Introduction to Bioinformatics

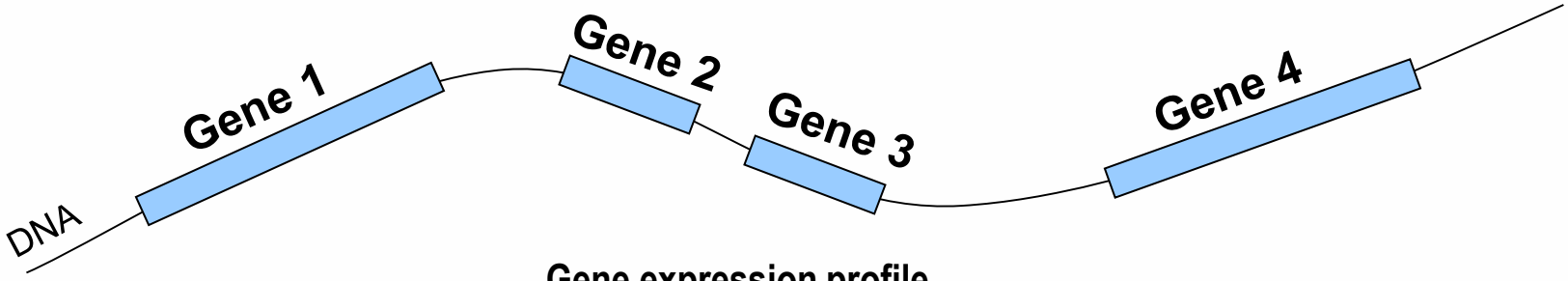
Microarray analysis

part 14

By: Mahdi Vasighi

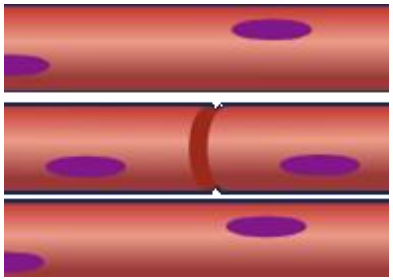
Analysis of genes

- Genes can be “turned on” or “off”. The process consisting in activate genes and suppress others is *gene regulation*.
- Gene regulation determines
  - The appearance and different function of different cells types
  - The ability of some cells to react quickly to environmental changes



Gene expression profile

Muscle cells



Gene 1 = off
Gene 2 = on
Gene 3 = off
Gene 4 = on
…

Gene 1 = on
Gene 2 = off
Gene 3 = off
Gene 4 = off
…

Pancreatic cells

Analysis of genes

How genes are regulated?

- Gene regulation can occur at any point in the process of expression but **often** occurs **during transcription**.
- Environmental signs or other cells activate proteins called **transcription factor**.
- They bind to the **regulatory regions** of genes, increasing or decreasing the level of transcription
- They control the amount of gene product produced by the gene in every moment.

Analysis of genes

Genetics:

study of single genes and their effects

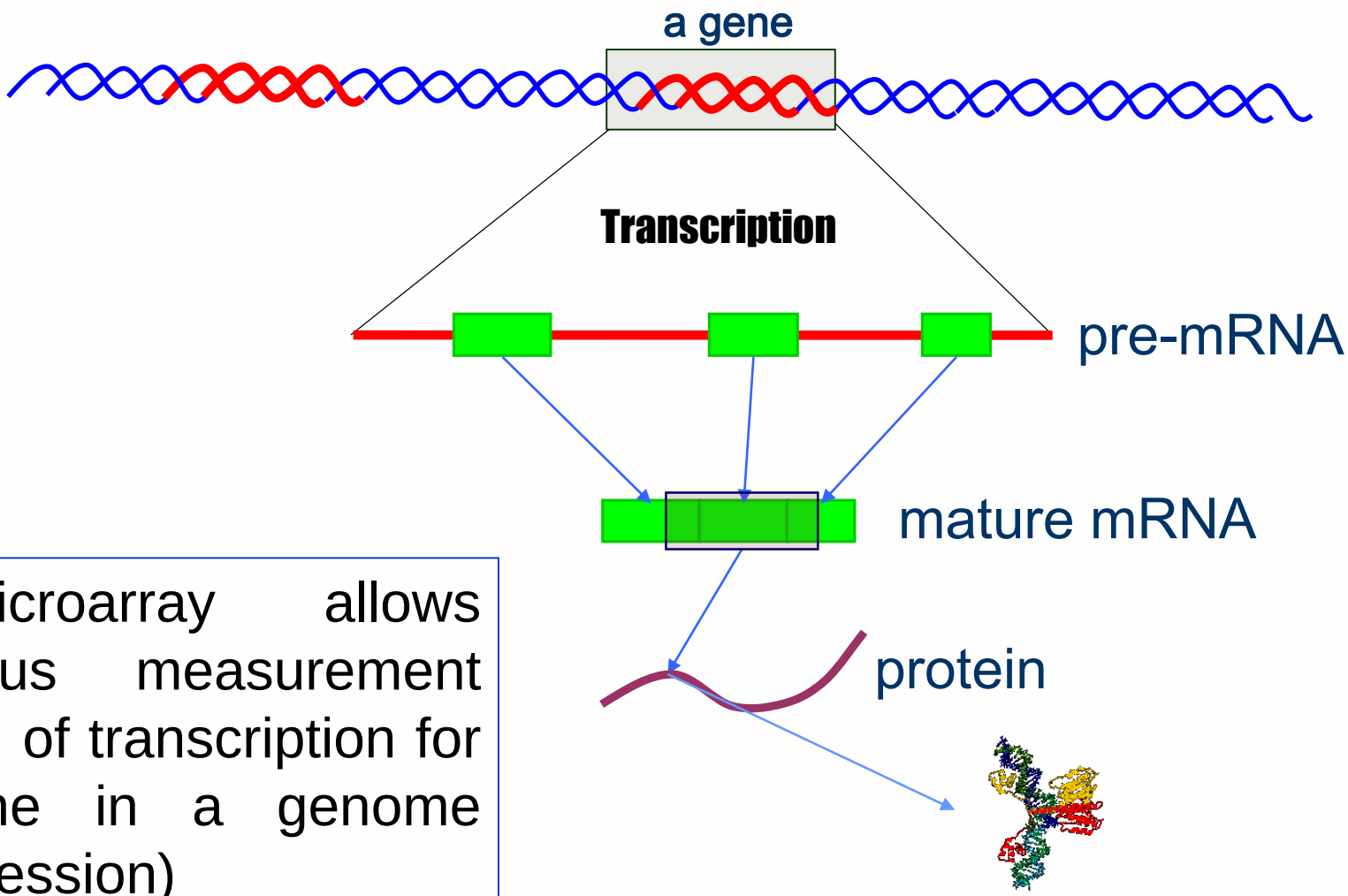
Genomics:

