CANCER GENOMICS

Complete coverage?

Does possession of the complete sequence of the human genome allow us to identify all the binding sites for a transcription factor? Until recently this has been logistically infeasible, but Wei *et al*. have now developed a technique to scan the entire genome for binding sites, and they have identified at least 98 previously unknown targets of the tumour suppressor and transcription factor p53.

Chromatin immunoprecipitation (ChIP) is a technique that extracts DNA fragments to which a transcription factor is bound. Having done this, these fragments need to be identified. One approach is to hybridize them to microarrays (ChIP on chip); this has been used successfully in yeast but mammalian genomes have proved too large. Alternatively, the fragments can be sequenced, but this has also proved unwieldy for mammalian genomes and only a few chromosomes have been studied in any one screen.

These authors have recently developed a new sequencing approach that is sufficiently efficient to screen the entire human genome — paired-end ditag (PET) sequencing. In this method, after cloning the precipitated fragments, the 5′ and 3′ ends of several clones are concatenated for efficient sequencing. Each pair of ends is then mapped to the genome to identify a potential binding site.

The authors chose to test their new method on p53 because of its importance in cancer and because many of its transcriptional targets are already known. They screened colorectal cancer cells that had been treated with 5-fluorouracil to activate p53 expression. After a comparison with expression data they identified 122 genes that are direct targets of p53, 98 of which were novel. Interestingly, they found a difference in the position of the binding site between genes that were upregulated and downregulated and even identified a second p53-binding site in the promoter of a well known p53-target gene, *CDKN1A*.

Several of the newly identified p53 targets are involved in cell motility, which is interesting as p53 is involved in suppressing metastasis. To further test the clinical significance of the p53 targets the authors looked at their expression in 251 breast tumour samples, some of which were p53 mutants and some wild type. The expression of the p53 targets clearly distinguished between the two types: many p53-downregulated genes were expressed at higher levels in p53-mutant tumours, and *vice versa*. Among the interesting biomarker candidates was the anti-apoptotic BCL2A1, which was identified here for the first time as being repressed by p53.

A further consequence of this study was a refinement of the consensus p53-binding sequence. Screening the genome *in silico* with this identified many more genes that are potential p53 targets. This might represent the complete set of potential p53 targets, whereas the results of the ChIP-PET screen are the genes that are actually regulated by p53 in colorectal cancer cells treated with 5-fluorouracil. Studies with ChIP-PET on other samples will show how the complete set relates to specific sets. ChIP-PET might therefore prove an important development both in cancer genomics and in genome biology in general.

Patrick Goymer

ORIGINAL RESEARCH PAPER Wei, C.-L. *et al.* A global map of p53 transcription-factor binding sites in the human genome. *Cell* **124**, 207–219 (2006)
FURTHER READING Ng, P. *et al.* Gene identification signature (GIS) analysis for transcriptome characterization and genome annotation. *Nature Methods* **2**, 105–111 (2005)