

REVIEW

Mapping chromatin modifications at the single cell level

Connor H. Ludwig and Lacramioara Bintu*

ABSTRACT

Understanding chromatin regulation holds enormous promise for controlling gene regulation, predicting cellular identity, and developing diagnostics and cellular therapies. However, the dynamic nature of chromatin, together with cell-to-cell heterogeneity in its structure, limits our ability to extract its governing principles. Single cell mapping of chromatin modifications, in conjunction with expression measurements, could help overcome these limitations. Here, we review recent advances in single cell-based measurements of chromatin modifications, including optimization to reduce DNA loss, improved DNA sequencing, barcoding, and antibody engineering. We also highlight several applications of these techniques that have provided insights into cell-type classification, mapping modification co-occurrence and heterogeneity, and monitoring chromatin dynamics.

KEY WORDS: DNA methylation, Chromatin regulation, Gene regulation, Histone modifications, Single cell detection

Introduction

Most cells in an organism share the same genetic code, yet they express different sets of genes and display distinct cellular phenotypes that give rise to complex tissues with specialized functions. At the heart of this incredible phenotypic diversity is an intricate network of transcription factors, non-coding RNAs and chromatin regulators (Allis and Jenuwein, 2016). Chromatin serves as a substrate for these regulatory factors: chemical modifications on histone proteins and DNA can introduce recognition motifs for these factors and alter chromatin accessibility and 3D organization. These chromatin modifications can affect gene expression, genome replication, and DNA repair, and, in turn, these processes can change the state of chromatin. As a result, chromatin modifications are essential for establishing and maintaining cellular identity as well as for regulating dynamic responses during development and disease.

The repertoire of distinct histone modifications is large, with approximately 15 known chemical moieties and hundreds of modifiable residues across five constituent histone proteins and their variants (Huang et al., 2014). The best-characterized histone modifications include acetylation and methylation of lysine residues, primarily on the N-terminal tails of histones H3 and H4 (Fig. 1A). Lysine acetylation (Kac) is positively correlated with transcription, whereas lysine methylation (Kme) is either positively or negatively correlated with transcription depending on the residue, degree of methylation (Lawrence et al., 2016) and modification status of neighboring residues (Kouzarides, 2007).

DNA can also be modified, although fewer DNA modifications are known compared with histone modifications. The most common

DNA modification in mammalian genomes is cytosine methylation (5mC), which occurs most often at CpG dinucleotides (Ehrlich et al., 1982; Stevens et al., 2013). This type of methylation generally correlates negatively with transcription at promoters (Jones et al., 1998). Interestingly, for actively transcribed genes, DNA methylation within the gene body is believed to prevent aberrant transcription initiation (Neri et al., 2017). Successive oxidation of 5mC yields other possible modifications of cytosine: 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (Fig. 1B). Whereas 5mC and, to a lesser extent, 5hmC are stable, 5fC and 5caC are more transient. Currently, there is debate as to whether these less stable modifications play a functional role beyond serving as intermediates in the active demethylation pathway (Wu and Zhang, 2017; Zhu et al., 2018).

Given the diversity of chromatin modifications, there has been a massive effort to measure them at a genome-wide level and in different cell types (Allis and Jenuwein, 2016; Martens and Stunnenberg, 2013; Roadmap Epigenomics Consortium et al., 2015). The development of bulk methods for detecting chromatin modifications has expanded our understanding of their roles during development and across cell types. However, these bulk methods often require substantial starting material (tens of millions of cells) and produce population-averaged signals. These limitations diminish their utility in cases where sample size is small and heterogeneous, for example in clinical settings and when studying primary cells during development or *in situ*.

By contrast, single cell methods overcome such limitations and can actually measure heterogeneity in small samples (Fig. 2A). Additionally, single cell-based detection of chromatin modifications allows us to track cell lineages in development or disease (Fig. 2B). Moreover, now that we know the identity of most molecular players in the chromatin regulatory network (Kouzarides, 2007), it is time to dissect the finer mechanistic details of their dynamic interactions. Because chromatin and gene regulation are stochastic and can lead to cell-to-cell heterogeneity, many questions concerning these processes would benefit from single cell approaches. For example, what combinations of chromatin modifications appear together in individual cells, and which signals result from cell-to-cell heterogeneity (Fig. 2C)? What type and number of modifications are necessary and sufficient to effect changes in gene expression? Finally, how and when does partitioning of chromatin during replication lead to epigenetic memory? Answering these latter questions requires a combination of single cell methods that detect chromatin state and RNA levels (Tanay and Regev, 2017).

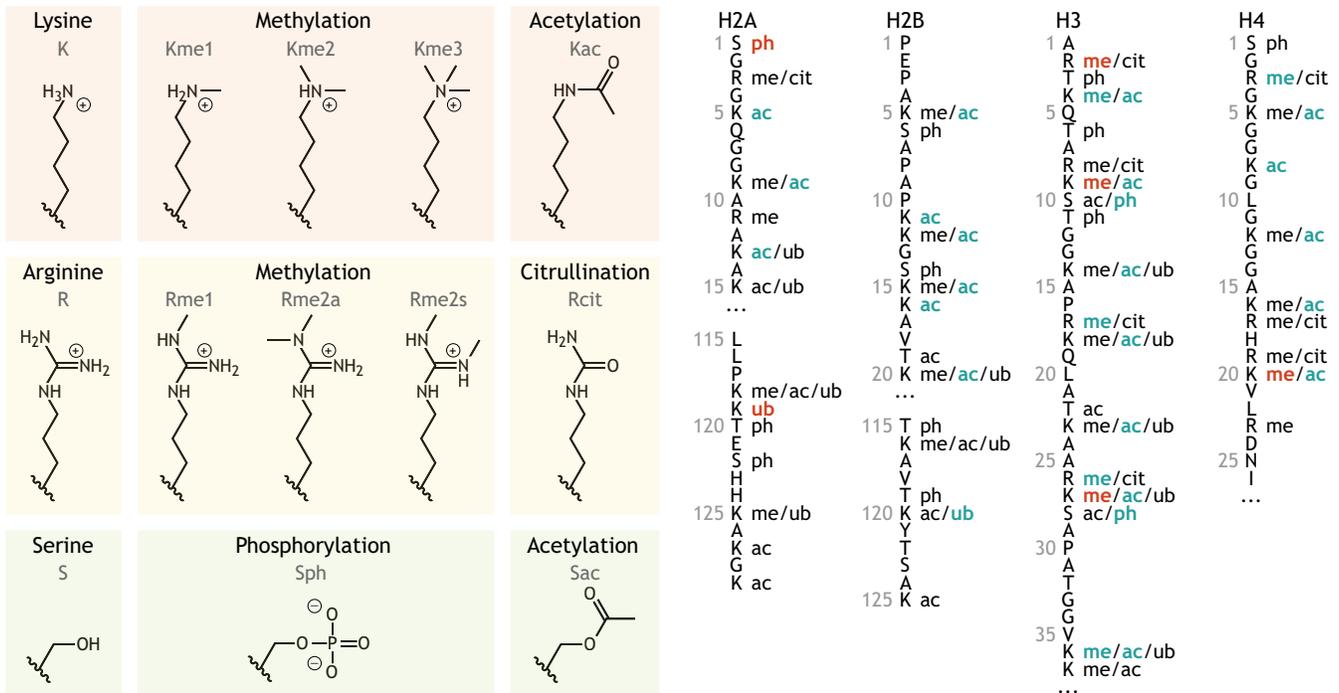
In this Review, we describe and assess (see Box 1, Table 1) methods for detecting and mapping chromatin modifications, focusing on technical advancements that have enabled adaptation to the single cell level and on ways of combining them with measurements of other cellular characteristics. Even though chromatin accessibility, three-dimensional chromatin structure and subnuclear localization contribute to cell identity and gene control (for recent reviews, see Shema et al., 2018; Tanay and Regev, 2017; Tycko et al., 2017), we will not cover these topics. Finally, we

Department of Bioengineering, Stanford University, Shriram Center, 443 Via Ortega, Rm 042, Stanford, CA 94305, USA.

*Author for correspondence (lbintu@stanford.edu)

 L.B., 0000-0001-5443-6633

A Histone modifications



B DNA modifications

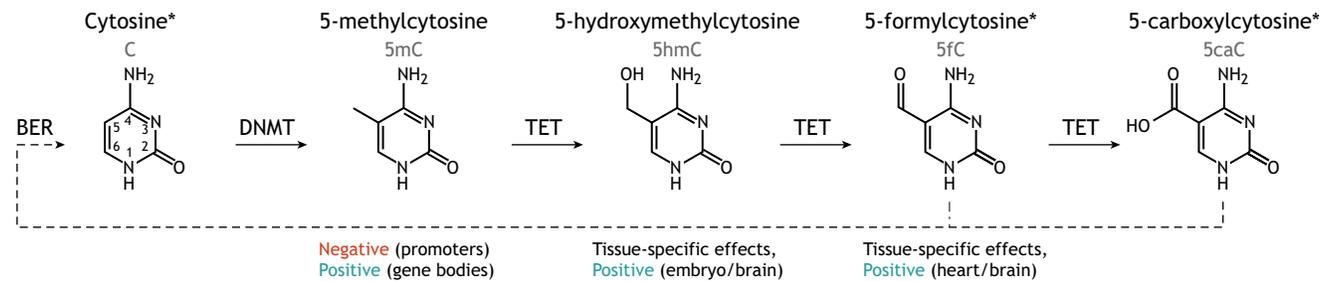


Fig. 1. Key types of chromatin modifications. (A) Some of the most common histone modifications (shown on the left) occur at lysine (K), arginine (R), serine (S) and threonine (not shown) residues. These occur on the protruding tails of various histones (right) and can be positively (blue) or negatively (red) associated with transcription. Residue number, single-letter amino acid codes, and modifications are shown. ac, acetylation; cit, citrullination; me, methylation; ph, phosphorylation; ub, ubiquitylation. Adapted from Huang et al. (2014) and Lawrence et al. (2016). (B) Cytosine, when in the form of deoxycytidine monophosphate in DNA, can be methylated by DNA methyltransferases (DNMTs) and further oxidized by TET enzymes to produce four distinct species with context-specific effects on transcription, as indicated. The final two species in this pathway may be converted back to unmodified cytosine through base excision repair (BER). Asterisks indicate nucleobases that are chemically converted to uracil upon bisulfite treatment.

describe biological applications and insights gained using these technologies and discuss how they could be improved and incorporated into global scientific initiatives.

Classical methods for measuring histone modifications

The detection of histone modifications relies primarily on chromatin fragmentation by sonication or DNA digestion, followed by immunoprecipitation of the resulting mononucleosomes with antibodies against a specific modification (in a process termed chromatin immunoprecipitation, ChIP). Readout of the genomic location of the modified nucleosomes is then performed using quantitative PCR of a specific genomic locus (ChIP-qPCR) or next-generation sequencing (NGS) for genome-wide detection (ChIP-seq; Fig. 3A).

Classical ChIP has some major limitations that preclude its adaptation to the single cell level robustly. First, crosslinking before fragmentation can lead to artifacts, decrease reproducibility and

impair immunoprecipitation. Thus, recent protocols have excluded this step, instead performing native ChIP (Kasinathan et al., 2014). Second, chromatin sonication requires large numbers of cells (hundreds of thousands to millions). The alternative to sonication, micrococcal nuclease (MNase) digestion, can be performed with smaller cell numbers or *in situ* (Skene et al., 2018). However, MNase treatment can digest away a portion of nucleosomal DNA (Henikoff et al., 2011), thus reducing the number of mapped reads from a single cell. Third, antibodies used for immunoprecipitation may exhibit low affinities for their respective targets, requiring increased input (i.e. large cell numbers); additionally, they may vary from lot to lot and have a low specificity (Kungulovski et al., 2014).

Genetically encoded histone modification-specific recognition proteins present an ideal alternative to antibodies, as they can be engineered for greater affinity and specificity and could improve reproducibility. One such approach employs ‘reader’ domains that have evolved to bind specific histone modifications (Yun et al.,

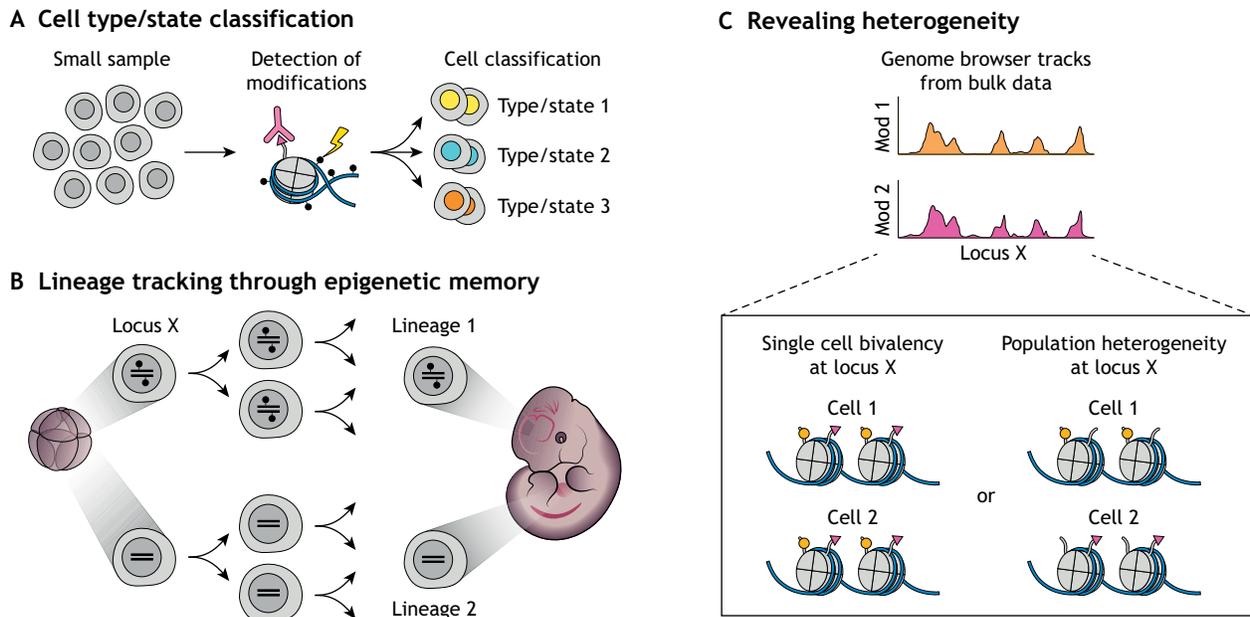


Fig. 2. Measuring chromatin modifications at the single cell level can provide mechanistic insights into biological questions. (A) Distinct single cell chromatin modification profiles enable computationally aided cell-type or cell-state classification. (B) Single cell methods allow cell lineages to be tracked by following epigenetic signatures maintained throughout differentiation and development. (C) Measuring chromatin modifications (e.g. Mod 1 and Mod2) at the single cell level can resolve averaging-associated ambiguities from bulk data to reveal potential population heterogeneity.

