Dock, et al. Clon Transgen 2016, 5:1
DOI: 10.4172/2168-9849.1000148

Cloning & Transgenesis
ISSN: 2168-9849

Research Article Open Access

# Maintenance of the *E.coli* dcm Methylation of the CMV Promoter, in Contrast to Hypomethylation of the Recognition Sequence of Transcription Factor NFkB in Transfected GBM Cells

**Dock H[2], Sjögren H-O[1], Salford LG[1], Widegren B[1]+ and Xue Z-T[1]***

[1]*The Rausing Laboratory, Division of Neurosurgery, Department of Clinical Sciences, Lund University, BMC-D10, SE-221 84 Lund, Sweden*
[2]*Division of Clinical Sciences, Department of Clinical and Experimental Medicine, Faculty of Health Science, Linköping University, S-581 85 Linköping, Sweden*
[+]*Deceased*

## Abstract

The human cytomegalovirus (CMV) immediate early promoter has been extensively used to drive target gene expression in transgenic mammalian cells. DNA methylation of the CMV promoter has been shown to be the reason for a reduced promoter activity and silencing of the target gene. We have established an *in vitro* model system, in which human brain cancer cells (glioblastoma multiforme, GBM) were transfected with pAdTrack-CMV-GFP plasmid, isolated from a dcm positive (dcm+) *E. coli* strain. We found that in two CCTGG sequences located at position from -304 to -300 nt and from -497 to -493 nt of the CMV promoter region, the internal C was methylated in all analyzed clones, i.e., the *E. coli* dcm methylation pattern is maintained in the CMV promoter region after its integration into the human genome. In contrast, we found that the recognition sites for the transcription factor NFkB and certain other transcription factors in the enhancer region of the CMV promoter (from -107 to -270 nt) were hypomethylated. This might explain why the CMV promoter maintained an active mode, driving the GFP expression despite the demonstrated methylation of the CMV promoter.

We noticed that the CCTGG sequence is also contained in the binding sequence motif of transcription factor NFkB. Hence we have comprehensively studied transcription factors through a database searching, and the responsive elements that contain dcm methylation sequences CCW(A/T)GG. A list of transcription factors and the corresponding regulated genes are presented.

**Keywords:** Adenovirus; Gene therapy; Promoter methylation; Plasmid DNA; Transcription factor

## Introduction

DNA methylation has attracted great interest due to its important role in regulation of the gene expression in most eukaryotes [1-3]. DNA methylation in eukaryotes mainly occurs in cytosine residues in palindromic CpG dinucleotide. In vertebrates such CpG sites are not evenly distributed in the genome but are present in only one fifth of its predicted random frequency. However, upstream regions of many genes have high CpG frequencies, known as CpG islands. CpG islands have also been shown to be present within or between genes [4]. Evidence obtained from the study of DNA methylation in eukaryotes indicates that active gene expression as a rule requires unmethylated CpG sites especially in the promoter region. When CpG sites are methylated, transcription of the gene will be switched off [5-7]. An example of the important implications of the DNA methylation is the noticed aberrant DNA methylation of CpG islands associated with tumorigenesis and tumor progression [8,9]. Activation of many oncogenes has been proposed to be caused by demethylation during carcinogenesis [10,11]. In cancer cells, regional hypermethylation events occurred accompanying a global hypomethylation. Loss of gene expression is frequently caused by hypermethylation of the promoter region of the gene, such as tumor suppressor genes [12-14]. Since changes in gene function due to DNA methylation are not due to changes in the DNA sequence, they therefore belong to the "epigenetic" pathway. Recent research achievements have shown that cancer has a common basis in early events of epigenetic alterations of so called "tumor-progenitor genes" present in the cancer progenitor cells [15]. Analysis of DNA methylation therefore has a potential impact on cancer risk assessment, chemoprevention and gene therapy.

