SUPPLEMENTAL MATERIAL

SUPPLEMENTAL MATERIALS AND METHODS

Supplemental Table 1 - Primer composition.

A) For restriction endonuclease analysis		
Region	Forward	Reverse
5'-flanking	5'-TGGAGTGGTCCTGATGTGGTAGTGG-3' (nt 1022 – 1046)	5'-ACCCAGAGATAAATATAGCCAACGC-3' (nt 1496 – 1520)
Exon 1	5'-TTCCTGTCCACCTTCAGGGCTTCG-3' (nt 1686 – 1709)	5'-GCGCTCAATGTACTGGATGGCG-3' (nt 1990 – 2011)
Exon 3	5'-TCCATCGTGGACAGCATCACG-3' (nt 3266 – 3286)	5'-TAAGGAGTCAGCTAAATTCCCTCGC-3' (nt 3507 – 3531)
B) For bisulphite analysis		
Region	Forward	Reverse
5'-flanking (plus strand)	5'-AGGAAAGAGAAGG$^{C}/_{T}$TAAGTGG-3' (nt 1106 – 1126)	5'-ACCCCACCCT$^{G}/_{A}$CA$^{G}/_{A}$$^{G}/_{A}$CC-3' (nt 1385 – 1401)

Supplemental Table 2 - Features of the endonucleases used. The / and the * indicate, respectively, the cutting site and the methylation sensitivity of each enzyme.

	Recognition sequence	Cutting(s) in 5'-flanking	Cutting(s) in exon 1	Notes
BglI	GCCNNNN/NGGC	1302, 1372	nd	no effect of methylation (in recognition sites investigated)
PvuII	CAG/C*TG	1195	none	inhibited by methylation in C*
AvaI	C/YCGRG	1232	nd	coincident with HpaII site (1233) but no effect of methylation
EcoNI	CCTNN/NNNAGG	1374	nd	no effect of methylation
EcoRI	G/AATTC	none	none	
HpaII	C*/C*GG	1233	1715, 1748, 1861	inhibited if either the external C* or the internal C* is methylated
MspI	C*/CGG	1233	1715, 1748, 1861	inhibited if the external C* is methylated

Rreb1. Zn-finger DNA binding factor that binds to a ras-responsive element. Like CTCF, Rreb1 has also been shown to bind divergent sequences. It has been shown to be involved in the differentiation response to ras. **Entrez Gene link for Rreb1**	No studies on the influence of CpG methylation	No studies on the influence of non-CpG methylation
Nfe2l2 (NRF-2). Transcriptional activator of a number of mitochondrial genes related to respiratory chain. It acts specifically in mithocondrial biogenesis, even during myogenesis, together with Sp1 (***Kraft et al.***, *2006, Am. J. Physiol Cell Physiol, 290, C1119-C1127*). **Entrez Gene link for Nfe2l2**	There are experimental results indicating that Nfe2l2 promoter polymorphisms may influence the CpG methylation of the p14 gene in human gastric mucosa (***Arisawa et al.***, *2008, Oncol. Rep., 19, 211-216*). Site-specific methylation of two CpG binding sites for transcription factor Nfe2l2 in the TOMM70 subunit of the outer mitochondrial membrane translocase abolished Nfe2l2 binding, thereby down-regulating the expression of a luciferase reporter in HeLa cells (***Blesa et al.***, *2008, Gene, 427, 58-64*).	No studies on the influence of non-CpG methylation
Snai1. It is a member of the Snail family of Zn-finger DNA binding factors, whose main direct target is E-cadherin, that is involved in the epithelial-mesenchymal transition in embryos of several organisms. Snail members compete for DNA binding with the basic-Helix-Loop-Helix transcription factors, sharing the same consensus binding site. They are cyclic genes, expressed periodically, that coordinate the segmentation and the presomitic mesoderm morphogenesis. **Entrez Gene link for Snai1**	In cell lines of human hepatoma, the ability of Snail genes both to induce promoter CpG methylation, by recruitment of HDAC1 and Dnmt1, and consequently to downregulate the E-cadherin gene has been demonstrated (***Lim et al.***, *2008, Gastroenterology, 135, 2128-40, 2140*).	No studies on the influence of non-CpG methylation
Cebpb (C/EBPβ, CCAAT / enhancer binding protein-β). It is a transcription factor that controls the differentiation and proliferation of several cell types. **Entrez Gene link for Cebpb**	The binding of Cebpb has been shown to be inhibited by CpG methylation in the promoter of the methionine adenosyltransferase gene, a key gene in the synthesis of S-adenosylmethionine (the main cellular methyl donor) in Hepa-1 cells; DNA hypomethylation favours Cebpb binding and expression of the gene (***Ikeda et al.***, *2008, Int. J. Biochem. Cell Biol., 40, 1956-1969)*.	No studies on the influence of non-CpG methylation

