

SEPTEMBER 08, 2006

Focusing in on Cancer's Complexity

In the first large-scale screen of genetic changes in cancer cells, researchers have found that a typical breast or colorectal tumor results from mutations in about 90 genes, with different sets of mutations producing the same type of cancer. But the many different genetic routes to malignancy share common features that point toward new means of cancer prevention, diagnosis, and treatment.

Previous genetic studies of cancer have concentrated on specific genes or on chromosomal regions. In the September 8, 2006, issue of *Science*, Howard Hughes Medical Institute (HHMI) investigators Bert Vogelstein at Johns Hopkins University and Sanford D. Markowitz at Case Western Reserve University School of Medicine, together with a team of researchers from The Kimmel Cancer Center at Johns Hopkins and other institutions, report on a radically new way of identifying genes involved in cancer.

"Scientists who have seen these data have told us that it keeps them up all night thinking."

— Bert Vogelstein

They screened the most well-annotated human genes, a total of more than 13,000 genes that all major genomic centers agree encode proteins. They first looked for mutations in 22 cancerous breast and colorectal tumors. From that list, 191 genes appeared to be particularly important. Scientists who have seen these data have told us that it keeps them up all night thinking, said Vogelstein. It will hopefully open up a large number of opportunities in many areas of cancer research.

The team found far more mutated genes in tumor cells than they had expected. The average breast or colorectal cancer cell was predicted to have an average of 90 mutations that alter protein structure. However, not all 90 were likely to contribute equally to the development of cancers. Through subsequent validation studies, the researchers identified an average of 11 genes in each cancer that were most likely to be directly responsible for its biologic properties. Extrapolating to the total number of genes in the human genome, an average of about 17 genes are expected to have critical involvement in the development of each cancer.

The researchers also were surprised by the heterogeneity of the cancers. Different genes were mutated in cancers of the same type, and the genes contributing to breast cancer were different from those mutated in colorectal cancers. It presents a whole new view of the neoplastic process, said Vogelstein, and explains the heterogeneity that clinicians have long noted to exist among cancer patients.

Despite the complexity of the results, a closer examination of the data has started to reveal an underlying order. Many of the genes that are mutated are involved in pathways thought to be important in cancer, such as cell adhesion, movement, and signaling. Each of these pathways relies on multiple genes, and flaws in any of the genes in a pathway may have similar consequences.

By taking a systems biology approach to connect these genes, we suspect that the complexity will be less than it appears at first sight, said Vogelstein. The same 10 or 20 pathways may be altered in every cancer, though the particular mutated genes in these pathways will be different. The picture will become much clearer as the function of these genes and the ways they interact are better worked out.

This kind of study could not have been done a few years ago, said Tobias Sjöblom, an HHMI research associate in Vogelstein's lab, who is the lead author of the *Science* article. But the availability of the human genome sequence and improvements in sequencing and bioinformatics technologies have made it possible to examine the genome of cancer cells in a comprehensive and unbiased manner, he said.

Still, a massive amount of work was involved. It was a straightforward process once all the mechanistic details had been worked out and the bioinformatic infrastructure was in place, said Sjöblom, but very laborious. The research team formulated 135,483 sets of DNA primers for the polymerase chain reactions needed to sequence the tumor cell genes. They then looked at 11 tumors for each type of cancer, along with two normal samples as a control. The result was almost a half billion letters of DNA sequence that had to be screened for suspicious mutations.

Successive rounds of computer analysis focused attention on smaller and smaller subsets of nucleotides. The hard work was to remove all the junk so that you were left with the true mutations, Sjöblom said. In the final stages, visual inspection of the sequences was required to confirm each mutation. According to Vogelstein, the eye is better than a computer for some types of pattern recognition.

Once the list of mutations was winnowed down, the chromosomal regions containing those mutations were resequenced in the tumors and matched to normal DNA samples to validate each mutation. This process resulted in 1,307 confirmed somatic mutations in 1,149 genes. These genes then were analyzed in 48 additional breast or colorectal tumors, which turned up an additional 365 mutations in 236 of the genes. Altogether, 921 and 751 somatic protein-altering mutations were identified in breast and colorectal

cancers, respectively, most of which were changes in single nucleotides.

The researchers then used statistical techniques to identify the changes in a given gene that were more likely to contribute directly to the cancers' properties. This identified 122 genes in breast cancer tumor cells and 71 genes in colorectal cancers, which the researchers called CAN-genes (candidate cancer genes). Surprisingly, only two genes appeared on both lists. Furthermore, even the types of mutations differed between breast cancer and colorectal cancer. For example, 59 percent of the colorectal cancer mutations went from a C:G base pair to an T:A pair, whereas this was the case for only 35 percent of the breast cancer mutations. These differences may be due to different kinds of carcinogens, different types of repair mechanisms, or different exposures to endogenous mutagens, said Vogelstein. This is a very fertile area of epidemiologists.

Even within each type of cancer, each tumor had its own distinct collection of mutated genes. No cancer had more than six mutated *CAN*-genes in common with any other cancer. This finding also was unexpected, but it's consistent with clinical observations, said Vogelstein, because clinicians have observed for years that each cancer behaves in a unique way.

The complexity of the results may seem discouraging, Vogelstein notes. If anyone thought cancer was simple, they were wrong, he said. On the other hand, once you get the picture in focus, you can start to figure out what's going on. Many of the genes they identified were not previously known to be involved in cancer, and each gene offers potential insights into the disease. The first thing we'll probably delve into is diagnostics, as that's been one of the themes of our lab, Vogelstein said. In particular, they will be looking to find evidence of the mutated genes in blood or other clinical specimens to help identify cancers before they cause symptoms.

Therapeutics based on the newly discovered genes are a ways off, in Vogelstein's estimation. But once the key pathways necessary for cancer are identified, researchers can look for ways to reverse the effects of the activated genes, said Vogelstein. "We now have a whole new set of targets to guide drug development."