

FEBRUARY 01, 1996

So Many Genes, So Little Time

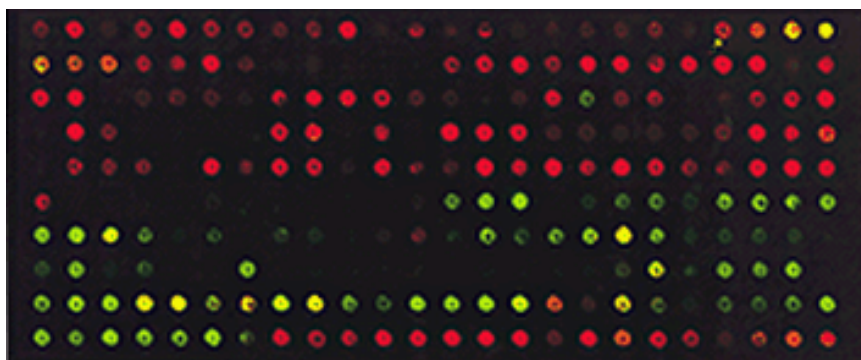


Image Title: In order to determine which genes a yeast spore inherits from its two parents, Patrick Brown's research team created a "microarray" of yeast DNA. Each of the 240 spots in this array contains DNA from a different gene from Chromosome V of baker's yeast. - Patrick O. Brown

Researchers with the Human Genome Project and various commercial ventures are racing to sequence the genomes of organisms in an effort to catalog all their genes. But genes and partial gene sequences are being found at a rate far faster than anyone can decipher their function.

Two groups of Hughes researchers have been attacking that bottleneck with automated programs designed to shed light on the function of newly found genes and the expression patterns of known genes.

Patrick O. Brown, a Hughes investigator at Stanford University School of Medicine, has developed a technique that can tell him instantaneously which of the many thousands of genes present in a cell are switched on at a given moment in time.

Bert Vogelstein, a Hughes investigator at the Johns Hopkins Oncology Center, in collaboration with Kenneth W. Kinzler, also of the Johns Hopkins Oncology Center, has devised a "genetic bar code" to identify unknown genes

and measure gene expression.

Brown's technique, which was published in the October 20, 1995 issue of *Science*, will allow researchers to gather important data about a gene's biology much more rapidly than was previously possible. "Researchers have done a great job of determining the sequences of a huge number of our genes," Brown said. "The problem is, we have the sequences of so many genes socked away in our data bases, but we don't know what they're there for or what the body wants to use them for."

Using the plant *Arabidopsis thaliana* (mouse-ear cress) that has become the model plant for genetic studies, Brown's group has applied the new technique to compare the levels of expression of certain genes in root tissues with those in leaf tissues. Their results have been found to match those obtained using the traditional northern blot method for measuring gene expression. Northern blots of a single gene may take hours or days. Using Brown's new method, a single scientist could gather expression data on more than 1,000 genes in the same time.

The new method hinges upon the use of a speedy robot designed to imprint small glass slides with arrays of up to 20,000 precisely placed microscopic DNA samples. Each sample in the array carries a known DNA sequence corresponding to a particular gene.

After washing the array with a fluorescently labeled mixture prepared from the messenger RNA (mRNA) of the cells under study, the scientists can tell which genes are being expressed in the cells and at what level. "The brighter a dot glows, the higher that gene's level of expression," Brown said.

Brown's team is now collaborating with scientists from the NIH to use the technique to study nearly 1,000 genes whose expression patterns are thought to influence tumor development.

Brown is optimistic that his microarray technique will prove beneficial to clinicians. Detecting gene expression changes in white blood cells, for example, might provide an important new window on the disease defenses of the immune system. "White blood cells circulate through our bodies with little 'antennas' out, asking, 'How's everything going out there? Should we be kicking into gear to deal with a potentially threatening situation?' They have a sensing system that allows them to react to pathological conditions — and what they sense is reflected in gene expression," Brown said.

"If we can figure out how these changes in gene expression reflect the body's condition, we could use white blood cells as little spies wandering through the body and reporting back to us," he said.

A different variation of "genes as reporters" was detailed in the same issue of *Science* by Bert Vogelstein and his colleagues at the Johns Hopkins Oncology Center. Their approach, analogous to a genetic bar code, is called SAGE (serial analysis of gene expression). Vogelstein and Kinzler have assigned a simple nine base pair sequence to each gene, which can be thought

of as being that gene's "bar code." Powerful computers and gene sequencers are employed to read the bar codes in tissue samples.

To demonstrate the technology, the scientists have tested SAGE on pancreas and liver tissue. Once perfected, it took only a few days to obtain thousands of bar codes and identify genes that were specifically expressed in each of the two tissues. In addition to detecting expressed genes of known function, the researchers discovered several new genes which had not been observed previously.

Vogelstein and Kinzler plan to begin using the new technique to compare gene expression patterns in colon cancer cells to those of normal colon cells in an attempt to identify genes that are expressed only in cancer cells. If successful, this approach may lead to better forms of cancer therapy and better diagnostic tests for colon cancer.