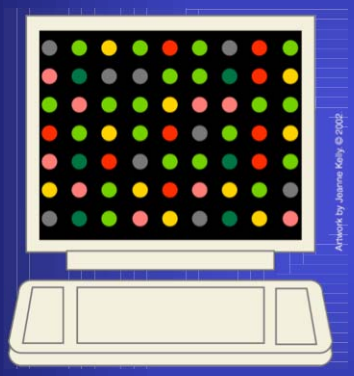
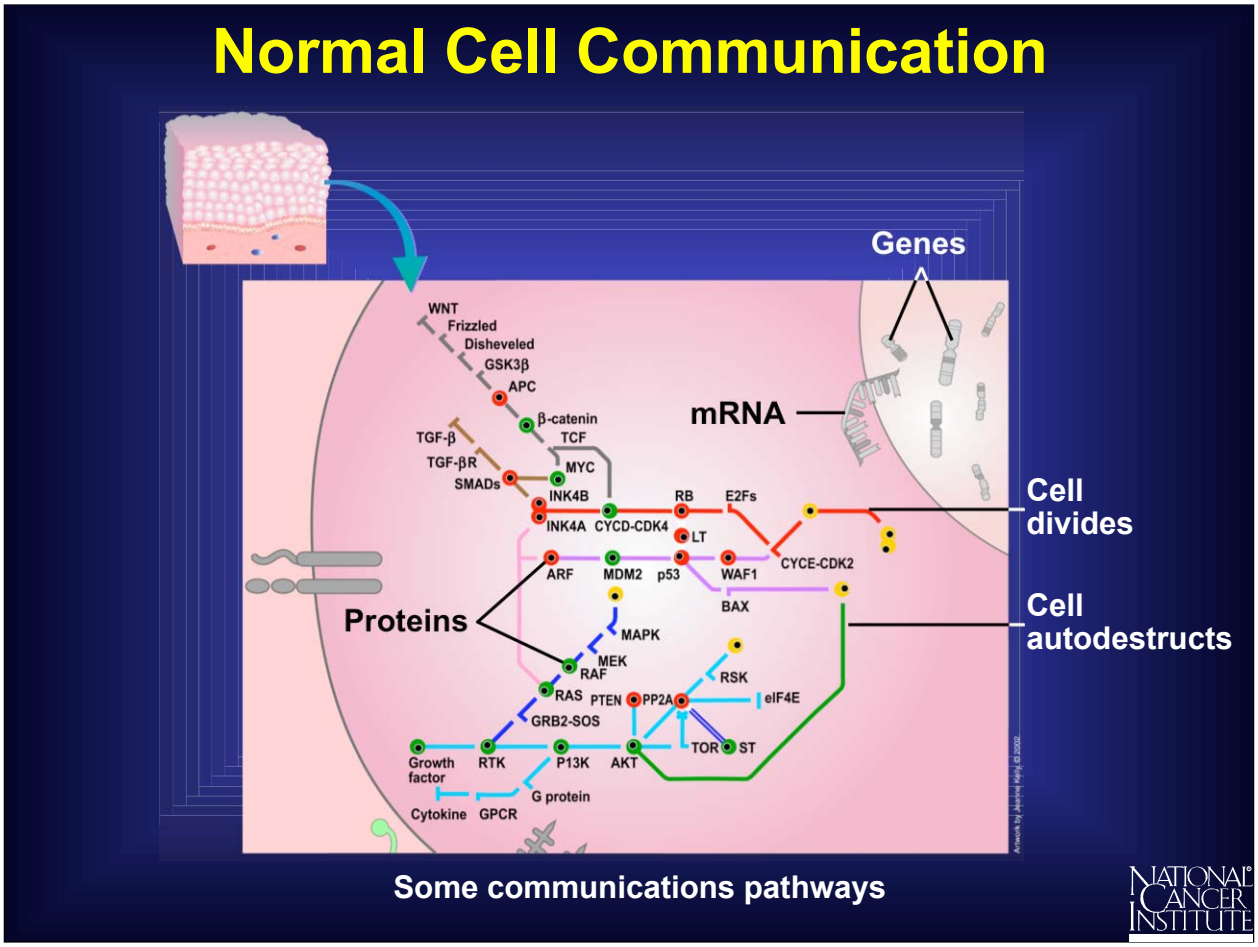
Understanding Cancer and Related Topics

Understanding Molecular Diagnostics

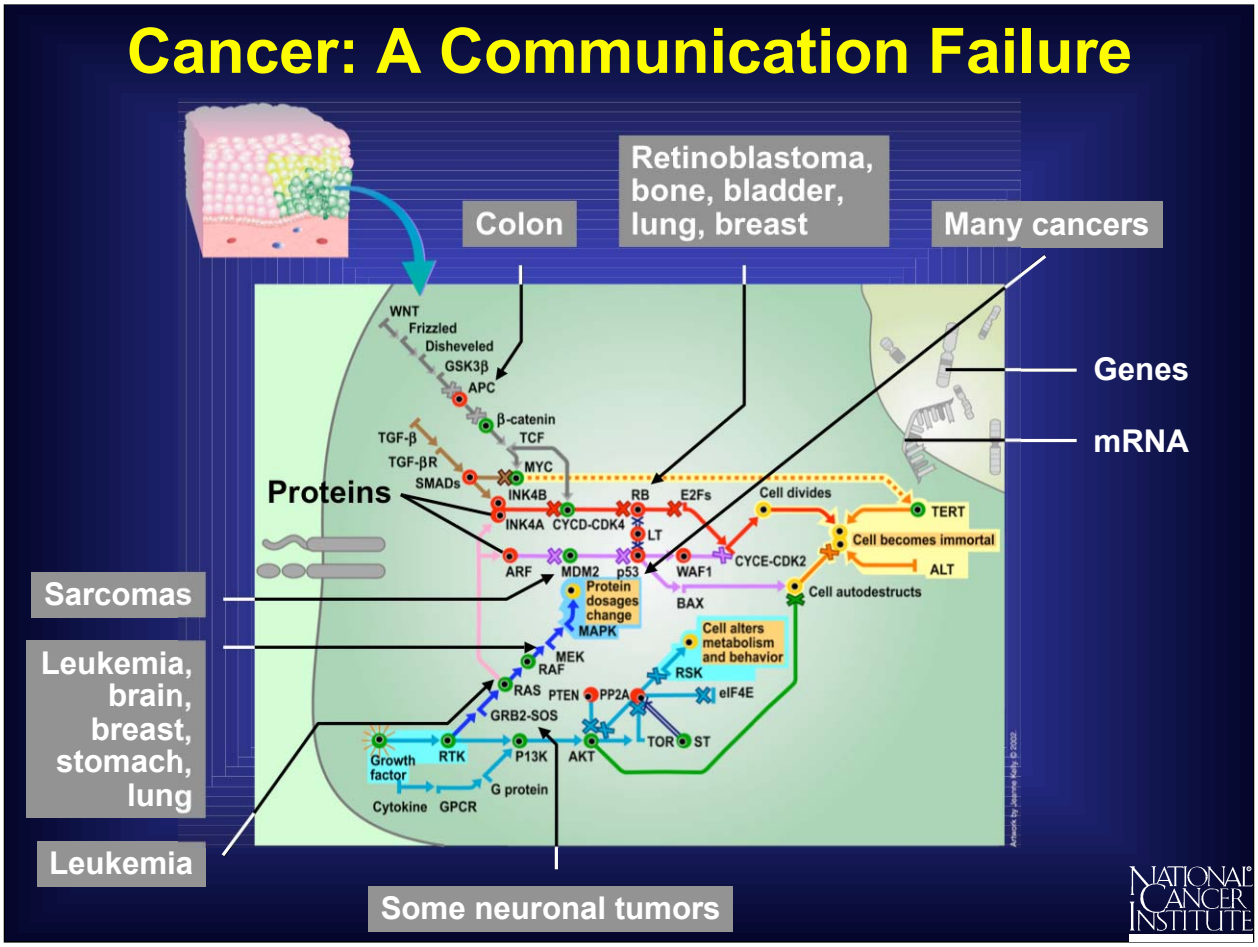


Developed by:
Donna Kerrigan, M.S.
Jeanne Kelly
Brian Hollen

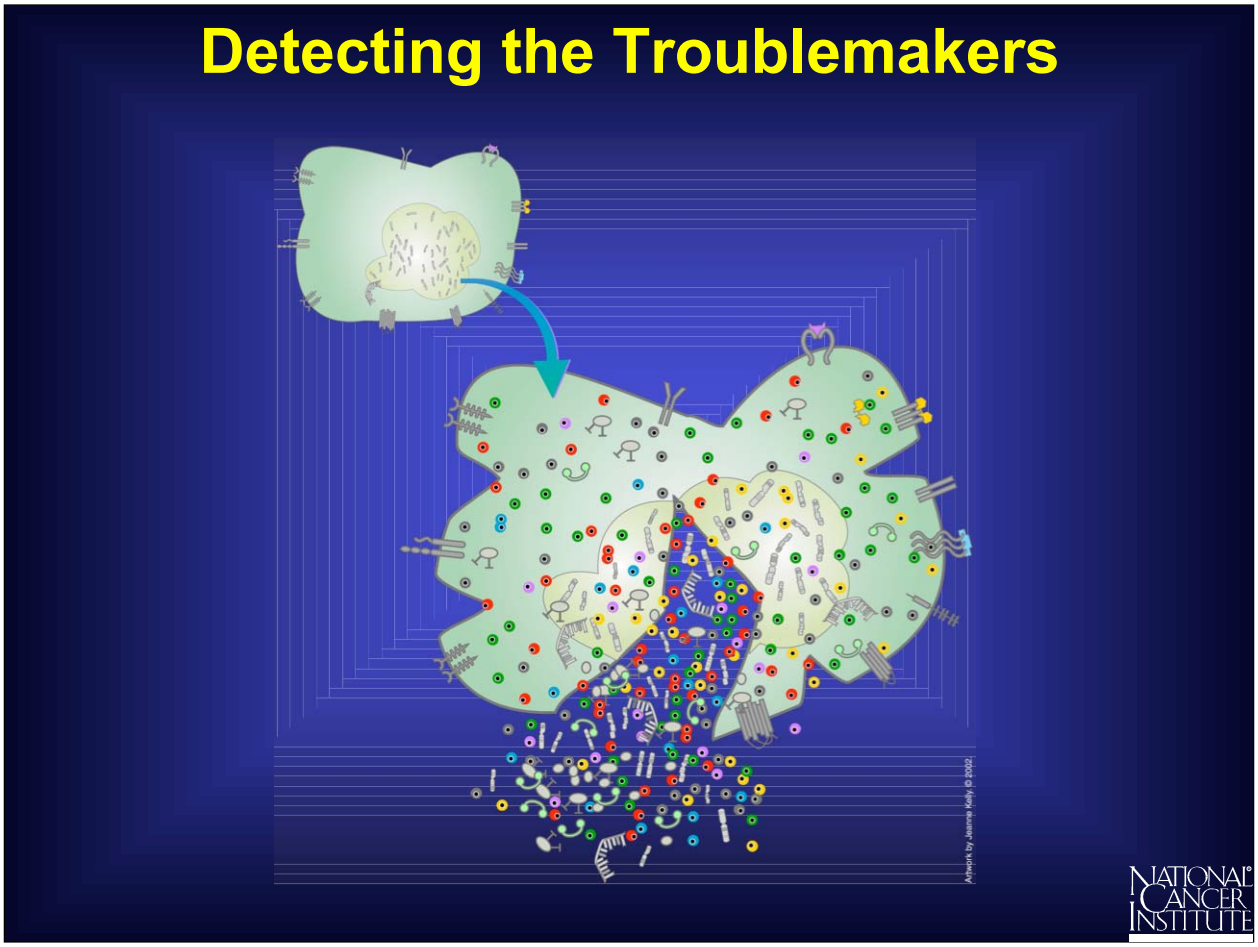
Molecular diagnostics is a new discipline that captures genomic and proteomic expression patterns and uses the information to distinguish between normal, precancerous, and cancerous tissues at the molecular level. This tutorial explains how this data is being used to create new tools for better cancer detection and diagnosis. The tutorial also shows how the new discipline should improve treatment planning for patients.



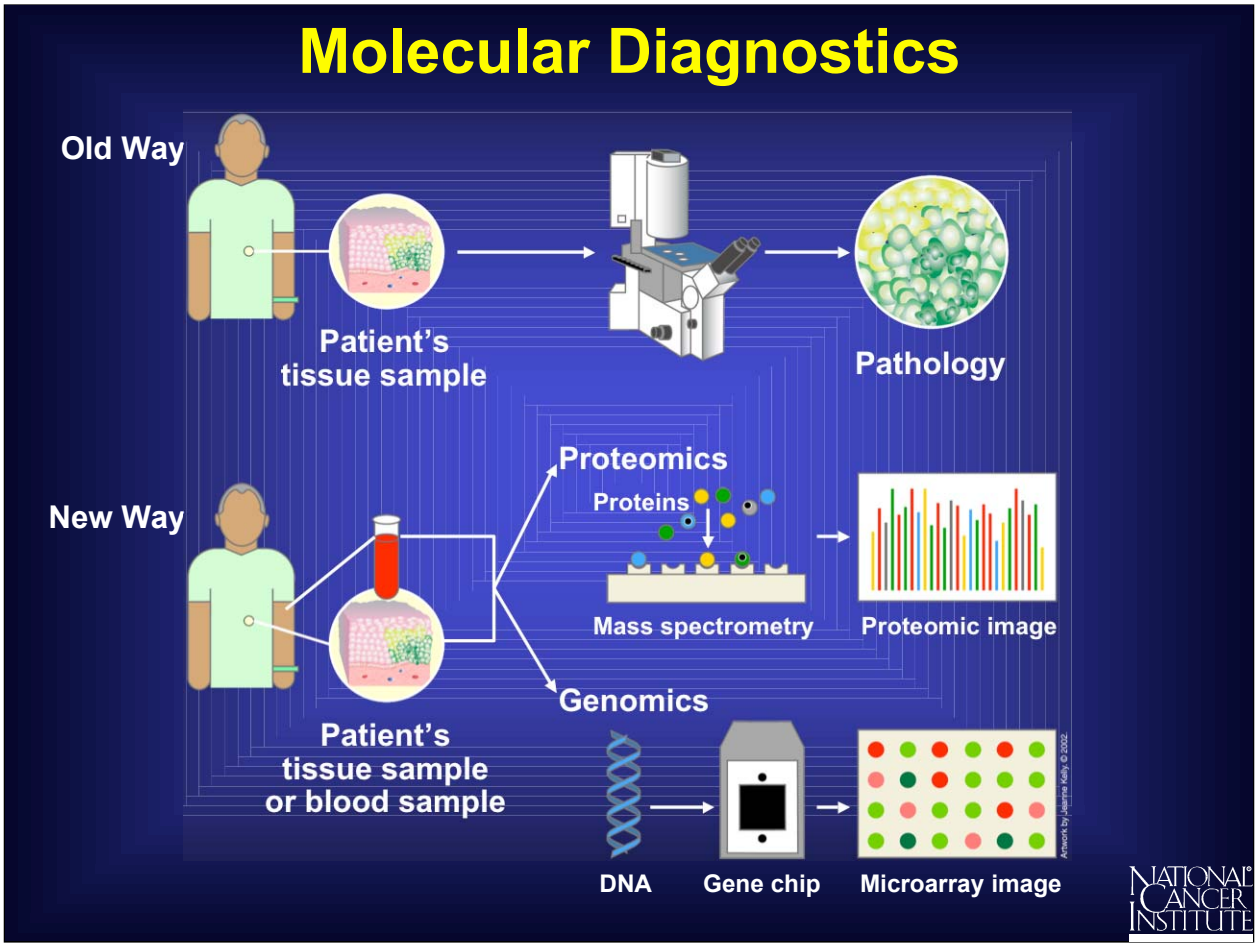
Crucial to all normal cell growth is a communication network that functions properly. This network is an intricate collection of pathways built with interactive proteins. Along these pathways, precise protein-to-protein signaling enables a very carefully guarded regulation of growth. Here are some examples of cellular pathways.



The genetic changes involved in cancer result in altered proteins that disrupt the cell's communication network. In cancer, altered proteins along many different pathways cause signals to be garbled, intercepted, amplified, or misdirected. These changes hijack what was once normal communication and use it to achieve uncontrolled tumor growth. Here are some examples of how cancer disrupts normal cellular pathways.

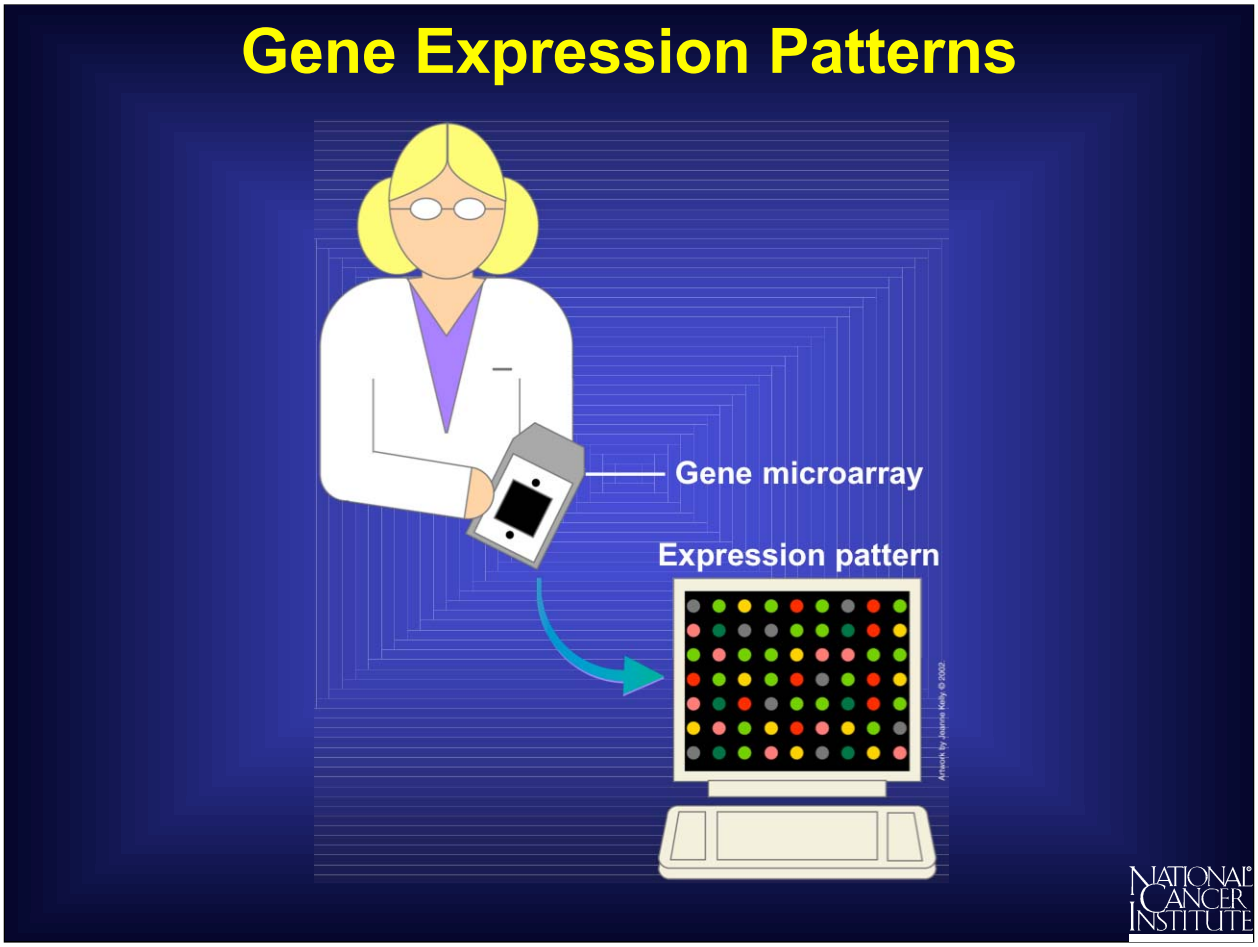


The challenge for cancer detection and diagnosis is to locate the renegade genes and proteins--the deranged, defective, and dominating molecules--that hijack communication in once-normal cells. This requires opening the cell and analyzing the biomolecules inside. The earlier this detection and diagnosis can occur, the better.

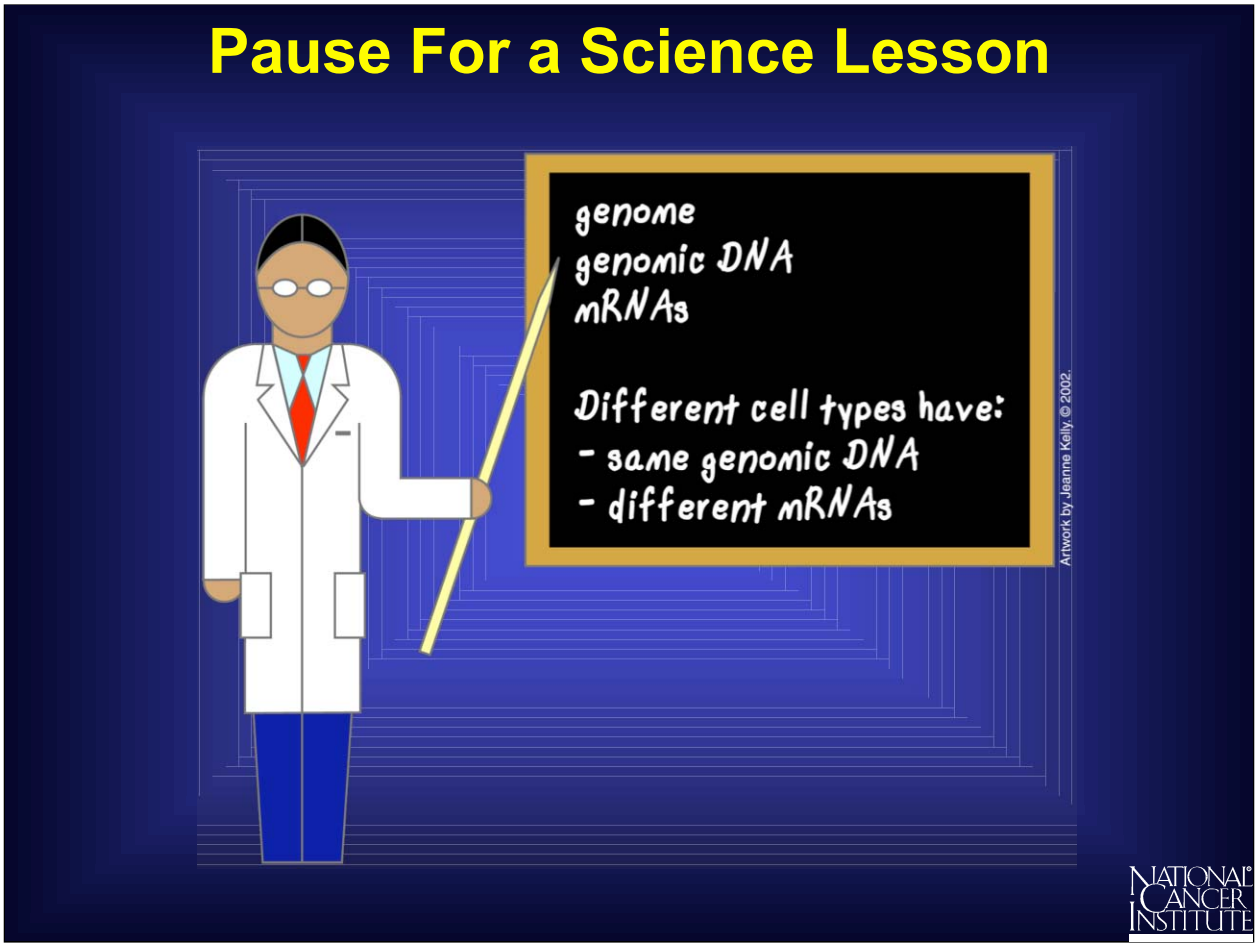


Before molecular diagnostics, clinicians categorized cancer cells according to their pathology, that is, according to their appearance under a microscope.

