Preface

Different types of stem cells—broadly defined as cells with the capacity to maintain their own undifferentiated population as well as to produce one or more mature cell types—play critical roles in embryos, adult tissues, and some types of cancer. Consequently, the term "stem cell biology" might best be considered a thread connecting researchers across a wide spectrum of interests rather than a single, cohesive field of study. Indeed, the functions and regulation of stem cells constitute a key focus of both basic and applied biological sciences, including the fields of developmental biology, regenerative medicine, cancer biology, and many other areas.

Stem cells exhibit a range of properties that differ among the many types of stem cells, including their metabolic profiles, physical characteristics, gene expression profiles, and proliferation rates. In spite of these differences, many types of stem cells share common markers, epigenetic features, and phenotypic properties. Two major questions are how developmental potency is maintained in stem cells and how cells remodel their identities to transform into mature cell types. Considerable progress has been made on each of these topics in recent years, as we have come to understand that numerous signaling pathways and metabolic stimuli impinge on an ever-expanding array of transcriptional and posttranscriptional regulatory mechanisms ranging from sequence-specific transcription factors to epigenetic mechanisms to noncoding RNAs. These regulators collectively control cell type-specific gene expression at the transcriptional and posttranscriptional levels.

The emergence of new single-cell technologies for identifying regulatory elements, mapping transcription factor binding, and quantifying gene expression on a genome-wide scale has enabled a surge of progress toward understanding of stem cell properties and differentiation pathways. As many of these technologies are still quite young, we are likely seeing only the first wave of breakthroughs in this domain. In addition, sensitive new methods for quantifying changes in protein abundance or identity as cells differentiate, as well as metabolite profiles and flux add additional layers to our understanding of stem cells and their progeny. A major challenge for future studies will be to better decipher these dense multifactorial datasets to understand when and how cell fate decisions are made, when they become irreversible, and how these processes can be manipulated for therapeutic purposes.

This volume of *Current Topics in Developmental Biology* entitled "Stem Cell Proliferation and Differentiation" discusses numerous advances in the field of stem cells, as well as some of the current challenges in stem cell biology and areas of focus for future research. Owing to the enormous breadth of the stem cell field, not all topics in this area could reasonably be covered in one volume. Instead, we focus on a few specific areas that have seen exciting new developments or considerable progress in recent years.

1. Chromatin regulation and dynamics

Work in multiple model systems has identified conserved mechanisms of epigenetic regulation and the protein machines responsible for remodeling the epigenome. More recently, an explosion of studies has examined how chromatin structure controls stem cell state and how chromatin is remodeled during differentiation to lock in gene expression programs defining mature cells. In Chapter 1, Klein and Hainer discuss the unique features of chromatin structure in some types of stem cells, the numerous forms of chromatin regulation, and how chromatin structure is remodeled to control cell state.

Notably, while the roles of covalent histone modifications in regulation of stem cell state have been studied extensively, the general mechanisms by which histone modifications increase or decrease target gene expression and shape the gene regulatory network of pluripotent stem cells have only recently come into focus. One such mechanism involves the recruitment of proteins that regulate the placement or positions of histones on DNA. Of particular note, nucleosome remodeling complexes containing members of the SWI/SNF family of ATPases have been shown to exhibit diverse roles in regulation of stem cell identity. In addition, key roles for histone variants such as histones H3.3 and H2AZ in pluripotency have been identified in recent years.

With the discoveries of critical roles of chromatin remodeling complexes and histone variants in pluripotency, along with recent developments in protocols for directed differentiation of pluripotent stem cells, we are now in position to understand where and when in the differentiation process chromatin remodeling factors function and how they help direct specification of different cell types. Although considerable focus has been directed at uncovering regulatory nodes impacting developmental regulatory transcription factors, recent studies have revealed that most chromatin regulatory complexes bind and regulate the expression of hundreds to thousands of

genes in different cell types, suggesting that a focus on individual targets of any one regulatory complex may be of limited utility in understanding how these factors function during development. In addition, recent studies have revealed important noncatalytic roles for numerous chromatin regulatory factors in pluripotency and development. Therefore, a renewed focus on the mechanisms underlying the gene regulatory functions of chromatin regulatory complexes will likely lead to fundamentally new insights.

