

Chapter 6

USING RNA INTERFERENCE TO VALIDATE TARGET GENES IDENTIFIED BY COUPLING CHROMATIN IMMUNOPRECIPITATION WITH CpG-ISLAND MICROARRAYS

Matthew J. Oberley¹, Antonis Kirmizis²,
William R. Schelman¹, Peggy J. Farnham^{3,4}

1 University of Wisconsin Medical School, Health Sciences Learning Center
750 Highland Avenue, Madison, Wisconsin 53705-2221

2 The Wellcome Trust/Cancer Research UK
Gurdon Institute and Department of Pathology, University of Cambridge
Tennis Court Road, Cambridge, CB2 1QR, United Kingdom

3 UC Davis Genome Center, Department of Medical Pharmacology
Genome and Biomedical Sciences Facility
451 East Health Sciences Drive, Davis, CA 95616-8816

Contact:

Peggy J. Farnham, Ph.D.
UC Davis Genome Center, Department of Medical Pharmacology
Genome and Biomedical Sciences Facility
451 East Health Sciences Drive, Davis, CA 95616-8816
E-mail: pjfarnham@ucdavis.edu

6.1 Introduction

The formaldehyde-based chromatin immunoprecipitation (ChIP) assay has revolutionized the study of transcriptional mechanisms since its first demonstration in *D. melanogaster* by Solomon et al. in 1988^[1], providing a means by which protein-DNA interactions can be analyzed without artificial perturbation of the cellular milieu. Because of the large size and sequence complexity of the mammalian genome, successful ChIP of a mammalian site-specific transcription factor was not accomplished for another decade^[2,3]. Although a relatively recent addition to the repertoire of methods used to study transcriptional regulation, ChIP is now the “gold standard” by which a transcription factor’s binding to a particular promoter region can be determined. Briefly, ChIP assays are performed by treating cells with formaldehyde to physically link protein and DNA components of chromatin. An antibody to a protein of interest is then used to immunoprecipitate the DNA to which the protein is crosslinked. The immunoprecipitated DNA is then analyzed by PCR using primers specific for certain promoters that are thought to be good candidate target genes. However, a significant disadvantage of ChIP-PCR is that only a limited number of genes can be studied because informed guesses must be made about where a particular factor binds in the genome. Scanning ChIP, a procedure in which the ChIP DNA pool is examined by PCR using primers located at prescribed intervals throughout a genomic region, can partially alleviate these problems^[4-6]. However, this approach is very labor intensive and can cover only a fraction of the genome. Thus, although the ChIP assay allows investigators to study DNA/protein interactions in living cells, the standard methods of analysis significantly limit the scope of the experimental design.

To overcome the limitations of standard ChIP assays, high throughput microarray-based interrogation of the immunoprecipitated DNA (ChIP-chip) was developed in the laboratory of Richard Young^[7] and has been used extensively for the past several years to analyze transcriptional regulation in yeast model systems. For example, ChIP-chip in yeast was used to identify promoters bound by specific transcription factors and to correlate actively transcribed regions with specific histone post-translational modifications^[7-15]. More recently, the ChIP-chip technique has been adapted to mammalian systems. To this end, arrays have been spotted with PCR products that either span specific regions like the β -globin locus or that represent collections of promoter regions for known genes^[13,16-20] or CpG islands^[6,21-27]. Further, oligonucleotide tiling array platforms have also been developed to identify the chromosomal binding locations of transcription factors such as Myc, Sp1, p53, and SUZ12^[6,28].

The combination of chromatin immunoprecipitation and genomic microarrays has become a powerful tool that provides a means of rapidly identifying a large number of target genes of a particular transcription factor. Recent studies using such ChIP-chip approaches have revealed hundreds of target genes for certain transcription fac-

tors, suggesting that up to 5-10% of mammalian promoters may be bound by a given transcription factor^[16,20,21,24]. Somewhat surprisingly, when entire chromosomes were analyzed, the number of target sites grew even larger due to the finding that the transcription factors bound not only to promoter regions, but also to transcribed regions and to 3' ends of genes. Extrapolation of this data to the entire genome suggests that thousands of binding sites might be occupied by certain transcription factors. Given these results, it becomes critical to ensure that the identified binding regions are: 1) reproducibly identified by the microarray analyses, 2) specifically precipitated by antibody-epitope interactions, and 3) biologically relevant sites.

