nature protocols
https://doi.org/10.1038/s41596-024-01058-z

Protocol

Check for updates

Multiplexed chromatin immunoprecipitation sequencing for quantitative study of histone modifications and chromatin factors

Banushree Kumar[1,2,3,5], Carmen Navarro[1,2,3,5], Philip Yuk Kwong Yung[1,2,5], Jing Lyu[1,2], Angelo Salazar Mantero[1,2], Anna-Maria Katsori[1,2], Hannah Schwämmle[1,2], Marcel Martin[4] & Simon J. Elsässer[1,2] ✉

Abstract

ChIP–seq is a widely used technique for studying histone post-translational modifications and DNA-binding proteins. DNA fragments associated with a specific protein or histone modification epitope are captured by using antibodies, sequenced and mapped to a reference genome. Albeit versatile and popular, performing many parallel ChIP–seq experiments to compare different conditions, replicates and epitopes is laborious, is prone to experimental variation and does not allow quantitative comparisons unless adequate spike-in chromatin is included. We present a detailed protocol for performing and analyzing a multiplexed quantitative chromatin immunoprecipitation-sequencing experiment (MINUTE-ChIP), in which multiple samples are profiled against multiple epitopes in a single workflow. Multiplexing not only dramatically increases the throughput of ChIP–seq experiments (e.g., profiling 12 samples against multiple histone modifications or DNA-binding proteins in a single experiment), but also enables accurate quantitative comparisons. The protocol consists of four parts: sample preparation (i.e., lysis, chromatin fragmentation and barcoding of native or formaldehyde-fixed material), pooling and splitting of the barcoded chromatin into parallel immunoprecipitation reactions, preparation of next-generation sequencing libraries from input and immunoprecipitated DNA and data analysis using our dedicated analysis pipeline. This pipeline autonomously generates quantitatively scaled ChIP–seq tracks for downstream analysis and visualization, alongside necessary quality control indicators. The entire workflow requires basic knowledge in molecular biology and bioinformatics and can be completed in 1 week. MINUTE-ChIP empowers biologists to perform every ChIP–seq experiment with an appropriate number of replicates and control conditions, delivering more statistically robust, exquisitely quantitative and biologically meaningful results.

Key points

- MINUTE-ChIP is a sensitive multiplexed ChIP method that quantitatively profiles relative differences in genome-wide epigenetic patterns across multiple samples and conditions. In a MINUTE-ChIP workflow, samples are ligated in a one-pot reaction to UMI-containing adapters before pooling and splitting into multiple parallel immunoprecipitation reactions.

- The multiplexed MINUTE-ChIP sequencing files can be readily analyzed by using a comprehensive analysis pipeline that generates output files for direct, quantitative comparison across conditions.

Key references

Kumar, B. & Elsässer, S. J. *Cell Rep*. **28**, 3274–3284.e5 (2019): https://doi.org/10.1016/j.celrep.2019.08.046

Kumar, B. et al. *Nat. Cell Biol*. **24**, 845–857 (2022): https://doi.org/10.1038/s41556-022-00916-w

Weigert, R. et al. *Nat. Cell Biol*. **25**, 579–591 (2023): https://doi.org/10.1038/s41556-023-01114-y

A full list of affiliations appears at the end of the paper. ✉e-mail: simon.elsasser@scilifelab.se

Introduction

Development of the protocol and comparison with other methods

ChIP–seq

Chromatin immunoprecipitation with high-throughput sequencing (ChIP–seq)[1–4] has been the method of choice for profiling histone post-translational modifications (PTMs) and protein binding to the genome. The method relies on the immunoprecipitation (antibody-based capture) of a target protein or a modified histone epitope from fragmented cellular chromatin and subsequent massive parallel sequencing of the captured DNA fragments. Naturally, the quality of the sequencing data largely relies on the use of specific and efficient antibodies. Inherent to immunoprecipitations, a certain fraction of the sequenced reads represents nonspecifically bound and pulled-down molecules, which are randomly sampled from the input fragments and produce a 'technical background'. Moreover, histone PTMs or transcription factor binding may also occur outside of preferred target regions, contributing to the 'biological background'. Although a myriad of different ChIP–seq protocols are practiced to date, they generally group into two alternative basic workflows: 'native' ChIP–seq[2,5,6] and cross-linking ChIP–seq. Native ChIP–seq relies on mild lysis of cells and enzymatic fragmentation of chromatin, most commonly with micrococcal nuclease (MNase), yielding mono- to oligonucleosome-sized fragments. In contrast, cross-linking ChIP–seq relies on formaldehyde fixation of cells or tissue followed by harsh lysis and mechanical fragmentation of the chromatin by sonication[1,3]. Formaldehyde fixation is used to cross-link DNA-associated proteins and complexes to DNA, enabling profiling of transcription factors and other chromatin-associated proteins that do not remain stably bound to DNA under the mild lysis and immunoprecipitation conditions used for native ChIP. Cross-linking can, however, lead to epitope occlusion and amplification of unspecific binding events. Because DNA-protein cross-links by formaldehyde are thought to be rare events, the technical background in the form of nonspecifically bound DNA chromatin fragments, in fact, dominates many cross-linking ChIP–seq datasets. In some instances, transcription factors such as CCCTC-binding factor (CTCF) have also been successfully mapped under native conditions in which salt concentrations were kept deliberately low[7].

ChIP–seq normalization

Mapping ChIP–seq reads to the respective reference genome yields the genome-wide 'landscape' of the underlying epitope, characterized by regions of higher mapped read density (typically referred to as 'enriched' regions or peaks) and lower read density (background regions). However, a biological interpretation of such ChIP–seq signals, especially when comparing different conditions, requires a strategy to normalize the samples to make the signals comparable. It is important to be aware that any normalization method operates on a set of fundamental assumptions about the sample and data composition. Whenever possible, these assumptions should be validated on an individual basis when choosing the normalization method.

Raw sequencing output yields an arbitrary number of total sequences for each library (dependent only on how much was loaded on a sequencer), and hence the most common strategy for normalizing ChIP–seq data is to correct for unequal dataset size and resulting mapped read depth by using either reads per kilobase per million reads (RPKM) or 1× genome coverage (also known as reads per genomic content (RPGC)) normalization. We prefer the latter method because it allows an intuitive interpretation of the units; for example, on a tracks' Y axis, a value of 5 RPGC indicates a fivefold higher read density than the genome average, thus a fivefold 'enrichment'. Conversely, values <1 indicate that the read density is below the genome average. One underlying assumption for choosing RPKM or RPGC normalization is that there is no genome-wide loss or gain of the protein occupancy or PTM under investigation because sequencing depth normalization would remove such global differences. However, if the technical background contributes to most of the reads of a dataset and such background is assumed constant between samples, then the use of RPKM or RPGC normalization methods may be justifiable even if minor global changes (e.g., in transcription factor binding) are expected.

Several normalization approaches have been proposed with the aim of preserving true biological variation. Assuming the existence of genomic regions that contain no specific signal (i.e., pure technical background), it has been proposed to exclusively base normalization on such low-signal regions. How such low-signal regions are defined varies greatly between strategies, which sometimes take into account a certain signal threshold or a control (e.g., IgG isotype) immunoprecipitation[8–10]. Such regions must, by definition, be entirely devoid of 'biological background' contributed by low levels of the probed histone modification or weak transcription factor binding. Hence, if applied without explicit validation of the 'constant background' hypothesis, background-normalization strategies risk introducing, rather than removing, bias.

Quantitative ChIP–seq

In the past decade, new developments have recognized the importance of using spike-ins that provide an invariable internal standard for any genome-wide sequencing experiment[11], allowing the capturing of quantitative differences between samples. ChIP with reference exogenous genome (ChIP-Rx) was the first of a series of methods that mix cells or chromatin from a different species into the sample as an internal standard. The resulting datasets can then be scaled relative to the constant signal from the spike-in[12] by using software packages that facilitate the processing of ChIP-Rx data[13,14]. Because an exact and constant proportion of spike-in across all samples is difficult to achieve (because of the chance of measurement error in cell count or protein determination, as well as pipetting error), input sequencing is often performed to empirically determine the actual proportion of spike-in chromatin in each sample before immunoprecipitation, and samples are then scaled accordingly[13]. Recombinant nucleosomes have also been used as a spike-in reference for histone PTMs, enabling relative and even absolute scaling of the signal[15]. Because epitopes from the sample and spike-in are allowed to compete for the available antibody binding sites in a biochemical equilibrium reaction, the scaling of the samples to be compared is proportional to the global PTM or DNA-bound protein fraction in each sample, respectively. Methods that produce a measurement proportional to the true amount of the studied epitope are considered 'quantitative'; hence, ChIP-Rx is the first example of a quantitative ChIP–seq protocol. Limitations of ChIP-Rx include the fact that the epitope to be probed (e.g., the histone PTM or the transcription factor) must be present in the spike-in chromatin, requiring sometimes a spike-in of a closely related species (e.g., human with mouse and vice versa), which in turn can lead to ambiguities in assigning reads between the sample and spike-in genome. Furthermore, ChIP-Rx does not explicitly account for and correct for variations in technical background between samples (resulting, for example, from variable lysis, fragmentation or immunoprecipitation conditions), and comparisons within an experimental batch are more reliable than between batches.

Multiplexed ChIP–seq

Multiplexed ChIP–seq methods, such as indexing-first chromatin immunoprecipitation[16], multiplexed, indexed T7 ChIP–seq (Mint-ChIP)[17], barcoded high-throughput ChIP–seq (Bar-ChIP)[18], restriction enzyme–based labeling of chromatin in situ (RELACS)[19], multiplexed indexed unique molecule T7 amplification end-to-end sequencing (MINUTE-ChIP)[20], emerged on the same principles of mixing genomes, but in this case, multiple samples are mixed after barcoding the fragmented chromatin with a sample-specific DNA barcode. Advantages of these 'barcoding-first' ChIP–seq methods include that multiple conditions and replicates can be pooled into a single immunoprecipitation reaction and carried through all the subsequent steps of the protocol together into a single sequencing library, minimizing technical variation between samples. Furthermore, the pool can be split into multiple immunoprecipitation reactions, allowing to probe multiple epitopes from the same pool. As part of the analysis pipeline, sample-specific barcodes are sequenced along with the genomic fragments and used to demultiplex and assign reads to their sample of origin. Sequencing of the input pool yields an empirical determination of the sample mixing ratios, which can then be used to quantitatively scale the ChIP–seq data for each sample in the pool[17]. The resulting quantification is proportional to the true epitope density over a wide dynamic range, while the technical background is low because unspecific binding is distributed equally across all samples[20]. As a result, multiplexed quantitative ChIP provides an unbiased and assumption-free quantification across samples. However, a critical requirement of barcoding-first methods is the efficient and robust tagging of the fragmented chromatin with sample-specific barcodes, and challenges exist because this step must be performed in crude chromatin preparations and must not disrupt the native chromatin complexes.

Our MINUTE-ChIP[20] protocol presented here is based on Mint-ChIP[17] and uses enzymatic fragmentation of native chromatin with MNase, followed by a one-pot DNA end repair and ligation reaction to introduce double-stranded DNA adapters carrying the sample-specific barcode and T7 RNA polymerase promoter (Fig. 1a). We additionally introduced a unique molecular identifier (UMI) on the adapter, which we use for removing duplicated sequences. The T7 promoter enables linear amplification of fragments with a single adapter ligation event, providing a great advantage over PCR-based strategies used, for example, in Bar-ChIP and RELACS, which require adapters on both ends[16,18,19]. We have determined a typical ligation efficiency of crude chromatin to be in the range of a few percent of all fragments; hence, double-ligated fragments are rare[20]. Through a linear amplification strategy, full-length cDNA copies of the single-ligated chromatin fragments are produced: DNA fragments are transcribed to RNA in vitro from the T7 promoter. After purification of the RNA, we ligate an RNA adapter to the 3′ end of the T7 RNA polymerase products to act as a priming site for the subsequent reverse transcription (RT) reaction. Compared to random hexamer priming, as used in Mint-ChIP, the addition of a 3′ adapter ensures that the cDNA product contains the full length of the original chromatin fragment. This not only preserves the footprint information of the probed nucleosome/factor, but also enables a stringent size-selection of the cDNA product to reduce carryover of smaller adapter side products. A low-cycle PCR of the cDNA yields the final library for paired-end Illumina sequencing[20] (Fig. 1a).

Applications

We have applied MINUTE-ChIP to compare histone PTM landscapes of human[21] and mouse[20,22,23] cultured cells from up to eight different conditions by barcoding and pooling up to 12 samples (including replicates). From the same pool, we have probed up to six histone PTMs in parallel, yielding a total of 12 × 6 = 72 ChIP–seq tracks in a single experiment[24]. Because of its quantitative nature, MINUTE-ChIP has been proven particularly valuable for the comparison of epigenomic landscapes across conditions that have different global levels of histone PTMs[20–22,24]. Importantly, we and others have extensively validated that MINUTE-ChIP accurately captures global differences in histone PTMs across samples, by performing standard curve experiments[20] (Fig. 2b) and comparing MINUTE-ChIP quantitation to complementary methods such as western blot, immunofluorescence and mass spectrometry[20–22,24]. We have also demonstrated that native MINUTE-ChIP can quantitatively profile non-histone proteins such as RNA polymerase II[23]. By using the protocol described here with native or cross-linked material, MINUTE-ChIP can be adapted for the typical range of ChIP–seq applications, including histone PTMs and DNA- and chromatin-binding factors. Although not tested so far, we believe that MINUTE-ChIP can be adapted to primary cell types and animal or human tissue material, provided that the chromatin can be extracted and fragmented either by MNase digestion or sonication.

Advantages and limitations

MINUTE-ChIP allows users to evaluate multiple histone PTMs and/or chromatin factors across multiple samples in a single experiment. Even if only a few conditions are to be compared, the possibility of including three or more replicates per condition greatly increases the confidence with which quantitative comparisons can be made between conditions. MINUTE-ChIP is a time-efficient method because multiplexing reduces the number of reactions to be handled in parallel.

