Stem Cell Epigenetics: Looking Forward

Epigenetics in Adult SCs

Salvador Aznar Benitah
The Institute for Research in Biomedicine, Barcelona

The integrity of tissues is maintained by adult stem cells during adulthood. However, recent work indicates that tissues often contain more than one population of stem cells that are located at distinct niches and display different behaviors, making the role of stem cells in tissue homeostasis far more complex than previously thought. Generally, each population of stem cells predominantly contributes to its compartment in steady-state conditions and tends to display its multipotency upon damage repair. Interestingly, this heterogeneity might also underlie the different types of tumors that can arise in given tissue. Because the discovery of stem cell heterogeneity is rather recent, I am interested in exploring its significance in health and disease. For instance, which epigenetic factors determine when and how different stem cell populations are established? What mechanisms prevent their intermingling in healthy tissues but allow for their multipotency during repair?

Importantly, the tissue-specific disruption of epigenetic factors often results in tissue aging and carcinogenesis, and several chromatin-remodeling factors are mutated in human cancers. Does the perturbation of epigenetic factors occur independently of other aging mechanisms, or do they converge? How do deregulated epigenetic factors cooperate with oncogenes and tumor suppressors to promote tumorigenesis? I believe that the field of epigenetics will offer great insights toward these unknowns in the future.

### Chromatin Regulator Relocations

Adrian Bracken
Trinity College Dublin

Recent genome-wide approaches to map histone posttranslational modifications (PTMs) and chromatin regulators (CRs) during differentiation and reprogramming have revealed dynamic changes in epigenetic landscapes with striking clarity. However, these descriptive approaches have failed to provide many new mechanistic insights into how CRs are localized, or displaced, so precisely during these transitions.

CRs are not drivers of differentiation or reprogramming. Instead, cell fate decisions are driven by master transcription factors. What we don't know from genome-wide mapping studies is how CRs follow the instruction of master transcription factors. CRs such as the BAF and Polycomb families form multiprotein complexes containing numerous protein domains. These domains function to interact with cofactors, histone PTMs, noncoding RNAs, and DNA without any apparent sequence specificity. I am interested in exploring how these domains combine to stabilize the precise relocalization of CRs during cell fate transitions.

It's exciting to think how new, efficient DNA editing approaches such as CRISPR will allow a surgical dissection of the role of the many domains within protein components of CR complexes, as well as the putative DNA sequences required for their relocalization. These approaches promise to shed new light on how CRs relocate to different target genes during cell fate changes, which will provide a more complete understanding of differentiation and reprogramming.

Obligatory Role of Histone Marks
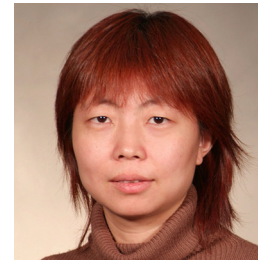
Yali Dou
University of Michigan

The past decade or so has witnessed an explosion of information depicting the epigenetic landscape of stem cells by concerted efforts from the chromatin field. High-resolution genome-wide studies reveal a "hyperactive" chromatin state in stem cells with different sets of histone modifications demarcating distinct regulatory elements such as "poised" enhancers and promoters. These studies also show dynamic, reciprocal changes in histone modifications upon cell fate alteration, illustrating an intriguing correlation between histone marks and stem cell functions. In the next decade, we need to establish whether histone modifications serve an obligatory role, as opposed to a "correlative" one, in stem cell self-renewal and differentiation. The advent of pharmacological inhibitors against specific histone-modifying enzymes allows the assessment of dynamic requirements of histone modifications in these processes, distinguished from genetic studies that always lead to concurrent alterations of protein interaction network. The results should enlighten us about whether the histone modifications underlie "epigenetic memory" per se and/or whether they play a role in setting up the initial ground rules on which downstream effector proteins act. It will also be important to examine whether histone modifications shape the cellular architecture in a way that efficiently engages reprogramming factors or lineage-specific transcription factors and, thus, function as a deterministic factor in these processes.

