TP 4: Transcriptómica

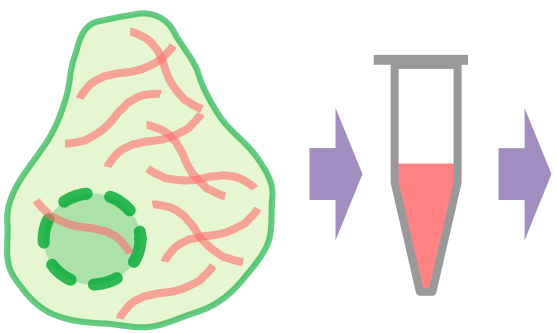
Biología Celular y Molecular
Lic. en Ciencias Biológicas
Universidad Nacional del Comahue

Datos de RNAseq

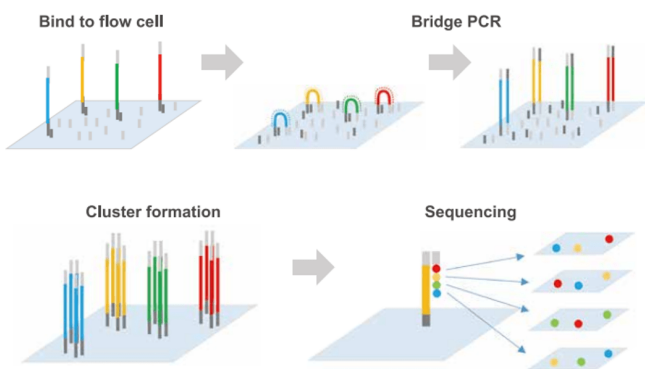
Generación

https://blogs.iu.edu/ncgas/2021/09/17/introduction-to-illumina-sequencing/

Extracción de RNA



Secuenciación



Datos de secuencia

archivo.FASTA

```
AGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTC
TAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCGTACAA
CTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTT
ACTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCGTA
CTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCG
TAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGT
CTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCGTACAA
CAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCG
AAGGAAGAAGACTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGG
```

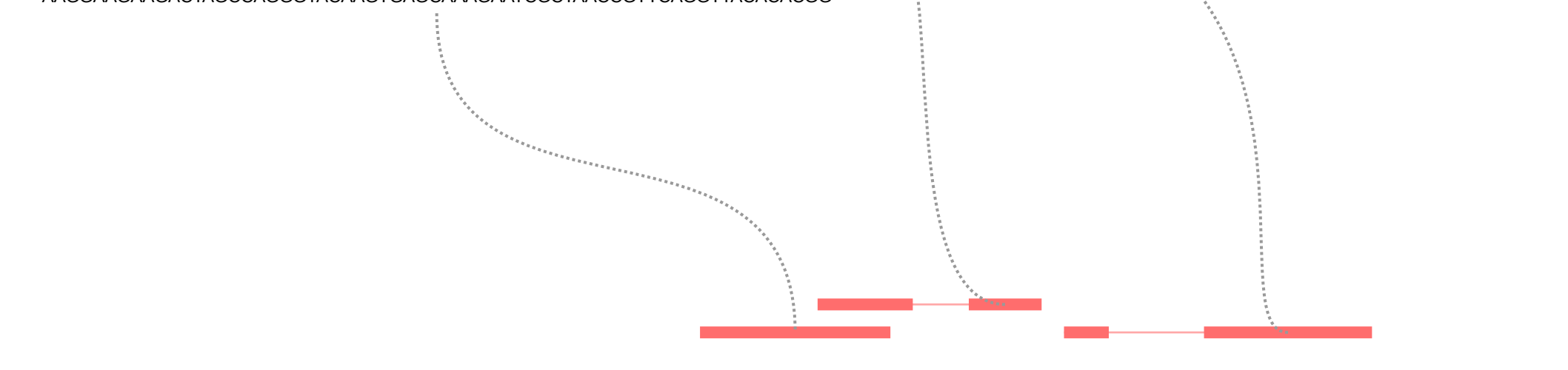
Datos de RNAseq

Alineamiento

TAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCGTACAA

ACTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCGTA

AAGGAAGAAGACTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGG



Datos de RNAseq

Alineamiento

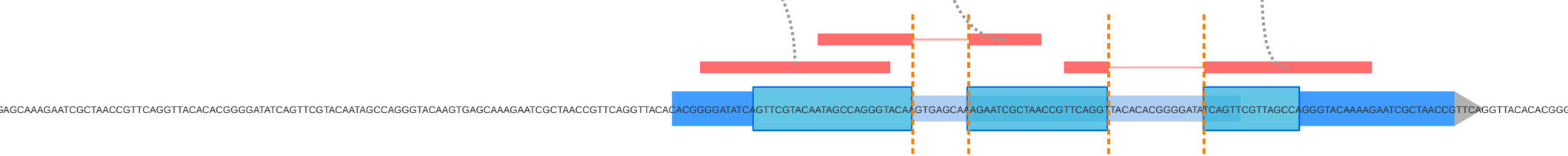
Mediante variados sofware se obtiene un MODELO DEL GEN:

MAPEADO + PREDICCIÓN DE ORFs

TAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCGTACAA

ACTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCGTA

AAGGAAGAAGACTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGG



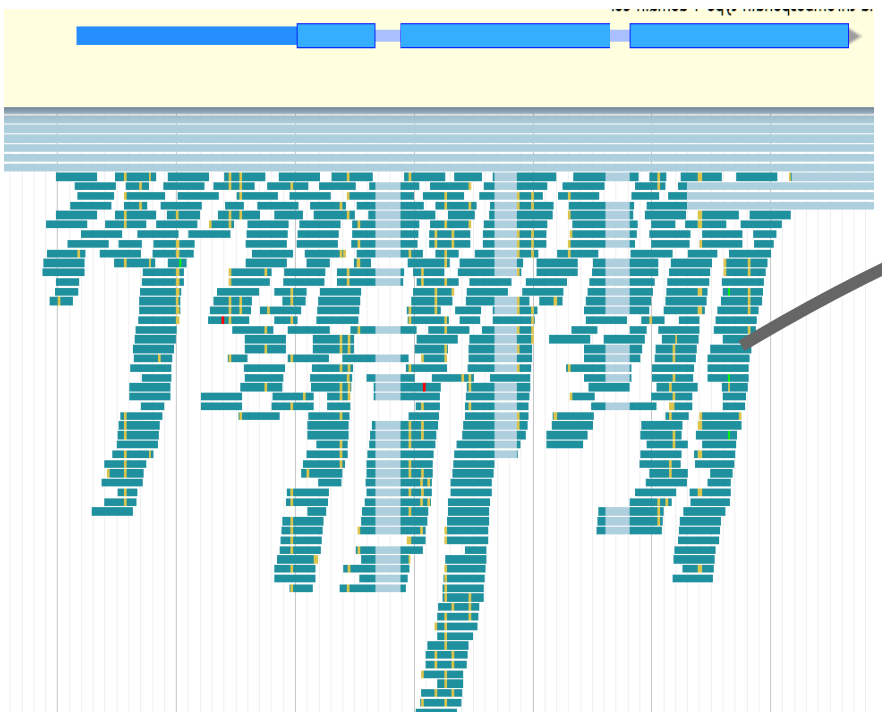
Identificación de regiones codificantes

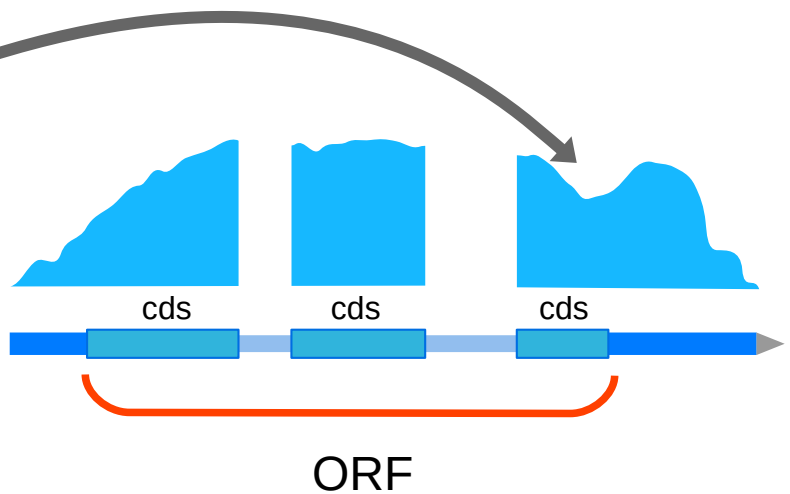
Datos de RNAseq

Alineamiento

Mediante variados sofware se obtiene un MODELO DEL GEN:

MAPEADO + PREDICCIÓN DE ORFs



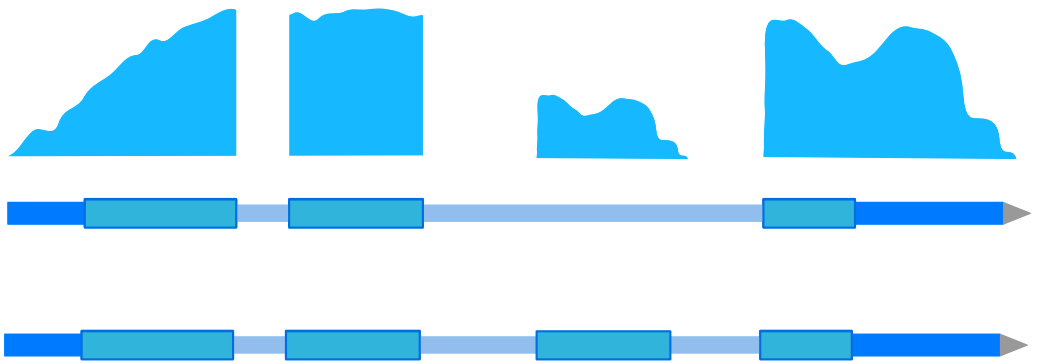


Datos de RNAseq

Modelos de genes

Mediante variados sofware se obtiene un MODELO DEL GEN:

MAPEADO + PREDICCIÓN DE ORFs



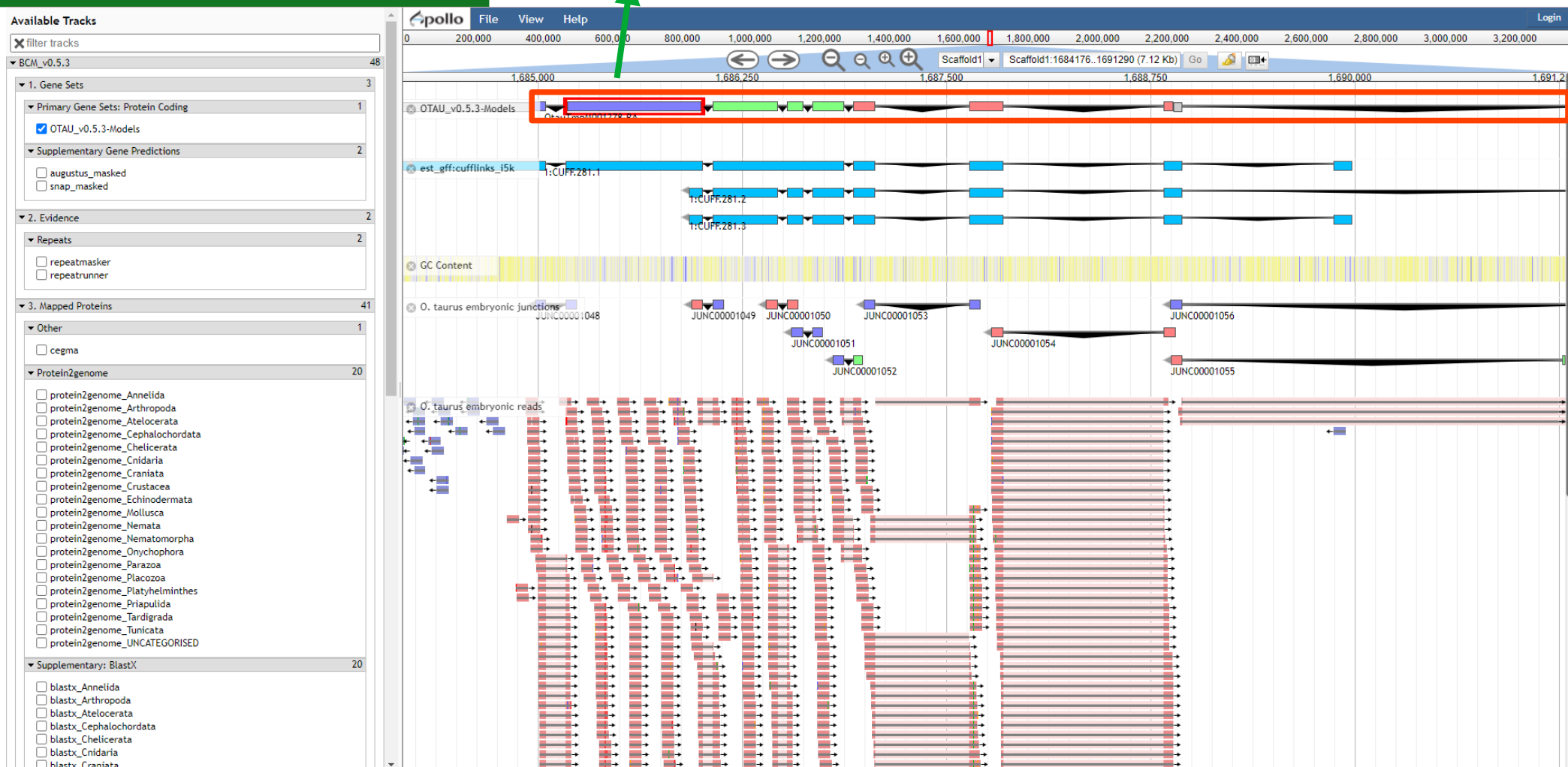
Isoformas

Se pueden obtener mas de un modelo de gen para una misma región del genoma.

