

# Chromatin profiling in model organisms

Tony D. Southall and Andrea H. Brand

## Abstract

The correct control of gene expression is essential for the proper development of organisms. Abnormal expression of genes can lead to cancerous growth and certain diseases. To understand how gene expression is controlled on a genome-wide scale, methods for assaying transcription factor binding sites are required. There are two prevailing techniques for mapping protein–chromatin interactions, ChIP (chromatin immunoprecipitation) and DamID (DNA adenine methyltransferase identification). Both of these methods, when combined with microarray technology, can provide powerful insights into transcription factor function, higher order chromatin structure and gene regulatory networks. *In vivo* chromatin profiling studies are now being performed on model organisms, targeting specific tissues to help generate more accurate maps of protein–DNA interactions.

**Keywords:** ChIP-chip; DamID; genetic regulatory networks; model organism; chromatin profiling

## INTRODUCTION

Musicians in an orchestra have to play notes and chords in a timely and coordinated manner to produce a beautiful piece of music. Similarly, transcription factors must regulate the expression of genes in a timely and coordinated manner to enable cells to differentiate and to perform their proper developmental or physiological function, be it differentiating into a neuron, transporting ions or secreting insulin. In humans, specific diseases and the onset of certain cancers are associated with the disruption of the normal function of transcription factors, histone-modifying enzymes or DNA repair proteins.

To gain a comprehensive understanding of how gene regulatory networks control cellular behaviour, physical mapping of chromatin-associated proteins and chromatin modifications is required. There are two main techniques for mapping these interactions, chromatin immunoprecipitation (ChIP) and DNA adenine methyltransferase identification (DamID). These methods, when coupled with microarray technology provide a powerful approach to profile chromatin modifications or protein–chromatin interactions on a genome-wide scale.

Many model organism genomes, such as mouse, *Drosophila* and *Caenorhabditis elegans*, are now fully sequenced and annotated; this information allows high density DNA microarrays to be designed that tile the entire genome of these organisms. This technology allows the mapping of all binding sites of a chromatin-associated protein in a single experiment. While *in vitro* analysis of protein–DNA interactions can provide important insights, *in vivo* studies give a more accurate representation of the endogenous binding of proteins to DNA. This review provides an overview of *in vivo* chromatin profiling techniques, focusing in particular on their use in model organisms.

## ChIP-chip

Chromatin immunoprecipitation (ChIP) was first described in a study examining the binding of histone H4 to the *Drosophila hsp70* gene [1]. Solomon and colleagues cross-linked proteins to genomic DNA using formaldehyde, fragmented the DNA and immunoprecipitated the histone–DNA complex using an antibody to histone H4. The DNA was

Corresponding author. Andrea H. Brand, The Gurdon Institute and Department of Physiology, Development and Neuroscience, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, UK. Tel: 44 1223 334141; Fax: 44 1223 334089; E-mail: ahh@mole.bio.cam.ac.uk

**Tony Southall** is a Post-Doctoral Researcher working on the transcriptional regulation of neural stem cell self-renewal and differentiation in *Drosophila* at the Gurdon Institute and the Department of Physiology, Development and Neuroscience at the University of Cambridge.

**Andrea Brand** is the Herchel Smith Professor of Molecular Biology at the Gurdon Institute and the Department of Physiology, Development and Neuroscience at the University of Cambridge and heads a lab investigating nervous system development.

then probed on a Southern blot to identify regions of the hsp70 gene that histone H4 is associated with. The immunoprecipitated DNA can also be analysed by quantitative PCR [2], however the combination of ChIP with DNA microarrays (ChIP-chip) allows the rapid mapping of protein–DNA interactions over multiple genomic loci in a single experiment [3]. Figure 1A is a schematic of how ChIP-chip is performed.

ChIP-chip has been extensively used in yeast for studying genome-wide transcription factor binding [3, 4] and dynamics of transcription factor binding [5, 6]. Furthermore, the analysis of chromatin structure, histone modification and DNA repair has benefited from ChIP-chip technology [7–12]. The use of antibodies that can recognize specific modifications on histones (e.g. methylation or acetylation) has allowed the correlation of gene activity with histone modifications [8] and the identification of histone landscapes that correspond to distinct regions of the genomic scaffold, such as promoter regions or coding sequence [9].

