UNIVERSITY OF NAPLES
"FEDERICO II"

DEPARTMENT OF BIOLOGY AND CELLULAR AND MOLECULAR PATHOLOGY "L. CALIFANO"

DOCTORATE IN:

Pathology and Molecular Pathophysiology

PROGRAM N. XXI 2005-2008

Transcription influences repair-induced DNA methylation

Coordinator
Ch.mo Prof. V. E. Avvedimento
Supervisor
Ch.mo Prof. V. E. Avvedimento

Author
Annalisa Morano

INDEX

ABSTRACT:.....**Pag. 4**

INTRODUCTION:.....**Pag. 5**

- **DNA DAMAGE**.....**Pag. 5**
- **DOUBLE-STRAND BREAKS AND RECOMBINATION-DIRECTED REPAIR**.....**Pag. 7**
- **DNA METHYLATION**.....**Pag. 11**
- **METHYLATION AND TRANSCRIPTION**.....**Pag. 19**

AIM:

- **MECHANISM OF DNA METHYLATION INDUCED BY HOMOLOGOUS REPAIR**.....**Pag. 23**

RESULTS:.....**Pag. 24**

- **DNA METHYLATION AND HOMOLOGOUS RECOMBINATION**.....**Pag. 24**

  I. **Recombination assay** .....**Pag. 25**

  II. **Methylation reduces GFP expression but not recombination Frequency**.....**Pag. 30**

III. Effect of the integration site on the expression of the recombinanat gene. Analysis of individual DR-GFP clones……Pag. 34

IV. DNA Methyltransferase I inhibits the expression of recombinant GFP genes……Pag. 40

V. CpG methylation before and after repair: Analysis of individual molecules……Pag. 42

VI. Dnmt1 Is Associated with Recombinant Chromatin…….Pag. 51

- HOMOLOGOUS RECOMBINATION AND TRANSCRIPTION…….Pag. 53

I. Transcription of GFP gene is influenced by DNA methylation induced by homologous repair……Pag. 53

II. Homologous recombination and methylation induced by homologous repair are dependent on transcription……Pag. 56

DISCUSSION:……Pag. 60

- DOUBLE STRAND BREAK AND DNA METHYLATION……Pag. 60
- REPAIR AND METHYLATION……Pag. 63
- REPAIR AND GENE SILENCING……Pag. 64
- TRANSCRIPTION AND GENE SILENCING……Pag. 65

MATERIALS AND METHODS……Pag. 68

BIBLIOGRAPHY……Pag. 76

ABSTRACT

This work is aimed at the dissection of the molecular mechanism(s) linking DNA damage and gene silencing. To this end, we have developed a genetic system that allows a rapid assessment of homologous-directed repair (HR) of a single DNA double strand break (DSB). Briefly, we induced a DBS in the genome of HeLa or mouse embryonic stem (ES) cells using the I-SceI restriction endonuclease. Homologous recombination repair by gene conversion, initiated at the site of the double strand break, converts 2 inactivated tandem repeated green fluorescent protein (GFP) genes (DR-GFP) in an intact functional gene. The efficiency of HR, under our conditions, is approximately 2%–4% and can be easily quantified by analyzing GFP^+ cells.

Half of these recombinants expressed GFP poorly, because GFP gene was silenced. Silencing was rapid and associated with HR and DNA methylation of the recombinant gene, since HeLa DR-GFP treatment with 5-aza-2'-deoxycytidine, a DNA demethylating drug, significantly increased the fraction of GFP expressing cells. Methylation did not alter recombination frequency in both cell types. ES cells deficient in DNA methyl-transferase 1 yielded as many recombinants as wild-type cells, but most of these recombinants expressed GFP robustly.

Bisulfite analysis of GFP DNA molecules revealed that approximately half of the HR repaired molecules were *de novo* methylated, principally at the 3'-end of the DSB in a range of ~300bp. The other half GFP molecules were hypomethylated. Uncleaved and non-homologous repaired molecules did not show changes of the methylation profile. DNA methyl-transferase 1 bound specifically to HR GFP DNA, as revealed by chromatin immunoprecipitation and RNA analysis. HR induced novel methylation profiles on top of the old patterns and contributed to the silencing of GFP expression.

Inhibition of transcription by α-amanitin for a very short period (6-24 h during I-SceI cleavage) significantly reduced the frequency of recombination. Surprisingly, the 2 classes of recombinants were better separated in terms of GFP expression. Methylation analysis showed that the methylated molecules were hypermethylated, whereas the hypomethylated GFP gene molecules were un-methylated, relative to the untreated samples. Taken together, our data support a mechanistic link between HR, DNA methylation and transcription.

We propose that stalled RNA polymerase molecules slow down homologous recombination by interfering possibly with DNA polymerase complex or strand

invasion. At the same time, the presence of RNA polymerase II transcription complex signal to DNMT1 the coding strand and facilitates strand selective DNA methylation. Overall, these data highlight a new and unexpected opportunity in understanding the mechanisms of silencing of damaged genes.

INTRODUCTION

DNA DAMAGE

The long-term survival of a species is driven by genetic changes, while the survival of individuals depends on the accurate transmission of the genetic information from a mother to daughter cells.

The DNA is continuously subjected to threats of different kind coming from intracellular and extracellular environment, which may induce an alteration of its primary code. Examples are the modifications of the bases by the oxidative metabolism, exposure to mutagens and errors during DNA replication.

DNA "damage" is therefore any change that introduces in the double helix structure a deviation from the normal such as: 1. changes in single bases which alter the sequence ; 2. structural distortions that create physical obstacles to the processes of replication and transcription, and which can cause deletions, fusions, aneuploidy and translocations. All these conditions may result in decrease of fitness and degenerative diseases such as cancer and aging in multicellular organisms, while causing death of unicellular organisms.

Most of DNA spontaneous changes are temporary because the damages are immediately corrected by appropriate repair machines. These repair systems recognize a wide range of DNA distortions: the modified bases are corrected by excision-recombination mechanisms, while specific repair systems are activated in case of annealing errors (*mut* system in E.Coli and MMR system in eukariotes). Defects in the

functioning of one of these systems may cause genomic instability resulting in an increase of the tumorigenic potential of the cell.

Cells have evolved complex signalling networks to carefully monitor the integrity of the genome during DNA replication, and to initiate cell cycle arrest, repair, or apoptotic responses if errors are detected, probably to eliminate those cells that have potentially catastrophic mutations. Cancer cells, on the other hand, undergo an array of genetic changes including mutations in the DNA repair pathways that allow them to escape these controls and barriers.

The mechanisms involved in the response to injury include:

a) damage recognition and activation of the checkpoints, which block the cell cycle progression and activate the repair systems;

b) removal of the damage and restoration of the double helix continuity, in order to prevent the transmission of damaged chromosomes or which replication is incomplete;

c) apoptosis induction to remove heavily damaged or seriously dysregulated cells (Sancar et al., 2004).

Specific signal molecules, able to move along the damage sites and often involved in more than a repair mechanism, allow the cell to continuously monitor changes in the DNA structure to ensure proper cell cycle progression from G1 to M phase. Modification of these DNA-associated proteins is intrinsic to pathway activation or DNA repair.

Fundamental parts of the DNA damage response are checkpoints, signaling cascades that regulate key aspects of the cellular metabolism by interacting with the cell cycle machine (Zhou et al.,2000). Chromatin modification is directly implicated in the development of these signaling cascades. DNA damage recognition and processing require chromatin modifications and events aiming at the coordination of checkpoint signaling with DNA repair or apoptosis. The ultimate goal is the preservation of genomic integrity through the coupling of repair to other

essential cellular activities such as DNA replication, gene expression, cell cycle progression and life or death decisions. For this reason, recently, many efforts in the description and in the understanding of the mechanistic processes of DNA damage-associated histone modifications have been particularly intensified (Kinner et al., 2008).

DOUBLE-STRAND BREAKS AND RECOMBINATION-DIRECTED REPAIR

DNA double-strand break (DBS) is one of the most serious threats to cells. DSBs can arise, directly or indirectly, due to exposure to chemical or radiological agents or at stalled or collapsed replication forks. A DSB must be repaired quickly and with sufficient accuracy to protect against detrimental chromosomal rearrangements, mutations or cell death.

The DSBs generate when two complementary strands of DNA break simultaneously in sites that are close enough that base-pairing and chromatin structure are insufficient to keep the DNA ends juxtaposed. So physically dissociated, these ends threaten to inappropriately recombine with other genomic sites and the risk to induce chromosomal translocations is high. If not repaired correctly, DSBs can lead to genomic instability and increased risk of cancer and degenerative diseases. In mammalian cells faulty DSBs repair compromises tissue and organ function; however, among multicellular eukaryotes, physiologic DSBs are found only in the vertebrate immune system V(D)J.

Because of threats posed by DSBs, eukaryotic cells have evolved complex, highly conserved systems to rapidly and efficiently detect these lesions, signal their presence, and bring about their repair.

There are at least two repair pathways which can repair DSBs: (1) homologous recombination (HR)-mediated repair and (2) non-homologous end-joining (NHEJ)- and/or micro-homology-mediated

recombination. These pathways are largely distinct from one another but function in complementary way (Essers et al., 2000).

The essence of HR is an exchange of information between two similar sequences. When provoked by a DSB, HR can serve not only for healing DNA strand discontinuities, but for restoring any genetic information that otherwise may have been lost due to exonucleolytic activity. During homologous recombination the damaged chromosome forms a synapse (synaptonemal complex) with a molecule of undamaged DNA by which it shares a remarkable sequence homology. Thus the genetic information lost on one allele can be picked up on the other remained intact. HR via DSB repair proceeds via two Holliday junction intermediates, and an event can resolve either as a crossover or as a gene conversion with no associated crossover. Providing a high homology requirement for recombination protects the integrity of the mammalian genome, particularly because of the abundance of similar repeated sequences (i.e. Alu family sequences) (Waldman 2008). HR repair is a templated repair process and is therefore error free.

In contrast, in non-homologous recombination, two ends of DNA that do not share sequence homology in terminal portions are ligated to each other without formation of synapses. NHEJ involves the direct religation of broken termini without use of the sister chromatid as a template. NHEJ does not use long stretches of homology, but the processing of the DNA ends can, at least in some cases, be influenced by terminal microhomology, the alignment of few homologous nucleotides, typically 1–4 nt. It should be noted that NHEJ proceeds even if there is no terminal microhomology. NHEJ often results in insertions or deletions of nucleotides at the repair site, leading to mutations within the genome.

Both pathways have been highly conserved during the evolution of the eukaryotic world (Cromie et al., 2001). Simple eukaryotes such as yeasts *S. Cerevisiae* e *S. Pombe* use homologous recombination to repair the

DSBs induced by ionizing radiation (Jasin, 1996), while non-homologous recombination is little used, except in case of absence of homology to the broken chromosome, or when the machinery of homologous recombination does not work. (Haber et al., 2000). In contrast, mammals use more frequently the non-homologous recombination.

In eukaryotes, homologous recombination is restricted to late S or G2 phase, where DSBs are often repaired at long regions (> 100bp) of homology using homologous recombination (although single-strand annealing can also occur); non-homologous DNA end joining (NHEJ) is instead the dominant pathway for the repair of DSBs in multicellular eukaryotes throughout the cell cycle.

NHEJ is distintive for the amount of the nuclease, polymerase, and ligase activities that are used. These activities permits NHEJ to function on the wide range of substrate configurations that can arise when double-strand breaks occur, particularly at sites of oxidative damage or ionizing radiation, but NHEJ does not return the local DNA to its original sequence. Pathologically, the imprecisions of NHEJ contribute to mutations that arise over time. Physiologic double-strand break processes use the imprecisions of NHEJ in generating antigen receptor diversity (Lieber, 2008).

It is not yet clear what determines whether a DSB is repaired by NHEJ *versus* homologous recombination during DNA replication. Recently, it was reported that generation of DSBs associated with DNA replication stresses such as stalled replication forks closely related to cancer incidences and that these DNA replication-related DSBs are repaired through the HR pathway (McCabe et al., 2006). This finding suggests the importance of HR repair for cancer prevention.

DSB activates signaling responses, termed cell-cycle checkpoints, which monitor DNA damage and transduce signals to coordinate repair and cell cycle progression. One of the key players of the cell-cycle checkpoints is

the tumor suppressor protein p53. p53 is activated and posttranscriptionally modified in response to DNA damage. These modifications include phosphorylation by ataxia teleangiectasia mutated (ATM), a central signaling kinase in the response to DNA damage. After DNA damage, p53 activates genes involved in DNA repair, cell cycle control and apoptosis, and takes part in the maintenance of the genome integrity (Shyloh., 2003). When DSBs are generated, ATM protein kinase is activated and relocates through an interaction with Rad50/Mre11/NBS1 complex in mammals and Mre11-Rad50-Xrs2 in yeast. Then ATM phosphorylates istone H2AX and many other substrate proteins including Artemis, DNA-PKcs kinase, MDC1, NBS1, p53 and Chk2. ATM-phosphorylated proteins activate cell cycle checkpoints, NHEJ repair pathway, and HR-related pathways (Kobayashi et al., 2008). Moreover, proteins involved in HR pathway are often ubiquitinated and this seems to be essential for HR repair (Spence et al., 2000).

One type of homologous repair is gene conversion**,** an event in DNA recombination which occurs at high frequency during meiotic division but which also occurs in somatic cells. It is a form of non-reciprocal recombination that can either maintain genetic identity or promote genetic diversity (Santoyo et al., 2005). During gene conversion DNA sequence information is transferred from one DNA helix, which remains unchanged, to another DNA helix, whose sequence is altered. Every gene conversion event takes as its substrate two DNA sequences that are homologous but not identical, because of sequence mismatches, and yields two identical DNA sequences. This conversion of one allele to the other is due to base mismatch repair during recombination: if one of the four strands during meiosis pairs up with one of the four strands of a different chromosome, as can occur if there is sequence homology, mismatch repair can alter the sequence of one of the chromosomes, so that it is identical to the other.

Gene conversion can also result from the DNA repair of DSBs. Here a break in both strands of DNA is repaired from an intact homologous region. Resection of the DNA strands near the break site leads to stretches of single stranded DNA that can invade the homologous DNA strand. The intact DNA can then function as a template to copy the lost information on the other strand. During this repair process a double Holliday structure is formed. Depending on how this structure is resolved, either cross-over or gene conversion products result. Gene conversion acts to "homogenize" the DNA sequences composing the gene pool of a species. Over time, gene conversion events yield a homogenous set of DNA sequences, both for allelic forms of a gene and for multigene families.

Gene conversions were first observed over 80 years ago, and then extensively studied, especially in yeast (Ezawa et al., 2006). In humans, gene conversions between multigene family members have been described in several protein coding genes. For example, gene conversions occur between human genes coding for β-globins, opsins, ubiquitins and also for the large palindromic sequences found in the human Y chromosome (Rozen et al., 2003; Benovoy et al., 2008).

DNA METHYLATION

DNA methylation is a covalent, postreplicative modification of genomic DNA. Changes in the methylation pattern are associated with specific developmental and differentiation stages (Razin and Shemer, 1995), imprinting (Reik and Walter, 1998), X chromosome inactivation (Latham, 1996) and carcinogenesis (Counts and Goodman, 1995).

Cancer cells and tissues exhibit genome wide hypomethylation and regional hypermethylation. In vertebrates, the preferred substrates for methylation are cytosines located within the dinucleotide CpG, whereas in plants and fungi also cytosines located outside of this sequence

context can be methylated (Colot and Rossignol, 1999). In mammals, methylation takes place exclusively on the C5 carbon of the cytosine belonging to the CpG (cytosine-guanine) dinucleotide by specific enzymes called DNA methyltransferases. CpG methylation of DNA is characterized by the formation of a C–C covalent bond between the 5′-C of cytosine and the –CH3 group of S-adenosylmethionine. Removal of the methyl-group from the methylated cytosine of DNA is one of the ways of DNA demethylation.

The mammalian genome contains a very little amount (2-3%) of 5'-methylcytosines (Patra et al., 2008). The distribution in the genome of CpG dinucleotides is quite asymmetric; often they are grouped in genomic regions known as "CpG islands". In mammalian genomes, CpG islands are typically 300-3,000 base pairs in length. The formal definition of a CpG island is a region with at least 200 bp, in which the GC percentage is greater than 50% and the observed/expected CpG ratio is greater than 60%. The "p" in CpG notation refers to the phosphodiester bond between the cytosine and the guanine. They are in and near approximately 40% of promoters of mammalian genes (about 70% in human promoters). In vertebrates CpG islands typically occur at or near the transcription start site of genes, particularly housekeeping genes. Normally a C (cytosine) base followed immediately by a G (guanine) base (a CpG) is rare in vertebrate DNA because the cytosines tend to be methylated in such an arrangement. Because 5-methylcytosine is chemically very similar to thymidine, CpG sites are frequently mutated and become rare in the genome. This methylation helps distinguish the newly synthesized DNA strand from the parent strand, which aids in the final stages of DNA proofreading after duplication. About 80% of the CpG dinucleotides that are not associated with CpG islands are heavily methylated; in contrast, in normal cells the dinucleotides in CpG islands, especially those associated with gene promoters, are usually unmethylated, whether or not the gene is being

transcribed (Bird, 2002), with the important exceptions of inactivated X-chromosome and imprinted genes. Aberrant CpG methylation has been observed in several tumors (Baylin et al., 2000); in fact, some CpG islands are hypermethylated in tumor cells (Yan et al., 2003).

DNA methylation contributes to the formation of a nuclease resistant chromatin, which results in a transcriptional silent state of the genes (Bird and Wolff, 1999). The methylation of DNA can also maintain wide non-coding regions of the genome of higher organisms in a transcriptional inert state. So at the variance with a more simple organism such yeast and *drosophila*, in mammals the control of gene expression occurs by epigenetic modification too.

In biology, the term "epigenetics" refers to heritable changes in phenotype or gene expression caused by mechanisms other than changes in the underlying DNA sequence. Epigenetic changes are preserved when cells divide. Most epigenetic changes only occur within the course of one individual organism's lifetime, but some epigenetic changes are inherited from one generation to the next. However, there is no change in the underlying DNA sequence of the organism, but non-genetic factors cause the organism's genes to behave differently.

The best example of epigenetic changes in eukaryotic biology is the process of cellular differentiation. During morphogenesis, totipotent stem cells become the different pluripotent cell lines of the embryo which in turn become fully differentiated cells, by activating some genes while inhibiting others.

A gene silenced may be reactivated with an opposite modification, such as demethylation (Bhattacharya et al., 1999), while gene silencing induced by mutations is irreversible.

Altered expression of a gene can be caused by aberrant epigenetic modifications of the chromatin. There are two major epigenetic gene silencing mechanisms that account for a growing number of diseases: cytosine DNA methylation and covalent histone modification.

Because the phenotype of a cell or individual depends on how its genes are transcribed, heritable transcription states can give rise to epigenetic effects.

There are several layers of regulation of gene expression and one is through the remodeling of chromatin. Chromatin is the complex of DNA and the histone proteins with which it associates. Recent data have shown that methylation of DNA and deacetylation of histones H3 and H4 leads to inactivation/repression, while selective acetylation of histones H1, H3, H4, methylation of H3-K4, and DNA demethylation are associated with activation of nucleosomes and gene transcription (Patra and Bettuzzi, 2007).

Two models have been proposed to explain the silencing effects of methylation on transcription. The first model suggests that the methylation of cytosine located at the level of promoters interferes with the binding of transcriptional factors that require contact with cytosine in the major groove of the double helix, while the second model suggests that methylated sites are recognised by positive or negative trans-acting proteins, which modulate gene expression. Many of these factors work by sequestering genes in highly condensed chromatin structures. The repeated observation that actively transcribed genes are typically in an open configuration suggests that regulation of chromatin structure plays a fundamental role in gene expression (Domìnguez-Bendala and McWhir, 2004).

The search for proteins with different ability to bind methylated or unmethylated DNA led to the discovery of two proteins called MeCP1, a protein complex of 450 kDa, and MeCP2, a single polypeptide of 55 kDa (Lewis, 1992). In particular, MeCP2 has both a methyl-CpG-binding domain (MBD) that a domain of repression of transcription (TRD), which enables it to monitor gene expression even at a distance of hundreds of bases (Hendrich and Bird, 1998). MeCP2, in turn recruits HDACs and histone methyltransferases, resulting in an inactive chromatin structure

(Boyes and Bird, 1991). Some methylated DNA binding proteins (MeCP2, MBD1, MBD2, MBD3 and MBD4) selectively bind CpG and/or methylated CpG sequences, contributing to remodeling of nucleosomes and chromatin structure. Under these conditions, chromatin would be closed as a consequence of histones deacetylation caused by recruitment of histone deacetylases (HDACs). These events precede and inhibit binding of transcription factors, including RNA polymerase; under these conditions, DNA demethylation could cause a reduction of the repression potential of a gene.

DNA methylation in mammalian cells is regulated by a family of highly related DNMTs. The DNA-methyltransferase recognized in humans and mice are:

- cytosine DNA methyltransferase-1(Dnmt1);
- cytosine DNA methyltransferase-3a (Dnmt3a);
- cytosine DNA methyltransferase-3b (Dnmt3b).

(Bestor et al., 1988; Okano et al., 1998).

Dnmt1 is ubiquitously expressed in proliferating cells and, in vitro, prefers hemimethylated DNA over non-methylated DNA as substrate. This property of mammalian DNA methyltransferases is in stark contrast with the activity of bacterial DNA-methyltransferases, which do not discriminate between methylated and unmethylated target sequences. Inactivation of Dnmt1 in ES (embryonic stem) cells and mice leads to extensive demethylation of all analyzed sequences (Li et al., 1992), but doesn’t inhibit the proliferation (Hong et al., 1996). All this suggests the role of Dnmt1 as a "maintenance methyltransferase”, responsible for copying the parental-strand methylation pattern into the newly synthesized strand after each round of replication.

Dnmt1 is therefore responsible for propagation and maintenance of established methylation patterns during embryonal development and cell division. Its expression is properly cell cycle-regulated in normal cells

(Leonhardt et al., 1992). Dnmt1 is recruited to replication foci through protein interactions involving the Dnmt1-associated protein-binding region (Rountree et al., 2000), the proliferating cell nuclear antigen-binding region (Chuang et al., 1997) and the replication foci targeting sequence. Dnmt1 and PCNA accumulate to DNA damage sites induced by UVA radiation, where colocalize with γ-H2AX (Mortusewicz et al., 2005).

Dnmt1, has a transcriptional repression domain that binds histone deacetylase I (HDAC I) (Fuks et al., 2000) and could thus, together with chromatin assembly factor 1 (CAF-1) , contribute to the re-establishment of chromatin structures after histone modifications (Green et al., 2003). Finally, Dnmt1 may also participate in the identification of the template strand in several repair pathways as was suggested for MMR repair system (Kim et al., 2004). Interestingly, a low level of Dnmt1 overexpression leads to cell transformation, whereas a high level is toxic (Wu et al., 1993).

Dnmt1 participates in the repression of the transcription of promoters containing binding sites for E2F, establishing a close correlation between DNA methylation and gene-specific transcriptional repression (Robertson et al., 2000). Dnmt1 is not only involved in maintaining the methylation of DNA, but also directly into giving in effect a hereditary transcriptional silencing on specific genomic regions during replication. The mechanisms through which Dnmt1 causes cellular transformation and through which inhibition of Dnmt1 reverses cellular transformation are unknown. The most obvious mechanism is that aberrant expression of Dnmt1 causes methylation and silencing of tumor suppressor genes (McCabe et al., 2006). Knock down of Dnmt1 by either antisense or siRNA results in demethylation and activation of tumor suppressor genes, such as p16 and p21 (Robert et al., 2003). These data suggest that increased Dnmt1 levels and activity affect the methylation status of genes critical to tumor formation (You et al., 2008).

