

# Ingeniería Genética II

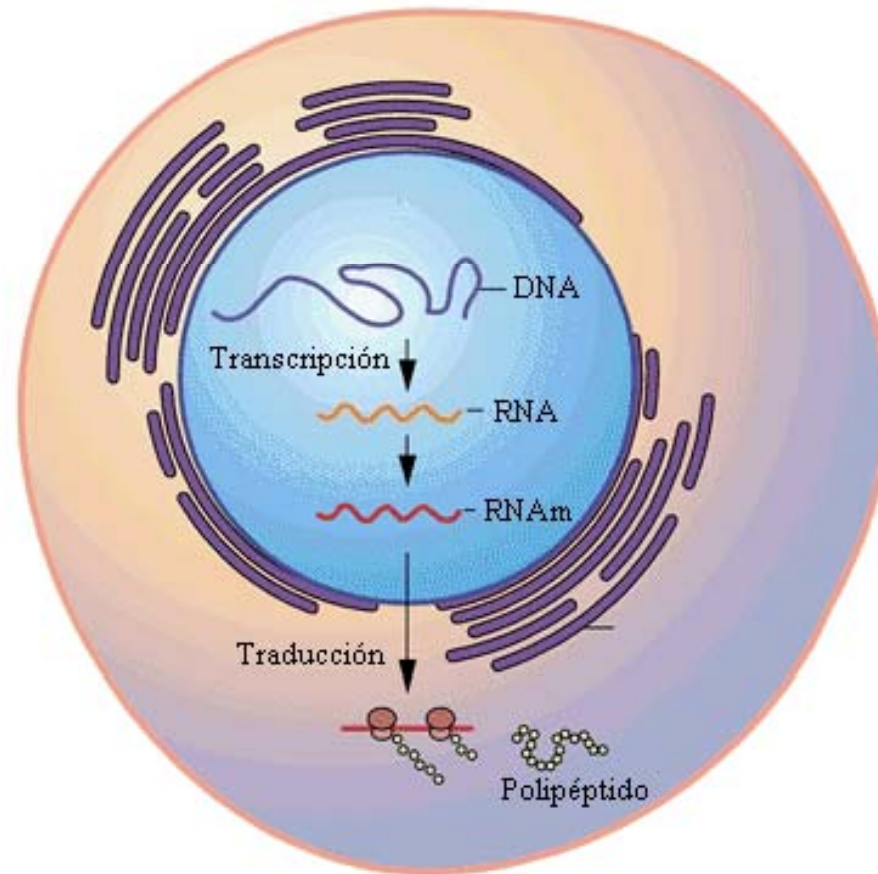
## Unidad V

### *Análisis de transcritos*

# Estudios de transcriptos

- En toda la **materia viva**, la **interacción** del **ambiente** con el **genoma** modula la producción de **transcriptos**.
- Los **transcriptos** son los **mensajes codificados** en las moléculas genómicas, los cuales tienen **diversas funciones celulares**.
- Algunos transcriptos son necesarios para la **producción de proteínas**, mientras que otros son **regulatorios** o presentan funciones específicas diferenciales.
- Los **transcriptos** forman la **primera expresión del fenotipo** y por ende su determinación y cuantificación pueden ser de gran utilidad para la caracterización celular.

# Estudios de transcriptos



# Estudios de transcriptos

Algunas variantes de RNA que aparecen en una célula



**mRNA** (*messenger*, mensajeros para producción de proteínas)

**rRNA** (*ribosome*, ribosomas para la traducción de proteínas)

**tRNA** (*transfer*, transferentes de aminoácidos para la traducción de proteínas)

**srpRNA** (*Signal Recognition Particle*, participa en el tráfico de proteínas a membranas).

**snRNA** (*small nuclear*, participan en el procesamiento de otros RNAs)

**aRNA** (*antisense*, actúa como regulador)

**miRNA** (*micro*, actúa como regulador)

**siRNA** (*small interfering*, actúa como regulador)

**Retrotransposones**

**RNA viral endógeno**

**y otros....**

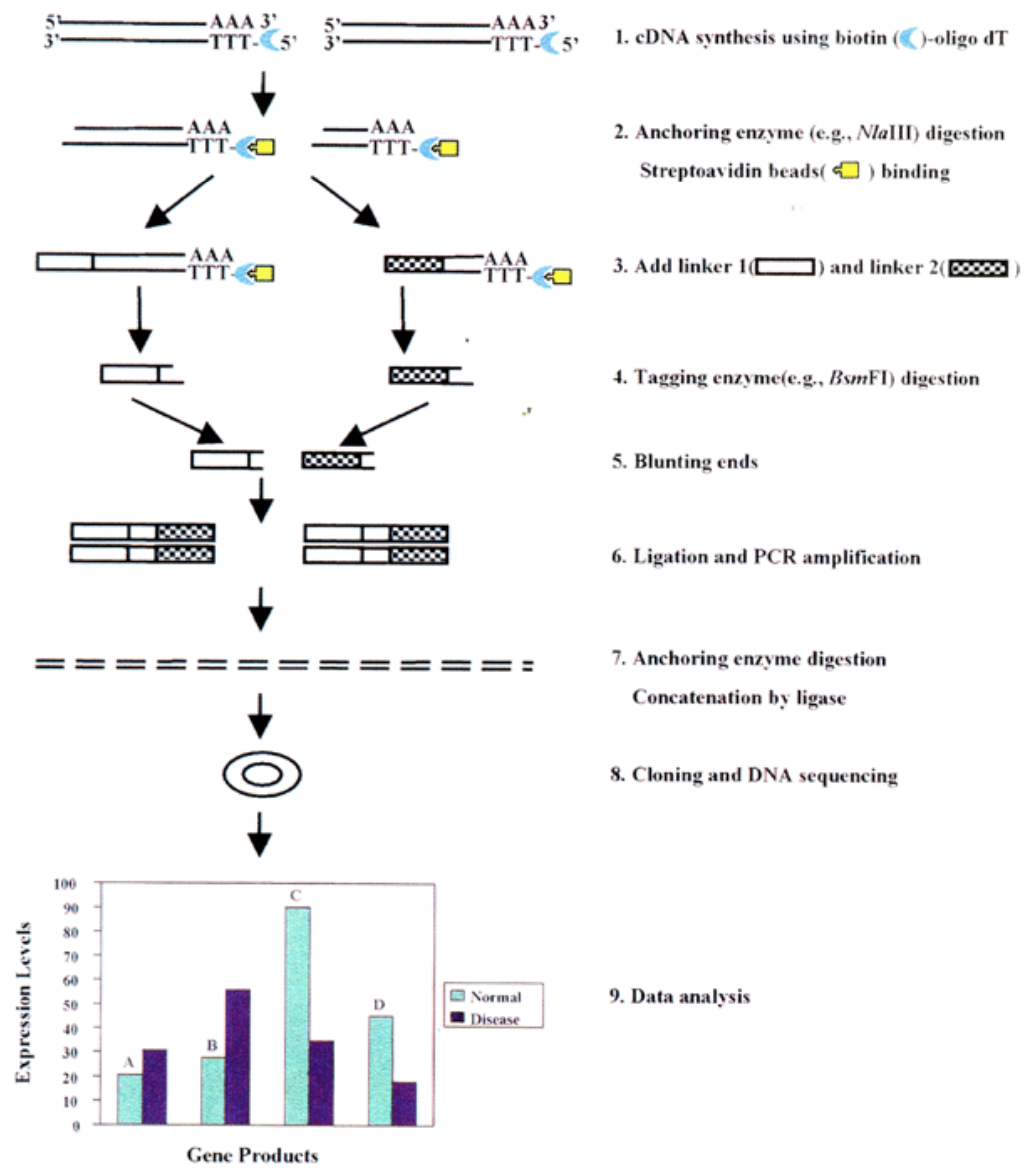
# Estudios de transcriptos

- Por **ejemplo**, el **genoma humano** tiene  **$3 \times 10^9$  pb**, alrededor de  **$1 \times 10^{14}$  células** y cada una de ellas con alrededor de  **$3 \times 10^5$  transcriptos**.
- Si analizáramos el **transcriptoma global** de un ser humano (en un momento dado), deberíamos considerar que su dimensión aproximada es de  **$8,4 \times 10^{23}$  moléculas distintas\***.
- Y por supuesto es necesario tener en cuenta que la **vida media** de cada una de esas moléculas es **distinta**, y que su **presencia o ausencia varía minuto a minuto**.
- En síntesis, analizar un genoma es una tarea mucho más sencilla que abarcar toda la posible diversidad transcriptómica que ese genoma puede producir.

