Chapter 8

Serial Analysis of Binding Elements for Transcription Factors

Jiguo Chen

Abstract

The ability to determine genome-wide location of transcription factor binding sites (TFBS) is crucial for elucidating gene regulatory networks in human cells during normal development and disease such as tumorigenesis. To achieve this goal, we developed a method called serial analysis of binding elements for transcription factors (SABE) for globally identifying TFBS in human or other mammalian genomes. In this method, a specific antibody targeting a DNA-binding transcription factor of interest is used to pull down the transcription factor and its bound DNA elements through chromatin immunoprecipitation (ChIP). ChIP DNA fragments are further enriched by subtractive hybridization against non-enriched DNA and analyzed through generation of sequence tags similar to serial analysis of gene expression (SAGE). The SABE method circumvents the need for microarrays and is able to identify immunoprecipitated loci in an unbiased manner. The combination of ChIP, subtractive hybridization, and SAGE-type methods is advantageous over other similar strategies to reduce the level of intrinsic noise sequences that is typically present in ChIP samples from human or other mammalian cells.

Key words: Serial analysis of binding elements (SABE), transcription factor binding sites (TFBS), chromatin immunoprecipitation (ChIP), subtractive hybridization, serial analysis of gene expression (SAGE), functional genomics, protein–DNA interaction, transcriptional regulation, gene expression, human genome.

1. Introduction

A major challenge in the post-genome era is to elucidate global gene transcriptional regulatory networks in human normal and cancer cells (1). Transcription factors control gene expression through binding-specific regulatory sequences and recruiting chromatin-modifying complexes and the general transcription machinery to initiate RNA synthesis (2). Alterations in gene expression required to co-ordinate various biological processes

Philippe Collas (ed.), *Chromatin Immunoprecipitation Assays,* Methods in Molecular Biology 567,
DOI 10.1007/978-1-60327-414-2_8,

such as the cell cycle and normal development and pathological states such as tumorigenesis are in part a consequence of changes in the DNA binding status of various transcription factors. Consequently, sensitive technologies, to accurately and efficiently identify bona fide regulatory elements for specific transcription factors in vivo on a genome-wide scale, will be needed to elucidate human gene regulatory networks.

Global localization analysis of binding sites for sequence-specific transcription factors in vivo can be performed using chromatin immunoprecipitation (ChIP) and determining the genomic location of the ChIP-enriched DNA by microarray hybridization (ChIP-on-chip) (3, 4). This method circumvents the limitations of traditional methods. When coupled with gene expression and other relevant information, ChIP-on-chip assays can be extremely useful in analyzing yeast transcriptional regulatory networks, in which the promoters are well characterized (1, 5). This technique has been broadly used to identify the genomic sites bound by regulators of transcription in yeast and other eukaryotic cells (6, 7). Limited analysis of human transcription factor binding sites using ChIP-on-chip strategies have also been performed with selected promoters of genes of interest (8, 9), with CpG microarrays (10) or with selected chromosomes (11). However, comparable strategy for globally analyzing binding sites of transcription factors to the human genome is currently impracticable due to the enormous size and complexity, and also because regulatory elements are often found at vast distances either upstream or downstream from the core promoter. In fact, only 20–30% of the transcription factor binding sites localize to known promoter regions (11, 12). A solution to this limitation is to use microarrays that interrogate the entire genome. Problems with such "whole-genome tiling" microarrays are cost, reproducibility, and statistical analysis (13).

To overcome these limitations and allow interrogation of entire mammalian genome in an unbiased manner, we developed a novel approach to study genome-wide location analysis of transcription factors in human genome in vivo. This technology, called serial analysis of binding elements (SABE) (12, 14), involves specific ChIP (15), enrichment of ChIP DNA by subtractive hybridization (16), and generation of sequence tags similar to serial analysis of gene expression (SAGE) (17). Similar approaches were developed independently by different groups, attesting to the utility of this approach (12, 18–22). Termed SACO (for serial analysis of chromatin occupancy) (18), STAGE (sequence analysis of genomic enrichments) (19), GMAT (genome-wide mapping technique) (21), or ChIP-PET (22), these techniques including SABE circumvent the need for microarrays to identify immunoprecipitated loci. Compared with tiling genomic microarrays, these methods are considerably more affordable. Although

whole-genome tiling arrays will undoubtedly become less expensive, this cost differential is likely to continue for the foreseeable future.

Our approach for generating sequence tags using SABE is different from those similar techniques (SACO, STAGE, or GMAT) in that SABE tags are generated as random 18-mers produced from ChIP DNA fragments. The advantage of this is that the "tag resolution" is not limited by the presence of a four-cutter restriction enzyme site in the ChIP DNA that is used to "anchor" the tags, which makes this technology truly unbiased. Moreover, SABE does not require cloning, re-cloning, and library construction steps of ChIP DNA described in ChIP-PET method (22), which are labor- and time-consuming and cause potential bias. In addition, our technique recruits a subtractive hybridization step, which is essential to reduce the intrinsic noise resulting from isolation of repetitive sequences during ChIP in mammalian cells (23).

2. Materials

2.1. Plasmids

1. Plasmids pTet-Off and pTRE2hyg are used for the construction of tetracycline-inducible cell line expressing transcription factor of interest. Both plasmids are available from Clontech (cat. No. 631017 and 631014, respectively).
2. Plasmid p3FLAG is a mammalian vector for stable expression of fusion protein with a triple FLAG epitope on the N-terminal. p3FLAG was constructed by inserting a triple FLAG epitope (5′-CTAGACC ATG GAC TAC AAA GAC CAT GAC GGT GAT TAT AAA GAT CAT GAC ATC GAT TAC AAG GAT GAC GAT GAC AAG-3′) (start code underlined) into *Nhe*I site of pcDNA3.1/*myc*-His(-)B (Invitrogen cat. No. V855-20). p3FLAG also has c-*Myc* and 6-His epitopes on its C-terminal to meet different purposes. Two similar plasmids to p3FLAG are commercially available (p3xFLAG-CMV^{TM}-10 for N-terminal Met-3xFLAG expression and p3xFLAG-*myc*-CMV^{TM}-26 for N-terminal Met-3xFLAG, C-terminal c-*Myc* (dual tagged) expression, Sigma-Aldrich E4401 and E6401, respectively).
3. Plasmid pZERO-2a is a modified version of cloning vector pZERO-2 (Invitrogen cat. No. K2600-01) specific for SABE library construction. pZERO-2a was made by creating a unique *Aat*II site (GACGTC) between *Spe*I and *Eco*RI of the multiple cloning site of pZERO-2 through site-directed mutagenesis (i.e., GCCGCC to GACGTC). Like pZERO-2, plasmid

pZERO-2a allows direct selection of positive recombinants via disruption of the lethal gene, *ccd*B. Expression of *ccd*B results in the death of cells containing non-recombinant vector.

