PROSPECT
Journal of Cellular Biochemistry 107:6–10 (2009)

Journal of Cellular Biochemistry

From Sextant to GPS: Twenty-Five Years of Mapping the Genome With ChIP

David A. Wacker[1] and Tae Hoon Kim[1,2*]
[1] *School of Medicine, Yale University, New Haven, Connecticut 06510*
[2] *Department of Genetics, Yale University, New Haven, Connecticut 06510*

ABSTRACT
Since its inception, ChIP technology has evolved immensely. Technological advances have improved its specificity and sensitivity, its scale has expanded to a genome-wide level, and its relative ease of use has made it a virtually ubiquitous tool. This year marks the 25th anniversary of the development of ChIP. In honor of this milestone, we briefly revisit its history, offer a review of recent articles employing ChIP on a genome-wide scale, and lay out our views for the future of ChIP. J. Cell. Biochem. 107: 6–10, 2009. © 2009 Wiley-Liss, Inc.

KEY WORDS: ChIP-CHIP; ChIP-SEQ; ChIP-PET; ChIP-SAGE; GENOME-WIDE

Chromatin immunoprecipitation (ChIP) was first reported by David Gilmour and John Lis in 1984 using UV crosslinking to investigate the in vivo RNA polymerase occupancy at bacterial genes [Gilmour and Lis, 1984]. The following year, Solomon and Varshavsky [1985] described the use of formaldehyde as a crosslinker in probing chromatin structures in vivo. The main difference between the two initial reports was the choice of crosslinkers. Although UV is a more general crosslinking reagent than formaldehyde, UV-induced protein–DNA crosslinks are not reversible, unlike formaldehyde crosslinks which can be reversed by mild heating [Solomon and Varshavsky, 1985]. Thus, DNA from formaldehyde-crosslinked chromatin can be readily manipulated for molecular analysis, making formaldehyde the crosslinker of choice for ChIP. The following decades saw maturation of ChIP into one of the most critical methods in molecular biology and functional genomics.

The basic method of ChIP involves treating living cells with formaldehyde which fixes proteins to their DNA substrates inside the cells. The chromosomes are then extracted and fragmented by physical shearing or enzymatic digestion. The specific DNA sequences associated with a particular protein complex are isolated by immuno-affinity purification using a specific antibody against the protein (Fig. 1). The purified DNA fragments are then assayed by a variety of molecular techniques, such as Southern blot or polymerase chain reaction (PCR), to determine association of particular DNA sequences with the protein of interest [Orlando and Paro, 1993; Hecht et al., 1996].

In 2000, two groups introduced a method to detect protein–DNA interaction sites scattered throughout millions of base pairs of sequences using DNA microarrays (ChIP-chip) [Ren et al., 2000; Iyer et al., 2001]. The commercial availability of high-density oligonucleotide arrays representing the entire human genome have facilitated comprehensive mapping of protein-DNA interaction sites by ChIP-chip [Kim et al., 2005] (Fig. 1). This genome-wide approach to investigating protein–DNA interactions was extended by the adaptation of serial analysis of gene expression (SAGE) [Velculescu et al., 1995] technique to analysis of ChIP DNA [Roh et al., 2004; Wei et al., 2006]. Subsequently, technological advances in high throughput sequencing [Johnson et al., 2007] (ChIP-seq) have opened a new chapter in ChIP-based analysis of gene control and epigenomics. Direct sequencing of ChIP DNA has also made it possible to interrogate a significant fraction of repeat elements in the genome that was technically inaccessible using DNA microarrays.

ChIP-based genome-wide analysis has been applied to a plethora of transcription factors and other DNA-binding proteins. Of particular note are studies mapping the binding of transcription factor IID (TFIID) [Kim et al., 2005], estrogen receptor (ER) [Carroll et al., 2006], p53 [Wei et al., 2006], neuron-restrictive silencing factor (NRSF) [Johnson et al., 2007], and CCCTC-binding factor (CTCF) [Kim et al., 2007].