2011). Although binding affinities for their targets are often lower than those of antibodies (Kungulovski et al., 2014), reader domains have known sequences and thus can be evolved to improve affinity and specificity (Tekel et al., 2018). Similarly, fluorescent modification-specific intracellular antibodies (mintbodies) are small, GFP-tagged, single-chain variable fragments that can be expressed *in vivo* to allow live imaging of histone modification dynamics spatiotemporally (Sato et al., 2013) (Fig. 3J). Although only anti-H3K9ac and anti-H4K20me1 mintbodies have been developed so far (Sato et al., 2016), a future library of these tools could enable multiplexed quantification of chromatin dynamics in live cells.

Box 1. Metrics of single cell technologies

In order to compare different single cell technologies, we use here a set of metrics that were previously defined (Tanay and Regev, 2017) – complexity, accuracy, throughput and efficiency – to which we will add a fifth: portability. Complexity describes the number of molecules that can be identified per cell. The complexity metric has multiple axes: the number of chromatin modifications, the number of genomic loci (Fig. 6), and the ability to follow modifications and loci dynamically over time and space. Accuracy refers to the signal-to-noise ratio of the assay, with high accuracy requiring high detection efficiency of the correct chromatin modifications and a low error rate. Throughput is defined as the number of cells that can be processed in an assay, and efficiency refers to the fraction of those cells that yields usable data. Finally, additional metrics, such as reproducibility between users and biological replicates, as well as accessibility, cost, and ease-of-use of equipment and reagents, can be subsumed under the broader category of portability. In a sense, how easily can a technology be ported to a new physical research environment or biological question with minimal troubleshooting and financial cost? As most current single cell technologies for mapping chromatin modifications are still in their infancy, each of them usually scores highly in one or two of these categories (Table 1).

Single cell measurements of histone modifications

In order to decipher the role of histone modifications in different cellular functions, we would ideally need to measure multiple histone modifications in single cells, link this information to defined genetic loci, and measure the dynamics over time in live cells. Although this type of multiplexed, genome-wide, live dynamics of histone modifications has not yet been achieved at the single cell level, recent advances – some of which are adaptations of classical ChIP – have allowed histone modifications to be probed in various ways at the single cell level.

Detecting and mapping single histone modifications over multiple loci

At present, single cell ChIP-seq (scChIP-seq) was the first technique that offers multi-locus reporting of a histone modification at the single cell level (Rotem et al., 2015) (Fig. 3B). To overcome the limitations associated with sonication, scChIP-seq uses microfluidics to isolate cells into individual droplets, in which chromatin is digested and barcoded before classical immunoprecipitation. Approximately 100 cells can be processed per assay, generating 500-10,000 unique reads per cell. In order to extract meaningful insights from the data despite limited coverage and low cell numbers, the authors of the above study clustered single cell data belonging to genomic regions likely to be similarly modified, such as a particular enhancer and the promoters it drives. This classification of genomic regions was performed using publicly available bulk ChIP-seq data. In this manner, the authors identified three distinct H3K4me2-associated epigenetic signatures within heterogeneous mouse embryonic stem cells (mESCs) grown in serum, which primes these cells for differentiation, and showed that these states correlate with the expression of pluripotency and chromatin factors (Rotem et al., 2015). The low cell yield and coverage, together with the need for a specialized microfluidics device, might explain why scChIP-seq has not yet been applied widely. However, as the authors suggest, the cell yield can be increased by expanding the number of barcodes using beads

Table 1. Assessment of single cell methods for detecting DNA and histone modifications

Method	Throughput	Efficiency	Complexity	Accuracy	Portability	Reference(s)
Histone modifications						
scChIP-seq	100 cells	5000-10,000 reads/cell (66-79% reads mapped)	One modification Multi-locus Endpoint	DR: NR FP: NR	Low (microfluidics)	Rotem et al., 2015
scChIL-seq	Five cells	10,000-100,000 reads/cell	One modification Multi-locus Endpoint+Spatial	DR: 10-93% FP: NR	Moderate	Harada et al., 2019
scCUT&Tag	1000 cells	1000-100,000 reads/cell	One modification Multi-locus Endpoint	DR: NR FP: NR	Moderate (nanowells)	Kaya-Okur et al., 2019
scChIC-seq	100-300 cells	1000-100,000 reads/cell	One modification Multi-locus Endpoint	DR: 47-61% FP: NR	High	Ku et al., 2019
Co-ChIP*	1-10×10 ⁶ cells	NR	Two modifications Multi-locus Endpoint	DR: NR FP: NR	Moderate	Weiner et al., 2016
SMD*	1-10×10 ⁶ cells	80% reads mapped	Two modifications Multi-locus Endpoint	DR: 75% FP: 0.1%	Low (TIRF)	Shema et al., 2016
EpiTOF	1-10×10 ⁶ cells	NR	<60 modifications Whole cell Endpoint	DR: NR FP: NR	Low (FACS-MS)	Cheung et al., 2018
ISH-PLA	Whole tissue section or culture	NR	One modification One locus Endpoint	DR: 66% FP: NR	Low (confocal)	Gomez et al., 2013
Mintbodies	Organism or cell culture	NR	One modification Whole cell Dynamics	DR: NR FP: NR	Low (live imaging)	Sato et al., 2013 Sato et al., 2016
DNA modifications						
scRRBS-seq	10-100 cells	80-100% cells (20% reads mapped)	5mC/5hmC Multi-locus (max 10% of CpGs) Endpoint	DR: NR FP: <0.8%	High	Guo et al., 2013 Wang et al., 2015 Ma et al., 2018
scBS-seq	10-100 cells	100% cells (20.1% reads mapped)	5mC/5hmC Multi-locus (max 48.4% of CpGs) Endpoint	DR: >92% FP: <2.3%	High	Smallwood et al., 2014 Clark et al., 2017
scAba-seq	100-1000 cells	43.4% cells (mapping efficiency NR)	5hmC Multi-locus Endpoint	DR: 10% FP: 2%	High	Mooijman et al., 2016
SMRT*	NR	NR	6mA Multi-locus Endpoint	DR: NR FP: NR	Low	Flusberg et al., 2010 Suzuki et al., 2016
Nanopore*	~1×10 ⁷ cells (100 µg starting DNA; 20 ng loaded)	NR	6mA, 5mC, 5hmC Multi-locus Endpoint	DR: NR FP: NR	High	Laszlo et al., 2013 Schreiber et al., 2013 Simpson et al., 2017 Rand et al., 2017

DR, detection rate; FP, false positives; NR, not reported.

*Single molecule, not single cell.

(Klein et al., 2015; Macosko et al., 2015), and higher coverage could be achieved by optimizing ligation efficiency and amplification.

Very recently, three new single cell methods for detecting histone modifications were published: single cell chromatin integration labeling followed by sequencing (scChIL-seq) (Harada et al., 2019) (Fig. 3C), single cell cleavage under targets and tagmentation (scCUT&Tag) (Kaya-Okur et al., 2019) (Fig. 3D), and single cell chromatin immunocleavage followed by sequencing (scChIC-seq) (Ku et al., 2019) (Fig. 3E). All three methods replace traditional immunoprecipitation and sonication with *in situ* (inside nuclei) reactions: either via antibody-directed, transposase-mediated integration of a DNA tag and fragmentation (for scChIL-seq and scCUT&Tag), or via DNA cleavage specifically around nucleosomes containing the target modification (for scChIC-seq). These *in situ* reactions remove the need to solubilize chromatin and

thus reduce material loss. scChIL-seq and scCUT&Tag convolute chromatin accessibility with immune recognition of the histone modification, as the transposase is biased toward accessible chromatin (Buenrostro et al., 2015). Overall, this accessibility bias results in low Pearson's correlation coefficients between bulk ChIP-seq and ChIL-seq, especially for repressive histone modifications, such as H3K27me₃, that lie within less accessible chromatin ($r=0.26-0.31$). This bias is also observed for the MNase action in scChIC-seq for H3K27me₃, albeit to a lesser extent ($r=0.67$). Although the application of these methods is in its infancy, they have unique advantages over scChIP-seq: scChIL-seq provides spatial information by imaging the inserted probe; scCUT&Tag and scChIC-seq have simplified workflows, are higher throughput owing to expanded barcodes, and are more cost-effective. Therefore, we foresee further development and applications of these techniques.

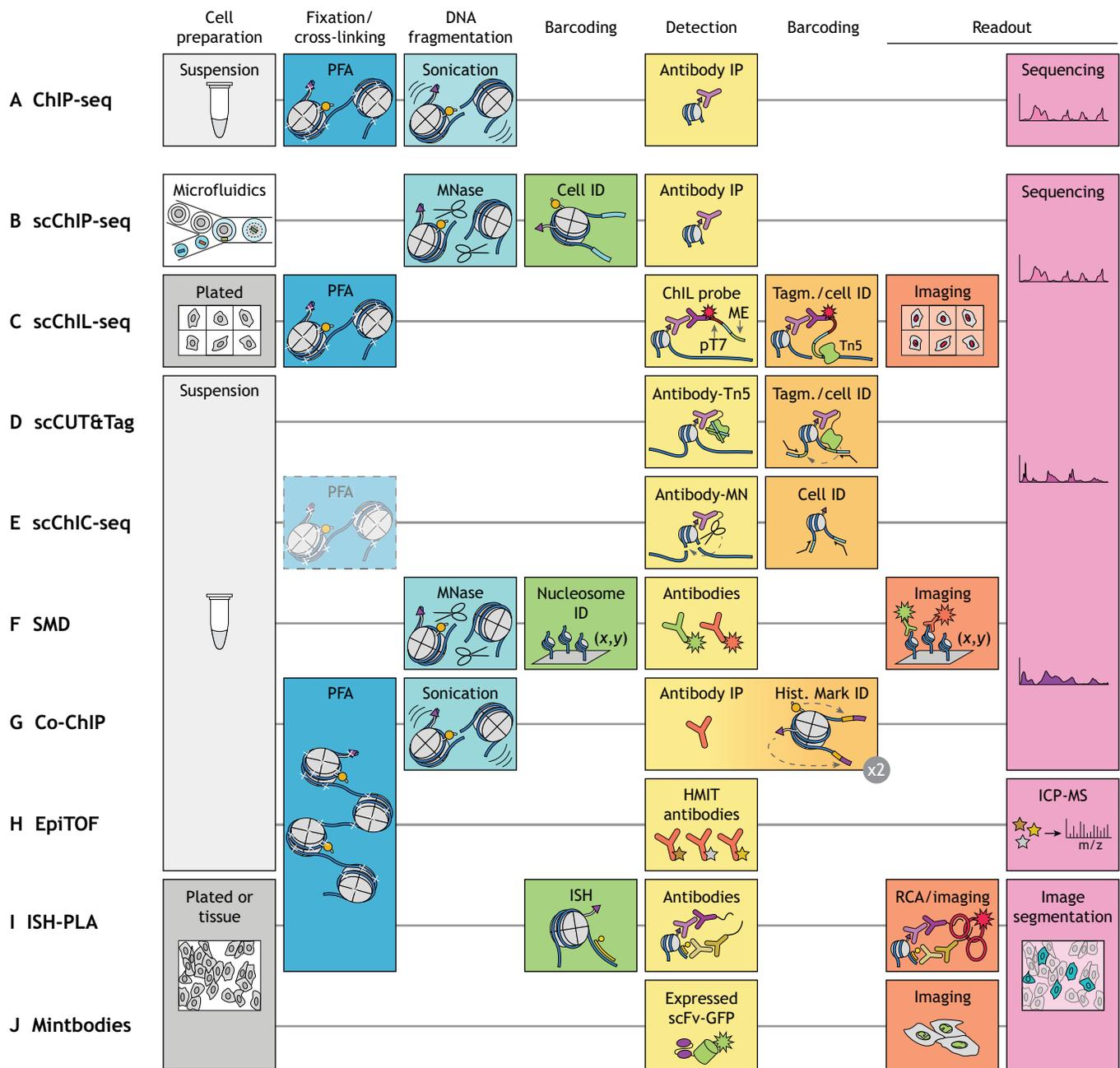


Fig. 3. See next page for legend.

Detecting and mapping co-occurring histone modifications

Although mapping one modification in single cells is informative, many loci have multiple modifications that contribute to their function (Strahl and Allis, 2000). For example, co-occurrence of histone modifications with opposing effects, specifically H3K4me3 (active) and H3K27me3 (repressive), is thought to poise a locus for fast transcriptional activation or silencing (Bernstein et al., 2006). Such bivalency is thought to be especially important during lineage specification and development. A recently developed imaging-based technology employed genetically encoded chromatin-sensing multivalent probes (cAMPs) to examine this bivalency (Delachat et al., 2018) at the whole-cell level. The probe – a fluorescent fusion protein containing micromolar-affinity readers of the H3K4me3 and H3K27me3 modifications – produces a stable signal at locally

clustered bivalent chromatin, revealing localization changes upon treatment with methyltransferase inhibitors.

