Genoma: Representación estática de una sistema biológico.
Conjunto completo de genes en un organismo o sus organelos.

Transcriptoma: Representación de la expresión génica en estado fisiológico determinado.
Set completo de mARNs, presentes en una célula, tejido u órgano.

Proteoma: Visión integrada de las proteínas en un proceso biológico.
Son las proteínas codificadas por un genoma, presentes en una célula, tejido u órgano.

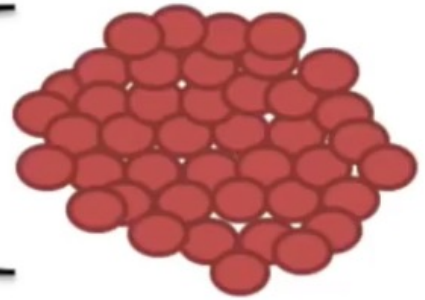
RNAseq

= a normal neural cell

A bunch of normal neural cells.

= a mutated neural cell

A bunch of mutated neural cells.

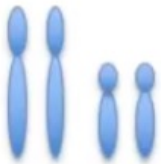


https://statquest.org

= a normal neural cell

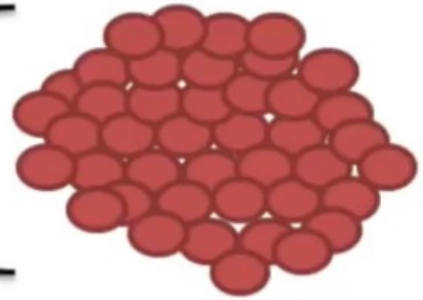
A bunch of normal neural cells.

Each cell has a bunch of chromosomes..



= a mutated neural cell

A bunch of mutated neural cells.

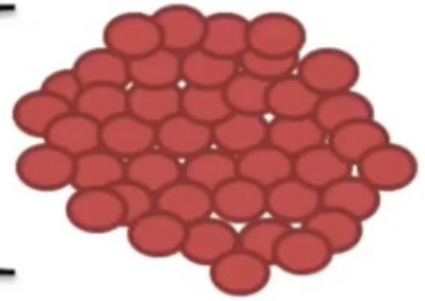


= a normal neural cell

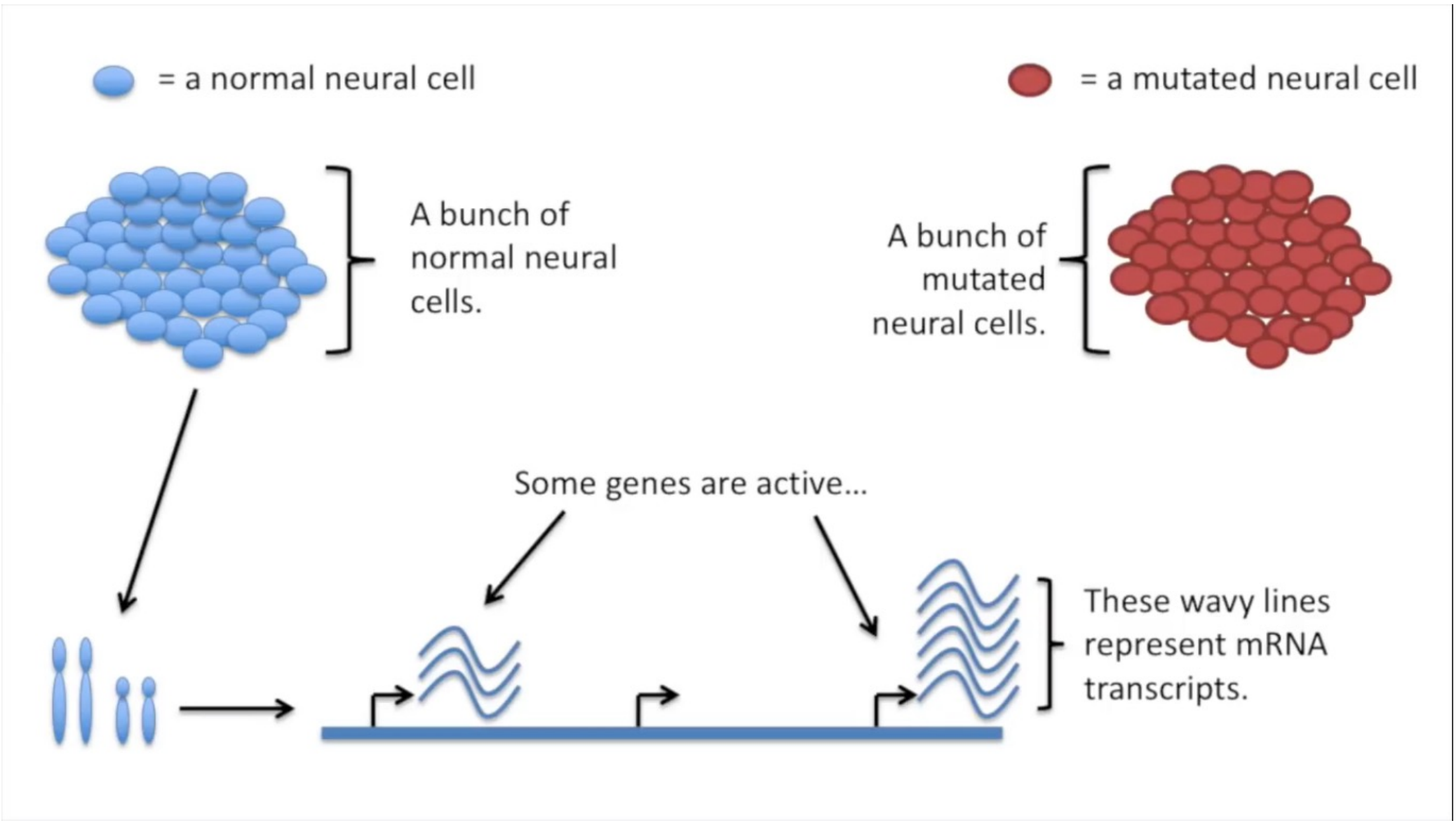
= a mutated neural cell

A bunch of normal neural cells.

A bunch of mutated neural cells.



Each chromosome has a bunch of genes…

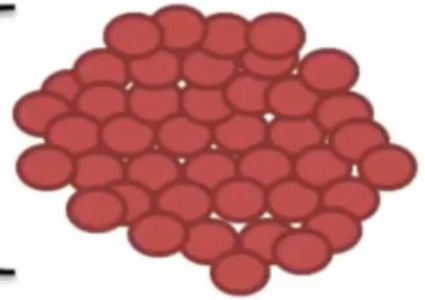

= a normal neural cell
= a mutated neural cell
A bunch of normal neural cells.
A bunch of mutated neural cells.
Some genes are active…
These wavy lines represent mRNA transcripts.