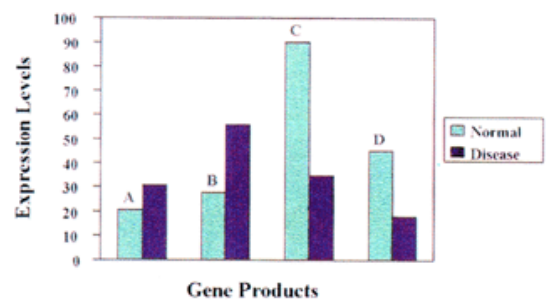
# Estudios de transcriptos

- Conociendo la **secuencia** de un **genoma**, es **posible estudiar** el **transcriptoma** mediante técnicas y procedimientos estándar.
- Entre ellos, valen mencionarse los **ensayos de hibridación** (*Northern Blots, Dot Blots, Slots Blots*), procedimientos de **amplificación específica** como la RT-PCR (*Reverse Transcription PCR*), **bibliotecas transcriptómicas** (generadoras de bases de datos EST), **RACEs** y **SAGE** (*Serial Analysis of Gene Expression*).
- Sin embargo, cuando desean hacerse **detecciones globales** de la célula, los **ensayos anteriores** se transforman en demasiado **complejos** y **costosos** en tiempo y recursos.
- Ante este escenario, surgieron **técnicas high throughput** para el **análisis transcriptómico**.

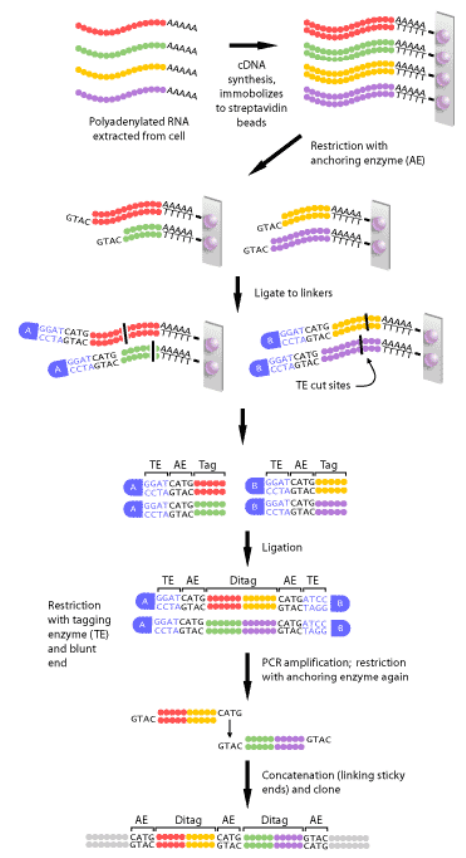
# Estudios de transcriptos



## SAGE



## *NlaIII* CATG/ *BsmFI* GGGAC(N)<sub>10</sub>/(N)<sub>14</sub>



# Estudios de transcriptos

Procedimientos actuales para el análisis transcriptómico

Mediante **Hibridación**

**MICROARREGLOS**

Preferentemente para mRNAs

Mediante **Secuenciación**

**RNA Seq**

Para todo tipo de RNA



**Microarreglos  
para el análisis  
global de  
transcriptos**

# Estudios de transcriptos

## Microarreglos

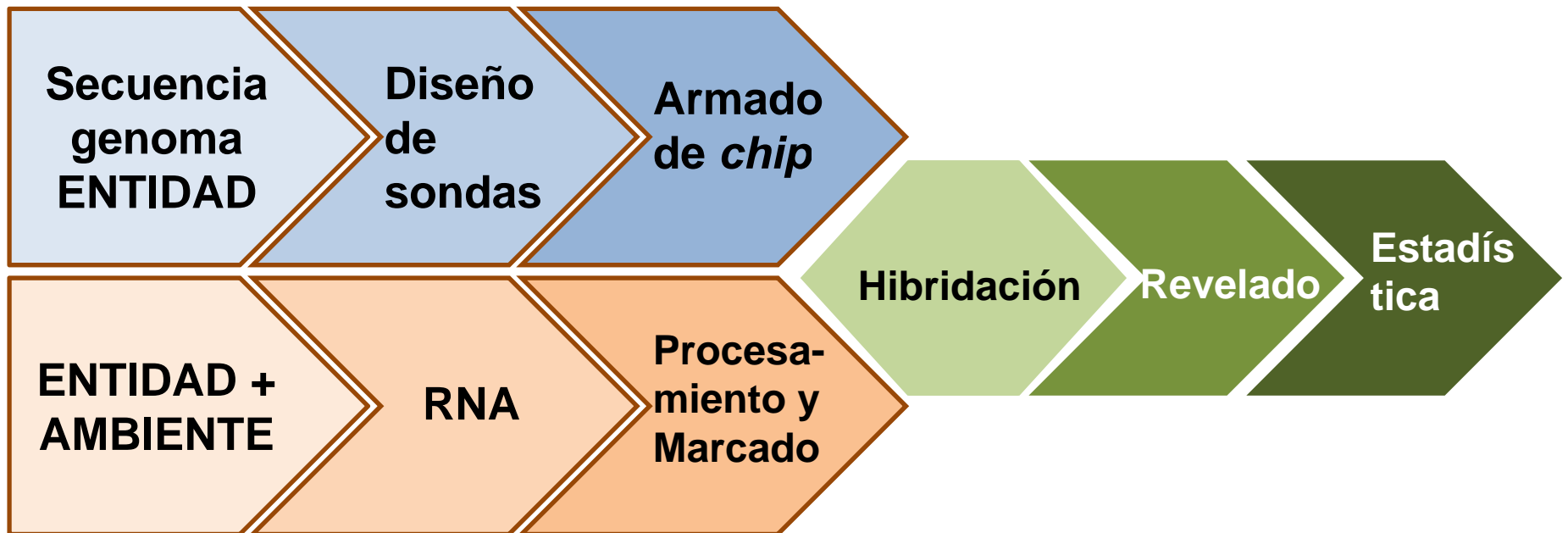
- La **tecnología de microarreglos** permite la **hibridación simultánea** de miles de secuencias conocidas inmovilizadas con ácidos nucleicos presentes en una muestra problema.
- Si el **microarreglo** se diseña de modo tal de **contener la información posible de transcriptos** que podría expresar una **célula** (principalmente mRNAs), se transforma en una **excelente plataforma** para el **análisis transcriptómico**.
- Básicamente, los **microarreglos** para análisis de transcriptos son **similares** a los que se utilizan para estudiar **diversidad poblacional de genomas**, pero concentrados en el **análisis de mRNAs**.

