Genome Biology

OPINION

Open Access

CrossMark

Single-cell epigenomics: powerful new methods for understanding gene regulation and cell identity

Stephen J. Clark[1], Heather J. Lee[1,2]*, Sébastien A. Smallwood[1,3], Gavin Kelsey[1,4] and Wolf Reik[1,2,4,5]

Abstract

Emerging single-cell epigenomic methods are being developed with the exciting potential to transform our knowledge of gene regulation. Here we review available techniques and future possibilities, arguing that the full potential of single-cell epigenetic studies will be realized through parallel profiling of genomic, transcriptional, and epigenetic information.

Introduction

Epigenetics involves the study of regulatory systems that enable heritable changes in gene expression within genotypically identical cells. This includes chemical modifications to DNA and the associated histone proteins, as well as changes in DNA accessibility and chromatin conformation [1]. Until recently, our understanding of these epigenetic modifications has depended entirely upon correlations between bulk measurements in populations of cells. These studies have classified epigenetic marks as being associated with active or repressed transcriptional states, but such generalizations often conceal a more complex relationship between the epigenome and gene expression.

Arguably, and as for many biological questions, investigation of epigenetic regulation in general is most usefully studied at the single-cell level, where intercellular differences can be observed leading to a more refined understanding compared with bulk analysis [2]. Additionally, the development of single-cell technologies is key to investigating the profound remodeling of the epigenome during the early stages of embryonic development, including in human samples where cell numbers are very limited and where epigenetic heterogeneity may be most pronounced.

High-throughput sequencing has revolutionized the field of epigenetics with methods for genome-wide mapping of DNA methylation, histone modifications, chromatin accessibility, and chromosome conformation (Table 1). Initially, the input requirements for these methods meant that samples containing hundreds of thousands or millions of cells were required; but in the last couple of years this has changed with numerous epigenetic features now assayable at the single-cell level (Fig. 1). Combined single-cell methods are also emerging that allow analyses of epigenetic–transcriptional correlations thereby enabling detailed investigations of how epigenetic states are associated with phenotype.

In this article, we review current and emerging methods for mapping epigenetic marks in single cells and the challenges these methods present. We subsequently discuss applications of these technologies to the study of development and disease.

Single-cell methodologies and future technological developments
Cytosine methylation and other DNA modifications
DNA methylation of cytosine (5mC) residues can be mapped genome-wide using several methods such as methylation-specific restriction enzymes [3], affinity purification [4], or by using bisulfite conversion followed by sequencing (BS-seq) [5]. The latter is considered the gold-standard method as it allows single base resolution and absolute quantification of DNA methylation levels. While investigation of DNA methylation at the single-cell level was motivated by important biological questions, until recently it was unfeasible due to the large amount of DNA degradation caused by the bisulfite conversion, which was traditionally performed after preparing adapter-tagged libraries.

* Correspondence: heather.lee@babraham.ac.uk
[1]Epigenetics Programme, Babraham Institute, Cambridge CB22 3AT, UK
[2]Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK
Full list of author information is available at the end of the article

Clark *et al. Genome Biology* (2016) 17:72

Page 2 of 10

Table 1 Survey of current and emerging single-cell epigenetics techniques

Technique	Epigenomic feature	Method	Approach	Single cell
Cytosine modification	5mC	BS-seq	Bisulfite converts C but not 5mC (or 5hmC) to U so only methylated sites are sequenced as "C"	Yes [6–8]
	5mC	MeDIP-seq	Immunoprecipitation of 5mC DNA followed by sequencing	Not currently possible
	5mC	Methyl-seq	Restriction enzyme specific for 5mC followed by sequencing	Possible
	5hmC	oxBS-seq	5hmC is oxidized to 5caC so that only 5mC survives bisulfite conversion. Readout is pure 5mC and subtraction from BS-seq determines 5hmC	Not possible for measuring 5hmC due to the need for subtraction
	5hmC	TAB-seq	Maps 5hmC by enzymatic oxidation prior to bisulfite treatment: only 5hmC survives conversion	Possible
	5hmC	hMeDIP-seq	Immunoprecipitation of 5hmC DNA followed by sequencing	Not currently possible
	5hmC	Aba-seq	Restriction enzyme specific for 5hmC	Possible
Protein–DNA interaction	Histone modification	ChIP-seq	Immunoprecipitation of DNA bound to a specific histone variant or transcription factor	Yes [17]
	Transcription factor binding	DamID	Cells are transfected with a fusion of a transcription factor gene and Dam protein which methylates adenine residues in proximity to the binding site. 6 mA specific restriction digest is used to map	Yes for nuclear lamina interactions [18]
Chromatin structure	Nucleosome positioning	MNase-seq	Microcococal nuclease digestion of chromatin and sequencing of the product which are regions protected by nucleosomes	Possible
	Nucleosome positioning	NOME-seq	GpC methylation of DNA not protected by nucleosomes followed by BS-seq	Possible
	DNA accessibility	DNase-seq	DNaseI digestion of open chromatin into small fragments suitable for library preparation and sequencing	Yes [23]
	DNA accessibility	FAIRE-seq	Chromatin is crosslinked, sonicated, and then purified by phenol–chloroform extraction. The aqueous layer contains only DNA not associated with protein	Not currently possible
	DNA accessibility	ATAC-seq	Tn5 transposase enzyme fragments and attaches adapters to open chromatin	Yes [21, 22]
Three-dimensional organization	Chromosome conformation	HiC	DNA is crosslinked, then restriction digested to fragment before ligation and reversal of the crosslinks. Resulting fragments are hybrids from separate genomic locations that were in close proximity in three-dimensional space. Paired-end sequencing is used to link the two regions	Yes [29]