Datos de RNAseq

Browser Apollo

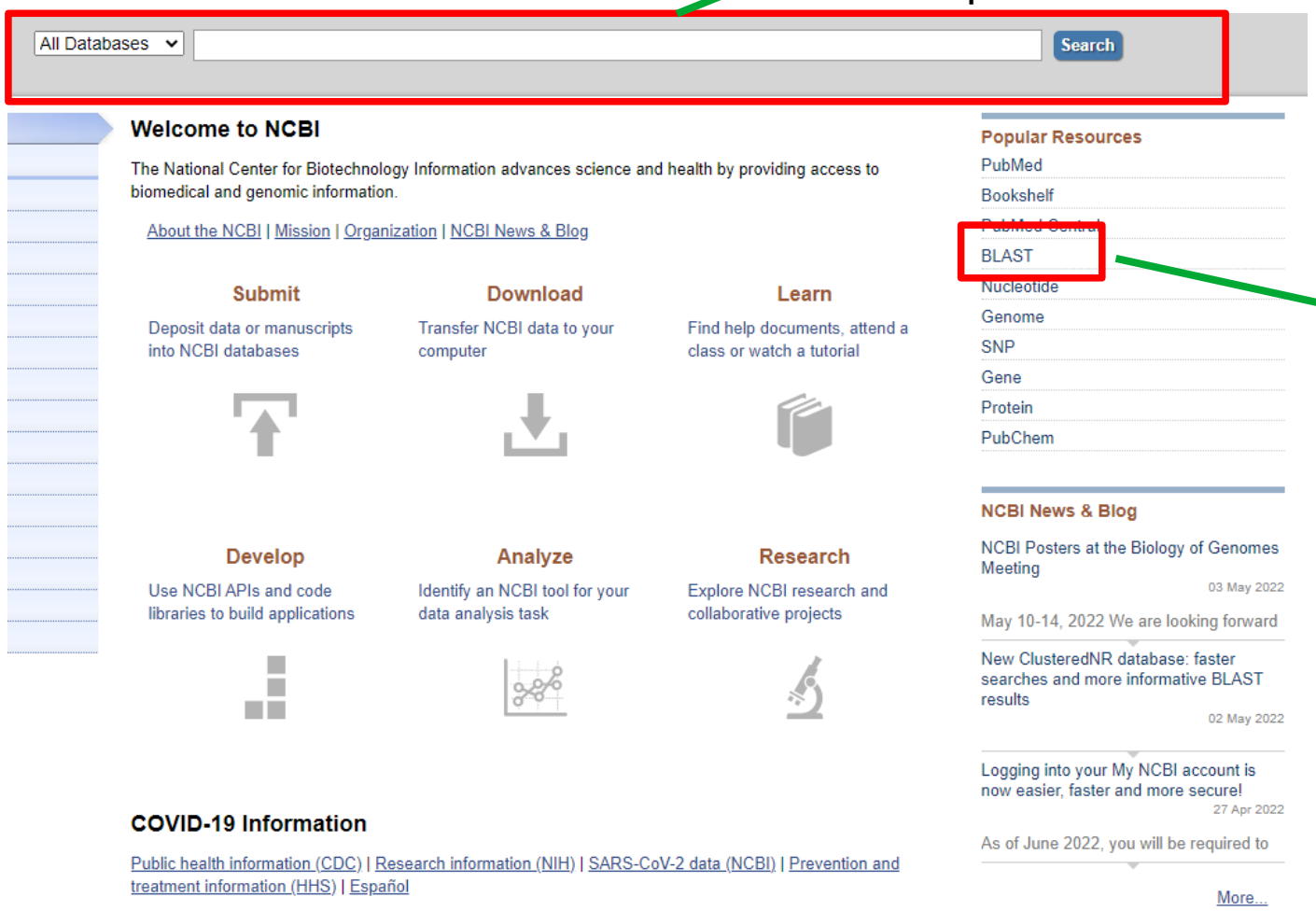
Se puede extraer la secuencia del gen



Datos de RNAseq

NCBI

Pueden buscar desde papers, hasta genomas y transcriptomas

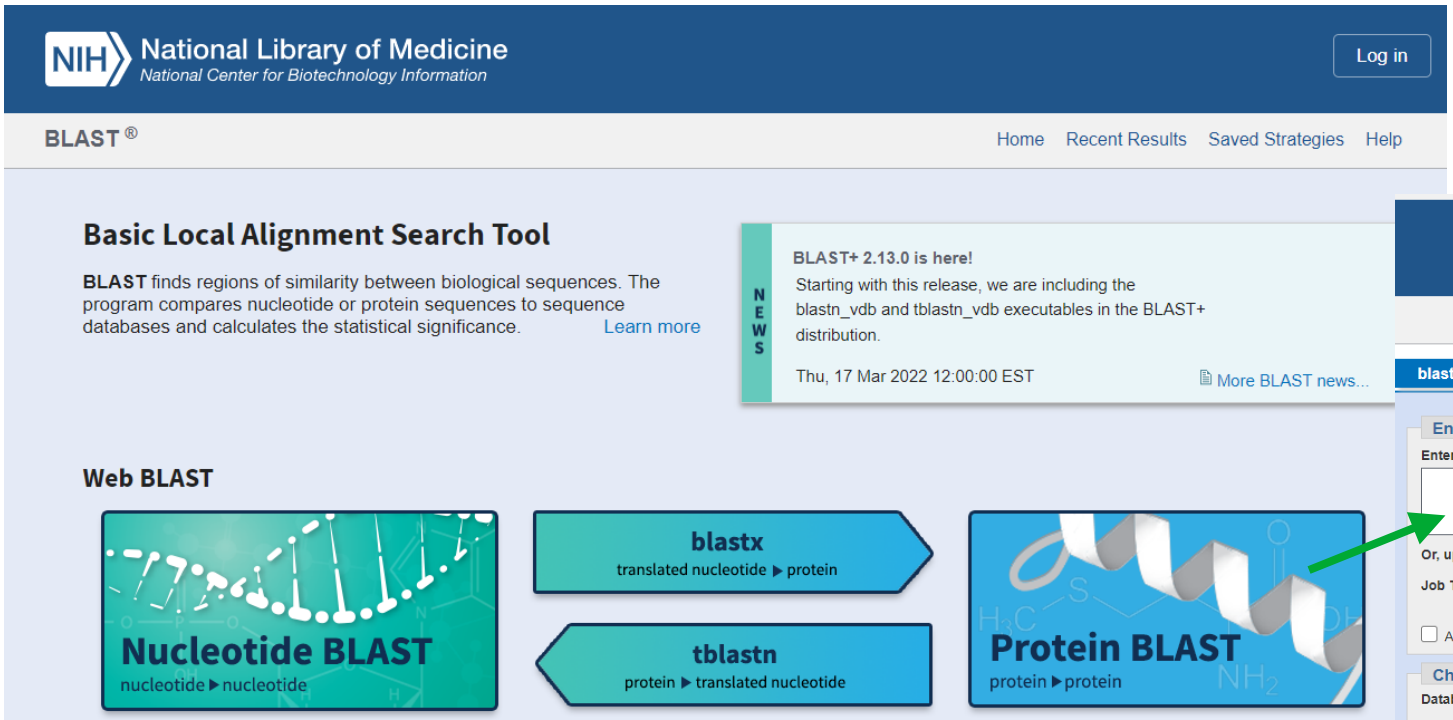


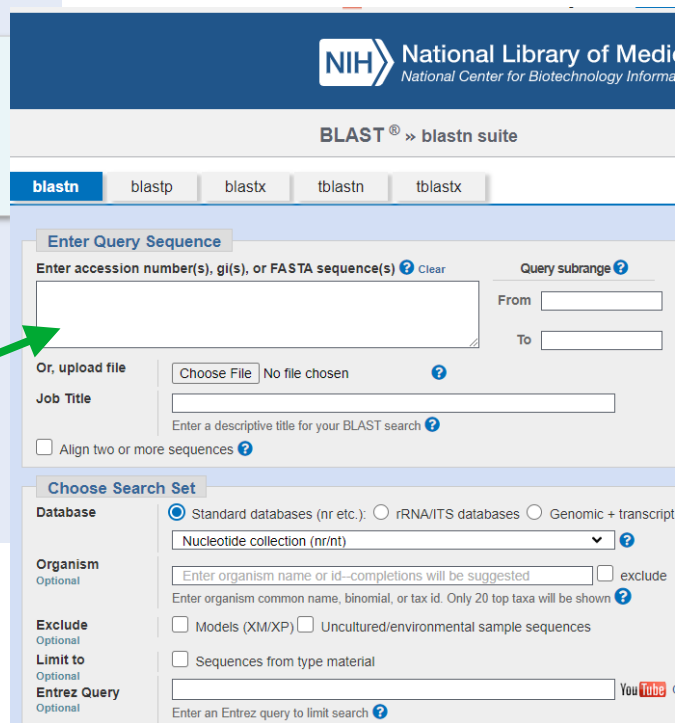
Una de las herramientas más populares es el BLAST

Datos de RNAseq

Blast

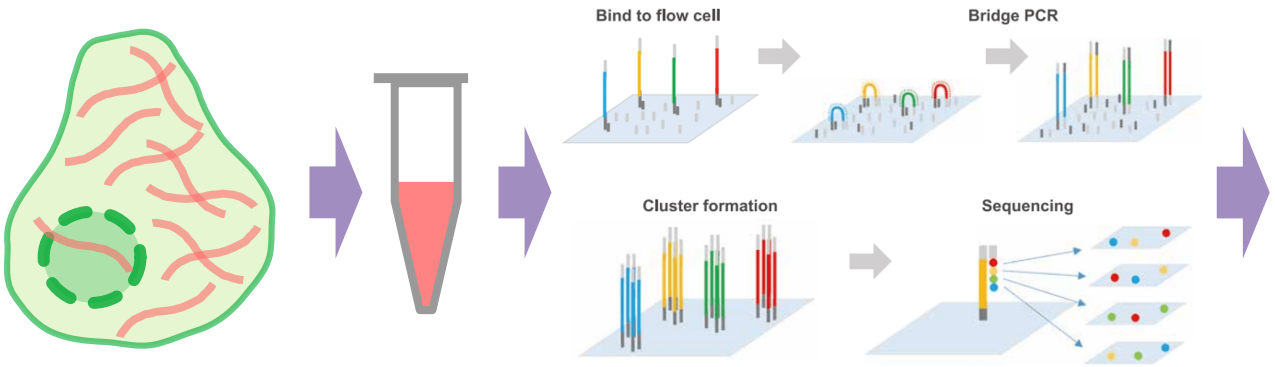
“Basic local alignment search tool”



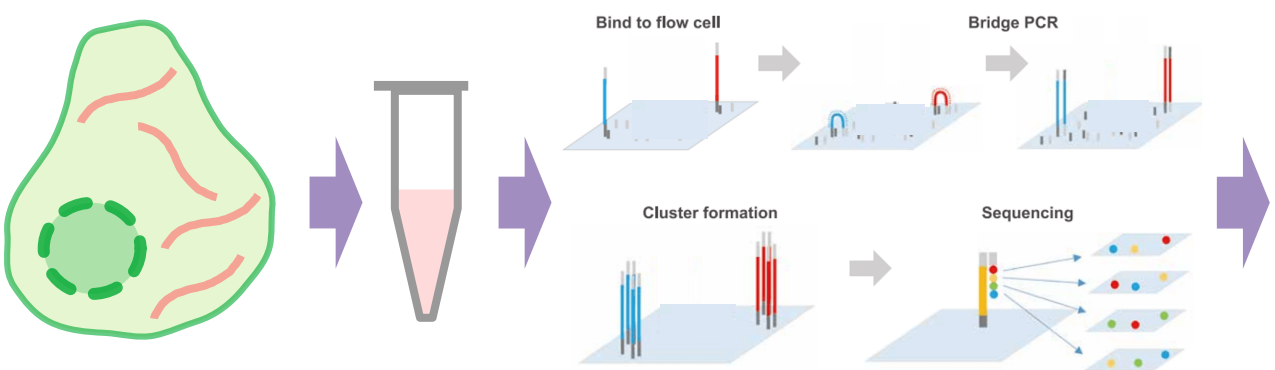


Datos de RNAseq

Información de expresión



AGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTC
TAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCGTACAA
CTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTT
ACTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCGTA
CTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCG
TAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGT
CTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCGTACAA
CAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCG
AAGGAAGAAGACTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGG



TAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCGTACAA

ACTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGGGGATATCAGTTCGTA

AAGGAAGAAGACTAGCCAGGGTACAAGTGAGCAAAGAATCGCTAACCGTTCAGGTTACACACGG

Datos de RNAseq

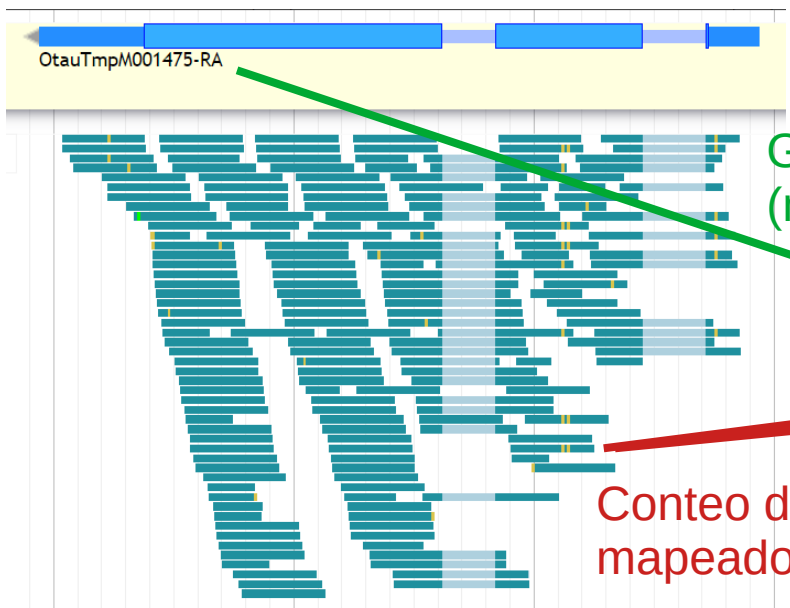
Información de expresión

Datos de RNAseq

Información de expresión

Muestras

...



Gen (modelo)

Conteo de reads mapeados

	gene	ctrlRNAi_M.BRN_S19	ctrlRNAi_M.BRN_S20	ctrlRNAi_M
1	OTAU000001	166	155	
2	OTAU000002	955	975	
3	OTAU000003	2318	2228	
4	OTAU000004	2	1	
5	OTAU000005	29	26	
6	OTAU000006	97	78	
7	OTAU000007	0	0	
8	OTAU000008	1841	1902	
9	OTAU000009	730	549	
10	OTAU000010	490	535	
11	OTAU000011	1	4	
12	OTAU000012	21	14	
13	OTAU000013	25	29	
14	OTAU000014	39	35	
15	OTAU000015	3625	3178	
16	OTAU000016	1130	1115	
17	OTAU000017	638	607	
18	OTAU000018	304	242	
19	OTAU000019	15	10	
20	OTAU000020	0	0	

Datos de RNAseq

Matriz de diseño

Factores

Reads totales

IDs de las muestras

	sampleID	Tissue	Treatment	Sex	genome_unique	X_unique
1	ctrlRNAi_M.BRN_S19	BRN	ctr	M	10748437	8861358
2	ctrlRNAi_M.BRN_S20	BRN	ctr	M	10129825	8339020
3	ctrlRNAi_M.BRN_S21	BRN	ctr	M	10571624	8676801
4	ctrlRNAi_M.BRN_S22	BRN	ctr	M	10899420	9015026
5	ctrlRNAi_M.BRN_S23	BRN	ctr	M	10275664	8460658
6	ctrlRNAi_M.BRN_S24	BRN	ctr	M	10281103	8359537
7	ctrlRNAi_M.THE_S43	THE	ctr	M	10175292	9123613
8	ctrlRNAi_M.THE_S44	THE	ctr	M	4640668	4110648
9	ctrlRNAi_M.THE_S45	THE	ctr	M	10431983	9411583
10	ctrlRNAi_M.THE_S46	THE	ctr	M	10239974	9337273
11	ctrlRNAi_M.THE_S47	THE	ctr	M	10954425	9844328
12	ctrlRNAi_M.THE_S48	THE	ctr	M	9431346	8379494
13	ctrlRNAi_M.CHE_S67	CHE	ctr	M	10063065	8960920
14	ctrlRNAi_M.CHE_S68	CHE	ctr	M	10213914	9003514
15	ctrlRNAi_M.CHE_S69	CHE	ctr	M	14606729	13274821
16	ctrlRNAi_M.CHE_S70	CHE	ctr	M	10461943	9431847
17	ctrlRNAi_M.CHE_S71	CHE	ctr	M	15317489	13661185
18	ctrlRNAi_M.CHE_S72	CHE	ctr	M	13292199	11776005
19	ctrlRNAi_M.GEN_S91	GEN	ctr	M	17698786	15493944
20	ctrlRNAi_M.GEN_S92	GEN	ctr	M	9542977	8330501

...

	gene	ctrlRNAi_M.BRN_S19	ctrlRNAi_M.BRN_S20	ctrlRNAi_M
1	OTAU000001	166	155	
2	OTAU000002	955	975	
3	OTAU000003	2318	2228	
4	OTAU000004	2	1	
5	OTAU000005	29	26	
6	OTAU000006	97	78	
7	OTAU000007	0	0	
8	OTAU000008	1841	1902	
9	OTAU000009	730	549	
10	OTAU000010	490	535	
11	OTAU000011	1	4	
12	OTAU000012	21	14	
13	OTAU000013	25	29	
14	OTAU000014	39	35	
15	OTAU000015	3625	3178	
16	OTAU000016	1130	1115	
17	OTAU000017	638	607	
18	OTAU000018	304	242	
19	OTAU000019	15	10	
20	OTAU000020	0	0	

Tabla de conteos

Datos de RNAseq

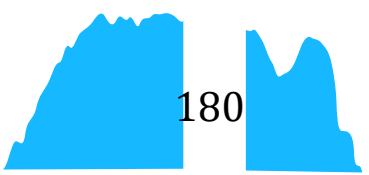
Normalización

Conteos “crudos”

