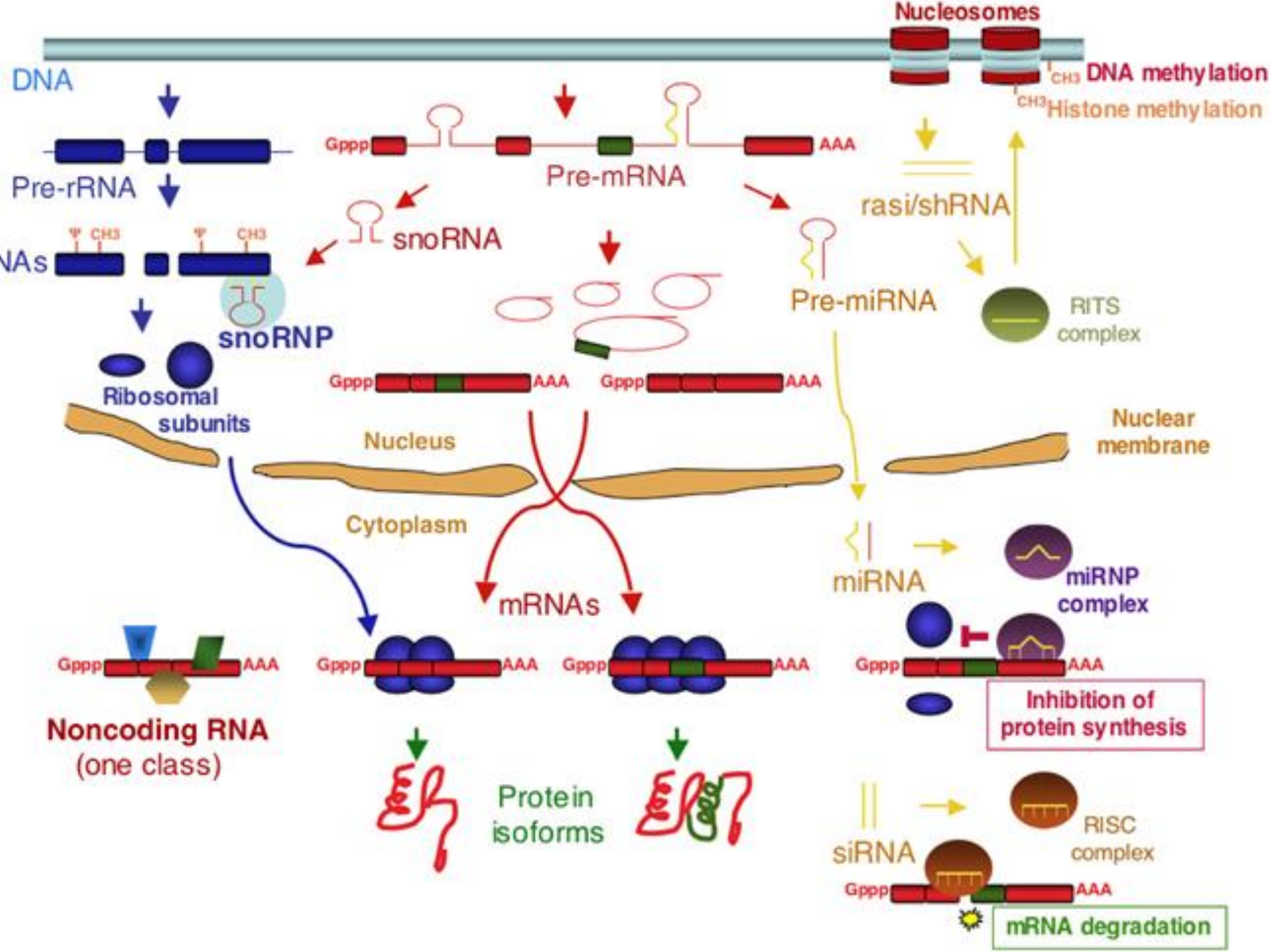
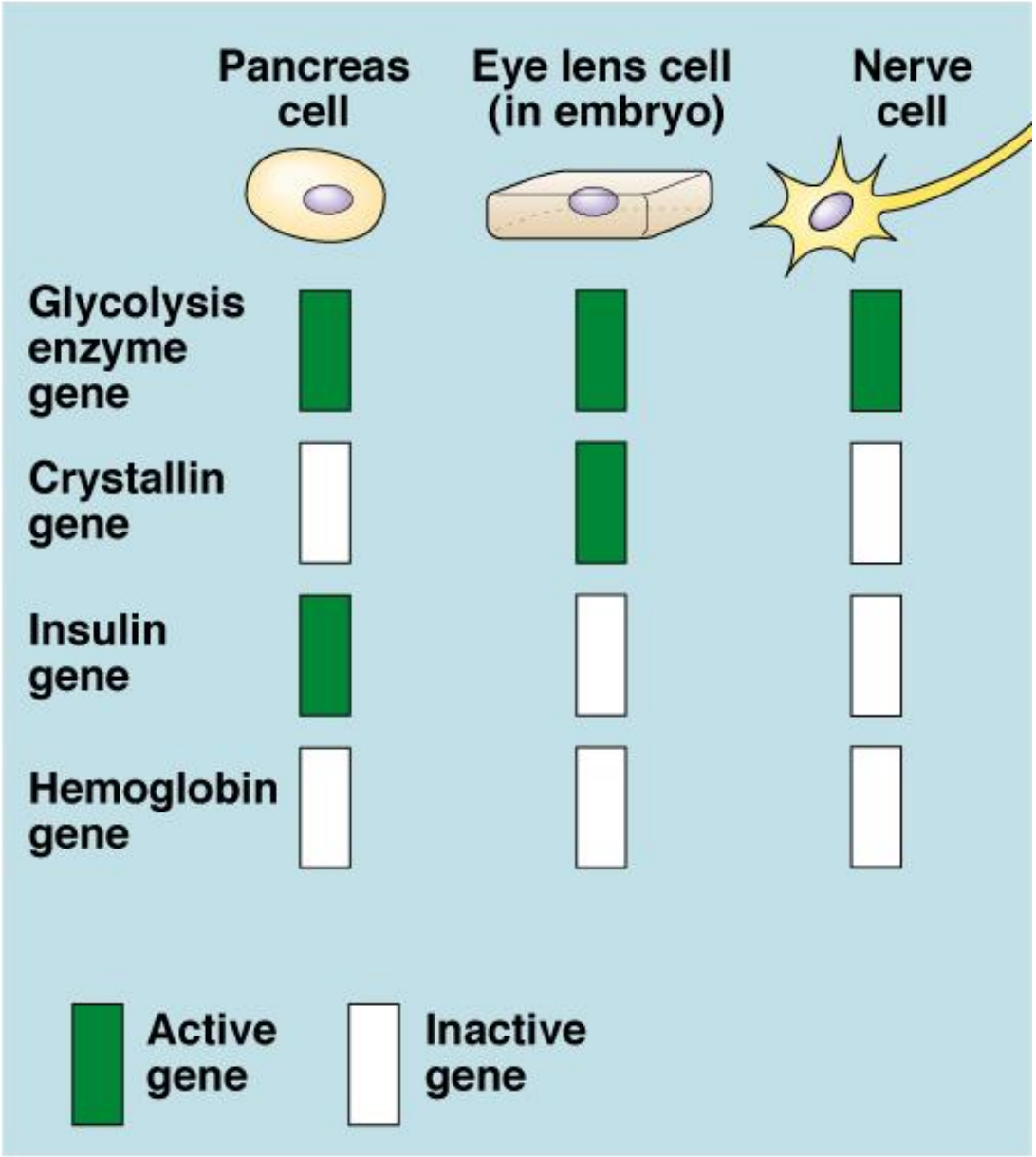
Estudio de Transcriptomas

Lo que el genoma tiene para decir


Nucleosomes
DNA
CH3 DNA methylation
CH3 Histone methylation
Pre-rRNA
Gppp
AAA
Pre-mRNA
rasi/shRNA
snoRNA
Pre-miRNA
RITS complex
snoRNP
Ribosomal subunits
Gppp
AAA
Gppp
AAA
Nucleus
Nuclear membrane
Cytoplasm
miRNA
miRNP complex
mRNAs
Gppp
AAA
Gppp
AAA
Gppp
AAA
Gppp
AAA
Inhibition of protein synthesis
Noncoding RNA
(one class)
Protein isoforms
RISC complex
siRNA
Gppp
AAA
mRNA degradation

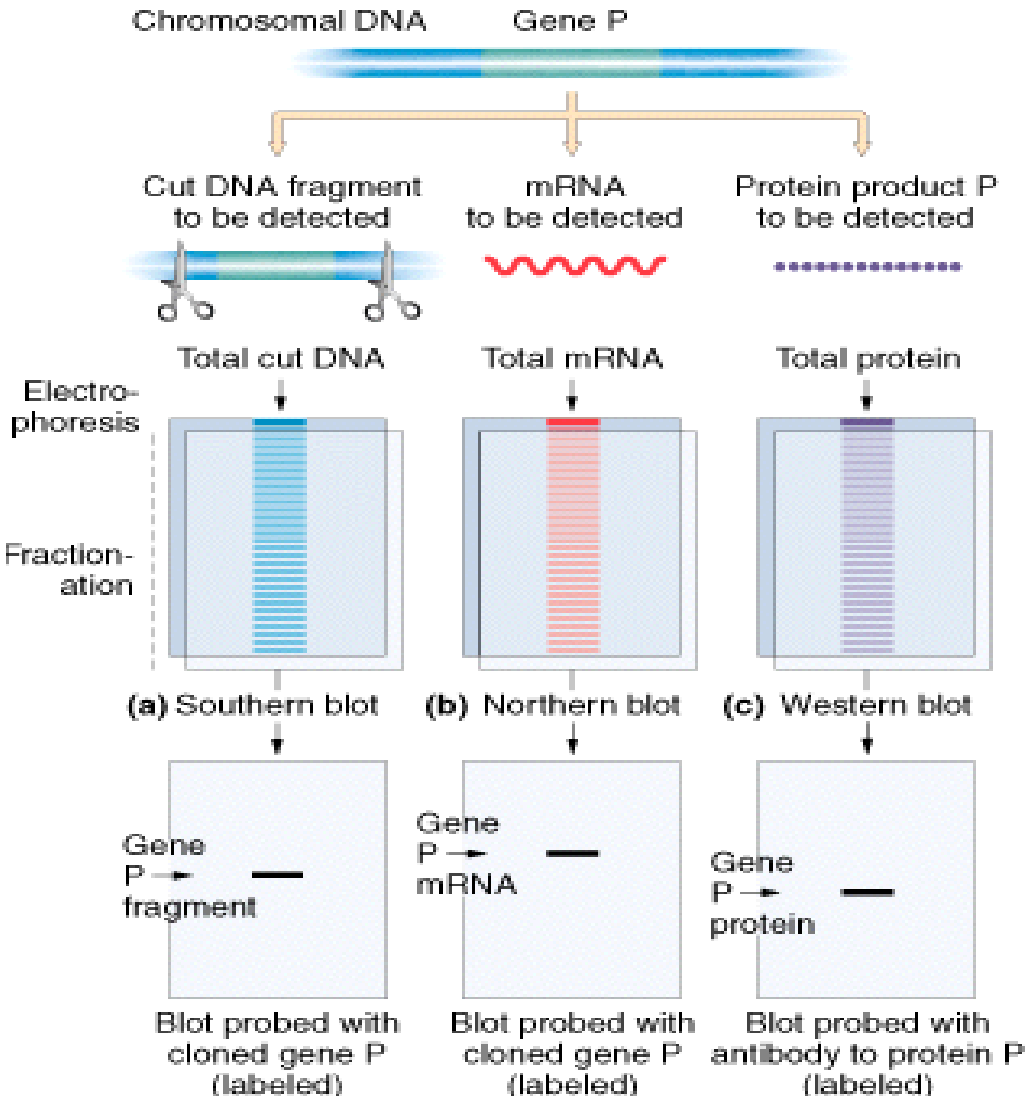
Expresión diferencial

- No todos los genes se expresan
  - existen distintos perfiles de ARNs y proteínas
- La expresión diferencial:
  - durante el desarrollo y diferenciación celular
  - en la homeostasis



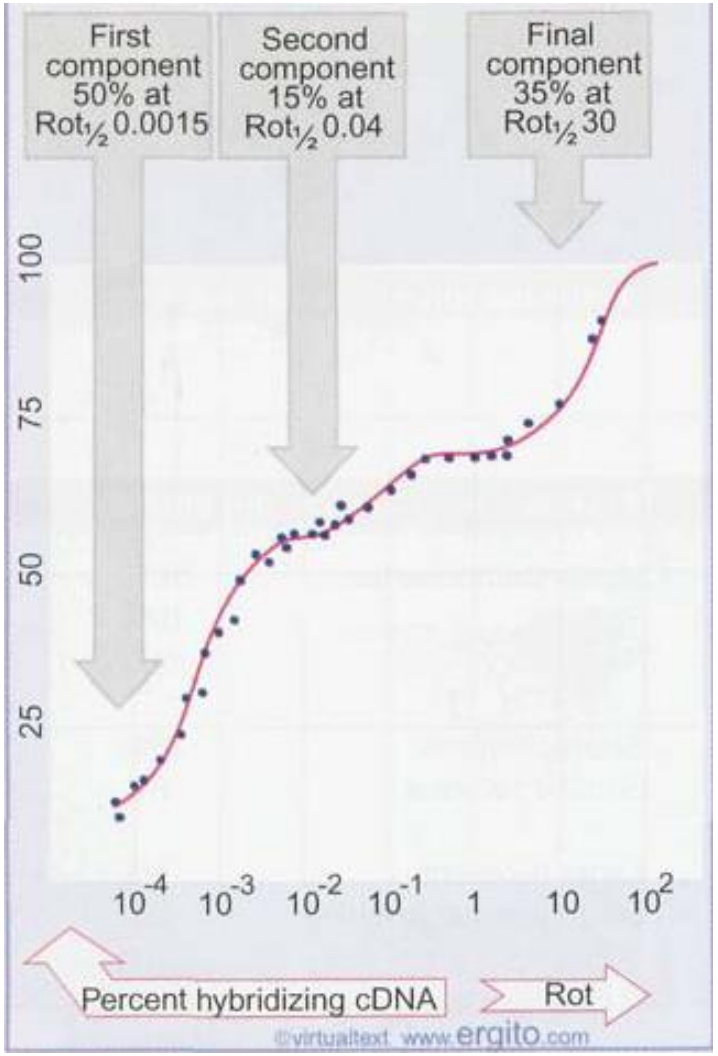
Estudios clásicos de expresión:

- Se focalizan en un número limitado de genes
  - *Northern blot*
  - RT PCR.

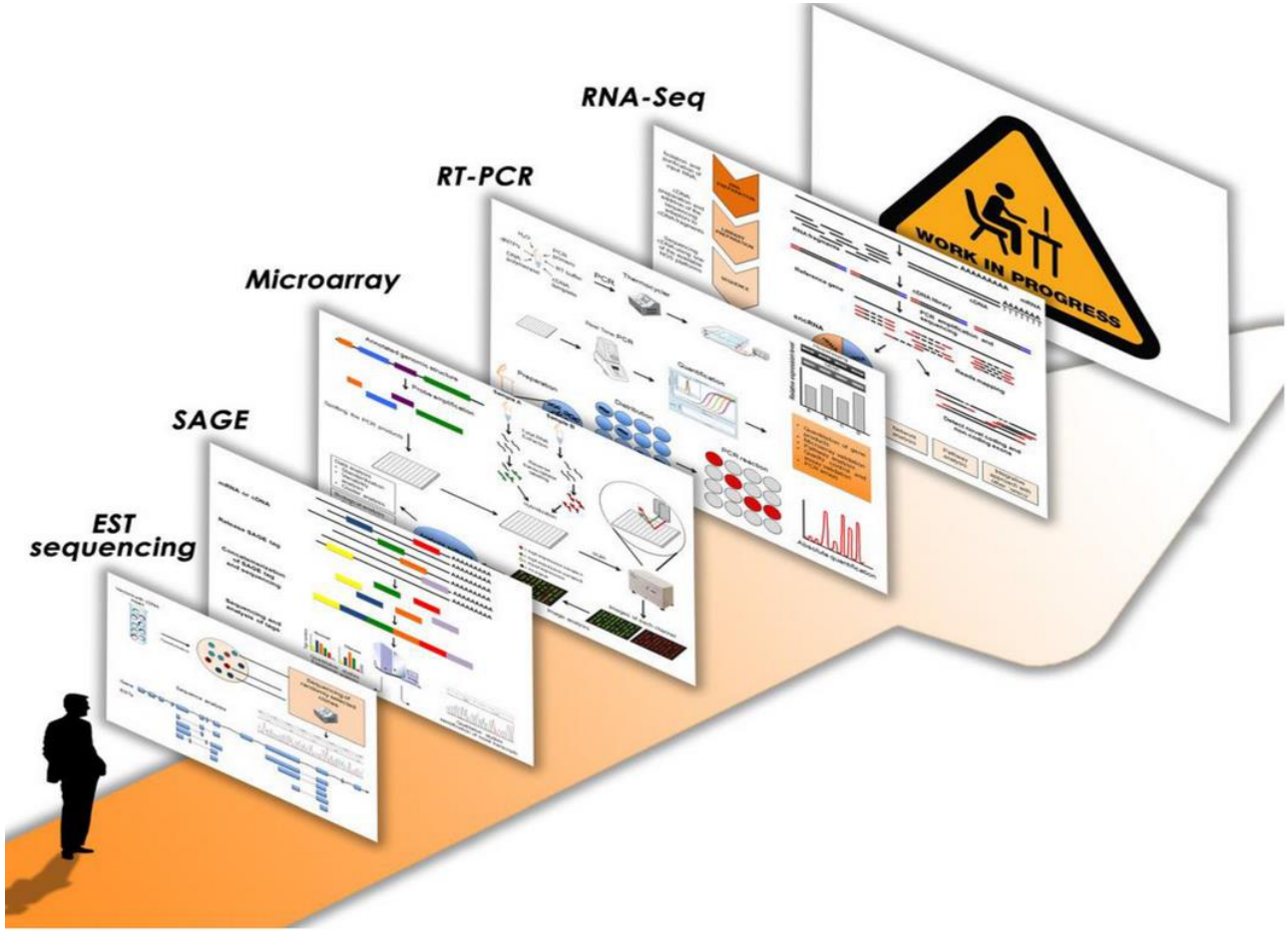


Curvas Rot

- Los ARN están presentes a diferentes abundancias
- Tipos de ARN (Okamuro & Goldberg)
  - Superabundantes
    - 15-90% de la masa de mRNA
    - <10 transcriptos diferentes
    - >5000 moléculas por célula
  - Abundantes
    - 50-75% de la masa de mRNA
    - ~200-1000 transcriptos diferentes (5% de la diversidad de mensajeros)
    - 500-2500 moléculas por célula
  - Raros
    - <25% de la masa
    - 95% de la diversidad
    - 1-10 por célula



