# Tissue specific differentially methylated regions (TDMR): Changes in DNA methylation during development

**Fei Song**[a], **Saleh Mahmood**[a,e], **Srimoyee Ghosh**[b], **Ping Liang**[b], **Domminic J. Smiraglia**[b], **Hiroki Nagase**[b,c,d], and **William A. Held**[a]

a *Department of Molecular and Cellular Biology, Roswell Park Cancer Institute, Buffalo, NY 14263, USA*

b *Department of Cancer Genetics, Roswell Park Cancer Institute, Buffalo, NY 14263, USA*

c *Life Science, Advanced Research Institute for Sciences and Humanities, Nihon University School of Medicine, Tokyo 173-8610, Japan*

d *Division of Cancer Genetics, Department of Advanced Medical Science, Nihon University School of Medicine, Tokyo 173-8610, Japan*

## Abstract

Tissue specific differentially methylated regions (TDMRs) were identified and localized in the mouse genome using second generation virtual RLGS (vRLGS). Sequenom MassARRAY quantitative methylation analysis was used to confirm and determine the fine structure of tissue specific differences in DNA methylation. TDMRs have a broad distribution of locations to intragenic and intergenic regions including both CpG islands, and non-CpG islands regions. Somewhat surprising, there is a strong bias for TDMR location in non-promoter intragenic regions. Although some TDMRs are within or close to repeat sequences, overall they are less frequently associated with repetitive elements than expected from a random distribution. Many TDMRs are methylated at early developmental stages, but unmethylated later, suggesting active or passive demethylation, or expansions of populations of cells with unmethylated TDMRs. This is notable during postnatal testis differentiation where many testis-specific TDMRs become progressively "demethylated". These results suggest that methylation changes during development are dynamic, involve demethylation and methylation, and may occur at late stages of embryonic development or even postnatally.

### Keywords

Correspondence to: William A. Held.

[e]Current Address: IMMCO Diagnostics, 60 Pineview Dr. Buffalo, NY 14228 USA

Song, Fei Fei.Song@RoswellPark.org 716-845-2367, Fax 716-845-5908

Mahmood, Saleh smahmood@immco.com 716-691-0091 ext134, Fax 716-691-0466

Ghosh, Srimoyee Srimoyee.Ghosh@RoswellPark.org 716-845-1737, Fax 716-845-1698

Liang, Ping Ping.Liang@RoswellPark.org 716-845-1556, Fax 716-845-1698

Smiraglia, Domminic J. Dominic.Smiraglia@RoswellPark.org 716-845-1347, Fax 716-845-1698

Nagase, Hiroki Hiroki.Nagase@RoswellPark.org 716-845-1546, Fax 716-845-1698

Held, William A. William.Held@RoswellPark.org 716-845-3301, Fax 716-845-5908

## Introduction

The concept of differentially methylated regions in the mammalian genome has changed markedly in recent years. CpG islands were previously thought to be almost entirely unmethylated except within imprinted regions and on the inactive X chromosome [1]. Early studies [2,3] using Restriction Landmark Genomic Scanning (RLGS) analysis of DNA from different tissues identified tissue specific differentially methylated regions (TDMR). However, the sequence and location of TDMRs in the genome were largely unknown. The advent of the mouse and human genome project and virtual RLGS [4-6] has made it possible to more rapidly identify the DNA sequence of RLGS fragments and the locations of TDMRs [7-10]. Comparison of TDMRs in human and mouse has established that the DNA sequence of many TDMR regions are conserved and that the methylation profile in mouse and human is conserved as well [9,11,12].

TDMRs have also been identified using DNA methylation array methods. Studies using antibody to 5-methyl cytosine to capture methylated DNA found strong CpG island promoters (high CpG density) were mostly unmethylated, weak CpG island promoters (low CpG density) were preferential targets for *de novo* methylation, and promoters of most germline-specific genes were methylated in somatic tissues [13,14]. Issa and colleagues, using methylated CpG island amplification in combination with microarrays, found that among more than 5000 autosomal genes with dense CpG island promoters, approximately 4% were methylated in normal peripheral blood [15]. Schilling and Rehli performed global methylation analysis of human testis, brain, and monocytes and found a significant association between tissue-specific promoter methylation and gene expression [16]. Tissue-specific CpG island methylation at developmental gene loci were identified using a CpG island array [17]. In addition, high throughput bisulfite DNA sequencing of regions of human chromosomes 6, 20, and 22 found that 17% of 873 analyzed genes are differentially methylated in the 5′ promoter region and that in about one-third, methylation is inversely correlated with transcription [11].

Thus, the recent analysis of genome-wide methylation patterns, by a variety of methods, indicate that there are more extensive differences in DNA methylation between differentiated tissues than previously thought. For the most part, the CpG density and fine structure of TDMRs have not been well defined nor is it clear how these differences are established during development. We previously identified and confirmed a number of TDMRs that were identified by RLGS and virtual RLGS technology [5,8,12]. In this study, we have identified and confirmed additional TDMRs using second-generation virtual RLGS software (vRLGS; [6] along with Sequenom MassARRAY quantitative methylation analysis [18]. We also present initial studies that characterize the fine structure of some of the TDMRs as well as tissue and developmental stage-specific methylation differences in TDMRs.

## Results

### Identification of TDMRs using second generation virtual RLGS (vRLGS)

In addition to DNA from adult male C57BL6/J mouse tissues analyzed previously (liver, brain, kidney, muscle, colon, and testis), DNA from C57BL6/J ES cells (Stewart Bruce 4, passage 13 and 17) were analyzed by RLGS using both the *Not*I-*Pst*I-*Pvu*II and *Not*I-*Pvu*II-*Pst*I enzyme combinations and vRLGS (see Supplementary Figure 1 and [6,8]. Figure 1 shows a *Not*I-*Pst*I-*Pvu*II RLGS profile of ES cells indicating that several Pvu-TDMR loci were methylated in ES cells that were unmethylated in Testis. ES cells from two different passages (13 and 17) and testis DNA from two different mice were analyzed as shown. The absence of RLGS spots indicates methylation since if the *Not*I site is methylated it will not be digested with *Not*I or end-labeled. Second generation virtual RLGS was able to identify the genomic location of 68 of the 150 TDMRs [8]. Tissue specific methylation was confirmed at 34 loci, mostly using

Sequenom MassARRAY quantitative methylation analysis [18] (4 were confirmed by other methods; See Figure 2 and Supplementary Figures 2 and 3). As shown in Table 1, the TDMR locations are distributed throughout the genome. Among the confirmed TDMRs, 32% are in CpG islands and 12% are in promoter regions. A very high proportion (68%) of the confirmed TDMRs were located in non-promoter intragenic regions, which is significantly higher than expected from a random distribution (p<0.00005; see legend for Supplementary Table 1). Array based methylation analysis as well as bisulfite sequencing methods generally exclude repetitive regions. As shown here, 21% of the confirmed TDMRs (*Not*I site) are within repetitive sequences, mostly LTRs. Table 2 lists the confirmed TDMRs, the RLGS based tissue methylation profile, the UCSC genomic position (RLGS fragment), CpG island location, their genomic position relative to the nearest gene, nearby gene, gene function, and Human homology (500 bp region centered on mouse *Not*I site). A complete listing of 68 TDMRs identified by vRLGS is provided in the Supplementary Table 1. Analysis of gene ontology indicates that developmentally related genes are significantly over-represented using the list of Virtual RLGS TDMR loci or the confirmed list of TDMR loci (p<0.05). Zinc binding proteins, including Zinc Finger proteins are also over-represented (p<0.01).