a way to study many genes, or even every gene in an organism and their interactions with environmental factors.

All at once!!!

Microarray technology

Analysis of genes



DNA microarray allows simultaneous measurement of the level of transcription for every gene in a genome (gene expression)

DNA Microarray

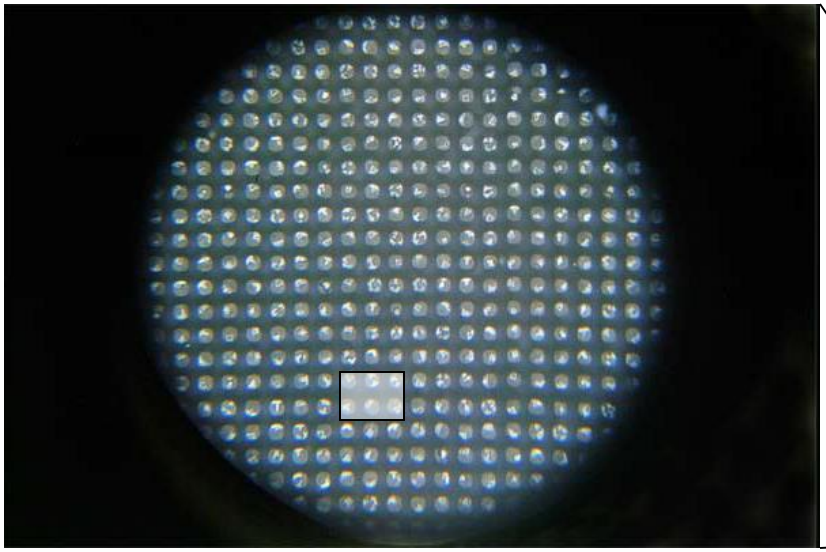
- An experimental format based on the synthesis or attachment of probes, which represent genes on a solid substrate (glass, plastic, silica ...)
- There are currently two platforms/types of DNA microarrays that are commercially available:

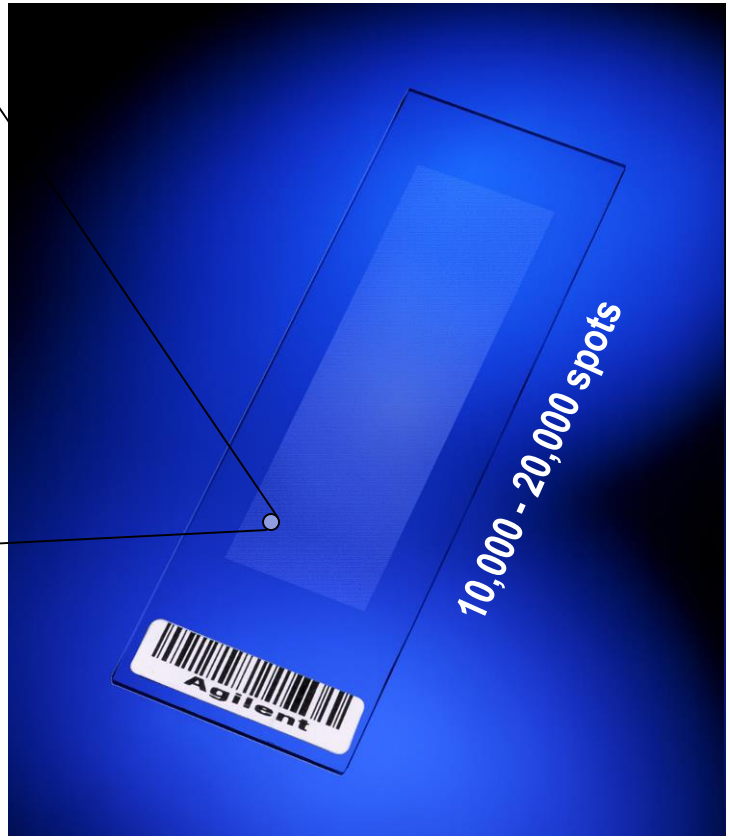
1. **Glass cDNA microarrays** which involves the micro spotting of pre-fabricated cDNA fragments on a glass slide.

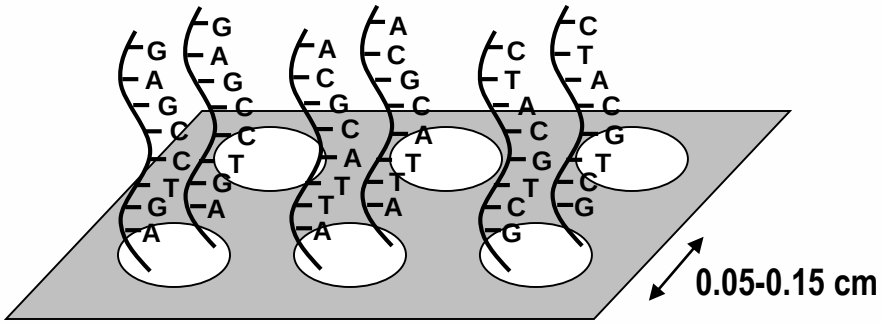
2. **High-density oligonucleotide microarrays** (**Affymetrix**) often referred to as a "chip" which involves in situ oligonucleotide synthesis.