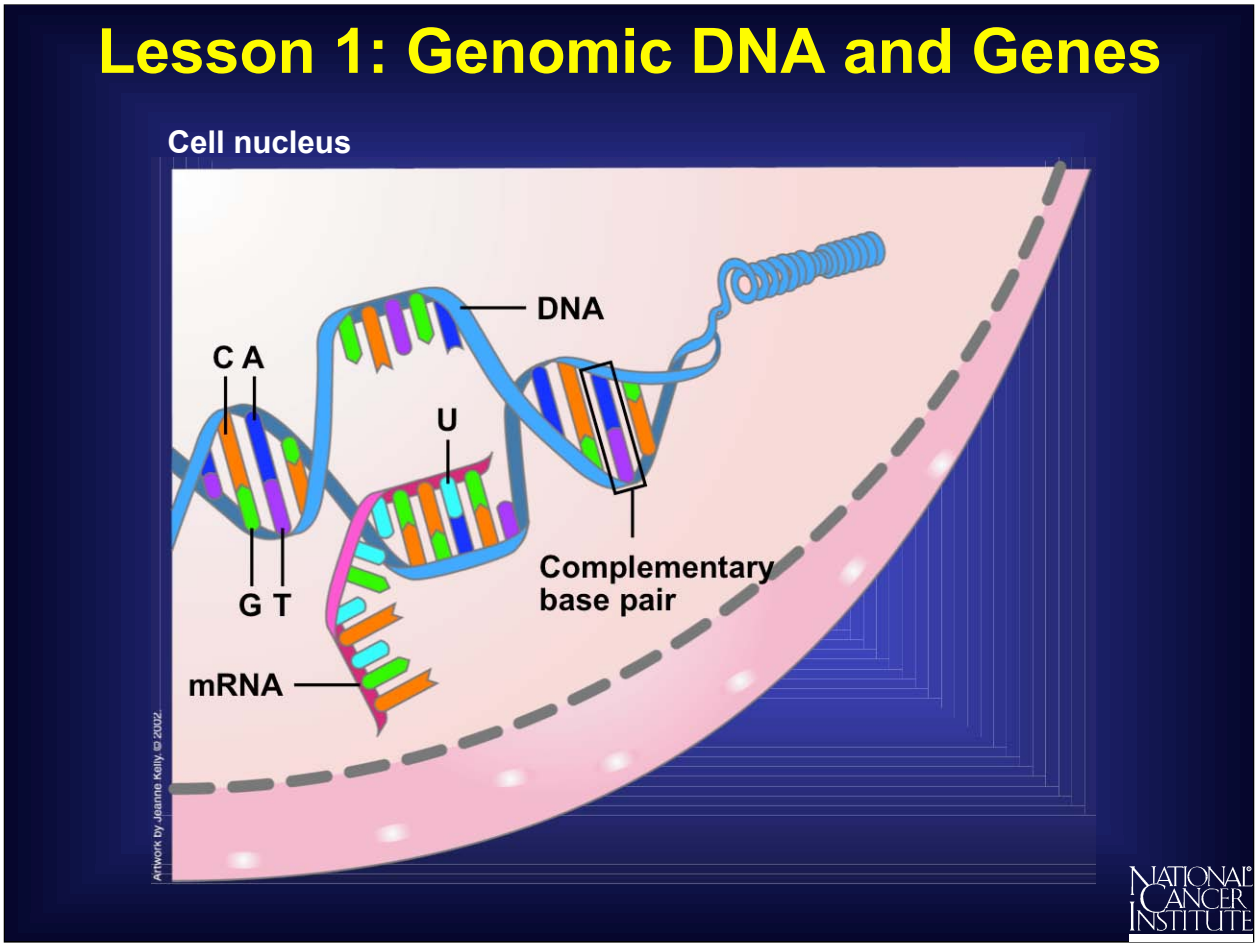
Borrowing from two new disciplines, genomics (gee-no-micks) and proteomics (proh-tee-oh-mics), molecular diagnostics categorizes cancer using technology such as mass spectrometry and gene chips. Genomics is the study of all the genes in a cell or organism, while proteomics is the study of all the proteins. Molecular diagnostics determines how these genes and proteins are interacting in a cell. It focuses upon patterns--gene and protein activity patterns--in different types of cancerous or precancerous cells. Molecular diagnostics uncovers these sets of changes and captures this information as expression patterns. Also called " molecular signatures," these expression patterns are improving the clinicians' ability to diagnose cancer. Soon all cancers may be diagnosed this way.



A major emphasis in molecular diagnostics is placed upon the expression patterns of genes. The challenge is to find the genes active in cancer and separate them from all of the others in a cell. Fortunately, new technology makes success possible. DNA microarrays, sometimes called "gene chips," allow researchers to "see" the expression of hundreds or thousands of genes all at once.

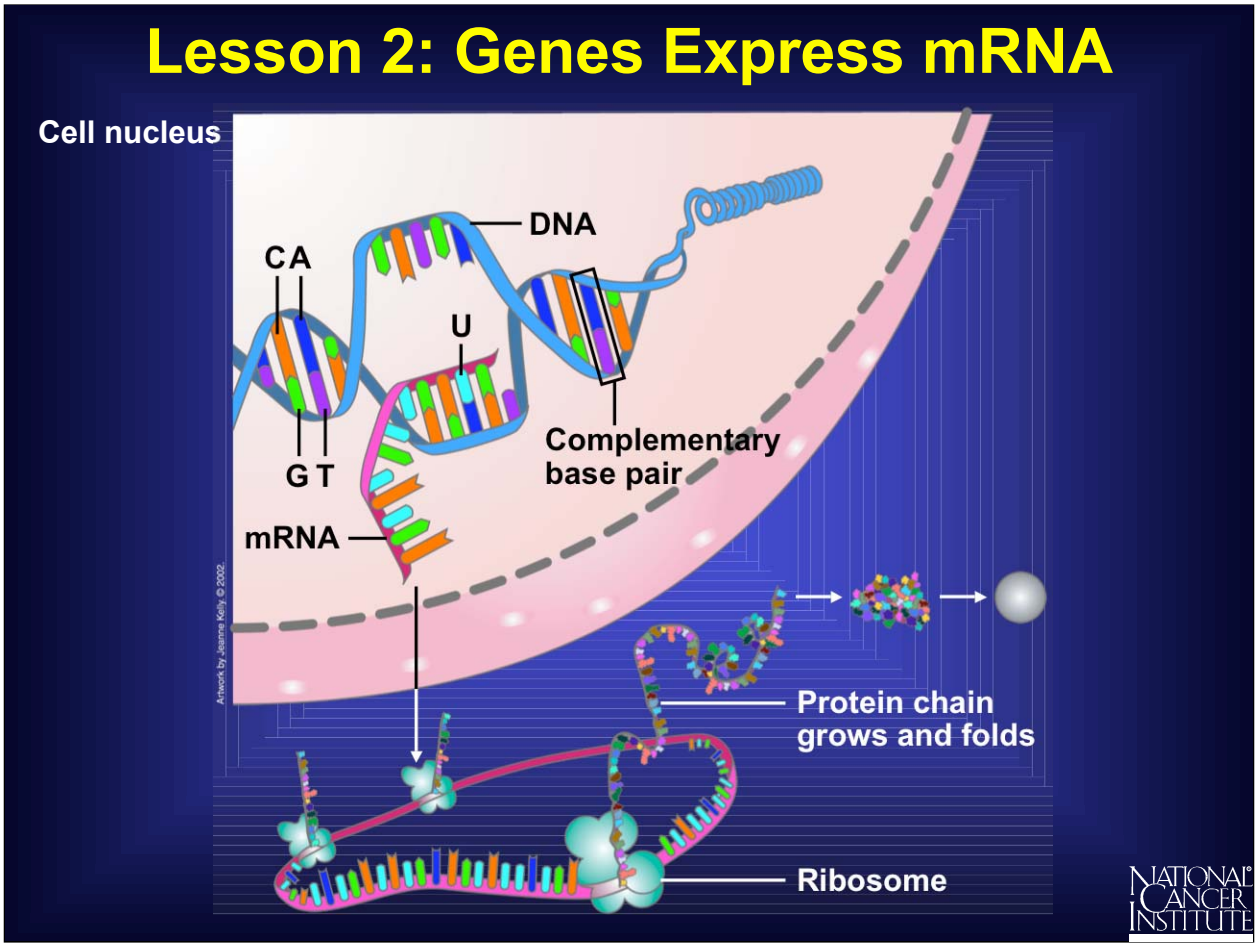


To appreciate how molecular diagnostics uses DNA microarrays to measure gene activity, you first must understand a little about the human genome, genomic DNA, and messenger RNA's link to its expression. You should also remember that gene expression patterns will vary with cell type throughout the human body.

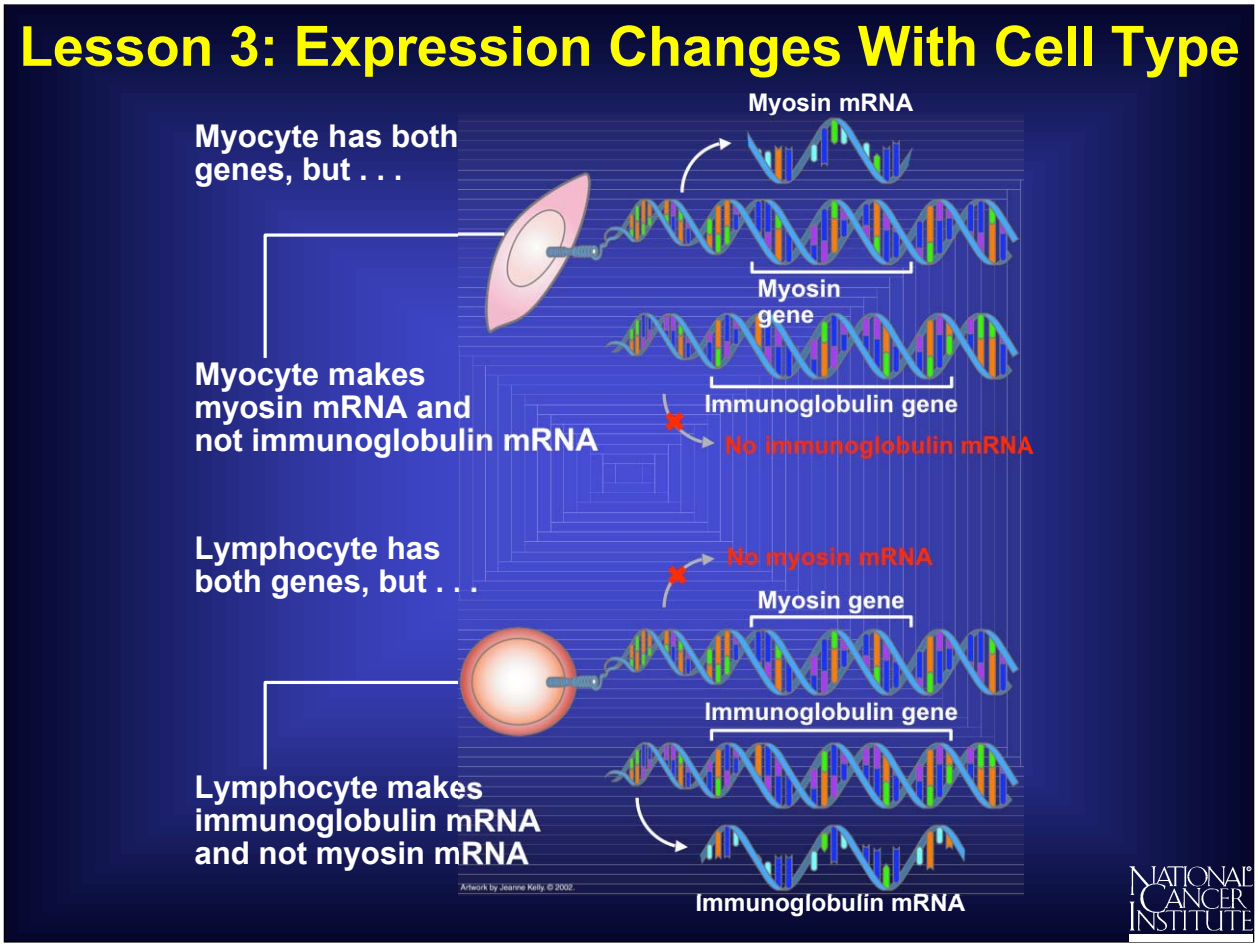


The human genome is the *complete* set of genetic "instructions" for human life. This information resides in genes within very large molecules of deoxyribonucleic acid (DNA). The DNA in the human genome that is passed on to offspring as information necessary for survival is called genomic DNA.

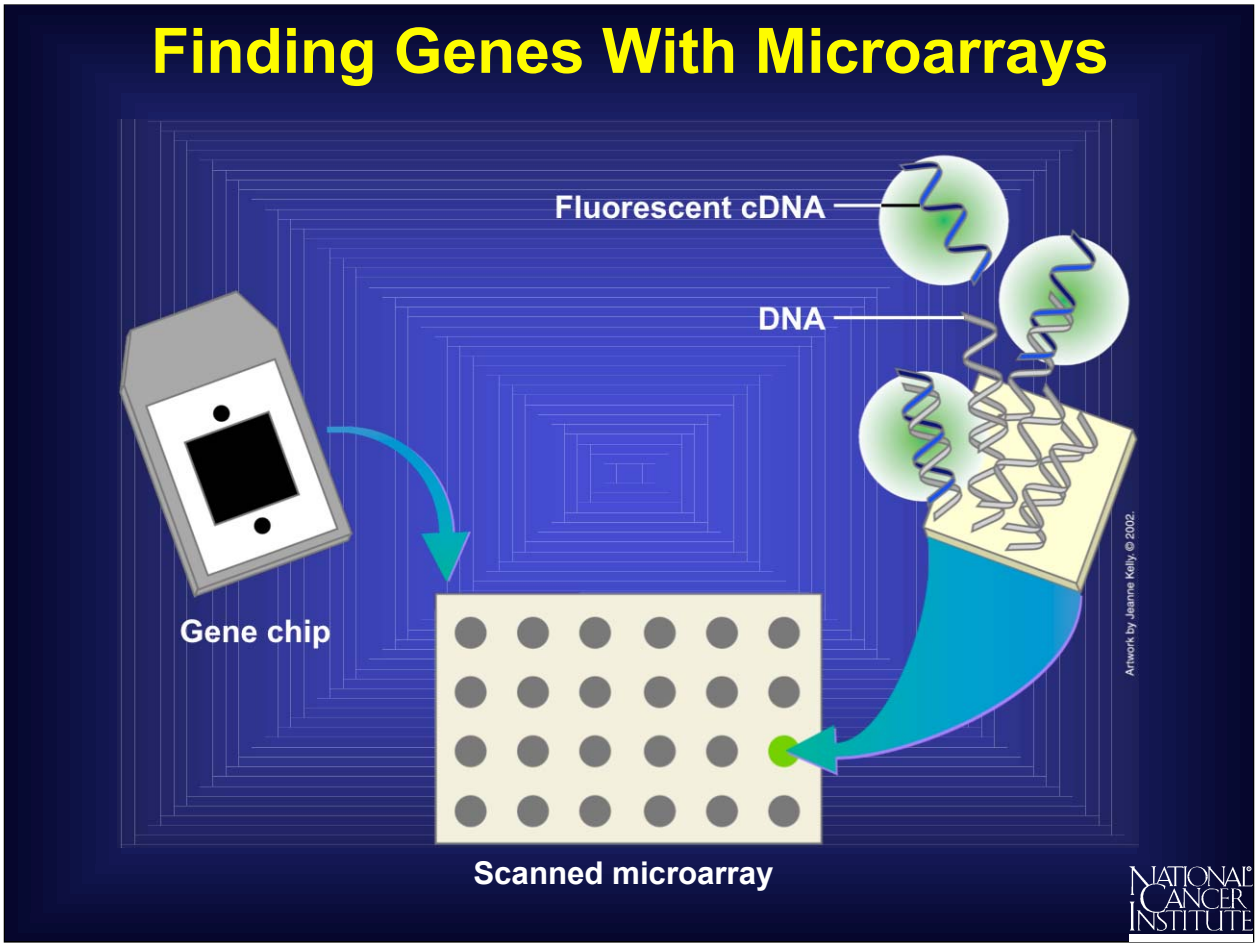
The large DNA molecules are composed of two strands twisted around each other to form a "double helix." Each strand is constructed from millions of chemical building blocks called "bases." DNA contains only four different bases (abbreviated A, T, G, C), but they can be arranged in any sequence. The order of the bases determines the message the gene contains, just as the letters of the alphabet can recombine to form new words. The base sequences of the two strands of each DNA molecule are related to each other by the following rule: A only matches with T, and C only with G. Therefore, the base sequence on one strand dictates the order of the bases on the other strand. This is called "base pairing" and enables the genome to copy itself and pass genomic DNA to the next generation.



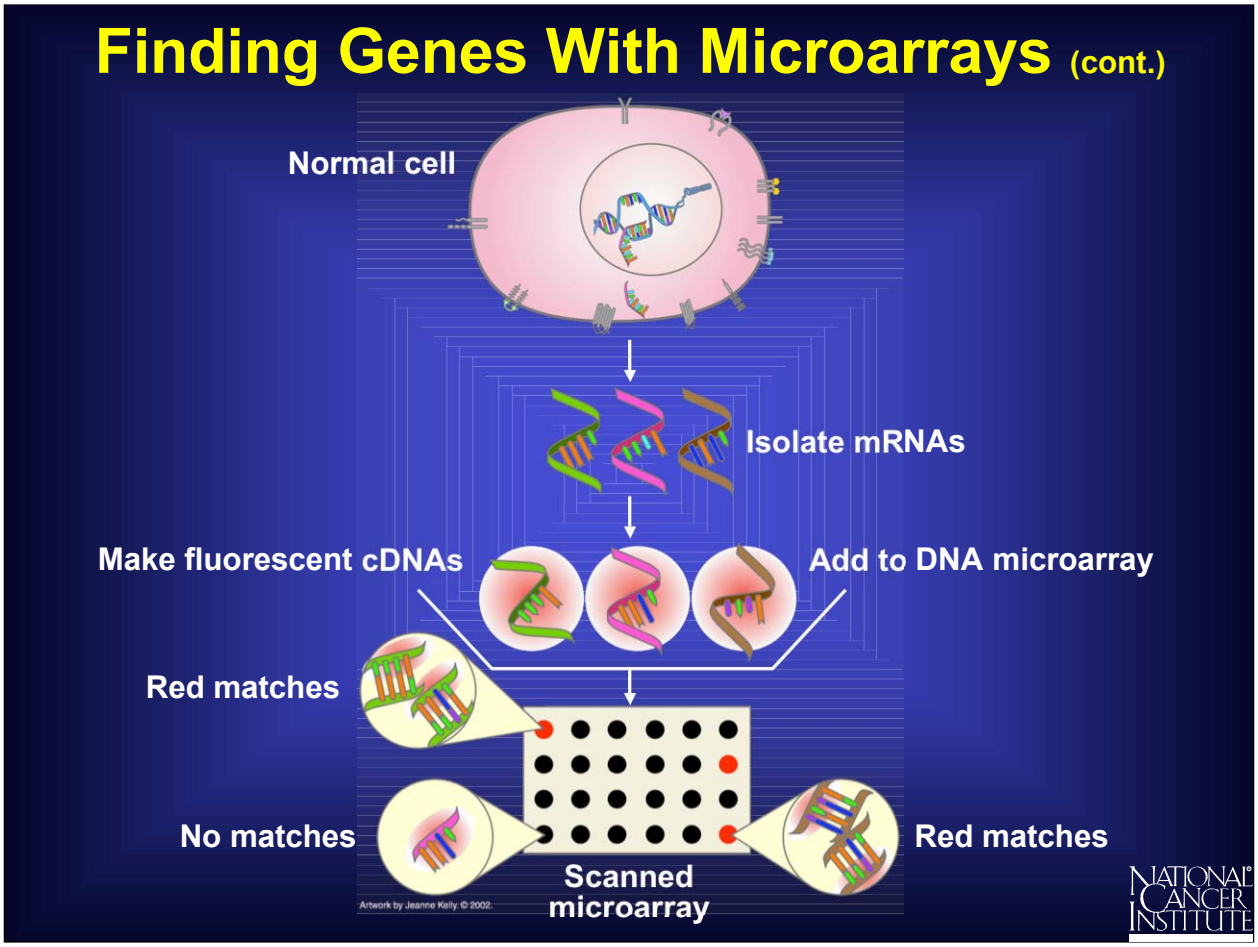
When a gene expresses itself, it "switches on" to produce a protein. The gene does so by first directing the synthesis of an intermediary molecule called messenger ribonucleic acid (mRNA). To transfer a gene's information from DNA to mRNA, base pairing is used. However, there is one change. RNA uses a new rule: A matches only with U, and C only with G. So the base sequence of an mRNA molecule resembles that of the DNA molecule from which it was copied, except the base U appears anywhere the base T would appear in DNA.