The issues surrounding the reproducibility of array hybridization are related to the algorithms by which "positive hits" are determined during the data analysis. The significance of enrichment (as identified by increased hybridization intensity) of any particular segment of DNA on a DNA microarray can be measured using a single array error model^[7], a median percentile ranking^[8], or a sliding window approach^[28]. Each of these methods has advantages and disadvantages, the discussion of which is beyond the scope of this chapter (see Buck and Lieb, 2004^[29] for more information). Regardless of the method used to identify positive spots, it is essential to demonstrate that these spots represent DNA fragments that have been enriched during the ChIP procedure. Confirmation of DNA-factor binding is typically performed using gene-specific primers and quantitative or semi-quantitative PCR on independent ChIP samples. This follow-up PCR analysis validates that the experimental and computational protocols that were used can accurately measure the relative amounts of sequences in a pool of antibody-precipitated DNA. However, this type of follow-up analysis does not directly address issues related to antibody specificity or biological importance of the target. This chapter will focus on an expanded set of follow-up analyses which can be performed to demonstrate that the precipitated samples are in fact bound by the transcription factor in the cell and that the binding is biologically relevant. To this end, we propose that RNA interference (RNAi) represents an excellent method for both target validation as well as investigation of the biological significance of the observed binding.

6.2 Target validation

As noted above, confirming the reproducibility of precipitation of a given genomic fragment is a critical first step in target gene identification. However, the ChIP assay may non-specifically enrich certain regions within complex mammalian genomes. This could be due to non-specific interactions of certain genomic DNA segments with experimental reagents (such as Staph A cells, magnetic beads, and/or the antibody itself) or to non-specific antibody-antigen interactions. Past studies have used artificial systems to demonstrate that the targets identified via ChIP-chip are bound to and regulated by the transcription factor under study. For example, electrophoretic mobility shift assays have been used as an independent means to exam-

ine transcription factor binding to identified sequences. While this *in vitro* assay is able to demonstrate that a factor can bind to a given sequence, it does not address the role that chromatin, protein-protein interactions, stoichiometry, and the cellular milieu may play in the recruitment of a factor to DNA. Thus, the gel shift assay does not accurately reflect *in vivo* conditions and may lead to an excess of false negatives and/or false positives. Another method used to validate targets is to clone the identified promoter sequences into a reporter plasmid, co-express the transcription factor and reporter construct, and determine if the promoter is activated or repressed by the factor. In this case, the promoter being studied is not located in its normal chromatin context and expression of the transcription factor at supra-physiological levels might lead to non-specific effects (e.g. via squelching of co-activators/repressors). Thus, while these studies might indicate that a specific protein-DNA interaction may occur in an isolated DNA fragment, they do not address whether the interactions and their regulation occur at the endogenous loci. In addition, because many factors operate as obligate complexes for activity, false negatives may result from stoichiometric imbalances due to over-expression of a single factor. For these reasons, we believe that other assays are best for validating the results of ChIP-chip experiments.

Strategies to assess the validity of ChIP-chip results can be grouped into two classes: 1) those in which changes are made in the specific antibody used in the ChIP assay and 2) those in which functional consequences of altering the expression of the target transcription factor are analyzed. Examples of the first class of approaches include experimental alterations such as the omission of the primary antibody from the assay (sometimes called a “No Antibody” control or, if only a secondary antibody is used, an “IgG” control), the use of an antibody that recognizes an epitope not present in any cellular protein (e.g. a HA tag for mammalian cells or a Myc tag for yeast cells), the use of multiple antibodies which each recognize a different epitope on the same factor (e.g. [20]), and the use of a blocking peptide to obstruct the antibody’s epitope recognition site prior to the immunoprecipitation[23]. Although these are all important controls that can be utilized to identify false positives based on nonspecific interactions, they do not address the biological significance of the true positives. The second class of validation strategies relies on the experimental alteration of the expression level of the factor of interest. This has been accomplished using a number of methods including: i) comparing immunoprecipitated DNA from unrelated cell lines or tissue types where one cell or tissue expresses the factor of interest and the other does not[30]; ii) comparing mouse embryo fibroblasts in which the factor of interest has been deleted relative to the parental strain[31]; iii) comparing cell lines in which the abundance of a factor has been modulated through either receptor-mediated activation of signal transduction cascades or through experimentally-mediated differentiation (e.g. [24]); and iv) examining the effect of a factor following its removal through the use of RNA interference (RNAi) ([6,23,32-34].