Definitions (with examples)

Methylation of specific C moieties. For a specific C moiety, the ratio between the number of clones where that moiety is methylated and the number of all analyzed clones (for that moiety), expressed as a percentage.

Example of calculation based on the exemplificative table shown below:

Methylation of C number 1 = 5 methylated clones / 5 analyzed clones = 100%
Methylation of C number 2 = 4 methylated clones / 5 analyzed clones = 80%
Methylation of C number 3 = 3 methylated clones / 5 analyzed clones = 60%
Methylation of C number 4 = 2 methylated clones / 5 analyzed clones = 40%
Methylation of C number 5 = 2 methylated clones / 5 analyzed clones = 40%
Methylation of C number 6 = 2 methylated clones / 5 analyzed clones = 40%

Group of C. A variable number of C grouped together according to a criterion. Examples are: a) all C of the whole 5'-flanking region analyzed, b) all CpN (all CpG or CpC or CpT or CpA) (these groups are called according to the kind of dinucleotide), c) all C in a cluster (as defined by cluster analysis and dendrogram) (these groups are called by numbering the cluster: cluster 1, 2, 3, 4, 5, 6), d) all C non-grouping by cluster analysis and dendrogram (these were called single-non-grouped-C (sng-C)), e) all C included in a cluster (as defined by cluster analysis and dendrogram) and spatially close together (without any other C, included in a different cluster, in the middle) (this group is called sub-cluster of adjacent C (adj-C) and numbered according to the cluster number followed by a letter: adj-C 1a, 1b, 2a, 2b, 2c, etc.....).

Overall methylation. It is a measurement, in percentage, of the total methylation level of a specific group of C; it is calculated as the number of methylated C of that group, divided by the total number of analyzed C (including all other C not belonging to that group); the calculation includes clone replicates.

Example of calculation based on the example table shown below:

Overall methylation of all sites = 18 methylated sites / 30 analyzed sites = 60.0%
Overall methylation of cluster made up of C 1-3-5 = 10 methylated sites / 30 analyzed sites = 33.3%

Overall methylation of cluster made up of C 2-4-6 = 8 methylated sites / 30 analyzed sites = 26.7%

Overall methylation of different groups can be added. In the example, the sum of the overall methylation of the cluster made up of C 1-3-5 and of the cluster made up of C 2-4-6 is 60%, equal to the overall methylation of all the sites.

Exemple table upon which the examples of calculation are based (6 sites (C) analyzed in 5 clones)