To date, the greater part of ChIP-chip experiments have been performed on cell lines, however, there are an increasing number of tissue and organismal studies in several different model systems. Recently, Ercan and colleagues mapped the binding of two components of the dosage compensation complex (which equalizes the expression of X-linked genes between sexes) at high resolution throughout the whole genome of *C. elegans* by performing ChIP on embryos [13]. In *Drosophila* there have been several studies using embryonic material, including the analysis of the location of heat-shock factors [14] and of dorso-ventral patterning transcription factors on a genome-wide scale [15]. In vertebrates, there have been studies of histone modifications in zebrafish embryos [12] and a recent comparison of transcription factor binding in humans and mice used liver tissue for ChIP profiling [16]. Last but not least, this technique is now established in *Arabidopsis* and has been used to map the binding of TGA2 at high resolution across the whole genome [17].

ChIP-chip allows identification of regions of DNA that are bound by an endogenous protein and is especially effective for mapping chromatin proteins with specific modifications. However, it relies on the availability of a highly specific antibody against the protein of

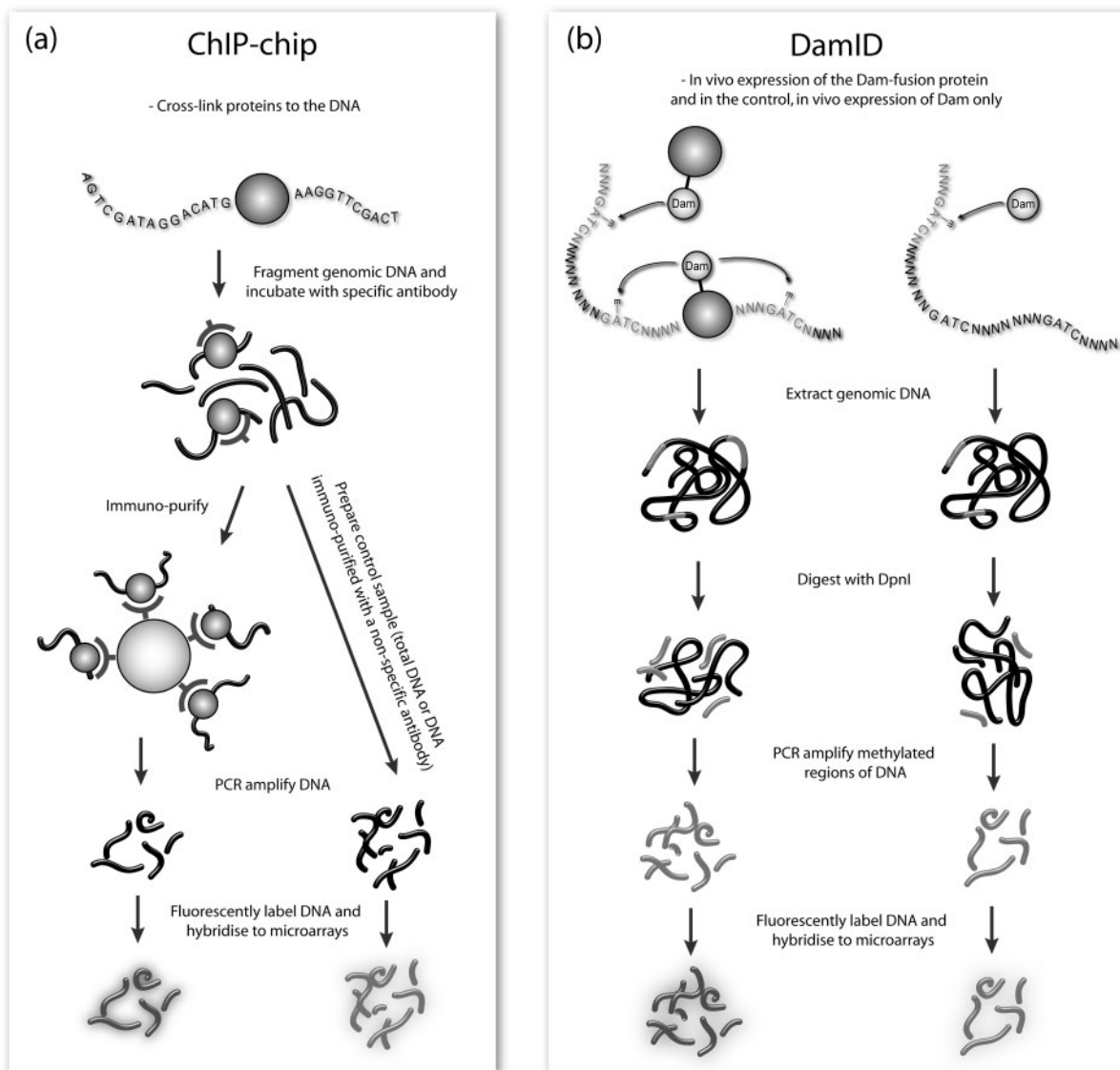
interest, which is not required by the next technique to be discussed.