# Estudios de transcriptos

## Microarreglos

Flujo de trabajo para el análisis transcriptómico

1 color



# Estudios de transcriptos

## Microarreglos

Flujo de trabajo para el análisis transcriptómico



Secuencia  
genoma  
ENTIDAD

Diseño  
de  
sondas

Armado  
de *chip*

Adquirida  
por un  
**Proyecto  
Genoma**

En base a la  
**interpretación  
bioinformática**  
de genes

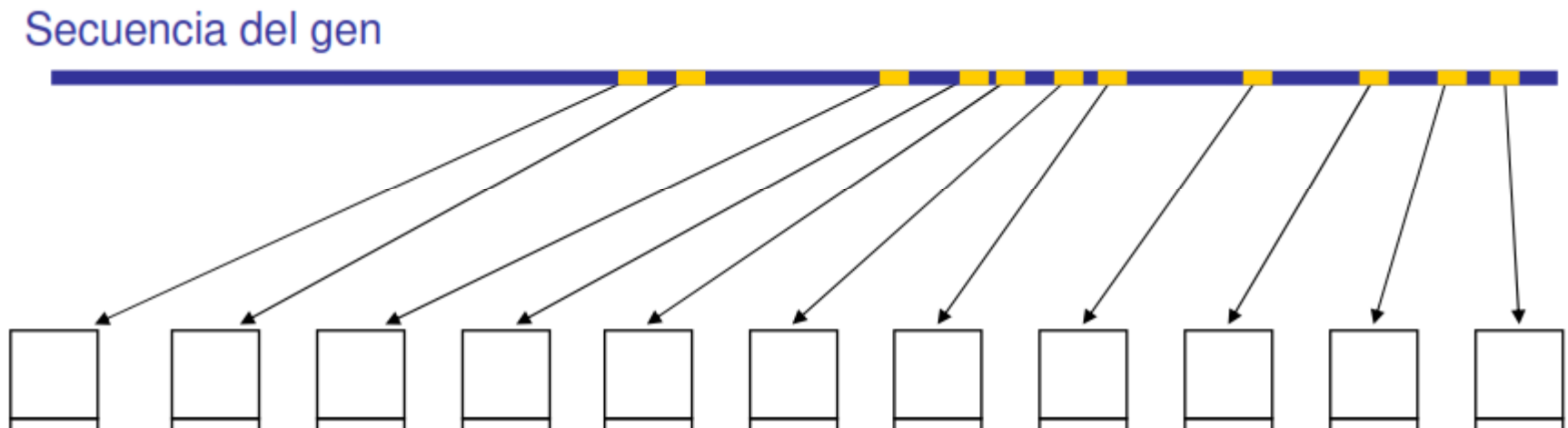
Por **síntesis de  
oligonucleótidos**  
*in situ* o por  
**depósito** directo  
de **EST**.

**Procedimientos  
equivalentes** a los  
utilizados para el  
amado de  
**microarreglos**  
utilizados en  
**análisis poblacional**

# Estudios de transcriptos

## Microarreglos

- Se diseñan **varias sondas** para **un mismo gen**, representando a los distintos exones si es un eucariota.
- A su vez, las distintas **sondas presentan superposición** para así aumentar la seguridad de la detección.



# Estudios de transcriptos

## Microarreglos

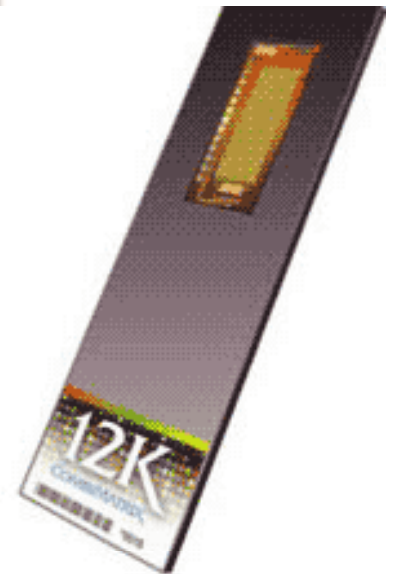
Ejemplo de sondas para detectar el transcripto de un gen

ctcagcttaagtcatggaattctagaggatgtatctcacaagtaggatcaag

|   |     |
|---|-----|
| ctcagcttaagtcatggaattctag               | PM1 |
| ctcagcttaagt <sup>g</sup> atggaattctag  | MM1 |
| tcagcttaagtcatggaattctaga               | PM2 |
| tcagcttaagtc <sup>t</sup> tggaattctaga  | PM2 |
| attctagaggatgtatctcacaagt               | PM3 |
| attctagaggatc <sup>t</sup> tatctcacaagt | MM3 |
| aggatgtatctcacaagtaggatca               | PM4 |
| aggatgtatctc <sup>t</sup> caagtaggatca  | MM4 |

# Estudios de transcriptos

## Microarreglos



Microarreglos

# Estudios de transcriptos

## Microarreglos

Flujo de trabajo para el análisis transcriptómico

1 color

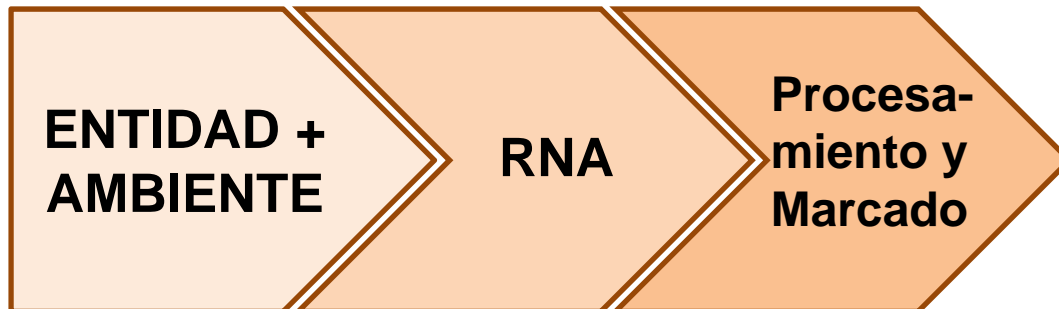
Definición de la muestra y las preguntas biológicas



Aislamiento de RNA total o subpoblaciones (preferentemente mRNA)



Preparación del transcriptoma marcado




Esta es la parte fundamental, dado que involucra el **tratamiento** de la **muestra**




# Estudios de transcriptos

## Microarreglos



**Procesamiento y Marcado**




Existen diferentes aproximaciones...

**1** Marcación del cDNA

**2** Marcación del RNA

# Estudios de transcriptos

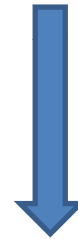
## Microarreglos



Procesamiento y Marcado

### 1 Marcación del cDNA


Se utilizan alrededor de 100 ng de mRNA



- a) Síntesis de cDNA con *primer* polyT y un dNTP biotinilado.
- b) Tratamiento con RNAsa H y RNAsa A.
- c) Fragmentación química.

