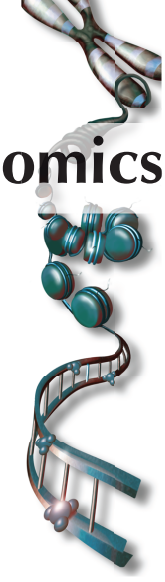
Epigenomics in the single cell era, an important read out for genome function and cell identity

Inês Pinheiro, Maria-Elena Torres-Padilla, Geneviève Almouzni

Editorial

For reprint orders, please contact: reprints@futuremedicine.com

Epigenomics



Epigenomics in the single cell era, an important read out for genome function and cell identity

Inês Pinheiro[1], Maria-Elena Torres-Padilla[2,3] & Geneviève Almouzni*,[1]

[1]Institut Curie, CNRS, PSL Research University, LabEx DEEP, Sorbonne Université, Nuclear Dynamics Unit, Equipe Labellisée Ligue Contre le Cancer, 75248 Paris Cedex 05, France

[2]Institute of Epigenetics and Stem Cells (IES), Helmholtz Zentrum München – German Research Center for Environmental Health, Munich 81377, Germany

[3]Faculty of Biology, Ludwig-Maximilians Universität, 82152 Martinsried, Germany

*Author for correspondence: genevieve.almouzni@curie.fr

"epigenomics research in an era where mapping cellular trajectories and mechanisms offers unique opportunities to address fundamental questions related to cell identity of key importance for medical applications."

First draft submitted: 16 March 2021; Accepted for publication: 29 April 2021; Published online: 11 June 2021Could you please provide 3-8 keywords?single-cell technologies, epigenomics, cell fate, cellular trajectories, disease c

Keywords: cell fate • cellular trajectories • disease onset • epigenomics • single-cell technologies • epigenetics • machine learning

While major advances in single cell transcriptomics have occurred, the epigenomics arena has only more recently gained momentum. Indeed, methods to track changes in different molecules present in biological samples have developed both in scale – providing a global view in each case – and in terms of resolution, allowing us today to disentangle heterogeneity within samples and access at the level of a single cell. The application of omics methods, comprising analysis of genomes, epigenomes, transcriptomes, proteomes and metabolomes, among other functional genomics techniques, individually or in combination, provides a wealth of valuable datasets. The treatment of these data with the development of artificial intelligence tools opens unprecedented avenues. The possibility to gain a holistic view of cells functioning within organs and during their formation will allow us to tackle fundamental questions of normal organism development and also to reveal early changes associated with disease onset and progression. In this editorial, we discuss the place of epigenomics research in an era where mapping cellular trajectories and mechanisms offers unique opportunities to address fundamental questions related to cell identity of key importance for medical applications.

Single-cell epigenomics & the cell state

The recent years have witnessed the dawn of an unprecedented development of single-cell technologies, which will be crucial to unravel not only cell diversity within organs but also to uncover a myriad of unknown cell types, urging the scientific community to revisit the very concept of cell identity [1]. Analyses of differentially expressed genes in cell populations by the most widely used single-cell method, single-cell RNA-sequencing, have enabled investigators to cluster cells in groups, some of which can contain intrinsic diversity (possibly corresponding to different phenotypic and functional cell states within a same cell type) as a result of developmental cues including cell cycle, stress signals, circadian rhythm, chemical and mechanical spatial stimuli among other signals [2]. This creates challenges when distinguishing previously undescribed cell types from a known cell type in a different state. In addition, this has recreated interest towards redefining the cell fate decision process as a continuum rather than as a binary process. As transcription can occur in bursts and can be highly responsive to external and internal stimuli, combination of different modalities can only outperform cell clustering based on one single method. Single-cell methods describing more stable, inherited epigenetic features might represent an appropriate complementary modality to help us distinguish cell types from cell states [3]. Interrogating RNA and chromatin using multi-omics

Future Medicine

methods will precisely allow us to measure these different molecular layers from the same cell [4]. Furthermore, they will provide the necessary knowledge to enable researchers to go beyond the description of cellular states towards understanding the mechanisms enabling the generation of those identities. While these methods are progressively being developed, data analysis tools such as multi-omics factor analysis [5] enable the integration of different data modalities and the characterization of the relationship between multiple molecular layers, therefore contributing to a comprehensive definition of cell types and states.

