HHS Public Access

Author manuscript

Nat Methods. Author manuscript; available in PMC 2020 April 28.

Published in final edited form as:

Nat Methods. 2019 April ; 16(4): 323–325. doi:10.1038/s41592-019-0361-7.

Single cell chromatin immunocleavage sequencing (scChIC-Seq) to profile histone modification

Wai Lim Ku[1,4], **Kosuke Nakamura**[1,2,4], **Weiwu Gao**[1,3,4], **Kairong Cui**[1], **Gangqing Hu**[1], **Qingsong Tang**[1], **Bing Ni**[3,#], **Keji Zhao**[1,#]

[1.]Laboratory of Epigenome Biology, Systems Biology Center, NHLBI, NIH, Bethesda, MD 20892, USA

[2.]National Institute of Health Sciences, 3-25-26 Tonomachi, Kawasaki-ku, Kawasaki, Kanagawa 210-9501, Japan

[3.]Department of Pathophysiology and High Altitude Pathology, Third Military Medical University, Chongqing, China

[4.]These authors contributed equally to this work.

Abstract

We report a single-cell chromatin immunocleavage sequencing (scChIC-seq) methodology for analyzing histone modifications, which involves targeting of the micrococcal nuclease (MNase) by tethering it to an antibody and selective PCR amplification of cleaved target sites. We show that the protocol reliably detects the H3K4me3 and H3K27me3 target sites in single human white blood cells (WBC), resulting data for successful identification of unique blood cell types based on clustering analysis.

Introduction and results

Recent studies have revealed a potential association of cellular heterogeneity in gene expression with that in the chromatin state of individual cells within the population[1-3].

Several single-cell epigenomic techniques have been reported recently, including scBS-seq [4], scATAC-seq [5,6], scDNase-seq [2], scNOME-seq [7,8] and scMNase-seq[3]. However, although ChIP-Seq[9] has been a crucial technique in evaluating chromatin states and a number of sensitive ChIP-seq derivatives[10-15] are available, a sensitive single-cell ChIP-seq method is still lacking [16]. Laemmli lab previously reported an alternative strategy to detect binding sites of transcription factors in the genome by targeting micrococcal nuclease (MNase) conjugated with protein A (PA) through a specific antibody (Ab), termed chromatin immunocleavage (ChIC) [17]. Recently, Henikoff lab combined ChIC with sequencing to

[#]Correspondence: Bing Ni, nibingxi@126.com; Keji Zhao, zhaok@nhlbi.nih.gov, Phone: 301-496-2098, Fax : 301-402-0971.

ACCESSION NUMBERS

The scChIC-seq data were deposited in the Gene Expression Omnibus database with accession number GSE105012

Analysis code. Codes for analysis of the figures are available at https://github.com/wailimku/scChIC-seq.git

detect genome-wide transcription factor binding sites and histone modifications on native chromatin in a small number of cells (CUT&RUN) [18].

In this study, we developed a single-cell chromatin immunocleavage sequencing method (scChIC-seq), which measures the epigenetic profiles at a single-cell level (Fig. 1a, Extended Fig. 1a). In scChIC-seq, chromatin is cleaved at sites of histone modifications or TF binding by MNase that is recruited to specific chromatin regions by a specific antibody either through direct covalent conjugation with the antibody (**Ab-MNase)** or through protein A-antibody interaction (**Ab+PA-MNase)** (Fig. 1a). The direct covalent conjugation between antibody and MNase eliminates the Ab and PA interaction step. On chromatin, MNase cleaves DNA around the nucleosome with the histone modification into small fragments. To minimize DNA loss in library preparation, both target and non-target DNA fragments are recovered and ligated to the adaptors. Since the targets are smaller fragments compared to non-target DNA, they are preferentially amplified by selective PCR conditions and isolated by agarose gel electrophoresis and sequenced on NGS platforms. In comparison to CUT&RUN [18], our scChIC-seq assays works well (1) with either covalent antibody-MNase conjugates or the complex between antibody and protein A-Mnase; (2) with either uncross-linked cells or cells cross-linked by formaldehyde to covalently stabilize the TF binding; and (3) without the need to isolate the soluble target sites. We feel that ChIC sequencing reflects better the nature of the protocol and thus we term our protocol as scChIC-seq following the original nomenclature of Laemmli lab's publication [17].

We first applied the scChIC-seq protocol to various numbers of NIH3T3 cells (100, 300, 1,000, and 3,000) using the covalent H3K4me3 Ab-MNase conjugate and reproducibly detected peaks of H3K4me3 at gene promoters (Fig. 1b, Extended Data Figs. 1b and 1c, Supplemental Table S1). Global analysis indicated that the scChIC-seq reads are enriched around transcription start sites (TSS) of genes (Extended Data Fig. 2a). The read densities from scChIC-seq and bulk cell ChIP-seq were highly correlated (r = 0.9) (Extended Data Figs. 2b-f, Supplemental Table S2) and the peaks identified by the two methods were highly overlapped (about 80-85%) (Extended Data Figs. 2g-i), indicating that the scChIC-seq protocol is capable of profiling H3K4me3 with a small number of cells.

Next, we compared the strategies of using the covalent H3K4me3 Ab-MNase conjugate and H3K4me3 Ab+PA-MNase complex in scChIC-seq and found that while RIPA buffer worked well for binding and washing steps for the covalent Ab-MNase conjugate, a buffer containing 400 mM NaCl and 1% Triton X-100 was optimal for the Ab+PA-MNase complex (Extended Data Figs. 3 and 4). Both strategies revealed similar H3K4me3 peaks on chromatin (Extended Data Fig. 5a) and provided results consistent with the bulk cell H3K4me3 ChIP-seq (Extended Data Fig. 5b-e), although the covalent Ab-MNase conjugate provides lower backgrounds and higher signal-to-noise ratios (Extended Data Figs 3-5, Supplemental Method).

scChIC-seq was capable of detecting specific H3K4me3 signals in different cell types including NIH3T3 cells, mouse embryonic stem cells (mESCs) and naïve CD4 T cells (Extended Data Fig. 6a-d, Supplemental Table S2). scChIC-seq also detected the target sites

of the repressive histone mark H3K27me3 using H3K27me3 Ab+PA-MNase with 3000 NIH3T3 cells (Extended Data Fig. 7a, b).

Next, we tested whether scChIC-seq is applicable to profiling H3K4me3 at a single-cell level from a heterogeneous cell population. Using human white blood cells (WBCs) and the H3K4me3 Ab-MNase conjugate, 285 single-cell libraries were constructed and sequenced, which provided approximately 100,000 unique reads per cell. The H3K4me3 reads pooled from the single cells show a similar pattern with that of bulk cell H3K4me3 ChIP-seq data (Fig. 1c) and were enriched around TSSs (Extended Data Fig. 8a). By pooling the reads from single cells, we identified 24,819 peaks, 61% of which overlapped with the known H3K4me3 peaks identified from bulk WBCs (Extended Data Fig. 8b). The pooled reads are significantly correlated with the bulk cell ChIP-seq reads (r=0.77) (Extended Data Fig. 8c). On average, 50% of the single-cell reads were mapped to the known peak regions (Extended Data Fig. 8d), and about 5,000 peaks were detected from one single-cell, with a sensitivity of 10% from all 242 informative cells and 18% from the top 10% cells (Extended Data Fig. 8e). These results indicate that the scChIC-seq protocol is a sensitive method for profiling H3K4me3 in single primary cells.

Next, we applied scChIC-seq to profiling H3K27me3 at a single-cell level in human WBCs. 106 single-cell libraries were constructed and sequenced, which provided approximately 131K unique reads per cell (Supplemental Table S1). The H3K27me3 reads pooled from the single cells show a similar pattern with that of bulk cell ENCODE data (Extended Data Fig. 9a). We identified 21,465 peaks from pooled single-cell data, 50% of which overlapped with the known H3K27me3 peaks identified from bulk cell H3K27me3 ChIP-seq data in WBCs (Extended Data Fig. 9b). The reads pooled from the 84 single cells after filtering cells with small number of reads (Supplemental Methods) were significantly correlated with the bulk cell ChIP-seq reads (r=0.67) (Extended Data Fig. 9c). On average, 47% of the single-cell reads were mapped to the known peak regions (Extended Data Fig. 9d), and about 3000 peaks were detected from one single-cell, with a sensitivity of 9.5% from all 84 cells (Extended Data Fig. 9e). These results indicate that the scChIC-seq protocol is applicable to profiling H3K27me3 in single primary cells.

