Published online 17 September 2021

Nucleic Acids Research, 2021, Vol. 49, No. 21 ***e125***
https://doi.org/10.1093/nar/gkab813

FACT-seq: profiling histone modifications in formalin-fixed paraffin-embedded samples with low cell numbers

Linxuan Zhao[1,†], Pengwei Xing[1,†], Vamsi Krishna Polavarapu[1,†], Miao Zhao[1], Blanca Valero-Martínez[1], Yonglong Dang[1], Nagaprathyusha Maturi[2], Lucy Mathot[1], Inês Neves[2], Irem Yildirim[2], Fredrik Johansson Swartling[1], Tobias Sjöblom[1], Lene Uhrbom[2] and Xingqi Chen[1,3,*]

[1]Department of Immunology, Genetics and Pathology, Uppsala University, 75108 Uppsala, Sweden, [2]Department of Immunology, Genetics and Pathology, Uppsala University and Science for Life Laboratory, Rudbeck Laboratory, SE-75185 Uppsala, Sweden and [3]Beijer Laboratories, Uppsala University, Uppsala, Sweden

Received May 19, 2021; Revised August 16, 2021; Editorial Decision September 03, 2021; Accepted September 06, 2021

ABSTRACT

The majority of biopsies in both basic research and translational cancer studies are preserved in the format of archived formalin-fixed paraffin-embedded (FFPE) samples. Profiling histone modifications in archived FFPE tissues is critically important to understand gene regulation in human disease. The required input for current genome-wide histone modification profiling studies from FFPE samples is either 10–20 tissue sections or whole tissue blocks, which prevents better resolved analyses. But it is desirable to consume a minimal amount of FFPE tissue sections in the analysis as clinical tissues of interest are limited. Here, we present FFPE tissue with antibody-guided chromatin tagmentation with sequencing (FACT-seq), the first highly sensitive method to efficiently profile histone modifications in FFPE tissues by combining a novel fusion protein of hyperactive Tn5 transposase and protein A (T7–pA–Tn5) transposition and T7 *in vitro* transcription. FACT-seq generates high-quality chromatin profiles from different histone modifications with low number of FFPE nuclei. We proved a very small piece of FFPE tissue section containing ∼4000 nuclei is sufficient to decode H3K27ac modifications with FACT-seq. H3K27ac FACT-seq revealed disease-specific super enhancers in the archived FFPE human colorectal and human glioblastoma cancer tissue. In summary, FACT-seq allows decoding the histone modifications in archival FFPE tissues with high sensitivity and help researchers to better understand epigenetic regulation in cancer and human disease.

INTRODUCTION

Epigenetic profiling is critically important for improving our understanding of the molecular mechanisms of human disease and cancer (1–5). In recent decades, the technologies of chromatin immunoprecipitation with sequencing (ChIP-seq) and its variations have helped us to generate large-scale epigenetic profiles in both basic research, epigenetic cancer biomarker discovery and preclinical studies (6–8). The majority of biopsies in both basic research and translational cancer studies are preserved in the format of archived formalin-fixed paraffin-embedded (FFPE) samples (9–11); it was reported that large amounts of FFPE specimens are newly archived every year worldwide (12,13). Accordingly, profiling epigenetic regulation in archived FFPE tissue can be invaluable for translational cancer research and potential epigenetic cancer biomarker discovery. The successful application of ChIP-seq to FFPE tissues, pathology tissue chromatin immunoprecipitation (PAT-ChIP) (14,15), fixed-tissue chromatin immunoprecipitation sequencing (FiT-seq) (7), fixed-tissue ChIP-seq for H3K27 acetylation (H3K27ac) profiling (FiTAc-seq) (16) and similar technologies (17), makes it possible to map histone modifications in clinically archived FFPE tissue; however, the required input for these technologies from FFPE samples is either 10–20 tissue sections or whole tissue blocks (7,14–16,18), which prevents better resolved analyses and their applications. However, it is desirable to consume a minimal amount of clinical FFPE tissue samples in the analysis, as clinical materials are very limited and difficult to collect. In addition, sonication is applied in these available

*To whom correspondence should be addressed. Tel: +46 18471 4072; Email: xingqi.chen@igp.uu.se
†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

technologies and can potentially introduce sequence bias (19). To overcome the sensitivity limitation of ChIP-seq technology, a series of high-sensitivity technologies (20–40), including ULI-NChIP (41), STAR-ChIP (40), MOWChIP (31), Drop-ChIP (33), ChIL-seq (23), CUT&RUN (26), scChIC-seq (28) and many others, have been developed and widely used to map the epigenetic profiles and transcription factor binding sites. Among these highly sensitive technologies, strategies based on the fusion of hyperactive-Tn5 transposase and protein A (pA–Tn5), Cleavage Under Targets and Tagmentation (CUT&Tag) (20,42), combinatorial barcoding and targeted chromatin release (CoBATCH) (22), and the other technologies (21–26,43), have been well adopted for different fresh or mild fixed cells, and fresh or frozen tissues due to its high efficiency (34–37). Because of the extensive fixation and paraffin embedding in the FFPE tissues, the DNA breaks and nicks are dominant in the extracted DNA from the FFPE tissue (44,45), which makes it impossible to directly apply currently available, highly sensitive epigenetic profiling technologies in the FFPE tissue blocks. To this end, we developed FFPE tissue with antibody-guided chromatin tagmentation with sequencing (FACT-seq), the first highly sensitive method to efficiently profile the histone modifications of FFPE samples with low cell numbers by combining a novel fusion protein of hyperactive Tn5 transposase and protein A (T7−pA−Tn5) transposition and T7 *in vitro* transcription.

MATERIALS AND METHODS

Cell culture

GM12878 cells were grown in RPMI 1640 (Gibco, 31870-025), 2 mM L-glutamine (Gibco, 25030-024), 15% fetal bovine serum (Gibco, 10270-106) and 1% Pen/Strep (Gibco, 15140-122).

Animals

The mouse tissues were from the 8-week-old FVBN mice, housed in individually ventilated cages (3–5 animals per cage) in accordance with Uppsala University regulations on mice with appropriate organic bedding, paper house enrichments, food and water *ad libitum* and 12/12-h light/dark cycle. All experiments were performed in accordance with national guidelines and regulations, and with the approval of the animal care and use committees at Uppsala University.

Mouse tissue collection

8-Week-old mice were sacrificed via inhalation euthanasia, and mouse organs (livers and kidneys) were collected. For frozen samples, mouse organs were snap-frozen on dry ice and stored at −80°C. For FFPE sample, mouse organs were fixed with formalin overnight, and then washed with phosphate-buffered saline (PBS) and kept in 70% ethanol for paraffin embedding. Fixed mouse organs were routinely processed, and paraffin embedded.

Transposase adaptor sequences

The oligonucleotides for Tn5, pA–Tn5 and T7−pA−Tn5 transposase adaptors were synthesized at INTERGATED DNA TECHNOLOGIES (IDT), and sequences of oligonucleotide used in pA–Tn5 and Tn5 are as follows:

Tn5MErev, 5′-[phos] CTGTCTCTTATACACATCT-3′,
Tn5ME-A, 5′TCGTCGGCAGCGTCAGATGTGTAT AAGAGACAG-3′,
Tn5ME-B, 5′-GTCTCGTGGGCTCGGAGATGTGTA TAAGAGACAG-3′.
Sequences of oligonucleotide used in T7−pA−Tn5 are as follows:
Tn5MErev, 5′-[phos]CTGTCTCTTATACACATCT-3′,
T7-Tn5ME, 5′-CATGAGATT AATACGACTCACTATA GGGAGAAGATGTGTATAAGAGACAG-3′.

PCR primers for sequencing library preparation

The PCR primers were synthesized at INTERGATED DNA TECHNOLOGIES (IDT), and sequences of primers were used by referring to the previous report (46).

Hyperactive pA–Tn5 and Tn5 production

Hyperactive pA–Tn5 and Tn5 were produced as previously described (20,47). In brief, psfpATn5 and pTXB1-Tn5 plasmid (Addgene, 60240) were introduced into T7 Express LysY/Iq *Escherichia coli* strain (New England Biolabs, C3013) separately. 10 ml of overnight cultured *E. coli* was inoculated to 500 ml LB medium. After incubation for 1.5 h at 37°C, bacteria were incubated about 2.5 h at RT. When the OD600 = 0.9, pA–Tn5 and Tn5 protein were induced by adding 0.25 mM IPTG for 4 h. *E. coli* pellet was resuspended in lysis buffer (20 mM HEPES–KOH (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016) pH 7.2, 0.8 M NaCl (Invitrogen, AM9759), 1mM EDTA (Invitrogen, AM9260G), 10% glycerol (Sigma-Aldrich, G9012), 0.2% Triton X-100 (Sigma-Aldrich, T8787), complete proteinase inhibitor (11697498001, Roche)) and lysed by sonication. 10% PEI was added to supernatant of lysate to remove bacterial genomic DNA. 10 ml chitin resin (New England Biolabs, S6651L) was added to the supernatant and incubated with rotating for 1 h at 4°C. The resin washed by lysis buffer extensively. In order to cleave Tn5 or pA–Tn5 protein from intein, lysis buffer containing 100 mM DTT was added to the resin and stored in 4°C. After 48 h, protein was eluted by gravity flow and collected in 1 ml fractions. 1 ul of each fraction was added to detergent compatible Bradford assay (Thermo Fisher Scientific, 23246) and peaked fractions were pooled and dialyzed against 2× dialysis buffer (100 mM HEPES–KOH (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016) at pH 7.2, 0.2 M NaCl (Invitrogen, AM9759), 0.2 mM EDTA (Invitrogen, AM9260G), 2 mM DTT (Thermo Fisher scientific, 20291), 0.2% Triton X-100 (Sigma-Aldrich, T8787), 20% glycerol (Sigma-Aldrich, G9012)). Dialyzed Tn5 or pA–Tn5 protein were concentrated by using ultracel 30-K column (Millipore, UFC903024) and the quantity of Tn5 and pA–Tn5 were measured by Bradford assay and visualized on

NuPAGE Novex 4–12% Bis–Tris gel (Thermo Fisher Scientific, NP0321) followed by Coomassie blue staining.

pA–Tn5 transposase assembly

Oligonucleotides (Tn5ME-A, Tn5ME-B, Tn5MErev) were resuspended in oligo annealing buffer (10 mM Tris–HCl pH 8.0 (Invitrogen, 15568-025), 50 mM NaCl (Invitrogen, AM9759), 1 mM EDTA (Invitrogen, AM9260G)) to a final concentration of 100 μM each. Equimolar amounts of Tn5MErev/Tn5ME-A and Tn5MErev/Tn5ME-B were mixed in separate 200 μl PCR tubes. Then, the adaptors were annealed on the PCR machine with the following PCR program (95°C for 5 min first, then the temperature was slowly ramped down to 25°C with the rate of −0.1°C/s, 25°C for 5 min). The pA–Tn5 transposase was assembled with the following components: 0.04 vol. Tn5MErev/Tn5ME-A, 0.04 vol Tn5MErev/Tn5ME-B, 0.4 vol. 100% glycerol (Sigma-Aldrich, G9012), 0.3116 vol. 2× dialysis buffer (100 mM HEPES−KOH (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016) at pH 7.2, 0.2 M NaCl (Invitrogen, AM9759), 0.2 mM EDTA (Invitrogen, AM9260G), 2 mM DTT (Thermo Fisher scientific, 20291), 0.2% Triton X-100 (Sigma-Aldrich, T8787), 20% glycerol (Sigma-Aldrich, G9012)), 0.0362 vol pure pA–Tn5 (55.55 μM), 0.1722 vol water (Invitrogen, AM9932). The reagents were mixed thoroughly but gently, and the mixture was left on the bench at RT for 1 h to allow annealing of oligos to pA–Tn5.

Tn5 transposase assembly

Oligonucleotides (Tn5ME-A, Tn5ME-B, Tn5MErev) were resuspended in oligo annealing buffer (10 mM Tris−HCl pH 8.0 (Invitrogen, 15568-025), 50 mM NaCl (Invitrogen, AM9759), 1 mM EDTA (Invitrogen, AM9260G)) to a final concentration of 100μM each. Equimolar amounts of Tn5MErev/Tn5ME-A and Tn5MErev/Tn5ME-B were mixed in separate 200 μl PCR tubes. Then, the adaptors were annealed on the PCR machine with the following PCR program (95°C for 5 min first, then the temperature was slowly ramped down to 25°C with the rate of −0.1°C/s, 25°C for 5 min). The Tn5 transposase was assembled with the following components: 0.04 vol. Tn5MErev/Tn5ME-A, 0.04 vol Tn5MErev/Tn5ME-B, 0.4 vol. 100% glycerol (Sigma-Aldrich, G9012), 0.3048 vol 2× dialysis buffer (100 mM HEPES−KOH (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016) at pH 7.2, 0.2 M NaCl (Invitrogen, AM9759), 0.2 mM EDTA (Invitrogen, AM9260G), 2 mM DTT (Thermo Fisher Scientific, 20291), 0.2% Triton X-100 (Sigma-Aldrich, T8787), 20% glycerol (Sigma-Aldrich, G9012)), 0.043 vol. pure Tn5 (46.55 μM), 0.1722 vol. water (Invitrogen, AM9932). The reagents were mixed thoroughly but gently, and the mixture was left on the bench at RT for 1 h to allow annealing of oligos to Tn5.

T7−pA−Tn5 transposase assembly

Oligonucleotides (T7-Tn5ME, Tn5MErev) were resuspended in oligo annealing buffer (10 mM Tris−HCl pH 8.0 (Invitrogen, 15568-025), 50 mM NaCl (Invitrogen, AM9759), 1 mM EDTA (Invitrogen, AM9260G)) to a final concentration of 100 μM each. Equimolar amounts of Tn5MErev/T7-Tn5ME were mixed in separate 200 μl PCR tubes. Then, the adaptors were annealed on the PCR machine with the following PCR program (95°C for 5 min first, then the temperature was slowly ramped down to 25°C with the rate of −0.1°C/s, finally 25°C for 5 min). The T7−pA−Tn5 transposase was assembled with the following components: 0.08 vol. Tn5MErev/T7-Tn5ME, 0.4 vol. glycerol (Sigma-Aldrich, G9012), 0.3116 vol. 2× dialysis buffer (100 mM HEPES−KOH (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016) at pH 7.2, 0.2 M NaCl (Invitrogen, AM9759), 0.2 mM EDTA (Invitrogen, AM9260G), 2 mM DTT (Thermo Fisher scientific, 20291), 0.2% Triton X-100 (Sigma-Aldrich, T8787), 20% glycerol (Sigma-Aldrich, G9012)), 0.0362 vol. pure pA–Tn5 (55.55 μM), 0.1722 vol. water (Invitrogen, AM9932). The reagents were mixed thoroughly but gently, and the mixture was left on the bench at RT for 1 h to allow annealing of oligos to pA–Tn5.

Activity assay of Tn5, pA–Tn5 and T7−pA−Tn5 transposase

The activity of the assembled Tn5, pA–Tn5 and T7−pA−Tn5 were checked as described below. The mixture of 10 μl of 2× TD Buffer (20 mM Tris−HCl pH 7.6 (Invitrogen, 15567-027), 10 mM $MgCl_2$ (Invitrogen, AM9530G), 20% dimethyl formamide (Sigma-Aldrich, D4551)), 50 ng human genomic DNA (Promega, G3041), 1 μl of 2 μM assembled Tn5 or pA–Tn5 and 6.5 μl of water (Invitrogen, AM9932) was incubated at 55°C for 7 min. After the incubation, the mixture was purified by Qiagen MiniElute PCR Purification kit (Qiagen, 28004) and eluted in 10 μl of elution buffer. Then purified DNA was mixed with 2 μl 6× loading dye (Thermo Fisher scientific, R0611) and run on a 1% agarose gel (Lonza, 50004) to check the length distribution of DNA fragment.

CUT&Tag for GM12878 cells

CUT&Tag was performed by following the previous report (42). In brief, cells were harvested, counted, and centrifuged for 5 min at 500 × g at RT. Aliquots of cells (100 000 cells/1.5 ml Lo-Bind tube (Sarstedt, 72.706.600)) were washed twice in washing buffer (20 mM HEPES(K+) pH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016); 150 mM NaCl (Invitrogen, AM9759); 0.5 mM Spermidine (Sigma-Aldrich, S2626); 1× Protease inhibitor cocktail (Sigma-Aldrich, 11873580001)) and finally resuspended with 50 μl washing buffer and kept on ice. Concanavalin A coated magnetic beads (Bangs Laboratories, BP531) were prepared as described (20) and 10 μl of prepared beads were added per tube. 50 μl of cell suspension with 10 μl activated beads were mixed thoroughly but gently, and the mixture was left on the bench at RT for 10 min to allow binding of cells to the Concanavalin A beads. The tubes were put on a magnetic stand for 5 min to allow the mixtures to clear and the liquid was discarded. Appropriate primary antibody (H3K27ac Monoclonal Antibody (Thermo Fisher scientific, MA5-23516) and Tri-Methyl-Histone H3 (Lys27) (Cell Signaling Technology, 9733S))

was diluted 1:100 in Antibody buffer (20 mM HEPES(K+) pH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016); 150 mM NaCl (Invitrogen, AM9759); 0.5 mM Spermidine (Sigma-Aldrich, S2626); 1× Protease inhibitor cocktail (Sigma-Aldrich, 11873580001); 0.05% Digitonin (Millipore, 300410); 2 mM EDTA (Invitrogen, AM9260G), 0.1% BSA (Miltenyi Biotech MACS, 130-091-376)). Bead-bound cells were resuspended in 50 μl antibody buffer (20 mM HEPES(K+) pH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016); 150 mM NaCl (Invitrogen, AM9759); 0.5 mM spermidine (Sigma-Aldrich, S2626); 1× Protease inhibitor cocktail (Sigma-Aldrich, 11873580001); 0.05% Digitonin (Millipore, 300410); 2 mM EDTA (Invitrogen, AM9260G), 0.1% BSA (Miltenyi Biotech MACS, 130-091-376)) pre-mixed with primary antibody (1:100) with gentle pipetting. Then the tubes were placed on a nutator and incubated at 4°C overnight. The primary antibody was removed by placing the tube on the magnet stand to clear and discard all of the liquid. An appropriate secondary antibody (Guinea Pig anti-Rabbit IgG antibody (Antibodies-Online, ABIN101961) for a rabbit primary antibody or Rabbit anti-Mouse antibody (Abcam, ab46540) for a mouse primary antibody) was diluted 1:100 in 100 μl of Dig-washing buffer and bead-bound cells were resuspended in 100 μl Dig-washing buffer (20 mM HEPES(K+) pH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016); 150 mM NaCl (Invitrogen, AM9759); 0.5 mM Spermidine (Sigma-Aldrich, S2626); 1× protease inhibitor cocktail (Sigma-Aldrich, 11873580001); 0.05% digitonin (Millipore, 300410)) pre-mixed with secondary antibody (1:100) and incubated at RT on the nutator for 1 h. Bead-bound cells were washed using the magnet stand 3 times in 0.8–1 ml Dig-washing buffer to remove unbound antibodies. 1:100 dilution of pA–Tn5 adapter complex (∼20 nM) was prepared in Dig-300 buffer (20 mM HEPES(K+) pH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016); 300 mM NaCl (Invitrogen, AM9759); 0.5 mM Spermidine (Sigma-Aldrich, S2626); 1× protease inhibitor cocktail (Sigma-Aldrich, 11873580001); 0.01% digitonin (Millipore, 300410)). After removing the liquid on the magnet stand, 100 μl of diluted pA–Tn5 adapter complex was added to the bead-bound cells with gentle mixing. The tubes were incubated at RT for 1 h on the nutator. Bead-bound cells were washed 3 times with 1 mL Dig-300 buffer to remove unbound pA–Tn5 protein. After washing, the liquid was removed on the magnetic stand, bead-bound cells were resuspended in 100 μl tagmentation buffer (20 mM HEPES(K+) pH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016); 300 mM NaCl (Invitrogen, AM9759); 0.5 mM spermidine (Sigma-Aldrich, S2626); 1× protease inhibitor cocktail (Sigma-Aldrich, 11873580001); 0.01% digitonin (Millipore, 300410); 10 mM $MgCl_2$ (Invitrogen, AM9530G)). The mixture was incubated at 37°C for 1 h. To stop tagmentation, 3.3 μl 0.5M EDTA (Invitrogen, AM9260G), 1 μl 10% SDS (Invitrogen, 1553-035) and 0.84 μl 20 mg/ml Proteinase K (Thermo Fisher scientific, EO0491) were added to each tube. Then the tubes were mixed by full speed vortex ∼2 s, and incubated for 1 h at 50°C to release DNA. The incubation mixture was purified with MiniElute PCR Purification kit (Qiagen, 28004) and DNA was eluted in 21 μl of elution buffer. To amplify libraries, 21 μl DNA was mixed with 2 μl of a universal i5 and i7 primer (10 μM), using a different barcode for each sample. 25 μl of NEB Next HiFi 2× PCR master mix (New England Biolabs, M0541S) was added to the tubes and mixed well. The samples were placed in a Thermocycler with a heated lid using the following cycling conditions: 72°C for 5 min (gap filling); 98°C for 30 s; 13 cycles of 98°C for 10 s, 63°C for 10 s; final extension at 72°C for 1 min and hold at 4°C. Post-PCR clean-up was performed by adding 1.3× volume of SPRI select beads (Beckman Coulter, B23317), and libraries were incubated with beads for 15 min at RT, washed twice gently in 80% ethanol, and eluted in 25 μl of elution buffer from Qiagen MiniElute PCR Purification kit (Qiagen, 28004). The concentration of the library was measured by high sensitivity DNA Bioanalyzer (Agilent, 5067-4626). The CUT&Tag libraries were sequenced on Illumina NovaSeq 6000 sequencer or MiniSeq sequencer with paired end sequencing.

