

# Fast genomic $\mu$ ChIP-chip from 1,000 cells

John Arne Dahl, Andrew H Reiner and Philippe Collas

Address: Institute of Basic Medical Sciences, Department of Biochemistry, Faculty of Medicine, University of Oslo, 0317 Oslo, Norway.

Correspondence: Philippe Collas. Email: philippe.collas@medisin.uio.no

Published: 10 February 2009

*Genome Biology* 2009, **10**:R13 (doi:10.1186/gb-2009-10-2-r13)The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2009/10/2/R13>

Received: 14 November 2008

Revised: 16 January 2009

Accepted: 10 February 2009

© 2009 Dahl *et al.*; licensee BioMed Central Ltd.This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Abstract

Genome-wide location analysis of histone modifications and transcription factor binding relies on chromatin immunoprecipitation (ChIP) assays. These assays are, however, time-consuming and require large numbers of cells, hindering their application to the analysis of many interesting cell types. We report here a fast microChIP ( $\mu$ ChIP) assay for 1,000 cells in combination with microarrays to produce genome-scale surveys of histone modifications.  $\mu$ ChIP-chip reliably reproduces data obtained by large-scale assays: H3K9ac and H3K9m3 enrichment profiles are conserved and nucleosome-free regions are revealed.

## Background

Chromatin immunoprecipitation (ChIP) has been widely used to analyze the location of post-translationally modified histones or transcription factors in a genome *in vivo* [1-4]. ChIP analysis of DNA-protein interactions has led to significant advances in the understanding of gene regulation and of how epigenetic phenomena are regulated to affect gene expression, DNA repair and replication [5,6]. In a typical ChIP assay, large numbers of cells are used, DNA and proteins are cross-linked and chromatin is sheared to fragments of approximately 400-500 bp. Antibodies to the protein of interest are coupled to beads and used to pull down protein-DNA complexes. Chromatin is eluted from the complexes, cross-links are reversed and ChIP DNA is purified. A limited number of genomic sequences associated with the precipitated protein can be identified by PCR. Alternatively, high-throughput sequencing or hybridization to DNA microarrays (ChIP-chip) enables genome-scale mapping [7].

The range of biological applications of ChIP assays has been limited by the requirement for large cell numbers (approximately  $10^7$  cells per immunoprecipitation) and the length of the procedure (typically 3-5 days). To remedy to these limita-

tions, a few ChIP-PCR strategies have recently been reported. A 'carrier ChIP' protocol [8] entails immunoprecipitation of chromatin from 100-1,000 mouse cells by mixing with millions of *Drosophila* cells; however, the assay takes several days and is unsuitable for genome-wide analysis due to excess of *Drosophila* carrier DNA that would interfere with such analysis. A one-day 'fast ChIP' assay [9] simplifies the procedure but has only been demonstrated for large cell samples and PCR assessment of relatively few loci. We have reported a downscaled Q<sup>2</sup>ChIP assay [10] for analysis of multiple proteins in 100,000 cells and, subsequently, a microChIP ( $\mu$ ChIP) protocol [11,12] for as few as 100 cells. Again, however, only few loci could be examined with these procedures. Concomitantly, another microChIP assay was reported for 10,000-100,000 cells, which allows genome-wide analysis by ChIP-chip [13]. This assay represents an advancement in ChIP applications, but it remains labor intensive, takes over 4 days and has been validated for 100,000 cells and the top 30% of enriched promoters only. Except for this single attempt to downscale the genome-wide approach, ChIP-chip typically starts out with  $10^7$ - $10^8$  cells and yields amplified DNA ready for labeling and hybridization after 4-5 days. Our ultimate goal is to enable genome-scale investigation of his-

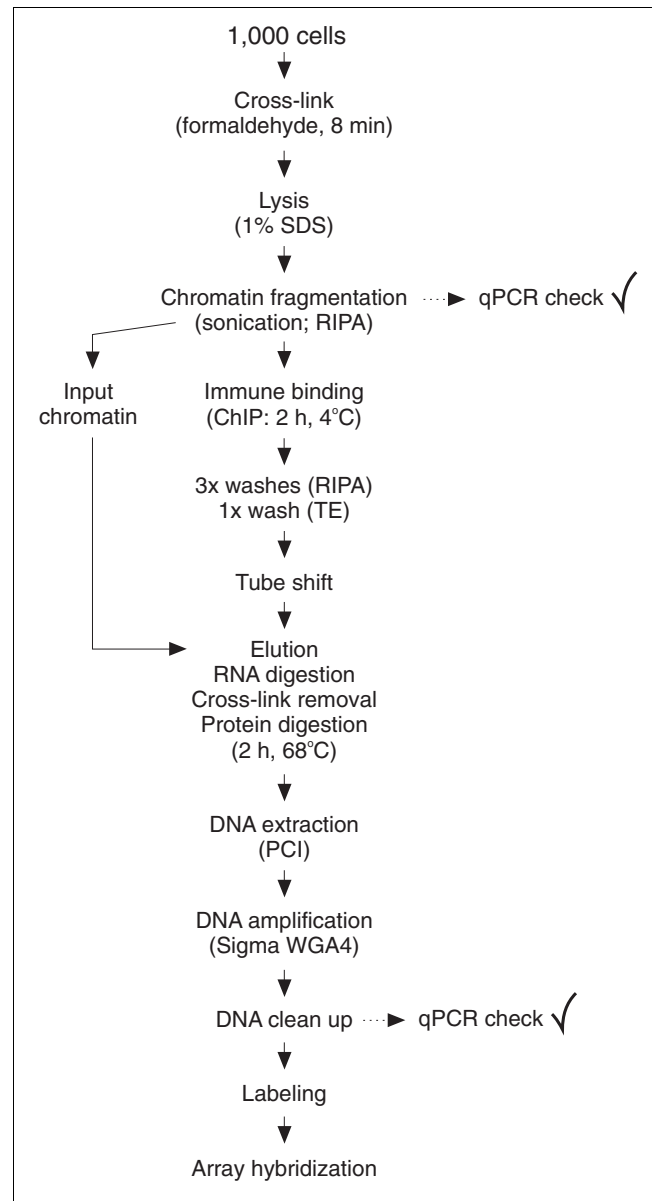
tone modifications in very small cell samples, such as sorted stem cell populations, human tumor biopsies and embryonic cells. Therefore, we wished to move beyond these limitations, reduce the cell numbers and time required, and enhance the robustness of the assay.

We report here the optimization and validation of a one-day  $\mu$ ChIP assay that enables genome-wide surveys of epigenetic histone modifications from 1,000 cells using microarrays. Typically, reliable resolution of ChIP location analysis is ensured through gel electrophoresis assessment of a sample of the fragmented chromatin to determine average DNA fragment length. However, this would require many more cells than what we used in this study. Thus, we devised a PCR-based approach and formulated an equation to allow an estimation of chromatin fragmentation in small cell samples, a step critical for reliable resolution of  $\mu$ ChIP-chip analysis. Using  $\mu$ ChIP-chip, we investigated the enrichment, on promoter regions, of acetylated lysine 9 and trimethylated lysine 9 of histone H3 (H3K9ac and H3K9m3, respectively) associated with transcriptionally active and silent promoters [14]. Four distinct classes of genes were identified based on differential marking by these modifications.  $\mu$ ChIP-chip also demonstrates a nucleosome-free region immediately upstream of the transcription start site (TSS) for active genes, and shows that silenced genes exhibit a more closed chromatin conformation. Furthermore, construction of a metagene and correlation analysis reveal mutually exclusive occupancy profiles for H3K9ac and H3K9m3.

