DNA Methylation Changes at Promoters of Endothelial Cell-Enriched Genes during *in vitro* Differentiation

Anna Kop

Master of Science

Department of Laboratory Medicine and Pathobiology
University of Toronto

2011

Abstract

This study examined DNA methylation patterns at promoters of endothelial cell (EC)-enriched genes during differentiation of mouse ES cells towards the EC. We have previously shown that eNOS, CD31, VE-cadherin and vWF, which have an EC-enriched pattern of gene expression are differentially methylated between EC and vascular smooth muscle cells. Given that differential promoter DNA methylation is functionally important we asked when these distinct patterns are established. Using the hanging drop method to differentiate ES cells, followed by FACS, we isolated early (EB-day4 VEGFR2-positive) and late (EB-day7 CD31-positive) endothelial progenitor cells. Though current paradigms suggest that lineage-restricted genes are methylated in ES cells, we show heterogeneous promoter DNA methylation. We show DNA demethylation at the CD31 promoter in EB-day 7 CD31-positive cells. In contrast, the eNOS promoter is still heavily methylated in EB-day 7 CD31 positive cells compared with murine EC where there is no DNA methylation.
Acknowledgments

I would like to thank my supervisor, Dr. Philip Marsden for the opportunity to work and grow in his lab. During the course of my graduate studies, I learned many skills that I will take with me into the working world.

Thank you to my supervisory committee, Dr. Michelle Bendeck, Dr. Michael Sefton and Dr. William Stanford for their input and advice, and for making me think on my feet.

I’ve had the pleasure of working with some truly great people in the lab. Thank you for your humour, reassurance and advice.

An exponential number of thank yous are due to my family and friends, whose encouraging words have made this possible.
Accreditation of work

The author completed the design and the experiments for this project.

Gene expression data and the list of differentially methylated, EC-enriched genes in human cells were kindly provided by Apurva Shirodkar and other Marsden lab members that contributed to the manuscript, “Cell-restricted patterns of DNA methylation in endothelial cell-enriched genes”. Gene expression data of Oct4 and EC-enriched genes during murine in vitro differentiation was kindly provided by Dr. Jason Fish. The LUMA data in HUVEC and VSMC were kindly provided by Apurva Shirodkar. The sodium bisulfite genomic sequencing data at the eNOS promoter in hES cells was kindly provided by Alisha Jamal. The sodium bisulfite genomic sequencing data at the eNOS promoter in human syncytiotrophoblasts was kindly provided by Megan Thompson.

In addition to this thesis, the author contributed to the design and experiments for the manuscript titled, “Transcriptional regulation of eNOS in the in vivo setting: role of a shear-response element” (in preparation).
### Table of Contents

Acknowledgments ........................................................................................................ iii

Accreditation of work .................................................................................................... iv

Table of Contents ........................................................................................................ v

List of Tables ................................................................................................................ vii

List of Figures ............................................................................................................... viii

List of Abbreviations and Acronyms .......................................................................... x

Chapter 1 Literature Review ......................................................................................... 1

1.1 Epigenetics .............................................................................................................. 2

1.1.1 Introduction ....................................................................................................... 2

1.1.2 DNA methylation ............................................................................................ 2

1.1.3 DNA methyltransferase-deficient mice and ES cells .................................... 4

1.1.4 DNA hydroxymethylation .............................................................................. 8

1.1.5 DNA demethylation ....................................................................................... 8

1.1.6 Epigenetic changes during differentiation of mouse ES cells ...................... 9

1.1.7 CpG islands .................................................................................................... 12

1.1.8 Histone modifications .................................................................................... 13

1.1.9 Long non-coding RNA ................................................................................. 13

1.2 Vascular Endothelium .......................................................................................... 17

1.2.1 Introduction .................................................................................................... 17

1.2.2 Endothelial cell-enriched genes ................................................................... 17

1.2.3 Transcriptional regulation of the endothelial cell-enriched genes by shear stress ................................................................. 21

1.2.4 In vitro differentiation of ES cells into endothelial cells ............................ 22

1.3 Methods to quantify DNA methylation ............................................................... 28

1.3.1 Sodium bisulfite genomic sequencing .......................................................... 28

1.3.2 Pyrosequencing ............................................................................................ 29

1.3.3 LINE-1 assay ............................................................................................... 33

1.3.4 LUMA assay ............................................................................................... 34

1.4 Objective ............................................................................................................... 34
Chapter 2 DNA methylation changes at promoters of endothelial cell-enriched genes during \textit{in vitro} differentiation ................................................................. 35

2.1 Rationale and Hypothesis ........................................................................... 36
2.2 Methods ........................................................................................................ 39
  2.2.1 Cell culture .............................................................................................. 39
  2.2.2 mES cell differentiation .......................................................................... 39
  2.2.3 Cell sorting ............................................................................................... 40
  2.2.4 DNA extraction ........................................................................................ 44
  2.2.5 Sodium bisulfite conversion of DNA ....................................................... 44
  2.2.6 Sodium bisulfite genomic sequencing ..................................................... 44
  2.2.7 PCR product pyrosequencing ................................................................ 45
  2.2.8 LINE-1 .................................................................................................... 48
  2.2.9 LUMA ...................................................................................................... 48
  2.2.10 Statistics ................................................................................................ 48

2.3 Results ............................................................................................................ 56
  2.3.1 LINE-1 hypomethylation in mES cells and hypermethylation in all other
        mouse cell types ....................................................................................... 56
  2.3.2 LINE-1 and CmCGG methylation does not change upon human ES cell
        differentiation .......................................................................................... 56
  2.3.3 Increase in CG methylation at the Nanog promoter upon differentiation ....57
  2.3.4 The VEGFR2 promoter is unmethylated in all cell types ......................... 57
  2.3.5 The promoters of eNOS, CD31, vWF and VE-cadherin are differentially
        methylated between mouse EC and VSMC ............................................ 57
  2.3.6 The promoters of EC-enriched genes are methylated in mouse ES cells ....57
  2.3.7 The promoters of eNOS, CD31 and VE-cadherin are heterogeneously
        methylated in ES cells ............................................................................. 57
  2.3.8 The eNOS promoter is heterogeneously methylated in extraembryonic cells..58
  2.3.9 Changes in promoter DNA methylation of eNOS, CD31, VE-cadherin and
        vWF during \textit{in vitro} differentiation ......................................................... 58

2.4 Discussion ....................................................................................................... 73
2.5 Limitations ..................................................................................................... 78
2.6 Conclusions .................................................................................................... 78
2.7 Future Directions ........................................................................................... 79

References ........................................................................................................... 82
List of Tables

Table 1: CpG island classification of murine promoters of endothelial cell-enriched genes        14
Table 2: A summary of histone posttranslational modifications and their effects on gene expression                                                                                             15
Table 3: Differential DNA methylation at promoters of human EC-enriched genes in human EC and VSMC                                                                                           16
Table 4: Transcriptional regulation of the endothelial cell-enriched genes by shear stress       27
Table 5: Murine sodium bisulfite sequencing primers                                                                                                                                     46
Table 6: Murine PCR pyrosequencing primers                                                                                                                                             47
# List of Figures

## Introduction
- Figure 1. The enzymatic conversion between cytosine, 5-methylcytosine and 5-hydroxymethylcytosine ................................................................. 6
- Figure 2. Mechanisms of transcriptional repression by DNA methylation ................................................................. 7
- Figure 3. Passive and active mechanisms of DNA demethylation ........................................................................... 11
- Figure 4. Expression of endothelial cell-enriched genes in five human primary cell types .... 18
- Figure 5. Increase in eNOS promoter DNA methylation in EC from the high probability region of the murine aortic arch ................................................................. 25
- Figure 6. Expression of Oct4 and EC-enriched genes during *in vitro* differentiation. ........................ 26
- Figure 7. The conversion of cytosine but not 5mC or 5hmC to uracil with sodium bisulfite treatment ........................................................................... 31
- Figure 8. Sodium bisulfite genomic sequencing. ................................................................................................. 32

## Methods
- Figure 9. The *in vitro* differentiation of murine ES cells into “early stage” and “late stage” endothelial progenitor cells. ................................................................. 38
- Figure 10. Mouse ES cell and EB morphology ................................................................................................. 41
- Figure 11. Brief summary of cell culture and methods .................................................................................. 42
- Figure 12. Cell sorting for “early” and “late” stages in endothelial differentiation ......................................... 43
- Figure 13: Multiple sequence alignment between human and mouse genomic DNA at the eNOS promoter ........................................................................... 49
- Figure 14: Multiple sequence alignment between human and mouse genomic DNA at the shear stress response element (SSRE) in the eNOS promoter ......................................... 50
- Figure 15: Multiple sequence alignment between human and mouse genomic DNA at the CD31 exon1 ........................................................................... 51
- Figure 16: Multiple sequence alignment between human and mouse genomic DNA at the VE-cadherin promoter ........................................................................... 52
- Figure 17: Multiple sequence alignment between human and mouse genomic DNA at the vWF exon1 ........................................................................... 53
- Figure 18: Multiple sequence alignment between human and mouse genomic DNA at the VEGFR2 promoter ........................................................................... 54
- Figure 19: Multiple sequence alignment between human and mouse genomic DNA at the Nanog promoter ........................................................................... 55
Results

Figure 20. LINE-1 is hypomethylated in mouse ES cells and hypermethylated in all other cell types ................................................................. 59

Figure 21. LINE-1 hypermethylation in mouse extraembryonic cells .......................... 60

Figure 22. LINE-1 and CmCGG methylation does not change upon human ES cell differentiation ............................................................... 61

Figure 23. Changes in promoter DNA methylation of Nanog and VEGFR2 during in vitro differentiation, and in terminally differentiated cells ................................................................. 62

Figure 24. Differential promoter DNA methylation of EC-enriched genes in mouse endothelial and vascular smooth muscle cells ........................................................................ 63

Figure 25. Unmethylated CG sites at the eNOS and VE-cadherin promoters in murine thoracic aortic EC ................................................................. 64

Figure 26. Intermediate to high levels of promoter DNA methylation in murine ES cells .......................... 65

Figure 27. The eNOS promoter is heterogeneously methylated in mouse and human ES cells 66

Figure 28. The EC-enriched genes CD31 and VE-cadherin are heterogeneously methylated in mouse ES cells ........................................................................ 67

Figure 29. The eNOS promoter is heterogeneously methylated in mouse and human extraembryonic cells ........................................................................ 68

Figure 30. Dense promoter DNA methylation of EC genes in EB-day 4 VEGFR2 -positive and negative cells. ........................................................................ 69

Figure 31. Promoter DNA heterogeneity in EC-enriched genes in EB-day 4 VEGFR2 -positive and -negative cells ........................................................................ 70

Figure 32. Demethylation at the CD31 and VE-cadherin promoter in EB-day 7 CD31 -positive cells ........................................................................ 72
## List of Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>5mC</td>
<td>5-methylcytosine</td>
</tr>
<tr>
<td>5hmC</td>
<td>5-hydroxymethylcytosine</td>
</tr>
<tr>
<td>A</td>
<td>adenine base</td>
</tr>
<tr>
<td>AID</td>
<td>activation-induced cytidine deaminase</td>
</tr>
<tr>
<td>ANOVA</td>
<td>analysis of variance</td>
</tr>
<tr>
<td>APC</td>
<td>allophycocyanin</td>
</tr>
<tr>
<td>BrdU</td>
<td>bromodeoxyuridine</td>
</tr>
<tr>
<td>C</td>
<td>cytosine base</td>
</tr>
<tr>
<td>CD31/Pecam1</td>
<td>platelet/ endothelial cell adhesion molecule 1</td>
</tr>
<tr>
<td>ChIP</td>
<td>chromatin immunoprecipitation</td>
</tr>
<tr>
<td>CpG, CG</td>
<td>cytosine - phosphate - guanine</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>Dnmt</td>
<td>DNA methyltransferase</td>
</tr>
<tr>
<td>EB</td>
<td>embryoid body</td>
</tr>
<tr>
<td>EC</td>
<td>endothelial cell</td>
</tr>
<tr>
<td>eNOS</td>
<td>endothelial nitric oxide synthase</td>
</tr>
<tr>
<td>ES</td>
<td>embryonic stem</td>
</tr>
<tr>
<td>FACS</td>
<td>fluorescence-activated cell sorting</td>
</tr>
<tr>
<td>G</td>
<td>guanine base</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>H3K4me3</td>
<td>histone H3 lysine 4 trimethylation</td>
</tr>
<tr>
<td>H3K27me3</td>
<td>histone H3 lysine 27 trimethylation</td>
</tr>
<tr>
<td>HAT</td>
<td>histone acetyltransferase</td>
</tr>
<tr>
<td>HDAC</td>
<td>histone deacetylase</td>
</tr>
<tr>
<td>HUVEC</td>
<td>human umbilical vein endothelial cell</td>
</tr>
<tr>
<td>HMVEC</td>
<td>human microvascular endothelial cell</td>
</tr>
<tr>
<td>IAS</td>
<td>inner antisense primer</td>
</tr>
<tr>
<td>IS</td>
<td>inner sense primer</td>
</tr>
<tr>
<td>Klf2</td>
<td>kruppel-like factor 2</td>
</tr>
<tr>
<td>LIF</td>
<td>leukemia inhibitory factor</td>
</tr>
<tr>
<td>LINE</td>
<td>long interspersed nucleotide elements</td>
</tr>
<tr>
<td>LUMA</td>
<td>luminometric methylation analysis</td>
</tr>
<tr>
<td>MBD4</td>
<td>methyl CG binding domain protein 4</td>
</tr>
</tbody>
</table>
Chapter 1
Literature Review
1 Literature Review

1.1 Epigenetics

1.1.1 Introduction

Epigenetics is defined as chromatin-based mechanisms that regulate gene transcription without changes in DNA sequence per se and include: DNA methylation, histone density and posttranslational modifications, and RNA-based mechanisms. Normal epigenetic marks are critical for mammalian development. Mice that do not maintain DNA methylation marks after cell division or lack one of the histone deacetylases die during embryonic development. Importantly, mutations in epigenetic regulation in humans leads to disease. For example, Rett syndrome is a disorder caused by mutations in methyl CG binding protein 2 (MeCP2) and leads to mental retardation in females.

Three giants of the epigenetic world, Dr. Adrian Bird, Dr. Howard Cedar and Dr. Aharon Razin share the 2011 Canada Gairdner International Award. Together they have shown that DNA methylation patterns are inherited by daughter cells and allele-specific methylation patterns exist at imprinted genes. They also identified methyl binding proteins and CpG islands and linked the protein MeCP2 with Rett Syndrome.

1.1.2 DNA methylation

DNA methylation, 5mC or the fifth-base pair, is a transcriptional repressive mark that occurs predominantly in the context of CpG dinucleotides in mammals. It is estimated that 70–80% of all CG in the mammalian genome are methylated and that this form of cytosine is the most mutagenic base.

DNA methylation is a covalent modification and is thought to be a stable mark with functions in long-term cellular memory. DNA methylation has well-established roles in X-chromosome inactivation in XX females, genomic imprinting and in silencing retroviral elements, heterochromatin and repetitive DNA. It is a modification that occurs after DNA synthesis and is catalyzed by DNA methyltransferases (DNMT). DNMT enzymes catalyze the transfer of a methyl group from the methyl donor, S-adenosyl methionine (SAM) to the carbon 5 position on the cytosine base (Fig. 1). The reaction converts cytosine to 5-methylcytosine (5mC) and SAM to S-adenosyl homo-cysteine (SAH).
Cytosine methylation represses transcription by several mechanisms: by directly hindering the binding of a transcription factor (TF), or indirectly through the recruitment of methyl-binding proteins or co-repressors, which act as competitors for the TF binding site (Fig. 2). The presence of a CG within the consensus sequence of a TF binding site may regulate transcription. The presence of at least one 5mC in the binding sites of E2F, ATF/CREB-like, HIF, and c-Myc have been shown to decrease or turn off transcription of their target genes. Certain TF, such as Sp1 are insensitive to 5mC within their consensus sequence. These TF can be indirectly restricted from accessing the DNA through competitive inhibition. Methyl binding proteins such as MeCP1, MeCP2, MBD1, MBD2 and MBD4 bind to symmetrically methylated CG in CG-rich and -poor regions to form stable and weak complexes, respectively. MeCP2 can add another layer of transcriptional repression through the recruitment histone deacetylases (HDAC) resulting in reduced chromatin accessibility (Fig. 2).

