

DNA Chips: Genes to Disease

Using Microarrays to Study Genes Involved in Lung Cancer

STUDENT GUIDE

Background

Microarray analysis is a powerful new research tool that enables technicians to view and interpret at one time, on one small surface, the extent to which thousands of genes have been expressed in cells. Researchers developed and continue to refine the technology by merging strides in genomics, computer science, and nanotechnology.

Detecting patterns or changes in transcription in cells is a way to understand both normal and abnormal aspects of cell function. A researcher who wanted to look for changes in transcription in a specific cancer tissue could use microarray analysis. As the first step in this process, a gene chip would be created. DNA chip, microarray, gene chip, and genome chip are all terms that describe a solid matrix, such as a glass slide, that is imprinted with a precisely arranged pattern of spots, each made up of many copies of a specific oligonucleotide representing part of a genome (e.g., a human genome).

As the next step, the DNA chip would be used to analyze complementary DNAs (cDNAs) that were made from mRNA isolated from cancerous and noncancerous parts of the same tissue. The cancerous and noncancerous DNA samples are flagged with dyes and applied to the prepared chip. The extent to which each flagged gene adheres to its complement on the chip directly indicates the extent to which transcription occurred. Computer analysis of the DNA chip reveals which genes were transcribed in the cancerous tissue and which in the normal tissue, and thus indicates which genes might be important in the development of the cancer. The use of a microarray in this application allows suspect genes to be identified years sooner that would have been possible with previous technologies that were unable to analyze so many genes so precisely at one time.

Gene Expression = Transcription into RNA and Translation into Protein



Induced (Expressed) Gene:

Gene X $\xrightarrow{\text{Transcription}}$ Lots of mRNA X

Repressed (not Expressed) Gene:

Gene Z $\xrightarrow{\text{X}}$ no mRNA Z

Gene Expression and Cancer

A single microarray can contain more than 30,000 spots of DNA, each representing a different gene in an organism. In this laboratory, you will use a DNA microarray ("gene chip") to study the expression of six different genes in normal lung cells and lung cancer cells. These results will show what genes have been transcribed (expressed) to produce more messenger RNA in lung cancer cells than in normal lung cells and what genes have been repressed from producing mRNA in the lung cancer cells (i.e., transcription is reduced).

Scientists have found that some genes are not transcribed as much in cancer cells as in normal cells. These repressed genes may play an important role in allowing the cancer cells to spread and grow. Other genes are transcribed more in cancer cells than normal cells. These genes may also play an important role in making the cells cancerous. There are also many genes that are transcribed at the same level in both cancer cells and normal cells. These genes probably do not play a significant role in causing cells to become cancerous. There are also some genes that may not be expressed at all in normal or cancerous lung cells. Can you think of any examples of these?

Additional Background

Dr. Malcolm Campbell's yeast microarray animation
<http://www.bio.davidson.edu/courses/genomics/chip/chip.html>

Longer, interactive DNA microarray animation
<http://gcat.davidson.edu/Pirelli/index.htm>

The Genome consortium for Active Teaching
www.bio.davidson.edu/GCAT

DNA Chips: From Genes to Disease at GCAT
<http://www.bio.davidson.edu/projects/GCAT/HsChips/Hschips.html>

Realistic DNA Chip Animation
<http://gslc.genetics.utah.edu/units/biotech/microarray/>

Mrs. Kathleen Gabric's Microarray Background Introduction
<http://www.hinsdale86.org/staff/kgabric/labsOnline/Microarrayer2.doc>

Data Analysis Web page for DNA Chips: From Genes to Disease
(Quantifying Gene Chip Colors-math exercise)
http://www.bio.davidson.edu/projects/GCAT/HsChips/hs_kit_math_module_v2.pdf

HHMI Microarray Resources (BioInteractive)
www.hhmi.org/biointeractive/genomics/genechipdata/index.html

Using Microarrays to Study Genes Involved in Cancer: A Paper Microarray Exercise

Objective

To offer students an interactive way to visualize how microarrays are used to study gene expression.

Below is a brief description of how microarrays are used in research labs. Following that, is a simplified procedure for a paper microarray activity that mimics the procedure of the wet lab. In this paper activity, you will experience the main concepts of working with DNA microarrays.