Muestra 1

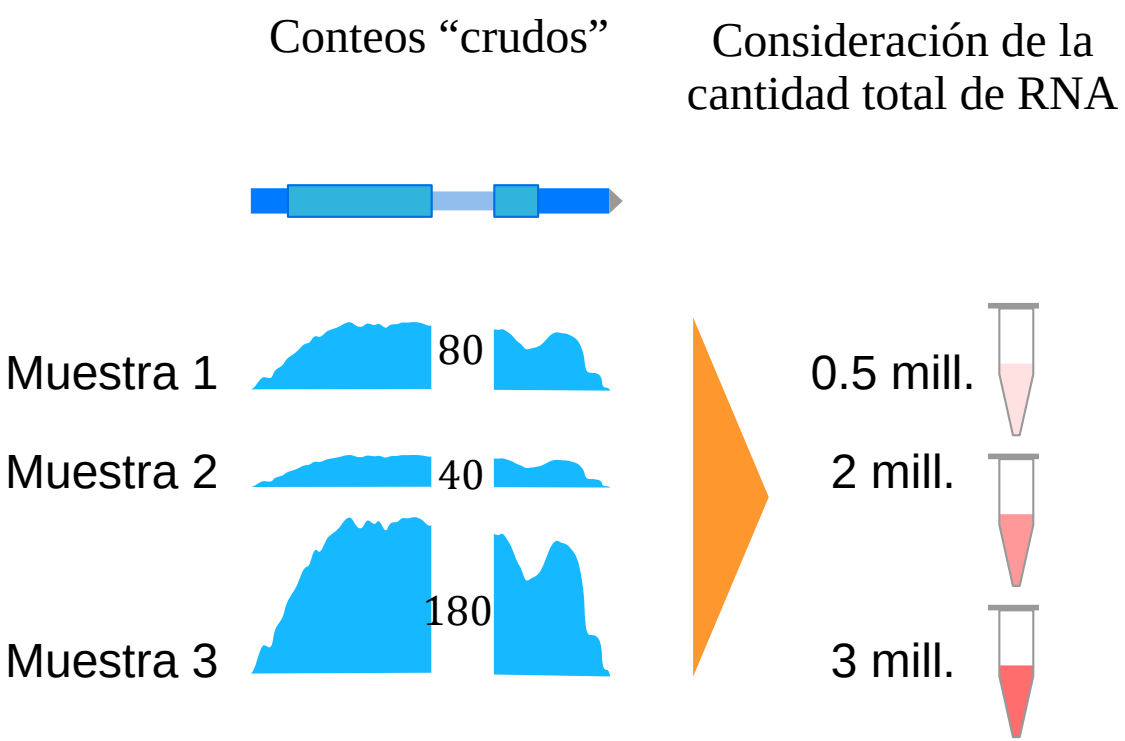
Muestra 2

Muestra 3



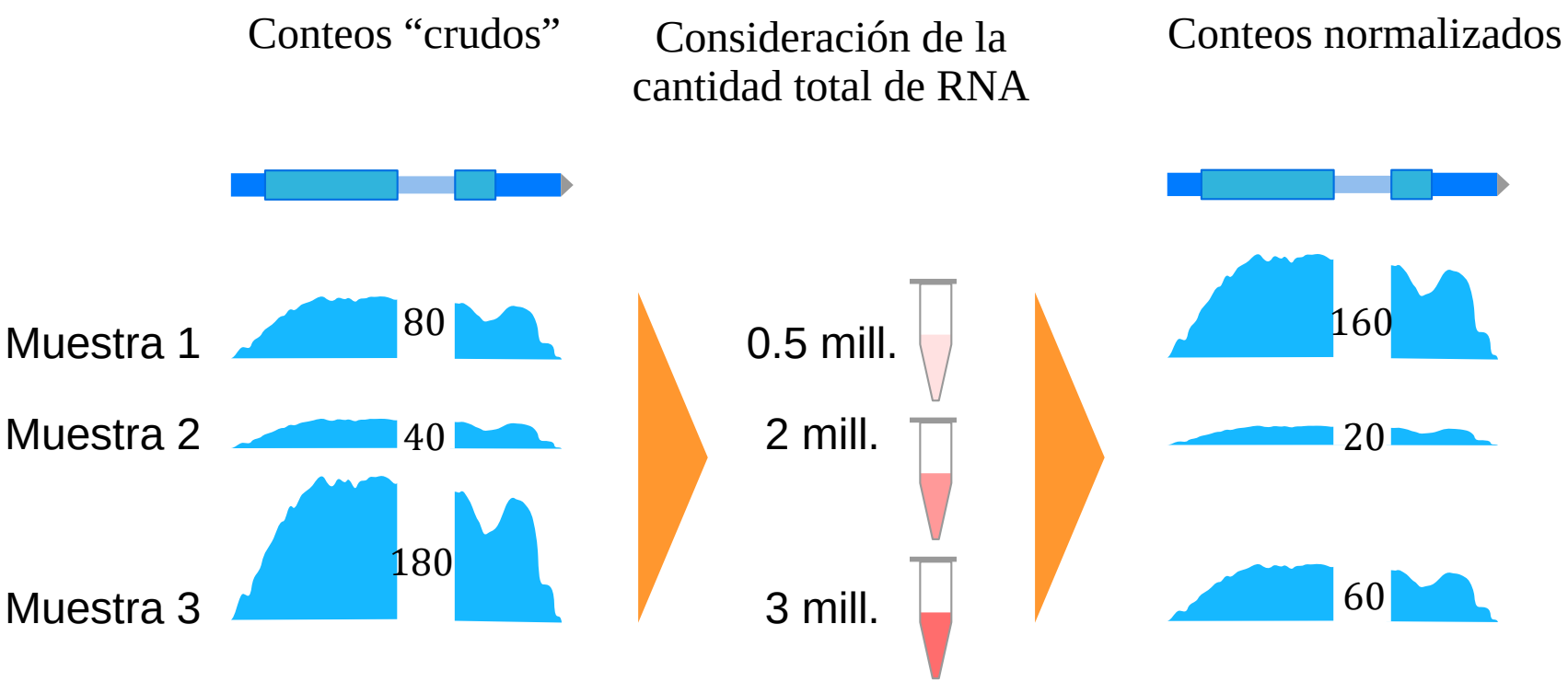
Datos de RNAseq

Normalización



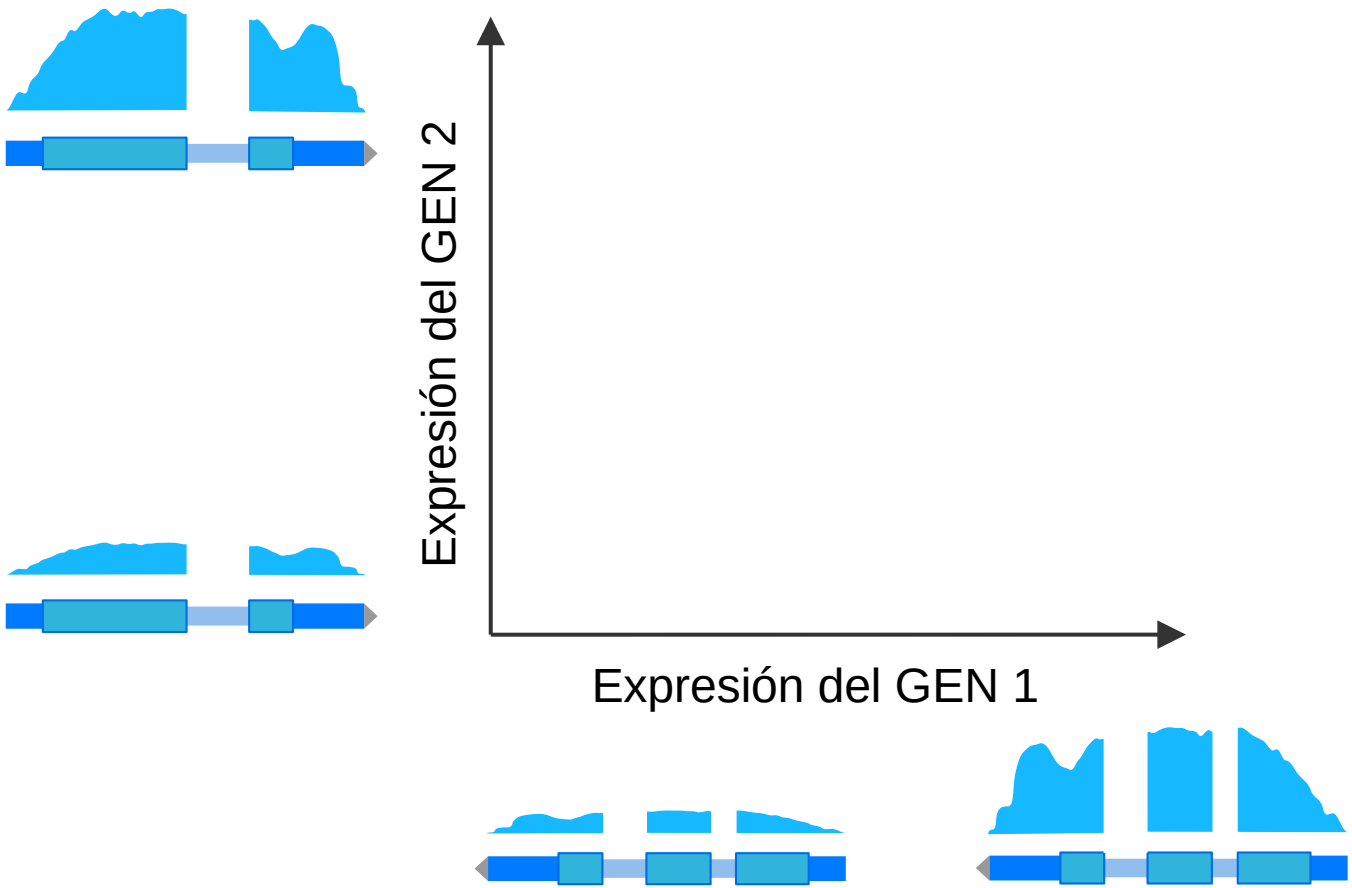
Datos de RNAseq

Normalización



Análisis de datos de RNAseq

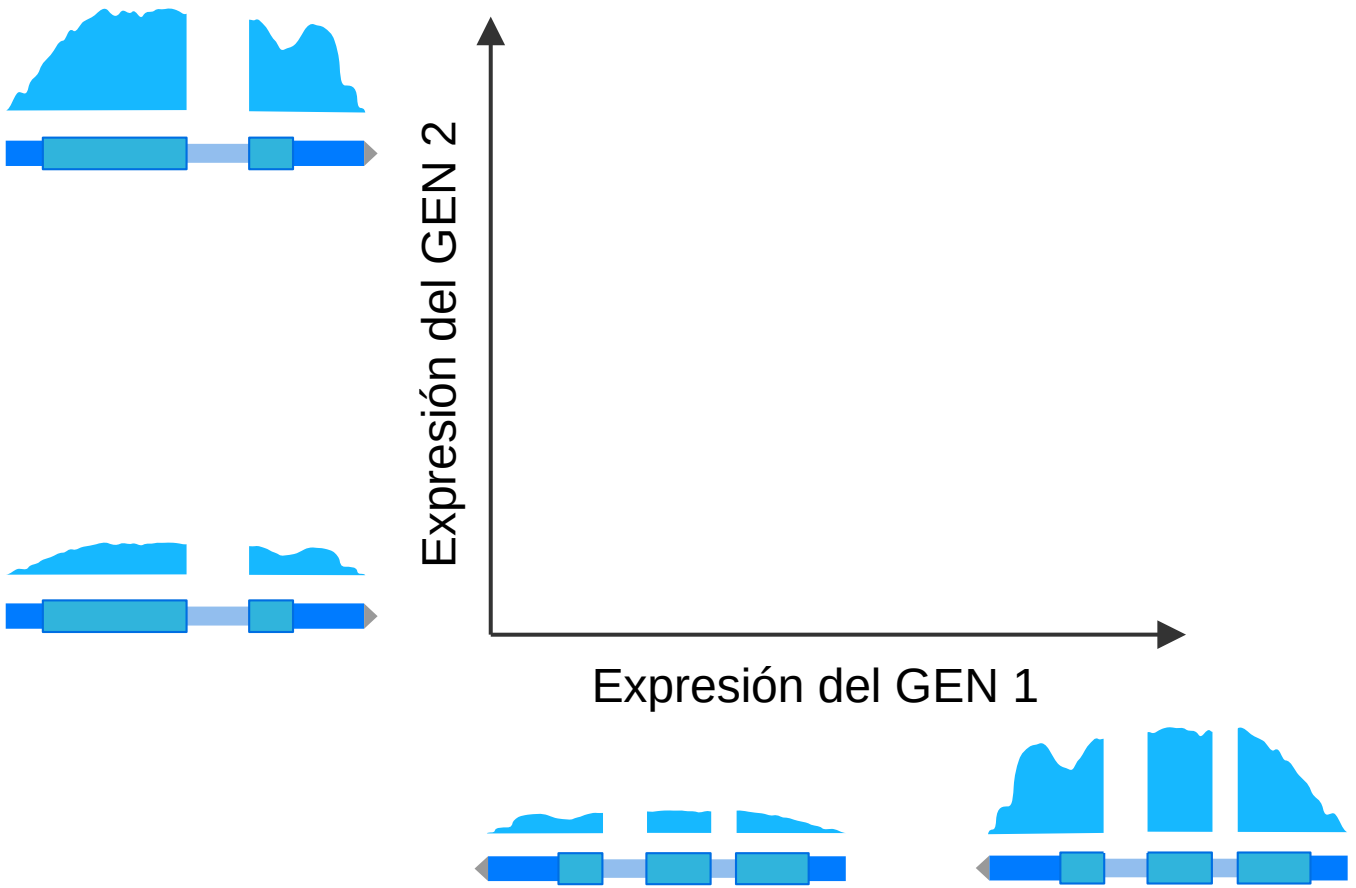
Expresión génica



Análisis de datos de RNAseq

Expresión génica

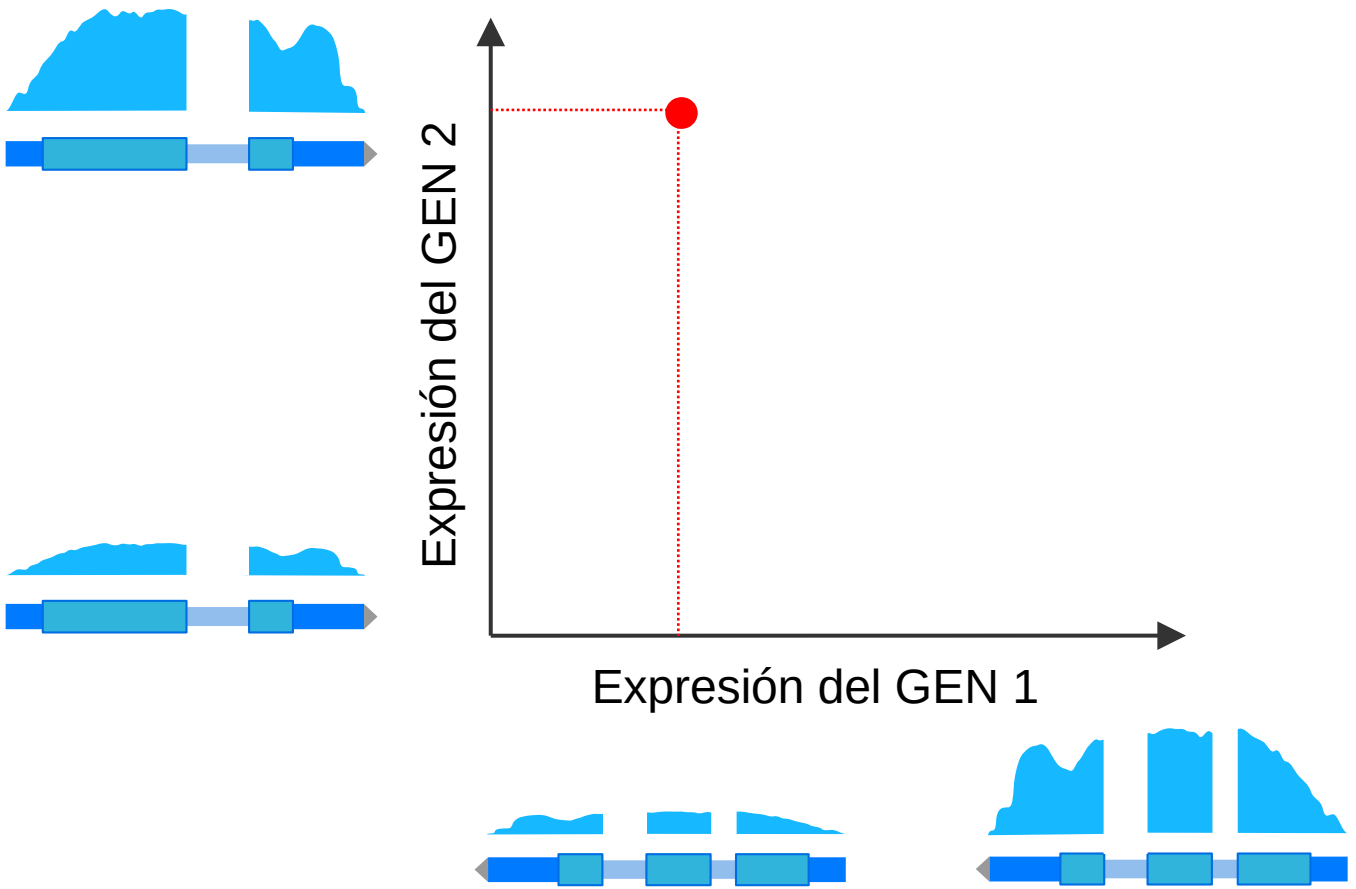
Posición/coordenadas = Patrón de expresión génica



