Trends in Biotechnology

CellPress REVIEWS

Review

Single-Cell DNA Methylation Profiling: Technologies and Biological Applications

Ino D. Karemaker[1] and Michiel Vermeulen[1,*]

DNA methylation is an epigenetic modification that plays an important role in gene expression regulation, development, and disease. Recent technological innovations have spurred the development of methods that enable us to study the occurrence and biology of this mark at the single-cell level. Apart from answering fundamental biological questions about heterogeneous systems or rare cell types, low-input methods also bring clinical applications within reach. Ultimately, integrating these data with other single-cell data sets will allow deciphering multiple layers of gene expression regulation within each individual cell. Here, we review the approaches that have been developed to facilitate single-cell DNA methylation profiling, their biological applications, and how these will further our understanding of the biology of DNA methylation.

Highlights

Recent technological innovations have made it possible to study DNA methylation at the single-cell level.

Many different strategies have been developed, accommodating a wide variety of research questions.

Combining single-cell DNA methylation analysis with other omics approaches, different levels of information can be integrated from each individual cell.

Maps of DNA methylation in early embryonic development have been drawn using these new techniques.

The development of low-input methods holds promise for clinical applications.

DNA Methylation as an Epigenetic Modification

Epigenetic modifications are changes in the genetic material that cause a heritable phenotype without changing the DNA sequence itself. DNA methylation is an important epigenetic mark that comprises the coupling of a methyl group (CH_3) to deoxyribonucleosides. Although methylation can take place at multiple positions on any of the bases, **5-methylcytosine** (5meC; see Glossary) is the most abundant methylated DNA base in vertebrates [1]. Methylation of cytosines generally occurs in the context of a CpG dinucleotide, a self-complementary DNA sequence in which a cytosine is followed by a guanine on the same strand. The tendency of methylated cytosine to mutate into thymine causes the frequency of CpGs in the genome to be much lower than expected by chance; instead, CG dinucleotides typically cluster together in CpG-dense regions called **CpG islands** (CGIs) [2]. CGIs are generally associated with promoters, and methylation of CGI promoters can regulate transcription of associated genes *in cis* [3]. Methylation of DNA has been shown to be involved in various cellular processes, including X chromosome inactivation, genomic imprinting, and silencing of transposable elements [4–6]. Being such an important modification, it should not come as a surprise that aberrant DNA methylation has been implicated in a number of diseases, most notably cancer [7].

As such, DNA methylation has been the subject of many investigations, giving rise to an ever-increasing toolbox to study this epigenetic mark. Some of the latest additions to this toolbox are methods that make it possible to study DNA methylation at the single-cell level. Similar to recent developments in other areas of single-cell epigenomics, these tools present exciting new opportunities as they allow profiling of DNA methylation in individual cells and in unprecedented detail [8,9]. An important question that can now be addressed is whether our current models on the biology of DNA methylation, which are predominantly built on bulk methylation analyses of heterogeneous populations, hold true when scrutinised on the single-cell level. Another elusive issue is how concomitant methylation dynamics on different genomic loci converge to determine the biology of that cell, or how DNA methylation is mechanistically linked to the regulation of gene expression. While any kind of heterogeneous population or tissue hence profits from

[1]Department of Molecular Biology, Faculty of Science, Radboud Institute for Molecular Life Sciences, Oncode Institute, Radboud University Nijmegen, Nijmegen, The Netherlands

*Correspondence: michiel.vermeulen@science.ru.nl (M. Vermeulen).

single-cell techniques, systems that benefit the most are those that suffer from limited availability of material and could therefore not be studied in detail before. Examples of such systems include early mammalian development and patient-derived samples; single-cell DNA methylation profiling thus opens the door to exciting new fields of research, in both fundamental and clinical contexts.

The Traditional DNA Methylation Profiling Toolbox

A large variety of traditional tools exist to study DNA methylation, based on a multitude of techniques such as chromatography, mass spectrometry, **ELISA**, restriction digestion, immunoprecipitation, and bisulfite conversion [10,11]. Continuous technical progress also means continuous development of new techniques, such as those based on single-molecule imaging [12] or nanopore technology [13]. Although every method has its own advantages and disadvantages, the toolbox as a whole is well suited to address a wide range of research questions.

While some methods are predominantly used in a particular biological or technical context, others have become more mainstream. For instance, over the past decades bisulfite sequencing has turned into the gold standard for the genome-wide analysis of DNA methylation. Upon treatment of DNA with sodium bisulfite, unmethylated cytosines are deaminated into uracil, while methylated cytosines remain unaltered. When after **PCR** amplification the bisulfite-converted DNA is analysed by Sanger sequencing, unmethylated cytosines are read as thymine, whereas methylated cytosines are read as cytosine [14]. This technique hence produces a readout at single base-pair resolution, which becomes particularly powerful when combined with deep sequencing to cover the entire genome [15,16]. Although **whole-genome bisulfite sequencing** (WGBS) can provide 1X coverage for approximately 95% of CpGs in the genome [17], reaching this coverage requires very deep sequencing indeed, making it a costly affair [18]. Therefore, more cost-effective methods such as **reduced representation bisulfite sequencing** (RRBS) have been developed (Box 1). RRBS utilises restriction enzyme digestion and size fractionation to reduce the complexity of the DNA, thereby generating a small but reproducible sequencing library [19]. The methylation-insensitive restriction enzyme *Msp*I, which cuts at CCGG sites, can be used to enrich for CpG-dense regions, which are then preferentially sequenced [20]. While RRBS thus is an excellent and cheaper choice when studying CGIs, coverage drops to around 10% of total CpGs. In particular, regions of low CpG density such as enhancers are usually poorly covered in RRBS, necessitating the use of expensive WGBS to study the prominent methylation dynamics occurring at these sites [21].

Bisulfite-Based Single-Cell Methods

Even though bisulfite sequencing quickly became established as the method of choice for bulk DNA methylation analysis, single-cell adaptations faced the major hurdle of bisulfite-induced DNA degradation, which initially prohibited the development of low-input methods. In 2013, the first single-cell RRBS (scRRBS; Table 1) protocol was established through the integration of all experimental steps up to and including the bisulfite conversion into a single-tube reaction, followed by two rounds of PCR amplification and deep sequencing [22,23]. Despite representing an important breakthrough by, for the first time, enabling DNA methylation analysis on the level of single cells, this method also has some limitations. Apart from the persisting matter of DNA degradation caused by bisulfite conversion, which is only party circumnavigated by the experimental setup, scRRBS provides relatively poor coverage for imprinting loci, and suffers from a restricted overlap in coverage between individual cells [22,23]. Another concern is the PCR-induced amplification bias that is a consequence of this strategy, which was later tackled

