Emerging molecular approaches in stem cell biology

Amritha Jaishankar and Kent E. Vrana
Department of Pharmacology, Pennsylvania State University College of Medicine, Hershey, PA, USA

Keywords: stem cells; transcriptomics; epigenomics; proteomics

Stem cells are characterized by their ability to self-renew and differentiate into multiple adult cell types. Although substantial progress has been made over the last decade in understanding stem cell biology, recent technological advances in molecular and systems biology may hold the key to unraveling the mystery behind stem cell self-renewal and plasticity. The most notable of these advances is the ability to generate induced pluripotent cells from somatic cells. In this review, we discuss our current understanding of molecular similarities and differences among various stem cell types. Moreover, we survey the current state of systems biology and forecast future needs and direction in the stem cell field.

Introduction
Stem cells have the remarkable ability to self-renew, as well as to differentiate into multiple cell types in response to extracellular signals. Thus, they are not only valuable research tools to understand cellular reprogramming and human disease, but hold tremendous promise in regenerative medicine. Since the creation of human embryonic stem cells a decade ago, tremendous progress has been made in the field. Although this technology has yet to yield human therapeutics, earlier this year, the world’s first human embryonic stem (ES) cell–based therapy was approved for clinical trial in patients with spinal cord injuries (Geron Corp., Menlo Park, CA, USA). Also, the factors required to create induced pluripotent stem (iPS) cells from terminally differentiated somatic cells have been identified within the past 2 years (1–4). These iPS cells have the potential to revolutionize patient-specific, cell-based therapy. While the generation of iPS cells is a giant leap for stem cell biology, it highlights some of the basic processes that still remain beyond our understanding. Specifically, while the factors required for reprogramming are now known, the mechanism by which these genes control the process and why these genes alone are sufficient for the production of iPS cells remains unknown.

Therefore, to realize the true potential of stem cells, and in order to safely manipulate cellular signaling, there is a need for unprecedented knowledge of the regulation of these cells. Such knowledge will be derived from the integration of gene expression, protein expression, and epigenetics. Specifically, we need to illuminate the networks that regulate pluripotency and self-renewal.

This review examines the current status of stem cell biology with an emphasis on recent progress and emerging molecular technologies that will help address crucial questions as we stand at the threshold of realizing the promise that stem cells hold in the field of regenerative medicine.

Stem cells, pluripotency, and self-renewal
There exist several different types of stem cells, each of which has unique properties and advantages. Comparisons between these stem cells, with varying self-renewal and plasticity, provide insights into the molecular mechanisms controlling these processes.

Adult stem cells are undifferentiated cells, found among differentiated cells within a tissue or organ, which are capable of limited self-renewal and are multipotent (may differentiate to form several types of cells). Hematopoietic stem cells (HSCs) are stem cell progenitors that have contributed to most of our current knowledge about adult stem cells. HSCs are multipotent cells predominantly found in the bone marrow that give rise to the different kinds of blood cells. Owing to vast amounts of research, HSCs have been used in therapeutic transplantation for decades. Bone marrow also harbors mesenchymal stem (MS) cells that are multipotent cells that give rise to chondrocytes, osteocytes, and adipocytes. In addition, adult stem cells are found in most organs of the body, including the brain (neural stem cells), liver, lungs, heart, intestines, skin, and muscles (5). Although adult stem cells cannot be expanded in culture indefinitely, their use lacks the ethical constraints surrounding the use of embryonic stem cells. Moreover, autologous transplantation of these cells (when possible) circumvents the problem of immune rejection, making them a preferred choice for cell-based therapies.

Human umbilical cord blood has been used in transplantation for over two decades now, as it is considered to be a valuable source of cord blood stem cells that comprise both hematopoietic and non-hematopoietic stem cells (6). Cord blood stem cells are of particular interest in cell therapy due to their plasticity, lower immunogenic properties, abundance of progenitors and immature cell types, as well as the ease with which they can be obtained and banked (7). Several disorders (including sickle cell anemia, thalassemia, severe combined immune deficiency, aplastic anemia, Fanconi’s anemia, and glycogen storage diseases) are currently treated with transplantation of cord blood. Also, the ability to generate cells of neural lineage from cord blood stem cells holds tremendous promise for treatment of neurodegenerative diseases (8). Embryonal carcinoma (EC) cells were first derived from human teratocarcinomas in 1977 (9), about two decades after the establishment of mouse EC cells (10). These cells were extensively studied for their potential to differentiate into multiple cell types, but the use of human EC cells is limited by their aneuploid nature. However, this research helped establish culture conditions and provided the framework for derivation of embryonic stem (ES) cells.