Grant sponsor: Rita Allen Foundation; Grant sponsor: Sidney Kimmel Foundation for Cancer Research; Grant sponsor: Yale Comprehensive Cancer Center; Grant sponsor: Yale University School of Medicine.
*Correspondence to: Prof. Tae Hoon Kim, Department of Genetics, Yale University School of Medicine, Cedar Street, SHM I-142B, P.O. Box 208005, New Haven, CT 06520-8805. E-mail: taehoon.kim@yale.edu
Received 15 December 2008; Accepted 17 December 2008 • DOI 10.1002/jcb.22060 • 2009 Wiley-Liss, Inc.
Published online 29 January 2009 in Wiley InterScience (www.interscience.wiley.com).

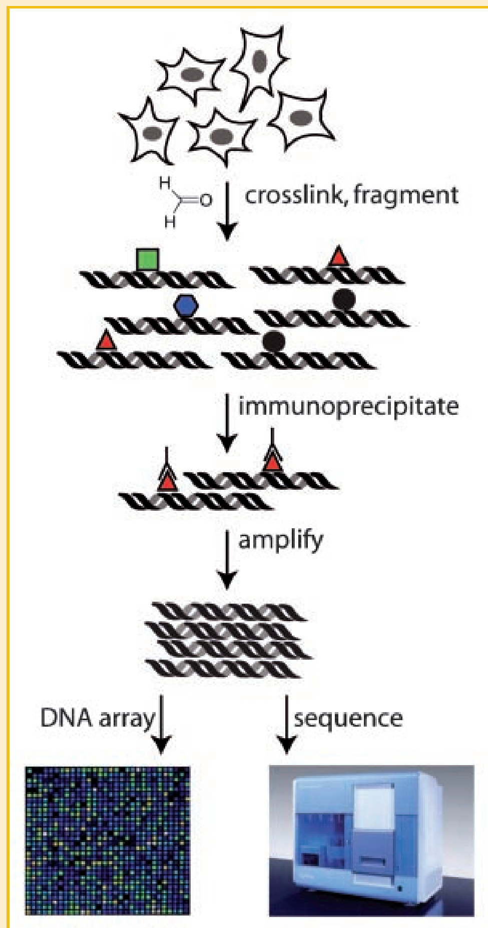

Fig. 1. Overview of genome-wide ChIP techniques. Cultured or primary cells are treated with formaldehyde to chemically crosslink DNA-binding proteins to their binding loci in situ. Nuclear DNA is then purified and fragmented to yield DNA segments with associated proteins (squares, triangles and circles, here) still intact. Proteins of interest (triangles, here) are selected for by immuno-affinity precipitation, and crosslinks are reversed. The resulting co-precipitated DNA is then analyzed by whole-genome microarray (ChIP-chip) or high-throughput sequencing (ChIP-seq).

TFIID has long been known as a general transcription factor critical to preinitiation complex (PIC) formation. In an effort to identify novel PIC sites, one recent study employed genome-wide ChIP-chip to interrogate human primary lung fibroblasts (IMR90 cells) for TFIID binding [Kim et al., 2005]. Over 12,000 sites were identified, of which 87% were within 2.5 kb of the 5′ end of a known mRNA. Interestingly, the correlation of TFIID binding sites with other genomic landmarks identified 368 novel gene candidates, and 1,239 putative new promoters, including 35 non-gene transcription units. Furthermore, an analysis of 8,960 TFIID binding sites at known genes was highly suggestive of significant alternative promoter usage throughout the genome, as over 4,000 of these TFIID sites map to genes with at least two sites present. The results of this study have also furthered our understanding of promoter makeup and function. Of over 9,000 TFIID-bound active promoters, less than 10% contain a TATA-box, an element previously thought to be widely utilized.