Two other methodologies have been used to not only identify co-occurring modifications on single nucleosomes but also map their genomic locations (Shema et al., 2016; Weiner et al., 2016). Although these techniques are single molecule-based, with each nucleosome originating from a single cell, they are not single cell at the genome-wide level, as there is no mechanism for grouping different nucleosomes together based on cell of origin in current workflows. The first technology providing modification co-occurrence information at the single nucleosome level is an imaging-based method combining visual detection of histone modifications with on-slide sequencing (Shema et al., 2016) (Fig. 3F). In this method, single nucleosomes are immobilized on

Fig. 3. Detecting and mapping histone modifications at the single cell level. (A) ChIP-seq is the classical approach to mapping histone modifications in bulk samples. It involves cell preparation, cross-linking of protein/DNA components by paraformaldehyde (PFA), chromatin fragmentation by sonication, immunoprecipitation (IP), and sequencing. Variations of this approach enable single cell level measurements to be made. (B) scChIP-seq employs microfluidics to encapsulate individual cells in droplet reaction chambers for DNA fragmentation by micrococcal nuclease (MNase) and ligation of cell-specific identifiers. Barcoded cells are pooled for IP, library preparation, and sequencing, as in bulk ChIP-seq. (C) scChIL-seq is performed on fixed single cells deposited in individual microwells. A primary antibody against a histone modification of interest is detected by a chromatin integration labeling (ChIL) probe comprising a secondary antibody conjugated to TAMRA fluorophore-labeled double-stranded DNA. This probe is imaged (via TAMRA) to determine the subnuclear localization of the modified nucleosomes and integrated into the sequence adjacent to the modified nucleosome via Tn5 transposase-binding of the mosaic end (ME) in a process called tagmentation (Tagm.). This resulting sequence is amplified by *in situ* transcription from a pT7 sequence included in the DNA oligo; this RNA library is reverse transcribed (not shown), and well-specific (cell-specific) barcodes are introduced during PCR amplification (not shown). (D) scCUT&Tag uses an antibody-Tn5 or a protein A-Tn5 fusion that associates with a pre-incubated antibody against a specific histone modification to insert pre-complexed adapters near modified nucleosomes upon magnesium addition. Insertion is performed on a pool/suspension of cells, which are then individually sorted into a chip array of nanowells with unique indexing primer pairs. In each nanowell, DNA is fragmented upon Tn5 dissociation and, upon PCR amplification, cell-specific barcodes are introduced. (E) scChIC-seq can be used on fixed or unfixed cells and involves incubation with an antibody-MNase or protein A-MNase fusion that associates with a pre-incubated antibody against a specific histone modification. Single cells are sorted into tubes and calcium is added to trigger MNase-mediated fragmentation. PCR amplification strongly selects for small, mononucleosomal fragments and enables cell-specific barcode incorporation. (F) Single molecule decoding of combinatorially modified nucleosomes (SMD) identifies the positions of bound nucleosomes and detects antibody-labeled histone modifications via total internal reflection fluorescence (TIRF) microscopy, followed by dissociation of histone proteins and on-slide sequencing of the remaining, bound DNA. (G) Co-ChIP employs two rounds ($\times 2$) of IP on bulk samples. After the first IP reaction, each nucleosome is uniquely barcoded via adapter ligation to effectively 'record' detection of the first histone modification at the DNA level (dashed arrow). Subsequently, chromatin is released, pooled, and divided for a second set of IPs. The resulting DNA is PCR-amplified with primers unique to the second modification to record its presence (dashed arrow). NGS enables mapping of nucleosomes that contain particular combinations of modifications to the genome. (H) EpiTOF employs heavy metal isotope-tagged (HMIT) antibodies and inductively coupled plasma mass spectrometry (ICP-MS) to report global levels of up to 60 targets in a single cell. (I) ISH-PLA involves DNA *in situ* hybridization (ISH) of a biotin-labeled oligo to a target sequence. Samples are then incubated with primary antibodies against biotin and a histone modification of interest. DNA oligo-labeled secondary antibodies serve as primers for rolling circle amplification (RCA). These primers hybridize to circular DNA to enable successive rounds of amplification by DNA polymerase, and the amplified DNA is detected with fluorescent single-stranded DNA probes. (J) The use of genetically encoded mCherry antibodies against histone modifications of interest enables the global measurement of temporal and spatial chromatin dynamics in live cells. scFv-GFP, single-chain variable fragment fused to GFP.

glass slides and imaged using fluorescent antibodies against two histone modifications at a time. Successive rounds of antibody removal and incubation can provide co-occurrence information for more than two histone modifications. This approach allowed the authors to measure the fractions of nucleosomes containing particular combinations of modifications in different cell types or after treatment with epigenetic inhibitors. Furthermore, dissociation of the histones and single molecule sequencing on the remaining slide-bound DNA enables mapping of each nucleosome to its genomic location. The addition of a cell-barcoding step could advance this technique to the genome-wide, single cell level.

However, the current workflow and required infrastructure/equipment limit the accessibility of this technology.

The second technology, termed Co-ChIP, has been used to characterize the co-occurrence of 70 histone modification pairs (14 primary modifications against five secondary modifications) at the single nucleosome level (Weiner et al., 2016) (Fig. 3G). Co-ChIP largely proceeds similarly to classical ChIP, whereby chromatin is cross-linked, sonicated and subjected to two sets of immunoprecipitations. However, after each immunoprecipitation step, each particular histone modification is effectively recorded to its nucleosomal DNA by adaptor ligation or PCR barcoding, thus allowing mapping of co-occurring modifications on single nucleosomes. Using this method, the authors discovered a new combination of modifications (H3K9me1 and H3K27ac) that is characteristic of super-enhancers and characterized bivalent domains (marked by H3K4me3 and H3K27me3) in different cell types. One can envision adding cell barcoding to this workflow in order to profile combinations of modifications across different regulatory regions in the same cell.

A recently published mass cytometry platform, termed epigenetic landscape profiling using cytometry by time-of-flight (EpiTOF; Cheung et al., 2018), greatly extends the number of histone modifications that can be measured simultaneously in a given cell (Fig. 3H). This multiparametric detection method is based on cytometry by time-of-flight (CyTOF), which can label up to 60 targets with heavy metal isotope-tagged antibodies and measure single cell global levels by mass spectrometry (Bandura et al., 2009). EpiTOF was used to measure cell type-specific histone modification patterns that could be used to predict immune cell identity (Cheung et al., 2018). Additionally, it was demonstrated that intercellular variability of histone modifications and histone variants increases in human immune cells with age. The throughput and target complexity of EpiTOF supersedes that of other epigenetic technologies; however, readouts of global levels preclude a precise understanding of chromatin state at individual loci, as mass spectrometry cannot be easily coupled with sequencing.

Monitoring spatial and temporal distributions of chromatin modifications

DNA *in situ* hybridization and proximity ligation (ISH-PLA) can provide imaging-based, locus-specific histone modification information, as well as cell phenotype and cell-cell interaction information (Gomez et al., 2013) (Fig. 3I). This method uses two oligo-tagged antibodies, one to recognize a genomic locus and the other a histone modification. When both are in proximity, rolling circle amplification from the two oligos locally produces DNA that can be imaged. ISH-PLA has been used in histological samples to demonstrate that epigenetic signatures can be used for lineage tracing and identifying cell types in the context of atherosclerosis, even in the absence of expression of classical marker genes (Gomez et al., 2013). However, ISH-PLA has not been widely adopted, perhaps because of its requirement for special equipment, its non-quantitative readout, and the low volume of information (i.e. single locus, single modification) that can be extracted from each experiment.

Measurements of histone modifications at a single locus have also been performed in live cells, although the locus needs to contain repetitive sequences in order to produce a detectable signal. The repetitive regions assayed were either engineered to contain a ~ 200 copy tandem gene (Stasevich et al., 2014) (Fig. 4A) or were part of natural satellite repeats (Lungu et al., 2017). In the tandem gene system, antibody-derived, fluorescently labeled antigen-binding fragments (Fabs) were introduced to measure, at the single cell

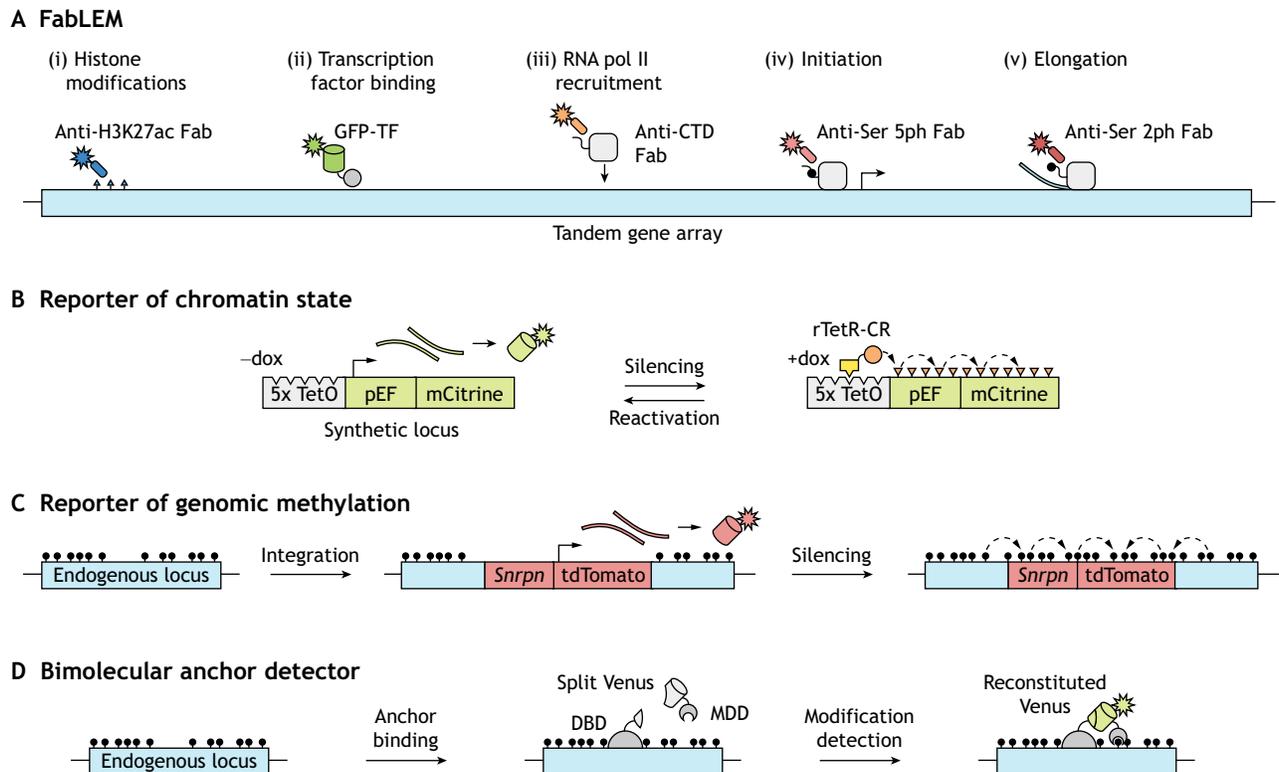


Fig. 4. Monitoring chromatin modifications and transcription dynamics. (A) Introducing fluorescent antigen-binding fragments (Fabs) into cells for live endogenous modification labeling (LEM) can elucidate chromatin and transcription dynamics to examine their interplay. For example, one can track histone modifications (i), GFP-labeled transcription factors (ii), and phosphorylated serine residues (Ser) in the carboxy-terminal domain (CTD) of RNA pol II that are related to recruitment, initiation and elongation (iii-v). Figure adapted from Stasevich et al. (2014). (B) Gene reporter systems can also be used to follow chromatin dynamics. In the example shown, doxycycline (dox) mediates the recruitment of a chromatin regulator (CR; orange circle) fused to a reverse tetracycline repressor (rTetR; yellow box) to TetO sites upstream of a constitutively expressed fluorescent reporter gene. Recruitment induces gene silencing and deposition of chromatin modifications, and CR release allows gene reactivation. (C) Methylation at an endogenous locus can be monitored by integration of a reporter gene that consists of a minimal imprinted gene promoter (*Snrpn*) and the coding sequence for a fluorescent protein (tdTomato). DNA methylation spreads from the endogenous locus into the reporter, leading to its silencing. (D) The direct readout of DNA or histone methylation at a locus of interest is possible upon reconstitution of a split fluorescent protein (Venus), the constituents of which are fused to a programmable DNA-binding domain (DBD) and a modification detection domain (MDD).

level, the effect of H3K27ac on RNA polymerase II recruitment, initiation, and elongation dynamics at the gene array with temporal resolution on the order of tens of seconds. This live endogenous modification labeling revealed that H3K27ac acts to increase accessibility for incoming transcription factors and outgoing RNA polymerase II, as evidenced by increased elongation rates.

Gene reporter systems that indirectly monitor single cell chromatin dynamics, avoiding the need for antibodies or reader domains, have also been utilized. For instance, gene silencing and reactivation dynamics can be quantified via flow cytometry and live imaging with single cell tracking upon recruitment or release of chromatin regulators at a fluorescent reporter (Bintu et al., 2016) (Fig. 4B). Using this approach, it was demonstrated that various chromatin regulators associated with repressive histone methylation, DNA methylation or histone deacetylation impart distinct kinetics and permanence of epigenetic memory. Notably, the authors proposed a three-state model of gene expression – with active, reversibly silent, and irreversibly silent states – to which they could fit single cell data to classify each chromatin regulator based on calculated kinetic parameters.