FACT-seq for GM12878 cells

Cells were harvested, counted, and centrifuged for 5 min at 500 × g at RT. Aliquots of cells (100 000cells/1.5 ml Lo-Bind tube (Sarstedt, 72.706.600)) were washed twice in washing buffer (20 mM HEPES(K+) PH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016) pH 7.5; 150 mM NaCl (Invitrogen, AM9759); 0.5 mM spermidine (Sigma-Aldrich, S2626); 1× protease inhibitor cocktail (Sigma-Aldrich, 11873580001)) and finally resuspended with 50 μl washing buffer and kept on ice. Concanavalin A coated magnetic beads (Bangs Laboratories, BP531) were prepared as described (20) and 10 μl of prepared beads were added per tube. 50 μl of cell suspension with 10 μl activated beads were mixed thoroughly but gently, and the mixture was left on the bench at RT for 10 min to allow binding of cells to the Concanavalin A beads. The tubes were put on a magnetic stand for 5 min to allow the mixtures to clear and the liquid was discarded. Appropriate primary antibody (H3K27ac Monoclonal Antibody (Thermo Fisher Scientific, MA5-23516) and tri-methylhistone H3 (Lys27) (Cell Signaling Technology, 9733S)) was diluted 1:100 in antibody buffer (20 mM HEPES(K+) pH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016); 150 mM NaCl (Invitrogen, AM9759); 0.5 mM spermidine (Sigma-Aldrich, S2626); 1× protease inhibitor cocktail (Sigma-Aldrich, 11873580001); 0.05% digitonin (Millipore, 300410); 2 mM EDTA (Invitrogen, AM9260G), 0.1% BSA (Miltenyi Biotech MACS, 130-091-376)). Bead-bound cells were resuspended in 50 μl Antibody buffer pre-mixed with primary antibody (1:100) with gentle pipetting. Then the tubes were placed on a nutator and incubated at 4°C overnight. The primary antibody was removed by placing the tube on the magnet stand to clear and discard all the liquid. An appropriate secondary antibody (Guinea Pig anti-Rabbit IgG antibody (Antibodies-online, ABIN101961) for a rabbit primary antibody or Rabbit anti-Mouse antibody (Abcam, ab46540) for a mouse primary antibody) was diluted 1:100 in 100 μl of Dig-washing buffer and bead-bound cells were resuspended in 100 μl Dig-washing buffer (20 mM HEPES(K+) pH

7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016); 150 mM NaCl (Invitrogen, AM9759); 0.5 mM Spermidine (Sigma-Aldrich, S2626); 1× protease inhibitor cocktail (Sigma-Aldrich, 11873580001); 0.05% Digitonin (Millipore, 300410)) pre-mixed with secondary antibody (1:100) and incubated at RT on the nutator for 1 h. Bead-bound cells were washed using the magnet stand 3 times in 0.8–1 ml Dig-washing buffer to remove unbound antibodies. 1:100 dilution of T7−pA−Tn5 adapter complex (∼20 nM) was prepared in Dig-300 buffer (20 mM HEPES(K+) pH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016); 300 mM NaCl (Invitrogen, AM9759); 0.5 mM spermidine (Sigma-Aldrich, S2626); 1× protease inhibitor cocktail (Sigma-Aldrich, 11873580001); 0.01% digitonin (Millipore, 300410)). After removing the liquid on the magnet stand, 100 μl of diluted T7−pA−Tn5 adapter complex was added to the bead-bound cells with gentle mixing. The tubes were incubated at RT for 1 h on the nutator. Bead-bound cells were washed 3 times in 0.8–1 mL Dig-300 buffer to remove unbound T7−pA−Tn5 protein. After washing, the liquid was removed on the magnetic stand, bead-bound cells were resuspended in 100 μl tagmentation buffer (20 mM HEPES(K+) pH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016); 300 mM NaCl (Invitrogen, AM9759); 0.5 mM spermidine (Sigma-Aldrich, S2626); 1× protease inhibitor cocktail (Sigma-Aldrich, 11873580001); 0.01% digitonin (Millipore, 300410); 10 mM $MgCl_2$ (Invitrogen, AM9530G)). The mixture was incubated at 37°C for 1 h. To stop tagmentation, 3.3 μl 0.5M EDTA (Invitrogen, AM9260G), 1 μl 10% SDS (Invitrogen, 1553-035) and 0.84 μl 20 mg/ml Proteinase K (Thermo Fisher scientific, EO0491) were added to each tube. Then the tubes were mixed by full speed vortexing ∼2 s, and incubated for 1 h at 50°C to release DNA. The incubation mixture was purified with MiniElute PCR Purification kit (Qiagen, 28004) and DNA was eluted in 20 μl of elution buffer. Then 20 μl of 2× PCR master mix (New England Biolabs, M0541S) was added to the samples. The mixture was in a thermo cycler at 72°C for 8 mins. The samples were purified with Qiagen MiniElute PCR Purification kit (Qiagen, 28004) first and eluted in 50 μl of water. Then SPRI select beads (Beckman Coulter, B23317) with 1:1 ratio was added into each tube. Libraries were incubated with beads for 15 min at RT, washed twice gently in 80% ethanol, and eluted in 26 μl of water.

Next, the *in vitro* transcription (IVT) was performed with T7 high yield RNA synthesis kit (New England Biolabs, E2040S) overnight. After the *in vitro* transcription, 5 μl DNase 1 (from Zymo, R1013) and 5 μl of DNA digestion buffer (from Zymo, R1013) were added to the tubes and the tubes were incubated at RT for 15 min to digest the remaining DNA. Then the reactions were purified using ZYMO RNA Clean & Concentration kit (Zymo, R1013) and eluted in 15 μl of nuclease-free water (Invitrogen, AM9932). 100 ng IVT RNA was used for each library preparation. The IVT RNA was reverse transcribed into cDNA with random primers (Thermo Fisher scientific, SO142) and SMART MMLV kit by following the manufactory protocol (TaKaRa, 639524). In brief, the mixture was incubated at 42°C for 60 min and 70°C for 15 min, then 2 μl of RNase H buffer and 0.2 μl RNase H enzyme (Thermo Fisher scientific, EN0201) were added and incubated at 37°C for 20 min. The cDNA was purified using RNA XP beads purification with 1:1.8 ratio of sample to beads (Beckman Coulter, A63987) and eluted in 24.2 μl water. Next, the single-stranded cDNA was converted to double-stranded cDNA through pre-PCR (98°C for 10 s, 63°C for 30 s, 72°C for 1 min, 10°C hold in one cycle) which was performed by adding 25 μl of 2× PCR master mix (New England Biolabs, M0541S) and 0.8 μl of reverse primer. Then samples were purified using MiniElute PCR Purification kit (Qiagen, 28004) and eluted in 20 μl water. The sequencing library was prepared with Tn5 tagmentation. In short, the double strand DNA samples were subjected to the tagmentation by adding 25 μl of 2× TD-buffer (20 mM Tris–HCl pH 7.6 (Invitrogen, 15567-027), 10 mM $MgCl_2$ (Invitrogen, AM9530G), 20% dimethyl formamide (Sigma-Aldrich, D4551)), 0.5 μl 2 μM normal Tn5, 4.5 μl nuclease-free water (Invitrogen, AM9932) and incubated at 55°C for 7 min, then samples were purified using Qiagen MiniElute PCR Purification kit (Qiagen, 28004) and eluted in 20 μl of elution buffer. The library amplification PCR was performed by adding 25 μl of 2× PCR master mix (New England Biolabs, M0541S), 0.4 ul of barcodes forward primer i5 (25 μM), 0.4 μl of barcodes i7 reverse primer (25 μM), 4.2 μl of nuclease-free water (Invitrogen, AM9932) to the sample, with the following PCR protocol (72°C 5 min first, 20 cycles of 98°C for 10 s, 63°C for 30 s, 72°C for 1 min), then samples were purified using Qiagen MiniElute kit (Qiagen, 28004) and eluted in 20 μl of water. The DNA library was purified with 8% PAGE gel (40% acrylamide:bis-acrylamide (Invitrogen, HC2040), 10× TBE buffer, 10% ammonium persulfate (Invitrogen, HC2005), TEMED (Invitrogen, HC2006)) purification and the DNA from 220–1000 bp was selected by gel cutting. The concentration of the library was measured by high sensitivity DNA Bioanalyzer (Agilent, 5067-4626). The FACT-seq libraries were sequenced on Illumina NovaSeq 6000 sequencer with paired end sequencing.

CUT&Tag for frozen tissue

The single nuclei were isolated from the frozen 8 weeks' mouse kidney with dounce homogenization by following the standard isolation protocol from Omni-ATAC (48). Frozen tissue was taken in the ice-cold 1 ml of 1× unstable Homogenization Buffer (5 mM $CaCl_2$ (Alfa Aesar, J63122), 3 mM $Mg(Ac)_2$ (Sigma-Aldrich, M5661), 10 mM Tris pH 7.8 (Invitrogen, 15568-025), 0.01667 mM PMSF (Sigma-Aldrich, P7626), 0.1667 mM β-mercaptoethanol (Sigma-Aldrich, M-6250), 320 mM Sucrose (Sigma-Aldrich, 84097-250), 0.1 mM EDTA (Invitrogen, AM9290G), 0.1% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50)) and incubated for 5 min on ice. The tissue was homogenized 10 times with loose pestle and 20 times with tight pestle. Then, the homogenized tissue was split and transferred to two new 1.5 ml Lo-Bind tube (Sarstedt, 72.706.600). The homogenized tissue was centrifuged for 1 min at 100 × g at 4°C. Then 400 μl of the homogenized sample was transferred and mixed with 400 μl of 50% OptiPrep Density Gradient Solution (5 mM $CaCl_2$ (Alfa Aesar, J63122), 3 mM $Mg(Ac)_2$ (Sigma-Aldrich, M5661), 10 mM Tris pH 7.8 (Invitrogen, 15568-025), 0.01667 mM PMSF (Sigma-

Aldrich, P7626), 0.1667 mM β-mercaptoethanol (Sigma-Aldrich, M-6250), 50% OptiPrep Density Gradient Medium (Sigma-Aldrich, D1556-250)), to make a final concentration of 25% of OptiPrep Density Gradient Solution with homogenized tissue. After preparation of tissue mixture, a fresh 2 ml Lo-Bind vial (Sarstedt, 72.695.400) was taken and 25% of OptiPrep Density Gradient Solution mixed with the tissue, 29% of OptiPrep Density Gradient Solution (5 mM $CaCl_2$ (Alfa Aesar, J63122), 3 mM $Mg(Ac)_2$ (Sigma-Aldrich, M5661), 10 mM Tris pH 7.8 (Invitrogen, 15568-025), 0.01667 mM PMSF (SIGMA-ALDRICH, P7626), 0.1667 mM β-mercaptoethanol (Sigma-Aldrich, M-6250), 160 mM sucrose (Sigma-Aldrich, 84097-250), 29% OptiPrep Density Gradient Medium (Sigma-Aldrich, D1556-250)) and 35% of OptiPrep Density Gradient Solution (5 mM $CaCl_2$ (Alfa Aesar, J63122), 3 mM $Mg(Ac)_2$ (Sigma-Aldrich, M5661), 10 mM Tris pH 7.8 (Invitrogen, 15568-025), 0.01667 mM PMSF (Sigma-Aldrich, P7626), 0.1667 mM β-mercaptoethanol (Sigma-Aldrich, M-6250), 160mM sucrose (Sigma-Aldrich, 84097-250), 35% OptiPrep Density Gradient Medium (Sigma-Aldrich, D1556-250)) were layered on the top of each other. The layered vial was centrifuged at 3000 × g for 20 min at 4°C. After gradient centrifugation, the top 1000 μl was discarded and the 200 μl of the nuclei region was carefully collected to a fresh 1.5 ml Lo-Bind tube (Sarstedt, 72.706.600). Then 1 ml of CUT&Tag washing buffer (20 mM HEPES(K+) PH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016) pH 7.5; 150 mM NaCl (Invitrogen, AM9759); 0.5 mM Spermidine (Sigma-Aldrich, S2626); 1× protease inhibitor cocktail (Sigma-Aldrich, 11873580001)) was added to each tube, mixed well and centrifuged at 4°C 600 × g for 10 min. After discarding the supernatant, the nuclei pellets were resuspended with 500 μl of washing buffer and counted by the cell counter. Aliquots of nuclei (100 000 nuclei/1.5 ml Lo-Bind tube (Sarstedt, 72.706.600)) were washed once in washing buffer and finally resuspended with 50 μl of washing buffer and kept on ice. The CUT&Tag libraries were prepared with same protocol of CUT&Tag in GM12878 cells. Antibodies were used for CUT&Tag libraries preparation are listed as follows: anti-H3K27ac antibody (abcam, ab4729), anti-H3K27me3 antibody (Cell Signaling Technology, 9733S), anti-H3K36me3 antibody (abcam, ab9050) and anti-H3K4me1 antibody (abcam, ab176877)). The CUT&Tag libraries were sequenced on Illumina NovaSeq 6000 sequencer with paired end sequencing.

Human CRC, GBM sample collections and FFPE block preparation

The regional ethical research committee at the Uppsala University approved the study of human CRC (Dnr 2015/419 and 2018/490) and human GBM (2007/353). The FFPE tissue blocks of human colorectal cancer (CRC) and glioblastoma multiforme (GBM) were prepared at the Department of Clinical Pathology, Uppsala University Hospital, Uppsala, Sweden, according to standard procedures. Briefly, tissue from surgical specimens of samples were fixed in buffered formalin for 24–72 hs. The pieces were then examined by a pathologist, excised and placed in plastic cassettes. The fixed tissue was then dehydrated in an automated system (Tissue-Tek® VIP®) where the tissue was immersed in ethanol of varying concentrations (70%, 95%, 99.5%) followed by xylene and finally paraffin (Histowax®, Histolab) over a period of ∼12 h. Finally, the paraffin embedded tissue piece was oriented in a cassette, liquid paraffin was poured over it and allowed to set, forming the FFPE block. The FFPE block was then sectioned on a microtome at a thickness of 10 μm for CRC and 20 μm for GBM.

Nuclei isolation from FFPE tissue sections

One curved tissue section (mouse FFPE kidney and liver tissue sections (20-μm thick sections), human GBM tissue section (20-μm thick section), human CRC tissue section (10-μm thick section)) was deparaffined with 1 ml of xylene (HistoLab, 02070) 5 min, thrice. Rehydration was done by sequential ethanol washing, started with 100% ethanol 5 min twice, 95%, 70%, 50%, 30% ethanol, 5 min each. After deparaffinization and rehydration, tissue was washed with 1 ml water, then 1 ml PBS containing 0.5 mM $CaCl_2$ (Alfa Aesar, J63122). Then the tissue was subjected to microdissection under a stereo microscope. 1 ml enzymatic cocktail (3 mg/ml of Collagenase (Sigma-Aldrich, C9263) and 300 U/ml of hyaluronidase (Merk Millipore, HX0154-1)) in PBS contains 0.5 mM $CaCl_2$ (Alfa Aesar, J63122)) was added to the dissected tissue. Then the mixture was incubated at 37°C for 16 h by adding 100 μg of Ampicillin (Serva, 69-52-3) and 50 μg of sodium azide (Merck Millipore, 26628-22-8). After the enzyme digestion, 400 μl NST buffer (146 mM NaCl (Invitrogen, AM9759) (Invitrogen, 00648496), 10 mM Tris pH 7.8 (Invitrogen, 15568-025), 1 mM $CaCl_2$ (Alfa Aesar, J63122), 21 mM of $MgCl_2$ (Invitrogen, AM9530G), 0.05% BSA (Miltenyi Biotech MACS, 130-091-376), 0.2% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50)) was added to the mixture, and the tube was centrifuged at 3000 × g for 10 min. After the centrifugation, the supernatant was aspirated and discarded, then the pellet was resuspended in NST buffer containing 0.1% DNase free RNase A (Thermo Fisher scientific, EN0531), and 10% fetal bovine serum (Life Technologies, 10108-105). The mixture was passed through the 27 G needle syringe 30 times and filtered with a 30 μM filter (Miltenyi Biotech MACS,130-098-458). Then the passthrough nuclei suspension were centrifuged at 3000 × g for 10 min, and the nuclei were resuspended in 1× PBS, checked and counted.

Nuclei microdissection from small pieces of FFPE tissue on glass slide

A mouse kidney FFPE tissue block was sectioned into 20-μm slides using microtome. Then two adjacent sections were taken and made identical cuts using a razor blade under stereo microscope. Both the sections were deparaffinized with xylene (HistoLab, 02070) for 5 min three times, followed by sequential rehydration. Rehydration was done by sequential ethanol washing, started with 100% ethanol 5 mins twice, 95%, 70%, 50%, 30% ethanol, 5 min each. After deparaffinization and rehydration, tissue was washed with water, then PBS containing 0.5 mM $CaCl_2$ (Alfa Aesar,

J63122). Then one of the sections was treated with 200 μl of Hoechst 33342 solution (with 1:100, H3570, Thermo Fisher Scientific) for 5 min and washed twice using PBS for 5 min each. After washing the slide, 10 μl mounting media was added and coved with coverslip. Then the slide's top tissue area was visualized and counted under microscope. Another tissue slide´s top area was carefully separated, scraped from the slide in small pieces using 19 G needle and transferred to a vial containing PBS for nuclei isolation by following nuclei isolation steps stated as above. After collecting the nuclei, H3K27ac FACT-seq was performed as described below.

Nuclei microdissection from 5-, 7- and 10-μm-thick mouse FFPE kidney tissue section

A small area of mouse kidney (~3 × 5 mm^2) FFPE tissue block was sectioned into 5-, 7- and 10-μm-thick curls using microtome. Nuclei isolation and FACTs-seq sequencing libraries preparation were performed by following same steps stated for 20-μm-thick mouse FFPE kidney tissue sections.

Epitope retrieval

After nuclei isolation, a proper amount of FFPE nuclei (100 000–1 500 000) were transferred to a new 1.5 ml Lo-Bind tube (Sarstedt, 72.706.600). The nuclei were centrifuged at 3000 × g for 5 min at RT and the supernatant was discarded. To profile histone modification H3K27ac, the FFPE nuclei (either from mouse FFPE tissue blocks or from human clinical samples) were resuspended with 50 μl of Epitope Retrieval Buffer-1 (10 mM EDTA (Invitrogen, AM9260G) pH 8, 50 mM Tris–HCl pH 8, 0.1% SDS (Invitrogen, 1553-035)) and transferred to a PCR tube. The nuclei suspension was incubated on a PCR machine at 50°C for 1 h. After incubation, 10 μl 10% Triton X-100 (or 30 μl 10% Triton X-100 for higher concentration (0.3%) of SDS) was added to the tube and the mixture was transferred to a 1.5 ml Lo-Bind tube. And the tube was placed on a shaker and incubated at 37°C for 30 min with 500 rpm to quench SDS. After quenching, the mixture was centrifuged at 3000 × g for 5 min under RT and the pellet was washed once by the FACT-seq Antibody buffer (20 mM HEPES(K+) pH 7.6 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016), 150 mM NaCl (Invitrogen, AM9759), 2 mM EDTA (Invitrogen, AM9260G), 0.5 mM Spermidine (Sigma-Aldrich, S2626), 0.05% digitonin (Millipore, 300410), 0.01% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50), 1× protease inhibitors (Sigma-Aldrich, 11873580001), 1% BSA (Miltenyi Biotech MACS, 130-091-376)). After washing, the nuclei pellet was resuspended with 200 μl of FACT-seq Antibody buffer (20 mM HEPES(K+) pH 7.6 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016), 150 mM NaCl (Invitrogen, AM9759), 2mM EDTA (Invitrogen, AM9260G), 0.5 mM spermidine (Sigma-Aldrich, S2626), 0.05% Digitonin (Millipore, 300410), 0.01% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50), 1× protease inhibitors (Sigma-Aldrich, 11873580001), 1% BSA (Miltenyi Biotech MACS, 130-091-376)) and the nuclei number was counted by the cell counter.

To profile histone modification of H3K27me3, different epitope retrieval buffers and temperatures were tested. Following conditions were used: 65°C for 1 h incubation with 10 mM EDTA, pH 8, 50 mM Tris–HCl pH 8, 0.1% SDS, 0.1% sodium deoxycholate; 65°C for 1 h incubation with 10 mM EDTA, pH 8, 50 mM Tris–HCl pH 8, 0.1% SDS; 50°C for 1 h incubation with 10 mM EDTA, pH 8, 50 mM Tris–HCl pH 8, 0.1% SDS; 50°C for 1 h incubation with 10 mM EDTA, pH 8, 50 mM Tris–HCl pH 8, 0.3% SDS. In brief, the FFPE nuclei (from mouse FFPE tissue blocks) were first resuspended with 50 μl Epitope Retrieval Buffer stated as above, then 10 μl 10% Triton X-100 was added to the tube and the mixture was transferred to a 1.5 ml Lo-Bind tube. And the tube was placed on a shaker and incubated at 37°C for 30 min with 500 rpm shaking to quench SDS. After quenching, the mixture was centrifuged at 3000 × g for 5 min under RT and the pellet was washed once with the FACT-seq Antibody buffer (20 mM HEPES(K+) pH 7.6 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016), 150 mM NaCl (Invitrogen, AM9759), 2 mM EDTA (Invitrogen, AM9260G), 0.5 mM spermidine (Sigma-Aldrich, S2626), 0.05% digitonin (Millipore, 300410), 0.01% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50), 1× protease inhibitors (Sigma-Aldrich, 11873580001), 1% BSA (Miltenyi Biotech MACS, 130-091-376)). After washing, the nuclei pellet was resuspended with 200 μl of FACT-seq Antibody buffer (20 mM HEPES(K+) pH 7.6 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016), 150 mM NaCl (Invitrogen, AM9759), 2 mM EDTA (Invitrogen, AM9260G), 0.5 mM spermidine (Sigma-Aldrich, S2626), 0.05% digitonin (Millipore, 300410), 0.01% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50), 1× protease inhibitors (Sigma-Aldrich, 11873580001), 1% BSA (Miltenyi Biotech MACS, 130-091-376)) and the nuclei number was counted by the cell counter. To check the nuclei integrity, 30 000–50 000 nuclei after epitope retrieval were stained with 1:2000 DAPI (Invitrogen, H3570) in 1× PBS for 15 min at RT. Then nuclei were transferred to the glass slides using Cytospin (Double Cellfunnel (THARMAC, 306-12), Filter cards (THARMAC, 307–500)) with 600 × g for 6 min. Finally, the slides were mounted using SlowFade Gold antifade reagent (Invitrogen, S36940) and imaged with Zeiss Axio Imager Z2.

For histone modifications of H3K27ac, H3K36me3 and H3K4me1, both mild epitope retrieval (50°C for 1 h incubation with 10 mM EDTA, pH 8, 50 mM Tris–HCl pH 8, 0.1% SDS) and harsh epitope retrieval conditions (65°C for 1 h incubation with 10 mM EDTA, pH 8, 50 mM Tris–HCl pH 8, 0.1% SDS, 0.1% sodium deoxycholate) were applied to perform FACT-seq.