General Microarray Analysis Procedure

1. Obtain a microarray slide containing 70 bp oligonucleotide (DNA) sequences, each of which represents a gene sequence in the genome of your favorite organism. Most scientists purchase these slides already prepared. The DNA is bound in spots approximately 100 microns in diameter. Each slide contains thousands of microscopic spots of DNA (each spot corresponds to a different gene sequence in the genome of the cell type or organism you are examining; humans have ~25,000 genes). While the function may be known for some of these gene sequences, many genes have unknown functions.
2. Extract mRNA from your experimental organism and your control organism for comparison. For example, corn growing under drought conditions vs. corn growing in normal conditions, or tumor cells vs. normal cells. Each sample will contain thousands of different mRNA sequences representing all of the genes expressed in those cells.
3. Prepare fluorescently labeled cDNA copies of this mRNA. Label the cDNA created from each sample of mRNA with different fluorescent nucleotides (either green Cy3 dye or red Cy5 dye). Denature the cDNA to produce single-stranded DNA prior to the next step.
4. Hybridize the microarray slide with both the green- and red-labeled cDNAs. Each cDNA will bind to the spots that have complementary sequences. Stringent conditions are used to ensure that the various cDNAs are entirely complementary to the microarray spot sequences to which they hybridize.
5. Wash the slide to remove excess fluorescent cDNAs not bound to spots.
6. Read the microarray using an instrument that measures the fluorescence of each spot at the two different wavelengths (for green Cy3 and red Cy5). Two images are created for each spot, but the instrument is connected to a computer that integrates the data into a single image. The colors of the spots are as follows:
 - green = The spot is bound to Cy3-labeled cDNA. These spots represent genes that are expressed in the tissue of cells whose mRNA was reverse transcribed into cDNA labeled with Cy3 (the green dye).
 - red = The spot is bound to Cy5-labeled cDNA. These spots represent genes that are expressed in the tissue of cells whose mRNA was reverse transcribed into cDNA labeled with Cy5 (the red dye).
 - yellow = The spot is bound to BOTH Cy3- and Cy5-labeled cDNA. These represent genes such as "housekeeping genes" that are required by all cells.
7. Analyze the data to determine which genes (represented by spots on the slide) are expressed in each cell sample and which are expressed in both. The next step in functional genomic studies is to study in more detail those genes that are differentially expressed in control vs. experimental conditions.

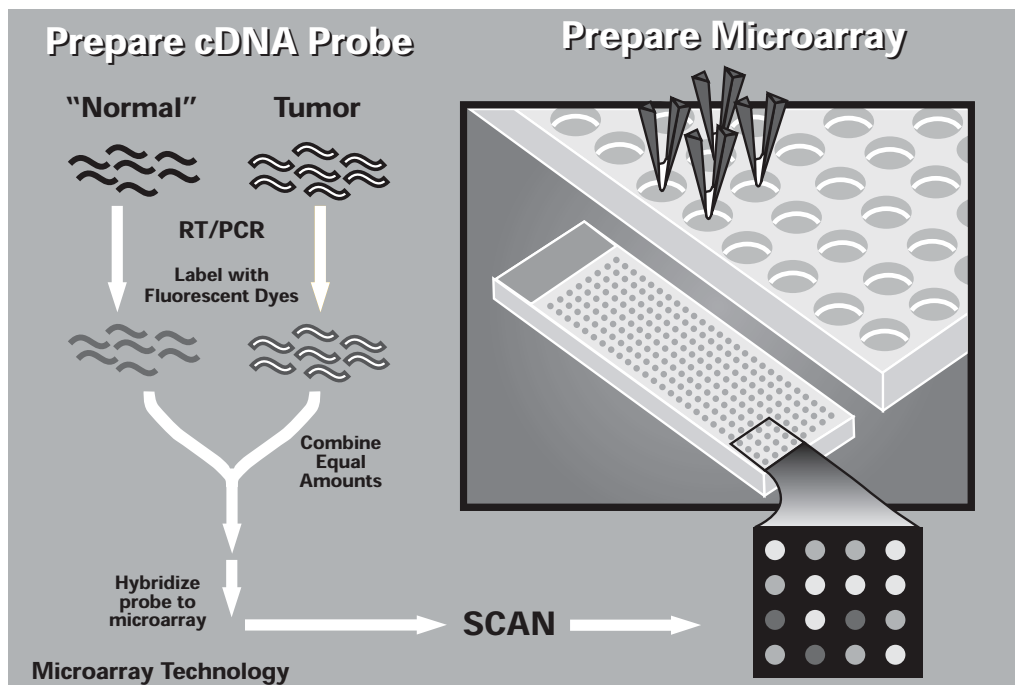


Figure 1. Using Microarray Technology to Study Gene Expression in Normal and Tumor Cells
(Daryl Leja, National Human Genome Research Institute)