C cytosine, *5caC* 5-carboxylcytosine, *5hmC* hydroxymethylcytosine, *5mC* methylcytosine, *U* uracil

The first single-cell method for measuring genome-wide 5mC used a reduced representation bisulfite sequencing (scRBBS) approach based on enrichment of CpG dense regions (such as CpG islands) via restriction digestion, and it allows the measurement of approximately 10 % of CpG sites [6]. scRRBS is powerful because it allows assessment of a large fraction of promoters with relatively low sequencing costs, but its limitation is poor coverage of many important regulatory regions such as enhancers.

To develop true whole-genome single-cell approaches [7, 8] technological developments have been based on a

Clark *et al. Genome Biology* (2016) 17:72

Page 3 of 10

Fig. 1 Epigenomics and the spectrum of single-cell sequencing technologies. The diagram outlines the single-cell sequencing technologies currently available. A single cell is first isolated by means of droplet encapsulation, manual manipulation, fluorescence-activated cell sorting (*FACS*) or microfluidic processing. The first examples of single-cell multi-omic technologies have used parallel amplification or physical separation to measure gene expression (*scRNA-seq*) and DNA sequence (*scDNA-seq*) from the same cell. Note that single-cell bisulfite conversion followed by sequencing (*scBS-seq*) is not compatible with parallel amplification of RNA and DNA, as DNA methylation is not conserved during in vitro amplification. Single-cell epigenomics approaches utilize chemical treatment of DNA (bisulfite conversion), immunoprecipitation or enzymatic digest (e.g., by DNaseI) to study DNA modifications (*scBS-seq* and *scRRBS*), histone modifications (*scChIP-seq*), DNA accessibility (*scATAC-seq*, *scDNase-seq*), chromatin conformation (*scDamID, scHiC*)

post-bisulfite adapter-tagging (PBAT) approach in which bisulfite conversion is performed before library preparation so that DNA degradation does not destroy adaptor-tagged fragments [9]. As a result, methylation in up to 50 % of the CpG sites in a single cell can now be measured and this has allowed, for example, the detection of high variability between single cells in distal enhancer methylation (not usually captured by scRRBS) in mouse embryonic stem cells (ESCs) [7].

Building on this method has allowed BS-seq and RNA-seq in parallel from the same single cell (scM&T-seq) [10]. This was made possible by way of

Clark *et al. Genome Biology* (2016) 17:72

Page 4 of 10

a method for physical separation of poly-A mRNA from DNA (genome and transcriptome sequencing or G&T-seq [11]), and this now allows intricate investigations of links between epigenetic and transcriptional heterogeneity within a particular cell and tissue type.

Hydroxymethylated cytosine (5hmC) is also thought to have a role in epigenetic gene regulation and has been analyzed in bulk samples using modified bisulfite sequencing methods [12, 13], 5hmC-specific restriction enzymes [14], or immunoprecipitation [15]. Of the currently established methods, TET-assisted bisulfite sequencing (TAB-seq) [12] and Aba-seq [14] could potentially be adapted to single cells. In TAB-seq, 5hmC is first enzymatically glucosylated in order to prevent its recognition by TET1, which is then used to oxidize 5mC to 5-formylcytosine and 5-carboxylcytosine which, along with unmodified cytosines, are sensitive to bisulfite conversion. These initial enzymatic steps could be performed in a single-tube reaction immediately before processing by single-cell BS-seq (scBS-seq). In Aba-seq, 5hmc is glucosylated prior to digestion with AbaSI, an enzyme that recognizes 5-glucosylhydroxymethylcytosine, and then prepared for sequencing by adapter ligation. Importantly, both of these techniques would be compatible with DNA purified using G&T-seq [11] thus allowing parallel measurements of 5hmC and poly-A RNA within the same single cell.

Histone modifications and transcription factor binding

Histones can carry a diversity of covalent modifications that are associated with different genomic features and transcriptional states [16]. Mapping of histone marks is typically carried out using chromatin immunoprecipitation followed by sequencing (ChIP-seq). Performing ChIP-seq at the single-cell level is extremely challenging due to background noise caused by nonspecific antibody pull-down, which increases as the level of target antigen decreases. This was overcome recently by performing the immunoprecipitation step on chromatin from a pool of single cells that had already undergone micrococcal nuclease (MNase) digestion and barcoding, so that the pull-down is effectively performed on thousands of cells. This approach used a droplet-based microfluidics setup to process large numbers of cells in parallel [17], and because only a limited number of valid sequencing reads are obtained per single cell a large number of cells has to be sequenced in order to evaluate intercellular variability.

Protein–DNA interactions in single cells have been mapped using DamID, in which a cell line expresses low levels of a fusion protein of *Escherichia coli* deoxyadenosine methylase (Dam) and the protein under study. Dam methylates DNA on adenine residues adjacent to sites of protein binding. Methylated sites are then cut by the methylation-sensitive restriction enzyme DpnI, followed by ligation of sequencing adapters. This technique has been successfully employed to study interactions with the nuclear lamina in single cells [18]. Currently resolution is in the order of 100 kb, which to some extent limits its applications, but future optimizations could see improvements such that it could be used for mapping transcription factor binding sites in single cells. In addition, single-cell DamID could also support genome-wide analysis of histone modifications by using Dam fusion with specific histone readers or modifiers.