Specific chromatin regulators seem to be crucial for cellular plasticity and maintenance of cell identity, such as CAF-1 [6,7], and other histone chaperones and histone variants variants [8]. Considering this, epigenomic modifications, chromatin accessibility and chromatin conformation [3,9] are all important features that, together with additional single-cell functional genomics readouts, will contribute to refining the definition of cell identities, types and states and to follow their fate. They will also enable the urgent need to identifying the molecular mechanisms behind cell fate decisions for their functional characterisation and manipulation. These are pre-requisites for any efforts of cell fate manipulation, for example towards regenerative therapies.

Epigenome & transcription: solving the hierarchy toward earlier detection of diseases?

Integration of data obtained from the various single-cell omics methods will be a powerful tool to understand the relationships between the different types of biological molecules in a cell. Moreover, recent development of the combination of single-cell epigenomics and single-cell transcriptomics from the same single cell will allow us to refine long-standing questions, such as the relationship between chromatin features and transcriptional statuses [4]. Revisiting these fundamental questions with the newest technologies will contribute to identify the earliest signs of changes in cell trajectories during development or of deviation toward disease onset. In this regard, learning lessons from *in vivo* developmental processes, which direct and enable dramatic changes in cellular plasticity, remains a priority. Over recent years this approach has enabled namely the surge of organoids as 'petri-dish' model systems, which will continue to be powerful to model e.g. disease states.

Using bulk methods has already provided examples of dynamic and hierarchical events occurring during re-programming, where transcription factors can induce changes in local chromatin states before alteration of the transcriptional status [10]. More specifically in a pathological context, metastasis-inducing pathways can modulate histone chaperones, altering the chromatin at promoters of metastasis-inducing transcription factor genes, resulting in the acquisition of more aggressive traits [11]. More recently, an epigenetic gene-environment link was established between mutations in the weak oncogene *KRAS* and the appearance of malignant pancreatic ductal adenocarcinoma [12]. Mutations on *KRAS*, together with tissue injury by environmental factors, establish a cancer-associated chromatin accessibility profile before malignant cell transformation occurs.

These findings not only highlight the importance of mechanistic studies to understand tissue injury and the role of environmental factors, but also place chromatin alterations as an early event in cellular transformation and possibly early disease detection.

The ever-expanding observations of altered regulation of chromatin modifications and structure in disease highlight the importance of defining chromatin profiles in healthy tissues and during disease onset and progression, even more so with the newest single-cell technologies. But as these technologies advance equally fast for various cellular readouts, what is the added value of the information we can retrieve from single-cell characterization of chromatin states? Can chromatin regulation provide important mechanistic insights into cell plasticity during disease progression? Can chromatin states add an important layer of information as biomarkers for better patient stratification and more effective and personalized therapies?

Chromatin & disease

In this section, we choose examples to illustrate the importance of chromatin profiling in disease.

Cancer

Even though chromatin defects are prominent in cancer cells, it is not always obvious whether the observed alterations should be considered as drivers or consequences of cellular transformation. In any event, epigenetic alterations have been continually explored in oncology as biomarkers for disease detection and therapy response or as therapeutic targets [13,14]. Single-cell technologies, with the resolution that they can bring, offer new opportunities to identify 'needles in the haystack' to detect in a cell population through epigenetic and chromatin features the cells that show early defects. This advance will help to obtain better biomarkers and better patient stratification.

With these powerful technologies, there is much hope for a new ground to understand the role of the epigenome in cancer onset and development, to better define the evolutionary relationship between events and cells, and the earliest possible alterations occurring in disease.

