Thesis for the Master's degree in Molecular Biosciences
Main field of study in Molecular Biology

**Anita Løvstad Sørensen**

# DNA methylation patterns of mesenchymal and hematopoietic stem cells

60 study points

**Department of Molecular Biosciences**
Faculty of mathematics and natural sciences
**UNIVERSITY OF OSLO 06/2008**

# TABLE OF CONTENTS

## SUMMARY

Mesenchymal stem cells (MSCs) have been identified in several adult tissues and have the ability to differentiate into multiple cell types. However, not all MSC types differentiate efficiently into all lineages. We show here that there is an epigenetic basis for this restricted differentiation capacity. DNA methylation was determined on lineage-specific promoters by bisulfite sequencing, and genome-wide by methyl DNA immunoprecipititation (MeDIP) with promoter array hybridization in mesenchymal stem cells (MSCs) of various tissues. Bisulfite sequencing shows that the adipogenic *FABP4* and *PPARG2* promoters are differentially methylated in ASCs and bone marrow (BM) MSCs relative to Wharton's jelly (WJ) MSCs and muscle progenitor cells (MPCs). In contrast, the myogenin (*MYOG*) promoter is hypomethylated in WJMSCs and MPCs relative to ASCs and BMMSCs. In hematopoietic stem cells (HSCs), all promoters are hypermethylated. Differential methylation correlates with distinct differentiation capacities: ASCs and BMMSCs differentiate efficiently into adipocytes but not into multinucleated myogenin positive myocytes, whereas MPCs display poor adipogenic differentiation but robust myogenic capacity. The endothelial *CD31* gene is methylated in ASCS, BMMSCs, WJMSCs and MPCs, in agreement with their poor endothelial differentiation potential. *CD31* is however unmethylated in HSCs, in which its expression can be induced. Methylation patterns in adipocytes, muscle and endothelial cells argue that ASCs (and BMMSCs) are epigenetically pre-programmed for adipogenesis, while MPCs have a methylation pattern predictive of myogenic potential. Bisulfite sequencing corroborates genome-wide methylation profiling. MeDIP reveals similarity in methylation profiles between MSCs from adipose tissue, bone marrow and muscle, reflecting the mesodermal origin of these cells. Our results put forward the hypothesis that DNA methylation patterns on lineage-specific promoters may predict lineage differentiation capacity.

# INTRODUCTION

## 1. STEM CELLS

Stem cells are unspecialized cells that have the ability to self-renew (self-renewal capacity) and to give rise to at least one more committed progenitor or differentiated cell type (differentiation capacity). Stem cells can be classified according to their differentiation potential as totipotent, pluripotent, multipotent and unipotent; this differentiation potential, or "stemness", is likely to exist as a continuum, or gradient, in the organism or tissue in which they reside (**Fig. I-1**). In mammals, only the zygote and cells of the cleavage-stage embryo are totipotent (up to the 4-8 cell-stage) and can generate an entire organism, including extraembryonic tissues. Embryonic stem cells (ESCs) are derived *in vitro* from culture of the inner cell mass, the group of cells that give rise to the embryo itself (Evans and Kaufman, 1981; Thomson *et al.*, 1998). ESCs are pluripotent and can give rise to all cell types derived from the three embryonic germ layers – mesoderm, endoderm and ectoderm, in chimeric mice. These three germ layers are the embryonic source of all cells of the body. Multipotent stem cells can generate several cell types within a lineage, and their differentiation capacity if often restricted to one germ layer. Unipotent stem cells contribute only to one mature cell type, and should perhaps be referred to as progenitor cells.

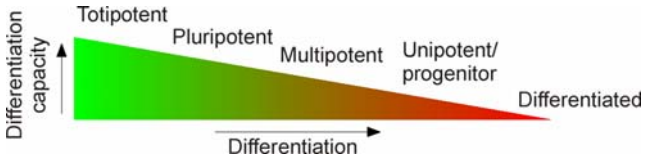



**Fig. I-1.** A gradient of stemness. As differentiation proceeds, the differentiation capacity of stem and progenitor cells along the gradient decreases.

Multipotent somatic adult stem cells, often commonly referred to as adult stem cells, are found in many different organs and tissues such as bone marrow, central nervous system, muscle, liver, dermis, epidermis, gastrointestinal tract, retina, adipose tissue, dental pulp and blood (Serafini and Verfaillie, 2006; Keating, 2006) (**Fig. I-2**). The origin of somatic stem

cells in mature tissues remains elusive and it is often difficult to distinguish adult, tissue-specific stem cells from progenitor cells on the basis of expression of specific surface markers. The term somatic stem cell plasticity describes the ability of tissue-specific stem cells to acquire, under certain microenvironmental conditions, the fate of cell types different from the tissue of origin. Somatic stem cells are rare in most tissues. As an example, 1 in ~15,000 cells in the bone marrow is a hematopoietic stem cell (HSC). The primary functions of tissue-specific somatic stem cells are to maintain the steady state functioning of a tissue (homeostasis) and with limitations, to replace cells that die because of disease or tissue injury (Baksh *et al.*, 2004; Serafini and Verfaillie, 2006).

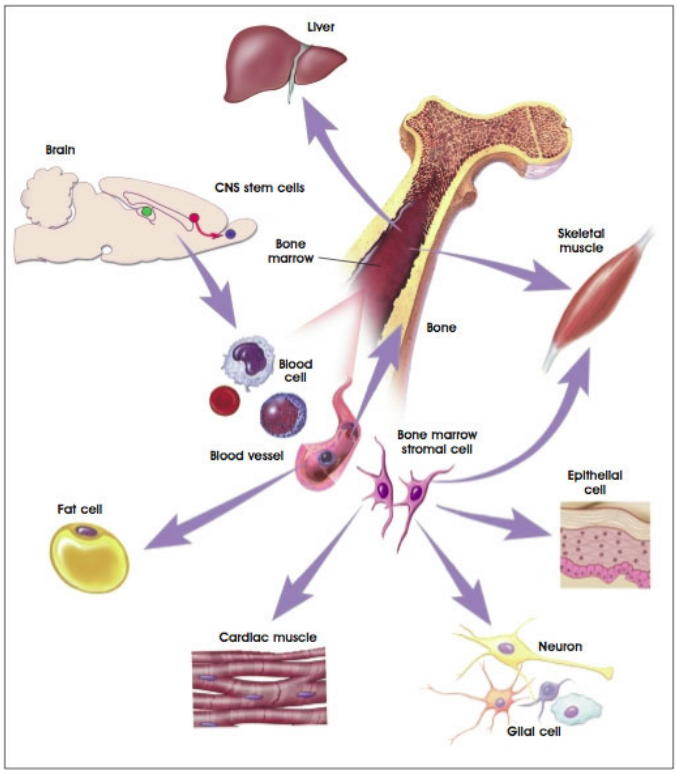



**Fig. I-2.** Multilineage differentiation capacity of bone marrow and neuronal stem cells (source: http://stemcells.nih.gov).

### 1.1. Mesenchymal stem cells

Mesenchymal stem cells (MSCs), also called stromal stem cells or mesenchymal stromal cells reside in the stroma of most organs. The stroma is the connective framework of a tissue, while the mesenchyme is connective tissue of mesodermal origin. MSCs were first described as bone-forming progenitors in bone marrow (Friedenstein *et al.*, 1968) Early cultures of non-adipocytic cells from brown adipose tissue (MASTERS, 1965) and an early report of "fibroblast-like transformation of human bone marrow fat cells *in vitro*" (Miyoshi *et al.*, 1966) suggest that also other laboratories were at that time studying a mesenchymal

population of cells that could be cultured and whose phenotype was able to change. MSCs have since been described and characterized from various adult mesenchymal tissues such as bone marrow (Pittenger *et al.*, 1999), adipose tissue (Zuk *et al.*, 2001; Katz *et al.*, 2005; Boquest *et al.*, 2005), umbilical cord (Wang *et al.*, 2004), cord blood (Lee *et al.*, 2004a), amniotic fluid (in 't Anker *et al.*, 2003), peripheral blood (Gronthos *et al.*, 1994), dermal tissue and skeletal muscle (Williams *et al.*, 1999).

The diversity of tissue sources from which MSCs have been isolated, and the use of various isolation methods have led to the proposal of three criteria that define multipotent MSCs by the International Society for Cellular Therapy. First MSCs should adhere to plastic and form fibroblast-like colonies (CFU-Fs) in standard culture conditions. Second, ≥95% of the MSC population must express CD105 (endoglin), CD73 and CD90 (Thy-1) on the surface, measured by flow cytometry. Additionally, less than 2% of these cells should express the hematopoietic markers CD45, CD34, CD14 or CD11b, CD79α and CD19 and HLA class II. Adult MSCs are reported to express intermediate levels of major histocompatibility complex (MHC) class I but do not express human leukocyte antigen (HLA class II) unless stimulated by, e.g., interferon-γ (Le *et al.*, 2003). MSCs are devoid of the endothelial cell marker CD31 (Chamberlain *et al.*, 2007). Third, the cells should be able to differentiate into osteoblasts, adipocytes and chondroblasts *in vitro*.

Indeed, MSCs can differentiate *in vitro* and *in vivo* along certain mesenchymal lineages such as adipocytic, osteoblastic and chondrocytic lineages. They may also give rise to skeletal and cardiac muscle cells and to endothelial cells (Pittenger *et al.*, 1999; Serafini and Verfaillie, 2006). A subset of MSCs may also give rise to neuronal or hepatic differentiated cell types (Sanchez-Ramos *et al.*, 2000; Weng *et al.*, 2003; Boquest *et al.*, 2005; Krabbe *et al.*, 2005).

Plastic-adherent cell populations isolated from bone marrow and adipose tissue contain cells with various differentiation capacities (Pittenger *et al.*, 1999; Zuk *et al.*, 2002; Katz *et al.*, 2005; Boquest *et al.*, 2005; Kucia *et al.*, 2005). This heterogeneity, together with the ability of MSCs to change in response to their environment, contributes to the difficulty to identify a unique MSC phenotypic fingerprint. The monoclonal antibody STRO-1 has been used to identify clonogenic bone marrow MSCs (Gronthos *et al.*, 1994), but MSCs also express a number of surface markers, none of which are specific to MSCs (Dominici *et al.*, 2006). While surface marker and gene expression profiles are being mapped by many research groups for various types of MSCs, variability in their differentiation potential has sparked studies on the identification of other markers, such as epigenetic marks associated with lineage-specification genes, that may define MSC subsets (Eckfeldt *et al.*, 2005). Some of this work is ongoing in our laboratory and the present work addresses the epigenetic makeup of human MSCs from several tissues.

### 1.2. Bone marrow mesenchymal stem cells

Multipotent MSCs derived from the stroma of bone marrow (BMMSCs) have the capacity to contribute to the regeneration of mesenchymal tissues such as bone, cartilage, muscle, ligament, tendon, adipose and stroma (Pittenger *et al.*, 1999). *In vitro*, BMMSCs can differentiate into adipogenic, chondrogenic and osteogenic lineages as well as into skeletal myocytes, neurons and endothelial cells (Wakitani *et al.*, 1995; Woodbury *et al.*, 2000; Reyes *et al.*, 2001). The immunophenotype of BMMSCs has been well characterized but is not necessarily unique for this cell type (Pittenger *et al.*, 1999; Chamberlain *et al.*, 2007). Moreover, BMMSCs secrete growth factors, interleukins, cytokines and chemokines, some of which are only expressed after stimulation. MSCs also express growth factor receptors, adhesion molecules and matrix molecules including fibronectin, laminin, collagen and

proteoglycans (Minguell *et al.*, 2001). These data sustain the view that BMMSCs contribute to the formation and function of a stromal microenvironment in bone marrow. This microenvironment produces inductive and regulatory signals for development of hematopoietic progenitors and other non-mesenchymal stromal cells (Klein, 1995; Minguell *et al.*, 2001).

## 1.3. Mesenchymal stem cells from adipose tissue

### *1.3.1. Phenotypic characteristics of adipose stem cells*

The stromal-vascular fraction of subcutaneous adipose tissue contains a cell population that displays multilineage developmental plasticity *in vitro* and *in vivo* (Zuk *et al.*, 2001). Multipotent adipose stem cells (ASCs) differentiate into adipogenic (**Fig. I-3A**), osteogenic (**Fig. I-3B-D**), chondrogenic (**Fig. I-3E**) and myogenic cell types, as well as along endothelial (**Fig. I-3F**) and neuronal (**Fig. I-3G**) lineages (Zuk *et al.*, 2001; Halvorsen *et al.*, 2001; Zuk *et al.*, 2002; Erickson *et al.*, 2002; Safford *et al.*, 2002; Gimble and Guilak, 2003; Boquest *et al.*, 2005; Strem *et al.*, 2005).

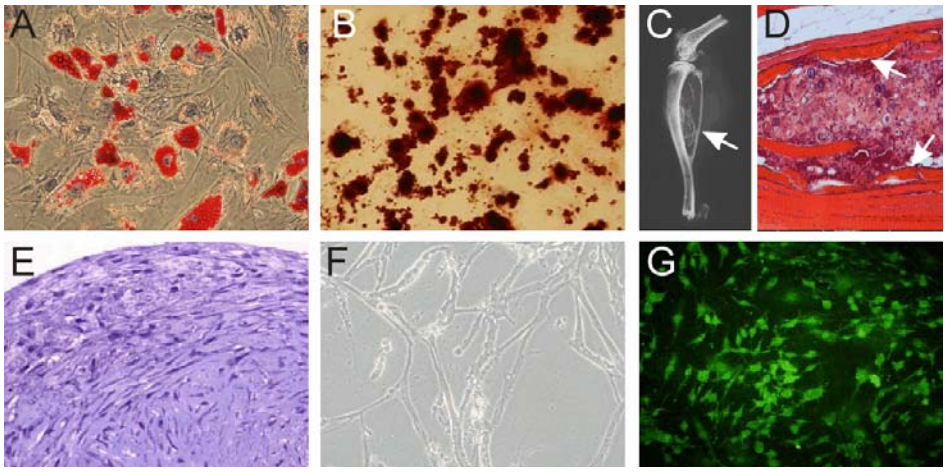



**Fig. I-3.** *In vitro* and *in vivo* differentiation of human ASCs. (A) Adipogenic differentiation (Oil Red-O staining). (B) Osteogenic differentiation (Alizarin red staining). (C,D) Ectopic bone formation in mouse muscle (arrows). (E) Chondrogenic differentiation (Toluidine blue staining). (F) Endothelial differentiation. (G) Neurogenic differentiation (immunostaining for IF-200). (A. Boquest, and P. Collas, unpublished data).

ASCs are similar but not identical to BMMSCs in various aspects. In culture, ASCs express surface markers similar to those expressed by BMMSCs, including CD105, CD73, CD90, CD44 and Stro-1, and do not express CD45 and CD31 (De Ugarte *et al.*, 2003; Boquest *et al.*, 2005; Strem *et al.*, 2005; Fraser *et al.*, 2006). BMMSCs however express different adhesion molecules with function in homing and mobilization of HSCs (De Ugarte *et al.*, 2003; Strem *et al.*, 2005; Fraser *et al.*, 2006), suggesting different homing properties for BMMSCs and ASCs. Notably, ASCs express the adhesion molecule integrin very late antigen CD49d (VLA4) but not CD106, whereas BMMSCs express CD106 but not CD49d (Boquest *et al.*, 2005). Another difference of practical significance is yield: in human bone marrow, only 0.001-0.01% of nucleated cells form CFU-Fs (Pittenger *et al.*, 1999) while adipose tissue harbors a large number of cells with CFU-F ability (Zuk *et al.*, 2002). Our laboratory has purified uncultured ASCs before plating, using antibodies to select against $CD45^+$ and $CD31^+$ cells from the stromal-vascular fraction (Boquest *et al.*, 2005). Up to 20 million ASCs per 300 ml lipoaspirate can be harvested in this manner.

ASCs constitute an attractive source of MSCs because they are abundant, easily purified, they do not require culture for isolation, unlike BMMSCs (Boquest *et al.*, 2005), and they show higher CFU-F ability than BMMSCs (Strem *et al.*, 2005; Kern *et al.*, 2006). ASCs also share the perceived therapeutical advantages of BMMSCs because they are also non-immunogenic (McIntosh *et al.*, 2006; Uccelli *et al.*, 2007) and they secrete cytokines supporting angiogenic and hematopoietic responses (Kilroy *et al.*, 2007).

The similarity of surface proteins expressed in MSCs from bone marrow or fat indicates that identifying a marker associated with the sub-population of highest differentiation potential among MSCs requires more than immunophenotyping. Gene expression profiling indicates that the transcriptome of ASCs and BMMSCs is similar (Lee *et al.*, 2004b; Boquest *et al.*, 2005), although minor differences have been claimed (Wagner *et*

*al.*, 2005). Despite intensive efforts, therefore, a marker that distinguishes MSCs with di-lineage, tri-lineage or higher differentiation capacity remains to be found. Another layer of analysis involves epigenetics: The present work addresses the question of whether the differentiation capacity of MSCs can be related to distinct epigenetics marks on genes that are expressed, repressed, or repressed with a potential for activation.

### *1.3.2. Expansion and differentiation capacity of adipose stem cells*

*In vitro*, ASCs can be expanded as polyclones or single cell-derived clones in long-term culture and retain a normal DNA content (Meza-Zepeda *et al.*, 2007).ASCs we have used in our laboratory senesce after 30-50 population doublings, in consistency with other reports (Stenderup *et al.*, 2003; Bonab *et al.*, 2006). The culture medium is supplemented with 10-20% fetal calf serum (FCS), although proliferation capacity is enhanced with bovine fibroblast growth factor (FGF) and epidermal growth factor (EGF). Polyclonal ASCs cultures were used in the present work.

Evidence suggests that ASCs can improve the function of damaged tissues *in vivo*. ASCs have been reported to differentiate into adipocytes, bone (A. Boquest and P. Collas, unpublished data; **Fig. I-3C,D**) and cartilage (Fraser *et al.*, 2006). ASCs may also incorporate into myofibers in injured muscle (Bacou *et al.*, 2004) and engraft within an infracted myocardium (Planat-Benard *et al.*, 2004). However, whether ASCs can truly form functional tissues of non-mesodermal lineages (Cousin *et al.*, 2003; Kim *et al.*, 2003; Kang *et al.*, 2003) remains to be carefully addressed.

ASCs may also differentiate toward the endothelial lineage. *In vitro*, ASCs can form a cellular network expressing endothelial markers (Planat-Benard *et al.*, 2004; Miranville *et al.*, 2004; Rehman *et al.*, 2004; Cao *et al.*, 2005). Freshly isolated ASCs express low levels of endothelial transcripts such as *CD144* (*CDH5*) and *CD31* (*PECAM1*) (Boquest *et al.*, 2005),

suggesting a potential for endothelial differentiation. ASCs can also release angiogenic factors (Rehman *et al.*, 2004; Nakagami *et al.*, 2005). After transplantation, ASCs promote re-vascularization of ischemic tissue either directly (Planat-Benard *et al.*, 2004; Miranville *et al.*, 2004; Miyahara *et al.*, 2006) or by secretion of angiogenic cytokines (Rehman *et al.*, 2004; Nakagami *et al.*, 2005; Kilroy *et al.*, 2007). We have shown that the differentiation capacity of human ASCs to endothelial cells *in vitro* is limited, as judged by the absence of significant upregulation of *CD144* and *CD31* (Boquest *et al.*, 2007). This limitation may be imposed by the highly methylated state of the *CD31* promoter in these cells (Boquest *et al.*, 2007).

### 1.4. Wharton's jelly mesenchymal stem cells

The mucoid connective tissue of the umbilical cord, the Wharton's jelly (WJ), is a rich source of MSCs. WJMSCs have been isolated from three relatively indistinct regions in the Whartons's jelly, namely the perivascular zone, the intervascular zone and the subamnion (Troyer and Weiss, 2008a) (**Fig. I-4**). Cells in human Wharton's jelly are not evenly distributed. The most immature cells are located in the intervascular and subamniotic regions, having greater competence to resume proliferation, whereas the perivascular zone mainly contains differentiated myofibroblasts (Nanaev *et al.*, 1997; Karahuseyinoglu *et al.*, 2007). MSCs have also been isolated from umbilical cord blood and umbilical vein subendothelium. It is unknown if MSCs isolated from different compartments of the umbilical cord represent different cell populations (Troyer and Weiss, 2008b).

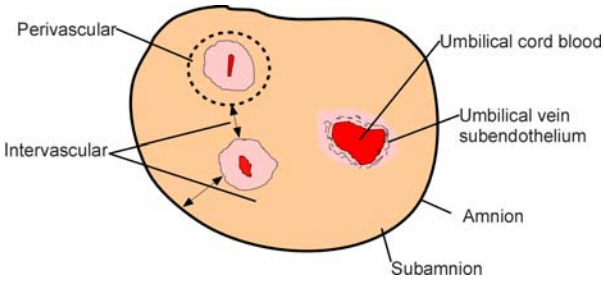



**Fig. I-4.** Compartments of the umbilical cord. Wharton's Jelly is the connective tissue surrounding the umbilical vessels and includes the intervascular and subamnion regions.

Human WJMSCs are CD45, CD34, CD14, CD31, and HLA class II negative while they express CD73, CD90, CD105, CD44 and HLA-class I, in consistency with MSCs from bone marrow and adipose tissue (Troyer and Weiss, 2008a). WJMSCs also express a low amount of Wnt-signaling molecules and higher telomerase activity than other somatic cells (Karahuseyinoglu *et al.*, 2007; Baksh *et al.*, 2007). These factors are key regulators of self-renewal and pluripotency in stem cells suggesting that WJMSCs use similar regulatory mechanisms to ESCs, especially when dissociated from their microenvironment. WJMSCs expand faster and to a greater extent than adult-derived MSCs, suggesting that WJMSCs may be a more "primitive" stromal cell population (Karahuseyinoglu *et al.*, 2007).

WJMSCs have been induced to differentiate *in vitro* into adipocytes, chondrocytes, osteocytes, endothelial cells (Wu *et al.*, 2007), and cells with characteristics of neurons (Fu *et al.*, 2006), cardiomyocytes and skeletal myocytes (Wang *et al.*, 2004; Conconi *et al.*, 2006), and photoreceptor cells (Lund *et al.*, 2007). Yet further work is needed to determine whether WJMSCs can engraft and display multipotency *in vivo*. The present work examines DNA methylation patterns on lineage-specific genes in human WJMSCs and to what extent these patterns correlate with their reported differentiation capacity.

### 1.5. Muscle-derived stem cells

In adult skeletal muscle, satellite cells are considered to be important for postnatal muscle growth and regeneration (Le and Rudnicki, 2007). Muscle satellite cells are quiescent, mononuclear cells located beneath the basal lamina of skeletal muscle fibers (**Fig. I-5**). Their self-renewal capacity has been demonstrated by transplanting single intact myofibers into radiation-ablated muscles to show that the donor satellite cells generate new satellite cells in the host (Collins *et al.*, 2005). In response to muscle damage or disease-induced degeneration, they are activated and undergo a terminal differentiation program, giving rise

to multinucleated muscle fibers (Schultz and McCormick, 1994; Collins *et al.*, 2005) (see **Fig. I-9**). The descendents of activated satellite cells, the muscle precursor cells (MPCs) or myoblasts, undergo multiple divisions before fusing and undergoing terminal differentiation. Satellite cells can differentiate into adipocytes and osteocytes *in vitro* (Asakura *et al.*, 2001). Although surface markers expressed by mouse satellite cells have been identified, human satellite cells are less characterized (Peault *et al.*, 2007). In consistency with other MSCs however, they do not to express CD34 in culture. Satellite cells also express several transcription factors required for myogenic differentiation (see below).

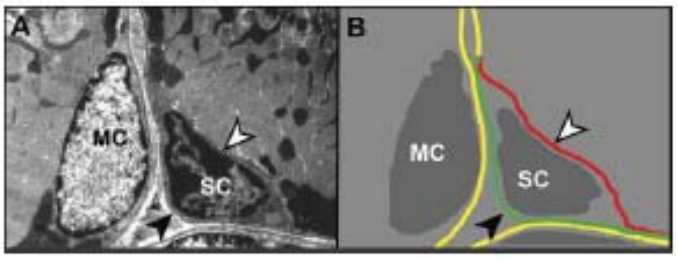



**Fig. I-5.** Satellite cells reside in a specialized niche in adult skeletal muscle. (A) Electron micrograph of skeletal muscle showing one myocyte nucleus (MC) and a satellite nucleus (SC). (B) Drawing of (A) emphasizing that the satellite cell lies between the basal lamina (black arrowhead and green line) and the sarcolemma (white arrowhead and red line). Taken from (Shi and Garry, 2006).

Alternative sources of stem cells have been identified in skeletal muscle, including side-population (SP) cells, muscle-derived stem cells (MDSCs), blood-derived circulating $AC133^{+}$ cells and vessel-associated stem cells and pericytes. SP cells are isolated by fluorescence-activated cell sorting (FACS) based on their capacity to exclude the DNA-binding dye Hoechst 33342 via the ABC transporter, and are present in several adult tissues (Zhou *et al.*, 2001). Muscle SP cells are unable to differentiate *in vitro* but they give rise to both differentiated cells and satellite cells when transplanted intramuscularly (Asakura *et al.*, 2002). Unlike satellite cells, SP cells can migrate in the blood stream after intravenous injection, home into muscle and contribute to muscle regeneration (Bachrach *et al.*, 2006). MDSCs have been purified from adult muscle by successive replating combined with FACS

sorting (Huard *et al.*, 2003). They have high self-renewal and proliferation capacity and can contribute to muscle regeneration (Payne *et al.*, 2005). It remains unclear to what extent MDSCs are different from satellite cells, which has led to authors using either terminology. Muscle-derived $AC133^+$ cells display myogenic differentiation capacity. $AC133^+$ cells are a subpopulation of blood-derived circulating cells expressing the hematopoietic marker AC133 (Torrente *et al.*, 2004). Pericytes associated with blood vessels of skeletal muscle can also regenerate diseased muscle and contribute to long term muscle regeneration, as well as give rise to other mesodermal lineages (Minasi *et al.*, 2002; Dellavalle *et al.*, 2007).

Cells with myogenic potential have also been derived from other tissues including bone marrow, brain and adipose tissue (Cossu and Biressi, 2005). However, the myogenic potential of BMMSCs, ASCs and neuronal stem cells remains questionable (see Results). The normal contribution of non-muscle cells to growing or regenerating muscle is probably minor and satellite cells are considered to play a major role in this process (Buckingham, 2006; Le and Rudnicki, 2007).

### 1.6. Hematopoietic stem cells

Hematopoietic stem cells (HSCs) are the best characterized multipotent stem cells. Human HSCs have been isolated from adult bone marrow, peripheral blood, umbilical cord, fetal liver and fetal bone marrow (Serafini and Verfaillie, 2006). HSCs give rise to all cell types of the lymphoid and myeloid lineages. A single HSC can self-renew and give rise to multiple progeny cells (Szilvassy *et al.*, 1989; Nolta *et al.*, 1996) that retain full lineage potential but have limited self-renewal capacity.

Two classes of HSCs can be detected on the basis of surface antigens and self-renewal capacity. Long-term repopulating HSCs can reconstitute all blood cell lineages after transplantation in a lethally irradiated mouse over a long time. Most cells in this class are

multipotent progenitors (Bryder *et al.*, 2006). Short-term HSCs are immediately able to regenerate all blood cell types in an irradiated mouse, but are not able to self-renew over long term (Morrison and Weissman, 1994; Serafini and Verfaillie, 2006).

Human HSCs express CD45, CD34, c-kit and Thy1 but not other lineage markers (Verfaillie *et al.*, 2002; Bryder *et al.*, 2006). During early development, HSCs can leave their tissue of origin, circulate and relocate to an available niche elsewhere (Quesenberry *et al.*, 2005). Adult HSCs can also leave the bone marrow, circulate and return to bone marrow (Bhattacharya *et al.*, 2006). Thus, based on their differentiation, migration properties and surface marker expression, HSCs represent a lineage clearly distinct from MSCs.

### 1.7. The stem cell niche

Stem cells reside within a defined microenvironment referred to as the stem cell niche (Fuchs *et al.*, 2004; Jones and Wagers, 2008) (**Fig. I-6**). Intercellular signals provided by the niche regulate stem cell proliferation, localization, differentiation and survival. The niche is composed of the stem cells themselves and differentiated cell types that interact with the stem cells and with each other. Extracellular matrix proteins provide structure, organization and mechanical signals. Via blood vessels, systemic signals as well as inflammatory and other cells can be recruited to the niche, while neurons provide neural input (**Fig. I-6**). The niche is also believed to provide a milieu protective against toxins and irradiation (Jones and Wagers, 2008).

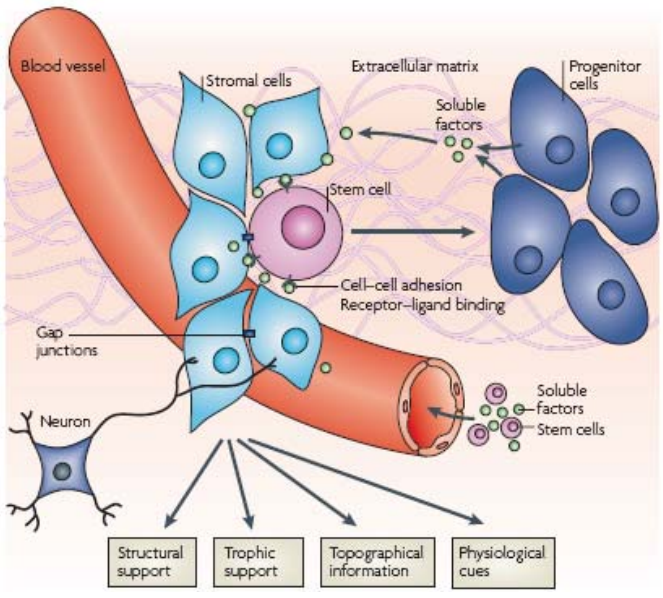



**Fig. I-6**. A stem cell niche. The various cell types making up the niche and influencing stem cell fate are shown. Taken from (Jones and Wagers, 2008).

In the niche, the self-renewal and differentiation properties of stem cells are enabled by asymmetric cell division, a distinct feature of stem cells (Fuchs *et al.*, 2004). Asymmetric cell division yields one stem cell daughter and one daughter cell committed to differentiation (**Fig. I-7**). Regulation of the balance between symmetric division (giving rise to two identical daughter stem cells) and asymmetric division is important for maintaining the proper stem cell number and providing the surrounding tissue with differentiated cells. Two models have been proposed to achieve asymmetric cell division (Knoblich, 2008). The first model relies on asymmetric distribution of fate determinants in the cell, which become segregated at mitosis (**Fig. I-7A**). A second model is based on extracellular signals which lead to orientation of the mitotic spindle perpendicular to the niche surface. At mitosis, only one daughter cell maintains contact with the niche and retains the ability to self-renew, while the other daughter cell acquires a committed phenotype (**Fig. I-7B**). The stem cell can also divide symmetrically, parallel to the niche, generating two stem cells.

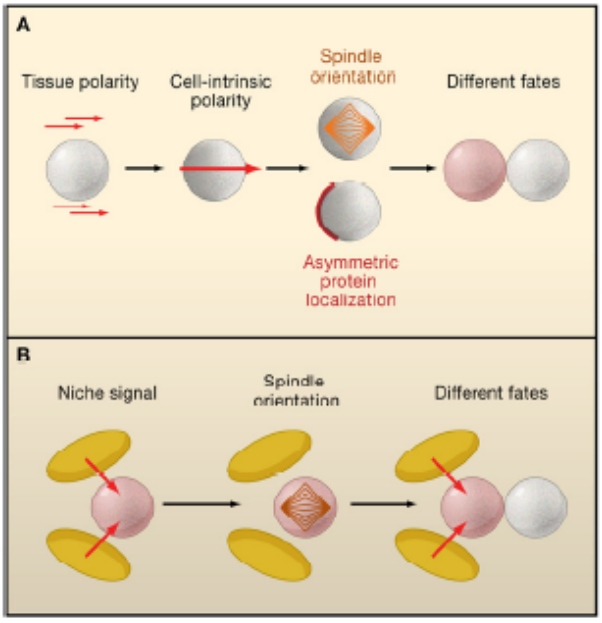



**Fig. I-7.** Regulation of stem cell self-renewal in the niche. (A) An axis of polarity distributes cell fate determinants asymmetrically at mitosis. The mitotic spindle is oriented along the same polarity axis ensuring asymmetric segregation of the determinants. (B) A stem cell may also depend on a signal coming from the niche (red arrows). By orienting the spindle perpendicular to the niche surface, only one daughter cell receives the signal and maintains the ability to self-renew, while the other cell may differentiate. Taken from (Knoblich, 2008).

Cell-cell adhesion proteins such as as cadherins and catenins are important for retaining stem cells in the niche (Song *et al.*, 2002; Jones and Wagers, 2008). Several signaling pathways are also essential for stem cell function in the niche (Fuchs *et al.*, 2004).

Notably, Wnt signaling plays a role both in proliferation and lineage specification (Kleber and Sommer, 2004). Notch signaling and the bone morphogenetic protein (BMP)/transforming growth factor β (TGFβ) superfamily are involved in stem cell plasticity (Fuchs *et al.*, 2004) The JAK-STAT pathway is also involved in maintenance and differentiation of stem cells (Zhao *et al.*, 2002). Thus, the balance of stemness, transient proliferation and differentiation involves many signaling pathways in the niche (Fuchs *et al.*, 2004).

Niches in the *Drosophila* germarium and testis, the subventricular zone of the brain, the hair follicle, the intestinal crypt and the bone marrow have been well characterized (Fuchs *et al.*, 2004). In bone marrow, HSCs are in close proximity to osteoblasts which, together with endothelial cell, play a role as HSC regulators (Porter and Calvi, 2008). Describing MSC niches has however been more difficult. Yet, it is within the cellular microenvironment in the bone marrow that BMMSCs are presumed to exist (Baksh *et al.*, 2004). Several studies also support the perivascular nature of the MSC niche and that this localization throughout the body would provide access to all tissues (Shi and Gronthos, 2003; da Silva *et al.*, 2006). ASCs are reported to associate with perivascular cells in adipose tissue (Zannettino *et al.*, 2008), thus pericytes may in fact be MSCs (Doherty and Canfield, 1999). The various sub-compartments of stem cells in adult tissues are likely to be at the origin of the difficulty of identifying MSC markers, and of differences in differentiation potential reported in various studies.

## 2. ADIPOGENIC AND MYOGENIC DIFFERENTIATION

### 2.1. Adipogenic differentiation

Differentiation of MSCs into adipocytes is a two-phase process (Rosen and MacDougald, 2006). The first phase, determination, regulated by activation of several signalling pathways,

involves the commitment of a pluripotent stem cell to the adipocyte lineage. The stem cell is altered to a pre-adipocyte, which cannot be distinguished morphologically from its precursor but has lost the potential to differentiate into other cell types. In the second phase, terminal differentiation, the pre-adipocyte differentiates into a mature adipocyte (Rosen and MacDougald, 2006).

Adipogenic differentiation is regulated by a network of transcription factors and co-regulators that coordinate expression of many genes and proteins (**Fig. I-8**). Peroxisome proliferator-activated receptor γ (PPARγ) and CCAAT-enhancer-binding protein α (C/EBPα) are the master regulators of adipogenesis (Farmer, 2006). In culture, preadipocytes become cell density-inhibited. Upon adipogenic stimulation, the cells re-enter the cell cycle and progress through two rounds of division (Otto and Lane, 2005). During this time, the cells express adipogenic transcription factors and cell cycle regulators that enable expression of PPARγ and C/EBPα. The committed cells differentiate and acquire the machinery for lipid synthesis and transport, insulin sensitivity and secretion of adipocyte proteins.

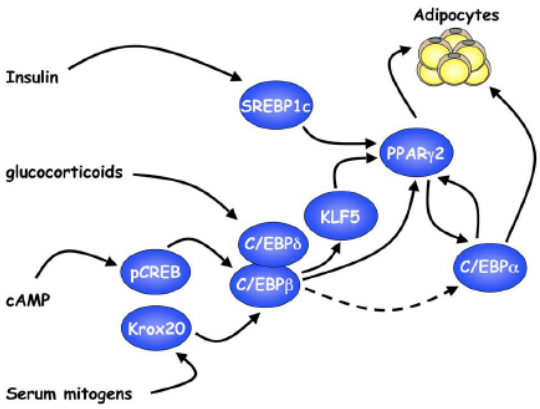



**Fig. I-8.** Induction of adipogenesis by a cascade of transcription factors. These can be activated *in vitro* by addition of insulin, glucocorticoids, cAMP-elevating agents and serum. Taken from (Farmer, 2006).

PPARγ is a member of the nuclear-receptor superfamily and is necessary and sufficient to initiate adipogenesis (Tontonoz *et al.*, 1994; Koutnikova *et al.*, 2003). Two isoforms of PPARγ, PPARγ1 and 2 are induced during adipogenesis, however PPARγ1 is also expressed in other cell types (Rosen and MacDougald, 2006). PPARγ activates genes involved in fatty acid binding, uptake and storage, including *aP2/FABP4* (fatty acid binding protein 4) and *LPL* (lipoprotein lipase).