Glossary

5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC): oxidised derivatives of 5meC. During active DNA demethylation, ten-eleven translocation (TET) methylcytosine dioxygenases catalyse the oxidation of 5meC into first 5hmC, then 5fC and 5caC, which can be converted into unmethylated cytosine.
5-methylcytosine (5meC): cytosine base with a methyl group (CH_3) covalently coupled to the 5′ position of its pyrimidine ring. The most abundant methylated DNA base in vertebrates.
Copy-number variation (CNV): structural variation in the genome in which specific regions of DNA are duplicated, with the number of repeats varying per individual.
CpG island (CGI): genomic region with a high frequency of CpG dinucleotides. Although not an official definition, CGIs are generally described to harbour a ratio of observed to expected CpGs that is at least 0.6.
ELISA: biochemical assay in which an enzyme-linked reaction is used to quantify the antibody-based detection of a substance in a sample.
Embryonic stem cells (ESCs): pluripotent cells derived from the inner cell mass of a blastocyst-stage embryo. ESCs can differentiate into all different cell types of an organism and can be kept in culture indefinitely, making them a convenient *in vitro* model for early embryonic development.
Formalin fixed and paraffin embedded (FFPE): a common method for clinical sample preservation. Tissue is fixed in its current state through treatment with formalin, and then embedded into paraffin for storage and downstream applications.
Intracytoplasmic sperm injection (ICSI): direct injection of a single sperm cell into the cytoplasm of a mature oocyte.
Methylation-sensitive restriction enzyme (MSRE): bacterial DNA-cutting enzyme that recognises a specific DNA sequence dependent on its methylation status.
PCR: molecular biology technique to amplify DNA through repeated cycles

in quantitative RRBS (Q-RRBS) through the introduction of **unique molecular identifiers** (UMIs) [24]. Lastly, coverage dropped to 40% of the CpG sites that can be detected by bulk RRBS – or 4% of the entire genome [22,23].

Coverage rate was improved to approximately 18% of all CpGs through coupling of post-bisulphite adaptor tagging (PBAT), where bisulfite conversion precedes adaptor ligation [25], to PCR amplification of the resulting tagged fragments followed by deep sequencing in single-cell bisulfite sequencing (scBS-seq) [26]. However, this method also does not facilitate full genome-wide coverage, leaving allele- or strand-specific methylation differences untraceable. In addition, the use of several rounds of primer binding means that methylation calls should be weighted by cell rather than read count, and that strand-specific information is lost [27]. An alternative method for single-cell WGBS (scWGBS) also took advantage of post-bisulphite adaptor ligation, but this time without the requirement for the preamplification step which is part of scBS-seq [28]. While this provides benefits such as the preservation of strandedness and reduced amplification bias, it comes at the expense of lower library complexity. This makes scWGBS particularly suitable for high-throughput analysis at low sequencing coverage, since deeper sequencing results in a higher rate of PCR duplicates [28]. Alternatively, a whole-genome shotgun bisulfite sequencing protocol termed single-cell PBAT (scPBAT) has been proposed [29]. Instead of using PCR to amplify the input material, scPBAT is based on a PBAT method where multiplexing increases the yield of the library [30]. This avoids the risk of amplification bias, but limits the use of this method to the analysis of repetitive regions, which are naturally more abundant in the genome [29].

With the rudiments of bisulfite sequencing-based single-cell methylation analysis thus having been explored, adaptations continue to be introduced. An example is the recently published single-nucleus methylcytosine sequencing (snmC-seq), which follows a now commercially available protocol that optimises recovery during library preparation from bisulfite-converted, single-stranded DNA [31]. This method makes use of random priming and extension, after which the samples are tailed and ligated to the second adapter in a single step [31]. Yet the question remains to what extent improved recovery of bisulfite-converted DNA compensates for the loss of information during the bisulfite conversion itself. Additionally, recent advances facilitate single-cell DNA methylation analysis in a high-throughput manner. For instance, single-cell combinatorial indexing for methylation analysis (sci-MET) uses a combinatorial indexing strategy to discriminate single cells after WGBS [32]. Key to this approach is to make the transposomes unresponsive to bisulfite treatment by loading them with oligonucleotides that are depleted of cytosines [32]. Further innovations are the introduction of multiplexing for single-cell RRBS [33] and the implementation of RRBS on a microfluidics device in microfluidic diffusion-based RRBS (MID-RRBS) [34]. These developments denote important steps towards

of DNA synthesis by a thermostable DNA polymerase enzyme.
Quantitative real-time PCR (qRT-PCR): variation on conventional PCR in which fluorescent dyes are used to quantify DNA in real time.
Reduced representation bisulfite sequencing (RRBS): targeted genome-wide analysis of DNA methylation through deep sequencing after bisulfite conversion, in which only a reduced, representative sample of the whole genome is sequenced.
Unique molecular identifiers (UMIs): short oligonucleotide barcodes that are added to a PCR, so that after amplification each copy can be traced back to a single parent molecule.
Whole-genome bisulfite sequencing (WGBS): genome-wide analysis of DNA methylation though deep sequencing after bisulfite conversion.

Box 1. A Quick Comparison between WGBS and RRBS

High-throughput bisulfite sequencing to study genome-wide DNA methylation comes in two flavours: WBGS and RRBS (Figure I). While both techniques are based on bisulfite conversion of the DNA to distinguish unmethylated from methylated cytosines, they differ slightly in the way that the samples are prepared for analysis by deep sequencing. In WGBS, sonication is usually the method of choice to shear the DNA into random fragments. These fragments are then blunted and an adenosine nucleotide is added to the 3′ end in processes called end repair and A-tailing, respectively. The overhanging adenosine serves as a binding site for sequencing adapters, which are ligated onto the DNA fragments. After this, fragments of the right size for sequencing are selected and subjected to bisulfite conversion, in which unmethylated cytosines are converted into uracil, while methylated cytosines remain unchanged. Bisulfite-converted fragments are amplified by PCR and sequenced, where at least 500 million reads are needed to provide enough coverage of the whole genome. For RRBS, the main adaptation to this protocol is that enzymatic cleavage is used to create DNA fragments that are CG rich at the ends. This means that RRBS libraries can be sequenced less deep to obtain a relatively high coverage of CG-dense regions. In summary, WGBS provides genome-wide coverage at a high cost, while RRBS provides less but targeted coverage at a reduced cost. Both methods are complementary, and which method is most suitable for a particular purpose is ultimately dictated by the research question.