In contrast to the maintenance methyltransferase Dnmt1, the *de novo* methyltransferases Dnmt3a and Dnmt3b are responsible for establishing new DNA methylation patterns during development and show a low and tissue-specific expression. They are strongly expressed in ES cells, early embyos and developing germ cells, but are expressed at low levels in differentiated somatic cells. Genetic studies have demonstrated that Dnmt3a and Dnmt3b are essential for *de novo* methylation in ES cells and post-implantation embyos, as well as for *de novo* methylation of imprinted genes in the germ cells (Okano et al., 1999). Although Dnmt3a and Dnmt3b function primarily as *de novo* methyltransferases to establish methylation patterns, they may also play role in maintaining methylation patterns. Similarly to Dnmt1, even Dnmt3b associates with histone deacetylase 1 (HDAC1) and RP58, a repressive transcriptional factor that binds the DNA found in transcriptional silent heterochromatic sites (Fuks et al., 2001).

However, these classifications are oversimplified as Dnmt1 is also known to possess *de novo* methylation activity (Hermann et al., 2004) and is the most abundant methyltransferase in somatic cells (Robertson et al., 1999).

In summary the mechanisms of gene silencing induced by methylation are:

- Chromatin-independent mechanisms:
  - interference with transcription factors;
  - MeCP2 contacts with the transcriptional machinery.
- Chromatin-dependent mechanisms:
  - DNMT1 association to histone deacetylase;
  - MBD recruitment of histone deacetylase;
  - MBD recruitment of ATP-dependent nucleosome remodelling enzymes.

(Ballestar and Esteller, 2002).

Methylation might also suppress homologous recombination: this possibility provides an attractive explanation as to how repeat-rich genomes can be stable, despite the increased number of opportunities for chromosome rearrangements.

Current models derived from studies in yeast are based on the central concept that meiotic recombination is initiated by a enzymatically-induced DNA double-strand break that has the same probability of occurring on one or the other of the two interacting chromatids. If methylation serves only to prevent nucleases from cutting methylated DNA, the non-methylated chromatid should undergo normal recombination initiation, where only one parent is methylated. As a result, only a two-fold reduction of crossovers is expected. The 50-fold reduction observed suggests instead that methylation acts primarily at steps that follow to the initial double-strand break. Methylation should also affect somatic recombination, usually initiated by accidental DNA double-strand breaks. The possible suppressing effect of methylation on somatic recombination appears to be particularly important in mammals. Indeed, many cancer cells show chromosomal rearrangements that might be caused by homologous recombination between repeats, and cancer cells are often hypomethylated (Baylin et al., 1998).

Understanding the mechanism by which DNA methylation is involved in the damage response, how it participates in the remodeling of chromatin and induces gene silencing is crucial for the possible development of drugs able to make function again the tumor suppressor genes often hypermethylated in many tumors.

A total of two types of DNA methylation can be distinguished. Stable methylation is inherited through generations in a male- or female-specific fashion and is responsible for both mono- and biallelic imprinting.

Metastable methylation is variable and generates different methylation patterns among individual cells and cell types. It is modified by environment and changes during the lifetime of individual cells.

Metastable methylation as the consequence of DNA damage repair is the subject of this study.

METHYLATION AND TRANSCRIPTION

Although DNA methylation is associated with silencing of several genes, such as tumor suppressor genes in cancer, the relationship between DNA methylation and gene transcription is complex.

Numerous studies, mainly in mammals, have revealed a strong correlation between the methylated state of DNA and gene silencing. Genes with methylated promoters are not expressed. Together these studies argue in favor of an inhibitory effect of methylation on transcription initiation. Although methylation can directly prevent the binding of some transcription factors to promoter sequences, its effect on transcription initiation seems to be indirect, depending on proteins that have an affinity for methylated CpGs (Nan et al., 1998).

Experiments with methylated templates microinjected into the nuclei of mammalian cells or in *Xenopus* oocytes indicate that transcriptional repression occurs in vivo only after chromatin assembly (Kass et al., 1993). Further experiments with *Xenopus* oocytes have shown that methylation in the coding region can trigger the time-dependent formation of a repressive nucleoprotein structure that spreads to the promoter (Kass et al., 1997).

The discovery that MeCP2 can recruit histone deacetylases, which are known to mediate the formation of repressive chromatin states, suggests

that, in gene silencing, a primary role of MeCP2 is to contribute to the formation of these states. Moreover, treatment with the histone deacetylase inhibitor TrichostatinA leads to a restoration of transcriptional competence on methylated chromatin templates. This indicates that methylation *per se* does not affect transcription through chromatin templates. Although DNA methylation can trigger the remodeling of chromatin into repressive states, these states can exist independently of methylation. This is obvious in organisms like *Drosophila* or yeasts, which lack methylation but, at the same time, display stable epigenetic repressed states. The available data on the formation of epigenetic repressive states in vertebrates suggest that chromatin changes are sufficient in themselves to ensure a silenced state and that methylation is used to reinforce the stability and efficiency of this state. Therefore, while methylation can trigger the formation of repressed chromatin, it can also be imposed on pre-existing repressed chromatin.

Transcription of genes by RNA-Polymerase II is a complex process that requires a highly coordinated and multistep process utilizing a large number of basal and transactivating factors. Furthermore, there exists a dynamic association of mRNA processing factors with differently modified forms of the polymerase throughout the transcription cycle (Komarnitsky et al, 2000). More specifically, the phosphorylation of the C-terminal domain in RNA-Polymerase II at serine 5 has been associated with transcription initiation. The principles and mechanisms underlying transcription are remarkably similar between eukaryotes and prokaryotes despite the increased complexity of eukaryotic transcription machinery (Hahn, 2004). The typical RNA polymerase II transcription cycle begins with the binding of activators upstream of the core promoter (including the TATA box and transcription start site). This event leads to the recruitment of the adaptor complexes such as SAGA or other mediators, which in turn facilitate binding of general transcription factors (GTFs;

Thomas and Chiang, 2006). Pol II is positioned at the core promoter by a combination of TFIID, TFIIA, and TFIIB to form the preinitiation complex. TFIIH then melts 11–15 bp of DNA in order to position the single strand template in the Pol II cleft to initiate RNA synthesis. The carboxy-terminal domain (CTD) of Pol II is phosphorylated by the TFIIH subunit during the first 30 bp of transcription and loses its contacts with GTFs before it proceeds onto the elongation stage. Meanwhile, the phosphorylated CTD begins to recruit the factors that are important for elongation and mRNA processing (Li et al., 2007).

RNA-Polymerase II interacts with chromatin remodeling enzymes, such as BRG1, a member of the SWI/SNF chromatin remodeling complex; with HATs and with chromatin-modifying enzymes, such as SET1. MLL1, a human equivalent of yeast SET1, is known to associate with highly expressed transcripts (Guenther et al., 2005).

RNA-Polymerase II transmits the change in promoter accessibility caused by transcription factor binding, recruitment of HAT and histone acetylation down the gene, thus translating early changes in promoter activity to more stable changes in chromatin structure (Orphanides and Reinberg, 2000). Histone acetylation enables initial recruitment of RNA-Polymerase II to the methylated promoter. The early progression of RNAP II along the gene either facilitates or directly recruits DNA demethylases, which demethylate the transcribed region, followed by demethylation of the promoter, a prerequisite for strong gene expression. The demethylated promoter significantly increases its association with RNAP II and the acetylation of histone tails, resulting in high levels of protein expression. It is possible, therefore, that cessation of transcription might lead to remethylation of the transcribed gene. Transcription and epigenetic programming due to demethylation might act coordinately in a positive feedback loop to maintain a gene in an active state (D'Alessio et al., 2007).

To study the effects of methylation in the coding sequence of a gene on transcription, we used a potent inhibitor of RNA polymerase II called alpha-amanitin, a cyclic nonribosomal peptide of eight amino acids. It is possibly the most deadly of all the amatoxins, toxins found in several members of the Amanita genus of mushrooms, one being the death cap (Amanita phalloides) as well as the destroying angel, a complex of similar species, principally *A. virosa* and *A. Bisporiga*. α-amanitin can also be used to determine which types of RNA polymerase are present, depending on their relative sensitivity to this drug. RNA polymerase I is insensitive, RNA pol II is highly sensitive, and RNA pol III is slightly sensitive.

The active principle of the ‘‘death cap’’ mushroom, α-amanitin, blocks both transcription initiation and elongation. The cocrystal structure suggests that α-amanitin interferes with a protein conformational change underlying the transcription mechanism. The α-amanitin binding site is beneath a ‘‘bridge helix’’ extending across the cleft between the two largest RNA-Polymerase II subunits, Rpb1 and Rpb2, in a ‘‘funnel’’-shaped cavity in the pol II structure. Bridge helix residues directly contact the DNA base paired with the first base in the RNA strand. Most pol II mutations affecting α-amanitin inhibition map to this site. After the addition of α-amanitin to a transcribing pol II complex, a phosphodiester bond can still be formed, but the rate of translocation of pol II on DNA is, however, reduced from several thousand to only a few nucleotides per minute. It may be explained by a constraint on bridge helix movement, and this movement is required for DNA translocation. Evidences coming from biochemical studies of transcription, from structure-activity relationship studies and from cocrystal structure determination suggest that binding of α-amanitin to pol II permits nucleotide entry to the active site and RNA synthesis but prevents the translocation of DNA and RNA required to empty the site for the next round of synthesis, thus inhibiting the further translocation (Bushnell et al., 2002).

AIM:

MECHANISM OF DNA METHYLATION INDUCED BY HOMOLOGOUS REPAIR

The ultimate goal of this study is to identify a mechanistic relationship between DNA methylation and homologous repair. To this end we will analyze the contribution of transcription to homologous repair and DNA methylation.

We have used a system pioneered by M. Jasin (Jasin et al.,1996, Pierce et al., 1999), in which recombination between partial duplications is initiated by a specific DSB in one copy. Recombination products can be detected by direct analysis of the DNA flanking the DSB or by the appearance of the product of the recombined gene. We have found that gene conversion profoundly modifies the methylation pattern of the repaired DNA and that this methylation silences the recombined gene. Dnmt1 is specifically associated with the chromatin of homologous repaired green fluorescent protein (GFP). We have shown that DNA methylation, induced by HR, marks the repaired DNA segments and

protects cells against unregulated gene expression following DNA damage.

Our data suggest that methylation induced by damage is strand-specific, because we invariantly find 2 populations arising from recombination, with different and discrete methylation profiles. Since RNA-Polymerase II transmits the change in promoter accessibility and recruits histone de-acetylases, thus translating early changes in promoter activity to more stable changes in chromatin structure (Orphanides and Reinberg, 2000), we have specifically analyzed if transcription is essential for the establishment of methylation–induced repair.

We show that active transcription is essential for recombination, but also dictates the methylation profiles.

RESULTS

DNA METHYLATION AND HOMOLOGOUS RECOMBINATION

DNA damage induced by oxidative stress or by constitutive expression of oncogenes is linked to gene silencing (El-Osta, 2004). To explore the molecular mechanism(s) linking DNA damage and gene silencing and to find a possible mechanistic relation, we used a new reporter system, by which we can monitor the homologous recombination repair of the double-strand breaks. The system is based on the use of a plasmid, called DR-GFP, modified so to express the green fluorescent protein (GFP) only after an event of recombination by gene conversion. The recombination repairs the cutting in a unique chromosomal site due to a site-specific DSB induced after the expression of a rare endonuclease, the enzyme I-SceI.

We can monitor:

1. the occurrence of recombination at this site ;
2. the expression profile of recombined and non-recombined

units;

3. the structure and the epigenetic modifications of the locus.

Furthermore, this system allows us to monitor how the frequency of recombination or the expression of recombinant units change following the treatment with drugs that affect DNA methylation or chromatin configuration.

I. Recombination assay

Our recombination assay relies on the two inactivated tandem repeated (DR)-GFP plasmid originally developed by M. Jasin at Rockfeller University in New York (Pierce et al., 1999), which contains two mutated GFP genes oriented as direct repeats and separated by a drug selection marker, the puromycin N-acetyltransferase gene (**Figure 1**). An upstream cytomegalovirus (CMV) enhancer fused to the chicken b-actin promoter provides a strong and insulated transcriptional unit. The upstream (5') GFP gene (cassette I) carries a recognition site for I-SceI, a rare-cutting endonuclease, encoded by a mitochondrial intron of *Saccharomyces cerevisiae.* This enzyme allows the induction of a site-specific DSB in the DR-GFP plasmid, as it recognizes a sequence of 18 bp absent in several eukaryotic genomes tested (Jasin, 1996). The I-SceI recognition sequence was incorporated into a BcgI restriction site, naturally present in the functional GFP, by substituting 11 bp of the wild-type gene. These substituted base pairs supply two inframe stop codons that terminate

translation, thereby inactivating cassette I. The downstream (3') GFP (cassette II) is inactivated by upstream and downstream truncations, leaving only ~502 bp of GFP not-functional product.

Two homologous recombination products are possible with DR–GFP, a short tract gene conversion (STGC) product or a deletion product. The STGC product results from a noncross-over gene conversion within the limited amount of homology (~812 bp), whereby the 3' GFP sequence acts as a donor of wild-type sequence information to the broken I-SceI GFP gene. The deletion product could result from a conservative recombination event with an associated cross-over, a long tract gene conversion, or from the nonconservative single-strand annealing pathway in which the sequence between the two GFP repeats is degraded. Whereas the STGC event restores an intact GFP gene, a deletional event retains only the 5' portion of the GFP gene that would encode a carboxy-terminal truncation. Only events that restore an intact GFP gene would be scored with the DR–GFP substrate (Pierce et al., 1999).

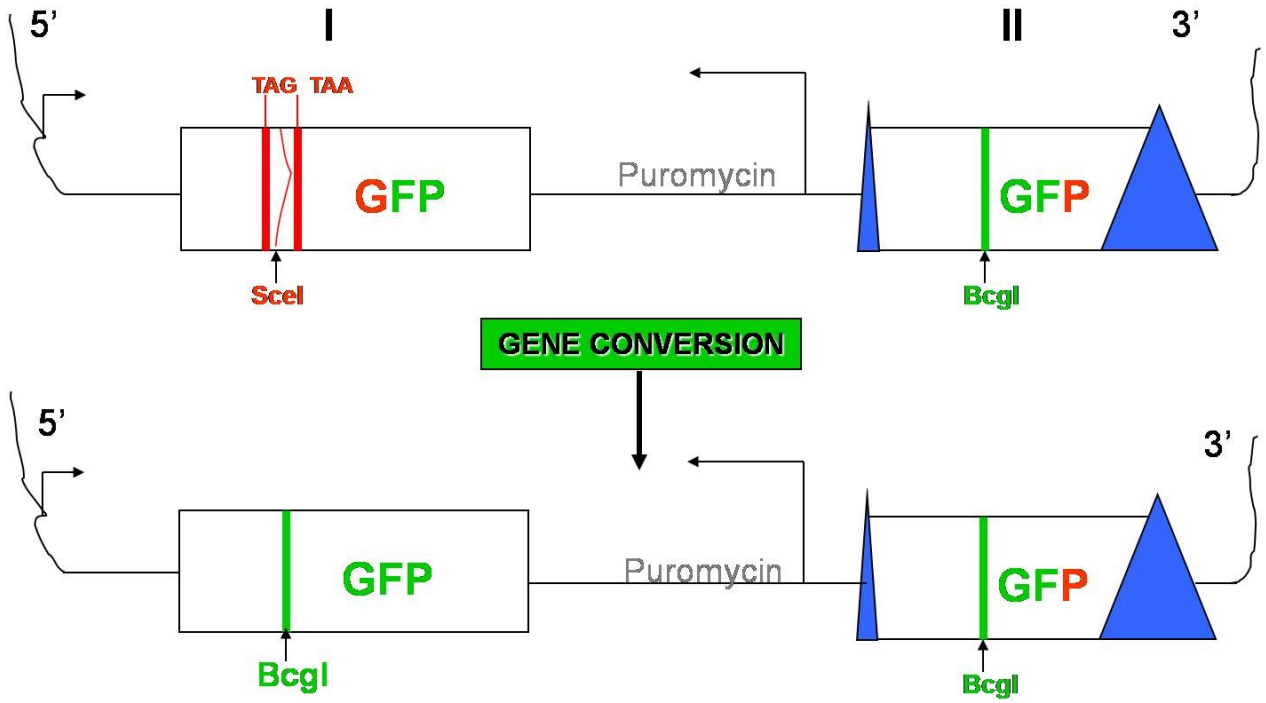


Figure 1: Formation of a functional GFP gene by recombination initiated at a specific double strand break. Cartoon showing the events, initiated at the I-SceI site in the DR-GFP reporter plasmid, leading to a functional GFP gene. Translation termination codons at the I-SceI site in the 5' end GFP cassette are indicated in red. Bcg I site in the homologous wild type GFP sequence in the 3' end cassette is indicated in green. The triangle at the 5' and 3' ends of the 3'-GFP cassette indicate the deletions in the NH and COOH termini of the protein. Short-tract gene conversion of the I-SceI site to BcgI, initiated by the DSB, produces a functional GFP gene.

HeLa cells (human epithelial cells from a cervical carcinoma transformed by human papillomavirus 18 [HPV18]) were stably transfected with the DR-GFP plasmid and selected in the presence of puromycin. Puromycin-resistant pools of cells carrying DR-GFP at various loci were then transiently transfected with a vector expressing I-SceI (Richardson et al., 1998). The resultant DSB induced homologous recombination. GFP^+ cells, derived from I-SceI transfected cell cultures, arise from homology-directed repair of the DSB at the I-SceI site.

Fluorescence-activated cell sorter (FACS) analysis was used to reveal the percentage of cells expressing GFP (**Figure 2A.1**). The structure of the GFP locus was determined by PCR analysis and Southern blot. We used a 3' end primer (**black**) present only in cassette I but not in cassette II. We

distinguished between the recombinant and non-recombinant units by using two 5' primers, the **UnRec** primer (**red**), which amplifies only non-recombinant units, and the **Rec** primer (**green**), which amplifies only recombined units (**Figure 2A.2**). Figure 2 shows GFP^+ cells 48h after I-SceI transfection. DNA and RNA were extracted and subjected to PCR or RT-PCR, normalized for transfection efficiency. As expected, the **Rec** primer amplifies a 436-bp fragment 48 hours after I-SceI transfection (**Figure 2A.3**). This PCR product indicates a wild-type GFP gene generated by gene conversion at the I-SceI site. In contrast to the **UnRec** PCR product (438bp), the **Rec** product was detected only after exposure of cells to ISceI.

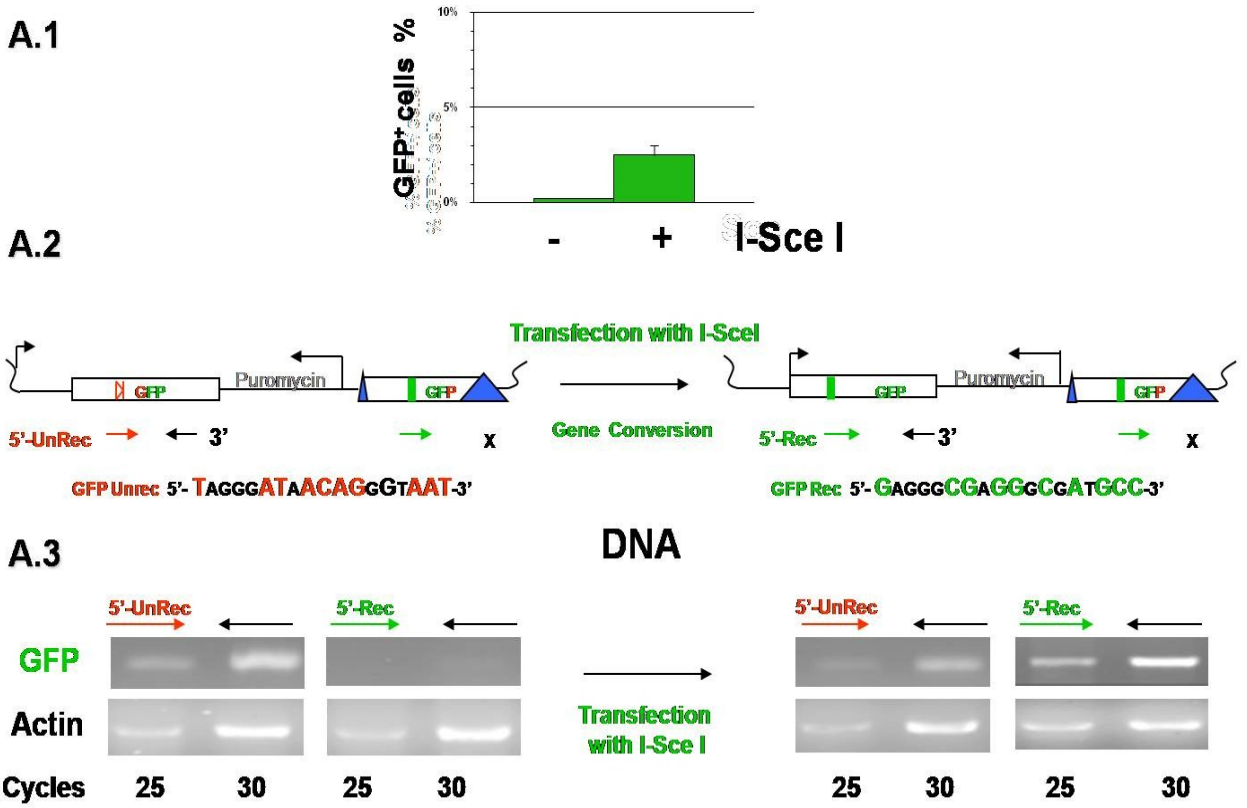


Figure 2A: Recombination assay. HeLa cells were transfected with I-SceI expression vector (or with a control plasmid) and with LacZ as shown in Materials and Methods. 48h after transfection, GFP+ cells were analyzed by FACS **(A.1).** The histogram shows three independent experiments.
PCR on genomic DNA was performed at different cycles (25 or 30) and DNA concentrations, both with the oligos for GFP and for a reference marker (Actin) **(A.3).**
The **red** primer (**5'-UnRec**), centered on I-SceI site, amplifies the non-recombinant units, the **green** primer (**5'-Rec**), centered on BcgI site, amplifies the recombinant units only after the reconstitution of a functional GFP. Both primers produce a ~500bp product when coupled with a common **3'** primer **(black).** The **3'** primer recognizes a sequence found in cassette I but not in cassette II. The different bases in the cassette I are indicated in capital letters **(A.2).**

We then measured the expression of GFP mRNA by reverse transcription (RT)-PCR. Recombined GFP mRNA was detected only in cells transfected with I-SceI (**Figure 2B**). To verify the presence of *bona fide* recombined GFP mRNA, we cleaved the double-stranded cDNA prior to PCR with BcgI enzyme (**Figure 2B, quality check**), which specifically cuts the recombined cassette, in which the I-SceI site was substituted with the BcgI after the DSB repair, ablating amplification with the **rec** primer. Note that the 3' end primer cannot support the amplification of the **rec** primer

unless I-SceI site is substituted with BcgI sequence. BcgI-sensitive PCR is a quality check of PCR recombined products.

