

Methods in
Molecular Biology 791

Springer Protocols



Trygve O. Tollefsbol *Editor*

Epigenetics Protocols

Second Edition

 Humana Press

Epigenetics Protocols

Second Edition

Edited by

Trygve O. Tollefsbol

Department of Biology, University of Alabama at Birmingham, Birmingham, AL, USA

 **Humana Press**

PARSE EX.1005.002

Editor

Trygve O. Tollefsbol
Department of Biology
University of Alabama at Birmingham
Birmingham, AL 35294-1170, USA
trygve@uab.edu

ISSN 1064-3745 e-ISSN 1940-6029
ISBN 978-1-61779-315-8 e-ISBN 978-1-61779-316-5
DOI 10.1007/978-1-61779-316-5
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011935043

© Springer Science+Business Media, LLC 2011

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Humana Press, c/o Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Humana Press is part of Springer Science+Business Media (www.springer.com)

Using ChIP-Seq Technology to Generate High-Resolution Profiles of Histone Modifications

Henriette O'Geen, Lorigail Echipare, and Peggy J. Farnham

Abstract

The dynamic modification of DNA and histones plays a key role in transcriptional regulation through altering the packaging of DNA and modifying the nucleosome surface. These chromatin states, also referred to as the epigenome, are distinctive for different tissues, developmental stages, and disease states and can also be altered by environmental influences. New technologies allow the genome-wide visualization of the information encoded in the epigenome. For example, the chromatin immunoprecipitation (ChIP) assay allows investigators to characterize DNA–protein interactions in vivo. ChIP followed by hybridization to microarrays (ChIP-chip) or by high-throughput sequencing (ChIP-seq) are both powerful tools to identify genome-wide profiles of transcription factors, histone modifications, DNA methylation, and nucleosome positioning. ChIP-seq technology, which can now interrogate the entire human genome at high resolution with only one lane of sequencing, has recently surpassed ChIP-chip technology for epigenomic analyses. Importantly, for the study of primary cells and tissues, epigenetic profiles can be generated using as little as 1 µg of chromatin. In this chapter, we describe in detail the steps involved in performing ChIP assays (with a focus on characterizing histone modifications in primary cells) either manually or using the IP-Star ChIP robot, followed by a detailed protocol to prepare successful libraries for Illumina sequencing. Critical quality control checkpoints are discussed. Although not a focus of this chapter, we also point the reader to several methods by which massive ChIP-seq data sets can be analyzed to extract the tremendous information contained within.

Key words: Chromatin immunoprecipitation, ChIP-seq, Next generation sequencing, Epigenomics, Histone modifications, IP-Star, ChIP robot

1. Introduction

Although the genetic information encoded in our DNA plays a major role in specifying our individual phenotypes, it is becoming increasingly clear that epigenetic information is also an important contributor to our mental and physical attributes (1–8). Our epigenome is defined as methylated DNA and modified histone

proteins (around which both methylated and unmethylated DNA are wrapped). DNA methylation and histone modifications undergo global changes during transitions in developmental states and in diseases such as cancer and therefore are major contributors to the dynamic nature of chromatin. Histone modifications such as acetylation (e.g., acetylation of lysine 9 of histone H3 which is called H3K9Ac) are typically associated with open and accessible chromatin regions, while histone methylation can be associated with either open or compacted (also referred to as heterochromatic) chromatin regions, depending on the specific histone amino acid that is methylated (9–12). For example, mono- or trimethylation of lysine 4 on histone H3 (H3K4me1 or H3K4me3) and trimethylation of histone H3 on lysine 36 (H3K36me3) are associated with open chromatin. However, each of these marks represents a unique category of open chromatin, with H3K4me3 marking gene promoter regions, H3K4me1 marking transcriptional enhancers, and H3K36me3 marking transcribed regions of the genome. In contrast, trimethylation of lysines 9 and 27 on histone H3 (H3K9me3 and H3K27me3, respectively) is associated with compacted chromatin regions resulting in repression of target genes. Although both of these modifications mark repressive chromatin, H3K9me3 and H3K27me3 regulate distinct sets of target genes; H3K27me3 predominantly represses homeobox transcription factors and H3K9me3 predominantly targets the genes of zinc finger transcription factors (13). Knowing the genome-wide pattern of single histone modifications, such as the six marks described above, provides a great deal of information about cell identity and disease state (10, 14–17). Accordingly, these six marks have been selected to provide “roadmaps” of the epigenomic profiles of primary cells by the Roadmap Epigenome Mapping Centers (<http://www.roadmapepigenomics.org/>). However, it is also becoming increasingly clear that different combinations of histone marks can provide even more detailed information. For example, the presence of both the open chromatin mark H3K4me3 and the compacted chromatin mark H3K9me3 at a promoter can identify imprinted genes (18). We are just beginning to understand the interrelationships between specific histone modifications and transcriptional regulation and more insights will certainly be forthcoming with the analysis of more and more epigenetic profiles.

Currently, the method of choice to study the epigenome is the chromatin immunoprecipitation (ChIP) assay. To perform a ChIP assay, proteins such as histones or transcription factors are covalently crosslinked to their genomic DNA substrates in living cells. This provides an opportunity to take a snapshot of histone or other protein–DNA interactions in a given cell type, in cells taken at different developmental stages, or in cells altered by disease. After isolation and fragmentation of chromatin, the protein–DNA complexes are captured using antibodies specific to the histone or transcription factor of interest. After reversal of crosslinks, the

ChIP DNA is then purified and analyzed by either hybridization to microarrays (ChIP-chip) or by high-throughput sequencing (ChIP-seq). While multiple DNA microarrays are needed to cover the entire human genome, resulting in high costs for comprehensive studies, ChIP-seq offers the possibility to interrogate the entire genome in one sequencing run. Therefore, ChIP-seq has generally replaced ChIP-chip for comprehensive epigenomic studies. To date there are four high-throughput sequencing platforms that have been used for ChIP-seq (see ref. 19 for review): (1) 454 Genome Sequencer FLX from Roche (20, 21), (2) Genome Analyzer GA2 from Illumina (9, 18, 22), (3) Sequencing by Oligo Ligation/ Detection (SOLiD) from Applied Biosystems (23, 24), and (4) true Single Molecular Sequencing (tSMS) by Helicos (19, 25). Unlike the first three methods that use an amplification step, the Helicos platform sequences unamplified DNA. Another difference among the platforms include read length; the Illumina, SOLiD, and Helicos platforms produce short read lengths of on average 35–50 bp whereas the 454 platform produces longer reads of 200–400 bp. Most ChIP-seq studies to date have used the Illumina sequencing platform (sometimes called Solexa sequencing). Therefore, this chapter describes in detail the steps needed to prepare ChIP samples and libraries for high-throughput sequencing using the Illumina GA2 platform and includes descriptions of quality control steps necessary to ensure a successful ChIP-seq experiment (see Fig. 1).

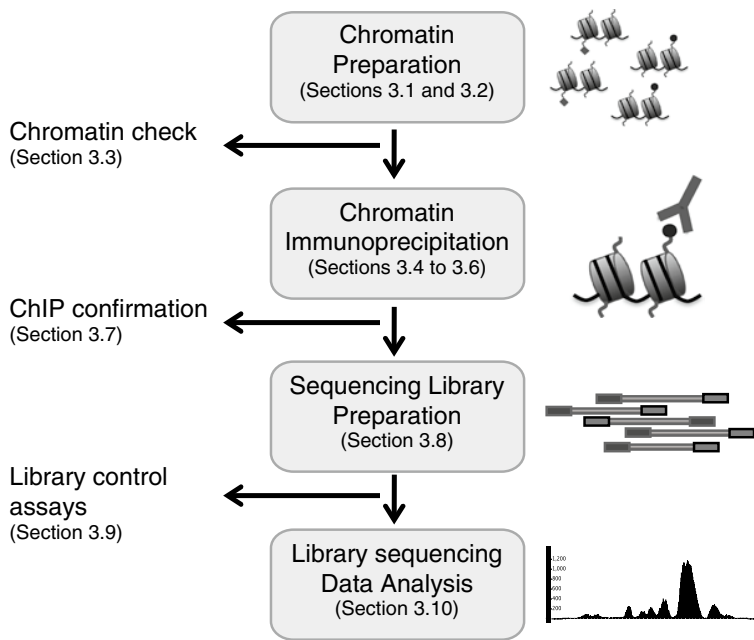


Fig. 1. ChIP-seq diagram summarizing the main experimental steps and quality check points.