Comparison of antibodies used for CUT&Tag, FACT-seq and ENCODE ChIP-seq in mouse tissue

Histone marker	Antibodies used in FACT-seq and CUT&Tag	Antibodies used in ENCODE ChIP-seq
H3K27ac	ab4729, abcam	ab4729, abcam
H3K27me3	9733S, Cell Signaling technology	07-449, abcam
H3K4me1	ab176877, abcam	ab8895, abcam
H3K36me3	ab9050, abcam	ab9050, abcam

FACT-seq for mouse FFPE nuclei

Aliquots of nuclei either with or without epitope retrieval (100,000 cells/0.5 mL Qubit tube (Invitrogen, Q32856)) were washed once with FACT-seq Antibody buffer (20 mM HEPES(K+) pH 7.6 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016), 150 mM NaCl (Invitrogen, AM9759), 2 mM EDTA (Invitrogen, AM9260G), 0.5 mM spermidine (Sigma-Aldrich, S2626), 0.05% digitonin (Millipore, 300410), 0.01% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50), 1× protease inhibitors (Sigma-Aldrich, 11873580001), 1% BSA (Miltenyi Biotech MACS, 130-091-376)). After washing, the nuclei were resuspended in 200 μl antibody buffer with 1:100 diluted primary antibodies (anti-H3K27ac antibody (abcam, ab4729), anti-H3K27me3 antibody (Cell Signaling Technology, 9733S), anti-H3K36me3 antibody (abcam, ab9050) and anti-H3K4me1 antibody (abcam, ab176877)) and incubated overnight at 4°C with slow rotation. Next day the nuclei were centrifuged 5 min at 2000 × g, washed once with 200 μl of FACT-seq Dig-washing buffer (20 mM HEPES(K+) pH 7.6 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016), 150 mM NaCl (Invitrogen, AM9759), 0.5 mM Spermidine (Sigma-Aldrich, S2626), 0.05% digitonin (Millipore, 300410), 0.01% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50), 1× protease inhibitors (Sigma-Aldrich, 11873580001), 1% BSA (Miltenyi Biotech MACS, 130-091-376)), and finally resuspended in 200 μl of Dig-washing buffer with 1:100 diluted secondary antibody (Guinea Pig anti-Rabbit IgG antibody (Antibodies-Online, ABIN101961) for a rabbit primary antibody or Rabbit anti-Mouse antibody (Abcam, ab46540) for a mouse primary antibody) and incubated for 1h at RT with slow rotation. Then, nuclei were centrifuged 5 mins at 2000 × g, washed three times with 200 μl of Dig-washing buffer and resuspended in 200 μl of Dig-300 buffer (20 mM HEPES(K+) pH 7.6 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016), 300 mM NaCl (Invitrogen, AM9759), 0.5 mM spermidine (Sigma-Aldrich, S2626), 0.05% Digitonin (Millipore, 300410), 0.01% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50), 1× protease inhibitors (Sigma-Aldrich, 11873580001), 1% BSA (Miltenyi Biotech MACS, 130-091-376)) with 1:100 diluted T7−pA−Tn5 and incubated for 1 h with slow rotation at RT. After T7−pA−Tn5 binding, the nuclei were centrifuged 5 mins at 600 × g, washed three times with 200 μl of Dig-300 buffer and resuspended in 200 μl of FACT-seq tagmentation buffer (20 mM HEPES(K+) pH 7.6 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016), 300 mM NaCl (Invitrogen, AM9759), 0.5 mM spermidine (Sigma-Aldrich, S2626), 0.05% digitonin (Millipore, 300410), 0.01% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50), 1× protease inhibitors (Sigma-Aldrich, 11873580001), 10 mM $MgCl_2$ (Invitrogen, AM9530G)), and incubated for 1 h at 37°C. After that, tagmentation was stopped by addition of 6.7 μl 0.5M EDTA (Invitrogen, AM9260G), 22 μl 10% SDS (Invitrogen, 1553-035) and 2.2 μl 20 mg/ml Proteinase K (Thermo Fisher scientific, EO0491) to each tube, and the mixture was mixed by full speed vortexing for ∼2 s. Then the first Proteinase K (Thermo Fisher scientific, EO0491) digestion was performed by incubating the mixture on a shaker at 65°C with 1200 rpm shaking for 2 hs. After digestion, the reverse-crosslinking was performed overnight under 72°C (for H3K27ac) or 80°C (H3K27me3) with 1200 rpm shaking. Next day, 2.2 μl 20 mg/ml Proteinase K (Thermo Fisher scientific, EO0491) was added to each tube after the mixture was cooled down to RT and the secondary Proteinase K (Thermo Fisher scientific, EO0491) digestion was performed by incubating the mixture on a shaker at 65°C with 1200 rpm shaking for 1 h. After digestion, 67 μl of water and 300 μl of phenol were added to each tube and the solution was mixed by full speed vortexing for ∼2 s. Then the mixture was centrifuged for 15 min at 4°C with 16 000 × g. After centrifugation, the liquid from aqueous layer was transferred to new 1.5 ml Lo-Bind tube (Sarstedt, 72.706.600) and the equal amount of chloroform was added to each tube and mixed by gentle pipetting. The mixture was centrifuged for 15 mins at 4°C with 16 000 × g. After centrifugation, the liquid from aqueous layer was transferred to a new 1.5 ml Lo-Bind tube (Sarstedt, 72.706.600) and 2.5×–3× volume of absolute ethanol and proper amount of 5 M NaCl (Invitrogen, AM9759) were added to each tube to make the final NaCl concentration 200 mM. After mixing, the DNA was precipitated at –80°C for 2–3 h. After precipitation, the DNA was centrifuged for 20 min at 4°C with 16 000 × g. Then, the supernatant was discarded, and the DNA was washed once with 700 μl 70% ethanol. After washing and drying up, the DNA was eluted in 20 μl of water. Then 20 μl of 2× PCR master mix (New England Biolabs, New England Biolabs, M0541S) was added to the samples. The mixture was in a thermo cycler at 72°C for 8 min. The samples were purified with Qiagen MiniElute PCR Purification kit (Qiagen, 28004) first and eluted in 50 μl of water. Then SPRI select beads (Beckman Coulter, B23317) with 1:1 ratio was added into each tube. Libraries were incubated with beads for 15 min at RT, washed twice gently in 80% ethanol, and eluted in 26 μl of water.

Next, the IVT was performed with T7 high yield RNA synthesis kit (New England Biolabs, E2040S) overnight. After the IVT, 5 μl DNase 1 (from Zymo, R1013) and 5 μl of DNA digestion buffer (from Zymo, R1013) were added to the tubes and the tubes were incubated at RT for 15 min to digest the remaining DNA. Then the reactions were purified using ZYMO RNA Clean & Concentration kit (Zymo, R1013) and eluted in 15 μl of nuclease-free water (Invitrogen, AM9932). 100 ng IVT RNA was used for each library preparation. The IVT RNA was reverse transcribed into cDNA with random primers (Thermo Fisher scientific, SO142) and SMART MMLV kit by following the manufactory protocol (TaKaRa, 639524). In brief, the mixture was incubated at 42°C for 60 min and 70°C for 15 min, then 2 μl of RNase H buffer and 0.2 μl RNase H enzyme (Thermo Fisher scientific, EN0201) were added and incubated at 37°C for 20 min. The cDNA was purified using RNA XP beads purification with 1:1.8 ratio of sample to beads (Beckman Coulter, A63987) and eluted in 24.2 μl water. Next, the single-stranded cDNA was converted to double-stranded cDNA through pre-PCR (98°C for 10 s, 63°C for 30 s, 72°C for 1 min, 10°C hold in one cycle) which was performed by adding 25 μl of 2× PCR master mix (New England Biolabs, M0541S) and 0.8 μl of reverse primer.

Then samples were purified using MiniElute PCR Purification kit (Qiagen, 28004) and eluted in 20 μl water.

The sequencing library was prepared with Tn5 tagmentation. In short, the double strand DNA samples were subjected to the tagmentation by adding 25 μl of 2× TD-Buffer (20 mM Tris–HCl pH 7.6 (Invitrogen, 15567-027), 10mM MgCl2 (Invitrogen, AM9530G), 20% Dimethyl Formamide (Sigma-Aldrich, D4551)), 0.5 μl 2 μM normal Tn5, 4.5 μl nuclease-free water (Invitrogen, AM9932) and incubated at 55°C for 7 min, then samples were purified using Qiagen MiniElute PCR Purification kit (Qiagen, 28004) and eluted in 20 μl of elution buffer. The library amplification PCR was performed by adding 25 μl of 2× PCR master mix (New England Biolabs, M0541S), 0.4 ul of barcodes forward primer i5 (25 μM), 0.4 μl of barcodes i7 reverse primer (25 μM), 4.2 μl of nuclease-free water (Invitrogen, AM9932) to the samples, with the following PCR protocol (72°C 5 min first, 20 cycles of 98°C for 10 s, 63°C for 30 s, 72°C for 1 min), then samples were purified using Qiagen MiniElute kit (Qiagen, 28004) and eluted in 20 μl of water. The DNA library was purified with 8% PAGE gel (40% Acrylamide:bis-acrylamide (Invitrogen, HC2040), 10× TBE buffer, 10% Ammonium persulfate (Invitrogen, HC2005), TEMED (Invitrogen, HC2006)) purification and the DNA from 220–1000 bp was selected by gel cutting. The concentration of the library was measured by high sensitivity DNA Bioanalyzer (Agilent, 5067-4626). At last, the FACT-seq libraries were sequenced on the Illumina NovaSeq 6000 sequencer.

FACT-seq for FFPE nuclei from clinical samples

Aliquots of nuclei with epitope retrieval from human GBM or human CRC samples (100 000 cells/0.5 ml Qubit tube (Invitrogen, Q32856)) were washed once with FACT-seq Antibody buffer (20 mM HEPES(K+) pH 7.6 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016), 150 mM NaCl (Invitrogen, AM9759), 2 mM EDTA (Invitrogen, AM9260G), 0.5 mM spermidine (Sigma-Aldrich, S2626), 0.05% digitonin (Millipore, 300410), 0.01% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50), 1× protease inhibitors (Sigma-Aldrich, 11873580001), 1% BSA (Miltenyi Biotech MACS, 130-091-376)). H3K27ac FACT-seq sequencing libraries were prepared with same methods as in the mouse samples.

FACT-seq sensitivity assay

1000, 10k and 100k purified mouse kidney FFPE mouse nuclei were split into different 0.5 ml Qubit tube (Invitrogen, Q32856)s with duplicate after epitope retrieval and all of them were washed once with FACT-seq Antibody buffer (20 mM HEPES(K+) pH 7.6 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016), 150 mM NaCl (Invitrogen, AM9759), 2 mM EDTA (INVITROGEN, AM9260G), 0.5 mM Spermidine (Sigma-Aldrich, S2626), 0.05% Digitonin (Millipore, 300410), 0.01% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50), 1× protease inhibitors (Sigma-Aldrich, 11873580001), 1% BSA (Miltenyi Biotech MACS, 130-091-376)). The H3K27ac FACT-seq sequencing libraries were prepared with same methods in the section of FACT-seq for mouse FFPE nuclei.

Recovery rate of FFPE nuclei isolation and epitope retrieval

The mouse kidney sections was stained with DAPI (Invitrogen, H3570) and scanned by the Zeiss Axio Imager Z2. The nuclei number was counted by the Fiji plugin (49): the Trainable Weka Segmentation (50). After scanning, the tissue section was proceeded with FFPE nuclei isolation. After the nuclei isolation, the total nuclei number was calculated accordingly. If it's the case that single nuclei suspension was originated from a small area of the tissue section, then the nuclei were transferred to the glass slides using Cytospin (Double Cellfunnel (THARMAC, 306-12), stained with DAPI (Invitrogen, H3570), imaged by the Zeiss Axio Imager Z2 and finally counted manually by the Fiji software. In summary, the recovery rates of FFPE nuclei isolation were 42.04% (from small-cropped mouse kidney region) and 53.49% (from a whole mouse kidney section). During epitope retrieval, a proper number of nuclei (100 000–1 500 000) were transferred out and resuspended with the epitope retrieval buffer. After the epitope retrieval, quenching and washing, the nuclei number was counted again by the cell counter. The recovery rates of epitope retrieval were calculated.

Count nuclei in tissue sections

The scanned FFPE whole kidney section was imported into ImageJ (Fiji) (49). One hundred twenty-six grids were added in the picture to separate the entire image into 126 small regions. Image thresholding was performed by using the Trainable Weka Segmentation (50) plugin. Fifty nuclei and 75 parts of the background were picked up randomly to train the classifier. After training, the classifier model was used to identify the nuclei in 27 small grids. Then, the identified nuclei regions were proceeded with the default 'Watershed' and 'Analyze Particles' function in Fiji to count the nuclei number in individual pictures. Eventually, the total number of nuclei was estimated based on the nuclei counted from the 27 grids.

CUT&Tag in FFPE samples with concanavalin A beads

Aliquots of purified mouse kidney FFPE mouse nuclei (100 000 nuclei/1.5 ml Lo-Bind tube (Sarstedt, 72.706.600)) were centrifuged for 5 min at 2500 × g at RT and washed once in CUT&Tag washing buffer (20 mM HEPES(K+) pH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016); 150 mM NaCl (Invitrogen, AM9759); 0.5 mM Spermidine (Sigma-Aldrich, S2626); 1× protease inhibitor cocktail (Sigma-Aldrich, 11873580001)) and finally resuspended with 50 μl of washing buffer and kept on ice. Concanavalin A coated magnetic beads (Bangs Laboratories, BP531) were prepared as described (20) and 10 μl of prepared beads were added per tube. 50 μl of nuclei suspension with 10 μl activated beads were mixed thoroughly but gently, and the mixture was left on the bench at RT for 10 min to allow binding of nuclei to the Concanavalin A beads. The tubes were put on a magnetic stand for 5 min to allow the mixtures to clear and the liquid was discarded. Appropriate primary antibody (anti-H3K27me3 antibody (Cell Signaling Technology, 9733S)) or anti-H3K27ac antibody (abcam, ab4729)) was diluted

1:100 in Antibody buffer (20 mM HEPES(K+) pH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016); 150 mM NaCl (Invitrogen, AM9759); 0.5 mM Spermidine (Sigma-Aldrich, S2626); 1 × Protease inhibitor cocktail (Sigma-Aldrich, 11873580001); 0.05% Digitonin (Millipore, 300410); 2 mM EDTA (Invitrogen, AM9260G), 0.1% BSA (Miltenyi Biotech MACS, 130-091-376)). Bead-bound nuclei were resuspended in 50 μl Antibody buffer pre-mixed with primary antibody (1:100) with gentle pipetting. Then the tubes were placed on a nutator and incubated at 4°C overnight. The primary antibody was removed by placing the tube on the magnet stand to clear and discard all the liquid. An appropriate secondary antibody (Guinea Pig anti-Rabbit IgG antibody (Antibodies-online, ABIN101961)) was diluted 1:100 in 100 μlof CUT&Tag Dig-washing buffer (20 mM HEPES(K+) pH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016); 150 mM NaCl (Invitrogen, AM9759); 0.5 mM spermidine (Sigma-Aldrich, S2626); 1× protease inhibitor cocktail (Sigma-Aldrich, 11873580001); 0.05% digitonin (Millipore, 300410)) and bead-bound nuclei were resuspended in 100 μl Dig-washing buffer pre-mixed with secondary antibody (1:100) and incubated at RT on the nutator for 1 h. Bead-bound nuclei were washed using the magnet stand 3 times in 0.8–1 ml Dig-washing buffer to remove unbound antibodies. The 1:100 dilution of pA–Tn5 adapter complex (∼20 nM) was prepared in Dig-300 buffer (20 mM HEPES(K+) pH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016); 300 mM NaCl (Invitrogen, AM9759); 0.5 mM spermidine (Sigma-Aldrich, S2626); 1× protease inhibitor cocktail (Sigma-Aldrich, 11873580001); 0.01% digitonin (Millipore, 300410)). After removing the liquid on the magnet stand, 100 μlof diluted pA–Tn5 adapter complex was added to the bead-bound nuclei with gentle mixing. The tubes were incubated at RT for 1 h on the nutator. Bead-bound nuclei were washed 3 times in 0.8–1 ml Dig-300 buffer to remove unbound pA–Tn5 protein. After washing, the liquid was removed on the magnetic stand, bead-bound nuclei were resuspended in 100 μltagmentation buffer (20 mM HEPES(K+) pH 7.5 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016); 300 mM NaCl (Invitrogen, AM9759); 0.5 mM Spermidine (Sigma-Aldrich, S2626); 1× protease inhibitor cocktail (Sigma-Aldrich, 11873580001); 0.01% digitonin (Millipore, 300410); 10 mM $MgCl_2$ (Invitrogen, AM9530G)). The mixture was incubated at 37°C for 1 h. To stop tagmentation and start the reverse-crosslinking, 10 μl0.5 M EDTA (Invitrogen, AM9260G), 33 μl10% SDS (Invitrogen, 1553-035) and 3.3 μl 20 mg/ml Proteinase K (Thermo Fisher scientific, EO0491) were added to each tube. Then the tubes were mixed by full speed vortexing ∼2 s, and incubated for 2 h at 65°C with 1200 rpm shaking. Then, the reverse-crosslinking was performed overnight under 72°C with 1200 rpm shaking. Next day, 3.3 μl 20 mg/ml Proteinase K (Thermo Fisher scientific, EO0491) was added to each tube after the mixture was cooled down to RT and the secondary Proteinase K (Thermo Fisher scientific, EO0491) digestion was performed by incubating the mixture on a shaker at 65°C with 1200 rpm shaking for 1 h. After digestion, equal volume of phenol (Thermo Fisher scientific, 17914) is added to each tube and the solution was mixed by full speed vortexing for ∼2 s. Then the mixture was centrifuged for 15 min at 4°C with 16 000 × g. After centrifugation, the liquid from aqueous layer was transferred to new 1.5 ml Lo-Bind tubes (Sarstedt, 72.706.600) and the equal amount of chloroform (Sigma-Aldrich, C2432) was added to each tube and mixed by gentle pipetting. The mixture was centrifuged for 15 min at 4°C with 16 000 × g. After centrifugation, the liquid from aqueous layer was transferred to a new 1.5 ml Lo-Bind tube (Sarstedt, 72.706.600) and 2.5×–3×– volume of absolute ethanol and proper amount of 5 M NaCl (Invitrogen, AM9759) were added to each tube to make the final NaCl (Invitrogen, AM9759) concentration 200 mM. After mixing, the DNA was precipitated at –80°C for 3–4 h. After precipitation, the DNA was centrifuged for 20 min at 4°C with 16 000 × g. Then, the supernatant was discarded, and the DNA was washed once with 700 μl 70% ethanol. After washing and drying up, the DNA was eluted in 21 μl of water. To amplify libraries, 21 μlDNA was mixed with 2 μl of a universal i5 and i7 primer (10 μM), using a different barcode for each sample. 25 μl of NEB Next HiFi 2× PCR master mix (New England Biolabs, M0541S) was added to the tubes and mixed well. The samples were placed in a Thermocycler with a heated lid using the following cycling conditions: 72°C for 5 min; 98°C for 30 s; 14 cycles of 98°C for 10 s, 63°C for 10 s; final extension at 72°C for 1 min and hold at 8°C. Post-PCR clean-up was performed by adding 1.3× volume of SPRI select beads (Beckman Coulter, B23317), and libraries were incubated with beads for 15 min at RT, washed twice gently in 80% ethanol, and eluted in 25 μl of Elution buffer from Qiagen MiniElute PCR Purification kit (Qiagen, 28004). The libraries were finally run on a 1.5% agarose gel (Lonza, 50004) using 120 V for 40 min to check the library pattern.