Studies comparing samples obtained from cells or tissues that differ in the abundance of a factor of interest can not only address the issue of “false positives” but may also provide insight into the biological significance of the factor/promoter interactions. Indeed, several investigators have used this technique to correlate the binding of a transcription factor at a particular promoter to the mRNA levels produced from that same locus in the presence versus absence of the factor^[16,19, 24,25]. An example of this was demonstrated in an elegant study by *Li et al.*^[13] who identified Myc binding sites in the human genome and then examined the expression levels of 453 of the identified targets in 46 different tissues and cell lines. The results of their exhaustive efforts showed a striking correlation between expression levels of the identified Myc targets and the reported levels of Myc mRNA in each of the tissues or cell lines. However, not all tissues with significant levels of Myc mRNA had the same patterns of target gene expression, nor was location analysis performed on each tissue or cell line. Therefore, while compelling, this study did not definitively prove that Myc binding is directly responsible for the differences in mRNA levels. For example, Myc could regulate other transcription factors that then directly regulate certain of the mRNAs examined. Further, Myc may be bound at a certain promoter but not functionally affect the activity of the general transcriptional machinery at that promoter. In a related study, *Mao et al.*^[24] validated Myc targets identified with ChIP-CpG arrays using the HL60 cell line. Differentiation of these cells with DMSO results in a large reduction in the levels of the Myc protein, thus allowing a comparison of Myc binding and mRNA expression level of the putative target genes. Using this model system, they showed a correlation between Myc expression and its binding at the target loci as well as a correlation between Myc binding and transcriptional activation of the identified target genes.

While the studies described above utilized ideal cell culture systems to study transcription factor binding and activity under physiological conditions, there are, unfortunately, only a handful of such cell lines in which the levels of known transcription factors have been characterized during differentiation. In addition, the differentiation process alters the expression of many transcriptional regulators, thus resulting in changes in transcriptional activation that are unrelated to the factor of interest. To circumvent this problem, it would be ideal to compare two cells or tissues which differ only in the expression of a single transcription factor. One way to achieve this is to use mice which are nullizygous for a certain transcription factor. However, this method is not without its own drawbacks since the nullizygous embryos undergo selective pressures throughout development and compensatory changes in gene expression may occur due to the loss of the factor. Further, human factors cannot be studied using this method. Although human factors can be targeted in cultured cells using homologous recombination, this method requires significant effort to create targeting vectors and to eliminate both alleles encoding the factor.

In contrast to the methods described above which are often labor-intensive, RNA interference (RNAi) represents a fairly simple and powerful reverse genetic method for the specific elimination of a single transcription factor. This technique is performed by introducing small interfering RNA (siRNA) duplexes into cells, inducing the degradation of mRNAs having sequence identity with the siRNAs^[35]. Although conceptually similar to using homologous recombination to eliminate specific transcription factors in mouse model systems, RNAi provides several advantages over the use of nullizygous mice. For example, transient reduction of mRNA and protein levels using RNAi can be accomplished rapidly in cell culture, with >90% reduction of the target protein in 3 to 6 days. Also, plasmids coding for short hairpin (sh) RNAs produced using an RNA Polymerase III promoter can be utilized to produce stable knock-downs of target proteins (e.g. ^[33]). Although in the following sections we will focus on the use of RNAi to validate potential array targets, a detailed discussion of RNAi is beyond the scope of this chapter. For a complete overview of the theory and methodology of RNAi, please refer to Hannon, 2002^[36]; McManus and Sharp, 2002^[37]; and Novina and Sharp, 2004^[35].

