



Although the list of genes mutated, lost or amplified in cancer is growing, finding which of these changes are relevant is a complex task. Could viruses that induce cancer in humans be one route to interpreting these changes?

A consortium of scientists used a stringent yeast 2 hybrid (Y2H) screen to examine the interactome of ~13,000 human open reading frames (ORFs) with ORFs from tumour-causing DNA viruses, including human papilloma virus (HPV) and Epstein–Barr virus (EBV). The authors also expressed the viral ORFs individually in human diploid fibroblasts to examine their effects on the host transcriptome, and to compare protein–protein interactions that were identified in this screen with those found in the Y2H screen. Interactions in the fibroblasts were detected using tandem affinity purification and mass spectrometry (TAP–MS). Comparisons of the Y2H data with the TAP–MS data indicated significant overlap and, overall, 92 transcription factors were either bound or altered by viral ORFs.

The production of a detailed network of 58 viral proteins that alter the function of 86 transcription factors enabled the authors to examine downstream transcriptional effects. For example, by focusing on the low- and high-risk HPV E6 proteins, the authors found that low-risk E6 proteins, but not high-risk E6 proteins, target mastermind-like protein 1 (MAML1) and inhibit NOTCH signalling. Further analyses indicated that the NOTCH pathway is also targeted by viral proteins from the remaining four DNA tumour viruses examined in this study.

Comparison of a stringent set of 947 host genes affected by viral ORFs with data from the COSMIC Classic database of probable human causal cancer genes indicated that 16 proteins targeted by viral ORFs are encoded by candidate cancer-causing genes, with a significant number of these being tumour suppressor proteins. Can these data be used to identify potential cancer-causing

genes from the many thousands of genes found to be mutated by large gene-sequencing projects? The authors ranked 10,543 somatically mutated genes from eight different cancers sequenced in 12 different studies on the basis of their probable effects on cell biology (identified using the PolyPhen2 program) and how often they were mutated. They found that of the top 947 genes, 23 were also present in the COSMIC Classic database, indicating that, in terms of identifying potential cancer-causing genes, these approaches are comparable. Moreover, the authors looked at the top ranked probable cancer genes that overlap between the sequencing studies and the viral–host interactome studies. Of the 43 overlapping genes, five were known cancer-causing genes, and 12 of the genes functioned in the regulation of apoptosis according to Gene Ontology (GO) annotation, indicating that combining these data increases the specificity of cancer gene identification. When compared with genome-wide association studies (GWASs) or approaches involving the identification of somatic copy number alterations (SCNAs), the viral–host gene approach was found to have a higher specificity for potential cancer-causing genes. Interestingly, although genes were found in common between the SCNA deletion data and the host-interactome data, there was no substantial overlap between genes identified by viral–host interactions and genes amplified by SCNAs, suggesting that viral proteins target tumour suppressor genes in preference to potential oncogenes.

The authors conclude that their approach makes it more likely that cancer-causing genes can be identified from high-throughput sequencing data, and that proteins encoded by DNA tumour viruses and genetic alterations selected for during tumour development can affect common pathways.

Nicola McCarthy

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

“ proteins encoded by DNA tumour viruses and genetic alterations selected for during tumour development can affect common pathways ”