## DamID

An alternative chromatin profiling method is DNA adenine methyltransferase identification (DamID). First described by Bas van Steensel and Steven Henikoff [18, 19], it entails the tethering of an *Escherichia coli* DNA adenine methyltransferase (Dam) to the DNA or chromatin-associated protein of choice. When the fusion protein is expressed, it binds to chromatin and the Dam methylates the adenine in surrounding GATC sites. This methylation can extend to several kilobases from the binding site [19] and leaves a signature on the DNA that can be analysed by post-processing of the extracted genomic DNA using the DpnI restriction enzyme (only cuts at methylated GATC sites) followed by PCR-mediated amplification [20, 21] (Figure 1B). Dam is a highly active enzyme, which can result in non-specific methylation of DNA when the protein is not tethered. Therefore, to compensate for this, a parallel control experiment is performed expressing the Dam protein alone. After genomic DNA processing, specific binding of the protein is detected as enrichment in the experimental sample over the control, either by quantitative PCR or most commonly using tiling microarrays.

DamID was initially developed for use in *Drosophila* [18], using a non-induced heat-shock promoter to drive expression either in cell culture or in the organism by producing transgenics. Chromatin profiling of Myc, Max, Mad/Mnt [22], Hairy [23], GAGA [24] and HP1 [24–26] have been performed using DamID in *Drosophila* cell culture. The system has been adapted for use in mammalian cell culture [27] using a non-induced expression vector and also by using a lentiviral system [28].

Later, whole organism studies in *Drosophila* utilized a non-induced GAL4 responsive promoter [29] to drive expression in transgenic embryos [23, 30, 31]. Tiling arrays covering the entire euchromatic genome allowed a comprehensive mapping of the transcription factors Prospero and Even-skipped in the developing *Drosophila* embryo [30, 31]. These studies showed that Prospero acts as binary switch between neural stem cell renewal and differentiation [30] and that Even-skipped directly regulated genes controlling the electrical properties of neurons [31]. Additionally, DamID has also been



**Figure 1:** ChIP-chip and DamID experimental design. **(A)** ChIP-chip technique. Protein-DNA complexes are cross-linked using formaldehyde. The cell extract is then sonicated to fragment the DNA (to ~1 kb or smaller) and incubated with an antibody (usually a polyclonal antibody) to your protein of interest (small circles). The protein-DNA complex is then immunopurified to enrich the sample for DNA fragments corresponding to the binding regions of the protein of interest. A control sample is also prepared, either using the total DNA (from before immunopurification) or by immunopurifying with a non-specific antibody. The experimental and control samples can then be labelled with different fluorophores and hybridised to microarrays. **(B)** DamID technique. A fusion protein consisting of DNA adenine methyltransferase (Dam) and the protein of interest is expressed at low levels. The DNA surrounding the site of binding is methylated and after extraction, the genomic DNA is cut with DpnI (cuts at methylated GATC sites). The methylated DNA (shown as the grey DNA) is then amplified using ligation-mediated PCR and labelled ready for hybridisation to microarrays. A parallel experiment, expressing Dam only, is performed to control for non-specific methylation. Figure adapted from [43].

successfully set up and utilized in *Arabidopsis* to test targets of LHP1 [32].