2. Materials

2.1. Crosslinking Reagents

1. Crosslinking reagent: formaldehyde solution (37% w/w).
2. Stopping reagent: glycine (electrophoresis grade).
3. Wash solution: phosphate-buffered saline (PBS).

2.2. Chromatin Preparation Reagents

1. Protease inhibitor stock solutions (store in small aliquots at -20°C): 10 mg/ml aprotinin (in water), 10 mg/ml leupeptin (in water), 100 mM PMSF (in isopropanol).
2. Cell lysis buffer (store at room temperature): 5 mM PIPES pH 8, 85 mM KCl. Add igepal fresh each time to give a final concentration of 1% (10 $\mu\text{l}/\text{ml}$). Warm buffer in 37°C water bath and vortex briefly to help mixing of igepal. After mixing has occurred, place buffer containing igepal on ice to allow solution to cool down and then add protease inhibitors [PMSF (10 $\mu\text{l}/\text{ml}$ f.c.), aprotinin (1 $\mu\text{l}/\text{ml}$ f.c.), and leupeptin (1 $\mu\text{l}/\text{ml}$ f.c.)].
3. Nuclei lysis buffer (store at room temperature): 50 mM Tris-HCl pH 8, 10 mM EDTA, and 1% (w/v) SDS. Place buffer on ice right before use to avoid precipitation of SDS and add protease inhibitors [PMSF (10 $\mu\text{l}/\text{ml}$ f.c.), aprotinin (1 $\mu\text{l}/\text{ml}$ f.c.), and leupeptin (1 $\mu\text{l}/\text{ml}$ f.c.)] just prior to use.
4. Bioruptor UCD-200 (Diagenode) or equivalent is used for sonication.

2.3. Chromatin Check Reagents

1. Elution buffer (store at room temperature): 50 mM NaHCO_3 , 1% (w/v) SDS.
2. DNase-free RNase A (Fermentas; 10 mg/ml).
3. QIAquick PCR purification kit (QIAGEN).
4. A NanoDrop 1000 is used to determine the concentration of double-stranded DNA samples. This instrument is invaluable for measuring low DNA concentrations (e.g., 10 ng/ μl) and for small sample volumes (as little as 1 μl of sample can be measured).

2.4. Chromatin Immunoprecipitation Reagents

Note, the reagents listed in steps 2 and 3 of Subheading 2.4 are required for manual ChIP assays; for automated ChIP assays, use the Auto ChIP kit for the IP-Star (Diagenode).

1. ChIP grade antibodies specific for the following six histone modifications:
 - (a) H3K4me3: Anti-Tri-Methyl-Histone H3 (Lys4) (C42D8) rabbit monoclonal antibody (CST #9751S).
 - (b) H3K9ac: Anti-acetyl-Histone H3 (Lys9) rabbit antibody (Millipore #07-352).

- (c) H3K27me3: Anti-Tri-Methyl-Histone H3 (Lys27) (C36B11) rabbit monoclonal antibody (CST #9733S).
 - (d) H3K9me3: Anti-Tri-Methyl-Histone H3 (Lys9) rabbit antibody (CST #9754S).
 - (e) H3K36me3: Anti-Tri-Methyl-Histone H3 (Lys36) rabbit antibody (CST #9763S).
 - (f) H3K4me1: Anti-Mono-Methyl-Histone H3 (Lys4) rabbit antibody (Diagenode #pAb-037-050).
2. Protease inhibitor stock solutions (store in small aliquots at -20°C): 10 mg/ml aprotinin (in water), 10 mg/ml leupeptin (in water), 100 mM PMSF (in isopropanol).
 3. IP dilution buffer (store at 4°C): 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% (v/v) igepal, 0.25% (w/v) deoxycholic acid, 1 mM EDTA pH 8. Add protease inhibitors [PMSF (10 $\mu\text{l}/\text{ml}$ f.c.), aprotinin (1 $\mu\text{l}/\text{ml}$ f.c.), and leupeptin (1 $\mu\text{l}/\text{ml}$ f.c.)] just prior to use. This buffer is used to adjust the salt and SDS concentrations for the immuno-precipitation step.

2.5. Capture the Antibody/Chromatin Complexes and to Reverse Crosslinks

Note, the reagents listed below are required for manual ChIP assays; for automated ChIP assays, use the Auto ChIP kit for the IP-Star (Diagenode).

1. Magnetic protein G beads (Cell Signaling Technology) and magnetic rack. Do not use magnetic beads that have been blocked with foreign DNA, such as herring sperm or salmon sperm DNA. This may result in sequencing of the blocking DNA, resulting in lower quality ChIP-seq data. Although protein G binds antibodies from a variety of species (rabbit, mouse, goat, etc.) with high affinity, magnetic protein A beads can be used if desired.
2. IP wash buffer 1 (store at 4°C): same as IP dilution buffer, but without protease inhibitors.
3. IP wash buffer 2 (store at room temperature): 100 mM Tris-HCl pH 9, 500 mM LiCl, 1% (v/v) igepal, 1% (w/v) deoxycholic acid.
4. IP wash buffer 3 (store at room temperature): 100 mM Tris-HCl pH 9, 500 mM LiCl, 150 mM NaCl, 1% (v/v) igepal, 1% (w/v) deoxycholic acid.
5. Elution buffer (store at room temperature): 50 mM NaHCO_3 , 1% (w/v) SDS.
6. 5 M NaCl.

2.6. DNA Purification

1. DNase-free RNase A (Fermentas; 10 mg/ml).
2. QIAquick PCR purification kit (QIAGEN).

2.7. ChIP Confirmation

1. SYBR-Green qPCR mix, such as SYBR Green JumpStart Taq ReadyMix (SIGMA).
2. Positive and negative control primer sets (see Note 1).

2.8. Sequencing Library Preparation

1. End-It DNA END Repair Kit (Epicentre).
2. Klenow (3'–5' exo minus) (NEB; 5,000 U/ml).
3. 100 mM dATP.
4. LigaFast DNA ligase (Promega; 3 U/μl).
5. Oligo-only kits for single end or paired end read sequencing are available from Illumina (#FC-102-1003 and PE-102-1003, respectively). Alternatively, adapter oligos and PCR primers compatible with Illumina sequencing can be purchased elsewhere; HPLC purification is recommended. The paired end DNA oligonucleotides are more universal since the resulting library can be sequenced with either single end or paired end sequencing primers. The following stock solutions are prepared: 15 μM Paired End Adapter Oligo mix, 25 μM Paired End PCR primer 1.01, and 25 μM Paired End PCR primer 2.01.

Paired End DNA oligonucleotide sequences (Oligonucleotide sequences© 2006 and 2008 Illumina, Inc. All rights reserved).