6.3 Use of RNA interference to validate novel targets and to study in vivo transcriptional mechanisms.

Many tools exist to assist in the design of effective siRNAs and shRNAs, such as http://www.protocolonline.org/prot/Research_Tools/Online_Tools/SiRNA_Design/. For designing siRNAs, we have adhered to the following basic requirements: i) the duplex RNA should be 21 nucleotides in length with a 3' overhang of TT to avoid exonuclease digestion; ii) the sequence chosen should avoid: introns, the 5' and 3' untranslated regions, and regions of the mRNA within 75 bp of the start codon; iii) the sequence should be low to medium GC content (30-50%), with the 5' end of the sense strand containing a higher GC content for stability, and the 3' end containing a higher AU content to promote unwinding by RISC (RNA Induced Silencing Complex); iv) it has been suggested that there should be a G or C at position 1 of the sense strand, a U at position 10 of the sense strand, and an A at position 19 of the sense strand (see Mittal, 2004 for details); and iv) to avoid off target effects, the chosen sequence(s) should be BLASTed against the mammalian genome under study to ensure the chosen region is unique.

The efficacy of any given short interfering RNA (siRNA) is determined not only by its effectiveness in mediating the degradation of the target mRNA but also by the efficiency by which it is delivered to the cell. To ensure optimal transfection efficiency and to avoid off-target effects, it is important to optimize the transfection conditions for each duplex RNA construct in every cell type used. Transient localization of siRNAs into cells can be accomplished either through electroporation or by using lipophilic carrier agents. Introduction of the siRNAs into cells can be monitored by labeling the siRNA with a fluorophore and then assessing the uptake via flow

cytometry. Using this technique, have found that certain transfection reagents work more effectively than others in a given cell line. For example, as shown in Figure 1, the siRNAs were efficiently transported into the SW480 colon cancer cells using either reagent 1 or reagent 2, but only reagent 2 was effective in introducing the siRNAs into the HT29 colon cancer cells. Due to differences in transfection efficiency, it is important to carefully titrate the amount of siRNA used in each experiment so that a high enough concentration is obtained for effective degradation of the target mRNA. However, introduction of excessive amounts of siRNA is not desirable. Semizarov *et al.*[38] found that transfection of human cells with concentrations of siRNAs of 100nM or more resulted in the reproducible activation of stress response genes such as *gadd45* and *bcl2*. Activation of a cellular stress response would confound the interpretation of the gene expression changes observed upon introduction of the siRNAs. Therefore, to avoid off-target effects, many investigators recommend using concentrations of siRNA that do not exceed 20nM. In addition, it is important to include a control siRNA to monitor non-specific effects. The control siRNA can either be a scrambled version of the siRNA used to target the factor or a sequence specific to an mRNA that is not produced in the cells (e.g. GFP in mammalian cells). A more rigorous approach towards eliminating the potential for off-target effects of an siRNA involves using two RNA duplexes specific for different areas of the target mRNA[39].

To overcome some of the difficulties associated with transient introduction of siRNAs including high cost, off-target effects and limited time periods for assays, constitutive expression of shRNAs using the RNA Polymerase III, U6, or H1 small nuclear RNA promoters can be employed[40]. While the introduction of plasmids encoding shRNAs can usually be easily accomplished using traditional transfection techniques, transfection of certain cell types, including stem cells, has proven to be more difficult. For this reason, several methods based on viral transduction have been established including the use of the Molony murine leukemia virus, the murine stem cell virus, or the lentiviral based systems. For proteins whose long-term loss would be detrimental to cell survival, inducible shRNA systems have also been developed to allow the controlled removal of the target protein (for example, see [41]).

After the introduction of the siRNA into the cell is optimized, it is critical to assess whether the constructs are effective in reducing the mRNA and protein levels of the target. Protein levels are usually compared using Western analysis in treated and untreated cells. As shown in *Figure 6-1B*, the level of SUZ12 protein was not reduced using reagent 1 (as expected since this reagent was not effective in introducing the siRNAs into the HT29 cells; see *Figure 6-1A*). However, the levels of SUZ12 were reduced using reagent 2, demonstrating that the introduced siRNAs were effective in mediating the degradation of the SUZ12 mRNA. Once it has been demonstrated that the introduction of the siRNA into the cells reduces the level of the transcription factor of interest, the technique can then be combined with ChIP for target val-

idation. For example, introduction of a siRNA or shRNA specific to a transcription factor can be used to verify that the antibody used in the ChIP is specific for that protein. This is especially important when studying a transcription factor that is a member of a highly related family of factors because the antibody might show cross-reactivity with other family members. However, if targeting the mRNA of the factor by siRNA results in knock-down of only that family member and in loss of signal detected by ChIP using an antibody to that protein, then clearly the signal detected is specific for the protein of interest and not another family member. Using this method, we have found that siRNA specific to E2F6 abrogated the ability of E2F6 antibodies to enrich target promoters (Figure 6-2B), but did not prevent E2F1 antibodies from enriching targets^[23]. In this case, RNAi-ChIP provided a clear demonstration of antibody specificity in the ChIP assay. In addition, the experiments also demonstrated that removal of E2F6 by siRNA resulted in altered mRNA levels of the target genes (Figure 6-2A). Thus, the combination of RNAi, ChIP, and RT-PCR provide target validation and confirmation of functional binding.

