

# Biophysics is reshaping our perception of the epigenome: from DNA-level to high-throughput studies

Austė Kanapeckaitė,<sup>1,4,\*</sup> Neringa Burokienė,<sup>2</sup> Asta Mažeikienė,<sup>3</sup> Graeme S. Cottrell,<sup>4</sup> and Darius Widera<sup>4</sup>

<sup>1</sup>Algorithm379, Laisvės g. 7, LT 12007, Vilnius, Lithuania; <sup>2</sup>Clinics of Internal Diseases, Family Medicine and Oncology, Institute of Clinical Medicine, Faculty of Medicine, Vilnius University, M. K. Ciurlionio str. 21/27, LT-03101 Vilnius, Lithuania; <sup>3</sup>Department of Physiology, Biochemistry, Microbiology and Laboratory Medicine, Institute of Biomedical Sciences, Faculty of Medicine, M. K. Ciurlionio str. 21/27, LT-03101 Vilnius, Lithuania; and <sup>4</sup>Reading School of Pharmacy, Whiteknights, Reading, UK, RG6 6UB

**ABSTRACT** Epigenetic research holds great promise to advance our understanding of biomarkers and regulatory processes in health and disease. An increasing number of new approaches, ranging from molecular to biophysical analyses, enable identifying epigenetic changes on the level of a single gene or the whole epigenome. The aim of this review is to highlight how the field is shifting from completely molecular-biology-driven solutions to multidisciplinary strategies including more reliance on biophysical analysis tools. Biophysics not only offers technical advancements in imaging or structure analysis but also helps to explore regulatory interactions. New computational methods are also being developed to meet the demand of growing data volumes and their processing. Therefore, it is important to capture these new directions in epigenetics from a biophysical perspective and discuss current challenges as well as multiple applications of biophysical methods and tools. Specifically, we gradually introduce different biophysical research methods by first considering the DNA-level information and eventually higher-order chromatin structures. Moreover, we aim to highlight that the incorporation of bioinformatics, machine learning, and artificial intelligence into biophysical analysis allows gaining new insights into complex epigenetic processes. The gained understanding has already proven useful in translational and clinical research providing better patient stratification options or new therapeutic insights. Together, this offers a better readiness to transform bench-top experiments into industrial high-throughput applications with a possibility to employ developed methods in clinical practice and diagnostics.

**WHY IT MATTERS** Epigenetic research holds great promise to advance our understanding of biomarkers and processes involved in health and disease. This in-depth biophysical research overview highlights how bioinformatics, machine learning, and artificial intelligence allow new insights into gene expression modulation and potential cancer biomarkers. Moreover, this review reveals how epigenetics and epigenomics are shifting from molecular-biology-driven to multidisciplinary strategies relying more on biophysical analysis. Importantly, the biophysical applications have already proven useful in translational and clinical research, providing better patient stratification and new therapeutic insights.

## INTRODUCTION

The genome encompasses multiple layers of biological information that extend beyond the DNA sequence. This is realized through the epigenetic regulome, which is interlinked with the genome and may control various phenotypic manifestations that can determine normal cellular function and/or various pathological states. Diverse epigenetic regulatory mechanisms influence long-term alterations in the transcriptional potential of

a cell by further blurring the lines between the genetic and epigenetic processes as well as the environmental influences that shape these interactions (1). The realization of this complex interplay shifted the focus toward epigenetics and epigenomics in modern biomedical research (2).

Broadly, epigenetic and epigenomic research focuses on mitotically heritable gene expression regulation, which is independent of changes in DNA sequence (3). Epigenetics and epigenomics cover a broad spectrum of DNA alterations, modifications to histones, complex DNA and protein interactions, and an intricate autoregulatory architecture. Because links to epigenetic alterations have been identified in many human disorders such as cancer, autoimmune

Submitted April 26, 2021, and accepted for publication September 24, 2021.

\*Correspondence: [auste.kan@algorithm379.com](mailto:auste.kan@algorithm379.com)

Editor: Jörg Enderlein.

<https://doi.org/10.1016/j.bpr.2021.100028>

© 2021 The Authors.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



pathologies, neurodegenerative diseases, and metabolic imbalances, the importance of epigenetic regulatory factors and their therapeutic potential have become more apparent in recent years (4). Furthermore, recent technical advances have made epigenetic analysis methods more affordable, enabling not only qualitative but also quantitative evaluation of epigenomic changes both on the gene and genome level (3,5–7).

Epigenetics draws from many different disciplines and this, no doubt, contributed to the fast growth of this interdisciplinary field (5–7). This is also reflected in our text mining of more than 30 million PubMed records indicating that there was a continuous growth in epigenetics research outputs in the last 20 years (Fig. 1). Although epigenetic profiling techniques were developed with a focus on molecular biology or biochemistry, it is important to highlight how biophysical research contributed to sequence-level, structural, and functional analyses as evidenced by the steadily growing number of biophysics publications in the context of epigenetics (Fig. 1). Thus, this review concentrates on the growing scope and capabilities of biophysical analysis methods in epigenetics (Table 1). Importantly, artificial intelligence (AI) and machine learning (ML) are becoming an essential part of the bio-

physical analysis pipeline in epigenomics because the increasing data volumes from structural, microscopy, or other measurement techniques require a robust methodology to harmonize and process the data. As a result, we discuss emerging concepts, both theoretical and experimental, in AI and ML as well as bioinformatics that are used for the exploration of the epigenome.

The discussion is split into several hierarchical levels that follow the organization of the epigenetic network. We provide an overview of the biophysical methods used to study heritable yet reversible epigenetic changes, such as DNA methylation and other base modifications. This is followed by an exploration of the available methods for the assessment of histone modifications and local chromatin architecture. Finally, biophysical techniques are introduced for higher-order chromatin hubs that exert transcriptional regulation through spatiotemporal changes in chromatin (Fig. 2). All this ties in with a quick overview of emerging bioinformatics and AI and ML techniques.

Overall, this review not only highlights key advances in the field but also underscores the need for standardized biophysical protocols and research practices in epigenetics. As multidisciplinary approaches are taking center stage in epigenetics and epigenomics, biophysics has the potential to expand research capabilities from

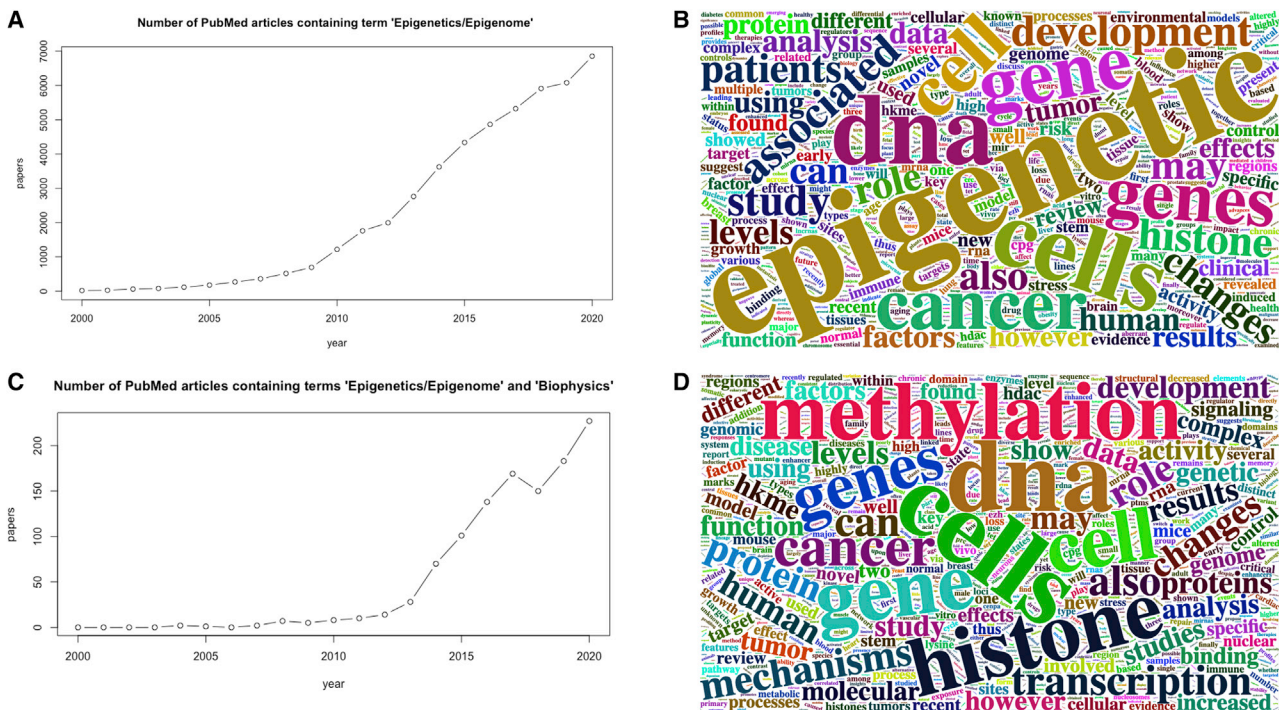


FIGURE 1 PubMed text mining. (A) 41,558 PubMed articles containing term “Epigenetics/Epigenome” in their abstracts between 2000 and 2020 were filtered from the global search; (B) word cloud visualization of the most common terms in the abstracts of (A) covering the peak range between 2013 and 2021. (C) 1065 of PubMed articles containing terms “Epigenetics/Epigenome” and “Biophysics” in their abstracts between 2000 and 2020 were filtered from the global search; (D) word cloud visualization of the most common terms in the abstracts of (C) covering the peak range between 2015 and 2020.

**TABLE 1 Comparison of epigenetics and epigenomics methods in the context of biophysical analyses**

Method	Analysis scale or level	Type of analysis	Comments	References
DNA-level modification analysis				
PCR-based bisulfite sequencing	single-base resolution	DNA methylation	semiquantitative approach	Li (8); Harrison and Parle-McDermott (9); Li and Reinberg (10)
Methylation-specific PCR	single-base resolution	DNA methylation	semiquantitative approach	Li (8); Harrison and Parle-McDermott (9); Li and Reinberg (10)
Bisulfite-based pyrosequencing	single-base resolution	DNA methylation	quantitative approach	Li (8); Harrison and Parle-McDermott (9); Li and Reinberg (10)
HRM	single-base resolution	DNA methylation	quantitative; genetic variation assessment	Harrison and Parle-McDermott (9)
SMRT sequencing	single-base resolution	DNA methylation	quantitative; allows sequencing native DNA through single-molecule long-read sequencing	Flusberg et al. (11)
Nanopore sequencing	single-base resolution	DNA methylation	allows sequencing native DNA through single-molecule long-read sequencing	Li (8); Shim et al. (12); Gilboa et al. (13); Simpson et al. (14)
scBS-seq	single-base resolution	DNA methylation	provides information on individual cell DNA methylation, can be extended to WGBS	Li (8); Smallwood et al. (15)
Methyl-TROSY	single-base resolution	DNA methylation	allows the investigation of methylated DNA interactions with proteins	Abramov et al. (16)
SERS	single-base resolution	DNA methylation	allows direct and multiplexed detection of DNA base modifications	Heck et al. (17); Heck et al. (18); Zrimsek et al. (19); Barhoumi and Halas (20); Guerrini et al. (21)
Global modification analysis				
WGBS	single-base resolution/ global analysis	DNA methylation	quantitative approach; a well-established protocol to detect methylated cytosines in genomic DNA	Li (8); Ziller et al. (22)
DIP-seq	genomic locations	DNA methylation and other modifications	semiquantitative approach; purification technique to enrich for methylated DNA sequences	Lentini et al. (23); Skvortsova et al. (24); Thomson et al. (25)
MeDIP	genomic locations/ genome-wide methylation patterns	DNA methylation	semiquantitative approach; large-scale (chromosome- or genome-wide) purification technique to enrich for methylated DNA sequence; abundance resolution in ~100 bp	Li (8); Lentini et al. (23); Skvortsova et al. (24); Thomson et al. (25); Harrison and Parle-McDermott (9)
epiGBS/RRBS	single-base resolution/ genome-wide analysis	DNA methylation	only detects 1–3% of genome; comparative analysis of DNA methylation and genetic variation in hundreds of samples de novo	Li (8); Van Gorp et al. (26)
m6A-LAIC-seq	single-base resolution/ transcriptome-wide analysis	RNA methylation	quantitative deconvolution of methylated versus nonmethylated transcripts	Molinie et al. (27)
sc-GEM	single-base resolution/ genome-wide analysis	DNA methylation	methylation profiling of target DNA sites using SCRAM in combination with single-cell RT-qPCR and single-cell genotyping by next-generation sequencing	Cheow et al. (28)
SPR	genomic locations	DNA methylation	label-free, real-time, sensitive, and specific method for accurate detection of regional DNA methylation	Carrascosa et al. (29)

(Continued on next page)