ES cells are derived from the inner cell mass of the blastocyst and are characterized by their unlimited self-renewal and pluripotency (ability to give rise to all the cells of the embryo and adult). The derivation of mouse ES cells in 1981 (11,12) and non-human primate ES cells in 1995 (13)
In the nucleus of a eukaryotic cell, the DNA is transcribed into heterogeneous nuclear RNA (hnRNA), which is then converted to the mature messenger RNA (mRNA). The mRNA is then exported into the cytoplasm where protein synthesis occurs on the ribosomes. microRNAs (miRNAs) mediate specific mRNA silencing. Some of the protein that is synthesized, such as transcription factors and chromatin modifiers, return to the nucleus and interact with DNA and mRNA, thereby regulating gene expression. As shown, numerous techniques are currently used to study DNA methylation and histone modifications (epigenetics) as well gene expression (transcriptomics) and protein expression (proteomics).

**Stem cells and nuclear reprogramming**

In order to understand cellular reprogramming involved in creating stem cells (iPS cells, for instance), we need to unravel the sequence of nuclear events occurring during the process. The nucleus is of central importance as it contains nearly all of the genetic information and direct effectors of temporal regulation of gene expression (transcription factors and chromatin remodeling factors). The orchestration of nuclear events in response to numerous signaling pathways culminates in precise regulation of cellular reprogramming, differentiation, or self-renewal.

Recent groundbreaking studies have shown that direct reprogramming of somatic cells can yield iPS cells, which are pluripotent ES cell–like cells. Takahashi et al. first reported the generation of mouse iPS cells from adult murine fibroblasts by ectopic expression of the transcription factors Oct4, Sox2, Klf4, and Myc (2). Subsequently, human iPS cells were also established from human somatic cells by two different four-factor combinations: Oct4, Sox2, Klf4 and Myc or Oct4, Sox2, Nanog, and Lin28. The promise of this technology in regenerative medicine, drug discovery, and basic research is evident. However, all of the reprogramming genes are believed to be associated with cancers (34). The biggest challenge this technology faces, therefore, is to overcome the need to use oncogenes and viral vectors (which can be oncogenic in their own right) to reprogram cells. Rapid progress has been made in addressing these technical challenges posed by iPS cell generation. Myc has already been shown to be non-essential for iPS cell generation (4,35). Recently, it was shown that two factors (Oct4 and Klf4/c-myc) are sufficient for reprogramming adult neural stem cells (36). This raises an interesting question: Are fewer factors required to reprogram cells that already possess greater plasticity? In an independent study, Huangfu et al. showed that primary human fibroblasts could be reprogrammed with a combination of two factors (Oct4 and Sox2) and valproic acid (a histone deacetylase inhibitor) (37). This study is particularly promising because it supports the possibility of effecting reprogramming with just chemical compounds or small molecules. It also illustrates the need to better understand the epigenetic status of these cells. Earlier this year, Kim et al. showed that Oct4 alone is sufficient to generate pluripotent stem cells from adult mouse neural stem cells (38).

The ability to reprogram somatic cells to an embryonic state with defined factors is indeed remarkable, and while this technology opens up novel applications for stem cells in understanding human biology is to understand which genes and factors is indeed remarkable, and while this technology opens up novel applications for stem cells in understanding human biology is to understand which genes and factors account for differences in self-renewal, directed differentiation, and cellular reprogramming in stem cells.
of neurotransmitter biosynthesis and with depression to generate serotonin—to be marvelous surrogate models. Imagine, in addition to thinking of iPS cells as single nucleotide polymorphisms (SNPs) cells also allow us to analyze the effect of therapies. Moreover, patient-specific iPS and ultimately autologous cell-based enables development of targeted drugs to understand human disease and demonstrated (40). Such disease-specific diseases, such as Parkinson’s, Huntington’s, effector genes and proteins.

For instance, it would be useful to not transform, but directly activate our own resident adult stem cells (such as the neural stem cells in the brain). This will require a more detailed knowledge of the nuclear effector genes and proteins.

