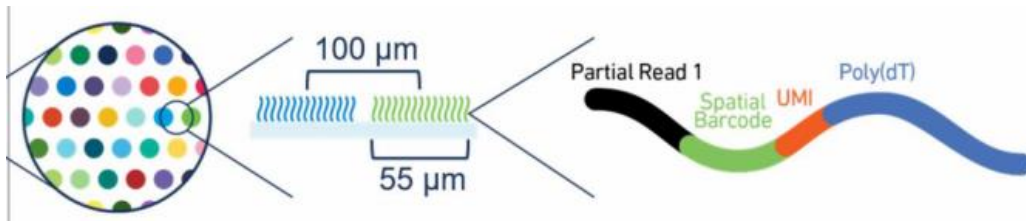
PASO A PASO...

- Se coloca la muestra sobre el área de captura de VISIUM
- Se realiza tinción (ej hematoxilina eosina, o inmunofluorescencia)
- Se toma fotografía de la muestra en microscopio.



PROTOCOLO SIMILAR A RNASEQ PERO SE COMIENZA SOBRE LOS SITIOS DE CAPTURA VISIUM

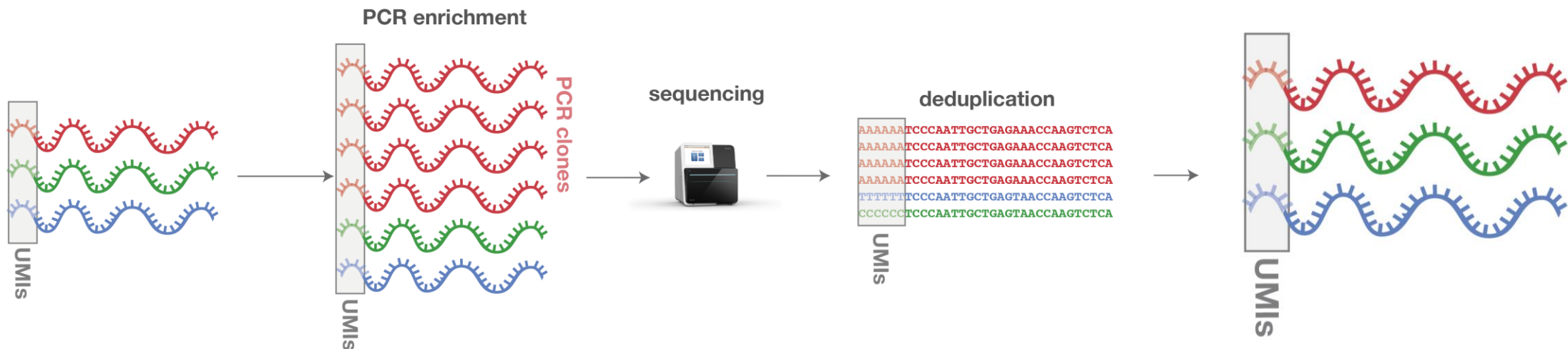
- Permeabilización (Los ARN presentes en la muestra se transfieren a los sitios de captura)
- Retrotranscripción, Síntesis de Segunda hebra. (Se utilizan los oligos de captura para estos pasos, lo que introduce un barcode espacial y UMIs)
- Amplificación del ADNc
- Preparación de librerías para secuenciación masiva
- Secuenciación (Se recomiendan mínimo 25.000 lecturas por sitio de captura)



PARA QUÉ SIRVEN LAS SECUENCIAS CONOCIDAS?

Estas secuencias se mantendrán en los fragmentos que se envían a secuenciar.

- Barcode – Único por cada sitio de captura: Marca los transcritos que vienen de un mismo sitio
- UMI – Único por cada oligonucleótido de captura: Marca cada ARNm presente en un sitio. Útil para corregir por sesgo de amplificación por PCR



