A microfluidic device for epigenomic profiling using 100 cells

Zhenning Cao[1,5], Changya Chen[2,5], Bing He[2], Kai Tan[2,3] & Chang Lu[4]

The sensitivity of chromatin immunoprecipitation (ChIP) assays poses a major obstacle for epigenomic studies of low-abundance cells. Here we present a microfluidics-based ChIP-seq protocol using as few as 100 cells via drastically improved collection of high-quality ChIP-enriched DNA. Using this technology, we uncovered many new enhancers and super enhancers in hematopoietic stem and progenitor cells from mouse fetal liver, suggesting that enhancer activity is highly dynamic during early hematopoiesis.

ChIP coupled with deep sequencing (ChIP-seq) is the technology of choice for examining *in vivo* genome-wide chromatin modifications[1]. A major limitation of conventional ChIP-seq protocols is the large number of cells required (~10^7 cells). Various strategies have been developed to improve the protocol over the past few years. Nano-ChIP-seq examines histone modification using 5,000 cells[2]. Single-tube linear DNA amplification profiles trimethylation of histone 3 lysine 4 (H3K4me3) using 10,000 cells and estrogen receptor α binding using 5,000 cells[3]. The use of a histone or mRNA carrier increases recovery of ChIP DNA and allows transcription factor ChIP-seq using 10,000 cells[4]. Indexing-first ChIP (iChIP) indexes and pools many chromatin samples for sequencing followed by demultiplexing based on sample-specific barcodes[5].

Several microfluidic ChIP protocols were reported recently for studying specific loci using ChIP coupled with quantitative PCR (qPCR)[6–8]. However, no effective strategies have been developed for high-efficiency collection of ChIP DNA and suppressed nonspecific adsorption at the same time. Meeting both requirements is critical for ChIP-seq when using a small number of cells.

The sensitivity of ChIP-seq is largely limited by the collection efficiency of ChIP DNA. A diploid mammalian cell contains 4–8 pg of DNA, yet previous ChIP-seq protocols can only obtain tens of picograms of DNA from 10,000 cells[2,3,5]. Here we introduce microfluidic oscillatory washing–based ChIP-seq (MOWChIP-seq), which allows genome-wide analysis of histone modifications using as few as 100 cells. The combined use of a packed bed of beads for ChIP and effective oscillatory washing for removing nonspecific adsorption is the key to extremely high yields of highly enriched DNA.

We used multilayer soft lithography to fabricate a poly (dimethylsiloxane) (PDMS) device, featuring a simple microfluidic chamber (~710 nl in volume) for high-efficiency ChIP. The microfluidic chamber has one inlet and one outlet, and the outlet has an on-chip pneumatic microvalve that can be partially closed by exerting a pressure at a port (refs. 7,9; **Fig. 1a** and **Supplementary Figs. 1** and **2**). Magnetic beads (~2.8 μm in diameter) coated with a ChIP antibody are flowed into the microfluidic chamber and form a packed bed while the pneumatic microvalve is partially closed. Sonicated chromatin fragments are then flowed through the chamber and adsorbed onto the bead surface. The gaps among the immunoprecipitation (IP) beads are smaller than 2 μm and facilitate rapid and high-efficiency adsorption of target chromatin fragments under the small diffusion length. The IP beads are then washed by oscillatory washing (**Supplementary Video 1**) in two different washing buffers to remove nonspecifically adsorbed chromatin fragments. Finally, the IP beads are flowed out of the chamber and collected for off-chip processing. The entire on-chip process takes ~1.5 h.

We found that the quality and amount of ChIP DNA were affected by several parameters. We optimized them by using MOWChIP-qPCR to examine fold enrichment at known positive and negative loci for H3K4me3 (primer sequences listed in **Supplementary Table 1**) in a human lymphoblastoid cell line, GM12878 (**Fig. 1b–d**). The fold enrichment reached a peak value at an intermediate bead amount; using too many beads increased nonspecific adsorption and trapping (**Fig. 1b**). We obtained the highest fold enrichment at an intermediate antibody concentration used for coating IP beads (**Fig. 1c**). This is likely due to insufficient antibody coating of the beads at low concentration, which lowers the amount of captured chromatin. On the other hand, an excessive amount of antibody on the beads may promote binding of nonspecific chromatin. The high-efficiency adsorption by the packed IP beads also led to increased nonspecific adsorption and physical trapping. Oscillatory washing was essential for the high quality of ChIP DNA (**Fig. 1d**). At the same time, excessive washing needs to be avoided to reduce DNA loss. Using optimized conditions that balanced both DNA yield and quality, we were able to obtain ~1.3 ng ChIP DNA from 10,000 cells (5.3% of the total chromatin) and ~180 pg from 1,000 cells (6.2% of the total chromatin) for H3K4me3 after DNA purification (**Supplementary Fig. 3**). This yield was almost two orders of magnitude higher than that reported in previous work[2] and

[1]Department of Biomedical Engineering and Mechanics, Virginia Tech, Blacksburg, Virginia, USA. [2]Interdisciplinary Graduate Program in Genetics, University of Iowa, Iowa City, Iowa, USA. [3]Department of Internal Medicine, University of Iowa, Iowa City, Iowa, USA. [4]Department of Chemical Engineering, Virginia Tech, Blacksburg, Virginia, USA. [5]These authors contributed equally to this work. Correspondence should be addressed to C.L. (changlu@vt.edu) or K.T. (kai-tan@uiowa.edu).