2.2. Cell Culture and Medium

1. Inducible cell line expressing transcription factor of interest tagged with 3xFLAG epitope.
2. RPMI 1640 or Dulbecco Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS).
3. Doxycycline (Sigma, St. Louis, MO) is dissolved in water at 50 mg/mL, stored in aliquots at 4°C, and used in cell culture at a concentration of 1 μg/mL.

2.3. Reagents

1. Anti-FLAG M2 affinity gel (Sigma-Aldrich, cat. No. A2220).
2. Normal mouse IgG-agarose (Sigma-Aldrich, cat. No. A0919).
3. Yeast tRNA (1 μg/μL) (Invitrogen, cat. No. # 15401-029).
4. Protease inhibitor cocktail (PIC, 100X, Sigma-Aldrich, cat. No. P8340).
5. RNase A (20 μg/μL, Invitrogen, cat. No. 12091-021).
6. Proteinase K (20 μg/μL, Invitrogen, cat. No. 25530-049).
7. Phenol:chloroform:isoamyl alcohol mixture (25:24:1).
8. Chloroform:isoamyl alcohol mixture (24:1).
9. QIAquick PCR purification kit (Qiagen, cat. No. 28106).
10. Micro Bio-Spin Chromatography Column (Bio-Rad, cat. No. 732-6204).
11. DNA polymerase I, Klenow fragment (NEB, cat. No. M0210L).
12. T4 DNA ligase (NEB, cat. No. #M0202L).
13. Taq DNA polymerase (NEB, cat. No. #M0267L).
14. MmeI (NEB, cat. No. #R0637L).
15. TaiI (Fermentas, cat. No. #ER1142).
16. AatII (NEB, cat. No. #R0117L).
17. 30% acrylamide (29:1) (Bio-Rad, cat. No. 161-0121).
18. 10 bp DNA ladder (Invitrogen, cat. No. 10821-015).
19. SYBR green I nuclear acid gel stain (Invitrogen, cat. No. S7567).
20. Dynabeads M-280 streptavidin (Invitrogen, cat. No. 112-05D).

2.4. Buffers

1. 10X PBS: 80 g NaCl, 2 g KCl, 14.4 g $Na_2HPO_4 \cdot 7H_2O$, 2.4 g KH_2PO_4, H_2O to 1 L. Adjust pH to 7.2, autoclave, and store at RT.
2. Hypotonic buffer: 10 m*M* HEPES, pH 7.4, 10 m*M* KCl, 1.5 m*M* $MgCl_2$, 1X PIC. Add PIC fresh before use.

3. ChIP lysis buffer: 50 m*M* HEPES, pH 7.4, 150 m*M* NaCl, 1 m*M* EDTA, 1% Triton X-100, 1X PIC. Add PIC fresh before use.
4. ChIP high salt buffer: 50 m*M* HEPES, pH 7.4, 500 m*M* NaCl, 1 m*M* EDTA, 1% Triton X-100, 1X PIC. Add PIC fresh before use.
5. ChIP wash buffer: 50 m*M* HEPES, pH 7.4, 250 m*M* LiCl, 1 m*M* EDTA, 1X PIC. Add PIC fresh before use.
6. TE buffer: 10 m*M* Tris-HCl, pH 8.0, 1 m*M* EDTA, 1X PIC. Add PIC fresh before use.
7. Elution buffer: 50 m*M* Tris-HCl, pH 8.0, 10 m*M* EDTA, 1% SDS.
8. 5X Hybridization buffer: 2.5 *M* NaCl, 250 m*M* HEPES, pH 8.3, 1 m*M* EDTA.
9. 10X TBE buffer: 890 m*M* Tris-HCl, pH 8.3, 890 m*M* boric acid, 20 m*M* EDTA.
10. PAGE gel diffusion buffer: 0.5 *M* ammonium acetate, 10 m*M* magnesium acetate, 1 m*M* EDTA, pH 8.0, 0.1% SDS.
11. 2X wash/binding buffer: 2 *M* NaCl, 10 m*M* Tris-HCl, pH 7.5, 1 m*M* EDTA.

2.5. Linkers and Primers

1. Linker LK-A: sense, 5′-AGCACTCTCCAGCATAT-CACTCCAACGT-3′; Anti-sense, 5′- ACGTTGGAGTGA-TATGCTGGAGAGTGCT amino-3′.
2. Linker LK-B: sense, 5′-ACCTGCCGACTATCCAAT-CATCCAACGT-3′; Anti-sense, 5′-ACGTTGGATGATTG-GATAGTCGGCAGGT amino-3′ (*see* **Note 1**).
3. Primer-A: 5′-Biotin-AGCACTCTCCAGCATATCAC-3′.
4. Primer-B: 5′-Biotin-ACCTGCCGACTATCCAATCA-3′ (*see* **Note 2**).

3. Methods

The SABE method involves serial enzymatic reactions and DNA manipulations; therefore, a good practice is to monitor the accuracy and efficiency of each step. Overall, there are several key factors to consider when performing SABE. First, for any selected transcription factor of interest, information of at least one well-defined target gene and binding site for that particular transcription factor is needed. This information of a known target gene is used to design PCR primers to monitor the efficiency of ChIP and subtractive hybridization. Without this information, it is hard to know whether the final ChIP DNA is really enriched or not

because ChIP with mammalian cells will definitely bring down a lot of background DNA. The information from in vitro binding assays (EMSA, DNA foot printing, SELEX, etc.) may not necessarily reflect in vivo binding of transcription factors and, therefore, cannot be used for this purpose. Second, it is estimated that at least 43% of the human genome is occupied by repetitive elements (24, 25). ChIP provides only a partial enrichment of specific DNAs, and consequently, the signal-to-noise ratio is too low to make direct analysis of target genes practical. To address this problem, SABE employs a subtractive hybridization step modified from representational difference analysis (16) that enables selective amplification of ChIP-enriched DNA over reference (non-enriched) DNA. This step is essential to reduce the intrinsic noise resulting from isolation of repetitive sequences during ChIP in mammalian cells. Third, the quality of antibody used for immunoprecipitation is very important. Transcription factors generally express at low level in living cells and have a weaker affinity for DNA than histone proteins; therefore, ChIP application of transcription factors is particularly demanding because the antibody must be capable of recognizing the native protein as part of a cross-linked protein–DNA complex. Many antibodies, even those that work well for Western blots, fail this more rigorous test. Different antibodies may also produce significantly different data sets (11). Triple FLAG epitope and the corresponding anti-FLAG M2 antibody provide the most sensitive antigen–antibody detection system to date. Detection of fusion proteins containing 3xFLAG is 20–200 times more sensitive than other tags such as c-myc, 6xHis, GST, or HA and is ideal for ChIP assays of low-level expression transcription factors in mammalian cells (http://www.sigmaaldrich.com/). There are several advantages in using a universal antibody–IP system with transcription factor of interest tagged with 3xFLAG. First is that many transcription factors show poor antigenicity and do not have good antibodies for efficient IP. Second is that some target genes show much less binding capacity than the others to the same transcription factor (26). To get enrichment of these weaker binding sites by ChIP, the transcription factor of interest has to be over-expressed to enhance the binding to these sites. Although over-expression of an epitope-tagged protein may cause artifactual interactions, this concern can be addressed by a subsequent verification step. Third is that using a universal antibody–IP system will produce a unique background related to IP process, which can be easily distinguished from bona fide IP products when applying to different transcription factors.