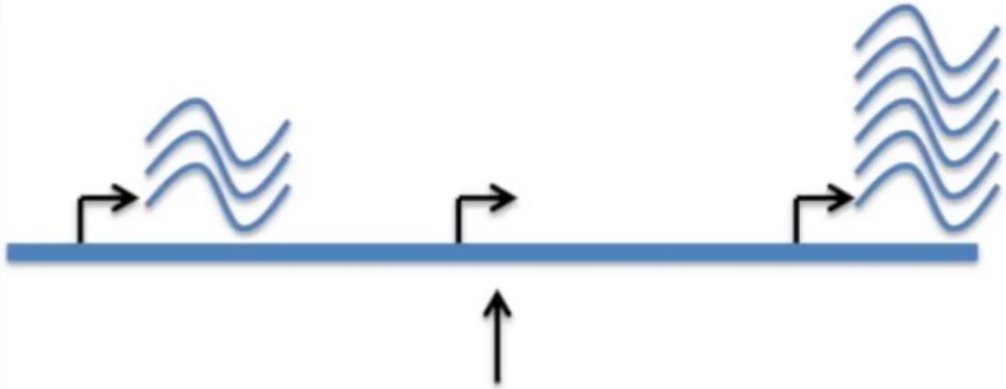
= a normal neural cell

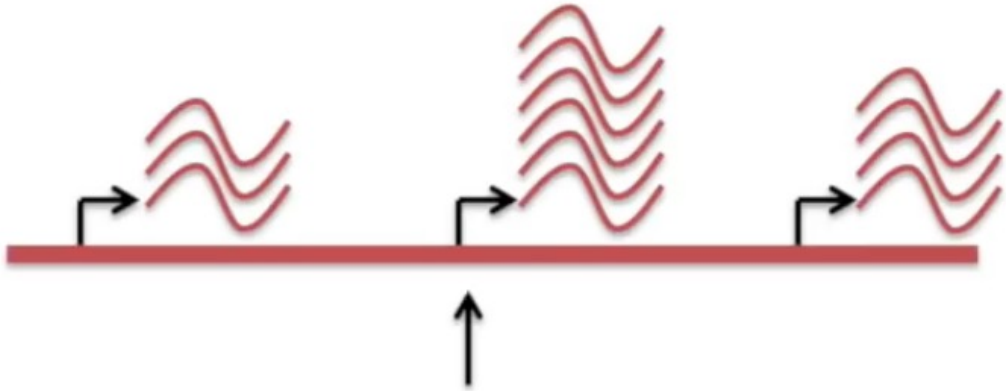
= a mutated neural cell

A bunch of normal neural cells.

A bunch of mutated neural cells.

