BOOKMARKING THE GENOME: MAINTENANCE OF EPIGENETIC INFORMATION

Sayyed K. Zaidi1, Daniel W. Young1,2, Martin Montecino3, Andre J. van Wijnen1, Janet L. Stein1, Jane B. Lian1, Gary S. Stein1,4

1Department of Cell Biology and Cancer Center, University of Massachusetts Medical School and Cancer Center, 55 Lake Avenue North, Worcester, MA 01655 USA; 2Centro de Investigaciones Biomédicas, Facultad de Ciencias Biológicas y Facultad de Medicina, Universidad Andrés Bello, Santiago, Chile; 3Present Address: Wolf Greenfield, Boston, MA 02210-2206; 4Corresponding author: Gary S. Stein, PhD, Department of Cell Biology, University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655, Tel: (508) 856-5625, Fax: (508) 856-6800, Email: gary.stein@umassmed.edu

Abstract
Mitotic inheritance of gene function is obligatory to sustain biological control. Emerging evidence suggests that epigenetic mechanisms are linked to transmission of cell fate, lineage commitment, and maintenance of cellular phenotype in progeny cells. Mechanisms of epigenetic memory include gene silencing by DNA methylation, transcriptional regulation by histone modifications, regulation of gene expression by non-coding small RNA molecules, and retention of regulatory machinery on target gene loci for activation and repression. We will focus on regulatory implications of epigenetic memory for physiological control as well as for the onset and progression of disease.

Introduction
Cell division requires functional changes that include transcriptional silencing by reorganization of regulatory machinery within the nucleus and structural alterations by remodeling of cellular architecture (1,2). Selective mitotic retention of regulatory signatures is necessary for continuity of function in progeny cells. For example, genes expressed immediately after cell division must be epigenetically marked and poised for transcription. Similarly, genes required exclusively during development and silenced postnatal, or genes selectively expressed in specific lineages and permanently repressed in others must epigenetically acquire functional memory, and retain or relinquish transcriptional status. Equally important, persistence of cell transformation and tumor progression are dependent on continuity of cancer-related gene expression.

Historically, nuclear transplantation studies established retention of competency for pluripotency and epigenetic plasticity (3) (Figure 1). Genetic control has been shown to be essential but insufficient to accommodate requirements for gene expression (4). Contributions of epigenetic control are evident and new dimensions are emerging. DNA methylation and histone modifications play key roles in epigenetically transmitting transcriptional memory (5-11). Non-coding small RNA molecules and mitotic retention of gene regulatory machinery are major epigenetic components controlling gene transcription in parent and progeny cells (12-18). These epigenetic mechanisms can be synergistic, antagonistic, or mutually exclusive in regulating the distribution of regulatory information during cell division, development and differentiation.

We will present an overview of mechanisms operative in cellular epigenetic memory–cohorts of combinatorial and cumulative epigenetic marks that are conveyed through generations for cell fate and lineage commitment. DNA methylation and histone modifications in regulating the interphase gene expression are covered in detail in this Mini Review Series (also see Figure 2). We will focus on mitotic
inheritance of non-DNA encoded regulatory parameters in symmetric and asymmetrical cell divisions.

**DNA methylation: A strategy for transcriptional silencing**

DNA methylation is a well-established epigenetic mechanism for both transient and irreversible transcriptional silencing (9-11). Three DNA methyltransferases, designated DNMT1, DMNT3a, and DNMT3b are responsible for dynamic transcriptional silencing by addition of methyl groups to CpG islands in gene promoters (19,20). Consequently, compromised binding of transcription factors to gene promoters and/or altered nucleosomal occupancy within the promoter contribute to biological control (21).

DNA methylation permits silenced states of genes to be inherited through multiple cellular divisions (9). Gene imprinting and allelic exclusion are examples of developmental gene silencing through successive mitoses. An allele-specific imprinting control region (ICR) within the promoter regions of certain genes regulates their mono-allelic expression. A differentially methylated domain (DMD) present within the ICR is epigenetically altered by DNA methylation (and histone modifications), which is perpetuated through cell divisions to maintain the silenced state of the gene in progeny cells.