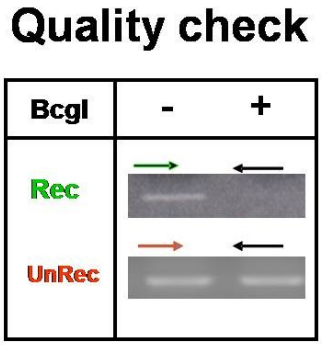


Figure 2B: Recombination assay. RNA analysis was performed on cDNA synthesized from total RNA. RT-PCR analysis was always carried out in the linear range of the reaction. Recombined and un-recombined PCR products were also controlled by treating the cDNA with BcgI before amplification (Quality check). Only the recombinant fragment is cleaved (See Materials and Methods).

II. Methylation reduces GFP expression but not recombination Frequency

To study the epigenetic modifications of the locus before or after homologous-directed repair or DNA double strand break, we used HeLa cells stably transfected with the DR-GFP plasmid. DSBs are efficient substrates of homologous recombination. The low yield of GFP^+ cells after DSB generation raised the possibility that some wild-type GFP recombinants were silenced, possibly by methylation. Accordingly, we asked if inhibiting methylation increased the yield of cells that expressed GFP. HeLa mass culture was transfected with I-SceI expression vector and, after two days, the pool was split and treated with 5μM 5-aza-2'-deoxycytidine (5-AzadC) for 48 hours to block or reverse DNA methylation (Juttermann et al., 1994). FACS analysis showed a significant increase of GFP^+ cells after the 5-AzadC treatment (**Figure 3A**). In the GFP^+ population, only low expressor cells were induced by 5-AzadC.

5-AzadC did not enhance the yield of GFP^+ cells by stimulating homologous recombination. PCR analysis, performed as described in **Figure 2A.3**, showed clearly that treatment with 5-AzadC after I-SceI exposure did not increase the number of GFP recombinant genes (**Figure 3A**). This experiment was repeated with DNA and RNA derived from independent transfections with identical results and on isolated clones. The effects of 5-AzadC on the intensity and on the distribution of GFP fluorescence can be appreciated in the dot-plot shown in **Figure 3** (bottom panel).

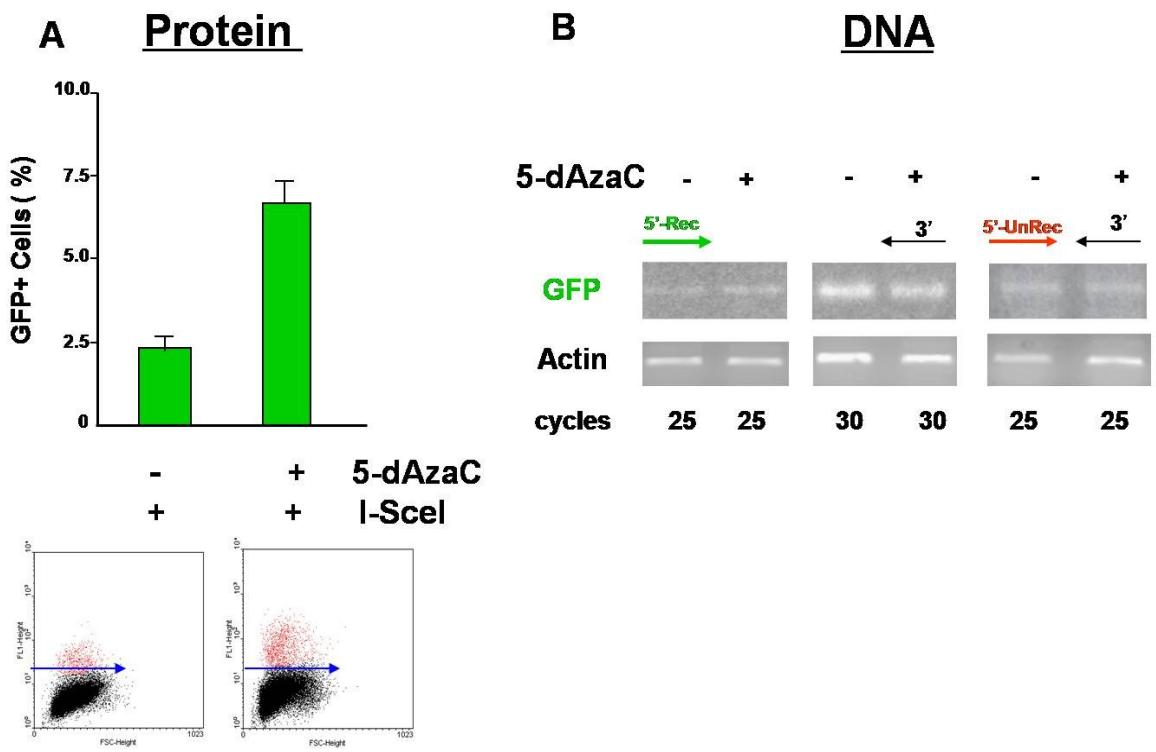
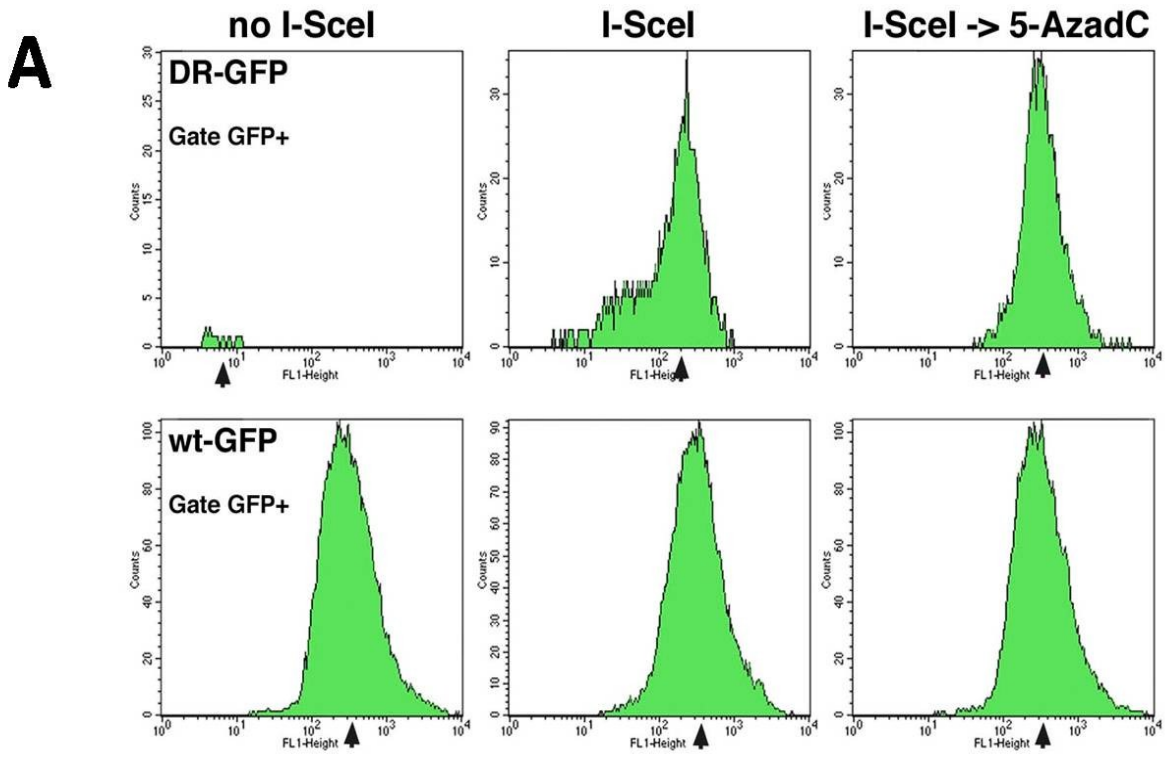


Figure 3: 5-AzadC increases GFP⁺ clones, but does not influence the frequency of recombination. HeLa cells stable transfected with DR-GFP plasmid were transiently transfected with I-SceI plasmid (or with a control plasmid) and treated 12-24-48h (24h after the transfection) with 5 μM 5-AzadC. **A.** GFP expression was analyzed by FACS as % of positive cells (as indicated in the ordinate axis of the histogram). Here is shown only the 24h treatment. The cells without I-SceI (control cells) are not shown. The tranfection efficiency was 70 ± 10%. The bottom panel shows the dot-plot from a representative experiment. **B.** GFP clones were analyzed by PCR with **red (UnRec)** or **green (Rec)** primers coupled with the common **3' (black)** and with βActin-L /βActin-R primers as control.

We considered the possibility that inhibition of recombinant GFP expression was not induced by homology-directed repair, but resulted instead from subsequent transgene silencing, often observed in cultured cells (Pikaart et al., 1998). Accordingly, we monitored the expression of a wild-type GFP transgene driven by a CMV promoter (wild-type GFP) in cells transfected with the I-SceI vector and treated with 5-AzadC, as described in **Figure 3**. **Figure 4A** shows that in contrast to the expression of recombinant GFP, which is bimodal in distribution, wild-type GFP expression is unimodal. Furthermore, unlike recombinant GFP, wild-type GFP expression was not enhanced by 5-AzadC.

To monitor the timing of silencing of recombinant GFP genes and to visualize the effect(s) of 5-AzadC, we separated high- (HR-H) and low-expressing (HR-L) cells, as shown in **Figure 4B**. The separated cells were grown for the times indicated in figures, and parallel cultures were treated with 5-AzadC for 24 h. GFP expression was monitored by FACS.

Figure 4C shows that: (1) only the HR-L fraction was silenced with time; (2) silencing was rapid and reached a plateau two weeks after I-SceI transfection; (3) 5-AzadC stimulated GFP expression in the HR-L population at all time points tested but did not affect expression of the HR-H population. Wild-type GFP expression declined only slightly during the two-week period tested.



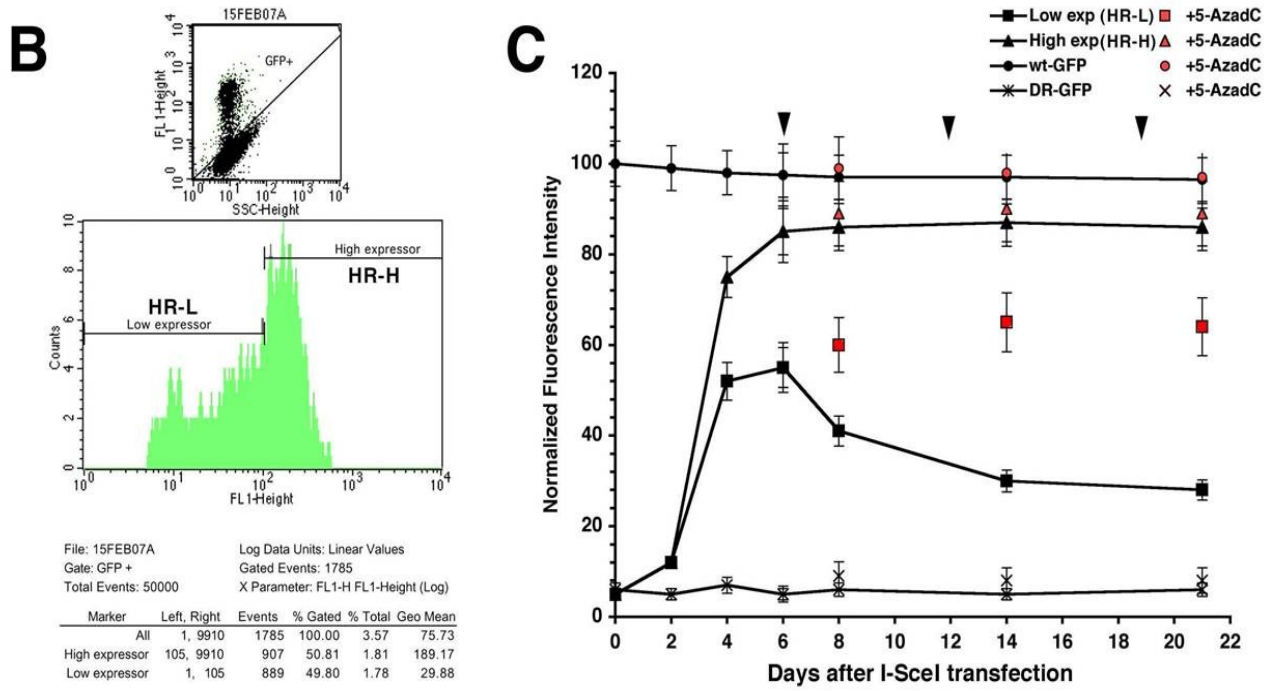
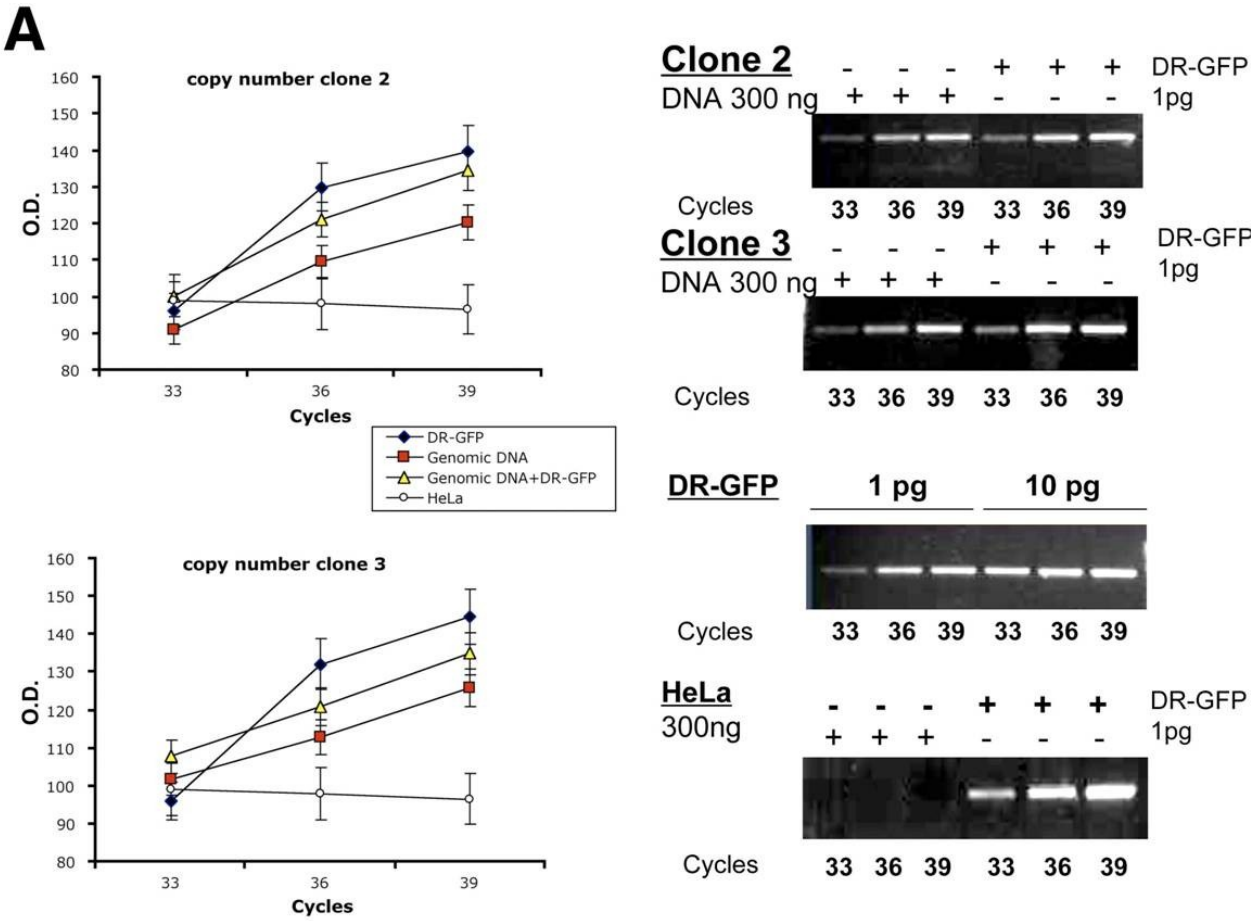


Figure 4: GFP Silencing in Recombinant Clones. Hela cells carrying the DR-GFP or wild-type GFP plasmids were transfected with I-SceI and a control vector, as described in Materials and Methods. **A.** shown is the fluorescence analysis of cell lines stably expressing wild-type GFP. Treatment with 5-AzadC was carried out as described in Figure 3. The cells were analyzed 7d following 5-AzadC treatment. The GFP$^+$ gate includes 95% of GFP-expressing cells. The mean fluorescence (arrow) and the % of GFP$^+$ cells were: (1) DR-GFP (no-I-SceI) 7.4 and 0.4 %; (I SceI) 203.56 and 4.7%; and (I-SceI + 5-AzadC) 335.31 and 7.6%, respectively and (2) wild-type GFP (no-I-SceI) 357.51 and 95.4 %; (I-SceI) 361.12 and 96.2%; and (I-SceI + 5-AzadC) 389.68 and 89.6%, respectively. **B.** Hela cells carrying the DR-GFP plasmid were transfected with I SceI and a control vector. GFP$^+$ cells were sorted by FACS and divided in two pools 4d after transfection: HR-L and HR-H GFP expressors. The upper panel shows the gate used to select GFP$^+$ cells. The lower panel shows only GFP$^+$ cells and the gates used for sorting HR-L and HR-H. **C.** GFP fluorescence in the mass culture was monitored by FACS before transfection and 2d following transfection. Sorted cells were monitored at 4, 6, 8, 14, and 21d after transfection. Parallel cultures at 6, 12, and 19d posttransfection were treated for 24 h with 5-AzadC (arrows). 5-AzadC was washed away and GFP fluorescence was determined 24 h later. DR-GFP and wild-type GFP represent cell lines derived from DR-GFP pools transfected with pSVbGal plasmid or from a stable line expressing the wild-type GFP gene, respectively.

III:Effect of the integration site on the expression of the recombinant gene. Analysis of individual DR-GFP clones

The data shown above suggest that recombination products induced by I-SceI cleavage are silenced by methylation. These results were obtained from pools of cells carrying DR-GFP integrated randomly in the genome and did not distinguish among individual clones. For example, integration of the DR-GFP at a euchromatic site may yield unmethylated, active recombinant units, whereas a heterochromatic location may favour methylation and silencing. We therefore asked if the integration locus influenced the expression of GFP recombination products, and by inference, their methylation status.

We isolated several HeLa DR-GFP clones and controlled the insertions by copy number. **Figures 5A and 5B** shows the PCR analysis and the relative quantification by RT-PCR of DNA extracted from clone 2 and from clone 3. The estimated copy number was 1-3 for clone 2 and 3-4 for clone 3.


A
copy number clone 2
O.D.
Cycles
DR-GFP
Genomic DNA
Genomic DNA+DR-GFP
HeLa
copy number clone 3
Clone 2
DNA 300 ng
DR-GFP
1pg
Cycles
33 36 39 33 36 39
Clone 3
DR-GFP
1 pg
10 pg
HeLa
300ng

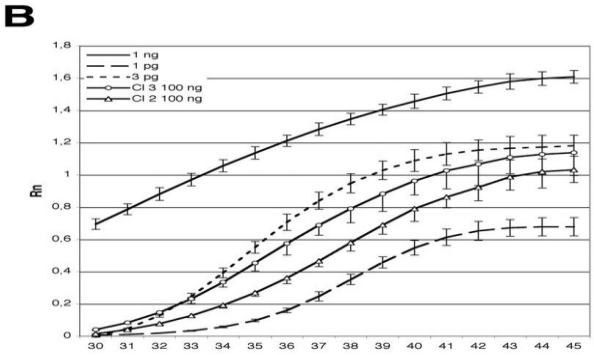

B
1 ng
1 pg
3 pg
Cl 3 100 ng
Cl 2 100 ng
Rn

Figure 5. DR-GFP Copy Number in Clones 2 and 3. A. qPCR analysis of DNA extracted from clones 2 and 3 is shown. Reference curves were generated with 1, 3, and 10 pg of DR-GFP plasmid (2 pg represents approximately 1 copy/haploid genome in 300 ng of genomic DNA for an unique 15-kb DNA sequence). Shown here are curves generated with 1 pg of DR-GFP or with 300 ng of genomic DNA isolated from clones 2 and 3 or control transfected Hela cells (right panels). PCR of genomic DNA mixed with 1 pg of DR-GFP indicates the sensitivity of our assay (left panels). **B.** qPCR carried out on 100 ng of total genomic DNA of clone 2 or 3 is presented. RT-PCR was performed on a 7500 RT-PCR System (Applied Biosystems) using the SYBR Green-detection system. Reference curves were generated for 1, 3, 5, 10, and 1,000 ng of DR-GFP. The mean value and standard deviation of the Rn of six replicates of 1 pg, 3 pg, and 1 ng of DR-GFP and 100 ng of genomic DNA are plotted (Rn: normalized reporter = emission intensity of SYBR Green/emission intensity of passive reference) .The estimated copy number was 1–3 for clone 2 and 3–4 for clone 3.

Hela DR-GFP clones were transfected with I-SceI. **Figure 6A** shows the fluorescent mean intensity as dot plots in red and the fraction of GFP-expressing cells in three individual clones. Both the frequency (ordinate) and the fluorescence intensity (abscissa) segregated in discrete peaks. The GFP^+ clones were high (clone1), middle (clone 2) and low (clone 3) expressors, in terms of GFP fluorescence intensity, normalized to the equal number of GFP positive cells.

We repeatedly transfected the individual clones with I-SceI and determined GFP fluorescence intensity after normalization for transfection efficiency. The results, shown in **Figure 6B**, indicate differences in GFP expression from experiment to experiment. Repeated transfection experiments indicated that the range of variability of GFP expression was limited and that each clone was grossly characterized by the high or low range of GFP expression profile. This effect was **independent of I-SceI** and was clone specific, probably due to the specific integration site. In all cases, I-SceI induced GFP expression and these GFP variations are I-SceI-dependent in the range of variability of the integration site.