Avances metodológicos

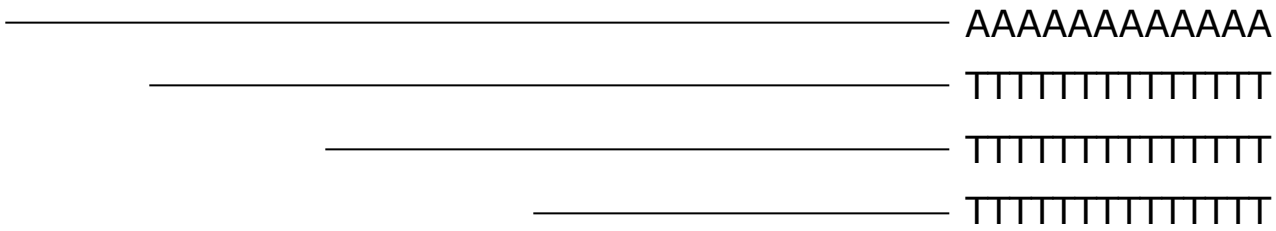


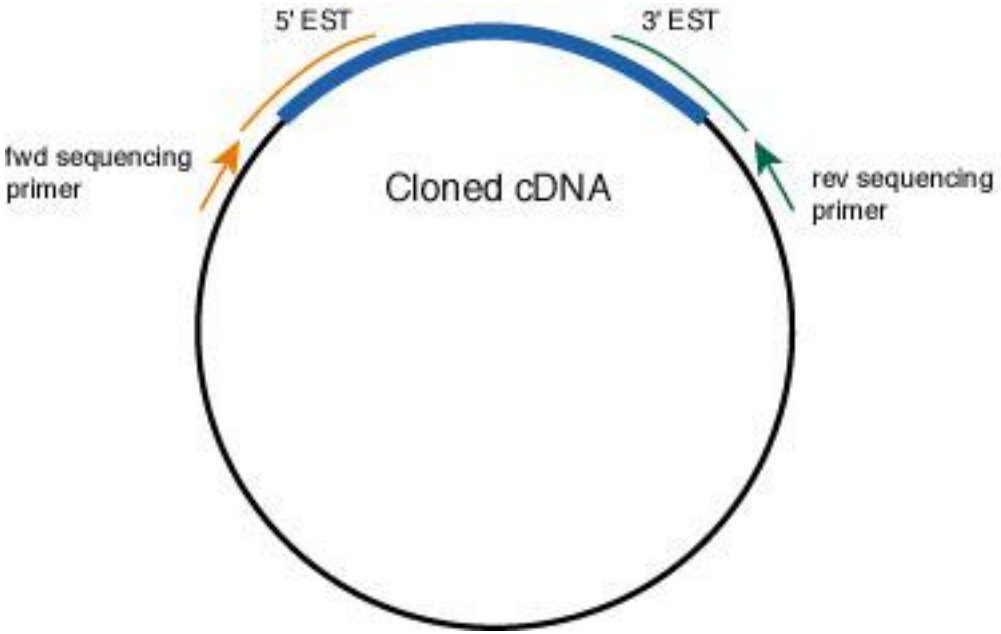
Expressed Sequence Tag (EST)

- Adams M. *et al*. Science 252:1651-1656 en 1991
- Lecturas de secuencias cortas, al azar y de un único pasaje por clon derivadas de bibliotecas de cDNA
- Alternativa mas barata al secuenciado genómico
- Propensas a errores y requieren un procesamiento bioinformático intensivo para dar información global del transcriptoma.
- Descubrimiento de genes nuevos, y determinación de estructura génica
- Se pueden usar para identificar el resto del gen en el genoma (anotación de genes nuevos)
- Pueden ser marcadas para localizarse en el cromosoma (marcadores físicos)(FISH).
- Verificación de las secuencias UTRs

Producción de ESTs

cDNAs parciales



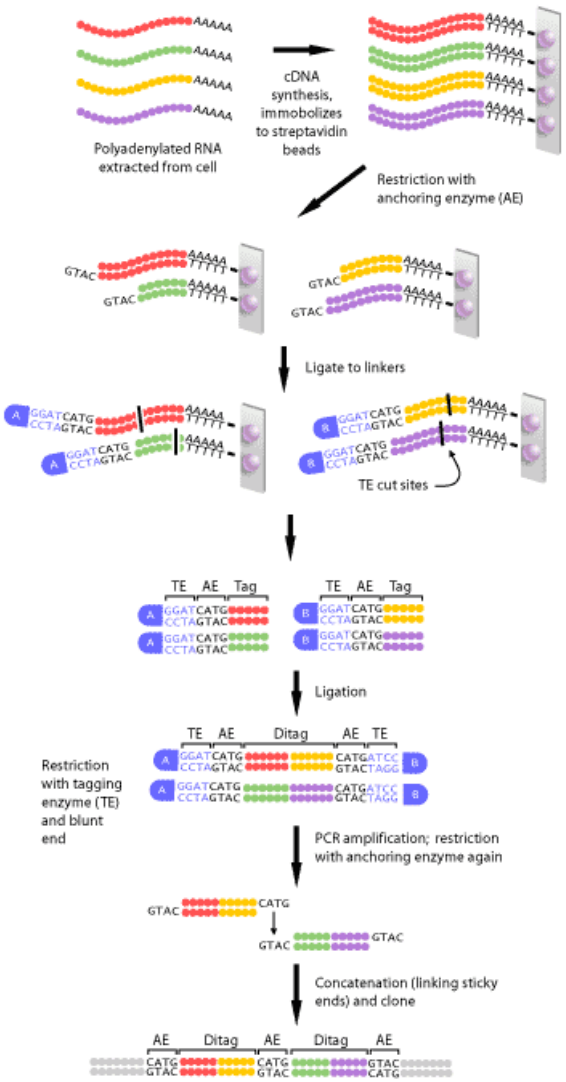


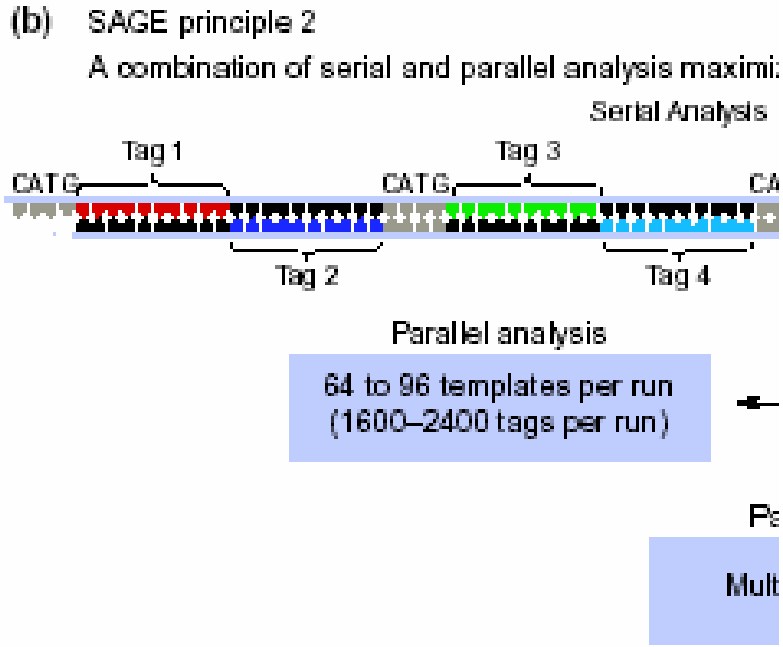
Problemas de los ESTs

- Las secuencias pueden contener errores tanto de cambio de base como de inserción/deleción
- Genes de baja expresión son difíciles de encontrar en la biblioteca de cDNA (normalización, secuenciado en masa)
- cDNAs difíciles de clonar
- No se representa el transcrito completo:
  - Solo un 11% de secuencias completas
  - 65% de regiones 3’
  - 25% de regiones 5’

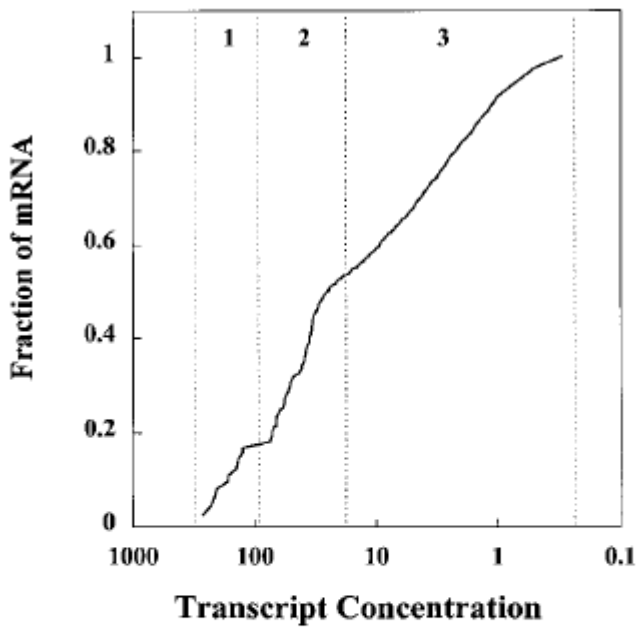
SAGE (Serial analysis of gene expression)

- 1995
- Secuenciado de pequeños tags
- Cada gen es identificado por un tag único:
  - Análisis cuantitativo de la expresión
- Requiere de un genoma secuenciado sobre el cual alinear los tags
- Búsqueda de nuevos genes





- Técnica muy sensible

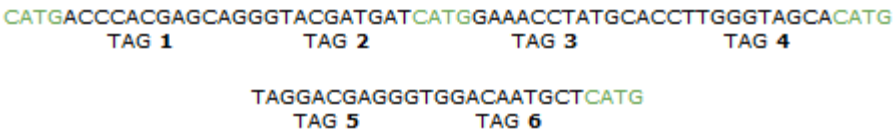


B

Component	Virtual Rot (SAGE) %mRNA	Virtual Rot (SAGE) Copies/cell	Rot (Reassociation) %mRNA	Rot (Reassociation) Copies/cell
1	17	180	23	200
2	38	40	51	30
3	45	2.5	26	1.5

Figure 3. Virtual Rot

Precesamiento de los datos

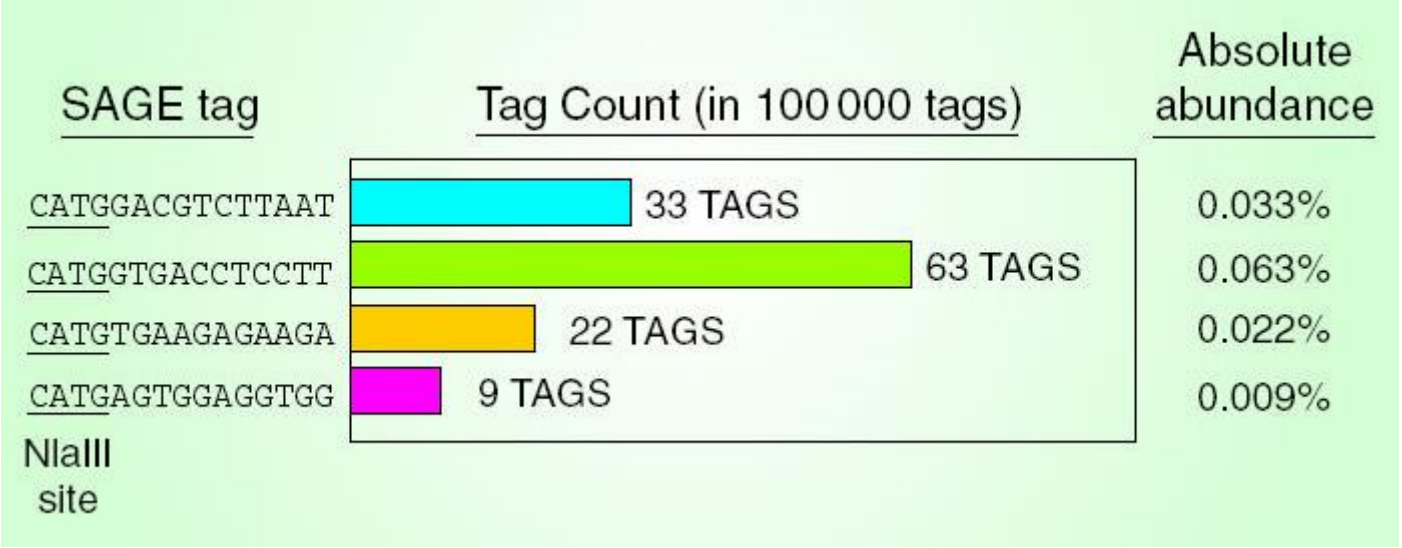


Tag_Sequence	Count
ATCTGAGTTC	1075
GCGCAGACTT	125
TCCCCGTACA	112
TAGGACGAGG	92
GCGATGGCGG	91
TAGCCCAGAT	83
GCCTTGTTTA	80
GCGATATTGT	66
TACGTTTCCA	66
TCCCGTACAT	66
TCCCTATTAA	66
GGATCACAAT	55
AAGGTTCTGG	54
CAGAACCGCG	50
GGACCGCCCC	48

Tag_Sequence	Count	Gene Name
ATATTGTCAA	5	translation elongation factor 1 gamma
AAATCGGAAT	2	T-complex protein 1, z-subunit
ACCGCCTTCG	1	no match
GCCTTGTTTA	81	rpa1 mRNA fragment for r ribosomal protein
GTTAACCATC	45	ubiquitin 52-AA extension protein
CCGCCGTGGG	9	SF1 protein (SF1 gene)
TTTTTGTTAA	99	NADH dehydrogenase 3 (ND3) gene
GCAAAACCGG	63	rpL21
GGAGCCCGCC	45	ribosomal protein L18a
GCCCGCAACA	34	ribosomal protein S31
GCCGAAGTTG	50	ribosomal protein S5 homolog (M(1)15D)
TAACGACCGC	4	BcDNA.GM12270

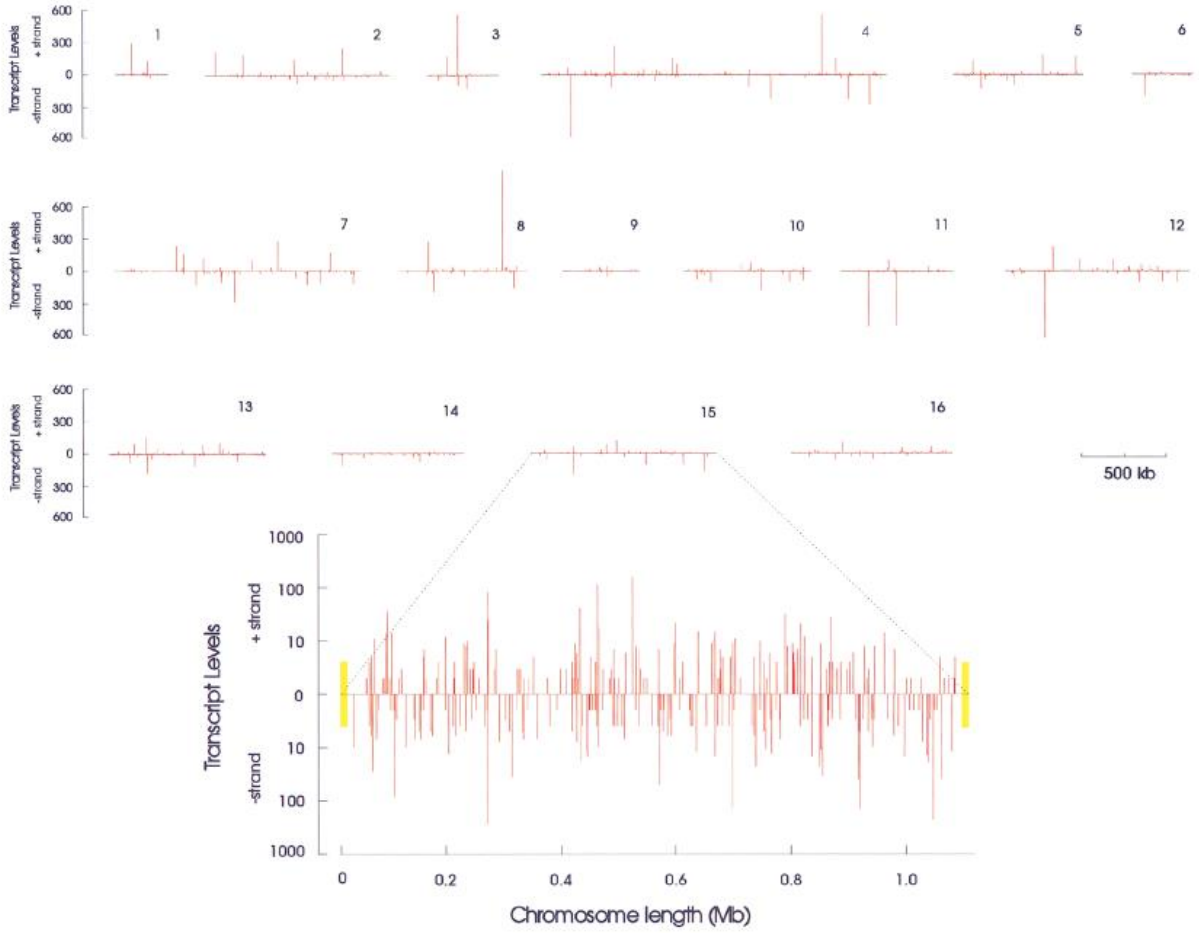
Análisis cuantitativo

- El número de veces que se secuencia un determinado tag puede ser usado como una medida de la cantidad relativa de mRNA original



Cell, Vol. 88, 243–251, January 24, 1997, Copyright ©1997 by Cell Press

Characterization of the Yeast Transcriptome



Problemas del SAGE

- Los tags son MUY pequeños (13 – 14 bp).
  - Si el gen no es conocido es casi imposible reconocer motivos funcionales
  - Especificidad: Múltiples genes pueden compartir el mismo tag. Esto ha sido mejorado aumentando el largo.
- Las enzimas de tipoIIS (generalmente la BsmFI) puede dar largos un poco diferentes lo cual genera problemas al interpretar los di-tags
- Algunos mRNAs pueden no tener el sitio de reconocimiento de la enzima
- Técnicamente complicado

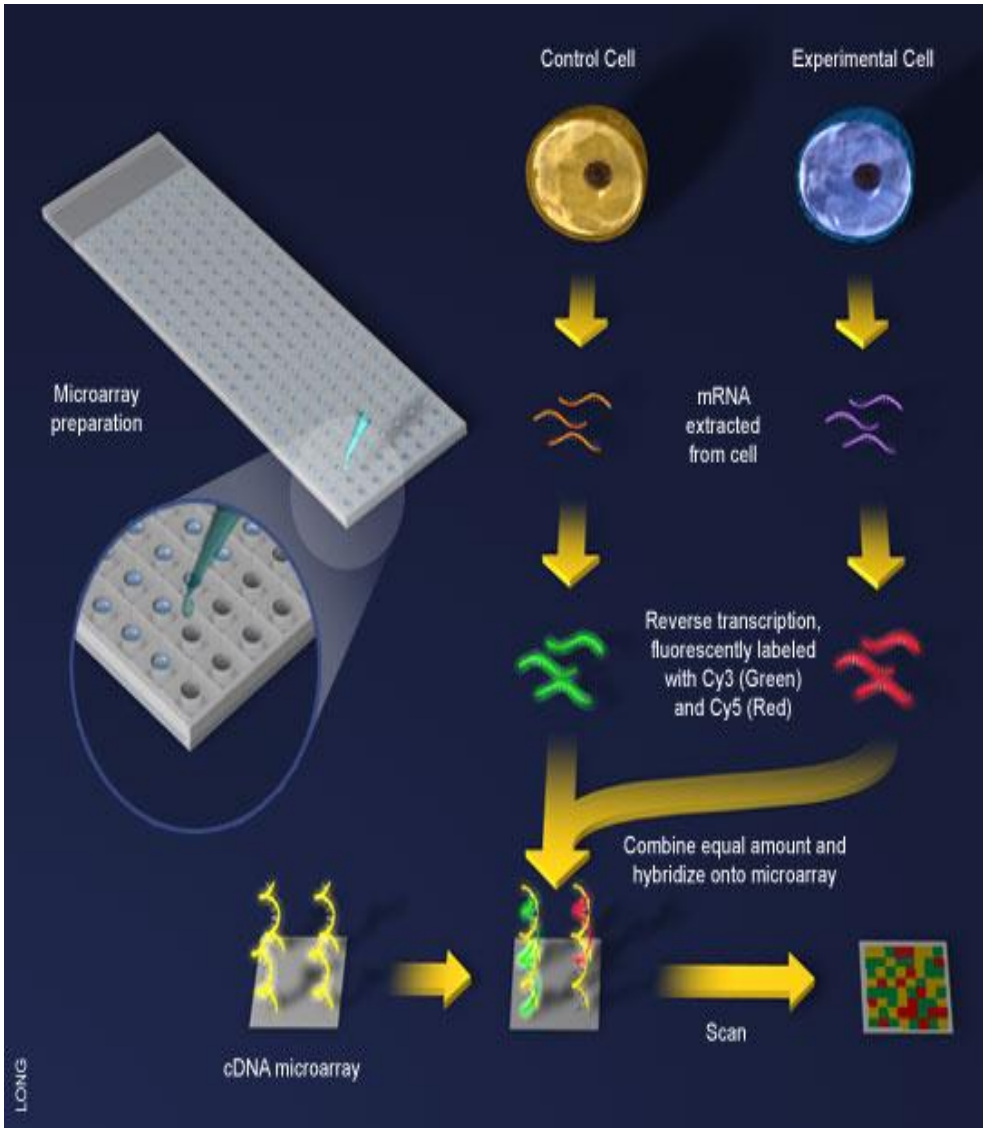
Análisis a nivel masivo: Microarreglos de ADN