## Results and discussion

### Optimization and validation of $\mu$ ChIP-chip

We established and validated  $\mu$ ChIP-chip (Figure 1) by monitoring promoter association of H3K9ac and H3K9m3. These modifications were immunoprecipitated from pluripotent human embryonal carcinoma NCCIT cells by large scale Q<sup>2</sup>ChIP or by  $\mu$ ChIP (1,000 cells). Notably, we set up a quantitative PCR (qPCR) approach and formulated an equation that enables assessment of chromatin fragmentation in small cell samples within a range of DNA fragment sizes suitable for ChIP. This was critical to ensure proper resolution of  $\mu$ ChIP-chip analysis and, to our knowledge, is the only strategy to overcome this task with small cell numbers. Plotting average DNA fragment length against qPCR signal intensities from large scale sonicated samples reveals a linear relationship within the examined range (300-600 bp) of fragmentation (Additional data file 1). This is a useful average fragment size window for most ChIP applications. The linear relationship is described through the equation ( $y = 0.0012x - 0.0059$ ), where  $y$  is the relative PCR signal intensity of the sample and  $x$  is the average DNA fragment length (Additional data file 1). Starting with 1,000 cells, sonication regimes of  $3 \times 30$  s resulted in a relative PCR signal intensity of 0.502. The equation estimates an average DNA fragment length of approximately 420 bp; hence, this condition is suited for  $\mu$ ChIP-chip analysis.



**Figure 1**  
Flow chart of the  $\mu$ ChIP-chip procedure.

Average fragment size was validated by agarose gel electrophoresis (Additional data file 1).

As part of the optimization of  $\mu$ ChIP-chip, we carried out a comparison of different ChIP DNA isolation procedures. The phenol-chloroform isoamylalcohol DNA extraction method used in this study proved superior to the MinElute (QIAGEN, Valencia, CA, USA; catalogue number 28004) and NucleoSpin Extract II (Machery-Nagel, Bethlehem, PA, USA; catalogue number 740609.10) DNA purification columns in that it recovered two to three times more DNA than the commercial kits, as determined by qPCR (data not shown). Hence, it enables a two- to three-fold further reduction in cell numbers

for  $\mu$ ChIP-chip relative to the commercial kits. Furthermore, an RNase digestion step was tailored to the fast downscaled ChIP procedure to remove RNA that would otherwise interfere with downstream amplification and array hybridization. Subsequently, DNA amounts in ChIP and input samples were measured with a Qubit fluorometer (Additional data file 2) to aid in determining the ChIP-DNA amplification conditions. DNA amount recovered from Q<sup>2</sup>ChIP inputs averaged 1,080 ng whereas  $\mu$ ChIP inputs averaged 6.9 ng. These measurements were in line with an estimated DNA content of 6.6 pg per cell [15,16]. Q<sup>2</sup>ChIP recovered 3.2% and 4.3% of input DNA with antibodies to H3K9ac and H3K9m3, respectively (Additional data file 2). Higher recovery was observed with antibodies against H3K9m2 and H3K4m3 (data not shown), arguing that  $\mu$ ChIP-chip is likely to also be effective with antibodies to other modified histones, which precipitate at least as well as those used here. DNA amounts in  $\mu$ ChIP samples were estimated from Q<sup>2</sup>ChIP DNA recoveries and  $\mu$ ChIP input to get a hint of the amount used for whole genome amplification (WGA). As determined by the assessed DNA amounts, Q<sup>2</sup>ChIP samples and inputs were amplified with the WGA2 kit (Sigma-Aldrich, St. Louis, MO, USA) whereas the WGA4 kit, optimized for very little DNA, was used for  $\mu$ ChIP DNA amplification.

To assess the validity of small cell number ChIP-chip and the reproducibility of this assay, correlation analysis of  $\log_2$  ChIP/input ratios between Q<sup>2</sup>ChIP-chip and  $\mu$ ChIP-chip, and between replicates, was carried out with values resulting from Maxfour calculations [17]. This algorithm scores each promoter by finding the highest average  $\log_2$  ratio among ten consecutive probes per tiled region ('MaxTen'). To fully validate  $\mu$ ChIP-chip, correlation analysis was carried out with MaxTen scores for all tiled regions, as opposed to the reported top 30% enriched promoters in an earlier study [13]. Both H3K9ac and H3K9m3  $\mu$ ChIP-chips robustly reproduced the large scale results ( $R = 0.80-0.94$ ; Figure 2a). We then compared results from  $\mu$ ChIP-chip biological replicates with each antibody to demonstrate high reproducibility (Figure 2b). Moreover, comparison between H3K9ac and H3K9m3 MaxTen values showed no correlation when assessed by large scale or  $\mu$ ChIP-chip (data not shown), as expected from the mutually exclusive occupancy of these modifications.

#### **$\mu$ ChIP-chip and Q<sup>2</sup>ChIP-chip enrichment profiles for H3K9ac and H3K9m3**

Examples of normalized  $\log_2$  ChIP/input signal ratios for two segments of chromosome 12 are shown in Figure 3a. For both H3K9ac and H3K9m3, the data show high similarity of enrichment profiles between  $\mu$ ChIP-chip and Q<sup>2</sup>ChIP-chip, and high reproducibility of profiles between  $\mu$ ChIP-chip replicates (Figure 3a). A detailed view of  $\log_2$  ratios for tiled regions selected for their enrichment in either H3K9ac (*POU5F1* and *SOX2* promoters), H3K9m3 (*TRIM40*), both marks (promoter and exon 1 of *NANOG*), or none of these marks (*ESR1*), illustrates the high similarity of enrichment