# Estudios de transcriptos

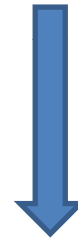
## Microarreglos



Procesamiento y Marcado

### 2 Marcación del RNA

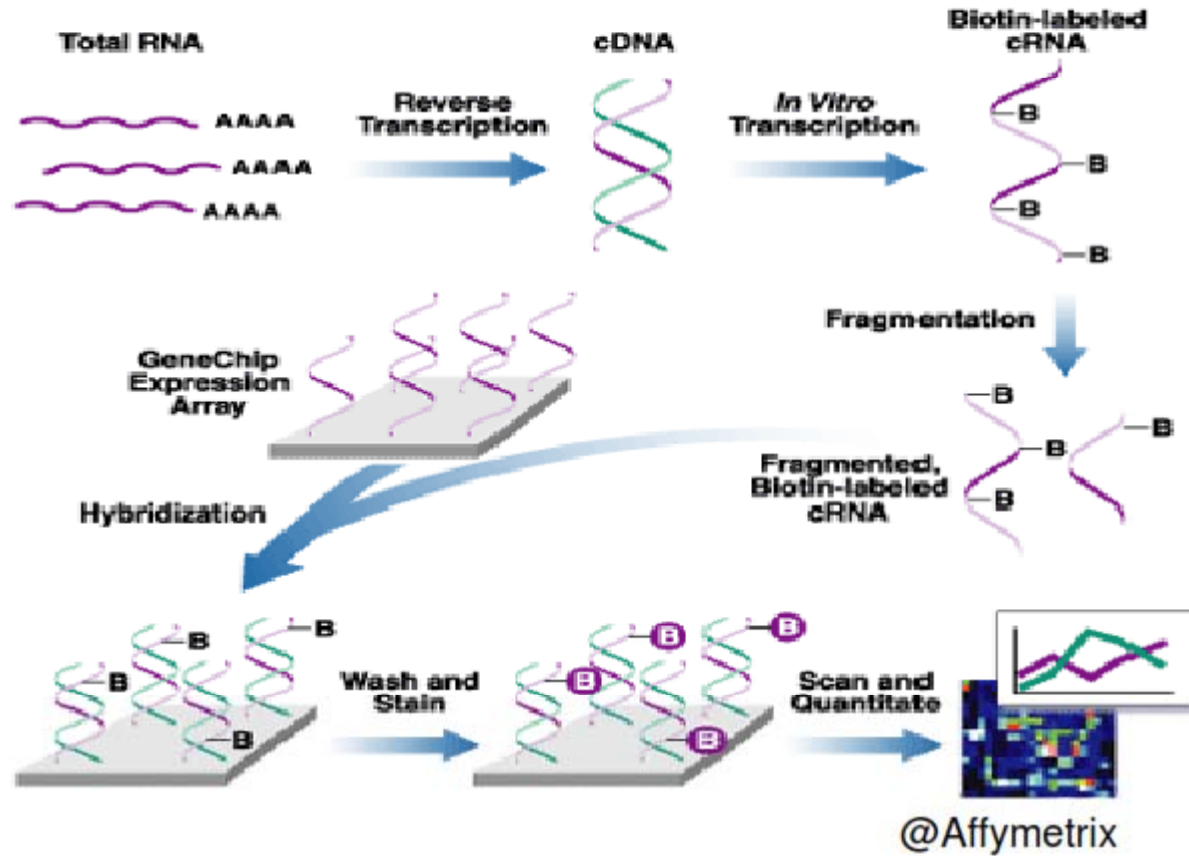
Se utilizan entre 10-100 ng de mRNA



- a) Síntesis de cDNA con *primer* polyT y dNTPs normales.
- b) Síntesis de la segunda cadena de DNA.
- c) Síntesis de cRNA con dUTP biotinilado.
- d) Fragmentación del cRNA marcado

# Estudios de transcriptos

## Microarreglos



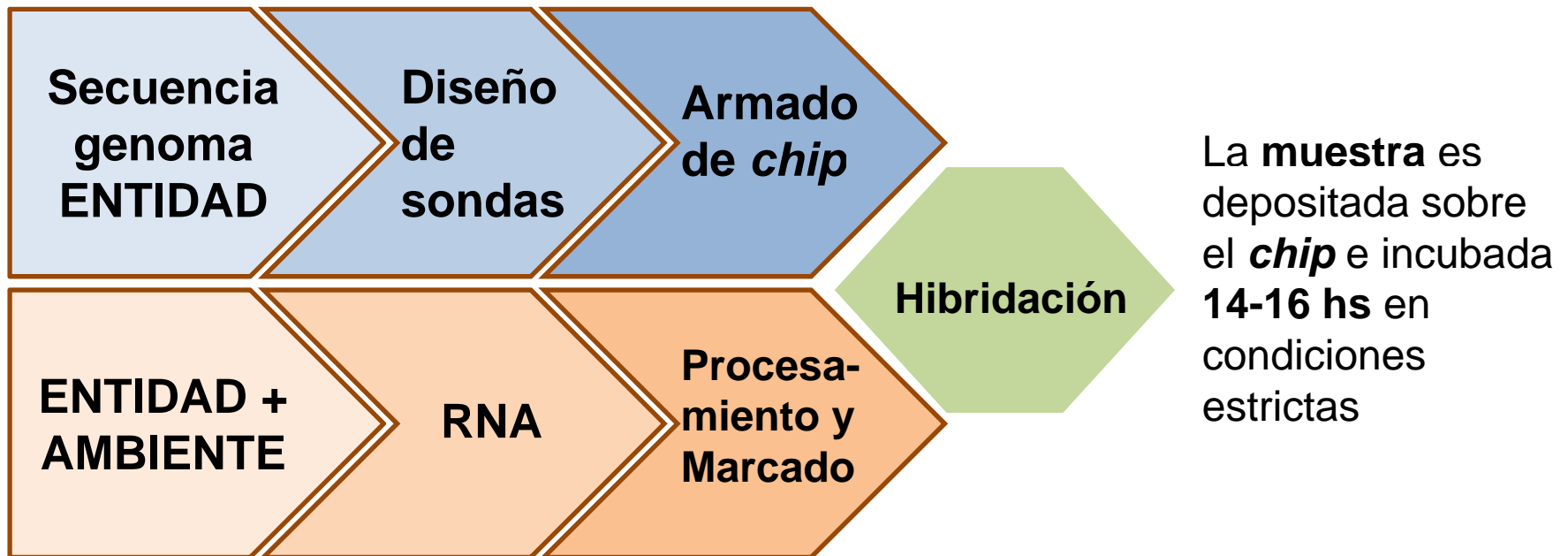
Este procedimiento involucra la amplificación de la muestra