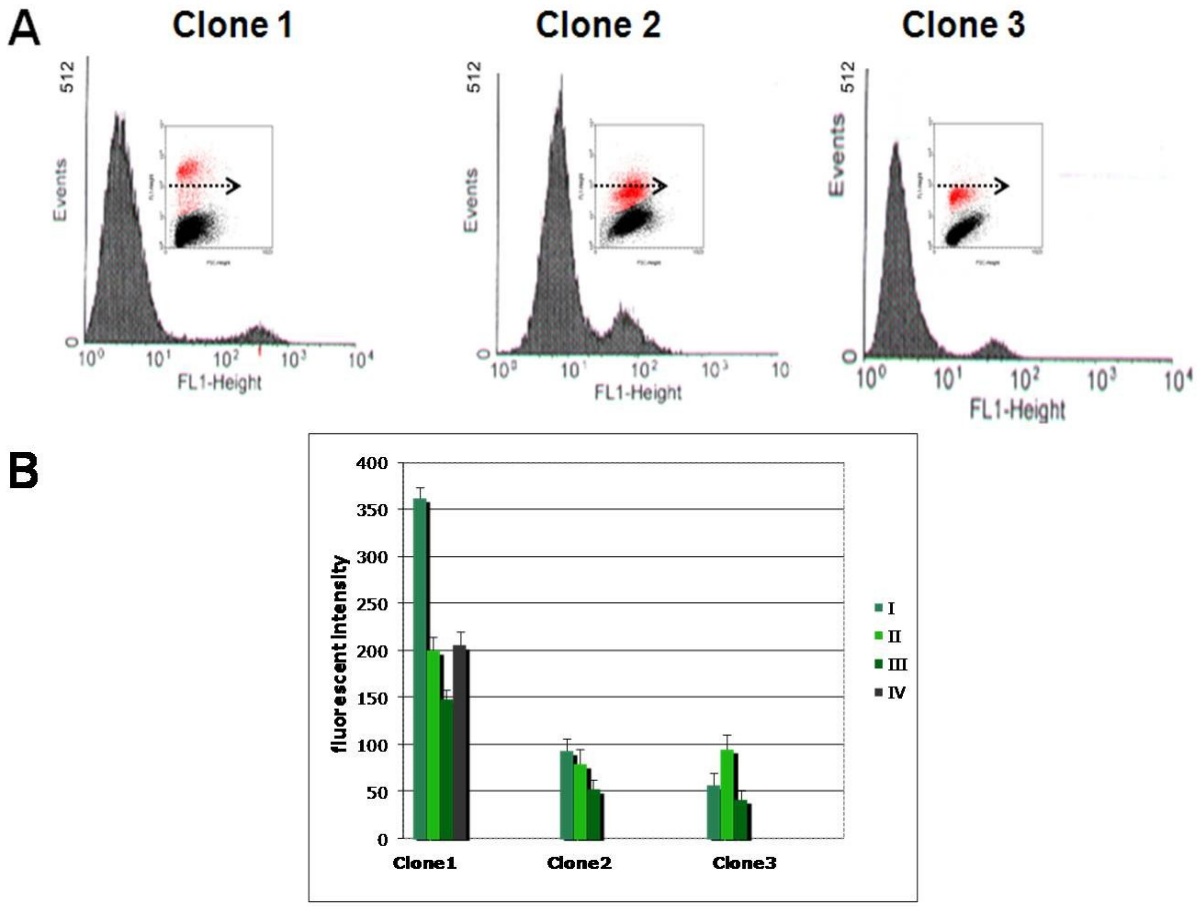


Figure 6. Recombination in Individual DR-GFP Clones. Single clones were isolated from pooled cultures of Hela carrying inactive copies of DR-GFP. The three clones analyzed indicated as 1, 2, and 3 contained a single insertion with one or four copies of DR-GFP as seen by qPCR. These clones were transfected in several independent experiments with I-SceI and pSVbGal vectors and scored for GFP^+ cells. **A.** FACS analysis of three clones 72 h after I-SceI transfection is shown. The ordinate shows the number of cells and the abscissa the intensity of the fluorescence, respectively. The inset shows the dot plot of the bivariate analysis. **B.** The experiment illustrated in (**A**) was repeated several times, and the mean of intensity of fluorescence of GFP^+ cells is shown in the histogram. The roman numerals indicate an individual experiment.

We showed above that inhibition of methylation with 5-AzadC significantly increased the number of GFP^+ cells in a pool of cells carrying DR-GFP at different loci (**Figures 3A and 4C**). We now asked if 5-AzadC affected GFP expression in an individual clone. FACS analysis of clone 3 shows that GFP^+ recombinants appear only after I-SceI exposure (**Figure 7A and 7B**). Transient treatment with 5μM 5-AzadC prior to DSB formation did not increase the number of GFP positive cells (**Figure 7B and 7C**). As was the case with the pooled DR-GFP transfectants, 5-AzadC added after I-SceI transfection significantly enhanced the yield of cells expressing GFP at high levels (**Figure 7D**).

The same experiments were performed with clones 1 and 2 with similar results (data not shown). Note that in clones with a single integration site, 5-AzadC stimulates expression levels to the level of the HR-H average. This effect is not evident in the pool of DR-GFP clones (**Figure 4C**). These results in single clones agree with those obtained from the pool of clones and indicate that methylation following homologous repair of DSBs suppresses expression of a fraction of recombinant GFP genes. Additionally, the bimodal GFP expression distribution characteristic of the mass culture was also seen in clones carrying DR-GFP inserted at a single chromosomal location.

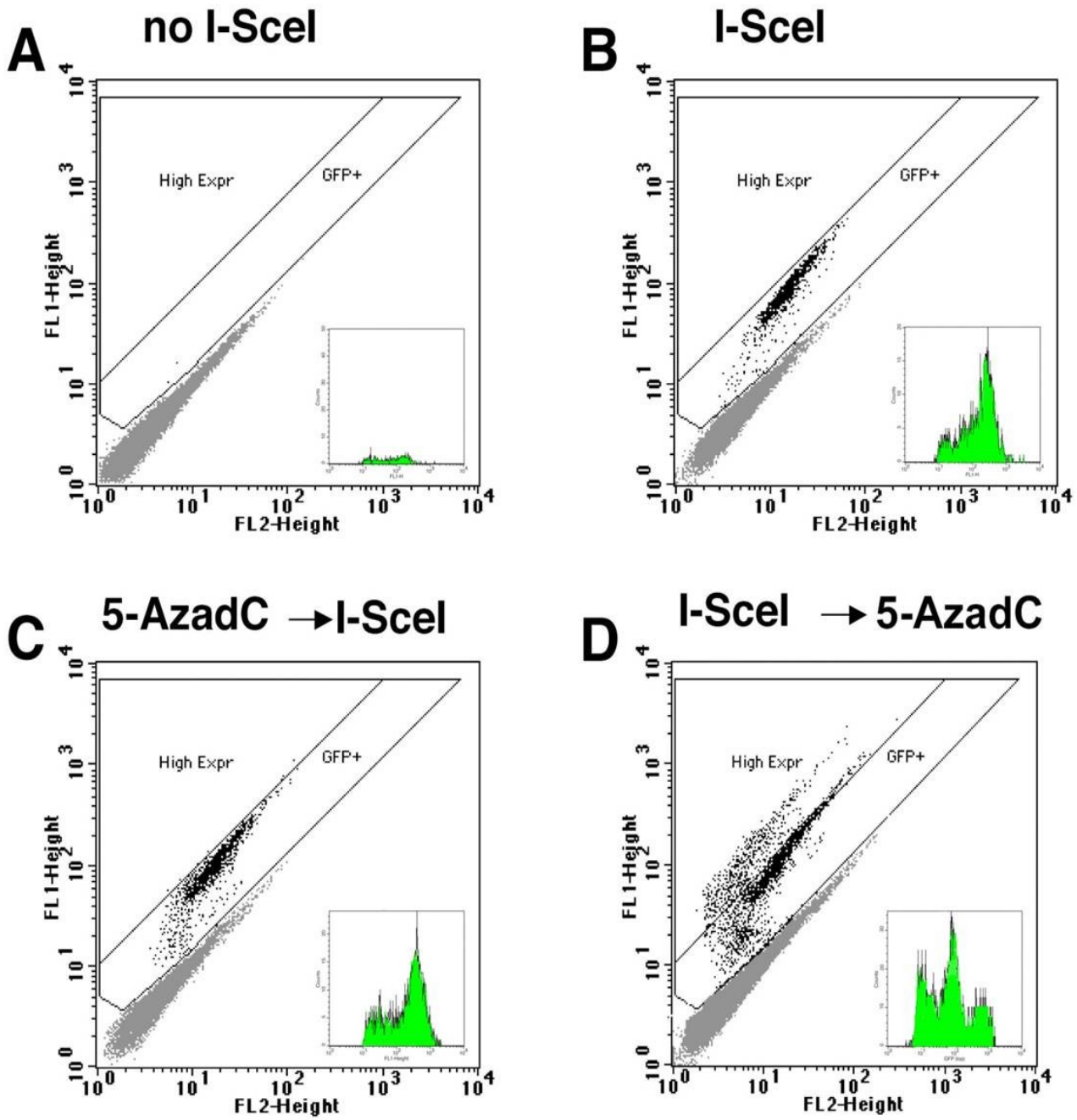


Figure 7. DNA Methylation Is Induced by I-SceI Cleavage/Recombination. Dot plots of GFP expression in clone 3 after I-SceI transfection are shown. The insets show the mean of fluorescence intensity versus number of GFP^+ cells. **A.** Cells transfected with a control plasmid (pSVbGal). **B.** Cells transfected with I-SceI. **C.** Cells treated with 5-AzadC 48 h before I-SceI transfection. **D.** Cells transfected with I-SceI and treated with 5-AzadC 48h later. All FACS analyses were performed 5d after transfection. The effects of 5-AzadC were analyzed also at 2–4 and 6d after transfection, and the results were similar. The efficiency of transfection was 70% ±10%.

IV.DNA Methyltransferase I inhibits the expression of recombinant GFP genes

Stimulation of recombinant GFP gene expression by 5-AzadC suggested that a significant fraction of recombinant genes was silenced by methylation. We confirmed this conclusion in another system in which global methylation was profoundly impaired by inactivation of DNA methyltransferase I (Dnmt1). Dnmt1 is responsible for methylation maintenance in the mouse genome (Li et al., 1992).

We transfected a Dnmt1 -/- ES cell line (Hong et al., 1996) with DR-GFP by electroporation to ensure single integration sites. The pool of puromycin-resistant clones was then transfected with I-SceI and analyzed as described above for HeLa cells. Our results indicate that the frequency of HR was the same in wild-type and Dnmt1 -/- ES cells, as shown by PCR and quantitative (q)PCR (**Figure 8A**).

FACS analysis indicates that the percentage of Dnmt1 -/- cells that expressed GFP at elevated levels was higher than wild-type cells (**Figure 8B and 8C**). Also, similarly to HeLa cells, we found that the low-expressor clones were more represented in mutant cells relative to the wild-type, accounting for the difference in the rate of GFP expression between the wild type and Dnmt1 -/- cells. Finally, treatment with 5-AzadC increased the fraction of wild-type ES high expressors but did not amplify the expression of GFP in Dnmt1 -/- cells (**Figure 8B and 8C**). These data suggest that Dnmt1-dependent methylation silences GFP expression in recombinant clones.

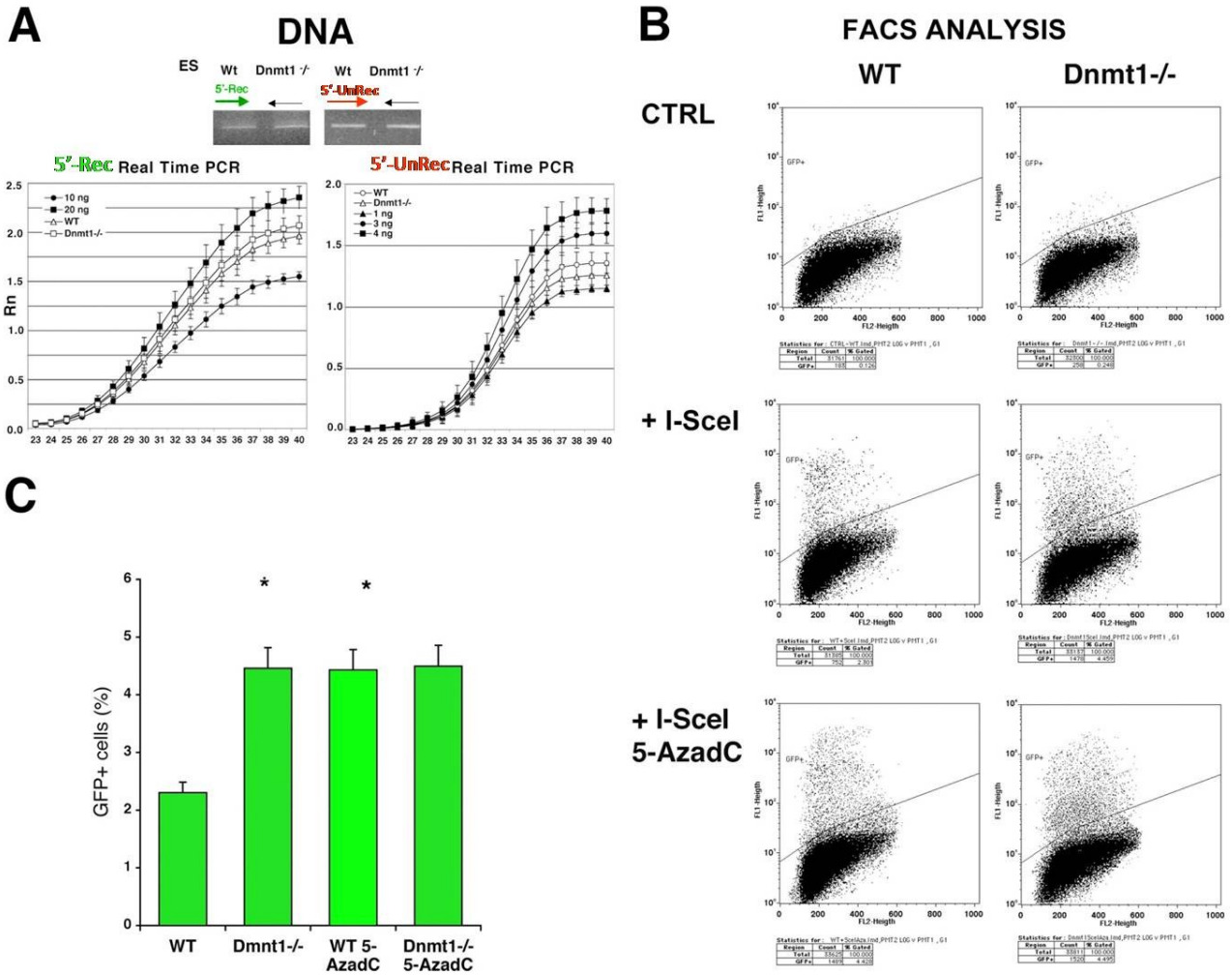


Figure 8. Dnmt1 Inhibits the Expression of Recombinant GFP Genes. Wild-type or Dnmt1 -/- ES cells carrying DR-GFP were transfected with the I-SceI expression vector and PSVbGal, grown 4d, and analyzed for GFP recombination and expression. **A.** Genomic DNA from the two cell lines was PCR-amplified with nonrecombinant (**5'-unrec**) and recombinant (**5'-rec**) primers. The specificity of the products and the linearity of the reactions were controlled as described previously. qPCR of the same samples was carried out as described previously. **B.** FACS analysis of cells transfected with I-SceI is shown. The gating of GFP$^+$ cells was created to exclude up the 99.5% of wild-type untransfected ES cells. The same gating applied to Dnmt1 -/- cells shows a significant increase in the population expressing GFP. Following I-SceI transfection, wt and Dnmt1 -/- ES cells were treated with 5μM 5-AzadC as described before. Treatment with 5-AzadC increased the fraction of cells expressing GFP in wildtype ES but did not enhance the expression of GFP in the Dnmt1 -/- cells. **C.** The histogram shows the fraction of GFP$^+$ cells derived from three experiments. To obtain reliable values of differential GFP fluorescence in ES and Dnmt1 -/- cells, we compared the percentage of GFP$^+$ cells, normalized for the transfection efficiency in six experiments (three in duplicate), with the Wilcoxon Kruskal-Wallis Test, *, $p < 0.012$ versus wild type.

V.CpG methylation before and after repair: Analysis of individual molecules

The data shown above formally prove that methylation following I-SceI expression modifies the expression profile of recombined units. The type of analysis done so far heavily relies on the expression of GFP in recombinant clones. Re-arranged, cleaved and repaired units, which reconstitute I-SceI site, cannot be analyzed by our assay, because these units are not functional.

To get a direct picture of the epigenetic modifications of the I-SceI locus before and after the I-SceI cleavage, we isolated single molecules from the mass culture of ES cells and analyzed directly the methylation profile by bisulfite treatment of genomic DNA before PCR. Bisulfite converts cytosines, but not 5-methylcytosines to thymines. Cytosines detected by direct sequence analysis, therefore, represent methylated residues.

Genomic DNA, extracted from ES cells wild type before and after several I-SceI transfections, was treated with bisulfite and amplified with three different pairs of primers (Materials and Methods). PCR products obtained were cloned and sequenced. **Figure 9A** shows the DNA (+ strand) methylation patterns of all classes found in the mass population of ES cells: (1) Before I-SceI cleavage (uncut); (2) recombinant GFP^+ molecules (HR) isolated by cell sorting for HR-H or HR-L GFP expression; (3) molecules containing a rearranged I-SceI site generated by NHEJ. The methylation status of the HR molecules corresponded with the GFP expression levels of the sorted cells. Relative to the uncut parent, molecules from HR-L cells were heavily methylated, mostly in a segment of approximately 300bp downstream to the DSB. Many of these modified CpGs represent *de novo* methylation sites. In contrast, molecules from HR-H cells were significantly undermethylated, both upstream and downstream to the DSB (**Figure 9A**). The ratio of the two classes was 1:1.

Note that HR repair in this system is a short-tract strand-conversion event, since cassette II is deleted at both upstream and downstream ends. We suggest that the length of the segment showing an altered methylation pattern in the recombinants is limited by the extent of homology between cassettes I and II (~400 bp downstream to the I-SceI/BcgI site).

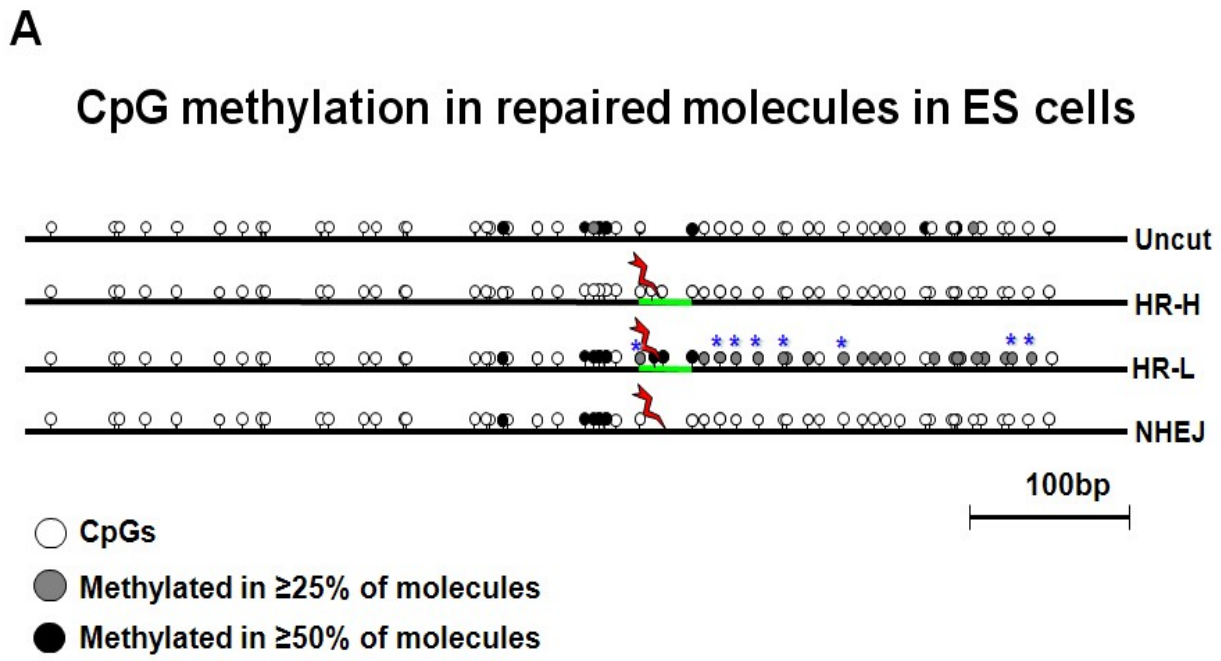


Figure 9A. DNA Methylation in Repaired DNA Molecules. CpG methylation in repaired molecules from ES cells. DNA molecules derived from pooled ES DR-GFP cultures transfected with the I-SceI expression vector or a control plasmid were subjected to bisulfite analysis. The number of molecules in each class was: (1) uncut, 40 from cells transfected with control plasmid; (2) HR-H, 25 homologous recombinant molecules from high expressor cells sorted by FACS (23) or picked randomly from mass culture; (3) HR-L, 30 recombinant molecules from low expressor cells sorted by FACS (28) or picked randomly from mass cultures; (4) molecules rearranged at the I-SceI site (NHEJ). The frequency (%) of each class was derived from several independent experiments with mass culture and fluorescent-sorted cells. HR-H 3 ± 0.5; HR-L 3 ± 1; NHEJ 2 ± 0.3. All CpGs (white circles) flanking the I-SceI site are shown. Gray circles, CpGs methylated in 25% of molecules; black circles, CpGs methylated in 50% of molecules.

The results illustrated in **Figure 8B** suggested that Dnmt1 was responsible for methylation at the DSB. We therefore examined molecules derived from Dnmt1 -/- ES cells before and after exposure to I-SceI. Note that the Dnmt1 -/- mutation increased the expression level of GFP^+ recombinants but not the recombination rate (**Figure 8A**).

As shown in **Figure 9B**, only undermethylated recombinant molecules were generated in Dnmt1 -/- ES cells. We found some methylcytosines on the 3'-end of the I-SceI site in mutant cells, suggesting that Dnmt1 can be substituted in the maintenance of methylation and that repair-coupled methylation can be carried out by Dnmt3a and Dnmt3b. This finding supports our opinion that methylation of the recombined molecules, shown in **Figure 9A**, was catalyzed by Dnmt1.

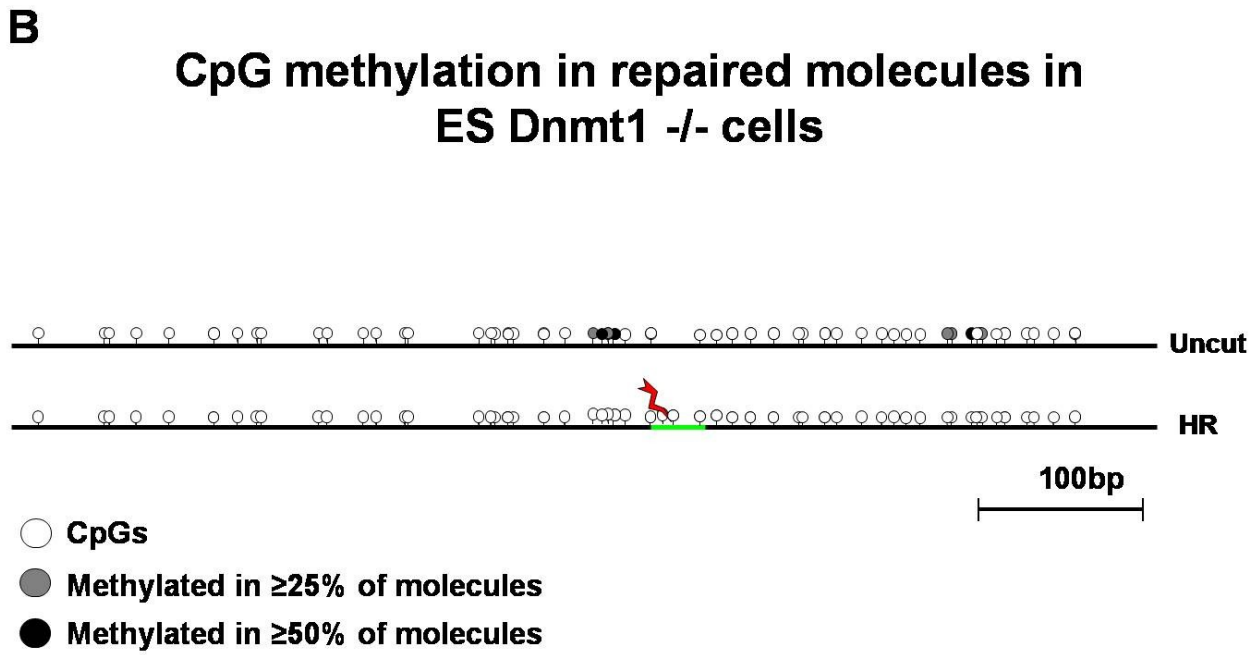


Figure 9B. DNA Methylation in Repaired DNA Molecules. CpG methylation in repaired molecules from ES Dnmt1 -/- cells is shown. DNA molecules were isolated from ES Dnmt1 -/- cells carrying DR-GFP, 16 from control cells and 40 from cells exposed to I-SceI. The frequency of GFP^+ cells was 5 ± 1. All CpGs (white circles) flanking the I-SceI site are shown. Gray circles, CpGs methylated in 25% of molecules; black circles, CpGs methylated in 50% of molecules.