Classical methods for measuring DNA modifications

Classically, mapping DNA modifications relies on one of the following techniques: (1) chemical conversion of unmodified

nucleobases to another nucleobase (Frommer et al., 1992); (2) recognition of the modified nucleobases by restriction enzymes (Jelinek and Madzo, 2016); or (3) recognition of the modified nucleobases by antibodies (Weber et al., 2005). The last two methods are less popular: the second is restricted to detecting modifications at the subset of DNA loci that contain the particular enzyme recognition sequence, and the third method is not ideal because existing antibodies do not discriminate well between DNA modifications, and it has low resolution (a few hundred base pairs, limited by fragment length).

The most commonly used method employs sodium bisulfite to convert unmethylated cytosines to uracil, followed by sequencing (BS-seq) (Frommer et al., 1992) (Fig. 5A). Although BS-seq can detect modifications at the single base level, it suffers from three major drawbacks. First, bisulfite treatment reduces all possible cytosine modifications to one of two states: converted or unconverted. Specifically, both 5mC and 5hmC, which can have opposing effects on gene expression (Zhu et al., 2018), are protected from conversion (Huang et al., 2010), whereas C, 5fC and 5caC are substrates for conversion. Second, bisulfite treatment is harsh: because conversion is much more efficient on single-stranded versus double-stranded DNA (Shapiro et al., 1974), the required acidic conditions and high temperatures ultimately degrade approximately 90% of the input DNA (Grunau et al., 2001). Moreover, as sequencing adapters are traditionally added to DNA prior to bisulfite treatment, a substantial

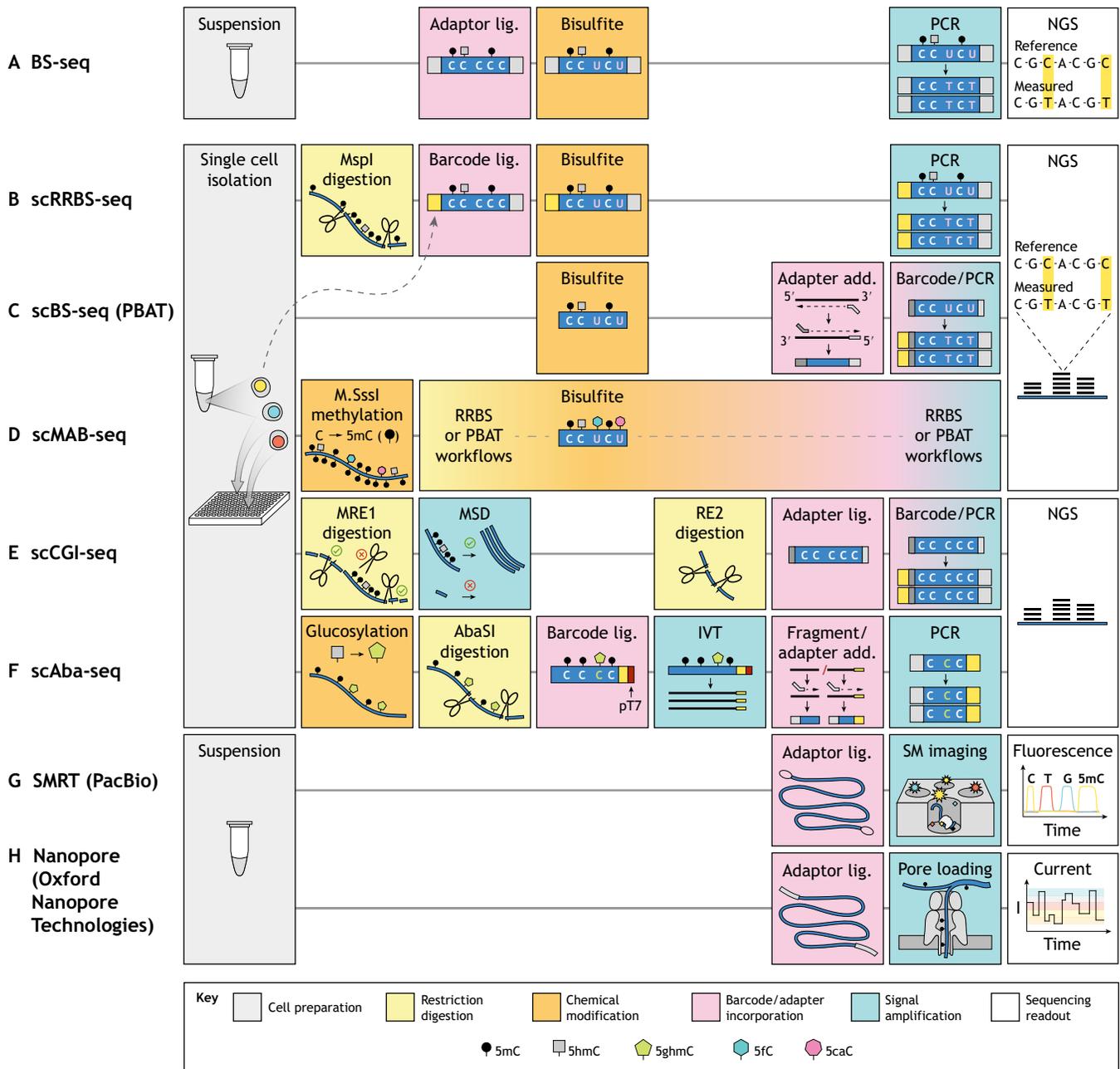


Fig. 5. Detecting and mapping DNA modifications at the single cell level. (A) BS-seq is the most commonly used bulk method for mapping 5mC/5hmC with single nucleotide resolution. This approach involves cell preparation, DNA adaptor ligation, bisulfite conversion, PCR amplification, and sequencing. Bisulfite converts only C, 5fC and 5caC to uracil, which are read as thymines in next-generation sequencing (NGS), whereas 5mC and 5hmC resist conversion and are read as cytosines. This technique has been modified in various ways to enable single cell-based measurements. (B) scRRBS-seq maps 5mC/5hmC and uses a one-tube protocol to minimize sample loss. The restriction endonuclease MspI releases CpG-rich DNA fragments for bisulfite treatment, increasing the amount of CpG information that can be obtained with fewer sequencing reads. (C) scBS-seq provides whole-genome 5mC/5hmC information with less material loss than traditional methods by using post-bisulfite adaptor tagging (PBAT) to avoid DNA damage that otherwise prevents PCR amplification. (D) scMAB-seq uses the methyltransferase M.SssI to convert all unmodified cytosines to 5mC such that the only nucleobases converted to uracil by bisulfite are 5fC and 5caC. Both RRBS and PBAT workflows can be employed to map 5fC/5caC. (E) scCGI-seq uses a cocktail of methylation-sensitive restriction enzymes (MRE1) to fragment unmethylated regions while preserving the length of methylated ones. Multiple strand displacement (MSD) preferentially amplifies longer fragments, and a second cocktail of restriction enzymes (RE2) digests these long pieces into smaller fragments appropriate for NGS. (F) scAba-seq involves glucosylation of 5hmC by T4 phage β -glucosyltransferase, followed by glucosylation-dependent digestion by the AbaSI restriction endonuclease, cell-specific barcoded adaptor ligation, and *in vitro* transcription amplification (IVT) from the T7 promoter (pT7). The transcripts are fragmented and subjected to RNA-seq library preparation. (G) Single molecule real-time sequencing (SMRT) is a third-generation, sequencing-by-synthesis technique that uses extremely small imaging volumes to boost the fluorescent signal from nucleotide incorporation into a single molecule (SM) template. Special adapters circularize the template to enable many reads of the same molecule, and modified nucleobases can be distinguished from unmodified ones based on differential kinetics. (H) Nanopore sequencing is a third-generation approach that reads the current as nucleobases pass through a transmembrane pore. Special adapters are required for pore loading, and nucleobase (modified or unmodified) identity can be predicted from the associated current reading. Both SMRT and Nanopore sequencing can read much longer molecules than NGS.

portion of sequences fail to amplify as a result of DNA damage. Third, the conversion of unmodified cytosines, which normally account for ~20% of the genome, to uracil greatly reduces sequence complexity. This constrains PCR primer design for library preparation and potentially hinders downstream sequencing and genome mapping (Callinan and Feinberg, 2006).

Single cell measurements of DNA modifications

Current methods for detecting DNA modifications at the single cell level are either direct – using bisulfite conversion, methylation sensitive restriction enzymes, or third-generation sequencing – or indirect – using fluorescent reporters sensitive to DNA methylation (for a recent review, see Karemaker and Vermeulen, 2018). These methods primarily focus on 5mC and 5hmC, the two most common DNA modifications, but a subset could be adapted to detect other modifications.

Bisulfite-based measurements of DNA modifications

Bisulfite conversion followed by sequencing was the first method to map DNA modifications to many loci at the single cell level and is currently the most developed in terms of throughput and base-pair resolution. As previously discussed, bisulfite conversion of unmodified cytosines cannot discriminate between 5mC and 5hmC (Huang et al., 2010). This limitation must be considered, especially when interpreting mechanistic results in cell types with higher levels of 5hmC, such as embryonic stem cells and neurons (Wu and Zhang, 2017).

The first single cell, multi-locus method for DNA methylation mapping used a modified version of reduced representation bisulfite sequencing (RRBS-seq) (Gu et al., 2011; Meissner et al., 2005), performing all reaction steps for a single cell in one tube in order to reduce DNA loss (Guo et al., 2013) (Fig. 5B). RRBS-seq involves enrichment of CpG-rich regions by digesting DNA with the restriction endonuclease MspI at C[^]CG_^G sites (approximately once per 256 bp), followed by ligation of sequencing adapters. This guarantees that each adapter-tagged fragment contains at least two CpG sites and allows for comparison of methylation profiles at select CpGs with fewer overall sequencing reads. However, single cell RRBS-seq (scRRBS-seq) can only detect a theoretical maximum of 10% of CpGs and still suffers from reduced material recovery, as bisulfite conversion is performed after adapter ligation. Nonetheless, the authors used this method to show that DNA demethylation in mouse zygotes occurs more rapidly at gene regions compared with intergenic regions after fertilization.

Single cell bisulfite sequencing (scBS-seq) has also been extended to the whole genome (Smallwood et al., 2014) by using post-bisulfite adaptor tagging (PBAT) (Miura et al., 2012). By performing adapter addition after the DNA-damaging bisulfite conversion step, this method avoids accidental fragmentation of adapter-tagged DNA molecules, thereby increasing the amount of DNA that can be successfully amplified during library preparation (Fig. 5C). The whole-genome nature of scBS-seq provides fivefold more CpG information than scRRBS-seq at the same sequencing depth (Smallwood et al., 2014). The reproducibility of the scBS-seq protocol was benchmarked using oocytes, which are highly homogeneous and have well-defined DNA methylation profiles. Using this approach, it was also demonstrated that mESCs grown in serum have increased variability of DNA methylation at active enhancers, suggesting that enhancers are the first triggers in cell fate changes.

Several recent alterations to the basic scBS-seq and scRRBS-seq protocols have increased throughput, decreased amplification

bias and improved data analysis. For example, workflow optimization, such as a reduction in the number of protocol steps and tube transfers (Farlik et al., 2015), has improved material recovery and enabled automation via liquid-handling robots (Clark et al., 2017) or microfluidic devices (Ma et al., 2018). An emerging, cost-effective strategy aims to sequence a larger number of cells at low coverage in order to construct composite reference methylome maps for specific tissues (Farlik et al., 2015) or to classify cell types and lineages (Luo et al., 2017; Mulqueen et al., 2018). Traditional single cell bisulfite workflows require an elevated number of PCR cycles to ensure adequate input material for sequencing, resulting in amplification bias and read duplication. One study examining the methylation status of highly abundant, repetitive elements (Kobayashi et al., 2016) removed PCR amplification, instead increasing starting material via sample multiplexing. However, this method is not generalizable to non-repetitive sequences. Alternatively, an RRBS-based method was used to preserve PCR amplification while incorporating unique molecular identifiers, which act as fragment-specific barcodes, to identify PCR-related duplications and SNP-independent allele-specific methylation (Wang et al., 2015). Finally, one group improved priming efficiency in PBAT, implemented additional quality control steps during data analysis, and found high CpG concordance at neighboring sites within a 1 kb window (Hui et al., 2018).

Although BS-seq is generally used to detect 5mC and 5hmC, this chemical treatment has been incorporated in a workflow to indiscriminately detect 5fC and 5caC at a genome-wide level (Wu et al., 2017). This approach – termed single cell methylase-assisted bisulfite sequencing (scMAB-seq, Fig. 5D) – uses the methyltransferase M.SssI to convert CpG cytosines to 5mC, resulting in only 5fC/5caC conversion to uracil upon bisulfite treatment. In order to improve material recovery, the authors successfully developed two workflows based on either PBAT or RRBS. Whereas the former captures more of the genome and enables more quantitative analysis, the RRBS-based scMAB-seq approach improves measurement consistency owing to the rareness of 5fC and 5caC and their tendency to cluster in the genome. The application of this approach to study the zygotic paternal genome following replication and cell division directly demonstrated for the first time in single cells that 5fC and 5caC are diluted during these processes.

Endonuclease-based measurements of DNA modifications

Endonuclease-based sequencing methods avoid random DNA loss and thus improve coverage and efficiency. Genome-wide CpG island (CGI) methylation sequencing for single cells (scCGI-seq) is one such method that enriches for sequences with high CpG content, providing 72.7% CGI coverage per cell (Han et al., 2017) (Fig. 5E). Following manual isolation of single cells, a cocktail of methylation-sensitive restriction endonucleases fragments DNA in unmethylated regions and preserves longer fragments in methylated regions (e.g. CGIs), allowing them to be amplified via multiple strand displacement. The use of a small pool of restriction enzymes with specific recognition sequences precludes this technique from directly providing single nucleotide information at most CpGs. Rather, this method relies on the observation that neighboring CpG sites within a 1 kb window have a high probability of sharing the same methylation state (Hui et al., 2018). Hierarchical clustering of CGI methylation revealed by scCGI-seq yielded distinct clusters according to cell type, with greater CGI/promoter hypermethylation and repeat region hypomethylation measured in the cancer cell line K562 compared with the non-cancer cell line GM12878 (Han et al.,

2017). This strategy could be extended to mapping 5hmC, although more 5hmC-sensitive restriction enzymes need to be identified to provide greater coverage.