Análisis de datos de RNAseq

Expresión génica

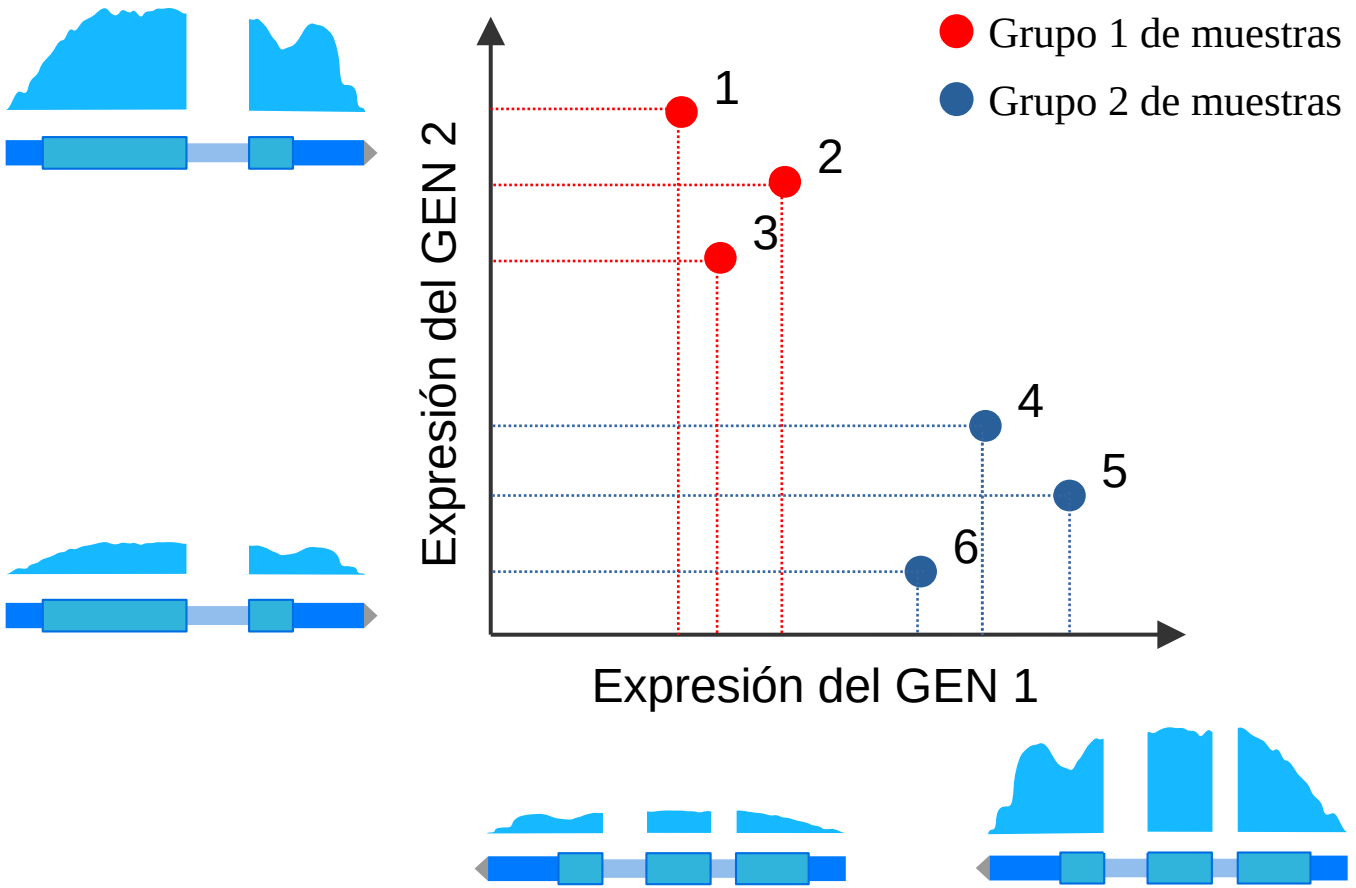
Posición/coordenadas = Patrón de expresión génica



Análisis de datos de RNAseq

Expresión génica

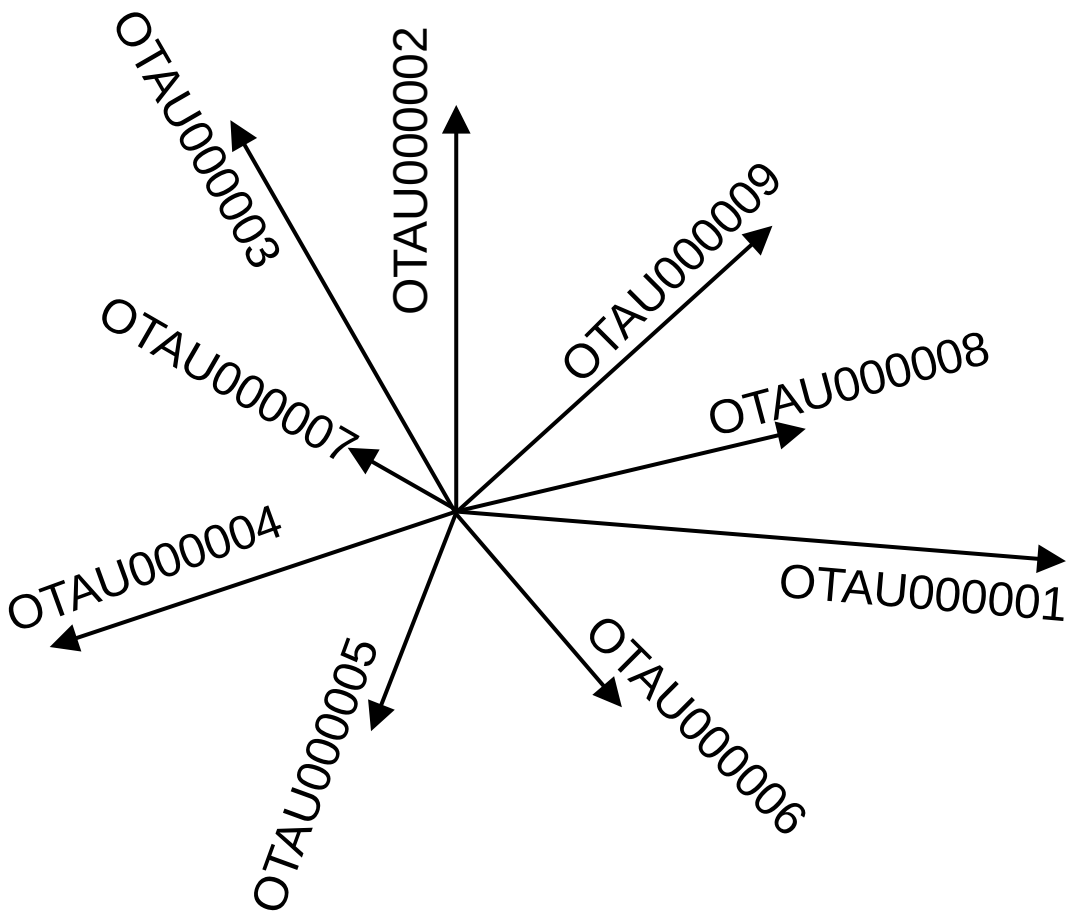
Posición/coordenadas = Patrón de expresión génica



Análisis de datos de RNAseq

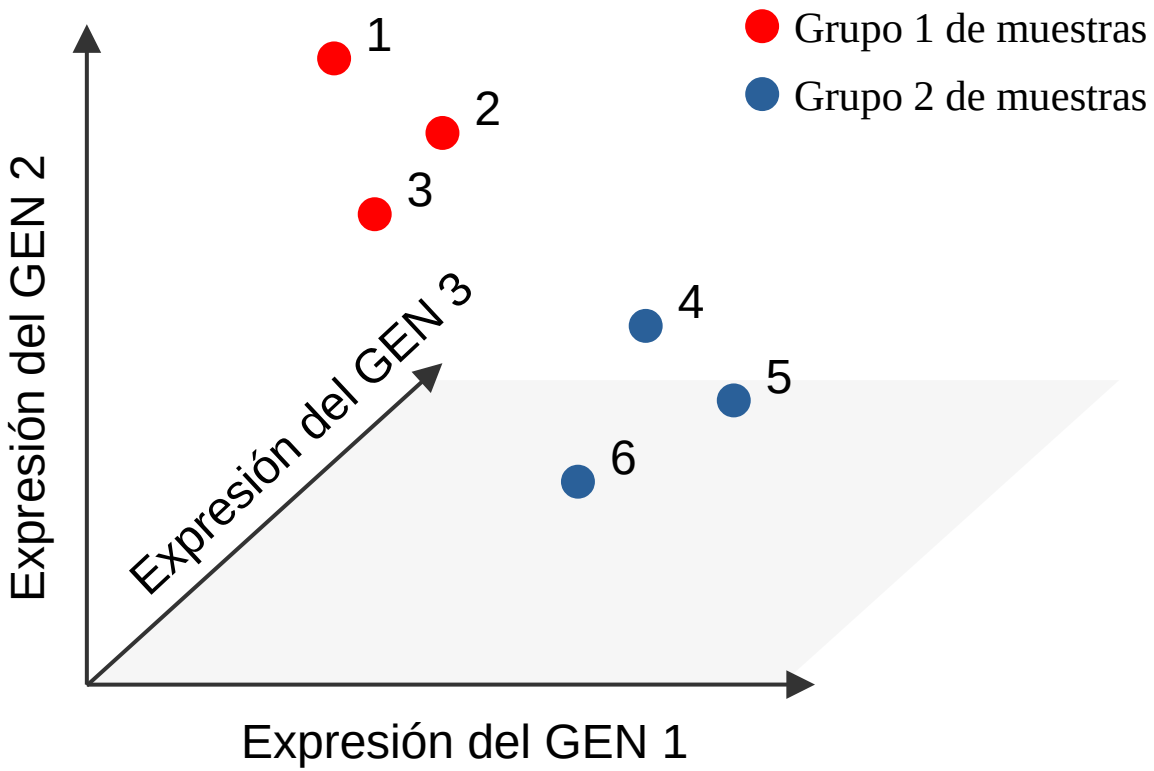
Expresión génica

	gene	ctrlRNAi_M.BRN_S19	ctrlRNAi_M.BRN_S20	ctrlRNAi_M
1	OTAU000001	166	155	
2	OTAU000002	955	975	
3	OTAU000003	2318	2228	
4	OTAU000004	2	1	
5	OTAU000005	29	26	
6	OTAU000006	97	78	
7	OTAU000007	0	0	
8	OTAU000008	1841	1902	
9	OTAU000009	730	549	
10	OTAU000010	490	535	
11	OTAU000011	1	4	
12	OTAU000012	21	14	
13	OTAU000013	25	29	
14	OTAU000014	39	35	
15	OTAU000015	3625	3178	
16	OTAU000016	1130	1115	
17	OTAU000017	638	607	
18	OTAU000018	304	242	
19	OTAU000019	15	10	
20	OTAU000020	0	0	