**Table 1. Continued**

Method	Analysis scale or level	Type of analysis	Comments	References
Spectrometric tracking or optical mapping for 5mC, 5hmC, 5fC, and 5caC	genome-wide	various DNA modifications	fluorescent reporters based semiquantitative assessment	Heck et al. (17); Gilboa et al. (13); Song et al. (30)
MS	genome-wide/genomic region	various DNA modifications	quantitative approach	Fernandez et al. (31); Ehrich et al. (32); Lin et al. (33)
Histone and chromatin modification analysis				
Immunoprecipitation (ChIP)-based methods	genome-wide/target protein specific	histone and chromatin modifications	evaluation of the global histone-modification state; the method allows capturing the respective DNA sequence	Li (8); Kimura (34)
DNase-seq	genome-wide/target protein specific	chromatin modifications	method for mapping active gene regulatory elements across the genome; however, there is evidence for the sequence-specific DNase I cutting biases that may not reflect chromatin accessibility to transcription factors	He et al. (35)
DNase-FLASH	genome-wide/target protein specific	chromatin modifications	demarcation of the chromatin structure surrounding human promoters	Vierstra et al. (36)
Nucleosome x-ray	genome-wide/target protein specific	histone/nucleosome modifications	structural biophysics method to assess nucleosome positioning, stretching of nucleosomal DNA, metal ion binding, and interaction determination with other proteins	Ekundayo et al. (37); Schalch et al. (38); Tan and Davey (39)
Nucleosome cryo-EM	genome-wide/target protein specific	histone/nucleosome modifications	structural biophysics method to assess nucleosome-based recognition and histone-modification mechanisms	Boopathi et al. (40)
Nucleosome NMR	genome-wide/target protein specific	histone/nucleosome modifications	method for the secondary structure and intramolecular dynamics analysis of histones and nucleosomes	Shi et al. (41); Moriwaki et al. (42); Gao et al. (43); Zhou et al. (44)
Nucleosome MS	genome-wide/target protein specific	histone/nucleosome modifications	quantitative analysis method to study the higher-order structures, such as nucleosome diversity, in the epigenome	Saikusa et al. (45)
SCAN	genome-wide	chromatin modifications	method uses the fluorescent probes to track changes in histones or survey the epigenomic marks across genomes	Heck et al. (17); Murphy et al. (46); Hyun et al. (47)
ChIP-String	genome-wide	chromatin modifications	method allows characterizing the chromatin state at several hundred positions in the genome by pulling down chromatin fragments labeled with antibodies to assess specific histone modifications	Ram et al. (48)

(Continued on next page)

**Table 1. Continued**

Method	Analysis scale or level	Type of analysis	Comments	References
Fluorescence-based chromatin mapping	genome-wide	chromatin modifications	chromatin fragments with a nucleosome are attached to a microscope slide and the fluorescent antibody interaction with the tethered structure can be used to evaluate different histone states	Harris et al. (49); Shema et al. (50)
SPRi	genome-wide/target protein specific	histone/nucleosome modifications	3D carbene-chip-based SPRi to assess modified histone peptides and nucleic acids; method involves a complicated preparation of the sample	Zhao et al. (51)
Higher-order chromatin structure analysis Chromosome conformation capture (3C, 4C, 5C)	spatial chromatin organization	interacting chromatin loci	These methods quantify the number of interactions between genomic loci that are close in 3D space. Quantification can be integrated with PCR or deep sequencing to produce genome-wide interaction maps.	Dostie et al. (52); Gavrillov et al. (53); Ohlsson and Göndör (54)
MST	chromatin-level interactions	protein and chromatin interactions	Thermophoretic movement-based rapid quantification of the molecular interactions between proteins and DNA. This method can be adapted to capture histone complexes or nucleosomal milieu influence on the recognition of histone modifications.	Schubert and Längst (55); Corbeski et al. (56)
Hi-C	chromosome level	chromosome conformation capture	method allows simultaneously capturing chromosome conformation and DNA methylome in a single assay; however, the method is complicated and expensive	Bonev and Cavalli (57); Gibcus and Dekker (58)
Single-molecule localization microscopy	chromatin level / chromosome level	macromolecular complex interactions	multiple variations exist for different visualization protocols to track chromosome subdomains and territories; subcellular structures and single-molecule dynamics can be analyzed at a nanoscale level	Manders et al. (59); Visser et al. (60); Zwettler et al. (61); Hausmann et al. (62)
FISH	chromatin level / chromosome level	macromolecular complex interactions	used to visualize genomic loci in three dimensions that can be integrated into chromatin tracing and multiplexing protocols	Wang et al. (63); Hu and Wang (64)
Confocal microscopy	chromatin level / chromosome level	macromolecular complex interactions	multiple different protocols and technical set-ups exist to track chromatin conformational dynamics	Poirier et al. (65); Schueder et al. (66); Jeffet et al. (67)
cryo-EM	chromatin level	macromolecular complex interactions	allows capturing structures from a single nucleosome in several dynamic states to higher-order chromatin formation events	Saksouk et al. (68); Machida et al. (69); Cai et al. (70); Merk et al. (71); Song et al. (72); Grant et al. (73); Punjani et al. (74)

*(Continued on next page)*



**Table 1. Continued**

Method	Analysis scale or level	Type of analysis	Comments	References
BE-ICP MS	chromatin level	macromolecular complex interactions	method allows measuring nucleosome electrostatics and assessing nucleosome formation as well as electrostatic forces governing chromatin fiber formation	Gebala et al. (75)
iDAPT	chromatin level	chromatin and interacting protein analysis	Biochemical enrichment using a bifunctional transposase-peroxidase method combined with the bioinformatic analysis of the genomic and proteomic profiles of open chromatin. This method allows capturing sequence-based characteristics of open chromatin and interacting proteins	Lee et al. (76)

DNase-FLASH, DNase I-released fragment-length analysis of hypersensitivity; epiGBS, reference-free reduced-representation bisulfite sequencing; FISH, DNA fluorescence in situ hybridization; iDAPT, integrative DNA and protein tagging; m6A-LAIC-seq, m6A-level and isoform-characterization sequencing; RRBS, reduced representation bisulfite sequencing; RT-qPCR, real-time qPCR; SCAN, single-chromatin molecule analysis at the nanoscale; scBS-seq, single-cell genome-wide or reduced-representation bisulfite sequencing; WGBS, whole-genome bisulfite sequencing.

singular experiments to high-throughput screens with the applications in translational and clinical research.

### Sequence nonaltering DNA modifications that reverberate through all regulatory networks: what can biophysics tell us?

#### *The landscape of DNA sequence modifications*

A major epigenetic mechanism involving direct chemical modification to the DNA sequence is achieved through DNA methylation. Historically, DNA methylation was identified in mammals as early as the discovery of DNA (79). This epigenetic modification is the first epigenetic regulatory layer that plays a role in gene expression, which is now recognized as a significant epigenetic factor influencing chromosomal structure, DNA conformation, DNA stability, and the interactions between DNA and proteins (80) (Figs. 1 and 2).

Methylation is achieved via the action of enzymes that add a methyl group at the 5'-position of cytosine bases (5mC) in DNA. Other modifications, which include the oxidation of 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) and the addition of a methyl group to adenine (A) generating N6-methyladenine (6mA), are also being recognized as important epigenetic regulators (81,82). However, out of 43 catalogued DNA modifications found in natural DNA (83), 5mC is the most frequent alteration to the DNA sequence in plants and animals (82,84,85). Because of the high 5mC frequency in the genome (more than 4% of the human genome cytosines are reported to be modified in this way), this modification became the most studied one—to such an extent that it is considered the fifth base of DNA (82,84,85). Although the methylation at the five position of DNA cytosine in the vertebrate genomes is maintained by the synergetic catalytic actions of three DNA methyltransferases (DNMTs), demethylation is believed to occur through several metabolic routes along which, for example, a family of mammalian TET proteins has been recently linked to this process. Mammalian DNMT3A and DNMT3B illustrate well the intricate nature of methylation maintenance because these enzymes are also known to be capable of directly removing the hydroxymethyl moiety from 5hmC in vitro, and based on the experimental work, it has been suggested that DNMTs in combination with TET might actively demethylate 5mC *in vivo* as well (86–88).

Epigenetic changes extend beyond single-base modifications, as some cytosine nucleotides are followed by a guanine nucleotide forming a linear alternating DNA sequence (5'-C-phosphate-G-3' or CpG). These stretches of DNA are known as CpG sites and they tend to occur at high frequency in genomic regions

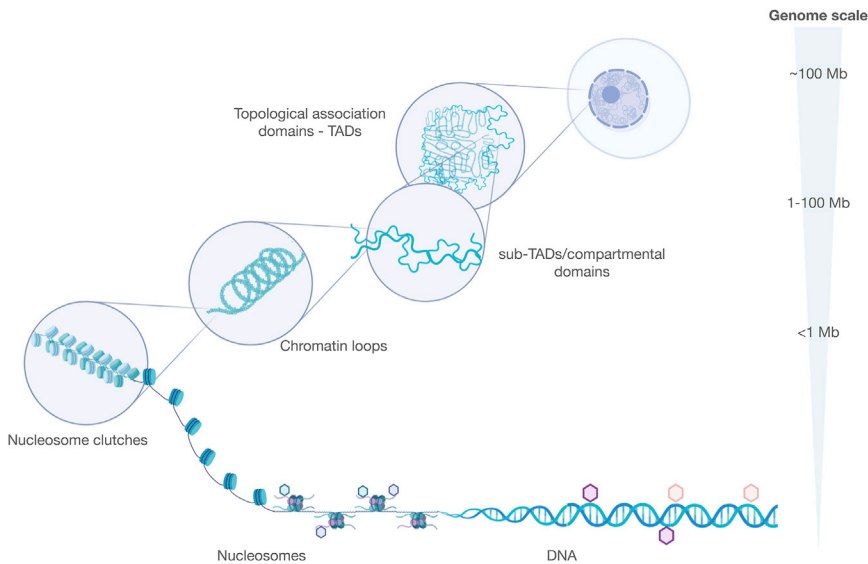


FIGURE 2 Eukaryotic chromatin organization (adapted from Harabula and Pombo (77) and Magaña-Acosta and Valadez-Graham (78)). Chromosome territories represent the highest organizational level followed by topological architecture domains (TADs) and sub-TADs. Sub-TADs are formed by the chromatin loops, which are composed of nucleosome clutches. Nucleosome aggregates are composed of individual nucleosomes wrapping around DNA. At each organizational level, the architectural domains or structures interact with different proteins and nucleosome tails, and DNA might have various modifications (exemplified by the hexagons in the diagram). The corresponding genome content in megabases (Mb) is provided next to the diagram.

called CpG islands (CGIs) (89). Although CpG dinucleotides are described as rare in genomes, 75–85% of them are methylated in mammals (90,91). CpG sites become especially relevant because around 60% of the promoters in the human genome contain CGIs, and these regions are typically unmethylated (87,91). This changes with aging and cellular perturbation events that promote further genomic destabilization through hypermethylation (91). Moreover, recent genome-wide methylome surveillance studies highlighted that methylation exerts a broad influence on gene control. In this context, methylation in the immediate vicinity of the transcriptional start site can prevent initiation. However, if it occurs in the gene body, it might have an opposite effect. Similarly, it became evident that methylation in centromeric CpG regions can influence chromosomal stability and successful segregation during mitosis (92,93). Another level of complexity is added by the facts that CGIs are not the only regions that can be methylated and that dinucleotide methylation patterns are unique to different tissues (92). In addition, it is also important to stress that methylation effects are not binary, fluctuating between on and off states, but rather a spectrum dependent on many factors. It has been previously reported that 40% methylation was sufficient to silence expression and that specific positioning of CpGs in promoter regions can influence gene expression (87,94). Other modifications to cytosines can also have an impact on the expression profile; for example, it has been shown that hydroxymethylation and methylation events in the genome can balance cell lineage commitment and pluripotency (52).

The importance of DNA methylation and other modifications is emphasized by the growing number of hu-

man diseases that are linked to these epigenetic marks (95). Aberrant methylation patterns have been identified in various cancers and other common human diseases, such as autoimmune diseases, metabolic imbalances, and psychological disorders. Therefore, these epigenetic changes are explored as potential biomarkers or therapeutic targets (95,96). In the following sections, we review some of the key research methods to analyze epigenetic alterations with a focus on classical and emerging biophysical techniques (Table 1).

#### *Molecular biology methods with ideas from biophysics to capture DNA sequence modifications*

The majority of the exploratory tools for DNA methylation belong to the realm of molecular biology, and several excellent reviews provide an in-depth perspective on the progress in epigenetics (8), as well as the most impactful developments over the past 10 years (2). Classically, DNA methylation analysis is performed by employing bisulfite conversion, in which sodium bisulfite reaction leads to cytosine deamination and conversion to uracil, whereas methylated cytosines are protected. The downstream analyses include polymerase chain reaction (PCR), quantitative PCR (qPCR), or sequencing to capture different methylation levels. PCR and qPCR offer a semiquantitative approach, and pyrosequencing allows quantitative evaluation of methylation in a particular genomic region (9,10). Other methods, such as high-resolution melt analysis (HRM) and immunoprecipitation techniques, allow the detection of DNA methylation levels through qualitative assessment. HRM can be used for a rapid genetic variation assessment, whereas methylated DNA immunoprecipitation captures genome-wide methylation

patterns (9). All these approaches are variations of PCR-, next-generation-, and/or microarray-based sequencing. Despite molecular biology driving significant advancements in our understanding of the epigenome and its regulation, these methods lack the ability to capture the full scope of the regulatory kinetics and fail to provide a quantitative assessment of epigenetic processes.

Furthermore, the analytical accuracy can be compromised by a significant variation in the readouts. This was demonstrated for DNA immunoprecipitation (DIP) sequencing (DIP-seq), which is a common method used to profile DNA modifications. Initially, DIP-seq was used as an enrichment and profiling method for genome-wide 5mC distribution studies in which antibody-based enrichment of methylated DNA fragments (MeDIP) was coupled with hybridization to DNA microarrays (MeDIP-chip) or high-throughput sequencing (MeDIP-seq) (23–25). DIP-seq has recently been applied to capture the genomic locations of other modifications, namely 5hmC, 5fC, 6mA, and 5caC (23). However, follow-up studies for DIP profiles using independent methods highlighted that DIP-seq analysis results in preferential enrichment of regions with low CG content (23). A study conducted by Lentini and colleagues concluded that DIP-based assays require both matched input (sample that has not been immunoprecipitated in all DNA recovery steps) and immunoglobulin G controls, as well as non-antibody-based techniques (23). Such examples demonstrate how cytosine profiling data can be misinterpreted if adequate precautions are not taken. In addition, it becomes evident that although there are various analytical techniques in molecular biology, the reliability varies, and even the established methods need to be scrutinized.