## Confirmation of TDMR methylation using Sequenom MassARRAY quantitative methylation analysis

The TDMR methylation status in different tissues was initially inferred from RLGS analysis. A diploid spot intensity indicating the *Not*I site was unmethylated was given a value of 2, and the complete absence of the spot, indicating complete methylation, a value of 0. Partially methylated spots would have a value between 0 and 2 (see Table 2 and Supplementary Table 1). Sequenom MassARRAY quantitative methylation analysis [18] was performed to confirm RLGS analysis and to provide some information on the CpG density and extent of the differentially methylated region. Thirty-four out of 68 loci were confirmed (Table 2,Supplementary Table 3), 30 by Sequenom MassARRAY quantitative methylation analysis and 4 by other methods (MSP or regular bisulfite sequencing). The Sequenom MassARRAY methylation pattern was not consistent with the RLGS analysis for 4 loci suggesting the vRLGS identification was incorrect. Technical problems prevented the confirmation of the remaining loci, primarily due to the very high CpG density. Several examples of Sequenom MassARRAY quantitative methylation analysis of confirmed loci are presented in Figure 2. Consistent with RLGS, the Sequenom MassARRAY quantitative methylation analysis indicated Pst6 and Pst4 are completely methylated in all tissues except testis. Pst4 is within a 5′ CpG island promoter region for Spesp1 (Sperm equatorial segment protein 1) that is expressed at high levels in testis [19] (http://symatlas.gnf.org/SymAtlas/). Pst6 is located in a 3′ exon CpG island for Hspa1l (Heat shock 70KDa protein 1-like) that is expressed virtually exclusively in testis (http://symatlas.gnf.org/SymAtlas/) even though the promoter region is not differentially methylated (see Figure 3). Pvu6, located in an intergenic region, is unmethylated in testis and partially unmethylated in muscle, but completely methylated in the other tissues. Pvu35 is located in an intron of Casz1 (Castor homolog 1, a zinc finger gene). RLGS analysis indicated Pvu35 methylation in most tissues but partial methylation in kidney (Table 2). Partial methylation is consistent with heterogeneity of methylation within a cell type or heterogeneity of cell types within a tissue and methylation differences between different cell types. Additional RLGS and Sequenom MassARRAY data are presented for Pvu74 and Pvu75, RLGS loci that are partially unmethylated in a number of tissues but almost completely methylated in ES cells (Supplementary Figures 2 and 3). Somewhat surprisingly, based on RLGS analysis, we found that 58% of the TDMR loci that are unmethylated in one or more adult tissues are methylated in ES cells (see Table 2).

### TDMRs associated with repetitive sequences

Methylation analysis using microarrays necessarily excludes repetitive sequences so it is not possible to directly obtain information concerning the differential methylation of these sequences. Due to sequence divergence within repetitive sequences and/or the presence of a portion of unique sequence within the RLGS fragment, it is possible to resolve many repetitive sequences that contain a *Not*I site using RLGS and to obtain information on differential methylation. Among the 34 confirmed TDMRs (Table 2), 7 contain methylated *Not*I sites within repetitive elements, six corresponding to LTR repetitive elements, and one corresponding to a (CCG)n simple repeat (Pvu24). All of these sequences are unmethylated in testis, but Pvu53 is partially unmethylated in colon, liver, and kidney as well (see Table 2 and Supplementary Figure 4). These results indicate that some TDMRs are within repetitive sequences although at a significantly lower than expected frequency based on a Monte Carlo simulation (simulated mean=16; z-score=-2.9; p<0.0001). Excluding the TDMRs that are within repetitive sequences, the average observed distance from the *Not*I site of the TDMR to a repeat sequence is 380 bp. This is not significantly different than the predicted value of 381 bp based on a random distribution of *Not*I sites (z-score: 1.18; p=0.120). Taken together, our data indicate that some TDMRs do occur within repetitive elements, but at much lower frequency than expected by a random distribution.

### TDMR methylation fine structure

In order to understand the functional relevance of TDMRs, it is essential that the methylation fine structure be determined. Both Pst3 and Pst4 are located within CpG island promoter regions for genes (Ddx4 and Spesp1, respectively) that are highly expressed in testis (http://symatlas.gnf.org/SymAtlas/). The region of the Pst3 TDMR that is unmethylated in testis is confined to the 5′ promoter CpG island and the 5′ flanking region of Ddx4 (Supplementary Figure 5; see also[8]). The downstream regions (intron and 3′ exon) are similarly methylated in all tissues including testis. The promoter regions of several genes upstream or downstream from Ddx4 that show high, but not exclusive testis expression are not differentially methylated (Supplementary Figure 6). In contrast to Pst3, the Pst4 TDMR region that is unmethylated in testis includes the 5′ promoter CpG island of Spesp1 and the 3′ exon more than 8 kb downstream (Supplementary Figure 7). Analysis of 2 intron regions, located approximately midway between exon 1 and 2, and 2 CpG rich regions about 2-3 kb bp upstream of the Spesp1 promoter were largely methylated in all tissues including testis (data not shown). The Pst6 TDMR is also within a gene that is almost exclusively expressed in testis, Hspa1l. The differentially methylated region is restricted to one of two small CpG islands in the 3′ exon, extending somewhat 5′ in the exon (Figure 3). These results indicate that the intragenic locations of the TDMRs may vary considerably, from the entire exonic and CpG island promoter region (Pst4), to a relatively small 5′ promoter CpG island region (Pst3), or 3′ region only (Pst6).

Several regions close to the Pst6 TDMR were analyzed for methylation using Sequenom MassARRAY quantitative methylation analysis. The average CpG methylation for each region is shown in Figure 3. Data is shown for testis, ES cells, and liver, which is representative of somatic tissue (similar results were obtained for muscle, brain, kidney, and colon; data not shown; see Figure 2). Although Hspa1l is expressed exclusively in testis, the promoter CpG island region is completely unmethylated in somatic and ES cells as well as testis tissues. Thus, in contrast to the testis specific expression of Ddx4 (Pst3) and Spesp1 (Pst4), promoter methylation status does not appear to be associated with the testis specific gene expression of Hspa1l and repression in somatic tissue. The testis specific gene, Hspa1l is homologous to the adjacent Hspa1a gene that is transcribed in the opposite direction. The region of Hspa1a that is homologous to the Pst6 TDMR is also differentially methylated, even though Hspa1a is not expressed to any appreciable extent in testis (http://symatlas.gnf.org/SymAtlas/).

## Changes in TDMR methylation during development

As noted previously, we were somewhat surprised that RLGS indicated almost 60% of the TDMRs were methylated in ES cells, which suggests that these TDMRs become demethylated during differentiation to adult tissues. To investigate this further, we isolated DNA from 10d embryo (head or body), 15d embryo, and neonatal tissues (brain, liver, kidney, intestine, and testis), and performed Methylation Specific PCR (MSP) to determine TDMR methylation status at different developmental stages (Figure 4). Pst3 and Pvu8, which are unmethylated in adult testis, are largely methylated in 15d embryo, including testis, but are partially unmethylated in neonatal testis. Pst21, which is unmethylated in adult liver, is fully methylated in 15d embryo including liver, and is partially unmethylated in neonatal liver. Pst46, which is unmethylated in adult kidney, muscle, and testis, is partially unmethylated in a number of embryonic tissues but appears to be mostly methylated in E15 and neonatal kidney and testis. These results suggest that some TDMRs may become demethylated during development. Somewhat surprisingly, some TDMRs are still largely methylated in neonatal tissues, suggesting that demethylation may occur late in development. However, we cannot exclude the possibility that the methylation differences reflect expansion of specific cell populations during development.

In mice, gametogenesis in testis is initiated shortly after birth, with the first wave occurring in a synchronous fashion [20,21]. Testis consists of sertoli cells and gonocytes at birth. Meiotic prophase begins around day 9 and meiosis is almost complete by day 20. By day 20, the gonocytes have differentiated into secondary spermatocytes and spermatids. Since a number of TDMRs that were unmethylated in adult testis showed almost complete or partial methylation in neonatal testis, we examined methylation of a number of TDMRs in neonatal, 10d, 20d, and adult testis using Sequenom MassARRAY quantitative methylation analysis. We examined TDMRs that are methylated in most somatic tissues but are unmethylated in adult testis (Figure 5). These included TDMRs that are located within repeat sequences (Pst2, Pvu1) close to a repeat sequence (Pst5) and a number of unique sequence TDMRs (Pst3, Pst4, Pst6, Pvu4, Pvu8). All of these TDMRs had a similar pattern of progression, from almost fully methylated in neonatal testis to almost completely unmethylated in adult testis. This suggests a common mechanism for demethylation during testis development of TDMRs associated with repeats and those associated with unique sequences. In contrast Pvu80, which is partially methylated in most somatic tissues, was unmethylted at all stages of testis development (Table 2, Supplementary Figure 8).

Analysis of somatic tissue (Supplementary Figure 8) indicated that Pvu80 is relatively unmethylated at early developmental stages with a progressive increase in methylation at later developmental stages. This suggests that considerable Pvu80 methylation in somatic tissues occurs at later developmental stages, even postnatally. Alternatively, the increased methylation may be the consequence of the expansion of specific cell populations during the later stages of development.

## Tissue and developmental stage specific differences in methylation

Pvu8 and Pst10 both identify a TDMR (*Not*I site) within the 3′ exon of Zfp206, a C2H2 type Zinc Finger protein involved in transcriptional regulation (Figure 6). In adult tissues the region downstream (Pvu8) and upstream (Pst10) from the *Not*I site is unmethylated in testis, partially unmethlyted in brain and methylated in other tissues (muscle, kidney, liver, and colon; see Figure 6 and data not shown). Sequenom MassARRAY quantitative methylation analysis of the TDMR, the promoter, and exon 2 regions indicates there are tissue and developmental stage specific differences in methylation in these regions. The promoter region is unmethylated in ES cells, partially methylated in e15 brain and testis, and completely methylated in adult brain and testis. Exon 2 is mostly unmethylated in ES cells and adult testis, but almost completely

methylated in E15 brain and testis, and adult brain. In contrast, the TDMR region is mostly unmethylated only in testis. These results suggest that there are dynamic changes in methylation status in these regions during development.