Paper Microarray Analysis—Student Procedure

You are part of a research group studying human genes involved in cancer. You have assisted with the Human Genome Project and have identified genes likely to be involved in cancer by means of genetic database comparisons and other computer analyses. You are aware of some genes with similar sequences in other organisms, and you know the proteins they code for. On this basis, you can predict likely functions of the human genes you have identified. However, you find many sequences with unknown functions. Your research group believes that many of these unknown genes play a role in either the prevention of cancer in normal cells or the proliferation of cancer cells in abnormal tissues.

In these post-genomics studies (research done after the human genome was sequenced), your group has decided to use microarrays to compare gene expression in normal cells vs. abnormal, cancerous cells. Your goal is to identify the genes that are expressed differently.

1. Each research group will work with two different tissue samples (one normal, and one cancerous). First, you must extract the mRNA from each sample. A problem with mRNA is that it is very unstable. Also, the mRNA from the different samples cannot be distinguished. In order to distinguish the two sets of mRNA, the mRNA must be converted into labeled cDNA (complimentary DNA), through a process similar to the transcription of mRNA from DNA. In this process, called "reverse transcription," the cDNA is copied from the mRNA template. In preparation for hybridizing to microarrays, the cDNA from the two types of tissue is labeled with different fluorescent nucleotides (in this kit, either blue dye or red dye.)*

**Note: In a real microarray the dyes are green and red, and yellow is the color that indicates expression of both genes. In this paper lab and the simulated wet lab in this kit, blue is used rather than green, and purple will indicate expression of both genes.*

To review the process of converting mRNA to cDNA, complete the following problem in the space provided. You extracted the following mRNA sequence (among thousands of other mRNAs) from cancerous cells:

5'-CCUAUUGGAAUCGG-3'

What is the cDNA sequence that would be synthesized from this mRNA? Designate which end is 3' and which is 5'.

Note: Remember that because of the chemical nature of individual nucleotides, DNA is synthesized from the 5' to the 3' end. Also, remember that complementary DNA strands pair with each other in antiparallel fashion, such that the 5' end of one strand pairs with the 3' end of the other strand.

2. Your group has done an excellent job carefully extracting mRNA and preparing fluorescently labeled cDNA from your tissue samples! Your teacher will give you copies of these labeled, single-stranded cDNAs. The cDNA from normal cells was labeled with Cy3 (blue), and the cDNA from the cancerous cells was labeled with Cy5 (red).

3. Your research group of 4–6 scientists has obtained a microarray slide containing 6 spots of DNA oligonucleotides representing different human genes with unknown functions. How does this compare with an actual microarray slide?

4. Mix the fluorescently labeled cDNA from the two cell samples (if they aren't already mixed). Hybridize the microarray slide with these labeled cDNAs. Each cDNA will bind to the spots that have complementary sequences. The DNA oligonucleotides are shorter than the cDNA sequence, so the oligos will bind with only a portion of the cDNA sequence. However, for binding to occur, the entire sequence of the oligo must be complementary to a sequence in the cDNA. Try to find the complementary sequence for all of the cDNAs. Neatly stacking your hybridized cDNAs will keep the microarray sequence in view. You may tape the hybridized cDNA onto the microarray slide using a small piece of removable tape.
5. Wash the slide to remove excess fluorescent cDNA not bound to spots. (Simply remove the unbound cDNAs left on the slide sheet.)
6. Read the microarray using an instrument that measures the fluorescence of each spot at the two different wavelengths for blue or red. Analyze the data to determine which genes (as represented by the cDNA-bound oligos on the spots on the slide sheet) are expressed in each tissue sample and which are expressed in both. On the microarray slide sheet, use markers to color each spot. The colors of the spots will be as follows:
blue = The spot is bound to Cy3-labeled cDNA. What do these spots represent? (These are green in an actual microarray.)
red = The spot is bound to Cy5-labeled cDNA. What do these spots represent?
purple = The spot is bound to both Cy3- and Cy5-labeled cDNA. These represent genes such as "housekeeping genes" that are required by all cells. These are yellow in an actual microarray.
7. Your group has obtained interesting results that may be useful in determining how cancer cells differ from normal cells! The next step is to study those genes that appear to be important in your experimental cells (in this case, the cancer cells.)
 - a. Which unknown gene sequences (#1–6) appear to belong to genes used in all cells?