RECEIVED 4 FEBRUARY; ACCEPTED 4 JUNE; PUBLISHED ONLINE 27 JULY 2015; DOI:10.1038/NMETH.3488

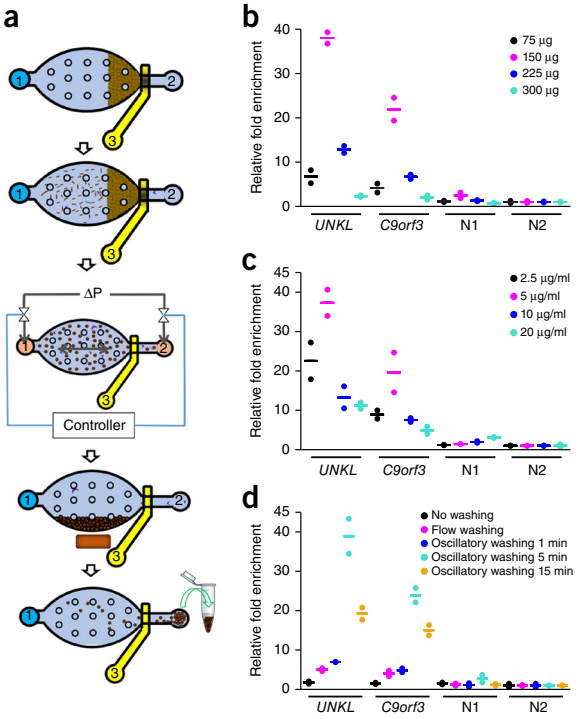

Figure 1 | Overview of the MOWChIP-seq protocol and its optimization. (**a**) Illustration for the five major steps. (i) Formation of a packed bed of IP beads; (ii) ChIP by flowing the chromatin fragments through the packed bed; (iii) oscillatory washing; (iv) removal of the unbound chromatin fragments and debris by flushing the chamber; (v) collection of the IP beads. The microfluidic chamber contains supporting pillars (shown as small circles) that prevent collapsing. The inlet and outlet are labeled 1 and 2, respectively; the outlet has a microvalve that can be partially closed by exerting a pressure at a port (3). (**b**–**d**) Optimization of the MOWChIP-seq protocol. Major parameters of the protocol were optimized by checking for IP fold enrichment of known positive (*UNKL* and *C9orf3*) and negative loci (N1 and N2). IP was done against H3K4me3 in GM12878 cells. All experiments were conducted in duplicate, and the horizontal lines represent the mean. Parameters optimized include the amount of beads in device chamber (**b**); concentration of antibody used for coating IP beads (**c**); and washing duration with each of the two washing buffers (**d**). The relative fold enrichment was normalized against that of N2. 1,000-cell samples were used in **b**–**d**. 150 µg IP beads were used in **c**,**d**. The antibody concentration for coating was 5 µg/ml for **b** and **d**. The duration of oscillatory washing was 5 min for **b**,**c**. Flow washing in **d** was implemented by flowing each washing buffer unidirectionally for 3 min under 1.5 µl/min.

within the range of the theoretical limit (2.2–7.8% of the genome is marked by H3K4me3 on the basis of Encyclopedia of DNA Elements (ENCODE) data[10]). We computed the fraction of reads in peaks[11]: the values were 35.6% and 21.6% for our 10,000- and 1,000-cell data, respectively. These were substantially higher than the 1% guideline recommended by ENCODE, suggesting low background in our recovered chromatin. As a result, we used ChIP DNA directly for sequencing library construction without preamplification.

