Review Article
doi: 10.5115/acb.2010.43.1.1
pISSN 2093-3665 eISSN 2093-3673
Anatomy & Cell Biology acb

Genomics and proteomics in stem cell research: the road ahead

Sung-Min Ahn, Richard Simpson[1], Bonghee Lee[2]
LCDI-BRC Joint Genome Center, Lee Gil Ya Cancer and Diabetes Institute, Gachon University of Medicine and Science, Incheon, Korea, [1]Joint Proteomics Laboratory, Ludwig Institute for Cancer Research & The Walter and Eliza Hall Institute of Medical Research, University of Melbourne, Melbourne, Australia, [2]Center for Genomics and Proteomics & Stem Cell Core Facility, Lee Gil Ya Cancer and Diabetes Institute, Gachon University of Medicine and Science, Incheon, Korea

Abstract: Stem cell research has been widely studied over the last few years and has attracted increasing attention from researchers in all fields of medicine due to its potential to treat many previously incurable diseases by replacing damaged cells or tissues. As illustrated by hematopoietic stem research, understanding stem cell differentiation at molecular levels is essential for both basic research and for clinical applications of stem cells. Although multiple integrative analyses, such as genomics, epigenomics, transcriptomics and proteomics, are required to understand stem cell biology, proteomics has a unique position in stem cell research. For example, several major breakthroughs in HSC research were due to the identification of proteins such as colony-stimulating factors (CSFs) and cell-surface CD molecules. In 2007, the Human Proteome Organization (HUPO) and the International Society for Stem Cell Research (ISSCR) launched the joint Proteome Biology of Stem Cells Initiative. A systematic proteomics approach to understanding stem cell differentiation will shed new light on stem cell biology and accelerate clinical applications of stem cells.

Key words: Stem cells, genomics, epigenomics, transcriptomics, proteomics
Received February 1, 2010; Revised March 3, 2010; Accepted March 4, 2010

Stem cells: definition and clinical relevance

Currently, stem cells are widely and intensively studied in all fields of medicine and science due to their potential to treat many incurable diseases. At the same time, there have been a wide range of social and ethical concerns on using human embryonic stem cells (ES cells) for research or therapy purposes because their derivation requires the destruction of embryos.

Stem cells are functionally defined as undifferentiated, primitive cells that retain the capability of indefinitely reproducing themselves ("self-renewal") and also have the ability to generate multiple types of cells upon proper signals from internal and external cues ("pluripotency" or "multipotency"). Stem cells are largely categorized into two groups, ES cells and adult stem cells. Although these two types of stem cells have different sets of advantages and disadvantages, they both may be valuable sources for the future cell therapy.

ES cells have many advantages as a cell source for regenerative therapy. Given the appropriate environmental conditions, they can be coaxed into forming most cell types in the body and can provide unlimited amounts of cells for cell therapy. However, their applications for human treatment require special safety precautions. One concern is the potential risk of genomic abnormality (Draper *et al.*, 2004; Maitra *et al.*, 2005) and tumor formation (Vogel 2005). Other concerns associated with ES cell-mediated therapy are immune rejection of transplanted tissue, difficulty in achieving homogeneous cell populations after differentiation, and ethical controversy resulting from the destruction of embryos. Many scientists believe that there still exists a great deal of social and scientific uncertainty surrounding ES cell-mediated cell therapy, which requires more extensive and

Corresponding author:
Bonghee Lee
Address: 7-45 Songdo-dong, Yeonsu-gu, LCDI, Incheon, Korea [406-840]
Tel: +82-32-899-6582, Fax: +82-32-899-6519, E-mail: bhlee@gachon.ac.kr

thorough study.

In contrast, adult stem cells have been used for cell therapy for quite a long time. For example, bone marrow transplant has been used to treat leukemia more than 10 years. Thus far, more than 70 different diseases and injuries have been successfully treated with adult stem cells. These include a number of blood-related abnormalities, immunological dysfunctions, cancers, stroke, brain trauma, juvenile diabetes, Parkinson's disease, blindness, spinal cord injuries and so on. More than 1200 clinical trials with adult stem cells have been approved by the U.S. Food and Drug Administration (FDA). There are several reasons that account for the wide use of adult stem cells in clinical applications. These include very little chance of tumor formation, no ethical problems, the possibility of autologous transplantation, which causes no immune rejection, and easy differentiation into a certain specific lineage of cells.

However, the prevailing belief that adult stem cells do not form tumors needs thorough reevaluation. There has been a recent report that adipose tissue-derived human mesenchymal stromal cell (MSC) populations can immortalize and transform spontaneously. This MSC-TMC (Transformed Mesenchymal cells) transition happened after in vitro culture and expansion (Rubio *et al.,* 2005). Furthermore, several other reports argue that some cancer cells are derived from adult stem cells in the body (Marx 2003; Pardal *et al.,* 2003; Clarke & Fuller, 2006). Currently, safety issues and techniques for in vitro expansion of the cells may be the two most important topics to develop additional successful cell therapies using adult stem cells.

Future studies should be focused on overcoming three major obstacles that hinder successful stem cell therapy. First, obtaining autologous cell sources is required to avoid immune rejection. This issue is especially serious for ES cell-mediated cell therapy. One way to get around the immune problem is to generate pluripotent, ESC-like cells from the patient's own cells either by cellular reprogramming or by Somatic Cell Nuclear Transfer (SCNT). Recently, there have been great achievements in generating induced pluripotent stem cells (iPS cells) (Takahashi *et al.,* 2007; Yu *et al.,* 2007), although more work is needed to bring iPS cells to patient treatment. At the same time, the recent success of SCNT in rhesus macaque primates (Byrne *et al.,* 2007) raises a hope that it may soon also be possible in humans. Establishing a stem cell therapeutic bank (Taylor *et al.,* 2005) and autologous cell therapy using adult stem cells would be other options to solve the immune rejection problem.

The second issue is preventing tumor formation. This issue is especially serious for clinical use of ES cells. Since tumor formation is thought to be caused by residual undifferentiated ES cells, the development of methods that remove these remnant ES cells is of great importance.

The third issue is establishing efficient protocols to differentiate ES cells into the cell types of interest. This will increase the efficacy of cell therapies, reduce side effects caused by the presence of unwanted cell types, and minimize tumor formation resulting from remnant undifferentiated ES cells.

Multiple aspects of understanding stem cell differentiation

Each cell type in the human body, including stem cells, has a unique information architecture maintained by cell-specific transcriptional regulatory states that lead to differential expression of genes (Davidson 2006). According to the International Human Genome Consortium, the number of genes in the human genome is estimated to be between 20,000 and 25,000. The 274 different kinds of cells in the human body are defined by combinatorial expression of this limited number of genes (Wenick & Hobert, 2004). The transcriptional regulatory circuitry in cells is maintained by the dynamic interplay of multiple regulatory mechanisms, including epigenetic, transcriptional, posttranscriptional, and posttranslational regulation of gene expression (Cheng *et al.,* 2005). Therefore, understanding stem cell differentiation requires an overall understanding of the elaborate web of regulations at the levels of the genome, epigenome, transcriptome, microRNAome, and proteome.

