2011-07-21

Global DNA Demethylation During Erythropoiesis: A Dissertation

Jeffrey R. Shearstone
University of Massachusetts Medical School

Let us know how access to this document benefits you.

Repository Citation

GLOBAL DNA DEMETHYLATION DURING ERYTHROPOIESIS

A Dissertation Presented

By

JEFFREY R. SHEARSTONE

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
In partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

July 21, 2011

Cancer Biology Program

GLOBAL DNA DEMETHYLATION DURING ERYTHROPOIESIS

A Dissertation Presented
By

JEFFREY R. SHEARSTONE

The signatures of the Dissertation Defense Committee signifies
completion and approval as to style and content of the Dissertation

_____
Merav Socolovsky, Ph.D., M.B.B.S., Thesis Advisor

_____
Job Dekker, Ph.D., Member of Committee

_____
Schahram Akbarian, M.D., Ph.D., Member of Committee

_____
Michael Green, M.D., Ph.D., Member of Committee

_____
Laurie Jackson-Grusby, Ph.D., Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets
the requirements of the Dissertation Committee

_____
Craig Peterson, Ph.D., Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that
the student has met all graduation requirements of the school

_____
Anthony Carruthers, Ph.D.,
Dean of the Graduate School of Biomedical Sciences

Cancer Biology Program

July 21, 2011

ACKNOWLEDGEMENTS

I would like to offer my sincerest thanks to the following people:

My thesis research advisor Merav Socolovsky for funding and for the freedom to pursue a line of research that, initially, only hinted at global demethylation.

My fellow lab members who have been both enjoyable to work with and outstanding scientific resources: Ramona Pop, Ermelinda Porpiglia, Miroslav „Miro" Koulnis, Daniel Hidalgo, Qichang Shen, and Kelly Hallstrom. I would especially like to thank Ramona Pop whose knowledge, constant availability, and collaboration has allowed my project to progress smoothly.

The Cancer Biology Department for stipend support through training grant NIH CA T32-130807. Special thanks to the department members who „talked me off the ledge" on several occasions, their advice was invaluable.

The UMass flow cytometry core run by Richard Konz. Specifically, Ted Giehl and Barbara Gosselin, along with Marc Barnard and Ceyhan Karabas-Akalin, performed over 80 sorting experiments in support of this research with virtually no equipment malfunctions. They have been outstanding.

My TRAC members, Job Dekker, Schahram Akbarian, and Craig Peterson, for their time, knowledge, availability, and constructive feedback throughout this project. Special thanks to Craig for being a superb TRAC chair and to my additional dissertation examiners, Michael Green and Laurie Jackson-Grusby (Children's Hospital Boston), for their time and willingness to participate. I also want to thank my QE committee members, Leslie Shaw, Peter Newburger, Job Dekker, and Craig Peterson.

Our collaborators, Dr. Alex Meissner, Dr. Christoph Bock, and their associates at Harvard University and the Broad Institute for generating and analyzing the reduced representation bisulfite sequencing data.

Liying Yan and the staff at EpigenDx, Inc in Worcester for assistance with assay design and pyrosequencing.

En Li of the Novartis Institutes for Biomedical Research for supplying Dnmt1, Dnmt3a, Dnmt3b, Dnmt3b1:PC constructs.

My friends, family, and wife. Their love and support means the world to me, and their willingness to listen and offer an outside viewpoint has given me invaluable perspective during graduate school. I am especially indebted to my wife, who has fully backed this pursuit since it was first suggested at the Olympics nobody attended. I couldn't have done it without her. In this respect, the „receipt" now belongs to both of us.

ABSTRACT

In the mammalian genome, 5"-CpG-3" dinucleotides are frequently methylated, correlating with transcriptional silencing. Genome-wide waves of demethylation are thought to occur only twice during development, in primordial germ cells and in the pre-implantation embryo. They are followed by *de novo* methylation, setting up a pattern that is inherited throughout development. No global methylation changes are thought to occur during further somatic development, although methylation does alter at gene-specific loci, contributing to tissue-specific patterns of gene expression. Here we studied DNA methylation in differentiating mouse erythroblasts *in vivo* using several approaches including genomic-scale, reduced representation bisulfite sequencing (RRBS). Surprisingly, demethylation at the erythroid-specific β-globin locus was coincident with a wave of global DNA demethylation at most genomic elements, including repetitive elements and genes silenced in erythropoiesis. Over 30% of total methylation is irreversibly lost during erythroid differentiation. Demethylation occurred through a passive mechanism, requiring the rapid DNA replication triggered with the onset of erythroid terminal differentiation. Global loss of DNA methylation was not associated with a global increase in transcription, as determined by GeneChip analysis. We propose that global demethylation is a consequence of cellular mechanisms required for the rapid demethylation and induction of β-globin and other erythroid genes. Our findings demonstrate that, contrary to previously held dogma, DNA demethylation can occur globally during somatic cell differentiation, providing a new experimental model for the study of global demethylation in development and disease.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF THIRD PARTY COPYRIGHTED MATERIAL

Figure 1.3 was adapted by permission from Macmillan Publishers Ltd: Nature, May 24;447(7143):425-32, copyright 2007.

Figure 1.4 has been used by permission from New England Journal of Medicine, Mar 13;358(11):1148-59, copyright 2008.

LIST OF MULTIMEDIA OBJECTS OR FILES

The entire reduced representation bisulfite sequencing data set generated for erythroblast subsets in Chapter III can be found online at:

http://erythrocyte-demethylation.computational-epigenetics.org

LIST OF ABBREVIATIONS

5mC	5-methyl cytosine
5hmC	5-hydroxymethyl cytosine
5-aza	5-aza-2''-deoxycytidine
Aphi	Aphidicolin
BER	Base excision repair
BrdU	5-bromo-2'-deoxyuridine
CpG	Cytosine-guanine dinucleotide
DMR	Differentially methylated region
Dnmt	DNA methyltransferase
ELISA	Enzyme linked Immunosorbent assay
Epo	Erythropoietin
ES cell	Embryonic stem cell
HS	Hypersensitivity site
IF	Immunofluorescence microscopy
IRES	Internal ribosome entry site
LCR	Locus control region
LINE	Long interspersed nuclear element retrotransposon
LUMA	Luminometric methylation assay
Mbd	Methyl binding domain
MFI	Median fluorescent intensity
Mim	Mimosine
NER	Nucleotide excision repair
PGC	Primordial germ cell
QRT-PCR	Quantitative real time PCR
RRBS	Reduced representation bisulfite sequencing
SAM	S-adenosyl methionine
shRNA	Short hairpin RNA
SINE	Short interspersed nuclear element retrotransposon

PREFACE

The work presented in Chapter II was a collaboration with Ramona Pop, a post-doctoral fellow in the Socolovsky laboratory at University of Massachusetts Medical School. My primary contributions to her work are the DNA methylation studies presented in Figure 2.6 and described in the results section "The Transition from S0 to S1 Coincides with S-Phase-Dependent DNA Demethylation at the β-Globin LCR". I also contributed to Figure 2.5A and Figure 2.S3A-B. This work was published in the journal PLoS Biology in September of 2010: Ramona Pop, Jeffrey R. Shearstone, Qichang Shen, Ying Liu, Kelly Hallstrom, Miroslav Koulnis, Joost Gribnau, and Merav Socolovsky. "A Key Commitment Step in Erythropoiesis is Synchronized with the Cell Cycle Clock through Mutual Inhibition between PU.1 and S-Phase Progression". PLoS Biol (2010) 8(9) e1000484.

The work presented in Chapter III constitutes the major portion of my thesis research. Ramona Pop performed the work in Figure 3.11 and contributed to Figure 3.12A and 3.13A. Alexander Meissner and Christoph Bock of Harvard University and the Broad Institute performed the reduced representation bisulfite sequencing and analysis presented in Figure 3.5, 3.6A-C, 3.14, and 3.16C. This chapter is an extended version of work that is currently under review with the following title and authors: Jeffrey R. Shearstone, Ramona Pop, Christoph Bock, Alexander Meissner, and Merav Socolovsky. "Global DNA Demethylation during Erythropoiesis In Vivo".

The CD71/Ter119 flow cytometric method for identifying erythroid progenitors and erythroblasts, as presented in Figure 1.1, Figure 2.1, and Figure 3.1 is currently in

press at the Journal of Visualized Experiments and will be published in August 2011 with the following title and authors: Miroslav Koulnis*, Ramona Pop*, Ermelinda Porpiglia*, Jeffrey R. Shearstone*, Daniel Hidalgo, and Merav Socolovsky. "Identification and analysis of mouse erythroid progenitors using the CD71/Ter119 flow-cytometric assay." J Vis Exp (2011) *Contributed equally.

I have also performed experiments of general relevance to our laboratory that are not discussed in this thesis. These experiments will provide a basis for future lines of research, as well as support current projects in the lab. For example, I have conducted Affymetrix GeneChip genome-wide mRNA profiling on S0, S1, S3, EpoR knockout, S0 „mid" wild type, and S0 „mid" EpoR knockout cell populations. This work has been done for triplicate biological replicates allowing statistical analysis of gene expression changes within key erythroid differentiation subsets. I have also analyzed and consolidated this data into a spreadsheet from which lab members can query the expression level of their gene(s) of interest.

Also not presented in this thesis are microRNA profiling experiments that I conducted on sorted S0 and S1 subsets. By combining microRNA expression, predicted mRNA targets from the TargetScan database, and differentially regulated mRNA from the GeneChip data described above, I computationally identified groups of miRNAs that synergize to suppress gene expression during erythropoiesis. This work was not pursued further in the wet lab because other projects took priority, but the data and resulting analysis should provide a starting point for new lines of research in the future.

CHAPTER I

Introduction

DNA Methylation and Gene Regulation

The basic building block of chromatin is comprised of genomic DNA spooled around a histone protein octamer. Histones or DNA can be modified by the enzymatic addition of various chemical groups, resulting in recruitment or exclusion of associated proteins that have profound effects on local chromatin structure and gene expression. Because these modifications fundamentally do not alter the DNA sequence and because they have the potential to be mitotically and/or meiotically heritable, they have been termed „epigenetic‟ modifications. One such epigenetic mark, DNA methylation (5mC), is simply the presence of a methyl group on the 5‟-carbon ring of cytosine. In mammals, the vast majority (>99%) of methylation is found to occur symmetrically at the dinucleotide palindrome cytosine-guanine (CpG), although asymmetrical CpHpG and CpHpH (H=A, T, or C) methylation has been detected in embryonic stem (ES) cells [1, 2]. DNA methylation has long been correlated with transcriptional inactivity and heterochromatin formation. The repressive effect of DNA methylation on gene transcription has been directly tested using transient transfections and stable integration at defined genomic loci [3-7]. Furthermore, DNA methylation is essential for X-chromosome inactivation, parental-specific silencing of imprinted genes, suppression of transposable elements, and silencing of gene-specific loci [8-15].

Proper establishment and maintenance of 5mC patterns are required for mammalian development and normal functioning of somatic cells. These processes are mediated through the DNA methyltransferase (Dnmt) family of proteins; the *de novo* methyltransferases Dnmt3a and Dnmt3b establish 5mC patterns, while the maintenance

methyltransferase Dnmt1 and accessory factors act to preserve existing patterns during DNA replication and subsequent cell division [16-25]. Knockout studies in mice show that DNA methyltransferases are essential genes, with death of Dnmt1$^{-/-}$ or Dnmt3a$^{-/-}$3b$^{-/-}$ embryos before day 11, Dnmt3b$^{-/-}$ embryos before birth, and Dnmt3a$^{-/-}$ mice at 4 weeks of age [13, 26-28]. The requirement for methylation in somatic cells has been well documented. For example, absence of Dnmt1 prevents differentiation of ES cells into embryoid bodies, teratomas, cardiomyocytes, or hematopoietic progenitors [26, 29, 30]. Inactivation of Dnmt1 in embryonic fibroblasts results in demethylation and p53-dependent apoptosis [31]. Similarly, maintenance of methylation by Dnmt1 is required for proliferation and survival of the T-cell lineage, survival of postnatal central nervous system neurons, and self-renewal of somatic stem cells of the epidermal and hematopoietic lineages [32-35].

Methylation generally occurs at 70-80% of all CpG sites, a consistent finding across a wide-range of somatic cell types [1, 36]. Somatic genomes are not methylated uniformly but contain methylated regions interspersed with unmethylated domains. For example, CpG islands, the CpG-dense (about 1 CpG per 10 bp) regions found at 70% of promoters [37], are almost always unmethylated regardless of developmental stage, tissue type, or transcriptional activity [12, 38-42]. In contrast, CpGs that are found infrequently (about 1 CpG per 100 bp) are highly methylated. Sites of low CpG frequency occur predominantly in repetitive genomic regions, including satellite DNA and retroelements, but are also common at 30% of promoters, exons, introns, 3" untranslated regions, and at the less CpG-rich regions flanking CpG-islands called „island shores" [1, 42, 43].

Importantly, despite these broad similarities in the overall 5mC abundance and distribution, the detailed DNA methylation pattern of each somatic cell type is distinct [1, 38-40, 42, 44].

Differentially methylated regions (DMRs) are short stretches of DNA (about 250 bp) that distinguish one somatic cell type from another on the basis of 5mC. The combination of genome-wide methylation analysis and transcriptional profiling has led to the characterization of DMRs that correlate with tissue-specific gene expression *in vivo*. DMRs are found infrequently at high CpG density promoters (about 5% of all DMRs), suggesting these regions are not major targets for developmental regulation [38, 40, 42-45]. Surprisingly, DMRs occurred most frequently at intermediate CpG density promoters and showed a high inverse relationship with gene expression [40, 46]. Consistent with this result, more detailed studies found the majority of DMRs reside at island shores, within 2 kb of a CpG island [43, 47]. Furthermore, these DMRs were tightly conserved between human and mouse tissues [38, 43, 47]. Collectively, these results show that intermediate CpG density promoters and island shores are regions of dynamic methylation that contribute significantly to gene expression and somatic cell identity.

DNA methylation exerts its repressive effects directly and indirectly. Many studies have demonstrated that 5mC is able to directly prevent the binding of transcription factors to DNA, as is the case for E2f, Creb, Usf1, Ap-2, and Myc/Max [48-56]. Similarly, DNA methylation at Ctcf-binding sites directly prevents Ctcf from binding, thereby altering Igf2 expression at the imprinted H19/Igf2 locus [48, 57].

Notably, Mll1 (mixed lineage leukemia 1), a histone H3 lysine 4 (H3K4)-specific methyltransferase, can only bind its target CpGs when they are unmethylated, suggesting how DNA methylation might inhibit the addition of a histone modification associated with active chromatin [51].

DNA methylation indirectly silences transcription by recruiting methyl-CpG binding proteins which associate with various chromatin modifiers. These proteins include members of the methyl binding domain (Mbd) family, such as MeCP2, Mbd1, Mbd2, and Mbd4; and the Kaiso family, which use a three-zinc-finger motif to specifically bind methylated DNA [58-61]. MeCP2 is found complexed with histone 3 lysine 9 (H3K9) methyltransferase activity and also binds the Sin3a, NCoR, and Ski co-repressor complexes which contain histone deacetylase activity [62-65]. Mbd1 associates with the H3K9-specific methyltransferases Setdb1 and Suv39h1 [66, 67]. Mbd2 interacts with the NuRD co-repressor complex that has both histone deacetylation and chromatin remodeling ATPase activity [68-71]. Additionally, Sin3A and NCoR co-repressor binding has been implicated in Mbd4 and Kaiso-dependent transcriptional repression [72, 73]. Interestingly, methyl-CpG binding proteins have distinct sequence preferences, providing a means to target 5mC-facilitated repression to specific genomic regions [60, 73-75]. Taken together, methyl-CpG binding proteins link DNA methylation to heterochromatin-associated histone modifications.

Conversely, histone methylation or the enzymes which modify histones are able to coordinate DNA methylation. For example, Dnmt3L-mediated recruitment of Dnmt3a and Dnmt3b to DNA is blocked by H3K4 methylation [76]. Methylation of histone 4

arginine 3 (H4R3) serves as a direct binding target of Dnmt3a, an interaction leading to 5mC addition at the γ-globin promoter [77]. Additionally, the H3K9 methyltransferases G9a, Suv39h1 and Setdb1, and the H3K27 methyltransferase EZH2, can mediate DNA methylation by direct interaction with Dnmt3a and Dnmt3b [65, 78-82].

The close connection between histone modification and DNA methylation is illustrated by the finding that the presence of H3K4 methylation and absence of H3K9 methylation are better predictors of unmethylated CpGs than CpG-density alone [39, 42]. Clearly, these epigenetic marks play an important and interrelated role in gene expression regulation. A model has been proposed where histone modifications act as flexible repressors, while DNA methylation results in highly stable inactivation [83]. As discussed in the following sections, it is less clear as to how DNA methylation is removed to permit gene activation.

β-globin Expression and Definitive Erythropoiesis

β-globin is an essential subunit of the oxygen-transporting hemoglobin complex found in erythrocytes. The β-globin locus contains an ordered cluster of 5 human genes (ε, ^{G}γ, ^{A}γ, δ, β) or 4 mouse genes (εy, βh1, βmaj, βmin) that are sequentially switched on/off during development [84]. In mouse, primitive erythropoiesis occurs in the yolk sac at embryonic day 7.5 (E7.5) and is defined by expression of the embryonic β-globins εy and βh1 [85]. Primitive erythropoiesis is transient, lasting only a few days, and results in nucleated erythrocytes. Definitive erythropoiesis begins at E11.5 in the fetal liver and is marked by expression of the adult globins βmaj and βmin [86]. Near the time of birth, the

site of definitive erythropoiesis moves to the bone marrow, where it persists throughout life. Primitive and definitive erythroblasts are considered two distinct cell lineages [84].

Full transcriptional activation of β-globin is dependent on the locus control region (LCR), a segment of DNA found 6 kb to 20 kb upstream of the embryonic εγ gene [87, 88]. The LCR is distinguished by five, 200-300 bp DNase I hypersensitivity sites (HS1-5), which acquire a relaxed chromatin state during β-globin activation [84, 89, 90]. This chromatin change exposes transcription factor binding motifs within the hypersensitivity sites and, in one proposed model, facilitates the looping of the LCR into regions proximal to β-globin promoters [91].

Epigenetic modifications play a key role in the lineage-specific regulation of the β-globin gene cluster [92-94]. For example, differences in the CpG methylation profile of the LCR or at 5" regions of β-globin genes have been observed in primitive and definitive erythroid lineages [95]. Additionally, in a large number of studies, gene activating histone marks are enriched at the LCR and β-globin loci of primitive [85, 96, 97] and definitive precursors [92, 94, 96-99], as well as in erythroid cell lines [96, 98, 100-103]. Unfortunately, the current understanding of epigenetic changes at the β-globin LCR is based on the differences between heterogeneous, late-stage erythroblasts of the primitive and definitive lineages, rather than the progressive changes that occur during the maturation of a single lineage. Moreover, the mechanism by which LCR demethylation occurs is completely unknown.

The developmental stage of definitive erythroid progenitors is traditionally defined morphologically or by colony formation in soft agar. Differentiation of these

progenitors is accompanied by temporally regulated changes in cell surface protein expression, decreasing cell size, increasing hemoglobin expression, nuclear condensation, and ultimately, enucleation. Recently, a flow cytometry method, based on the cell surface markers CD71 and Ter119, has been developed to identify subsets within the fetal liver that form an erythroid developmental sequence [104, 105]. Fetal liver cells may be classified into 6 such subsets, each at an increasing stage of differentiation (S0-S5) (Figure 1.1). S0 is the earliest developmental subset, composed largely of early erythroid progenitors, while S5 cells contain nucleated, hemoglobinized erythroblasts and enucleated erythrocytes. Therefore, flow cytometric sorting of CD71/Ter119-staged fetal liver cells has the potential to address the timing of epigenetic alterations, such as CpG methylation, at the β-globin gene cluster during erythropoiesis.

Passive DNA Demethylation

Demethylation that requires DNA replication is termed „passive", while replication-independent demethylation is classified as „active". Each mode of demethylation has been implicated in both gene-specific and global demethylation processes. Active demethylation is well established in plants, but some have questioned its prominence in mammals [106]. Conversely, some are doubtful that passive demethylation plays a major role in gene-specific demethylation [107]. In the remainder of this chapter, evidence for both mechanisms will be presented with an emphasis on the events that drive global demethylation in the early embryo, primordial germ cells, and cancer.

During semi-conservative DNA replication, unmethylated deoxycytidine is incorporated into the nascent strand, creating hemi-methylated DNA at CpG sites where the parent strand is methylated. Dnmt1, which is targeted to replication foci and has a 20 to 40-fold preference for hemi-methylated DNA, is able to restore methylation symmetry by catalyzing the transfer of a methyl group from an S-adenosylmethionine (SAM) donor [17, 18, 22, 108]. Inhibition of this „maintenance" methylation leads to a passive loss of up to 50% of 5mC per round of DNA replication (Figure 1.2A).

Targeting of Dnmt1 to replication foci is promoted by an interaction with proliferating cell nuclear antigen (Pcna) [17]. In Dnmt1 mutants deficient for Pcna binding, the efficiency of methylation was reduced by 2-fold [109, 110]. Another critical co-factor is Uhrf1 (ubiquitin-like, containing PHD and RING finger domains 1), which binds specifically to hemi-methylated DNA and recruits Dnmt1 to these sites by a direct association [111, 112]. Loss of maintenance methylation observed during gliomagenesis is associated with a loss of Dnmt1-Uhrf-Pcna interactions and a decline in recruitment of Dnmt1 to DNA. Loss of these interactions was mediated by phosphorylation of Dnmt1 at serine 127, which prevented binding to Uhrf1, and additional phosphorylation at serine 143, further abrogated Pcna binding [113]. Similarly, a peptide of the Pcna binding domain from cyclin dependent kinase inhibitor 1A can disrupt the Dnmt1-Pcna interaction [17]. Therefore, the combined action of Pcna, Uhrf1, and Dnmt1 is critical to the high fidelity of 5mC maintenance and disruption of any of these interactions could lead to passive demethylation.

Maintenance methylation also depends on the level and activity of Dnmt1, which increases during S-phase, decreases after S-phase, and reaches its lowest level during G1 [114-116]. Dnmt1 protein stability is regulated by a complex system of post-translational modifications; Usp7 and Uhrf1 deubiquitnate and ubiquitnate Dnmt1, respectively [114, 117]; Tip60 promotes Dnmt1 degradation through acetylation, while stability is induced by Hdac1-mediated deacetylation [114]; and methylation of Dnmt1 by Set7 at Lys-142 leads to its degradation, while Lsd1 demethylates and stabilizes Dnmt1 [118, 119]. Additional post-translational modifications, such as phosphorylation and SUMOylation, seem to positively affect Dnmt1 activity [120, 121]. Dnmt1 activity is also controlled by its localization, as is the case in the early embryo where the maternal specific splice variant of Dnmt1 is excluded from the nucleus during early cleavage divisions [122, 123]. In human embryonic kidney (HEK)-293 cells phosphorylation near the nuclear localization signal motif of Dnmt1 enhances its localization in the nucleus [124]. Finally, patients with neurodegenerative disorders exhibit a mislocalization of Dnmt1 in the cytoplasm and a significant loss of global DNA methylation [125].

Maintenance methylation can also be inhibited by the binding of protein factors to DNA. For example, Sp1 binding sites induced localized promoter demethylation of an in vitro methylated plasmid injected into *Xenopus* embryos, and Gal4 binding sites could be demethylated in the presence of exogenous transcription factors designed to target those sites [126]. Importantly, in both cases DNA replication was required. Similarly, DNA replication and EBNA-1 binding was required for demethylation of an episomal oriP sequence in PC-3 prostate cancer cells [127]. Furthermore, episomal lacO sites become

demethylated upon binding of LacI in the human embryonic kidney 293 cell line [128]. A decrease in LacI binding affinity, caused by increasing amounts of IPTG, resulted in a slower loss of methylation. Collectively, these studies suggest that binding of transcription factors can protect hemi-methylated CpGs from the maintenance activity of Dnmt1, a mechanism recently proposed in neural precursor cell differentiation [129].

Recently it was discovered that the Tet family of proteins are 5mC hydroxylases that catalyze the conversion of 5mC to 5-hydroxymethycytosine (5hmC) [130]. 5hmC accumulates during global DNA demethylation of the paternal genome in the early embryo, discussed later in more detail, and is found at relatively high frequency in ES cells [131-133]. Interestingly, DNA sequences containing 5hmC are not substrates for the maintenance methyltransferase Dnmt1 [134]. Therefore, a defect in maintenance methylation could result from a mechanism where 5mC is first converted to 5hmC, followed by dilution through DNA replication, but this has yet to be shown to occur *in vivo*.

Although Dnmt1 is responsible for the majority of maintenance methylation, some have suggested that the *de novo* methyltransferases Dnmt3a and Dnmt3b participate in this process [135]. Unlike Dnmt1, Dnmt3a and 3b have equal affinity for both hemi- and un-methylated CpGs [21]. In ES cells lacking Dnmt3, global methylation was slowly but progressively lost with continued passage, even though Dnmt1 remained stably expressed [16, 136]. These cells also exhibit an increase in the frequency of hemi-methylated sites as compared to their wild-type ES cell counterparts [136]. Similarly, inactivation of Dnmt3b leads to retroelement and minor satellite repeat hypomethylation

in murine embryonic fibroblasts [137]. Furthermore, combined knockdown of Dnmt1 and Dnmt3b enhanced the methylation loss observed by Dnmt1 knockdown alone [138]. These results are consistent with a model where, soon after DNA replication, Dnmt3a and Dnmt3b correct errors that result from Dnmt1 infidelity [135]. Therefore, a reduction in Dnmt3a or 3b activity might contribute to inhibited maintenance methylation.

Maintenance methylation activity is also modulated by the cellular steady state concentration of SAM, the methyl donor utilized by DNA methyltransferases in a reaction that forms 5mC and S-adenosyl homocysteine (SAH). SAM is also decarboxylated (dcSAM) and then used as an aminopropyl group donor for the synthesis of polyamines. Both SAH [139, 140] and dcSAM [141] act as competitive inhibitors of DNA methyltransferases and disruption in the balance of SAM relative to these by-products is linked to hypomethylation [142-147]. SAM is derived from methionine, whose formation is in turn dependent on the levels of methyl group donors such as choline, folate, and vitamin B-12. Folate is also required for the de novo synthesis of deoxynucleotides, thereby linking the metabolic pathways involved in DNA synthesis and DNA methylation [148]. Diets deficient in methionine, folate, choline and B-12 lead to reduced levels of SAM and genomic hypomethylation in rat hepatocytes [149-151]. Conversely, maternal dietary supplementation with these factors can permanently alter offspring coat color by increased methylation at the agouti gene promoter [152].

Active DNA Demethylation

Erasure of DNA methylation in a manner that does not require DNA replication is, by default, called active. Active demethylation has been extensively documented in

plants and animals, with the majority of examples acting locally, at gene-specific regions [107, 153]. In contrast, active genome-wide demethylation has only been documented in a handful of physiological situations; in plants during endosperm development and in mammals during pre-implantation embryo development and primordial germ cell specification, as will be discussed in later sections (Figure 1.3) [154-157]. In some of these cases, such as the promoter regions of interleukin-2 and interferon-γ genes in T-cells, the authors have formally excluded DNA replication as a contributing factor to demethylation using rapamycin and mitomycin C, respectively, to block cell cycle progression [158, 159]. Many other examples rely on fast kinetics and non-dividing/post-mitotic nature of the model system as proof of active demethylation. The exact mechanism of active demethylation in mammals is still unknown and this area of research continues to be contentious [106]. Five potential mechanisms have been proposed, each with varying amounts of supportive evidence (Figure 1.2B)

In one proposed mechanism the methyl group is directly cleaved from cytosine. Mbd2 (methyl binding domain 2) was identified as having this ability [160]. However, this finding could not be repeated when attempted by two independent groups and Mbd2 knockout mice are viable with no alteration in 5mC levels [69, 71, 161].

The nucleotide excision repair (NER) pathway has also been implicated in demethylation. In two independent reports, Gadd45a (growth arrest and DNA-damage-inducible gene 45 alpha) overexpression promoted gene specific and/or global demethylation [162, 163]. Conversely, knockdown of Gadd45a induced hypermethylation. Gadd45a associated with Ercc5 and other members of the NER

pathway and knockdown of Ercc5 led to hypermethylation or inhibited demethylation. However, attempts to repeat many of these results have proved unsuccessful and biochemical evidence is lacking, leaving the role of the NER pathway in active demethylation in question [164].

More recently, Elp3 and other Elp family members were identified as necessary for genome demethylation in the paternal pronucleus of the murine zygote [165]. Elp proteins are components of the RNA polymerase elongator complex, suggesting a link between transcription and DNA demethylation. Demethylation required the Fe-S radical SAM domain of Elp3, but otherwise the mechanism by which Elp proteins act is entirely unknown.

Methyl groups are known to be removed through a base excision repair (BER) pathway. This has been well documented in plants, where direct excision of methyl-cytosine is performed by the DME/ROS1 family of bifunctional glycosylases [166-168]. These enzymes first remove the 5mC base to form an apurinic/apyrimidinic (AP) site. Next, using their AP lyase activity, they cleave the abasic site. AP endonucleases remove the sugar group and the resulting single nucleotide gap is then filled with unmethylated cytosine. In mammals, no such 5mC glycosylase orthologues have been identified. However, the T/G glycosylases Mbd4 (methyl binding domain 4) and Tdg (thymine-DNA glycosylase) are able to act as a 5mC glycosylases *in vitro*, albeit with a 30-fold reduced activity compared to their T/G mismatch activity [169, 170]. Furthermore, recent work suggests a 5mC glycosylase activity for Mbd4 that is dramatically improved upon its phosphorylation [171].

A more likely BER-based mechanism in mammals begins by deamination of 5mC to form a thymine, followed by T/G mismatch repair via the T/G glycosylase activity of Mbd4 or Tdg. The Aicda (activation-induced cytidine deaminase) and Apobec (apolipoprotein B mRNA editing enzyme, catalytic polypeptide) protein families are cytidine deaminases that act on 5mC *in vitro* [172]. Aidca and Apobec2 are required for demethylation of genomic and exogenous DNA in zebrafish embryos and co-expression of Mbd4 enhanced this effect [173]. Importantly, T/G mismatch intermediates were detected, supporting deamination as one step in the demethylation process. Other evidence linking Aidca to demethylation comes from Aidca[-/-] primordial germ cells, where global methylation levels were found to be slightly elevated, and nuclear reprogramming in interspecies heterokaryons, where siRNA knockdown of Aidca prevented Nanog and Oct4 promoter demethylation [174, 175]. In addition to Aidca and Apobec, the catalytic domains of Dnmt3a and Dnmt3b have deaminase activity *in vitro* when SAM availability is limited [176]. Evidence that active demethylation is mediated by Tdg comes from studies in Tdg[-/-] MEFs where Tdg was required for gene-specific hypomethylation and recruitment of the DNA repair enzymes Xrcc1 and Apex1 [177]. Furthermore, Tdg knockdown attenuated pS2 promoter methylation loss *in vivo* [176].

The ten-eleven translocation (Tet) family of proteins is capable of oxidizing hemi- or fully-methylated 5mC to produce 5hmC [130, 131]. A variety of active demethylation mechanisms that utilize a 5hmC intermediate have been proposed [153]. For example, direct removal of 5hmC by a 5hmC glycosylase could induce BER-mediated replacement with cytosine [178]. Alternatively, deamination of 5hmC to form hydroxymethyluracil

(5hmU) followed by 5hmU removal by a 5hmU glycosylase could invoke the BER pathway. This mechanism is strongly supported by a very recent study that found Tet1, when introduced into HEK-293 cells, could induce 5mC loss of exogenous template and gene-specific loci [179]. Template DNA containing 5hmC, but not 5mC, was demethylated, an effect that could be prevented by chemical inhibition of the BER pathway [179]. Furthermore, Aidca and Apobec family members significantly increased 5hmC demethylation [179]. Finally, expression of Tet1, introduction of 5hmC template DNA, or overexpression of Apobec in HEK293 cells each led to an increase in cellular 5hmU [179].

Global DNA Demethylation in the Early Embryo

Prior to fertilization, the overall DNA methylation level of sperm is equal to somatic cells [180], while the egg shows a lower initial level of methylation [156, 157, 181-185]. The zygote that results from fertilization undertakes several cell divisions in the formation of the blastocyst, during which the parental DNA undergoes genome-wide methylation loss that affects single copy genes, long interspersed repeated elements (LINE), intracisternal A particle elements (IAP), early retrotransposons (Etn), and centromeric satellites (Figure 1.3) [156, 157, 181-186].

Immunofluorescence microscopy (IF) using an anti-5-methylcytosine antibody (5mC-IF) has been the primary tool for measuring DNA methylation loss in the early embryo. In these studies, virtually all methylation appears to be lost during blastocyst formation. In contrast, about 50% of methylation is lost when measured by non-5mC-IF assays, such as those that utilize bisulfite conversion of genomic DNA [187]. Only

imprinted regions appear to be spared from demethylation [123, 188-190]. Genome-wide DNA methylation profiling of egg, sperm, and blastocyst has yet to be done, but is necessary for a comprehensive understanding of the magnitude and sequence specificity of methylation loss during blastocyst formation.

It is noteworthy that the timing and mechanism of paternal and maternal DNA demethylation during blastocyst formation are considerably different. Maternal 5mC loss occurs gradually over several post-zygotic cell divisions [186, 191]. During this time, the highly abundant, maternally derived isoform of Dnmt1 (Dnmt1o) is localized to the cytoplasm [122, 123]. Therefore, maternal 5mC loss is thought to occur through a passive mechanism, i.e. DNA replication in the absence of nuclear Dnmt1. Yet 50% of global 5mC is maintained in non-5mC-IF assays and imprinted regions are protected from demethylation, suggesting that some maintenance methylation activity is still present. Consistent with this idea, recent studies have shown that the somatic Dnmt1 isoform (Dnmt1s) is present in the nucleus of the early embryo, albeit at a far lower level than cytoplasmic Dnmt1o [123, 192, 193]. Antibody neutralization or siRNA knockdown of Dnmt1s resulted in a loss of 5mC at IAP repeats and a H19 imprinted region at the morula stage [193]. Furthermore, conditional knockout experiments demonstrated that Dnmt1s and Dnmt1o, but not Dnmt3, is required to maintain imprints in the preimplantation embryo [123]. These results also show that the extent of methylation loss in a passive mechanism is critically sensitive to Dnmt1 abundance.

In contrast to the maternal pronucleus, the paternal pronucleus begins to demethylate in the one-cell embryo, shortly after fertilization [191]. *In vitro* fertilization

experiments established that 5mC loss was initiated 4 hours post-fertilization and was completed by 8 hours post-fertilization, just prior to fusion of parental pronuclei [186]. Similar demethylation timing has been observed in rat, pig, bovine, and human [194, 195]. These 5mC-IF results are supported by bisulfite sequencing at several endogenous genes, including Igf2, a foreign transgene TKZ751, and LINE and Etn regions [165, 182].

There are several lines of evidence supporting the idea that paternal pronuclear 5mC loss is via an active mechanism. First, methylation loss is observed before S-phase of the first cell cycle [182, 191]. Second, the demethylation observed by IF and at the Igf2 and TKZ751 loci is so severe that it could not simply result from a lack of maintenance methylation [182]. Finally, demethylation still occurred in the presence of the DNA replication inhibitor aphidicolin [191, 196]. Active demethylation of the paternal pronucleus is thought to proceed through the BER pathway, since 5mC loss is correlated with the appearance of repair markers γH2A.X, Parp, Apex1, and DNA-bound Xrcc1 by IF [184, 197]. Additionally, Parp or Apex1 inhibitors were able to prevent 5mC loss [197]. Elp3 is also required for demethylation, but the molecular mechanism by which it acts is unknown [165].

However, the evidence for active paternal DNA demethylation is far from conclusive. For example, the exact timing of DNA replication in the mouse zygote varies across laboratory and methodology [184, 198, 199], which has raised the possibility that paternal DNA may have at least partially replicated in previous studies [133]. Importantly, emerging research using fluorescent antibodies against 5hmC calls conclusions from previous 5mC-IF data into question. 5mC reduction coincides with

5hmC accumulation specifically in the paternal, and not maternal, pronucleus prior to S-phase of the first cell cycle [133, 200]. Accumulation of 5hmC occurred in the presence of aphidicolin, which is consistent with the 5mC loss observed in aphidicolin. Finally, knockdown of the Tet3, which oxidizes 5mC to form 5hmC, diminished the loss of 5mC and prevented 5hmC accumulation. These data strongly suggest that loss of 5mC staining in previous 5mC-IF studies is simply a reflection of its conversion into 5hmC, rather than a *bona fide* active demethylation event. 5hmC has been speculated to act as a substrate or as an intermediate in an active, repair-coupled demethylation process [201]. Even more provocative is the finding DNA sequences containing 5hmC are not substrates for the maintenance methyltransferases Dnmt1 [134]. Therefore, 5mC loss in the paternal pronucleus could be the result of a passive mechanism, whereby 5mC is first converted to 5hmC, followed by dilution through DNA replication [133].

The significance of global methylation loss in the early embryo is still under investigation. One theory is that genome-wide reprogramming might play an important role in preventing the transgenerational accumulation of potentially detrimental epialleles, which could cause chronic disease and limit life span [202, 203]. Instead, current evidence suggests that global methylation loss is required for the activation of genes leading to pluripotency. For example, Nanog is a known pluripotency regulator that is completely methylated in sperm, but lacks any 5mC in the blastocyst or embryonic stem (ES) cells [204]. Targeted methylation of an exogenous Nanog promoter was sufficient to silence its expression in ES cells, emphasizing the importance of demethylation in Nanog gene activation.

The majority of supporting evidence for global demethylation-induced activation of pluripotency-specific genes is derived from work investigating reprogramming of somatic cells to an ES or ES cell-like state [205]. While methylation of Oct4 and Nanog promoters silences gene expression in somatic cells [206-208], their demethylation is associated with gene activation in ES cells derived from nuclear transfer [42, 209]. Failure to demethylate Oct4, Nanog, and other pluripotency genes is associated with an intermediate reprogramming state in induced pluripotent stem (iPS) cells [210-212]. Importantly, in these partially reprogrammed cells, 5-aza-2-deoxycytidine inhibition or siRNA knockdown of Dnmt1 led to global 5mC loss and induced a rapid and stable transition to a fully reprogrammed iPS state [212]. Similarly, by using a hypomorphic allele of Dnmt1 in donor fibroblasts, the rate of ES cell generation after nuclear transfer to the oocyte was improved [209]. Finally, initiation of reprogramming towards pluripotency in heterokaryons, generated by fusion of human fibroblasts to mouse ES cells, required 5mC loss at the Nanog and Oct4 promoters and their subsequent activation [174]. These examples show that somatic cell reprogramming is dependent on activation of key pluripotency genes through DNA methylation loss. Therefore, it is likely that global demethylation acts similarly in the pre-implantation embryo; reprogramming sperm and egg-derived genomes to allow for pluripotent gene activation.

Global DNA Demethylation in Primordial Germ Cells

At embryonic day E3.5, just prior to implantation, the blastocyst undergoes a wave of global methylation that is mediated by the de novo DNA methyltransferases Dnmt3a and Dnmt3b [27]. The majority of pluripotent epiblast cells originating in the

inner cell mass develop high levels of global methylation prior to E7.25, which is then maintained during lineage specification into embryonic and adult somatic cells, with only the pattern at specific loci being altered depending on the particular cell type. However, at E7.25 a subset of <50 epiblast cells become committed to the primordial germ cell (PGC) lineage, i.e. cells that can only differentiate into eggs or sperm [213, 214]. The PGCs proliferate and migrate into the genital ridges between E10.5 and E11.5, then continue to proliferate until E13.5, making approximately 10 cell divisions in total. PGCs will erase the vast majority of their DNA methylation marks, a process that begins as early as E8.5 and continues through E13.5 (Figure 1.3) [214-218].

Using several complementary assays, early work established that PGC lose methylation at a variety of sequence types, including single copy genes, retroviral repetitive elements, and the CCpGG motif recognized by the isoschizomers MspI and HpaII [156, 157, 185]. Importantly, and in contrast to the early embryo, PGCs completely lose DNA methylation within imprinted regions [214, 215, 219]. Additionally, the magnitude of global methylation loss is greater than that seen in the early embryo. For example, PGCs lost 74% and 55% of initial methylation at LINE and IAP regions, respectively, compared to only 57% and 36% in the early embryo [185]. Recently, these observations have been extended using whole-genome bisulfite sequencing of male and female PGCs [175]. Both male and female PGCs were drastically hypomethylated, showing a median methylation level of 16.3% and 7.8%, respectively, compared to 70-75% for the entire embryo or ES cells. Methylation loss was observed within introns,

intergenic regions, repeats, exons, and promoters, confirming that PGC demethylation is truly global.

The mechanism of 5mC loss in PGCs is largely unknown, but evidence exists to support both a passive and/or active mechanism. In support of the passive mechanism, IF suggests Dnmt1 is briefly down-regulated at E7.25, but then re-expressed at E8.5 [218]. In addition, mRNA for the Dnmt1 co-factor Uhrf1 is also transiently down-regulated during this period [220]. However, other studies show Dnmt1 mRNA level is constant throughout early PGC specification and found at equal or greater levels than in surrounding somatic cells [220, 221]. Furthermore, IF demonstrates Dnmt1 protein is highly expressed from E8.5 to 12.5 [214, 218]. Given the brevity and timing of Dnmt1 and Uhrf1 down-regulation, these events cannot account for the majority of demethylation. Experiments utilizing aphidicolin or other cell cycle inhibiting drug to prevent DNA replication during PGC development are yet to be performed. Therefore, 5mC loss through a passive, DNA replication-dependent mechanism cannot be formally ruled out.

Several lines of evidence suggest active demethylation in PGCs. First, methylation is lost despite the presence of Dnmt1 [153]. Second, BrdU incorporation and 5mC-IF has shown that from E7.75 to E9.0 the majority of PGCs (>85%) are in G1 or G2/M phases of the cycle, yet 5mC IF staining decreases during this time [218, 222]. Third, bisulfite sequencing of E12.5 and E13.5 PGCs showed an almost complete loss of 5mC at imprinted and single-copy genes [214]. Since the doubling time of cells at this

stage is about 14-16 hours, the extent of methylation loss cannot be accounted for by a passive mechanism alone [213, 223].

The mechanism for active demethylation in PGCs is not known. However, activation of the BER pathway correlates with the dramatic loss in methylation occurring between E11.5 and E13.5 [197]. One possible trigger for BER is deamination of 5mC by Aidca or Apobec resulting in a T/G mismatch. In a genome-wide bisulfite sequencing experiment, knockout of Aicda resulted in only a minor effect on global methylation levels in E13.5 PGCs [175]. While 5mC deamination by Aidca does not appear to play a major role in PGC demethylation, other deaminases such as the Apobec family might be acting redundantly. The detection of Tet1 and Tet2 in E11.5 PGCs raises the possibility that 5mC might be first modified to 5hmC prior to removal by an active repair-coupled, or even passive DNA replication-dependent, pathway that requires a 5hmC intermediate [197]. Whether demethylation occurs by an active or passive mechanism, re-methylation is thought to be prevented by the down-regulation of Dnmt3a and the down-regulation and nuclear exclusion of Dnmt3b, in the early stages of PGC specification [214, 218, 220, 221].

A critical function of methylation loss in PGCs is the removal of parental imprints, so that new imprints can be established according to the sex of the developing embryo [213-215, 219, 224]. DNA methylation, when located at DMRs near or within imprinted genes, regulates expression of the imprinted gene from either the paternal or maternal allele. This parental-origin-specific monoallelic expression makes the presence of both paternal and maternal genomes a requirement for development. In the female

germline, imprints are re-established in the postnatal ovary during the growth phase of meiotic prophase I, while male germline imprints are progressively re-established in PGCs after E14.5 and are complete in the newborn [225]. Imprinting is coincident with global methylation and both of these events are dependent on Dnmt3a, Dnmt3b, and Dnmt3L [225-227].

While lacking supporting evidence, global methylation loss in the PGCs, like that of the early embryo, could be a way to limit aberrant epiallele inheritance [203]. Another proposed role is the activation of germ-line specific genes. In support of this function, induction of Ddx4, Sycp3, Ant4, and Dazl occurs in PGC at E11.5-E13.5 and is coincident with DNA hypomethylation of flanking DNA, while in somatic cells these genes are hypermethylated and not expressed [228-230]. Ddx4, Sycp3 and Dazl are prematurely expressed in E9.5 embryos derived from Dnmt1 knockouts or hypomorphs [229]. 5-aza-2-deoxycytidine treatment of fibroblasts induces Dazl expression [228]. Furthermore, *in vitro* reporter assays show that methylation of Ant4 and Dazl promoter regions suppresses their transcription [228, 230]. DNA demethylation has been linked to germ-line specific activation of other genes as well [231-233].

Global DNA Demethylation in Cancer

Cancer-related global loss of 5mC was first observed in chemically induced hepatocellular carcinomas in rats [234]. Shortly thereafter, two seminal studies established global demethylation as a common occurrence in primary human tumors; the first compared 5mC in colon and lung cancers to matched adjacent tissue at three gene loci unrelated to cancer [235]; the second measured total 5mC by liquid chromatography

in a wide variety of benign, malignant and metastatic cancers and normal tissues [236]. Over the last 28 years, an abundance of studies have established that global hypomethylation is a hallmark of many cancers (Figure 1.4) [237, 238].

The magnitude of 5mC loss during oncogenesis differs by the type of cancer, ranging between 20-50% [237]. For example, relative to matched controls, breast and hepatocellular carcinomas lose up to 50% of methylation [239, 240], while colorectal neoplasias only lose 10-20% of total 5mC [241-243]. Therefore, the overall 5mC loss in cancer is similar to that of the early embryo, but not as severe as in PGC development. Similar to the early embryo and PGCs, hypomethylation in cancer is observed at an extensive variety of sequence elements including cancer-linked and non-cancer related gene loci, centromeric satellite DNA, and retroelements such as LINEs, short interspersed repetitive elements (SINEs), and long-terminal repeats (LTRs) [237, 238, 244]. Unbiased, genome-wide DNA methylation profiling of cancer show results consistent with these earlier findings [43, 245, 246].

DNA methylation loss appears to play a causal role in cancer development. Reduction of methylation in Dnmt1 hypomorphic mice led to development of T-cell lymphomas at 4 to 8 months of age that harbored centromeric hypomethylation and a single-copy gain in chromosome 15 [247]. Similarly, DNA demethylation resulting from the introduction of a hypomorphic Dnmt1 allele into Apc$^{Min/+}$ mice increased the formation of intestinal microadenomas and hepatocellular carcinoma, which correlated with an elevated loss of the wild-type Apc allele [248]. When the Dnmt1 hypomorph was introduced into Nf1$^{+/-}$ p53$^{+/-}$ mice, soft tissue sarcomas occurred at an earlier age by

promoting the loss of heterozygosity (LOH) required for tumor development [249].

Finally, Dnmt1$^{-/-}$ ES cells showed an increase in Hprt and tk viral transgene mutation rate

due to genomic rearrangements or deletion. These observations suggest that

hypomethylation promotes carcinogenesis by decreasing genomic stability and is

consistent with a model where DNA methylation suppresses the potentially destabilizing

effect of recombination at repetitive and transposable elements [8, 250-252].

Hypomethylation of retroelements also leads to their activation and transposition,

which could result in perturbed gene expression at the site of insertion [8, 253, 254]. For

example, disruption of gene expression by LINE transposition to the Myc and Apc loci is

found in breast and colon cancer, respectively [255, 256]. Furthermore, demethylation

and activation of retroelements are correlated events in cancer [257-260]. Direct evidence

of this mechanism has been demonstrated in the Dnmt1 hypomorph-induced lymphomas

described above, where many were found to contain an IAP somatic insertion in the

middle of the Notch1 genomic locus, leading to an oncogenic form of Notch1 [261].

In addition to repetitive elements, single copy gene loci also exhibit

hypomethylation-linked expression of genes that contribute to malignant progression

[237, 238]. For example, activation of synuclein-γ, which is highly expressed in breast

carcinomas and can induce breast cancer cells to proliferate and metastasize, is correlated

to methylation loss at a CpG island within exon 1 in a variety of cancers [262, 263].

Furthermore, treatment of synuclein-γ negative ovarian and breast cancer cells with 5-

aza-2-deoxycytidine resulted in removal of exon 1 methylation and gene activation [262].

Imprinted gene expression is also altered by hypomethylation. By genetically eliminating,

then re-establishing Dnmt1 expression, ES cells were generated that had normal global DNA methylation but lacked 5mC associated with imprinting [264]. Significantly, adult chimeric mice derived from these cells developed tumors in multiple tissues and injected embryonic fibroblasts formed tumors in recipient mice. In colorectal cancer, loss of imprinting of Igf2, an important autocrine growth factor in cancer, was linked to hypomethylation of a differentially methylated region on the normally silent allele [265].

It is notable that global DNA methylation loss occurs despite coincident CpG island hypermethylation, and subsequent silencing, of tumor-suppressor genes. Hypermethylation has been shown to affect genes involved in cell cycle control, invasion, DNA repair, cell signaling, and apoptosis [266]. For example, abnormal CpG island methylation in the promoter region of the VHL tumor-suppressor was associated with its silencing in renal carcinomas [267]. Similarly, E-cadherin and BRCA1 gene silencing has been linked to aberrant DNA methylation in sporadic breast carcinomas [268-271]. The cause of CpG island hypermethylation, and whether CpG island hypermethylation and global hypomethylation are mechanistically linked, remains unknown.

Despite a causal role in disease, the mechanism of demethylation in any cancer, whether passive or active, is largely unknown. A role for putative active demethylation genes such as Gadd45a, Mbd4, Tdg, Aicda, Apobec and Tet family members has not been reported. Another proposed mechanism would act through a cellular deficiency in the methyl donor SAM, which could induce Dnmt3-mediated active demethylation or simply reduce maintenance methylation activity [176, 272]. DNA demethylation is known to be an early event during carcinogenesis that often escalates with disease

progression [113, 239, 240, 248, 273-275]. This slow progression, and incomplete loss of methylation, suggests a passive mechanism requiring many rounds of cell division. Yet, there is no evidence of a link between DNA methyltransferase abundance and genome hypomethylation in cancer [238]. In fact, in many cancers the level of Dnmt1 and Dnmt3 increase [237, 240, 243, 276-279], likely due to an increase in cellular proliferation [243, 276].

Only very recently has evidence of a mechanism for global DNA methylation loss in cancer been presented. In glioma, decreased methyltransferase activity resulted from decreased recruitment of Dnmt1 to DNA, even though Dnmt1 protein levels remained constant [113]. This defect was mediated by Dnmt1 phosphorylation which prevented its interaction with Pcna and Uhrf1. Clearly, the role of DNA methyltransferases, replication machinery components, and their post-translational modifications, as well as the involvement of cell proliferation in cancer-linked global hypomethylation is an important area of future study.

Figure 1.1. Definitive erythropoiesis in the fetal liver.

Flow cytometric profile of freshly isolated E14.5 fetal liver stained with fluorescent antibodies against CD71 and Ter119 then detected by flow cytometry (left panel). Fetal liver cells may be classified into 6 subsets, each at an increasingly mature stage of differentiation (S0-S5). Examples of cells from each of subsets S0 to S4/5 (right panel). Erythroid maturation is associated with decreasing cell size, nuclear condensation and positive staining for hemoglobin with diaminobenzidine (brown). Cell density in the flow-cytometric profile is represented by color, from highest (red) to lowest (blue).
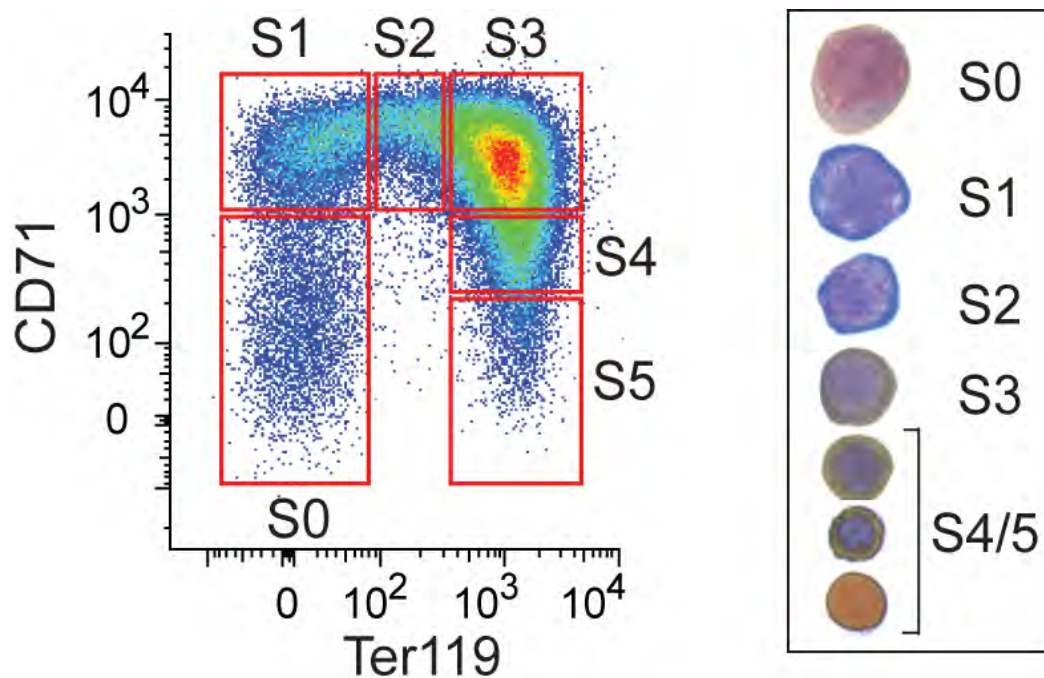
Figure 1.2. General mechanisms of DNA demethylation.

       A. Passive demethylation. Following DNA replication, the nascent strand is methylated to maintain the pattern of the parental strand. Inhibition of this „maintenance" methylation leads to a passive loss of up to 50% of 5mC per round of DNA replication.