Recently, iPS cells have been derived from patients with a variety of genetic diseases, such as Parkinson’s, Huntington’s, Down syndrome, and muscular dystrophy (39). Also, differentiation of amyotrophic lateral sclerosis (ALS)—specific iPS cells into motor neurons has recently been demonstrated (40). Such disease-specific iPS cells provide a unique opportunity to understand human disease and enables development of targeted drugs and ultimately autologous cell-based therapies. Moreover, patient-specific iPS cells also allow us to analyze the effect of single nucleotide polymorphisms (SNPs) on disease progression. For example, in addition to thinking of iPS cells as potential therapeutics, they will also prove to be marvelous surrogate models. Imagine, for instance, using iPS cells from a patient with depression to generate serotonin-producing neuronal cells. Characterization of neurotransmitter biosynthesis and release from these cells will capture genetic nuances unique to that patient. Similar approaches can be taken with Alzheimer’s disease or any complex polygenic neurodegenerative disorder. Also, it is now possible to compare adult stem cells and iPS cells derived from the same person (or better yet, derive iPS cells from those adult stem cells), thereby circumventing the issue of genetic background differences seen with the use of ES cells. Such comparisons would help illuminate differences that truly underlie the disparity in pluripotency and self-renewal of these cells.

Stem cells, epigenomes, and integratomes

In addition to the genomic and proteomic studies described above, rapid advancements have been made in assessing genome-wide epigenetic states (epigenomes). Such assessments of global DNA methylation and histone states will be crucial to further our understanding of cellular reprogramming events in iPS cells and to develop drugs to precisely manipulate this process. Moreover, it is increasingly evident that the missing pieces of the stem cell pluripotency puzzle will come from the integration of recent epigenetic data with existing gene expression and proteomic data. Together, these will help decipher the sequence of nuclear events and enable us to construct a regulatory network that holds the key to understanding the molecular mechanism of self-renewal and differentiation in stem cells.

DNA methylation and histone modifications are epigenetic processes that have been extensively studied in ES cells (41). Numerous techniques have been developed to detect DNA methylation, including bisulfite sequencing, methylation specific PCR (MSP), restriction landmark genomic scanning (RLGS) and, more recently, genome-wide methylation profiling using microarray-based approaches (Figure 1). Bibikova et al used a microarray-based method to analyze 1536 CpG sites from 371 genes and showed that human ES cells have a unique DNA methylation profile (42). Another popular high-throughput technique has been to combine methylated DNA immunoprecipitation (MeDIP) with DNA microarrays to map genome-wide methylation (43). In an “integratome” approach, Fouse et al used MeDIP in combination with microarrays to integrate promoter DNA methylation patterns with chromatin modifications (histone H3 lysines 4 and 27 trimethylation) and promoter occupancy of pluripotent regulators (polycomb group proteins and Oct4/Nanog) in regulation of gene expression in mouse ES cells (44). As with DNA methylation, numerous techniques have also been developed to study genome-wide histone modifications, such as chromatin immunoprecipitation coupled with microarrays (ChIP-chip) or with massively parallel short-tag based sequencing (ChIP-seq) (41, 45–47). In addition, ChIP-seq was also recently used to map the binding sites of 13 transcription factors (including Oct4, Nanog, Sox2, Klf4, and myc) as well as two transcription regulators (p300 and Suz12) in mouse ES cells (48). An extension of all these emerging technologies to human ES cells and iPS cells will enable generation of a map of the epigenome. Moreover, understanding how this pattern changes during differentiation of ES cells will enable us to control cellular reprogramming.