VISIUM

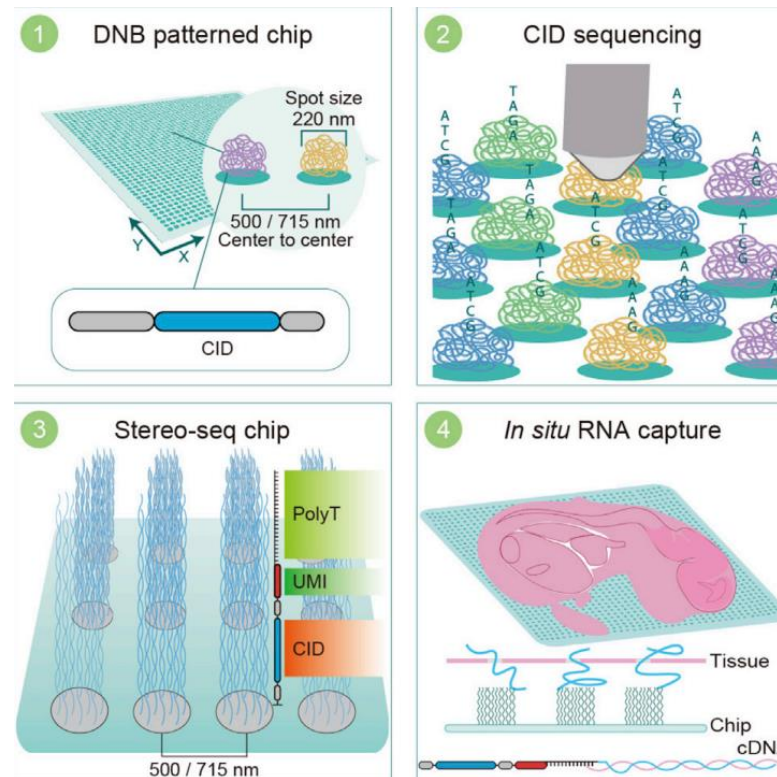
- Costo: U\$\$ 2000 por slide
- Resolución: Regional - 50 μ m
- Baja eficiencia de captura (superior al 7%)
- Captura de ARNm poliadenilados
- Recientemente salió:

Visium HD – Protocolo similar pero con sitios de captura de 2x2 μ m



STEREO-SEQ (BGI)

- Similar a Visium, pero con sitios de captura con DNA nanoballs (BGI)
- Los sitios de captura son más pequeños (220nm) y con distancia de 500nm al siguiente. Resolución subcelular!



STEREO-SEQ

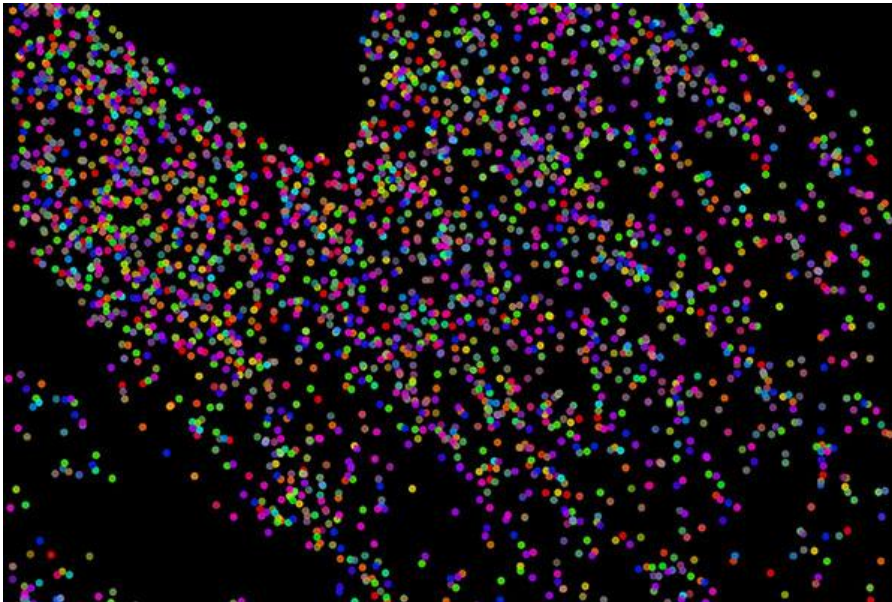
- Resolución: Subcelular - 220nm
- Baja eficiencia de detección (7 UMIs por sitio de captura en corteza cerebral de ratón) --- aprox: 1500 UMI por célula
- Captura de ARN poliadenilados



RESUMIENDO...