       B. Active demethylation. Methylation loss can occur in the absence of DNA replication through a variety of proposed mechanisms. Proteins implicated in each mechanism are listed.
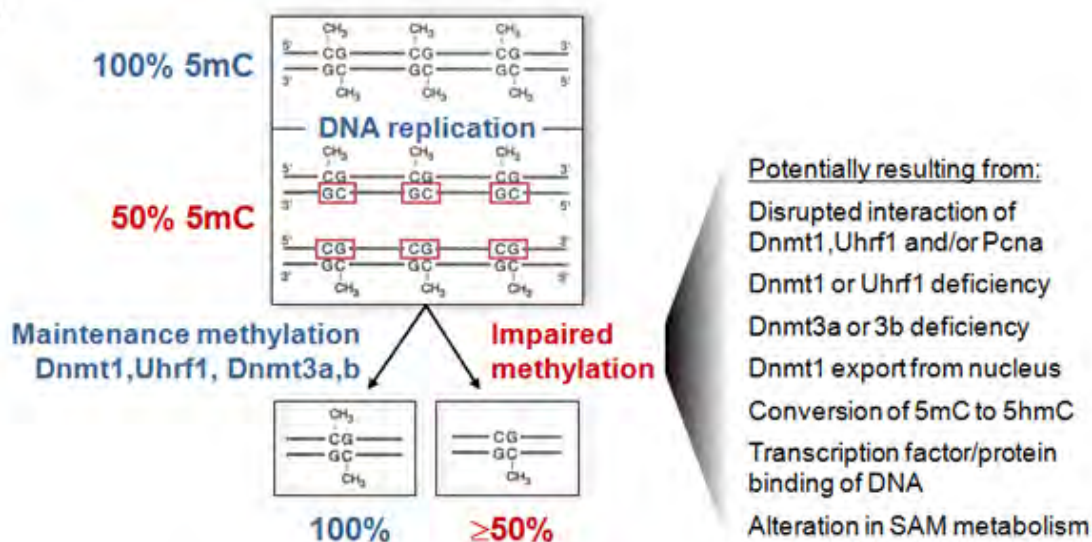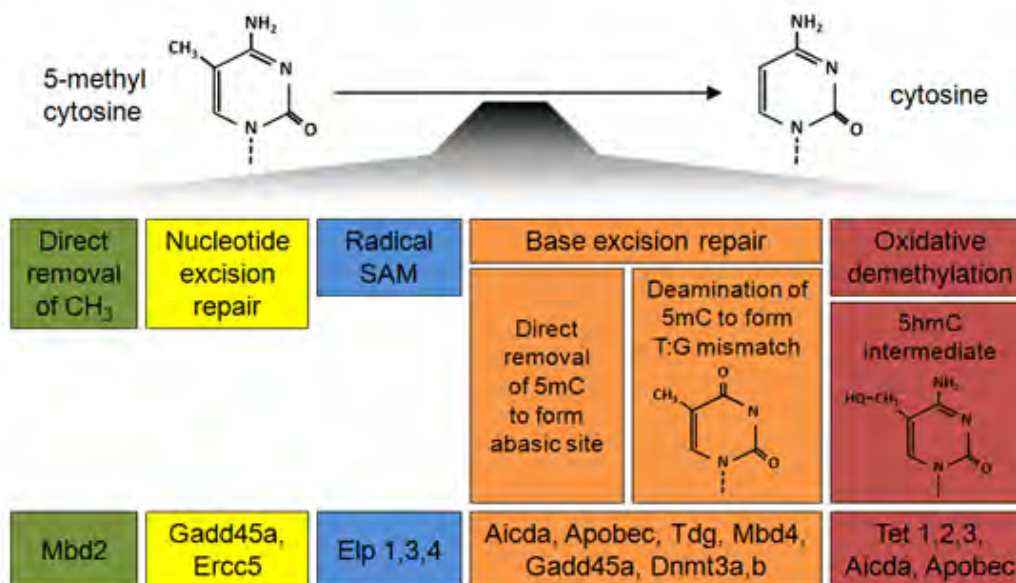
Figure 1.3. Genome-wide demethylation in mammalian development.

Global DNA demethylation has been documented at only two specific times during the mammalian lifespan. The first wave of demethylation occurs in the pre-implantation embryo following fertilization. The male and female pronuclei demethylate via an active and passive mechanism, respectively. The second wave of demethylation occurs during primordial germ cell specification, potentially through an active mechanism. Sex-specific genomes (♂= male, ♀= female) undergo demethylation and methylation at different times. While DNA methylation is known to be altered at gene-specific loci within somatic lineages, global methylation levels are thought to remain stable (black line). Adapted by permission from Macmillan Publishers Ltd: Nature, May 24;447(7143):425-32, copyright 2007.
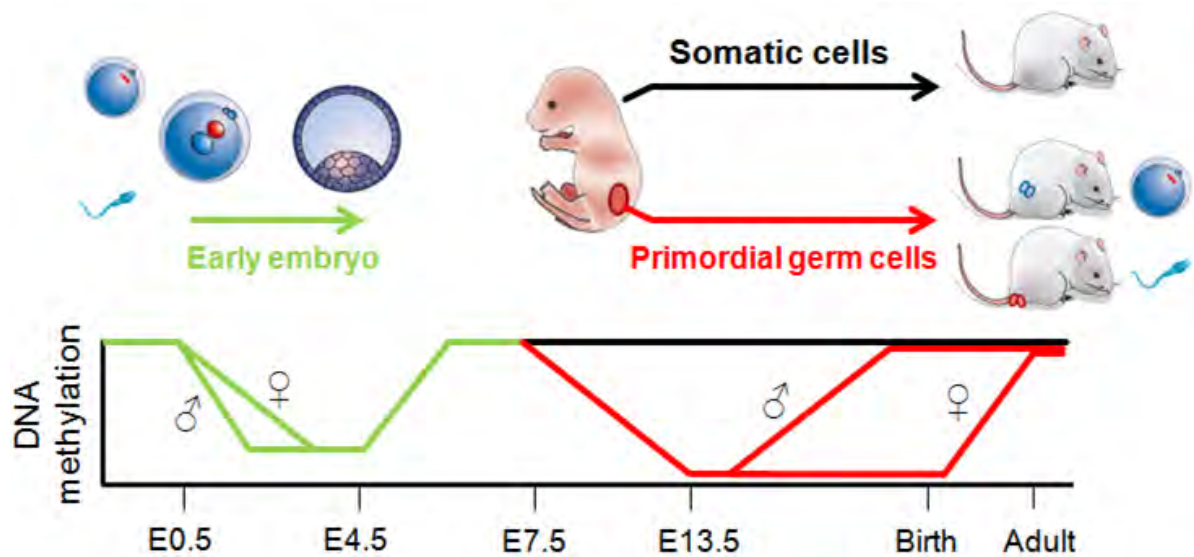
Figure 1.4. Genome-wide demethylation is observed in many cancers.

A multistage model of carcinogenesis in skin is shown [274]. In conjunction with phenotypic cellular changes and the accumulation of genetic defects, there is a progressive loss of total DNA methylation content in the development of the disease. The mechanism(s) of methylation loss is cancer is unknown. Used by permission from New England Journal of Medicine, Mar 13;358(11):1148-59, copyright 2008.
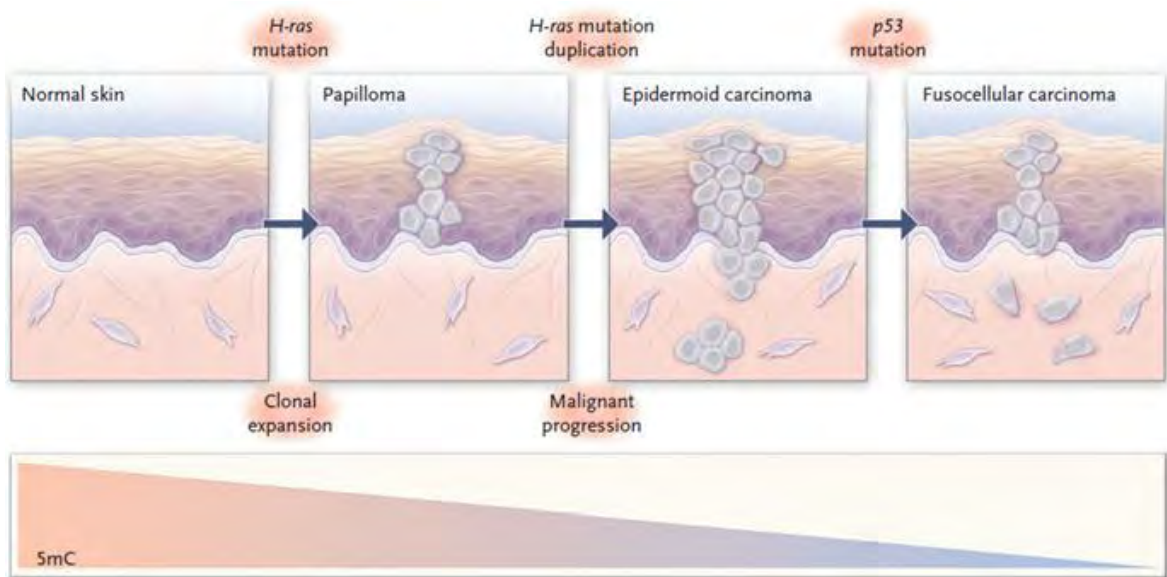
CHAPTER II

A Key Commitment Step in Erythropoiesis is Synchronized with the Cell Cycle Clock

through Mutual Inhibition between PU.1 and S-Phase Progression

Statement of Contribution

The work presented in Chapter 2 was a collaboration with Ramona Pop, a post-doctoral fellow in the Socolovsky laboratory at University of Massachusetts Medical School. My primary contribution was to the DNA methylation studies presented in Figure 2.6 and described in the results section "The Transition from S0 to S1 Coincides with S-Phase-Dependent DNA Demethylation at the β-Globin LCR" and in discussion section "An all-or-none switch in chromatin state at the β-globin LCR". I also contributed to Figure 2.5A and Figure 2.S3A-B.

This work was published in the journal PLoS Biology in September of 2010:

Ramona Pop, Jeffrey R. Shearstone, Qichang Shen, Ying Liu, Kelly Hallstrom, Miroslav Koulnis, Joost Gribnau, and Merav Socolovsky. "A Key Commitment Step in Erythropoiesis is Synchronized with the Cell Cycle Clock through Mutual Inhibition between PU.1 and S-Phase Progression". PLoS Biol (2010) 8(9) e1000484.

Abstract

Hematopoietic progenitors undergo differentiation while navigating several cell division cycles, but it is unknown whether these two processes are coupled. We addressed this question by studying erythropoiesis in mouse fetal liver in vivo. We found that the initial up-regulation of cell surface CD71 identifies developmentally matched erythroblasts that are tightly synchronized in S-phase. We show that DNA replication within this but not subsequent cycles is required for a differentiation switch comprising rapid and simultaneous committal transitions whose precise timing was previously unknown. These include the onset of erythropoietin dependence, activation of the erythroid master transcriptional regulator GATA-1, and a switch to an active chromatin conformation at the β-globin locus. Specifically, S-phase progression is required for the formation of DNase I hypersensitive sites and for DNA demethylation at this locus. Mechanistically, we show that S-phase progression during this key committal step is dependent on down-regulation of the cyclin-dependent kinase $p57^{KIP2}$ and in turn causes the down-regulation of PU.1, an antagonist of GATA-1 function. These findings therefore highlight a novel role for a cyclin-dependent kinase inhibitor in differentiation, distinct to their known function in cell cycle exit. Furthermore, we show that a novel, mutual inhibition between PU.1 expression and S-phase progression provides a „„synchromesh""'mechanism that „„locks""'the erythroid differentiation program to the cell cycle clock, ensuring precise coordination of critical differentiation events.

Introduction

Hematopoietic progenitors execute a cell division program in parallel with a differentiation program in which lineage choice is followed by lineage-specific gene expression. In many differentiation models, cell cycle exit, driven by cyclin-dependent kinase inhibitors (CDKI), is a prerequisite for terminal differentiation, establishing a key interaction between the cell cycle and differentiation programs [280-282]. However, it is unclear how the cell cycle and differentiation programs might be linked prior to cell cycle exit. Such links are presumably required to ensure the correct number of differentiated progeny. In addition, it has been speculated that the reconfiguration of chromatin at sites of lineage-specific genes, a necessary step preceding lineage-specific gene expression, may be innately dependent on DNA replication [283, 284]. An intriguing possibility is that the clockwork-like mechanisms regulating orderly cell cycle transitions may also be used, in the context of differentiating cells, to coordinate key steps in differentiation.

Here we studied differentiation of the enucleated red blood cell lineage, which first arises from hematopoietic stem cells in the fetal liver on embryonic day 11 (E11). It replaces a transient, nucleated yolk-sac erythrocyte lineage and persists throughout life. Although many of the committal events that lead to the erythroid phenotype are known, their precise timing in erythroid differentiation, and the manner in which they are coordinated with each other and/or with the cell cycle machinery, is poorly understood. Thus, survival of erythroid progenitors requires both the hormone erythropoietin (Epo), and its receptor, EpoR, a class I cytokine receptor expressed by erythroid progenitors [285]. However, the precise time in erythroid differentiation when progenitors become

dependent on Epo had not been defined. The master transcriptional regulator GATA-1 is responsible for the erythroid gene expression profile, in combination with a number of additional transcriptional regulators, including FOG-1, EKLF, SCL/Tal-1, LMO2, Ldb1, E2A, and Zbtb7a [286-289]. Though GATA-1 functional activation must precede erythroid gene induction, its precise timing in primary differentiating progenitors is not known. GATA-1 functions are antagonized by PU.1, an Ets transcription factor that acts as a master regulator in the myeloid and B-cell lineages. The mutual inhibition between PU.1 and GATA-1 is thought to underlie cell fate choice in multipotential progenitors [290-293]. PU.1 has been implicated in erythroleukemia [292, 294], but its physiological function in erythropoiesis is not known.

Erythroid gene induction by GATA-1 requires an „open chromatin‟ conformation in the vicinity of erythroid-specific genes. The erythroid-specific β-globin locus is one of the best studied models of lineage-specific gene expression [295, 296]. The active locus is characterized by early replication during S-phase, higher sensitivity to DNase I digestion, low levels of DNA methylation, and post-translational histone tail modifications associated with actively transcribed genes. Conversely, the same locus in non-erythroid cells is DNase I resistant, replicates late in S-phase, and contains histone tail modifications characteristic of silent chromatin.

In spite of the detailed knowledge contrasting chromatin states in erythroid cells with non-erythroid cells, the precise time during erythroid differentiation when chromatin reconfiguration occurs is not known. Furthermore, it is not known whether this reconfiguration involves a number of sequential stepwise alterations occurring over a

number of cell cycles/differentiation stages or whether the many changes entailed in chromatin activation occur simultaneously.

Here we studied erythroid differentiation using a flow-cytometric assay that identifies sequential stages in erythroid differentiation directly within primary hematopoietic tissue. We found that in mouse fetal liver in vivo, up-regulation of CD71 marks cells that are synchronized in S-phase of a single cell cycle, corresponding to the last generation of erythroid colony-forming cells, approximately three cell cycles prior to terminal cell cycle exit. A number of differentiation milestones, whose precise timing in erythroid development was previously unknown, occur during early S-phase of this cycle. These include the onset of Epo dependence, activation of GATA-1 function, and the opening up of chromatin at the β-globin locus. We show that S-phase progression during this specific cell cycle is dependent on down-regulation of $p57^{KIP2}$ and is required for execution of these differentiation milestones, including the reconfiguration of chromatin at the β-globin locus. Further, this S-phase dependent rapid differentiation transition is regulated by PU.1 through a newly identified, mutual antagonism between S-phase progression and PU.1 expression that coordinates the precise locking of the differentiation program to the cell cycle clock as cells enter a terminal differentiation phase.

Results

Up-regulation of cell-Surface CD71 marks the onset of EpoR dependence in erythroid progenitors

Mouse fetal liver between E11 and E15 is primarily an erythropoietic tissue. Cell surface markers CD71 and Ter119 may be used to identify differentiation-stage specific subsets, directly in primary tissue [104, 297, 298]. Here we divided freshly harvested fetal liver cells into six CD71/Ter119 subsets that we termed S0 to S5 and that form a developmental sequence (Figure 2.1A). Cells isolated from subsets S1 to S5 show morphological features characteristic of erythroid maturation, including decreasing cell and nuclear size, nuclear condensation, and hemoglobin expression (Figure 2.1A, right panel). The precise proportion of fetal liver cells within each of the CD71/Ter119 subsets is a function of embryonic age, with the majority of cells being in the early, S0 and S1 subsets in E12. The more mature, S3 to S5 subsets are gradually populated with cells during subsequent embryonic days (E13 to E15) [104].

The EpoR$^{-/-}$ fetal liver is small and lacks morphologically identifiable hemoglobinized erythroblasts of the enucleated (definitive) lineage [285]. Here we found that EpoR$^{-/-}$ fetal liver does not contain subsets S1 to S5 (Figure 2.1B). This suggested that in the definitive erythropoietic lineage that gives rise to adult-type enucleated red cells, EpoR becomes essential on or prior to the transition from S0 to S1; subsets S1 to S5 are composed almost entirely of Epo-dependent erythroblasts. Of note, the small number (<5%) of Ter119^{+} cells in the EpoR$^{-/-}$ fetal liver are all nucleated erythrocytes of the transient yolk-sac (primitive) lineage (Figure 2.S1A).

The majority of S0 cells are erythroid progenitors at the CFU-e stage

Erythroid progenitors have traditionally been identified by their in vitro colony-forming potential. "Colony forming unit-erythroid" (CFU-e) are defined as cells that give rise to colonies containing 8 to 32 hemoglobinized cells after 2–3 days of in vitro culture in Epo [299]. We investigated the colony-forming potential of cells sorted from each of the S0 to S3 subsets (Figure 2.1C). CFU-e potential was exclusive to S0 and S1 and was lost with the transition to S2. Cells in S2 and S3 gave rise to small, 2 to 4 cell clusters (Figure 2.1C, right panel).

The frequency of CFU-e obtained from sorted S0 cells was 65%–70% of the frequency from sorted S1 (Figure 2.1C). S1 consists entirely of Epo-dependent cells of similar maturation, with CFU-e potential (Figure 2.1A–C). Assuming similar plating efficiency for sorted S0 and S1 (of <30%, Figure 2.1C), this suggested that CFU-e make up 65%–70% of the S0 subset. This is in agreement with our finding that fetal liver cells expressing non-erythroid lineage markers, which were limited to S0, formed up to 30% of this subset (Figure 2.1D, Figure 2.S1B). Non-erythroid colony-forming progenitors were also restricted to S0, where they formed less than 5% of all colony-forming cells (Figure 2.1C). Our conclusion that 65%–70% of S0 cells are CFU-e was further supported by single cell RT-PCR, which showed that 68% of S0 cells expressed EpoR mRNA (Figure 2.S1C). In all the experiments that follow, "S0" refers to S0 cells from which cells expressing non-erythroid markers were excluded by flow-cytometric gating or sorting.

S1 cells are synchronized in S-phase of a single cell cycle

To examine the cell cycle status of erythroid subsets S0 to S5 in vivo, we injected pregnant female mice with the nucleotide analogue bromodeoxyuridine (BrdU) and harvested fetal livers 30 min post-injection. We sorted cells from each of S0 to S5 and stained them with antibodies directed at BrdU (Figure 2.1E, F). Cells that incorporated BrdU were in S-phase of the cell cycle at the time of harvesting. Subsets S4 to S5 showed a rapid decline in the number of S-phase cells, consistent with cell cycle exit of terminally differentiating cells. Unexpectedly, we noted that <90% of S1 cells were in S phase, as compared with <50% of cells in S0 (Figure 2.1E, F). In addition, the intensity of the BrdU fluorescence within S1 cells was approximately 50% higher than in S0, suggesting a higher rate of DNA synthesis (Figure 2.S1D). Similar experiments with EpoR$^{-/-}$ fetal liver showed that EpoR appears to have no effect on progenitor cell cycle status (Figure 2.S1E).

Consistent with the higher number of S-phase cells in S1, we found a corresponding increase in the E cyclins in S1 compared with S0 (Figure 2.1G). Strikingly, we noted >30-fold decrease in the CDKI p57^{KIP2} mRNA, but no significant change in the mRNA of other members of the CIP/KIP CDKI family; there was induction in p27^{KIP1} later in differentiation, in subsets S2 and S3 (Figure 2.1G and Figure 2.S1F). [300, 301]. The p57^{KIP2} protein also decreased at the S0 to S1 transition (Figure 2.1G lower panel).

The finding that nearly all S1 cells were in S-phase could be due to an unusual cell division cycle with short or no gap phases. Alternatively, S1 cells may be

synchronized in S-phase of the cycle. The latter explanation would require that cells spend only a brief period of a few hours in S1, lasting through part or all of a single S-phase. The preceding G1 phase of this same cell cycle would have occurred prior to the transition from S0 to S1. The G2 and M-phases of this same cycle would occur as cells upregulate Ter119 and transition into S2.

To investigate these possibilities, we isolated S0 cells by flow cytometry, labeled them with the cell-tracking dye carboxyfluorescein diacetate succinimidyl ester (CFSE), and followed their Epo dependent differentiation into S1 in vitro (Figure 2.1H). By 10 h, 53% of S0 cells transitioned into S1 in the absence of cell division, as indicated by a single CFSE peak for S1 (solid red histogram, t=10 h) that was identical in intensity to that of the CFSE peak for S0 (blue histogram, t=10 h; median CFSE fluorescence for both S1 and S0 peaks=4,400). This suggested that the transition from S0 to S1 occurred in the absence of cell division, within a single cell cycle. Four hours later, at t=14 h, essentially all S1 cells had divided once, as indicated by the halving of the CFSE signal (red histogram at t=14 h, CFSE fluorescence=2,100). The simultaneous division of S1 cells suggested they were synchronized in their cell cycle phase. By contrast, only a portion of S0 cells, which were presumably asynchronous in their cell cycle phase, had divided at this time, resulting in a biphasic CFSE peak (blue histogram, t=14 h).

Taken together, these results suggest that the most mature CFU-e progenitor ("CFU-e.2", Figure 2.1I), capable of giving rise to an eight-cell colony, traverses S0, S1, and enters S2 within a single cell cycle. This progenitor arises in S0, becomes Epo dependent, and upregulates CD71, transitioning into S1 during S-phase of its cell cycle.

Up-regulation of Ter119 occurs at approximately the same time that it completes its cycle and divides, giving rise to progeny that lack CFU-e activity in S2 (Figure 2.1C). These conclusions are consistent with essentially all S1 cells being in S-phase (Figure 2.1E, F), and with our finding that nearly all S1 cells are sensitive to hydroxyurea, a drug that specifically targets S-phase cells (Figure 2.S2A). These conclusions are consistent with a number of other observations: the loss of CFU-e activity with Ter119 expression (Figure 2.1C, [302]), the short time span (<15 h) that freshly sorted S0 cells require to transition through S1 and into S2 (compare with an estimated cell cycle length of 16 h for a CFU-e cell that will undergo three cell divisions in 48 h, giving rise to an eight cell colony), and with early work suggesting that Epo dependence first occurs in early S-phase of a specific CFU-e cell generation [303]. These conclusions are also consistent with the finding that EpoR$^{-/-}$ embryos have normal numbers of CFU-e [285]: though EpoR$^{-/-}$ embryos lack S1 cells, all the CFU-e in S1 first arise as Epo-independent cells in S0, where they are presumably retained in the EpoR$^{-/-}$ fetal liver.

S-Phase progression is required for the transition from S0 to S1

There are two ways to explain how up-regulation of CD71, a differentiation event, might coincide with S-phase, a cell cycle event. These events may have each been initiated in parallel by a common upstream regulator, such as the EpoR, since both occur at the time that cells become EpoR dependent. Alternatively, there may be a direct mechanistic link between the differentiation and cell cycle programs. To distinguish these possibilities, we examined whether a block to S-phase progression would interfere with CD71 up-regulation (Figure 2.2). We incubated sorted S0 cells in vitro for 10 h in the

presence of Epo, and either in the presence or absence of aphidicolin, an inhibitor of DNA polymerase that arrests S-phase progression [304]. At t=10 h, cells were washed free of aphidicolin and incubated in Epo alone for an additional 10 h (Figure 2.2A). In the initial 10 h of incubation, there was an Epo-dependent transition of cells from S0 to S1 (Figure 2.2C, rows 1 and 5). However, the presence of aphidicolin blocked this transition (Figure 2.2C, rows 2 & 3, t=10 h). Both S-phase and the transition into S1 resumed once the cells were washed free of aphidicolin (Figure 2.2C, rows 2 & 3, t=20 h). These observations suggested that the transition from S0 to S1 occurred during S-phase and required both Epo and S-phase progression.

We also examined the effect of mimosine, a plant amino acid that blocks cell cycle progression in late G1 [305]. We incubated sorted S0 cells in Epo and in the presence or absence of mimosine. By 4 h of incubation, the majority of cells were arrested in G1. However, a small fraction of cells (12%) could be seen in S-phase at t=4 h (Figure 2.2C, row 4, BrdU/7AAD at t =4 h). Presumably, at the time mimosine was added, these cells were advanced in their cell cycle beyond the point at which mimosine exerts its block. BrdU/7AAD analysis showed that these cells were in the early half of S-phase and expressed the highest CD71 levels within the S0 subset (Figure 2.2D, cells marked in red). By t=10 h, no S-phase cells were seen in S0, presumably because they have now transitioned into S1, where a similar number of cells (15%) had newly appeared (Figure 2.2C, row 4, BrdU/7AAD for S0 at t=10 h, and CD71/Ter119 for S1 at t=10 h). These observations were consistent with the onset of CD71 up-regulation

occurring in early S-phase in S0, culminating in the transition to S1 later within that same S-phase.

CD71, the transferrin receptor, is required during erythroid differentiation in order to facilitate cellular uptake of iron for hemoglobin synthesis. CD71 is also expressed, albeit at lower levels, on all cycling cells. We therefore examined whether, in the context of S1 cells, CD71 might be required specifically for S-phase progression. We used RNAi to prevent CD71 up-regulation in S0 cells during their incubation in Epo (Figure 2.S2B,C). The failure of these cells to upregulate CD71 did not interfere with the number of cells in S-phase (Figure 2.S2B). Therefore, the link between S-phase progression and CD71 up-regulation in S1 cells is not due to a cell cycle function for this gene.

The S0 to S1 transition is marked by down-regulation of PU.1 and GATA-2 and precedes induction of erythroid-specific genes

To investigate the link between S-phase and the erythroid differentiation program, we examined expression of erythroid transcriptional regulators and erythroid-specific genes in freshly sorted fetal liver subsets and in fetal brain (Figure 2.3A). We found that the GATA-1 mRNA was present in S0 cells, at 200-fold higher levels than in fetal brain (Figure 2.3A) and 40-fold higher level than in Mac-1$^+$ cells (Figure 2.S3A). It increased a further 2-fold with the transition from S0 into S1 and continued to increase in S2 and S3. Of note, total RNA per cell decreased 4-fold over the course of differentiation from S2 to S4 (Figure 2.S3B), suggesting an overall modest increase in GATA-1 mRNA per cell over this period. Other erythroid transcriptional activators and GATA-1 associated factors, including EKLF, NF-E2 [306], SCL/Tal-1, and Lmo2, showed a similar

expression pattern to that of GATA-1 (Figure 2.3A). Therefore, expression of GATA-1 and of other activators of the erythroid transcriptional program precedes the transition from S0 to S1. By contrast, we found that PU.1, a repressor of GATA-1 function, and GATA-2, a target of GATA-1-mediated repression [307], were both down-regulated approximately 30-fold and 20-fold, respectively, at the S0 to S1 transition, becoming undetectable with further differentiation (Figure 2.3A). Prior to its down-regulation, the level of PU.1 in S0 cells was comparable to that of myeloid Mac-1[+] cells (Figure 2.S3A). PU.1 protein levels also declined with the transition from S0 to S1 (Figure 2.S3C).

EpoR$^{-/-}$ fetal liver cells, though apparently arrested at the S0 stage (Figure 2.1B), have a similar expression pattern of transcriptional regulators to wild-type S1 (Figure 2.3A). Therefore, down-regulation of PU.1 and GATA-2 at the S0 to S1 transition, as well as the preceding induction of GATA-1, are independent of EpoR signaling.

We examined expression of several erythroid-specific GATA-1 target genes: β-globin (Hbb-b1); the first enzyme of heme synthesis, aminolevulinic acid synthase 2 (ALAS2); and the anion exchanger Band 3 (Slc4a1), a major erythrocyte membrane protein [308]. There was a modest increase in their expression at the S0 to S1 transition, followed by a 30- to 100-fold induction during subsequent differentiation in S2 and S3 (Figure 2.3A). Expression of the EpoR gene, itself a GATA-1 target, increased 10-fold above its S0 level with the transition to S1 (Figure 2.S3D). Taken together, induction of erythroid GATA-1 target genes and repression of GATA-2 suggest that GATA-1 function is activated at the S0 to S1 transition. The modest increase in GATA-1 mRNA at this time suggests that its activation may be principally a result of PU.1 down-regulation.

S-Phase arrest at the S0 to S1 transition blocks induction of erythroid-specific genes

We had found that S-phase progression at the transition from S0 to S1 was required for CD71 up-regulation (Figure 2.2). We therefore examined whether S-phase progression at this time was also required for induction of erythroid-specific genes. We cultured sorted S0 cells in Epo for 10 h, a period sufficient for 25%–50% of cells to transition into S1 (Figures 1H, 2C), and examined the effect of adding aphidicolin to the culture. Cells were then washed free of aphidicolin, continuing incubation in Epo alone. Cells incubated in Epo alone for the entire period showed 50- to 100-fold induction in the mRNAs for β-globin, Band 3, and ALAS2 (Figure 2.3B, red curves). By contrast, cells that were subject to aphidicolin treatment during the initial 10 h showed reduced mRNA induction by the end of the culture period (Figure 2.3B, blue curves). The reduced mRNA levels corresponded closely to the levels predicted had there been a 10 h delay in the time course of induction for each of the genes (Figure 2.3B, black curves). Therefore, induction of erythroid-specific genes was likely blocked during the incubation period in aphidicolin.

We also examined whether S-phase arrest interferes with erythroid gene induction if applied at the S1 stage of differentiation. We sorted S1 cells and incubated them in Epo, either in the presence or absence of aphidicolin. Unlike S0 cells, aphidicolin mediated S-phase arrest of S1 did not interfere substantially with their induction of erythroid specific genes, as shown by the unperturbed induction of β-globin, Alas2, and Band 3 (Figure 2.3C, Figure 2.S3E) or with the up-regulation of Ter119 (Figure 2.S3F). Therefore, S-phase progression is required for activation of erythroid-specific genes,

specifically at the S0 to S1 transition, but not a few hours later when the cells have traversed into S1. The lack of effect of aphidicolin on mRNA induction in S1 suggests its effects in S0 are not due to non-specific suppression of transcription.

S-Phase arrest at the S0 to S1 transition blocks down-regulation of PU.1 and GATA-2 and arrests erythroid morphological maturation

Transcripts for PU.1 and GATA-2 are markedly down-regulated at the transition from S0 to S1 (Figure 2.3A). We examined whether S-phase arrest interferes with their down-regulation. Sorted S0 cells were incubated in Epo for 4 h, at which time, just prior to their transition into S1 (Figure 2.2C), aphidicolin was added to the cultures for a period of 10 h. Cells were then washed free of aphidicolin and incubated in Epo for a further 10 h. Aphidicolin halted the down-regulation of both PU.1 and GATA-2, which resumed once the cells were washed free of the drug (Figure 2.3D,E). Similar results were obtained in cells treated with mimosine (Figure 2.S3G). Therefore, S-phase progression is required for down-regulation of PU.1 and GATA-2 at the S0 to S1 transition. Of note, GATA-1, Nfe2, and Lmo2 mRNAs, which did not change significantly during the transition from S0 to S1 (Figure 2.3A), were not altered significantly by the aphidicolin treatment (Figure 2.3E, Figure 2.S3H).

We also examined the effects of aphidicolin or mimosine treatment on morphological maturation of S0 cells cultured in Epo. Following 10 h in Epo in the presence of aphidicolin or mimosine, cells appeared larger than cells incubated in Epo alone. This suggested that, while S-phase progression and the erythroid differentiation program had both arrested, cell growth was not perturbed (Figure 2.3F). Cells were then

washed free of aphidicolin or mimosine and cultured in Epo alone. By 20 h, erythroid

maturation had resumed in cells that were initially incubated in cell cycle blocking drugs,

as judged by decreasing cell size, nuclear condensation, and decreased nuclear to

cytoplasmic ratio, but was nevertheless delayed when compared with control cells. These

results are consistent with the effect of S-phase arrest on gene expression (Figure 2.3B,

D, E) and suggest that S-phase progression at the S0 to S1 transition is a key requirement

for activation of the erythroid differentiation program.

Preventing p57^{KIP2} down-regulation blocks S-phase progression at the S0 to S1

transition and arrests erythroid differentiation

Expression of p57^{KIP2} mRNA decreases over 30-fold at the S0 to S1 transition,

and this is associated with down-regulation of the p57^{KIP2} protein (Figure 2.1G). To

examine the effect of preventing p57^{KIP2} down-regulation, we generated a point mutant of

p57^{KIP2}, p57T329A, analogous to a proteolysis-resistant human p57^{KIP2} mutant [309].

Sorted S0 cells were infected with bicistronic retroviral vectors expressing either wild-

type p57^{KIP2} or p57T329A, linked through an internal ribosomal entry site (IRES) to a

human CD4 (hCD4) reporter; control cells were infected with retroviral vector expressing

the IRES-hCD4 construct only (MICD4). To allow expression of the transduced p57^{KIP2},

infected cells were cultured for 15 h in stem-cell factor (SCF) and interleukin 3 (IL-3),

cytokines that sustain viability of progenitors but, unlike Epo, do not support

differentiation from S0 to S1. Infected S0 cells were then transferred to Epo for 14 h

(Figure 2.3G). Expression of either wild-type (unpublished data) or mutant p57^{KIP2}, but

not expression of MICD4, resulted in a block to S-phase progression and inhibited the

transition from S0 to S1 (Figure 2.3G). Further, PU.1 mRNA was >3-fold higher in cells expressing p57^{KIP2} compared with control cells expressing vector only (Figure 2.3H), suggesting that, as in the case of aphidicolin-mediated S-phase arrest, p57^{KIP2}-mediated S-phase arrest prevents down-regulation of PU.1 at the transition from S0 to S1. Erythroid morphological maturation, but not cell growth, of p57T329A-transduced cells was also arrested (Figure 2.S3I).

Taken together, up-regulation of CD71, which defines the transition from S0 to S1, identifies a key differentiation transition within the last generation of CFU-e ("CFU-e.2", Figure 2.3I). It marks the onset of EpoR dependence and occurs exclusively during S-phase of the cell cycle. Induction of GATA-1 and other activators of the erythroid transcriptional program precede this transition, whereas induction of erythroid-specific genes such as β-globin and Ter119 follows it. The S0 to S1 transition coincides with rapid down-regulation of p57^{KIP2}, PU.1, and GATA-2. Both Epo and S-phase progression are required for up-regulation of CD71. S-phase progression at the S0 to S1 transition requires the down-regulation of p57^{KIP2} and is in turn required for the down-regulation of PU.1 and GATA-2 and the subsequent activation of erythroid-specific genes. By contrast, S-phase arrest in S1 cells does not affect erythroid gene activation (Figures 2.3C, 2.S3E–F).

Persistently elevated PU.1 arrests S-phase progression and blocks erythroid differentiation

Both PU.1 and GATA-2 were rapidly and dramatically down-regulated at the transition from S0 to S1 (Figure 2.3A, Figure 2.S3A). We examined the effect of

preventing this down-regulation by expressing either PU.1 (Figure 2.4A–D) or GATA-2 (Figures 2.4E, 2.S4C, D) in S0 cells using retroviral constructs and a similar strategy to that described above for $p57^{KIP2}$. Following infection, S0 cells were cultured for 15 h in IL-3 and SCF and then transferred to Epo for 24 h, when CD71/Ter119 and cell cycle profiles were examined (Figure 2.4A–C). We divided the PU.1 expression profile at t=24 h into 7 sequential hCD4 gates labeled (i) to (vii) (Figure 2.4A), each containing cells with increasing levels of the hCD4 reporter and, therefore, increasing levels of PU.1. By measuring PU.1 protein directly in fixed and permeabilized cells using a PU.1-specific antibody and flow-cytometry, we found that hCD4 protein expression was a reliable reporter of exogenous PU.1 protein expression in our system (Figures 2.4D, 2.S4A–B); expression of transduced PU.1 was also measured by qPCR (Figure 2.S4D). Sequential hCD4 gates were also obtained for control cells expressing the empty MICD4 vector. PU.1 expression blocked transition from S0 to S1, with the number of cells transitioning into S1 declining as PU.1 expression increased (Figure 2.4B, upper panels). PU.1 expression also resulted in a decrease in the number of S-phase cells, with cells arresting principally at the transition from G1 to S-phase, though there was also an increase in the number of cells within G2 or M (Figure 2.4B, lower panels). The decrease in the number of cells in S1 was paralleled by decreased S-phase cell number, suggesting a direct correlation between the PU.1-mediated block of the transition from S0 to S1, and its inhibitory effect on S-phase (Figure 2.4C). Therefore, PU.1 inhibits both S-phase and erythroid differentiation at the S0 to S1 transition.

Since the down-regulation of both PU.1 and p57^{KIP2} are required for S-phase

progression and for the transition from S0 to S1 (Figure 2.3G, Figure 2.4B, C), we

examined whether PU.1 may be a regulator of p57^{KIP2}. However, we found that

exogenous expression of PU.1 did not prevent down-regulation of p57^{KIP2} (Figure

2.S4E). Therefore, PU.1"s inhibitory effect on S-phase is not mediated via p57KIP2.

In contrast to PU.1, expression of GATA-2 in S0 cells did not prevent transition

into S1, though it somewhat reduced the subsequent transition from S1 to S2 (Figure

2.4E). GATA-1 overexpression in S0 cells had the opposite effect, of promoting the

transition from S1 to S2. There was no significant effect of either GATA-1 or GATA-2

on the cell cycle profile (Figure 2.4E).

The S0 to S1 transition coincides with a switch in the timing of replication of the β-globin locus

A long-standing hypothesis suggests that DNA replication may provide an

opportunity for the restructuring of chromatin at tissue specific gene loci [283, 284].

Given the requirement for DNA replication for the transition from S0 to S1, we asked

whether chromatin change may be taking place at this time. The β-globin gene locus

(Figure 2.5A) is a well-studied model of tissue-specific gene expression. The features that

characterize the open chromatin conformation at the actively transcribed locus in

erythroid cells have been established, but the time during development when the active

chromatin conformation is acquired is not known. We therefore set out to examine

whether the S0 to S1 transition might coincide with an alteration in the structure or

function of chromatin at this locus.

The timing of replication of the β-globin locus is correlated with its chromatin state. In higher eukaryotes the timing of replication of genes correlates with their transcriptional activity [310]. Housekeeping genes replicate early in S-phase, whereas silent chromatin and heterochromatin replicate late. The β-globin locus replicates in mid to late S-phase in non-erythroid cells and early in S-phase in erythroid cells [311]. We examined the timing of replication of the β-globin locus in S0 and S1 cells sorted from fresh fetal liver. Individual alleles were identified using fluorescence in situ hybridization (FISH) with a probe directed at the β-major gene. Cells in S-phase were identified by positive staining for BrdU incorporation. Nuclei from at least 100 S-phase cells from either S0 or S1 were examined in each of two experiments (Figure 2.5B). Using this approach, two single dots ("SS") suggest that neither of the β-globin alleles had yet replicated. Nuclei in which both alleles have replicated contain a pattern of two double dots ("DD"). Replication of only one allele results in one single and one double dot (SD) [311]. We found that the number of cells with a DD pattern increased from only 15% in S0 to over 50% in S1 (Figure 2.5B), suggesting a switch in the timing of replication from late to early S-phase. In addition, an average of 36% of S0 cells, but only 21% of S1, had an SD pattern, consistent with a switch from late, asynchronous replication in S0 to early, synchronous replication in S1 [311].

The S0 to S1 transition coincides with the onset of DNase I hypersensitivity at the β-globin locus control region (LCR)

A key indicator of open chromatin at the β-globin LCR is the presence of hypersensitivity (HS) sites (Figure 2.5A). We prepared nuclei from freshly sorted S0 or

S1 cells and tested their sensitivity to DNase I digestion. Following digestion, we measured remaining DNA using quantitative PCR, with amplicons within HS2, HS3, and HS4 [312]. Results were expressed as a ratio to the DNase I resistant, non-expressing neural gene, Nfm. We found that S0 cells were relatively resistant to DNase I, while S1 cells were hypersensitive at all tested HS sites (Figure 2.5C). Therefore, the S0 to S1 transition coincides with the onset of DNase I hypersensitivity at the β-globin LCR.

We also examined E12.5 EpoR$^{-/-}$ whole fetal livers, which do not contain S1 cells (Figure 2.1B). We found that EpoR$^{-/-}$ fetal livers were resistant to DNase I, whereas whole fetal livers from wild-type or heterozygous littermates showed the expected hypersensitive sites (Figure 2.5D). We therefore concluded that DNase I hypersensitivity develops at the S0 to S1 transition, synchronously with the onset of EpoR dependence.

S-phase progression is required for the onset of DNase I hypersensitivity at the β-globin LCR

Since the transition from S0 to S1 coincides with, and requires, S-phase progression, we examined whether development of DNase I hypersensitivity at the β-globin LCR also requires S-phase progression. We incubated sorted S0 cells in Epo, in the presence or absence of aphidicolin, for 10 h. Over this period 25%–50% of S0 cells transition into S1, a process arrested by aphidicolin (Figures 1H, 2C). At the end of a 10-h incubation period, nuclei were prepared and digested with varying concentrations of DNase I. There was a clear increase in DNase I sensitivity in cells incubated in Epo alone, relative to cells incubated in Epo and aphidicolin (Figure 2.5E). Therefore, the

development of DNase I hypersensitivity at the S0 to S1 transition is dependent on S-phase progression.

Changes in post-translational histone tail modifications associated with the transition from S0 to S1

The switch in timing of replication and in DNase I hypersensitivity at the S0 to S1 boundary suggested the β-globin LCR was undergoing structural changes. To investigate these, we used chromatin immunoprecipitation (ChIP) to determine specific histone tail modifications at the β-globin LCR in freshly sorted S0, S1, and in fetal brain. We used ChIP-qPCR for amplicons at the β-globin LCR HS sites, or at a control, neural gene, Nfm. Changes in histone modifications were expressed as a ratio, between S0 and either S1 or fetal brain (Figure 2.5F, G). Figure 2.5F summarizes data pooled from seven experiments with various immunoprecipitating antibodies as indicated. A comparison of S1 with S0 shows a 7-fold decrease in trimethylation of histone 3 lysine 27 (H3K27me3, $p = 0.019$, paired t test), a mark associated with silent chromatin, and a 2.5-fold increase in histone 3 lysine 4 dimethylation, a mark associated with active chromatin (H3K4me2, $p = 0.032$), at the HS2 site of the β-globin LCR. A similar trend for these two modifications was also found at other HS sites ($p = 0.0006$ and $p = 0.011$ for H3K4me2 and H3K27me3, respectively, pooling all HS sites). An increase in acetyl marks in histones H3 and H4 associated with active chromatin was also seen consistently across the HS sites tested, though it did not reach statistical significance. Of note, no significant changes in histone marks were found between S0 and S1 at the Nfm gene. Further, there

was no significant change in total histone occupancy of the HS sites between S0 and S1, as determined by ChIP with antibodies directed against total H3 and H4 (Figure 2.5F).

We noted that H3K27me3, associated with silent chromatin, and H3K4me2, associated with active chromatin, were both enriched in S0 compared with fetal brain (Figure 2.5G, lower panel). These results were suggestive of bivalent chromatin at the β-globin LCR in S0, and loss of the repressive H3K27me3 mark with transition into S1 (Figure 2.5G, upper panel, 2.5F).

The transition from S0 to S1 coincides with S-phase-dependent DNA demethylation at the β-globin LCR

We examined DNA methylation of six CpG dinucleotides, three each at the HS1 and HS2 sites of the β-globin LCR (Figures 5A, 6A). Genomic DNA was prepared from sorted hematopoietic cell subsets from fresh fetal liver, including S0, S1, megakaryocytic CD41$^+$, myeloid Mac-1$^+$, and Lin$^-$Sca1$^+$Kit$^+$ (LSK) cells, enriched for hematopoietic stem-cells. We also examined EpoR$^{-/-}$ fetal livers depleted of cells expressing lineage markers, and fetal brain. DNA methylation at each of the six CpGs was obtained following bisufite conversion of genomic DNA, PCR amplification at HS1 and HS2, and pyrosequencing. In fetal brain methylation levels were high, at <60%–80%, for all six CpG dinucleotides. Methylation levels were lower in all hematopoietic cell subsets (Figure 2.6A). Methylation levels were largely similar in all hematopoietic, Epo-independent cell subsets examined: LSK, Mac-1$^+$, CD41$^+$, S0, and EpoR$^{-/-}$ cells. The onset of Epo dependence in S1 was associated with a marked reduction in DNA

methylation in all six CpG dinucleotides, with the level of methylation dropping to virtually undetectable levels in S1 for four of the six CpGs.

We found that DNA demethylation also took place in freshly sorted S0 cells allowed to differentiate in vitro (Figure 2.6B, C). Demethylation in vitro occurred earlier at the HS1A, B, C, and HS2C than at HS2A, B (Figure 2.6B, C, red lines), in agreement with results in vivo (Figure 2.6A). Demethylation in vitro was arrested at all CpGs if either aphidicolin or mimosine were added to the incubation medium, and resumed when these drugs were removed (Figure 2.6B, C). Therefore, DNA demethylation, initiated at the transition from S0 to S1, is dependent on S-phase progression. These results are suggestive of a passive demethylation process, due to loss of maintenance methylation at nascent DNA.

PU.1 down-regulation is required for the onset of DNA demethylation at the transition from S0 to S1.

We examined HS1 and HS2 DNA methylation levels in S0 cells transduced with PU.1 (as in Figure 2.4A) and incubated in Epo for 24 h. DNA methylation was significantly higher at 3 of the 6 CpGs in S0 cells transduced with PU.1-ICD4, compared with control cells transduced with MICD4 (Figure 2.6D). Therefore, PU.1 expression, along with its inhibitory effect on erythroid differentiation, also impaired DNA demethylation, possibly due to its inhibitory effect on S-phase in these cells (Figure 2.4C).

Discussion

We have identified a committal step in erythropoiesis in which the cell cycle clock is precisely synchronized with and coordinates an erythroid differentiation switch. It takes place during S-phase of the last CFU-e generation, at the transition from S0 to S1, when S-phase progression is required for several distinct committal differentiation events, including the onset of Epo dependence, a switch in chromatin at the β-globin locus into an open conformation, and activation of GATA-1 function with consequent transcription of GATA-1 target genes (Figure 2.7A). The transition from S0 to S1 can be replicated in vitro, where sorted S0 cells develop into a differentiation state characteristic of S1 within 10 to 12 h.

S-phase progression at the S0 to S1 transition and the ensuing differentiation switch are dependent on the down-regulation of p57^{KIP2} (Figures 2.3G, 2.7B), a novel finding since, to date, the principal known role of CDKIs in differentiating cells had been to mediate terminal differentiation secondary to cell cycle exit [280-282]. Unlike other CDKIs, p57^{KIP2} is required for the development of multiple tissues [313], suggesting that its novel role in erythropoiesis, triggering S-phase progression during a committal differentiation event, may be replicated in other systems. Of note, p57^{KIP2} is the only CDKI of the CIP/KIP family to be down-regulated at the S0/S1 transition.

The synchronization of S-phase progression with several rapid and committal differentiation transitions suggests they are coregulated. A key mediator of this co-regulation is PU.1, whose expression declines at the transition from S0 to S1. We have identified a novel cross-antagonism between S-phase progression and PU.1 expression.

We show that S-phase arrest, caused by high levels of p57^{KIP2} or by cell cycle blocking

drugs, prevents down-regulation of PU.1 (Figure 2.3E, H); conversely, failure to

downregulate PU.1 arrests S-phase progression (Figure 2.4B, C). Either maneuver blocks

erythroid differentiation, including a block of chromatin reconfiguration at the β-globin

locus and blocked expression of erythroid-specific genes (Figures 2.3B, 2.3D–H, 2.S3G,

2.S3I, 2.4B, 2.5E, 2.6C–D).

      We propose that the mutual inhibition between PU.1 and S-phase progression at

the S0 to S1 transition (Figure 2.7B) simultaneously controls the transition in both

differentiation and cell cycle states. Its function is analogous to a synchromesh

mechanism in an automotive transmission, matching the speeds of two rotating gears

before allowing them to lock together during a gear-shift. The mutual antagonism

between S-phase progression and PU.1 expression ensures that PU.1 down-regulation

does not occur prior to the cell‟s entry into S-phase; conversely, S-phase entry cannot

occur before conditions for PU.1 down-regulation are in place. In this manner the cell

cycle and differentiation programs can only proceed when precisely synchronized.

      Once cells have transitioned from S0 into S1, S-phase progression is no longer

required for expression of erythroid genes (Figure 2.3C, 2.S3E–F). Therefore, the

synchromesh mechanism is specific to the transition from S0 to S1, when committal

decisions bring about an irreversible terminal differentiation phase. Our findings reveal a

key organizational feature in erythroid differentiation and have implications for

differentiation of other lineages, where similar synchronization events may occur.

Further, the synchromesh mechanism we describe may be a target in leukemogenesis,

consistent with reports that high levels of PU.1 promote erythroleukemia [294]. Similarly, although no reports at present implicate p57^{KIP2} specifically in erythropoiesis, mutations in p57^{KIP2} are implicated in the familial Beckwith-Weidemann syndrome, which predisposes to pediatric tumors [301].

The cross-antagonism between PU.1 and S-phase progression is lineage and differentiation stage-specific

PU.1, whose physiological function in erythropoiesis had not been clear, plays a pivotal role at the S0 to S1 transition, through its cross-antagonism with S-phase progression. This cross-antagonism is lineage and differentiation stage-specific, since it presumably does not operate in myeloid and B-cell lineages where PU.1 is an essential transcriptional activator. Similarly, within the erythroid lineage, this mutual antagonism must be activated specifically in the last generation of CFU-e. Its premature activation at an earlier CFU-e cycle may be predicted to result in premature transition into S1 and consequently, in a reduced number of differentiated progeny. This prediction helps explain previous observations, where erythroid cells from PU.1-null embryos were found to differentiate prematurely and to have reduced self-renewal capacity [314]. These observations are consistent with the PU.1-null phenotype mimicking premature down-regulation of PU.1. The mutual inhibition between PU.1 and S-phase may also explain findings in the T-cell lineage, where exogenous expression of PU.1 at the pro-T cell stage was found to block both thymocyte expansion and differentiation [315].

The cross-antagonism between PU.1 and GATA-1

Previous work documented cross-antagonism between PU.1 and GATA-1, showing them to interfere with each other's transcriptional functions through a variety of mechanisms including direct physical binding [290-293]. We propose that the activation of erythroid terminal differentiation at the S0/S1 boundary is due to functional activation of GATA-1 (Figure 2.7B). Though present in S0 cells prior to the transition into S1 (Figure 2.3A), GATA-1 function is inhibited by PU.1. Down-regulation of PU.1 at the S0 to S1 transition alleviates this inhibition, allowing GATA-1-mediated activation of erythroid gene induction. Among its known targets, GATA-1-mediated transcriptional repression of GATA-2 [307] would account for our observation that GATA-2 is down-regulated at the S0 to S1 transition (Figure 2.3A). This scheme places the decrease in GATA-2 downstream of the cross-antagonism between PU.1 and GATA-1 (Figure 2.7B) and explains why exogenous high levels of GATA-2, unlike PU.1, do not block the transition from S0 to S1 (Figure 2.4).

Based largely on immortalized progenitor-like cells, the antagonism between GATA-1 and PU.1 was proposed to underlie a binary cell fate choice in cells expressing both GATA factors and PU.1. An increase in GATA-1 would result in PU.1 suppression and the erythro-megakaryocytic cell fates, whereas an increase in PU.1 would suppress GATA-1 and give rise to the myelo-lymphocytic lineages [290-293]. However, our data show that CFU-e cells, considered committed erythroid progenitors, express PU.1 at levels equivalent with those found in cells of the myeloid lineage (Figure 2.S3A). Our results are consistent with previous reports of PU.1 expression in early erythroid

progenitors, including the expression of a GFP reporter "knocked in" to the PU.1 gene locus in S0 (CD71^{low}Ter119^{negative}) fetal liver cells [314, 316].

The biochemical nature of commitment to the erythroid lineage is unknown at present. It is possible that CFU-e cells prior to PU.1 down-regulation, which are also expressing GATA-1 and GATA-2, are in fact multipotential cells that may give rise to either myeloid or erythro-megakaryocytic lineages. In this case, the cross-antagonism between PU.1 and GATA-1 would simultaneously be responsible for a lineage choice, as well as facilitate activation of the erythroid gene expression program at the S0 to S1 transition should this choice be in favor of the erythroid lineage. Our ability to isolate CFU-e cells expressing high levels of PU.1 prior to their transition into S1 should facilitate further study of this issue.

The role of S-phase

Why is S-phase progression coupled to the erythroid differentiation program at the S0 to S1 boundary? Linking developmental transitions to cell cycle phases may serve as a strategy for their correct developmental timing [317] and may ensure the correct number of differentiated progeny. Another possibility is that S-phase progression plays a direct role in the re-configuration of chromatin at erythroid-specific gene loci. DNA replication was proposed to provide an opportunity for structural changes in chromatin, since the passage of the replication fork transiently disrupts nucleosomes [283, 284]. Indeed, S-phase is essential for activation or silencing of some genes in yeast [318, 319] and metazoa [317, 320-323], though it is not known that this is due to a requirement in the reconfiguration of chromatin. However, S-phase is not required for activation of other

developmental genes [324-328]. Further, in recent years the structure of chromatin was found to be much more dynamic outside S-phase than originally suspected [329]. It is therefore unclear whether there is an innate requirement for DNA replication in the reconfiguration of chromatin during activation of lineage-specific genes, or what specific aspects of chromatin restructuring might require S-phase.

Here we found that S-phase is required for DNA demethylation and for formation of DNase I hypersensitive sites. The requirement for DNA replication suggests that DNA demethylation is passive, due to a decrease in maintenance methylation of the nascent DNA strand [330]. This raises the possibility that formation of DNase I hypersensitive sites may require DNA replication because it might be contingent on DNA demethylation. Alternatively, DNase I hypersensitivity may require S-phase progression in order to lift a direct repressive effect of PU.1 on chromatin [292].

The role of EpoR

Our examination of EpoR$^{-/-}$ fetal liver shows that the EpoR becomes essential for erythroid differentiation at the S0/S1 boundary. The principal function of EpoR at this time is its pro-survival signaling: EpoR$^{-/-}$ erythroid progenitors undergo apoptosis but their cell cycle status is unaltered, suggesting that EpoR signaling is not required for S-phase progression (Figure 2.S1E). These findings are consistent with the established role of EpoR as a survival factor that does not affect the erythroid cell cycle [331]. EpoR signaling is probably also dispensable for down-regulation of PU.1 at the S0 to S1 transition, since both PU.1 and GATA-2 are low in EpoR$^{-/-}$ cells (Figure 2.3A). In spite of both S-phase progression and PU.1 down-regulation being apparently unimpaired,

EpoR$^{-/-}$ cells fail to develop DNase I HS sites and fail to undergo DNA demethylation at the β-globin LCR (Figures 2.5D, 6A). It has been reported that exogenous expression of Bcl-xL facilitates Epo-independent differentiation of erythroblasts [332] arguing against a direct requirement for EpoR signaling in chromatin reconfiguration. Therefore, EpoR$^{-/-}$ cells may be undergoing rapid apoptosis prior to the time when the chromatin change would have otherwise taken place.

Other than its survival function, EpoR is probably directly required for CD71 up-regulation, via Stat5 [333, 334]. However, EpoR signaling results in CD71 up-regulation only if S-phase is allowed to proceed (Figure 2.2).

Thus, while the onset of Epo dependence occurs synchronously with committal chromatin and transcriptional events in erythroid differentiation, there is apparently no direct requirement for EpoR signaling in these events, other than ensuring cell survival. The principal function of Epo in erythropoiesis is to determine the number of differentiated erythrocytes, via Epo concentration [104]. The S0 to S1 transition may have evolved as the time of onset of Epo dependence as it represents a biochemical commitment to erythroid differentiation, setting in motion chromatin and transcriptional transformations that lead to expression of erythroid- specific genes. This therefore represents the earliest time in erythroid differentiation when Epo may regulate cell number specifically within the erythroid lineage, with minimal lateral effects on other hematopoietic cells.

An all-or-none switch in chromatin state at the β-globin LCR

The β-globin LCR had long been studied as a model of chromatin at sites of lineage-specific genes. However, the time in erythroid differentiation when the locus switches from a "closed" to an "open" conformation had not been clearly defined. Further, it was not known whether activation of the locus develops in a step-wise fashion over several cell cycles and differentiation stages or whether it occurs rapidly in a single step.