Transcription factors (TFs) are good examples of how multiple regulatory mechanisms act in concert. TFs are core elements of the transcriptional regulatory circuitry (Boyer *et al.,* 2005). In theory, it is possible to turn a brain cell into a liver cell by transferring all TFs in a liver cell to a brain cell while inactivating all brain-specific TFs (Reik 2007). A recent study that shows induction of pluripotent stem cells from fibroblasts by defined TFs also supports this idea (Takahashi & Yamanaka, 2006). Given their importance in cell specification and function, it is not surprising that multiple regulatory mechanisms are involved in the regulation of TFs. For example, when embryonic stem cells are induced to

differentiate, Oct-3/4, a main regulator of human embryonic stem cell pluripotency and self-renewal (Boyer *et al.,* 2005), is epigenetically repressed through histone modifications, followed by DNA methylation at the promoter region (Feldman *et al.,* 2006). Once transcribed, mRNAs of TFs may be posttranscriptionally regulated by microRNAs (Bartel 2004). According to a recent computational analysis, many TFs are target hub genes of microRNAs, and TF-microRNA pairs may coregulate large sets of common targets (Shalgi *et al.,* 2007). Posttranslational modifications (PTMs) provide another layer of regulation. It is well known that TFs, such as STATs and Smad 2/3, require PTMs for their nuclear localization and regulatory activities (O'Shea *et al.,* 2002; Elliott & Blobe, 2005).

In this section, we aim to provide a concise review of multiple aspects of understanding and studying stem cell-differentiation: genomics, epigenomics, transcriptomics, microRNAomics, and proteomics. In the next section, we will specifically discuss the potential roles of proteomics in the field of stem cell biology.

Regulatory genome

The human genome encodes 20,000~25,000 protein-coding genes that account for only 1.5% of the nearly three billion base pairs of the genome (Stein 2004). In addition to protein-coding regions, a substantial portion of the non-protein-coding regions in the genome, which are poorly defined, is transcribed (The EPC 2004). Non-coding RNAs (i.e. transcripts of non-protein-coding genes) include microRNAs, small regulatory RNAs, and tens of thousands of longer transcripts with unknown functions (Mattick & Makunin, 2006).

In addition to coding sequences, the genome contains regulatory architecture called cis-regulatory modules (CRMs) that consist of a set of TF-binding sites (TFBSs) located in the vicinity of the gene being regulated (Blanchette *et al.,* 2006). As previously described, differentiation requires the institution of new transcriptional regulatory states. The CRMs, which are "information processing devices hardwired into the genome," read the given transcriptional regulatory states, process that information, and instruct the basal transcription apparatus to turn the gene on or off (Davidson 2006). Therefore, the human genome is a self-executing set of instructions that not only encodes proteins and non-coding RNAs, but also directs where, when, and how much each of these products will be expressed, ultimately producing the complex body plan of the adult organism (Howard & Davidson, 2004). One of the most widely and intensively studied examples of the regulatory genome in stem cells is the regulatory network of Oct4, Sox2, Nanog and Klf4 in embryonic stem cells (Nichols *et al.,* 1998; Avilion *et al.,* 2003; Chambers *et al.,* 2007). Activation of these genes is key to maintaining pluripotency in embryonic stem cells. Another group of genes is also reported as related to neural stem cells and their progenitor cells (Wang *et al.,* 2008). These examples of the regulatory genome have been a focus of stem cell research because of their importance in maintaining pluripotency. Scientists are still trying to identify novel regulatory genome components that are involved in maintaining self renewal of stem cells or lead to differentiation.

Epigenetic aspects of stem cell differentiation

Genetic information is stored in the DNA sequence using a four digit code (i.e., A, G, T, C), and the genome is the complete set of information in an organism's DNA (Alberts 2002). Epigenetic information is heritable information that is not encoded in the DNA sequence but still affects gene expression (Russo *et al.,* 1996). In eukaryotes, the DNA molecule is wrapped around a core of eight histones (two H2A, H2B, H3, and H4 histones) to form nucleosomes, which, in turn, are packaged into chromatin (Luger *et al.,* 1997). Epigenetic information is stored by covalent modifications of the DNA or its packaging histones. Therefore, an epigenome can be defined as the sum of heritable covalent modifications of the chromatin, 'the dynamic template of the genetic information'. The human genome is organized into accessible (euchromatic) or inaccessible (heterochromatic) subdomains by epigenetic changes (Arney & Fisher, 2004) (Fig. 1).

"Development is, by definition, epigenetic" (Reik 2007). Epigenetic mechanisms provide both stable and flexible gene regulation required for development. Histone modifications confer short-term, flexible regulation while DNA methylation confers long-term, stable silencing (Reik 2007). Since one genome can generate many epigenomes, the dynamic genetic templates created by the combination of a genome and a variety of epigenomes provide the basis of cell-type-specific, unique gene expression profiles.

DNA methylation: DNA methylation in mammalian cells refers to covalent modification of DNA by methylation of the cytosine, which occurs predominantly in CpG dinucleotides. DNA methylation has been implicated in a wide range of physiological and pathological processes such as genomic

doi: 10.5115/acb.2010.43.1.1

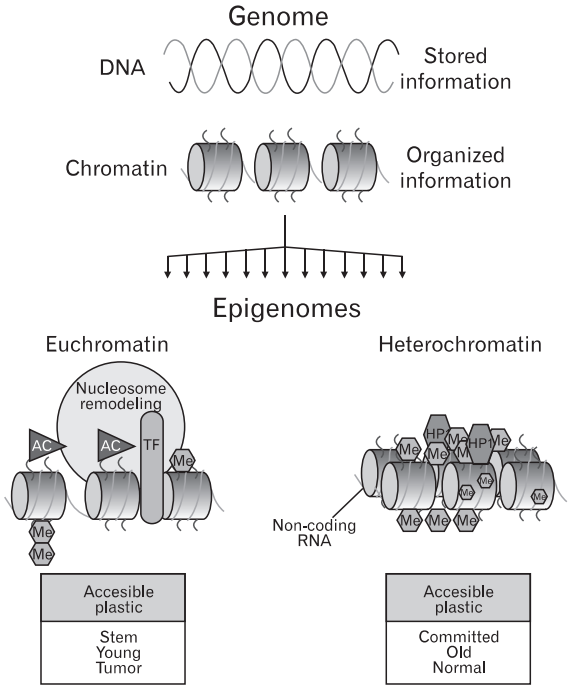

Fig. 1. One genome, many epigenomes. Every somatic cell in the human body has almost the same genetic information. Epigenetic changes define accessibility to the genetic information, thus creating cell-type-specific, dynamic genetic templates. In general, stems cells have more open, accessible chromatin structures than differentiated cells.

imprinting, X chromosome inactivation, tissue-specific gene expression, cellular differentiation, carcinogenesis and aging (Jones & Takai, 2001; Bird 2002). Recent findings in animal cloning indicate that epigenetic reprogramming is important for successful cloning (Kang *et al.,* 2001), which in turn support the idea that DNA methylation provides a specific cellular memory function in development (Bird 2002).