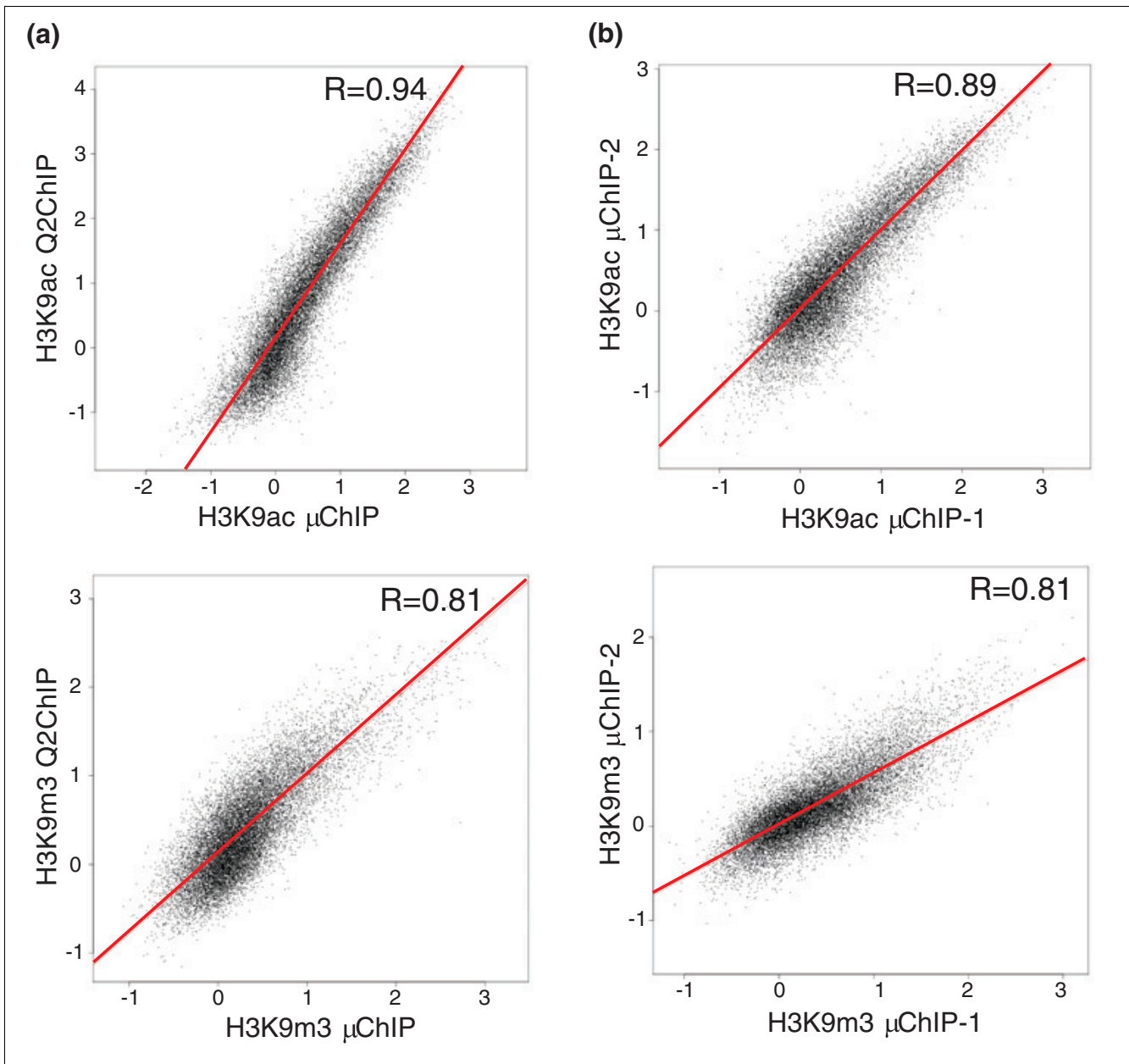
profiles also within a tiled region, and confirms the reproducibility between the two methods (Figure 3b). H3K9ac and H3K9m3 peaks detected with a false discovery rate (FDR) of  $\leq 0.05$  robustly overlapped between both ChIP assays (Figure 3b, red areas). Note, however, that due to limitations in the peak-calling software, the exact position of a peak may vary from array to array, and a broad peak may be called as a multiple peaks [17].

#### **Verification of $\mu$ ChIP-chip data by quantitative PCR**

qPCR analysis of WGA-amplified  $\mu$ ChIP DNA and non-amplified Q<sup>2</sup>ChIP DNA verified the array data (Figure 4). As expected from their expression in NCCIT cells, the *POU5F1*, *NANOG* and *SOX2* promoters were acetylated on H3K9 in the absence of H3K9m3. The *UBE2B* housekeeping promoter was enriched in H3K9ac but not in H3K9m3, similarly to *KNTC1* and *FLJ11021*, also expressed in NCCIT cells. Conversely, promoters of genes not expressed in NCCIT cells were either enriched in H3K9m3 without H3K9ac (*TSH2B*, *H1t*, *ZNF323*, *KCNA1*, *TRIM40*) or enriched in neither H3K9m3 nor H3K9ac (*LDHC*, *ESR1*, *OXT*, *GPR109A*) (Figure 4). Altogether, the data show that: promoters examined are enriched in either H3K9ac, H3K9m3 or none of these modifications, but not in both; active promoters are marked by H3K9ac in the absence of H3K9m3; and inactive promoters are marked by either H3K9m3 in the absence of H3K9ac, or by none of these marks, supporting the existence of both H3K9m3-dependent and independent gene repression mechanisms [18].

#### **Metagene analysis of H3K9ac and H3K9m3 enrichment**

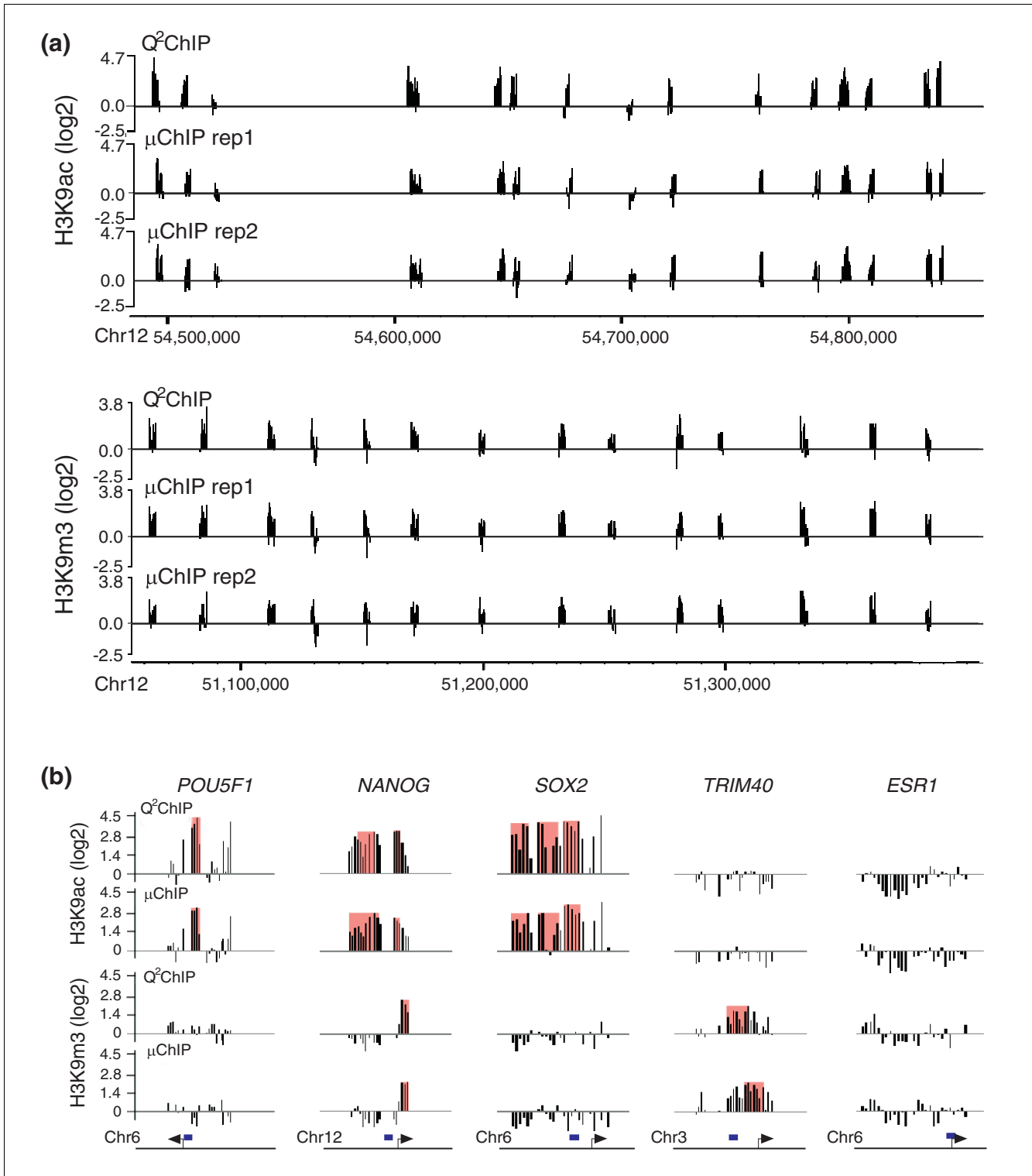
We next compared the average promoter enrichment profiles for H3K9ac and H3K9m3 over 2.7 kb within the tiled regions in large-scale ChIP-chip, and determined whether these profiles were maintained in  $\mu$ ChIP-chip. We created a composite metagene from the collection of genes enriched (by the detection of one or more peaks with  $FDR \leq 0.05$ ; see Materials and methods) within the tiled region in either H3K9ac or H3K9m3 by Q<sup>2</sup>ChIP-chip. These two sets of genes were the basis for similar metagene analysis of  $\mu$ ChIP-chip enrichment. Analysis of modified histone occupancy by large scale assays revealed distinct enrichment profiles for acetylated and trimethylated H3K9 (Figure 5a). H3K9ac showed a bimodal distribution with a pronounced dip immediately upstream of the TSS, suggesting a nucleosome-free region for active genes, as marked by acetylation; in contrast, H3K9m3 was more evenly distributed across the regions examined, with most prominent enrichment in the upstream half of the region and only a slight decrease in signal intensity around the TSS (Figure 5a). These profiles were rigorously conserved in  $\mu$ ChIP-chip, both when we examined the same genes found to be enriched by either modification by Q<sup>2</sup>ChIP (Figure 5b), and when we examined all genes enriched in H3K9ac or H3K9m3 based on  $\mu$ ChIP peak detection (data not shown). These findings support evidence that transcribed genes have a nucleosome-free region immediately upstream of the TSS,



