Genome-scale ChIP-chip analysis using 10,000 human cells

Luis G. Acevedo1, A. Leonardo Iniguez3, Heather L. Holster3, Xinmin Zhang3, Roland Green3, and Peggy J. Farnham1

1Department of Pharmacology and the Genome Center, University of California-Davis, Davis, CA, 2Cell and Molecular Biology Program, University of Wisconsin-Madison, and 3NimbleGen Systems Inc., Madison, WI, USA

BioTechniques 43:791-797 (December 2007)
doi 10.2144/000112625

The technique of chromatin immunoprecipitation (ChIP) is a powerful method for identifying in vivo DNA binding sites of transcription factors and for studying chromatin modifications. Unfortunately, the large number of cells needed for the standard ChIP protocol has hindered the analysis of many biologically interesting cell populations that are difficult to obtain in large numbers. New ChIP methods involving the use of carrier chromatin have been developed that allow the one-gene-at-a-time analysis of very small numbers of cells. However, such methods are not useful if the resultant sample will be applied to genomic microarrays or used in ChIP-sequencing assays. Therefore, we have miniaturized the ChIP protocol such that as few as 10,000 cells (without the addition of carrier reagents) can be used to obtain enough sample material to analyze the entire human genome. We demonstrate the reproducibility of this MicroChIP technique using 2.1 million feature high-density oligonucleotide arrays and antibodies to RNA polymerase II and to histone H3 trimethylated on lysine 27 or lysine 9.

INTRODUCTION

We and others (1–5) have developed protocols for studying DNA-protein interactions and histone modifications in living cells or tissues (see also www.epigenome-noe.net/research-tools/protocols.php). In particular, the method of chromatin immunoprecipitation (ChIP) has provided many new insights into gene regulation (6,7). Briefly, ChIP involves crosslinking the transcription factors to their DNA binding sites by treatment of cells with formaldehyde and then preparation of chromatin by sonication of the treated cells. An immunoprecipitation is then performed using the crosslinked chromatin, resulting in the collection of all the binding sites in the genome for the factor of interest. This sample can then be analyzed by PCR to study a particular gene(s) or applied to microarrays for analysis of many genes (8–12). The majority of ChIP studies have used large numbers of cells grown in culture. However, even using the standard amount of 10⁷ cells, there is not enough precipitated sample for hybridization to a single microarray. Investigators have solved this problem by pooling ChIP samples (13,14). However, the problem is exacerbated when considering the number of arrays needed to analyze the entire human genome; pooling samples is simply not practical under these conditions. Therefore, most ChIP-chip protocols incorporate an amplification step that converts the small amount of material present in a ChIP sample into the microgram quantities required for hybridization to an array. Using such amplification techniques, investigators can routinely obtain enough sample material from 10⁶ to 10⁷ cells for analysis of the human genome.

We are interested in applying the ChIP technology to study very small numbers of human cells, such as those obtained from tumor biopsies, fractionation of mixed populations by cell sorting, or differentiation of embryonic stem cells. Although one study has been published in which 1000 cells were used for a ChIP assay, the protocol required mixing a small number of mammalian cells with large numbers of Drosophila cells (15).

The success of this protocol was demonstrated by analysis of a few promoters using PCR reactions. Unfortunately, the use of carrier chromatin is not appropriate if the ChIP samples are to be analyzed by sequencing (the carrier DNA would constitute most of the sequenced tags) or by hybridization (due to cross-reaction of the carrier DNA to the oligonucleotides on the array). Therefore, we have taken the approach of altering the ChIP-chip assay such that only 10,000 cells (without carrier DNA) are now required. Importantly, enough sample material can be obtained from 10,000 cells to analyze the entire human genome using microarray technologies.