- Permiten el estudio simultaneo y comparativo del transcriptoma
- Siguen en mismo principio que el *northern blot* pero paralelizado masivamente
- Aunque la presencia de un mensajero no implica presencia de la proteína un cambio en la cantidad de mensajero esta mostrando una respuesta fisiológica:
  - Análisis comparativos de diferentes situaciones
- Permiten tener un pantallazo del transcriptoma a nivel global:
  - vías metabólicas completas
- Permiten también realizar inferencias funcionales (*clustering*)
- Aplicaciones clínicas:
  - Farmacogenómica
  - Marcadores pronósticos

Tipos de Arrays

- De cDNA
  - Se construyen imprimiendo un biblioteca de cDNA en una placa
  - Son de hibridación comparativa (“de dos canales”)
  - Generalmente personalizados (diseño personal de la biblioteca a imprimir)
  - Fragmentos largos de ADN (productos de PCR)
  - Alrededor de 10.000 “*features*” por cm2
- De oligos (Chips de ADN)
  - Fragmentos cortos de ADN (Alrededor de 25pb)
  - Muy alta densidad (millones de *features*)
  - Son impresos “*in situ*”:
    - Fotolitografía
    - Eliminan la necesidad de construir una biblioteca
  - De un solo canal (un chip por muestra)

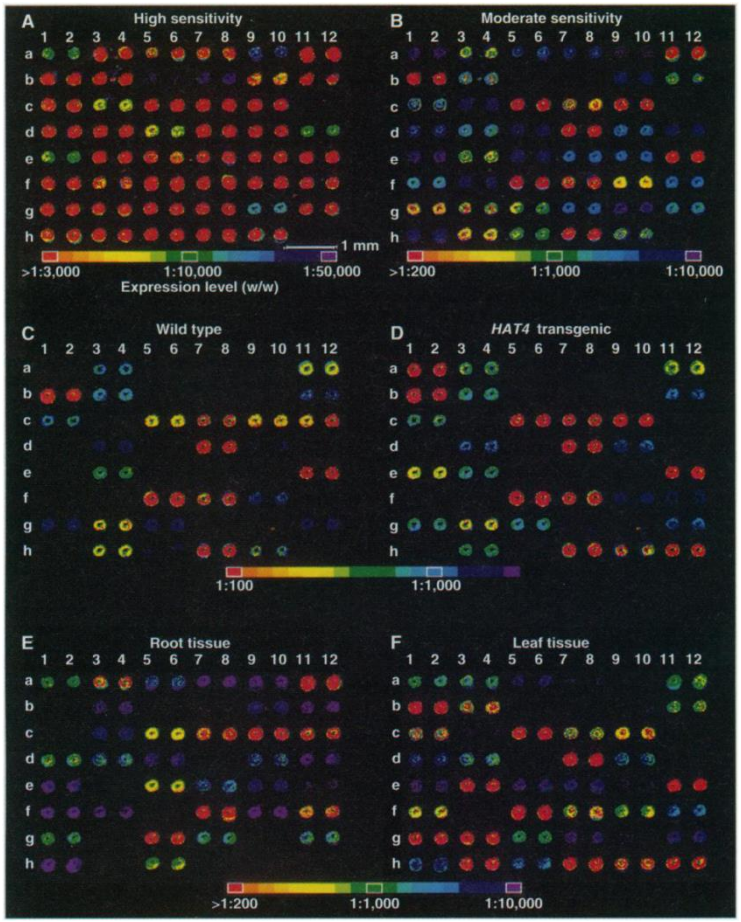
cDNA arrays


Control Cell
Experimental Cell
Microarray preparation
mRNA extracted from cell
Reverse transcription, fluorescently labeled with Cy3 (Green) and Cy5 (Red)
Combine equal amount and hybridize onto microarray
Scan
cDNA microarray
LONG

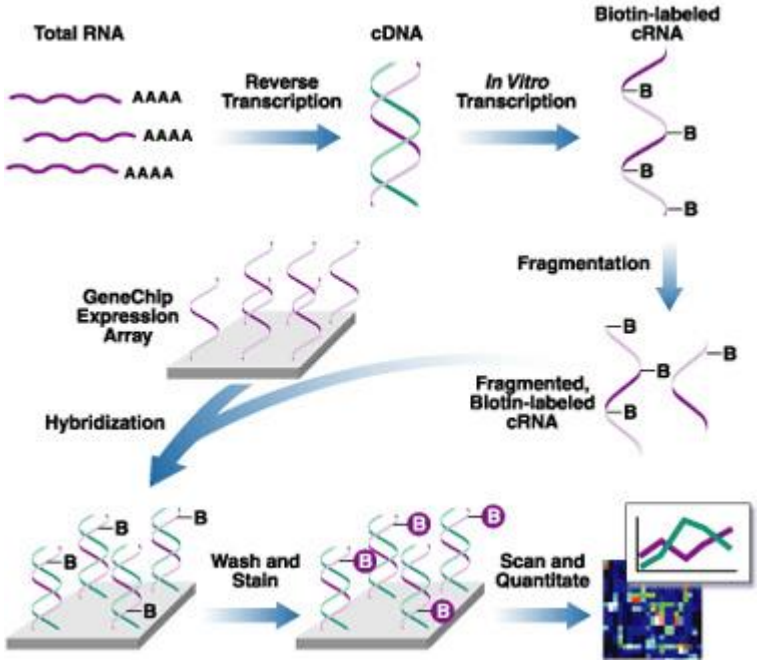
Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray

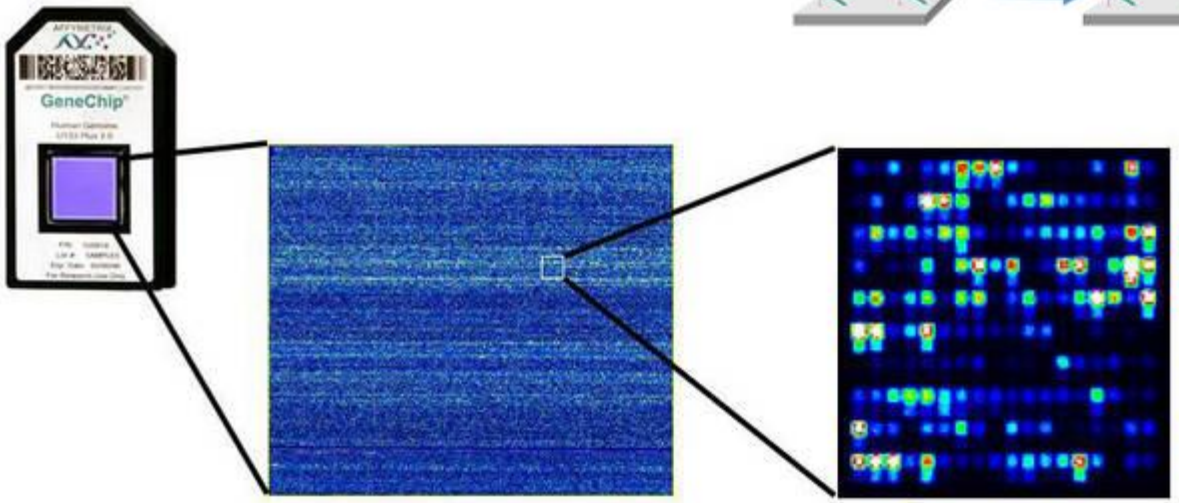
Mark Schena,* Dari Shalon,*† Ronald W. Davis,
Patrick O. Brown‡

- 1995
- 48 sondas de cDNA por duplicado
- Hoy en día se producen *arrays* de decenas de miles de *features*

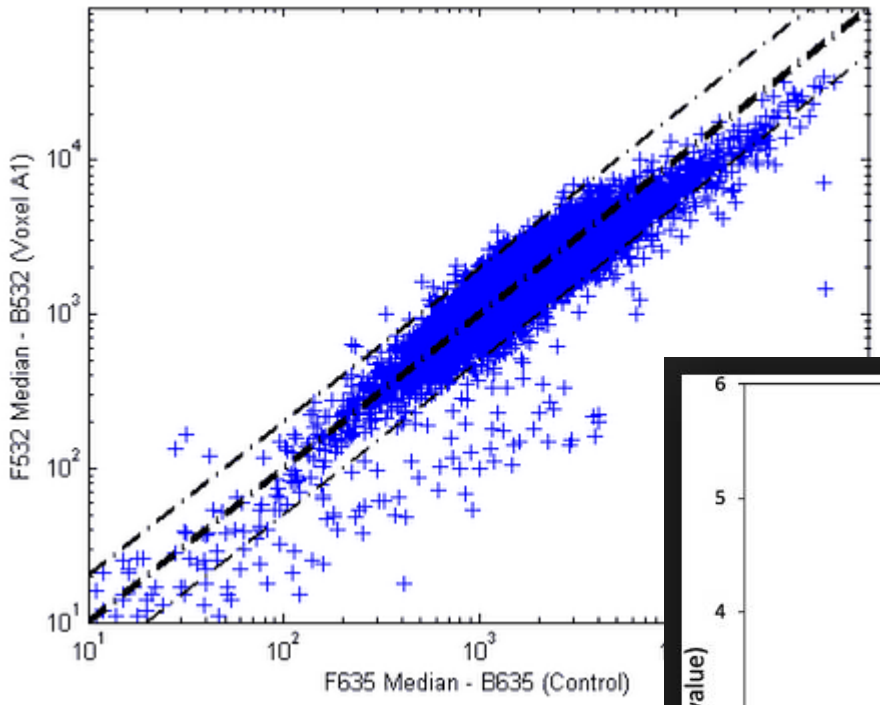


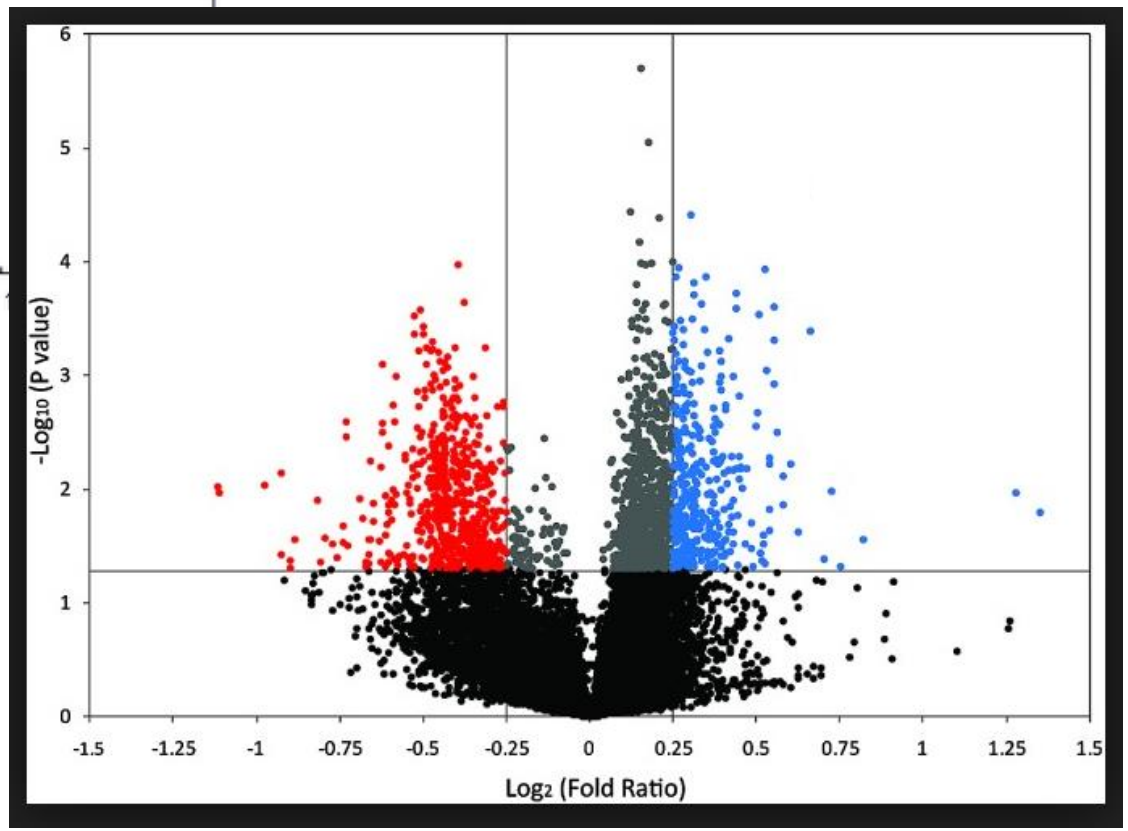
Affimetrix


Total RNA
cDNA
Biotin-labeled cRNA
AAAA
AAAA
AAAA
Reverse Transcription
In Vitro Transcription
B
B
B
B
Fragmentation
GeneChip Expression Array
B
B
B
B
Fragmented, Biotin-labeled cRNA
Hybridization
B
B
B
B
Wash and Stain
B
B
B
B
Scan and Quantitate