An informative example is the RAS protein-specific guanine nucleotide-releasing factor 1 (Rasgrf1) gene, the product of which is an important component of the Ras pathway (22-24). The ICR region of the locus contains a repeated element that is essential and sufficient for the establishment and maintenance of DNA methylation at the DMD for germ-line imprinting of the paternal allele. The CCCTC-binding factor (CTCF), a transcriptional repressor, occupies the DMD on the un-methylated maternal allele and prevents interaction of the enhancer with the promoter, thus silencing the maternal allele (25). CTCF binding to the paternal DMD is prevented by DNA methylation, allowing specific expression of the paternal allele. Thus, DNA methylation plays a critical role in gene imprinting and allelic exclusion at the Rasgrf1 gene locus during development and ensures transmission of ‘transcriptional memory’ through cell divisions.

Although the role of DNA methylation in inheritable silencing of transcriptional states is well understood (9), some key questions remain to be addressed. How are DNA methyltransferases targeted to gene promoters? Recent reports show that H3K9 and H3K27 histone modifications may play a role in directed DNA methylation. It is yet to be resolved whether these histone modifications are a pre-requisite for targeting DNMTs to gene promoters. Equally important is understanding of the extent to which enzymatic demethylation and replication-mediated loss of methyl groups are operative in reactivation of genes that are inheritably repressed. The genomic distribution of methylated marks is emerging as a resource for epigenetic signature. It is important to discriminate between methylation signatures that are inheritable through mitoses and methylation patterns confined to a single cell cycle.

**Histone modifications: Determinants of accessibility**

Post-translational modification of nucleosomal histones at the amino-terminal tails is a well-studied epigenetic mechanism (5-8,26); histone phosphorylation and acetylation were among the first parameters of epigenetic control (27-29). Histone modifications are dynamic and enzymes that add or remove post-translational histone modifications have been identified (5-8). Modified histones disrupt contacts between
nucleosomes and DNA and render chromatin more accessible for recruitment of non-histone, transcriptional regulatory proteins (5-8,26). Additional complexity is offered by the ordered and sequential nature of histone modifications that either attenuate or accentuate ongoing transcription. Relative turnover of histone modifications, choice of modified residues, and unique functional consequences of each histone modification have led to the proposal of a histone code that governs gene regulatory events. An in-depth discussion of the contributions of histone modifications to context dependent regulation of gene expression is covered in this Mini Review Series.

Emerging perspective suggests that histone modifications play a role in gene bookmarking during mitosis and in inheritance of chromatin state during differentiation. Histone modifications often work in concert with other epigenetic marks (e.g., DNA methylation) to ensure the inheritance of specific gene expression patterns through cell generations (30-32). Histone modifications have been implicated in a number of epigenetic phenomena including the inheritable silencing of heterochromatin (33-35). Complexity of heterochromatin formation is reflected by its absence in human embryonic stem cells (hESC). In contrast, the sperm heterochromatin is highly compact (33). In each scenario, it is important to maintain the state of heterochromatin. Methylation of Lysine 9 of histone H3 (H3K9-me) plays a crucial role in the transmission of the heterochromatin (34,35). Once methylated, Lysine 9 recruits heterochromatin protein 1 (HP1) to the chromatin. HP1 recruits additional H3K9-methylating activity that modifies nucleosomes on the nascent strand, ensuring the transmission of the H3K9-me mark (36). It is likely that other epigenetic modifications and histone variants contribute to inheritable heterochromatin silencing (37). One of the mechanisms that maintain heterochromatin during interphase is the incorporation of a variant of histone H2A (macro H2A) into DNA to be packaged as heterochromatin. But, mechanisms that prevent heterochromatin formation in hESC as well as initiate and maintain heterochromatin during and following lineage commitment are minimally understood.