DNA Microarray

cDNA microarrays





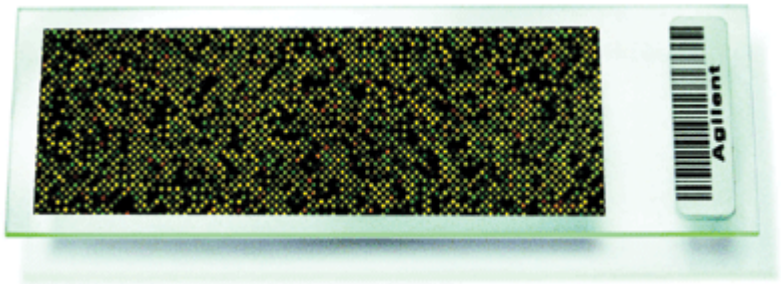


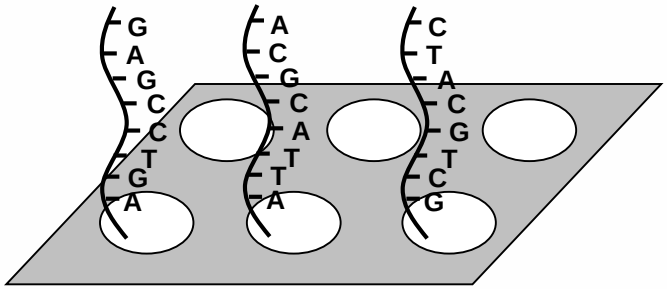
glass slide / polymer membrane

Each spot represents one gene

DNA Microarray

How does it work?





The level of hybridization between specific probes and target molecules is generally indicated by means of fluorescence:

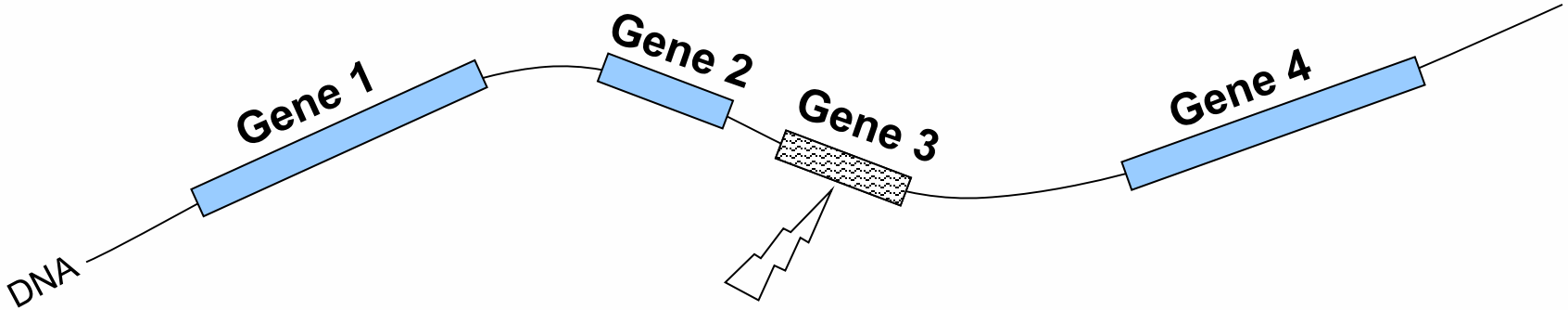
- Fluorescence intensity is measured by image analysis.
- The measured intensity indicates the level of expression of the gene corresponding to the probe in the test sample

DNA Microarray

Two color microarrays example

Now we are going to use the power of genomics to answer a very important question:

What's the difference between a healthy cell and a cancer cell?

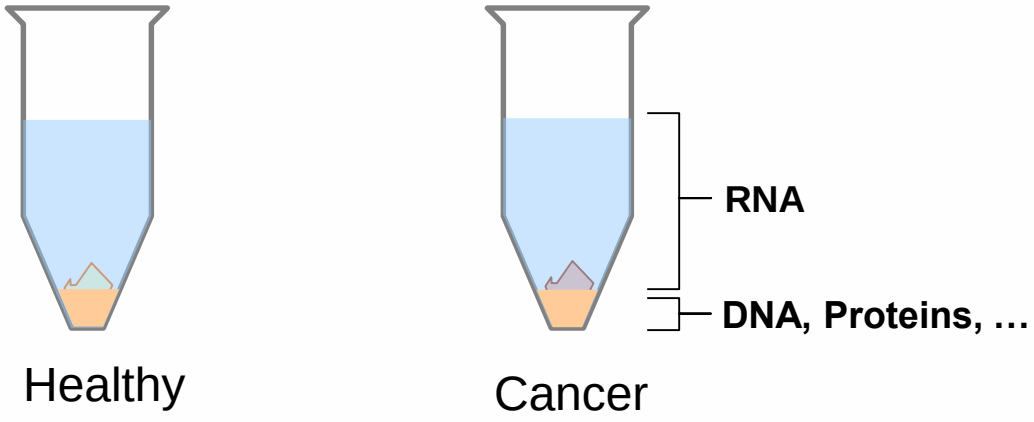


Cancer is basically a disease of "genes gone bad"

DNA Microarray

Two color microarrays example

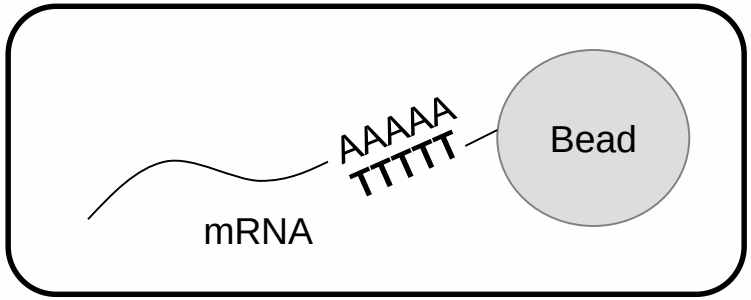
1. Collecting healthy and cancer tissue
2. Adding some organic solvents to extract RNA content
3. Dissolving tissue using shaker an RNA will be released
4. The samples has been placed in the centrifuge.