# Estudios de transcriptos

## Microarreglos

Flujo de trabajo para el análisis transcriptómico

1 color

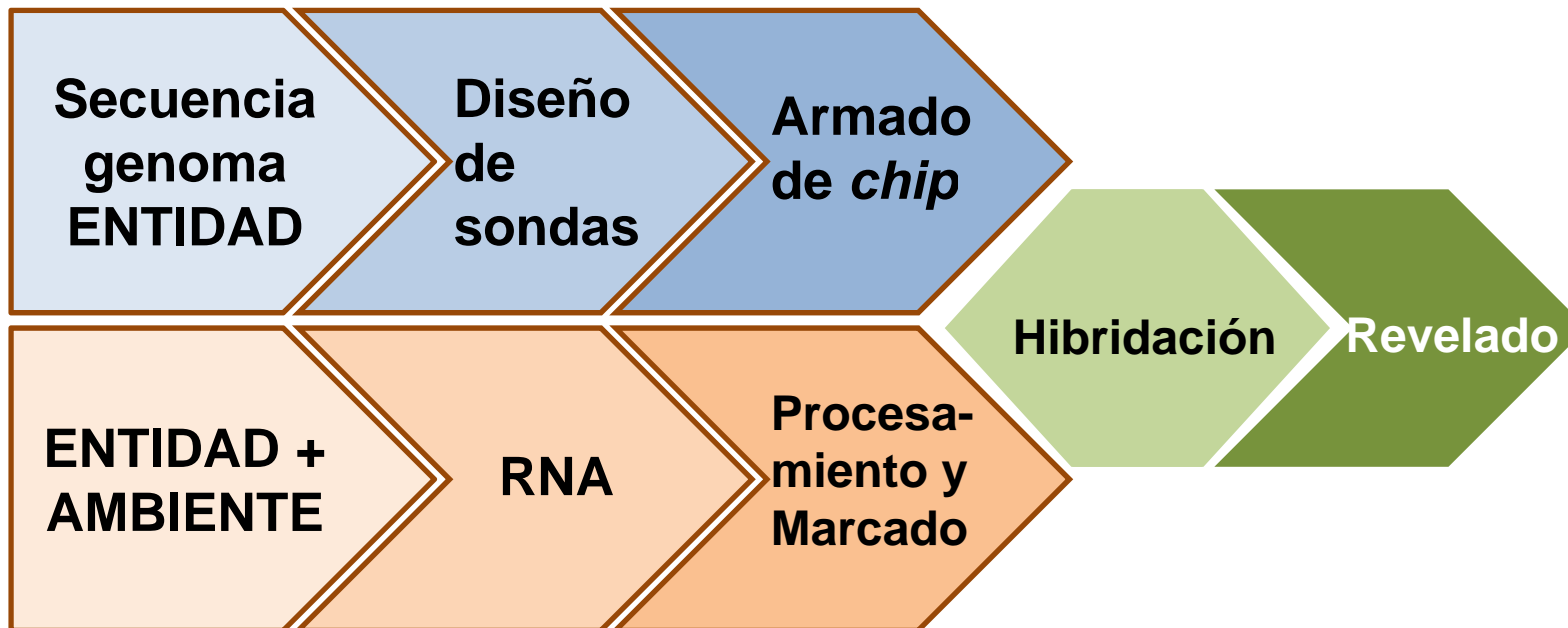


# Estudios de transcriptos

## Microarreglos

Flujo de trabajo para el análisis transcriptómico

1 color



# Estudios de transcriptos

## Microarreglos

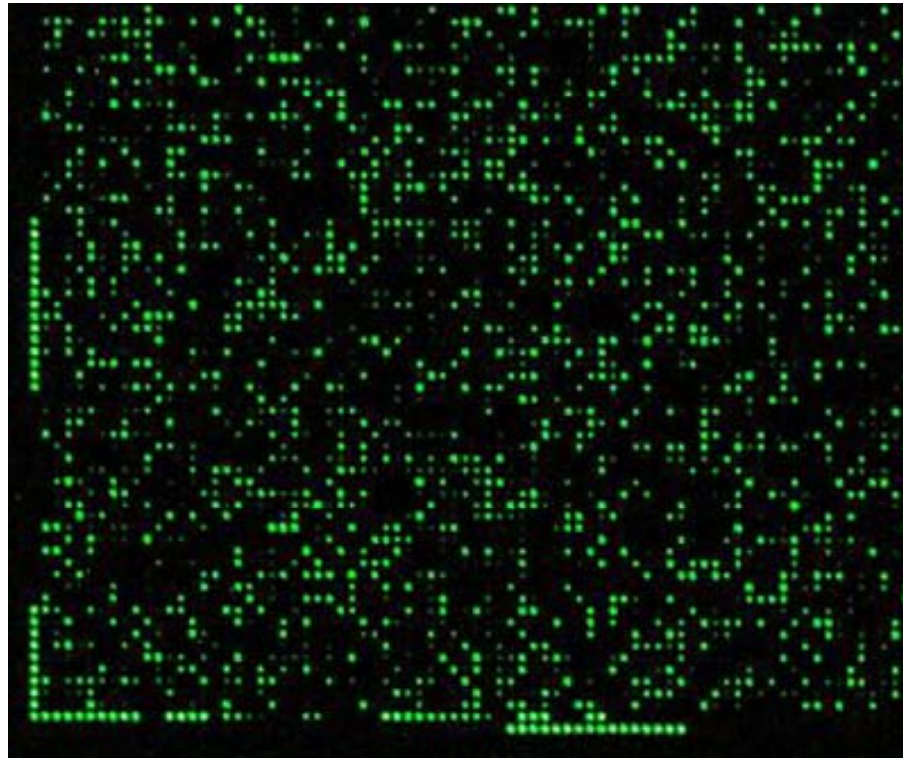
- **Luego** de la **hibridación** se realizan **lavados** para eliminar los DNAs imperfectamente apareados (soluciones de concentración salina decreciente y presencia de detergentes).
- **Después**, se agrega **estreptavidina marcada con un fluoróforo** en presencia de BSA.
- La **estreptavidina** quedará **unida** a las **biotinas** presentes en las muestras derivadas del transcriptoma.
- **Finalmente**, se **excita** el microarreglo con **luz** apropiada y se **registra** la emisión de **fluorescencia** con un escáner.

# Estudios de transcriptos

## Microarreglos

Resultado de un microarreglo revelado

1 color



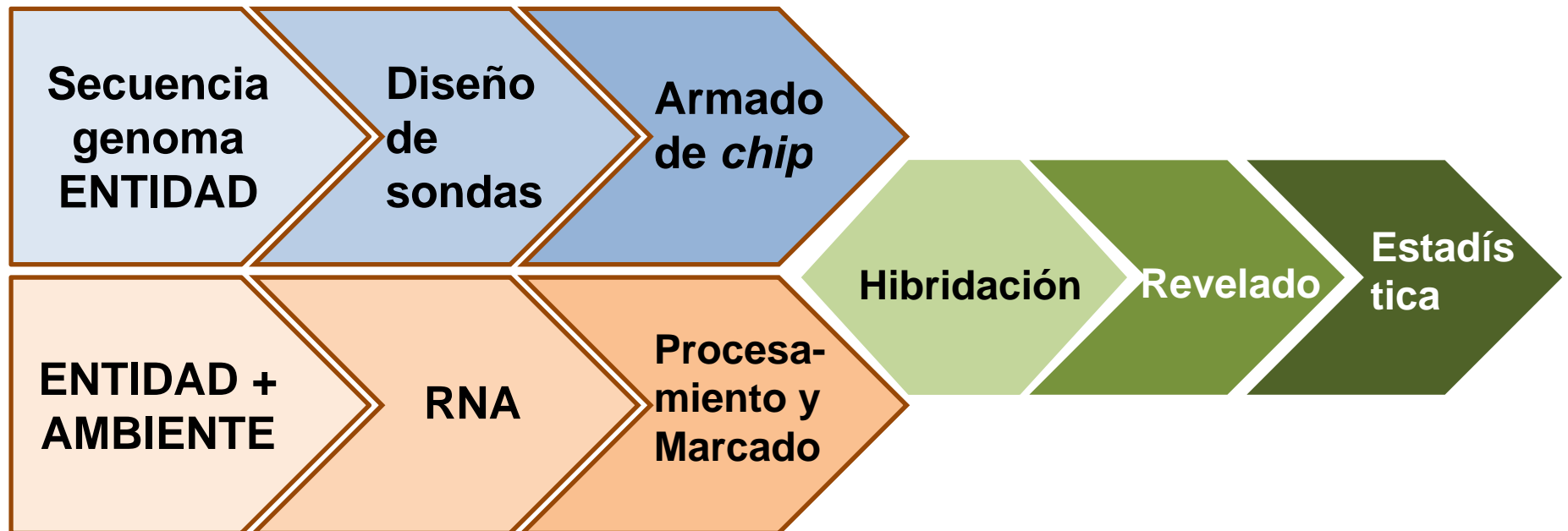


# Estudios de transcriptos

## Microarreglos

Flujo de trabajo para el análisis transcriptómico

1 color



# Estudios de transcriptos

## Microarreglos

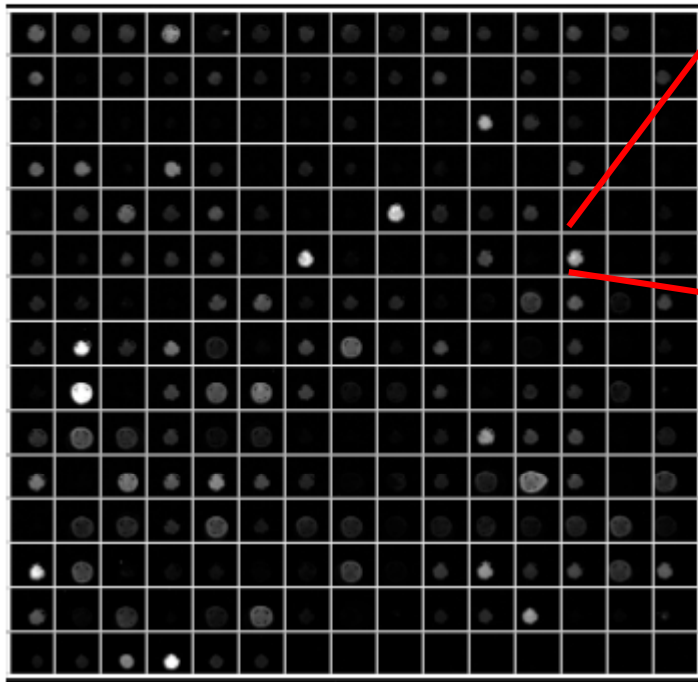
- **Luego de digitalizarse** la imagen del **microarreglo**, es necesario realizar un **análisis estadístico** para definir cuáles son los **transcriptos presentes** en la muestra.
- Esto involucra **comparar** las **señales** obtenidas en cada celda respecto de los **controles**, y de analizar en su **conjunto** todas las **sondas disponibles** para la determinación de **cada gen** (recordemos que se encontraban varias sondas superpuestas de distintas regiones del ORF).
- En función de todo lo anterior, se define el **transcriptoma** de la **muestra** analizada referenciando el rigor estadístico de la determinación.

# Estudios de transcriptos

## Microarreglos

El microarreglo es analizado en cada uno de los *spots*.