The introduction of other molecular research tools or modifications to the existing methods allowed researchers to improve the accuracy and broaden the detection spectrum. The shortcomings of bisulfite sequencing, such as the inability to detect other base modifications, led to the emergence of new sequencing and analysis methods for a wide range of modifications. One such example is reduced-representation bisulfite sequencing allowing for DNA methylation and genetic variation comparative analysis at a low cost for a large number of de novo samples. This genotyping method relies on bisulfite-converted DNA sequencing, which enables mapping, variant calling, and distinction of single-nucleotide polymorphisms. This information can be subsequently referenced against known methylation variations (26). Another recently introduced method, m6A-level and isoform-characterization sequencing, allows a quantitative deconvolution of methylated versus nonmethylated

transcripts (27). This new sequencing approach does not fragment RNA before anti-m6A RNA immunoprecipitation, as it would be performed following the m6A-seq protocol. Instead, Molinie and co-workers (27) developed a protocol allowing sequencing of intact full-length transcripts of m6A-positive and m6A-negative fractions post-RNA immunoprecipitation. These improved protocols illustrate the emergence of novel techniques in epigenetics to quantify differential methylation of transcript isoforms.

However, despite all the progress, some fundamental limitations still needed to be addressed to gain quantitative insights. By incorporating ideas from biophysics, such as interaction or process kinetics modeling, it became possible to develop analytical methods capturing epigenetic modifications on a base level. One such example is PacBio's Single Molecule, Real-Time (SMRT) sequencing, which relies on sequencing by synthesis, in which base modifications affect the kinetics of the polymerase and alterations to nucleotides can be inferred comparing a modified sequence to an *in silico* model or unaltered template (11). As can be expected, this method suffered from a high signal/noise ratio. However, the explored ideas were echoed in other projects. Excellent reviews by Kurdyukov and Bullock as well as by Young and colleagues provide resources on the molecular techniques and sequencing methods used for the methylation analysis (97,98). Overall, conventional molecular biology techniques combined with biophysical assessment of kinetics and/or interactions tracking are first conceptual steps toward better precision and integration into high-throughput studies.

#### *Zooming in on single-base alternations with the help from biophysics*

Versatility is key in epigenetics and epigenomics research and helps to capture the full scope of the regulatory complexity. Therefore, biophysical methods are also beginning to make inroads into routine laboratory techniques to study epigenetics (Table 1). In this context, nanopore-based sequencing methods are employed to chart different epigenetic modifications, capturing not only the epigenetic tissue heterogeneity but also cellular level complexity. For example, solid-state nanopores were used to explore methylation by selective labeling of methylation sites with methyl-CpG-binding domain protein 1 (MBD1) (99). This technique relies on detection of methylated CpG dinucleotides and subsequent labeling of the sequence with a 75-amino-acid region of the methyl-DNA binding protein. Protein tethering to the methylated region induces a threefold increase in the blockage current when passing through the solid-state nanopore. This



nanopore-based methylation assay is an alternative to bisulfite conversion, fluorescent labeling, or PCR (12) and allows a more direct evaluation of the methylation state. Another single-molecule quantification method for the detection of multiple unmethylated CpGs relies on linking of DNA to a synthetic cofactor using DNA methyltransferases. Electro-optical nanopore detection is then used to assess and quantify unmethylated CpGs at the single-molecule level (13). Gilboa's proof-of-principle study showed that this electro-optical method can be used to analyze longer stretches of double-stranded DNA (around 10 kbp), foregoing PCR amplification or bisulfite modification. Biophysical detection methods for methylated dinucleotides have also led to commercial products, namely the Oxford Nanopore Technologies MinION sequencer, which turned nanopore sequencing devices into electrolytic current detectors sensitive to base modifications (14). Solutions for faster sample handling without the need of extensive preprocessing can find their way into precision medicine applications (100,101).

Advances in single-cell sequencing also found their application in epigenetics. One example of single-cell genome-wide analysis is reduced-representation bisulfite sequencing (15), which is used to map relationships between epigenetic, genomic, and transcriptional levels. This method relies on the isolation of single cells and treatment of the genomic DNA with sodium bisulfite leading to DNA modification and fragmentation. After the DNA conversion, fragments undergo random priming and PCR amplification before the sequencing. After the deep sequencing, it is possible to identify single cytosine methylation events at high resolution. Another protocol (scM&T-seq) allows the performance of parallel single-cell genome-wide methylome and transcriptome sequencing. scM&T-seq uses genome and transcriptome sequencing (G&T-seq) methodology to convert DNA using bisulfite without affecting the transcriptome. The authors of this method claim that the developed protocol enables discovering associations between transcriptional and epigenetic variation (102). However, single-cell sequencing combined with bisulfite processing is still an expensive method, and chemical processing introduces additional variations to the sample. Therefore, sequencing methods are combined with other analysis platforms to improve detection limits.

A recent study introduced a method to genotype single cells and capture DNA methylation at multiple loci (28). High-throughput microfluidic platform (single-cell analysis of genotype, expression, and methylation; sc-GEM) relies on single-cell restriction analysis of methylation (SCRAM) in combination with single-cell real-time qPCR as well as on single-cell genotyping by

next-generation sequencing. As a basis for sc-GEM, Cheow and colleagues introduced SCRAM in their earlier work to measure the DNA methylation state and assess multiple target sites in single cells (103). The SCRAM pipeline begins with the isolation and lysis of single cells, followed by digestion of genomic DNA with a methylation-sensitive restriction endonuclease in which the main focus is the amplification of multiple targets by two rounds of PCR to capture the methylation profile of target DNA sites. The authors claim that sc-GEM overcomes the shortcoming of bisulfite-based single-cell DNA methylation assays such as DNA degradation, providing a more reliable examination of methylation.

Thus, microfluidics and/or nanopore-based systems provide a framework that can be easily integrated with molecular biology or biophysical techniques to analyze methylation. Moreover, parallelization of sequencing and genomic mapping can enrich the analysis by monitoring genetic and epigenetic perturbations. However, it is necessary to highlight that various new single-cell sequencing methods have to be rigorously tested to ensure that the assessments are precise and sensitive.

#### *Light-based methods to capture DNA methylation and other alteration patterns*

Optical detection and quantification are widely used in biomedical research. The broad applicability of various microscopy techniques not only offers a sensitive method to detect alterations at the molecular level but also provides means to quantify them (Table 1). Not surprisingly, a growing number of optical protocols and analytical platforms are used to study DNA epigenetic modifications. Specifically, optical detection techniques for DNA methylation have been developed by employing different fluorescence, Raman spectroscopy, electrochemiluminescence, and colorimetric readouts, as well as surface plasmon resonance (SPR). Perhaps one of the most interesting applications is SPR, which has the potential to become a real-time and sensitive monitoring method to assess multiple samples and detect clinically important biomarkers (29). Furthermore, it has been recently demonstrated that an SPR-based assay can be used to screen DNA methyltransferase inhibitors (104). A thorough review by Nazmul Islam and colleagues details the progress in DNA methylation optical research, in which the authors overview various optical analysis strategies and associated challenges. Moreover, they believe that many of the discussed proof-of-concept ideas are suitable to become point-of-care DNA methylation biosensors in the future (105).

Other imaging techniques, such as single-molecule optical detection of DNA modifications, have the

potential to detect and quantify separate molecules through fluorescent reporters. These methods are especially useful for multiplex experiments. A review by Heck outlines the labeling of DNA modifications, such as 5mC, 5hmC, 5fC, and 5caC, for spectrometric tracking or optical mapping (17). Although cytosine alterations are charted by the decrease in their content and the assessment is only semiquantitative, there are significant advances in pushing resolution limits. For example, human blood was analyzed by combining 5hmC labeling with optical genome mapping using nanochannels in the first whole-genome single-molecule epigenetic profiling study (106). This type of mapping may help to unveil locus-specific 5hmC patterns for diagnostic purposes. This approach has been commercialized by BioNano Genomics, Inc. (13). Optical tracking was also used to measure 5mC and 5hmC abundance in CpG sites using single-molecule fluorescence resonance energy transfer to explore a potential involvement of these structures in gene regulation in the mouse genome (30). Thus, undoubtedly, optical detection of epigenetic changes at a single-base level has great commercialization potential if these attempts can grow beyond laboratory experiments.

#### *Biophysics bridges sequence modifications and function in epigenetics research*

Biophysical methods for single-base analyses extend far beyond sequencer modifications or optical techniques (Table 1). The classical biophysical toolbox has been extensively applied and modified to capture methylation-specific events and even connect those changes with structural and functional effects. An interesting comparative study of the temperature dependences of  $^1\text{H}$  NMR, ultraviolet absorption, and Raman scattering spectra revealed that CpG motif methylation and the surrounding base composition influence the enthalpy and entropy of DNA duplex formation (107). Thus, thermodynamic analysis could be one avenue to improve our understanding of DNA methylation effects. Specifically, measuring the thermodynamic parameter changes can help assess the methylation in real time and infer single-nucleotide interactions. Although in practical terms, this might not be used for diagnostics, it could greatly enhance our understanding of the intricate nature of how sequence composition and methylation affect the formation of a DNA duplex (108,109).

However, there are also challenges in adopting biophysical analysis techniques for epigenetics; for example, although NMR is a versatile and important tool in studying biomolecules, nucleic acid NMR studies have proven to face more challenges in chem-

ical shift dispersion, spectral quality, and sample preparation. Thus, research is ongoing to address structure-dynamics-function relationships for DNA methylation and interacting proteins (16,110). Abramov and co-workers optimized methyl-transverse relaxation-optimized spectroscopy (methyl-TROSY)-based NMR and demonstrated the importance of labeling DNA molecules with methyl-group probes at specific sites to obtain high-quality data for large molecular complexes (16). The authors explored methyl-labeled DNA (5mC and 153-bp Widom DNA molecule) with nucleosomal proteins reaching high spectral resolution and sensitivity with methyl-TROSY.

The fluorescence-based techniques discussed above provide high sensitivity, but the accessible spectral window may limit multiplexing options (17). These limitations can be resolved by other biophysical techniques, such as surface-enhanced Raman scattering (SERS). SERS is a label-free analyte detection method sensitive enough to capture individual molecules. Moreover, SERS tagging (e.g., gold or silver nanoparticles with dye coating) can be easily multiplexed because Raman spectra have very narrow bandwidths (18,19). This method relies on the strong Raman signal enhancement occurring in close proximity to gold or silver nanostructures during SPR (17). SERS has already been applied to directly detect 6mA, 5mC, 5hmC, and 8-oxo-guanine (20). Other extensions of SERS were used to analyze single-base mismatches, as well as to detect 5mC and 6mA in DNA duplexes, permitting the quantification of hybridization events (21). Taken together, it is evident that label-free multiplexing SERS can be valuable for the evaluation of DNA base modifications. However, as Heck points out, more work needs to be done to go beyond the proof-of-concept stage by establishing robust sample preparation and analysis procedures (17).

#### *DNA modification assessment with mass spectrometry*

Understanding DNA epigenetic modifications in their biological context requires establishing and quantifying epigenetic alteration patterns. Mass spectrometry (MS) allows charting epigenetic marks across the epigenome in a high-throughput fashion to surveil global DNA methylation patterns (Table 1).

Quantification of 5mC and 5hmC can be performed using ultraperformance liquid chromatography and MS (triple quadrupole) detection for accurate and fast quantification of relative global 5mC and 5hmC levels (required genomic material 1  $\mu\text{g}$ ) (31). Quantitative high-throughput evaluation of DNA methylation can also be achieved using matrix-assisted laser desorption/ionization time-of-flight (TOF)-MS. The method relies on a single-base-specific cleavage

reaction to determine methylation sites as well as methylation ratios for a region of interest. The high-throughput nature of this approach can be especially useful for quantitative analysis of methylation to accurately classify histopathology samples (32).

Target fragmentation assay (TFA), a novel MS method, allowed ultrasensitive DNA methylation profiling because TFA selectively captures cleaved target DNA fragments using magnetic separation followed by MS (33). Although this method is highly sensitive (the detection limit is as low as 0.056 amol) and does not require sample amplification to detect nonenzymatic hydrolysates of the target sequences (33), it still needs to become high throughput. Moreover, TFA requires thorough testing to assess its quantification capabilities.

The need for benchmarking newly developed methods is exemplified by a recent study, which was able to reveal methylation levels of DNA, but the success depended on the cation exclusion levels (13). The presented analysis involved a nanoelectrospray TOF-MS (nano-ESI-TOF-MS) combined with CpG methylation insensitive nuclease restriction to characterize DNA methylation (<400 bp, *in vitro* samples).

Although the advantages of MS are evident when it comes to methylation level assessments, its applicability, accuracy, and scalability still pose challenges (111,112). This becomes especially apparent when considering a simple example of electrospray TOF-MS (ESI-MS) pipeline, in which samples must undergo extensive desalting in contrast to matrix-assisted laser desorption/ionization TOF-MS. Moreover, additional sample preparation such as chromatographic separation of DNA fragments may also be needed. Thus, the promising experimental work still needs to be further benchmarked, and the variations within the techniques require developing consistent protocols.