Chapter 2: Epigenome Editing
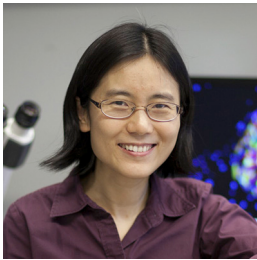
Danwei Huangfu
Sloan-Kettering Institute

Understanding how different cells interpret the same DNA sequence in different ways is central to biomedical research. Epigenetics may help us understand why a muscle cell looks and acts differently from a liver cell, and ultimately this understanding may lead to the generation of specific cell types of therapeutic interest. In recent years there has been growing interest in epigenetics, and tremendous progress has been made. We now have ways to describe the epigenome with ever-increasing precision, and new technologies are being developed to characterize the epigenome at single-base and single-cell resolution. However, it remains challenging to make causal links between changes in the epigenome with the cellular or organismal phenotypes. To address causality, current methods typically use nonspecific approaches, such as blocking an epigenetic modifier through genetic or pharmacological manipulations. Further development of the field will benefit from introducing epigenetic changes to specific genomic loci coupled with precise spatial and temporal control. Genetic studies are already benefiting from new tools such as programmable site-specific nucleases (e.g., CRISPR/Cas) for locus-specific gene editing and optogenetic tools for temporal and spatial manipulation. It is likely that in the next 10 years we will see the development of a series of analogous or perhaps better tools to precisely edit the epigenome, which will enable us to comprehensively decipher its functional relevance.

Going Beyond Self-Renewal

Natalia Ivanova
Yale University

The ability of embryonic stem cells (ESCs) to self-renew and give rise to a full repertoire of differentiated lineages has captivated biologists for decades. While the regulatory network that lies beneath self-renewal is becoming increasingly defined, the molecular circuitry that defines developmental potency of the cells remains largely unknown.

Genome-scale mapping in ESCs and differentiated cells suggests that chromatin holds the key to understanding the pluripotent state. ESCs maintain low levels of DNA methylation, and gains in DNA methylation that occur during differentiation result in lineage-specific gene silencing. Furthermore, regulatory elements of developmental genes are marked by both activating and silencing histone marks, indicating that epigenetic priming may be required for differentiation. Although the enzymatic complexes that regulate DNA methylation and modification of histones have been identified, how these complexes are recruited to specific target sites in ESCs is unclear. Do the core pluripotency factors guide the placement of epigenetic marks? When and how is the epigenetic landscape of the pluripotent state assembled and subsequently taken apart during differentiation?

The field needs the tools for analysis and visualization of chromatin at a single-cell level in culture and in developing embryos to answer these questions. With the knowledge obtained we should be able to safely and efficiently guide pluripotent cells toward specific cell fates for use in medicine and research.

Metabolomics in Stem Cell Aging

Haruhiko Koseki
RIKEN Center for Integrative Medical Sciences

Aging modifies tissue stem cell function, but the underlying mechanisms are largely unknown. Tissue stem cells are often kept in a mitotically dormant state, likely to avoid activation of cellular senescence mechanisms by signals from the niche and to protect against other environmental insults. Despite such protection, tissue stem cells can sense environmental stimuli and may alter their properties to maintain tissue homeostasis. I am interested in the potential impacts of metabolites on epigenetic pathways in stem cells because they may provide a link between the environment and cellular senescence.

One of the main sources of environmental stimuli directed to tissue stem cells may be metabolites that come from foods, commensal microbiota, and various somatic cells. Such metabolites may affect stem cell function by acting as substrates for metabolic pathways, co-enzymes for various regulatory mechanisms, including epigenetic regulatory systems, and ligands for transcription factors. Epigenetic regulators such as Polycomb are known to suppress cellular senescence by robustly repressing critical targets. Epigenetic pathways may also form an interface to buffer against metabolic insults on cellular senescence mechanisms. To validate this model, we will need to overcome structural complexities of metabolites and functional complexities in epigenetic regulation. Systematic studies approaching this problem from both angles and mathematical approaches to unify these fields may be necessary.