## Discussion

### Identification, confirmation, and locations of TDMRs

RLGS was previously used to identify 150 tissue specific differentially methylated regions (TDMRs) in the mouse genome [8]. Based on this observation and the number of *Not*I sites in the genome, we previously projected that 5% or more of the CpG islands are TDMRs. Due to the limited number of tissues examined in these earlier studies, we indicated that this is likely to be an underestimate. Recent analysis of CpG island methylation using an improved set of CpG island clones found that 6-8% of the CpG islands were methylated in the tissues examined [17]. This value is also likely to be an underestimate of the total number of CpG islands methylated in the genome due to limitations in the number of tissues surveyed and heterogeneity of cell types within a tissue.

Second-generation vRLGS [6] was used to determine the DNA sequences and locations of 68 TDMRs in the mouse genome, of which, tissue specific methylation was confirmed for 34 loci using primarily Sequenom MassARRAY quantitative methylation analysis. In spite of the strong bias of RLGS toward CpG island genomic regions, TDMRs were found to be distributed throughout the genome including CpG islands, promoter regions, exons, introns, and intergenic regions (Table 1). Somewhat surprisingly, we found a significantly higher fraction of TDMRs in non-promoter intragenic regions than expected from a random distribution (P<0.00001; see Supplementary Table 1). Analysis of tissue specific methylation using a CpG island array also found methylated CpG islands located disproportionately remote from TSS of the associated gene [17]. Some TDMRs were found to be within repetitive sequences, mostly LTRs. Nevertheless, our data indicates that they are much less frequently associated with repetitive elements than expected from a random distribution. A recent report, however, indicates that a combination of repeat structure, DNA sequence and unusual predicted DNA structure can be correlated with CpG island methylation [22]. Most TDMRs examined in this study (28/34) have homology to the corresponding region in the human genome (Table 2). Gene ontology analysis of nearby genes indicates a higher than expected frequency of developmentally associated genes and those that encode Zn binding proteins. This is consistent with a recent report that genes associated with tissue specific CpG island methylation were developmental gene loci [17]. In addition, our results indicate dynamic changes in methylation at these loci during development.

### TDMRs and gene regulation

Methylation of CpG island promoter regions is associated with gene silencing [23,24] whereas methylation of insulator regions may be associated with up regulation of gene expression [25,26]. Our studies identified 4 TDMRs associated with gene promoter regions (Table 2), and in each, methylation is inversely associated with gene expression ([8], Supplementary Figures 3 and 7). In addition, many genes known as cancer testis antigens that are highly expressed in testis have CpG island promoter regions that are unmethylated in testis but are methylated in somatic tissues that do not express them [27]. However, examination of the fine structure of DNA methylation of three TDMRs associated with genes that have exclusive or high level expression in the testis indicates that each has a somewhat different methylation distribution relative to the associated gene. Pst3 and Pst4 are located in CpG island promoter regions for Ddx4 and Spesp1 respectively (Supplementary Figure 5 and Supplementary Figure 7). Somatic methylation of Ddx4 appears to be restricted to the CpG island promoter region. Promoter regions of other genes with high testis expression upstream and downstream from Ddx4 are

not differentially methylated (Supplementary Figure 6). In contrast, somatic methylation of Spesp1 includes the promoter CpG island region and the 3′ exon, approximately 8 kb downstream of the 5′ promoter CpG island. However, the intron region is not differentially methylated. Thus, for TDMRs located in the promoter regions of testis specific genes, there is a strong inverse correlation between methylation and gene expression but the methylated regions may be different. Pst6 is within one of two weak CpG islands located in the 3′ exon of Hspa1l, a gene exclusively expressed in testis (Figure 3; http://symatlas.gnf.org/SymAtlas/). A homologous gene, *Hspa1a* (~80% homology), is located just upstream but transcribed in the opposite direction. The CpG rich promoter of *Hspa1l* and the proximal portion of the CpG island associated with the *Hspa1a* gene are not differentially methylated. However, the region of *Hspa1a* CpG island that has homology to region of Pst6, is differentially methylated even though *Hspa1a* is not expressed in the testis. Interestingly, these regions contain apparent binding sites for GCNF (germ cell nuclear factor, http://genome.ucsc.edu/), an orphan receptor that is present in germ cells and has been reported to directly interact with Dnmt 3a and 3b [28], as well as to recruit binding of MBD2 and MBD3 to the Oct4 promoter [29]. However, it is unclear whether and how GCNF and the Pst6 TDMR relate to the testis tissue-specific expression of Hspa1l. These results suggest that the property that confers differential methylation may be related to primary DNA sequence and may be conserved, but that additional factor(s) may be required to impart testis specific gene expression of Hspa1l.

Other TDMRs located within or near gene promoter regions (Pst61, Pvu74, Pvu75) are unmethylated in most somatic tissues as well as testis, but methylated in a single tissue (of those tested). Pst61 is located in an alternative promoter region for Gata2, is methylated in liver, and has very low expression in liver [8]. Pvu74 appears to be in an alternative promoter region of Cadherin 22 (Cdh22), is methylated in ES cells, and expressed at low levels in ES cells (Supplementary Figure 2). Pvu75 is located in the promoter region of Cacna1e, a voltage sensitive calcium channel protein, is methylated in ES cells, and is expressed at low level in ES cells (Supplementary Figure 3). Thus for these TDMRs, promoter methylation is associated with low gene expression. However, the TDMRs are unmethylated in some tissues that also have low gene expression. Thus, it is not clear whether methylation of these TDMRs has a primary role in gene regulation. It should also be emphasized that tissues are composed of multiple cell types and that RLGS and Sequenom MASSarray methylation analysis present an average methylation for all cell types. High expression of a gene expressed in cell type that makes up a small fraction of the total tissue would obscure any association between methylation and gene expression. These issues can only be resolved by using purified cell populations.

## TDMR locations and methylation fine structure

The locations of TDMRs appear to vary considerably. For example, for the testis specific TDMRs, the differentially methylated region for Pst3 is confined to the CpG island promoter region, Pst4 includes the CpG island promoter region as well as the 3′ exon, and Pst6 is confined to a region close to one of two weak 3′ exon CpG islands. The CpG islands associated with the Rhox gene cluster of 12 related homeobox genes on the X chromosome are differentially methylated in a stage and lineage specific manner indicating long range gene silencing for an entire cluster of genes [30]. In colon cancer, coordinate epigenetic silencing may occur across an entire chromosome band [31]. In this study, we have not observed methylation that extends through several genes.

## "Demethylation" during gameotogenesis in testis

Our results indicate that several TDMRs that are unmethylated in adult testis are almost completely methylated in neonatal testis. These include TDMRs associated with or very close to repeat sequences (Pvu1, Pst2, Pst5), and those associated with unique sequences in gene promoter regions (Pst3, Pst4) and in 3′ exons (Pvu4, Pvu8, Pst6). At birth, testis consists

primarily of undifferentiated gonocytes and somatic sertoli cells. The first cycle of gametogenesis occurs synchronously with differentiation to undifferentiated type A spermatogonia around day 6 and initiation of the first meiotic prophase around day 9. The major cell types at day 11 are leptotene and zygotene spermatocytes that reach late pachytene by day 18 [21]. Our results indicate that the TDMRs are almost completely methylated at birth in the testis, partially demethylated at day 10 around the start of meiosis, and almost completely demethylated by the end of meiosis at day 20. A recent study, using purified spermatogenic cells, found both *de novo* methylation and demethylation in spermatogonia and spermatocytes in early meioic prophase that was largely completed by end of pachytene spermatocyte [32]. Their results indicated that the methylated regions were all associated with unique sequences whereas the unmethylated regions were all associated with LTR repeats. Our results indicate a similar timing of demethylation, but that demethylation occurs in TDMRs associated with both unique sequences and those associated with LTR repeats. The similarity in kinetics of demethylation of unique sequence and repeat TDMRs suggests a common mechanism (Figure 5).

### Changes in TDMR methylation during development

One of the surprising findings of this study, is that many of the TDMRs are methylated at early developmental stages but are unmethylated in adult tissues suggesting active or passive demethylation during development. Almost 60% of the TDMRs are methylated in ES cells (Table 2), which are derived from the embryo inner cell mass at the blastocyst stage, after the period of demethylation following fertilization, but before the period of *de novo* methylation that occurs at or after gastrulation and implantation. Other investigators also found differences in methylation status between ES cells and differentiated cells using RLGS [33]. More recent genome-wide studies also concluded that promoter DNA methylation contributes to ES cell gene regulation [34,35]. However, since ES cells are grown in tissue culture, it is not clear whether they are completely "normal", especially with respect to DNA methylation.