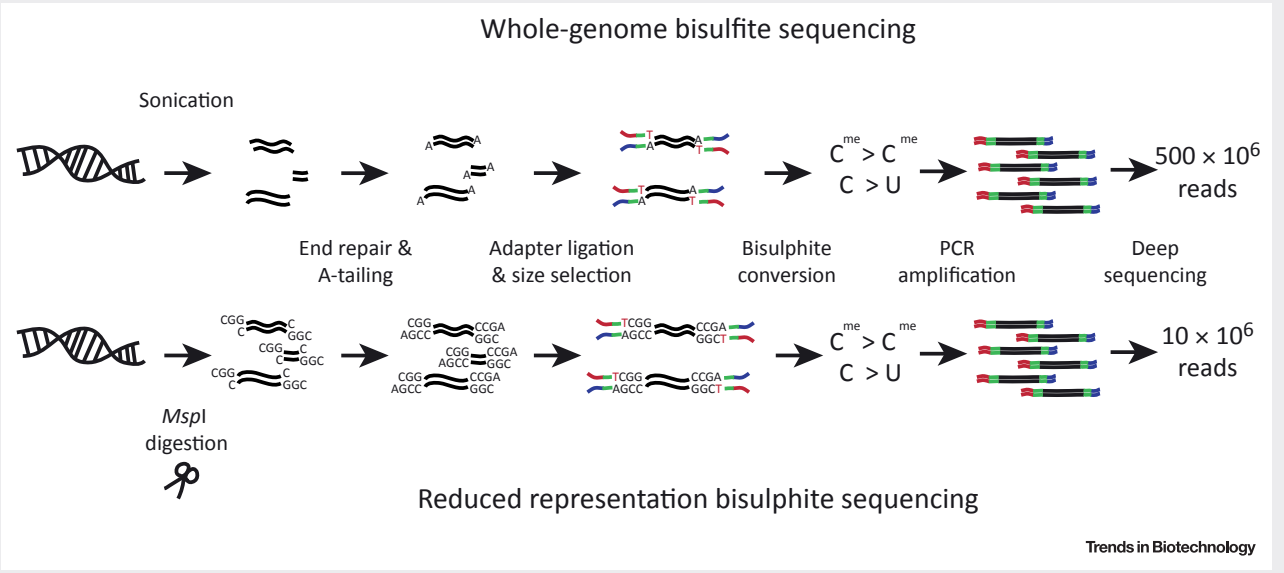

Figure I. A Schematic Comparison of Whole-Genome Bisulfite Sequencing (upper part) and Reduced Representation Bisulfite Sequencing (lower part).

low-input, high-throughput methylation profiling but also leave room for improvement; for example, MID-RRBS does not yet allow library preparation on the microfluidics device [34]. Finally, a different class of advancements are improved computational methods that increase the amount of information that can be recovered from low-depth sequencing [28,35]. By contrast, genome-wide or even CGI-wide coverage is superfluous for studies focussing only on particular loci, in which case single-cell locus-specific bisulfite sequencing (SLBS) provides a fine alternative [36].

Bisulfite-Free Single-Cell Methods

Even so, bisulfite treatment remains relentlessly harsh and conversion rates can vary, causing inconsistency across samples and fuelling the search for bisulfite-free single-cell methods. An early study utilised restriction digestion by **methylation-sensitive restriction enzymes** (MSREs) coupled to PCR amplification in a single reaction mixture on a microreaction slide for high-throughput DNA methylation analysis of single cells [37]. Although relatively affordable and easy to implement, this restriction enzyme-based single-cell methylation assay (RSMA) suffered from some drawbacks, most prominently its nonquantitative nature which in diploid genomes prohibits the distinction between fully methylated sequences and samples in which only one allele is methylated [37]. A similar approach was obtained by combining methylation-sensitive restriction digestion with multiplexed **quantitative real-time PCR** (qRT-PCR), carried out in a microfluidics device to enable high throughput [38,39]. Although more quantitative by nature, this single-cell restriction analysis of methylation (SCRAM) is still not sensitive enough to distinguish between heterozygously and homozygously methylated alleles in diploid cells [39]. More recently, genome-wide CGI methylation sequencing for single cells (scCGI-seq) extended the use of MSREs from a limited number of loci to CGIs at the genome scale through the introduction of multiple displacement amplification, in which CGI-containing sequences are selectively amplified and subjected to deep sequencing [40]. While this approach increases the coverage overlap between single cells, fewer CpGs across the genome are covered in total compared with bisulfite-based methods, especially in CpG-poor

Table 1. Single-Cell DNA Methylation Profiling Methods, Sorted by the Year That the Study Was Published

Method	Basis	Comments	Average CpG coverage	Refs
RSMA	Restriction enzyme (RE) based	Readout by gel electrophoresis	Not applicable	[37]
SCRAM	RE based	Readout by qRT-PCR	Not applicable	[38,39]
scRRBS	Bisulfite based	First genome-wide method for single-cell DNA methylation analysis	1×10^6	[22,23]
scBS-seq	Bisulfite based	First protocol where adaptor ligation precedes bisulfite treatment; recently updated [27]	3.7×10^6	[26]
Q-RRBS	Bisulfite based	Introduction of UMIs for scRRBS	$0.5–1 \times 10^6$	[24]
scWGBS	Bisulfite based	Coverage based on sequencing depth; cumulative coverage >90%	$0.5–2.5 \times 10^6$	[28]
SLBS	Bisulfite based	Locus specific	Not applicable	[36]
RGM	Fluorescent reporter system	Allows for visualisation of DNA methylation	Not applicable	[41]
scM&T-seq	Multiomics	Combined with RNA-seq	2.5×10^6	[56]
scTrio-seq	Multiomics	Combined with RNA-seq and CNV analysis	$0.8–1.5 \times 10^6$	[50]
scMT-seq	Multiomics	Combined with RNA-seq	0.5×10^6	[57]
scAba-seq	Glucosylation based	For detection of 5hmC	$2–4.4 \times 10^5$ unique 5hmC sites	[54]
scGEM	Multiomics	SCRAM combined with qRT-PCR and genotyping by next-generation sequencing	Not applicable	[60]
scPBAT	Bisulfite based	Amplification-free	Not specified	[29]
scCGI-seq	RE based	Up to ±75% cumulative coverage of CGIs	Not specified	[40]
scMAB-seq	Bisulfite based with M.SssI treatment	For detection of 5fC and 5-caC	$1–5 \times 10^5$	[55]
scCOOL-seq	Multiomics	Combined with analysis of chromatin state, nucleosome positioning, CNVs, and ploidy	3.8×10^6	[61]
scNOMe-seq	Multiomics	Combined with analysis of chromatin accessibility	1.3×10^6	[58]
snmC-seq	Bisulfite based	Uses single nuclei; optimises recovery during library preparation	Not specified; 4.7%–5.7% coverage of the entire genome	[31]
scNMT-seq	Multiomics	Combined with RNA-seq and chromatin accessibility	Not specified; estimated $1–1.5 \times 10^6$	[59]
sci-MET	Bisulfite based	Uses combinatorial indexing for discrimination of single cells	Not specified	[32]
MID-RRBS	Bisulfite based	Implementation of scRRBS on a microfluidics device	$3.5–23.1 \times 10^4$	[34]

regions [40]. In general, the major disadvantages of MSRE-based DNA methylation analysis methods are their dependency on naturally occurring MRSEs and their reduced resolution. Whereas bisulfite-based methods often provide single-nucleotide resolution, employers of MSRE-based approaches must be satisfied with resolution based on fragment size.

Of course, it ultimately depends on the research question whether the circumvention of bisulfite conversion outweighs this loss of resolution.

An altogether different means of assessing DNA methylation at the single-cell level was achieved through development of the reporter of genomic methylation (RGM) method [41]. In contrast to both bisulfite- and MSRE-based methods, which concentrate on DNA profiling of static states, RGM aims to trace dynamic changes in endogenous methylation states through a fluorescence-based DNA methylation reporter system. This tactic allows for visualisation of methylation dynamics even at single-cell resolution but is inherently locus specific and relatively laborious, compromising its suitability for high-throughput applications. Nevertheless, this nicely illustrates that many different techniques can be adapted to study DNA methylation at the single-cell level (Figure 1).