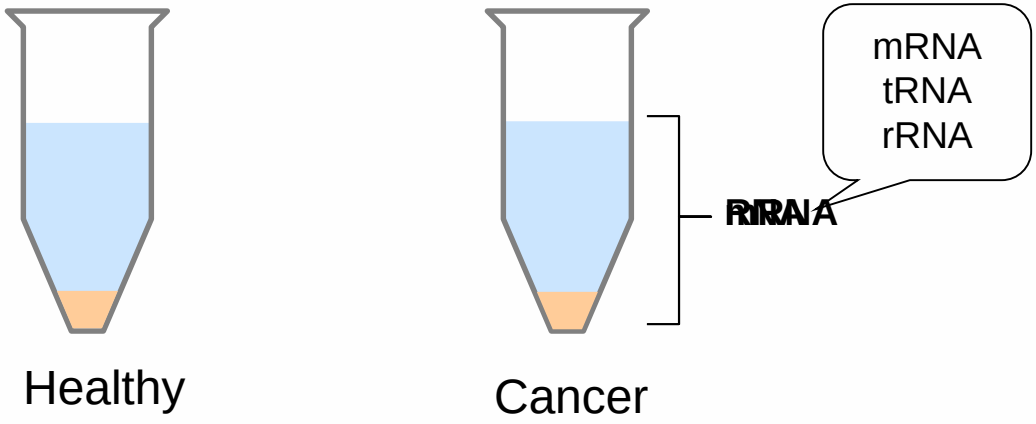


DNA Microarray

Two color microarrays example

5. Isolating m-RNA using columns with poly-T-tail beads

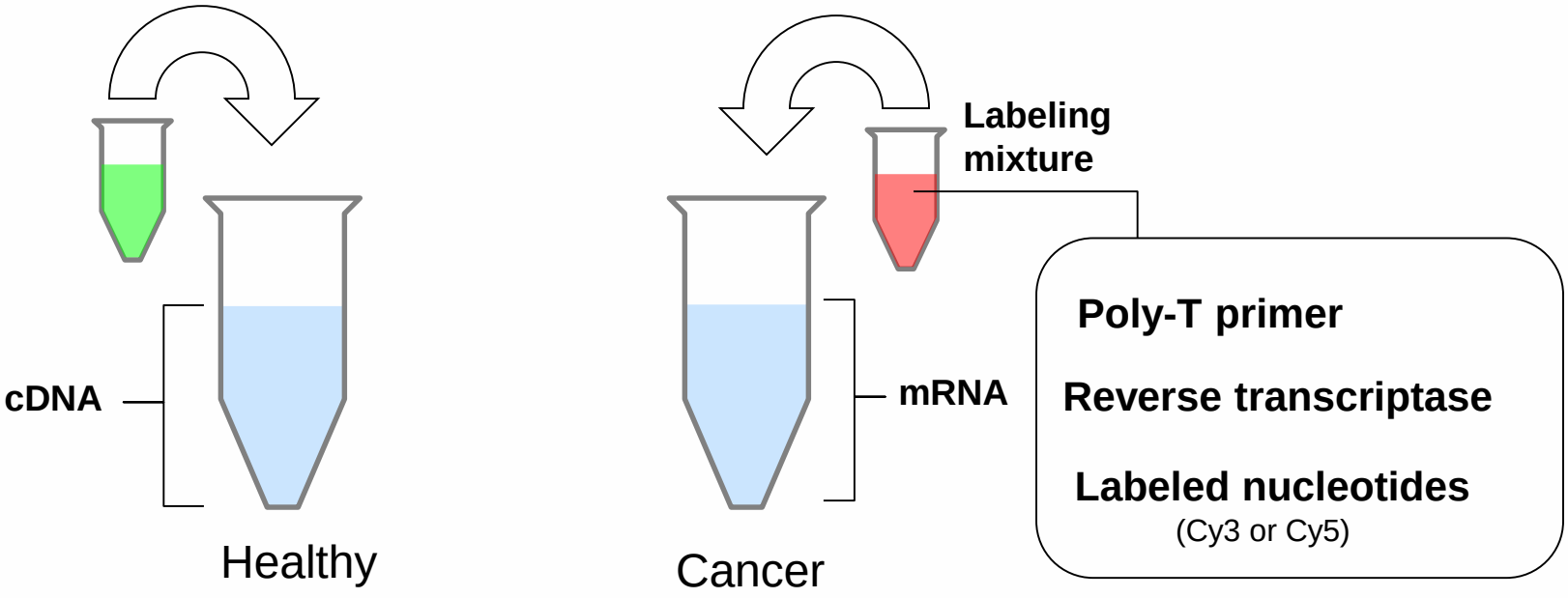




DNA Microarray

Two color microarrays example

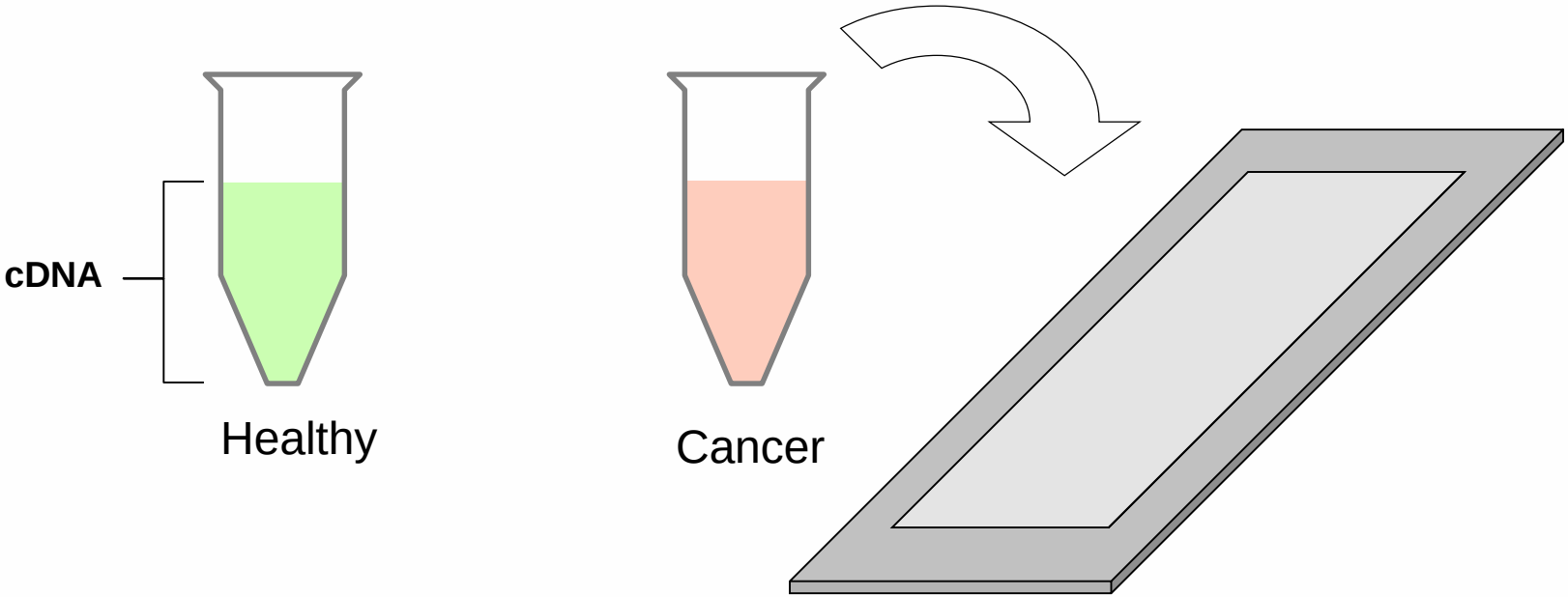
5. Isolating m-RNA using columns with poly-T-tail beads
6. Using labeling mixture to make labeled DNA copy (cDNA)



DNA Microarray

Two color microarrays example