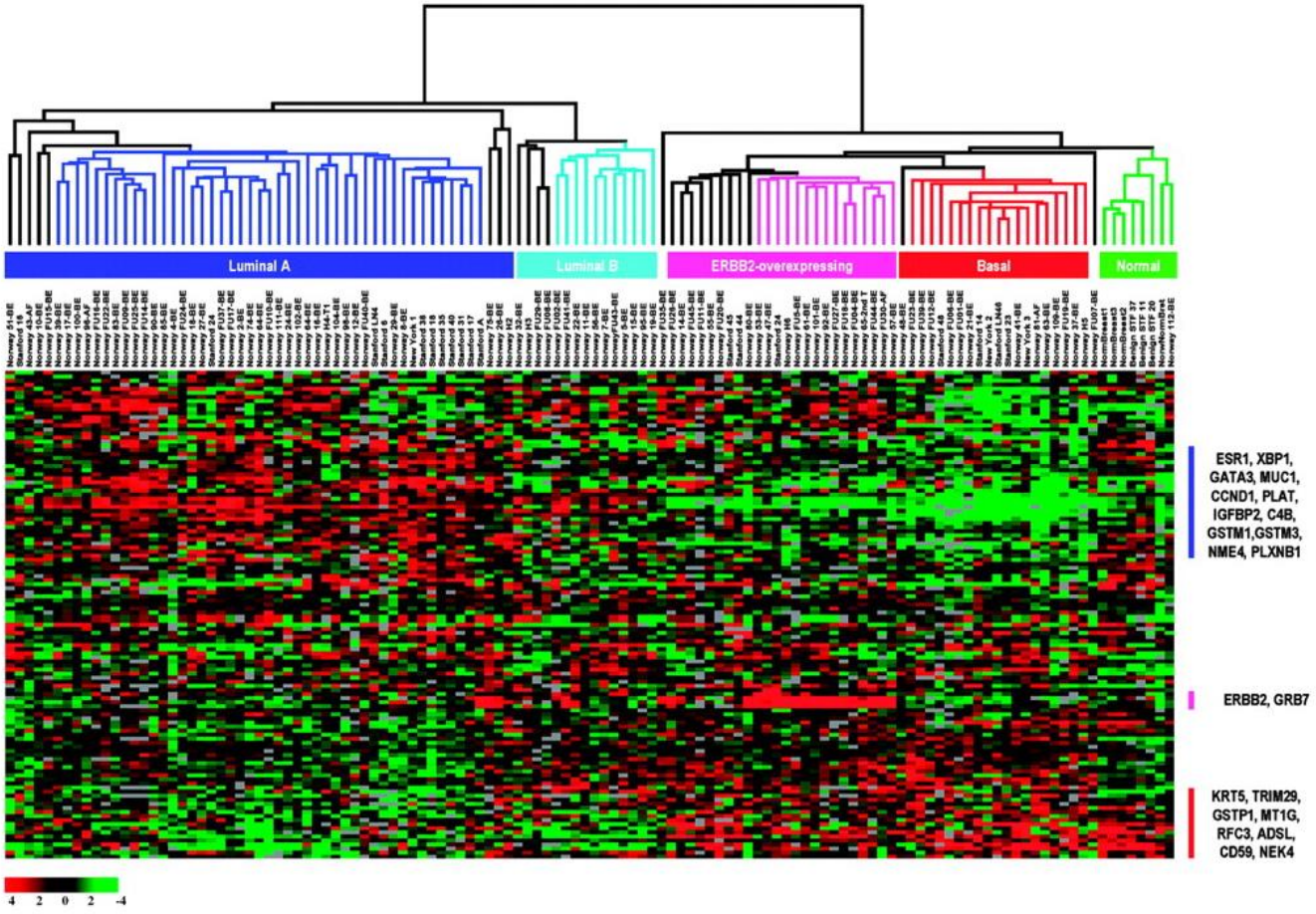

GeneChip

Análisis de niveles de expresión



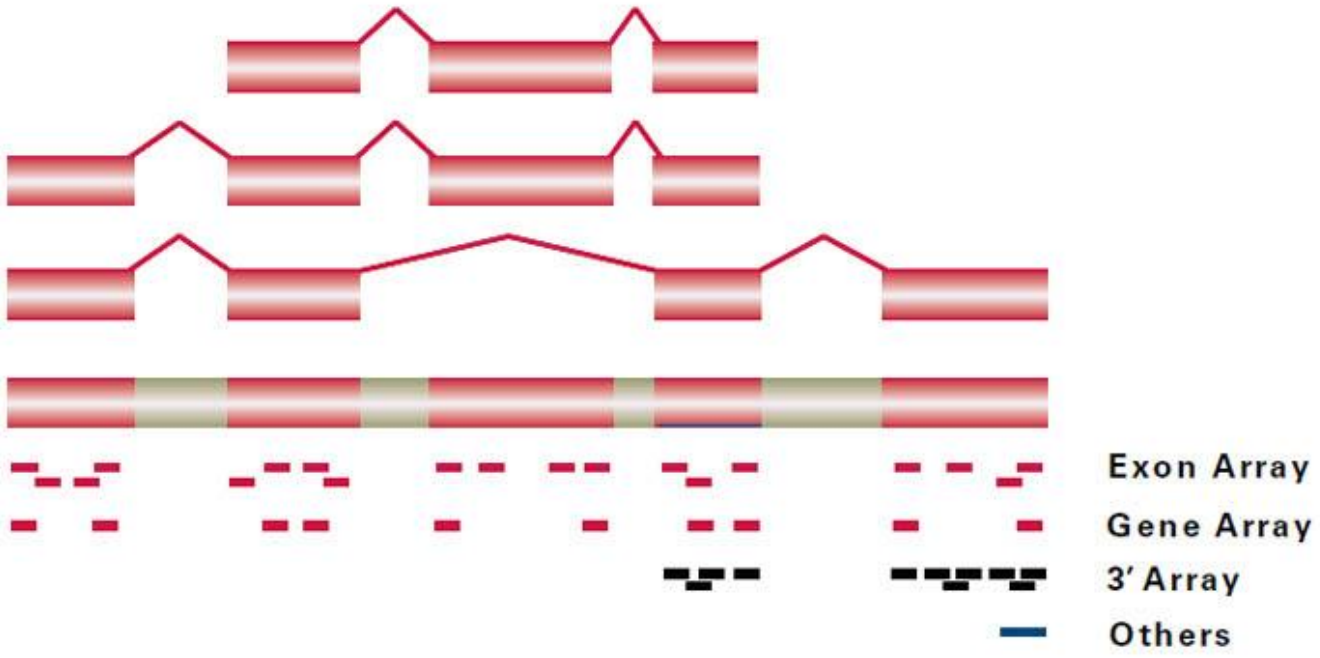


Análisis de niveles de expresión



Arrays de exones

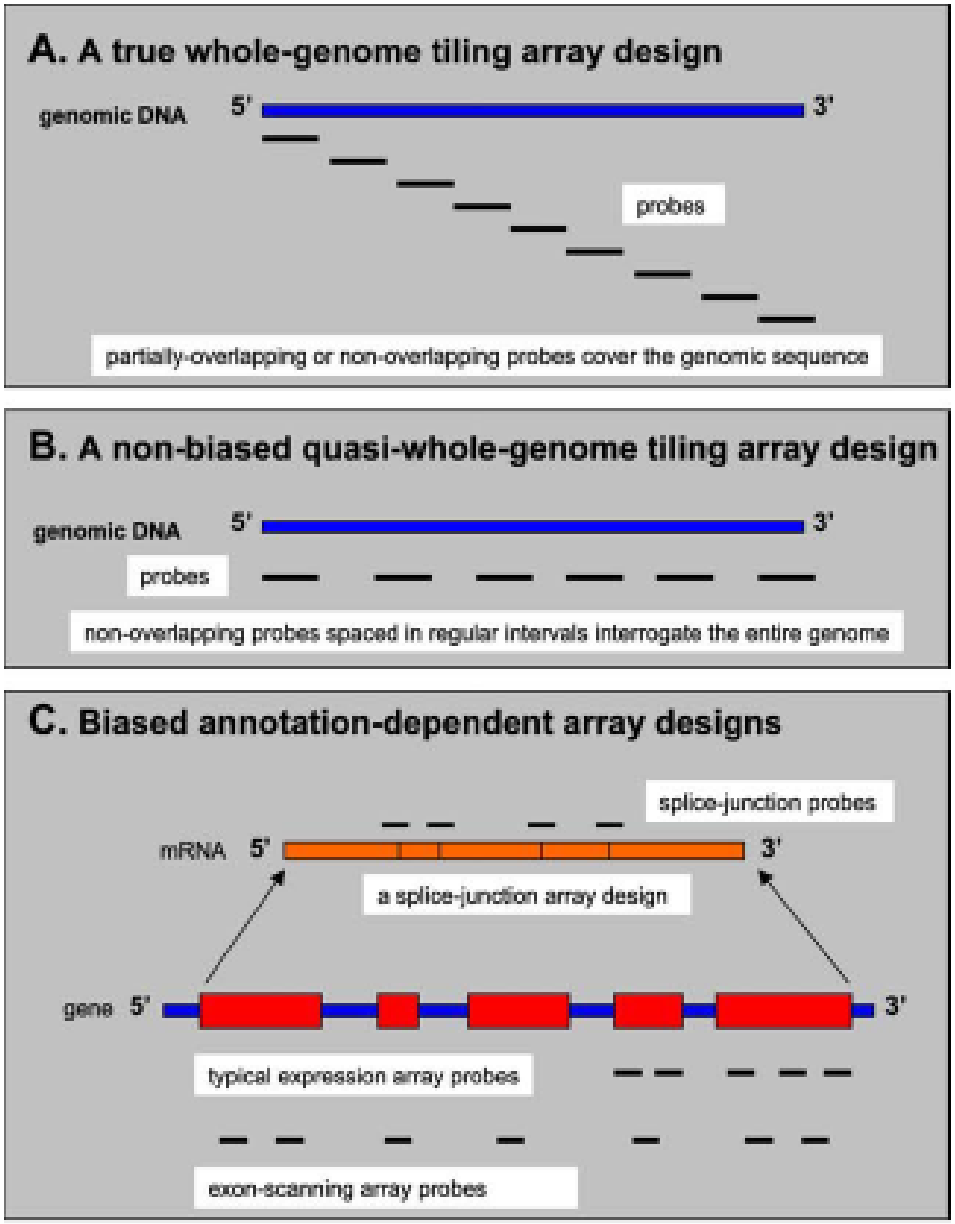
- Permiten evaluar patrones de splicing



Veo solo lo que quiero ver…

- Desarrollo de los "*tiling arrays*" (*Whole genome arrays*, WGA)
  - La idea es abarcar el genoma entero
  - No es necesario conocimiento previo de que es lo que se transcribe
- Permitió avances significativos en el conocimiento de los transcriptomas
- Permitió abarcar un número mucho mayor de preguntas

Tiling arrays



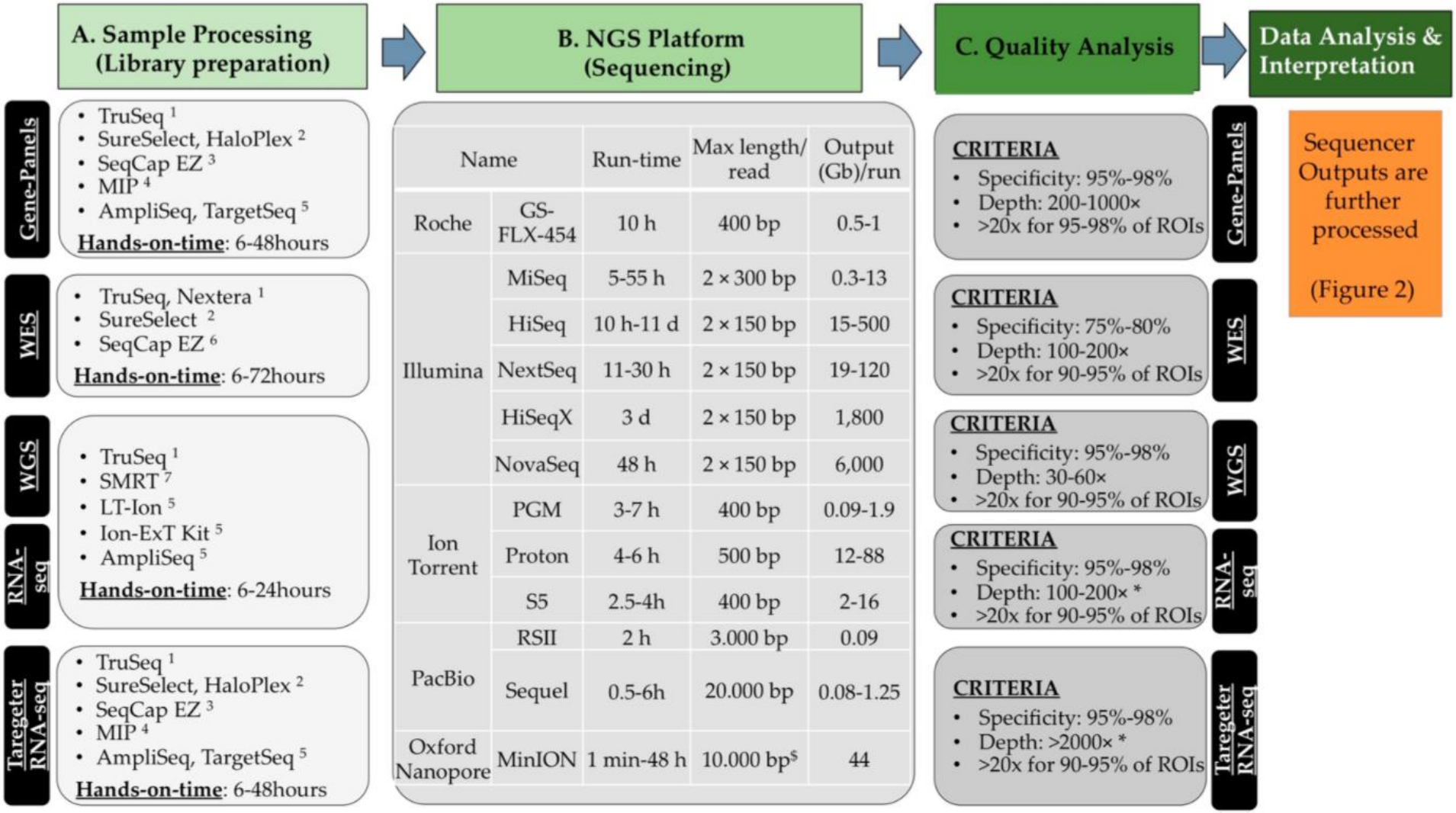
Aplicaciones

Global Identification of Human Transcribed Sequences with Genome Tiling Arrays

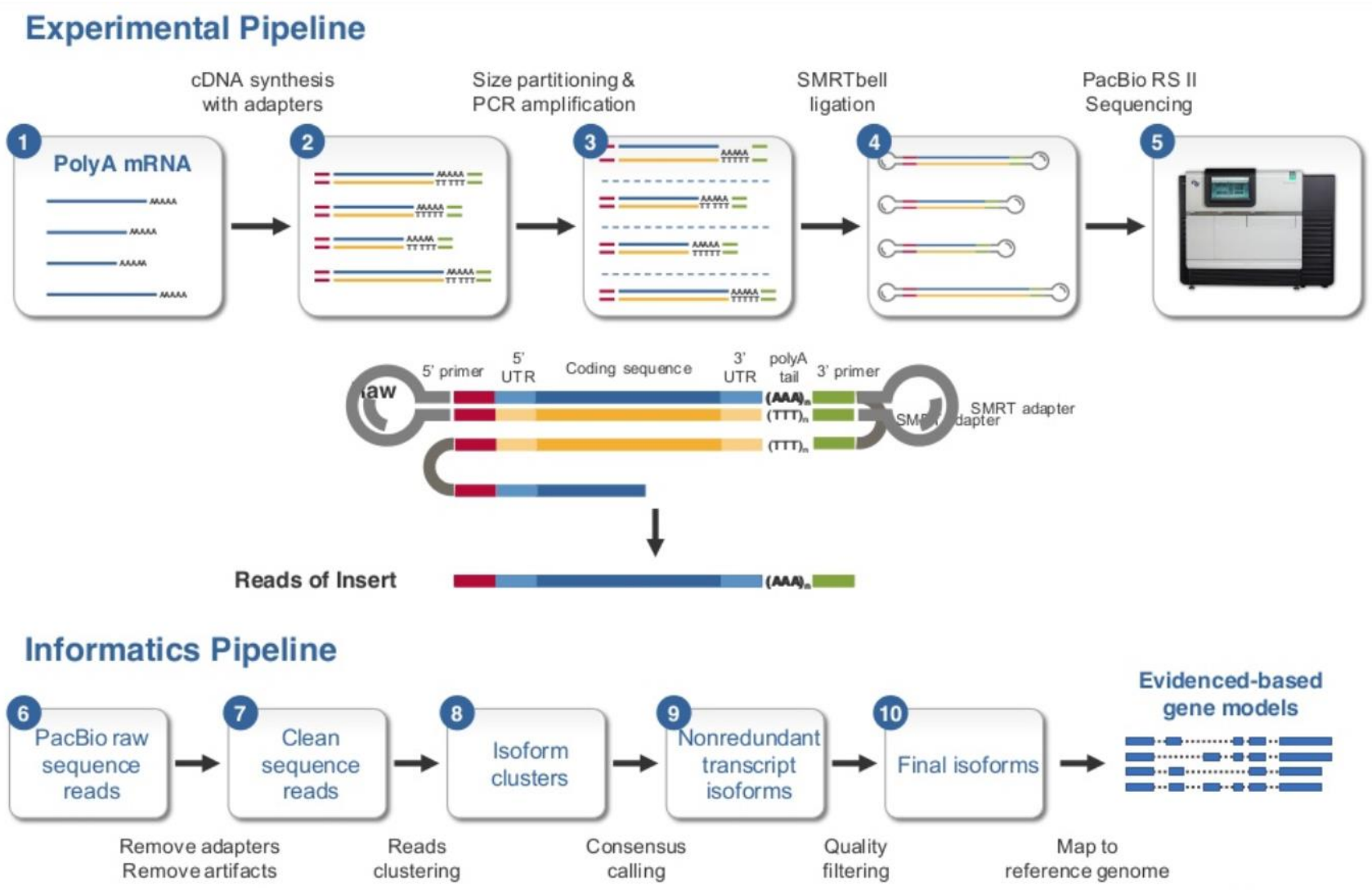
Paul Bertone,[1*] Viktor Stolc,[1,2*] Thomas E. Royce,[3] Joel S. Rozowsky,[3] Alexander E. Urban,[1] Xiaowei Zhu,[1] John L. Rinn,[3] Waraporn Tongprasit,[4] Manoj Samanta,[2] Sherman Weissman,[5] Mark Gerstein,[3†] Michael Snyder[1,3†]

- 13000 regiones transcritas previamente anotadas
- 8000 regiones transcritas sin anotación previa
  - 1500 serian exones previamente no reconocidos
  - 1500 en hebra anti sentido en los intrones
  - 5700 lejos de genes conocidos

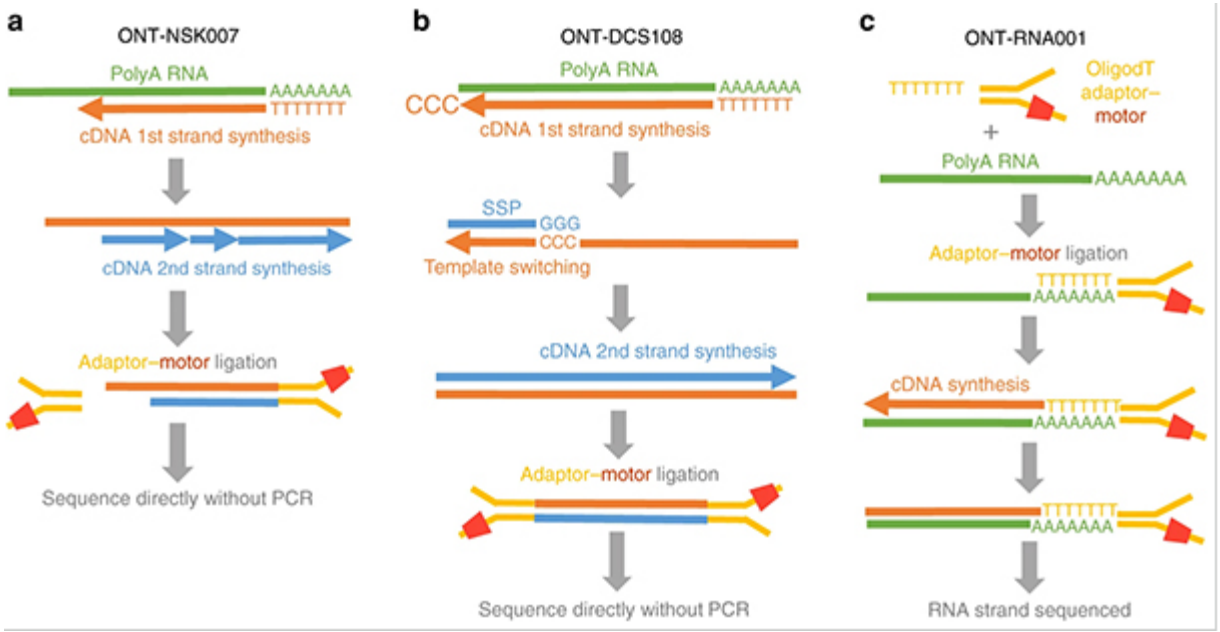
Plataformas de secuenciación masiva

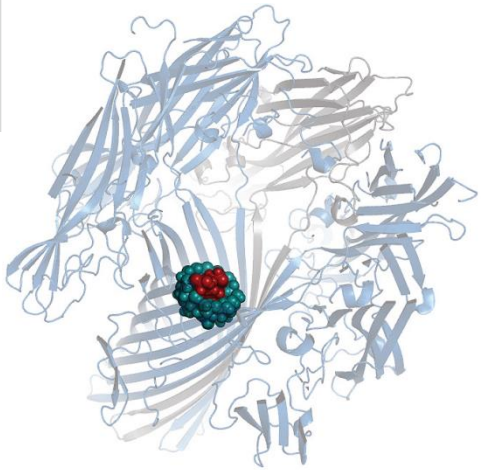


Secuenciadores de tercera generación: PacBio



Secuenciadores de tercera generación: Nanopore





Capacidad de secuenciar el transcrito entero

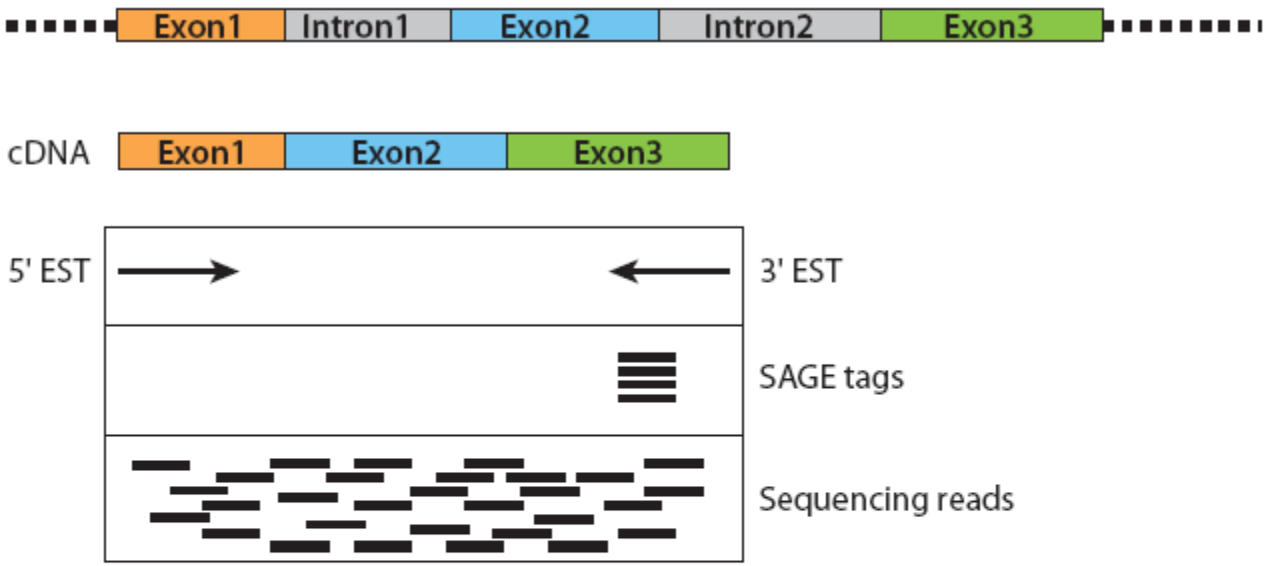
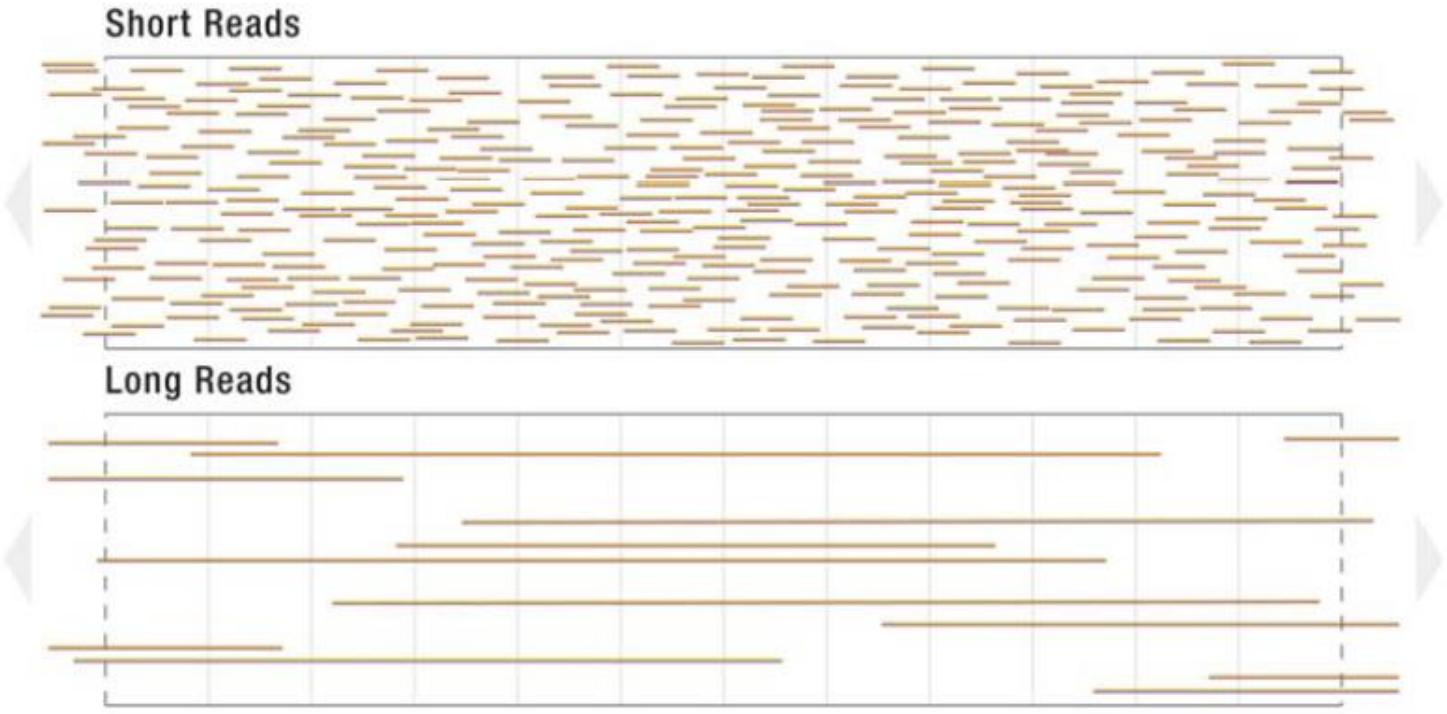