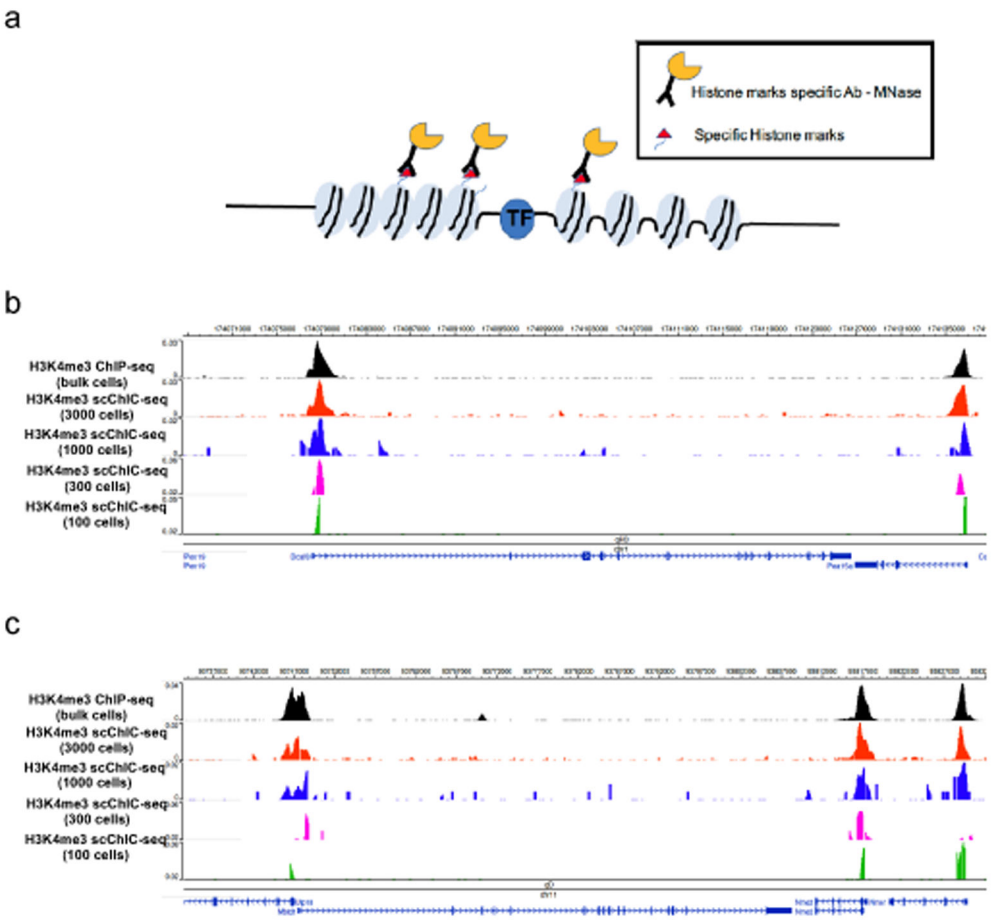
Next, we performed clustering analysis by applying the SC3 software [19] to a set of 12,779 informative H3K4me3 peaks derived from the 242 single cell H3K4me3 libraries (Supplemental Methods). Seven clusters were generated and 5 of them had significant cluster marker peaks (FDR <0.05) (Fig. 2a). By comparing with cell-specific peaks identified from the bulk cell H3K4me3 ChIP-seq data of T, NK, B cells and monocytes (Supplemental Methods, Supplemental Table S2), we identified the cell types represented by these clusters (Figs. 2a and 2b). In total,15, 12, 41 and 43 cells are confidently annotated to be monocyte, B, T, and NK cells respectively. The H3K4me3 data of 12 cells from each group were displayed in Genome Browser, which showed patterns consistent with the cell-specific H3K4me3 signals detected by bulk cell ChIP-seq at several cell-specific genes (Fig. 2c).

We found that cell-to-cell variation in H3K4me3 is positively correlated to that in gene expression in both T cells and monocytes (Extended Data Figs. 10a,c); co-methylated

H3K4me3 peak pairs are positively correlated to co-expression of the corresponding pair of genes (Extended Data Figs. 10b,d) for both T cells and Monocytes. Furthermore, highly variable peaks and highly co-methylated peaks of H3K4me3 may provide information to identify different sub-cell types within the whole underlying cell population (supplemental figures 10 and 11 and Supplemental Results).

In summary, our results indicate that scChIC-seq works well as a single-cell technique for profiling genome-wide histone modifications in a complex primary cell population, which may find broad applications in elucidating chromatin states in rare primary and patients samples.

Extended Data



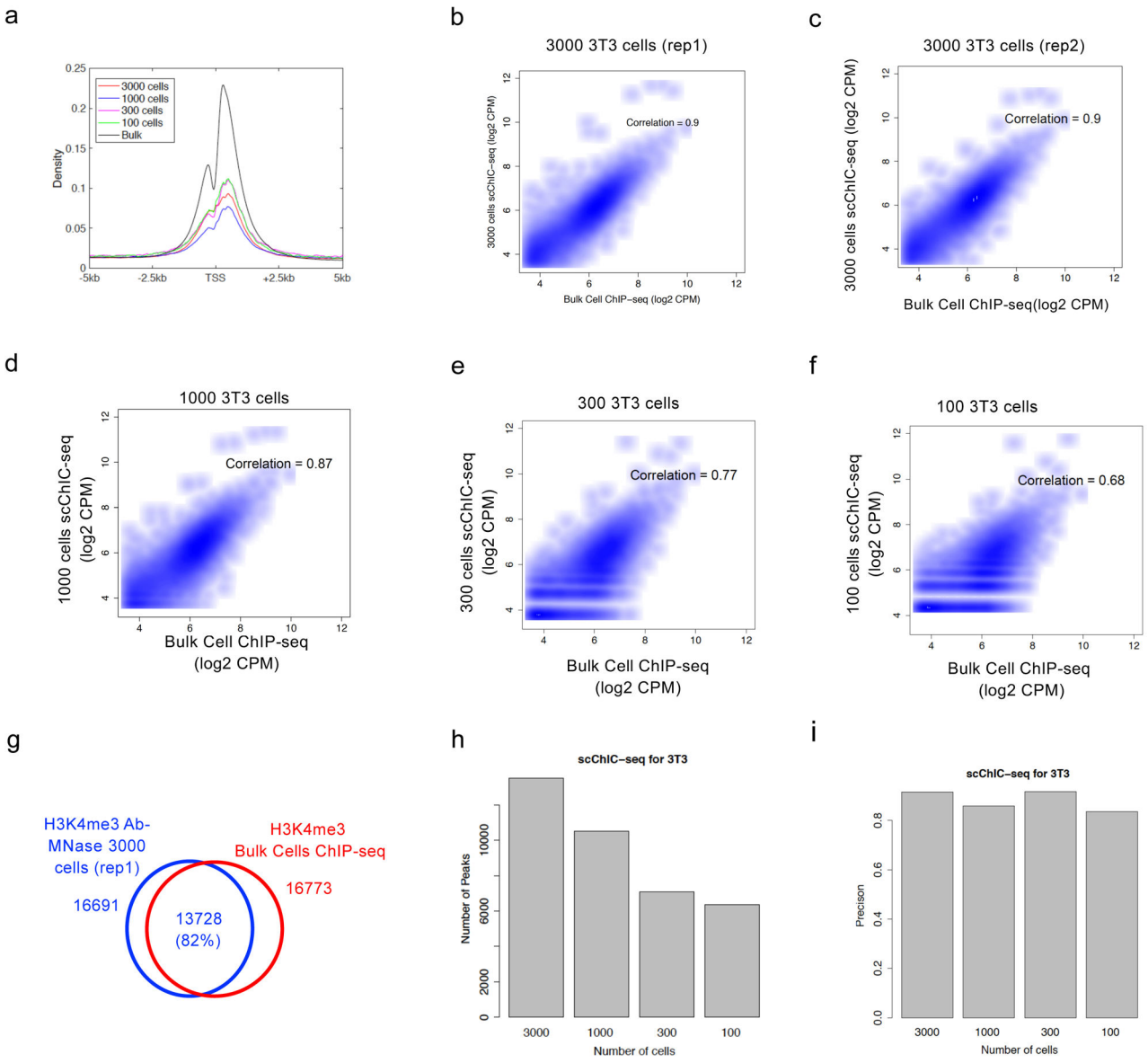
Extended data Figure 1.