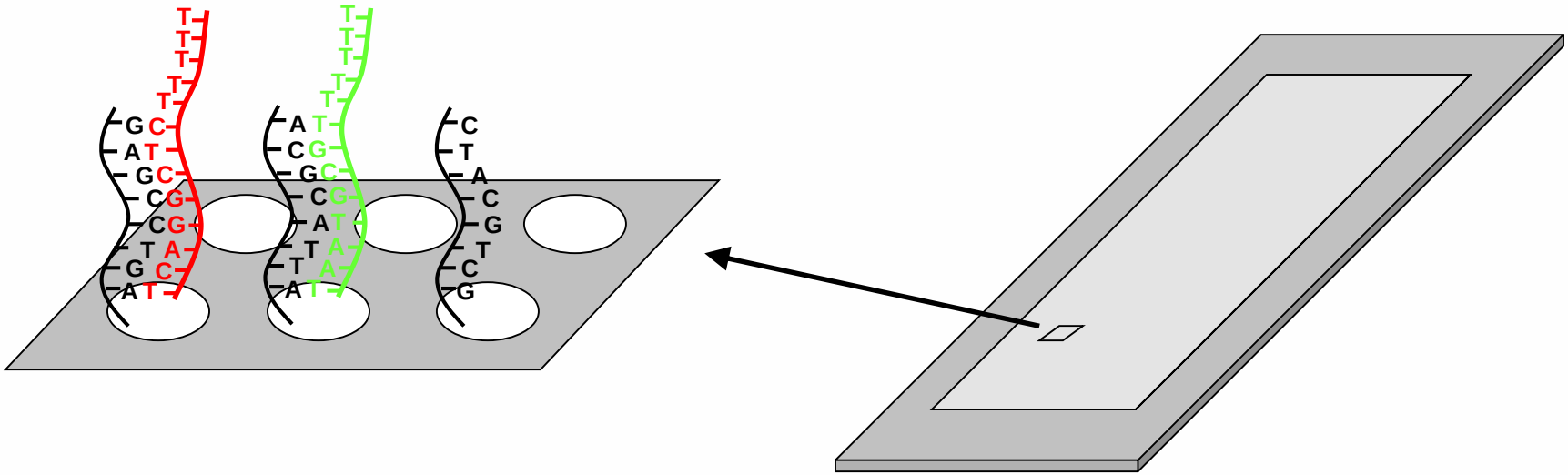
5. Isolating m-RNA using columns with poly-T-tail beads
6. Using labeling mixture to make labeled DNA copy (cDNA)
7. Apply cDNA on Microarray



DNA Microarray

Two color microarrays example

5. Isolating m-RNA using columns with poly-T-tail beads
6. Using labeling mixture to make labeled DNA copy (cDNA)
7. Apply cDNA on Microarray
8. Washing extra cDNA that didn't bind to the slide