MSP analysis of several TDMRs at different embryonic developmental stages also indicates that they are methylated at early developmental stages (E10 and E15) and are still partially methylated in neonatal mice (Figure 4), suggesting that demethylation occurs relatively late in development. Sequenom MassARRAY quantitative methylation analysis of testis specific TDMRs (Figure 5) indicates that they are almost fully methylated at birth and that demethylation occurs postnatally during the first synchronous wave of germinal differentiation. Sequenom MassARRAY quantitative methylation analysis of Pst10/Pvu8 TDMR indicates more methylation in E15 brain and testis than in the adult tissues (Figure 6). These results indicate that there is a progressive demethylation of some TDMRs during later stages of development or that there is a proliferation of cells that are unmethylated at these sites. Currently, we cannot distinguish these possibilities. In contrast, methylation at the Pvu80 TDMR indicates that substantial methylation may occur after birth in some tissues.

A very recent report provides additional information and resources regarding genome-wide tissue-specific DNA methylation analysis. Methylation profiles of DNA (mPod) for human tissue-specific differentially methylated regions utilizes MeDIP (Methylated DNA Immunoprecipitation) to analyze genome-wide methylation of 16 different human tissues, including sperm [36]. The results of this study [36] indicate that tissue specific DNA methylation, including CpG islands, is relatively common and is consistent with our current and previous studies [8]. Also consistent with out results, many TDMRs (27%) were found to be testis specific. The human studies [36] revealed a small but significant negative correlation between promoter methylation and gene expression across a range of CpG densities. Interestingly, a small but significant positive correlation between gene body methylation and gene expression was found although the basis for this is currently unclear. Unmethylated

regions, had a clear association with active chromatin signatures, but methylated promoters did not have clear associations with repressive histone modifications[36]. The data can be accessed through the Ensembl genome browser [36].

In another recent report, DNA methylation maps of mouse pluripotent and differentiated cells were generated using representational bisulphate sequencing and single-molecule based sequencing [37]. These studies [37] indicate that DNA methylation patterns can be correlated with histone methylation patterns and along with the results presented here indicate dynamic changes in DNA methylation during development that may involve both methylation and demethylation. In addition, there are many changes in methylation that occur relatively late in development. Additional studies will be necessary to determine the mechanisms involved in these processes and how changes in cell populations relate to changes in tissue specific DNA methylation.

## Materials and Methods

### Growth of ES cells

ES cells were grown by the Roswell Park Gene Targeting and Transgenic Core Resource under standard conditions [38] on irradiated embryonic fibroblasts as a feeder layer. ES cell medium contained DMEM and 10% fetal calf serum. Cultures were incubated at 37° C in humidified air with 5% $CO_2$. ES cells were then cultured on gelatin (0.2%) for 2 days with LIF (1U/ml) in the absence of feeder cells before being collected. Contamination of feeder cells was estimated to be no more than 5%.

### Collection of tissues, DNA and RNA preparations

Tissues from adult and 15d embryo were collected according to an IACUC approved protocol. Embryonic tissues were collected using a dissecting microscope (Leica MZ-125). Tissues and cells were immediately snap-frozen in liquid nitrogen and stored at -80° C until use. The DNA for RLGS was isolated from tissues of 12-week-old C57BL/6J male mice, 15 day embryos and ES cells by using the protocols described [39-41]. DNA for Sequenom methylation analysis was prepared according to Qiagen protocol. TRIzol was used to extract RNA from the same samples [42]. The RNA was quantified by a spectrophotometer and aliquots were checked for integrity by electrophoresis in denaturing agarose gels [43].

### RLGS and vRLGS

RLGS was performed according to published protocols [5,8,39-41] using the enzyme combinations *Not*I-*Pst*I-*Pvu*II and *Not*I-*Pvu*II-*Pst*I. Two independently derived RLGS profiles were analyzed for each tissue. For these studies we used the vRLGS software developed by Smiraglia and co-workers[6]. In short, we aligned the actual autoradiograms of both the enzyme combinations with the vRLGS profiles and identified the spots that matched up very closely between the vRLGS and the real ones (Supplemental Figure 1).

### MSP

MSP was performed as described using bisulfite-treated DNAs [44]. The methylated and unmethylated primers were designed by using METHPRIMER [45]. The primers were chosen to include the *Not*I landmark within the amplification product and also as many CpG dinucleotides in the product as possible. The melting temperatures of the primer pairs were constrained to be between 50°C and 60°C. The MSP primers were synthesized by Integrated DNA Technologies (Coralville, IA). The PCR reactions were carried out for 40 cycles and analyzed on a 2% of agarose gel. The sequences of the primers used are available upon request.

## Sequenom MassARRAY quantitative methylation analysis

Sequenom MassARRAY quantitative methylation analysis [18]was performed using the MassARRAY Compact System (www.sequenom.com). This system utilizes mass spectrometry (MS) for the detection and quantitative analysis of DNA methylation using Homogeneous MassCLEAVE (hMC) base-specific cleavage and matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) MS [18]. DNA (1 ug) was converted with sodium bisulfite using the EZ DNA methylation kit (Zymo Research, Orange, California) according to the manufacturer's instructions. The primers were designed using *Methprimer* [45]. Each reverse primer has a T7-promotor tag for in vitro transcription (5′-cagtaatacgactcactataggggagaaggct-3′) and the forward primer is tagged with a 10mer to balance TM (5′aggaagagag-3′). The primer pairs were designed to span the restriction landmark or closely adjacent region or CG rich region as indicated. Amplification of 1 ul bisulfite treated DNA (~20 ng/ml) was performed using HotStar Taq Polymerase (Qiagen) in a 5 ul reaction volume using PCR primers at a 200 nM final concentration. PCR amplification was performed with the following parameters: 94°C for 15 minute hot start, followed by denaturing at 94°C for 20 seconds, annealing at 56°C for 30 seconds, extension at 72°C for one minute for 45 cycles, and final incubation at 72°C for three minutes. After Shrimp Alkaline Phosphatase treatment, 2 ul of the PCR products were used as a template for in vitro transcription and RNase A Cleavage for the T-reverse reaction as per manufacturer's instructions (Sequenom hMC). The samples were desalted and spotted on a 384-pad SpectroCHIP (Sequenom) using a MassARRAY nanodispenser (Samsung), followed by spectral acquisition on a MassARRAY Analyzer Compact MALDI-TOF MS (Sequenom). The resultant methylation calls were performed by the EpiTyper software v1.0 (Sequenom) to generate quantitative results for each CpG site or an aggregate of multiple CpG sites. A minimum of two independently derived tissue DNAs were analyzed. The average methylation was calculated as mean value of the CpGs methylation value and expressed as percent methylation. The non-applicable reading and its corresponding site were eliminated in calculation. The sequences of the primers used are available upon request.

## Quantitative, Real Time RT-PCR

The iScript cDNA kit from Bio-Rad was used to make cDNAs according to the manufacturer's protocol. One microliter of each tissue cDNA was used per quantitative PCR. PCR primers were designed to bridge the exon–intron boundaries within the gene of interest to exclude possible contamination by genomic DNA (except for Hspa1l). The SYBR Green primers were designed by the web program (http://sourceforge.net/) and purchased from Integrated DNA Technologies. A test RT-PCR was performed to check for a single PCR product before running the quantitative PCRs. The quantitative PCRs were run on the Bio-Rad MyiQ Cycler for SYBR Green according to the manufacturer's recommendations for each probe. The appropriate master mixes were used for each application. The resulting PCR cycle time (Ct) values were collected by using the software provided for the iCycler, and the data were then analyzed in Microsoft EXCEL to determine ΔCt (test Ct - GAPDH Ct). The reverse transcription reactions were performed in triplicate with tissue RNA from three separate animals unless otherwise noted. PCR products were analyzed by agarose gel electrophoresis after 40 cycles for correct product size. Melt curves and standard curves were performed for SYBR Green reactions. The sequences of the primers used are available upon request.