Figure 1

Gene model coverage by various sequencing-based methods for transcriptome analysis. Sanger-based expressed sequence tags (ESTs) are generated from the 3′ or 5′ end of the transcript, whereas SAGE tags represent short sequences at its 3′ end. Randomly primed short reads generated by next-generation sequencers detect bases throughout the length of the transcript.

Ensamblaje

- Le elección de la plataforma va a estar sujeta a la disponibilidad de genoma completo



Anotación de regiones codificantes

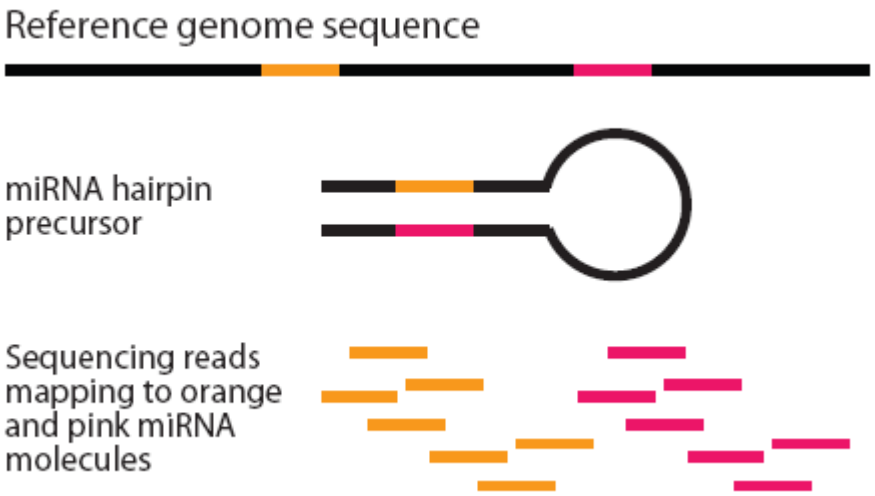
- Se obtienen secuencias de todos los exones del gen
  - Descripción de nuevos exones
  - Descripción de nuevas variantes de splicing (pair ends sequencing)

Figure 3

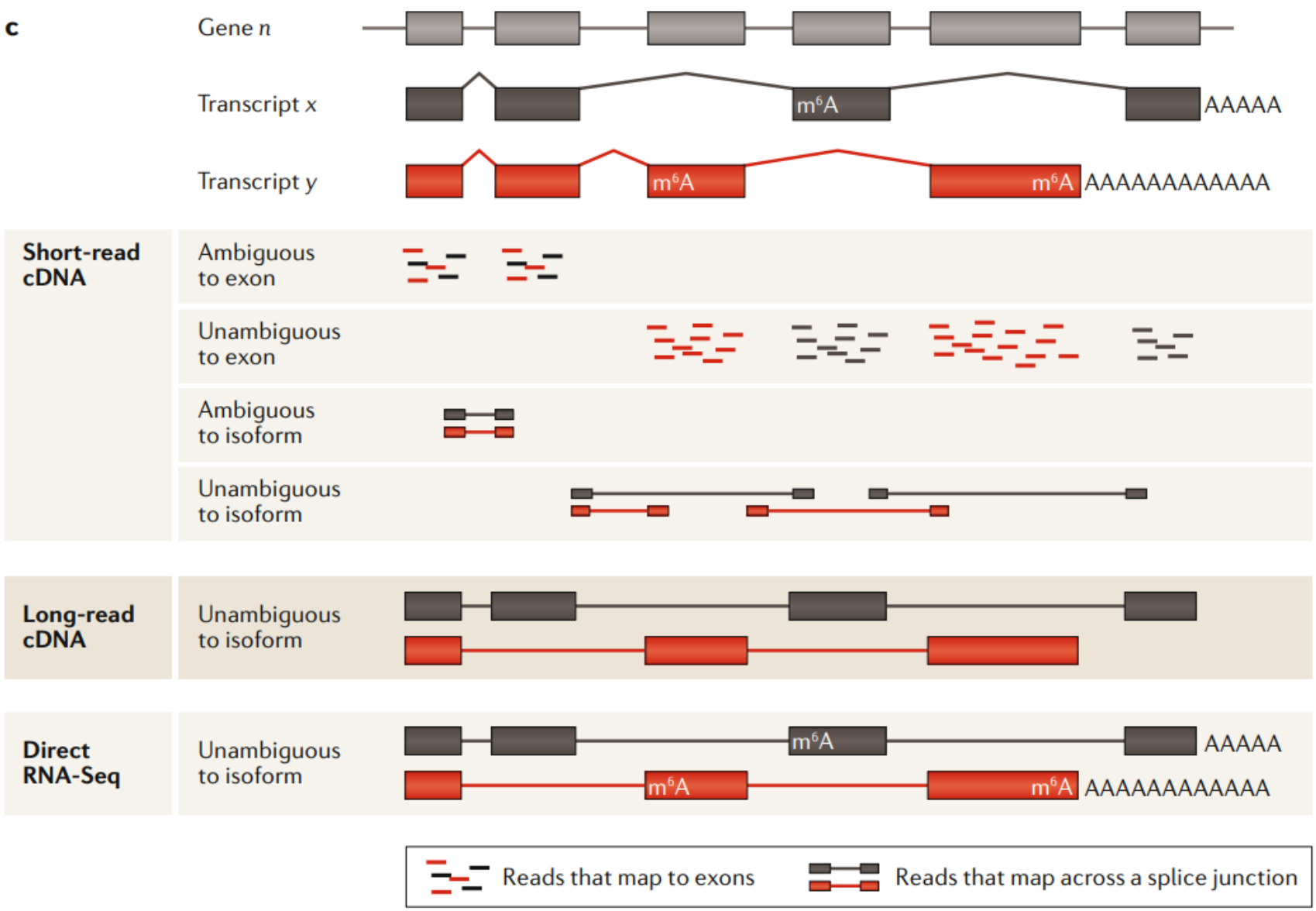
Protein-coding gene annotation using transcriptome sequencing data. This figure illustrates how novel exons and introns can be discovered by mapping transcriptome sequencing reads to an annotated reference genome sequence.

RNAs no codificantes

- Secuenciado de regiones no anotadas no codificantes



Identificación de variantes

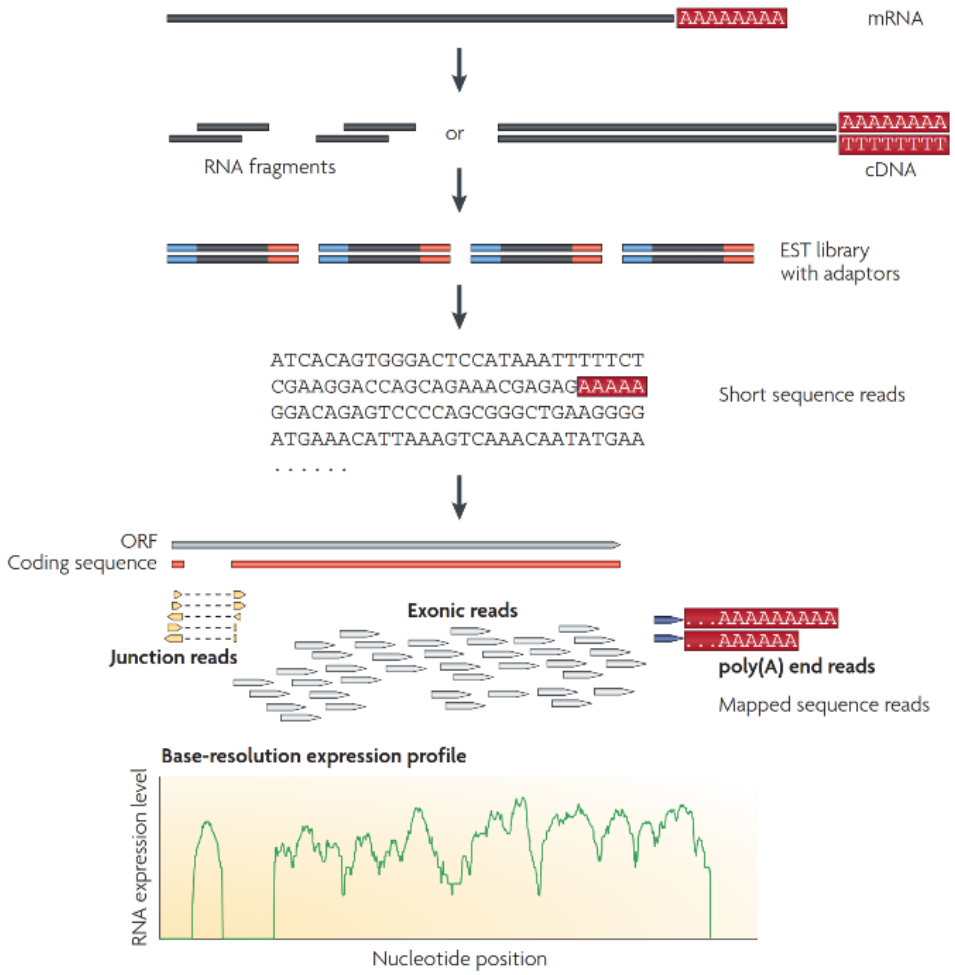


Segunda vs tercera generación

Sequencing technology	Platform	Advantages	Disadvantages	Key applications
Short-read cDNA	Illumina, Ion Torrent	• Technology features very high throughput: currently 100–1,000 times more reads per run than long-read platforms • Biases and error profiles are well understood (homopolymers are still an issue for Ion Torrent) • A huge catalogue of compatible methods and computational workflows are available • Analysis works with degraded RNA	• Sample preparation includes reverse transcription, PCR and size selection adding biases to all methods • Isoform detection and quantitation can be limited • Transcript discovery methods require a de novo transcriptome alignment and/or assembly step	Nearly all RNA-seq methods have been developed for short-read cDNA sequencing: DGE, WTA, small RNA, single-cell, spatialomics, nascent RNA, translatome, structural and RNA–protein interaction analysis, and more are all possible
Long-read cDNA	PacBio, ONT	• Long reads of 1–50 kb capture many full-length transcripts • Computational methods for de novo transcriptome analysis are simplified	• Technology features low-to-medium throughput: currently only 500,000 to 10 million reads per run • Sample preparation includes reverse transcription, PCR and size selection (for some protocols), adding biases to many methods • Degraded RNA analysis is not recommended	Sequencing is particularly suited to isoform discovery, de novo transcriptome analysis, fusion transcript discovery, and MHC, HLA or other complex transcript analysis
Long-read RNA	ONT	• Long reads of 1–50 kb capture many full-length transcripts • Computational methods for de novo transcriptome analysis are simplified • Sample preparation does not require reverse transcription or PCR-reducing biases • RNA base modifications can be detected • Poly(A) tail lengths can be directly estimated from single-molecule sequencing	• Technology features low throughput: currently only 500,000 to 1 million reads per run • Sample preparation and sequencing biases are not well understood • Degraded RNA analysis is not recommended	• Sequencing is particularly suited to isoform discovery, de novo transcriptome analysis, fusion transcript discovery, and MHC, HLA or other complex transcript analysis • Ribonucelotide modifications can be detected

RNA-seq

- El conteo de lecturas es cuantitativo lo que nos da idea de cantidad de expresión



RNA-seq

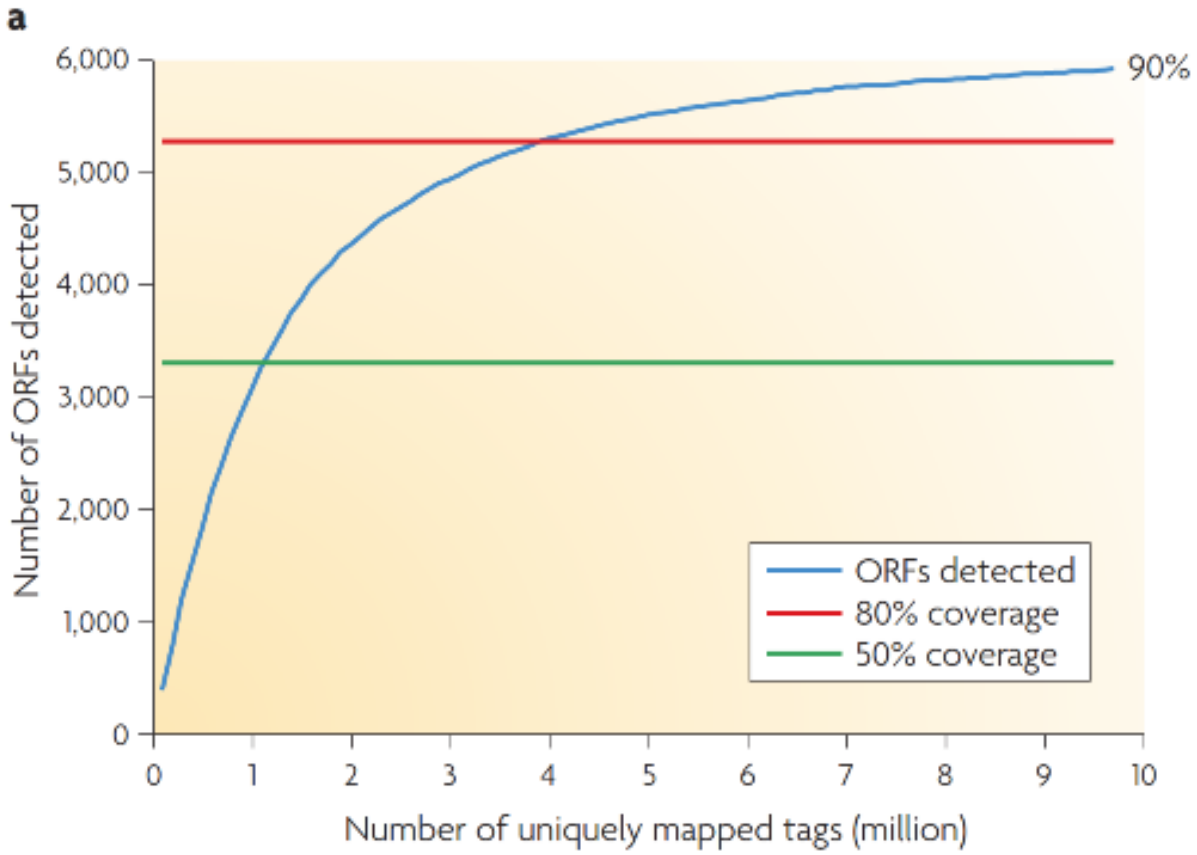
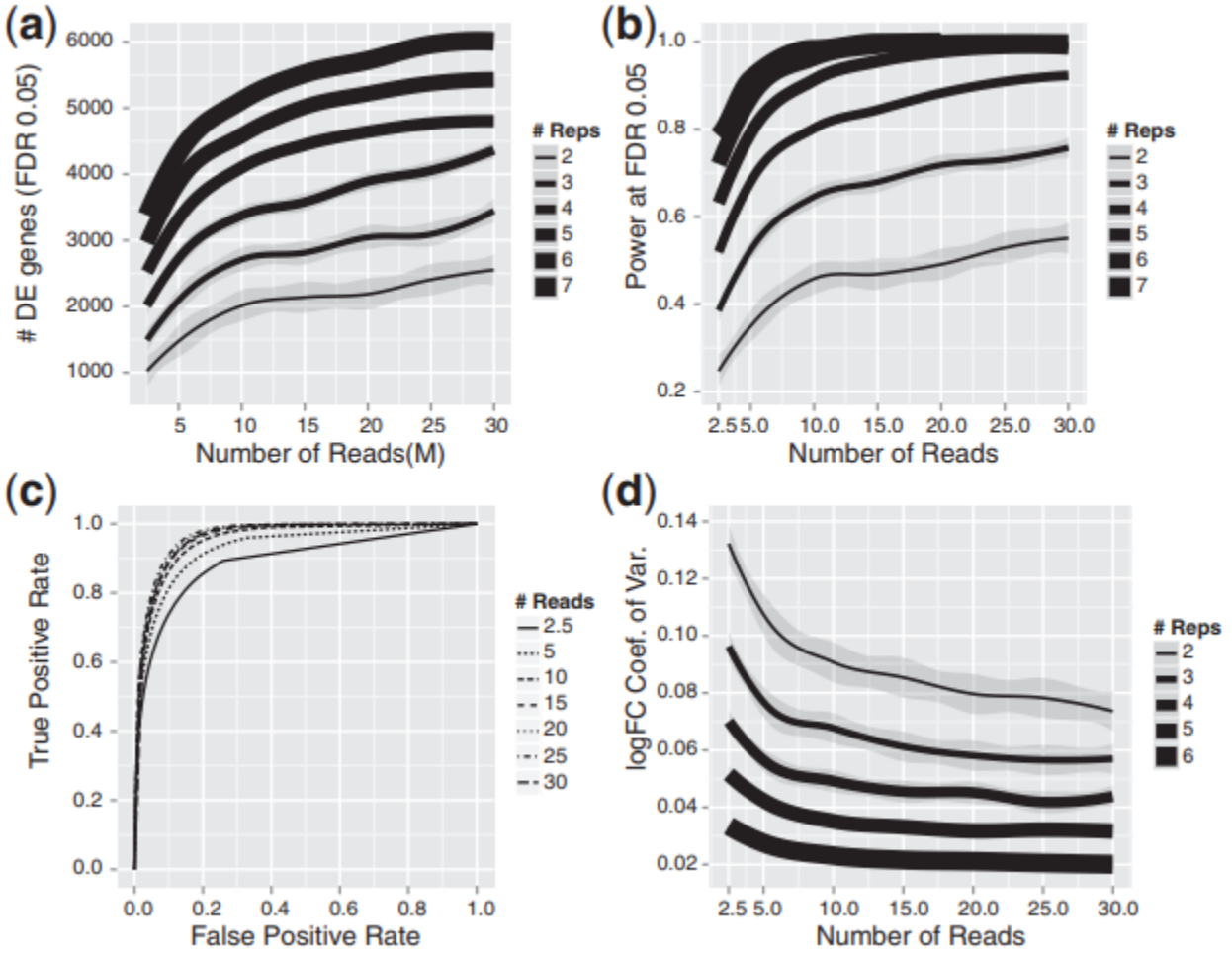


Figure 5 | **Coverage versus depth. a** | 80% of yeast genes were detected at 4 million uniquely mapped RNA-Seq reads, and coverage reaches a plateau afterwards despite the increasing sequencing depth. Expressed genes are defined as having at least four independent reads from a 50-bp window at the 3′ end. Data is taken from REF. 18.