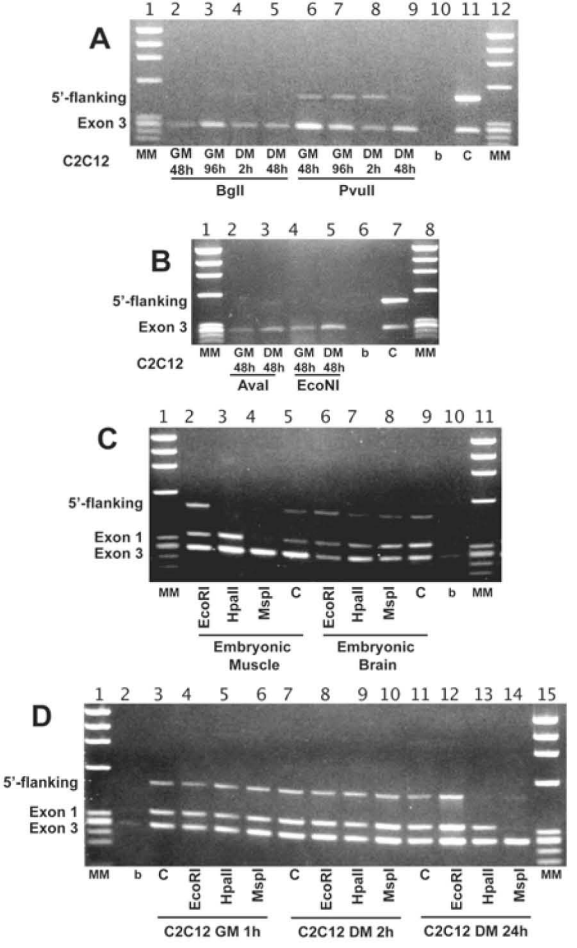
Sites (C): clones	*1*	*2*	*3*	*4*	*5*	*6*
1	methylated	methylated	methylated	methylated	methylated	methylated
2	methylated	unmethylated	methylated	unmethylated	unmethylated	unmethylated
3	methylated	methylated	methylated	methylated	methylated	unmethylated
4	methylated	methylated	unmethylated	unmethylated	unmethylated	unmethylated
5	methylated	methylated	unmethylated	unmethylated	unmethylated	methylated

SUPPLEMENTAL RESULTS

Supplemental Table 3 – Statistical significance of differentiation and expression differences. Based on the results shown in Figure 1, the significance of differences between each pair of experimental conditions was evaluated by Bonferroni's post test for differentiation (A) and myogenin expression (B). Non-significant differences are indicated as n.s. EmM = Embryonic muscle, EmB = Embryonic brain.

A	C2C12 GM	C2C12 DM	C2C12 DM+DH	C2T18 GM
C2T18 DM	p<0.001	p<0.001	p<0.001	p<0.001
C2T18 GM	p<0.01	n.s.	n.s.	
C2C12 DM+DH	p<0.01	p<0.05		
C2C12 DM	p<0.01			

B	C2C12 GM48h	C2C12 GM96h	C2C12 DM	C2C12 DM+DH	C2T18 GM48h	C2T18 GM96h	C2T18 DM	EmB
EmM	p<0.001	p<0.001	p<0.001	p<0.01	p<0.001	p<0.001	n.s.	p<0.001
EmB	n.s.	p<0.001	p<0.001	p<0.001	p<0.001	p<0.001	p<0.001	
C2T18 DM	p<0.001	p<0.001	p<0.001	p<0.01	p<0.001	p<0.001		
C2T18 GM96h	p<0.001	p<0.01	n.s.	p<0.01	p<0.01			
C2T18 GM48h	p<0.001	p<0.01	p<0.01	p<0.001				
C2C12 DM+DH	p<0.001	p<0.01	p<0.01					
C2C12 DM	p<0.001	p<0.05						
C2C12 GM96h	p<0.001							



Supplemental Figure 1 – Analysis of CpG and non-CpG methylation patterns by restriction enzymes.

A) and B) PCR amplification of the 5'-flanking and exon 3 myogenin regions after cutting with the methylation-insensitive nuclease BglI (A, lanes 2-5) and the methylation-sensitive nuclease PvuII (A, lanes 6-9), the methylation-insensitive nucleases AvaI (B, lanes 2-3) and EcoNI (B, lanes 4-5). C) and D) PCR amplification of the 5'-flanking region, exon 1 and exon 3 of myogenin after cutting with the EcoRI nuclease (with no recognition sites within the analyzed regions), the methylation sensitive HpaII and MspI nucleases, in embryonic muscle and embryonic brain (C), as well as in C2C12 GM 1h, C2C12 DM 2 h and C2C12 DM 24 h (D). No endonuclease used has recognition sites within exon 3. EcoRI has no recognition sites within 5'-flanking and exon 1 and was used, together with uncut DNA, as a positive control of amplification. MM: = DNA molecular weight marker, b = blank, C = uncut control DNA. Only the experiments with completely digested parallel control samples (data not shown) made of PCR amplified products, without methylation and from the same myogenin zones as the experimental samples, were taken into account.