Measurement of H3K4me3 profiles by the scChIC-seq using 100, 300, 1,000 and 3,000 NIH3T3 cells

a. H3K4me3 Ab-MNase conjugates are added to the pre-treated cells, such that the conjugates could be bound to the locations with the H3K4me3 histone modification mark.

b. A genome browser snapshot showing the H3K4me3 profiles around the locus of Dcaf8 for 100, 300, 1,000 and 3,000 cells.

c. A genome browser snapshot showing the H3K4me3 profiles around the locus of Mbtd1 for 100, 300, 1,000 and 3,000 cells



Extended data Figure 2.

scChIC-seq detects H3K4me3 profiles in a small number of cells

a. A TSS profile plot for H3K4me3 measured by ChIP-seq using bulk cells (black) and by scChIC-seq using 100 (green), 300 (magenta), 1,000 (blue) and 3,000 (red) cells.

b. A scatter plot of the H3K4me3 read density of ChIP-seq (bulk cell) versus that of scChIC-seq (3,000 cells rep1) at the enriched regions identified using the H3K4me3 ChIP-seq (bulk cells). The correlation was computed using the Pearson correlation coefficient.

c. Scatter plots for the H3K4me3 scChIC-seq data (3,000 cells rep2) versus H3K4me3 ChIP-seq in NIH-3T3 cells.

d. Scatter plots for the H3K4me3 scChIC-seq data (1,000 cells) versus H3K4me3 ChIP-seq in NIH-3T3 cells.

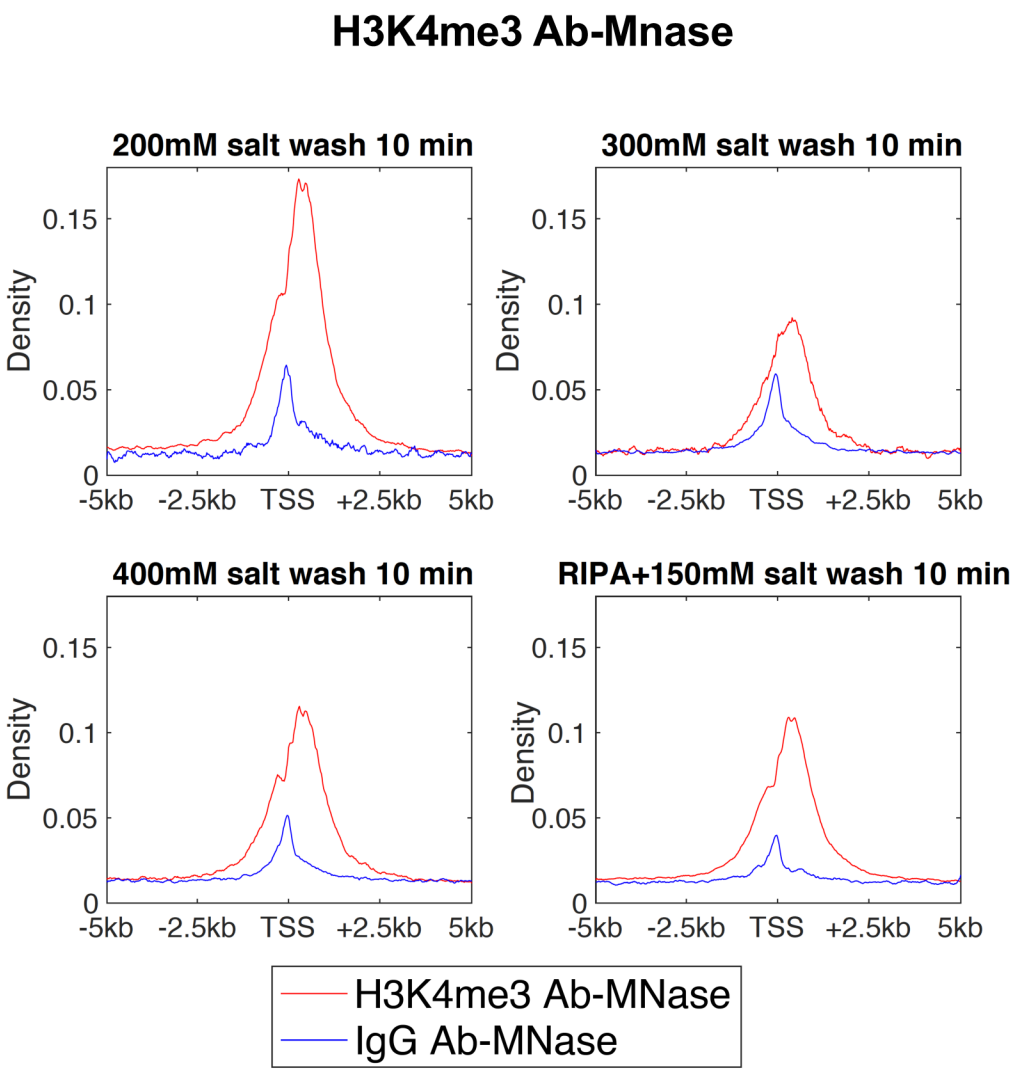
e. Scatter plots for the H3K4me3 scChIC-seq data (300 cells) versus H3K4me3 ChIP-seq in NIH-3T3 cells.

f. Scatter plots for the H3K4me3 scChIC-seq data (100 cells) versus H3K4me3 ChIP-seq in NIH-3T3 cells.

g. A Venn diagram showing the overlap of the enriched regions of H3K4me3 profiles measured by ChIP-seq using bulk cells (red) and by scChIC-seq using 3,000 cells rep1 (blue).

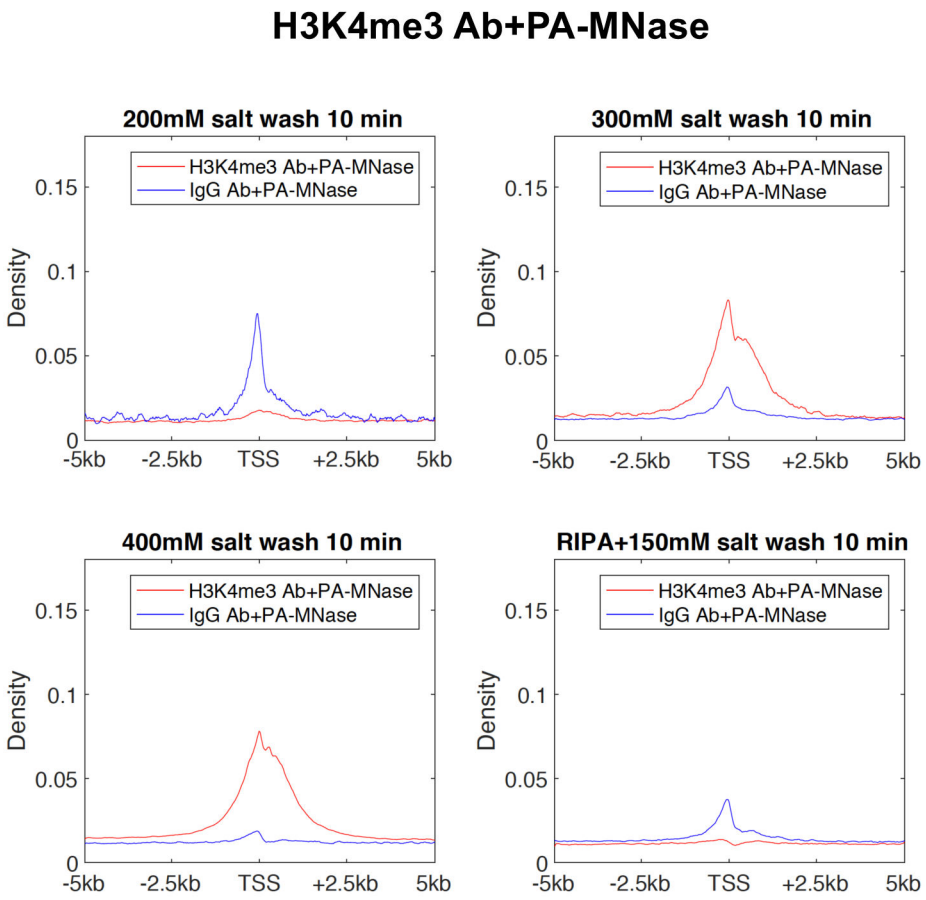
h. A histogram showing the fractions of enriched regions identified by scChIC-seq that are overlapped with those identified by bulk cell H3K4me3 ChIP-seq. In each scChIC-seq library (100, 300, 1,000 and 3,000 cells), we computed the precision, which is the fraction of these enriched regions that are overlapped with that identified by bulk cell H3K4me3 ChIP-seq.

