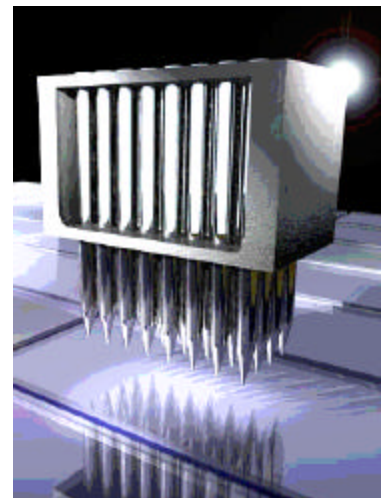**Lecture 2: Microarrays**

Central dogma of molecular biology: The state of a cell at any given time is governed by which of its genes are expressed at that time.

- Transcription, in which expressed DNA sequences are transcribed into mRNA.

- What mRNAs are present in the cell and in what quantities => inferences regarding the state of the cell.

- *Transcriptome:* The complete collection of the organism's mRNAs .

- Why not study the proteins?

    - The function of a protein is determined not just by its amino acid sequence, but also the specific structure it folds up into.

    - Proteins are difficult to purify, harder to study with current technology.

- *DNA microarrays*: technology for studying mRNA levels.

    - Monitoring the expression patterns of large numbers of genes at once,

    - Improvement over the tedious "one gene per experiment" paradigm.

    - Tiny glass slide on which genes (purified single-stranded cDNA sequences in solution) have been robotically printed in an (approximately) rectangular array; each spot on the array corresponds to a single gene.

Typical DNA microarray experiment:

1. Take a small glass slide (*chip*).

2. The surface of the chip has been divided into series of square cells to form a rectangular grid.

3. Onto each square cell, stick a tiny amount of liquid which contains DNA corresponding to a gene of known sequence.

4. Different cells will have different genes. Separately, prepare a solution that contains a mixture of mRNAs whose sequences are unknown.

5. Add to this solution a substance that fluoresces when excited by light.

cellular contents   ?   mRNA
              (isolate & purify)

   ?   cDNA   ?   sample
   (reverse   (add flourescent
   transcription)   dye)

6. Pour the solution onto the chip. The mRNA molecules will diffuse over the chip and, wherever they find a matching (*i.e.*, complementary) DNA sequence, *e.g.*, one taken from the gene from which the mRNA was transcribed, they will stick. Without a match, the solution will not stick to the chip.

7. Wash away any solution not sticking to the chip.

8. Use a laser scanner to detect and measure the fluorescent signal being emitted at each cell.

- Comparative microarray experiments: different chips containing the same set of genes will be exposed to different mRNA samples.

- By comparing the intensity levels of the fluorescent signals across the multiple mRNA samples, a scientist will be able to understand how the expression profile of a set of genes differs across the different mRNA samples.
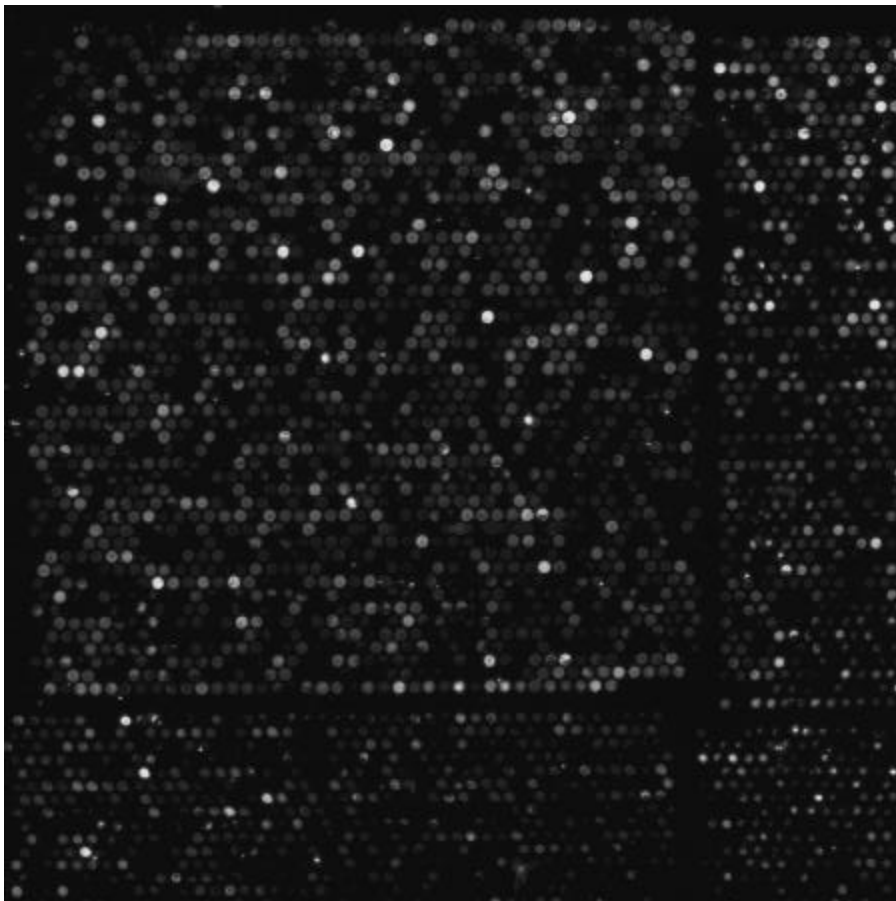
Examples of microarray experiments

1. Tissue-specific gene expression

Cells from different tissues perform different functions.