MATERIALS AND METHODS

Cells, Tissues, Antibodies, and PCR

Cells used for the miniaturization optimization are HuH7 human hepatocellular carcinoma cells. Cells were grown at 37°C in a humidified 5% CO₂ incubator in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal bovine serum and 100 U/mL of penicillin and streptomycin. The cells were crosslinked and chromatin prepared as described in the MicroChIP Protocol (see Supplementary Material available online at www.BioTechniques.com). For the experiments in this study, the chromatin ranged from 0.5 to 3 kb in size, with a median size of ~1 kb. Antibodies to RNA polymerase II (8WG16; Covance, Denver, PA, USA), H3me3K27 (07–449; Upstate, Medford, PA, USA), H3me3K27 (07–449; Upstate, Medford, PA, USA), and H3me3K9 (8898; Abcam, Cambridge, MA, USA) were used in the ChIP assays. Primers used to confirm the binding of factors to the promoter region of the described genes have the following sequence: RNAPII: 5′-agatgaaacgccaggtctcaac-3′, 5′-aggttacgactttgtctcte-3′; SOAT: 5′-ccctacttcagggacacca-3′, 5′-ccaaagaaacccaggaacca-3′; EVX1: 5′-ccgggtcctcacttctaa-3′, 5′-aaagccaaacggcagcatat-3′, and ZNF44: 5′-ggtttccacaaactc-3′, 5′-aacactccagccagataga-3′, ZNF333: 5′-acagggagaaggccctagc-3′, 5′-tggcgcactataccttg-3′. PCR conditions were 95°C for 3 min, 30 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 45 s, and then 72°C for 5 min.
Short Technical Reports

MicroChIP Protocol
(For 10⁴–10⁵ Cells)

This entire protocol can be performed in ~4 days, assuming that the Staph A cells are prepared in advance (indicated as Day 0). A brief summary of the protocol is provided here; further details, including a step-by-step protocol, catalog numbers for reagents, and composition of each solution is provided in the MicroChIP Protocol (see Supplementary Material).

Day 0: Staph A cells, which are used to precipitate the antibody-chromatin complexes, are prepared in large batches, aliquoted, and stored in liquid nitrogen prior to the pre-treatment (blocking) required for the ChIP assays that are performed on Day 2. Instructions for preparation and storage of the Staph A cells are provided in the MicroChIP Protocol (see Supplementary Material).

Day 1: The cells are crosslinked and the chromatin is sonicated and prepared for use in ChIP assays. The crosslinked cells can be used immediately or stored at -80°C prior to use in the chromatin preparation step. The chromatin can also be prepared and used immediately or stored at -80°C prior to use in the ChIP assay. A portion of the chromatin should be purified by QIAquick column (Qiagen, Valencia, CA, USA) and quantified by NanoDrop (NanoDrop Technologies, Wilmington, DE, USA). Also, Staph A cells, which will be used to precipitate the antibody-chromatin complexes, are treated overnight with nonspecific protein (BSA) and DNA (herring sperm) to reduce background in the ChIP assay.

Day 2: Antibodies are added to the chromatin and allowed to form complexes.

Day 3: The antibody-chromatin complexes are purified, the “Input” DNA sample is prepared, and the size of the chromatin is determined. Although one could determine the size of the sonicated chromatin prior to the ChIP assays, this would require using the equivalent amount of chromatin as needed for several ChIP assays. Because this MicroChIP Protocol has been developed for the analysis of limiting amounts of starting material, we recommend that the size of the chromatin be determined using the supernatants from the ChIP samples.

Day 4: To obtain enough DNA to analyze by PCR or to hybridize to an array, the ChIP sample must be amplified. A single round of amplification provides enough sample from a MicroChIP assay for many PCR reactions or to probe one or two arrays. To probe a set of arrays that represent the entire genome, a second round of amplification must be performed. PCR reactions are used to confirm positive and negative controls for each set of amplifications.

ChIP-chip Assays

Microarray hybridization was performed by the NimbleGen Service Facility in Iceland or at NimbleGen Systems in Madison, WI, USA, using standard protocols for ChIP-chip arrays, and data analysis was performed using the Maxfour algorithms developed by the Farnham laboratory (9). Briefly, the Maxfour method selects the highest consecutive probes per promoter (4 in a row were used for the RNAPII arrays and 10 in a row were used for the H3me3K27 arrays) and averages them to assign a value to each promoter for ranking analysis. The choice as to the number of consecutive probes to consider for peak calling varies depending on the average peak width of the binding sites for a particular factor. The peaks