i. A histogram showing the number of enriched regions identified by scChIC-seq libraries of using 100, 300, 1,000 and 3,000 cells NIH-3T3 cells.



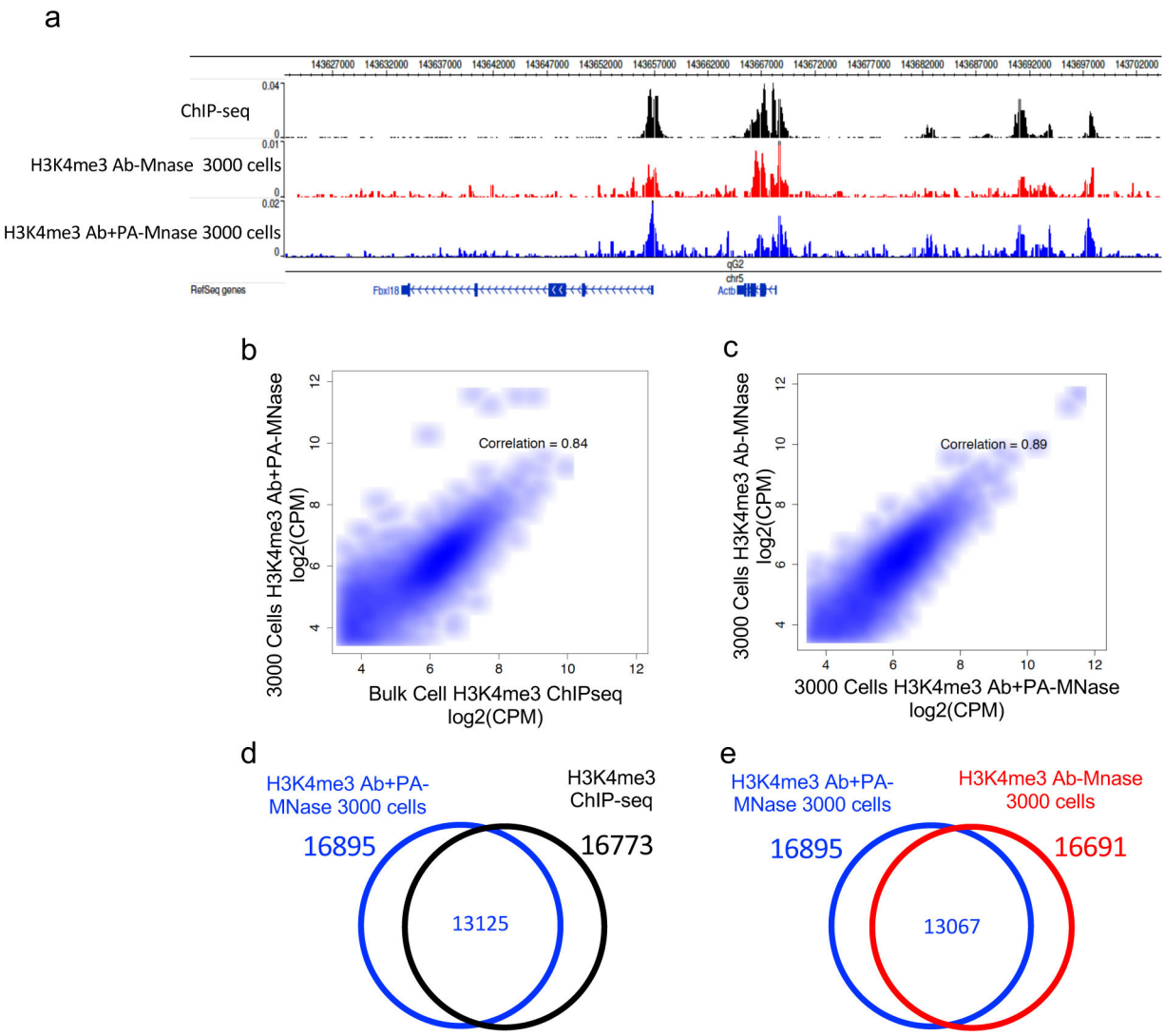
Extended data Figure 3.

TSS profile plots for the H3K4me3 profiles detected by scChIC-seq using H3K4me3 Ab-MNase conjugate (red) and IgG Ab-MNase (blue) with different washing conditions A) 200 mM NaCl, B) 300 mM NaCl, C) 400 mM NaCl, and D) RIPA buffer + 150 mM NaCl.



Extended data Figure 4.

TSS profile plots for the H3K4me3 profiles detected by scChIC-seq using H3K4me3 Ab + PA-MNase conjugate (red) and IgG Ab + PA-MNase (blue) with different washing conditions A) 200 mM NaCl, B) 300 mM NaCl, C) 400 mM NaCl, and D) RIPA buffer + 150 mM NaCl.



Extended data Figure 5.

Comparison between the H3K4me3 profiles obtained by scChIC-seq assays using H3K4me3-MNase or H3K4me3 Ab + PA-MNase.

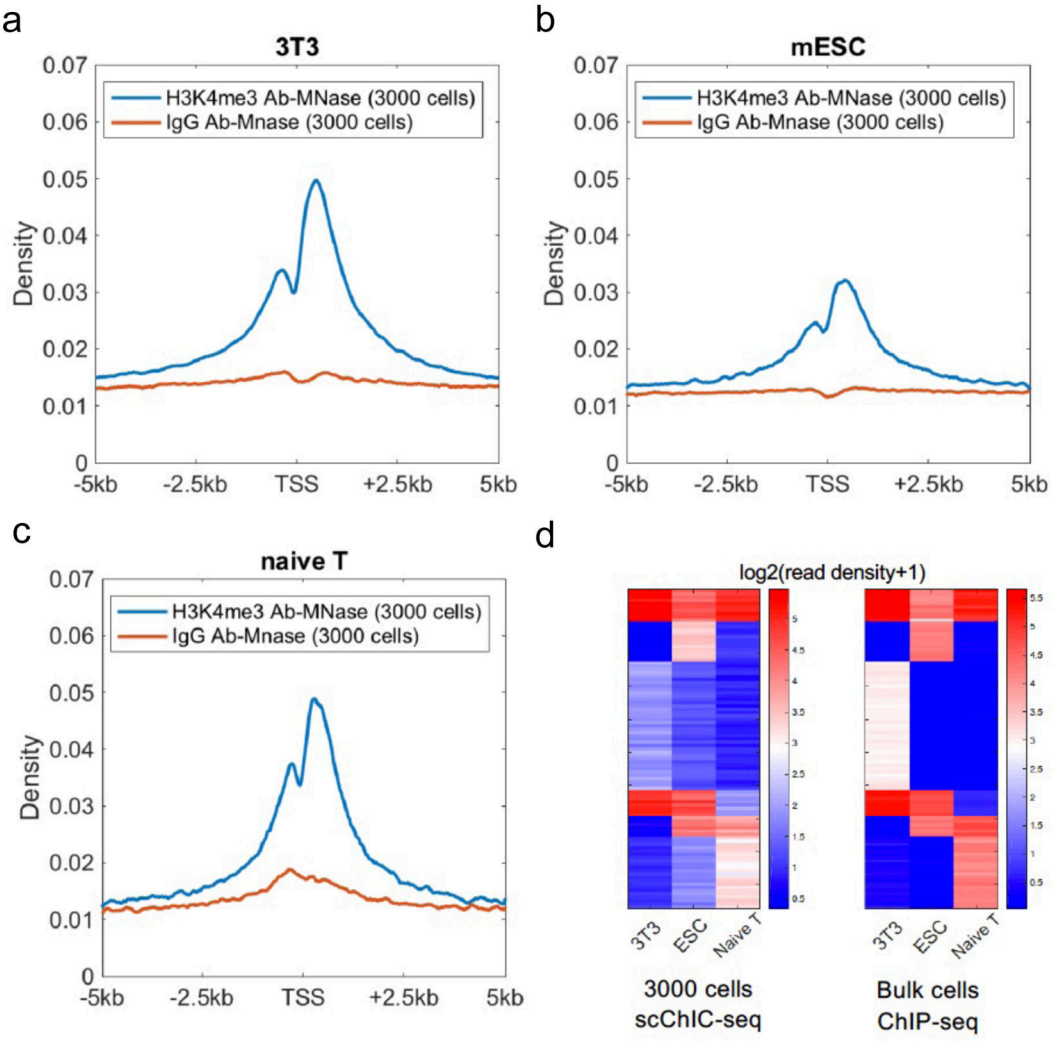
a. A genome browser snapshot showing the H3K4me3 profiles identified by ChIP-seq using bulk cell data (black), scChIC-seq using H3K4me3 Ab-MNase (blue) and scChIC-seq using H3K4me3 Ab + PA-MNase (red).