Panels A and B. After digestion with the methylation-insensitive BglI, AvaI and EcoNI nucleases, not inhibited by cytosine methylation (see supplemental Tab. 2S for restriction endonuclease features), no amplified PCR product relative to the 5'-flanking region was present; this assay represents a negative control and also shows that the 5'-flanking region is fully accessible to endonucleases and that an amplified product is always obtainable from the exon 3 region (with no recognition sites for any of the endonucleases used). It is noteworthy that the AvaI and HpaII recognition sites coincide, though the former is not inhibited by DNA methylation. Moreover, the use of the methylation-sensitive endonuclease PvuII not only revealed the methylated experimental conditions (C2C12 GM 48 h and 96 h, C2C12 DM 2 h, those in which an amplified product from the 5'-flanking region was obtained), but also showed the presence of non-CpG methylation.

Panels C and D. Amplified products from the 5'-flanking region and exon 1, after HpaII or MspI treatment, revealed methylated experimental conditions (embryonic brain, C2C12 GM 1h and C2C12 DM 2 h); amplified products were also obtained after treatment by MspI nucleases, again confirming the presence of non-CpG methylation in the CCGG site of 5'-flanking as well as in all three CCGG sites of exon 1. By contrast, the 5'-flanking region was found to be unmethylated in other experimental conditions (embryonic muscle and C2C12DM 24h), as the fact that no amplicon was detected upon treatment with either HpaII or MspI shows. The presence of an amplicon relative to the exon 1 region after HpaII treatment, though not after MspI treatment, revealed the presence of CpG methylation even in the three CCGG sites of exon 1.

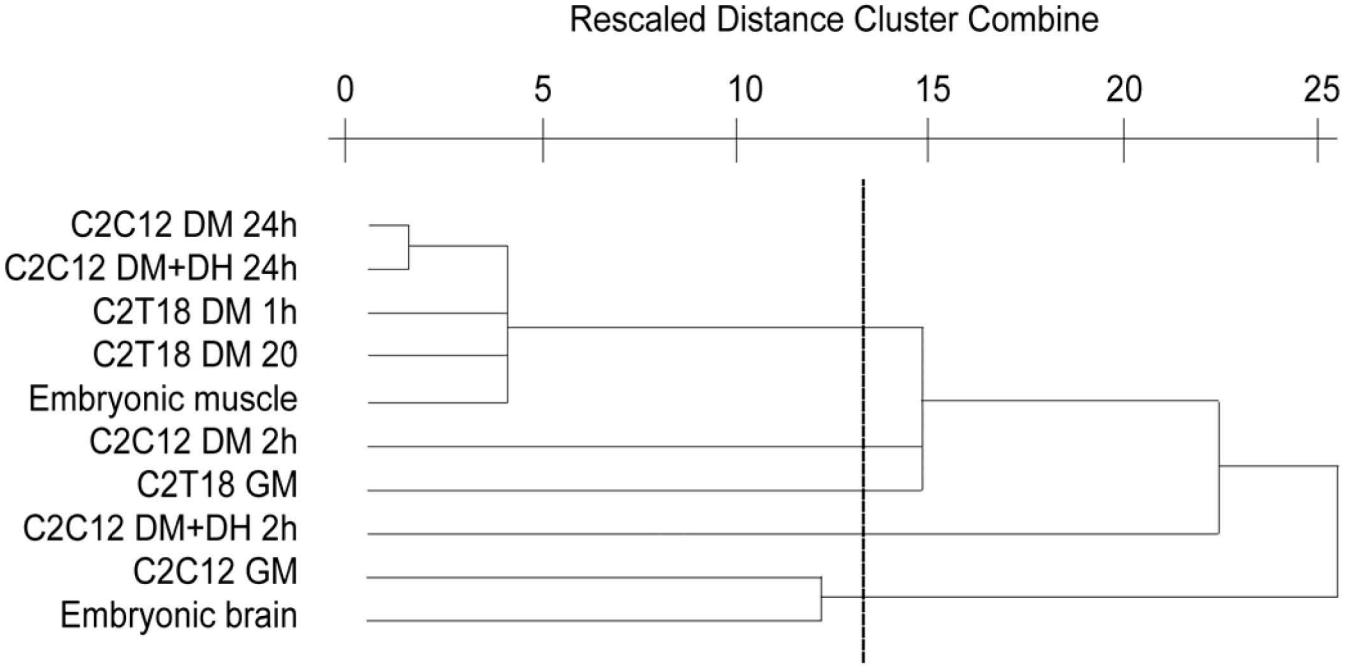
Supplemental Table 4 – Association between methylation and experimental conditions (temporal variations) and cytosine position (spatial variation) in C2C12 cell line, C2T18 clone and embryonic tissues. The mean square contingency Φ^2 is used as an association index not dependent on the number of observations. Kendall's τ-c is reported as a non-parametric measure of association; sign and absolute values indicate, respectively, direction and intensity of association. Its value varies within the interval [-1, 1]. A) Association between methylation levels and experimental conditions; B) association between methylation levels and cytosine position.