Our findings show that, strikingly, the locus transitions to an active conformation rapidly, within S-phase of a single cell cycle. Further, several distinct functional and biochemical changes that characterize the active chromatin conformation appear to develop simultaneously. We found marked differences between S0 and S1 cells in DNA methylation and in DNase I hypersensitivity at the LCR. These transformations could be reproduced when purified S0 cells transitioned into S1 in vitro (Figures 2.5, 6). Both DNA demethylation and DNase I hypersensitivity required S-phase progression for their development. Further, we also found that development of histone-tail modifications characteristic of active chromatin, as well as the switch in the timing of replication of the locus from late to early S-phase, both coincide with the transition from S0 to S1 (Figure 2.5). Therefore, our findings support an „„all or none‟‟‟model for the state of chromatin, previously hypothesized based on the probabilistic nature of developing DNase I hypersensitivity in a range of mutated chicken β-globin enhancer constructs [335].

Previous work showed that while the highest levels of DNase I accessibility at the β-globin LCR are attained in mature erythroid progenitors, the β-globin LCR is already

poised for expression in earlier multipotential progenitors, contributing to low-level β-globin transcription ("priming") [336, 337]. The β-globin LCR was found to already contain DNase I hypersensitive sites in cell lines resembling early hematopoietic progenitors [338]. Here we find that the β-globin LCR appears poised for change prior to the transition from S0 to S1. Thus, LSK and S0 cells have similar DNA methylation levels that are substantially lower than in fetal brain, suggesting chromatin already primed for expression at the LSK stage (Figure 2.6A). Histone tail modifications in the LCR similarly suggest that chromatin in S0 is poised for change, as it is enriched with both H3K4me2, a mark associated with active chromatin, and with H3K27me3, a mark found in silent chromatin (Figure 2.5G). The LCR is therefore marked as a bivalent domain, which may denote chromatin that is silent but primed for activation [339, 340]. Regardless of the precise state of chromatin readiness in earlier hematopoietic progenitors, however, our results show a clear switch in chromatin conformation at the S0 to S1 transition.

The clear switch we identified at the β-globin LCR occurs in synchrony with other switch-like transformations at the transition from S0 to S1, including the onset of Epo dependence and activation of GATA-1 function. Our ability to identify this transition with precision in vivo and manipulate it genetically in vitro should facilitate further study of the pivotal link between the cell cycle clock and the committal chromatin decisions that bring about the erythroid phenotype.

Materials and Methods

Flow cytometry

Fetal livers (E12.5–E14) were mechanically dissociated and immunostained as described [341]. Immunofluorescence was measured on an LSRII (BD Biosciences, CA) and data analyzed using FloJo (Tree Star, CA). Cells were sorted on a FACSAria, FACSVantage (BD Biosciences), or MoFlo (Beckman Coulter) cell sorters. In a small number of experiments StemSep columns (StemCell Technologies) were used.

Fetal liver cells were simultaneously stained for CD71, Ter119 and a Lineage-cocktail with anti-CD3, Gr-1, B220, CD41 and Mac-1 antibodies (BD Biosciences). Background fluorescence in each channel ("fluorescence minus one" - FMO) was determined with isotype-matched antibodies substituting for each specific antibody, one channel at a time. Cells were sorted on a FACSAria, FACSVantage (BD Biosciences) or MoFlo (Beckman Coulter) cell sorters. Dead cells were excluded using 7-aminoactynomycin (7-AAD, Viaprobe, BD Biosciences) or DAPI (Roche). Alternatively, fetal liver cells were labeled with biotin-conjugated antibodies against CD71, Ter119, Gr-1, Mac-1 and CD41, and S0 cells were purified through a StemSep magnetic column as per manufacturer's instructions (StemCell Technologies, Vancouver, BC, Canada).

For detection of PU.1 protein, fetal liver cells freshly sorted, or after retroviral infection were fixed and permeabilized using the BD Cytofix/Cytoperm kit (BD Pharmigen), stained for PU.1 with anti-PU.1 Alexa Fluor® 647 (#2240, Cell Signaling) and analyzed by flow cytometry. FMO control with Rabbit (DA1E) mAb IgG Isotype Control Alexa Fluor® 647 (#2985, Cell Signaling) was used. The following antibodies

were used: Rat Anti-Mouse Gr1-biotin monoclonal RB6-8C5, BD Pharmingen; Rat Anti-Mouse CD71-biotin monoclonal C2F2, BD Pharmingen; Rat Anti-Mouse Ter119-biotin monoclonal Ter119, BD Pharmingen; PE Anti-human CD4 monoclonal RPA-T4, BD Pharmingen; FITC Rat Anti-Mouse CD41 MWReg30, BD Pharmingen; FITC Rat Anti-Mouse CD45R/B220 RA3-6B2, BD Pharmingen; FITC Hamster Anti-Mouse CD3e 145-2C11, BD Pharmingen; FITC Rat Anti-Mouse CD411b/Mac-1 M1/70, BD Pharmingen; FITC Rat Anti-Mouse Ly-6G and Ly-6C (Gr-1) RB6-8C5, BD Pharmingen; Rat Anti-Mouse CD41-biotin monoclonal MWReg30, GenWay Biotech,Inc; PU.1 (9G7) Rabbit mAb Alexa Fluor® 647 (#2240, Cell Signaling).

In vitro culture

Freshly harvested fetal liver cells were sorted and cultured in medium containing 20% fetal calf serum and 2 U/ml Epo (Amgen) for up to 48 h.

Cell cycle analysis

BrdU (100 μl of 10 mg/ml) was injected intra-peritonealy to pregnant mice and embryos were harvested 30–50 min later. In vitro, cells were pulsed with BrdU for 30 min. BrdU incorporation was detected using BrdU flow kit (BD Biosciences). Cell tracking with CFSE (carboxyfluorescein diacetate succinimidyl ester) was performed on sorted S0, incubated with 2.5 μM CFSE (Invitrogen) for 10 min at 37°C.

Retroviral transduction

Retroviral transduction was by spin infection of sorted S0 cells at 2,000 rpm, 37°C on fibronectin coated dishes in 5 μg/ml polybrene (Sigma). Transduced cells were

incubated overnight in the presence of 100 ng/ml SCF and 10 ng/ml IL3 (Peprotech, Rocky Hill, NJ) and were then transferred to Epo-containing medium for the indicated times.

cDNA for PU.1 (provided by D. Tennen, Beth Israel Deaconess Medical Center), GATA-1 and GATA-2 (Open Biosystems) were subcloned into MSCV-IRES-hCD4 retroviral vector. p57T329A has a 1 bp mutation at position T329A (based on the human mutant p57 T310A,[309]) which was created by PCR with the following primers: wild-type forward: 5''-ATGTCAGAATTCACAGCGATGGAACG-3'' and mutant reverse: 5''-ATGTCAGCGGCCGCTCATCTCAGACGTTTGCGCGGGGC-CTGCTCC-3''. The mutation was verified by sequencing. CD71 shRNA (source ID V2MM_13414) was obtained through the UMass Medical School shRNA core facility and subcloned into the LMP-IRES-GFP retroviral vector (Open Biosystems). High-titer viral supernatants were prepared by co-transfecting the desired plasmid and the pCL-Eco packaging vector into Phoenix cells.

Quantitative RT-PCR

Quantitative RT-PCR was performed as described [341]. RNA was prepared using the All Prep DNA/ RNA Micro Kit (Qiagen), and measured with RiboGreen RNA reagent kit (Thermo Scientific) on the 3300 NanoDrop Fluoro-spectrometer. Reverse-transcription was done using the SuperScript II/III (Invitrogen) with random hexamer primers. The ABI 7300 sequence detection system, TaqMan reagents and TagMan MGB probes (Applied Biosystems, San Diego, CA) were used following the manufacturer's instruction. The ΔCt method was used to estimate mRNA using β-actin as a control. The

following TagMan MGB probes were used: β-actin (actin, beta) Mm02619580_g1;

BAND 3 (solute carrier family 4 anion exchanger) Mm01245920_g1; ALAS2

(aminolevulinic acid synthase 2, erythroid) Mm01260713_m1; PU.1 (SFFV proviral

integration1) Mm00488140_m1; GATA-2 (GATA binding protein 2) Mm00492300_m1;

GATA-1 (GATA binding protein 1) Mm01352636_m1; NF-E2 (nuclear factor, erythroid

derived 2) Mm00801891_m1; Tal1 (T-cell acute lymphocytic leukemia 1)

Mm00441665_m1; LMO2 (LIM domain only 2) Mm00493153_m1; EKLF1 (Kruppel-

like factor 1, erythroid) Mm00516096_m1; P57 (cyclin-dependent kinase inhibitor 1C)

Mm00438170_m1; P21 (cyclin-dependent kinase inhibitor 1A) Mm00432448_m1; P27

(cyclin-dependent kinase inhibitor 1B) Mm00438168_m1; EpoR (erythropoietin

receptor) Mm01175894_m1.

DNase I hypersensitivity assays

DNase I hypersensitivity assays were performed as described [312] with

modifications to amplicons. Nuclei were prepared by incubating cells for 5 minutes in

hypotonic buffer (="RSB", 10 mM Tris pH 7.5, 10 mM NaCl, 3 mM MgCl2), then

adding 0.2% NP40 and briefly vortexing. Nuclei were visually inspected, washed in

DNase I digestion buffer (= "DDB", 15 mM Tris pH 7.5, 15 mM NaCl, 60 mM KCl, 0.2

mM EDTA, 0.2 mM EGTA, 0.15 mM spermine, 0.5 mM spermidine with added 1mM

DTT and 1 mM PMSF). DNase I digestion with increasing doses of DNase I (Sigma or

Roche) was then carried out in DDB supplemented with 1mM $CaCl_2$, at 37ºC for 10

minutes, and stopped by the addition of an equal volume of cold stop buffer (50 mM Tris

pH 7.5, 150 mM NaCl, 15 mM EDTA, 0.3% SDS). Samples were treated with proteinase

K overnight and DNA was extracted with the DNeasy Tissue kit (Qiagen) or phenol/chloroform extraction.

The following primers were used for real-time quantitative PCR, in both DNase I hypersensitivity and ChIP-qPCR assays, in conjunction with SYBR Green PCR Master Mix (Applied Biosystems): β-actin Forward, GTTGGGTGACCCCCAGAAT; β-actin Reverse, TTGTGGACACTGCCCCATT; HS1 Forward, TTATCTACTCATGAAGG-AGCAATGATG; HS1 Reverse, GTCAGCTGGGTGGAGTCACA; HS2 Forward, GGGTGTGTTCAGCCTTGTGA; HS2 Reverse, TCCCTGTGGACTTCCTCCTAGA; HS3 Forward, GATGTGTCTATCAGAGGTCCCATATG; HS3 Reverse, TCTCCCCA-CCCTTTGTCCTA; HS4 Forward, TGAACAATTTGCCCTCTCTTACATC; HS4 Reverse, GGCTTCCTGGTCCAGTAGATAGTATT; Nfm Forward, TGCAGGA-TGAGGTGGCTTTC; Nfm Reverse, TGGTGCATGTTCTGGTCTGAGT

ChIP-qPCR

Sorted S0, sorted S1, or fetal brain cells (at least 10^6 cells per sample) were cross-linked in 1% formaldehyde for 10 min at room temperature and stopped by adding 150 mM glycine (pH 7). Cells were lysed in immunoprecipitation (IP) buffer (150 mM NaCl, 50 mM Tris-HCl, pH7.5, 5 mM EDTA, 0.5% NP-40, 1% Triton X-100), supplemented with 20 mM sodium butyrate and protease inhibitors (Roche). Chromatin was sonicated 30 times for 10 sec pulses with Misonix Sonicator 3000's microprobe. For immuno-precipitation, samples were diluted to 500 µl with IP buffer and incubated overnight with antibodies for specific covalent modifications of histones (see list below), followed by a 3-4 hour incubation with Dynabeads protein G (Invitrogen) at 4 C. Cross-links were

reversed at 65 C overnight in elution buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1% SDS). DNA was purified with Zymo Research mini-spin columns; 2 µg herring sperm DNA was used as a carrier. Purified DNA was measured by qPCR using the same amplicons as in the DNase I hypersensitivity assay. The following antibodies were used: Anti H3 acetyl K9, K14, Upstate 06-599; Anti H4 acetyl K5, 8, 12, 16, Upstate 06-866; Anti H3 acetyl K9, Abcam ab4441; Anti H3 dimethyl K4, Abcam ab7766; Anti H3 trimethyl K4, Abcam ab8580; Anti H3 dimethyl K9, Abcam ab1220; Anti H3 trimethyl K27, Active Motif 39535; Anti H3, Abcam ab1791; Anti H4, Abcam ab31827

DNA methylation

Genomic DNA (gDNA) was isolated using the DNA/RNA Micro Kit (Qiagen). Between 50 and 150 ng of gDNA were treated with sodium bisulfite using the EZ DNA Methylation-Direct Kit (Zymo Research, Orange, CA) to generate bisulfite-converted DNA (bcDNA). Specific primers were designed to the sense strand of the bcDNA sequence, within the HS1 and HS2 sites (see primer sequences below), and used for PCR amplification, with 2 µl of bcDNA as template, in a 30 µl reaction containing 3 µl of 10x PCR buffer (Qiagen), 1.8 µl of 25 mM $MgCl_2$, 0.6 µl of 10 mM (each) dNTP, 0.6 µl of 10 µM forward primer, 0.6 µl of 10 µM reverse primer, and 0.15 µl HotStarTaq *Plus* DNA Polymerase (Qiagen). One of the PCR primers was biotinylated to convert the PCR product to single-stranded DNA templates. For the HS1 amplicon, PCR cycling conditions were as follows: 95ºC for 15 min, 45 x (95ºC for 30 sec, 56ºC for 30 sec, 72ºC for 30 sec), then 72ºC for 5 min. The HS2 amplicon used 61ºC for annealing.

Calculation of CpG methylation and verification of complete bisulfite conversion of template DNA was performed on the PCR product by EpigenDx (Worcester, MA) using the Pyrosequencing PSQ96 HS System (Qiagen) as per manufacturer″s instructions. Pyrosequencing for allele quantification is a real-time, sequencing-based DNA analysis that quantifies multiple, and consecutive CpG sites individually. HS1 primers: Forward, GGTATATTTGAAAATTTTGGTAATAGT; Reverse, CAACAAATAAAAACCAAA-ACTACACA. HS2 primers: Forward, GGTTTTTTGGTAGATAGGTTATATGTG; Reverse, AAAACTAAACACACCCACAAAATA.

Cytospin preparation

Fetal liver cells (freshly sorted, or after retroviral infection and incubation in Epo for 24-48 hours) were cytospun onto coated glass slides at 800 rpm for 5 minutes (Single Cytofunnel, Fisher Brand). The slides were air-dried, methanol fixed, stained with diaminobenzidine hydrochloride followed by Wright-Giemsa stain (Sigma). Digital images were taken with a Zeiss microscope using a SPOT Flex Camera (Diagnostic Instruments, Inc) and analyzed with version 4.5 Spot imaging software.

Single cell PCR analysis

S0 cells were individually sorted from fresh fetal liver directly into 96 well plates containing lysis buffer (0.4% NP-40, 25 µM DTT, 0.5 U RNase Out and 65 µM dNTPs). Reverse transcription was performed directly on cell lysates according to the manufacturer″s instructions (Qiagen Sensiscript RT kit or Invitrogen SuperScript III First Strand Synthesis System kit). EpoR and β-actin were amplified from 2 µL of cDNA via

nested PCR. Each experiment included spleen cDNA (positive control), thigh muscle cDNA and no template (negative controls). The following primers were used: β-actin Forward, CTAGGCACCAGGGTGTGATGG; β-actin Reverse: TCTCTTTGATGTCAC-GCACGA; EpoR Forward, AACTTCCAGCTGTGGCTGCTG; EpoR Reverse, CCTTG-TCCCAGACTCAGAG; Nested PCR β-actin Forward, CGAGGCCCAGAGCAAGA-GAG; Nested PCR β-actin Reverse, CGGTTGGCCTTAGGGTTCAG; Nested PCR EpoR Forward: CCTACTTGGTATTGG; Nested PCR EpoR Reverse: CCTCACCTTCCAGCTTTGAG

Fluorescent in situ hybridization

Cell preparation: BrdU pulse labeling was done in vivo by injecting pregnant females with BrdU intraperitoneally. Sorted S0 and S1 (0.5 to 1 x 10^6 cells for each) were washed twice in 5 ml cold PBS/2% sucrose, and fixed by adding 5 ml of cold Methanol/Acetic acid mixture (3:1) drop-wise, while vortexing. Cells were fixed for 10 min at 4ºC, washed x3 in Methanol/ Acetic acid and stored at -20ºC until used for FISH. Alternatively, sorted S0 and S1 cells were suspended at 50,000 cells in 200 ml PBS/2% sucrose and cytospins prepared by spinning 5 min at 800 rpm, allowed to dry briefly (10-20 seconds), and fixed in a 50 ml Falcon containing 20 ml Methanol/Acetic acid at 4ºC for 10 min; slides were poly-L lysine coated.

Prior to FISH, fixed cells on slides were treated in 70% ethanol, subjected to pepsin digestion (0.01%) and post-fixed in 3.7% paraformaldehyde. Cells were dehydrated in successive 3 min washes of 70%, 90% and 100% ethanol and air-dried and used for probe hybridization on the same day.

Probe to the β-globin major gene was labeled with digoxigenin and resuspended in hybridization mix in the presence of tRNA and mouse COT DNA, used at 20 ng/µl, 50 µl per slide. Hybridization was carried out in a humidified chamber at 37% overnight in hybridization mix containing 10 mg/ml sonicated Salmon Sperm DNA, 50% formamide, 2X SSC, 50 mM sodium phosphate buffer and 10% dextran sulphate. Immunostaining was performed by labeling slides with sheep anti-digoxigenin (Roche, *1 333 089*) followed by fluorescein-conjugated rabbit anti-sheep IgG (402104, EMD Bioscience, NJ, USA) and then FITC-conjugated goat anti-rabbit IgG together with mouse anti-BrdU IgG (M0744, Dako, CA, USA) followed by Texas Red-conjugated donkey anti-mouse IgG (Jackson immunoresearch). Nuclei were counterstained in DAPI and visualized using fluorescence microscopy.

Quantitative western blot analysis

Sorted cells from each fetal liver subset were lysed by rocking for 15 min at 4°C in lysis buffer (1% NP40, 50 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 10% glycerol supplemented with protease inhibitors); following centrifugation for 15 min at 4°C, supernatant was quantified by the BCA Protein Assay Kit (Pierce). Lysates were analyzed on a 4-12% gradient polyacrylamide gel (Invitrogen), transferred to a nitrocellulose membrane, and probed with Rabbit polyclonal to beta actin (ab8227, Abcam), and with Mouse monoclonal to $p57^{KIP2}$ (sc-56341, Santa Cruz) overnight at 4ºC. Membranes were probed simultaneously with anti-Rabbit IRDye 680 (red channel) and anti-Mouse-IRDye 800CW (green channel) and analyzed for both β-actin and $p57^{KIP2}$

using the Odyssey infrared imaging system (LI-COR Biosciences). Target protein bands

were quantified using the Odyssey software v3 (LI-COR).

Figure 2.1. Up-regulation of CD71 coincides with the onset of EpoR dependence and with S-phase of the last generation of CFU-e.

A. Fetal liver subsets S0 to S5 form an erythroid developmental sequence. Freshly isolated E14.5 fetal liver was mechanically dissociated and labeled for cell-surface CD71 and Ter119. Cytospin preparations from each subset (right panel) were stained with Giemsa and diaminobenzidine. Scale bar = 20 μm

B. CD71/Ter119 profiles for E12.5 EpoR$^{-/-}$ and wild-type littermate fetal livers. Erythroid differentiation of EpoR-/- cells is blocked at the transition from S0 to S1. Ter119^{+} cells in EpoR$^{-/-}$ liver are nucleated yolk-sac erythrocytes (Figure 2.S1A). Representative of more than four experiments.

C. Erythroid and non-erythroid colony forming potential of fetal liver subsets. 100,000 cells sorted from each of S0 to S3 were plated in methylcellulose in the presence of Epo, IL-3, SCF, and IL-6. Colonies were scored on days 3 (CFU-e) and 10 (CFU-GM). Data are mean ± SE of three independent experiments. Pictures of colonies on day 2 are shown (right panel, lens magnification x20 for all subsets). There was no statistically significant difference between S0 and S1 (p = 0.2, paired t test).

D. Non-erythroid lineage-marker expression in S0 to S5. Wild-type or EpoR$^{-/-}$ fetal livers were labeled with CD71 and Ter119 to identify subsets S0 to S5, and with a cocktail of non erythroid lineage markers containing Mac-1, Gr-1, CD41, B220, and CD3, or with isotype control antibody. See also Figure 2.S1B.

E. Representative cell cycle analysis for S0 to S5. Pregnant mice were injected with a pulse of BrdU, and fetal livers were harvested 30 to 50 min post-injection. Cells from each of subsets S0 to S5 were sorted by flow-cytometry and labeled for BrdU incorporation and DNA content (7AAD).

F. Summary of six independent cell cycle analysis experiments as described in (E). Data are mean ± SE. The difference between the number of S-phase cells in S0 and S1 is significant at p<0.0001 (paired t test).

G. Upper panel: Quantitative RT-PCR analysis of mRNA expression for cyclins A2, B1, E1, E2, D2, and for p57^{KIP2}. Data (mean ± SE of three experiments) were normalized to β-actin mRNA in each sample and expressed as S1/S0 fold change (no change = 0). Fold changes are indicated. Lower panel: p57^{KIP2} protein in sorted S0, S1, and S2+3 subsets. Quantitative western blotting with antibodies directed at p57^{KIP2} and β-actin; Near infra-red (NIR) fluorescence–conjugated secondary antibodies. S0 cell transduced with retroviral vector encoding either p57T329A-ICD4 or empty vector (MICD4) were harvested 24 h post-infection and used as positive and negative controls, respectively.

H. CFSE cell tracking of S0 cells as they transition into S1 and S2 in vitro. Sorted S0 cells were pulsed with CFSE and incubated in Epo for 18 h. The time points examined during in vitro incubation are indicated. Upper panel: CD71/Ter119 profiles. The S0 (blue), S1 (red), and S2 (green) gates are indicated with the percentage of cells in each gate. Middle panel: Corresponding CFSE histograms for cells in each of the S0 (blue), S1 (red), and S2 (green) subsets. Lower panel: Median fluorescence intensity of CSFE for the corresponding histograms and colors shown in the middle panel. Representative of 4 similar experiments.

I. Representation of the transition from S0 to S1. S0 contains several CFU-e generations. The last generation of CFU-e, noted as "CFU-e.2", arises in S0 as an EpoR-independent cell. The onset of EpoR dependence and up-regulation of CD71 (= transition to S1) occur during S-phase of this cell generation. Up-regulation of Ter119 (= transition into S2) occurs as the CFU-e.2 cell divides, giving rise to non-CFU-e progeny in S2. Other than the cell cycle corresponding to CFU-e.2, the timing of other cell cycles with respect to differentiation events is not known. See also Figure 2.S1.
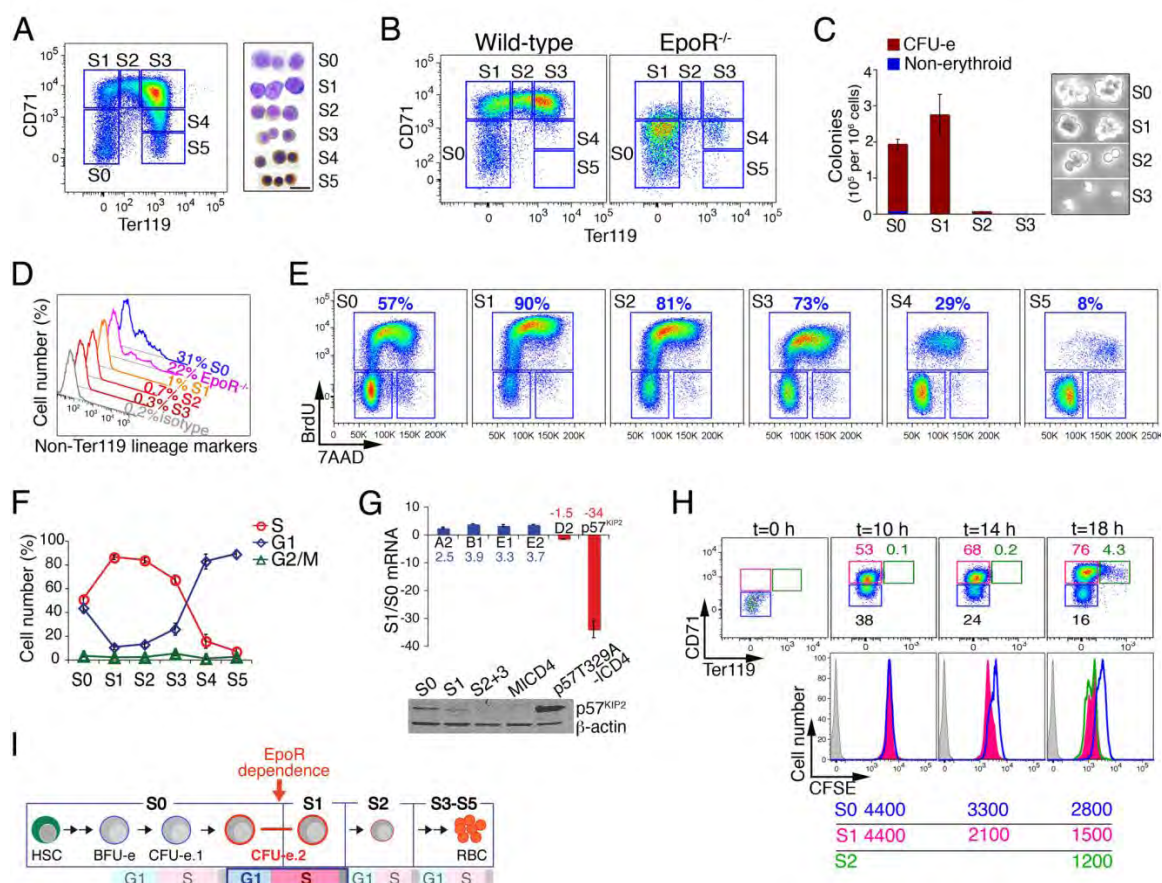
Figure 2.2. The S0 to S1 transition requires S-phase progression.

A. Design of experiments illustrated in sections (B–D). Flow-cytometrically sorted S0 cells were incubated in Epo for 20 h. In the first 10 h, cells were also in the presence or absence of a cell cycle blocking drug, either aphidicolin (Aphi) or mimosine (Mim). DMSO was added to control cells in the Aphi experiments. Aphi or Mim were removed by washing at t = 10 h. Cell cycle status and CD71/Ter119 expression were examined at 4, 10, and 20 h.

B. CD71/Ter119 expression and BrdU/7AAD cell cycle profile of freshly sorted S0 cells at t = 0. Cells were incubated in the presence of BrdU for 30 min prior to fixation, permeabilization, and staining with antibodies for CD71, Ter119, and BrdU.

C. CD71/Ter119 expression (left columns) and corresponding cell cycle profile of S0 and S1 cells (right columns) at the indicated time points. The presence or absence of Epo, Aphi, or Mim is indicated above each row of histograms. The indicated percentages correspond to the fraction of cells in S1 (left columns) or in S-phase of the cycle (right columns). Note that Ter119 signal is reduced in fixed and permeabilized cells compared with equivalent non-fixed cells (e.g., Figure 2.1H and Figure 2.S2A), since the Ter119 epitope is partially detergent soluble. Data representative of five experiments.

D. "Back-gating" analysis for CD71 expression of S0 cells that are either in G1 (blue) or in S-phase (red), at t = 4 h in the presence of Epo+Mim. The same BrdU/7AAD profile as in section (C), row 4, t = 4 h. See also Figure 2.S2.
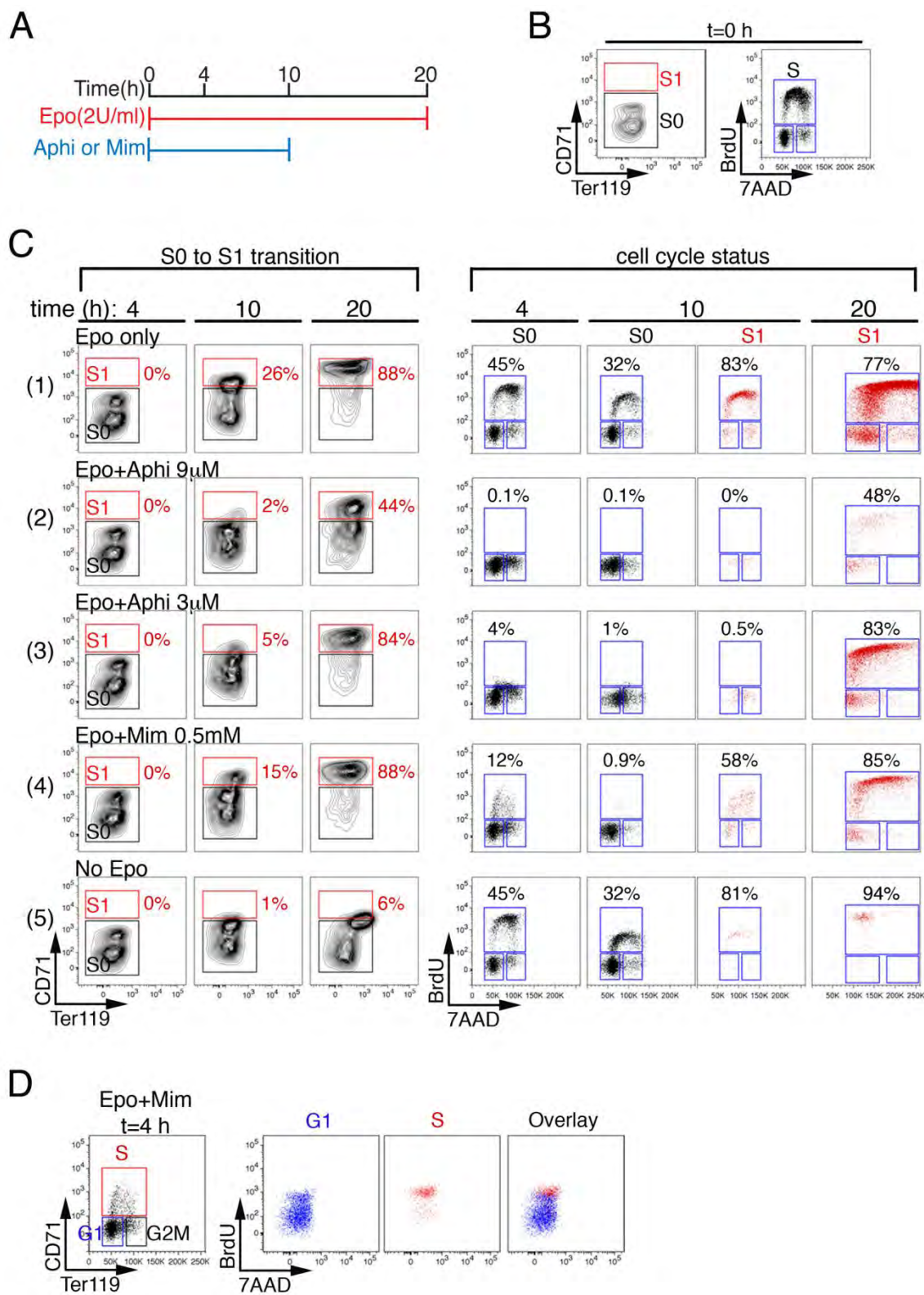
Figure 2.3. Block of S-phase progression at the S0 to S1 transition arrests the

erythroid differentiation program.

A. Expression of transcriptional regulators (upper panel) and erythroid-specific genes (lower panel) in sorted fetal liver subsets S0 to S3, in fetal brain, and in lineage marker-depleted EpoR$^{-/-}$ fetal liver. mRNA measured by quantitative RT-PCR, normalized to the β-actin mRNA, and expressed as a ratio to the S0 subset. Data are mean ± SD of 2 (for EpoR$^{-/-}$ and brain) or 3 (for S0 to S3) independent experiments.

 B, C. Effect of aphidicolin-mediated S-phase arrest on erythroid-specific gene expression. Sorted S0 (in B) or S1 (in C) cells were incubated in Epo, and in either aphidicolin (3 μM) or DMSO, for the first 10 h. Aphidicolin and DMSO were removed by washing at t = 10 h. mRNA was measured by qRT-PCR, normalized to β-actin, and expressed as a ratio to mRNA at t = 10 h in Epo+DMSO. Duplicate independent experiments shown, fitted with exponential curves. Black curves are the calculated time course for a 10 h delay in induction for each gene, obtained by shifting the respective red curves (describing time course for Epo+DMSO) by 10 h.

D. Effect of aphidicolin-mediated S-phase arrest on PU.1 and GATA-2. Experiment and mRNA measurement as described for (B), except that aphidicolin and/or DMSO were applied at t = 4 h and removed at t = 14 h. Data from two independent experiments.

E. mRNA expression of transcriptional regulators GATA-1, GATA-2, and PU.1 at the end of a 10 h incubation in Epo aphidicolin, compared with cells incubated in Epo+DMSO. Experimental design and mRNA measurement as described in section (D), with data pooled from nine independent experiments (each with a distinct symbol). Bar indicates position of mean. Differences between Epo+aphidicolin and Epo control samples are significant (paired two-tailed t test) for GATA-2 (p = 0.004) and for PU.1 (p = 0.002).

F. Effect of aphidicolin or mimosine-mediated S-phase arrest on erythroblast morphology during differentiation of S0 cells in vitro. Sorted S0 were incubated in Epo, and for the first 10 h of incubation, also in the presence of either aphidicolin or mimosine. Control cells were incubated in Epo alone throughout. Cytospin preparations of cells at 10 and 20 h of incubation are shown, stained with Giemsa. Morphological maturation (decreasing cell and nuclear size), compared with control cells, was arrested at t = 10 h, and remained delayed when the block to cell cycle progression was removed (t = 20 h).

G. S-phase arrest by overexpression of a non-degradable mutant of p57^{KIP2} (p57T329A). Sorted S0 cells were transduced with retroviral vector expressing p57T329A linked to IRES-hCD4 reporter, or with control vector (MICD4). Cells were cultured for 15 h in IL-3 and SCF and transferred to an Epo containing medium at t = 0, to allow their transition from S0 to S1. CD71/Ter119 profiles of infected, hCD4 positive S0 at t = 0 and t = 14 h (upper panels) and BrdU/7AAD cell cycle profiles (lower panels)

at t = 14 h are shown. Representative of 4 experiments. Transduction efficiency of p57[KIP2] exceeded 90% in all experiments.

H. Transcriptional regulators GATA-1, GATA-2, and PU.1 in cells expressing p57T329A. Experiment as described in (G). mRNAs measured by qRT-PCR following 24 h of incubation in Epo, and expressed relative to cells transduced with control vector (MICD4) after normalization to β-actin. Duplicate independent experiments are shown.

I. Representation of erythroid gene expression at the transition from S0 to S1. GATA-1 and other activators of the erythroid transcriptional program are induced at an unknown time preceding the S0/S1 boundary, and increase modestly with further differentiation. P57[KIP2], PU.1, and GATA-2 are expressed in S0 and are markedly down-regulated at the S0 to S1 transition. Erythroid specific Ter119, β-globin, ALAS2, and Band3 are induced subsequent to the S0/S1 boundary. S-phase arrest at this stage (dashed black line) results in arrest of all subsequent events, including PU.1 and GATA- 2 down-regulation, CD71 and Ter119 expression, erythroid-specific gene induction, and morphological maturation. See also Figure 2.S3.
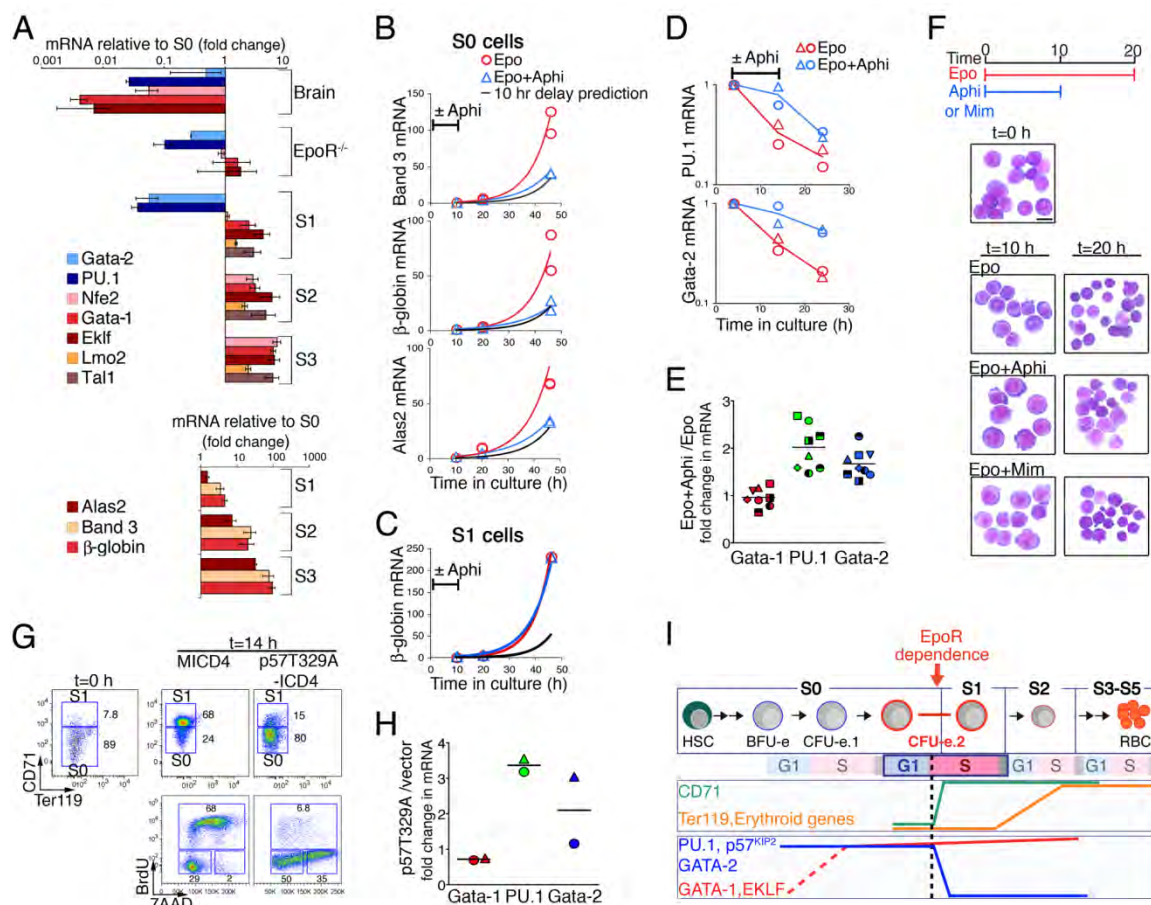
Figure 2.4. PU.1, but not GATA-2, inhibits the transition from S0 to S1.

A–C. Effect of exogenous PU.1 on the transition from S0 to S1. Sorted S0 cells were transduced with retroviral vector expressing PU.1-IRES-hCD4 (PU.1-ICD4) or control vector (MICD4) and were incubated in IL-3 and SCF for 15 h before being transferred to Epo for 24 h.

A. Expression profiles of the hCD4 reporter in cells transduced with either PU.1-ICD4 or with control MICD4, at 24 h of Epo culture. Vertical, narrow gates each containing cells of relatively uniform hCD4 expression are shown and numbered (i to vii) and are used in the analysis shown in sections (B) and (C) below.

B. CD71/Ter119 (upper panels) and cell cycle (lower panels) analysis of cells in individual hCD4 gates (ii) to (v), at t = 24 h. Higher hCD4 indicates higher PU.1 expression in cells transduced with PU.1-ICD4.

C. Summary of histogram data in (B), correlating cell cycle and differentiation data to hCD4 expression. Each data point corresponds to one of the hCD4 vertical gates marked in (A). Data are representative of five independent experiments.

D. Linear correlation between PU.1 expression levels as measured by flow-cytometry using a PU.1-specific antibody, and hCD4 levels, in cells transduced either with PU.1-ICD4 or with control MICD4. Summary of data shown in Figure 2.S4A, B. Data are representative of 2 independent experiments.

E. Effect of exogenous GATA-1 or GATA-2 on S0 cell differentiation and cell cycle. Experimental design as in section (A–C). CD71/Ter119 and cell cycle profiles are shown for hCD4+ cells transduced with the indicated retrovirus. Flow-cytometric expression of retroviral constructs is shown in Figure 2.S4C.
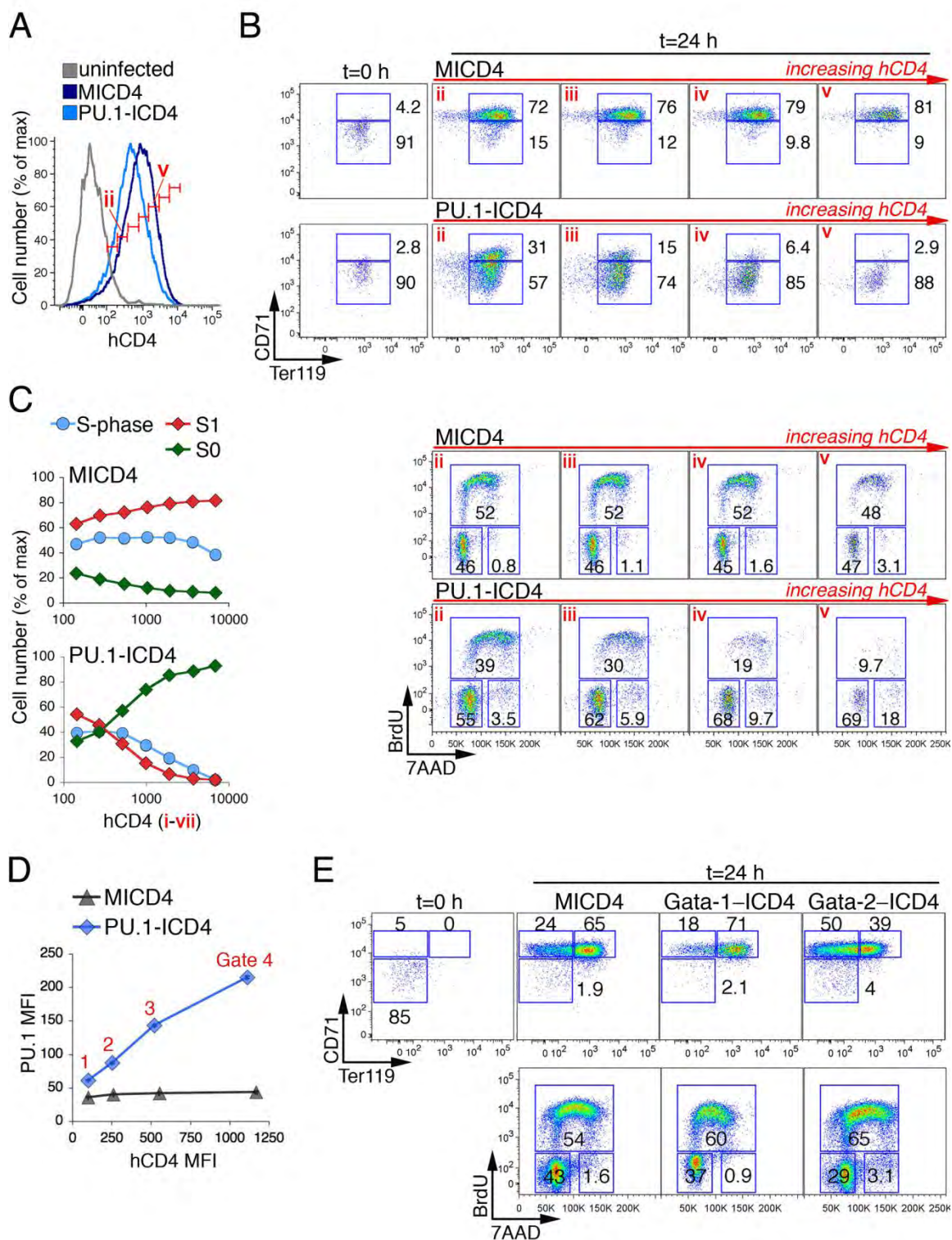
Figure 2.5. The S0 to S1 transition coincides with an S-phase dependent switch in the state of chromatin at the β-globin locus.

A. A map of the mouse β-globin locus. β-globin genes are indicated with horizontal arrowheads. Vertical arrows indicate DNase I hypersensitivity sites (HSs). Solid arrows for HS1 to HS4 indicate sites examined in experiments below. Expanded HS1 and HS2 sites show locations of CpG dinucleotides (labeled HS1A, B, C, and HS2A, B, C).

B. FISH analysis of the timing of replication at the β-globin locus. Pregnant female mice were injected with BrdU 30 min prior to harvesting of fetal livers. Sorted S0 and S1 cells were fixed and stained for BrdU (red) to identify S-phase cells. Cells were hybridized with a probe to the β-major gene (green), and the number of hybridization spots in 100 BrdU-positive cells in consecutive fields for each of S0 or S1 were counted by fluorescence microscopy, in two independent experiments. Cells were scored as DD (two double dots, indicating both alleles have replicated), SD (one single and one double dot, indicating only one allele has replicated), or SS (two single dots, indicating neither allele has replicated). Examples of nuclei with each of the patterns are shown.

C. DNase I sensitivity in S0 or S1 cells. Nuclei were prepared from sorted S0 and S1 cells and digested with increasing concentrations of DNase I for 10 min. DNA was extracted and quantitative PCR used to measure remaining DNA at each of HS2, HS3, and HS4 using 150 bp amplicons. DNA measurements were normalized to DNA amplified at the neuronal gene Nfm. Representative of three independent experiments.

D. DNase I sensitivity of whole fetal liver from E12.5 EpoR$^{-/-}$ and littermate wild-type controls. Method as in (C). Representative of three independent experiments.

E. Effect of S-phase arrest on development of DNase I hypersensitivity. Sorted S0 cells were incubated in Epo for 10 h to allow transition to S1, in the presence of aphidicolin (Aphi) or DMSO (control). Nuclei were prepared and DNase I sensitivity measured as described in (C). DNA in each sample was normalized to Nfm and expressed as a ratio of DNA in cells incubated in Epo+aphidicolin to DNA in cells incubated in Epo+DMSO. Individual data points are pooled from 3 independent sorting and digestion experiments; curves are fitted second order polynomials.

F, G. ChIP-qPCR in sorted S0 and S1 cells and in fetal brain. ChIP was performed with the indicated antibodies and with control, isotype-matched antibody. qPCR was of 150 bp amplicons at the LCR HSs and at the β-actin and Nfm genes. Data are expressed as enrichment over input DNA in S1 relative to S0 (in F and in G, upper panel) or S0 relative to fetal brain (in G, lower panel). Each sample was normalized to β-actin after subtraction of background (ChIP background was the signal with isotype control antibodies, which was, 10% of the signal obtained with specific antibodies).

F. Summary of seven independent ChIP-qPCR experiments. Data are means of at least 2 to 4 experiments for each antibody/amplicon combination (SE is provided when at least 3 experiments are averaged for a given antibody/amplicon). * indicates statistically significant difference between S0 and S1 at the HS2 site (p = 0.019 and 0.032 for H3K27me3 and H3K4me2, respectively). Changes in H3K27me3 and H3K4me2 over all HSs tested were significant at p = 0.011 and p = 0.0006, respectively (paired t test used for all significance tests).

G. Representative ChIP-qPCR experiment that included sorted S0, S1, and fetal brain. Data are mean ± SE of three replicates, expressed as a ratio of S1 to S0 (upper panel) and S0 to fetal brain (lower panel).
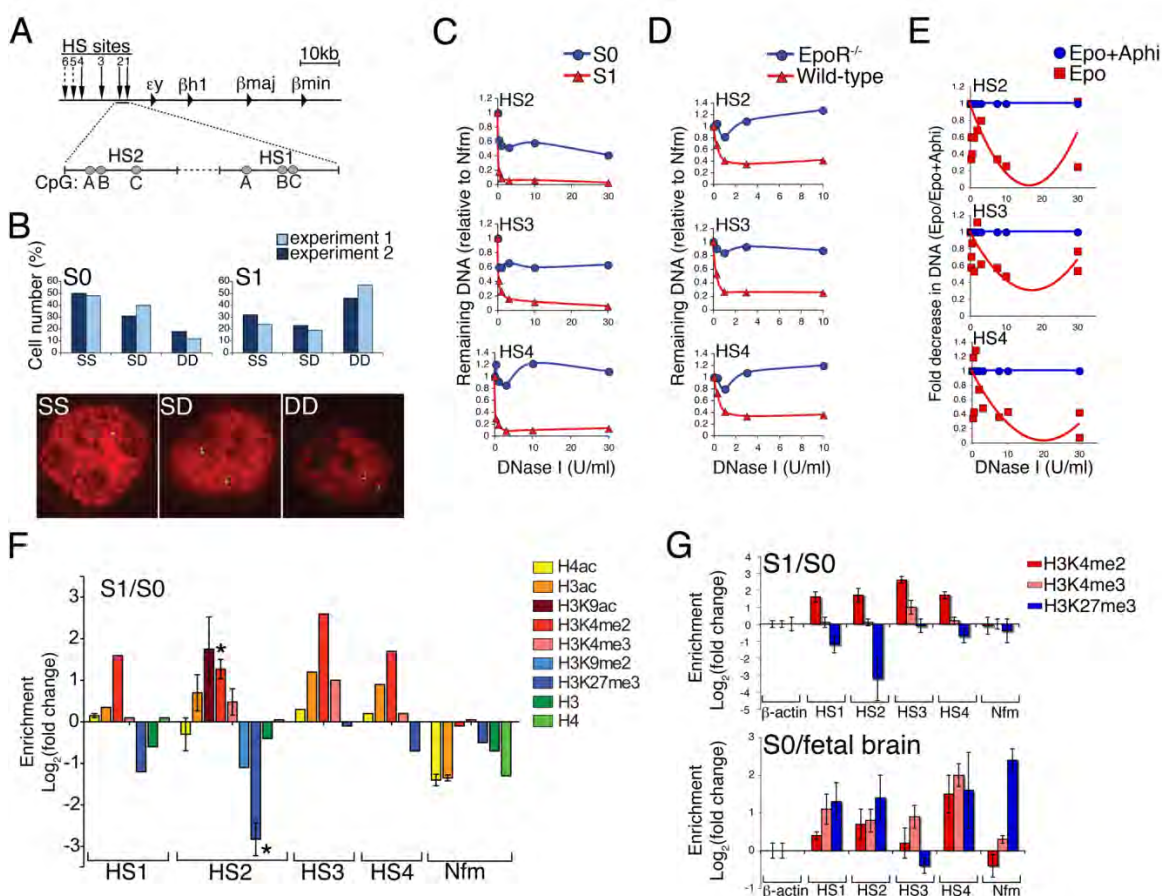
Figure 2.6. The transition from S0 to S1 is marked by the onset of S-phase

dependent, DNA demethylation at HS1 and HS2.

A. Methylation levels at each of the 6 CpG dinucleotides in HS1 (HS1A, B, C) and HS2 (HS2A, B, C; Figure 2.5A), in each of the indicated cell populations. Hematopoietic cells were sorted flow-cytometrically from freshly isolated fetal liver. Brain = fetal brain; LSK = $Lin^-Sca1^+Kit^+$; Mac-1 = $CD71^{low}Ter119^-Mac-1^+$; CD41 = $CD71^{low}Ter119^-CD41^+$; $EpoR^{-/-}$ = Lin^- cells from $EpoR^{-/-}$ fetal liver. Methylation levels were assessed following PCR-amplification of bisufite-converted genomic DNA at each of the HS1 and HS2 loci followed by pyrosequencing. Each data point is the mean ± SE of 2 to 4 independent sorting and pyrosequencing experiments.

B. Arrest of S-phase progression by aphidicolin prevents DNA demethylation. Aphidicolin was added for 10 h at t = 0 (HS1A, B, C, and HS2C, left panel) or at t = 4 h (HS2A, B, right panel) to sorted S0 cells incubated in Epo for 48 (left panel) or 24 h (right panel). Control cells were incubated in Epo only. CpG methylation was measured at the indicated time points as described in (A). Data are from two independent experiments.

C. Arrest of S-phase entry by mimosine prevents DNA demethylation. Sorted S0 cells were incubated in Epo for 24 h, in the presence or absence of mimosine between t = 4 and t = 14 h. CpG methylation was measured as described in (A). Data are mean ± SE of three independent sorting and pyrosequencing experiments.

D. Preventing PU.1 down-regulation prevents DNA demethylation of HS1 and HS2 at the S0 to S1 transition. S0 cells were transduced with retroviral vectors expressing either PU.1-ICD4 or MICD4 as described in Figure 2.4A–C. CpG methylation levels were measured as described in (A) following 24 h incubation in Epo. Data are mean ± SE of three independent experiments (cells were transduced with PU.1-ICD4 at >90% efficiency in one experiment, and in two additional experiments hCD4+ cells were sorted before the start of Epo incubation).
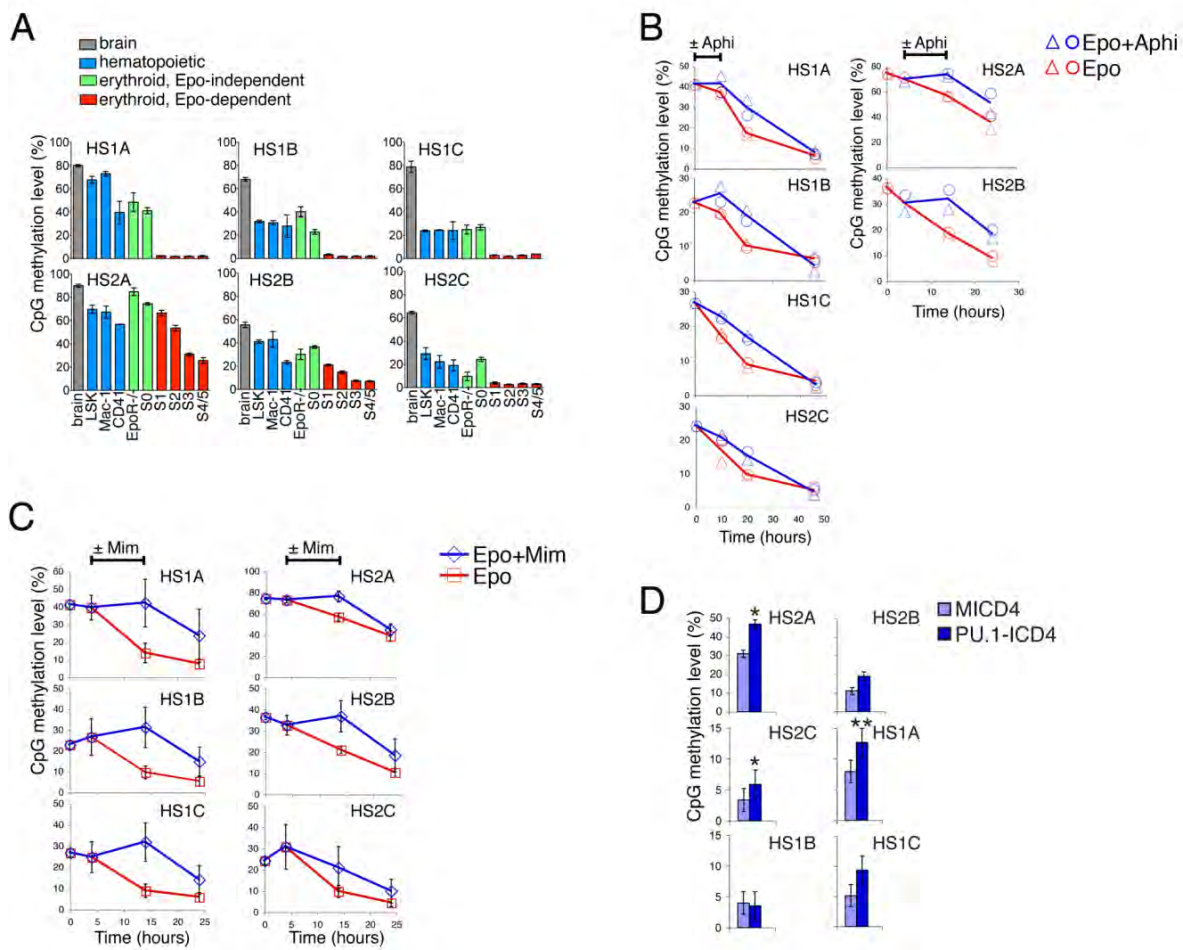
Figure 2.7. Regulatory events at the transition from S0 to S1.