- b. Which unknown gene sequences (#1–6) might belong to cancer-preventing genes?
- c. Which unknown gene sequences (#1–6) might be from genes that cause cells to become cancerous?
- d. Are all of the genes expressed at the same level? How do you know this? What could this mean?
- e. What additional questions do you have regarding your microarray results?
- f. In the space below, describe further research that your group would like to accomplish using microarrays. Your study can be related to the cancer study that you just carried out, or it can be unique. Be sure to describe what samples you will use and what will be spotted on the microarray.

Single-stranded, Blue-labeled cDNA from normal cells

Run a BLUE line down the end of each strip (or copy the strips onto blue paper).

Cut out each cDNA.

Combine these cDNAs with one sheet of the red cDNAs from tumor cells for each student or team.

3'-ACCCATCGGAACCTTAG-5'	3'-ACCCATCGGAACCTTAG-5'
3'-ACCCATCGGAACCTTAG-5'	3'-ACCCATCGGAACCTTAG-5'
3'-TTATCCCGGAAATT-5'	3'-TTATCCCGGAAATT-5'
3'-AGCGTAAATTCCATT-5'	3'-AGCGTAAATTCCATT-5'
3'-GAGAGGTAGGAATATCAATT-5'	3'-GAGAGGTAGGAATATCAATT-5'
3'-ATAGCGGCCCGCGCG-5'	3'-ATAGCGGCCCGCGCG-5'
3'-GAGAGGTAGGAATATCAATT-5'	3'-GAGAGGTAGGAATATCAATT-5'
3'-ATATATTAAACAAGTTGCC-5'	3'-ATATATTAAACAAGTTGCC-5'

Single-stranded, Red-labeled cDNA from abnormal tumor cells

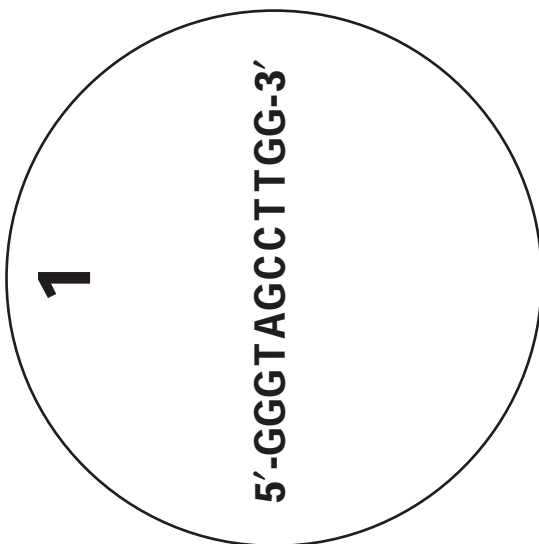
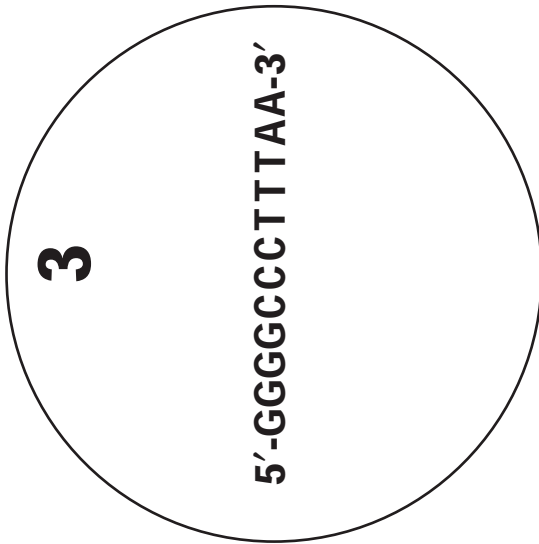
Run a RED line down the end of each strip (or copy the strips onto red paper).

Cut out each cDNA.

Combine these cDNAs with one sheet of the blue cDNA from normal cells for each student or team.