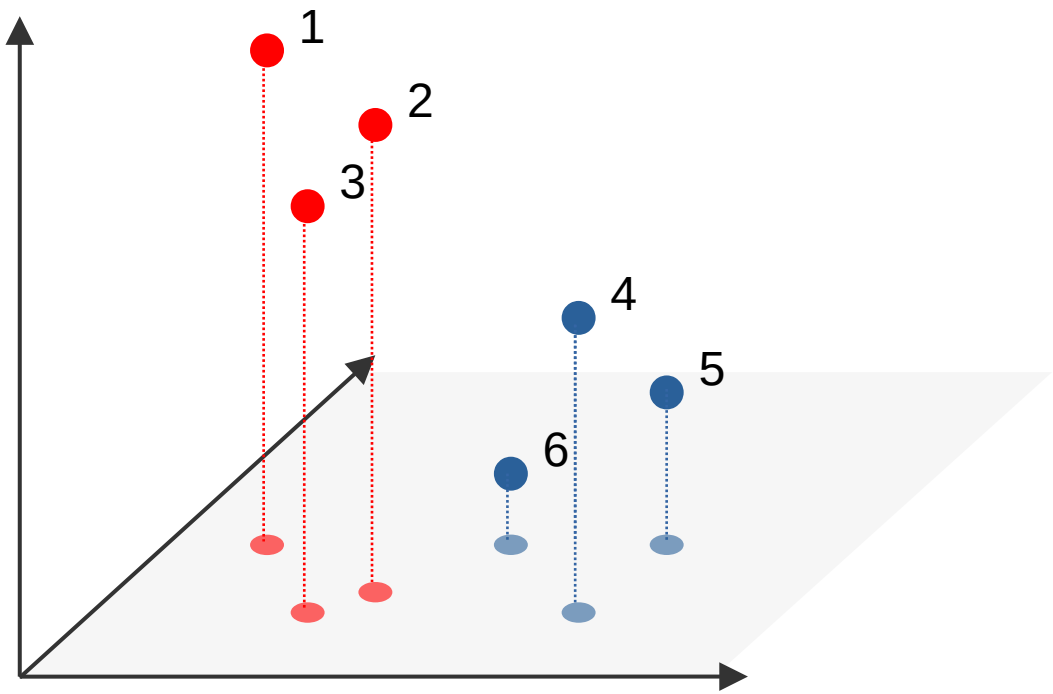
Análisis de datos de RNAseq

Expresión génica



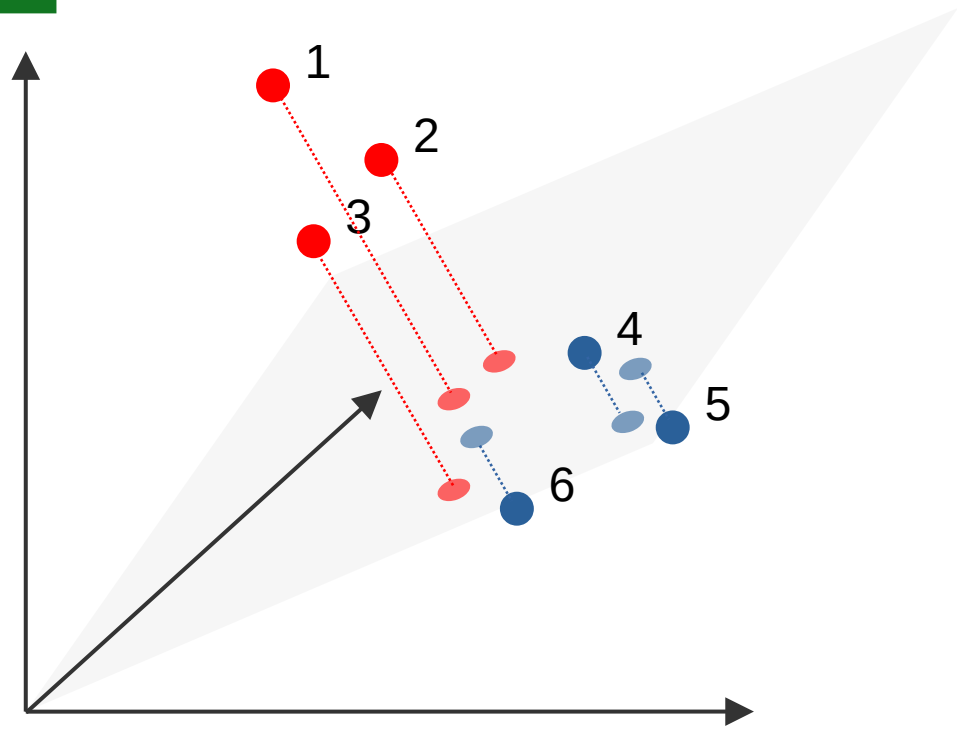
Análisis de datos de RNAseq

Expresión génica



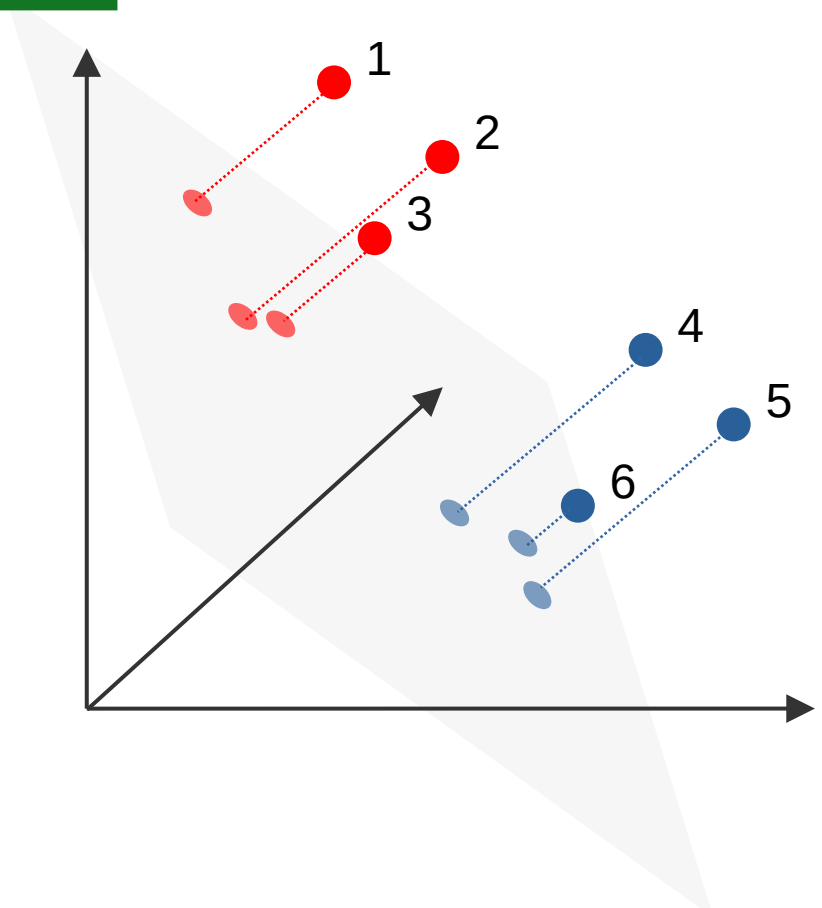
Análisis de datos de RNAseq

Expresión génica



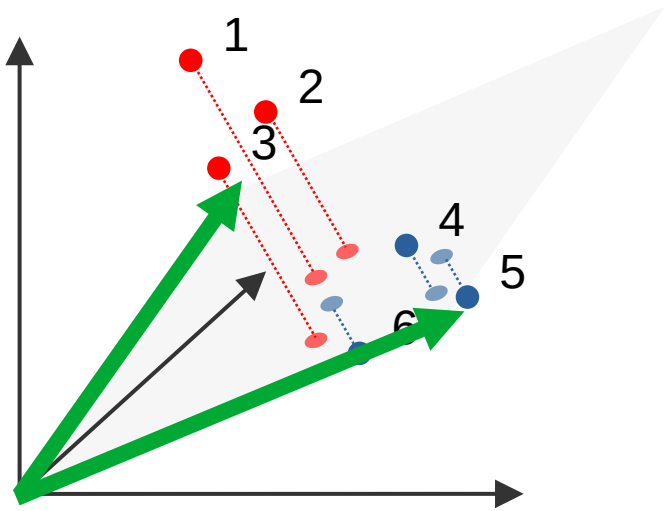
Análisis de datos de RNAseq

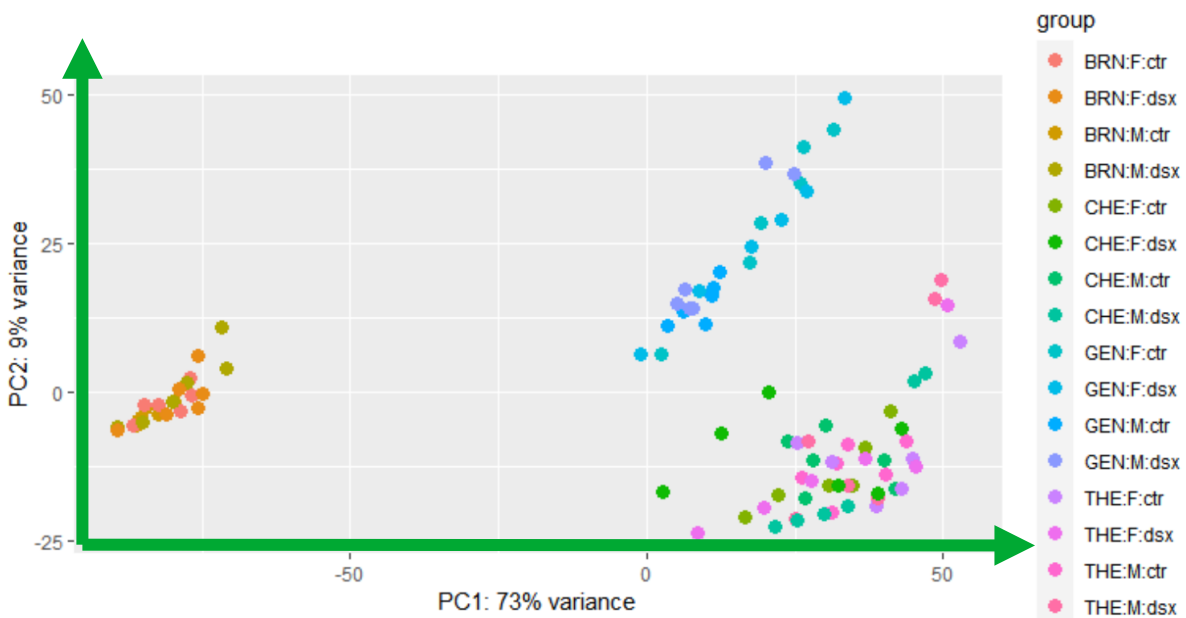
Expresión génica



Análisis de datos de RNAseq

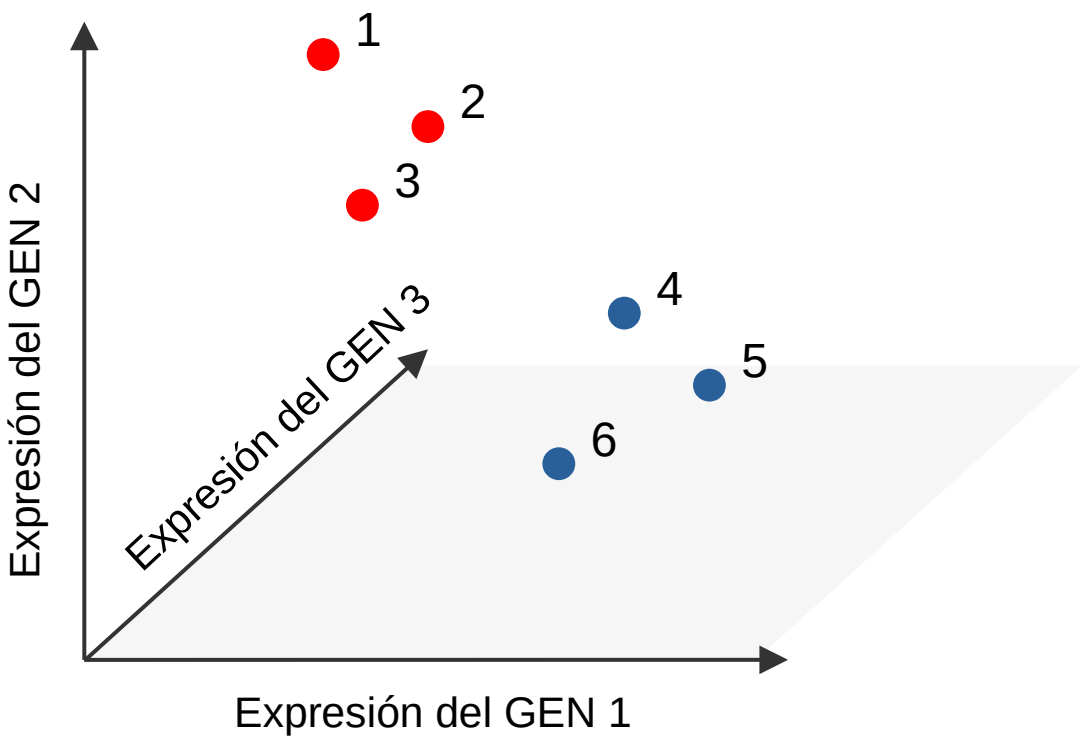
Comp. principales (PCA)



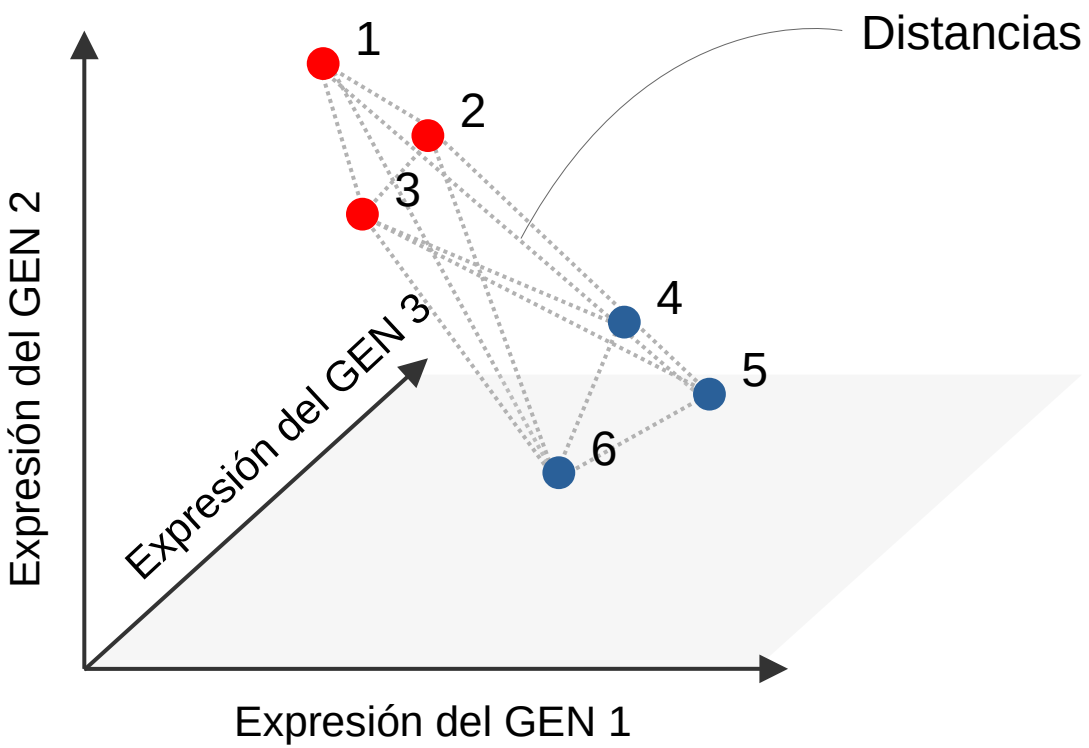


Análisis de datos de RNAseq

Distancias

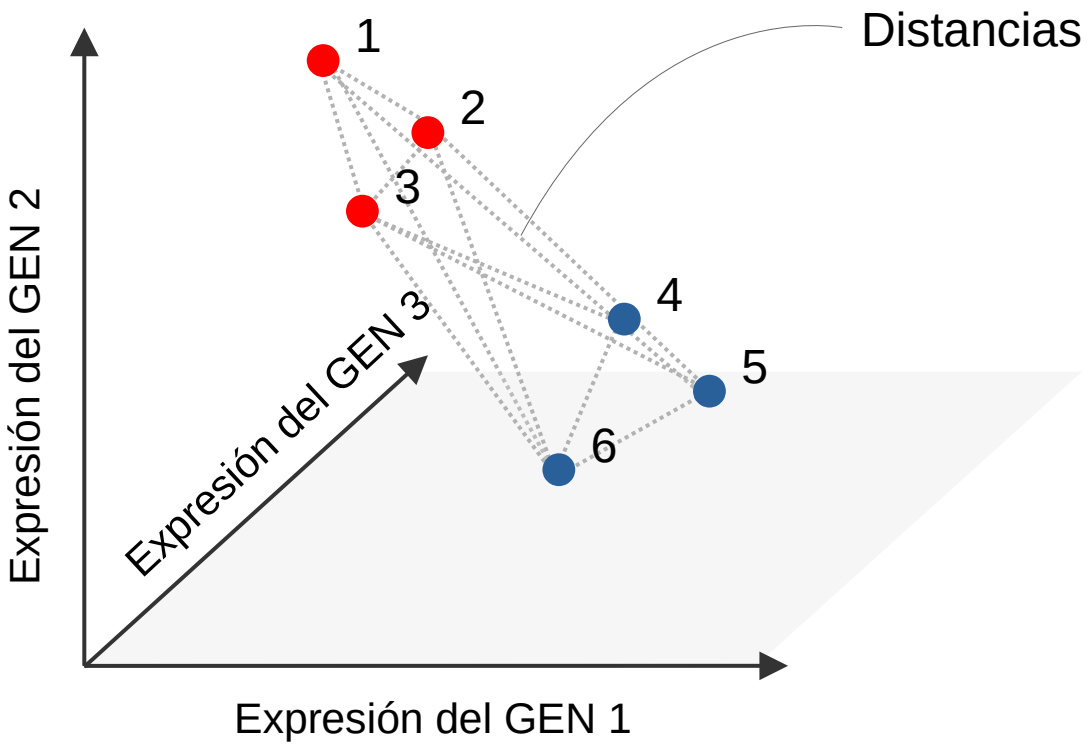


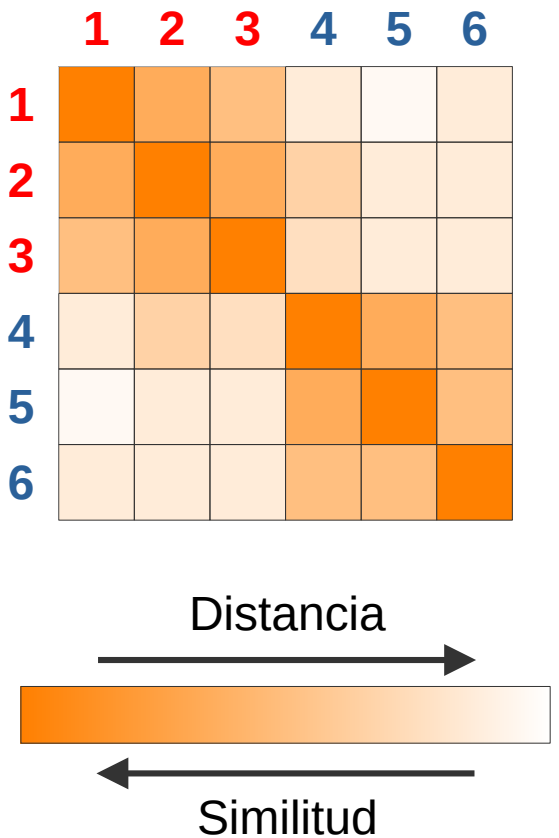
Distancias


Distancias
1
2
3
4
5
6
Expresión del GEN 2
Expresión del GEN 3
Expresión del GEN 1

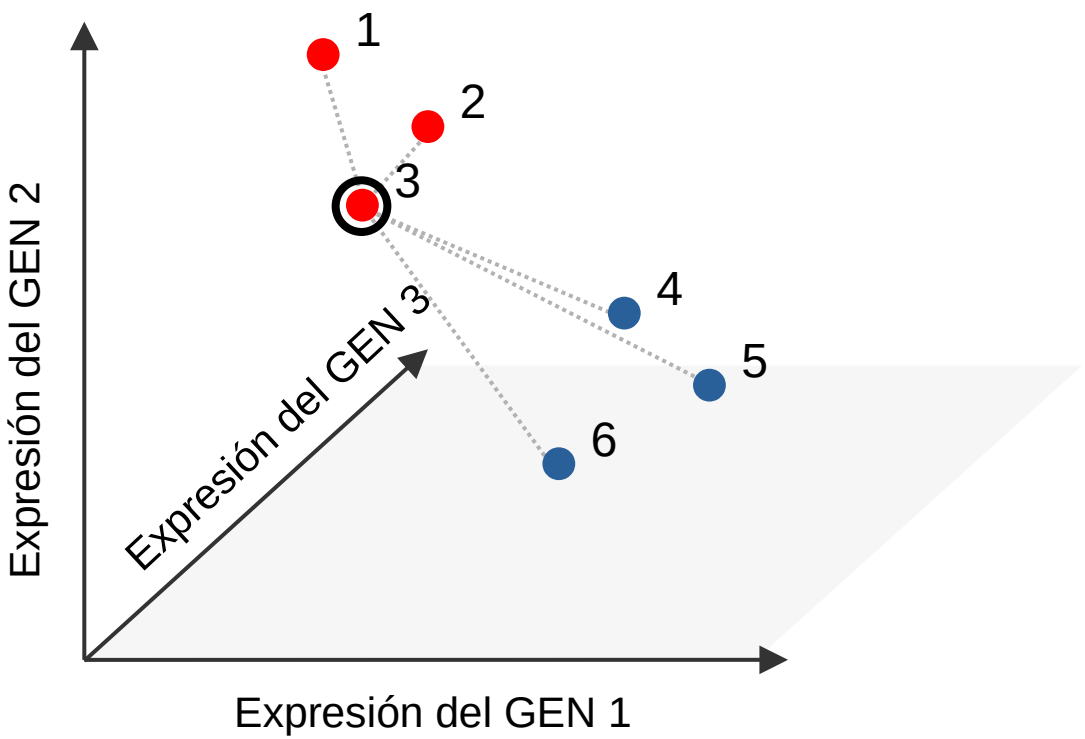
Análisis e interpretación de datos de RNAseq