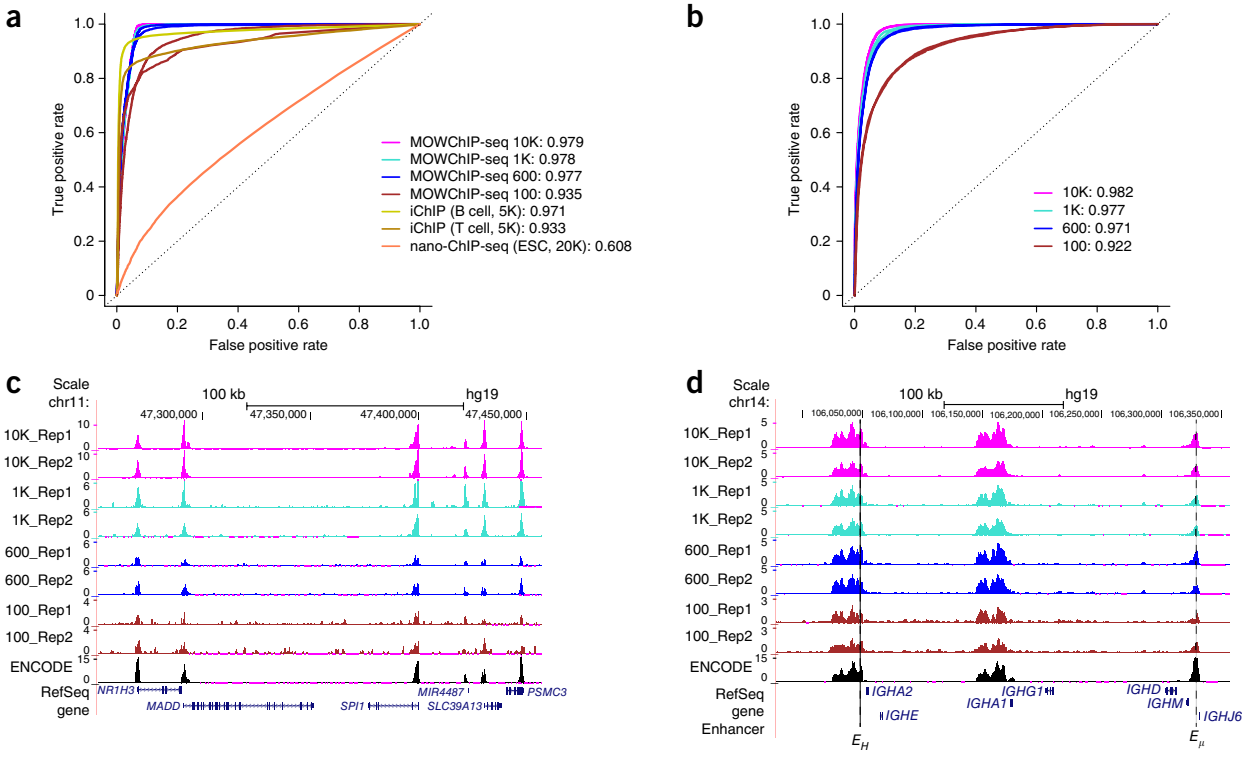

Figure 2 | MOWChIP-seq generates high-quality data using small quantities of chromatin. The performance of MOWChIP-seq was compared to that of two other methods: nano-ChIP-seq[2] and iChIP[5]. (**a**) Receiver operating characteristic (ROC) curves for H3K4me3 data. ROC curves were constructed by comparing the ChIP-seq data generated by various methods to published gold-standard data generated using conventional protocols with millions of cells. Values shown are average area under the ROC curve (AUC) of two replicate experiments (preceded by the number of cells used). ESC, embryonic stem cell. (**b**) ROC curves for H3K27ac data generated by MOWChIP-seq. (**c**) Normalized H3K4me3 MOWChIP-seq signals at the *SPI1* gene locus using data generated with various sample sizes. ENCODE data were generated using millions of cells and shown for comparison. (**d**) Normalized H3K27ac MOWChIP-seq signals at the immunoglobulin heavy chain locus. Known B cell enhancers are indicated at the bottom of the figure.

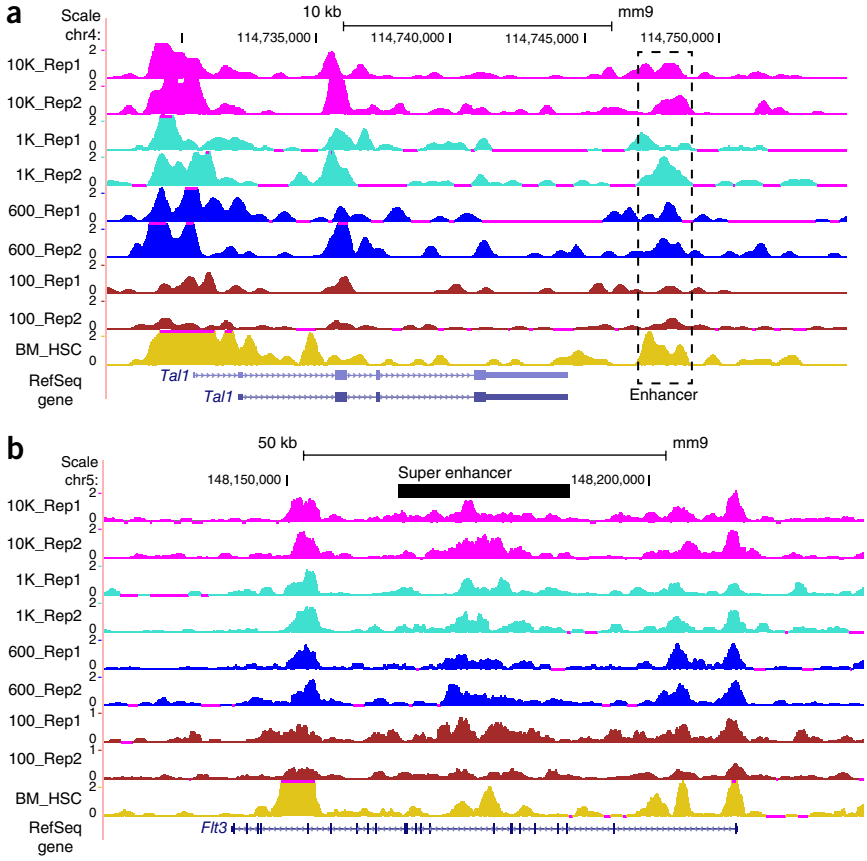

Figure 3 | Epigenomics-aided discovery of new enhancers and super enhancers in mouse fetal liver HSPCs. (**a**) Normalized H3K27ac MOWChIP-seq signals at the known *Tal* +19 enhancer. BM_HSC denotes data on bone marrow hematopoietic stem cells generated by Lara-Astiaso *et al.*[5]. (**b**) Normalized H3K27ac MOWChIP-seq signals at the super enhancer of the *Flt3* gene that plays a critical role in hematopoiesis.

We used MOWChIP-seq to profile H3K4me3 and H3K27 acetylation (H3K27ac) marks with various amounts of chromatin from GM12878 cells. We prepared sonicated chromatin using 10,000 cells and used this stock to generate aliquots containing chromatin samples of various sizes. For all four cell sample sizes, replicate experiments were highly correlated (average $r = 0.933$ and 0.894 for H3K4me3 and H3K27ac, respectively; **Supplementary Fig. 4**). Using published ChIP-seq data generated by conventional protocols as the gold standard, we compared the performance of MOWChIP-seq to that of two other methods, nano-ChIP-seq[2] and iChIP[5] (Online Methods and **Supplementary Table 2**). We then used the receiver operating characteristic (ROC) curve to quantify the data quality. For H3K4me3, MOWChIP-seq with 100–600 cells showed performance comparable to that of iChIP with 5,000 cells and superior to that of nano-ChIP-seq with 20,000 cells (**Fig. 2a**). For H3K27ac, MOWChIP-seq also produced data with excellent quality using as few as 100 cells (**Fig. 2b**). Normalized MOWChIP-seq signals (Online Methods) for H3K4me3 at the *SPI1* gene locus (**Fig. 2c**) and H3K27ac at the immunoglobulin heavy chain locus (**Fig. 2d**) showed consistency among samples of various sizes. As expected, the data quality decreased with decreasing number of cells. Nevertheless, all data had good quality and reproducibility that enabled analysis of important genome-wide features (**Supplementary Table 3**).