As a disadvantage, MINUTE-ChIP may require a higher sequencing volume than individual ChIP–seq libraries because adapter-only molecules contribute to the sequenced reads. This is particularly problematic for the library generated from the input pool, whereas excess free adapter molecules are washed away efficiently during the immunoprecipitation steps. By using

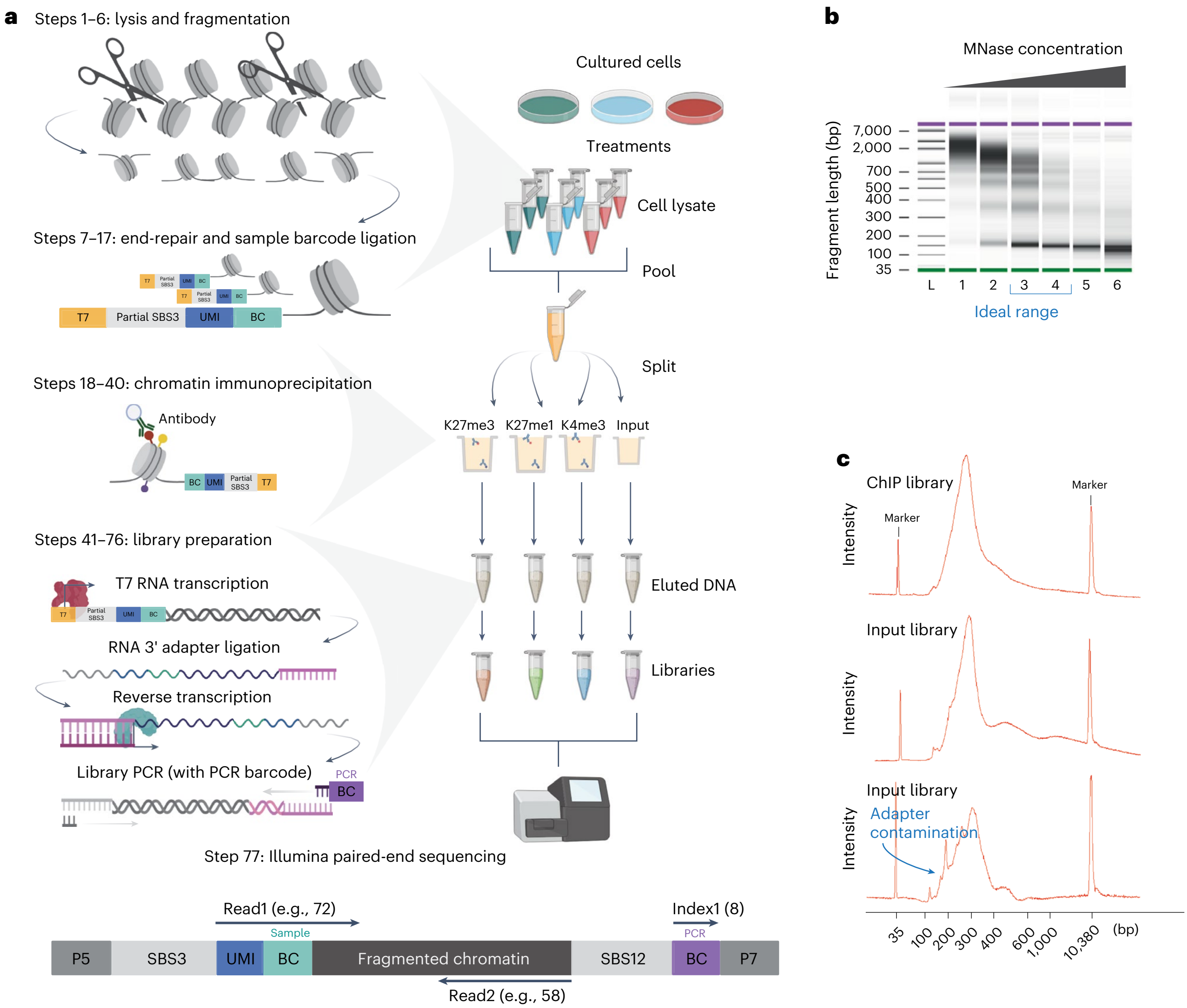

Fig. 1 | MINUTE-ChIP protocol workflow and quality control examples. **a**, Scheme of the MINUTE-ChIP workflow described in this protocol. Cell lysis and fragmentation of the chromatin is followed by a one-pot DNA end-repair and ligation reaction to introduce double-stranded DNA adapters carrying the sample-specific barcode, a unique molecular identifier (UMI) and T7 RNA polymerase promoter. The barcoded chromatin is then pooled into a single tube and split up for the desired chromatin immunoprecipitations. After DNA purification, libraries are generated via linear amplification and a final low-cycle PCR to introduce a PCR barcode. The libraries can be pooled and sequenced by using standard Illumina sequencing primers. The PCR barcode is read as Index Read 1 and used for demultiplexing on the Illumina system (bcl2fastq or BaseSpace), while the Sample barcode and UMI are read as part of Read1. Example cycle numbers are given for a 100-cycle NextSeq2000 kit. **b**, Example results from an MNase titration experiment on mouse embryonic stem (mES) cell. We recommend optimizing the MNase concentration for the cell type of interest to obtain mono- and dinucleosomal chromatin fragments as seen in lanes 3 and 4. Suitable starting concentrations are given in the Supplementary Methods. **c**, Bioanalyzer capillary electrophoresis trace of the final libraries. A peak centered at ~300 bp corresponds to a mono-nucleosomal insert (including UMI, sample barcode and adapters for Illumina sequencing) and should be present in both the ChIP and input libraries (top and middle). The bottom panel shows an input library contaminated with adapter monomers and dimers, which typically spike at 150 and 200 bp, respectively. BC, barcode; L, ladder Panel **a** created with BioRender.com.

appropriate size selection during DNA purification steps, the adapter amount in the input can usually be reduced below 30% in the input library and usually lies below 10% in ChIP libraries.

Two critical variables of a MINUTE-ChIP experiment are the quality of the antibody and the abundance of the epitope. Although traditional ChIP–seq experiments are sometimes performed with tens of millions of cells as input to overcome a low-efficiency antibody capture, the barcoding of such a large number of cells is not feasible. Hence, MINUTE-ChIP does not

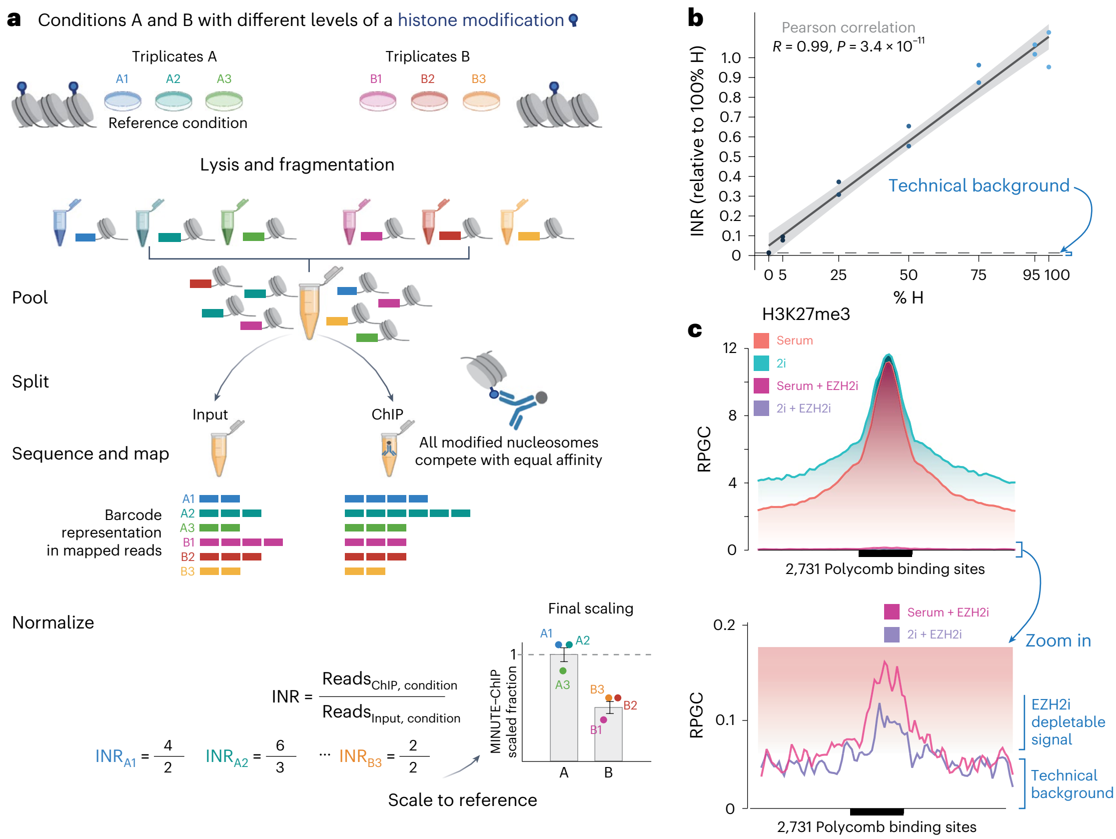

Fig. 2 | MINUTE-ChIP quantification principle and example data. a, Scheme explaining the principles of MINUTE-ChIP quantitation. Two conditions are to be compared, and three replicates are prepared from each condition. Chromatin is barcoded and pooled. The barcode distribution after ChIP in principle reflects the quantitative differences of the probed histone modification between the two conditions but must be normalized to the barcode distribution in the input. Barcodes are counted in the input and H3K27me3 ChIP libraries, and a ratio (input-normalized read count, INR) is calculated for each barcode, dividing ChIP by input reads. Finally, condition A is set as reference with a global coverage of 1, and condition B is scaled accordingly to 0.5. The bars indicate the INR of two groups of replicates; dashed line denote reference group INR; error bars indicate s.d. **b**, Example data from ref. 20 showing that the calculated INR correlates very well with the actual H3K27me3 abundances in each sample. mES cells were treated with EZH2 inhibitor EPZ-6438 (resulting in depletion of H3K27me3, 'L') or JMJD3/UTX inhibitor GSK-J4 (resulting in very high H3K27me3, 'H'). L and H cells were mixed in defined ratios, barcoded in two replicates and pooled. To quantify MINUTE-ChIP data after anti-H3K27me3 ChIP, the INR values are plotted against the mixing ratio. Linear regression (black line) with 95% confidence interval (gray ribbon) and coefficient of determination (R^2) is given. The replicates of the EPZ-6438 sample (0% H) have an INR of ~0.03, indicating that 97% of the signal can be depleted with EZH2i, and only maximally 3% of the signal can be attributed to technical background (dashed line, nonspecifically captured DNA molecules). Underlying data points are available as source data. Raw data are available in the GEO database with accession number GSE133056 [20]. **c**, Example data from ref. 20 showing a quantitative H3K27me3 profile in mES cells at 2,731 PcG target sites, in serum (red) and 2i (green) culture conditions and after treatment with EZH2 inhibitor EPZ-6438 (EZH2i) for 7 days (magenta and blue, respectively). EZH2i treatment strongly depletes the signal from >10 RPGC to ~0.1 RPGC (see the magnification in the lower graph), demonstrating that the peak signal at Polycomb binding sites is up to 99% specific and H3K27me3-dependent. Panel **a** created using BioRender.com. Panel **c** adapted with permission from ref. 20, Elsevier.

perform well for low-abundant protein targets, for proteins that indirectly interact with the DNA and/or for proteins that do not cross-link well.

Although, in principle, cell number could be substantially reduced for abundant targets, such as most histone PTMs, we do not consider MINUTE-ChIP to be particularly well suited for low-cell-number samples (e.g., ≤10,000 cells). This is because enzymatic chromatin digestion of small cell numbers becomes difficult to tune to the optimal fragment size, and barcoding also appears to become less efficient because the concentration of adapters must be reduced accordingly. Nevertheless, we have performed successful MINUTE-ChIP experiments with 100,000 cells per sample, and the creators of Mint-ChIP provide a modified version of their original method for 50,000 cells[25]. Thus, our protocol may provide a useful starting point for further development of low-input applications.

Overview of the procedure

Here, we provide a step-by-step protocol for performing and analyzing a MINUTE-ChIP experiment from either native or cross-linked chromatin. In brief, the protocol includes four parts:

1. Sample preparation (Steps 1–9), in which cells are lysed, and chromatin is fragmented, end-repaired and ligated to adapters with unique barcodes for each sample.
2. Pooling samples and immunoprecipitation (Steps 10–40), in which barcoded samples are pooled and split into several ChIP reactions and an aliquot for the input library.
3. Sequencing library preparation (Steps 41–77), in which first a linear amplification via in vitro transcription to RNA and RT from the RNA 3′ end to cDNA is performed, and subsequently a low-cycle PCR is performed to generate the Illumina sequencing library.
4. Data analysis (Steps 78–89), in which FASTQ files obtained from the sequencing platform are processed through our minute pipeline to yield BAM and quantitatively scaled bigWig files for downstream analysis and visualization.

Experimental design

If planned carefully, a MINUTE-ChIP can yield a large number of datasets and quantitative comparisons in a single experiment. We recommend including up to 16 samples of typically one to two Mio cells in a MINUTE-ChIP experiment. When a quantitative comparison of conditions with robust statistics (e.g., using statistical packages such as DESeq2[26]) is desired, we recommend including at least three biological replicates of each condition, allowing the comparison of up to four to five conditions in one experiment.

Recommended controls

Other samples, such as additional negative or positive controls might be added, even without replicates. For example, whenever possible, we recommend including a negative control where the histone PTM or the protein of interest has been depleted (Figs. 2c, 3d and 4c–e), either genetically (i.e., genetic knockout or knockdown) or pharmacologically or by using a cell type that does not express the protein of interest. These negative controls allow researchers to empirically determine and account for the technical background of the experiment (Figs. 2b,c, 3d and 4c). An important consideration in this respect is that the depth of data generated for each sample scales with the abundance of the epitope in the sample. As a consequence, the read coverage of those samples depleted in the probed modifications will necessarily be scarce but still accurately reflect the quantitative difference (see EZH2i tracks in Fig. 3e).

Considerations for immunoprecipitation

After pooling, multiple immunoprecipitation reactions can be performed in parallel from aliquots of the same pool. We recommend allocating at least four Mio cells to each ChIP reaction as a starting point, meaning that up to four epitopes can be probed from a pool of 16 samples. However, strong antibodies and abundant epitopes (such as H3K27me3) may yield high-quality data with less than four Mio cells, whereas immunoprecipitations for other targets, such as transcription factors, may benefit from a higher input amount. Generally, a small aliquot of the pool (e.g., 10% of the volume used for one ChIP reaction) is used to generate the input library. For quantitation, it is essential to sequence the input library, because this will establish the initial barcode distribution in the pool (Fig. 2a). Although the theoretical proportions of samples with exactly equal input cell counts would be 1:1, variability in barcoding performance always leads to over- and underrepresentation of some sample barcodes in a pool. The analysis pipeline accounts for this imbalance in the input pool as discussed in Data analysis.