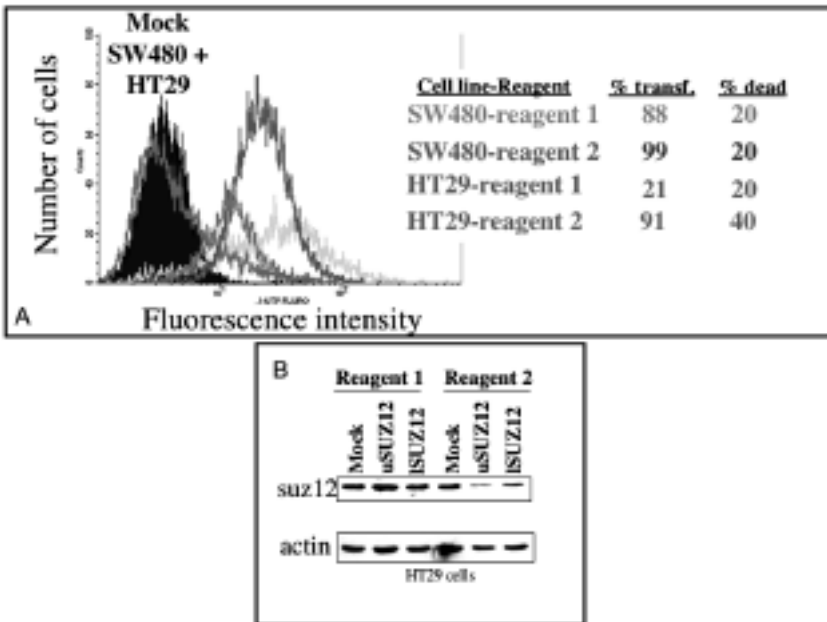


Figure 6-1. Correlation of transfection efficiency and RNAi-mediated protein knockdown. **A.** Flow cytometric analysis of SW480 and HT29 cells either mock transfected (purple line), or transfected with a duplex RNA specific for human SUZ12 that was labeled with a fluorophore (Label IT siRNA Tracker, Mirus Co., Madison, WI) to measure cellular uptake of the siRNA. SW480 cells were effectively transfected with both transfection reagents tested (green and pink lines), while HT29 cells were only effectively transfected with reagent 2 (light gray and dark gray lines). **B.** Western analysis of SUZ12 protein levels in HT29 cells after transfection with unlabeled (uSUZ12) and labeled (ISUZ12) SUZ12 siRNA. Details of the specific cell-culture, duplex RNA sequence, transfection procedures, and western analysis are described in *Kirmizis et al.*, 2004.