DNA methylation can affect gene transcription through two main mechanisms. Firstly, site-specific DNA methylation may interfere with the binding to DNA of transcription factors (TFs) that activate gene transcription (Watt & Molloy, 1988). Secondly, DNA methylation attracts methyl-CpG-binding domain (MBD) proteins that in turn recruit corepressor complexes involved in transcriptional silencing. Comparative sequence analysis using the sequence of MeCP2, the prototypical MBD protein, led to the identification of additional family members (MBD1, MBD2, MBD3, and MBD4) (Wade 2001). Interestingly, each MBD protein seems to have exclusive binding sites, which suggests that MBD proteins are not functionally redundant (Klose *et al.,* 2005).

As previously mentioned, DNA methylation is involved in the long term silencing of gene expression, which plays an important role in the regulation of pluripotency-associated genes (Reik 2007). For example, transcriptional repression of Oct-3/4 is first followed by histone H3 methylation on Lys 9, which leads to local heterochromatization and DNA methylation at the promoter region (Feldman *et al.,* 2006). The importance of this mechanism is manifested in cancers. According to one recent study, Oct-3/4 is expressed in adult human stem cells, tumor cells and cell lines, but not in differentiated cells (Tai *et al.,* 2005). This evidence is consistent with the cancer stem cell hypothesis, which argues that tumors arise from small populations of cancer stem cells that originate from the transformation of normal stem cells (Pardal *et al.,* 2003). According to the epigenetic progenitor origin of cancer theory, a modified version of the cancer stem cell hypothesis, the first step to carcinogenesis is epigenetic disruption of stem cells or progenitor cells, which is demonstrated as changes in DNA methylation (Feinberg *et al.,* 2006). Thus, understanding stem cell differentiation from an epigenetic perspective may provide new insights on carcinogenesis. Recently, detailed analysis of DNA methylation in the maintenance of pluripotency has become a popular area of study to better understand the basic mechanisms underlying this process in ES cells as well as in iPSCs (Bhutani *et al.,* 2009).

Histone code: As previously described, histones are not merely DNA packaging molecules but also important carriers of epigenetic information. Epigenetic information is stored in histones by chemical modifications of the N-terminal tails of core histones, including acetylation, phosphorylation, methylation, and ubiquitination.

Histone modifications can affect the chromatin template and therefore transcription through three main mechanisms: Cis-effects, trans-effects, and histone replacements. Firstly, histone modifications can alter intranucleosomal interactions by changing the physical properties of modified histone tails, such as electrostatic charges (cis-effects). These intranucleosomal structural changes can affect the accessibility of the DNA contained on that nucleosome to transcriptional factors. For example, phosphorylation of the linker histone H1, which reduces the electrostatic binding of H1 to DNA in chromatin, can repress transcription (Dou & Gorovsky, 2000). Secondly, histone modifications are recognized (or read) by modification-specific binding partners (trans-effects), which mediate down-stream chromatin-modulating events. For example, methylation of lysine 9 on histone H3 is read by the

chromodomain of heterochromatin protein 1 (HP1), which is involved in both maintenance and somatic inheritance of transcriptionally silent heterochromatin (Bannister *et al.,* 2001). Lastly, histone modifications can recruit ATP-dependent chromatin remodelers that shuffle histone variants into and out of chromatin. The replacement of histone variants plays important roles in transcriptional regulation. For example, the replacement of H3 with the histone variant H3.3 is associated with active chromatin and coupled to transcription (Ahmad & Henikoff, 2002; Schwartz & Ahmad, 2005).

The "histone code" hypothesis has evolved from the recognition of the combinatorial nature of histone modifications that can generate synergistic or antagonistic interaction affinities for chromatin-associated proteins (Jenuwein & Allis, 2001). According to this hypothesis, histone modifications can act sequentially, in that a given modification of a specific histone residue is required for subsequent modifications of the same histone or another histone molecule (Margueron *et al.,* 2005). For example, ubiquitination of histone H2B on lysine 123 is required for the methylation of histone H3 on lysine 4, which leads to transcriptional repression (Dover *et al.,* 2002). The histone code hypothesis also implies that multiple signaling pathways converge on chromatin through histone modifications, the combination or summation of which are 'read' by chromatin-associated proteins (Nightingale *et al.,* 2006). The presence of local binary switches (e.g., the 'methyl/phos switch' on H3 Lys 9/Ser 10) and modification cassettes (i.e. the combination of modifications in short clusters with distinct biological readouts) support this hypothesis (Fischle *et al.,* 2003). One of the most studied modifications of the histones is acetylation. There are a few reports describing histone acetylation as important in human embryonic stem cell pluripotency (Lee *et al.,* 2004, Bernstein *et al.,* 2006). This epigenomic control has been a focus of stem cell research for many years, with DNA methylation and miRNA studies of transcriptional regulation in maintaining self-renewal and regulating differentiation.

Transcriptomic approach to understand stem cell differentiation

Microarray technology has revolutionized the field of transcriptomics, enabling biologists to perform global gene expression analyses (Chang *et al.,* 2006). After the isolation of various stem cells, including ESCs, NSCs, and HSCs, one of the first questions addressed by microarray technology was identifying whether there are intrinsic molecular programs shared by all stem cells that define "stemness." The search for stemness is based on the hypothesis that stem cells have to respond in similar ways to regulate self-renewal and differentiation, irrespective of their lineal origin (Cai *et al.,* 2004).

The first studies comparing the transcriptomic profiles of stem cells raised both enthusiasm and confusion. Ramalho-Santos et al. (Ramalho-Santos *et al.,* 2002) and Ivanova et al. (Ivanova *et al.,* 2002) compared transcriptional profiles of ESCs, HSCs, and NSCs, each identifying a list of genes that are likely to reveal core stem cell properties. Interestingly, the number of overlapping genes between the two studies was minimal (only 15 genes), and most of the stemness genes from both groups were not exclusively expressed in NSCs. These findings raised two fundamental questions: 1) whether stemness can be defined at the genetic level (Fortunel *et al.,* 2003), and 2) whether a specific combination of genes rather than individual genes endows stemness (Burns & Zon, 2002). These studies not only provided a first glimpse of stemness networks, but also revealed critical issues in the transcriptomic profiling approach to defining stemness.

microRNA

Basic concepts: Non-coding RNAs or nonprotein-coding RNAs control various levels of gene expression in physiology and development, including chromatin architecture/epigenetic memory, transcription, and RNA splicing, editing, translation and turnover (Mattick & Makunin, 2006). MicroRNAs (miRNAs) are a class of non-coding RNAs that are about 22 nucleotides in length and play important roles in the regulation of target genes (Bartel 2004). According to current estimates, miRNAs account for ~1% of predicted genes in higher eukaryotic genomes and regulate up to 10~30% of genes (Cui *et al.,* 2006).

miRNAs regulate gene expression by either of two post-transcriptional mechanisms, depending on sequence complementarity; miRNA induces cleavage of the mRNA (100% complementarity) or represses translation by binding to the 3' UTR (less complementary) (Pasquinelli *et al.,* 2005). In plants, miRNAs base pair with target mRNAs by precise or nearly precise complementarity, leading to direct cleavage and destruction of the targets. In contrast, most animal miRNAs are less complementary to their target mRNAs and inhibit

doi: 10.5115/acb.2010.43.1.1

protein synthesis while preserving the stability of the target mRNAs. According to the combinatorial rheostat model of gene expression micromanagement by miRNAs, especially in animals, the combinatorial binding of miRNAs with different effective concentrations in different cell types leads to resistance changes (i.e., resistance to translation), thereby fine-tuning the level of protein that is produced (Bartel & Chen, 2004).