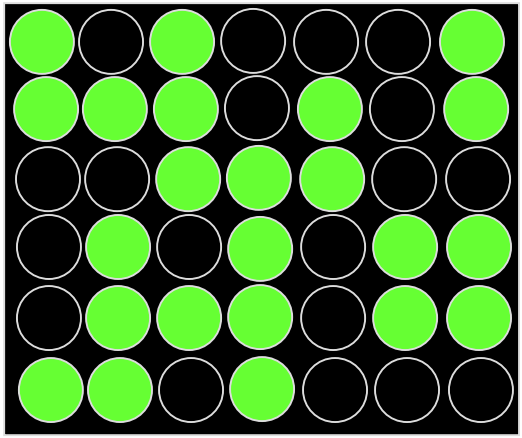
DNA Microarray

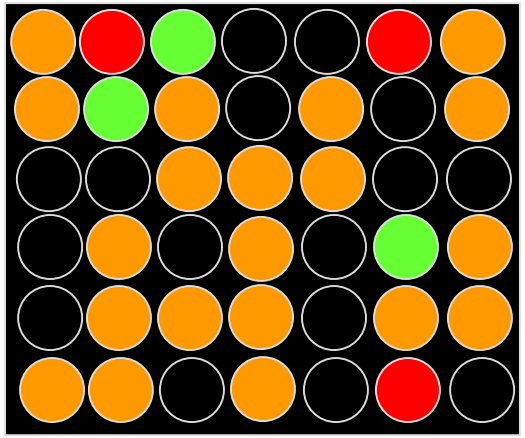
Two color microarrays example

9. The final step of the laboratory process is to produce an image of the surface of the hybridized array.

once with a green laser (for Cy3; excitation wavelength is 550 nm and emission wavelength is 581 nm) and once with a red laser (for Cy5; excitation wavelength is 649 nm and emission wavelength is 670 nm).

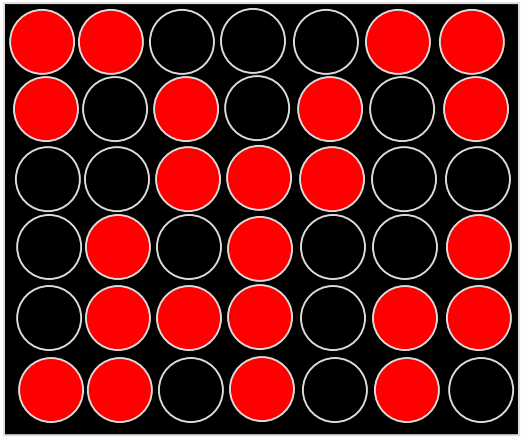
Scan Green





Merged

Scan Red

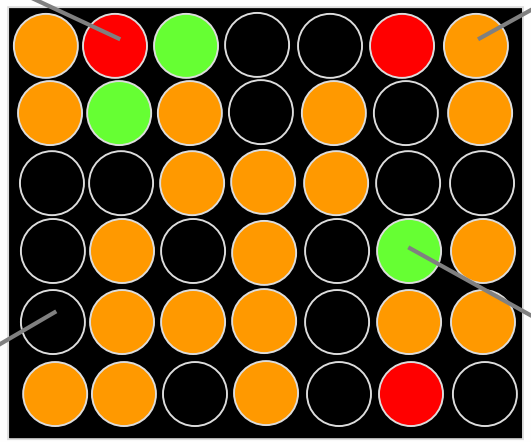


DNA Microarray

Two color microarrays example

A red spot contains a gene that produced more mRNA in the cancer than the healthy cells.

A yellow spot contains a gene that hybridized to both green and red cDNA, which means that the gene is active in both cancer and healthy cells

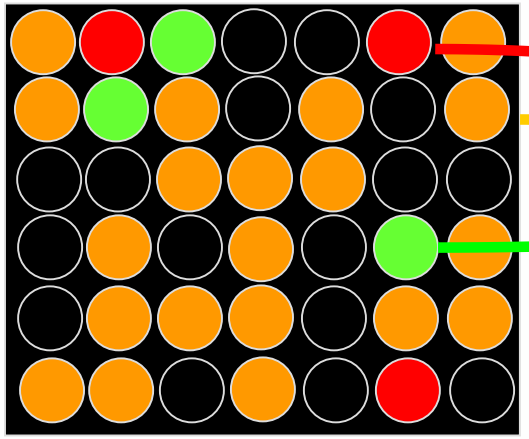


A black spot contains a gene which is not active in healthy and cancer cells

A green spot contains a gene which the expression is turn down in the cancer cells.

