



# Combing the genome for cancer genes

DOI:  
10.1038/nrc2279

New approaches make it possible to search for cancer-causing mutations on a genome-wide scale. Two recent studies have characterized the cancer genome — one at the level of individual nucleotide changes and the other at the level of copy-number variation — and showed that there are still many more cancer-associated genes to be discovered, a large number of which are specific to subsets of tumours.

Vogelstein and colleagues sequenced all the genes in the Reference Sequence database (a consolidated collection consisting of most human gene sequences) from 11 breast and 11 colorectal tumours. Of the 18,191 genes they sequenced, 1,718 had a non-silent mutation in at least one tumour sample. However, most of these are likely to be passenger mutations — ones that are not causally involved in tumorigenesis but which have arisen in clones that subsequently acquired causal mutations. The authors used a two-stage design that included investigation of 24 to 120 additional tumour samples to help identify genes that are most likely to be causal. By this method, they identified 280 candidate cancer genes.

In addition to analysing these candidates on a gene-by-gene basis, the authors classified them into

pathways. Mutations affecting the phosphatidylinositol 3-kinase (PI3K) pathway were particularly common although, interestingly, different parts of the pathway were mutated in breast and colorectal tumours. The overall picture of the genomic landscape of cancer that emerged from this study was that, although there may be some genes that are mutated in many tumours (such as *TP53*), most of the candidate cancer genes were mutated at relatively low frequencies — typically in fewer than 5% of tumours.

Meyerson and colleagues used single nucleotide polymorphism (SNP) arrays to look at copy-number alterations in 371 lung cancer samples. Alterations in copy number at each of the 238,000 SNPs were assessed by measuring the intensity of hybridization compared with a sample of normal genomic DNA.

Most chromosomal arms underwent deletion or amplification in a large proportion of samples, with 10 large segments being frequently amplified and 16 deleted. However, with these large segments, it is difficult to pinpoint the gene whose deletion or amplification is causal, so the authors then focused on smaller gains and losses. Common deletions occurred in the regions containing known cancer-associated

genes *CDKN2A*, *CDKN2B*, *PTEN* and *RBI*. Another commonly deleted region contained the gene *PTPRD*, which encodes a tyrosine phosphatase; the authors found *PTPRD* mutations in several samples that they sequenced.

Some of the amplifications were in known oncogenes such as *EGFR*, *KRAS* and *ERBB2*, but others were not; for example, most common amplification involved a region containing *MBIP* and *NKX2-1*. RNA interference knockdown of *NKX2-1* but not of *MBIP* reduced the ability of cells to grow in an anchorage-independent manner, suggesting that *NKX2-1* is a novel oncogene. Interestingly, *NKX2-1* is a lineage-determining factor for lung alveolar cells.

These two studies demonstrate that many cancer-associated genes remain to be identified, and that many of them will be important in a relatively small proportion of tumours. Genomic technologies such as sequencing and SNP arrays offer efficient means of identifying these genes.

Patrick Goymer

**ORIGINAL RESEARCH PAPERS** Wood, L. D. *et al.* The genomic landscapes of human breast and colorectal cancers. *Science* 11 October 2007 (doi:10.1126/science.1145720) | Weir, B. A. *et al.* Characterizing the cancer genome in lung adenocarcinoma. *Nature* 4 November 2007 (doi:10.1038/nature06358)