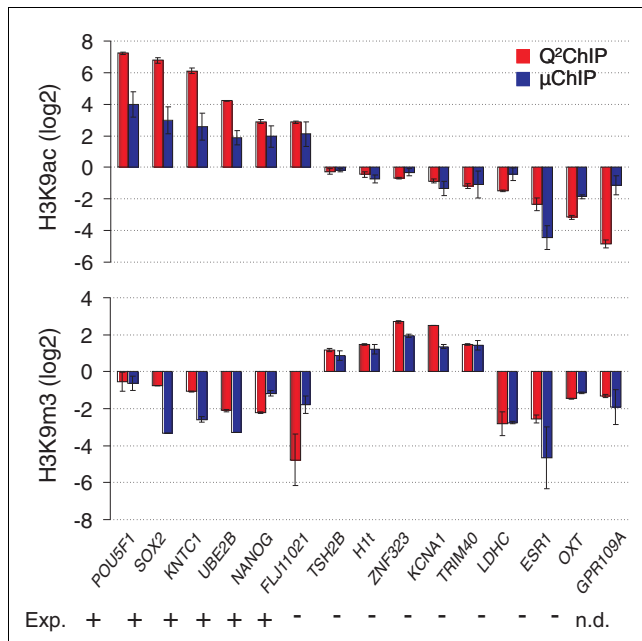
**Figure 2**  
 Reproducibility and specificity of  $\mu$ ChIP-chip. **(a)** Two-dimensional scatter plots comparing  $\mu$ ChIP-chip with Q<sup>2</sup>ChIP-chip using antibodies to H3K9ac and H3K9m3. **(b)** Comparison of two  $\mu$ ChIP-chip experiments using antibodies to H3K9ac and H3K9m3. MaxTen values from each experiment are plotted on a log<sub>2</sub> scale.

whereas most transcriptionally silenced genes do not [19]. H3K9m3-marked genes that lack a nucleosome-free region immediately 5' of the TSS may belong to a group of silent genes that do not recruit a pre-initiation complex, as the absence of pre-initiation complex recruitment in unexpressed genes has been shown to coincide with a lack of nucleosome depletion [20]. Members of a second and smaller group of silent genes do recruit a pre-initiation complex [20], and may be repressed by an H3K9m3-independent mechanism.

Distinct classes of genes based on differential enrichment in H3K9ac and H3K9m3 were evidenced by metagene profiles from  $\mu$ ChIP-chip data on genes solely enriched in either H3K9ac (Figure 6a, left panel), H3K9m3 (Figure 6a, right panel) or both (Figure 6b) within the tiled region. Additionally, in concordance with the qPCR data, we identified a fourth group of genes not enriched in either of these modifications. We also found two groups of genes marked by H3K9ac: those devoid of H3K9m3 (Figure 6a, left panel) and



**Figure 3**  
 Enrichment profiles for H3K9ac and H3K9m3 expressed as log<sub>2</sub> ChIP/input ratios in Q<sup>2</sup>ChIP-chip and μChIP-chip assays. **(a)** Comparison of enrichment profiles over a 360-kb region (H3K9ac) and a 300-kb region (H3K9m3) of chromosome 12 obtained by Q<sup>2</sup>ChIP-chip and in two μChIP-chip replicates. **(b)** Detailed profiles of H3K9ac and H3K9m3 enrichment on the 3-kb tiled regions (shown in bottom panels) of the *POU5F1*, *NANOG*, *SOX2*, *TRIM40* and *ESR1* promoters. Log<sub>2</sub> ChIP/input ratios are shown in black and H3K9ac or H3K9m3 peaks (FDR ≤ 0.05) are superimposed (shaded in red). Tiled regions (bottom panel) also show the region examined by ChIP-qPCR (blue bars) in Figure 4.



**Figure 4**  
ChIP-qPCR analysis of H3K9ac and H3K9m3 association with the tiled regions of indicated genes. Experiments are based on the same cell batches as those examined by microarray. Data are expressed as  $\log_2$  of fold-enrichment relative to input (mean  $\pm$  standard deviation) in three independent experiments. Promoter regions covered by amplicons are shown in Figure 3b, bottom panels. The expression pattern of each gene was determined in duplicate transcriptome analysis of NCCIT cells using Affymetrix UI33A GeneChips (not shown).

those also enriched in H3K9m3 primarily upstream of the acetylated region (Figure 6b). Examination of individual genes confirmed the distribution of each mark throughout the tiled regions. H3K9m3 profiles also appear different based on the analyzed set of genes. Genes solely enriched in H3K9m3 display a relatively even distribution of this modification (Figure 6a, right panel), whereas genes also marked by H3K9ac show lower levels of H3K9m3 in the 3' half of the assessed region (Figure 6b). Metagene analysis of genes harboring both H3K9ac and H3K9m3 peaks within the tiled region reveals mutually exclusive occupancy profiles for these marks, and predominantly contain trimethylation upstream of the region marked by acetylation (Figure 6b). It will be interesting to investigate whether differential marking of genes by H3K9ac and H3K9m3 has implications on transcriptional regulation and hierarchy [21-23].