In another recent study, the human colon cancer cell line HCT116 was probed by ChIP coupled to paired end tag sequencing (ChIP-PET) for p53 binding sites [Wei et al., 2006]. From this, 542 high-probability p53 binding sites were identified, as were 122 p53-regulated genes, 98 of them not previously known to be associated with p53. To explore the clinical relevance of these 122 p53-regulated genes in controlling cancers, their expression levels were assayed in 251 primary breast cancer specimens, 58 of them with p53 mutations. The p53-mutant samples demonstrated weaker expression of p53 up-regulated genes, and stronger expression of p53 down-regulated genes than their p53-intact counterparts. Interestingly, the tumors showing deregulation of p53-controlled genes were of significantly higher grade and metastatic potential than normally regulated tumors.

Estrogen receptor and neuron-restrictive silencing factor both control transcription in a cell type-specific manner. NRSF silences neuron-specific genes in non-neuronal cells, while ER facilitates cellular responses to estrogens in numerous tissue types. Global binding maps for these two factors have recently been generated using ChIP-seq with Jurkat T cells (NRSF) [Johnson et al., 2007] and ChIP-chip with MCF-7 breast cancer cells (ER) [Carroll et al., 2006]. Over 1,900 NRSF binding sites were identified, including 22 regions lacking a canonical NRSF consensus motif. Analysis of these non-canonical regions revealed that the NRSF binding site actually consists of two non-palindromic half sites separated by a spacer, which canonically is 11 basepairs, but when extended results in the non-canonical sites [Johnson et al., 2007]. While NRSF tended to occur in or near genes, only 4% of 3,665 ER binding sites identified occurred within a 1 kb region surrounding transcription start sites. Furthermore, the number of ER binding sites on a given chromosome tended to correlate more with the length of the chromosome, than the number of genes present. These results confirm earlier smaller-scale experiments [Carroll et al., 2005], suggesting that ER binding sites tend not to be promoter proximal, but rather act as distal regulatory elements. Consistent with this, motif analysis followed by ChIP-qPCR confirmation has identified several factors which co-localize with ER, including Forkhead (FoxA1), AP-1, Oct, and C/EBP, and these motifs also do not enrich to promoter-proximal areas.

Finally, CTCF is a ubiquitously expressed nuclear protein which has been shown under various contexts to act as a transcriptional activator [Vostrov and Quitschke, 1997], repressor [Lobanenkov et al., 1990], and insulator [Bell et al., 1999]. A ChIP-chip survey of CTCF binding in primary lung fibroblasts (IMR90) revealed that of almost 14,000 binding sites found, half were located near or in genes, while the other half distributed to intergenic DNA. Consistent with CTCF's function as an insulator, binding sites were depleted in regions containing clusters of co-regulated genes, but enriched at genes with multiple alternative promoters, where activators for one promoter may need to be isolated from other potential targets. Furthermore, analysis of the layout of CTCF sites relative to nearby genes suggested a fundamental containment scheme, whereby one or more genes (average 2.5) were isolated from neighboring regions by a CTCF binding site on either side. In all, 74% of genes are located in such CTCF-pair-defined domains. In addition to these discoveries, motif alignment of the identified binding sites revealed a 20-bp

consensus sequence for CTCF binding, which was present in over 75% of the sites.

In addition to mapping DNA-binding proteins across the genome, ChIP-chip, ChIP-seq, and other genome-wide interrogation techniques have also been used to identify regions associated with particular histone modifications, such as methylation and acetylation. Two ChIP-seq studies provide an especially complete map of histone marks in human primary $CD4^{+}$ T-cells [Barski et al., 2007; Wang et al., 2008] (Fig. 2). Several of the findings of these studies confirm on a whole-genome scale what was already known or suspected, for example di- and trimethylation of H3K4 are often found within 1 kb of the transcription start site of active genes. However many novel findings also arose. While H3K27 and K9 di- and trimethylation were confirmed as repressive marks, monomethylation of these residues proved to be associated with active promoters. Additionally, H2BK5 monomethylation, a previously unexplored mark, was found to associate downstream of active transcription start sites. Exploration of CTCF binding sites and known active enhancers revealed enrichment for H3K4 methylation and H2A.Z at these elements [Barski et al., 2007].

