



## REGULATORY ELEMENTS

# Predicting functional modules

The precise spatiotemporal patterns of gene expression that are required during the development of multicellular organisms are largely governed by the binding of transcription factors to *cis*-regulatory sequences. Although many methods have been successful in identifying specific *cis*-regulatory modules (CRMs), it remains difficult to systematically describe the location and function of CRMs in a specific developmental process. This paper presents a novel computational approach that tackles this challenge and has led to the most complete, quantitative description to date of the control network for fruitfly embryo anteroposterior patterning.

Kazemian and colleagues developed a regression-based model that estimates the potential of a genomic region to control the spatially restricted expression of a nearby gene, which they term the 'pattern-generating potential' (PGP)

of the region. The model integrates information on the DNA-binding specificity of transcription factors, sequence conservation, patterns of expression of transcription factors and gene expression data sets. The inclusion of target gene expression information is a key difference from previous models of this type as it allows quantitative prediction of the function of CRMs. Considering patterns of expression of the transcription factors themselves ensures that the predictions are appropriate for a specific developmental time point and cell type.

The authors chose a well-characterized developmental process — fruitfly embryo anteroposterior patterning — for training and testing their model, and this approach should be applicable in any system for which there is adequate information on transcription factor-binding specificity and gene expression patterns.

The CRMs predicted by the PGP model are supported by data from previous genome-wide chromatin immunoprecipitation (ChIP) studies, and the authors suggest that their modelling approach can be more informative than ChIP because many of the sites bound by transcription factors *in vivo* may not be regulatory in a specific context. The results from this study also suggest that many genes seem to have multiple CRMs that drive the same expression pattern — whether this is a widespread phenomenon in embryonic development will be an interesting question to explore.

Mary Muers

**ORIGINAL RESEARCH PAPER** Kazemian, M. *et al.* Quantitative analysis of the *Drosophila* segmentation regulatory network using pattern generating potentials. *PLoS Biol.* **8**, e1000456 (2010)  
**FURTHER READING** Farnham, P. J. Insights from genomic profiling of transcription factors *Nature Rev. Genet.* **10**, 605–616 (2009) | Alexander, R.P. *et al.* Annotating non-coding regions of the genome. *Nature Rev. Genet.* **11**, 559–571 (2010)