DamID has the advantage that a specific antibody is not required; Dam itself is small (32 kDa) and it can

be fused to either end of a protein [18]. DamID can detect indirect interactions of proteins with DNA. However, because Dam is highly active, expression levels of the fusion protein have to be kept at a very

low level, enough to methylate wild-type sites but not so much that the system becomes saturated from non-specific methylation. Presently, this means that in transgenic organisms, expression of Dam has to be very low and ubiquitous. The disadvantages and possible improvements of the technique are discussed later.

## DNA MICROARRAYS FOR CHROMATIN PROFILING

Early ChIP-chip and DamID experiments made use of spotted cDNA arrays [19, 22]. This progressed to using tiling arrays that cover all intergenic regions of the yeast genome [3] and specific regions of the *Drosophila* genome [8, 14, 24] generated by spotting PCR amplicons. The invention of maskless array synthesis [33, 34] has allowed the generation of high density oligonucleotide arrays with up to ~390,000 features per array (Nimblegen Systems). Using this technology, high resolution whole genome ChIP and DamID has been performed on human cells [35], *Arabidopsis* [17], *Drosophila* embryos [30] and *C. elegans* embryos [13]. Full genome arrays for mouse, rat, dog, chicken, yeast and *E. coli* are also available. Agilent Technologies generate high density spotted oligonucleotide tiling arrays for chromatin profiling studies. These have included the mapping of key transcription factors involved in human embryonic stem cell maintenance [36] and of transcription factors governing dorso-ventral patterning in *Drosophila* [15]. For zebrafish, Wardle and colleagues [12] have designed promoter tiling arrays in conjunction with Agilent for ChIP-chip experiments. The scale and availability of tiling arrays is becoming less of a limiting factor for comprehensive chromatin profiling experiments with the present and continuing advancement of microarray synthesis technologies.

## ChIP VERSUS DamID

ChIP-chip and DamID have their respective advantages and disadvantages, some of which have been highlighted in previous reviews [37, 38]. Advantages of ChIP include its capacity to map proteins with specific post-translational modifications (e.g. histones), its potential to pinpoint binding to within a few hundred base pairs and that, in tissues or whole organisms, it detects the interactions of the

endogenous protein (which is expressed at its correct level and location). The disadvantages are the necessity for a specific antibody to the protein, false positive results from cross-linking artefacts and the technical problems inherent in trying to detect transient or indirect interactions between proteins and the DNA.

DamID does not require a protein-specific antibody, it works well for detecting indirect interactions with chromatin, it produces wide and robust signatures in tiling array experiments and also allows a lot of flexibility in analysing the effects of protein mutations and truncations on DNA-binding activity. However, there are disadvantages, which include the risk that the Dam fusion may act differently to the endogenous protein and the present inability to target expression spatially and temporally in transgenic organisms (discussed in more detail below). This restriction currently means that the fusion protein is ubiquitously expressed at low levels in the organism resulting in its presence in tissues where it is not usually expressed. Furthermore, if a particular time window is being studied, previous interactions of the protein with DNA may still be detected due to the constant accumulation of adenine methylation throughout the lifespan of the organism.

Therefore, if a specific antibody is available, ChIP-chip provides a robust technique for chromatin profiling. However, if there is no antibody available, DamID presents a good option for mapping the binding of a protein to chromatin and also provides flexibility to study mutations and truncations of the protein of interest.

## THE FUTURE FOR CHROMATIN PROFILING TECHNOLOGIES IN MODEL ORGANISMS

The standard ChIP technique requires a protein-specific antibody, however there is a method that can circumvent this and provide superior signal-to-noise ratios. This involves generating a tagged version of the protein that can be biotinylated, allowing the protein-DNA complex to be pulled down by the strong biotin-avidin interaction [39]. The strength of the biotin-avidin interaction permits much stronger washes and has been shown to improve the sensitivity of ChIP-chip experiments [40]. This robust system could be exploited in model organisms by either introducing the tag

into the endogenous locus by homologous recombination or by inserting a transgene that contains the tagged gene under its own promoter.