The imprecise base pairing between the typical animal miRNA and a target mRNA suggests that animal miRNAs may regulate a wide repertoire of different mRNAs. This hypothesis was supported by a recent study using microarray analysis, in which transfection of tissue-specific miRNAs downregulated a large number of target mRNAs (e.g., miR-124 downregulated 174 genes) (Lim *et al.,* 2005). Interestingly, there are some apparent trends in the types of genes regulated by miRNAs. For example, in *Drosophila,* the target genes of miR-277 are enriched for functionally related enzymes involved in the catabolism of leucine, isoleucine and valine, suggesting that miR-277 acts as a 'metabolic switch' of this biochemical pathway (Stark *et al.,* 2003). Certain miRNAs (e.g., miR-34a-c) are important components of signaling pathways, such as the p53 network (He *et al.,* 2007). miRNAs are also known to be involved in the regulation of various biological processes, including apoptosis (Xu *et al.,* 2004), animal development (Wienholds & Plasterk, 2005), physiological function (e.g., insulin secretion) (Poy *et al.,* 2004), and differentiation (Chen *et al.,* 2004).

miRNA and stem cell differentiation: From the perspective of stem cell biology, miRNAs are attractive regulatory molecules, as they provide an efficient means for coordinating the action of many functionally related target genes (Cheng *et al.,* 2005). The highly tissue-specific expression of most miRNAs also suggests that miRNAs play important roles in differentiation and maintenance of cell identity (Wienholds *et al.,* 2005). As previously described, TFs are core elements of the transcriptional regulatory circuitry that determines cell specifications (Davidson 2006). Therefore, the fact that many TFs are target hub genes of miRNAs underscores the roles of miRNAs in the transcriptional regulatory circuitry and cell specifications. miRNAs involved in differentiation not only target TFs such as Nanog, Smad1, and c-Myb (Luzi *et al.,* 2007; Xiao *et al.,* 2007), but also form elaborate regulatory loops with TFs, which provides a mechanism for determining stable end states. For example, Johnston et al. (Johnston *et al.,* 2005) showed that two miRNAs and their TF targets act in a double-negative feedback loop to control neuronal cell fate decision, and miRNAs function as developmental switches in such regulatory loops. Yoo and Greenwald (Yoo & Greenwald, 2005) also showed a similar feedback loop in which a miRNA works as a developmental switch for mutually exclusive expression of two cell-fate-determining genes. In this model, the protein encoded by the first gene switches on a miRNA, which in turn switches off expression of the second gene (Plasterk 2006).

Recently, several reports has been published regarding novel miRNAs and their function in both maintaining pluripotency and self renewal (Sanosaka *et al.,* 2009; Liu & Zhao, 2009, Judson *et al.,* 2009).

Stem cell proteomics: the road ahead

Why stem cell proteomics?

Major breakthroughs in HSC research were made by the identification of proteins such as colony-stimulating factors (CSFs) and cell-surface CD molecules. Proteins, key players in the cell, have diverse features that are not predictable from gene sequences or transcript levels. For example, posttranslational modifications (PTMs), protein-protein interactions, and subcellular locations affect the function and activity of proteins, but these are not predictable using genomics or transcriptomics technology.

Both basic and clinically oriented stem cell research are confronted with many open questions that can be most efficiently answered by proteomics. For instance, the cell surface proteins and signaling cascades of stem cells and their differentiated progenies are largely unknown, as are the differentiation-specific proteins that can be used as biomarkers of the intermediate or terminal steps of cell differentiation, or discriminate tumorigenic cells from the pool (Krijgsveld *et al.,* 2008).

According to Metcalf, one of the pioneers of CSF research, the most formidable technical challenge in identifying CSFs was their purification (Metcalf 1991). Given the current status of proteomics technology, a project like that of CSF purification and identification can now be completed in six months rather than years or decades. The biggest challenge we now face is the systematic and optimized use of proteomics technology to decipher stem cell biology.

Stem cell proteomics: the road ahead

Finding CDs: the membrane proteome and beyond

Membrane proteins perform most of the specific functions of the plasma membrane (Alberts 2002). According to bioinformatic analyses, 20~30% of all open reading frames in the genome are predicted to encode membrane proteins (Wallin & von Heijne, 1998). Membrane proteins represent more than two-thirds of drug targets (Hopkins & Groom, 2002), and they are essential to understand the underlying mechanisms of various biological functions. For example, our understanding of immune function is based on recognition of only 10~20% of leukocyte membrane proteins (Zola & Swart, 2003). In stem cell research, membrane proteins are important not only for understanding stem cell biology, but also for their clinical application.

The first membrane proteomics initiative: the Human Leukocyte Differentiation Antigens Workshop (HLDA), arguably the first systematic membrane proteomics consortium, was initiated in 1982, before the word 'proteomics' was even coined in 1995. HLDA aimed to bring order to the chaos caused by the generation of large numbers of monoclonal antibodies reactive against leukocyte cell-surface molecules, each with unknown molecular targets and different associated nomenclatures (Zola *et al.,* 2005). Through the organized efforts of HLDA, many cell-surface molecules with important functions in leukocyte biology were identified and characterized under the CD ("cluster of differentiation") system, many of which are currently used widely as diagnostic reagents (e.g., CD3, CD4), and as the basis for therapeutic agents (e.g., CD3, CD20) (Zola & Swart, 2003). In 2005, the HLDA council changed the name of the organization to Human Cell Differentiation Molecules (HCDM) to reflect their broader objectives: to extend the focus from leukocytes to other cell types and to broaden the organization's scope from cell-surface molecules to any molecule whose expression reflects differentiation (Zola *et al.,* 2005). Therefore, CD molecules now include proteins involved in differentiation as well as membrane proteins.

After the development of protein sequencing technology, such as Edman degradation (Edman 1960) and mass spectrometry (Hunt *et al.,* 1986), monoclonal antibodies no longer represent the primary tool for discovering new proteins, but combining knowledge of the human genome sequence, powerful bioinformatics, and cloning technologies allows for new opportunities in antibody proteomics. In addition, CD antibody microarrays, used for high-throughput screening of CD antigens in cells and tissues, provides a good example of how the conventional CD approach can be used in conjunction with high-throughput proteomics technology (Woolfson *et al.,* 2006).