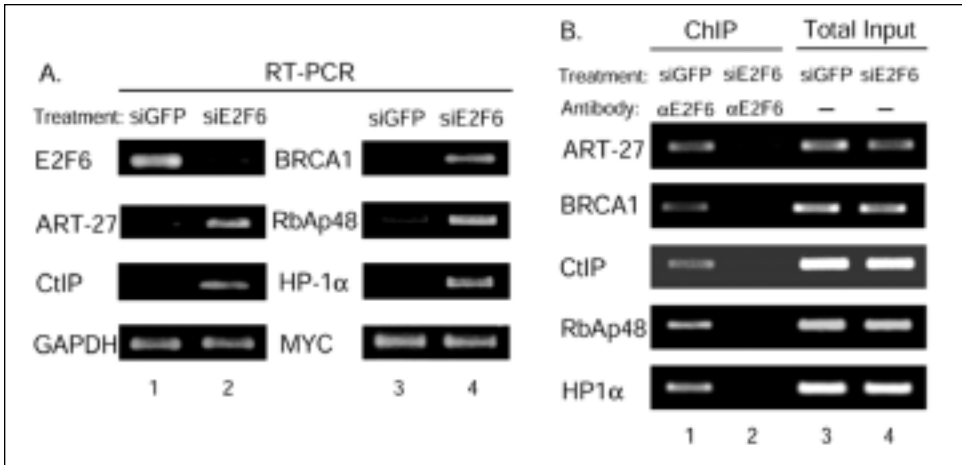


Figure 6-2. Use of RNAi to validate target genes identified via ChIP-chip assays. **A.** RNAi followed by RT-PCR demonstrates that depletion of E2F6 from cells results in the increased transcription from ChIP-CpG-identified targets art27, brca1, ctip, rbap48, and hp1alpha. **B.** RNAi targeting of E2F6 followed by chromatin immunoprecipitation prevents antibody-mediated enrichment of E2F6 target sequences, demonstrating antibody specificity. Figure reproduced with permission from Oberley et al., 2003. Details of specific cell-culture, RT-PCR, transfection procedures, and ChIP analysis are described in Oberley et al., 2003.

In addition to validation of the targets identified by ChIP-CpG arrays, RNAi-ChIP can also be used to study the mechanisms by which the transcription factor of interest regulates specific target genes. The post-translational covalent modification of histones is an important characteristic of chromatin structure that is associated with gene expression. For example, methylation of lysine 27 on histone H3 has been associated with transcriptional repression^[6,32]. The mammalian Polycomb Repressive Complex 2 (PRC2) methylates lysine 27 and, to a lesser extent, lysine 9 of histone H3 in vitro^[32,42]. This complex contains enhancer of zeste 2 (EZH2), suppressor of zeste 12 (SUZ12), embryonic ectoderm development (EED), and RbAP46/48. In a recent study, we identified human targets of PRC2 by using antibodies specific to SUZ12 in ChIP-CpG assays^[6]. We were subsequently able to validate the identified targets using RNAi-ChIP by demonstrating that the CpG islands were specifically enriched as a result of the antibody-epitope interaction and that other members of the PRC2 complex were also present on the identified loci in a SUZ12-dependent manner. In addition, we also showed that lysine 27 of histone H3 was methylated at the PRC2 target genes and that this methylation was dependent on the presence of SUZ12 (Figure 6-3 and Kirmizis et al., 2004^[6]). Thus, RNAi was used not only to validate the ChIP-CpG targets, but also to elucidate a functional outcome (histone methylation) of the SUZ12 binding.



Figure 6-3. RNAi followed by ChIP allows the identification of specific histone modifications mediated by SUZ12. SW480 cells were either mock-transfected or transfected with SUZ12 siRNA and allowed to incubate for 72 hours. Cells were then re-plated and re-transfected with either SUZ12 siRNA or no RNA duplex for another 72 hours and then harvested for ChIP analysis using antibodies specific to tri-methyl lysine 27 of histone H3, Histone H3, or IgG. Depletion of SUZ12 results in the specific depletion of the tri-methylated lysine 27 signal, while the overall amount of histone H3 on the target genes remains unchanged. Details of the specific cell-culture, duplex RNA sequence, transfection procedures, and ChIP analysis are described in *Kirmizis et al., 2004*.

6.4 Summary

Recent contributions from numerous investigators have demonstrated that ChIP-CpG microarray analysis is a powerful means to identify a large set of target genes for a specific transcription factor. The quality and reproducibility of the hybridization methodology has eliminated the need to individually analyze all of the identified targets using subsequent standard ChIP assays. However, we suggest that the initial identification of target promoters should be validated, not with a repetition of the simple ChIP assay, but rather by the addition of an assay which can allow evaluation of the functional consequences of the DNA-protein interactions. The inclusion of RNAi in the experimental design provides a mechanism to monitor the effects of removing the factor from a binding site on gene expression and chromatin structure. To date, these follow-up assays have been performed on only a subset of the target genes identified by ChIP-chip assays. However, we anticipate that future experiments will incorporate custom-designed mRNA or promoter arrays to allow a comprehensive follow-up analysis of the identified targets.

Acknowledgements

This work was supported by DOD grant (BC020760) and NIH grant R01 CA045240-16 to PJF.