DNA Microarray

Two color microarrays example

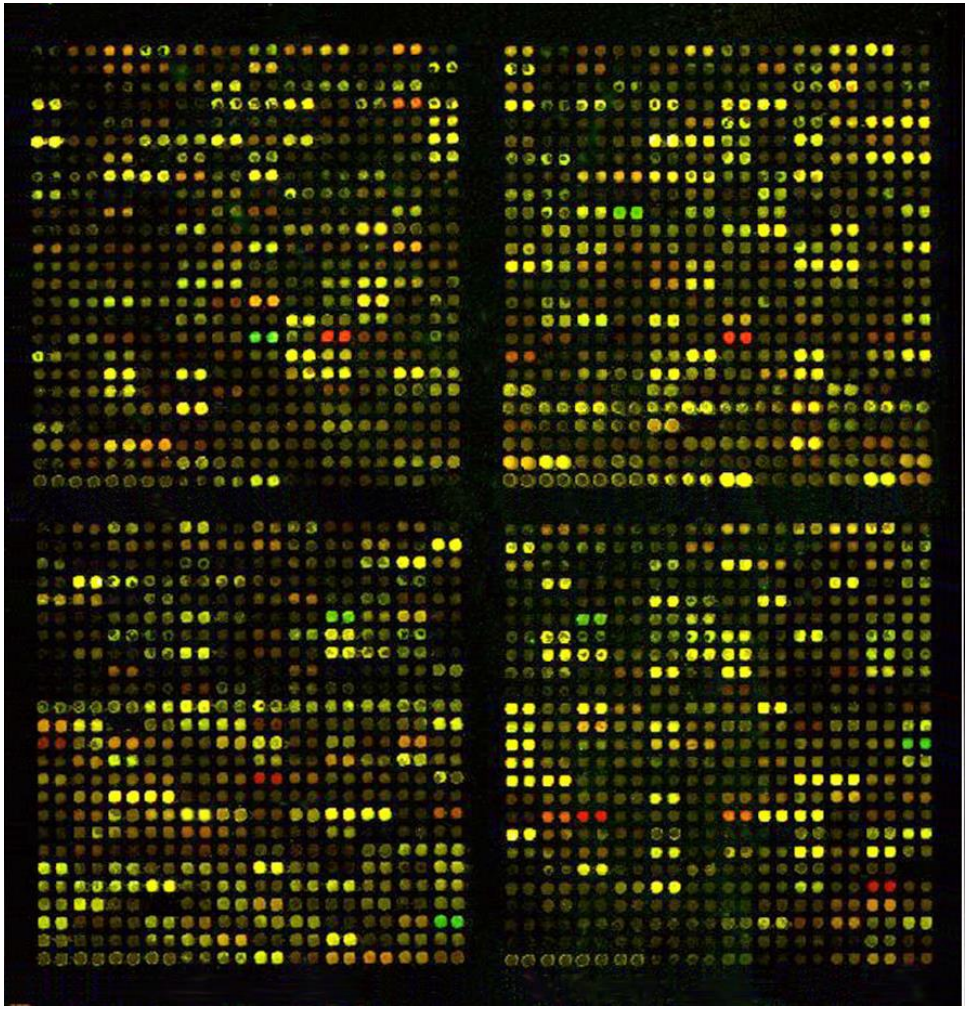


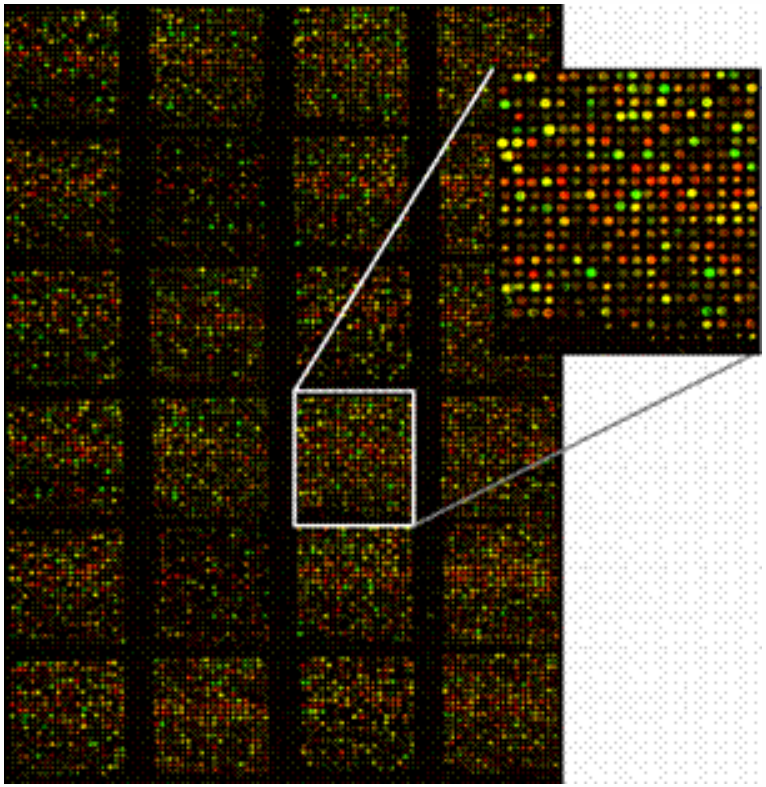
Cy3	Cy5	$\frac{Cy5}{Cy3}$	$\log_2\left(\frac{Cy5}{Cy3}\right)$
200	10000	50.00	5.64
4800	4800	1.00	0.00
9000	300	0.03	-4.91

$$T_i = \frac{R_i}{G_i} \quad \text{or} \quad \log ratio_i = \log_2\left(\frac{R_i}{G_i}\right)$$

DNA Microarray

Two color microarrays example





DNA Microarray

cDNA microarrays

Advantages of cDNA microarrays:

- lower cost
- hybridisation does not need specialised equipment.
- Data capture can be carried out using equipment that is very often already available in the laboratory.
- High detection sensitivity due to longer target sequences (2 kbp).

Disadvantages of cDNA microarrays:

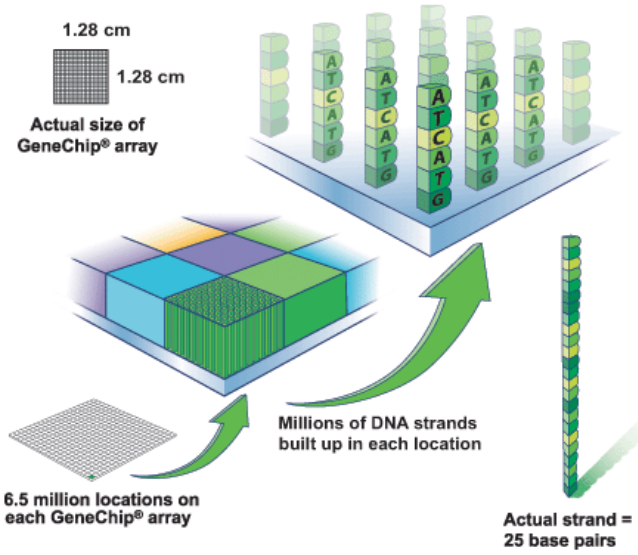
- Intensive labour requirement before microarray fabrication.
- Printing devices required thus making microarrays more expensive.
- Cross-hybridization.

DNA Microarray

Affymetrix microarrays

Hundreds of thousands of oligonucleotide probes packed at extremely high densities in affymetrix microarray. The probes designed to maximize performance of analysis.