Numerous studies illustrate the growing importance of epigenetic alterations in cancer. Bulk methods to define DNA methylation profiles of triple-negative breast cancer samples provided a possible prognostic signature of patient survival [15]. Single-cell ChIP-seq allowed the identification of rare populations of cells in breast cancer samples, more precisely cells with molecular signatures related with transformation toward therapy resistance in untreated tumour samples [16]. More recently, single-cell analyses revealed that overexpression of the histone variant CENP-A, a common feature of various cancers, can actually drive distinct cell fates and response to therapy depending on the status of p53 [17]. Indeed, chromatin marks represent attractive readouts for cancer screening as they provide stable cell-type specific information [18], improving patient stratification and contributing to more adequate disease management.

Besides their role as stable biomarkers, chromatin regulators can also play a mechanistic role in cancer progression, placing them as attractive drug targets in cancer therapy. There is potential for the combined use of drugs based on tumor chromatin profiling, sequential treatments increasing efficacy and drug tolerability, as well as new generations of more specific epi-drugs [13]. These are exciting avenues that will allow us to move toward precision approaches to cancer treatments.

Persistent viral infections

The understanding of HIV infections can also profit tremendously from the monitoring of chromatin states at the single-cell level. Successful combined antiretroviral therapy (cART) allows the control of the disease, but the virus persists in cellular reservoirs, where it integrates into the host's genome and remains in a latent state dependent on epigenetic mechanisms [19]. The chromatin features present at viral integration sites include CpG methylation of the HIV promoter, recruitment of histone deacetylases and presence of repressive histone marks such as H3K27me3 [14]. Full characterization of these marks is central to the total success of cART, as reactivation of the latent virus pool is needed to render the virus accessible to therapy and potentially achieve complete cure. Importantly, the still expanding atlas of HIV reservoir cell types also raises the issue of whether the epigenetic latency mechanisms might be heterogeneous, depending on the type of host reservoir cell. Beyond these examples, the potential is tremendous in many infectious diseases to understand the host–pathogen interaction.

Diabetes

Diabetes mellitus is characterized by a dysregulation of the glucose levels in blood, due to loss or dysfunction of insulin producing β-cells in the pancreas, ultimately leading to extremely serious complications. Use of single-cell technologies can be essential to fully characterize the disease and to identify rare cell populations of regenerative potential. In fact, integration of data from Assay for Transposase-Accessible Chromatin sequencing and RNA-seq from FACS-sorted cell populations reveals that α-cells can transdifferentiate into β-cells [20,21], and have a more flexible epigenome with areas of open chromatin corresponding to β-cell signature genes. Such α- to β-cell transition can be seen as a promising approach for β-cell regeneration and therapeutic solutions for diabetes [22].

All these examples highlight the importance that epigenomics can have in our understanding of diseases. Besides, as the epigenome also reflects individual lifestyle habits and exposure to a variety of environmental conditions, it represents a crucial readout in the development of personalized medical treatments.

Importantly, we still need to further investigate the role of the epigenome in the functioning of our cells, its relationship with the environment and their relationship with other cellular processes. Integrating epigenomics data with other functional genomics approaches, or analysis of the epigenome together with other modalities from the same cell in multi-omics approaches will provide a new understanding of the epigenome and facilitate the translation of this knowledge into the clinics. To fully take advantage of this wealth of information, it will be important to invest in method benchmarking as well as in data analyses tools and infrastructure, data curation and data harmonization so that such rich and costly datasets can be effectively combined and compared. There is no doubt that acting in a coordinated manner can be beneficial to all parts [17].

Coordinated action based on international research consortia represents a real asset if we wish to harness the potential of new technologies, benchmarking existing ones, analyzing the most relevant samples and integrating datasets originating from different methods. The emergence of research consortia or networks focused on epigenetics, such as EpiGeneSys, the International Human Epigenome Consortium or the 4D Nucleome, and other

10.2217/epi-2021-0153

recent consortia focused on building cell atlases of organs at a single-cell resolution (Human Cell Atlas) [23] or identifying single-cell events at the origin of cellular changes from healthy tissues toward disease (LifeTime and 37TrillionCells) [24] illustrate the efforts, the investment and the coordination needed in the field.