## Computational data analysis

The association of TMDRs with repetitive elements and annotated genes in the mouse genome was determined using in-house PERL scripts with manual validation based on the related genome annotation data available from the UCSC genome website at http://genome.ucsc.edu. The sequence similarity to human genome was calculated based on

the 500 bp mouse genome sequences flanking the NotI sites (i.e.: 250 bp on each side) of TDMRs. The function classification and statistical over-representation of gene function categories represented by TDMR associated genes were analyzed through a combinatory use of the DAVID [46] (Dennis et al., 2003) and GOstat [47] programs. To examine statistical significance of the distribution of TDMR in association with repetitive elements, we performed Monte Carlo Simulation using a set of 8565 vRLGS mouse genomic fragments [6] computationally generated using enzyme combinations identical to those used in the experimental RLGS, i.e.: *Not*I plus *Pst*I or *Pvu*II for the 1st digestion and *Pvu*II or *Pst*I for 2nd digestion. Only the fragments predicted to be resolved by RLGS are used. For vRLGS fragments, we identified the *Not*I site position and their association with repetitive elements by assigning a value of zero for being located within a repetitive element or a positive number to represent the distance to the closest repetitive elements. 1000 random samplings of 34 *Not*I loci from these 8565 sites were performed. From these permutations, means values representing the number of *Not*I loci located within and the distance to repetitive elements are obtained and were used to calculate the Z-score and the p value for assessing the statistical significance of the observed association of TDMRs with repetitive elements. A similar analysis was performed to evaluate the distribution of TDMRs in relation to gene context.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

## References

1. Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. Trends Genet 1997;13:335–340. [PubMed: 9260521]

2. Kawai J, Hirotsune S, Hirose K, Fushiki S, Watanabe S, Hayashizaki Y. Methylation profiles of genomic DNA of mouse developmental brain detected by restriction landmark genomic scanning (RLGS) method. Nucleic Acids Res 1993;21:5604–5608. [PubMed: 8284204]

3. Watanabe S, Kawai J, Hirotsune S, Suzuki H, Hirose K, Taga C, et al. Accessibility to tissue-specific genes from methylation profiles of mouse brain genomic DNA. Electrophoresis 1995;16:218–226. [PubMed: 7774562]

4. Rouillard JM, Erson AE, Kuick R, Asakawa J, Wimmer K, Muleris M, et al. Virtual genome scan: a tool for restriction landmark-based scanning of the human genome. Genome Res 2001;11:1453–1459. [PubMed: 11483587]

5. Matsuyama T, Kimura MT, Koike K, Abe T, Nakano T, Asami T, et al. Global methylation screening in the Arabidopsis thaliana and Mus musculus genome: applications of virtual image restriction landmark genomic scanning (Vi-RLGS). Nucleic Acids Res 2003;31:4490–4496. [PubMed: 12888509]

6. Smiraglia DJ, Kazhiyur-Mannar R, Oakes CC, Wu YZ, Liang P, Ansari T, et al. Restriction landmark genomic scanning (RLGS) spot identification by second generation virtual RLGS in multiple genomes with multiple enzyme combinations. BMC Genomics 2007;8:446. [PubMed: 18053125]

7. Shiota K. DNA methylation profiles of CpG islands for cellular differentiation and development in mammals. Cytogenet Genome Res 2004;105:325–334. [PubMed: 15237220]

8. Song F, Smith JF, Kimura MT, Morrow AD, Matsuyama T, Nagase H, et al. Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. Proc Natl Acad Sci U S A 2005;102:3336–3341. [PubMed: 15728362]

9. Ching TT, Maunakea AK, Jun P, Hong C, Zardo G, Pinkel D, et al. Epigenome analyses using BAC microarrays identify evolutionary conservation of tissue-specific methylation of SHANK3. Nat Genet 2005;37:645–651. [PubMed: 15895082]

10. Oakes CC, La Salle S, Smiraglia DJ, Robaire B, Trasler JM. A unique configuration of genome-wide DNA methylation patterns in the testis. Proc Natl Acad Sci U S A 2007a;104:228–233. [PubMed: 17190809]

11. Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, et al. DNA methylation profiling of human chromosomes 6, 20 and 22. Nat Genet 2006;38:1378–1385. [PubMed: 17072317]

12. Kitamura E, Igarashi J, Morohashi A, Hida N, Oinuma T, Nemoto N, et al. Analysis of tissue-specific differentially methylated regions (TDMs) in humans. Genomics 2007;89:326–337. [PubMed: 17188838]

13. Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, et al. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 2005;37:853–862. [PubMed: 16007088]

14. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 2007;39:457–466. [PubMed: 17334365]

15. Shen L, Kondo Y, Guo Y, Zhang J, Zhang L, Ahmed S, et al. Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters. PLoS Genet 2007;3:2023–2036. [PubMed: 17967063]

16. Schilling E, Rehli M. Global, comparative analysis of tissue-specific promoter CpG methylation. Genomics 2007;90:314–323. [PubMed: 17582736]

17. Illingworth R, Kerr A, Desousa D, Jorgensen H, Ellis P, Stalker J, et al. A novel CpG island set identifies tissue-specific methylation at developmental gene loci. PLoS Biol 2008;6:e22. [PubMed: 18232738]

18. Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. Proc Natl Acad Sci U S A 2005;102:15785–15790. [PubMed: 16243968]

19. Su AI, Cooke MP, Ching KA, Hakak Y, Walker JR, Wiltshire T, et al. Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci U S A 2002;99:4465–4470. [PubMed: 11904358]

20. Nebel BR, Amarose AP, Hacket EM. Calendar of gametogenic development in the prepuberal male mouse. Science 1961;134:832–833. [PubMed: 13728067]

21. Maratou K, Forster T, Costa Y, Taggart M, Speed RM, Ireland J, et al. Expression profiling of the developing testis in wild-type and Dazl knockout mice. Mol Reprod Dev 2004;67:26–54. [PubMed: 14648873]

22. Bock C, Paulsen M, Tierling S, Mikeska T, Lengauer T, Walter J. CpG island methylation in human lymphocytes is highly correlated with DNA sequence, repeats, and predicted DNA structure. PLoS Genet 2006;2:e26. [PubMed: 16520826]

23. Baylin SB, Esteller M, Rountree MR, Bachman KE, Schuebel K, Herman JG. Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. Hum Mol Genet 2001;10:687–692. [PubMed: 11257100]

24. Ehrlich M. Expression of various genes is controlled by DNA methylation during mammalian development. J Cell Biochem 2003;88:899–910. [PubMed: 12616529]

25. Bell AC, Felsenfeld G. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature 2000;405:482–485. [PubMed: 10839546]

26. Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature 2000;405:486–489. [PubMed: 10839547]

27. De Smet C, Lurquin C, Lethe B, Martelange V, Boon T. DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. Mol Cell Biol 1999;19:7327–7335. [PubMed: 10523621]

28. Sato N, Kondo M, Arai K. The orphan nuclear receptor GCNF recruits DNA methyltransferase for Oct-3/4 silencing. Biochem Biophys Res Commun 2006;344:845–851. [PubMed: 16631596]

29. Gu P, Le Menuet D, Chung AC, Cooney AJ. Differential recruitment of methylated CpG binding domains by the orphan receptor GCNF initiates the repression and silencing of Oct4 expression. Mol Cell Biol 2006;26:9471–9483. [PubMed: 17030610]

30. Oda M, Yamagiwa A, Yamamoto S, Nakayama T, Tsumura A, Sasaki H, et al. DNA methylation regulates long-range gene silencing of an X-linked homeobox gene cluster in a lineage-specific manner. Genes Dev 2006;20:3382–3394. [PubMed: 17182866]

31. Frigola J, Song J, Stirzaker C, Hinshelwood RA, Peinado MA, Clark SJ. Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band. Nat Genet 2006;38:540–549. [PubMed: 16642018]

32. Oakes CC, La Salle S, Smiraglia DJ, Robaire B, Trasler JM. Developmental acquisition of genome-wide DNA methylation occurs prior to meiosis in male germ cells. Dev Biol 2007b;307:368–379. [PubMed: 17559830]

33. Shiota K, Kogo Y, Ohgane J, Imamura T, Urano A, Nishino K, et al. Epigenetic marks by DNA methylation specific to stem, germ and somatic cells in mice. Genes Cells 2002;7:961–969. [PubMed: 12296826]

34. Fouse SD, Shen Y, Pellegrini M, Cole S, Meissner A, Van Neste L, et al. Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. Cell Stem Cell 2008;2:160–169. [PubMed: 18371437]

35. Bibikova M, Laurent LC, Ren B, Loring JF, Fan JB. Unraveling epigenetic regulation in embryonic stem cells. Cell Stem Cell 2008;2:123–134. [PubMed: 18371433]

36. Rakyan V, Down T, Thorne N, Flicek P, Kulesha E, Graf S, et al. An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). Genome Res. 2008

37. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature. 2008

38. Nagy, A.; Gertsenstein, M.; Vintersten, K.; Behringer, R. A laboratory manual. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, New York: 1991. Manipulating the mouse Embryo.