Biological Applications

While many single-cell DNA methylation papers focussed primarily on technical improvements, new biological insights have also been obtained. Most studies up to now have investigated either early embryonic development or tumour tissues. Here, we categorise these observations based on their biological context and briefly discuss them.

Early Mammalian Development

To begin with, single-cell DNA methylation analyses in preimplantation embryos or their *in vitro* model system **embryonic stem cells** (ESCs) provided confirmation of some previously established findings. For example, scRRBS showed that demethylation after fertilisation happens more quickly in the paternal pronucleus than in the maternal one [22,42]. In addition, the role of maternal TRIM28 in DNA methylation was confirmed on the single-cell scale using SCRAM [38,43]. Similarly, single-cell analyses found global methylation of metaphase II oocytes [26,44] and hypomethylation in mouse ESCs (mESCs) grown in the presence of two inhibitors (2i) compared with serum [26,28,45]. Although not completely novel, these findings do validate the methods and tell us that the results hold true on the single-cell level and thus are not merely an effect of averaging over a population, which is valuable information in itself.

Moreover, truly novel findings were reported through the use of single-cell DNA methylation analysis. Findings include characterisations of DNA methylation by scWGBS in mESCs upon different kinds of induced differentiation [28] or in human oocytes during maturation [46], and a comparison between the first polar body and the metaphase II oocyte within the same female mouse gamete, which showed that their methylomes are highly similar [22]. On a somewhat bigger scale, scRRBS was used to map DNA methylation in human early embryos [47]. While most data were obtained from experiments on whole embryos, this study also included analyses on single female and male pronuclei at different time points after **intracytoplasmic sperm injection** (ICSI), as well as a couple of single metaphase II oocytes and single sperm cells. This revealed that although suffering from large cell-to-cell variation, also in humans demethylation of the paternal pronucleus occurs faster than that of the maternal pronucleus [47]. A recent paper used scBS to study the heterogeneity of DNA methylation between individual cells of human preimplantation embryos in more detail, whilst also addressing issues of aneuploidy and parental-specific DNA methylation [48]. In addition, a better chart of DNA methylation in murine germ cells was obtained through identification of repetitive elements that stay methylated during germ cell development [29], and of mESC DNA through the identification of motifs that are associated with methylation variability [35]. In summary, these studies signify important steps towards reconstructing how DNA methylation shapes early embryonic development.

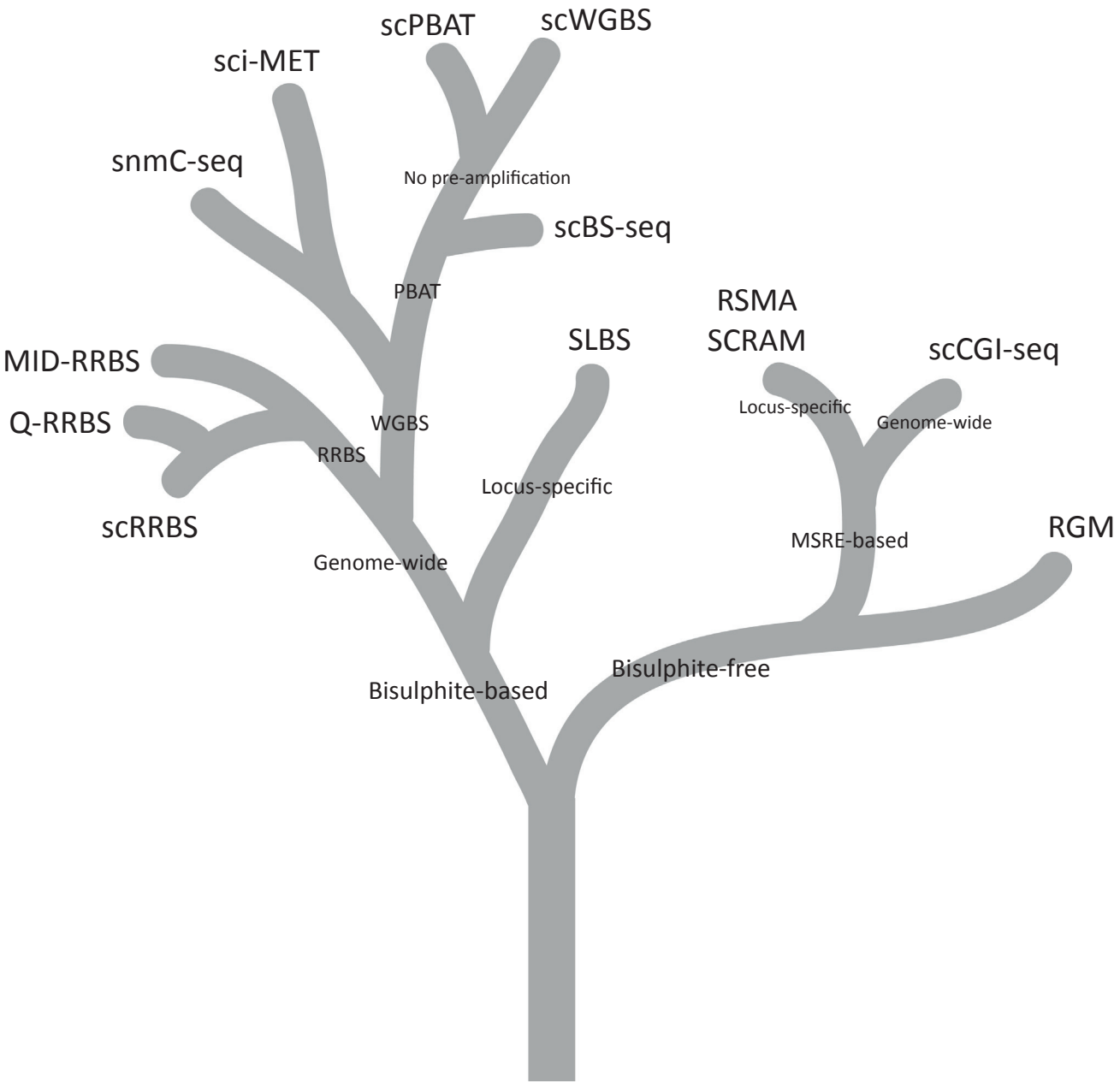

Figure 1. The Family of Single-Cell DNA Methylation Analysis Methods. Schematic representation of the wide variety of single-cell DNA methylation profiling techniques. Every branch point of the tree depicts different alternatives of a particular property. MID-RRBS, microfluidic diffusion-based RRBS; MSRE, methylation-sensitive restriction enzymes; Q-RRBS, quantitative RRBS; RGM, reporter of genomic methylation; RRBS, reduced representation bisulfite sequencing; RSMA, restriction enzyme-based single-cell methylation assay; scBS-seq, single-cell bisulfite sequencing; scCGI-seq, genome-wide CpG islands methylation sequencing for single cells; sci-MET, single-cell combinatorial indexing for methylation analysis; scPBAT, single-cell post-bisulfite adaptor tagging; SCRAM, single-cell restriction analysis of methylation; scRRBS, single-cell RRBS; scWGBS, single-cell whole-genome bisulfite sequencing; SLBS, single-cell locus-specific bisulfite sequencing; snmC-seq, single-nucleus methylcytosine sequencing.

