DNA Microarray

Mr. U.P. Sirdeshmukh

7. DNA Microarrays for Gene Expression

We have outlined the principles of the **DNA microarray** in Chapter 8, DNA Sequencing, while discussing its use in sequencing DNA and in the diagnostic detection of particular DNA sequences. In the previous experiments, DNA immobilized on the chip hybridizes to the target DNA fragments in the sample to be analyzed. Since DNA microarrays work by hybridization, they can also be used to monitor RNA. Microarrays are fairly expensive and analysis of the data is highly labor intensive, despite computerized analysis. If only one or a few genes are the objects of interest, other methods such as Northern hybridization to detect mRNA or using a reporter gene to measure the level of transcription are more appropriate.

DNA microarrays can be used to detect gene expression by hybridizing the array to messenger RNA.

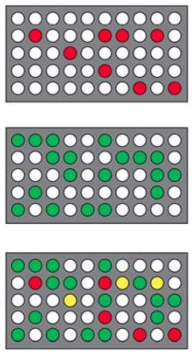
For total transcriptome analysis, the solid support (i.e., the "chip") has DNA sequences complementary to all possible mRNA molecules that a cell might express. The DNA is robotically printed onto a nylon membrane or a glass slide. Current technology can print about 100,000 spots of DNA per cm^2, with glass slides capable of carrying higher densities than nylon membranes. Next, mRNA is extracted from

DNA microarray or DNA array Chip carrying array of DNA segments used to simultaneously detect and identify many short RNA or DNA fragments by hybridization. Also known as DNA chip or oligonucleotide array.
Ribosome footprint The protected area of an mRNA due to the binding of a ribosome.
Ribosome profiling A method to analyze the transcriptome, especially actively translated mRNAs, using the areas protected due to the binding of ribosomes.

cells and labeled, either with a radioactive isotope or more often with a fluorescent dye. Next, the labeled mRNA is placed on the DNA array in conditions that favor binding of complementary sequences. After binding to the chip, the intensity of the label in each spot correlates to the amount of that particular mRNA. Most gene expression studies compare two different conditions, one "control" set or untreated cells, and one "experimental" set where the cells are exposed to a different environment. Both mRNA samples can be hybridized to the chip at the same time if two different fluorescent dyes (e.g., Cy3, which is green and Cy5, which is red) are used for each mRNA set. Red spots will show genes expressed under "control" conditions and green spots will show genes expressed under "experimental" conditions. When the same mRNA is expressed in both conditions that spot will fluoresce yellow (Figs. 21.23 and 21.24). Determining the intensity of green, red, and yellow for each spot is accomplished by computer analysis, which determines the mean of the pixels or median value for the pixels, and normalizes these to a set of internal controls. Rather

FIGURE 21.23
DNA Chip Showing Detection of mRNA by Fluorescent Dyes

DNA chips can monitor many different mRNAs at one time. Each spot on the grid has a different DNA sequence attached. To determine which genes are expressed under which conditions, mRNA is isolated. In this case, mRNA isolated from cells grown under two different conditions is labeled with two different fluorescent dyes. Under condition one (red dye), eight different mRNAs hybridized to DNA spots on the chip. Under condition two, 19 different mRNAs were seen (green dye). Since two different color dyes were used, both samples can be analyzed on the same chip. In this case, the mRNAs that are expressed under both conditions give yellow spots.



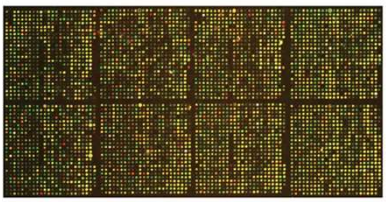
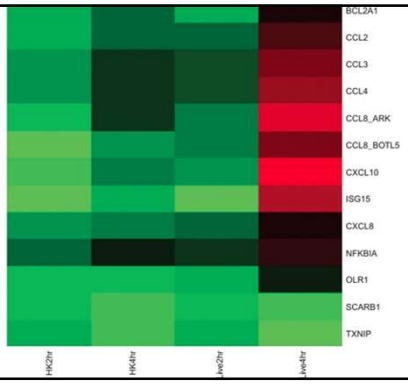