We applied MOWChIP-seq to study the epigenome of hematopoietic stem and progenitor cells (HSPCs) isolated from mouse fetal liver (FL). Previously, histone modifications had been mapped in hematopoietic stem cells from adult bone marrow (BM) but not from any earlier stage, such as FL. We mapped H3K4me3 and H3K27ac using purified FL HSPCs (**Supplementary Fig. 5**). Similarly to in **Figure 2**, chromatin samples of various sizes were divided up from a stock generated by sonicating 10,000 cells. For all four cell sample sizes, replicate experiments were highly correlated (average $r = 0.864$ and 0.881 for H3K4me3 and H3K27ac, respectively; **Supplementary Fig. 6**). We also computed the correlation of our data with three published data sets on BM HSPC data. Although our data showed lower correlation with published data (average $r = 0.677$ and 0.745 for H3K4me3 and H3K27ac, respectively; **Supplementary Fig. 6**), the correlations are still reasonable because the HSPCs being compared, though highly related, are not identical. Histone modification signals of the promoter regions were also correlated with gene expression levels (r ranged from 0.47 to 0.60; **Supplementary Fig. 7**). Taken together, these results suggest that our FL HSPC data are of high quality.

Using MOWChIP-seq data, we predicted active enhancers in FL HSPCs using the signature of H3K4me3^{low} + H3K27ac^{high}. In total, 4,446 enhancers were shared among all four data sets, which we used as the final set of enhancers in this study (**Supplementary Fig. 8**). We identified many known transcriptional enhancers that are active in FL HSPCs, such as the *Tal1* +19 enhancer[12] (**Fig. 3a**), *Erg* +85 enhancer and *Runx1* +24 enhancer[13,14] (**Supplementary Fig. 9**). DNA motif analysis revealed that the set of enhancers were enriched for binding motifs of 45 transcription factors, including many well-known hematopoietic transcription factors such as ERG, ETV6, FLI1, PU.1 and RUNX1 (**Supplementary Table 4**). By comparing our enhancer set to the enhancer catalog covering 16 blood cell types[5], we found that 58% (2,561) of enhancers identified in this study were unique to FL HSPCs (**Supplementary Table 5**), suggesting that enhancer activity is highly dynamic during early hematopoiesis.

'Super enhancer' is a newly discovered class of enhancers that are typically much longer than single enhancers[15]. They play a critical role in regulating genes that determine lineage identity. Using our epigenomic data, we discovered 131 super enhancers in FL HSPCs (**Supplementary Fig. 10** and **Supplementary Table 6**). Target genes of predicted super enhancers were enriched for genes involved in hematopoiesis ($P = 6.5 \times 10^{-3}$, hypergeometric test; **Supplementary Table 7**). Example target genes included many known regulators of hematopoiesis, such as *Erg*, *Etv6*, *Fli1*, *Flt3*, *Runx1* and *Spi1*. For example, *Flt3* plays an important role in hematopoiesis, especially for FL HSPCs[16] (**Fig. 3b**).

In order to use MOWChIP-seq directly with 100–600 cells, we created a special protocol for sonicating a tiny number of cells (Online Methods). Our modified cross-linking and sonication procedures generated the desired chromatin size distribution for sequencing after library preparation (**Supplementary Fig. 11**). We generated additional MOWChIP-seq data using the modified protocol with starting cell numbers of 100 and 600. Data generated using the two protocols (i.e., samples from stock chromatin vs. samples starting directly with 100 or 600 cells) were highly correlated (average $r = 0.843$ over two histone marks; **Supplementary Fig. 12**). Data generated using the two protocols also had similar quality as measured by ROC curves (**Fig. 2** and **Supplementary Fig. 13**).

Our microfluidic technology is fundamentally different from other high-sensitivity ChIP technologies that rely on superior amplification[2,3,17] and indexing-pooling[5] schemes, and thus may potentially complement other methods. Our technology paves the way for epigenomic studies involving extremely low number of cells.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. NCBI Gene Expression Omnibus: MOWChIP-seq data are deposited under accession number GSE65516.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank the Genomics Research Laboratory of Virginia Bioinformatics Institute and Genomics Division of Iowa Institute of Human Genetics for providing sequencing service. We thank L. Van Tol and University of Iowa Institute for Clinical and Translational Science for providing computing support. This work was supported by US National Institutes of Health grants EB017855 (K.T. and C.L.), CA174577 (C.L.), EB017235 (C.L.), HG006130 (K.T.) and GM104369 (K.T.) and a seed grant from Virginia Tech Institute for Critical Technology and Applied Science (C.L.).

AUTHOR CONTRIBUTIONS

C.L. designed the microfluidic device. C.L., K.T. and Z.C. developed the MOWChIP-seq technology. Z.C. generated the MOWChIP-qPCR and MOWChIP-seq data. K.T., C.C. and B.H. designed the biological experiments, isolated the primary cells from mice and conducted data analysis. All authors wrote the manuscript together.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Park, P.J. *Nat. Rev. Genet.* **10**, 669–680 (2009).
2. Adli, M., Zhu, J. & Bernstein, B.E. *Nat. Methods* **7**, 615–618 (2010).
3. Shankaranarayanan, P. *et al. Nat. Methods* **8**, 565–567 (2011).
4. Zwart, W. *et al. BMC Genomics* **14**, 232 (2013).
5. Lara-Astiaso, D. *et al. Science* **345**, 943–949 (2014).
6. Wu, A.R. *et al. Lab Chip* **9**, 1365–1370 (2009).
7. Geng, T. *et al. Lab Chip* **11**, 2842–2848 (2011).
8. Wu, A.R. *et al. Lab Chip* **12**, 2190–2198 (2012).
9. Hong, J.W., Studer, V., Hang, G., Anderson, W.F. & Quake, S.R. *Nat. Biotechnol.* **22**, 435–439 (2004).
10. Kundaje, A. *et al. Nature* **518**, 317–330 (2015).
11. Landt, S.G. *et al. Genome Res.* **22**, 1813–1831 (2012).
12. Göttgens, B. *et al. EMBO J.* **21**, 3039–3050 (2002).
13. Thoms, J.A. *et al. Blood* **117**, 7079–7089 (2011).
14. Nottingham, W.T. *et al. Blood* **110**, 4188–4197 (2007).
15. Hnisz, D. *et al. Cell* **155**, 934–947 (2013).
16. Buza-Vidas, N. *et al. Blood* **113**, 3453–3460 (2009).
17. Jakobsen, J.S. *et al. BMC Genomics* **16**, 46 (2015).