If only one or a small number of epitopes will be analyzed, different aliquots of the pool can be used to probe with different antibodies for the same epitope (e.g., comparing antibody products from different vendors). This allows researchers to independently compare and validate antibody specificity from the same starting material. It is also possible to flash-freeze aliquots of the pool, providing an option to save a fraction of the pool for additional ChIP reactions at a later point.

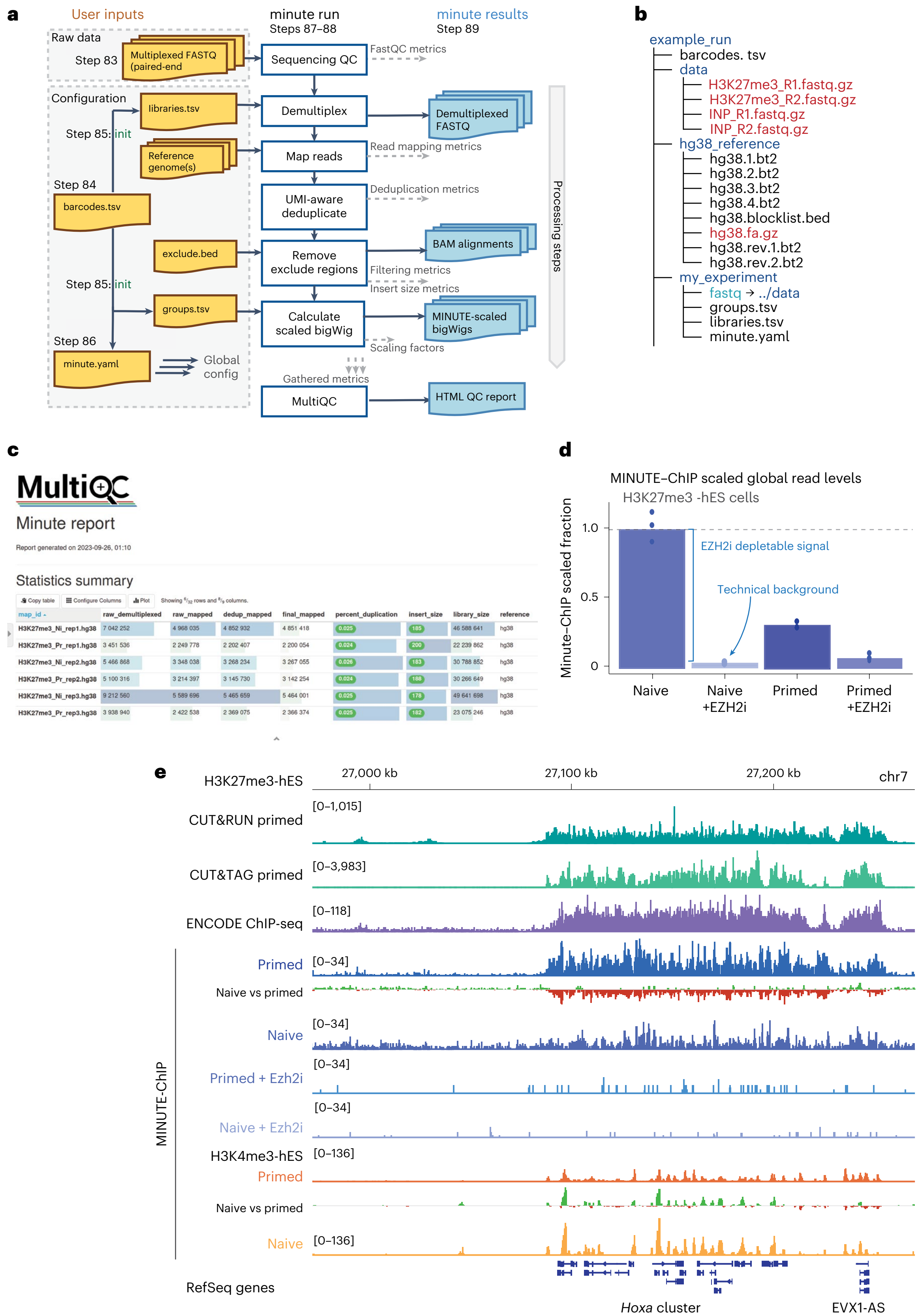
a
User inputs
minute run
Steps 87–88
minute results
Step 89
Raw data
Step 83
Multiplexed FASTQ (paired-end
Sequencing QC
FastQC metrics
Configuration
libraries.tsv
Demultiplex
Demultiplexed FASTQ
Step 85: init
Reference genome(s)
Map reads
Read mapping metrics
Step 84
barcodes.tsv
UMI-aware deduplicate
Deduplication metrics
exclude.bed
Remove exclude regions
BAM alignments
Filtering metrics
Insert size metrics
Step 85: init
groups.tsv
Calculate scaled bigWig
MINUTE-scaled bigWigs
Scaling factors
Step 86
minute.yaml
Global config
Gathered metrics
MultiQC
HTML QC report
Processing steps
b
example_run
barcodes. tsv
data
H3K27me3_R1.fastq.gz
H3K27me3_R2.fastq.gz
INP_R1.fastq.gz
INP_R2.fastq.gz
hg38_reference
hg38.1.bt2
hg38.2.bt2
hg38.3.bt2
hg38.4.bt2
hg38.blocklist.bed
hg38.fa.gz
hg38.rev.1.bt2
hg38.rev.2.bt2
my_experiment
fastq → ../data
groups.tsv
libraries.tsv
minute.yaml
c
MultiQC
Minute report
Report generated on 2023-09-26, 01:10
Statistics summary
Copy table
Configure Columns
Plot
Showing 6/32 rows and 8/9 columns.
map_id | raw_demultiplexed | raw_mapped | dedup_mapped | final_mapped | percent_duplication | insert_size | library_size | reference
H3K27me3_Ni_rep1.hg38 | 7 042 252 | 4 968 035 | 4 852 932 | 4 851 418 | 0.025 | 185 | 46 588 641 | hg38
H3K27me3_Pr_rep1.hg38 | 3 451 536 | 2 249 778 | 2 202 407 | 2 200 054 | 0.024 | 200 | 22 239 862 | hg38
H3K27me3_Ni_rep2.hg38 | 5 466 868 | 3 348 038 | 3 268 234 | 3 267 055 | 0.026 | 183 | 30 788 852 | hg38
H3K27me3_Pr_rep2.hg38 | 5 100 316 | 3 214 397 | 3 145 730 | 3 142 254 | 0.024 | 188 | 30 266 649 | hg38
H3K27me3_Ni_rep3.hg38 | 9 212 560 | 5 589 696 | 5 465 659 | 5 464 001 | 0.025 | 178 | 49 641 698 | hg38
H3K27me3_Pr_rep3.hg38 | 3 938 940 | 2 422 538 | 2 369 075 | 2 366 374 | 0.025 | 182 | 23 075 246 | hg38
d
MINUTE–ChIP scaled global read levels
H3K27me3 -hES cells
Minute-ChIP scaled fraction
1.0
0.5
0
EZH2i depletable signal
Technical background
Naive
Naive +EZH2i
Primed
Primed +EZH2i
e
H3K27me3-hES
27,000 kb
27,100 kb
27,200 kb
chr7
CUT&RUN primed
[0–1,015]
CUT&TAG primed
[0–3,983]
ENCODE ChIP-seq
[0–118]
Primed
[0–34]
Naive vs primed
Naive
[0–34]
Primed + Ezh2i
[0–34]
Naive + Ezh2i
[0–34]
MINUTE-ChIP
H3K4me3-hES
[0–136]
Primed
Naive vs primed
Naive
[0–136]
RefSeq genes
Hoxa cluster
EVX1-AS

Fig. 3 | Data analysis, quality control and visualization using the minute pipeline. a, Simplified scheme with the main steps of the minute bioinformatics pipeline. Input files are shown in yellow, and output files are shown in blue. Sequenced paired-end reads are demultiplexed, aligned to the reference genome, deduplicated on the basis of their UMIs and optionally filtered to remove alignments to blacklisted regions of the genome. Final read counts are used to calculate the final MINUTE-scaled bigWig files. A MultiQC HTML report is also generated. **b**, File structure after initializing a minute run on the example human ES (hES) cell data[21]. **c**, A screenshot of the MultiQC report obtained after running the full minute workflow on example hES cell data[21]. The number of demultiplexed, aligned, deduplicated and filtered reads per library are shown, together with percentage duplicate reads, median insert size and estimated library size. **d**, Example MINUTE-ChIP quantification, modified from ref. 21, showing global levels of H3K27me3 in naive or primed hES cells, cultured with or without EZH2 inhibitor, as determined by MINUTE-ChIP INRs relative to the naive culture condition. The bar plot shows pooled replicates, and dots indicate individual replicates. **e**, Example data[21] showing a genome browser view of the HOXA cluster with quantitatively scaled tracks from MINUTE-ChIP (GSE181241) performed in hES cells, with difference tracks showing gains and losses between naive and primed culture conditions. For comparison, ChIP (GSM1436879), CUT&RUN (GSM3677833) and CUT&Tag (GSM3536498) data from primed hES cells is shown. Raw data are available in the GEO database under accession number GSE181241. Chr, chromosome.

Input preparation

The protocol relies on optimal fragmentation of the starting material. Therefore, we strongly recommend performing titration of the MNase nuclease for the specific cell type by running an optimization experiment in which different MNase concentrations are tested and the fragmented chromatin is analyzed on an agarose gel or BioAnalyzer before the experiment (Supplementary Methods, Fig. 1b,c). The ideal digestion conditions yield a mono- to trinucleosomal ladder to avoid overdigestion, which can lead to poor adapter ligation efficiency in the subsequent step. It is also possible to enzymatically digest formaldehyde-fixed chromatin to mononucleosomes, but depending on the cell type, condensed regions of the genome may be resilient to MNase digestion. Hence, we provide here an alternative sample preparation protocol for the fragmentation of cross-linked material by using sonication (Fig. 4a).

Data analysis

Data analysis workflow

Data processing from raw paired-end FASTQ files to the final quality control results and scaled data tracks is fully automated by minute[27], our open-source, user-friendly Snakemake[28] workflow (Fig. 3a). In brief, users provide a sample barcode table, in which the sample name, barcode sequence and replicate groups are specified, along with a reference genome identifier. Then, the pipeline automates all necessary analysis steps from primary sequencing output to scaled ChIP–seq tracks for visualization and downstream analysis.

Cutadapt[29] is used for demultiplexing and extracting a paired-end FASTQ file for each sample in each ChIP or input library, followed by an alignment to the specified reference genome with Bowtie2[30] and an ad-hoc UMI-aware deduplication strategy based on the Je suite[31]. Optionally, a blacklist of regions (typically, artefact-prone regions in the target genome) for which all alignments will be discarded by using bedtools[32], can be specified. Based on the input sample barcode representation, the libraries are scaled to generate quantitative and directly comparable bigWig tracks by using deepTools[33].

Input files

A minute run requires several input files (Fig. 3a).

1. *Raw data: paired-end FASTQ files*. These are the FASTQ files obtained from the Illumina sequencing platform, typically downloaded from BaseSpace.
2. *Reference genome*. This is the FASTA genome sequence of the relevant organism. In addition, a matching bowtie2-index file is necessary to run the alignment (see the Procedure for information on how to build it).
3. *An annotation of excluded loci in the reference genome in BED format*. This is usually a blacklist obtained from a trusted source (e.g., ENCODE).
4. Minute run configuration files:
   - *libraries.tsv*. A tab-separated file that contains the following columns: (i) library name, (ii) replicate number, (iii) barcode sequence and (iv) corresponding FASTQ file basename. Note that all library names in libraries.tsv must be unique

- *groups.tsv*. A tab-separated file that matches each library in libraries.tsv with the replicate numbers (or the string 'pooled' to produce a pooled track from all replicates), the corresponding input library (for normalization), a scaling group value