A method for specifically detecting 5hmC (Mooijman et al., 2016), called single cell Aba-seq, first glucosylates 5hmC to 5ghmC for recognition by the restriction enzyme AbaSI, which cleaves the DNA downstream of the modification to generate fragments for ligation of cell-specific barcodes for NGS (Fig. 5F). This methodology suffers from a low detection efficiency of ~10% largely owing to incomplete glucosylation. Nevertheless, scAba-seq was used in mESCs and in developing mouse embryos to show that each chromosome in a particular cell has high 5hmC density on one DNA strand and low density on the other, suggesting the lack of a specific replication-dependent maintenance mechanism for 5hmC. Moreover, this 5hmC strand bias can be used to identify sister cells, as the sum of their strand biases must add to unity.

Although the increasing biological importance of 5hmC, 5fC and 5caC has spurred the development of single nucleotide-resolution detection methods at the bulk and single cell levels, there exists the need for a method to detect 5mC alone (without 5hmC) at similar resolution (Clark et al., 2017).

Third-generation sequencing of DNA methylation

Third-generation sequencing methods, such as single molecule real-time sequencing (SMRT, commercialized by PacBio) and Nanopore sequencing (commercialized by Oxford Nanopore Technologies), promise detection of DNA modifications directly, without amplification, on single DNA molecules (Fig. 5G,H). However, the distinction between single molecule detection versus single cell detection must be noted. The use of physical cell separation or sample barcoding enables methods such as scBS-seq to provide epigenetic information for many loci (thousands or more) within the same cell. By contrast, although single molecule information is inherently derived from single cells, without unique sample identifiers it only provides information from one genomic locus. Nonetheless, the advantage of both SMRT and Nanopore sequencing is that reads from single DNA molecules are long (up to 2.2 Mb for Nanopore sequencing; Payne et al., 2018), with read length limited by the method of DNA purification (Laver et al., 2015); thus, these methods can provide unprecedented information about a large genomic locus from a particular cell.

As with NGS, SMRT sequencing uses imaging of fluorescent nucleotides upon incorporation during DNA synthesis (Fig. 5G). Unlike NGS, which relies on local amplification of the template to increase the fluorescent signal, SMRT uses zero-mode waveguides (subwavelength holes in a metal film) to decrease the imaging volume, reducing background fluorescence from unincorporated nucleotides to enable single molecule detection of incorporation events (Levene et al., 2003). When the DNA template contains epigenetic modifications, the kinetics of nucleotide incorporation slows, allowing the detection of individual 6mA (adenosine methylated at carbon six), 5hmC or 5mC events on a synthetic template with known sequence (Flusberg et al., 2010). However, the change in signal – the time between two incorporation events – is low (requiring 500× coverage for differentiating 5hmC from 5mC; Suzuki et al., 2016), depends on the surrounding sequence, and is spread over the several bases that contact the DNA polymerase. With these limitations, SMRT-seq can only reliably detect 6mA (Flusberg et al., 2010) and not 5mC/5hmC (the most common mammalian DNA modifications) on genomic sequences.

Nanopore sequencing relies on changes in electrical current through a pore to read the sequence of long DNA molecules as well as their epigenetic modifications (Fig. 5H). Initial experiments using test DNA oligos of known sequence have shown that detection efficiency of 5mC and 5hmC can be as high as 80–97% with a single-pass read (Laszlo et al., 2013; Schreiber et al., 2013). However, the current depends on multiple nucleotides that are in the pore concurrently, rendering the magnitude of the signal dependent on the surrounding sequence. Thus, accurate detection on random genomic DNA requires training the detection algorithm on a library of synthetic oligos consisting of all possible sequence combinations for the number of bases that span the pore height and with different modifications on each cytosine. Although this feat has not yet been achieved, significant steps have been taken to this end by training hidden Markov models on smaller libraries and applying them to genomic DNA samples from bacteria (Rand et al., 2017) and human cells (Simpson et al., 2017). Indeed, this approach was used to analyze data from individual DNA molecules in cancer cells to show that the methylation status of CpGs is highly correlated within 1 kb regions from a single cell (Simpson et al., 2017), in agreement with previous results.

Real-time reporters of DNA methylation

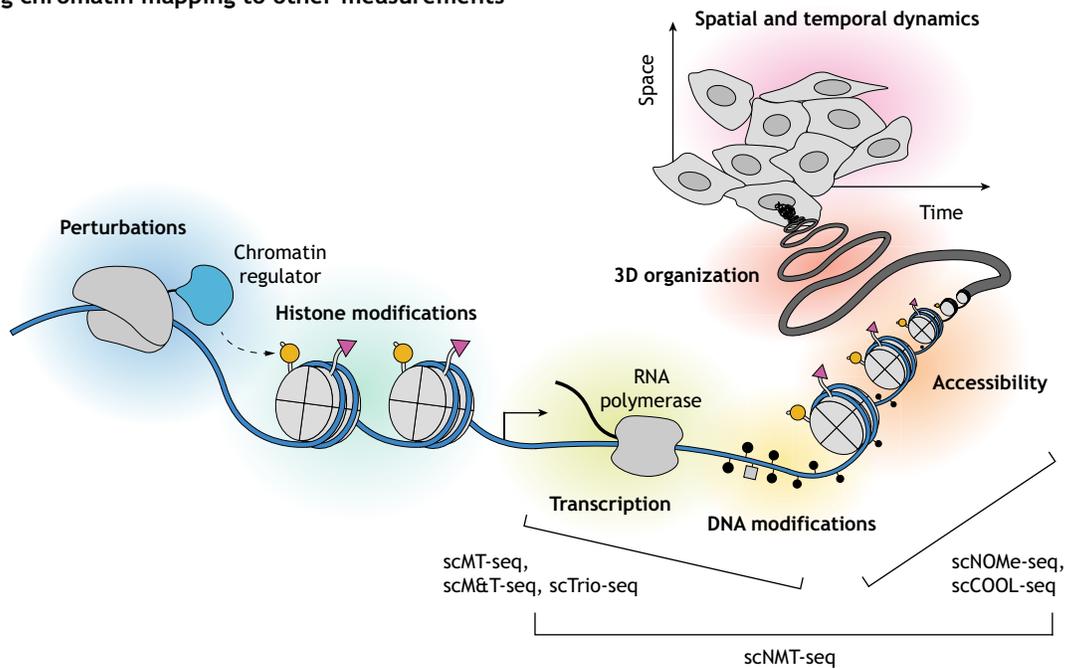
DNA modification detection methods are arguably most informative when paired with DNA sequencing, which optimally yields single nucleotide-resolution data. However, these approaches often capture a snapshot of the methylome, which is itself dynamic. A reporter of genomic methylation (RGM) sacrifices single nucleotide resolution in order to indirectly monitor real-time methylation state information at a specific locus (Stelzer et al., 2015) (Fig. 4C). The reporter consists of a minimal imprinted gene promoter, which is highly sensitive to proximal methylated DNA in a non-tissue-specific manner and drives the expression of a fluorescent protein. Stable CRISPR-mediated integration of this reporter near an endogenous region of interest enables the methylation state of that region to control expression of the fluorophore. Using this reporter, the dynamics of super enhancers were studied during mouse development and cell reprogramming. In the future, this method could be adapted to use different fluorescent proteins to simultaneously report the methylation status at multiple genes.

Recently developed real-time chromatin modification sensors (Lungu et al., 2017) directly detect a chromatin modification at a repetitive locus of interest using bimolecular fluorescence complementation (Fig. 4D), while simultaneously providing subnuclear localization. These sensors, termed bimolecular anchor detectors (BiADs), consist of a programmed DNA-binding domain ('anchor', e.g. zinc finger, TALE, dCas9) and a chromatin modification detector, each of which are fused to a portion of the monomeric Venus fluorescent protein, which only produces a signal upon proximity-mediated reconstitution. The authors of the above study validated their BiAD sensors through detection of 5mC and H3K9me3 at centromeric satellites, demonstrated BiAD accessibility to constitutive heterochromatin and condensed chromatin throughout mitosis, and suggested extending their method to simultaneous detection of multiple modifications.

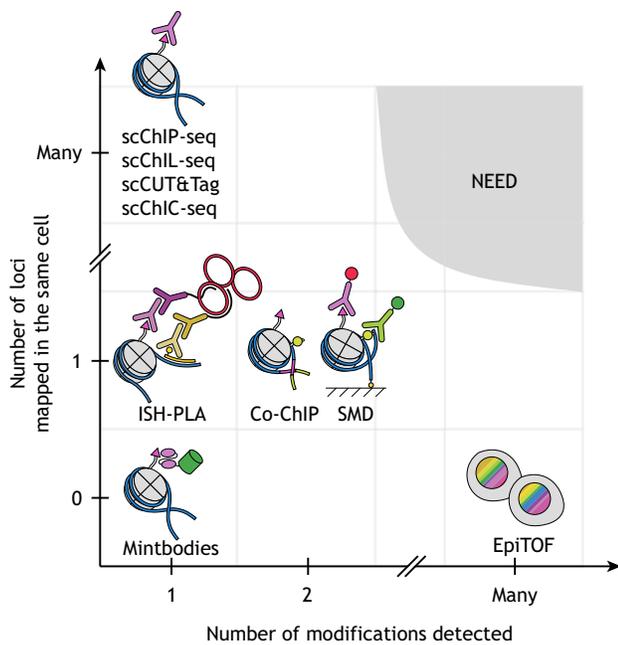
Linking chromatin modifications to other measurements

There has been a recent development of single cell methodologies that combine detection of chromatin modifications with measurements of gene expression and/or chromatin organization in order to elucidate the relationship between them (Fig. 6A).

A Linking chromatin mapping to other measurements



B Complexity of mapping histone modifications



C Complexity of mapping DNA modifications

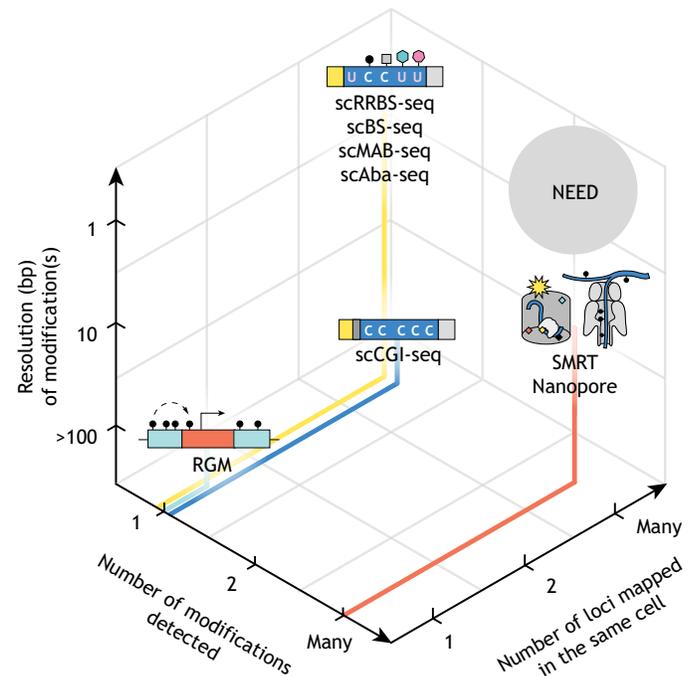


Fig. 6. Future developments in single cell detection of chromatin modifications. (A) An ideal method would monitor chromatin modifications in conjunction with other processes (e.g. accessibility, transcription, 3D organization) over time and space and in response to perturbations. A number of existing methods measure two or more cell characteristics, as indicated. (B) Single cell methods for detecting and mapping histone modifications are plotted according to two components of the complexity metric: the number of modifications detected (x-axis) and the number of loci mapped in a single cell (y-axis). (C) Single cell methods for detecting and mapping DNA modifications are plotted according to three components of the complexity metric: the number of modifications detected, the number of loci mapped in a single cell, and the mapping resolution (bp). The plots (B,C) highlight that there currently exists a need for methods that map many chromatin modifications at many loci.

DNA methylation and chromatin accessibility

Single cell nucleosome occupancy and methylome sequencing (scNOME-seq) provides chromatin accessibility and DNA methylation information, which, taken together, are particularly useful for characterizing regulatory regions (Pott, 2017). In this approach, a viral GpC methyltransferase, M.CviPI, is used to convert chromatin accessibility into a pattern of Gp5mC (Kelly et al., 2012), a type of methylation that is rare in eukaryotes. This enzyme preferentially methylates GpC cytosines in non-nucleosomal DNA when added to extracted nuclei. These methyltransferase-treated nuclei are deposited by fluorescence-activated cell sorting (FACS) into individual wells and subsequently subjected to the scBS-seq protocol, revealing both endogenous CpG methylation and nucleosome occupancy as a lack of GpC methylation. Using this approach, it was shown that promoter accessibility and gene-body CpG methylation are positively correlated with gene expression annotations, whereas promoter CpG methylation is negatively correlated with known expression, consistent with previous findings (Schübeler, 2015). Furthermore, both DNA methylation and accessibility could be used to independently classify cell types.