## Conclusion

Our results demonstrate that fast genome-scale analysis of promoter occupancy by modified histones is possible from as few as 1,000 cells. This represents an improvement over previous ChIP-chip protocols, which require significantly more cells and take at least 4 days up to hybridization [13,24]. Specific steps of the Acevedo *et al.* [13] microChIP protocol that

differ from our  $\mu$ ChIP-chip assay are detailed in Additional data file 3. Although performant, in its current form  $\mu$ ChIP-chip has limitations. First, we cannot formally exclude that the DNA amplification step does not introduce bias. A recent comparison of three widely used amplification procedures, including WGA, reports that all procedures introduce some bias [25]. Notably however, WGA-based amplification resulted in the most accurate performance. We show here by qPCR that amplification of ChIP DNA from 1,000 cells introduces little bias; however, when establishing  $\mu$ ChIP-chip, amplification of DNA from 100 cells produced inconsistencies, presumably due to lower signal/noise ratios [12] and more rounds of amplification were required owing to the minute amount of template DNA used. Development of improved amplification protocols may result in successful application of  $\mu$ ChIP-chip to fewer than 1,000 cells. Secondly,  $\mu$ ChIP-chip may also be suitable for non-histone proteins, although this remains to be tested. For analysis of low-abundance or transiently bound proteins, cell numbers might need to be increased compared to histone  $\mu$ ChIP-chip. Further, we formulate an equation that allows an estimation of the average DNA fragment length produced by sonication of chromatin from minute cell samples. It is imperative to assess chromatin fragmentation prior to ChIP location analysis to ensure good resolution and valid analysis. To our knowledge, this is the only strategy to overcome this task with small cell numbers. Moreover, array data show high reproducibility between biological replicates and conservation of H3K9ac and H3K9m3 enrichment profiles in the large scale and  $\mu$ ChIP-chip assays. We demonstrate that  $\mu$ ChIP-chip can be applied to reveal nucleosome-free regions in as few as 1,000 cells. In addition, metagene analysis reveals distinct occupancy profiles for each histone modification in the tiled regions, which are maintained in  $\mu$ ChIP-chip, and identify four distinct groups of genes in human embryonal carcinoma cells.  $\mu$ ChIP-chip therefore makes genome-wide epigenetic analyses amenable to small cell samples, such as rare stem cell subpopulations, cells from the early embryo or human biopsies.

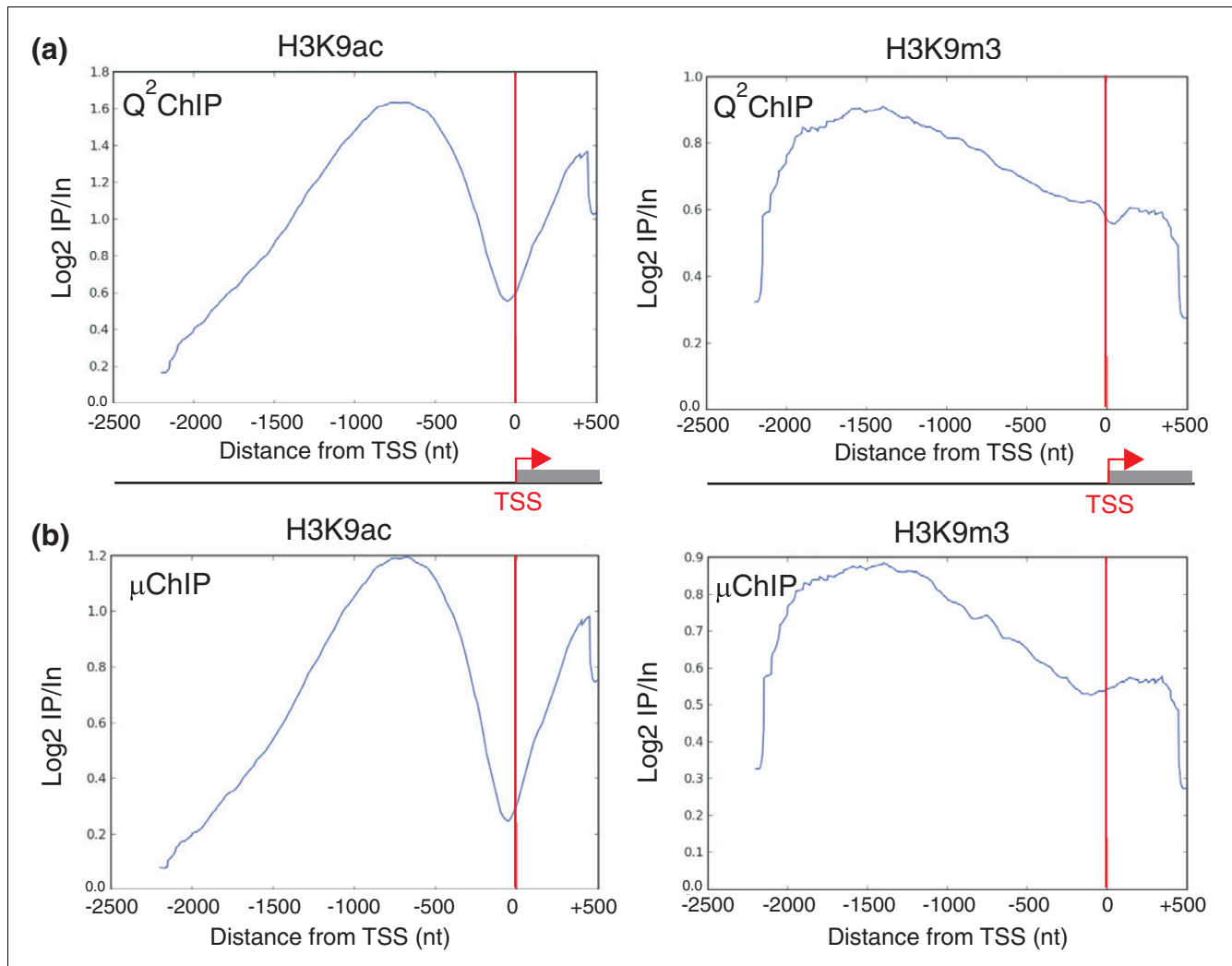
## Materials and methods

### Materials

Pluripotent human embryonal carcinoma NCCIT cells were cultured as described [10]. Antibodies against H3K9ac were from Upstate (Millipore Inc., Billerica, MA, USA; catalogue number 06-942) and antibodies to H3K9m3 were from Diagenode (Liège, Belgium; catalogue number pAb-056-050). All other reagents were from Sigma-Aldrich unless otherwise indicated.