A subset of genes is ‘marked’ by specific histone modifications and/or variants to designate inheritable expression post-mitotically or during differentiation (38). Inheritance of regulatory machinery for myogenesis provides an example where both histone modifications and variants contribute to epigenetic memory (38,39). Nuclear transplant experiments have shown that the methylation status of the muscle-specific MyoD promoter remains unaltered in non-muscle cell lineages (39). Instead, Gurdon and co-workers have shown that the association of histone H3 variant (H3.3) with the MyoD promoter dictates epigenetic memory. H3.3 is associated with the MyoD gene in embryos that display ‘transcriptional memory’ but not in those where memory has been lost. Mechanistically, blocking the methylation of H3.3 at lysine 4 leads to the loss of transcriptional memory, while over-expression of H3.3 contributes to memory. Long-term retention of histone H3.3 with the MyoD promoter has been established as a component to lineage-restricted transcriptional memory. Consequently, histone H3.3 stabilizes transcription during development (38,39).

These observations indicate that mechanisms exist during mitosis to maintain open chromatin architecture on selective genes for expression following cell division. However, the extent to which the histone code that operates in interphase is retained during mitosis remains to be determined.
Important open-ended questions relate to the existence of a mitotic histone code. Are both histone modifications and histone variants involved in mitotic gene bookmarking? Does the entire cohort of histone modifications persist? Is the stoichiometry of histone protein variants present in nucleosomes on a gene locus a critical determinant for epigenetic transmission of transcriptional state? Does genomic and biological context determine whether a gene will be bookmarked? What components of enzymology are involved?

**Non-coding RNA molecules**

Non-coding RNAs range from very small (21–25 nucleotide microRNAs) to very large, such as the Air transcript, which exceeds 100 kb in length (40, 41-44). Presence of RNA interference (RNAi) in lower eukaryotes and micro-RNA (miRs) in higher eukaryotes indicate biological roles for non-coding RNA; these include transcriptional regulation and translational inhibition that are accomplished through direct base-pairing with the DNA or RNA target, by mimicking the structure of other nucleic acids, or by functioning as a component of an RNA–protein complex (41-44).

A well-studied example of RNA involvement in heritable epigenetic regulation is X-chromosome inactivation by Xist RNA in mammals (45-50). XIST RNA is a large, non-coding, and alternatively spliced RNA that associates with the X chromosome from which it is transcribed (45). The Xist RNA is essential for inactivation of one copy of the X chromosome when more than one copy is present (48). Inactivation of the paternal X chromosome occurs in extra-embryonic tissues and consistent with this the paternal Xist copy is expressed preferentially in female pre-implantation embryos (49). The Xist RNA coats the X-chromosome that is silenced and together with recruitment of the chromatin remodeling and DNA methylation machinery, leads to stable, inheritable silencing of one copy of X-chromosome (50). A complete understanding of mechanisms that mediate, preserve and propagate X-chromosome inactivation will require additional experimental evidence. Histone modifications also contribute to X-chromosome inactivation (47), but specific roles of these modifications are not well established. Adapter proteins that integrate independent epigenetic mechanisms for X-chromosome inactivation are not known. It also remains to be established whether all RNA-mediated silencing is heritable, or the extent to which RNA control of gene expression is reversible or permanent.

**Mitotic Retention of Gene Regulatory Machinery: Sustaining the Family Legacy**

Mitotic remodeling of cellular and nuclear architecture includes dynamic re-localization of transcription factors (e.g., Ets1, Oct2, B-Myb, Sp factors) and transient degradation of key regulatory proteins (e.g., cyclins) (1,2,51-54). After mitosis, structural and functional integrity of cell’s regulatory machinery must be re-established to accommodate cell cycle and growth control as well as phenotype. Enzymatically and chemically sensitive sites on mitotic chromosomes mark active genes (18,55), indicating that some regulatory complexes remain bound to the condensed chromatin for rapid reactivation of genes following mitosis. Mitotic retention of the TFIID complex at gene promoters supports a ‘bookmarking’ mechanism to resume active transcription upon exit from mitosis (17,56).