Each probe designed to be perfectly match (PM) to a target sequence, a partner probe is generated that is identical except for a single base mismatch (MM) in its centre.

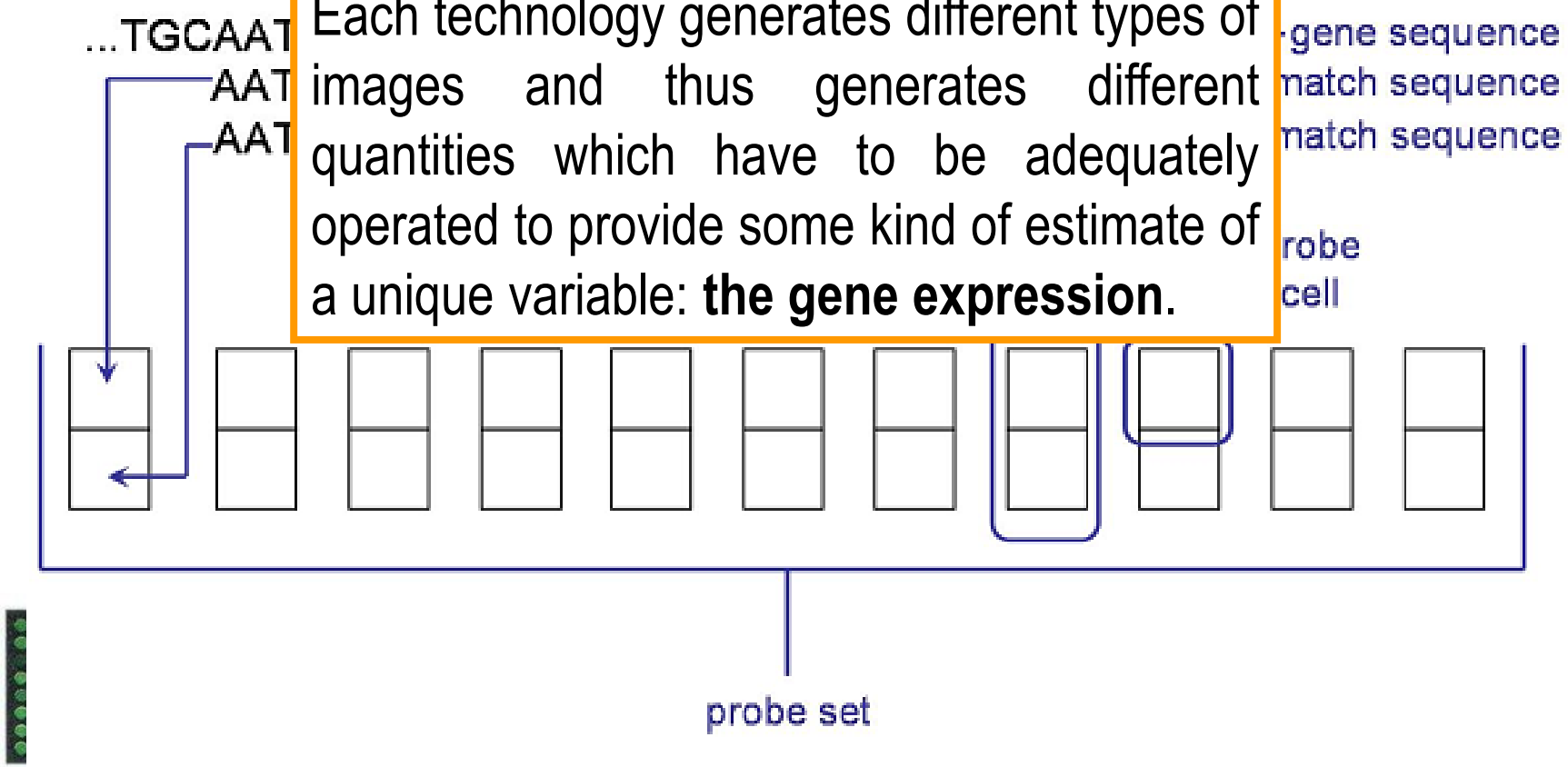


DNA Microarray

Spotted or **cDNA microarrays**

Oligonucleotide chips or Affymetrix

Each technology generates different types of images and thus generates different quantities which have to be adequately operated to provide some kind of estimate of a unique variable: **the gene expression**.



DNA Microarray

Spotted or cDNA microarrays

several quantities are produced for each spot. Basically they consist on

(i) **signal** measures, Red (R) or Green (G), for each channel.

(ii) **background** measures, R_b , G_b, intended to measure fluorescence not due to hybridization

expression ratio $$M = \frac{R}{G}$$

background-corrected expression ratio $$M = \frac{R - R_b}{G - G_b}$$

this representation may be unhelpful when one has to represent up-regulation and down-regulation.?

It is very common to use the base 2 logarithm of this quantity as the final outcome of relative expression.

Any disadvantages of using expression ratios or transformations of the ratios for data analysis?

DNA Microarray

$$Avg.diff = \frac{1}{|A|}\sum_{j \in A}(PM_j - MM_j)$$

Oligonucleotide chips or Affymetrix

Each probe is a probe pair made of a perfect match (PM) probe that corresponds to the original DNA chain and a mismatch (MM) probe whose central nucleotide has been changed.

Hybridizes with the mismatch probe should not represent "real expression". Affymetrix suggested to combine both measures in a background corrected expression measure.

Gene Expression Omnibus http://www.ncbi.nlm.nih.gov/geo

DNA Microarray

Potentials

Differing expression of genes over time, between tissues, and disease states.

Drug discovery and toxicology studies

Identification of complex genetic diseases

Pitfalls

Detect transcription mRNA level, not translation protein level

Many factors (variations) can affect the result

DNA Microarray

The microarray data analysis process