2. Long noncoding RNAs and RNA-binding proteins

Important noncoding functions of RNAs in cell type specification have been discovered in recent years. Indeed, critical developmental roles for multiple classes of noncoding RNAs, both small (e.g., microRNAs, piRNAs, tRNA fragments) and large (e.g., lncRNAs, enhancer RNAs) have been described in multiple systems. In Chapters 2 and 3, we explore the functions of lncRNAs and RNA-binding proteins (RBPs) in regulation of the cell state in stem and progenitor cells.

Although individual lncRNAs were identified decades ago, the extent to which lncRNAs are expressed, their differential expression patterns in stem cells and differentiation cell types, and their molecular roles in gene expression and development are only now beginning to be understood. Numerous factors have obscured the general and cell type-specific functions of lncRNAs. First, the discovery that transcription is pervasive throughout the genome—which is 98% noncoding in humans—has revealed that mRNAs corresponding to coding genes constitute only a small fraction of the number of unique transcripts within any individual cell, although coding transcripts are often much more abundant than most lncRNAs. Many of these pervasive transcripts are not stable and are therefore present at very low levels, raising the question of what, if any, function they serve in the cell. However, many characterized lncRNAs are expressed from individual promoters, have gene structures that are similar to coding genes (other than the lack of significant coding potential), have relatively high stability, and (as with mRNAs) can be many kilobases in length. Although sequence conservation of lncRNAs is typically very low, their locations vis-à-vis coding genes and predicted secondary structures are often more highly conserved. A second barrier to our understanding of lncRNA functions is the lack of straightforward outputs (such as proteins) for which we have effective tools to assay their functions.

Despite these obstacles, in Chapter 2, Aich and Chakraborty describe substantial progress in our understanding of where and how lncRNAs contribute to developmental potency and cell fate. The expression of a number of lncRNAs expressed in pluripotent stem cells is regulated by pluripotency transcription factors OCT4, SOX2, and NANOG, and several lncRNAs have been shown to perform functions critical for pluripotency and/or cell viability. In addition, some lncRNAs have been shown to help regulate cell fate upon differentiation of pluripotent stem cells. Besides regulation of pluripotency, numerous lncRNAs have been shown to be critical for later stages of embryonic development. For example, multiple studies have identified lncRNAs with key roles in organogenesis, including both heart and brain development.

How do lncRNAs function in regulation of gene expression and cell fate? Several molecular mechanisms have been described by which lncRNAs contribute to gene regulation. In some cases, lncRNAs appear to impact chromatin structure near promoter regions of target genes. Consistent with these findings, a number of chromatin regulatory factors have been shown to interact broadly with coding and noncoding RNAs. However, the degree to which chromatin remodeling complexes are directed to specific loci by individual lncRNAs remains unclear and is an active area of current research. A second mechanism by which lncRNAs control gene expression is through interactions with other regulatory RNAs, such as miRNAs, or in some cases by regulating the enzymatic activities of proteins. Finally, some lncRNAs appear to alter chromatin structure directly through interactions with DNA, either in the form of R-loops (RNA/DNA hybrids) or through triplex structures.

A rapidly progressing area of stem cell and developmental biology centers on the identification and characterization of RBPs that play critical roles in regulation of cell fate. Along with the transcriptional regulatory functions of noncoding RNAs, including lncRNAs discussed above, Li, Kishta, and Wang describe in Chapter 3 an ever-expanding list of posttranscriptional mechanisms controlling the functions of developmental regulators, as well as an expanding network of RBPs that mediate each type of regulation. The development of efficient technologies for crosslinking proteins to RNA and proteomics methods for identification of RNA-binding proteins has uncovered numerous RNA-binding proteins in pluripotent stem cells. Several recently identified RBPs have been shown to play key roles in maintenance of the pluripotent state. While some RBPs control the identities of proteins expressed from RNAs to which they bind (e.g., through alternative

splicing), many others control protein levels through numerous posttranscriptional mechanisms, including control of RNA export, stability, and translation.