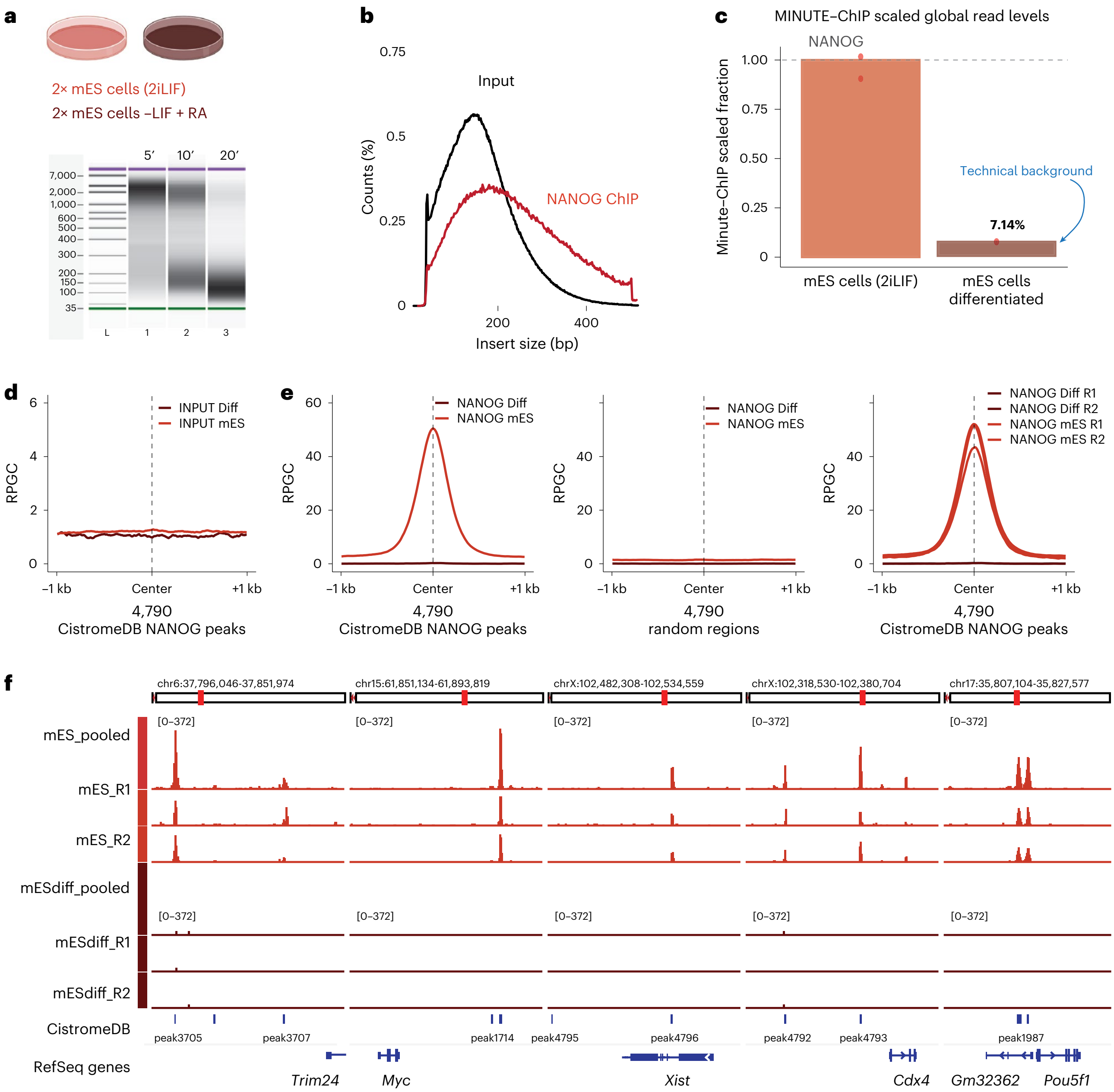

Fig. 4 | Demonstration data for formaldehyde-fixed MINUTE-ChIP, profiling pluripotency transcription factor NANOG. a, Top: experimental scheme of the demonstration experiment (GSE244908). Duplicates of mES cells grown in naive culture conditions or alternatively treated with retinoic acid to induce exit from pluripotency and differentiation. Cells were fixed by using 1% (vol/vol) formaldehyde and sonicated to fragment the chromatin. Bottom: example results from a sonication-based fragmentation of fixed mES cells. We recommend optimizing the sonication duration to obtain a chromatin shear between 150 and 300 bp, as seen in lane 3. **b**, Insert size distribution of reads of both NANOG and input samples obtained by Picard and summarized with MultiQC. Most fragments are sheared to 100–200 bp, but NANOG-bound fragments are on average slightly larger. **c**, Global, genome-wide levels of NANOG binding as determined by input-normalized total read counts (INR relative to mES cells) in mES cells or differentiated mES cells. The bar plot shows pooled replicates, and dots indicate individual replicates. **d**, Input profile plots for both mES cells and differentiated cells over a NANOG peak set obtained from the Cistrome database[37,38]. **e**, NANOG profile plots from mES cells and differentiated cells over the NANOG peakset[37,38], pooled (left) and replicates (right), and on randomized regions of the same size (middle). **f**, Genome browser examples of genomic regions with differential occupancy of Nanog for pooled as well as individual replicates of the different conditions. Diff, differentiated; LIF, leukemia inhibitory factor.

(usually matching each ChIP) and the reference genome for alignment. This reference genome name must match a genome defined in the minute.yaml file. The columns to provide are (i) ChIP library name, (ii) replicate number, (iii) corresponding input library name, (iv) scaling group and (v) genome. The first row of each scaling group in groups.tsv defines the reference condition

- minute.yaml

To simplify the initialization process, minute has a utility command `minute init` that infers a suitable pair of libraries.tsv and groups.tsv from a single barcodes.tsv file (Table 1) and the FASTQ files directory. barcodes.tsv is a text file in tab-separated value format describing the sample barcodes used, one barcode per row. The first row defines the reference condition across all ChIPs. The reference condition is always scaled to 1× genome coverage (RPGC), and the remaining samples are scaled quantitatively relative to the reference condition (see Scaling strategy).

Output files

As final output files (Fig. 3a), the minute pipeline provides (i) quantitatively scaled bigWig files that can be used for visualization in any genome browser tool, such as IGV[34], or for downstream analysis in any bigWig suite, such as wigglescout (https://github.com/cnluzon/wigglescout); (ii) demultiplexed FASTQ files for each sample; (iii) BAM alignments; and (iv) quality control indicators summarized in a MultiQC[35] HTML report (Fig. 3c). These outputs are structured as follows, in final and reports directories.

The final directory contains:

- *bam*. Alignments per sample after deduplication and filtering out the BED exclude regions (if provided)
- *bigwig*. bigWig genome track files you can visualize on genome browser tools like IGV or further analyze with other tools like deepTools or wigglescout
- *demultiplexed*. FASTQ files in which each condition is on a separate file

The reports directory contains:

- *fastqc*. A directory with the reports generated by FastQC on the pooled original FASTQ files
- *multiqc_report.html*. An html report gathering a summary of all relevant statistics and quality control indicators rendered by MultiQC
- *stats_summary.txt*. A table with the most important values per sample (demultiplexing, alignment, deduplication and insert size statistics)

Table 1 | Sample data barcodes.tsv content

Name	Replicate	Barcode	Reference
Ni	1	CTACCAGG	hg38
Ni	2	CATGCTTA	hg38
Ni	3	GCACATCT	hg38
Ni-EZH2i	1	TGCTCGAC	hg38
Ni-EZH2i	2	AGCAATTC	hg38
Ni-EZH2i	3	AGTTGCTT	hg38
Pr	1	CCAGTTAG	hg38
Pr	2	TTGAGCCT	hg38
Pr	3	ACCAACTG	hg38
Pr-EZH2i	1	GGTCCAGA	hg38
Pr-EZH2i	2	GTATAACA	hg38
Pr-EZH2i	3	TTCGCTGA	hg38

Contents of a barcodes.tsv (excluding header) for the example data. The sample barcodes included are triplicates for hES cell H3K27me3 in naive (Ni) and primed (Pr) conditions and triplicates for EZH2-inhibitor treatment on both (Ni-EZH2i and Pr-EZH2i). In this case, the reference sample is naive (Ni). Columns represent (i) sample name, (ii) replicate number, (iii) barcode sequence for demultiplexing, and (iv) reference genome. Reference genome must contain a matching entry in the minute.yaml configuration file. A tsv formatted version of this table is available to download from SciLifeLab Data Repository.

- *scalinginfo.txt*. Scaling factors used to generate the bigWig files
- *pooled_stats_summary.txt*. A table with the alignment and insert size values per replicate group

Scaling strategy

Quantitative scaling of MINUTE-ChIP data relies on the prerequisite that chromatin fragments from all samples compete for available antibody binding sites during immunoprecipitation (Fig. 2a). In an example scenario in which condition A1 has roughly twice as many H3K27me3-modified nucleosomes as sample B, chromatin fragments with the A barcode will have twice the chance of being captured by anti-H3K27me3 beads. After sequencing ChIP and input libraries, we can thus calculate the ratio of mapped reads in the ChIP sample to the mapped reads in the input, to account for the initial barcode distribution for each sample. We termed this ratio the INR, defined as: $INR_{ChIP, A1} = reads_{ChIP, A1}/reads_{input, A1}$. The INR is directly proportional to the abundance of the epitope in each sample, as validated by a standard curve, in which the INR values correlated well with the known quantities[20] (Fig. 2b). INRs are then used to produce scaled bigWig tracks in which one condition, defined as the reference condition, will be scaled to 1× genome coverage (RPGC), whereas all remaining tracks are scaled relative to the reference according to their respective INR. In the above scenario, if sample A is the reference, sample B will be effectively scaled to 0.5× genome coverage.

Materials

Biological materials

- Mouse ES cells: RW.4 originally isolated from the 129X1/SvJ (male) mouse strain (provided by the Karolinska Center for Transgene Technologies)
- Human ES cells: H9 cell line (female) (Wicell, WA09)

We anticipate that the protocol can be performed on other mouse and human ES cell line strains.

▲ **CAUTION** Before working with hES cells, ensure that all national regulations are followed and that required institutional ethics and review board approval is obtained and adhered to. The work presented here was performed under ethical permit 2011/745:31/3 issued by the Swedish Ethical Review Authority.

▲ **CAUTION** The cell lines used in your research should be regularly checked to ensure they are authentic and are not infected with mycoplasma.

Reagents

- Roche cOmplete Mini EDTA-free protease inhibitor cocktail (Sigma, 11836170001)
- Ethanol absolute ≥99.8% (vol/vol) (VWR, 28021.310)
  ▲ **CAUTION** Pure ethanol is toxic and highly flammable. Store in a fireproof cabinet and use it with caution.
- Nuclease-free water (ThermoFisher, AM9932)
- Pierce 16% (wt/vol) formaldehyde, methanol-free (ThermoFisher, 28906)
  ▲ **CAUTION** Formaldehyde is toxic if inhaled, ingested or absorbed through the skin; always wear a laboratory coat and gloves and use formaldehyde in a chemical hood. It presents an environmental threat, so dispose of waste according to institutional guidelines.
- PBS (Sigma, D5837)
- Tris-HCl (1 M), pH 8.0, RNase free (ThermoFisher, AM9855G)
- Trizma base (Sigma, cT1503)
- Hydrochloric acid, 37% (wt/vol) (HCl; Sigma, 320331)
  ▲ **CAUTION** Concentrated HCl is highly corrosive to the skin and is a respiratory irritant. Wear a laboratory coat and gloves when handling the solution in a chemical hood. Thoroughly wash any skin exposed to this chemical.

- EDTA, 0.5 M, pH 8.0 (ThermoFisher, AM9260G)
  ▲ **CAUTION** EDTA is toxic if swallowed. Wear gloves when handling.
- EGTA (Sigma, E3889)
  ▲ **CAUTION** EGTA is toxic if swallowed. Wear gloves when handling.
- SDS, 20% (wt/vol) (VWR, 0837-200ML)
  ▲ **CAUTION** SDS is corrosive to the skin and is a respiratory irritant. Wear gloves when handling the solution. Thoroughly wash any skin exposed to this chemical.
- TE buffer (Sigma, 93302-100ML)
- Triton X-100 (Sigma, T8787-50ML)
  ▲ **CAUTION** Triton X-100 is a skin and eye irritant. Wear gloves when handling.
- Tween 20 (Millipore, 655204)
- DTT (Sigma, D0632-5G)
  ▲ **CAUTION** DTT is toxic if inhaled or absorbed through the skin. Always wear a laboratory coat and gloves and use in a fume hood.
- Sodium deoxycholate (DOC; Sigma, 30970-100G)
  ▲ **CAUTION** DOC is harmful if swallowed or absorbed through the skin. It may cause respiratory and eye irritation. Wear gloves and avoid breathing dust.
- Magnesium chloride hexahydrate ($MgCl_2 \cdot 6H_2O$; Sigma, 2670)
- Calcium chloride dihydrate ($CaCl_2 \cdot 2H_2O$; Sigma, C8106)
- Sodium chloride solution, 5 M (NaCl; Sigma, S5150-1L)
- Nonidet P-40 substitute (Sigma, 74385)
- MNase (NEB, M0247S)
- End-It DNA end-repair kit (Lucigen, ER81050)
- Fast-Link DNA ligation kit (Lucigen, LK6201H)
- T4 DNA polymerase (NEB, M0203L)
- T4 polynucleotide kinase (ThermoFisher, EK0032)
- T4 DNA ligase (ThermoFisher, EL0011)
- ATP solution (ThermoFisher, R0441)
- dNTP mix (ThermoFisher, R1122)
- Proteinase K (Life Technologies, 25530015)
- HiScribe T7 quick high yield RNA synthesis kit (NEB, E2050S)
- T4 RNA ligase 2, truncated (NEB, M0242L)
- RNaseOUT (ThermoFisher, 10777019)
- SuperScript III First-Strand Synthesis SuperMix (ThermoFisher, 108080-400)
- RNase H (ThermoFisher, EN0202)
- NEBNext Ultra II Q5 master mix (NEB, M0544)
- Dynabeads protein A/G beads (ThermoFisher, 10001/3D)
- Agencourt AMPure XP (Beckman Coulter, A63881)
- Agencourt RNAClean XP (Beckman Coulter, A63987)
- Qubit double-stranded DNA high-sensitivity (HS) assay kit (Life Technologies, cat.no. Q32854)
- Qubit RNA HS assay kit (Life Technologies, Q32852)
- Agilent high sensitivity DNA kit (Agilent, 5067-4626)
- NextSeq 500/550 high output kit v2 (75 cycles) (Illumina, 20024906)
- Mouse feeder removal microbeads (Milteny Biotech, 130-095-531)

Oligonucleotides

- The required oligonucleotides are listed in Supplementary Table 1. All oligos were ordered from IDT in their lyophilized form

Antibodies

- H3 (Active Motif, 39763; RRID: AB_2650522)
- H3K4me3 (Millipore, 04-745; RRID: AB_1163444)
- H3K27me3 (Millipore, 07-449; RRID: AB_310624)
- H3K27me3 (Cell Signaling Technology, 9733; RRID: AB_2616029)

- H3K27me3 (Diagenode, C15410195; RRID: AB_2753161)
- H3K27me2 (Cell Signaling Technology, 9728; RRID: AB_1281338)
- H3K27me1 (Active Motif, 61015; RRID: AB_2715573)
- H3K4me1 (Cell Signaling Technology, 5326; RRID: AB_10695148)
- H4K20me3 (Millipore, 07-463; RRID: AB_310636)
- H4K20me1 (Abcam, ab9051; RRID: AB_306967)
- H3K9me3 (Active Motif, 39161; RRID: AB_2532132)
- H2Aub (Cell Signaling Technology, 8240; RRID: AB_10891618)
- NANOG (Abcam, ab80892; RRID: AB_2150114)

Equipment

- Scientific Industries Vortex-Genie (or equivalent)
- Eppendorf ThermoMixer (or equivalent; Eppendorf, 2231000574)
- PCR tubes
- DNA/protein 1.5-ml low-bind tubes (Sarstedt, 72.706.700/600)
- 0.5-ml Bioruptor Plus microtubes (Diagenode, C30010013)
- Agilent 2100 Bioanalyzer system (or equivalent)
- Invitrogen Qubit fluorometer (or equivalent)

Hardware

- Preferably, a Linux x86_64 system with eight cores and 16 GB of random access memory (RAM). minute is implemented in Snakemake and, as such, can run on other systems as well. It can also be configured to run on high-performance computing environments. For more details about Snakemake support of other non-Linux-based systems, refer to Snakemake documentation at https://snakemake.readthedocs.io. The detailed pipeline architecture and full parameter description can be found in the *bioRxiv* preprint[27] and minute official documentation (https://minute.readthedocs.io/en/latest/)

Software

▲ **CRITICAL** To ensure reproducibility and portability, minute software dependencies are organized in a dedicated conda environment.