PE Adapters

5' P-GATCGGAAGAGCGGTTCAGCAGGAATGCCGAG

5' ACACTCTTTCCCTACACGACGCTCTTCCGATCT

PE PCR Primer 1.01

5' AATGATACGGCGACCACCGAGATCTACACTCTTTCC
CTACACGACGCTCTTCCGATCT

PE PCR Primer 2.01

5' CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATT
CCTGCTGAACCGCTCTTCCGATCT

6. 2% Agarose precast E-Gel® (Invitrogen #G501802), loading dye such as TrackIt™ Cyan/Orange Loading Buffer (Invitrogen #10482-028), 50 or 100 bp DNA markers (e.g., TrackIt™ 50 bp DNA Ladder, Invitrogen #10488-043).
7. QIAquick Gel Extraction Kit (QIAGEN #28704).
8. Phusion DNA polymerase (NEB #F531).
9. QIAquick PCR purification kit (QIAGEN #28104) and MinElute PCR purification kit (QIAGEN #28004).
10. Agencourt AMPure system (Beckman Coulter Genomics #A29152).

2.9. Library Control Assay

1. DNA High Sensitivity Kit (Agilent Cat# 5067–4626) for use with the Agilent 2100 Bioanalyzer.

2. SYBR-Green qPCR mix, such as SYBR Green JumpStart Taq ReadyMix (SIGMA #S4438).
3. Positive and negative control primer sets (see Note 1).

2.10. Sequencing Reagents

These reagents will be supplied by the sequencing facility.

3. Methods

3.1. Preparation of Crosslinked Cells

1. Cell cultures should be healthy and not density-arrested prior to crosslinking. For primary cells or tissues, the samples can be snap frozen in liquid nitrogen immediately after collection or snap frozen after crosslinking. The amount of cells needed for ChIP-seq will vary depending on the antibody used and the abundance of the histone mark of interest. In general between 100,000 and 500,000 cells are used per histone antibody. In a chemical hood, prepare 1% formaldehyde solution in PBS and add directly to frozen cell pellet. Resuspend cell pellet by pipetting up and down. Alternatively for cultured cells, add formaldehyde (37% stock) directly to tissue culture media to a final concentration of 1%.
2. Rotate primary cells in a tightly closed tube or rock cultured cells on a shaking platform for 10 min at room temperature. Do not crosslink for longer periods since this may cause cells to form aggregates that do not sonicate efficiently.
3. Stop crosslinking reaction by adding glycine to a final concentration of 0.125 M. We use a 10× (1.25 M) stock solution. Continue to rotate/rock at room temperature for 5 min.
4. For primary cells or other cells crosslinked in suspension, centrifuge cells at 430 rcf for 5 min at 4°C, discard the solution, wash the pellet twice with ice-cold 1× PBS (mix by pipetting, pellet cells by centrifugation at 430 rcf for 5 min at 4°C and discard wash solution). For adherent cells, pour off media and rinse plates twice with ice-cold 1× PBS and pour off wash solution. Using a cell scraper, transfer adherent cells from the culture dish to a 15-ml conical tube on ice. Centrifuge the crosslinked cells at 430 rcf for 5 min at 4°C. It is important to carefully aspirate supernatants so as to not lose cells. Note: media containing formaldehyde should be treated as hazardous waste.
5. Cells may be used immediately for a chromatin preparation or snap frozen in liquid nitrogen and stored at -80°C.

3.2. Preparation of Chromatin

1. If using frozen crosslinked cells thaw them on ice; keep all cells and chromatin samples on ice at all times. Prepare the cell lysis buffer (1 ml cell lysis buffer per 1×10^7 cells): add Igepal

- (10 μ l per ml cell lysis buffer, agitate at 37°C to dissolve, cool on ice), then add protease inhibitors [PMSF (10 μ l/ml), aprotinin (1 μ l/ml), and leupeptin (1 μ l/ml)]. Resuspend cell pellet in freshly prepared ice-cold cell lysis buffer by pipetting. The final volume of cell lysis buffer should be sufficient so that there are no clumps of cells. Incubate on ice for 15 min.
2. Homogenize cells using a glass dounce homogenizer (type B) to break open the cells and release nuclei. Homogenize cells on ice with 20 strokes. Omit this step when processing less than one million cells.
 3. Centrifuge cells at 430 rcf for 5 min at 4°C.
 4. Discard the supernatant and resuspend the nuclear pellet in nuclei lysis (NL) buffer plus protease inhibitors. Be careful not to use too much NL buffer as it may lead to dilute chromatin; we suggest ~20 μ l/10⁶ cells. Incubate on ice for 30 min.
 5. An optional flash-freezing step may help break open nuclei more efficiently. This step is critical if the homogenizing in step 2 is omitted. After incubation of nuclei in NL buffer for 30 min, flash freeze samples in liquid nitrogen, thaw at room temperature (once thawed, immediately transfer to ice; do not allow samples to warm up to room temperature), and proceed to sonication.
 6. Sonicate cells in a coldroom and/or on ice to achieve average chromatin length of 200–500 bp (see Note 2). Larger chromatin fragments can negatively influence data quality and can lead to failure of the ChIP-seq experiment. Therefore, before processing large quantities of cells, sonication conditions should be optimized for each cell type (see Note 3).
 7. Transfer sonicated samples into an Eppendorf tube and centrifuge using a microcentrifuge at 10,000 rcf for 10 min at 4°C. Carefully transfer the supernatant (sonicated chromatin) to a new tube while avoiding cell debris. Keep sonicated chromatin at 4°C while performing quantification and determining chromatin size; then proceed with the ChIP assays (see Note 4).

3.3. Determination of Chromatin Size and Concentration

1. Take an aliquot of chromatin sample from Subheading 3.2, step 7 prepared above. A typical size determination uses chromatin from 100,000 to 200,000 cells (see Note 5).
2. Add ChIP elution buffer to a total volume of 100 μ l and then 12 μ l 5 M NaCl to give a final salt concentration of 0.54 M. Boil samples in a water bath for 20 min to reverse crosslinks.
3. Allow sample to cool down, add 1 μ l DNase-free RNase (10 mg/ml), and incubate for 20 min at 37°C. This step is important because the presence of RNA results in false estimation of chromatin size.

4. Purify DNA using a PCR purification kit, elute DNA in 25 μ l water. Measure chromatin concentration by NanoDrop and calculate the chromatin yield (see Note 6).
5. Run remaining chromatin on a 1.2% agarose gel to visualize average size. If the chromatin is larger than $\sim$ 600 bp, adjust the sonication conditions by adding more pulses and repeat steps 1–5.

3.4. Chromatin Immunoprecipitation

ChIP is usually done within the same day as the chromatin preparation to avoid any concern about the quality of chromatin (see Note 4). The steps detailed in Subheadings 3.4 and 3.5 are for manual ChIP assays. However, ChIP reactions can be automated using a ChIP robot (e.g., the IP-Star from Diagenode) (see Note 7).

1. Measure volume of chromatin and divide chromatin as needed. The amount of chromatin needed for each ChIP reaction varies depending on the histone modification. For histone marks covering a small portion of the genome displaying sharp peaks, such as H3K4me3, we typically use 1 μ g chromatin. Based on our experience, we prefer to use 5 μ g chromatin for spreading histone marks such as H3K9me3 or H3K36me3 that cover large portions of the genome.
2. Optional: An IgG negative control sample can be included along with experimental antibodies. However, oftentimes chromatin cannot be spared for a control ChIP when using small amounts of primary cells. In this case, one can rely on regions bound by the activating histone marks as negative controls for the repressive histone marks and vice versa (Fig. 2a).