Synthesizing the available information from these two studies, the authors identify 17 histone modifications that concurrently appear at 3,286 active promoters (Fig. 2) [Wang et al., 2008]. While the significance of this "backbone" of modifications is not fully clear, the elevated expression levels observed of genes with the backbone, and the relative absence of promoters with only 16 of the 17 modifications, suggests that these form a single functional unit which establishes a transcription-friendly environment. In addition, recent studies indicate monomethylation of H3K4 seems to be a general mark for enhancers in the human genome [Heintzman et al., 2007; Robertson et al., 2008]. Therefore, various epigenetic modifications cluster into relatively simple classes of functional elements in the genome.

Perhaps the greatest advances using ChIP-based and other genome-wide analyses have been made in the field of gene expression control in stem cells. A number of recent, genome-wide studies have analyzed the role of histone and DNA modifications in controlling transcription of the genes that promote or inhibit differentiation [Mikkelsen et al., 2007; Pan et al., 2007; Zhao et al., 2007; Fouse et al., 2008; Meissner et al., 2008; Orford et al., 2008]. These studies have unmasked global trends in the control of several classes of genes in embryonic stem (ES) cells and their differentiated progeny.

Studies on mouse ES cells using ChIP-seq [Mikkelsen et al., 2007], human hES3 cells using ChIP-PET [Zhao et al., 2007], and human H1 ES cells using ChIP-chip [Pan et al., 2007] have all recently confirmed the role of histone H3K4 and H3K27 trimethylation (H3K4me3 and H3K27me3) in regulating expression and the potential expression of genes in these lines. Promoters associated with only H3K4me3 demonstrated a high likelihood of expression of their products, indeed approximately 80% of genes in this group were actively expressed in hES3 cells [Zhao et al., 2007]. Conversely, promoters associated with H3K27me3, or with neither K4 nor K27 trimethylation ("neither") show very low expression, with less than 1% of "neither" genes expressing in hES3 cells [Zhao et al., 2007]. Interestingly, these studies demonstrated that the "bivalent" state, in which both H3K4 and H3K27 are trimethylated at the same promoter [Azuara et al., 2006; Bernstein et al., 2006], occurs throughout the stem cell genome, and is associated with a moderately expressed, poised state which can quickly resolve to frank activation (H3K4me3-only) or deeper, more permanent silencing (H3K27 me3-only or neither mark) [Mikkelsen et al., 2007; Pan et al., 2007; Zhao et al., 2007]. In a genome-wide survey of bivalently marked promoters in mES cells upon differentiation to neural progenitor cells, only 8% retained a bivalent mark, while 46% were promoted to H3K4me3 only, and 40% were left with either H3K27me3-only or neither mark. The fate of each individual bivalent gene was most related to its function: genes with functions specific to the new cell type were activated, while unrelated genes were suppressed [Mikkelsen et al., 2007; Pan et al., 2007].

Because of the global nature of these studies, a break down of each promoter class by ontological family was possible. Promoters

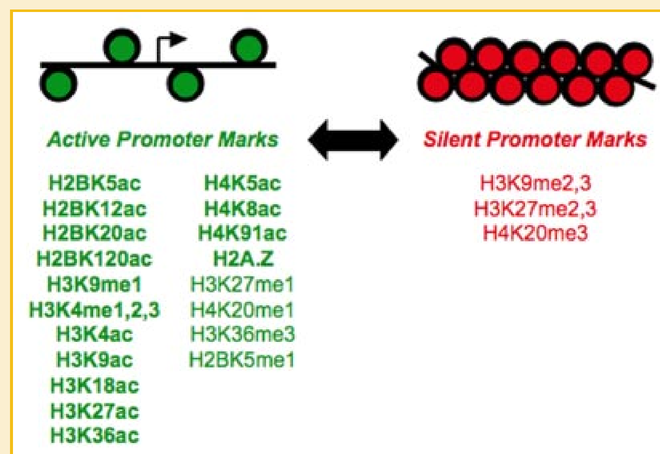

Fig. 2. Enrichment of active and silent promoters for specific histone marks. Outline of the distribution of histone marks to active (left) versus silent (right) promoters. The 17 "backbone" active chromatin marks are shown in bold. me1, me2, me3 = mono-, di-, tri-methyl, respectively. ac = acetyl. Based on data from Barski et al. [2007] and Wang et al. [2008].