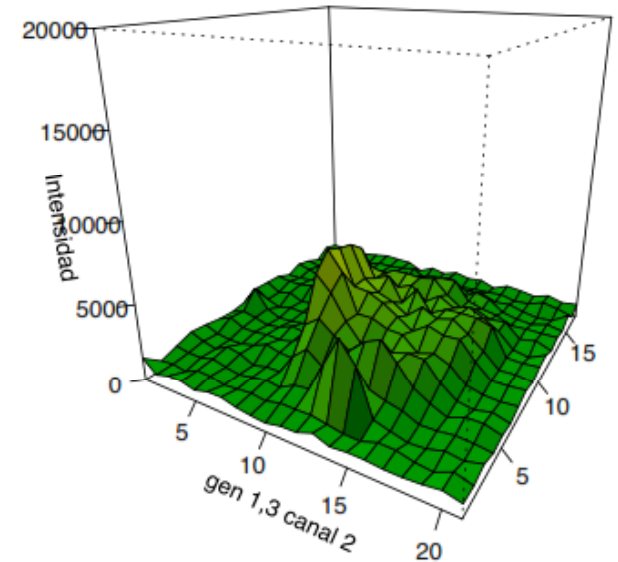
Selección de 1 *Spot*



Grilla del microarreglo



Análisis de la intensidad de señal

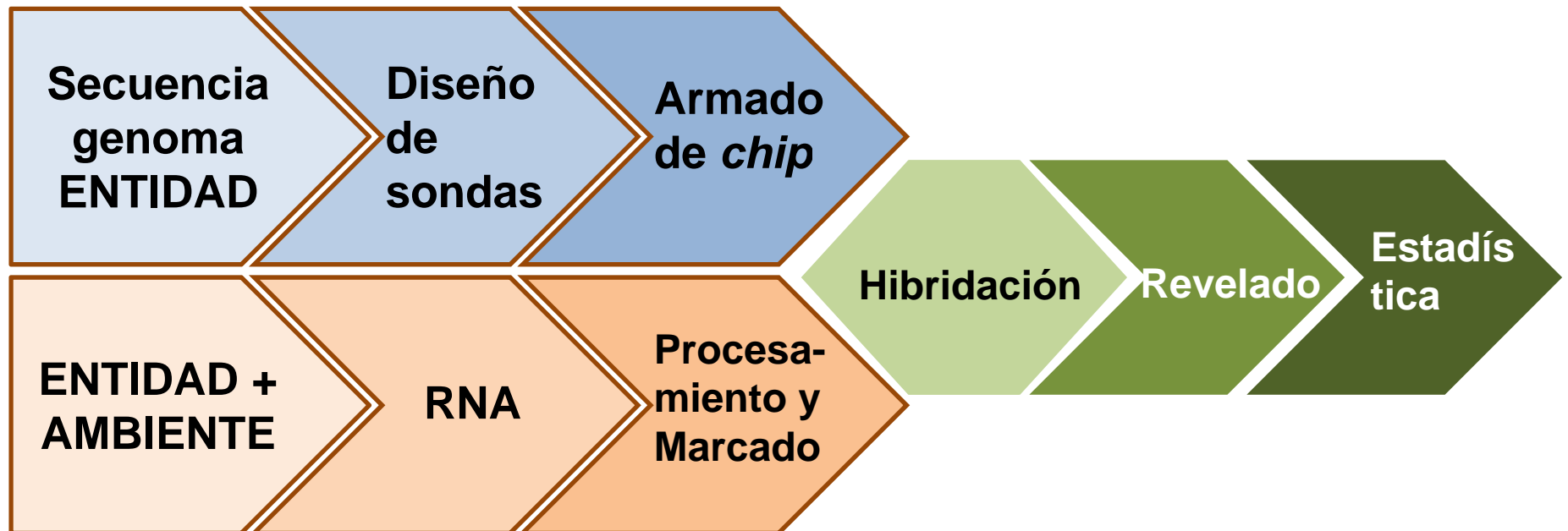


# Estudios de transcriptos

## Microarreglos

Flujo de trabajo para el análisis transcriptómico

2 colores



# Estudios de transcriptos

## Microarreglos

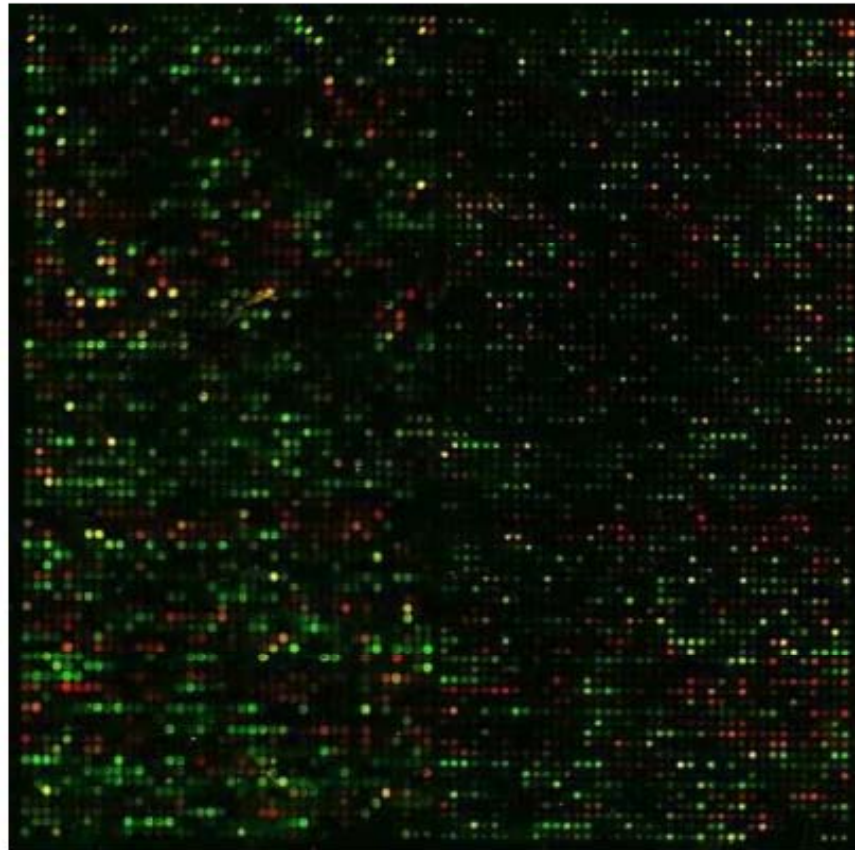
Flujo de trabajo para el análisis transcriptómico

2 colores

- La gran diferencia de los **análisis de 2 colores** respecto a los de 1 color es
- que **dos muestras de RNA** se analizan **simultáneamente**, dado que se busca definir cuáles transcriptos son diferentes en presencia y cantidad.
- Ante esta situación, la diferencia en el flujo de trabajo está centrada en el
- procesamiento de las **muestras**, las cuales se **marcan directamente** con nucleótidos marcados con **fluoróforos distintos**.
  - El **revelado** se hace con la excitación con **dos láseres** y luego se **superponen las imágenes**.