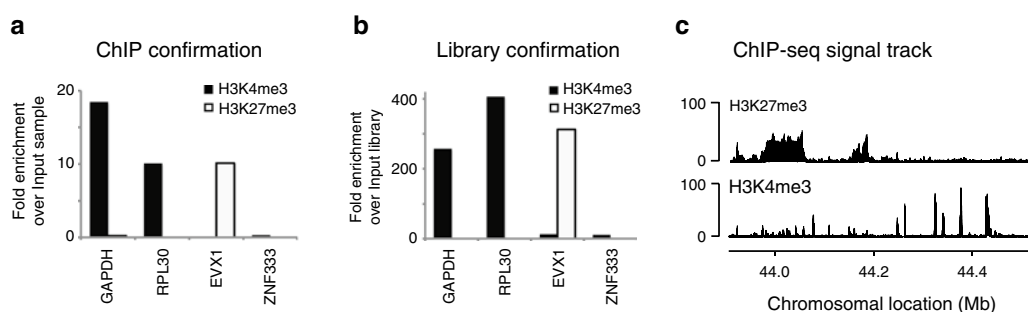


Fig. 2. ChIP-seq experiments using the IP-Star ChIP robot. ChIP assays using antibodies specific for H3K4me3 and H3K27me3 were performed using 1 μ g chromatin from Ntera2 cells and the IP-Star ChIP robot (Diagenode). Libraries were prepared as outlined in Subheading 3.8. Quantitative PCR confirms specific enrichment over input in (a) the ChIP samples and (b) the ChIP-seq libraries. Black bars represent H3K4me3 enrichment and white bars represent H3K27me3 enrichment. Primer sets used are shown on the x-axis. GAPDH and RPL30 are positive control primer sets for H3K4me3 and negative control primer sets for H3K27me3. EVX1 was used as a positive primer set for H3K27me3 and a negative primer set for H3K4me3. The ZNF333 primer set was used as a negative control for both histone marks. (c) ChIP-seq binding patterns for H3K4me3 and H3K27me3 obtained from samples prepared using the IP-Star are shown for a region on chromosome 17 encompassing the HOXB gene cluster. The peak height is plotted along the y-axis; chromosomal coordinates (hg18 coordinates) are shown on the x-axis. Samples were sequenced at the DNA Technologies Core at UC Davis (http://genomecenter.ucdavis.edu/dna_technologies/).

3. Save volume corresponding to 500 ng of chromatin to prepare an input sample (often also referred to as total DNA). Store the reserved amount at -20°C until the next day and then reverse the crosslinks in the input chromatin at the same time as the crosslinks in the ChIP samples are reversed.
4. Dilute chromatin fivefold with ice-cold IP Dilution buffer (1 volume chromatin and 4 volumes IP Dilution buffer) containing protease inhibitors.
5. Add an antibody specific to the histone mark of interest to capture the protein/chromatin complexes (see Note 8). Although antibody amounts are determined empirically, we typically use between 1 and 5 μg antibody per ChIP assay. Always record catalogue number and lot number of antibodies used.
6. Incubate 8–16 h on a rotating platform at 4°C .

3.5. Capture of Antibody/Chromatin Complexes and Reversal of Crosslinks

Step 1 is carried out at 4°C , whereas steps 2–9 are carried out at room temperature.

1. Add 15 μl magnetic protein G beads to each ChIP sample ranging from 1 to 5 μg chromatin starting material and incubate on a rotating platform for 2 h at 4°C .
2. At room temperature, allow beads to settle for 1 min in a magnetic separation rack. Carefully remove the supernatant without disturbing magnetic beads.
3. Wash magnetic beads two times with IP Dilution buffer (take tubes out of magnetic rack and mix by pipetting). Efficient washing is critical to reduce background. Avoid cross contamination of samples and loss of magnetic beads.
4. Wash magnetic beads two times with IP wash buffer 2 (take tubes out of magnetic rack and mix by pipetting). Discard all wash solution after final wash.
5. Wash once with the higher stringency IP wash buffer 3. Discard wash solutions.
6. Elute antibody/chromatin complexes by adding 100 μl elution buffer per ChIP sample. Shake samples on vortexer for 30 min.
7. Allow beads to settle for 1 min in a magnetic separation rack. Carefully transfer the supernatant containing antibody/chromatin complexes to a siliconized tube.
8. Add 12 μl of 5 M NaCl per 100 μl elution buffer mix to give a final concentration of 0.54 M NaCl.
9. At this point, thaw the input sample from the previous day (Subheading 3.4, step 3). Dilute 1 volume input sample with 4 volumes ChIP elution buffer (e.g., add 80 μl ChIP elution

buffer to 20 µl input sample). Add 12 µl of 5 M NaCl per 100 µl elution buffer mix.

10. Incubate all samples in a 67°C water bath overnight to reverse formaldehyde crosslinks.

3.6. DNA Purification

1. Allow samples to cool, add 1 µl of RNaseA; incubate at 37°C for 20 min.
2. Purify DNA with a PCR clean up kit, one column per sample. Elute each sample with 40 µl EB buffer.
3. Assess ChIP enrichments by quantitative PCR (qPCR) before proceeding to preparation of Solexa libraries.

3.7. ChIP Confirmation

Enrichment of histone marks in the ChIP samples are determined by quantitative real-time PCR (qPCR). The input sample is diluted with EB to give a final concentration of 2 ng/µl and serves as a reference. Prepare a master reaction mix for each library with triplicate reactions per primer set. Add extra reagents for 10% of the total number of reagents to account for loss of volume. Add 14 µl of reaction mix to each PCR reaction well. Add 2 µl primer mix to each well.

Recipe for one reaction:

1 µl	Undiluted ChIP sample or diluted Input sample (2 ng/µl)
4.5 µl	Nuclease-free H2O
7.5 µl	2× SYBR Green mix (containing polymerase)
2 µl	5 µM target primer mix (containing both Forward and Reverse primers)
15 µl	Total reaction volume

Amplify using the following PCR protocol:

- 3 min at 95°C.
40 cycles of 30 s at 95°C, then 30 s at 60°C.

Include a 70–95°C melting curve at the end of the qPCR program, reading all points or every 0.2°C.

Analyze the qPCR results by first manually determining the cycle threshold for each reaction across the plate within the linear range of the amplification curve. Calculate the average cycle threshold for each triplicate reaction of each sample. The relative DNA amount is then calculated for any given primer set as 2 to the power of the cycle threshold (cT) difference between input chromatin and ChIP samples, where cT is the average value.

Relative DNA amount = 2^(cT input – cT sample)

The enrichment is then calculated by comparing relative enrichment for the target and a negative control. This is accomplished by dividing the relative DNA amount of each sample for a target primer

set by the corresponding value for a negative control primer set. The resulting quotient represents the fold enrichment. The fold enrichment will vary depending on the histone marks as well as the location of the chosen target primer set (see Fig. 2a for an example).

3.8. Preparation of the Sequencing Library

The library protocol is based on the Illumina Sample Preparation Kit for Genomic DNA with some modifications. This protocol describes the preparation of libraries from ChIP DNA that are compatible with the Illumina sequencing platforms. Libraries are prepared from the ChIP sample as well as matching input DNA from the same cell type (see Note 9). Boiled chromatin samples should not be used since single-stranded DNA will lower the library preparation efficiency.

3.8.1. End-Repair

End-repair is performed using the “End-It DNA End Repair Kit” from Epicentre. This step ensures that all DNA fragments are converted to 5'-phosphorylated blunt-ended DNA. The entire ChIP DNA volume from Subheading 3.6 is used. Combine and mix the following components in a siliconized Eppendorf tube:

1–34 µl	ChIP DNA from Subheading 3.6 or 200 ng input DNA
5 µl	10× End-Repair Buffer
5 µl	10 mM ATP
5 µl	2.5 mM dNTP Mix
1 µl	End-Repair Enzyme Mix
50 µl	Total reaction volume

Incubate at room temperature for 45 min and purify DNA using a PCR purification kit (such as QIAquick PCR purification kit), elute in 34 µl EB buffer.