Antibody proteomics: Antibody proteomics, one of the major initiatives of the Human Proteome Organization (HUPO), is defined as the systematic generation and use of protein-specific antibodies to functionally explore the proteome (Uhlen & Ponten, 2005). Agaton et al. (Agaton *et al.,* 2003) showed that Protein Epitope Signatures Tags (PrESTs), unique epitopes present in native proteins, can be used as antigens and affinity ligands for cost-effective generation of highly selective, mono-specific antibodies or affinity-purified polyclonal antibodies. Currently, this approach has been extensively used to construct a comprehensive, antibody-based protein atlas of expression and localization profiles in human normal and cancer tissues (Uhlen *et al.,* 2005). As monoclonal antibodies have been systematically validated for leukocyte biology through HLDA (currently HCDM), the resources developed by the antibody proteomics initiative can be systematically used for stem cell research. For that purpose, well-defined panels of stem cells and their derivatives are required. Once prepared, these panels of cells can be efficiently screened using high-throughput cell microarray (tissue microarray) methods. This approach allows for detection of novel CD molecules involved in the differentiation of a variety of stem cells.

MS-based membrane proteomics: As previously mentioned, the antibody-based approach has both success and promise in identifying novel proteins, especially membrane proteins. However, the antibody-based approach has intrinsic limitations because it is an indirect method for characterizing proteins, and there can be cross-reactivity in antibody-antigen interactions (Bentley *et al.,* 1994). Also, the antibody-based approach cannot provide much information about posttranslational modifications that may be important for protein function.

In contrast, a mass spectrometry (MS)-based approach allows direct sequencing of peptides and proteins, and it can be used either as a method complementing the antibody-based approach or for direct identification and characterization of CD molecules. Firstly, MS can be used to identify molecular targets of antibodies (e.g. identification of immunoprecipitated proteins). Furthermore, even epitope mapping of antibodies is possible using MS (Yu *et al.,* 1998). Secondly, a MS-based approach, coupled with membrane

doi: 10.5115/acb.2010.43.1.1

protein preparation and separation technologies, can be used to directly identify and characterize membrane proteins from cells and tissues (Wu *et al.,* 2003; Wu & Yates, 2003). In 2005, AOHUPO (Asia Oceania Human Proteome Organization) launched the Membrane Proteomics Initiative (MPI), which aims to develop methods for characterizing membrane proteomes, and to characterize the proteomics of specific membrane systems. Once phase I (technology development) and phase II (large-scale analysis of membrane proteomes) are finished, the methods and experiences from MPI can be readily utilized for more subtle areas, such as stem cell proteomics.

Posttranslational modifications (PTMs): The high complexity of the human proteome, consisting of ∼300,000∼3,000,000 distinct protein forms, results from the diversification of genetic information at both the mRNA level and the protein level (Walsh 2006). Genetic information stored in ∼30,000 genes is diversified through alternative splicing of primary mRNA transcripts at the mRNA level (Black 2003). Once translated, proteins are diversified through post-translational covalent changes termed PTMs (Krishna & Wold, 1993), and many important regulatory steps depend on proteins' PTMs rather than on their expression levels (Levchenko 2005).

PTMs play important roles in stem cell biology. Elliott et al. (Elliott *et al.,* 2004) showed that ∼21% of the proteins identified in murine R1 ES cells had PTMs, and several of them, including Ras-GTPase activating protein binding protein 1 and phosphoglycerate kinase, had not been previously associated with PTMs. Unwin et al. (Unwin *et al.,* 2006), revealed the importance of posttranslational control as a regulatory factor in primary hematopoietic stem cells. From the proteomics perspective, assaying changes in PTMs during stem cell differentiation in a high-throughput fashion is a key to understanding the underlying mechanisms of stemness and differentiation.

The determination of PTMs is one of the main challenges in proteomics research. However, recent developments in affinity-based enrichment and extraction methods, multidimensional separation technologies, and mass spectrometry (MS) now allow systematic investigation of PTMs. For example, histone PTMs have been extensively characterized by MS-based proteomics approaches (Pesavento *et al.,* 2004; Thomas *et al.,* 2006). Furthermore, emerging top-down MS technology captures not only individual PTMs, but also their combinations present on the same protein molecule (Garcia *et al.,* 2007; Taverna *et al.,* 2007). Proteomics also allows for quantitative analysis of subproteomes with a specific PTM (e.g. the phosphoproteome (Reinders & Sickmann, 2005) and the methylproteome (Ong *et al.,* 2004). For example, Kratchmarova et al. (Kratchmarova *et al.,* 2005) used MS-based approaches to quantitatively monitor tyrosine-phosphorylated proteins upon the differentiation of mesenchymal stem cells into bone-forming cells in response to growth factors. Recently, Van Hoof et al. (2009) successfully reported phospho-protein profiling during early differentiation of human embryonic stem cells, which enables the identification of key protein activations in specific signaling pathways of the early differentiation process. This paper provides strong evidence of how PTM-based proteomics can contribute to understanding the mechanism of adaptor signaling at the stem cell level. This technique will likely be widely applied to mechanism-based research focusing on self-renewal and differentiation processes of stem cells.

Protein-protein interactions (PPI): A single protein may have diverse functions that are determined in the context of the network or functional module it belongs to (Sharan *et al.,* 2007). The importance of PPI is well illustrated in the famous Myc-Max-Mad network in which different combinations of TFs lead to different functional outcomes (Baudino & Cleveland, 2001). With the help of proteomics and bioinformatics, a comprehensive map of protein interactions in human was recently created, which describes the sum of all potential protein interactions in humans (Gandhi *et al.,* 2006). Since each of the ∼274 different cell types in the human body has different protein expression profiles and thus different protein networks, understanding the unique protein network architecture of each cell type is a key to understanding stem cell differentiation.

As previously described, the transcriptional regulatory circuitry, in which TFs are core elements, determine cell specification and function. Therefore, it is important to understand protein interactions of transcription factors in order to understand the transcriptional regulatory circuitry. Kim et al. (Kim *et al.,* 2006) showed that a small protein interaction network starting from Nanog, a central TF in ES cells, might serve as a functional module for maintaining ES cell pluripotency. The diverse roles of TFs in both stem cells and differentiated cells also indicate that their corresponding protein networks might differ. For example, in neural crest stem cells, Sox10, a high-mobility-group TF, maintains

pluripotency and inhibits neuronal differentiation (Kim *et al.,* 2003). On the other hand, Sox10 is involved in the terminal differentiation of myelin-forming oligodendrocytes (Stolt *et al.,* 2002). The fact that Sox10 interacts with many TFs, forming diverse protein networks, may partly explain its contradictory roles in stem cells and differentiated cells (Wissmuller *et al.,* 2006).

Finding stem cell growth factors: secretome As mentioned earlier, understanding differentiation is important both for basic research and clinical applications. In HSC research, the first major technical advance towards understanding hemopoiesis was the introduction of solid-state cultures of bone marrow and spleen cells (Bradley & Metcalf, 1966; Ichikawa *et al.,* 1966). Then, the discovery of hemopoietic growth factors or colony-stimulating factors (CSFs) led to molecular level understanding of hemopoietic differentiation, followed by clinical applications of CSFs (Metcalf 1991). The extensive purification required to achieve complete purification of each CSF illustrates the importance and usefulness of proteomics technology in discovering growth factors present in minute amounts. In the most difficult case, enrichment of one million-fold was required for purification and characterization of a CSF (Metcalf 1991).

As illustrated in HSC research, finding growth factors in the secretome is a critical step towards understanding and controlling stem cell-differentiation at the molecular level. Recently, a proteomics approach revealed the endodermal secretome network guiding stem cell cardiopoiesis from ES cells (Arrell *et al.,* 2008). Evidence from normal and tumor cells also illustrate the linkage between signal transduction pathways and the secretome (Jacobs *et al.,* 2006; Khwaja *et al.,* 2006).