39. Hatada I, Hayashizaki Y, Hirotsune S, Komatsubara H, Mukai T. A genomic scanning method for higher organisms using restriction sites as landmarks. Proc Natl Acad Sci U S A 1991;88:9523–9527. [PubMed: 1946366]

40. Hayashizaki, Y.; Watanabe, S. Restriction Landmark Genomic Scanning (RLGS). In: Hayashizaki, Y.; Watanabe, S., editors. A laboratory manual. Spring-Verlag; Tokyo: 1997. p. 1-179.

41. Costello JF, Smiraglia DJ, Plass C. Restriction landmark genome scanning. Methods 2002;27:144–149. [PubMed: 12095273]

42. Chomczynski P. A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. Biotechniques 1993;15:536–537.

43. Sambrook, J.; Russel, DW. A laboratory manual. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, New York: 2001. Molecular Cloning.

44. Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci U S A 1996;93:9821–9826. [PubMed: 8790415]

45. Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics 2002;18:1427–1431. [PubMed: 12424112]

46. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 2003;4:P3. [PubMed: 12734009]

47. Falcon S, Gentleman R. Using GOstats to test gene lists for GO term association. Bioinformatics 2007;23:257–258. [PubMed: 17098774]

ES cell RLGS (*Not*I-*Pst*I-*Pvu*II)



Comparison of RLGS Profile of testis and ES cell
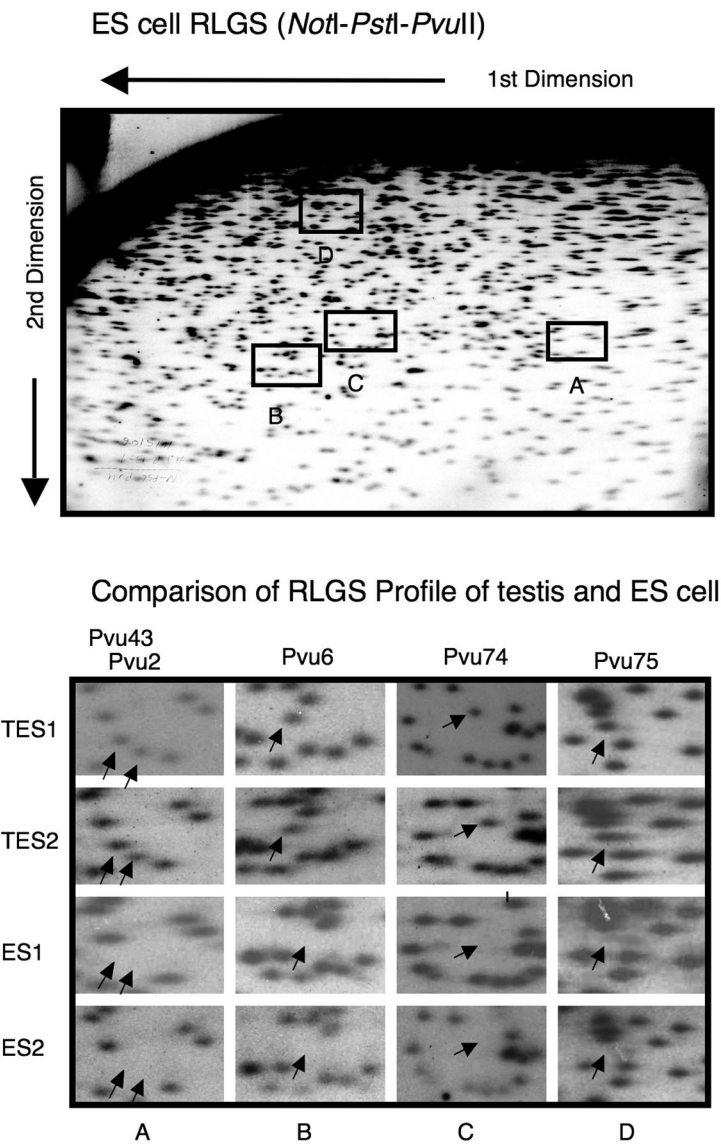