## Biophysical perspective can offer new insights into histone modifications and their architecture

### *The epigenetic landscape of chromatin*

DNA assembles into chromatin, in which gene regulation is directly tied to higher-order structure organization (10,113,114). Nucleosomes serve as fundamental chromatin units that wrap 146-bp DNA stretches around histone octamers (H3, H4, H2A, and H2B) (10,115–119) (Fig. 2). Whereas histone lysine residues are tightly associated with the negatively charged phosphate groups of DNA because of their positive charge, histone acetyltransferases can change this through the transfer of an acetyl group to histone tail lysine residues, leading to chromatin opening through negative-negative charge repulsion. In contrast, deacetylases remove acetyl groups from histones packing

chromatin regions more tightly. Consequently, post-translational modifications (PTMs) of histones can affect the access of transcription regulatory machinery to DNA and thus can exert multiple effects on both short- and long-term gene expression. A good illustration of the effects of histone PTMs can be found in actively transcribed genes, which are marked by trimethylated H3K4 (H3K4me3) and acetylated H3K27 (H3K27ac) at their transcription sites (34,120). It is also important to highlight that histone PTMs encompass a wide spectrum of alterations, which may also include methylation, phosphorylation, ubiquitylation, and sumoylation (10,68,115,116,121,122). As a result, the “histone code” is a term often used to describe these PTMs that govern gene expression through multiple systemically coordinated interactors (123–125).

Overall, nucleosomes and their role in the organization of chromatin architecture influence transcriptional activation, RNA processing, DNA damage repair, replication, and nuclear organization. Nucleosomal ordering is the result of a convergence of multiple cellular factors in a complex regulatory environment (126). Failure to maintain chromatin ordering is a known cause of neurodevelopmental or neurodegenerative diseases as well as immune and metabolic disorders (127). Therefore, the diversity of this epigenome layer highlights the need for precise tools to uncover the underlying network dynamics.

### *Classical molecular biology approaches with new ideas from biophysics to capture nucleosome and chromatin changes*

The majority of methods used to study histone function rely on molecular biology or biochemical techniques, such as chromatin immunoprecipitation (ChIP)-based methods (Table 1) (34). ChIP sequencing (ChIP-seq) allows to capture genome-wide chromatin modifications *in vivo* but the method typically needs larger samples ( $\sim 10^7$  cells). To overcome the technical limitations, new methodologies are continuously developed, as showcased by a new protocol that was introduced for microfluidics-based ChIP-seq and requires only as few as 100 cells (128). To make ChIP more efficient, the authors used multilayer soft lithography to build a poly(dimethylsiloxane) (PDMS) device with a simple microfluidic chamber ( $\sim 710$  nL in volume). The incorporation of the microfluidic chamber that supports magnetic bead flow ( $\sim 2.8$   $\mu\text{m}$  in diameter) allowed adsorption of sonicated chromatin fragments. Specifically, ChIP antibody-coated beads captured fragments of chromatin flowing through the chamber. After capture, the IP beads were washed by oscillatory washing to remove nonspecifically adsorbed chromatin fragments and collected for off-chip processing.

Another challenging aspect of chromatin studies is the capture and analysis of histone PTMs. To address this need, a method was developed using a streamlined semisynthesis of DNA-barcoded nucleosome libraries (129). These libraries with distinct combinations of nucleosomal PTMs are assembled and treated with various effectors or subjected to the nuclear proteome. ChIP is used to isolate the products of the interactions, and the samples are subsequently subjected to multiplexed DNA-barcode sequencing. The authors of this method anticipate that the high-throughput nature and sensitivity of the technology can uncover various interactions and signaling events at the chromatin level. Another study on high-resolution mapping of transcription factor binding sites on native chromatin highlighted that the choice of chemical treatment (e.g., formaldehyde) can introduce noise into the data by fixing some transient interactions (130). Instead, micrococcal nuclease was suggested for the digestion of chromatin without cross-linking before affinity purification of the transcription factors and paired-end sequencing. The use of micrococcal nuclease was based on an earlier study in which the researchers demonstrated that this nuclease digests only exposed DNA not protected by nucleosomes or other binding proteins (131). For example, the occupied regions of genomes from affinity-purified naturally isolated chromatin method uses micrococcal nuclease-digested non-cross-linked chromatin to establish high-resolution binding sites of transcription factors and provides a sensitive and specific profiling of direct protein-DNA interactions (130). Similarly, other nucleases were proposed to map transcription factor occupancy and predict the surrounding nucleosome architecture. DNase I-released fragment-length analysis of hypersensitivity was shown to enable capturing the locations of DNA-nucleosome interface, such as sites around human promoters (36). Yet, a number of methods still need to go through thorough testing. An analysis of sequencing of DNase I hypersensitive sites (DNase-seq) revealed sequence-specific DNase I cutting bias, which had not been adequately accounted for in previous footprinting studies. Therefore, the specific protein-DNA interaction might not be correctly assessed, leading to incorrect assignment of transcription factor binding (35).

#### *Nucleosome and chromatin architecture analysis using structural and quantitative biophysics methods*

To better understand how structural factors drive the histone assembly into nucleosomes, it is necessary to go beyond charting sequences and interactors. The quantitative and structural exploration of nucleosomes and chromatin organization may permit a bet-

ter architectural understanding of these epigenome elements. This has been attempted with x-ray studies, NMR, MS, cryogenic electron microscopy (cryo-EM), and even optical analyses (Table 1) (37,38,42,45,65,69,132).

Nucleosome x-ray studies provided the first glimpses into the structure and function of nucleosomes. For example, nucleosome positioning, stretching of nucleosomal DNA, metal ion binding, as well as interaction with other proteins allowed to advance and gain a better understanding of epigenomic processes (38,39). This structural analysis method recently allowed the capture of crystal structures of tetranucleosomes (variants of Widom 601 DNA reconstituted with *Xenopus laevis* histones expressed in *Escherichia coli*) with ~11-bp DNA linkers (5.8 and 6.7 Å resolution). The study showed how minimal intramolecular nucleosome-nucleosome interactions result in a fiber composition that resembles a flat ribbon (37). However, electron microscopy was needed to obtain high-resolution structures. Moreover, additional chemical manipulation was necessary to assess the diversity of nucleosome-nucleosome interactions. This exemplifies how challenging it is to study these chromatin architecture elements and connect that information with higher-order structures.

X-ray crystallography can be substituted by cryo-EM, and this is also true for nucleosomal studies (133). A number of recent studies demonstrated how powerful this method can be in capturing better-resolved nucleosomes. For example, cryo-EM was used to obtain structures for the nuclear receptor-binding SET domain (NSD) family proteins, NSD2 and NSD3, bound to mononucleosomes (reconstituted with 147-bp Widom 601 DNA and 187-bp DNA sequences). It was possible to capture how the binding of NSD2 and NSD3 to mononucleosomes leads to DNA unwrapping close to the linker region. This study offered new insights into NSD2 and NSD3-mediated nucleosome-based recognition and histone-modification mechanisms (134). Another cryo-EM study revealed the differential flexibility of DNA ends around CENP-A nucleosome (CENP-A 601 nucleosome core particle (NCP)) (40). As the resolution for x-ray NCP structures ranges from 1.9 to 7 Å (135–137), an interesting comparative study for x-ray and cryo-EM was performed for a 200 kDa NCP reconstituted with Widom 601 DNA, which also provided a high-resolution x-ray crystal structure. The authors were able to achieve an overall resolution of 3.9 Å, demonstrating that phase-plate cryo-EM could be an important tool to determine novel near-atomic resolution structures of complex samples (138).

Another valuable tool for structural analysis is NMR, which permits assessment of conformational flexibility



and has been exploited for nucleosome studies (43,44). An excellent review by van Emmerik and Ingen provides a detailed summary of the state-of-the-art NMR studies covering nucleosomal DNA, histone complexes, nucleosomes, and nucleosomal arrays (139). The authors suggest that by increasing both the depth and breadth of nucleosome NMR studies, it will become possible to capture new information on the dynamic landscape of nucleosomes and the interacting proteins. Recently, NMR helped to determine the first structure of isolated full-length H2A-H2B heterodimers (42). Another key study on the secondary structure and intramolecular dynamics of human histone H4 (hH4) in the nucleosome was performed using solid-state NMR. Structure dynamics from nano- to microsecond as well as micro- to millisecond timescales were explored, revealing motion diversity of the hH4 protein (41). In contrast to other structure analysis methods that require specifically prepared or “frozen” samples, NMR can be used to assess the flexibility of observed structures, which could allow monitoring changes of the higher-order chromatin structures and DNA accessibility.

MS studies also provided new insights into multimolecule interactions, and this quantitative analysis method has been used to study higher-order structures in the epigenome. Structural diversity of nucleosomes was recently explored using native MS in which nucleosomes were analyzed by observing charge states of nucleosomes reconstituted with DNA (varying length) and using positive-mode ESI-MS (45). The results indicated that histone-free DNA regions have an effect on the structural diversity of nucleosomes. The generated data allowed the authors to postulate that histone tails with PTMs might behave differently in nucleosomes with long DNA strands compared to nucleosomes with unmodified histones. They further concluded that this might play a role in the biological function of the histone tail, which is linked to the regulation of DNA transcription, replication, and repair. Another report described a quantitative protein-DNA and protein-nucleosome binding assay based on affinity-purified nuclear extracts and isobaric chemical labeling, which enabled MS analysis; this approach helped capture apparent binding affinities of monomeric and multimeric transcription factors as well as chromatin remodeling complexes to DNA (132). As evident from these examples, MS can be adapted into assays for exploration of regulatory interactions.

Nucleosomes compact the genome by wrapping the DNA around a histone octamer, influencing many crucial regulatory processes ranging from gene expression to DNA repair (139). Thus, by incorporating more structural biophysical methods, we can gain new insights into the dynamics of various interactions.

#### *Optical techniques for histone modifications*

In recent decades, continuously evolving microscopy such as super-resolution microscopy played a central role in revealing that the spatial genome organization is not random and that temporal interactions govern many complex regulatory elements. Super-resolution microscopy methods are typically categorized into structured illumination microscopy, single-molecule localization microscopy, and stimulated emission-depletion microscopy (140–142). These optical techniques are especially useful because they offer multiplex tracking of molecules of interest, which can be used for biomolecular interaction studies (143). For example, super-resolution methods such as cryo-electron tomography and super-resolution light microscopy, as well as other highly effective imaging methods, including fluorescence anisotropy and fluorescence correlation spectroscopy, have made significant contributions in advancing our understanding of the chromatin organization and dynamics (144–150). Because these techniques rely on preserving chromatin integrity and require minimally invasive preparation methods, it is possible to investigate chromatin organization in live cells (143). The analysis of the dynamics of chromatin and its components under *in vivo* conditions was made possible by fluorescently tagging the histones in human osteosarcoma cells (151). Several key reviews provide an in-depth overview of the current optical detection methods for chromatin analysis (143,46).

As discussed above, one of the main advantages of light-based detection methods is the ability to capture changes at a molecular level. One such example is single-chromatin molecule analysis at the nanoscale, which is an affinity-based detection method using fluorescent tags to track changes in histones or survey epigenomic marks across genomes (e.g., diseased state versus healthy) (17,47,152). Moreover, light-based techniques offer an avenue to quantify relative epigenetic status and zoom in on specific genomic loci which could aid in biomarker studies (48). In this context, a ChIP-String method was used to characterize the chromatin state at several hundred positions in the genome using a pull-down of chromatin fragments labeled with 126 chromatin regulator antibodies, 17 histone-modification antibodies and two immunoglobulin G control antibodies (48). Downstream analysis was performed using an nCounter (a platform for detecting and counting large sets of molecules; NanoString Inc., Seattle, WA).

However, multiplexing studies can be difficult when using fluorescent probes because of the limited detection and emission range. Recently, an innovative method was devised to overcome this limitation.



Chromatin fragments with a nucleosome were attached to a microscope slide and the fluorescent antibody interaction with the tethered structure was used to evaluate different histone states. After probe data acquisition, nucleosomal DNA was sequenced on the slide to determine genomic location (49,50). The usefulness of the discussed methods lies in their ability to analyze unamplified DNA sequences. However, improving the accuracy of quantification and probe binding kinetics is still needed (17).

Another optical method, namely SPR, also holds promise for high-throughput multicomplex analyses. By using SPR to quantify binding affinities, it has been recently demonstrated that modified histone peptides and nucleic acids can be assessed using three-dimensional (3D) carbene-chip-based SPR imaging (SPRi). Even though the high-throughput nature of SPRi allows studying the interactome in vitro, it still has drawbacks, as it requires a “drying process” to minimize carbene blocking for immobilization. However, such treatment is not suitable for proteins that are sensitive to water loss, and the carbene-based immobilization strategy is thus not optimal (51).

Since the advent of super-resolution microscopy, it became evident that 3D folding has important consequences across the genome and epigenome. These new optical analysis methods were crucial in establishing the existence of a range of chromatin compaction states that are heterogeneous in nature (143). Despite some technical limitations, biophysics helps advancing imaging and quantification of various chromatin interactions at the earliest stages of DNA compaction. Thus, advancing our understanding of the chromatin and associated interactome dynamics can shed new light on disease mechanisms, ranging from oncological disorders to metabolic and immune pathologies (2).

### **New methods in biophysics for higher-order chromatin structure studies**

#### *The epigenomic landscape above the chromatin*

Chromatin ordering within the nuclear space is a complex process that must accommodate not only packaging of the genome but also proper gene expression and replication of DNA. Studies exploring the 3D organization of the genome and genomic loci interactions have challenged the long-established dogma of a stringent hierarchical organization. The transcriptional or chromatin state is responsible for the formation of small compartmental domains, and the different levels of compaction result from created inter- or intrachain contacts (153,154). These new insights encouraged the development of new analytical methodologies

that can capture transient events or new structural elements within the genome and epigenome.