- Métodos basados en hibridación

Múltiples fotos



- Métodos basados en secuenciación

Múltiples secuencias

```

Y103.fastq
@Y103_1_HISEQ301.HN.WGKB.CX.X2.1101.6397.2187.1.N:0.CCAGTT orig_bc=AAAAAAAAAAAA new_bc=AAAAAAAAAAAA bc_diffs=0
TAGCTAGGGGTGCGAGCGTGTTCGGGAATTACTGGGCGTAAAGAGCGTCGTAGGCGGTTTGTCCGCGTGTCTGTGAATCCCGCAGCTCAACTCGGGGCTTCGACGGCGATACGG
CGAAACTGAGTATCGCAGCGAGTATCGAATATCTCTGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGCGGGTCTCTGGCGAGTAAGTACGCG
TGAGGAGCCAAAGCGTGGGTAGCGAACAGG
+
#####
@Y103_1_HISEQ301.HN.WGKB.CX.X2.1101.11507.2246.1.N:0.CCAGTT orig_bc=AAAAAAAAAAAA new_bc=AAAAAAAAAAAA bc_diffs=0
TAGCTAGGGGGCAAGCGTATTCGGGAATTACTGGGTGTAAAGGAGCGCTAGACGGCAGCGCAAGTCTGATGTGAAACCCAGGGCTTAACCTGGGAGTCATTGGAAACT
GTCTGCGTCAAGTATCGCGGACAGGTAGAGCGGAATCTTAGTGTAGCGGTGAAATGCGTAGATATTAGGAAGAACACCAAGTGGCGAAGCGGCTCTAGCAGCGGTAACTGACGC
TTGAGGTCCGAAAGCGTGGGAGCAACAGG
+
#####
@Y103_2_HISEQ301.HN.WGKB.CX.X2.1101.18481.2208.1.N:0.CCAGTT orig_bc=AAAAAAAAAAAA new_bc=AAAAAAAAAAAA bc_diffs=0
TAGCTAGGGTGGCAGCGTGTTCGGGAATTACTGGGCGTAAAGAGCGTCGTAGGCGGTTTGTCCGCGTGTCTGTGAATCCCGCAGCTCAACTCGGGGCTTCGACGGGATACGG
CGCAGACTCGAGTACTCGCAGGGGAGCGTAGCAATCTCTGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGCGGGTCTCTGGCGAGTAAGTACGCG
TGAGGAGCCAAAGCGTGGGTAGCGAACAGG
+
#####
@Y103_3_HISEQ301.HN.WGKB.CX.X2.1101.5935.2268.1.N:0.CCAGTT orig_bc=AAAAAAAAAAAA new_bc=AAAAAAAAAAAA bc_diffs=0
TAGCAAGGGGGGTGAGCGTGTTCGGGAATTACTGGGCGTAAAGCGCAGTAGGCGGATGTCAGTATAGAGGTGAAATCTGGAGCTCAACTCGGAACTGCCTTTAATACTG
CCAGTCTCGAGTCCGGAAGCGTAGTGAGTCAACTCTAGTGTAGAGGTGGAATTCGTAGATATTAGGAAGAACACCAAGTGGCGAAGCGGGTCTACTGGTCCGGTATGACGC
TGAGGTCCGAAAGCGTGGGAGCAACAGG
+
#####
@Y103_4_HISEQ301.HN.WGKB.CX.X2.1101.9217.2438.1.N:0.CCAGTT orig_bc=AAAAAAAAAAAA new_bc=AAAAAAAAAAAA bc_diffs=0
TAGCAAGGGTGCAGCGTATTCGGGAATTACTGGGCGTAAAGCGCGCTAGTGGTGTGTGAAGTGAAGTGAAGCCCGGGCTCAACTCGGGAAGTGCATCCAAACT
GGCAAGCTAGAGTAGGCGAGGGGTGGTGAATTCTCTGTGTAGCGGTGAAATGCGTAGATATTAGGAAGGAACACCAAGTGGCGAAGGGCGACCACTGGGCTAATCTGACA
CTGAGGTCCGAAAGCGTGGGAGCAACAGG
+
#####
@Y103_5_HISEQ301.HN.WGKB.CX.X2.1101.5325.2570.1.N:0.CCAGTT orig_bc=AAAAAAAAAAAA new_bc=AAAAAAAAAAAA bc_diffs=0
TAGCTAGGGTGGCAGCGTGTTCGGGAATTACTGGGCGTAAAGAGCTGTAGGCGGTTTGTCCGCGTGTCTGTGAATCCCGCAGCTCAACTCGGGGCTTCGACGGGATACGG
CGAAACTGAGTATCGCAGGGGAGCGTAGCAATCTCTGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGCGGGTCTCTGGCGAGTAAGTACGCG
TGAGGAGCCAAAGCGTGGGTAGCGAACAGG
+
#####

```

Los resultados de ambos tipos de protocolos son muy diferentes...