### ChIP assays

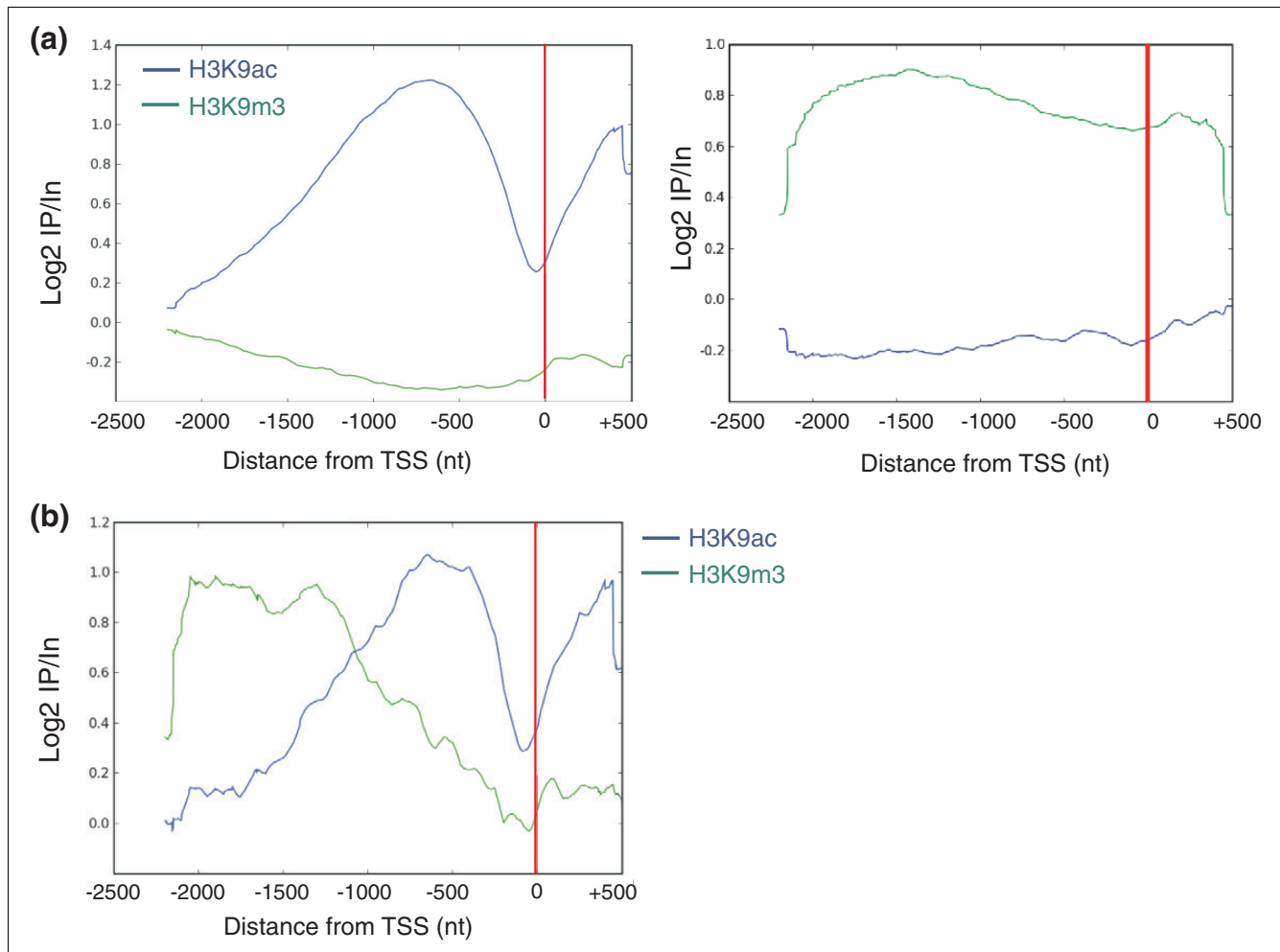
The Q<sup>2</sup>ChIP assay, referred to as large scale ChIP, was performed as described [10]. Chromatin was prepared from  $2 \times 10^6$  cells, diluted to  $2 A_{260}$  units and aliquoted into 100  $\mu$ l per ChIP.

**Figure 5**

Metagenome analysis of the distribution of H3K9ac and H3K9m3 in Q<sup>2</sup>ChIP-chip and μChIP-chip assays. H3K9ac and H3K9m3 occupancy profiles detected by (a) Q<sup>2</sup>ChIP-chip and (b) μChIP-chip in the 2.7 kb tiled regions. In (b), the same genes showing enrichment in either mark by Q<sup>2</sup>ChIP-chip (genes shown in (a)) were included in the analysis.

μChIP was carried out from 1,000 cells per ChIP [11,12] with modifications and optimization to enable genome-wide analysis. A troubleshooting guide is presented in Additional data file 4. Primary antibodies (2.4 μg) were coupled to Dynabeads Protein A (10 μl; Dynal Biotech, Invitrogen, Oslo, Norway) in RIPA buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% SDS, 0.1% Na-deoxycholate, 140 mM NaCl) for 2 h at 4°C. Tubes were in strips handled in a magnetic rack (Diagenode, catalogue number kch-816-001). Concomitantly, cells were detached by a few sharp blows to the flask in the presence of 20 mM of the histone deacetylase inhibitor sodium butyrate, and aliquots of 1,000 cells were suspended in 500 μl phosphate-buffered saline/butyrate. Proteins and DNA were cross-linked with 1% formaldehyde for 8 minutes and cross-linking was stopped with 125 mM glycine. Cells were centrifuged at 620 g in a swing-

out rotor for 10 minutes at 4°C and washed twice in 0.5 ml ice-cold phosphate-buffered saline/butyrate by gentle vortexing and centrifugation as above. Approximately 20 μl buffer was left with the pellet after removal of the last wash. Cells were lysed by addition of 120 μl room temperature lysis buffer (50 mM Tris-HCl, pH 8, 10 mM EDTA, 1% SDS, protease inhibitor cocktail (Sigma-Aldrich, catalogue number P8340), 1 mM PMSF, 20 mM butyrate) and thorough vortexing. Following a 3-minute incubation on ice and additional vortexing, nuclei were centrifuged at 860 g for 10 minutes and the supernatant discarded, leaving approximately 20 μl lysis buffer in the tube. RIPA buffer (120 μl containing protease inhibitor cocktail, 1 mM PMSF, 20 mM butyrate) was added, the tube was vortexed thoroughly and cells sonicated for 3 × 30 s on ice with 30 s pauses, using a probe sonicator (Labsonic-M, 3-mm probe; cycle 0.5, 30% power; Sartorius AG, Göttingen, Ger-

**Figure 6**

$\mu$ ChIP-chip identifies differential enrichment in H3K9ac and H3K9m3. **(a)** H3K9ac and H3K9m3 occupancy profiles detected by  $\mu$ ChIP-chip on tiled regions with peaks in H3K9ac only (left) and H3K9m3 only (right). Genes containing at least one peak in H3K9m3 or H3K9ac, respectively, were removed from the analysis. **(b)** H3K9ac and H3K9m3 occupancy profiles exclusively on tiled regions containing both marks.

many) to produce fragments of approximately 400-500 bp. The sample was centrifuged at 12,000 g for 10 minutes at 4°C and the supernatant was transferred into a 0.2-ml PCR tube containing beads pre-incubated with antibodies, leaving approximately 10  $\mu$ l of supernatant behind. A chromatin sample identical to that used in the ChIP sample was prepared to represent the input and transferred into a 1.5 ml tube.

Proper chromatin fragmentation from 1,000 cells was ensured using a qPCR assay further developed from that previously described [12]. Since the reliability of ChIP depends on control of chromatin fragmentation, we accomplished this for our genome-scale location analysis by formulating an equation that enables an estimation of average fragment length in minute cell samples. A linear relationship between average DNA fragment length and qPCR signal intensities within a fragmentation range useful for ChIP (300-600 bp) was described through this equation, which allows calcula-

tion of DNA fragment length after experimental determination of the relative PCR signal in a sonicated 1,000-cell sample (Additional data file 1).