- A 64-bit Linux operating system (this protocol was tested under Ubuntu 22.04)
- conda 24.1.2 (https://docs.conda.io), used to reproduce the software environment on which minute has been tested
- An active minute conda environment. The software dependencies installed on a functional minute conda environment are as follows: minute 0.6.0 (https://github.com/elsasserlab/minute), python 3.10.8 (https://www.python.org/), snakemake 7.32.4 (https://snakemake.readthedocs.io/en/stable/), samtools 1.16.1 (https://github.com/samtools/samtools), cutadapt 4.8 (https://cutadapt.readthedocs.io/en/stable/), bowtie2 2.5.1 (https://bowtie-bio.sourceforge.net/bowtie2/index.shtml), je-suite 2.0.RC (https://github.com/gbcs-embl/Je), deeptools 3.5.5 (https://deeptools.readthedocs.io/en/develop/), Picard 3.1.1 (https://broadinstitute.github.io/picard/), FastQC v0.12.1 (https://github.com/s-andrews/FastQC), bedtools v2.31.1 (https://bedtools.readthedocs.io/en/latest/), MultiQC 1.13 (https://multiqc.info/), R 4.2.1 (https://www.r-project.org), ggplot2 3.5.1 (https://ggplot2.tidyverse.org/), dplyr 1.1.4 (https://dplyr.tidyverse.org/)

Reagent setup

25% (wt/vol) Triton X-100

Dilute 12.5 g of 100% (wt/vol) Triton X-100 to 25% (wt/vol) in 50 ml of nuclease-free water. The solution can be stored at room temperature (21–25 °C) for ≥6 months.

12.5% (wt/vol) DOC

Dissolve 2.5 g of DOC in 20 ml of nuclease-free water (12.5%, wt/vol). The buffer can be stored at room temperature for ≥6 months, protected from light.

1 M $CaCl_2$ solution
Dissolve 11.1 g of $CaCl_2$ in 100 ml of nuclease-free water. The solution can be stored at room temperature for ≥6 months.

0.5 M EGTA solution
Dissolve 19 g of EGTA in 70 ml of distilled water, adjust pH to 7.5 with KOH pellets and then bring the solution to 100 ml with Milli-Q water. The solution can be stored at room temperature for ≥6 months.

2× lysis buffer
Prepare 2× lysis buffer by mixing 100 mM Tris-HCl (pH 8.0), 0.2% (wt/vol) Triton X-100, 0.1% (wt/vol) DOC and 10 mM $CaCl_2$ in Milli-Q water. The buffer can be stored at 4 °C for 1 month.

3× stop buffer
Prepare 3× stop buffer by mixing 150 mM Tris-HCl (pH 8.0), 450 mM NaCl, 3% (wt/vol) Triton X-100, 150 mM EGTA, 150 mM EDTA and 0.3% (wt/vol) DOC in Milli-Q water. The buffer can be stored at 4 °C for ≥6 months.

Wash buffer RIPA
Mix 0.1% (wt/vol) DOC, 0.1% (wt/vol) SDS, 1% (wt/vol) Triton X-100, 10 mM Tris-HCl (pH 8.0), 1 mM EDTA and 140 mM NaCl in Milli-Q water. The buffer can be stored at 4 °C for ≥6 months.

Wash buffer RIPA high salt
Mix 0.1% (wt/vol) DOC, 0.1% (wt/vol) SDS, 1% (wt/vol) Triton X-100, 10 mM Tris-HCl (pH 8.0), 1 mM EDTA and 360 mM NaCl in Milli-Q water. The solution can be stored at 4 °C for ≥6 months.

Wash buffer LiCl
Mix 250 mM LiCl, 0.5% (vol/vol) Nonidet P-40 substitute, 0.5% (wt/vol) DOC, 1 mM EDTA and 10 mM Tris-HCl (pH 8.0). The solution can be stored at 4 °C for ≥6 months.

ChIP elution buffer
Prepare ChIP elution buffer by mixing 1× TE, 0.1% (wt/vol) SDS and 300 mM NaCl in Milli-Q water. The buffer can be stored at room temperature for 1 month.

1× sonication buffer
Mix 50 mM Tris-HCl (pH 8.0), 0.5% (wt/vol) SDS and water. The buffer can be stored at 4 °C for 1 month.

Sonication dilution buffer
Mix 50 mM Tris-HCl (pH 8.0) and 2% (wt/vol) Triton X-100 in Milli-Q water. The solution can be stored at 4 °C for 1 month.

3 M Tris-HCl solution
Dissolve 90.85 g of Trizma base in 150 ml of distilled water, adjust pH to 8.0 with concentrated (~37%, wt/vol) HCl and then bring the solution to 250 ml with distilled water. The solution can be stored at room temperature for ≥6 months.

70% (vol/vol) ethanol
Mix 7 ml of absolute ethanol with distilled water to a final 10 ml. Make fresh on every day of the protocol, to avoid evaporation.

Cell line culture and maintenance
Culture conditions for the respective cell lines can be found in the original publications[20,21]. For a summary, see Box 1.

BOX 1

Culture conditions for cell lines

In brief, mES cells are cultured feeder-free in 0.1% (w/vol) gelatin-coated dishes. For the serum condition, maintain cells in Knockout DMEM medium (Life Technologies, 10829018) supplemented with 2 mM Glutamax (Life Technologies, 31331028), 0.1 mM non-essential amino acids (Sigma, M7145), 15% ES cell-grade FBS (Sigma, F7524), 0.1 mM β-mercaptoethanol and leukemia inhibitory factor (Millipore, ESG1107). For the 2i (naive) condition, grow cells in serum-free ESGRO complete basal medium (Millipore, SF002-500) with 0.1 mM leukemia inhibitory factor, 1 μM MEK inhibitor PD0325901 (Sigma, PZ0162) and 3 μM GSK3 inhibitor CHIR99021 (Sigma, SML1046). For all conditions, culture cells at at 37°C, 5% O_2 and 5% CO_2 and passage in 48-h intervals, using ACCUTASE (Sigma, A6964) for detachment.

Thaw primed female H9 hES cells (Wicell, WA09) by using Nutristem hES cell XF medium (Biological Industries, 05–100–1 A) and culture on tissue culture plates pre-coated with 10% (vol/vol) LN521 (Biolamina, LN521-02). Passage cells when 70–80% confluency is reached. Thaw naive H9 cells (previously converted to the naive stem cell state by using NaïveCult induction medium (STEMCELL Technologies, 05580) onto plates with high-density inactivated mouse embryonic fibroblasts (Gibco, A24903) by using NaïveCult expansion medium (t2iLGö; STEMCELL Technologies, 05590) supplemented with 10 μM ROCKi (Merck, Y-27632). Passage cells every 4–5 days. Culture all hES cells at 37 °C, 5% O_2 and 5% CO_2.

Sample preparation

Harvest cultured cells by detaching them from the culture plate by using accutase and then neutralizing with fresh medium (use the same volume as used for culturing). Mouse feeder removal microbeads should be used according to the manufacturer's protocol to reduce the amount of mouse embryonic fibroblasts in the naive hES cell samples.

For the native condition, pellet the cells by centrifugation at 300*g* for 5 min. Wash the cell pellets twice with ice-cold PBS to remove any residual culture medium. In the second PBS wash, count the cells and aliquot volumes of cell suspension corresponding to 1 million cells into 1.5-ml tubes before spinning down. Try to remove as much of the PBS as possible and proceed to flash-freeze the cell pellet in liquid nitrogen. The frozen cell pellets can be stored at −80 °C for 6 months. On the day of the experiment, thaw the cells on ice and proceed with resuspension until there are no visible clumps.

For the cross-linked condition, add fresh 16% (wt/vol) stock of formaldehyde to a final concentration of 1% (wt/vol). Incubate the cells for 10 min at room temperature on a nutating mixer. Quench the cross-linker by adding 3 M Tris-HCl (pH 7.8) to a final concentration of 0.75 M and incubating for 10 min at room temperature on a nutating mixer. Pellet the cells by centrifugation at 300*g* and 4 °C for 5 min. Wash the cell pellets twice with ice-cold PBS and count the cells in the final wash. Continue with dividing the cell suspension into aliquots equivalent to 5 million cells and freezing the cell pellets as described above for the native condition.

Preparation of adapter oligos

1. Reconstitute the lyophilized 'adapter' oligos at 500 μM each (200 μl/100 nmol) in 1× TE buffer (pH 7.4).
2. Mix equal volumes of matching sense and antisense adapter oligos and anneal in a thermocycler by using the program indicated below. The adapter annealing step has a large impact on the efficiency of the protocol, so be sure to allow for slow cooling down from 5 degrees above Tm to 5 degrees below melting temperature (Tm).
   - Heat to 95 °C and maintain the temperature for 3 min
   - Cool to 69 °C, ramping down the temperature by 1 degree/10 s
   - Maintain the temperature at 69 °C for 5 min
   - Cool to 40 °C, ramping down the temperature by 1 degree/10 s
   - Cool to 4 °C indefinitely
3. Dilute the annealed, double-stranded adapters (concentrated to 250 μM) to a working concentration of 50 μM in 1× TE buffer (pH 7.4) and divide into smaller-volume aliquots (to avoid several freeze/thaw cycles). The aliquots of annealed adapters can be stored at −20 °C for ≥2 years.

Preparation of PCR primers

1. Resuspend the lyophilized 'PCR primer' oligos at 100 µM each with 1× TE buffer (pH 7.4).
2. For each pair of barcoded primers, prepare a primer mix at 10 µM as follows:
   - 25 µl of forward primer (always the same PCR_Forward primer)
   - 25 µl of reverse primer (with barcode PCR_BC1-10)
   - 200 µl of 1× TE
3. The resulting primer pair mixes (each 10 µM) can be stored at −20 °C for ≥2 years. Dividing into aliquots is not necessary, because repeated freeze/thaw cycles are possible.

Preparation of the RNA 3′ adapter (RA3) and RT_primer

1. Resuspend the lyophilized oligos at 100 µM in 1× TE buffer (pH 7.4).
2. Further dilute the 100 µM RA3 stock to 10 µM and the 100 µM RT_primer stock to 20 µM in 1× TE buffer (pH 7.4).
3. These primer stocks can be stored at −20 °C for ≥2 years. Divide into smaller-volume aliquots (to avoid several freeze/thaw cycles). The aliquots of annealed adapters can be stored at −20 °C for ≥2 years.

Software setup

Download and install conda and bioconda

Conda can be installed by using the following steps (see https://conda.io/projects/conda/en/latest/index.html for official documentation):

```
mkdir -p ~/miniconda3

wget https://repo.anaconda.com/miniconda/Miniconda3-latest-
Linux-x86_64.sh -O ~/miniconda3/miniconda.sh

bash ~/miniconda3/miniconda.sh -b -u -p ~/miniconda3
rm -rf ~/miniconda3/miniconda.sh
```

After this, conda can be initialized by:

```
~/miniconda3/bin/conda init bash
```

Set up up the bioconda channels (see https://bioconda.github.io/ for more details). This is a one-time action:

```
conda config --add channels defaults
conda config --add channels bioconda
conda config --add channels conda-forge
conda config --set channel_priority strict
```

Create the minute conda environment

After setting up conda, create the minute environment:

```
conda create -n minute minute
```

Activate the minute environment

```
conda activate minute
```

Ensure that the minute conda environment is active by running `conda env list`. The active environment will have a '*' next to it. If you perform this workflow at different times or close the terminal that you were using, you will need to reactivate the conda environment.

Download the example dataset

We provide an example dataset (https://doi.org/10.17044/scilifelab.25348405.v1) to illustrate the functioning of the pipeline. We recommend first running through the data-analysis protocol with this sample data. You can download all the necessary files and put them in a directory, example_run/example_data, with the following commands:

```
mkdir -p example_run/example_data
cd example_run/example_data

wget -O H3K27me3_R1.fastq.gz https://figshare.scilifelab.se/ndownloader/files/44890294
wget -O H3K27me3_R2.fastq.gz https://figshare.scilifelab.se/ndownloader/files/44890291
wget -O INP_R1.fastq.gz https://figshare.scilifelab.se/ndownloader/files/44890288
wget -O INP_R2.fastq.gz https://figshare.scilifelab.se/ndownloader/files/44890285
wget -O barcodes.tsv https://figshare.scilifelab.se/ndownloader/files/44890255
wget -O minute.yaml https://figshare.scilifelab.se/ndownloader/files/44890264

cd ..
```

The files downloaded at this step are the following:

1. Two pairs of FASTQ files: H3K27me3_R1.fastq.gz and H3K27me3_R2.fastq.gz; INP_R1.fastq.gz and INP_R2.fastq.gz. R1 and R2 correspond to paired-end sequencing reads. This is the equivalent raw data to the files that you will obtain from the Illumina platform after sequencing.
2. A barcodes.tsv file used to specify the samples within each pool. It is used by the minute `init` command to initalize the main configuration files, libraries.tsv and groups.tsv.
3. A minute.yaml configuration file.

Procedure

Part 1: sample preparation

Lysis and fragmentation

● **TIMING** 1 h

▲ **CRITICAL** The optimal conditions for lysis and fragmentation should be determined on a separate day, as described in the Supplementary Methods.

1. Prepare the desired volume of ice-cold 2× lysis buffer (25 µl per sample) supplemented with 100× protease inhibitor to a final concentration of 2× and use it to dilute MNase according to the optimal concentration determined by titration (as described in the Supplementary Methods).
2. Prepare an end-repair and ligation mix on ice by using the components provided in the end-repair and ligation kit and keep it on ice until needed (Step 7).