b. A scatter plot of the H3K4me3 read density of ChIP-seq versus that of scChIC-seq using H3K4me3 Ab + PA-MNase.

c. A scatter plot of the H3K4me3 read density of scChIC-seq using H3K4me3 Ab + PA-MNase versus that of scChIC-seq using H3K4me3 Ab-MNase.

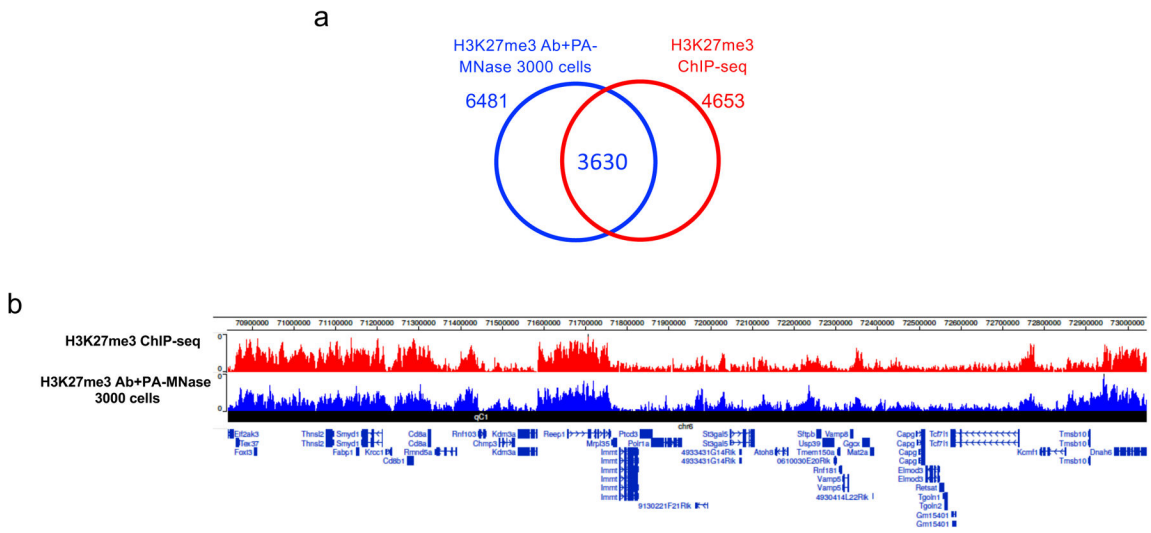
d. A Venn diagram showing the overlap of the enriched regions of H3K4me3 profiles measured by ChIP-seq using bulk cells (black) and by scChIC-seq using H3K4me3 Ab + PA-MNase and 3,000 cells (blue).

e. A Venn diagram showing the overlap of the enriched regions of H3K4me3 profiles measured by scChIC-seq using H3K4me3 Ab + PA-MNase and 3,000 cells (blue) and by scChIC-seq using H3K4me3 Ab-MNase and 3,000 cells (red).



Extended data Figure 6.

TSS profile plots of the H3K4me3 profiles around TSS for a) 3T3 cells, b) mouse ESC cells and c) mouse naïve CD4 T cells. In each cell type, the H3K4me3 TSS profiles (blue) are compared to the control IgG (red). d) Two heatmaps showing the clusters of the H3K4me3 enriched regions measured by ChIP-seq using bulk cells (right panel) and scChIC using 3,000 cells (left panel).

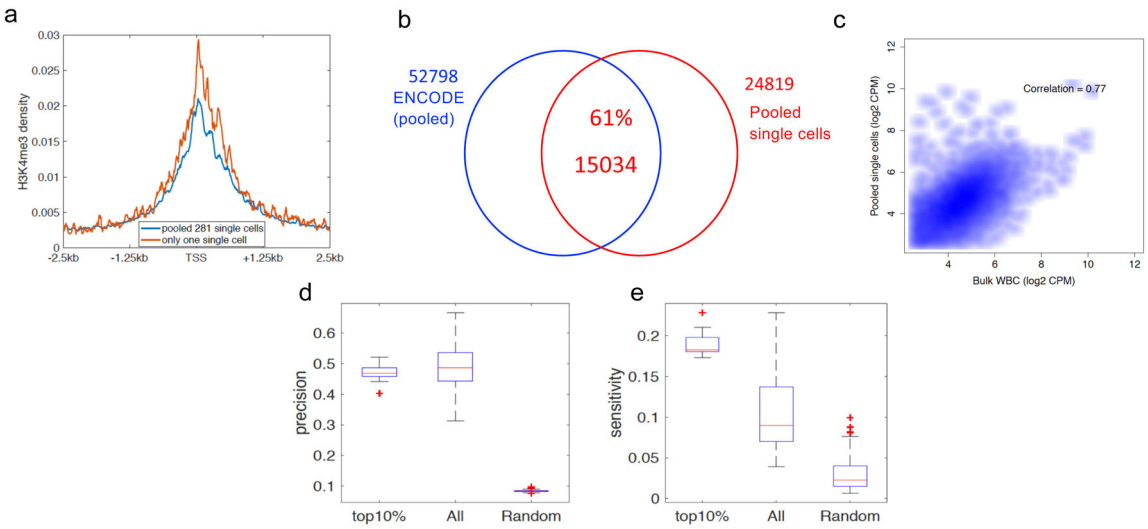


Extended data Figure 7.

Application of scChIC-seq to detect the profiles of H3K27me3 profile.

a. A Venn diagram showing the overlap of the enriched regions of H3K27me3 profiles measured by ChIP-Seq using bulk cells (red) and by scChIC-seq using 3,000 cells (blue).

b. A genome browser snapshot showing profiles of H3K27me3 and Brd4 detected by scChIC and ChIP-seq. Genome browser snapshots showing the H3K27me3 profiles detected by ChIP-seq using bulk cells (top panel in red), by scChIC-seq using 3,000 cells (second panel in blue).



Extended Data Figure 8.