Table 1. Quantitation of ChIP Samples

<table>
<thead>
<tr>
<th>Number of cells</th>
<th>Chromatin input</th>
<th>ChIP output</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 × 10⁷</td>
<td>200 μg</td>
<td>150 ng⁴</td>
</tr>
<tr>
<td>1 × 10⁶</td>
<td>20 μg</td>
<td>10 ng⁴</td>
</tr>
<tr>
<td>5 × 10⁵</td>
<td>10 μg</td>
<td>1.3 ng⁴</td>
</tr>
<tr>
<td>1 × 10⁵</td>
<td>2 μg</td>
<td>300 pg⁵</td>
</tr>
<tr>
<td>1 × 10⁴</td>
<td>200 ng</td>
<td>30 pg⁵</td>
</tr>
</tbody>
</table>

⁴Output measured using picogreen analysis of a RNAPII ChIP sample.
⁵Output estimated based on larger scale ChIP samples.
for RNAPII binding sites are narrow and thus we analyze 4 probes in a row (∼400 bp); the H3me3K27 signals are very broad and thus we analyze 10 probes in a row (∼1 kb). Information concerning all of the arrays used in the ChIP-chip experiments is provided in Supplementary Table S1.

RESULTS AND DISCUSSION

There is widespread interest in the molecular characterization of differences between tumor and normal tissue. Of particular concern is the identification of transcriptionally active genomic regions that have been converted to transcriptionally silent domains (or vice versa) during neoplastic transformation. These genomic regions can be identified using the ChIP assay and antibodies to RNA polymerase II (to identify active regions) or to histone H3 trimethylated on lysine 27 or lysine 9 (to identify silenced regions). To adapt the ChIP assay such that it is amenable to the analysis of active and silenced genomic regions using the small amount of tissue obtained from tumor biopsies, we have miniaturized the ChIP assay, developed a modified amplification process that allows representative amplification of picograms of ChIP sample, and demonstrated the robustness of this MicroChip Protocol by applying the amplicons to high-density oligonucleotide arrays containing 2.1 million features (Figure 1).

Modification of the Amplicon Preparation Method

Usually the success of a ChIP assay is monitored by PCR analysis of the sample, using primers specific for a genomic region known to be bound by a particular transcription factor or modified histone. Generally 2% or 20% of the enriched material is analyzed in the PCR confirmation step when using 10^7 or 10^6 cells, respectively, in a ChIP assay. However,
when 10,000–100,000 cells are used,
the amount of precipitated DNA is
quite small (Table 1) and therefore
the success of a “miniaturized” ChIP
assay is most easily monitored using
amplified samples. Therefore, to adapt
our ChIP protocol for small numbers
of cells, we first had to develop an
amplification method appropriate for
very small amounts of precipitated
DNA. We have recently shown that a
whole genome amplification method
(WGA) (16) gives very reliable results using
the Whole Genome Amplification 2 kit
(WGA2) (Sigma-Aldrich, St. Louis,
MO, USA). Unfortunately, the reagents
in the WGA2 kit are optimized to
give representative amplification of
samples containing 10–100 ng of
DNA. We estimate that 10,000 cells
corresponding to ~200 ng of input
DNA) would provide only 30 pg of an
RNAPII ChIP sample. (Although the
amount of precipitated DNA will vary
depending on the antibody used and
the efficiency of washing the immuno
precipitate, the amount of ChIP sample
from 10,000 cells will always be <10
ng.) Therefore, the WGA2 kit is not
appropriate for amplification of ChIP
samples from such small numbers of
cells. Fortunately, a similar amplifi
cation method is available (WGA4)
that allows representative amplification
beginning with 10–100 pg of DNA.
Before using this method to amplify
ChIP samples from 10,000 cells, we
first optimized the WGA4 amplifi
cation protocol using chromatin from
a standard ChIP assay that was diluted
to the MicroChIP scale. We prepared
amplimers from 10 ng (using WGA2)
and 50 pg (using WGA4; Sigma-
Aldrich) of the same ChIP sample and
tested the amplimers in a PCR assay.
As shown in Figure 2A, signals from
RNAPII and H3me3K27 samples
were similar using 10 ng (left panel)
or 50 pg (middle panel) to prepare
the amplimers. Therefore, we were
confident that we could use the WGA4
amplification method to analyze ChIP
samples from small numbers of cells.