Recent evidence that phenotypic transcription factors are associated with target gene loci on mitotic chromosomes adds a tissue-specific dimension to this concept (57-61).
The tandemly organized and developmentally regulated Globin gene loci exhibit DNase I-hypersensitive sites that are linked to erythroid lineage-restricted gene expression (62,63). These hypersensitive sites are retained for more than 20 generations (64). Mechanistically, NF-E2, a Globin gene regulator, remains bound to mitotic chromosomes (65), but GATA1 erythroid transcription factor dissociates from condensed chromatin (65,66). NF-E2 recruits chromatin-remodeling factors that include TAFII130 and CBP to support re-expression of the Globin gene post mitotically (65). Thus, NF-E2 supports the persistent hypersensitive state of globin genes. This is complemented by active histone modifications (e.g., H3 acetylation, H3-K4 dimethylation and K79 dimethylation) in distal regulatory domains of globin gene loci that are transcriptionally competent (65). Collectively, an ‘epigenetic memory’ mechanism mediates efficient and lineage-restricted reactivation of Globin gene transcription following mitosis.

Differentiation of mesenchymal progenitor cells into osteoblasts, adipocytes, or myoblasts is dictated by specific sets of master transcription factors (i.e., Runx2 in osteoblasts, C/EBPα in adipocytes, and MyoD and myogenin in myoblasts)(67-69). In interphase, these proteins reside in distinct nuclear microenvironments where they interact with co-regulatory proteins as well as cognate DNA elements to carry out biological functions (70). Studies from our group and others have shown that these regulatory proteins are also associated with phenotypic target genes on mitotic chromosomes in their respective lineages (57-61). During mitosis, these proteins localize to nucleolar organizing regions (NORs), the precursors of interphase nucleoli, and occupy RNA Pol I-transcribed ribosomal RNA genes as well as RNA Pol II-transcribed phenotypic genes that are involved in control of the cell cycle and differentiation. Post-mitotically, phenotype regulatory proteins are equally partitioned to progeny cells, and consistent with their anti-proliferative and differentiation-promoting roles, down-regulate rRNA gene transcription, which is intimately linked to cell growth, thus coordinating cell proliferation and growth with lineage maintenance (57-61). The occupancy and regulation of genes that are transcribed by both RNA Pol I and Pol II by these proteins during interphase and mitosis reflects coordinate genetic and epigenetic control of cell proliferation, growth and differentiation.

Asymmetric Partitioning of Transcription Factors: Separate But Not Equal

Lineage progenitors as well as stem cells must undergo asymmetric cell division to sustain the availability of a progenitor pool for two important cell types: pluripotent cells as well as lineage committed cells (71,72). Most informative observations regarding the asymmetric cell division have come from lower eukaryotes (73-75). For example, in budding yeast, the Ace2 chromatin-remodeling factor is asymmetrically accumulated in the nucleus of the daughter cell (76). Ace2, which is translated in G2 cells, is retained in the cytoplasm. In late mitosis, a dephosphorylation event leads to unmasking of the nuclear localization signal and nuclear translocation of the protein. Once in the daughter nucleus, Ace2 is phosphorylated by the daughter-cell specific Cbk1 kinase, which results in masking the export signal and daughter cell-specific nuclear retention of Ace2 (76). It remains to be explored whether daughter-cell specific retention of Ace2 results in resumption of gene expression that is restricted to the progeny.

Asymmetric cell division has also been observed during Drosophila neurogenesis, where the progenitor
undergoes several rounds of asymmetric division (74). Each division produces to a large cell that remains a neuroblast, and a smaller ganglion mother cell that eventually differentiates into neurons and glia of the central nervous system. In parallel, the sensory organ precursor also undergoes asymmetric cell divisions to give rise to five distinct cell types of the peripheral nervous system. In both these progenitors, Numb, an attenuator of Notch signaling, and the Prospero transcription factor distribute asymmetrically to only one of the daughter nuclei (77-79). This asymmetry is achieved through the actions of a protein complex that localizes to the apical cortex of dividing neuroblasts and regulates orientation of the mitotic spindle, thus controlling the polarity of these proteins.

Similar mechanisms may be operative in mammalian cells for maintenance of stem/progenitor cell pools, while giving rise to distinct cell types from a single progenitor. However, experimental evidence to establish the existence of asymmetrical distribution of regulatory complexes in mammals is lacking. It also remains to be determined how phenotypic properties of progeny cells that arise from an asymmetric cell division are established. A principal limitation is the lack of an in vitro model system where in vivo mechanisms that underlie asymmetric distribution can be recapitulated.