Cell's functions are determined by which proteins are produced by the cell

Depends on which genes are expressed by the cell.

Microarrays: identify which genes are preferentially expressed in which tissues.

2. Developmental genetics

- Different genes express at different stages of its developmental process.

- It has been found that there is a subset of genes involved in early development that

are used and reused at different stages in the development of the organism, in different

order in different tissues, with each tissue having its own combination.

- Growth factors that can later be involved in cancer.

- Microarrays are used to track the changes in gene expression profile

- Transforming growth factor-b (TGFb) pathways in Drosophila and C. elegans.

3. Genetic diseases:

- Result of mutations in a gene or a set of genes: *mutant gene*.

- Genes express inappropriately or do not express at all.

- Cancer p53 tumor suppressor gene inactivated

 - Identify which genes are differentially expressed in diseased Vs normal cells.

- Development of drugs that are specific to the difference between diseased Vs normal

4. Complex diseases

- Combination of polymorphisms predisposing an individual to a serious problem.

- The risk of such an individual contracting a complex disease tends to be amplified by non-genetic factors such as environmental influences, diet and lifestyle.

- CAD, multiple sclerosis, diabetes and schizophrenia have a genetic component

- Identify the genetic markers that may predispose an individual to a complex disease.

5. Pharmacological agents

- Gene expression patterns maybe altered by external stimulus: pharmacological agent

- Identify those genes which express differently in response to such exposure.

- The simplest such experiment is one in which a sample of cells is exposed to the pharmacological agent and permitted to reach a steady state of transcription. The mRNA levels in the treated cells can then be compared to those in a control sample.

- *Temporal study*. to monitor the gradual change in gene expression profiles

- *Toxicity*: changes in gene expression.

    Class of toxic agents : *signature*

    Make a prediction regarding the potential toxicity of a new chemical.

    If successful, this procedure could potentially reduce animal testing.

6. Plant breeding Some goals:

- Induce greater tolerance for environmental stressors - extreme weather conditions

- Boost the resistance of plants to infections

- Reduce insect predation, to maximize the productivity of plants

- Improve the quailty of plant products, to

- Increase the nutrition level of foods processed from plants, and to

- Develop characteristics of plant products that are valued by consumers

- Identify the genes responsible for various traits.

7. Environmental monitoring

- Disruptions of pathways by genes expressing differently affect the organism's health.

- Environmental stressors: contamination of air, food and water.

- Environmental changes that constitute hazards to health.

- Detection of pathogens in food and water.

    Microarray containing the DNA of a number of different pathogens DNA would be

    extracted from an environmental sample

    Scientists can assess whether or not there is a hazard to health.

A very simple hypothetical microarray experiment

- Cancerous liver tissue and some normal liver tissue from a liver cancer patient.

- Which genes are expressed differently in the two.

    - extracting mRNA from each tissue

    - mRNA corresponding to any genes that were expressed would be present.

- Reverse transcribe the mRNA to cDNA, add fluorescent dye to each sample.

- These two labeled samples are called *probes* because they are used to probe the collection of spots on a microarray.

- DNA microarray containing the entire human genome

- Two such microarrays are prepared.

- Flood each microarrays with one probe probe.

- Allow enough time for any cDNA in the samples to hybridize

- Wash off any excess probe from the microarrays and dry them.

- Scan the microarrays and measure the intensity level of fluorescence at each spot.

- Which genes are differentially expressed in cancerous tissue versus normal tissue.
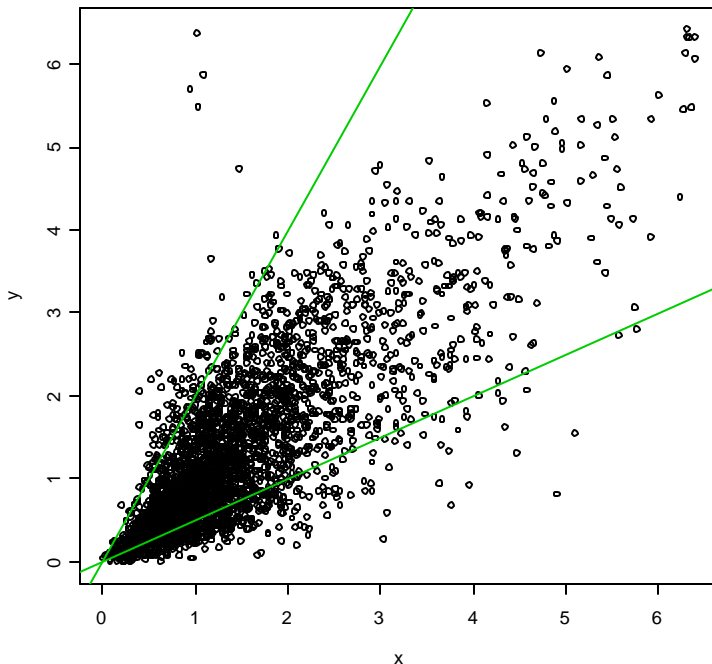
X_g and Y_g intensities measured for the gth gene in the normal tissue and cancerous

$R_g = Y_g / X_g$.