C/EBPα is a pleiotropic transcriptional activator of adipocyte-specific genes including *LEP* (leptin) by binding C/EBP elements (Otto and Lane, 2005). C/EBPα plays a role in maintaining PPARγ expression in adipocytes (Wu *et al.*, 1999). Two other members of the C/EBP family, C/EBPβ and C/EBPδ induce C/EBPα and PPARγ expression (Clarke *et al.*, 1997). In turn, PPARγ, C/EBPβ and C/EBPδ co-activate C/EBPα. The cAMP regulatory element-binding protein CREB also participates in induction of C/EBPβ (Zhang *et al.*, 2004), while C/EBPδ induction is facilitated by glucocorticoids (Cao *et al.*, 1991). These processes explain why inducers of cAMP (isobutylmethylxanthine) and glucocorticoids (dexamethasone) are needed in the adipogenic differentiation medium (**Fig. I-8**). Other regulators of adipogenesis include SREBP1c (sterol regulatory element binding protein 1c) which in response to insulin induces PPARγ (Kim *et al.*, 1998) and Krox20, induced by exposure to mitogens, which promotes expression of C/EBPβ (Chen *et al.*, 2005). Transcription factors of the E2F family have also been shown to regulate adipogenesis (Fajas *et al.*, 2002). Thus, adipogenesis involves a complex balance of factors.

### 2.2. Myogenic differentiation

Similarly to adipogenesis, myogenesis involves a series of differentiation steps leading to committed progenitors and to multinucleated myogenic cells. In adult muscle, satellite cells are mitotically quiescent and become activated in response to stress elicited by weight bearing, muscle damage or disease-induced muscle degeneration. Microenvironment-secreted growth factors also promote satellite cell activation (Le and Rudnicki, 2007). One key signaling molecule in satellite cells is sphingosine-1-phosphate, required for entry into the cell cycle (Nagata *et al.*, 2006).

Myogenic progenitors express specific transcription factors at defined stages of myogenesis (Buckingham, 2006; Peault *et al.*, 2007) (**Fig. I-9**). Pax7 is expressed in satellite

cells and myogenic precursor cells *in vivo*, and in primary myoblasts *in vitro*. Pax7 is essential for postnatal maintenance and self-renewal of satellite cells. Pax7 and its paralog Pax3 (Montarras *et al.*, 2005) are key upstream regulators of myogenesis and induce expression of the myogenic determination gene *MyoD* (Buckingham, 2006; Peault *et al.*, 2007). Myf5 is expressed at low level in quiescent satellite cells and continues to be expressed upon activation (Buckingham, 2007). Co-expression of MyoD and Myf5 is required for activation of myogenin. The transcription factors myogenin and MRF4 (Myf6) control muscle differentiation, leading to cell fusion and multinucleated myofibers (Chen and Goldhamer, 2003). Notably however, differentiation of satellite cells is preceded by down-regulation of the *Pax* genes, a process regulated by myogenin itself (with specificity for Pax7) (Olguin *et al.*, 2007).

The human muscle-derived cells used in this study were shown by quantitative reverse transcription polymerase chain reaction (RT-qPCR) to express *MYF5*, *MYOD* and *MYOG* at the transcript level, while at the protein level no myogenin expression was observed **(**see Results**)**. Because *PAX3* and *PAX7* were not expressed, these cells were at or beyond the myoblast stage (**Fig. I-9**).

We refer to these cells as MPCs.

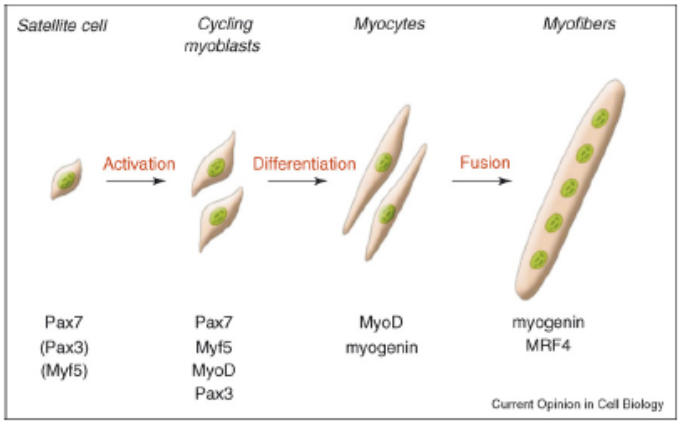



**Fig. I-9.** Myogenic differentiation. Distinct sets of genes are expressed at specific stages of myogenesis, enabling the distinction of four progenitor cell types. Taken from (Le and Rudnicki, 2007).

# 3. EPIGENETIC REGULATION OF GENE EXPRESSION: ROLE OF DNA METHYLATION

## 3.1. Introduction to epigenetics

Genomic DNA contains the central core of genetic information of the cell. Throughout development, distinct patterns of gene expression are set up in somatic cells, and stably inherited through cell division. These patterns must be specified by information apart from the primary DNA sequence (Meehan, 2003). Superimposed upon the DNA sequence is a layer of heritable epigenetic information. Epigenetic mechanisms refer to heritable modifications of DNA and chromatin that do not affect DNA sequence (Collas *et al.*, 2007). Epigenetic modifications fall into two main categories: DNA methylation and histone modifications. In addition to epigenetic modifications, the positioning of transcriptional activators, transcriptional repressors, other chromatin remodelling complexes and small interfering RNAs on target genes also regulate gene expression (Kawasaki and Taira, 2004; Matzke and Birchler, 2005).

In the nucleus, DNA is packed into chromatin. The nucleosome is the fundamental unit of chromatin and consists of 147 base pairs of DNA wrapped around an octamer composed of 2 subunits of each of histone H2A, H2B, H3 and H4. The linker DNA between the nucleosomes is associated with histone H1. Histones and in particular their amino-terminal tails protruding from nucleosomes are subject to post-translational modifications such as acetylation, methylation, phosphorylation, ubiquitylation, sumoylation, ADP-ribosylation, deimination and proline isomerization (Kouzarides, 2007) (**Fig. I-10A)**. In addition, chromatin may be modified by dynamic replacement of core histones by histone variants such as the deposition of histone H3.3 on transcriptionally active promoters (Mito *et al.*, 2005; Mito *et al.*, 2007).

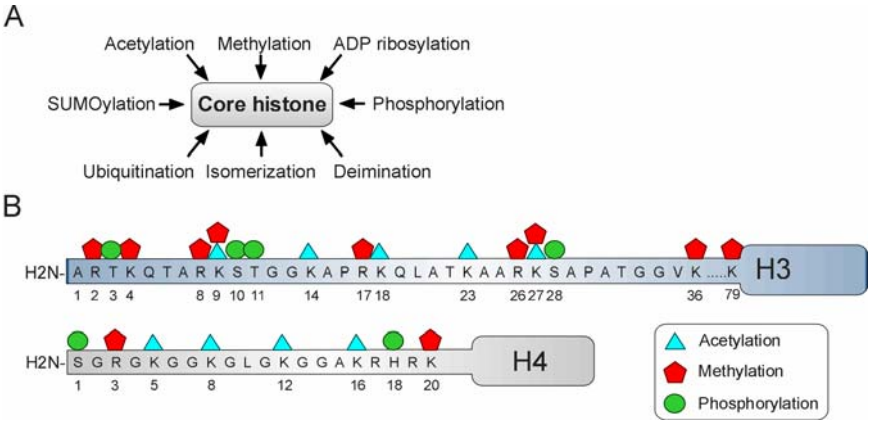



**Fig. I-10.** Post-translational histone modifications. (A) Core histones can be methylated, acetylated, phosphorylated, ubiquitinated or sumoylated to modulate gene expression. (B) Site and nature of known modifications on the amino-terminal tails of H3 and H4.

Epigenetic histone modifications are so far best characterized for H3 and H4, in particular lysine acetylation and methylation (**Fig. I-10B**). Lysine acetylation almost always correlates with transcriptionally active chromatin while lysine methylation can have different effect depending on which residue is modified. Methylation of histone H3 lysine 4 (H3K4) and H3K36 is associated with transcribed chromatin. In contrast, methylation of H3K9, H3K27 and H4K20 generally correlates with repression (Bernstein *et al.*, 2007). In particular, di- and tri-methylation of H3K9 (H3K9m2, H3K9m3) and tri-methylation of H3K27 (H3K27m3) elicit the formation of repressive heterochromatin through the recruitment of heterochromatin protein 1 (HP1) (Lachner *et al.*, 2001) and polycomb group (PcG) proteins, respectively (Cao *et al.*, 2002). Acetylation of histone tails neutralizes the charge interaction between the DNA backbone and the histone tail. Acetylated lysines are recognized by bromodomains within nucleosome remodelling complexes, creating a chromatin conformation accessible for transcription activators (Bernstein *et al.*, 2007). Acetylation of H3K9 (H3K9ac) and H4K16 (H4K16ac), together with di- and tri-methylation of H3K4 (H3K4m2, H3K4m3), are exclusively found in euchromatin, often in association with transcriptionally active genes (Struhl, 1998; Santos-Rosa *et al.*, 2002).

Histone tail modifications are mediated by specific enzymes. Numerous histone acetyl transferases (HATs), histone methyltransferases (HMTs), histone deacetylases (HDACs) and

histone demethylases (HDMases) have been identified (Kouzarides, 2007). Enzymes have also been identified for phosphorylation, ubiquitination, sumoylation, ADP-ribosylation, deimination and proline isomerization (Kouzarides, 2007). Combinations of histone modifications create an enormous potential for functional responses; however, not all these modifications mark the same histone at the same time, and the modifications are dynamic and rapidly changing.

### 3.2. DNA methylation

Cytosine methylation on DNA in general is associated with long-term gene silencing. In vertebrates, DNA methylation occurs almost exclusively on cytosine-phosphate-guanine (CpG) dinucleotides. DNA methylation consists of the addition of a methyl group to position 5 of cytosine in the CpG dinucleotide (**Fig. I-11A**). CpG methylation is symmetrical – it occurs on both strands (**Fig. I-11B**) – and targets isolated CpGs, clustered CpGs, or clustered CpGs within a CpG island. A CpG island is defined as a sequence in which the observed/expected CpG frequency is greater than 0.6, with a G+C content greater than 50%. The expected number of CpGs in a given 200-bp window is calculated as the number of C's in the window multiplied by the number of G's in the window, divided by window length (Gardiner-Garden and Frommer, 1987). This 200-bp window is moved across the sequence of interest at 1 bp intervals. Takai and Jones published in 2002 a more refined definition where regions greater than 500 bp, with a G+C content >55 % and an observed CpG/expected CpG ratio of 0.65, provide a more accurate association with 5' regions of genes and excludes most Alu repeats (Takai and Jones, 2002). CpG islands are often found in the 5' regulatory regions of vertebrate housekeeping genes; they are often protected from methylation, enabling constitutive expression of these genes. A consequence of this protection from de novo methylation is that CpG islands can remain unmethylated even when

their associated gene is silent (Bird *et al.*, 1987; Weber *et al.*, 2007). Changes in CpG island methylation are often (but not always) associated with disease. CpG islands in the promoter of tumor suppressor genes are unmethylated in normal cells, whereas a hallmark of cancer is *de novo* methylation of these CpG islands, resulting in repression of tumor suppressor genes and triggering of an uncontrolled cell cycle (Robertson, 2005).

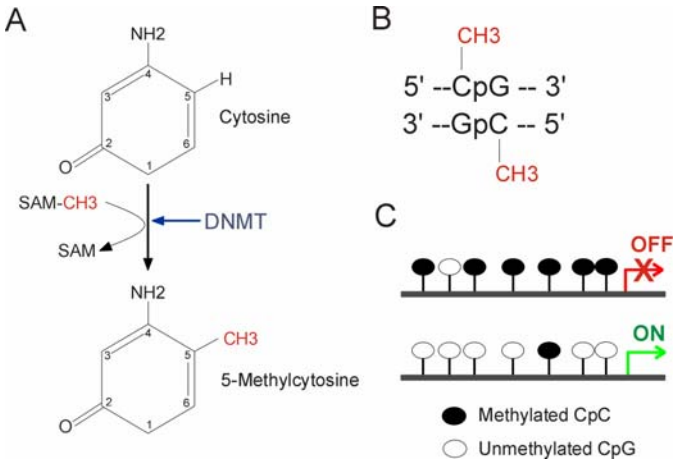



**Fig. I-11.** Principles of DNA methylation. (A) Mechanism of DNA methylation. (B) CpG methylation is symmetrical. (C) DNA methylation contributes to long-term gene silencing.

CpG methylation is catalyzed by DNA methyltransferases (DNMTs). The maintenance DNA methyltransferase DNMT1 specifically recognizes hemi-methylated DNA after replication and methylates the daughter strand, ensuring fidelity in the methylation profile after replication (Jaenisch and Bird, 2003). In contrast to DNMT1, DNMT3a and DNMT3b are implicated in *de novo* DNA methylation that takes place during embryonic development and differentiation (Turek-Plewa and Jagodzinski, 2005), as a means of shutting down genes whose activity is no longer required as cells differentiate. DNMT2 has to date no clear ascribed function in DNA methylation (Hermann *et al.*, 2003), but may have cytoplasmic transfer RNA methyltransferase activity (Goll *et al.*, 2006; Rai *et al.*, 2007). DNMT3L is a DNMT-related protein with no DNA methyltransferase activity, but which physically associates with DNMT3a and DNMT3b and modulates their catalytic activity (Suetake *et al.*, 2004). DNA methylation largely contributes to long-term gene silencing (**Fig. I-11C**) (Klose and Bird, 2006; Hoffman and Hu, 2006) and as such it is essential for development (Reik *et al.*, 2001; Morgan *et al.*, 2005).

DNA methylation patterns of somatic cells are established according to developmental program. In early mammalian development, the paternal genome is actively demethylated shortly after fertilization whereas the maternal genome is subsequently demethylated through several rounds of replication (Reik *et al.*, 2001). Thereafter, genome-wide re-methylation occurs rapidly in the blastocyst and results in the methylation pattern found in adult somatic cells. DNA methylation is also implicated in X chromosome inactivation (Hellman and Chess, 2007) and genomic imprinting (Tremblay *et al.*, 1995). DNA demethylation and methylation during embryogenesis are tightly controlled and are important for viability of the embryo. It is the best evidence of demethylation *in vivo* so far. However, the mechanisms by which it occurs are unknown (Reik *et al.*, 2001; Cortazar *et al.*, 2007). A thymine DNA glycosylase (removing thymine from G·T mispairs) has been suggested to have a 5-mC DNA glycosylase activity and function as a DNA demethylase, yet its role remains debated (Cortazar *et al.*, 2007). In adult cells, alterations of DNA methylation are often associated with disease and are a hallmark of cancer (Jones and Baylin, 2002).

### 3.3. Mechanisms of DNMT targeting to DNA

Studies of *de novo* DNA methylation in cell culture model systems have suggested at least three possible means by which *de novo* methylation might be targeted (**Fig. I-12**). First, DNMT3a and b themselves might recognize DNA or chromatin via specific domains (**Fig. I-12A**). In mouse cells, DNMT3 enzymes have been found to partially localize to regions of pericentromeric heterochromatin with the conserved PWWP (Pro-Trp-Trp-Pro) motif required for targeting to these regions. The PWWP domain may interact with DNA in a sequence independent manner (Qiu *et al.*, 2002). Second, DNMT3s may be recruited through interactions with transcriptional repressors or other factors, which in turn are targeted to DNA or chromatin (**Fig. I-12B**). For example, the transcription factor Myc associates with

DNA methyltranseferase activity and a direct interaction between DNMT3a and Myc is required for efficient repression of the Myc target gene $P21^{cip1}$ (Brenner *et al.*, 2005). Other studies have shown that the nucleolar remodelling complex NoRC represses ribosomal DNA transcription and associates with DNMT3b and DNMT1 (Santoro *et al.*, 2002; Santoro and Grummt, 2005). Third, short interfering RNAs may induce *de novo* DNA methylation and gene silencing (**Fig. I-12C**). For instance in plants, RNAi-mediated gene silencing results in transcriptional silencing often as a result of *de novo* methylation of the silenced gene (Matzke and Birchler, 2005). *De novo* DNA methylation during RNAi silencing in mammalian cells has also been reported (Morris *et al.*, 2004; Kawasaki and Taira, 2004); however these results remain controversial to date (Klose and Bird, 2006). Thus, an RNAi mechanism may also be responsible for DNMT targeting, at least in plants.

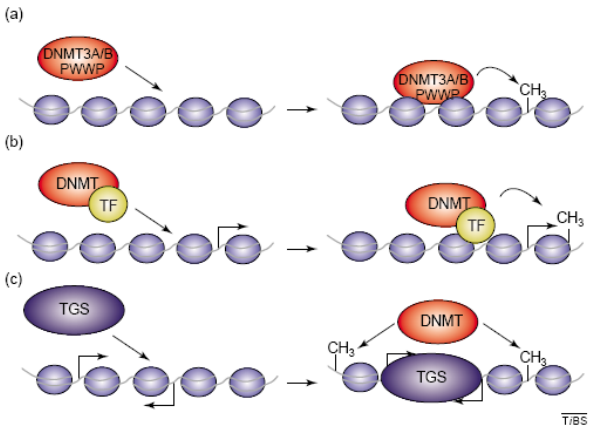



**Fig. I-12.** Targeting *de novo* methylation. (A) The PWWP domain of DNMT3a and DNMT3b is required to target the methyltransferases to regions of pericentromeric heterochromatin. (B) Transcription factors (TF) have the capacity to interact with DNMTs to recruit methyltranseferase activity. (C) *De novo* DNA methylation might be targeted by transcriptional gene silencing (TGS) pathways that respond to RNAi signals. Taken from (Klose and Bird, 2006).

### 3.4. Mechanisms of DNA methylation-dependent transcriptional gene silencing

Several mechanisms have been proposed for DNA methylation-dependent repression of gene expression (Klose and Bird, 2006). (i) Cytosine methylation can directly interfere with transcription factor binding to cognate DNA sequences, and thereby inhibit transcription (**Fig. I-13A**). (ii) However, repression seems to occur largely indirectly, via recruitment of methyl CpG-binding domain (MBD) proteins that induce chromatin changes (**Fig. I-13B**). A mammalian methyl-CpG-binding protein was identified a "long" time ago (1989) (Meehan *et al.*, 1989) and since, a family of five mammalian methyl-binding proteins (MBPs) sharing the

methyl-CpG-binding domain has been identified through bioinformatics (Hendrich and Bird, 1998). MeCP2, MBD1, MBD2 and MBD4 all recognize methyl-CpG. MBD3 however contains amino acid substitutions that prevent binding to methyl-CpG. Kaiso is a novel MBP that lacks the MBD but recognize methylated DNA through zinc-fingers (Prokhortchouk *et al.*, 2001). All MBPs can recruit chromatin remodelling co-repressor complexes to regions containing DNA methylation and mediate silencing of gene expression (Klose and Bird, 2006). (iii) DNMTs can, in addition to their enzymatic role, also interact with HMTs (Fuks *et al.*, 2003; Geiman *et al.*, 2004), HDACs (Fuks *et al.*, 2000; Fuks *et al.*, 2001; Geiman *et al.*, 2004) and the ATP-dependent chromatin remodelling protein hSNF2H (Geiman *et al.*, 2004), and thereby induce chromatin modifications and transcriptional repression (**Fig. I-13C**).

(iv) Lastly, there are indications that DNA methylation is involved in the maintenance, rather than initiation, of gene silencing (Weber and Schubeler, 2007) (**Fig. I-13D**). In a study of the epigenetic reprogramming of the *Oct4* gene during differentiation of mouse ES cells, DNA methylation was found to be a late event and dispensable to initiate silencing, but was required to stably prevent re-expression of the gene (Feldman *et al.*, 2006).

The relationship between DNA methylation and gene expression is complex (Jones and Takai, 2001) and recent evidence based on genome-wide CpG methylation profiling highlights the CpG content of promoters as a component of this complexity (Weber *et al.*, 2007). It has been shown that sequences outside promoters have a high degree of DNA methylation. Thus in mammals most DNA outside regulatory regions (intergenic DNA, in genes, both intronic and exonic regions, and in repeat elements) appears to be methylated (Eckhardt *et al.*, 2006; Weber and Schubeler, 2007). DNA methylation is consequently suggested to play a role in the global maintenance of the genome. For instance, a potential role for intragenic DNA methylation can be to inhibit cryptic transcriptional initiation outside

gene promoters. However, intragenic DNA methylation has also been suggested to reduce gene expression probably by reducing the capacity of RNA polymerase II (RNAPII) to transcribe through methylated regions (**Fig. I-13D**) (Klose and Bird, 2006). DNA methylation of transposable elements can lead to transcriptional inactivation and immobilization of these elements; thus, DNA methylation is important for maintenance of genome stability (global hypomethylation is associated with increased genome instability) (Weber and Schubeler, 2007).

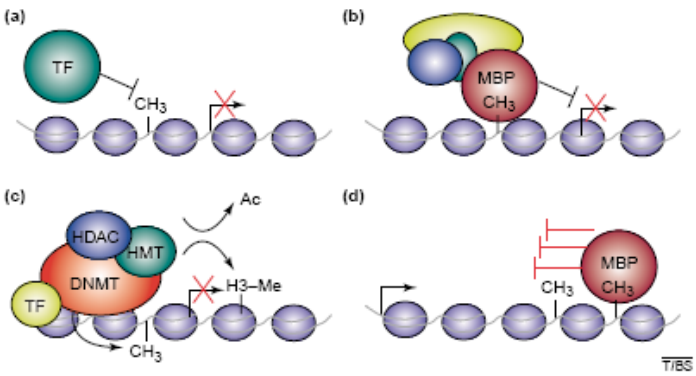



**Fig. I-13.** Mechanisms of DNA methylation mediated repression. (A) DNA methylation can inhibit transcription factors binding to target sites and thereby directly inhibits transcription activation. (B) MBPs directly recognize methylated DNA and recruit co-repressors to silence transcription and modify surrounding chromatin. (C) DNMTs can interact with HMTs and HDACs to induce chromatin modifications and gene repression. (D) DNA methylation within the body of genes can reduce transcriptional elongation rates. MBPs might be involved, either directly or by their effects on the surrounding chromatin. Taken from (Klose and Bird, 2006).

### 3.5. Methods for detecting DNA methylation

There are a number of methods for studying cytosine methylation, including bisulfite genomic sequencing, methylation-sensitive restriction digestion, or immunoprecipitation of MBPs or anti-5-methylcytosine (5-mC) (Clark *et al.*, 2006). In combination with DNA microarrays and high-throughput sequencing these are powerful tools for detection of DNA methylation on a genome-wide scale. Methylation-sensitive restriction enzyme-based methods utilize restriction enzymes, which either do not cleave their recognition site when it is methylated, or specifically digest methylated DNA. An important limitation is that these techniques are restricted to the analysis of methylation only within enzyme recognition sites.

Thus, two other strategies were used in this thesis, namely bisulfite genomic sequencing and immunoprecipitation of methylated DNA.

In the bisulfite genomic sequencing approach, denatured DNA is treated with sodium bisulfite and a series of reactions convert unmethylated cytosines to uracil while methylated cytosines remain as cytosines (**Fig. I-14A,B**) (Clark *et al.*, 2006). PCR amplification of converted DNA using primers to the region of interest replaces uracil with thymine and subsequent sequencing determines, by reading a thymidine or cytosine, the methylation state of the CpG dinucleotides in the original sequence (**Fig. I-14A**). This thesis has relied on bisulfite genomic sequencing to start unveiling the DNA methylation profile of tissue-specific genes in human MSCs and HSCs. For our work, reverse sequencing was used to unveil either an adenine (A) corresponding to an unmethylated cytosine, or a guanine (G), corresponding to a methylated cytosine (**Fig. I-14C**). Bacterial cloning of the PCR products generates several sequences to provide a quantitative assessment of the extent of methylation of a given CpG.

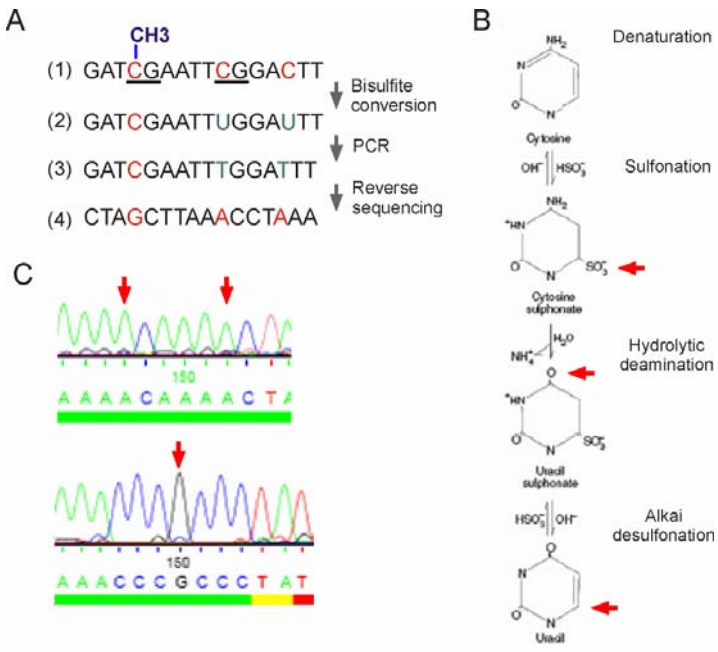
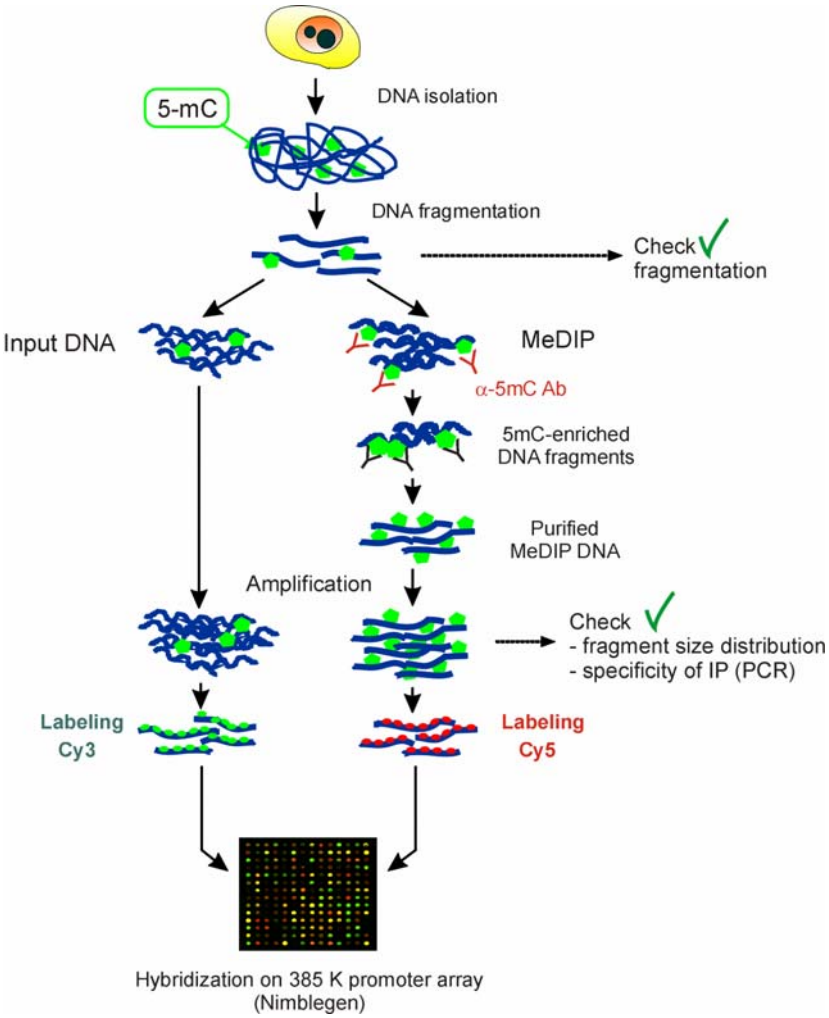



**Fig. I-14.** Bisulfite genomic sequencing. (A) Overview of the procedure. (1) A DNA sequence (here, random) contains methylated and unmethylated cytosines in CpG dinucleotides (underlined). (2) Bisulfite treatment converts unmethylated Cs to Us while the methylated C is not converted. (3) PCR changes Us to Ts and (4) reverse sequencing identifies which cytosine was methylated and which was not in the original (top) sequence. (B) Reactions involved in the bisulfite conversion process. (C) Examples of sequences obtained after bisulfite conversion, PCR and reverse sequencing. Arrows in the top panel point to two As representing two unmethylated cytosines. Arrow in the bottom panel points to a G, representing a methylated C in the original sequence.

A second powerful strategy for analyzing DNA methylation relies on the enrichment of methylated DNA sequences by affinity purification. This is achieved either by immunoprecipitating MBPs, or, in the case of the present work, immunoprecipitating methylated DNA using a monoclonal antibody that specifically recognizes 5-mC (Weber *et al.*, 2005; Zilberman and Henikoff, 2007). In the methylated DNA immunoprecipitation (MeDIP) assay (**Fig. I-15**), purified genomic DNA is fragmented to ~200-1,000 bp by sonication, and 5-mC enriched fragments are immunoprecipitated using anti-5mC antibodies. Fragmented input DNA remains untreated. Precipitated and input DNAs are amplified using a whole genome amplification kit and specificity of the immunoprecipitation is verified by PCR. For microarray-based analysis, such as that performed in this study, input DNA is labeled green with, e.g., Cy3 and MeDIP DNA is labeled red with Cy5 (**Fig. I-15**). In this study, MeDIP and the corresponding input DNA samples were co-hybridized onto a Nimblegen promoter array. A detailed account of the MeDIP procedure implemented in the laboratory as part of this work is provided in Materials and Methods and in the **Supplementary Manuscript**.

**Fig. I-15.** The MeDIP assay. Genomic DNA is purified from cells, fragmented to ~200-1,000 bp by sonication, and 5-mC enriched fragments are immunoprecipitated using anti-5mC antibodies (α-5mC Ab). Fragmented input DNA remains as is. Precipitated and input DNAs are amplified. Uniformity of amplified fragment size distribution is again assessed by agarose gel electrophoresis and specificity of the immunoprecipitation is verified by PCR. For array-based analysis, such as that performed in this study, MeDIP and input DNA are differentially labeled. MeDIP and corresponding input DNA samples are co-hybridized onto genomic (e.g., promoter) arrays.



# FRAMEWORK AND AIMS OF THE STUDY

MSCs, including ASCs, can differentiate into multiple cell types *in vitro* and *in vivo*. However, ASCs have a propensity to differentiate into primarily mesodermal lineages, and even so, their capacity to differentiate into non-adipogenic mesodermal pathways seems to be limited. With the aim of identifying an epigenetic basis for this restricted differentiation capacity, our research group has performed an analysis of DNA methylation at promoters of lineage-specific genes in ASCs and shown CpG hypomethylation of adipogenic promoters, in contrast to heavier methylation at non-adipogenic, lineage-specific promoters.

**The objective of this study was to test the hypothesis that the DNA methylation pattern on lineage-specific promoters in MSCs may constitute a predictor of differentiation potential.**

To test this hypothesis, we examined CpG methylation in selected promoters in different human MSC types such as adipose tissue-derived stem cells (ASCs), human muscle-derived progenitor cells (MPCs), bone marrow MSCs (BMMSCs) and Wharton's Jelly MSCs (WJMSCs). HSCs were also analyzed, together with differentiated adipocytes, skin fibroblasts, keratinocytes and DNA from a muscle biopsy.

The specific aims of the study were to:

1. Analyze by bisulfite genomic sequencing the DNA methylation profiles of adipogenic, myogenic and endothelial gene promoters in undifferentiated ASCs, BMMSCs, MPCs, WJMSCs and HSCs.
2. Assess promoter DNA methylation changes upon adipogenic and myogenic differentiation of ASCs and MPCs, and compare these profiles to mature adipocytes and muscle, respectively.

3. Evaluate at the phenotypic and transcriptional level the differentiation capacity of ASCs and MPCs towards adipogenic and myogenic lineages.
4. Implement and validate a methylated DNA immunoprecipitation (MeDIP) assay for genome-wide DNA methylation profiling of ASCs, BMMSCs, MPCs and HSCs.
5. Carry out preliminary analyses of DNA methylation profiling in ASCs, BMMSCs, MPCs and HSCs

In the course of this work, I have benefited from technical assistance from Kristin Vekterud. She has performed, specifically, bisulfite sequencing analyses of BMMSCs and WJMSCs (shown in Fig. R-2A), adipogenic-differentiated ASCs (shown in Figs. R-3A and R-8A), $CD31^+$ SVF cells and HUVECs (shown in Fig. R-8A). She has also performed adipogenic and myogenic differentiation of BMMSCs (shown in Figs. R-4A and R-5C). All other experiments and analyses were performed by me.

# MATERIALS AND METHODS

## Cells

ASCs were purified from the stromal vascular fraction of human adipose tissue as described earlier (Boquest *et al.*, 2006b). In short, stromal cells were isolated from liposuction material by collagenase and DNase digestion, sedimentation and straining. $CD45^+$ and $CD31^+$ cells were removed by magnetic cell sorting to eliminate hematopoietic and endothelial cells, respectively, resulting in $CD45^-CD31^-$ cells which were shown to have MSC properties (Boquest *et al.*, 2005). Purified ASCs were plated overnight in DMEM/F12 containing 50% fetal calf serum (FCS) to facilitate adhesion and further cultured in DMEM/F12/10% FCS. A pool of ASCs from three donors (healthy women, aged 24-40) was used in this study. Cells were passaged with a split ratio of 1:3 by trypsinization and used at passage (P) 12. This polyclonal culture was shown to senesce at P30 (Noer *et al.*, 2007).

$CD45^-CD31^+$ endothelial progenitor cells were purified from adipose tissue by magnetic cell sorting and consisted of the positively sorted $CD31^+$ cells of the stromal vascular fraction (see above) (Boquest *et al.*, 2005). These cells were not cultured but rather, were analyzed as freshly isolated cells for DNA methylation (Boquest *et al.*, 2005). The cells used in this study were isolated by Andrew Boquest in the laboratory prior to the start of this project.

$CD14^-CD34^+$ BMMSCs (a gift from Aboulghassem Shahdadfar, Institute of Immunology, Rikshospitalet-Radiumhospitalet Medical Center, Oslo) were isolated and cultured in DMEM/F12 supplemented with 20% FCS (Shahdadfar *et al.*, 2005). Cells were used at P4 in this study and shown to senesce between P10 and 14 (not shown).

HSCs were isolated from bone marrow (Steidl *et al.*, 2004). Mononuclear cells were isolated by Lymphoprep (GE Healthcare; ww.gehealthcare.com) density gradient

centrifugation and resuspended in RPMI 1640. $CD34^+$ cells were positively selected using magnetic beads (Direct CD34 Progenitor Cell Isolation Kit, Miltenyi Biotec; www.miltenyibiotec.com). Purity of the $CD34^+$ HSCs was >97%, determined by flow cytometry. Purified HSCs were donated by Aboulghassem Shahdadfar.

Primary cultures of MPCs derived from human muscle were purchased from Lonza (CC-2580 Muscle Myoblast Cell System; www.lonza.com). Cells were cultured in SkGM® Skeletal Muscle Medium (Lonza) and used at P7-P8 (this was necessary to obtain sufficient cell numbers, although cells reduced proliferation at P9; not shown).

WJMSCs were derived in the laboratory of Mark Kirkland (Deakin University, Geelong, Australia) from umbilical cord explants. The outer membrane was manually removed, eliminating most of the subamnion, and blood vessels were dissected out, removing most perivascular cells. The intervascular tissue was digested with collagenase and dispase, and cells were sedimented, strained and plated. Although the resulting WJMSCs have not been characterized in detail (M. Kirkland, personal communication), we anticipate that the WJMSC culture contained primarily intervascular cells, with few, if any, perivascular cells. Cells were cultured for ~10 passages, DNA was isolated (see below) and sent to our laboratory.

Human epidermal keratinocytes (Invitrogen; www.invitrogen.com) isolated from adult skin were cultured in Epilife® Medium containing keratinocyte growth complement (Invitrogen). Cells used were under P5. Human foreskin fibroblasts (SkFib) from a 12 year old male (American Type Culture Collection; www.atcc.org) were cultured in DMEM/10% FCS. Passage number was not known but the cells were not senescent. Differentiated SGBS (Simpson-Golabi-Behmel syndrome) human adipocytes (Wabitsch *et al.*, 2001). Human umbilical vein cells (HUVECs) (Skovseth *et al.*, 2007) were a gift from Gutthorm Haraldsen (Department of Pathology, Rikshospitalet-Radiumhospitalet Medical Center).