When moving from sequences and nucleosomes to higher organizational units (Fig. 2), it becomes necessary to evaluate chromatin accessibility and conformation. This has been mostly achieved through microscopy-based techniques and specific molecular biology assays, such as chromatin immunoprecipitation or nuclease accessibility of DNA (155). The global chromatin accessibility analysis provided information on what interactors can be found around chromatin and what sequences are preferred. Furthermore, sequencing allowed prediction of DNA-structural differences and how different genomic regions are organized and how they change over time. For example, chromosome conformation capture (3C) (53) and subsequent evolutions of the method to 4C (54) and 5C (52) helped evaluate spatial changes in nuclear organization through gene expression. These techniques allow prediction of the three-dimensional conformation of chromatin by cross-linking chromatin and restrictively digesting the extracted DNA. Gene-specific PCR or other high-throughput next-generation sequencing methods help identify these DNA regions coming into close proximity. Thus, tracking such organizational changes can elucidate what genomic regions become activated (10,156).

An important organizational level to consider is chromatin domains, which consist of several different architectural elements from topologically associated domains (TADs), lamina-associated domains, and nucleolar-associated domains to long-range contacts, such as chromosome domains (77,78,157–159) (Fig. 2). Sequencing and genome-scale studies highlight that the nuclear environment is important in regulating gene expression and that it can also play a role in various diseases, including developmental disorders and cancer (57,160). Yet, going beyond sequences and inferred associations is still difficult, and the research is relying more and more on biophysical tools to capture structural and organizational aspects of higher-order chromatin organization (156).

#### *Physicochemical aspects of the higher-order epigenetic structures*

Studies on chromatin organization and the complex architecture involving multiple elements underscored the structural and functional plasticity of chromatin and its role in nuclear organization, gene expression, and protein-chromatin interactions (37,156,158). Monya Baker provided an excellent overview of how new data and better analytical approaches changed the understanding of the gene activity from a linear DNA sequence to chromatin structures that dictate gene expression (161).

Studying thermodynamic interactions can facilitate the exploration of epigenetic processes on a larger scale. The simplest interaction analysis can be performed with microscale thermophoresis (MST), which offers a rapid method to characterize molecular interactions between proteins and DNA requiring microliter volumes with low concentrations of interactors. MST relies on precise temperature gradients for thermophoretic movement depending on the size, charge, and hydration shell. Thus, binding and/or association can be tracked by quantifying the fluorescence of a labeled target molecule, for which fluorescence can be affected by alterations in the chemical microenvironment (55). Shubert and Längst reported that binding affinities (picomolar and millimolar scale) permit the analysis of the epigenetic interactions occurring between proteins and protein-DNA complexes (55). MST can capture even more biomolecular interactions, such as binding events of histone complexes or the influence of the nucleosomal milieu on the recognition of histone modifications (56).

#### *Higher-order structure analysis employing structural biophysics approaches*

Earlier efforts to determine protein complexes and chromatin architecture generally focused on x-ray crystallography, but it was mostly constrained to domains or subunits posing technical challenges in achieving a high resolution (Table 1) (71,72). Recent breakthroughs in cryo-EM as well as computational analysis allowed the visualization of chromatin-modifying or tethered protein (71,73,74). For example, a cryo-EM study on heterochromatin protein 1 (HP1), which interacts with H3K9me3 maintaining constitutive heterochromatin, allowed capturing the heterochromatin formation mediated by HP1 (119,68,69,162). Machida and co-workers demonstrated that HP1 bridges the adjacent nucleosomes via histone H3 contacts to promote pair formation and DNA regions between dinucleosomes are exposed to solvent without any clear signs of interaction with HP1. Cryo-EM studies also revealed nucleosome aggregation into 30 nm fiber via stacked nucleosomal arrays (72). Nevertheless, the debate about the existence and precise nature of the 30 nm *in vivo* is still ongoing (10,113,146,163,164). A recent study of 3D organization of chromatin *in vivo* in *Schizosaccharomyces pombe* demonstrated the existence of porous chromatin structure without any indication of 30-nm fiber (70). Therefore, exploring and documenting different states of the higher-order structures of chromatin can help build a conceptual framework for these structures at high resolution.

It is also important to note that chromatin domain analysis builds on long-range electrostatics, eluci-

dating how these forces change the chromatin fiber organization (165–167). An exciting study proposed a method to directly measure nucleosome electrostatics through ion counting (75). Inductively coupled plasma MS (BE-ICP MS) is used to determine the number of ionsphere ions around the molecule complex—in this case, a nucleosome. BE-ICP MS allows a precise measurement of the magnitude of the surrounding electrostatic field, and the authors were not only able to directly measure nucleosome electrostatics but also showed that although nucleosome formation leads to reduction of the complex charge, the negative electrostatic field around the nucleosome is nevertheless maintained. Thus, the polyelectrolyte nature of the nucleosome cannot be underestimated when considering DNA compaction and binding factor interactions.

Parallelization of multiple techniques has been recently demonstrated with integrative DNA and protein tagging platform. This approach combines biochemical enrichment using a bifunctional transposase-peroxidase with bioinformatic analysis of the genomic and proteomic profiles of open chromatin (76). Similarly, another recent study multiplexed the analysis by profiling DNA methylation and chromatin architecture in single cells; specifically, the investigation evaluated if distal DNA sequences can have a coordinated methylation through chromatin folding (168). To simultaneously capture DNA methylation and chromatin conformation, Li and co-workers (168) developed Methyl-HiC, a method for *in situ* high-throughput chromosome conformation capture (Hi-C) and whole-genome bisulfite sequencing. The authors demonstrated that their method accurately characterized both >80% of the methylome and the main architectural features of a chromosome in a biological sample. A very interesting direction was also taken in this study by designing single-cell Methyl-HiC to resolve the epigenetic heterogeneity across different cells.

#### *Theoretical and experimental biophysics offers novel insights for 3D epigenome studies*

As discussed above, 3D genome organization can affect gene expression and change how other proteins access genomic regions. Although it might still be difficult to fully capture the complex nature of chromatin packing, biophysical modeling *in silico* allows expanding our theoretical and conceptual understanding of these processes.

Using Monte Carlo simulations, Yanao et al. identified the importance of asymmetric bend-writhe elasticity of DNA when higher-order structures are formed, providing insights into DNA wrapping around nucleosome cores (169). However, further investigations are

required to refine the models and take into consideration other parameters, such as more degrees of freedom for DNA movements as well as twisting, stretching, and occurring rigidities. Another study employed a Monte Carlo simulation to describe the nucleosome geometry and impact of repeat length on chromatin fiber compaction (170). It is postulated that nucleosomal units form higher-order chromatin structures referred as “30 nm chromatin fiber,” and their physicochemical properties depend on the mechanical forces and ionic conditions (114). More complex models were developed to explore energetically feasible conformations of the fibers, in which it was shown that the level of DNA supercoiling depends on the length of the internucleosome linker in the chromatin fiber (118). Norouzi and co-authors hypothesized that the topological polymorphism of chromatin fibers may regulate transcription, as different levels of DNA supercoiling occur around RNA polymerase. The authors were successful in confirming this prediction, with their developed model assessing genome-wide nucleosome repeat length distribution in active and silent yeast genes. The chromatin environment, containing numerous polyions in the form of DNA, histones, and ions, is also dependent on the electrostatic interactions that govern the structural interplay. Nikolay Korolev et al. described the necessary “electrostatic conditions” to achieve a nucleosome and chromatin organization (171). The reviewed electrostatic rules provide important insights into the role of acidic domains found in nuclear proteins, such as nucleoplasmin, high-mobility group proteins, and histone chaperones, as well as how other PTMs alter chromatin structure and dynamics (171). It is important to note that current theoretical models, for example, Poisson-Boltzmann mean-field calculations or all-atom models, show little agreement on nucleosome electrostatics (114,172,173) and remain to be experimentally tested.

### *3D imaging of the epigenome*

As we move past the primary epigenomic organization units, we might ask how structural components in chromatin can be imaged and tracked at a larger scale. New imaging-based 3D genomics techniques made this possible by tracing chromatin folding or nucleosome architecture. Spatial epigenome analysis to identify genomic contact regions started with sequencing-based Hi-C. TADs were one of the structures identified with Hi-C (57,58). Similar sequencing-based efforts led to further identification of lamina-associated domains and nucleolus-associated chromosomal domains (57,58).

Although light microscopy could not resolve higher-order chromatin structures, confocal microscopy

brought better resolution, for example, enabling the tracking of fluorescent DNA strand assembly into chromosomes (156,59). For example, confocal microscopy in combination with immunohistochemistry was used to track chromatin territories and subdomains (60). Zwettler and co-authors combined structured illumination microscopy and different expansion microscopy, as well as magnified analysis of the proteome to investigate the molecular organization of the synaptonemal complex, which is responsible for proper synapsis, i.e., recombination and segregation of homologous chromosomes (61). Although single-molecule localization microscopy reached ultrastructure levels, a new image processing software needed to be developed (61). Another macromolecular complex study used fluorescently labeled trinucleosome arrays via fluorescence resonance energy transfer to explore chromatin structure effects on the DNA access. The authors were able to define two intermediate conformational states and the microscopic rate constants for chromatin fiber spontaneously opening to DNA-processing protein complexes (65).

Chromatin tracing was introduced to track the 3D folding of chromatin using multiplexed fluorescence in situ hybridization (63). A recent review by Hu and Wang details the development and challenges of chromatin tracing (64). They discuss how the correct 3D genome organization is essential for its function and how image-based 3D genomics techniques allow the direct tracing of chromatin folding and capture of the nucleome architecture. Subcellular structures and single-molecule dynamics can also be imaged using single-molecule localization microscopy. Indeed, this technique was employed to investigate damage to DNA and assess molecular structures during repair (62). Another research team combined 3D DNA points accumulation for imaging in nanoscale topography with spinning disk confocal microscopy (66). The authors achieved multiplexed 3D super-resolution with sample depths up to  $\sim 10 \mu\text{m}$  as well as resolution of up to 20 nm planar and 80 nm axial. A review summarizing single-molecule optical genome mapping (67) provides an in-depth overview of the current methods used for detecting large-scale genomic rearrangements and simultaneous mapping of multiple genomic layers. Diverse applications of the optical techniques are discussed in the context of epigenomics, in which they can be used for bacterial strain identification, detection of genes responsible for antibiotic resistance, or elucidation of new disease-relevant structural variations.

Imaging of the chromatin architecture is moving forward at a fast pace and addressing technical challenges creates new research opportunities. Large volumes of imaging data and complex analyses also demonstrate why computational techniques and

algorithms are becoming one of the central pieces in epigenetic research.

### **Bioinformatics, ML, and AI: harmonizing biophysical observations with high-throughput computing**

As introduced earlier, technical advancements rely on the ability to evaluate the growing number of complex data sets and draw meaningful conclusions. Thus, computational algorithms have been increasingly employed to model and extrapolate experimental data. Bioinformatics represents a fast-growing field in multiomics integration (8), and it is worth highlighting that novel computational pipelines are becoming more available, allowing integration of multiomics data spanning genome, epigenome, transcriptome, proteome, metabolome, and even microbiome (8,22,174–177). Moreover, ML, AI, and deep learning (DL) have also found their applications in epigenetic analyses.

The high dimensionality of data requires new statistical and algorithmic approaches. A recent study aimed to analyze deep features of DNA methylation by designing a deep neural network that was based on several stacked binary restricted Boltzmann machines to learn the low-dimensional deep features of DNA methylation. The method's self-learning ability extracts the low-dimensional features allowing the separation of the normal and pathological samples. The authors claim that compared to other probabilistic mixture model-based methods, their deep neural network-based method performs significantly better when comparing error rates for sample differentiation (178). Another interesting application of DL was to study 6mA. Despite the large volume of data generated for this DNA signature in eukaryotes, the role of DNA 6mA remains elusive. To address the biotechnological limitations of the analysis of this base modification, a DL-based algorithm has been developed for predicting de novo DNA 6mA (179). ML analyses of methylation profiles have also been applied in plant studies to uncover tissue-specific expression patterns. Another study employed a mixture of six ML algorithms through "VotingClassifier," in which each data point is subjected to each ML model predictions. The predictions shared by the majority of models are used to form the final prediction (180). In the context of application development, the DL framework MethylNet was made accessible to epigenetics researchers to identify relevant CpG and perform such tasks as cell-type deconvolution, cancer subtype classification, or age regression (age estimation) (181).

DNA base modifications such as 5mC and 6mA have also been explored using a bidirectional recurrent neural network with long short-term memory to detect DNA

modifications from short-read bisulfite sequencing as well as long-read PacBio sequencing (182). The developed method, named DeepMod, was evaluated on three types of genomes (*E. coli*, *Chlamydomonas reinhardtii*, and *Homo sapiens*), showing average precision up to 0.99 for both synthetically introduced and naturally occurring 5mC modifications; slightly lower average precision (~0.9) was reached for 6mA on *E. coli* data. In another study, DeepSignal (DL method) was developed to detect 6mA and 5mC (*H. sapiens*, *E. coli*, and pUC19) from Nanopore sequencing reads at a read and genome level (183). DeepSignal consists of four modules in which two modules are employed to construct features from raw electrical signals of Nanopore reads using a convolutional neural network (CNN) (signal feature) and bidirectional recurrent neural network (sequence feature). Constructed features from these reads are then fed into a fully connected neural network to predict the methylation states (182). Although DL can be successfully applied to a genome-scale detection of DNA modifications, a successful design of the neural network architecture still requires in-depth testing and comparative studies to facilitate the epigenetic analysis. An excellent review by Xu and Seki discusses recent advances in the detection of base modifications using the Nanopore sequencer and summarizes associated statistical tests as well as ML methods (184). They also comment on the applications of Nanopore technology to study open chromatin, DNA replication, and RNA metabolism.