Immunoprecipitation and washes of the ChIP product were performed essentially as described [12]. Beads were released into the chromatin suspension and rotated at 40 rpm for 2 h at 4°C. The ChIP material was washed three times by 4-minute incubations in 100  $\mu$ l of RIPA buffer and once in 100  $\mu$ l of 10 mM Tris-HCl, pH 8.0, 10 mM EDTA (TE) buffer, and transferred into a new tube while in TE buffer. Elution buffer (20 mM Tris-HCl, pH 7.5, 5 mM EDTA, 20 mM butyrate, 50 mM NaCl, 1% SDS) (150  $\mu$ l) and 5  $\mu$ g RNase (Roche, Basel, Switzerland; catalogue number 11119915001) were added after removal of TE buffer. The same amount of RNase was added to the input, and ChIP and input samples were incubated at 37°C for 20 minutes on a Thermomixer (1,300 rpm; Eppendorf, Hamburg, Germany). Samples were briefly cen-



trifuged, 1  $\mu$ l of proteinase K (at 20  $\mu$ g/ $\mu$ l) was added and tube lids were replaced by new ones to prevent leakage resulting from softening of the plastic upon heating. DNA elution, cross-link reversal and proteinase K digestion were carried out in a single step for 2 h at 68°C on a Thermomixer. After capturing of beads, the supernatant was recovered, beads were incubated for another 5 minutes in 150  $\mu$ l elution buffer containing 50  $\mu$ g/ml proteinase K, and both supernatants were pooled. ChIP and input samples were made up to a final volume of 490  $\mu$ l in elution buffer without SDS. ChIP DNA was extracted with phenol-chloroform isoamylalcohol and chloroform isoamylalcohol, ethanol-precipitated in the presence of 10  $\mu$ l acrylamide carrier (Sigma-Aldrich, catalogue number A9099) and dissolved in 10  $\mu$ l MilliQ water.

### Whole genome amplification and clean up of ChIP DNA

ChIP and input DNA were amplified with WGA2 (Q<sup>2</sup>ChIP) or WGA4 ( $\mu$ ChIP) GenomePlex Whole Genome Amplification Kits (Sigma-Aldrich) as per the manufacturer's instructions; however, we omitted the lysis and DNA fragmentation steps. Starting from step 6 in the WGA procedures, library preparation was carried out and immediately followed by amplification for 14 or 25 cycles for Q<sup>2</sup>ChIP and  $\mu$ ChIP, respectively.

Amplified DNA was cleaned up using the QIAquick PCR purification kit (Qiagen, catalogue number 28104) as per the manufacturer's instructions except that five volumes of buffer PB (Qiagen, catalogue number 19066) were used instead of buffer PBI to ensure the absence of pH indicator in the sample (the pH indicator in buffer PBI may interfere with microarray applications). Furthermore, DNA was eluted in 30  $\mu$ l 1 mM Tris-HCl, pH 8.0. The kit is designed for purification of DNA fragments of 100-10,000 bp, and thus was well suited for ChIP and input DNA fragments. Following DNA purification, samples were quantified by NanoDrop (NanoDrop Technologies, Wilmington, DE, USA) and aliquots were diluted to 7.5 ng/ $\mu$ l in TE buffer for PCR-based quality assessment. Importantly, parallel ChIP experiments were carried out without amplification and were directly assessed by qPCR to serve as a reference for amplified samples as well as for array data. Quality of amplified samples was also evaluated by agarose gel electrophoresis. Typically, amplification produced 7.5-15  $\mu$ g DNA (depending on the WGA kit lot number) with an average size of approximately 400-500 bp. WGA amplification can therefore yield enough DNA to probe as many as seven arrays without further amplification.

### DNA labeling and array hybridization

ChIP and input DNA fragments were labelled with Cy5 and Cy3, respectively, and hybridized onto Nimblegen human HG18 RefSeq Promoter arrays. Arrays covered approximately 27,000 human RefSeq promoters, ranging from -2,200 to +500 bp relative to the transcription start site (TSS). Probes consisted of 385,000 50- to 75-mers tiled throughout non-repetitive genomic sequences at a median spacing of 100 bp.

Sequence source for probes was the UCSC Genome Browser. ChIP and input DNA labeling, hybridization and detection were performed using the services of Nimblegen (Madison, WI, USA).

### Data analysis

Signal intensity data were extracted from the scanned images of each array using NimbleScan software. Log<sub>2</sub> ChIP/input ratios were scaled and centered around zero by subtracting the bi-weight mean for the log<sub>2</sub> ratio values for all features on the array from each log<sub>2</sub> ratio value. Peaks were detected by searching for four or more probes with a signal above a cut-off value using a 500-bp sliding window. Cut-off values were a percentage of a hypothetical maximum defined as (mean + 6 [standard deviation]). Ratio data were randomized 20 times to evaluate the probability of false positives, and each peak was assigned a FDR score. Normalization and peak detection were performed by Nimblegen in accordance with their protocols. This process uses a cut-off range of 90% to 15%, with higher cut-offs corresponding to more stringent peak detection, as reflected in the FDR calculation. The Nimblegen protocol was recently evaluated as part of a comprehensive study that objectively analyzed the performance of a number of commercially available ChIP-chip array platforms and signal detection algorithms [26], and found to deliver reliable results.

For scoring the promoters before correlation analysis, we assigned an amplification value to each promoter by applying the Maxfour algorithm with a ten-probe window [27]. For each promoter, the corresponding probes' log<sub>2</sub> ratios were scanned in genome order with a ten-probe window. The highest ten-probe average was used as the amplification value for the promoter. Promoters represented by less than ten probes (1.5% of the total) were not included in the analysis.

### Metagene analysis of regions containing H3K9ac or H3K9m3

Metagene analysis of promoter occupancy was performed essentially as described [28]. Genes with a high probability of enrichment (FDR  $\leq$  0.05) in H3K9ac or H3K9m3 marks within the tiled region were collected and used to assemble a metagene of the average composite binding. Each region was interrogated for probes and these were mapped into a 2.7-kb wide window at the appropriate offsets based on strand orientation. Linear interpolation was used to estimate the fold enrichment at each base position within the 2.7-kb window. This interpolation left the 5' and 3' ends of the window under-represented. The metagene was created from this collection of functions by calculating the mean of the values mapped to each position by all the regions found to be enriched in H3K9ac or H3K9m3 Q<sup>2</sup>ChIP or  $\mu$ ChIP. If the offset corresponded to the exact location of a probe within a specific tiled region, values were directly measured; if not, values were linearly interpolated from the values of the two flanking probes [28]. Genes solely enriched by only one of the examined

marks were selected from the entire set of genes harboring the mark (peak detection with  $FDR \leq 0.05$ ) and then removing from that set all genes also possessing a peak for the other mark.

### Quantification of non-amplified ChIP DNA

Because the NanoDrop spectrophotometer does not allow accurate quantification of minute amounts of non-amplified ChIP DNA, we used a Qubit fluorometer (Invitrogen, Carlsbad, CA, USA; catalogue number Q32857) and a Quant-iT dsDNA HS kit (Invitrogen, catalogue number Q32851) for quantification. Ten percent of Q<sup>2</sup>ChIP DNA samples and whole  $\mu$ ChIP inputs were mixed with Quant-iT working solution to a final volume of 200  $\mu$ l, incubated for 2 minutes and analyzed by the Quant-iT DNA HS program on a Qubit fluorometer.

### Quantitative PCR

Immunoprecipitated DNA from three independent ChIPs was analyzed by duplicate qPCR [10] (Additional data file 5). qPCR data are expressed as mean ( $\pm$  standard deviation)  $\log_2$  values of enrichment relative to input DNA.