Undifferentiated pluripotent NCCIT human embryonal carcinoma cells (American Type Culture Collection) were derived from a testicular germ-cell tumor and cultured in RPMI 1640/10% FCS (Taranger *et al.*, 2005). Human T cells were purified from peripheral blood (Skålhegg *et al.*, 1994) and donated by Heidi K. Blomhoff (Institute of Basic Medical Sciences, University of Oslo). A human muscle biopsy from the deltoid of a 12 year old male was provided by Gisèle Bonne (Institut de Myologie, Paris, France). The biopsy was collected as part of a muscle dystrophy project ongoing in Bonne's laboratory.

**Adipogenic, myogenic and endothelial differentiation**

For adipogenic differentiation, ASCs, BMMSCs and MPCs were cultured to confluency in DMEM/F12/10% FCS and stimulated for 3 weeks with 0.5 mM 1-methyl-3 isobutylxanthine, 1 μM dexamethasone, 10 μg/ml insulin and 200 μM indomethacin. Cells were stained with Oil Red-O to visualize lipid droplets (Boquest *et al.*, 2005). Quantification of labeling was performed by extraction of the dye and measurement of absorbance at 500 nm ($A_{500}$) in triplicate cultures.

For myogenic differentiation of MPCs, cells were grown to 70% confluency and cultured for 6 days in DMEM/F12 with 2% horse serum (Sigma-Aldrich) and 1% penicillin. For ASCs and BMMSCs, cells at ~70% confluency were cultured for 5 weeks in DMEM/F12/5% horse serum, 50 μM hydrocortisone and 1% penicillin (Zuk *et al.*, 2001). Nuclei were stained using Hemacolor® (Merck; www.merck.com). The resulting purple color of nuclei was due to the interaction between eosin Y and an azure B-DNA complex.

**Immunofluorescence**

Cells were cultured and/or differentiated into myocytes on acid-washed glass coverslips, fixed with 3% paraformaldehyde for 15 min and permeabilized with 0.1% Triton X-100 for

15 min. Preparations were washed 2 x 5 min in PBST (PBS/0.01% Tween-20), proteins blocked in PBST/2% BSA for 15 min. Myogenin was detected using anti-myogenin antibody F5D (Santa Cruz; sc-12732; www.scbt.com) diluted 1:100 in PBST/2% BSA, and a Cy3-conjugated anti-mouse IgG (Jackson ImmunoResearch; 15-165-044; www.jacksonimmuno.com) diluted 1:200 in PBST/2% BSA. DNA was counterstained with DAPI. Samples were observed on an Olympus BX51 epifluorescence microscope and images taken with an F-View CCD camera and Analysis 2.0 software (Soft Imaging System; ww.soft-imaging.net). For counting of myogenin-positive nuclei, at least 100 cells were counted in triplicates.

**Bisulfite genomic sequencing**

DNA was purified from cultured cells by phenol-chloroform-isoamylalcohol extraction and ethanol precipitation, and bisulfite-converted as described (Noer *et al.*, 2006) according to a procedure I have set up in the laboratory. MethylEasy™ and MethylEasy *Xceed*™ kits (Human Genetic Signatures; www.geneticsignatures.com) were used indifferently, the MethylEasy *Xceed*™ kit being faster than MethylEasy™. Converted DNA was amplified by PCR using primers designed with Methprimer (www.urogene.org/methprimer/index1.html). We designed primers to include the proximal promoter regions of four adipogenic genes (*LEP*, *LPL*, *PPARG2*, *FABP4*), one myogenic gene (*MYOG*) and one endothelial gene (*CD31/PECAM-1*). Primers to the 5' end of exon 1 of *MYOG* were also designed. Primer sequences and amplicon sizes are given in **Supplementary Table 1**. The regions examined contained the transcription start site (TSS) or were immediately upstream of the TSS (Ensembl; www.ensembl.org/Homo_sapiens) (see **Fig. R-1**). PCR conditions were 95°C for 7 min and 35-37 cycles of 95°C 1 min, 54/58°C 2 min and 72°C 2 min, followed by 10 min at

72°C. PCR products were cloned into *E. coli* by TOPO TA cloning (Invitrogen) and reverse-sequenced (MWG Biotech; www.mwg-iotech.com).

Methylation data are shown as filled (methylated CpG) or empty (unmethylated CpG) circles for each bacterial clone (rows). Each circle represents one CpG. Average methylation in promoter regions were compared pair-wise between cell types using a Fisher's exact test and two-tailed *P* values. Numbers of methylated cytosines for a given CpG were compared between cell populations using unpaired *t*-tests and two-tailed *P* values.

**Reverse transcription (RT)-PCR**

Quantitative RT-PCR (RT-qPCR) was carried from 0.5 μg total RNA (Qiagen RNeasy; www.qiagen.com) using the Iscript cDNA synthesis kit (BioRad) and IQ SYBR® Green (Noer *et al.*, 2006). PCR conditions were 95°C for 3 min and 40 cycles of 95°C for 30 sec, 60°C for 30 sec and 72°C for 30 sec. qPCR data were analyzed (Pfaffl, 2001) using *GAPDH* as normalization control. Alternatively, 30 cycles were performed (end-point PCR) and products were resolved by 1% agarose gel electrophoresis. RT-PCR primers used are listed in **Supplementary Table 1**.

**Methylated DNA immunoprecipitation (MeDIP)**

The MeDIP protocol was implemented in the laboratory as one of the aims of this work (**Fig. I-15**). The protocol was adapted from that of Weber et al. (Weber *et al.*, 2005; Weber *et al.*, 2007) and posted on the Epigenomic Network of Excellence website (http://www.epigenome-noe.net/researchtools/protocol.php?protid=33). The protocol has been written in detail for publication (**Supplementary Manuscript)**.

In short, genomic DNA was purified by phenol-chloroform-isoamylalcohol extraction and ethanol precipitation, and fragmented to ~200-1,000 bp (enriched in 300-500 bp

fragments) by sonication (see Results). 5-methylcytosine (5-mC)-enriched fragments were immunoprecipitated using anti-5mC antibodies (Eurogentec cat. no. BI-MECY-1000; www.eurogentec.com). Precipitated and input DNA was amplified by 14 PCR cycles using the WGA2 Whole Genome Amplification kit (Sigma-Aldrich) and amplified DNA was cleaned up using the Qiagen MiniElute PCR Purification kit (www.qiagen.com).

Following amplification, uniformity of fragment size distribution was again assessed by agarose gel electrophoresis and specificity of immunoprecipitation was verified by PCR using primers (**Supplementary Table 1**) to genes known to be unmethylated (housekeeping gene ubiquitin-conjugating enzyme E2B; *UBE2B*) or methylated (*H19* Imprinting Control Region; *H19*ICR) in somatic cells (see Results). For hybridization to microarrays, input DNA was labeled with Cy3 (green) and MeDIP DNA was labeled with Cy5 (red).

**Microarrays**

Cy5-labeled methylated DNA-enriched fragments and Cy3-labeled input DNA fragments were co-hybridized onto Nimblegen human HG18 RefSeq Promoter arrays (cat. no. C4226-00-01; www.nimblegen.com). Array design was built to cover over 24,000 human promoters, ranging from -2,200 to +500 bp relative to the TSS. Probes consisted of 385,000 50 to 85-mers tiled throughout non-repetitive genomic sequences at an average spacing of 100 bp. Repeat sequences (centromeres) were masked. Sequence source for the probes was the UCSC Genome Browser (http://genome.ucsc.edu/). MeDIP and input DNA labeling, hybridization and detection were performed using the services of Nimblegen.

**MeDIP data analysis**

Methylation data were delivered by Nimblegen as xls and gff files. The latter were visualized using the Nimblegen SignalMap data browser version 1.9. The browser enabled interpretation

of tracks that included: scaled $\log_2$ ratios ($\log_2$ MeDIP/input signals), *P*-values, peaks, transcript localization and orientation, TSS, tiled region and nucleotide number.

Signal intensity (raw) data: signal intensity data were extracted from the scanned images of each array using the NimbleScan extraction software. Signal intensities were saved in txt files, which is the raw data format for DNA methylation experiments.

Scaled $\log_2$ ratio data: each feature on the array has a corresponding scaled $\log_2$ ratio, which is the ratio of the input signals for the "MeDIP" and "input" samples co-hybridized onto the array. The $\log_2$ ratio was computed and scaled to center the ratio data on zero. Scaling was performed by subtracting the bi-weight mean for the $\log_2$ ratio values for all features on the array from each log2 ratio value. Scaled $\log_2$ ratios were provided in gff files.

*P*-value data: from the scaled $\log_2$ ratio data, a 750 bp window was placed around each consecutive probe and a one-sided Kolmogorov-Smirnov test was applied to determine if the probes were drawn from a significantly more positive distribution of intensity $\log_2$ ratios than those in the rest of the array (the Kolmogorov–Smirnov test is a goodness-of-fit test used to determine if two probability distributions differ, or, in the case of our data, whether a probability distribution (that of the $\log_2$ ratio) differs from a hypothesized distribution). The resulting score for each probe was converted into a $-\log_{10}$ *P*-value from the windowed test around that probe. *P*-value data files were in gff format.

Peak data: peak data files were generated from *P*-value data files, by detecting peaks by searching for at least 2 probes above a *P*-value minimum cut-off of 2. Peaks within 500 bp of each other were merged. Peak data files were in gff format.

CpG islands: to relate methylation profiles to CpG density, a CpG island track was provided. Peak height reflected CpG density in the island, while peak width defined the length of the CpG island.

Primary transcripts and TSS: positioning of primary transcripts, their orientation, and positioning of the TSS for each gene included on the array were provided.

The time course of this project and current limitations in access to bioinformatics have not allowed the analysis of methylation profiles deeper than what was enabled with SignalMap. The required bioinformatics expertise is being developed at the Norwegian Microarray Consortium (NMC; www.mikromatrise.no) in collaboration with our laboratory using the data generated for this project.

# RESULTS

### Lineage-specific promoter regions examined by bisulfite sequencing

We designed bisulfite-converted DNA-specific PCR primers to include the proximal promoter regions of four adipogenic genes (*LEP*, *LPL*, *PPARG2*, *FABP4*), one myogenic gene (*MYOG*) and one endothelial gene (*CD31*) (**Fig. R-1**).

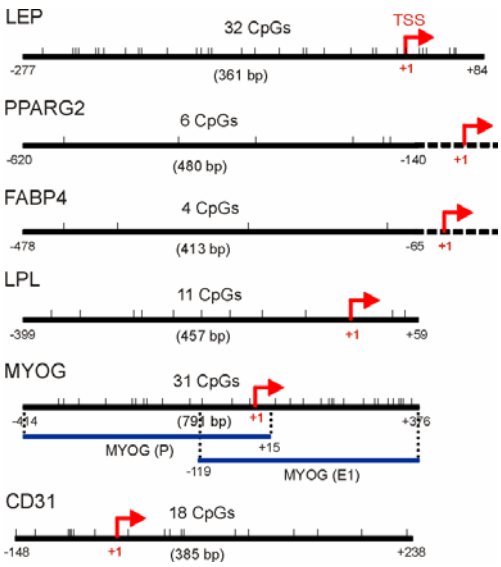



**Fig. R-1.** Genomic regions examined by bisulfite sequencing. TSS, transcription start site (+1). Numbers indicate nucleotide positioning relative to TSS. For *MYOG*, two amplicons were generated, overlapping 3 CpGs. Length of the regions examined is indicated in bp.

In the *LEP* (GenBank U43589) promoter, we analyzed 32 CpGs covering 361 bp ranging from nucleotides -277 to +84 relative to the TSS (+1). This region has been shown to be regulated by DNA methylation (Melzner *et al.*, 2002). In the *PPARG2* (GenBank AB005520) promoter, 6 CpGs were examined between nucleotides -620 to -149 relative to the TSS. The *FABP4* (GenBank NM_001442) promoter region examined included 4 CpGs within nucleotides -478 to -65 upstream of the TSS. The *LPL* (GenBank X68111) promoter region examined spanned nucleotides -399 to +59 relative to the TSS and included 11 CpGs. The *MYOG* (GenBank X62155) region spanned nucleotides -414 to +376 relative to the TSS and included 31 CpGs. This region was covered by two amplicons (*MYOG*(P) [promoter] and *MYOG*(E1) [exon 1]) which shared 3 overlapping CpGs (**Fig. R-1**). We randomly chose to show all CpGs covered by *MYOG*(P) and remove the three 5'-most CpGs from *MYOG*(E1); these are shown in **Supplementary Figure 1**. The *CD31* (GenBank X96848) promoter region examined spanned nucleotides -148 to +238 relative to the TSS and included 18 CpGs.

**CpG methylation profiles of adipogenic promoters in ASCs, BMMSCs, MPCs, WJMSCs and HSCs**

We have previously established by bisulfite sequencing the DNA methylation profile of adipogenic promoters (*LEP*, *LPL*, *FABP4* and *PPARG2*) in freshly isolated and cultured ASCs, to show that these promoters are largely unmethylated (Noer *et al.*, 2006). Using a different polyclonal culture of ASCs, we corroborated, also by bisulfite sequencing, the hypomethylated state of these promoters (**Fig. R-2A**). We then compared these methylation profiles to those of BMMSCs and MPCs. Methylation profiles of *LEP*, *FABP4* and *PPARG2* were similar in ASCs and BMMSCs and revealed mosaic hypomethylation (**Fig. R-2A**). Four CpGs (No. 1, 2, 11, 20; No. 1 being the 5' most CpG examined) in *LEP* were more methylated in ASCs than in BMMSCs ($P<0.001$). Methylation in this region was also observed in our earlier work (Noer *et al.*, 2006).

In MPCs, the *LEP* promoter was also unmethylated, consistent with its localization in a CpG island (**Fig. R-2A; Supplementary Table 2**). *FABP4* was strongly methylated in MPCs while *PPARG2* was more methylated than in ASCs. In all promoters and in all cell types, methylation was mosaic between alleles, corroborating our previous findings (Noer *et al.*, 2006) and reflecting the individual history of each cell in culture. We also examined the methylation of the same promoters in WJMSCs and HSCs. In both cell types, *FABP4* and *PPARG2* were hypermethylated relative to ASCs and BMMSCs whereas *FABP4* methylation was similar to MPCs (**Fig. R-2A**; **Supplementary Table 2**). *LEP* was essentially unmethylated in WJMSCs but was hypermethylated in HSCs relative to all other cell types (**Fig. R-2A**; **upplementary Table 2**).

Comparison of percentages of methylation (**Fig. R-2B**; **Supplementary Table 2**) indicates that *LEP*, *FAPB4* and *PPARG2* were globally more methylated in MPCs and HSCs

than in ASCs or BMMSCs, both of which displayed similar methylation percentages. *FABP4* and PPARG2 were also more methylated in WJMSCs than in ASCs and BMMSCs. The *LPL* promoter was unmethylated in all cell types examined, also consistent with its localization in a CpG island. Extent and pattern of CpG methylation in adipogenic promoters, therefore, may differ depending on the cell type examined.

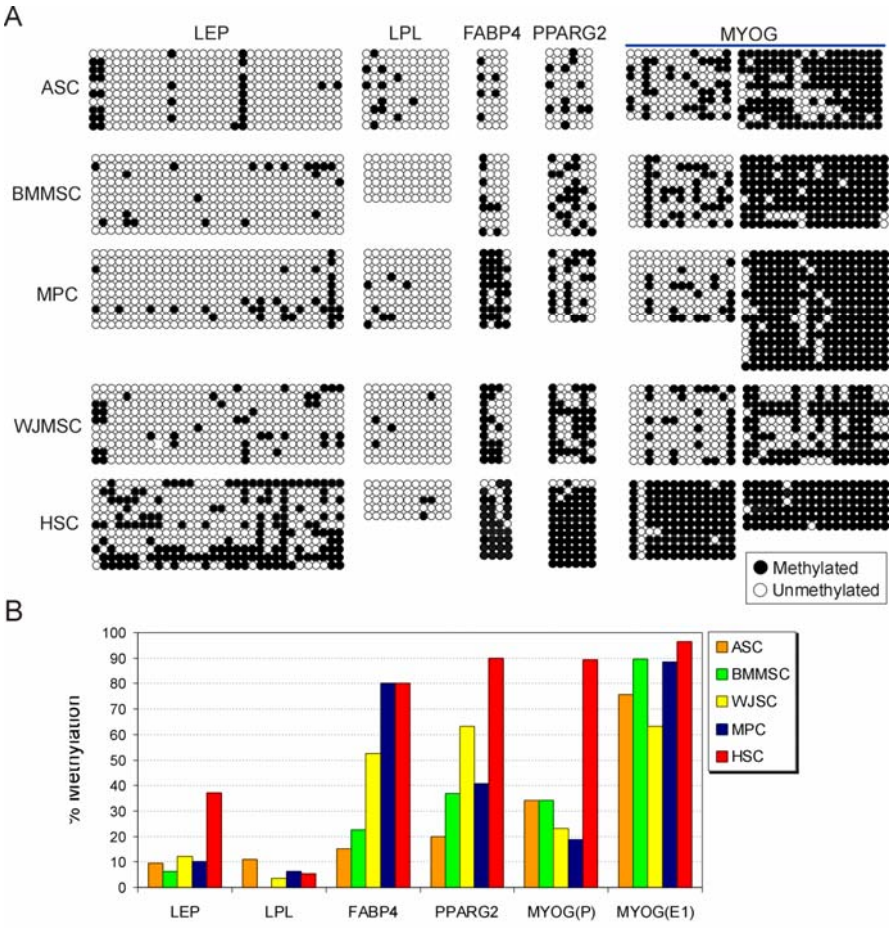



**Fig. R-2.** DNA methylation profile of adipogenic and myogenic promoters in ASCs, BMMSCs, MPCs, WJMSCs and HSCs. (A) Bisulfite sequencing analysis of CpG methylation in each promoter and on *MYOG* exon 1. Each dot depicts one CpG in the 5'–3' order and each row of dots represents one bacterial clone (i.e., one genomic allele). Positioning of each CpG in each region is shown in Fig. R1. (B) Percentage of methylated CpGs in each region and for each cell type, determined from data in (A). Statistical comparisons within genes and between cell types are shown in Supplementary Table 2.

### Methylation profiles in the *MYOG* promoter and 5' end of exon 1 of *MYOG*

CpG methylation of the *MYOG* promoter (*MYOG*(P)) was mosaic and similar in ASCs and BMMSCs **(Fig. R-2A)**. *MYOG*(P) was less methylated in MPCs and WJMSCs, but was

nearly 100% methylated in HSCs, except for CpG No. 2 which was unmethylated as in all other cell types (**Fig. R-2A,B; Supplementary Table 2)**.

Interestingly, the *MYOG*(E1) region examined was consistently more methylated than the proximal promoter (**Fig. R-2A**). *MYOG*(E1) was strongly methylated in ASCs, BMMSCs, MPCs and HSCs, but displayed regions of low methylation in 60% of the alleles in WJMSCs (**Supplementary Table 2**). Furthermore, comparison of sequencing data for the three overlapping CpGs in the *MYOG*(P) and *MYOG*(E1) amplicons (**Supplementary Fig. 1A**) showed that methylation was overall conserved, although some variation was detected for some CpGs in ASCs, WJMSCs (and fibroblasts) (**Supplementary Fig. 1B**). Thus, the *MYOG* promoter is more methylated in ASCs and BMMSCs than in MPCs and WJMSCs, and is nearly fully methylated in HSCs. One CpG (-335 nt from the TSS), however, is consistently unmethylated in all cell types.

**Methylation patterns are conserved in *in vitro* differentiated ASCs and SGBS adipocytes**

To determine the extent of epigenetic commitment of ASCs to adipogenesis, we compared the methylation profiles of differentiated ASCs to those of SGBS adipocytes. Methylation profiles and extent of methylation were similar for adipogenic and myogenic genes in undifferentiated ASCs, differentiated ASCs and SGBS adipocytes (**Fig. R-3A-C**). As suggested earlier (Noer *et al.*, 2006), the hypomethylated state of adipogenic promoters may reflect a pre-programming of ASCs to adipogenesis.

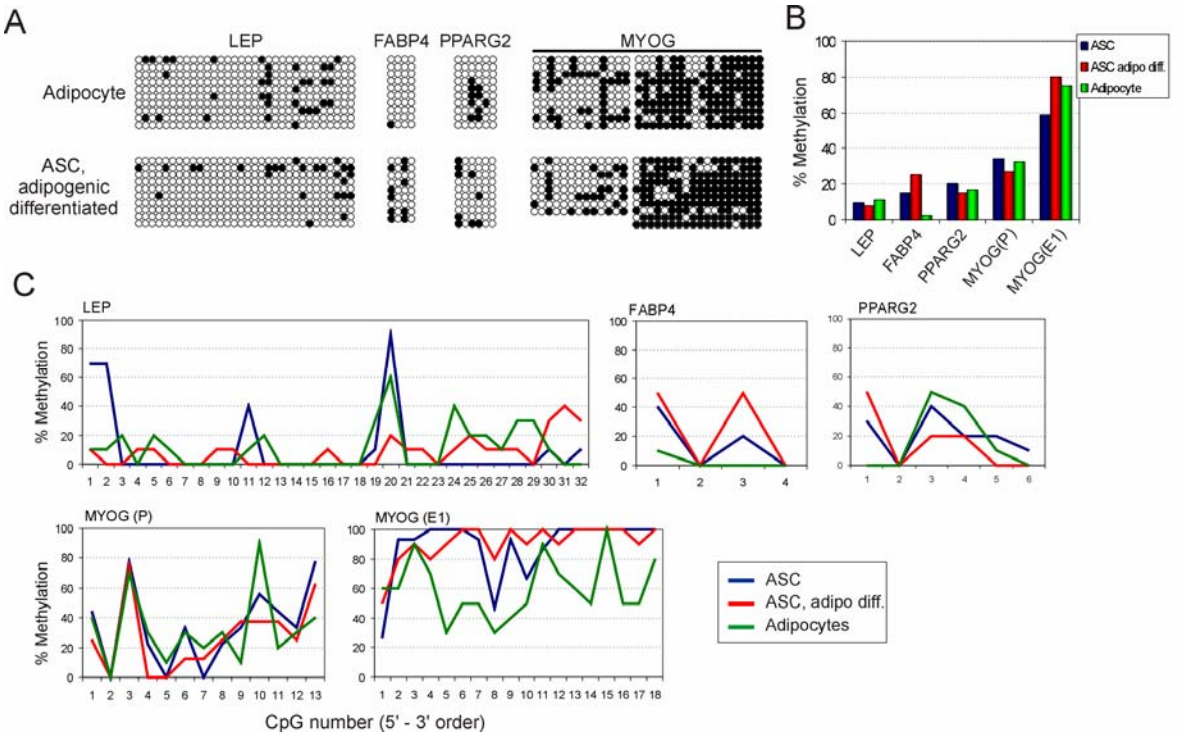



**Fig. R-3.** CpG methylation profiles in undifferentiated ASCs, *in vitro* adipogenic-differentiated ASCs and SGBS adipocytes. (A) Bisulfite sequencing data for SGBS adipocytes and adipogenic-differentiated ASCs, at indicated loci. (B) Percentage of methylation in each region and for each cell type, determined from (A). (C) Methylation profile across the regions examined, determined from (A).

### *In vitro* adipogenic differentiation capacity of ASCs, BMMSCs and MPCs

To put the methylation profiles identified in MSCs in a functional context, we determined the adipogenic differentiation capacity of ASCs, BMMSCs and MPCs. Cells were stimulated for three weeks toward adipogenesis and lipid droplets were stained with Oil Red-O. Both ASCs and BMMSCs efficiently differentiated into adipocytes, as shown by most cells containing lipid droplets (**Fig. R-4A)**. In contrast, MPCs showed poor adipogenic differentiation. Quantification of Oil Red-O staining supported the visual picture (**Fig. R-4B**) and confirmed the strong adipogenic potential of ASCs and BMMSCs, but not of MPCs.

Adipogenic differentiation was also evaluated by RT-PCR analysis of expression of adipogenic genes (*LEP, LPL, FABP4, PPARG2*) and, as a control for lineage-specificity, myogenic genes (*PAX7, PAX3, MYF5, MYOD1, MYOG)* after adipogenic induction. The adipogenic genes were not expressed at detectable levels in any of the undifferentiated cell

types (**Fig. R-4C**). Upon adipogenic differentiation, *LPL, FABP4* and *PPARG2* were upregulated (*PPARG2* to a smaller extent) in ASCs and BMMSCs (**Fig. R-4D**). *FABP4* and *LPL* transcripts (*LPL* at low level) were also detected in MPCs (**Fig. R-4D**), reflecting the differentiation of a minor proportion of the cells (see **Fig. R-4A**). Upregulation of *LEP* was not detected in any cell type. None of the myogenic genes examined were expressed in undifferentiated ASCs and BMMSCs and they were not induced after adipogenic differentiation (**Fig. R-4C,D**). *MYF5, MYOD* and *MYOG* were expressed in undifferentiated MPCs and their expression was maintained after adipogenic stimulation. *PAX7* and *PAX3* were not expressed in undifferentiated MPCs but surprisingly, both were upregulated in adipogenic-stimulated cells, though at low levels (**Fig. R-4C,D**). As expected, PCR without RT to control of DNA contamination using *MYOG* and *GAPDH* primers gave no products **(Fig. R-4C**, -RT**)**.

Expression of adipogenic genes was also examined by quantitative RT-PCR for ASCs and BMMSCs, and confirmed end-point RT-PCR data (**Fig. R-4E**). *LPL* and *FABP4* were strongly upregulated after adipogenic stimulation while *PPARG2* showed more modest upregulation. In addition, *LEP* expression was detected by quantitative RT-PCR and showed a 4- to 5-fold upregulation after adipogenic stimulation; however CT values were over 30, indicative of low mRNA levels (data not shown) and accounting for the lack of *LEP* mRNA detection by end-point RT-PCR. Overall, our results indicate that ASCs and BMMSCs exhibit strong adipogenic potential *in vitro*, in contrast to MPCs.

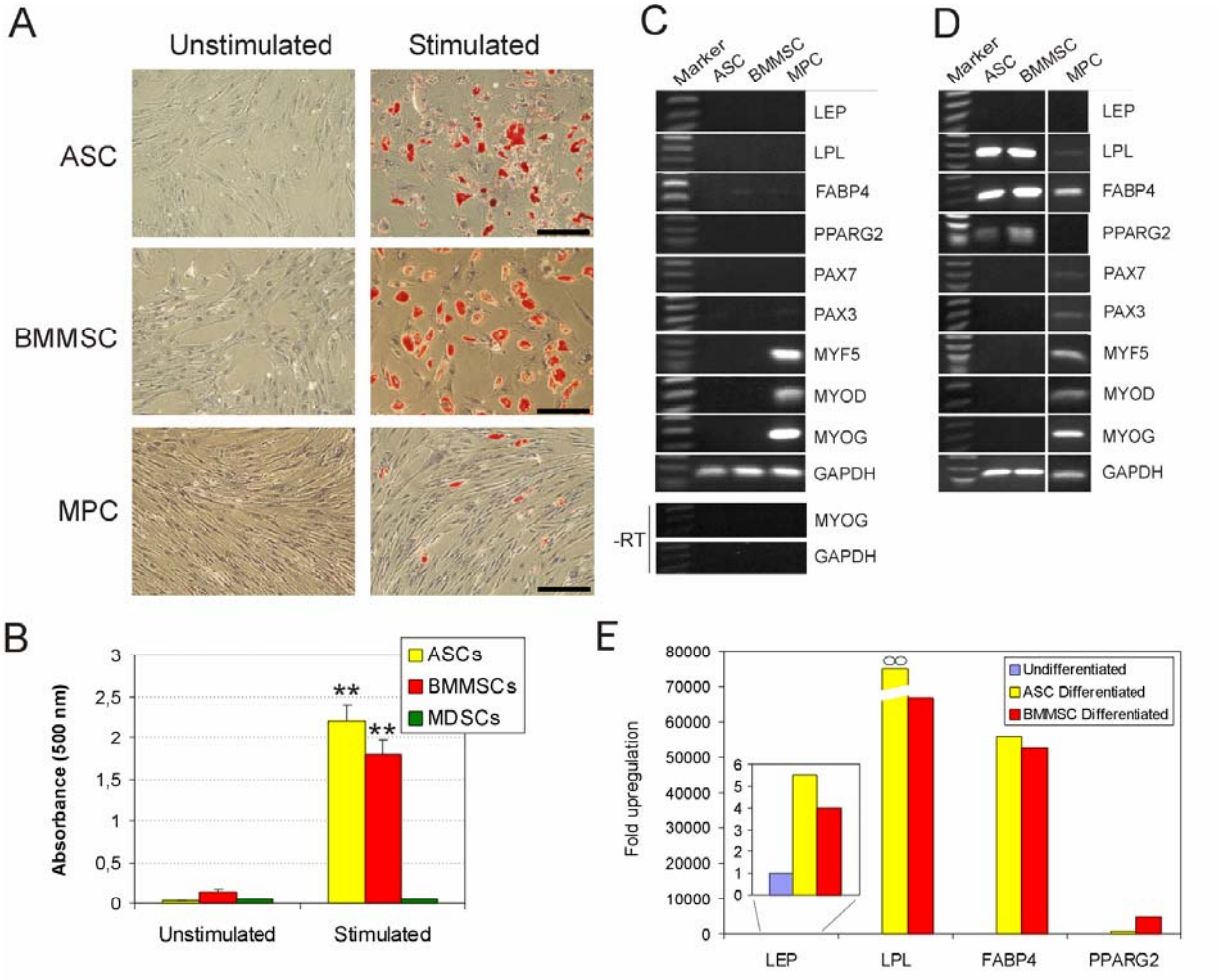



**Fig. R-4.** *In vitro* adipogenic differentiation of ASCs, BMMSCs and MPCs. (A) Oil Red-O staining of intracellular lipid droplets in ASCs, BMMSCs and MPCs before and after three weeks of adipogenic stimulation. Bars, 100 µm. (B) Quantification of Oil Red-O incorporation in undifferentiated cells and after adipogenic differentiation (mean±SD $A_{500}$ of a triplicate experiment; $^{**}P<0.01$ relative to undifferentiated cells; *t*-test). (C) RT-PCR analysis of gene expression in undifferentiated ASCs, BMMSCs and MPCs. –RT, no reverse transcription reaction control. (D) RT-PCR analysis of gene expression in ASCs, BMMSCs and MPCs after adipogenic stimulation as in (A). In (C) and (D), size markers are shown. (E) Real-time RT-PCR analysis of expression of indicated genes in adipogenic differentiated ASCs and BMMSCs, relative to their undifferentiated counterparts (level 1).

## In vitro myogenic differentiation of ASCs, BMMSCs and MPCs

The predicted natural differentiation pathway of MPCs is myogenesis. Myogenic potential of MPCs was shown by their ability to differentiate into elongated multinucleated cells after 6 days of myogenic sitmulation in horse serum (**Fig. R-5A**, Day 6). Control MPCs cultured for 6 days in proliferation medium (containing FCS instead of horse serum) also showed an

elongated growth pattern, but there were no multinucleated cells (**Fig. R-5A**, Day 6, control). In contrast to MPCs, ASCs and BMMSCs stimulated under the same conditions showed no signs of myogenic differentiation after 5 weeks (**Fig. R-5B**).

We next analyzed myogenin protein expression by immunofluorescence using anti-myogenin antibodies **(Fig. R-5C)**. Undifferentiated MPCs, ASCs and BMMSCs did not express myogenin. Differentiation however induced myogenin expression in ~60% of nuclei ($P$<0.001 relative to Day 0; **Fig. R-5C**) in MPCs. Day 6 control MPCs also displayed some myogenin expression, although in a lower proportion of nuclei (28%; $P$<0.002; **Fig. R-5D**). We observed no myogenin expression in stimulated ASCs and BMMSCs (**Fig. R-5C**), supporting the view that these MSC types have poor myogenic potential, at least under the conditions tested.

To further evaluate the extent of myogenic differentiation of MSCs, we performed RT-PCR analysis of adipogenic and myogenic gene expression (**Fig. R-5E**; mRNA levels in undifferentiated cells are shown in **Fig. R-4C**). Undifferentiated MPCs did not express *PAX7* or *PAX3* but expressed *MYF5*, *MYOD* and *MYOG*. The same transcript profiles were detected after myogenic stimulation (**Fig. R-5E**). Quantitative RT-PCR confirmed these data and showed small changes in gene expression upon myogenic differentiation **(Fig. R-5F)**. In addition, ASCs and BMMSCs stimulated toward myogenesis upregulated *LEP* and *LPL* transcripts (*FABP4* was also upregulated in BMMSCs), but not muscle-specific markers **(Fig. R-5E)**. This may reflect a default differentiation pathway for these cells, possibly induced by confluency, as ASCs and BMMSCs grown to confluency spontaneously differentiate into adipocytes (data not shown). No adipogenic transcripts were detected in differentiated MPCs. Together with the phenotypic observations, these results argue that whereas MPCs have myogenic differentiation capacity *in vitro*, ASCs and BMMSCs are refractory to differentiation into myogenin-expressing multinucleated cells.

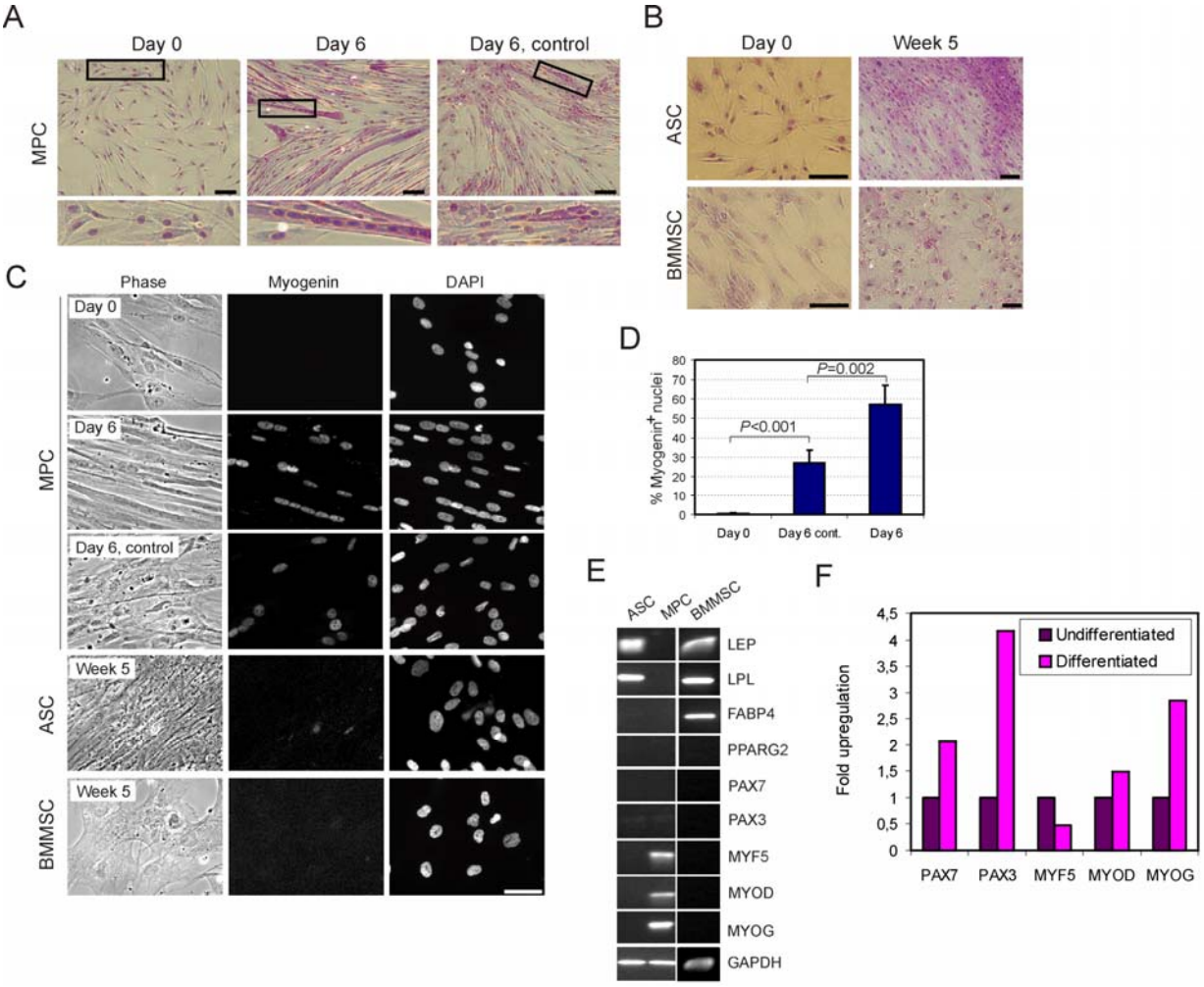



**Fig. R-5.** *In vitro* myogenic differentiation of MPCs, ASCs and BMMSCs. (A) Myogenic differentiation of MPCs. Cells at day 0 (prior to induction of differentiation), day 6 in differentiation medium and day 6 in control (proliferation) medium were stained with Hemacolor. Boxed areas are enlarged (bottom). Note the multinucleated cells at day 6 of stimulation. Bars, 100 μm. (B) Lack of myogenic differentiation of ASCs and BMMSCs after 5 weeks in myogenic differentiation medium; cells were stained with Hemacolor. Bars, 100 μm. (C) Immunofluorescence analysis of myogenin expression in MPCs, ASCs and BMMSCs induced to differentiate. Time points of analysis are indicated. Nuclei were counterstained with DAPI. Bar, 50 μm. (D) Proportion of myogenin-positive nuclei in myogenic-differentiated MPCs (n=100 per time point in each of a triplicate area). (E) RT-PCR analysis of expression of indicated genes after myogenic stimulation of ASCs (5 weeks), BMMSCs (5 weeks) and MPCs (6 days). (F) Real-time RT-PCR analysis of expression of indicated genes in differentiated versus undifferentiated MPCs (level 1).

