

 GENE EXPRESSION

Of ChIPs and flies

Whether a gene is active or silent depends on a sequence-specific interaction between transcription factors and *cis*-regulatory modules (CRMs). A new genome-wide study in *Drosophila melanogaster* embryos reveals that transcription factor binding is much more extensive than previously thought and that the level of binding, rather than simply its presence or absence, might be functionally important.

The complexity of gene regulation at the transcriptional level is well illustrated during early embryonic development. For this reason, Li, MacArthur and colleagues turned to the *D. melanogaster* blastoderm stage. Using ChIP–chip (chromatin immunoprecipitation combined with microarray analysis) they determined, on a genome-wide scale, *in vivo* binding sites for six transcription

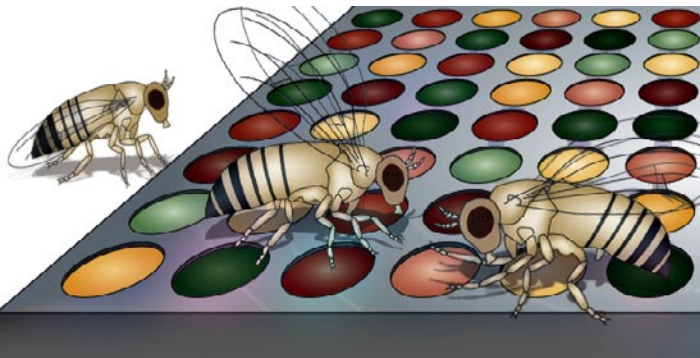
factors that have a key role in anterior–posterior axis patterning. The first and most striking observation was the extent of the binding: most factors were seen binding to thousands of sites — many more than expected on the basis of their known targets. There was also extensive overlap between the binding sites, although each factor bound with qualitatively different preferences.

But could all these sites indicate functional regulation? The levels of site occupancy turned out to be crucial — the few hundred most highly bound regions were associated with genes that are known to be transcribed, show patterned expression in the blastoderm, and tend to be highly conserved among *Drosophila* species. However, thousands of other regions, which were bound at lower levels, are not found close to genes that are transcribed in the early embryo, and tend to be found either in less well-conserved non-coding DNA or within protein-coding regions. The authors suggest a number of possible explanations for the low-level binding in other regions: it might regulate expression at another stage in development; it might have a structural function within the chromatin; or it might not be functional at all — low-level binding has been proposed to occur simply on thermodynamic grounds.

The authors admit that determining which regions are truly functional *in vivo* will be difficult. For example, comparative genomics is frequently used to identify conserved sequences, taking their conservation as a sign of functional constraint. In this current analysis, however, specific recognition sites for five of the six factors, although conserved relative to unconstrained DNA, are no more conserved than the immediately flanking DNA, making it impossible to determine whether the conservation reflects functionality or something else.

As well as providing a genome-wide glimpse of how gene expression is regulated in *cis*, this study contains a caution against using ChIP data without further evidence of functionality. Importantly, it is now also clear that CRM occupancy must be thought of in quantitative terms, rather than as an on–off state, to fully explore the function of these sites.

Magdalena Skipper



ORIGINAL RESEARCH PAPER Li, X.-Y. & MacArthur, S. et al. Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLoS Biol.* **6**, e27 (2008)
FURTHER READING Kapranov, P. et al. Genome-wide transcription and the implications for genomic organization. *Nature Rev. Genet.* **8**, 413–423 (2007) | Schones, D. E. & Zhao, K. Genome-wide approaches to studying chromatin modifications. *Nature Rev. Genet.* **9**, 179–191 (2008)