

FUNCTIONAL GENOMICS

Affinity for enhancer prediction

Differences in the DNA-binding affinities of transcription factors and in the tissue-specific combinations in which transcription factors bind regulatory elements underlie the temporal and spatial patterns of gene expression that characterize development. Although transcription-factor binding sites within promoter and 3' UTR sequences have received much attention, similar studies of distal enhancer elements have been hindered by both experimental and computational difficulties. In their recent study, Hallikas, Palin and their colleagues have tackled both problems. They have developed a high-throughput method for determining the DNA-binding specificities of representative members of three transcription-factor families and a computational tool to predict the location of novel mammalian enhancers.

To address the issue of varying affinities of transcription factors for different DNA sequences, the authors developed a microwell-plate-based, high-throughput method in which the DNA-binding domains of a transcription factor of interest were fused to a luciferase reporter. The fusion proteins were incubated with biotinylated oligonucleotides that had the highest affinity to the transcription factor of interest. Binding was detected by measuring luciferase activity that could be captured by immobilized streptavidin. By adding competing unlabelled oligonucleotides that carried single base substitutions to the reaction, the authors could determine the relative affinity of the given transcription factor. The utility of the assay was demonstrated for members of the GLI family, and for TCF4 and ETS1 transcription factors.

Parallel to the experimental approach, the authors developed a computational tool — called

enhancer element locator (EEL) for genome-wide enhancer analysis. EEL aligns the order of transcription-factor binding sites on orthologous DNA sequences by taking into account clustering, affinity and site conservation. Moreover, a system of correction factors is used to account for cooperativity between transcription factors at a given enhancer.

Having tested EEL on well-defined *Drosophila* enhancers and on mammalian *MyoD*, the authors carried out pairwise genome-wide alignments of human genes with their orthologues from the mouse, rat, chicken and pufferfish. Several potential enhancers were chosen for further study — for example, to test their ability to drive expression of reporter genes. The confirmation rate was impressive for the highest scoring enhancers, although there were also a few surprises. Some enhancer-reporter gene constructs recapitulated only a part of the expression pattern of the corresponding gene, indicating the regulation of these genes by multiple enhancers.

The expression of reporter genes in tissues where the studied transcription factors are potentially not active could in turn be explained by the presence of other transcription-factor binding-site modules within the enhancers tested.

The experimental and computational approach described by Hallikas, Palin and colleagues are genome-wide and can be easily generalized to any transcription factor for which binding specificity is known. Together with other recently developed methods to study DNA binding on a global scale, they will help us to dissect the spatio-temporal complexity of gene regulation.

Magdalena Skipper

ORIGINAL RESEARCH PAPER Hallikas, O. & Palin, K. *et al.* Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell* **124**, 47–59 (2006)

FURTHER READING Wasserman, W. W. & Sandelin, A. Applied bioinformatics for the identification of regulatory elements. *Nature Rev. Genet.* **5**, 276–287 (2004)