### Abbreviations

ChIP: chromatin immunoprecipitation; FDR: false discovery rate; H<sub>3</sub>K<sub>9</sub>ac: acetylated lysine 9 of histone H<sub>3</sub>; H<sub>3</sub>K<sub>9</sub>m<sub>3</sub>: trimethylated lysine 9 of histone H<sub>3</sub>;  $\mu$ ChIP: microChIP; qPCR: quantitative PCR; TSS: transcription start site; WGA: whole-genome amplification.

### Authors' contributions

JAD designed the study, performed experiments, contributed to analysis design, made figures and wrote parts of the manuscript. AHR performed bioinformatics analyses, made figures and wrote parts of the methods. PC designed the study, wrote the manuscript, made figures and supervised the work. All authors read and approved the final manuscript.

### Additional data files

The following additional data are available with the online version of this paper. Additional data file 1 is a figure describing the steps behind the equation formulated to estimate, using qPCR, chromatin fragment length after a given sonication regime of 1,000 cells. Additional data file 2 a table providing values of DNA recovery from Q<sup>2</sup>ChIP and  $\mu$ ChIP. Additional data file 3 is information on a technical comparison between  $\mu$ ChIP-chip and a previously published protocol. Additional data file 4 is a troubleshooting guide for  $\mu$ ChIP-chip. Additional data file 5 is a table listing ChIP qPCR primers used in this study.

### Acknowledgements

This work was supported by the FUGE, YFF, STAMCELLER and STOR-FORSK programs of the Research Council of Norway, and by the Norwegian Cancer Society.

### References

1. Dedon PC, Soultis JA, Allis CD, Gorovsky MA: **Formaldehyde cross-linking and immunoprecipitation demonstrate developmental changes in H1 association with transcriptionally active genes.** *Mol Cell Biol* 1991, **11**:1729-1733.
2. Hebbes TR, Clayton AL, Thorne AW, Crane-Robinson C: **Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken  $\beta$ -globin chromosomal domain.** *EMBO J* 1994, **13**:1823-1830.
3. O'Neill LP, Turner BM: **Immunoprecipitation of chromatin.** *Methods Enzymol* 1996, **274**:189-197.
4. Solomon MJ, Larsen PL, Varshavsky A: **Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene.** *Cell* 1988, **53**:937-947.
5. Bernstein BE, Meissner A, Lander ES: **The mammalian epigenome.** *Cell* 2007, **128**:669-681.
6. Bird A: **Perceptions of epigenetics.** *Nature* 2007, **447**:396-398.
7. Collas P, Dahl JA: **Chop it, CHIP it, check it: the current status of chromatin immunoprecipitation.** *Front Biosci* 2008, **13**:929-943.
8. O'Neill LP, Vermilyea MD, Turner BM: **Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations.** *Nat Genet* 2006, **38**:835-841.
9. Nelson JD, Denisenko O, Sova P, Bomszyk K: **Fast chromatin immunoprecipitation assay.** *Nucleic Acids Res* 2006, **34**:e2.
10. Dahl JA, Collas P: **Q<sup>2</sup> ChIP, a quick and quantitative chromatin immunoprecipitation assay unravels epigenetic dynamics of developmentally regulated genes in human carcinoma cells.** *Stem Cells* 2007, **25**:1037-1046.
11. Dahl JA, Collas P: **MicroChIP - A rapid micro chromatin immunoprecipitation assay for small cell samples and biopsies.** *Nucleic Acids Res* 2008, **36**:e15.
12. Dahl JA, Collas P: **A rapid micro chromatin immunoprecipitation assay ( $\mu$ ChIP).** *Nat Protoc* 2008, **3**:1032-1045.
13. Acevedo LG, Iniguez AL, Holster HL, Zhang X, Green R, Farnham PJ: **Genome-scale ChIP-chip analysis using 10,000 human cells.** *Biotechniques* 2007, **43**:791-797.
14. Vakoc CR, Mandat SA, Olenchok BA, Blobel GA: **Histone H3 lysine 9 methylation and HPI gamma are associated with transcription elongation through mammalian chromatin.** *Mol Cell* 2005, **19**:381-391.
15. Hess HH, Thalheimer C: **DNA and RNA and the cytoarchitecture of human frontal cortex.** *J Neurochem* 1971, **18**:1281-1290.
16. Forslund O, Antonsson A, Higgins G, Ly H, Delius H, Hunziker A, de Villiers EM: **Nucleotide sequence and phylogenetic classification of candidate human papilloma virus type 92.** *Virology* 2003, **312**:255-260.
17. O'Geen H, Nicolet CM, Blahnik K, Green R, Farnham PJ: **Comparison of sample preparation methods for ChIP-chip assays.** *Biotechniques* 2006, **41**:577-580.
18. Kouzarides T: **Chromatin modifications and their function.** *Cell* 2007, **128**:693-705.
19. Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z, Furey TS, Crawford GE: **High-resolution mapping and characterization of open chromatin across the genome.** *Cell* 2008, **132**:311-322.
20. Oszolak F, Song JS, Liu XS, Fisher DE: **High-throughput mapping of the chromatin structure of human promoters.** *Nat Biotechnol* 2007, **25**:244-248.
21. Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R, Young RA: **Core transcriptional regulatory circuitry in human embryonic stem cells.** *Cell* 2005, **122**:947-956.
22. Kim J, Chu J, Shen X, Wang J, Orkin SH: **An extended transcriptional network for pluripotency of embryonic stem cells.** *Cell* 2008, **132**:1049-1061.
23. Loh YH, Wu Q, Chew JL, Vega VB, Zhang W, Chen X, Bourque G, George J, Leong B, Liu J, Wong KY, Sung KW, Lee CW, Zhao XD,

- Chiu KP, Lipovich L, Kuznetsov VA, Robson P, Stanton LW, Wei CL, Ruan Y, Lim B, Ng HH: **The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells.** *Nat Genet* 2006, **38**:431-440.
24. Lee TI, Johnstone SE, Young RA: **Chromatin immunoprecipitation and microarray-based analysis of protein location.** *Nat Protoc* 2006, **1**:729-748.
  25. Ponzilli R, Boutros PC, Katz S, Stojanova A, Hanley AP, Khosravi F, Bros C, Jurisica I, Penn LZ: **Optimization of experimental design parameters for high-throughput chromatin immunoprecipitation studies.** *Nucleic Acids Res* 2008, **36**:e144.
  26. Johnson DS, Li W, Gordon DB, Bhattacharjee A, Curry B, Ghosh J, Brizuela L, Carroll JS, Brown M, Flicek P, Koch CM, Dunham I, Bieda M, Xu X, Farnham PJ, Kapranov P, Nix DA, Gingeras TR, Zhang X, Holster H, Jiang N, Green RD, Song JS, McCuine SA, Anton E, Nguyen L, Trinklein ND, Ye Z, Ching K, Hawkins D, et al.: **Systematic evaluation of variability in ChIP-chip experiments using predefined DNA targets.** *Genome Res* 2008, **18**:393-403.
  27. Acevedo LG, Bieda M, Green R, Farnham PJ: **Analysis of the mechanisms mediating tumor-specific changes in gene expression in human liver tumors.** *Cancer Res* 2008, **68**:2641-2651.
  28. Wardle FC, Odom DT, Bell GW, Yuan B, Danford TW, Wieltte EL, Herbolsheimer E, Sive HL, Young RA, Smith JC: **Zebrafish promoter microarrays identify actively transcribed embryonic genes.** *Genome Biol* 2006, **7**:R71.