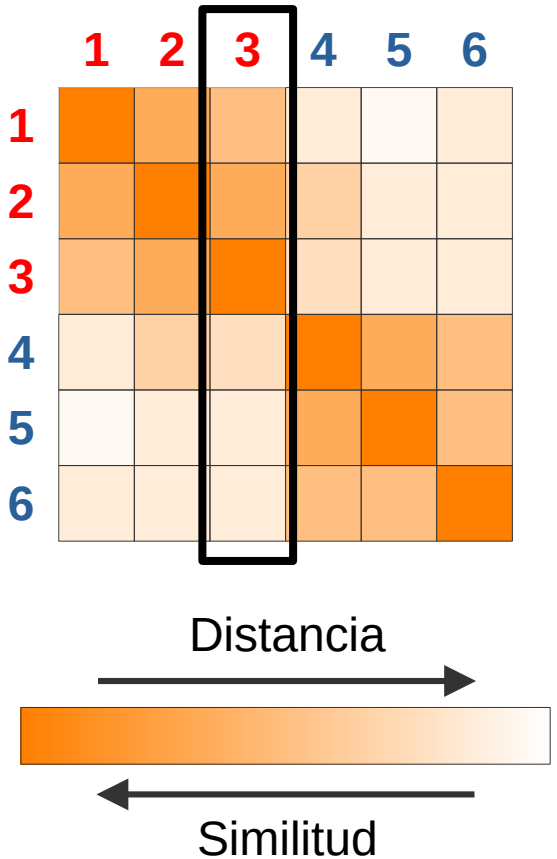
Heatmaps



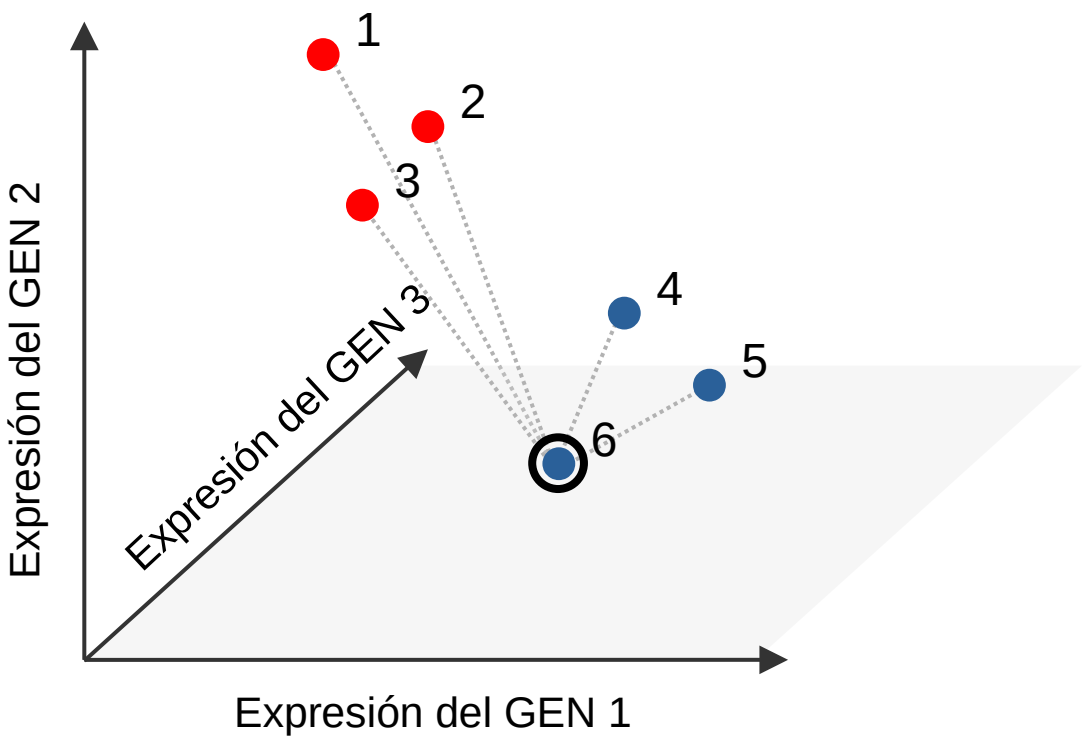


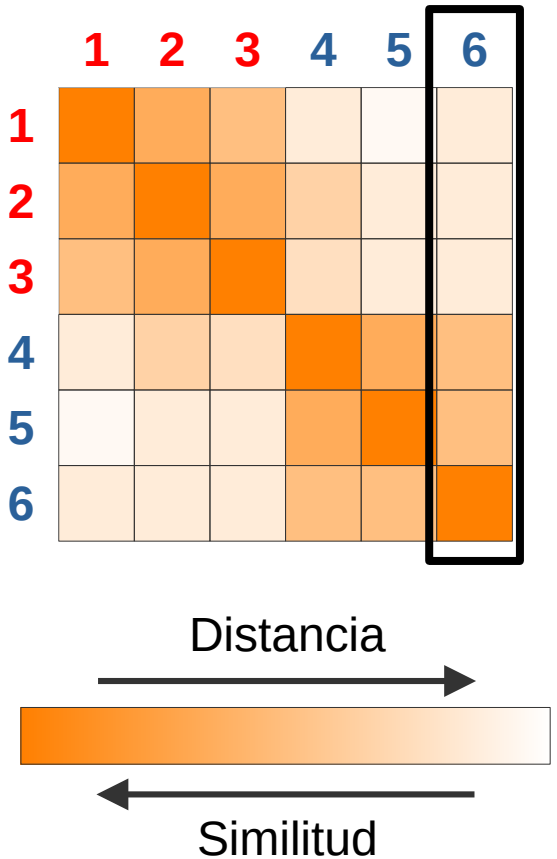
Heatmaps


1
2
3
4
5
6
Expresión del GEN 2
Expresión del GEN 3
Expresión del GEN 1

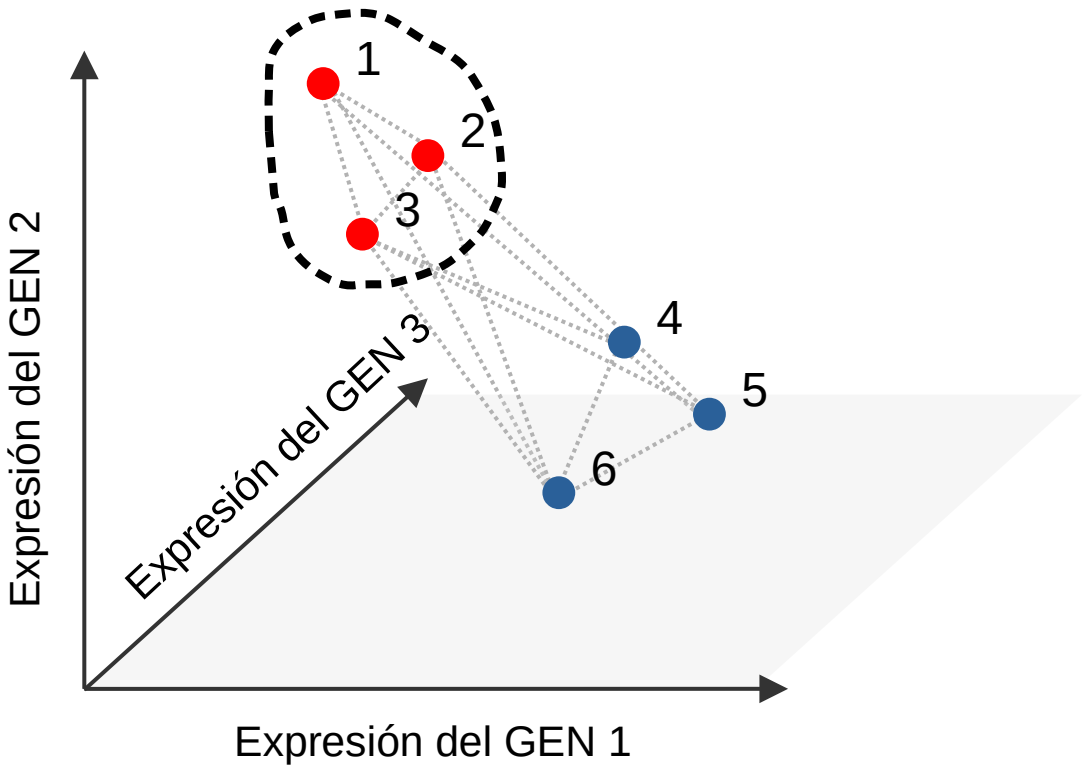

1 2 3 4 5 6
1
2
3
4
5
6
Distancia
Similitud

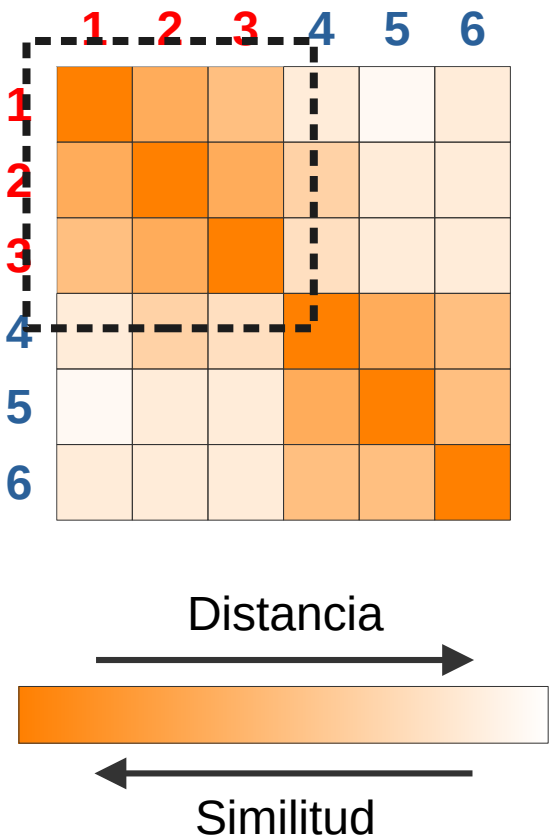
Heatmaps


1
2
3
4
5
6
Expresión del GEN 2
Expresión del GEN 3
Expresión del GEN 1


1 2 3 4 5 6
1
2
3
4
5
6
Distancia
Similitud

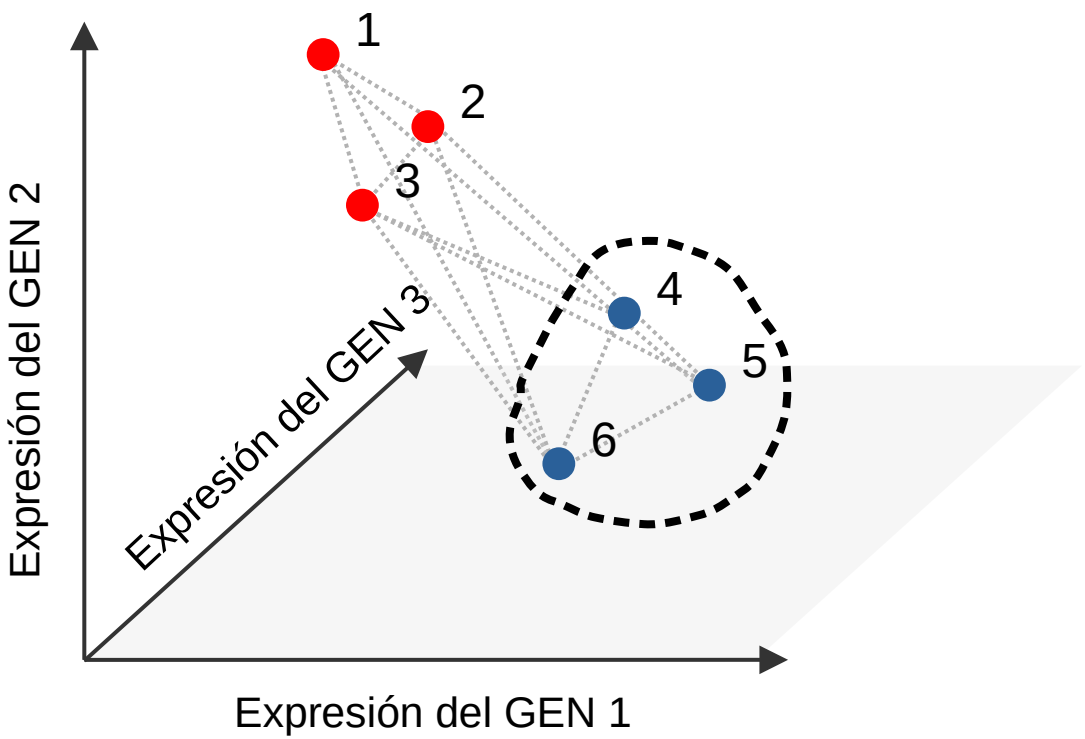
Heatmaps

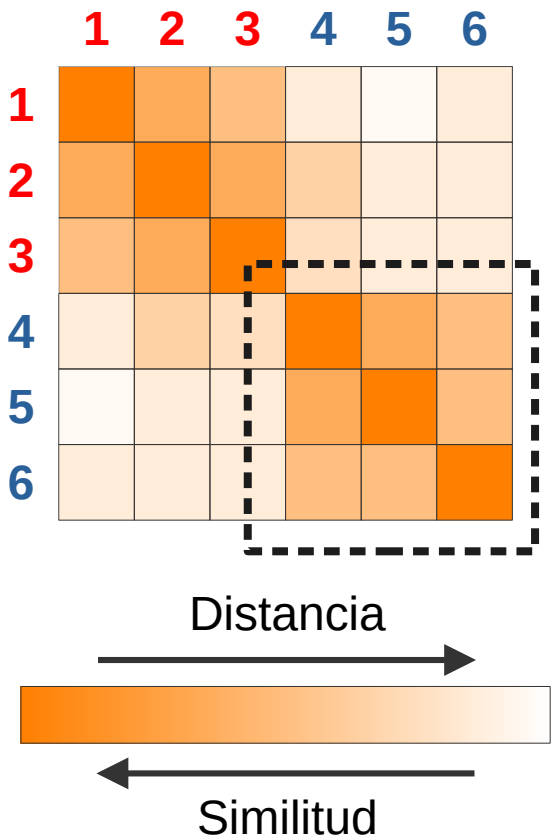

1
2
3
4
5
6
Expresión del GEN 2
Expresión del GEN 3
Expresión del GEN 1


1 2 3 4 5 6
1
2
3
4
5
6
Distancia
Similitud

Análisis e interpretación de datos de RNAseq

Heatmaps



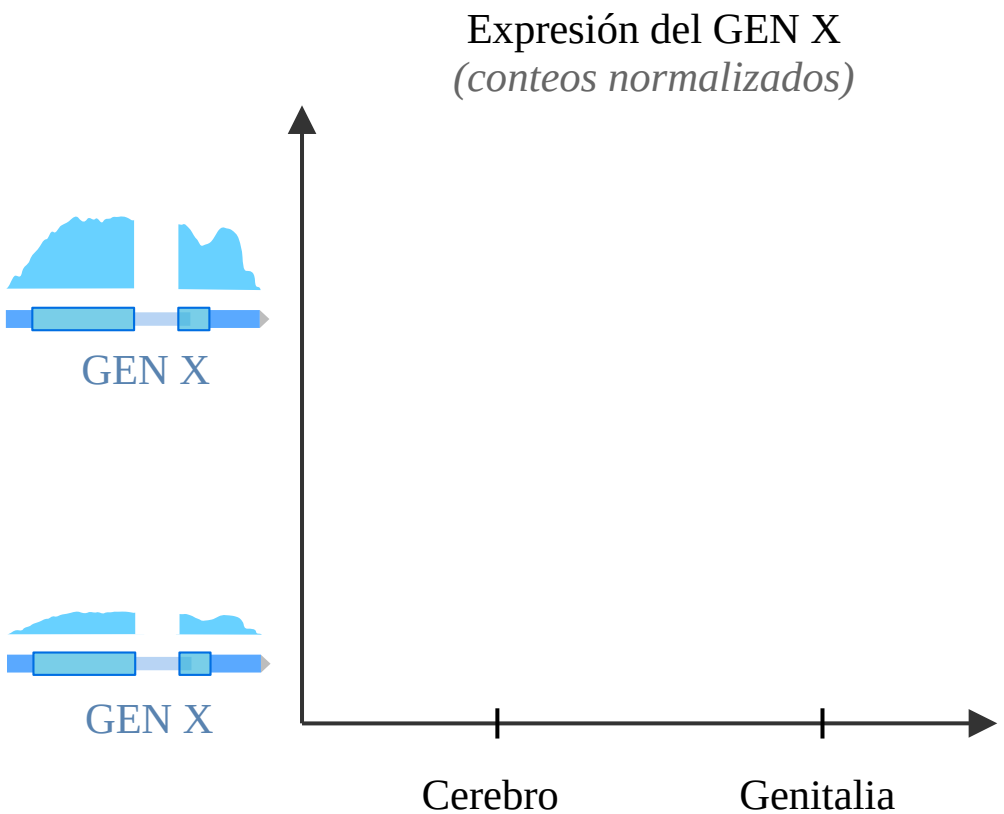