CUT&Tag in FFPE samples with centrifugation

Aliquots of purified mouse kidney FFPE mouse nuclei with or without epitope retrieval (100 000 cells/0.5 ml Qubit tube (Invitrogen, Q32856)) were washed once with FACT-seq Antibody buffer (20 mM HEPES(K+) pH 7.6 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016), 150 mM NaCl (Invitrogen, AM9759), 2 mM EDTA (INVITROGEN, AM9260G), 0.5 mM Spermidine (Sigma-Aldrich, S2626), 0.05% Digitonin (Millipore, 300410), 0.01% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50), 1× Protease inhibitors (Sigma-Aldrich, 11873580001), 1% BSA (Miltenyi Biotech MACS, 130-091-376)). After washing, the nuclei were resuspended in 200 μl Antibody buffer with 1:100 diluted primary antibody (anti-H3K27ac antibody (Abcam, ab4729) and anti-H3K27me3 antibody (Cell Signaling Technology, 9733S)) and incubated overnight at 4°C with slow rotation. Next day the nuclei were centrifuged 5 mins at 2000 × g, washed once with 200 μl of FACT-seq Dig-washing buffer (20 mM HEPES(K+) pH 7.6 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016), 150 mM NaCl (Invitrogen, AM9759), 0.5 mM spermidine (Sigma-Aldrich, S2626), 0.05% digitonin (Millipore, 300410), 0.01% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50), 1× protease inhibitors (Sigma-Aldrich, 11873580001), 1% BSA (Miltenyi Biotech MACS, 130-091-376)), and finally resuspended in 200 μl

of Dig-washing buffer with 1:100 diluted secondary antibody (Guinea Pig anti-Rabbit IgG antibody (AntibodiesonlineABIN101961)) and incubated for 1 h at RT with slow rotation. Then, nuclei were centrifuged 5 min at 2000 × g, washed three times with 200 μl of Dig-washing buffer and resuspended in 200 μl of Dig-300 buffer (20 mM HEPES(K+) pH 7.6 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016), 300 mM NaCl (Invitrogen, AM9759), 0.5 mM spermidine (Sigma-Aldrich, S2626), 0.05% digitonin (Millipore, 300410), 0.01% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50), 1× protease inhibitors (Sigma-Aldrich, 11873580001), 1% BSA (Miltenyi Biotech MACS, 130-091-376)) with 1:100 diluted pA–Tn5 and incubated for 1 h with slow rotation at RT. After pA–Tn5 binding, the nuclei were centrifuged 5 min at 600 × g, washed three times with 200 μl of Dig-300 buffer and resuspended in 200 μl of FACT-seq tagmentation buffer (20 mM HEPES(K+) pH 7.6 (HEPES: Sigma-Aldrich, H3375; KOH: Sigma-Aldrich, 484016), 300 mM NaCl (Invitrogen, AM9759), 0.5 mM Spermidine (Sigma-Aldrich, S2626), 0.05% Digitonin (Millipore, 300410), 0.01% IGEPAL® CA-630 (Sigma-Aldrich, 13021-50), 1× protease inhibitors (Sigma-Aldrich, 11873580001), 10 mM $MgCl_2$ (Invitrogen, AM9530G)), and incubated for 1 h at 37°C. After that, tagmentation was stopped by addition of 6.7 μl 0.5 M EDTA (Invitrogen, AM9260G), 22 μl 10% SDS (Invitrogen, 1553-035) and 2.2 μl 20 mg/ml Proteinase K (Thermo Fisher scientific, EO0491) to each tube, and the mixture was mixed by full speed vortexing for ∼2 s. Then the first Proteinase K (Thermo Fisher scientific, EO0491) digestion was performed by incubating the mixture on a shaker at 65°C with 1200 rpm shaking for 2 h. After digestion, the reverse-crosslinking was performed overnight under 72°C with 1200 rpm shaking. Next day, 2.2 μl 20 mg/ml Proteinase K (Thermo Fisher scientific, EO0491) was added to each tube after the mixture was cooled down to RT and the secondary Proteinase K (Thermo Fisher scientific, EO0491) digestion was performed by incubating the mixture on a shaker at 65°C with 1200 rpm shaking for 1 h. After digestion, 67 μl of water and 300 μl of phenol were added to each tube and the solution was mixed by full speed vortexing for ∼2 s. Then the mixture was centrifuged for 15 min at 4°C with 16 000 × g. After centrifugation, the liquid from aqueous layer was transferred to new 1.5 ml Lo-Bind tube (Sarstedt, 72.706.600) and the equal amount of chloroform was added to each tube and mixed by gentle pipetting. The mixture was centrifuged for 15 min at 4°C with 16 000 × g. After centrifugation, the liquid from aqueous layer was transferred to new 1.5 ml Lo-Bind tube (Sarstedt, 72.706.600) and 2.5×–3× volume of absolute ethanol and proper amount of 5 M NaCl (Invitrogen, AM9759) were added to each tube to make the final NaCl (Invitrogen, AM9759) concentration 200 mM. After mixing, the DNA was precipitated at –80°C for 2–3 h. After precipitation, the DNA was centrifuged for 20 min at 4°C with 16 000 × g. Then, the supernatant was discarded, and the DNA was washed once with 700 μl 70% ethanol. After washing and drying up, the DNA was eluted in 21 μl of water. To amplify libraries, 21 μl DNA was mixed with 2 μl of a universal i5 and i7 primer (10 μM), using a different barcode for each sample. 25 μl of NEB Next HiFi 2× PCR master mix (New England Biolabs, M0541S) was added to the tubes and mixed well. The samples were placed in a Thermocycler with a heated lid using the following cycling conditions: 72°C for 5 min; 98°C for 30 s; 13 cycles of 98°C for 10 s and 63°C for 10 s; final extension at 72°C for 1 min and hold at 8°C. Post-PCR clean-up was performed by adding 1.3× volume of SPRI select beads (Beckman Coulter, B23317), and libraries were incubated with beads for 15 min at RT, washed twice gently in 80% ethanol, and eluted in 25 μl of Elution buffer from Qiagen MiniElute PCR Purification kit (Qiagen, 28004). The libraries were finally run on a 1.5% agarose gel using 120 V for 40 min to check the library pattern. The positive control groups on the gel were from the previously prepared corresponding CUT&Tag libraries from GM12878 cell line.

Immunostaining

30 000–50 000 FFPE mouse nuclei either with or without epitope retrieval were transferred to the glass slides using Cytospin (Double Cellfunnel (THARMAC, 306-12), Filter cards (THARMAC, 307–500)) with 600 × g for 6 min. Then the slides containing FFPE nuclei were rinsed in 1 × PBS and permeabilized in PBS with 0.5% Triton-X 100 (Sigma-Aldrich, T8787) for 10 min at RT. Then, the slides were blocked with blocking buffer (1% BSA (Miltenyi Biotech MACS, 130-091-376) and 0.1% Triton-X 100 (Sigma-Aldrich, T8787) in 1 × PBS) for 1 h at RT. Primary antibodies were diluted in blocking buffer 1:500 (anti-H3K27ac antibody (abcam, ab4729) and anti-H3K27me3 antibody (Cell Signaling Technology, 9733S)). The nuclei were covered by blocking buffer containing 1:500 primary antibodies and were incubated overnight at 4°C. After washing with 1× PBS containing 0.05% Tween-20 (Sigma-Aldrich, P416) for 3 times 10 min each, slides were incubated with 1:500 secondary antibodies in blocking buffer (goat anti-rabbit-Atto-488 (Sigma-Aldrich, 18772-1ml-F) or goat anti-rabbit-Atto-594 (Sigma-Aldrich, 77671-1ml-F) for 1 h at RT. The slides were washed with 1× PBS containing 0.05% Tween-20 (Sigma-Aldrich, P416) for 3 times 10 mins each. Then the nuclei were stained with 1:2000 DAPI (Invitrogen, H3570) in 1× PBS for 15 min at RT. Finally, the slides were mounted using SlowFade Gold antifade reagent (Invitrogen, S36940) and imaged with Zeiss Axio Imager Z2.

Timing for FACT-seq

Day 1 (1 h 40 min):

Step 1: Tissue sectioning (5 min);
Step 2: Deparaffinization and rehydration (1 h 5 min);
Step 3: Microdissection (30 min);
Step 4: Enzymatic digestion (16 h).

Day 2 (10 hs and 55 min):

Step 5: Nuclei isolation (1 h);
Step 6: Epitope retrieval + Quenching + Washing (1 h and 45 min);
Step 7: Nuclei Counting + Splitting (20 min);
Step 8: Primary antibody incubation (1 h and 30 min);

Step 9: Washing + Secondary antibody incubation (1 h and 15 min);
Step 10: Washing + T7 pA–Tn5 incubation (1 h and 30 min);
Step 11: Washing + T7 pA–Tn5 tagmentation (1 h and 30 min);
Step 12: Stop T7 pA–Tn5 tagmentation + 1st Proteinase K digestion (2 hs 5 min);
Step 13: Reverse crosslinking (8 hs or overnight).

Day 3 (2 hs and 45 min):

Step 14: 2nd Proteinase K digestion (1 h);
Step 15: DNA purification (15 min);
Step 16: Gap filling + DNA purification (25 min);
Step 17: Beads purification + adding in vitro transcription (IVT) reagents (1 h 5 min);
Step 18: IVT (12 hs or overnight).

Day 4 (7 h):

Step 19: RNA purification (50 min);
Step 20: Reverse transcription + RNase H digestion (2 h);
Step 21: RNA clean XP beads purification (1 h);
Step 22: Pre-PCR + DNA purification (20 min);
Step 23: Tn5 tagmentation + DNA purification (25 min);
Step 24: PCR amplification (1 h);
Step 25: DNA purification (15 min);
Step 26: Gel purification (1 h and 10 min);
Step 27: DNA elution (8 hs or overnight).

Day 5 (1h and 50 min):

Step 28: DNA sequencing library purification (40 min);
Step 29: Sequencing library quantification (1 h and 10 min).

Sequencing data analysis

Only reads one (R1) was involved in our analysis. We first adopted levenshtein distance algorithm to map T7 promoter sequence to each read in the adaptor trimming procedure, then T7 promoter sequences were trimmed from each read in the fastq file with in-house script custom script (https://github.com/pengweixing/FACT). The trimmed fastq file of human samples and mouse samples were mapped to the hg38 reference genome or mm10 reference genome respectively using bowtie2 v.2.3.5 (51) with the parameters –end-to-end –very-sensitive -I 10 -X 700. The aligned BAM files were sorted by samtools v.1.9 (52) and were filtered with alignment quality of >q2. The bigwig files were generated from BAM file using deeptools (53) with the parameters bamCoverage –normalizeUsing CPM. Transcription start sites (TSS) enrichments of sequencing libraries were calculated using deeptools (53) with the option computeMatrix, and heatmaps were plotted by deeptools (53) with the options plotHeatmap and plotProfile. The reads count within peaks were calculated by bedtools (54) with the option multicov. The peaks of sequencing libraries were visualized by IGV software (55). The VennDigram was plotted with R pacakge venn. Genomic annotation of peaks was performed using the ChIPseeker R package (56). For the peak calling, H3K27me3, H3K36me3 and H3K4me1 peaks were called using SICER (57) with following parameters gap size = 600 bp and window size = 200 bp. H3K27ac peaks were called using SICER with following parameters gap size = 400 bp and window size = 200 bp. The GM12878 ATAC-seq data set was downloaded from GEO under accession number, GSE76006 (58). The ENCODE ChIP-seq data sets (59) used for comparison were downloaded from GEO under accession numbers GSM733771 (GM12878, H3K27ac), GSM945196 (GM12878, H3K27me3), GSM1000092 (adult 8 weeks Mouse kidney, H3K27ac), GSM1000077 (adult 8 weeks Mouse kidney, H3K27me3), GSM1000063 (adult 8 week mouse kidney, H3K36me3), GSM769023 (adult 8 week mouse kidney, H3K4me1), GSM1000140 (adult 8 week mouse liver, H3K27ac). H3K27ac FiTAc-seq data used for comparison was downloaded from GEO under accession numbers GSM4186366 (H3K27ac FiTAc-seq_65°C_5 mins_rep1), GSM4186368 (H3K27ac FiTAc-seq_65°C_5mins_rep2). The RNA-seq data of human CRC and human GBM were downloaded from GEO with the accession numbers GSE158559 (CRC tissue), GSE119834 (GBM tissue). Super enhancer (SE) elements were mapped and quantified by MACS version 2 (60) and ROSE (61,62) software based on H2K27ac. A p-value threshold of enrichment of 1×10^{-5} was used during the peak calling with MACS. Genes were assigned to SEs by proximity. Clustering analysis of human GBM H3K27ac FACT-seq and human CRC H3K27ac FACT-seq was performed with an unsupervised non-negative matrix factorization algorithm which implemented on NMF software package (63). Library complexity for each sample was estimated with Preseq v3.1.2 (64). Ontological enrichment of SE-associated genes was analiyzed by following the protocle as described (65). GO analysis of SE target genes was calculated by g:Profiler (66) with significance threshold set to Benjamini–Hochberg false discovery rate (FDR) (0.05) and the pathway enrichment was annotated and visualized in Cytoscape (67) with Enrichment Map (68) and AutoAnnotate (69) by taking GO analysis results as input. Transcriptor factor enrichment for histone modifications peaks was performed with Homer v.4.11 software (70). The correlation matrix for each data set was clustered using the R package pheatmap. Oncogenes and tumor suppressor genes were downloaded from the COSMIC Cancer Gene Census databse (71). All of the scripts which involved in this study were deposited in (https://github.com/pengweixing/FACT).

RESULTS

DNA damage in FFPE samples hampers the application of CUT&Tag in FFPE samples

To profile chromatin structure in the FFPE samples, we established a protocol to isolate high-quality intact single nuclei from 20-μm-thick mouse FFPE kidney tissue sections (Supplementary Figure S1A, see Materials and Methods). When we performed H3K27me3 CUT&Tag and H3K27ac CUT&Tag with 100 000 isolated FFPE nuclei following the standard CUT&Tag protocol for fixed cells (see Methods), we did not obtain PCR amplicons (Figure 1A and B) potentially because (i) extensive fixation during FFPE

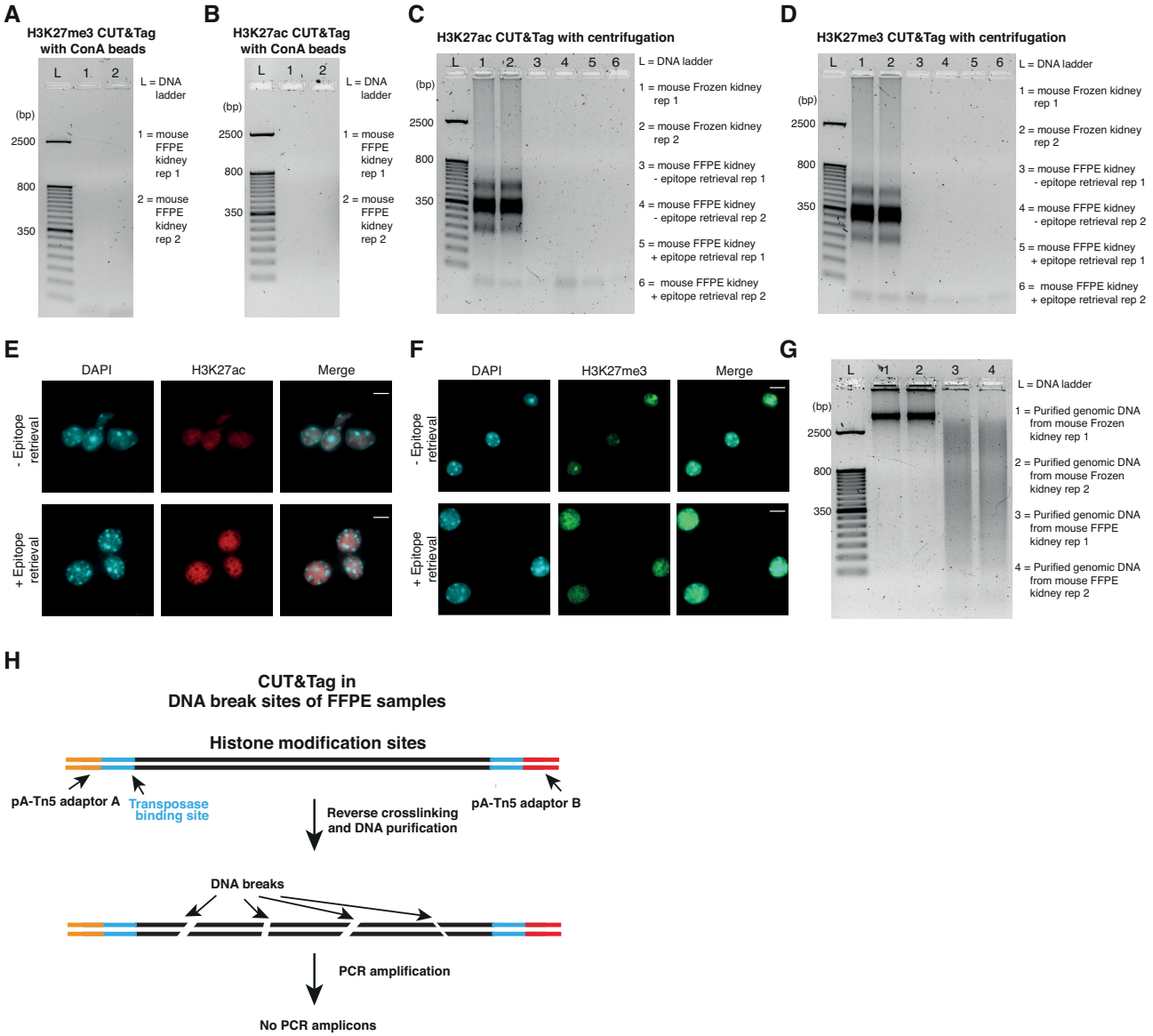

Figure 1. DNA damage in genomic DNA purified from FFPE samples hampers the application of CUT&Tag in FFPE samples. (**A**) No PCR amplicons are observed with a standard H3K27me3 CUT&Tag protocol in purified mouse FFPE kidney nuclei. (**B**) No PCR amplicons are observed with a standard H3K27ac CUT&Tag protocol in purified mouse FFPE kidney nuclei. (**C**) No PCR amplicons are observed with a modified H3K27ac CUT&Tag protocol using centrifugation instead of ConA beads, with and without epitope retrieval in purified mouse FFPE kidney nuclei. (**D**) No PCR amplicons are observed with a modified H3K27me3 CUT&Tag protocol using centrifugation instead of ConA beads, with and without epitope retrieval in purified mouse FFPE kidney nuclei. (**E**) Immunostaining of H3K27ac from purified mouse FFPE kidney nuclei with and without epitope retrieval conditions, Scale bar = 10 μm. (**F**) Immunostaining of H3K27me3 from purified mouse FFPE kidney nuclei with and without epitope retrieval conditions, Scale bar = 10 μm. (**G**) Comparison of genome DNA fragments purified from frozen mouse kidney nuclei and mouse FFPE kidney nuclei. (**H**) DNA damage in FFPE tissues hampers PCR amplification with the standard CUT&Tag protocol.

sample preparation changes the conformation of glycosylated proteins in the nucleus membranes and further interferes with the binding of lectin concanavalin A (ConA) magnetic beads and isolated FFPE nuclei in the CUT&Tag (ii) the epitope of H3K27me3 and H3K27ac in the isolated FFPE nuclei is masked during FFPE tissue preparation and (iii) there is a high degree of DNA damage to the extracted DNA from the FFPE isolated nuclei. To solve these potential issues, we first used a centrifugation strategy instead of ConA beads to enrich isolated FFPE nuclei and performed H3K27ac and H3K27me3 CUT&Tag with FFPE nuclei; however, we still did not obtain PCR amplicons (Figure 1C and D). Next, we established an epitope retrieval protocol for the isolated FFPE nuclei (see Materials and Methods). Even though the signal intensity of H3K27ac and H3K27me3 immunostaining from these nuclei was significantly improved with epitope retrieval (Figure 1E and F), no PCR product was generated from H3K27ac and H3K27me3 CUT&Tag with centrifugation and epitope retrieval steps (Figure 1C and D). During FFPE tissue preparation, the formaldehyde in formalin reacts with primary amines in chromatin to form

Schiff bases, and with amides to form hydroxymethyl compounds, resulting in large chromatin complexes (9). To decode histone modifications in FFPE samples, it is essential to break such chromatin complexes with reverse crosslinking (7,58), which is also applied to the CUT&Tag protocol for fixed cells that we used. However, DNA damage can be introduced during reverse cross-linking in FFPE samples. Indeed, we observed that many DNA breaks were generated during reverse crosslinking in the extracted DNA from the isolated FFPE nuclei (Figure 1G). In antibody-guided hyperactive Tn5 transposase-protein A (pA–Tn5) tagmentation based technologies (20–23,25–26,42), including CUT&Tag, the amplification of transposase insertion sites from target regions with polymerase chain reaction (PCR) is the key step (20). However, PCR amplification does not work efficiently for FFPE samples if there is the high degree of DNA damage present in these samples (Figure 1H). Thus, we conclude that standard CUT&Tag does not work well in the FFPE samples probably because of the high degree of DNA damage present in the FFPE sample after reverse crosslinking (Figure 1H).

Design of FACT-seq

The strategy in transposed-based chromatin profiling technologies is to use the inserted antibody-guided pA–Tn5 adaptors to mark the location of the histone modifications or transcription factors (TFs) binding (20,21). Since the adaptors of pA–Tn5 transposase are inserted into the genome after formalin fixation in CUT&Tag, breakpoints generated during reverse crosslinking are unlikely to occur at the adaptors of pA–Tn5 transposase. Thus, we assumed that the adaptors of pA–Tn5 transposase are still at the end of broken targeting chromatin sites after reverse crosslinking. We reasoned that if the T7 promoter sequence is added to the adaptors of pA-T5 transposase, even if there are DNA breaks in the middle of DNA fragment we could still decode the insertion sites of pA–Tn5 transposase by transferring broken DNA fragments into RNA molecules with *in vitro* transcription (IVT) and making DNA sequencing libraries from the IVT RNA molecules (Figure 2A). Recently, we had used such a strategy to successfully profile the chromatin accessibility in FFPE samples with high sensitivity (45), and we assumed the same principle could be applied in profiling histone modifications in FFPE samples. The activity of the Tn5 transposase is very robust with different sequence modifications on the adaptors (24,43,58,72–73), and the combination of linear amplification with the T7 promoter and Tn5 tagmentation for enriching target sequences is widely used in a series of high sensitivity technologies, including chromatin integration labelling with sequencing (ChIL-seq) (23,24), Linear Amplification via Transposon Insertion (LIANTI) (73), transposome hypersensitive sites sequencing (THS-seq) (72,74), multiplexed, indexed T7 ChIP-seq (Mint-ChIP) (32) and many others. Thus, we hypothesized that modifications to the adaptors of pA–Tn5 transposase with the T7 promoter would also be feasible. Therefore, we designed, produced and optimized a novel pA–Tn5 transposase with the T7 promoter sequence on the adaptor, termed T7−pA−Tn5 (Figure 2B, see Materials and Methods). T7−pA−Tn5 retains the activity of the standard pA–Tn5 (Supplementary Figure S1B, see Materials and Methods). While our work is ongoing, a new high-throughput and highly sensitive protein:DNA binding site mapping technology, Targeted Insertion of Promoters (TIP-seq) (43), was established, where a similar strategy of adding the T7 promoter sequence into the pA–Tn5 adaptor was also used and proved that the activity of pA–Tn5 is not affected by adding the T7 promoter sequence (43). To test that T7−pA−Tn5 adaptors are still at the end of broken DNA fragments after reverse crosslinking from the isolated FFPE nuclei, we performed IVT on the DNA extracted from mouse kidney FFPE nuclei after H3K27ac-guided T7−pA−Tn5 tagmentation, and we found that RNA molecules from the IVT contain both short- and long- range fragments (Supplementary Figure S1C). These results convinced us that T7−pA−Tn5 adaptors were still at the ends of broken DNA fragments, and those inserted sites of the T7−pA−Tn5 adaptor in histone modification sites could be read out by making a DNA sequencing library from IVT RNA molecules (Figure 2C). In the CUT&Tag or similar technologies (20,21), one insertion event or unpaired pA–Tn5 adaptor insertions from pA–Tn5 transposition could not be amplified with PCR (75). Since the IVT depends on only one insertion of the T7 promoter, one insertion event and unpaired adaptor insertions on targeting sites from T7−pA−Tn5 can be detected with FACT-seq during IVT (Figure 2A). Thus, we expect the library complexity of FACT-seq to be is higher than that of CUT&Tag or similar technologies.