Dnmt1 was the first mammalian DNA methyltransferase discovered. At the time, it was unclear whether other methyltransferases existed and if they had overlapping or distinct functions. Dnmt1 is the maintenance enzyme, which has high affinity for hemi-methylated DNA and functions to transfer the epigenetic signature to daughter cells. It interacts with both nuclear protein 95 (NP95/UHRF1), a methyl-binding protein and proliferating cell nuclear antigen (PCNA), a cofactor for DNA polymerase at replication fork. The presence of CG methylation in Dnmt1-null mouse ES cells, along with their ability to de novo methylate viral DNA suggested the existence of other Dnmt.

Dnmt3a and Dnmt3b are the de novo methyltransferases that preferentially catalyze the addition of a methyl group to unmethylated DNA. These enzymes have non-overlapping functions: Dnmt3b is responsible for the repression of centromeric satellites and Dnmt3a along with Dnmt3L, a catalytically inactive protein, is required for the methylation of maternal imprinted genes in oocytes. Dnmt3L interacts with the unmethylated lysine in the histone H3 lysine 4 (H3K4) tail and induces de novo methylation activity of Dnmt3a. The Lsh (lymphoid specific helicase) protein is also required for de novo methylation of satellite repeats and single copy genes. This may be mediated through the physical interaction with Dnmt3a and Dnmt3b.

Non-CpG nucleotides can be also be methylated, Dnmt1 has a function in de novo methylation, and Dnmt3a and Dnmt3b have a role in maintenance methylation. In the mammalian genome, non-CpG methylation is rare in adult tissues but represents 15-25% of the total methylation found in human and mouse ES cells, especially methylated cytosine followed by an adenine (CpA), methylated cytosine followed by an thymine (CpT) or CpHpG, where the H is
A, C or T\textsuperscript{31}. Also, Dnmt1 and Dnmt3a \textit{de novo} methylate DNA on intact nucleosomes\textsuperscript{34}. The \textit{de novo} activity of Dnmt1 is induced by the activity of Dnmt3a\textsuperscript{35}. Additionally, Dnmt3a and Dnmt3b are present in the somatic cell nucleus of adult mice where they have maintenance activity. They are tightly bound to methylated nucleosomes, especially at repetitive sequences and propagate the DNA methylation signal\textsuperscript{36,37}.

1.1.3 DNA methyltransferase-deficient mice and ES cells

1.1.3.1 Dnmt1 knockout

Dnmt1-null mice, which do not copy the DNA methylation code after cell division, do not survive to term\textsuperscript{24}. At embryonic day 9.5, Dnmt1-null embryos have a deformed neural tube and are unable to repress satellite and retroviral DNA integrated in the genome\textsuperscript{24}. Dnmt1-null ES cells have stable 5mC methylation\textsuperscript{24,38}. Dnmt1-null ES cells have a hypomethylated genome (20% CG methylation vs. 50-70% in wildtype ES cells), yet replicate normally as pluripotent cells. In contrast, they have minimal ability to differentiate \textit{in vitro} and do not contribute to embryoid bodies when co-cultured with wild type cells\textsuperscript{24,38}. Therefore, the maintenance of DNA methylation is not essential for the growth of undifferentiated ES cells but is required for the proliferation of differentiated cells\textsuperscript{24}.

1.1.3.2 Dnmt3a/3b knockout

Dnmt3a and Dnmt3b establish 5mC marks on unmethylated DNA\textsuperscript{25}. Dnmt3a-deficient mice develop to term but die after four weeks. In contrast Dnmt3b-deficient mice are embryonic lethal due to growth and neural tube defects. The double knockouts, Dnmt3a/3b-null mice die before E11.5: they lack somites and do not go through embryonic turning\textsuperscript{25}. Unlike Dnmt1-null ES cells, Dnmt3a/3b-null ES cells do not maintain a stable level of 5mC. Early passage (ie. p5) Dnmt3a/3b-null ES cells have normal methylation levels and demethylate at every cell division. After seventy-five passages, there is only 0.6% of global CG methylation remaining\textsuperscript{34}. DNA methylation loss with mitotic cell division can be rescued by the transfection of Dnmt3a or Dnmt3b cDNA\textsuperscript{34}. Low passage Dnmt3a/3b-null ES cells with 10% CG methylation are able to terminally differentiate, while late passage cells cannot. The presence of Dnmt1 and some threshold of genome-wide methylation is required for ES cell differentiation. At least one \textit{de novo} enzyme is required for maintaining normal global methylation levels.
1.1.3.3 Dnmt1/3a/3b knockout

Mouse ES cells that are deficient in all three of the functional Dnmt are known as triple knockout (TKO) cells. They show a lack of genome-wide CG methylation and an increase in the transcription of repetitive elements\textsuperscript{39}. Despite the lack of CG methylation, TKO ES cells express pluripotency genes, proliferate normally and have normal histone activating and repressing marks. Like their Dnmt1-/- mouse ES cell counterparts, they do not contribute substantially to embryoid bodies\textsuperscript{39}.
Figure 1. The enzymatic conversion between cytosine, 5-methylcytosine and 5-hydroxymethylcytosine.
A) DNA methyltransferases (Dnmt) 1, 3a and 3b catalyze the transfer of a methyl group (-CH$_3$) from the methyl donor S-adenosyl methionine (SAM) to the carbon 5 position on the cytosine (C) base. This converts cytosine to 5-methylcytosine (5mC) and SAM to S-adenosyl homo-cysteine (SAH). B) Tet-eleven translocation (Tet) 1-3 enzymes are Fe(II)-dependent dioxygenases which use 2-oxoglutarate (2-OG) and molecular oxygen (O$_2$) to catalyze the hydroxylation of 5mC to 5-hydroxymethylcytosine (5hmC) in DNA. C) One active DNA demethylation pathway has been shown to convert 5mC to C via a deamination, thymine glycosylation and base excision repair. D) It is hypothesized that 5hmC may be a substrate for active DNA demethylation, however no mechanism or stable intermediates are currently known.
**Figure 2. Mechanisms of transcriptional repression by DNA methylation.**

In the absence of DNA methylation, A) transcription factors (TF) are able to access the upstream regulatory regions and transcribe the gene. Cytosine methylation represses transcription by several mechanisms: by directly hindering the binding of a transcription factor (B), or indirectly through the recruitment of methyl-binding proteins ie. MeCP2 (C) or co-repressors, ie. histone deacetylase (HDAC) (D), which act as competitors for the TF binding site. The arrow marks the transcriptional start site.
1.1.4 DNA hydroxymethylation

The biological relevance of the sixth base pair was recently highlighted. 5-hydroxymethylcytosine (5hmC) was discovered in mammalian Purkinje neurons and has since been found in the central nervous system (brainstem, spinal cord), in organs (heart, kidney, liver) and in ES cells. While the average global 5mC levels remain constant in all adult tissues examined, 5hmC levels are highly variable and may represent a novel mechanism in the regulation of gene expression. ES cells contain about 4% of 5hmC, which is considered abundant and 70–80% 5mC in the mammalian genome. As ES cells differentiate, the global level of 5hmC decreases. Unlike 5mC, 5hmC is correlated with transcriptional activation and is found at euchromatic regions of the genome such as CG islands and gene bodies. Current detection methods using sodium bisulfite do not distinguish between 5mC and 5hmC.

The hydroxylation of 5mC to 5hmC is catalyzed by ten-eleven translocation (Tet) 1-3 enzymes, Fe(II)-dependent dioxygenases which use 2-oxoglutarate (2-OG) and molecular oxygen (O2) (Fig. 1). The Tet1 and Tet2 enzymes are expressed in ES and somatic cells while Tet3 is found primarily in oocytes.

5hmC is not recognized by many of the methyl-CG binding proteins, MBD1, MBD2 and MBD4, or by the maintenance DNA methyltransferases, DNMT1. However, 5hmC is recognized by UHRF, the co-factor for DNMT1 during maintenance methylation, with the same affinity as for 5mC.

1.1.5 DNA demethylation

The process of DNA demethylation occurs when 5mC is replaced with cytosine. Passive DNA demethylation occurs in proliferating cells when Dnmt1 is inactive or absent from the nucleus. Based on the semi-conservative model of DNA replication, without maintenance DNA methylation daughter cells inherit half of the DNA methylation marks after one cell cycle. After fertilization, the maternal pronucleus undergoes passive DNA demethylation after several cell divisions (Fig. 3).

Active DNA demethylation is cell cycle and DNA replication independent. There are few well-established examples of this process. The paternal pronucleus of the mouse undergoes active DNA demethylation within 8 hours of fertilization in the absence of cell division. A second example of active DNA demethylation occurs at the interleukin-2 locus during T cell activation.
via lipopolysaccharide injection in mice\textsuperscript{51}. Native T cells demethylate at 4 key CG sites without entry into the cell cycle\textsuperscript{51}.

The subject of DNA demethylation as a result of a single enzyme is controversial\textsuperscript{52}. MBD2 was argued to fulfill the role of the demethylase, but was subsequently dismissed by two groups not to be involved\textsuperscript{53,54}.

However, a group of enzymes working in sequence can actively demethylate 5mC. Activation-induced cytidine deaminase (AID) recognizes 5mC and deaminates it to form thymine, which results in a thymine-guanine mismatch\textsuperscript{55}. Next, MBD4, one of the methyl-binding proteins, recognizes the mismatch and excises the thymine base using its glycolase activity to form an abasic site. Gadd45 is a non-enzymatic component, and facilitates the interactions with AID and MBD4 (Fig. 3)\textsuperscript{55}. Base excision repair mechanisms may fill the site to replace 5mC with C\textsuperscript{55}.

Additionally, the conversion of 5mC to 5hmC by Tet proteins may provide mechanisms for passive and active DNA demethylation. Upon fertilization, the male pronuclear 5mC levels drop, and an accumulation of 5hmC is observed due to abundant Tet3 in the oocyte\textsuperscript{45,46}. 5mC in the maternal pronucleus is protected from demethylation and oxidation to 5hmC by PGC7/Stella\textsuperscript{46,56}.

5hmC is not recognized by Dnmt1\textsuperscript{49} and would therefore promote passive demethylation. Tet1 promotes active DNA demethylation in the adult brain through a base excision repair pathway involving AID/Apobec\textsuperscript{57}.

1.1.6 Epigenetic changes during differentiation of mouse ES cells

There is a general consensus that mouse ES cells have levels of 5mC that are comparable with terminally, somatic differentiated cells. Moreover, global 5mC levels do not change upon differentiation\textsuperscript{32,33,58}. However, current genomic arrays do not reliably differentiate between 5mC and 5hmC, so the older studies may need revisiting. Variability exists between cell lines and tissues and the method of 5mC quantification. Multiple groups estimate the genome-wide CG methylation as 50-70\% in mouse ES cells and 65-85\% in differentiated cells or tissues\textsuperscript{32,33}. Importantly, there are net changes as ES cells differentiate. A small percent of all CG sites change their methylation status upon differentiation: 8\% become methylated, while 2\% lose the repressive mark\textsuperscript{32} in the mouse genome, corresponding with the activation or repression of lineage-specific genes.

The current paradigm is that many tissue-specific genes have methylated promoters in ES cells while the promoters of pluripotency and housekeeping genes are unmethylated in ES cells\textsuperscript{59,60}. 
Upon ES cell differentiation, the promoters of pluripotency-associated genes such as Oct4 and Nanog show de novo methylation activity corresponding with gene repression\(^{61-64}\). In contrast, many tissue specific genes, especially those regulated by CpG poor promoters demethylate upon in vitro differentiation toward that lineage, corresponding with gene activation\(^{59}\).

Mouse ES cells contain “bivalent chromatin” domains: genomic regions marked by histone H3 lysine 4 trimethylation (H3K4me3) and histone H3 lysine 9 acetylation (H3K9ac) which are associated with transcriptional activation and H3 lysine 27 trimethylation (H3K27me3) which is associated with transcriptional repression\(^{65-67}\). The paradoxical presence of both activating and repressing marks was an important discovery. These domains are often found at tissue-specific genes that are expressed at low levels (or turned off) in ES cells\(^{65-67}\). The bivalency domains are mostly restricted to ES cells and are thought to keep genes poised for activation upon differentiation to a particular lineage\(^{65-67}\). The repressive chromatin marks have been shown to limit transcription since cells lacking polycomb repressive complex 1 or 2 (PRC1 or 2) components show an increased expression of lineage-restricted genes in ES cells\(^{66,68,69}\). Recent studies have shown that the majority of genes with bivalent chromatin domains also have 5hmC at their transcriptional start sites, which may also contribute to the poised state\(^{70}\).

Upon differentiation of ES cells, the bivalent domains are resolved into activating or repressive domains. Lineage-specific genes that are transcriptionally active after differentiation are marked by H3K4Me, increase in 5hmC, and lose 5mC and H3K27me3 marks. Lineage-specific genes that are repressed upon differentiation are heavily marked by 5mC and H3K27me3 with a loss of H3K4me3 and 5hmC at their promoters\(^{43,67,70}\).
Figure 3. Passive and active mechanisms of DNA demethylation.

A) Passive DNA demethylation occurs in proliferating cells when Dnmt1 is inactive or absent from the nucleus. DNA demethylation occurs when 5mC is replaced by C. After two cycles of cell division, there is 75% genome-wide 5mC loss.

B) Active DNA demethylation is replication independent and occurs without cell division. One mechanism to enzymatically demethylated DNA has been shown to include activation-induced cytidine deaminase (AID), methyl CG-binding domain protein 4 (MBD4) and Gadd45. 1) It involves the deamination (NH₂, blue) of 5mC to thymine by AID. 2) The thymine-guanine form mismatches and the thymine base is excised by the glycolase activity of MBD4 to 3) form an abasic site. 4) Base excision repair restores the CG base pairs (H, red). Both AID and MBD4 interact with Gadd45.
1.1.7 CpG islands

5mC is the most mutagenic base in the mammalian genome because of its susceptibility to deamination to thymine (T)\(^{11}\) (Fig. 1). The 5mC to T transition causes the genome to become CpG-depleted and TpG- or AT-enriched. CpG islands are regions of the genome that appear less affected by this phenomenon, where the number of CpG is exactly what is expected based on the abundance of C or G nucleotides. In other words, the number of observed (O) CpG dinucleotides is exactly the number of expected (E) CpG in that sequence (ie. ideally, O/E = 1). The genomic composition of CpG islands is somewhat arbitrary: CpG islands are defined as stretches of DNA that have are >500bp in length with a %C+G content >55% and an O/E >0.65\(^{71}\). Three classes of CpG islands have been described: strong, weak, and poor\(^{72}\). VEGFR2 is an example of a gene with a strong CpG island promoter (Table 1). In contrast, promoters of other genes such as eNOS are poor CpG islands. These regions are CpG-depleted, where the number of CpG is far lower than what is expected based on C+G composition (ie. O/E <<0.65).

Strong CpG islands are usually unmethylated and are sites of transcription initiation\(^{73}\). They are often found at promoters of housekeeping and developmentally important genes\(^{60}\). It appears that these genomic regions are protected from DNA methylation in several ways. RNA polymerase binds these regions and is able to recruit H3K4 methyltransferase. Any modifications on the H3K4 inhibit Dnmt3L binding and therefore inhibit de novo DNA methyltransferases. They are targets for CGBP (CpG-binding protein), a protein that is specific to double stranded, symmetrically unmethylated CpG sites\(^{74}\). This protein may function to compete with Dnmt3a and 3b. A third mechanism for the protection of CpG island methylation maybe the significant amount of 5hmC found at these regions\(^{60}\). 5hmC is associated with transcription and is derived from 5mC\(^{40,43}\). Since 5mC is not observed, this implies that 5mC is very efficiently converted to 5hmC at CpG islands\(^{75}\).

CpG islands are rarely methylated. Some physiological examples include genomic imprinting, X-chromosome inactivation and the repression of repetitive elements. Cancer shows one pathological example where the CpG islands of tumor suppressor genes are hypermethylated, thereby inhibiting their expression.
1.1.8 Histone modifications

DNA is condensed into chromatin using nucleosome units. DNA is wrapped around an octamer, two each of the core histones (H2A, H2B, H3, H4)\(^76\). Adjacent nucleosomes are separated by linker DNA, which associates with histone H1 to further DNA compaction. Core histones are post-translationally modified on their N-terminal tails to dynamically regulate chromatin accessibility. The most common modifications of histone tails include acetylation, methylation and phosphorylation on lysine (K), serine (S) or threonine (T) residues. Histone 3 lysine 27 trimethylation (H3K37me3) and H3K9me3 are signatures of heterochromatin, while H3K4me3 (at the promoter) and H3K36me3 (within the transcribed region) are signatures of transcribed genes. Other histone post-translational modifications and their effect on transcription is summarized in Table 2\(^77\). Additionally, histone density plays a key role in the regulation of transcription\(^78\).