Chromatin structure and chromosome organization

A raft of publications was seen in 2015 describing methods for mapping open chromatin in single cells. The first of these was based on the assay for transposase-accessible chromatin (ATAC-seq) which uses a Tn5 transposase enzyme to simultaneously fragment DNA and attach adapter sequences in a process called tagmentation [19]. Open chromatin regions can be defined by introducing the transposase into intact nuclei, where it acts on only DNA free of nucleosomes and transcription factors [20]. ATAC-seq was first adapted to single-cell resolution by employing a "combinatorial indexing" strategy in which the tagmentation is carried out on 96 pools of a few thousand nuclei, introducing a unique barcode to each pool. The 96 reactions are then pooled and split before a second barcode is introduced by polymerase chain reaction (PCR). The number of pools and cells per pool are optimized so that the probability that a particular barcode combination originates only from a single cell is kept sufficiently high [21]. In parallel a second single-cell ATAC-seq method has been described, which makes use of a commercially available microfluidics device to carry out the transposition reaction on individual cells [22]. This approach has resulted in a large increase in resolution compared with the combinatorial indexing method, mapping an average of 70,000 reads per cell compared with 3000, although throughput was substantially lower. Finally, investigation of open chromatin genomic regions has been achieved in single cells by employing a DNase-seq approach to map regions that are DNaseI hypersensitive. scDNase-seq provides an improved resolution of 300,000 mapped reads per single cell, albeit with a very low mapping efficiency (2 %) and even lower throughput [23]. Both of these methods could be combined with RNA-seq, either by way of physical separation [10] or parallel amplification [24].

In bulk samples, genome-wide nucleosome occupancy has been assayed by sequencing the products of MNase digestion [25] and by nucleosome occupancy and methylome sequencing (NOMe-seq) [26].

Clark *et al. Genome Biology* (2016) 17:72

Page 5 of 10

In NOMe-seq, a methyltransferase enzyme is used to methylate exposed GpC dinucleotides while DNA bound by nucleosomes is protected. Sequencing of the bisulfite-converted DNA can then be used to map nucleosome positions and this is particularly attractive for single-cell use since it will also give a readout of CpG methylation within the same single cell. Indeed nucleosome positioning has already been studied using locus-specific bisulfite PCR in the yeast *PHO5* gene, which revealed significant variability between cells that correlated with gene expression [27]. Single-cell nuclei prepared according to this method should be compatible with scBS-seq.

In addition to defining the linear chromatin organization of single cells, it is now possible to assess chromosome conformation at the single-cell level using a HiC-based method [28, 29]. Single-cell HiC is currently limited in its resolution but still allows description of the individual chromosome organization and compartmentalization, as well as interchromosomal interactions. This is a good example of how single-cell approaches can really provide cutting-edge tools, as regular HiC was traditionally performed on millions of cells resulting in an average of all chromosome organization within the cell population and hence some ambiguity in interpretation of the results.

Advances in equipment to perform single-cells methods

Development of single-cell approaches is intimately linked to the development of physical equipment and devices. The first step in any single-cell analysis is the isolation and lysis of single cells from culture or dissociated tissue. This can be performed manually with a pipette and a microscope but such methods cannot realistically be scaled up for higher-throughput requirements. Fluorescence-activated cell sorting (FACS) can be used to isolate many thousands of single cells into microtiter plates in a short time with the additional benefit of being able to select cells based on a subset of fluorescent markers. Microfluidics systems have been developed, such as the C1 from Fluidigm, in which cells are trapped in chambers in which lysis and RNA-seq library preparation can subsequently be carried out. One advantage of this system is that captured cells are photographed in the system meaning that morphology of cells and presence of doublets can be assessed post hoc; however, these devices are currently only low-to-medium throughput, typically processing only 96 cells at one time. Recently, an innovative approach to generating single-cell libraries using microfluidics has emerged, which allows a significant increase in library preparation throughput compared with other methods. Cells are encapsulated within aqueous droplets in flowing oil, in which early stages of library preparation including cell-specific barcoding are performed, before being pooled

for downstream reactions. In this way, thousands of cells are processed in parallel, with vastly reduced costs per cell and improved sensitivities compared with conventional tube-based methods. So far this approach has been applied to RNA-seq [30–32] and ChIP-seq [17] but in principle it could also be adapted to other methods such as BS-seq. Commercialization of droplet-sequencing technologies has already begun, meaning that these single-cell methodologies will be easily accessible and will be able to achieve their full potential.

In parallel we are witnessing a significant improvement in the field of single-molecule sequencing technologies with the potential to measure DNA modifications directly from native DNA and over tens of kilobases on the same molecule. This is particularly relevant since PCR amplification, bisulfite treatment, and other manipulations involved in library preparation can introduce technical artifacts, e.g., CG bias in BS-seq libraries. There are currently two single-molecule sequencing technologies on the market. The first of these, single-molecule real-time sequencing [33] as employed by the Pacific Biosciences RSII and Sequel machines works by real-time measurements of incorporating nucleotides and has been shown to discriminate cytosine from 5mC and 5hmC, although much effort is yet required to enable this analysis to be performed in a flexible way and routinely due to the subtle and context-specific effects of DNA modifications on incorporation kinetics [34]. The other technology, marketed by Oxford Nanopore, uses measurements of electrostatic charge as a DNA strand passes through a protein nanopore. While this technology is still in its infancy, in principle modified bases such as 5mC and derivatives could be detected [35]. These technologies currently require microgram quantities of DNA and therefore are not directly applicable to single cells; however, the use of cell-specific barcoding followed by pooling many thousands of single cells could allow analysis of individual cells.