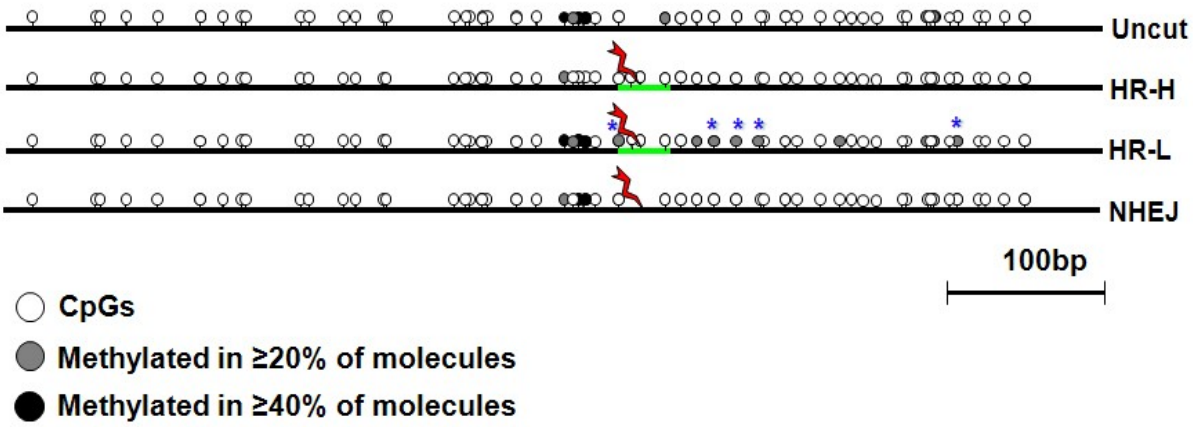
We then asked if the methylation changes following recombination in ES cells could be seen in the human HeLa cell line. **Figure 9C** and **Figure 9D** show the results of our analysis after the bisulfite treatment of genomic DNA extracted from the mass culture and from three individual clones as seen in ES cells.

Figure 9C shows that untreated Hela cell DNA was relatively undermethylated compared to ES cell DNA. Nevertheless, the fraction of hypermethylated HR-L cells as well as the frequency, profile, and length of the segment containing *de novo* methylated CpGs in HeLa cells was similar to that observed in mouse ES cells.

Recombinant molecules derived from individual clones exposed to I-SceI were likewise hypomethylated and hypermethylated in a 1:1 ratio, similar to those isolated from the pool of clones (clones 1, 2, and 3 in **Figure 9D**). These data shown are for the (+) strand, and were confirmed for the (-) DNA strand (data not shown).

C

CpG methylation in repaired molecules in HeLa cells



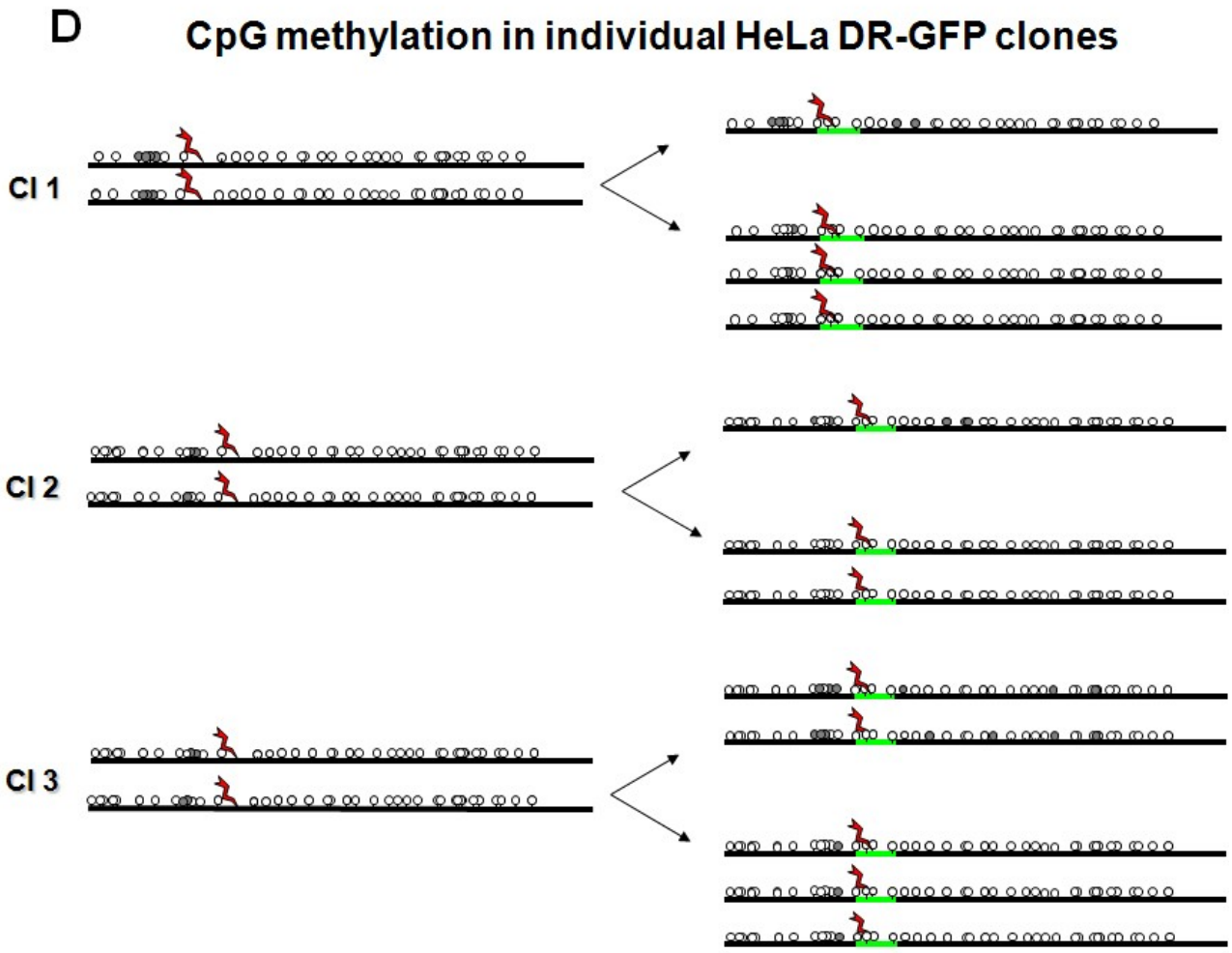


Figure 9C and D. DNA Methylation in Repaired DNA Molecules. C. CpG methylation in repaired molecules from Hela cells is shown. DNA molecules, derived from pooled Hela DR-GFP cultures transfected with the I-SceI expression vector or a control plasmid, were subjected to bisulfite analysis . The number of molecules in each class was : (1) 25 molecules from cells transfected with control plasmid; (2) 20 recombinant molecules from low expressor cells sorted by FACS; (3) 15 recombinant molecules from high expressor cells sorted by FACS; (4) six molecules rearranged at the I-SceI site (NHEJ). The frequency (%) of each class was derived from several independent experiments with mass culture and fluorescence sorted cells. HR-H 2 ± 0.5; HR-L 2 ± 1; NHEJ 2 ± 0.4. **D.** CpG methylation in repaired molecules derived from individual HeLa DR-GFP clones is shown. DNA molecules were derived from clones 1, 2, and 3 of Figure 6. DNA was isolated and subjected to bisulfite analysis. Shown on the left are nonrecombinant molecules amplified with the 5'-unrec primer (see Figure 2). Shown on the right are recombinant DNA molecules isolated from cells transfected for 4d with the I-SceI expression vector and selected for GFP expression. The arrows indicate hypermethylated DNA from clones expressing GFP at low levels and hypomethylated DNA from high GFP expressors. All CpGs (white circles) flanking the I–SceI site are shown. Gray circles, molecules methylated in 20% of molecules; black circles, molecules methylated in 40% of molecules.

The recombination event that generates the appearance of GFP^+ cells is a gene conversion phenomenon that affects a small region; it is caused by a DSB and copies the BcgI site of the cassette II into the I-SceI site of the cassette I. Later, we asked whether the repair process altered the methylation pattern of the cassette II and if this methylation was transferred, at least in some cases, to the gene recombined.

Figure 9E shows the methylated CpG dinucleotides of cassette II from ES and HeLa cells. In ES cells, this segment is more extensively and heavily methylated than the cassette I, and this profile does not change after exposure to I-SceI. On the other hand, the cassette II of Hela cells is ipomethylated than the cassette I, both in the control cells and in those transfected with I-SceI. It should be noted that the cassette II was also methylated in GFP^+ cells sorted by FACS. These data clearly show that changes in methylation profile of the cassette I after the DSB is not dictated by the methylation state of cassette II. Conversely, recombination with cassette I does not influence methylation pattern of cassette II.

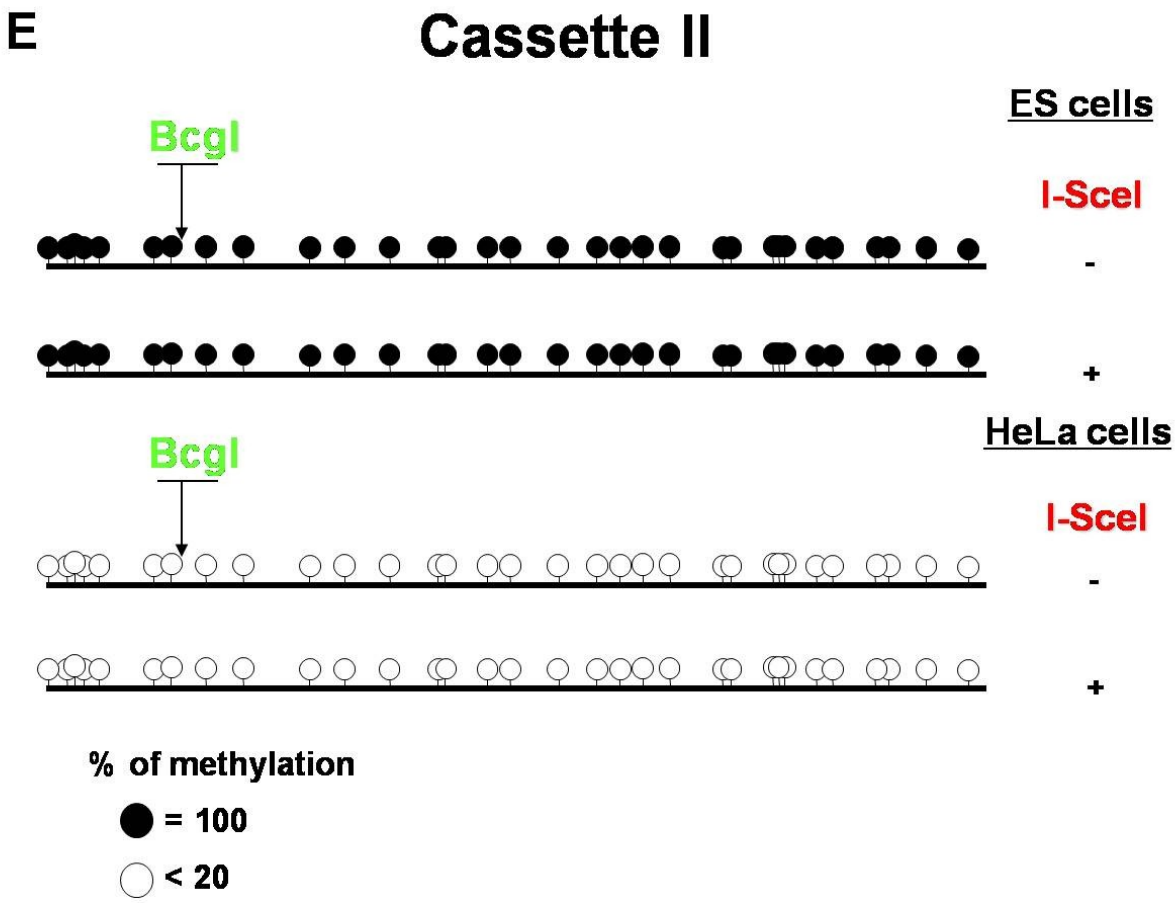


Figure 9E. DNA Methylation in Repaired DNA Molecules. Methylation of GFP cassette II is not influenced by recombination. DNA methylation pattern of cassette II in Hela DR-GFP (25 molecules) and ES DRGFP (30 molecules) cells after transfection with I-SceI or before transfection (7 molecules) is shown. The methylation pattern of cassette II was identical in FACS sorted ES or Hela cells. The molecules analyzed both in ES and Hela cells derived from at least five independent bisulfite reactions and ten independent PCRs for each group: (1) PSVbGal transfected cells; (2) GFP- and (3) GFP^+ high, and (4) low expressors from I-SceI transfected cells.

To get a more defined picture of the distribution of methylated CpGs in the area surrounding the I-SceI site in recombinant and parental GFP molecules, we divided the GFP segment in two regions centered on the I-SceI site: (1) a segment spanning -500 to -51 and (2) a segment at -50 to +420 relative to I-SceI site, respectively.

Figure 10 shows the distribution of methylated CpGs, grouped in three classes containing 0%–1%, 1.1%–6.5%, and 6.6%–50% for ES cells, and 0%–1%, 1.1%–3%, and 3.1%–25 % for HeLa cells, of methylated sites in these segments before or after HR. The distribution is Gaussian before I-SceI exposure in both GFP segments. After DSB and repair, only the segment located at +50 to +420, shows a bimodal distribution ($p < 0.001$) of methylated CpGs in Hela and ES cells. This pattern strikingly recalls the bimodal distribution in the pattern of GFP expression found following DSB-induced repair (**Figures 4 and 7**).

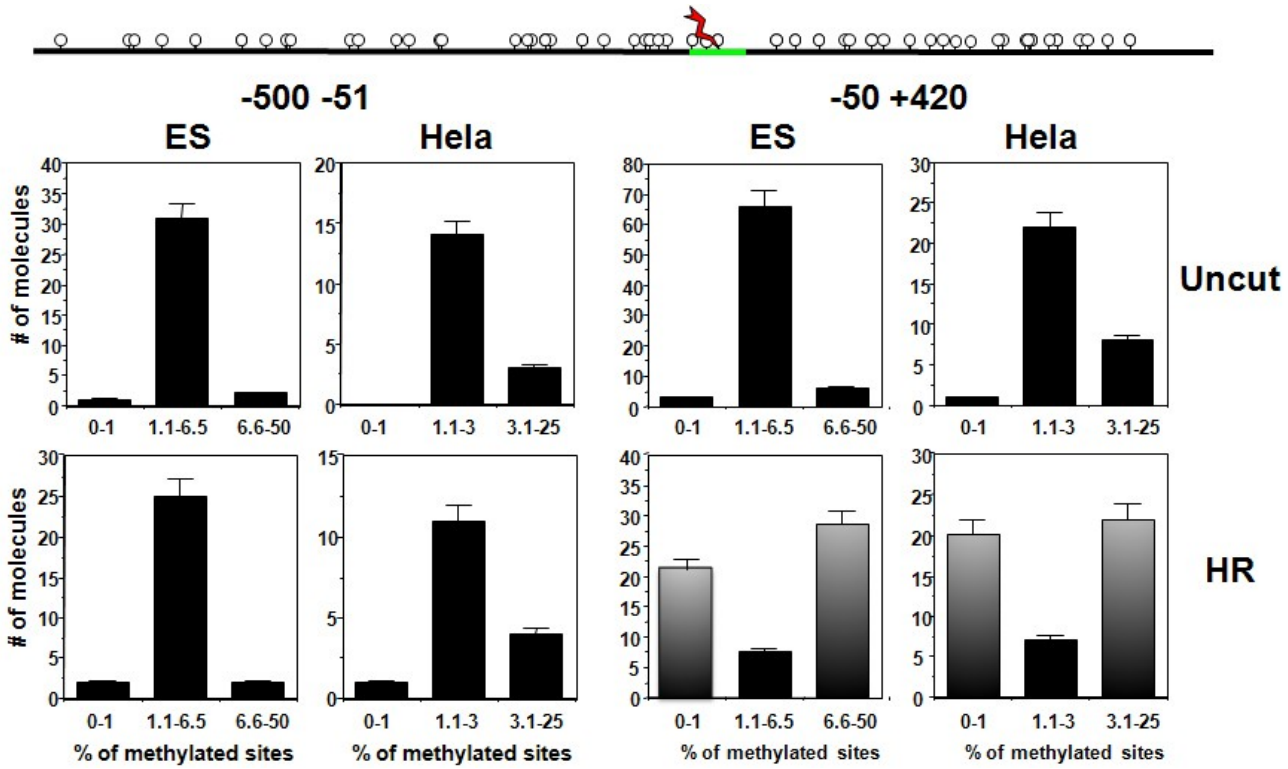


Figure 10. Distribution of Methylated CpGs before and after DSB Repair by Homologous Recombination. Statistical analysis of CpG methylation following a DSB. Cassette I was arbitrarily divided in two segments located at -500 to -51 and -50 to +420 bp relative to the I-SceI site. Methylation was measured as percent of methylated CpGs in each molecule relative to all CpGs present in the segment. GFP molecules were arbitrarily divided in three classes: unmethylated (0%–1% methylated sites), methylated (1.1%–6.5% in ES and 1.1%–3% in Hela), and hypermethylated 6.5%–50% in ES and 3.1%–25% in Hela). The methylated class contains all molecules between ±1 standard deviation. After DSB and homologous recombination, the normal (Gaussian) distribution of methylated sites change to bimodal distribution in both cell lines in the segment downstream to the break ($p < 0.001$ Shapiro-Wilk test).

The data shown in **Figures 6** and **7** summarize the statistical analysis of GFP DNA methylation before and after recombination. However, these data do not reveal the impact of recombination on the methylation pattern of individual GFP molecules.

To visualize changes in individual molecules, we performed ClustalW analysis on the complete collection of GFP molecules. The difference in DNA sequence between recombinant and nonrecombinant molecules may obscure changes due to methylation. To eliminate this problem and to better assess the impact of recombination on *de novo* methylation, we converted the I-SceI site into a BcgI restriction site in all nonrecombinant

sequences and repeated the ClustalW analysis on the total pool of sequences. The molecules now are identical in sequence and differ only in methylated CpGs.

ClustalW analysis of these molecules shows the methylation profiles and the degree of similarity among different molecules. Sequences containing the same methylated CpGs are clustered in branches of the dendrogram. Recombination profoundly altered the methylation pattern of GFP molecules in both wild type and Dnmt1 -/- ES cells. Before recombination the methylation patterns of ES cells and Dnmt1 -/- cells are essentially identical. After recombination, two methylated populations appear in ES cells, whereas Dnmt1 -/- cells yield only undermethylated products. In Hela cells, there are more classes, but the segregation is the same as found in ES cells (HR-L and HR-H).

The simplest interpretation of these data is that methylation is largely random in the culture but that there are preferred sites. Thus pre-existing patterns (before DSB-recombination) can be identified. After recombination, the old pattern is erased in half of the molecules, the high-expressors, or significantly modified in the other half, the low-expressors (see Cuozzo et al., 2007). Dnmt1 is essential for this modification.

VI:Dnmt1 Is Associated with Recombinant Chromatin

The data show so far that DSB repair by HR with consequent gene conversion is associated with significant methylation pattern changes in the area of the DSB. Furthermore, this methylation requires the activity of Dnmt1.

To find the molecular link between recombination and DNA methylation, we asked if Dnmt1 was associated with GFP DNA in the chromatin of cells exposed to I-SceI. Transfected Hela cells were treated with 1µM 5-AzadC and fragmented chromatin was precipitated with

specific antibodies to Dnmt1. Under these conditions, incorporated 5-AzadC ‘‘freezes’’ Dnmt1 on the DNA and amplifies the Dnmt1 signal (Juttermann et al., 1994; Schermelleh et al., 2005).

Figure 11A shows that Dnmt1 is specifically recruited to chromatin regions carrying recombined GFP DNA. It should be noted that nonrecombinant sequences are present in large excess relative to recombined GFP DNA in input chromatin DNA. The specificity of the assay is shown by the presence of Dnmt1 on chromatin of DNA segments heavily methylated in Hela cells (the MGMT and p16 genes) (**Figure 11B**), by the absence of the Dnmt1 signal with nonspecific antibodies (**Figure 11B**) and by the absence of signal with actin primers (**Figure 11A**, lower panel).