CellPress

Epigenetic Variability
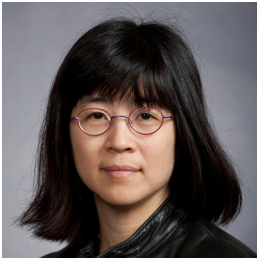
Louise Laurent
University of California, San Diego

Recent studies have discovered significant epigenetic variability among pluripotent stem cell lines, including abnormalities in imprinting and X chromosome inactivation. Aberrant epigenetic regulation is associated with many diseases, including cancer, and thus evidence of epigenetic instability in pluripotent stem cells has raised concerns regarding their potential use in cell transplantation therapy. Certain epigenetic patterns have been associated with specific derivation methods or phenotypic characteristics, including differentiation potential. These results suggest that epigenetic profiling may be a way to assess the quality of a pluripotent stem cell line, in terms of its safety for clinical applications and its utility as an in vitro model for development and disease.

In order for this to be accomplished, several critical areas must be further explored. The mechanistic pathways between epigenetic patterns and cellular phenotype, regulatory interactions among different epigenetic mechanisms, and mechanisms for targeting of epigenetic marks to specific regions of the genome are largely unknown. Perhaps most importantly, we do not yet know how closely the epigenetic marks in pluripotent stem cells and their derivatives reflect those of their cellular counterparts in vivo. I believe that for pluripotent stem cells to reach their full potential as clinical therapeutics, it is critical for us to continue to investigate basic questions in stem cell biology and epigenetics.

lncRNAs and Brain Epigenetics

Daniel A. Lim
University of California, San Francisco

The mammalian genome transcribes many thousands of unique long noncoding RNAs (lncRNAs). Some lncRNAs such as *Xist* have long been known for their role in epigenetics, and more recent studies have revealed additional lncRNAs that interact functionally with chromatin-modifying factors or other epigenetic regulators. lncRNA expression is highly tissue specific. In particular, the expansive set of unique lncRNAs in the brain is even more specific to different brain regions than mRNAs and may also be highly cell-type specific. This exquisite specificity of lncRNAs could be useful for cataloguing the diverse classes of neural cell types, as well as for the molecular diagnosis of neurological diseases.

Emerging evidence also indicates that lncRNAs play key roles in the epigenetics of neural stem cells and brain development. We have obtained only glimpses of the various molecular mechanisms of lncRNAs, and many more novel functions will likely soon be discovered. While epigenetic mechanisms are undoubtedly critical during brain development, it is also clear that some of the molecular mechanisms that underlie the epigenetics of development are also important in postmitotic neurons. For instance, active DNA methylation appears to play a role in memory formation. Given the emerging roles of lncRNAs in neurodevelopmental epigenetics, it will be exciting to learn whether lncRNAs also play key roles in learning, memory, and diseases such as neurodegeneration.

Epigenetics One Stem at a Time
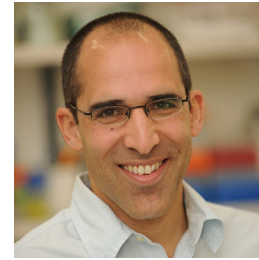
Eran Meshorer
The Hebrew University of Jerusalem

These are exciting times for the field of stem cell epigenetics. Enhanced abilities to program stem cells and reprogram somatic cells combined with the development of "beyond the primary sequence" methods that feed the sequencing monster are driving this field forward at an increasingly rapid pace. Multilevel maps of chromatin hypersensitivity, DNA methylation, chromatin factors, histone modifications, insulators, chromatin folding, and more are now routinely being generated toward a holistic view of the stem cell epigenome. But while epigenomic maps are insightful, they normally provide correlations, not function. Even at the most fundamental level, we cannot always say for certain if, say, an active chromatin mark is the cause, not the effect, of transcription. Advances in genome editing technologies should enable us to take a step back and test many of the predictions generated over the years. A drawback of current whole-genome methods is that they only reveal population averages, so existing epigenomic profiles fail to explain cell-to-cell variation. Why do some cells differentiate or reprogram better than others? Why do some cells become ectoderm while their neighbors become endoderm under similar conditions? The next challenge will be to take the epigenetic field one step further by relating chromatin plasticity to function at a single-cell, genome-wide level. Once this feat is achieved, the holy grail of programming cells a la carte will become ever more feasible.