ONLINE METHODS

Fabrication of the microfluidic ChIP device. The microfluidic chip consisted of a microfluidic chamber, connecting channels and a micromechanical valve (**Fig. 1** and **Supplementary Fig. 2**). The microfluidic chamber had an elliptic shape with a major axis of 6 mm, a minor axis of 3 mm and a depth of 40 µm. Micropillars were positioned inside the microfluidic chamber to prevent collapse. The on-chip micromechanical valve, which allowed partial closure, was employed to stop magnetic IP beads while allowing liquid flow.

The microfluidic device was fabricated out of poly (dimethylsiloxane) (PDMS) using multilayer soft lithography with minor modifications[7]. Briefly, two photomasks were generated with the microscale patterns designed using FreeHand MX (Macromedia) and printed on high-resolution (5,080 d.p.i.) transparencies. The patterns in the photomasks were replicated onto two masters (i.e., silicon wafers with photoresist patterns) for the control layer (~50 µm thick, SU-8 2025, Microchem) and the fluidic layer (~40 µm thick, SU-8 2025) with the photoresist spun on a 3-inch silicon wafer (978, University Wafer). Prepolymer PDMS (General Electric silicone RTV 615, MG chemicals) with a mass ratio of A:B = 5:1 was poured onto the fluidic layer master in a Petri dish to generate an ~5-mm-thick fluidic layer. PDMS at a mass ratio of A:B = 20:1 was spun onto the control layer master at 1,100 r.p.m. for 35 s, resulting in the thin PDMS control layer (~108 µm thick). Both layers of PDMS were partially cured at 80 °C for 30 min. The fluidic layer was then peeled off the master. The fluidic layer feature was aligned with and bonded to that of the control layer from the top. The two-layer PDMS structure was baked at 80 °C for 60 min, peeled off from the control layer master, and punched to produce the inlet and the outlet. The two-layer PDMS and a precleaned glass slide were treated with oxygen plasma cleaner (PDC-32G, Harrick Plasma) and immediately brought into contact against each other to form closed channels and chamber. Finally, the assembled chip was baked at 80 °C for 1 h to strengthen the bonding between PDMS and glass. Glass slides were cleaned in a basic solution (H_2O: 27% NH_4OH: 30% H_2O_2 = 5:1:1, volumetric ratio) at 75 °C for 2 h and then rinsed with ultrapure water and thoroughly blown dry.

Setup of the microfluidic device. The microfluidic chip was mounted on an inverted microscope (IX 71, Olympus), and the operation was monitored by a charge-coupled device (CCD) camera (ORCA-285, Hamamatsu) attached to the port of the microscope. Prior to experiments, the control channel was prefilled with water to prevent bubble formation in the fluidic channel. The reagents were introduced into the inlet via perfluoroalkoxy alkane (PFA) high-purity tubing (1622L, ID: 0.02 in. and OD: 0.0625 in., IDEX Health & Science) with the flow driven by a syringe pump (Fusion 400, Chemyx). The on-chip micromechanical valve was actuated by a solenoid valve (18801003-12V, ASCO Scientific) and a pressure source (either a gas cylinder or a compressed air outlet). A data acquisition card (NI SCB-68, National Instruments) and a LabVIEW (LabVIEW 2012, National Instruments) program were employed to control the switching of the solenoid valve. The applied pressure (35–40 p.s.i.) in the PDMS control channel deformed the thin PDMS membrane between the fluidic and control channels and closed the fluidic channel partially to stop beads while allowing liquid to flow. During oscillatory washing, the inlet and outlet of the microfluidic chamber were attached to two solenoid valves via PFA tubing, and the pressure pulses were applied via the two solenoid valves under the automation by the data acquisition card and the LabVIEW program.

Preparation of sonicated chromatin. *10,000-cell samples.* 10,000-cell samples were centrifuged at 1,600*g* for 5 min at room temperature in a swing bucket centrifuge with soft deceleration. Cells were then washed twice with 1.0 ml 1× PBS (14190-144, Sigma-Aldrich) at room temperature by centrifugation and resuspension. Cells were cross-linked for 5 min with 1 ml of 1% freshly prepared formaldehyde (28906, Thermo Scientific). Cross-linking was terminated by adding 0.05 ml of 2.5 M glycine (R000333, Covaris) and shaking for 5 min at room temperature. Cross-linked cells were pelleted and washed with precooled PBS buffer and resuspended in 130 µl of the sonication buffer (Covaris, 10 mM Tris-HCl, pH 8.1, 1 mM EDTA, 0.1% SDS and 1× protease inhibitor cocktail (R000306, Covaris)). Cross-linked cells were sonicated with a Covaris E220 sonicator for 14 min with 5% duty cycle, 105 peak incident power and 200 cycles per burst. The sonicated lysate was centrifuged at 14,000*g* for 10 min at 4 °C. Sonicated chromatin in the supernatant was transferred to a new 1.5-ml LoBind Eppendorf tube (17014013, Denville) for MOWChIP-seq. From this stock chromatin preparation, samples equivalent to 1,000, 600 and 100 cells were divided into aliquots and diluted to give a final volume of 50 µl for MOWChIP-seq. 10% of the sample was used as the input. After this procedure, we typically obtained ~2.7 pg DNA per cell from the pre-ChIP chromatin samples. DNA was extracted using the IPure kit from Diagenode (C03010012). DNA concentration was measured using a Qubit 2.0 fluorometer with dsDNA HS Assay kit (Q32851, Life Technologies).