DamID is open to several modes of modification; it has already been useful in mapping long range interactions of chromatin [41] and for studying the effects of single mutations in human CBX1 on its ability to bind chromatin [27]. Vogel and colleagues elegantly show that single specific mutations cause a partial loss of CBX1 binding whilst the double mutation causes a total loss. This principal of modifying, mutating or truncating a Dam fusion protein can be utilized to study indirect protein–DNA interactions and also for the identification of co-factors.

As previously mentioned, the current use of DamID in transgenic organisms involves a continuous ubiquitous low level expression of the fusion protein (due to the high activity of the methyltransferase). In *Drosophila*, the UAS/GAL4 system [29, 42] provides a high degree of flexibility for expressing transgenes both spatially and temporally. The levels of expression generated by this system are too high for Dam fusion proteins; if driven with GAL4, the background methylation is too high and is toxic to the flies (T. Southall and A. Brand, unpublished data). However, a method to combine DamID and the versatility of the GAL4 system would be a welcome advancement in the field, possibly by altering the activity of the Dam or by reducing the relative amount of protein expressed. This would allow chromatin profiling experiments to be performed in specific tissues or even within specific cell types without the need for dissection or a specific antibody.

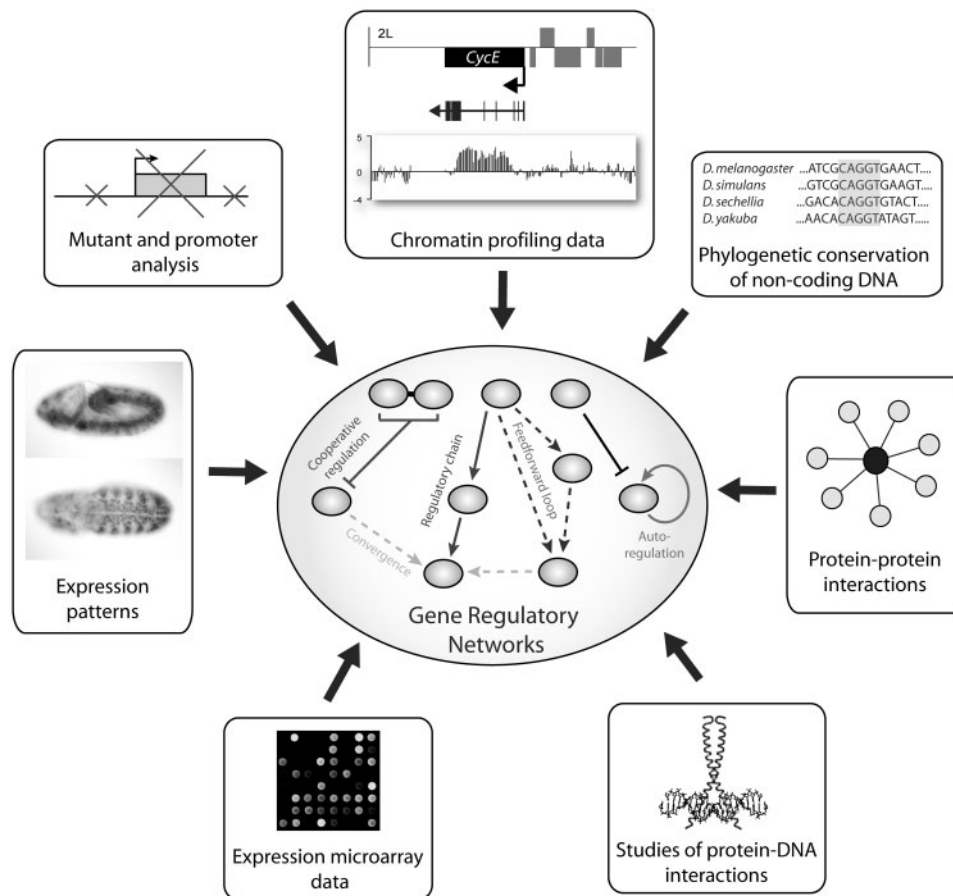
## BUILDING GENE REGULATORY NETWORKS

The wealth of expression array and chromatin profiling data is permitting researchers to build and study epigenetic and genetic regulatory networks [43]. Much of the groundwork in this field has been done in yeast [3, 5, 6, 44] elucidating networks that include autoregulatory loops, feed-forward loops, convergent targets and regulatory chains (Figure 2). These network features originate from transcription factors regulating their own expression and the expression of other transcription factors. To build these networks in model organisms, data can be collated from chromatin profiling