Most CSFs were purified from media conditioned by various human tumor cells, and it is noteworthy that tumors secreting CSFs include not only hemopoietic cancers such as a T-lymphoblast cell line (Gasson *et al.,* 1984), but also a human squamous cell line from a lower oral cavity (Nomura *et al.,* 1986) and a gastric cancer cell line (Baba *et al.,* 1995), which are not related to hemopoiesis. Therefore, systematic screening of stem and cancer cell media, combined with extensive proteomic purification and characterization, may be required for high-throughput characterization of the growth factor repertoire of each stem cell line. Recently, Kang et al. (Kang *et al.,* 2008) reported proteomic differences of various sized exosomes from human neural stem cells and its differentiated oligodendrocyte progeny. In the study, the authors separated exosomes with Field-Flow Fractionation by 10 nm size differences, and showed that this step was useful for reproducibly analyzing the proteomics changes. These promising results suggest that well-designed sampling strategy and analysis technologies should be further developed to achieve better yields in secretome-based proteomics.

Perspectives

In 2007, the HUPO and ISSCR joint initiative, Proteome Biology of Stem Cells, was established as a collaborative platform bringing together stem cell biologists and researchers in proteomics (Heck *et al.,* 2007; Krijgsveld *et al.,* 2008). The aim of the initiative is to effectuate the implementation of cutting edge proteomic technology in stem cell research to further our understanding of stem cell biology. This has been prompted primarily by major breakthroughs in stem cell biology, the potential of stem cells for biomedical application, and the awareness that proteomics may be able to accelerate this progress further and possibly open yet unexplored areas of research. Over the last 2 years, acting groups of the initiative have started workshops to optimize protocols for hESC sampling, MS analysis, and bioinformatics analyses. The initiative has successfully chosen hESC cells for initial study from the ES cell bank of the International Embryonic Stem Cell Consortium. Analysis of hESC membrane proteins will be started soon to standardize biomarker discovery, and this will help provide standard guidelines for stem cell proteomics. The voyage to stem cell proteomics has already begun.

Acknowledgements

This research was supported by a grant from the Korean Government (MOEHRD) KRF-2006-352-D001714 and SC-2110.

References

Agaton C, Galli J, Höidén Guthenberg I, et al. (2003). Affinity proteomics for systematic protein profiling of chromosome 21 gene products in human tissues. Mol Cell Proteomics 2: 405-414

Ahmad K, Henikoff S. (2002). The histone variant H3.3 marks

 doi: 10.5115/acb.2010.43.1.1

active chromatin by replication-independent nucleosome assembly. Molecular cell 9: 1191-1200

Alberts B. (2002). Molecular biology of the cell, 4th edn (New York, Garland Science)

Arney KL, Fisher AG. (2004). Epigenetic aspects of differentiation. J Cell Sci 117: 4355-4363

Arrell DK, Niederlander NJ, Faustino RS, Behfar A, Terzic A. (2008). Cardioinductive network guiding stem cell differentiation revealed by proteomic cartography of tumor necrosis factor alpha-primed endodermal secretome. Stem Cells 26: 387-400

Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev 17: 126-140

Baba M, Hasegawa H, Nakayabu M, et al. (1995). Establishment and characteristics of a gastric cancer cell line (HuGC-OOHIRA) producing high levels of G-CSF, GM-CSF, and IL-6: the presence of autocrine growth control by G-CSF. Am J Hematol 49: 207-215

Bannister AJ, Zegerman P, Partridge JF, et al. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature 410: 120-124

Bartel DP. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116: 281-297

Bartel DP, Chen CZ. (2004). Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. Nat Rev Genet 5: 396-400

Baudino TA, Cleveland JL. (2001). The Max network gone mad. Mol Cell Biol 21: 691-702

Bentley GA, Boulot G, Chitarra V. (1994). Cross-reactivity in antibody-antigen interactions. Res Immunol 145: 45-48

Bernstein BE, Mikkelsen TS, Xie X, et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125: 315-326

Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM. (2010). Reprogramming towards pluripotency requires AID-dependent DNA demethylation. Nature 463: 1042-1047

Bird A. (2002). DNA methylation patterns and epigenetic memory. Genes Development 16: 6-21

Black DL. (2003). Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem 72: 291-336

Blanchette M, Bataille AR, Chen X, et al. (2006). Genome-wide computational prediction of transcriptional regulatory modules reveals new insights into human gene expression. Genome Res 16; 656-668

Boyer LA, Lee TI, Cole MF, et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122: 947-956

Bradley TR, Metcalf D. (1966). The growth of mouse bone marrow cells in vitro. Aust J Exp Biol Med Sci 44: 287-299

Burns CE, Zon LI. (2002). Portrait of a stem cell. Dev Cell 3: 612-613

Byrne JA, Pedersen DA, Clepper LL, et al. (2007) Producing primate embryonic stem cells by somatic cell nuclear transfer. Nature 450: 497-502

Cai J, Weiss ML, Rao MS. (2004). In search of "stemness". Exp Hematol 32: 585-598

Chambers I, Silva J, Colby D, et al. (2007). Nanog safeguards pluripotency and mediates germline development. Nature 450: 1230-1234

Chang HY, Thomson JA, Chen X. (2006) Microarray analysis of stem cells and differentiation. Methods Enzymol 420: 225-254

Chen CZ, Li L, Lodish HF, Bartel DP. (2004). MicroRNAs modulate hematopoietic lineage differentiation. Science 303: 83-86

Cheng LC, Tavazoie M, Doetsch F. (2005). Stem cells: from epigenetics to microRNAs. Neuron 46: 363-367

Clarke MF, Fuller M. (2006). Stem cells and cancer: two faces of eve. Cell 124: 1111-1115

Cui Q, Yu Z, Purisima EO, Wang E. (2006). Principles of microRNA regulation of a human cellular signaling network. Mol Syst Biol 2: 46

Davidson EH. (2006). The regulatory genome: gene regulatory networks in development and evolution, New edn (Oxford Boston, Elsevier / Academic Press)

Dou Y, Gorovsky MA. (2000). Phosphorylation of linker histone H1 regulates gene expression in vivo by creating a charge patch. Molecular Cell 6: 225-231

Dover J, Schneider J, Tawiah-Boateng MA, et al. (2002). Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. J Biol Chem 277: 28368-28371

Draper JS, Smith K, Gokhale P, et al. (2004). Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. Nat Biotechnol 22: 53-54

Edman P. (1960). Phenylthiohydantoins in protein analysis. Ann N Y Acad Sci 88: 602-610

Elliott RL, Blobe GC. (2005). Role of transforming growth factor Beta in human cancer. J Clin Oncol 23: 2078-2093

Elliott ST, Crider DG, Garnham CP, Boheler KR, Van Eyk JE. (2004). Two-dimensional gel electrophoresis database of murine R1 embryonic stem cells. Proteomics 4: 3813-3832

Feinberg AP, Ohlsson R, Henikoff S. (2006). The epigenetic progenitor origin of human cancer. Nat Rev Genet 7: 21-33