Towards Clinical Purposes

Another favourite model system for single-cell DNA methylation studies comprises cancer cell lines or tissues and wild-type controls (Table 2), not only because of their heterogeneity but also for their potential clinical relevance. Indeed, rather than making significant contributions to our understanding of cancer biology, these studies explored and expanded the use of single-cell DNA methylation analysis in a clinical setting. For example, a number of healthy somatic tissues have been studied by single-cell DNA methylation analysis, such as mouse hepatocytes [36,49] and mouse and human neurons [31,36]. scWGBS and snmC-seq revealed heterogeneity and identified subpopulations in the mouse liver and the human brain, respectively [31,49], adding to our fundamental knowledge about these tissues. In addition, SLBS was successfully used to detect random changes in the methylation status of single CpG sites in mouse hepatocytes and neurons, suggesting that this method could also be exploited in the context of human tumours [36].

Furthermore, RSMA was used to examine DNA methylation status at regions that are known to be differentially methylated in cancer, thereby demonstrating its suitability for diagnostic purposes [37]. To learn more about cancer methylation patterns and their response to drugs, DNA methylome dynamics were measured on the single-cell level in K562 cells upon treatment with a commonly used epigenetic drug, as well as in HL60 cells upon induced differentiation [28]. K562 cells, together with GM12878 cells, were also subjected to scCGI-seq to investigate the methylation state of CpG-dense regions on a genome-wide scale [40]. This analysis showed that CGIs and promoters are hypermethylated in K562 cells as compared with GM12878 cells, while repeat regions are hypomethylated, in agreement with our current comprehension of cancer methylation patterns [40]. These studies hence added to our knowledge about the haematopoietic system in general, and leukaemia and its treatment in particular.

Although cancer cell lines make for a convenient model system, the question remains to what extent they accurately recapitulate primary tumours. Especially for clinical purposes, it is hence

Table 2. Wild-Type or Cancer Cell Lines or Tissues Used for Single-Cell DNA Methylation Analysis

Cell line or tissue	Species	Origin	Comments	Refs
Hepatocytes	Mouse	Liver	Demonstrates possible diagnostic purpose for SLBS	[36,49]
Neurons	Human	Brain	Identification of new neuronal subtypes	[31]
	Mouse		Demonstrates possible diagnostic purpose for SLBS	[36]
SW480	Human	Colorectal adenocarcinoma	Demonstrates possible diagnostic purpose for RSMA	[37]
K562	Human	Chronic myeloid leukaemia	Insight into single-cell methylome dynamics upon leukemic drug treatment	[28]
			Validation of current model of cancer methylation patterns	[40]
HL60	Human	Acute myeloid leukaemia	Insight into single-cell methylome dynamics upon vitamin D3 treatment	[28]
Hepatocellular carcinoma	Human	Patient derived	Identification of subpopulations; possible tool for developing targeted cancer treatment	[50]
HepG2		Hepatocellular carcinoma		
GM12878	Human	Lymphoblast	Validation of current model of cancer methylation patterns	[40]
Ovarian serous papillary carcinoma	Human	Patient derived	Unlocks patient-derived FFPE tissue as a source for clinical single-cell studies	[51]

important to be able to study patient material obtained from *in vivo* tumours. The first steps towards this goal have already been made: for example, hepatocellular carcinoma was investigated using single-cell triple-omics sequencing (scTrio-seq), which resulted in the identification of two subpopulations within the patient-derived sample [50]. One of these populations, constituting the minority of the tumour tissue, was found to be more invasive and more likely to evade the immune system [50]. This indicates that DNA methylation analysis, in conjunction with other data from the same single cells, has potential to identify the most dangerous cells in a tumour, assisting the development of targeted cancer treatment. Lastly, single cells were obtained by laser capture microdissection from patient-derived **formalin-fixed and paraffin-embedded** (FFPE) tissue, thereby unlocking an important resource for future single-cell studies [51]. Together, these data convincingly illustrate the potential of single-cell methylome studies for clinical purposes.

DNA Demethylation and Multiomics Approaches

All in all, numerous approaches, both bisulfite-based and bisulfite-free, exist for the analysis of DNA methylation on the single-cell level, accommodating a wide variety of research questions. Yet DNA methylation is a dynamic modification, and additional methods are needed to monitor these dynamics and integrate them with other data sets to obtain an accurate overview of the DNA methylation network and its biological consequences. One main disadvantage shared by the aforementioned methods is the incapability of distinguishing between 5meC and its oxidised derivatives **5-hydroxymethylcytosine** (5hmC), **5-formylcytosine** (5fC), and **5-carboxylcytosine** (5-caC) [52,53]. Recently, advances have been made to discriminate between the different forms of modified cytosine on the single-cell level in methods such as single-cell hydroxymethylation sequencing (scAba-seq) and single-cell methylase-assisted bisulfite sequencing (scMAB-seq) [54,55], thereby allowing mapping of active DNA demethylation. Even though these studies primarily showed that single-cell mapping of the oxidised derivatives of 5meC can be used for lineage reconstruction [54,55], these methods will allow for a deeper understanding of the dynamics of DNA methylation.

In addition, methods that integrate single-cell DNA methylation analysis with other single-cell omics approaches have been reported (Figure 2). Single-cell methylome and transcriptome sequencing (scM&T-seq and scMT-seq) combines genome-wide DNA methylation analysis with RNA sequencing [56,57], while single-cell nucleosome occupancy and methylome-sequencing (scNOMe-seq) combines analysis of endogenous methylation at CpGs with chromatin accessibility [58]. Techniques that also analyse both the methylome and transcriptome but go even one step further are scTrio-seq, which furthermore takes into account **copy-number variations** (CNVs) [50], single-cell nucleosome, methylation and transcription sequencing (scNMT-seq), which instead looks at chromatin accessibility [59], and single-cell analysis of genotype, expression and methylation (scGEM), which combines SCRAM with single-cell qRT-PCR to measure the transcriptome and next-generation sequencing for single-cell genotyping [60]. Moreover, single-cell Chromatin Overall Omic-scale Landscape Sequencing (scCOOL-seq) analyses in the same single cell chromatin state, nucleosome positioning, CNVs, and ploidy in addition to DNA methylation [61].