3.8.2. Addition of an “A” Base to the 3' End of DNA Fragments

Before starting, prepare stocks of 1 mM dATP from 100 mM dATP stock (e.g., add 5 µl of 100 mM dATP to 495 µl sterile RNase DNase free water), and store aliquots of 11 µl at –20°C. Once thawed, 1 mM dATP solution should not be refrozen. Combine and mix the following components in PCR tubes:

34 µl	DNA from Subheading 3.8.1
5 µl	10× Klenow buffer
10 µl	1 mM dATP
1 µl	Klenow fragment (3'–5' exo minus)
50 µl	Total reaction volume

Incubate for exactly 30 min at 37°C using a PCR machine. Purify DNA using a PCR purification kit (such as MinElute PCR purification kit), elute in 12 µl EB buffer.

3.8.3. Ligation of Adapters to DNA Fragments

The Paired End Adapter Oligo mix is diluted 1:10 in water before use to adjust for the small quantity of ChIP DNA.

Combine and mix the following components in a siliconized Eppendorf tube:

12 µl	DNA from Subheading 3.8.2
15 µl	2× DNA ligase buffer
1 µl	1:10 dilution of PE Adapter Oligo mix
2 µl	LigaFast DNA ligase
30 µl	Total reaction volume

Incubate for 15 min at room temperature. Purify DNA using a PCR purification kit (such as QIAquick PCR purification kit), elute in 19 µl EB buffer.

3.8.4. Size Selection of DNA Fragments

Size selection of the sample ensures removal of unused adapters and selection of proper fragment size for amplification and sequencing (see Note 10). We use precast agarose gels to minimize risk of contamination.

1. Dilute 10 µl of 6× Cyan/Orange with 50 µl EB buffer to obtain 1× Cyan/Orange buffer dye. Add 1 µl of 1× Cyan/Orange buffer dye to eluted DNA from Subheading 3.8.3.
2. Prerun e-gels according to manufacturer’s instructions.
3. Load 20 µl of appropriately diluted 50 or 100 bp DNA ladder (500 ng ladder per well) in one well per gel. Skip at least one well between marker and samples to avoid contamination.
4. Load 20 µl DNA with dye in each well. Skip at least one well between samples to avoid contamination.
5. Load 20 µl of EB buffer in each of the empty wells.
6. Run the gel until desired size separation is achieved (30 min for e-gels).
7. Take a gel picture to visualize sample.
8. Size select samples: using a fresh razor blade, excise two gel pieces of 200–400 bp and 400–600 bp (see Note 11). The DNA concentrations may be too low for the sample to be visible by eye; in this case, use markers as a guide. Keep exposure to UVlight to a minimum to reduce DNA damage. Alternatively, a non-UV transilluminator can be used. Gel slices can be stored at –20°C.
9. Solubilize gels at room temperature using the QIAGEN gel extraction buffer by shaking on the vortexer for 30 min.
10. Purify on one QIAquick column using QIAquick PCR Purification Kit. Elute in 25 µl EB buffer.

3.8.5. Amplification
of Adapter-Modified
DNA Fragments and
Gel Purification

Because we make libraries from both the small (200–400 bp) and the big (400–600 bp) size selected DNA fragments from Subheading 3.8.4, we prepare two amplification reactions (and two libraries) per ChIP sample. We also prepare a 200–400-bp and a 400–600-bp size-selected input library. Therefore, if all six histone modifications are analyzed, there will be 14 amplification reactions and 14 libraries (12 ChIP libraries and 2 input libraries). For these reactions, dilute Paired End primers 1:4 with sterile water.

Combine and mix the following components in PCR tubes:

23 µl	DNA from step 2
25 µl	2× Phusion DNA polymerase
1 µl	Paired End PCR primer 1.01 (1:4 dilution)
1 µl	Paired End PCR primer 2.01 (1:4 dilution)
50 µl	Total reaction volume

Amplify using the following PCR protocol:

- 30 s at 98°C.
- 15 cycles: 10 s at 98°C, 30 s at 65°C, 30 s at 72°C.
- 5 min at 72°C.
- Hold at 4°C.
- Perform 15 cycles of amplification.

3.8.6. Library Purification

Purify library samples from Subheading 3.8.5 using the Agencourt AMPure system following manufacturer’s instructions.

1. Mix Ampure beads thoroughly before addition.
2. Add 90 µl of Ampure beads to each 50 µl DNA sample from step 5. Pipette several times to ensure proper mixing.
3. Use magnetic rack to separate bead–DNA complexes and discard the supernatant. Allow beads to settle, this may take several minutes.
4. Wash bead–DNA complexes using 70% ethanol without disturbing the beads. Leave the tube in the magnetic rack and add 200 µl of 70% ethanol. After 30 s, discard 70% ethanol by pipetting.
5. Repeat wash one more time for a total of two washes.
6. Allow beads to air dry for 10–20 min.
7. Add 30 µl of EB buffer and elute DNA on the vortexer for 30 min.
8. Place tubes back in magnetic rack to collect DNA, transfer liquid to siliconized Eppendorf tubes. Store libraries at –20°C.

3.9. Library Quality Control Assays

3.9.1. Library Quantification

The constructed libraries are assessed on a Bioanalyzer using the DNA High Sensitivity chip. The High Sensitivity DNA chip allows sizing and quantification of DNA samples in the single-digit pg/ μ l concentration range; the sequencing flow cell is loaded according to the library concentration (check with your sequencing facility for their requirements). The Bioanalyzer also allows visualization of possible adapter contamination; adapter dimers are visible as a sharp peak at approximately 120 bp. Libraries having large adapter dimer peaks should not be sequenced (contaminating adapter dimers can be removed by an additional gel size selection step as described in 3.8.4). For libraries that are not quantifiable, five additional cycles can be performed as in Subheading 3.8.5. However, over-amplification should be avoided to reduce the risk of PCR artifacts.

3.9.2. Library Enrichment Confirmation

To verify that a ChIP library has maintained a specific enrichment of target sites, perform qPCR on the ChIP-seq library using both positive targets and negative control primer pairs. The input library serves as a control to normalize the qPCR data to determine the relative enrichment of a given target (see Note 9).

A. Real-time quantitative PCR (qPCR)

Analyze the ChIP-seq sample as well as the appropriately sized input library for reference. Prepare a master reaction mix for each library with triplicate reactions per primer set. Add extra reagents for 10% of the total number of reagents to account for loss of volume. Add 15 μ l of reaction mix to each PCR reaction well. Add 2 μ l of primer mix to each well.

2 μ l	2 ng library from Subheading 3.8.6
3.5 μ l	Nuclease-free H ₂ O
7.5 μ l	2 $\times$ SYBR Green mix (containing polymerase)
2 μ l	5 μ M target primer mix
15 μ l	Total reaction volume

Amplify using the following PCR protocol:

3 min at 95°C.

40 cycles of 30 s at 95°C then 30 s at 60°C.

Include a 70–95°C melting curve at the end of the qPCR program, reading all points or every 0.2°C.

B. Determine enrichment

Analyze library enrichments by qPCR as described in Subheading 3.7. The enrichment values vary depending on histone modification and placement of primers (see Fig. 2b, for an example). If enrichment is acceptable, then the sample can be provided to a sequencing facility (e.g., http://genomecenter.ucdavis.edu/dna_technologies/) for high-throughput sequencing; do not proceed with sequencing unless the positive targets are at least 20-fold enriched.