# Estudios de transcriptos

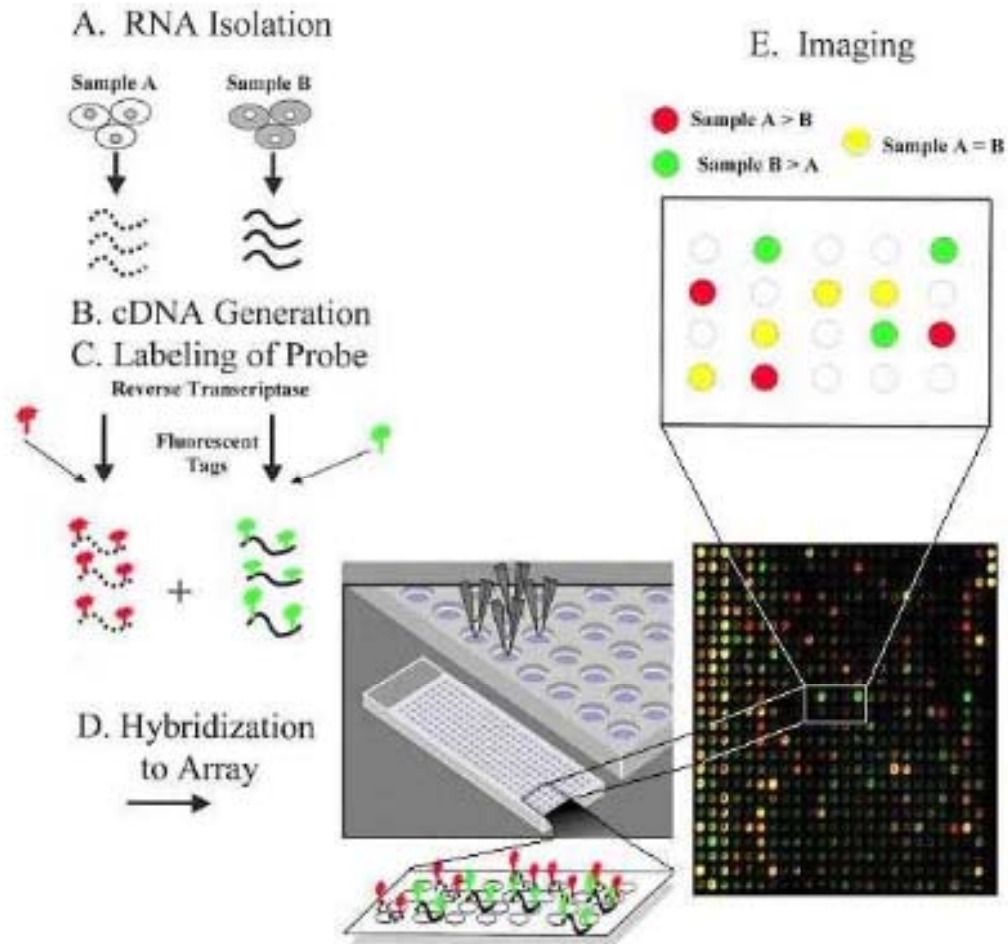
## Microarreglos



Microarreglo de 2 colores

# Estudios de transcriptos

## Microarreglos



# **RNA seq para el análisis global de transcriptos**



# Estudios de transcriptos

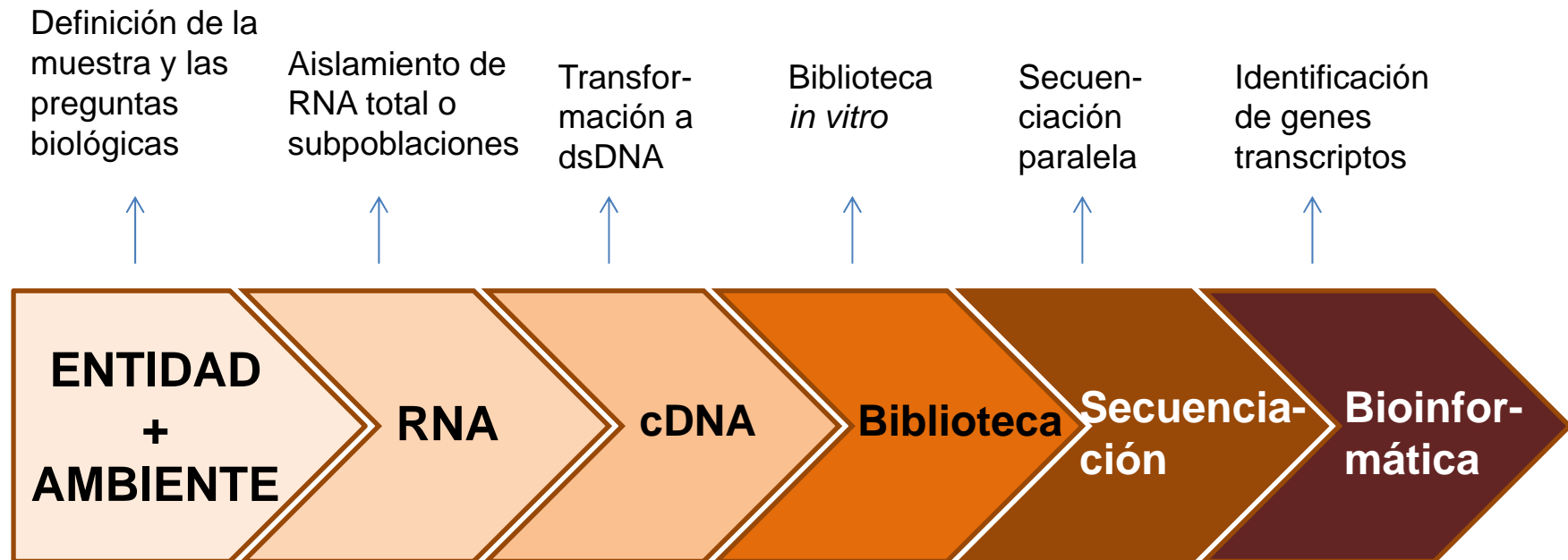
## RNA seq

- Dada la **complejidad** del **transcriptoma celular**, una mejor alternativa al estudio realizado mediante microarreglos es la **secuenciación total** del **RNA** existente en una célula.
- Esta aproximación, que es la ideal, se denomina **secuenciación de RNA** (**RNA seq**, de *RNA sequencing*) o **Whole Transcriptome Shotgun Sequencing** (WTSS).
- **RNA seq** permite **conocer** la **identidad**, la **unión de exones**, la presencia de **RNAs editados** e incluso la **cantidad relativa**, de los distintos transcriptos que se encuentran en una célula en un momento dado.
- **RNA seq** utiliza las **tecnologías de secuenciación** de segunda y tercera generación (**NGS**).

# Estudios de transcriptos

## RNA seq

Flujo de trabajo para el análisis transcriptómico



# Estudios de transcriptos

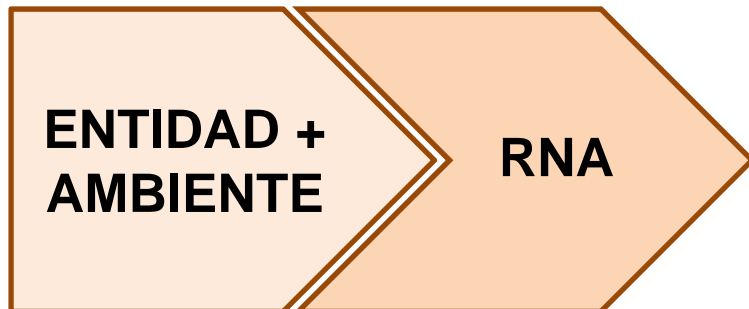
## Microarreglos

Flujo de trabajo para el análisis transcriptómico

1 color

Definición de la muestra y las preguntas biológicas

Aislamiento de RNA total o subpoblaciones



Esta **parte** es **central** en función del experimento que se desea realizar. Se puede trabajar con tan poco **RNA** como **10-100ng**.

