Some of the new ENCODE funding goes toward functional validation using newly developed tools, such as a high-throughput version of CRISPR/Cas9 to study the consequences of deleting tens of thousands of genomic elements at a time, in parallel, says Ren, who co-leads this effort with UCSF’s Yin Shen, PhD. Another group, led by Mark Gerstein, PhD, of Yale University in New Haven, CT, is developing a tool called FunSeq, available at http://funseq.gersteinlab.org, to prioritize the impact of mutations in noncoding regions.

Despite the utility of ENCODE resources, some researchers worry that the growing number of “big science” initiatives could deplete NIH funding for investigator-driven projects. “This top-down approach opens the door for insider influence” in choosing topics and soliciting grant applications, says Robert Waterston, MD, PhD, chair of genome sciences at the University of Washington in Seattle.

To be useful to a wide range of investigators, ENCODE is “fairly agnostic” about what biological samples are studied, Feingold says. “The primary goal is to make this resource freely available to the community to study the genetic basis of disease and gene regulation.” –Esther Landhuis ■

Cervical Cancer Analysis Reveals New Mutations

A comprehensive analysis of cervical cancer, a project of The Cancer Genome Atlas (TCGA), has revealed new molecular characteristics that might serve as biomarkers for identifying clinically important patient subgroups in this disease (Nature 2017 Jan 23 [Epub ahead of print]).

So far, only one targeted therapy, the VEGF inhibitor bevacizumab (Avastin; Genentech), which blocks angiogenesis, has received FDA approval for patients with metastatic or recurrent cervical cancer. To uncover other potential targets and better understand how this disease develops, scientists with TCGA used a variety of techniques to profile 228 cervical tumors, including whole-genome and whole-exome sequencing, as well as analyses of copy number, protein levels, mRNA and miRNA expression data, and DNA methylation.

The team reported that these tumors fell into three broad categories: adenocarcinomas, or squamous cell carcinomas with high or low expression of genes in the keratin family. Although previous studies had noted the split within squamous cell carcinomas, this study pinpointed differences “that may serve as key markers for these subgroups,” says co-author Christopher Vellano, PhD, of The University of Texas MD Anderson Cancer Center in Houston. For example, the two subgroups differed in expression of genes such as ARID1A, PIK3CA, and SPRR3.

Approximately 95% of cervical cancers result from persistent human papillomavirus (HPV) infection. However, the researchers pinpointed a subgroup of tumors that were usually HPV-negative and resembled endometrial cancer. This subgroup had high frequencies of ARID1A, KRAS, and PTEN mutations and “may require a different subset of therapeutic options,” says co-author Akinyemi Ojesina, MD, PhD, of the University of Alabama at Birmingham.

The team uncovered five mutant genes not previously implicated in cervical cancer: SHKBP1, ERBB3, CASP8, HLA-A, and TGFBR2. They also determined that most mutations in the profiled tumors appeared to have been caused by APOBEC proteins. Although these enzymes are part of the innate immune system, altering the genetic material of viruses to ward off infection, they are also a major source of mutagenesis if misregulated.

Overall, “we have identified new markers underlying cervical cancer subgroups which may translate to novel clinical therapies,” says Vellano. Additionally, he notes that CD274 and PDCD1LG2—encoding the programmed cell death ligands PD-L1 and PD-L2, respectively—were often amplified in the study tumors, suggesting a potential therapeutic role for immune checkpoint inhibitors.

Like previous TCGA analyses of endometrial and high-grade serous ovarian cancer, this study “is a tremendous reference guide for the rational development of new therapeutic strategies,” says Ursula Matulonis, MD, of Dana-Farber Cancer Institute in Boston, MA, who wasn’t connected to the work. She notes that several immunotherapies have already reached clinical trials for advanced cervical cancer, including the PD-1 blocker pembrolizumab (Keytruda; Merck), and the findings on CD274 and PDCD1LG2 amplification “justify further work” on these approaches.

Jocelyn Chapman, MD, of the University of California, San Francisco, who also wasn’t involved with the study, adds that it “gives us some ideas about where to pursue novel targets.” For example, the findings suggest that the normally beneficial APOBEC enzymes “have been hijacked by cancer,” she says.

–Mitch Leslie ■

Sequencing Errors Rife in Genome Databases

Many low-frequency somatic variants included in The Cancer Genome Atlas (TCGA) may actually be sequencing errors, not necessarily rare driver mutations, as often suspected. Rather, they could be artifacts of DNA damage introduced by routine sample preparation, according to a recent study (Science 2017;355:752–6).

“This is a very timely paper,” says Trevor Pugh, PhD, a cancer geneticist at Princess Margaret Cancer Centre in Toronto, Canada, who was not involved in the new study. “Today, we’re sequencing much, much more deeply than we used to, so we’re going to start confounding mutations like these oxidative-damage mutations with real tumor-driving variants.”

Archived samples are known to be riddled with mutagenic changes that could be confused for tumor-driving mutations, but fresh tumor samples were thought to be mostly fine. Then in 2013, a team from the Broad Institute, which included Pugh, was sequencing tumors from children with neuroblastoma and found hundreds to thousands of mutations when they expected just
10 to 20. The researchers discovered that the acoustic energy used to shear DNA extracted from tumor tissue was frequently turning guanine into 8-oxoguanine, a nucleotide that the sequencing machine read as a thymine (Nucleic Acids Res 2013;41:e67). These G-to-T transversions were not tumor-causing mutations but artifacts of the sonication process.