Putting DNA methylation in a broader cellular context through the use of other single-cell techniques, coupling genome-wide methylomes and transcriptomes showed that although the extent to which these are connected differs from cell to cell, heterogeneous expression of key pluripotency factors in mESCs is in general accompanied by heterogeneous methylation of distal regulatory elements [56]. When DNA methylation and other chromatin-related analyses are combined within the same single cells from mouse preimplantation embryos through

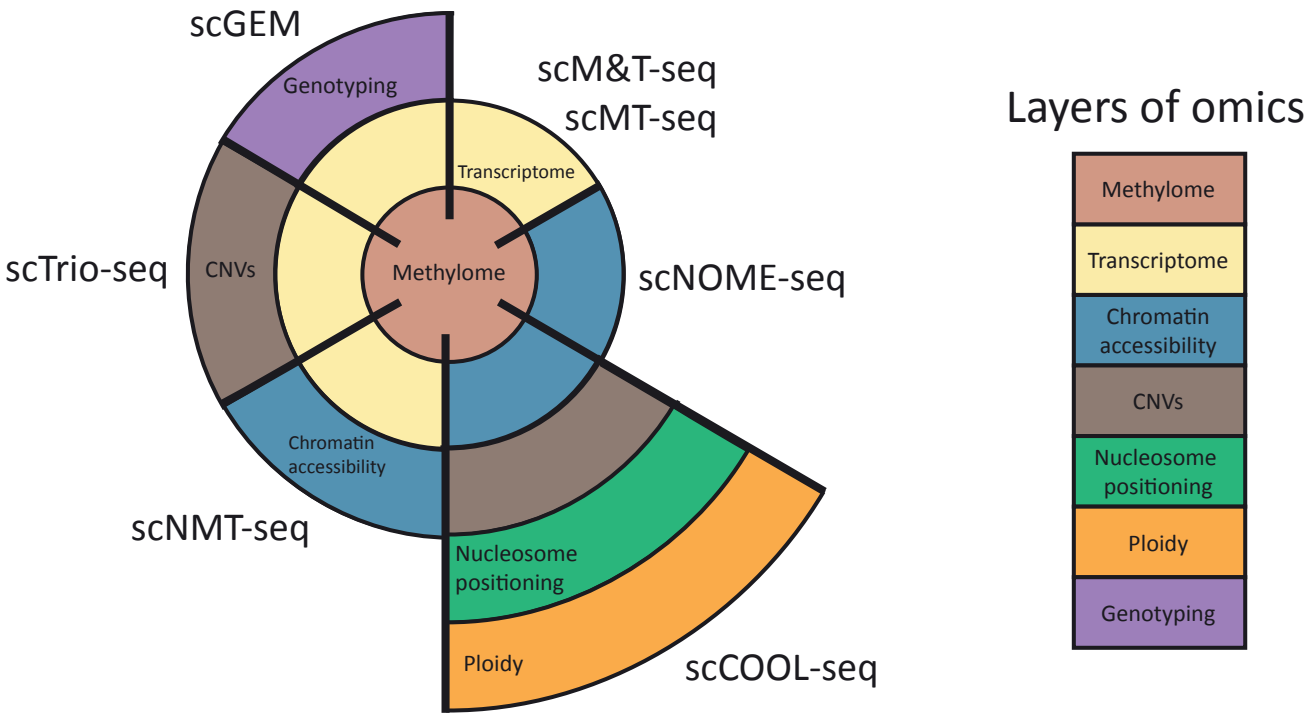

Figure 2. Multiomics Techniques for Dissecting Different Layers of Information from the Same Single Cell in Addition to the Methylome. Graphical representation of the different layers of omics that are addressed by each method on top of the methylome. CNV, copy-number variation; scCOOL-seq, single-cell Chromatin Overall Omic-scale Landscape Sequencing; scGEM, single-cell analysis of genotype, expression and methylation; scNMT-seq, single-cell nucleosome, methylation and transcription sequencing; scNOME-seq, single-cell nucleosome occupancy and methylome-sequencing; scTrio-seq, single-cell triple-omics sequencing.

scCOOL-seq, this produces a detailed overview of the relationship between DNA methylation and chromatin dynamics and how they together shape early mouse development [61]. These advancements reflect the enormous technological progress that has been made over the past years and help us piece together the complete picture of the biology of DNA methylation on a single-cell level.

Concluding Remarks and Future Perspectives

These are exciting times for the field of DNA methylation: recent technological advances have generated unprecedented possibilities for studying this epigenetic modification at the single-cell level. Despite the great progress that has been made, there always remains room for improvement (see Outstanding Questions).

Single-cell DNA methylation technologies can be categorised in several ways (Figure 3, Key Figure). Rather than there being one supreme method, the protocols are complementary and can be suited to investigate a wide variety of research questions. Firstly, a broad distinction can be made between bisulfite-based and bisulfite-free methods (Figure 3A). With bisulfite sequencing being the traditional method of choice for investigating DNA methylation in bulk, an obvious approach to developing single-cell DNA methylation techniques was through adaptation of both WGBS and RRBS for use in single cells. Continuous efforts are made to adapt and optimise these techniques, as data robustness and CpG coverage rate as well as cost and handling time could still be improved further. Meanwhile, the harsh nature of bisulfite treatment stimulates the search for alternative, bisulfite-free methods. Most of these alternative