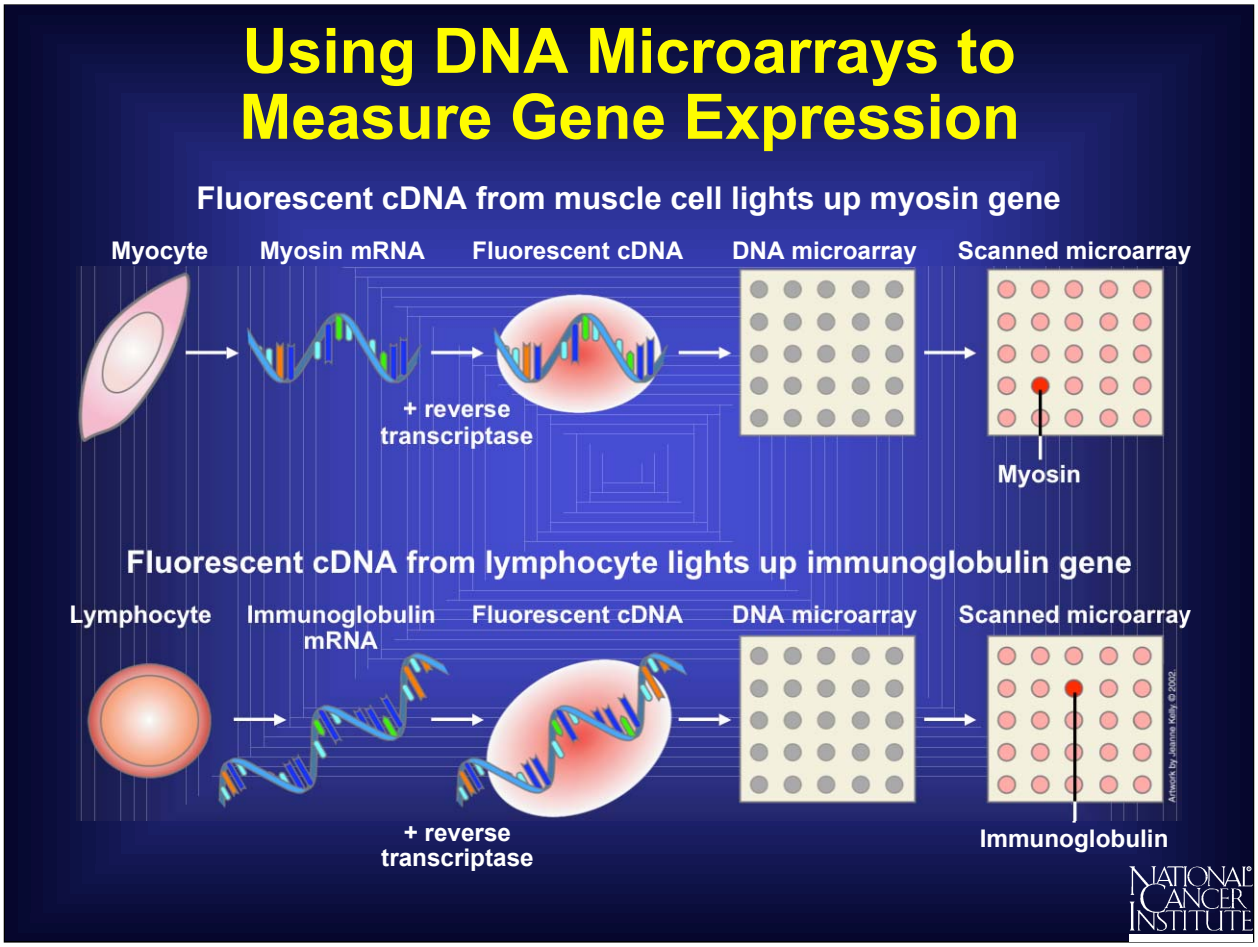
Different cell types express a distinct set of mRNAs from their genomes. For example, while a muscle cell (myocyte) and an immune cell (lymphocyte) both possess the same inherited genomic DNA, regulatory networks inside each cell type cause different subsets of these genes to be expressed as mRNA. And sometimes different cell types express the same mRNAs, but one cell type will produce more copies than the other.



At any given time in each cell in the body, thousands of different genes are active. Until recently, it has not been possible to capture and compare the patterns of gene expression present in different cells in any systematic way. DNA microarrays allow the comparison of thousands of genes that can be measured simultaneously, and the information gained using these arrays is dramatically changing cancer-treatment decisions.

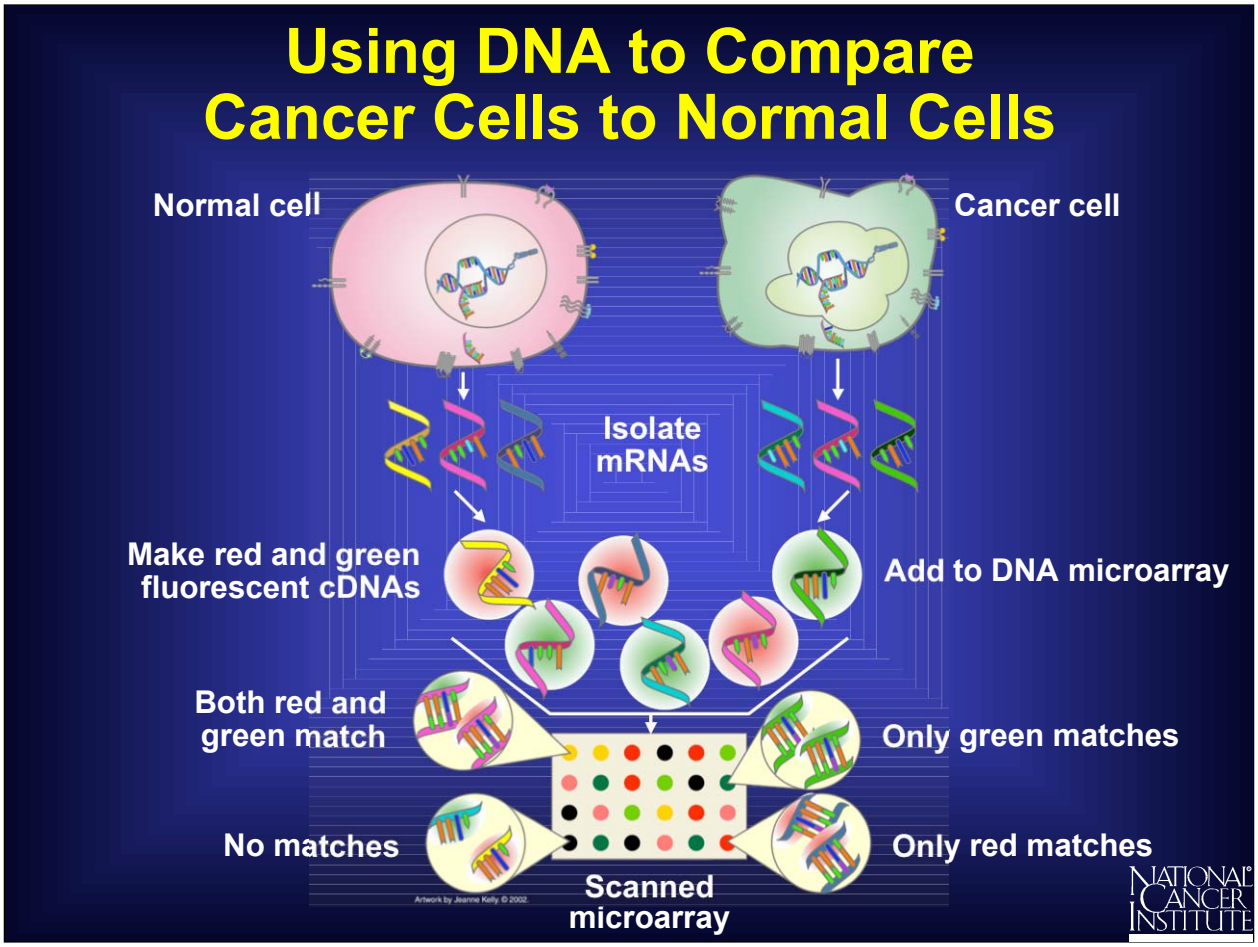


A DNA microarray is a thin-sized chip that has been spotted at fixed locations with thousands of single-stranded DNA fragments corresponding to various genes of interest. A single microarray may contain 10,000 or more spots, each containing pieces of DNA from a different gene. A single gene chip can even hold representative fragments from the entire human genome.

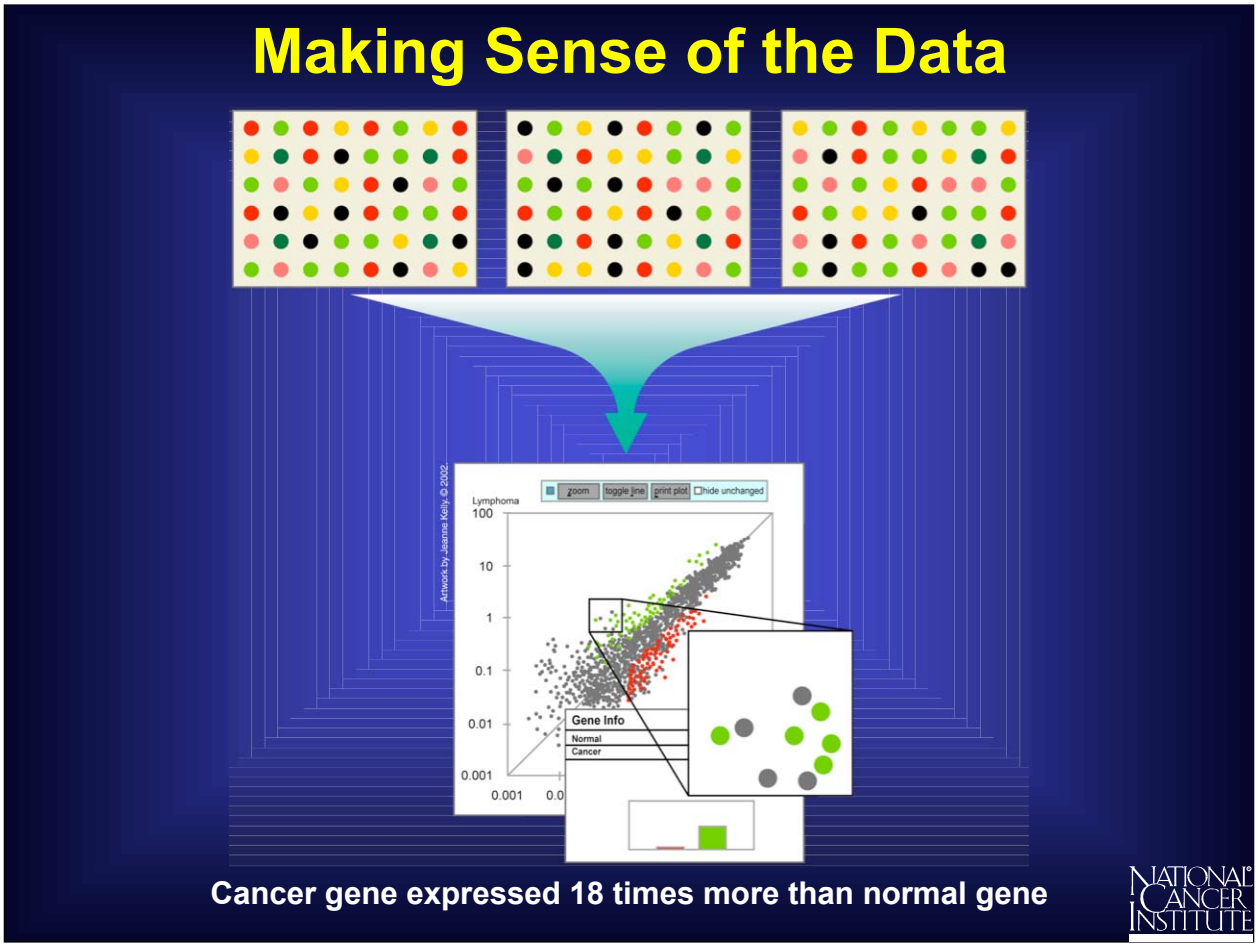


To determine which genes are being expressed in any given cell population, mRNA molecules, which are produced by active genes as they assemble proteins, are isolated from the cells and copied with a special enzyme called "reverse transcriptase." The enzyme copies the mRNA strand by using the DNA rule (see 8), and the copy is called a cDNA. Thus, each cDNA made using reverse transcriptase corresponds directly to a specific mRNA that was coming from an active gene in the cell.

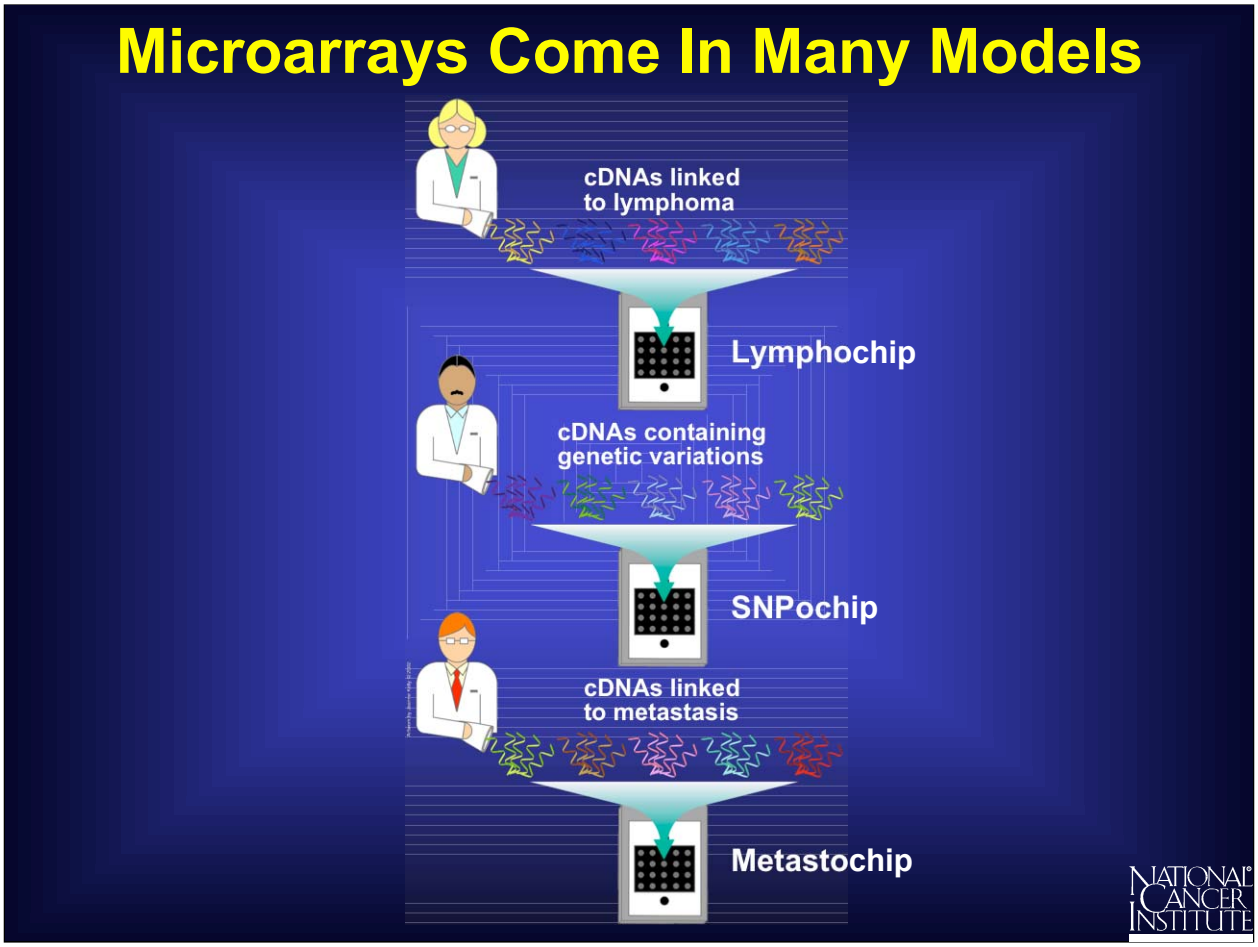
All the cDNAs are then attached to a fluorescent dye. When a DNA microarray is bathed with the fluorescent cDNAs, each cDNA molecule will bind by base-pairing to the spot where pieces of its specific matching gene are located. Therefore, each fluorescent spot in the microarray corresponds to a gene that was actively being transcribed into mRNA in the original cell.



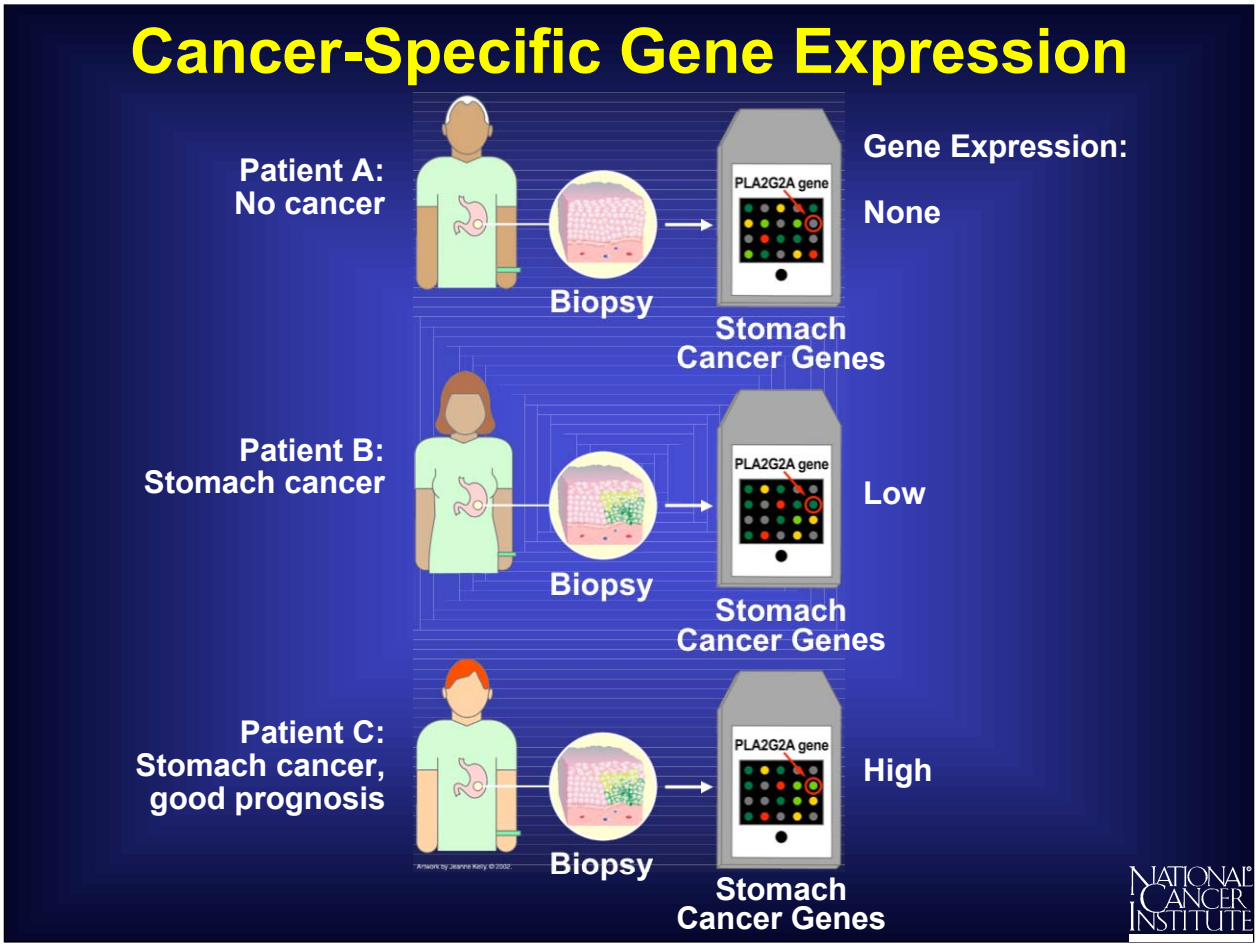
Besides being used to assess the activity of genes being expressed in a single sample, DNA microarrays can also be used to compare the patterns of gene expression in two different cell populations, such as a population of cancer cells with a population of normal cells. In this case, two different fluorescent dyes are used. For example, a red dye can label cDNAs derived from corresponding normal cells, and a green dye, those derived from cancer cells. When the red and green cDNAs are mixed together and placed on a DNA microarray, the green cDNAs will bind to genes expressed in cancer cells and the red cDNAs will bind to genes expressed in normal cells. Green spots represent more copies of a gene being made in a cancer cell, and red spots mean more copies are being made in the normal cell. Yellow spots, caused by mixing red and green fluorescence, represent genes whose expression is roughly the same in both cell types, and black spots, or absence of fluorescence, represent genes expressed in neither cell type.