Quality control of single-cell epigenomic libraries

Quality control of sequencing data is crucial in order to avoid technical artifacts. This is especially true of single-cell sequencing, which is technically noisy due to low amounts of starting material, often resulting in variable capture efficiencies. The large number of amplification cycles that are needed often means that reagent contamination or sample cross-contamination is a very real problem and so sequencing of negative controls is recommended. Mapping efficiency or coverage cut-offs are also useful in order to eliminate cells that have performed much worse than the average. The use of spike-in controls may also be useful for some methods, for example to measure underconversion and overconversion

Clark *et al. Genome Biology* (2016) 17:72

Page 6 of 10

by bisulfite. Another important consideration is the effect of cell dissociation on downstream analysis, since harsh enzymatic digestion of solid tissues is thought to influence single-cell transcriptomic studies [36]. Although epigenomic profiles are generally thought to be more stable than transcriptomes, the dissociation of cells should be performed as quickly and as mildly as possible to minimize the potential influence on single-cell libraries. Finally, batch effects can have a profound influence on single-cell datasets [37], so it is important to process samples in parallel wherever possible.

Computational challenges to analyzing single-cell epigenomic data

The main computational challenges in single-cell data arise from the technical variability in the methods; this is due to low and variable capture efficiencies and biases introduced during PCR. This is a problem since it can be difficult to determine whether an observed difference is due to biological or technical reasons. These have been discussed in detail with respect to single-cell RNA-seq [38] where technical variability can be measured and normalized by use of synthetic spike-ins and unique molecular identifiers that are introduced during reverse transcription [39]. Single-cell epigenomic methods would likely benefit from similar strategies. Additionally, local correlations in epigenetic marks and correlations between epigenetic features and the underlying genetic sequence mean that missing information can be imputed in order to reduce the effect of low coverage. Such methods are in development and will be much improved when combined single-cell technologies become more sophisticated.

Future applications of single-cell epigenomics
Single-cell approaches to refine our understanding of epigenetic regulation

As mentioned above, epigenetic modifications have been characterized as transcriptionally repressive or activating based on correlations made in bulk cell populations. However, growing evidence has exposed the naivety of this assumption and revealed the great complexity of epigenetic regulation. For example, 5mC has long been considered to be a transcriptionally repressive mark since promoter methylation is negatively correlated with gene expression. However, in some cases DNA methylation of gene bodies has been positively correlated with transcription, demonstrating that the genomic context can influence the biological outcome [40]. Furthermore, global DNA hypomethylation seen in naive ESCs is not associated with widespread transcriptional activation [41, 42], demonstrating that the strength of regulatory links between DNA methylation and transcription can also vary depending on the developmental stage and

cellular context. Since the discovery of 5hmC and other oxidized derivatives of 5mC, the situation has become even less clear, with inconsistent reports on the biological functions of these modifications [43–46].

Therefore, the use of single-cell approaches has the potential to refine our understanding of DNA modifications as regulatory epigenetic marks. The recent development of combined single-cell methods (e.g., scM&T-seq) will be invaluable to such studies [10]. In addition, the very low levels of 5hmC measured in bulk cell samples (e.g., less than 5 % of CpG sites in primed ESCs) indicate that only a few cells in that population have this modification at any particular cytosine residue. Therefore, parallel profiling of 5hmC and transcription will impact profoundly on our understanding of this epigenetic mark. In the future, it may even be possible to assay multiple epigenetic features (e.g., DNA methylation and chromatin accessibility) together with gene expression in the same single cell, leading to further refinements in our view of the epigenomic influence on the transcriptome.

According to the classical definition, epigenetic modifications must be heritable through cell divisions. While the mechanisms governing 5mC maintenance during DNA replication have been well described [47], the inheritance of other components of the epigenome is understood poorly. For example, the means by which histone modifications are conserved through DNA replication remain unclear [48]. This represents another application of single-cell approaches, in which one can imagine in vitro systems where mother and daughter cells can be sequenced to reveal the distribution of epigenetic marks between these two cells. When coupled with manipulations of epigenetic modulators (e.g., knockout models of histone-modifying enzymes), such an approach would allow the true nature of epigenetic propagation to be elucidated.

Single-cell approaches to understand developmental processes and improve regenerative medicine

Single-cell transcriptional profiling has revealed population substructure in various developmental contexts [31, 32, 49–52]. In combination with lineage-tracing experiments, this information can be used to decipher the cellular hierarchy underlying complex tissues, giving unprecedented information on the molecular mechanisms governing differentiation processes. Epigenetic mechanisms are conventionally thought to restrict cell-fate decisions during development [53], so single-cell epigenomics studies will add valuable detail to these tissue hierarchies. It is also not excluded that in certain situations epigenetic information could be instructive for cell-fate decisions, and finely timed combined single-cell profiling techniques may provide insights into this important question.