Components	Final concentration	Volume per reaction (µl)	16+1 reactions
10× blunting buffer	1×	10	170
10 mM ATP	1.5 mM	15	255
10 mM dNTP	1 mM	10	170
100× protease inhibitor	1×	1	17

Components	Final concentration	Volume per reaction (µl)	16+1 reactions
0.5 M EGTA	30 mM	6	102
End-repair enzyme mix	–	2	34
Fast-link DNA ligase	–	2	34
Total	–	46	782

3. Harvest the cells (see Reagent setup) and resuspend each cell pellet (1–2 million cells) in 25 µl of ice-cold 1× PBS.
4. To each sample from Step 3, add 25 µl of 2× lysis buffer-MNase at the optimized concentration (from Step 1) (final volume = 50 µl).
   ◆ TROUBLESHOOTING
5. Lyse the cells on ice for 20 min.
6. Perform MNase digestion by following option A if working with native cells or option B if working with cross-linked cells.
   (A) **Native condition**
      (i) Digest the samples at 37 °C in a thermo mixer set at 800 rpm for 10 min.
      (ii) Place the samples immediately back on ice to stop digestion.
   (B) **Cross-linked condition**
      (i) Resuspend each 5 M cell pellet in 100 µl of ice-cold 1× sonication buffer and transfer to 0.5-ml sonication tubes.
      (ii) Sonicate by using optimized conditions for your sonicator (see Supplementary Methods, 'fragmentation optimization').
      (iii) Centrifuge the lysate in a microcentrifuge (2,000*g*) for 30 s to collect all the liquid and place it on ice. Twenty microliters of this solution corresponds to 1 million cells.
      (iv) Split the lysate into the desired number of 20-µl aliquots (e.g., to use as technical replicates). Pipette up and down a few times before dividing into aliquots, to ensure that the lysate is mixed well.
      (v) Add 30 µl of ice-cold sonication dilution buffer to each 20 µl of lysate.

End-repair and sample barcode ligation

● **TIMING overnight**

7. Add 46 µl of end-repair and ligation mix (from Step 2) to each 50 µl of lysate, mix well by pipetting and place the tubes back on ice.
8. To each sample, add 4 µl of one of the adapter oligos (50 µM working concentration) BC1-16 (total volume per sample = 100 µl). Make sure to take note of which adapter was added to which sample.
9. Place the tubes in a thermoshaker and incubate under constant agitation at 800–1,000 rpm for 2 h at room temperature and then at 16 °C overnight with agitation at 800 rpm.

Part 2: pooling samples and immunoprecipitation

▲ **CRITICAL** Immunoprecipitation conditions such as the amount of antibody to use and the volume of chromatin pool to be added can be optimized for individual experiments. We recommend using one to six Mio cells per immunoprecipitation depending on the expected abundance of the epitope and strength of the antibody. For example, for 16 samples, the total pooled volume of ~2,400 µl (16 Mio cells) can be split equally into four immunoprecipitations of 500 µl, while 50 µl is set aside as input. Here, we provide our recommendation to prepare beads coupled to the desired antibody. However, users can follow their own routine practice.

Precouple antibody to beads

● **TIMING 1 h**

10. One hour into the incubation in Step 9, transfer 50 ul of well-mixed protein A bead slurry for each antibody to a low-protein-binding tube.

11. Place the tube on a magnetic stand and wait until the liquid is clear.
12. Keeping the tube on the magnet, carefully remove the liquid by pipetting. Wash the beads with 500 µl of PBS-0.1% (vol/vol) Tween.
13. Wash the beads one additional time, repeating Steps 11 and 12.
14. Carefully remove and discard the second wash and resuspend the beads in 200 µl of PBS-0.1% (vol/vol) Tween containing 3–5 µl of the desired antibody. Incubate the beads for 1 h at room temperature (or until the sample pool has been prepared in Step 18).
    ◆ TROUBLESHOOTING

Pooling samples

● **TIMING 10 min**

15. To an appropriate volume of 3× stop buffer, add 100× protease inhibitor to a final concentration of 3×. End the barcoding reaction by adding 50 µl of this solution to each sample from Step 9 (total volume per sample = 150 µl).
16. Pool all the samples into a single tube. Using a tabletop centrifuge, centrifuge at maximum speed (e.g., 12,000*g*) for 5 min at 4 °C to remove debris. Transfer the supernatant to a fresh tube.
17. Save an appropriate amount of the pool as input (usually 10% (vol/vol) of the volume used for the immunoprecipitations) at 4 °C until Step 30.
    ■ **PAUSE POINT** We strongly recommend continuing straight to the immunoprecipitation step, but it is possible to flash-freeze the pool to continue later. Before flash-freezing the pool, we recommend adding glycerol to a final concentration of 10% (vol/vol) and storing it at −80 °C for up to several months.

Immunoprecipitation

● **TIMING 6 h**

18. For each antibody-bead coupling reaction from Step 14, place the tube on a magnetic stand and wait until the liquid is clear.
19. Keeping the tube on the magnet, carefully remove the liquid by pipetting. Wash the beads quickly (<1 min) with 500 µl of PBS-0.1% (vol/vol) Tween to wash out any unbound antibody.
20. Repeat Steps 18 and 19 quickly (<1 min) with 500 µl of PBS-Tween 0.1% (vol/vol) while keeping the beads on the magnet; remove all wash buffer.
21. For each immunoprecipitation reaction, add the the desired volume of the chromatin pool (200–500 µl; see the Critical callout above) to the pre-coupled antibody-beads and resuspend well after removing from the magnet.
22. Perform the immunoprecipitation reactions at 4 °C for 4 h on a rotor.
    ◆ TROUBLESHOOTING
    ■ **PAUSE POINT** The immunoprecipitation can be performed overnight at 4 °C.
23. Place the samples on a magnetic stand and incubate for 1 min until the liquid is clear. Carefully pipette the supernatant into a clean tube. Store the supernatant at −20 °C (it can be useful for quality control and troubleshooting).
    ◆ TROUBLESHOOTING
24. Resuspend the beads in 500 µl of ice-cold RIPA buffer.
25. Place the tubes on a magnetic stand, wait until the liquid is clear and remove and discard the supernatant. This wash should take <1 min.
26. Repeat Steps 20 and 21 one more time, for a total of two washes.
27. Quickly wash the beads as described in Steps 20 and 21 with the following buffers:
    - Once with 500 µl of ice-cold RIPA/high-salt buffer
    - Twice with 500 µl of ice-cold LiCl wash buffer
28. At the end of the last wash, resuspend the beads in 200 µl of ice-cold 1× TE buffer and transfer the suspension to a new DNA low-binding tube.

29. Placing the tube on a magnetic stand, carefully remove the TE buffer. Then, resuspend the chromatin/antibody/bead conjugates by adding 100 µl of the following mix:

	Volume per reaction (µl)	4+1 reactions
ChIP elution buffer	97.5	487.5
Proteinase K	2.5	12.5

30. Bring the volume of the input sample from Step 17 to 100 µl by adding 2.5 µl of Proteinase K and an appropriate volume of ChIP elution buffer.
31. Perform protein digestion by following option A if working with native material or option B if working with cross-linked material
    (A) **Native condition**
        (i) Digest the proteins by incubating the immunoprecipitation bead suspension (Step 29) and the input (Step 30) at 63 °C for 1 h with agitation (800 rpm on a thermoshaker).
    (B) **Cross-linked condition**
        (i) Supplement the ChIP elution buffer with 1% (vol/vol) SDS and perform the protein digestion by incubating the samples at 63 °C overnight (~16 h) with agitation (800 rpm on a thermoshaker). These conditions will also reverse the cross-links.
32. Cool the samples to room temperature and centrifuge in a microcentrifuge (2,000*g*) for 30 s to collect any condensate.
33. Bring AMPure XP beads to room temperature 30 min before use and mix thoroughly before use. Prepare a fresh aliquot of 70% (vol/vol) ethanol on the day of the experiment.
34. Place the Proteinase K digestions from Step 32 on a magnetic stand and, once the liquid is clear, transfer 90 µl of the liquid to new tubes and add 90 µl of AMPure XP bead suspension directly to each new tube.
35. Mix by pipetting until the solution looks homogeneous and incubate for 10 min to allow the DNA to bind to the beads.
    ▲ **CRITICAL STEP** The incubation with 1.0× AMPure beads will allow users to select all DNA fragments with lengths greater than 100 bp, thereby reducing the carryover of excess adapters. This is especially important for the input sample.
36. Place the samples on a magnetic stand and incubate them for 5 min.
37. Remove the supernatant and wash the AMPure beads twice with 200 µl of 70% (vol/vol) ethanol, without removing the tubes from the magnet.
38. Air-dry at room temperature for 1 min and remove the tubes from the magnet.
39. Resuspend the beads in 9 µl of 1× TE buffer, wait for 1 min, place the tubes back on the magnet and transfer the eluted DNA to a new PCR tube.
40. Use 1 µl of the eluate to measure the immunoprecipitated DNA concentration with a Qubit DNA HS assay kit.
    ◆ **TROUBLESHOOTING**
    ■ **PAUSE POINT** The immunoprecipitation DNA can be stored at −20 °C for ≤3 weeks.

Part 3: sequencing library preparation

In vitro transcription, RNA cleanup and RNA adapter ligation

● **TIMING** 5.5 h

41. To each sample (8 µl of DNA), add 12 µl of the in vitro transcription (IVT) master mix as described below.

Components	Volume per reaction (µl)	4+1 reactions
DNA from the immunoprecipitation (Step 39)	8	–
2× NTP buffer mix	10	50
T7 RNA polymerase mix	2	10
Total	20	–

42. Incubate the samples at 37 °C for 4 h.
■ **PAUSE POINT** If it is not possible to continue directly to Step 41, the RNA samples can be held/stored at 4 °C overnight.
43. Add 1 µl of DNase I to each sample and incubate for 15 min at 37 °C.
44. Thirty minutes before their use, bring RNAclean XP beads to room temperature and mix thoroughly. Prepare a fresh aliquot of 70% (vol/vol) ethanol.
45. Add 38 µl of RNAclean XP beads to 21 µl of sample from Step 43 (1.8 times the volume of the sample).
46. Mix by pipetting until the solution looks homogeneous and then incubate for 10 min.
47. Place the tubes on a magnetic stand and incubate for 5 min to retrieve all the beads.
48. Remove the supernatant and wash the beads twice with 200 µl of freshly prepared 70% (vol/vol) ethanol.
49. Keeping the tubes on the magnetic stand, carefully remove the ethanol. Then, air-dry at room temperature for 1 min.
50. Remove the tubes from the magnet and resuspend the beads in 9 µl of 1× TE. Wait for 1 min, place the tubes back on the magnet and transfer the eluted RNA to a new tube.
51. Use 1 µl of the sample to measure the RNA concentration by using the Qubit RNA HS assay kit. This measurement serves as indication of successful amplification if the yield of RNA is higher than the input DNA amount.
◆ **TROUBLESHOOTING**
52. Use the remaining RNA (8 µl) in the ligation reaction as indicated in the table below and incubate it at 25 °C for 1 h.

Components	Volume per reaction (µl)	4+1 reactions
RNA from IVT (from Step 48)	8	–
RNA3 adapter (10 µM)	1	5
10× RNA ligase buffer	2	10
T4 RNA ligase truncated	1	5
RNaseOUT	1	5
Total	13	–

Reverse transcription and cDNA cleanup

● **TIMING 1.5 h**

53. After the ligation of the RA3 adapter, place the tubes on ice and add the following components of the reverse transcription (RT) reaction:

Components	Volume per reaction (µl)	4+1 reactions
RNA from Step 50	13	–
RT primer (20 µM) (see Reagent setup)	4	20
Annealing buffer	1	5
Total	18	–

54. Heat the samples to 65 °C for 5 min and then place the tubes on ice for 1 min. Then, add the room temperature mix prepared as follows:

Components	Volume per reaction (µl)	4+1 reactions
2× first-strand reaction mix	20	100
SIII/RNaseOUT enzyme mix	2	10
Total	40	–

55. Incubate the samples in a thermocycler with the following program:
    - 25 °C for 10 min
    - 50 °C for 50 min
    - 85 °C for 5 min
    - 4 °C indefinitely

56. Add 1 µl of RNase H and incubate the samples for 15 min at 37 °C.
57. Thirty minutes before use, bring RNAclean XP beads to room temperature and mix thoroughly.
58. Add 74 µl of RNAclean XP beads to 41 µl of the cDNA mixture (1.8 times the sample volume) from Step 54.
59. Mix by pipetting and incubate for 10 min.
60. Place the tubes on a magnetic stand and incubate for 5 min. While keeping the samples on the magnet, remove and discard the supernatant.
61. Wash the beads twice with 200 µl of 70% (vol/vol) ethanol.
62. Air-dry the samples at room temperature for 1 min.
63. Elute the DNA in 25 µl of 1× TE buffer and transfer the eluted cDNA to a new tube.
64. Use 1 µl of the eluate to measure cDNA concentration on a NanoDrop spectrophotometer or Qubit fluorometer.

◆ TROUBLESHOOTING

■ **PAUSE POINT** The cDNA sample can be stored at −20 °C for ≤48 h.

Library preparation

● **TIMING 2.5 h**

65. Take a volume of solution (from Step 59) corresponding to 200 ng of cDNA and make up to 22.5 µl with nuclease-free water (or all the eluted cDNA if the total amount is <200 ng).
66. Set up the PCR reaction as described in the table below, by using a different PCR primer pair (PCR_BC1-10; see Reagent setup and Supplementary Table 1) for each reaction.

Components	Volume per reaction (µl)
cDNA (≤200 ng)	22.5
2× PCR premix	25
PCR primer pair (10 µM each)	2.5
Total	50

▲ **CRITICAL STEP** It is important to note the primer pair used for each PCR, because each primer pair corresponds to a unique Index1 sequence for pooling on an Illumina sequencer and demultiplexing with bcl2fastq or BaseSpace.

67. Place the tubes in a thermocycler and perform a PCR reaction with the following cycling conditions:

Number of cycles	Denature	Anneal	Extend	Hold
1	95 °C, 3 min	–	–	–
4	95 °C, 15 s	65 °C, 30 s	72 °C, 30 s	–
10	95 °C, 15 s	–	72 °C, 1 min	–
1	–	–	72 °C, 5 min	–
1	–	–	–	4 °C, indefinitely

▲ **CRITICAL STEP** This part of the procedure allows us to size-select DNA fragments with lengths greater than 200 bp by using 0.7× AMPure beads. This part of the procedure is important to remove excess adapters.

68. Thirty minutes before use, bring AMPure XP beads to room temperature and mix thoroughly.
69. Add 35 µl of AMPure XP beads (0.7×) to 50 µl of PCR mixture (from Step 67).
70. Mix well by pipetting and incubate for 10 min.
71. Place the tube on a magnetic stand and wait until the liquid is clear. Then, remove the supernatant.
72. While keeping the tube on the magnetic stand, wash the beads twice with 200 µl of freshly prepared 70% (vol/vol) ethanol.