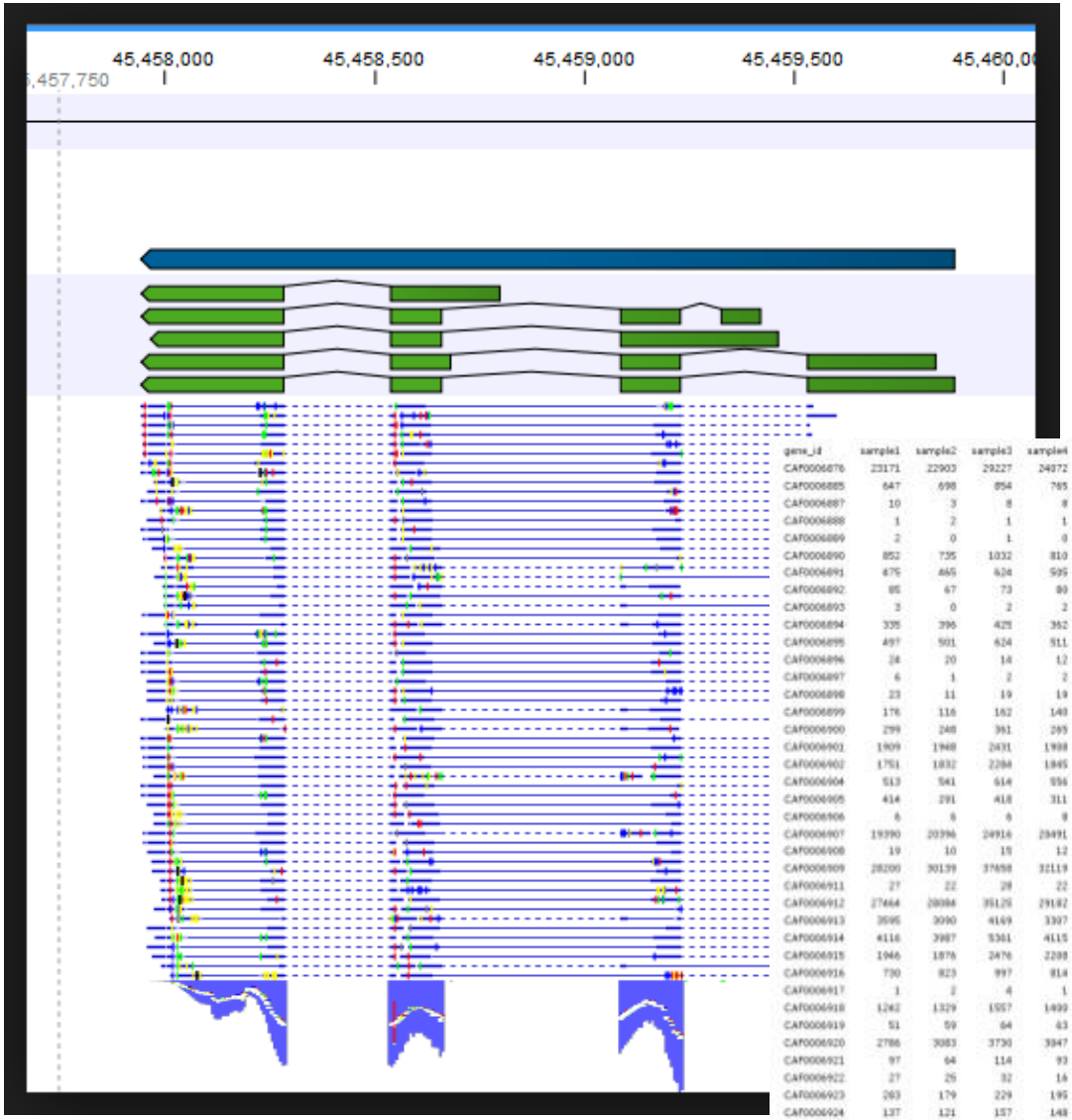
RNA-seq

- Replicas vs profundidad



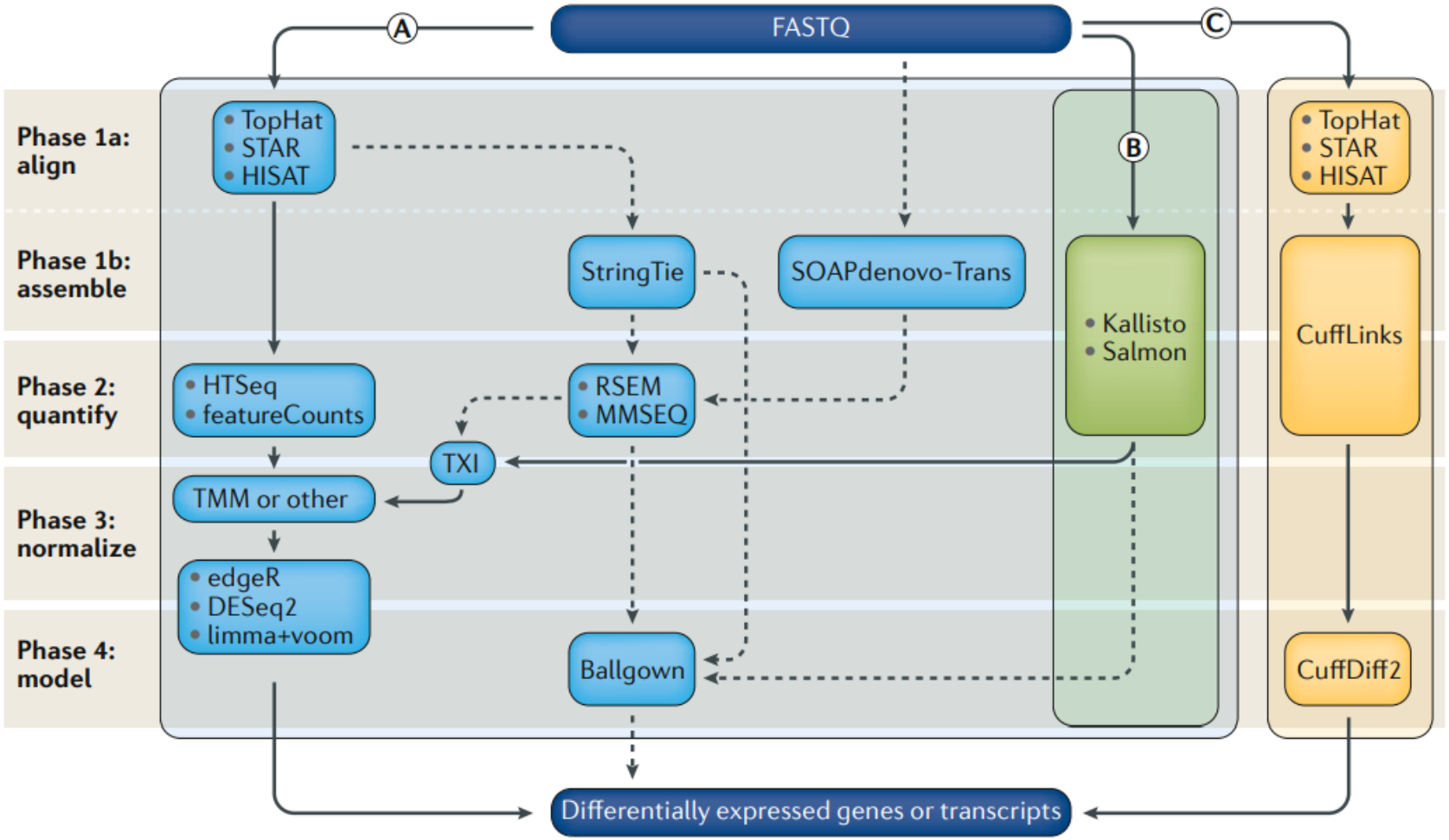
Liu et al. Bioinformatics 2014

RNA-seq



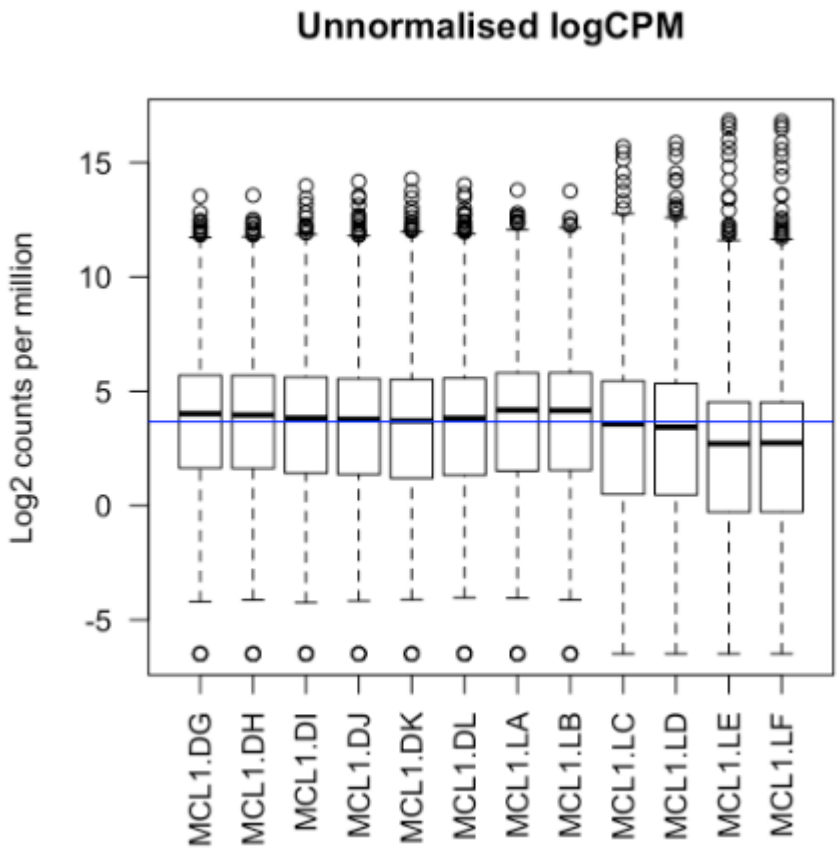
gene_id	sample1	sample2	sample3	sample4	sample5	sample6	sample7	sample8	sample9	sample10	sample11	sample12	sample13	sample14
CAF0006876	23171	22903	29227	24072	23151	26336	25252	24122	19527	26898	18883	24237	28640	22315
CAF0006885	647	698	854	765	797	816	868	767	532	761	583	854	748	720
CAF0006887	10	3	8	8	5	8	5	3	7	8	2	10	7	8
CAF0006888	1	2	1	1	0	0	0	0	1	0	1	0	0	2
CAF0006889	2	0	1	0	1	0	2	0	1	1	1	0	0	3
CAF0006890	852	735	1032	810	1476	1437	1575	1358	844	859	549	747	1320	945
CAF0006891	675	665	624	595	538	624	654	562	431	586	419	550	624	472
CAF0006892	85	67	73	80	151	91	114	93	81	65	47	84	91	76
CAF0006893	3	0	2	2	0	0	0	3	1	3	0	0	1	2
CAF0006894	335	396	425	362	414	455	456	405	315	418	293	416	479	388
CAF0006895	497	501	624	511	734	679	858	671	526	599	389	551	748	539
CAF0006896	24	20	14	12	41	20	35	20	17	29	7	25	22	15
CAF0006897	6	1	2	2	0	1	1	1	0	0	0	0	3	0
CAF0006898	23	11	19	19	24	26	23	19	21	17	8	20	20	19
CAF0006899	176	116	162	140	286	210	290	202	211	147	85	149	188	149
CAF0006900	299	248	361	265	409	364	412	382	261	301	193	273	392	273
CAF0006901	1909	1940	2431	1908	2556	2457	2893	2333	1953	2114	1575	1864	2913	2113
CAF0006902	1751	1832	2284	1845	1409	1884	1409	1675	1344	2231	1511	1988	1670	1784
CAF0006904	513	561	614	556	565	624	592	569	434	577	395	538	594	487
CAF0006905	414	291	418	311	638	495	633	502	414	387	233	385	493	367
CAF0006906	6	6	6	8	3	12	4	11	3	6	5	5	7	10
CAF0006907	19390	20396	24916	20491	16028	21635	19870	20193	15819	22492	15145	20599	18641	18245
CAF0006908	19	10	15	12	18	22	27	26	9	7	4	12	8	21
CAF0006909	28200	30139	37658	32119	42627	49176	45831	44885	24894	33766	23816	38757	46659	39146
CAF0006911	27	22	28	22	25	34	21	25	41	32	15	29	48	26
CAF0006912	27464	28084	35125	29182	26054	29358	29221	28087	21584	31347	22244	28215	28627	26157
CAF0006913	3595	3090	4149	3307	3826	3900	3976	3820	3384	4852	2623	3917	3914	3228
CAF0006914	4116	3987	5361	4115	4394	4369	4941	4036	3093	4719	3286	4448	5695	3829
CAF0006915	1946	1876	2476	2288	2729	2517	3088	2366	1893	2112	1469	1953	2379	1864
CAF0006916	730	823	997	814	764	947	837	798	688	951	633	891	881	786
CAF0006917	1	2	4	1	1	4	0	1	0	6	0	4	1	1
CAF0006918	1242	1329	1557	1400	1061	1491	1247	1444	969	1543	1045	1374	1321	1413
CAF0006919	51	59	64	63	35	50	50	55	29	55	46	53	48	46
CAF0006920	2786	3083	3730	3047	4723	5068	5199	4673	2311	3382	2556	3867	5462	4505
CAF0006921	97	64	114	93	156	155	158	118	152	105	55	133	75	81
CAF0006922	27	25	32	16	32	30	25	34	25	35	16	28	34	27
CAF0006923	263	179	229	195	520	384	484	435	407	223	138	229	365	289
CAF0006924	137	121	157	148	128	136	125	132	139	160	111	178	152	119
CAF0006925	28	13	27	18	46	36	58	18	38	19	12	27	33	18

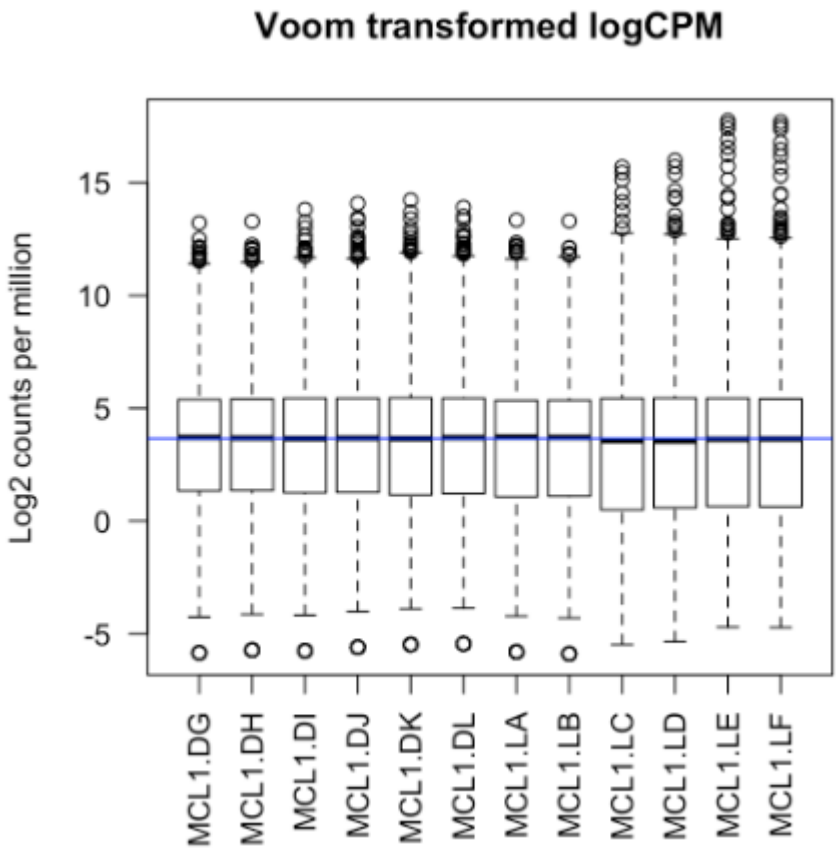
RNA-seq


FASTQ
A
B
C
Phase 1a: align
Phase 1b: assemble
Phase 2: quantify
Phase 3: normalize
Phase 4: model
• TopHat
• STAR
• HISAT
StringTie
SOAPdenovo-Trans
• Kallisto
• Salmon
• TopHat
• STAR
• HISAT
CuffLinks
• HTSeq
• featureCounts
• RSEM
• MMSEQ
TXI
TMM or other
• edgeR
• DESeq2
• limma+voom
Ballgown
CuffDiff2
Differentially expressed genes or transcripts

RNA-seq

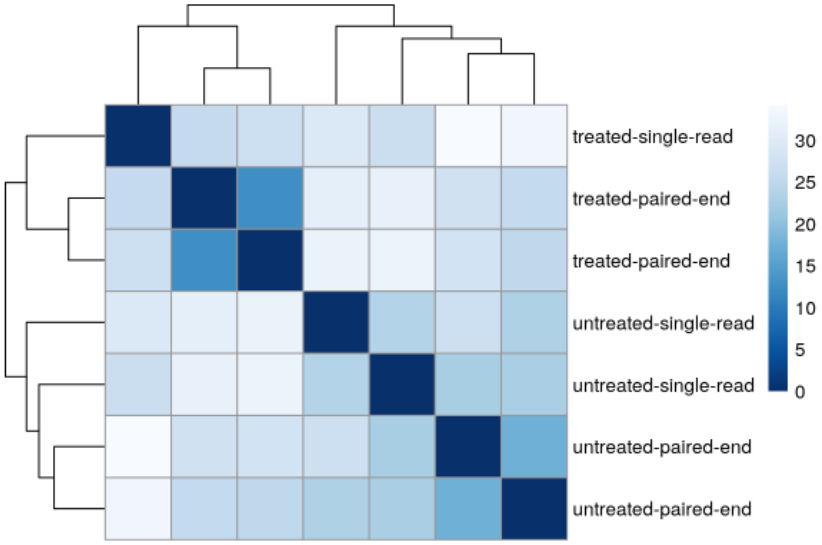
- Normalización:
  - Corrige efecto de profundidad diferente entre muestras

