

GENOMICS

Capturing transcription factors in the wild

An assay that snares transcription factors with genomic DNA fragments reveals how genetic and epigenetic factors shape binding behavior.

Knowing exactly where transcription factors (TFs) bind in the genome can unlock many mysteries of cell biology. However, this information can be hard to come by. Chromatin immunoprecipitation followed by sequencing (ChIP-seq) directly captures TF–DNA interactions *in vivo*, but it is technically challenging. *In vitro* screens with synthetic DNA oligonucleotides make it possible to define the motifs that various TFs can bind at high throughput, but these assays can be undermined by loss of context—adjacent genetic and epigenetic elements that influence *in vivo* binding.

Researchers led by Joseph Ecker at the Salk Institute for Biological Studies have now developed a strategy that may offer a happy medium. Their DNA affinity purification

sequencing (DAP-seq) approach combines *in vitro*-purified TFs with tagged fragments of genomic DNA. These chunks of DNA retain the original sequence milieu of the TF binding site, as well as any epigenetic modifications. This makes it possible to capture large numbers of DNA–TF interactions under more realistic conditions, and the bound sequences can then be sequenced and mapped back onto the genome.

Working with the thale cress plant, *Arabidopsis thaliana*, Ecker and colleagues were able to map binding sites for more than a thousand different TFs situated across roughly 9% of the genome, and they identified actual motifs for 529 of those TFs. Critically, direct comparisons of DAP-seq data with the ‘gold standard’ *in vivo* data collected via ChIP-seq showed that the new method performed very well.

The researchers also used DAP-seq data to uncover patterns of reduced TF binding

at genomic sites that are known to undergo methylation in leaf cells. They confirmed these findings with a parallel assay in which they used PCR to amplify the genomic DNA before performing DAP-seq, essentially erasing all epigenetic marks. This restored TF binding at roughly 180,000 sites that would otherwise have been inaccessible as a result of epigenetic modifications—and presumably transcriptionally silenced.

Ecker and colleagues have made their *Arabidopsis* data freely available online, and they conclude that such data should “provide a valuable resource to evaluate the impact of natural genetic and epigenomic variation on transcriptional networks controlling plant adaptation.”

Michael Eisenstein

RESEARCH PAPERS

O’Malley, R.C. *et al.* Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell* **165**, 1280–1292 (2016).