Transcriptional regulation is undeniably crucial to stem cell pluripotency. In addition, however, focus has recently shifted to understanding post-transcriptional regulation of gene expression. microRNAs (miRNAs) are small, endogenous, non-coding RNAs transcribed in the nucleus and processed by RNAi machinery in the cytoplasm. miRNAs bind to complementary target mRNA sequences (in RNA-inducing silencing complexes), thereby mediating mRNA silencing through translational repression or mRNA degradation (49). Recent studies have identified a unique subset of miRNAs expressed in ES cells (50–53). The role of miRNA in ES cell differentiation and self-renewal has been suggested by loss/gain of function studies of stem-cell–specific miRNA, as well as of proteins involved in miRNA processing (Dicer and DGCR8) (49). Additionally, one of the reprogramming factors, Lin28, has been shown to specifically block pri-ltet-
7g processing (54). The repression of let-7 expression has also been reported to be essential for self-renewal of breast cancer stem cells, thereby suggesting a role for this miRNA in stem cell self-renewal and differentiation (49,55). Marson et al. recently reported association of key ES cell transcription factors (Oct4, Nanog, Sox2, and Tcf3) with promoters of miRNA that are preferentially expressed in ES cells (56). Moreover, depletion of Oct4 resulted in down-regulation of these miRNAs (56). Thus, further characterization of the role of miRNA in stem cell self-renewal could enable the use of small-molecule inhibitors of miRNA to reprogram somatic cells. It would also be interesting to examine the effect of the remaining iPS genes on ES-specific miRNAs and to determine the targets of these miRNAs.

Recently, efforts to integrate miRNAs with transcriptional regulation were taken one step further when Back and Selbach independently applied a quantitative mass spectrometric approach called SILAC (stable-isotope labeling with amino acids in culture) and pulsed-SILAC, respectively, to investigate the effect of specific miRNAs on human and mouse proteomes (57,58). In parallel, microarrays were also used to quantify miRNA levels in response to miRNA induction or knockdown. Their data show for the first time that individual miRNAs can down-regulate the production of numerous proteins. Such an integrated approach in stem cells is likely to further our understanding of the regulatory circuit.

A network analysis of proteins that are differentially expressed between human ES and MS cells in combination with the transforming IPS cell lines is shown in Figure 2. These proteins were found to be differentially expressed (by a factor of two or more) in a nuclear proteome screen between the two cell types. It is interesting to note that several of the differentially expressed proteins do interact, and more remarkably, that their interactions converge on Myc. The pathway analysis also allows us to examine known miRNA interactions that occur in this regulatory network. Such analyses represent powerful bioinformatic tools that can be used to integrate existing genomic, proteomic, and epigenomic data to predict crucial regulators of gene expression in stem cells.

In conclusion, while enormous strides have been made in understanding stem cell biology, large-scale systems biology approaches will be required to understand how molecules interact in a complex, dynamic system such as within the human ES cell. Emerging technologies to obtain information from single molecules or single cells will help overcome the need for large quantities of starting materials and help circumvent issues with sample heterogeneity. They will also be crucial in achieving a more integrated approach to studying stem cell biology. Moreover, advances in vivo molecular imaging techniques will help visualize the dynamic changes that occur in a stem cell undergoing self-renewal or directed differentiation.

Acknowledgments

This work was supported by the Pennsylvania State College of Medicine and the Elliot S. Vesell endowment. The authors have no competing interests.

References

12. Ross, R., C. Kumar, J. Cox, H. Scholer, and J.M. Mann. 2008. Stable isotope labeling by amino acids in cell culture (SILAC) and proteome quanti...


Received: 13 February 2009; accepted: 20 February 2009.

Address correspondence to Kent E. Vrana, Penn State College of Medicine, Department of Pharmacology, 500 University Dr., R130 Hershey, PA 17033, USA. email: kvrana@psu.edu

Index to Advertisers

454 Life Sciences | 353
ACCURI CYTOMETERS | 366
AFLYMATRIX | 337
AGOWA GmbH | 365
Applied Biosystems | Cover 4
Applied Biosystems, Part of Life Technologies | 332
Biometra part of GE Healthcare | 324*
Bio-Rad Laboratories | Gatefold
BioTrove | 361
EPICENTRE Biotechnologies | 323
Erixon | 338

Febit Holding GmbH | 349
Finzymes | 375-376
ForteBIO | 347*
Fujifilm Life Sciences | 336*
Gibson | 326*
Integrated DNA Technologies | 362
Lonza | 377
Lonza Cologne AG | 330
MD Biosciences | 357
Millopol | 345
NuGen Technologies, Inc. | 328
PromoCell GmbH | 331

Protea Biosciences Inc | 372-374
QIAGEN Inc | 325
Roche NimbleGen | 335
Semrock | 359
TA Instruments | 343
Thermo Fisher Scientific/Erie | 341*
ThermoScientific/Genomics Group | 322
*North American Edition
**European/Rest of World Edition