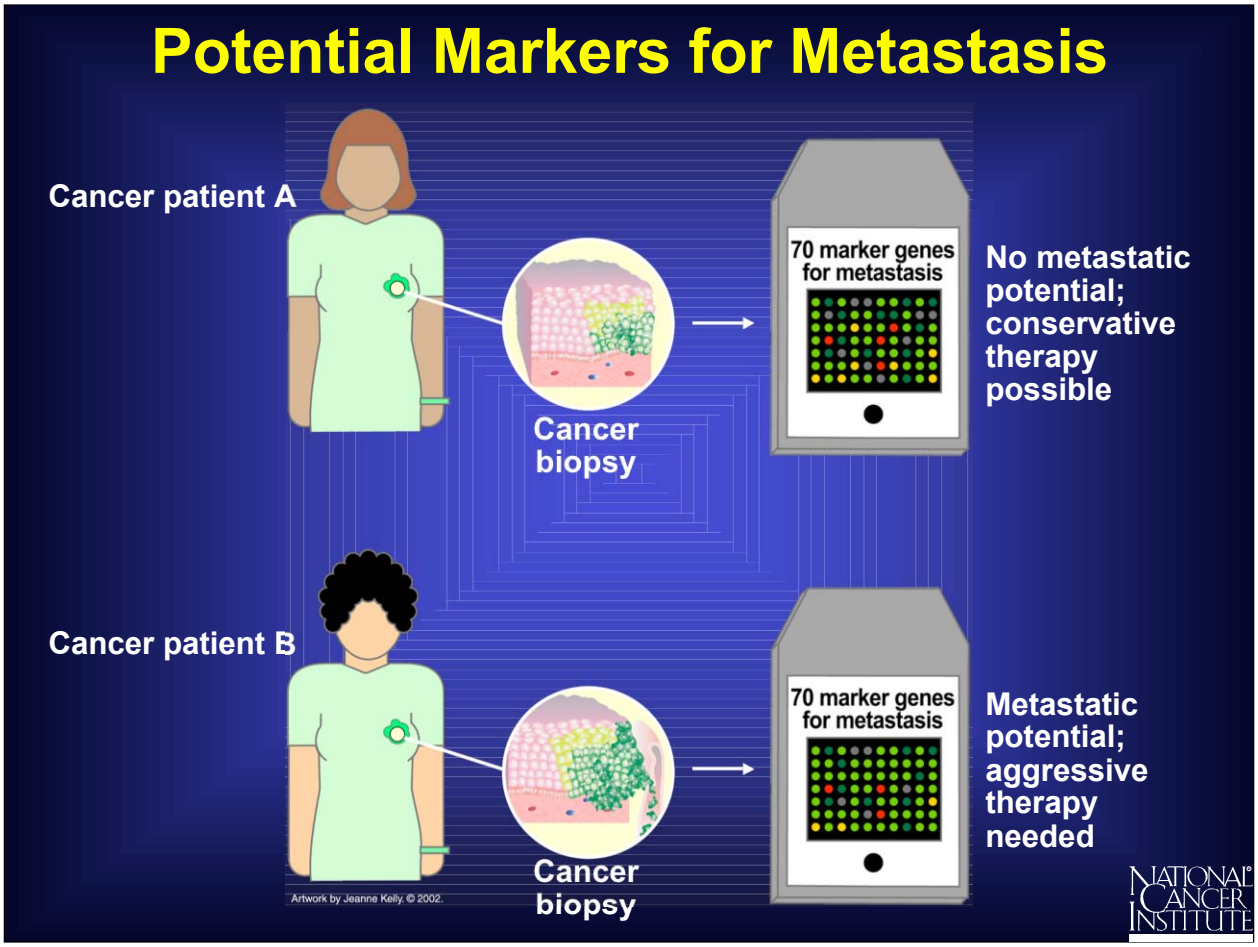
Microarrays make it possible to compare the relative expression of thousands of genes in cancer and normal cells by measuring the intensity and color of the fluorescence of each spot. A computer program can then analyze the data and make precise measurements of genes expressed at high levels and low levels. Most microarray studies now include "cluster analysis," which groups together genes that have similar levels of expression.



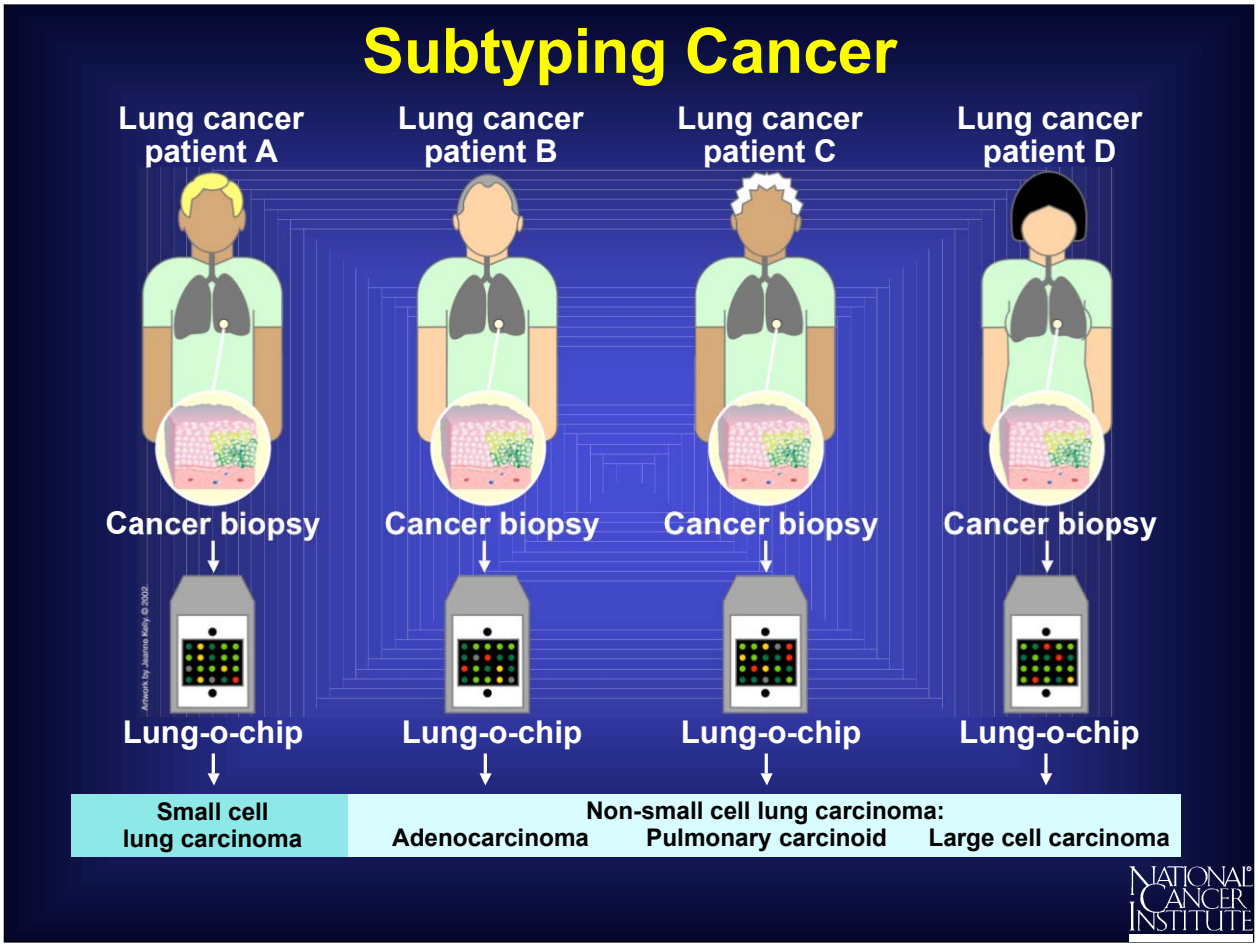
There are some standard microarrays for use in cancer research, such as the Lymphochip. Standard arrays include genes that have been reported as important in a certain type of cancer. Other arrays contain strands of DNA called single nucleotide polymorphisms, or SNPs ("snips"), which are common variations in DNA. Researchers also can create their own microarrays using whatever genes (or gene fragments) they think might be important to the questions they are trying to answer.



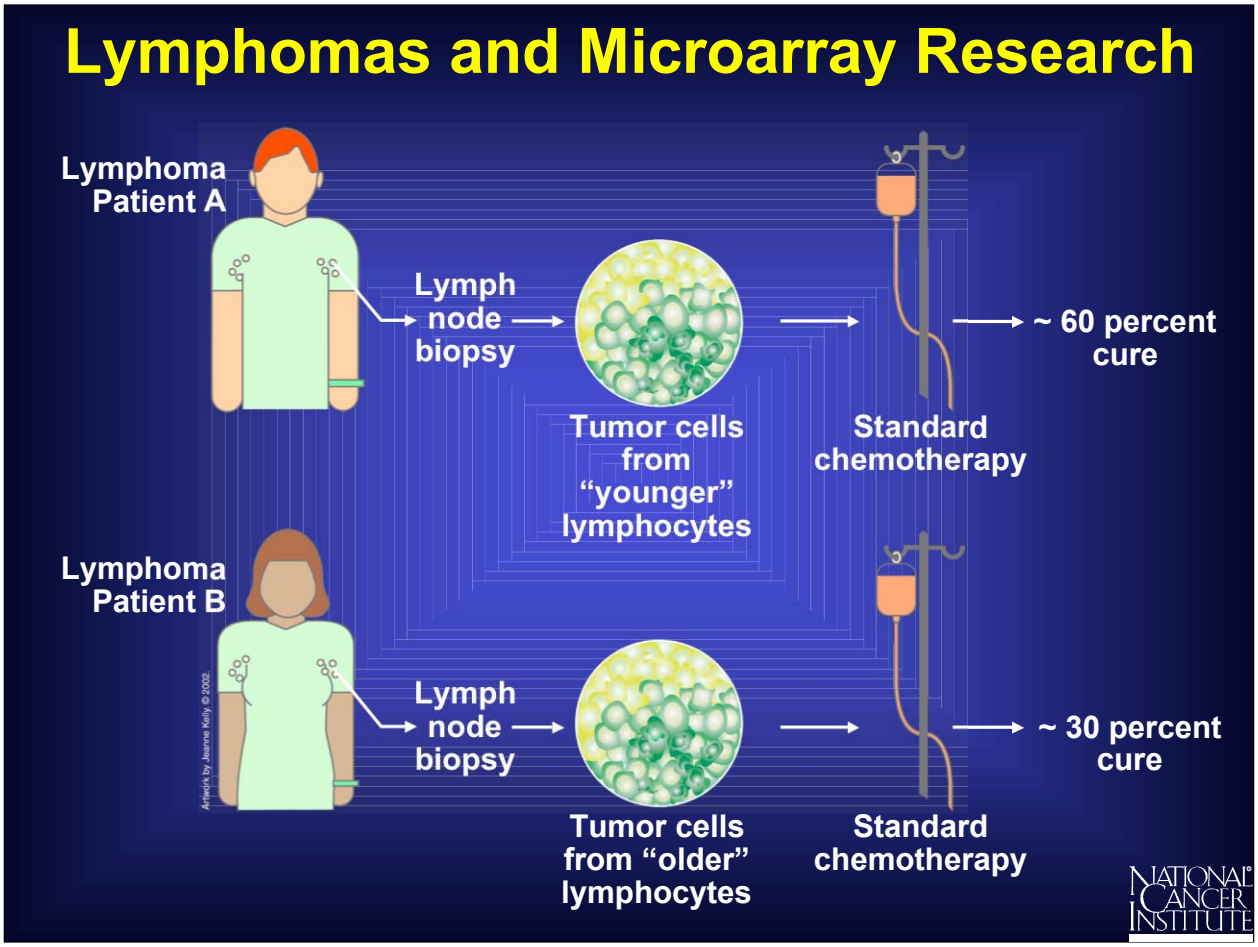
Microarrays are being used in cancer research with incredible results. For example, researchers used microarrays to compare gene activity patterns in normal stomach tissue and tissue from stomach tumors. They found that the tumor tissue, but not the normal tissue, expressed a gene called *PLA2G2A*. But they also found that cancer patients with high expression levels of *PLA2G2A* were more likely to survive for 5 years, compared with patients whose tumors produced lower levels.



Gene expression patterns are also being used as a tool for improving the diagnosis of breast cancer. Microarrays have been used to examine the expression patterns of some 25,000 genes in tissues from breast cancer patients. Computer cluster analysis of the patterns led to the identification of about 70 marker genes that can correctly identify about 90 percent of the women who would eventually develop metastases. While still experimental, such studies bring hope that clinicians will be able to act much sooner and use new technology to better monitor the women more likely to experience aggressive progression of their breast cancer.



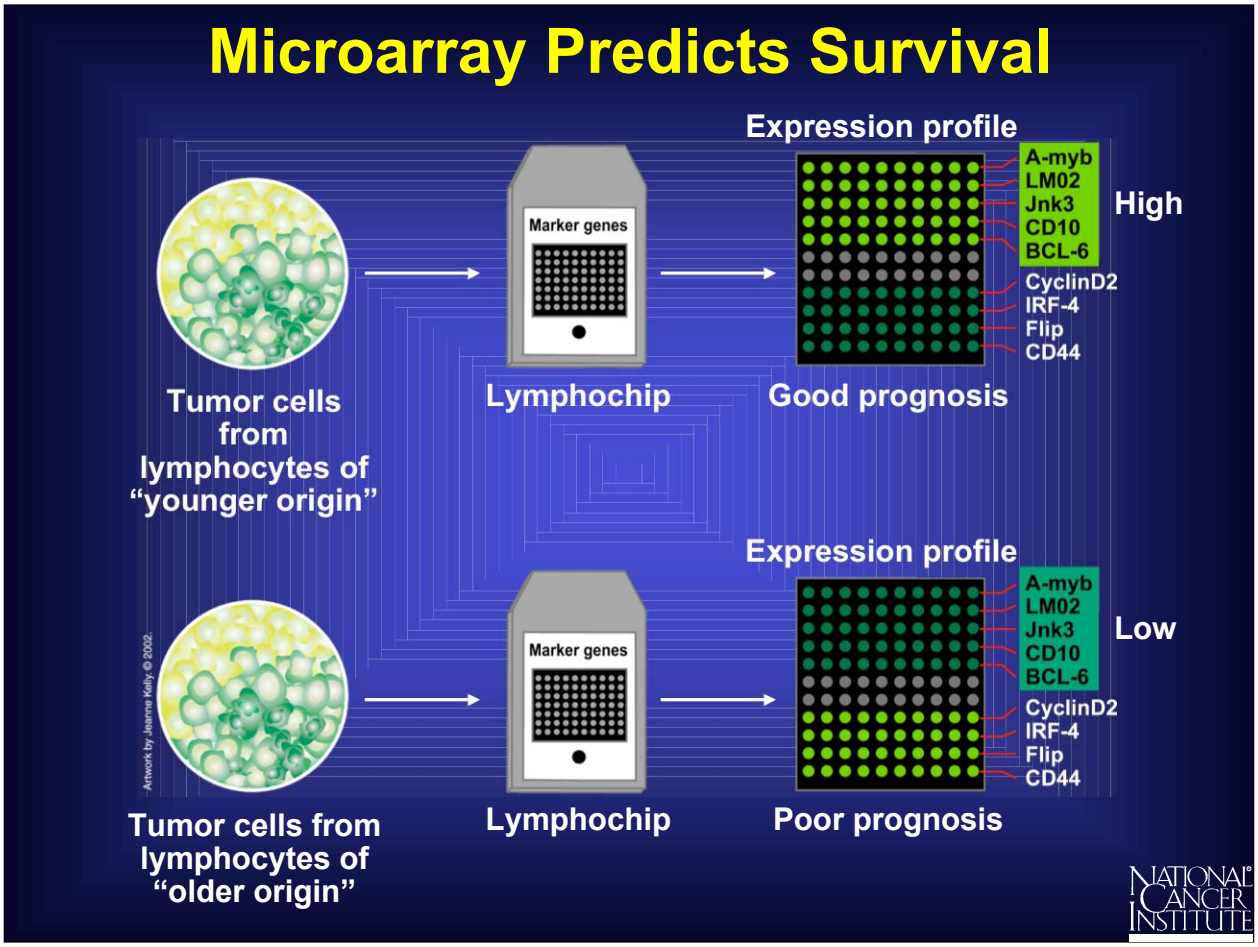
Researchers also use microarrays to detect differences in patterns of gene activity even within the same tumor type. In many cases, researchers have discovered that what had been considered a single type of cancer--based on how the cells looked under a microscope--was really two, three, or even more subtypes, each with a distinct gene expression pattern. For example, using microarrays, researchers have identified new types of leukemia and discovered that the most common type of lung cancer (lung adenocarcinoma) is actually four distinct types of cancer, each with its own gene expression pattern.



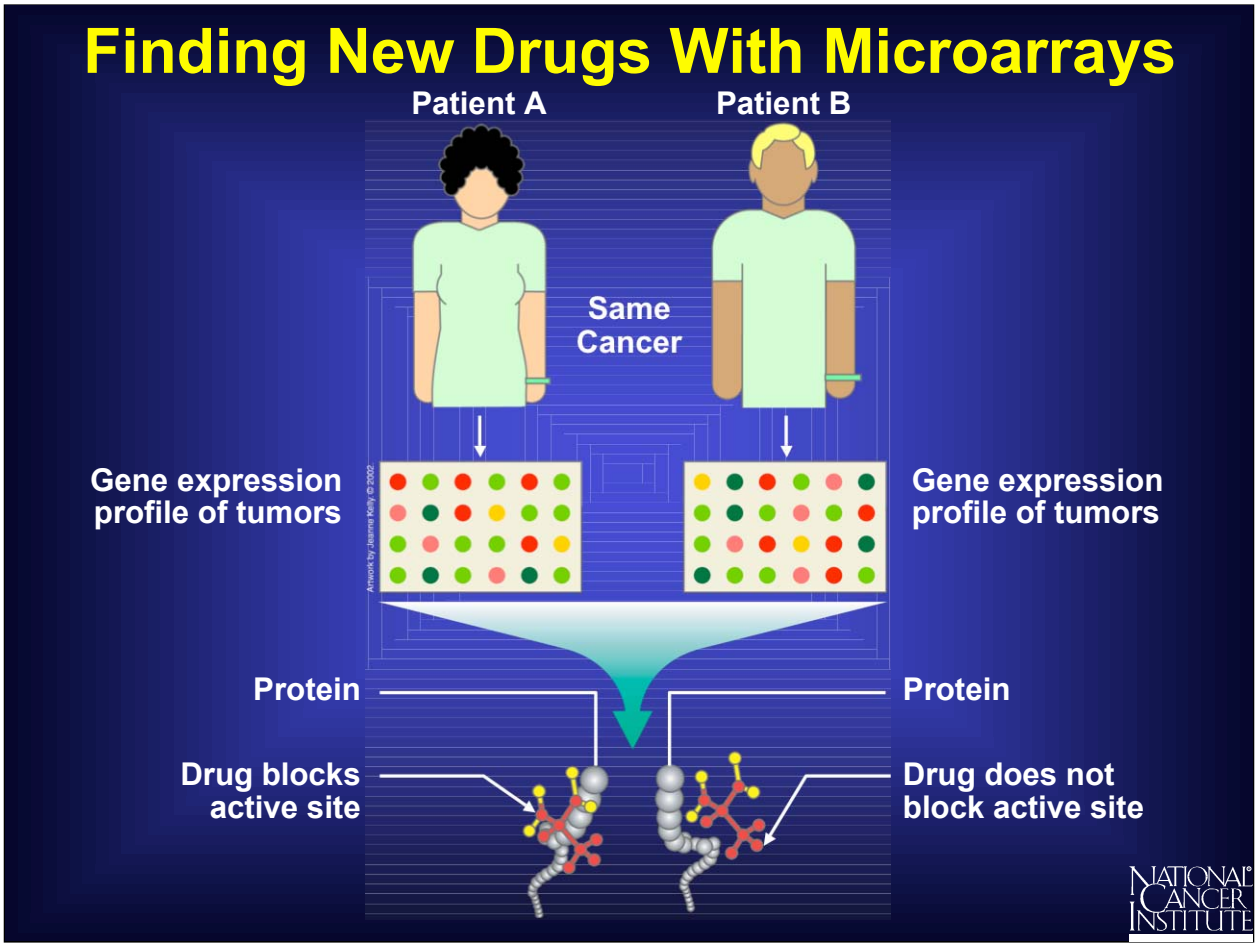
One of the first discoveries of cancer subtypes was in a blood cancer called diffuse large B cell lymphoma (the most common subtype of non-Hodgkin's lymphoma). Using a chip that contained fragments of 18,000 genes, researchers found two distinct cancer subtypes. These cancers looked the same under the microscope, but had different patterns of gene activity.

The subtypes were different in other ways, too. For one thing, they arose from different types of cells. Tumor cells of one cancer subtype arose from less differentiated ("younger") lymphocytes, while the other subtype arose from more differentiated ("older") lymphocytes.

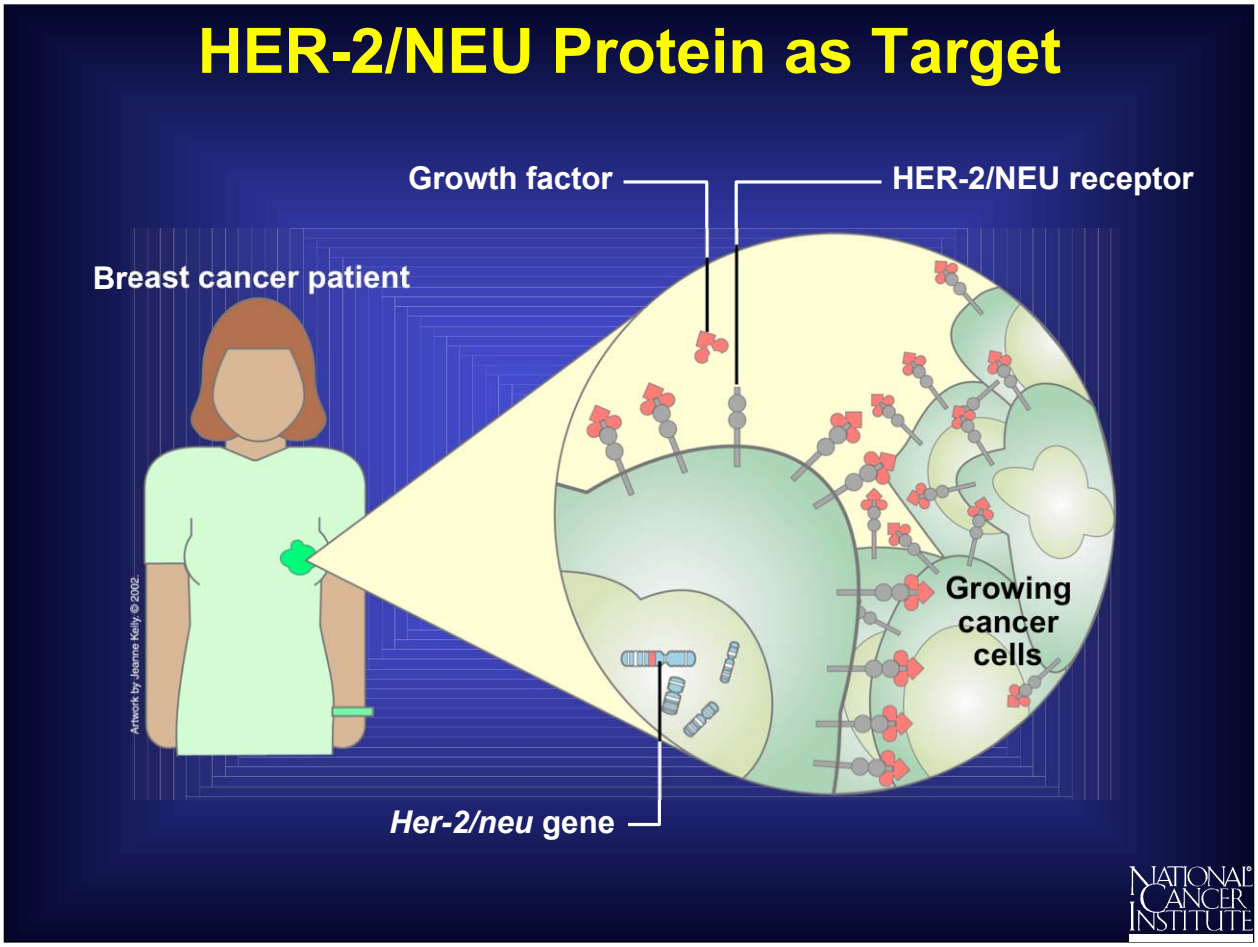
There was also a survival difference between the two types. About 75 percent of the people with the "younger" subtype responded to chemotherapy, compared with only 25 percent of the people with the "older" subtype.