Application of the scChIC-seq to profiling H3K4me3 in single human white blood cells

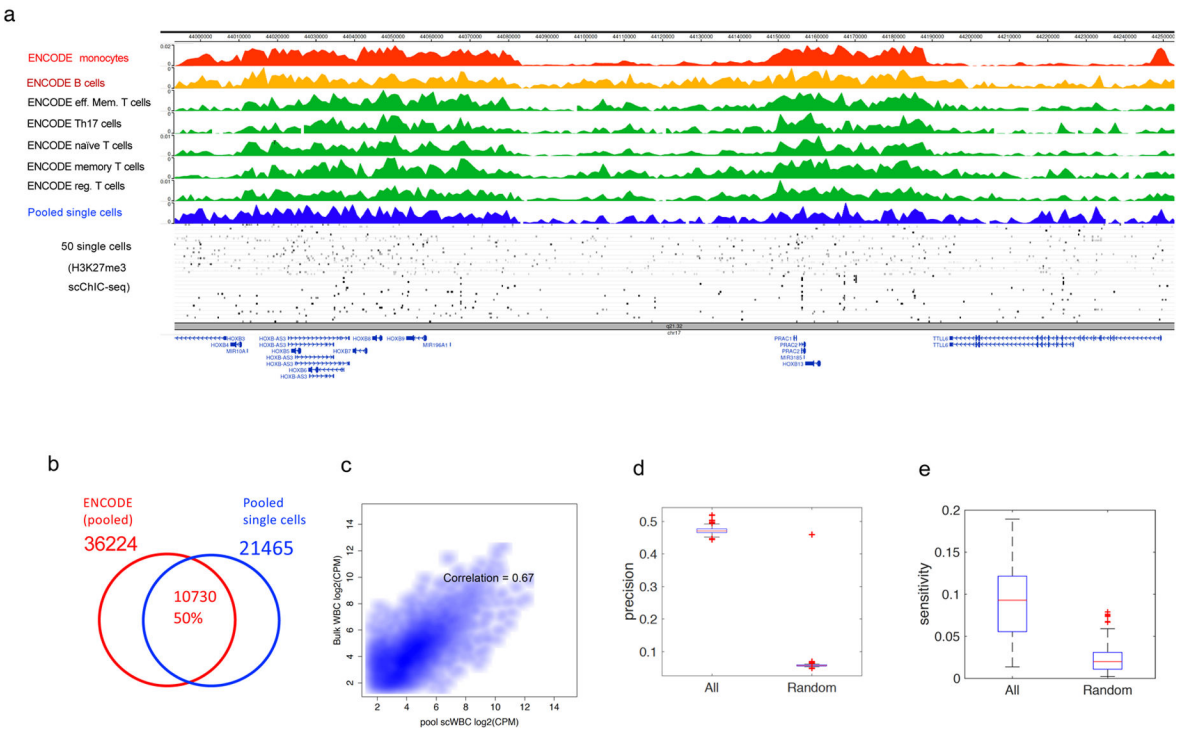
a. A TSS profile plot showing the H3K4me3 profile around TSS for a single cell (red) and for the aggregation of 281 pooled single cells.

b. A Venn diagram for comparison between the identified enriched regions from the data by bulk cell ChIP-seq and the pooled 281 single cells by scChIC-seq.

c. A scatter plot showing the correlation between the ChIP-seq and pooled top 40 single cell ChIC-seq data on the 52,798 combined H3K4me3 peaks for human white blood cells. The top 40 single cells are selected based on precision. The Pearson correlation between the ChIP-seq and pooled 281 single cell ChIC-seq data is 0.66.

d. A boxplot showing the precision from the top 10% single cells (about 48%) and all 242 single cells (48%). They are also compared to the random case, in which reads are randomly positioned in each cell. Precision is defined by the fraction of reads in single cells that are within the enriched regions identified using bulk cell ChIP-seq data. On each box, the central mark indicated the median. The bottom and top edges of the box indicated the 25th and 75th percentiles, respectively. A more detailed explanations of the boxplot could be found in Supplemental Methods.

e. A boxplot showing the sensitivity from the top 10% single cells (about 18%) and all 242 single cells (10%). They are also compared to the random case, in which reads are randomly positioned in each cell. Sensitivity is defined by the fraction of enriched regions identified using bulk cell ChIP-seq data that have single cell reads. On each box, the central mark indicated the median. The bottom and top edges of the box indicated the 25th and 75th percentiles, respectively. A more detailed explanations of the boxplot could be found in Supplemental Methods.



Extended data Figure 9.

Application of scChIC-seq to profiling H3K27me3 in single cells

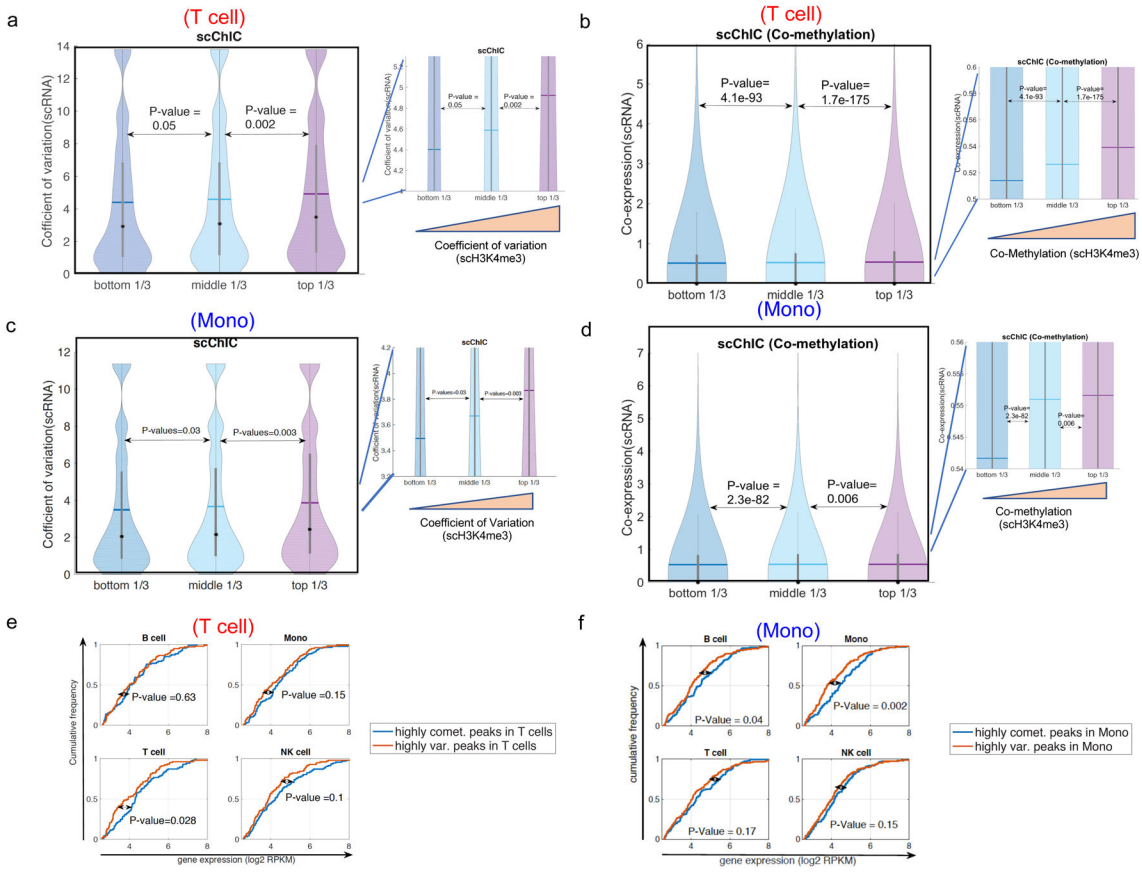
a. Genome browser snapshots showing the H3K27me3 profiles from bulk cell H3K27me3 ChIP-Seq data, from the pooled 106 single-cell ChIC-seq data and from 50 individual cells.

b. A Venn diagram showing the overlap between the identified enriched regions from the bulk cell H3K27me3 ChIP-seq data and the pooled 106 single cell scChIC-seq data.

c. A scatter plot showing the correlation between the bulk cell H3K27me3 ChIP-seq and pooled 84 single cell ChIC-seq data.

d. A boxplot showing the precision from all single cells (47%). They are also compared to the random case, in which reads are randomly positioned in each cell. Precision is defined by the fraction of reads in single cells that are within the enriched regions identified using bulk cell ChIP-seq data.

e. A boxplot showing the sensitivity for all single cells (9.5%). They are also compared to the random case, in which reads are randomly positioned in each cell. Sensitivity is defined by the fraction of enriched regions identified using bulk cell ChIP-seq data that have single cell reads.