## DNA methylation patterns of adipogenic and myogenic promoters are similar in *in vitro* differentiated MPCs and muscle DNA

To relate the DNA methylation patterns of adipogenic and myogenic genes in myogenic-differentiated MPCs to that of muscle, bisulfite sequencing was performed on myogenic-differentiated MPCs and on DNA isolated from a human deltoid biopsy from a 12-year old

male. **Figures R-6A-C** show that the *LEP*, *FABP4*, *PPARG2* and *MYOG* promoters displayed similar methylation in both cell and tissue types. However, *MYOG*(E1) was hypermethylated in differentiated MPCs relative to muscle DNA, which showed a mix of hypomethylated and methylated alleles (**Fig. R-6A,C**). This suggests incomplete myogenic differentiation of MPCs *in vitro*. Methylation profiles of undifferentiated MPCs, myogenic-differentiated MPCs and muscle DNA showed that differentiation maintained the *MYOG* methylation profile of undifferentiated cells (**Fig. R-6C**). These results suggest nonetheless that undifferentiated MPCs display epigenetic commitment to myogenesis, reflected by their DNA methylation pattern on the promoters examined.

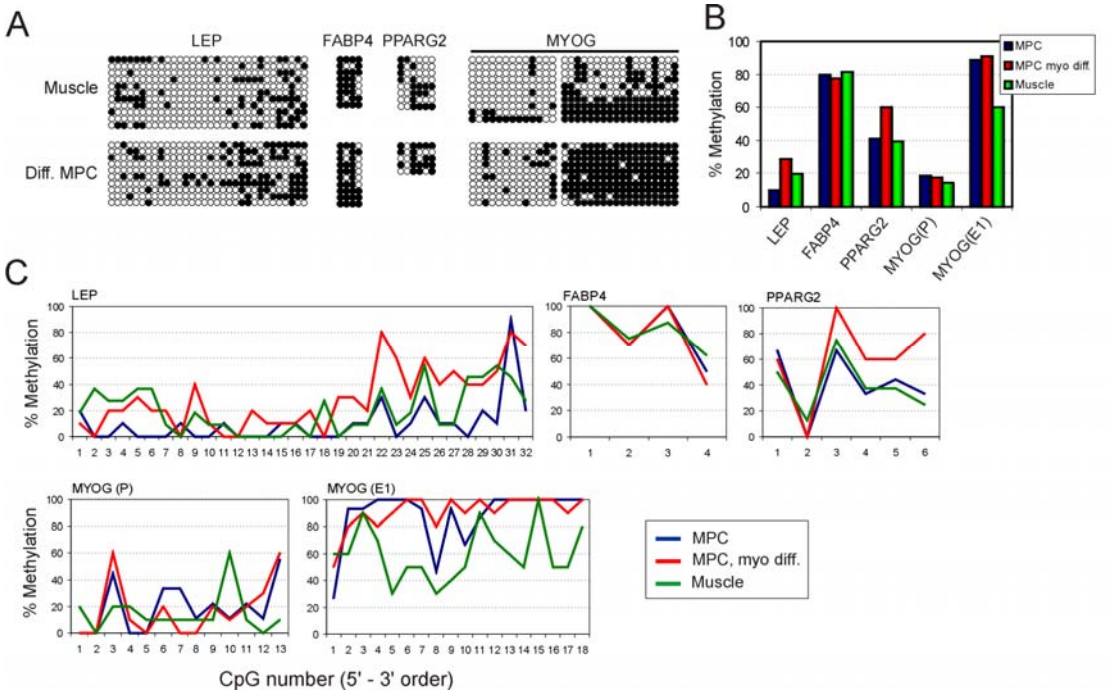



**Fig. R-6.** CpG methylation in undifferentiated MPCs, myogenic-differentiated MPCs and in a muscle biopsy. (A) Bisulfite sequencing data for muscle and myogenic-differentiated MPCs. (B) Percentage of methylated CpGs in each region, for each cell type, determined from (A). (C) Profile of CpG methylation determined from (A).

### Adipogenic and myogenic promoters show distinct CpG methylation patterns in keratinocytes and in skin fibroblasts

Keratinocytes and foreskin fibroblasts are differentiated cells with no direct functional link to adipogenesis or myogenesis. Thus, to gain further insight on the lineage-specificity of the

methylation patterns observed in MSCs and HSCs, we examined the methylation of *LEP*, *FABP4*, *PPARG2* and *MYOG* in keratinocytes and fibroblasts. These cells revealed different profiles (**Fig. R-7**). *LEP* was unmethylated in both cell types except for two CpGs methylated in fibroblasts. *FABP4* was largely methylated in both cell types. *PPARG2*, however, was methylated in keratinocytes but largely unmethylated in fibroblasts. *MYOG*(P) and *MYOG*(E1) were hypermethylated in keratinocytes relative to fibroblasts. *MYOG* displayed a profile different from all other cell types examined in this study, with regions of hypermethylation in the promoter and low methylation of the first 9 CpGs in exon 1 (**Fig. R-7A,C**). The relatively weaker promoter methylation in fibroblasts was not indicative of adipogenic or myogenic differentiation capacity (data not shown).

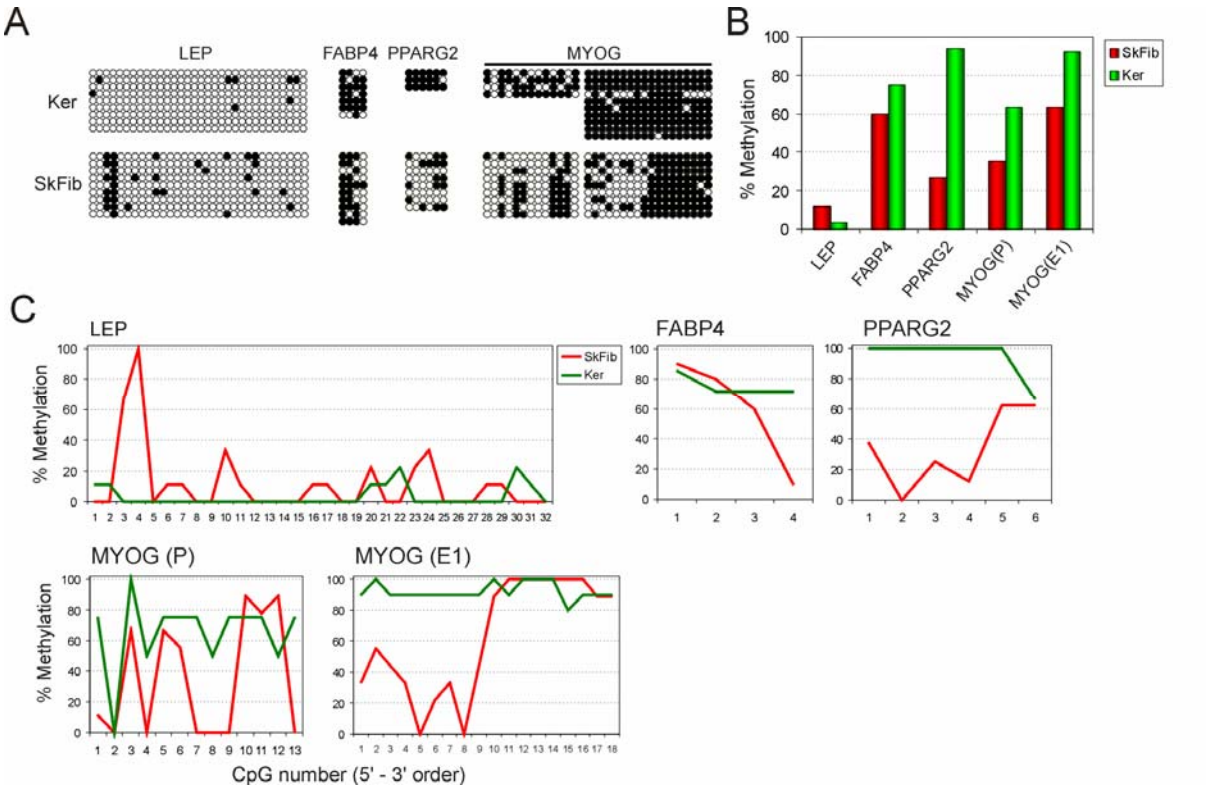



**Fig. R-7.** Methylation profiles in cultured primary keratinocytes and skin fibroblasts. (A) Bisulfite sequencing data for keratinocytes (Ker) and skin fibroblasts (SkFib). (B) Percentage of methylated CpGs determined from (A). (C) Profile of CpG methylation determined from (A).

**The endothelial *CD31* locus is methylated in MSCs and unmethylated in HSCs**

To ascertain the specificity of the methylation profiles obtained so far, we examined the methylation pattern in the endothelial *CD31* proximal gene promoter and in the first 238 nt of exon 1 in MSCs, HSCs and in differentiated cells. *CD31* is not known to be involved in adipogenic or myogenic differentiation, and thus was used as an "irrelevant" gene. The *CD31* proximal promoter (**Fig. R-1**) was methylated in ASCs, BMMSCs, MPCs and WJMSCs (**Fig. R-8A,B**; **Supplementary Table 2**), BMMSCs revealed 5 unmethylated CpGs in exon 1. In contrast to MSCs, *CD31* was unmethylated in HSCs, suggesting a different chromatin organization on this promoter in this cell type (**Fig. R-8A,B**). In differentiated cells, *CD31* was methylated in SGBS adipocytes, muscle, fibroblasts and T cells, although unmethylated alleles were also detected in T cells, reminiscent of the unmethylated state of HSCs (**Fig. R-8**). Lastly, we verified that *CD31* was largely unmethylated in HUVEC cells (**Fig. R-8**). These data were in agreement with our earlier results (Boquest *et al.*, 2007), which also showed that *CD31* appeared as hemimethylated in $CD31^+$ cells isolated from the stromal vascular fraction of human lipoaspirates (Boquest *et al.*, 2007) (see **Fig. R-8A**). This analysis indicates therefore that the *CD31* promoter is methylated in MSCs and unmethylated in HSCs ($P<0.001$; **Fig. R-8C; Supplementary Table 2**). The unmethylated state of the *CD31* promoter in HSCs reflects a chromatin organization which is suggestive of endothelial differentiation potential, a possibility currently being tested in our laboratory. The unmethylated state of 5 CpGs in exon 1 in BMMSCs also suggests a distinct chromatin composition at this site, and perhaps also a potential for endothelial differentiation.

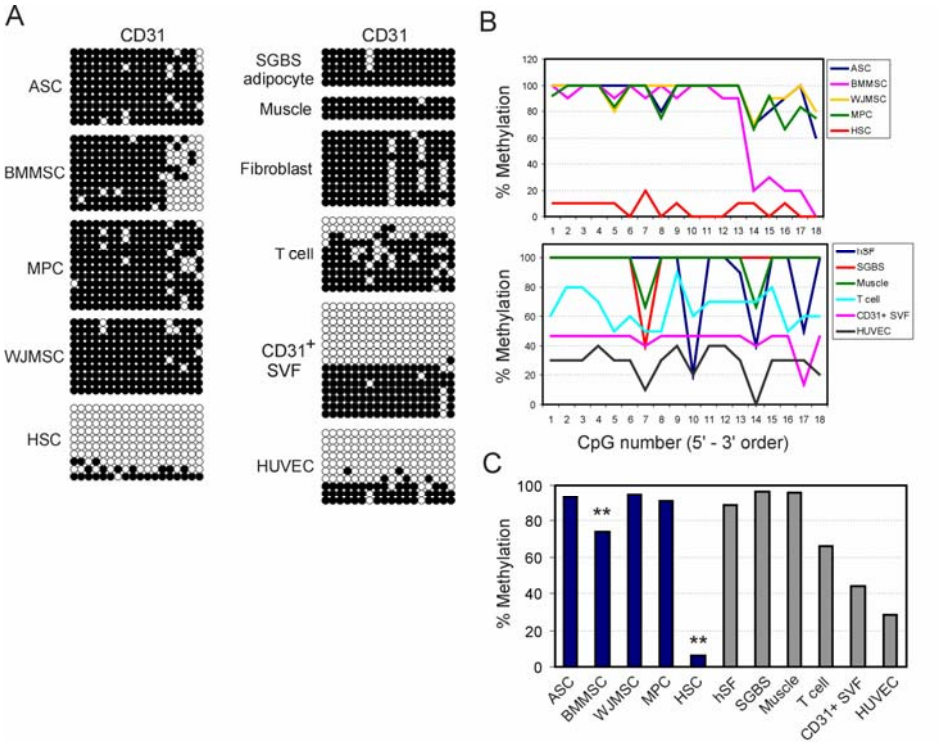



**Fig. R-8.** Methylation profile of *CD31*. (A) Bisulfite sequencing analysis of *CD31* in indicated cell types. (B) Percentage of methylation of each CpG determined from data in (A). (C) Percentage of methylation in MSCs (blue) and differentiated cells (gray). ** $P<0.01$ (Fisher's exact tests) relative to the other cell types. Methylation profile of $CD31^+$ SVF cells was taken from Boquest *et al*. (2007).

**Genome-wide DNA methylation profiling: establishment of a methylated DNA immunoprecipitation (MeDIP) assay**

To obtain a more global picture of the DNA methylation profile of ASCs, BMMSCs, MPCs and HSCs, we implemented a genome-wide strategy. We chose MeDIP because it can be combined with DNA hybridization to promoter arrays (Weber *et al.*, 2005) which our laboratory also uses in ChIP-on-chip experiments. Genomic DNA was fragmented and methylated fragments were immunoprecipitated using anti-5mC antibodies. Detection of genomic regions of interest in the methylated DNA fraction was done by PCR (for validation) and by hybridization to Nimblegen promoter arrays. The MeDIP procedure is outlined in **Figure I-15** and detailed in the **Supplementary Manuscript**.

We performed several quality control tests as indicated in **Figure I-15**. First, we assessed sonication efficiency, which varies with DNA concentration, sonicator model (probe or bath), sonicator settings and time of sonication, and ensured homogenous fragment sizes between samples. Electrophoresis analysis shows that ASCs, BMMSCs and MPCs uniformly

showed high molecular weight DNA before sonication, and fragments of ~300-1,000 bp after sonication (**Fig. R-9A**).

Following immunoprecipitation, the yield of MeDIP DNA is low (~300-450 ng) and incompatible with array hybridization (which requires 2 μg DNA per array). Thus, a DNA amplification step of MeDIP and input DNA fractions was performed. To ensure uniformity of MeDIP and input DNA fragment sizes after amplification, electrophoresis analysis was also carried out. Both input and immunoprecipitated DNA displayed the recommended fragment sizes (~200-1,000 bp) (**Fig. R-9B**).

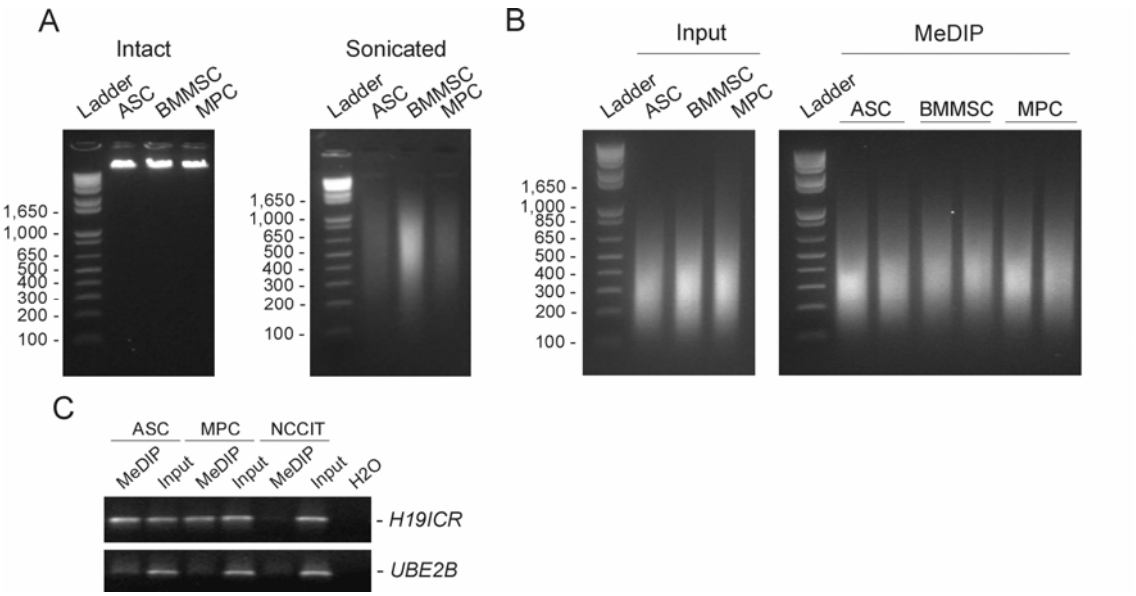



**Fig. R-9.** Quality assessment of DNA in the MeDIP procedure. (A) Assessment of DNA fragmentation by sonication. Intact and sonicated DNA from ASCs, BMMSCs and MPCs were analyzed by agarose gel electrophoresis. (B) Assessment of input and MeDIP DNA fragment size and uniformity after amplification. Input and MeDIP DNA samples from ASCs, BMMSCs and MPCs were amplified, purified and resolved by agarose gel electrophoresis. (C) PCR analysis of MeDIP specificity. Input and MeDIP DNA were analyzed by PCR using primers for the *H19* Imprinting Control Region (*H19*ICR), which is methylated in somatic cells but not in male germ-cell-derived embryonal carcinoma NCCIT cells, and for the *UBE2B* housekeeping gene promoter, which is unmethylated (Weber *et al.*, 2007). DNA from ASCs, BMMSCs and NCCIT cells was used.

Next, we verified by PCR that methylated DNA fragments were effectively immunoprecipitated by MeDIP. The imprinted *H19*ICR gene locus served as methylated control, and the housekeeping *UBE2B* gene promoter served as unmethylated control (Weber *et al.*, 2007). Amplified MeDIP and input DNA from ASCs, MPCs and embryonal carcinoma

NCCIT cells (derived from a male germ-cell tumor and in which *H19*ICR was expected to be unmethylated), were tested. ASC and MPC samples gave *H19*ICR products, reflecting the methylated state of this locus in these cells **(Fig. R-9C)**. However, MeDIP from NCCIT cells did not precipitate the *H19*ICR locus (Fig. R9C), confirming its unmethylated state in this cell type. *UBE2B* PCR was negative for all MeDIP samples, as expected from the unmethylated state of this housekeeping promoter (Weber *et al.*, 2007). These observations indicate that DNA fragment sizes are uniform between replicates and cell types, and suggest that the immunoprecipitation is specific.

**Preliminary analysis of MeDIP results: introduction to data generated through NimbleScan and SignalMap**

Input and duplicate MeDIP DNA samples from ASCs, BMMSCs, MPCs and HSCs (**Supplementary Fig. 2**) were sent to Nimblegen for labeling and hybridization to promoter arrays. The NimbleGen SignalMap data browser enabled a visualization of the IP/Input $\log_2$ ratio, *P*-values (expressed in –log10) computed from these ratios and methylation peak data generated from the *P*-values (see Materials and Methods).An overview of the information for chromosome 1 in one ASC sample is shown in **Figure R-10A** and a random segment shown in **Figure R-10B**. The genomic position of the regions examined is marked on top. Track 1 displays the IP/input $\log_2$ ratio, with each bar representing the hybridization signal to one probe relative to the input (the 0 line). The ratios, therefore, reflect areas of hypomethylation (below the 0 line) and hypermethylation (above the 0 line) relative to genome average. From the $\log_2$ ratios, a statistical test reports the significance of positive enrichment (methylation), surrounding each probe within a 750 bp window, against all other probes on the array. The resulting score for each probe is the *P*-value, shown in track 2. Peak data (track 3) are generated from the *P*-value data by searching for at least 2 probes above a *P*-value of 2.

Peaks within 500 bp of each other are merged, which explains why there are fewer peaks than methylated regions detected by the *P*-value. Positions of transcripts are localized on track 4 with the TSS on track 5. CpG islands are marked on track 6, with the height of the peak reflecting CpG density and the width the area covered by the island. The tiled regions on the array are positioned on track 7. The probe-less area between nucleotides 122,000,000 and 142,000,000 (**Fig. R-10A**) represents the centromere (no probes were made to DNA repeats).

Several methylation profiles were observed in ASCs, with, as expected, an anti-correlation between methylated areas and CpG islands (**Fig. R-10B**). The promoter regions shown displayed either strong or weak methylation, ith consistent unmethylation or hypomethylation over 200-400 bp around the TSS **(Fig. R-10B**). Unmethylated regions corresponded to CpG islands (**Fig. R-10B**), however unmethylated areas were also outside CpG islands (**Supplementary Fig. 3**). The different methylation patterns observed, on all chromosomes and each cell type (data not shown), reinforce the view that the MeDIP was specific.

**DNA methylation profiles are similar between MeDIP replicates, and between ASCs, BMMSCs, MPCs and HSCs**

To evaluate the technical reproducibility of the MeDIP procedure, duplicate MeDIPs were carried out from a single DNA preparation from ASCs, BMMCs, MPCs and HSCs. *P*-values on chromosome 1 showed similar methylation profiles for each replicate, arguing that MeDIP was reproducible (**Fig. R-11A**). Similar observations were made for all chromosomes (not shown). A scatter-plot analysis of reproducibility, currently being performed at the NMC (Andy Reiner), will provide a more robust indication of reproducibility.

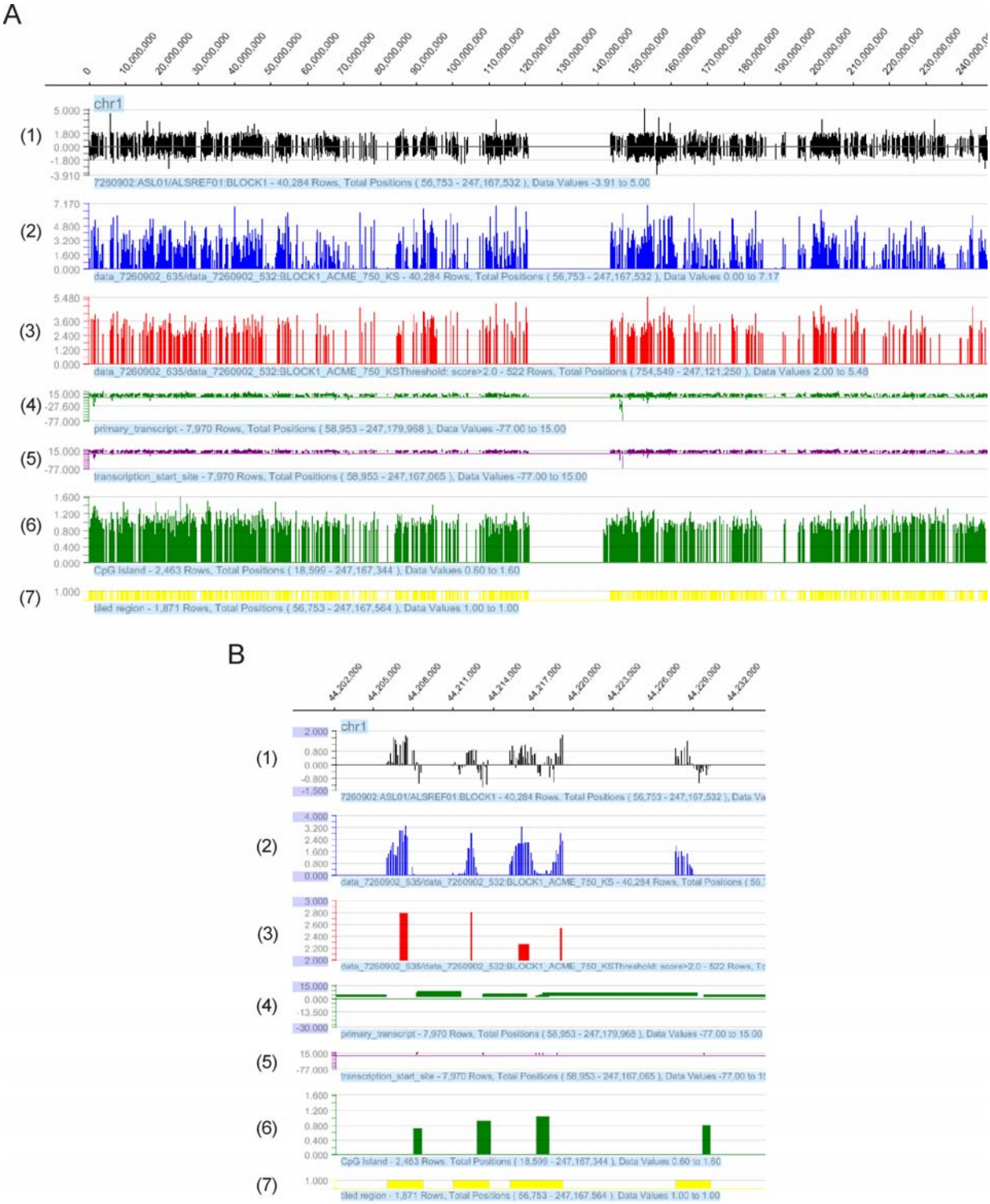



**Fig. R-10.** SignalMap data from a MeDIP from ASCs, for chromosome 1. (A) DNA methylation profile of ASCs. Top, chromosome positioning. Track (1): scaled IP/input $\log_2$ ratios. Track (2): *P*-values, expressed in –log10. Track (3): methylation peaks. Track (4): primary transcripts. Track (5): TSS. Track 6,: CpG islands. Track (7): tiled regions. The "blank" area covers the centromere and is devoid of probes. (B) Methylation profile for random 30-kb segment of chromosome 1. The same tracks as in (A) are shown. The genes represented are *DPH2*, *ATPF6*, *B4GALT2* and *CCD24* (left to right).

The methylation peaks, which provide a more stringent account of methylation, also revealed similarity between MeDIP replicates (**Fig. R-11B**). Yet, the profiles were not identical and some areas showed different peak profiles between replicates (e.g., region 160,000,000 to 180,000,000 in MPC-1 and MPC-2; **Fig. R-11B**). However, $-\log_{10}$ *P*-values (**Fig. R-11A**) for the same region showed that the profiles were nearly identical, albeit with slightly lower values for MPC-1 than for MPC-2. As methylation peaks are defined from $-\log_{10}$ *P*-values $\geq 2$, MPC-1 shows fewer peaks than MPC-2 (**Fig. R-11B**). Thus, differences in methylation peak profiles are mostly due to differences in *P*-values. They do not however necessarily reflect true differences in methylation patterns (see **Figs. R-11A, R-12**). More advanced statistical analysis, underway at the NMC, is again required to assess the technical reproducibility of the assay.

To start comparing methylation profiles between cell types, we scanned *P*-value and peak profiles for all chromosomes in all cell types. Methylation profiles were similar but not identical (**Figs. R-11A,B**). Similarities between ASCs, BMMSCs, MPCs and HSCs were also illustrated at all levels on a random segments of chromosome 1 (**Supplementary Fig. 4A,B**). The different cell types also displayed differences in methylation, as exemplified by the cluster of olfactory receptor family 2 gene members (**Supplementary Fig. 5**). Differential methylation was independent of CpG islands, absent from this region. These observations indicate that methylation profiles between ASCs, BMMSCs, MPCs and HSCs are similar but not identical. An analysis of similarly and differentially methylated genomic areas between cell types is underway at the NMC.

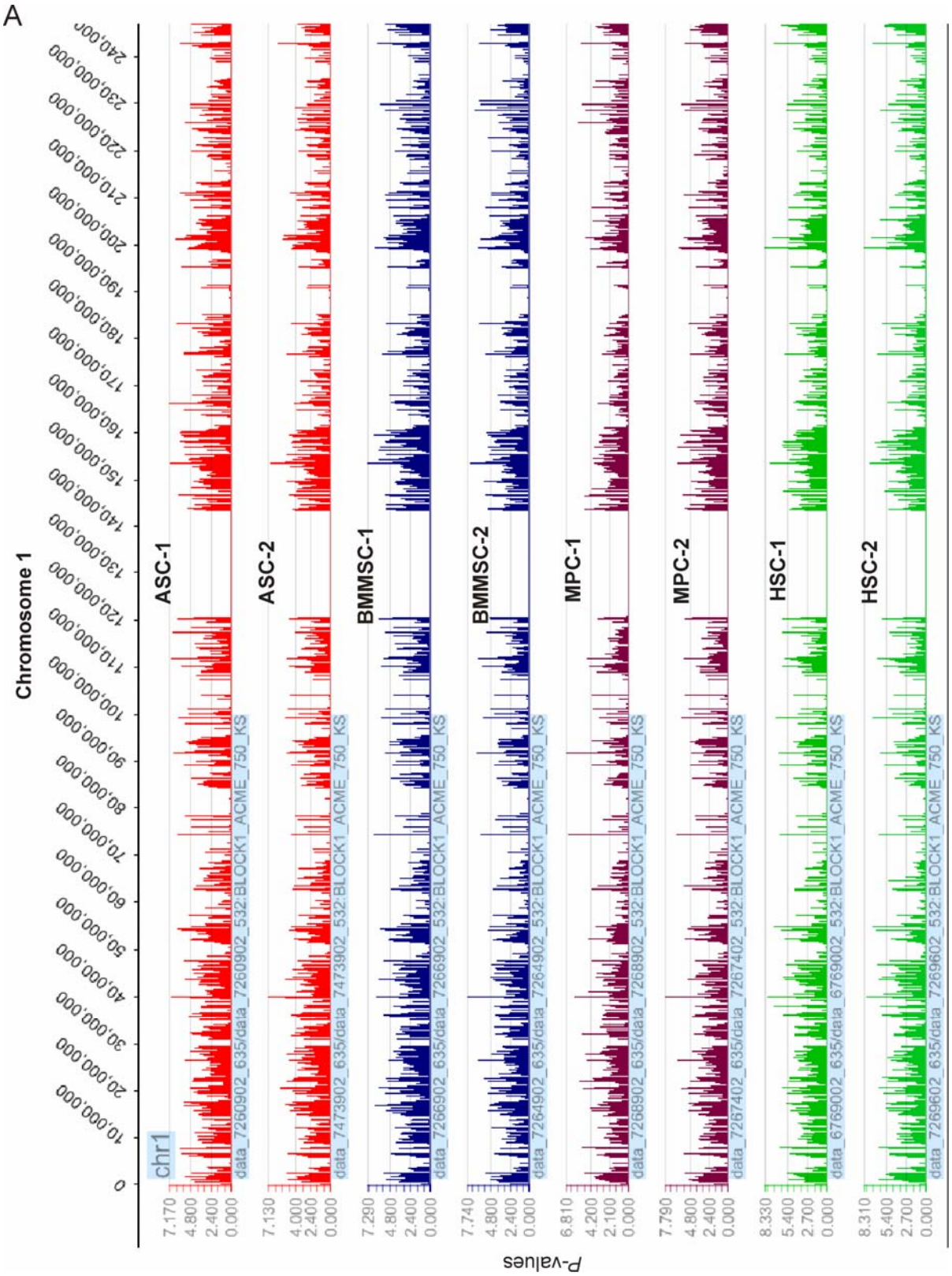



**Fig. R-11.** Methylation profiles in ASCs, BMMSCs, MPCs and HSCs. (A) *P*-values (expressed in -log10 values) for duplicate MeDIP samples for each cell type, shown here for chromosome 1. (B) Methylation peaks identified from *P*-values shown in (A); see next page.

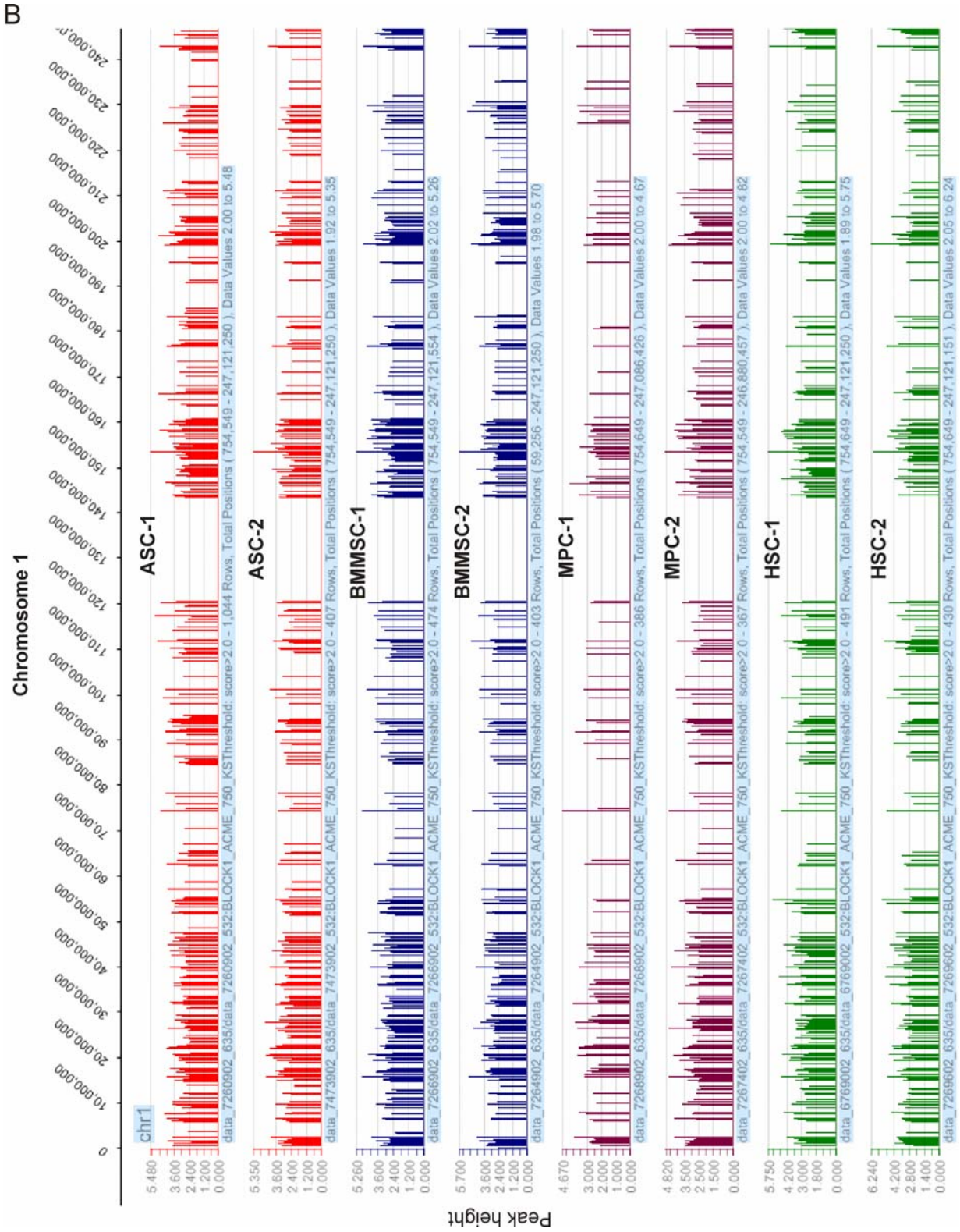



**Fig. R-11 (cont.).** Methylation profiles in ASCs, BMMSCs, MPCs and HSCs. (A) *P*-values (expressed in –log10 values) for duplicate MeDIP samples for each cell type, shown here for chromosome 1. (B) Methylation peaks identified from *P*-values shown in (A).

**Validation of the MeDIP data**

To validate the MeDIP procedure, we first assessed methylation at the levels of IP/input $\log_2$ ratios and *P*-values, for known unmethylated (*UBE2B, PEX13*) and methylated (*OXT, LDHC*) promoters, on the basis of data from Weber et al. (Weber *et al.*, 2007). For all cell types, **Figure R-12** revealed DNA methylation in the promoters expected to be methylated from PCR and MeDIP data in fibroblasts (Weber *et al.*, 2007); there was no probe on the array for *H19ICR* (see **Fig. R-9C**). Also, all cell types showed no methylation in the unmethylated promoters, including *UBE2B* (see also **Fig. R-9C** for our PCR data).