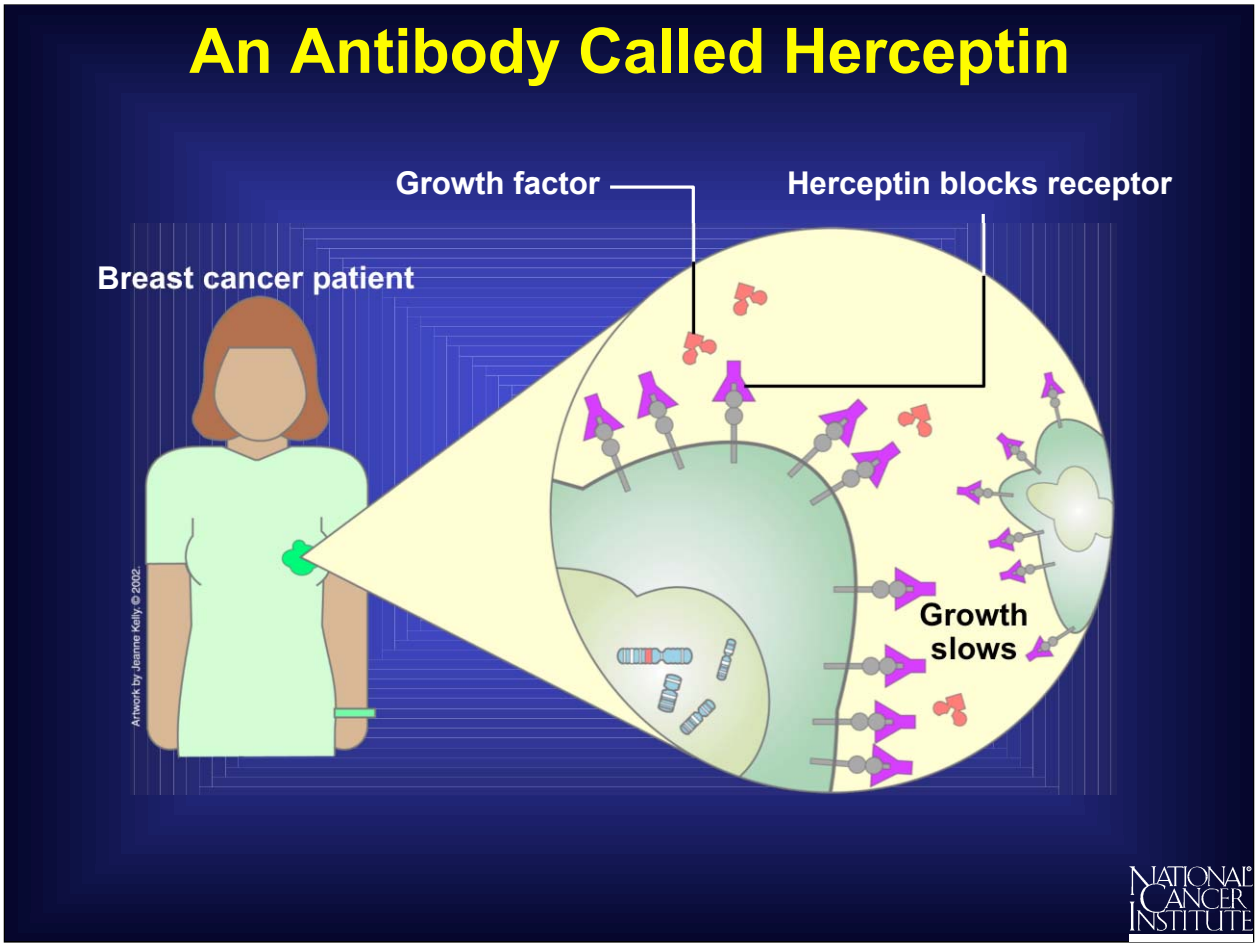
Two years later, another gene chip analysis yielded a third subtype of diffuse large B cell lymphoma. It also yielded another advance: Researchers found 17 genes that were strongly related to survival. The scientists created a formula to predict response to chemotherapy based on the expression patterns of these 17 genes. This formula divided B cell lymphoma patients into four groups. Two of the groups had about 72 percent survival rates 5 years after diagnosis. The third group had a 34 percent survival rate, and the fourth had a 15 percent survival rate. The 17-gene predictor formula was better than any current methods at identifying patients with the poorest prognosis. Identification like this is important because if patients predicted to have poor outcomes can be identified correctly early in treatment, they can receive more aggressive therapies right away instead of waiting until standard therapies fail.



Another important use of microarrays occurs in pharmacogenomics, an area of research that studies why certain drugs work in combination with particular genetic expression patterns, but not with others. This information is being used to design new drugs that target cancer cells without affecting normal ones.

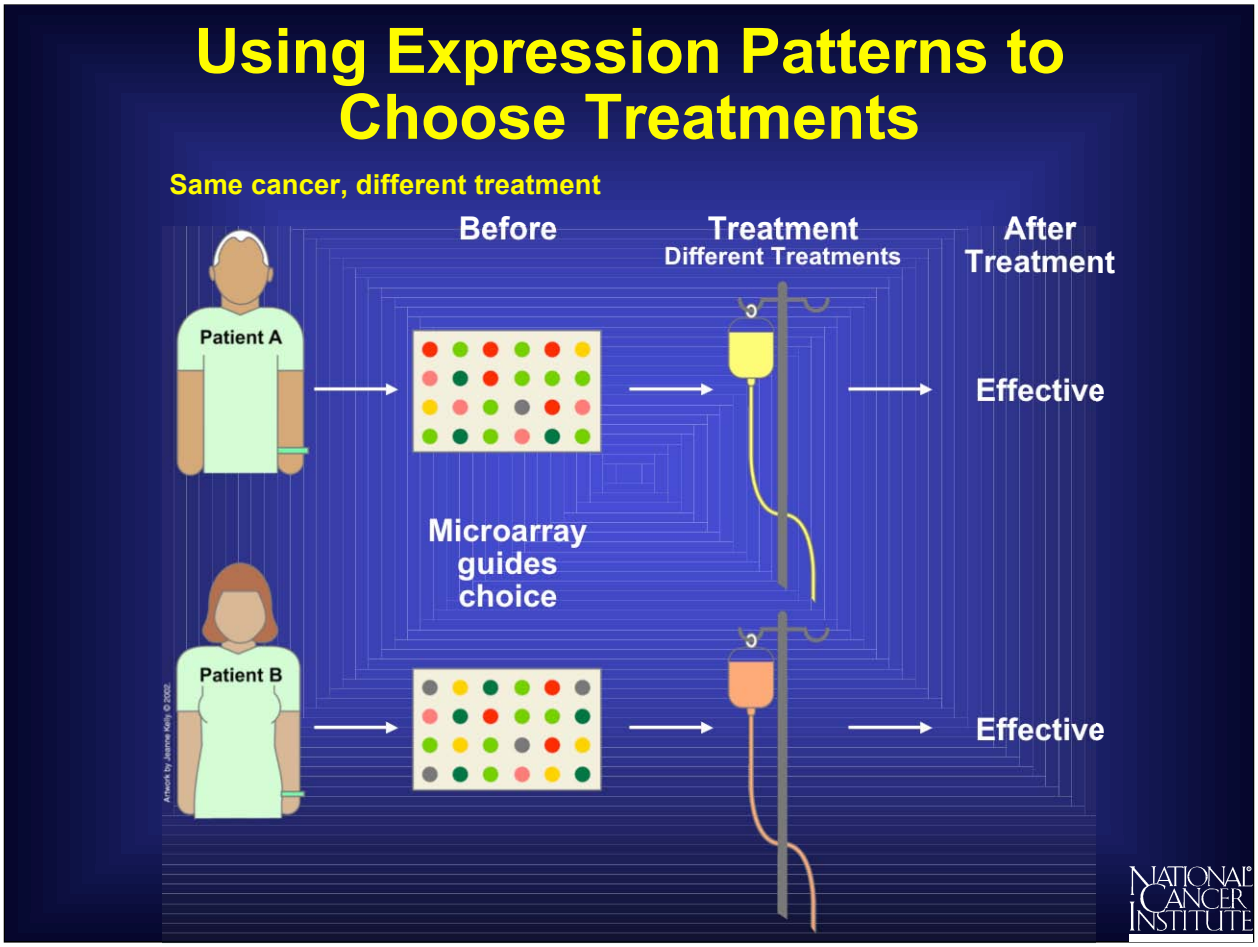


Herceptin is a good example of a pharmacogenomics success story. In the 1980s, researchers discovered that some women who had particularly fast-growing breast cancers expressed extra copies of a gene called *Her-2/neu*. The genes were producing many copies of a protein that appeared to be driving the growth of the cancer cells.



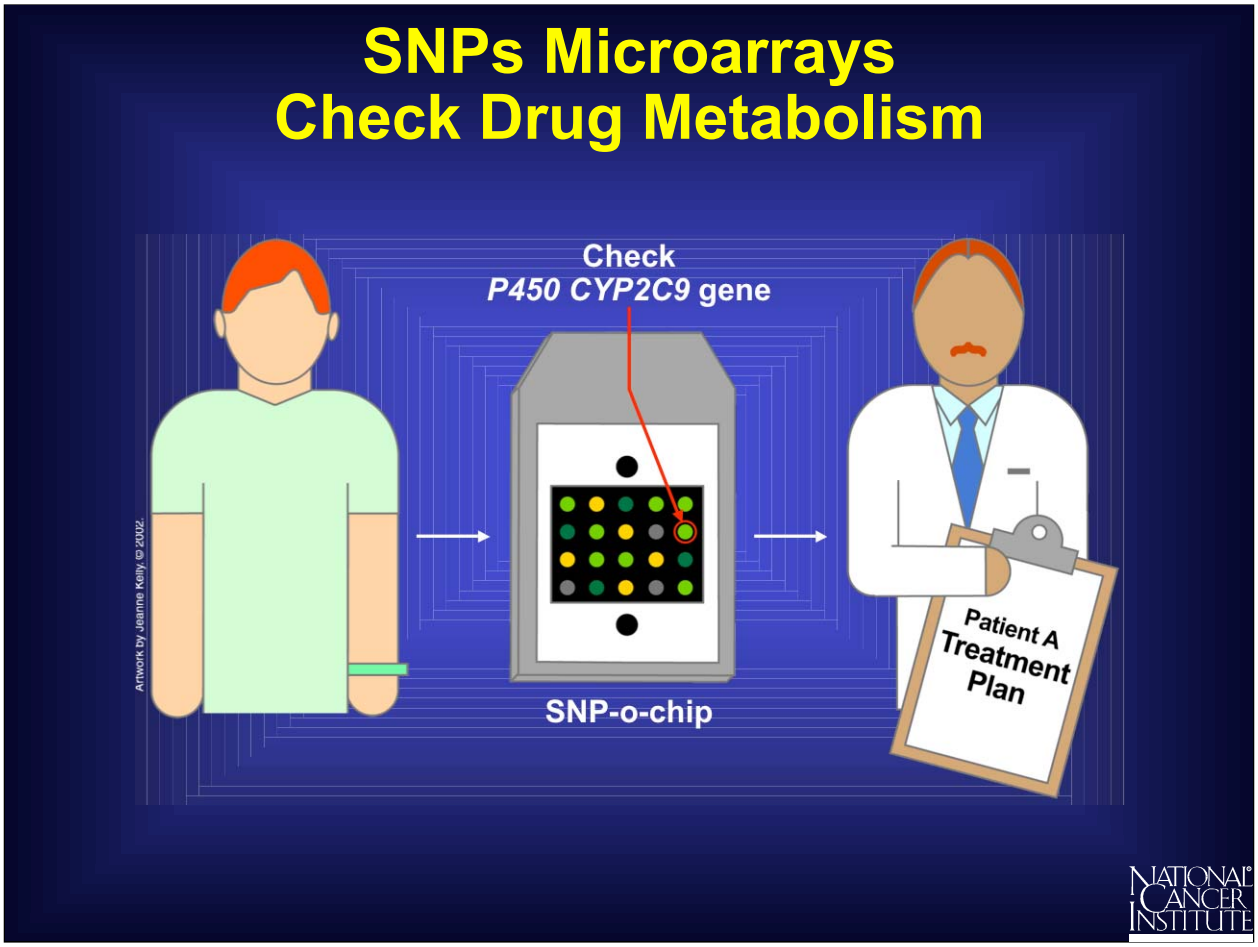
In the early 1990s, an antibody was developed that latches onto the her-2/neu proteins on the surface of a cancer cell. It stops the proteins from spurring on cancer-cell growth, and through this "stop signal" can also stop propagation of survival signals within the cancer cell. Some women who were given the experimental antibody saw their cancer growth slow or stop altogether when their Herceptin treatment was combined with cytotoxic chemotherapy.

Herceptin was approved by the Food and Drug Administration in late 1998 as therapy for women who test positive for high levels of the her-2/neu protein (about 25 to 30 percent of all breast cancer patients).

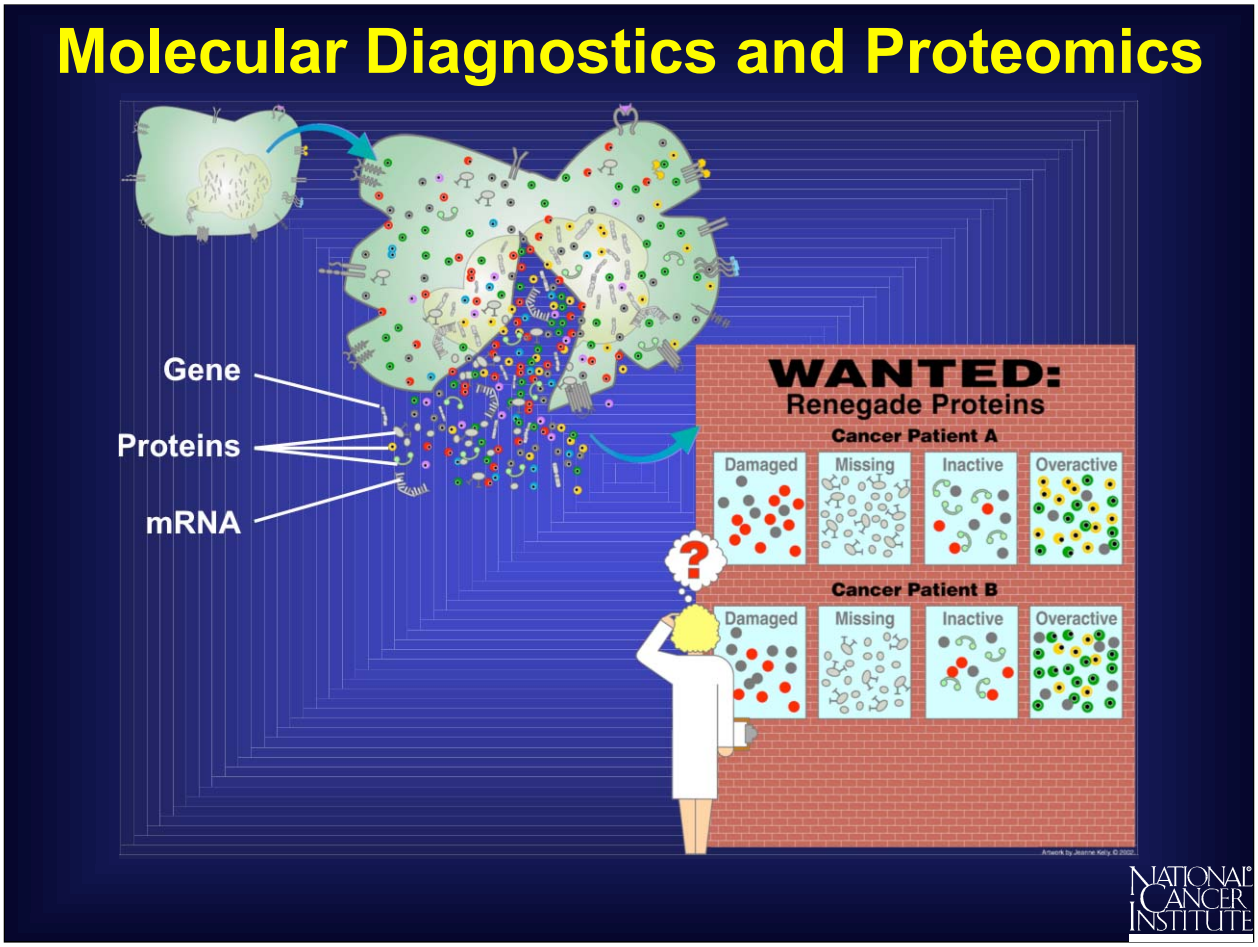


With the advent of microarrays, pharmacogenomics research has entered a whole new level of sophistication. This technology allows researchers to study patient populations, collecting expression patterns of tumors before any treatment is given.

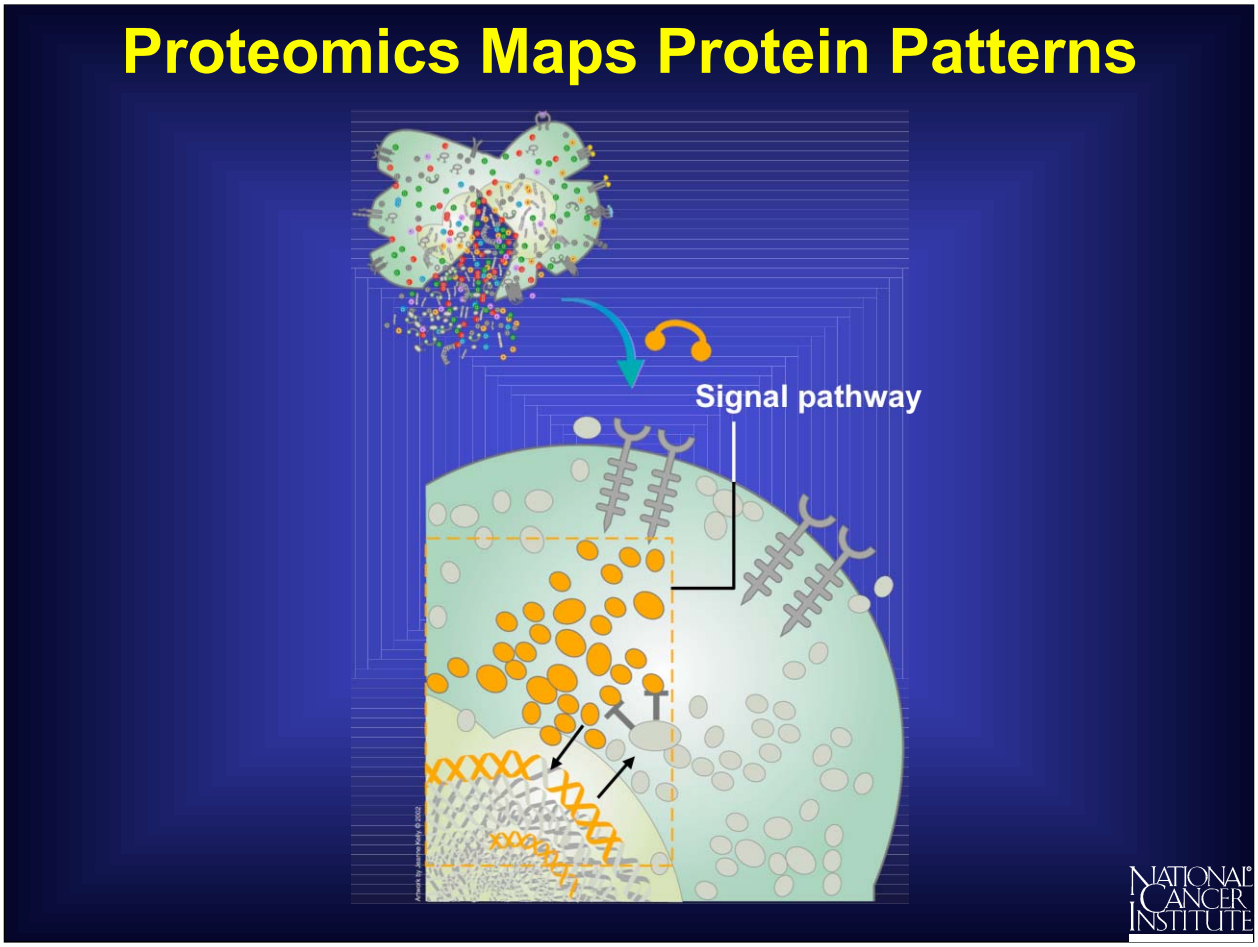
By analyzing treatment outcomes, scientists hope to identify expression patterns that can accurately predict a good response to treatment. In the future, *before* any treatment is given, doctors may be able to predict a patient's response to chemotherapy or radiation based on this research. This could eliminate much clinical "guesswork." One day, simply by looking at expression patterns found in a single blood sample or a small tissue biopsy, doctors may be able to tell a patient which of the available drugs or radiation treatments will work best and which will not. Clinicians will be better able to plan effective therapies at the proper doses for individual cancers.