73. Air-dry the beads at room temperature for 1 min.
74. Remove the tubes from the magnet, resuspend the beads in 15 µl of 1× TE buffer and incubate for 1 min to elute the DNA.
75. Place the tubes back on the magnetic stand and incubate until the liquid is clear. Transfer the supernatant to a new tube to yield your sequencing libraries.
76. Use 1 µl of the library to measure its concentration on the Qubit fluorometer and run 1 µl of the library on a Bioanalyzer to obtain library size distribution.
77. Follow the specific instructions for your Illumina sequencing platform to prepare the final dilutions of your sequencing libraries, and sequence the libraries by using standard Illumina sequencing primers provided in the run cartridge. Information can be found on the supplier's website (https://emea.illumina.com/informatics/infrastructure-pipeline-setup/sequencing-experiment-setup.html). A useful sequencing depth per sample barcode is 10–50 Mio reads (i.e., 120–600 Mio reads for a ChIP or input library of a pool of 12 sample barcodes, prepared as described above).
    ◆ **TROUBLESHOOTING**

Part 4: data analysis

● **TIMING 4 h plus processing time**

Environment setup

78. Activate the minute conda environment.

```
conda activate minute
```

79. Ensure that the minute conda environment is active by running `conda env list`. The active environment will have a '*' next to it.
    ▲ **CRITICAL STEP** The minute conda environment needs to be active to be able to run the software. If you perform this workflow at different times or close the terminal that you were using, you will need to reactivate the conda environment.

Reference genome setup

80. Download the human hg38 reference genome sequence to the hg38_reference directory under your working directory. If you are running minute with the provided example dataset and went through the configuration steps to download the files, your working directory should be named 'example_run':

```
mkdir hg38_reference
cd hg38_reference
wget https://hgdownload.soe.ucsc.edu/goldenPath/hg38/bigZips/hg38.fa.gz
```

81. (Optional) Download the hg38 reference blacklist annotation from ENCODE[36]. Although it is considered good practice to remove these regions from the downstream analysis, it is possible to skip this step and leave this empty in the configuration file minute.yaml (see Step 83).

```
wget https://github.com/Boyle-Lab/Blacklist/raw/master/lists/hg38-blacklist.v2.bed.gz
gunzip hg38-blacklist.v2.bed.gz
```

82. Build the bowtie2 index of the reference sequence, needed by bowtie2 for alignment to the reference genome, on the same directory where you downloaded the reference. The pipeline expects these files to be located in the same directory.

```
bowtie2-build hg38.fa.gz hg38
```

    ◆ **TROUBLESHOOTING**

Run configuration and execution

83. Copy your FASTQ files to a data directory. If you are running the workflow with the example data that you downloaded in the setup section, these will be in the example_data directory.

```
cd ..
mkdir data
cd data
cp ./example_data/*.fastq.gz .
```

▲ **CRITICAL STEP** FASTQ files obtained from certain Illumina platforms may end in `_R1_001.fastq.gz`, `_R2_001.fastq.gz`. Rename them with `_R1.fastq.gz`, `_R2.fastq.gz` before running the pipeline.

84. Fill in the sample barcodes.tsv (see Table 1 for a description of its contents). If you are running the example data, this file is also in the example_data directory:

```
cd ..
cp example_data/barcodes.tsv .
```

85. Initialize your minute run by specifying the barcodes.tsv file from Step 84, the directory where you copied the FASTQ files and which FASTQ pair is the input sample (the name without `_R1.fastq.gz/_R2.fastq.gz`):

```
minute init my_experiment --reads data --barcodes barcodes.tsv --input INP
```

This will create a directory named 'my_experiment' in your current working directory, and create inside it:

- A fastq directory pointing to the location of the original data directory
- A pair of files, groups.tsv, libraries.tsv, ready to use
- A template minute.yaml to configure the parameters of the run (Fig. 3b).

◆ **TROUBLESHOOTING**

86. Edit the template minute.yaml with relevant values for your experiment and the location of your bowtie2 index files and FASTA reference genome. Note that there should be an entry under references matching the values in your barcodes.tsv, groups.tsv. Your minute.yaml file should be located in the my_experiment directory (Fig. 2b). If you are running the example dataset, you can copy an already completed minute.yaml (Box 2) in the example_data directory:

```
cd my_experiment
cp ./example_data/minute.yaml .
```

▲ **CRITICAL STEP** Make sure that you fill in the correct UMI length for your experiment. Otherwise, reads will not be properly demultiplexed.

87. Run minute in dry-run mode to show a summary of what will be run before starting the workflow. Some possible errors in configuration files can be spotted already on the dry run:

```
minute run --dry-run
```

◆ **TROUBLESHOOTING**

88. Run minute:

```
minute run
```

◆ **TROUBLESHOOTING**

BOX 2

Sample minute.yaml configuration file

Provided sample minute.yaml configuration file. The entry 'references' must have at least one entry with a path to the corresponding FASTA file and bowtie2-index, together with an optional 'exclude' entry with the path to a specified blacklist. In this example, the path is relative to the directory where the minute pipeline is run.

```
# Configuration settings for the minute pipeline
# Paths (unless absolute) are relative to the
directory in which snakemake is run.
references:
 hg38:
  # Path to a reference FASTA file (may be
  gzip-compressed).
  # A matching Bowtie2 index must exist in the
  same location.
  fasta: "./hg38_reference/hg38.fa.gz"
  exclude: "./hg38_reference/hg38.blocklist.bed"

# Length of the 5' UMI
umi_length: 6
# Fragment length (insert size)
fragment_size: 150
# Max barcode errors allowed
max_barcode_errors: 1
```

89. Check the final results and reports. The output of a minute run is a final directory and a reports directory. The most relevant files are the final/bigwig ones containing scaled bigWig tracks for visualization and a MultiQC HTML report (Fig. 3c). See Output files for a more detailed description of the output files.
◆ **TROUBLESHOOTING**

Troubleshooting

Troubleshooting advice can be found in Table 2.

Table 2 | Troubleshooting table

Step	Problem	Possible reason	Solution
4	Digested chromatin is not a mono- and dinucleosomal population	Over- or under-digestion of chromatin by MNase	Ensure that the cell pellets are resuspended into a homogeneous solution
14, 22, 23, 40	The amount of immunoprecipitation DNA is too low to measure	Highly specific immunoprecipitation with little background	Can occur for low-abundant targets. Continue to IVT by using all immunoprecipitation material
		Low saturation of beads with antibody	Check that the best bead type for your antibody (e.g., protein A, protein G, protein A/G or species-specific antibody beads) has been used
			Ensure that the antibody is efficiently captured on the beads (it should be depleted from the supernatant; this can be checked by western blot detection of IgG) and prolong antibody coupling time to beads if necessary

Table 2 (continued) | Troubleshooting table

Step	Problem	Possible reason	Solution
		Poor immunoprecipitation efficiency	Optimize the amount of antibody-coupled beads to use for your target of interest. Assess immunoprecipitation efficiency by western blot by comparing ChIP target abundance before and after the immunoprecipitation reaction. Prolong the immunoprecipitation duration to overnight incubation at 4 °C, adjust the antibody amount or repeat the experiment by using an antibody with a higher binding affinity if needed
		Low abundance of the target protein in the samples	Consider using a higher volume of the pooled material
		Instability of histone modification	Consider adding appropriate inhibitors from the cell lysis step to preserve the histone modification of interest
		Stringent wash conditions	Reduce wash times (<1 min). Consider removing the high-salt wash buffer
		Low cross-linking efficiency of the target protein with formaldehyde	Consider higher formaldehyde concentration (up to 4% (wt/vol)) or additional cross-linking agents with longer-range linkers such as ethylene glycol bis(succinimidyl succinate) (EGS)
51	Non-measurable amount of RNA	Low amount of immunoprecipitation DNA	If the immunoprecipitation DNA was already low, then the presence of amplifiable DNA molecules may be a limiting factor. For troubleshooting, refer to 'The amount of immunoprecipitation DNA is too low to measure' above
		Inefficient end-repair/ligation of adapter barcoding in Steps 7–9	Check to see if chromatin was not over-digested (see Supplementary Methods, 'fragment optimization'). Check if the correct amount of adapters was added. Consider starting with more cells
		Inefficient in vitro transcription reaction	Check the expiration date, storage condition and freeze/thaw cycle of the IVT reagents and prolong the IVT incubation time at 37 °C for up to 16 h
64	The amount of cDNA is not detectable by Nanodrop fluorometer	Inefficient RNA 3′ adapter ligation or RT reaction	Check RNA ligation and RT reagents
		Degradation of RNA	Ensure that the IVT and RNA 3′ ligation reactions are supplemented with RNase inhibitors
77	Very low concentration of final library	The RT or PCR reaction did not work	Check the expiration date, storage condition and freeze/thaw cycle of RT and PCR reagents
		Oversaturated PCR reaction with too much cDNA	Take forward a maximum of 200 ng of cDNA per reaction
	The final profile of the library is not enriched at ~300 bp	Too few PCR cycles	Continue to Illumina sequencing as long as the concentration of the DNA is sufficiently high to load on a flow cell. Assume a 300-bp fragment size when calculating your molarity for preparing your final library dilution
			Increase libray amplification up to 12 PCR cycles. More PCR cycles also risk a higher duplication rate
		Under-digestion of input chromatin leads to predominantly large fractions	Perform a second PCR (will bias the library for smaller fragments) or perform double-sided size selection by using the Ampure XP beads
	Strong spike(s) in Bioanalyzer profile at ~150 and 200 bp	High amount of adapter contamination (adapter monomers and dimers)	It is still possible to sequence the library, but a large fraction of reads may be filtered out by the pipeline; hence, more sequencing depth is needed
			Perform size selection by using the Ampure XP beads
82	Failed to run bowtie2-build	The minute conda environment is not active	Activate the conda environment with conda activate minute
85	Failed to run minute init	The minute conda environment is not active	Activate the conda environment with conda activate minute
		wrong or missing --reads parameter	Ensure that the --reads parameter is specified and points to an existing directory
		wrong or missing --barcodes parameter	Ensure that the --barcodes parameter is specified and that the specified file exists
87, 88	Failed to run minute run/ minute run --dry-run	The minute conda environment is not active	Activate the conda environment with conda activate minute
		Missing configuration files	Check that the current working directory (pwd) is the minute_run directory
			Ensure that the minute_run directory contains libraries.tsv, minute.yaml, groups.tsv and a fastq directory with FASTQ files
		FASTQ files are incorrectly named	Check that your fastq/ directory contains paired-end FASTQ files with names ending in _R1.fastq.gz, _R2.fastq.gz
		Incorrect parameter setting	Check the detailed user guide on minute documentation for details on valid parameters: https://minute.readthedocs.io/en/latest/guide/

Table 2 (continued) | Troubleshooting table

Step	Problem	Possible reason	Solution
89	Very low number of reads per barcode	Incorrectly specified barcode sequence	Verify that the sequences provided on the barcodes.tsv file match the sample barcode sequence used to identify the affected sample
		Incorrect UMI length	UMI length is easily inferred from the 'per base sequence content' section in the MultiQC report
	High adapter content	Over-digestion of chromatin. Inefficient repair and adapter ligation. Inefficient size selection on Ampure XP beads. Insufficient enrichment of the ChIP target	Refer to the solutions above for the adapter ligation, library preparation and quality control. Consider increasing the input cell number per sample (although we do not recommend going beyond 5 Mio)
	Larger or smaller insert size than expected	Over- or under-digestion of chromatin by MNase	Titrate the amount of MNase and the duration of digestion required to achieve the desired digestion pattern. Over-digestion, in particular, must be avoided to ensure the success of the protocol
	High duplication rate/low diversity	Over-digestion of chromatin. Inefficient repair and adapter ligation. Insufficient enrichment of the ChIP target. Problems in any step of the library preparation	Refer to the solutions above for adapter ligation, immunoprecipitation troubleshooting and library preparation. High adapter content, especially in the input sample, may inhibit amplification of genomic fragments

Timing

Part 1: Steps 1–9, sample preparation (day 1)
Steps 1–6, lysis and fragmentation: 1 h
Steps 7–9, end-repair and ligation: overnight

Part 2: Steps 10–40, sample pooling and immunoprecipitation (days 1 and 2)
Steps 10–14, pre-coupling of the antibody to the magnetic beads: 1 h
Steps 15–17, pooling of samples: 10 min
Steps 18–22, immunoprecipitation: 6 h
Steps 23–28, washing: 30 min
Steps 29–32, elution from beads and protein digestion: 1 h/overnight (~16 h) for fixed materials
Steps 33–40, AMPure XP bead cleanup to isolate DNA: 30 min

Part 3: Steps 41–43, sequencing library preparation (days 2–3)
Steps 41–52, in vitro transcription, RNA cleanup and RNA adapter ligation
Steps 41–42, in vitro transcription: 4 h
Steps 43–52, RNAclean XP cleanup to isolate RNA and RNA 3′ adapter ligation: 1.5 h
Steps 53–64, RT and cDNA cleanup
Steps 53–55, RT: 1 h
Steps 56–64, RNAclean XP bead cleanup: 30 min
Steps 65–76, library preparation
Steps 65–67, library PCR: 1 h
Steps 68–76, AMPure XP bead cleanup: 30 min
Step 77, Illumina sequencing setup: variable depending on the resources available and the sequencing depth (1 h)

Part 4: Steps 78–89, data analysis: 4 h plus processing time

Anticipated results

Sample lysis and fragmentation (Steps 1–6) should yield mono- and dinucleosomal fragments of chromatin (Fig. 1b). The initial number of cells should be similar across conditions to ensure comparable fragmentation across all samples that go into the pool. It should, however, be noted that the relative quantitation of samples performed by the pipeline takes into account the actual input distribution of sample barcodes. Hence, the correct scaling is not affected by the over- or under-representation (e.g., due to variation in input cell counts) of individual samples in the pool. The fragmentation pattern can be further validated on the basis of the insert size distribution obtained after analysis in the MultiQC report (see example Fig. 4b).

The efficiency of the end-repair and ligation (Steps 7–12) cannot be directly measured but should result in a fraction of chromatin fragments being ligated to an adapter molecule. On an analytical gel performed from the ligation reaction, one can usually observe both the chromatin and excess unligated adapter but not a specific chromatin ligation product.

Pre-coupling the antibody to the magnetic beads (Steps 13–15) should result in the conjugation of the antibody to the magnetic beads for use in the subsequent ChIP reaction. By analyzing aliquots of the antibody dilution used for coupling and of the supernatant after coupling by SDS–PAGE, the successful binding of the antibody should be evident from a visible reduction or depletion of the heavy- and light-chain specific bands.