FACT-seq accurately decodes both repressive and active histone modifications with low cell numbers

To prove the combination of T7−pA−Tn5 transposition and T7 IVT is feasible to decode histone modifications, we first performed H3K27me3 FACT-seq with human B-cell (GM12878 cells) and sequenced the H3K27me3 FACT-seq sequencing library with ∼50 million sequencing reads for each technical replicate (see Materials and Methods). The two technical replicates of H3K27me3 FACT-seq had high similarity (Supplementary Figure S2A and B, $R = 0.99$). We found that 20 million sequencing reads were needed for H3K27me3 FACT-seq to obtain peak number saturation by downsampling the total reads with different ranges (Supplementary Figure S2C), and H3K27me3 FACT-seq peaks with 20 million sequencing reads could reach over 63.17% overlap with peaks from Encyclopedia of DNA Elements (ENCODE) GM12878 H3K27me3 ChIP-seq (Supplementary Figure S2D). In the following steps, we used H3K27me3 FACT-seq with a 20 million-read sequencing depth to perform a detailed comparison with ENCODE H3K27me3 ChIP-seq and standard H3K27me3 CUT&Tag (Supplementary Figure S2E–J, see Materials and MethodsMethods). FACT-seq clearly revealed that sequencing read coverage of H3K27me3 is highly correlated with standard H3K27me3 CUT&Tag (Figure 3A–C, $R = 0.85$). The genome-wide correlation of H3K27me3 FACT-seq and ENCODE H3K27me3 ChIP-seq (Supplementary Figure S2G, $R = 0.83$) and the correlation of H3K27me3 CUT&Tag and ENCODE H3K27me3 ChIP-seq (Supplementary Figure S2H, $R = 0.82$) were in the similar ranges.

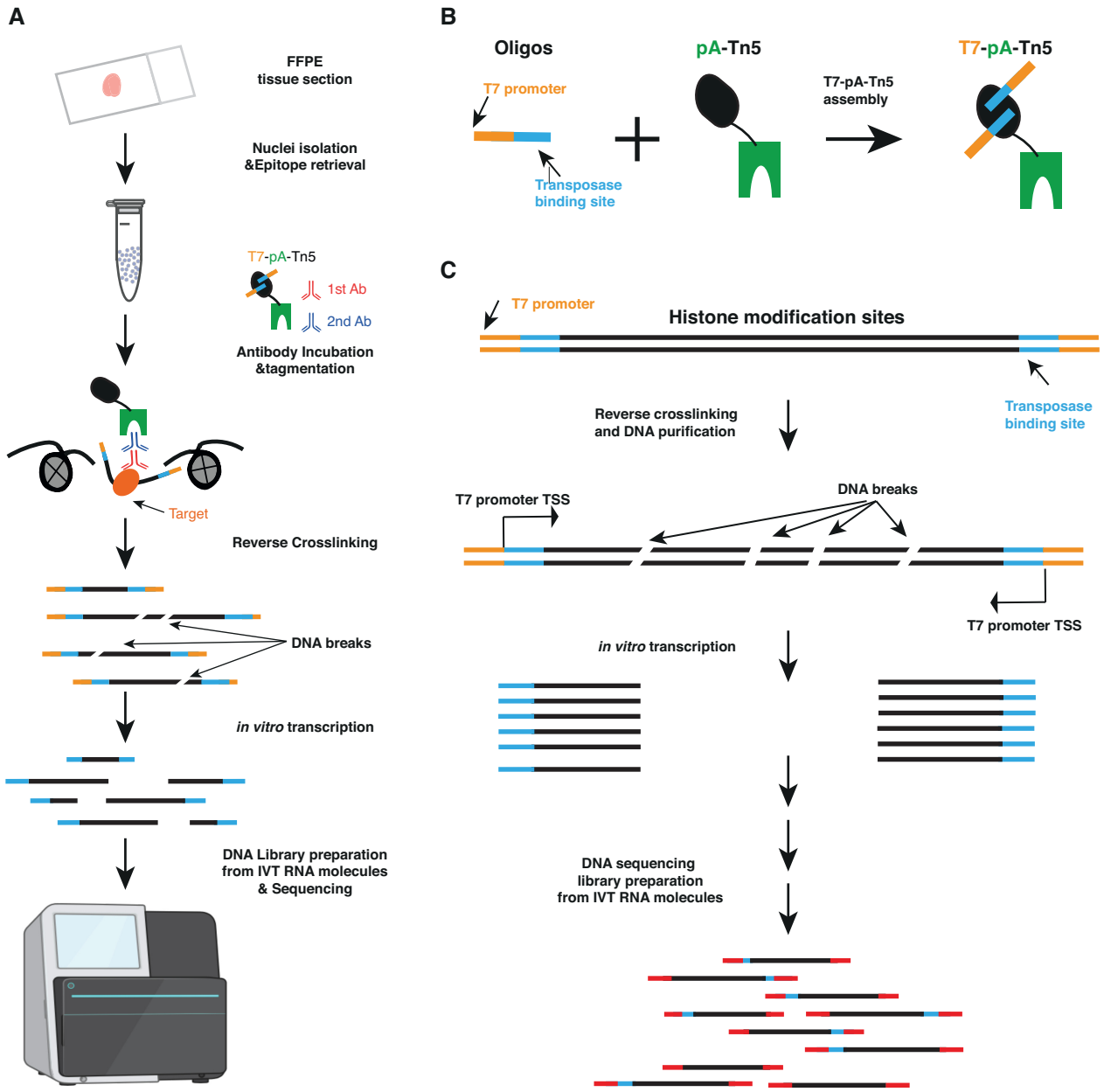

Figure 2. The design of FACT-seq. (**A**) Workflow of FACT-seq. (**B**) Design of T7−pA−Tn5 transposase. (**C**) FACT-seq maps of histone modifications in FFPE samples by combining pA–Tn5-mediated transposition and T7 *in vitro* transcription.

The distributions of peak sizes from H3K27me3 FACT-seq, H3K27me3 CUT&Tag, and ENCODE H3K27me3 ChIP-seq were also within similar ranges (Supplementary Figure S2I). Moreover, we used the peak centers of ENCODE H3K27me3 ChIP-seq for GM12878 cells as the reference points to calculate the sequencing signal enrichments from H3K27me3 FACT-seq and H3K27me3 CUT&Tag, and we observed clear enrichment of the H3K27me3 FACT-seq signal at ENCODE H3K27me3 ChIP-seq sites, which was very similar to the enrichment of H3K27me3 CUT&Tag at ENCODE H3K27me3 ChIP-seq sites (Supplementary Figure S2J). We also found that 71% of peaks from standard H3K27me3 CUT&Tag overlapped with H3K27me3 FACT-seq peaks (Supplementary Figure S3A), and the fractions of overlapping peaks between H3K27me3 FACT-seq and ENCODE H3K27me3 ChIP-seq were similar to those of overlapping peaks between H3K27me3 CUT&Tag and ENCODE H3K27me3 ChIP-seq (Supplementary Figure S3B and C). The genomic annotation of the overlapping peaks and unique peaks from H3K27me3 FACT-seq, H3K27me3 CUT&Tag and ENCODE H3K27me3 ChIP-seq showed that both overlapping peaks and exclusive peaks from different types of libraries were distributed in a similar pattern relative to the transcription start sites (TSS) (Supplementary Figure S3A–C). At the same time, we found that different transcription factors (TFs) were enriched at exclu-

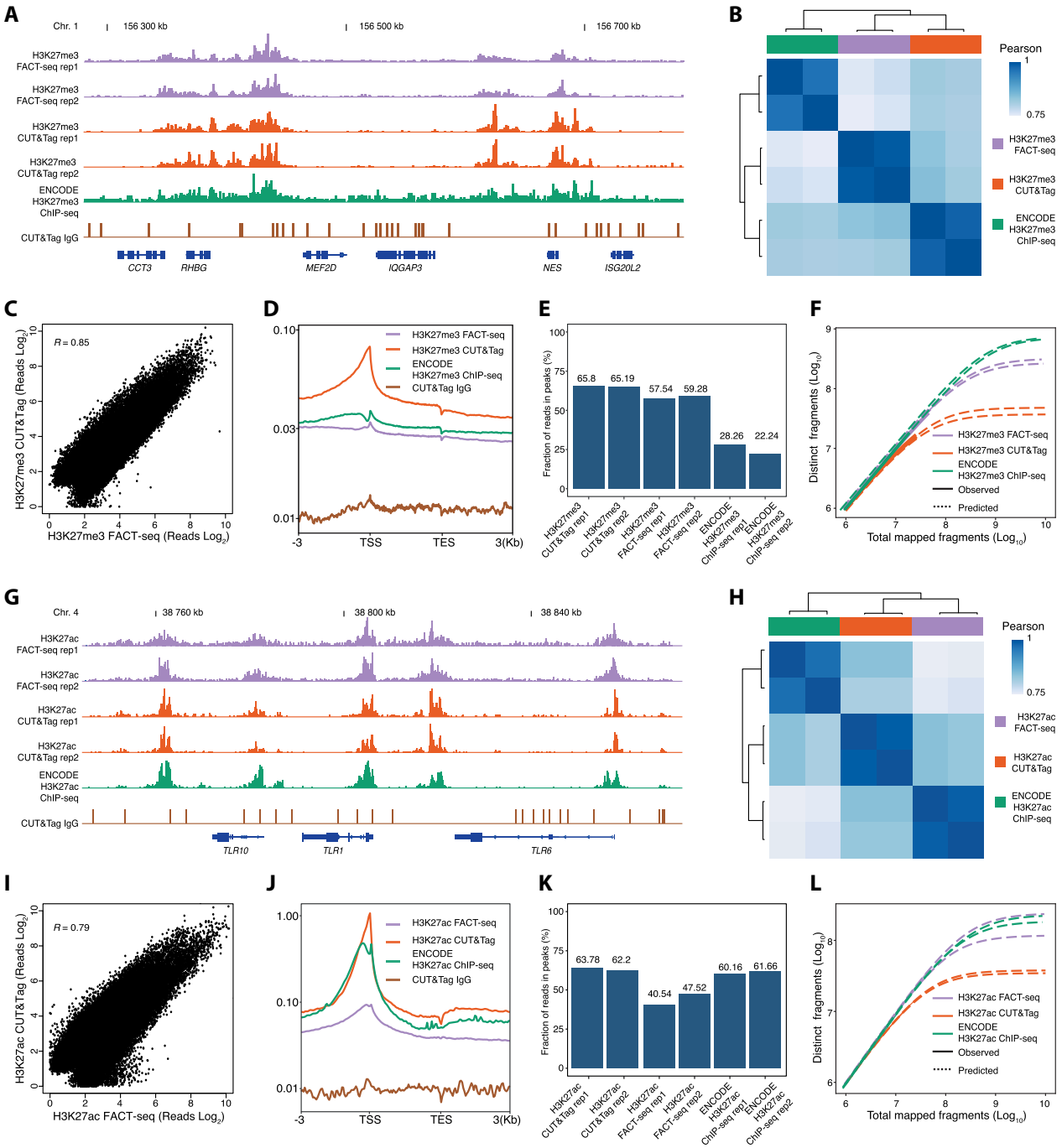

Figure 3. FACT-seq decodes genome-wide repressive and active histone modifications in human B-cells (GM12878 cells). (**A**) Genome browser tracks for H3K27me3 sequencing libraries in GM12878 cells: Results from FACT-seq, CUT&Tag, ENCODE ChIP-seq and CUT&Tag IgG control. The gene names are shown at the bottom. Chr. = Chromosome. (**B**) Pearson correlation heat map showing sample-by-sample unsupervised clustering on all peaks identified across different types of libraries. (**C**) Genome-wide correlation of H3K27me3 FACT-seq sequencing reads and H3K27me3 CUT&Tag sequencing reads from GM12878 cells. Each dot represents an individual peak. R = Pearson correlation. (**D**) Metaplots showing the distribution of sequencing reads across transcription start sites (TSSs) and transcription end sites (TESs) from H3K27me3 FACT-seq, H3K27me3 CUT&Tag, ENCODE H3K27me3 ChIP-seq, and H3K27me3 CUT&Tag IgG control in GM12878 cells. (**E**) Quantification of the fraction of reads in peaks (FRiP) for each sequencing library. (**F**) Library complexity for different sequencing libraries. (**G**) Genome browser tracks for H3K27ac sequencing libraries in GM12878 cells: Results from FACT-seq, CUT&Tag, ENCODE ChIP-seq, and CUT&Tag IgG control. The gene names are shown at the bottom. Chr. = Chromosome. (**H**) Pearson correlation heat map showing sample-by-sample unsupervised clustering on all peaks identified across different types of libraries. (**I**) Genome-wide correlation of H3K27ac FACT-seq sequencing reads and H3K27ac CUT&Tag sequencing reads from GM12878 cells. Each dot represents an individual peak. R = Pearson correlation. (**J**) Metaplots showing the distribution of sequencing reads across transcription start sites (TSS) and transcription end sites (TES) from H3K27ac FACT-seq, H3K27ac CUT&Tag, ENCODE H3K27ac ChIP-seq and H3K27ac CUT&Tag IgG control in GM12878 cells. (**K**) Quantification of the fraction of reads in peaks (FRiP) for different sequencing libraries. (**L**), Estimation of sequencing library complexity for each sequencing libraries.

sive peaks (Supplementary Figure S3A–C), e.g., FOXP and FOXA TFs were enriched at unique peaks from CUT&Tag sequencing libraries, MEF2 and STAT TFs were found at unique peaks of FACT-seq sequencing libraries, TEAD family TFs were identified in unique peaks of ENCODE ChIP-seq sequencing libraries. Since the same antibodies were used for H3K27me3 CUT&Tag and H3K27me3 FACT-seq, the enrichments of different TFs in the unique peaks of H3K27me3 CUT&Tag and H3K27me3 FACT-seq indicate that there is potential technical bias between CUT&Tag and FACT-seq. The different TFs enrichments in the unique peaks of ENCODE H3K27me3 ChIP-seq may be because of the different antibodies used and the technical bias between the ChIP-seq and transposase-based methods.

In addition, we also compared the signal distribution from H3K27me3 FACT-seq, H3K27me3 CUT&Tag and H3K27me3 ENCODE ChIP-seq in the genome by using TSS as reference points (see Materials and Methods, Figure 3D, Supplementary Figure S4A). Strikingly, we found that the signal distribution of H3K27me3 FACT-seq near the TSSs was more similar to that of ENCODE H3K27me3 ChIP-seq than that of H3K27me3 CUT&Tag (Supplementary Figure S4A). There was strong enrichment at TSSs sites from H3K27me3 CUT&Tag but not from H3K27me3 FACT-seq or H3K27me3 ENCODE ChIP-seq (Figure 3D, Supplementary Figure S4A). It has been reported that ATAC-seq signal contamination from standard CUT&Tag was observed (20). Thus, we assumed that the signal enrichment at TSSs sites from H3K27me3 CUT&Tag is potentially from ATAC-seq signals. H3K27me3 is a repressive chromatin marker that is usually located at distal regulatory elements in the genome but is also found at silenced promoters and bivalent promoters (76–78). To test our hypothesis, we first broke GM12878 cell ATAC-seq peaks (58) into promoter (±1 kb from TSSs) and nonpromoter regions, then calculated sequencing signal enrichments from H3K27me3 FACT-seq, H3K27me3 CUT&Tag and ENCODE H3K27me3 ChIP-seq in promoter and nonpromoter regions of the GM12878 ATAC-seq peaks (Supplementary Figure S4B–E). We found that there was much stronger signal enrichment around the ATAC-seq peaks from H3K27me3 CUT&Tag than H3K27me3 FACT-seq, and ENCODE H3K27me3 ChIP-seq at both the promoter and non-promoter regions of the ATAC peaks (Supplementary Figure S4B). The quantification of the fraction of reads within the ATAC-seq peaks from H3K27me3 CUT&Tag, H3K27me3 FACT-seq, and ENCODE H3K27me3 ChIP-seq showed that 20% of reads from H3K27me3 CUT&Tag were from the promoter regions of the ATAC-seq peaks, while only 13% of reads from H3K27me3 FACT-seq were located in the promoter regions of the ATAC-seq peaks, which was in the same range as ENCODE H3K27me3 ChIP-seq (9%) (Supplementary Figure S4C). A similar tendency was also observed in nonpromoter regions of the ATAC-seq peaks (Supplementary Figure S4D and E), where 18% of sequencing reads from H3K27me3 CUT&Tag were in nonpromoter regions of the ATAC-seq peaks, and 14% of sequencing reads from H3K27me3 FACT-seq are located in the non-promter regions of the ATAC-seq peaks. The contamination of the ATAC-seq signals in H3K27me3 FACT-seq and H3K27me3 CUT&Tag was potentially due to the nonspecific binding and insertion of pA–Tn5 and T7−pA−Tn5. Since insertion events in H3K27me3 FACT-seq and H3K27me3 CUT&Tag are controlled by antibody guidance and the presence of Mg2+ (20), the proportions of insertion events from pA–Tn5 and T7−pA−Tn5 nonspecific binding and insertion should be similar in FACT-seq and CUT&Tag. We suspected that the exponential amplification PCR amplification (79) in CUT&Tag could enlarge the proportion of DNA fragments from the nonspecific binding of pA–Tn5, which does not occur in FACT-seq because of the linear amplification that occurs in T7 IVT (80). In the CUT&Tag protocol, one insertion event or unpaired adaptor insertions from pA–Tn5 could not be amplified, and long fragments could not be sequenced with short read sequencing (75). However, one insertion event and unpaired adaptor insertions from targeting sites could be captured with FACT-seq during IVT (20). At the same time, long DNA fragments from histone modification sites become shorter when transferring target DNA fragments to RNA molecules with IVT and could be sequenced from IVT RNA sequencing libraries. Consequently, the proportion of DNA fragments in the sequencing library from nonspecific binding and insertion was lower in FACT-seq than in CUT&Tag. Thus, the proportion of contamination of ATAC-seq in H3K27me3 FACT-seq was lower. We also compared the fraction of reads in peaks (FRiP) from the sequencing libraries of H3K27me3 CUT&Tag, H3K27me3 FACT-seq and ENCODE H3K27me3 ChIP-seq, and found that the FRiP was in similar ranges for CUT&Tag (65.5% +/− 0.43%) and FACT-seq (58.41% ± 1.23%), but was much lower for ENCODE ChIP-seq (25.25% ± 4.26%) (Figure 3E). As we expected, the library complexity of H3K27me3 FACT-seq was much higher than that of CUT&Tag (Figure 3F), and the same conclusion was also proposed for TIP-seq (43). In summary, we conclude that FACT-seq can efficiently map H3K27me3 modifications with low cell numbers, and that the H3K27me3 FACT-seq library has a lower proportion of ATAC-seq signal contamination in both promoter and nonpromoter regions than CUT&Tag.

Next, we performed FACT-seq by targeting active histone modification H3K27ac from GM12878 (see Materials and Methods). Technical replicates of H3K27ac FACT-seq also had good reproducibility (Supplementary Figure S5A and B, $R = 0.97$). H3K27ac FACT-seq had the high similarity to standard H3K27ac CUT&Tag (Supplementary Figure S5C and D) in terms of the correlation of sequencing read coverage (Figure 3G–I, $R = 0.79$). In addition, the correlation between H3K27ac FACT-seq and ENCODE H3K27ac ChIP-seq was 0.61 (Supplementary Figure S5E), and the correlation between H3K27ac CUT&Tag and ENCODE H3K27ac (Supplementary Figure S5F) was 0.75. There was a large proportion of overlapping peaks among H3K27ac FACT-seq, H3K27ac CUT&Tag and ENCODE H3K27ac ChIP-seq (Supplementary Figure S5G–I). The genomic annotation of unique peaks from H3K27ac FACT-seq, H3K27ac CUT&Tag and ENCODE H3K27ac ChIP-seq showed that exclusive peaks from the different types of libraries were distributed in a similar pattern relative to the TSSs (Supplementary Figure S5G–I). We also

found that there are similar TFs, e.g. FRA1, FRA2, AP-1 and JUNB, were enriched in the unique peaks of H3K27ac FACT-seq and H3K27ac CUT&Tag, but different TFs, e.g., IRF family and ETV family TFs, were enriched at unique in the unique peaks of ENCODE H3K27ac ChIP-seq. Different TFs were identified in the unique peaks of ENCODE H3K27ac ChIP-seq are most likely because of technical bias or different antibodies used in the ENCODE H3K27ac ChIP-seq. Additionally, the peak sizes from H3K27ac FACT-seq, H3K27ac CUT&Tag, and ENCODE H3K27ac ChIP-seq had similar distributions (Supplementary Figure S6A). The sequencing reads from H3K27ac FACT-seq, H3K27ac CUT&Tag and ENCODE H3K27ac ChIP-seq all had strong enrichments at the TSSs(Figure 3J, Supplementary Figure S6B). The distribution of peaks from these three types of libraries using TSSs as reference points was also similar (Supplementary Figure S6C), and genomic annotation of peaks showed that the peaks from these three types of libraries had similar features (Supplementary Figure S6D). We noticed that the FACT-seq had lower signal enrichment for H3K27ac relative to CUT&Tag or ENCODE ChIP-seq at the TSSs (Figure 3J), and we assumed that the relatively lower signal recovery of H3K27ac from FACT-seq could be another reason that H3K27me3 FACT-seq resultied in a lower proportion of ATAC-seq signals at promoter and nonpromoter regions (Supplementary Figure S4B–E). Comparison of the FRiPs from the three different types of libraries showed that the proportion of signals within the peaks from H3K27ac FACT-seq (44.03% ± 4.94%) was lower than that from H3K27ac CUT&Tag (62.99% ± 1.12%) and ENCODE H3K27ac ChIP-seq (60.91% ± 1.06%) (Figure 3K). The sequencing library complexity of H3K27ac FACT-seq was within a similar range of as that of ChIP-seq and was much higher than that of CUT&Tag (Figure 3L).

In summary, we conclude that design of FACT-seq by combining the novel T7−pA−Tn5 transposition and T7 IVT, FACT-seq can feasibly map both active and silent histone modifications with low cell numbers, and the sequencing library complexity from FACT-seq is higher than that from CUT&Tag. Our results are also in agreement with the recently developed TIP-seq (43), showing that linear amplification from IVT with T7 promoter modified pA–Tn5 helps to obtain more unique reads in sequencing libraries.