Clark *et al. Genome Biology* (2016) 17:72

Page 7 of 10

Embryonic development involves global remodeling of the mammalian epigenome [38–40], including incorporation of maternal histones into the paternal genome following fertilization, and mechanisms leading to global DNA demethylation in both the preimplantation embryo and developing primordial germ cells. For this reason, many single-cell epigenomics techniques have been applied to embryonic development, taking mouse ESCs as a model system [6–8, 10, 29]. These studies have revealed intercellular epigenetic heterogeneity in cells poised for differentiation, which may have biological importance in lineage priming [54].

The near future will undoubtedly witness the application of single-cell epigenomic approaches in vivo. For example, mouse zygotes fertilized in vitro and embryos resulting from natural matings will be studied to understand epigenome dynamics during this critical stage of development. Due to the low cell numbers associated with these samples, FACS isolation of single cells is infeasible, so single cells will be manually picked after embryo dissociation. For the early stages of development, it should be possible to study every cell isolated from an embryo, while at later time points (E6.5 onwards) the increasing cell number may necessitate focused studies on specific cell lineages or on representative subpopulations of each lineage. A limitation to these studies will be the loss of spatial information upon embryo dissociation. Complementary studies including in vivo imaging of lineage-specific genes will be used to map cell types identified by single-cell sequencing back to the three-dimensional embryo [55]. By employing single-cell multi-omics, these studies will reveal the fundamental processes of cell-fate specification and establish an atlas of differentiation in which every tissue type can be traced back to its embryonic origins. This information will bring light to one of the most fascinating processes of biology, clarifying key questions such as whether cell-type-specific epigenetic marks are established during lineage priming prior to cell-fate commitment.

In addition, these experiments will have important applications in the clinic. For example, such information will assist efforts to reprogram cells from adult tissues into induced pluripotent stem cells (iPSCs). The inefficiency of this process is currently limiting the applicability of iPSCs to regenerative medicine, so single-cell gene expression analyses have been performed to decipher the molecular pathway to successful reprogramming [56–58]. Somatic cell reprogramming is known to be associated with dramatic nuclear remodeling [59, 60], so single-cell epigenomic studies will add an important layer of information. Furthermore, a detailed understanding of the mechanisms involved in cell-fate decisions in v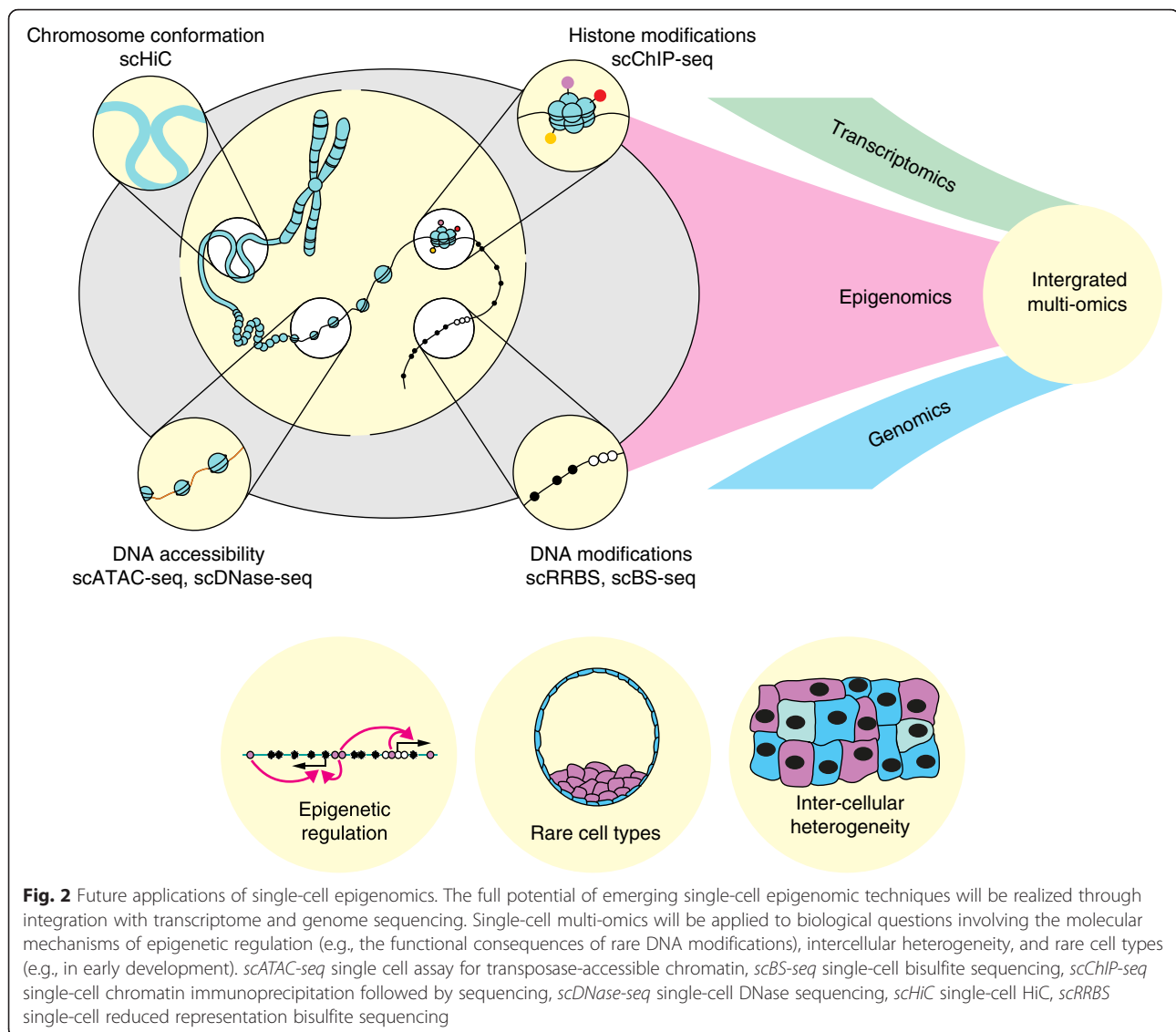ivo will improve our ability to generate specific cell types (from iPSCs or other stem cells) for therapeutic use in regenerative medicine.