3'-CCATATATATATGGC-5'	3'-CCATATATATATGGC-5'
3'-CCATATATATATGGC-5'	3'-CCATATATATATGGC-5'
3'-CCATATATATATGGC-5'	3'-CCATATATATATGGC-5'
3'-CCCCAGTAGTAG-5'	3'-CCCCAGTAGTAG-5'
3'-TTATCCCGGAAATT-5'	3'-TTATCCCGGAAATT-5'
3'-ATACCCGATCCCC-5'	3'-ATACCCGATCCCC-5'
3'-GAGAGTAGGAATCAATTT-5'	3'-GAGAGTAGGAATCAATTT-5'
3'-CCCCAGTAGTAG-5'	3'-CCCCAGTAGTAG-5'

Microarray Slide Sheet



The six genes that you will study using a “microarray”:

Gene 1. C4BPA—Complement component 4 binding protein, alpha

The protein this gene codes for helps initiate part of our immune system to kill pathogens.

Gene 2. ODC1—Ornithine decarboxylase 1

The protein this gene codes for is an enzyme in the polyamine biosynthesis pathway. The pathway catalyzes the conversion of ornithine to putrescine. Growth-promoting stimuli can cause the activity level of the enzyme to vary.

Gene 3. FGG—Fibrinogen, gamma polypeptide

The protein encoded by this gene is a part of fibrinogen, a protein found in the blood. When blood vessels are injured, fibrinogen is cleaved to form fibrin, the most abundant component of blood clots. In addition, other pieces of fibrinogen and fibrin control how cells adhere to other tissues and cells within the body, how they spread (the kind of spreading that involves “flattening” of the individual cell), and how they move in response to chemical signals.

Gene 4. HBG1—Hemoglobin, gamma A

HBG1 is one of two γ -globulin genes (HBG1 and HBG2) normally expressed in the fetal liver, spleen, and bone marrow. The two γ -chains coded for by these genes combine with two α -chains to form the fetal hemoglobin protein. Fetal hemoglobin is usually replaced by adult hemoglobin at birth.

Gene 5. SIAT9—Sialytransferase 9

The protein encoded by this gene catalyzes the formation of another protein called GM3. Ganglioside GM3 is known to play a role in inducing cell differentiation, controlling cell growth, and in maintaining the shape of cells called fibroblasts. It also plays a role in communication pathways within the cell, and in certain types of cell adhesion. Mutation of SAIT9 has been associated with a disease called Amish infantile epilepsy syndrome.

Gene 6. CYP24 (also called CYP24A1)—Cytochrome P450, family 24, subfamily A, polypeptide 1

The cytochrome P450 proteins are enzymes called monooxygenases that catalyze many reactions involved in drug metabolism and the making of cholesterol, steroids, and other lipids. The protein coded for by the CYP24 gene controls the level of vitamin D₃ (the physiologically active form of vitamin D) and thus plays a role in regulating calcium and the vitamin D endocrine system.

Procedure

Part 1: Prepare the simulated microarray slide.

First, you will prepare your DNA microarray by spotting each of the six different gene sequences onto a glass slide. For real microarrays, scientists actually print thousands of microscopic DNA spots onto a slide, one spot for each gene they want to examine. Your spots will be much larger than those in a regular microarray, and you will be able to view them without specialized equipment.

1. Do not touch the surface of your slide (handle it only by the edges). If the clear spots on the slide are not already labeled, use the permanent marker to number six of them (1–6). Also, write your group number on the frosted labeling area of the slide.
2. Bring your labeled slide to the waterbath area.
3. **Method 1:** Using the dropper bottles, carefully spot the appropriate gene solution onto each of the labeled spots of your slide. Be sure to place the correct DNA sequence in the correct spot (i.e., Gene 6 needs to be spotted on spot 6, etc.) and to place the same amount on each spot. Once the spots are hardened and dry, your microarray has been successfully “printed” with the 6 genes!

OR (your instructor will let you know which protocol to use)

Method 2: Set a micropipet to 30 μL . Measure 30 μL of solution from each of the numbered bottles and place onto the corresponding spots on your slide. Use a different tip for each spot! Be sure to replace the appropriate dropper bottle end into each bottle when you are finished.

Part 2: Hybridize your microarray with labeled cDNAs from normal lung tissue and lung cancer tissue.

mRNA was isolated from normal lung cells, and the cDNA created for this mRNA was labeled with a blue dye. mRNA from lung cancer cells was isolated, and its cDNA was labeled with a pink dye. You have been given a bottle (to be shared between groups) containing a solution of labeled cDNAs from lung cancer cells and normal cells mixed together. *You will not be able to see these dyes until you visualize your results at the end of the experiment.*

In a real microarray, the principle behind the hybridization step is as follows: You cannot see the color because the cDNA is very dilute. When added to the printed microarray slide, the labeled cDNAs in the solution will base pair with the complementary DNA for each gene spotted onto the microarray. As each cDNA binds to the appropriate DNA spot on the slide, the labeled cDNA becomes concentrated in that spot, allowing them to be visualized by means of a sophisticated device.