3D Genome Organization

Ana Pombo
Berlin Institute for Medical Systems Biology

The control of gene expression pathways in the context of tissues and whole organisms is increasingly understood as a complex process, with the role of chromatin folding and nuclear positioning emerging as key factors in gene expression outcomes. Genome sequencing projects had shown that the colinear distribution of genes along the DNA fiber matters for their expression. More recently, microscopy and chromatin conformation capture have shown that the same may be true in 3D space. However, we are still far from grasping what mechanisms underlie the spatial interdependence of gene expression states, how long-range chromatin proximity is established within the 3D nuclear environment, and which mechanisms sense preferred positioning. The picture emerging is one in which local biochemical environments are created within the nuclear landscape by various mechanisms, thereby promoting more effective, robust, or specialized nuclear functions.

A major goal in understanding transcriptional states globally is to identify not only active states of transcription but also the mechanisms that recruit the transcription machinery to silent genes to prime them for expression in response to differentiation or extracellular stimuli. A global and integrated understanding of the interplay between long-range chromatin contacts, chromatin states of protein and/or RNA occupancy and modification, and transcriptional and expression outcomes will be key to unlocking this puzzle.

The Epigenome in Disease

Maike Sander
University of California, San Diego

The advent of genome-wide DNA sequencing technology has led to a dramatic increase in our knowledge of how molecular modifications to DNA and chromatin affect gene expression. Over the last 10 years, we have also identified many of the writers, readers, and erasers that modify the "epigenetic code." Furthermore, we have learned that the epigenome is highly dynamic and undergoes significant change in response to environmental cues and during the aging process.

Now, one of the next frontiers is to explore how the cellular environment shapes the epigenome and thereby influences cellular phenotypes. Such knowledge will likely yield an understanding of the mechanisms underlying common diseases influenced by both environmental and genetic factors, such as diabetes and neurodegenerative diseases. Genome-wide approaches could help define genomic loci that are susceptible to environmental insults, and pluripotent stem cells (PSCs) could provide a model for studying environment-epigenome interactions in a human context. PSC-based approaches should also help shed light on how genetic variation influences susceptibility to environmental stress and predisposes to disease. To enable such studies, we will need to develop improved cell models so that mature disease-relevant cell types can be studied in meaningful ways ex vivo. Exploring the role of the epigenome in disease could reveal epigenetic targets as highly selective entry points for disease intervention and could thus lead to new therapies.

Model Organisms for Epigenetics

Guo-Liang Xu
Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences

The study of cell lineage commitment and conversion will continue to be at the cutting edge. However, the epigenetic mechanisms underlying this process are hard to study in mammals, because obtaining sufficient individual cells locked into a given cellular state can be challenging. The problem is compounded by the fact that commonly used model organisms lack some key epigenetic components. Notably, yeast, worms, and flies all lack DNA methylation, which is vitally important in mammals.

Therefore, there is a need to expand our tool box. Fortunately, numerous candidates across evolution are available. Take the single-celled eukaryote *Naegleria gruberi* as an example. Introduced by Chandler Fulton in the 1960s to study cell differentiation, this free-living organism is famous for its stunning ability to transform from a crawling, bacteria-eating amoeba into a fast-swimming flagellate within 1.5 hr. The amoeba's short doubling time (every 1.6 hr) adds another advantage. The promise of using protozoa is showcased by the recent discovery of the Tet-mediated oxidative demethylation pathway, based on the homology of mammalian Tet proteins to thymidine hydroxylases of trypanosomes. *Naegleria* has the Tet pathway; what role does it play? With new molecular techniques, barriers to using nonstandard models have started to dissipate. Such models can be expected to reveal new and potentially generalizable aspects of epigenetic regulation and can thus be anticipated to contribute to future cell reprogramming research.