**Miniaturization of the ChIP Assay**

Reduced sample recovery was a
major problem that we encountered
when using small numbers of cells in
the ChIP assay. Therefore, important
modifications were incorporated into
the ChIP protocol: (i) small volume,
low retention siliconized tubes were
used at each step; (ii) a bioruptor
sonicator was used that allowed a
low volume during sonication and no
probe contact with the sample; and
(iii) a freeze-thaw step was included
at the nuclear lysis step. We also found
that it is important to quantitate the
chromatin using a NanoDrop, to ensure
that ~200 ng is used for a MicroChIP
experiment. Finally, the procedure used
to determine the size of the sonicated
DNA in the ChIP assay (which is
often very useful as a quality control
measure) had to be modified. With the
standard assay, investigators usually
reserve 10% of the amount used for
a ChIP assay, reverse the crosslinks,
isolate the DNA, and process this as
Input DNA, which can be used to
check the size of the chromatin and as
a control for PCR assays. However, it
is not possible to visualize the DNA on
a gel if 10% of a MicroChIP is used.
Therefore, we collected all the super
naturants from the ChIP precipitation
assays (see Supplementary Material:
MicroChIP Protocol, Day 3, Step 4),
pooled them, reversed the crosslinks,
concentrated the sample, cleaned up
the DNA, and ran the DNA on a gel to
check the size of the chromatin.

![Figure 3. ChIP-chip assays using 10,000 cells and 2.1 million feature arrays.](image-url)
Using the modifications described above and in the MicroChIP Protocol (see Supplementary Material), we performed ChIP-chip assays with 10⁶ cells using antibodies to RNAPII and H3me3K27, to recognize active and inactive chromatin states, respectively. We prepared amplicons from the MicroChIPs using WGA4. We used primers for the promoter of the largest subunit of RNAPII as a positive control for the RNAPII ChIP and primers for the promoter of SOAT as a positive control for the H3me3K27 ChIP. Conversely, the promoter of the largest subunit of RNAPII served as a negative control for the H3me3K27 ChIP, and the promoter of SOAT served as the negative control for the RNAPII ChIP. As can be seen in Figure 2A, right panel, signals from RNAPII and H3me3K27 samples were similar for the MicroChIP/WGA4 samples as for the standard ChIP/WGA2 or standard ChIP/WGA4 (diluted) samples. These controls suggest that our miniaturization of the ChIP protocol was successful. We applied the MicroChIP/WGA4 amplicons to arrays representing 5 kb tiled regions for all known human promoters (~25,000 genes; NimbleGen Systems); hybridization signals from each promoter were ranked using Maxfour (9). As shown in Figure 2B, amplicons from two independent H3me3K27 MicroChIP experiments showed very reproducible results on promoter arrays. In contrast, a comparison of H3me3K27 targets to RNAII targets showed very little overlap (Figure 2C), as expected when comparing active to inactive marks of chromatin.

We then performed two independent standard ChIP/WGA2 assays using antibodies to RNAPII and H3me3K27 and hybridized these amplicons to the promoter arrays. Again, the promoters were ranked using the Maxfour program and then the top 10%, 20%, and 30% of the promoters were identified. As can be seen in Figure 2D, good reproducibility was obtained between the top targets bound by RNAPII when 10⁷ cells were used in two independent ChIP-chip experiments (STD ChIP vs. STD ChIP) and when 10⁵ cells were used in two independent MicroChIP experiments (MicroChIP vs. MicroChIP). In addition, the comparison between standard and MicroChIP ranked lists showed that the same promoters were identified as top-ranked targets using both methods. Similar comparisons are shown for H3me3K27 in Supplementary Figure S1. For all comparisons of the RNAPII and H3me3K27 ChIP-chip data, the overlaps are much greater than expected by random change (Figure 2D and Supplementary Figure S1). Examples of the raw data from the standard vs. MicroChIP arrays are shown in Figure 2E.