100- or 600-cell samples. The procedure was different for preparing sonicated chromatin from 100 or 600 cells directly. Cells were counted with a hematocytometer and then 100 or 600 cells were transferred to a 1.5-ml LoBind Eppendorf tube containing 10 µl of 10% FBS in PBS. Cells were then cross-linked for 5 min at room temperature by adding 0.625 µl of 16% formaldehyde to yield a final concentration of 1%. Cross-linking was quenched by adding 1.25 µl of 2.5 M glycine for 5 min at room temperature. The cross-linked sample was then diluted using 120 µl Covaris sonication buffer (to give a total volume of 130 µl) and sonicated with a Covaris E220 sonicator for 8 min with 5% duty, 105 peak incident power and 200 cycles per burst in a Covaris microtube (520045, Covaris). The sonicated lysate was centrifuged at 14,000*g* for 10 min at 4 °C. Sonicated chromatin in the supernatant was transferred to a new 1.5-ml LoBind Eppendorf tube for MOWChIP-seq. After this procedure, we typically obtained ~3.8 pg DNA per cell from the pre-ChIP chromatin samples. This per-cell yield was substantially higher than that obtained using the above procedure because we replaced washing of cross-linked cells (involving centrifugation and resuspension) with dilution by the sonication buffer to minimize chromatin loss.

Preparation of immunoprecipitation (IP) beads. Superparamagnetic Dynabeads Protein A (2.8 µm, 30 mg/ml, 10001D, Invitrogen) were used for immunoprecipitation. 150 µg (5 µl of the original suspension) of beads were washed twice with freshly prepared IP buffer (20 mM Tris-HCl, pH 8.0, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% (w/v) sodium deoxycholate, 0.1% SDS, 1% (v/v)

Triton X-100) and resuspended in 150 µl IP buffer containing antibody. Beads were gently mixed with the antibody at 4 °C on a rotator mixer at 24 r.p.m. for 2 h. Antibody-coated beads were washed twice with the IP buffer and resuspended in 5 µl IP buffer. We optimized the antibody concentration for the bead coating step on the basis of our ChIP-qPCR results. The optimal antibody concentration for MOWChIP-seq with anti-H3K4me3 antibody (07-473, Millipore) and anti H3K27ac antibody (ab4729, Abcam) was 3.3 µg/ml for ~100–600 cells, 5 µg/ml for 1,000 cells and 6.6 µg/ml for 10,000 cells. These conditions were equivalent to using 495, 750 and 990 ng of antibody in the preparation of 150 µg IP beads.

MOWChIP. The MOWChIP process involved several steps (**Fig. 1a** and **Supplementary Fig. 2**). The microfluidic device was first rinsed with the IP buffer for conditioning. The antibody-coated magnetic IP beads were then loaded into the microfluidic chamber via the combined effects of pressure-driven flow (provided by the syringe pump) and magnetic force generated by a cylindrical permanent magnet (NdFeB, D48-N52, 0.25 inch dia. and 0.5 inch thick, K&J Magnetics). The on-chip micromechanical valve was partially closed, and the IP beads were packed against the valve to form a packed bed. After the loading of the IP beads (~150 µg under optimal conditions), the IP buffer (with freshly added 1 mM PMSF (78830-1G, Sigma-Aldrich) and 1% protease inhibitor cocktail (P8340, Sigma-Aldrich)) containing sonicated chromatin fragments (with a total volume of either 50 or 130 µl) was flowed through the packed bed of IP beads at a flow rate of 1.5 or 3.5 µl/min, respectively. Under these flow rates, the immunoprecipitation step was finished after around 40 min.

After ChIP, a low-salt washing buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% (v/v) Triton X-100) was flowed into the microfluidic chamber. Oscillatory washing was conducted (for 5 min unless otherwise noted) to remove nonspecifically adsorbed or physically trapped materials from the bead surface. We prefilled the tubing with 10 µl washing buffer at each end of the microfluidic chamber and kept the on-chip valve open. Pressure pulses (each at 3 p.s.i., with a pulse width of 0.5 s and an interval of 0.5 s between two pulses) were applied alternatingly at either end of the microfluidic chamber. The duration and frequency of the pressure pulses were set in a LabVIEW program and implemented via the regulation of the two solenoid valves by the data acquisition card (**Supplementary Fig. 1b**). After the oscillatory movement, the IP beads were retained by the NdFeB magnet on one side of the chamber while the unbound chromatin fragments and other debris/waste were flushed out of the microfluidic chamber by a clean washing buffer flow at 2 µl/min. The process of oscillatory washing was repeated once using a high-salt washing buffer (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% (v/v) Triton X-100). Finally, the IP beads were flowed out of the microfluidic chamber under a flow rate of 50 µl/min and collected into a 1.5-ml LoBind Eppendorf tube containing 100 µl IP buffer. The optimal duration for washing was 5 min for both washing buffers.

Extraction of ChIP DNA and input DNA. Chromatin samples (either ChIP or input chromatin) were processed by IPure kit (C03010012, Diagenode) to extract DNA. Purified DNA was dissolved in 10 µl DNase-free water and used directly for ChIP-qPCR or for sequencing library construction. DNA concentrations were measured using a Qubit 2.0 fluorometer with dsDNA HS Assay kit (Q32851, Life Technologies).