Single-cell approaches to assess the complexity of cancer

Cancer is a highly heterogeneous disease with molecular characteristics that depend on the tissue of origin and differ between patients. Intratumor heterogeneity (within patients) is not fully understood, but includes regional differences that reflect tumor microenvironment, differences between primary and metastatic disease, and genetic diversity resulting from tumor evolution. Emerging single-cell sequencing technologies will reveal the full extent of intratumor heterogeneity and this will have many applications for clinical management as different cell types are likely to play distinct roles in disease initiation, metastasis, and drug resistance [61]. Already, single-cell DNA sequencing has found evidence of clonal evolution in multiple cancer types, and has identified founder mutations and subclonal mutations that have implications for cancer progression [62, 63]. Likewise, single-cell transcriptome profiling has been used to identify cell subpopulations within cancers, including cells with transcriptional programs suggesting stem cell activity [64–66]. These studies have extended our understanding of disease progression and have improved our ability to predict disease outcome.

The epigenome is known to be drastically remodeled in multiple malignancies, and therapeutics targeting DNA methyltransferases and histone deacetylases are used in several cancer types [67]. Typically, loss of DNA methylation is observed on a global scale while gains in DNA methylation occur in a more specific manner, and these changes are accompanied by abnormal nucleosome positioning and chromatin modifications. Descriptions of intertumor epigenetic heterogeneity have yielded clinically relevant information (e.g., stratification of triple-negative breast cancers into subgroups with differing prognosis [68]), but the full extent of epigenetic intratumor heterogeneity remains unknown and will rely upon single-cell analyses.

In the future, single-cell epigenomic studies will complement single-cell transcriptome and genome analysis in defining rare subpopulations of cells with clinically significant characteristics. For example, cancer stem cells could be characterized using these single-cell studies, such that targeted therapeutics can be designed to prevent disease recurrence following conventional therapy [69]. Single-cell epigenomic studies may also lead to the development of novel screening strategies based on circulating tumor cells and cell-free DNA, where patient material is severely limited. In particular, DNA methylation is an attractive target for cancer screening as it provides cell-type-specific information that is more stable than transcriptional profiles.

Clark *et al. Genome Biology* (2016) 17:72

Page 8 of 10

Fig. 2 Future applications of single-cell epigenomics. The full potential of emerging single-cell epigenomic techniques will be realized through integration with transcriptome and genome sequencing. Single-cell multi-omics will be applied to biological questions involving the molecular mechanisms of epigenetic regulation (e.g., the functional consequences of rare DNA modifications), intercellular heterogeneity, and rare cell types (e.g., in early development). *scATAC-seq* single cell assay for transposase-accessible chromatin, *scBS-seq* single-cell bisulfite sequencing, *scChIP-seq* single-cell chromatin immunoprecipitation followed by sequencing, *scDNase-seq* single-cell DNase sequencing, *scHiC* single-cell HiC, *scRRBS* single-cell reduced representation bisulfite sequencing

Conclusions

In conclusion, the field of single-cell epigenomics is in its infancy but with the rapid pace of technological development and the increasingly recognized importance of intercellular heterogeneity we anticipate enormous progress over the next few years. Methods are evolving such that researchers will soon be able to profile multiple epigenetic marks within the same single cell and do so in combination with transcriptional and genetic information (Fig. 2). Correlations between features at precise genomic locations will lead to a more refined appreciation of how epigenetic processes interact with one another to control gene expression. Ultimately this has the potential to transform our understanding of how the phenotype of the cell is maintained and how it is perturbed in disease—a subject that is fundamental to biology.

Authors' contributions

SJC, HJL, and SAS contributed equally to the manuscript, and prepared the manuscript, table and figures. GK and WR read and approved the complete manuscript.

Clark *et al. Genome Biology* (2016) 17:72

Page 9 of 10

Author details
[1]Epigenetics Programme, Babraham Institute, Cambridge CB22 3AT, UK. [2]Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK. [3]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, CH 4058 Basel, Switzerland. [4]Centre for Trophoblast Research, University of Cambridge, Cambridge CB2 3EG, UK. [5]Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 3EG, UK.