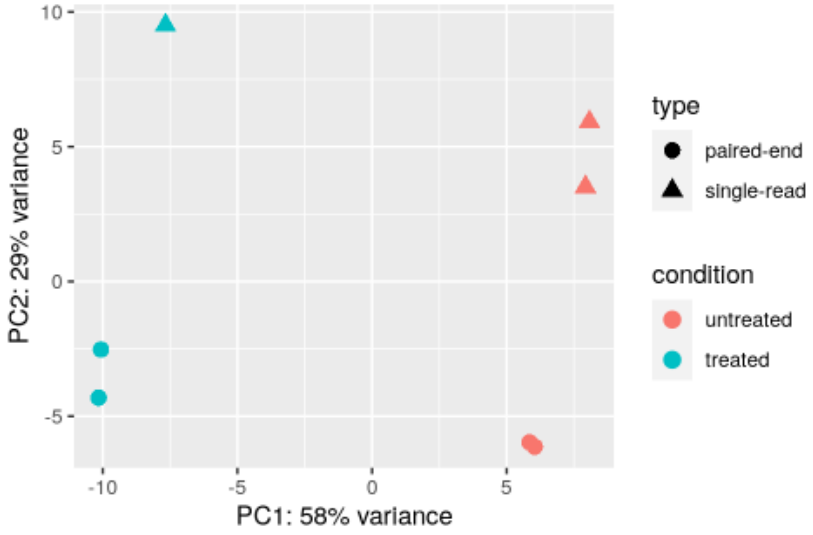




RNA-seq

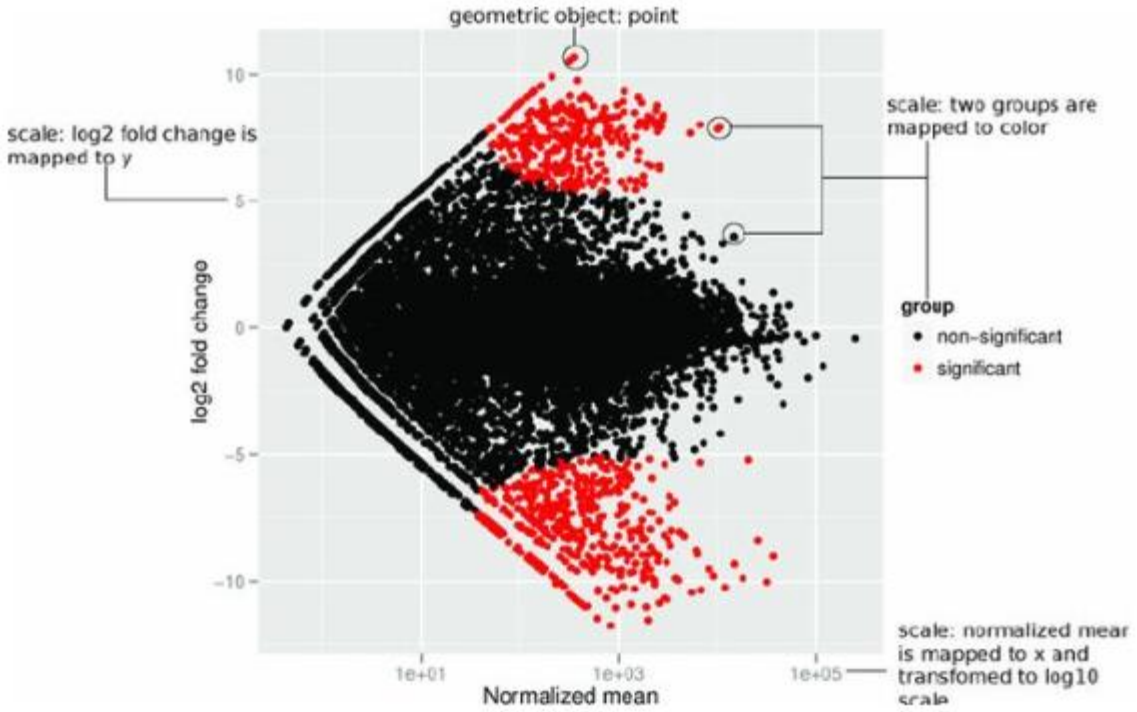
- Análisis exploratorios:
  - Correlación entre muestras





RNA-seq

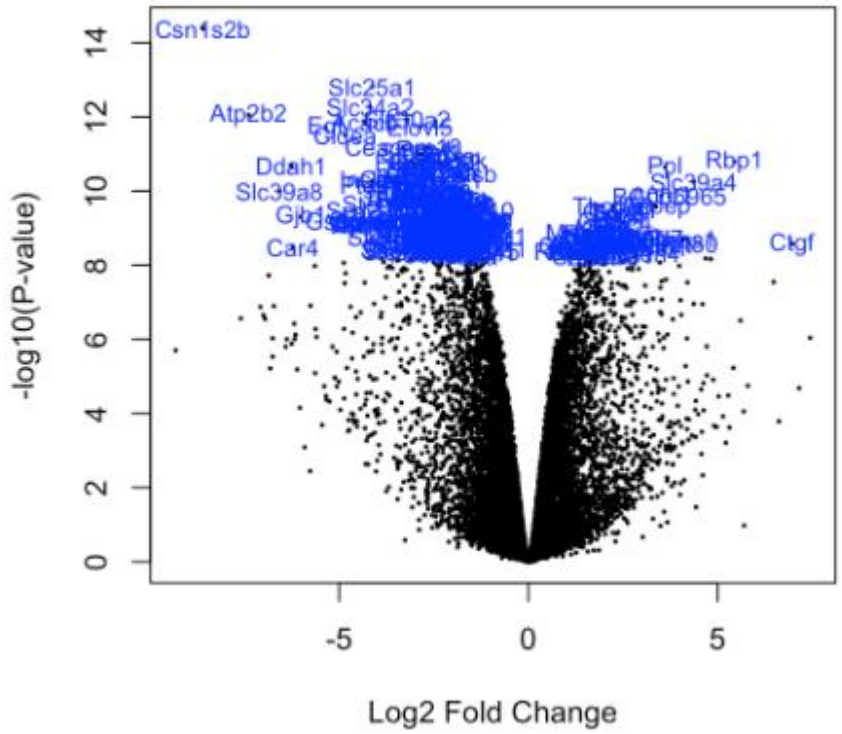
- Expresión diferencial
  - MA plot



Yin et al. Genome Biology 2012, 13:R77

RNA-seq

- Expresión diferencial
  - *Volcano plot*



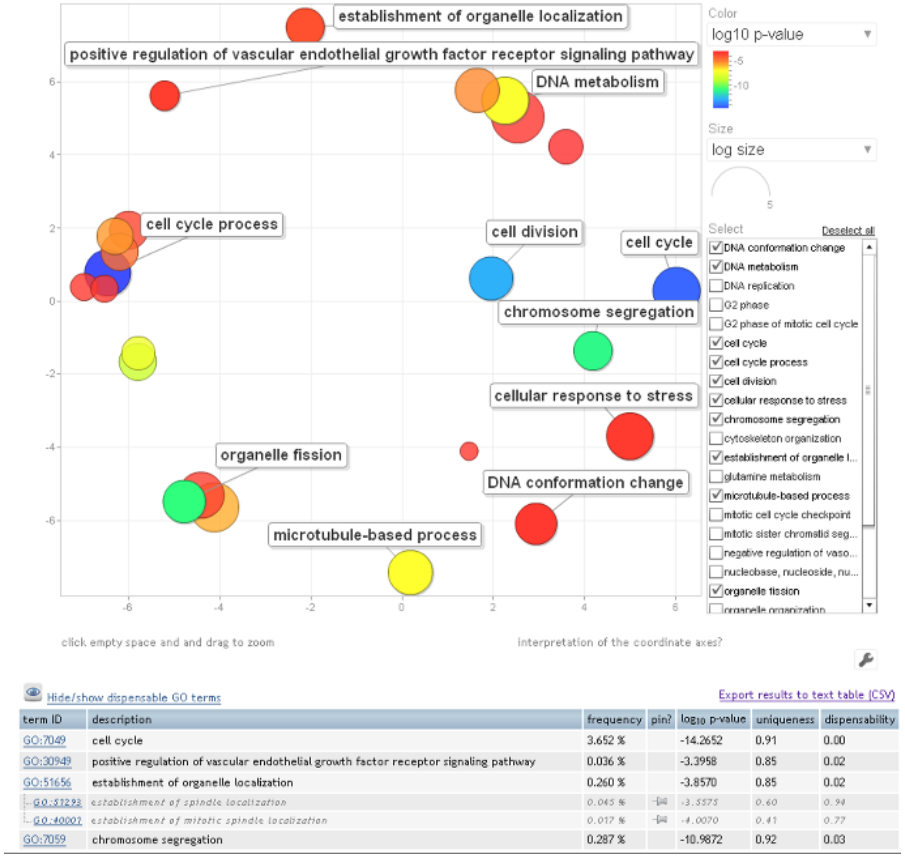
RNA-seq

	A	B	C	D	E	F	G
1	transcript_name	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj
2	30484730	3,144.32943360273	5.04026145518364	0.196928866418254	25.5943252345683	1.7644314301418e-144	4.54023495604088e-140
3	30561786	1,413.35213135671	3.41626574251214	0.13451370327947	25.3971577558488	2.71095008913392e-142	3.4879083846797e-138
4	30504829	880.771473715433	4.32396956861188	0.173416033458984	24.9340818283367	3.17818540613069e-137	2.72603556235183e-133
5	30544189	8,258.20272934674	3.24090885472581	0.130554314699064	24.8242186571643	4.91076326741851e-136	3.15909400993032e-132
6	30483127	1,295.09139872873	6.55681270831358	0.273206432704986	23.9994814301966	2.81566948590279e-127	1.44905614422501e-123
7	30551621	1,511.46596367277	3.3786619006878	0.142163737793915	23.7659895070121	7.50972165530951e-125	3.22066929390707e-121
8	30475802	1,898.44230304467	6.95150465936128	0.296772084097406	23.4237148029047	2.45035765738168e-121	9.00751474853505e-118
9	30480296	1,210.07334890955	3.48353842881072	0.154359693570149	22.5676687238797	9.00712662522495e-113	2.89714227900361e-109
10	30543987	710.020123515779	5.68496687104772	0.252361926081687	22.5270386833534	2.2553967057206e-112	6.44842978128917e-109
11	30517326	3,611.6823066722	3.51111066113919	0.156500582481462	22.4351283903694	1.78792563628964e-111	4.60069024730051e-108
12	30535740	559.981220962047	4.09061739312591	0.182682690256112	22.3919266099654	4.71750598855994e-111	1.1035533099784e-107
13	30544292	1,472.41459635489	4.80156003343999	0.216812606426031	22.1461293814487	1.13676216585688e-108	2.43759700431909e-105
14	30499835	457.358467856158	6.26052947903444	0.2955910524082	21.1796988712259	1.46968872158091e-99	2.90907924490154e-96
15	30548836	1,891.57726951951	5.50297046375875	0.263949523079027	20.8485713463901	1.57041671532408e-96	2.88642592276565e-93
16	30477139	1,224.48273150217	3.84512112082617	0.184655160791858	20.8232529453122	2.66467630015296e-96	4.57116337036906e-93
17	30518455	417.579205665653	4.7316312598701	0.229901300395872	20.5811417844206	4.05028224251837e-94	6.51386641653017e-91
18	30554782	755.049051445421	3.60242774699136	0.180563057495959	19.9510785702772	1.46689640228573e-88	2.22036342491861e-85
19	30487155	1,808.40178314046	4.28163601075905	0.215972721453325	19.8248926158222	1.81568244319052e-87	2.59561892378769e-84
20	30499588	608.66999182642	7.19339714094185	0.363520473282165	19.78814859035	3.76624508057103e-87	5.10068517964494e-84

RNA-seq

	A	B	C	D	E	F	G
1	transcript_name	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj
2	30484730	3,144.32943360273	5.04026145518364	0.196928866418254	25.5943252345683	1.7644314301418e-144	4.54023495604088e-140
3	30561786	1,413.35213135671	3.41626574251214	0.13451370327947	25.3971577558488	2.71095008913392e-142	3.48790838467 97e-138
4	30504829	880.771473715433	4.32396956861188	0.173416033458984	24.9340818283367	3.17818540613069e-137	2.72603556235183e-133
5	30544189	8,258.20272934674	3.24090885472581	0.130554314699064	24.8242186571643	4.91076326741851e-136	3.15909400993032e-132
6	30483127	1,295.09139872873	6.55681270831358	0.273206432704986	23.9994814301966	2.81566948590279e-127	1.44905614422501e-123
7	30551621	1,511.46596367277	3.3786619006878	0.142163737793915	23.7659895070121	7.50972165530951e-125	3.22066929390707e-121
8	30475802	1,898.44230304467	6.95150465936128	0.296772084097406	23.4237148029047	2.45035765738168e-121	9.00751474853505e-118
9	30480296	1,210.07334890955	3.48353842881072	0.154359693570149	22.5676687238797	9.00712662522495e-113	2.89714227900361e-109
10	30543987	710.020123515779	5.68496687104772	0.252361926081687	22.5270386833534	2.2553967057206e-112	6.44842978128917e-109
11	30517326	3,611.6823066722	3.51111066113919	0.156500582481462	22.4351283903694	1.78792563628964e-111	4.60069024730051e-108
12	30535740	559.981220962047	4.09061739312591	0.182682690256112	22.3919266099654	4.71750598855994e-111	1.1035533099784e-107
13	30544292	1,472.41459635489	4.80156003343999	0.216812606426031	22.1461293814487	1.13676216585688e-108	2.43759700431909e-105
14	30499835	457.358467856158	6.26052947903444	0.2955910524082	21.1796988712259	1.46968872158091e-99	2.90907924490154e-96
15	30548836	1,891.57726951951	5.50297046375875	0.263949523079027	20.8485713463901	1.57041671532408e-96	2.88642592276565e-93
16	30477139	1,224.48273150217	3.84512112082617	0.184655160791858	20.8232529453122	2.66467630015296e-96	4.57116337036906e-93
17	30518455	417.579205665653	4.7316312598701	0.229901300395872	20.5811417844206	4.05028224251837e-94	6.51386641653017e-91
18	30554782	755.049051445421	3.60242774699136	0.180563057495959	19.9510785702772	1.46689640228573e-88	2.22036342491861e-85
19	30487155	1,808.40178314046	4.28163601075905	0.215972721453325	19.8248926158222	1.81568244319052e-87	2.59561892378769e-84
20	30499588	608.66999182642	7.19339714094185	0.363520473282165	19.78814859035	3.76624508057103e-87	5.10068517964494e-84

- Análisis de sobrerrepresentación de términos de ontología



Hide/show dispensable GO terms

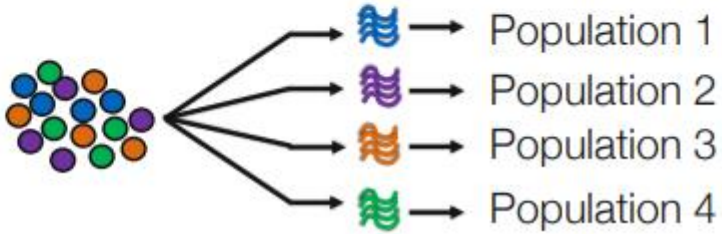
Export results to text table (CSV)

term ID	description	frequency	pin?	$\log_{10}$ p-value	uniqueness	dispensability
GO:7049	cell cycle	3.652 %		-14.2652	0.91	0.00
GO:30949	positive regulation of vascular endothelial growth factor receptor signaling pathway	0.036 %		-3.3958	0.85	0.02
GO:51656	establishment of organelle localization	0.260 %		-3.8570	0.85	0.02
GO:31293	*establishment of spindle localization*	*0.045 %*		*-3.3575*	*0.60*	*0.94*
GO:40001	*establishment of mitotic spindle localization*	*0.017 %*		*-4.0070*	*0.41*	*0.77*
GO:7059	chromosome segregation	0.287 %		-10.9872	0.92	0.03

RNA-seq de célula única

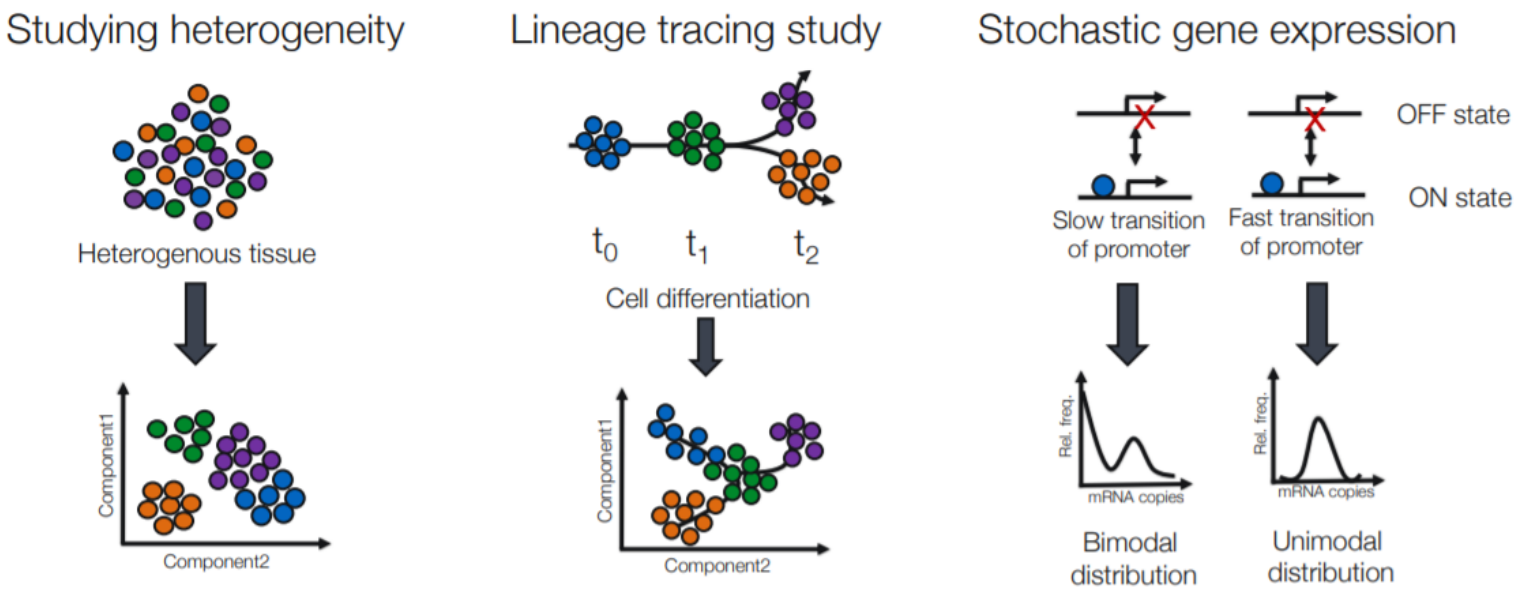
Bulk RNA-seq  average expression level

- comparative transcriptomics
- disease biomarker
- homogenous systems

scRNA-seq 

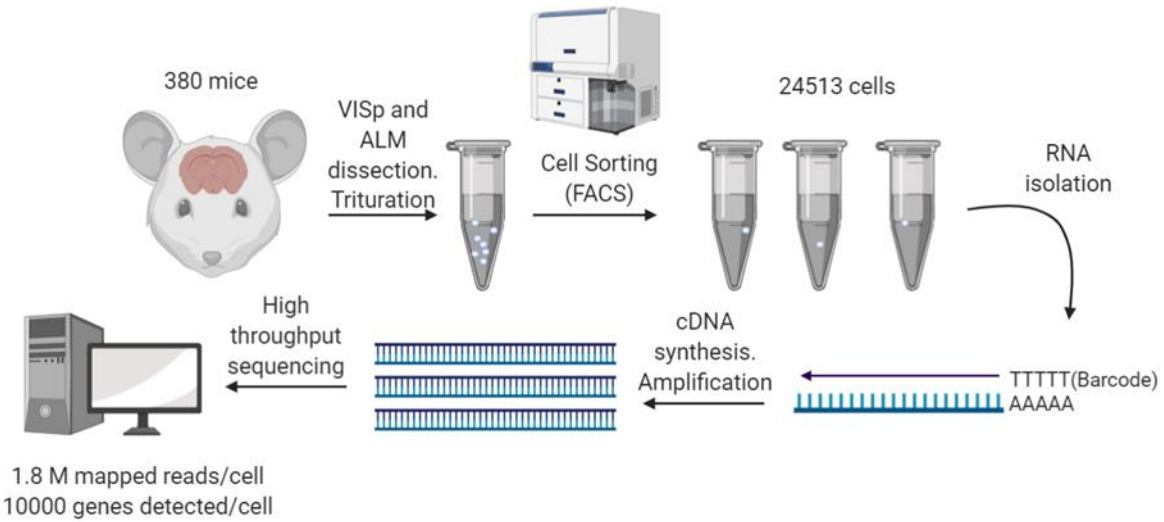
- define heterogeneity
- identify rare cell population
- cell population dynamics