3.10. Library Sequencing and Data Analysis

After the sequencing is performed, the short tags (~25–50 nts) are mapped to the human genome, the tags that map uniquely to only one location in the genome are selected, the unique tags are extended to the average size of the library fragments (~200 nt), and then the extended fragments are grouped into consecutive bins running the length of each chromosome. The binned data can be visualized using the UCSC browser (<http://www.genome.ucsc.edu/>) or the Affymetrix Integrated Genome Browser (http://www.affymetrix.com/partners_programs/-programs/developer/tools/download_igb.affx; see Fig. 2c). Target sites can be identified using a variety of peak calling methods (26–33). Most peak calling programs account for binding pattern in peak shape, but algorithms need to be adjusted for spreading histone marks covering larger regions of the genome. Sole-search is a peak-calling program that was initially developed to identify transcription factor binding sites, which typically display peaks (34). Applying this type of program to spreading histone marks such as H3K9me3 produces a large number of small peaks rather than identifying a binding region. Sole-search version 2 has been modified to address this issue and offers the choice of using the peak or histone method for peak calling (Blahnik et al., in preparation). We note that the histone method can be used to call peaks for both types of binding; it correctly identifies sharp peaks for H3K4me3, and also calls broad regions occupied by H3K9me3 (Fig. 3).

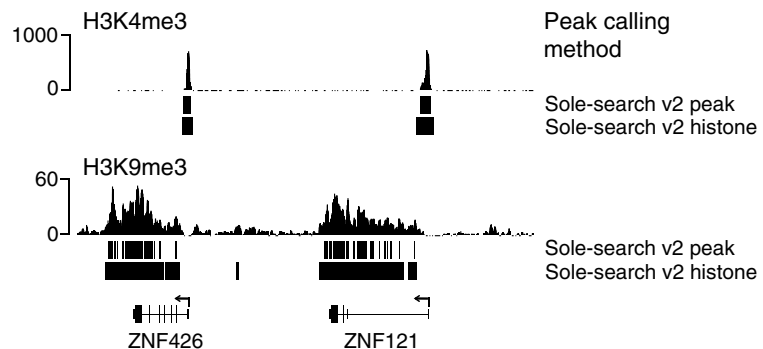


Fig. 3. Comparison of target identification for histone modifications having peak-like vs. spreading binding patterns. Libraries were prepared from ChIP samples using an antibody for H3K4me3 (K562 chromatin) and an antibody for H3K9me3 (Ntera2 chromatin) and were sequenced using the Illumina platform. As expected, sharp H3K4me3 peaks are observed proximal to the transcription start sites, whereas H3K9me3 covers larger chromatin regions; two zinc finger genes on chromosome 19 are shown to illustrate the different binding patterns (the number of tags is plotted along the y-axis and the hg19 chromosomal coordinates are shown on the x-axis). Targets were identified with Sole-search version 2 using the peak method as well as the histone method of the program; targets identified using both methods are depicted underneath each ChIP-seq signal track. The peak method (alpha value 0.01; FDR 0.001) works well for H3K4me3 but not for H3K9me3, whereas the histone method (alpha value 0.0001; FDR 0.01) can identify the sharp peaks in the H3K4me3 dataset as well as the broad binding regions of the H3K9me3 dataset.

The number of reads required to identify all sites bound by a particular histone mark depends on the characteristics of this mark. Similar to most site-specific transcription factors, the binding patterns of certain modified histones (H3K9ac and H3K4me3) can be identified as sharp peaks. For analysis of these modifications throughout the human genome, ten million sequenced tags should be sufficient. With the current Illumina GA2 platform, 20–40 million sequenced reads can be obtained routinely in a single sequencing run and 75% or more of the reads can be mapped to the human genome. Thus, one lane is usually sufficient to identify the regions bound by H3K9ac and H3K4me3. However, reads should ideally come from two independent ChIP samples, with the binding sites identified in each replicate having at least a 60% overlap. Therefore, a minimum of two lanes of sequencing (one lane each from two independent ChIP assays) are usually performed for each of the histone modifications. As noted above, certain histone modifications do not have a peak-like binding pattern, but instead spread over large chromatin regions. Spreading histone marks require more reads to identify significant enrichment over background. For spreading marks such as H3K36me3, H3K4me1, or H3K27me3, 20–40 million reads may be required. H3K9me3 is not only a spreading mark but it is also present on repetitive regions of the genome (such as centromeres). For this mark, up to 50% of the reads may map to more than one place in the human genome and are discarded from analysis. Thus, up to 40–80 million sequenced reads may be required to achieve 20–40 million mapped reads (see Fig. 4). For the Illumina GA2 machine this translates to ~2–3 lanes of sequencing. Fortunately, the newest Illumina technology (HiSeq) will greatly increase the number of reads/lane. However, HiSeq will generally provide more reads/lane than is required for many site-specific factors or certain modified histones. Therefore, multiplexing and barcoding of ChIP-seq libraries will become necessary to ensure the most cost-effective sequencing strategy (35).

4. Notes

1. Use the primer design program Primer3 (36) or another suitable program to design the target and control primers, making the product length 90–150 bp. If possible, design primers for at least two positive targets as well as for two negative control regions. Before testing the ChIP sample or library, it is important to determine that the primers work for the SYBR Green-based real-time PCR assay. To do this, use input DNA and run a melting curve following the real-time PCR reaction conditions and view the dissociation curve to ensure that the desired amplicon was detected, as seen by a single peak.

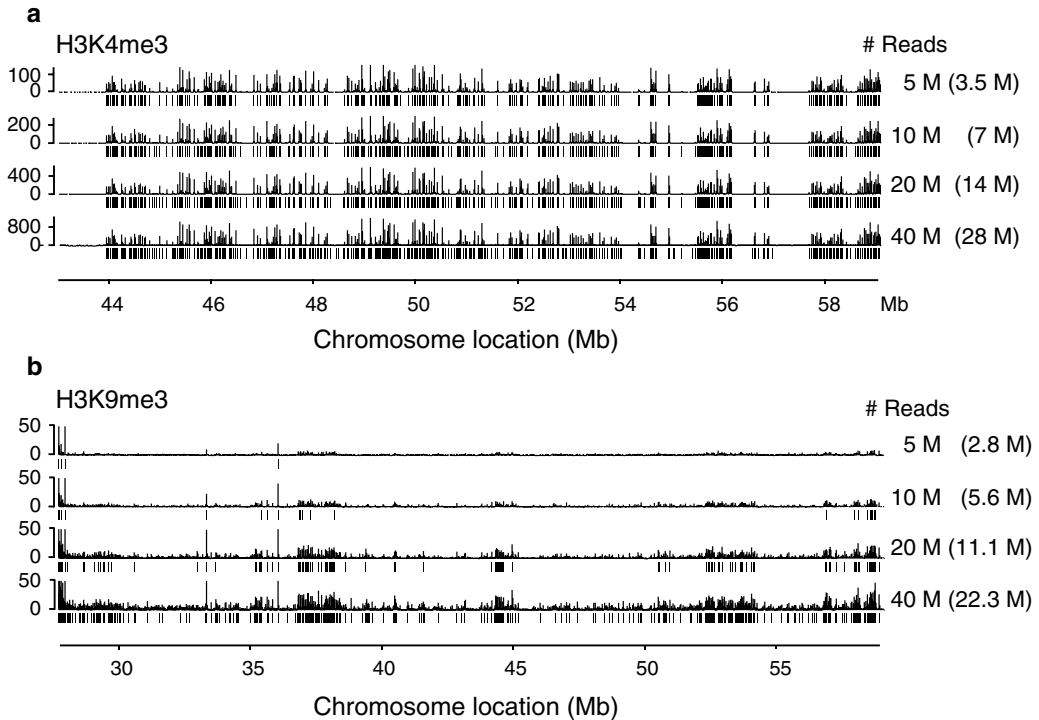


Fig. 4. Required read number varies depending on the histone mark. The ChIP-seq binding patterns from 5, 10, 20, and 40 million (M) sequenced reads for the libraries described in Fig. 3 are shown for a region on chromosome 19 (hg19 coordinates). The number of sequences that uniquely mapped to the genome is given in *brackets* and is representative for a typical ChIP-seq experiment. Peak height is plotted along the *y*-axis and the chromosomal location is shown on the *x*-axis. Called peaks are indicated as *black bars* below each track. While ten million reads are sufficient to identify most H3K4me3 sites in the human genome, more reads are required for the spreading mark H3K9me3. For example, in the region shown, only one peak is called with five million reads and only 20 peaks are identified with ten million reads; appropriate peak calling is achieved when ~40 million reads are analyzed.