# Estudios de transcriptos

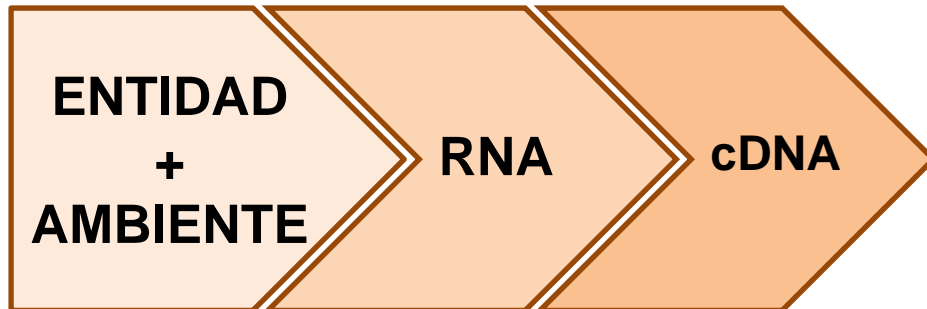
## RNA seq

Flujo de trabajo para el análisis transcriptómico

Definición de la muestra y las preguntas biológicas

Aislamiento de RNA total o subpoblaciones

Transformación a dsDNA

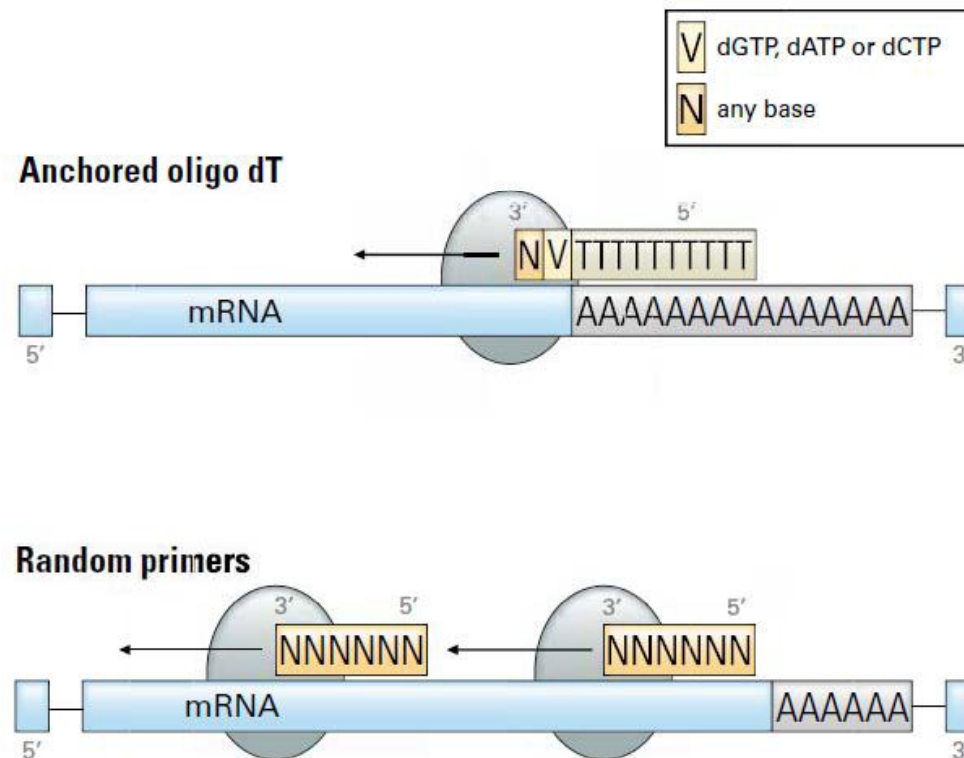


El RNA aislado es fragmentado químicamente y luego transformado en cDNA doble cadena.

# Estudios de transcriptos

## RNA seq

- El **RNA** luego de ser fragmentado es transformado a **cDNA** preferentemente con **hexámeros al azar (primera cadena de DNA)**.
- La **segunda cadena de DNA** es sintetizada con **DNA polimerasa I**.



# Estudios de transcriptos

## RNA seq

Flujo de trabajo para el análisis transcriptómico

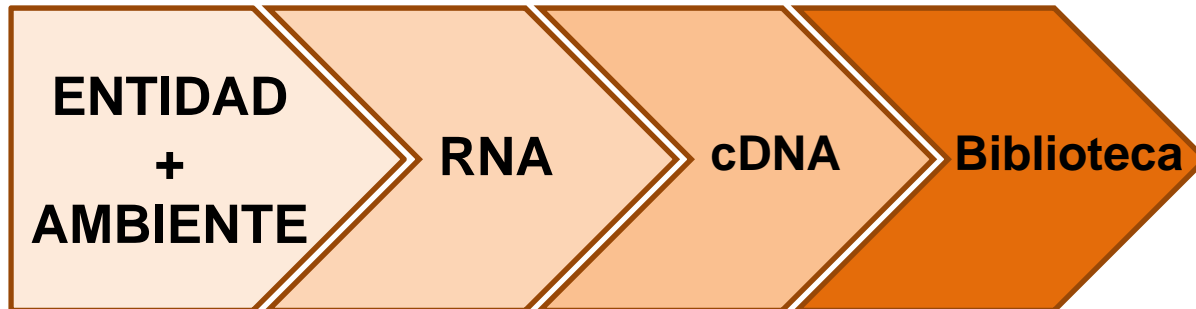
Definición de la muestra y las preguntas biológicas

Aislamiento de RNA total o subpoblaciones

Transformación a dsDNA

Biblioteca *in vitro*

En esta etapa se confecciona una **biblioteca *in vitro*** de acuerdo a la tecnología NGS seleccionada.



# Estudios de transcriptos

## RNA seq

- El **cDNA doble cadena** es **reparado** en sus extremos (DNA polimerasa I y Klenow). En algunos procedimientos, también se introducen  $\overline{A}$  en los extremos 3' para ligación  $\overline{TA}$  (Klenow)
- El **dsDNA** así obtenido es **ligado a los adaptadores** propuestos por la tecnología elegida.
- Los distintos **dsDNAs** obtenidos son **separados** por apareamiento de bases (microesferas, superficies planas).
- El **dsDNA** aislado es **amplificado** por algún procedimiento (em-PCR, *bridge-PCR*, *wildfire*) si se eligió **NGS de segunda generación**.

# Estudios de transcriptos

## RNA seq

Flujo de trabajo para el análisis transcriptómico

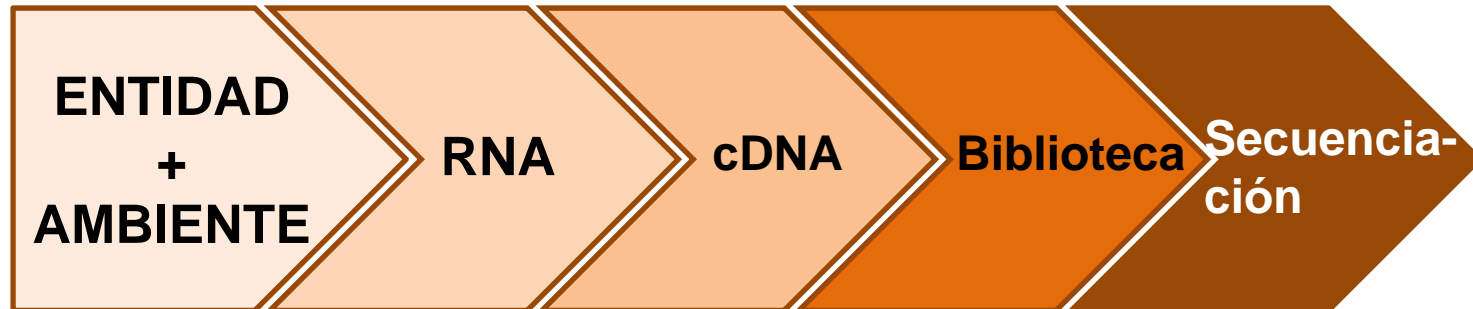
Definición de la muestra y las preguntas biológicas

Aislamiento de RNA total o subpoblaciones

Transformación a dsDNA

Biblioteca *in vitro*

Secuenciación paralela





# **Estudios de transcriptos**

## **RNA seq**

Cualquiera de las NGS descritas puede ser aplicada para determinar la secuencia del un transcriptoma

# Estudios de transcriptos

## RNA seq

| Technology   | Tiling microarray       | cDNA or EST sequencing      | RNA-Seq                    |
|--|-------------------------|-----------------------------|----------------------------|
| <b>Technology specifications</b>                           |                         |                             |                            |
| Principle  | Hybridization           | Sanger sequencing           | High-throughput sequencing |
| Resolution   | From several to 100 bp  | Single base                 | Single base                |
| Throughput   | High                    | Low                         | High                       |
| Reliance on genomic sequence                               | Yes                     | No                          | In some cases              |
| Background noise   | High                    | Low                         | Low                        |
| <b>Application</b>   |                         |                             |                            |
| Simultaneously map transcribed regions and gene expression | Yes                     | Limited for gene expression | Yes                        |
| Dynamic range to quantify gene expression level            | Up to a few-hundredfold | Not practical               | >8,000-fold                |
| Ability to distinguish different isoforms                  | Limited                 | Yes                         | Yes                        |
| Ability to distinguish allelic expression                  | Limited                 | Yes                         | Yes                        |
| <b>Practical issues</b>                                    |                         |                             |                            |
| Required amount of RNA                                     | High                    | High                        | Low                        |
| Cost for mapping transcriptomes of large genomes           | High                    | High                        | Relatively low             |