Computational microscopy and associated DL methods have been increasingly employed to compensate resolution limits, to reduce artifacts, or for image analysis (185–188). For example, high-density photoactivated localization microscopy and DL were used to gain insight into the spatiotemporal organization of the genome by capturing resolved super-resolution images of chromatin in living cells (189). A computational chromatin model for elongated ~45–90 nm wide chromatin structures or "blobs" was showed to form dynamically associating chromatin fragments. The devised Deep-photoactivated localization microscopy (or DL-based photoactivated localization microscopy) method used CNN and previously developed methods to obtain super-resolution images from stochastically blinking emitters, such as fluorescent molecules (189,190). Computational and experimental work demonstrated that the chromatin structure and dynamics are closely linked, and this may play a role in making the regions with a high local chromatin concentration more accessible (189). Chromatin architecture was also explored using a CNN-based approach to evaluate links to the transcriptional state of individual cells (188).

Connecting different layers of biological information, such as the epigenetic markers and phenotype, might



be challenging, but new bioinformatics models provide a means to map the intricate regulatory networks and their effects. For example, epigenome-wide association studies (EWASs) focus on elucidating the interactions between observable characteristics and the epigenome. EWASs require a rigorous study design in which statistical significance and power, confounding factors, and statistical assumptions need to be carefully considered (2,191). One of the key obstacles in EWASs is the epigenetic variability between different types of cells, which might reduce our ability to recover relevant biological information (2,192,193). To obtain accurate estimates of cell-type composition in a sample, a new statistical approach, factored spectrally transformed linear mixed model “EWASher” (FaST-LMM-EWASher) was developed. FaST-LMM-EWASher automatically corrects for cell-type composition without prior knowledge of the distribution (194). The model uses a linear mixed model (195) combined with principal components to compute the methylome similarity between every pair of samples. The deduced similarities are then used as the covariance in the mixed model to approximate cell-type composition.

ML has been applied not only in experimental analyses but also in the clinics. The dramatic growth in big data processing capabilities permits the processing of clinical data sets, and there are exciting opportunities to enrich general clinical readouts with epigenetic biomarker data. A review by Rauschert and colleagues offers insights into how ML models can be used to utilize DNA methylation information to diagnose disease states related to cancers, obesity, neurodevelopmental syndromes, and cardiovascular pathologies (196). However, the main challenge is to acquire reliable, organized, and longitudinal patient data to increase the predictive power. Moreover, information-rich patient data might also push for better regulations in how algorithms are selected and benchmarked to ensure that the outputs can be trusted (3,180,196,197). For example, epigenetic biomarker (<sup>m</sup>SEPT9 biomarker for colorectal cancer) was approved by the US Food and Drug Administration for the use in a diagnostic kit for blood-plasma-based analysis (198).

Considering all the progress in epigenetics, it becomes apparent that the epigenetic marks exploration cannot rely on just experimental observations. It is also important to integrate the generated data with epigenetic control levels and their interactions. In addition, the enthusiasm in ML and AI and big data requires rigorous testing to ensure that no “black box” situations are created in which it is hard to evaluate the conclusions. Better advocacy for regulatory oversight is also needed to ensure that the market is not flooded with tools that have very little value for making a real scientific and/or clinical impact.

## Future perspectives

The growing number of biophysical tools used to analyze the various levels of epigenetic processes offer an unprecedented insight into this regulatory network. Moreover, biophysical techniques in epigenetics expand qualitative and semiquantitative analyses into the exploration of structural interactions, chromatin architecture elements, or binding events. The growing computational power as well as multidisciplinary approaches will likely connect biophysical wet-lab tools with computational modeling. Moreover, by expanding our understanding of how the epigenome regulatory network is organized, we can develop better therapeutic solutions. The first steps to prepare high-throughput assays in epigenetics reflect the need for clinical application innovation to detect biomarkers and stratify patients (3,100,199). Bridging the gap between bench analyses and practical applications will facilitate future advancements in epigenetics.

## Image preparation and text mining methods

PubMed text mining (>30 million records) (200) and graph and word cloud rendering were performed using R programming language (v 4.0) in RStudio environment (201); the packages used for data mining and integration were easyPubMed (<https://cran.r-project.org/web/packages/easyPubMed/index.html>) and Rentrez (<https://cran.r-project.org/web/packages/rentrez/index.html>). Biorender (202) platform was used to generate the representation of the epigenome architecture.

## AUTHOR CONTRIBUTIONS

A.K. conceptualized and wrote the manuscript. G.S.C. and D.W. reviewed and edited the manuscript, as well as providing guidance. N.B. and A.M. provided clinical research and molecular biology comments and writing suggestions.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## REFERENCES

1. Aristizabal, M. J., I. Anreiter, ..., K. J. O'Donnell. 2020. Biological embedding of experience: a primer on epigenetics. *Proc. Natl. Acad. Sci. U S A.* 117:23261–23269.
2. Tost, J. 2020. 10 years of epigenomics: a journey with the epigenetic community through exciting times. *Epigenomic.* 12:81–85.
3. Cazaly, E., J. Saad, ..., J. Tang. 2019. Making sense of the epigenome using data integration approaches. *Front. Pharmacol.* 10:126.
4. Moosavi, A., and A. M. Ardekani. 2016. Role of epigenetics in biology and human diseases. *Iran. Biomed. J.* 20:246–258.



5. Attar, N. 2012. The allure of the epigenome. *Genome Biol.* 13:419.
6. Bird, A. 2007. Perceptions of epigenetics. *Nature.* 447:396–398.
7. Meissner, A. 2012. What can epigenomics do for you? *Genome Biol.* 13:420.
8. Li, Y. 2021. Modern epigenetics methods in biological research. *Methods.* 187:104–113.
9. Harrison, A., and A. Parle-McDermott. 2011. DNA methylation: a timeline of methods and applications. *Front. Genet.* 2:74.
10. Li, G., and D. Reinberg. 2011. Chromatin higher-order structures and gene regulation. *Curr. Opin. Genet. Dev.* 21:175–186.
11. Flusberg, B. A., D. R. Webster, ..., S. W. Turner. 2010. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods.* 7:461–465.
12. Shim, J., G. I. Humphreys, ..., R. Bashir. 2013. Detection and quantification of methylation in DNA using solid-state nanopores. *Sci. Rep.* 3:1389.
13. Gilboa, T., C. Torfstein, ..., A. Meller. 2016. Single-molecule DNA methylation quantification using electro-optical sensing in solid-state nanopores. *ACS Nano.* 10:8861–8870.
14. Simpson, J. T., R. E. Workman, ..., W. Timp. 2017. Detecting DNA cytosine methylation using nanopore sequencing. *Nat. Methods.* 14:407–410.
15. Smallwood, S. A., H. J. Lee, ..., G. Kelsey. 2014. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat. Methods.* 11:817–820.
16. Abramov, G., A. Velyvis, ..., L. E. Kay. 2020. A methyl-TROSY approach for NMR studies of high-molecular-weight DNA with application to the nucleosome core particle. *Proc. Natl. Acad. Sci. U S A.* 117:12836–12846.
17. Heck, C., Y. Michaeli, ..., Y. Eisenstein. 2019. Analytical epigenetics: single-molecule optical detection of DNA and histone modifications. *Curr. Opin. Biotechnol.* 55:151–158.
18. Heck, C., Y. Kanehira, ..., I. Bald. 2018. Placement of single proteins within the SERS hot spots of self-assembled silver nanolenses. *Angew. Chem. Int. Ed. Engl.* 57:7444–7447.
19. Zrimsek, A. B., N. Chiang, ..., R. P. V. Duyne. 2017. Single-molecule chemistry with surface- and tip-enhanced Raman spectroscopy. *Chem. Rev.* 117:7583–7613.
20. Barhoumi, A., and N. J. Halas. 2011. Detecting chemically modified DNA bases using surface-enhanced Raman spectroscopy. *J. Phys. Chem. Lett.* 2:3118–3123.
21. Guerrini, L., Ž. Krpetić, ..., D. Graham. 2015. Direct surface-enhanced Raman scattering analysis of DNA duplexes. *Angew. Chem. Int. Ed. Engl.* 54:1144–1148.
22. Ziller, M. J., K. D. Hansen, ..., M. J. Aryee. 2015. Coverage recommendations for methylation analysis by whole-genome bisulfite sequencing. *Nat. Methods.* 12:230–232, 1, 232.
23. Lentini, A., C. Lagerwall, ..., C. E. Nestor. 2018. A reassessment of DNA-immunoprecipitation-based genomic profiling. *Nat. Methods.* 15:499–504.
24. Skvortsova, K., E. Zotenko, ..., C. Stirzaker. 2017. Comprehensive evaluation of genome-wide 5-hydroxymethylcytosine profiling approaches in human DNA. *Epigenetics Chromatin.* 10:16.
25. Thomson, J. P., J. M. Hunter, ..., R. R. Meehan. 2013. Comparative analysis of affinity-based 5-hydroxymethylation enrichment techniques. *Nucleic Acids Res.* 41:e206.
26. van Gurp, T. P., N. C. A. M. Wagemaker, ..., K. J. F. Verhoeven. 2016. epiGBS: reference-free reduced representation bisulfite sequencing. *Nat. Methods.* 13:322–324.
27. Molinie, B., J. Wang, ..., C. C. Giallourakis. 2016. m(6)A-LAIC-seq reveals the census and complexity of the m(6)A epitranscriptome. *Nat. Methods.* 13:692–698.
28. Chew, L. F., E. T. Courtois, ..., W. F. Burkholder. 2016. Single-cell multimodal profiling reveals cellular epigenetic heterogeneity. *Nat. Methods.* 13:833–836.
29. Carrascosa, L. G., A. A. I. Sina, ..., M. Trau. 2014. Molecular inversion probe-based SPR biosensing for specific, label-free and real-time detection of regional DNA methylation. *Chem. Commun. (Camb.).* 50:3585–3588.
30. Song, C. X., J. Diao, ..., S. R. Quake. 2016. Simultaneous single-molecule epigenetic imaging of DNA methylation and hydroxymethylation. *Proc. Natl. Acad. Sci. USA.* 113:4338–4343.
31. Fernandez, A. F., L. Valledor, ..., M. F. Fraga. 2018. Quantification of global DNA methylation levels by mass spectrometry. *Methods Mol. Biol.* 1708:49–58.
32. Ehrich, M., M. R. Nelson, ..., D. van den Boom. 2005. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc. Natl. Acad. Sci. U S A.* 102:15785–15790.
33. Lin, X. C., T. Zhang, ..., J. H. Jiang. 2016. Mass spectrometry based ultrasensitive DNA methylation profiling using target fragmentation assay. *Anal. Chem.* 88:1083–1087.
34. Kimura, H. 2013. Histone modifications for human epigenome analysis. *J. Hum. Genet.* 58:439–445.
35. He, H. H., C. A. Meyer, ..., M. Brown. 2014. Refined DNase-seq protocol and data analysis reveals intrinsic bias in transcription factor footprint identification. *Nat. Methods.* 11:73–78.
36. Vierstra, J., H. Wang, ..., J. A. Stamatoyannopoulos. 2014. Coupling transcription factor occupancy to nucleosome architecture with DNase-FLASH. *Nat. Methods.* 11:66–72.
37. Ekundayo, B., T. J. Richmond, and T. Schalch. 2017. Capturing structural heterogeneity in chromatin fibers. *J. Mol. Biol.* 429:3031–3042.
38. Schalch, T., S. Duda, ..., T. J. Richmond. 2005. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature.* 436:138–141.
39. Tan, S., and C. A. Davey. 2011. Nucleosome structural studies. *Curr. Opin. Struct. Biol.* 21:128–136.
40. Boopathi, R., R. Danev, ..., J. Bednar. 2020. Phase-plate cryo-EM structure of the Widom 601 CENP-A nucleosome core particle reveals differential flexibility of the DNA ends. *Nucleic Acids Res.* 48:5735–5748.
41. Shi, X., C. Prasanna, ..., L. Nordenskiöld. 2018. Structure and dynamics in the nucleosome revealed by solid-state NMR. *Angew. Chem. Int. Ed. Engl.* 57:9734–9738.
42. Moriwaki, Y., T. Yamane, ..., Y. Nishimura. 2016. Solution structure of the isolated histone H2A-H2B heterodimer. *Sci. Rep.* 6:24999.
43. Gao, M., P. S. Nadaud, ..., C. P. Jaroniec. 2013. Histone H3 and H4 N-terminal tails in nucleosome arrays at cellular concentrations probed by magic angle spinning NMR spectroscopy. *J. Am. Chem. Soc.* 135:15278–15281.
44. Zhou, B. R., J. Jiang, ..., Y. Bai. 2015. Structural mechanisms of nucleosome recognition by linker histones. *Mol. Cell.* 59:628–638.
45. Saikusa, K., A. Osakabe, ..., S. Akashi. 2018. Structural diversity of nucleosomes characterized by native mass spectrometry. *Anal. Chem.* 90:8217–8226.
46. Boettiger, A., and S. Murphy. 2020. Advances in chromatin imaging at kilobase-scale resolution. *Trends Genet.* 36:273–287.
47. Hyun, B. R., J. L. McElwee, and P. D. Soloway. 2015. Single molecule and single cell epigenomics. *Methods.* 72:41–50.
48. Ram, O., A. Goren, ..., B. E. Bernstein. 2011. Combinatorial patterning of chromatin regulators uncovered by genome-wide location analysis in human cells. *Cell.* 147:1628–1639.
49. Harris, T. D., P. R. Buzby, ..., Z. Xie. 2008. Single-molecule DNA sequencing of a viral genome. *Science.* 320:106–109.