1.1.9 Long non-coding RNA

In addition to DNA methylation, histone post-translational modification and histone density, long non-coding RNA (lncRNA) or long intergenic RNA (lincRNA) provide a third epigenetic mechanism to regulate gene transcription. They are classified as >200 nucleotides in length (based on purification protocols) and function to activate or suppress gene transcription as functional RNA molecules\(^79\). Through stringent screening protocols, it has been well established that lncRNAs are highly abundant in the human and mouse genomes and are functionally important\(^80,81\). XIST is a lncRNA that is in part responsible for X-chromosome inactivation and establishing the correct gene dosage in XX female cells of higher species\(^82\).
Table 1: CpG island classification of murine promoters of endothelial cell-enriched genes.

We examined the CpG island classification of a number of EC-enriched genes in the mouse. These genes can be classified by the observed CG dinucleotides versus expected CG dinucleotides (Obs/Exp or O/E) ratio into strong (>0.65), weak (between 0.65 and 0.48) and poor (<0.48) CpG islands by examining a stretch of 1000bp surrounding the region of interest (ie. inner bisulfite amplicon).

TSS = transcriptional start site, Obs = Observed number of CG, Exp = Expected number of CG

<table>
<thead>
<tr>
<th>Gene</th>
<th>Inner Bisulfite Amplicon Location</th>
<th>#Obs CG</th>
<th>% C+G</th>
<th>Obs CG Exp CG</th>
<th>CpG Island</th>
<th>Reference for TSS</th>
<th>Genbank or Ensemble</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENOS</td>
<td>-230 to +57</td>
<td>27</td>
<td>59</td>
<td>0.34</td>
<td>Poor</td>
<td>83,84</td>
<td>AF091262</td>
</tr>
<tr>
<td>SSRE region (eNOS)</td>
<td>-1138 to -785</td>
<td>8</td>
<td>53</td>
<td>0.12</td>
<td>Poor</td>
<td>84</td>
<td>ENSMUSG00000028978</td>
</tr>
<tr>
<td>CD31</td>
<td>+101 to +368</td>
<td>13</td>
<td>49</td>
<td>0.23</td>
<td>Poor</td>
<td>85</td>
<td>AF412277S4</td>
</tr>
<tr>
<td>VE-cadherin</td>
<td>-230 to +249</td>
<td>17</td>
<td>55</td>
<td>0.25</td>
<td>Poor</td>
<td>86,87</td>
<td>ENSMUSG00000031871</td>
</tr>
<tr>
<td>VWF</td>
<td>+80 to +347</td>
<td>11</td>
<td>55</td>
<td>0.14</td>
<td>Poor</td>
<td>88,89</td>
<td>AF152417</td>
</tr>
<tr>
<td>VEGFR2</td>
<td>-246 to +90</td>
<td>61</td>
<td>61</td>
<td>0.67</td>
<td>Strong</td>
<td>90,91</td>
<td>ENSMUSG0000062960</td>
</tr>
<tr>
<td>Nanog</td>
<td>-159 to +29</td>
<td>19</td>
<td>49</td>
<td>0.32</td>
<td>Poor</td>
<td>92</td>
<td>ENSMUSG0000012396</td>
</tr>
</tbody>
</table>
Table 2: A summary of histone posttranslational modifications and their effects on gene expression

K = lysine, S = serine, T = threonine
Ac = acetylation, me3 = trimethylation, phos = phosphorylation, ub = ubiquitination.

<table>
<thead>
<tr>
<th>Histone Modification</th>
<th>Location</th>
<th>Effect on gene expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histone H3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K9ac</td>
<td>Promoter</td>
<td>On</td>
</tr>
<tr>
<td>K14ac</td>
<td>Promoter</td>
<td>On</td>
</tr>
<tr>
<td>K4me3</td>
<td>Promoter</td>
<td>On</td>
</tr>
<tr>
<td>K9me3</td>
<td>Promoter, heterochromatin</td>
<td>Off</td>
</tr>
<tr>
<td>K27me3</td>
<td>Promoter, heterochromatin</td>
<td>Off</td>
</tr>
<tr>
<td>K36me3</td>
<td>Transcribed region</td>
<td>On</td>
</tr>
<tr>
<td>S10phos</td>
<td>Promoter</td>
<td>On/off</td>
</tr>
<tr>
<td>T6phos</td>
<td>Promoter</td>
<td>Off</td>
</tr>
<tr>
<td>T11phos</td>
<td>Promoter</td>
<td>Off</td>
</tr>
<tr>
<td>Histone H4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K5ac</td>
<td>Promoter</td>
<td>On</td>
</tr>
<tr>
<td>K8ac</td>
<td>Promoter</td>
<td>On</td>
</tr>
<tr>
<td>K12ac</td>
<td>Promoter</td>
<td>On</td>
</tr>
<tr>
<td>K16ac</td>
<td>Promoter</td>
<td>On</td>
</tr>
<tr>
<td>K20me3</td>
<td>Heterochromatin</td>
<td>Off</td>
</tr>
<tr>
<td>Histone H2A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K119ub</td>
<td>Promoter</td>
<td>Off</td>
</tr>
<tr>
<td>Histone H2B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K120ub</td>
<td>Unknown</td>
<td>On</td>
</tr>
</tbody>
</table>
Table 3: Differential DNA methylation at promoters of human EC-enriched genes in human EC and VSMC.

We examined the promoters of EC-enriched genes in primary human cell endothelial and vascular smooth muscle cells. We observed the absence of DNA methylation in EC and the presence of DNA methylation in VSMC in many of the genes. As in the mouse, eNOS, CD31, VE-cadherin and vWF are all classified as CpG poor promoters, while VEGFR1 and VEGFR2 are classified as strong CpG island promoters (Shirodkar et al. in preparation).

- indicates absence of methylation
+ indicates presence of methylation (Shirodkar et al. in preparation).

<table>
<thead>
<tr>
<th>Gene</th>
<th>DNA methylation detected?</th>
<th>Human endothelial cells</th>
<th>Human vascular smooth muscle cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>eNOS</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>CD31</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>VE-cadherin</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>vWF</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ICAM-2</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>P-selectin</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Tie2</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>VEGFR1</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>VEGFR2</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>
1.2 Vascular Endothelium

1.2.1 Introduction

The vascular endothelium is a single layer of endothelial cells (EC), which line the lumen of all blood vessels. EC are critical in the regulation of transport of materials between the blood and tissues, in the maintenance of vascular tone, in inflammation, and in haemostasis\textsuperscript{93}. Endothelial dysfunction is thought to contribute to a diverse set of diseases, including atherosclerosis\textsuperscript{93}.

1.2.2 Endothelial cell-enriched genes

Aside from anatomical location, it is difficult to define the characteristics of an EC due to the heterogeneity in function and phenotype\textsuperscript{94}. It is difficult to pinpoint a single gene, which is expressed by every EC of the body. Microarray studies comparing transcripts from EC and non-EC can help define genes that are enriched in the endothelium. The number of EC-enriched genes depends on the cell types studied, on the stringency of the selection criteria and can vary from dozens to hundreds of genes\textsuperscript{95,96}. We have confirmed the EC-enriched expression of a number of genes using a microarray and quantitative reverse transcriptase PCR (qRT-PCR). Genes such as endothelial nitric oxide synthase (eNOS/Nos3), platelet/endothelial cell adhesion molecule (Pecam1/CD31), vascular endothelial-cadherin (VE-cadherin), von Willebrand factor (vWF), vascular endothelial growth factor receptor 2 (VEGFR2), among others are all highly enriched in endothelial cells (HUVEC and HMVEC) with little to no expression in other cell types assayed (hepatocytes, AoVSMC, keratinocytes) (Fig. 4).

eNOS is expressed in medium to large-sized arteries and is essential the vasculature for the maintenance of blood vessel tone\textsuperscript{97,98}. VE-cadherin is critical for the maintenance of vascular integrity through the formation of adherens junctions between endothelial cells\textsuperscript{99,100}. CD31 also has a cell-cell adhesion function and plays a role in the transendothelial migration of leukocytes and in vasculogenesis\textsuperscript{101,102}. vWF is vital for proper coagulation upon vascular injury and is stored in Weibel-Palade bodies in EC\textsuperscript{103}. Lastly, VEGF signaling through the VEGFR2 receptor are essential for vasculogenesis\textsuperscript{104,105}.

A newer concept is that of endothelial cell-exempt genes, a class of genes, which are expressed in many cell types, and which must be transcriptionally repressed for an EC to function and maintain identity. As a consequence, the dysregulation of these genes may lead to EC dysfunction.
Figure 4. Expression of endothelial-enriched genes in five human primary cell types. RT-qPCR analysis of steady-state transcript levels. Copy number was quantified by comparison with a standard curve and calculated per 100ng of total RNA based on calculated efficiency of exogenously synthesized luciferase mRNA. A) eNOS, B) CD31, C) VE-cadherin, D) vWF and E) VEGFR2 were highly enriched in endothelial cells (HUVEC and HMVEC) with little to no expression in other cell types (hepatocytes, AoVSMC, keratinocytes). Data represents mean±SEM (n=3), * and † denote statistical significance at p<0.05, as compared to HUVEC and HMVEC, respectively (Shirodkar et al. in preparation)
Figure 4. Expression of endothelial-enriched genes in five human primary cell types.
(Shirodkar et al. in preparation)
**Epigenetic regulation of endothelial cell-enriched genes**

There are several hundred of unique cell types in the human body, estimated at 411 in a healthy adult\(^6\). Each genetically identical cell has unique aspects to its morphology, function, life-span and gene expression profiles. Historically, it was thought that combinations of TF interact with DNA elements to define a cell-specific pattern of transcription. One of the best-described examples is MyoD, a TF that alone can induce the differentiation of many lineages to skeletal muscle cells\(^7\).

EC-enriched genes are activated by ubiquitously expressed TF such as Sp and Ets family members and AP1\(^8\). Though GATA TF are cell specific, they are not restricted in expression to vascular EC. Etsrp or Erg1 has been described to act as a regulator of vasculogenesis in the zebrafish\(^9\). However, there is no single regulatory DNA element common to our short list of EC-enriched genes and no master TF has been found to foster EC-specific gene transcription\(^8\). Yet, genes such as eNOS, VE-cadherin, CD31, vWF and VEGFR2 are highly enriched in EC compared with other cells (Fig. 4).

Using eNOS as a model EC-enriched gene, our lab has shown that epigenetic pathways regulate EC gene transcription. Transient transfections of the eNOS promoter-reporter constructs were shown to be transcriptionally active in both primary EC and non-EC\(^10\). In contrast, genomic integration of the eNOS promoter-reporter show an EC-restricted pattern of expression\(^11\). Together, this implies that 1) the availability of appropriate TF in non-EC is necessary to activate the promoter, and 2) chromatin structure has a key role in the regulation of eNOS transcription. The eNOS proximal promoter DNA is hypomethylated in human and mouse EC while non-expressing VSMC are hypermethylated\(^10\). Further, chromatin immunoprecipitation (ChIP) for methyl CG binding protein 2 (MeCP2) shows enrichment in VSMC\(^12\). Treating HeLa and VSMC with 5-azacytidine, a DNA methyltransferase inhibitor induced steady state eNOS mRNA\(^10\), supporting the importance of DNA methylation in transcriptional regulation of this gene.

The eNOS promoter has specific histone post-translational modifications that are absent in non-expressing VSMC. These marks include transcriptionally activating histone H3 lysine 9 acetylation (H3K9ac), histone H3 lysine 4 di- and trimethylation (H3K4me2/3), and histone H4 lysine 12 acetylation (H4K12ac)\(^12\). Inhibiting histone deacetylases (HDAC) using trichostatin A in VSMC resulted in an increase in steady state mRNA and histone H3 and H4 acetylation\(^12\). Epigenetic pathways are clearly important in the cell-specific expression of the eNOS gene.
We have extended our findings with eNOS to other EC-restricted genes. Genes such as CD31, VE-cadherin and vWF have differential DNA methylation at their promoters: EC are hypomethylated while VSMC are hypermethylated (Table 3). Importantly DNA methylation plays a functional role. Inhibiting DNMT with 5-azacytidine induced the expression of EC-enriched genes such as eNOS and CD31 in VSMC\textsuperscript{110} (Shirodkar et al, in preparation).

VEGFR2 has a strong CpG island promoter, which is unmethylated in both EC and VSMC (Table 1, 3). ChIP-ChIP data, validated with classical ChIP from our lab shows that this gene is highly enriched in acetylated H3 and H4 in EC compared to VSMC (Matt Yan, unpublished). In addition, the VEGFR2 promoter is targeted by Sp1 and Sp3, which in part provide a mechanism for EC-enriched expression\textsuperscript{113-115}.

It is clear that the phenomenon of epigenetic regulation is not exclusive to our shortlist of EC-enriched genes. In fact, there are hundreds of genes that are characterized by CpG poor promoters, which are differentially methylated in expressing and non-expressing cell types\textsuperscript{59}. It is important to add that there are genes which are highly methylated in one cell type and unmethylated in many other cell types\textsuperscript{59}. This implies that it may be necessary to keep certain genes repressed to maintain cell function and identity.

1.2.3 Transcriptional regulation of the endothelial cell-enriched genes by shear stress

EC are chronically exposed to hemodynamic forces such as laminar and disturbed shear stress and hydrostatic pressure\textsuperscript{116}. Shear stress is the frictional force of circulating blood flow on the endothelium and has key roles in embryonic cardiovascular development\textsuperscript{117} and in cardiovascular pathology at focal regions in the vascular tree\textsuperscript{116}.

There are many mechanotransducers and flow sensors on the surface of EC. One mechanosensory complex describes cell surface CD31 as the shear stress sensor, VE-cadherin at EC junctions as an adaptor protein, and VEGFR2 as the signal transducer\textsuperscript{118}.

eNOS steady state mRNA levels are regulated by transcription factor binding, by DNA methylation at the eNOS promoter and by shear stress. Promoter mutations in human EC show that shear-induced eNOS transcription relies on the shear stress response element (SSRE), the 5'-GAGACC-3' located at –990 to –984 from the transcriptional start site\textsuperscript{119}. NF-κB constituents p50 and p65 interact with the SSRE and are key for shear-induction\textsuperscript{119}. Additionally, the transcription of Kruppel-like factor 2 (Klf2) transcription factor is induced by flow\textsuperscript{120} and it
positively regulates eNOS transcription by binding to the consensus sequence at –655 to –650 in the human promoter\textsuperscript{121}.

Shear stress controls the transcriptional induction and regulation of eNOS \textit{in vivo}\textsuperscript{97,122}. As a late onset gene, robust transcription of eNOS is observed after unidirectional blood flow at embryonic day 9.5 in the mouse\textsuperscript{97}. eNOS is transcriptionally down-regulated in the high probability (HP) region of the mouse aortic arch that is characterized by disturbed blood flow patterns\textsuperscript{122}. Additionally, endothelium cultured under disturbed and laminar flow \textit{in vitro} mimics the pattern found in the HP and LP regions, respectively\textsuperscript{122}. Recently we have shown a connection between shear stress and eNOS promoter DNA methylation. EC isolated from the HP region of the mouse aortic arch are more highly methylated at the eNOS promoter than EC from the LP region (Fig. 5).

The EC-enriched gene VEGFR2 is transcriptionally induced by laminar shear stress\textsuperscript{123,124} while the steady state mRNA levels of VE-cadherin and CD31 do not change\textsuperscript{125} (Table 4).

1.2.4 \textit{In vitro} differentiation of ES cells into endothelial cells

Historically, stem cells were isolated from teratocarcinomas\textsuperscript{126}. Currently, ES cells are isolated from the inner cell mass of a pre-implantation blastocyst\textsuperscript{127-129}. ES cells help advance studies of genetic abnormalities and model disease, through the generation of knockout mice\textsuperscript{130-132}.