SABE method is shown in **Fig. 8.1**. An inducible human cell line expressing a transcription factor of interest tagged with 3xFLAG epitope is established. Cells are cross-linked in vivo using formaldehyde and lysed; DNA is sheared by sonication to produce fragments of 200–1,000 bp. Protein–DNA complexes are

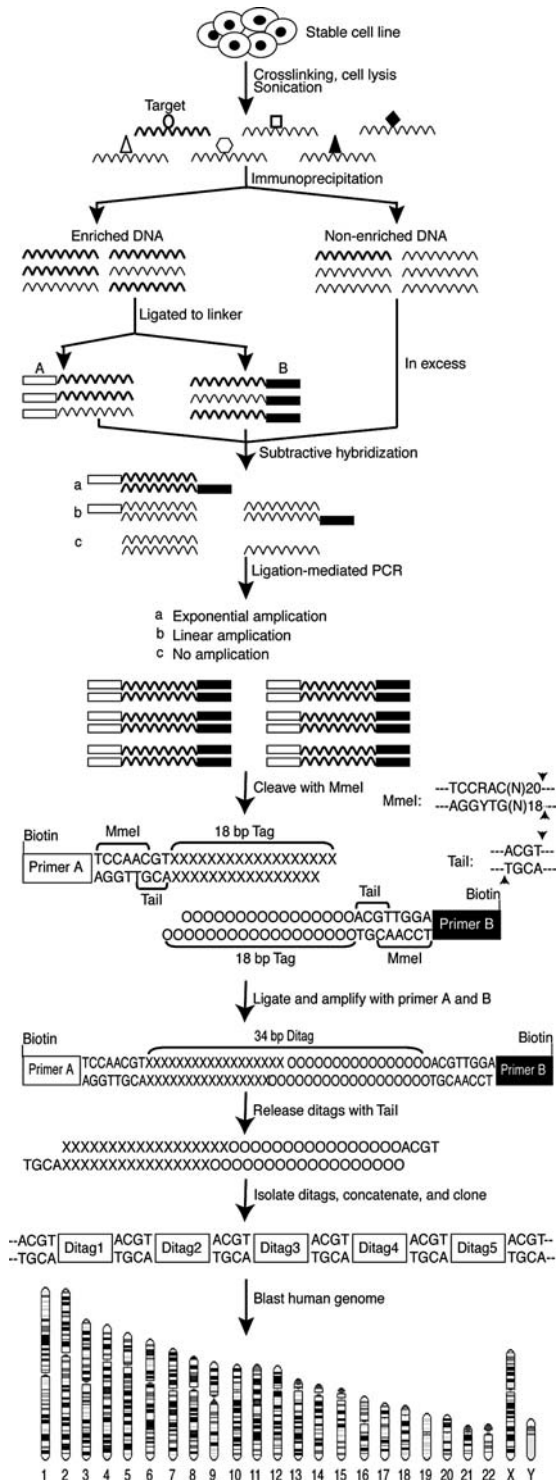

Fig. 8.1. Schematic representation of SABE method (12, 14). Inducible human cell line expressing transcription factor (TF) of interest are cross-linked, lysed, and sonicated. Protein–DNA complexes are immunoprecipitated using specific antibody. ChIP-enriched DNA is ligated to linkers and specific DNA selectively amplified by subtractive hybridization

then immunoprecipitated by using anti-FLAG M2 affinity beads. ChIP-enriched DNA is divided into two and ligated to either linker A or B, then hybridized to excessive amount of non-enriched DNA control (subtractive hybridization), followed by ligation-mediated PCR. After amplification, non-specific DNA sequences will be under-represented in the product mixture relative to specific DNA fragments. To analyze enriched DNA fragments, a strategy modified from SAGE is performed (17). The linkers A and B are designed with overlapping recognition sites for the type III endonuclease, *Mme*I and a 4 bp cutter, *Tai*I (**Fig. 8.1**). Additionally, to facilitate separation of the linkers from the final tag DNAs, the primers contain a 5′ biotin moiety. DNA fragments from subtractive hybridization and PCR amplification are digested with *Mme*I, and the 46 bp fragments, including 28 bp of the linker plus 18 bp of flanking tag sequence, are purified on 8% acrylamide gels. Because *Mme*I leaves a 2 bp 3′ overhang, to maximize information content of the tags, the digested fragments are ligated directly to form ditags, rather than trimming to create blunt ends (**Fig. 8.1**). The ligated ditags are amplified with primers A and B and then released by digestion with *Tai*I. *Tai*I was selected because it maximally overlaps with the *Mme*I site and is more efficient than *Nla*III, the anchoring enzyme used in SAGE (27). After digestion, the ditags can be separated from the biotin-tagged primer fragments by using streptavidin Dynabeads, further purified by electrophoresis, ligated to form concatemers, and directly cloned into pZERO-2a vector containing an *Aat*II site (GACGTC). Clones containing concatemers of 200–2,000 bp are analyzed by sequencing. Ditags can be identified in the sequencing data because each is 34 bp long separated by a *Tai*I sequence (ACGT). The final tag generated by SABE method is 18 bp long, including a 2 bp overlap generated by the *Mme*I digestion (**Fig. 8.1**). Tag sequences are used to blast the human genome database to identify its genomic location. Putative binding sites for the factor of interest can then be identified by analyzing flanking DNA on genes of particular interest for consensus sequences, with the rationale that the SABE tag must reside within a segment no greater than the length of the original sheared immunoprecipitated DNA fragments.

Fig. 8.1. (continued) and ligation-mediated PCR. Sequence tags are released by digestion with *Mme*I and ditags are produced by ligation, which are released by digestion with *Tai*I and separated from biotinylated linkers by using streptavidin magnetic beads. Ditags are concatemerized, cloned, and sequenced. Ditag sequences are 34 bp long and are separated by the *Tai*I recognition sequence (ACGT). Each tag sequence is 18 bp long and can be used to blast human genome database to decide its unique location.