Published online: 18 April 2016

References
1. Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. Cell. 2007;128:669–81.
2. Macaulay IC, Voet T. Single cell genomics. Advances and future perspectives. PLoS Genet. 2014;10:e1004126.
3. Brunner AL, Johnson DS, Kim SW, Valouev A, Reddy TE, Neff NF, et al. Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver. Genome Res. 2009;19:1044–56.
4. Down TA, Rakyan VK, Turner DJ, Flicek P, Li H, Kulesha E, et al. A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. Nat Biotechnol. 2008;26:779–85.
5. Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, et al. Shotgun bisulfite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. Nature. 2008;452:215–9.
6. Guo H, Zhu P, Wu X, Li X, Wen L, Tang F. Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. Genome Res. 2013;23:2126–35.
7. Smallwood SA, Lee HJ, Angermueller C, Krueger F, Saadeh H, Peat J, et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat Methods. 2014;11:817–20.
8. Farlik M, Sheffield NC, Nuzzo A, Datlinger P, Schönegger A, Klughammer J, et al. Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics. Cell Rep. 2015;10:1386–97.
9. Miura F, Enomoto Y, Dairiki R, Ito T. Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. Nucleic Acids Res. 2012;40:e136.
10. Angermueller C, Clark SJ, Lee HJ, Macaulay IC, Teng MJ, Hu TX, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. Nat Methods. 2016;13:229–32.
11. Macaulay IC, Haerty W, Kumar P, Li YI, Hu TX, Teng MJ, et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. Nat Methods. 2015;12:519–22.
12. Yu M, Hon GC, Szulwach KE, Song C-X, Zhang L, Kim A, et al. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell. 2012;149:1368–80.
13. Booth MJ, Branco MR, Ficz G, Oxley D, Krueger F, Reik W, et al. Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science. 2012;336:934–7.
14. Sun Z, Terragni J, Borgaro JG, Liu Y, Yu L, Guan S, et al. High-resolution enzymatic mapping of genomic 5-hydroxymethylcytosine in mouse embryonic stem cells. Cell Rep. 2013;3:567–76.
15. Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, et al. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature. 2011;473:398–402.
16. Zhou VW, Goren A, Bernstein BE. Charting histone modifications and the functional organization of mammalian genomes. Nat Rev Genet. 2011;12:7–18.
17. Rotem A, Ram O, Shoresh N, Sperling RA, Goren A, Weitz DA, et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. Nat Biotechnol. 2015;33:1165–72.
18. Kind J, Pagie L, de Vries SS, Nahidiazar L, Dey SS, Bienko M, et al. Genome-wide maps of nuclear lamina interactions in single human cells. Cell. 2015;163:134–47.
19. Adey A, Morrison HG, Asan, Xun X, Kitzman JO, Turner EH, et al. Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. Genome Biol. 2010;11:R119.
20. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013;10:1213–8.
21. Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, et al. Epigenetics: multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. Science. 2015;348:910–4.
22. Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, et al. Single-cell chromatin accessibility reveals principles of regulatory variation. Nature. 2015;523:486–90.
23. Jin W, Tang Q, Wan M, Cui K, Zhang Y, Ren G, et al. Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples. Nature. 2015;528:142–6.
24. Dey SS, Kester L, Spanjaard B, Bienko M, van Oudenaarden A. Integrated genome and transcriptome sequencing of the same cell. Nat Biotechnol. 2015;33:285–9.
25. Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, et al. Dynamic regulation of nucleosome positioning in the human genome. Cell. 2008;132:887–98.
26. Kelly TK, Liu Y, Lay FD, Liang G, Berman BP, Jones PA. Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. Genome Res. 2012;22:2497–506.
27. Small EC, Xi L, Wang J-P, Widom J, Licht JD. Single-cell nucleosome mapping reveals the molecular basis of gene expression heterogeneity. Proc Natl Acad Sci U S A. 2014;111:E2462–71.
28. Nagano T, Lubling Y, Yaffe E, Wingett SW, Dean W, Tanay A, et al. Single-cell Hi-C for genome-wide detection of chromatin interactions that occur simultaneously in a single cell. Nat Protoc. 2015;10:1986–2003.
29. Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature. 2013;502:59–64.
30. Rotem A, Ram O, Shoresh N, Sperling RA, Schnall-Levin M, Zhang H, et al. High-throughput single-cell labeling (Hi-SCL) for RNA-Seq using drop-based microfluidics. PLoS One. 2015;10:e0116328.
31. Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell. 2015;161:1187–201.
32. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell. 2015;161:1202–14.
33. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, et al. Real-time DNA sequencing from single polymerase molecules. Science. 2009;323:133–8.
34. Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, et al. Direct detection of DNA methylation during single-molecule real-time sequencing. Nat Methods. 2010;7:461–5.
35. Wanunu M, Cohen-Karni D, Johnson RR, Fields L, Benner J, Peterman N, et al. Discrimination of methylcytosine from hydroxymethylcytosine in DNA molecules. J Am Chem Soc. 2011;133:486–92.
36. Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC, Teichmann SA. The technology and biology of single-cell RNA sequencing. Mol Cell. 2015;58:610–20.
37. Hicks SC, Teng M, Irizarry RA. On the widespread and critical impact of systematic bias and batch effects in single-cell RNA-Seq data. Paper presented at the 2015 PQG Conference: Single-cell genomics: technology, analysis, and applications, Boston, MA, 5–6 November 2015.
38. Stegle O, Teichmann SA, Marioni JC. Computational and analytical challenges in single-cell transcriptomics. Nat Rev Genet. 2015;16:133–45.
39. Kivioja T, Vaharautio A, Karlsson K, Bonke M, Enge M, Linnarsson S, et al. Counting absolute numbers of molecules using unique molecular identifiers. Nat Methods. 2012;9:72–4.
40. Schubeler D. Function and information content of DNA methylation. Nature. 2015;517:321–6.
41. Habibi E, Brinkman AB, Arand J, Kroeze LI, Kerstens HH, Matarese F, et al. Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. Cell Stem Cell. 2013;13:360–9.