In this study, we are interested in the differentiation of mouse ES cells to all cell lineages followed by cell-surface sorting for EC-markers. The \textit{in vitro} differentiation of ES cells can occur on a 2D surface such as an extracellular matrix protein or within a 3D cell aggregates called embryoid bodies, EB\textsuperscript{133}. Methylcellulose is a semi-solid medium and is used to make colonies of differentiating cells, derived from single cells\textsuperscript{133}. Otherwise, 3D cell aggregates can be made from random collisions of ES cells in suspension culture (called cystic bodies), through the use of gravity (in hanging drops) or through forced aggregation of cells (microwells or aggrewells). Cystic bodies produce a variety of EB sizes, while the hanging drop method and microwells produces more uniformly-sized aggregates which are thought to differentiate synchronously\textsuperscript{133,134}. Despite differences in method of ES cell differentiation, specific EC markers are consistently expressed and therefore can be use to select for endothelial progenitor cells.

EB are used to model the early post implantation embryo and recapitulate many aspects of embryonic vasculogenesis\textsuperscript{135,136}. There are morphological similarities between EB and the early
embryo: a single trophoblast-like cell layer surrounds the EB and cavitation occurs in later stages of EB development\textsuperscript{133}. Especially relevant to the EC lineage, there is co-localization of CD31/Pecam1, VEGFR2/flk1 and VE-cadherin in cell clusters in early EB and the formation of cord-like structures after EB growth\textsuperscript{100,131,137}.

The cells that eventually give rise to blood cells and blood vessels first arise in the lateral plate mesoderm and become hemangioblasts\textsuperscript{138}. Hemangioblasts are a subset of mesodermal cells that co-express VEGFR2/flk1 and Brachyury\textsuperscript{139}. Hemangioblasts give rise to hematopoietic stem cells and angioblasts. Some of the angioblasts give rise to endothelial cells.

Our lab performed \textit{in vitro} differentiation of mouse ES cells through the cystic body method (\textbf{Fig. 6}, data from Dr. Jason Fish). Steady state mRNA were assayed for the expression of Oct3/4, one of the markers of mouse ES cells pluripotency, and several EC-enriched transcripts: VEGFR2/Flk1, VE-cadherin, eNOS/Nos3, CD31/Pecam1 and Tie2\textsuperscript{131,137}. As cells differentiate, the level of Oct4 decreases\textsuperscript{137}. The same pattern of expression is expected for other pluripotency markers such as Nanog\textsuperscript{137}. Mouse ES cells show a low basal level of CD31, vWF and Tie2 mRNA. By EB-day 4, VEGFR2 has relatively higher steady state mRNA than in ES cells, while the other EC-markers show low levels compared to ES cells. By EB-day 6-7, VEGFR2, VE-cadherin, eNOS, CD31, Tie2 show increased steady state mRNA (\textbf{Fig. 6}). Work by others clearly demonstrates the same temporal increase in EC-enriched gene expression during \textit{in vitro} differentiation\textsuperscript{131,137}. At EB-day 4, 10-20\% of EB express the VEGFR2 protein and at EB-day 6, 40-50\% express CD31 as examined by cryosectioning and immunological staining\textsuperscript{131}.

Embryoid bodies at day 3-4 express VEGFR2/flk1\textsuperscript{140}, the earliest marker of lineage commitment into endothelial and hematopoietic lineages\textsuperscript{138}, but do not express other EC markers. This receptor is a well-studied cell surface marker that has been used extensively to select EC derived from ES cells\textsuperscript{138,140,141}. EB-day 3-4 VEGFR2-positive cells eventually become EC, VSMC and hematopoietic cells\textsuperscript{140}. VEGFR2 expression is present in the mature endothelium, but not in mature hematopoietic cells\textsuperscript{135}. EB-day 4 cells that stain positive for cell surface VEGFR2 will be considered as “early” endothelial progenitor cells.

Day 6-8 EB positively stained for cell surface CD31 exhibit phenotypical and functional properties of EC\textsuperscript{137}. They express mRNA and protein of many mature EC genes such as VE-cadherin, vWF, CD31 and ICAM-2. Additionally, CD31-positive cells form tube structures on
matrigel and take up acetylated LDL\textsuperscript{137}. In this study, EB-day 7 cells that stain positive for cell surface CD31 will be considered as “late” endothelial progenitor cells.

Using mouse ES cells to model patterns of mammalian, especially human differentiation has many limitations. There are many differences between human and mouse ES cells, including the expression of cell surface antigens, gene expression profiles and mechanism to maintain pluripotency\textsuperscript{142}. Nevertheless, there are large similarities in gene expression profiles and function of ES cell-derived EC, including the isolation of CD31-positive cells for the selection of vascular EC\textsuperscript{143}. The \textit{in vitro} differentiation is not identical to the \textit{in vivo} environment. Some cell organization has been described in EB culture but there is no clear tissue development that is found in the embryo. The lack of orchestrated organization may affect differentiation through the lack of proper cell-cell contacts, paracrine signaling and biomechanical forces.
Figure 5. Increase in eNOS promoter DNA methylation in EC from the high probability region of the murine aortic arch. EC from the high probability (HP) and low probability (LP) regions of the aortic arch were isolated. DNA was treated with sodium bisulfite, PCR amplified and sequenced at CG sites to estimate the average % methylation at the eNOS promoter. HP EC are more methylated than LP EC at the eNOS promoter. Data is presented as the mean ± SEM (n=4). Two way ANOVA between cell types indicates that the groups are statistically different (*p<0.001).
Figure 6. Expression of Oct4 and EC-enriched genes during in vitro differentiation. RT-qPCR analysis of steady-state transcript levels. Mouse E14 ES cells were differentiated by the removal of LIF to formed into cystic bodies. A) The expression of Oct4 rapidly decreases as cells differentiate while B-F) the expression of EC markers, Flk1/VEGFR2, VE-cadherin, Nos3/eNOS, Pecam1/CD31 and Tie2 increases. Expression is shown the relative to GAPDH and TBP and normalized to expression in ES cells. ES= mouse embryonic stem cell, EB-d = embryoid body day (data: Dr. J. Fish)
Table 4: Transcriptional regulation of the endothelial cell-enriched genes by shear stress.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Transcriptionally regulated by shear stress?</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>eNOS</td>
<td>Yes – positive</td>
<td>119,122</td>
</tr>
<tr>
<td>CD31</td>
<td>No</td>
<td>122</td>
</tr>
<tr>
<td>VE-cadherin</td>
<td>No</td>
<td>125</td>
</tr>
<tr>
<td>VEGFR2</td>
<td>Yes – positive</td>
<td>123,124</td>
</tr>
<tr>
<td>Klf2</td>
<td>Yes – positive</td>
<td>120,121</td>
</tr>
</tbody>
</table>
1.3 Methods to quantify DNA methylation

1.3.1 Sodium bisulfite genomic sequencing

The sodium bisulfite genomic sequencing method is widely used in DNA methylation analysis. Genomic DNA is treated with sodium bisulfite to distinguish between methylated (5mC) and unmethylated cytosines (C) in the context of CpG dinucleotides. 5mC are protected from deamination, while C are converted to uracil (U) (Fig. 7A-B). Through subsequent PCR amplification and sequencing, 5mC retains its identity, while a C becomes converted to a thymine (T).

DNA is denatured using strong alkali or quick heating and cooling to generate single stranded DNA (ssDNA). The conversion reaction occurs in successive steps (Fig. 7A): sulphonation, hydrolytic deamination and alkali desulphonation. Sodium bisulfite reacts with the double bond between carbons 5 and 6 of C to form the cytosine sulphonate adduct. Next, hydrolytic deamination occurs as cytosine sulphonate is deaminated at carbon 4 to form the uracil sulphonate adduct. Finally, alkali desuphonation occurs as uracil sulphonate is converted to U.

In contrast to cytosine, (Fig. 7B) 5mC reacts one hundred times slower with bisulfite treatment compared to C. Therefore the reaction does not proceed at a quantifiable rate\textsuperscript{144}. The methyl group protects 5mC from becoming sulphonated and converted. While 5hmC does react with bisulfite (Fig. 7C), it produces the cytosine 5-methylene sulphonate and this product also doesn’t proceed with the remaining two steps of the conversion reaction. This leaves a very large and bulky product that is known to inhibit the processivity of heat stable DNA-dependent DNA polymerase (ie. Taq polymerase).

Since neither 5mC nor 5hmC are converted to uracil, they are indistinguishable using the sodium bisulfite method. However, as noted above, the bisulfite adduct of 5hmC delays PCR amplification at CG sites\textsuperscript{144}. This implies >99% of 5mC observed from sodium bisulfite genomic sequencing is true to its identity. Additionally, this implies that 5hmC-rich templates will be under represented using this technique and/or misread as a 5mC.

The MeDIP or hMeDIP method (methylated or hydroxymethylated- DNA immunoprecipitation) relies of the immunoprecipitation of chromatin fragments and can distinguish between 5mC and 5hmC through the reactivity with specific antibodies.
After sodium bisulfite conversion, all C are converted to U (Fig. 7-8). The initial unmethylated CG in the genomic sequence has been converted to UG and DNA remains single stranded since it cannot re-anneal due to the mismatches. DNA is PCR amplified at specific loci, ie. promoter regions with primers designed for bisulfite converted DNA. Following PCR, the UG pairs are replaced with TG as U complements with adenine (and A complements to T on its complementary strand). It is important to note that PCR primers can be designed to either the new (+) or the new (-) strand (Fig. 8).

A nested PCR is commonly used. After PCR amplification, the bisulfite PCR products can be ligated into plasmid DNA vectors, grown in bacteria, isolated and sequenced to analyze the methylation status of all CG in individual strands of DNA (clonal analysis). This method provides high resolution but may not represent the entire population of the sample due to small sampling size ie. 20 strands of DNA from a pool of >10^6 DNA strands. Alternatively, the PCR products can be directly pyrosequenced, as discussed in detail in the following section (Fig. 8).

### 1.3.2 Pyrosequencing

Pyrosequencing is a recent advance in sequencing technology that does not require labeled primers or gel electrophoresis. It is based on the release of pyrophosphate (PP$_i$) from a dNTP during its incorporation into a growing DNA molecule$.^{145}$ The presence of pyrophosphate in solution begins a series of enzymatic reactions and ends with the emission of light. The light intensity is directly proportional to the quantity of the dNTP incorporated by the polymerase. Since the dispensation order of the nucleotides is known, and can be automated, the template sequence can be found.

DNA is denatured and only one of the strands is sequenced using a single primer. First, a nucleotide in the form of dNTP is added to the sample. If it is complementary to the template, the DNA polymerase will extend the molecule by one base and release PP$_i$.[$^1$]
where:

\((\text{DNA})_n\) is a sequence of nucleotides of length n
PP\(_i\) is inorganic pyrophosphate
APS is adenosine phosphosulfate
ATP is adenosine triphosphate
AMP is adenosine monophosphate
dNTP is deoxynucleotide triphosphate

In the presence of PP\(_i\), ATP sulfurylase converts adenosine phosphosulfate (APS) to ATP and sulphate \([2]\). Subsequently, the ATP is required by luciferase to convert luciferin into oxyluciferin and produce light\(^{145}\). The light is captured by a charged-coupled device (CCD) camera which converts light intensity into an electric signal. Before the next nucleotide is dispensed, apyrase degrades unincorporated nucleotides from the prior addition step.

Since ATP is present in two of the three reactions above, another form of adenine is necessary for DNA synthesis and nucleotide detection. Deoxyadenosine-5’(α-thio) triphosphate (dATPαS) is used instead of dATP in \([1]\), given that this base can be incorporated into the DNA by the polymerase without interfering with luciferase\(^{146}\).

The pyrosequencing technology has wide applicability and it is used here to 1) sequence bisulfite-converted, PCR amplified DNA at specific loci ie. gene promoters and in the LINE-1 assay, and 2) sequence by synthesis the fragment ends generated by the DNA restriction enzyme endonuclease activity. This is called the LUMA assay and is described below.

To analyze bisulfite-converted and PCR amplified DNA with pyrosequencing, a biotinylated primer is used to purify and separate the individual DNA strands. Since the sequence is already known, the dNTP nucleotide order is simple except at CG sites, where a dTTP is first added, followed by a dCTP in a single nucleotide polymorphism (SNP) analysis. The ratio of the light intensity emitted from the T:C incorporation represents the ratio of unmethylated: methylated cytosines at that CG site. Pyrosequencing is a quantitative method. The output is the average percent methylation at one CG site across the entire pool of DNA strands in the sample. However, it is a low-resolution method and it is impossible to infer how many of the CG sites are methylated in a single DNA strand.
Figure 7. The conversion of cytosine but not 5mC or 5hmC to uracil with sodium bisulfite treatment.
A) Sulphonation occurs as bisulfite reacts with the double bond between carbons 5 and 6 of C to form the cytosine sulphonate adduct. Next, hydrolytic deamination occurs as cytosine sulphonate is deaminated at carbon 4 to form the uracil sulphonate adduct. Finally, alkali desuphonation occurs as uracil sulphonate is converted to uracil. 
B) 5mC is protected from deamination. The bisulfite conversion of 5mC is 100 times slower than C. 
C) 5hmC reacts rapidly with bisulfite to form a cytosine 5-methylenesulphonate adduct. The sodium bisulfite method cannot distinguish between 5mC and 5hmC. R = residue, the bond leading to the deoxyribose sugar in the DNA strand.
Figure 8. Sodium bisulfite genomic sequencing.
Methylated CG are shown in blue while unmethylated CG are in green. Double stranded genomic DNA is denatured with strong alkali and/or rapid heating and cooling. DNA is sodium bisulfite converted and all unmethylated C deaminate into U while 5mC are protected. The single stranded, bisulfite converted DNA is PCR amplified and the U converted into T. PCR products are further subloned and sequenced for clonal analysis of single DNA strands or are directly pyrosequenced for analysis of the average pool of DNA strands in the sample. C= cytosine, U= uracil, T= thymine bases
1.3.3 LINE-1 assay

Long interspersed nucleotide elements (LINE-1) or L1 elements are abundant repetitive elements. They are distributed on many chromosomes, can be found at >500,000 copies and comprise 19% of the human and mouse genomes\(^\text{147}\). In the mouse, they have a full repeat length of 7kb, although most copies are truncated at the 5’ end resulting in a wide range of sizes\(^\text{148}\). LINEs have an open reading frame that encodes for reverse transcriptase and endonuclease enzymes. Since they contain a 5’ consensus sequence for RNA polymerase II, they can be transcribed, reverse transcribed and then inserted randomly into the genome\(^\text{148}\). In fact, the original function of DNA methylation was thought to repress retrotransposition of repetitive DNA to maintain genome integrity\(^\text{149}\). The upstream regulatory region of repetitive elements is highly methylated limiting their transcription and transposition.

Since these elements are highly abundant and highly methylated, they are a useful tool for estimating genome-wide DNA methylation levels in mature or terminally differentiated cell types\(^\text{150}\). The LINE-1 assay utilizes bisulfite converted DNA that is PCR amplified under low stringency conditions using universal primers targeted to a consensus sequence\(^\text{150}\). Therefore, the resulting PCR products represent thousands of repeats. Known CG sites are analyzed within the PCR amplicon using pyrosequencing (SNP analysis) to estimate global DNA methylation levels.

One caveat in using the LINE-1 assay is the poor estimate of genome-wide methylation in ES cells which show retrotransposon activity. Unlike other repetitive elements such as satellite, simple and tandem repeats, LINE-1 are hypomethylated in mES cells\(^\text{43}\).
1.3.4 LUMA assay

LUMA assays genome-wide levels of CG methylation at 5’-CmCGG-3’ sites using isoschizomer restriction enzymes: *MspI* is an efficient cutter and is insensitive to methylation at the internal CG while *HpaII* will only cut when the DNA is unmethylated. It also works more efficiently in the presence of multiple adjacent *HpaII* sites. Digesting the DNA using *EcoRI* (at 5’-GAATTC-3’) acts as an internal control for mass amount of input DNA. Samples are analyzed using pyrosequencing to sequence the sticky ends through synthesis.