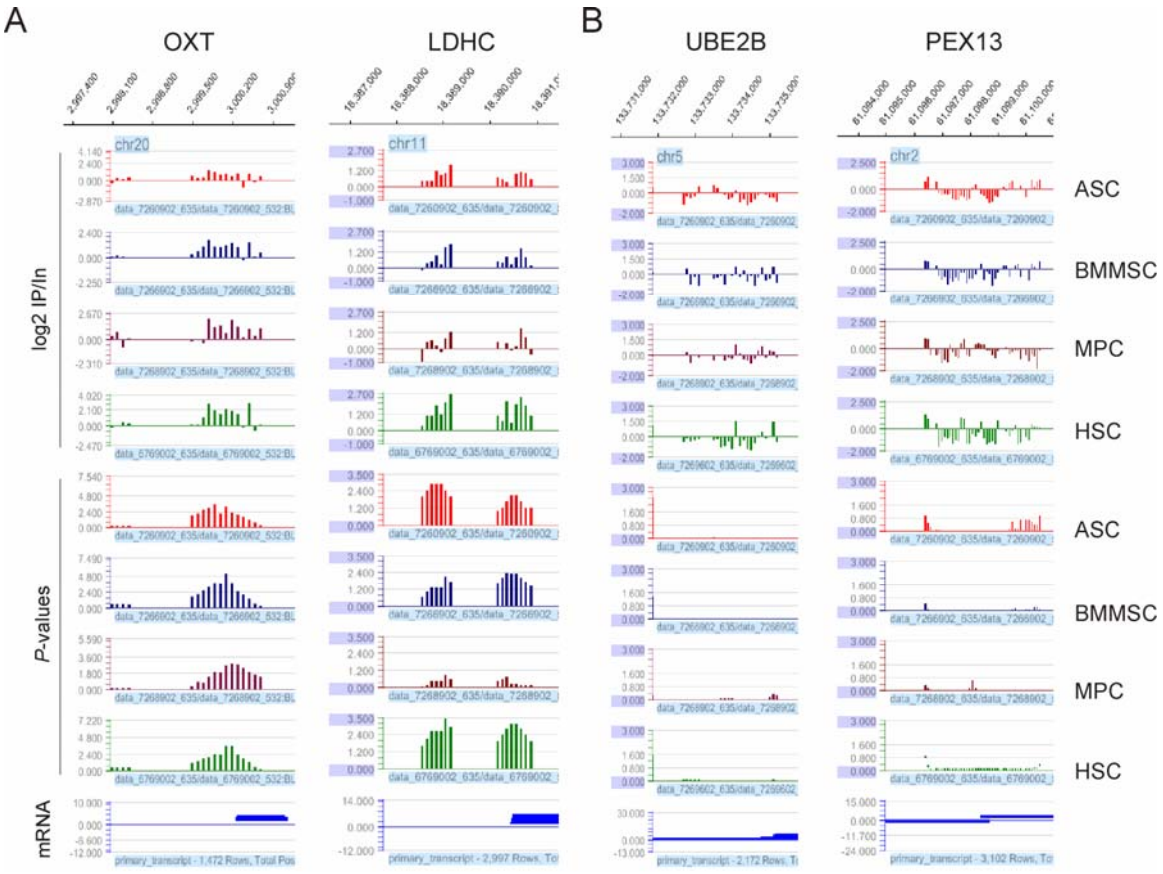



**Fig. R-12.** Validation of the MeDIP assay: MeDIP methylation profile of genes known to be methylated and unmethylated, in ASCs, BMMSCs, MPCs and HSCs. (A) Methylation profiles of two methylated promoters (*OXT, LDHC*). (B) Methylation profiles of two unmethylated promoters (*UBE2B, PEX13*). $\text{Log}_2$ IP/input ratios, *P*-values and transcripts are shown.

Next, we examined promoter regions that we studied by bisulfite sequencing. The $\log_2$ IP/input ratio and *P*-values for the *LEP*, *LPL*, *FABP4*, *PPARG2*, *MYOG* and *CD31* promoters

from ASCs are shown in **Figure R-13A**. Regions examined by bisulfite sequencing (red rectangles) were detected as unmethylated by MeDIP for *LEP*, *LPL* and *FABP4* (**Fig. R-13A**). MeDIP data for *PPARG2* revealed methylation at the 5' end of the tiled region with less methylation near the TSS. The region examined by bisulfite sequencing showed low methylation at the MeDIP level (**Fig. R-13A**), confirming our sequencing data (**Fig. R-2**). We also noticed that the *PPARG2* promoter displayed strong methylation upstream the region sequenced. The *MYOG* promoter showed no or weak methylation by MeDIP, a result also confirmed by sequencing. MeDIP analysis showed enhanced methylation 3' of the TSS (exon 1), also in agreement with sequencing data (compare **Fig. R-13A** with **Fig. R-2A**). Lastly, *CD31* showed methylation, as expected from sequencing. These results indicate therefore that MeDIP data for ASCs are supported by bisulfite sequencing.

We concluded from these preliminary analyses that immunoprecipitation of methylated DNA is specific and that the data appear to be reliable. Additional validation is currently being performed for 20 more genes by bisulfite sequencing.

**DNA methylation profiles in ASCs, BMMSCs, MPCs and HSCs: comparison between MeDIP and bisulfite sequencing**

Lastly, we next compared the methylation profiles obtained by MeDIP in ASCs, BMMSCs, MPCs and HSCs, to those obtained by bisulfite sequencing (**Fig. R-13B**). In the region covered by sequencing, MeDIP showed that *LEP* was not or weakly methylated in ASCs, BMMSCs and HSCs, but displayed more methylation in MPCs than expected from sequencing **(**see **Fig. R-2A)**. *LPL* was unmethylated which was also consistent with our previous results. *FABP4* was unmethylated in ASCs and BMMSCs in agreement with sequencing (see above); however, it also appeared unmethylated by MeDIP in MPCs and HSCs, in contrast to our sequencing data. A similar discrepancy was seen for *PPARG2*,

which was detected as weakly in MPCs and HSCs by MeDIP (**Fig. R-13B**) but was methylated based on bisulfite sequencing (**Fig. R-2A**). As discussed below, this discrepancy may be explained by differences in the number of CpGs in the immunoprecipitated region. MeDIP profiles of *PPARG2* in ASCs and BMMSCs were compatible with bisulfite sequencing. The *MYOG* promoter in ASCs showed weak methylation by MeDIP while exon 1 showed increased methylation; this was also in complete agreement with sequencing (**Fig. R-13B**). MPCs and HSCs also displayed MeDIP methylation profiles consistent with sequencing. However, discrepancy was detected in BMMSCs, which showed by MeDIP methylation in the *MYOG* promoter region, a result not anticipated by sequencing (compare **Figs. R-2A** and **R-13B**). Lastly, MeDIP indicates that the *CD31* promoter was methylated in ASCs and BMMSCs, in agreement with bisulfite sequencing, and to a lesser extent in MPCs. However, *CD31* was unmethylated in HSCs, again in agreement with our sequencing data (compare **Figs. R-8A** and **R-13B**).

Collectively, these results indicate that there is an overall solid correlation between MeDIP and bisulfite sequencing data in the regions examined. The differences can be explained by CpG density, which affects the immunoprecipitation (see Discussion). The results also argue that different MSC types can display distinct methylation profiles on lineage-specific promoters; however conclusions on whether DNA methylation patterns may predict differentiation capacity, on the basis of genome-wide MeDIP data, await thorough analysis.

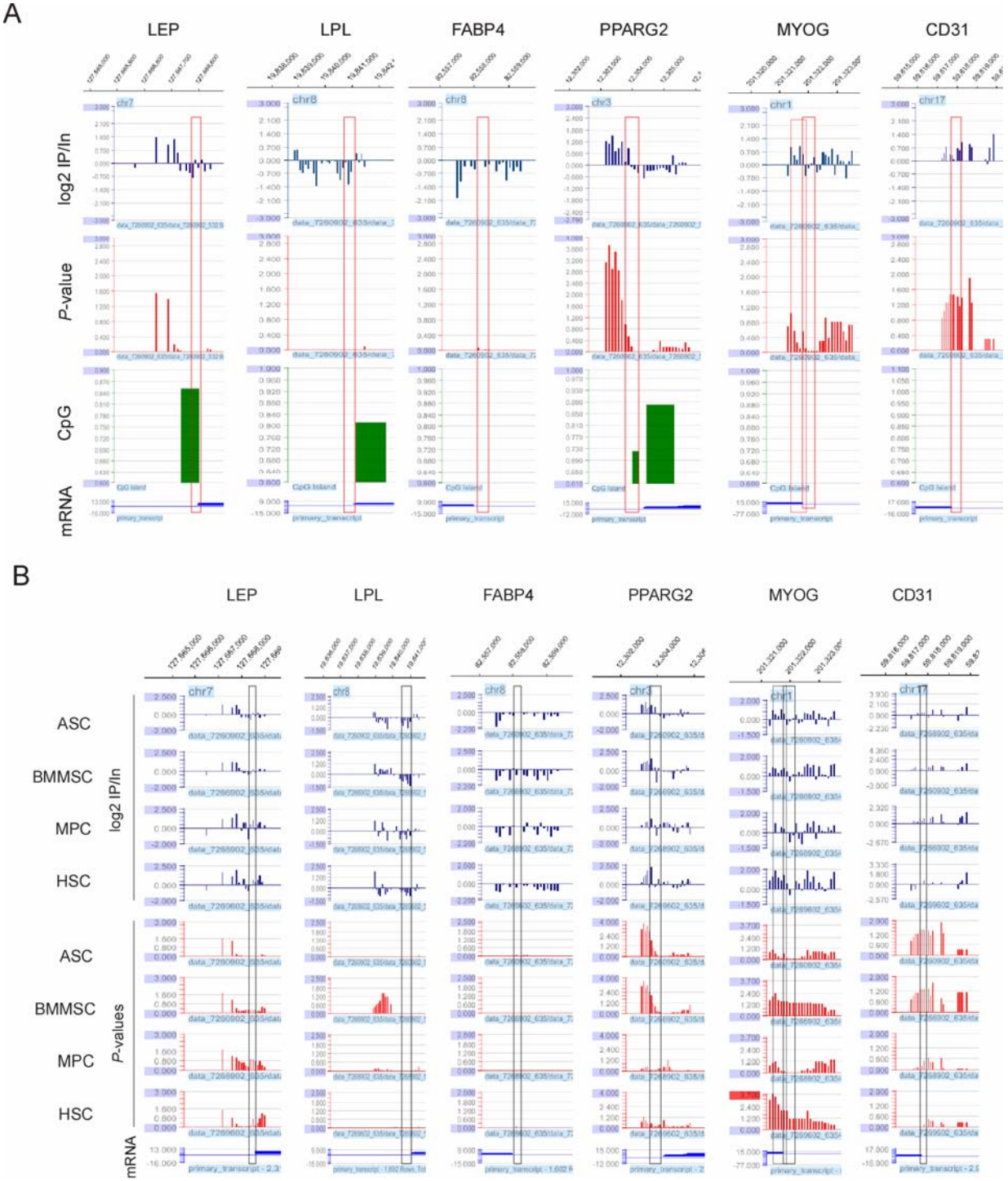



**Fig. R-13.** MeDIP methylation profiles of *LEP*, *LPL*, *FABP4*, *PPARG2*, *MYOG* and *CD31* in ASCs, BMMSCs, MPCs and HSCs. (A) Methylation patterns in ASCs. Red boxes delineate regions analyzed by bisulfite sequencing (Fig. R2A). A CpG island track (CpG) is shown. Note the robust fit between MeDIP and bisulfite sequencing data (compare with Figure 2A). (B) Methylation profiles for the same genes in all cell types examined. Black boxes delineate regions analyzed by bisulfite sequencing.

# DISCUSSION

This work tests the hypothesis that the DNA methylation pattern on promoters of lineage-specific genes in MSCs may be used as a predictor of differentiation capacity toward a given lineage. Indeed, MSCs from various origins have distinct differentiation capacities and similarly, different ASC clones have distinct differentiation potential (Boquest *et al.*, 2005). Using bisulfite sequencing and *in vitro* differentiation assays, we show here a relationship between the extent of CpG methylation of adipogenic and myogenic promoters and adipo- and myogenic differentiation capacity of MSCs derived from adipose tissue, bone marrow and muscle. This relationship, however, is dependent on the CpG content of the promoters examined. In addition, Wharton's Jelly MSCs display epigenetic preference for myogenic differentiation, while HSCs are clearly not epigenetically programmed for adipo- or myogenesis. Surprisingly however, HSCs and to a lesser extent BMMSCs show a methylation pattern suggestive of endothelial differentiation potential. Our bisulfite sequencing data are corroborated by a preliminary analysis of methylation profiling of ~24,000 promoters by MeDIP-chip.

## CpG methylation patterns in MSCs and HSCs

### Lineage-specific promoters show mosaic methylation

Except for *CD31*, all promoter regions examined showed mosaic methylation patterns. This is consistent with our earlier observations based on heterogeneous methylation patterns between ASC donors, single cell-derived ASC clones, and individual cells within a clone (Noer *et al.*, 2006; Boquest *et al.*, 2007). Mosaic CpG methylation has also been reported in stem cells from single intestinal crypts (Yatabe *et al.*, 2001), and may result from stochastic

methylation which accumulates in culture and from environmental, aging and health related factors (Esteller, 2005; Laird, 2005).

**Transcription factor binding elements are unmethylated**

CpG methylation is critical for maintenance of a silent chromatin state through recruitment of repressive chromatin modifiers to methylated DNA, and many transcription factors are unable to bind to their recognition site if these contain methylated CpGs. The consistent unmethylated state of specific CpGs suggests localization within binding sites for transcription factors. In the *LEP* promoter, three Sp1 elements and a CAAT enhancer binding protein (C/EBP) element are consistently unmethylated. We also found several key cis-elements in unmethylated regions of the *LPL* promoter. These include Oct-1 (nt -46), NF-Y (nt -65) and CT (CCTCCCCC, nt -91 to -83) elements essential for basal promoter activity (Yang and Deeb, 1998; Mead *et al.*, 2002) and a sterol-responsive element (SRE; nt -90 to -81), a peroxisome proliferator response element (PPRE) and Sp1/Sp3 motifs essential for *LPL* activation (Schoonjans *et al.*, 2000). Similarly, the *FABP4* and *PPARG2* promoters contain unmethylated binding sites for C/EBP and AP1 (*FABP4*) and a putative binding site for GATA1/2 (*PPARG2*). Lastly, we found binding sites for Sp1, AP-2, GATA and an E-box (Gumina *et al.*, 1997) in *CD31*, which may be unmasked by the unmethylated state of CpGs in HUVECs and HSCs.

These observations argue that in promoters, consistently unmethylated CpGs tend to reside within binding sites for transcription factors. These may, by binding to these elements, mask these CpGs and keep them unmethylated. Alternatively, CpGs may be protected from methylation to enable transcription factor targeting. Mutational promoter studies have indeed shown that manipulating the methylation state of specific CpGs affects factor binding and

transcription Campanero, 2000 CAMPANERO2000 /id; Comb, 1990 COMB1990 /id; Prendergast, 1991 PRENDERGAST1991B /id}.

## Relationship between promoter DNA methylation and gene activity

### CpG methylation and gene expression

Irrespective of MSC type, the *LEP* and *LPL* promoters are unmethylated, which is consistent with their localization in a CpG island. Surprisingly however, *LEP* showed significantly more methylation in HSCs. The contribution of promoter DNA methylation to tissue-specific gene expression is uncertain (Weber *et al.*, 2007) and most CpG islands remain unmethylated even in cell types that do not express the gene (Bird, 2002). Yet, methylation of CpG islands not associated with disease occasionally occurs. CpG islands can be differentially methylated in a tissue-specific manner, reflecting expression patterns (Song *et al.*, 2005).The CpG island of the *SERPINB5* promoter is unmethylated in cells expressing the gene, but is densely methylated in SERPINB5 negative cells (Futscher *et al.*, 2002). Furthermore, a recent MeDIP-chip study interrogating ~16,000 RefSeq human promoters identified 3% high CpG promoters (HCPs) with marked methylation (Weber *et al.*, 2007). Of note, the *LEP* promoter is an HCP, arguing that methylation of *LEP* in HSCs is one of the few instances where a CpG-rich promoter is methylated in a cell type where it is inactive. It will be interesting to investigate which HCP promoters in our on-going MeDIP work show differential methylation in MSCs, HSCs and differentiated cells.

The *FABP4* and *PPARG2* promoters are low CpG promoters (LCPs). Remarkably, as with HCPs, their methylation state could not be correlated to expression status because they were not expressed in MSCs or HSCs, although they displayed distinct unmethylated (ASCs), moderately methylated (BMMSCs, MPCs) or strongly methylated profiles (MPCs, HSCs). Methylation patterns were maintained in differentiated adipocytes and muscle relative to

undifferentiated cells. Further, genome-wide profiling of promoter DNA methylation and RNAPII binding indicates that LCPs can be methylated both when active and inactive (Weber *et al.*, 2007). Therefore, not only LCPs can be repressed even when unmethylated (as in ASCs), but a low concentration of methylated cytosines does also not preclude RNAPII binding and transcriptional activity. Thus, in MPCs, HSCs and WJMSCs, methylation of the *FABP4* and *PPARG2* promoters does not necessarily preclude initiation of transcription, which would be undetectable by RT-PCR.

The region of the endothelial *CD31* gene examined, which extended into exon 1, was strongly methylated in ASCs, BMMSCs, MPCs and WJMSCs, but was unmethylated in HSCs, as in $CD31^+$ cells and HUVECs. The *CD31* promoter itself belongs to the LCP category of Weber *et al*. (2007). Accordingly, we have shown that freshly isolated ASCs express *CD31* (at low level) although the promoter is methylated (Boquest *et al.*, 2007). So as with other LCPs, CpG methylation is compatible with low level transcription. However, in cells expressing high levels of *CD31* mRNA and protein ($CD31^+$ cells from the SVF and HUVECs), we found that *CD31* exhibits a hemimethylated profile (Boquest *et al.*, 2007). Therefore, although we at present do not know whether *CD31* is or can be expressed in HSCs (work in progress), available data suggest that *CD31* methylation does not preclude low level gene expression (without protein expression), but high level expression correlates with demethylation of the *CD31* region examined. Whether *CD31* is unmethylated in HSCs remains to be shown, however upregulation of *CD31* expression in blood-derived $CD34^+$ hematopoietic progenitors and contribution of these cells to the vasculature (Kung *et al.*, 2008) suggest that in endothelial cells differentiated from $CD34^+$ cells, *CD31* is likely to be unmethylated.

These observations provide evidence against DNA methylation as a primary regulator of tissue-specific gene expression not only in differentiated cells (Jones and Takai, 2001), but

also in adult stem and progenitor cells. Our on-going genome-wide methylation profiling study of MSCs, HSCs and differentiated cells combined with gene expression array analysis is expected to provide deeper insights on the relationship between promoter DNA methylation and actual gene expression, rather than merely RNAPII binding (Weber *et al.*, 2007).

**Regulation of lineage-specific gene expression by histone modification**

If DNA methylation is not the primary determinant of gene expression, what then regulates gene activation or activation potential, upon stem cell differentiation? Another epigenetic component, namely post-translational histone modification, is a key player. In particular, trimethylation of H3K27, a mark of facultative heterochromatin, has recently emerged as a candidate for a transcriptional "brake". Genome-wide and locus-specific ChIP analyses of mouse ESCs reveal that repressed but potentially active developmentally-regulated promoters are enriched in "bivalent" histone marks characterized by H3K4m3 and H3K27m3 (Azuara *et al.*, 2006; Bernstein *et al.*, 2006; Mikkelsen *et al.*, 2007). H3K4m3 is a mark of active genes whereas H3K27m3 is associated with inactive genes. To support this view, in undifferentiated ESCs, Ezh2 (a PcG protein methylating H3K27) was found to occupy genes poised for transcription (Lee *et al.*, 2006; Bracken *et al.*, 2006). Observations from our laboratory have also identified H3K4m3 and H3K27m3 on adipogenic and myogenic promoters in undifferentiated ASCs, and demethylation of H3K27 selectively on adipogenic promoters upon adipogenic differentiation (A. Noer, L. Lindeman and P. Collas, manuscript submitted). This finding further supports the view that in ASCs, adipogenic promoters are epigenetically pre-programmed for activation upon adipogenic stimulation.

Notably, these marks were identified on hypo- or moderately methylated promoters (including those examined in this study), indicating that targeting of trithorax and polycomb

proteins, which methylate H3K4 and H3K27 respectively, and resulting histone methylation are compatible with the unmethylated state of DNA. It will be interesting to unravel histone marks enriched on the *CD31* promoter in its hypermethylated state in ASCs, and in its unmethylated state in HSCs relative to $CD31^+$ cells in which it is active. These studies are underway.

## Relationship between promoter DNA methylation and differentiation potential

Despite the complexity of the relationship between DNA methylation and gene expression, our results opens for the suggestion that methylation patterns on lineage-specific promoters can predict differentiation *potential*. This relationship may also depend on CpG content (Weber *et al.*, 2007), because HCPs are mostly unmethylated regardless of their activation status, and the low CpG density of LCPs does not preclude promoter activation. This, therefore, would suggest that HCPs and LCPs constitute poor predictors of differentiation capacity. However, our bisulfite sequencing data suggest a positive relationship between CpG methylation of some of the promoters examined in undifferentiated MSCs, and differentiation capacity – as opposed to gene expression.

### Methylation patterns of *FABP4* and *PPARG2* predict adipogenic potential

Adipogenic differentiation capacity may be more evident on the basis of methylation of ICPs than HCPs. ASCs and BMMSCs, but not MPCs, showed robust adipogenic differentiation capacity, confirming previous studies (Reyes *et al.*, 2001; Zuk *et al.*, 2002; Gimble and Guilak, 2003; Boquest *et al.*, 2005). Both *LEP* and *LPL* promoters are unmethylated regardless of MSC type and of adipogenic potential. Thus in MSCs, methylation of these loci provide no information on predicting adipogenic differentiation. In HSCs however, the *LEP*

CpG island is 38% methylated and hypermethylated relative to MSCs, excluding the adipogenic potential of these cells.

The methylation states of *FABP4* and *PPARG2* (discussed above), however, were more informative. In particular, we found a correlation between the unmethylated state of *FABP4* in ASCs and BMMSCs and adipogenic differentiation capacity of these cells, in contrast to MPCs in which it was hypermethylated. Adipogenic potential of WJMSCs was not tested in our study, but has been reported *in vitro* (Karahuseyinoglu *et al.*, 2007). However, adipogenic differentiation of WJMSCs was achieved over a longer period compared to BMMSCs (Karahuseyinoglu *et al.*, 2007), suggesting a weaker adipogenic potential of WJMSCs. Because *FABP4* is more methylated in WJMSCs than ASCs or BMMSCs, it is possible however that this potential is provided by the unmethylated state of one CpG (No. 4) in all MSC types examined. In contrast, *FABP4* is fully methylated in HSCs, again presumably precluding adipogenic potential and reflecting a lineage clearly distinct from MSCs. It would be interesting to determine whether demethylation of *FABP4* and *PPARG2* by 5-azacytidine (which demethylates DNA) in HSCs would promote adipogenic potential.

**CpG methylation pattern of the MYOG promoter reveals myogenic potential of MPCs**

The *MYOG* promoter was hypomethylated in MPCs and WJMSCs relative to ASCs or BMMSCs, but was fully methylated in HSCs. One CpG (No.3, -314 nt from TSS) was strongly methylated in ASCs, BMMSCs and HSCs, but less methylated in WJMSCs and MPCs. Thus one may speculate that unmethylation of this CpG is involved in activation of the gene, in line with the myogenic potential of MPCs and, to a lesser extent, WJMSCs (Conconi *et al.*, 2006).

The *MYOG*(E1) (exon 1) region examined was heavily methylated in ASCs, BMMSCs, MPCs and HSCs but displayed strong mosaicism in WJMSCs. Methylation of *MYOG*(E) is in agreement with earlier studies showing DNA methylation outside regulatory regions, including exons (Rabinowicz *et al.*, 2003; Eckhardt *et al.*, 2006; Weber and Schubeler, 2007). Our MeDIP data also reveal genes with a non-methylated promoter and methylation 3' of the TSS (data not shown). Intragenic DNA methylation may inhibit cryptic transcription initiation (Weber and Schubeler, 2007). One CpG (No.1; nt +21) in *MYOG*(E1) was largely unmethylated in MPCs. Unmethylation of this CpG may be important for activation of the gene as it is adjacent to an E-box.

The *MYOG* promoter methylation pattern correlated with differentiation potential of MPCs, but not ASCs or BMMSCs, into multinucleated fibers. Interestingly, mouse satellite cells can differentiate into adipocytes *in vitro* (Asakura *et al.*, 2001; Wada *et al.*, 2002); however our data show the inability of human MPCs to do so. This may be because MPCs are more differentiated than satellite cells and already committed to myogenesis. This may also reflect however a programming of satellite cells themselves to myogenesis preferably over adipogenesis. In addition, WJMSCs have adipogenic differentiation capacity *in vitro* (Karahuseyinoglu *et al.*, 2007); however, adipogenic differentiation was achieved after a longer period compared to BMMSCs. These observations argue that MPCs, and putatively the more primitive satellite cells, are epigenetically programmed for myogenesis, presumably by DNA methylation on, at least, the *MYOG* promoter. Methylation patterns of undifferentiated MPCs were conserved in myogenic differentiated MPCs, supporting this view. Other MSC types, such as those from Wharton's Jelly, may also display methylation-dependent myogenic potential, in contrast to ASCs, BMMSCs and as expected, HSCs.

**Methylation of CD31 reflects limited endothelial differentiation capacity of MSCs**

Extensive methylation of *CD31* strongly suggests a lack of endothelial differentiation in ASCs (Boquest *et al.*, 2007), as well as BMMSCs, MPCs and WJMSCs. This is confirmed by the *CD31* demethylation in CD31 expressing cells. In contrast, HSCs were unmethylated in the *CD31* region examined. Endothelial differentiation of each of these cell types is currently being tested. BMMSCs, in contrast to other MSCs, showed unmethylated CpGs in *CD31* exon 1. These were outside binding elements for known transcription factors (Gumina *et al.*, 1997), perhaps because they locate to a genic rather than regulatory region. The significance of unmethylation of these CpGs may be related to the ability of BMMSCs to differentiate into endothelial cells *in vitro* and *in vivo* (Reyes *et al.*, 2001).

As expected, *CD31* was also strongly methylated in adipocytes, muscle, fibroblasts and T cells. Yet, the 20% unmethylated alleles in T cells were reminiscent of the unmethylated state of HSCs. These alleles may also emanate from contaminating immature T cells or $CD34^+$ cells (HSCs) with unmethylated *CD31*. Interestingly, $CD34^+$ peripheral blood mononuclear cells (hematopoietic precursors; (Ketterer *et al.*, 1998) have been shown to restore blood flow to ischemic limbs (Schmeisser *et al.*, 2001) and form microvessels in human skin substitutes (Kung *et al.*, 2008), suggesting the endothelial potential of HSCs. Endothelial differentiation capacity of HSCs and BMMSCs is being tested in our laboratory. The methylation profile of *CD31*, however, argues against the endothelial potential of MSCs, in agreement with reports that contribution of MSCs to vasculature is most likely indirect and mediated by secretion of angiogenic factors (Rehman *et al.*, 2004).

## Differentiation plasticity of somatic stem cells

The multilineage differentiation capacity of somatic stem cells makes them candidates for replacement cells in regenerative medicine (Verfaillie *et al.*, 2002; Fraser *et al.*, 2006).

However, such plasticity remains controversial and may vary with the cell type (including mode of isolation), tissue and the model systems studied.

MSCs exhibit differentiation capacity along mesodermal and non-mesodermal lineages and transplanted MSCs can contribute to various tissues (Cousin *et al.*, 2003; Kang *et al.*, 2003; Cowan *et al.*, 2004). Yet, little is known on the epigenetic commitment of MSCs to a specific cell or tissue type *in vivo*. It is to date not proven that MSCs form functional tissue of non-mesodermal lineages *in vivo*. MSCs may contribute to tissue repair indirectly by promoting the recruitment of endogenous cells or through paracrine mechanisms (Heil *et al.*, 2004; Miyahara *et al.*, 2006; Boquest *et al.*, 2006a). In addition, in several studies undifferentiated cells were treated with 5-azacytidine prior to differentiation (Wakitani *et al.*, 1995; Conconi *et al.*, 2006). Thus, genes that would normally be silenced may be demethylated and amenable to activation. To get more knowledge on the epigenetic commitment of MSCs to tissue *in vivo*, our laboratory is currently testing an in vivo model system that allows tracking and recovering MSCs transplanted in tissues in order to examine their phenotype and epigenetic commitment to a particular lineage.

Several hypotheses may account for the multilineage differentiation capacity of MSCs. One mechanism for pluripotency is transdifferentiation, whereby somatic stem cells from one germ layer might differentiate into cells of another germ layer. (ii) A second mechanism may involve a de-differentiation process, whereby somatic stem cells “revert” to a more primitive state, followed by re-differentiation along another lineage. A classical example is the replacement of lost anatomical parts in urodele amphibians, through dedifferentiation, proliferation and redifferentiation of epithelial cells in the wounded area (Brockes and Kumar, 2002). Whether such a mechanism also exists in humans is unclear, although the regenerative capacity of the liver may reflect the existence of such potential. (iii) Cell fusion can also alter cell fate, in a process where a stem cell may fuse with another

somatic cell to produce a heterokaryon in which genes from both cell types are expressed (Weimann *et al.*, 2003).

Another likely explanation for stem cell pluripotency is that a given tissue harbors multiple stem or progenitor cell types, each committed toward a preferred lineage. As mentioned earlier, bone marrow contains at least three distinct stem cell populations including HSCs, MSCs and endothelial progenitors. A rare MSC subset in bone marrow also seems to form tissues of all three germ layers (Jiang *et al.*, 2002). Similarly, adipose tissue contains different progenitor cell types, including ASCs and $CD31^+$ endothelial progenitors (Boquest *et al.*, 2005). Furthermore, even ASCs are thought to contain multiple cell types, because different single cell clones differentiate into adipogenic, osteogenic or neurogenic pathways with various efficiencies (Boquest *et al.*, 2005). The molecular determinants of these various stem cell sub-populations are not entirely determined. We have hypothesized that these may include a strong epigenetic component, including promoter DNA methylation. Our previous work (Noer *et al.*, 2006; Boquest *et al.*, 2006a; Boquest *et al.*, 2007) and this thesis illustrate a mosaicism in DNA methylation patterns of MSCs from adipose, bone marrow and muscle tissues, suggesting the existence of epigenetically non-identical cells even within a tissue-specific population.

## Perspectives: extension to genome-wide methylation profiling

Our conclusions on the use of DNA methylation as a predictor of differentiation capacity are at present based on a detailed analysis of CpG methylation in a handful of promoters and in a restricted area (300-200 bp). Extension of these data to genome-wide interrogation is expected to provide additional insights on this prediction.

As part of this work we have implemented a MeDIP approach to genome-wide methylation profiling of MSCs, HSCs and differentiated cells. Visual scanning of $\log_2$

IP/Input ratios and *P*-values shows high similarity between MeDIP duplicates and between cell types. Reproducibility between duplicates is more evident at the $\log_2$ IP/input ratios and at the *P*-value levels than at the methylation peak level, as a peak is defined by at least two consecutive *P*-values above 2. Therefore, similar methylation profiles at the *P*-value level may appear as different methylation peak profiles (**Fig. R11**). Once we have established a work-flow for these analyses, we will be able to experimentally adjust the offset for defining peaks based on values that are biologically relevant. Indeed, methylation profiles are clearly similar at the *P*-value level, which should also be rendered at the peak level.

The high conservation of global methylation patterns between somatic cells implies that few differences would be detected between MSC types. Bioinformatics analysis of these data is underway in collaboration with the NMC (Oslo). A critical component of this analysis in the frame of our study will be to bring out genes or gene cluster showing distinct methylation profiles between all cell types, and classify them based on ontology (functional grouping).

Comparison of methylation data produced by bisulfite genomic sequencing and MeDIP reveals apparent discrepancies, and thereby limitations to the MeDIP approach. Although most comparisons reveal consistency between the two approaches (**Fig. R13**), a classical example of apparent discrepancy is illustrated by the methylation pattern of *FABP4* and *PARG2* in HSCs (**Figs. R2A** and **R13B**). Whereas bisulfite sequencing reveals strongly methylated promoters, MeDIP does not detect any methylation. A simple explanation takes into account the number of methylated cytosines (recognized by the antibody and therefore immunoprecipitated) in the region examined, relative to the total number of CpGs. Thus, MeDIP handles 4 but all methylated CpGs (in the 413-bp *FABP4* promoter region examined) exactly as 4 methylated CpGs among a much larger number of CpGs in a 413-bp area. In the first instance, bisulfite sequencing indicates that all four CpGs are methylated, making the

promoter “methylated”. MeDIP however, considers the same area as unmethylated (*FABP4*) or poorly methylated (*PPARG2*) (**Fig. R13B**). It is therefore important to remember that the strength of MeDIP is to provide an average methylation level over the tiled region examined, relative to genome average methylation. Moreover, in the arrays used in this study, average probe spacing was 100 bp with an oligonucleotide length of 60 bp, leaving ~40 uncovered (and therefore not taken into consideration) nucleotides between each probe. These undetected nucleotides may contain significantly methylated CpGs, contributing to the methylation pattern reported by bisulfite sequencing.

We conclude thus far, therefore, that MeDIP and bisulfite sequencing analyses are highly complementary strategies which cannot always substitute for one another. In addition, MeDIP-chip data must be validated for large number of promoters, either by MeDIP-PCR (Weber and Schubeler, 2007) or, preferably, by an independent method such as bisulfite sequencing. To this end, we are currently validating ~20 methylated and unmethylated promoters by bisulfite sequencing. Once validation is completed and the appropriate work-flow determined, we will be in a position to (i) identify differentially methylated promoters between ASCs, BMMSCs, MPCs, WJMSCs and HSCs, (ii) determine whether our hypothesis on methylation pattern as a prediction of differentiation capacity is verified on a large scale, and importantly (iii) determine to what extent DNA methylation profiles on lineage-specific promoters are maintained or altered in the context of differentiation. On the basis of our present and earlier bisulfite sequencing data (Noer *et al.*, 2006; Boquest *et al.*, 2006a; Boquest *et al.*, 2007), we anticipate greater methylation stability after differentiation than between MSC types.

# ACKNOWLEDGEMENTS

I am grateful to Kristin Vekterud for collaboration on specific experiments. I am also thankful to Prof. Mark Kirkland (Deakin University, Geelong, Australia) for providing WJMSC DNA, Dr. Gisèle Bonne (Institut de Myologie, Paris, France) for the muscle biopsy, Drs. Jan Brinchann and Aboulghassem Shahdadfar (Rikshospitalet-Radiumhospitalet Medical Center, Oslo) for HSCs and BMMSCs, Prof. Christian Drevon (University of Oslo) for SGBS adipocytes, Prof. Heidi K. Blomhoff (University of Oslo) for T cells, Prof. Gutthorm Haraldsen (Rikshospitalet-Radiumhospitalet Medical Center, Oslo) for HUVEC cells, and Prof. Lorenzo P. Puri (Dulbecco Telethon Institute, Rome, Italy) for advice on culturing muscle-derived cells. We are indebted to Dr. Andy Reiner (Norwegian Microarray Consortium, Rikshospitalet-Radiumhospitalet Medical Center, Oslo) for preliminary statistical analysis of MeDIP data. This work was supported by grants from the Research Council of Norway (FUGE, STAMCELLE, and STORFORSK programs) to P. Collas.

I am thankful to my advisor (and husband) Prof. Philippe Collas for giving me the opportunity to become a master student in his group. Philippe, to benefit from your enormous motivation, knowledge and support has been a great pleasure. I want to thank everyone in "The Collas lab" for being such great colleagues and bringing joy into the daily work. Agate and Andrew, I appreciate everything you have taught me about ASCs and for sharing your knowledge with me, and Kristin, it is fun working together again. I want to thank Martina for the valuable collaboration and discussions we have had during the courses at Blindern. Finally, I want to thank my family Elena, Sebastian and Philippe for all your love, support and patience during the last two years.

# REFERENCES


Asakura,A., Komaki,M., and Rudnicki,M. (2001). Muscle satellite cells are multipotential stem cells that exhibit myogenic, osteogenic, and adipogenic differentiation. Differentiation *68*, 245-253.

Asakura,A., Seale,P., Girgis-Gabardo,A., and Rudnicki,M.A. (2002). Myogenic specification of side population cells in skeletal muscle. J. Cell Biol. *159*, 123-134.

Azuara,V. *et al.* (2006). Chromatin signatures of pluripotent cell lines. Nat. Cell Biol. *8*, 532-538.

Bachrach,E. *et al.* (2006). Muscle engraftment of myogenic progenitor cells following intraarterial transplantation. Muscle Nerve *34*, 44-52.