A. Multiple differentiation milestones coincide with early S-phase in the last CFU-e generation. Up-regulation of CD71 marks the transition from S0 to S1 in early S-phase of this cell cycle. It coincides with the onset of Epo dependence, down-regulation of transcriptional suppressor PU.1 and CDKI p57KIP2, and reconfiguration of chromatin at the β-globin locus, including a switch in the timing of replication from late to early S-phase and, within the LCR, the formation of DNase I hypersensitivity sites, the onset of DNA demethylation, and a loss of repressive histone marks from bivalent chromatin. The dashed black line marks the time at which inhibition of S-phase arrests PU.1 down-regulation, CD71 up-regulation, and the switch in chromatin. Induction of erythroid transcriptional activators, including GATA-1, precedes this step. Erythroid specific genes including β-globin, Band3, and ALAS2 are induced subsequently. Up-regulation of Ter119 occurs approximately with entry into the next cell cycle.

B. Causal relationships at the transition from S0 to S1. Red arrows mark novel causal relationships implicated in the transition from S0 to S1. Down-regulation of p57KIP2 is required for S-phase progression during the last CFU-e generation. Mutual antagonism between PU.1 and S-phase results in PU.1 down-regulation as cells progress into S-phase. Decreasing PU.1 allows for functional activation of GATA-1, which in turn represses GATA-2 and induces erythroid-specific genes such as β-globin. S-phase progression is also required for the formation of DNase I hypersensitive sites and for DNA demethylation bringing about a switch in chromatin conformation at the β-globin LCR and, together with Epo, for CD71 up-regulation.
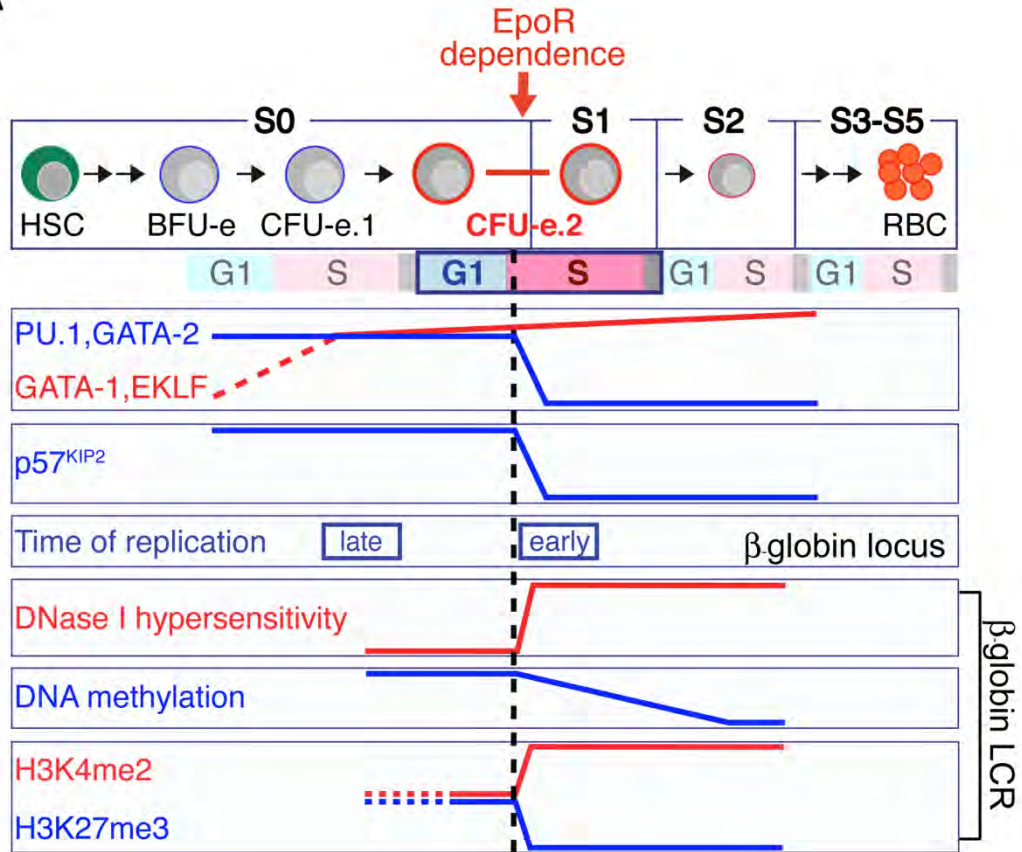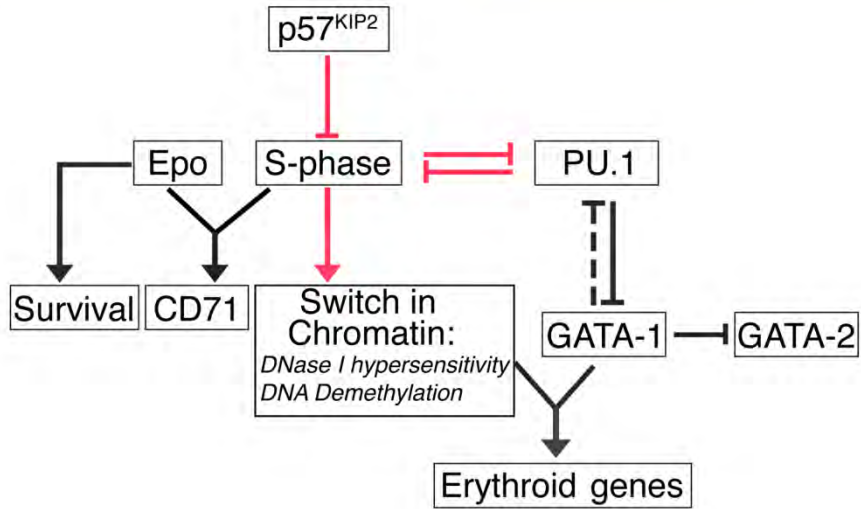
Figure 2.S1. Supplemental data to Figure 2.1.

A. Ter119[+] cells in EpoR[-/-] fetal livers are nucleated erythrocytes of the yolk-sac (primitive) lineage. Cytospin preparations of sorted Ter119 [+] cells from fetal livers of EpoR[-/-] (E12.5) and wild-type littermate. Yolk-sac erythrocytes are fully hemoglobinized, large nucleated cells (arrow; see brown coloration of hemoglobinized cells stained with diaminobenzidine). Basophilic (blue cytoplasm) erythroblast precursors of the definitive lineage form the majority of Ter119[+] cells in the wild type fetal liver but are absent from EpoR[-/-] fetal liver. Scale bar =20 μ.

B. Distribution of cells expressing non-erythroid lineage markers within fetal liver. The fraction of cells expressing each indicated marker is shown for embryonic ages E12.5 to E14.5 in wild-type embryos, and for EpoR-/- embryos on E12.5. The same data are represented for each lineage marker either as a fraction of whole fetal liver (left panel) or as a fraction of S0 (right panel).

C. The fraction of S0 cells expressing EpoR mRNA assessed by single cell RT-PCR. Single cell RT-PCR was carried out on 324 individual S0 cells. An mRNA signal (either EpoR or β-actin or both) was obtained for 159 cells (49%). 68% of cells with a positive mRNA signal were positive for the EpoR mRNA. Single S0 cells were sorted by flow-cytometry into single wells of a 96-well plate. Following reverse transcription, two rounds of PCR amplification were used to detect EpoR and β-actin expression in individual cells. Shown are representative examples of RT-PCR for EpoR (top gel) and β-actin (lower gel) for four individual cells (lanes 1 to 4). The cells in lanes 1 and 4 expressed both EpoR and β-actin. Neither EpoR nor β-actin signals were obtained for the cell in lane 3. The cell in lane 2 expressed only β-actin. Control lanes are RT-PCR on spleen cells (positive control, "c1"), thigh muscle (negative control, "c2"), and no template (negative control, "c3").

D. BrdU incorporation rate (measured as median fluorescence intensity, MFI) in wild-type S1 is higher than in other fetal liver subsets, and higher than in EpoR[-/-] fetal liver (latter computed for the few cells within the S1 gate). Difference between wild-type S0 and S1 is significant at $p<0.0001$ (paired t test). Data are mean ± SE of 7 independent experiments.

E. Left panels: EpoR does not regulate cell cycle status of erythroid progenitors. Representative CD71/Ter119 profiles (upper panels) and BrdU/7AAD (lower panel) for the S0 subset in EpoR[-/-] fetal liver and in wild-type littermates. The fractions (%) of S-phase cells (lower panels) is indicated. Right panel: EpoR regulates survival of erythroid progenitors. "LIVE/ DEAD" profiles of the same EpoR[-/-] and wild-type littermate embryos. The LIVE/DEAD dye (Molecular Probes) stains apoptotic cells with impaired membrane permeability prior to fixation and permeabilization assays. The fractions (%) of apoptotic cells are indicated.

F. qRT-PCR analysis of mRNA expression for p27^{KIP1}, p21^{CIP1}, and p57^{KIP2} in sorted fetal liver subsets S0 to S3 cells. Data were normalized to the β-actin mRNA in each sample and expressed as a ratio to the S0 subset. Data are mean ± SD of three independent experiments.
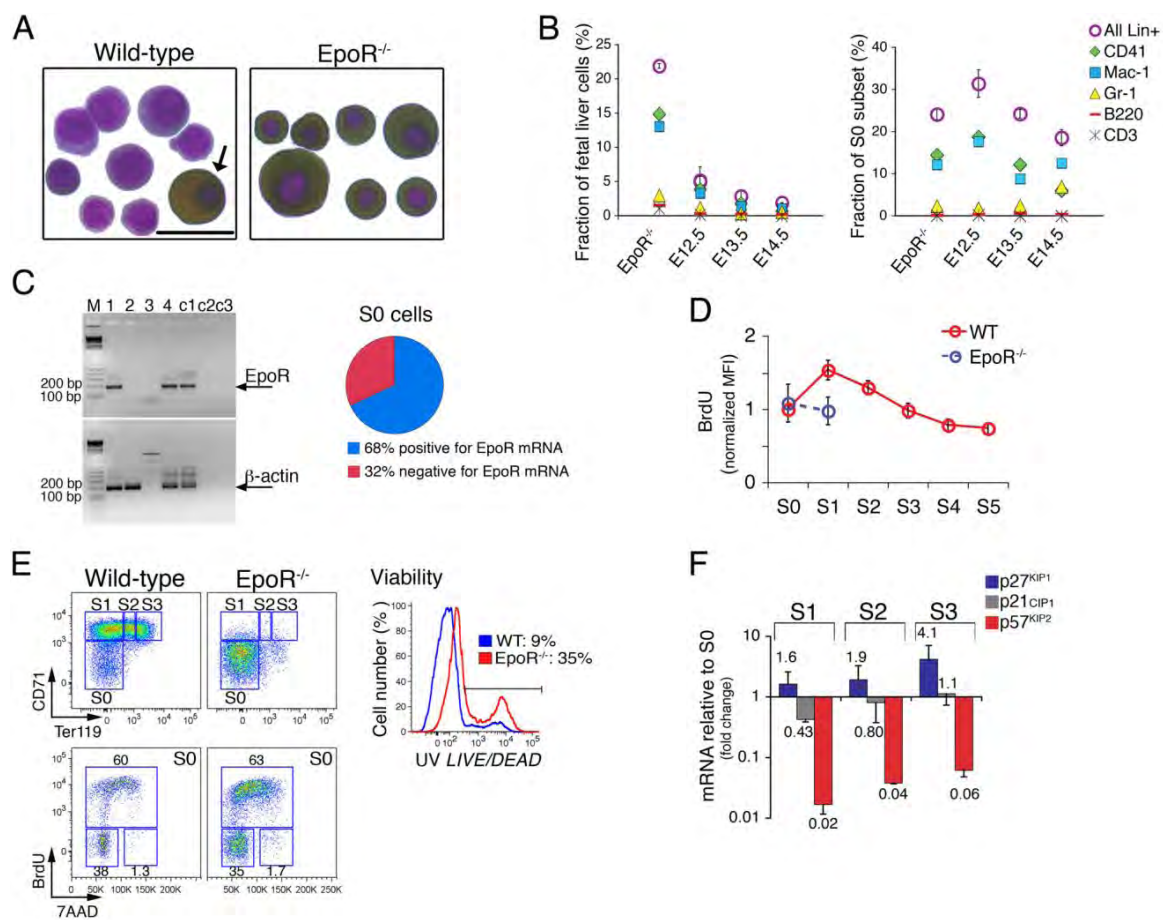
Figure 2.S2. Supplemental data to Figure 2.2.

A. S1 cells are sensitive to hydroxyurea (HU). S0 cells were sorted by flowcytometry (t= 0) and incubated in the presence or absence of Epo, and in the presence or absence of HU (5 mM), as indicated. Samples incubated in Epo alone up-regulated CD71 and Ter119 sequentially (see t =15 h, t= 24 h). Up-regulation of CD71, but not Ter119, is Epo dependent (see „no Epo""sample in which Epo was added at t= 15 h). Cells incubated in Epo and HU did not upregulate CD71 or Ter119, consistent with the transition into S1 (CD71 up-regulation) being S-phase dependent. Cells incubated in Epo alone for 15 h had transitioned into S1 and began to transition into S2. If HU was added at this point (t= 15 h), nearly all S1 cells are lost (due to HU toxicity, unpublished data), but many cells in S0 and S2 persist. This suggests that essentially all S1 cells are in S-phase.

B. Preventing CD71 up-regulation at the S0/S1 boundary does not interfere with the erythroid cell cycle. Sorted S0 cells were transduced with retroviral vectors containing an "IRES-GFP" reporter, expressing short hairpin RNA targeting CD71 (CD71shRNA) or "empty vector" control (LMPv). Cells were then cultured for 24 h in the presence of Epo. Shown are the CD71/Ter119 profiles (top panels) and corresponding BrdU/ 7AAD cell cycle profiles of cells at t= 24 h. Only retrovirally infected cells are shown, identified by the expression of GFP. The fraction (%) of cells in each gate is indicated. Representative of three experiments.

C. Cytospin preparations of cells in the experiment described in (B) at t= 24 h. Control cells, but not cells expressing shRNA to CD71, have started to express hemoglobin in their cytoplasm (brownish color, see arrow). Stained with Giemsa-diaminobenzidine; scale bar is 20 μ.

A

B

C
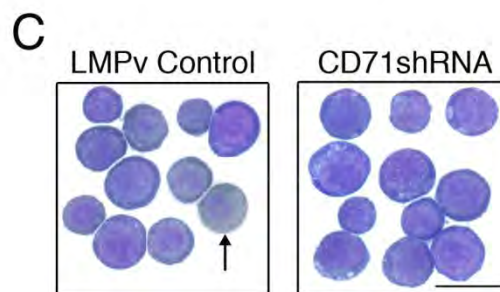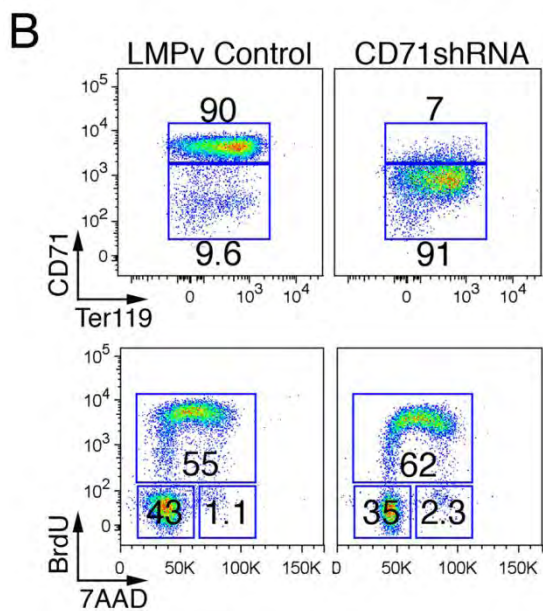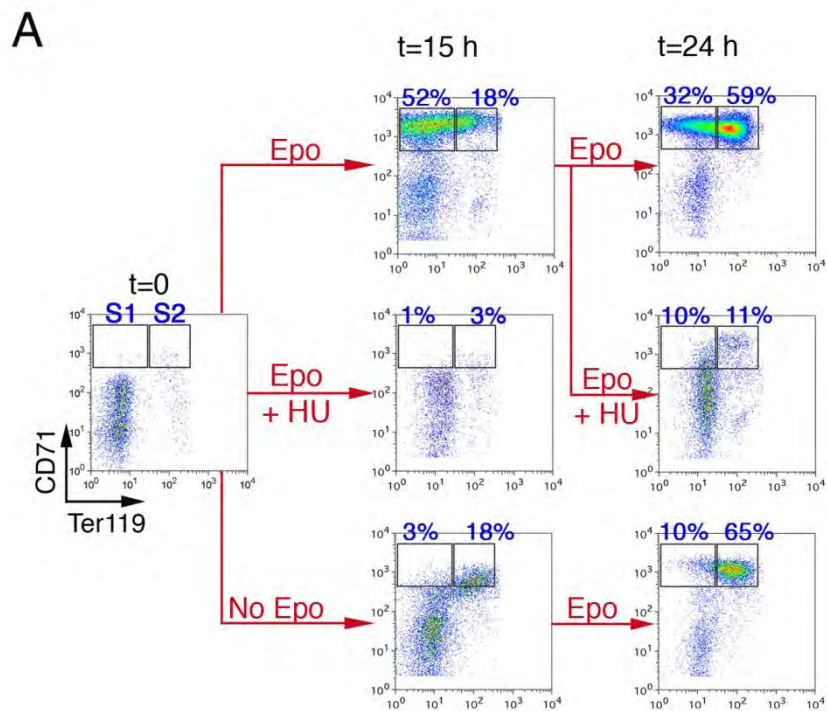
Figure 2.S3. Supplemental data to Figure 2.3.

A. Quantitative RT-PCR analysis of PU.1, Gata-2, and Gata-1 mRNAs in Mac-1$^+$, S0, or S1 cells sorted from fetal liver. mRNAs were normalized to the β-actin mRNA in each sample. Data are mean ± SD of 2 independent experiments.

B. Total RNA per cell in sorted fetal liver subsets. Total isolated RNA for each subset was measured using spectrophotometric optical density and divided by the number of sorted cells. Mean ± SE of five independent sort experiments.

C. Flow-cytometry histograms of PU.1 protein levels in sorted fetal liver subsets S0 to S3. Fresh fetal liver cells were fixed, permeabilized, and stained for CD71, Ter119, and PU.1. Bar graph indicates the PU.1 median fluorescence intensity of each subset. Representative of 2 independent experiments.

D. EpoR mRNA increases at the S0 to S1 transition. Results normalized to β-actin and expressed as a ratio to S0.

E, F. Experimental design for (E) and (F): sorted S1 cells (t= 0 h) were incubated in Epo and in the presence or absence of aphidicolin, for 10 h. Cells were then washed free of aphidicolin and Epo incubation continued for a further 10 h.

E. Arrest of S-phase progression in S1 does not affect the mRNA expression of erythroid-specific genes, β-globin, Alas2, and Band3. mRNA was measured by qRT-PCR, normalized to β-actin, and expressed as relative to mRNA at t= 0. Data are mean ± SE from 3 independent experiments.

F. Arrest of S-phase progression in S1 does not affect expression of Ter119. CD71/Ter119 profiles showing that Ter119 was up-regulated between t= 0 and t= 10 h regardless of the presence of aphidicolin. The fraction (%) of cells in S1 (left gate) or S2 (right gate) is indicated. Note that the larger cell size resulting from aphidicolin-mediated block of DNA replication (Figure 2.3F, main manuscript) is likely responsible for the higher Ter119 signal in the aphidicolin treated cells (cell surface Ter119 would be expected to increase in proportion to the square of the cell"s radius). Ten hours following the release of the block, there is no significant difference in Ter119 expression between treated and untreated cell samples (t= 20 h).

G. Effect of mimosine-mediated S-phase arrest on down-regulation of PU.1 and Gata-2 during the S0 to S1 transition. Experiment and mRNA measurement were as in Figure 2.3D, main manuscript, with the exception that mimosine was used in place of aphidicolin. Data are mean ± SE of three independent experiments.

H. Effect of aphidicolin-mediated S-phase arrest on Lmo2 and Nfe2 mRNAs during the S0 to S1 transition. Experiment and mRNA measurements as in Figure 2.3B, main manuscript; mRNA levels are expressed as a ratio to mRNA at t= 10 h in the "Epo only" control. Data are from two independent experiments.

I. Erythroblast morphology of S0 cells transduced with p57T329A-ICD4 or with control vector MICD4 at t= 32 h of incubation in Epo. Experiment as described in Figure 2.3G, main manuscript. Cytospins were stained with Giemsa-diaminobenzidine; scale bar= 20 μ.
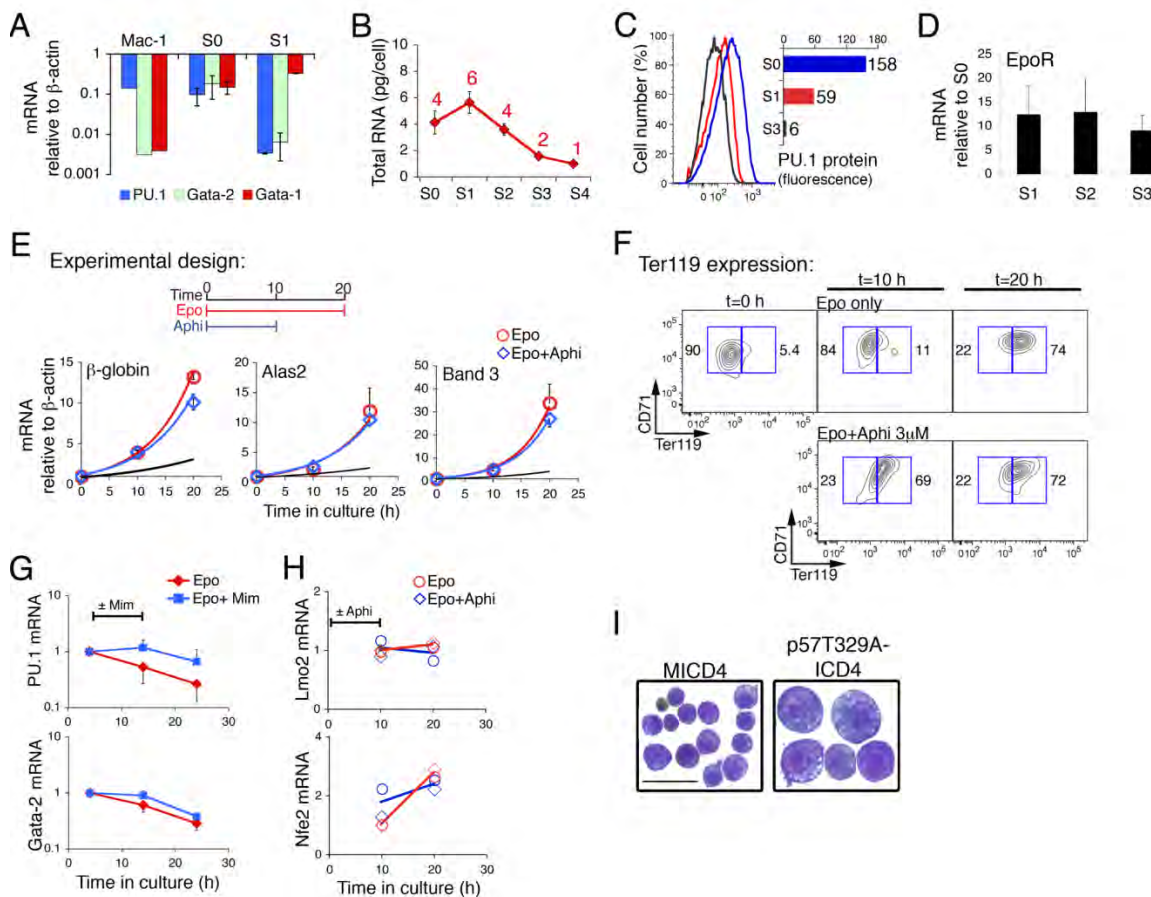
Figure 2.S4. Supplemental data to Figure 2.4.

A–B. PU.1 protein levels in cells transduced with PU.1-IRES-hCD4 is proportional to the level of the hCD4 reporter in the same cells. Please see also summary of these data in Figure 2.4D.

A. Expression profiles of the hCD4 reporter in cells transduced with either PU.1-ICD4, or MICD4 control vector, or uninfected control cells, following 24 h of culture in Epo. Vertical, narrow gates each containing cells of relatively uniform hCD4 expression, numbered 1 to 4, are shown and are used in the analysis in sections (B) and in Figure 2.4D.

B. Flow-cytometry histograms of cells transduced with MICD4 control vector (top panels) or with PU.1-ICD4 (lower panels) for each individual hCD4 gate (numbered 1 to 4), showing the percent of PU.1 positive cells. An overlay of the same flow cytometry histograms is shown on the right.

C. Expression of GATA-1-IRES-hCD4 (GATA-1-ICD4), GATA-2-ICD4, and MICD4 at 24 h of Epo culture. The hCD4 + cell gate is indicated in black. Associated with Figure 2.4E in the main text.

D. qRT-PCR analysis of Gata-1, Gata-2, and PU.1 mRNAs in S0 cells transduced with MICD4 control vector or gene-ICD4 at 24 h of Epo culture, as compared with endogenous levels in freshly sorted S0 and S1 subsets. mRNA was normalized to the β-actin mRNA and expressed as a ratio to the S0 subset. Data are mean ± SD of 3 independent experiments (the transduction efficiency was >90% for PU.1-ICD4 and >50% for Gata-1-ICD4 and Gata-2-ICD4).

E. Preventing down-regulation of PU.1 does not halt down-regulation of p57^{KIP2} mRNA. Quantitative RT-PCR analysis of p57^{KIP2} mRNA in S0 cells retrovirally transduced with PU.1-ICD4 or with control vector MICD4 as described in Figure 2.4A–C, and incubated for 24 h in Epo. mRNA measurements from two independent experiments. In the first experiment, hCD4-positive cells were sorted by flow-cytometry at t= 0 (first experiment). In the second experiment transduction efficiency exceeded 90%, as judged by hCD4 expression. mRNAs were also measured in freshly sorted S0 and S1 subsets.
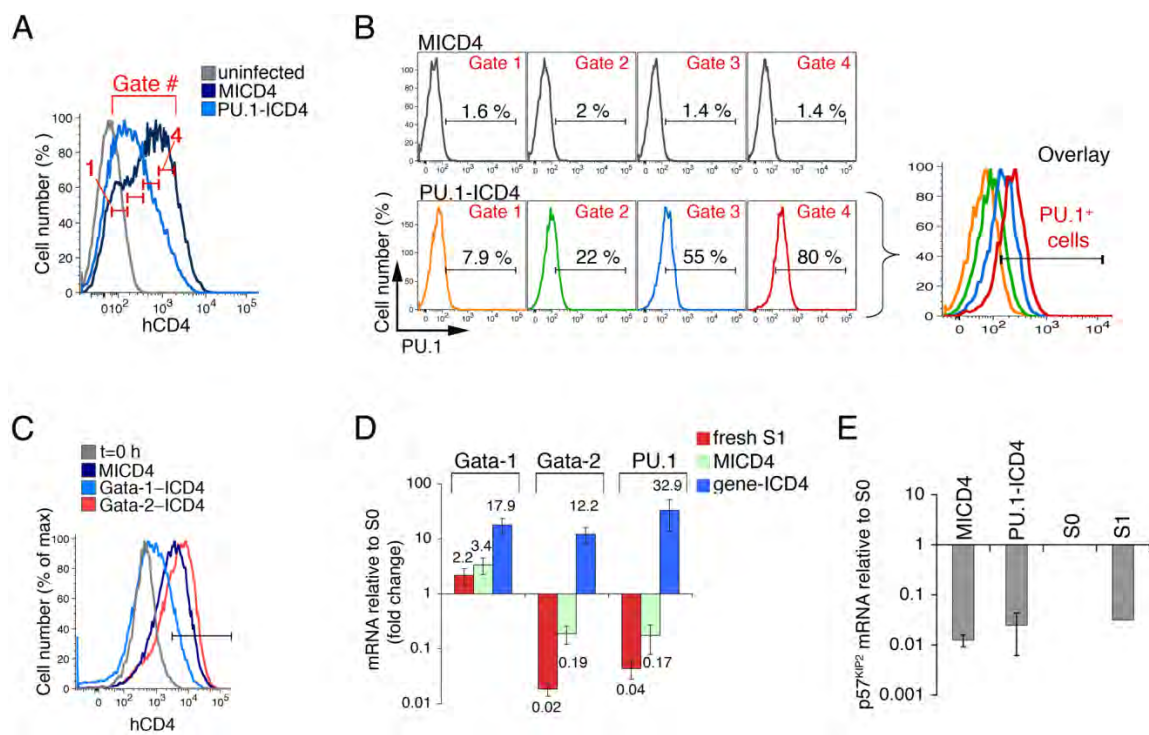
CHAPTER III

Global DNA Demethylation During Erythropoiesis

Statement of Contribution

I performed the work presented in Chapter 3 unless otherwise noted. Ramona Pop performed the work in Figure 3.11 and contributed to Figure 3.12A and 3.13A. Christoph Bock and Alexander Meissner performed the reduced representation bisulfite sequencing and analysis presented in Figure 3.5, 3.6A-C, 3.14, and 3.16C. Statistics in Figures 3.1, 3.2, and 3.3 were performed with assistance from Stephen Baker. This chapter is an extended version of work that is currently under review with the following title and authors:

Jeffrey R. Shearstone, Ramona Pop, Christoph Bock, Alexander Meissner, and Merav Socolovsky. "Global DNA Demethylation During Erythropoiesis In Vivo"

Abstract

In the mammalian genome, 5"-CpG-3" dinucleotides are frequently methylated, correlating with transcriptional silencing. Genome-wide waves of demethylation are thought to occur only twice during development, in primordial germ cells and in the pre-implantation embryo [330]. They are followed by *de novo* methylation, setting up a pattern that is inherited throughout development. No global methylation changes are thought to occur during further somatic development, although methylation does alter at gene-specific loci, contributing to tissue-specific patterns of gene expression. Here we studied DNA methylation in differentiating mouse erythroblasts *in vivo* using several approaches including genomic-scale reduced representation bisulfite sequencing (RRBS). Surprisingly, demethylation at the erythroid-specific β-globin locus was coincident with a wave of global DNA demethylation at most genomic elements, including repetitive elements and genes silenced in erythropoiesis. Over 30% of total methylation is irreversibly lost during erythroid differentiation. Demethylation occurred through a passive mechanism, requiring the rapid DNA replication triggered with the onset of erythroid terminal differentiation. Global loss of DNA methylation was not associated with a global increase in transcription, as determined by GeneChip analysis. We propose that global demethylation is a consequence of cellular mechanisms required for the rapid demethylation and induction of β-globin and other erythroid genes. Our findings demonstrate that, contrary to previously held dogma, DNA demethylation can occur globally during somatic cell differentiation, providing a new experimental model for the study of global demethylation in development and disease.

Introduction

Histones or DNA can be modified by the enzymatic addition of various chemical groups, resulting in recruitment or exclusion of associated proteins that have profound effects on local chromatin structure. Because these modifications fundamentally do not alter the DNA sequence and because they are mitotically and/or meiotically heritable, they have been termed „epigenetic" modifications. One such epigenetic mark, DNA methylation, is simply the presence of a methyl group on the 5"-carbon ring of cytosine (5mC). In mammals, the vast majority of DNA methylation (>99%) is found to occur symmetrically at the dinucleotide palindrome cytosine-guanine (CpG) [1, 2]. DNA methylation leads to transcriptional inactivity and/or heterochromatin formation and plays an important role in X-chromosome inactivation, parental-specific silencing of imprinted genes, suppression of transposable elements, and silencing of gene-specific loci.

The formation of enucleated red cells, a process known as definitive erythropoiesis, first occurs in the murine fetal liver between embryonic days 11 and 15 (E11-E15), and is dependent on the hormone erythropoietin (Epo). The cell surface markers CD71 and Ter119 can be used to identify differentiation-stage-specific erythroblasts directly within freshly harvested mouse fetal liver [104, 105, 342]. With this approach the fetal liver can be divided into six subsets, S0 to S5, which form an erythroid developmental sequence (Figure 3.1A). Subsets S1 to S5 are composed entirely of erythroid cells. When plated *in vitro*, S1 cells differentiate into S5 cells within 48 hours, undergoing three cell divisions [105]. The S0 subset is composed largely (70%) of early

erythroid progenitors prior to the onset of Epo dependence, which can be further enriched (>95%) by excluding non-erythroid hematopoietic cells with lineage-specific cell surface markers [105]. Erythroid maturation is associated with decreasing cell size, nuclear condensation and positive staining for hemoglobin with diaminobenzidine.

Proper establishment and maintenance of 5mC patterns are required for mammalian development and normal functioning of somatic cells. Methylation generally occurs at 70-80% of all CpG sites, a consistent finding across a wide-range of tissues. Importantly, despite tissue-specific alterations in DNA methylation within gene regulatory regions, the overall 5mC level has been thought to remain constant in all somatic cells. Physiological examples of genome-wide methylation loss have been limited to the pre-implantation embryo and primordial germ cells, where demethylation is thought to induce pluripotency related genes or reset parental imprints, respectively. Global DNA hypomethylation is also a well-documented feature of cancer that is thought to lead to genomic instability and aberrant gene activation. The mechanism and functional significance of global methylation loss remains largely unknown and is an area of intensive study.

In Chapter II we demonstrated that the transition of erythroid progenitors from S0 to S1 marks a key step in erythropoiesis that is coordinated with S-phase entry and requires DNA replication [105]. This transition comprises several rapid and synchronous commitment events that include the onset of Epo dependence, activation of the erythroid master transcriptional regulator GATA-1, and a switch in chromatin conformation at the locus control region (LCR) of the β-globin gene [105]. In this chapter, we report on a

completely unexpected finding that emerged while characterizing the S0 to S1 transition: DNA demethylation at the erythroid-specific β-globin locus is coincident with progressive and irreversible genome-wide methylation loss. We describe the magnitude, timing, and location of global demethylation during erythropoiesis, and provide evidence of its mechanism and functional significance. These findings demonstrate that DNA demethylation can occur globally during somatic cell differentiation, and establishes fetal liver erythropoiesis as a new experimental model for studying genome-wide methylation loss in development and in disease.

Results

DNA methylation is lost at the β-globin LCR and at loci of genes silenced in erythropoiesis

We examined DNA methylation at the β-globin LCR in freshly-sorted cells from subsets S0 to S4/5. Genomic DNA from each subset was subjected to bisulfite conversion, and methylation levels at each of 6 CpG dinucleotides in Hypersensitive Sites 1 and 2 (HS1, HS2) were measured by pyrosequencing of PCR products amplified from the HS1 and HS2 regions (Figure 2.6A, Figure 3.1B). All CpGs at HS1 and HS2 were undergoing demethylation, beginning with the transition from S0 to S1 and continuing throughout erythroid maturation (Figure 2.6A, Figure 3.1B) [105]. LCR demethylation preceded induction of the β-globin gene (Figure 3.1B) and correlated with indicators of active chromatin at the LCR, including increased DNaseI hypersensitivity (Figure 2.5C), increased H3K4 methylation, and decreased H3K27 methylation (Figure 2.5F, G) [105].

In parallel, we also examined the methylation status of genes whose expression declines during erythroid differentiation. We first investigated PU.1, a transcriptional repressor that stimulates self-renewal and prevents differentiation in erythroid progenitors [293, 314]. PU.1 mRNA declines dramatically ($\approx$30-fold) at the S0 to S1 transition (Figure 2.3A) [105] and is undetectable in the S4/5 subset. DNA methylation around the PU.1 promoter was previously shown to correlate with silenced PU.1 expression [343, 344]. Therefore, we hypothesized that methylation at this locus might increase during erythropoiesis. Surprisingly, each of the four CpG dinucleotides within the PU.1

promoter amplicon underwent significant and continuous demethylation as cells transitioned from S0 to S4/5 (Figure 3.1C, p<0.001). Using the average 5mC level of the amplicon, the PU.1 region underwent an absolute methylation loss of 24%, representing a 50% reduction in the initial S0 methylation level.

We similarly investigated the Fas gene, a pro-apoptotic cell surface receptor and known negative regulator of erythropoiesis, whose expression also declines with erythroid differentiation [104]. We found significant (p<0.001) and progressive demethylation at each of the 7 CpG dinucleotides within a regulatory region whose methylation was previously shown to correlate with silencing of Fas expression (Figure 3.1D) [345]. This region underwent a locus average absolute methylation loss of 12%, representing a 19% reduction in the initial S0 methylation level. Data was not obtained for the S4/5 subset.

DNA methylation is lost at differentially methylated regions of imprinted genes

The unexpected results at the PU.1 and Fas loci suggested that there may be global loss of DNA methylation during erythropoiesis. To explore this possibility further, we examined the methylation status of a number of imprinted genes. In somatic cells, imprinted genes are predominantly expressed from one parental allele as a result of parental-allele-specific epigenetic modifications that include stable methylation of differentially methylated regions (DMRs) [13]. We assayed the H19 DMR, one of four DMRs that coordinate the paternal-specific expression of the Igf2 gene [57, 346]. We amplified bisulfite converted genomic DNA from S0 and from S4/5 subsets with gene-specific primers in a region containing 10 CpG dinucleotides, ligated the PCR products

into a cloning vector and sequenced individual clones (Figure 3.2A, left panel). In the

case of S0 cells, clones formed a bimodal distribution, with approximately half being

fully methylated, and the other half almost completely unmethylated, a pattern consistent

with imprinting. However, in the S4/5 subset all but one clone had at least one or more

demethylated CpGs (Figure 3.2A, left panel); the average methylation level for all clones

declined from 61% in S0 to 42% in S4/5 (p=0.007, two-tailed Mann-Whitney test).

Pyrosequencing of the bulk PCR product for each erythroid subset at the H19 DMR was

consistent with cloning results, showing a significant (p<0.001) and progressive decrease

in the average locus methylation from 50% in S0 to 36% in S4/5 (Figure 3.2A, right

panel), representing a 28% reduction in the initial S0 methylation level.

The H19 DMR assay includes genomic regions that are binding sites for the

enhancer-blocking protein CTCF [48, 57]. CTCF prevents expression of Igf2 by binding

unmethylated DNA within the H19 DMR, thereby insulating the upstream Igf2 gene

promoter from the effects of the downstream enhancer. Therefore, demethylation of the

H19 DMR could potentially lead to CTCF binding on the active Igf2 allele and

subsequent Igf2 repression. Consistent with this hypothesis, we found that Igf2 mRNA is

down-regulated approximately 10-fold during differentiation from S0 to S3 using

Affymetrix GeneChips (data not shown). However, the observed Igf2 repression could be

due to another gene regulatory mechanism besides CTCF binding and this line of

research was not pursued futher.

We examined two additional imprinted loci by pyrosequencing and again found a

significant and progressive loss of average DNA methylation at each locus, as well as at

individual CpG dinucleotides within each locus (Figure 3.2B, C) [347-349]. Paternally

expressed gene 3 (Peg3) average locus methylation decreased from 43% in S0 to 36% in

S4/5 (p<0.001), a 15% drop in initial S0 methylation level. Similarly, Kcnq1ot1 average

locus methylation decreased from 44% to 37% (p<0.01).

Global 5mC content is reduced during erythropoiesis

Unexpected methylation loss within gene-specific regions is consistent with

genome-wide demethylation during erythropoiesis. To support and extend these

observations, we measured DNA methylation levels using three distinct assays of global

5mC content. First, we examined 5 CpG dinucleotides within the 5" tandem repeat region

of the long interspersed nuclear element (LINE-1) using bisulfite conversion with

pyrosequencing [350]. LINE-1 are repetitive retrotransposon elements of variable lengths

that are distributed at approximately 100,000 chromosomal sites in the mouse genome

[351]. Additionally, LINE-1 regions are expected to be highly and stably methylated in

order to silence retrotransposon activity [8]. These characteristics have made LINE-1

regions a widely used surrogate marker for global DNA methylation [352-354]. As

expected, LINE-1 regions were highly methylated in S0 cells, with a locus average

methylation of 90% (Figure 3.3A). Following the transition from S0 to S1, methylation

levels of the LINE-1 CpGs underwent a continuous decline, to a locus average of 70% in

S4/5 (p<0.001), with a maximal decline for CpG2 from 91% in S0 to 65% in S4/5

(p<0.001) (Figure 3.3A). On average the LINE-1 regions lost 22% of their initial S0

methylation levels.

Second, we used the LUminometric Methylation Assay (LUMA) [355] (Figure
3.3B) which compares genome-wide cleavage at CCpGG sites by each of the
isoschizomers HpaII and MspI, which are methylation sensitive and insensitive,
respectively [356]. The CCpGG site recurs approximately 2.5 x 10^6 times throughout the
diploid mouse genome, making LUMA an appropriate measure of global DNA
methylation [357]. Using LUMA, 70% of sites in S0 cells were methylated, a result
consistent with the expected level of global methylation in somatic cells [36]. S0 cells
underwent a global, progressive and significant (p<0.001) loss of DNA methylation upon
differentiation, reaching a methylation level 50% in S4/5 cells (Figure 3.3B).

Finally, we measured binding of a 5-methylcytosine-specific antibody to genomic
DNA purified from S0, S1 and S3 cells, using an Enzyme-Linked Immunosorbent Assay
(ELISA) (Figure 3.3C). We found 70% of CpG dinucleotides in S0 cells were methylated
using the ELISA, a finding consistent with the results from LUMA. This approach also
showed a global, progressive and significant (p<0.01) loss of DNA methylation with
erythroid differentiation (Figure 3.3C).

**Global demethylation begins at the S0 to S1 transition, is a general feature of
erythropoiesis, and persists in extruded nuclei.**

We used LINE-1 regions to assess the global methylation levels of other cell
types within the fetal liver. We isolated Lin⁻Sca1⁺cKit⁺ (LSK) cells, a population of cells
that are at an earlier developmental stage than S0 erythroid progenitors. LSK cells
contain early hematopoietic progenitors (including short and long-term hematopoietic
stem cells and hematopoietic progenitors) and the common myeloid and lymphoid

progenitors (CMP and LMP). We also isolated differentiated cells from other hematopoietic lineages; Mac1$^+$ cells, composed of granulocytes and macrophages, and CD41$^+$ cells, mostly composed of megakaryocytes. Finally, to determine if global methylation loss is dependent on EpoR signaling, we isolated erythroid progenitors from EpoR$^{-/-}$ fetal livers. LSK, CD41$^+$, Mac1$^+$ and EpoR$^{-/-}$ cells each contained a high level of DNA methylation, equivalent to the level found in S0 cells (Figure 3.4A). These data suggest that global demethylation is a unique property of the erythroid lineage that begins specifically at the EpoR-dependent transition from S0 to S1.

Next, we characterized methylation levels in primitive erythroid cells in the embryo and definitive erythroid cells of adult bone marrow. In mouse, primitive erythropoiesis occurs in the yolk sac at embryonic day 7.5 (E7.5), lasts only a few days, and results in nucleated erythrocytes. We isolated nucleated primitive erythrocytes from the circulation of E13.5 embryos and found that they possessed methylation levels even lower than that of S4/5 cells (Figure 3.4A). Near the time of birth, the site of definitive erythropoiesis moves from the fetal liver to the bone marrow, where it persists throughout life. We sorted increasingly differentiated erythroblasts from adult bone marrow and found progressive hypomethylation (Figure 3.4B). ProE and EryB/C populations had methylation levels that were equivalent to their S2 and S4/5 fetal liver counterparts, respectively. Together, these findings suggest that global demethylation is a general feature of erythropoiesis, neither limited to the definitive lineage nor the fetal liver.

Definitive erythroblasts enucleate by nuclear extrusion, generating reticulocytes and small, nucleated cells with a thin rim of cytoplasm, known as pyrenocytes, which are

eventually phagocytosed by macrophages. Pyrenocytes, like S5 cells, have exited the cell cycle (Figure 2.1E). We sorted these extruded nuclei by flow cytometry and determined LINE-1 methylation levels (Figure 3.4A). We found that pyrenocytes had methylation levels equivalent to S4/5 cells. This result suggests that global demethylation *in vivo* is an irreversible process that persists even after cell cycle exit and enucleation.

Genome-wide reduced representation bisulfite sequencing of erythroid subsets

To fully characterize DNA methylation changes during erythroid differentiation on a genomic scale, we used reduced representation bisulfite sequencing (RRBS) [42]. In this method, genomic DNA is digested by MspI, 40-220 bp fragments are selected for bisulfite conversion, and converted fragments are sequenced from each end. RRBS provides methylation data for approximately 1 million distinct CpGs, or in other words, 5% of all CpGs in the mouse genome. CpG coverage is not distributed evenly. For example, RRBS covers 70% of all promoter regions, 30-80% of CpG islands (depending on island definition), 25% of island shores, 25% of enhancers, and 30% of all 5 kb sliding, non-overlapping windows [358].

We performed RRBS for each of the sorted erythroblast subsets S0, S1, S3 and S4/5 in biological duplicate (Figure 3.5A). In the shown scatterplots (Figure 3.5A) data points represent 5 kb non-overlapping tiles with sufficient sequencing coverage, i.e. at least 5 CpGs, each with at least 5 sequencing reads. The coordinates for each data point are the mean methylation level for each such tile, in each of subsets S1 to S4/5 in turn, plotted against the mean methylation level of the same 5 kb tile in the least mature, S0 subset. The gradual fall of nearly all data points away from the middle diagonal with

increasing erythroid maturation indicates a progressive and orderly loss of DNA

methylation. Overall, the median methylation level genome-wide in S0 cells is 79%,

falling to 55% in the most mature, S4/5 subset, just prior to erythroblast enucleation

(Figure 3.5B). Therefore, erythroblasts lose 30% of their initial S0 methylation during

differentiation.

Next, each 5 kb window was classified, using statistical analysis of sample

replicates and the individual CpGs within each window, as either hyper- or hypo-

methylated in a comparison of S0 to each of S1, S3, and S4/5 subsets. A 5 kb window

was designated "no change" if the false discovery rate (FDR) >0.1; „hypomethylated" if

FDR <0.1, p <0.001, and methylation change >20%; „hypomethylated (weak)" if FDR

<0.1, methylation change >0% and not „hypomethylated" as defined above.

„Hypermethylation" criteria were defined similarly, but with the methylation change in

the opposite direction. By these criteria, less than 0.05% of the 62,844 unique 5 kb

windows became hyper or weakly hyper-methylated during erythroid differentiation into

S4/5 cells (Figure 3.5C, D). In contrast, 67% of 5 kb windows became hypo- or weakly

hypo-methylated (Figure 3.5C), indicating that the majority of the genome undergoes

some significant methylation loss.

Further RRBS analysis revealed that the loss of DNA methylation takes place

across a broad range of genomic elements. Methylation losses were observed at

retrotransposable elements, including LINEs, short interspersed nuclear elements

(SINEs), and long terminal repeats (LTRs) (Figure 3.6A); at CpG islands, defined by the

Gardiner-Garden criteria [41] (Figure 3.6B); at island shores, defined as 2-kb genomic

regions that are directly adjacent to *bona fide* CpG islands (Figure 3.6B) [359]; at

enhancers, defined based on histone H3K4 dimethylation peaks (Figure 3.6B) [42, 358];

and at imprinted regions manually curated from the literature (Figure 3.6B) [358]. The

largest losses in DNA methylation are in regions whose initial methylation level in S0 is

high. This includes promoters with low CpG frequency (<1%) (Figure 3.6C). Conversely,

promoters with high CpG frequency (>2%), having lower initial methylation levels in S0,

show a smaller methylation loss (Figure 3.6C).

The fractional loss across most genomic elements is similar; all elements lose

25%-30% of their initial S0 methylation level by the time cells have differentiated into

mature S4/5 erythroblasts (Figure 3.6D). The principal exception to this is imprinted

regions, where the median loss is 14% of the initial methylation in S0 (Figure 3.6D). All

of the RRBS data, including genome browser tracks with genomic locations, has been

placed on a supplementary website, http://erythrocyte-demethylation.computational-

epigenetics.org.

Genome-wide and β-globin LCR demethylation are linearly correlated but occur at

different rates

β-globin LCR and genome-wide demethylation are initiated at the transition from

S0 to S1 and continue throughout erythroid differentiation (Figure 2.6A and Figure 3.4A,

respectively). These similarities prompted us to investigate the relationship between

gene-specific and global demethylation events. We found a linear correlation throughout

erythroid maturation between DNA methylation levels at the β-globin LCR and

methylation levels of various genomic regions, including LUMA (R^2=0.95), the

pyrosequencing assays for LINE-1 (R^2=0.97), H19 DMR (R^2=0.98), or PU.1 (R^2=0.94), as well as the RRBS measurements at enhancers (R^2=0.98), CpG islands (R^2=0.96), imprinted genes (R^2=0.60), and 5 kb windows (R^2=0.98) (Figure 3.7). Therefore, the fraction of total methylation loss that occurs at a given differentiation stage is similar for all loci. These strong linear correlations suggest that a common mechanism is responsible for demethylation at both erythroid-specific genes and genome-wide loci.

The slopes of the various regression lines suggest that the rate of demethylation at nearly all global sites is only 30-50% of the rate at the β-globin LCR (Figure 3.7). This result is consistent with the observation that the β-globin LCR loses 65-80% of initial S0 methylation during differentiation to S4/5, while genome-wide levels are reduced by only 30% (Figure 3.6D). Therefore, while genome-wide and erythroid-specific demethylation might be driven by a single mechanism, methylation loss is significantly enhanced at the β-globin LCR.

Genes implicated in active and passive demethylation are differentially regulated during erythropoiesis

We examined whether regulators of DNA methylation were altered in the course of erythroid differentiation. Expression of the maintenance methyltransferase Dnmt1 appeared unaltered at the mRNA and protein levels throughout differentiation (Figure 3.8A, B). Its interacting partner Uhrf1, which targets Dnmt1 to hemi-methylated DNA, was also unchanged at the mRNA level (Figure 3.8A) [111, 112, 360]. By contrast, the *de novo* methyltransferases Dnmt3a and Dnmt3b, which may also contribute to maintenance methylation, were down-regulated 10- and 100-fold at the mRNA level, respectively,

with the progression from S0 to S3 (Figure 3.8A) [135]. Dnmt3a and 3b mRNA was undetectable in the S4/5 subsets and the mRNA for their interacting co-factor, Dnmt3L, was not detected in any erythroid subset. Dnmt3a and 3b protein decreased during the S0 to S1 transition and became virtually undetectable in the S3 subset (Figure 3.8B).

We also investigated regulators recently implicated in active demethylation [163, 173-175]. Transcripts for Apobec1, a cytidine deaminase, and Tdg, a T/G glycosylase, were present in erythroid cells, but not significantly up-regulated upon differentiation (Figure 3.8A). Aicda, another cytidine deaminase, was not detected in any erythroid subset. The T/G glycosylase Mbd4 showed a modest, but significant, 2-fold up-regulation during erythroid differentiation. Gadd45a, implicated in NER and BER-mediated demethylation, was progressively and significantly up-regulated during differentiation, reaching a 15-fold greater mRNA abundance in the S3 subset (Figure 3.8A).

Sustained expression of Dnmt3a or Dnmt3b does not prevent DNA demethylation

Down-regulation of Dnmt3a and Dnmt3b during differentiation may have contributed to global demethylation. We examined the effect of exogenous expression of these enzymes in differentiating erythroblasts *in vitro*. Erythroid differentiation may be reproduced *in vitro* by culturing freshly sorted S0 cells in the presence of Epo; the cells transition into S1 within 10 to 15 hours and continue differentiating for 48-72 hours. Dnmt3a, Dnmt3b, or the catalytically inactive mutant Dnmt3b1:PC (courtesy of En Li, Novartis) [16] were subcloned into a bicistronic retroviral vector where the gene of interest is linked through an internal ribosomal entry site (IRES) to green fluorescent protein (GFP). Sorted S0 cells were infected for 16 hours in the presence of interleukin-3

(IL3) and stem cell factor (SCF) to allow time for retroviral construct expression. IL3 and SCF sustain viability of S0 progenitors, but they do not support erythroid differentiation (Figure 3.9A). At 16 hours, cells were shifted to medium containing erythropoietin (Epo) and allowed to differentiate for 24 hours, at which point GFP-positive cells were sorted and assessed for methylation status of LINE-1 and the β-globin LCR loci. Retroviral supernatants of the empty vector or Dnmt3a showed high transduction efficiency relative to Dnmt3b1 supernatants. Therefore, sorting experiments were done with supernatant dilutions of 1:200 for the empty vector and 1:25 for Dnmt3a in order to obtain equivalent levels of GFP expression to cells transduced with Dnmt3b1 (Figure 3.9A).

Cells infected with exogenous Dnmt3a or Dnmt3b1 expressed the respective mRNA in all subsets at a constant level, while in the empty vector control Dnmt3 expression decreased 10-fold during differentiation (Figure 3.9B). Therefore, retroviral transduction prevented differentiation-induced down-regulation of Dnmt3a and Dnmt3b transcripts. However, exogenous expression of Dnmt3a or Dnmt3b1 did not prevent demethylation at global LINE-1 regions (Figure 3.9C). Similarly, exogenous expression of Dnmt3a did not prevent demethylation at the β-globin LCR (Figure 3.9C). Exogenous expression of Dnmt3b1 led to small but significant elevation in β-globin LCR methylation in the S3 subset (Figure 3.9C). However, this observation is not the result of DNA methyltransferase activity, since the catalytically inactive Dnmt3b1:PC mutant had the same effect (Figure 3.9C).

Unexpectedly, we noted that ectopic expression of Dnmt3b1, but not DNMT3a, had an inhibitory effect on erythroid differentiation. After 24 hours in Epo, approximately

50% of empty vector control or Dnmt3a-expressing cells had differentiated into S3 cells. In contrast, only 30% of Dnmt3b1 expressing cells had become S3 cells (Figure 3.9D). This effect was not dependent on DNA methylation activity as it was also observed in cultures expressing Dnmt3b1:PC (Figure 3.9D). Back-gating analysis demonstrates that GFP positive S3 cells from Dnmt3b1-tranduced cultures have a significantly higher number of DAPI (4',6-diamidino-2-phenylindole) positive cells as compared to Dnmt3a-expressing or empty vector control cells (Figure 3.9E). This finding suggests that sustained expression of Dnmt3b1 in differentiating erythroblasts leads to cell death.

Suppression of Gadd45a or Mbd4 does not prevent DNA demethylation

Up-regulation of Gadd45a or Mbd4 during differentiation may have contributed to global demethylation. We therefore examined the effect of blocking expression of these transcripts using retroviral short hairpin RNAs (shRNA) using an experimental design similar to that used for the Dnmt3 exogenous expression experiments described above. We transduced fresh S0 cells with a non-silencing control shRNA or shRNAs directed at Gadd45a or Mbd4 for 16 hours in the presence of SCF and IL-3. Cells were then cultured for 24 hours in the presence of Epo, followed by sorting of the GFP positive, Ter119$^+$ population (Ter119$^+$ =combined S1, S2, and S3 sorting gate). The infection rate using shRNA retrovirus was nearly 100% at 16 hours (data not shown).