Proof-of-concept of FACT-seq in FFPE samples

Harsh chemical treatments during the preparation of FFPE samples could mask protein epitopes in the tissue, and it is essential to retrieve these protein epitopes for binding of the antibody to their targets (81). We optimized a condition for epitope retrieval in the isolated FFPE nuclei (see Materials and Methods), where the H3K27ac signal intensity from the epitope retrieval condition was stronger than that without epitope retrieval (Figure 1E). We cut the mouse kidney into two parts, where one part was frozen down and the other part was prepared as an FFPE block (Supplementary Figure S7, see Materials and Methods). We performed H3K27ac CUT&Tag with 100 000 nuclei isolated from frozen mouse kidneys, and H3K27ac FACT-Seq with 100 000 nuclei (under the conditions of ± epitope retrieval) isolated from 20-μm thick FFPE mouse kidney tissue sections (see Materials and Methods). Both H3K27ac CUT&Tag and FACT-seq showed good reproducibility (Supplementary Figure S8A–C). The patterns of the H3K27ac signals were similar among H3K27ac CUT&Tag, H3K27ac FACT-seq (+ epitope retrieval), H3K27ac FACT-seq (- epitope retrieval) and ENCODE mouse kidney H3K27ac ChIP-seq (Figure 4A). The genome-wide correlation of H3K27ac CUT&Tag, H3K27ac FACT-seq (+ epitope retrieval), H3K27ac FACT-seq (- epitope retrieval) and ENCODE mouse kidney H3K27ac ChIP-seq clearly shows that H3K27ac FACT-Seq (+ epitope retrieval) had a higher correlation with ENCODE mouse kidney H3K27ac ChIP-seq (Figure 4A–C, $R = 0.81$) than with H3K27ac FACT-seq (- epitope retrieval) (Figure 4C, $R = 0.67$), and a similar tendency was observed for the pair comparison of H3K27ac FACT-seq (+ epitope retrieval) vs. H3K27ac CUT&Tag (Figure 4C, $R = 0.82$) and H3K27ac FACT-seq (− epitope retrieval) vs. H3K27ac CUT&Tag (Supplementary Figure S8D, $R = 0.72$). The quantification of the FRiP from different types of sequencing libraries clearly demonstrated that the FRiP of H3K27ac FACT-seq under epitope retrieval conditions (45.18% ± 0.88%)) was much higher FRiP than that of without epitope retrieval (17.56% ± 3.38%), but was in a similar range as that of H3K27ac CUT&Tag (51.68% ± 2.09%) (Figure 4D). As we expected, the library complexity from FACT-seq was also higher than that from CUT&Tag (Figure 4E). We further compared the genomic features from H3K27ac FACT-seq (+ epitope retrieval) of the FFPE sample, H3K27ac CUT&Tag of the frozen sample and ENCODE mouse kidney H3K27ac ChIP-seq and found sequencing read enrichments at the TSSs (Figure 4F, Supplementary Figure S8E), sequencing read enrichments at the ENCODE mouse kidney H3K27ac ChIP-seq peak site (Figure 4G) and genomic annotation of peaks with similar patterns from the three different libraries (Supplementary Figure S8F). At the peak level, we also found a large proportion of overlap in the paired comparison of FACT-seq (+ epitope retrieval) vs. H3K27ac CUT&Tag, and FACT-seq (+ epitope retrieval) versus ENCODE mouse kidney H3K27ac ChIP-seq (Supplementary Figure S9A–C). Further genomic annotation and TF enrichment analysis of the overlapping peaks and unique peaks showed that the exclusive peaks from different types of libraries were distributed in a similar pattern relative to TSSs (Supplementary Figure S9A–C). However, the different TF enrichments were found at those exclusive peaks (Supplementary Figure S9A–C). All of observations suggested that FACT-seq could decode H3K27ac profiles in FFPE samples with high sensitivity, but the epitope retrieval is essential there.

To further confirm that our FACT-seq worked effectively in FFPE samples, we had performed H3K27ac FACT-seq in mouse FFPE livers with 100 000 FFPE nuclei under the same epitope retrieval conditions in mouse FFPE kidneys and compared it with H3K27ac FiTAc-seq data from mouse FFPE liver sections and frozen mouse liver ENCODE H3K27ac ChIP-seq data (Supplementary Figure S10). Our comprehensive comparisons, including genome-wide correlations in peak enrichment of sequencing reads (Supplementary Figure S10A–C), FRiP (Supplementary Figure S10D), library complexity (Supplementary

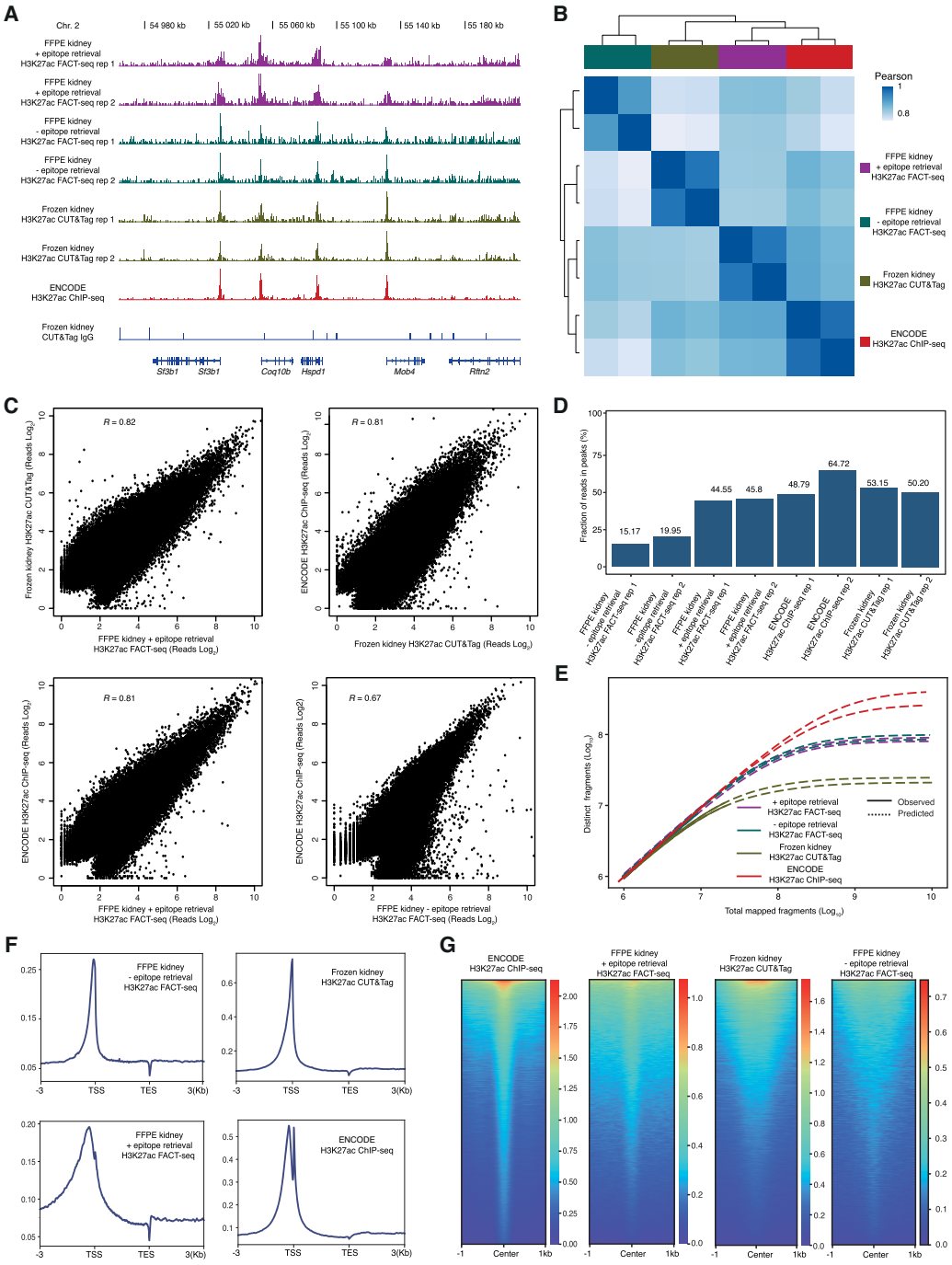

Figure 4. FACT-seq maps of genome-wide H3K27ac modifications in FFPE mouse kidney nuclei with high sensitivity. (**A**) Genome browser tracks for H3K27ac sequencing libraries in mouse kidney nuclei: Results from H3K27ac FACT-seq of mouse FFPE kidney nuclei (- epitope retrieval), H3K27ac FACT-seq of mouse FFPE kidney nuclei (+ epitope retrieval), H3K27ac CUT&Tag of frozen mouse kidney nuclei, ENCODE H3K27ac ChIP-seq of frozen mouse kidney nuclei, and CUT&Tag IgG control of mouse kidney nuclei. The gene names are shown at the bottom. Chr = Chromosome. (**B**) Pearson correlation heat map showing sample-by-sample unsupervised clustering on all peaks identified across different types of libraries. (**C**) Genome-wide correlation of H3K27ac sequencing libraries in mouse kidney nuclei under different conditions: H3K27ac FACT-seq of mouse FFPE kidney nuclei (- epitope retrieval), H3K27ac FACT-seq of mouse FFPE kidney nuclei (+ epitope retrieval), H3K27ac CUT&Tag of frozen mouse kidney nuclei, and ENCODE H3K27ac ChIP-seq of frozen mouse kidney nuclei. Each dot represents an individual peak. R = Pearson correlation. (**D**) Quantification of the fraction of reads in peaks (FRiP) for different sequencing libraries. (**E**) Estimation of sequencing library complexity for each sequencing libraries. (**F**) Metaplots showing the distribution of sequencing reads across transcription start sites (TSS) and transcription end sites (TES) from different sequencing libraries: H3K27ac FACT-seq of mouse FFPE kidney nuclei (- epitope retrieval), H3K27ac FACT-seq of mouse FFPE kidney nuclei (+ epitope retrieval), H3K27ac CUT&Tag of frozen mouse kidney nuclei, and ENCODE H3K27ac ChIP-seq of frozen mouse kidney nuclei. (**G**) Heatmaps showing the distribution of sequencing reads across peaks of ENCODE H3K27ac ChIP-seq of frozen mouse kidney nuclei from different sequencing libraries: H3K27ac FACT-seq of mouse FFPE kidney nuclei (- epitope retrieval), H3K27ac FACT-seq of mouse FFPE kidney nuclei (+ epitope retrieval), H3K27ac CUT&Tag of frozen mouse kidney nuclei, and ENCODE H3K27ac ChIP-seq of frozen mouse kidney nuclei.

Figure S10E), genomic annotation (Supplementary Figure S10F and G) and proportion of overlapping peaks (Supplementary Figure S10H) among the sequencing libraries of FACT-seq, FiTAc-seq and ENCODE ChIP-seq further demonstrated that FACT-seq could profile H3K27ac in FFPE samples with high sensitivity.

Next, we checked the sensitivity of H3K27ac FACT-Seq by using different numbers (ranging from 1000 to 10 000) of purified mouse FFPE kidney nuclei after the step of epitope retrieval (see Materials and Methods, Figure 5A–F, Supplementary Figure S11). Our multiangle comparison proved that FACT-seq resulted in similar H3K27ac profiles using as few as 1000 purified mouse FFPE kidney nuclei (Figure 5A−F, Supplementary Figure S11). However, we learned that a proportion of nuclei were lost during the nuclei isolation step and epitope retrieval. Then, we aimed to assess the lowest amounts of nuclei on the FFPE tissue section needed for FACT-seq. To do so, we first calculated the nucleus recovery rate during the nuclei isolation from FFPE tissue sections and epitope retrieval steps by using both one whole mouse FFPE kidney section (Figure 5G) and a small piece of FFPE tissue (Figure 5H). We concluded the nucleus recovery rate was ∼50% at the step of nuclei isolation (53.49% for the big FFPE tissue section, and 42.04% for the small FFPE tissue section) (Supplementary Figure S11C). The average of nucleus recovery rate after epitope retrieval was 55.95% (ranging from 34% to 98%) (Supplementary Figure S11D). Thus, the nucleus recovery rate from tissue sectioning to epitope retrieval was ∼25–30%. Taken together, we assumed that ∼4000 nuclei on the FFPE tissue section were needed for FACT-seq. To confirm our hypothesis, we counted 4000 nuclei under the microscope, dissected such small piece of FFPE tissue (Figure 5H), isolated nuclei and performed H3K27ac FACT-seq with all nuclei from this small piece of FFPE tissue. The comprehensive comparison (Figure 5A–F, Supplementary Figure S11E–G) of the H3K27ac FACT-seq data generated from different amounts of FFPE nuclei clearly demonstrated that we could accurately profile H3K27ac modifications in FFPE samples from a very small FFPE tissue section containing ∼4000 nuclei. We also performed H3K27ac FACT-seq on mouse FFPE kidney tissue sections of different thickness (5-, 7-, 10-μm) with an area of ∼3 × 5 mm^2. Our results showed that FACT-seq could be applied to a 10-μm-thick tissue section (Supplementary Figure S12) but was not optimal for 5- and 7-μm-thick FFPE tissue sections. We found that the majority of sequencing reads (80.57–94.39%) from 5- and 7-μm thick FFPE tissue sections could not be mapped to the reference genome, where a large proportion of unmapped reads (>50%) were from primer dimers (Supplementary Table S1). It has been reported that the diameter of the mammalian nucleus is 6–10 μm (82), and we suspected that there are potentially plenty of nonintact nuclei in 5- to 7- μm-thick FFPE tissue sections. The chromatin structure in those nonintact nuclei could be destroyed during the nuclei isolation, epitope retrieval and long FACT-seq procedures, resulting in low-quality histone modification profiles. Thus, we proposed that FFPE tissue sections with thicknesses greater than the diameter of nucleus should be used in FACT-seq. This observation is also in agreement with the finding in our recently established FFPE-ATAC technology (45), where 5- and 7- μm-thick FFPE tissue sections were not optimal for chromatin accessibility profiling in FFPE samples.

Furthermore, we performed FACT-seq by targeting the repressive histone modification H3K27me3 in FFPE samples (Figure 6). We used the same epitope retrieval conditions for active histone modification H3K27ac to perform epitope retrieval of the repressive histone modification H3K27me3 and found that the signal intensity of H3K27me3 also significantly increased with immunostaining (Figure 1F). However, the genome-wide correlation of mouse FFPE kidney H3K27me3 FACT-seq under this epitope retrieval condition and frozen mouse kidney H3K27me3 CUT&Tag was very low (Supplementary Figure S13A, $R = 0.65$). In addition, the sequencing signal enrichment of mouse FFPE kidney H3K27me3 FACT-seq with this epitope retrieval condition at TSSs was completely different from frozen mouse kidney H3K27me3 CUT&Tag and ENCODE frozen mouse kidney H3K27me3 ChIP-seq (Supplementary Figure S13B). Thus, we assumed that the epitope retrieval condition for active histone markers was not optimal for repressive histone markers. One possibility is that the solubility of heterochromatin and euchromatin is quite different in FFPE samples (83), and harsher chemical treatment conditions for epitope retrieval are needed for repressive histone modifications. Thus, we optimized the epitope retrieval conditions for H3K27me3 with purified FFPE nuclei by increasing the detergent concentration in the epitope retrieval solution and using higher temperatures during epitope retrieval (see Materials and Methods). We first chose conditions of epitope retrieval to check the nuclei integrity by using microscopy imaging of nuclei staining (see Materials and Methods), and then prepared mouse FFPE kidney H3K27me3 FACT-seq libraries in conditions where the nuclei were still intact. The sequencing libraries of mouse FFPE kidney H3K27me3 FACT-seq from different epitope retrieval conditions were compared to ENCODE frozen mouse kidney H3K27me3 ChIP-seq and frozen mouse kidney H3K27me3 CUT&Tag (Supplementary Figure S13A, Figure 6). With a series of optimizations, we found that the Pearson correlation between mouse FFPE kidney H3K27me3 FACT-seq and frozen mouse kidney H3K27me3 CUT&Tag increased to 0.76 (Figure 6A–C), while the correlation between mouse FFPE kidney H3K27me3 FACT-seq and ENCODE frozen mouse kidney H3K27me3 ChIP-seq reached 0.82 (Figure 6C) under the epitope retrieval condition of 0.1% sodium dodecyl sulfate (SDS) plus 0.1% sodium deoxycholate with a 1-hour incubation at 65°C. With such a harsher condition of epitope retrieval, mouse FFPE kidney H3K27me3 FACT-seq sequencing libraries have good reproducibility (Supplementary Figure S13D), and the signal enrichment at TSSs (Supplementary Figure S13B) and the genomic features of the mouse FFPE kidney H3K27me3 FACT-seq sequencing library (Supplementary Figure S13E, F) had high similarity to the ENCODE frozen mouse kidney H3K27me3 ChIP-seq and the frozen mouse kidney H3K27me3 CUT&Tag sequencing libraries. The FRiP from H3K27me3 FACT-seq under harsh epitope conditions was 18.29 ± 1.84% (Figure 6D), which was similar to that of ENCODE H3K27me3 ChIP-seq (15.32% ± 0.70%), but 2-3-fold lower than that

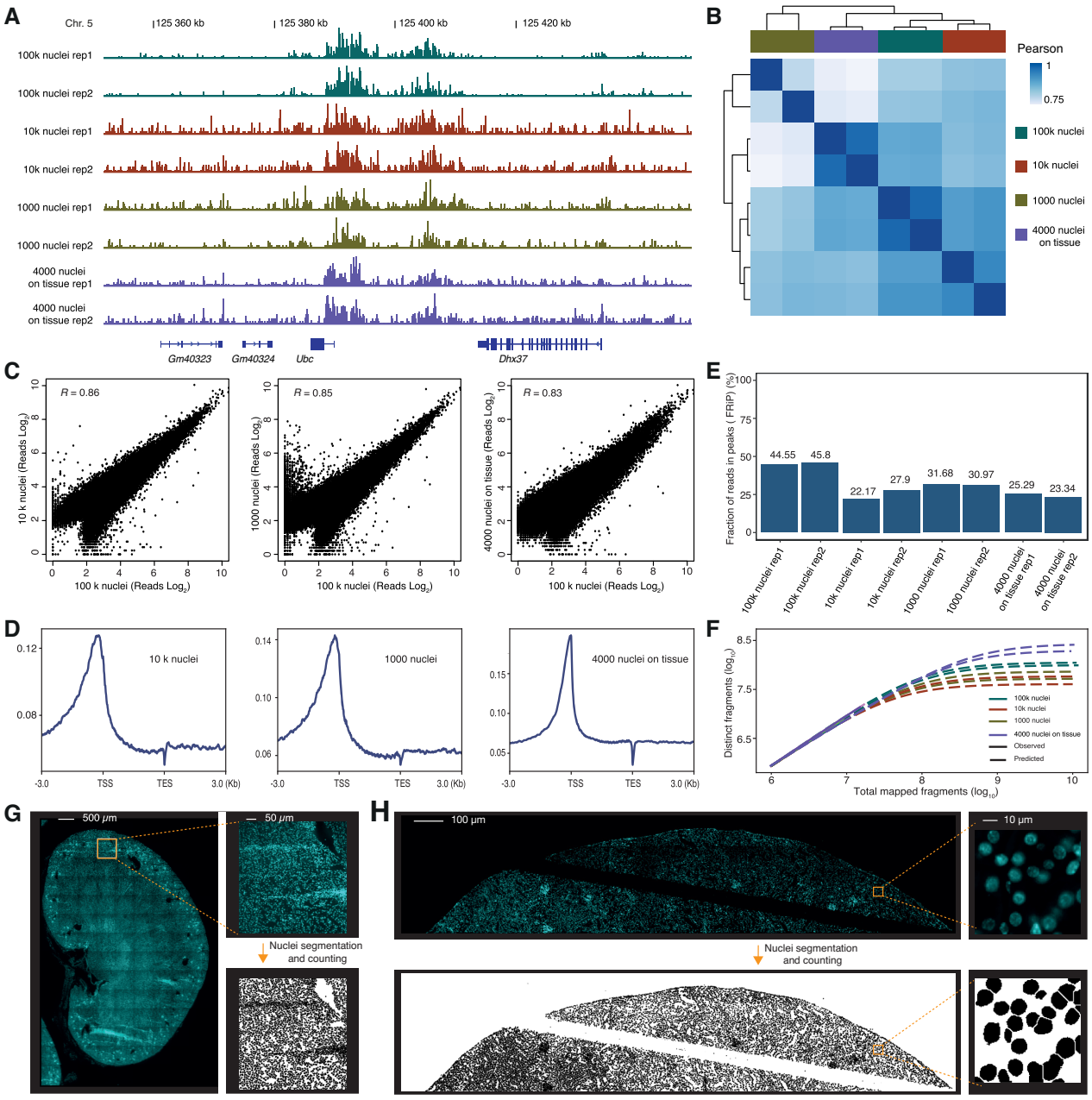

Figure 5. Sensitivity assay of H3K27ac FACT-seq. (**A**) Genome browser tracks for H3K27ac FACT-seq from mouse FFPE kidneys with different nuclei numbers collected after epitope retrieval (1000 nuclei, 10k nuclei and 100k nuclei) and 4000 nuclei in tissue sections. The gene names are shown at the bottom. Chr. = Chromosome. (**B**) Pearson correlation heatmap showing unsupervised clustering for all peaks identified across all sequencing libraries generated with different numbers of nuclei from tubes or tissue sections. Each sample represents an individual technical replicate. (**C**) Genome-wide correlation of H3K27ac FACT-seq data generated using 1000 isolated FFPE nuclei in tubes, 10k isolated FFPE nuclei, and 4000 nuclei in FFPE tissue sections with 100k isolated FFPE nuclei. Each dot represents an individual peak. Peaks analyzed were derived from a union peak set using data from all different sequencing libraries. R = Pearson correlation. (**D**) Sequencing reads enrichment across transcription start sites (TSSs) from mouse FFPE kidney H3K27ac FACT-seq with different numbers of nuclei. (**E**) Quantification of the fraction of reads in peaks (FRiP) for different sequencing libraries. (**F**) Estimation of sequencing library complexity for each sequencing libraries. (**G**) Image showing the nuclei quantification from the entire mouse FFPE kidney tissue section. (**H**) Image showing the nuclei quantification of 4000 nuclei from a small mouse FFPE kidney section for sequencing library preparation.