Extended data Figure 10.

Correlation between H3K4me3 scChIC-seq data and scRNA-seq data

a. A violin plot showing the relationship between variability in H3K4me3 and heterogeneity in gene expression for T cell. The p-value is computed using Wilconxon ranksum test. The central mark indicated the median.

b. A violin plot showing the relationship between co-methylation and co-expression for T cell. The p-value is computed using Wilconxon ranksum test. The central mark indicated the median.

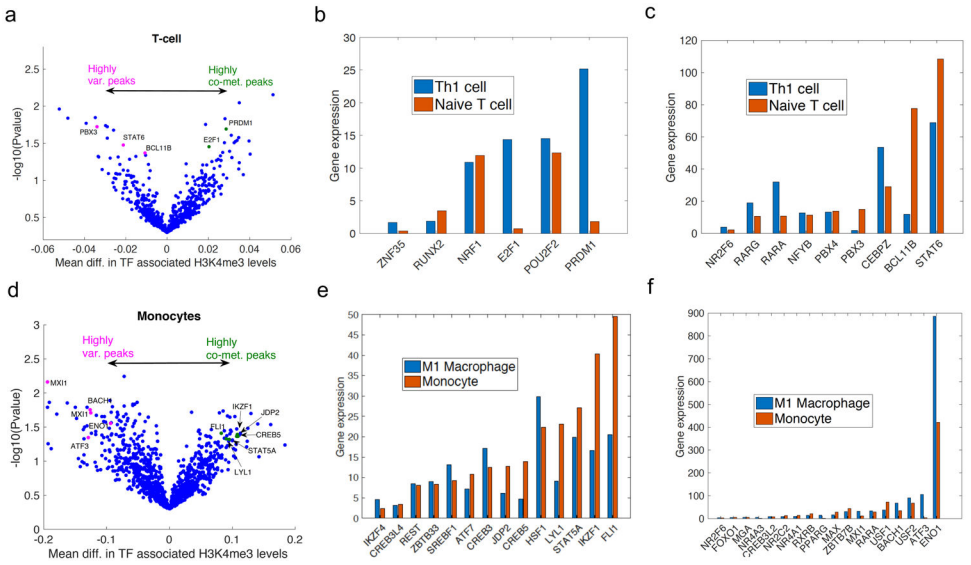
c. A violin plot showing the relationship between variability in H3K4me3 and heterogeneity in gene expression for monocytes.

d. A violin plot showing the relationship between co-methylation and co-expression for monocytes.

e. Two groups of annotated genes are selected from the highly co-methylated peaks (blue) and the highly variable peaks (red) using the T cells identified from the scChIC-seq data. Four cdf plots are plotted for the two groups of genes using the gene expression in B cells (top left), monocytes (top right), T cells (bottom left), and NK cells (bottom right). The p-values for the difference between the gene expression of the two groups are computed using Wilcoxon ranksum test.

f. Two groups of annotated genes are selected from the highly co-methylated peaks (blue) and the highly variable peaks (red) using the monocytes identified from the scChIC-se data. Four cdf plots are plotted for the two groups of genes using the gene expression in B cells

(top left), monocytes (top right), T cells (bottom left), and NK cells (bottom right). The p-values for the difference between the gene expression of the two groups are computed using Wilcoxon ranksum test.



Extended data Figure 11.

TFs enriched in the highly co-methylated and highly variable peaks are associated with cell-specific expression.

a. A volcano plot of the comparison between the enriched TFs that are specific to the highly co-methylated peaks and highly variable peaks in the T cells identified from the H3K4me3 scChIC-seq data. The y-axis is the negative log of p-value in the differential TFs analysis. X-axis is the difference between the mean value of two TF-bias corrected deviations vectors obtained from chromVAR[20].

b. The enriched TFs, which are specific to highly co-methylated peaks in T cells (Panel a), are preferentially expressed in Th1 cells. The bar plot show the gene expression levels (RPKM) of the enriched TFs in Th1 cells and naïve T cells.

c. The enriched TFs, which are specific to highly variable peaks in T cells (Panel a), are preferentially expressed in naïve T cells. The bar plot shows the gene expression levels (RPKM) of the enriched TFs in Th1 cells and naïve T cells.

d. A volcano plot of the comparison between the enriched TFs that are specific to the highly co-methylated peaks and highly variable peaks in monocytes identified from the H3K4me3 scChIC-seq data. The y-axis is the negative log of p-value in the differential TFs analysis. X-axis is the difference between the mean value of two TF-bias corrected deviations vectors obtained from chromVAR[20].

e. The enriched TFs, which are specific to highly co-methylated peaks in monocytes (Panel c), are preferentially expressed in monocytes. The bar plot shows the gene expression levels (RPKM) of the enriched TFs in monocytes and macrophages.

f. The enriched TFs, which are specific to highly variable peaks in monocytes (Panel c), are preferentially expressed in macrophages. The bar plot shows the gene expression levels (RPKM) of the enriched TFs in monocytes and macrophages.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the NHLBI DNA Sequencing Core facility, the NHLBI Flow Cytometry Core facility, and the NIH Biowulf High Performance Computing Systems for assistance with this work. This work was supported by the Division of Intramural Research of NHLBI, NIH.

References

1. Corces MR et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat Genet 48, 1193–1203, doi:10.1038/ng.3646 (2016). [PubMed: 27526324]
2. Jin W et al. Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples. Nature 528, 142–146, doi:10.1038/nature15740 (2015). [PubMed: 26605532]
3. Lai B et al. Principles of nucleosome organization revealed by single-cell micrococcal nuclease sequencing. Nature 562, 281–285, doi:10.1038/s41586-018-0567-3 (2018). [PubMed: 30258225]
4. Smallwood SA et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat Methods 11, 817–820, doi:10.1038/nmeth.3035 (2014). [PubMed: 25042786]
5. Cusanovich DA et al. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. Science 348, 910–914, doi:10.1126/science.aab1601 (2015). [PubMed: 25953818]
6. Buenrostro JD et al. Single-cell chromatin accessibility reveals principles of regulatory variation. Nature 523, 486–490, doi:10.1038/nature14590 (2015). [PubMed: 26083756]
7. Lay FD, Kelly TK & Jones PA Nucleosome Occupancy and Methylome Sequencing (NOMe-seq). Methods Mol Biol 1708, 267–284, doi:10.1007/978-1-4939-7481-8_14 (2018). [PubMed: 29224149]
8. Pott S Simultaneous measurement of chromatin accessibility, DNA methylation, and nucleosome phasing in single cells. Elife 6, doi:10.7554/eLife.23203 (2017).
9. Barski A et al. High-resolution profiling of histone methylations in the human genome. Cell 129, 823–837, doi:10.1016/j.cell.2007.05.009 (2007). [PubMed: 17512414]
10. Cao Z, Chen C, He B, Tan K & Lu C A microfluidic device for epigenomic profiling using 100 cells. Nat Methods 12, 959–962, doi:10.1038/nmeth.3488 (2015). [PubMed: 26214128]
11. Lara-Astiaso D et al. Immunogenetics. Chromatin state dynamics during blood formation. Science 345, 943–949, doi:10.1126/science.1256271 (2014). [PubMed: 25103404]
12. Brind'Amour J et al. An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. Nat Commun 6, 6033, doi:10.1038/ncomms7033 (2015). [PubMed: 25607992]
13. Blecher-Gonen R et al. High-throughput chromatin immunoprecipitation for genome-wide mapping of in vivo protein-DNA interactions and epigenomic states. Nat Protoc 8, 539–554, doi:10.1038/nprot.2013.023 (2013). [PubMed: 23429716]
14. Schmidl C, Rendeiro AF, Sheffield NC & Bock C ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. Nat Methods 12, 963–965, doi:10.1038/nmeth.3542 (2015). [PubMed: 26280331]
15. Adli M & Bernstein BE Whole-genome chromatin profiling from limited numbers of cells using nano-ChIP-seq. Nat Protoc 6, 1656–1668, doi:10.1038/nprot.2011.402 (2011). [PubMed: 21959244]
16. Rotem A et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. Nat Biotechnol 33, 1165–1172, doi:10.1038/nbt.3383 (2015). [PubMed: 26458175]
17. Schmid M, Durussel T & Laemmli UK ChIC and ChEC; genomic mapping of chromatin proteins. Mol Cell 16, 147–157, doi:10.1016/j.molcel.2004.09.007 (2004). [PubMed: 15469830]
18. Skene PJ & Henikoff S An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife 6, doi:10.7554/eLife.21856 (2017).
19. Kiselev VY et al. SC3: consensus clustering of single-cell RNA-seq data. Nat Methods 14, 483–486, doi:10.1038/nmeth.4236 (2017). [PubMed: 28346451]