marked by only H3K4me3 showed enrichment of housekeeping genes such as those associated with proliferation or metabolism [Mikkelsen et al., 2007; Pan et al., 2007; Zhao et al., 2007; Orford et al., 2008]. In contrast, promoters with neither methyl mark enriched for genes expressed only in highly differentiated cell types such as olfactory receptors and immune response factors [Pan et al., 2007; Zhao et al., 2007]. In accordance with their poised nature, bivalent promoters were most often found associated with genes for developmental transcription factors and morphogens [Mikkelsen et al., 2007; Pan et al., 2007; Zhao et al., 2007].

In addition to the role of histone trimethylation in establishing transcription control in ES cells, one ChIP-chip study has explored the role of H3K4 dimethylation (H3K4me2) in mouse hematopoetic progenitor cell lines (U-EML cells) [Orford et al., 2008]. In accordance with the findings described above, promoters associated with H3K4me3 and H3K4me2 drove gene expression at a high rate (79% expression). Genes associated with only H3K4me2 at their promoters were expressed at a much lower level (21% expression), and encompassed many genes related to hematopoiesis. Upon differentiation to an erythroid line (E-EML), H3K4me2+/H3K4me3− genes demonstrated a bimodal behavior, with erythroid-specific genes gaining H3K4 trimethylation, and those specific to myeloid and lymphoid cell types remaining H3K4me2+/H3K4me3− or losing H3K4 methylation altogether. Thus, H3K4 dimethylation presents another mechanism by which genes can be "poised" in multipotent progenitor cells. Interestingly, mouse ES cells lack H3K4me2+/H3K4me3− associated promoters, indicating that these genes must be poised at some point during the differentiation from stem cells to hematopoietic progenitor cells.

Methylation of CpG sites on DNA also plays an important role in regulation of gene expression, and two recent studies have provided insight into global trends of this modification in ES and differentiated cells. The first used genome-wide bisulphate sequencing to interrogate mouse ES cells [Meissner et al., 2008]; the second employed antibodies against methylated DNA to survey mouse ES cell promoters in a ChIP-chip assay [Fouse et al., 2008]. Promoters fell into one of two classes, those with high CpG density (HCP) which often contained a CpG island, or those with low CpG density (LCP). Of note, the likelihood of methylation at a specific CpG was inversely correlated with CpG density: at HCP promoters less than 1% of CpGs were methylated, whereas ~90% of CpGs at LCPs carried this mark [Meissner et al., 2008]. Ontologically, greater than 50% of the unmethylated genes coded housekeeping genes, while methylated genes were enriched for specialized products such as sensory receptors and cell signaling molecules [Fouse et al., 2008].

Strikingly, methylation of DNA and histones seems to be correlated: of genes with H3K4me3 (alone or bivalent), only 40–50% have DNA methylation, while 87% of genes lacking both histone marks show CpG methylation [Mikkelsen et al., 2007; Fouse et al., 2008]. Furthermore, the presence of H3K4me2 seems to be a strong predictor of decreased DNA methylation levels [Meissner et al., 2008]. This evidence is strongly suggestive of a global model in which histone and DNA modifications work synergistically to drive or repress transcription in a developmentally appropriate way. The generally unmethylated HCPs can be dynamically regulated by histone modifications, with more permanent silencing by DNA methylation occurring where appropriate upon differentiation. Concurrently, LCP regulation is less dependent on histone methylation, and more so on dynamic methylation and demethylation of DNA [Meissner et al., 2008].