References

1. Solomon MJ, Larsen PL and Varshavsky A: Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. *Cell* 1988; 53:937-947
2. Boyd KE, Wells J, Gutman J, Bartley SM and Farnham PJ: c-Myc target gene specificity is determined by a post-DNA binding mechanism. *Proc Natl Acad Sci U S A* 1998; 95:13887-13892
3. Grandori C, Mac J, Siebelt F, Ayer DE and Eisenman RN: Myc-Max heterodimers activate a DEAD box gene and interact with multiple E box-related sites in vivo. *Embo J* 1996; 15:4344-4357
4. Johnson KD, Christensen HM, Zhao B and Bresnick EH: Distinct mechanisms control RNA polymerase II recruitment to a tissue-specific locus control region and a downstream promoter. *Mol Cell* 2001; 8:465-471
5. Frank SR, Schroeder M, Fernandez P, Taubert S and Amati B: Binding of c-Myc to chromatin mediates mitogen-induced acetylation of histone H4 and gene activation. *Genes Dev* 2001; 15:2069-2082
6. Kirmizis A, Bartley SM, Kuzmichev A, Margueron R, Reinberg D, Green R and Farnham PJ: Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. *Genes Dev* 2004; 18:1592-1605
7. Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J, Schreiber J, Hannett N, Kanin E, Volkert TL, Wilson CJ, Bell SP and Young RA: Genome-wide location and function of DNA binding proteins. *Science* 2000; 290:2306-2309
8. Iyer VR, Horak CE, Scafe CS, Botstein D, Snyder M and Brown PO: Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* 2001; 409:533-538
9. Lieb JD, Liu X, Botstein D and Brown PO: Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. *Nat Genet* 2001; 28:327-334
10. Simon I, Barnett J, Hannett N, Harbison CT, Rinaldi NJ, Volkert TL, Wyrick JJ, Zeitlinger J, Gifford DK, Jaakkola TS and Young RA: Serial regulation of transcriptional regulators in the yeast cell cycle. *Cell* 2001; 106:697-708
11. Damelin M, Simon I, Moy TI, Wilson B, Komili S, Tempst P, Roth FP, Young RA, Cairns BR and Silver PA: The genome-wide localization of Rsc9, a component of the RSC chromatin-remodeling complex, changes in response to stress. *Mol Cell* 2002; 9:563-573
12. Ng HH, Robert F, Young RA and Struhl K: Genome-wide location and regulated recruitment of the RSC nucleosome-remodeling complex. *Genes Dev* 2002; 16:806-819
13. Li Z, Van Calcar S, Qu C, Cavenee WK, Zhang MQ and Ren B: A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells. *Proc Natl Acad Sci U S A* 2003; 100:8164-8169

14. Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J and Kouzarides T: Active genes are tri-methylated at K4 of histone H3. *Nature* 2002; 419:407-411
15. Zeitlinger J, Simon I, Harbison CT, Hannett NM, Volkert TL, Fink GR and Young RA: Program-specific distribution of a transcription factor dependent on partner transcription factor and MAPK signaling. *Cell* 2003; 113:395-404
16. Ren B, Cam H, Takahashi Y, Volkert T, Terragni J, Young RA and Dynlacht BD: E2F integrates cell cycle progression with DNA repair, replication, and G(2)/M checkpoints. *Genes Dev* 2002; 16:245-256
17. Horak CE, Mahajan MC, Luscombe NM, Gerstein M, Weissman SM and Snyder M: GATA-1 binding sites mapped in the beta-globin locus by using mammalian chIp-chip analysis. *Proc Natl Acad Sci U S A* 2002; 99:2924-2929
18. Martone R, Euskirchen G, Bertone P, Hartman S, Royce TE, Luscombe NM, Rinn JL, Nelson FK, Miller P, Gerstein M, Weissman S and Snyder M: Distribution of NF-kappaB-binding sites across human chromosome 22. *Proc Natl Acad Sci U S A* 2003; 100:12247-12252
19. Euskirchen G, Royce TE, Bertone P, Martone R, Rinn JL, Nelson FK, Sayward F, Luscombe NM, Miller P, Gerstein M, Weissman S and Snyder M: CREB binds to multiple loci on human chromosome 22. *Mol Cell Biol* 2004; 24:3804-3814
20. Odom DT, Zizlsperger N, Gordon DB, Bell GW, Rinaldi NJ, Murray HL, Volkert TL, Schreiber J, Rolfe PA, Gifford DK, Fraenkel E, Bell GI and Young RA: Control of pancreas and liver gene expression by HNF transcription factors. *Science* 2004; 303:1378-1381
21. Weinmann AS, Yan PS, Oberley MJ, Huang TH and Farnham PJ: Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev* 2002; 16:235-244
22. Ballestar E, Paz MF, Valle L, Wei S, Fraga MF, Espada J, Cigudosa JC, Huang TH and Esteller M: Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer. *Embo J* 2003; 22:6335-6345
23. Oberley MJ, Inman DR and Farnham PJ: E2F6 negatively regulates BRCA1 in human cancer cells without methylation of histone H3 on lysine 9. *J Biol Chem* 2003; 278:42466-42476
24. Mao DY, Watson JD, Yan PS, Barsyte-Lovejoy D, Khosravi F, Wong WW, Farnham PJ, Huang TH and Penn LZ: Analysis of Myc bound loci identified by CpG island arrays shows that Max is essential for Myc-dependent repression. *Curr Biol* 2003; 13:882-886
25. Shi H, Wei SH, Leu YW, Rahmatpanah F, Liu JC, Yan PS, Nephew KP and Huang TH: Triple analysis of the cancer epigenome: an integrated microarray system for assessing gene expression, DNA methylation, and histone acetylation. *Cancer Res* 2003; 63:2164-2171