2. Chromatin can be sheared by sonication or digested by micrococcal nuclease. Although both methods work well, this protocol is based on sonication. An alternative method using micrococcal nuclease is available from Cell Signaling Technology (<http://www.cellsignal.com>).
3. Sonication conditions should be optimized for each cell type and chromatin size determined before processing large quantities of cells, especially if the cells are collected from patients. We use the BioRuptor UCD-200 (Diagenode) on high setting for sonication. Wear hearing protection! Volumes between 0.5 and 2 ml are sonicated in 15 ml tubes, volumes between 0.1 and 0.3 ml are sonicated in 1.5 ml Eppendorf tubes, and volumes between 10 and 100 μ l are sonicated in 0.5 ml tubes. The pulse duration, intensity, and number will vary depending on the sonicator, the extent of cross-linking, and cell type.

Ideally, the least amount of input energy that gives satisfactory fragmentation should be used. We commonly sonicate 20–30 min (pulses of 30 s at setting high, with 1.5 min pauses in between pulses).

4. It is recommended to prepare chromatin and set up ChIP assays on the same day for best results. If necessary, unused chromatin can be snap frozen in liquid nitrogen and stored at -80°C for use at a later time.
5. In some cases the cell number is extremely limited. As an alternative to determining chromatin yield before the ChIP assay, chromatin can be isolated after incubation with magnetic protein G beads in Subheading 3.4, step 2. Do not discard the supernatant, but reverse crosslinks and purify DNA instead. This sample can substitute for the Input sample.
6. A typical chromatin yield ranges from 5 to 10 μg chromatin per one million tissue culture cells. In our experience, primary cells often give lower chromatin yields than cultured cells. If the yield drops below 1 μg chromatin per one million cells, the ChIP reaction mixture becomes too dilute and the ChIP quality may be significantly compromised.
7. While manual ChIP assays are very reproducible when performed by experienced investigators, the IP-Star[®] (Diagenode) offers an alternative by automating Subheadings 3.4 and 3.5. The ChIP robot can perform up to 16 ChIP assays at a time, requiring only manual set-up of the tubes and solutions. Automation of the ChIP assay can enhance consistency of ChIP results from one experiment to the other (especially for novice researchers) or between different users. We use the IP-Star in combination with the Auto ChIP kit (Diagenode) following the manufacturer's instructions. Typically, an automated ChIP assay starts with 1–5 μg chromatin (see Fig. 2 for PCR analysis of ChIPs performed using 1 μg of chromatin and the IP-Star); if higher chromatin amounts are required, it may be best to perform manual ChIP assays. Chromatin concentration is important since the chromatin sample has to be diluted to a total reaction volume of 100 μL with Buffer A (Diagenode) to ensure appropriate dilution of the SDS that is present in the nuclei lysis buffer. Therefore, the chromatin sample used per ChIP should not exceed 20 μL to allow fivefold dilution with buffer A; more dilute chromatin (higher volumes) may decrease ChIP enrichment and are thus not desirable. Volumes of antibody and Protein G magnetic beads used per assay are the same as described in the manual ChIP protocol. The main difference in the manual and the automated protocol is that in the IP-Star, protein G magnetic beads are coated with antibody before proceeding to 8 h of IP reaction.

8. The biggest limitations to histone ChIP assays are the specificity of the antibody and the variability between different lot numbers of the same antibody. Companies have acknowledged this problem and some are making an effort to provide antibody specificity information. However, the number of ChIP validated antibodies is still very small and it is extremely important to test histone antibodies for cross reactivity with other histone marks. Antibody specificity can be tested using histone dot blots (each dot on the membrane contains a histone peptide carrying a specific modification(s)). The dot blot can be probed with the histone antibody of interest using a Western blotting protocol. Antibodies with high specificity will only recognize one dot corresponding to the histone modification(s) that was used to raise the antibody. Antibody efficiency can vary significantly between different batches, resulting in variation of the quality of the resultant ChIP-seq data. It is therefore important to record antibody details, such as catalog number, lot number, batch of affinity purification, etc. and to test each new antibody batch before performing a ChIP assay.
9. An input library is also critical for determining a baseline genome for the identification of binding sites. It is important that the input library size matches the size of the ChIP-seq library. For each cell type, 20–40 million sequenced tags of an input library are required. The same input library used to determine enrichment can be used for sequencing.
10. Gel size selection is the most variable step among protocols used by different investigators. For example, some protocols incorporate a size selection of the ChIP sample, whereas others size select after adapter ligation or after amplification (some protocols size select at more than one of these steps). We have chosen to perform size selection only once at the step after adapter ligation for three reasons: (1) each size selection results in loss of sample, so we limit the protocol to only one size selection step, (2) selecting after adapter ligation allows the removal of unincorporated adapters from the amplification reaction (which would not be possible if size selection was done after the ChIP assay but before adapter ligation), and (3) by selecting prior to PCR, we only amplify ChIP fragments of the proper length. This is especially important when the sonicated sample contained a significant fraction of fragments bigger than 500 bp (which can sometimes occur with certain cell types).
11. We prepare two libraries for each sample because the efficiency of sonication varies for different chromatin regions. Compacted chromatin sonicates less efficiently than accessible chromatin regions. The smaller portion of sonicated chromatin (200–400 bp) contains more promoter fragments, while fragments from compacted chromatin are overrepresented in the bigger

fragment (400–600 bp). We proceed with both sized fragments through the library-making process and determine enrichment of positive controls by qPCR in both the large and small libraries for each sample. We typically find that the smaller sized library gives the highest enrichment for histone marks of open chromatin (e.g., H3K9Ac, H3K4me3, H3K4me1, and H3K36me3) whereas the larger sized library gives the highest enrichment for compacted chromatin (e.g., H3K9me3 or H3K27me3).