20. Schep AN, Wu B, Buenrostro JD & Greenleaf WJ chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. Nat Methods 14, 975–978, doi:10.1038/nmeth.4401 (2017). [PubMed: 28825706]

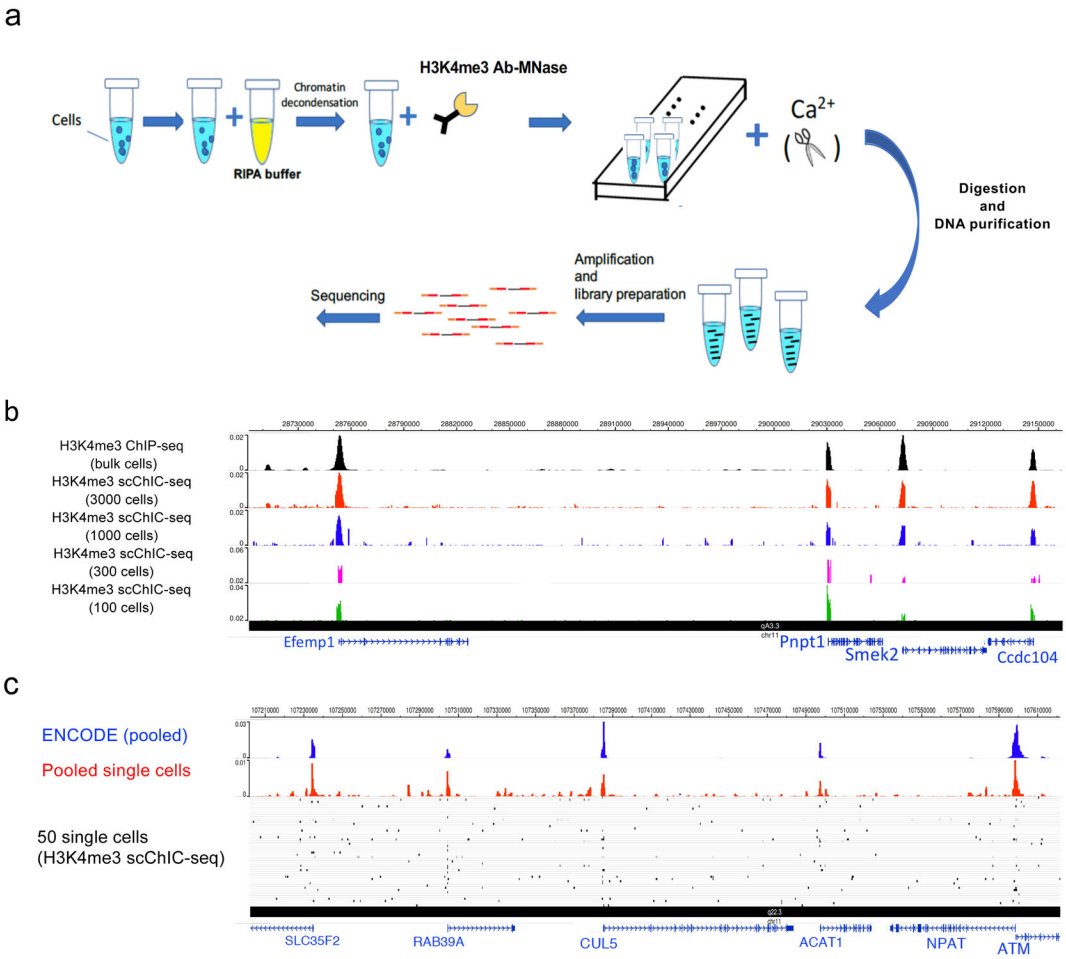


Figure 1.

scChIC-seq detects H3K4me3 profiles in a small number of cells and single cells

a. Experimental procedures of the scChIC-seq protocol. Following pre-treatment of fixed cells with RIPA buffer (with 0.2% SDS) for chromatin de-condensation, the Ab-MNase conjugates are added to allow Ab binding. Following washing of the unbounded and excess Ab-MNase conjugates in the nucleus, the MNase is activated by addition of calcium ion into the cell nucleus. Standard library preparation procedures are applied to the samples for library preparation and sequencing.

b. A genome browser snapshot showing panels of H3K4me3 profiles in NIH 3T3 cells obtained by scChIC-seq analysis using the direct conjugate between H3K4me3 Ab and MNase. The top panel in black refers to H3K4me3 profiles measured by ChIP-seq using bulk cells. H3K4me3 profiles measured by scChIC-seq using 100 (green), 300 (magenta), 1,000 (blue) and 3,000 (red) cells.

c. Genome browser snapshots showing the H3K4me3 profiles from pooled bulk cells ChIP-seq data (Supplemental Methods), pooled 281 single-cell ChIC-seq data and 50 individual cells. The ChIP-seq data sets are downloaded from ENCODE (top panel in blue). The H3K4me3 data from the pooled 281 single cells are displayed in the bottom panel.

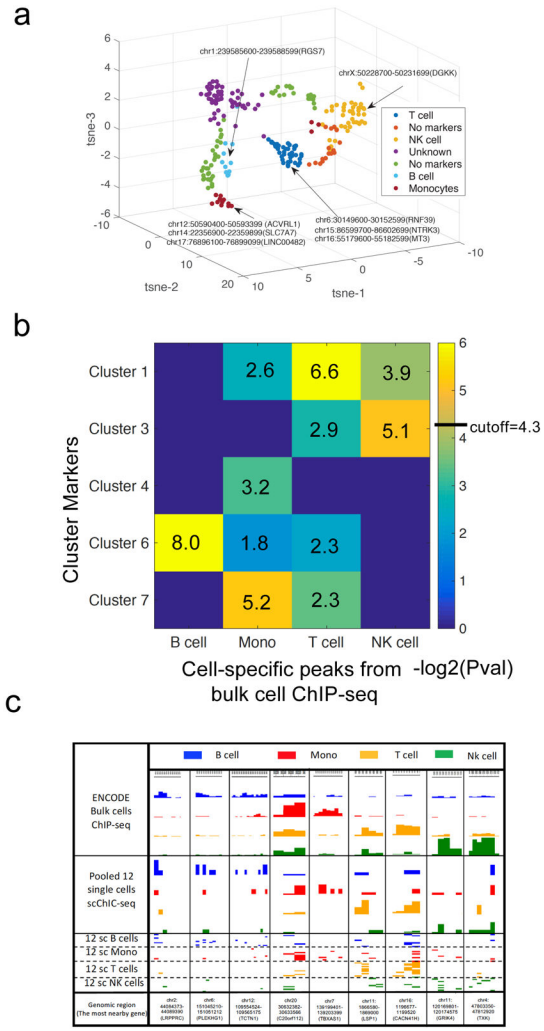


Figure 2.

Identification of sub-cell types in white blood cells based on clusters generated from single-cell H3K4me3 profiles

a. A t-SNE visualization of cells by applying the t-sne analysis on the consensus matrix obtained from the software SC3[19]. Cell type annotations of clusters are obtained via the analysis in **Fig. 2b**. Some of the cell type marker peaks which are also the marker peaks of clusters are explicitly shown next to the annotated clusters.

b. A heatmap showing negative log2 of p-values for the overlap comparison between the marker peaks of clusters and the cell type marker peaks.

c. Genome browser snapshots showing the H3K4me3 profiles from bulk cells ChIP-Seq data and single-cell scChIC-seq data. The ChIP-Seq data for B cells, monocytes, T cells and, NK cells are downloaded from ENCODE (top panel). The H3K4me3 scChIC-seq data from the 12 randomly selected single cells for each identified cell type are displayed in the bottom panels.