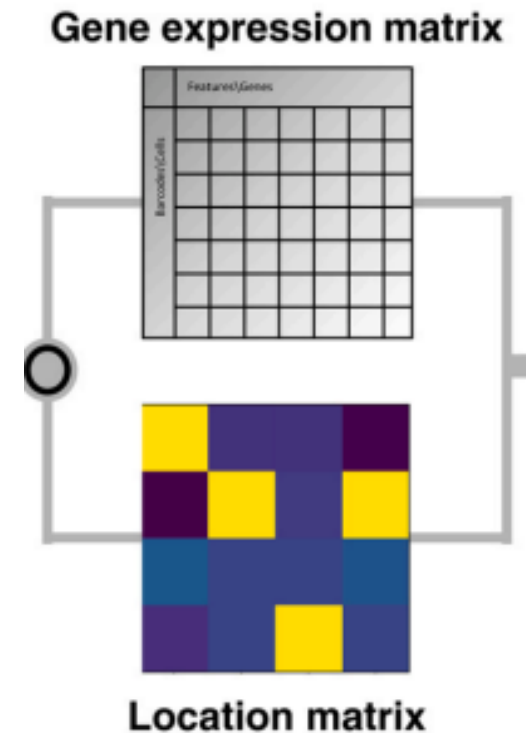
ANÁLISIS DE DATOS

- A pesar de que los resultados iniciales son muy diferentes, ambos tipos de métodos tienen pasos en común

Objetivo 1:

Generar matriz de expresión génica por posición

Generar matriz de coordenadas



CÓMO SE GENERAN ESTAS MATRICES?

- **Métodos basados en hibridación:**

- Identificación de coordenadas de señal en cada ronda de imagen.
- Decodificación de los barcodes.
- Generación de matrices de conteos por localización y matriz de coordenadas.
- Si además se quiere información que relacione posiciones con células (información de célula única):
Segmentación de células

Programas específicos permiten realizar estas tareas:

BarDensr
ISTDECO
Sparcle
Spa2vec

Y mas...

Programas gráficos de pago (ofrecidos por empresas que venden equipos y kits)



CÓMO SE GENERAN ESTAS MATRICES?

- **Métodos basados en secuenciación:**

- Debe combinar la imagen que se obtiene con tinción (previa al secuenciado) con las secuencias que se obtienen

- **Análisis de secuencias:**

- Análisis de calidad de las secuencias
- Alineamiento de secuencias frente a Genoma de referencia con herramienta que identifique los diferentes barcodes presentes (barcode identificador de posición y UMI).
- Conteo de lecturas asociadas a un mismo barcode identificador de posición (colapsando UMIs iguales). Obtención de Matriz de conteos

- **Análisis de imagen:**

- Dividir la imagen adquirida en cuadrantes y generar matriz de coordenadas.



Programas específicos que permiten realizar estas tareas:

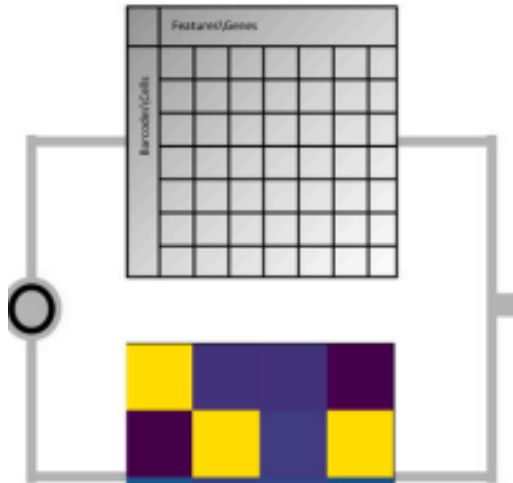
Spacemake
SpaceRanger (Visium)

Alternativas: Pipeline con programas clásicos de RNAseq

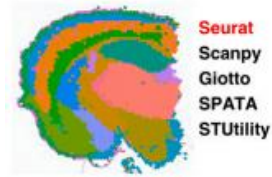


QUÉ INFORMACIÓN PODEMOS EXTRAER?