Key Figure

Single-Cell DNA Methylation Technologies and Applications

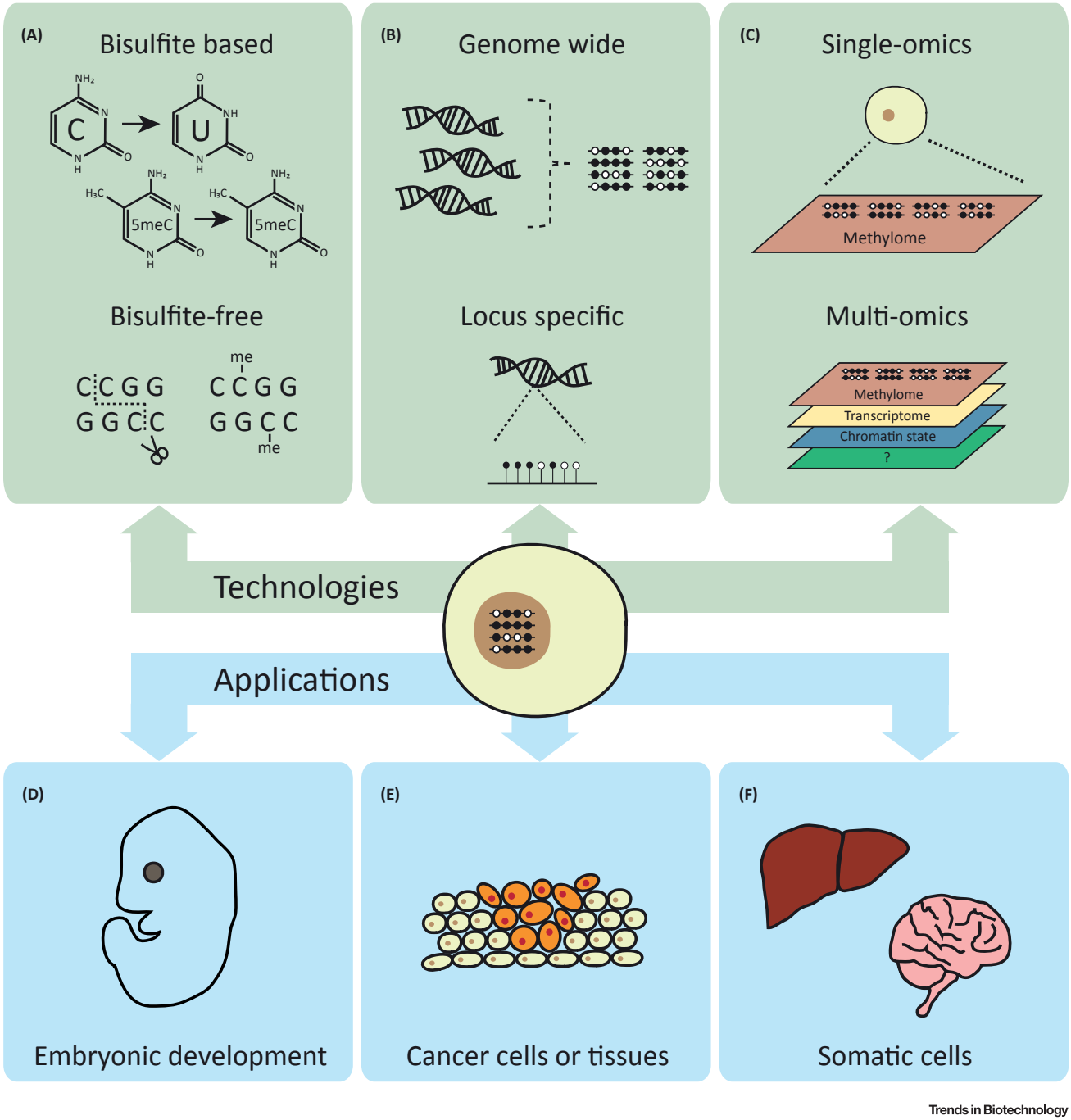

Figure 3. (A–C) Techniques developed for the analysis of single-cell DNA methylation can be categorised in various ways, such as their use of bisulfite conversion (A), their scale (B), and whether they integrate methylome analysis with additional single-cell omics methods (C). (D–F) Biological applications of single-cell DNA methylation analysis include the study of early embryonic development (D), cancer cell lines or tissues (E), and healthy somatic cells, representing any heterogeneous system (F).

methods are based on MSREs and are recently being adapted for use on a genome-wide scale. Secondly, we can distinguish genome-wide and locus-specific methods (Figure 3B). While genome-wide techniques mostly rely on deep sequencing, locus-specific methods can employ a different range of techniques as a read-out. This means that locus-specific methods are generally easier to implement and more accessible for any laboratory, especially as long as sequencing remains a relatively costly and time-consuming process. Lastly, a classification that since recently can be made concerns single-omics and multiomics methods (Figure 3C). Single-omics methods are usually more straightforward and easier to interpret, but more powerful approaches arise when combining several single-cell techniques, thereby connecting DNA methylation analysis with different layers of information from the same single cell. A couple of such multiomics methods have already been described [50,56–61], and it is likely that these are just the tip of the iceberg of what is possible. The development of such methods also emphasises the need for bioinformatics tools for data integration (such as [62]); indeed, caution must be taken when comparing multiple layers of information, since these are distinct biological pathways whose dynamics operate on different timescales. In addition, easy-to-use software and an improved level of proper bioinformatics training in general will make these techniques better accessible to the community and will aid correct processing and interpretation of the data, which is vital to decipher the biology that they contain.

Although most single-cell DNA methylation research has been technology driven rather than biology oriented, a modest start has been made in unlocking the potential for biological applications that these techniques contain (Figure 3D–F). So far, these findings mostly agree with our previous models of DNA methylation, but it will be interesting to see if, where, and how any discrepancies will arise. Apart from such observational studies, future efforts should also be directed to elucidate the mechanisms through which DNA methylation regulates gene expression. Multiomics investigations provide the first correlative data to answer this question but as yet fall short when it comes to proving causation. Perhaps the adaptation for use in single cells of other techniques such as clustered regularly interspaced short palindromic repeats-nuclease-dead Cas9 (CRISPR-dCas9)-based (epi)genome editing or high-resolution imaging will offer a solution, possibly in combination with more advanced integrative mathematical modelling. Lastly, developments should be directed towards clinical purposes. Some promising results have been obtained pointing to potential diagnostic applications, but more work should be done to optimise and standardise these methods. Ultimately, the detailed view that single-cell DNA methylation profiling brings benefits both fundamental and applied research and brings us closer to a fundamental understanding of any complex biological system.

Outstanding Questions

How can we develop easy-to-use single-cell DNA methylation techniques that provide robust data at high sequence coverage while minimising cost and handling time?

How can we use single-cell DNA methylation data to develop fast and reliable applications that can be used for clinical diagnostics?

How do we ensure that the data can be understood and verified by a wide audience, to avoid misinterpretation?

How do the different methods that exist now directly compare with each other, when performed in parallel on exactly the same sample?

Does our current understanding of DNA methylation hold true when examined on the single-cell level?

How many different omics techniques can we combine to obtain different levels of information from the same single cell simultaneously?

When integrating DNA methylation with other omics data, how are these different layers of information related?

How can we integrate multiomics data with biochemical evidence to elucidate how DNA methylation is mechanistically linked to regulation of gene expression?

Acknowledgements

Work in the Vermeulen lab is supported by a European Research Council Starting Grant (no. 309384) and the Netherlands Organisation for Scientific Research (NWO Gravitation Programme Cancer Genomics Netherlands). The Vermeulen lab is part of the Oncode Institute, which is partly funded by the Dutch Cancer Society (KWF).

Supplemental Information

Supplemental information associated with this article can be found, in the online version, at https://doi.org/10.1016/j.tibtech.2018.04.002.

References

1. Culp, L.A. *et al.* (1970) Methylated bases in DNA of animal origin. *Arch. Biochem. Biophys.* 136, 73–79
2. Bird, A.P. (1986) CpG-rich islands and the function of DNA methylation. *Nature* 321, 209–213
3. Deaton, A.M. and Bird, A. (2011) CpG islands and the regulation of transcription. *Genes Dev.* 25, 1010–1022
4. Mohandas, T. *et al.* (1981) Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. *Science* 211, 393–396
5. Li, E. *et al.* (1993) Role for DNA methylation in genomic imprinting. *Nature* 366, 362–365