Genomic DNA is digested with *MspI+EcoRI* or *HpaII+EcoRI*. *MspI* and *HpaII* leave a 5’-CG overhang while *EcoRI* cuts to produce a 5’-AATT overhang. The dispensation order of dinucleotide triphosphates is:

1. dATPdS
2. dCTP+cGTP
3. dTTP
4. dCTP+cGTP

The light intensity for the addition of dATP and dTTP (steps 1 and 3) are to fill in the overhang from the *EcoRI* digest and theoretically should be equivalent. The second step (2) adds dCTP and dGTP fill in the ends of the *HpaII* and/or *MspI* fragments and are used to calculate the percent global DNA methylation. The last step adds excess of dCTP and dGTP and acts as a control to ensure that all of the ends have been extended. The amount of light released is proportional to the number of overhangs produced by the respective restriction enzymes.

The *HpaII/MspI* ratio estimates the unmethylated fraction of the genome and is calculated as 

\[
\frac{HpaII}{EcoRI} / \frac{MspI}{EcoRI}
\]

The global percent methylation is calculated by:

Percent methylation = 1 - *HpaII/MspI* *100%

1.4 Objective

We have shown differential promoter DNA methylation in human EC and VSMC at promoters of EC-enriched genes. We aimed to follow promoter DNA methylation of EC-enriched genes as mouse ES cells differentiate to EB-day 4 VEGFR2 (+) and (-) and EB-day 7 CD31 (+) and (-) cells.
Chapter 2

DNA methylation changes at promoters of endothelial cell-enriched genes during *in vitro* differentiation
2 DNA methylation changes at promoters of endothelial cell-enriched genes during *in vitro* differentiation

2.1 Rationale and Hypothesis

Our lab is interested in the epigenetic regulation of EC gene expression. We used a microarray to screen several EC and non-EC primary human cell lines for genes that are highly enriched in EC. We found that the genes eNOS, CD31, VE-cadherin, vWF and VEGFR2, among others, show an EC-enriched pattern of expression. We have also shown that these EC-enriched genes show a differential pattern of promoter DNA methylation in human EC and vascular smooth muscle cells (VSMC) (Table 3). The promoters of these genes are unmethylated in expressing EC and densely methylated in non-expressing VSMC. Inhibition of DNA methyltransferases with 5-azacytidine induces the expression of eNOS and CD31 in VSMC\(^\text{110}\) (Shirodkar et al. in preparation).

During *in vitro* differentiation of mouse ES cells, EB-day 3-4 VEGFR2-positive cells and give rise to EC, VSMC and hematopoetic cells\(^\text{140}\). VEGFR2 is the earliest marker of lineage commitment towards EC\(^\text{138}\). Later stage endothelial progenitor cells can be isolated by sorting EB day 6-8 cells for cell surface CD31\(^\text{137,143}\). These cells exhibit some phenotypical and functional properties of EC.

This study asks: How does the DNA methylation pattern that we find in terminally differentiated EC and VSMC become established? To follow the changes in promoter DNA methylation from pluripotent ES cells to terminally differentiated cells, we use a murine *in vitro* model of differentiation. Mouse ES cells are differentiated into all germ layers by using spherical cell aggregates, or embryoid bodies (EB). They are allowed to grow for a total of 4 or 7 days, at which point they are sorted based on cell surface expression of VEGFR2 or CD31, respectively. The differentiation cascade begins with ES cells, followed by EB-day 4 VEGFR2 (+) cells and EB-day 7 CD31 (+) cells acting as “early” and “late” endothelial progenitor cells, respectively. EB-day 4 VEGFR2 (-) cells and EB-day 7 CD31 (-) cells act as points of comparison in the intermediate stages since they are non-EC. Mature EC and VSMC isolated from the thoracic aorta of the mouse are used as terminally differentiated cells in the cascade. The degree of global and promoter-specific DNA methylation is analyzed in these cell types (Fig. 9).
Hypothesis 1:

The proximal promoter regions of EC-enriched genes are hypomethylated in murine aortic EC. Conversely, the proximal promoter regions of EC-enriched genes are methylated in non-expressing murine cell types, such as VSMC.

Hypothesis 2:

The promoters of EC-enriched genes are methylated in mouse ES cells and demethylate during differentiation to murine EC.
Figure 9. The *in vitro* differentiation of murine ES cells into “early stage” and “late stage” endothelial progenitor cells. Mouse ES were differentiated into embryoid bodies (EB) using hanging drops. EB were grown for 4 days and then sorted for cell surface VEGFR2, a marker of cardiovascular progenitors. EB-day 4 VEGFR2 (+) cells represent an “early stage” in endothelial differentiation. EB grown for 7 days and then sorted for cell surface CD31. EB-day 7 CD31 (+) cells represent a “late stage” in EC differentiation, and most become EC while EB-day 7 CD31 (-) cells are non-EC.
2.2 Methods

2.2.1 Cell culture

Murine ES line E14 cells (generously provided by Dr. WL Stanford) were maintained at 37°C in humidified air with 5% CO₂. The cells were cultured in Dulbecco’s modified Eagle’s medium (D-MEM) with high glucose (4,500 mg/L) and without L-glutamine or sodium pyruvate. The ES Maintenance medium was supplemented with 15% fetal bovine serum (Gibco, ES qualified), 0.1mM MEM non-essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), 25 U/ml penicillin and 25 µg/ml streptomycin (Gibco), 55 mM beta-mercaptoethanol (Gibco), 2 mM glutamine (GlutaMAX, Gibco), 1000 U/ml leukemia inhibitory factor (LIF; Chemicon). The culture medium was replaced daily and ES cells were passaged every two days or as needed with 0.05% trypsin-EDTA onto 0.1% gelatin (Sigma) coated tissue culture plates.

DNA harvested from R1 mES cell lines was kindly provided by the lab of Dr. WL Stanford. Genomic DNA from J1, TKO (Dnmt1,3a,3b-null) and extraembryonic mouse stem cell lines XEN and TS was kindly provided by the lab of Dr. J. Rossant. H2 and H9 hES cell lines were kindly provided by Manuel Alvarez from the McLaughlin-McEwen Pluripotent Stem Cell Core.

Human umbilical vein endothelial cells (HUVEC) were isolated from fresh umbilical cords as previously described. Cryopreserved human aortic vascular smooth muscle cells (AoVSMC) cells were obtained from ScienCell and maintained as recommended. Human epidermal keratinocytes and human hepatocytes were obtained from Lonza (Lonza, Basel, CH) and maintained as recommended. Genomic DNA from primary HUVEC, AoVSMC, keratinocytes and hepatocytes cells was kindly provided by Apurva Shirodkar.

2.2.2 mES cell differentiation

ES cells were induced to differentiate upon LIF removal and formed embryoid bodies (EB) using the hanging drop method. Cells were counted on a Beckman Coulter counter or haemocytometer and ~500 ES cells were suspended in 20 ul drops on the lids of 150mm bacteriological grade dishes and left inverted for two days (Fig. 10-11). At day 2, EBs were transferred to 60mm bacteriological grade dishes (STARDISH, Phoenix Biomedical) and cultured in Differentiation media (identical to ES Maintenance media except without LIF). Media was replaced daily.
2.2.3 Cell sorting

Day 4 and 7 EB were washed in warm calcium and magnesium-free PBS for >10min, disaggregated using Accutase (Chemicon) for 10-30 minutes at 37°C, followed by gentle pipetting to form a single cell suspension. Cells were passed through a Pre-separation filter (Miltenyi 130-041-407) to remove cell clumps, and up to $10^7$ cells were stained with APC rat anti-mouse VEGFR2 antibody (1:100, BD 560070) or APC rat anti-mouse CD31 (1:100 BD 551262) in 100ul for 30 minutes at 4°C. A FACSaria cell sorter (BD Biosciences) was used to exclude dead cells (using 7-Aminoactinomycin D), exclude cell aggregates and to select for positively- and negatively-labeled fractions. Top and bottom 5% of the EB-day 4 VEGFR2-labelled pool were selected as VEGFR2-positive and -negative fractions, respectively. Top and bottom 15% of the EB-day 7 CD31-labelled pool were selected as the CD31-positive and -negative fractions, respectively (Fig. 12).
Figure 10. Mouse ES cell and EB morphology.
A) The E14 line of mouse ES cells are feeder-independent and cultured on gelatin. Pluripotent ES colonies (i) form compact hemispheres with round perimeters while colonies that have differentiated (ii) are flatter and have many cellular extensions. B) ES cells differentiated using the hanging drop method form spherical EB. Two EB are shown floating in suspension culture. Images captured under phase contrast. ES = embryonic stem, EB = embryoid body.
Embryoid Body (EB) Formation: Hanging Drop Method

1. Make a single-cell suspension of mouse ES cells (EB-day 0)
2. Dilute cells to 25 cells/ul and make 20ul drops on a non-adhesive surface ie. petri dish lid, 500 ES cells/ drop:
3. Invert the lid for two days allowing the cells to aggregate at the bottom of the drop due to gravity (EB-day 2):
4. Transfer and grow the aggregates in a petri dish or non-adherent plates for 2 days (EB-day 4) or 5 days (EB-day 7):

Cell Sorting, DNA extraction and Sodium Bisulfite Conversion

1. Enzymatically dissociate embryoid bodies into single cells:
2. Sort cells based on cell-surface marker expression (ie. EB-d4 VEGFR2):
3. Lyse cells and extract DNA
4. Sodium bisulfite convert DNA, PCR amplify, clone and sequence:

Methylated CG

\[
\text{m-TCCATCGCT-} \\
\text{m-AGGTAGCGA-} \\
\text{m-TUUATCGUT-} \\
\text{m-AGGTAGCGA-}
\]

Unmethylated CG

\[
\text{-TCCATCGCT-} \\
\text{-AGGTAGCGA-} \\
\text{-TUUATCGUT-} \\
\text{-AGGTAGCGA-}
\]

Figure 11. Brief summary of cell culture and methods. To differentiate mouse ES cells, the hanging drop method of EB formation was used (left). It relies on the force of gravity and cell-cell adhesion to form spherical aggregates. Following 2 or 5 days of culture in suspension, the EB are enzymatically dispersed into single cells (right). Next, cell are sorted using FACS for cell surface VEGFR2 or CD31. DNA was extracted, treated with sodium bisulfite and PCR amplified for cloning and sequencing.
Figure 12. Cell sorting for “early” and “late” stages in endothelial cell differentiation. Mouse ES cells were differentiated for 4 or 7 days and sorted for cell surface markers to select for “early” and “late” EC. FACS was used to sort for (A) the top 5% EB-day 4 VEGFR2 (+) cells (early EC), the bottom 5% EB-day 4 VEGFR2 (-) cells, (B) the top 15% EB-day 7 CD31 (+) cells (late EC) and bottom 15% EB-day 7 CD31 (-) cells.
2.2.4 DNA extraction

DNA from abundant samples (>10^6 cells) was extracted using a standard phenol/chloroform method with salt-ethanol precipitation. DNA from FACS-sorted cells was extracted using the QIAamp DNA Blood Mini Kit (Qiagen) according to the manufacturer’s protocol. The thoracic aorta was dissected from FVB/N mice (3-6 months), opened en face and incubated with Liberase Blendzyme 2 (Roche Applied Science) for 10 minutes at 37°C. The descending thoracic aorta was scraped to isolate EC and VSMC, which were subsequently pelleted and digested for 4 hours with proteinase K (Fermentas, 20mg/ml).

2.2.5 Sodium bisulfite conversion of DNA

Up to 500ng of genomic DNA was subjected to sodium bisulfite treatment using the EZ DNA Methylation Kit (Zymo Research, CA) following the manufacturer’s protocol. During the DNA denaturation step, the DNA was denatured for 30 minutes at 42°C instead of 15 minutes at 37°C. The conversion of C to T in a non-CG context was 99% or higher, as detected after PCR amplification and sequencing.

2.2.6 Sodium bisulfite genomic sequencing

An aliquot of the bisulfite-treated DNA (>10ng, 1-3ul of the 10ul elution) was subjected to nested PCR amplification in a 50ul reaction: 1x PCR buffer, 1.5mM MgCl₂, 200uM each dNTP, 10pmol of each primer and 1 unit of Platinum Taq DNA polymerase (Invitrogen). Thermocycler conditions were: 95°C 5 min; 35x (95°C 30 s; annealing temp 30 s; 72°C 30 s); 72°C 15 min. The first PCR was completed with outer primers (outer sense = OS, outer antisense = OAS). The 1st PCR products (2ul) were used in the subsequent 2nd PCR as template with the inner primers (inner sense = IS, inner antisense = IAS).

For analysis of individual DNA strands, 1st or 2nd PCR products were subcloned into the pCRII-TOPO vector (Invitrogen) and transformed in chemically competent TOP10 E. coli cells or subcloned into pCRII vector (Invitrogen) and transformed into electrocompetent DH10β E. coli cells. Individual plasmid clones (15-24) were prepared for sequencing using the QuickLyse Miniprep Kit (Qiagen).

Sodium bisulfite genomic primers are listed in Table 5. Primer for murine sequences were designed to assay the homologous genomic regions as those previously assayed in human cells (Table 3). Murine Nanog bisulfite primers were designed to assay the same region as in
Multiple sequence alignments were made between human and mouse genomic DNA at the promoters of eNOS, CD31, VE-cadherin, vWF, VEGFR2 and Nanog using ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2) (**Fig. 13-19**). The genomic sequences were in situ sodium bisulfite converted, repeat masked (http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker) and nested primers were designed using Oligo 7 Primer Analysis Software.

### 2.2.7 PCR product pyrosequencing

For pyrosequencing analysis, the 1st PCR product (2-5ul) was used as a template in the reaction. Pyrosequencing primers for bisulfite converted DNA were designed by EpigenDx (Worcester, MA, USA) and are proprietary, the assay ID and annealing temperatures are listed in **Table 6**. For eNOS, CD31 and VEGFR2, the 50ul reaction contained: 1x PCR buffer, 1.5mM MgCl2, 200uM each dNTP, 4pmol each of the Forward and Reverse-biotinylated primer and 1 unit of Platinum Taq DNA polymerase (Invitrogen). Thermocycler conditions are as follows: 95°C 15 min; 45 x (95°C 30 s; annealing temp 30 s; 72°C 30 s); 72°C 5 min. For vWF and VE-cadherin, the 50ul reaction contained: 1x PCR buffer, 1.5mM MgCl2, 200uM each dNTP, 2pmol of Forward primer, 0.8pmol of Reverse-tailed primer, 1.8pmol of Universal reverse biotinylated primer and 1 unit of Platinum Taq DNA polymerase (Invitrogen). After amplification, biotinylated PCR products were bound to sepharose beads, purified, washed and denatured in 0.2M NaOH and then washed with Pyrosequencing Vacuum Prep Tool (Pyrosequencing Inc.). The sequencing primer (0.3uM) was used with the single stranded template and analysed with a PSQ96HS Instrument (EpigenDx Inc, Worcester, MA, USA) in a SNP (C/T) analysis. The relative amount of 5mC was calculated as the percentage of cytosines divided by the sum of methylated and unmethylated cytosines at a particular CG site: 5mC% = 5mC/(5mC+unmethylated C)\(^{145}\).
Table 5: Murine sodium bisulfite sequencing primers

A nested PCR amplification designed for bisulfite converted DNA was used to amplify the promoters of Nanog and EC-enriched genes.