A major class of regulators of RNA binding and function are DEAD box helicases of the DDX family. DDX helicases play roles in transcription, RNA maturation, ribosome biogenesis, and translation, among other processes. Given their functions in essential cellular processes, it is perhaps not surprising that DDX family proteins play important roles in stem and progenitor cells. However, the distinct effects of mutation or depletion of various DDX proteins on reprogramming to pluripotency, differentiation, and cell fate may be unexpected. Additional studies are needed to understand how these factors contribute to the regulation of cellular identity.

Finally, an exciting new development in this field is the discovery of widespread covalent RNA modifications, and the finding that these modifications play important roles in stem and progenitor cells. For example, methylation at position six within the purine ring of adenine bases on RNA (m6A) has been shown to be critical for transition from the naïve to the primed state of pluripotent stem cells. While many of the targets of RNA modifications and the proteins performing these modifications are now beginning to be characterized, their precise roles in regulation of cellular state are not completely understood.

The onslaught of studies implicating lncRNAs and other noncoding RNAs in regulation of developmental potency has precipitated a recent explosion in the identification and characterization of RBPs in these processes. Additional efforts will be needed to understand the temporal and tissue specificity of factors regulating lncRNA function, as well as how different lncRNAs mediate specific effects within stem cell gene regulatory networks. In addition, a comprehensive examination of how RNA-binding proteins control the activities of developmental regulators, and how these activities are altered during cell state changes (such as stem cell differentiation) will illuminate this aspect of developmental gene regulation.

3. Distinct modes of pluripotency

A large fraction of pluripotent stem cell research focuses on pluripotent cells most similar to early blastocyst stage embryos, specifically embryonic stem cells and induced pluripotent stem cells. During the early stages of formation of the embryonic body plan, one of the first developmental stages

is formation of the epiblast, wherein the pluripotent cells of the inner cell mass of the blastocyst transition to a distinct form of pluripotent cells in the epiblast stage embryo. This transition has been described as progression from a "naïve" state to a "primed" state. In this process, naïve cells—a collection of spherical cells that grow in three dimensions *in vitro* and *in vivo*, which express the full repertoire of pluripotency transcription factors—morph into cells that grow as a monolayer that expresses a subset of pluripotency transcription factors, as well as a number of genes not expressed in naïve cells. In addition, these primed pluripotent cells exhibit a number of features more commonly found in mature cells.

Although many studies with stem cells involve directed differentiation from naïve pluripotent stem cells to specific differentiation pathways, *in vivo* pluripotent cells transition through the primed state, which can be modeled in culture in what are known as epiblast stem cells (EpiSCs). Studies with EpiSCs have revealed features observed in cells during gastrulation, where the primary germ layers begin to differentiate *in vivo*, and have informed numerous directed differentiation protocols that utilize cultured pluripotent stem cells. To facilitate the use of these cells as models for differentiation studies, in Chapter 4, Samanta and Kalantry describe the isolation and culture of these cells from embryos, along with their characterization in a special methods chapter.

4. Differentiation of pluripotent stem cells

One area of study that has seen enormous growth in the past several years is the development of new methods and protocols to more precisely differentiate pluripotent stem cells into mature cells or lineage-restricted progenitors that can contribute to functional tissues upon transplant into animal models. Robust methods for differentiation of pluripotent stem cells into distinct neuronal lineages, cardiac muscle, pancreatic beta cells, and many other cell types have been described. In Chapter 5, Magaletta, Siller, and Maehr take a deep dive into pluripotent stem cell-derived pharyngeal endoderm—a field that has seen significant progress in recent years in both human and mouse pluripotent stem cell models.