A nearly identical method, termed single cell chromatin overall omic-scale landscape sequencing (scCOOL-seq), has been used to provide information about nucleosome positioning, DNA methylation, copy number variation (CNV), and ploidy in mouse preimplantation embryos (Guo et al., 2017). The authors observed higher DNA methylation heterogeneity between embryos than within an embryo at the two-cell stage, demonstrating a high degree of synchronization during development. They also found that, from the late zygote to the four-cell stage, intergenic regions on paternal alleles are more methylated relative to maternal ones, whereas intragenic regions are less methylated. In addition, their analysis of mESCs revealed that promoters with increased heterogeneity in accessibility are generally hypomethylated, although the factors responsible for this phenomenon remain to be discovered.

DNA methylation and gene expression

At present, there are three methodologies that provide DNA methylation and gene expression information at the single cell level. Two approaches are both named single cell methylome and transcriptome-sequencing [scM&T-seq (Angermueller et al., 2016) and scMT-seq (Hu et al., 2016)], and the third, which also looks at DNA copy number variation, is called single cell triple omics sequencing (scTrio-seq) (Hou et al., 2016). The first approach, scM&T-seq, separates polyadenylated RNA from DNA with biotinylated oligo-dT primers for RNA-seq (Macaulay et al., 2015) and performs whole-genome methylome sequencing with scBS-seq. The second approach, scMT-seq, uses microcapillary pipetting to isolate the nucleus from the cytosol and enriches for CpG-dense regions with scRRBS-seq. Both techniques implement a widely used scRNA-seq protocol known as Smart-seq2 (Picelli et al., 2014) for transcriptome library preparation. The third approach, scTrio-seq, uses gentle lysis and centrifugation to physically separate cytosolic contents from the nucleus, thereby minimizing DNA contamination during RNA-seq library preparation, and also uses the scRRBS-seq method for methylome profiling.

scM&T-seq has been used to recapitulate the increased epigenetic and gene expression heterogeneity observed when ESCs are grown in serum versus 2i media (Angermueller et al., 2016). Importantly, this study showed that distinct clusters of the most variable genes arise by either methylome- or transcriptome-

based hierarchical clustering, supporting the utility of the dual methodology. Moreover, the data quality and results of using the combined technique are comparable to the data quality and results of performing scRNA-seq or scBS-seq individually. Because scM&T-seq utilizes the whole-genome bisulfite approach, 48.4% of CpG sites can theoretically be identified, in contrast to the 1% of sites that the reduced representation approach employs in both scMT-seq and scTrio-seq (Hu et al., 2016).

scMT-seq has been used in dorsal root ganglion sensory neurons, the large size of which facilitates micropipetting, to link methylation patterns with allelic expression (Hu et al., 2016). Specifically, this study revealed that transcription of non-CGI promoter genes negatively correlates with promoter methylation, whereas transcription of CGI promoter genes positively correlates with gene body methylation. The scMT-seq method additionally revealed a rare subset of highly expressed genes with hypermethylated, CpG-rich promoters.

scTrio-seq recapitulates correlations between gene expression and promoter hypomethylation and gene-body hypermethylation, and provides nearly complete single cell copy-number variation determination with 10-Mb resolution; this method allows classification of cancer cell subpopulations with high metastatic potential that bulk assays would fail to capture (Hou et al., 2016). However, scTrio-seq provides only approximately two-thirds the transcriptome coverage of scMT-seq (Hu et al., 2016), likely owing to the incomplete transfer of cytosolic material during the physical separation procedure. Nonetheless, this procedure may more flexibly accommodate multiple, downstream transcriptomic and methylomic pipelines, including whole genome scBS-seq (Hou et al., 2016).

DNA methylation, chromatin accessibility, and gene expression

Single cell nucleosome, methylation, and transcription sequencing (scNMT-seq) is the natural successor to the above methods, combining elements of the scM&T-seq method with the scNOME-seq workflow (Clark et al., 2018). Using this technique, the number and nature of associations between pairs of these three layers of molecular regulation in embryoid bodies was examined and revealed the following: (1) CpG methylation and transcription are generally negatively correlated at the genome-wide level; (2) accessibility and transcription are positively correlated in promoters, p300 targets, and super enhancers; and (3) CpG methylation and accessibility are generally negatively correlated in introns. Clustering of single cell accessibility profiles provided an estimate of accessibility heterogeneity for each gene, with those that were homogeneously open being highly expressed housekeeping genes and those with a mixed state being lowly expressed and more likely to contain a bivalent promoter. Finally, the authors used gene expression data to infer the position of each cell along a differentiation trajectory to examine accessibility and methylation dynamics as functions of pseudotime. As expected, methylation and accessibility become increasingly negatively correlated throughout differentiation.

Applications of single cell chromatin mapping technologies: insights into development and disease

Single cell chromatin modification mapping technologies have deepened our understanding of cell types and states as well as how these change throughout development and disease. Cell-type or cell-state classifications have both been used to benchmark these technologies and as a tool to study cell-fate decisions in development, aging and disease. Benchmarking has been performed in a few ways: (1) by mixing known cell types together

and showing that they can be computationally separated; (2) by using haploid cells with homogeneous chromatin modification profiles, such as oocytes, and showing that single cell data reproduce bulk results; or (3) by looking for increased heterogeneity in mESCs grown in serum versus 2i medium. Single cell detection of chromatin modifications has been used to classify cell types in the early embryo (Guo et al., 2013; Guo et al., 2017), smooth muscle (Gomez et al., 2013), cortex (Luo et al., 2017), the hematopoietic system (Cheung et al., 2018; Farlik et al., 2016; Hui et al., 2018) and the immune system (Cheung et al., 2018). Additionally, single cell measurements have shown that immune cells exhibit increased overall levels and heterogeneity of histone modifications with aging (Cheung et al., 2018). Similarly, hepatic cancer samples show heterogeneity in DNA methylation and contain a subpopulation of cells that expresses markers associated with immune evasion and malignancy (Hou et al., 2016). These results underscore the value of mapping chromatin modifications for diagnosis.

Classification of cell types can also be achieved with other single cell methods that rely on detection of transcripts or chromatin accessibility (Trapnell, 2015). scRNA-seq is currently the most widely used method for cell-type classification (Tanay and Regev, 2017), as gene expression is a good indicator of cell type, and one can recover a fair number of RNA molecules from each cell. However, because transcription occurs in bursts and is dynamic over short timescales, scRNA-seq data has intrinsic noise. Therefore, many cells are needed to extract trends, and care should be taken when interpreting these data as cell types rather than signaling states within a cell type. Given these dynamics in RNA levels, chromatin state might be a more stable signature of cell type. One method to globally assess chromatin state is the assay for transposase-accessible chromatin (ATAC-seq) (Buenrostro et al., 2013), which has been adapted to the single cell level (Buenrostro et al., 2015; Cusanovich et al., 2015) and applied to classify cell types (for a review, see Klemm et al., 2019). A recent study that built an immune cell atlas has shown that analysis of chromatin accessibility by single cell ATAC-seq at distal enhancers results in sharper cell classification than analysis based on RNA expression or accessibility of transcription start sites (Yoshida et al., 2019). Similarly, analysis of chromatin modifications at select genomic elements could yield improved cell-type classification once the number of reads recovered from each cell increases sufficiently.

Single cell detection of chromatin modifications can also be used for lineage-tracking during development and disease. Tracking relies either on strong maintenance of histone or DNA methylation at particular loci throughout cell generations or on the complete lack of maintenance of modifications, such as 5hmC, that leads to their strand-specific partitioning in daughter cells. 5mC is a stable modification, owing to its replication-dependent maintenance by DNMT1, and thus can be used to track lineages across generations (Farlik et al., 2016). Histone methylation can also persist for many generations in certain cell types. For instance, H3K4me2 is detected at tissue-specific promoters of smooth muscle cells even when these cells have differentiated into atherosclerotic tissue and no longer express those specific genes (Gomez et al., 2013). Alternatively, modifications that lack strong maintenance, such as 5hmC in early embryo development, can be used to identify sister cells, as the 5hmC signal from the original DNA is partitioned perfectly between sister cells upon DNA replication (Mooijman et al., 2016). Lineage tracking, either by monitoring naturally occurring chromatin modifications or by introducing orthogonal modifications from other organisms, can be used to interrogate mechanisms in development and to improve diagnosis.

Single cell measurements have also confirmed key results about histone modification combinations and DNA modification dynamics that were implied from bulk data and have started to unravel details that would otherwise be obscured by the stochastic nature of epigenetic events. In particular, modifications that are associated with opposing gene expression patterns (H3K4me3 and H3K9me3) were definitively shown to exist on the same nucleosomes (bivalency) in mESCs, as well as other cell types (Rotem et al., 2015; Weiner et al., 2016), as opposed to resulting from mixed cell populations. Histone modifications associated with synergistic effects on gene expression (H3K9ac and H3K4me3) have also been identified on the same nucleosomes, and their co-existence is enriched compared with levels that would be expected from measurements of each modification independently, suggesting positive feedback between these two modifications (Weiner et al., 2016). Moreover, multiple studies have uncovered an increased degree of heterogeneity in histone modifications and DNA methylation at enhancers in hepatocytes (Gravina et al., 2015; Gravina et al., 2016) and in mESCs primed for differentiation (Angermueller et al., 2016; Smallwood et al., 2014; Weiner et al., 2016), suggesting that changes in chromatin modifications could be drivers of cell differentiation.

Conclusions and perspectives

The rapidly increasing pace of method development in mapping single cell chromatin modifications reflects the appreciation of intrinsic stochasticity and heterogeneity in biology. This fast pace has been supported by advancements in microfluidics, automation, single molecule imaging, DNA synthesis, and sequencing. The major DNA and histone modifications have now been measured at the single cell level in multiple ways. So far, research on single cell detection of chromatin modifications has primarily focused on the development and validation of these techniques. This focus on technological development is justified, as single cell technologies at present have quite low complexity [low number of modifications and loci mapped per cell (Fig. 6B,C) and limited spatial and temporal resolution]. Owing to these limitations, genome-wide single cell technologies are currently much better at classifying cell types and states than at answering mechanistic questions. At the other extreme, measurements that follow a single locus over time, such as real-time reporters (Bintu et al., 2016; Stasevich et al., 2014; Stelzer et al., 2015), can shed some mechanistic insight on the role of chromatin in gene expression. However, it remains to be seen how these findings extend to different loci and cell types. In particular, different cell types grown in culture appear to be more similar in their epigenetic profiles to one another than to cells from within the tissue from which they were derived (Zhu et al., 2013). This finding suggests that using bulk data from *in vitro* cell culture to guide the analysis of sparse reads from single cell measurements performed on primary or clinical samples should be performed with care, as it may lead to misclassification.

We foresee that, as they mature and increase in complexity, single cell technologies could be used to detect multiple modifications, including combined DNA and histone modifications (e.g. bulk ChIP-BS-seq; Gao et al., 2016) at many loci, and they could additionally measure gene expression and chromatin structure over time and space (Fig. 6A). One way to achieve this ideal is by using recently developed multiplexed imaging technologies – such as Oligopaints for chromatin (Beliveau et al., 2015; Bintu et al., 2018; Boettiger et al., 2016) and MERFISH or intron seqFISH for RNA (Chen et al., 2015; Shah et al., 2018) – as endpoint measurements in time-lapse microscopy experiments. Finally, with the development of programmable DNA-binding domains, we can now perturb the epigenome at will (Cano-Rodriguez and Rots, 2016; Park et al.,

Box 2. Single cell technologies in global scientific initiatives

A number of ongoing global scientific efforts will benefit from advances in single cell technologies. These include the Human Cell Atlas (HCA; <https://www.humancellatlas.org/>), the Encyclopedia of DNA Elements (ENCODE; <https://www.encodeproject.org/>), Roadmap Epigenomics (<http://www.roadmapepigenomics.org/>), the 4D Nucleome (<https://www.4dnucleome.org/>) and the International Human Epigenome Consortium (IHEC; <http://ihec-epigenomes.org/>). Established in 2016, the HCA is an international scientific community that aims to characterize every cell type in the human body in terms of its DNA sequence, chromatin modifications, and RNA and protein expression. These efforts were catalyzed by the fact that single cell technologies are now ready to be deployed for cell-type classification and are ripe for high-throughput development. Mapping all healthy human cells at the single cell level in the body is an essential first step in understanding how these cells interact in space and time and will also illuminate how changes come about in disease. ENCODE, Roadmap Epigenomics, and the 4D Nucleome are collaborations between research groups that started in 2003, 2008 and 2015, respectively, and are funded by the National Institutes of Health. These groups contribute data and collaborate with the broader international community as part of the IHEC, which started in 2010. These international collections of research groups collectively focus on mapping the connections between DNA sequence, chromatin modifications, genome organization and gene expression.

2016; Thakore et al., 2016; Vora et al., 2016) and actually determine causation rather than correlation between chromatin and gene regulation. These improvements will increase our mechanistic understanding of chromatin and its connections with cell identity (Tanay and Regev, 2017).

Technological advances associated with single cell technologies can also be used to improve traditional bulk methods for detecting chromatin modifications. Protocol modifications that combine all reactions in one small volume can reduce the number of cells necessary for bulk analysis. For instance, CUT&RUN (Skene et al., 2018), a protocol that performs targeted chromatin digestion inside nuclei using MNase linked to an antibody against a specific histone modification, can be used with hundreds of cells instead of the millions required for classical ChIP. Additionally, barcoding can be used to index different samples instead of single cells, thus allowing robust, high-throughput analysis of many small samples, as demonstrated in the indexing-first chromatin IP approach (iChIP) (Lara-Astiaso et al., 2014). Finally, extensive panels of antibodies against histone modifications have been validated (Cheung et al., 2018; Weiner et al., 2016) and new, genetically encoded ones (mintbodies) have been developed (Hayashi-Takanaka et al., 2011; Sato et al., 2013; Sato et al., 2016). These well-characterized antibodies should be used for bulk assays, and newly developed antibodies should be benchmarked in a similar manner. Taken together, these improvements will enable the processing and analysis of smaller samples, allowing us to reduce costs, improve diagnosis, and answer biological questions using primary cells.