After 24 hours in Epo, Gadd45a was induced 30-fold in Ter119$^+$ cells expressing the non-silencing control hairpin (Figure 3.10A). In contrast, Gadd45a remained unchanged, or was even further suppressed, during differentiation in cells expressing shRNAs directed against Gadd45a (Figure 3.10A). Similarly, Mbd4 shRNA prevented

the mild 2-fold induction of Mbd4 upon differentiation and further reduced Mbd4 mRNA to only 2% of its S0 expression level (Figure 3.10A). Preventing Gadd45a or Mbd4 expression had no obvious effect on erythroid differentiation as judged by cell surface markers Ter119 and CD71 (data not shown). Knockdown of Gadd45a and Mbd4 did not prevent 5mC loss at LINE-1 regions or the β-globin LCR (Figure 3.10B). These findings suggest that Mbd4 and Gadd45a are not involved in erythroid DNA demethylation.

The rate of DNA replication and fraction of S-phase cells increases during erythropoiesis

We recently found that the onset of erythroid differentiation at the transition from S0 to S1 occurs during S-phase of a developmentally-specific cell cycle, and that essentially all S1 cells are in S-phase (Figure 3.11A, B) [105]. Further, when subjecting fetal livers to a brief pulse of the nucleotide analogue bromodeoxyuridine (BrdU) *in vivo*, S-phase cells at the onset of differentiation in S1 and S2 incorporate 50% more BrdU than S-phase cells in previous cycles in S0 (Figure 3.11A, C) [105]. Importantly, this increased BrdU incorporation is indicative of a higher rate of DNA synthesis [361, 362]. DNA replication rate remains slightly elevated in S3 cells relative to S0, and by S4/5 cells have exited the cell cycle just prior to enucleation (Figure 3.11). These alterations in DNA replication rate and proportion of S-phase cells during the S0 to S1 transition coincide with the initiation of DNA demethylation, suggesting a potential role for DNA replication in the global and gene-specific loss of 5mC observed during erythropoiesis.

DNA demethylation is prevented by arresting the cell cycle with mimosine

To investigate the potential role of DNA replication in demethylation, we sorted fresh S0 cells and incubated them in the presence of mimosine, a plant amino acid that blocks cell cycle progression in late G1 [305]. S0 cells cultured in Epo undergo gradual demethylation at the β-globin LCR and LINE-1 regions (Figure 3.12A, black circles, t=0 to 24 h). The addition of 1 mM mimosine for 10 hours prevented methylation loss at these loci (Figure 3.12A, red circles, t=4 to 14 h). Demethylation resumed once mimosine was removed (Figure 3.12A, red circles, t=14 to 24 h). We performed additional experiments with sorted S0 cells cultured in Epo only or Epo plus mimosine for 16 hours, allowing more time for demethylation. The presence of mimosine, and subsequent G1 arrest (Figure 3.12B), prevented methylation loss at the β-globin LCR, LINE-1, PU.1, and H19 DMR loci (Figure 3.12C). We previously have shown that the erythroid differentiation program also requires DNA replication at the S0 to S1 transition [105]. Therefore, while these results suggest a DNA replication-dependent, passive demethylation mechanism, the possibility remains that blocking DNA replication in S0 cells is simply preventing the initiation of an active demethylation mechanism linked to the erythroid differentiation program.

DNA demethylation requires rapid DNA replication

We examined the potential role of DNA replication in demethylation using aphidicolin, an inhibitor of DNA polymerase α that arrests S-phase progression [304]. Freshly sorted S0 cells cultured in Epo undergo gradual demethylation at the β-globin LCR and LINE-1 regions (Figure 3.13A, black circles, t=0 to 46 h). The addition of 3 μM

aphidicolin prevented methylation loss at these loci (Figure 3.13A, red circles, t=0 to 10 h) and demethylation resumed once aphidicolin was removed (Figure 3.13A, red circles, t=10 to 46 h). These results were consistent with similar experiments using mimosine, suggesting that DNA replication is a prerequisite for 5mC loss.

We tested this possibility further by slowing down the rate of DNA replication with aphidicolin in different erythroid subsets. We incubated sorted S0, S1, or S2 cells in Epo or Epo plus a range of aphidicolin concentrations for 16 hours. DNA replication rate was assessed by the amount of BrdU incorporated into cells during a brief, 30 minute pulse at the end of the incubation period (Figure 3.13B). The presence of „high" dose (1 μM) aphidicolin dramatically reduced the DNA replication rate (i.e. BrdU MFI of S-phase cells) in cultured S0 and S1 cells, and completely blocked DNA replication in S2 cells cultured for 16 hours (Figure 3.13B). High dose aphidicolin fully prevented methylation loss at the β-globin LCR, LINE-1, PU.1, and H19 DMR loci in all cultured subsets (Figure 3.13C), demonstrating that demethylation is dependent on DNA replication throughout erythroid differentiation.

We also noted that blocking DNA replication for 16 hours in each subset not only prevented global demethylation, but actually led to a higher methylation level than found in the freshly sorted (t=0 h) controls (Figure 3.13C). For example, S1 cells cultured with high dose aphidicolin for 16 hours have a LINE-1 methylation level that is not only greater than freshly sorted S1 cells, but also reach a 5mC level almost equivalent to freshly sorted S0 cells (Figure 3.13C). In contrast to global regions, demethylation at the β-globin LCR appeared to be irreversible (Figure 3.13C), possibly the result of

transcription factor binding and exclusion of Dnmt1. Therefore, blocking DNA replication allowed cells to recover global methylation that was lost during earlier stages of differentiation. This finding suggests that Dnmt1 is capable of restoring lost DNA methylation when DNA synthesis is blocked and, because Dnmt1 has limited *de novo* methylation activity, also suggests that methylation loss proceeds through a hemi-methylated intermediate.

In contrast, the presence of „low" dose (0.1 μM) aphidicolin did not substantially change the number of cells in S-phase, but it did slow down their rate of S-phase BrdU incorporation (Figure 3.13B). For example, in cultured S1 cells the proportion of S-phase cells was 74% both in the absence or presence of low dose aphidicolin, yet the DNA synthesis rate, as determined by S-phase BrdU MFI, was reduced from 3000 to 1800 relative fluorescence units (RFU) (Figure 3.13B). This slower DNA synthesis rate was sufficient to completely eliminate demethylation at the LINE-1, PU.1, and H19 DMR loci and partially reduce the loss of methylation at the β-globin LCR in all cultured subsets (Figure 3.13C).

The sensitivity of the demethylation process to DNA synthesis rate was also seen by genomic-scale RRBS, where freshly sorted S1 cells that were allowed to differentiate in 0.1 μM aphidicolin for 24 hours were protected from genome-wide demethylation (Figure 3.14A). Low dose aphidicolin prevented demethylation at various types of genomic regions, including the retroviral SINEs and LINEs, CpG islands, enhancers and imprinted genes (Figure 3.14D). Treatment of S1 cells with low dose aphidicolin for 24

hours does not alter the number of cells in S-phase, but does reduce the DNA synthesis rate, as measured by BrdU incorporation by roughly 50% (Figure 3.17A).

The observed delay in DNA demethylation resulting from aphidicolin treatment might simply be a reflection of higher methylation levels in regions that had not replicated in the 24 h culture period. In other words, many regions may not have had an opportunity to demethylate because aphidicolin delayed the replication fork passage through these regions. If this were the case, then we would expect to see a reduction in methylation within the regions that had replicated, even in the presence of low dose aphidicolin. Instead, the RRBS of 5 kb windows, which contain regions replicating at all different times throughout S-phase, argues against this possibility since low dose aphidicolin was able to block the vast majority of methylation loss (Figure 3.14A, B).

Collectively, these results demonstrate that DNA replication is required for global and gene-specific methylation loss at all stages of erythropoiesis. Furthermore, the lack of demethylation in S-phase arrested subsets strongly suggests that an active demethylation mechanism is not involved. Therefore, we conclude that demethylation is the result of a passive mechanism. Importantly, slowing the rate of DNA synthesis is sufficient to prevent global demethylation. This observation suggests that the elevated rate of DNA replication induced during the transition from S0 to S1 contributes to inadequate maintenance methylation.

Demethylation is accelerated by inefficient knockdown of Dnmt1

We attempted to prevent demethylation by retroviral overexpression of Dnmt1 (courtesy of En Li, Novartis), using an experimental design identical to that described for

Dnmt3 (Figure 3.9). Viral supernatant containing Dnmt1-IRES-GFP was used to infect sorted S0 cells for 16 hours prior to Epo-induced differentiation for 24 hours. Preliminary experiments showed virtually no GFP expression at 40 hours. We attributed this to the large, 4.9 kb, size of the Dnmt1 coding region, which when subcloned into the expression construct yields a transcript that is 8.5 kb. Since the packaging limit of this virus is 8.3 kb, poor infectivity is likely due to inefficient viral packaging of our Dnmt1 construct. We were able to marginally improve infectivity by concentrating the viral supernatant, resulting in differentiated S3 cells expressing low levels of GFP. However, Dnmt1 mRNA was not increased in these cells compared to control.

Next we attempted to knockdown (KD) Dnmt1 using retrovirally expressed shRNA linked by IRES to a GFP reporter. We transduced fresh S0 cells with a non-silencing control shRNA or shRNA directed at Dnmt1 for 16 hours in the presence of SCF and IL-3 to allow time for shRNA expression. The infection rate was nearly 100% at 16 hours (data not shown). Cells were shifted to differentiation media containing Epo for 24 hours and then harvested, without sorting, for subsequent mRNA and protein analysis. Dnmt1 mRNA was reduced 50% after 16 hours of culture and at 40 hours an 80% reduction was observed (Figure 3.15A, left panel). Surprisingly, Dnmt1 protein was lowered only 27% after 16 hours of culture and by 40 hours this knockdown was neutralized (Figure 3.15A, right panel). This result suggests that Dnmt1 protein in erythroid cells is only partially controlled by mRNA abundance and post-transcriptional mechanisms act to keep Dnmt1 protein levels constant.

If Dnmt1 is in excess in the nucleus, then the incomplete and transient knockdown observed above would be expected to have no impact on methylation levels. Therefore, we examined the effect of Dnmt1 KD on DNA methylation levels in sorted cells. As above, S0 cells were cultured with shRNA viral supernatants for 16 hours in IL-3 and SCF. Cells were then shifted to differentiation media containing Epo or Epo plus low dose aphidicolin for an additional 24 hours. After differentiation, GFP positive cells from S1 and S3 gates were sorted and analyzed. Dnmt1 mRNA showed a 70% reduction after 16 hours of culture and by 40 hours an 85% reduction was observed in differentiated S1 and S3 subsets (Figure 3.15B, upper panel), consistent with our previous results in unsorted cells (Figure 3.15A, left panel). Dnmt1 KD did not lead to substantial methylation loss in non-differentiating S0 cells at 16 hours (Figure 3.15B, lower panel), even though Dnmt1 protein abundance is reduced at this time (Figure 3.15A, left panel). This finding suggests that prior to differentiation, a time when DNA replication rate is comparatively low (Figure 3.11A), Dnmt1 protein is present in relative excess (Figure 3.15A).

In contrast, KD of Dnmt1 in differentiated S1 and S3 cells, which have an elevated DNA synthesis rate (Figure 3.11A), caused accelerated demethylation at the β-globin LCR and LINE-1 regions (Figure 3.15B, lower panel). This finding suggests that Dnmt1 availability during differentiation is limited, since even a slight deficiency (<27% reduction, Figure 3.15A) negatively affected DNA methylation levels. The addition of low dose aphidicolin to slow the rate of DNA replication caused a modest but discernable methylation increase in Dnmt1 KD and control cells, consistent with previous results

showing that demethylation is replication-dependent. These data support a model where the very high rate of DNA replication triggered in cells as they transition into S1 (Figure 3.11A) generates a quantitative deficit in maintenance methylation by Dnmt1, leading to passive, genome-wide demethylation.

Global methylation loss does not induce genome-wide transcriptional activation

DNA methylation is associated with gene silencing. Therefore, we investigated the possibility that global DNA hypomethylation leads to genome-wide transcriptional activation. First, we looked for LINE-1 activation. These regions are normally highly methylated in order to silence retrotransposon activity [8] and lose approximately 30% of their DNA methylation during erythropoiesis (Figure 3.3A and Figure 3.6A). However, the LINE-1 transcript showed only a modest, 1.7-fold induction that did not reach statistical significance (Figure 3.16A) (p>0.2, Student''s two-tailed, paired t-test). This result was consistent with our earlier finding of methylation loss in genes whose expression decreases in erythropoiesis, such as PU.1 and Fas (Figure 3.1C, D). Interestingly, the methyl-CpG-binding protein 2 (MeCP2) has been shown to repress LINE-1 transcription in transformed cell lines and in neurons [363, 364]. We found that MeCP2 mRNA was up-regulated by approximately 2-fold during differentiation from S0 to S3 using Affymetrix GeneChips (data not shown). This finding raises the possibility that a compensatory mechanism involving MeCP2 may be activated during differentiation to keep LINE-1 retrotransposons silenced.

We examined whether global demethylation affected global mRNA levels using the Affymetrix GeneChip gene expression platform, comparing early erythroblasts in S1

with later, more mature erythroblasts in S3 (Figure 3.16B, C). We analyzed this data

using two different approaches. In our first analysis, prior to obtaining RRBS results,

gene expression data was normalized based on a group of three housekeepers, a method

traditionally followed in qRT-PCR experiments. We found that 761 transcripts were

down-regulated during differentiation, while only 146 were up-regulated (Figure 3.16B).

A similar result was obtained by applying a genome-wide normalization approach to the

data, with 5862 down-regulated genes and 2268 up-regulated genes during differentiation

(Figure 3.16C). These results show that the erythroid gene expression profile becomes

increasingly restricted during differentiation, consistent with the cellular and functional

specification that occurs during this time. Additionally, these findings suggest that global

methylation loss is not leading to global transcript induction.

We tested this directly by combining gene expression profiling with the RRBS

results. In this analysis, each 5 kb genomic region was assigned to an Affymetrix gene

based on the closest upstream or downstream Ensembl-annotated transcription start site.

We observed that up-regulated genes had a median methylation level in S0 cells that was

lower than that of down-regulated genes, pointing to a correlation between genes poised

for induction and their levels of DNA methylation. However, both up-regulated and

down-regulated genes lost 30% of their S0 methylation levels during differentiation into

S4/5 cells (Figure 3.16C), consistent with the magnitude of methylation loss observed

genome-wide (Figure 3.6). Therefore, demethylation appears just as likely to be

associated with gene induction as it is with gene repression during erythroid

differentiation, suggesting that hypomethylation is not leading to transcript induction on a

genome-wide scale. However, this analysis utilized a rather simple approach in assimilating RRBS and GeneChip data. Additional, and more complex, analysis methods still hold the potential to reveal a correlation between DNA methylation and gene expression in our system. Importantly, our preliminary finding, that global demethylation is not leading to global gene activation, does not eliminate a potential role for loci-specific demethylation in gene activation.

Demethylation is required for activation of erythroid-specific GATA-1 target genes

We next asked whether inhibiting genome-wide demethylation would affect the erythroid differentiation process. We examined expression of several erythroblast-specific GATA-1 target genes: β-globin (Hbb-b1); aminolevulinic acid synthase 2 (Alas2), required for the first step of the heme synthesis pathway [365]; and solute carrier family 4, anion exchanger, member 1 (Slc4a1, formally known as Band3), a membrane protein that functions as both a chloride/bicarbonate exchanger and a binding site for cytoskeletal proteins [308, 366, 367]. These genes show a modest increase in their expression at the S0 to S1 transition, followed by a 30- to 100-fold induction during subsequent differentiation in S2 and S3 (Figure 2.3A and Figure 3.1B). Furthermore, GeneChip analysis of S0 and S3 subsets reveals these genes are within the top 25 most highly induced genes during differentiation (data not shown).

We cultured freshly sorted S0 cells with Epo or Epo plus low dose, 0.1 μM, aphidicolin for 24 hours followed by sorting of the differentiated Ter119$^+$ population (Ter119$^+$ =combined S1, S2, and S3 sorting gate). Aphidicolin did not change the number of cells in S-phase, but it did lower the rate of DNA replication, as determined by a 30-

minute pulse of BrdU at the end of the culture period (Figure 3.17A). Low dose

aphidicolin prevented global demethylation, and reduced the extent of methylation loss at

the β-globin LCR (Figure 3.17B). Delayed demethylation was associated with

compromised expression of erythroid genes; transcript levels for β-globin, Alas2, and

Band3, were all significantly lower, by 25-40% (Figure 3.17C). We were able to rescue

the aphidicolin-induced delay in DNA demethylation by adding 50 nM of 5-aza-2″-

deoxycytidine (5-aza), an inhibitor of DNA methyltransferases (Figure 3.17B) [368]. 5-

aza did not alter the number of cells in S-phase or the DNA synthesis rate (Figure 3.17A).

Importantly, the addition of 5-aza to aphidicolin-treated cells restored normal levels of

expression for β-globin, Alas2, and Band3 (Figure 3.17C). These findings suggest that

the slower induction of erythroid genes in the aphidicolin-treated cells was a result of a

failure to demethylate DNA.

Discussion

Location and extent of erythroid global demethylation

We have discovered that erythropoiesis is associated with progressive genome-wide DNA demethylation. Global demethylation is initiated with the onset of erythropoiesis, as cells transition from S0 to S1 (Figure 3.4), and is temporally coordinated with extensive demethylation at the β-globin LCR (Figure 3.7). Demethylation was truly global, with 67% of 5 kb genomic regions undergoing statistically significant methylation loss (Figure 3.5). Demethylation was spread across a wide-variety of genomic elements including CCpGG and CpG repeats (3.3B, C); genes down-regulated in erythropoiesis (Figure 3.1C, D and 3.16C), endogenous retroviruses (Figure 3.3A and 3.7); imprinted genes (Figure 3.2 and 3.6B); and CpG islands, island shores, promoters, and enhancers (Figure 3.6B, C). On average, most genomic regions lost 30% of initial S0 methylation levels during differentiation into S4/5 cells, with notable exceptions at the β-globin LCR which lost 65-80% and imprinted regions that lost only 14% (Figure 3.6D). Furthermore, global hypomethylation appears to be a general feature of erythropoiesis, neither limited to the definitive lineage nor the fetal liver (Figure 3.4).

Our findings are consistent with a previous study showing reduced levels of total methylation in murine placenta and yolk sac tissues [369]. However, their data was not sufficient to conclude that hypomethylation was the result of a demethylation process. Furthermore, due to the cellular heterogeneity of these tissues, it was unclear as to which cell types were hypomethylated. Our data suggests that hypomethylation in the placenta

and yolk sac is due to the large number of primitive erythrocytes present in these two tissues.

Erythropoiesis is the first example of global DNA demethylation in somatic cells, which until now, were thought to possess similar levels of total methylation with differences occurring at a relatively small number of gene-specific loci [36, 40, 47]. Instead, physiological examples of global demethylation have been limited to the pre-implantation embryo and primordial germ cells [203]. In disease, DNA hypomethylation is a hallmark of a large number of cancers [237, 238].

The assays employed to measure DNA methylation in the pre-implantation embryo have been largely qualitative, relying on immunofluorescence using an anti-5mC antibody or DNA hybridization-based methods. Furthermore, genome-scale methylation analysis, such as the RRBS used in our study, has not been performed in the early embryo. Therefore, it is difficult to directly compare the genomic locations affected by demethylation in the pre-implantation embryo with our data. However, quantitative methods have been used, mostly measuring methylation loss at repetitive retroviral elements [157, 183-185]. These studies show that the total methylation loss is approximately 50%, as compared to the 30% loss observed in our system. Notably, unlike in erythropoiesis and PGCs, imprinted regions are spared from demethylation in the pre-implantation embryo [123, 188-190].

In PGCs genome-scale methylation analysis has recently been performed [175]. Similar to erythropoiesis, PGCs undergo demethylation at a wide-variety of genomic elements, including exons, introns, promoters, intergenic regions, and retroviral LINE,

SINE and LTR regions. In contrast to erythropoiesis, total methylation loss in PGCs is much more dramatic, reaching levels of only 16% and 7.8% in male and female PGCs, respectively [156, 157, 175, 185]. This is roughly a 75% reduction in total 5mC, vastly exceeding the methylation loss observed in erythropoiesis. In PGCs, DNA methylation at imprinted regions is completely removed [214, 215, 219]. This is a notable difference compared to the relatively minor 15-30% of 5mC removed at imprinted regions during erythropoiesis (Figure 3.2 and 3.6).

The magnitude of total 5mC loss during oncogenesis differs by the type of cancer, but generally ranges between 10-50% [237, 238]. In addition, it is well known that the majority of cancers lose methylation at repetitive elements, including retrotransposons [244]. However, the precise locations of demethylation in cancer remain largely unknown. For example, methylation loss can occur at imprinted regions, but it is not known if this is a common feature of all cancers that exhibit global hypomethylation [265]. The absence of this information in the literature is mainly due to the large number of cancers, their molecular diversity, and a lack of unbiased genome-scale studies.

Mechanism of erythroid global demethylation

Demethylation, both globally and at the β-globin LCR, is completely dependent on DNA replication at all stages of differentiation (Figure 3.12 and 3.13), and therefore proceeds through a passive mechanism. Furthermore, knockdown of proteins with a proposed role in active demethylation, Mbd4 and Gadd45a, did not prevent DNA demethylation (Figure 3.8 and 3.10). Demethylation coincides with a dramatic increase in the rate of DNA replication during the S0 to S1 transition, which remains elevated

through the S3 differentiation stage (Figure 3.11A). Slowing the rate of DNA replication, without perturbing the number of cells in S-phase, was sufficient to prevent global demethylation and limit the extent of β-globin LCR methylation loss (Figure 3.13B, C and 3.14). We found that Dnmt3a and Dnmt3b are down-regulated during erythropoiesis, but their loss does not appear to be contributing to demethylation, since their exogenous expression failed to rescue demethylation (Figure 3.9). Therefore, Dnmt1 seems to be solely responsible for maintenance of methylation in erythroid cells. Consistent with this finding, transient and inefficient knockdown of Dnmt1 led to an acceleration of methylation loss, suggesting that Dnmt1 is present in limiting amounts during differentiation (Figure 3.15). Collectively, these results support a model where the very high rate of DNA replication triggered in cells as they transition into S1 generates a quantitative deficit in maintenance methylation by Dnmt1, leading to passive, genome-wide demethylation. To our knowledge the finding that increased replication rate is sufficient to induce global demethylation has not been described.

Like erythropoiesis, demethylation of maternal DNA in the pre-implantation embryo occurs gradually over several cell divisions [186, 191]. During these divisions the highly abundant and oocyte-specific Dnmt1 isoform, Dnmt1o, is localized to the cytoplasm, leaving only low levels of somatic Dnmt1 in the nucleus [123, 192, 193]. We have not yet investigated the subcellular localization of Dnmt1 during erythropoiesis, but Dnmt1 has been found by others to reside almost exclusively in the nucleus of somatic cells [193, 370]. Therefore, the mechanism driving passive demethylation in erythropoiesis is potentially quite different from that in the pre-implantation embryo; in

erythropoiesis the methylation enzyme remains relatively constant, but the DNA substrate increases, while in the pre-implantation embryo, presumably normal levels of substrate are present, but the enzyme decreases.

The mechanism of 5mC loss in PGCs is largely unknown. Correlative evidence, such as activation of the BER pathway, suggests that demethylation is through an active mechanism [197]. Also, Dnmt1 is highly expressed at the mRNA and protein levels in the nucleus from E8.5 to E12.5, a period of time during which 5mC is erased [214, 218]. This finding has been cited as supporting an active mechanism, since it is not consistent with known examples of passive demethylation [153], i.e. nuclear exclusion of Dnmt1 during replication in the pre-implantation embryo. However, experiments utilizing cell cycle inhibitors during PGC development have not been performed, so a DNA replication-dependent mechanism has not formally been ruled out. Furthermore, it is not known if PGCs have an elevated rate of intra-S-phase DNA synthesis of the kind we have described in erythropoiesis.

Could reduced maintenance methylation, on the order of magnitude described in our system, account for the methylation loss observed in PGCs? The „efficiency" of maintenance has been calculated by others as the fraction of 5mC retained per cell cycle [371]. By this metric, even in the absence of any methyltransferase activity, efficiency would still be 50% per cycle. We propose that a more appropriate calculation for maintenance efficiency is the fraction of 5mC retained on the nascent strand per cell cycle. Differentiating erythroblasts undergo 3 cell divisions and lose 30% of their initial methylation during this time, resulting in a maintenance efficiency of 76% per cell cycle.

PGCs undergo 10 divisions over 5 days during which time the median genome-wide methylation level is reduced from approximately 70% in ES or whole embryo to 16% in the male PGC [175]. Methylation loss of this magnitude, if it were due to passive demethylation only, would require a maintenance efficiency of 73% per cycle. This result is in close agreement with the observed maintenance efficiency in erythropoiesis of 76% per cycle. Therefore, passive DNA demethylation of the kind described for erythropoiesis could account for global methylation loss in PGCs.

The mechanism of demethylation in cancer is largely unknown, but indirect evidence suggests a passive mechanism. First, DNA demethylation is known to be an early event during carcinogenesis that often escalates with disease progression [113, 239, 240, 248, 273-275]. This slow progression, and incomplete loss of methylation, suggests a passive mechanism requiring many rounds of cell division. Second, hypomethylation in cancer has been correlated with an increased number of proliferating cells, determined by markers such as Pcna, Ki-67, histone H4, BrdU immunohistochemistry, or total DNA content [239, 243, 248, 276, 372, 373]. Transitional cell carcinomas of the urinary bladder have been associated with reduced Dnmt1 abundance relative to the proliferative marker Pcna [372]. In glioma, decreased methyltransferase activity resulted from decreased recruitment of Dnmt1 to DNA, even though protein levels remained constant [113]. Together, these findings suggest that an imbalance between maintenance methylation and proliferation likely leads to hypomethylation in cancer. Like cancer, demethylation in erythropoiesis occurs progressively over several cell divisions and is associated with an increase in the number of proliferating cells (Figure 3.11B).

Importantly, our data demonstrate that the rate of intra-S-phase DNA synthesis is a contributing factor to passive demethylation and raises the possibility that a similar mechanism could be responsible for global hypomethylation in cancer.

Functional significance of erythroid global demethylation

We have started to study the functional significance of global DNA demethylation during erythropoiesis. Given the known association of 5mC with transcriptional repression, we thought that global methylation loss might be linked to global up-regulation of mRNA. This does not appear to be the case, as demethylation was just as likely to be associated with gene repression as it was with gene induction during differentiation (Figure 3.16). However, this analysis utilized a rather simple approach that potentially lacks the sensitivity necessary to reveal a correlation. For example, in some cases the closest 5 kb region to a transcription start site might reside within the gene-body. This could confound results, since in contrast to the majority of genomic regions, gene-body methylation is associated within highly expressed genes [46]. Our analysis might be improved by only utilizing 5 kb regions upstream of, or at, the gene transcription start site. Alternatively, CpG island shore regions, in which a lack of 5mC is strongly correlated with gene expression, could be employed in lieu of 5 kb regions [43, 47]. More complex methods for associating putative regulatory regions to genes have also been recently proposed [374].

Nevertheless, our finding is in agreement with other systems, where global loss of methylation is not sufficient to generate global gene induction [31, 375-377]. In addition to DNA demethylation there are several enzyme classes required for gene activation,

including chromatin modifiers and tissue-specific transcription factors. Since erythroblasts become increasingly specialized during the differentiation process, it is likely that the factors necessary for global gene activation are restricted. Alternatively, nuclear condensation and enucleation of mature erythroblasts may protect these cells from excessive transcription and other unwanted effects that might otherwise result from global demethylation. Interestingly, we observed that up-regulated genes had lower initial methylation levels in S0 cells compared to down-regulated genes, pointing to a correlation between genes poised for induction and their levels of DNA methylation.

Global demethylation in the pre-implantation embryo, PGCs, and cancer has been shown to affect single copy gene regions. In fact, activation of specific genes important for cellular identity is one of the major proposed functions for global demethylation in these systems. Global demethylation in the pre-implantation embryo is believed to be part of a mechanism required to activate genes leading to pluripotency, such as Nanog and Oct4 [42, 174, 204, 205, 209-212]. Similarly, a critical function of methylation loss in PGCs is the removal of parental imprints [213-215, 219, 224] and activation of germ-line specific genes [228-230]. The functional consequence of global demethylation in cancer is more diverse, but single copy gene loci have been shown to exhibit hypomethylation-linked expression [237, 238, 262, 265]. Importantly, in PGCs, pre-implantation embryo, and cancer this function was deduced by a temporal correlation with genome-wide demethylation, due to an inability to modulate the rate of demethylation in these systems.

We have also demonstrated strong temporal correlations between global demethylation and gene specific demethylation at the β-globin LCR (Figure 3.7).

Furthermore, we have demonstrated that the β-globin LCR and global demethylation are linked through a DNA replication-dependent mechanism (Figure 3.12, 3.13, and 3.14). The ability to slow the rate of demethylation, using DNA synthesis inhibitors, and accelerate demethylation, through addition of 5-aza, has allowed us to dissect the functional significance of global demethylation even further. In initial experiments using these tools, we provided evidence that demethylation is required for full transcriptional activation of the erythroid-specific GATA-1 target genes β-globin, Alas2, and Band3 (Figure 3.17). These genes are massively up-regulated in normal erythroblast differentiation and critically important to red cell function. For example, hemoglobin, of which β-globin is a major component, eventually comprises >95% of red cell mass and even a 10% reduction has clinical manifestations and is considered anemia [378]. We propose that genome-wide demethylation is part of a mechanism required for rapid demethylation, and activation, of specific erythroid gene loci. This model is also supported by a recent study which showed de-repression of myeloerythroid genes in hematopoietic stem cells that lose DNA methylation as a result of expressing hypomorphic Dnmt1 [32].

Accelerated demethylation at the β-globin LCR

Demethylation during erythropoiesis has a hierarchical structure, with the highest demethylation rate found at the erythroid-specific β-globin LCR (65-80% loss), slower demethylation in most other genomic sites (30% loss), and the lowest rate at imprinted loci (15% loss) (Figure 3.6D). Correlation data (Figure 3.7) and response to cell cycle inhibitors (Figure 3.12, 3.13, and 3.14) suggest that demethylation in each of these

regions is driven by a single, DNA-replication dependent mechanism. Despite this common mechanism, the efficiency of maintenance methylation at the β-globin LCR is only 20-40% per cell cycle as compared to 76% per cycle genome-wide. One explanation for this dramatically reduced maintenance fidelity could be the binding of erythroid-specific transcription factors at the β-globin LCR. DNA replication-dependent demethylation is known to be induced by the binding of protein factors to DNA [16, 126-128]. For example, DNA replication and EBNA-1 binding was required for demethylation of an episomal oriP sequence in PC-3 prostate cancer cell line [127]. Similarly, episomal lacO sites become demethylated upon binding of LacI in the human embryonic kidney 293 cell line, and a decrease in LacI binding affinity, caused by increasing amounts of IPTG, resulted in a slower loss of methylation. [128].

Gata-1 and Eklf are transcription factors essential for erythrocyte differentiation [379-381]. These factors directly bind DNA within the hypersensitivity sites of the β-globin LCR to activate the β-globin gene [103, 382-384]. NF-E2, another transcription important for β-globin activation, specifically targets HS2 of the LCR [103, 382-384]. We found that Gata-1, Eklf, and NF-E2 mRNAs were present in S0 cells at 200-fold higher levels than in fetal brain (Figure 2.3A). They increased a further 2-fold with the transition from S0 into S1 and continued to increase in S2 and S3.

We propose that the increase in DNA replication rate during the S0 to S1 transition results in an increase in the number of hemi-methylated sites genome-wide and at the β-globin LCR. Within genome-wide regions, these sites have the potential to become fully methylated by the delayed action of Dnmt1 [370]. In contrast, binding of

erythroid-specific transcription factors at the β-globin LCR protects hemi-methylated CpGs from the maintenance activity of Dnmt1. This model is supported by data from S-phase arrest of partially differentiated cells in culture, where non-erythroid loci regained methylation lost at previous stages of differentiation, while the β-globin LCR did not (Figure 3.13C). This mechanism has recently been proposed in neural precursor cell differentiation; where NFI plays a critical regulator role in the epigenetic switch towards astrocytogenesis by binding to astrocytic gene promoters and protecting them from Dnmt1 activity [129].

Materials and Methods

Isolation of erythroid progenitors by flow cytometry

Single cell suspensions were prepared by mechanically dissociating whole fetal livers, obtained from E12.5 to E13.5 Balb/C mouse embryos as described [105]. Cells were stained for CD71, Ter119, and a cocktail containing lineage-specific antibodies directed at CD3, Gr-1, B220, CD41 and Mac-1 (CD11b). Cells positive for lineage-cocktail markers were excluded from analysis. Isotype-matched antibodies were used to set up background staining in each channel. Erythroid subsets were sorted using gates illustrated in Figure 3.1A, on a BD FACSAria (BD Biosciences) using a 100 μm nozzle. Dead cells were excluded from analysis using DAPI (Roche). Alternatively, to isolate S0 cells only, fetal liver cells were labeled with biotin-conjugated anti-CD71 at 1:250 and anti-Ter119 at 1:100 and subjected to magnetic separation using EasySep (StemCell Technologies). The LSRII (BD Biosciences) cytometer was used for analysis in experiments where cell sorting was not required. All FACS data was analyzed using the FlowJo software (Tree Star Inc., CA). Antibodies used in flow cytometry analysis: PE or biotin rat anti-mouse CD71 (C2 clone), APC or biotin rat anti-mouse Ter119, FITC rat anti-mouse CD41 (MWReg30), FITC rat anti-mouse CD45R/B220 (RA3-6B2), FITC hamster anti-mouse CD3e (145-2C11), FITC rat anti-mouse CD11b (M1/70), FITC rat anti-mouse Ly-6G and Ly-6C (RB6-8C5) (BD Biosciences).

Measurement of DNA methylation by pyrosequencing

Genomic DNA (gDNA) was isolated from flow-cytometrically sorted fetal liver subsets S0 to S4/5 using the All Prep DNA/RNA Micro Kit (Qiagen). For sodium bisulfite conversion, between 50 and 150 ng of gDNA were treated using the EZ DNA Methylation-Direct Kit (Zymo Research). Alternatively, for samples containing small cell numbers (≈10,000), cells were treated with proteinase-K and used directly for bisulfite conversion. For each genomic locus, the region of interest was amplified by PCR using primers specific to the bisulfite-converted gDNA. DNA methylation assays for β-globin LCR HS1 (ADS937), β-globin LCR HS2 (ADS938), H19 DMR (ADS438), PEG3 (ADS183), Kcnq1ot1 (ADS912), PU.1 (ADS1090), and LINE-1 (ADS685) were designed by EpigenDx (Worcester, MA). Assay design reports, which include the target region sequence, linearity/validation data and primer sequences, are available from EpigenDx. As a control, the efficiency of bisulfite conversion was assessed using primers to the non-converted DNA sequence. In this case, there was no amplification, indicating that the DNA was fully converted. The PCR product was visualized on a 1% agarose gel to verify amplification specificity, followed by pyrosequencing (EpigenDx) using sequencing primers internal to the amplification primers. Cytosine dispensation in non-CpG positions was used to verify that the PCR product amplified genomic DNA that was fully converted. The methylation status of each locus was analyzed individually as a T/C SNP using QCpG software (Qiagen) by EpigenDx.

Measurement of DNA methylation levels by cloning

Genomic DNA was isolated and bisulfite converted as described above.

Converted DNA was amplified by PCR using primers specific for the H19 DMR (FW,

aggaaagaaaaaggttggtgagaaaatagag; RV, aaataacccacaacattaccatttataaattcc). PCR product

was cloned into pCR2.1-TOPO vector using the TA Cloning kit (Invitrogen) and used to

transform DH5α E.coli (Invitrogen). Plasmid DNA was prepared from individual

colonies, sequenced using M13 forward or reverse primers, and analyzed using the

Bisulfite Sequencing DNA Methylation Analysis (BISMA) toolkit [385]. DNA

methylation levels of individual CpGs obtained by this strategy strongly correlated with

those of the pyrosequencing assay described above ($R^2 = 0.85$).

Measurement of DNA methylation levels using LUMA

Genome-wide methylation of CCpGG sites was detected using the LUminometric

Methylation Assay (LUMA) [355]. In parallel reactions, 100 ng of gDNA was digested

with either methylation sensitive HpaII or methylation insensitive MspI. EcoRI was

included in both reactions as an internal control for the amount of template DNA. MspI

and HpaII leave 5"CG overhangs, while EcoRI produces 5"AATT overhangs. The

quantity of overhangs present in each enzymatic reaction was measured using

pyrosequencing by EpigenDx.

Measurement of DNA methylation levels by ELISA with an anti-meCpG antibody

Purified gDNA was first quantified using the DNA Quant-iT™ dsDNA HS Assay

Kit (Invitrogen), followed by methyl-cytosine quantification using the Methylamp™

Global DNA Methylation Quantification Ultra Kit (Epigentek). gDNA (100 ng) for each

sorted subset was immobilized to the bottom of a 96-well plate. Binding of an antibody

raised against 5-methylcytosine was quantified by ELISA and compared to binding of a

fully methylated control DNA.

Measurement of DNA methylation by reduced representation bisulfite sequencing

RRBS was performed as previously described [42]. Genomic DNA (gDNA) was

isolated using the All Prep DNA/RNA Micro Kit (Qiagen). For all RRBS experiments,

two independent biological replicate pools were investigated. Genomic DNA from 4 or 7

sorting experiments was pooled to yield each biological replicate for the S0, S1, S3 and

S4/5 populations (Figure 3.5 and 3.6). Genomic DNA from 2 sorting experiments was

pooled to yield each biological replicate in experiments utilizing aphidicolin treated S1

cells in culture (Figure 3.14). RRBS data for these samples are available at

http://erythrocyte-demethylation.computational-epigenetics.org/

GeneChip analysis

Experiment-matched S1 (n=3) and S3 (n=3) cells were isolated by sorting E12.5

fetal liver cells on three different days. Total RNA (20 ng per sample) was converted to

cDNA, linearly amplified and biotinylated using Ovation reagents (Nugen, San Carlos,

CA). Samples were hybridized to Mouse Genome 430 2.0 Arrays (Affymetrix, Santa

Clara, CA). For the analysis presented in Figure 3.16B, microarray suite 5 (MAS5)

processed sample data were normalized to the average of 18SRNA (AFFX-

18SRNAMur/X00686_M_at), GAPDH (AFFX-GapdhMur/M32599_3_at) and β-actin

(1419734_at) expression values using BRB-ArrayTools developed by Dr. Richard Simon and BRB-ArrayTools Development Team [386]. Genes were identified as differentially expressed if their random variance t-test p-value was less than 0.005 and absolute fold change value larger than 2. The analysis presented in Figure 3.16C used R/Bioconductor and the limma library, with FDR<0.05 and $\log_2$(fold-change)<-1 for down-regulated genes and $\log_2$(fold change) >1 for up-regulated genes. Data was normalized using gcRMA and expression values were subsequently scaled into the interval from 0 to 10, with 0 corresponding to genes expressed at or below the 0.5% quantile and 10 corresponding to gene expressed at or above the 99.5% quantile.

Cell cycle analysis

For determination of DNA replication *in vivo* (Figure 3.11), mice pregnant with E13.5 or E14.5 embryos were injected intraperitoneally 30 min prior to harvesting of fetal livers with 100 μl of 10 mg/ml BrdU; erythroid subsets were fixed, permeabilized and analyzed using the BrdU flow kit (BD Bioscience). Simultaneously, CD71 and Ter119 antibodies were used to identify differentiation subsets and 7AAD (BD Biosciences) was used as a marker of total DNA content. For determination of DNA replication *in vitro*, cells in culture were pulsed with 10 μM BrdU for 30 min. BrdU incorporation was detected as described above.

Culture of primary fetal liver cells *in vitro*

Freshly sorted fetal liver subsets were incubated for up to 48 hours in IMDM (L-glutamine, 25mM HEPES) (Gibco) supplemented with 20% fetal calf serum, 1%

penicillin/streptomycin and 10^{-4} M β-mercapthoethanol, in the presence of 2 U/ml Epo. Aphidicolin (Sigma), L-mimosine (Sigma), 5-Aza-2'-deoxycytidine (Sigma) were added to the cells at 0.1-1.0 μM, 1 mM, or 0.05 μM respectively.

Retroviral transduction of Dnmt3

cDNAs for Dnmt3a, Dnmt3b1 and Dnmt3b1:PC were kindly provided by Dr. En Li (Novartis Institutes for BioMedical Research) and subcloned into the MSCV-IRES-GFP retroviral vector MSIG [387]. Viral supernatants were prepared by co-transfecting the desired plasmid and the pCL-Eco packaging vector [388] into 293T cells using FuGENE6 (Roche). One milliliter of viral supernatant, supplemented with 5 μg of polybrene (Sigma), was used to infect $5x10^5$ S0 cells by spinning at 2000 rpm at 30ºC for 1 hr in fibronectin-coated dishes. Cells were then incubated overnight (16 h) in the presence of 100 ng/ml SCF (Peprotech, Rocky Hill, NJ), and 10 ng/ml IL-3 (Peprotech) to allow expression of retrovirus. Expression was examined 16 hours and 40 hours post infection by flow-cytometry for GFP, and at 40 hours by QRT-PCR for DNMT mRNA.

Knockdown experiments

Short hairpin RNA targeting Gadd45a (clone1= V2MM_13565, ID SM2169-e-9; clone 2= VSHS_151136, ID SH2954-H-4), MBD4 (V2MM_16858, ID SM2374-d-10) or Dnmt1 (V2MM_46797, SM2437-d-12) were subcloned into LMP microRNA-adapted retroviral vector containing an "IRES-GFP" reporter (Open Biosystems, Huntsville, AL). Similarly, a nonsilencing negative control shRNA (RHS4971, Open Biosystems) that is processed by the endogenous RNAi pathway but will not target any mRNA sequence in

mammals, was subcloned into LMP. Viral supernatants were generated as described above. Sorted S0 cells were transduced with retroviral vectors as described above.

Quantitative RT-PCR

Total RNA was prepared from cells sorted from each fetal liver subset, using the All Prep DNA/RNA Micro Kit (Qiagen) with on column DNase treatment. Reverse-transcription was conducted using Superscript II (Invitrogen) with random hexamer primers. The ABI 7300 sequence detection system, TaqMan reagents and TagMan MGB probes (Applied Biosystems) were used and several dilutions of each template were used to ensure detection in the linear range of the assay. A „no template" and „no reverse-transcriptase" controls were included. The threshold cycle (Ct) for housekeeping genes GAPDH or β-actin was subtracted from the Ct for genes of interest to yield a relative expression value. QRT-PCR probes used: Gapdh (Mm99999915_g1), β-actin (Mm02619580_g1), β-globin (Mm01611268_g1), PU.1 (Mm00488140_m1), Dnmt1 (Mm00599784_m1), Uhrf1 (Mm00477873_g1), Dnmt3a (Mm00432870_m1), Dnmt3b (Mm01240113_m1), Mbd4 (Mm00521972_m1), Gadd45a (Mm00432802_m1), Tdg (Mm00834243_g1), Aidca (Mm00507774_m1), Apobec1 (Mm00482894_m1), Alas2 (Mm01260713_m1), Band3 (Mm00441492_m1), LINE-1 5" Region (FW,GCTACCTTGCCAGCAGAGTCTT; RV, AGTTCCGCGCGATTGGATT; probe, ACCCTAAGACCTCTGGTGAGT), LINE-1 3" Region (FW, TTCCACCTCACACCAGTCAGA; RV, CACCAACAATGGAGGAGTGTTC; probe, TTCAGGTGACAGCAGATG).

Western blot analysis

Sorted cells from each fetal liver subset were lysed by rocking for 15 min at 4°C in 1% NP40, 50 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 10% glycerol supplemented with protease inhibitors (Roche); following centrifugation for 15 min at 4°C, supernatant was quantified by the BCA Protein Assay Kit (Pierce). Phoenix cells, a 293T-derived cell line, transiently transduced with retroviral vectors encoding either Dnmt3a or Dnmt3b were used as positive controls. Lysates (10 μg protein per lane) were analyzed on a 3-8% Tris-acetate gradient polyacrylamide gel (Invitrogen), transferred to a nitrocellulose membrane, and probed with mouse anti-Dnmt3a (ab13888) (Abcam, Cambridge, MA), mouse anti-Dnmt3b (ab13604) (Abcam), or mouse anti-Dnmt1 (ab92453) (Abcam) overnight at 4°C. Following washing, membranes were probed with peroxidase donkey anti-mouse IgG (Jackson ImmunoResearch, West Grove, PA) followed by washing, then detection with Lumi-LightPLUS western blotting substrate (Roche) (Figure 3.8B, left panel). Membranes were then re-probed with rabbit anti-beta actin (ab8227) (Abcam) and mouse anti-DNMT antibodies, followed by detection with anti-Mouse-IRDye 800CW (green channel) (LI-COR Biosciences) and anti-Rabbit IRDye 680 (red channel) (LI-COR). Membranes were simultaneously analyzed for β-actin and either Dnmt1, Dnmt3a, or Dnmt3b using the Odyssey infrared imaging system (LICOR). Target protein bands were quantified using the Odyssey software (Figure 3.8B, right panel).

Statistical analysis

Methylation levels on multiple CpGs (in Figure 3.1, 3.2, and 3.3) were analyzed using a general linear mixed model with CpG and/or subset as fixed effects and individual experiments as the random effect. Normal distribution of model error was confirmed by Kolmogorov-Smirnov test on residuals. „P" values were adjusted using Sidak multiple comparison correction. Other statistical tests are described directly in the main text.
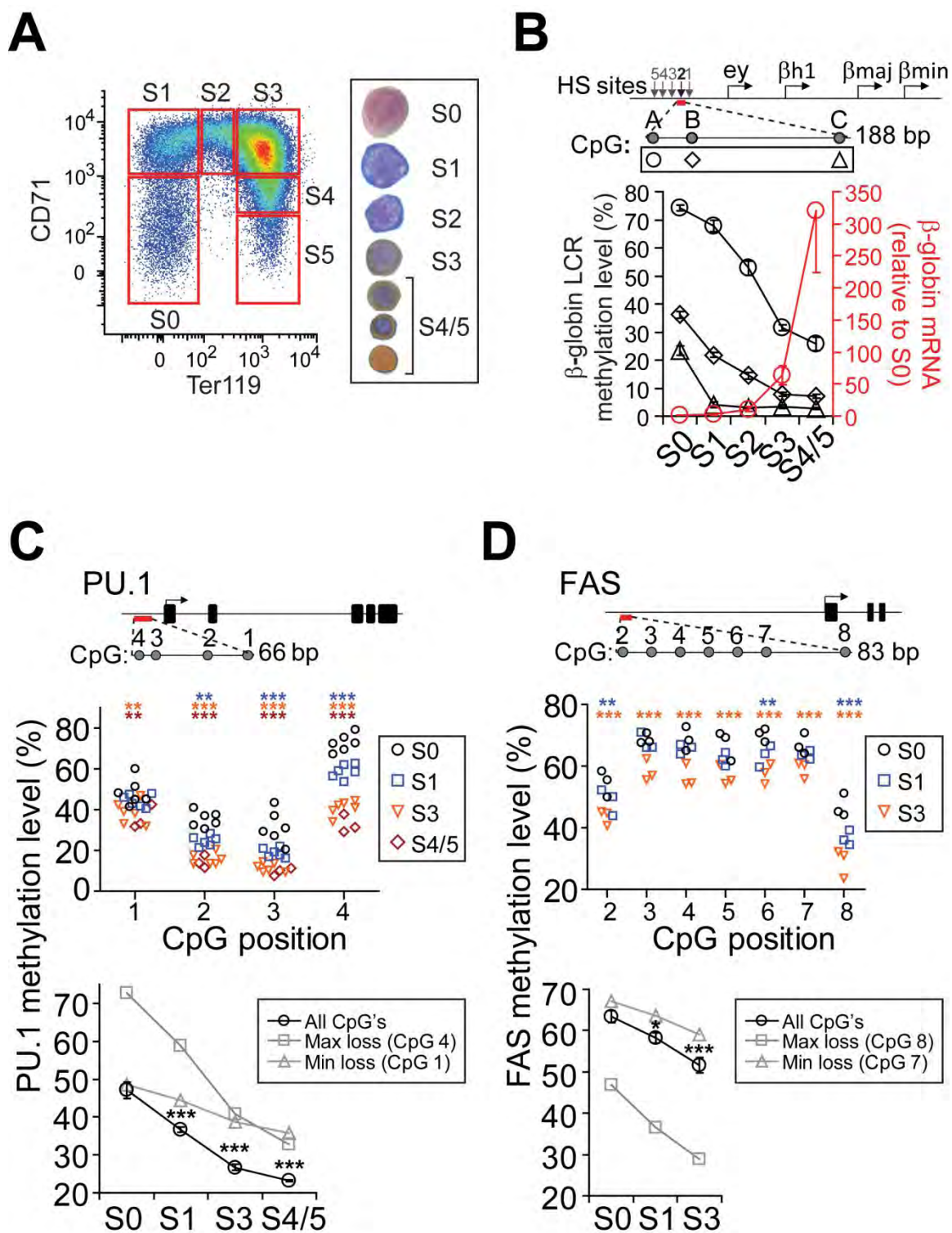
Figure 3.1. DNA methylation loss at the β-globin LCR and at loci of genes down-regulated with erythroid differentiation.

       A. Erythroid maturation subsets in fetal liver. CD71/Ter119 flow cytometric profile of freshly isolated fetal liver (left panel), and examples of cells from each of subsets S0 to S4/5 (right panel). Erythroid maturation is associated with decreasing cell size, nuclear condensation and positive staining for hemoglobin with diaminobenzidine (brown-orange). Cell density in the flow-cytometric profile is represented by color, from highest (red) to lowest (blue).

       B. DNA methylation (in black) and β-globin mRNA expression (in red) in erythroid subsets S0 to S4/5 sorted from fresh fetal liver. DNA methylation is shown for 3 CpGs labeled A to C within hypersensitivity site 2 (HS2) of the β-globin LCR (see map, not to scale). Results are mean ± s.e.m. of independent replicate sorting experiments; number of experiments (n) for each subset is n=11 (S0, S1), n=5 (S2), n=7 (S3) and n=3 (S4/5). β-globin mRNA data is the mean of four replicates, expressed relative to the Gapdh mRNA and normalized to expression in S0. As expected, DNA methylation loss at the β-globin LCR precedes β-globin gene activation.

       C. DNA methylation of a region upstream of the PU.1 gene transcription start site (marked in red and expanded, see locus map). Upper panel shows methylation level at each of four CpGs for each of subsets S0, S1, S3 and S4/5 by pyrosequencing of PCR products generated from bisulfite converted gDNA; each symbol represents a single sorting experiment. Lower panel shows a summary of the data in upper panel: the mean methylation level of all four CpGs for all experiments at each subset (black symbols, mean ± s.e.m.), and the mean methylation level for all experiments at CpG4 (which shows the largest loss in methylation) and CpG1 (showing the smallest loss in methylation). Statistics for each erythroblast subset are based on a comparison to S0. ***P<0.001, **P<0.01, *P<0.05 by linear mixed model analysis of variance with Sidak correction for multiple tests.

       D. DNA methylation of a regulatory region upstream of the Fas gene transcription start site (marked in red and expanded, see locus map). Data and statistics represented as described in „C".

Figure 3.2. DNA methylation loss at imprinted regions.

      A. DNA methylation at the imprinted H19 DMR locus in erythroid subsets. Left panel shows individual clones (rows, n=34 and 36) were sequenced from S0 and S4/5, respectively (left panel). Columns represent the position of individual CpGs (see locus map). The red line divides the clones into two halves. Black boxes represent methylated cytosines, white boxes represent unmethylated cytosines. Average methylation level for all clones is 61% for S0, and 42% for S4/5 (p=0.007, two-tailed Mann-Whitney test). Right panel shows methylation levels at each individual CpG for each of subsets S0, S1, S3 and S4/5 by pyrosequencing of PCR products generated from bisulfite converted gDNA. Data and statistics as described in Figure 3.1C.

      B. DNA methylation at the imprinted PEG3 locus. Data and statistics as described in Figure 3.1C.

      C. DNA methylation at the imprinted Kcnq1ot1 locus. Data and statistics as described in Figure 3.1C.
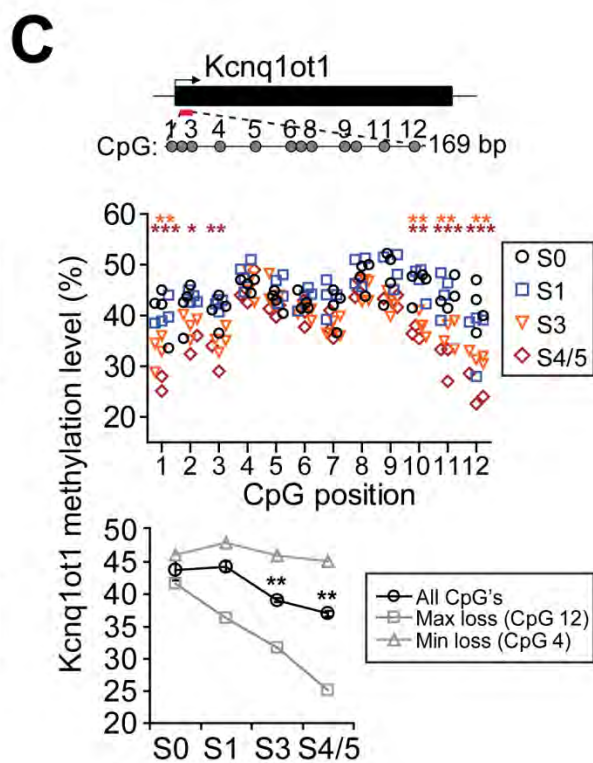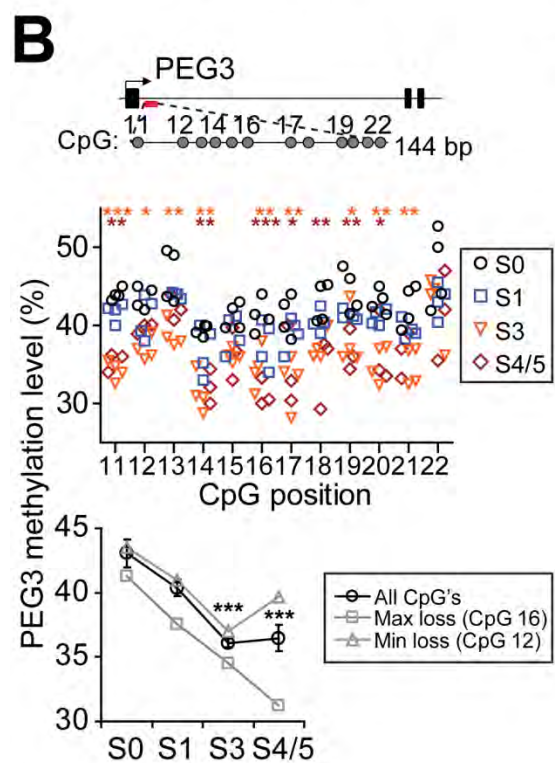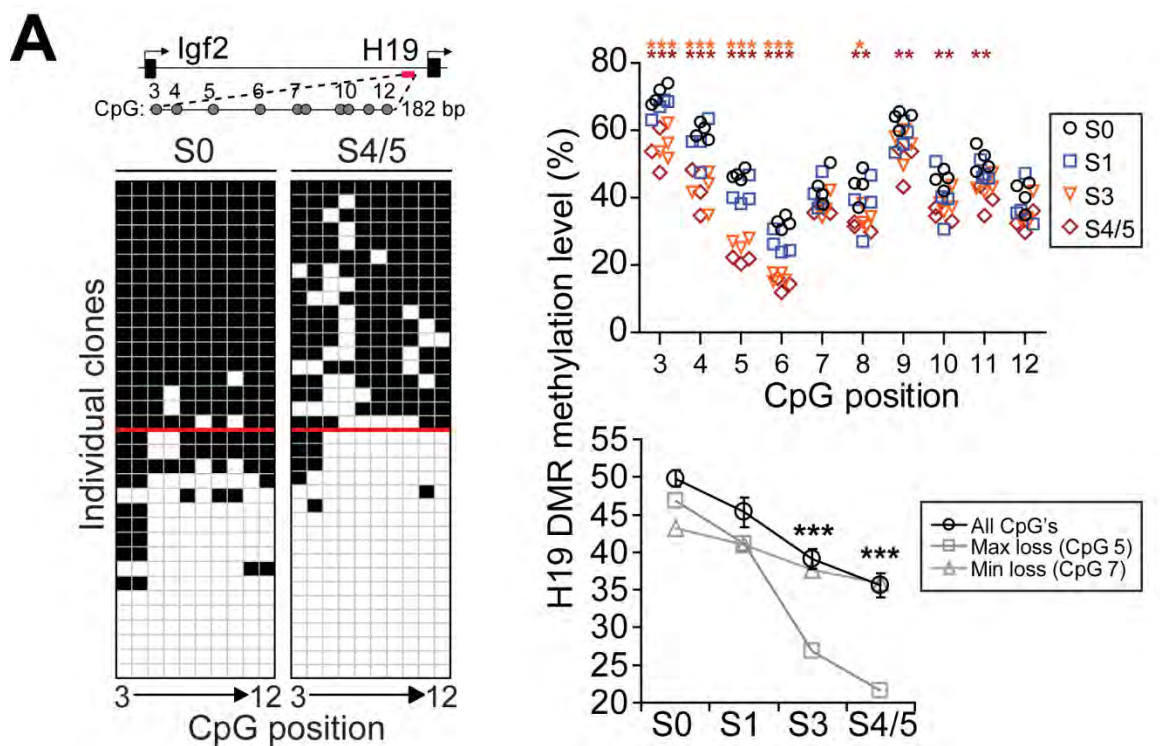
Figure 3.3. DNA methylation loss in global methylation assays.