Although ChIP-based methods are widely employed, there are several technical challenges. The one major limitation of ChIP is a requirement for highly specific antibodies against the desired target. In many cases, such antibodies may not exist or be readily available. Alternative means, such as introduction of an epitope tag or DNA adenine methyltransferase identification (DamID) methods can only be used in engineered cell systems. DamID detects potential binding sites for a protein by analyzing DNA adenine methylation in cells that contain the protein of interested fused to *E. coli* DNA adenine methyltransferase [van Steensel and Henikoff, 2000]. Most recently, DamID has been successfully applied to identify human genomic regions that associate with nuclear lamina which define large repressive chromosomal domains in the nucleus [Guelen et al., 2008]. Another major limitation is selective reactivity of formaldehyde. Although formaldehyde is quite capable of immobilizing histones on DNA, it is not able to crosslink some transcription factors, such as the *lac* repressor to its binding site in vitro [Solomon and Varshavsky, 1985]. Due this limitation, other crosslinking reagents have been used in combination with formaldehyde to increase the efficiency of crosslinking of transcription factors to their cognate sites in the genome [Kurdistani et al., 2002].

In addition to the technical challenges, there are analytical challenges due to the lack of a centralized analytical repository of ChIP-based genomic data. These challenges should be overcome with more coordinated efforts at the three main genomic databases to enable more streamlined data submission and analysis methods. Over the next few years, we expect to see an even greater explosion of ChIP-based analyses of genome function from individual laboratories and from the ENCODE and modENCODE consortia. Surely, there will be more dynamic maps of chromatin states and regulation capturing even more temporal and molecular details of genome regulation. In addition, a greater diversity of cell types and tissues will be analyzed by genome-wide ChIP. We also expect to see greater integration of information obtained from genome-wide ChIP and other genomic and nuclear features such as sequence variants and subnuclear localization. Lastly, we will see more mechanism-based analysis of genome function coupling other functional genomic resources (such as siRNAs) with various implementation of ChIP.

In the last 25 years, we have seen ChIP grow from an obscure, but promising technique for probing single loci in vivo, to a crucial workhorse of molecular genomics. So much so that we must apologize to the numerous researchers whose work couldn't be mentioned in this short perspective due to space limitations. Undoubtedly, the next 25 years will be just as fruitful.

ACKNOWLEDGMENTS

We thank Katharine Cecchini for comments on the manuscript. Work in the laboratory is supported by Rita Allen Foundation, Sidney Kimmel Foundation for Cancer Research, Yale Comprehensive Cancer Center, and Yale University School of Medicine.

REFERENCES

Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merkenschlager M, Fisher AG. 2006. Chromatin signatures of pluripotent cell lines. Nat Cell Biol 8:532–538.

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. Cell 129:823–837.

Bell AC, West AG, Felsenfeld G. 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell 98:387–396.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125:315–326.

Carroll JS, Liu XS, Brodsky AS, Li W, Meyer CA, Szary AJ, Eeckhoute J, Shao W, Hestermann EV, Geistlinger TR, Fox EA, Silver PA, Brown M. 2005. Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. Cell 122:33–43.

Carroll JS, Meyer CA, Song J, Li W, Geistlinger TR, Eeckhoute J, Brodsky AS, Keeton EK, Fertuck KC, Hall GF, Wang Q, Bekiranov S, Sementchenko V, Fox EA, Silver PA, Gingeras TR, Liu XS, Brown M. 2006. Genome-wide analysis of estrogen receptor binding sites. Nat Genet 38:1289–1297.

Fouse SD, Shen Y, Pellegrini M, Cole S, Meissner A, Van Neste L, Jaenisch R, Fan G. 2008. Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. Cell Stem Cell 2:160–169.

Gilmour DS, Lis JT. 1984. Detecting protein-DNA interactions in vivo: Distribution of RNA polymerase on specific bacterial genes. Proc Natl Acad Sci USA 81:4275–4279.

Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, van Steensel B. 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature 453:948–951.

Hecht A, Strahl-Bolsinger S, Grunstein M. 1996. Spreading of transcriptional repressor SIR3 from telomeric heterochromatin. Nature 383:92–96.

Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, Ren B. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet 39:311–318.

Iyer VR, Horak CE, Scafe CS, Botstein D, Snyder M, Brown PO. 2001. Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. Nature 409:533–538.

Johnson DS, Mortazavi A, Myers RM, Wold B. 2007. Genome-wide mapping of in vivo protein-DNA interactions. Science 316:1497–1502.

Kim TH, Barrera LO, Zheng M, Qu C, Singer MA, Richmond TA, Wu Y, Green RD, Ren B. 2005. A high-resolution map of active promoters in the human genome. Nature 436:876–880.

Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, Zhang MQ, Lobanenkov VV, Ren B. 2007. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell 128:1231–1245.

Kurdistani SK, Robyr D, Tavazoie S, Grunstein M. 2002. Genome-wide binding map of the histone deacetylase Rpd3 in yeast. Nat Genet 31:248–254.

Lobanenkov VV, Nicolas RH, Adler VV, Paterson H, Klenova EM, Polotskaja AV, Goodwin GH. 1990. A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5′-flanking sequence of the chicken c-myc gene. Oncogene 5:1743–1753.

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R, Lander ES. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454:766–770.

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448:553–560.

Orford K, Kharchenko P, Lai W, Dao MC, Worhunsky DJ, Ferro A, Janzen V, Park PJ, Scadden DT. 2008. Differential H3K4 methylation identifies developmentally poised hematopoietic genes. Dev Cell 14:798–809.

Orlando V, Paro R. 1993. Mapping Polycomb-repressed domains in the bithorax complex using in vivo formaldehyde cross-linked chromatin. Cell 75:1187–1198.

Pan G, Tian S, Nie J, Yang C, Ruotti V, Wei H, Jonsdottir GA, Stewart R, Thomson JA. 2007. Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. Cell Stem Cell 1:299–312.

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, Young RA. 2000. Genome-wide location and function of DNA binding proteins. Science 290:2306–2309.

Robertson AG, Bilenky M, Tam A, Zhao Y, Zeng T, Thiessen N, Cezard T, Fejes AP, Wederell ED, Cullum R, Euskirchen G, Krzywinski M, Birol I, Snyder M, Hoodless PA, Hirst M, Marra MA, Jones SJ. 2008. Genome-wide relationship between histone H3 lysine 4 mono- and tri-methylation and transcription factor binding. Genome Res 18:1906–1917.

Roh TY, Ngau WC, Cui K, Landsman D, Zhao K. 2004. High-resolution genome-wide mapping of histone modifications. Nat Biotechnol 22:1013–1016.

Solomon MJ, Varshavsky A. 1985. Formaldehyde-mediated DNA-protein crosslinking: A probe for in vivo chromatin structures. Proc Natl Acad Sci USA 82:6470–6474.

van Steensel B, Henikoff S. 2000. Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. Nat Biotechnol 18:424–428.

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. 1995. Serial analysis of gene expression. Science 270:484–487.

Vostrov AA, Quitschke WW. 1997. The zinc finger protein CTCF binds to the APBbeta domain of the amyloid beta-protein precursor promoter. Evidence for a role in transcriptional activation. J Biol Chem 272:33353–33359.

Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh TY, Peng W, Zhang MQ, Zhao K. 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet 40:897–903.

Wei CL, Wu Q, Vega VB, Chiu KP, Ng P, Zhang T, Shahab A, Yong HC, Fu Y, Weng Z, Liu J, Zhao XD, Chew JL, Lee YL, Kuznetsov VA, Sung WK, Miller LD, Lim B, Liu ET, Yu Q, Ng HH, Ruan Y. 2006. A global map of p53 transcription-factor binding sites in the human genome. Cell 124:207–219.

Zhao XD, Han X, Chew JL, Liu J, Chiu KP, Choo A, Orlov YL, Sung WK, Shahab A, Kuznetsov VA, Bourque G, Oh S, Ruan Y, Ng HH, Wei CL. 2007. Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell Stem Cell 1:286–298.