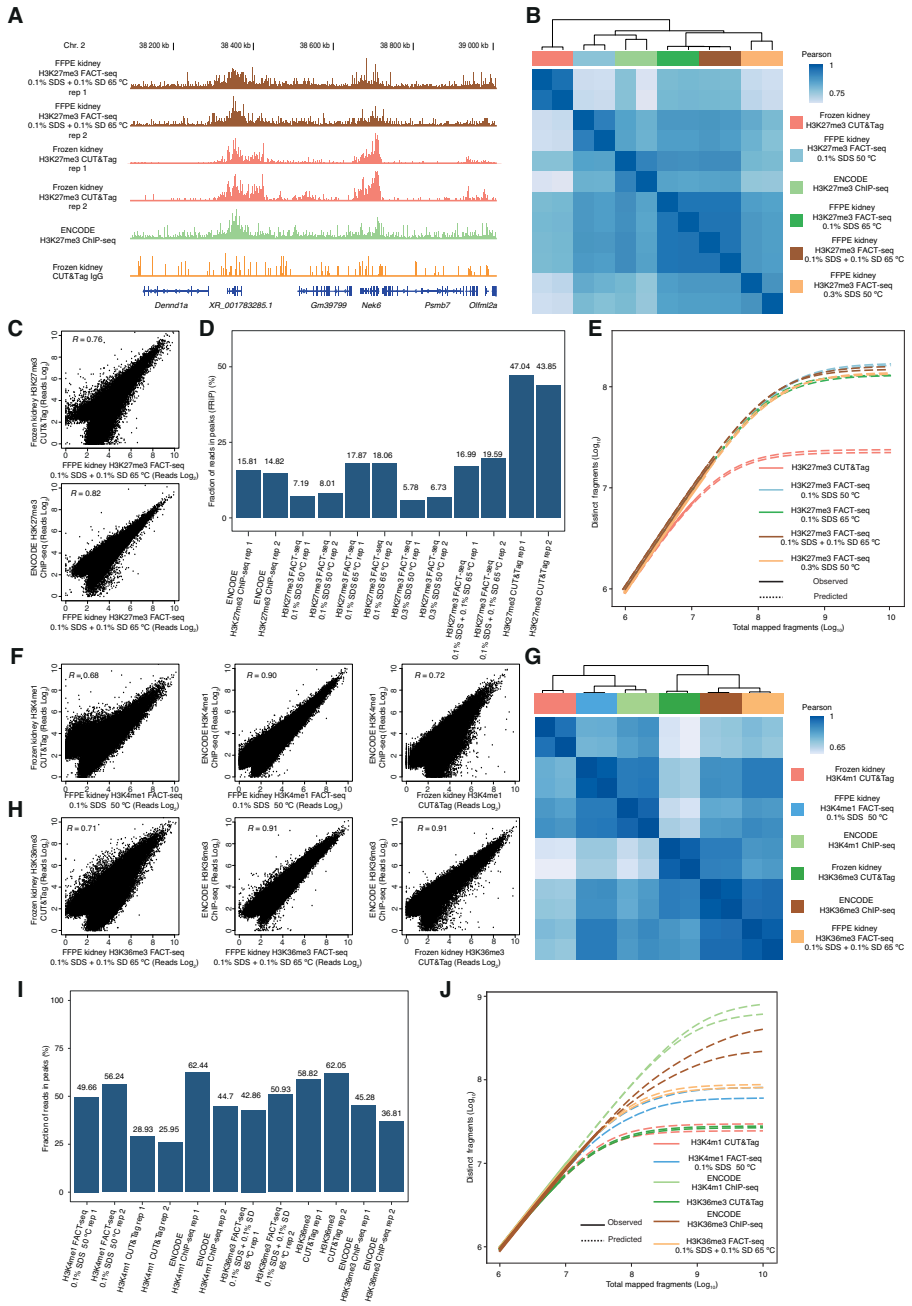

Figure 6. FACT-seq maps of genome-wide H3K27me3, H3K36me3 and H3K4me1 modifications in FFPE mouse kidney nuclei with high sensitivity. (**A**) Genome browser tracks for H3K27me3 sequencing libraries in mouse kidney nuclei: results from H3K27me3 FACT-seq of mouse FFPE kidney nuclei, H3K27me3 CUT&Tag of frozen mouse kidney nuclei, ENCODE H3K27me3 ChIP-seq of frozen mouse kidney nuclei, and CUT&Tag IgG control of frozen mouse kidney nuclei. The gene names are shown at the bottom. Chr. = Chromosome. SDS = Sodium dodecyl sulfate, SD = Sodium Deoxycholate. (**B**) Pearson correlation heat map showing sample-by-sample unsupervised clustering of all peaks identified across different types of libraries for H3K27me3. (**C**) Genome-wide correlation of H3K27me3 sequencing libraries in mouse kidney nuclei under different conditions: H3K27me3 FACT-seq of mouse FFPE kidney nuclei, H3K27me3 CUT&Tag of frozen mouse kidney nuclei, and ENCODE H3K27me3 ChIP-seq of frozen mouse kidney nuclei. Each dot represents one peak. R = Pearson correlation. (**D**) Quantification of the fraction of reads in peaks (FRiP) for different sequencing libraries of H3K27me3. (**E**) Comparison of library complexity for H3K27me3 CUT&Tag from frozen mouse kidneys and H3K27me3 FACT-seq from FFPE mouse kidneys with different epitope retrieval conditions. (**F**) Correlation of H3K4me1 sequencing libraries generated from different methods: H3K4me1 FACT-seq of mouse FFPE kidney nuclei, H3K4me1 CUT&Tag of frozen mouse kidney nuclei, and ENCODE H3K4me1 ChIP-seq of frozen mouse kidney nuclei. Each dot represents an individual peak. Peaks analyzed were derived from a union peak set using data from all different sequencing libraries. R = Pearson correlation. (**G**) Pearson correlation heat map showing sample-by-sample unsupervised clustering of all peaks identified across different types of libraries for H3K36me3 and H3K4me1. (**H**) Correlation of H3K36me3 sequencing libraries generated with different methods: H3K36me3 FACT-seq of mouse FFPE kidney nuclei, H3K36me3 CUT&Tag of frozen mouse kidney nuclei, and ENCODE H3K36me3 ChIP-seq of frozen mouse kidney nuclei. Each dot represents an individual peak. Peaks analyzed were derived from a union peak set using data from all different sequencing libraries. R = Pearson correlation. (**I**) Quantification of the fraction of reads in peaks (FRiP) for different sequencing libraries of H3K36me3 and H3K4me1. (**J**) Comparison of library complexity for different sequencing libraries of H3K36me3 and H3K4me1.

of H3K27me3 CUT&Tag in frozen samples (45.45% ± 2.26%). The library complexity of FACT-seq for different epitope retrieval conditions was similar (Figure 6E), and much higher than that of the CUT&Tag sequencing libraires. There was also a large proportion of overlapping peaks between mouse FFPE kidney H3K27me3 FACT-seq and ENCODE frozen mouse kidney H3K27me3 ChIP-seq and frozen mouse kidney H3K27me3 CUT&Tag (Supplementary Figure S14A–C). Thus, we concluded that FACT-seq could be used to profile the repressive histone marker H3K27me3 in FFPE samples but harsh epitope retrieval conditions are needed.

Finally, we tested whether we could use the same epitope retrieval conditions for different histone modifications. We used harsh epitope retrieval condition (optimal epitope retrieval condition for H3K27me3) for H3K27ac FACT-seq in mouse FFPE kidneys, and we found that signal enrichment around TSSs from H3K27ac FACT-seq with harsh epitope retrieval conditions completely disappeared (Supplementary Figure S15A). We also profiled the other two histone modifications in mouse FFPE kidney samples, H3K36me3 and H3K4me1, with both mild (optimal epitope retrieval condition for H3K27ac) and harsh (optimal epitope retrieval condition for H3K27me3) epitope retrieval conditions. Our comprehensive characterization of H3K36me3 FACT-seq and H3K4me1 FACT-seq (Figure 6F–J, Supplementary Figure S15B–J) by comparison with the sequencing libraries of CUT&Tag and ENCODE ChIP-seq, clearly indicated that harsh epitope retrieval conditions are needed for H3K36me3, and mild epitope retrieval conditions are optimal for H3K4me1. Thus, we proposed that epitope retrieval conditions for FACT-seq in FFPE samples may be histone modification dependent but could be optimized based on these mild and harsh conditions and judged with sequencing results.

Taken together, accurate mapping of different histone modifications, H3K27ac, H3K27me3, H3K36me3 and H3K4me1, with low numbers of isolated FFPE nuclei from FFPE tissue sections demonstrates that FACT-seq can serve to profile histone modifications from FFPE samples with high sensitivity.

FACT-seq reveals the disease-specific super enhancers in archived FFPE human colorectal and human glioblastoma cancer tissue

Finally, we applied the FACT-seq method to archived FFPE human colorectal (CRC) and human glioblastoma multiforme (GBM) cancer tissue, where we performed H3K27ac FACT-seq on one CRC and four GBM FFPE samples. The sequencing libraries of H3K27ac FACT-seq from both the CRC and GBM samples showed good technical reproducibility in terms of the genome wide sequencing read correlation (*R* values ranging from 0.89 to 0.94, Supplementary Figure S16A) and in the proportion of overlapping peaks (Supplementary Figure S16B). In addition, sequencing reads from all H3K27ac FACT-seq libraries were strongly enriched at the TSS (Supplementary Figure S16C). Furthermore, we downloaded published RNA-seq data of CRC (84) and GBM (85), and correlated gene expression levels with the H3K27ac signal intensity for CRC and GBM (Figure 7A, Supplementary Figure S16D, E). In brief, genes were first categorized into four ranks (from high gene expression to low gene expression: top 25%, 25–50%, 50–75% and 75–100%) (86) using the RNA-seq data of both human CRC (84) and human GBM (85), and the signal intensity of H3K27ac at the TSSs (±3 kb) was compared among these four categories (Figure 7A, Supplementary Figure S16D, E). The results clearly showed that there was higher H3K27ac signal enrichment at high ranks of gene expression in both human CRC and human GBM, which further suggests that our H3K27ac FACT-seq data from human CRC and human GBM are reliable. In addition, the genomic annotation of H3K27ac peaks (Supplementary Figure S16F and G), FRiP (from 15.45% to 31.88%) (Figure 7B) and library complexity (Figure 7C) from those clinical archived FFPE samples were all in reasonable ranges. When we clustered all H3K27ac FACT-seq peaks from the one CRC and four GBM samples with the non-negative matrix factorization (NMF) method, we found that two clusters were the best to characterize all H3K27ac FACT-seq peaks (Figure 7D, Supplementary Figure S16H), where samples from CRC were in one cluster, and samples from GBM were in the other (Figure 7D). Because H3K27ac has been used for super enhancer identification (61), we used H3K27ac FACT-seq to identify super enhancers using the published algorithm (61,62). With this approach, we identified 1206 super enhancers in the CRC samples (Figure 7E, Supplementary Table S2), and 492 super enhancers in the GBM samples (Figure 7F, Supplementary Table S3). We also compared the list of super enhancers in our study with the reported list of super enhancers from other GBM clinical samples (87), and we found that 310 of 482 super enhancers identified in our samples overlapped with the reported super enhancers in GBM clinical samples (Supplementary Figure S16I, Supplementary Table S4), which further proved the accuracy of FACT-seq in profiling H3K27ac histone modifications from FFPE samples. Among the list of super enhance genes, Tumor suppressor genes and oncogenes (71) were identified in GBM and CRC (Figure 7E, F, Supplementary Table S2, S3). The tumor suppressor gene, *ERBB* receptor feedback inhibitor (*ERRFI1*), ranked in the top 10 for CRC (Figure 7E and G), but was not found in the super enhancer list of GBM. *ERRFI1* is reported to be a negative regulator of epidermal growth factor receptor (*EGFR*) (88), and has been investigated with regard to drug resistance in CRC (89). Interestingly, *EGFR* was ranked in the top 4 for GBM (Figure 7F and H), and was also listed as a super enhancer of CRC (Figure 7E), but with a rank of 120. The high rank of *EGFR* in the super enhancer list of GBM properly reflects the important function of *EGFR* in GBM progression (90,91). In addition, we also found that another CRC-relevant oncogene, MET (92), is listed in the top 10 CRC super enhancer list (Figure 7E, Supplementary Figure S16J, Supplementary Table S2). Furthermore, we performed the ontological enrichment (65) of super enhancer associated genes in GBM and CRC (Figure 7I and J). Interestingly, tissue migration epithelial, regulation intracellular stimulus and other ontological pathways (Figure 7I) were identified in the list of CRC super enhancers, but ontological pathways from nervous system development (Figure 7J) are enriched in the super enhancer list of GBM. Thus, we

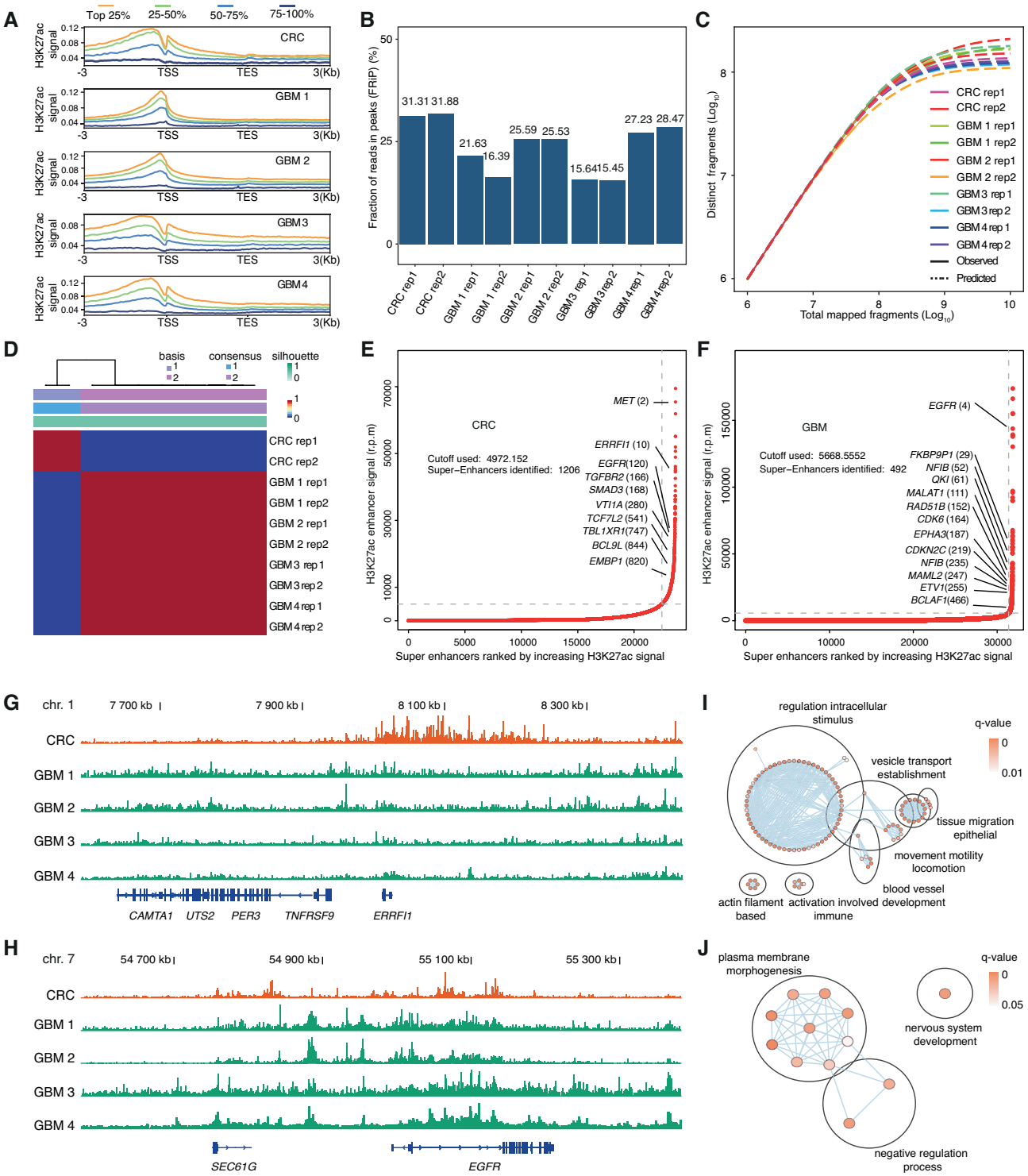

Figure 7. Tumour-specific super enhancers are identified by H3K27ac FACT-seq from clinical archived FFPE tumour samples. (**A**) Metaplots showing H3K27ac FACT-seq signal enrichment at gene regions (±3 kb TSS) from different ranking categories of gene expression in human colorectal cancer (CRC) and human glioblastoma multiforme (GBM) tissue samples. (**B**) Fraction of reads in peaks (FRiP) of H3K27ac FACT-seq for archived FFPE human CRC and 4 cases of archived FFPE human GBM tissue samples. (**C**), Comparison of library complexities for H3K27ac FACT-seq for human CRC and human GBM. (**D**) Nonnegative matrix factorization (NMF) clustering of H3K27ac FACT-seq libraries for human CRC and human GBM. (**E, F**) Super enhancer ranking plots of human CRC (**E**) and human GBM cancer tissue (**F**) identified from H3K27ac FACT-seq. Name and ranks of the oncogenes and tumor suppressor genes are indicated in the plots. (**G**) Genome browser track of a representative super-enhancer (*ERRFI1*) from H3K27ac FACT-seq of human CRC (**H**), Genome browser track of a representative super-enhancer (*EGFR*) from H3K27ac FACT-seq of human GBM. (**I**), Gene pathways identified from ontological enrichment of super enhancer associated genes in human CRC. (**J**), Gene pathways identified from ontological enrichment of super enhancer associated genes in human GBM.

concluded that FACT-seq serves to identify disease-specific super enhancers using low numbers of nuclei prepared from a single clinically archived FFPE tissue section.

DISCUSSION

FFPE tissue samples represent a large source of material for epigenetic analysis in both basic research and clinical translational studies (93). The successful application of ChIP-seq in FFPE tissues, PAT-ChIP (14,15), (FiT-seq) (7), FiTAc-seq) (16) and other similar technologies (17), makes it possible to map histone modifications in clinically archived FFPE tissue; however, a large amount of input materials from clinical samples are required for these technologies (7,18). Thus far, the material has not been widely used in epigenetic studies because of the lack of sufficiently sensitive detection technologies (7,16). Combining a novel fusion protein of hyperactive Tn5 transposase and protein A (T7−pA−Tn5) transposition and T7 *in vitro* transcription, we developed the first highly sensitive technology for profiling histone modifications in clinically archived FFPE samples. We demonstrated that FACT-seq is a robust tool to decode genome-wide histone modifications using a tiny FFPE tissue containing ∼4000 nuclei. FACT-seq is a sonication-free method, and could thus minimize potential sequence bias introduced by sonication procedures. In CUT&Tag and similar technologies, one insertion event or unpaired pA–Tn5 adaptor insertions from pA–Tn5 transposition could not be amplified with PCR (75). Since the IVT depends on only one insertion of the T7 promoter, one insertion event and unpaired adaptor insertions at the targeting sites from T7−pA−Tn5 could be detected with FACT-seq during IVT. Thus, the library complexity of FACT-seq is higher than that of CUT&Tag or similar technologies. Recently, a similar technology TIP-seq (43), was established by an independent group, who also proved that the linear amplification feature of T7 *in vitro* transcription on pA–Tn5 could increase the sequencing library complexity of targeting protein:DNA binding sites.

In our currently established FACT-seq protocol, 5 working days are needed from cutting the FFPE tissue section to sequencing library readiness, which is in the same time scale of other histone modification profiling technologies for FFPE samples (7,16), but is much longer than that of high sensitivity technologies working with non-FFPE samples (20–39,75). There is a great need to reduce the timing in the future by further optimizing the protocol. We also learned that epitope retrieval is essential in FACT-seq but histone marker dependent, thus, it will be very helpful to further optimize a generic epitope retrieval condition for all histone modifications. Furthermore, we noticed that some proportion of nuclei was lost during nuclei isolation and epitope retrieval, and a better nuclei isolation protocol for FACT-seq is needed in the future.

The high sensitivity of FACT-seq, consuming minimal amounts of clinical materials, makes it possible to combine epigenetic profiling with other clinical-pathological parameters for clinical diagnosis. The successful profiling of the disease specific super enhancer from the clinically archived CRC and GBM FFPE samples makes FACT-seq a powerful tool in preclinical studies and for precision medicine. In addition, FACT-seq has the potential to extend our current understanding of the cancer epigenome by combining it with other omics data from the same FFPE materials. FACT-seq can have broad applications both in basic research and clinical settings. Most of available high sensitivity technologies for profiling histone modifications involving non-FFPE biological materials are already at the single-cell level (20–39), thus, it would be of great interest to push the resolution of FACT-seq to the single cell level and to investigate epigenetic heterogeneity in clinical samples in the near future.

DATA AVAILABILITY

T7 promoter sequences were trimmed from each read in the fastq file with in-house script custom script (https://github.com/pengweixing/FACT).

The raw sequencing data was deposited to NCBI GEO with accession number GSE171758.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank members of the Chen labs for discussion. Part of this work was facilitated by the Protein Science Facility at Karolinska Institute, Stockholm.

FUNDING

Swedish Research Council [VR-2016-06794, VR-2017-02074 to X.C.]; Åke Wibergs stiftelse [M20-0007 to X.C.]; Beijer Foundation (to X.C.); Jeassons Foundation (to X.C.); Petrus och Augusta Hedlunds Stiftelse (to X.C.); Göran Gustafsson's prize for younger researchers (to X.C.); Vleugel Foundation (to X.C.); Linnéstiftelsen for medicinsk forskning (to X.C.); Uppsala University (to X.C.); Swedish Cancer Society (CAN 2018/772 to T.S.); Protein Science Facility at Karolinska Institute, Stockholm (in part). Funding for open access charge: Swedish research Council.

Conflict of interest statement. X.C., V.K.P. and L.Z. have filed patent applications related to the work described here. The title of the patent application is 'Method of preparing DNA from formalin-fixed-paraffin-embedded (FFPE) tissue samples'. The Swedish Provisional Application was filed on 28 June 2021, Patent Application No. 2150823-9 in Sweden. The authors declare no competing financial interests.

REFERENCES

1. Corces,M.R., Granja,J.M., Shams,S., Louie,B.H., Seoane,J.A., Zhou,W., Silva,T.C., Groeneveld,C., Wong,C.K., Cho,S.W. *et al.* (2018) The chromatin accessibility landscape of primary human cancers. *Science*, **362**, eaav1898.
2. Qu,K., Zaba,L.C., Satpathy,A.T., Giresi,P.G., Li,R., Jin,Y., Armstrong,R., Jin,C., Schmitt,N., Rahbar,Z. *et al.* (2017) Chromatin accessibility landscape of cutaneous T cell lymphoma and dynamic response to HDAC inhibitors. *Cancer Cell*, **32**, 27–41.
3. Fraga,M.F., Ballestar,E., Villar-Garea,A., Boix-Chornet,M., Espada,J., Schotta,G., Bonaldi,T., Haydon,C., Ropero,S., Petrie,K. *et al.* (2005) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat. Genet.*, **37**, 391–400.