Bacou,F., el Andalousi,R.B., Daussin,P.A., Micallef,J.P., Levin,J.M., Chammas,M., Casteilla,L., Reyne,Y., and Nougues,J. (2004). Transplantation of adipose tissue-derived stromal cells increases mass and functional capacity of damaged skeletal muscle. Cell Transplant. *13*, 103-111.

Baksh,D., Song,L., and Tuan,R.S. (2004). Adult mesenchymal stem cells: characterization, differentiation, and application in cell and gene therapy. J. Cell Mol. Med. *8*, 301-316.

Baksh,D., Yao,R., and Tuan,R.S. (2007). Comparison of proliferative and multilineage differentiation potential of human mesenchymal stem cells derived from umbilical cord and bone marrow. Stem Cells *25*, 1384-1392.

Bernstein,B.E., Meissner,A., and Lander,E.S. (2007). The mammalian epigenome. Cell *128*, 669-681.

Bernstein,B.E. *et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Bhattacharya,D., Rossi,D.J., Bryder,D., and Weissman,I.L. (2006). Purified hematopoietic stem cell engraftment of rare niches corrects severe lymphoid deficiencies without host conditioning. J. Exp. Med. *203*, 73-85.

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

Bird,A.P., Taggart,M.H., Nicholls,R.D., and Higgs,D.R. (1987). Non-methylated CpG-rich islands at the human alpha-globin locus: implications for evolution of the alpha-globin pseudogene. EMBO J. *6*, 999-1004.

Bonab,M.M., Alimoghaddam,K., Talebian,F., Ghaffari,S.H., Ghavamzadeh,A., and Nikbin,B. (2006). Aging of mesenchymal stem cell in vitro. BMC. Cell Biol. *7*, 14-21.

Boquest,A.C., Noer,A., and Collas,P. (2006a). Epigenetic programming of mesenchymal stem cells from human adipose tissue. Stem Cell Rev. *2*, 319-329.

Boquest,A.C., Noer,A., Sorensen,A.L., Vekterud,K., and Collas,P. (2007). CpG methylation profiles of endothelial cell-specific gene promoter regions in adipose tissue stem cells suggest limited differentiation potential toward the endothelial cell lineage. Stem Cells *25*, 852-861.

Boquest,A.C., Shahdadfar,A., Brinchmann,J.E., and Collas,P. (2006b). Isolation of stromal stem cells from human adipose tissue. Methods Mol. Biol. *325*, 35-46.

Boquest,A.C., Shahdadfar,A., Fronsdal,K., Sigurjonsson,O., Tunheim,S.H., Collas,P., and Brinchmann,J.E. (2005). Isolation and transcription profiling of purified uncultured human stromal stem cells: alteration of gene expression after in vitro cell culture. Mol. Biol. Cell. *16*, 1131-1141.

Bracken,A.P., Dietrich,N., Pasini,D., Hansen,K.H., and Helin,K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev. *20*, 1123-1136.

Brenner,C. *et al.* (2005). Myc represses transcription through recruitment of DNA methyltransferase corepressor. EMBO J. *24*, 336-346.

Brockes,J.P. and Kumar,A. (2002). Plasticity and reprogramming of differentiated cells in amphibian regeneration. Nat. Rev. Mol. Cell Biol. *3*, 566-574.

Bryder,D., Rossi,D.J., and Weissman,I.L. (2006). Hematopoietic stem cells: the paradigmatic tissue-specific stem cell. Am. J. Pathol. *169*, 338-346.

Buckingham,M. (2007). Skeletal muscle progenitor cells and the role of Pax genes. C. R. Biol. *330*, 530-533.

Buckingham,M. (2006). Myogenic progenitor cells and skeletal myogenesis in vertebrates. Curr. Opin. Genet. Dev. *16*, 525-532.

Cao,R., Wang,L., Wang,H., Xia,L., Erdjument-Bromage,H., Tempst,P., Jones,R.S., and Zhang,Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039-1043.

Cao,Y., Sun,Z., Liao,L., Meng,Y., Han,Q., and Zhao,R.C. (2005). Human adipose tissue-derived stem cells differentiate into endothelial cells in vitro and improve postnatal neovascularization in vivo. Biochem. Biophys. Res. Commun. *332*, 370-379.

Cao,Z., Umek,R.M., and McKnight,S.L. (1991). Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells. Genes Dev. *5*, 1538-1552.

Chamberlain,G., Fox,J., Ashton,B., and Middleton,J. (2007). Concise review: mesenchymal stem cells: their phenotype, differentiation capacity, immunological features, and potential for homing. Stem Cells *25*, 2739-2749.

Chen,J.C. and Goldhamer,D.J. (2003). Skeletal muscle stem cells. Reprod. Biol. Endocrinol. *1*, 101-110.

Chen,Z., Torrens,J.I., Anand,A., Spiegelman,B.M., and Friedman,J.M. (2005). Krox20 stimulates adipogenesis via C/EBPbeta-dependent and -independent mechanisms. Cell Metab *1*, 93-106.

Clark,S.J., Statham,A., Stirzaker,C., Molloy,P.L., and Frommer,M. (2006). DNA methylation: bisulphite modification and analysis. Nat. Protoc. *1*, 2353-2364.

Clarke,S.L., Robinson,C.E., and Gimble,J.M. (1997). CAAT/enhancer binding proteins directly modulate transcription from the peroxisome proliferator-activated receptor gamma 2 promoter. Biochem. Biophys. Res. Commun. *240*, 99-103.

Collas,P., Noer,A., and Timoskainen,S. (2007). Programming the genome in embryonic and somatic stem cells. J. Cell Mol. Med. *11*, 602-620.

Collins,C.A., Olsen,I., Zammit,P.S., Heslop,L., Petrie,A., Partridge,T.A., and Morgan,J.E. (2005). Stem cell function, self-renewal, and behavioral heterogeneity of cells from the adult muscle satellite cell niche. Cell *122*, 289-301.

Conconi,M.T., Burra,P., Di,L.R., Calore,C., Turetta,M., Bellini,S., Bo,P., Nussdorfer,G.G., and Parnigotto,P.P. (2006). CD105(+) cells from Wharton's jelly show in vitro and in vivo myogenic differentiative potential. Int. J. Mol. Med. *18*, 1089-1096.

Cortazar,D., Kunz,C., Saito,Y., Steinacher,R., and Schar,P. (2007). The enigmatic thymine DNA glycosylase. DNA Repair (Amst) *6*, 489-504.

Cossu,G. and Biressi,S. (2005). Satellite cells, myoblasts and other occasional myogenic progenitors: possible origin, phenotypic features and role in muscle regeneration. Semin. Cell Dev. Biol. *16*, 623-631.

Cousin,B., Andre,M., Arnaud,E., Penicaud,L., and Casteilla,L. (2003). Reconstitution of lethally irradiated mice by cells isolated from adipose tissue. Biochem. Biophys. Res. Commun. *301*, 1016-1022.

Cowan,C.M., Shi,Y.Y., Aalami,O.O., Chou,Y.F., Mari,C., Thomas,R., Quarto,N., Contag,C.H., Wu,B., and Longaker,M.T. (2004). Adipose-derived adult stromal cells heal critical-size mouse calvarial defects. Nat. Biotechnol. *22*, 560-567.

da Silva,M.L., Chagastelles,P.C., and Nardi,N.B. (2006). Mesenchymal stem cells reside in virtually all post-natal organs and tissues. J. Cell Sci. *119*, 2204-2213.

De Ugarte,D.A., Alfonso,Z., Zuk,P.A., Elbarbary,A., Zhu,M., Ashjian,P., Benhaim,P., Hedrick,M.H., and Fraser,J.K. (2003). Differential expression of stem cell mobilization-associated molecules on multi-lineage cells from adipose tissue and bone marrow. Immunol. Lett. *89*, 267-270.

Dellavalle,A. *et al.* (2007). Pericytes of human skeletal muscle are myogenic precursors distinct from satellite cells. Nat. Cell Biol. *9*, 255-267.

Doherty,M.J. and Canfield,A.E. (1999). Gene expression during vascular pericyte differentiation. Crit Rev. Eukaryot. Gene Expr. *9*, 1-17.

Dominici,M., Le,B.K., Mueller,I., Slaper-Cortenbach,I., Marini,F., Krause,D., Deans,R., Keating,A., Prockop,D., and Horwitz,E. (2006). Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. Cytotherapy. *8*, 315-317.

Eckfeldt,C.E., Mendenhall,E.M., and Verfaillie,C.M. (2005). The molecular repertoire of the 'almighty' stem cell. Nat. Rev. Mol. Cell Biol. *6*, 726-737.

Eckhardt,F. *et al.* (2006). DNA methylation profiling of human chromosomes 6, 20 and 22. Nat. Genet. *38*, 1378-1385.

Erickson,G.R., Gimble,J.M., Franklin,D.M., Rice,H.E., Awad,H., and Guilak,F. (2002). Chondrogenic potential of adipose tissue-derived stromal cells in vitro and in vivo. Biochem. Biophys. Res. Commun. *290*, 763-769.

Esteller,M. (2005). Aberrant DNA methylation as a cancer-inducing mechanism. Annu. Rev. Pharmacol. Toxicol. *45*, 629-656.

Evans,M.J. and Kaufman,M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154-156.

Fajas,L., Landsberg,R.L., Huss-Garcia,Y., Sardet,C., Lees,J.A., and Auwerx,J. (2002). E2Fs regulate adipocyte differentiation. Dev. Cell *3*, 39-49.

Farmer,S.R. (2006). Transcriptional control of adipocyte formation. Cell Metab. *4*, 263-273.

Feldman,N., Gerson,A., Fang,J., Li,E., Zhang,Y., Shinkai,Y., Cedar,H., and Bergman,Y. (2006). G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nat. Cell Biol. *8*, 188-194.

Fraser,J.K., Wulur,I., Alfonso,Z., and Hedrick,M.H. (2006). Fat tissue: an underappreciated source of stem cells for biotechnology. Trends Biotechnol. *24*, 150-154.

Friedenstein,A.J., Petrakova,K.V., Kurolesova,A.I., and Frolova,G.P. (1968). Heterotopic of bone marrow.Analysis of precursor cells for osteogenic and hematopoietic tissues. Transplantation *6*, 230-247.

Fu,Y.S., Cheng,Y.C., Lin,M.Y., Cheng,H., Chu,P.M., Chou,S.C., Shih,Y.H., Ko,M.H., and Sung,M.S. (2006). Conversion of human umbilical cord mesenchymal stem cells in Wharton's jelly to dopaminergic neurons in vitro: potential therapeutic application for Parkinsonism. Stem Cells *24*, 115-124.

Fuchs,E., Tumbar,T., and Guasch,G. (2004). Socializing with the neighbors: stem cells and their niche. Cell *116*, 769-778.

Fuks,F., Burgers,W.A., Brehm,A., Hughes-Davies,L., and Kouzarides,T. (2000). DNA methyltransferase Dnmt1 associates with histone deacetylase activity. Nat. Genet. *24*, 88-91.

Fuks,F., Burgers,W.A., Godin,N., Kasai,M., and Kouzarides,T. (2001). Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. EMBO J. *20*, 2536-2544.

Fuks,F., Hurd,P.J., Deplus,R., and Kouzarides,T. (2003). The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. Nucleic Acids Res. *31*, 2305-2312.

Futscher,B.W., Oshiro,M.M., Wozniak,R.J., Holtan,N., Hanigan,C.L., Duan,H., and Domann,F.E. (2002). Role for DNA methylation in the control of cell type specific maspin expression. Nat. Genet. *31*, 175-179.

Gardiner-Garden,M. and Frommer,M. (1987). CpG islands in vertebrate genomes. J. Mol. Biol. *196*, 261-282.

Geiman,T.M., Sankpal,U.T., Robertson,A.K., Zhao,Y., Zhao,Y., and Robertson,K.D. (2004). DNMT3B interacts with hSNF2H chromatin remodeling enzyme, HDACs 1 and 2, and components of the histone methylation system. Biochem. Biophys. Res. Commun. *318*, 544-555.

Gimble,J. and Guilak,F. (2003). Adipose-derived adult stem cells: isolation, characterization, and differentiation potential. Cytotherapy *5*, 362-369.

Goll,M.G., Kirpekar,F., Maggert,K.A., Yoder,J.A., Hsieh,C.L., Zhang,X., Golic,K.G., Jacobsen,S.E., and Bestor,T.H. (2006). Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. Science *311*, 395-398.

Gronthos,S., Graves,S.E., Ohta,S., and Simmons,P.J. (1994). The STRO-1+ fraction of adult human bone marrow contains the osteogenic precursors. Blood *84*, 4164-4173.

Gumina,R.J., Kirschbaum,N.E., Piotrowski,K., and Newman,P.J. (1997). Characterization of the human platelet/endothelial cell adhesion molecule-1 promoter: identification of a GATA-2 binding element required for optimal transcriptional activity. Blood *89*, 1260-1269.

Halvorsen,Y.D., Franklin,D., Bond,A.L., Hitt,D.C., Auchter,C., Boskey,A.L., Paschalis,E.P., Wilkison,W.O., and Gimble,J.M. (2001). Extracellular matrix mineralization and osteoblast gene expression by human adipose tissue-derived stromal cells. Tissue Eng *7*, 729-741.

Heil,M., Ziegelhoeffer,T., Mees,B., and Schaper,W. (2004). A different outlook on the role of bone marrow stem cells in vascular growth: bone marrow delivers software not hardware. Circ. Res. *94*, 573-574.

Hellman,A. and Chess,A. (2007). Gene body-specific methylation on the active X chromosome. Science *315*, 1141-1143.

Hendrich,B. and Bird,A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol. Cell Biol. *18*, 6538-6547.

Hermann,A., Schmitt,S., and Jeltsch,A. (2003). The human Dnmt2 has residual DNA-(cytosine-C5) methyltransferase activity. J. Biol. Chem. *278*, 31717-31721.

Hoffman,A.R. and Hu,J.F. (2006). Directing DNA methylation to inhibit gene expression. Cell Mol. Neurobiol. *26*, 425-438.

Huard,J., Cao,B., and Qu-Petersen,Z. (2003). Muscle-derived stem cells: potential for muscle regeneration. Birth Defects Res. C. Embryo. Today *69*, 230-237.

in 't Anker,P.S., Scherjon,S.A., Kleijburg-van der,K.C., Noort,W.A., Claas,F.H., Willemze,R., Fibbe,W.E., and Kanhai,H.H. (2003). Amniotic fluid as a novel source of mesenchymal stem cells for therapeutic transplantation. Blood *102*, 1548-1549.

Jaenisch,R. and Bird,A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet. *33 Suppl*, 245-254.

Jiang,Y. *et al.* (2002). Pluripotency of mesenchymal stem cells derived from adult marrow. Nature *418*, 41-49.

Jones,D.L. and Wagers,A.J. (2008). No place like home: anatomy and function of the stem cell niche. Nat. Rev. Mol. Cell Biol. *9*, 11-21.

Jones,P.A. and Baylin,S.B. (2002). The fundamental role of epigenetic events in cancer. Nat. Rev. Genet. *3*, 415-428.

Jones,P.A. and Takai,D. (2001). The role of DNA methylation in mammalian epigenetics. Science *293*, 1068-1070.

Kang,S.K., Lee,D.H., Bae,Y.C., Kim,H.K., Baik,S.Y., and Jung,J.S. (2003). Improvement of neurological deficits by intracerebral transplantation of human adipose tissue-derived stromal cells after cerebral ischemia in rats. Exp. Neurol. *183*, 355-366.

Karahuseyinoglu,S., Cinar,O., Kilic,E., Kara,F., Akay,G.G., Demiralp,D.O., Tukun,A., Uckan,D., and Can,A. (2007). Biology of stem cells in human umbilical cord stroma: in situ and in vitro surveys. Stem Cells *25*, 319-331.

Katz,A.J., Tholpady,A., Tholpady,S.S., Shang,H., and Ogle,R.C. (2005). Cell surface and transcriptional characterization of human adipose-derived adherent stromal (hADAS) cells. Stem Cells *23*, 412-423.

Kawasaki,H. and Taira,K. (2004). Induction of DNA methylation and gene silencing by short interfering RNAs in human cells. Nature *431*, 211-217.

Keating,A. (2006). Mesenchymal stromal cells. Curr. Opin. Hematol. *13*, 419-425.

Kern,S., Eichler,H., Stoeve,J., Kluter,H., and Bieback,K. (2006). Comparative analysis of mesenchymal stem cells from bone marrow, umbilical cord blood or adipose tissue. Stem Cells *24*, 1294-1301.

Ketterer,N. *et al.* (1998). High CD34(+) cell counts decrease hematologic toxicity of autologous peripheral blood progenitor cell transplantation. Blood *91*, 3148-3155.

Kilroy,G.E. *et al.* (2007). Cytokine profile of human adipose-derived stem cells: expression of angiogenic, hematopoietic, and pro-inflammatory factors. J. Cell Physiol *212*, 702-709.

Kim,D.H., Je,C.M., Sin,J.Y., and Jung,J.S. (2003). Effect of partial hepatectomy on in vivo engraftment after intravenous administration of human adipose tissue stromal cells in mouse. Microsurgery *23*, 424-431.

Kim,J.B., Wright,H.M., Wright,M., and Spiegelman,B.M. (1998). ADD1/SREBP1 activates PPARgamma through the production of endogenous ligand. Proc. Natl. Acad. Sci. U. S. A *95*, 4333-4337.

Kleber,M. and Sommer,L. (2004). Wnt signaling and the regulation of stem cell function. Curr. Opin. Cell Biol. *16*, 681-687.

Klein,G. (1995). The extracellular matrix of the hematopoietic microenvironment. Experientia *51*, 914-926.

Klose,R.J. and Bird,A.P. (2006). Genomic DNA methylation: the mark and its mediators. Trends Biochem. Sci. *31*, 89-97.

Knoblich,J.A. (2008). Mechanisms of asymmetric stem cell division. Cell *132*, 583-597.

Koutnikova,H., Cock,T.A., Watanabe,M., Houten,S.M., Champy,M.F., Dierich,A., and Auwerx,J. (2003). Compensation by the muscle limits the metabolic consequences of lipodystrophy in PPAR gamma hypomorphic mice. Proc. Natl. Acad. Sci. U. S. A *100*, 14457-14462.

Kouzarides,T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Krabbe,C., Zimmer,J., and Meyer,M. (2005). Neural transdifferentiation of mesenchymal stem cells--a critical review. APMIS *113*, 831-844.

Kucia,M., Reca,R., Jala,V.R., Dawn,B., Ratajczak,J., and Ratajczak,M.Z. (2005). Bone marrow as a home of heterogenous populations of nonhematopoietic stem cells. Leukemia *19*, 1118-1127.

Kung,E.F., Wang,F., and Schechner,J.S. (2008). In vivo perfusion of human skin substitutes with microvessels formed by adult circulating endothelial progenitor cells. Dermatol. Surg. *34*, 137-146.

Lachner,M., O'Carroll,D., Rea,S., Mechtler,K., and Jenuwein,T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116-120.

Laird,P.W. (2005). Cancer epigenetics. Hum. Mol. Genet. *14*, R65-R76.

Le,B.K., Tammik,C., Rosendahl,K., Zetterberg,E., and Ringden,O. (2003). HLA expression and immunologic properties of differentiated and undifferentiated mesenchymal stem cells. Exp. Hematol. *31*, 890-896.

Le,G.F. and Rudnicki,M.A. (2007). Skeletal muscle satellite cells and adult myogenesis. Curr. Opin. Cell Biol. *19*, 628-633.

Lee,J.H., Hart,S.R., and Skalnik,D.G. (2004a). Histone deacetylase activity is required for embryonic stem cell differentiation. Genesis *38*, 32-38.

Lee,R.H., Kim,B., Choi,I., Kim,H., Choi,H.S., Suh,K., Bae,Y.C., and Jung,J.S. (2004b). Characterization and expression analysis of mesenchymal stem cells from human bone marrow and adipose tissue. Cell Physiol Biochem. *14*, 311-324.

Lee,T.I. *et al.* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301-313.

Lund,R.D. *et al.* (2007). Cells isolated from umbilical cord tissue rescue photoreceptors and visual functions in a rodent model of retinal disease. Stem Cells *25*, 602-611.

MASTERS,E.M. (1965). Monolayer cultures of brown fat cells. Proc. Soc. Exp. Biol. Med. *119*, 44-46.

Matzke,M.A. and Birchler,J.A. (2005). RNAi-mediated pathways in the nucleus. Nat. Rev. Genet. *6*, 24-35.

McIntosh,K. *et al.* (2006). The immunogenicity of human adipose-derived cells: temporal changes in vitro. Stem Cells *24*, 1246-1253.

Mead,J.R., Irvine,S.A., and Ramji,D.P. (2002). Lipoprotein lipase: structure, function, regulation, and role in disease. J. Mol. Med. *80*, 753-769.

Meehan,R.R. (2003). DNA methylation in animal development. Semin. Cell Dev. Biol. *14*, 53-65.

Meehan,R.R., Lewis,J.D., McKay,S., Kleiner,E.L., and Bird,A.P. (1989). Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. Cell *58*, 499-507.

Melzner,I., Scott,V., Dorsch,K., Fischer,P., Wabitsch,M., Bruderlein,S., Hasel,C., and Moller,P. (2002). Leptin gene expression in human preadipocytes is switched on by maturation-induced demethylation of distinct CpGs in its proximal promoter. J. Biol. Chem. *277*, 45420-45427.

Meza-Zepeda,L.A., Noer,A., Dahl,J.A., Micci,F., Myklebost,O., and Collas,P. (2007). High-resolution analysis of genetic stability of human adipose tissue stem cells cultured to senescence. J. Cell. Mol. Med. *11*, doi: 10.1111/j.1582-4934.2007.00146x.

Mikkelsen,T.S. *et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.

Minasi,M.G. *et al.* (2002). The meso-angioblast: a multipotent, self-renewing cell that originates from the dorsal aorta and differentiates into most mesodermal tissues. Development *129*, 2773-2783.

Minguell,J.J., Erices,A., and Conget,P. (2001). Mesenchymal stem cells. Exp. Biol. Med. *226*, 507-520.

Miranville,A., Heeschen,C., Sengenes,C., Curat,C.A., Busse,R., and Bouloumie,A. (2004). Improvement of postnatal neovascularization by human adipose tissue-derived stem cells. Circulation *110*, 349-355.

Mito,Y., Henikoff,J.G., and Henikoff,S. (2005). Genome-scale profiling of histone H3.3 replacement patterns. Nat. Genet. *37*, 1090-1097.

Mito,Y., Henikoff,J.G., and Henikoff,S. (2007). Histone replacement marks the boundaries of cis-regulatory domains. Science. *315*, 1408-1411.

Miyahara,Y. *et al.* (2006). Monolayered mesenchymal stem cells repair scarred myocardium after myocardial infarction. Nat. Med. *12*, 459-465.

Miyoshi,I., Irino,S., and Hiraki,K. (1966). Fibroblast-like transformation of human bone marrow fat cells in vitro. Exp. Cell Res. *41*, 220-223.

Montarras,D., Morgan,J., Collins,C., Relaix,F., Zaffran,S., Cumano,A., Partridge,T., and Buckingham,M. (2005). Direct isolation of satellite cells for skeletal muscle regeneration. Science *309*, 2064-2067.

Morgan,H.D., Santos,F., Green,K., Dean,W., and Reik,W. (2005). Epigenetic reprogramming in mammals. Hum. Mol. Genet. *14*, R47-R58.

Morris,K.V., Chan,S.W., Jacobsen,S.E., and Looney,D.J. (2004). Small interfering RNA-induced transcriptional gene silencing in human cells. Science *305*, 1289-1292.

Morrison,S.J. and Weissman,I.L. (1994). The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. Immunity. *1*, 661-673.

Nagata,Y., Partridge,T.A., Matsuda,R., and Zammit,P.S. (2006). Entry of muscle satellite cells into the cell cycle requires sphingolipid signaling. J. Cell Biol. *174*, 245-253.

Nakagami,H. *et al.* (2005). Novel autologous cell therapy in ischemic limb disease through growth factor secretion by cultured adipose tissue-derived stromal cells. Arterioscler. Thromb. Vasc. Biol. *25*, 2542-2547.

Nanaev,A.K., Kohnen,G., Milovanov,A.P., Domogatsky,S.P., and Kaufmann,P. (1997). Stromal differentiation and architecture of the human umbilical cord. Placenta *18*, 53-64.

Noer,A., Boquest,A.C., and Collas,P. (2007). Dynamics of adipogenic promoter DNA methylation during clonal culture of human adipose stem cells to senescence. BMC. Cell Biol. *8*, 18-29.

Noer,A., Sørensen,A.L., Boquest,A.C., and Collas,P. (2006). Stable CpG hypomethylation of adipogenic promoters in freshly isolated, cultured and differentiated mesenchymal stem cells from adipose tissue. Mol. Biol. Cell *17*, 3543-3556.

Nolta,J.A., Dao,M.A., Wells,S., Smogorzewska,E.M., and Kohn,D.B. (1996). Transduction of pluripotent human hematopoietic stem cells demonstrated by clonal analysis after engraftment in immune-deficient mice. Proc. Natl. Acad. Sci. U. S. A *93*, 2414-2419.

Olguin,H.C., Yang,Z., Tapscott,S.J., and Olwin,B.B. (2007). Reciprocal inhibition between Pax7 and muscle regulatory factors modulates myogenic cell fate determination. J. Cell Biol. *177*, 769-779.

Otto,T.C. and Lane,M.D. (2005). Adipose development: from stem cell to adipocyte. Crit Rev. Biochem. Mol. Biol. *40*, 229-242.

Payne,T.R., Oshima,H., Sakai,T., Ling,Y., Gharaibeh,B., Cummins,J., and Huard,J. (2005). Regeneration of dystrophin-expressing myocytes in the mdx heart by skeletal muscle stem cells. Gene Ther. *12*, 1264-1274.

Peault,B., Rudnicki,M., Torrente,Y., Cossu,G., Tremblay,J.P., Partridge,T., Gussoni,E., Kunkel,L.M., and Huard,J. (2007). Stem and progenitor cells in skeletal muscle development, maintenance, and therapy. Mol. Ther. *15*, 867-877.

Pfaffl,M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. *29*, e45.

Pittenger,M.F., Mackay,A.M., Beck,S.C., Jaiswal,R.K., Douglas,R., Mosca,J.D., Moorman,M.A., Simonetti,D.W., Craig,S., and Marshak,D.R. (1999). Multilineage potential of adult human mesenchymal stem cells. Science *284*, 143-147.

Planat-Benard,V. *et al.* (2004). Plasticity of human adipose lineage cells toward endothelial cells: physiological and therapeutic perspectives. Circulation *109*, 656-663.

Porter,R.L. and Calvi,L.M. (2008). Communications between bone cells and hematopoietic stem cells. Arch. Biochem. Biophys. *473*, 193-200.

Prokhortchouk,A., Hendrich,B., Jorgensen,H., Ruzov,A., Wilm,M., Georgiev,G., Bird,A., and Prokhortchouk,E. (2001). The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. Genes Dev. *15*, 1613-1618.

Qiu,C., Sawada,K., Zhang,X., and Cheng,X. (2002). The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. Nat. Struct. Biol. *9*, 217-224.

Quesenberry,P.J., Colvin,G., and Abedi,M. (2005). Perspective: fundamental and clinical concepts on stem cell homing and engraftment: a journey to niches and beyond. Exp. Hematol. *33*, 9-19.

Rabinowicz,P.D., Palmer,L.E., May,B.P., Hemann,M.T., Lowe,S.W., McCombie,W.R., and Martienssen,R.A. (2003). Genes and transposons are differentially methylated in plants, but not in mammals. Genome Res. *13*, 2658-2664.

Rai,K., Chidester,S., Zavala,C.V., Manos,E.J., James,S.R., Karpf,A.R., Jones,D.A., and Cairns,B.R. (2007). Dnmt2 functions in the cytoplasm to promote liver, brain, and retina development in zebrafish. Genes Dev. *21*, 261-266.

Rehman,J., Traktuev,D., Li,J., Merfeld-Clauss,S., Temm-Grove,C.J., Bovenkerk,J.E., Pell,C.L., Johnstone,B.H., Considine,R.V., and March,K.L. (2004). Secretion of angiogenic and antiapoptotic factors by human adipose stromal cells. Circulation *109*, 1292-1298.

Reik,W., Dean,W., and Walter,J. (2001). Epigenetic reprogramming in mammalian development. Science *293*, 1089-1093.

Reyes,M., Lund,T., Lenvik,T., Aguiar,D., Koodie,L., and Verfaillie,C.M. (2001). Purification and ex vivo expansion of postnatal human marrow mesodermal progenitor cells. Blood *98*, 2615-2625.

Robertson,K.D. (2005). DNA methylation and human disease. Nat. Rev. Genet. *6*, 597-610.

Rosen,E.D. and MacDougald,O.A. (2006). Adipocyte differentiation from the inside out. Nat. Rev. Mol. Cell Biol. *7*, 885-896.

Safford,K.M., Hicok,K.C., Safford,S.D., Halvorsen,Y.D., Wilkison,W.O., Gimble,J.M., and Rice,H.E. (2002). Neurogenic differentiation of murine and human adipose-derived stromal cells. Biochem. Biophys. Res. Commun. *294*, 371-379.

Sanchez-Ramos,J. *et al.* (2000). Adult bone marrow stromal cells differentiate into neural cells in vitro. Exp. Neurol. *164*, 247-256.

Santoro,R. and Grummt,I. (2005). Epigenetic mechanism of rRNA gene silencing: temporal order of NoRC-mediated histone modification, chromatin remodeling, and DNA methylation. Mol. Cell Biol. *25*, 2539-2546.

Santoro,R., Li,J., and Grummt,I. (2002). The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. Nat. Genet. *32*, 393-396.

Santos-Rosa,H., Schneider,R., Bannister,A.J., Sherriff,J., Bernstein,B.E., Emre,N.C., Schreiber,S.L., Mellor,J., and Kouzarides,T. (2002). Active genes are tri-methylated at K4 of histone H3. Nature *419*, 407-411.

Schmeisser,A., Garlichs,C.D., Zhang,H., Eskafi,S., Graffy,C., Ludwig,J., Strasser,R.H., and Daniel,W.G. (2001). Monocytes coexpress endothelial and macrophagocytic lineage markers and form cord-like structures in Matrigel under angiogenic conditions. Cardiovasc. Res. *49*, 671-680.

Schoonjans,K., Gelman,L., Haby,C., Briggs,M., and Auwerx,J. (2000). Induction of LPL gene expression by sterols is mediated by a sterol regulatory element and is independent of the presence of multiple E boxes. J. Mol. Biol. *304*, 323-334.

Schultz,E. and McCormick,K.M. (1994). Skeletal muscle satellite cells. Rev. Physiol Biochem. Pharmacol. *123*, 213-257.

Serafini,M. and Verfaillie,C.M. (2006). Pluripotency in adult stem cells: state of the art. Semin. Reprod. Med. *24*, 379-388.

Shahdadfar,A., Fronsdal,K., Haug,T., Reinholt,F.P., and Brinchmann,J.E. (2005). In vitro expansion of human mesenchymal stem cells: choice of serum is a determinant of cell proliferation, differentiation, gene expression, and transcriptome stability. Stem Cells *23*, 1357-1366.

Shi,S. and Gronthos,S. (2003). Perivascular niche of postnatal mesenchymal stem cells in human bone marrow and dental pulp. J. Bone Miner. Res. *18*, 696-704.

Shi,X. and Garry,D.J. (2006). Muscle stem cells in development, regeneration, and disease. Genes Dev. *20*, 1692-1708.

Skålhegg,B.S., Tasken,K., Hansson,V., Huitfeldt,H.S., Jahnsen,T., and Lea,T. (1994). Location of cAMP-dependent protein kinase type I with the TCR-CD3 complex. Science *263*, 84-87.

Skovseth,D.K., Kuchler,A.M., and Haraldsen,G. (2007). The HUVEC/Matrigel assay: an in vivo assay of human angiogenesis suitable for drug validation. Methods Mol. Biol. *360*, 253-268.

Song,F., Smith,J.F., Kimura,M.T., Morrow,A.D., Matsuyama,T., Nagase,H., and Held,W.A. (2005). Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. Proc. Natl. Acad. Sci. U. S. A *102*, 3336-3341.

Song,X., Zhu,C.H., Doan,C., and Xie,T. (2002). Germline stem cells anchored by adherens junctions in the Drosophila ovary niches. Science *296*, 1855-1857.

Steidl,U. *et al.* (2004). Primary human CD34+ hematopoietic stem and progenitor cells express functionally active receptors of neuromediators. Blood *104*, 81-88.

Stenderup,K., Justesen,J., Clausen,C., and Kassem,M. (2003). Aging is associated with decreased maximal life span and accelerated senescence of bone marrow stromal cells. Bone *33*, 919-926.

Strem,B.M., Hicok,K.C., Zhu,M., Wulur,I., Alfonso,Z., Schreiber,R.E., Fraser,J.K., and Hedrick,M.H. (2005). Multipotential differentiation of adipose tissue-derived stem cells. Keio J. Med. *54*, 132-141.

Struhl,K. (1998). Histone acetylation and transcriptional regulatory mechanisms. Genes Dev. *12*, 599-606.

Suetake,I., Shinozaki,F., Miyagawa,J., Takeshima,H., and Tajima,S. (2004). DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. J. Biol. Chem. *279*, 27816-27823.

Szilvassy,S.J., Fraser,C.C., Eaves,C.J., Lansdorp,P.M., Eaves,A.C., and Humphries,R.K. (1989). Retrovirus-mediated gene transfer to purified hemopoietic stem cells with long-term lympho-myelopoietic repopulating ability. Proc. Natl. Acad. Sci. U. S. A *86*, 8798-8802.

Takai,D. and Jones,P.A. (2002). Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc. Natl. Acad. Sci. U. S. A *99*, 3740-3745.

Taranger,C.K., Noer,A., Sorensen,A.L., Hakelien,A.M., Boquest,A.C., and Collas,P. (2005). Induction of dedifferentiation, genome-wide transcriptional programming, and epigenetic reprogramming by extracts of carcinoma and embryonic stem cells. Mol. Biol. Cell *16*, 5719-5735.

Thomson,J.A., Itskovitz-Eldor,J., Shapiro,S.S., Waknitz,M.A., Swiergiel,J.J., Marshall,V.S., and Jones,J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Tontonoz,P., Hu,E., and Spiegelman,B.M. (1994). Stimulation of adipogenesis in fibroblasts by PPAR gamma 2, a lipid-activated transcription factor. Cell *79*, 1147-1156.

Torrente,Y. *et al.* (2004). Human circulating AC133(+) stem cells restore dystrophin expression and ameliorate function in dystrophic skeletal muscle. J. Clin. Invest *114*, 182-195.

Tremblay,K.D., Saam,J.R., Ingram,R.S., Tilghman,S.M., and Bartolomei,M.S. (1995). A paternal-specific methylation imprint marks the alleles of the mouse H19 gene. Nat. Genet. *9*, 407-413.

Troyer,D.L. and Weiss,M.L. (2008a). Concise Review: Wharton's Jelly-derived Cells are a Primitive Stromal Cell Population. Stem Cells *26*, 591-599.

Troyer,D.L. and Weiss,M.L. (2008b). Wharton's jelly-derived cells are a primitive stromal cell population. Stem Cells *26*, 591-599.

Turek-Plewa,J. and Jagodzinski,P.P. (2005). The role of mammalian DNA methyltransferases in the regulation of gene expression. Cell Mol. Biol. Lett. *10*, 631-647.

Uccelli,A., Frassoni,F., and Mancardi,G. (2007). Stem cells for multiple sclerosis: promises and reality. Regen. Med. *2*, 7-9.

Verfaillie,C.M., Pera,M.F., and Lansdorp,P.M. (2002). Stem cells: hype and reality. Hematology Am. Soc. Hematol. Educ. Program. *10*, 369-391.

Wabitsch,M., Brenner,R.E., Melzner,I., Braun,M., Moller,P., Heinze,E., Debatin,K.M., and Hauner,H. (2001). Characterization of a human preadipocyte cell strain with high capacity for adipose differentiation. Int. J. Obes. Relat Metab Disord. *25*, 8-15.

Wada,M.R., Inagawa-Ogashiwa,M., Shimizu,S., Yasumoto,S., and Hashimoto,N. (2002). Generation of different fates from multipotent muscle stem cells. Development *129*, 2987-2995.

Wagner,W. *et al.* (2005). Comparative characteristics of mesenchymal stem cells from human bone marrow, adipose tissue, and umbilical cord blood. Exp. Hematol. *33*, 1402-1416.

Wakitani,S., Saito,T., and Caplan,A.I. (1995). Myogenic cells derived from rat bone marrow mesenchymal stem cells exposed to 5-azacytidine. Muscle Nerve *18*, 1417-1426.

Wang,H.S., Hung,S.C., Peng,S.T., Huang,C.C., Wei,H.M., Guo,Y.J., Fu,Y.S., Lai,M.C., and Chen,C.C. (2004). Mesenchymal stem cells in the Wharton's jelly of the human umbilical cord. Stem Cells *22*, 1330-1337.