Ta = annealing temperature
OS = outer sense, OAS = outer antisense
IS = inner sense, IAS = inner antisense

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence</th>
<th>Amplicon Location</th>
<th>Amplicon Size (bp)</th>
<th>Ta (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>eNOS</strong></td>
<td>OS</td>
<td>5’ AGATAGGAGAGGAGTAAGGTGAATT 3’</td>
<td>-409 to +73</td>
<td>482</td>
</tr>
<tr>
<td></td>
<td>OAS</td>
<td>5’ CCCTAAACCACAAAATAACCCTAACTC 3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IAS</td>
<td>5’ GGGTTCAGTTTATTAGTTTTAGTTTT 3’</td>
<td>-230 to +57</td>
<td>287</td>
</tr>
<tr>
<td></td>
<td>IAS</td>
<td>5’ AACCCCAAACCTACCCACACCTCTCC 3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SSRE region (eNOS)</strong></td>
<td>OS</td>
<td>5’ ATAGGGTTTTTTATTTTGTG 3’</td>
<td>-1178 to -776</td>
<td>403</td>
</tr>
<tr>
<td></td>
<td>OAS</td>
<td>5’ AAAAAATATAATTCTATACTATCTACAAC 3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IAS</td>
<td>5’ GTTTTTGTTTTAATTTAGTTGGAATTT 3’</td>
<td>-1138 to -785</td>
<td>354</td>
</tr>
<tr>
<td></td>
<td>IAS</td>
<td>5’ AAAAAAACCTACCTCTTCCTACAACAAAC 3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CD31</strong></td>
<td>OS</td>
<td>5’ AGGAGGTATGATTATATTATTATG 3’</td>
<td>-27 to +462</td>
<td>490</td>
</tr>
<tr>
<td></td>
<td>OAS</td>
<td>5’ TAAAACACTCAACTAATTCTCTATT 3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IAS</td>
<td>5’ AGGAGGAAGAAGTTTAAAATTTAGAAGTTT 3’</td>
<td>+101 to +368</td>
<td>293</td>
</tr>
<tr>
<td></td>
<td>IAS</td>
<td>5’ CACAAAAACCCAAAAACCTACTATAAAAC 3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>VE-cadherin</strong></td>
<td>OS</td>
<td>5’ AGGGATGGTGTTTTGTTATTTTGGTAT 3’</td>
<td>-312 to +306</td>
<td>619</td>
</tr>
<tr>
<td></td>
<td>OAS</td>
<td>5’ GACATCTACTTTACAAAATACATAAA 3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IAS</td>
<td>5’ TTGAGGTATGAGTTGGAATTTATTTT 3’</td>
<td>-230 to +249</td>
<td>480</td>
</tr>
<tr>
<td></td>
<td>IAS</td>
<td>5’ TTAATCTCTACCTACTAATAACTATAAA 3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>vWF</strong></td>
<td>OS</td>
<td>5’ ATTTGGAGTTTAGTTATTAGTTAGGAGG 3’</td>
<td>+1 to +531</td>
<td>531</td>
</tr>
<tr>
<td></td>
<td>OAS</td>
<td>5’ CCTAAAAACACCAAAAAATATAT 3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IAS</td>
<td>5’ TTTTGGATTTTATTTAGTTAGGATTTAGG 3’</td>
<td>+80 to +347</td>
<td>268</td>
</tr>
<tr>
<td></td>
<td>IAS</td>
<td>5’ ACAAAATCACCTTACTTATAAAAC 3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>VEGFR2</strong></td>
<td>OS</td>
<td>5’ TTTTTGTTGTTTTTGTAGTTT 3’</td>
<td>-270 to +393</td>
<td>666</td>
</tr>
<tr>
<td></td>
<td>OAS</td>
<td>5’ CTTTTACTCAAACAAAAATCTCTCTAT 3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IAS</td>
<td>5’ TGGGGGTATGTTTTGTTATTAGGATTTAGG 3’</td>
<td>-246 to +90</td>
<td>349</td>
</tr>
<tr>
<td></td>
<td>IAS</td>
<td>5’ GAAACACACTACCTACTATATAAAAC 3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nanog</strong></td>
<td>OS</td>
<td>5’ GGGGTATTTTGTTAGTTGGATTTAAGTTTT 3’</td>
<td>-240 to +203</td>
<td>443</td>
</tr>
<tr>
<td></td>
<td>OAS</td>
<td>5’ ACAAACACTACCAAAACACCTACCTACAAC 3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IAS</td>
<td>5’ GGGTATTTTGTTAGTTGGATTTAACTCTAC 3’</td>
<td>-159 to +29</td>
<td>189</td>
</tr>
</tbody>
</table>
Table 6: Murine PCR pyrosequencing primers.

FP = forward primer, FS = forward primer sequencing
RPB = reverse primer biotinylated, RPT = reverse primer tailed
Uni RPB = universal reverse biotinylated primer
Ta = annealing temperature
Primers were designed by EpigenDx Inc.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primers</th>
<th>Sequence</th>
<th>Amplicon Location</th>
<th>Amplicon Size (bp)</th>
<th>Ta (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>eNOS</td>
<td>Forward Rev–biotin</td>
<td>ADS251FP, ADS251RPB, ADS251FS1, ADS251FS2,</td>
<td>-235 to +29</td>
<td>256</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>Sequencing</td>
<td>ADS251FS3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CD31</td>
<td>Forward Rev–biotin</td>
<td>+102 to +366</td>
<td>266</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>Sequencing</td>
<td>ADS1723FP, ADS1723RPB, ADS1723FS1, ADS1723FS2,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ADS1723FS3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VE-cadherin</td>
<td>Forward Reverse-tailed Uni RPB</td>
<td>-241 to -74</td>
<td>170</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>Sequencing</td>
<td>ADS1886FP, ADS1886RPT, Uni RPB, ADS1886FS1,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ADS1886FS2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>vWF</td>
<td>Forward Reverse-tailed Uni RPB</td>
<td>+65 to +247</td>
<td>184</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>Sequencing</td>
<td>ADS1722FP, ADS1722RPT, Uni RPB, ADS1722FS1,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ADS1722FS2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VEGFR2</td>
<td>Forward Rev–biotin</td>
<td>-270 to -77</td>
<td>244</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>Sequencing</td>
<td>ADS1721FP, ADS1721RPB, ADS1721FS1, ADS1721FS2,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ADS1721FS3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>muLINE-1</td>
<td>For–biotin Reverse</td>
<td>abundant</td>
<td>113</td>
<td>50.5</td>
</tr>
<tr>
<td></td>
<td>Sequencing</td>
<td>ADS685FPB, ADS685RP, ADS685RS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>huLINE-1</td>
<td>For–biotin Reverse</td>
<td>~170</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sequencing</td>
<td>huLINE-1 FPB</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>huLINE-1 RP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>huLINE-1 RS</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.2.8 LINE-1

DNA was bisulfite converted as described above and 1-2ul (of the 10ul elution) was used for the LINE-1 PCR in a 50ul reaction: 1x PCR buffer, 1.5mM MgCl$_2$, 200uM each dNTP, 3pmol of Forward-biotinylated primer, 3 pmol of Reverse primer and 1 unit of Platinum Taq DNA polymerase (Invitrogen). Thermocycler conditions are as follows: 95ºC 15 min; 45 x (95ºC 30 s; annealing temp 30 s; 72ºC 30 s); 72ºC 5 min. PCR products run on a 1% agarose-TAE gel yield many products represented by a ladder-like smear. PCR products are further analyzed by pyrosequencing to analyze the following bisulfite converted sequence 5'-YGTTYGATTYGAGATTYGAGTTTYG-3' using a SNP analysis of the 5 CG sites (GenBank M13002, genomic sequence: 5’-CGCTCGACTCGAGACTCGAGCCCCG-3’) with a PSQ96HS instrument (EpigenDx Inc, Worcester, MA, USA). LINE-1 pyrosequencing primers were designed by EpigenDx (Worcester, MA, USA) and are listed in Table 6.

2.2.9 LUMA

Genomic DNA (500ng) was digested with *HpaII+EcoRI* and *MspI+EcoRI* (5U/reaction, New England Biolabs) in two separate 20ul reactions for 4 hours at 37ºC. Digests were subsequently inactivated for 20 minutes at 65ºC. Fragment ends were sequenced using luminometric polymerase extension with a PSQ96HS instrument (EpigenDx Inc, Worcester, MA, USA). The unmethylated fraction is calculated by comparing the ratio of the light intensities emitted from the addition of dCTP and dGTP between reactions containing *HpaII* and *MspI*. The *HpaII/Mspl* ratio was calculated as ($HpaII/EcoRI$)/($MspI/EcoRI$). The methylated fraction is calculated by:

%methylation = 1 - $HpaII/Mspl$. Templates with varying levels of CG methylation (0, 50 and 100%) were used as controls.

2.2.10 Statistics

All experiments were performed at least three times (unless stated otherwise) and data is presented as the mean ± standard error of the mean (SEM). Statistical tests were performed using a Student’s *t* test or analysis of variance (ANOVA). A *p* value less than 0.05 was considered to be statistically significant.
**Human vs. mouse eNOS/NOS3 promoter alignment**

**Figure 13: Multiple sequence alignment between human and mouse genomic DNA at the eNOS promoter.** ClustalW2 was used to do a cross-species promoter DNA alignment. Genomic location of primers are underlined (human) or highlighted in grey (mouse). CG are highlighted in yellow, * indicates a conserved nucleotide. OS = outer sense, OAS = outer antisense, IS = inner sense, IAS = inner antisense.
Human vs. mouse eNOS/NOS3 promoter alignment (region surrounding the SSRE)

**Figure 14**: Multiple sequence alignment between human and mouse genomic DNA at the eNOS promoter near the shear stress response element (SSRE). ClustalW2 was used to do a cross-species promoter DNA alignment. Genomic location of primers are underlined (human) or highlighted in grey (mouse). CG are highlighted in yellow, * indicates a conserved nucleotide. OS = outer sense, OAS = outer antisense, IS = inner sense, IAS = inner antisense.
Figure 15: Multiple sequence alignment between human and mouse genomic DNA at the CD31 exon1. ClustalW2 was used to do a cross-species promoter DNA alignment. Genomic location of primers are underlined (human) or highlighted in grey (mouse). CG are highlighted in yellow, * indicates a conserved nucleotide. OS = outer sense, OAS = outer antisense, IS = inner sense, IAS = inner antisense.
Human vs. mouse VE-cadherin/CDH5 promoter alignment

**Figure 16: Multiple sequence alignment between human and mouse genomic DNA at the VE-cadherin promoter.** ClustalW2 was used to do a cross-species promoter DNA alignment. Genomic location of primers are underlined (human) or highlighted in grey (mouse). CG are highlighted in yellow, * indicates a conserved nucleotide. OS = outer sense, OAS = outer antisense, IS = inner sense, IAS = inner antisense.
**Human vs. mouse vWF promoter alignment**

- **Human**
  - TTGGAGGTACTTCTTATAATACCATTCTTTTTCATTTTGTAAATTAAAAGGAGGC
  - ACCATTTCCTTTCATTGTTTCTTTTGGTAATTAAAA
  - GGAGGC
  - OS

- **Mouse**
  - CTGGAAGGGCT
  - CG
  - ACTACTGCTTCCTGCCATTGTTTCC
  - CG
  - CTGGTAATTAAGAGAAGGC
  - IS>

- **Human**
  - TGTTCAGAGTTGGGCTTCTTCATCCTCCCAGAGCGAGCTTTCACAGCTACTAGCTCCGGCTGGCTACAT
  - -GCCACAGCTTGGGCTTCTTCATCCTCCCAGAGCGAGCTTTCACAGCTACTAGCTCCGGCTGGCTACAT +98

- **Mouse**
  - AACAGCAAGACAGTCGGAGCTGTAGCAGACCTGATTG
  - GACCTATTGGAGCATTTGCA
  - ++

**Figure 17: Multiple sequence alignment between human and mouse genomic DNA at the vWF exon1.** ClustalW2 was used to do a cross-species promoter DNA alignment. Genomic location of primers are underlined (human) or highlighted in grey (mouse). CG are highlighted in yellow, * indicates a conserved nucleotide. OS = outer sense, OAS = outer antisense, IS = inner sense, IAS = inner antisense.
Human vs. mouse KDR/VEGFR2 promoter alignment

<table>
<thead>
<tr>
<th>Human</th>
<th>Mouse</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGGAGGACCGGTGCTCTTCGAGCGCTCTGTGGATGCTCTCCCCAAATTTGCGGGGACC</td>
<td>TGGAGGACCGACGTGGGGGCTC---GCTCTGTGGATGCTCTCCCCAAATTTGCGGGGACC</td>
</tr>
<tr>
<td><strong>18</strong></td>
<td><strong>18</strong></td>
</tr>
</tbody>
</table>

Figure 18: Multiple sequence alignment between human and mouse genomic DNA at the VEGFR2 promoter. ClustalW2 was used to do a cross-species promoter DNA alignment. Genomic location of primers are underlined (human) or highlighted in grey (mouse). CG are highlighted in yellow, * indicates a conserved nucleotide. OS = outer sense, OAS = outer antisense, IS = inner sense, IAS = inner antisense.
### Human vs. mouse Nanog promoter alignment

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Human</strong></td>
<td>TTAAAGTTTTATCCCATCTTCTGTTATTTAAAGTTGGAAA</td>
<td>-207</td>
</tr>
<tr>
<td></td>
<td>CTAGAGATC--GCCAGGGTCTG--GAAGTACGACGCTGGTT--AAGAGATG--ATTTAAGATG</td>
<td>-282</td>
</tr>
<tr>
<td><strong>Mouse</strong></td>
<td>GAAACCTAG----AAGTATTTGTT-----GCTTGGTGGTCTCTAGGTTCT--GTTGCTC--</td>
<td>-159</td>
</tr>
<tr>
<td></td>
<td>GAAATGAGGTAAAGCTTCTTTTTGGGGGGGGGAGATCTTTTAGACAGAGAGGATGGCGGCC</td>
<td>-222</td>
</tr>
<tr>
<td><strong>Human</strong></td>
<td>--GTTTTTCTAGTTCCACCTAGTCCTGTTACTCTGAGCTACTTTTTGGATTACATAG</td>
<td>-103</td>
</tr>
<tr>
<td><strong>Mouse</strong></td>
<td>TAAGCTTTCCC--CTCCCTCCAGCTCTGTCACCTTACGCATTCAATAGGT</td>
<td>-165</td>
</tr>
<tr>
<td><strong>Human</strong></td>
<td>GCCTTGTTGAGACTGTAGACGTAGACGAGAAATCTACGAGGATCTGTTGACTGAGGGGAG</td>
<td>-46</td>
</tr>
<tr>
<td><strong>Mouse</strong></td>
<td>TCCATGGTGGACCTTCGAGGGATCTAAGTGACCTACAGGGGTGGTGGGCTG</td>
<td>-222</td>
</tr>
<tr>
<td><strong>Human</strong></td>
<td>GGG------GTGTGGCAGCCAGGA----GGGGTGAGGTGTTCTAG--GTTGGCTAG--</td>
<td>-9</td>
</tr>
<tr>
<td><strong>Mouse</strong></td>
<td>GTGCCCTGCTGGCTTGGGAAGATCGGGCAGCTGAACTTCTCTGACGACG</td>
<td>-46</td>
</tr>
<tr>
<td><strong>Human</strong></td>
<td>---TAGAGCCCTCATTAAACTCTAGAGACTCCAGGATTTTAAATCGTTGCTGACTGAGG</td>
<td>+49</td>
</tr>
<tr>
<td><strong>Mouse</strong></td>
<td>GAGCAGGACCTCTCTCTTAATCTATCATGAGTGAGGATGGTGCTGGGCCGAGCTTCGAT</td>
<td>+13</td>
</tr>
<tr>
<td><strong>Human</strong></td>
<td>CTCTGTTGCC--CTATGTTTATATT-GTGGAGG----CAACTCAGTTTATCCTCAGAGAAGCTTGG</td>
<td>+100</td>
</tr>
<tr>
<td><strong>Mouse</strong></td>
<td>TTGTTGTTGTCTGGCTCTTTTTCTGAGGGCTGGCAGCTCCTTCTCAGGGGAG</td>
<td>+73</td>
</tr>
<tr>
<td><strong>Human</strong></td>
<td>ATACTTTTCTCTTG----AGGTCTTATTTCTCTTCAATAGCAGCTTCACTACGAGCACTTACGAGT</td>
<td>+149</td>
</tr>
<tr>
<td><strong>Mouse</strong></td>
<td>ATAAATTTTGGCATAGCATTAAACGCTCTCCTTACTGCTTACTGATCCTGTGATGCTGGAGGAGA</td>
<td>+133</td>
</tr>
<tr>
<td><strong>Human</strong></td>
<td>CTTAAAAGCTG--CCTTAAACCTTTTCTTTCAAGTCACCCCTCCTCACTTTGTCTTCCTTCTTCTTCT</td>
<td>+206</td>
</tr>
<tr>
<td><strong>Mouse</strong></td>
<td>TTTTGGTTGTTCTGCTGCTTTTAAGACATCACTCTCTCTCCCTACAGCTTCTTCTTCTTCTTCT</td>
<td>+179</td>
</tr>
<tr>
<td><strong>Human</strong></td>
<td>CTTATACTACATGACCTGAGTGAGTGGATCTCAGGATCTCAGGCTTGCTGTCAGGAA----A</td>
<td>+263</td>
</tr>
<tr>
<td><strong>Mouse</strong></td>
<td>ATG----ACACTGACATGAGTGGAGTCCTTGCTCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT</td>
<td>+238</td>
</tr>
<tr>
<td><strong>Human</strong></td>
<td>GCAATGACGTCTTAAGAAGCTAGCACTCTAGTGGATTCTGAGAGAGAAGGAAATATGCACTCTG</td>
<td>+323</td>
</tr>
<tr>
<td><strong>Mouse</strong></td>
<td>GCAGATGAAATTTTCTGGGAAGAAGCCTGCTGACTGAGGATTTTTCTTGTCTTCTGAGGAGG</td>
<td>+295</td>
</tr>
</tbody>
</table>

**Figure 19:** Multiple sequence alignment between human and mouse genomic DNA at the Nanog promoter. ClustalW2 was used to do a cross-species promoter DNA alignment. Genomic location of primers are underlined (human) or highlighted in grey (mouse). CG are highlighted in yellow, * indicates a conserved nucleotide. OS = outer sense, OAS = outer antisense, IS = inner sense, IAS = inner antisense.
2.3 Results

2.3.1 LINE-1 hypomethylation in mES cells and hypermethylation in all other mouse cell types

The long interspersed element (LINE-1) assay for global DNA methylation relies on the repression of interspersed and abundant repetitive elements through DNA methylation. Genomic DNA was sodium bisulfite converted, PCR amplified and pyrosequenced using a SNP analysis at 5 CG sites. The mouse ES cells deficient in Dnmt1, 3a and 3b or triple knockout cells (TKO) show <5% LINE-1 methylation. Wild type mouse ES cell lines E14 and R1 show 30-40% methylation while the line J1 shows ~10-15% methylation at LINE-1 elements (Fig. 20A). As ES cells differentiate to EB days 4 and 7, they gain 5mC marks at these repetitive DNA sequences to 60-80%. EC and VSMC isolated from the mouse descending thoracic aorta show high (>80%) levels of DNA methylation at LINE-1 repeats.