	A) Methylation - experimental condition			B) Methylation - cytosine position		
	C2C12	C2T18	embryonic tissues	C2C12	C2T18	embryonic tissues
Φ^2	0.186	0.160	0.398	0.065	0.069	0.041
Kendall's τ -c	-0.292	-0.050	-0.619	-0.089	-0.040	-0.106

A - The mean square contingency Φ^2 indicates the existence of a connection between the methylation state and the experimental condition: when the experimental condition varies, the methylation state also varies. Indeed, the mean square contingency is zero if the two characters are independent, but is positive if there is dependency. Its absolute value is directly proportional to dependency's strength. Kendall's τ-c specifies the dependency type more accurately, using the additional information concerning the distribution order of the dependent variables: the aforesaid relationship is negative, which means that during the progression from the first to the last experimental condition the variable "methylation state" decreases (the experimental conditions have been arranged in order of increasing demethylation).

B - The same connection indexes were used to evaluate the connection between the methylation state and the cytosine position. According to the values checked, this connection is weaker or almost absent. The relationship, however, remains negative, which means that the variable "methylation state" hardly decreases in the progression from the first cytosine of the sequence to the last.

A +B - The absolute values of both the mean square contingency Φ^2 and Kendall's τ-c are markedly higher in A than in B, indicating that the variation of the methylation pattern is more closely related to the experimental conditions (temporal variations) than to a gradient of demethylation according to position of the cytosine in the 5'-flanking region sequence, from the distal to the proximal (near exon 1).



Supplemental Figure 2 Cluster analysis of experimental conditions.

The cluster analysis shown is based on bisulphite data. Using the cut-off value indicated, a main cluster (of C2C12 at the late experimental times, 24h with or without DH, both C2T18 DM, 1h and 20 min, and EmM) was selected. This is a cluster of experimental conditions that display high terminal differentiation and myogenin expression, most of which is demethylated. Another cluster was composed of C2C12 GM, the least myogenin expressing undifferentiated muscle condition, and EmB, both of which displayed the highest methylation levels. The other three experimental conditions did not group and were separated from the other conditions; these experimental conditions display either high terminal differentiation and myogenin expression (C2C12 DM 2h with or without DH) or high myogenic potential only partially inhibited by a high serum level (C2T18 GM), both with intermediate levels of methylation. It is noteworthy that C2C12 GM and EmB are the first two separating conditions (according to the cluster distance index), followed by the separation of C2C12 DH 2 h. The 2 most non-expressing and undifferentiated conditions (i.e. C2C12 GM and EmB) and the most stimulated experimental condition at the early time (i.e. C2C12 DM+DH 2 h) displayed an extremely specific methylation pattern. Experimental conditions with similar terminal differentiation and myogenin expression levels grouped in clusters if classified according to their methylation status, a result that highlights the specificity of the methylation patterns. This specificity might arise from a spatial variation of cytosine methylation. This spatial pattern may be either an ordered linear pattern along the 5'-flanking region of myogenin, or a more complicated non-ordered pattern with no linear correlation between the methylation levels of each cytosine and its position along the 5'-flanking region of the gene. The analyses shown in Figure 2 and Table 4S suggest that the latter hypothesis is more likely to be correct.