6. Walsh, C. *et al.* (1998) Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.* 20, 116–117
7. Robertson, K.D. (2005) DNA methylation and human disease. *Nat. Rev. Genet.* 6, 597–610
8. Schwartzman, O. and Tanay, A. (2015) Single-cell epigenomics: techniques and emerging applications. *Nat. Rev. Genet.* 16, 716–726
9. Clark, S.J. *et al.* (2016) Single-cell epigenomics: powerful new methods for understanding gene regulation and cell identity. *Genome Biol.* 17, 72
10. Laird, P.W. (2010) Principles and challenges of genome-wide DNA methylation analysis. *Nat. Rev. Genet.* 11, 191–203
11. Kurdyukov, S. and Bullock, M. (2016) DNA methylation analysis: choosing the right method. *Biology (Basel)* 5, 3
12. Song, C.-X. *et al.* (2016) Simultaneous single-molecule epigenetic imaging of DNA methylation and hydroxymethylation. *Proc. Natl. Acad. Sci. U. S. A.* 113, 4338–4343
13. Rand, A.C. *et al.* (2017) Mapping DNA methylation with high-throughput nanopore sequencing. *Nat. Methods* 14, 411–413
14. Frommer, M. *et al.* (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. U. S. A.* 89, 1827–1831
15. Cokus, S.J. *et al.* (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 452, 215–219
16. Lister, R. *et al.* (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* 133, 523–536
17. Harris, R.A. *et al.* (2010) Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nat. Biotechnol.* 28, 1097–1105
18. Beck, S. (2010) Taking the measure of the methylome. *Nat. Biotechnol.* 28, 199–218
19. Meissner, A. *et al.* (2005) Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* 33, 5868–5877
20. Gu, H. *et al.* (2011) Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat. Protoc.* 6, 468–481
21. Bogdanović, O. *et al.* (2016) Active DNA demethylation at enhancers during the vertebrate phylotypic period. *Nat. Genet.* 48, 417–426
22. Guo, H. *et al.* (2013) Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res.* 23, 2126–2135
23. Guo, H. *et al.* (2015) Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing. *Nat. Protoc.* 10, 645–659
24. Wang, K. *et al.* (2015) Q-RRBS: a quantitative reduced representation bisulfite sequencing method for single-cell methylome analyses. *Epigenetics* 10, 775–783
25. Miura, F. *et al.* (2012) Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. *Nucleic Acids Res.* 40, e136
26. Smallwood, S.A. *et al.* (2014) Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat. Methods* 11, 817–820
27. Clark, S.J. *et al.* (2017) Genome-wide base-resolution mapping of DNA methylation in single cells using single-cell bisulfite sequencing (scBS-seq). *Nat. Protoc.* 12, 534–547
28. Farlik, M. *et al.* (2015) Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics. *Cell Rep.* 10, 1386–1397
29. Kobayashi, H. *et al.* (2016) Repetitive DNA methylome analysis by small-scale and single-cell shotgun bisulfite sequencing. *Genes Cells* 21, 1209–1222
30. Miura, F. and Ito, T. (2015) Highly sensitive targeted methylome sequencing by post-bisulfite adaptor tagging. *DNA Res.* 22, 13–18
31. Luo, C. *et al.* (2017) Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex. *Science* 604, 600–604
32. Mulqueen, R.M. *et al.* (2017) Scalable and efficient single-cell DNA methylation sequencing by combinatorial indexing. *BioRxiv* http://dx.doi.org/10.1101/157230
33. Charlton, J. *et al.* (2018) Global delay in nascent strand DNA methylation. *Nat. Struct. Mol. Biol.* 25, 327–332
34. Ma, S. *et al.* (2018) Cell-type-specific brain methylomes profiled via ultralow-input microfluidics. *Nat. Biomed. Eng.* 2, 183–194
35. Angermueller, C. *et al.* (2017) DeepCpG: accurate prediction of single-cell DNA methylation states using deep learning. *Genome Biol.* 18, 67
36. Gravina, S. *et al.* (2015) Single-cell, locus-specific bisulfite sequencing (SLBS) for direct detection of epimutations in DNA methylation patterns. *Nucleic Acids Res.* 43, e93
37. Kantlehner, M. *et al.* (2011) A high-throughput DNA methylation analysis of a single cell. *Nucleic Acids Res.* 39, e44
38. Lorthongpanich, C. *et al.* (2013) Single-cell DNA-methylation analysis reveals epigenetic chimerism in preimplantation embryos. *Science* 341, 1110–1112
39. Cheow, L.F. *et al.* (2015) Multiplexed locus-specific analysis of DNA methylation in single cells. *Nat. Protoc.* 10, 619–631
40. Han, L. *et al.* (2017) Bisulfite-independent analysis of CpG island methylation enables genome-scale stratification of single cells. *Nucleic Acids Res.* 45, e77
41. Stelzer, Y. *et al.* (2015) Tracing dynamic changes of DNA methylation at single-cell resolution. *Cell* 163, 218–229
42. Mayer, W. *et al.* (2000) Demethylation of the zygotic paternal genome. *Nature* 403, 501–502
43. Messerschmidt, D.M. *et al.* (2012) Trim28 is required for epigenetic stability during mouse oocyte to embryo transition. *Science* 335, 1499–1502
44. Lucifero, D. *et al.* (2002) Methylation dynamics of imprinted genes in mouse germ cells. *Genomics* 79, 530–538
45. Ficz, G. *et al.* (2013) FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell* 13, 351–359
46. Yu, B. *et al.* (2017) Genome-wide, single-cell DNA methylomics reveals increased non-CpG methylation during human oocyte maturation. *Stem Cell Rep.* 9, 397–407
47. Guo, H. *et al.* (2014) The DNA methylation landscape of human early embryos. *Nature* 511, 606–610
48. Zhu, P. *et al.* (2017) Single-cell DNA methylome sequencing of human preimplantation embryos. *Nat. Genet.* 50, 12–19
49. Gravina, S. *et al.* (2016) Single-cell genome-wide bisulfite sequencing uncovers extensive heterogeneity in the mouse liver methylome. *Genome Biol.* 17, 150
50. Hou, Y. *et al.* (2016) Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res.* 26, 304–319
51. Li, Q. *et al.* (2017) Heterogeneous DNA methylation status in same-cell subpopulations of ovarian cancer tissues. *Tumor Biol.* 39, 1–7
52. Tahiliani, M. *et al.* (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324, 930–935
53. Ito, S. *et al.* (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 466, 1129–1133
54. Mooijman, D. *et al.* (2016) Single-cell 5hmC sequencing reveals chromosome-wide cell-to-cell variability and enables lineage reconstruction. *Nat. Biotechnol.* 34, 852–856

55. Wu, X. *et al.* (2017) Simultaneous mapping of active DNA demethylation and sister chromatid exchange in single cells. *Genes Dev.* 31, 511–523

56. Angermueller, C. *et al.* (2016) Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat. Methods* 13, 229–232

57. Hu, Y. *et al.* (2016) Simultaneous profiling of transcriptome and DNA methylome from a single cell. *Genome Biol.* 17, 88

58. Pott, S. (2017) Simultaneous measurement of chromatin accessibility, DNA methylation, and nucleosome phasing in single cells. *Elife* 6, e23203

59. Clark, S.J. *et al.* (2017) scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat. Commun.* 9, 781

60. Cheow, L.F. *et al.* (2016) Single-cell multimodal profiling reveals cellular epigenetic heterogeneity. *Nat. Methods* 13, 833–836

61. Guo, F. *et al.* (2017) Single-cell multi-omics sequencing of mouse early embryos and embryonic stem cells. *Cell Res.* 27, 1–22

62. Welch, J.D. *et al.* (2017) MATCHER: Manifold alignment reveals correspondence between single cell transcriptome and epigenome dynamics. *Genome Biol.* 18, 138