**Figure 1.**
RLGS analysis of Testis and ES cell DNA. A full RLGS profile of ES cell DNA and portions of RLGS profiles (*Not*I-*Pst*I-*Pvu*II) are shown. Arrows indicate position of Pvu 43, 2, 6, 74, and 75 (see Table 2). Testis DNAs are from two different mice (C57BL/6J) and ES1 and ES2 correspond to C57BL/6J Stewart Bruce 4 ES cells, passage 13 and 17 respectively.
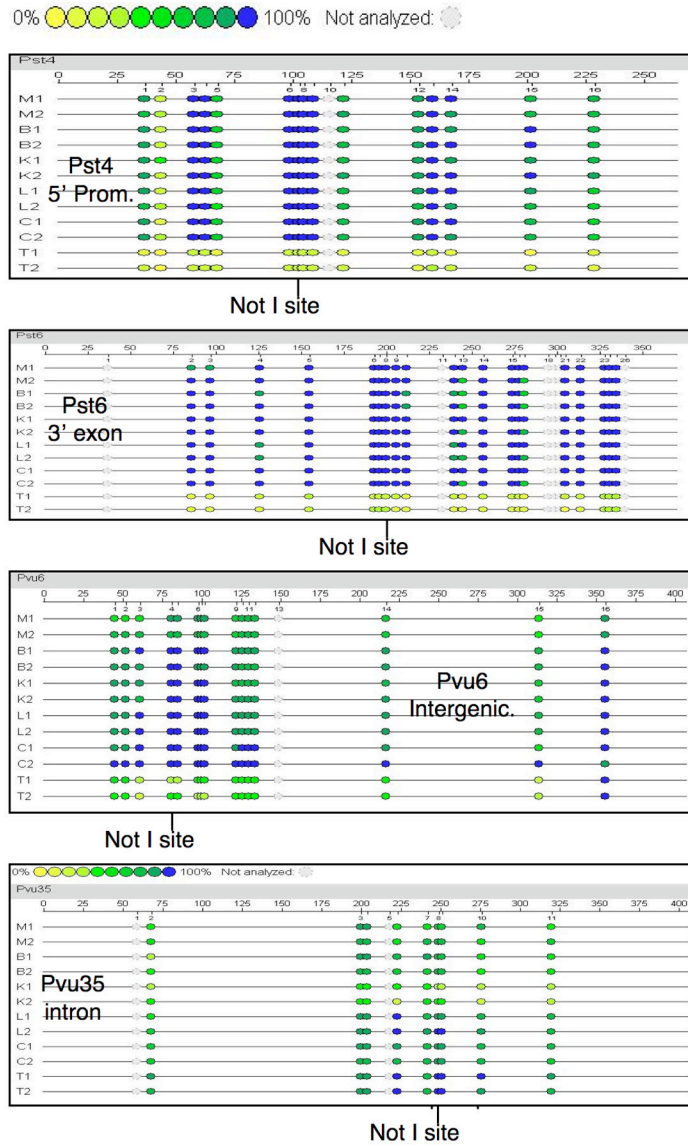
**Figure 2.**
TDMRs confirmed by Sequenom MassARRAY methylation analysis. The figure shows the analysis of two sets of mouse tissue samples including muscle (M), brain (B), kidney (K), liver (L), colon (C) and testis (T). The colored circles indicate the degree of methylation with yellow representing 0% methylation and blue representing 100%. The position of the Restriction Landmark *Not*I site is indicated. The four TDMRs (Pst4, Pst6, Pvu6, Pvu35) shown are located at 5′ promoter, 3′ exon, intergenic and intron, respectively.
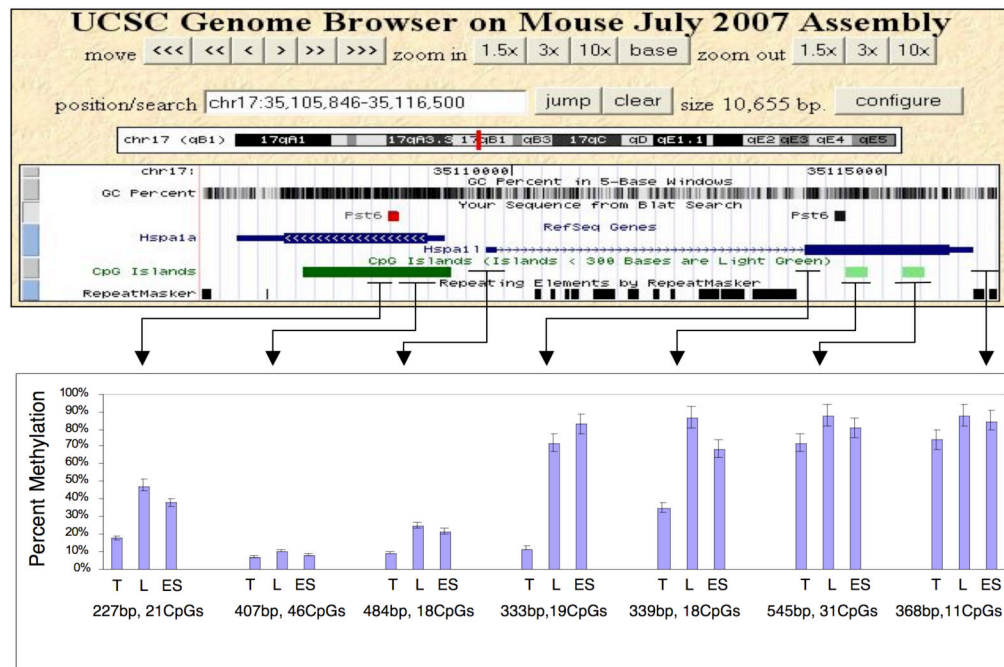
**Figure 3.**
Methylation fine structure associated with the Pst6 TDMR. The location of Pst6 TDMR within Hspa1l and a homologous region within the Hspa1a gene are indicated in mouse July, 2007 UCSC assembly. The GC percent, positions of CpG islands and repeat sequences are also shown. The lower panel indicates the percentage of methylation in adult testis (T), liver (L) and ES cells. Other somatic tissues (muscle, brain, kidney, and colon) were essentially identical to liver (see Figure 2). The length (bp) and number of CpGs in each region analyzed by Sequenom MassARRAY methylation analysis are also shown.
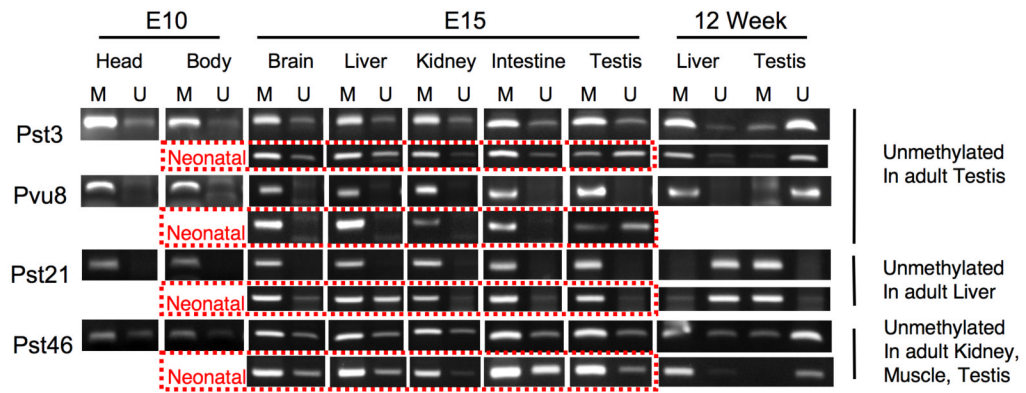
**Figure 4.**
MSP analysis of methylation status of TDMRs during development. MSP primers that would amplify methylated (M) or Unmethylated (U) genomic regions that contain, or were close to, the TDMR *Not*I restriction landmark site were designed by using METHPRIMER (http://www.urogene.org/methprimer/). DNAs from BAC clones (bacterial artificial chromosome) containing the TDMR region were used as a positive control for the U primers (not shown). Sss1 methylase-treated DNA was used as a positive control for the M primers (not shown). Results from 2 independently derived DNA samples from 12 wk liver and testis are shown.
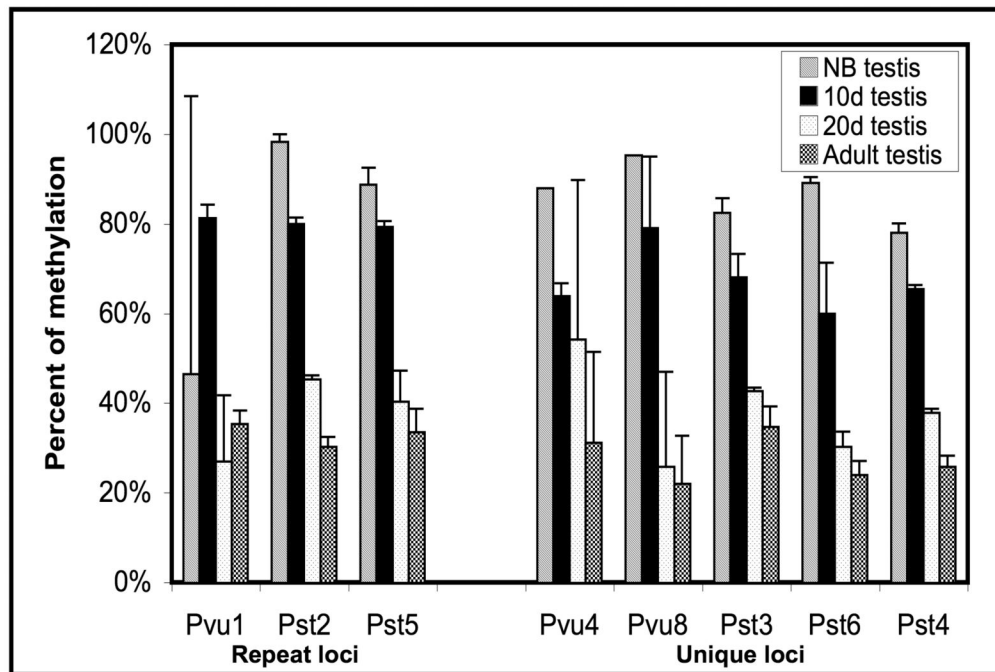
**Figure 5.**
Demethylation of repeat and unique sequences during postnatal testis development. The Bar graph shows a summary of the methylation status obtained from Sequenom MassArray methylation analysis of testis samples at different stages of development from new born (NB), 10 day old (10d), 20 day old (20d), and adult mice (12 wk) at each locus. The regions analyzed are the following: Pvu1, 7 CpGs (400bp); Pst2, 13 CpGs (225bp); Pst5, 19 CpGs (425bp); Pvu4, 43 CpGs (700bp); Pvu8, 25 CpGs (500bp); Pst3, 59 CpGs (525bp); Pst6, 26 CpGs (350bp); Pst4, 16 CpGs (250bp). The y- axis shows the percent of methylation obtained as an average level of methylation of the CpG dinucleotides for each locus for two independent determinations. The x-axis shows loci of interest (loci associated with either repeat or unique sequences). The error bars indicate standard deviation of the mean. It is to be noted that for the NB samples Pvu4 and Pvu8 loci there are no error bars as these were derived from only one set of samples.
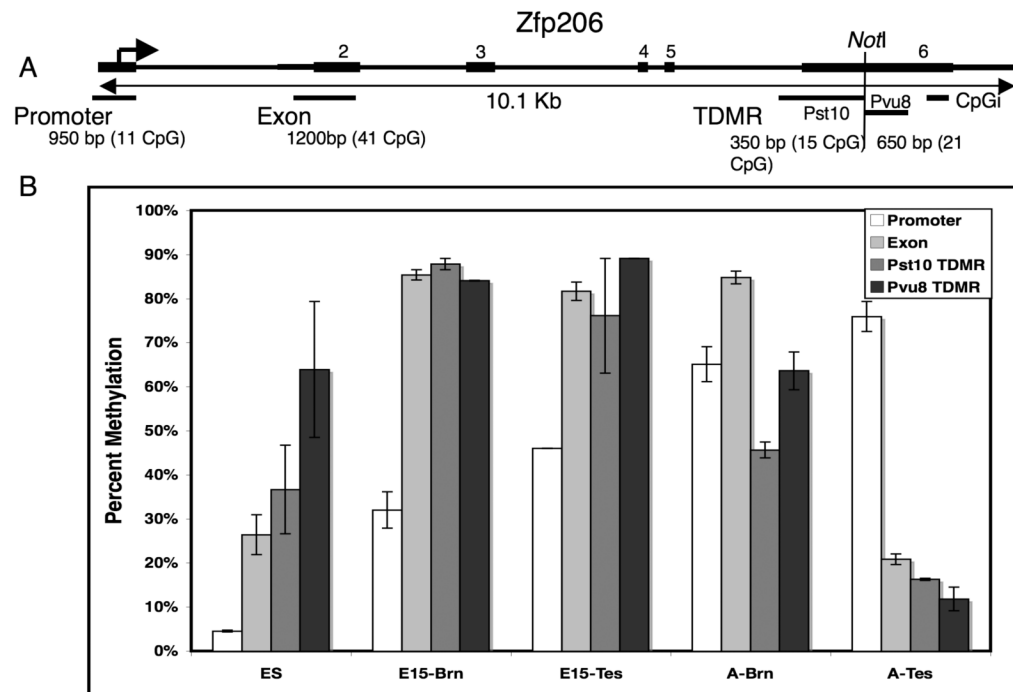
**Figure 6.**
Tissue and developmental stage specific differences in methylation. (A) A diagram of 10.1 kb Zfp206 zinc finger gene region is shown. The positions of exons (black bars), the *Not*I restriction landmark site, and the transcription start site (arrow) are indicated. The locations of a CpG island and the regions analyzed by Sequenom massARRAY methylation analysis (promoter, exon2, two TDMRs [Pst 10 and Pvu 8]) are indicated by black bars under the gene diagram. The size of the regions (bp) and number of CpG analyzed for methylation by Sequenom massARRAY are also shown. (B) The bar graph indicates the percent of methylation in ES cells, 15 day embryo and adult tissues in the indicated regions. The methylation values were calculated as the mean of two independent determinations. The error bars indicate the standard deviation of the mean.