Feldman N, Gerson A, Fang J, et al. (2006). G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nat Cell Biol 8: 188-194

Fischle W, Wang Y, Allis CD. (2003) Binary switches and modification cassettes in histone biology and beyond. Nature 425: 475-479

Fortunel NO, Otu HH, Ng HH, et al. (2003). Comment on " 'Stemness': transcriptional profiling of embryonic and adult stem cells" and "a stem cell molecular signature". Science 302: 393

Gandhi TK, Zhong J, Mathivanan S, et al. (2006). Analysis of the human protein interactome and comparison with yeast, worm and fly interaction datasets. Nat Genet 38: 285-293

Garcia BA, Pesavento JJ, Mizzen CA, Kelleher NL. (2007). Pervasive combinatorial modification of histone H3 in human cells. Nat Methods 4: 487-489

Gasson JC, Weisbart RH, Kaufman SE, et al. (1984). Purified human granulocyte-macrophage colony-stimulating factor: direct action on neutrophils. Science 226: 1339-1342

He L, He X, Lim LP, et al. (2007). A microRNA component of the p53 tumour suppressor network. Nature 447: 1130-1134

Heck AJ, Mummery C, Whetton AD, et al. (2007). Proteome biology of stem cells. Stem Cell Res 1: 7-8

Hopkins AL, Groom CR. (2002). The druggable genome. Nat Rev Drug Discov 1: 727-730

Howard ML, Davidson EH. (2004). cis-Regulatory control circuits in development. Dev Biol 271: 109-118

Hunt DF, Yates JR 3rd, Shabanowitz J, Winston S, Hauer CR. (1986). Protein sequencing by tandem mass spectrometry. Proc Natl Acad Sci U S A 83: 6233-6237

Ivanova NB, Dimos JT, Schaniel C, Hackney JA, Moore KA, Lemischka IR. (2002). A stem cell molecular signature. Science 298: 601-604

Jacobs JM, Waters KM, Kathmann LE, et al. (2008). The mammary epithelial cell secretome and its regulation by signal transduction pathways. J Proteome Res 7: 558-569

Jenuwein T, Allis CD. (2001). Translating the histone code. Science 293: 1074-1080

Johnston RJ Jr, Chang S, Etchberger JF, Ortiz CO, Hobert O. (2005). MicroRNAs acting in a double-negative feedback loop to control a neuronal cell fate decision. Proc Natl Acad Sci U S A 102: 12449-12454

Jones PA, Takai D. (2001). The role of DNA methylation in mammalian epigenetics. Science 293: 1068-1070

Judson RL, Babiarz JE, Venere M, Blelloch R. (2009). Embryonic stem cell-specific microRNAs promote induced pluripotency. Nat Biotechnol 27: 459-461

Kang DJ, Oh SO, Ahn SM, Lee BH, Moon MH. (2008). Proteomic analysis of exosomes from human neural stem cells by flow field-flow fractionation and nanoflow liquid chromatography-tandem mass spectrometry. J Proteome Res 7: 3475-3480

Kang YK, Koo DB, Park JS, et al. (2001). Aberrant methylation of donor genome in cloned bovine embryos. Nat Genet 28: 173-177

Khwaja FW, Svoboda P, Reed M, Pohl J, Pyrzynska B, Van Meir EG. (2006). Proteomic identification of the wt-p53-regulated tumor cell secretome. Oncogene 25: 7650-7661

Kim H, Hahn M, Grabowski P, et al. (2006). The Bacillus subtilis spore coat protein interaction network. Mol Microbiol 59: 487-502,

Kim J, Lo L, Dormand E, Anderson DJ. (2003). SOX10 maintains multipotency and inhibits neuronal differentiation of neural crest stem cells. Neuron 38: 17-31

Klose RJ, Sarraf SA, Schmiedeberg L, McDermott SM, Stancheva I, Bird AP. (2005). DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. Mol Cell 19: 667-678

Kratchmarova I, Blagoev B, Haack-Sorensen M, Kassem M, Mann M. (2005). Mechanism of divergent growth factor effects in mesenchymal stem cell differentiation. Science 308: 1472-1477

Krijgsveld J, Whetton AD, Lee BH, et al. (2008). Proteome biology of stem cells: a new joint HUPO and ISSCR initiative. Mol Cell Proteomics 7: 204-205

Krishna RG, Wold F. (1993). Post-translational modification of proteins. Adv Enzymol Relat Areas Mol Biol 67: 265-298

Lee JH, Hart SRL, Skalnik DG. (2004). Histone deacetylase activity is required for embryonic stem cell differentiation. Genesis 38: 32-38

Levchenko A. (2005). Proteomics takes stem cell analyses to another level. Nat Biotechnol 23: 828-830

 doi: 10.5115/acb.2010.43.1.1

Lim LP, Lau NC, Garrett-Engele P, et al. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature 433: 769-773

Liu C, Zhao X. (2009). MicroRNAs in adult and embryonic neurogenesis. Neuromolecular Med 11: 141-152

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389: 251-260

Luzi E, Marini F, Carbonell SS, Tognarini I, Galli G, Brandi ML. (2008). Osteogenic differentiation of human adipose tissue-derived stem cells is modulated by the miR-26a targeting the SMAD1 transcription factor. J Bone Miner Res 23: 287-295

Maitra A, Arking DE, Shivapurkar N, et al. (2005). Genomic alterations in cultured human embryonic stem cells. Nat Genet 37: 1099-1103

Margueron R, Trojer P, Reinberg D. (2005). The key to development: interpreting the histone code? Curr Opin Genet Dev 15: 163-176

Marx J. (2003). Cancer research. Mutant stem cells may seed cancer. Science 301: 1308-1310

Mattick JS, Makunin IV. (2006). Non-coding RNA. Hum Mol Genet 15 Spec No 1: R17-29

Metcalf D. (1991). The Florey Lecture, 1991. The colony-stimulating factors: discovery to clinical use. Philos Trans R Soc Lond B Biol Sci 333: 147-173

Nichols J, Zevnik B, Anastassiadis K, et al. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell 95: 379-391

Nightingale KP, O'Neill LP, Turner BM. (2006). Histone modifications: signalling receptors and potential elements of a heritable epigenetic code. Curr Opin Genet Dev 16: 125-136

Nomura H, Imazeki I, Oheda M, et al. (1986). Purification and characterization of human granulocyte colony-stimulating factor (G-CSF). EMBO J 5: 871-876

Ong SE, Mittler G, Mann M. (2004). Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. Nat Methods 1: 119-126

O'Shea JJ, Gadina M, Schreiber RD. (2002). Cytokine signaling in 2002: new surprises in the Jak/Stat pathway. Cell 109(Suppl): S121-131

Pardal R, Clarke MF, Morrison SJ. (2003). Applying the principles of stem-cell biology to cancer. Nat Rev Cancer 3: 895-902

Pasquinelli AE, Hunter S, Bracht J. (2005). MicroRNAs: a developing story. Curr Opin Genet Dev 15: 200-205