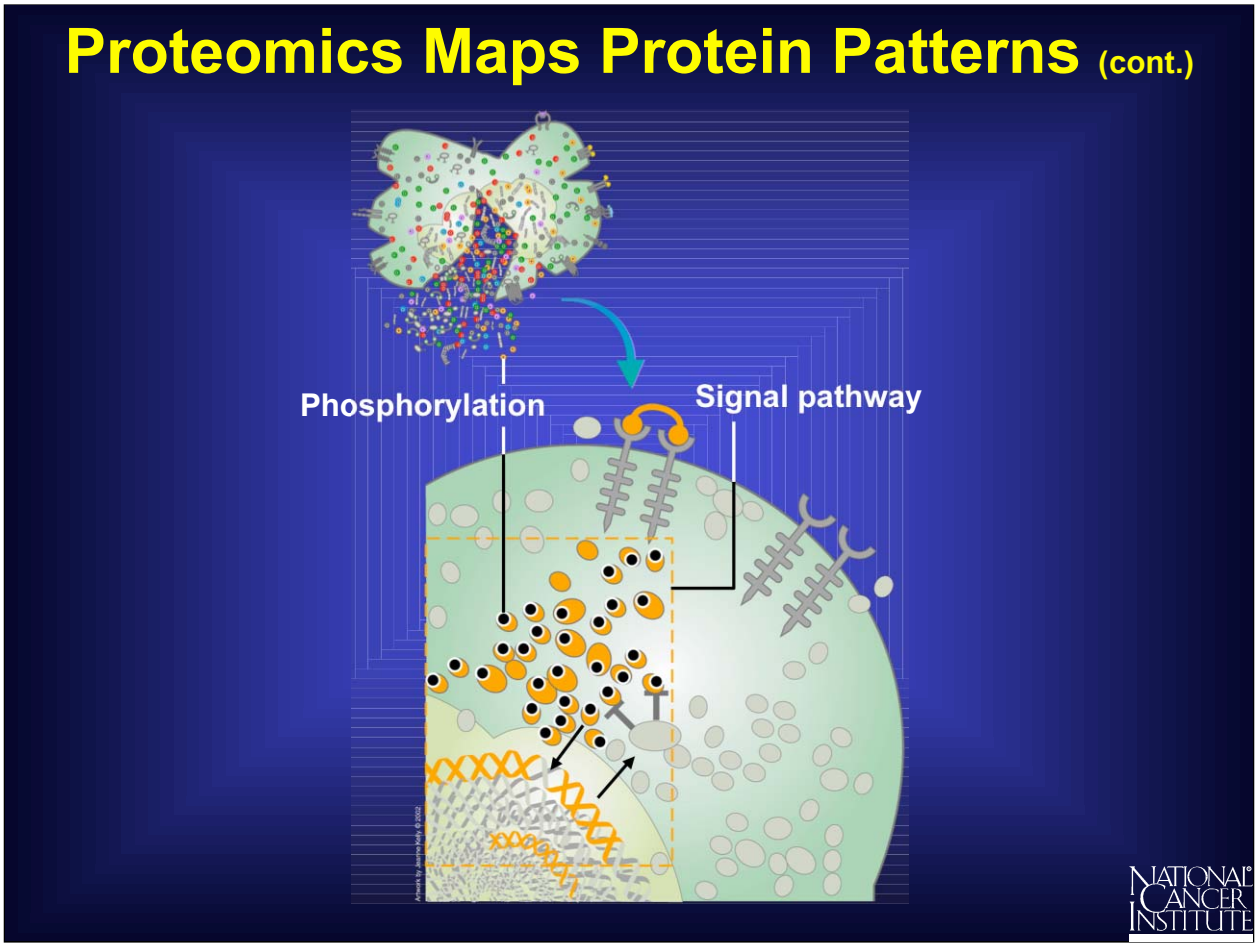
Microarrays will also be useful in studying drug metabolism. Drugs are broken down in the body by proteins, and not all of these proteins are identical from person to person. Genetic variations, called single nucleotide polymorphisms (SNPs), can result in subtle differences in proteins that translate to major differences in how the protein functions. For example, as many as 40 percent of people have a SNP that causes a deficiency in a protein called cytochrome P450 CYP2C9, which is important to drug metabolism. Deficiencies in this protein can impair the metabolism of as many as 1 in 5 medications currently on the market.



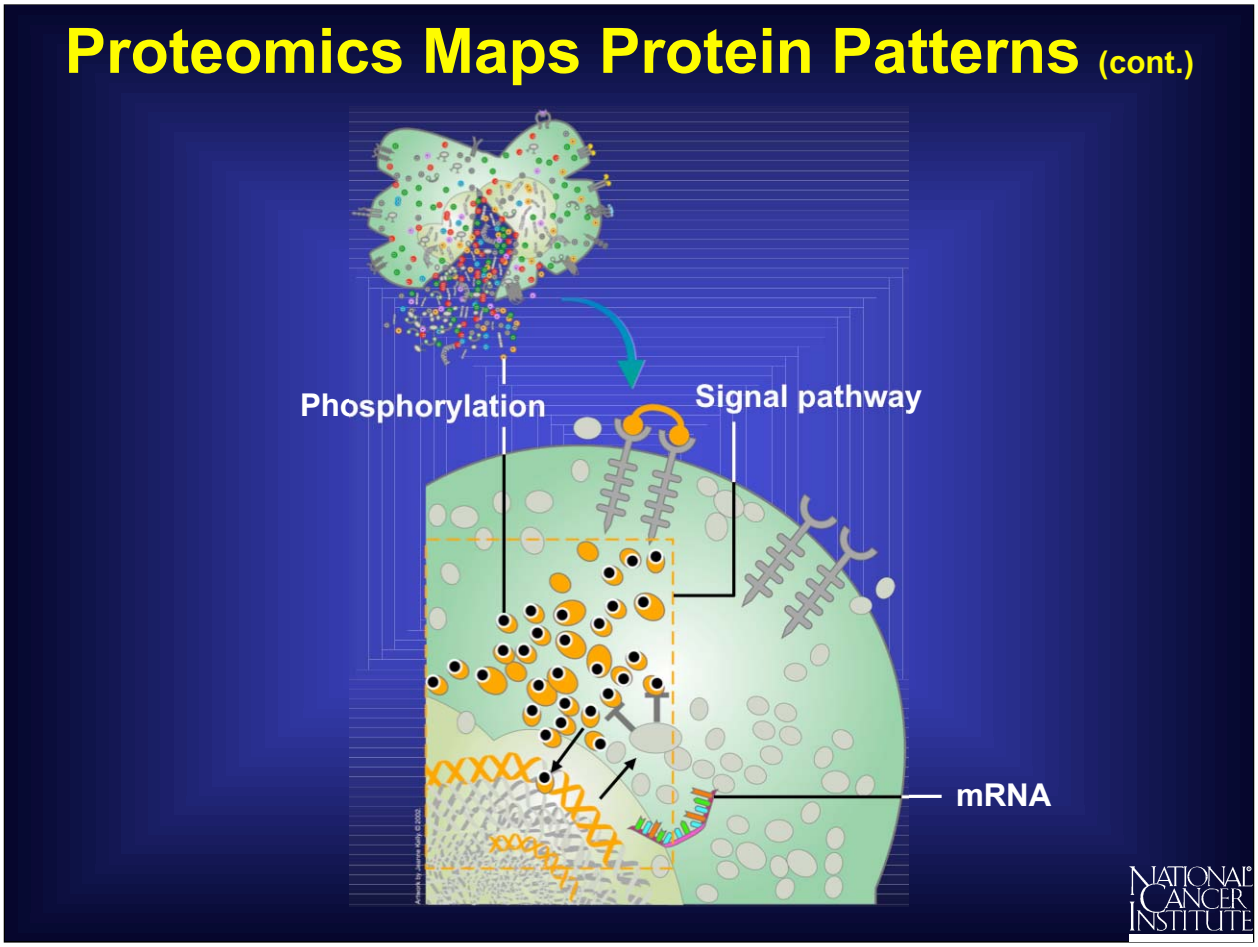
Another major focus of molecular diagnostics looks beyond genes to an investigation of the proteome. The term "proteome" refers to all the proteins produced by a cell, and the study of these proteins is proteomics. Cancer proteomics is enabling mapping of the patterns of proteins involved when normal cellular pathways are hijacked in support of malignant growth. In cancerous tissue, some of the proteins critical for normal communication are damaged, inactive, overactive, or missing entirely. The full set of deranged and dominating proteins at work disrupting cellular communications may vary from one cancer type to another. They may also vary somewhat from one patient to another within a cancer type.

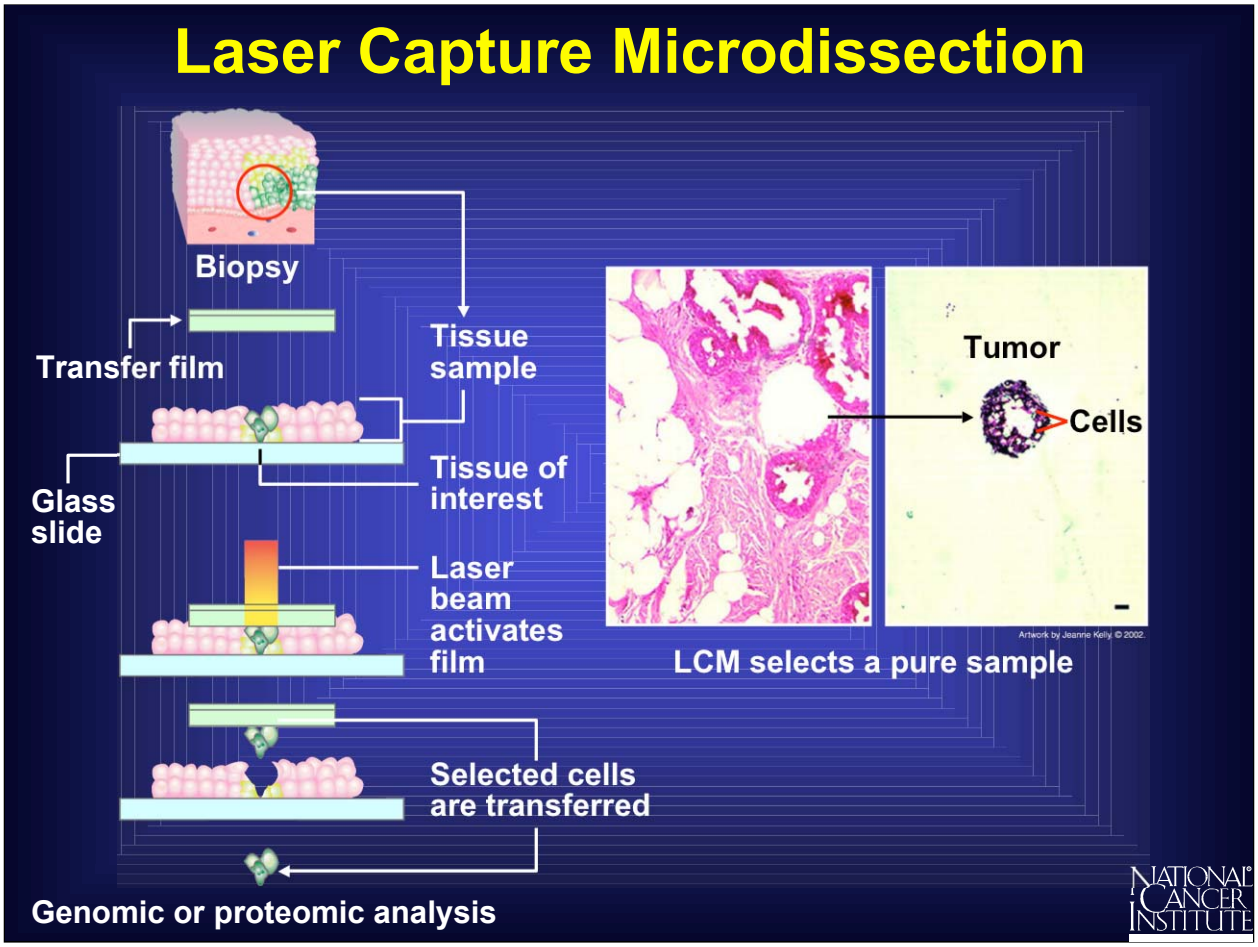


Cancer researchers are devising clever ways to capture data on each cancer's characteristic unruly sets of proteins. Biopsy or blood samples from cancer patients are the starting point. Many different methods are being used to extract and analyze the various sets of interactive proteins.



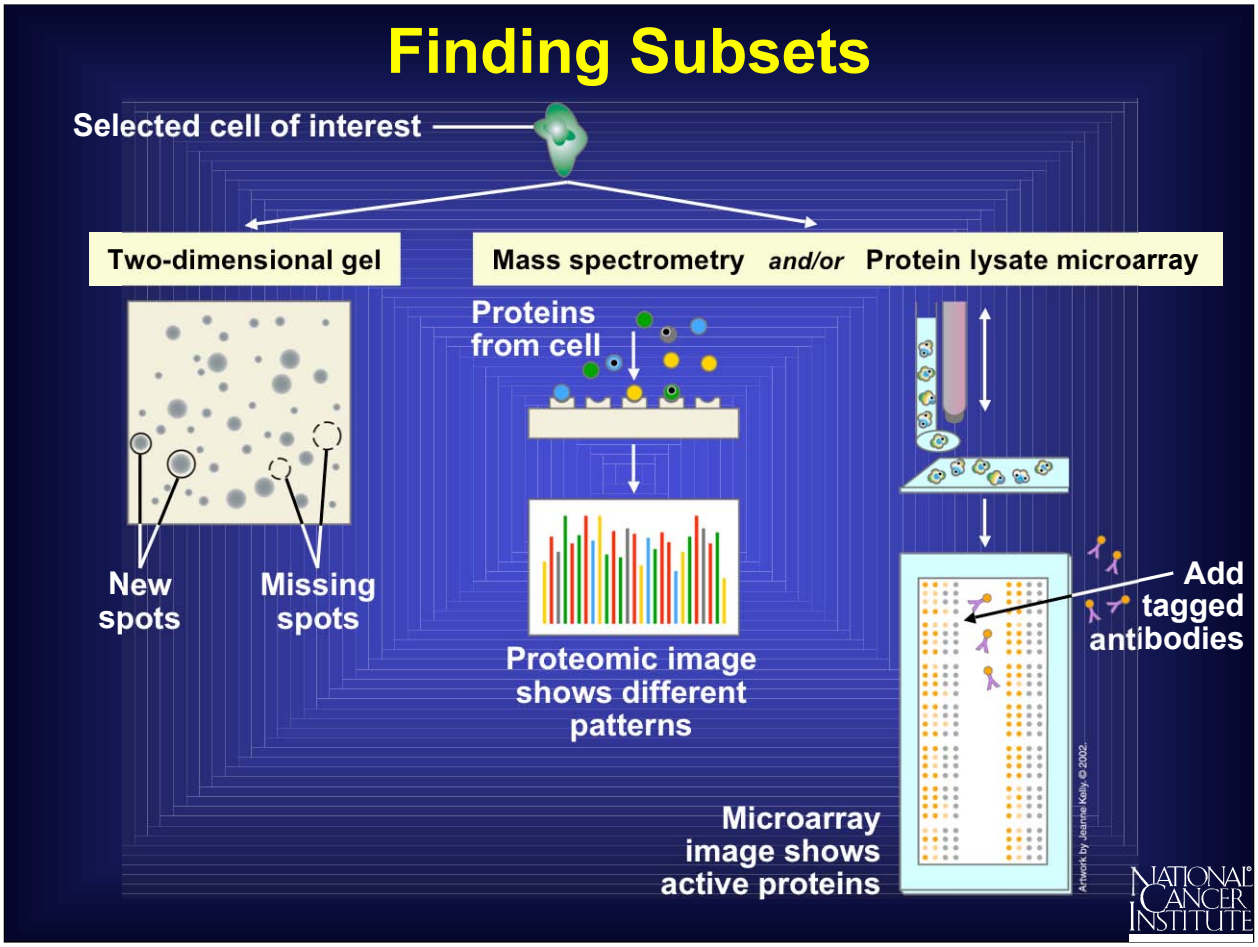
For many proteins in the cell, the addition of a small molecule called phosphate acts as a switch that activates the protein. This process is called phosphorylation. To preserve a cell's phosphorylation state and capture an accurate pattern of proteins interacting in a cancer cell, proteomics researchers must handle the biopsy or blood specimens carefully. They do this by seeing that samples are quickly treated with enzymes to block the removal of phosphates from proteins. This enables researchers to identify a protein pattern almost identical to what was in the cell at the time of collection.


Proteomics Maps Protein Patterns (cont.)
Phosphorylation
Signal pathway
mRNA
NATIONAL CANCER INSTITUTE

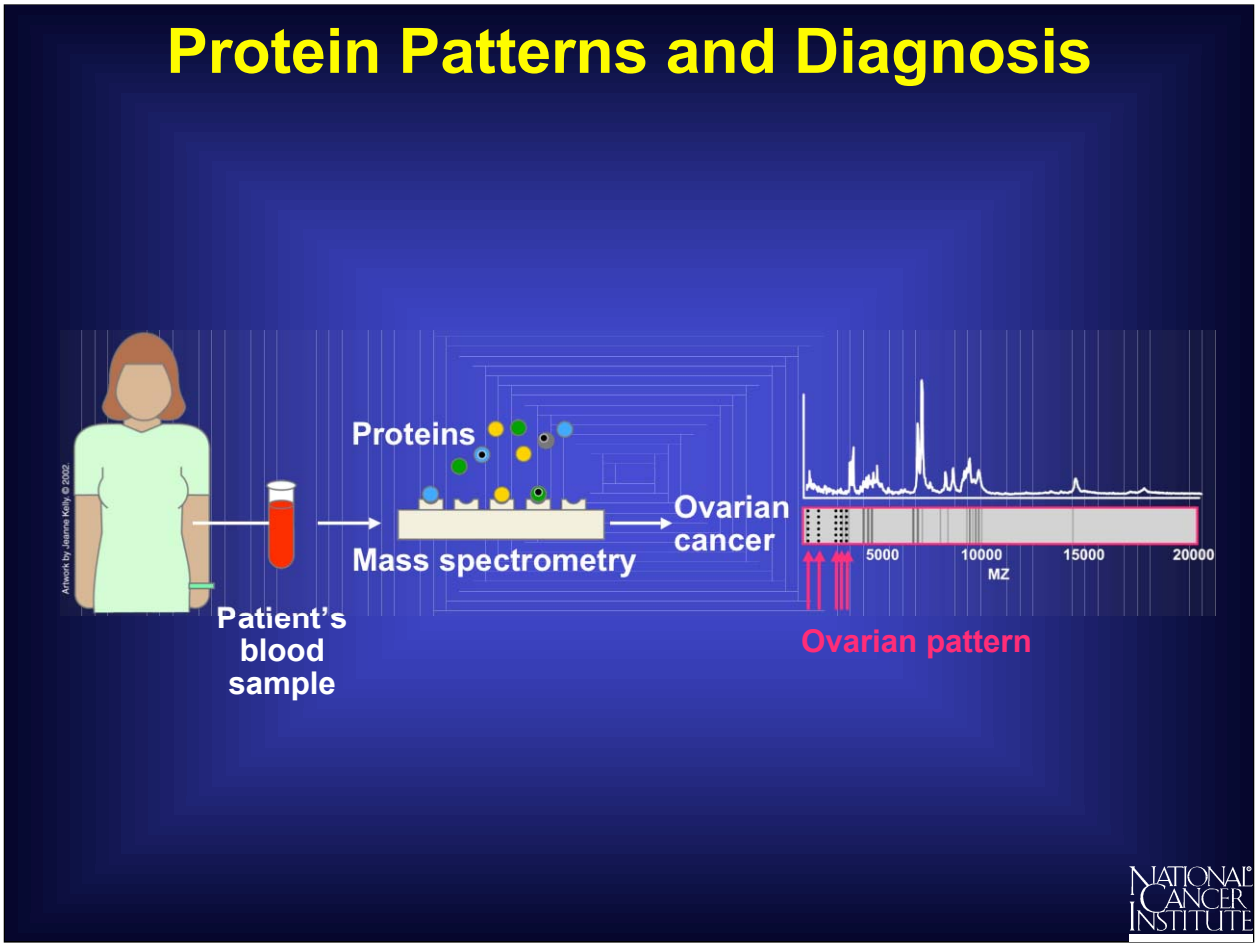


One method being used to minimize cellular damage and capture accurate protein patterns in cancer and normal cells is laser capture microdissection. Researchers use a low-energy laser beam and special transfer film to lift a desired cell out of the tissue section, leaving all unwanted cells behind. They can then collect all the proteins that were present in the selected cells, map the protein pattern, and store the information in a computer database.

Laser capture microdissection offers several advantages for studying proteins. For example, with this technique researchers are able to capture separate sets of cells from normal, precancerous, cancerous, and even stromal tissue, all from the same patient's biopsy sample.