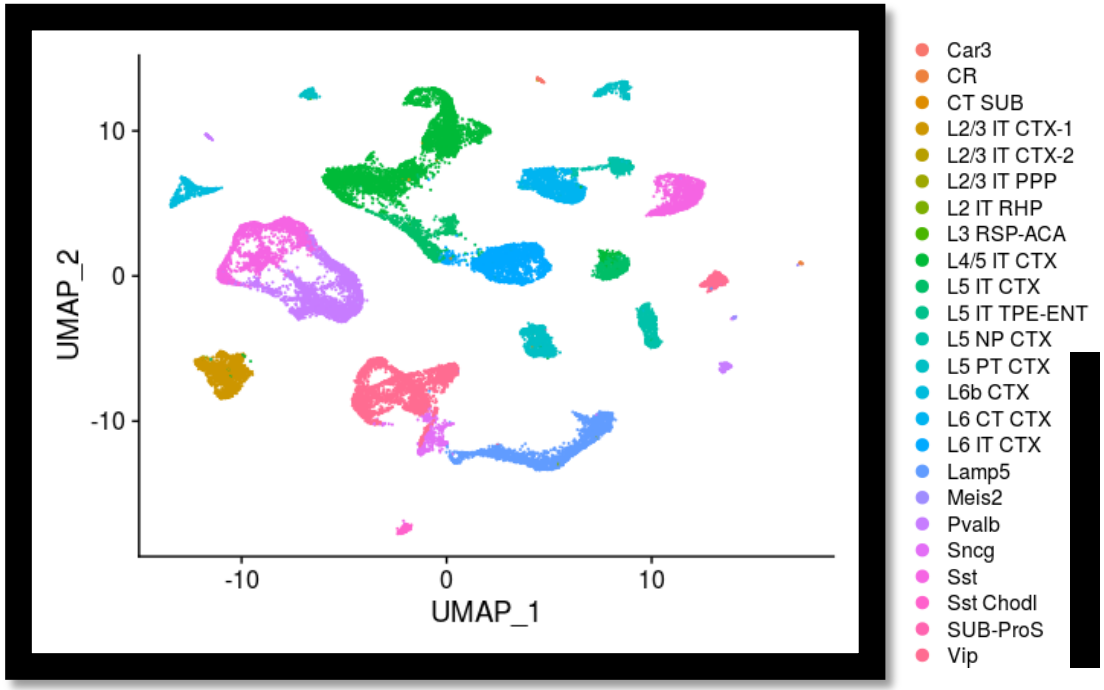
RNA-seq de célula única



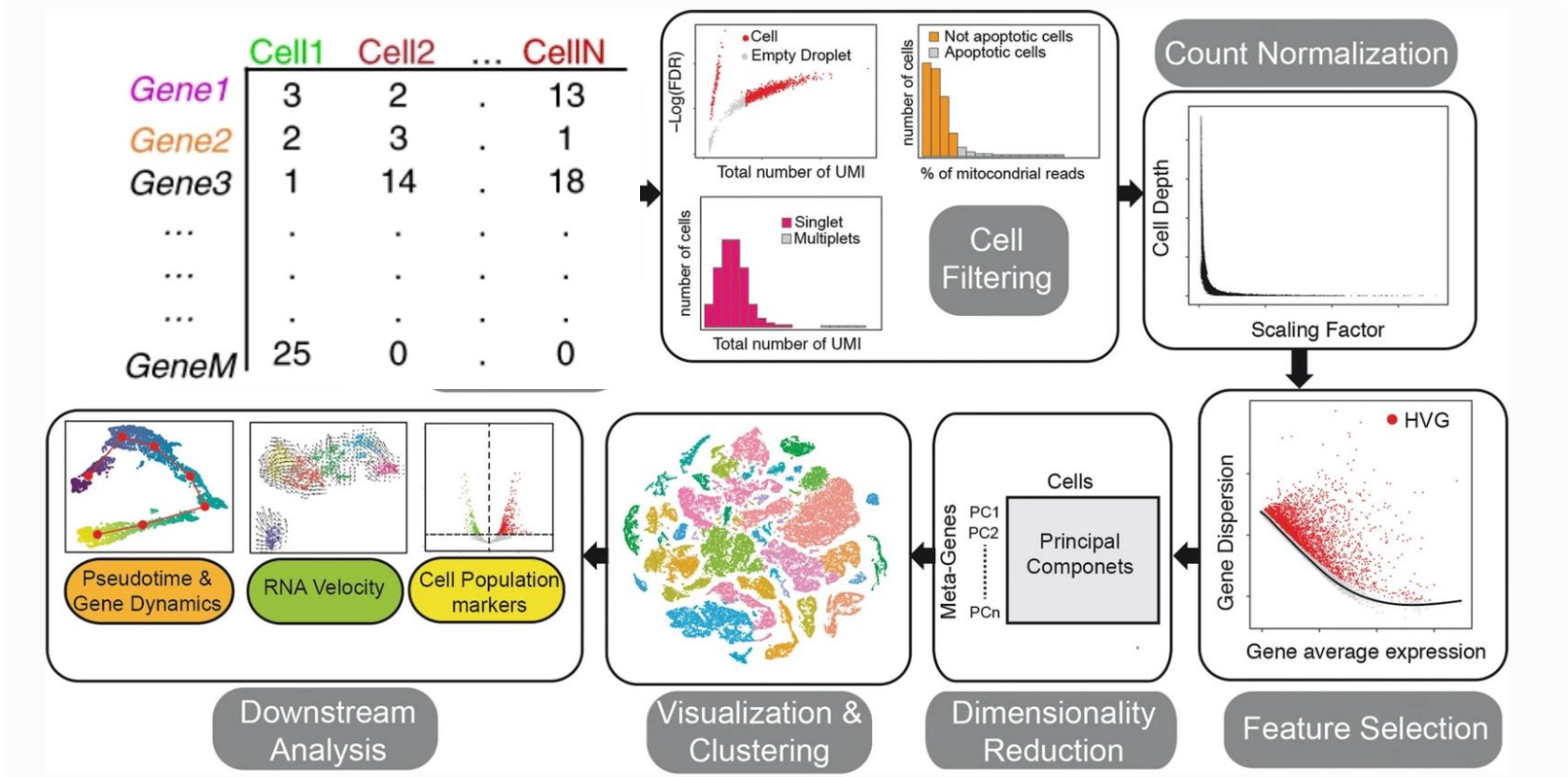
Liu S and Trapnell C. Single-cell transcriptome sequencing: recent advances and remaining challenges, F1000 Research 2016 (doi: 10.12688/f1000research.7223.1)
Junker and van Oudenaarden; Every Cell Is Special: Genome-wide Studies Add a New Dimension to Single-Cell Biology, Cell 2014 (doi: 10.1016/j.cell.2014.02.010)

RNA-seq de célula única



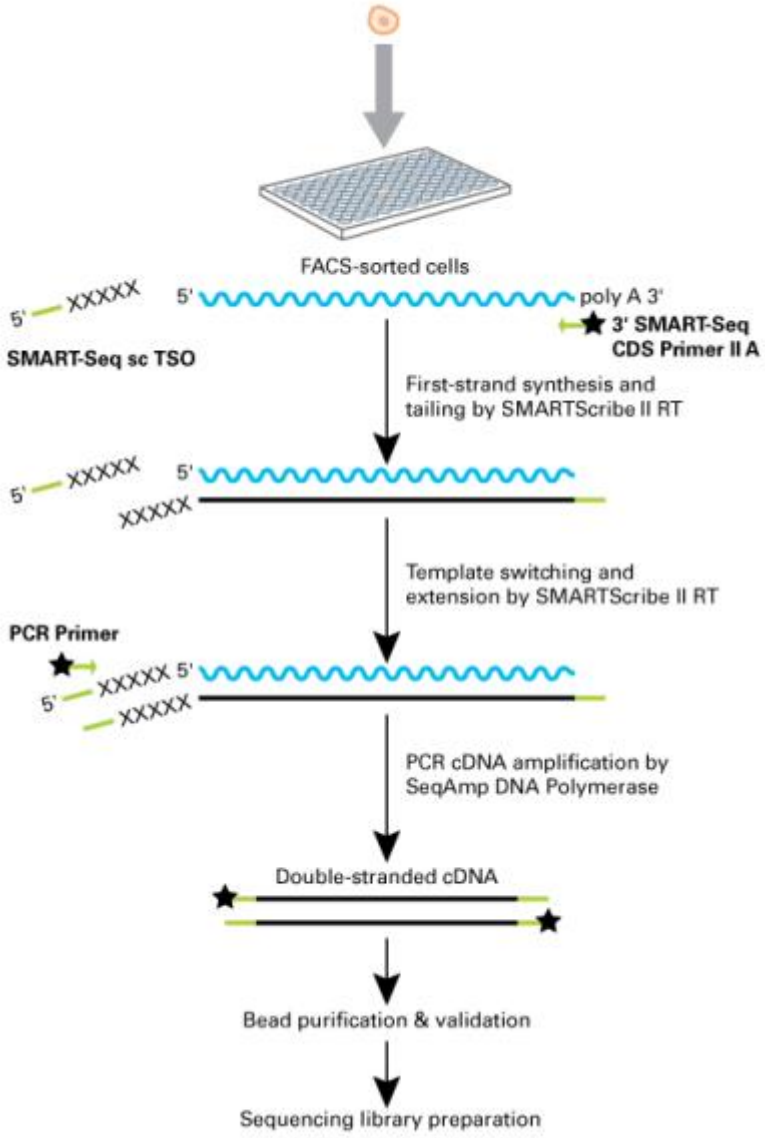


RNA-seq de célula única



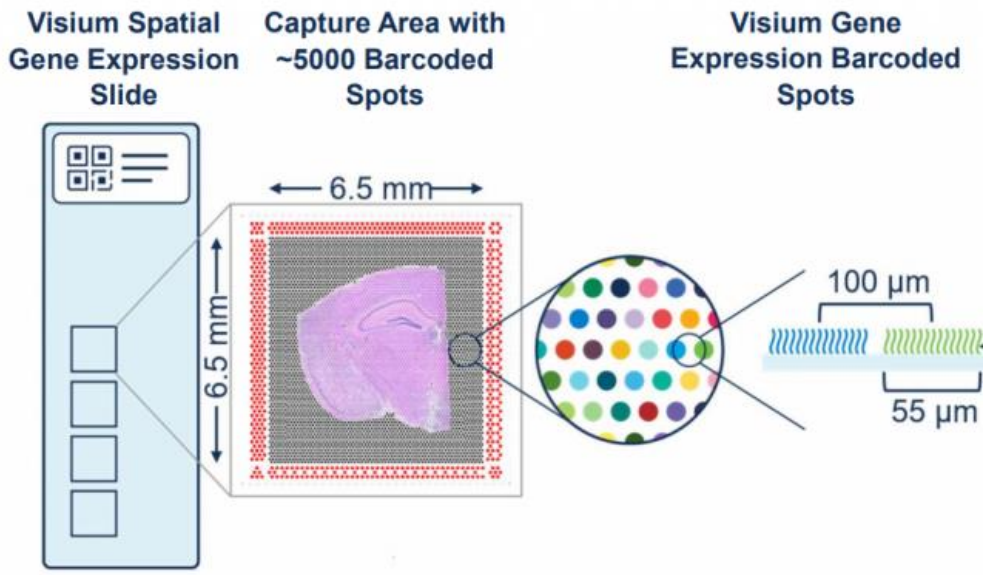
RNA-seq de célula única

- Smart-seq

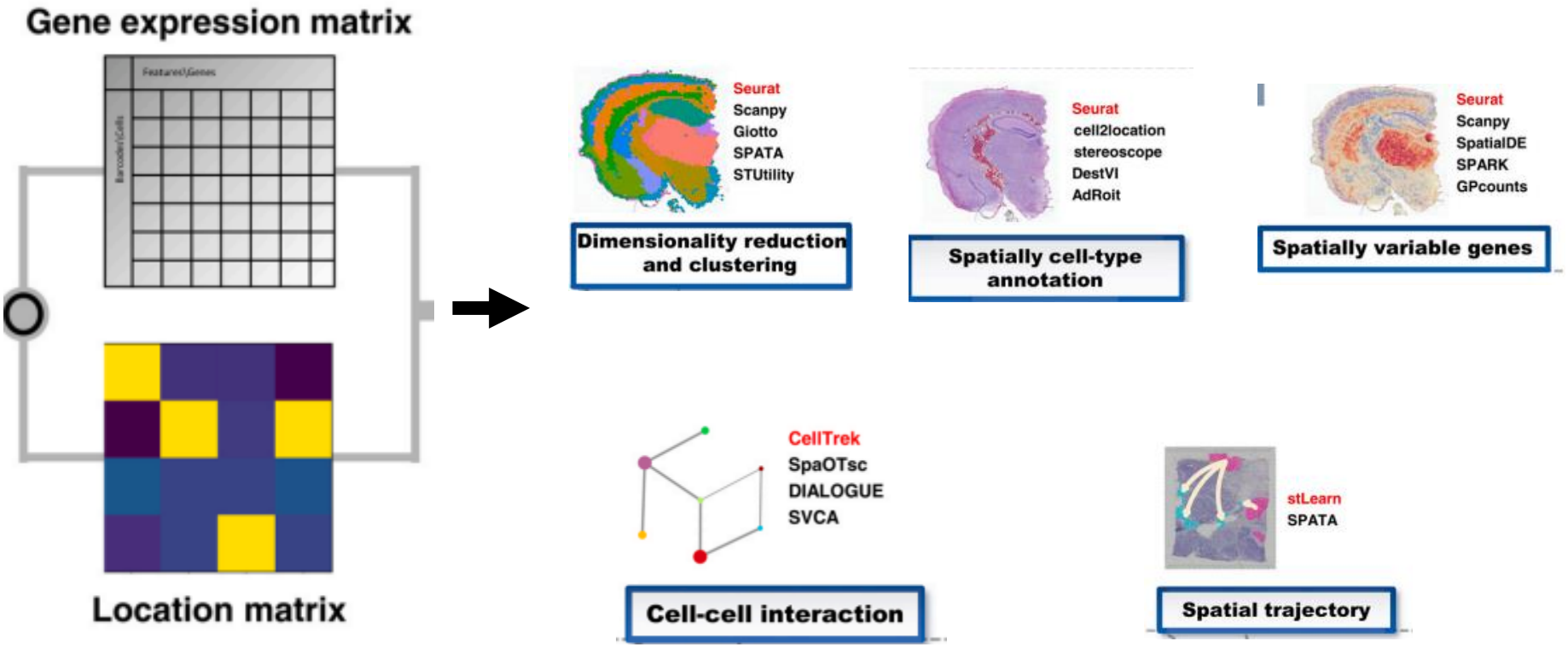


Transcriptómica espacial

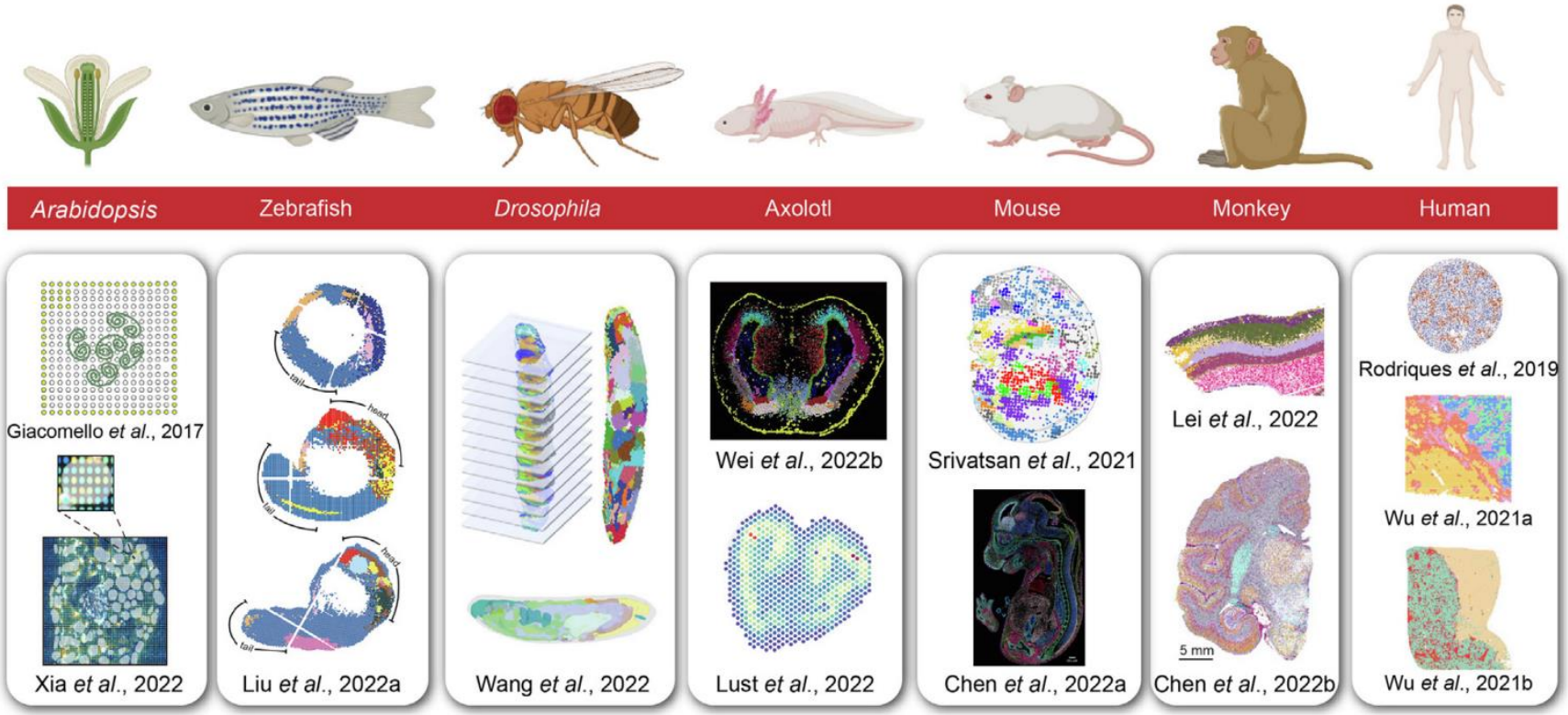
- Combina RNA-seq de pequeña región (célula única) con coordenadas
- Visium (10x)
  - Método más utilizado de transcriptómica especial
  - Lisis y transferencia de ARNs a sitios de captura
  - La resolución viene dada por el tamaño del sitio de captura.



Transcriptómica espacial



Transcriptómica espacial



Cheng et al, 2022

Preguntas????