References

1. Crews, D. Epigenetics, brain, behavior, and the environment, *Hormones (Athens)* **9**, 41–50.
2. Grayson, D. R., Chen, Y., Dong, E., Kundakovic, M., and Guidotti, A. (2009) From trans-methylation to cytosine methylation: evolution of the methylation hypothesis of schizophrenia, *Epigenetics* **4**, 144–149.
3. Iwamoto, K., and Kato, T. (2009) Epigenetic profiling in schizophrenia and major mental disorders, *Neuropsychobiology* **60**, 5–11.
4. Jiang, Y., Langley, B., Lubin, F. D., Renthal, W., Wood, M. A., Yasui, D. H., Kumar, A., Nestler, E. J., Akbarian, S., and Beckel-Mitchener, A. C. (2008) Epigenetics in the nervous system, *J Neurosci* **28**, 11753–11759.
5. McCarthy, M. M., Auger, A. P., Bale, T. L., De Vries, G. J., Dunn, G. A., Forger, N. G., Murray, E. K., Nugent, B. M., Schwarz, J. M., and Wilson, M. E. (2009) The epigenetics of sex differences in the brain, *J Neurosci* **29**, 12815–12823.
6. Jones, P. A., and Baylin, S. B. (2007) The epigenomics of cancer, *Cell* **128**, 683–692.
7. Ellis, L., Atadja, P. W., and Johnstone, R. W. (2009) Epigenetics in cancer: targeting chromatin modifications, *Mol Cancer Ther* **8**, 1409–1420.
8. Lund, A. H., and van Lohuizen, M. (2004) Epigenetics and cancer, *Genes Dev* **18**, 2315–2335.
9. Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007) High-resolution profiling of histone methylations in the human genome, *Cell* **129**, 823–837.
10. Kouzarides, T. (2007) Chromatin modifications and their function, *Cell* **128**, 693–705.
11. Heintzman, N. D., Hon, G. C., Hawkins, R. D., Kheradpour, P., Stark, A., Harp, L. F., Ye, Z., Lee, L. K., Stuart, R. K., Ching, C. W., Ching, K. A., Antosiewicz-Bourget, J. E., Liu, H., Zhang, X., Green, R. D., Lobanenkov, V. V., Stewart, R., Thomson, J. A., Crawford, G. E., Kellis, M., and Ren, B. (2009) Histone modifications at human enhancers reflect global cell-type-specific gene expression, *Nature* **459**, 108–112.
12. Heintzman, N. D., Stuart, R. K., Hon, G., Fu, Y., Ching, C. W., Hawkins, R. D., Barrera, L. O., Calcar, S. V., Qu, C., Ching, K. A., Wang, W., Weng, Z., Green, R. D., Crawford, G. E., and Ren, B. (2007) Distinct predictive chromatin signatures of transcriptional promoters and enhancers in the human genome, *Nature Genetics published online Feb 4, 2007*.
13. O'Geen, H., Squazzo, S. L., Iyengar, S., Blahnik, K., Rinn, J. L., Chang, H. Y., Green, R., and Farnham, P. J. (2007) Genome-Wide Analysis of KAP1 Binding Suggests Autoregulation of KRAB-ZNFs, *PLoS Genet* **3**, e89.
14. Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Bruggmann, S. A., Goodnough, L. H., Helms, J. A., Farnham, P. J., Segal, E., and Chang, H. Y. (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs, *Cell* **129**, 1311–1323.
15. Chen, L., and Daley, G. Q. (2008) Molecular basis of pluripotency, *Hum Mol Genet* **17**, R23–27.
16. Hemberger, M., Dean, W., and Reik, W. (2009) Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal, *Nat Rev Mol Cell Biol* **10**, 526–537.
17. Lanzuolo, C., and Orlando, V. (2007) The function of the epigenome in cell reprogramming, *Cell Mol Life Sci* **64**, 1043–1062.
18. Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R. P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., Nusbaum, C., Lander, E. S., and Bernstein, B. E. (2007) Genome-wide maps of

- chromatin state in pluripotent and lineage-committed cells, *Nature* **448**, 553–560.
19. Park, P. J. (2009) ChIP-seq: advantages and challenges of a maturing technology, *Nature Rev Genet*, 669–680.
 20. Wei, C. L., Wu, Q., Vega, V. B., Chiu, K. P., Ng, P., Zhang, T., Shahab, A., Yong, H. C., Fu, Y., Weng, Z., Liu, J., Zhao, X. D., Chew, J. L., Lee, Y. L., Kuznetsov, V. A., Sung, W. K., Miller, L. D., Lim, B., Liu, E. T., Yu, Q., Ng, H. H., and Ruan, Y. (2006) A global map of p53 transcription-factor binding sites in the human genome, *Cell* **124**, 207–219.
 21. Bhinge, A. A., Kim, J., Euskirchen, G. M., Snyder, M., and Iyer, V. R. (2007) Mapping the chromosomal targets of STAT1 by Sequence Tag Analysis of Genomic Enrichment (STAGE), *Genome Res* **17**, 910–916.
 22. Hawkins, R. D., Hon, G. C., Lee, L. K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L. E., Kuan, S., Luu, Y., Klugman, S., Antosiewicz-Bourget, J., Ye, Z., Espinoza, C., Agarwahl, S., Shen, L., Ruotti, V., Wang, W., Stewart, R., Thomson, J. A., Ecker, J. R., and Ren, B. (2010) Distinct epigenomic landscapes of pluripotent and lineage-committed human cells, *Cell Stem Cell* **6**, 479–491.
 23. Motallebipour, M., Ameur, A., Reddy Bysani, M. S., Patra, K., Wallerman, O., Mangion, J., Barker, M. A., McKernan, K. J., Komorowski, J., and Wadelius, C. (2009) Differential binding and co-binding pattern of FOXA1 and FOXA3 and their relation to H3K4me3 in HepG2 cells revealed by ChIP-seq, *Genome Biol* **10**, R129.
 24. Koerber, R. T., Rhee, H. S., Jiang, C., and Pugh, B. F. (2009) Interaction of transcriptional regulators with specific nucleosomes across the *Saccharomyces* genome, *Mol Cell* **35**, 889–902.
 25. Goren, A., Ozsolak, F., Shores, N., Ku, M., Adli, M., Hart, C., Gymrek, M., Zuk, O., Regev, A., Milos, P. M., and Bernstein, B. E. (2010) Chromatin profiling by directly sequencing small quantities of immunoprecipitated DNA, *Nat Methods* **7**, 47–49.
 26. Fejes, A. P., Robertson, G., Bilenky, M., Varhol, R., Bainbridge, M., and Jones, S. J. M. (2008) FindPeaks 3.1: a tool for identifying areas of enrichment from massively parallel short-read sequencing technology, *Bioinformatics* **24**, 1729–1730.
 27. Xu, H., Wei, C.-L., Lin, F., and Sung, W.-K. (2008) An HMM approach to genome-wide identification of differential histone modification sites from ChIP-seq data, *Bioinformatics* **24**, 2344–2349.
 28. Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nussbaum, C., Myers, R. M., Brown, M., Li, W., and Liu, X. S. (2008) Model-based analysis of ChIP-Seq (MACS), *Genome Biology* **9**, R137.
 29. Johnson, D. S., Mortazavi, A., Myers, R. M., and Wold, B. (2007) Genome-wide mapping of in vivo protein-DNA interactions., *Science* **316**, 1497–1502.
 30. Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., Delaney, A., Thiessen, N., Griffith, O. L., He, A., Marra, M., Snyder, M., and Jones, S. (2007) Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing, *Nat Methods* **4**, 1–7.
 31. Jothi, R., Cuddapah, S., Barski, A., Cui, K., and Zhao, K. (2008) Genome-wide identification of in vivo protein-DNA binding sites from ChIP-seq data, *Nucleic Acids Res* **36**, 5221–5231.
 32. Rozowsky, J., Euskirchen, G., Auerbach, R. K., Zhang, Z. D., Gibson, T., Bjornson, R., Carriero, N., Snyder, M., and Gerstein, M. B. (2009) PeakSeq enables systematic scoring of ChIP-seq experiments relative to controls, *Nat Biotechnol* **27**, 66–75.
 33. Valouev, A., Johnson, D. S., Sundquist, A., Medina, C., Anton, E., Batzoglou, S., Myers, R. M., and Sidow, A. (2008) Genome-wide analysis of transcription factor binding sites based on ChIP-seq data, *Nature Methods* **5**, 829–834.
 34. Blahnik, K. R., Dou, L., O'Geen, H., McPhillips, T., Xu, X., Cao, A. R., Iyengar, S., Nicolet, C. M., Ludaescher, B., Korf, I., and Farnham, P. J. (2010) Sole-search: An integrated analysis program for peak detection and functional annotation using ChIP-seq data, *Nucleic Acids Res* **38**, e13.
 35. Lefrançois, P., Euskirchen, G. M., Auerbach, R. K., Rozowsky, J., Gibson, T., Yellman, C. M., Gerstein, M., and Snyder, M. (2009) Efficient yeast ChIP-Seq using multiplex short-read DNA sequencing, *BMC Genomics* **10**, 37.
 36. Rozen, S., and Skaletsky, H. J. (2000) Primer3 on the WWW for general users and for biologist programmers, in *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (Krawetz, S., and Misener, S., Eds.), pp 365–386, Humana Press, Totowa, NJ.