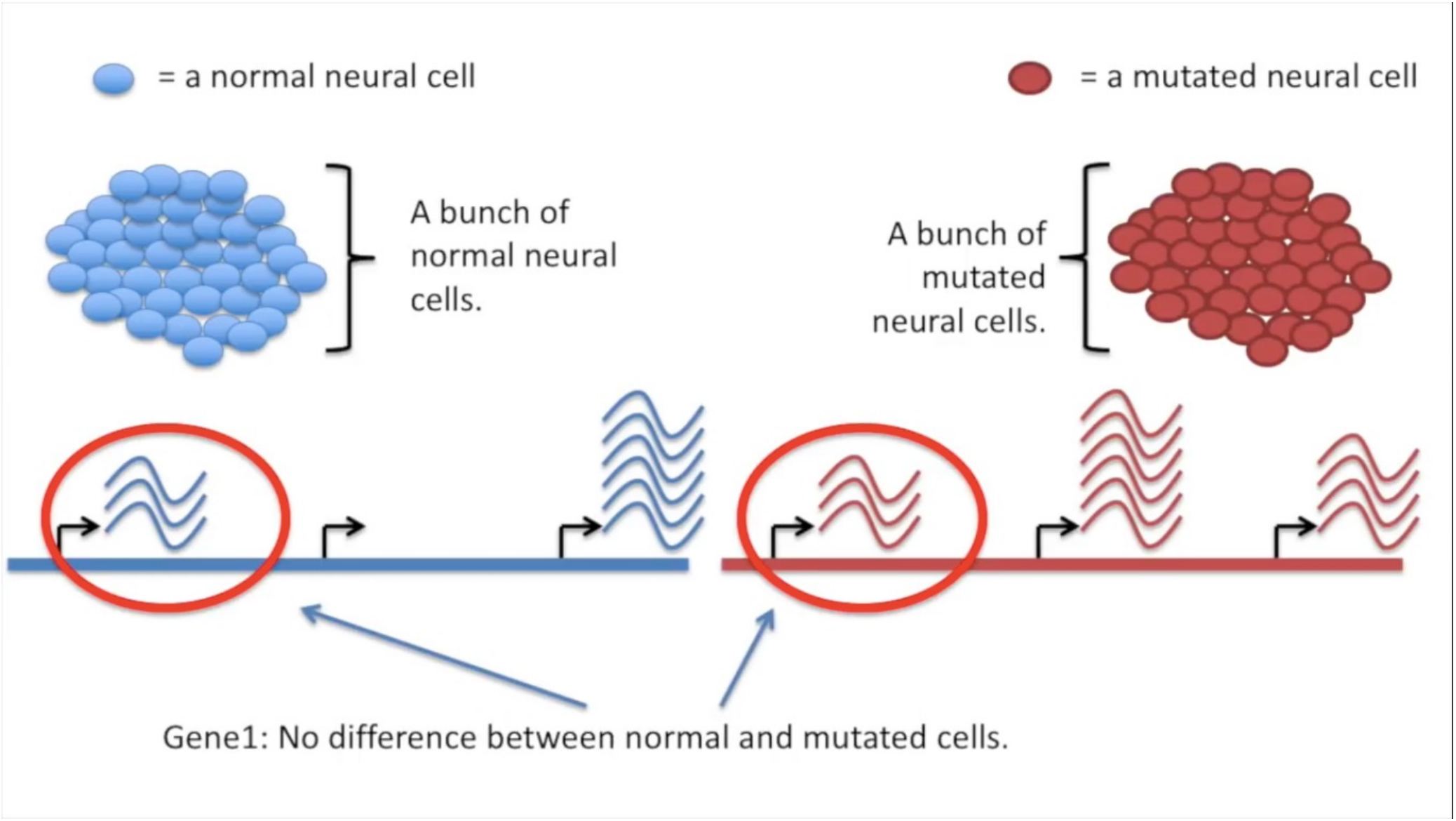


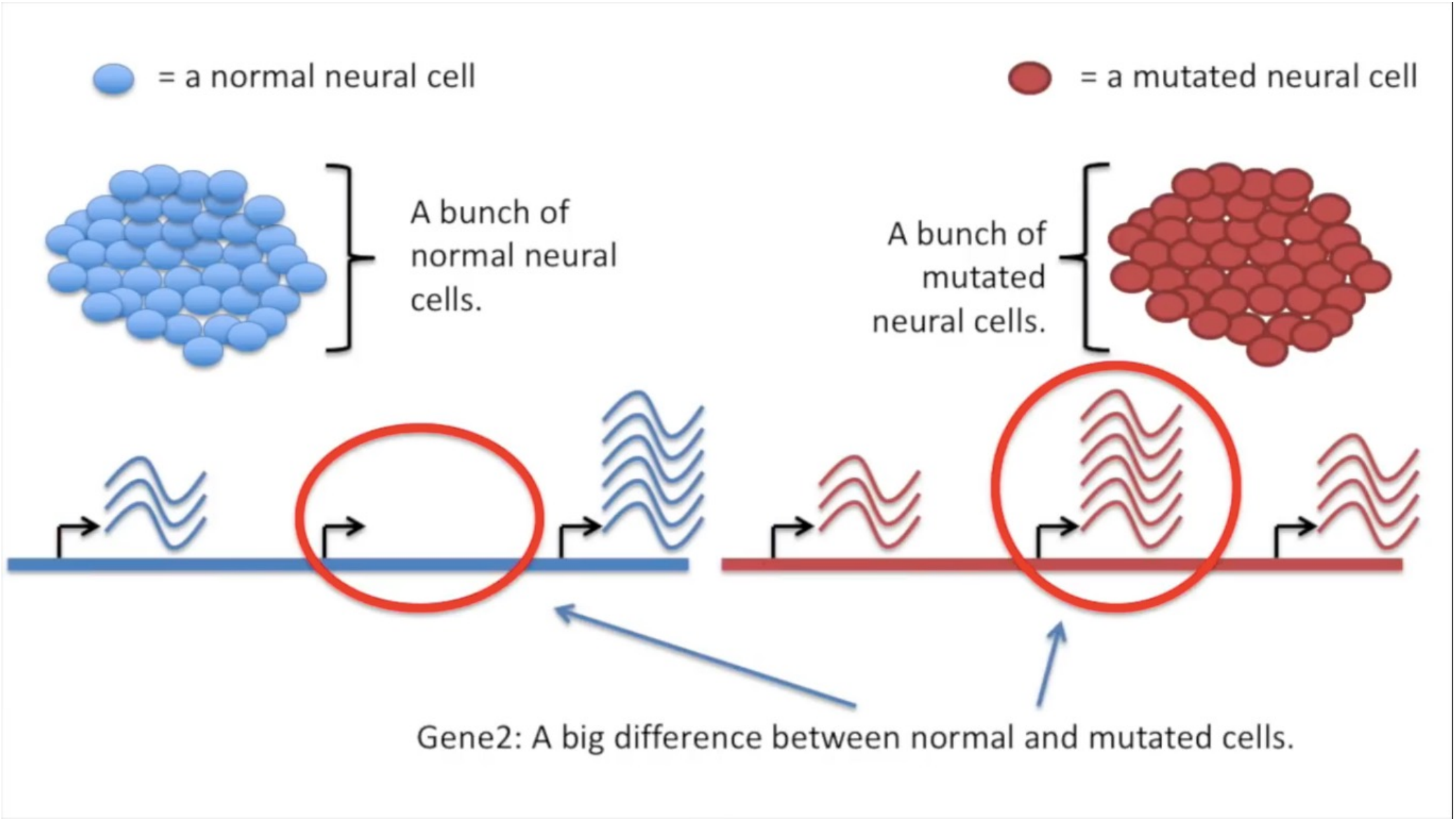


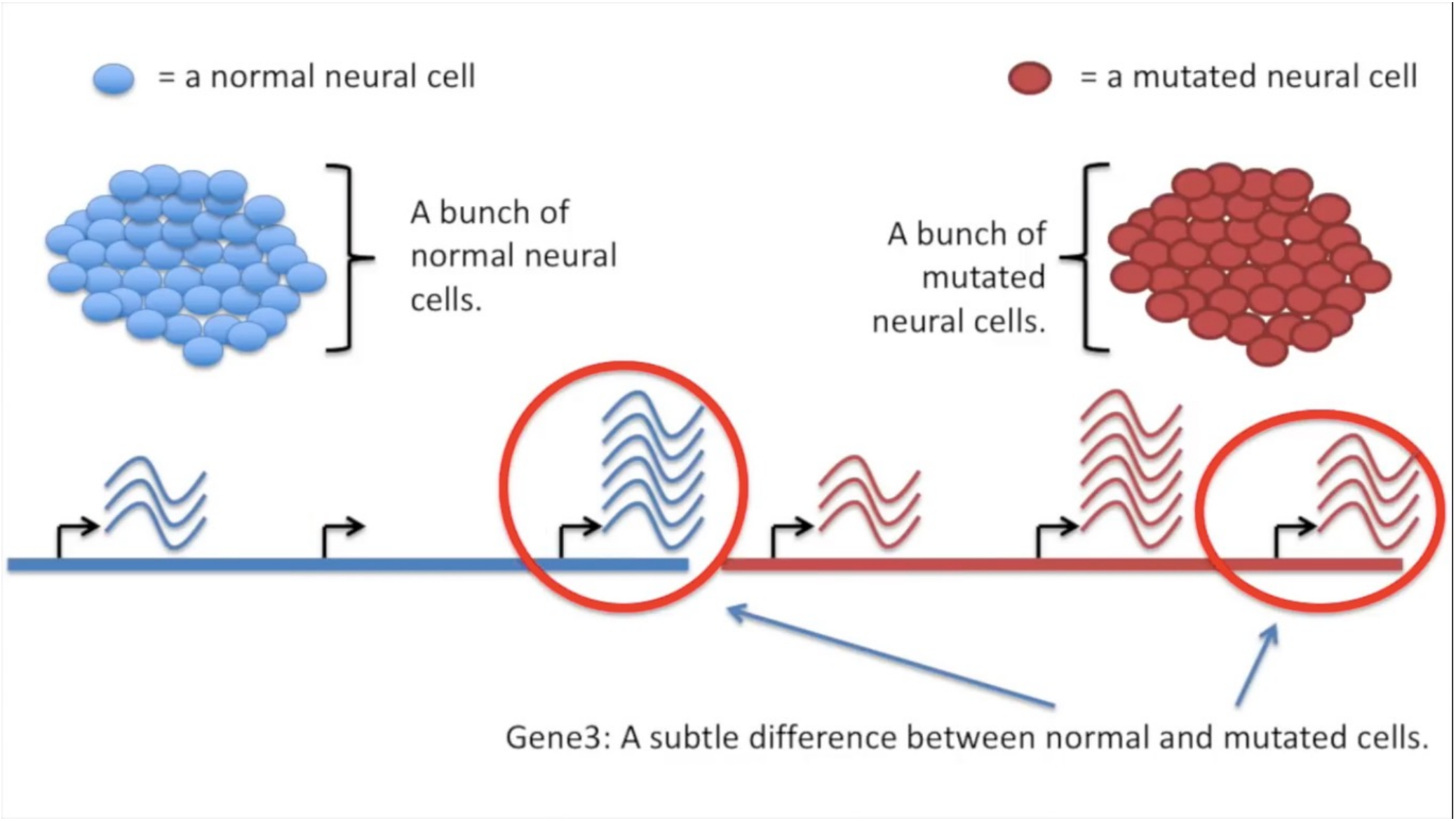


We can use RNA-seq to measure gene expression in normal cells…

… then use it to measure gene expression in mutated cells…


= a normal neural cell
= a mutated neural cell
A bunch of normal neural cells.
A bunch of mutated neural cells.
Gene1: No difference between normal and mutated cells.


= a normal neural cell
= a mutated neural cell
A bunch of normal neural cells.
A bunch of mutated neural cells.
Gene2: A big difference between normal and mutated cells.


= a normal neural cell
= a mutated neural cell
A bunch of normal neural cells.
A bunch of mutated neural cells.
Gene3: A subtle difference between normal and mutated cells.

3 Main Steps for RNA-Seq:

1) Prepare a sequencing library

2) Sequence

3) Data analysis

Preparing an RNA-seq library

Step 1: Isolate the RNA

Preparing an RNA-seq library

Step 1: Isolate the RNA

Step 2: Break the RNA into small fragments.

We do this because RNA transcripts can be thousands of bases long, but the sequencing machine can only sequence short (200-300 bp) fragments

Preparing an RNA-seq library

Step 1: Isolate the RNA

Step 2: Break the RNA into small fragments.

Step 3: Convert the RNA fragments into double stranded DNA.

Double stranded DNA is more stable than RNA and can be easily amplified and modified. This leads us to the next step…

Preparing an RNA-seq library

Step 1: Isolate the RNA

Step 2: Break the RNA into small fragments.

Step 3: Convert the RNA fragments into double stranded DNA.

Step 4: Add sequencing adaptors.

Preparing an RNA-seq library

Step 1: Isolate the RNA

Step 2: Break the RNA into small fragments.

Step 3: Convert the RNA fragments into double stranded DNA.

Step 4: Add sequencing adaptors.

The adaptors do two things:

1) Allow the sequencing machine to recognize the fragments.

2) Allow you to sequence different samples at the same time, since different samples can use different adaptors.

