Alex Hartemink

Department of Computer Science
Duke University.

Toward a Mechanistic Understanding of Transcriptional Regulation: A Systems Perspective on Genome Occupancy

A key goal of regulatory genomics is to predict gene expression---and more immediately, transcript production---from genomic sequence. One significant milestone toward this goal will be to accurately predict promoter occupancy from genomic sequence. We adopt a systems perspective to model the competitive binding of multiple factors along the genome, with an eye toward a more mechanistic understanding of genome occupancy and its dynamics. Hundreds of different factors adorn the eukaryotic genome, binding to it in large number. These DNA binding factors (DBFs) include nucleosomes, transcription factors (TFs), and other proteins and protein complexes, such as the origin recognition complex (ORC). DBFs compete with one another for binding along the genome, yet many current models of genome binding do not consider different types of DBFs together simultaneously. Additionally, binding is a stochastic process that results in a continuum of binding probabilities at any position along the genome, but many current models consider genomic positions to be either binding sites or not binding sites. We present COMPETE, a model that allows a multitude of DBFs, each at different concentrations, to compete with one another for binding sites along the genome. The result is an occupancy profile, a probabilistic description of the DNA occupancy of each factor at each position along the genome.