Análisis e interpretación de datos de RNAseq

Expresión diferencial

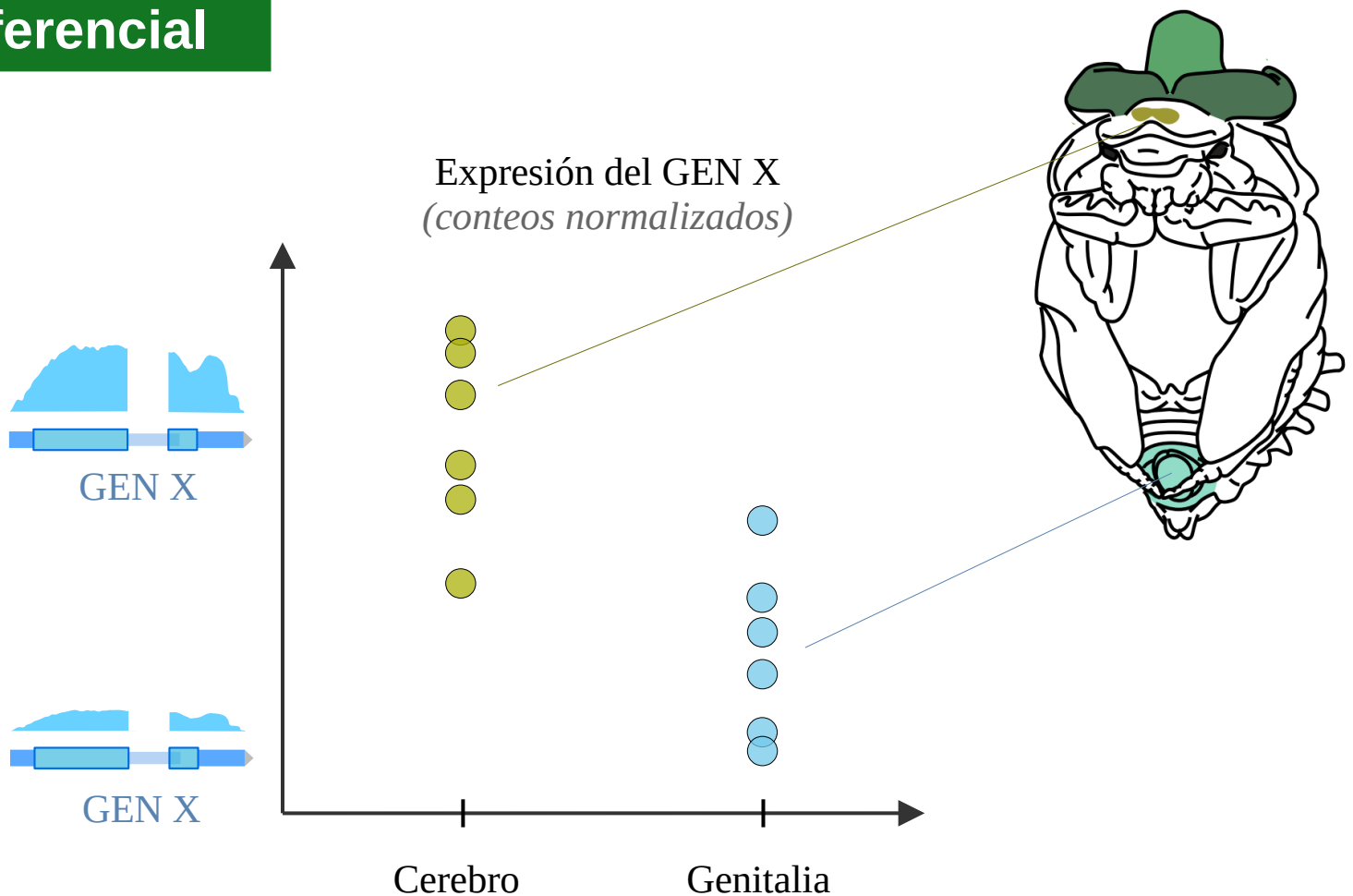
Análisis e interpretación de datos de RNAseq

Expresión diferencial



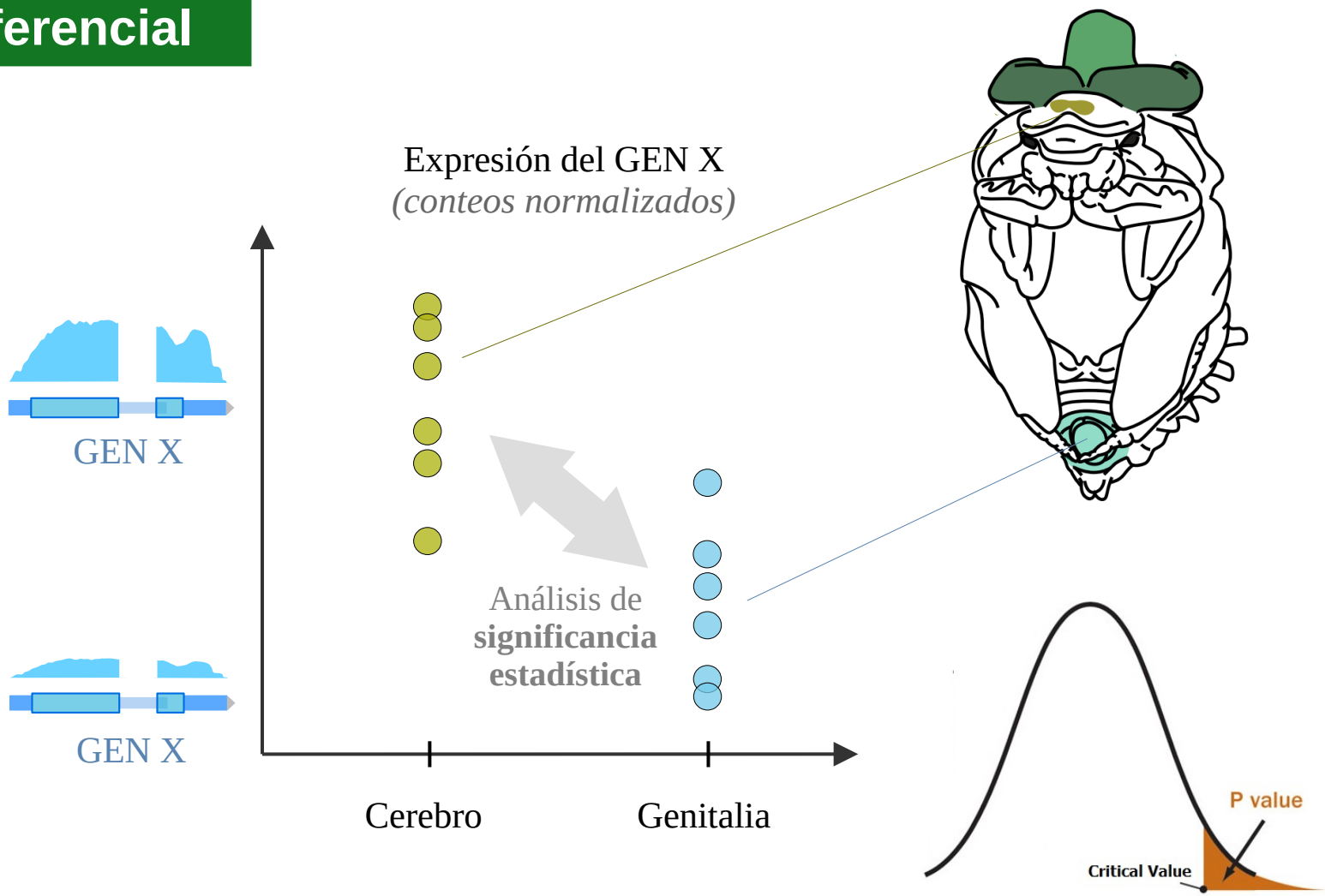
Análisis e interpretación de datos de RNAseq

Expresión diferencial



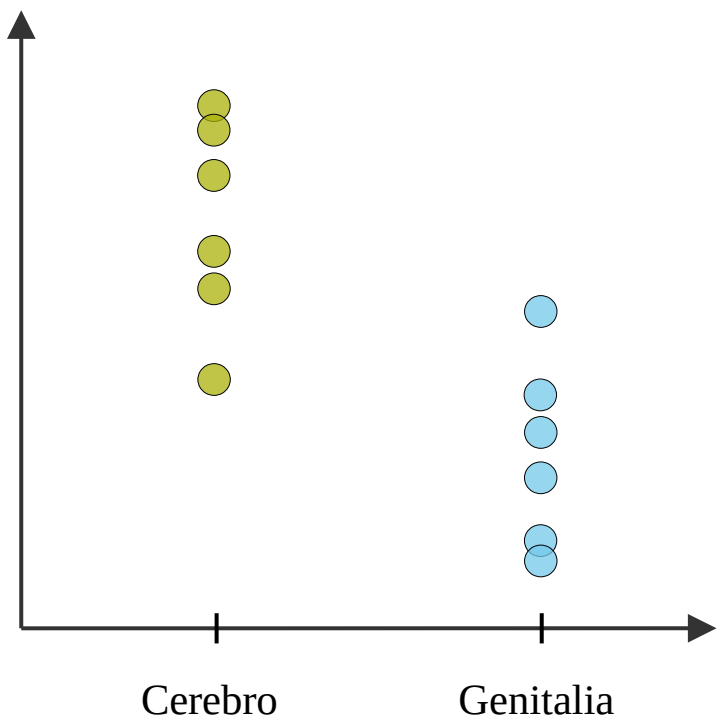
Análisis e interpretación de datos de RNAseq

Expresión diferencial



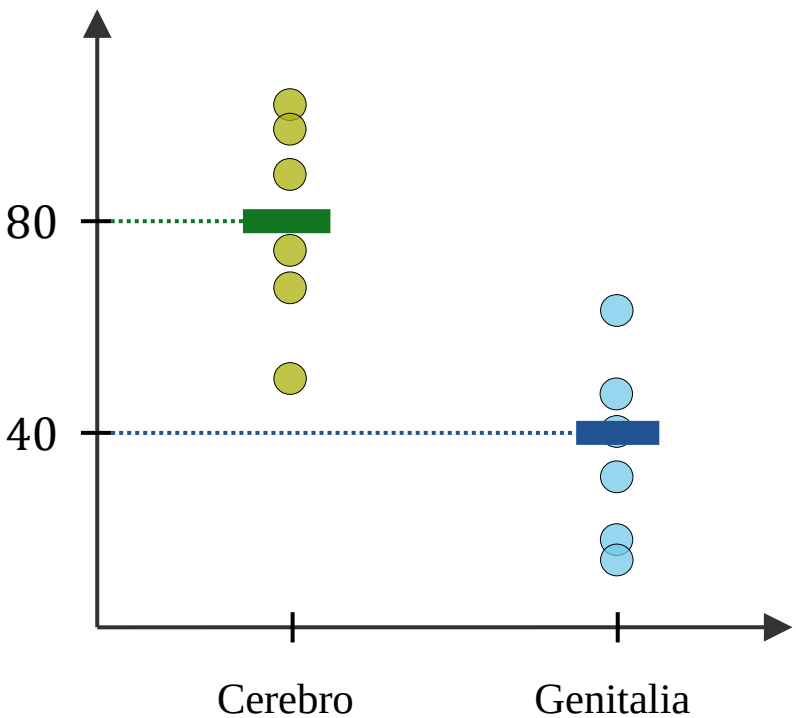
Análisis e interpretación de datos de RNAseq

Expresión diferencial



Análisis e interpretación de datos de RNAseq

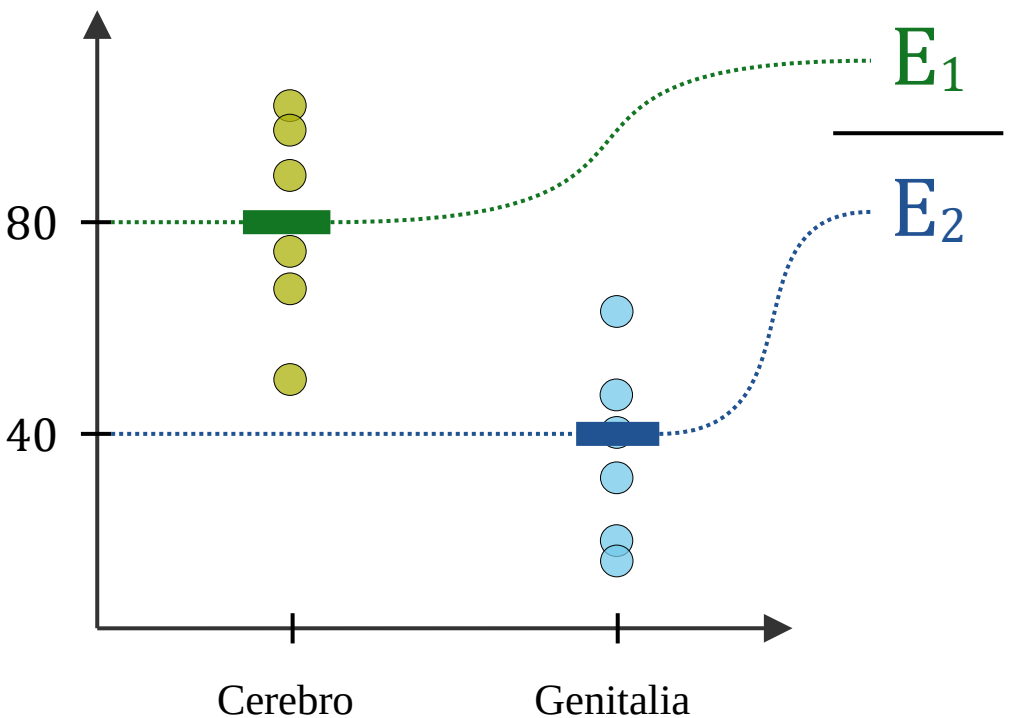
Expresión diferencial



Análisis e interpretación de datos de RNAseq

Expresión diferencial

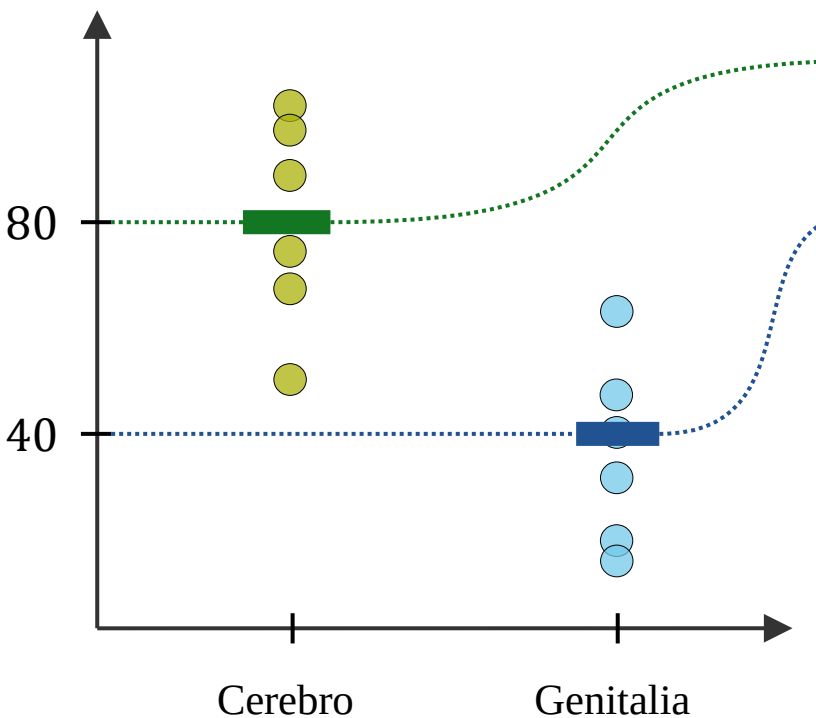
Relación de cambio
(Fold Change)