50. Shema, E., D. Jones, ..., B. E. Bernstein. 2016. Single-molecule decoding of combinatorially modified nucleosomes. *Science*. 352:712–721.
51. Zhao, S., M. Yang, ..., H. Li. 2017. Kinetic and high-throughput profiling of epigenetic interactions by 3D-carbene chip-based surface plasmon resonance imaging technology. *Proc. Natl. Acad. Sci. U S A*. 114:E7245–E7254.
52. Dostie, J., T. A. Richmond, ..., J. Dekker. 2006. Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res*. 16:1299–1309.
53. Gavrilov, A., E. Eivazova, ..., Y. Vassetzky. 2009. Chromosome conformation capture (from 3C to 5C) and its ChIP-based modification. *Methods Mol. Biol.* 567:171–188.
54. Ohlsson, R., and A. Göndör. 2007. The 4C technique: the 'Rosetta stone' for genome biology in 3D? *Curr. Opin. Cell Biol.* 19:321–325.
55. Schubert, T., and G. Längst. 2015. Studying epigenetic interactions using MicroScale Thermophoresis (MST). *AIMS Biophys.* 2:370–380.
56. Corbeski, I., V. Horn, ..., H. van Ingen. 2018. Microscale thermophoresis analysis of chromatin interactions. *Methods Mol. Biol.* 1837:177–197.
57. Bonev, B., and G. Cavalli. 2016. Organization and function of the 3D genome. *Nat. Rev. Genet.* 17:661–678.
58. Gibcus, J. H., and J. Dekker. 2013. The hierarchy of the 3D genome. *Mol. Cell*. 49:773–782.
59. Manders, E. M. M., H. Kimura, and P. R. Cook. 1999. Direct imaging of DNA in living cells reveals the dynamics of chromosome formation. *J. Cell Biol.* 144:813–821.
60. Visser, A. E., R. Eils, ..., J. A. Aten. 1998. Spatial distributions of early and late replicating chromatin in interphase chromosome territories. *Exp. Cell Res.* 243:398–407.
61. Zwettler, F. U., M. C. Spindler, ..., M. Sauer. 2020. Tracking down the molecular architecture of the synaptonemal complex by expansion microscopy. *Nat. Commun.* 11:3222.
62. Hausmann, M., C. Neitzel, ..., M. Falk. 2020. Single molecule localization microscopy analyses of DNA-repair foci and clusters detected along particle damage tracks. *Front. Phys.* 8:1.
63. Wang, S., J. H. Su, ..., X. Zhuang. 2016. Spatial organization of chromatin domains and compartments in single chromosomes. *Science*. 353:598–602.
64. Hu, M., and S. Wang. 2021. Chromatin tracing: imaging 3D genome and nucleome. *Trends Cell Biol.* 31:5–8.
65. Poirier, M. G., E. Oh, ..., J. Widom. 2009. Dynamics and function of compact nucleosome arrays. *Nat. Struct. Mol. Biol.* 16:938–944.
66. Schueder, F., J. Lara-Gutiérrez, ..., R. Jungmann. 2017. Multiplexed 3D super-resolution imaging of whole cells using spinning disk confocal microscopy and DNA-PAINT. *Nat. Commun.* 8:2090.
67. Jeffet, J., S. Margalit, ..., Y. Ebnstein. 2021. Single-molecule optical genome mapping in nanochannels: multidisciplinary at the nanoscale. *Essays Biochem.* 65:51–66.
68. Saksouk, N., E. Simboeck, and J. Déjardin. 2015. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin*. 8:3.
69. Machida, S., Y. Takizawa, ..., H. Kurumizaka. 2018. Structural basis of heterochromatin formation by human HP1. *Mol. Cell*. 69:385–397.e8.
70. Cai, S., C. Chen, Z. Y. Tan, Y. Huang, J. Shi, and L. Gan. 2018. Cryo-ET reveals the macromolecular reorganization of *S. pombe* mitotic chromosomes in vivo. *Proc Natl Acad Sci U S A [Internet]*. 115:10977–10982.
71. Merk, A., A. Bartesaghi, ..., S. Subramaniam. 2016. Breaking cryo-EM resolution barriers to facilitate Drug discovery. *Cell*. 165:1698–1707.
72. Song, F., P. Chen, ..., G. Li. 2014. Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* 376–380.
73. Grant, T., A. Rohou, and N. Grigorieff. 2018. cisTEM, user-friendly software for single-particle image processing. *eLife*. 7:e35383.
74. Punjani, A., J. L. Rubinstein, ..., M. A. Brubaker. 2017. cryo-SPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods*. 14:290–296.
75. Gebala, M., S. L. Johnson, ..., D. Herschlag. 2019. Ion counting demonstrates a high electrostatic field generated by the nucleosome. *eLife*. 8:e44993.
76. Lee, J. D., J. A. Paulo, ..., P. P. Pandolfi. 2021. Dual DNA and protein tagging of open chromatin unveils dynamics of epigenomic landscapes in leukemia. *Nat. Methods*. 18:293–302.
77. Harabula, I., and A. Pombo. 2021. The dynamics of chromatin architecture in brain development and function. *Curr. Opin. Genet. Dev.* 67:84–93.
78. Magaña-Acosta, M., and V. Valadez-Graham. 2020. Chromatin remodelers in the 3D nuclear compartment. *Front. Genet.* 11:600615.
79. Moore, L. D., T. Le, and G. Fan. 2013. DNA methylation and its basic function. *Neuropsychopharmacology*. 38:23–38.
80. Wang, S., and W. Wu. 2018. DNA methylation alterations in human cancers. In *Epigenetics in Human Disease*. T. O. Tollefsbol, ed. Elsevier, pp. 109–139.
81. Klungland, A., and A. B. Robertson. 2017. Oxidized C5-methyl cytosine bases in DNA: 5-Hydroxymethylcytosine; 5-formylcytosine; and 5-carboxycytosine. *Free Radic. Biol. Med.* 107:62–68.
82. Kumar, S., V. Chinnusamy, and T. Mohapatra. 2018. Epigenetics of modified DNA bases: 5-methylcytosine and beyond. *Front Genet.* 9:640.
83. Sood, A. J., C. Viner, and M. M. Hoffman. 2019. DNAmdb: the DNA modification database. *J. Cheminform.* 11:30.
84. Breiling, A., and F. Lyko. 2015. Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond. *Epigenetics Chromatin*. 8:24.
85. Charlet, J., C. E. Duymich, ..., P. A. Jones. 2016. Bivalent regions of cytosine methylation and H3K27 acetylation suggest an active role for DNA methylation at enhancers. *Mol. Cell*. 62:422–431.
86. Chen, C.-C., K.-Y. Wang, and C.-K. J. Shen. 2013. 5-methylcytosine demethylation activities of the mammalian DNA methyltransferases. *J. Biol. Chem.* 288:9084–9091.
87. Maeder, M. L., J. F. Angstman, ..., J. K. Joung. 2013. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nat. Biotechnol.* 31:1137–1142.
88. Wang, K. Y., C. C. Chen, and C. K. Shen. 2014. Active DNA demethylation of the vertebrate genomes by DNA methyltransferases: deaminase, dehydroxymethylase or demethylase? *Epigenomics*. 6:353–363.
89. Jabbari, K., and G. Bernardi. 2004. Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. *Gene*. 333 (Suppl):143–149.
90. Ehrlich, M., M. A. Gama-Sosa, ..., C. Gehrke. 1982. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells. *Nucleic Acids Res.* 10:2709–2721.
91. Ficz, G. 2015. New insights into mechanisms that regulate DNA methylation patterning. *J. Exp. Biol.* 218:14–20.
92. Jones, P. A. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* 13:484–492.

93. Moarefi, A. H., and F. Chédin. 2011. ICF syndrome mutations cause a broad spectrum of biochemical defects in DNMT3B-mediated de novo DNA methylation. *J. Mol. Biol.* 409:758–772.
94. Farthing, C. R., G. Ficiz, ..., W. Reik. 2008. Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. *PLoS Genet.* 4:e1000116.
95. Robertson, K. D. 2005. DNA methylation and human disease. *Nat. Rev. Genet.* 6:597–610.
96. Jin, Z., and Y. Liu. 2018. DNA methylation in human diseases. *Genes Dis.* 5:1–8.
97. Kurdyukov, S., and M. Bullock. 2016. DNA methylation analysis: choosing the right method. *Biology (Basel).* 5:3.
98. Yong, W.-S., F.-M. Hsu, and P.-Y. Chen. 2016. Profiling genome-wide DNA methylation. *Epigenetics Chromatin.* 9:26.
99. Li, L., B. F. Chen, and W. Y. Chan. 2015. An epigenetic regulator: Methyl-CpG-binding domain protein 1 (MBD1). *Int. J. Mol. Sci.* 16:5125–5140.
100. Costantino, S., P. Libby, ..., F. Paneni. 2018. Epigenetics and precision medicine in cardiovascular patients: from basic concepts to the clinical arena. *Eur. Heart J.* 39:4150–4158.
101. Kanherkar, R. R., S. E. Stair, ..., A. B. Csoka. 2017. Epigenetic mechanisms of integrative medicine. *Evid. Based Complement. Alternat. Med.* 2017:4365429.
102. Angermueller, C., S. J. Clark, ..., W. Reik. 2016. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat. Methods.* 13:229–232.
103. Cheow, L. F., S. R. Quake, ..., D. M. Messerschmidt. 2015. Multiplexed locus-specific analysis of DNA methylation in single cells. *Nat. Protoc.* 10:619–631.
104. Xia, Y., L. Wu, ..., J. Wang. 2019. Sensitive surface plasmon resonance detection of methyltransferase activity and screening of its inhibitors amplified by p53 protein bound to methylation-specific ds-DNA consensus sites. *Biosens. Bioelectron.* 126:269–274.
105. Nazmul Islam, M., S. Yadav, ..., M. J. A. Shiddiky. 2017. Optical biosensing strategies for DNA methylation analysis. *Biosens. Bioelectron.* 92:668–678.
106. Gabrieli, T., H. Sharim, ..., Y. Eisenstein. 2018. Epigenetic optical mapping of 5-hydroxymethylcytosine in nanochannel arrays. *ACS Nano.* 12:7148–7158.
107. Římal, V., O. Socha, ..., H. Štěpánková. 2015. Spectroscopic study of cytosine methylation effect on thermodynamics of DNA duplex containing CpG motif. *J. Spectrosc.* 2015:1–8.
108. Hörberg, J., and A. Reymer. 2018. A sequence environment modulates the impact of methylation on the torsional rigidity of DNA. *Chem. Commun. (Camb.).* 54:11885–11888.
109. Temiz, N. A., D. E. Donohue, ..., J. R. Collins. 2012. The role of methylation in the intrinsic dynamics of B- and Z-DNA. *PLoS One.* 7:35558.
110. Taqi, M. M., S. K. T. S. Wärmländer, ..., G. Bakalkin. 2012. Conformation effects of CpG methylation on single-stranded DNA oligonucleotides: analysis of the opioid peptide dynorphin-coding sequences. *PLoS One.* 7:e39605.
111. Manduzio, H., A. Martelet, ..., F. Fenaille. 2010. Comparison of approaches for purifying and desalting polymerase chain reaction products prior to electrospray ionization mass spectrometry. *Anal. Biochem.* 398:272–274.
112. Null, A. P., L. T. George, and D. C. Muddiman. 2002. Evaluation of sample preparation techniques for mass measurements of PCR products using ESI-FT-ICR mass spectrometry. *J. Am. Soc. Mass Spectrom.* 13:338–344.
113. Felsenfeld, G., and J. D. McGhee. 1986. Structure of the 30 nm chromatin fiber. *Cell.* 44:375–377.
114. Schiessel, H. 2003. The physics of chromatin. *J. Phys. Condens. Matter.* 15:R699–R774.
115. Kaplan, N., I. K. Moore, ..., E. Segal. 2009. The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature.* 458:362–366.
116. Kouzarides, T. 2007. Chromatin modifications and their function. *Cell.* 128:693–705.
117. Lowary, P. T., and J. Widom. 1997. Nucleosome packaging and nucleosome positioning of genomic DNA. *Proc. Natl. Acad. Sci. U S A.* 94:1183–1188.
118. Norouzi, D., A. Katebi, ..., V. B. Zhurkin. 2015. Topological diversity of chromatin fibers: interplay between nucleosome repeat length, DNA linking number and the level of transcription. *AIMS Biophys.* 2:613–629.
119. Vakoc, C. R., S. A. Mandat, ..., G. A. Blobel. 2005. Histone H3 lysine 9 methylation and HP1 $\gamma$  are associated with transcription elongation through mammalian chromatin. *Mol. Cell.* 19:381–391.
120. Ngo, V., Z. Chen, ..., W. Wang. 2019. Epigenomic analysis reveals DNA motifs regulating histone modifications in human and mouse. *Proc. Natl. Acad. Sci. U S A.* 116:3668–3677.
121. Castillo, J., G. López-Rodas, and L. Franco. 2017. Histone post-translational modifications and nucleosome organisation in transcriptional regulation: Some open questions. *Adv. Exp. Med. Biol.* 966:65–92.
122. Taylor, B. C., and N. L. Young. 2021. Combinations of histone post-translational modifications. *Biochem. J.* 478:511–532.
123. Jenuwein, T., and C. D. Allis. 2001. Translating the histone code. *Science.* 293:1074–1080.
124. Materese, C. K., A. Savelyev, and G. A. Papoian. 2009. Counterion atmosphere and hydration patterns near a nucleosome core particle. *J. Am. Chem. Soc.* 131:15005–15013.
125. Moshkin, Y. M. 2015. Chromatin-A global buffer for eukaryotic gene control. *AIMS Biophys.* 2:531–554.
126. Davis, I. J., and S. G. Pattenden. 2018. Chromatin accessibility as a strategy to detect changes associated with development, disease, and exposure and susceptibility to chemical toxins. In *Toxicogenetics: Core Principles and Applications*. S. D. McCullough and D. C. Dolinoy, eds. Elsevier, pp. 85–103.
127. Mirabella, A. C., B. M. Foster, and T. Bartke. 2016. Chromatin deregulation in disease. *Chromosoma.* 125:75–93.
128. Cao, Z., C. Chen, ..., C. Lu. 2015. A microfluidic device for epigenomic profiling using 100 cells. *Nat. Methods.* 12:959–962.
129. Nguyen, U. T. T., L. Bittova, ..., T. W. Muir. 2014. Accelerated chromatin biochemistry using DNA-barcoded nucleosome libraries. *Nat. Methods.* 11:834–840.
130. Kasinathan, S., G. A. Orsi, ..., S. Henikoff. 2014. High-resolution mapping of transcription factor binding sites on native chromatin. *Nat. Methods.* 11:203–209.
131. Henikoff, J. G., J. A. Belsky, ..., S. Henikoff. 2011. Epigenome characterization at single base-pair resolution. *Proc. Natl. Acad. Sci. USA.* 108:18318–18323.
132. Makowski, M. M., C. Gräwe, ..., M. Vermeulen. 2018. Global profiling of protein-DNA and protein-nucleosome binding affinities using quantitative mass spectrometry. *Nat. Commun.* 9:1653.
133. Callaway, E. 2020. Revolutionary cryo-EM is taking over structural biology. *Nature.* 578:201.
134. Li, W., W. Tian, ..., Z. Wang. 2021. Molecular basis of nucleosomal H3K36 methylation by NSD methyltransferases. *Nature.* 590:498–503.
135. Davey, C. A., D. F. Sargent, ..., T. J. Richmond. 2002. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* 319:1097–1113.
136. Luger, K., A. W. Mäder, ..., T. J. Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature.* 389:251–260.