Mitotic retention of gene regulatory machinery, particularly on chromosomes, offers a new dimension to epigenetic regulation during cell division, both symmetric and asymmetric. Together with DNA methylation and histone modifications, transcription factor association with mitotic gene loci or their retention during mitosis provides architectural epigenetic signatures, with several regulatory implications. For example, transcription factor association may ‘tag’ genes for rapid reactivation post-mitotically. Another implication is that mitotic association of chromatin remodeling proteins (e.g., Ace2) with gene loci may render genes more accessible for regulatory proteins in G1 cells. Additionally, mitotic association of phenotypic regulatory proteins with genes in an allele-specific manner in omnipotent and pluripotent cells can facilitate asymmetric distribution of transcription factors to cells destined to commit to a particular lineage. In committed cells, equal partitioning of phenotypic transcription factors can then lead to maintenance of the lineage. Symmetric distribution of regulatory factors has been demonstrated (57), but whether this is the rule or exception must be established. Other key questions that remain include: How are genes that do not retain cognate regulatory factors during mitosis marked for activation or suppression in progeny cells? How are genes epigenetically regulated in a selective manner during asymmetric cell division? Are all, some, or none of co-regulatory proteins of a particular transcription factor present at mitotic gene loci? Despite open-ended questions, this epigenetic mechanism offers a flexible, yet cell type specific and context dependent dimension to gene regulation.

**Epigenetic Memory Sustains the Disease Phenotypes**

Cancer development not only depends on genetic alterations, but on an abnormal cellular memory or epigenetic changes that convey heritable gene expression patterns critical for tumor initiation and progression (80-83). These aberrant epigenetic mechanisms result in global changes in chromatin packaging and in localized gene promoter changes that influence the transcription of genes important to cancer.

The hematopoietic master regulator RUNX1 provides an example of de-
regulated epigenetic control during mitosis. RUNX1 is required for definitive hematopoiesis and is frequently translocated in hematological malignancies (84). One of the most prevalent chromosomal translocations involving the RUNX1 gene is the 8;21 translocation, present in over 15% of all patients with acute myelogenous leukemia (85-87). The t(8;21) results in expression of a chimeric protein AML1-ETO from the RUNX1 gene locus. AML1-ETO shares the DNA binding domain of normal RUNX1 protein, but lacks the carboxy terminus, which is replaced by nearly full length ETO protein (85-87). AML1-ETO localizes to distinct nuclear microenvironments and interferes with RUNX1 function and results in a blast-like phenotype of myeloid progenitors (88-90). RUNX1 associates with mitotic chromosomes and is equally distributed to the progeny cells (91). RUNX1 also occupies rRNA gene promoters during interphase and downregulates the expression of pre-rRNA transcripts. However, during mitosis and interphase in leukemic cells that express AML1-ETO from one allele, rRNA genes are occupied by the chimeric protein instead. AML1-ETO occupancy of rRNA genes upregulates rRNA genes, concomitant with growth advantage of leukemic cells (91). This example provides mechanistic insights into inheritable progression and maintenance of a disease phenotype and offers novel opportunities to specifically target AML1-ETO during mitosis.

**Targeting the Inherited Epigenome in Cancer: Limitations and Promises**

Genome-wide screens of histone modifications and DNA methylation have provided unbiased means to define diagnostic epigenetic signatures for tumor type and offered promising therapeutic targets (80-83). Drugs that inhibit DNA methyltransferase activity (such as azacytidine (5-azacytidine), decitabine (5-aza-2-deoxycytidine), fazarabine (1-D-arabinofurasonyl-5-azacytosine), and dihydro-5-azacytidine) or histone deacetylase function (e.g., hydrozamates such as SAHA, cyclic peptide benzamide such as MS-275, and aliphatic acids such as phenyl butarate) are in clinical trials (92). Global effects of these inhibitors that may impede normal cellular functions are a limitation. A confounding consideration is plasticity that may be a function of DNA hemimethylation, which can influence the composite epigenetic signature as well as the optimal approach for drug design. How is the enzymology targeted specifically to these silenced genes? Are certain histone modifications unique to gene loci of oncogenes and tumor suppressors?

