**A Catalog of Cancer Genes That’s Done, or Just a Start**

FEB. 6, 2014

Inside



Launch media viewer

A tissue section showing cancer cells inside the lungs. National Cancer Institute



Carl Zimmer

Cancer is a disease of genes gone wrong. When certain genes mutate, they make cells behave in odd ways. The cells divide swiftly, they hide from the immune system that could kill them and they gain the nourishment they need to develop into tumors.

Scientists started identifying these cancer genes in the 1970s and their list slowly grew over the years. By studying them, scientists came to understand how different types of cancer develop and in some cases they were even able to develop gene-targeting drugs. Last May, for example, the Food and Drug Administration [approved](http://www.cancer.gov/cancertopics/druginfo/fda-erlotinib-hydrochloride) a drug known as Tarceva to treat lung cancer in which a gene called EGFR has mutated.

The National Institutes of Health, hoping to speed up the identification of cancer genes, [started](http://cancergenome.nih.gov/newsevents/newsannouncements/news_12_13_2005) an ambitious project in 2005 called the [Cancer Genome Atlas](http://cancergenome.nih.gov). They analyzed 500 samples from each of over 20 types of cancer and found a wealth of new genes. The data have helped scientists discover more of the tricks cancer cells use to thrive at our expense.

“The Cancer Genome Atlas has been a spectacular success, there’s no doubt about that,” said [Bruce Stillman](http://www.cshl.edu/Research-Administration/bruce-stillmans-bio), the president of Cold Spring Harbor Laboratory.

But now, as the Atlas project is coming to an end, researchers at the Broad Institute of M.I.T. and Harvard have published a study in the journal Nature that has scientists debating where cancer research should go next. They estimated that scientists would need to examine about 100,000 cancer samples —10 times as many as the $375 million Cancer Genome Atlas has gathered — to find most of the genes involved in 50 cancer types.

“We now know what it would take to get a complete catalog,” said [Eric S. Lander](http://www.broadinstitute.org/about/bios/bio-lander.html), the founding director of the Broad Institute and a co-author of the new [study](http://www.nature.com/nature/journal/v505/n7484/full/nature12912.html). “And we now know we’re not close to done. We have a lot left to learn.”

Traditionally, scientists have identified cancer genes by comparing healthy cells with cancerous ones. If they find a statistically unusually high number of cells with mutations in a particular gene, they can then examine it to see if it really does help drive cancer — or if it is just carrying a harmless mutation.

Dr. Lander and his colleagues suspected this method could miss some genes. While some cancer genes affect most cells of a given type of cancer, other genes are only involved in a fraction of them. (EGFR, the gene treated with Tarceva, is mutated in only about 10 percent of cases of nonsmall cell lung cancer.) Small samples of cancer cells might not contain the less common mutations.

The Broad researchers suspected that they could catch some of these missing genes by looking at several cancer types at once, because some genes are not limited to a single type of cancer.

For their new study, the scientists examined cancer samples from the Cancer Genome Atlas, as well as cancer samples from the Broad’s own collection. All told, they analyzed 4,742 samples from 21 types of cancer.

The new study detected many of the genes that other scientists have previously linked to those 21 types of cancer. But they also found new genes that had been overlooked before. All told, they identified 33 genes that they consider strong candidates for playing a role in cancer — a potential increase of the catalog of cancer genes of 25 percent.

“This was eye-opening to me,” said Dr. Lander.

Dr. Lander and his colleagues began to wonder how many genes could be found if scientists looked at more cancer samples. Was the cancer catalog almost finished, or only just begun?

“We were able to ask for the first time, ‘Are we there yet?'” said Dr. Lander.

They extrapolated from their own results to gauge how many more samples scientists would need to look at to find most cancer genes involved in at least 2 percent of cancers of a given type.

To find most cancer genes involved in the 50 most common types of cancer, the researchers estimated that they would have to analyze 100,000 samples. In other words, the atlas has gotten us a tenth of the way to the finish line.

[Dr. Harold Varmus](http://www.cancer.gov/aboutnci/director/biography), the director of the National Cancer Institute, said the study has raised valuable questions. “The paper provides some models about what we might think about doing next,” he said. He said the agency is now considering testing Dr. Lander’s hypothesis on a few types of cancer by gathering more samples.

Dr. Lander and his colleagues argue for finishing off the cancer gene catalog. “Completing the genomic analysis of this disease should be a biomedical imperative,” they wrote in their new paper.

In an interview, Dr. Lander said knowing most genes involved in cancer would be a powerful weapon against the disease. “How could we think of beating cancer in the long term without having the whole catalog?” he said. “It would be crazy not to have the information.”

But Dr. Stillman of Cold Spring Harbor Laboratory said completing the atlas has to be weighed against other needs. “Whether we need to know every cancer gene, I’d like to see an argument for how that’s going to help the advancement of new therapy,” he said.

For many researchers, the question comes down to whether extending the atlas project would be the best use of existing research funds. “There’s no question that it would be valuable. The question is whether it’s worth it,” said Dr. [Bert Vogelstein](http://www.hhmi.org/scientists/bert-vogelstein), a Howard Hughes Medical Institute Investigator at Johns Hopkins University.

Some scientists say it might make more sense to study common cancer genes that have already been identified, instead of searching for relatively rare genes that might not turn out to be helpful in fighting cancer.

Also in question is who would pay for advancing the cancer catalog project. “We still don’t know how much money we’re going to have this year,” said Dr. Varmus of the National Cancer Institute’s budget. “We’re not going to set off tomorrow and do 100,000 complete genomes.”

Dr. Lander argued that the project could be done for a reasonable cost, and might also be supported by philanthropic organizations or international partners. In any case, he said, he welcomed a debate about when science will finish the cancer gene catalog.

“If people say, ‘I would rather not know that for five years, or 10 years,’ that’s a reasonable argument,” said Dr. Lander. “But I would rather know that sooner.”

**Correction: February 6, 2014**

Because of an editing error, an earlier version of this article misstated a finding of the new study.  The researchers estimated that scientists would need to examine about 100,000 cancer samples — not genes — to find most of the genes involved in 50 cancer types.