Construction of sequencing libraries. Sequencing libraries were prepared using ThruPLEX-FD kit (Rubicon Genomics). This kit reduces the assay time and the risk of contamination by using a single tube and eliminating intermediate purification steps. The process involved template preparation, library synthesis and library amplification. Adaptor-based PCR amplification (98 °C for 20 s, 72 °C for 50 s for each cycle) was used during library amplification. We used 11 cycles for input DNA, ~12 or 13 cycles for ChIP DNA from 10,000 cells, and ~14–17 cycles for ChIP DNA from 1,000 or fewer cells. The libraries were purified using Ampure XP beads (A63880, Beckman Coulter). Library fragment size was determined using high-sensitivity DNA analysis kit (5067-4626, Agilent) on an Agilent 2200 TapeStation. The Kapa library quantification kit (KK4809, Kapa Biosystems) was used to determine effective library concentrations. The final concentrations of libraries submitted for sequencing were ~2 nM. The libraries were sequenced on an Illumina HiSeq 2500 with single-end 50-nt reads. Typically, 15–20 million reads were generated per library.

Cell culture. GM12878 cells were obtained from Coriell Institute for Medical Research. Species of origin of the cell line was confirmed by PCR targeting the gene encoding glucose-6-phosphate dehydrogenase. The donor subject has a single base-pair (G-to-A) transition at nucleotide 681 in exon 5 of the *CYP2C19* gene (*CYP2C19*2*), which creates an aberrant splice site. Donor origin of the cell line was confirmed using PCR against the point mutation. The cell line was tested for mycoplasma contamination using ABI MycoSEQ mycoplasma detection assay (Applied Biosystems). Cells were propagated in RPMI 1640 (11875-093, Gibco) plus 15% FBS (26140-079, Gibco), 100 U penicillin (15140-122, Gibco) and 100 mg/ml streptomycin (15140-122, Gibco) at 37 °C in a humidified incubator containing 5% CO_2. Cells were subcultured every 2 d to maintain them in exponential growth phase.

Mouse strain, embryo dissection and cell sorting by FACS. The University of Iowa Office of the Institutional Animal Care and Use Committee review board approved these studies. Wild-type C57BL/6 (Stock No. 000664) and B6129S6F1 (Stock No. 101043) mice were purchased from the Jackson Laboratory. To obtain embryonic day 14.5 (E14.5) fetal liver (FL), we mated B6129SF1 females with C57BL/6 males (6–9 weeks old) late in the afternoon, and females were checked the following morning for the presence of a vaginal plug, which was designated as E0.5. FLs were dissected from E14.5 embryos. Single-cell suspensions were prepared by dissociating mechanically and expelling the cells through a 40-µm nylon filter (352340, Falcon), followed by red blood cell lysis (ACK Lysing Buffer, 10-548E, Lonza). Cells were resuspended in 1 ml staining buffer (2% FBS in PBS) per 10^8 cells. Fluorescence-activated cell sorting (FACS) was performed as previously described[18,19] with a few modifications. To remove nonspecific binding, we added anti-mouse CD16/CD32 (Fc Block, 101302, BioLegend) to single-cell suspensions and incubated for 10 min at 4 °C. Next, cells were stained with a cocktail

doi:10.1038/nmeth.3488

of antibodies against lineage markers (Ly-6G/Ly-6C (108417), CD45R/B220 (103225), CD3 ε (100321), TER-119b (116215), CD4 (100529), CD8a (100723), CD19 (115521)), Kit (17-1171-83) and Sca-1 (12-5981-83). Lineage antibodies were purchased from BioLegend. Kit and Sca-1 antibodies were purchased from eBioscience. Stained samples were first subjected to yield sort for Lineage$^-$Sca-1$^+$Kit$^+$ (LSK) cells and collected into a 12 × 75-mm polystyrene tube containing 500 µl 1× IMDM (12440-053, Gibco) + 20% FBS. Collected cells were then subjected to purity sort using the same gating strategy and sorted into a 1.5-ml DNA LoBind tube containing 0.8 ml 1× IMDM + 50% FBS. On average, ~10,000 FL LSK cells could be obtained per mouse embryo.

MOWChIP-qPCR data analysis. Real-time PCR was done using iQ Sybr Green Supermix (1708882, Bio-Rad) on an CFX96 real-time PCR machine (Bio-Rad) with C1000Tm thermal cycler base. All PCR assays were performed using the following thermal cycling profile: 95 °C for 10 min followed by 40 cycles of 95 °C for 15 s, 58 °C for 40 s, 72 °C for 30 s. Primer concentrations were 400 nM. All primers were ordered from Integrated DNA Technologies. The results were represented as relative fold enrichment, which is the ratio of percent input between a positive locus and a negative locus. Percent input was computed using the following equation:

$$\text{Percent input} = 100 \times 2\text{^}\left(\text{Ct}_{\text{input}} - \frac{\log(\text{DF})}{\log 2} - \text{Ct}_{\text{IP}}\right)$$

where Ct_{input} and Ct_{IP} are the threshold cycle values of input and ChIP DNA, respectively; dilution factor (DF) is defined as (sample volume of input + sample volume of IP)/(sample volume of input).

MOWChIP-seq reads mapping and normalization. Sequencing reads were mapped to the mouse genome (mm9) and human genome (hg19) using Bowtie2 (v2.2.2)[20] with default parameter settings. Uniquely mapped reads from both ChIP and input samples were used to compute a normalized signal for each 100-nt bin across the genome. Normalized signal is defined as follows:

$$\text{Normalized signals} = \text{IP}\left(\frac{\text{Reads in each bin}}{\text{Total uniquely mapped reads}} \times 1{,}000{,}000\right) - \text{Input}\left(\frac{\text{Reads in each bin}}{\text{Total uniquely mapped reads}} \times 1{,}000{,}000\right)$$

Peak-calling of MOWChIP-seq data. Only uniquely mapped reads were used for peak-calling. Two peak-callers were used with the following parameter settings: MACS (P value < 10^{-5})[21] and SPP (z score > 4)[22] with other parameters set at default values. The final set of high-confidence peaks comprised those that were called by both methods.