4. Seligson,D.B., Horvath,S., Shi,T., Yu,H., Tze,S., Grunstein,M. and Kurdistani,S.K. (2005) Global histone modification patterns predict risk of prostate cancer recurrence. *Nature*, **435**, 1262–1266.
5. Akhtar-Zaidi,B., Cowper-Sal-lari,R., Corradin,O., Saiakhova,A., Bartels,C.F., Balasubramanian,D., Myeroff,L., Lutterbaugh,J., Jarrar,A., Kalady,M.F. *et al.* (2012) Epigenomic enhancer profiling defines a signature of colon cancer. *Science*, **336**, 736–739.
6. Landt,S.G., Marinov,G.K., Kundaje,A., Kheradpour,P., Pauli,F., Batzoglou,S., Bernstein,B.E., Bickel,P., Brown,J.B., Cayting,P. *et al.* (2012) ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res.*, **22**, 1813–1831.
7. Cejas,P., Li,L., O'Neill,N.K., Duarte,M., Rao,P., Bowden,M., Zhou,C.W., Mendiola,M., Burgos,E., Feliu,J. *et al.* (2016) Chromatin immunoprecipitation from fixed clinical tissues reveals tumor-specific enhancer profiles. *Nat. Med.*, **22**, 685–691.
8. Jin,W., Tang,Q., Wan,M., Cui,K., Zhang,Y., Ren,G., Ni,B., Sklar,J., Przytycka,T.M., Childs,R. *et al.* (2015) Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples. *Nature*, **528**, 142–146.
9. Fox,CH, J.F.,Whiting J and Roller,PP. (1985) Formaldehyde fixation. *J. Histochem. Cytochem.*, **33**, 845–853.
10. Haile,S., Corbett,R.D., Bilobram,S., Bye,M.H., Kirk,H., Pandoh,P., Trinh,E., MacLeod,T., McDonald,H., Bala,M. *et al.* (2019) Sources of erroneous sequences and artifact chimeric reads in next generation sequencing of genomic DNA from formalin-fixed paraffin-embedded samples. *Nucleic Acids Res.*, **47**, e12.
11. Wang,Y.K., Moorhead,M., Karlin-Neumann,G., Falkowski,M., Chen,C.N., Siddiqui,F., Davis,R.W., Willis,T.D. and Faham,M. (2005) Allele quantification using molecular inversion probes (MIP). *Nucleic Acids Res.*, **33**, e183.
12. Waldron,L., Simpson,P., Parmigiani,G. and Huttenhower,C. (2012) Report on emerging technologies for translational bioinformatics: a symposium on gene expression profiling for archival tissues. *BMC Cancer*, **12**, 124.
13. Glimelius,B., Melin,B., Enblad,G., Alafuzoff,I., Beskow,A., Ahlstrom,H., Bill-Axelson,A., Birgisson,H., Bjor,O., Edqvist,P.H. *et al.* (2018) U-CAN: a prospective longitudinal collection of biomaterials and clinical information from adult cancer patients in Sweden. *Acta Oncol*, **57**, 187–194.
14. Fanelli,M., Amatori,S., Barozzi,I. and Minucci,S. (2011) Chromatin immunoprecipitation and high-throughput sequencing from paraffin-embedded pathology tissue. *Nat. Protoc.*, **6**, 1905–1919.
15. Fanelli,M., Amatori,S., Barozzi,I., Soncini,M., Dal Zuffo,R., Bucci,G., Capra,M., Quarto,M., Dellino,G.I., Mercurio,C. *et al.* (2010) Pathology tissue-chromatin immunoprecipitation, coupled with high-throughput sequencing, allows the epigenetic profiling of patient samples. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 21535–21540.
16. Font-Tello,A., Kesten,N., Xie,Y., Taing,L., Vareslija,D., Young,L.S., Hamid,A.A., Van Allen,E.M., Sweeney,C.J., Gjini,E. *et al.* (2020) FiTAc-seq: fixed-tissue ChIP-seq for H3K27ac profiling and super-enhancer analysis of FFPE tissues. *Nat. Protoc.*, **15**, 2503–2518.
17. Zhong,J., Ye,Z., Clark,C.R., Lenz,S.W., Nguyen,J.H., Yan,H., Robertson,K.D., Farrugia,G., Zhang,Z., Ordog,T. *et al.* (2019) Enhanced and controlled chromatin extraction from FFPE tissues and the application to ChIP-seq. *BMC Genomics*, **20**, 249.
18. Fanelli,M., Amatori,S., Barozzi,I., Soncini,M., Dal Zuffo,R., Bucci,G., Capra,M., Quarto,M., Dellino,G.I., Mercurio,C. *et al.* (2010) Pathology tissue-chromatin immunoprecipitation, coupled with high-throughput sequencing, allows the epigenetic profiling of patient samples. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 21535–21540.
19. Teytelman,L., Ozaydin,B., Zill,O., Lefrancois,P., Snyder,M., Rine,J. and Eisen,M.B. (2009) Impact of chromatin structures on DNA processing for genomic analyses. *PLoS One*, **4**, e6700.
20. Kaya-Okur,H.S., Wu,S.J., Codomo,C.A., Pledger,E.S., Bryson,T.D., Henikoff,J.G., Ahmad,K. and Henikoff,S. (2019) CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.*, **10**, 1930.
21. Carter,B., Ku,W.L., Kang,J.Y., Hu,G.Q., Perrie,J., Tang,Q.S. and Zhao,K.J. (2019) Mapping histone modifications in low cell number and single cells using antibody-guided chromatin tagmentation (ACT-seq). *Nat. Commun.*, **10**, 3747.
22. Wang,Q.H., Xiong,H.Q., Ai,S.S., Yu,X.H., Liu,Y.X., Zhang,J.J. and He,A.B. (2019) CoBATCH for high-throughput single-cell epigenomic profiling. *Mol. Cell*, **76**, 206.
23. Harada,A., Maehara,K., Handa,T., Arimura,Y., Nogami,J., Hayashi-Takanaka,Y., Shirahige,K., Kurumizaka,H., Kimura,H. and Ohkawa,Y. (2019) A chromatin integration labelling method enables epigenomic profiling with lower input. *Nat. Cell Biol.*, **21**, 287–296.
24. Handa,T., Harada,A., Maehara,K., Sato,S., Nakao,M., Goto,N., Kurumizaka,H., Ohkawa,Y. and Kimura,H. (2020) Chromatin integration labeling for mapping DNA-binding proteins and modifications with low input. *Nat. Protoc.*, **15**, 3334–3360.
25. Skene,P.J., Henikoff,J.G. and Henikoff,S. (2018) Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.*, **13**, 1006–1019.
26. Skene,P.J. and Henikoff,S. (2017) An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife*, **6**, e21856.
27. Grosselin,K., Durand,A., Marsolier,J., Poitou,A., Marangoni,E., Nemati,F., Dahmani,A., Lameiras,S., Reyal,F., Frenoy,O. *et al.* (2019) High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. *Nat. Genet.*, **51**, 1060–1066.
28. Ku,W.L., Nakamura,K., Gao,W., Cui,K., Hu,G., Tang,Q., Ni,B. and Zhao,K. (2019) Single-cell chromatin immunocleavage sequencing (scChIC-seq) to profile histone modification. *Nat. Methods*, **16**, 323–325.
29. Patty,B.J. and Hainer,S.J. (2021) Transcription factor chromatin profiling genome-wide using uliCUT&RUN in single cells and individual blastocysts. *Nat. Protoc.*, **16**, 2633–2666.
30. Brind'Amour,J., Liu,S., Hudson,M., Chen,C., Karimi,M.M. and Lorincz,M.C. (2015) An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. *Nat. Commun.*, **6**, 6033.
31. Cao,Z.N., Chen,C.Y., He,B., Tan,K. and Lu,C. (2015) A microfluidic device for epigenomic profiling using 100 cells. *Nat. Methods*, **12**, 959–962.
32. van Galen,P., Viny,A.D., Ram,O., Ryan,R.J., Cotton,M.J., Donohue,L., Sievers,C., Drier,Y., Liau,B.B., Gillespie,S.M. *et al.* (2016) A multiplexed system for quantitative comparisons of chromatin landscapes. *Mol. Cell*, **61**, 170–180.
33. Rotem,A., Ram,O., Shoresh,N., Sperling,R.A., Goren,A., Weitz,D.A. and Bernstein,B.E. (2015) Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat. Biotechnol.*, **33**, 1165–1172.
34. Hainer,S.J., Boskovic,A., McCannell,K.N., Rando,O.J. and Fazzio,T.G. (2019) Profiling of pluripotency factors in single cells and early embryos. *Cell*, **177**, 1319–1329.
35. Wu,S.J., Furlan,S.N., Mihalas,A.B., Kaya-Okur,H.S., Feroze,A.H., Emerson,S.N., Zheng,Y., Carson,K., Cimino,P.J., Keene,C.D. *et al.* (2021) Single-cell CUT&Tag analysis of chromatin modifications in differentiation and tumor progression. *Nat. Biotechnol.*, **39**, 819–824.
36. Bartosovic,M., Kabbe,M. and Castelo-Branco,G. (2021) Single-cell CUT&Tag profiles histone modifications and transcription factors in complex tissues. *Nat. Biotechnol.*, **39**, 825–835.
37. Liu,B., Xu,Q., Wang,Q., Feng,S., Lai,F., Wang,P., Zheng,F., Xiang,Y., Wu,J., Nie,J. *et al.* (2020) The landscape of RNA Pol II binding reveals a stepwise transition during ZGA. *Nature*, **587**, 139–144.
38. Zheng,X., Yue,S., Chen,H., Weber,B., Jia,J. and Zheng,Y. (2015) Low-cell-number epigenome profiling aids the study of lens aging and hematopoiesis. *Cell Rep.*, **13**, 1505–1518.
39. Ai,S.S., Xiong,H.Q., Li,C., Luo,Y.J., Shi,Q., Liu,Y.X., Yu,X.H., Li,C. and He,A.B. (2019) Profiling chromatin states using single-cell itChIP-seq. *Nat. Cell Biol.*, **21**, 1164.
40. Zhang,T., Tseng,C., Zhang,Y., Sirin,O., Corn,P.G., Li-Ning-Tapia,E.M., Troncoso,P., Davis,J., Pettaway,C., Ward,J. *et al.* (2016) CXCL1 mediates obesity-associated adipose stromal cell trafficking and function in the tumour microenvironment. *Nat. Commun.*, **7**, 11674.
41. Cao,J., Spielmann,M., Qiu,X., Huang,X., Ibrahim,D.M., Hill,A.J., Zhang,F., Mundlos,S., Christiansen,L., Steemers,F.J. *et al.* (2019) The single-cell transcriptional landscape of mammalian organogenesis. *Nature*, **566**, 496–502.

42. Kaya-Okur,H.S., Janssens,D.H., Henikoff,J.G., Ahmad,K. and Henikoff,S. (2020) Efficient low-cost chromatin profiling with CUT&Tag. *Nat. Protoc.*, **15**, 3264–3283.
43. Bartlett,D.A., Dileep,V., Henikoff,S. and Gilbert,D.M. (2021) High throughput genome-wide single cell protein:DNA binding site mapping by targeted insertion of promoters (TIP-seq). bioRxiv doi: https://doi.org/10.1101/2021.03.17.435909, 19 March 2021, preprint: not peer reviewed.
44. Chin,H.G., Sun,Z., Vishnu,U.S., Hao,P., Cejas,P., Spracklin,G., Esteve,P.O., Xu,S.Y., Long,H.W. and Pradhan,S. (2020) Universal NicE-seq for high-resolution accessible chromatin profiling for formaldehyde-fixed and FFPE tissues. *Clin Epigenet.*, **12**, 143.
45. Polavarapu,V.K., Xing,P., Zhang,H., Zhao,M., Mathot,L., Zhao,L., Rosen,G., Swartling,F.J., Sjoblom,T. and Chen,X. (2021) Profiling chromatin accessibility in formalin-fixed paraffin-embedded samples. *Genome Res.*, https://doi.org/10.1101/gr.275269.121.
46. Buenrostro,J.D., Wu,B., Litzenburger,U.M., Ruff,D., Gonzales,M.L., Snyder,M.P., Chang,H.Y. and Greenleaf,W.J. (2015) Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature*, **523**, 486–490.
47. Picelli,S., Bjorklund,A.K., Reinius,B., Sagasser,S., Winberg,G. and Sandberg,R. (2014) Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.*, **24**, 2033–2040.
48. Corces,M.R., Trevino,A.E., Hamilton,E.G., Greenside,P.G., Sinnott-Armstrong,N.A., Vesuna,S., Satpathy,A.T., Rubin,A.J., Montine,K.S., Wu,B. *et al.* (2017) An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods*, **14**, 959–962.
49. Schindelin,J., Arganda-Carreras,I., Frise,E., Kaynig,V., Longair,M., Pietzsch,T., Preibisch,S., Rueden,C., Saalfeld,S., Schmid,B. *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods*, **9**, 676–682.
50. Arganda-Carreras,I., Kaynig,V., Rueden,C., Eliceiri,K.W., Schindelin,J., Cardona,A. and Sebastian Seung,H. (2017) Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics*, **33**, 2424–2426.
51. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
52. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G., Durbin,R. and Processing,Genome Project Data, S. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
53. Ramirez,F., Dundar,F., Diehl,S., Gruning,B.A. and Manke,T. (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.*, **42**, W187–W191.
54. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
55. Thorvaldsdottir,H., Robinson,J.T. and Mesirov,J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.*, **14**, 178–192.
56. Yu,G., Wang,L.G. and He,Q.Y. (2015) ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*, **31**, 2382–2383.
57. Xu,S., Grullon,S., Ge,K. and Peng,W. (2014) Spatial clustering for identification of ChIP-enriched regions (SICER) to map regions of histone methylation patterns in embryonic stem cells. *Methods Mol. Biol.*, **1150**, 97–111.
58. Chen,X., Shen,Y., Draper,W., Buenrostro,J.D., Litzenburger,U., Cho,S.W., Satpathy,A.T., Carter,A.C., Ghosh,R.P., East-Seletsky,A. *et al.* (2016) ATAC-see reveals the accessible genome by transposase-mediated imaging and sequencing. *Nat. Methods*, **13**, 1013–1020.
59. Consortium,E.P. (2004) The ENCODE (ENCyclopedia Of DNA Elements) project. *Science*, **306**, 636–640.
60. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
61. Whyte,W.A., Orlando,D.A., Hnisz,D., Abraham,B.J., Lin,C.Y., Kagey,M.H., Rahl,P.B., Lee,T.I. and Young,R.A. (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, **153**, 307–319.
62. Loven,J., Hoke,H.A., Lin,C.Y., Lau,A., Orlando,D.A., Vakoc,C.R., Bradner,J.E., Lee,T.I. and Young,R.A. (2013) Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell*, **153**, 320–334.
63. Gaujoux,R. and Seoighe,C. (2010) A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics*, **11**, 367.
64. Daley,T. and Smith,A.D. (2013) Predicting the molecular complexity of sequencing libraries. *Nat. Methods*, **10**, 325–327.
65. Reimand,J., Isserlin,R., Voisin,V., Kucera,M., Tannus-Lopes,C., Rostamianfar,A., Wadi,L., Meyer,M., Wong,J., Xu,C.J. *et al.* (2019) Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat. Protoc.*, **14**, 482–517.
66. Raudvere,U., Kolberg,L., Kuzmin,I., Arak,T., Adler,P., Peterson,H. and Vilo,J. (2019) g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.*, **47**, W191–W198.
67. Shannon,P., Markiel,A., Ozier,O., Baliga,N.S., Wang,J.T., Ramage,D., Amin,N., Schwikowski,B. and Ideker,T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2504.
68. Merico,D., Isserlin,R., Stueker,O., Emili,A. and Bader,G.D. (2010) Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One*, **5**, e13984.
69. Kucera,M., Isserlin,R., Arkhangorodsky,A. and Bader,G.D. (2016) AutoAnnotate: A Cytoscape app for summarizing networks with semantic annotations. *F1000Res*, **5**, 1717.
70. Heinz,S., Benner,C., Spann,N., Bertolino,E., Lin,Y.C., Laslo,P., Cheng,J.X., Murre,C., Singh,H. and Glass,C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell*, **38**, 576–589.
71. Sondka,Z., Bamford,S., Cole,C.G., Ward,S.A., Dunham,I. and Forbes,S.A. (2018) The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. *Nat. Rev. Cancer*, **18**, 696–705.
72. Sos,B.C., Fung,H.L., Gao,D.R., Osothprarop,T.F., Kia,A., He,M.M. and Zhang,K. (2016) Characterization of chromatin accessibility with a transposome hypersensitive sites sequencing (THS-seq) assay. *Genome Biol.*, **17**, 20.
73. Chen,C., Xing,D., Tan,L., Li,H., Zhou,G., Huang,L. and Xie,X.S. (2017) Single-cell whole-genome analyses by Linear Amplification via Transposon Insertion (LIANTI). *Science*, **356**, 189–194.
74. Lake,B.B., Chen,S., Sos,B.C., Fan,J., Kaeser,G.E., Yung,Y.C., Duong,T.E., Gao,D., Chun,J., Kharchenko,P.V. *et al.* (2018) Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat. Biotechnol.*, **36**, 70–80.
75. Slatko,B.E., Gardner,A.F. and Ausubel,F.M. (2018) Overview of next-generation sequencing technologies. *Curr Protoc Mol Biol*, **122**, e59.
76. Kimura,H. (2013) Histone modifications for human epigenome analysis. *J. Hum. Genet.*, **58**, 439–445.
77. Kouzarides,T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.
78. Bernstein,B.E., Mikkelsen,T.S., Xie,X., Kamal,M., Huebert,D.J., Cuff,J., Fry,B., Meissner,A., Wernig,M., Plath,K. *et al.* (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, **125**, 315–326.
79. Saiki,R.K., Scharf,S., Faloona,F., Mullis,K.B., Horn,G.T., Erlich,H.A. and Arnheim,N. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, **230**, 1350–1354.
80. Liu,C.L., Schreiber,S.L. and Bernstein,B.E. (2003) Development and validation of a T7 based linear amplification for genomic DNA. *BMC Genomics*, **4**, 19.
81. Shi,S.R., Shi,Y. and Taylor,C.R. (2011) Antigen retrieval immunohistochemistry: review and future prospects in research and diagnosis over two decades. *J. Histochem. Cytochem.*, **59**, 13–32.
82. Webster,M., Witkin,K.L. and Cohen-Fix,O. (2009) Sizing up the nucleus: nuclear shape, size and nuclear-envelope assembly. *J. Cell Sci.*, **122**, 1477–1486.
83. Becker,J.S., McCarthy,R.L., Sidoli,S., Donahue,G., Kaeding,K.E., He,Z., Lin,S., Garcia,B.A. and Zaret,K.S. (2017) Genomic and

proteomic resolution of heterochromatin and its restriction of alternate fate genes. *Mol. Cell*, **68**, 1023–1037.
84. Zhu,X., Tian,X., Ji,L., Zhang,X., Cao,Y., Shen,C., Hu,Y., Wong,J.W.H., Fang,J.Y., Hong,J. *et al.* (2021) A tumor microenvironment-specific gene expression signature predicts chemotherapy resistance in colorectal cancer patients. *NPJ Precis Oncol*, **5**, 7.
85. Mack,S.C., Singh,I., Wang,X.X., Hirsch,R., Wu,Q.L., Villagomez,R., Bernatchez,J.A., Zhu,Z., Gimple,R.C., Kim,L.J.Y. *et al.* (2019) Chromatin landscapes reveal developmentally encoded transcriptional states that define human glioblastoma. *J. Exp. Med.*, **216**, 1071–1090.
86. Song,M.J., Kim,M., Choi,Y., Yi,M.H., Kim,J., Park,S.J., Yong,T.S. and Kim,H.P. (2017) Epigenome mapping highlights chromatin-mediated gene regulation in the protozoan parasite Trichomonas vaginalis. *Sci. Rep.*, **7**, 45365.
87. Xu,L., Chen,Y., Huang,Y., Sandanaraj,E., Yu,J.S., Lin,R.Y., Dakle,P., Ke,X.Y., Chong,Y.K., Koh,L. *et al.* (2021) Topography of transcriptionally active chromatin in glioblastoma. *Sci. Adv.*, **7**, eabd4676.
88. Cairns,J., Fridley,B.L., Jenkins,G.D., Zhuang,Y., Yu,J. and Wang,L. (2018) Differential roles of ERRFI1 in EGFR and AKT pathway regulation affect cancer proliferation. *EMBO Rep.*, **19**.
89. Kim,S.C., Shin,Y.K., Kim,Y.A., Jang,S.G. and Ku,J.L. (2018) Identification of genes inducing resistance to ionizing radiation in human rectal cancer cell lines: re-sensitization of radio-resistant rectal cancer cells through down regulating NDRG1. *BMC Cancer*, **18**, 594.
90. Saadeh,F.S., Mahfouz,R. and Assi,H.I. (2018) EGFR as a clinical marker in glioblastomas and other gliomas. *Int. J. Biol. Marker*, **33**, 22–32.
91. Gimple,R.C., Kidwell,R.L., Kim,L.J.Y., Sun,T., Gromovsky,A.D., Wu,Q., Wolf,M., Lv,D., Bhargava,S., Jiang,L. *et al.* (2019) Glioma stem cell-specific superenhancer promotes polyunsaturated fatty-acid synthesis to support EGFR signaling. *Cancer Discov.*, **9**, 1248–1267.
92. Luraghi,P., Reato,G., Cipriano,E., Sassi,F., Orzan,F., Bigatto,V., De Bacco,F., Menietti,E., Han,M., Rideout,W.M. 3rd *et al.* (2014) MET signaling in colon cancer stem-like cells blunts the therapeutic response to EGFR inhibitors. *Cancer Res.*, **74**, 1857–1869.
93. Gaffney,E.F., Riegman,P.H., Grizzle,W.E. and Watson,P.H. (2018) Factors that drive the increasing use of FFPE tissue in basic and translational cancer research. *Biotech. Histochem.*, **93**, 373–386.