Laurence Ettwiller, PhD, and her colleagues from New England Biolabs (NEB), a molecular biology reagents company in Ipswich, MA, have now extended those findings and quantified the prevalence of such erroneous variants in two widely used sequencing datasets: the 1000 Genomes Project and TCGA. The researchers compared the reads of the two complementary strands from each sequencing run to detect aberrant transversions introduced by DNA damage and scored the degree of mismatching in a metric dubbed the Global Imbalance Value (GIV).

Based upon the GIV, Ettwiller’s team found that 41% of the datasets in the 1000 Genomes Project contained damaged samples. In TCGA, 73% of the 1,800 sequenced tumor and healthy matched samples revealed damage so extensive that at least half of all the G-to-T variants were not true mutations. Other nucleotide imbalances such as C-to-T occurred at lower but still appreciable frequencies.

According to Pugh, analytic tools like MuTect and VarScan can correct the problem, although not perfectly. To eliminate the false variants, the NEB researchers used a mix of enzymes that repaired the DNA damage before sequencing. “But,” says Ettwiller, “we don’t know whether or not this cocktail of enzymes will actually work on the TGCA dataset,” because of differences in experimental setup.

NEB markets the DNA-repair mix used in the study, so the authors have an inherent financial conflict, yet that doesn’t bother Alexander Dobrovic, PhD, a molecular geneticist from the Olivia Newton-John Cancer Research Institute in Melbourne, Australia. “They clearly have a product to sell, but it’s a useful product,” he says. “We’ll be using that ourselves.”

Another workaround: molecular barcoding, which involves adding unique tags to each stretch of DNA so that errors introduced during library construction can be detected among duplicate sequences and remedied computationally. “I like molecular barcoding because you’re able to directly measure the type and degree of DNA damage,” Pugh says. “You’re reading out exactly what you have in the tube and then correcting for it.”

However, repairing DNA was not the study’s primary aim. “The goal of the paper was to alert the community to a potential problem,” says Tom Evans, PhD, an enzymologist at NEB. “Solutions will come later.” –Elie Dolgin

**Protein Turnover Provides Pancreatic Cancer Target**

Pancreatic adenocarcinoma is such an aggressive and deadly cancer that fewer than 1 in 3 patients live long enough to see the one-year anniversary of their diagnosis. For the majority of patients, surgery is not an option and their tumors have evolved to the point that they’re not responsive to chemotherapy.

Finding new drugs for pancreatic cancer has proven elusive. But according to new research, one of the molecular adaptations that makes these cells more aggressive and treatment-resistant also presents a vulnerability that can be exploited for therapeutic purposes (Nature 2017;542:362–6).

“It’s a beautiful paper,” says Ben Stanger, MD, PhD, from the University of Pennsylvania Perelman School of Medicine in Philadelphia, who was not involved in the study. “It reveals that cancer cells have differential vulnerabilities in the epithelial versus the mesenchymal state.”

In the study, a team from The University of Texas MD Anderson Cancer Center in Houston identified and characterized highly aggressive malignant cell populations that emerge during pancreatic cancer progression.

These highly mobile and invasive cells no longer depend on KRAS signaling and rely on the aberrant activation of mesenchymal programs regulated by the chromatin remodeling factor SMARCB1. Mouse models showed that Smarcb1 ablation could intensify cancer spread; conversely, restoring Smarcb1 slowed tumor growth and restored the cells to their less invasive, epithelial form.

These findings were supported by an analysis of surgically resected specimens from 134 patients with pancreatic ductal adenocarcinoma for whom follow-up data were available. Those whose tumors had high levels of SMARCB1 lived, on average, for around 14 months after their diagnosis. In contrast, those whose tumors had low expression levels had a median survival of just 3.4 months.

Those are the patients where mesenchymal subpopulations are prominent,” says Giannicola Genovese, MD, the study’s first author. “As a result, they do the worst.”

Gene expression profiling revealed that the reduction of SMARCB1 expression leads to an increase in MYC-related activity that drives protein metabolism and the stress response pathways that help the cell tolerate the increased protein turnover.

Therein lies the cancer’s Achilles’ heel. Treatment with the drug AUY922 (luminespiib; Vernalis), which blocks HSP90, reduced cancer growth in Smarcb1-deficient mice but not in Smarcb1-proficient animals. What’s more, the therapeutic effect of AUY922 was enhanced with the addition of drugs targeting the endoplasmic reticulum–stress response pathway. “We identified two ways to target the vulnerability,” says senior study author Giulio Draetta, MD, PhD. “One is we...
Sequencing Errors Rife in Genome Databases

Cancer Discov 2017;7:344-345. Published OnlineFirst February 27, 2017.

Updated version
Access the most recent version of this article at:
doi:10.1158/2159-8290.CD-NB2017-029

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link http://cancerdiscovery.aacrjournals.org/content/7/4/344.2. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.