26. Wells J, Yan PS, Cechvala M, Huang T and Farnham PJ: Identification of novel pRb binding sites using CpG microarrays suggests that E2F recruits pRb to specific genomic sites during S phase. *Oncogene* 2003; 22:1445-1460
27. Kondo Y, Shen L, Yan PS, Huang TH and Issa JP: Chromatin immunoprecipitation microarrays for identification of genes silenced by histone H3 lysine 9 methylation. *Proc Natl Acad Sci U S A* 2004; 101:7398-7403
28. Cawley S, Bekiranov S, Ng HH, Kapranov P, Sekinger EA, Kampa D, Piccolboni A, Sementchenko V, Cheng J, Williams AJ, Wheeler R, Wong B, Drenkow J, Yamanaka M, Patel S, Brubaker S, Tammana H, Helt G, Struhl K and Gingeras TR: Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell* 2004; 116:499-509
29. Buck MJ and Lieb JD: ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. *Genomics* 2004; 83:349-360
30. Oberley MJ, Tsao J, Yau P and Farnham PJ: High-throughput screening of chromatin immunoprecipitates using CpG-island microarrays. *Methods Enzymol* 2004; 376:315-334
31. Townsend MJ, Weinmann AS, Matsuda JL, Salomon R, Farnham PJ, Biron CA, Gapin L and Glimcher LH: T-bet regulates the terminal maturation and homeostasis of NK and Valpha14i NKT cells. *Immunity* 2004; 20:477-494
32. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS and Zhang Y: Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 2002; 298:1039-1043
33. Cao R and Zhang Y: SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Mol Cell* 2004; 15:57-67
34. Wang H, An W, Cao R, Xia L, Erdjument-Bromage H, Chatton B, Tempst P, Roeder RG and Zhang Y: mAM facilitates conversion by ESET of dimethyl to trimethyl lysine 9 of histone H3 to cause transcriptional repression. *Mol Cell* 2003; 12:475-487
35. Novina CD and Sharp PA: The RNAi revolution. *Nature* 2004; 430:161-164
36. Hannon GJ: RNA interference. *Nature* 2002; 418:244-251
37. McManus MT and Sharp PA: Gene silencing in mammals by small interfering RNAs. *Nat Rev Genet* 2002; 3:737-747
38. Semizarov D, Frost L, Sarthy A, Kroeger P, Halbert DN and Fesik SW: Specificity of short interfering RNA determined through gene expression signatures. *Proc Natl Acad Sci U S A* 2003; 100:6347-6352
39. Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, Li B, Cavet G and Linsley PS: Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* 2003; 21:635-637

40. Miyagishi M and Taira K: RNAi expression vectors in mammalian cells. *Methods Mol Biol* 2004; 252:483-491
41. Gupta S, Schoer RA, Egan JE, Hannon GJ and Mittal V: Inducible, reversible, and stable RNA interference in mammalian cells. *Proc Natl Acad Sci U S A* 2004; 101:1927-1932
42. Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P and Reinberg D: Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev* 2002; 16:2893-2905