In the luminometric methylation assay (LUMA), isoschizomers (HpaII and MspI) are used to estimate global DNA methylation at the 5'-CmCGG-3' sequence. Genomic DNA was digested with EcoRI and HpaII or MspI and the fragment ends were sequenced by synthesis. Pluripotent mouse ES J1 and E14 cell lines show low levels of CmCGG methylation while TKO cells lack CmCGG methylation (Fig. 20B).

LINE-1 methylation was also assayed in two extraembryonic endoderm (XEN) and trophoblast stem (TS) cells isolated from E3.5 mouse blastocyst. Both extraembryonic cell types are highly methylated at LINE-1 (Fig. 21).

2.3.2 LINE-1 and CmCGG methylation does not change upon human ES cell differentiation

Two lines of human ES cells (H2 and H9) and terminally differentiated cell types (HUVEC, VSMC, hepatocytes and keratinocytes) were assayed for LINE-1 and CmCGG methylation (Fig. 22). Genome CG hypermethylation is observed in all cell types. H2 and H9 hES cells show comparable levels of global DNA methylation as in HUVEC, VSMC, keratinocytes and hepatocytes as detected by the LINE-1 and LUMA assays.
2.3.3 Increase in CG methylation at the Nanog promoter upon differentiation

The proximal promoter of Nanog is a poor CpG island (Table 1) and differentially methylated between expressing ES cells and non-expressing cells. Sodium bisulfite genomic sequencing was used to show that the Nanog promoter is unmethylated in ES cells and increases in methylation upon differentiation (Fig. 23A).

2.3.4 The VEGFR2 promoter is unmethylated in all cell types

The VEGFR2 gene has a strong CpG island promoter (Table 1) and a high ratio of the observed to expected CG sites based on nucleotide composition. Regardless of cell type or stage in the differentiation cascade, the proximal promoter region is completely unmethylated (Fig. 23B-C).

2.3.5 The promoters of eNOS, CD31, vWF and VE-cadherin are differentially methylated between mouse EC and VSMC

The promoter DNA methylation of EC-enriched genes eNOS, CD31, vWF and VE-cadherin was assayed. DNA was bisulfite converted, PCR amplified and pyrosequenced using a SNP analysis at CG sites. The promoters of the EC-enriched genes eNOS, CD31, vWF and VE-cadherin are methylated in murine VSMC and show little methylation in murine EC (Fig. 24A-D). Two-way ANOVA was used to determine that the two groups (EC vs. VSMC) are statistically different (*p<0.001). Sodium bisulfite genomic sequencing was used to confirm that the promoters of eNOS and VE-cadherin are entirely unmethylated in EC (Fig. 25).

2.3.6 The promoters of EC-enriched genes are methylated in mouse ES cells

PCR product pyrosequencing was used to quantify the degree of eNOS, CD31, VE-cadherin and vWF promoter methylation in ES cells (Fig. 26A-D). On average, the promoters of these genes have 40-80% methylation in E14 and R1 mouse ES cells.

2.3.7 The promoters of eNOS, CD31 and VE-cadherin are heterogeneously methylated in ES cells

Sodium bisulfite genomic sequencing at the eNOS promoter was used to confirm the findings of pyrosequencing and, to examine the DNA methylation pattern at a higher resolution. The eNOS promoter is heterogeneously methylated in mouse and human ES cells (Fig. 27). ES cells have a random CG methylation pattern between individual DNA strands. The frequency
histogram shows is a wide distribution in the number of CG sites methylated and their abundance in the sample. The promoters of CD31 and VE-cadherin also show a heterogeneous pattern of DNA methylation (Fig. 28 A-B). In contrast, the vWF promoter is highly methylated in ES cells (Fig. 28C).

2.3.8 The eNOS promoter is heterogeneously methylated in extraembryonic cells

Extraembryonic endoderm (XEN) and trophoblast (TS) stem cells were derived from E3.5 mouse. Syncytiotrophoblast cells from the first trimester were collected from placentas of healthy pregnancies using a floating villus technique. Sodium bisulfite genomic sequencing was used to show that all 3 extraembryonic cell types contain a heterogeneously methylated eNOS promoter (Fig. 29).

2.3.9 Changes in promoter DNA methylation of eNOS, CD31, VE-cadherin and vWF during in vitro differentiation

The promoters of eNOS, CD31, vWF and VE-cadherin were assayed at an “early” stage in EC-differentiation, the EB-day 4 VEGFR2 positive cells and in non-EC, the EB-day 4 VEGFR2 negative cells using PCR product pyrosequencing. Both groups showed high (> 60%) levels of methylation at all CG sites (Fig. 30 A-D). The degree of methylation overlaps and no difference is observed between the two groups. Sodium bisulfite genomic sequencing confirms that individual DNA strands in each of the four promoters are heavily methylated (Fig. 31 A-D). Additionally, promoter DNA heterogeneity is still observed 4 days after the differentiation of ES cells.

At a “later” stage in EC-differentiation, the EB-day 7 CD31 positive cells show a large reduction in CD31 promoter DNA methylation in comparison to the non-EC, the EB-day 7 CD31 negative cells. A moderate degree of VE-cadherin promoter DNA demethylation is observed. Though the methylation of eNOS is significantly lower in EB-day 7 CD31-positive versus negative cells, the differences are modest. Importantly, the eNOS promoter is still heavily methylated in EB-day 7 CD31 positive cells compared with murine aortic EC where there is no DNA methylation (Fig. 32). Two-way ANOVA was used to determine that the two groups (EB-day 7 CD31 positive vs. EB-day 7 CD31 negative) are statistically different for all 4 genes studied (*p<0.001).
Figure 20. LINE-1 is hypomethylated in mouse ES cells and hypermethylated in all other cell types. A) LINE-1 methylation assays 5 CGs across thousands of copies of the repetitive element. Terminally differentiated EC and VSMC show high degree of methylation while pluripotent ES cells have low levels of LINE-1 methylation. B) LUMA data was generated from pyrosequencing to fill in fragment ends from MspI and HpaII digests. Totipotent mouse ES cells show low levels of CCGG methylation. Data is presented as the mean ± SEM (n=3-4), E13, R1, J1 wild type mES cells and TKO (Dnmt1,3a,3b-null mES cells), n=1.
Figure 21. LINE-1 hypermethylation in mouse extraembryonic cells. LINE-1 methylation assays 5 CGs across thousands of copies of the repetitive element. Mouse TS and XEN cells are more highly methylated than ES cells. TS = trophoblast stem cells, XEN = extraembryonic endoderm stem cells.
Figure 22. LINE-1 and CmCGG methylation does not change upon human ES cell differentiation. A) LINE-1 methylation assays 4 CGs across thousands of copies of the repetitive element. Pluripotent human ES cell lines H2 and H9 and terminally differentiated HUVEC, VSMC, keratinocytes and hepatocytes show high degree of methylation. B) LUMA data was generated from pyrosequencing to fill in fragment ends from MspI and HpaII digests. Human ES cells and terminally differentiated HUVEC and VSMC show high degree of global DNA methylation. (n=1, LINE-1 and LUMA data for HUVEC, n=3)
Figure 23. Changes in promoter DNA methylation of Nanog and VEGFR2 during in vitro differentiation, and in terminally differentiated cells. Sodium bisulfite genomic sequencing indicates that A) the Nanog promoter is unmethylated in ES cells (grey bars) and increases upon differentiation. In contrast, B-C) the VEGFR2 promoter is unmethylated in all cell types. B) PCR product pyrosequencing and C) sodium bisulfite genomic sequencing at the VEGFR2 promoter (n=1).
Figure 24. Differential promoter DNA methylation of EC-enriched genes in mouse EC and VSMC. DNA was treated with sodium bisulfite, PCR amplified and sequenced at CG sites to estimate the average % methylation. A-D) The promoters of eNOS, CD31, VE-cadherin and vWF are unmethylated in EC (violet) and densely methylated in VSMC (black). Data is presented as the mean ± SEM (n=3-7). Two way ANOVA between cell types indicates that the groups are statistically different (*p<0.001).
Figure 25. Unmethylated DNA at the eNOS and VE-cadherin promoters in murine thoracic aortic EC. Sodium bisulfite genomic sequencing was used to confirm results from PCR product pyrosequencing, indicating that every CG site in the promoters of A) eNOS and B) VE-cadherin is unmethylated in endothelial cells. The arrow represents the transcriptional start site, eNOS n>3, VE-cadherin n=1.
Figure 26. Intermediate to high levels of promoter DNA methylation in murine ES cells. DNA was treated with sodium bisulfite, PCR amplified and sequenced at CG sites to estimate the average % methylation. Pyrosequencing of the PCR product shows that the promoters of A) eNOS, B) CD31, C) VE-cadherin and D) vWF are on average 40-80% methylated in murine ES cells. Data is presented as the mean ± SEM (n=3)
Figure 27. The eNOS promoter is heterogeneously methylated in mouse and human embryonic stem (ES) cells. The eNOS promoter shows DNA strands (single lines) with varying number of methylated CG in mouse (A) and human (B) ES cells. This data has been confirmed in multiple ES clones (E14, R1, H9). Sodium bisulfite genomic sequencing of n ≥15 DNA strands was used to assess methylation. Filled circles represent 5mC while open circles are unmethylated C, mES n=4, hES n=2.
Figure 28. The EC-enriched genes, CD31 and VE-cadherin are heterogeneously methylated in mouse ES cells. EC-enriched genes A) CD31 and B) VE-cadherin are heterogeneously methylated and shows DNA strands (single lines) with varying number of methylated CGs. In contrast, the promoter of C) vWF is heavily methylated and most strands have all four CG sites methylated. Sodium bisulfite genomic sequencing of n ≥15 DNA strands was used to assess methylation. Filled circles represent 5mC while open circles are unmethylated cytosines, n=3.
Figure 29. The eNOS promoter is heterogeneously methylated in mouse and human extraembryonic cells. The eNOS promoter shows DNA strands (single lines) with varying number of methylated CG. Sodium bisulfite genomic sequencing was used to assess methylation. Filled circles represent 5mC while open circles are unmethylated C. XEN and TS cells were derived from E3.5 mouse. Syncytiotrophoblast cells from the first trimester were collected from placentas of healthy pregnancies using a floating villus technique. All three cell types show that the eNOS promoter is heterogeneously methylated in extraembryonic cells. XEN = extraembryonic endoderm stem cells, TS = trophoblast stem cells.
Figure 30. Dense promoter DNA methylation of EC-enriched genes in EB-day 4 VEGFR2 -positive and -negative cells. DNA was treated with sodium bisulfite, PCR amplified and sequenced at CG sites to estimate the average % methylation. Pyrosequencing of the PCR product shows that the promoters of A) eNOS, B) CD31, C) VE-cadherin and D) vWF are heavily methylated in EB-day 4 VEGFR2 -positive and -negative cells. No difference is observed between cell types. Data is presented as the mean ± SEM (n=3-4).
Figure 31. A-B. Promoter DNA heterogeneity in EC-enriched genes in EB-day 4 VEGFR2-positive and negative cells. Sodium bisulfite genomic sequencing was used to confirm results from PCR product pyrosequencing, indicating that the promoters of A) eNOS and B) CD31 are heavily methylated in EB-day 4 VEGFR2-positive (left) and negative (right) cells. No difference is observed between cell types. Some promoter DNA mosaicism is evident, n=2.
Figure 31. C-D. Promoter DNA heterogeneity in EC-enriched genes in EB-day 4 VEGFR2-positive and negative cells. Sodium bisulfite genomic sequencing was used to confirm results from PCR product pyrosequencing, indicating that the promoters of C) VE-cadherin and D) vWF are heavily methylated in EB-day 4 VEGFR2-positive (left) and negative (right) cells. No difference is observed between cell types. Some promoter DNA mosaicism is evident, n=2.
Figure 32. Demethylation at the CD31 and VE-cadherin promoter in EB-day 7 CD31–positive cells. DNA was treated with sodium bisulfite, PCR amplified and sequenced at CG sites to estimate the average % methylation. A-D) The promoters of eNOS, CD31, VE-cadherin and vWF are hypermethylated in EB-day 7 CD31 (−) cells. Less methylation is observed in EB-day 7 CD31 (+) cells at the CD31 and VE-cadherin promoters. Data is presented as the mean ± SEM (n=3-4). Two way ANOVA between cell types indicates that the groups are statistically different (p<0.001).
2.4 Discussion

The main goal of this study was to determine the pattern of EC-enriched gene promoter DNA methylation as pluripotent cells differentiated to terminally differentiated, somatic cells. This is an important question for basic research and for potential applications in regenerative medicine and gene therapy targeting EC dysfunction. The major findings of the current thesis are: i) a conserved pattern of promoter DNA methylation of EC-enriched genes in human and mouse EC and VSMC, ii) heterogeneous promoter DNA methylation at promoters of EC-enriched genes in multipotent cells and iii) CD31 promoter DNA demethylation in EB-day 7 CD31-positive cells.

LINE-1 hypomethylation in mouse ES cells and hypermethylation in all other cell types

It is thought that ES cells have comparable levels of global 5mC as somatic cells and that this does not change upon differentiation\(^{32,33,58}\). However, variability exists between cell and tissue type and also with the method of quantification. Genome-wide CG methylation is estimated at 50-70% in mouse ES cells and 65-85% in differentiated cells and tissues\(^{32,33,38,41}\). Only a small percent of the total CG sites change their methylation status upon differentiation: 8% become methylated, while 2% demethylate, corresponding with cell type-specific gene repression and activation\(^{32}\).

LINE-1 methylation was assayed to estimate global changes in DNA methylation. Since LINE are abundant and distributed throughout the genome, it is a good “random sampling” to estimate global DNA methylation. With the exception of LINE-1 methylation in mouse ES cells, all other cell types, mouse and human, are hypermethylated at this repetitive sequence. Additionally, assaying CmCGG in pluripotent and differentiated human cell lines confirms the finding that global CG methylation does not change as cells differentiate.

The LINE-1 methylation is a poor representative of global DNA methylation in multipotent cell types since LINE-1 are hypomethylated, have retrotransposition activity and do not show the same extent of hypermethylation as other repeat types in ES cells\(^{39,155,156}\). For example, LINE-1 5mC levels are ~40-50% in mES cells while other repetitive regions of the genome show ~80% methylation in ES cells\(^{23,39}\). An additional assay at an alternative repetitive element is required to confirm that global CG methylation levels do not change as mouse ES cells differentiate.
Increase in CG methylation at the Nanog promoter upon differentiation

Nanog promoter DNA methylation is a widely used marker of pluripotency, epigenetic reprogramming and extent of ES cell differentiation\textsuperscript{61-63}. The proximal promoter region of the Nanog gene contains a differentially methylated region that is unmethylated in expressing ES cells and methylated in EB-day 5 cells and in non-expressing cells\textsuperscript{61-63}. Likewise, we show that the Nanog promoter is unmethylated in ES cells and methylated in all other cell types, which is in agreement with studies showing an increase in promoter methylation upon ES differentiation.

