

Modeling transcriptional regulation

Systems models of the ways transcription factor networks operate and evolve are essential for understanding cell identity, developmental commitment and regulatory variation. Terminologies from different techniques and disciplines may need to be adapted or put aside to make and test these models effectively.

Four contexts are crucial for understanding the function of proximal (promoter) and distal (enhancer) *cis* regulatory elements in the genome: firstly, the genes they regulate; secondly, the cell types in which they are active; thirdly, the set of transcription factors that bind the motifs within the element; and, fourthly, the chromatin loops in which the regulation occurs (*Cell* doi:10.1016/j.cell.2014.11.021; 11 December 2014).

On page 8, Sebastian Pott and Jason Lieb argue that the definition of a ‘super-enhancer’ (*Cell* **153**, 307–319, 2013) is useful in identifying some distal *cis* elements responsible for cell type-specific transcriptional regulation but is not consistent enough to restrict our search for key regulatory elements to those that fit the current definition. They highlight the discrepancy between the number of enhancer elements tested functionally and the large number of putative elements identified by chromatin immunoprecipitation and sequencing (ChIP-seq) and claim that relatively few elements have been characterized by both criteria. However, to be fair, enhancer trapping in cells (*Nat. Genet.* **46**, 685–692, 2014) and *in vivo* (*Nat. Genet.* **43**, 379–386, 2011), together with analysis of reporter constructs in transgenic animals (*Nat. Genet.* **42**, 806–810, 2010 and *Nat. Genet.* **44**, 89–93, 2011), identifies a large number of elements characterized for both the chromatin marks and function of enhancers. Many of the binding motifs for transcription factors within enhancers have been identified; but the majority of motifs still need to be discovered (*Nature* **489**, 91–100, 2012 and *Nature* **507**, 462–470, 2014). This Perspective (page 8) argues that we should suspend judgment about the importance of the length of a

complex enhancer region and binding of its coactivator Med1 until we have more multi-level information about its operation in different cell types and contexts. For example, there are simple enhancers with exceptionally high coactivator binding.

From the perspective of this journal, interested in finding predictive models with which to test the effects of regulatory variation in eukaryotic genomes, the most important context in which an enhancer element finds itself is the network of mutual regulation among transcription factors and the genes that encode them. Bearing in mind other layers of chromatin modification, noncoding RNA and other signals, we bet that the framework of operation of the switching networks of the genome can be seen in mutual transcription factor interactions. We therefore single out for special attention the proximal and distal *cis* elements of transcription factor genes that are bound by other transcription factors. Davidson and Levine (*Proc. Natl. Acad. Sci. USA* **105**, 20063–20066, 2008) lay out the logic of gene circuits that confer cell identity, enable developmental transitions and transmit cell-cell communication. Densely connected networks of multiple mutually activating transcription factors and their genes have the attractor property we expect of cell identity. These networks are beginning to be found experimentally (*Cell* **144**, 296–309, 2011). The discovery of regulatory variation within the motifs of enhancers in such a network offers the potential for adaptive evolution within a network that should otherwise be relatively robust to mutation (*Nat. Genet.* **46**, 136–143, 2014). We are betting that these dense mutual networks are a general feature of the transcriptional commitment of many cell types. ■