The continued development and dissemination of single cell technologies for mapping chromatin modifications will also be essential for global scientific efforts (see Box 2). Although these projects are currently dominated by bulk data, the importance of single cell measurements in the context of cellular heterogeneity is coming into focus (Dekker et al., 2017) and is likely to increase as these technologies improve throughput and portability. These advances in single cell technologies will greatly enhance our understanding of mammalian cell biology and help us to develop cell-based diagnostics and therapeutics.

Acknowledgements

We thank members of the Bintu lab for providing feedback on the manuscript.

Competing interests

The authors declare no competing or financial interests.

Funding

This work was supported by the National Institutes of Health (MIRA R35GM128947 to L.B.) and Burroughs Wellcome Fund (Career at the Scientific Interface Award to L.B.). Deposited in PMC for release after 12 months.

References

- Allis, C. D. and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* **17**, 487-500. doi:10.1038/nrg.2016.59
- Angermueller, C., Clark, S. J., Lee, H. J., Macaulay, I. C., Teng, M. J., Hu, T. X., Krueger, F., Smallwood, S. A., Ponting, C. P., Voet, T. et al. (2016). Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat. Methods* **13**, 229-232. doi:10.1038/nmeth.3728
- Bandura, D. R., Baranov, V. I., Ornatsky, O. I., Antonov, A., Kinach, R., Lou, X., Pavlov, S., Vorobiev, S., Dick, J. E. and Tanner, S. D. (2009). Mass cytometry: Technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Anal. Chem.* **81**, 6813-6822. doi:10.1021/ac901049w
- Beliveau, B. J., Boettiger, A. N., Avendaño, M. S., Jungmann, R., McCole, R. B., Joyce, E. F., Kim-Kiselak, C., Bantignies, F., Fonseka, C. Y., Erceg, J. et al. (2015). Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes. *Nat. Commun.* **6**, 7147. doi:10.1038/ncomms8147
- 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. doi:10.1016/j.cell.2006.02.041
- Bintu, L., Yong, J., Antebi, Y. E., McCue, K., Kazuki, Y., Uno, N., Oshimura, M. and Elowitz, M. B. (2016). Dynamics of epigenetic regulation at the single-cell level. *Science* **351**, 720-724. doi:10.1126/science.aab2956
- Bintu, B., Mateo, L. J., Su, J.-H., Sinnott-Armstrong, N. A., Parker, M., Kinrot, S., Yamaya, K., Boettiger, A. N. and Zhuang, X. (2018). Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* **362**, eaau1783. doi:10.1126/science.aau1783
- Boettiger, A. N., Bintu, B., Moffitt, J. R., Wang, S., Beliveau, B. J., Fudenberg, G., Imakaev, M., Mirny, L. A., Wu, C. T. and Zhuang, X. (2016). Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418-422. doi:10.1038/nature16496
- Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. and Greenleaf, W. J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213-1218. doi:10.1038/nmeth.2688
- 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. doi:10.1038/nature14590
- Callinan, P. A. and Feinberg, A. P. (2006). The emerging science of epigenomics. *Hum. Mol. Genet.* **15** Suppl. 1, RR95-R101. doi:10.1093/hmg/ddl095
- Cano-Rodriguez, D. and Rots, M. G. (2016). Epigenetic editing: on the verge of reprogramming gene expression at will. *Curr. Genet. Med. Rep.* **4**, 170-179. doi:10.1007/s40142-016-0104-3
- Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. and Zhuang, X. (2015). Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090. doi:10.1126/science.aaa6090
- Cheung, P., Vallania, F., Warsinske, H. C., Donato, M., Schaffert, S., Chang, S. E., Dvorak, M., Dekker, C. L., Davis, M. M., Utz, P. J. et al. (2018). Single-cell chromatin modification profiling reveals increased epigenetic variations with aging. *Cell* **173**, 1385-1397.e14. doi:10.1016/j.cell.2018.03.079
- Clark, S. J., Smallwood, S. A., Lee, H. J., Krueger, F., Reik, W. and Kelsey, G. (2017). Genome-wide base-resolution mapping of DNA methylation in single cells using single-cell bisulfite sequencing (scBS-seq). *Nat. Protoc.* **12**, 534-547. doi:10.1038/nprot.2016.187
- Clark, S. J., Argelaguet, R., Kapourani, C. A., Stubbs, T. M., Lee, H. J., Alda-Catalinas, C., Krueger, F., Sanguinetti, G., Kelsey, G., Marioni, J. C. et al. (2018). scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat. Commun.* **9**, 781. doi:10.1038/s41467-017-02088-w
- Cusanovich, D. A., Daza, R., Adey, A., Pliner, H. A., Christiansen, L., Gunderson, K. L., Steemers, F. J., Trapnell, C. and Shendure, J. (2015). Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* **348**, 910-914. doi:10.1126/science.aab1601
- Dekker, J., Belmont, A. S., Guttman, M., Leshyk, V. O., Lis, J. T., Lomvardas, S., Mirny, L. A., O'Shea, C. C., Park, P. J., Ren, B. et al. (2017). The 4D nucleome project. *Nature* **549**, 219-226. doi:10.1038/nature23884

- Delachat, A. M.-F., Guidotti, N., Bachmann, A. L., Meireles-Filho, A. C. A., Pick, H., Lechner, C. C., Deluz, C., Deplancke, B., Suter, D. M. and Fierz, B. (2018). Engineered multivalent sensors to detect coexisting histone modifications in living stem cells. *Cell Chem. Biol.* **25**, 51-56.e6. doi:10.1016/j.chembiol.2017.10.008
- Ehrlich, M., Gama-Sosa, M. A., Huang, L.-H., Midgett, R. M., Kuo, K. C., McCune, R. A. and Gehrke, C. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells. *Nucleic Acids Res.* **10**, 2709-2721. doi:10.1093/nar/10.8.2709
- Farlik, M., Sheffield, N. C., Nuzzo, A., Datlinger, P., Schönegger, A., Klughammer, J. and Bock, C. (2015). Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics. *Cell Rep.* **10**, 1386-1397. doi:10.1016/j.celrep.2015.02.001
- Farlik, M., Halbritter, F., Müller, F., Choudry, F. A., Ebert, P., Klughammer, J., Farrow, S., Santoro, A., Ciauro, V., Mathur, A. et al. (2016). DNA methylation dynamics of human hematopoietic stem cell differentiation. *Cell Stem Cell* **19**, 808-822. doi:10.1016/j.stem.2016.10.019
- Flusberg, B. A., Webster, D. R., Lee, J. H., Travers, K. J., Olivares, E. C., Clark, T. A., Korfach, J. and Turner, S. W. (2010). Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat. Methods* **7**, 461-465. doi:10.1038/nmeth.1459
- Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., Molloy, P. L. and Paul, C. L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci.* **89**, 1827-1831. doi:10.1073/pnas.89.5.1827
- Gao, Y., Xiong, X., Wong, S., Charles, E. J., Lim, W. A. and Qi, L. S. (2016). Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. *Nat. Methods* **13**, 1043-1049. doi:10.1038/nmeth.4042
- Gomez, D., Shankman, L. S., Nguyen, A. T. and Owens, G. K. (2013). Detection of histone modifications at specific gene loci in single cells in histological sections. *Nat. Methods* **10**, 171-177. doi:10.1038/nmeth.2332
- Gravina, S., Ganapathi, S. and Vijg, J. (2015). Single-cell, locus-specific bisulfite sequencing (SLBS) for direct detection of epimutations in DNA methylation patterns. *Nucleic Acids Res.* **43**, e93-e93. doi:10.1093/nar/gkv366
- Gravina, S., Dong, X., Yu, B. and Vijg, J. (2016). Single-cell genome-wide bisulfite sequencing uncovers extensive heterogeneity in the mouse liver methylome. *Genome Biol.* **17**, 150. doi:10.1186/s13059-016-1011-3
- Grunau, C., Clark, S. J. and Rosenthal, A. (2001). Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res.* **29**, e65e. doi:10.1093/nar/29.13.e65
- Gu, H., Smith, Z. D., Bock, C., Boyle, P., Gnirke, A. and Meissner, A. (2011). Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat. Protoc.* **6**, 468-481. doi:10.1038/nprot.2010.190
- Guo, H., Zhu, P., Wu, X., Li, X., Wen, L. and Tang, F. (2013). Single-Cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res.* **23**, 2126-2135. doi:10.1101/gr.161679.113
- Guo, F., Li, L., Li, J., Wu, X., Hu, B., Zhu, P., Wen, L. and Tang, F. (2017). Single-cell multi-omics sequencing of mouse early embryos and embryonic stem cells. *Cell Res.* **27**, 967-988. doi:10.1038/cr.2017.82
- Han, L., Wu, H. J., Zhu, H., Kim, K. Y., Marjani, S. L., Riemer, M., Euskirchen, G., Zi, X., Yang, J., Han, J. et al. (2017). Bisulfite-independent analysis of CpG island methylation enables genome-scale stratification of single cells. *Nucleic Acids Res.* **45**, e77. doi:10.1093/nar/gkx026
- 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. doi:10.1038/s41556-018-0248-3
- Hayashi-Takanaka, Y., Yamagata, K., Wakayama, T., Stasevich, T. J., Kainuma, T., Tsurimoto, T., Tachibana, M., Shinkai, Y., Kurumizaka, H., Nozaki, N. et al. (2011). Tracking epigenetic histone modifications in single cells using Fab-based live endogenous modification labeling. *Nucleic Acids Res.* **39**, 6475-6488. doi:10.1093/nar/gkr343
- Henikoff, J. G., Belsky, J. A., Krassovsky, K., MacAlpine, D. M. and Henikoff, S. (2011). Epigenome characterization at single base-pair resolution. *Proc. Natl. Acad. Sci. USA* **108**, 18318-18323. doi:10.1073/pnas.1110731108
- Hou, Y., Guo, H., Cao, C., Li, X., Hu, B., Zhu, P., Wu, X., Wen, L., Tang, F., Huang, Y. et al. (2016). Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res.* **26**, 304-319. doi:10.1038/cr.2016.23
- Hu, Y., Huang, K., An, Q., Du, G., Hu, G., Xue, J., Zhu, X., Wang, C. Y., Xue, Z. and Fan, G. (2016). Simultaneous profiling of transcriptome and DNA methylome from a single cell. *Genome Biol.* **17**, 88. doi:10.1186/s13059-016-0950-z
- Huang, Y., Pastor, W. A., Shen, Y., Tahiliani, M., Liu, D. R. and Rao, A. (2010). The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One* **5**, e8888. doi:10.1371/journal.pone.0008888
- Huang, H., Sabari, B. R., Garcia, B. A., David Allis, C. and Zhao, Y. (2014). SnapShot: histone modifications. *Cell* **159**, 458-458.e1. doi:10.1016/j.cell.2014.09.037
- Hui, T., Cao, Q., Wegrzyn-Woltosz, J., O'Neill, K., Hammond, C. A., Knapp, D. J. H. F., Laks, E., Moksa, M., Aparicio, S., Eaves, C. J. et al. (2018). High-resolution single-cell DNA methylation measurements reveal epigenetically distinct hematopoietic stem cell subpopulations. *Stem Cell Rep.* **11**, 578-592. doi:10.1016/j.stemcr.2018.07.003
- Jelinek, J. and Madzo, J. (2016). DREAM: a simple method for DNA methylation profiling by high-throughput sequencing. In *Chronic Myeloid Leukemia: Methods and Protocols* (ed. S. Li and H. Zhang), pp. 111-127. New York, NY: Springer New York.
- Jones, P. L., Jan Veenstra, G. C., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J. and Wolffe, A. P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.* **19**, 187-191. doi:10.1038/561
- Karemaker, I. D. and Vermeulen, M. (2018). Single-cell DNA methylation profiling: technologies and biological applications. *Trends Biotechnol.* **36**, 952-965. doi:10.1016/j.tibtech.2018.04.002
- Kasinathan, S., Orsi, G. A., Zentner, G. E., Ahmad, K. and Henikoff, S. (2014). High-resolution mapping of transcription factor binding sites on native chromatin. *Nat. Methods* **11**, 203-209. doi:10.1038/nmeth.2766
- 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. doi:10.1038/s41467-019-09982-5
- Kelly, T. K., Liu, Y., Lay, F. D., Liang, G., Berman, B. P. and Jones, P. A. (2012). Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. *Genome Res.* **22**, 2497-2506. doi:10.1101/gr.143008.112
- Klein, A. M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D. A. and Kirschner, M. W. (2015). Droplet Barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187-1201. doi:10.1016/j.cell.2015.04.044
- Klemm, S. L., Shipony, Z. and Greenleaf, W. J. (2019). Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* **20**, 207-220. doi:10.1038/s41576-018-0089-8
- Kobayashi, H., Koike, T., Sakashita, A., Tanaka, K., Kumamoto, S. and Kono, T. (2016). Repetitive DNA methylome analysis by small-scale and single-cell shotgun bisulfite sequencing. *Genes Cells* **21**, 1209-1222. doi:10.1111/gtc.12440
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* **128**, 693-705. doi:10.1016/j.cell.2007.02.005
- 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. doi:10.1038/s41592-019-0361-7
- Kungulovski, G., Kycia, I., Tamas, R., Jurkowska, R. Z., Kudithipudi, S., Henry, C., Reinhardt, R., Labhart, P. and Jeltsch, A. (2014). Application of histone modification-specific interaction domains as an alternative to antibodies. *Genome Res.* **24**, 1842-1853. doi:10.1101/gr.170985.113
- Lara-Astiaso, D., Weiner, A., Lorenzo-Vivas, E., Zaretzky, I., Jaitin, D. A., David, E., Keren-Shaul, H., Mildner, A., Winter, D., Jung, S. et al. (2014). Immunogenetics. Chromatin state dynamics during blood formation. *Science* **345**, 943-949. doi:10.1126/science.1256271
- Laszlo, A. H., Derrington, I. M., Brinkerhoff, H., Langford, K. W., Nova, I. C., Samson, J. M., Bartlett, J. J., Pavlenok, M. and Gundlach, J. H. (2013). Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA. *Proc. Natl. Acad. Sci.* **110**, 18904-18909. doi:10.1073/pnas.1310240110
- Laver, T., Harrison, J., O'Neill, P. A., Moore, K., Farbos, A., Paszkiewicz, K. and Studholme, D. J. (2015). Assessing the performance of the Oxford Nanopore Technologies MinION. *Biomol. Detect. Quantif.* **3**, 1-8. doi:10.1016/j.bdq.2015.02.001
- Lawrence, M., Daujat, S. and Schneider, R. (2016). Lateral thinking: how histone modifications regulate gene expression. *Trends Genet.* **32**, 42-56. doi:10.1016/j.tig.2015.10.007
- Levene, H. J., Korfach, J., Turner, S. W., Foquet, M., Craighead, H. G. and Webb, W. W. (2003). Zero-mode waveguides for single-molecule analysis at high concentrations. *Science* **299**, 682-686. doi:10.1126/science.1079700
- Lungu, C., Pinter, S., Broche, J., Rathert, P. and Jeltsch, A. (2017). Modular fluorescence complementation sensors for live cell detection of epigenetic signals at endogenous genomic sites. *Nat. Commun.* **8**, 649. doi:10.1038/s41467-017-00457-z
- Luo, C., Keown, C. L., Kurihara, L., Zhou, J., He, Y., Li, J., Castanon, R., Lucero, J., Nery, J. R., Sandoval, J. P. et al. (2017). Elements in mammalian cortex. *Science* **604**, 600-604. doi:10.1126/science.aan3351
- Ma, S., De La Fuente Revenga, M., Sun, Z., Sun, C., Murphy, T. W., Xie, H., González-Maeso, J. and Lu, C. (2018). Cell-type-specific brain methylomes profiled via ultralow-input microfluidics. *Nat. Biomed. Eng.* **2**, 183-194. doi:10.1038/s41551-018-0204-3
- Macaulay, I. C., Haerty, W., Kumar, P., Li, Y. I., Hu, T. X., Teng, M. J., Goolam, M., Saurat, N., Coupland, P., Shirley, L. M. et al. (2015). G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat. Methods* **12**, 519-522. doi:10.1038/nmeth.3370

- Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M. et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202-1214. doi:10.1016/j.cell.2015.05.002
- Martens, J. H. A. and Stunnenberg, H. G. (2013). BLUEPRINT: Mapping human blood cell epigenomes. *Haematologica* **98**, 1487-1489. doi:10.3324/haematol.2013.094243
- Meissner, A., Gnirke, A., Bell, G. W., Ramsahoye, B., Lander, E. S. and Jaenisch, R. (2005). Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* **33**, 5868-5877. doi:10.1093/nar/gki901
- Miura, F., Enomoto, Y., Dairiki, R. and Ito, T. (2012). Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. *Nucleic Acids Res.* **40**, e136. doi:10.1093/nar/gks454
- Mooijman, D., Dey, S. S., Boisset, J.-C., Crosetto, N. and Van Oudenaarden, A. (2016). Single-cell 5 hmC sequencing reveals chromosome-wide cell-to-cell variability and enables lineage reconstruction. *Nat. Biotechnol.* **34**, 852-856. doi:10.1038/nbt.3598
- Mulqueen, R. M., Pokholok, D., Norberg, S. J., Torkenczy, K. A., Fields, A. J., Sun, D., Sinnamon, J. R., Shendure, J., Trapnell, C., O'Roak, B. J. et al. (2018). Highly scalable generation of DNA methylation profiles in single cells. *Nat. Biotechnol.* **36**, 428-431. doi:10.1038/nbt.4112
- Neri, F., Rapelli, S., Krepelov, A., Incarnato, D., Parlato, C., Basile, G., Maldotti, M., Anselmi, F. and Oliviero, S. (2017). Intragenic DNA methylation prevents spurious transcription initiation. *Nature* **543**, 72-77. doi:10.1038/nature21373
- Park, M., Keung, A. J. and Khalil, A. S. (2016). The epigenome: the next substrate for engineering. *Genome Biol.* **17**, 183. doi:10.1186/s13059-016-1046-5
- Payne, A., Holmes, N., Rakyán, V. and Loose, M. (2018). BulkVis: a graphical viewer for Oxford Nanopore bulk FAST5 files. *Bioinformatics* **34**, e1841. doi:10.1093/bioinformatics/bty841
- Picelli, S., Faridani, O. R., Björklund, A. K., Winberg, G., Sagasser, S. and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171-181. doi:10.1038/nprot.2014.006
- Pott, S. (2017). Simultaneous measurement of chromatin accessibility, DNA methylation, and nucleosome phasing in single cells. *Elife* **6**, e23203. doi:10.7554/eLife.23203
- Rand, A. C., Jain, M., Eizenga, J. M., Musselman-Brown, A., Olsen, H. E., Akeson, M. and Paten, B. (2017). Mapping DNA methylation with high-throughput nanopore sequencing. *Nat. Methods* **14**, 411-413. doi:10.1038/nmeth.4189
- Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J. et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317-330. doi:10.1038/nature14248
- Rotem, A., Ram, O., Shores, 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. doi:10.1038/nbt.3383
- Sato, Y., Mukai, M., Ueda, J., Muraki, M., Stasevich, T. J., Horikoshi, N., Kujirai, T., Kita, H., Kimura, T., Hira, S. et al. (2013). Genetically encoded system to track histone modification in vivo. *Sci. Rep.* **3**, 2436. doi:10.1038/srep02436
- Sato, Y., Kujirai, T., Arai, R., Asakawa, H., Ohtsuki, C., Horikoshi, N., Yamagata, K., Ueda, J., Nagase, T., Haraguchi, T. et al. (2016). A genetically encoded probe for live-cell imaging of H4K20 monomethylation. *J. Mol. Biol.* **428**, 3885-3902. doi:10.1016/j.jmb.2016.08.010
- Schreiber, J., Wescoe, Z. L., Abu-Shumays, R., Vivian, J. T., Baatar, B., Karplus, K. and Akeson, M. (2013). Error rates for nanopore discrimination among cytosine, methylcytosine, and hydroxymethylcytosine along individual DNA strands. *Proc. Natl. Acad. Sci.* **110**, 18910-18915. doi:10.1073/pnas.1310615110
- Schübeler, D. (2015). Function and information content of DNA methylation. *Nature* **517**, 321-326. doi:10.1038/nature14192
- Shah, S., Takei, Y., Zhou, W., Lubeck, E., Yun, J., Eng, C.-H. L., Koulana, N., Cronin, C., Karp, C., Liaw, E. J. et al. (2018). Dynamics and spatial genomics of the nascent transcriptome by intron seqFISH. *Cell* **174**, 363-376.e16. doi:10.1016/j.cell.2018.05.035
- Shapiro, R., DiFate, V. and Welcher, M. (1974). Deamination cytosine derivatives by bisulfite. mechanism of the reaction. *J. Am. Chem. Soc.* **96**, 906-912. doi:10.1021/ja00810a043
- Shema, E., Jones, D., Shores, N., Donohue, L., Ram, O. and Bernstein, B. E. (2016). Single-molecule decoding of combinatorially modified nucleosomes. *Science* **352**, 717-721. doi:10.1126/science.aad7701
- Shema, E., Bernstein, B. E. and Buenrostro, J. D. (2018). Single-cell and single-molecule epigenomics to uncover genome regulation at unprecedented resolution. *Nat. Genet.* **51**, 19-25. doi:10.1038/s41588-018-0290-x
- Simpson, J. T., Workman, R. E., Zuzarte, P. C., David, M., Dursi, L. J. and Timp, W. (2017). Detecting DNA cytosine methylation using nanopore sequencing. *Nat. Methods* **14**, 407-410. doi:10.1038/nmeth.4184
- 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. doi:10.1038/nprot.2018.015
- Smallwood, S. A., Lee, H. J., Angermueller, C., Krueger, F., Saadeh, H., Peat, J., Andrews, S. R., Stegle, O., Reik, W. and Kelsey, G. (2014). Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat. Methods* **11**, 817-820. doi:10.1038/nmeth.3035
- Stasevich, T. J., Hayashi-Takanaka, Y., Sato, Y., Maehara, K., Ohkawa, Y., Sakata-Sogawa, K., Tokunaga, M., Nagase, T., Nozaki, N., McNally, J. G. et al. (2014). Regulation of RNA polymerase II activation by histone acetylation in single living cells. *Nature* **516**, 272-275. doi:10.1038/nature13714
- Stelzer, Y., Shivalila, C. S., Soldner, F., Markoulaki, S. and Jaenisch, R. (2015). Tracing dynamic changes of DNA methylation at single-cell resolution. *Cell* **163**, 218-229. doi:10.1016/j.cell.2015.08.046
- Stevens, M., Cheng, J. B., Li, D., Xie, M., Hong, C., Maire, C. L., Ligon, K. L., Hirst, M., Marra, M. A., Costello, J. F. et al. (2013). Estimating absolute methylation levels at single-CpG resolution from methylation enrichment and restriction enzyme sequencing methods. *Genome Res.* **23**, 1541-1553. doi:10.1101/gr.152231.112
- Strahl, B. D. and Allis, C. D. (2000). The language of covalent histone modifications. *Nature* **403**, 41-45. doi:10.1038/47412
- Suzuki, Y., Korch, J., Turner, S. W., Tsukahara, T., Taniguchi, J., Qu, W., Ichikawa, K., Yoshimura, J., Yurino, H., Takahashi, Y. et al. (2016). Agln: measuring the landscape of CpG methylation of individual repetitive elements. *Bioinformatics* **32**, 2911-2919. doi:10.1093/bioinformatics/btw360
- Tanay, A. and Regev, A. (2017). Scaling single-cell genomics from phenomenology to mechanism. *Nature* **541**, 331-338. doi:10.1038/nature21350
- Tekel, S. J., Barrett, C., Vargas, D. and Haynes, K. A. (2018). Design, construction, and validation of histone-binding effectors in vitro and in cells. *Biochemistry* **57**, 4707-4716. doi:10.1021/acs.biochem.8b00327
- Thakore, P. I., D'ippolito, A. M., Song, L., Safi, A., Shivakumar, K., Kabadi, A. M., Reddy, T. E., Crawford, G. E. and Gersbach, C. A. et al. (2016). Highly specific epigenome editing by CRISPR/Cas9 repressors for silencing of distal regulatory elements. *Nat. Methods* **12**, 1143-1149. doi:10.1038/nmeth.3630
- Trapnell, C. (2015). Defining cell types and states with single-cell genomics. *Genome Res.* **25**, 1491-1498. doi:10.1101/gr.190595.115
- Tycko, J., Van, M. V., Elowitz, M. B. and Bintu, L. (2017). Advancing towards a global mammalian gene regulation model through single-cell analysis and synthetic biology. *Curr. Opin. Biomed. Eng.* **4**, 174-193. doi:10.1016/j.cobme.2017.10.011
- Vora, S., Tuttle, M., Cheng, J. and Church, G. (2016). Next stop for the CRISPR revolution: RNA-guided epigenetic regulators. *FEBS J.* **283**, 3181-3193. doi:10.1111/febs.13768
- Wang, K., Li, X., Dong, S., Liang, J., Mao, F., Zeng, C., Wu, H., Wu, J., Cai, W. and Sun, Z. S. (2015). Q-RRBS: A quantitative reduced representation bisulfite sequencing method for single-cell methylome analyses. *Epigenetics* **10**, 775-783. doi:10.1080/15592294.2015.1075690
- Weber, M., Davies, J. J., Wittig, D., Oakeley, E. J., Haase, M., Lam, W. L. and Schübeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* **37**, 853-862. doi:10.1038/ng1598
- Weiner, A., Lara-Astiaso, D., Krupalnik, V., Gafni, O., David, E., Winter, D. R., Hanna, J. H. and Amit, I. (2016). Co-ChIP enables genome-wide mapping of histone mark co-occurrence at single-molecule resolution. *Nat. Biotechnol.* **34**, 953-961. doi:10.1038/nbt.3652
- Wu, X. and Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat. Rev. Genet.* **18**, 517-534. doi:10.1038/nrg.2017.33
- Wu, X., Inoue, A., Suzuki, T. and Zhang, Y. (2017). Simultaneous mapping of active DNA demethylation and sister chromatid exchange in single cells. *Genes Dev.* **31**, 511-523. doi:10.1101/gad.294843.116
- Yoshida, H., Lareau, C. A., Ramirez, R. N., Rose, S. A., Maier, B., Wroblewska, A., Desland, F., Chudnovskiy, A., Mortha, A., Dominguez, C. et al. (2019). The cis-regulatory atlas of the mouse immune system. *Cell* **176**, 897-912.e20. doi:10.1016/j.cell.2018.12.036
- Yun, M., Wu, J., Workman, J. L. and Li, B. (2011). Readers of histone modifications. *Cell Res.* **21**, 564-578. doi:10.1038/cr.2011.42
- Zhu, J., Adli, M., Zou, J. Y., Verstappen, G., Coyne, M., Zhang, X., Durham, T., Miri, M., Deshpande, V., De Jager, P. L. et al. (2013). Genome-wide chromatin state transitions associated with developmental and environmental cues. *Cell* **152**, 642-654. doi:10.1016/j.cell.2012.12.033
- Zhu, Q., Stöger, R. and Alberio, R. (2018). A lexicon of DNA modifications: their roles in embryo development and the Germline. *Front. Cell Dev. Biol.* **6**, 24. doi:10.3389/fcell.2018.00024