Promoter DNA methylation of EC-enriched genes is conserved between the mouse and human

We have shown that transcription of the eNOS gene is regulated by epigenetic pathways\textsuperscript{110,112}. Recent studies in our lab have shown that eNOS and other EC-enriched genes contain regions of differential DNA methylation in human EC versus VSMC (Shirodkar et al. in prep). This study extends the findings of differential promoter DNA methylation in EC-enriched genes to the mouse by assaying DNA methylation of the homologous regions in murine EC and VSMC isolated from the descending thoracic aorta. Promoter DNA methylation of tissue-specific genes is highly conserved between human and mouse tissues\textsuperscript{157,158}. Similarly we have shown that the EC-enriched transcripts eNOS, CD31, VE-cadherin and vWF have hypomethylated promoters in EC while non-expressing murine VSMC are hypermethylated. These findings indicate that some genes with cell type-specific patterns of expression and differential promoter DNA methylation are conserved between species. Parallel with our findings in human cells, we argue that DNA methylation likely has a role in the transcriptional regulation of EC-enriched genes in the mouse.

The promoters of EC-enriched genes are methylated in mouse ES cells

In several genomic studies of gene promoters in ES cells, CG tend to have a bimodal distribution and are either methylated or unmethylated\textsuperscript{32,59,60}. Many pluripotency-associated, housekeeping and developmentally important genes have strong CpG island promoters and are unmethylated in ES cells\textsuperscript{32,59,60}. We have shown that the strong CpG island promoter of VEGFR2 is unmethylated in all cell types examined. In contrast, many tissue-specific genes have methylated promoters in ES cells\textsuperscript{32,59,60}. Here we show that the promoter regions of the EC-enriched genes eNOS, CD31, VE-cadherin and vWF have 40-80% CG methylation in ES.
cells. This pattern does not follow the bimodal paradigm: it is not completely methylated or unmethylated.

The promoters of EC-enriched genes are heterogeneously methylated in ES, EB-day 4 and in extraembryonic cells

We assayed the promoters of EC-enriched genes using high-resolution sodium bisulfite sequencing. We found that the eNOS promoter shows a heterogeneous pattern of DNA methylation in human and mouse ES cells and in cells of extraembryonic origin. Additionally, we found that EB-day 4 VEGFR2-positive and -negative cells also show a mosaic pattern of eNOS promoter DNA methylation. We extended these findings with eNOS to other EC-enriched genes by showing that the promoters of CD31 and VE-cadherin are also heterogeneously methylated in ES, EB-day 4 VEGFR2-positive and -negative cells. Each of these cell types is multipotent, and thus has the ability to differentiate towards several lineages. ES cells can become cells of any embryonic germ layer. Trophoblast stem cells differentiate to syncytiotrophoblast and cytotrophoblast cells, while VEGFR2-positive cells have the potential to become EC, VSMC or haematopoietic cells\(^{140}\). We believe that promoter DNA heterogeneity at cell-type specific genes maybe a signature of multipotency. Our findings have been confirmed in multipotent mesenchymal stem cells that show DNA heterogeneity at promoters of lineage restricted genes\(^{159}\).

Promoter DNA heterogeneity may be \(i\) established by the activity of Tet proteins and \(ii\) a result of dynamic pattern of DNA methylation in ES cells. ES cells show abundant Tet1 and Tet2 activity and are enriched in 5hmC\(^ {40,44}\). The factors regulating Tet activity and 5mC substrate selectivity are currently unknown. We hypothesize that the 5mC present in EC-enriched gene promoters in ES cells is dynamically converted to 5hmC by Tet1 and Tet2. 5hmC and Tet activity has been implicated in DNA demethylation\(^ {57}\). The activity of Tet proteins on 5mC at promoters of EC-enriched genes may result in the conversion of some 5mC to 5hmC and eventually lead to DNA demethylation at certain CG sites. Further work is required to demonstrate that Tet proteins are responsible for establishing promoter DNA heterogeneity.

Additionally, our data with promoter DNA heterogeneity in ES and other multipotent cells questions the static and permanent nature of DNA methylation. ES cells maybe re-setting the DNA methylation code with every cell division. The clonal expansion of a single ES cell will generate one of two outcomes. The first outcome is the recapitulation of the EC-enriched promoter DNA heterogeneity we observed in a dish of ES cells. Each clonal colony would
contain greater than three different alleles at the eNOS promoter and therefore be different from the founder population. The second outcome is that clones of single ES cells would be stably inherited from the founding cell. A heterogeneous pattern in the clones implies that the establishment and/or inheritance of DNA methylation is dynamic and can re-set with mitotic division. In contrast, inheritance of 1-2 alleles at the eNOS promoter per clonal colony implies that the DNA methylation code is identically copied from mother to daughter cell. We think that the DNA methylation code at promoters of EC-enriched genes in multipotent cells maybe re-setting with mitotic division.

Another unusual feature of ES cell gene regulation is the bivalent chromatin domains which are often found at tissue-specific genes\(^{65,67}\). These domains are marked by H3K4me3 and H3K27me3, which are associated with transcriptional activation and repression, respectively\(^{67}\). Bivalent chromatin domains were recently identified in hemangioblasts\(^{69}\) and extraembryonic stem cells\(^{151}\) and may, generally, be a feature of multipotent cells. They have been implicated in keeping transcription of cell- or tissue-specific genes off or at low levels in ES cells, and keeping the cells in a poised state for activation upon differentiation\(^{65,67}\). Promoter DNA heterogeneity at EC-enriched genes found in multipotent cells may be correlated with bivalent chromatin domains. Taken together, promoter DNA heterogeneity may be a feature of “stemness”. Further work is needed to correlate this phenomenon with bivalent chromatin.

**Changes in promoter DNA methylation of eNOS, CD31, VE-cadherin and vWF during *in vitro* differentiation**

Mouse ES cells were *in vitro* differentiated to all embryonic germ layers by the removal of LIF and formation of spherical aggregates, or embryoid bodies. EB are used to model the early embryo\(^{136}\). Clusters of cells with co-localization of VE-cadherin, CD31 and VEGFR2 are visible early in EB development, while cord-like structures become evident as EB mature\(^{100,131,137}\). The *in vitro* method of differentiation is not identical to the *in vivo* environment. Though similarities are evident between EB and the embryo, there is a lack of cell organization, tissue and organ architecture. As a result, differentiation pathways will be affected due to lack of external signaling such as biomechanical forces and paracrine signaling.

In our system, cells were sorted for cell surface VEGFR2 four days and for CD31 seven days after initiation of differentiation to select for “early” and “late” endothelial progenitor cells, respectively. VEGFR2 mRNA is evident in EB-day 3-4. By EB-day 4, we also observe an increase in steady state VEGFR2 mRNA. EB-day 4 VEGFR2 positive cells differentiate into
hematopoietic, endothelial and smooth muscle cells\textsuperscript{140}, while EB-day 4 VEGFR2 negative cells never differentiate into EC. In another study, EB-day 8 CD31-positive cells have an endothelial-like behavior through the incorporation of acetylated low-density lipoprotein, tube formation on matrigel and expression of EC-enriched mRNA and proteins VE-cadherin, vWF and ICAM2\textsuperscript{137}. By EB-day 7, we also show robust expression of VE-cadherin. It should be noted that the increase in eNOS, CD31 and Tie2 are minor compared to the >250-fold increase in VE-cadherin mRNA.

We noticed that the level of 5mC increased from ES cells to EB-day 4 cells at the promoters of eNOS, CD31 and VE-cadherin. This may be due to the decrease of Tet proteins that is observed ES cells differentiate\textsuperscript{40}. As a result of lower Tet activity, fewer 5mC gets oxidized to 5hmC, causing an increase in 5mC.

It is interesting to note that two of the promoters, CD31 and VE-cadherin have undergone DNA demethylation in EB-day 7 CD31 positive cells. EB-day 7 cells selected for cell surface expression of CD31 have a hypomethylated CD31 promoter. This confirms that DNA methylation at the promoter is a functional mechanism of regulating CD31 gene expression. The transcription of VE-cadherin is largely driven by the transcription factor Erg which is expressed in day 4 EB\textsuperscript{160}. The knockdown of Erg during differentiation results in reduced vasculature formation, implying an essential role in the differentiation into vascular EC\textsuperscript{160}. Although we observe a massive increase in VE-cadherin steady state mRNA by EB-day 7, it is interesting to note that the VE-cadherin promoter DNA still has \textasciitilde 40\% methylation in EB-day 7 CD31 positive cells.

Although eNOS and vWF promoter DNA is significantly lower in EB-day 7 CD31 positive cells versus negative cells, the differences are modest and the promoters are still hypermethylated relative to mouse EC. We have shown \textit{in vivo} that eNOS is a late onset gene and is transcribed in the mouse after the onset of robust unidirectional blood flow\textsuperscript{97}. Additionally, we have shown links between shear stress and eNOS promoter DNA methylation. Namely, EC exposed to disturbed blood flow in the high probability (HP) region of the mouse aortic arch show higher levels of eNOS promoter methylation, which correlates with lower eNOS promoter activity\textsuperscript{122}. We have shown that mutating the shear stress response element (SSRE) in the eNOS promoter causes transcriptional repression through promoter DNA hypermethylation (unpublished data). Since the \textit{in vitro} differentiation took place in the EB which are not exposed to fluid forces, shear stress may be required to drive EB-day 7 CD31 positive cells to the EC lineage, and to cause
demethylation at the eNOS promoter. Further experimental work is required to test the hypothesis that biomechanical forces will result in EC-enriched gene promoter demethylation.

EB-day 4 VEGFR2–negative and EB-day 7 -negative cells develop into non-EC and maintain the hypermethylation status at eNOS, CD31, VE-cadherin and vWF promoters. This implies that the steady state mRNA of EC-enriched genes observed at EB-day 7 is likely restricted to EB-day 7 CD31 positive cells.

2.5 Limitations

The findings described in this study are based on a population of cells differentiating from ES cells to EB-day 4 and EB-day 7 cells. The average DNA methylation patterns described may not represent how individual cells act in the population. For example, an EB-day 7 cell that is CD31-positive, expresses CD31, has an unmethylated CD31 promoter may not be expressing eNOS, VE-cadherin and vWF.

Though it is not addressed in our work, it is likely that cell-cell interactions and signaling affect the differentiation program. ES cells differentiated on a 2D surface, attached to a substrate, differentiate into a limited number of hematopoietic cells implying that the 3D environment is important. Likewse, is it not clear if EB-day 4 VEFR2-positive and EB-day 7 CD31-positive cells have an inherent ability to differentiate towards the EC lineage or do so because of external factors. For example, VE-cadherin acts as a negative regulator for VEGFR2 receptor internalization.

Our study of promoter DNA methylation relied on estimating CG methylation by assaying small regions of the gene. There are obvious limitations to this type of assay and can be extended and confirmed by using genomic arrays that span a larger region.

2.6 Conclusions

This study has provided evidence for the conservation of promoter DNA methylation patterns of EC-enriched genes between mouse and human. Using hanging drops to differentiate ES cells, followed by cell culture and FACS, we showed that EB-day 7 CD31 positive cells demethylate at the CD31 promoter. Further studies are required to evaluate if mechanical forces can drive EB-day 7 CD31 positive cells to demethylate at the eNOS promoter. We have also shown DNA heterogeneity at promoters of lineage-restricted genes multipotent cells, namely ES, extraembryonic stem and EB-day 4 VEGFR2-positive and -negative cells. Further studies are
required to evaluate if promoter DNA heterogeneity at cell type-enriched genes is stably or dynamically inherited.

2.7 Future Directions

Driving cells with endothelial cell-potential to terminal differentiation using physical forces

Is fluid shear stress the missing factor in driving EB-day 7 CD31-positive cells to terminally differentiate to vascular EC? A parallel plate apparatus can be used to expose cells to the same magnitude that is found in large arteries, i.e. 15 dyn/cm$^2$ for 24-48hr. Cells grown under static conditions can serve as a control. eNOS promoter DNA demethylation and increased steady state mRNA can provide evidence for any commitment that the sheared cells have made toward the EC lineage.

Is the promoter DNA methylation pattern locked in ES cells, or is it dynamic?

To determine if EC-enriched gene promoter DNA methylation patterns are fixed or in flux, culture single ES cells (limited dilution or FACS sorted) and compare ES cell colonies to a dish (with >10$^6$) of ES cells. If clones are heterogeneous, then the ES cells are in flux and this implies a dynamic mode of DNA methylation inheritance. If clones are homogeneous, but every colony is different, then ES cell promoter methylation is fixed and copied to daughter cells.

Is promoter DNA heterogeneity at lineage-restricted genes functional? Is it correlated with chromatin bivalency and 5hmC enrichment?

To ask whether heterogeneous promoter DNA is functional, a series of experiments can be performed:

1. ES cells deficient in all three Dnmts (Dnmt1,3a,3b-null TKO cells) lack methylation. Yet, they proliferate normally as pluripotent cells$^{39}$. They can be cultured and then assayed for the gene expression of eNOS, CD31, VE-cadherin, vWF and VEGFR2. An increase in the steady state mRNA compared to J1 wild type ES cells will indicate that the heterogeneous or hypermethylated pattern at the promoter is maintaining the genes in a repressed state.
2. Perform a chromatin immunoprecipitation (ChIP) in mouse and/or human ES cells for H3K4me3 and H3K27me3 at the promoters of eNOS, CD31, VE-cadherin and vWF. Perform MeDIP and hydroxyMeDIP at the same promoters to evaluate if heterogeneous promoter DNA methylation is correlated with bivalent chromatin and 5hmC enrichment.

Functional assays with EB-day 7 CD31-positive cells

Characterization of EB-day 7 CD31-positive cells would provide clues about the level of “maturity” and equivalent in vivo location of these in vitro-derived cells. Mature mouse arterial, venous, and microvascular endothelial cells or equivalent cell lines could be used as controls. Cells could be: 1) immunostained for CD31, VE-cadherin and eNOS, 2) tested for acetylated low density lipoprotein (LDL) uptake and 3) tube formation on matrigel.

Which Dnmt is responsible for the methylation of EC-enriched gene promoters in ES cells?

DNA methyltransferases 1, 3a and 3b are expressed in ES cells. This aim will show whether one or multiple Dnmt are responsible for the methylation of EC-enriched gene promoters in ES cells. Murine Dnmt1 -/- ES cells24, Dnmt3a -/-, Dnmt3b -/-, and Dnmt3a,3b -/- ES25 cells will be cultured. Genomic DNA will be sodium bisulfite converted and analyzed at promoter using PCR product pyrosequencing. Promoter hypomethylation in Dnmt-null ES cells compared to J1 mouse ES cells (wild-type control) will be indicative of enzyme activity. The promoter of vWF will act as a particularly good control for the loss of methylation due to its hypermethylation in ES cells.

Is the promoter DNA heterogeneity found in ES cells DNA-replication dependent?

Primary human cells are harvested one day post-confluence for epigenetic studies to ensure low rates of cell division and sufficient time for the epigenetic code to be replicated by DNMT1 into daughter cells. In contrast, ES cells are asynchronous, constantly dividing, spontaneously differentiate at high densities and have active Dnmt1, Dnmt3a and Dnmt3b enzymes. It is possible that the heterogeneity found at promoters of EC-enriched genes is due to the incomplete addition of methylation marks to nascent DNA. To show that promoter DNA heterogeneity is not replication dependent, ES cells need to be pulse-labelled with bromodeoxyuridine (BrdU) to label newly synthesized DNA, fixed and stained with propidiu
iodide (PI) to label nucleic acid. Subsequently, equal numbers of cells will be FACS sorted into 6 cell cycle fractions: G1, S1, S2, S3, S4 and G2 based on PI staining, where the G2 cells contain twice the DNA content of G1 cells (G is the growth phase and S is DNA synthesis phase). DNA will be extracted, BrdU-immunoprecipitated and quantified at the eNOS promoter by using qPCR with genomic primers for the eNOS promoter.
References