3.1. Cell Culture, Cross-Linking, and Sonication

1. Grow 5×10^8 cells expressing the transcription factor of interest tagged with 3xFLAG epitope. The cells should be 80–90% confluent.
2. Collect cells. For adherent cells, aspirate the growth medium from the cells, scrape the cells into 50 mL conical centrifuge tubes using fresh medium, centrifuge for 5 min at 450*g* at room temperature (25°C), and then discard the supernatant. For suspension cells, collect the cells in 50 mL conical centrifuge tubes, centrifuge for 5 min at 450*g* at room temperature, and then discard the supernatant.
3. Re-suspend the cells in 45 mL pre-warmed culture media and collect all cells into one 50 mL conical centrifuge tube. Add 1.25 mL 37% formaldehyde solution to the cell suspension (final concentration: 1% formaldehyde). Incubate at room temperature for 10 min, with occasional inversion, to cross-link the protein of interest with DNA (*see* **Note 3**).
4. Add 5 mL 1.25 *M* glycine to the fixed culture and incubate at room temperature for 5 min, with occasional inversion.
5. Centrifuge cells for 5 min at 420*g* at 4°C and discard supernatant. Wash cells twice with 40 mL ice-cold 1X PBS, spin down cells for 5 min at 420*g* at 4°C after each wash, and discard supernatant. Place cell pellet on ice.
6. Re-suspend cell pellet in 5 mL ice-cold hypotonic buffer. Pass the cells through 27 1/2 gauge needle 10 times on ice to extract the nuclei. Collect the nuclei by centrifuging for 10 min at 10,000*g* at 4°C (*see* **Note 4**).
7. Discard the supernatant and re-suspend the nuclei in 6 mL lysis buffer. Incubate on ice for 30 min (*see* **Note 5**).
8. Shear chromatin by sonicating cell lysate for 10 min with cycles of 10 s of sonication followed by 50 s of pause with a sonicator. Keep cell lysate on ice during sonication. The final size of sheared DNA should be around 200–1,000 bp with average ~500 bp (*see* **Note 6**).
9. Centrifuge the suspension at 12,000*g* for 10 min at 4°C. Transfer supernatant (soluble cell lysate) into a new 15 mL tube. Place the tube on ice.

3.2. Pre-cleaning and Immunoprecipitation

1. Thoroughly suspend the ANTI-FLAG M2 affinity agarose gel and normal control mouse IgG-agarose gel in the vial, in order to make a uniform suspension of the resin. Immediately transfer 400 μL (for 6 mL of cell lysate) of the resin from each agarose gel in its suspension buffer to a separate new 1.5 mL tube to allow a homogenous dispersion of the resin. For resin transfer, use a clean, plastic pipette tip with the end enlarged to allow the resin to be transferred.

2. Centrifuge the resins for 30 s at 6,000*g* using a fixed angle rotor. In order to let the resin settle in the tube flatly, wait for 1–2 min before handling the samples. Aspirate the supernatant with a 27G1/2 needle.
3. Wash the packed gel with 1 mL lysis buffer. Repeat the wash twice. Be sure that the wash buffer is removed and no resin is discarded.
4. Add the normal control mouse IgG-agarose gel to the 6 mL soluble cell lysate to pre-clean the cell lysate. Add ANTI-FLAG M2 affinity agarose gel to 1 mL lysis buffer with 0.1% BSA and 1 μg/μL yeast tRNA to block the gel. Incubate both tubes on the rotating platform at 4°C for at least 1 h.
5. Collect pre-cleaned cell lysate and ANTI-FLAG M2 affinity agarose gel separately by centrifugation for 30 s at 6,000*g* at 4°C. Note pre-cleaned cell lysate is the supernatant in one tube and ANTI-FLAG M2 affinity agarose gel is the pellet in another tube.
6. Transfer ANTI-FLAG M2 affinity agarose gel to pre-cleaned cell lysate. Dilute the cell lysate with 1 volume (6 mL) of lysis buffer. Incubate at 4°C on a rotating platform overnight. Immunoprecipitation may be carried out for a longer time for convenience.

3.3. Washing, Elution, and Reversal of Cross-Link

1. Centrifuge the cell lysate with resin for 5 min at 3,000*g* at 4°C. Transfer the supernatants to a new 15 mL tube and keep as the non-enriched control.
2. Transfer the resin to a new 1.5 mL tube with fresh ChIP lysis buffer. Wash the resin three times sequentially with 1 mL each of the following pre-cooled buffers, all containing 1X PIC:ChIP lysis buffer; ChIP high salt buffer; ChIP wash buffer; and TE buffer. Pellet the resin during each wash by centrifugation for 30 s at 6,000*g* at 4°C and carefully aspirate the supernatant with a 27G1/2 needle.
3. Add 400 μL of elution buffer to the washed resin. As a control, transfer 360 μL of non-enriched control into a 1.5 mL tube and add 40 μL of 10% SDS. Incubate overnight at 65°C in a hybridization oven with rotation to revert the cross-link. This step may be carried out for a longer time for convenience.

3.4. Purification of ChIP DNA

1. Pellet the resin by centrifugation for 30 s at 6,000*g*. Transfer the supernatant to a new tube.
2. Add 3 μL of RNase A (20 μg/μL) to each tube. Incubate samples for 1 h at 37°C. Add 20 μL of proteinase K (20 μg/μL) to each tube. Incubate at 50°C for another hour.

3. Extract twice with 1 volume of phenol:chloroform:isoamyl alcohol mixture (25:24:1). Centrifuge for 3 min at 16,000*g* (13,000 rpm in an Eppendorf centrifuge with a 24-place fixed angle rotor) at 4°C. Transfer the DNA solution (upper aqueous phase) into a new tube after each extraction. Extract once with 1 volume of chloroform:isoamyl alcohol mixture. Centrifuge again for 3 min at 16,000 *g* at 4°C. Transfer the DNA solution into a new tube.
4. Add 1/10 volume of 3 *M* NaAc (pH 5.3). Add 3 volumes of cold 95–100% ethanol and mix briefly. Incubate at –20°C for at least 2 h.
5. Centrifuge at 16,000*g* for 20 min at 4°C. Pour off the supernatant, add 1 mL cold 70% ethanol, vortex briefly, and centrifuge again at the same speed for 3 min at 4°C. Carefully remove the supernatant with a pipette.
6. Let the pellet dry for a couple of minutes and re-suspend the pellet in 50 μL of TE; incubate at 65°C for 10 min.
7. Measure the DNA yield and purity using a spectrophotometer. The yield using anti-FLAG M2 affinity gel generally is 50–100 μg. Adjust both ChIP-enriched DNA and non-enriched DNA concentration to 1 μg/μL. The DNA can be stored for several months at –20°C.
8. Test specific enrichment of ChIP DNA over non-enriched DNA using known target and binding sites information for the transcription factor of interest. This information will be used in the subtractive hybridization step (*see* **Note 7**).