Análisis e interpretación de datos de RNAseq

Expresión diferencial

Relación de cambio
(Fold Change)

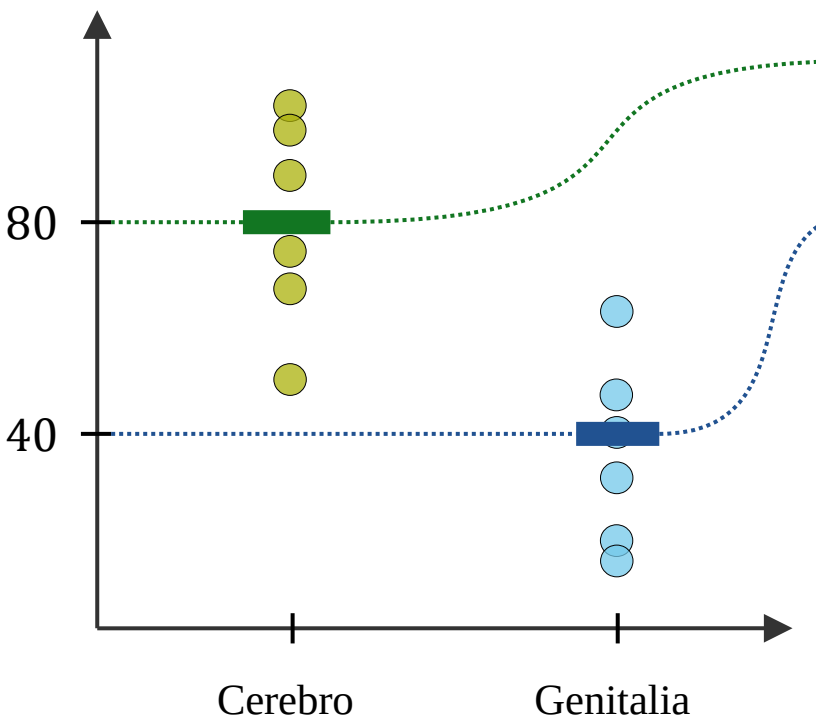


$$\frac{E_1}{E_2} = C_{12} = 2$$

Análisis e interpretación de datos de RNAseq

Expresión diferencial

Relación de cambio
(Fold Change)



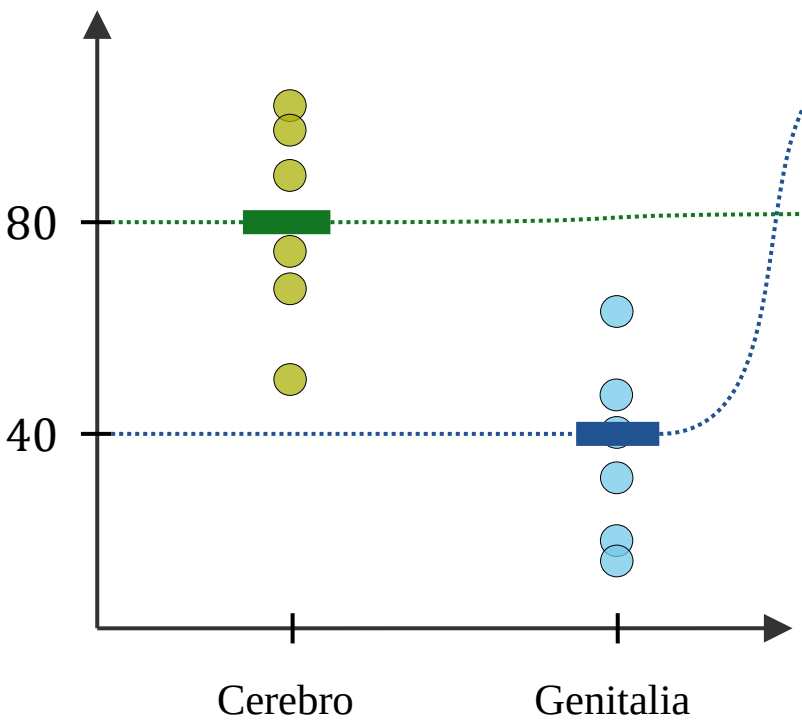
$$\frac{E_1}{E_2} = C_{12} = 2$$

$$Log_2(C_{12}) = 1$$

Log (2) Fold Change
(LFC)

Análisis e interpretación de datos de RNAseq

Expresión diferencial



Relación de cambio
(Fold Change)

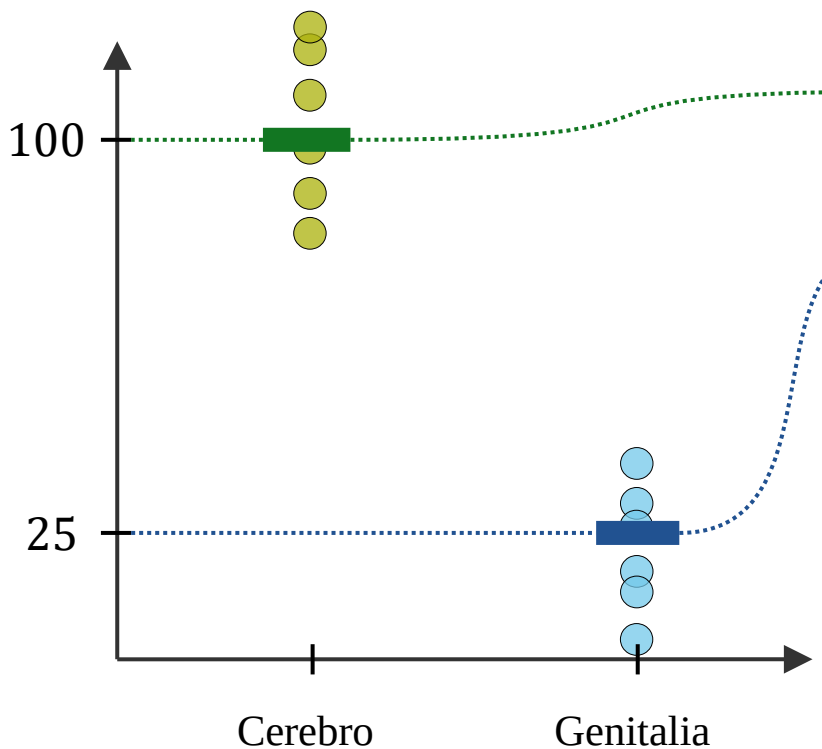
$$\frac{E_2}{E_1} = C_{21} = \boxed{0.5}$$

$$Log_2(C_{21}) = -1$$

Log (2) Fold Change
(LFC)

Análisis e interpretación de datos de RNAseq

Expresión diferencial



Relación de cambio
(Fold Change)

$$\frac{E_1}{E_2} = C_{12} = 4$$

$$Log_2(C_{12}) = 2$$

Log (2) Fold Change
(LFC)

Análisis e interpretación de datos de RNAseq

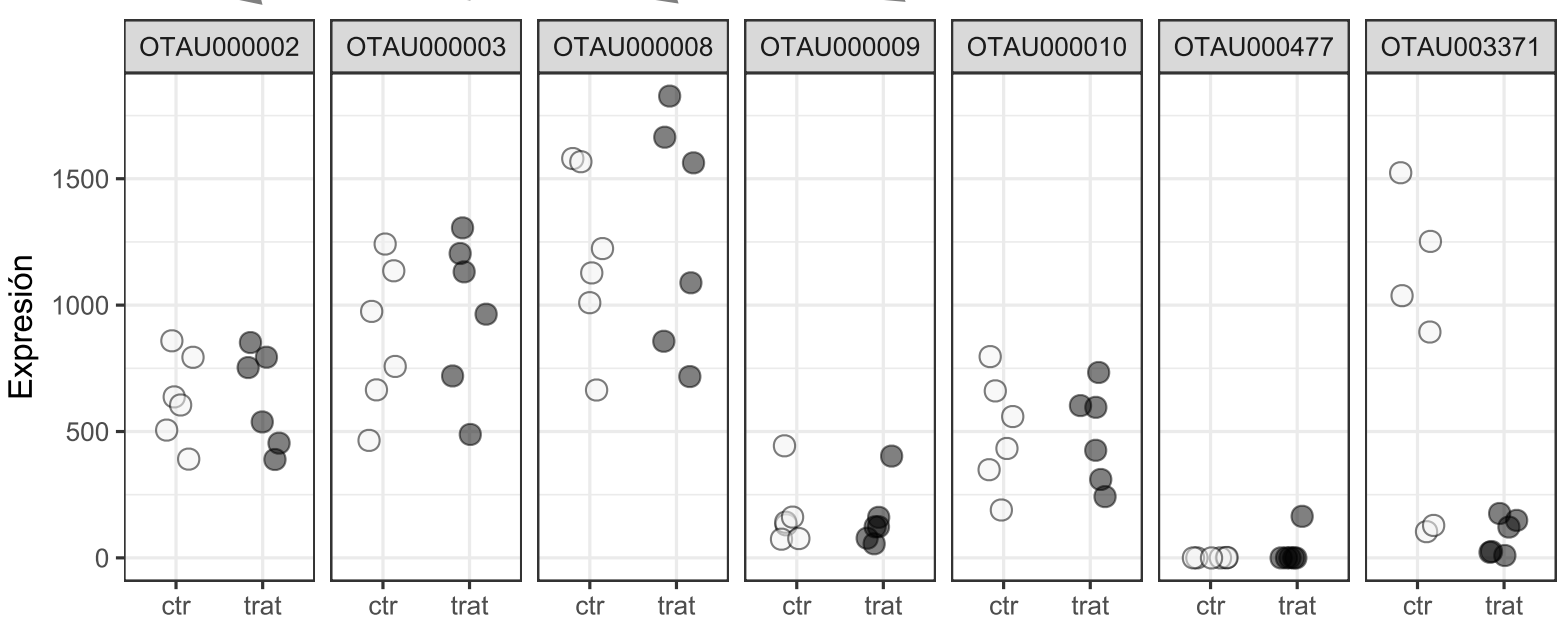
Expresión diferencial

	gene	ctrlRNAi_M.BRN_S19	ctrlRNAi_M.BRN_S20	ctrlRNAi_M
1	OTAU000001	166	155	
2	OTAU000002	955	975	
3	OTAU000003	2318	2228	
4	OTAU000004	2	1	
5	OTAU000005	29	26	
6	OTAU000006	97	78	
7	OTAU000007	0	0	
8	OTAU000008	1841	1902	
9	OTAU000009	730	549	
10	OTAU000010	490	535	
11	OTAU000011	1	4	
12	OTAU000012	21	14	
13	OTAU000013	25	29	
14	OTAU000014	39	35	
15	OTAU000015	3625	3178	
16	OTAU000016	1130	1115	
17	OTAU000017	638	607	
18	OTAU000018	304	242	
19	OTAU000019	15	10	
20	OTAU000020	0	0	

Análisis e interpretación de datos de RNAseq

Expresión diferencial

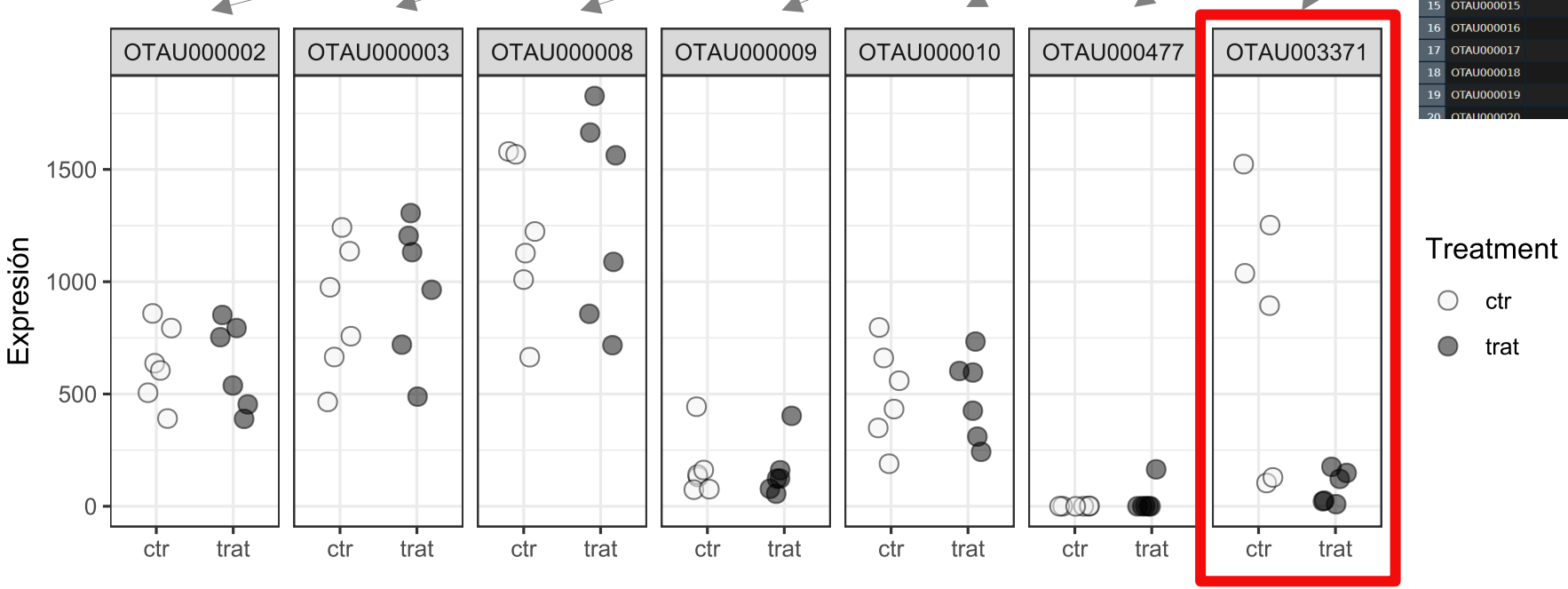
	gene	ctrlRNAi_M.BRN_S19	ctrlRNAi_M.BRN_S20	ctrlRNAi_M
1	OTAU000001	166	155	
2	OTAU000002	955	975	
3	OTAU000003	2318	2228	
4	OTAU000004	2	1	
5	OTAU000005	29	26	
6	OTAU000006	97	78	
7	OTAU000007	0	0	
8	OTAU000008	1841	1902	
9	OTAU000009	730	549	
10	OTAU000010	490	535	
11	OTAU000011	1	4	
12	OTAU000012	21	14	
13	OTAU000013	25	29	
14	OTAU000014	39	35	
15	OTAU000015	3625	3178	
16	OTAU000016	1130	1115	
17	OTAU000017	638	607	
18	OTAU000018	304	242	
19	OTAU000019	15	10	
20	OTAU000020	0	0	



Treatment

- ctr
- trat

Análisis e interpretación de datos de RNAseq

Expresión diferencial

	gene	ctrlRNAi_M.BRN_S19	ctrlRNAi_M.BRN_S20	ctrlRNAi_M
1	OTAU000001	166	155	
2	OTAU000002	955	975	
3	OTAU000003	2318	2228	
4	OTAU000004	2	1	
5	OTAU000005	29	26	
6	OTAU000006	97	78	
7	OTAU000007	0	0	
8	OTAU000008	1841	1902	
9	OTAU000009	730	549	
10	OTAU000010	490	535	
11	OTAU000011	1	4	
12	OTAU000012	21	14	
13	OTAU000013	25	29	
14	OTAU000014	39	35	
15	OTAU000015	3625	3178	
16	OTAU000016	1130	1115	
17	OTAU000017	638	607	
18	OTAU000018	304	242	
19	OTAU000019	15	10	
20	OTAU000020	0	0