Clark *et al. Genome Biology* (2016) 17:72

Page 10 of 10

42. Ficz G, Hore TA, Santos F, Lee HJ, Dean W, Arand J, et al. FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. Cell Stem Cell. 2013;13:351–9.

43. Branco MR, Ficz G, Reik W. Uncovering the role of 5-hydroxymethylcytosine in the epigenome. Nat Rev Genet. 2012;13:7–13.

44. Raiber EA, Beraldi D, Ficz G, Burgess HE, Branco MR, Murat P, et al. Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase. Genome Biol. 2012;13:R69.

45. Shen L, Wu H, Diep D, Yamaguchi S, D'Alessio AC, Fung HL, et al. Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. Cell. 2013;153:692–706.

46. Song C-X, Szulwach KE, Dai Q, Fu Y, Mao S-Q, Lin L, et al. Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. Cell. 2013;153:678–91.

47. Smith ZD, Meissner A. DNA methylation. roles in mammalian development. Nat Rev Genet. 2013;14:204–20.

48. Margueron R, Reinberg D. Chromatin structure and the inheritance of epigenetic information. Nat Rev Genet. 2010;11:285–96.

49. Jaitin DA, Kenigsberg E, Keren-Shaul H, Elefant N, Paul F, Zaretsky I, et al. Massively parallel single-cell RNA-Seq for marker-free decomposition of tissues into cell types. Science. 2014;343:776–9.

50. Treutlein B, Brownfield DG, Wu AR, Neff NF, Mantalas GL, Espinoza FH, et al. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature. 2014;509:371–5.

51. Zeisel A, Muñoz-Manchado AB, Codeluppi S, Lönnerberg P, La Manno G, Juréus A, et al. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science. 2015;347:1138–42.

52. Shin J, Berg DA, Zhu Y, Shin JY, Song J, Bonaguidi MA, et al. Single-cell RNA-seq with waterfall reveals molecular cascades underlying adult neurogenesis. Cell Stem Cell. 2015;17:360–72.

53. Hemberger M, Dean W, Reik W. Epigenetic dynamics of stem cells and cell lineage commitment. Digging Waddington's canal. Nat Rev Mol Cell Biol. 2009;10:526–37.

54. Lee HJ, Hore TA, Reik W. Reprogramming the methylome. Erasing memory and creating diversity. Cell Stem Cell. 2014;14:710–9.

55. Achim K, Pettit JB, Saraiva LR, Gavriouchkina D, Larsson T, Arendt D, et al. High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. Nat Biotechnol. 2015;33:503–9.

56. Zunder ER, Lujan E, Goltsev Y, Wernig M, Nolan GP. A continuous molecular roadmap to iPSC reprogramming through progression analysis of single-cell mass cytometry. Cell Stem Cell. 2015;16:323–37.

57. Kim DH, Marinov GK, Pepke S, Singer ZS, He P, Williams B, et al. Single-cell transcriptome analysis reveals dynamic changes in lncRNA expression during reprogramming. Cell Stem Cell. 2015;16:88–101.

58. Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, Ganz K, et al. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell. 2012;150:1209–22.

59. Apostolou E, Hochedlinger K. Chromatin dynamics during cellular reprogramming. Nature. 2013;502:462–71.

60. Nashun B, Hill PW, Hajkova P. Reprogramming of cell fate: epigenetic memory and the erasure of memories past. EMBO J. 2015;34:1296–308.

61. Alizadeh AA, Aranda V, Bardelli A, Blanpain C, Bock C, Borowski C, et al. Toward understanding and exploiting tumor heterogeneity. Nat Med. 2015;21:846–53.

62. Navin NE. The first five years of single-cell cancer genomics and beyond. Genome Res. 2015;25:1499–507.

63. Francis JM, Zhang C-Z, Maire CL, Jung J, Manzo VE, Adalsteinsson VA, et al. EGFR variant heterogeneity in glioblastoma resolved through single-nucleus sequencing. Cancer Discov. 2014;4:956–71.

64. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science. 2014;344:1396–401.

65. Lawson DA, Bhakta NR, Kessenbrock K, Prummel KD, Yu Y, Takai K, et al. Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. Nature. 2015;526:131–5.

66. Dalerba P, Kalisky T, Sahoo D, Rajendran PS, Rothenberg ME, Leyrat AA, et al. Single-cell dissection of transcriptional heterogeneity in human colon tumors. Nat Biotechnol. 2011;29:1120–7.

67. Baylin SB, Jones PA. A decade of exploring the cancer epigenome—biological and translational implications. Nat Rev Cancer. 2011;11:726–34.

68. Stirzaker C, Zotenko E, Song JZ, Qu W, Nair SS, Locke WJ, et al. Methylome sequencing in triple-negative breast cancer reveals distinct methylation clusters with prognostic value. Nat Commun. 2015;6:5899.

69. Beck B, Blanpain C. Unravelling cancer stem cell potential. Nat Rev Cancer. 2013;13:727–38.