     A. Methylation of LINE-1 retrotransposons (marked in red and expanded, see locus map). LINE-1 retrotransposons are present at approximately 100,000 copies throughout the mouse genome. Data and statistics as described in Figure 3.1C.

     B. Global DNA methylation level of CCpGG motifs. Methylation levels of CCpGG motifs, present at 2.5×10^6 copies in the mouse genome, were evaluated by digestion of genomic DNA with the isoschizomers HpaII and MspI restriction enzymes which are methylation sensitive and insensitive, respectively. Results are mean ± s.e.m. of n=3 (S0, S1, S3) and n=2 (S4/5). Statistics for each erythroblast subset are based on a comparison to S0. **P<0.01, and ***P<0.001 by linear mixed model analysis of variance.

     C. Measurement of global DNA methylation level by ELISA with an antibody directed at 5-methylcytosine. Results are mean ± s.e.m. of n=3. Statistics for each erythroblast subset are based on a comparison to S0. *P<0.05, **P<0.01 by linear mixed model analysis of variance.
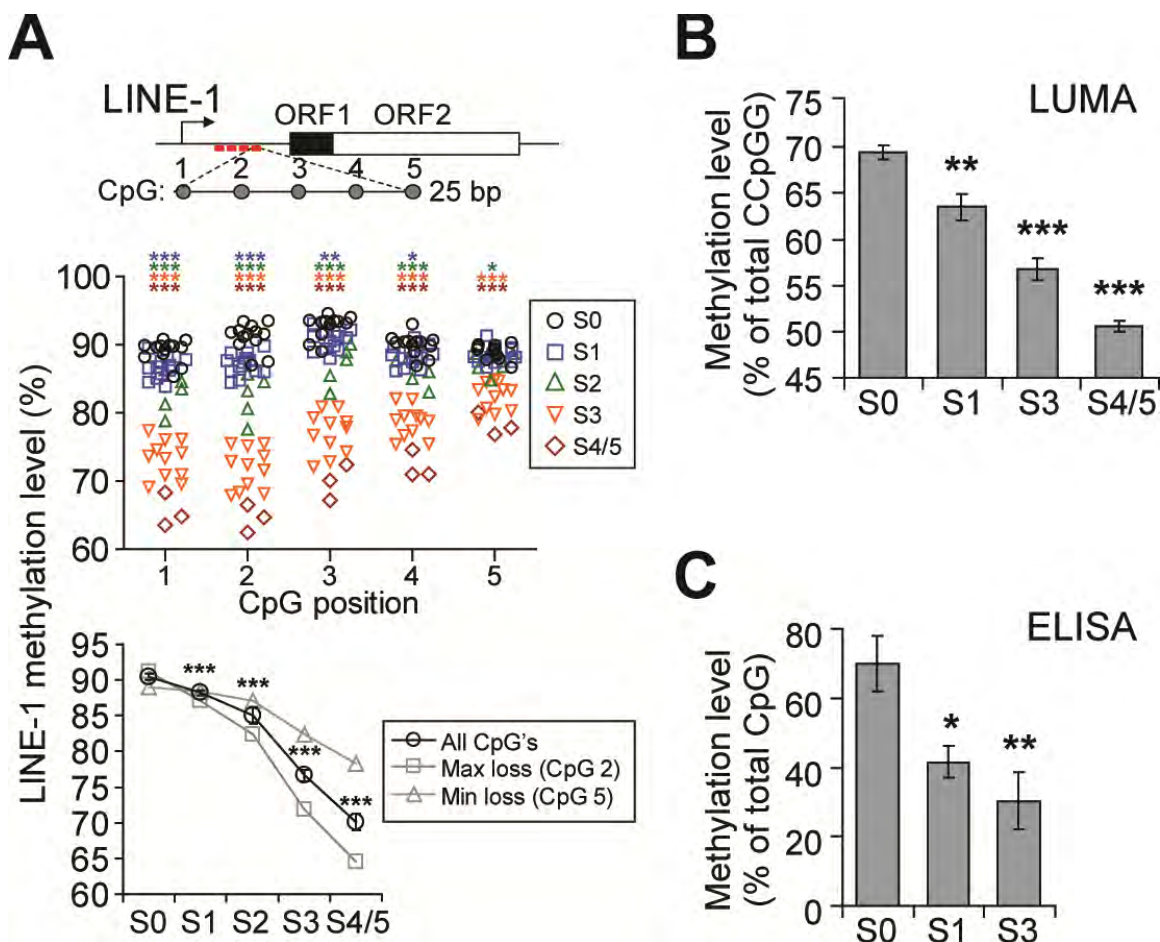
Figure 3.4. Global methylation levels in erythroid and non-erythroid subsets.

A. DNA methylation at LINE-1 retrotransposons in a variety of cell lineages of the early embryo. All cell types were isolated from E12.5 or E13.5 embryos. „LSK" (Lin⁻Sca1⁺Kit⁺) cells contain early hematopoietic progenitors (short and long-term hematopoietic stem cells, and hematopoietic progenitor cells) and the common myeloid and lymphoid progenitors (CMP and LMP). „Mac1⁺" cells are granulocytes and macrophages. „CD41⁺" are mostly megakaryocytes. „EpoR⁻/⁻" are Lin- cells from EpoR⁻/⁻ fetal liver at E12.5. Cells from EpoR⁻/⁻ fetal livers do not differentiate into red cells and resemble the S0 subset in their CD71/Ter119 profile. Nuclei, also known as pyrenocytes, are extruded nuclei from definitive erythrocytes. Yolk sac red cells are the differentiated progeny of the primitive erythropoietic lineage and were isolated from the circulation of E13.5 embryos. Results are mean of n≥5 (S0, S1, S2, S3, S4/5), n=2 (Brain, LSK, Mac1⁺, CD41⁺, Nuclei), n=3 (EpoR⁻/⁻), n=4 (yolk sac); error bars are s.e.m.

B. LINE-1 DNA methylation levels for erythropoiesis in adult bone marrow. ProE and EryB/C subsets of adult bone marrow are equivalent to S2 and S4/5 cells, respectively. Results are means of n=2 (ProE, EryA, EryB, EryC), n=8 (S2), or n=5 (S4/5).
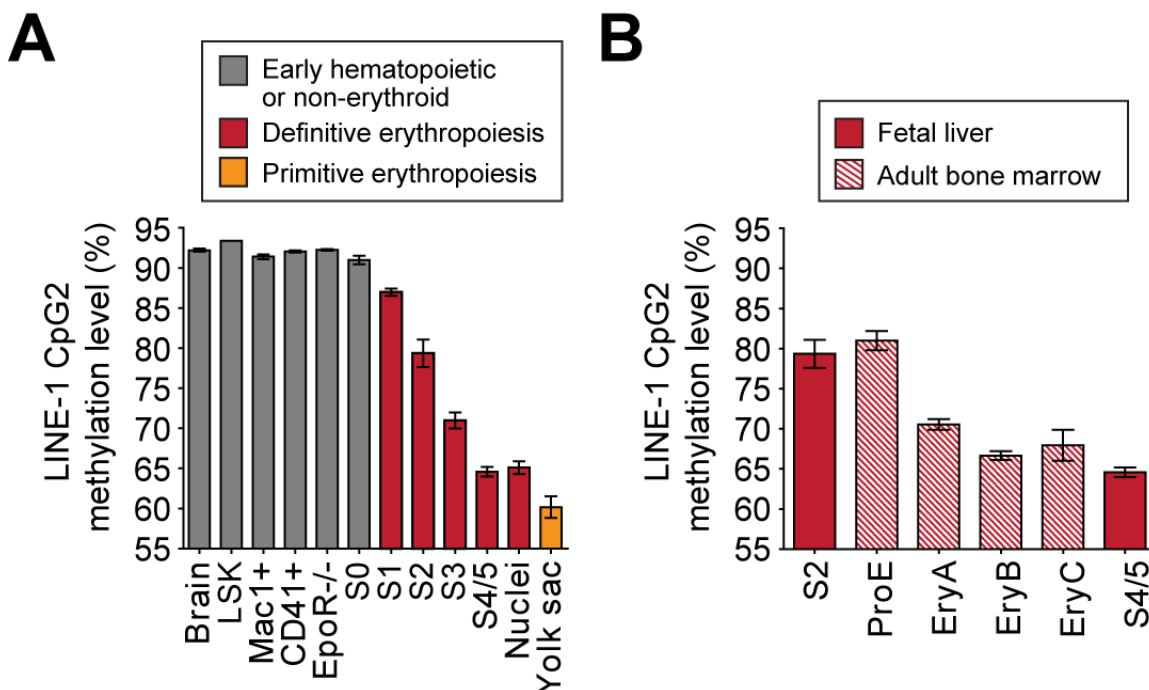
Figure 3.5. Genome-wide reduced representation bisulfite sequencing reveals global loss of DNA methylation during erythropoiesis.

A. Scatterplots of the DNA methylation level of 5 kb, non-overlapping, sliding windows of S0 progenitors (x-axis) vs. increasingly mature S1, S3 and S4/5 erythroblasts (y-axis). Each data point (n=62,844) represents the average CpG methylation in each 5 kb window. Only 5 kb windows that included at least 5 CpGs with at least 5 sequencing reads each for each sample were included in the analysis. Red line represents a constant methylation level between each subset comparison. Data shown is for a single, representative sample.

B. Box plots of replicate averaged RRBS data from „A". Replicate data (n=2) for each sample in „A" were averaged. Box plots, representing n=62,844 unique 5 kb tiles per sample, display the median (black line) and mean (cross) ± central quartiles (boxes); whiskers show the 2.5 to 97.5 percentiles. Median DNA methylation level for each subset is indicated at the top of each box plot.

C. Statistical analysis of 5 kb RRBS data. The methylation level of each 5 kb window in the S0 subset was compared to S1, S3, or S4/5 subset and categorized as hypomethylated, hypermethylated, or no change. Utilizing sample replicate (n=2) and the individual CpG data within each window (n>5), a false discovery rate (FDR) and p-value for each 5 kb window was generated by t-test. A 5 kb window was designated "no change" if the FDR>0.1; „hypomethylated" if FDR<0.1, P<0.001, and methylation change >20%; and „hypomethylated (weak)" if FDR<0.1, methylation change >0% and not „hypomethylated" as defined above. „Hypermethylation" criteria were defined similarly, but with the methylation change in the opposite direction. Percentages are based on n=62,844 unique 5 kb windows per sample comparison.

D. Absolute number of 5 kb windows in each category described in „C"
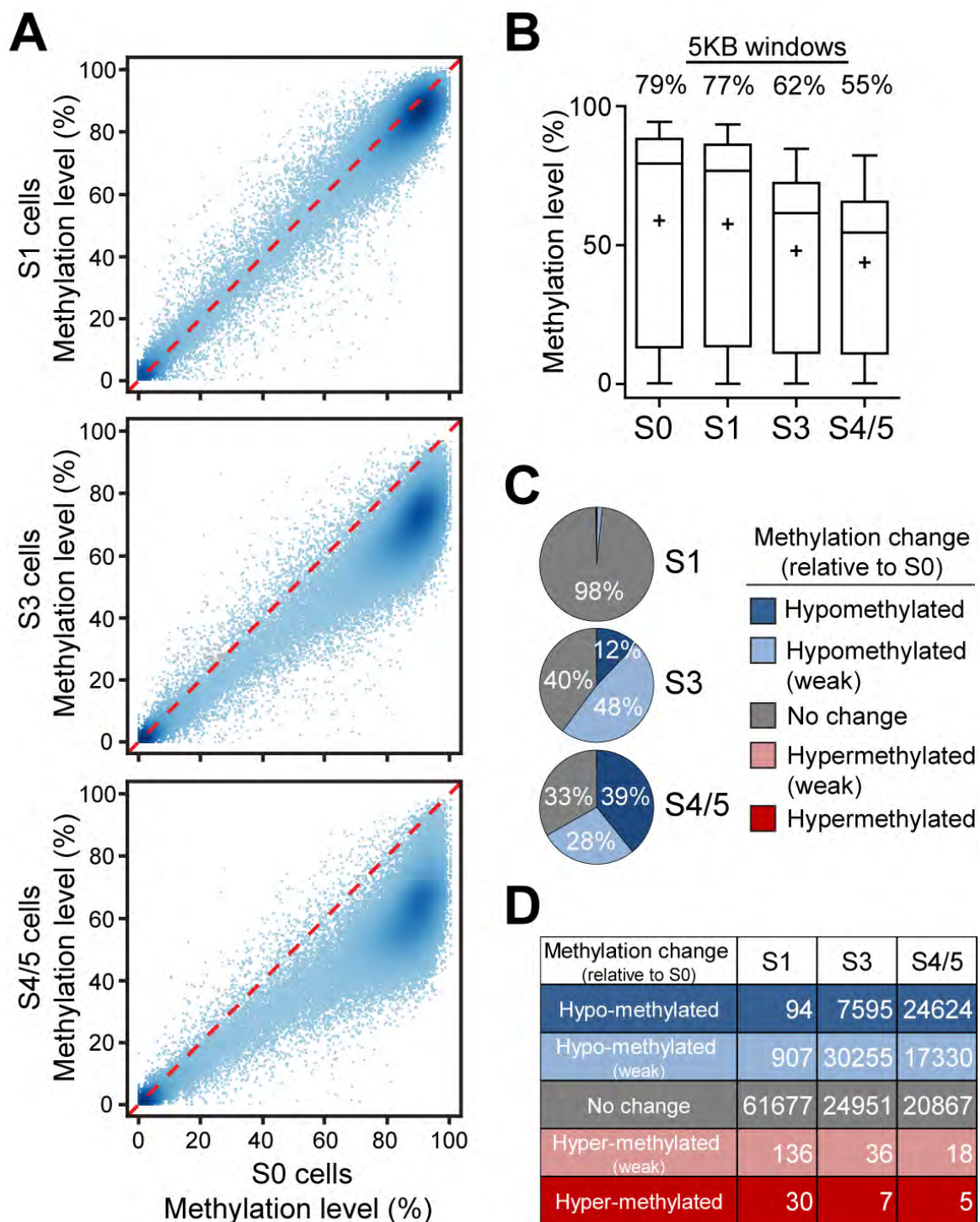
Figure 3.6. DNA methylation loss at various sequence motifs using RRBS.

A. Genome-wide RRBS data for each subset S0 to S4/5 was analyzed at repeat elements; LINE (n=5680 individual genomic sites), SINE (n=163), and LTR (n=3305). Each data point is the mean of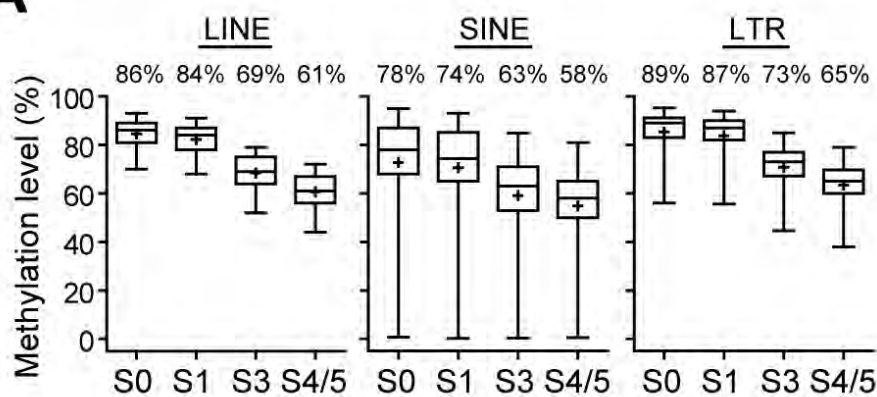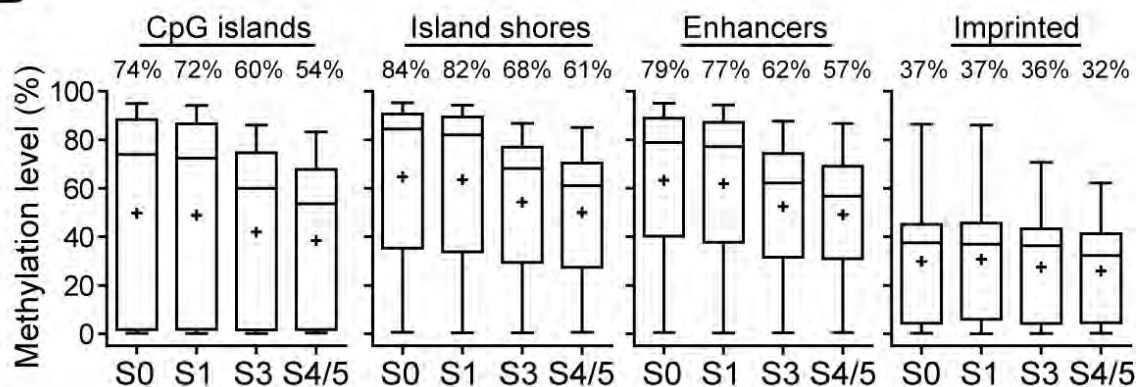 duplicate samples from independent sorting experiments. Data points are shown only for genomic regions with sufficient RRBS coverage (≥5 covered CpGs with at least 5 valid sequencing reads per CpG). Box plots display the median (black line) and mean (cross) ± central quartiles (boxes); whiskers show the 2.5 to 97.5 percentiles. Median DNA methylation level for each subset is indicated at the top of each box plot.

B. Genome-wide RRBS data for each subset S0 to S4/5 was analyzed at functional genomic elements; CpG islands (n=34,558 individual genomic sites), CpG island shores (n=4,302), putative enhancer elements (n=2070), and imprinting-associated regulatory elements (n=72). CpG islands were mapped using the Gardiner-Garden criteria [41], CpG island shores were defined as 2-kb genomic regions that are directly adjacent to *bona fide* CpG islands [358, 359], enhancers are based on histone H3K4me2 peaks [42, 358], and imprinting-associated regulatory elements were manually curated from the literature [358]. Data points and box plots as in „A".

C. Genome-wide RRBS data for each subset S0 to S4/5 was analyzed at promoter regions. Promoter regions are defined as regions -5kb to +1kb surrounding Ensembl-annotated transcription start sites. Promoter regions were analyzed based on CpG frequency; <1% (n=753), 1%< CpG <2% (n=8217), and >2% (n=13,878). Data points as in „A". Boxes correspond to center quartiles with the median marked by a black bar, and whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box.

D. Fraction of S0 methylation lost during erythroid differentiation. Methylation level in S4/5 cells was compared to the initial level of methylation in S0 cells for various assays. Upper grouping, proportion of methylation lost at β-globin LCR regions assayed by bisulfite sequencing in Figure 3.1B. Middle grouping, proportion of methylation lost at various sequence motifs using RRBS data in „A", „B", and Figure 3.5B. Lower grouping, proportion of methylation lost using LUMA assay from Figure 3.3B.

A

LINE
86% 84% 69% 61%

SINE
78% 74% 63% 58%

LTR
89% 87% 73% 65%

Methylation level (%)

S0 S1 S3 S4/5   S0 S1 S3 S4/5   S0 S1 S3 S4/5

B

CpG islands
74% 72% 60% 54%

Island shores
84% 82% 68% 61%

Enhancers
79% 77% 62% 57%

Imprinted
37% 37% 36% 32%

Methylation level (%)

S0 S1 S3 S4/5   S0 S1 S3 S4/5   S0 S1 S3 S4/5   S0 S1 S3 S4/5

C

Promoters CpG <1%

Promoters 1%< CpG <2%

Promoters CpG >2%

Methylation level (%)

S0 S1 S3 S4/5   S0 S1 S3 S4/5   S0 S1 S3 S4/5
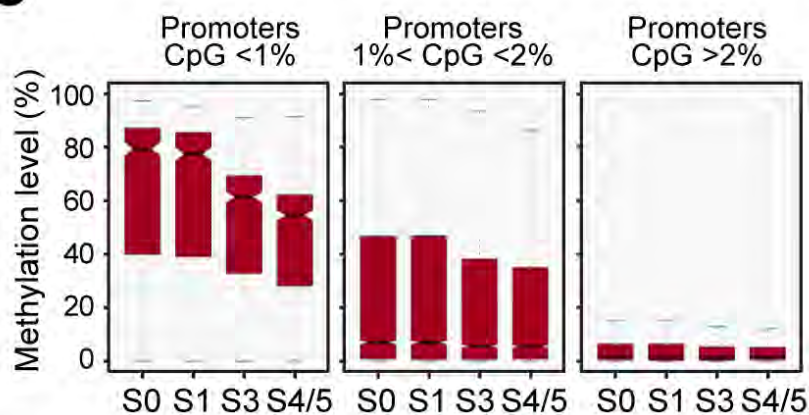
D

Region	Fraction meCpG lost
HS2 CpG B	80%
HS2 CpG A	65%
5 kb window	31%
LINE	29%
Enhancer	28%
Island shores	28%
CpG islands	28%
LTR	27%
SINE	26%
Imprinted	14%
LUMA	27%

Figure 3.7. Linear correlation between DNA methylation levels at the β-globin LCR and at various genomic regions.

For each of subsets S0 to S4/5 the DNA methylation levels at the β-globin LCR (x-axis) were compared to DNA methylation levels at various genomic regions (y-axis) measured by LUMA, pyrosequencing (P), or RRBS (R). Data for each assay is pooled from Figures 3.1-3.6. In the case of pyrosequencing comparisons to β-globin LCR, data from matched samples were used. The slope of this correlation suggests that demethylation at β-globin LCR is 2- to 3-fold greater than global demethylation. Data is mean (LUMA, pyrosequencing) or median (RRBS) of 2 to 12 replicates.
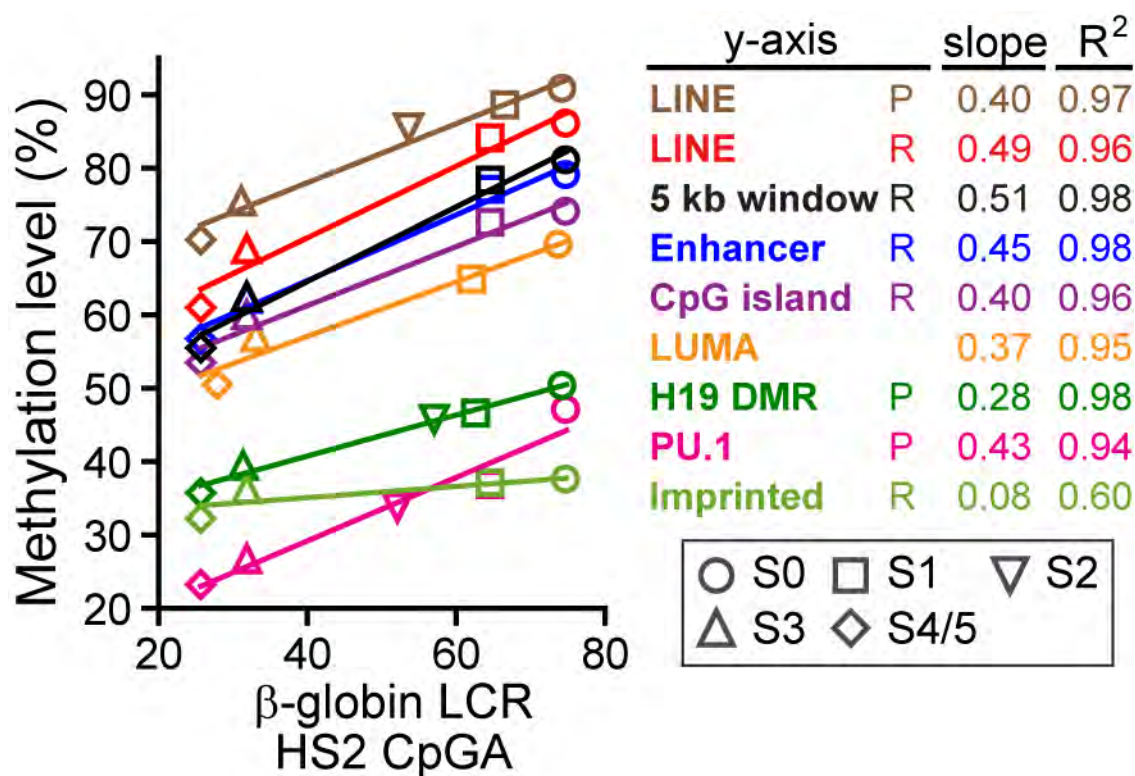
Figure 3.8. Decreased expression of Dnmt3 and increased expression of Gadd45a and Mbd4 with erythroid differentiation.

A. mRNA expression of genes involved in maintaining and establishing DNA methylation (left panel) and in active DNA demethylation (right panel). Activation-induced cytidine deaminase (Aicda) was undetectable in any erythroid subset. Dnmt3a and 3b were undetectable in S4/5. Data (qRT-PCR) is mean ± s.e.m relative to β-actin mRNA and normalized to expression in S0. n=4 replicate experiments. Similar results were obtained when expression was measured relative to Gapdh (not shown). ***P<0.001, and **P<0.01 by 2-tailed Student''s t-test with unequal variance in a comparison of S0 and S3 subset data.

B. Western blotting for DNA methyltransferases in freshly sorted erythroid subsets (left panels) and corresponding quantitation of the same membranes using near-infrared fluorescence detection (right panels), expressed as an absolute value (top) or normalized to β-actin (bottom). An equal mass of cell lysate was loaded in each lane. Control Phoenix cells were used either untransfected (Ctrl) or were transfected with vectors encoding Dnmt3a1, Dnmt3a2, Dnmt3b1, or Dnmt3b3 as indicated.
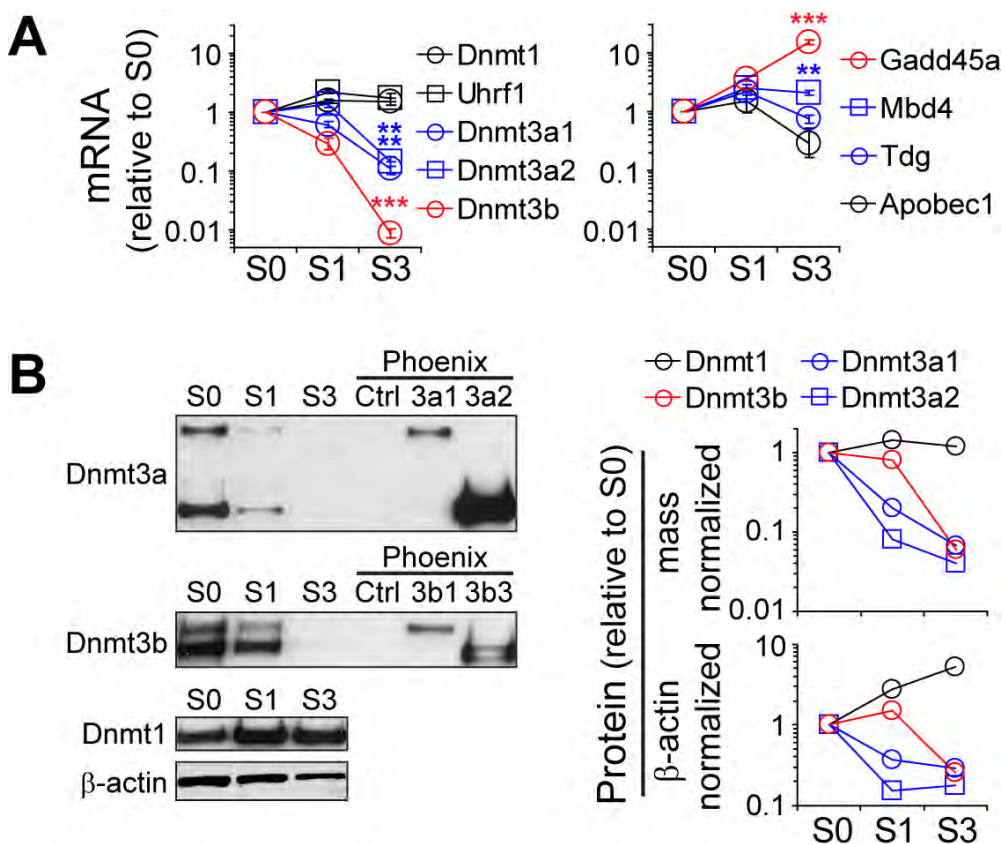
Figure 3.9. Exogenous expression of Dnmt3a or Dnmt3b does not prevent DNA demethylation.

A. Exogenous expression of DNA methyltransferases in freshly sorted S0 cells. S0 cells were transduced with retroviral constructs expressing DNMT-IRES-GFP or 'empty' vector expressing GFP only (MSIG) and cultured in interleukin 3 (IL-3) and stem cell factor (SCF) for 16 hours to allow time for Dnmt and GFP expression. Cells were then transferred to Epo for 24 hours to allow for differentiation into the S1 and S3 subsets. Total GFP expression at the end of the incubation period is shown for each of the retroviral supernatants used (upper-right panel). GFP$^+$ cells in each of S0, S1, and S3 gates (lower panel) at the end of the incubation period were isolated by sorting and used for mRNA or DNA methylation analysis described in „B‟ and „C‟. Sorting and analysis experiments in „B-E‟ were performed using supernatant dilutions of 1:200 for the empty vector and 1:25 for Dnmt3a in order to obtain equivalent GFP expression to cells transduced with Dnmt3b1 and Dnmt3b1:PC. Flow cytometry profiles are representative of 3 to 6 replicate experiments for each viral supernatant.

B. Dnmt3 expression is maintained throughout differentiation in retrovirally infected cells. Expression of Dnmt3a (left panel) and Dnmt3b (right panel) mRNAs in GFP$^+$ cells sorted from each of gates S0, S1 and S3, at the end of a 24 hr culture in Epo in an experiment as illustrated in 'A'.

C. DNA methylation is unchanged by exogenous expression of Dnmt3. Methylation levels of LINE-1 elements and of the β-globin LCR in GFP$^+$ cells sorted from each of gates S0, S1 and S3, at the end of a 24 h culture period in Epo, in the experimental strategy illustrated in 'A'. Results are mean ± s.e.m. n=3 for infection with the MSIG, Dnmt3a, or the Dnmt3b1 supernatants or n=2 for infection with the Dnmt3b1:PC supernatant. Statistical significance was determined by comparing Dnmt3 infection to MSIG within each sorting gate. ***P<0.001, **P<0.01 by 2-tailed Student‟s t-test with unequal variance.

D. Exogenous expression of Dnmt3b1 or Dnmt3b1:PC prevents erythroid differentiation. The percent of cells was calculated in each of gates S0, S1 and S3, at the end of a 24 h culture period in Epo, in an experimental strategy illustrated in 'A' above. Results are mean ± s.e.m. n=6 (MSIG), n=5 (Dnmt3a), n=6 (Dnmt3b1), and n= 3 (Dnmt3b1:PC). Statistics as described in „C‟.

E. Dnmt3b expression induces cell death in S3 cells. GFP$^+$ cells in each of gates S0, S1 and S3, at the end of a 24 h culture period in Epo as illustrated in 'A' above, were analyzed for DAPI staining. Results are mean ± s.e.m. n=2. Statistics as described in „C‟.
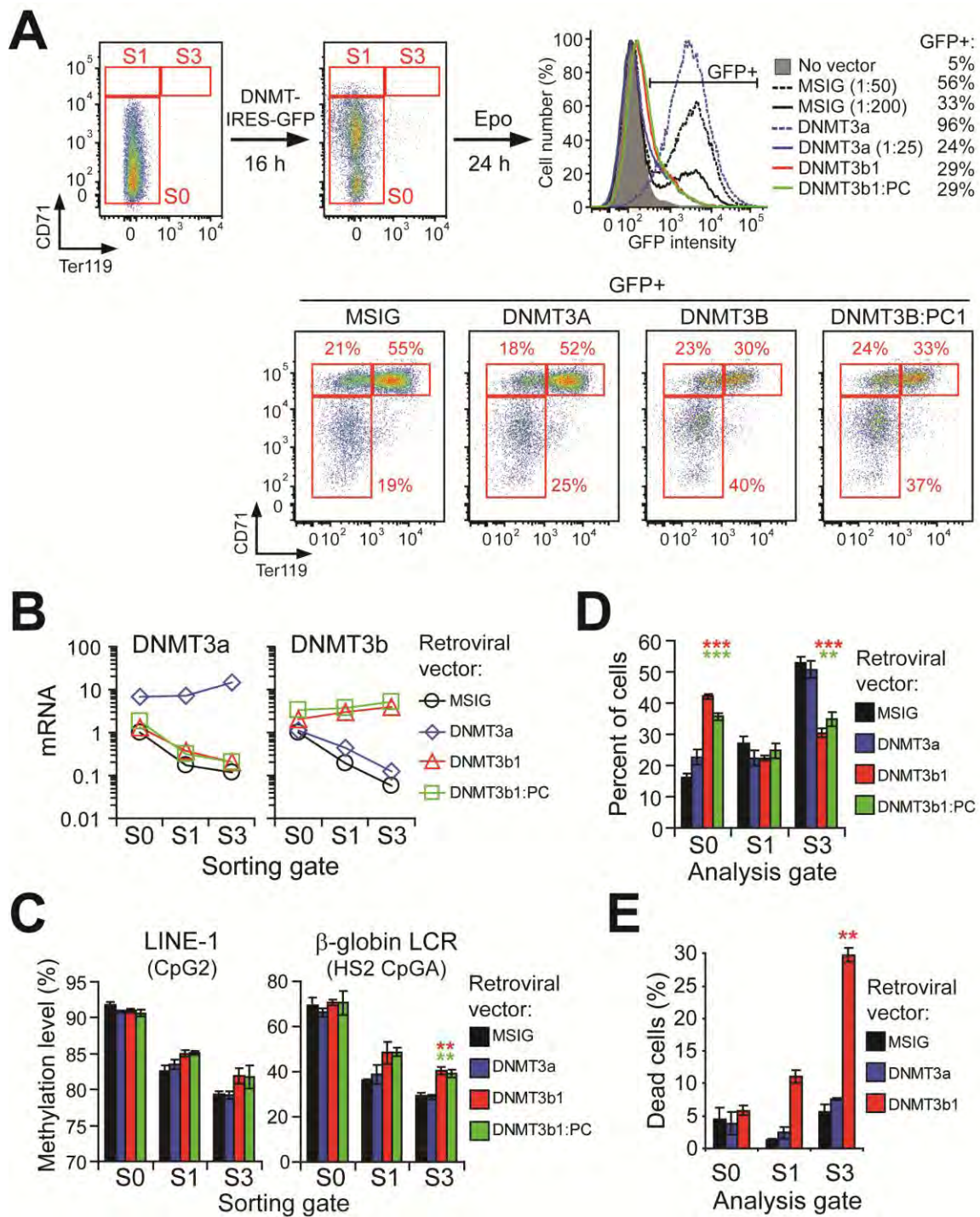
Figure 3.10. Knockdown of Gadd45a or Mbd4 does not prevent DNA

demethylation.

   A. Efficiency of shRNA-mediated knockdown. Freshly sorted S0 cells were transduced with a retroviral construct expressing shRNA (linked by IRES to GFP) directed against either Mbd4 (shMbd4) or Gadd45a (two different hairpins, shG45a-1 or shG45a-2), or a non-silencing shRNA control (shCtrl). Cells were cultured with viral supernatant in IL-3 and SCF for 16 hours to allow time for shRNA and GFP expression. Cells were then transferred to Epo for 24 hours to allow for differentiation into „CD71^{+}" cells (CD71^{+} = combined S1, S2, and S3 subsets). GFP^{+}CD71^{+} cells were sorted and evaluated for mRNA levels of Gadd45a (left panel) or Mbd4 (right panel). Data (qRT-PCR) is expressed relative to Gapdh mRNA and normalized to expression in S0 t=0.

   B. DNA methylation levels. Samples, as described in „A" were assayed for DNA methylation at LINE-1 regions or the β-globin LCR.
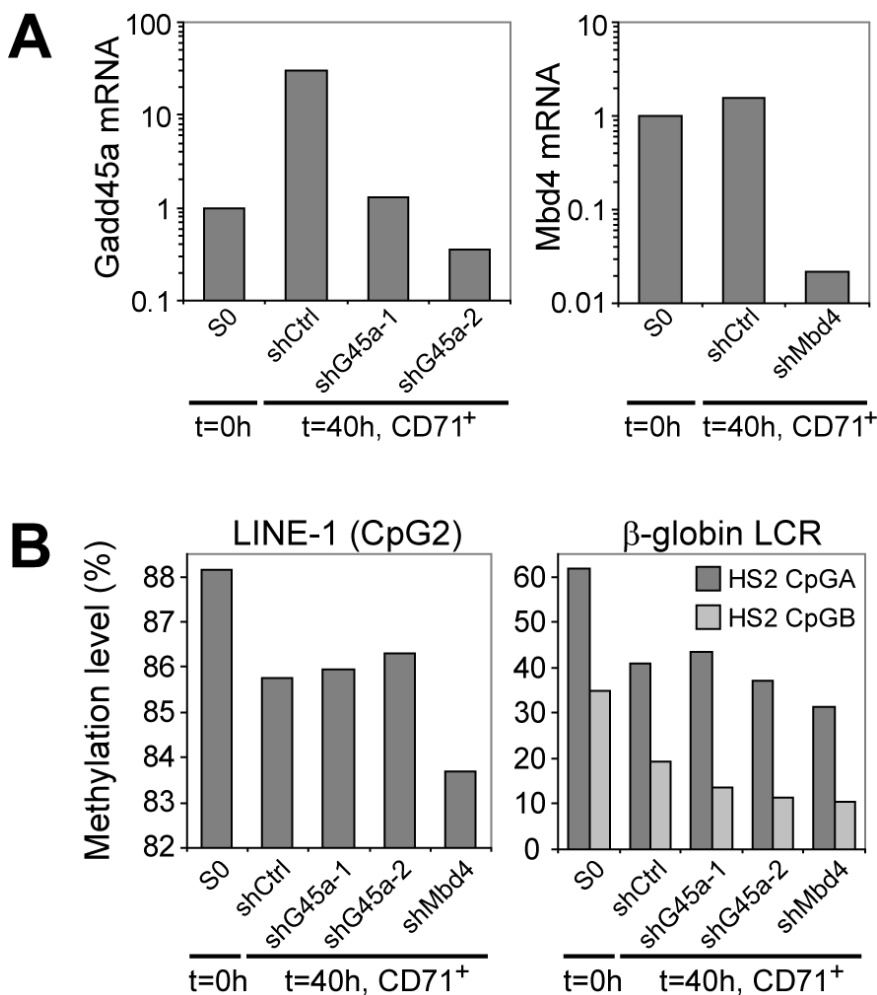
Figure 3.11. Cell cycle status of erythroid subsets S0 to S4/5.

A. Fetal livers were labeled *in vivo* with BrdU by injecting pregnant mice 30 minutes prior to harvest of embryos. Cells were fixed and then stained for CD71, Ter119, BrdU, or total DNA content (7-AAD). BrdU and 7-AAD staining is shown for each of S0 to S4/5 subsets. Red gates show cells in the G1 (bottom left), S (top), or G2/M (bottom right) phases of the cycle. For each subset, both the fraction of cells in S-phase, and the median fluorescent intensity (MFI) of the BrdU signal within the S-phase gates, are indicated. MFI is an indicator of DNA replication rate. Representative of 8 to 15 replicate experiments. Color mapping represents the highest to lowest cell frequency as spectrum from red to blue, respectively.

B. Percent of cells in S-phase in each subset. Data is mean ± s.e.m of 8 to 15 replicate experiments per subset.

C. Median fluorescent intensity of S-phase cells. Data is mean ± s.e.m of 8 to 15 replicate experiments per subset.
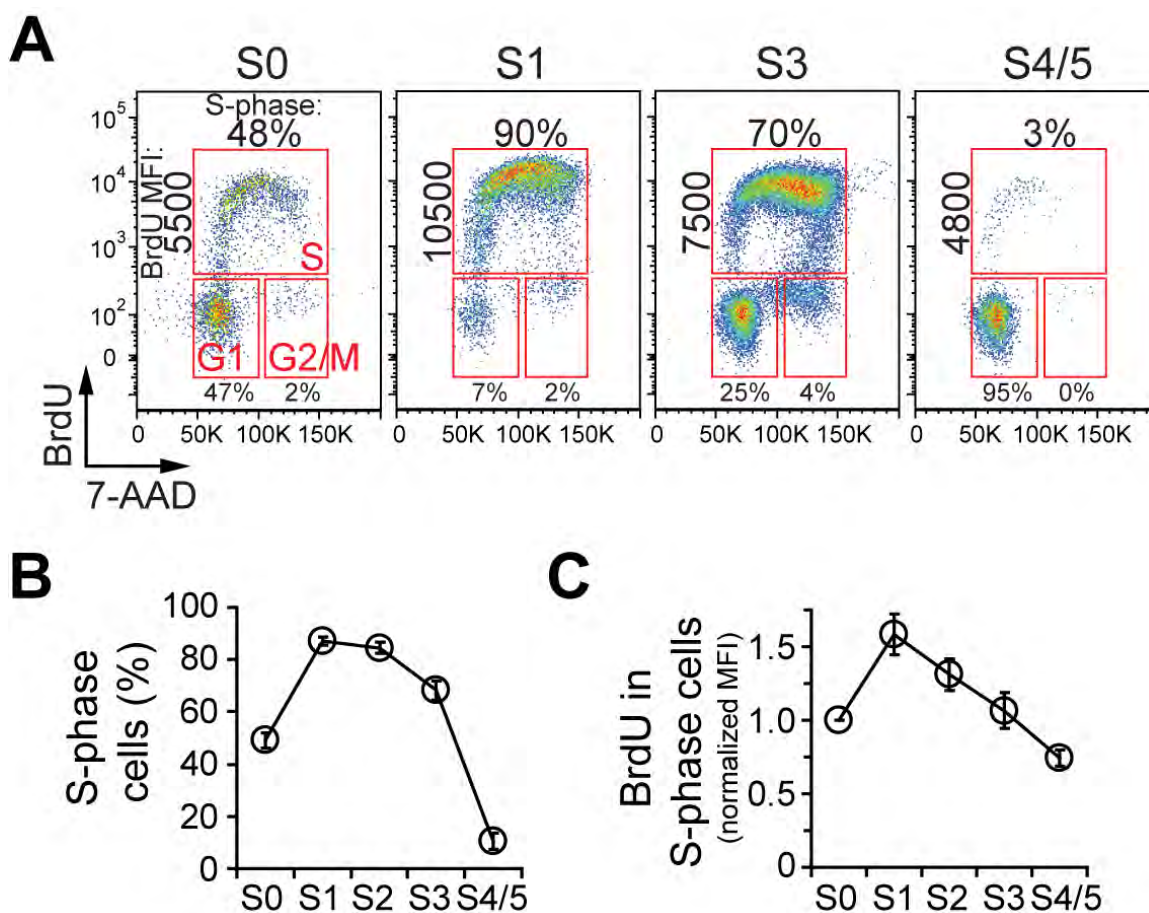
Figure 3.12. Mimosine treatment prevents DNA methylation loss at LINE-1 regions,

and at the β-globin LCR, PU.1 and H19 DMR loci.

A. Cell cycle block with 1mM mimosine (Mim) reversibly arrests DNA demethylation at LINE-1 elements and at the β-globin LCR during erythroid differentiation *in vitro*. Sorted S0 cells were cultured in Epo for 24 hr (black line); Mim was added at t=4 and removed at t=14 hr (red line). Data is mean ± s.e.m from three independent experiments.

B. Mimosine prevents DNA replication in cultured S0 cells. Sorted S0 cells were cultured in Epo for 16 hr in the presence or absence of 1 mM mimosine. BrdU was added in the last 30 minutes of culture. BrdU incorporation and total DNA content (7-AAD) were measured by flow cytometry. Upper red gate contains cells in S-phase that are actively synthesizing DNA at 16 hr. Lower left and right gates contain G1 and G2/M phase cells, respectively, or cells not actively synthesizing DNA at 16 hr. Contour density indicates relative cell frequency.

C. Methylation level at the LINE-1, β-globin LCR, PU.1, and H19 DMR loci for the experiment described in „B". Red dotted line is the mean methylation level of freshly sorted S0 cells at t=0. Formatting and y-axis scaling of plots was chosen to be equivalent to that of Figure 3.13C. Exp, experiment. Ave, average of replicate experiments.
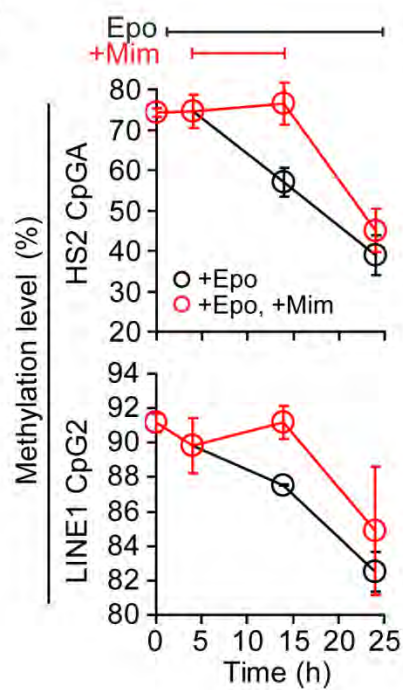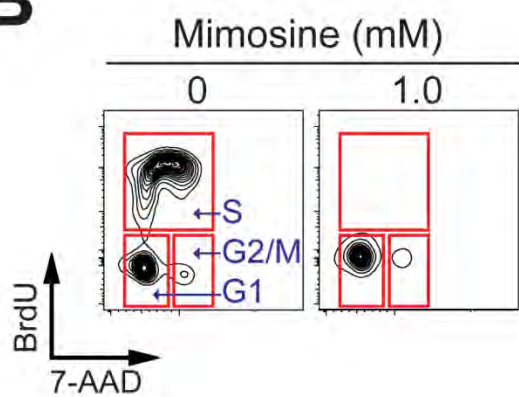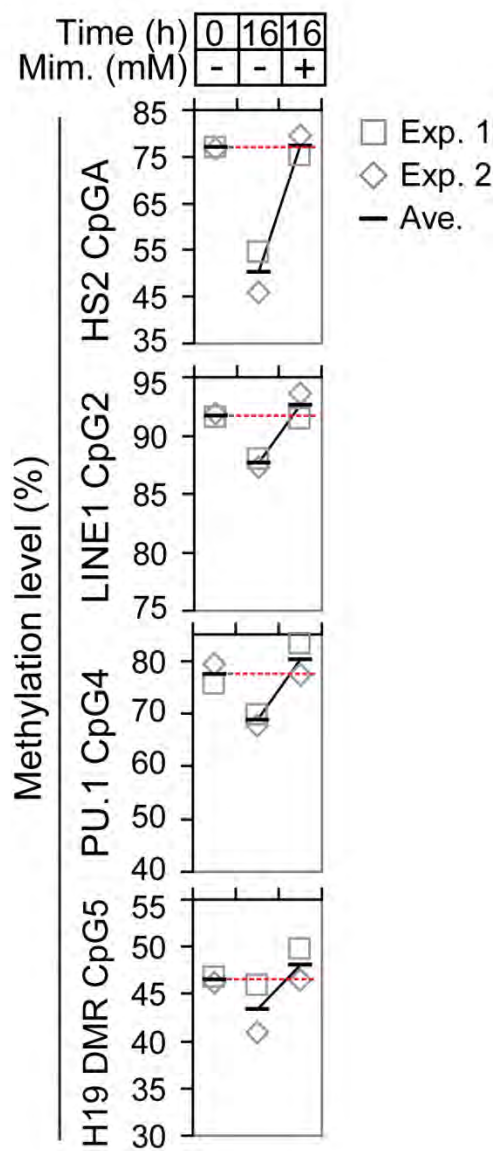
Figure 3.13. DNA methylation loss at LINE-1 regions, and at the β-globin LCR,

PU.1 and H19 DMR loci, is dependent on DNA replication.

 A. DNA replication arrest with aphidicolin reversibly prevents demethylation. Sorted S0 cells were cultured with Epo (black line) or Epo plus 1 μM aphidicolin (red line) for 10 h. Aphidicolin was removed and culture continued for an additional 36 hr in Epo. DNA methylation was determined at the β-globin LCR and global LINE-1 regions. Exp, experiment. Ave, average of replicate experiments. Aphi, aphidicolin.

 B. Decreasing DNA replication rate with increasing amounts of aphidicolin. Freshly sorted S0, S1, or S2 cells were incubated in Epo or Epo plus 0.1, 0.5, or 1.0 μM aphidicolin for 16 h. BrdU was added in the last 30 min of culture. BrdU incorporation and total DNA content (7-AAD) were measured by flow cytometry. MFI of BrdU in S-phase cells is shown, as is the fraction of cells in S-phase (%). Contour density indicates relative cell frequency.

 C. Demethylation is inhibited by decreasing the DNA replication rate. For the experiment described in „B‟, a portion of cells were removed just prior to BrdU addition. Bisulfite pyrosequencing was used to determine methylation levels at the β-globin LCR, LINE1, PU.1, and H19 DMR. Red dotted line is the mean methylation level of freshly sorted S0, S1, or S2 cells at t=0.
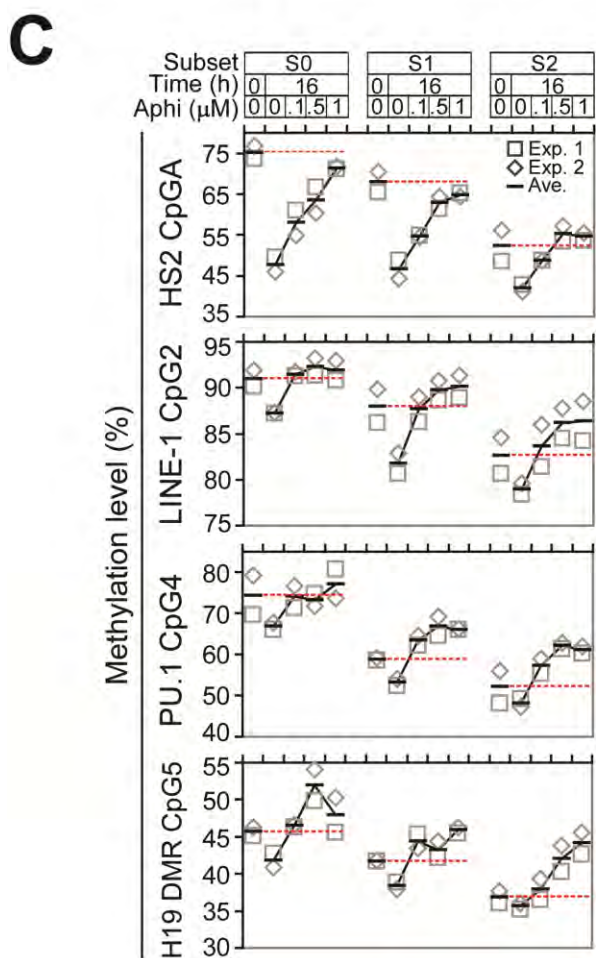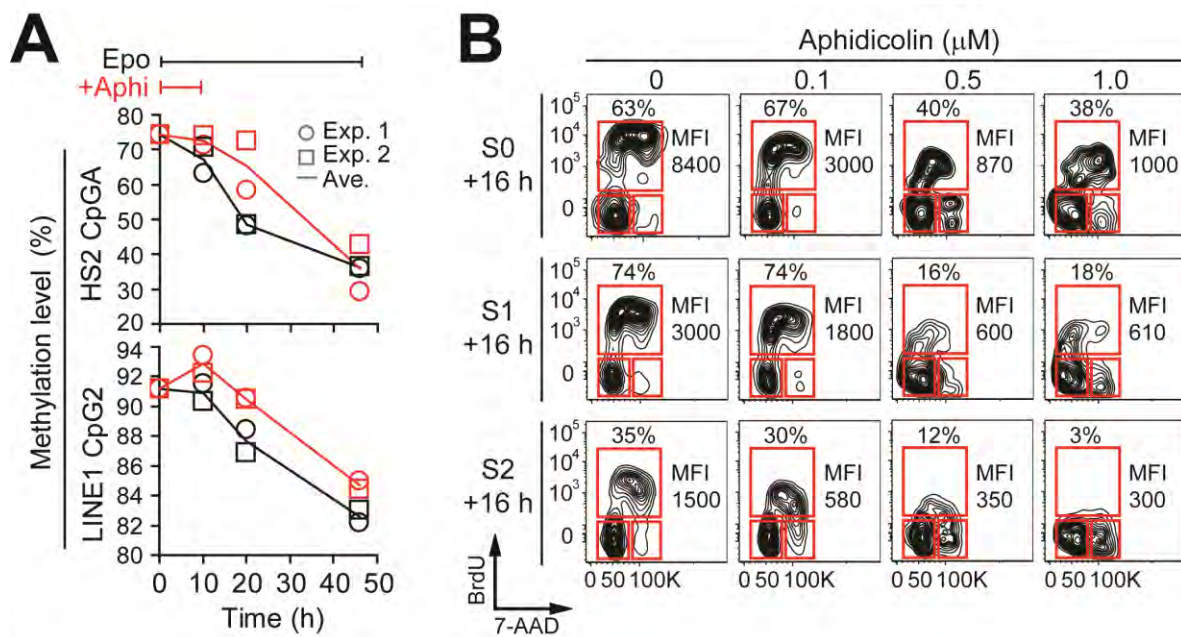
Figure 3.14. Treatment of differentiating S1 cells with low dose aphidicolin prevents genome-wide demethylation.

A. Freshly sorted S1 cells were cultured in the absence (left panel) or presence (right panel) of low dose aphidicolin (0.1 µM) for 24 hours (y-axis) and compared with their methylation levels at the time of sorting (t=0, x-axis) using RRBS. Scatterplots show the average CpG methylation level of 5 kb, non-overlapping, sliding windows (n=62,844). Data points are shown only for genomic regions with sufficient RRBS coverage (≥5 covered CpGs with at least 5 valid sequencing reads per CpG). Red line represents a constant methylation level between each comparison. Data shown is for a single, representative sample.

B. Box plots of replicate averaged RRBS data from „A". Replicate data (n=2) for each sample in „A" were averaged. Box plots, representing n=62,844 unique 5 kb tiles per sample, display the median (black line) and mean (cross) ± central quartiles (boxes); whiskers show the 2.5 to 97.5 percentiles. Median DNA methylation level for each subset is indicated at the top of each box plot.