experiments, expression arrays, expression patterns, binding-site motif discovery, mutant analysis and protein–protein interaction data (Figure 2). The sheer quantity of data involved provides a considerable bioinformatic challenge. Strong collaborations between biologists and bioinformaticians will be essential for integrating and optimizing the potential of these data.

Although studies of yeast regulatory networks have been invaluable, there lie great challenges and rewards in mapping transcriptional networks in metazoans. In model organisms, such as mouse, *Drosophila* and *C. elegans*, transcriptional networks controlling development and disease can be studied, however this does mean working with larger, more complicated genomes. There are approximately 260 yeast transcription factors, some of which have the potential to bind to up to 200 target promoters [44]. This in itself permits the existence of highly complex networks, but is put into perspective, in considering that in humans (where intergenic sequences form a large part of the genome) there are nearly 2000 transcription factors [45] of which some, such as CREB, are predicted to bind to 19 000 loci [46]. Despite this enormous challenge, progress has already been made by the ENCODE (ENCyclopedia Of DNA Elements) Project [47, 48]. This project aims to identify all functional elements in the human genome by the generation and integration of functional data from numerous and diverse experiments. These experiments include ChIP–chip examination of specific factors and chromatin modifications, full genome transcriptome analysis and genome sequence comparisons with other species. The ENCODE pilot study, covering 1% of the genome, is now completed [48] and has enhanced the knowledge of human genome function in several key areas.

Studies of gene regulatory networks in *C. elegans* have been facilitated by the development of high-throughput, gene-centred protein–DNA interaction mapping methods. Using a Gateway compatible yeast one-hybrid system [49], Deplancke and colleagues [50] were able to map the 283 interactions of 117 proteins with 72 *C. elegans* digestive tract gene promoters. In parallel, the *C. elegans* Promoterome is being developed [51], which presently consists of 6000 promoters cloned into Gateway vectors. This provides an excellent resource for expression pattern analysis,



**Figure 2:** Building gene regulatory networks in model organisms. Data from different experimental approaches can be used to assemble gene regulatory networks. Chromatin profiling provides positional information, expression arrays and expression patterns can give a direction of regulation, protein–protein interactions (e.g. from yeast-2-hybrid screens) can identify factors that work cooperatively to regulate expression. *In vitro* protein–DNA interaction studies and conservation of non-coding DNA can identify transcription factor binding motifs. Also, analysis of mutations in genes encoding proteins that regulate transcription and the manipulation of promoter sequences provide important information for building these networks.

yeast one-hybrid and promoter manipulation experiments.

Together with projects like ENCODE, the discovery and mapping of gene regulatory networks in model organisms will make an important contribution to deciphering human regulatory networks and for their future manipulation for therapeutic purposes. Chromatin profiling methods and microarray technologies are helping to accelerate the studies of transcriptional control and regulatory networks. However, the improvement of these techniques is making the acquisition of data less of a limiting factor. The next great challenge will be the interpretation of this information and its integration with other data sets.

### Key Points

- ChIP-chip and DamID are techniques to map the interaction of proteins with DNA.
- Combining ChIP or DamID with high density DNA microarrays enables mapping of chromatin factors on a genome-wide scale.
- Chromatin profiling data can be integrated with other data sets, such as expression profiling microarrays and mutant analysis, to build gene regulatory networks in model organisms.

### References

1. Solomon MJ, Larsen PL, Varshavsky A. Mapping protein–DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. *Cell* 1988;**53**:937–47.
2. Hecht A, Strahl-Bolsinger S, Grunstein M. Spreading of transcriptional repressor SIR3 from telomeric heterochromatin. *Nature* 1996;**383**:92–96.