Construction of receiver operating characteristic (ROC) curves. Using ROC curves, we compared the performance of MOWChIP-seq to that of two state-of-the-art methods nano-ChIP-seq[2] and iChip[5]. We focused on promoter regions (defined as 2,000 bp upstream and 500 bp downstream of a transcription start site (TSS)). We obtained published ChIP-seq data generated using conventional protocol with a large sample size (typically 10 million cells per sample) as the gold standard. The gold-standard true positives were defined as the set of high-confidence promoter peaks identified as described in the peak-calling section. The set of promoter regions that did not overlap with any peaks was defined as the gold-standard negative set. Using the gold-standard sets, the following quantities were defined to compute the ROC curve: true positives (TPs), peaks that were supported by the gold-standard positive set; false positives (FPs), peaks that were not supported by the gold-standard positive set; false negatives (FNs), gold-standard positives that were not called peaks in an experiment; true negatives (TNs), peaks that were not called in an experiment and were in the gold-standard negative set. True positive rate (TPR) was defined as TP/(TP + FN) and false positive rate (FPR) was defined as FP/(FP + TN). ROC curves were generated by computing TPR and FPR values on prediction sets obtained by varying the peak-calling threshold.

The gold-standard data sets used for constructing the ROC curves for MOWChIP-seq, nano-ChIP-seq and iChIP are summarized in **Supplementary Table 2**. Briefly, we generated H3K4me3 and H3K27ac data using GM12878 cells. The corresponding gold-standard data were generated by the ENCODE Consortium. The authors of nano-ChIP-seq generated H3K4me3 data using mouse ESCs. The corresponding gold-standard data were from Marson *et al.*[23] and Goren *et al.*[24]. The authors of iChIP generated H3K4me3 data using mouse CD4 T and B cells. The corresponding gold-standard data were from Wei *et al.*[25] and Heinz *et al.*[26].

Correlation analysis of MOWChIP-seq data with other published ChIP-seq data sets. To evaluate the quality of our FL HSPC data, we selected four published data sets of H3K4me3 and H3K27ac using BM LSK[2], BM LT-HSC[5], B cell[5] and macrophage[5]. For a given histone modification, normalized ChIP-seq signals in all promoter regions in the genome were extracted. Promoter regions were defined as ±2 kb around TSSs. TSS annotation was based on RefSeq. Averaged signals across the promoter region were used. Promoter regions with zero signals in both data sets were excluded for computing Pearson correlation coefficient.

Prediction of enhancers and super enhancers using epigenomic data. We used 'H3K4me3^{low} + H3K27ac^{high}' to define enhancers. Specifically, enhancers were predicted using the CSI-ANN algorithm[27] and normalized H3K4me3 and H3K27ac MOWChIP-seq signals across the genome. The coordinates of the predicted enhancers and H3K27ac MOWChIP-seq data were then used as the input to predict super enhancers using the ROSE software by the Young lab (http://bitbucket.org/young_computation/rose/). We set the parameters to allow enhancers within 15,000 bp to be stitched together. In addition, we excluded the constituent enhancers located within ±2,000 bp from annotated TSSs.

Transcription factor motif enrichment analysis. We compiled a set of 1,207 transcription factor binding motifs from three major public databases—JASPAR[28], UniPROBE[29] and Transfac[30]—and motifs of ten hematopoietic transcription factors[31]. We used the program CentriMo[32] to identify over-represented motifs in a given set of enhancer sequences. Default parameters of CentriMo were used.

Gene Ontology (GO) term enrichment analysis of super enhancer targets. Genes closest to the super enhancers were used as their targets. Database for Annotation, Visualization, and Integrated Discovery (DAVID)[33] was used for GO analysis of the target genes. Nominal *P* values were corrected for multiple testing using the method by Benjamini and Hochberg[34]. GO terms with a corrected *P* value of 0.05 were regarded as significant.

Assumptions of statistical tests. All statistical tests were performed using large sample sizes, and underlying distribution assumptions were met. Sample sizes were reported in figure legends. All reported *P* values were corrected for multiple testing.

18. McKinney-Freeman, S. *et al. Cell Stem Cell* **11**, 701–714 (2012).
19. McKinney-Freeman, S.L. *et al. Blood* **114**, 268–278 (2009).
20. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. *Genome Biol.* **10**, R25 (2009).
21. Zhang, Y. *et al. Genome Biol.* **9**, R137 (2008).
22. Kharchenko, P.V., Tolstorukov, M.Y. & Park, P.J. *Nat. Biotechnol.* **26**, 1351–1359 (2008).
23. Marson, A. *et al. Cell* **134**, 521–533 (2008).
24. Goren, A. *et al. Nat. Methods* **7**, 47–49 (2010).
25. Wei, G. *et al. Immunity* **30**, 155–167 (2009).
26. Heinz, S. *et al. Mol. Cell* **38**, 576–589 (2010).
27. Firpi, H.A., Ucar, D. & Tan, K. *Bioinformatics* **26**, 1579–1586 (2010).
28. Portales-Casamar, E. *et al. Nucleic Acids Res.* **38**, D105–D110 (2010).
29. Newburger, D.E. & Bulyk, M.L. *Nucleic Acids Res.* **37**, D77–D82 (2009).
30. Matys, V. *et al. Nucleic Acids Res.* **34**, D108–D110 (2006).
31. Wilson, N.K. *et al. Cell Stem Cell* **7**, 532–544 (2010).
32. Whitington, T., Frith, M.C., Johnson, J. & Bailey, T.L. *Nucleic Acids Res.* **39**, e98 (2011).
33. Huang, D.W., Sherman, B.T. & Lempicki, R.A. *Nat. Protoc.* **4**, 44–57 (2009).
34. Benjamini, Y. & Hochberg, Y. *J. R. Stat. Soc. Series B Stat. Methodol.* **57**, 289–300 (1995).

doi:10.1038/nmeth.3488