Note: You must wear gloves and goggles. The hybridization solution contains 0.4 M sodium hydroxide (NaOH), which is caustic and causes burns. Do not get it in your eyes, on your skin, or on clothing. If you feel an itching sensation, wash that area of your skin in plenty of running water. Be sure to wash your hands after the lab. If you get hybridization solution in your eyes, flood them with water and seek medical attention. Also seek medical attention if you ingest any.

4. Carefully drop 1–2 drops of “Hybridization Solution” from the dropper bottle onto each spot. Do not allow the dropper bottle to touch the DNA spots!

Part 3: Visualize your labeled microarray results.

5. Record your results by writing a description of the color of each spot or drawing the results below. Your teacher may also take a photo of your slide (be sure your group number is on the slide).

Cleanup

6. Use a paper towel to wipe off the six spots on your slide. Rinse your slide in water and dry it using a paper towel. Wear gloves and eye protection; there is NaOH on the slide!

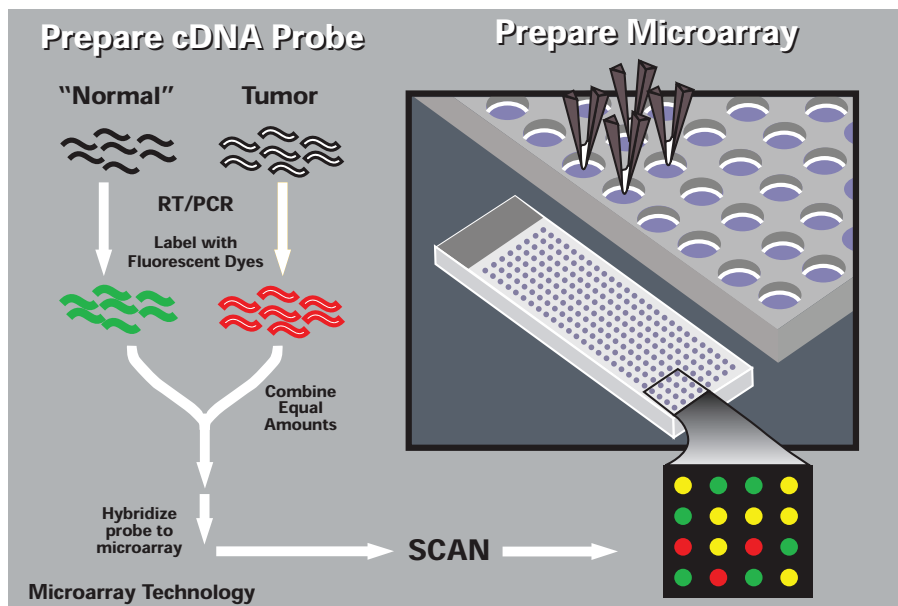
Extensions of the Microarray Unit

- Learn how to quantitate the results of your microarray by going to the following Web site: http://www.bio.davidson.edu/projects/GCAT/HSChips/hs_kit_math_module_v2.pdf.
- Learn more about medical applications of microarrays in the diagnosis of two types of leukemia, AML and ALL, from an activity in DNA Chips: A Genetics Lab in the Palm of Your Hand (in Modern Biology for High School Classrooms—Snapshots of Science and Medicine Magazine): <http://science.education.nih.gov/newsnapshots/index.html>.
- Debate the ethical issues of DNA microarray use for medical and genetic testing. Affymetrix has supported the development of several activities that can be accessed at the following Web site: <http://www.affymetrix.com/corporate/outreach/educator.affx>.
- The genes that we used in this activity are actual genes. To learn more about these genes, you can search for each gene name in the following database: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>.
- Analyze the gene expression profiles for the six genes used in the lung cancer activity by using the Stanford Microarray Database (<http://genome-www5.stanford.edu/MicroArray/SMD/>) or by looking at the gene sequences in the following database: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>.
- The actual microarray study of these genes is described in the following paper:
Garber, M. E., Troyanskaya, O. G., Schluens, K., Petersen, S., Thaesler, Z., Pacyna-Gengelbach, M., van de Rijn, M., Rosen, G. D., Perou, C. M., Whyte, R. I., Altman, R. B., Brown, P. O., Botstein, D., and Petersen, I. (2002) *Diversity of gene expression in adenocarcinoma of the lung*. Proc. Natl. Acad. Sci. 99(2): 1098.