Upregulated   Downregulated

The five steps are:

   (1)preparing the microarray

   (2)preparing the probe

   (3)hybridizing the probe to the microarray and washing the microarray

   (4)scanning the microarray

   (5)interpreting the scanned image.

Microarray preparation

   arrayer. This process is called *arraying* or *spotting*.

   produce a regular grid of thousands of spots.

   The DNA in the spots is bonded to the glass

            cDNA microarray,

            oligonucleotide microarray.

      ESTs are commonly arrayed.

The DNA spotted on oligonucleotide microarrays are synthesized chains of

oligonucleotides corresponding to part of a known gene or putative ORF;

   Each oligonucleotide is usually only about 25 base pairs long.

   A gene is represented by several different oligonucleotides;

   The oligonucleotides are carefully chosen for maximal specificity.

The selection of DNA targets to be spotted on the microarray determines which genes can be studied in the experiment in which it is used.

Robotic technologies for arraying microarrays.

- Robotic arm to touch and spot nanoscale droplets of the solution containing the cDNA or oligoneucleotide.

- Ink jet technology to eject the solution onto the surface of the glass slide without the robot actually touching it.

- Other technologies concurrently synthesize oligoneucleotides on the slide in situ, using either photolithography (Affymetrix) or ink jet technology (Rosetta Inpharmatics).

Probe preparation

- mRNA is purified from total cellular contents. captured mRNA degrades very quickly.

- Immediately reverse transcribed into more stable cDNA.

- Not all mRNAs are reverse transcribed with the same efficiency.

- As this effect is gene-specific, the fluorescence intensity that is measured may not be a true reflection of mRNA level.

- Should not compare fluor. intensities for different genes across a single sample.

- Compare fluorescence intensities across several samples.

- Fluorescent dyes, called fluors or fluorophores, chemicals that fluoresce when exposed to a specific wavelength of light.

- The number of fluor molecules which label each cDNA depends on its length and also possibly its sequence composition.

The hybridization step

The probe is poured onto the microarray and allowed to diffuse uniformly all over it. Then it is sealed in a hybridization chamber for enough time for the hybridization reactions to complete.

The microarray is then removed from the hybridization chamber and washed

The microarray is dried

Scanning the microarray

Now the microarray is scanned to determine the amount of probe bound to each spot.

Interpreting the scanned image

The end product of a microarray experiment is a scanned image

Image processing software will convert the image into quantitative intensity measurements, which will then be analyzed for gene expression differences.

Multi-channel cDNA microarrays

- Probe out of two or more mRNA samples mixed together.

- Each mRNA sample in the mixture is labeled with a different fluorescent dye.

- The slide is scanned several times, once for each dye.

- Advantages: Difficult to control the amount of DNA that is spotted onto the slides

Could vary from array to array for the same gene.

Some economy is gained as data on expression levels of several mRNA

- Drawbacks: Dye effect

Gene-specific dye effects.

Oligonucleotide microarrays

- Affymetrix proprietary to this day

- A gene is represented by a set of 20 oligonucleotides, *perfect match probes-PM*

- Paired with an artificially created mismatch probe -*MM* changing the middle base

MM's: internal control for hybridization specificity.

- Each PM-MM pair is referred to as a *probe pair* --- entire set= *probe set*. The Series

of multiple independent detectors- greatly improving measurement.

- Manufactured by synthesizing the oligonucleotides directly onto a silicon chip.

- Very a large number of oligonucleotides can be arrayed at the same time.

Bead based arrays

- A bead-based fiber-optic microarray is a bundle of optical fibers.

- Microscopic wells are etched onto the end of each fiber.

- These wells hold the target DNA sequences in bead form.

- The array is exposed to the fluorescently labeled probe.

- Wherever the probe finds a matching (*i.e.*, complementary) DNA sequence on the

microarray, hybridization takes place. Without a match, the probe does not hybridize to

the target. The array is illuminated with a lamp. This triggers fluorescence in the tagged samples, which causes a signal to be passed through the optical fiber to a detector, which indicates which target DNA sequences match some sequence in the probe.

- The throughput of three-dimensional bead-based microarrays is a great deal higher than conventional two-dimensional microarrays. In fact, the number of DNA sequences tested could be in the hundreds of thousands, or even, millions range.

Supplementary reading

The Chipping Forecast, a special supplementary issue of the journal *Nature Genetics*, Online at http://www.nature.com/ng/chips_interstitial.html.