Gene expression matrix



Location matrix



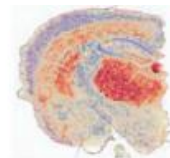
Seurat
Scanpy
Giotto
SPATA
STUtility

Dimensionality reduction and clustering



Seurat
cell2location
stereoscope
DestVI
AdRoit

Spatially cell-type annotation



Seurat
Scanpy
SpatialDE
SPARK
GPcounts

Spatially variable genes



CellTrek
SpaOTsc
DIALOGUE
SVCA

Cell-cell interaction



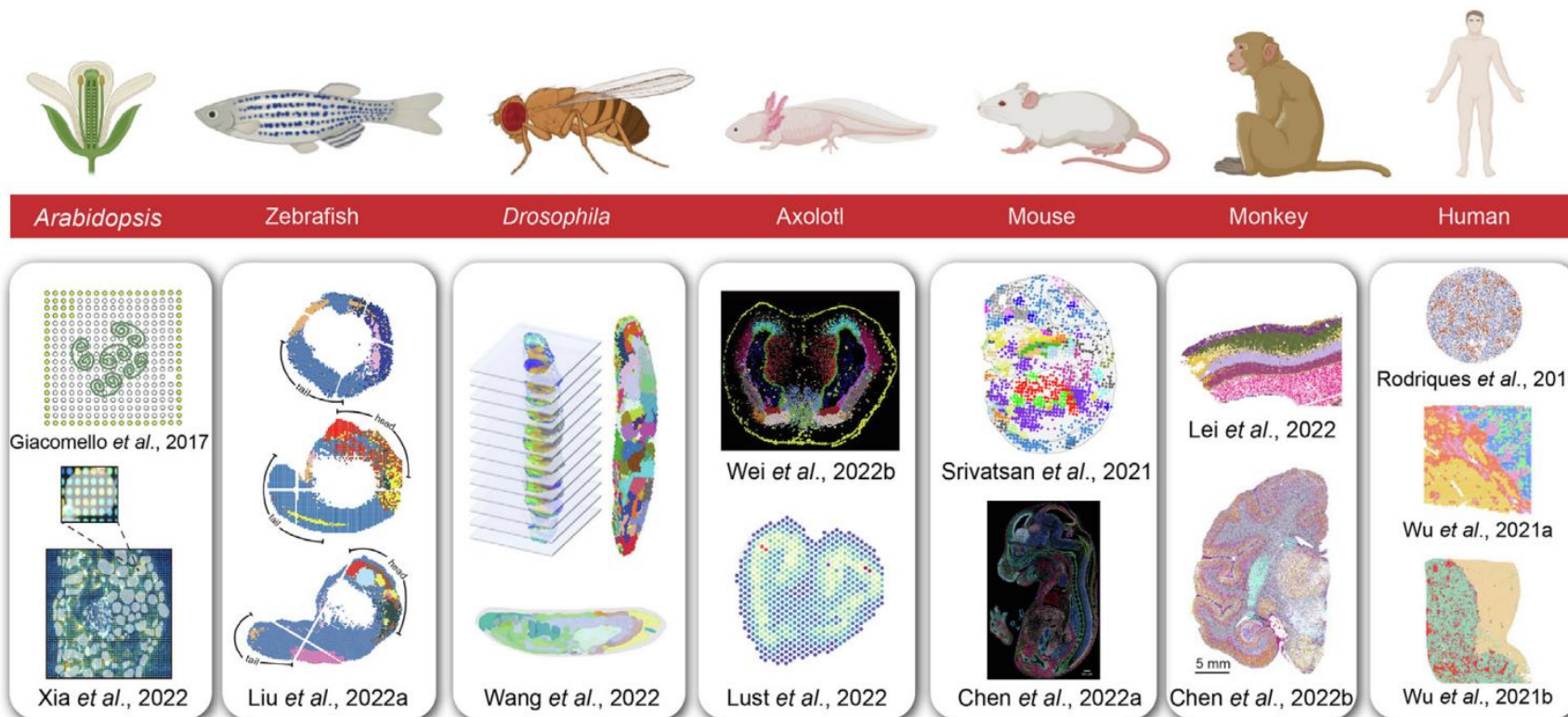
stLearn
SPATA

Spatial trajectory



PARA QUÉ SE HAN APLICADO ESTOS MÉTODOS HASTA AHORA?