**ChIP-chip Analysis of 10,000 Cells**

Having confidence that the miniaturization protocol and new amplification method gave reliable data, we then further reduced the number of cells to 10,000 per sample (corresponding to 200 ng of sample) and performed ChIP assays using antibodies to H3me3K9, RNAPII, and H3me3K27. Our ultimate goal was to identify transcription factor binding sites and chromatin modifications throughout the entire human genome, using small populations of cells. Using the entire ChIP sample in a WGA4 amplification protocol, we could obtain ~5 to 7 μg of amplicons from 10,000 cells. However, to probe a 10-array set (representative of the entire human genome) we needed ~40 μg of DNA. Therefore, it was clear that we would need to perform a second amplification step, known as a round 2 amplicon. For this, a second round of amplification was performed using the WGA3 kit (which is similar to the WGA2 method, but is used for second round amplification) and 10 ng of the amplified material from the WGA4 round 1. The round 2 amplicons were confirmed in PCR assays using ZNF44 and ZNF333 primers as positive controls for the H3me3K9 samples, RNAPII primers as a positive control for the RNAPII sample, and SOAT primers as a positive control for the H3me3K27 sample (Figure 3A). The amplicons were all clearly highly enriched for the appropriate targets, but not for the negative controls. Also, none of the promoters were enriched by the IgG sample.

We then probed a high-density genomic array that contains 2.1 million 50-mer probes that represent part of chromosome 16, all of chromosomes 17, 18, and 19, and part of chromosome 20 (NimbleGen HG18_Tile_09; HD2 TILE 25, 26, and 27). These experiments were performed in triplicate, using three independent samples of H3me3K9 amplicons, each prepared from 10,000 cells. As shown in Figure 3B and Supplementary Figure S2, the reproducibility of the MicroChIP-chip protocol is very high when the probe enrichments from any two of the ChIP-chip experiments are compared (r = 0.9). Also, the individual probes show very similar patterns when an entire chromosome is viewed (Figure 3C) and when individual peaks are examined (Figure 3D and Supplementary Figure S3).

**Conclusion**

ChIP is now a fairly common technique for studying chromatin structure or transcription factor binding patterns. However, researchers typically require large numbers of cells to perform this assay. We have modified the physical aspects of the ChIP assay, incorporated an amplification method that allows linear representation of 10–100 pg of DNA, and used 2.1 million feature arrays to demonstrate that transcription factor binding and histone modifications can be studied on a genome-wide scale using only 10,000 human cells. Several previous studies (15,17,18) have reported modifications to the ChIP assay that allow analysis of a small number of cells. However, it is unclear if the samples produced from the previous protocols could be reproducibly studied using microarrays (especially those samples obtained by the addition of large amounts of cells from other organisms). In contrast, we have demonstrated that our approach can reproducibly identify genes bound by transcription factors (RNAPII) or modified histones using ChIP-chip assays. We have also used this protocol to identify binding sites for KAP1, a co-repressor for site-specific zinc finger-containing transcription factors.
(data not shown). We anticipate that the MicroChIP protocol reported here will allow the study of cell populations that are available in limited quantities, such as human tumor biopsies, archival frozen or fixed tissues, or stem cell populations. However, our ultimate goal is to couple the ChiP assay with techniques, such as laser capture microscopy, that involve even smaller numbers of cells. Further modifications to allow the analysis of 1000 cells are in progress.

ACKNOWLEDGMENTS

This work was supported in part by Public Health Service grants CA45250, HG003129, R21CA128471, and DK067889. We thank the members of the Farnham laboratory for helpful discussion and data analysis. L.G.A. and P.J.F conceived and designed the experiments; L.G.A. performed the ChiP experiments, prepared the amplicons, and analyzed the data; A.L.I., H.L.H., and X.Z. were involved in the array hybridizations; R.G. contributed reagents, materials, and analysis tools; and L.G.A. and P.J.F. wrote the manuscript.

COMPETING INTERESTS STATEMENT

Several of the authors (A.L.I., H.L.H., X.Z., and R.G.) are employees of NimbleGen Systems Inc., and we have used NimbleGen arrays in our studies.

REFERENCES