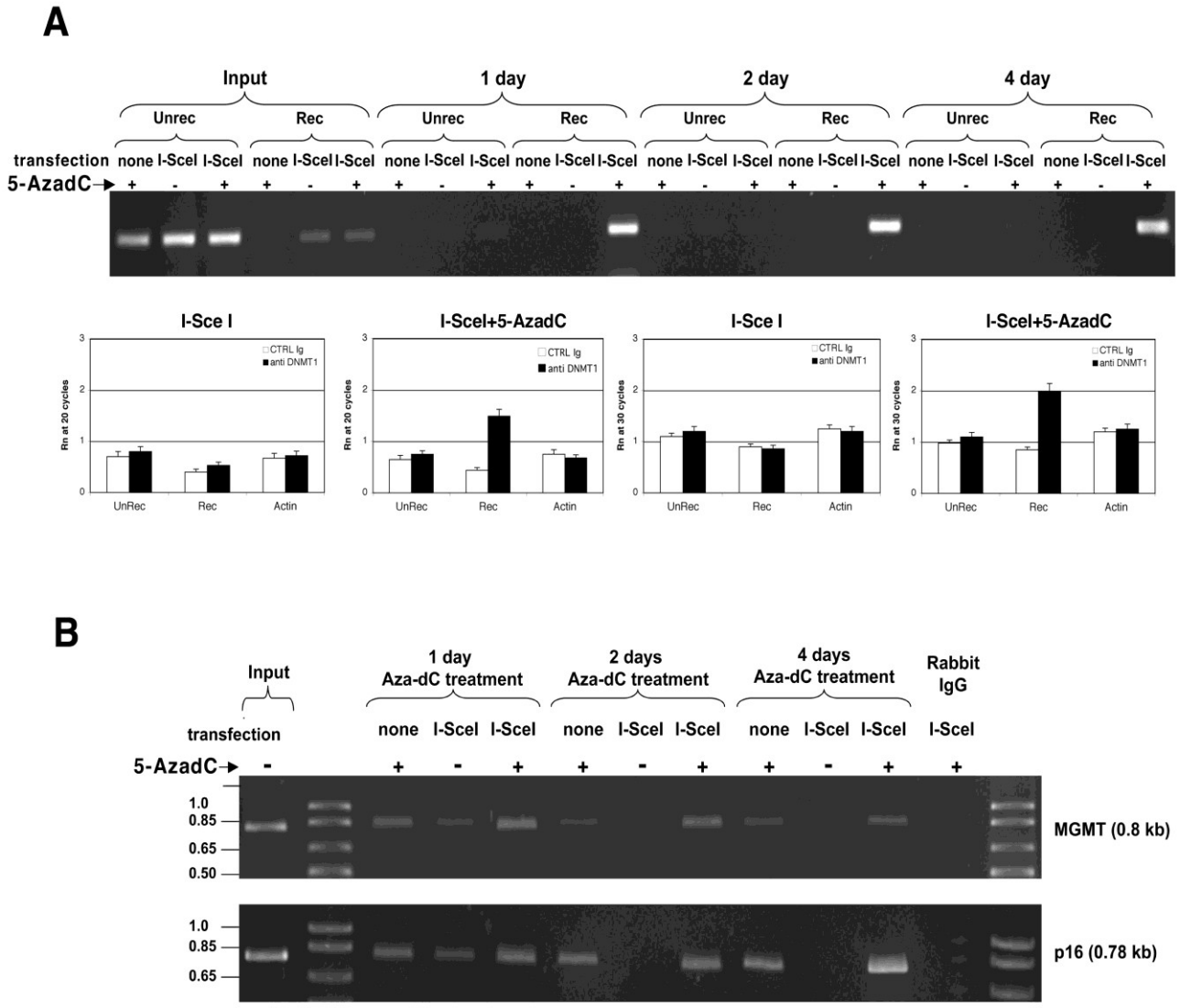


Figure 11. Dnmt1 Selectively Binds Recombinant GFP Chromatin. Hela cells carrying DR-GFP were transfected with I-SceI and treated 24 h later with 1 μM 5-AzadC for 1, 2, and 4d. Chromatin immunoprecipitation (ChIp) was carried out as described in Materials and Methods. **A.**PCR of immunoprecipitated DNA with antibodies to Dnmt1 is shown. None indicates chromatin derived from cells transfected with control plasmid, (-) or (+) indicates the treatment with 5-AzadC. Rec, unrec indicate the primers used for amplification. The lower panel shows the statistical analysis of PCR reactions carried out at 25 and 30 cycles when the reactions with the three sets of primers were in the linear range. Immunoprecipitations were carried out with nonspecific immunoglobulin G (Control immunoglobulin G) and anti-Dnmt1 specific antibodies. The primers used were: (1) unrec; (2) rec; and (3) actin (*, $p < 0.01$ versus control immunoglobulin G). **B.** The same conditions as **A.** MGMT and p16 indicate the primers used for amplification.

HOMOLOGOUS RECOMBINATION AND TRANSCRIPTION

I:Transcription of GFP gene is influenced by DNA methylation induced by homologous repair

The data shown above indicate that following homologous–directed repair of the GFP cassette 2 types of molecules are generated: 1. hypomethylated and 2. hypermethylated in GFP segment repaired containing the I-SceI site. In fact that methylation of a short segment of DNA flanking the DSB (**Figure 9**) is sufficient to silence GFP expression in a significant fraction of cells (HR-L) (**Figures 4C and 9**). Since the CMV promoter and chicken β-actin enhancer that drive GFP expression are located ~1,000 bp from the BcgI/I-SceI site and are insulated from surrounding genomic regions, the link between methylation and silencing is not readily evident.

To explore this question, we asked if methylation inhibited transcription initiation and/or elongation. We performed RT-PCR analysis of RNA with primers derived from the upstream intron (close to the transcription initiation site), from the beginning of the GFP gene, and from the I-SceI (control cells) or BcgI (HR-L and HR-H cells) sites (**Figure 12A**). Since PCR reactions performed with different primers cannot be directly compared, we measured amplification of the PCR signal in a particular region of the gene after 5-AzadC treatment. This value indicates how methylation affects transcription near the promoter and at downstream regions. The results of **Figure 12** suggest that RNA derived from both upstream and downstream regions of the GFP gene was significantly reduced by methylation in the HR-L population. Methylation did not affect RNA synthesis in HR-H or in nonrecombinant (ctrl) clones. Finally, 5-AzadC stimulation was greater in the region of the BcgI site than upstream (**Figure 12A and 12B**). It should be noted that the difference

between 5’ and 3’ end transcript levels may be artificially amplified by the fact that only the 5’ end primers are selective for recombinant RNA, as seen in **Figure 1**. Since sorted cells may contain copies of unrecombined DR-GFP, these units can generate nonrecombinant transcripts, which are not stimulated by 5-AzadC (see ctrl). As a result, the differential levels (- or + 5-AzadC) of 5’ end may appear lower than the 3’ end transcripts. Despite this limitation, we find a significant and reproducible increase of 5’ end transcript by 5-AzadC ($p < 0.01$).

Our data indicate that CMV promoter activity is inhibited by methylation at the DSB and suggest further that elongation may also be hindered by methylation of the repaired segment. We propose that this inhibition is triggered by changes in the chromatin domain that includes the repaired DSB. Nucleosome structure is known to affect both transcription initiation and elongation (Li et al., 2007).

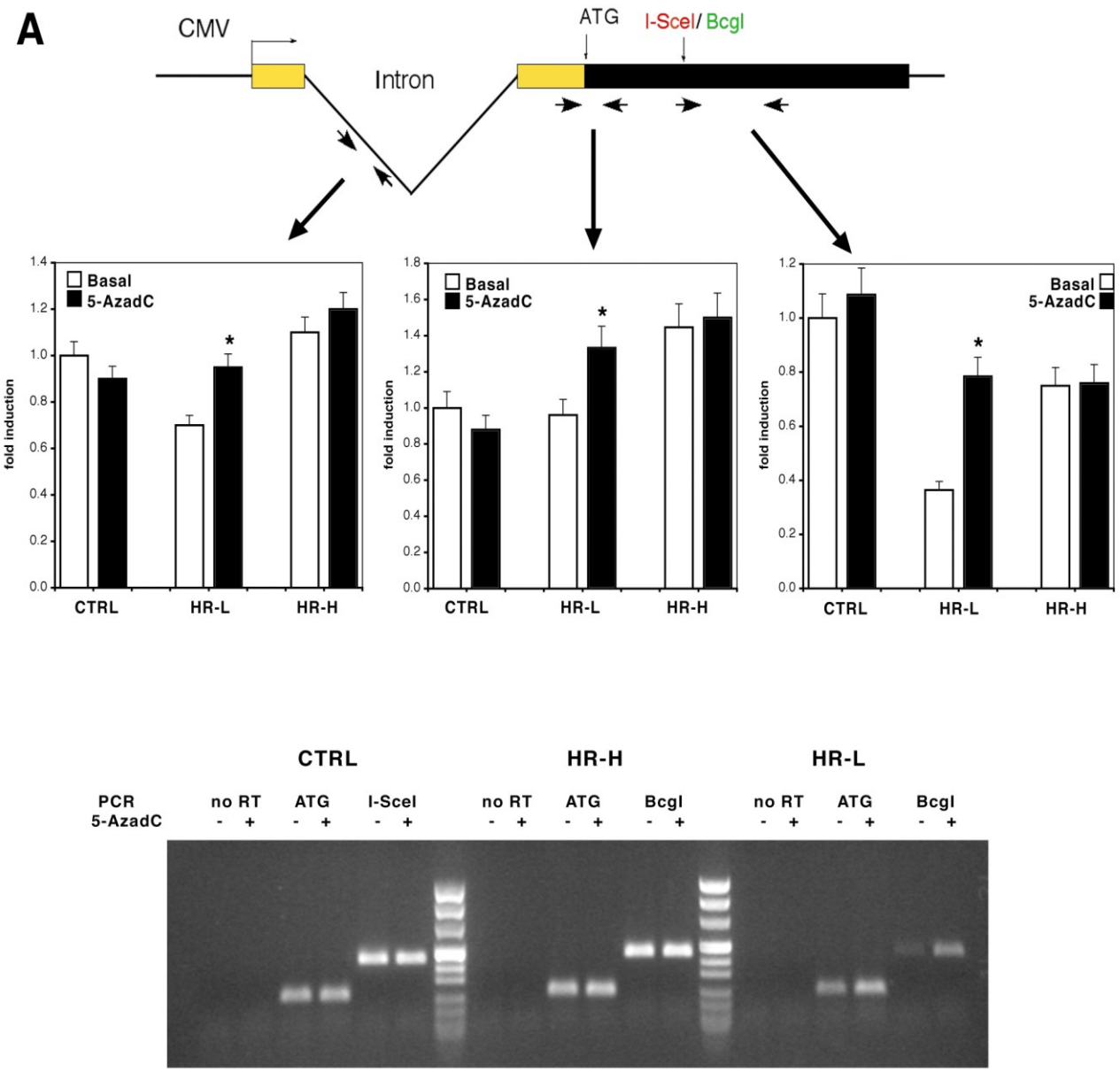


Figure 12. Mapping of GFP Transcription in Recombinant and Nonrecombinant Cells with and without 5-AzadC Treatment. A. A schematic of the DR-GFP transcriptional unit shows the location of the CMV promoter, intron, and GFP coding sequence. Primers used for quantitative RT-PCR and RT-PCR are indicated by arrows. Different sets of primers were derived from the intron (738–757, forward 5′- CGTTACTCCCACAGGTGAGC-3′; 966–948, reverse 5′-CGCCCGTAGCGCTCACAGC-3′), AUG (1,666–1,685, forward 5′-TACAGCTCCTGGGCAACGTG-3′; 1,911–1,892, reverse, 5′-TCCTGCTCCTGGGCTTCTCG-3′), and BcgI/I-SceI (described in Figure 1A) segments of GFP gene. Control (DR-GFP cells transfected with pBluescript), HR-L, and HR-H cells were treated with 40 μM 5-AzadC for 48 h. Total RNA was extracted 24 h later and subjected to quantitative RT-PCR with the primers indicated. The data, derived from three independent cDNAs, are shown as fold induction by 5- AzadC over the basal control, normalized to GADPH and b-actin. The primers used to amplify the control samples were those indicated as I-SceI **unrec** (Figure 1A) **.B.** Shown is RT-PCR with the same cDNAs indicated in **(A)** at 30 cycles.

II.Homologous recombination and methylation induced by homologous repair are dependent on transcription

The data shown above indicate that following or during the homologous repair hypermethylated ad hypomthylated GFP molecules accumulate in 1: 1 ratio. We speculate that during recombination only one strand will be methylated by DNMT1. This intermediate, with half methylated strand, is very difficult to isolate, since it is transient and diluted in a vast majority of unrecombined or NHEJ-repaired GFP DNA.

To address the nature of methylation induced by repair and to find out a possibile mechanism explaining the accumulation of methylated and unmethylated molecules, we treated for a short time, after I-SceI transfection, the cells with low dose of α-amanitin. α-amanitin is a selective RNA polymerase II inhibitor, which binds RNA polymerase II and arrests transcription by stalling the RNA polymerase complex on the DNA.

We treated the HeLa cells 24 h after I-SceI transfection for 6-24 and 48 h with 2,5 μM α-amanitin. Later the cells were washed and cultured for additional 5 days, before analysis. Note that under these conditions we monitor only permanent effects induced by the short term treatment with the drug.

We first determined the rate of recombination by running PCR reactions with specific primers for the recombinant (**Rec**) and for the unrecombinant (**UnRec**) on total DNA extracted from the pool of clones or isolated clones before or after I-SceI transfection. The specificity of the reaction is shown by the selective amplification of recombinant primers only after I-SceI transfection (**Figure 13A**). α-amanitin treatment reduced the rate of recombination by 30% compared to the control (**Figure 13B**).

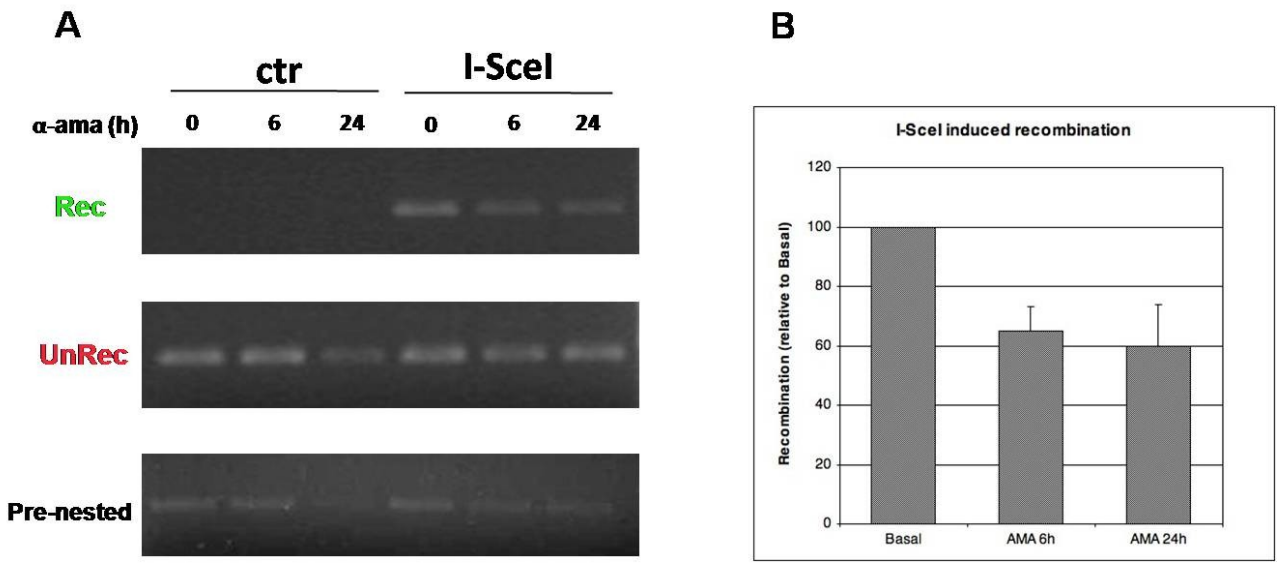


Figure 13. I-SceI- induced recombination after α-amanitin treatment. A. Genomic DNA extracted from HeLa cells transfected with I-SceI expression vector or a control plasmid (ctr) and treated with α-amanitin at indicated times was PCR amplified with **Pre-nested** primers (see Materials and Methods). 1/50 of the first PCR was amplified again with the oligos for Rec and UnRec. All the PCRs were carried out in the linear range of the reaction. **B.** The histogram shows the frequency of I-SceI induced recombination obtained by comparing the densitometric spots of the Rec and UnRec amplifications, derived from a nested PCR on the amplification with Pre-nested primers, normalized on the basal value, i.e. not treated.

Fluorescence analysis revealed a significant change in the fluorescence intensity and in the number of GFP positive cells. **Figure 14A** shows a representative experiment where low and high GFP expressor are shown before or after α-amanitin treatment. The intensity of fluorescence of the high expressors was increased by α-amanitin treatment, whereas the low expressors were reduced in number by the same treatment. Statistical analysis presented in the middle and lower panels (**Figure 14B-C**) shows that in all the experiments performed the high expressors displayed increased GFP signal. These effects were visible in 6 and 24h treated cells. The reduction of GFP positive cells was evident in the lower expressor fraction and most likely was dependent on the reduction of the recombination frequency.

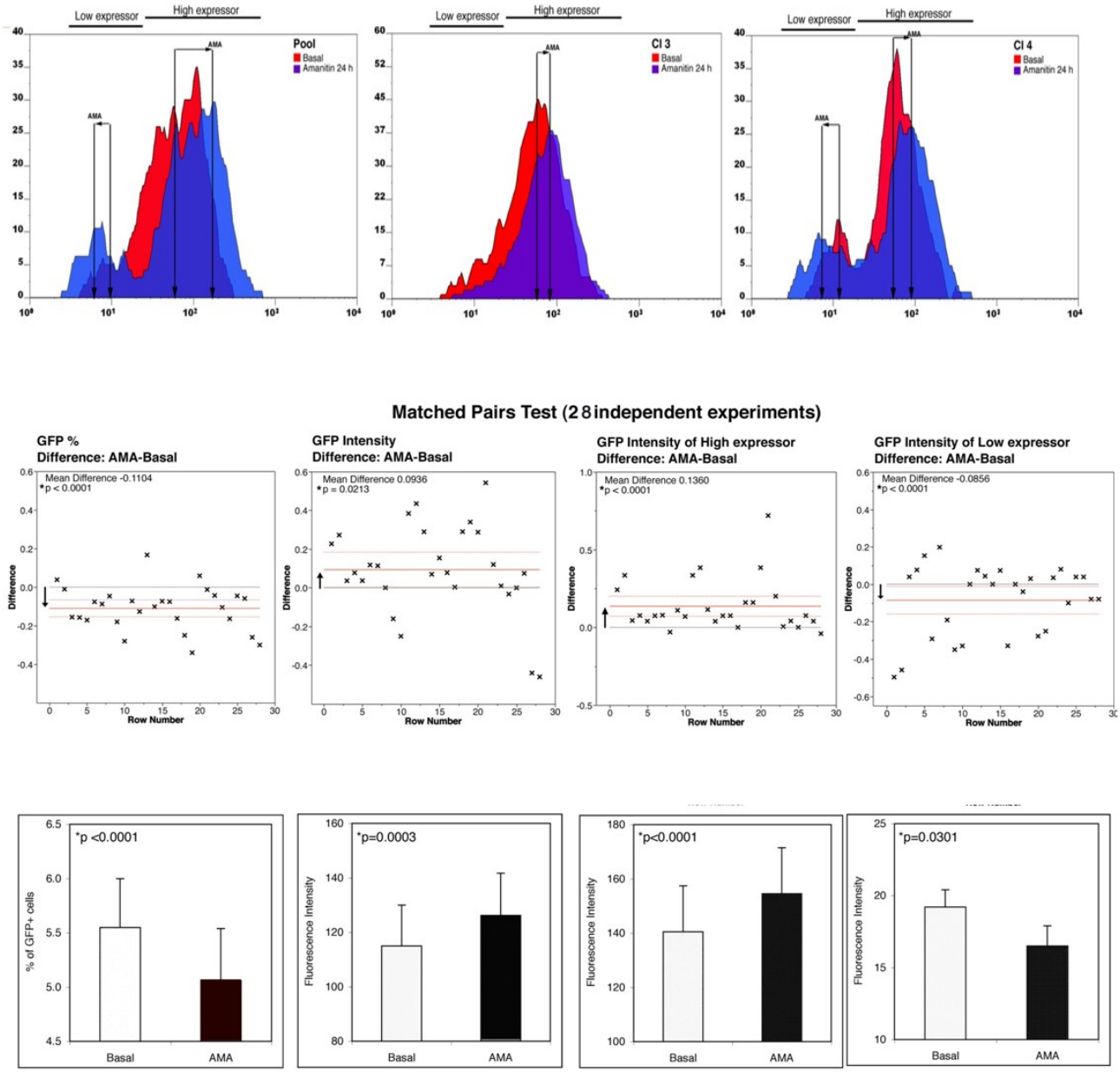


Figure 14. Effects of α-amanitin on repair-induced GFP expression. HeLa cells were trasfected with I-SceI expression vector and treated 24h after with 2,5 μM α-amanitin for 24h. The samples were analyzed 5d later. **A.** FACS plot from a representative experiment shows the distribution of GFP fluorescence in the population of GFP positive cells. RED, blue and violet indicate untreated or 6 or 24 h a-amanitin treated samples, respectively. **B** and **C.** Statistical analysis of GFP fluorescence distribution derived from 28 independent experiments. High expressor and low expressor distribution was determined relative to untreated cells. The median and the control, set to 0, are indicated by a red and a black line, respectively in the middle dot panels **(B)** The statistical significance for each class (p) is is shown on the top of the histograms in the lower panel **(C)**.

To directly visualize the effects of α-amanitin treatment on the methylation status of GFP DNA we carried out bisulfite analysis. The data shown in **Figure 15** indicate that the GFP DNA molecules from the cells subjected to α-amanitin treatment were both hypermethylated and hypomethylated relative to untreated controls. These data suggest that slowing down recombination enhances DNMT-1 methylation on a fraction of GFP molecules. Under these conditions the methylation distribution profile of GFP cells in high (hypomethylated) and low (hypermethylated) cells becomes markedly accentuated.

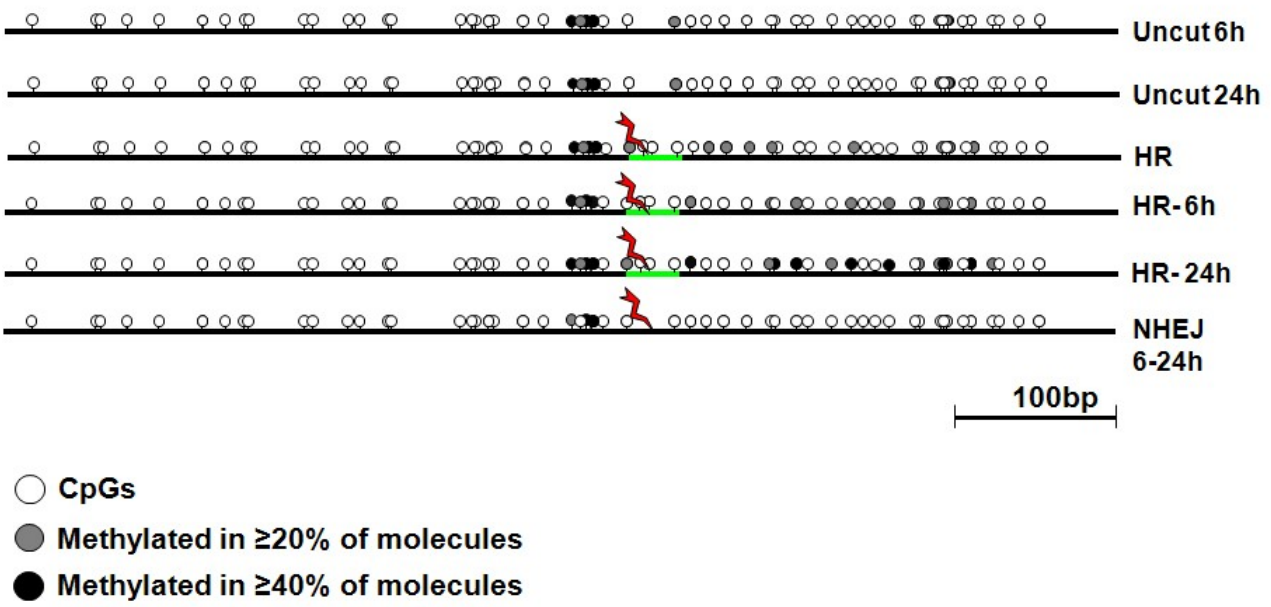


Figure 15. GFP Methylation in Repaired Molecules after α-amanitin treatment. CpG methylation in repaired molecules from HeLa cells is shown. DNA molecules, derived from pooled Hela DR-GFP cultures transfected with the I-SceI expression vector or a control plasmid and treated 24h after, with α-amanitin for 6 or 24h, were subjected to bisulfite analysis (see Matherials and Methods). The number of molecules in each class was : (1) 20 molecules from cells transfected with control plasmid and from cells treated 6h with α-amanitin and (2) 25 molecules from cells treated 24h with α-amanitin ; (3) 15 recombinant molecules trasfected with I-SceI and (4) 15 molecules from cells treated with 6h α-amanitin and (5) 17 molecules from cells treated with 24h α-amanitin; (6) 15 molecules rearranged at the I-SceI site (NHEJ) in 6 and 24h treatment. Gray circles indicate methylation of at least 20% of molecules; black circles indicate methylation in at least 40% of molecules.