The remaining immunoprecipitation and library-preparation steps of the protocol yield highly variable amounts of material depending on the target PTM/protein abundance, immunoprecipitation efficiency and specificity and amount of background binding. In our experience, most histone modifications yield measurable amounts of immunoprecipitation DNA; however, cross-linked target proteins and rare histone PTMs may yield DNA amounts too low to measure. Regardless, it is critical that the in vitro transcription (Steps 32–34) results in a measurable amount of RNA owing to the amplification of the immunoprecipitation material. The amount of cDNA produced by RT (Steps 43–53) is measured by using a Nanodrop fluorometer and should be sufficient to approximate the amount of template (≤200 ng) required for library PCR (Steps 54–56). The final library should have a fragment size distribution of 200–350 bp (see example Fig. 1c) and ≤2 µg of DNA.

The processing and analysis pipeline (Steps 65–75) generates many useful quality control parameters (such as the fragment size profile of each sample and immunoprecipitation, the mapping and duplication rates) that can give indications for troubleshooting the experimental workflow. Table 3 provides a range of expected values for different steps of the protocol.

As part of a MultiQC summary (Fig. 3c), the pipeline provides statistics on how many reads were assigned to each sample barcode from each library and the deduplication statistics, including the extrapolation of the total number of unique molecules (diversity per library size) for each sample barcode. If replicates are performed, the unique molecules from each replicate are additive; hence, the sequencing depth of the combined (pooled) track is proportionally higher. Furthermore, the MultiQC summary generates paired-end fragment-size density plots (Fig. 4b), which can provide information on the selection of fragments during immunoprecipitation and additional standard metrics like paired-end mappability statistics.

A MINUTE-ChIP–specific global quantification plot is produced for each scaling group (Figs. 3d and 4c). The plot shows the relative global levels for each of the probed PTMs and/or chromatin-bound factors present in each sample/replicate, as compared to the user-defined reference condition. The global quantification is an important quality control metric, because the genome-wide tracks are scaled according to these values (effectively, the average genome

Table 3 | Typical values of quality control indicators

Quality control measurement	Typical range	Comment
DNA yield after immunoprecipitation	2–200 ng	Depends on epitope abundance and antibody efficiency
RNA yield after IVT	1–70 µg	Range can vary greatly and is not indicative of success of ChIP
cDNA yield after RT	0.1–5 µg	Usually limited by the capacity of the RT reaction
Library PCR yield	1–2 µg	Usually stable when input cDNA amount is normalized to 200 ng
Picard average insert size	150–200 bp	Expected distribution is mostly mononucleosomal
Library diversity (per sample)	10–150 Mio	Depends on adapter ligation efficiency, epitope abundance and antibody efficiency

coverage of each sample track will equal the global value plotted in this graph). Replicates should agree on the global quantification, and a large variation between replicates should call for an investigation of other quality contro parameters that could explain a technical problem during the initial stages of the sample preparation (e.g., insufficient lysis, severe under- or over-digestion of one of the replicates or a barcode mixup). Global differences between conditions revealed in the MINUTE-ChIP experiment should correlate with other experimental observations obtained, for example, by western blot, immunofluorescence staining or mass spectrometry.

As discussed in the 'Experimental design' section, including negative control samples is extremely useful for QC of the antibody used; in addition, it provides additional valuable insights for a quantitative interpretation of the ChIP–seq tracks and deriving mechanistic hypotheses. As illustrated by our demonstration data (Fig. 3d,e), pharmacological depletion of H3K27me3 abrogates up to 97% of the H3K27me3 ChIP–seq signal observed in untreated naive hES cells. This provides empirical proof that MINUTE-ChIP can operate with very little technical background and shows that H3K27me3 broadly covers the genome (Fig. 2c), even outside of classical Polycomb domains. Such insight cannot be gained from ChIP–seq, CUT&RUN or CUT&Tag experiments unless appropriate spike-in strategies are used.

MINUTE-ChIP signal always reflects the relative nucleosome-level abundance of histone PTMs or chromatin-bound fraction of a transcription factor and can accurately capture genome-wide changes (i.e., gains or losses), as exemplified by pair-wise comparisons of naive and primed pluripotent mES cells, as well as naive and primed pluripotent hES cells[20,21] (Fig. 3d,e). The same quantitative principles apply to transcription factors, but extra care must be taken to ensure that technical factors (e.g., uneven cross-linking efficiency between samples) do not confound the quantitative comparison. Immunoprecipitation efficiency is usually much lower for cross-linked protein targets, leading to lower diversity libraries (often 5–10 Mio unique reads per sample/replicate). Thus, it is advisable to increase the input cell number for low-abundant protein targets. Here, we provide an example dataset showing a NANOG MINUTE-ChIP experiment with cross-linked and sonicated chromatin from naive and differentiated mES cells that do not express NANOG as negative control (Fig. 4). Sonication typically results in a wider fragment size distribution around 150–200 bp (Fig. 4a,b). Technical background is generally larger for cross-linked starting material, here ~7% (Fig. 4c). NANOG ChIP shows high specific enrichment over known NANOG peaks but not in random regions of the genome (Fig. 4d–f), and replicates agree well (Fig. 4e,f).

In summary, MINUTE-ChIP provides the possibility to perform complex experiments with quantitative comparisons supported by a sufficient number of replicates and control conditions to arrive at robust and trustable results in a single workflow.

Data availability

The main data referred to in this protocol are available in the supporting primary research papers[20,21]. Demultiplexed FASTQ files and scaled bigWig tracks of MINUTE-ChIP data used in this protocol are available in the Genome Expression Omnibus (GEO) database under accession numbers GSE133056 (Fig. 2b), GSE126252 (Fig. 2c), GSE181241 (Fig. 3c–e) and GSE244908 (Fig. 4b–f). Refer to Supplementary Table 2 for a detailed dataset description. An example dataset for running the data analysis workflow is provided at https://figshare.scilifelab.se/articles/dataset/MINUTE-ChIP_example_data/25348405. Source data are provided with this paper.

Code availability

The primary analysis for MINUTE-ChIP sequencing data described in this protocol is performed by using minute, a Snakemake workflow with source code available on GitHub at https://github.com/elsasserlab/minute. Thorough documentation of minute's full functionality is available at https://minute.readthedocs.io.

Received: 29 September 2023; Accepted: 29 July 2024;
Published online: 3 October 2024

References

1. Robertson, G. et al. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat. Methods* **4**, 651–657 (2007).
2. Barski, A. et al. High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).
3. Johnson, D. S., Mortazavi, A., Myers, R. M. & Wold, B. Genome-wide mapping of in vivo protein-DNA interactions. *Science* **316**, 1497–1502 (2007).
4. Mikkelsen, T. S. et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
5. Cuddapah, S. et al. Native chromatin preparation and Illumina/Solexa library construction. *Cold Spring Harb. Protoc.* **2009**, pdb.prot5237 (2009).
6. O'Neill, L. P. & Turner, B. M. Immunoprecipitation of native chromatin: NChIP. *Methods* **31**, 76–82 (2003).
7. Kasinathan, S., Orsi, G. A., Zentner, G. E., Ahmad, K. & Henikoff, S. High-resolution mapping of transcription factor binding sites on native chromatin. *Nat. Methods* **11**, 203–209 (2014).
8. Jin, H. et al. ChIPseqSpikeInFree: a ChIP-seq normalization approach to reveal global changes in histone modifications without spike-in. *Bioinformatics* **36**, 1270–1272 (2020).
9. Polit, L. et al. CHIPIN: ChIP-seq inter-sample normalization based on signal invariance across transcriptionally constant genes. *BMC Bioinforma.* **22**, 407 (2021).
10. Liang, K. & Keleş, S. Normalization of ChIP-seq data with control. *BMC Bioinforma.* **13**, 199 (2012).
11. Chen, K. et al. The overlooked fact: fundamental need for spike-in control for virtually all genome-wide analyses. *Mol. Cell. Biol.* **36**, 662–667 (2015).
12. Orlando, D. A. et al. Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. *Cell Rep.* **9**, 1163–1170 (2014).
13. Descostes, N., Tsirigos, A. & Reinberg, D. ChIPSeqSpike: A R/Bioconductor package for ChIP-Seq data scaling according to spike-in control. Preprint at *bioRxiv* https://doi.org/10.1101/269118 (2018).
14. Blanco, E., Di Croce, L. & Aranda, S. SpikChIP: a novel computational methodology to compare multiple ChIP-seq using spike-in chromatin. *NAR Genom. Bioinform.* **3**, lqab064 (2021).
15. Grzybowski, A. T., Chen, Z. & Ruthenburg, A. J. Calibrating ChIP-Seq with nucleosomal internal standards to measure histone modification density genome wide. *Mol. Cell* **58**, 886–899 (2015).
16. Lara-Astiaso, D. et al. Chromatin state dynamics during blood formation. *Science* **345**, 943–949 (2014).
17. van Galen, P. et al. A multiplexed system for quantitative comparisons of chromatin landscapes. *Mol. Cell* **61**, 170–180 (2016).
18. Chabbert, C. D., Adjalley, S. H., Steinmetz, L. M. & Pelechano, V. Multiplexed ChIP-Seq using direct nucleosome barcoding: a tool for high-throughput chromatin analysis. *Methods Mol. Biol.* **1689**, 177–194 (2018).
19. Arrigoni, L. et al. RELACS nuclei barcoding enables high-throughput ChIP-seq. *Commun. Biol.* **1**, 214 (2018).
20. Kumar, B. & Elsässer, S. J. Quantitative multiplexed ChIP reveals global alterations that shape promoter bivalency in ground state embryonic stem cells. *Cell Rep.* **28**, 3274–3284.e5 (2019).
21. Kumar, B. et al. Polycomb repressive complex 2 shields naïve human pluripotent cells from trophectoderm differentiation. *Nat. Cell Biol.* **24**, 845–857 (2022).
22. Weigert, R. et al. Dynamic antagonism between key repressive pathways maintains the placental epigenome. *Nat. Cell Biol.* **25**, 579–591 (2023).
23. Shao, R. et al. Distinct transcription kinetics of pluripotent cell states. *Mol. Syst. Biol.* **18**, e10407 (2022).
24. Stolz, P. et al. TET1 regulates gene expression and repression of endogenous retroviruses independent of DNA demethylation. *Nucleic Acids Res.* **50**, 8491–8511 (2022).
25. Walter, L. D., Van Galen, P., Bernstein, B. E. & Epstein, C. B. Mint-ChIP3: a low-input ChIP-seq protocol using multiplexed chromatin and T7 amplification v1. Available at https://www.protocols.io/view/mint-chip3-a-low-input-chip-seq-protocol-using-mul-rm7vznk84vx1/v1 (2018).
26. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
27. Navarro, C., Martin, M. & Elsässer, S. minute: a MINUTE-ChIP data analysis workflow. Preprint at *bioRxiv* https://doi.org/10.1101/2022.03.14.484318 (2022).
28. Mölder, F. et al. Sustainable data analysis with Snakemake. *F1000Res.* **10**, 33 (2021).
29. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
30. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
31. Girardot, C., Scholtalbers, J., Sauer, S., Su, S.-Y. & Furlong, E. E. M. Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. *BMC Bioinforma.* **17**, 419 (2016).
32. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
33. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
34. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinforma.* **14**, 178–192 (2013).
35. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
36. Amemiya, H. M., Kundaje, A. & Boyle, A. P. The ENCODE blacklist: identification of problematic regions of the genome. *Sci. Rep.* **9**, 9354 (2019).
37. Zheng, R. et al. Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. *Nucleic Acids Res.* **47**, D729–D735 (2019).
38. King, H. W. & Klose, R. J. The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells. *eLife* **6**, e22631 (2017).

Acknowledgements

S.J.E. was funded by Vetenskapsrådet, Sweden (2020–04313); the Ming Wai Lau Center for Reparative Medicine, Sweden; the Ragnar Söderbergs Stiftelse, Sweden; Stiftelsen för Strategisk Forskning (FFL7); and the Knut och Alice Wallenbergs Stiftelse, Sweden. C.N. was funded by Karolinska Institutet KI Research Foundation grants (2022-01826). We thank the National Academic Infrastructure for Supercomputing in Sweden (NAISS) funded by the Swedish Research Council through grant agreement no. 2022-06725. The Supercomputing Cluster UPPMAX was used under projects NAISS 2023/6-19, NAISS 2023/22-84, SNIC 2022/6-14 and NAISS 2024/22-108. M.M. was supported by the Knut and Alice Wallenberg Foundation as part of the National Bioinformatics Infrastructure Sweden at SciLifeLab. We thank members of the Elsässer laboratory, especially R. Caridha and E. Soto Afonso for feedback on the protocol.

Author contributions

B.K., P.Y.K.Y., C.N. and S.J.E. wrote the protocol. J.L., A.S.M., A-M.K., H.S. and M.M. gave input to the protocol. B.K., C.N. and H.S. contributed example data and figures. C.N. and M.M. wrote the analysis software.

Competing interests

B.K., P.Y.K.Y. and S.J.E. hold shares in Epigenica AB, Sweden, a company operating in epigenomics. B.K. and C.N. are employees of Epigenica AB. B.K., P.Y.K.Y. and S.J.E. are inventors of patents in multiplexed ChIP–seq technology owned by Epigenica AB.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41596-024-01058-z.

Correspondence and requests for materials should be addressed to Simon J. Elsässer.

Peer review information *Nature Protocols* thanks Pradeepa Madapura, Judith Zaugg and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Related links

Key references using this protocol

Kumar, B. & Elsässer, S. J. *Cell Rep.* **28**, 3274–3284.e5 (2019): https://doi.org/10.1016/j.celrep.2019.08.046
Kumar, B. et al. *Nat. Cell Biol.* **24**, 845–857 (2022): https://doi.org/10.1038/s41556-022-00916-w
Weigert, R. et al. *Nat. Cell Biol.* **25**, 579–591 (2023): https://doi.org/10.1038/s41556-023-01114-y

[1]Science for Life Laboratory, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Solna, Sweden. [2]Ming Wai Lau Centre for Reparative Medicine, Stockholm node, Karolinska Institutet, Solna, Sweden. [3]Epigenica AB, A Working Lab, Solna, Sweden. [4]Department of Biochemistry and Biophysics, National Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Stockholm University, Solna, Sweden. [5]These authors contributed equally: Banushree Kumar, Carmen Navarro, Philip Yuk Kwong Yung.