Preparing an RNA-seq library

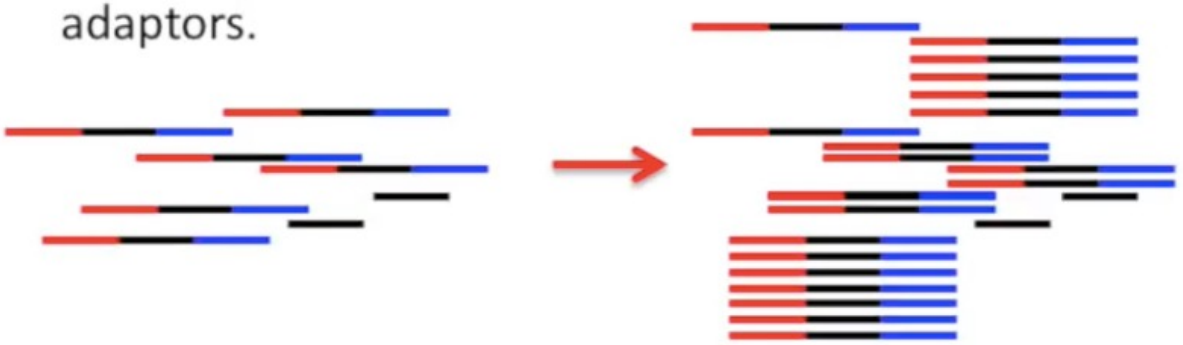
Step 1: Isolate the RNA

Step 2: Break the RNA into small fragments.

Step 3: Convert the RNA fragments into double stranded DNA.

Step 4: Add sequencing adaptors.

Step 5: PCR amplify.



Preparing an RNA-seq library

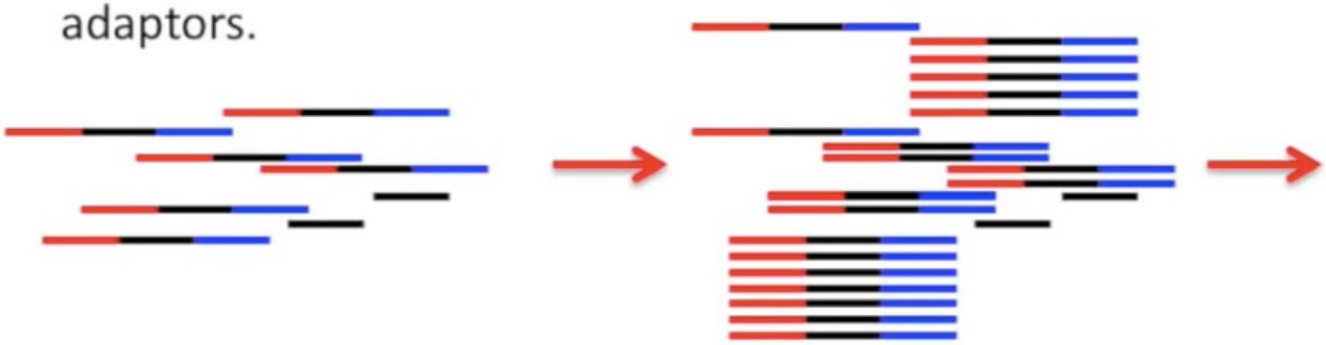
Step 1: Isolate the RNA

Step 2: Break the RNA into small fragments.

Step 3: Convert the RNA fragments into double stranded DNA.

Step 4: Add sequencing adaptors.

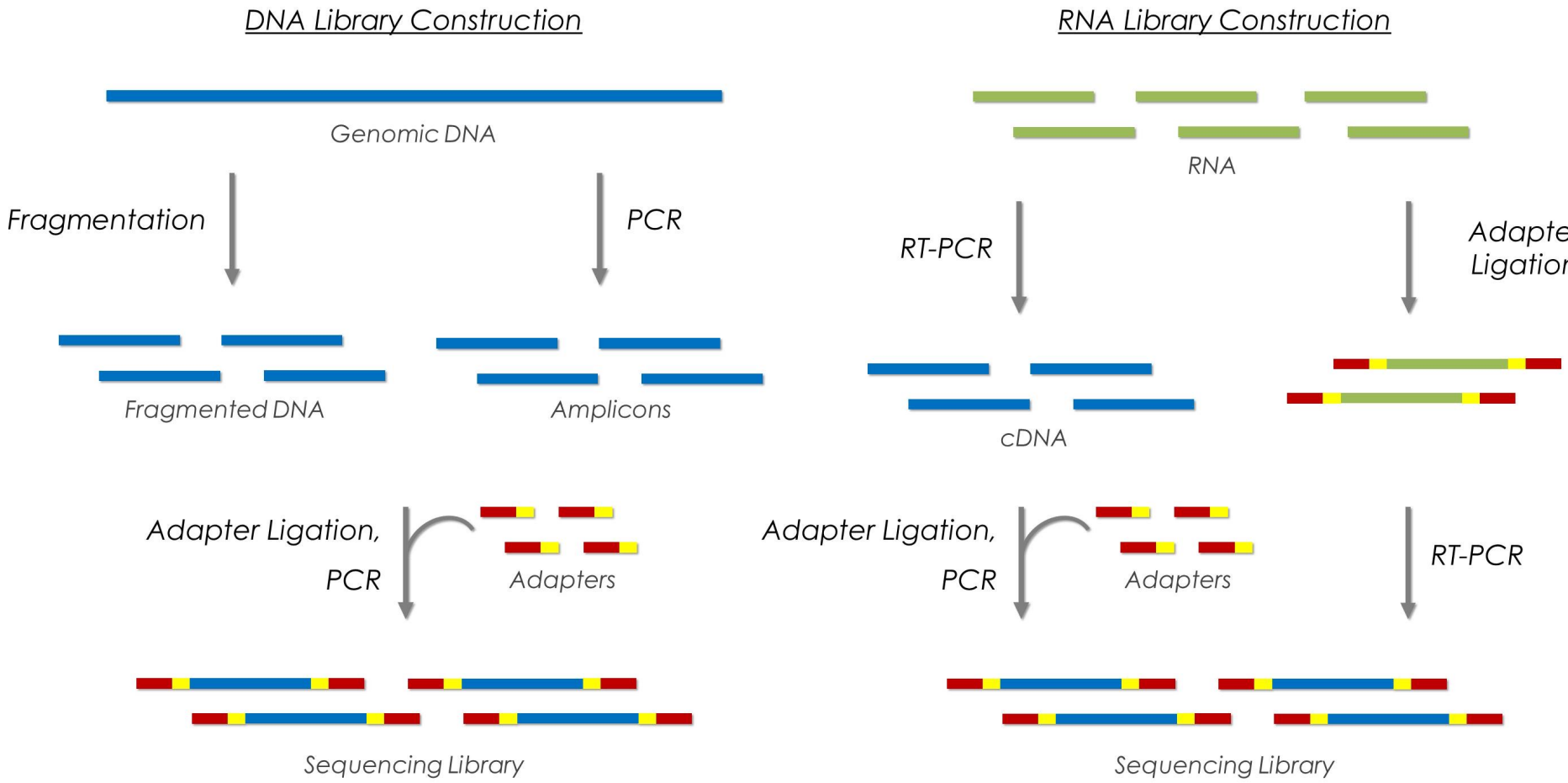
Step 5: PCR amplify.