3.5. Blunting of ChIP-Enriched DNA

1. To blunt ChIP-enriched DNA, set up the following reaction mix: 50 μL of ChIP DNA, 1 μg/μl, 30 μL of 10X EcoPol buffer, 1 μL dNTP mix (10 m*M* each), 10 μL Klenow fragment (5 U/μL), and 209 μL water.
2. Mix by pipetting and incubate at RT for 15 min. Stop the reaction by adding 6 μL of 0.5 *M* EDTA and heating at 75°C for 20 min.
3. Extract once with phenol:chloroform:isoamyl alcohol mixture. Centrifuge for 3 min at 16,000*g* at 4°C. Transfer the DNA solution to a new tube. Extract once again with chloroform:isoamyl alcohol mixture and centrifuge for 3 min at 16,000*g* at 4°C. Transfer the DNA solution to a new tube.
4. Add 1/10 volume of 3 *M* NaAc. Add 3 volumes of cold 95–100% ethanol and centrifuge at 16,000*g* for 20 min at 4°C. Wash with 1 mL cold 70% ethanol. Dry the pellet and re-suspend the DNA pellet in 30 μL TE. Incubate at 65°C for 10 min.
5. Measure the DNA yield and adjust the DNA concentration to 1 μg/μL. The DNA can be stored for several months at –20°C.

3.6. Ligation of ChIP-Enriched DNA and Subtractive Hybridization

1. Separate ChIP-enriched DNA into two equal amounts and set up two ligation reactions, one with linker A and another with linker B. Set up the ligation mix as follows, also include mock ligation (without linker) as negative control: 6 μL of Blunt ChIP DNA (1 μg/μL), 6 μL of 10X DNA ligase buffer, 1 μL of LK-A (45 μ*M*) or LK-B, 3 μL of T4 DNA ligase, and 44 μL of water. Mix by pipetting and incubate for at least 2 h at 16°C. Longer ligation may be optimal. The ligation reaction can be left overnight at 16°C.
2. Recover the DNA using a QIAquick PCR purification kit according to the manufacturer's direction. Briefly, add 5 volumes (300 μL) of buffer PB1 to each of the ligation reaction (60 μL) and mix. To bind DNA, apply the samples to the QIAquick columns and centrifuge for 60 s at 10,000*g*. Wash the columns with 0.75 mL buffer PE. Place each column in a clean 1.5 mL tube and elute DNA in 30 μL buffer EB. Measure DNA concentration and purity. Adjust DNA concentration to 0.1 μg/μL. Test ligation efficiency by PCR using primer A or B. Efficient ligation with linkers will produce a significant amount of PCR products compared with control ligation without linker. The DNA can be stored for several months at –20°C.
3. Set up two hybridization solutions with either LK-A or LK-B ligated DNA as follows: 4 μL of 5X hybridization buffer, 12 μL of LK-A or LK-B DNA (0.1 μg/μL), and 4 μL of non-enriched DNA (1 μg/μL). Overlay with mineral oil, denature at 98°C for 1.5 min, and then hybridize at 65°C for 1.5 h.
4. Mix the two hybridization solutions (LK-A DNA and LK-B DNA), add 8 μL more heat-denatured non-enriched DNA and 2 μL of 5X hybridization buffer. Hybridize overnight at 65°C (*see* **Note 8**).
5. In the final 30 μL hybridization reaction, add the following: 20 μL of 10X PCR reaction buffer (NEB), 6 μL dNTP (10 m*M*), and 142 μL water. Incubate at 85°C for 3 min, and then bring down to 72°C before adding 2 μL of Taq DNA polymerase. Incubate at 72°C for another 10 min.
6. Purify the DNA using a QIAquick PCR purification kit. Briefly, add 5 volumes (1000 μL) of buffer PB1 to the DNA solution (200 μL) and mix. To bind DNA, apply the mixed solution to two QIAquick columns, each with 600 μL and centrifuge for 60 s at 10,000 *g*. Wash the columns with 0.75 mL buffer PE. Place each column in a clean 1.5 mL tube. Elute DNA in 50 μL elution buffer. These are linker-ligated DNAs (LK-DNAs). The final eluted DNA can be stored for several months at –20°C.

3.7. Optimizing PCR Condition and Linker-Mediated PCR

1. Make a serial two-fold dilution of LK-DNA template for a total of 20 dilutions. Set up the PCR reactions as follows: 4 μL (with various concentration) of LK-DNA template, 10 μL of 10X PCR buffer, 3 μL dNTP (10 m*M*), 2 μL primer A, 2 μL primer B, 1 μL Taq DNA polymerase, 78 μL water. Run the PCR as follows: 95°C 3 min; then 95°C 1 min, 58°C 1 min, and 72°C 2 min for 30 cycles; 72°C 10 min, and hold at 4°C.
2. Run 15 μL of each PCR product on a 2% agarose gel. The PCR product should be a smear ranging from 100 to 2,000 bp with an average size of 500 bp. Determine the minimal amount of template DNA required to yield maximum amount of PCR products. Set this amount of template DNA as optimal concentration for the following PCR reactions. Generally the optimal amount of template is 0.1–1 μL.
3. Set up large-scale PCR reactions using optimal template concentration determined at the last step: total 20 PCR reactions are needed for this step; each PCR reaction contains: 4 μL of ChIP-enriched DNA template (optimal concentration), 10 μL of 10X PCR buffer, 3 μL dNTP (10 m*M*), 2 μL primer A, 2 μL primer B, 1 μL Taq DNA polymerase, and 78 μL water.
4. Run PCR as follows: 95°C 3 min; then 95°C 1 min, 58°C 1 min, and 72°C 2 min for 30 cycles; 72°C 10 min; and then hold at 4°C.
5. Collect the PCR products (total 2,000 μL) in a 15 mL tube, and purify using a QIAquick PCR purification kit according to the handbook. Briefly, add 5 volumes (10 mL) of buffer PB1 to the PCR solution (2,000 μL) and mix. To bind DNA, add the mixed solution to six QIAquick columns, each with 600 μL and centrifuge for 60 s at 10,000*g*. Add the remaining solutions to the columns until all solutions have been added to the columns. Repeat the centrifuge step after each loading. Wash the columns with 0.75 mL buffer PE. Place each column in a clean 1.5 mL tube. Elute DNA with 50 μL of buffer EB. Collect the elution from all columns. Measure PCR yield and purity with a spectrophotometer. Adjust DNA concentration to 0.1 μg/μL. Generally the DNA yield will be 20–30 μg. The DNA can be stored for several months at –20°C.