Weber,M., Davies,J.J., Wittig,D., Oakeley,E.J., Haase,M., Lam,W.L., and Schubeler,D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat. Genet. *37*, 853-862.

Weber,M., Hellmann,I., Stadler,M.B., Ramos,L., Paabo,S., Rebhan,M., and Schubeler,D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat. Genet. *39*, 457-466.

Weber,M. and Schubeler,D. (2007). Genomic patterns of DNA methylation: targets and function of an epigenetic mark. Curr. Opin. Cell Biol. *19*, 273-280.

Weimann,J.M., Johansson,C.B., Trejo,A., and Blau,H.M. (2003). Stable reprogrammed heterokaryons form spontaneously in Purkinje neurons after bone marrow transplant. Nat. Cell Biol. *5*, 959-966.

Weng,Y.S., Lin,H.Y., Hsiang,Y.J., Hsieh,C.T., and Li,W.T. (2003). The effects of different growth factors on human bone marrow stromal cells differentiating into hepatocyte-like cells. Adv. Exp. Med. Biol. *534*, 119-128.

Williams,J.T., Southerland,S.S., Souza,J., Calcutt,A.F., and Cartledge,R.G. (1999). Cells isolated from adult human skeletal muscle capable of differentiating into multiple mesodermal phenotypes. Am. Surg. *65*, 22-26.

Woodbury,D., Schwarz,E.J., Prockop,D.J., and Black,I.B. (2000). Adult rat and human bone marrow stromal cells differentiate into neurons. J. Neurosci. Res. *61*, 364-370.

Wu,K.H., Zhou,B., Lu,S.H., Feng,B., Yang,S.G., Du,W.T., Gu,D.S., Han,Z.C., and Liu,Y.L. (2007). In vitro and in vivo differentiation of human umbilical cord derived stem cells into endothelial cells. J. Cell Biochem. *100*, 608-616.

Wu,Z., Rosen,E.D., Brun,R., Hauser,S., Adelmant,G., Troy,A.E., McKeon,C., Darlington,G.J., and Spiegelman,B.M. (1999). Cross-regulation of C/EBP alpha and PPAR gamma controls the transcriptional pathway of adipogenesis and insulin sensitivity. Mol. Cell. *3*, 151-158.

Yang,W.S. and Deeb,S.S. (1998). Sp1 and Sp3 transactivate the human lipoprotein lipase gene promoter through binding to a CT element: synergy with the sterol regulatory element binding protein and reduced transactivation of a naturally occurring promoter variant. J. Lipid Res. *39*, 2054-2064.

Yatabe,Y., Tavare,S., and Shibata,D. (2001). Investigating stem cells in human colon by using methylation patterns. Proc. Natl. Acad. Sci. U. S. A. *98*, 10839-10844.

Zannettino,A.C., Paton,S., Arthur,A., Khor,F., Itescu,S., Gimble,J.M., and Gronthos,S. (2008). Multipotential human adipose-derived stromal stem cells exhibit a perivascular phenotype in vitro and in vivo. J. Cell Physiol *214*, 413-421.

Zhang,J.W., Klemm,D.J., Vinson,C., and Lane,M.D. (2004). Role of CREB in transcriptional regulation of CCAAT/enhancer-binding protein beta gene during adipogenesis. J. Biol. Chem. *279*, 4471-4478.

Zhao,S. *et al.* (2002). JAK2, complemented by a second signal from c-kit or flt-3, triggers extensive self-renewal of primary multipotential hemopoietic cells. EMBO J. *21*, 2159-2167.

Zhou,S. *et al.* (2001). The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. Nat. Med. *7*, 1028-1034.

Zilberman,D. and Henikoff,S. (2007). Genome-wide analysis of DNA methylation patterns. Development *134*, 3959-3965.

Zuk,P.A., Zhu,M., Ashjian,P., De Ugarte,D.A., Huang,J.I., Mizuno,H., Alfonso,Z.C., Fraser,J.K., Benhaim,P., and Hedrick,M.H. (2002). Human adipose tissue is a source of multipotent stem cells. Mol. Biol. Cell *13*, 4279-4295.

Zuk,P.A., Zhu,M., Mizuno,H., Huang,J., Futrell,J.W., Katz,A.J., Benhaim,P., Lorenz,H.P., and Hedrick,M.H. (2001). Multilineage cells from human adipose tissue: implications for cell-based therapies. Tissue Eng. *7*, 211-228.

# LIST OF ABBREVIATIONS

| | |
|---|---|
| ADP | Adenosine diphosphate |
| ATP | Adenosine triphosphate |
| ADSC | Adipose derived stem/stromal cell |
| ASC | Adipose stem cell |
| bp | Base pair |
| BMMSC | Bone marrow mesenchymal stem cell |
| BSA | Bovine serum albumin |
| cAMP | Cyclic adenosine monophosphate |
| C/EBP | CAAT enhancer binding protein |
| CFU-F | Colony forming unit fibroblast |
| CHIP | Chromatin immunoprecipitation |
| CpG | Cytosine-phosphate guanine dinucleotide |
| DAPI | 4',6-diamidino-2-phenylindole |
| DNA | Deoxyribonucleic acid |
| DNMT | DNA methyl transferase |
| ESC | Embryonic stem cell |
| FABP4 | Fatty acid binding protein 4 |
| FACS | Fluorescence-activated cell sorting |
| FCS | Fetal calf serum |
| HAT | Histone acetyl transferase |
| HCP | High CpG promoter |
| HDAC | Histone deacetylase |
| HDMase | Histone demethylase |
| H3K | Histone 3 lysine |
| H4K | Histone 4 lysine |
| HMT | Histone methyl transferase |
| HSC | Hematopoietic stem cell |
| HUVEC | Human umbilical vein endothelial cell |
| IgG | Immunoglobulin G |
| IP | Immunoprecipitate |
| m2 | Dimethylation |
| m3 | Trimethylation |
| LCP | Low CpG promoter |
| LEP | Leptin |
| LPL | Lipoprotein lipase |
| MBD | Methyl CpG binding domain |
| MBP | Methyl binding protein |
| 5-mC | 5-methylcytosine |
| MDSC | Muscle derived stem cell |
| MeDIP | Methyl DNA immunoprecipitation |
| MPC | Muscle progenitor cell |
| mRNA | Messenger ribonucleic acid |
| MSC | Mesenchymal stem cell |
| MYOG | Myogenin |
| *MYOG*(E) | *Myogenin* exon |
| *MYOG*(P) | *Myogenin* promoter |
| NCCIT | Human teratocarcinoma cell line |

| | |
|---|---|
| NMC | Norwegian microarray consortium |
| nt | Nucleotide |
| P | Passage |
| PBST | Phosphate buffered saline Tween-20 |
| PcG | Polycomb group protein |
| PCR | Polymerase chain reaction |
| PLA | Processed lipoaspirate cell |
| PPARγ | Peroxisome proliferator- activated receptor γ |
| RNA | Ribonucleic acid |
| RNAi | RNA interference |
| RNAPII | RNA polymerase II |
| RT-qPCR | Reverse transcriptase quantitative polymerase chain reaction |
| SGBS | Simpson-Golabi-Behmel syndrome |
| SP | Side population |
| TSS | Transcription start site |
| WJMSC | Wharton's Jelly mesenchymal stem cell |

# LIST OF SUPPLEMENTARY MATERIAL

**Supplementary Table 1.**
Bisulfite sequencing PCR and RT-PCR primers used in this study.

**Supplementary Table 2.**
*P* values from Fisher's exact tests on the percentage of methylated CpGs identified by bisulfite sequencing at indicated loci in ASCs, BMMSCs, WJMSCs, MPCs and HSCs.

**Supplementary Fig. 1.**
Bisulfite genomic sequencing results for the three overlapping CpGs shared between MYOG(P) and MYOG(E1) amplicons.

**Supplementary Fig. 2.**
Assessment of DNA fragment sizes in input DNA and MeDIP samples sent for Nimblegen promoter array analysis.

**Supplementary Fig. 3.**
Methylation profile of 3,400 kb fragment of chromosome 1.

**Supplementary Fig. 4.**
Similarity of methylation profiles between ASCs, BMMCs, MPCs and HSCs.

**Supplementary Fig. 5.**
A region differentially methylated in ASCs, BMMCs, MPCs and HSCs.

**Supplementary Manuscript.**
Sørensen, A.L. and Collas, P. 2009. DNA methylation profiling by methylated DNA immunoprecipitation (MeDIP). Methods Mol. Biol. Submitted.

# SUPPLEMENTARY MATERIAL

**Supplementary Table 1.** Bisulfite sequencing PCR and RT-PCR primers used in this study.

| Gene | F: forward primer<br>R: reverse primer | Product size (bp) |
|---|---|---|
| **Bisulfite primers** | | |
| *CD31* | F: TATGAGAAGGGTTTTTAAATTTAGG | 449 |
| | R: ACACTCACAAAACAAAAAAATCAAC | |
| *FABP4* | F: GGTAATTTTTGAGATAGGAGTGTTT | 413 |
| | R: CCAATTAAAAATAAAATCCAATCATTT | |
| *LEP* | F: GTTTTTGGAGGGATATTAAGGATTT | 361 |
| | R: TAACCTACCAAAAAAAACCAACAA | |
| *LPL* | F: GGGAGGATTGTAAGTGATAAATAGG | 457 |
| | R: CAACTAAAAATAAACAACTTTCCCTT | |
| *MYOG*(P) | F:GTTTTTAAGGTTTTTTTTAGTTTAAGTTTA | 429 |
| | R: TACTAAATACCATTTAAACCCTCCC | |
| *MYOG*(E1) | F: GGGGAATTATATTTAATTTATTGTAAA | 495 |
| | R: ACCTCATTCACCTTCTTAAACCTAC | |
| *PPARG2* | F: GTTGAAGTTTTTAAGAAAGTAAATT | 480 |
| | R:AAAAAAAATATTACCACACTATCTC | |
| **RT-PCR primers** | | |
| *CD31* | F: AGCAGCATCGTGGTCAACATA | 105 |
| | R: GATGGAGCAGGACAGGTTCAG | |
| *FABP4* | F: TCAGTGTGAATGGGGATGTGAT | 310 |
| | R: TTCAATGCGAACTTCAGTCCAG | |
| *GAPDH* | F: TTGCCATGGGTGGAATCATA | 148 |
| | R: TCGGAGTCAACGGATTTGGT | |
| *LEP* | F: TTTCACACACGCAGTCAGTCT | 61 |
| | R: CCAGGAATGAAGTCCAAACC | |
| *LPL* | F: CCTGAAGTTTCCACAAATAAGACC | 321 |
| | R: ATGCCGTTCTTTGTTCTGTAGAT | |
| *MYF5* | F: TCTGTGGCATATACATTTGATACAT | 177 |
| | R: GAGAGAGCAGGTGGAGAACTACT | |
| *MYOD1* | F: GGCGGAACTGCTACGAAGG | 113 |
| | R: GATGCGCTCCACGATGCT | |
| *MYOG* | F: CGTCTCCTGTTCTGGTCTCTTCC | 245 |
| | R: CCCTTTTCCCTGCCTGTCC | |
| *PAX3* | F: GCCGACTTGGAGAGGAAGGAG | 184 |
| | R: GCTGTTCTGCTGTGAAGGTGGTT | |
| *PAX7* | F: ATTCTTTGCCGCTACCAGGA | 177 |
| | R: GTCACAGTGCCCATCCTTCA | |
| *PPARG2* | F: CTTCCATTACGGAGAGATCCAC | 125 |
| | R: AAGCGATTCCTTCACTGATACAC | |
| **Primers used for PCR on genomic DNA from MeDIP and input samples** | | |
| *H19*ICR | F: GAGCCGCACCAGATCTTCAG | |
| | R: TTGGTGGAACACACTGTGATCA | |
| *UBE2B* | F: CTCAGGGGTGGATTGTTGAC | |
| | R: TGTGGATTCAAAGACCACGA | |

**Supplementary Table 2.** *P*-values from Fisher's exact tests on the percentage of methylated CpGs identified by bisulfite sequencing at indicated loci in ASCs, BMMSCs, WJMSCs, MPCs and HSCs[a].

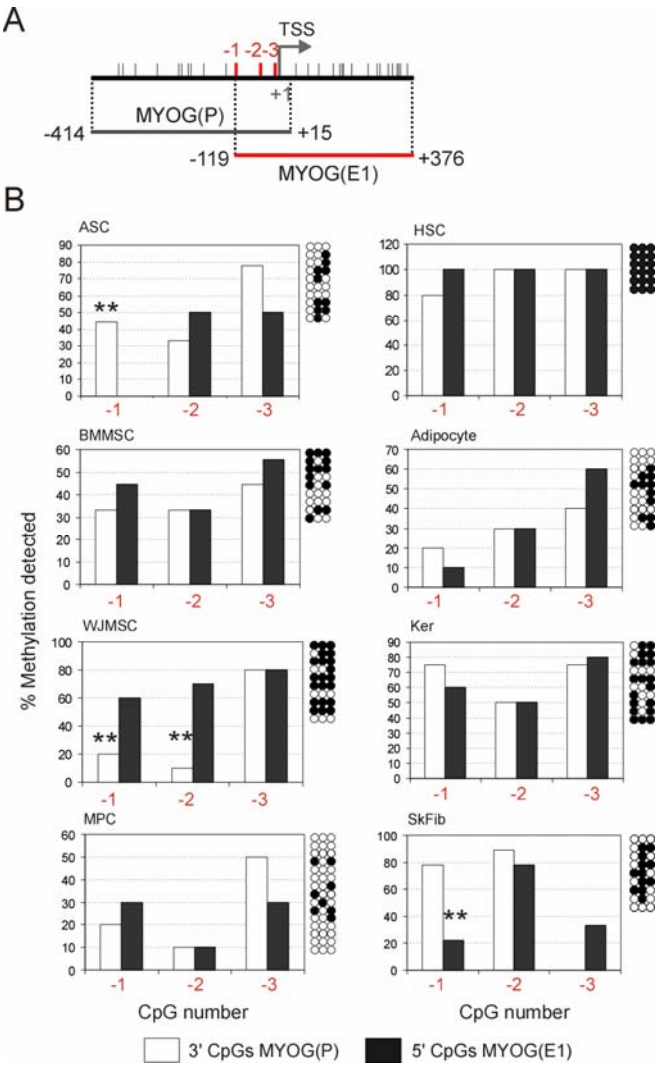
| | ASC | BMMSC | WJMSC | MPC | HSC |
|---|---|---|---|---|---|
| **LEP** | | | | | |
| % methylation | 9.7 | 9.7 | 12.6 | 10 | 40 |
| *P* value | >0.1 | ○[b] | ○ | ○ | |
| | <0.001 | | | | ○ |
| | | >0.1 | ○ | ○ | |
| | | <0.001 | | | ○ |
| | | | >0.1 | ○ | |
| | | | <0.001 | | ○ |
| | | | | <0.001 | ○ |
| **LPL** | | | | | |
| % methylation | 10.9 | 0 | 3.6 | 6.4 | 5.5 |
| *P* value | <0.01 | ○ | | | |
| | >0.1 | | ○ | ○ | ○ |
| | | | >0.1 | ○ | ○ |
| | | | | >0.1 | ○ |
| **FABP4** | | | | | |
| % methylation | 15.0 | 22.5 | 52.5 | 80.0 | 80.0 |
| *P* value | >0.1 | ○ | | | |
| | <0.001 | | ○ | ○ | ○ |
| | | <0.01 | ○ | ○ | ○ |
| | | | <0.02 | ○ | ○ |
| | | | | >0.1 | ○ |
| **PPARG2** | | | | | |
| % methylation | 20.0 | 36.7 | 63.3 | 40.7 | 90.3 |
| *P* value | >0.05 | ○ | | | |
| | <0.001 | | ○ | | ○ |
| | <0.03 | | | ○ | |
| | | <0.01 | ○ | | |
| | | >0.1 | | ○ | |
| | | <0.001 | | | ○ |
| | | | <0.03 | ○ | |
| | | | <0.001 | | ○ |
| **MYOG(P)** | | | | | |
| % methylation | 34.2 | 34.2 | 23.1 | 18.8 | 89.2 |
| *P* value | >0.1 | ○ | | | |
| | >0.05 | | ○ | | |
| | <0.01 | | | ○ | ○ |
| | | >0.05 | ○ | | |
| | | <0.01 | | ○ | ○ |
| | | | >0.1 | ○ | |
| | | | <0.001 | | ○ |
| | | | | <0.001 | ○ |

Supplementary Table 2 is continued on the next page.

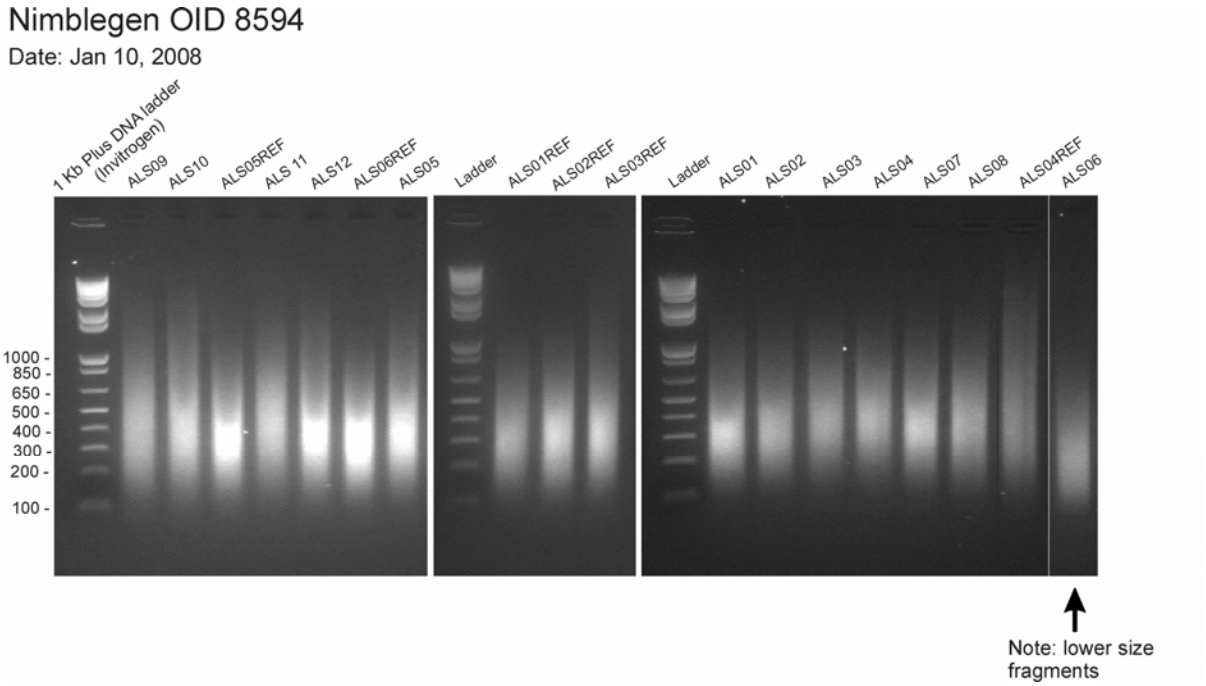
**Supplementary Table 2 (cont.)**. *P*-values from Fisher's exact tests on the percentage of methylated CpGs identified by bisulfite sequencing at indicated loci in ASCs, BMMSCs, WJMSCs, MPCs and HSCs[a].

| | | | | | |
|---|---|---|---|---|---|
| **MYOG(E1)** | | | | | |
| % methylation | 75.4 | 89.5 | 63.3 | 88.3 | 96.3 |
| *P* value | <0.01 | ○ | ○ | ○ | ○ |
| | | <0.001 | ○ | | |
| | | >0.05 | | ○ | ○ |
| | | | <0.001 | ○ | ○ |
| | | | | <0.03 | ○ |
| **CD31** | | | | | |
| % methylation | 93.3 | 73.9 | 95.0 | 90.0 | 6.1 |
| *P* value | <0.001 | ○ | | | ○ |
| | >0.1 | | ○ | ○ | |
| | | <0.001 | ○ | ○ | ○ |
| | | | >0.1 | ○ | |
| | | | <0.001 | | ○ |
| | | | | <0.001 | ○ |

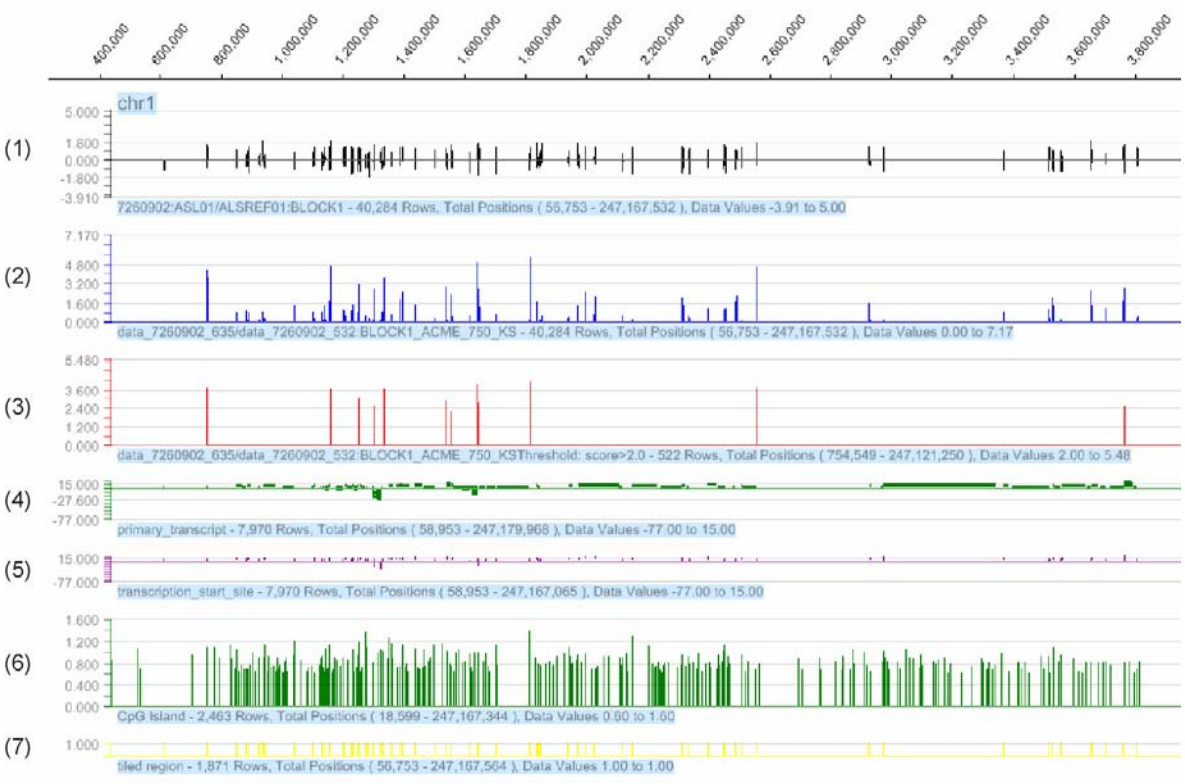
[a] Fisher's exact tests were calculated using the GaphPad software (www.graphpad.com/quickcalcs/contingency1.cfm).
[b] Denotes the cell type with which comparison is made.



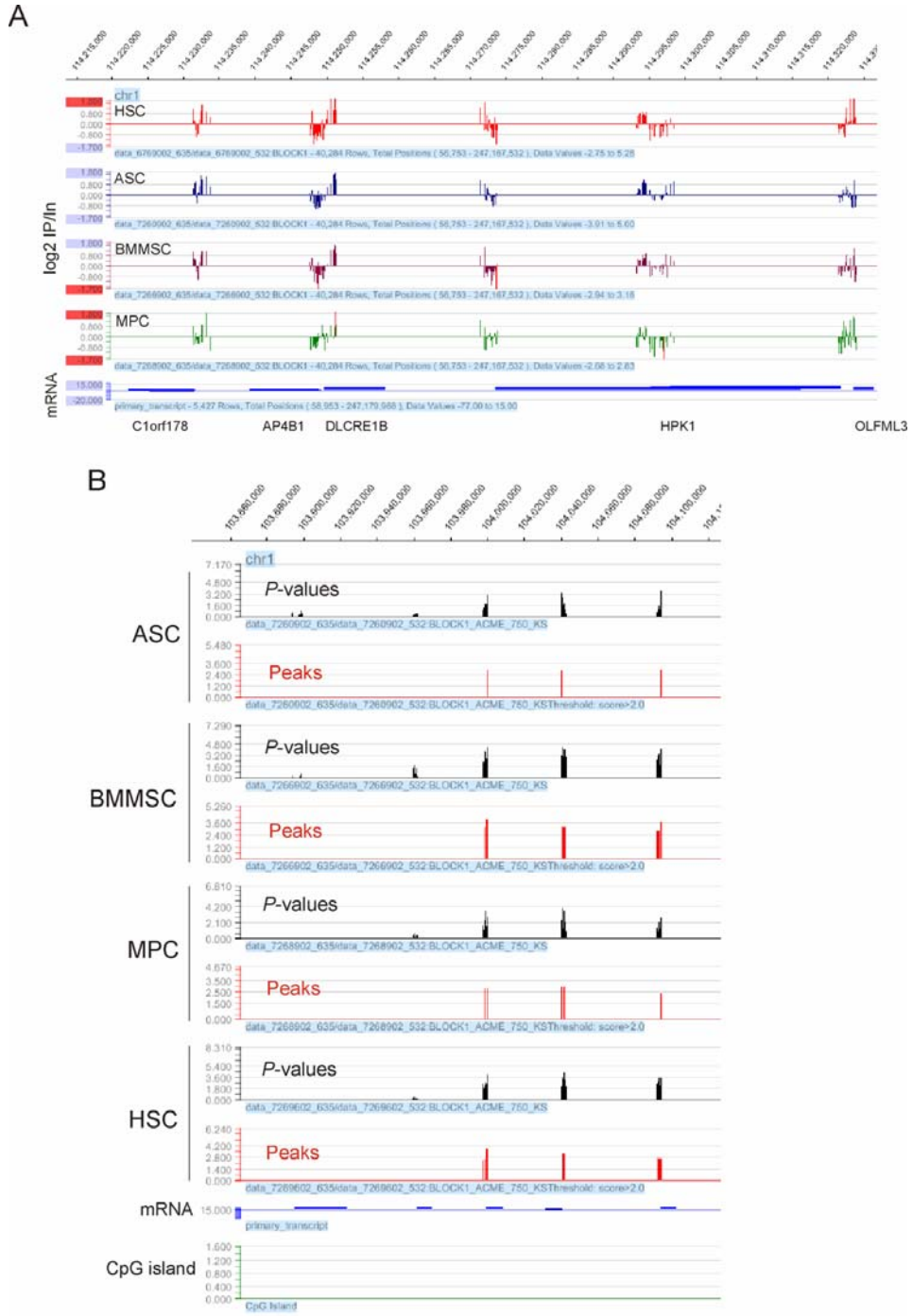

**Supplementary Fig. 1.** Bisulfite genomic sequencing results for the three overlapping CpGs shared between MYOG(P) and MYOG(E1) amplicons. (A) Map of the *MYOG* region examined. The three overlapping CpGs are shown in red and belong to the MYOG(E1) amplicon (red bar). (B) Bisulfite sequencing results (lollipop diagram) and comparison of the extent of methylation at each CpG in 3' region of MYOG(P) and 5' region of MYOG(E1). ** $P<0.01$.



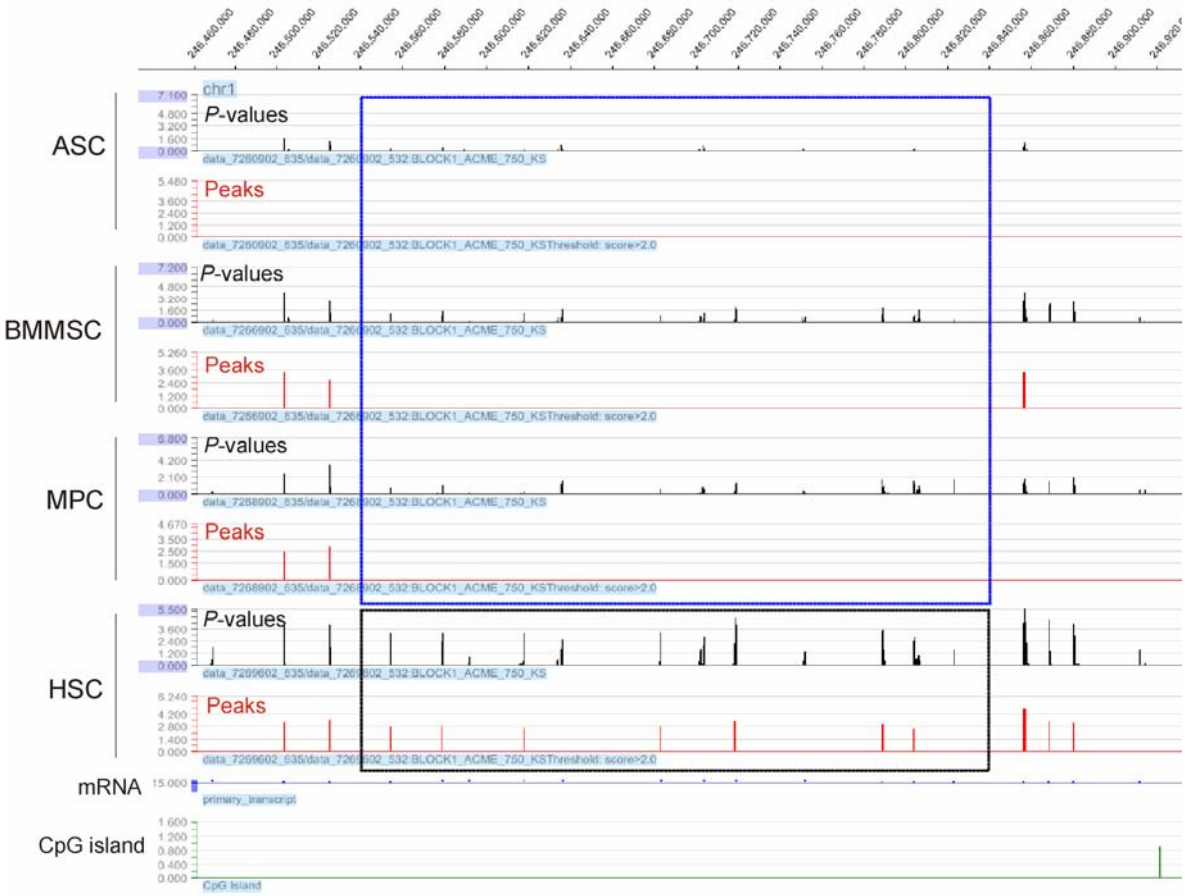

**Supplementary Fig. 2.** Assessment of DNA fragment sizes in input DNA and MeDIP samples sent to Nimblegen for promoter array hybridization. Input and MeDIP DNA was amplified as in Fig. R9 and resolved by agarose gel electrophoresis. Sample names (top) are those used as identification number in the Nimblegen data sheet. Samples ALS01REF, ALS02REF, ALS03REF, ALS01-04, ALS07 and ALS08 are those shown in Fig. R9B.




**Supplementary Fig. 3.** Methylation profile of 3,400 kb fragment of chromosome 1. Top, chromosome positioning (nucleotide number). Track (1): scaled $\log_2$ IP/input ratios. Track (2): *P*-values, expressed in –log10. Track (3): methylation peaks. Track (4): primary transcripts. Track (5): TSS. Track 6,: CpG islands. Track (7): tiled regions. Note the absence of CpG methylation in CpG island areas.




**Supplementary Fig. 4.** Similarity of methylation profiles between ASCs, BMMCs, MPCs and HSCs. (A) $\log_2$ IP/Input ratios in a randomly chosen 105 kb segment of chromosome 1, containing 5 loci. Note the similarity of the hybridization profiles between cell types (only one MeDIP replicate is shown for each cell type; replicates not shown are identical to those shown here). (B) P-values and methylation peaks in another randomly chosen 400 kb segment of chromosome 1. Unmethylated and methylated promoters are shown. Profiles are nearly identical for all cell types. This segment does not contain any CpG island (lower track).




**Supplementary Fig. 5.** A 460 kb segment of chromosome 1 showing different methylation profiles in ASCs, BMMSCs and MPCs on the one hand, and HSCs on the other hand. P-values and methylation peaks are shown. The olfactory receptor family 2 gene members are boxed in blue (unmethylated) and black (methylated).

**Supplementary manuscript**
**Submitted to Methods Mol. Biol.**

# Immunoprecipitation of Methylated DNA

**Anita L. Sørensen and Philippe Collas**

Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway

Corresponding author: Philippe Collas, Institute of Basic Medical Sciences, Department of Biochemistry, University of Oslo, PO Box 1112 Blindern, 0317 Oslo, Norway. Phone: (47) 2285 1066. Fax: (47) 2285 1058; Email: philippe.collas@medisin.uio.no

Running title: *Methylated-DNA immunoprecipitation*

## Abstract

DNA methylation contributes to the regulation of long-term gene repression by enabling the recruitment of transcriptional repressor complexes to methylated cytosines. Several methods for detecting DNA methylation at the gene-specific and genome-wide levels have been developed. Methylated DNA immunoprecipitation, or MeDIP, consists in the selective immunoprecipitation of methylated DNA fragments using antibodies to 5-methylcytosine. The genomic site of interest can be detected by PCR, hybridization to DNA arrays or by direct sequencing. This Chapter describes the MeDIP protocol and quality control tests that should be performed throughout the procedure.

**Key Word**s: DNA methylation, immunoprecipitation, 5-methylcytosine antibody, microarray

## 1. Introduction

DNA methylation consists in the post-replicative addition of a methyl group to the 5 position of a cytosine in a cytosine-phosphate-guanine (CpG) dinucleotide (**Fig.** 1**A**). CpG methylation alters the interaction of DNA with proteins which in turn may modulate transcription: it is either impaired, by methylation of activator sites, or enhanced, by methylation of insulators and silencers (***1***). CpG methylation in vertebrates is symmetrical (it occurs on both DNA strands) (**Fig. 1B**) and targets isolated or clustered CpGs. In plants, cytosines are methylated both symmetrically (CpG or CpNpG) and asymmetrically (CpNpNp), where N is any nucleotide. *Drosophila melanogaster* only exhibits DNA methylation in early stages of development, while *Saccharomyces cerevisiae* shows no DNA methylation

CpG methylation is catalyzed by DNA methyltransferases (DNMTs). The maintenance DNA methyltransferase DNMT1 specifically recognizes hemi-methylated DNA after replication and methylates the daughter strand, ensuring fidelity in the methylation profile after replication (***2***). In contrast to DNMT1, DNMT3a and DNMT3b are implicated in *de novo* DNA methylation that takes place during embryonic development and cell differentiation (***3***), as a means of shutting down genes whose activity is no longer required as cells differentiate (e.g., that of pluripotency-associated genes). The fourth DNMT, DNMT2,

has to date no clear ascribed function in DNA methylation but has been shown to have cytoplasmic transfer RNA methyltransferase activity (***4,5***).

DNA methylation is a hallmark of long-term gene silencing (***6,7***) (**Fig. 1C**). The methyl groups create target sites for methyl-binding proteins which induce transcriptional repression by recruiting co-repressor complexes, histone deacetylases or histone methyltransferases (***7***). DNA methylation is essential for development (***8-11***), X chromosome inactivation (***12***) and genomic imprinting (***13-15***). The relationship between DNA methylation and gene expression is complex (***1***) and recent evidence based on genome-wide CpG methylation profiling highlights promoter CpG content as a component of this complexity (***16***).

Several approaches have been developed to analyze DNA methylation profiles. Protocols relying on bisulfite conversion of DNA have been recently reviewed and the widely used bisulfite genomic sequencing approach has been extensively improved and described (***17***). An alternative to bisulfite sequencing is the immunoprecipitation of methylated DNA, referred to as methylated DNA immunoprecipitation, or MeDIP (***18***). The principle of MeDIP is simple: genomic DNA is randomly fragmented by sonication and methylated fragments are selectively immunoprecipitated using an antibody to 5-methyl cytosine (5mC). Detection of a gene of interest in the methylated DNA fraction can be done by polymerase chain reaction (PCR), hybridization to genomic arrays (MeDIP-chip) or high-throughout sequencing (MeDIP-seq) (***19,20***). We have used MeDIP-chip for the analysis of DNA methylation profiles in various mesenchymal stem cell (MSC) types. This chapter describes the MeDIP assay it is performed in our laboratory, including control tests that can be performed along the way (**Fig. 2**). The protocol is derived from that established in Dirk Schübeler's laboratory (***16,18***).