FIGURE 21.24
Hybridization of mRNA to a 19,200 Element Array

RNA from related human colon carcinoma and reference cell lines was reverse transcribed and the cDNA was labeled with Cy5 (red) and Cy3 (green), respectively. The cDNA was then hybridized to a microarray containing 19,200 distinct human cDNA clones. Genes expressed by the cancer cells are shown in red and those from the normal cells are green. Yellow spots indicate expression in both cell lines. *Credit: Hegde et al. (2000) A concise guide to cDNA microarray analysis. Biotechniques 29: 548–562. The Institute for Genomic Research, Rockville, MD.*

than simply presenting a table of numbers, the computerized analysis is often presented as a "heat map" grid. The gene sequences for the control set of data are listed on the x-axis and the experimental genes are listed along the y-axis. Each square of the grid is colored, where red indicates an increase in expression and blue represents a decrease in expression over the control. Shading of either red or blue from light to dark indicates relative increases or decreases of gene expression (Fig. 21.25).

than simply presenting a table of numbers, the computerized analysis is often presented as a "heat map" grid. The gene sequences for the control set of data are listed on the x-axis and the experimental genes are listed along the y-axis. Each square of the grid is colored, where red indicates an increase in expression and blue represents a decrease in expression over the control. Shading of either red or blue from light to dark indicates relative increases or decreases of gene expression (Fig. 21.25).



Heat Map Representation of Microarray Data

The amounts of various mRNA transcripts (listed at right) in cow alveolar macrophages were measured using microarray analysis. The amount of each transcript was determined in four different conditions: HKc2hr compares the mRNA amount in uninfected macrophages vs infected with heat-killed *Mycobacterium bovis*, the causative agent of bovine tuberculosis, for 2 h; HK4hr compares the same after 4 h of exposure; Live2hr used live *M. bovis* for 2 h; and Live4hr used live *M. bovis* for 4 h. The green represents down-regulation and red represents up-regulation. Each value was normalized. *Credit: Widdison, S., Watson, M., & Coffey, T. (2011). Early response of bovine alveolar macrophages to infection with live and heat-killed Mycobacterium bovis. Developmental & Comparative Immunology, 35(5), 580–591.*

In practice, two types of DNA microarray are used for binding mRNA, arrays of cDNA or arrays of oligonucleotides. For a cDNA array, cDNA is generated by PCR amplification of each gene in the organism. One problem is the existence of gene families (e.g., the globin family) whose individual members are highly homologous and may cross-hybridize. To avoid this, it is normal to use sequences from the 3′ end of the cDNA, which often include part of the 3′-UTR of the mRNA transcript. Non-coding sequences diverge much more than coding sequences, and so are much less likely to cross-hybridize. Each amplified cDNA is attached to the nylon sheet or glass slide for use.

cDNA arrays use the cDNA versions of whole genes.

The immobilized DNA molecules in **oligonucleotide arrays** are synthetic segments of single-stranded DNA usually 20–25 nucleotides long. Oligonucleotides are synthesized for each gene in a genome (see Chapter 5: Manipulation of Nucleic Acids). Determining the sequence for each oligonucleotide requires some investigation. A particular sequence of n bases will occur simply by chance every 4^n bases. For a mammalian genome with 3×10^9 bases, n must be at least 16 for a sequence to be unique. It is safer to make oligonucleotides longer than this minimum and, for example, the GeneChip arrays made by Affymetrix Corporation use 25-mers. Determining which 25 nucleotide sequence to use for each gene is also a consideration. The sequence of the oligonucleotide must not create any stem-loop or cloverleaf structures. In addition, the oligonucleotide must not hybridize with stable mRNA secondary structures. To overcome this obstacle, multiple different oligonucleotides for one gene are included at different locations on the chip. These also serve as controls (Box 21.02).

oligonucleotide array DNA array used to simultaneously detect and identify many short RNA or DNA fragments by hybridization. Also known as DNA array or DNA chip.