Pesavento JJ, Kim YB, Taylor GK, Kelleher NL. (2004). Shotgun annotation of histone modifications: a new approach for streamlined characterization of proteins by top down mass spectrometry. J Am Chem Soc 126: 3386-3387

Plasterk RH. (2006). Micro RNAs in animal development. Cell 124: 877-881

Poy MN, Eliasson L, Krutzfeldt J, et al. (2004). A pancreatic islet-specific microRNA regulates insulin secretion. Nature 432: 226-230

Ramalho-Santos M, Yoon S, Matsuzaki Y, Mulligan RC, Melton DA. (2002). "Stemness": transcriptional profiling of embryonic and adult stem cells. Science 298: 597-600

Reik W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. Nature 447: 425-432

Reinders J, Sickmann A. (2005). State-of-the-art in phosphoproteomics. Proteomics 5: 4052-4061

Rubio D, Garcia-Castro J, Martín MC, et al. (2005). Spontaneous human adult stem cell transformation. Cancer Res 65: 3035-3039

Russo VEA, Martienssen RA, Riggs AD. (1996). Epigenetic mechanisms of gene regulation (Plainview, N.Y., Cold Spring Harbor Laboratory Press)

Sanosaka T, Namihira M, Nakashima K. (2009). Epigenetic mechanisms in sequential differentiation of neural stem cells. Epigenetics 4: 89-92

Schwartz BE, Ahmad K. (2005). Transcriptional activation triggers deposition and removal of the histone variant H3.3. Genes Dev 19: 804-814

Shalgi R, Lieber D, Oren M, Pilpel Y. (2007). Global and Local Architecture of the Mammalian microRNA-Transcription Factor Regulatory Network. PLoS Comput Biol 3: e131

Sharan R, Ulitsky I, Shamir R. (2007). Network-based prediction of protein function. Mol Syst Biol 3: 88

Stark A, Brennecke J, Russell RB , Cohen SM. (2003). Identification of Drosophila MicroRNA targets. PLoS Biol 1: E60

Stein LD. (2004). Human genome: end of the beginning. Nature 431: 915-916

Stolt CC, Rehberg S, Ader M, et al. (2002). Terminal differentiation of myelin-forming oligodendrocytes depends on the transcription factor Sox10. Genes Dev 16: 165-170

Tai MH, Chang CC, Kiupel M, Webster JD, Olson LK,

Trosko JE. (2005). Oct4 expression in adult human stem cells: evidence in support of the stem cell theory of carcinogenesis. Carcinogenesis 26: 495-502

Takahashi K, Tanabe K, Ohnuki M, et al. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131: 861-872

Takahashi K, Yamanaka S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126: 663-676

Taverna SD, Ueberheide BM, Liu Y, et al. (2007). Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. Proc Natl Acad Sci U S A 104: 2086-2091

Tay YM, Tam WL, Ang YS, et al. (2008). MicroRNA-134 modulates the differentiation of mouse embryonic stem cells where it causes post-transcriptional attenuation of nanog and LRH1. Stem Cells 26: 17-29

Taylor CJ, Bolton EM, Pocock S, Sharples LD, Pedersen RA, Bradley JA. (2005). Banking on human embryonic stem cells: estimating the number of donor cell lines needed for HLA matching. Lancet 366: 2019-2025

The ENCODE Project Consortium. (2004). The ENCODE (ENCyclopedia Of DNA Elements) Project. Science 306: 636-640

Thomas CE, Kelleher NL, Mizzen CA. (2006). Mass spectrometric characterization of human histone H3: a bird's eye view. J Proteome Res 5: 240-247

Uhlen M, Bjorling E, Agaton C, et al. (2005). A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol Cell Proteomics 4: 1920-1932

Uhlen M, Ponten F. (2005). Antibody-based proteomics for human tissue profiling. Mol Cell Proteomics 4: 384-393

Unwin RD, Smith DL, Blinco D, et al. (2006). Quantitative proteomics reveals posttranslational control as a regulatory factor in primary hematopoietic stem cells. Blood 107: 4687-4694

Van Hoof D, Muñoz J, Braam SR, et al. (2009). Phosphorylation dynamics during early differentiation of human embryonic stem cells. Cell Stem Cell 5: 214-226

Vogel G. (2005). Cell biology. Ready or not? Human ES cells head toward the clinic. Science 308: 1534-1538

Wade PA. (2001). Methyl CpG binding proteins: coupling chromatin architecture to gene regulation. Oncogene 20: 3166-3173

Wallin E, von Heijne G. (1998). Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. Protein Sci 7: 1029-1038

Walsh C. (2006). Posttranslational modification of proteins: expanding nature's inventory (Englewood, Colo., Roberts and Co. Publishers)

Wang ZX, Teh CH, Chan CM, et al. (2008). The transcription factor Zfp281 controls embryonic stem cell pluripotency by direct activation and repression of target genes Stem Cells 26: 2791-2799

Watt F, Molloy PL. (1988). Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. Genes Dev 2: 1136-1143

Wenick AS, Hobert O. (2004). Genomic cis-regulatory architecture and trans-acting regulators of a single interneuron-specific gene battery in C. elegans. Dev Cell 6: 757-770

Wienholds E, Kloosterman WP, Miska E, et al. (2005). MicroRNA expression in zebrafish embryonic development. Science 309: 310-311

Wienholds E, Plasterk RH. (2005). MicroRNA function in animal development. FEBS Lett 579: 5911-5922

Wissmuller S, Kosian T, Wolf M, Finzsch M, Wegner M. (2006). The high-mobility-group domain of Sox proteins interacts with DNA-binding domains of many transcription factors. Nucleic Acids Res 34: 1735-1744

Woolfson A, Ellmark P, Chrisp JS, A Scott M, Christopherson RI. (2006). The application of CD antigen proteomics to pharmacogenomics. Pharmacogenomics 7: 759-771

Wu CC, MacCoss MJ, Howell KE , Yates JR 3rd. (2003). A method for the comprehensive proteomic analysis of membrane proteins. Nat Biotechnol 21: 532-538

Wu CC, Yates JR 3rd. (2003). The application of mass spectrometry to membrane proteomics. Nat Biotechnol 21: 262-267.

Xiao C, Calado DP, Galler G, et al. (2007). MiR-150 controls B cell differentiation by targeting the transcription factor c-Myb. Cell 131: 146-159

Xu P, Guo M, Hay BA. (2004). MicroRNAs and the regulation of cell death. Trends Genet 20: 617-624

Yoo AS, Greenwald I. (2005). LIN-12/Notch activation leads to microRNA-mediated down-regulation of Vav in C. elegans. Science 310: 1330-1333

Yu J, Vodyanik MA, Smuga-Otto K, et al. (2007). Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. Science 318: 1917-1920

Yu L, Gaskell SJ, Brookman JL. (1998). Epitope mapping

doi: 10.5115/acb.2010.43.1.1

of monoclonal antibodies by mass spectrometry: identification of protein antigens in complex biological systems. J Am Soc Mass Spectrom 9: 208-215

Zola H, Swart B, Nicholson I, et al. (2005). CD molecules 2005: human cell differentiation molecules. Blood 106: 3123-3126

Zola H, Swart BW. (2003). Human leucocyte differentiation antigens. Trends Immunol 24: 353-354