3.8. MmeI Digestion, Isolation of Sequence Tag, and Ditag Formation

1. Set up a MmeI digest reaction as follows: 200 μL of PCR product (at 0.1 μg/μL), 40 μl of NEB buffer, 440 μL of 10X SAM, and 110 μL water.
2. Mix the reaction before adding MmeI enzyme. Then add 10 μL MmeI (2 U/μL) and mix very gently by pipetting six to eight times. Incubate at 37°C for 2 h. The reaction can be left overnight at 37°C (*see* **Note 9**).

3. Set up an 8% PAGE gel (16 × 20 cm) with a 15-well comb (well width: 6.5 mm, thickness: 1.0 mm) in a Bio-Rad PROTEAN II xi cell as follows: 10.7 mL of 30% acrylamide (29:1), 2 mL of 10X TBE, 200 μL APS (10%), 40 μL TEMED, and 27.06 mL water.
4. Add 50 μL of 50% glycerol (do not use loading dye) to the MmeI digestion (400 μL) and mix. Load the reaction directly to the gel, each lane with 60 μL. Total eight lanes are needed. Also include one lane with un-cut control, one with loading dye (bromophenol blue) only and one with 1 μg of 10 bp DNA ladder.
5. Run the gel for 2–4 h at 200 V with water-cooling until bromophenol blue is three-fourths down the gel.
6. Stain the gel with SYBR Green I at a dilution of 1:10,000 in 1X TBE buffer for 30 min with gentle agitation. Visualize the bands under a standard UV trans-illuminator and take a photo as a record. A strong 46 bp band should be seen.
7. Make a hole through the bottom of a 0.5 mL Eppendorf tube using an 18-gauge needle.
8. Using a new razor blade, excise the 46 bp band from the gel. Collect the gel slices from two lanes into one 0.5 mL tube with a hole and place the tube on a 2 mL screwed tube. Total four tubes are needed. Centrifuge for 1 min at 16,000*g*. The excised bands will be broken into small pieces and collected in the 2 mL tube.
9. Add 1 mL of gel diffusion buffer to the 2 mL tube containing the gel pieces. Incubate at 65°C for 2 h to elute the DNA from the gel with agitation.
10. Pass the gel solution through a Micro Bio-Spin Chromatography Column by centrifuging at 3 min at 16,000*g* to remove any residual polyacrylamide. Collect the DNA solution in 1.5 mL tubes.
11. Fill up the tubes with 1-butanol and mix. Centrifuge 1 min at 16,000*g*. Discard the upper phase containing 1-Butanol. Repeat this step until the volume in each tube is reduced to 200 μL. Transfer all the DNA solutions from four tubes to a new 1.5 mL tube and reduce the volume the DNA solution to 400 μL with 1-Butanol.
12. Extract the DNA solution twice with 1 volume of phenol:-chloroform:isoamyl alcohol mixture. Centrifuge for 3 min at 16,000 *g* at 4°C. Transfer the DNA solution to a new tube. Extract once again with 1 volume of chloroform:isoamyl alcohol mixture and centrifuge for 3 min at 16,000 *g* at 4°C. Transfer the DNA solution to a new tube.
13. Add 1/10 volume of 3 *M* NaAc, add 3 volumes of cold 95–100% ethanol, and vortex. Incubate at –20°C for 2 h.

14. Centrifuge at 16,000*g* at 4°C for 30 min. Carefully remove and discard the supernatant. Wash the DNA pellet twice, each with 800 μL of cold 70% ethanol. Air-dry and re-suspend the DNA pellet in 20 μL of water. These are the 46 bp long MmeI sequence tags. The DNA can be stored for several months at –20°C.
15. Set up a ligation reaction as follows: 17 μL of purified tags, 2 μL of 10X ligation buffer, and 1 μL T4 DNA ligase. Mix gently and incubate overnight at 16°C.
16. Add 180 μL of TE to the ligation reaction. Heat at 65°C for 10 min to inactivate the DNA ligase.

3.9. Optimizing PCR Condition and PCR Amplification of Ditags

1. Make a serial two-fold dilution of ligated ditags for a total of 20 dilutions. Set up the PCR reactions as follows: 10 μL (various concentration) of ligated ditag template, 10 μL of 10X PCR buffer, 3 μL dNTP (10 m*M*), 2 μL primer A, 2 μL primer B, 1 μL Taq DNA polymerase, and 72 μL water.
2. Run the PCR as follows: 95°C 3 min; then 95°C 30 s, 58°C 30 s and 72°C 10 s for 30 cycles; 72°C 10 min; and finally hold at 4°C.
3. After PCR, set up an 8% PAGE gel as indicated before and analyze the PCR products. A clear 90 bp ditag band should be seen. Determine the minimal amount of template DNA (ditags) required to yield a significant 90 bp band. Set this amount of template DNA as optimal concentration and proceed to scale-up PCR. Generally the optimal template amount is 1 μL.
4. Set up the PCR reaction as follows using optimal ditag template concentration determined in last step. Total 20 reactions are needed. One reaction contains: 1 μL of ligated ditag template at optimal concentration, 10 μL of 10X PCR buffer, 3 μL dNTP (10 m*M*), 2 μL primer A, 2 μL primer B, 1 μL Taq DNA polymerase, and 81 μL water.
5. Run the PCR as follows: 95°C 3 min; then 95°C 30 s, 58°C 30 s and 72°C 10 s for 30 cycles; 72°C 10 min, and hold at 4°C.
6. Collect the PCR products into five 1.5 mL tubes, each containing 400 μl. Extract once with 1 volume of phenol:chloroform:isoamyl alcohol mixture and once with chloroform:isoamyl alcohol mixture. Spin for 3 min at 4°C at 16,000*g* after each extraction. Transfer the supernatant to new tubes.
7. Add 1/10 volume of 3 *M* NaAc, add 3 volumes of cold 95–100% ethanol, and vortex. Incubate at –20°C for 2 h.
8. Centrifuge at 16,000 *g* at 4°C for 30 min. Carefully remove and discard the supernatant. Wash the DNA pellet with 70% ethanol. Air-dry and re-suspend the pellet in 20 μL of water. Collect all DNA solutions into one tube. Measure the yield and purity. Adjust DNA concentration to 0.1 μg/μL. The DNA can be stored at –20°C for several months.