DISCUSSION

DOUBLE STRAND BREAK AND DNA METHYLATION

In this work we demonstrate that there is a causal relation between double-strand breaks (DBS), homologous recombination, DNA methylation and gene silencing. We used a fluorescence-based recombination assay to monitor homology-directed repair following the introduction of a site-specific DSB after the expression of the I-SceI plasmid. The recombination system was developed by Jasin et al., in which direct repeats of defective GFP genes recombine to yield a wild-type product by gene conversion.

We found both hypermethylated and hypomethylated recombinant DNA products, in about a 1:1 ratio. The hypermethylated recombinants expressed low levels of GFP compared to the hypomethylated recombinants. Silencing of these GFP genes was due to DNA methylation.

Treatment with the demethylating agent, 5-AzadC, resulted in a significant increase of the number and rate of GFP expression, without changes in rate recombination. In 2 independent cell lines, HeLa and ES cells, methylation did not appear to grossly influence the frequency of recombination, assayed as the fraction of recombinant over non-recombinant DNA fragments. 5-AzadC treatment after exposure to I-SceI induced a large scale DNA demethylation and reactivation of global gene expression (Juttermann et al., 1994).

In this study, the inhibition of DNA methylation selectively stimulated the expression of recombinant GFP relative to non-recombinant GFP gene. Demethylation prior to DSB generation was ineffective. DNA methylation and histone deacetylation act as a layer for epigenetic gene silencing and, in this circumstance, DNA methylation appears to be the dominant layer. We demonstrate also that the pre-I-SceI status of the integration site can be dissociated from the effects following I-SceI

cleavage and recombination, since treatment with the demethylating drug induced GFP expression only if administered after I-SceI transfection, i.e., after DSB.

We translated this notion in another system, where the major DNA methyltransferase activity had been knocked down (ES mouse Dnmt1 -/-). These cells, unable to methylate as wild type, were unable to silence recombinant GFP as the wild type. Murine ES cells lacking Dnmt1 -/- yielded more clones expressing high levels of GFP than wild-type controls. Also, treatment with 5-AzadC stimulated GFP expression in the wild type ES much better than in Dnmt1 -/-.

Silencing of GFP homologous recombinants differed from transgene silencing observed by Pikaart et al., 1998. Expression of a transfected wild-type GFP was only slightly reduced over a three-week period, whereas HR-L clones were silenced during the two weeks following exposure to I-SceI (Figure 4C). Likewise, we saw no silencing of uncleaved or wild-type GFP genes, since their expression was not stimulated by 5-AzadC (Figures 4 and 12).

Single clones did not display the same methylation pattern following I-SceI exposure. Since the I-SceI vector persisted for about one week, cells were subjected to repeated cycles of cleavage and repair. The resultant methylation pattern varied from cell to cell, although preferred modification sites were found (Figure 9).

The methylation patterns generated by gene conversion are not seen in the parental DNA. The analysis of methylated cytosines revealed that the striking difference between all types of GFP DNA molecules (uncleaved, cleaved and re-sealed reconstituting I-SceI site, recombinant, re-arranged) was the exposure to I-SceI. We found that DNA molecules derived from cells expressing I-SceI were methylated at specific locations around the I-SceI site, in the 200 bases flanking the site. Recombinant molecules contained the same methylcytosines found in molecules

exposed to I-SceI, indicating that the DSB, rather than recombination, is indeed linked to methylation. The methylation pattern in recombinant molecules varied considerably from high to no methylation at all, depending on the type and extension of recombination. Molecules unexposed to I-SceI had an average methylation density of 1%–3% (Figure 10). About half of the converted molecules were hypermethylated (7%–50% of methylated sites). The remainder were hypomethylated (0%–1%).

To illustrate the two aspects of DNA methylation, the data of our study on the methylation status of GFP molecules are presented either as percent of methylated CpGs in different GFP segments (Figure 10) or as profiles of methylation, showing which specific CpGs are methylated and the degree of similarity among different molecules (Figure 9). The analysis of single molecules is highly informative, allowing us to identify common and unique methylation profiles between recombinant and nonrecombinant molecules. HR, which occurs stochastically on the nonrecombinant GFP molecules, erases or induces methylation at the DSB and generates the bimodal methylation patterns shown in Figure 10.

Hypomethylated and hypermethylated molecules in ~1:1 ratio were also found in recombinants derived from single clones, indicating that hypomethylated or hypermethylated GFP sequences did not segregate in individual clones. Thus, both HR-L and HR-H recombinants are generated in both mass cultures or in clones containing a single integration site (Figures 4 and 7). Note that the GFP expression levels of wildtype GFP transfectants show a Gaussian distribution, totally unlike the bimodal distribution observed after recombination. This finding implies that DNA repair generates one hypermethylated and one hypomethylated strand and replication yields double-stranded molecules that retain the methylation pattern of the precursor strand (see Cuozzo et al., 2007).

GFP^+ cells are generated from short-tract gene conversion of the I-SceI site on GFP cassette I to the BcgI site on GFP cassette II. Cassette II, which

provides the template strand, has only 395 bp of homology downstream to the I-SceI site of cassette I. This may explain why the altered methylation pattern in the recombinant molecules is limited to a small region flanking the DSB. The methylation status of cassette II did not dictate the modification of the recombined GFP gene. Cassette II was hypermethylated in ES and hypomethylated in Hela cells, whereas the methylation patterns of the repaired DNA of cassette I were similar in both cell types. Furthermore, HR of cassette I did not change the methylation pattern of cassette II (Figure 9).

REPAIR AND METHYLATION

DNA cleaved at the DSB in cassette I is filled in with newly synthesized DNA templated by cassette II, with gene conversion at the I-SceI site, or by a sister chromatid exchange without gene conversion. We propose that only one of the newly synthesized strands is methylated by Dnmt1, which is recruited to the DSB by proliferating cell nuclear antigen (Chuang et al., 1997) and marks the HR molecules (Figure 11). Replication of the hemimethylated intermediate yields hypermethylated and hypomethylated DNA products in a 1:1 ratio. In the case of gene conversion, the length of the hyper- or hypomethylated track does not exceed the length of the cassette II template homology.

We find that Dnmt1 can act as a *de novo* methyltransferase, i.e., that it can recognize unmethylated DNA as a substrate. Although it is commonly thought that Dnmt1 cannot promote *de novo* methylation, in fact the preference of Dnmt1 for hemimethylated versus unmethylated DNA is only 7 to 21 fold (Pradhan et al.,1999). The possibility that the known *de novo* methyltransferases, Dnmt3a and Dnmt3b act at the DSB seems unlikely. First, we find that Dnmt1 -/- ES cells show a greater frequency of GFP^+ recombinants than wild-type cells. Second, these methyltransferases

are not recruited to sites of DNA damage (Mortusewicz et al., 2005). Third, all GFP recombinants were unmethylated in Dnmt1 -/- ES cells, although the parental molecules were normally methylated (Figure 9B). Fourth, Dnmt1 was specifically associated with HR GFP molecules (Figure 11). The similarity in the methylation profiles of parental DNA molecules in Dmnt1 -/- cells to wild-type cells implies that Dnmt3a and Dnmt3b can partially substitute for Dnmt1 during normal replication, but cannot replace Dnmt1 in methylation induced by HR (Figure 11A and 11B).

REPAIR AND GENE SILENCING

Alteration of the methylation pattern following gene conversion was restricted to 200–300 bp flanking the DSB. Changes in methylation did not extend to the CMV promoter that drives GFP expression or to cassette II.

The data shown in Figure 12 indicate that methylation at the BcgI site downregulates transcription initiation and probably partially blocks transcription elongation. We propose that local methylation initiates chromatin compaction and, ultimately, silencing of genes near the DSB (Padjen et al., 2005). Since the GFP gene is flanked by a puromycin-resistance gene transcribed on the other strand, which remains active, and is driven by a strong and insulated CMV promoter, the mechanism of inhibition appears to be selective. The nature and the extent of propagation of the silencing signal remains to be explored.

TRANSCRIPTION AND GENE SILENCING

To explore the molecular mechanism underlying the effect of methylation on transcription, we used a potent inhibitor of RNA polymerase II. α-amanitin blocks RNA polymerase II on DNA by preventing transcription initiation and elongation. We found that short pulses of α-amanitin during I-SceI action reduced the amount of GFP^+ cells, while their fluorescence intensity significantly increased. α-amanitin influenced the rate of recombination by reducing the accumulation of homologous repaired GFP. Also, the drug markedly accentuated the methylation profiles of the 2 classes of GFP expression found in the recombinant GFP population. In the pool of clones or in isolated clones the methylated GFP molecules (low expressors) became heavily methylated, whereas hypomethylated clones (high expressors) became non-methylated.

Since the appearance of these *de novo* methylation profiles suggest that the 2 DNA strand are differently methylated during homologous repair, we believe that RNA polymerase II may signal the DNMT1 complex the selective location where to methylate. We propose that the location of RNA polymerase targets DNMT1 to the (-) strand. The stalled transcription complex facilitates the methylation of the (-) strand and decrease the scattered methylation of the (+) strand. This explains the fluorescence data and why the α-amanitin treatment reduce the expression of GFP of the low expressors and increases the expression of the high expressors (Figure 14). Methylation analysis in Figure 15 support this model.

At the same time, stalled transcription forks might also interfere with DNA polymerase complex or strand invasion, because, despite the short

period of treatment, recombination rate was reduced by α-amanitin. Further studies will highlight the mechanisms underlying this phenomenon.

In conclusion, these results argue for a cause-effect relation between HR repair and *de novo* DNA methylation. Moreover, the existing link between RNA polymerase II- dependent transcription and homologous recombination opens a new and unexpected opportunity for the analysis of the mechanism(s) of DNA methylation of damaged genes. Our data suggest that hypermethylation of short DNA tracts in tumors or in aging may be the consequence of gene conversion. We note that silencing of tumor suppressor genes by methylation is characteristic of tumor cells (Jones et al., 2002). Indeed, evidence of enhanced DSB formation in hyperplastic precancerous cells precedes the genomic instability and loss of p53 characteristic of more advanced tumors (Pierce et al., 2001; Gorgoulis et al., 2005). Selection of methylated silenced alleles will eliminate unmethylated ones from the population (Figure 16).

Our data also imply that gene imprinting may be linked to homologous recombination events. In this perspective it is noteworthy that a systematic genome-wide analysis reveals that human imprinted chromosomal regions are historical hot spots of recombination (Sandovici et al., 2006).

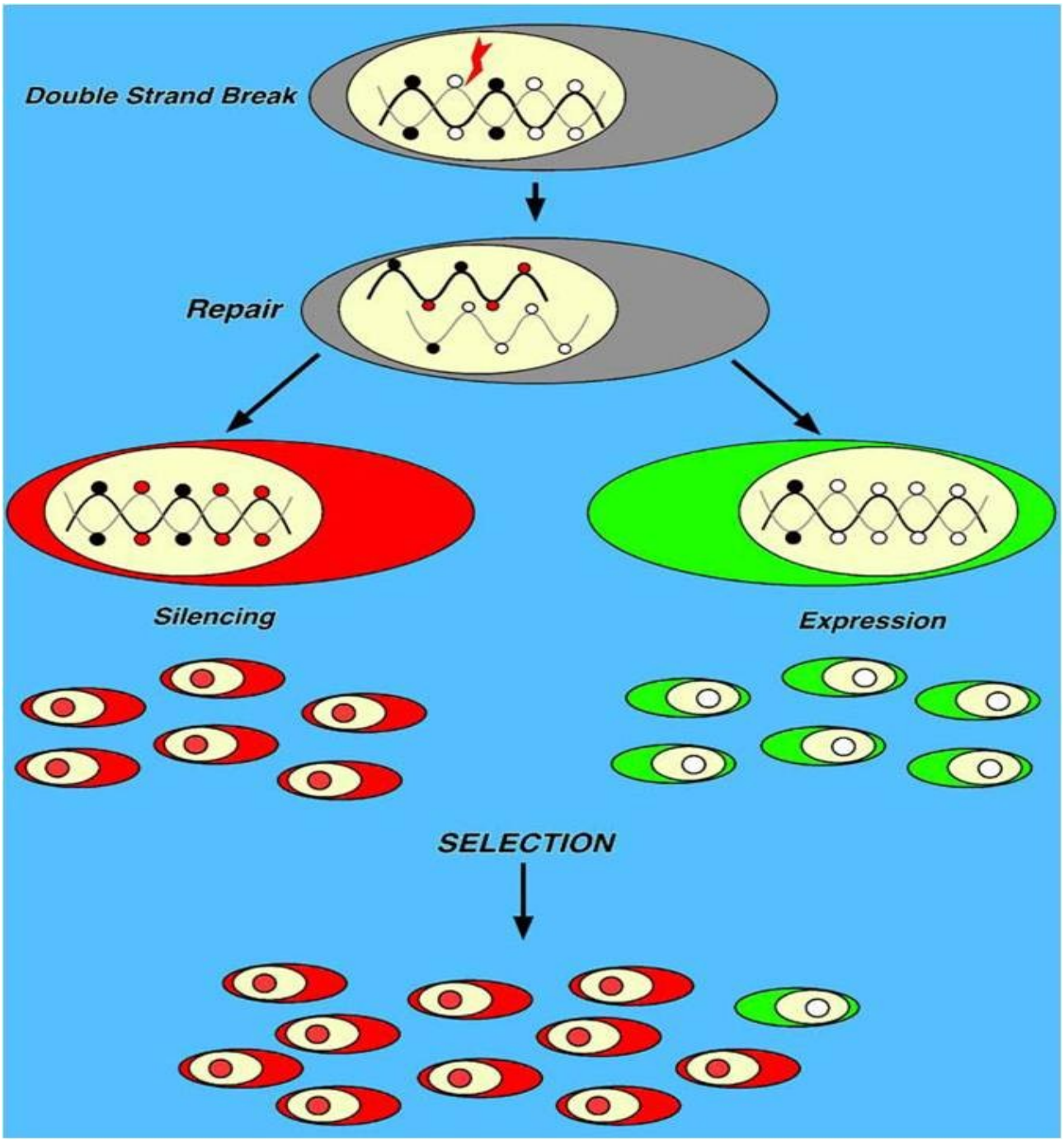


Figure 16. Biological Consequences of Recombination-Induced Methylation Switch. The drawing illustrates the sequence of events leading to silencing or expression of HR DNA segments. Red circles represent *de novo* methylated CpGs induced by HR. Black circles represent methylated CpGs before HR. Since silencing depends on the location of *de novo* methylated CpGs and DNA damage is random, HR-induced methylation is also random. If the expression of the repaired gene is harmful, only cells inheriting the silenced copy will survive. Conversely, if the function of the repaired gene is beneficial, cells inheriting the undermethylated copy will have a selective advantage.

MATERIALS AND METHODS

Cell culture

Hela-DR-GFP stable lines were cultured in RPMI medium supplemented with 10% fetal bovine serum (Invitrogen,), 1% penicillin-streptomycin, and 2 mM glutamine. ES-DR-GFP wild-type and ES stable lines were cultured in DMEM medium supplemented with 15% fetal bovine serum (Invitrogen), 50 U of penicillin streptomycin/ml, 0.1 mM β-mercaptoethanol, 0.1 mM nonessential amino acids (Invitrogen), and 500 U/ml of leukemia inhibitory factor (Invitrogen). Cells were grown on gelatine-coated dishes without feeder cells and cultured for amaximum of 20 generations (Hong et al., 1996).

The two cell lines were grown at 37 °C in 5% CO2.

Stable and transient transfections.

DR-GFP plasmid was 14,735 bp containing: CMV IE enhancer (1–385); chicken beta actin promoter (386–751); chicken beta-actin first intron (752–1,622); rabbit beta-globin second intron (1,623–1,670); rabbit beta-globin third exon (1,671–1,724); EGFP with a STOP (Pierce et al., 1999) codon at I-SceI site (1,740–2,756, I-SceI at 2,135); SV40 splice/polyadenylation signal (2,757–3,023); polyadenylation signal from phosphoglycerate kinase gene (3,025–3,607) for the puromycin resistance gene (3,600–4,200); a truncated EGFP gene sequence (5,609–6,138); and 6,450 bp of mouse genome (Jasin, 1996 and A. Porcellini, unpublished data).

Hela-DR-GFP stable lines were transfected with lipofectamine as recommended by the manufacturer (Invitrogen) with 2 μg of circular pDR-GFP plasmid and selected in the presence of puromycin (2 mg/ml). Puromycin-resistant colonies (approximately 200 clones) were seeded at 3 x 10^5 cells per 60-mm plate and transfected with 2.5 μg pCbASce

plasmid DNA on the following day by lipofectamine transfection (Invitrogen). Cells were harvested 2 or 4 d after transfection.

Wild-type ES and Dnmt1 -/- lines (Li et al., 1992) were transfected by electroporation in tissue culture medium with 50 μg circular pDR-GFP plasmid and plated in the presence of puromycin at 0.5 mg/ml. Puromycin-resistant colonies were transfected by electroporation with 10 μg pCbASce. At 24 h posttransfection, cells were expanded for a week in selective medium, aliquoted, and frozen at 135 °C. Aliquots of transiently transfected cells were cultured for 2 weeks and analyzed. Pools of clones (ES wild type, Dnmt1 -/-, or Hela) were generated in two or three independent transfections and frozen in aliquots. Transient transfections with I-SceI were carried at different times of culture after the primary transfection (for the isolation of single clones, we used cells cultured for more than 1 month). We used the same conditions of growth (~40% confluency starting from freshly frozen aliquots). Transfection efficiency was measured by assaying β-galactosidase activity of an included pSVbGal vector (Promega). Normalization by FACS was performed using antibodies to β-gal or pCMV-DsRed-Express (Clontech). pEGFP (Clontech) was used as GFP control vector.

Drug treatments.

Hela-DR-GFP cells were transfected with control vector or pCbASce plasmid. The cells were plated at low confluence 24h later and incubated with 5-AzadC (5 μM) for 48 h or with α-amanitin (2,5 μM) for 6 or 24h (The drugs come both from Sigma). The cells were analyzed 48 h later. When the drug was used before transfection, the treatment was terminated 48 h before transfection. The analysis of the cells was carried out 5d after transfection. Drug treatments were performed under the same growth conditions (40% confluency from freshly thawed aliquots).

RNA and DNA extraction.

Total RNA was extracted using Triazol (Gibco/Invitrogen).

Genomic DNA extraction was performed with following protocol: cellular pellet was resuspended in 10 mM TRIS (pH 7.8) and 50 mM NaCl solution. After addition of 1% SDS the sample was gently mixed. Proteinase K, at a final concentration of 20 µg/ml, was added and the mixture was incubated at 55 °C overnight. The following day, hot 1.5 M NaCl (70 °C) was added to the mixture, and the DNA was extracted by phenol/chloroform. DNA was ethanol precipitated, dried, and resuspended in TE buffer.

qRT-PCR, qPCR, RT-PCR, and PCR.

cDNA was synthesized in a 20 µl reaction volume containing 2 µg of total RNA, four units of Omniscript Reverse Transcriptase (Qiagen), and 1 µl random hexamer (20 ng/µl) (Invitrogen). mRNA was reversetranscribed for 1 h at 37 °C, and the reaction was heat inactivated for 10 min at 70 °C. The products were stored at -20 °C until use.

PCR was performed in a 50 µl reaction mixture containing 2 µl of synthesized cDNA product or 0.5 µg of genomic DNA, 5 µl of 10X PCR buffer, 1.5 mM MgCl2, 0.5 mM dNTP, 1.25 unit of Taq polymerase (Roche), and 0,2 µM of each primer.

The primer sequences that were used for the different mRNAs or DNAs were:

unrec 5'-GCTAGGGATAACAGGGTAAT-3';

rec 5'-GAGGGCGAGGGCGATGCC – 3';

reverse common oligo 5'-TGCACGCTGCCGTCCTCG-3' (443bp amplified fragment);

b-actin/R 5'-AAAGCCATGCCAATCTCATC-3';

b-actin/L 5'-GATCATTGCTCCTCCTGAGC-3' (250bp amplified fragment);

GAPDH forward oligonucleotide 5’-TTCACCACCACCATGGAGAAGGCT-3’;
GAPDH reverse 5’-ACAGCCTTGGCAGCACCAGT-3’ (346bp amplified fragment);
intron forward (738–757), 5’-CGTTACTCCCACAGGTGAGC-3’;
intron reverse (966–948), 5’-CGCCCGTAGCGCTCACAGC-3’;
AUG forward (1,666–1,685), 5’-TACAGCTCCTGGGCAACGTG-3’;
AUG reverse (1,911–1,892), 5’-TCCTGCTCCTGGGCTTCTCG-3’;
Prenested F (2100-2122),5’-gacgtaaacggccacaagttca-3’;
Prenested R (2659-2671),5’-ttctcgttggggtctttgctca-3’.
Amplifications were performed in a Primus thermocycler(MWG/Biotech) using the following program: 95 °C /5 min X 1 cycle; 95 °C /45 s, 57 °C /30 s, and 72 °C /2 min X 28–35 cycles; 72 °C/10 min X 1 cycle. The number of cycles were selected and validated by running several control reactions and determining the linear range of the reaction. A total of 15 µl of the PCR products were applied to a 1.2% agarose gel and visualized by ethidium bromide staining. Densitometric analysis was performed using a phosphoimager. Each point was determined for at least three independent reactions.

Quantitative (q)RT-PCR and qPCR were performed three times in six replicates on a 7500 RT-PCR System (Applied Biosystems) using the SYBR Green-detection system with the following program: 95 °C /5 min X 1 cycle, 95 °C /45 sec and 62 °C /45 min X 40 cycles. Reference curves were generated for 1, 3, 5, 10, and 1,000 ng of DR-GFP plasmid in six replicates for copy number assay and from 0.1, 0.5, 1.0, and 5.0 µl of each cDNA for GAPDH in the relative assay.

BcgI digestion of DNA or cDNA prior to PCR.

cDNA prepared from 2 µg of total RNA extracted from Hela-DR-GFP transfected with

pCbASce or control plasmid was digested with BcgI restriction enzyme as recommended by the manufacturer (New England Biolabs). We used 2.5 µl of cleavage reaction to perform PCR with primers that specifically amplify unrecombinant (**unrec**) and recombinant (**rec**) GFP sequence.

FACS analysis.