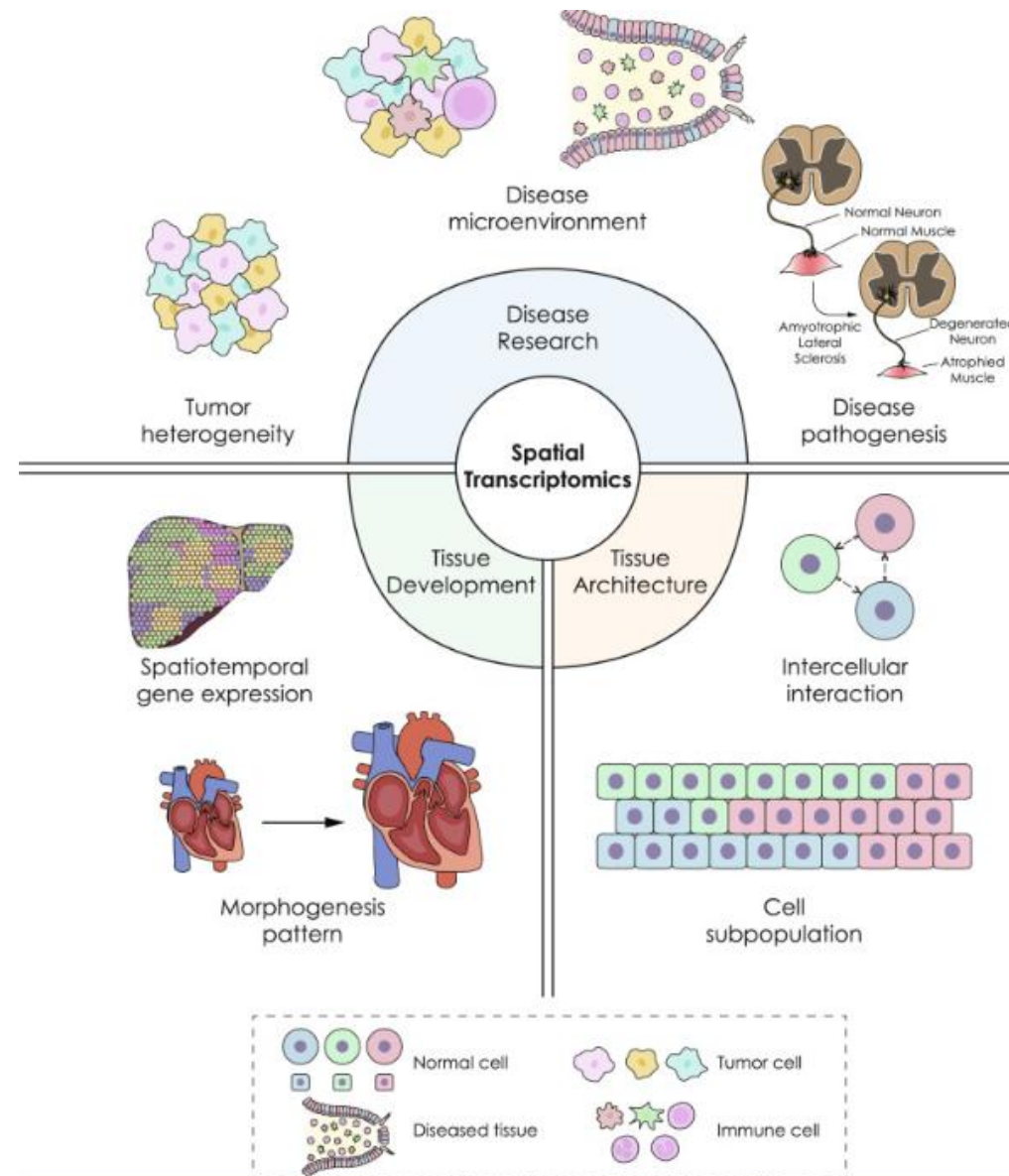
- Generación de Atlas Celular: Descripción de tipos celulares y localización de los mismos en diferentes tejidos