Consistent with these concerns, it has become clear over the past decade or so that targeting DNA methylation or histone modifications alone is not sufficient for therapeutic purposes. Because these mechanisms are cumulative and context-dependent, it is likely that they function in concert with each other or with other epigenetic mechanisms to contribute to cancer phenotype. Consequently, the strategy for effective therapy may be complex and multi-factorial.

Small non-coding RNA molecules and their relatively specific inhibition of gene expression offer attractive options for use as diagnostic markers and for targeting selective gene set. Although some studies have identified roles of small RNA molecules as epigenetic carriers, additional studies are needed to define their activity as a principal epigenetic mechanism.

In addition to its contribution to physiologically responsive gene expression during development and tissue remodeling, retention of regulatory proteins with gene loci during cell division can be translated to targeted therapy with enhanced specificity
and reduced off-target effects. In contrast, global inhibition of histone modification or DNA methylation does not provide a viable option for selectivity. During mitosis, minimal components of transcriptional regulatory complexes may be present, thus unmasking epitopes to generate a druggable target. Such a mechanism can favorably influence pharmacological kinetics, i.e., minimal drug concentration will be required. In addition, mitotic association of regulatory proteins, combined with global, genome-wide assessment of histone modifications and DNA methylation, may provide a diagnostic ‘epigenetic signature’.

**Concluding Remarks: Illuminating the Dark Side of Gene Regulation by Understanding Epigenetic Mechanisms**

With increased capabilities for genome-wide expression profiling and accruing signatures for gene expression that are linked to biological control and pathology, gaps in our ability to mechanistically account for transcriptional activation, suppression, and persistence of expression status during cell division have become evident. Elucidation of control that is outside of DNA sequences can provide additional levels of regulation. Epigenetic mechanisms that include DNA methylation, histone modifications, non-coding RNA molecules, and mitotic retention of transcription factors individually and as combinatorial signatures have been shown to be informative in predicting the biological outcome as well as in diagnosing the diseased state. The definition of epigenetic control is evolving by necessity to accommodate recognition for an expanding repertoire of non-DNA encoded information that establishes and sustains regulatory machinery. Understanding the control of these non-genomic but inheritable parameters can yield novel dimensions to regulation of cell fate and lineage commitment. These observations, in turn, will provide important insights into development, differentiation, and tissue remodeling as well as an appreciation for regulatory mechanisms that are compromised with the onset and progression of disease. Finally, while abrogation of epigenetic mechanisms in disease can result in dysfunction similar to that caused by point mutations or chromosomal translocations, reversibility renders them viable therapeutic targets.

**Acknowledgements**

The authors thank Patricia Jamieson for her assistance in preparation of this article. This work was supported by National Institutes of Health grants P01 AR048818, P01 CA082834 and 5 P30 DK32520. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.
References

**Figure Legends**

**Figure 1:** Timeline of important discoveries in genetic and epigenetic regulation of gene expression. A simplified timeline of some key events and discoveries in the fields of genetic and epigenetic gene regulation is presented. Important discoveries in genetic components of gene expression are in black text color, while the epigenetic events are labeled with blue text color. Not drawn on scale.

**Figure 2:** Epigenetic mechanisms of gene expression. Genetic and epigenetic mechanisms contribute to physiological control of mammalian gene expression. While the genetic information encodes for necessary components of gene expression, epigenetic control alters the phenotype without changing the genotype of a cell and provides an additional level of control. Shown here are some well-defined epigenetic mechanisms i.e., DNA methylation, histone modifications, and non-coding RNA-mediated modulation of gene expression. These mechanisms contribute to inheritable transcriptional memory of some of the genes through successive cell divisions and contribute to maintenance of cellular phenotype. Recent studies show that association of transcriptional regulatory machinery with target genes on mitotic chromosomes is a novel epigenetic mechanism that poise genes involved in key cellular processes such as growth, proliferation, and lineage commitment, for expression in progeny cells.