Q&A: Michael Stratton on What's Next in Sequence

Tackling the challenges and opportunities in whole-genome analyses of cancer

After scientists completed sequencing the human genome in 2003, "the world paused, wondering whether DNA sequencing had delivered all that it was going to deliver in the form of reference genomes and sets of genes and sets of predicted proteins," says Michael Stratton, MD, PhD, director of the Wellcome Trust Sanger Institute. But since then, DNA sequencing has roared ahead at an astonishing pace, adding and deepening genomic information for many species.

Stratton, a molecular biologist who led the team that cloned *BRCA2* in 1995, founded Sanger's Cancer Genome Project in 2000. He was appointed director of the 900-person Institute last year. To learn more about today's major initiatives in sequencing cancer, *Cancer Discovery's* Suzanne Rose spoke with Stratton.

What's the current focus in cancer genome sequencing?

The field is moving quickly to systematic exploration of the complete cancer genome, providing complete catalogs of somatic mutations of all classes (point mutations, rearrangements, insertions, and deletions) including driver mutations, which cause the cells to convert from normal cells into cancer cells, or passenger mutations, which are along for the ride.

Is this a continuation of the Cancer Genome Project?

Yes, but when we started, we couldn't sequence very much. We were using conventional modalities of sequencing, based on prior PCR, to interrogate handfuls of genes or handfuls of cancer samples.

Now that next-generation sequencing technologies have come along, we are exploring the cancer genome in a broader way. We're sequencing all the genes in thousands of cancer samples, and we're looking much more into the architecture of the cancer genome because the patterns of mutations—substitutions, deletions, and rearrangements—have to be caused by *something.* Characterizing cancer genomes to get catalogs of mutations allows us to look at mutation patterns and detect processes of mutagenesis that we weren't aware of before.

You mentioned passenger and driver mutations. What has been learned about them?

Until recently, we've really only had a rudimentary view of the thousands of passenger mutations present in most cancer genomes. By looking into those in more detail, we're seeing different patterns between different cancer types and among individual cancers within a cancer type. This variation tells us that there must have been different causes in these different cancer types or individual cancers. Some patterns of mutations we know already. For example, we know that you find a particular pattern of mutations in malignant melanomas due to the impact of ultraviolet light. We don't have a clear idea of what is causing many mutation patterns, but we are beginning to peer into a world that we had only a rather misty view of previously.

Driver mutations and new cancer genes have been emerging at a steady rate over the last few years, and some of them make good targets for new drugs, such as *BRAF* and vemurafenib (Zelboraf; Genentech). However, a lot of the emerging cancer genes are recessive cancer genes (tumor suppressor genes), which aren't easy to use as targets for novel cancer therapeutics because they are already inactivated.

Wellcome Collection London

Sequencing the genomes of thousands of cancer samples has shown how heterogeneous the disease really is. "If we had wanted to create cancer in a way that we could treat, this is not how we would've designed it," says Michael Stratton.

But knowing the whole cancer landscape is what's important. Because we're beginning to see so many combinations of mutated cancer genes even in one type of cancer, we are seeing clearly now how heterogeneous a disease cancer really is. The implications of this heterogeneity are not completely clear, however. On the one hand, it may turn out that all these different combinations of mutated cancer genes feed into the same biologic pathways, making the number of biologic states in cancer cells much smaller and potentially making it easier to target them. On the other hand, these different combinations of mutated cancer genes may cause so much biologic diversity that we begin to understand why cancers have such heterogeneous responses to therapy.

If we had wanted to create cancer in a way that we could treat, this is not how we would've designed it.

What are the major challenges that we now face in sequencing cancers?

Around the world in the International Cancer Genome Consortium, which includes The Cancer Genome Atlas in the United States, we have to build up really large sample collections—ultimately tens of thousands of samples—that we can use in sequencing experiments.

Building on this, we have to deliver a "legacy product" that cancer researchers will mine for decades to come. For that, we need whole-genome sequencing of tens of thousands of human cancers. The genome is finite. We can explore everything, and that's what we should do.

We should do it to high-enough coverage so that we find most of the mutations in each of these tens of thousands of cancer samples, and we should complement these data as much as possible with expression data and with epigenetic data, particularly methylation data. Of course, all sorts of levels of information could be added—from proteomics and so on. But practically speaking, in this first phase of large-scale cancer genomics, that's what we can deliver and, in my opinion, it will transform our perception of the disease.

PhD students entering the field in future years will wonder how cancer researchers ever tried to understand and treat this disease when we had so little understanding of the diversity of the biologic abnormalities that drive it. "Know your enemy" is a robust maxim, and our burgeoning insights into the cancer genome are dramatically improving our understanding of the foe! ■

Q&A: Michael Stratton on What's Next in Sequence

Cancer Discovery 2011;1:460.