137. Richmond, T. J., J. T. Finch, ..., A. Klug. 1984. Structure of the nucleosome core particle at 7 Å resolution. *Nature*. 311:532–537.
138. Chua, E. Y. D., V. K. Vogirala, ..., S. Sandin. 2016. 3.9 Å structure of the nucleosome core particle determined by phase-plate cryo-EM. *Nucleic Acids Res.* 44:8013–8019.
139. van Emmerik, C. L., and H. van Ingen. 2019. Unspinning chromatin: revealing the dynamic nucleosome landscape by NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* 110:1–19.
140. Betzig, E., G. H. Patterson, ..., H. F. Hess. 2006. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*. 313:1642–1645.
141. Heilemann, M., S. van de Linde, ..., M. Sauer. 2008. Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angew. Chem. Int. Ed. Engl.* 47:6172–6176.
142. Hell, S. W., and J. Wichmann. 1994. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* 19:780–782.
143. Birk, U. J. 2019. Super-resolution microscopy of chromatin. *Genes (Basel)*. 10:493.
144. Bhattacharya, D., S. Talwar, ..., G. V. Shivashankar. 2009. Spatio-temporal plasticity in chromatin organization in mouse cell differentiation and during *Drosophila* embryogenesis. *Biophys. J.* 96:3832–3839.
145. Boettiger, A. N., B. Bintu, ..., X. Zhuang. 2016. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature*. 529:418–422.
146. Li, G., and P. Zhu. 2015. Structure and organization of chromatin fiber in the nucleus. *FEBS Lett.* 589:2893–2904.
147. Ou, H. D., S. Phan, ..., C. C. O'Shea. 2017. ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science*. 357:eaag0025.
148. Prieto, E. I., and K. Maeshima. 2019. Dynamic chromatin organization in the cell. *Essays Biochem.* 63:133–145.
149. Ricci, M. A., C. Manzo, ..., M. P. Cosma. 2015. Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. *Cell*. 160:1145–1158.
150. Xu, J., H. Ma, ..., Y. Liu. 2018. Super-resolution imaging of higher-order chromatin structures at different epigenomic states in single mammalian cells. *Cell Rep.* 24:873–882.
151. Reymann, J., D. Baddeley, ..., U. Birk. 2008. High-precision structural analysis of subnuclear complexes in fixed and live cells via spatially modulated illumination (SMI) microscopy. *Chromosome Res.* 16:367–382.
152. Murphy, P. J., B. R. Cipriani, ..., P. D. Soloway. 2013. Single-molecule analysis of combinatorial epigenomic states in normal and tumor cells. *Proc. Natl. Acad. Sci. USA*. 110:7772–7777.
153. Baumann, K. 2017. Genome organization: a vision of 3D chromatin organization. *Nat. Rev. Mol. Cell Biol.* 18:532.
154. Rowley, M. J., and V. G. Corces. 2018. Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* 19:789–800.
155. Tsompana, M., and M. J. Buck. 2014. Chromatin accessibility: a window into the genome. *Epigenetics Chromatin*. 7:1–16.
156. Woodcock, C. L., and R. P. Ghosh. 2010. Chromatin higher-order structure and dynamics. *Cold Spring Harb. Perspect. Biol.* 2:a000596.
157. Babu, D., and M. J. Fullwood. 2015. 3D genome organization in health and disease: emerging opportunities in cancer translational medicine. *Nucleus*. 6:382–393.
158. Bianchi, A., and C. Lanzuolo. 2015. Into the chromatin world: role of nuclear architecture in epigenome regulation. *AIMS Biophysics*. 2:585–612.
159. Pennisi, E. 2015. Inching toward the 3D genome: maps of DNA's loops folds advance - but may disagree. *Science*. 347:10.
160. Schoenfelder, S., and P. Fraser. 2019. Long-range enhancer-promoter contacts in gene expression control. *Nat. Rev. Genet.* 20:437–455.
161. Baker, M. 2011. Genomics: genomes in three dimensions. *Nature*. 470:289–294.
162. Bannister, A. J., P. Zegerman, ..., T. Kouzarides. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*. 410:120–124.
163. Langowski, J. 2006. Polymer chain models of DNA and chromatin. *Eur. Phys. J. E Soft Matter*. 19:241–249.
164. Tremethick, D. J. 2007. Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell*. 128:651–654.
165. Bai, Y., M. Greenfeld, ..., D. Herschlag. 2007. Quantitative and comprehensive decomposition of the ion atmosphere around nucleic acids. *J. Am. Chem. Soc.* 129:14981–14988.
166. Gebala, M., S. Bonilla, ..., D. Herschlag. 2016. Does cation size affect occupancy and electrostatic screening of the nucleic acid ion atmosphere? *J. Am. Chem. Soc.* 138:10925–10934.
167. Jacobson, D. R., and O. A. Saleh. 2017. Counting the ions surrounding nucleic acids. *Nucleic Acids Res.* 45:1596–1605.
168. Li, G., Y. Liu, ..., B. Ren. 2019. Joint profiling of DNA methylation and chromatin architecture in single cells. *Nat. Methods*. 16:991–993.
169. Yanao, T., S. Sano, and K. Yoshikawa. 2015. Chiral selection in wrapping, crossover, and braiding of DNA mediated by asymmetric bend-writhe elasticity. *AIMS Biophys.* 2:666–694.
170. Stehr, R., N. Kepper, ..., G. Wedemann. 2008. The effect of inter-nucleosomal interaction on folding of the chromatin fiber. *Biophys. J.* 95:3677–3691.
171. Korolev, N., O. V. Vorontsova, and L. Nordenskiöld. 2007. Physicochemical analysis of electrostatic foundation for DNA-protein interactions in chromatin transformations. *Prog. Biophys. Mol. Biol.* 95:23–49.
172. Elbahnsi, A., R. Retureau, ..., C. Oguey. 2018. Holding the nucleosome together: a quantitative description of the DNA-histone interface in solution. *J. Chem. Theory Comput.* 14:1045–1058.
173. West, S. M., R. Rohs, ..., B. Honig. 2010. Electrostatic interactions between arginines and the minor groove in the nucleosome. *J. Biomol. Struct. Dyn.* 27:861–866.
174. Andergassen, D., C. P. Dotter, ..., Q. J. Hudson. 2015. Allelome-PRO, a pipeline to define allele-specific genomic features from high-throughput sequencing data. *Nucleic Acids Res.* 43:e146.
175. Van De Geijn, B., G. Mcvicker, ..., J. K. Pritchard. 2015. WASP: Allele-specific software for robust molecular quantitative trait locus discovery. *Nature Methods*. 12:1061–1063.
176. Hu, Y., Q. An, ..., Y. Guo. 2018. Single cell multi-omics technology: Methodology and application. *Front Cell Development Biol.* 6:28.
177. Richard Albert, J., T. Koike, ..., M. C. Lorincz. 2018. Development and application of an integrated allele-specific pipeline for methylomic and epigenomic analysis (MEA). *BMC Genomics*. 19:463.
178. Si, Z., H. Yu, and Z. Ma. 2016. Learning deep features for DNA methylation data analysis. *IEEE Access*. 4:2732–2737.
179. Tan, F., T. Tian, ..., H. Hakonarson. 2020. Elucidation of DNA methylation on N 6-adenine with deep learning. *Nat. Mach. Intell.* 2:466–475.
180. N'Diaye, A., B. Byrns, ..., C. J. Pozniak. 2020. Machine learning analyses of methylation profiles uncovers tissue-specific gene expression patterns in wheat. *Plant Genome*. 13:e20027.
181. Levy, J. J., A. J. Titus, ..., B. C. Christensen. 2020. MethylNet: an automated and modular deep learning approach for DNA methylation analysis. *BMC Bioinformatics*. 21:108.

182. Liu, Q., L. Fang, ..., K. Wang. 2019. Detection of DNA base modifications by deep recurrent neural network on Oxford Nanopore sequencing data. *Nat. Commun.* 10:2449.
183. Ni, P., N. Huang, ..., J. Wang. 2019. DeepSignal: detecting DNA methylation state from Nanopore sequencing reads using deep-learning. *Bioinformatics.* 35:4586–4595.
184. Xu, L., and M. Seki. 2020. Recent advances in the detection of base modifications using the Nanopore sequencer. *J. Human Genet.* 65:25–33.
185. Prakash, K., B. Diederich, ..., L. Schermelleh. 2021. Super-resolution structured illumination microscopy: past, present and future. *Philos. Trans. R Soc. A Math Phys. Eng. Sci.* 379:20200143.
186. Boland, M. A., E. A. K. Cohen, ..., M. A. A. Neil. 2021. Improving axial resolution in Structured Illumination Microscopy using deep learning. *Philos. Trans. R Soc. A Math Phys. Eng. Sci.* 379:20200298.
187. Lachetta, M., H. Sandmeyer, ..., M. Müller. 2021. Simulating digital micromirror devices for patterning coherent excitation light in structured illumination microscopy. *Philos. Trans. R Soc. A Math. Phys. Eng. Sci.* 379.
188. Rajpurkar, A. R., L. J. Mateo, ..., A. N. Boettiger. 2021. Deep learning connects DNA traces to transcription to reveal predictive features beyond enhancer-promoter contact. *Nat. Commun.* 12:3423.
189. Nehme, E., L. E. Weiss, ..., Y. Shechtman. 2018. Deep-STORM: super-resolution single-molecule microscopy by deep learning. *Optica.* 5:458.
190. Barth, R., K. Bystricky, and H. A. Shaban. 2020. Coupling chromatin structure and dynamics by live super-resolution imaging. *Sci. Adv.* 6:eaaz2196.
191. Rakyan, V. K., T. A. Down, ..., S. Beck. 2011. Epigenome-wide association studies for common human diseases. *Nature Rev. Genetics.* 12:529–541.
192. Koestler, D. C., B. C. Christensen, ..., E. A. Houseman. 2013. Blood-based profiles of DNA methylation predict the underlying distribution of cell types: a validation analysis. *Epigenetics.* 8:816–826.
193. Reinius, L. E., N. Acevedo, ..., J. Kere. 2012. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS One.* 7:e41361.
194. Zou, J., C. Lippert, ..., J. Listgarten. 2014. Epigenome-wide association studies without the need for cell-type composition. *Nat. Methods.* 11:309–311.
195. Listgarten, J., C. Lippert, ..., D. Heckerman. 2012. Improved linear mixed models for genome-wide association studies. *Nature Methods.* 9:525–526.
196. Rauschert, S., K. Raubenheimer, ..., R. C. Huang. 2020. Machine learning and clinical epigenetics: A review of challenges for diagnosis and classification. *Clin. Epigenetics.* 12:1–11.
197. Holder, L. B., M. M. Haque, and M. K. Skinner. 2017. Machine learning for epigenetics and future medical applications. *Epigenetics.* 12:505–514.
198. Payne, S. R. 2010. From discovery to the clinic: the novel DNA methylation biomarker mSEPT9 for the detection of colorectal cancer in blood. *Epigenomics.* 2:575–585.
199. Rasool, M., A. Malik, ..., S. H. Gan. 2015. The role of epigenetics in personalized medicine: challenges and opportunities. *BMC Genomics.* 8:S5.
200. PubMed [Internet]. <https://pubmed.ncbi.nlm.nih.gov/>.
201. RStudio | Open source & professional software for data science teams - RStudio [Internet]. <https://rstudio.com/>.
202. BioRender [Internet]. <https://biorender.com/>.