Pharyngeal endoderm makes up critical components of the thymus, parathyroid, and thyroid, and defects in pharyngeal endoderm development contribute to a number of congenital diseases. Using information accumulated about key developmental regulators and signaling pathways necessary

for pharyngeal endoderm development *in vivo*, several groups have developed methods for generation of *in vitro* cell types analogous to those that make up thyroid, parathyroid, and thymus. More detailed characterization of these cells using new genomic and single-cell methods will lead to a more complete understanding of how these tissue types develop. In turn, a better understanding of pharyngeal endoderm differentiation should lead to refinement of differentiation methods to increase their fidelity and robustness, facilitating the use of pluripotent stem cells in therapies.

5. Metabolic control of cell state

Although most efforts in developmental and stem cell biology have focused primarily on the signaling pathways, transcriptional regulators, and (more recently) RNA-based mechanisms that impinge on cell type-specific gene regulatory networks, major roles for metabolites in cell type specification have emerged. Early efforts to understand how different cells differentially employ metabolic pathways used for energy and biomolecule production have focused mainly on tumors. However, as discussed by Ruohola-Baker and colleagues in Chapter 6, it has recently become evident that alterations in metabolism accompany changes in stem cell state, and in some cases help direct cell fate changes.

Several mechanisms govern how metabolic inputs impact cell fate in different types of stem cells. For example, hematopoietic stem cells, which are largely quiescent, utilize anaerobic glycolysis during periods of self-renewal, which is critical for maintenance of their undifferentiated state. In contrast, the transition of rapidly dividing naïve pluripotent stem cells to primed cells is met with a transition from a combined metabolic program of glycolysis and oxidative phosphorylation to one in which glycolysis is used exclusively.

Interestingly, changes in metabolism when cells transition from stem cells to differentiation often impact the epigenome. This is perhaps not surprising when one considers that the cofactors utilized in deposition of multiple epigenetic marks, such as histone methylation, histone acetylation, and DNA methylation, are key metabolites that often reside at the interface of multiple metabolic pathways. Consequently, changes in the concentrations of these metabolites are sometimes met with alterations in the epigenome, leading to changes in gene expression. Undoubtedly, new examples of metabolic regulation of cell fate will be uncovered in coming years as this field continues to progress.

6. Conclusions and perspectives

In recent years, the stem cell community has made tremendous advances in mapping the gene regulatory networks and signaling pathways regulating cell fate, both in pluripotent stem cells and tissue-restricted stem cells. The ability to isolate and characterize stem cell populations *in vitro* and direct their differentiation in cell culture has facilitated these advances. Consequently, one of the ongoing questions is how closely *in vitro* stem cell models reflect the processes of cell type specification that occur *in vivo*. Parallel advances in genomics methodologies have enabled characterization of populations of cells *in vivo* and *in vitro* at the single-cell level, facilitating a more detailed understanding of the similarities and differences among cell populations derived from *in vitro* differentiation of pluripotent stem cells and developing tissues.

Given these increasingly powerful tools for identifying and characterizing the epigenome and gene expression profiles of cells, one major challenge going forward will be to understand how differences between pluripotent stem cell-derived cells and *in vivo* cell populations arise and how they can be eliminated. Two major barriers to the therapeutic use of cells derived from pluripotent stem cell differentiation are incomplete differentiation and differentiated cells that incompletely mimic mature cells *in vivo*. Whether the increasingly detailed maps of the epigenomes and transcriptomes of *in vitro* differentiated cells will lead to better stem cell therapies remains an open question. These concerns aside, stem cell models have uncovered many important new insights into cell type specification that have informed our understanding of mammalian development. There is every indication that stem cell models will continue to represent a powerful conduit for discoveries that enrich our understanding of embryogenesis, as well as models for numerous congenital diseases.

THOMAS G. FAZZIO

Department of Molecular, Cell, and Cancer Biology, University of Massachusetts Medical School, Worcester, MA, United States