3. Ren B, Robert F, Wyrick JJ, *et al.* Genome-wide location and function of DNA binding proteins. *Science* 2000;**290**: 2306–09.
4. Iyer VR, Horak CE, Scafe CS, *et al.* Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* 2001;**409**:533–8.
5. Harbison CT, Gordon DB, Lee TI, *et al.* Transcriptional regulatory code of a eukaryotic genome. *Nature* 2004;**431**: 99–104.
6. Simon I, Barnett J, Hannett N, *et al.* Serial regulation of transcriptional regulators in the yeast cell cycle. *Cell* 2001; **106**:697–708.
7. Glynn EF, Megee PC, Yu HG, *et al.* Genome-wide mapping of the cohesin complex in the yeast *Saccharomyces cerevisiae*. *PLoS Biol* 2004;**2**:E259.
8. Schubeler D, MacAlpine DM, Scalzo D, *et al.* The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev* 2004;**18**:1263–71.
9. Liu CL, Kaplan T, Kim M, *et al.* Single-nucleosome mapping of histone modifications in *S. cerevisiae*. *PLoS Biol* 2005;**3**:e328.
10. Raisner RM, Hartley PD, Meneghini MD, *et al.* Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. *Cell* 2005;**123**:233–248.
11. Meier A, Fiegler H, Munoz P, *et al.* Spreading of mammalian DNA-damage response factors studied by ChIP-chip at damaged telomeres. *EmboJ* 2007;**26**:2707–18.
12. Wardle FC, Odom DT, Bell GW, *et al.* Zebrafish promoter microarrays identify actively transcribed embryonic genes. *Genome Biol* 2006;**7**:R71.
13. Ercan S, Giresi PG, Whittle CM, *et al.* X chromosome repression by localization of the *C. elegans* dosage compensation machinery to sites of transcription initiation. *Nat Genet* 2007;**39**:403–8.
14. Birch-Machin I, Gao S, Huen D, *et al.* Genomic analysis of heat-shock factor targets in *Drosophila*. *Genome Biol* 2005;**6**: R63.
15. Zeitlinger J, Zinzen RP, Stark A, *et al.* Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the *Drosophila* embryo. *Genes Dev* 2007;**21**: 385–90.
16. Odom DT, Dowell RD, Jacobsen ES, *et al.* Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat Genet* 2007;**39**:730–2.
17. Thibaud-Nissen F, Wu H, Richmond T, *et al.* Development of *Arabidopsis* whole-genome microarrays and their application to the discovery of binding sites for the TGA2 transcription factor in salicylic acid-treated plants. *PlantJ* 2006;**47**:152–62.
18. van Steensel B, Henikoff S. Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. *Nat Biotechnol* 2000;**18**:424–8.
19. van Steensel B, Delrow J, Henikoff S. Chromatin profiling using targeted DNA adenine methyltransferase. *Nat Genet* 2001;**27**:304–8.
20. Greil F, van der Kraan I, Delrow J, *et al.* Distinct HP1 and Su(var)3-9 complexes bind to sets of developmentally coexpressed genes depending on chromosomal location. *Genes Dev* 2003;**17**:2825–38.
21. Greil F, Moorman C, van Steensel B. DamID: mapping of in vivo protein-genome interactions using tethered DNA adenine methyltransferase. *Methods Enzymol* 2006;**410**: 342–59.
22. Orian A, van Steensel B, Delrow J, *et al.* Genomic binding by the *Drosophila* Myc, Max, Mad/Mnt transcription factor network. *Genes Dev* 2003;**17**:1101–14.
23. Bianchi-Frias D, Orian A, Delrow JJ, *et al.* Hairy transcriptional repression targets and cofactor recruitment in *Drosophila*. *PLoS Biol* 2004;**2**:E178.
24. Sun LV, Chen L, Greil F, *et al.* Protein-DNA interaction mapping using genomic tiling path microarrays in *Drosophila*. *Proc Natl Acad Sci USA* 2003; **100**:9428–33.
25. de Wit E, Greil F, van Steensel B. Genome-wide HP1 binding in *Drosophila*: developmental plasticity and genomic targeting signals. *Genome Res* 2005;**15**: 1265–73.
26. de Wit E, Greil F, van Steensel B. High-resolution mapping reveals links of HP1 with active and inactive chromatin components. *PLoS Genet* 2007;**3**:e38.
27. Vogel MJ, Guelen L, de Wit E, *et al.* Human heterochromatin proteins form large domains containing KRAB-ZNF genes. *Genome Res* 2006;**16**:1493–504.
28. Song S, Cooperman J, Letting DL, *et al.* Identification of cyclin D3 as a direct target of E2A using DamID. *Mol Cell Biol* 2004;**24**:8790–802.
29. Brand AH, Perrimon N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 1993;**118**:401–15.
30. Choksi SP, Southall TD, Bossing T, *et al.* Prospero acts as a binary switch between self-renewal and differentiation in *Drosophila* neural stem cells. *Dev Cell* 2006;**11**:775–89.
31. Pym EC, Southall TD, Mee CJ, *et al.* The homeobox transcription factor Even-skipped regulates acquisition of electrical properties in *Drosophila* neurons. *Neural Develop* 2006;**1**:3.
32. Germann S, Juul-Jensen T, Letarnc B, *et al.* DamID, a new tool for studying plant chromatin profiling in vivo, and its use to identify putative LHP1 target loci. *PlantJ* 2006;**48**: 153–63.
33. Pease AC, Solas D, Sullivan EJ, *et al.* Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc Natl Acad Sci USA* 1994;**91**:5022–6.
34. Singh-Gasson S, Green RD, Yue Y, *et al.* Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. *Nat Biotechnol* 1999;**17**: 974–8.
35. Kim TH, Barrera LO, Zheng M, *et al.* A high-resolution map of active promoters in the human genome. *Nature* 2005;**436**:876–80.
36. Boyer LA, Lee TI, Cole MF, *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 2005;**122**:947–56.
37. Bulyk ML. DNA microarray technologies for measuring protein-DNA interactions. *Curr Opin Biotechnol* 2006;**17**: 422–30.
38. Orian A. Chromatin profiling, DamID and the emerging landscape of gene expression. *Curr Opin Genet Dev* 2006;**16**: 157–64.