3.10. Tail Digestion and Purification of Ditags

1. Set up TaiI digest reaction as follows: 100 μL PCR product (0.1 μg/μL), 20 μl of buffer R^+, 10 μL TaiI (10 units/μL), and 70 μL water.
2. Mix gently. Incubate at 65°C for 2 h or overnight at 65°C.
3. Aliquot 100 μL (10 μg/μL) of Dynabeads M-280 streptavidin into a clean 1.5 mL tube. Add 200 μL of 1X wash/binding buffer and vortex to suspend beads. Apply a magnet field to the side of the tube for 1–2 min. Remove and discard the supernatant. Repeat wash once.
4. To 200 μL of TaiI-digested ditags, add an equal volume of 2X wash/binding buffer and mix. Then transfer the solution to the tube containing magnetic beads. Vortex to suspend the particles and incubate at room temperature for 10 min with agitation.
5. Apply a magnet field. Transfer the supernatant into a new tube.
6. Set up a 12% PAGE gel (16 × 20 cm) using a Bio-Rad PROTEAN II xi cell system as follows: 16 mL of 30% acrylamide (29:1), 2 mL of 10X TBE, 200 μL APS (10%), 40 μL TEMED, and 21.76 mL water.
7. Add 50 μL of 50% glycerol (do not use loading dye) to the ditag solution (400 μL) and mix. Load the solution directly to the gel, each lane with 60 μL. Total of eight lanes are needed. Also include one lane with un-cut control, one with loading dye (bromophenol blue) only and one with 1 μg of 10 bp DNA ladder.
8. Run the gel at 200 V for 2 h. Purify the 34 bp ditag band as described in **Section 3.8.** Dissolve the final ditag in 20 μL of water. The precipitated DNA can be stored at –20°C for several months.

3.11. Ligation of Ditags to Form Concatemers and Isolation

1. Set up a ligation reaction as follows: 17 μL of purified ditags, 2 μL of 10X ligation buffer, and 1 μL T4 DNA ligase. Mix gently and incubate overnight at 16°C.
2. Load the ligation solution onto a 1% agarose gel and run the gel.
3. Excise the concatemers of 200–2,000 bp from the gel. Collect the gel slices into 1.5 mL tube.
4. Purify the concatemers using a QIAquick gel purification kit according to the manufacturer's directions. Briefly, weigh the gel slices and add 3 volumes of buffer QG to 1 volume of gel (100 mg ~100 μL). Incubate at 50°C for 10 min. Add 1 gel volume of isopropanol to the sample and mix. To bind DNA, add the samples to QIAquick columns and centrifuge for 60 s at 10,000*g*. Add 0.5 mL of buffer QG to the column and centrifuge again for 60 s. Wash the columns with 0.75 mL buffer PE. Place each column in a clean 1.5 mL tube. Elute DNA with 50 μL of buffer EB. Measure the DNA concentration and adjust to 10 ng/μL. The DNA can be stored at –20°C for several months.

3.12. Cloning Concatemers and Colony PCR Analysis

1. Transform plasmid pZERO-2a into an F′ *E. coli*strain (e.g., JM109) and spread to LB-Kanamycin plate. Select one single colony and grow in 500 mL of LB medium containing 50 μg/mL Kanamycin and purify plasmid DNA using CsCl gradient ultracentrifugation (*see* **Note 10**).
2. Digest 1 μg of CsCl-purified pZERO-2a plasmid with AatII. Extract the DNA with phenol:chloroform:isoamyl alcohol mixture and chloroform:isoamyl alcohol mixture. Precipitate the DNA with ethanol and dissolve it in 100 μL of TE (10 ng/μL).
3. Set up the ligation reaction as follows: 5 μL of digested vector (10 ng/μL), 5 μL of purified concatemers (10 ng/μL), 1 μL of 10X ligation buffer, 8 μL water, and 1 μL T4 DNA ligase.
4. Incubate at 16°C for 30 min. Longer ligation may be optimal. Transform 10 μL of ligation solution into 100 μL of DH5α competent cells. Plate all transformation mix on LB-Kanamycin plates.
5. Analyze Kanamycin-resistant colonies by colony PCR using M13 forward and reverse primers. Pick up clones bearing an insert between 200 and 2,000 bp.

3.13. Sequencing and Sequence Analysis

1. Grow selected clones and sequence these clones using T7 primer.
2. Analyze the sequencing data. Typically, each clone contains 10–30 ditags. Ditags are 34 bp long and separated by the TaiI recognition site, ACGT. The final tag generated by the SABE method is 18 bp long, including a 2 bp overlap generated by MmeI digestion. Tag sequences are used to blast the human genome database to identify its genomic location. Putative binding sites for the transcription factor of interest can be identified by analyzing flanking sequences for consensus sequences (*see* **Note 11**).

4. Notes

1. Two linkers are used to prevent the formation of pan-like structure during subtractive hybridization and LM-PCR. Both linkers are modified with an amino group at the 3′ end to prevent self-ligation. Linkers should be obtained PAGE-purified after synthesis from oligo company (i.e., Integrated DNA Technologies for linker syntheses).
2. Both primers are biotinylated at the 5′ end to facilitate isolation of ditags. Primers should be obtained PAGE purified.

3. The aim of cross-linking is to fix the transcription factor of interest to its chromatin binding sites. Cross-linking is a time-critical procedure and the optimal length of cross-linking depends on the cell type and transcription factor of choice. Too much cross-linking may mask epitopes for efficient immunoprecipitation and too little cross-linking may lead to incomplete fixation. If uncertain, perform a time-course experiment and run a conventional ChIP assay to optimize cross-linking conditions.
4. Cell lysis can be observed by the addition of the Trypan blue solution to an aliquot of cells. The dye is excluded from the intact cells, but stains the nuclei of lysed cells. Lysis should be 80–90%. If the lysis is not sufficient, perform several more strokes until lysis is complete. If nuclear lysis or clumps of nuclei are visualized, the cell disruption was too vigorous or too many strokes were performed.
5. Foaming during the sonication step can result in insufficient shearing of chromatin DNA. To avoid this, use 6 mL total volume in a 15 mL conical tube and keep sonicator tip 0.5–1 cm deep in cell lysate sample during sonication.
6. Sonication efficiency will vary depending on sonicator, cell type, and extent of cross-linking and will have to be optimized to yield the desired final average length of DNA for each specific cell type. Ideally, the average DNA size of sheared sample should be confirmed by 2% agarose gel electrophoresis stained with ethidium bromide.
7. For all DNA enriched by ChIP experiments, the efficiency of immunoprecipitation must be determined by quantitative real-time PCR analyses (i.e., ratio of the amount of enriched (immunoprecipitated) DNA over that of the non-enriched (left-over) DNA). For this purpose, the knowledge of at least one well-defined binding site for the transcription factor of interest is needed. This knowledge of a known target gene is used to design the primers for real-time PCR and optimize immunoprecipitation conditions. The ratio of enrichment should also be normalized to the level observed at a control region, which is defined as 1.0. In general, if more than 10 fold of enrichment can be achieved by immunoprecipitation step, the following subtractive hybridization step can further increase the signal-to-noise ratio. However, please note that the transcription factor of interest may have different affinity to its individual binding sites. For example, ChIP recovers several 100-fold more p21 and MDM2 promoter DNA while recovers substantially weaker or background p53 binding elements for Bax, AIP1, and PIG3 (26). For this reason, if there are more than one known target genes for the

transcription factor of interest available, choose the one that can achieve higher signal-to-noise ratio. Always test the quality of antibody and optimize ChIP conditions (cross-linking, sonication, immunoprecipitation, etc.) using this knowledge of known target gene(s) for the specific transcription factor and cell type of your choice. If a satisfactory ratio of ChIP enrichment cannot be achieved using its own antibody, consider making a construct of transcription factor tagged with 3xFLAG epitope and using anti-FLAG M2 affinity agarose beads for ChIP, which has been proven to have the highest affinity compared with other epitope tags.