C. Genome-wide RRBS data for each sample in „A" was analyzed at repeat elements; LINE (n=5680 individual genomic sites) and SINE (n=163). Data points and box plots as in „B".

D. Genome-wide RRBS data for each sample in „A" was analyzed at functional genomic elements; CpG islands (n=34,558 individual genomic sites), putative enhancer elements (n=2070), and imprinting-associated regulatory elements (n=72). Data points and box plots as in „B".

A Epo — Epo+Aphi (0.1 μM)

S1 t= 24 h Methylation level (%)

S1 t= 0 h, methylation level (%)

B 5KB windows
76.8 68.6 75.0

Methylation level (%)

0 h   24 h   24 h +Aphi

C SINE
74.3 68.0 73.0

Methylation level (%)

0 h   24 h   24 h +Aphi

LINE
84.0 76.0 82.0

0 h   24 h   24 h +Aphi

D CpG islands
72.4 65.9 70.3

Methylation level (%)

0 h   24 h   24 h +Aphi

Enhancers
77.1 69.4 75.5

0 h   24 h   24 h +Aphi

Imprinted regions
37.0 32.3 34.3

0 h   24 h   24 h +Aphi
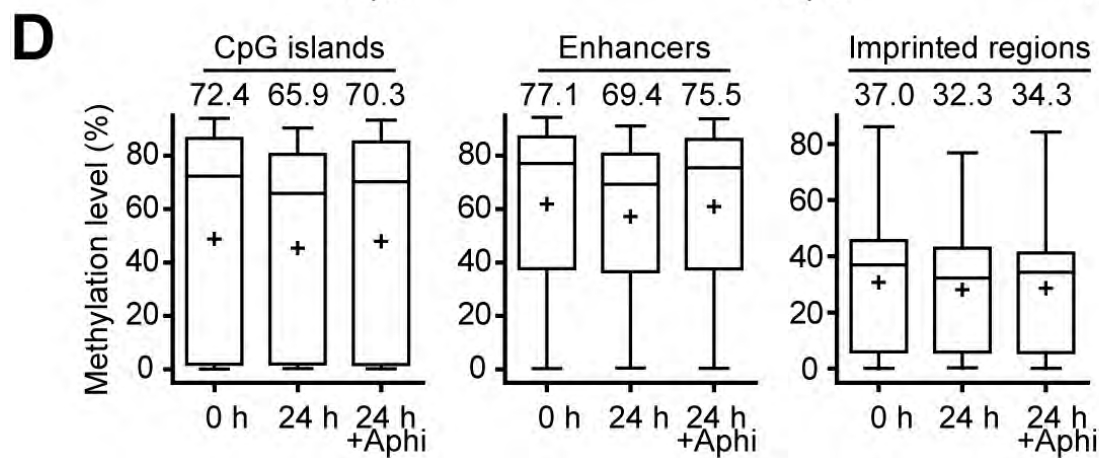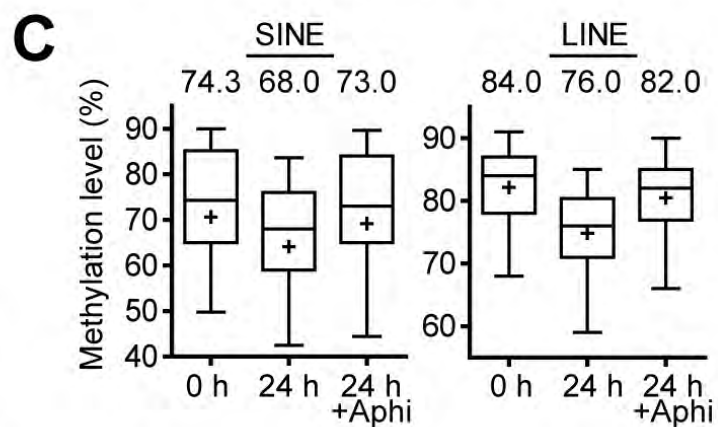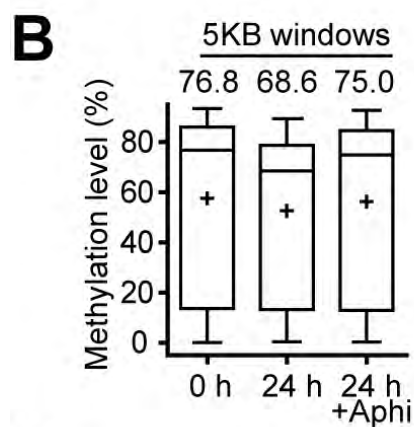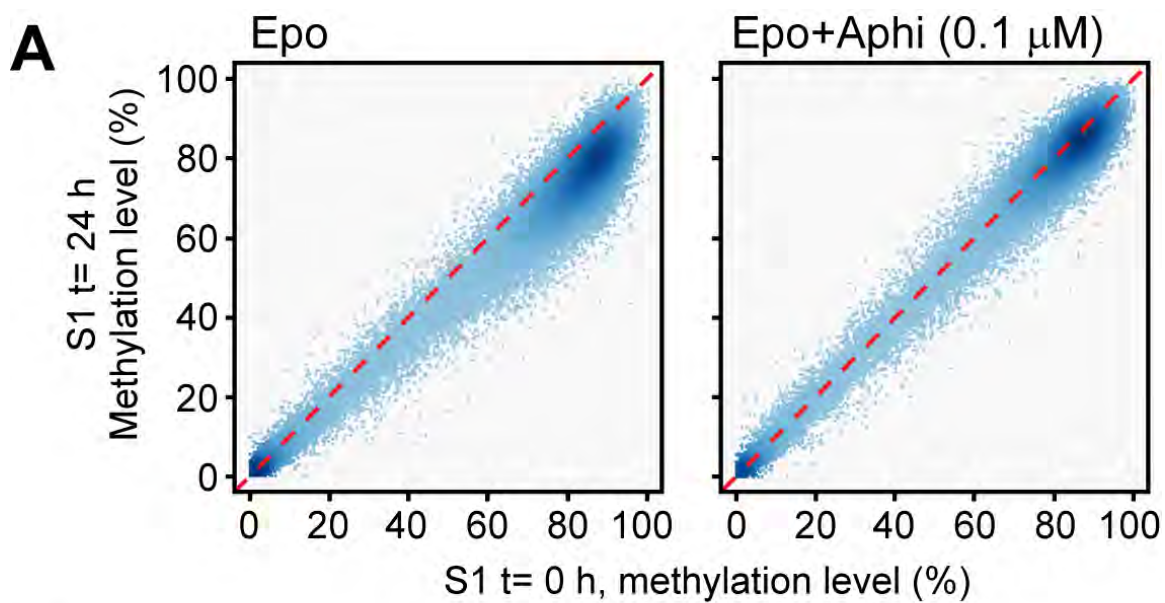
Figure 3.15. Accelerated demethylation induced by inefficient knockdown of Dnmt1.

A. Efficiency of shRNA-mediated knockdown. Freshly sorted S0 cells were transduced with retroviral constructs containing shRNAs (linked by IRES to GFP) against Dnmt1 (D1) or a non-silencing shRNA control (Ctrl). Cells were cultured in IL-3 and SCF for 16 hours to allow time for shRNA and GFP expression. Cells were then transferred to Epo for 24 hours to allow for differentiation. Cells (not sorted) were evaluated for Dnmt1 mRNA by qRT-PCR (left panel) or for Dnmt1 protein by western (right panel). For western blot, an equal mass of cell lysate was loaded in each lane. Dnmt1 and β-actin was quantified using near-infrared fluorescence detection. For both mRNA and protein, Dnmt1 levels have been expressed relative to β-actin and normalized to the non-silencing control at t=16 h.

B. Effect of Dnmt1 knockdown on LINE-1 methylation. S0 cells were cultured with shRNA viral supernatants described in „A" for 16 hours in IL-3 and SCF. Cells were then shifted to differentiation media containing Epo or Epo plus low dose aphidicolin (Aphi) for an additional 24 hours. After differentiation, GFP positive cells from S1 or S3 gates (as defined in Figure 3.9A) were sorted and analyzed for Dnmt1 mRNA by qRT-PCR (top panel) or LINE-1 methylation (bottom panel). Dnmt1 mRNA has been expressed relative to Gapdh and normalized to the non-silencing control at t=16 h. Results are representative of 2 or 3 experiments for data points with or without aphidicolin, respectively.
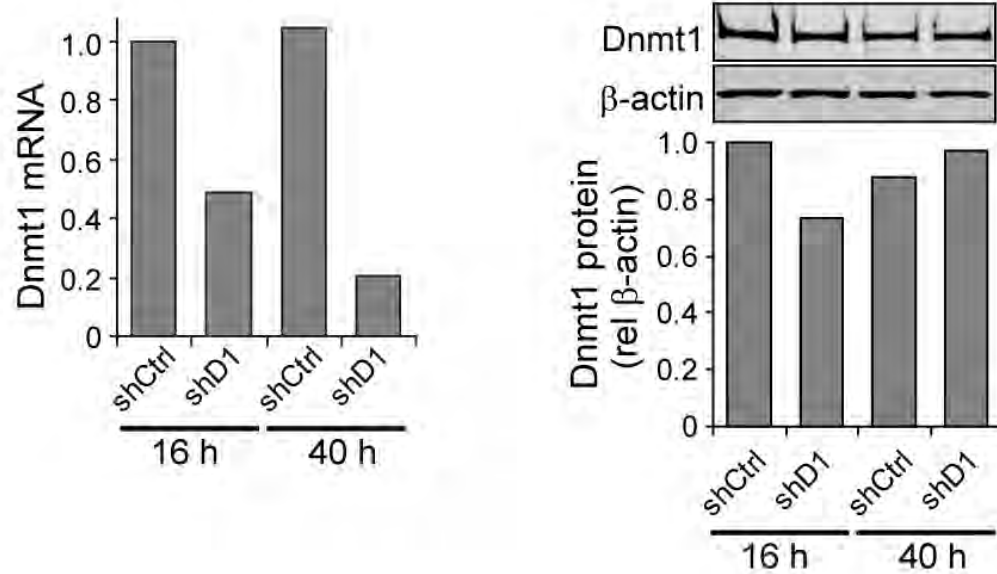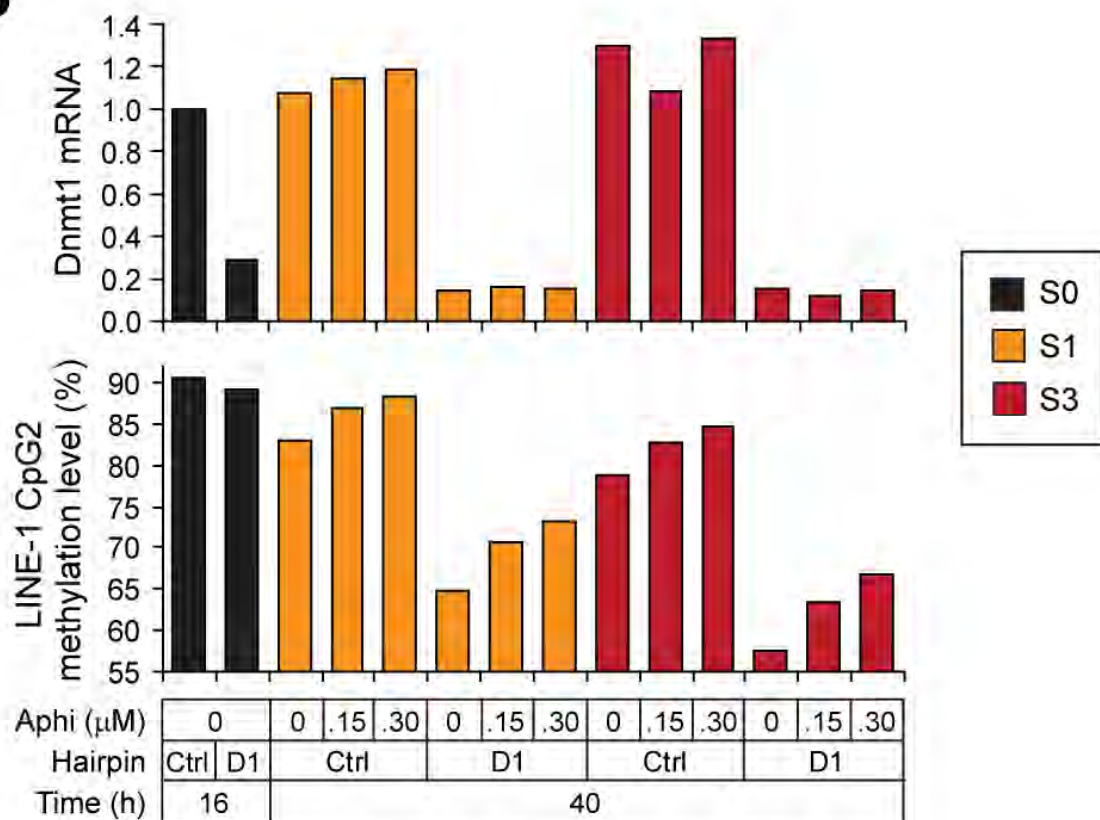
Figure 3.16. Global methylation loss does not induce genome-wide transcriptional activation.

A. Expression of LINE-1 RNA during erythroid differentiation. QRT-PCR for LINE-1 RNA in freshly sorted erythroid subset. Probes were designed to the 5" and 3" ends of the full length LINE-1 RNA (GenBank: M13002.1). Results were normalized to the indicated housekeeping gene and then expressed relative to the S0 subset. Data are mean ± s.e.m. of 5" and 3" probes for 4 biological replicates.

B. No global increase in gene expression as a result of global demethylation. Gene expression in S1 (x-axis) and S3 (y-axis) subsets was measured using Affymetrix GeneChips. Data was processed using microarray suite 5 (MAS5) algorithm and then normalized to the average housekeeper gene expression (see methods).Transcripts displayed have a random variance t-test of $p < 0.005$ and absolute fold change >2; n=3. Genes up-regulated (n=146) or down-regulated (n=761) in S3 relative to S1 are in green and red respectively.

C. DNA methylation (5 kb tiling regions) for genes that exhibit increased vs. decreased expression between erythroblast stages S1 and S3. GeneChip data described in „B" was re-analyzed using the gcRMA algorithm (see methods). Differentially regulated transcripts were selected based on FDR <0.05 and absolute fold change >2; n=3. The methylation levels for up-regulated (n=2268) or down-regulated (n=5862) transcripts are represented as box plots. Boxes correspond to center quartiles with the median marked by a black bar, and whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box.

A

LINE-1 RNA (relative to S0)

- -○- Actin normalized
- -□- GAPDH normalized

0.9
1.5
1.7
0.5

S0 S1 S3

B

S3 mRNA (RFU)

146 genes up

761 genes down

S1 mRNA (RFU)

C

Methylation level (%)

Upregulated genes
n=2268

S0 S1 S3 S4/5

Downregulated genes
n=5862

S0 S1 S3 S4/5
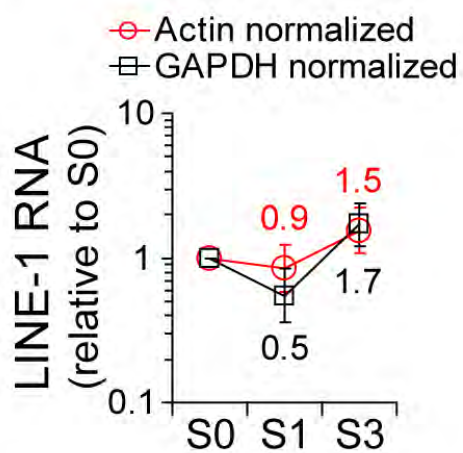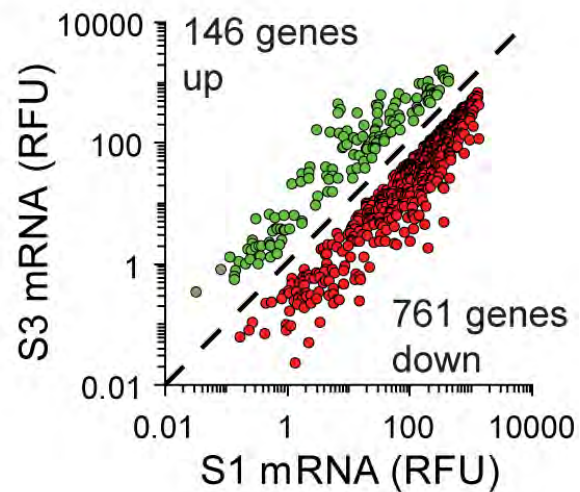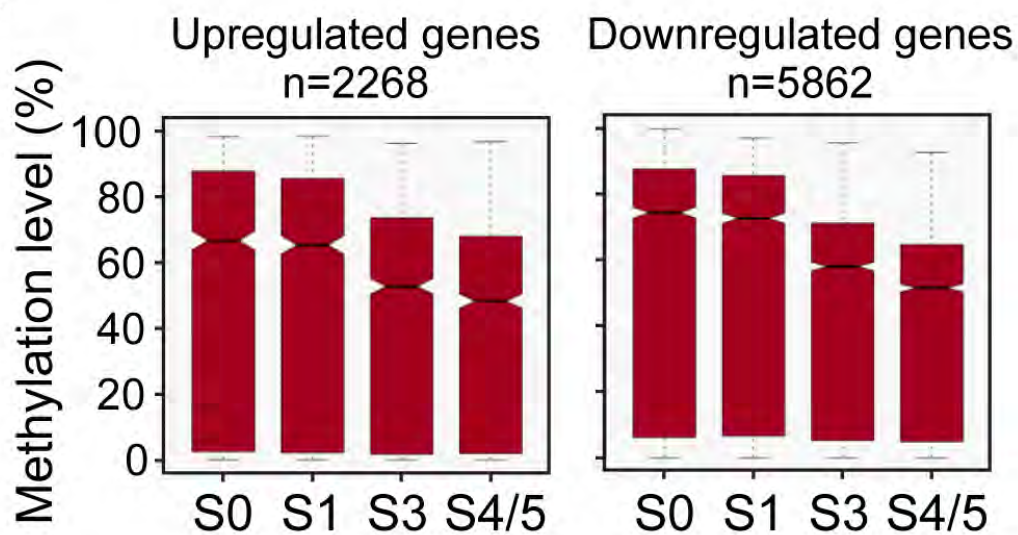
Figure 3.17. Rapid demethylation is required for rapid erythroid gene induction.

       A. Cell cycle status of aphidicolin and 5-aza-2-deoxycytidine treated cells. Freshly sorted S0 cells were allowed to differentiate for 24 hour in Epo into CD71$^+$ cells (CD71$^+$= combined S1, S2, and S3 cells), in the presence or absence of low dose (0.1 µM) aphidicolin (Aphi) and/or 0.05 µM 5-aza-2''-deoycytidine (5-aza), as indicated. Cell cycle status was assessed by pulsing with BrdU for 30 minutes at the end of incubation. BrdU incorporation and total DNA content (7-AAD) for CD71$^+$ cells were measured by flow cytometry. MFI of BrdU in S-phase cells is shown, as is the fraction of cells in S phase (%). Low dose aphidicolin did not alter the percent of cells in S-phase but did slow down the rate of DNA synthesis. 5-aza had no significant cell cycle effect.

       B. DNA methylation levels in CD71$^+$ cells treated in panel „A‟. At 24 h, CD71$^+$ cells were sorted by flow cytometry and DNA methylation determined by bisulfite pyrosequencing. Treatment of S0 cells with aphidicolin prevents methylation loss at LINE-1 elements and delays methylation loss at the β-globin LCR. 5-aza is able to reverse the effect of aphidicolin on DNA methylation. Results are mean ± s.e.m of 2 independent experiments.

       C. Expression of erythroid genes in CD71+ cells treated as in panels „A‟. QRT-PCR of β-globin, Alas2, and Band3 for CD71+ cells isolated at 24 hr (same samples as in „B‟) was normalized to housekeeper genes (mean of Gapdh and β-actin mRNAs) and expressed relative to each aphidicolin negative control. Results are the mean value ± s.e.m. of 2 independent experiments. The aphidicolin-induced delay in erythroid gene expression is reversed in the presence of 5-aza.

A Epo | Epo+Aphi | Epo+5-aza | Epo+Aphi +5-aza

BrdU / 7-AAD

B Methylation level (%)

β-globin LCR HS2 CpG A

β-globin LCR HS2 CpG B

LINE-1 CpG2

Aphi	-	-	+	-	+
5-aza	-	-	-	+	+
time (h)	0		24		

C mRNA

	Aphi	5-aza
	-	-
	+	-
	-	+
	+	+

β-globin Band3 Alas2
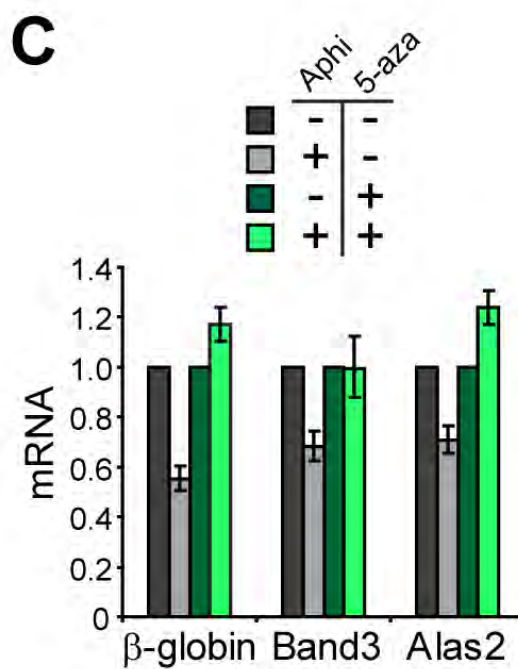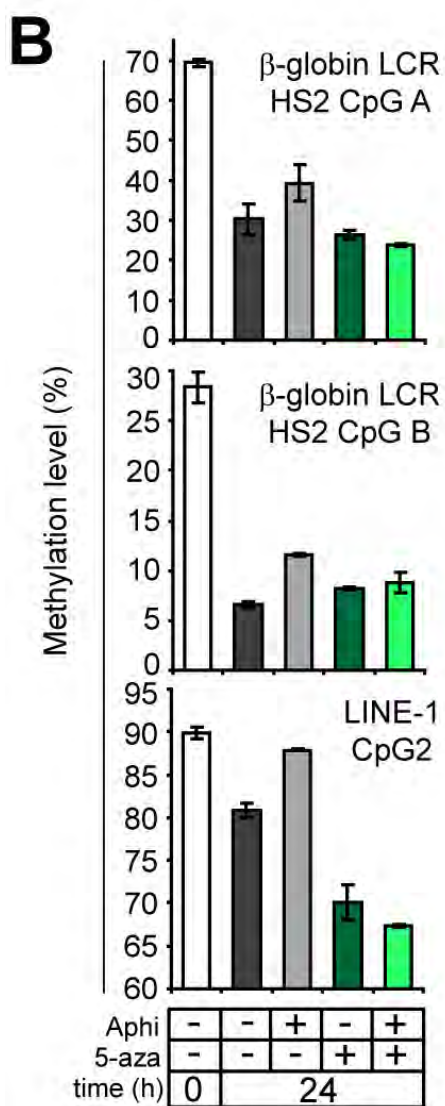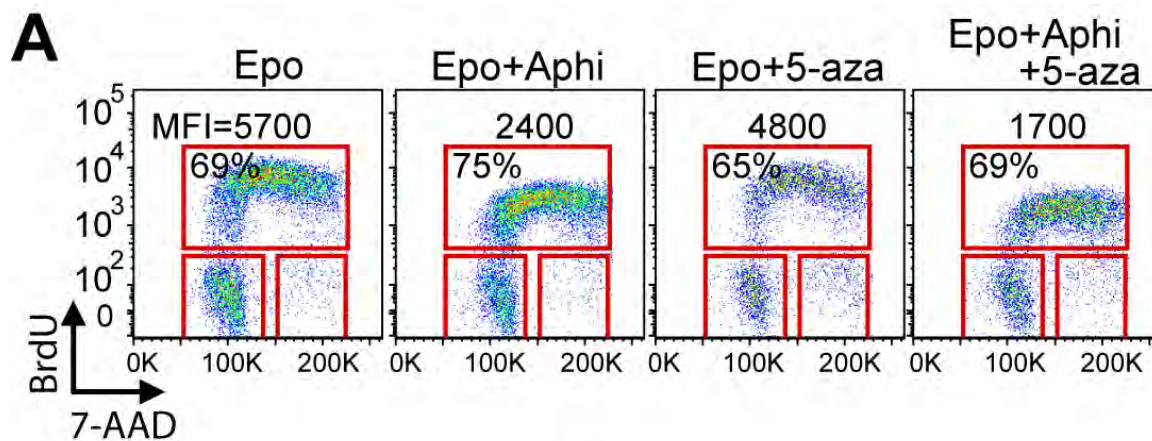
CHAPTER IV

Concluding Remarks and Future Directions

Concluding Remarks

We have utilized a recently developed flow cytometry method to classify fetal liver cells into 6 distinct subsets, each at an increasingly mature stage of erythroid differentiation. We found that up-regulation of CD71 identifies cells that are synchronized in S-phase of a single cell cycle and is associated with a number of differentiation milestones, whose precise timing in erythroid development was previously unknown. These include the onset of Epo dependence, activation of GATA-1 function, and the opening up of chromatin at the β-globin locus. Specifically, our work identified the timing of, and provided mechanistic insight into, DNA demethylation at the β-globin LCR *in vivo*; demonstrating that it is initiated with up-regulation of CD71, continues in subsequent maturation stages, precedes up-regulation of β-globin, and is dependent on S-phase progression (Chapter II). Prior to this work, little was known about β-globin LCR demethylation during progressive erythroid progenitor maturation in primary cells. Furthermore, the mechanism by which β-globin LCR demethylation occurred was completely unknown.

Successful characterization of DNA methylation at the β-globin LCR encouraged us to investigate DNA methylation at other loci that are important for erythropoiesis. This pursuit eventually led us to a completely unexpected discovery: Global DNA methylation is progressively and irreversibly lost during red cell differentiation. This finding is significant because in somatic cells global methylation levels were previously thought to remain constant. In normal development, only primordial germ cells and the pre-implantation embryo have been identified as undergoing waves of genome-wide

demethylation. In disease, global hypomethylation has long been known to be associated with cancer. Despite decades of research, fundamental aspects of global demethylation in these model systems are still poorly understood. Unfortunately, these are especially difficult systems in which to investigate global methylation loss, due to severe limitations in the accessibility of cells at various developmental stages and the small number of cells that can be obtained for study.

Documentation and characterization of global demethylation in erythropoiesis comprise the major components of my dissertation research (Chapter III). We used many orthogonal approaches, including genome-scale reduced representation bisulfite sequencing, to map the extent and magnitude of DNA methylation loss in four progressive stages of erythropoiesis. We also identified two important components of the mechanism responsible for erythroblast demethylation. First, that it is dependent on DNA replication at all stages of differentiation. Therefore, methylation loss is passive, resulting simply from inefficient maintenance methylation, as opposed to resulting from active enzymatic removal. Second, we discovered that loss of methylation requires the accelerated rate of intra-S-phase DNA synthesis that is coincident with erythroid lineage commitment. This requirement has not previously been identified as a contributor to methylation loss in any system, but could be relevant to the hypomethylation that is observed in cancer, where the mechanism remains largely unknown. Functionally, we show that inhibition of genome-wide demethylation results in slower demethylation at the β-globin LCR and slower erythroid gene induction, an effect that could be reversed by the addition of the demethylation agent 5-aza. This last finding strengthens existing

evidence in PGCs and the pre-implantation embryo for a functional link between global demethylation and a requirement for demethylation at specific loci for gene activation.

Our progress in elucidating the mechanism and functional significance of global demethylation in erythropoiesis is largely due to our model system, which offers significant advantages over other models of global demethylation. First, using *in vivo* tissues, we can reproducibly identify and isolate pure cell populations with discrete levels of methylation. The number and purity of these cells allows for greater experimental flexibility. Second, isolated erythroid subsets undergo comparable demethylation when placed in *ex vivo* culture. These traits allow us the ability to genetically or chemically manipulate these cells *in vivo* and *ex vivo*.

In conclusion, our work demonstrates that global hypomethylation can occur in somatic cells, overturning previous held dogma. Furthermore, we demonstrate that elevated rates of DNA replication can lead to DNA methylation loss required for erythroid-specific gene activation. Finally, our work establishes fetal liver erythropoiesis as a new experimental model that will facilitate the study of genome-wide methylation loss associated with development and disease.

Future Experiments

Understanding the mechanism of demethylation

Direct detection of methylation loss on the nascent DNA strand. Passive

demethylation is defined as DNA replication-dependent loss of methylation on the newly

synthesized DNA strand. We have demonstrated that the methylation loss we observe is

dependent on DNA replication. Additionally, partially differentiated cells in which the

DNA replication has been inhibited are able to recover methylation lost during earlier

stages of differentiation, suggesting the presence of a hemi-methylated intermediate.

However, we have yet to directly demonstrate the presence of a hemi-methylated

intermediate or lower methylation on nascent DNA. The addition of these data would

strengthen the finding of passive demethylation.

Hemi-methylation can be detected by two slightly different methods. One method

uses a hairpin linker, targeted and ligated to restriction-enzyme-cleaved genomic DNA,

to maintain attachment of complementary strands during subsequent bisulfite conversion

and PCR amplification [389]. Another method takes advantage of the fact that HpaII will

not cut the sequence CCpGG in either the fully- or hemi- methylated state [389]. Cut

DNA is bisulfite converted and amplified by PCR using primers designed to flank the

cutting site. Molecules without any methylation will not be amplified, allowing for the

calculation of fully-, hemi-, and un- methylated fractions. We would expect to find an

increase in hemi-methylated DNA in later stages of differentiation as compared to freshly

sorted S0 cells.

Nascent DNA can be isolated by incorporation of BrdU, followed by purification using an anti-BrdU antibody [390, 391]. Methylation of precipitated and bulk DNA can then be determined by bisulfite conversion, PCR, and pyrosequencing. To isolate nascent DNA in erythropoiesis, BrdU would first be injected into adult mice pregnant with E12.5 to E14.5 embryos, and shortly thereafter developmentally-staged erythroid populations would be isolated by flow cytometry. If demethylation is passive, then it is expected that within each subpopulation (S1 to S4/5), the nascent BrdU-precipitated DNA should have a lower methylation level than the bulk DNA. As a control, nascent DNA in the S0 population should have an equal amount of 5mC compared to bulk DNA.

Dnmt1 overexpression. The maintenance methylation deficiency we observed in erythropoiesis could be due to inadequate levels of Dnmt1. Therefore, successful overexpression of Dnmt1 is critical to our understanding of the mechanism of 5mC loss. We have attempted to prevent demethylation by retroviral over expression of Dnmt1. However, due to the packaging limit of this virus and the large size of Dnmt1, poor infectivity was observed. We were able to marginally improve infectivity by concentrating the viral supernatant, resulting in differentiated S3 cells expressing low levels of GFP. However, Dnmt1 mRNA was not increased in these cells compared to control, possibly the result of a high level of endogenous Dnmt1 mRNA and inadequate time to express the exogenous transcript above endogenous levels. Delivery of Dnmt1 could be improved by utilizing an alternative viral delivery system with a greater packaging size limit, such as an erythroid-specific lentiviral system described previously [392, 393].

Alternatively, we have initiated experiments in which we enrich for infected cells using culture conditions that allow for erythroblast expansion without differentiation. This system uses the glucocorticoid steroid dexamethasone to maintain freshly isolated fetal liver cells in an „S1-like" state (data not shown) in culture for up to 8 days [394]. Differentiation is induced by dexamethasone withdrawal. We have completed experiments that confirm global and β-globin LCR methylation levels are maintained in cells expanded in dexamethasone-containing media, and that demethylation is induced upon differentiation. We have also infected these S1-like cells with Dnmt1-IRES-GFP retrovirus and, through sequential rounds of sorting for GFP positive cells, obtained a population of cells expressing Dnmt1 mRNA at a level 5-fold greater than controls. Additional experiments using this system need to be performed to determine if Dnmt1 protein is also overexpressed and the effect of elevated Dnmt1 expression on global demethylation.

Dnmt1 localization, activity, and co-factor association. Total protein levels of the maintenance methyltransferase Dnmt1 were slightly increased during erythroid differentiation, consistent with increased numbers of cells in S-phase. In the pre-implantation embryo, Dnmt1 is shuttled to the cytoplasm, leading to passive demethylation during subsequent cell cycles. We have not yet investigated the subcellular localization of Dnmt1 during erythropoiesis. The subcellular localization of Dnmt1 should be determined by nuclear and cytoplasmic protein fractionation of erythroid subsets S0 to S4/5 followed by western blotting. Alternatively, anti-Dnmt1 immunofluorescence microscopy of erythroid subsets could be utilized.

Post-translational modifications, such as phosphorylation and SUMOylation, have the potential to affect Dnmt1 enzymatic activity [120, 121]. Additionally, the possibility exists that there is a direct inhibitor of Dnmt1 that is induced during erythroid differentiation. To test if diminished Dnmt1 enzymatic activity plays a role in passive methylation loss during erythropoiesis, the methyltransferase activity of nuclear lysates from S0, S1 and S3 subsets should be determined.

Dnmt1 is targeted to replication foci and hemi-methylated DNA by direct interactions with Pcna and Uhrf1, respectively. All three factors are up-regulated during S-phase and co-localize at replication foci [111, 112]. Depletion of Uhrf1 or preventing Dnmt1-Pcna binding reduces Dnmt1 association with chromatin and leads to a reduction in DNA methylation levels [109-112]. Similarly, phosphorylation of Dnmt1 leads to a disruption of Dnmt1-Uhrf1-Pcna interactions and hypomethylation in gliomagenesis [113]. Therefore, an imbalance in the ratios of these proteins [372] or a disruption in their interactions could contribute to the reduced maintenance methylation observed in erythropoiesis. To investigate the former possibility, nuclear protein levels of Pcna and Uhrf1 should be determined in each erythroid subset and then compared directly to one another and to Dnmt1. The latter possibility could be tested by monitoring phosphorylation of Dnmt1 during differentiation or by immunoprecipitation of Dnmt1 followed by western blotting with Dnmt1, Uhrf1, and Pcna.

The mechanism driving increased intra-S-phase DNA synthesis rate. The increased DNA replication rate observed in S1 to S3 cells is a contributing factor in genome-wide demethylation during erythropoiesis. This rapid DNA synthesis rate may

result from an increase in the number of origins firing simultaneously, or alternatively, may represent accelerated replication fork progression. A better understanding of the mechanism driving increased intra-S-phase DNA synthesis rate is critical to our general goals of characterizing the S0 to S1 transition. More specifically, this line of research will aid our understanding of the mechanism of global DNA methylation loss. Since our data suggest that Dnmt1 abundance is limiting, if the number of origins firing is found to increase, then it would raise the possibility that the quantitative defect in maintenance methylation results from lowered occupancy of Dnmt1 at origins. Inter-origin distance and replication fork velocity can be determined by sequential pulse labeling of replicating DNA using two nucleotide analogues, followed by dynamic molecular DNA combing and fluorescent detection of the nucleotide analogues [395-397].

SAM, SAH, and dcSAM abundance. SAM is the methyl donor utilized by DNA methyltransferases in a reaction that forms 5mC and S-adenosyl homocysteine (SAH). SAM is also decarboxylated (dcSAM) and then used as an aminopropyl group donor for the synthesis of polyamines. Both SAH [139, 140] and dcSAM [141] act as competitive inhibitors of DNA methyltransferases. Disruption in the balance of SAM relative to its by-products SAH and dcSAM is linked to hypomethylation [142-147]. Therefore, it is possible that a contributing factor to passive global demethylation in erythropoiesis is an alteration in SAM, SAH, and dcSAM ratios. Intracellular levels of these metabolites in S0, S1 and S3 subsets can be measured by high performance liquid chromatography [398].

Gata-1 and Eklf occupancy at the β-globin LCR. Demethylation during erythropoiesis has a hierarchical structure, with the highest demethylation rate found at the erythroid-specific β-globin LCR (65-80% loss) and slower demethylation in most other genomic sites (30% loss). One explanation for this dramatically reduced maintenance fidelity at the β-globin LCR could be the binding of erythroid-specific transcription factors, such as Eklf and Gata-1, leading to protection of hemi-methylated CpGs from the maintenance activity of Dnmt1. These factors are present in S0 cells and their expression increases in differentiating erythroblasts, but it is unknown at which developmental stage they bind to the β-globin LCR. If transcription factor binding is found to be correlated with genome-wide demethylation, this would suggest that the two processes are linked. Additionally, immunoprecipitation of genomic DNA using Gata-1 antibodies, in conjunction with DNA methylation analysis at the β-globin LCR, could determine if Gata-1 binding is enriched at demethylated DNA.

Understanding the functional significance of demethylation

Demethylation in nuclear condensation and enucleation. Maturing erythroblasts undergo extensive chromatin condensation leading to a 10-fold reduction in the nuclear volume [399]. Extrusion of the pycnotic nucleus occurs by an asymmetric cell division following terminal cell cycle exit. [400]. Global levels of several acetylated histones are reduced during culture of primary fetal liver erythroblasts, and ectopic expression of the histone acetyltransferase Gcn5, whose level gradually decreases during erythropoiesis, blocks condensation and enucleation [401]. Consistent with this finding, histone deacetylase 2, whose level gradually increases during erythropoiesis, is required for

chromatin condensation and enucleation [393] and chemical inhibition of histone deacetylases blocks enucleation [393, 402]. These studies establish a critical role for global histone deacetylation in chromatin condensation and enucleation in erythroid cells. Importantly, they raise the possibility that other global epigenetic alterations, such as DNA methylation loss, might also be required for condensation and enucleation during erythropoiesis.

We have begun to investigate this possibility by utilizing „ImageStream" flow cytometry (Amnis Corporation, Seattle WA). This technology combines the single cell resolution of traditional multicolor flow cytometry with a CCD camera, allowing us to quickly capture and analyze the morphological features of tens of thousands of individual cells [403]. By using CD71, Ter119, and a cell permeable nuclear stain, Draq5, we have been able to accurately quantify the nuclear volume of distinct erythroid subsets in suspension. Using this method, we found that blocking global demethylation in S1 cells using low dose aphidicolin prevents the decrease in nuclear size associated with differentiation (data not shown). However, addition of 5-aza, and subsequent methylation loss, was unable to reverse this effect. Additionally, acceleration of demethylation by using 5-aza alone did not accelerate nuclear condensation. However, in this experiment, aphidicolin might also be preventing other events that are required for condensation. Therefore, a potential role for DNA demethylation in nuclear condensation and enucleation is still unclear and warrants additional investigation.

Demethylation requirement for activation of erythroid-specific target genes. We have shown that slowing DNA replication prevents global demethylation and reduces the

extent of methylation loss at the β-globin LCR. Delayed demethylation was associated with delayed expression of the erythroid genes β-globin, Alas2, and Band3. This effect was reversed by adding 5-aza, an inhibitor of DNA methyltransferases. This result suggested that the slower induction of erythroid genes was a result of a failure to demethylate DNA, providing a functional link between global demethylation and a requirement for demethylation at a subset of specific loci. This result could be strengthened by the following additional experiments.

First, the experiment above used 5-aza, which has known cytotoxic effects. We used nanomolar concentrations of 5-aza, well below the micromolar range that causes genotoxic damage [404], and as a result we observed minimal effects on cell viability. Nonetheless, the possibility exists that 5-aza may have indirect effects that influenced our experimental results. Therefore, a similar experiment should be performed utilizing shRNA knockdown of Dnmt1, in place of 5-aza, to reverse the demethylation delay caused by slowing the cell cycle. Second, Affymetrix GeneChip mRNA profiling of the samples described above would reveal the exact set of genes affected by global methylation levels. Once identified, a detailed methylation analysis at selected loci can be performed.

New lines of research

Differentiation block resulting from Dnmt3b expression. Dnmt3a and Dnmt3b were both dramatically down-regulated at the onset of erythropoiesis. Surprisingly, ectopic expression of Dnmt3b1, but not Dnmt3a, had an inhibitory effect on erythroid differentiation. This effect was not dependent on DNA methylation activity, as it was also

observed in cells expressing the catalytically inactive mutant Dnmt3b1:PC. This inhibitory effect was reminiscent of the PU.1-mediated arrest of erythroid differentiation described in Chapter II. In this case, PU.1 was shown to simultaneously prevent the transition from S0 to S1 and S-phase progression.

A role for Dnmt3b1 in restricting proliferation has been described previously; Dnmt3b1-deficient murine embryonic fibroblasts undergo spontaneous immortalization, and reintroduction of Dnmt3b1 results in reduced proliferation [137]. Interestingly, Dnmt3b and PU.1 have been shown to directly associate with one another in hematopoietic progenitor cells [405]. Furthermore, Dnmt3b associates with the PU.1-Sin3a-Hdac1-MeCP2 repressor complex *in vitro* and is able to repress PU.1 mediated transcriptional activation [405]. Therefore, the effect of Dnmt3b on erythropoiesis could be the result of inhibited proliferation, possibly through an interaction with PU.1. Therefore, further investigation into the cell cycle characteristics of Dnmt3b-expressing erythroblasts, as well as the potential association between Dnmt3b and PU.1 in S0 cells and erythroblasts expressing exogenous Dnmt3b, is required.

5-hydroxymethylcytosine in erythroid cells. 5-hydroxymethylcytosine was first identified in bacteriophage a half a century ago and has recently been discovered in relatively high abundance in ES and brain cells [130, 406, 407]. 5hmC is formed directly from 5mC through the action of the Tet family of 5mC hydroxylases [130]. This epigenetic modification is the subject of intensive investigation because it might reflect an intermediate step in the process of 5mC demethylation. Alternatively, 5hmC could influence chromatin structure and local transcriptional activity by recruiting selective

5hmC binding proteins or excluding methyl-binding domain proteins that normally recognize 5mC [179, 408].

The bisulfite sequencing assays employed in our work cannot distinguish between 5mC and 5hmC. Therefore, we currently do not know if 5hmC is present during erythropoiesis. GeneChip analysis detects Tet2 and Tet3 transcripts in S0, S1, and S3 subsets (data not shown), suggesting that 5hmC might be found in these cells. Importantly, DNA sequences containing 5hmC are not substrates for the maintenance methyltransferase Dnmt1 [134]. If 5hmC were to be found in erythroid cells, its presence could provide new mechanistic insight into the DNA replication dependent methylation loss that we have observed. Additionally, the finding in and of itself would be novel, since 5hmC has not been reported for the erythroid lineage. There are a number of techniques that detect both 5mC and 5hmC including thin-layer chromatography, high performance liquid chromatography, mass spectrometry and single-molecule real-time sequencing [130, 406, 409].

BIBLIOGRAPHY

1. Lister, R., et al., *Human DNA methylomes at base resolution show widespread epigenomic differences.* Nature, 2009. **462**(7271): p. 315-22.

2. Malone, C.S., et al., *CmC(A/T)GG DNA methylation in mature B cell lymphoma gene silencing.* Proc Natl Acad Sci U S A, 2001. **98**(18): p. 10404-9.

3. Schubeler, D., et al., *Genomic targeting of methylated DNA: influence of methylation on transcription, replication, chromatin structure, and histone acetylation.* Mol Cell Biol, 2000. **20**(24): p. 9103-12.

4. Keshet, I., J. Lieman-Hurwitz, and H. Cedar, *DNA methylation affects the formation of active chromatin.* Cell, 1986. **44**(4): p. 535-43.

5. Boyes, J. and A. Bird, *Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein.* EMBO J, 1992. **11**(1): p. 327-33.

6. Vardimon, L., et al., *Expression of a cloned adenovirus gene is inhibited by in vitro methylation.* Proc Natl Acad Sci U S A, 1982. **79**(4): p. 1073-7.

7. Stein, R., A. Razin, and H. Cedar, *In vitro methylation of the hamster adenine phosphoribosyltransferase gene inhibits its expression in mouse L cells.* Proc Natl Acad Sci U S A, 1982. **79**(11): p. 3418-22.

8. Yoder, J.A., C.P. Walsh, and T.H. Bestor, *Cytosine methylation and the ecology of intragenomic parasites.* Trends Genet, 1997. **13**(8): p. 335-40.

9. Futscher, B.W., et al., *Role for DNA methylation in the control of cell type specific maspin expression.* Nat Genet, 2002. **31**(2): p. 175-9.

10. Sado, T., et al., *X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation.* Dev Biol, 2000. **225**(2): p. 294-303.

11. Panning, B. and R. Jaenisch, *DNA hypomethylation can activate Xist expression and silence X-linked genes.* Genes Dev, 1996. **10**(16): p. 1991-2002.

12. Bird, A., *DNA methylation patterns and epigenetic memory.* Genes Dev, 2002. **16**(1): p. 6-21.

13. Li, E., C. Beard, and R. Jaenisch, *Role for DNA methylation in genomic imprinting.* Nature, 1993. **366**(6453): p. 362-5.

14. Walsh, C.P., J.R. Chaillet, and T.H. Bestor, *Transcription of IAP endogenous retroviruses is constrained by cytosine methylation.* Nat Genet, 1998. **20**(2): p. 116-7.

15. Bourc'his, D. and T.H. Bestor, *Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L.* Nature, 2004. **431**(7004): p. 96-9.

16. Chen, T., et al., *Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b.* Mol Cell Biol, 2003. **23**(16): p. 5594-605.

17. Chuang, L.S., et al., *Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1.* Science, 1997. **277**(5334): p. 1996-2000.

18. Leonhardt, H., et al., *A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei.* Cell, 1992. **71**(5): p. 865-73.

19. Liu, Y., et al., *Multiple domains are involved in the targeting of the mouse DNA methyltransferase to the DNA replication foci.* Nucleic Acids Res, 1998. **26**(4): p. 1038-45.

20. Lyko, F., et al., *Mammalian (cytosine-5) methyltransferases cause genomic DNA methylation and lethality in Drosophila.* Nat Genet, 1999. **23**(3): p. 363-6.

21. Okano, M., S. Xie, and E. Li, *Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases.* Nat Genet, 1998. **19**(3): p. 219-20.

22. Pradhan, S., et al., *Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation.* J Biol Chem, 1999. **274**(46): p. 33002-10.

23. Gowher, H. and A. Jeltsch, *Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites.* J Mol Biol, 2001. **309**(5): p. 1201-8.

24. Yokochi, T. and K.D. Robertson, *Preferential methylation of unmethylated DNA by Mammalian de novo DNA methyltransferase Dnmt3a.* J Biol Chem, 2002. **277**(14): p. 11735-45.

25. Hsieh, C.L., *In vivo activity of murine de novo methyltransferases, Dnmt3a and Dnmt3b.* Mol Cell Biol, 1999. **19**(12): p. 8211-8.

26. Lei, H., et al., *De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells.* Development, 1996. **122**(10): p. 3195-205.

27. Okano, M., et al., *DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development.* Cell, 1999. **99**(3): p. 247-57.

28. Li, E., T.H. Bestor, and R. Jaenisch, *Targeted mutation of the DNA methyltransferase gene results in embryonic lethality.* Cell, 1992. **69**(6): p. 915-26.

29. Jackson, M., et al., *Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells.* Mol Cell Biol, 2004. **24**(20): p. 8862-71.

30. Gaudet, F., et al., *A short DNA methyltransferase isoform restores methylation in vivo.* J Biol Chem, 1998. **273**(49): p. 32725-9.

31. Jackson-Grusby, L., et al., *Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation.* Nat Genet, 2001. **27**(1): p. 31-9.

32. Broske, A.M., et al., *DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction.* Nat Genet, 2009. **41**(11): p. 1207-15.

33. Lee, P.P., et al., *A critical role for Dnmt1 and DNA methylation in T cell development, function, and survival.* Immunity, 2001. **15**(5): p. 763-74.

34. Sen, G.L., et al., *DNMT1 maintains progenitor function in self-renewing somatic tissue.* Nature, 2010. **463**(7280): p. 563-7.

35. Fan, G., et al., *DNA hypomethylation perturbs the function and survival of CNS neurons in postnatal animals.* J Neurosci, 2001. **21**(3): p. 788-97.

36. Ehrlich, M., et al., *Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells.* Nucleic Acids Res, 1982. **10**(8): p. 2709-21.

37. Saxonov, S., P. Berg, and D.L. Brutlag, *A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters.* Proc Natl Acad Sci U S A, 2006. **103**(5): p. 1412-7.

38. Eckhardt, F., et al., *DNA methylation profiling of human chromosomes 6, 20 and 22.* Nat Genet, 2006. **38**(12): p. 1378-85.

39. Brunner, A.L., et al., *Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver.* Genome Res, 2009. **19**(6): p. 1044-56.

40. Weber, M., et al., *Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome.* Nat Genet, 2007. **39**(4): p. 457-66.

41. Gardiner-Garden, M. and M. Frommer, *CpG islands in vertebrate genomes.* J Mol Biol, 1987. **196**(2): p. 261-82.

42. Meissner, A., et al., *Genome-scale DNA methylation maps of pluripotent and differentiated cells.* Nature, 2008. **454**(7205): p. 766-70.

43. Irizarry, R.A., et al., *The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores.* Nat Genet, 2009. **41**(2): p. 178-86.

44. Illingworth, R., et al., *A novel CpG island set identifies tissue-specific methylation at developmental gene loci.* PLoS Biol, 2008. **6**(1): p. e22.

45. Song, F., et al., *Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression.* Proc Natl Acad Sci U S A, 2005. **102**(9): p. 3336-41.

46. Ball, M.P., et al., *Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells.* Nat Biotechnol, 2009. **27**(4): p. 361-8.

47. Doi, A., et al., *Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts.* Nat Genet, 2009. **41**(12): p. 1350-3.

48. Bell, A.C. and G. Felsenfeld, *Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene.* Nature, 2000. **405**(6785): p. 482-5.

49. Iguchi-Ariga, S.M. and W. Schaffner, *CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation.* Genes Dev, 1989. **3**(5): p. 612-9.

50. Campanero, M.R., M.I. Armstrong, and E.K. Flemington, *CpG methylation as a mechanism for the regulation of E2F activity.* Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6481-6.

51. Birke, M., et al., *The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation.* Nucleic Acids Res, 2002. **30**(4): p. 958-65.

52. Klages, S., B. Mollers, and R. Renkawitz, *The involvement of demethylation in the myeloid-specific function of the mouse M lysozyme gene downstream enhancer.* Nucleic Acids Res, 1992. **20**(8): p. 1925-32.

53. Comb, M. and H.M. Goodman, *CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2.* Nucleic Acids Res, 1990. **18**(13): p. 3975-82.

54. Kovesdi, I., R. Reichel, and J.R. Nevins, *Role of an adenovirus E2 promoter binding factor in E1A-mediated coordinate gene control.* Proc Natl Acad Sci U S A, 1987. **84**(8): p. 2180-4.

55. Watt, F. and P.L. Molloy, *Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter.* Genes Dev, 1988. **2**(9): p. 1136-43.

56. Prendergast, G.C., D. Lawe, and E.B. Ziff, *Association of Myn, the murine homolog of max, with c-Myc stimulates methylation-sensitive DNA binding and ras cotransformation.* Cell, 1991. **65**(3): p. 395-407.

57. Hark, A.T., et al., *CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus.* Nature, 2000. **405**(6785): p. 486-9.

58. Hendrich, B. and A. Bird, *Identification and characterization of a family of mammalian methyl-CpG binding proteins.* Mol Cell Biol, 1998. **18**(11): p. 6538-47.

59. Nan, X., R.R. Meehan, and A. Bird, *Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2.* Nucleic Acids Res, 1993. **21**(21): p. 4886-92.

60. Prokhortchouk, A., et al., *The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor.* Genes Dev, 2001. **15**(13): p. 1613-8.

61. Filion, G.J., et al., *A family of human zinc finger proteins that bind methylated DNA and repress transcription.* Mol Cell Biol, 2006. **26**(1): p. 169-81.

62. Jones, P.L., et al., *Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription.* Nat Genet, 1998. **19**(2): p. 187-91.

63. Nan, X., et al., *Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex.* Nature, 1998. **393**(6683): p. 386-9.

64. Kokura, K., et al., *The Ski protein family is required for MeCP2-mediated transcriptional repression.* J Biol Chem, 2001. **276**(36): p. 34115-21.

65. Fuks, F., et al., *The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation.* J Biol Chem, 2003. **278**(6): p. 4035-40.

66. Sarraf, S.A. and I. Stancheva, *Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly.* Mol Cell, 2004. **15**(4): p. 595-605.

67. Fujita, N., et al., *Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression.* J Biol Chem, 2003. **278**(26): p. 24132-8.

68.	Boeke, J., et al., *The minimal repression domain of MBD2b overlaps with the methyl-CpG-binding domain and binds directly to Sin3A.* J Biol Chem, 2000. **275**(45): p. 34963-7.

69.	Ng, H.H., et al., *MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex.* Nat Genet, 1999. **23**(1): p. 58-61.

70.	Zhang, Y., et al., *Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation.* Genes Dev, 1999. **13**(15): p. 1924-35.

71.	Wade, P.A., et al., *Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation.* Nat Genet, 1999. **23**(1): p. 62-6.

72.	Kondo, E., et al., *The thymine DNA glycosylase MBD4 represses transcription and is associated with methylated p16(INK4a) and hMLH1 genes.* Mol Cell Biol, 2005. **25**(11): p. 4388-96.

73.	Yoon, H.G., et al., *N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso.* Mol Cell, 2003. **12**(3): p. 723-34.

74.	Clouaire, T., et al., *Recruitment of MBD1 to target genes requires sequence-specific interaction of the MBD domain with methylated DNA.* Nucleic Acids Res, 2010. **38**(14): p. 4620-34.

75.	Klose, R.J., et al., *DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG.* Mol Cell, 2005. **19**(5): p. 667-78.

76.	Ooi, S.K., et al., *DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA.* Nature, 2007. **448**(7154): p. 714-7.

77.	Zhao, Q., et al., *PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing.* Nat Struct Mol Biol, 2009. **16**(3): p. 304-11.

78.	Vire, E., et al., *The Polycomb group protein EZH2 directly controls DNA methylation.* Nature, 2006. **439**(7078): p. 871-4.

79.	Lehnertz, B., et al., *Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin.* Curr Biol, 2003. **13**(14): p. 1192-200.

80.	Li, H., et al., *The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A interact directly and localize to promoters silenced in cancer cells.* J Biol Chem, 2006. **281**(28): p. 19489-500.

81.	Epsztejn-Litman, S., et al., *De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes.* Nat Struct Mol Biol, 2008. **15**(11): p. 1176-83.

82.	Feldman, N., et al., *G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis.* Nat Cell Biol, 2006. **8**(2): p. 188-94.

83.	Cedar, H. and Y. Bergman, *Linking DNA methylation and histone modification: patterns and paradigms.* Nat Rev Genet, 2009. **10**(5): p. 295-304.

84.	Stamatoyannopoulos, G., *Control of globin gene expression during development and erythroid differentiation.* Exp Hematol, 2005. **33**(3): p. 259-71.

85.	Kingsley, P.D., et al., *"Maturational" globin switching in primary primitive erythroid cells.* Blood, 2006. **107**(4): p. 1665-72.

