

SYSTEMS BIOLOGY

Cancer gene discovery goes viral

Examining how proteins from tumor viruses affect their cellular targets helps researchers zero in on the human genes that drive cancer.

Viruses can infiltrate our cells, manipulate cellular proteins and alter gene expression. Some viruses are innocuous, others can kill, and a handful can cause cancer. Led by seven senior investigators including Marc Vidal of Harvard Medical School, a team of virologists, molecular biologists and bioinformaticians set out to learn how proteins from DNA tumor viruses affect targets in the host cell. Driving their work was the notion that proteins of DNA tumor viruses and cancer-causing mutations in the human genome have common mechanisms that lead to disease.

The researchers were inspired by seminal discoveries using tumor viruses in the 1980s and 1990s. Among these was the finding by Vidal's former postdoc advisor, Edward Harlow, that an adenovirus protein interacts with human retinoblastoma, a cell-cycle regulator that causes cancer when mutated. "There were potentially amazing convergent phenomena between what the virus has evolved to touch, to manipulate, to perturb and what the genetics points to," says Vidal.

In a modern twist, the team expanded the insight to a global scale, asking how viral proteins rewire the cell's network of protein interactions and gene-regulatory relationships. They chose members of four distinct DNA tumor virus families—human papillomavirus, Epstein-Barr virus, adenovirus and polyomavirus—and subjected dozens of their encoded proteins to yeast two-hybrid assays to test for binary interactions with 13,000 human gene products. They also identified complexes of viral and host proteins in human fibroblasts by pulling down tandem epitope-tagged viral proteins and analyzing them by mass spectrometry.

To understand the effect of viral proteins on downstream gene expression, the researchers collected microarray data for each protein expressed in fibroblasts and then searched for patterns among the most perturbed genes by clustering. Most of the clusters were enriched for genes associated with one or more functions, such as response to DNA damage, with potential

roles in cancer. By looking at the regulatory regions of genes in each cluster, they identified candidate transcription factors that control expression of these genes.

The researchers synthesized the data into a single network model, which shows how viral proteins can perturb the activity of specific host transcription factors, through physical interaction or by affecting expression, to regulate groups of genes likely to contribute to cancer. Another key finding was nearly 1,000 high-confidence viral host targets enriched for known cancer-causing genes. Intersecting this list with the thousands of predicted functional mutations from several tumor DNA sequencing projects narrowed candidates down to 43 genes that are strong candidates for driving the disease. One of the clearest hits implicates members of the Notch signaling pathway—also among the first hits to come up in tumor DNA sequencing studies.

Vidal takes pains to point out the critical role of collaborative science in the project, in this case under the aegis of a Center of Excellence in Genome Science grant funded by the US National Human Genome Research Institute. "Working together in an interdisciplinary way is obviously more likely to give you a result, or an interpretation of a result, that you wouldn't expect," he says.

The approach may miss perturbations that result from multiple viral proteins acting in the same cell or the effects of different underlying human genotypes. Vidal also concedes that using a single cell line is just a start and that not all proteins of DNA tumor viruses are involved in cancer ("[viruses] have [their] own passenger stuff," he says). What is clear is that orthogonal approaches to tumor DNA sequencing will be critical for sifting through the scores of mutations that tumors accumulate, to find the handful that actually drive progression of cancer. Identifying driver genes will likely require the use of tools from the tiny virus to large-scale functional screens in model organisms.

Tal Nawy

RESEARCH PAPERS

Rozenblatt-Rosen, O. *et al.* Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. *Nature* **487**, 491–495 (2012).