**Table 1**

No legend

| | Number | Percent (%) | Confirmed Loci | Confirmed Loci (%) |
|---|---|---|---|---|
| RLGS loci surveyed | 3200 | ND | ND | ND |
| Tissue Specific loci | 150 | ND | ND | ND |
| Virtual RLGS identification | 68 | 100 | 34 | 100 |
| CpG island | 31 | 46 | 11 | 32 |
| 5′ promoter | 19 | 28 | 4 | 12 |
| CpG island & 5′ promoter | 14 | 21 | 3 | 9 |
| Non promoter Intragenic | 37 | 54 | 23 | 68 |
| 3′ exon | 12 | 18 | 8 | 24 |
| Other exon | 8 | 12 | 6 | 18 |
| Intron | 17 | 25 | 9 | 26 |
| Intergenic | 12 | 18 | 7 | 21 |
| Repeat associated TDMRs | 18 | 26 | 7 | 21 |

**Table 2**

Confirmed TDMRs. The table provides the loci; the RLGS spot intensity in ES cells, liver, kidney, brain, colon, muscle, and testis (2= dipoild, 1= haploid, 0= absent), which is an indication of the relative methylation at the NotI site of the TDMR; UCSC genomic position; location relative to CpG island; location relative to nearby gene; location and type of nearest repeat sequence relative to TDMR *Not*I site; Gene symbol; gene function; and human homology using a 500 bp mouse sequence centered on the *Not*I site, and whether Human sequence is within a CpG island. Pst10 and Pvu8 identified the same *Not*I site.

| RLGS (0=Methylated; 2=Unmethylated) | | | | | | | Genomic position (UCSC mm9,July, 2007) | CpGi* | Gene Context | GeneID | Gene function | Human Ortholog** |
| ES | L | K | B | C | M | T | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 2 | chr13:23,870,412-23,870,679 | N | intergenic | | | None |
| 1 | 0 | 0 | 0 | 0 | 0 | 2 | chr13:113,442,268-113,442,637 | Y | 5' promoter | Ddx4 | Enzyme | F, CpGi |
| 0 | 0 | 0 | 0 | 0 | 0 | 2 | chr9:62,130,029-62,130,377 | Y | 5' promoter | Spesp1 | Sperm Membrane Protein | F, CpGi |
| 0 | 0 | 0 | 0 | 0 | 0 | 2 | chr2:166,334,892-166,335,070 | N | intergenic | | | None |
| 0 | 0 | 0 | 0 | 0 | 0 | 2 | chr17:35,114,326-35,114,476 | Y | 3' exon | Hspa1l | Adaptor/Regulator | F |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | chr13:47,740,395-47,740,624 | N | intergenic | | | F |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | chr2:86,339,069-86,339,456 | N | intergenic | | | None |
| 1 | 0 | 0 | 0 | 0 | 0 | 2 | chr17:23,745,919-23,746,668 | Y | 3' exon | Zscan10 | Transcription Factor | P, CpG |
| 1 | 2 | 0.2 | 0 | 0 | 0 | 0 | chr19:6,506,916-6,507,202 | N | intron | Nrxn2 | Receptor | F |
| 1 | 0 | 0.5 | 0 | 0 | 0 | 2 | chr12:72,418,916-72,419,073 | Y | 3' exon | Dact1 | Adaptor/Regulator | F |
| 0 | 0 | 0.5 | 0 | 0 | 0 | 2 | chr7:53,047,584-53,048,525 | N | exon | Lmtk3 | Kinase | F, CpGi |
| 0 | 0.1 | 2 | 0 | 0 | 1 | 2 | chr2:72,650,905-72,651,173 | N | intron | AK019889 | Unknown | F |
| 0 | 0 | 2 | 0 | 0 | 2 | 2 | chr17:19,068,300-19,068,483 | N | intron | Vmn2r99 | Receptor | None |
| 2 | 0 | 2 | 2 | 2 | 2 | 2 | chr6:88,143,544-88,143,892 | Y | 5' promoter | Gata2 | Transcription Factor | F, CpGi |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | chrX:82,550,596-82,550,800 | N | intergenic | | | None |
| 0 | 0 | 0 | 0 | 0 | 0 | 2 | chr11:115,330,714-115,330,950 | N | exon | Slc16a5 | Channel/Transporter | F |
| 0.5 | 0 | 0 | 0 | 0 | 0 | 1 | chr7:13,556,457-13,556,671 | Y | 3' exon | Zfp324 | Transcription Factor | F, CpGi |
| 1.5 | 0 | 0 | 0 | 0 | 0 | 1.5 | chr6:48,570,188-48,570,357 | N | 3' exon | Zfp775 | Transcription Factor | F, CpGi |
| 0 | 0 | 0 | 0 | 0 | 0 | 1.5 | chr2:166,335,063-166,335,276 | N | intergenic | | | P |
| 0 | 0.2 | 0 | 0.1 | 0 | 0 | 2 | chr12:112,266,963-112,267,177 | N | intergenic | | | P |
| 1.5 | 0 | 0 | 0 | 0 | 0 | 2 | chr17:23,746,661-23,746,884 | Y | 3' exon | Zscan10 | Transcription Factor | F, CpGi |
| 1.5 | 0 | 0 | 0 | 0 | 0 | 2 | chr10:80,740,143-80,740,491 | N | intron | Tjp3 | Adaptor/Regulator | P |
| 0 | 0 | 0 | 0 | 0 | 0 | 2 | chr1:167,973,707-167,974,203 | N | intergenic | | | P |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | chr14:115,954,195-115,954,552 | N | intron | Gpc5 | Membrane Protein | None |
| 2 | 0 | 0 | 0 | 0 | 0 | 2 | chr8:111,318,491-111,318,829 | N | exon | Zfhx3 | Transcription Factor | F |

| RLGS (0=Methylated; 2=Unmethylated) | | | | | | | Genomic position (UCSC mm9,July, 2007) | CpGi* | Gene Context | GeneID | Gene function | Human Ortholog** |
| ES | L | K | B | C | M | T | | | | | | |
| 0 | 0 | 2 | 0 | 0.1 | 0 | 0 | chr4:148,294,717-148,294,921 | N | intron | Casz1 | Adaptor/Regulator | **P** |
| 2 | 0 | 0.1 | 0.1 | 1 | 0 | 2 | chr8:124,860,388-124,860,558 | Y | 3′ exon | Zfpm1 | Adaptor/Regulator | **P, CpGi** |
| 0 | 1 | 1 | 0 | 0 | 0 | 2 | chr15:82,413,163-82,413,352 | N | intron | U58494 | Unknown | **None** |
| 1 | 2 | 0 | 2 | 0 | 0 | 2 | chr14:120,788,365-120,788,475 | N | exon | Mbnl2 | Adaptor/Regulator | **F, CpGi** |
| 1.5 | 0.1 | 2 | ND | 2 | 1 | 2 | chr1:43,589,525-43,589,684 | Y | 3′ exon | Adra1b | Receptor | **F, CpGi** |
| 0 | 1 | 1 | 0.5 | 1 | 1 | 1 | chr2:165,006,494-165,006,766 | N | exon | Cdh22 | Adaptor/Regulator | **F, CpGi** |
| 0 | 1 | 1 | 0.2 | 0.5 | 1 | 2 | chr1:156,572,334-156,572,941 | N | 5′ promoter | Cacna1e | Channel/Transporter | **P, CpGi** |
| 0 | 2 | 2 | 2 | 0.5 | 2 | 0.1 | chr17:48,673,331-48,673,623 | N | intron | Unc5cl | Adaptor/Regulator | **F** |
| 1.5 | 0.5 | 0.5 | 1 | 0.1 | 0.5 | 2 | chr17:47,809,157-47,809,384 | Y | exon | Usp49 | Protease | **F, CpGi** |

Y, yes;

ll length homology; P, rtial length homology; CpGi, human region containing CpG islands