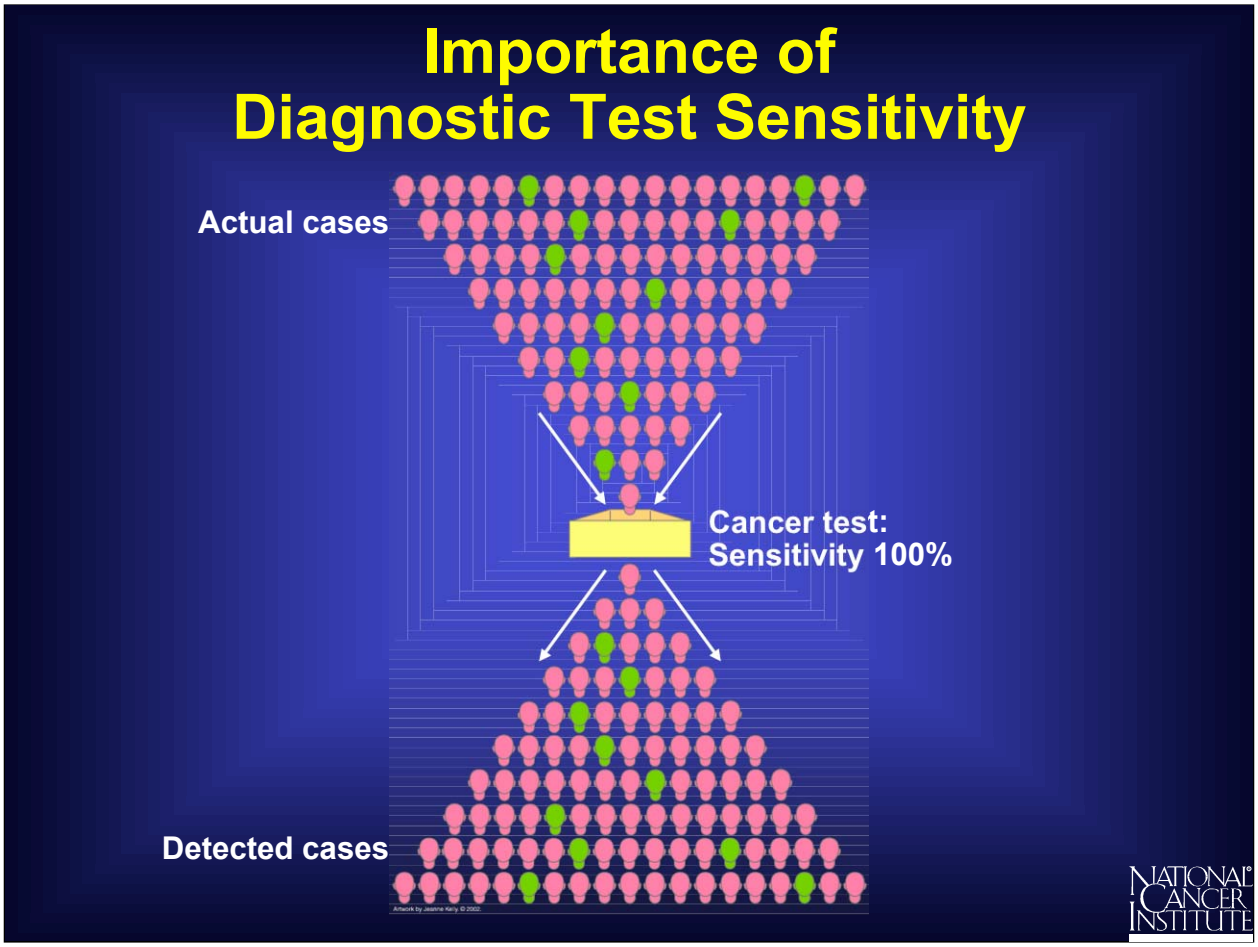


After laser capture microdissection, scientists can extract the proteins and use 2-D gels, protein lysate microarrays, and/or mass spectrometry to separate and identify subsets of proteins that may then serve as unique markers for a particular cancer cell type.

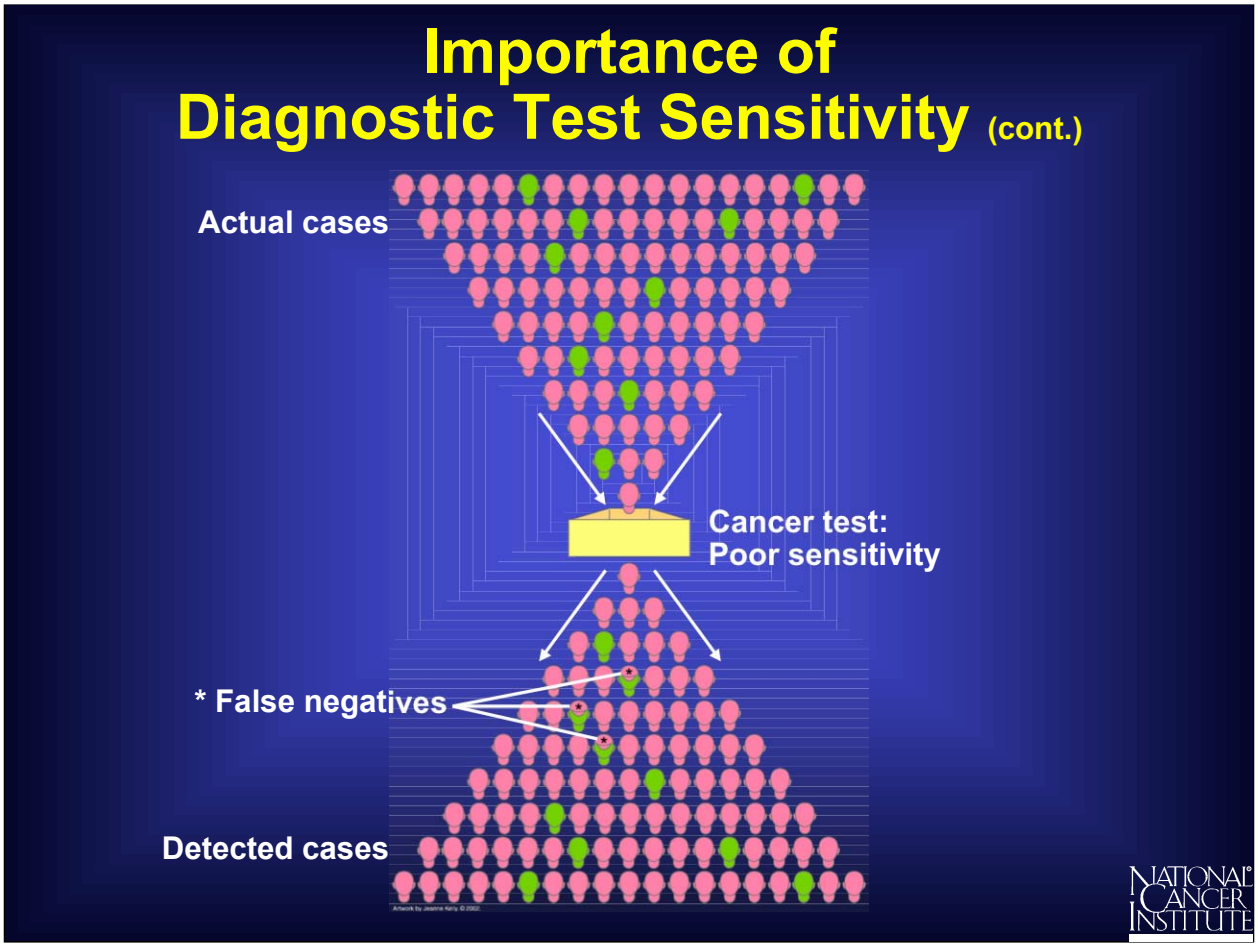


Scientists have learned that it is not necessary to ***identify*** every protein active in cancer and captured by laser microdissection. To accomplish a molecular diagnosis, all that may be needed is to separate and preserve a unique subset, a ***pattern of proteins*** shared by all patients with the same cancer type.

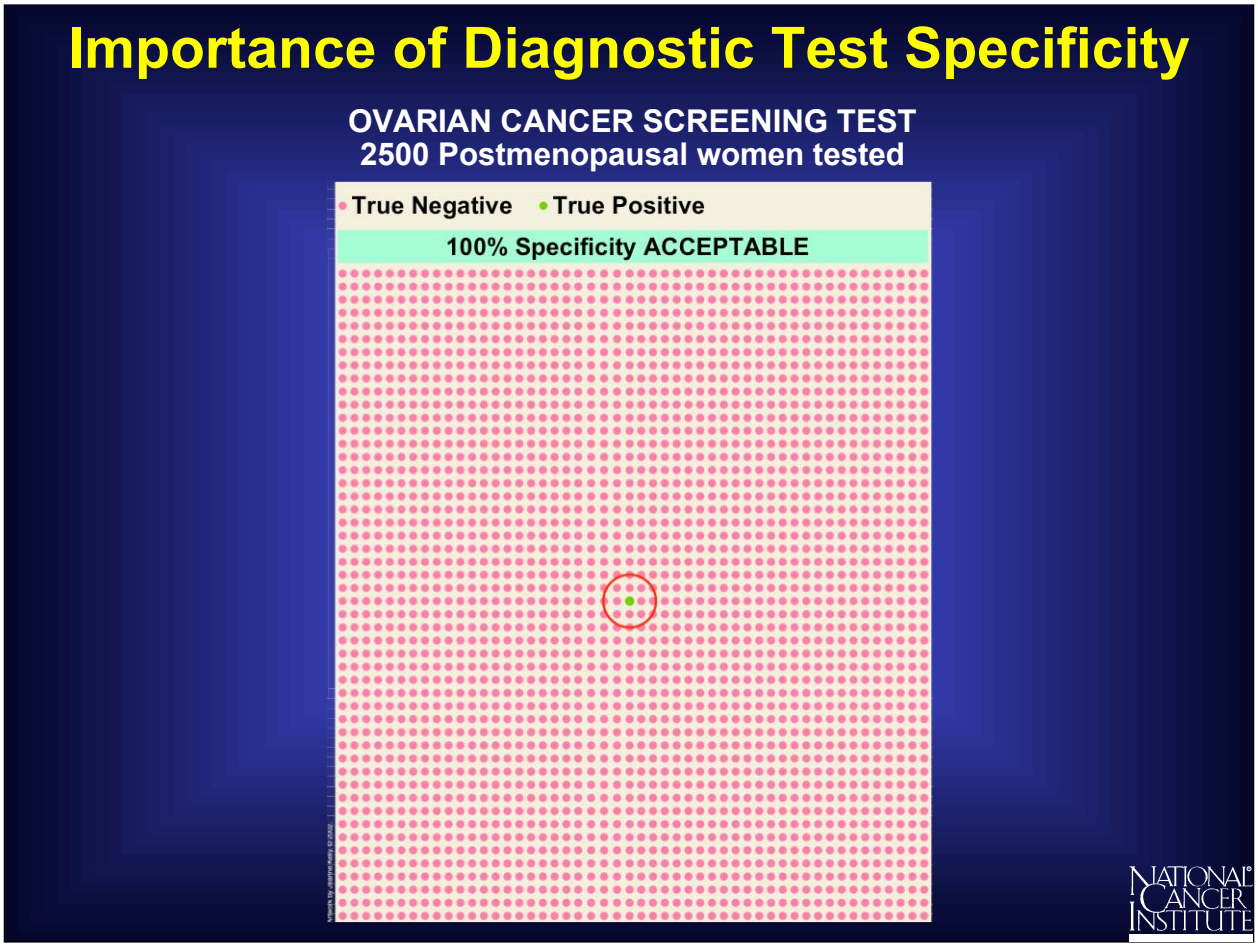
For example, researchers collected blood samples from a group of patients with diagnosed ovarian cancer. They used mass spectrometry to collect all the protein profiles that appeared in the patients' serum samples. While there was some variation from patient to patient, pooling enough samples of confirmed ovarian cancer enabled clinicians to use cluster analysis to determine what subset of proteins consistently served as markers for the presence of cancer. Identification of this subset was successful, and today we are on our way to completing a method of molecular diagnosis for ovarian cancer.



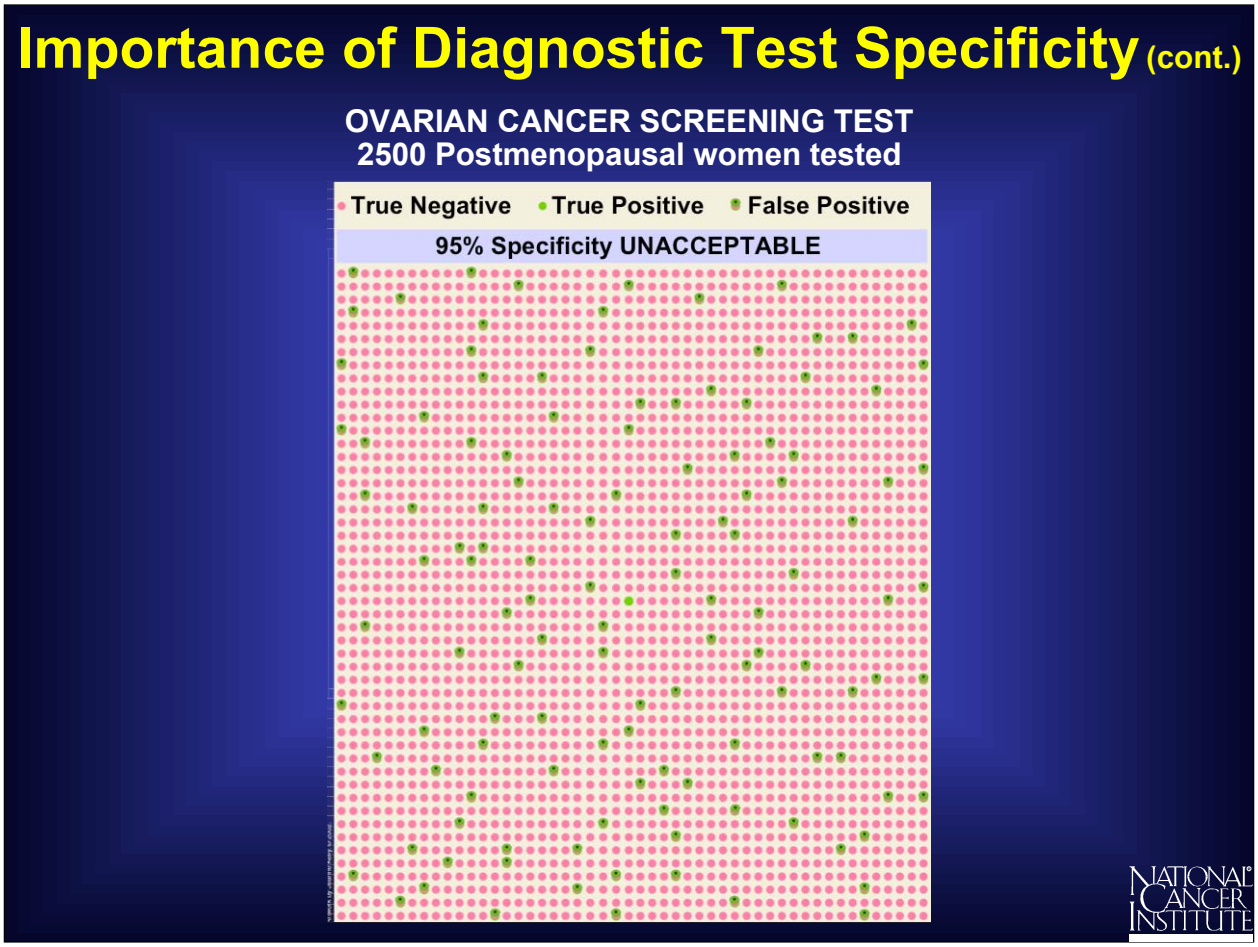
Molecular diagnostics for ovarian cancer must correctly detect all the women who have this cancer without incorrectly identifying people who do not have the disease.



The first hurdle is a test's sensitivity. Sensitivity answers this question: What percentage of people with a given type of cancer will have their cancer detected when a given screening test is used? If a test is not very sensitive, there will be many "false negative" results, and persons with cancer will go undetected. A test that yields many false negatives will obviously not be very good at reducing cancer deaths, giving a false sense of security to people who actually have the disease.

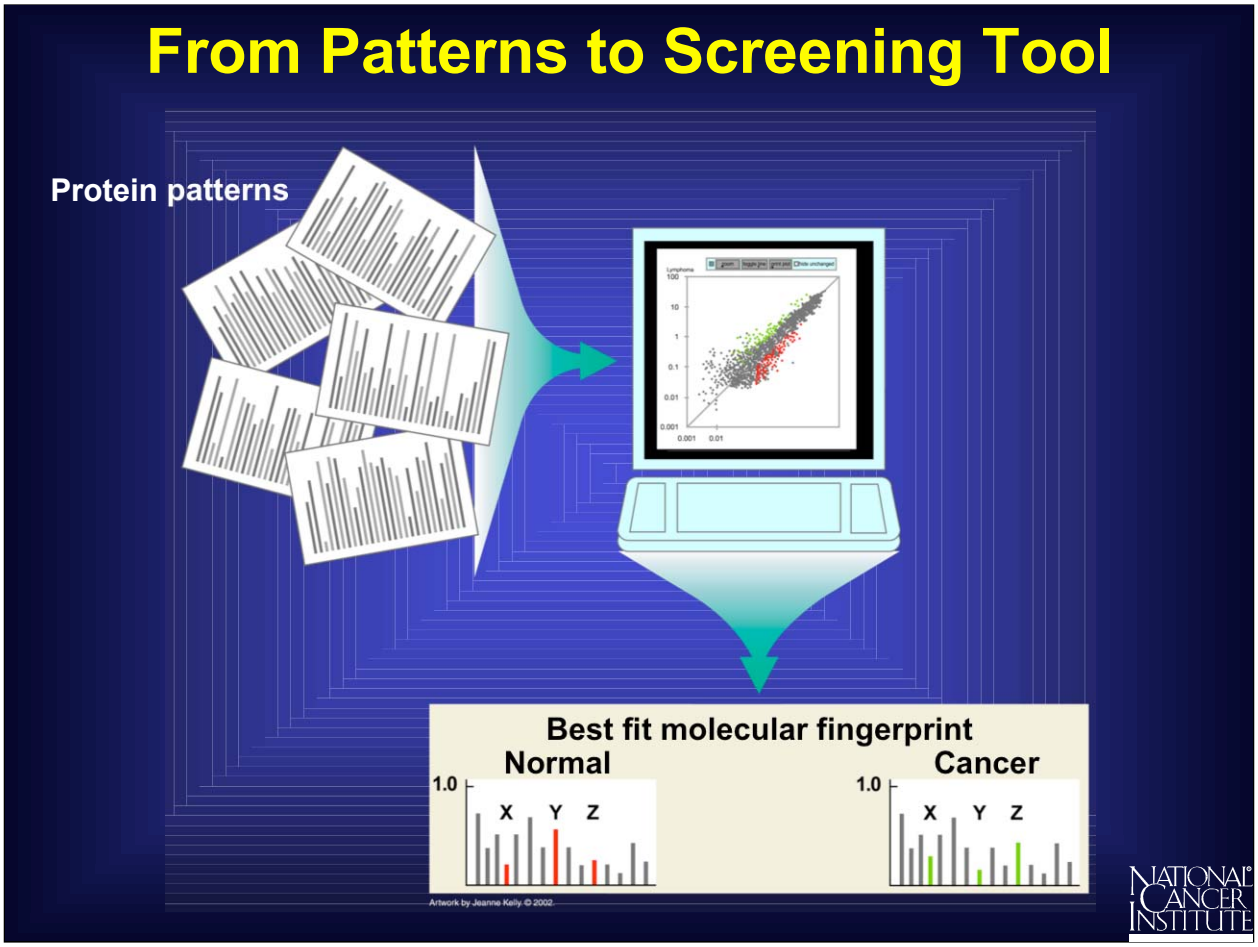


The next hurdle is a test's specificity. Specificity answers this question: What percentage of the people who do not have cancer are correctly identified as being free of disease? If a cancer test is not very specific, it will yield many "false positive" results; that is, a person will test positive even though they are cancer free. Such error can lead to unnecessary and costly follow-up procedures and cause anxiety in the person misdiagnosed.

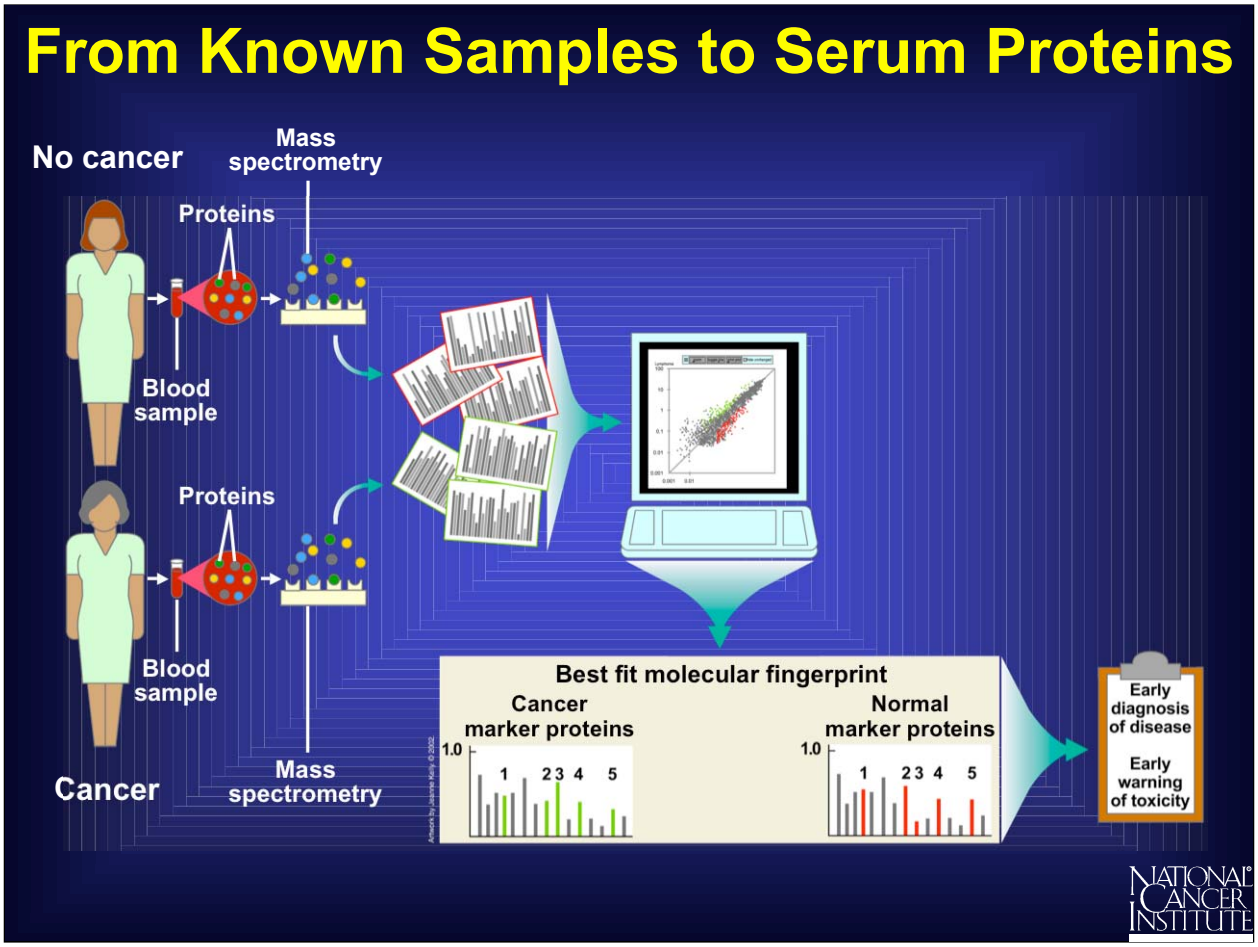


High specificity is important for an ovarian cancer test because this cancer is a relatively rare disease, occurring only in about 1 of 2500 postmenopausal women.