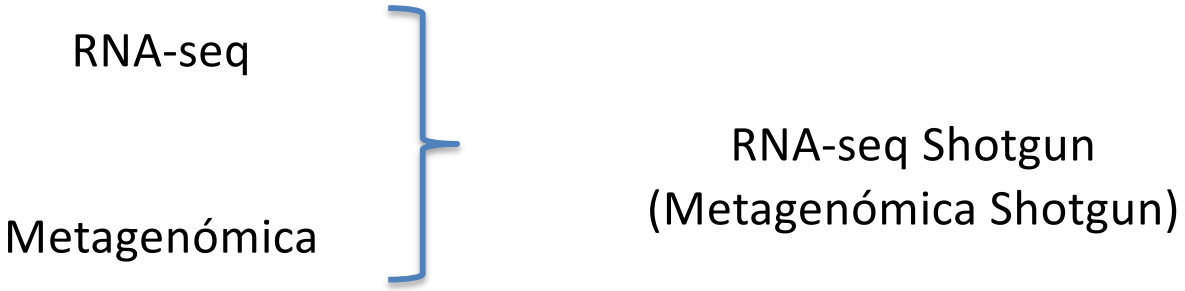
Step 6: QC

1) Verify library concentration

2) Verify library fragment lengths



http://media.biocompare.com/



- **Análisis metagenómico mejorado:** Se han desarrollado nuevos métodos, como PT-seq, para analizar la epigenética de los fosforotioatos en las comunidades microbianas, lo que proporciona una visión más profunda de las funciones e interacciones microbianas.
- **Procesamiento de datos mejorado:** Herramientas como Lemur y Magnet se han optimizado para elaborar perfiles taxonómicos ligeros y precisos a partir de conjuntos de datos metagenómicos de "Shotgun" de lectura larga, reduciendo los falsos positivos y mejorando la precisión de los datos.
- **Perfiles completos del microbioma:** Los estudios han utilizado la metagenómica Shotgun para elaborar perfiles de comunidades microbianas en diversos entornos, como lagos periurbanos y el intestino humano, revelando importantes conocimientos ecológicos y funcionales.
- **Integración con la multiómica:** Integración de metagenómica Shotgun con otros datos ómicos (por ejemplo, la metabolómica) para proporcionar una visión holística de los ecosistemas microbianos y sus rutas metabólicas.

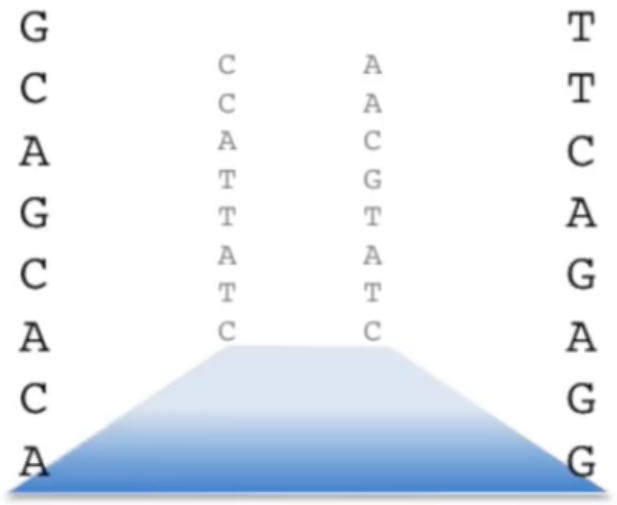
SECUENCIACIÓN DE NUEVA GENERACIÓN (NGS)

Examples of NGS systems

abm

ILLUMINA

Actually, there are about 400,000,000 fragments laid out vertically in a grid.



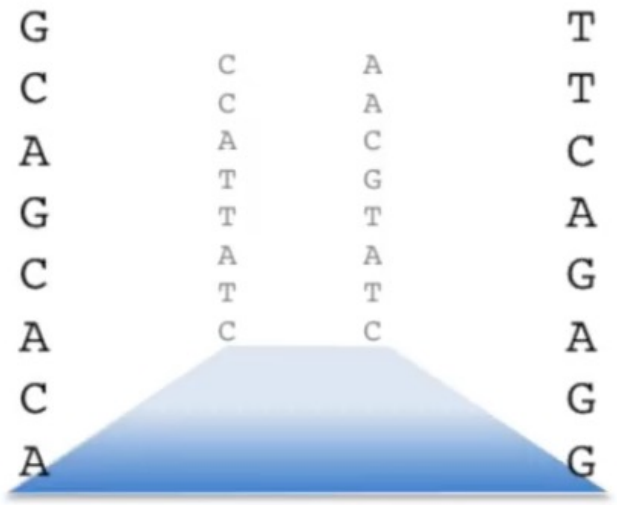
The machine has fluorescent probes that are color coded according to the type of nucleotide they can bind to.

= A

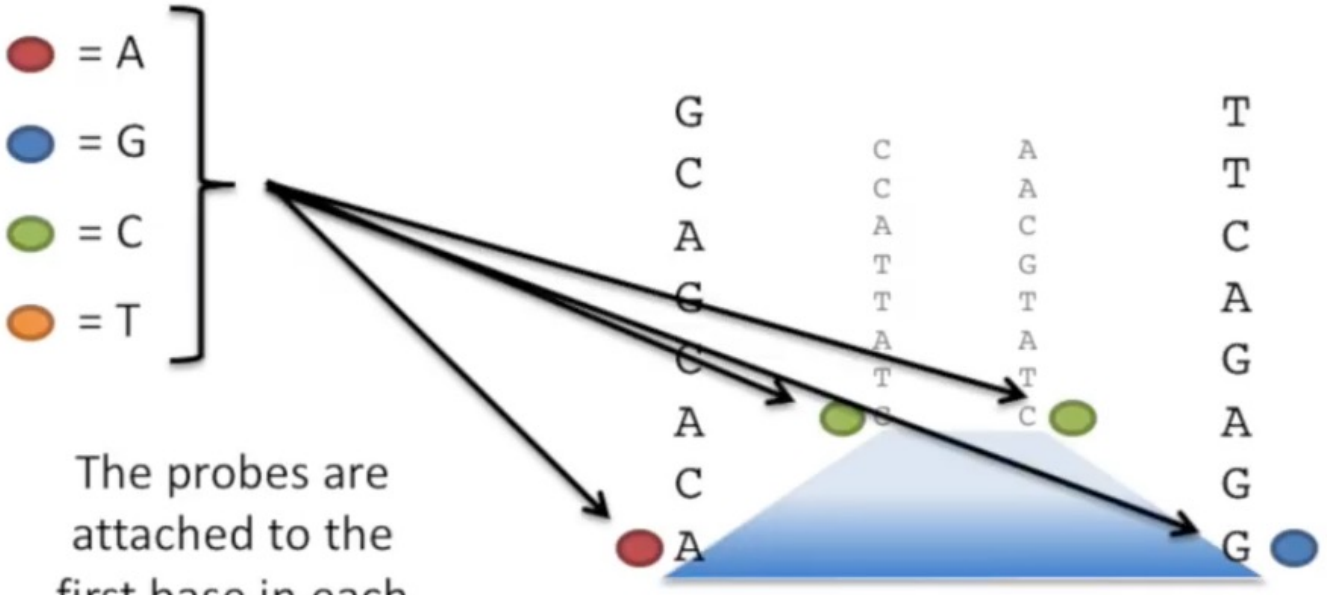
= G

= C

= T

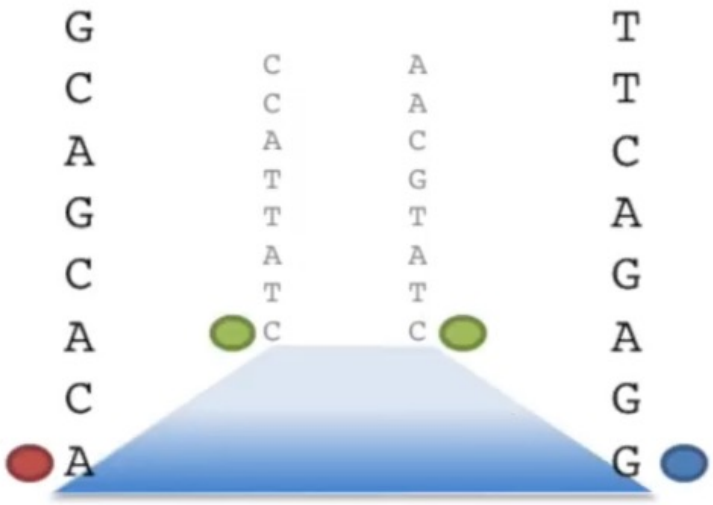


The machine has fluorescent probes that are color coded according to the type of nucleotide they can bind to.