Cheng *et al.*, 2022

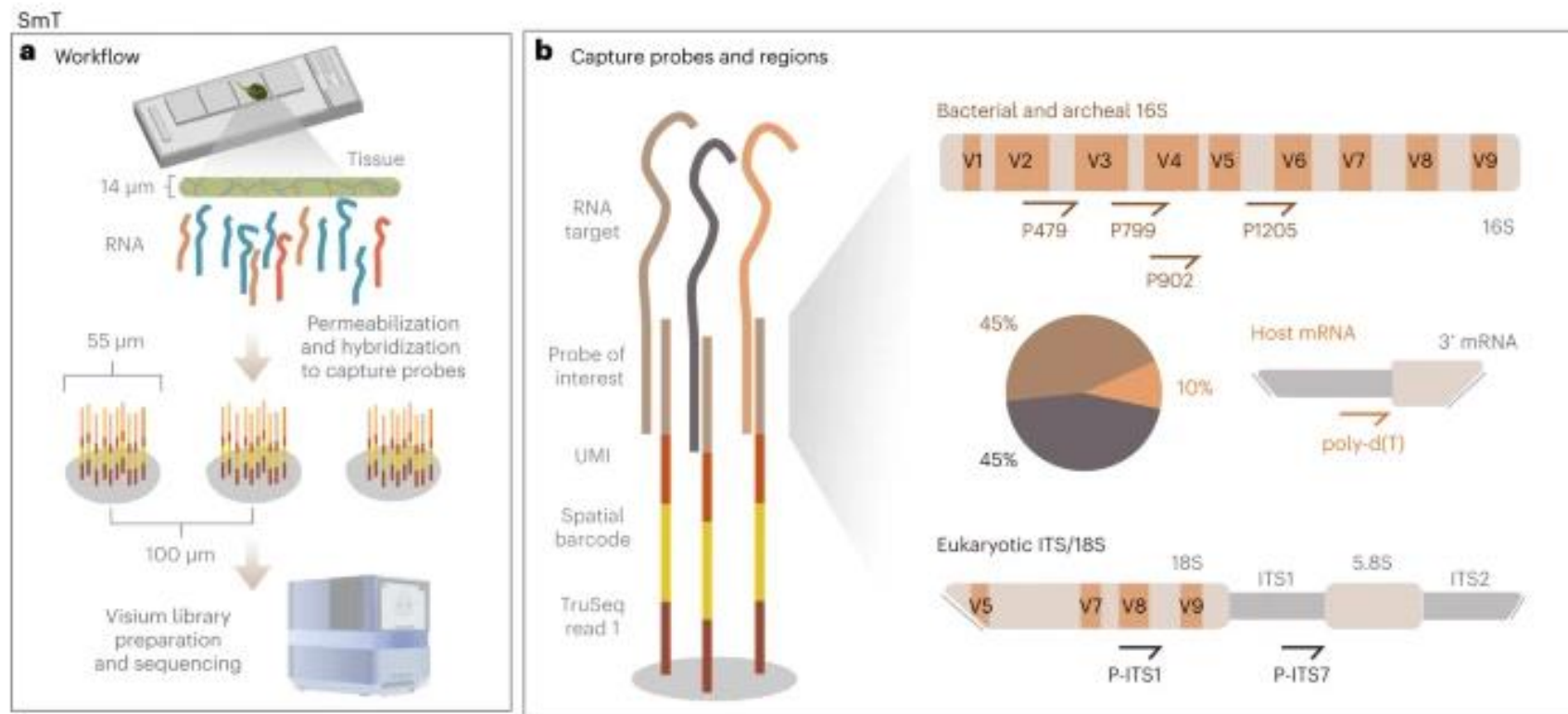


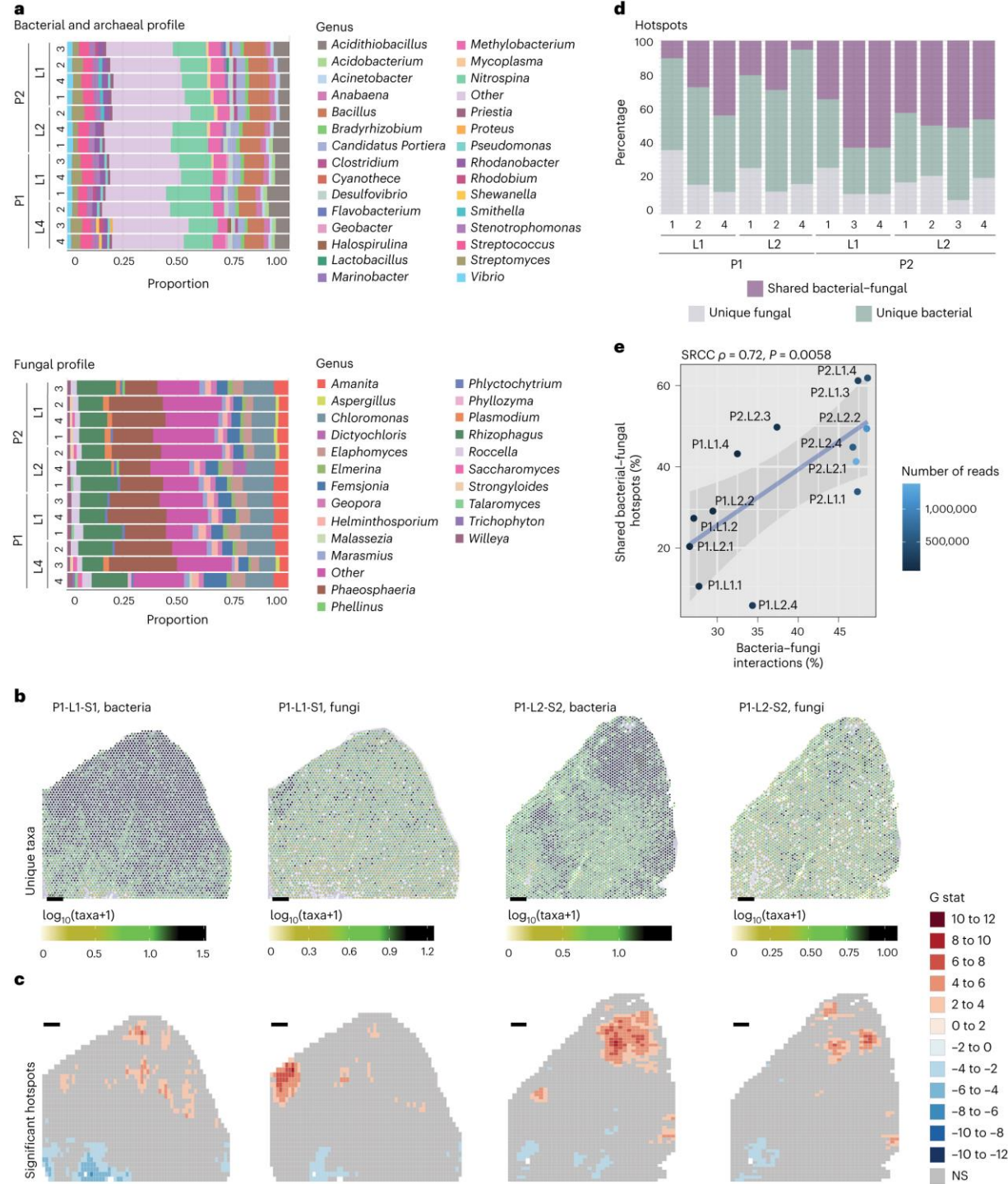
- Investigación de enfermedades
- Desarrollo celular
- Interacciones celulares



Y TAMBIÉN SPATIAL METATRANSCRIPTOMICS

Fig. 1: Overview of the method.





**LAS APLICACIONES SON MUCHISIMAS Y SE
ESPERA QUE LAS TÉCNICAS MEJOREN EN LOS
PRÓXIMOS AÑOS...**

DÓNDE SE APLICA SPATIAL TRANSCRIPTOMICS?