86. Whitelaw, E., et al., *The globin switch at the level of mRNA in the developing mouse.* Prog Clin Biol Res, 1989. **316A**: p. 323-33.
87. Grosveld, F., et al., *Position-independent, high-level expression of the human beta-globin gene in transgenic mice.* Cell, 1987. **51**(6): p. 975-85.
88. Li, Q., et al., *Locus control regions.* Blood, 2002. **100**(9): p. 3077-86.
89. Tuan, D., et al., *The "beta-like-globin" gene domain in human erythroid cells.* Proc Natl Acad Sci U S A, 1985. **82**(19): p. 6384-8.
90. Forrester, W.C., et al., *A developmentally stable chromatin structure in the human beta-globin gene cluster.* Proc Natl Acad Sci U S A, 1986. **83**(5): p. 1359-63.
91. Tolhuis, B., et al., *Looping and interaction between hypersensitive sites in the active beta-globin locus.* Mol Cell, 2002. **10**(6): p. 1453-65.
92. Litt, M.D., et al., *Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus.* Science, 2001. **293**(5539): p. 2453-5.
93. Mavilio, F., et al., *Molecular mechanisms of human hemoglobin switching: selective undermethylation and expression of globin genes in embryonic, fetal, and adult erythroblasts.* Proc Natl Acad Sci U S A, 1983. **80**(22): p. 6907-11.
94. Yin, W., et al., *Histone acetylation at the human beta-globin locus changes with developmental age.* Blood, 2007. **110**(12): p. 4101-7.
95. Hsu, M., et al., *CpG hypomethylation in a large domain encompassing the embryonic beta-like globin genes in primitive erythrocytes.* Mol Cell Biol, 2007. **27**(13): p. 5047-54.
96. Forsberg, E.C., et al., *Developmentally dynamic histone acetylation pattern of a tissue-specific chromatin domain.* Proc Natl Acad Sci U S A, 2000. **97**(26): p. 14494-9.
97. Miles, J., et al., *Intergenic transcription, cell-cycle and the developmentally regulated epigenetic profile of the human beta-globin locus.* PLoS ONE, 2007. **2**(7): p. e630.
98. Litt, M.D., et al., *Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci.* EMBO J, 2001. **20**(9): p. 2224-35.
99. Bulger, M., et al., *A complex chromatin landscape revealed by patterns of nuclease sensitivity and histone modification within the mouse beta-globin locus.* Mol Cell Biol, 2003. **23**(15): p. 5234-44.
100. Demers, C., et al., *Activator-mediated recruitment of the MLL2 methyltransferase complex to the beta-globin locus.* Mol Cell, 2007. **27**(4): p. 573-84.
101. Kim, A. and A. Dean, *Developmental stage differences in chromatin subdomains of the beta-globin locus.* Proc Natl Acad Sci U S A, 2004. **101**(18): p. 7028-33.
102. Kim, A., C.M. Kiefer, and A. Dean, *Distinctive signatures of histone methylation in transcribed coding and noncoding human beta-globin sequences.* Mol Cell Biol, 2007. **27**(4): p. 1271-9.
103. Kim, A., et al., *Nucleosome and transcription activator antagonism at human beta-globin locus control region DNase I hypersensitive sites.* Nucleic Acids Res, 2007. **35**(17): p. 5831-8.

104. Socolovsky, M., et al., *Negative Autoregulation by FAS Mediates Robust Fetal Erythropoiesis.* PLoS Biol, 2007. **5**(10): p. e252.

105. Pop, R., et al., *A key commitment step in erythropoiesis is synchronized with the cell cycle clock through mutual inhibition between PU.1 and S-phase progression.* PLoS Biol, 2010. **8**(9).

106. Ooi, S.K. and T.H. Bestor, *The colorful history of active DNA demethylation.* Cell, 2008. **133**(7): p. 1145-8.

107. Niehrs, C., *Active DNA demethylation and DNA repair.* Differentiation, 2009. **77**(1): p. 1-11.

108. Hermann, A., R. Goyal, and A. Jeltsch, *The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites.* J Biol Chem, 2004. **279**(46): p. 48350-9.

109. Schermelleh, L., et al., *Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation.* Nucleic Acids Res, 2007. **35**(13): p. 4301-12.

110. Spada, F., et al., *DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells.* J Cell Biol, 2007. **176**(5): p. 565-71.

111. Bostick, M., et al., *UHRF1 plays a role in maintaining DNA methylation in mammalian cells.* Science, 2007. **317**(5845): p. 1760-4.

112. Sharif, J., et al., *The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA.* Nature, 2007. **450**(7171): p. 908-12.

113. Hervouet, E., et al., *Disruption of Dnmt1/PCNA/UHRF1 interactions promotes tumorigenesis from human and mice glial cells.* PLoS One, 2010. **5**(6): p. e11333.

114. Du, Z., et al., *DNMT1 stability is regulated by proteins coordinating deubiquitination and acetylation-driven ubiquitination.* Sci Signal, 2010. **3**(146): p. ra80.

115. Robertson, K.D., et al., *Differential mRNA expression of the human DNA methyltransferases (DNMTs) 1, 3a and 3b during the G(0)/G(1) to S phase transition in normal and tumor cells.* Nucleic Acids Res, 2000. **28**(10): p. 2108-13.

116. Szyf, M., et al., *Cell cycle-dependent regulation of eukaryotic DNA methylase level.* J Biol Chem, 1985. **260**(15): p. 8653-6.

117. Qin, W., H. Leonhardt, and F. Spada, *Usp7 and Uhrf1 control ubiquitination and stability of the maintenance DNA methyltransferase Dnmt1.* J Cell Biochem, 2011. **112**(2): p. 439-44.

118. Esteve, P.O., et al., *Regulation of DNMT1 stability through SET7-mediated lysine methylation in mammalian cells.* Proc Natl Acad Sci U S A, 2009. **106**(13): p. 5076-81.

119. Wang, J., et al., *The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation.* Nat Genet, 2009. **41**(1): p. 125-9.

120. Goyal, R., et al., *Phosphorylation of serine-515 activates the Mammalian maintenance methyltransferase Dnmt1.* Epigenetics, 2007. **2**(3): p. 155-60.

121. Lee, B. and M.T. Muller, *SUMOylation enhances DNA methyltransferase 1 activity.* Biochem J, 2009. **421**(3): p. 449-61.

122. Carlson, L.L., A.W. Page, and T.H. Bestor, *Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting.* Genes Dev, 1992. **6**(12B): p. 2536-41.

123. Hirasawa, R., et al., *Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development.* Genes Dev, 2008. **22**(12): p. 1607-16.

124. Hodge, D.R., et al., *IL-6 enhances the nuclear translocation of DNA cytosine-5-methyltransferase 1 (DNMT1) via phosphorylation of the nuclear localization sequence by the AKT kinase.* Cancer Genomics Proteomics, 2007. **4**(6): p. 387-98.

125. Desplats, P., et al., *Alpha-synuclein sequesters Dnmt1 from the nucleus: a novel mechanism for epigenetic alterations in Lewy body diseases.* J Biol Chem, 2011. **286**(11): p. 9031-7.

126. Matsuo, K., et al., *An embryonic demethylation mechanism involving binding of transcription factors to replicating DNA.* EMBO J, 1998. **17**(5): p. 1446-53.

127. Hsieh, C.L., *Evidence that protein binding specifies sites of DNA demethylation.* Mol Cell Biol, 1999. **19**(1): p. 46-56.

128. Lin, I.G., et al., *Modulation of DNA binding protein affinity directly affects target site demethylation.* Mol Cell Biol, 2000. **20**(7): p. 2343-9.

129. Namihira, M., et al., *Committed neuronal precursors confer astrocytic potential on residual neural precursor cells.* Dev Cell, 2009. **16**(2): p. 245-55.

130. Tahiliani, M., et al., *Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1.* Science, 2009. **324**(5929): p. 930-5.

131. Ito, S., et al., *Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification.* Nature, 2010. **466**(7310): p. 1129-33.

132. Koh, K.P., et al., *Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells.* Cell Stem Cell, 2011. **8**(2): p. 200-13.

133. Iqbal, K., et al., *Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine.* Proc Natl Acad Sci U S A, 2011. **108**(9): p. 3642-7.

134. Valinluck, V. and L.C. Sowers, *Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1.* Cancer Res, 2007. **67**(3): p. 946-50.

135. Jones, P.A. and G. Liang, *Rethinking how DNA methylation patterns are maintained.* Nat Rev Genet, 2009. **10**(11): p. 805-11.

136. Liang, G., et al., *Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements.* Mol Cell Biol, 2002. **22**(2): p. 480-91.

137. Dodge, J.E., et al., *Inactivation of Dnmt3b in mouse embryonic fibroblasts results in DNA hypomethylation, chromosomal instability, and spontaneous immortalization.* J Biol Chem, 2005. **280**(18): p. 17986-91.

138. Leu, Y.W., et al., *Double RNA interference of DNMT3b and DNMT1 enhances DNA demethylation and gene reactivation.* Cancer Res, 2003. **63**(19): p. 6110-5.

139. Hoffman, D.R., et al., *S-Adenosylmethionine and S-adenosylhomocystein metabolism in isolated rat liver. Effects of L-methionine, L-homocystein, and adenosine.* J Biol Chem, 1980. **255**(22): p. 10822-7.

140. Cox, R., C. Prescott, and C.C. Irving, *The effect of S-adenosylhomocysteine on DNA methylation in isolated rat liver nuclei.* Biochim Biophys Acta, 1977. **474**(4): p. 493-9.

141. Heby, O., L. Persson, and S.S. Smith, *Polyamines, DNA methylation and cell differentiation.* Adv Exp Med Biol, 1988. **250**: p. 291-9.

142. Frostesjo, L., et al., *Interference with DNA methyltransferase activity and genome methylation during F9 teratocarcinoma stem cell differentiation induced by polyamine depletion.* J Biol Chem, 1997. **272**(7): p. 4359-66.

143. Tsuji, T., et al., *Induction of epithelial differentiation and DNA demethylation in hamster malignant oral keratinocyte by ornithine decarboxylase antizyme.* Oncogene, 2001. **20**(1): p. 24-33.

144. Yamamoto, D., et al., *Ornithine decarboxylase antizyme induces hypomethylation of genome DNA and histone H3 lysine 9 dimethylation (H3K9me2) in human oral cancer cell line.* PLoS One, 2010. **5**(9): p. e12554.

145. Yi, P., et al., *Increase in plasma homocysteine associated with parallel increases in plasma S-adenosylhomocysteine and lymphocyte DNA hypomethylation.* J Biol Chem, 2000. **275**(38): p. 29318-23.

146. Caudill, M.A., et al., *Intracellular S-adenosylhomocysteine concentrations predict global DNA hypomethylation in tissues of methyl-deficient cystathionine beta-synthase heterozygous mice.* J Nutr, 2001. **131**(11): p. 2811-8.

147. James, S.J., et al., *Elevation in S-adenosylhomocysteine and DNA hypomethylation: potential epigenetic mechanism for homocysteine-related pathology.* J Nutr, 2002. **132**(8 Suppl): p. 2361S-2366S.

148. James, S.J., et al., *Mechanisms of DNA damage, DNA hypomethylation, and tumor progression in the folate/methyl-deficient rat model of hepatocarcinogenesis.* J Nutr, 2003. **133**(11 Suppl 1): p. 3740S-3747S.

149. Wainfan, E., et al., *Rapid appearance of hypomethylated DNA in livers of rats fed cancer-promoting, methyl-deficient diets.* Cancer Res, 1989. **49**(15): p. 4094-7.

150. Shivapurkar, N. and L.A. Poirier, *Tissue levels of S-adenosylmethionine and S-adenosylhomocysteine in rats fed methyl-deficient, amino acid-defined diets for one to five weeks.* Carcinogenesis, 1983. **4**(8): p. 1051-7.

151. Wainfan, E. and L.A. Poirier, *Methyl groups in carcinogenesis: effects on DNA methylation and gene expression.* Cancer Res, 1992. **52**(7 Suppl): p. 2071s-2077s.

152. Cooney, C.A., A.A. Dave, and G.L. Wolff, *Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring.* J Nutr, 2002. **132**(8 Suppl): p. 2393S-2400S.

153. Wu, S.C. and Y. Zhang, *Active DNA demethylation: many roads lead to Rome.* Nat Rev Mol Cell Biol, 2010. **11**(9): p. 607-20.

154.   Hsieh, T.F., et al., *Genome-wide demethylation of Arabidopsis endosperm.* Science, 2009. **324**(5933): p. 1451-4.
155.   Gehring, M., K.L. Bubb, and S. Henikoff, *Extensive demethylation of repetitive elements during seed development underlies gene imprinting.* Science, 2009. **324**(5933): p. 1447-51.
156.   Kafri, T., et al., *Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line.* Genes Dev, 1992. **6**(5): p. 705-14.
157.   Monk, M., M. Boubelik, and S. Lehnert, *Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development.* Development, 1987. **99**(3): p. 371-82.
158.   Kersh, E.N., et al., *Rapid demethylation of the IFN-gamma gene occurs in memory but not naive CD8 T cells.* J Immunol, 2006. **176**(7): p. 4083-93.
159.   Bruniquel, D. and R.H. Schwartz, *Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process.* Nat Immunol, 2003. **4**(3): p. 235-40.
160.   Bhattacharya, S.K., et al., *A mammalian protein with specific demethylase activity for mCpG DNA.* Nature, 1999. **397**(6720): p. 579-83.
161.   Hendrich, B., et al., *Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development.* Genes Dev, 2001. **15**(6): p. 710-23.
162.   Schmitz, K.M., et al., *TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation.* Mol Cell, 2009. **33**(3): p. 344-53.
163.   Barreto, G., et al., *Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation.* Nature, 2007. **445**(7128): p. 671-5.
164.   Jin, S.G., C. Guo, and G.P. Pfeifer, *GADD45A does not promote DNA demethylation.* PLoS Genet, 2008. **4**(3): p. e1000013.
165.   Okada, Y., et al., *A role for the elongator complex in zygotic paternal genome demethylation.* Nature, 2010. **463**(7280): p. 554-8.
166.   Choi, Y., et al., *DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis.* Cell, 2002. **110**(1): p. 33-42.
167.   Gong, Z., et al., *ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase.* Cell, 2002. **111**(6): p. 803-14.
168.   Law, J.A. and S.E. Jacobsen, *Establishing, maintaining and modifying DNA methylation patterns in plants and animals.* Nat Rev Genet, 2010. **11**(3): p. 204-20.
169.   Zhu, B., et al., *5-Methylcytosine DNA glycosylase activity is also present in the human MBD4 (G/T mismatch glycosylase) and in a related avian sequence.* Nucleic Acids Res, 2000. **28**(21): p. 4157-65.
170.   Zhu, B., et al., *5-methylcytosine-DNA glycosylase activity is present in a cloned G/T mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex.* Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5135-9.
171.   Kim, M.S., et al., *DNA demethylation in hormone-induced transcriptional derepression.* Nature, 2009. **461**(7266): p. 1007-12.

172. Morgan, H.D., et al., *Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming.* J Biol Chem, 2004. **279**(50): p. 52353-60.

173. Rai, K., et al., *DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45.* Cell, 2008. **135**(7): p. 1201-12.

174. Bhutani, N., et al., *Reprogramming towards pluripotency requires AID-dependent DNA demethylation.* Nature, 2010. **463**(7284): p. 1042-7.

175. Popp, C., et al., *Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency.* Nature, 2010. **463**(7284): p. 1101-5.

176. Metivier, R., et al., *Cyclical DNA methylation of a transcriptionally active promoter.* Nature, 2008. **452**(7183): p. 45-50.

177. Cortazar, D., et al., *Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability.* Nature, 2011. **470**(7334): p. 419-23.

178. Cannon, S.V., A. Cummings, and G.W. Teebor, *5-Hydroxymethylcytosine DNA glycosylase activity in mammalian tissue.* Biochem Biophys Res Commun, 1988. **151**(3): p. 1173-9.

179. Guo, J.U., et al., *Hydroxylation of 5-Methylcytosine by TET1 Promotes Active DNA Demethylation in the Adult Brain.* Cell, 2011. **145**(3): p. 423-34.

180. Popp, C., et al., *Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency.* Nature. **463**(7284): p. 1101-5.

181. Howlett, S.K. and W. Reik, *Methylation levels of maternal and paternal genomes during preimplantation development.* Development, 1991. **113**(1): p. 119-27.

182. Oswald, J., et al., *Active demethylation of the paternal genome in the mouse zygote.* Curr Biol, 2000. **10**(8): p. 475-8.

183. Kim, S.H., et al., *Differential DNA methylation reprogramming of various repetitive sequences in mouse preimplantation embryos.* Biochem Biophys Res Commun, 2004. **324**(1): p. 58-63.

184. Wossidlo, M., et al., *Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes.* EMBO J, 2010. **29**(11): p. 1877-88.

185. Lane, N., et al., *Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse.* Genesis, 2003. **35**(2): p. 88-93.

186. Santos, F., et al., *Dynamic reprogramming of DNA methylation in the early mouse embryo.* Dev Biol, 2002. **241**(1): p. 172-82.

187. Clark, S.J., et al., *High sensitivity mapping of methylated cytosines.* Nucleic Acids Res, 1994. **22**(15): p. 2990-7.

188. Tremblay, K.D., K.L. Duran, and M.S. Bartolomei, *A 5' 2-kilobase-pair region of the imprinted mouse H19 gene exhibits exclusive paternal methylation throughout development.* Mol Cell Biol, 1997. **17**(8): p. 4322-9.

189. Olek, A. and J. Walter, *The pre-implantation ontogeny of the H19 methylation imprint.* Nat Genet, 1997. **17**(3): p. 275-6.

190. Nakamura, T., et al., *PGC7/Stella protects against DNA demethylation in early embryogenesis.* Nat Cell Biol, 2007. **9**(1): p. 64-71.

191. Mayer, W., et al., *Demethylation of the zygotic paternal genome.* Nature, 2000. **403**(6769): p. 501-2.

192. Cirio, M.C., et al., *Preimplantation expression of the somatic form of Dnmt1 suggests a role in the inheritance of genomic imprints.* BMC Dev Biol, 2008. **8**: p. 9.

193. Kurihara, Y., et al., *Maintenance of genomic methylation patterns during preimplantation development requires the somatic form of DNA methyltransferase 1.* Dev Biol, 2008. **313**(1): p. 335-46.

194. Beaujean, N., et al., *Non-conservation of mammalian preimplantation methylation dynamics.* Curr Biol, 2004. **14**(7): p. R266-7.

195. Dean, W., et al., *Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos.* Proc Natl Acad Sci U S A, 2001. **98**(24): p. 13734-8.

196. Kishigami, S., et al., *Epigenetic abnormalities of the mouse paternal zygotic genome associated with microinsemination of round spermatids.* Dev Biol, 2006. **289**(1): p. 195-205.

197. Hajkova, P., et al., *Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway.* Science, 2010. **329**(5987): p. 78-82.

198. Yamauchi, Y., M.A. Ward, and W.S. Ward, *Asynchronous DNA replication and origin licensing in the mouse one-cell embryo.* J Cell Biochem, 2009. **107**(2): p. 214-23.

199. Ajduk, A., Y. Yamauchi, and M.A. Ward, *Sperm chromatin remodeling after intracytoplasmic sperm injection differs from that of in vitro fertilization.* Biol Reprod, 2006. **75**(3): p. 442-51.

200. Wossidlo, M., et al., *5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming.* Nat Commun. **2**: p. 241.

201. Wossidlo, M., et al., *5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming.* Nat Commun, 2011. **2**: p. 241.

202. Cropley, J.E., et al., *Germ-line epigenetic modification of the murine A vy allele by nutritional supplementation.* Proc Natl Acad Sci U S A, 2006. **103**(46): p. 17308-12.

203. Feng, S., S.E. Jacobsen, and W. Reik, *Epigenetic reprogramming in plant and animal development.* Science, 2010. **330**(6004): p. 622-7.

204. Farthing, C.R., et al., *Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes.* PLoS Genet, 2008. **4**(6): p. e1000116.

205. Yamanaka, S. and H.M. Blau, *Nuclear reprogramming to a pluripotent state by three approaches.* Nature, 2010. **465**(7299): p. 704-12.

206. Gidekel, S. and Y. Bergman, *A unique developmental pattern of Oct-3/4 DNA methylation is controlled by a cis-demodification element.* J Biol Chem, 2002. **277**(37): p. 34521-30.

207. Hattori, N., et al., *Epigenetic regulation of Nanog gene in embryonic stem and trophoblast stem cells.* Genes Cells, 2007. **12**(3): p. 387-96.

208. Li, J.Y., et al., *Synergistic function of DNA methyltransferases Dnmt3a and Dnmt3b in the methylation of Oct4 and Nanog.* Mol Cell Biol, 2007. **27**(24): p. 8748-59.

209. Blelloch, R., et al., *Reprogramming efficiency following somatic cell nuclear transfer is influenced by the differentiation and methylation state of the donor nucleus.* Stem Cells, 2006. **24**(9): p. 2007-13.

210. Chan, E.M., et al., *Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells.* Nat Biotechnol, 2009. **27**(11): p. 1033-7.

211. Wernig, M., et al., *A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types.* Nat Biotechnol, 2008. **26**(8): p. 916-24.

212. Mikkelsen, T.S., et al., *Dissecting direct reprogramming through integrative genomic analysis.* Nature, 2008. **454**(7200): p. 49-55.

213. Mochizuki, K. and Y. Matsui, *Epigenetic profiles in primordial germ cells: global modulation and fine tuning of the epigenome for acquisition of totipotency.* Dev Growth Differ, 2010. **52**(6): p. 517-25.

214. Hajkova, P., et al., *Epigenetic reprogramming in mouse primordial germ cells.* Mech Dev, 2002. **117**(1-2): p. 15-23.

215. Lee, J., et al., *Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells.* Development, 2002. **129**(8): p. 1807-17.

216. Sato, S., et al., *Erasure of methylation imprinting of Igf2r during mouse primordial germ-cell development.* Mol Reprod Dev, 2003. **65**(1): p. 41-50.

217. Yamazaki, Y., et al., *Reprogramming of primordial germ cells begins before migration into the genital ridge, making these cells inadequate donors for reproductive cloning.* Proc Natl Acad Sci U S A, 2003. **100**(21): p. 12207-12.

218. Seki, Y., et al., *Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice.* Dev Biol, 2005. **278**(2): p. 440-58.

219. Tada, M., et al., *Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells.* EMBO J, 1997. **16**(21): p. 6510-20.

220. Yabuta, Y., et al., *Gene expression dynamics during germline specification in mice identified by quantitative single-cell gene expression profiling.* Biol Reprod, 2006. **75**(5): p. 705-16.

221. Kurimoto, K., et al., *Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice.* Genes Dev, 2008. **22**(12): p. 1617-35.

222. Seki, Y., et al., *Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice.* Development, 2007. **134**(14): p. 2627-38.

223. Gehring, M., W. Reik, and S. Henikoff, *DNA demethylation by DNA repair.* Trends Genet, 2009. **25**(2): p. 82-90.

224. Sasaki, H. and Y. Matsui, *Epigenetic events in mammalian germ-cell development: reprogramming and beyond.* Nat Rev Genet, 2008. **9**(2): p. 129-40.

225. Kato, Y., et al., *Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse.* Hum Mol Genet, 2007. **16**(19): p. 2272-80.

226. Hata, K., et al., *Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice.* Development, 2002. **129**(8): p. 1983-93.

227. Kaneda, M., et al., *Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting.* Nature, 2004. **429**(6994): p. 900-3.

228. Linher, K., et al., *An epigenetic mechanism regulates germ cell-specific expression of the porcine Deleted in Azoospermia-Like (DAZL) gene.* Differentiation, 2009. **77**(4): p. 335-49.

229. Maatouk, D.M., et al., *DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages.* Development, 2006. **133**(17): p. 3411-8.

230. Suzuki, M., et al., *A new class of tissue-specifically methylated regions involving entire CpG islands in the mouse.* Genes Cells, 2007. **12**(12): p. 1305-14.

231. Hisano, M., et al., *Methylation of CpG dinucleotides in the open reading frame of a testicular germ cell-specific intronless gene, Tact1/Actl7b, represses its expression in somatic cells.* Nucleic Acids Res, 2003. **31**(16): p. 4797-804.

232. Geyer, C.B., et al., *Ontogeny of a demethylation domain and its relationship to activation of tissue-specific transcription.* Biol Reprod, 2004. **71**(3): p. 837-44.

233. Xie, W., et al., *Regulation of ALF gene expression in somatic and male germ line tissues involves partial and site-specific patterns of methylation.* J Biol Chem, 2002. **277**(20): p. 17765-74.

234. Lapeyre, J.N. and F.F. Becker, *5-Methylcytosine content of nuclear DNA during chemical hepatocarcinogenesis and in carcinomas which result.* Biochem Biophys Res Commun, 1979. **87**(3): p. 698-705.

235. Feinberg, A.P. and B. Vogelstein, *Hypomethylation distinguishes genes of some human cancers from their normal counterparts.* Nature, 1983. **301**(5895): p. 89-92.

236. Gama-Sosa, M.A., et al., *The 5-methylcytosine content of DNA from human tumors.* Nucleic Acids Res, 1983. **11**(19): p. 6883-94.

237. Wild, L. and J.M. Flanagan, *Genome-wide hypomethylation in cancer may be a passive consequence of transformation.* Biochim Biophys Acta, 2010. **1806**(1): p. 50-7.

238. Wilson, A.S., B.E. Power, and P.L. Molloy, *DNA hypomethylation and human diseases.* Biochim Biophys Acta, 2007. **1775**(1): p. 138-62.

239. Soares, J., et al., *Global DNA hypomethylation in breast carcinoma: correlation with prognostic factors and tumor progression.* Cancer, 1999. **85**(1): p. 112-8.

240. Lin, C.H., et al., *Genome-wide hypomethylation in hepatocellular carcinogenesis.* Cancer Res, 2001. **61**(10): p. 4238-43.

241. Pufulete, M., et al., *Folate status, genomic DNA hypomethylation, and risk of colorectal adenoma and cancer: a case control study.* Gastroenterology, 2003. **124**(5): p. 1240-8.

242. Feinberg, A.P., et al., *Reduced genomic 5-methylcytosine content in human colonic neoplasia.* Cancer Res, 1988. **48**(5): p. 1159-61.

243. Bariol, C., et al., *The relationship between hypomethylation and CpG island methylation in colorectal neoplasia.* Am J Pathol, 2003. **162**(4): p. 1361-71.

244. Belancio, V.P., A.M. Roy-Engel, and P.L. Deininger, *All y'all need to know 'bout retroelements in cancer.* Semin Cancer Biol, 2010. **20**(4): p. 200-10.

245. Rauch, T.A., et al., *High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer.* Proc Natl Acad Sci U S A, 2008. **105**(1): p. 252-7.

246. Weber, M., et al., *Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells.* Nat Genet, 2005. **37**(8): p. 853-62.

247. Gaudet, F., et al., *Induction of tumors in mice by genomic hypomethylation.* Science, 2003. **300**(5618): p. 489-92.

248. Yamada, Y., et al., *Opposing effects of DNA hypomethylation on intestinal and liver carcinogenesis.* Proc Natl Acad Sci U S A, 2005. **102**(38): p. 13580-5.

249. Eden, A., et al., *Chromosomal instability and tumors promoted by DNA hypomethylation.* Science, 2003. **300**(5618): p. 455.

250. Chen, R.Z., et al., *DNA hypomethylation leads to elevated mutation rates.* Nature, 1998. **395**(6697): p. 89-93.

251. Hsieh, S.Y., et al., *High-frequency Alu-mediated genomic recombination/deletion within the caspase-activated DNase gene in human hepatoma.* Oncogene, 2005. **24**(43): p. 6584-9.

252. Strout, M.P., et al., *The partial tandem duplication of ALL1 (MLL) is consistently generated by Alu-mediated homologous recombination in acute myeloid leukemia.* Proc Natl Acad Sci U S A, 1998. **95**(5): p. 2390-5.

253. Liu, W.M. and C.W. Schmid, *Proposed roles for DNA methylation in Alu transcriptional repression and mutational inactivation.* Nucleic Acids Res, 1993. **21**(6): p. 1351-9.

254. Hata, K. and Y. Sakaki, *Identification of critical CpG sites for repression of L1 transcription by DNA methylation.* Gene, 1997. **189**(2): p. 227-34.

255. Miki, Y., et al., *Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer.* Cancer Res, 1992. **52**(3): p. 643-5.

256. Morse, B., et al., *Insertional mutagenesis of the myc locus by a LINE-1 sequence in a human breast carcinoma.* Nature, 1988. **333**(6168): p. 87-90.

257. Roman-Gomez, J., et al., *Promoter hypomethylation of the LINE-1 retrotransposable elements activates sense/antisense transcription and marks the progression of chronic myeloid leukemia.* Oncogene, 2005. **24**(48): p. 7213-23.

258. Szpakowski, S., et al., *Loss of epigenetic silencing in tumors preferentially affects primate-specific retroelements.* Gene, 2009. **448**(2): p. 151-67.

259. Ting, D.T., et al., *Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers.* Science, 2011. **331**(6017): p. 593-6.

260. Daskalos, A., et al., *Hypomethylation of retrotransposable elements correlates with genomic instability in non-small cell lung cancer.* Int J Cancer, 2009. **124**(1): p. 81-7.

261. Howard, G., et al., *Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice.* Oncogene, 2008. **27**(3): p. 404-8.

262. Gupta, A., et al., *Hypomethylation of the synuclein gamma gene CpG island promotes its aberrant expression in breast carcinoma and ovarian carcinoma.* Cancer Res, 2003. **63**(3): p. 664-73.

263. Liu, H., et al., *Loss of epigenetic control of synuclein-gamma gene as a molecular indicator of metastasis in a wide range of human cancers.* Cancer Res, 2005. **65**(17): p. 7635-43.

264. Holm, T.M., et al., *Global loss of imprinting leads to widespread tumorigenesis in adult mice.* Cancer Cell, 2005. **8**(4): p. 275-85.

265. Cui, H., et al., *Loss of imprinting in colorectal cancer linked to hypomethylation of H19 and IGF2.* Cancer Res, 2002. **62**(22): p. 6442-6.

266. Herman, J.G. and S.B. Baylin, *Gene silencing in cancer in association with promoter hypermethylation.* N Engl J Med, 2003. **349**(21): p. 2042-54.

267. Herman, J.G., et al., *Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma.* Proc Natl Acad Sci U S A, 1994. **91**(21): p. 9700-4.

268. Graff, J.R., et al., *E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas.* Cancer Res, 1995. **55**(22): p. 5195-9.

269. Dobrovic, A. and D. Simpfendorfer, *Methylation of the BRCA1 gene in sporadic breast cancer.* Cancer Res, 1997. **57**(16): p. 3347-50.

270. Esteller, M., et al., *Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors.* J Natl Cancer Inst, 2000. **92**(7): p. 564-9.

271. Catteau, A., et al., *Methylation of the BRCA1 promoter region in sporadic breast and ovarian cancer: correlation with disease characteristics.* Oncogene, 1999. **18**(11): p. 1957-65.

272. Fang, J.Y. and S.D. Xiao, *Folic acid, polymorphism of methyl-group metabolism genes, and DNA methylation in relation to GI carcinogenesis.* J Gastroenterol, 2003. **38**(9): p. 821-9.

273. Alvarez, H., et al., *Widespread Hypomethylation Occurs Early and Synergizes with Gene Amplification during Esophageal Carcinogenesis.* PLoS Genet, 2011. **7**(3): p. e1001356.

274. Fraga, M.F., et al., *A mouse skin multistage carcinogenesis model reflects the aberrant DNA methylation patterns of human tumors.* Cancer Res, 2004. **64**(16): p. 5527-34.

275. Watts, G.S., et al., *DNA methylation changes in ovarian cancer are cumulative with disease progression and identify tumor stage.* BMC Med Genomics, 2008. **1**: p. 47.

276. Cadieux, B., et al., *Genome-wide hypomethylation in human glioblastomas associated with specific copy number alteration, methylenetetrahydrofolate*

reductase allele status, and increased proliferation. Cancer Res, 2006. **66**(17): p. 8469-76.

277. Eads, C.A., et al., *CpG island hypermethylation in human colorectal tumors is not associated with DNA methyltransferase overexpression.* Cancer Res, 1999. **59**(10): p. 2302-6.

278. Mizuno, S., et al., *Expression of DNA methyltransferases DNMT1, 3A, and 3B in normal hematopoiesis and in acute and chronic myelogenous leukemia.* Blood, 2001. **97**(5): p. 1172-9.

279. Kanai, Y., et al., *DNA methyltransferase expression and DNA methylation of CPG islands and peri-centromeric satellite regions in human colorectal and stomach cancers.* Int J Cancer, 2001. **91**(2): p. 205-12.

280. Buttitta, L.A. and B.A. Edgar, *Mechanisms controlling cell cycle exit upon terminal differentiation.* Curr Opin Cell Biol, 2007. **19**(6): p. 697-704.

281. Miller, J.P., et al., *Interweaving the cell cycle machinery with cell differentiation.* Cell Cycle, 2007. **6**(23): p. 2932-8.

282. Zhu, L. and A.I. Skoultchi, *Coordinating cell proliferation and differentiation.* Curr Opin Genet Dev, 2001. **11**(1): p. 91-7.

283. Weintraub, H., *Assembly of an active chromatin structure during replication.* Nucleic Acids Res, 1979. **7**(3): p. 781-92.

284. Wolffe, A.P., *Implications of DNA replication for eukaryotic gene expression.* J Cell Sci, 1991. **99 (Pt 2)**: p. 201-6.

285. Wu, H., et al., *Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor.* Cell, 1995. **83**: p. 59-67.

286. Cantor, A.B. and S.H. Orkin, *Transcriptional regulation of erythropoiesis: an affair involving multiple partners.* Oncogene, 2002. **21**(21): p. 3368-76.

287. Fujiwara, T., et al., *Discovering Hematopoietic Mechanisms through Genome-wide Analysis of GATA Factor Chromatin Occupancy.* Mol Cell, 2009. **36**(4): p. 667-81.

288. Tripic, T., et al., *SCL and associated proteins distinguish active from repressive GATA transcription factor complexes.* Blood, 2009. **113**(10): p. 2191-201.

289. Yu, M., et al., *Insights into GATA-1-Mediated Gene Activation versus Repression via Genome-wide Chromatin Occupancy Analysis.* Mol Cell, 2009. **36**(4): p. 682-95.

290. Chou, S.T., et al., *Graded repression of PU.1/Sfpi1 gene transcription by GATA factors regulates hematopoietic cell fate.* Blood, 2009. **114**(5): p. 983-94.

291. Nerlov, C., et al., *GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription.* Blood, 2000. **95**(8): p. 2543-51.

292. Rekhtman, N., et al., *PU.1 and pRB interact and cooperate to repress GATA-1 and block erythroid differentiation.* Mol Cell Biol, 2003. **23**(21): p. 7460-74.

293. Zhang, P., et al., *PU.1 inhibits GATA-1 function and erythroid differentiation by blocking GATA-1 DNA binding.* Blood, 2000. **96**(8): p. 2641-8.

294. Moreau-Gachelin, F., et al., *Spi-1/PU.1 transgenic mice develop multistep erythroleukemias.* Mol Cell Biol, 1996. **16**(5): p. 2453-63.

295. Bulger, M., et al., *ChIPs of the beta-globin locus: unraveling gene regulation within an active domain.* Curr Opin Genet Dev, 2002. **12**(2): p. 170-7.

296. Palstra, R.J., W. de Laat, and F. Grosveld, *Beta-globin regulation and long-range interactions.* Adv Genet, 2008. **61**: p. 107-42.

297. Socolovsky, M., et al., *Ineffective erythropoiesis in Stat5a(-/-)5b(-/-) mice due to decreased survival of early erythroblasts.* Blood, 2001. **98**(12): p. 3261-73.

298. Zhang, J., et al., *Role of Ras signaling in erythroid differentiation of mouse fetal liver cells: functional analysis by a flow cytometry-based novel culture system.* Blood, 2003. **102**(12): p. 3938-46.

299. Stephenson, J.R., et al., *Induction of colonies of hemoglobin-synthesizing cells by erythropoietin in vitro.* Proc. Natl. Acad. Sci. USA, 1971. **68**(7): p. 1542-1546.

300. Lee, M.H., I. Reynisdottir, and J. Massague, *Cloning of p57KIP2, a cyclin-dependent kinase inhibitor with unique domain structure and tissue distribution.* Genes Dev, 1995. **9**(6): p. 639-49.

301. Matsuoka, S., et al., *p57KIP2, a structurally distinct member of the p21CIP1 Cdk inhibitor family, is a candidate tumor suppressor gene.* Genes Dev, 1995. **9**(6): p. 650-62.

302. Kina, T., et al., *The monoclonal antibody TER-119 recognizes a molecule associated with glycophorin A and specifically marks the late stages of murine erythroid lineage.* Br. J. Haematol., 2000. **109**(2): p. 280-287.

303. Landschulz, K.T., et al., *Onset of erythropoietin response in murine erythroid colony-forming units: assignment to early S-phase in a specific cell generation.* Blood, 1992. **79**(10): p. 2749-58.

304. Ikegami, S., et al., *Aphidicolin prevents mitotic cell division by interfering with the activity of DNA polymerase-alpha.* Nature, 1978. **275**(5679): p. 458-60.

305. Lalande, M., *A reversible arrest point in the late G1 phase of the mammalian cell cycle.* Exp Cell Res, 1990. **186**(2): p. 332-9.

306. Shivdasani, R.A. and S.H. Orkin, *Erythropoiesis and globin gene expression in mice lacking the transcription factor NF-E2.* Proc Natl Acad Sci U S A, 1995. **92**(19): p. 8690-4.

307. Grass, J.A., et al., *GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling.* Proc Natl Acad Sci U S A, 2003. **100**(15): p. 8811-6.

308. Surinya, K.H., T.C. Cox, and B.K. May, *Transcriptional regulation of the human erythroid 5-aminolevulinate synthase gene. Identification of promoter elements and role of regulatory proteins.* J Biol Chem, 1997. **272**(42): p. 26585-94.

309. Kamura, T., et al., *Degradation of p57Kip2 mediated by SCFSkp2-dependent ubiquitylation.* Proc Natl Acad Sci U S A, 2003. **100**(18): p. 10231-6.

310. Aladjem, M.I., et al., *Replication initiation patterns in the beta-globin loci of totipotent and differentiated murine cells: evidence for multiple initiation regions.* Mol Cell Biol, 2002. **22**(2): p. 442-52.

311. Simon, I., et al., *Developmental regulation of DNA replication timing at the human beta globin locus.* Embo J, 2001. **20**(21): p. 6150-7.

312. McArthur, M., S. Gerum, and G. Stamatoyannopoulos, *Quantification of DNaseI-sensitivity by real-time PCR: quantitative analysis of DNaseI-hypersensitivity of the mouse beta-globin LCR.* J Mol Biol, 2001. **313**(1): p. 27-34.

313. Zhang, P., et al., *Altered cell differentiation and proliferation in mice lacking p57KIP2 indicates a role in Beckwith-Wiedemann syndrome.* Nature, 1997. **387**(6629): p. 151-8.

314. Back, J., et al., *PU.1 determines the self-renewal capacity of erythroid progenitor cells.* Blood, 2004. **103**(10): p. 3615-23.

315. Anderson, M.K., et al., *Constitutive expression of PU.1 in fetal hematopoietic progenitors blocks T cell development at the pro-T cell stage.* Immunity, 2002. **16**(2): p. 285-96.

316. Hromas, R., et al., *Hematopoietic lineage- and stage-restricted expression of the ETS oncogene family member PU.1.* Blood, 1993. **82**(10): p. 2998-3004.

317. Ambros, V., *Cell cycle-dependent sequencing of cell fate decisions in Caenorhabditis elegans vulva precursor cells.* Development, 1999. **126**(9): p. 1947-56.

318. Aparicio, O.M. and D.E. Gottschling, *Overcoming telomeric silencing: a trans-activator competes to establish gene expression in a cell cycle-dependent way.* Genes Dev, 1994. **8**(10): p. 1133-46.

319. Miller, A.M. and K.A. Nasmyth, *Role of DNA replication in the repression of silent mating type loci in yeast.* Nature, 1984. **312**(5991): p. 247-51.

320. Edgar, L.G. and J.D. McGhee, *DNA synthesis and the control of embryonic gene expression in C. elegans.* Cell, 1988. **53**(4): p. 589-99.

321. Fisher, D. and M. Mechali, *Vertebrate HoxB gene expression requires DNA replication.* Embo J, 2003. **22**(14): p. 3737-48.

322. Forlani, S., et al., *Relief of a repressed gene expression state in the mouse 1-cell embryo requires DNA replication.* Development, 1998. **125**(16): p. 3153-66.

323. Weigmann, K. and C.F. Lehner, *Cell fate specification by even-skipped expression in the Drosophila nervous system is coupled to cell cycle progression.* Development, 1995. **121**(11): p. 3713-21.

324. Chiu, C.P. and H.M. Blau, *Reprogramming cell differentiation in the absence of DNA synthesis.* Cell, 1984. **37**(3): p. 879-87.

325. Edgar, B.A. and P.H. O'Farrell, *The three postblastoderm cell cycles of Drosophila embryogenesis are regulated in G2 by string.* Cell, 1990. **62**(3): p. 469-80.

326. Harris, W.A. and V. Hartenstein, *Neuronal determination without cell division in Xenopus embryos.* Neuron, 1991. **6**(4): p. 499-515.

327. Hartenstein, V. and J.W. Posakony, *Sensillum development in the absence of cell division: the sensillum phenotype of the Drosophila mutant string.* Dev Biol, 1990. **138**(1): p. 147-58.

328. de Nooij, J.C. and I.K. Hariharan, *Uncoupling cell fate determination from patterned cell division in the Drosophila eye.* Science, 1995. **270**(5238): p. 983-5.

329. Henikoff, S., *Nucleosome destabilization in the epigenetic regulation of gene expression.* Nat Rev Genet, 2008. **9**(1): p. 15-26.
330. Reik, W., W. Dean, and J. Walter, *Epigenetic reprogramming in mammalian development.* Science, 2001. **293**(5532): p. 1089-93.
331. Koury, M.J. and M.C. Bondurant, *Erythropoietin retards DNA breakdown and prevents programmed death in erythroid progenitor cells.* Science, 1990. **248**: p. 378-381.
332. Dolznig, H., et al., *Apoptosis protection by the Epo target Bcl-X(L) allows factor-independent differentiation of primary erythroblasts.* Curr Biol, 2002. **12**(13): p. 1076-85.
333. Kerenyi, M.A., et al., *Stat5 regulates cellular iron uptake of erythroid cells via IRP-2 and TfR-1.* Blood, 2008. **112**(9): p. 3878-88.
334. Zhu, B.M., et al., *Hematopoietic-specific Stat5-null mice display microcytic hypochromic anemia associated with reduced transferrin receptor gene expression.* Blood, 2008. **112**(5): p. 2071-80.
335. Boyes, J. and G. Felsenfeld, *Tissue-specific factors additively increase the probability of the all-or-none formation of a hypersensitive site.* Embo J, 1996. **15**(10): p. 2496-507.
336. Bottardi, S., et al., *Developmental stage-specific epigenetic control of human beta-globin gene expression is potentiated in hematopoietic progenitor cells prior to their transcriptional activation.* Blood, 2003. **102**(12): p. 3989-97.
337. Hu, M., et al., *Multilineage gene expression precedes commitment in the hemopoietic system.* Genes Dev, 1997. **11**(6): p. 774-85.
338. Jimenez, G., et al., *Activation of the beta-globin locus control region precedes commitment to the erythroid lineage.* Proc Natl Acad Sci U S A, 1992. **89**(22): p. 10618-22.
339. Bernstein, B.E., et al., *A bivalent chromatin structure marks key developmental genes in embryonic stem cells.* Cell, 2006. **125**(2): p. 315-26.
340. Sharov, A.A. and M.S. Ko, *Human ES cell profiling broadens the reach of bivalent domains.* Cell Stem Cell, 2007. **1**(3): p. 237-8.
341. Liu, Y., et al., *Suppression of Fas-FasL coexpression by erythropoietin mediates erythroblast expansion during the erythropoietic stress response in vivo.* Blood, 2006. **108**(1): p. 123-33.
342. Socolovsky, M., et al., *Ineffective erythropoiesis in Stat5a(-/-)5b(-/-) mice due to decreased survival of early erythroblasts.* Blood, 2001. **98**(12): p. 3261-73.
343. Ushmorov, A., et al., *Epigenetic processes play a major role in B-cell-specific gene silencing in classical Hodgkin lymphoma.* Blood, 2006. **107**(6): p. 2493-500.
344. Ivascu, C., et al., *DNA methylation profiling of transcription factor genes in normal lymphocyte development and lymphomas.* Int J Biochem Cell Biol, 2007. **39**(7-8): p. 1523-38.
345. Gazin, C., et al., *An elaborate pathway required for Ras-mediated epigenetic silencing.* Nature, 2007. **449**(7165): p. 1073-7.

346. Lopes, S., et al., *Epigenetic modifications in an imprinting cluster are controlled by a hierarchy of DMRs suggesting long-range chromatin interactions.* Hum Mol Genet, 2003. **12**(3): p. 295-305.

347. Kuroiwa, Y., et al., *Peg3 imprinted gene on proximal chromosome 7 encodes for a zinc finger protein.* Nat Genet, 1996. **12**(2): p. 186-90.

348. Li, L.L., et al., *Organization and parent-of-origin-specific methylation of imprinted Peg3 gene on mouse proximal chromosome 7.* Genomics, 2000. **63**(3): p. 333-40.

349. Smilinich, N.J., et al., *A maternally methylated CpG island in KvLQT1 is associated with an antisense paternal transcript and loss of imprinting in Beckwith-Wiedemann syndrome.* Proc Natl Acad Sci U S A, 1999. **96**(14): p. 8064-9.

350. Loeb, D.D., et al., *The sequence of a large L1Md element reveals a tandemly repeated 5' end and several features found in retrotransposons.* Mol Cell Biol, 1986. **6**(1): p. 168-82.

351. Silver, L.M., *Repetitive "Non-Functional" DNA Families*, in *Mouse Genetics: Concepts and Applications*. 1995, Oxford University Press. p. 105-111.

352. Chalitchagorn, K., et al., *Distinctive pattern of LINE-1 methylation level in normal tissues and the association with carcinogenesis.* Oncogene, 2004. **23**(54): p. 8841-6.

353. Estecio, M.R., et al., *LINE-1 hypomethylation in cancer is highly variable and inversely correlated with microsatellite instability.* PLoS One, 2007. **2**(5): p. e399.

354. Yang, A.S., et al., *A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements.* Nucleic Acids Res, 2004. **32**(3): p. e38.

355. Karimi, M., et al., *LUMA (LUminometric Methylation Assay)--a high throughput method to the analysis of genomic DNA methylation.* Exp Cell Res, 2006. **312**(11): p. 1989-95.

356. Cedar, H., et al., *Direct detection of methylated cytosine in DNA by use of the restriction enzyme MspI.* Nucleic Acids Res, 1979. **6**(6): p. 2125-32.

357. Bestor, T.H., S.B. Hellewell, and V.M. Ingram, *Differentiation of two mouse cell lines is associated with hypomethylation of their genomes.* Mol Cell Biol, 1984. **4**(9): p. 1800-6.

358. Gu, H., et al., *Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling.* Nat Protoc, 2011. **6**(4): p. 468-81.

359. Bock, C., et al., *CpG island mapping by epigenome prediction.* PLoS Comput Biol, 2007. **3**(6): p. e110.

360. Avvakumov, G.V., et al., *Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1.* Nature, 2008. **455**(7214): p. 822-5.

361. Dolbeare, F., et al., *Flow cytometric measurement of total DNA content and incorporated bromodeoxyuridine.* Proc Natl Acad Sci U S A, 1983. **80**(18): p. 5573-7.

362.	Gratzner, H.G., *Monoclonal antibody to 5-bromo- and 5-iododeoxyuridine: A new reagent for detection of DNA replication.* Science, 1982. **218**(4571): p. 474-5.

363.	Muotri, A.R., et al., *L1 retrotransposition in neurons is modulated by MeCP2.* Nature, 2010. **468**(7322): p. 443-6.

364.	Yu, F., et al., *Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription.* Nucleic Acids Res, 2001. **29**(21): p. 4493-501.

365.	Bishop, D.F., A.S. Henderson, and K.H. Astrin, *Human delta-aminolevulinate synthase: assignment of the housekeeping gene to 3p21 and the erythroid-specific gene to the X chromosome.* Genomics, 1990. **7**(2): p. 207-14.

366.	Bennett, V. and P.J. Stenbuck, *The membrane attachment protein for spectrin is associated with band 3 in human erythrocyte membranes.* Nature, 1979. **280**(5722): p. 468-73.

367.	Kopito, R.R. and H.F. Lodish, *Primary structure and transmembrane orientation of the murine anion exchange protein.* Nature, 1985. **316**(6025): p. 234-8.

368.	Christman, J.K., *5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy.* Oncogene, 2002. **21**(35): p. 5483-95.

369.	Razin, A., et al., *Variations in DNA methylation during mouse cell differentiation in vivo and in vitro.* Proc Natl Acad Sci U S A, 1984. **81**(8): p. 2275-9.

370.	Easwaran, H.P., et al., *Replication-independent chromatin loading of Dnmt1 during G2 and M phases.* EMBO Rep, 2004. **5**(12): p. 1181-6.

371.	Chen, Z.X. and A.D. Riggs, *Maintenance and regulation of DNA methylation patterns in mammals.* Biochem Cell Biol, 2005. **83**(4): p. 438-48.

372.	Kimura, F., et al., *Decrease of DNA methyltransferase 1 expression relative to cell proliferation in transitional cell carcinoma.* Int J Cancer, 2003. **104**(5): p. 568-78.

373.	Gloria, L., et al., *DNA hypomethylation and proliferative activity are increased in the rectal mucosa of patients with long-standing ulcerative colitis.* Cancer, 1996. **78**(11): p. 2300-6.

374.	McLean, C.Y., et al., *GREAT improves functional interpretation of cis-regulatory regions.* Nat Biotechnol, 2010. **28**(5): p. 495-501.

375.	Hamm, C.A., et al., *Global demethylation of rat chondrosarcoma cells after treatment with 5-aza-2'-deoxycytidine results in increased tumorigenicity.* PLoS One, 2009. **4**(12): p. e8340.

376.	Michalowsky, L.A. and P.A. Jones, *Gene structure and transcription in mouse cells with extensively demethylated DNA.* Mol Cell Biol, 1989. **9**(3): p. 885-92.

377.	Hatada, I., et al., *Genome-wide demethylation during neural differentiation of P19 embryonal carcinoma cells.* J Hum Genet, 2008. **53**(2): p. 185-91.

378.	Beutler, E. and J. Waalen, *The definition of anemia: what is the lower limit of normal of the blood hemoglobin concentration?* Blood, 2006. **107**(5): p. 1747-50.

379.	Pevny, L., et al., *Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1.* Nature, 1991. **349**(6306): p. 257-60.

380. Perkins, A.C., A.H. Sharpe, and S.H. Orkin, *Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF.* Nature, 1995. **375**(6529): p. 318-22.

381. Nuez, B., et al., *Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene.* Nature, 1995. **375**(6529): p. 316-8.

382. Johnson, K.D., et al., *Cooperative activities of hematopoietic regulators recruit RNA polymerase II to a tissue-specific chromatin domain.* Proc Natl Acad Sci U S A, 2002. **99**(18): p. 11760-5.

383. Kim, K. and A. Kim, *Sequential changes in chromatin structure during transcriptional activation in the beta globin LCR and its target gene.* Int J Biochem Cell Biol, 2010. **42**(9): p. 1517-24.

384. Shyu, Y.C., et al., *Chromatin-binding in vivo of the erythroid kruppel-like factor, EKLF, in the murine globin loci.* Cell Res, 2006. **16**(4): p. 347-55.

385. Rohde, C., et al., *BISMA--fast and accurate bisulfite sequencing data analysis of individual clones from unique and repetitive sequences.* BMC Bioinformatics, 2010. **11**: p. 230.

386. Simon, R., et al., *Analysis of gene expression data using BRB-ArrayTools.* Cancer Inform, 2007. **3**: p. 11-7.

387. Socolovsky, M., et al., *Fetal anemia and apoptosis of red cell progenitors in Stat5a$^{-/-}$5b$^{-/-}$ mice: a direct role for Stat5 in bcl-X$_L$ induction.* Cell, 1999. **98**: p. 181-191.

388. Naviaux, R.K., et al., *The pCL vector system: rapid production of helper-free, high-titer, recombinant retroviruses.* J Virol, 1996. **70**(8): p. 5701-5.

389. Laird, C.D., et al., *Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules.* Proc Natl Acad Sci U S A, 2004. **101**(1): p. 204-9.

390. Contreas, G., M. Giacca, and A. Falaschi, *Purification of BrdUrd-substituted DNA by immunoaffinity chromatography with anti-BrdUrd antibodies.* Biotechniques, 1992. **12**(6): p. 824-6.

391. Vassilev, L. and G. Russev, *Purification of nascent DNA chains by immunoprecipitation with anti-BrdU antibodies.* Nucleic Acids Res, 1988. **16**(21): p. 10397.

392. Moreau-Gaudry, F., et al., *High-level erythroid-specific gene expression in primary human and murine hematopoietic cells with self-inactivating lentiviral vectors.* Blood, 2001. **98**(9): p. 2664-72.

393. Ji, P., et al., *Histone deacetylase 2 is required for chromatin condensation and subsequent enucleation of cultured mouse fetal erythroblasts.* Haematologica, 2010. **95**(12): p. 2013-21.

394. Dolznig, H., et al., *Expansion and Differentiation of Immature Mouse and Human Hematopoietic Progenitors.* 2004. p. 323-343.

395. Conti, C., et al., *Replication fork velocities at adjacent replication origins are coordinately modified during DNA replication in human cells.* Mol Biol Cell, 2007. **18**(8): p. 3059-67.

396. Michalet, X., et al., *Dynamic molecular combing: stretching the whole human genome for high-resolution studies.* Science, 1997. **277**(5331): p. 1518-23.

397. Rimmele, P., et al., *Spi-1/PU.1 oncogene accelerates DNA replication fork elongation and promotes genetic instability in the absence of DNA breakage.* Cancer Res, 2010. **70**(17): p. 6757-66.

398. Della Ragione, F., et al., *[A new method for determination of adenosylmethionine, adenosyl-homocysteine and decarboxylated adenosylmethionine].* Boll Soc Ital Biol Sper, 1980. **56**(3): p. 250-6.

399. Ji, P., M. Murata-Hori, and H.F. Lodish, *Formation of mammalian erythrocytes: chromatin condensation and enucleation.* Trends Cell Biol, 2011.

400. Ji, P., S.R. Jayapal, and H.F. Lodish, *Enucleation of cultured mouse fetal erythroblasts requires Rac GTPases and mDia2.* Nat Cell Biol, 2008. **10**(3): p. 314-21.

401. Jayapal, S.R., et al., *Down-regulation of Myc is essential for terminal erythroid maturation.* J Biol Chem, 2010. **285**(51): p. 40252-65.

402. Popova, E.Y., et al., *Chromatin condensation in terminally differentiating mouse erythroblasts does not involve special architectural proteins but depends on histone deacetylation.* Chromosome Res, 2009. **17**(1): p. 47-64.

403. McGrath, K.E., T.P. Bushnell, and J. Palis, *Multispectral imaging of hematopoietic cells: where flow meets morphology.* J Immunol Methods, 2008. **336**(2): p. 91-7.

404. Palii, S.S., et al., *DNA methylation inhibitor 5-Aza-2'-deoxycytidine induces reversible genome-wide DNA damage that is distinctly influenced by DNA methyltransferases 1 and 3B.* Mol Cell Biol, 2008. **28**(2): p. 752-71.

405. Suzuki, M., et al., *Site-specific DNA methylation by a complex of PU.1 and Dnmt3a/b.* Oncogene, 2006. **25**(17): p. 2477-88.

406. Kriaucionis, S. and N. Heintz, *The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain.* Science, 2009. **324**(5929): p. 929-30.

407. Wyatt, G.R. and S.S. Cohen, *The bases of the nucleic acids of some bacterial and animal viruses: the occurrence of 5-hydroxymethylcytosine.* Biochem J, 1953. **55**(5): p. 774-82.

408. Valinluck, V., et al., *Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2).* Nucleic Acids Res, 2004. **32**(14): p. 4100-8.

409. Flusberg, B.A., et al., *Direct detection of DNA methylation during single-molecule, real-time sequencing.* Nat Methods, 2010. **7**(6): p. 461-5.