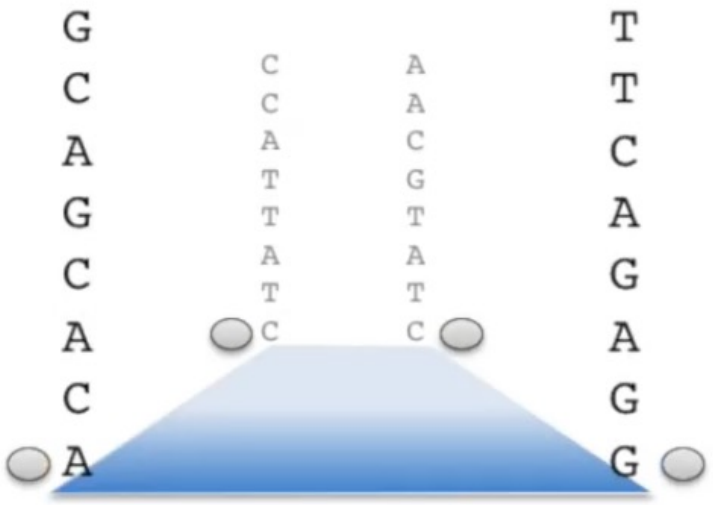


The probes are attached to the first base in each sequence.

Once the probes have attached, the machine takes a picture of the flow cell from above that looks like this…



Then the machine washes the color off of the probes….



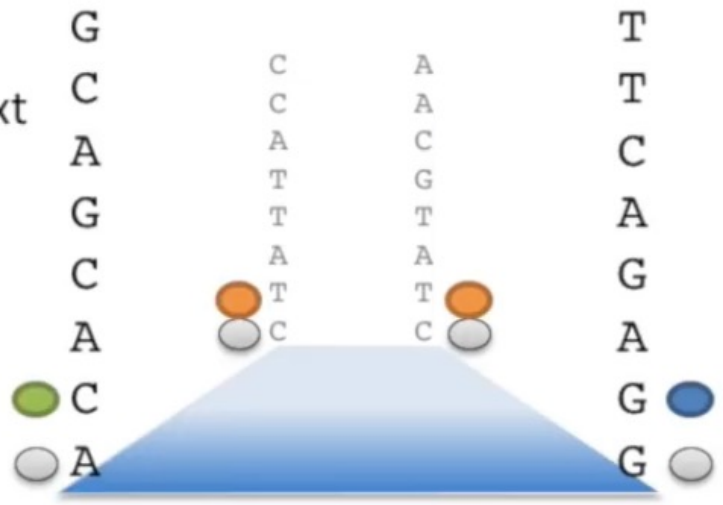
= A

= G

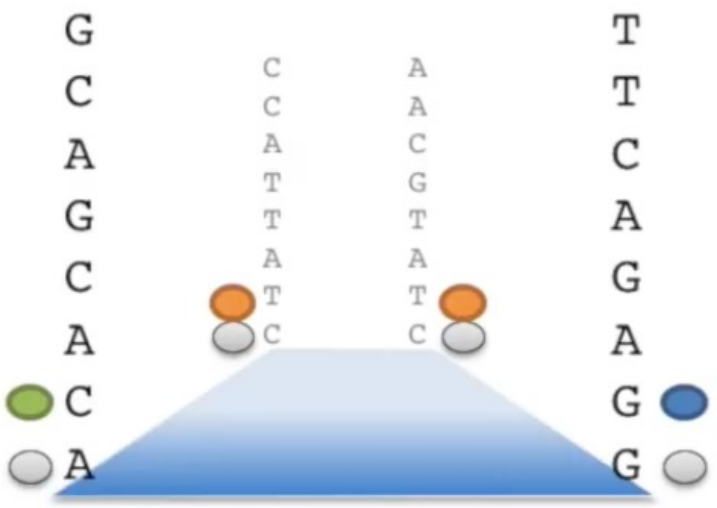
= C

= T

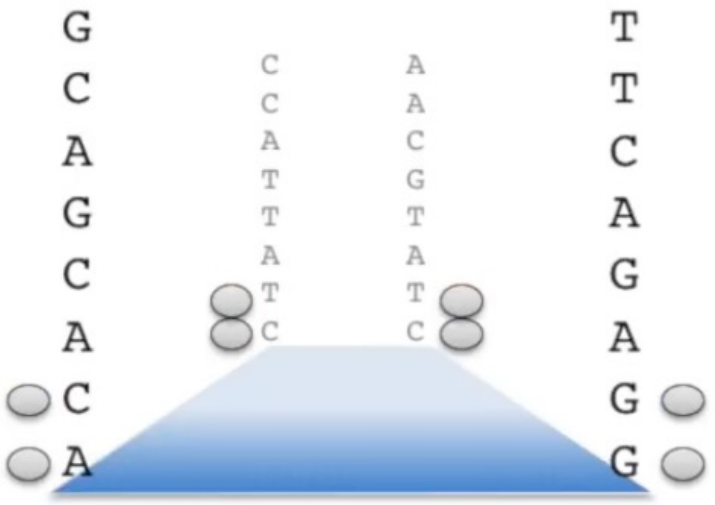
Then probes are bound to the next base in each fragment.



The machine takes a picture from above…



Then the machine washes the color off of the probes….



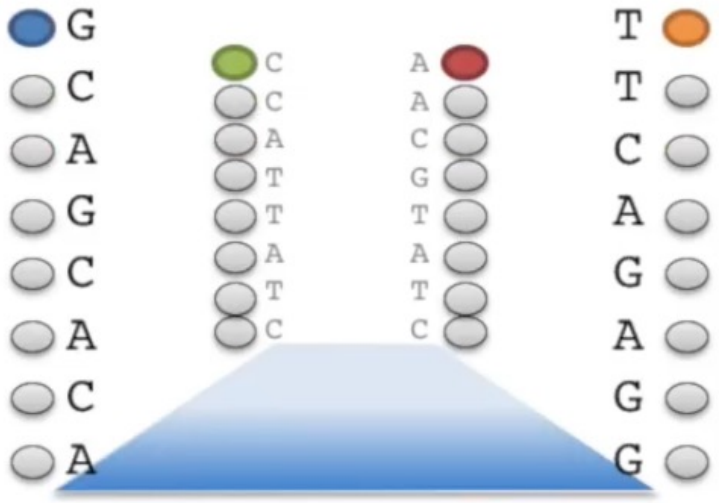
= A

= G

= C

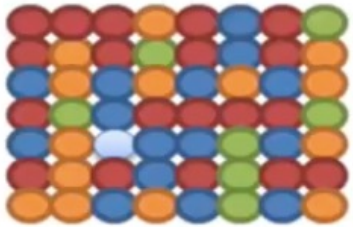
= T

And the process repeats until the machine has determined each sequence of nucleotides.



This is how it works with 4 DNA fragments.

This matrix still isn't 400,000,000 DNA fragments, but it illustrates one type of problem that can occur.



The raw data…

```
@NS500177:196:HFTTTAFXX:1:11101:10916:1458 2:N:0:CGCGGCTG
ACACGACGATGAGGTGACAGTCACGGAGGATAAGATCAATGCCCTCATTAAAGCAGCCGGTGTAA
+
AAAAAEEEEEEEEEEE//AEEEAEEEEEEEEEEE/EE/<<EE/AAEEAEE///EEEEAEEEAEA<
```

Each sequencing “read” consists of 4 lines of data.

A typical sequence run with 400,000,000 reads will generate a file containing 1.6 billion lines of data!!!

Count reads per gene

Once we know the chromosome and position for a read, we can see if it falls within the coordinates of a gene (or some other interesting feature.)