Supplemental Table 5 - Characteristics of transcription factors with putative recognition sites that overlap the short CpC-rich peculiar elements.

TRANSCRIPTION FACTOR	INFLUENCE OF CpG METHYLATION	INFLUENCE OF non-CpG METHYLATION
SP1. Widely expressed Zn-finger DNA binding factor involved in transcriptional modulation during development and differentiation. It targets multiple biological functions. **Entrez Gene link for SP1**	It was first demonstrated that the methylation of the CpG dinucleotide within the Sp1 recognition sequence (GGGCGG) does not interfere with the DNA binding of this factor (***Harrington et al.***, *1988, Proc. Natl. Acad. Sci. U. S. A, 85, 2066-2070;* ***Holler et al.***, *1988, Genes Dev., 2, 1127-1135*). However, several studies subsequently demonstrated the influence of CpG methylation of 5'-flanking of several genes on Sp1 binding and, consequently, the effect on their transcription: a) in the p21 promoter of human lung cancer cell line H719 (***Zhu et al.***, *2003, Mol. Cell Biol., 23, 4056-4065*); b) in the ABCC6 transporter in several mouse tissues (***Douet et al.***, *2007, Biochem. Biophys. Res. Commun., 354, 66-71*); c) in the cyclin D1 gene in rat leukemic cell lines (***Kitazawa et al.***, *1999, J. Biol. Chem., 274, 28787-28793*); d) in the p57 when muscle differentiation and concomitant rapid DNA hypomethylation are induced (***Figliola et al.***, *2008, J. Mol. Biol., 380, 265-277*).	The methylation of the outer C of the lower strand (CCGCCC) of the Sp1 binding site of the CpG island promoter in some retinoblastoma tumors inhibits Sp1 binding; in particular, the methylation of both cytosines ($^{m}C^{m}CG$) has been shown to inhibit binding by 95% (***Clark et al.***, *1997, Gene, 195, 67-71*). The authors suggested that this kind of methylation, by blocking Sp1 binding and overcoming its protective effect against de novo methylation, may trigger hypermethylation of the CpG island, which in turn leads to promoter silencing. Non-CpG methylation within the two continuous Sp1-binding sites of 5'-flanking of cyclin D1 gene in normal rat uterine stromal cells of the basal layer has been reported (***Kitazawa et al.***, *1999, J. Biol. Chem., 274, 28787-28793*).
CTCF. Conserved Zn-finger DNA binding factor with recognition sites overlapping those of Sp1-like proteins. It binds to evolutionarily divergent target sequences by using different combinations of its Zn-fingers. CTCF binding sites usually contain 3 regularly spaced CCCTC motifs. The multiple sequence specificity corresponds to multiple functions in gene regulation. CTCF has been shown to act as both a classical transcription factor and as a chromatin insulator, according to the current idea of CTCF as an active member of the eukaryotic gene regulation and not just as an inert barrier. **Entrez Gene link for CTCF**	In vitro binding of CTCF is reduced by CpG methylation (***Chao et al.***, *2002, Science, 295, 345-347;* ***Filippova.***, *2008, Curr. Top. Dev. Biol., 80, 337-360*)	In vitro binding of CTCF is abolished by non-CpG methylation (***Chao et al.***, *2002, Science, 295, 345-347*).
	The binding sites for CTCF map to DNA methylation-free domains throughout the genome (***Mukhopadhyay*** *et al., 2004, Genome Res., 14, 1594-1602*). When CTCF is bound to methylation-free targets, it has been shown to prevent the spreading of methylation and play a critical role in the maintenance of methylation-free regions, even affording protection against de novo methylation (***Filippova****, 2008, Curr. Top. Dev. Biol., 80, 337-360*).	