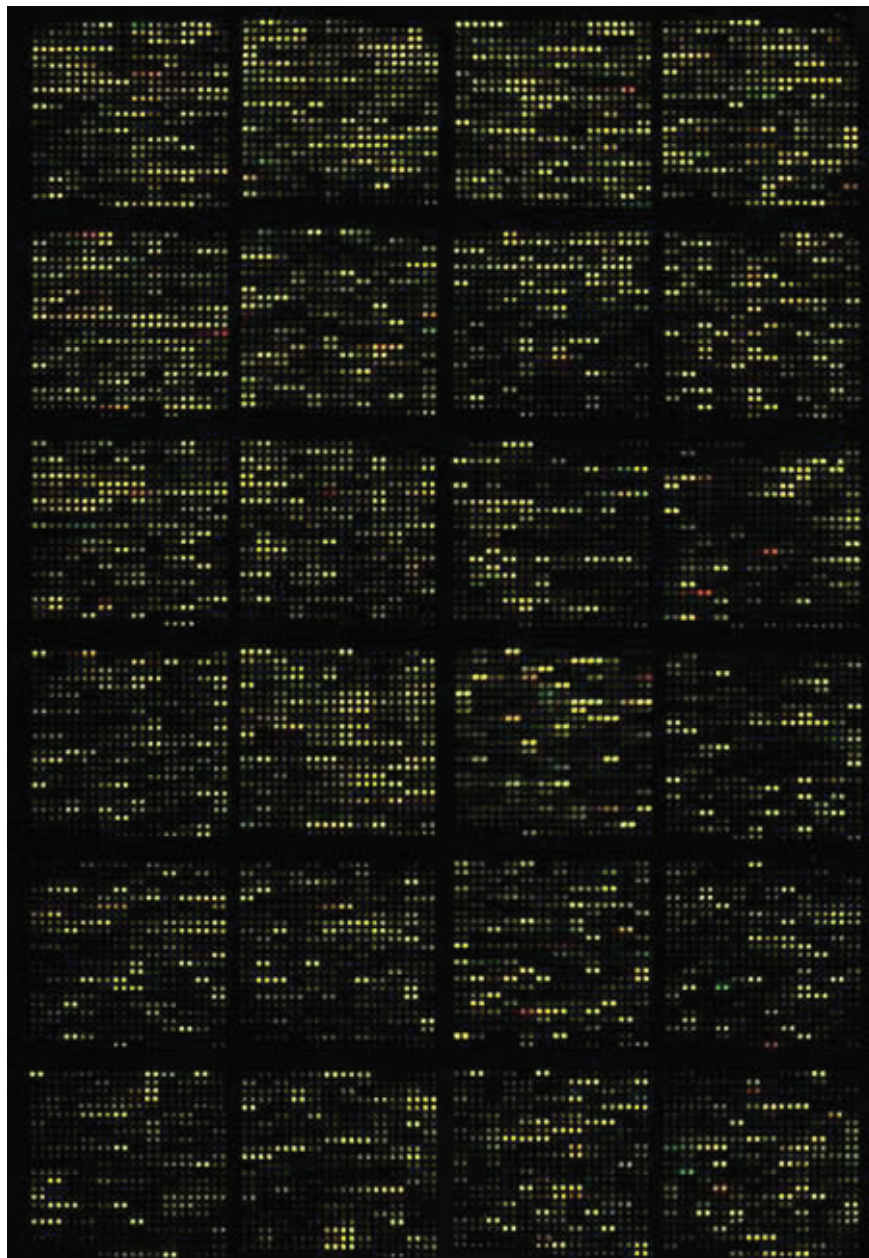
Analysis of Results

1. Which gene(s) were expressed (transcribed) in the lung cancer cells? How do you know?

2. Which gene(s) were not expressed in the lung cancer cells? How do you know?



Using Microarray Technology to Study Gene Expression in Normal and Tumor Cells
(Daryl Leja, National Human Genome Research Institute)



*Courtesy,
 A. Malcolm Campbell,
 Davidson College and
 Genome Consortium
 for Active Teaching*