## 2. Materials

### *2.1. Materials*

1. 1.5-ml centrifuge tubes
2. Magnetic rack for 1.5-ml tubes (Invitrogen cat. no. MPC-E)
3. Probe sonicator (Sartorius Labsonic M sonicator fitted with 3-mm diameter probe, or similar)
4. Thermomixer (Eppendorf, model no. 5355-28402, or similar)
5. Table top centrifuge (Eppendorf, model no. 5424, or similar)
6. Minicentrifuge (Merck Eurolab Galaxy Mini, model no. C1211, or similar)
7. Vortex (VWR International, model no. 444-1372, or similar)
8. Rotator (Science Lab, model no. Stuart SB3, or similar), placed at 4$^{o}$C
9. Thermal cycler (PCR machine) with accessories

### *2.2. Reagents*

1. Anti-5mC antibody (Eurogentec cat. no. BI-MECY-1000; Diagenode cat. no. Mab-5MECYT-500)
2. Dynabeads® M-280 sheep anti-mouse IgG (Invitrogen, cat. no. 112.01)
3. GenomePlex Complete Whole Genome Amplification (WGA) Kit (Sigma-Aldrich, cat .no. WGA2-50RXN)
4. MinElute PCR Purification Kit (Qiagen cat. no. 28004)
5. 500 mM EDTA (Sigma-Aldrich, cat. no. E9884)
6. 1 M Tris-HCl, pH 8.0 (Sigma-Aldrich, cat. no. T3253)
7. 5 M NaCl (Sigma-Aldrich, cat. no. S5150)

8. Triton X-100 (Sigma-Aldrich, cat. no. T8787)
9. Bovine serum albumin (BSA) (Sigma-Aldrich cat. no.A7906)
10. Glycogen (Sigma-Aldrich, cat. no. G1767)
11. Proteinase K (Sigma-Aldrich, cat. no. P2308)
12. Phosphate buffered saline (PBS) (Sigma-Aldrich, cat. no. P4417)
13. RNase, DNase-free (Roche, cat. no. 11 119 915 001)
14. SDS (Sigma-Aldrich, cat. no. L4509)
15. 3 M NaAc (Sigma-Aldrich, cat. no. S8750)
16. Phenol-chloroform isoamylalcohol (25:24:1; Invitrogen, cat. no. 15593-031)
17. Chloroform isoamylalcohol (24:1; Sigma-Aldrich, cat. no. C0549)
18. 96% ethanol at -20°C
19. 70% ethanol at -20°C
20. PCR Master Mix (Promega, cat. no. M7505)
21. Deionized (e.g., MilliQ) water
22. Crushed ice

### *2.3. Buffers*

1. Cell lysis buffer: 10 mM Tris-HCl pH 8.0, 0.1 M EDTA pH 8.0, 0.5% SDS.
2. 10x immunoprecipitation (IP) buffer: 1.4 M NaCl, 100 mM Na-Phosphate (pH 7.0), 0.5% Triton X-100.
3. Na-phosphate: 39 ml of 2 M $NaH_2PO_4$ (276 mg/ml), 61 ml 2 M $Na_2HPO_4$ (284 mg/ml), 100 ml MilliQ water. This makes a 1 M Na-phosphate solution at pH 7.0.
4. PBS: dissolve 1 tablet in 200 ml MilliQ water.
5. PBS-BSA solution: 0.05 g BSA in 50 ml PBS (i.e., 0.1% BSA).
6. TE buffer: 10 mM Tris-HCl, pH 8.0, 10 mM EDTA.
7. Proteinase K stock solution: 20 mg/ml Proteinase K in MilliQ water.

## 3. Methods

### *3.1. Purification of Genomic DNA*

The procedure described here is for DNA purification from $10^6$ cells harvested and washed by your standard protocol to obtain enough DNA for a duplicate MeDIP (*see* **Note 1**).

1. Suspend the cell pellet ($10^6$ cells) in 400 µl of cell lysis buffer in a 1.5-ml centrifuge tube.
2. Add 1.2 µl Proteinase K solution (20 mg/ml stock) per 100 µl cell lysis buffer.
3. Incubate at 55°C for 1 h.
4. Add another 0.6 µl of Proteinase K solution per 100 µl cell lysis buffer and incubate at 37°C overnight.
5. Add one volume of phenol-chloroform isoamylalcohol, centrifuge at 15,000 *g* for 5 min and transfer the aqueous phase to a clean tube.
6. Add one volume of chloroform isoamylalcohol, centrifuge at 15,000 *g* for 5 min and transfer the aqueous phase to a clean tube.
7. Precipitate the DNA by adding 0.1 volumes of 3 M NaAc and 2.5 volumes of 96% ethanol at -20°C; mix and incubate for 1 h at -20°C.
8. Centrifuge at 20,000 *g* for 15 min at 4°C and remove the supernatant.
9. Add 0.5 ml of 70% ethanol to wash the pellet, vortex and centrifuge at 20,000 *g* for 10 min at 4°C. Remove the supernatant.
10. Collect the DNA pellets (from each tube) into one tube.

11. Dissolve the DNA in 100 µl TE buffer per $10^6$ cells (~50 µg DNA).

### *3.2. RNAse Treatment of Genomic DNA*

It is imperative to treat the genomic DNA with RNase because the antibody also recognizes 5-methylcytidine in the context of RNA.

1. Place 100 µl genomic DNA (~50 µg DNA) into a 1.5-ml tube. The amount of DNA to be RNase-treated should be that needed to continue with MeDIP.
2. Add 6 µl RNase solution (final concentration, 30 µg/ml) and incubate for 2 h at 37°C.

### *3.3. Fragmentation of Genomic DNA*

1. Dilute the RNase-treated DNA in a total of 200 µl in TE pH 8.0 in a 1.5-ml tube placed on ice.
2. Sonicate on ice for 3 x 30 sec, with 30 sec pauses on ice between each sonication session, using the probe sonicator. With the Labsonic M sonicator, we use the following settings: cycle 0.5, 30% power (*see* **Note 2**).
3. Repeat for each DNA sample (if relevant) while leaving the sonicated samples on ice.
4. To assess fragmentation, resolve 4 µl of sonicated DNA by 1.6% agarose gel electrophoresis and ethidium bromide staining (**Fig. 3A**).
5. If necessary, continue with sonication until the desired DNA fragment length is achieved.
6. Precipitate the sonicated DNA by adding 1 µl of glycogen, 400 mM NaCl and 2 volumes of 100% ethanol, mix and incubate at -80°C for 1 h. Thaw the tubes and centrifuge at 20,000 *g* for 15 min at 4°C. Decant the ethanol. it is important to remove all the ethanol. The pellet may be left to dry at room temperature for 15 min.
7. Dissolve the DNA in 60 µl MilliQ $H_2O$ and measure DNA concentration.

### *3.4. Immunoprecipitation of Methylated DNA*

1. Dilute 4 µg sonicated DNA in 450 µl TE buffer. Remember to store the rest of the DNA for input.
2. Denature for 10 min in boiling water and immediately chill on ice for 10 min.
3. Add 51 µl of 10x IP buffer.
4. Add 10 µl of anti-5mC antibody and incubate for 2 h at 4°C on a rotator set to 40 rpm.
5. Pre-wash 40 µl of Dynabeads with 800 µl PBS-BSA for 5 min at room temperature with shaking, 800 rpm on Thermomixer.
6. Place on magnetic rack to collect the beads; remove the PBS-BSA and repeat the wash (step 5) with 800 µl PBS-BSA
7. Collect the beads with a magnetic rack and resuspend in 40 µl of 1x IP buffer.
8. Add Dynabeads to the sample and incubate for 2 h at 4°C on a rotator set at 40 rpm.
9. Place the tube on a magnetic rack to collect the beads and wash with 700 µl 1x IP buffer for 10 minutes at room temperature on a Thermomixer at 950 rpm.
10. Repeat wash step 9 once.
11. Transfer the content of the tube to a clean 1.5-ml tube. This tube shift step eliminates any non-specifically bound DNA stuck on the tube wall which may give rise to background in the analysis.
12. Place the tube on the magnetic rack, collect the beads and wash once with 700 µl IP buffer at room temperature on a Thermomixer at 950 rpm.
13. Collect the beads with the magnetic rack and resuspend in 250 µl Proteinase K digestion buffer.

14. Add 3.5 µl proteinase K solution (20 mg/ml stock).
15. Incubate for 3 h at 50°C on a Thermomixer at 950 rpm.
16. Extract DNA once with 250 µl phenol-chloroform isoamylalcohol and once with 250 µl chloroform isoamylalcohol.
17. Precipitate DNA by adding 20 µl of 5 M NaCl stock (400 mM final concentration), 1 µl glycogen and 500 µl 100% ethanol; mix and incubate at -80°C for 1 h. Centrifuge at 20,000*g* for 15 min at 4°C. Make sure to remove all ethanol after centrifugation.
18. Dissolve the DNA in 15 µl $H_2O$ overnight at 4°C.
19. Measure DNA concentration with a Nanodrop and store at -20°C.

### *3.5. Amplification of Precipitated DNA*

Following immunoprecipitation, the yield of MeDIP DNA is low (300-450 ng in our hands) and incompatible with hybridization to microarrays (Nimblegen promoter arrays require 2 µg DNA per array). A genomic DNA amplification step, followed by a clean up, are therefore introduced in the protocol. For MeDIP and input DNA amplification, we use the Sigma WGA2 GenomePlex Complete Whole Genome Amplification Kit but omit the DNA fragmentation step (*see* **Note 3**).

#### *3.5.1. Amplification of precipitated and input DNA*

1. Place 11 µl (i.e., 100 ng) of MeDIP DNA into a 0.6-ml tube, and add 2 µl of 1x library preparation buffer provided with the WGA2 kit, and 1 µl of library stabilization solution.
2. Repeat steps 1 and 2 with the input DNA sample.
3. Vortex thoroughly, centrifuge briefly and denature by placing in a thermal cycler at 95°C for 2 min.
4. Place the samples on ice, centrifuge briefly and return the samples on ice.
5. Add 1 µl of library preparation enzyme solution provided with the WGA2 kit, vortex well and centrifuge briefly.
6. Place samples in a thermal cycler and run the following program: 16°C for 20 min, 24°C for 20 min, 37°C for 20 min, 75°C for 5 min and hold at 4°C.
7. Remove samples from the thermal cycler, centrifuge briefly and either freeze and store at -20°C (for up to 3 days) or proceed with amplification.
8. Prepare a PCR master mix by adding to the 15-µl reaction from step 7: 7.5 µl of 10x amplification master mix, 47.5 µl nuclease-free water and 5 µl of WGA DNA polymerase (provided with the WGA2 kit).
9. Vortex thoroughly and centrifuge briefly.
10. Incubate samples in a thermal cycler with the following program: 95°C for 3 min and 14 cycles of 94°C for 15 sec (dematuration) and 65°C for 5 min (annealing/extension). Hold the reaction at 4°C when completed. The amplified sample can be stored at -20°C, similarly to genomic DNA.

#### *3.5.2. Clean up the amplified DNA*

The amplified DNA needs to be cleaned up regardless of the mode of analysis that follows. We use the MinElute PCR Purification Kit protocol from Qiagen as per manufacturer's instructions. The kit is designed for purification of DNA fragments ranging from 70 bp to 4 kb and thus is well suited for fragmented MeDIP and input DNA fragments.

1. Add 5 volumes of PBI buffer provided with the kit to the amplified DNA sample, i.e., 75 µl buffer to the 15 µl sample of step 10 in **Subheading 3.5.1**, and mix.

2. Make sure the color of the mixture is yellow, indicative of the right pH. If it is not yellow, at 10 µl of 3 M sodium acetate, pH 5.0, and mix. The mixture will turn yellow.
3. Place a MinElute column in a provided 3 ml collection tube in a rack, add the DNA sample from step 2 and centrifuge for 1 min to bind DNA to the membrane in the column.
4. Discard the flow-through and return the column back into the same tube.
5. Add 750 µl buffer PE provided with the kit (wash buffer) and centrifuge for 1 min.
6. Discard the flow-through, return the column to the tube and centrifuge again for 1 min at full speed.
7. Place the MinElute column into a clean 1.5 ml centrifuge tube.
8. Add 10 µl MilliQ water to the center of the membrane, let the column stand for 1 min and centrifuge for 1 min. The eluate volume should be ~9 µl starting from 10 µl elution solution (*see* **Note 4**).

### *3.6. Analysis of Precipitated DNA*

Analysis of DNA methylation can be performed by PCR, hybridization to genomic arrays or by high-throughout sequencing (***16,18,21-23***).

#### *3.6.1. Assessment of DNA fragment size distribution*

Before any analysis of MeDIP DNA we routinely check that the range of fragment sizes is conserved between multiple samples. This is done by 1.2% agarose gel electrophoresis of an aliquot of amplified and cleaned up DNA. **Figure 3B** shows that all input and MeDIP DNA samples display fragment sizes uniformly ranging from 200 to ~850 bp, with most fragments around 300 bp.

#### *3.6.2. PCR analysis*

It is recommended to verify the specificity of the MeDIP by PCR analysis of immunoprecipitated and input DNA samples prior to fluorescent labeling and array hybridization, primarily due to the labor and costs involved. Both input (positive control) and MeDIP DNA samples should be analyzed using primer pairs to loci known to be methylated or unmethylated (such as housekeeping genes). As methylated control locus, we use the *H19* imprinted control region (*H19*ICR) with the following primer pair: 5'-GAGCCGCACCAGATCTTCAG-3' and 5'-TTGGTGGAACACACTGTGATCA-3' (annealing temperature 60°C). As unmethylated control locus, we use primers to the promoter of the housekeeping *UBE2B* gene with the following primer pair: 5'-CTCAGGGGTGGATTGTTGAC-3' and 5'-TGTGGATTCAAAGACCACGA-3' (annealing temperature 60°C). **Figure 3C** illustrates the result of MeDIP and input DNA sample PCR analysis using the above primers for three different cell types. Note that in NCCIT embryonal carcinoma cells, the *H19*ICR is unmethylated (**Fig. 3C**). A list of candidate methylated and unmethylated loci and respective PCR tests after MeDIP has been recently published (***16***).

#### *3.6.3. Microarray-based analysis*

Several commercial platforms exist for hybridization of MeDIP samples. Choice of platform depends on the experimental objective (e.g., investigation of CpG islands specifically, or promoters), array design, probe density, previous experience and cost. We have used Nimblegen human HG18 RefSeq Promoter arrays (cat. no. C4226-00-01; www.nimblegen.com) in an investigation of methylation profiles in the promoters of various cell types (see also Ref. (***16***) for an earlier version of these arrays). **Figure 4** illustrates the confirmation that MeDIP-array results match the PCR data (see **Fig. 3C** for the *UBE2B*

promoter methylation) and the PCR and array data of Weber and colleagues (***16***). In addition, **Figure 5** shows the MeDIP methylation profiles of several adipogenic, myogenic and endothelial gene promoters in human adipose tissue stem cells. These profiles are in agreement with CpG methylation patterns reported by bisulfite genomic sequencing (***24-26***).

### *3.6.4. High-throughout Sequencing*

By similarity to chromatin immunoprecipitation (ChIP)-sequencing approaches (***27***), it is possible to analyze MeDIP DNA in an unbiased manner by direct quantitative high-throughput sequencing (***23***).

## 4. Notes

1. We find that $10^6$ cells are usually sufficient for one duplicate MeDIP. Note that the extent of DNA recovery may vary between cell types.

2. Sonication should produce DNA fragments of ~300 to 1,000 bp (**Fig. 3A**). DNA fragments of less than 200 bp will be difficult to label when analyzed by microarray. The sonication protocol reported here is suitable for a variety of cell lines and primary cell types such as NCCIT cells, 293T cells, skin fibroblasts, keratinocytes, adipose-, bone marrow- and muscle-derived MSCs or hematopoietic stem cells). Optimization of sonication conditions may be required for a different cell type and other sonicator models. Samples should not foam as this reduces sonication efficiency.

3. DNA amplification: The Sigma WGA2 DNA amplification procedure includes a three-step process: DNA fragmentation, library generation and PCR amplification. The first two steps, fragmentation and library generation, are recommended by the manufacturer to be carried out without interruption. However, in the MeDIP assay, the DNA to be amplified is already fragmented, so we omit the fragmentation step of the WGA2 protocol and start at the library preparation step. The WGA2 kit recommends starting with 10-100 ng DNA; however we consistently start with 100 ng DNA. At the end of the amplification procedure as described in **Subheading 3.5.1.**, we obtain 20 µl of DNA at ~500 ng/µl, or a total of approximately 10 µg DNA.

4. We elute DNA with MilliQ water, as performed by Farnham and colleagues (***28***). However, it should be possible to use the Qiagen elution buffer (EB) buffer provided with the clean up kit. We do not recommend eluting with TE buffer because EDTA may inhibit further enzymatic reaction, notably during sample labeling for array hybridization.

## 5. References


1. Jones, P. A. and Takai, D. (2001) The role of DNA methylation in mammalian epigenetics. *Science* **293**, 1068-1070.

2. Jaenisch, R. and Bird, A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* **33 Suppl**, 245-254.

3. Turek-Plewa, J. and Jagodzinski, P. P. (2005) The role of mammalian DNA methyltransferases in the regulation of gene expression. *Cell Mol. Biol. Lett.* **10**, 631-647.

4. Rai, K., Chidester, S., Zavala, C. V., Manos, E. J., James, S. R., Karpf, A. R., Jones, D. A. and Cairns, B. R. (2007) Dnmt2 functions in the cytoplasm to promote liver, brain, and retina development in zebrafish. *Genes Dev.* **21**, 261-266.

5. Goll, M. G., Kirpekar, F., Maggert, K. A., Yoder, J. A., Hsieh, C. L., Zhang, X., Golic, K. G., Jacobsen, S. E. and Bestor, T. H. (2006) Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science* **311**, 395-398.

6. Hoffman, A. R. and Hu, J. F. (2006) Directing DNA methylation to inhibit gene expression. *Cell Mol. Neurobiol.* **26**, 425-438.

7. Klose, R. J. and Bird, A. P. (2006) Genomic DNA methylation: the mark and its mediators. *Trends Biochem. Sci.* **31**, 89-97.

8. Morgan, H. D., Santos, F., Green, K., Dean, W. and Reik, W. (2005) Epigenetic reprogramming in mammals. *Hum. Mol. Genet.* **14**, R47-R58.

9. Young, L. E. and Beaujean, N. (2004) DNA methylation in the preimplantation embryo: the differing stories of the mouse and sheep. *Anim. Reprod. Sci.* **82**, 61-78.

10. Mann, J. R. (2001) Imprinting in the germ line. *Stem Cells* **19**, 287-294.

11. Razin, A. and Shemer, R. (1995) DNA methylation in early development. *Hum. Mol. Genet.* **4**, 1751-1755.

12. Hellman, A. and Chess, A. (2007) Gene body-specific methylation on the active X chromosome. *Science* **315**, 1141-1143.

13. Tremblay, K. D., Saam, J. R., Ingram, R. S., Tilghman, S. M. and Bartolomei, M. S. (1995) A paternal-specific methylation imprint marks the alleles of the mouse H19 gene. *Nat. Genet.* **9**, 407-413.

14. Sapienza, C., Peterson, A. C., Rossant, J. and Balling, R. (1987) Degree of methylation of transgenes is dependent on gamete of origin. *Nature* **328**, 251-254.

15. Reik, W., Collick, A., Norris, M. L., Barton, S. C. and Surani, M. A. (1987) Genomic imprinting determines methylation of parental alleles in transgenic mice. *Nature* **328**, 248-251.

16. Weber, M., Hellmann, I., Stadler, M. B., Ramos, L., Paabo, S., Rebhan, M. and Schubeler, D. (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* **39**, 457-466.

17. Clark, S. J., Statham, A., Stirzaker, C., Molloy, P. L. and Frommer, M. (2006) DNA methylation: bisulphite modification and analysis. *Nat. Protoc.* **1**, 2353-2364.

18. Weber, M., Davies, J. J., Wittig, D., Oakeley, E. J., Haase, M., Lam, W. L. and Schubeler, D. (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* **37**, 853-862.

19. Zilberman, D. and Henikoff, S. (2007) Genome-wide analysis of DNA methylation patterns. *Development* **134**, 3959-3965.

20. Jacinto, F. V., Ballestar, E. and Esteller, M. (2008) Methyl-DNA immunoprecipitation (MeDIP): hunting down the DNA methylome. *Biotechniques* **44**, 35, 37, 39.

21. Penterman, J., Zilberman, D., Huh, J. H., Ballinger, T., Henikoff, S. and Fischer, R. L. (2007) DNA demethylation in the Arabidopsis genome. *Proc. Natl. Acad. Sci U. S. A* **104**, 6752-6757.

22. Zilberman, D., Gehring, M., Tran, R. K., Ballinger, T. and Henikoff, S. (2007) Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nat. Genet.* **39**, 61-69.

23. Jacinto, F. V., Ballestar, E., Ropero, S. and Esteller, M. (2007) Discovery of epigenetically silenced genes by methylated DNA immunoprecipitation in colon cancer cells. *Cancer Res.* **67**, 11481-11486.

24. Boquest, A. C., Noer, A., Sorensen, A. L., Vekterud, K. and Collas, P. (2007) CpG methylation profiles of endothelial cell-specific gene promoter regions in adipose tissue stem cells suggest limited differentiation potential toward the endothelial cell lineage. *Stem Cells* **25**, 852-861.

25. Noer, A., Sørensen, A. L., Boquest, A. C. and Collas, P. (2006) Stable CpG hypomethylation of adipogenic promoters in freshly isolated, cultured and differentiated mesenchymal stem cells from adipose tissue. *Mol. Biol. Cell* **17**, 3543-3556.

26. Noer, A., Boquest, A. C. and Collas, P. (2007) Dynamics of adipogenic promoter DNA methylation during clonal culture of human adipose stem cells to senescence. *BMC Cell Biol.* **8**, 18-29.

27. Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T. K., Koche, R. P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., Nusbaum, C., Lander, E. S. and Bernstein, B. E. (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553-560.

28. Acevedo, L. G., Iniguez, A. L., Holster, H. L., Zhang, X., Green, R. and Farnham, P. J. (2007) Genome-scale ChIP-chip analysis using 10,000 human cells. *Biotechniques* **43**, 791-797.

## Acknowledgements

The basis for this MeDIP protocol has been the procedure established in Dirk Schübeler's laboratory (Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland) by Michaël Weber and Dirk Schübeler and posted on the Epigenome Network of Excellence website (http://www.epigenome-noe.net/researchtools/protocol.php?protid=33). We are also grateful to Dirk Schübeler for discussion and advice.

## Figure Legends

Fig. 1. Principles of DNA methylation. (A) Mechanism of DNA methylation. (B) CpG methylation is symmetrical and occurs on both DNA strands. (C) DNA methylation correlates with long-term gene repression.

Fig. 2. The MeDIP assay. Genomic DNA is purified from cells, fragmented to ~300-1,000 bp by sonication, and 5-mC enriched fragments are immunoprecipitated using an anti-5mC antibody. Precipitated and input DNA are amplified. For array-based analysis, the input DNA is labeled with Cy3 while and MeDIP (precipitated) DNA is labeled with Cy5.

Fig. 3. Quality assessments of DNA during the MeDIP assay. (A) Assessment of fragmentation by sonication. Both intact (non-fragmented) and sonicated DNA samples from adipose stem cells (ASCs), bone marrow mesenchymal stem cells (BMMSCs) and muscle progenitor cells (MPCs) were analyzed by electrophoresis in 1.2% agarose and ethidium bromide staining. (B) Assessment of input and MeDIP DNA fragment size distribution and uniformity after amplification. Input and MeDIP DNA samples from ASCs, BMMSCs and MPCs were amplified, purified and resolved by electrophoresis in 1.2% agarose and ethidium bromide staining. Note the uniformity of the fragment sizes. Such samples are ready for processing for labeling and microarray hybridization. (C) PCR analysis of specificity of the MeDIP assay. Input and MeDIP DNA was analyzed by PCR using primers specific for the human *H19* Imprinting Control Region (*H19* ICR), which is methylated in somatic cells but not in male germ-cell-derived embryonal carcinoma NCCIT cells, and for the human *UBE2B* (ubiquitin-conjugating enzyme E2B) locus, which is unmethylated. DNA from ASCs, BMMSCs and NCCIT cells was used in this analysis. PCR products were resolved by agarose gel electrophoresis and stained with ethidium bromide.

Fig. 4. Validation of the MeDIP assay: MeDIP methylation profile of genes known to be methylated and unmethylated, in adipose stem cells (ASCs), bone marrow MSCs (BMMSCs), muscle progenitor cells (MPCs) and hematopoietic stem cells (HSCs). (A) Methylation profiles of two known methylated promoters (*OXT, LDHC*). (B) Methylation profiles of two known unmethylated housekeeping gene promoters (*UBE2B, PEX13*). $Log_2$ IP/input ratios, *P*-values and transcripts are shown.

Fig. 5. MeDIP methylation profiles of the *LEP*, *LPL*, *FABP4*, *PPARG2*, *MYOG* and *CD31* promoters in adipose stem cells. Rectangle boxes delineate genomic regions analyzed earlier by bisulfite sequencing (***24-26***). A CpG density track (CpG) is also shown.

Fig. 1.

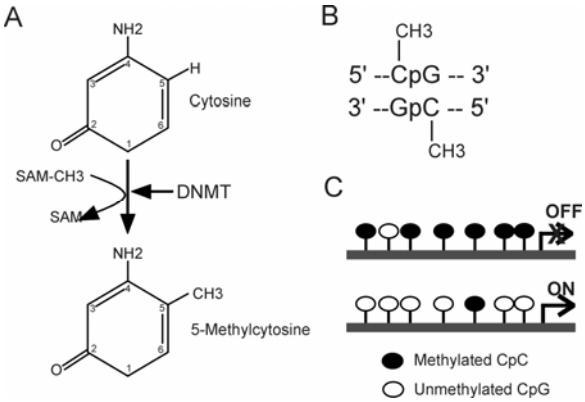

A
NH2
H
Cytosine
O
SAM-CH3
DNMT
SAM
NH2
CH3
5-Methylcytosine
O
B
CH3
5' --CpG -- 3'
3' --GpC -- 5'
CH3
C
OFF
ON
Methylated CpC
Unmethylated CpG

Fig. 2.

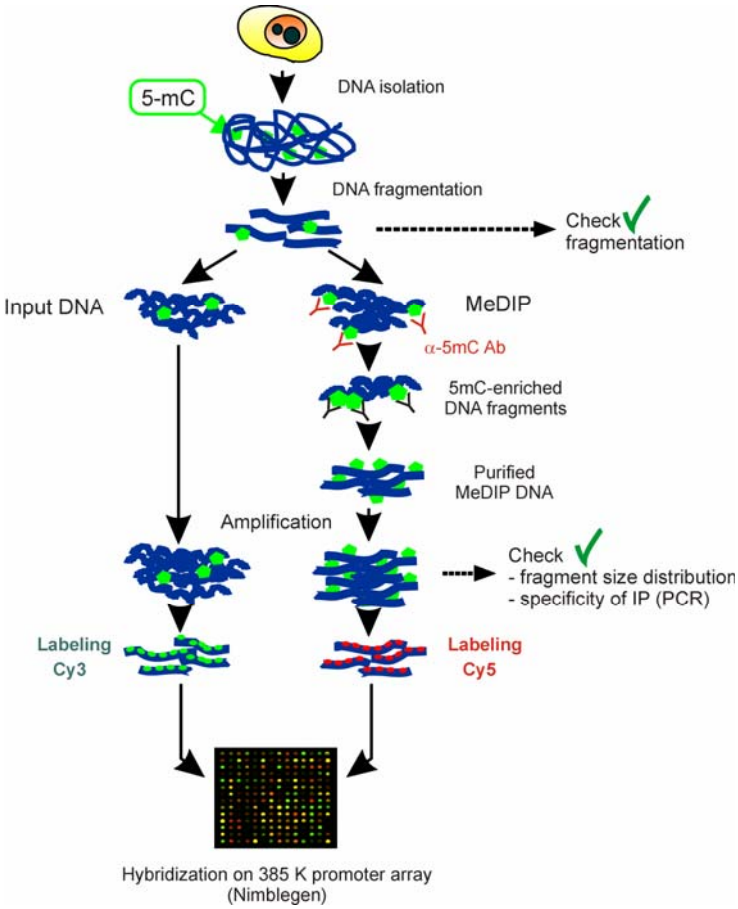

5-mC
DNA isolation
DNA fragmentation
Check
fragmentation
Input DNA
MeDIP
α-5mC Ab
5mC-enriched
DNA fragments
Purified
MeDIP DNA
Amplification
Check
- fragment size distribution
- specificity of IP (PCR)
Labeling
Cy3
Labeling
Cy5
Hybridization on 385 K promoter array
(Nimblegen)

Fig. 3.

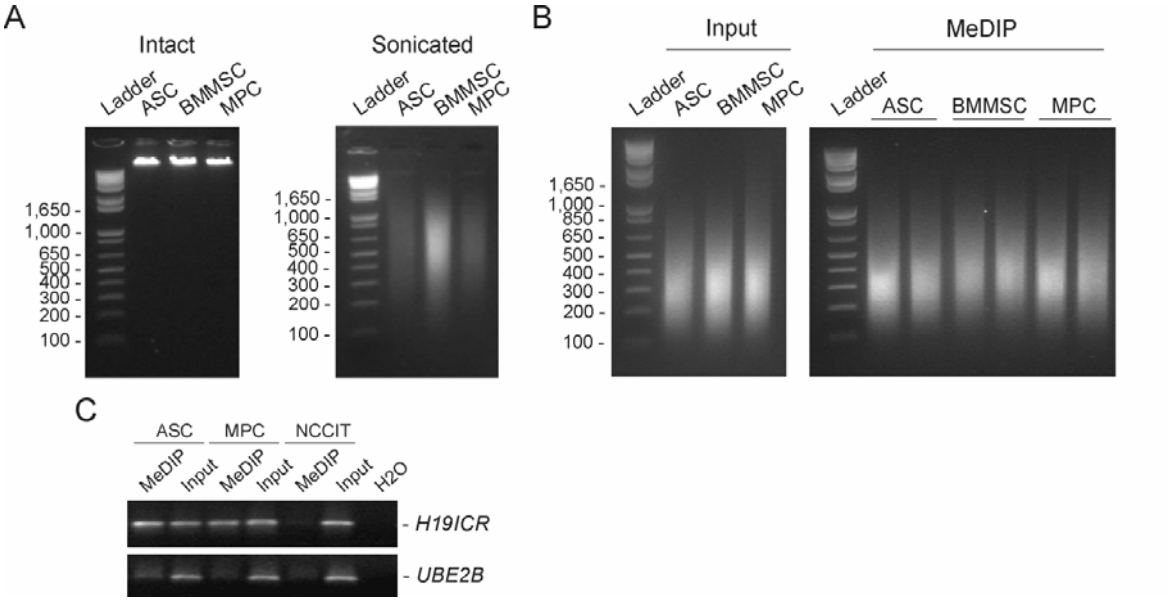

A
Intact
Sonicated
Ladder
ASC
BMMSC
MPC
1,650 -
1,000 -
650 -
500 -
400 -
300 -
200 -
100 -
B
Input
MeDIP
C
ASC
MPC
NCCIT
MeDIP
Input
H2O
- H19ICR
- UBE2B

Fig. 4.

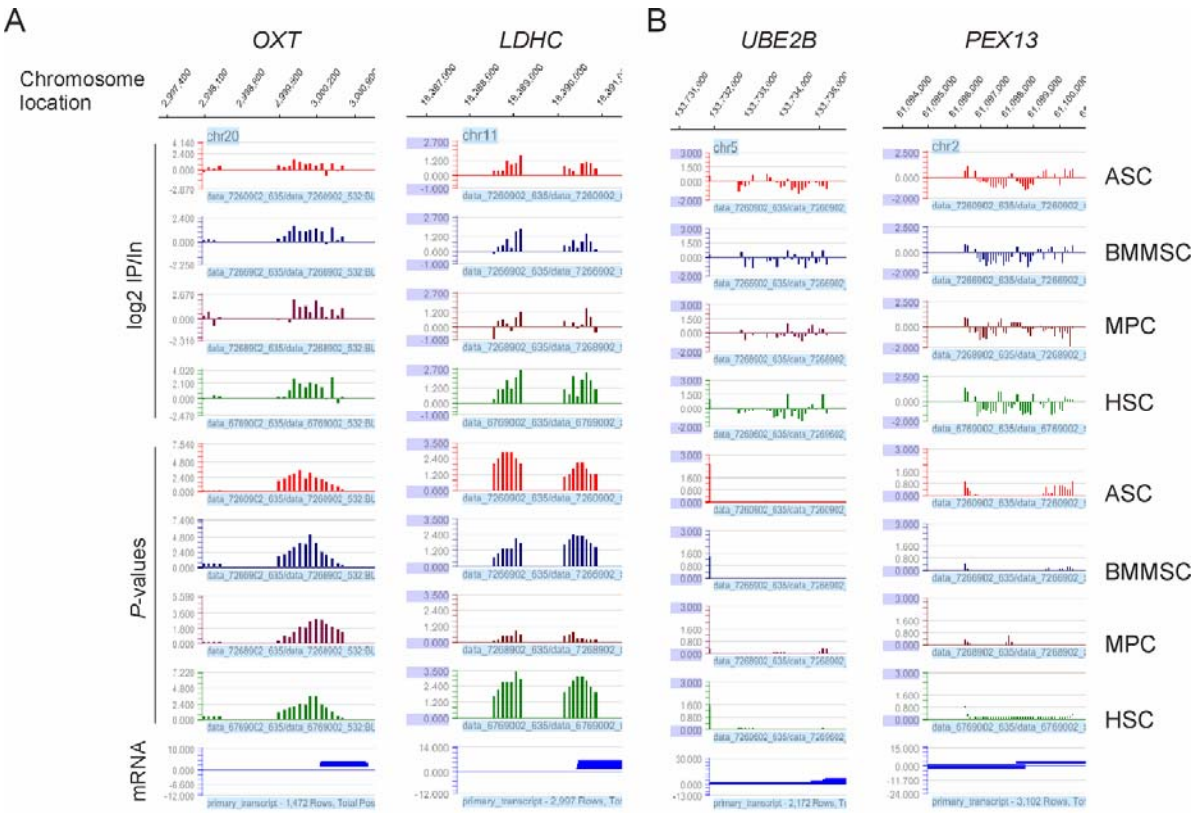
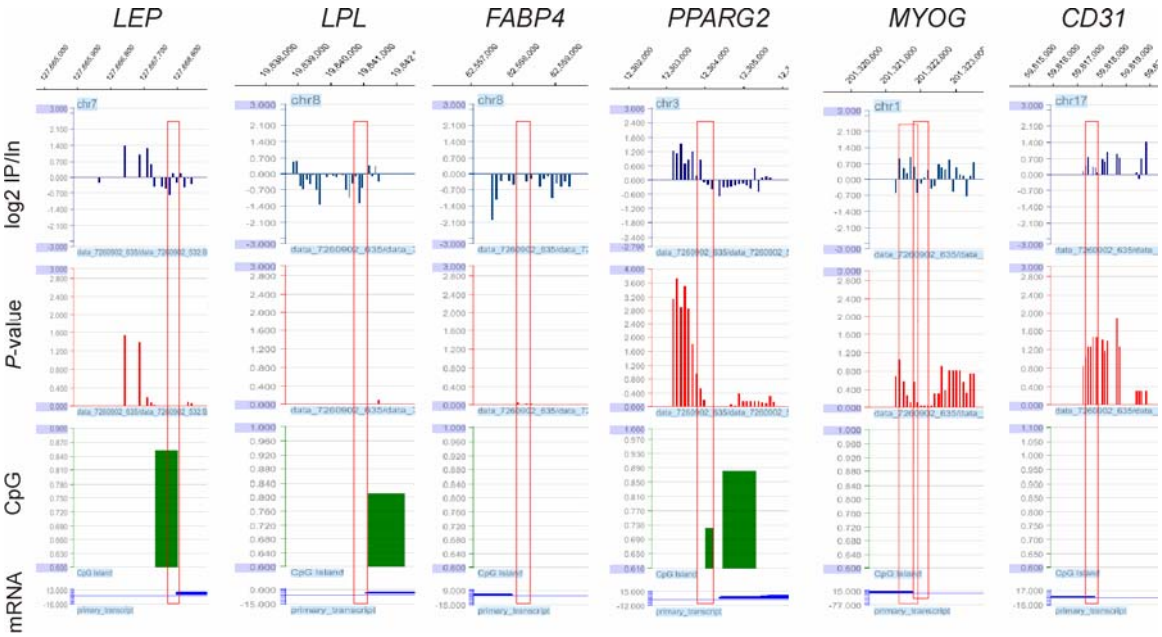

A
OXT
LDHC
B
UBE2B
PEX13
Chromosome location
chr20
chr11
chr5
chr2
log2 IP/In
P-values
mRNA
ASC
BMMSC
MPC
HSC

Fig. 5.


LEP
LPL
FABP4
PPARG2
MYOG
CD31
log2 IP/In
P-value
CpG
mRNA
chr7
chr8
chr8
chr3
chr1
chr17
CpG Island
primary_transcript