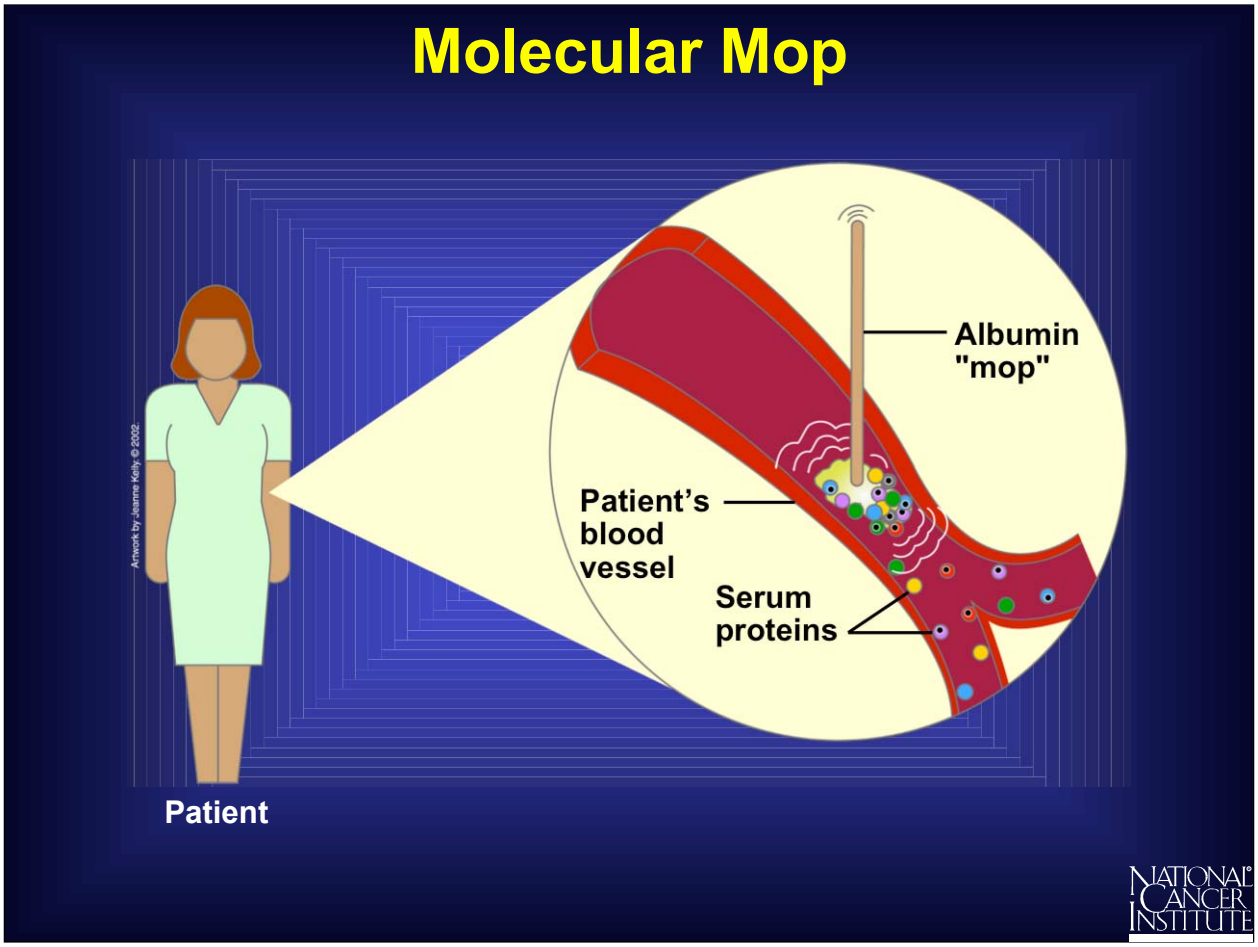
The ovarian cancer diagnostic test is being developed and tested first among women at risk of **recurrence** for ovarian cancer. If this use is successful, the test will next be extended to women at high risk for developing the disease.



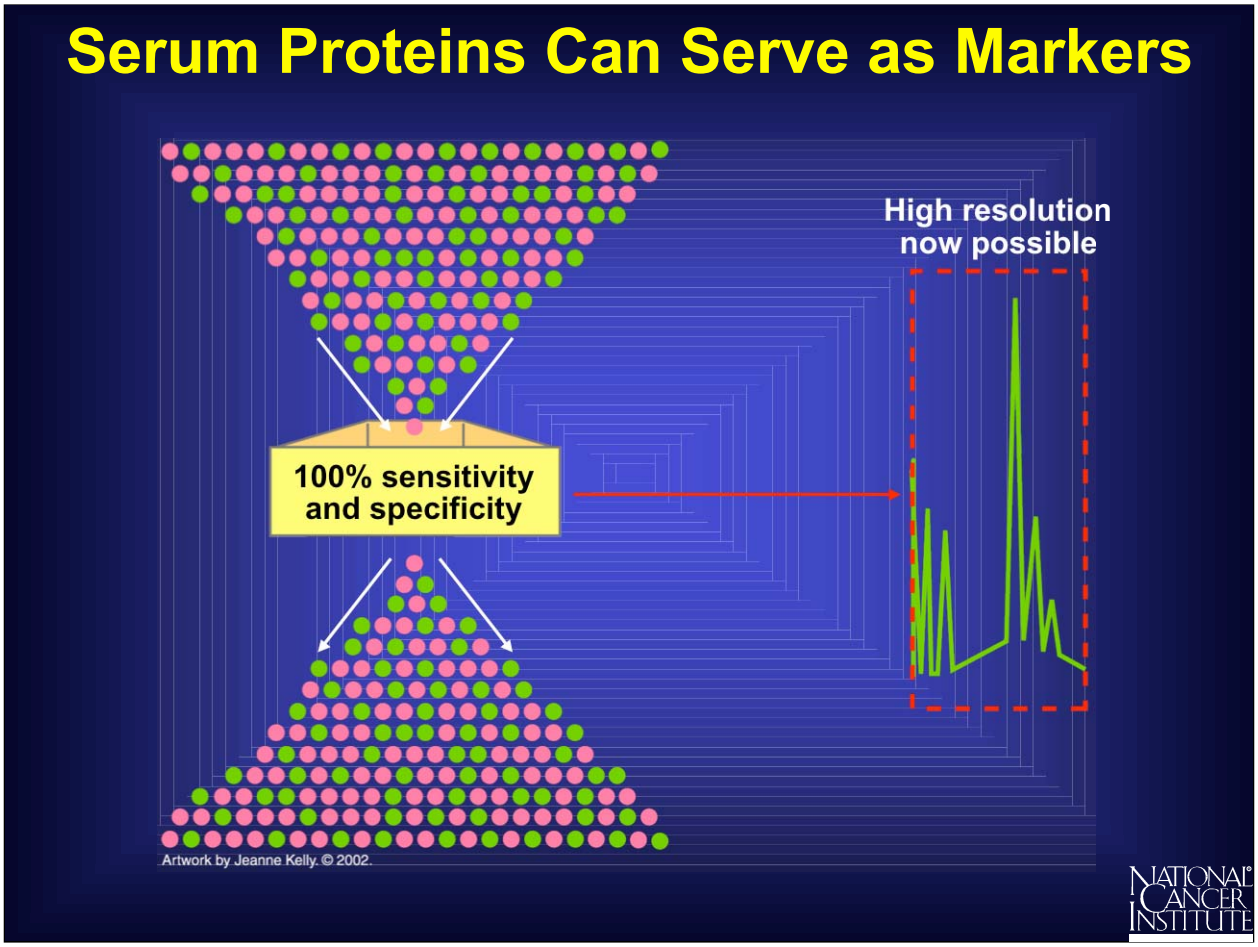
Researchers used the protein patterns they had collected from the ovarian cancer patients along with an artificial intelligence computer program to create and train a new bioinformatics-based screening tool. The program was designed to recognize the presence or absence of ovarian cancer in blood samples of unknown cancer status.



Scientists next asked if a new bioinformatics-based screening tool for ovarian cancer could be created using low-molecular-weight proteins in blood serum as the biomarkers. Could this tool accurately handle specimens of unknown cancer status? If so, researchers could devise a minimally invasive test for the detection of early-stage ovarian cancer. If so, their bioinformatics-based technology might also have broad application to the early diagnosis of many other cancers as well.

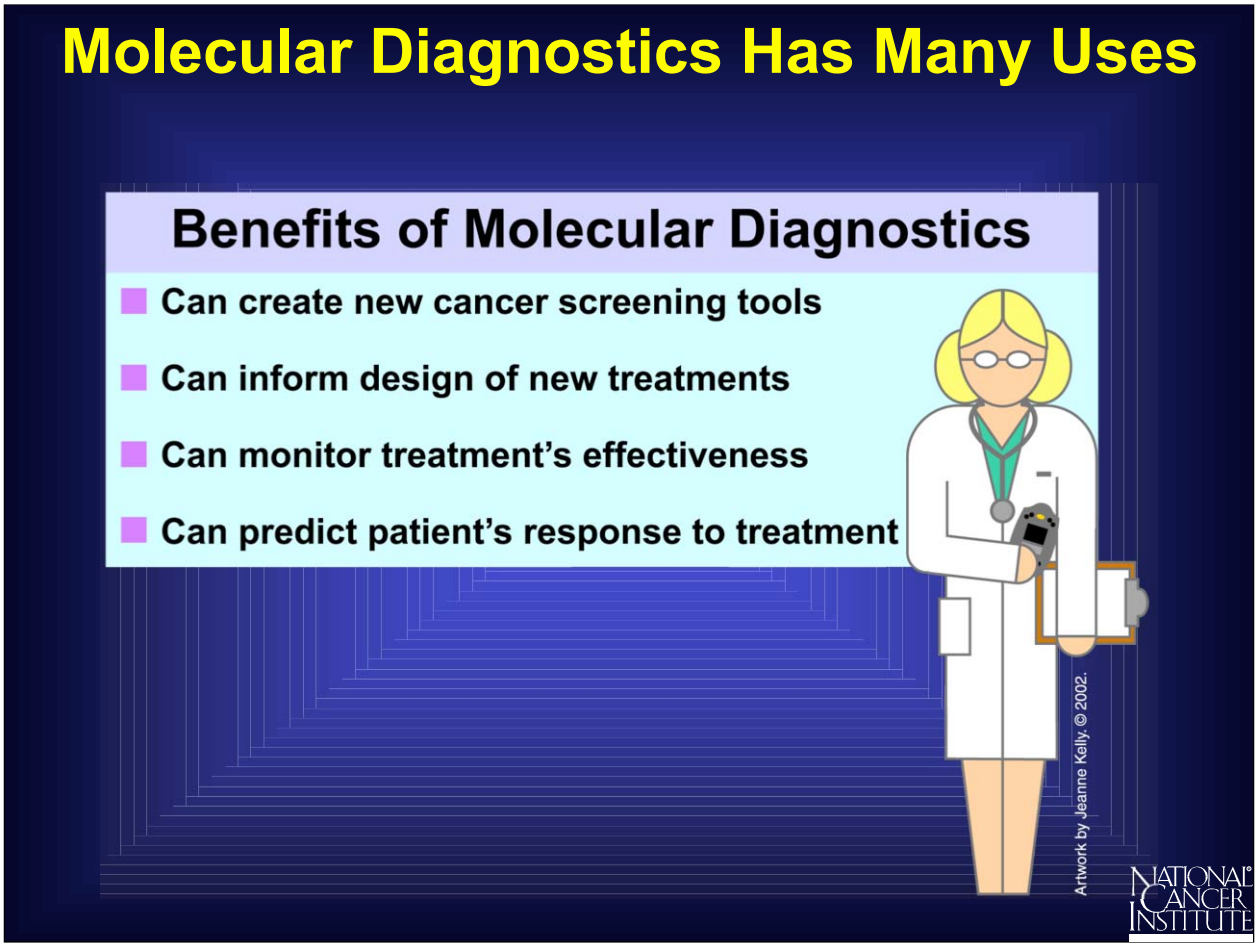


A major problem in the identification of cancer biomarkers is the very low concentrations of markers coming from tissues with small, early-stage cancer lesions. Fortunately, investigators have found a new way to amplify and concentrate these biomarkers in the blood. When scientists searched for protein patterns in the serum of ovarian cancer patients, they came upon an interesting discovery. The albumin molecule and other long-lived carrier proteins that circulate in the bloodstream act as molecular mops, grabbing a lot of low-molecular-weight proteins as they are degraded and prepared for elimination from the blood. The "mop" actually helped them collect and amplify, more than 100-fold, the low-abundance serum protein patterns needed for analysis.

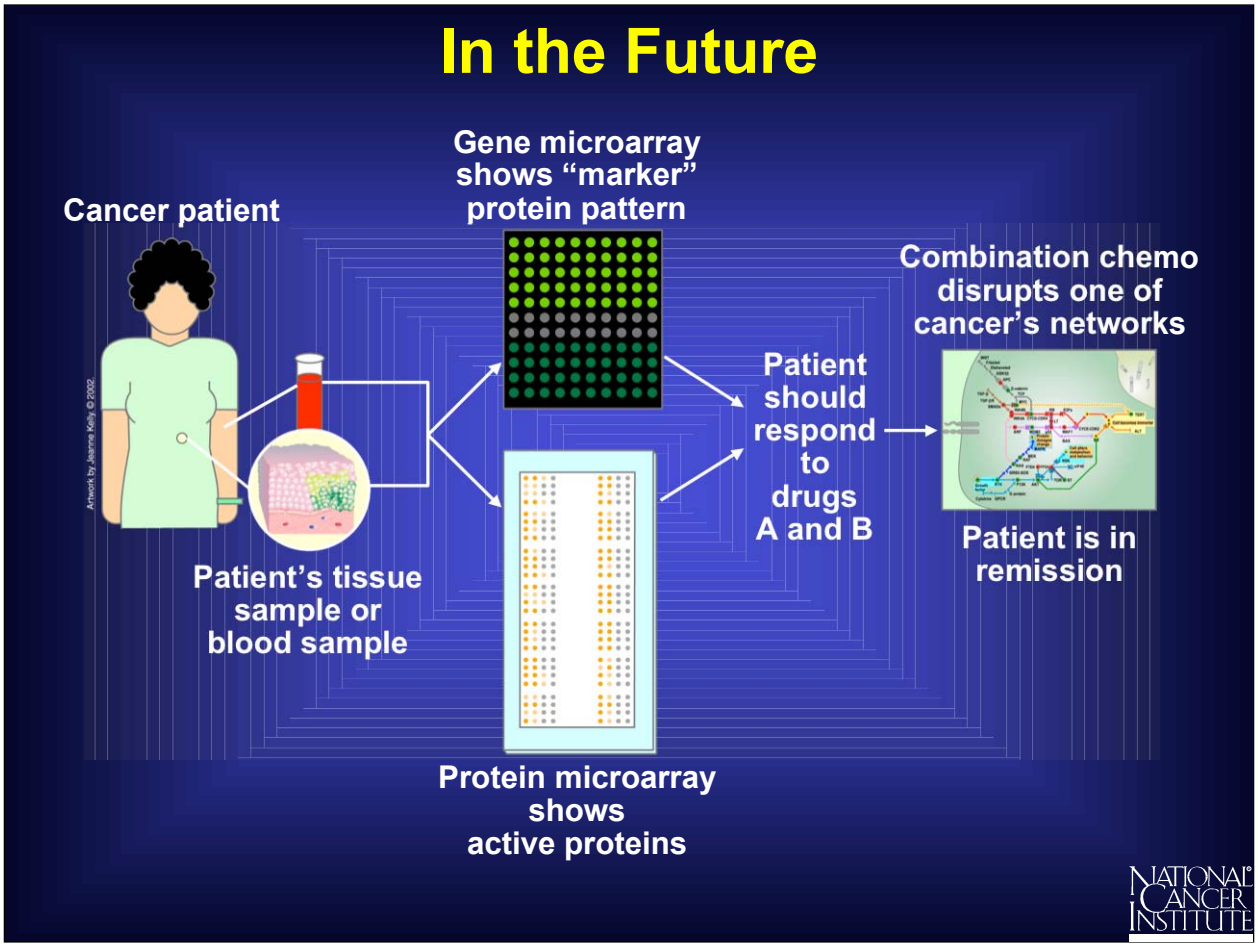


The researchers succeeded in showing that a subset of low-molecular-weight proteins from the serum of patients formed predictive patterns that could be used experimentally to mark the presence of ovarian cancer.

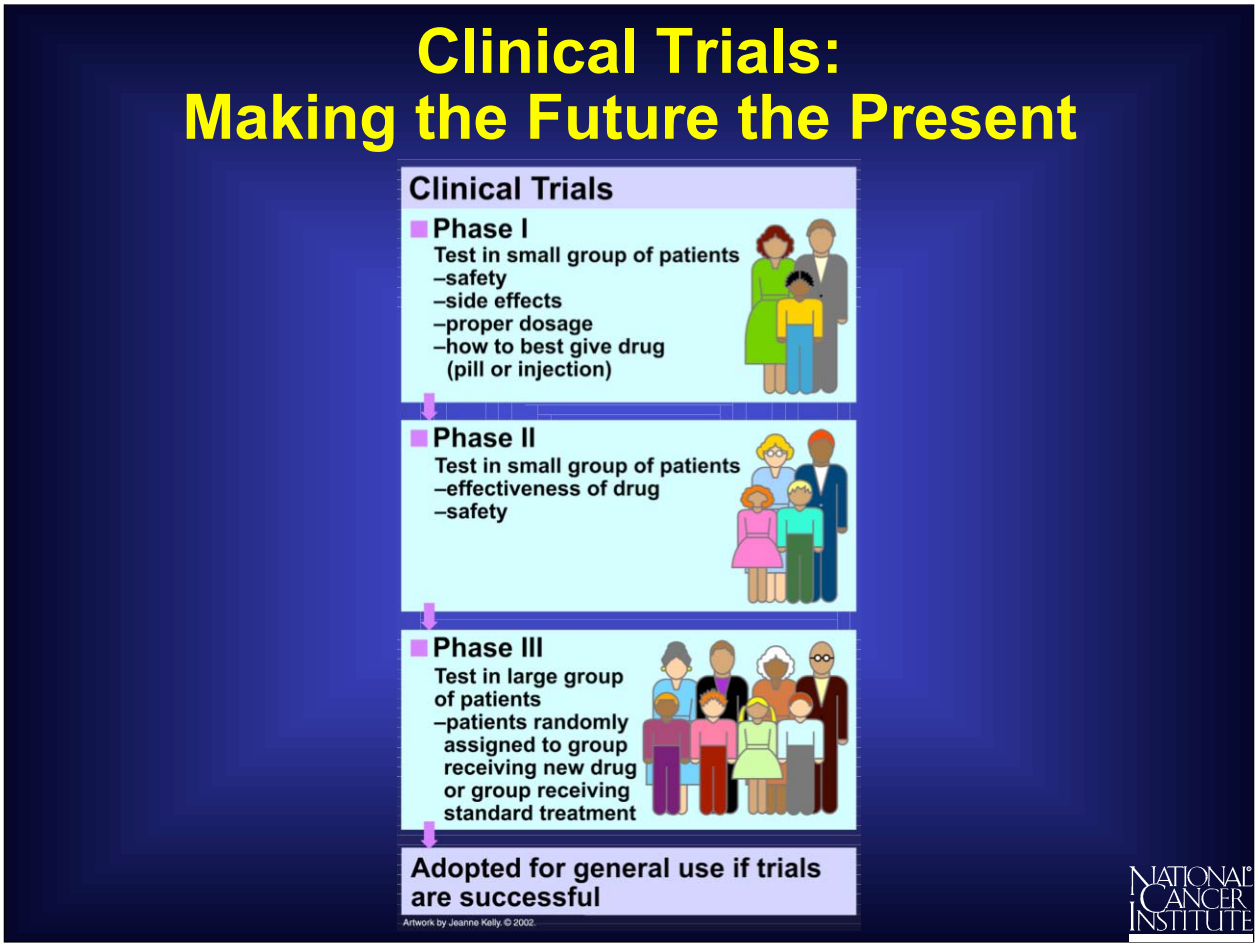
They detected a diagnostic pattern of five proteins in the blood of women with ovarian cancer that is not found in the blood of other women. Experimentally, in a limited subset of test patients, the test has had a sensitivity of 100 percent and a specificity of 100 percent.



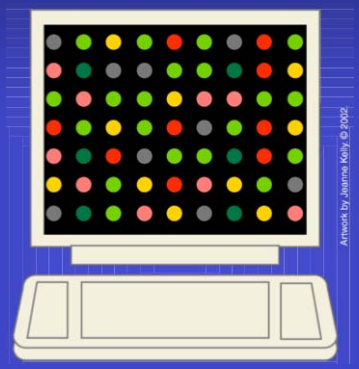
Clinicians have discovered that molecular diagnostics has many uses beyond the creation of new screening and diagnostic tools. Expression patterns also can provide information for the design of new cancer treatments, monitor the treatment's effectiveness as it is studied in a clinical trial, and even predict the patient's response to a new treatment.



Picture this: A cancer patient visits her oncologist, gives a few drops of blood or a biopsy specimen, and is told that her genetic expression pattern shows she has a certain subtype of disease. Meanwhile, another expression pattern predicts that her genetic profile should respond well to chemotherapy regimens A and B (with minimal side effects). During her treatments, protein expression patterns are used to make sure that her treatment is effectively disrupting the targeted cellular pathway in her tumor. After treatment, more gene and protein expression patterns verify that the cancer is in remission.



Molecular diagnostics and its newly developed techniques for examining the molecular signatures of cancer cells--protein as well as gene patterns--offer great promise for revolutionizing our approaches to screening, diagnosis, and classification for many different kinds of cancer. Participation in the clinical trials that will validate these new approaches will lead to significant improvements in care for cancer patients in the not-too-distant future.



We would like to hear from you . . .

If you have questions about this tutorial's content, suggestions for new topics, or other feedback on the Web site, please send an e-mail to kerrigad@mail.nih.gov.

If you have questions about this tutorial's artwork or want permission to use it, please send an e-mail to beankelly@verizon.net.