39. Viens A, Mechold U, Lehmann H, *et al.* Use of protein biotinylation in vivo for chromatin immunoprecipitation. *Anal Biochem* 2004;**325**:68–76.
40. van Werven FJ, Timmers HT. The use of biotin tagging in *Saccharomyces cerevisiae* improves the sensitivity of chromatin immunoprecipitation. *Nucleic Acids Res* 2006;**34**: e33.
41. Cleard F, Moshkin Y, Karch F, *et al.* Probing long-distance regulatory interactions in the *Drosophila melanogaster* bithorax complex using Dam identification. *Nat Genet* 2006;**38**:931–5.
42. McGuire SE, Roman G, Davis RL. Gene expression systems in *Drosophila*: a synthesis of time and space. *Trends Genet* 2004;**20**:384–91.
43. van Steensel B. Mapping of genetic and epigenetic regulatory networks using microarrays. *Nat Genet* 2005;**37**(Suppl):S18–24.
44. Lee TI, Rinaldi NJ, Robert F, *et al.* Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* 2002;**298**:799–804.
45. Messina DN, Glasscock J, Gish W, *et al.* An ORFeome-based analysis of human transcription factor genes and the construction of a microarray to interrogate their expression. *Genome Res* 2004;**14**:2041–7.
46. Euskirchen G, Royce TE, Bertone P, *et al.* CREB binds to multiple loci on human chromosome 22. *Mol Cell Biol* 2004;**24**:3804–14.
47. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* 2004;**306**:636–40.
48. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 2007;**447**:799–816.
49. Deplancke B, Dupuy D, Vidal M, *et al.* A gateway-compatible yeast one-hybrid system. *Genome Res* 2004;**14**:2093–101.
50. Deplancke B, Mukhopadhyay A, Ao W, *et al.* A gene-centered *C. elegans* protein-DNA interaction network. *Cell* 2006;**125**:1193–205.
51. Dupuy D, Li QR, Deplancke B, *et al.* A first version of the *Caenorhabditis elegans* Promoterome. *Genome Res* 2004;**14**: 2169–75.