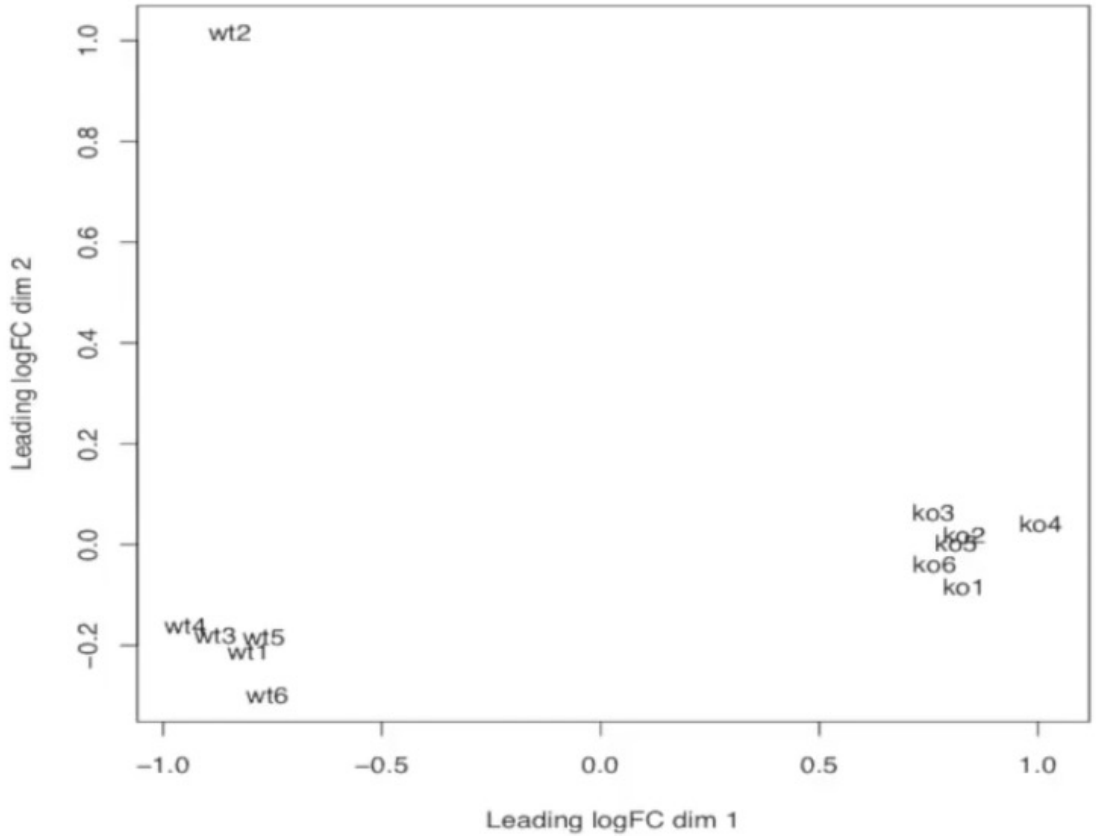
Xkr4 – Chromosome 1, position: 3204563-3661579
Rp1 – Chromosome 1, position: 4280927-4399322

etc.. (for all 20,000 genes in the genome)

Gene	Sample1	Sample2	Sample3…
A1BG	30	5	13…
A1BG-AS1	24	10	18…
A1CF	0	0	0…
A2M	5	9	7…
A2M-AS1	3563	5771	4123…
A2ML1	13	8	7…
...	...	...	...

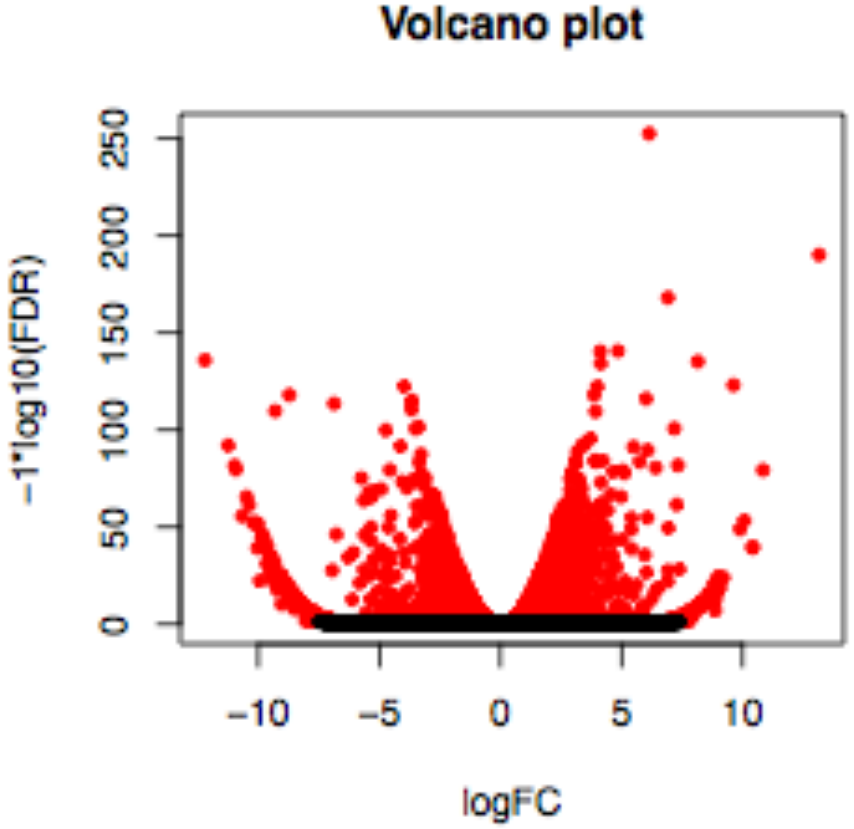
If this were a Single Cell RNA-seq experiment, we would have 20,000 rows (genes) by 800+ columns (samples), giving us at least 16 million values to keep track of

This is a PCA plot from a real RNA-seq experiment done on neural cells.
The “wt” samples are “normal”.
The “ko” samples are samples that were mutated by the researchers.



Step 2) Identify differentially expressed genes between the “normal” and “mutant” samples.

This is typically done using R with either **edgeR** or **DESeq2**, and the results are generally displayed using this sort of graph.