Financial & competing interest disclosure

The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

1. Morris SA. The evolving concept of cell identity in the single cell era. *Development* 146(12), 1-5 (2019).
2. Mayr U, Serra D, Liberali P. Exploring single cells in space and time during tissue development, homeostasis and regeneration. *Development* 146(12), (2019).
3. Ludwig CH, Bintu L. Mapping chromatin modifications at the single cell level. *Development* 146(12), (2019).
4. Chappell L, Russell AJC, Voet T. Single-Cell (Multi)omics Technologies. *Annu. Rev. Genomics Hum. Genet.* 19(1), 15–41 (2018).
5. Argelaguet R, Clark SJ, Mohammed H *et al.* Multi-omics profiling of mouse gastrulation at single-cell resolution. *Nature* 576(7787), 487–491 (2019).
6. Cheloufi S, Elling U, Hopfgartner B *et al.* The histone chaperone CAF-1 safeguards somatic cell identity. *Nature* 528(7581), 218–224 (2015).
7. Ishiuchi T, Enriquez-Gasca R, Mizutani E *et al.* Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat. Struct. Mol. Biol.* 22(9), 662–671 (2015).
8. Yadav T, Quivy JP, Almouzni G. Chromatin plasticity: a versatile landscape that underlies cell fate and identity. *Science* 361(6409), 1332–1336 (2018).
9. Stuart T, Satija R. Integrative single-cell analysis. *Nat. Rev. Genet.* 20(5), 257–272 (2019).
10. Stadhouders R, Vidal E, Serra F *et al.* Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. *Nat. Genet.* 50(2), 238–249 (2018).
11. Gomes AP, Ilter D, Low V *et al.* Dynamic incorporation of histone H3 variants into chromatin is essential for acquisition of aggressive traits and metastatic colonization. *Cancer Cell* 36(4), 402–417.e13 (2019).
12. Alonso-Curbelo D, Ho YJ, Burdziak C *et al.* A gene–environment-induced epigenetic program initiates tumorigenesis. *Nature* 590(7847), 642–648 (2021).
13. Morel D, Jeffery D, Aspeslagh S, Almouzni G, Postel-Vinay S. Combining epigenetic drugs with other therapies for solid tumours — past lessons and future promise. *Nat. Res.* 17(2), 91–107 (2020).
14. Berdasco M, Esteller M. Clinical epigenetics: seizing opportunities for translation. *Nat. Rev. Genet.* 20(2), 109–127 (2019).
15. Stirzaker C, Zotenko E, Song JZ *et al.* Methylome sequencing in triple-negative breast cancer reveals distinct methylation clusters with prognostic value. *Nat. Commun.* 6(1), 1–11 (2015).
16. Grosselin K, Durand A, Marsolier J *et al.* High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. *Nat. Genet.* 51(6), 1060–1066 (2019).
17. Jeffery D, Gatto A, Podsypanina K *et al.* CENP-A overexpression promotes distinct fates in human cells, depending on p53 status. *Commun. Biol.*. 4(1), 1–18 (2021).
18. Clark SJ, Lee HJ, Smallwood SA, Kelsey G, Reik W. Single-cell epigenomics: powerful new methods for understanding gene regulation and cell identity. *Genome Biol.* 17(1), 72 (2016).
19. Darcis G, Van Driessche B, Van Lint C. HIV Latency: Should We Shock or Lock? *Trends Immunol.* 38(3), 217–228 (2017).
20. Ackermann AM, Wang Z, Schug J, Naji A, Kaestner KH. Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. *Mol. Metab.* 5(3), 233–244 (2016).
21. Kamies R, Martinez-Jimenez CP. Advances of single-cell genomics and epigenomics in human disease: where are we now? *Mamm. Genome* 31(5–6), 170–180 (2020).
22. Tritschler S, Theis FJ, Lickert H, Böttcher A. Systematic single-cell analysis provides new insights into heterogeneity and plasticity of the pancreas. *Mol. Metab.* 6(9), 974–990 (2017).
23. Regev A, Teichmann SA, Lander ES *et al.* The human cell atlas. *Elife* 6, e27041 (2017).
24. Rajewsky N, Almouzni G, Gorski SA *et al.* LifeTime and improving European healthcare through cell-based interceptive medicine. *Nature* 587(7834), 377–386 (2020).