8. The ratio of non-enriched over ChIP-enriched DNA in subtractive hybridization solution is dependent on the ratio of enrichment obtained from the immunoprecipitation step. Typically, 10-fold of ChIP enrichment will be needed to begin the subtractive hybridization step.
9. (a) Make 10 times *S*-adenosylmethionine (10X SAM) solution (500 μ*M*) from its original concentration (32 m*M*) freshly before use. (b) Reaction using MmeI should be done at or near stoichiometric concentration as indicated (1 μg DNA/1 μL MmeI). Excessive amounts of MmeI block cleavage.
10. (a) Plasmid pZERO-2a cannot grow in *E. coli*strains without a *lac*I^q gene (e.g., DH5α); (b) plasmid DNA purified by other methods (i.e., Qiagen plasmid purification kit) contains small amount of *E. coli* genomic DNA, which may be cloned and mistakenly selected for sequencing.
11. Due to the quality of performance for each SABE step, final clones may contain primer dimers and linker sequences. Final clones may also contain *E. coli* genomic sequences if using plasmid DNA purified by methods other than CsCl ultracentrifugation. It is worth noting that there are about 30–40% of the final SABE tags that cannot be assigned unique locations to the human genome due to multiple hits. This is probably because of the repetitive elements in the human genome, whose lengths range from several hundreds to several thousands of base pairs (24, 25).

References

1. Wyrick, J. J. and Young, R. A. (2002) Deciphering gene expression regulatory networks. *Curr. Opin. Genet. Dev.* **12**, 130–136.
2. Ptashne, M. and Gann, A. (1997) Transcriptional activation by recruitment. *Nature* **386**, 569–577.
3. Iyer, V. R., Horak, C. E., Scafe, C. S., Botstein, D., Snyder, M. and Brown, P. O. (2001) Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* **409**, 533–538.
4. Ren, B., Robert, F., Wyrick, J. J., et al. (2000) Genome-wide location and function

of DNA binding proteins. *Science* **290**, 2306–2309.

5. Lee, T. I., Rinaldi, N. J., Robert, F., et al. (2002) Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**, 799–804.
6. Boyd, K. E. and Farnham, P. J. (1997) Myc versus USF: discrimination at the cad gene is determined by core promoter elements *Mol. Cell. Biol.* **17**, 2529–2537.
7. Zeller, K. I., Haggerty, T. J., Barrett, J. F., Guo, Q., Wonsey, D. R., Dang, C. V. (2001) Characterization of nucleophosmin (B23) as a Myc target by scanning chromatin immunoprecipitation. *J. Biol. Chem.* **276**, 48285–48291.
8. Ren, B., Cam, H., Takahashi, Y., et al. (2002) E2F integrates cell cycle progression with DNA repair, replication, and G(2)/M checkpoints. *Genes Dev.* **16**, 245–256.
9. Odom, D. T., Zizlsperger, N., Gordon, D. B., et al. (2004) Control of pancreas and liver gene expression by HNF transcription factors. *Science* **303**, 1378–1381.
10. Weinmann, A. S., Yan, P. S., Oberley, M. J., Huang, T. H. and Farnham, P. J. (2002) Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev.* **16**, 235–244.
11. Cawley, S., Bekiranov, S., Ng, H. H., et al. (2004) Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell* **116**, 499–509.
12. Chen, J. and Sadowski, I. (2005) Identification of the mismatch repair genes PMS2 and MLH1 as p53 target genes by using serial analysis of binding elements. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 4813–4818.
13. Colvis, C. M., Pollock, J. D., Goodman, R. H., et al. (2005) Epigenetic mechanisms and gene networks in the nervous system. *J. Neurosci.* **25**, 10379–10389.
14. Chen, J. (2006) Serial analysis of binding elements for human transcription factors. *Nat. Protoc.* **1**, 1481–1493.
15. Orlando, V. (2000) Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation. *Trends Biochem. Sci.* **25**, 99–104.
16. Lisitsyn, N. and Wigler, M. (1993) Cloning the differences between two complex genomes. *Science* **259**, 946–951.
17. Velculescu, V. E., Zhang, L., Vogelstein, B. and Kinzler, K. W. (1995) Serial analysis of gene expression. *Science* **270**, 484–487.
18. Impey, S., McCorkle, S. R., Cha-Molstad, H., et al. (2004) Defining the CREB regulon: a genome-wide analysis of transcription factor regulatory regions. *Cell* **119**, 1041–1054.
19. Kim, J., Bhinge, A. A., Morgan, X. C. and Iyer, V. R. (2005) Mapping DNA–protein interactions in large genomes by sequence tag analysis of genomic enrichment. *Nat. Methods* **2**, 47–53.
20. Labhart, P., Karmakar, S., Salicru, E. M., et al. (2005) Identification of target genes in breast cancer cells directly regulated by the SRC-3/AIB1 coactivator. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 1339–1344.
21. Roh, T. Y., Ngau, W. C., Cui, K., Landsman, D. and Zhao, K. (2004) High-resolution genome-wide mapping of histone modifications. *Nat. Biotechnol.* **22**, 1013–1016.
22. Wei, C. L., Wu, Q., Vega, V. B., et al. (2006) A global map of p53 transcription-factor binding sites in the human genome. *Cell* **124**, 207–219.
23. Blais, A. and Dynlacht, B. D. (2005) Constructing transcriptional regulatory networks. *Genes Dev.* **19**, 1499–1511.
24. Lander, E. S., Linton, L. M., Birren, B., et al. (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921.
25. Venter, J. C., Adams, M. D., Myers, E. W., et al. (2001) The sequence of the human genome. *Science* **291**, 1304–1351.
26. Kaeser, M. D. and Iggo, R. D. (2002) Chromatin immunoprecipitation analysis fails to support the latency model for regulation of p53 DNA binding activity in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 95–100.
27. Saha, S., Sparks, A. B., Rago, C., et al. (2002) Using the transcriptome to annotate the genome. *Nat. Biotechnol.* **20**, 508–512.