Muestra 1

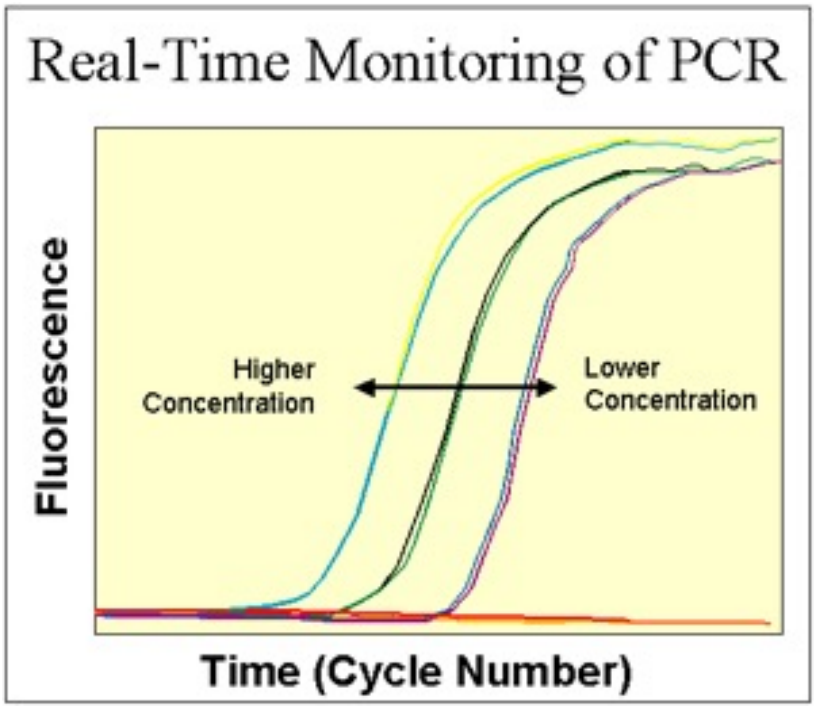
Muestra 2

Muestra 1

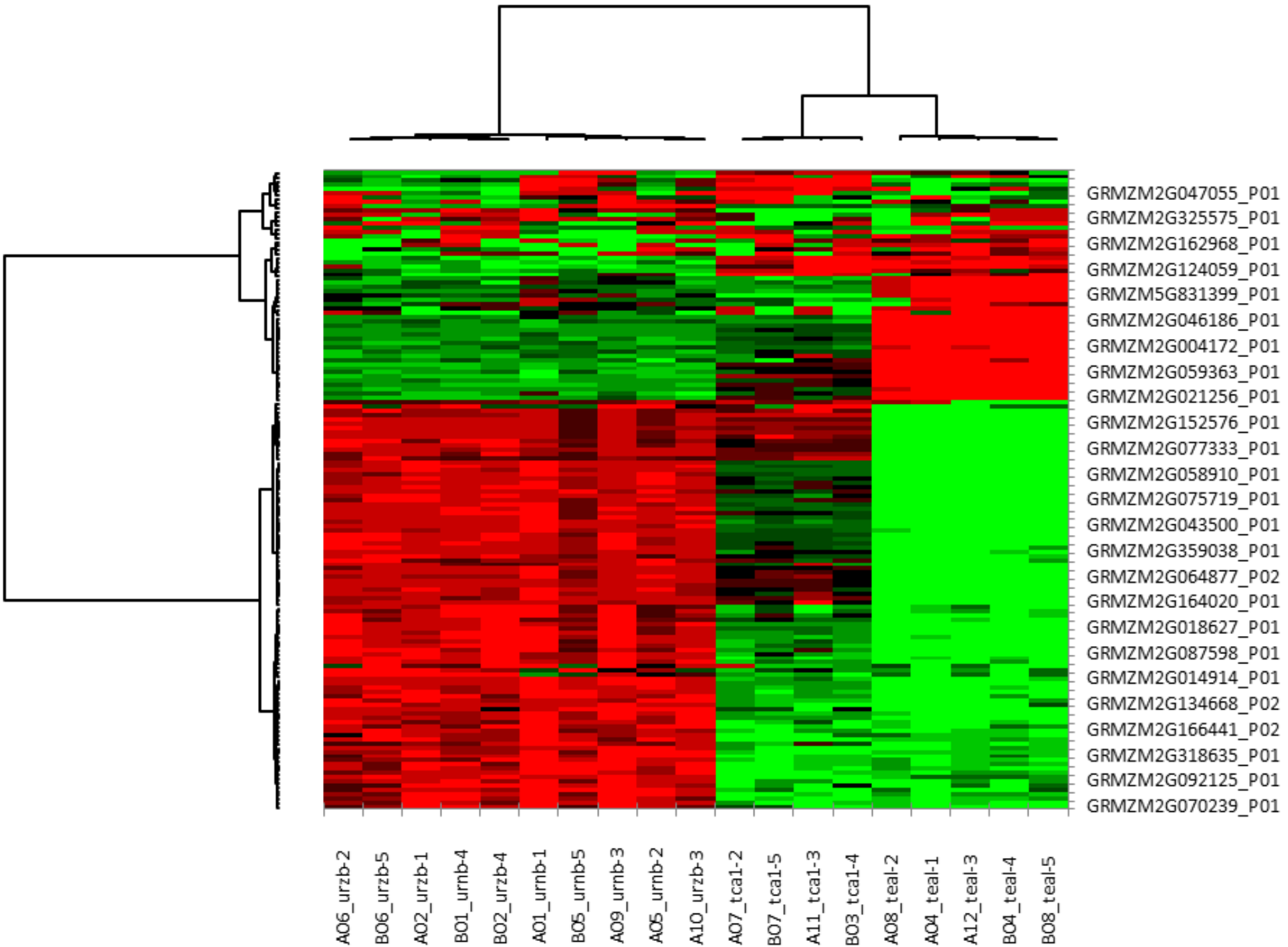
```
CCCACATCTTCTCCATCTCCGACAACGCCTATCAGTACATGCTGACAGGTGAGAGGCCCTGGAA
    ATCTTCTCCATCTCCGACAACGCCTATCAGTACATGCTGACAGGTGAGAG
    ATCTTCTCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAG
    ATCTTCTCCATCTCCGACAACGCCTATCAGTACATGCTGACAGGTGAGAG
    ATCTTCTCCATCTCCGACAACGCCTATCAGTACATGCTGACAGGTGAGAG
    ATCTTCTCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAG
     TCTTCTCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAGG
     TCTTCTCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAGG
     TCTTCTCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAGG
     TCTTCTCCATCTCCGACAACGCCTATCAGTACATGCTGACAGGTGAGACT
     TCTTCTCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAGG
     TCTTCTCCATCTCCGACAACGCCTATCAGTACATGCTGACAGGTGAGAGG
     TCTTCTCCATCTCCGACAACGCCTATCAGTACATGCTGACAGGTGAGAGG
     TCTTCTCCATCTCCGACAACGCAATATCAGTACATGCTGACAGGTGAGAGG
     TCTTCTCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAGG
     TCTTCTCCATCTCCGACAACGCCAATCAGTACATGCTGACAGGTGAGAGG
      CTTCTCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAGGC
      CTTCTCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAGGC
      CTTCTCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAGGC
      CTTCTCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAGGC
       TTCTCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAGGCC
       TTCTCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAGGCC
        TCTCCATCTCCGACAACGCCTATCAGTACATGCTGACAGGTGAGAGGCCC
         CTCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAGGCCCT
         CTCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAGGCCCT
         CTCCATCTCCGACAACGCCTATCAGTACATGCTGACAGGTGAGAGGCCCT
          TCCATCTCCGACAACGCCCATCAGTACATGCTGACAGGTGAGAGGCCCTG
```

Muestra 2

```
ACGGATCAATGTAATGAACCGTGGGGATGACACCCTCCTGCATCTGGCAGCCAGTCATGGA
ACGGATCAATGTAATGAACCGTGGGGATGACACCC
ACGGATCAATGTAATGAACCGTGGGGATGACACCC
ACGGATCAATGTAATGAACCGTGGGGATGACACCC
ACGGATCAATGTAATGAACCGTGGGGATGACAC
```



HEAT MAP


GRMZM2G047055_P01
GRMZM2G325575_P01
GRMZM2G162968_P01
GRMZM2G124059_P01
GRMZM5G831399_P01
GRMZM2G046186_P01
GRMZM2G004172_P01
GRMZM2G059363_P01
GRMZM2G021256_P01
GRMZM2G152576_P01
GRMZM2G077333_P01
GRMZM2G058910_P01
GRMZM2G075719_P01
GRMZM2G043500_P01
GRMZM2G359038_P01
GRMZM2G064877_P02
GRMZM2G164020_P01
GRMZM2G018627_P01
GRMZM2G087598_P01
GRMZM2G014914_P01
GRMZM2G134668_P02
GRMZM2G166441_P02
GRMZM2G318635_P01
GRMZM2G092125_P01
GRMZM2G070239_P01
A06_urzb-2
B06_urzb-5
A02_urzb-1
B01_urnb-4
B02_urzb-4
A01_urnb-1
B05_urnb-5
A09_urnb-3
A05_urnb-2
A10_urzb-3
A07_tca1-2
B07_tca1-5
A11_tca1-3
B03_tca1-4
A08_teal-2
A04_teal-1
A12_teal-3
B04_teal-4
B08_teal-5