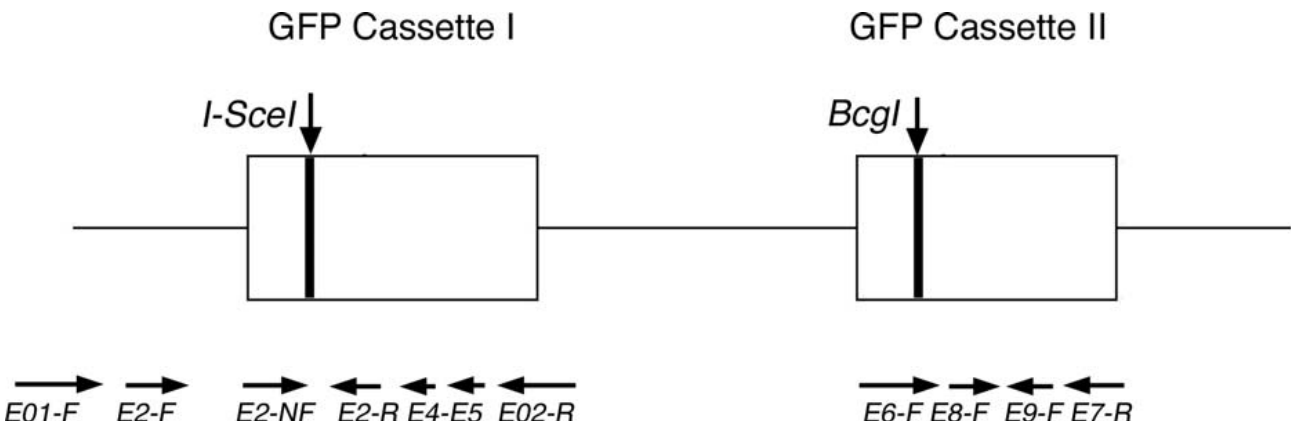
For the FACS analysis, Hela-DRGFP and ES–DR-GFP cells were transfected with pCbASce or control plasmid and after 2 or 4d were treated with trypsin and collected. After two washes with PBS, cells were resuspended in 500 µl of PBS 1X. A total of 6 X 10^4 cells were analyzed on a 9600 Cyan System (Dako Cytometrix).

Bisulfite DNA preparation, PCR, and sequence analysis.

Sodium bisulfite analysis was carried out essentially as described by Frommer et al., 1992. A total of 8 µg sample of total DNA was digested with EcoRI restriction enzyme (New England Biolabs) and denatured in 0.3M NaOH for 15 min at 37 °C in a volume of 100 µl. We then added 60 µl of 10 mM hydroquinone and 1.04 ml of 3.6 M NaHS3 (pH 5). Reaction mixtures were incubated at 50 °C for 16 h in the dark. DNA was desalted and concentrated using Geneclean (Qbiogene/Bio101), denatured with 0.3M NaOH for 15 min at 37 °C, neutralized with 3M NH4OAc (pH 7), and ethanol precipitated. An aliquot of DNA was amplified with the following modified primers: primers:

E01-F, 5’-GTGTGATTGGTGGTTTTAGAGT-3’;
E02-R, 5’-CCATCCTCAATATTATAACAAAT-3’;
E2-F, 5’-GGAGTTGTTTATTGGGGTGGTGTTTATTTTGGT-3’;
E2-NF, 5’-TGGATGGTGATGTAAATGGTTATAAGTTT-3’;
E2-R, 5’-GTTTGTGTTTTAGGATGTTGTTG-3’;
E4-R, 5’-ACTTATACAACTCATCCATACCAAAAATAATCC-3’;
E5-R, 5’-ACTTATACAACTCATCCATACCGAAAATAATCC-3’;
E6-F, 5’-GGTTGTTATGAATAAAGGTGGTTATAAGA-3’;

E7-R, 5’-CTCACTCATTAAACACCCCAAACTTTACAC-3’;
E8-F, 5’-GAAGATTTTTPyGATTTGTAGTTTAAGTTTTAGG-3’;
and E9-R, 5’-GAAGATTTTTPyGATTTGTAGTTTAAGTTTTAGG-3’.
All PCR reactions were carried out in 100 µl reaction mixtures containing 5 µl bisulfite-treated genomic DNA, 200 µM dNTPs, 10 pmol of each primer, 1 mM MgCl2, 50 mM KCl, 10 mM Tris, 5% dimethyl sulfoxide, and 2 U of Taq polymerase (Stratagene) in a Primus thermocycler (MWG/Biotech) under the following conditions: 95 °C/5 min and 70 °C/2 min X 1 cycle, 97 °C/1 min, 54 °C/2 min, and 72 °C/1 min X 5 cycles; 95 °C/0.5 min, 52 °C/2 min, and 72 °C/1 min X 25 cycles; and 72 °C/3 minX 1 cycle. We used 2µl of the first PCR to perform the second PCR (PCR nested or seminested) under the same PCR conditions.
PCR products were cloned into pGEM-T Easy Vector (Promega), and at least 20 independent clones for each fragment were sequenced with the T7 primer.
Sequence analysis and alignment were performed using MegAlign software (a module of the Lasergene Software Suite for sequence analysis by DNASTAR) for MacOSX. Statistical analysis was performed using the JMP 6.0.3 statistical analysis software by S.A.S.



Chromatin ImmunoPrecipitation.

We seeded 7 X 10^5 cells per well of Hela-DR-GFP in a six-well plate one day before transfection. Cells were transfected with 2 μg of I-SceI plasmids using Lipofectamine 2000 (Invitrogen). Cells were plated in a 150-mm plate 24 h later and incubated with 1 μM of 5-AzadC (Sigma) for 24 h, 48 h, or 96 h. The medium and 5-AzadC were changed every 24 h. Cells were harvested and GFP positive cells were scored by FACS. Remaining cells were stored at -80 °C.

ChIp-IT Express kit (Active Motif) was used for ChIp experiments. Because cells were treated with 5-AzadC, formaldehyde cross-linking and reverse cross-linking steps were omitted.

Chromatin isolation and enzymatic shearing.

A total of ~1 X 10^7 cells were resuspended in 0.5-ml ice-cold lysis buffer supplemented with 2.5 μl protease inhibitor cocktail (ChIp-IT Express kit, Active Motif) and 2.5 μl of 100 mM PMSF. After 30 min incubation on ice, cells were homogenized with a Dounce homogenizer (ten strokes), and nuclei were collected by centrifugation at 2,000 g/5 min. The nuclei were resuspended in 0.5 ml of digestion buffer (provided by the kit) and incubated at 37 °C for 10 min with 25 μl of enzymatic shearing cocktail (200 U/ml, ChIp-IT Express kit, Active Motif). The reaction was stopped by adding 10 μl of 0.5 M EDTA. The sheared chromatin was stored at -80 °C. An aliquot of sheared chromatin was further treated with proteinase K, phenol/chloroform extracted, and precipitated to determine DNA concentration and shearing efficiency (input DNA).

Capture of chromatin on magnetic beads.

The ChIp reaction was set up according to the manufacturer's instructions. Briefly, the sheared chromatin (corresponding to 18 µg of DNA) was mixed with protein G magnetic beads and 2 µl of anti-Dnmt1 antibody (New England Biolabs) or 0.8 µg of normal rabbit IgG (Santa Cruz). The reaction mixture was incubated at 4 °C for 48 h. Beads were washed with wash buffer and immunoprecipitated DNA was recovered.

PCR analysis for unrec and rec.

PCR for **rec** was performed in a 20 µl reaction mixture containing 5 µl of recovered DNA, 0.2 mM dNTP, 1.25 units of HotStar Taq DNA polymerase (Qiagen), and 5 µM of each primer. Amplifications were performed using following conditions: 95 °C/15 min X 1 cycle; 95 °C/45 sec, 65 °C/30 sec, and 72 °C/1 min X 40 cycles; and 72 °C 10 min X 1 cycle.

PCR for **unrec** was performed in a 20 µl reaction mixture containing 1 µl of recovered DNA. Amplifications were performed using following conditions: 95 °C/15 min X 1 cycle; 95 °C/45 sec, 52 °C/30 sec, and 72 °C/1 min X 30 cycles; and 72 °C 10 min X 1 cycle.

PCR for Input was performed in a 20 µl reaction mixture containing 100 ng of input DNA. Amplifications were performed with 30 cycles.

PCR analysis for positive and negative controls.

To validate the Dnmt1 ChIp reactions, the following primers were used for PCR:

hMGMT, 813 bp, (-726 to +87); MGMT-F, 5'-GAGTCAGGCTCTGGCAGTGT-3';

MGMT-R, 5'-GAGCTCCGCACTCTTCCGG-3';

pP16, 778 bp, (-656 to +132); p16-F, 5'-GCAGTCCGACTCTCCAAAAG-3';

p16-R, 5'-AGCCAGTCAGCCGAAGGC-3';

b-actin, 250 bp; b-actin/R: 5'-AAAGCCATGCCAATCTCATC-3';

and b-actin/L: 5'-GATCATTGCTCCTCCTGAGC-3'.

PCR was performed in a 20 µl reaction mixture containing 5 µl (1 µl for actin) of recovered DNA, 0.2 mM dNTP, 1.25 unit of HotStarTaq DNA

polymerase (Qiagen), and 5 µM of each primer. Amplifications were performed using following conditions: 95 °C/15 min X 1 cycle; 95 °C/45 sec, 50 °C/30 sec, and 72 °C/1 min X 40 cycles; and 72 °C/10 min X 1 cycle.

BIBLIOGRAPHY

Baylin S.B., Herman J.G., Graff J.R., Vertino P.M. and Issa J.P. (1998).Alterations in DNA methylation: a fundamental aspect of neoplasia. **Adv Cancer Res** 72: 141–196.

Ballestar E. and Esteller M. (2002). The impact of chromatin in human cancer:linking DNA methylation to gene silencing. **Carcinogenesis review** 23(7), 1103-1109.

Bhattacharya S.R., Ramchandani S., Cervoni N. and Szyf M. (1999). A mammalian protein with specific demethylase activity for mCpG DNA. **Nature** 397, 579-583.

Benovoy D. and Drouin G. (2008). Ectopic gene conversions in the human genome. **Genomics** article in press

Bestor T., Laudano A., Mattaliano R. and Ingram V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. **J Mol Biol** 203:971–983.

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

Boyes J. and Bird A. (1991). DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. **Cell** 64: 1123–1134.

Bushnell D.A., Cramer P. and Kornberg R.D. (2002). Structural basis of transcription: α-Amanitin–RNA polymerase II cocrystal at 2.8 Å resolution. **PNAS** 99(3), 1218–1222.

Chuang L.S., Ian H.I., Koh T.W., Ng H.H., Xu G. and Li B.F. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. **Science** 277: 1996–2000.

Chen R.Z., Petterson U., Beard C., Jackson-Grusby L. and Jaenisch R. (1998). DNA hypomethylation leads to elevated mutation rates. **Nature** 395, 89-93.

Colot V. and Rossignol J.L. (1999). Eukaryotic DNA methylationas an evolutionary device. **BioEssays** 21:402–411.

Counts J.L. and Goodman J.I. (1995) Alterations in DNA methylation may play a variety of roles in carcinogenesis. ***Cell*** 83: 13–15.

Cromie G.A., Connelley J.C. and Leach D.R.F. (2001) Recombination at double-strand breaks and DNA ends:conserved mechanism from phage to human. **Moll. Cell.** 8, 1163-1174.

Cuozzo C., Porcellini A., Angrisano T., Morano A., Lee B., Di Pardo A., Messina S., Iuliano R., Fusco A., Santillo M.R., Muller M.T., Chiariotti L., Gottesman M.E. and Avvedimento E.V. (2007). DNA Damage, Homology-Directed Repair, and DNA Methylation. **PLoS Genetics** 3(7), 1144-1162

D'Alessio Ana C., Weaver Ian C. G. and Szyf M. (2007). Acetylation-Induced Transcription Is Required for Active DNA Demethylation in Methylation-Silenced Genes. **Mol. Cell. Biol.** 27(21), 7462–7474

Domìnguez-Bendala J. and McWhir J. (2004). Enhanced gene targeting frequency in ES cells with low genomic methylation levels. **Transgenic Research** 13: 69–74.

El-Osta A. (2004). Coordination of epigenetic events. **Cell Mol Life Sci**. 61(17), 2135-2137

Essers J., Van Steeg H., De Wit J., Swagemarkers S.M.A., Vermeij M., Hoeijimakers J.H.J. and Kanaar R. (2000). Homologous and non-homologous recombination differentially affect DNA damage repair in mice. **EMBO J.** 19,1703-1710.

Ezawa K., Oota S. and Saitou N. (2006). Genome-wide search of gene conversions in duplicated genes of mouse and rat, **Mol. Biol. Evol.** 23, 927–940.

Frommer M., McDonald L.E., Millar D.S., Collis C.M., Watt F. et al. (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. **Proc Natl Acad Sci U S A** 89: 1827–1831.

Fuks F., Burgers W.A., Brehm A., Hughes-Davies L. and Kouzarides T. (2000) DNA methyltransferase *Dnmt*1 associates with histone deacetylase activity. **Nat Genet 24**: 88–91.

Fuks F., Burgers W.A., Godin N., Kasai M. and Kouzarides T. (2001). Dnmt3 binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. **EMBO J.** 20, 2536-2544.

Gorgoulis V.G., Vassiliou L.V., Karakaidos P., Zacharatos P., Kotsinas A. et al. (2005). Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. **Nature** 434: 907–913.

Green C.M. and Almouzni G. (2003). Local action of the chromatin assembly factor CAF.1 at sites of nucleotide excision repair in vivo. **EMBO J.** 22, 5163-5174.

Guenther M. G., Jenner R. G., Chevalier B., Nakamura T., Croce C. M., Canaani E. and Young R. A. (2005). Global and Hox-specific roles for the MLL1 methyltransferase. **Proc. Natl. Acad. Sci. USA** 102:8603–8608.

Haber J.E. (2000) Partners and pathways repairing a double-strand break. **Trends Genet.** 16, 259- 264.

Hahn S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. **Nat. Struct. Mol. Biol.** 11, 394–403.

Hendrich B. And Bird A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins**. Mol. Cell. Biol.** 18(11), 6538-6547.

Hermann A., Gowher H. and Jeltsch A. (2004). Biochemistry and biology of mammalian DNA methyltransferases. **Cell Mol Life Sci** 61:2571–2587.

Hong L., Suk P., Masaki Okano, Ruth Juttermann, Kendrick A. Goss, Rudolf Jaenisch and En Li. (1996). De novo DNA cytosine methyltransferase activities in mouse embrionic stem cells. **Develpment.** 122, 3195-3205

Jasin M. (1996). Genetic manipulation of genomes with rare-cutting endonucleases. **Trends Genet.** 12, 224-228.

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

Jüttermann R., Li E. and Jaenisch R. (1994). Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation. **Proc Natl Acad Sci U S A** 91: 11797–11801.

Kass S.U., Goddard J.P. and Adams R.L. (1993). Inactive chromatin spreads from a focus of methylation. **Mol Cell Biol** 13:7372–7379.

Kass S.U., Landsberger N. and Wolffe A.P. (1997). DNA methylation directs a time-dependent repression of transcription initiation. **Curr Biol** 7:157–165.

Kass S.U., Pruss D. and Wolffe A.P. (1997). How does DNA methylation repress transcription? **Trends Genet** 13:444–449.

Kim M., Trinh B. N., Long T. I., Oghamian S. and Laird, P. W. (2004). Dnmt1 deficiency leads to enhanced microsatellite instability in mouse embryonic stem cells. **NucleicAcids Res.** 32, 5742–5749

Kinner Andrea, Wenqi Wu, Christian Staudt and George Iliakis. (2008). γ-H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin. **Nucleic Acids Research**, 36 (17), 5678–5694.

Kobayashi Junya, Iwabuchi Kuniyoshi, Miyagawa Kiyoshi, Sonoda Eiichiro, Suzuki Keiji, Takata Minoru and Tauchi Hiroshi. (2008). Current topics in DNA double-strand break repair. **J. Radiat. Res.** 49(2), 93-103.

Komarnitsky, P., Cho E. J. and Buratowski S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. **Genes Dev.** 14:2452–2460.

Latham KE (1996) X chromosome imprinting and inactivation in the early mammalian embryo. **Trends Genet 12**: 134–138.

Leonhardt H., Page AW., Weier H.U. and Bestor T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. **Cell** 71: 865–873.

Lewis J.D., Meehan R.R., Henzel W.J., Maurer-Fogy I., Jeppesen P., Klein F. and Bird A. (1992) Purification,sequence and cellular localization of a novel chromosomal protein that binds to methylated DNA. **Cell** 69, 905-914.

Li E., Bestor T.H. and Jaenisch R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. **Cell** 69: 915–926.

Li B., Carey M. and Workman J.L. (2007). The role of chromatin during transcription. **Cell** 128: 707–719.

Lieber Michael R. (2008) The Mechanism of Human Nonhomologous DNA End Joining. **JBC Papers**, 283(1), 1-5

McCabe N., Turner N.C., Lord C.J., Kluzel K., Bialkowska A., Swift S., Giavara S., O'Connor M.J., Tutt A.N., Zdzienicka M.Z., Smith G.C. and Ashworth A. (2006). Deficiency in the repair of DNA damage by homologous recombination and sensitivity to poly(ADP-ribose) polymerase inhibition. **Cancer Res.** 66, 8109-8115.

McCabe M.T., Low J.A., Daignault S., Imperiale M.J., Wojno K.J. and Day ML. (2006). Inhibition of DNA methyltransferase activity prevents

tumorigenesis in a mouse model of prostate cancer. **Cancer Res** 66: 385–392.

Mortusewicz O., Schermelleh L., Walter J., Cardoso M.C. and Leonhardt H. (2005). Recruitment of DNA methyltransferase 1 to DNA repair sites. **PNAS** 102(25), 8905-8909.

Nan X., Cross S. and Bird A. (1998). Gene silencing by methyl-CpG-binding proteins. **Novartis Found Symp** 214:6–16.

Nan X., Ng H.H., Johnson C.A., Laherty C.D., Turner B.M., Eisenman R.N. and Bird A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. **Nature** 393:386–389.

Okano M., Xie S. and Li E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. **Nat Genet** 19: 219–220.

Okano M., Bell D.W., Haber D.A. and En Li. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. **Cell** 93, 305-308.

Orphanides, G. and D. Reinberg. (2000). RNA polymerase II elongation through chromatin. **Nature** 407:471–475.

Padjen K., Ratnam S. and Storb U. (2005). DNA methylation precedes chromatin modifications under the influence of the strain-specific modifier Ssm1. **Mol Cell Biol** 25: 4782–4791.

Patra S.K. and Bettuzzi S. (2007). Epigenetic DNA methylation regulation of genes coding for lipid raft-associated components: A role for raft

proteins in cell transformation and cancer progression (Review). **Oncology Reports**, 17, 1279–1290.

Patra S.K., Patra a., Rizzi F., Ghosh T.C. and Bettuzzi S. (2008). Demethylation of (Cytosine-5-C-methyl) DNA and regulation of transcription in the epigenetic pathways of cancer development. **Cancer Metastasis Rev** 27:315–334.

Pierce A.J., Johnson R.D., Thompson L.H. and Jasin M (1999). XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. **Genes Dev** 13: 2633–2638.

Pierce A.J., Stark J.M., Araujo F.D., Moynahan M.E., Berwick M. et al. (2001) Double-strand breaks and tumorigenesis. **Trends Cell Biol** 11: S52–S59.

Pikaart M.J., Recillas-Targa F. and Felsenfeld G. (1998). Transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. **Genes Dev** 15: 2852–2862.

Pradhan S., Bacolla A., Wells R.D. and Roberts R.J. (1999). Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. **J Biol Chem** 274: 33002–33010.

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

Reik W. and Walter J. (1998). Imprinting mechanisms in mammals. **Curr Opin Genet Dev 8**: 154–164.

Richardson C., M.E. Moynahan and M. Jasin (1998). Double strand break repair by interchromosomal recombination: Suppression of chromosomal translocation. **Genes & Dev.** 12, 3831-3842

Robert M.F., Morin S., Beaulieu N., Gauthier F., Chute I.C., Barsalou A. et al. (2003). DNMT1 is required to maintain CpG methylation and aberrant gene silencing in human cancer cells. **Nat Genet** 33: 61–65.

Robertson K.D., Uzvolgyi E., Liang G., Talmadge C., Sumegi J., Gonzales F.A et al. (1999). The human DNA methyltransferases (DNMTs)1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. **Nucleic Acids Res** 27:
2291–2298.

Robertson K.D., Ait-Si-Ali S., Yokochi T., Wade P.A., Jones P.L. and Wolffe A.P. (2000). DNMT1 forms a complex with Rb, E2F1 e HDAC1 and repress transcription from E2F-responsive promoters. **Nature Genet.** 25, 338-342.

Rountree M.R., Bachman K.E. and Baylin S.B. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. **Nat Genet** 25: 269–277.

Rozen S., Skaletsky S., Marszalek J.D., Minx P.J., Cordum H.S., Waterston R.H., Wilson R.K. and Page D.C. (2003). Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. **Nature** 423, 873–876.

Sancar A., Lindsey-Boltz L.A., Ünsal-Kaçmaz K. and Linn S. (2004). Molecular Mechanisms of Mammalian DNA Repair and the DNA Damage Checkpoints. **Annu. Rev. Biochem.** 73, 39-85.

Sandovici I., Kassovksa-Bratinova S., Vaughan J.E., Stewart R., Leppert M. et al. (2006). Human imprinted chromosomal regions are historical hot-spots of recombination. **PLoS Genet** 2: e101. doi:10.1371/journal.pgen.0020101

Santoyo G. and Romero D. (2005). Gene conversion and concerted evolution in bacterial genomes, **FEMS Micro. Rev.** 29, 169–183.

Schermelleh L., Spada F., Easwaran H.P., Zolghadr K., Margot J.B. et al. (2005). Trapped in action: Direct visualization of DNA methyltransferase activity in living cells. **Nat Methods** 2: 751–756.

Shiloh Y. (2003). ATM and related protein kinases: safe-guarding genome integrity. **Nat Rev Cancer,** 3, 155-168.

Spence J., Gali R.R., Dittmar G., Sherman F., Karin M. and Finley D. (2000). Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain. **Cell.** 102, 67-76.

Thomas M.C. and Chiang C.M. (2006). The general transcription machinery and general cofactors. **Crit. Rev. Biochem. Mol. Biol.** 41, 105–178.

Waldman Alan S.. (2008). Ensuring the fidelity of recombination in mammalian chromosomes. **BioEssays** 30:1163–1171.

Wu J., Issa J.P., Herman J., Bassett Jr D.E., Nelkin B.D. and Baylin S.B. (1993). Expression of an exogenous eukaryotic DNA methyltransferase

gene induces transformation of NIH 3T3 cells. **Proc Natl Acad Sci USA** 90: 8891–8895.

Xu G.L., Bestor T.H., Bourchis D., Hsieh C.L., Tommerup N., Bugge M., Hulten M., Qu X., Russo J.J. and Viegas-Pèquignot E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. **Nature** 402, 187-191.

Yan PS, Shi H, Rahmatpanah F, Hsiau TH, Hsiau AH, Leu YW et al. (2003). Differential distribution of DNA methylation within the RASSF1A CpG island in breast cancer. **Cancer Res** 63: 6178–6186.

You J.S., Kang J.K., Lee E.K., Lee J.C., Lee S.H., Jeon Y.J., Koh D.H., Ahn S.H., Seo D-W, Lee H.Y., Cho E-J and Han J-W (2008). Histone deacetylase inhibitor apicidin downregulates DNA methyltransferase 1 expression and induces repressive histone modifications via recruitment of corepressor complex to promoter region in human cervix cancer cells**. Oncogene** 27, 1376–1386

Zhou B.B. and Elledge S.J. (2000). The DNA damage response: putting checkpoints in perspective. **Nature**, 408, 433–439.