In gene therapy, the human cytomegalovirus (CMV) immediate early gene promoter with enhancer (CMV-PE) has been widely used to drive the target gene expression. However, when this and other promoters are used in *in vivo* approaches for gene therapy, the transgene silencing has become an issue for a successful therapy [16]. This raises the question whether silencing of the transgene under control of the CMV promoter could be due to methylation of the CMV-PE? It has been shown that methylation of the CMV-PE at CpG sites efficiently blocks promoter activity in *in vitro* transfected cells, in fish embryos and rat muscle [17-19]. In our present work, we have established an *in vitro* model system in which human GBM cells were transfected with the pAdTrack-CMV-GFP gene construct (Clontech). The methylation status of the CMV promoter is determined by bisulfite sequencing. We demonstrate that when a plasmid, containing the immediate early CMV promoter isolated from *E. coli* (dcm+), is transfected into human GBM cells, the *E. coli* dcm methylation with the two CCTGG sequences in the CMV promoter region is maintained after stable integration in the human genome. The GFP expression was monitored by flow cytometric analysis and a stable level of GFP expression was observed for 2-5 months after the transfection. In contrast, we find that the recognition sites for the NFkB and certain other transcription factors in the enhancer region of the CMV promoter are hypomethylated, possibly providing a reason for maintenance of an active mode of the CMV promoter in our *in vitro* cultured human GBM cells.

***Corresponding author:** Xue Z-T, The Rausing Laboratory, Division of Neurosurgery, Department of Clinical Sciences, Lund University, BMC-D10, SE-221 84 Lund, Sweden, Tel: 46 462229260, Fax: 46462224606; E-mail: Zhongtian.xue@med.lu.se

**Received** February 09, 2016; **Accepted** March 06, 2016; **Published** March 11, 2016

**Citation:** Dock H, Sjögren H-O, Salford LG, Widegren B and Xue Z-T (2016) Maintenance of the *E.coli* dcm Methylation of the CMV Promoter, in Contrast to Hypomethylation of the Recognition Sequence of Transcription Factor NFkB in Transfected GBM Cells. Clon Transgen 5: 148. doi:10.4172/2168-9849.1000148

## Materials and Methods

### DNA and cell line

Plasmid DNA pAdTrack-CMV-GFP (Clontech) was isolated from *E.coli* DH5-α (dcm+) strain and purified using Qiagen DNA purification column. A human GBM cell line GA49 was previously established in our laboratory under GMP conditions from malignant tissue of a patient, subjected to surgery at the department of Neurosurgery. After transfection with the plasmid DNA and stable integration (2-5 months), the genomic DNA was isolated by using a DNA purification kit (Gentra, USA).

### Co-transfection and flow cytometric analysis of GFP expression

Human GBM cells that had been grown in the Iscove's Modified Dulbecco's Medium (IMDM-20 containing 20% FCS) to about 80% confluency were co-transfected with pAdTrack-CMV-GFP plasmid DNA and pLXSN plasmid DNA (Clontech) that contains the neomycin resistance gene as a selection marker using the FuGene 6 transfection system (Roche, Germany). 12 µl of FuGene 6 were mixed with 0.5 ml of medium, then 4 µg of pAdTrack-CMV-GFP DNA and 2 µg of pLXSN DNA were added. The mixture was kept for 2 h at room temperature then transferred into a flask and incubated for 2 h at 37°C 5% $CO_2$. Then 4.5 ml pre-warmed medium was added to the flask. Transient expression of the GFP was checked by fluorescence microscopy after incubation for 24 hrs. Stable expression of the GFP was assessed by flow cytometric analysis after each passage during cultivation under selection with 500 g/ml Geneticin G-418 (Life Technologies). The transfected cells were collected during 2-5 months post-transfection for DNA isolation and further methylation analysis.

### Bisulfite sequencing

Bisulfite reaction of DNA is a modification based on the method reported previously [20]. One microgram DNA in a volume of 30 µl distilled water was denatured by NaOH (final concentration, 0.2 M) for 10 min at 37°C. Freshly prepared sodium bisulfite (Sigma) 3 M, 600 µl at pH 5, and 60 µl of 40 mM hydroquinone (Sigma) were added. Samples were incubated in a thermal cycler with the following program: 95°C, 3 min; 55°C, 16 h and every 60 min 95°C, 3 min. DNA were purified by using the DNA purification spin column (Qiagen, UK). Modification was completed by precipitation of the eluted DNA with NaOH (final concentration, 0.3 M) at room temperature for 15 min. Modified DNA after centrifugation and washing with 70% ethanol was dissolved in 50 µl of distilled water. Several pairs of designed PCR primers were used for methylation specific PCR (MSP). They are listed as the following:

Forward primers:

1. CMV22: 5'-GTATTTTAGTTGTGGTTTG-3'
2. CMV23: 5'-GTTTAAATTTATTAATGTATT-3'
3. CMV24: 5'-AATGGTTTGTCTGGTTGA-3'
4. CMV24B: 5'-AATGGTTTGTTTGGTTGA-3'
5. CMV25: 5'-TTAATGGGTGGAGTATTTA-3'
6. CMV26: 5'-GTATGTTTTTTATTGATGTT-3'
7. CMV27: 5'-AAGTTTTTATTTTATTGATGTT-3'

Reverse primers:

8. CMV3: 5'-ACCAAAATAAACACCACC-3'
9. CMV31: 5'-ACAAATAAACTTCAAAATCA-3'
10. CMV26R: 5'-AACATCAATAAAAAACATAC-3

The PCR mixture contained 1X PCR buffer, dNTP (each at 0.2 mM), primers (0.4 µM each per reaction), 5% dimethyl sulfoxide and 1 µl of bisulfite-modified DNA (20 ng) in a final volume of 25 µl. The program for PCR was as follows: 94°C for 3 min, then 35 cycles of 94°C for 30 sec, 50°C for 45 sec, 72°C for 60 sec, and finally 5 min at 72°C. The PCR products were cloned in *E.coli* using pGEM-T easy cloning kit (Promega, USA). Six clones from different DNA sources were randomly selected, according to the sizes of the inserts checked on an agarose gel. The cloned DNA was sequenced, using the capillary electrophoresis automated sequencing approach (Applied Biosystem, USA) after thermocycle sequencing reaction using 3,1 version kit. As controls, the pAdTrack-CMV-GFP constructs were passed through dcm+ DH5α or dcm- INV110 bacteria (Invitrogen, USA) and bisulfite sequenced to confirm the methylation status of $C^mCWGG$ sites.

### Database search for transcription factor responsive elements with CCWGG sites

We used the Transcriptional regulatory element database according to Cold Spring Harbor Laboratory (http://rulai.cshl.edu/TRED) [21,22] to search and list the transcription factors and their recognition sequences containing CCWGG sites.

## Results

### Reporter GFP expression in the transfected GBM cells

A human GBM cell line GA49 was previously established from malignant tissue of a patient as described in the Methods. We have used this cell line to establish an *in vitro* transfection system in which the GBM cells were transfected with the commonly used gene construct pAdTrack-CMV-GFP plasmid [23,24]. Since pAdTrack-CMV-GFP vector contains no selection marker in eukaryotic cells, another vector, pLXSN plasmid that contains the neomycin resistance gene under control of the SV40 promoter was used in our transfection experiments called co-transfection. Moreover, the SV40 promoter contains no restriction enzyme EcoRII sites (CCWGG). As a reporter gene, the GFP expression can be transiently detected by fluorescence microscopy during 1-2 days after the transfection. The stable transfection in which the pAdTrack-CMV-GFP plasmid DNA is integrated into the genome was observed for one month after transfection. The GFP expression, driven by the CMV promoter in the stably transfected cells was measured by flow cytometric analysis. The time course of the GFP expression in the transfected cells during the post-transfection period (from day 100 to day 162) is shown in Supplementary Figures 1A and 1B. The gated GFP cells constituted 13%-37% of the total cell number. We used cells cultured for 141 days in which the GFP positive cells represented about 25% of the total cell number for further study.

### Methylation status of the CMV promoter in the transfected GBM cells

We isolated DNA from the stably transfected cells to determine the methylation status of the CMV promoter by bisulfite sequencing. The CMV promoter region from the transcription start to the nucleotide -573 of the CMV promoter was chosen for analysis of the methylation status (Supplementary Figure 2A).

Six clones were randomly selected from the transformants of the PCR products and were sequenced. The results showed that 4 clones contained full-length fragments with the expected size and 2 clones contained truncated fragments, 240 bp shorter at the 5' end. It can be seen that the methylation occurred not only strikingly at C of the CpG dinucleotide but also at non-CpG sequences (CpA, CpT, and CpC)

including two C$^{m}$CTGG sequences located at the nucleotide position -497 to -493 (Cytosine's position no.18 C, C18) and -304 to -300 (C73). The internal Cs of these sequence motifs from different clones were all methylated (Supplementary Figures 2B-2D). It has been known that CC(A/T)GG is the recognition site of the *E.coli* dcm methylase in the dcm+ bacteria such as DH5-α. As a control, the pAdTrack-CMV-GFP plasmid DNA isolated from DH5α showed C$^{m}$CTGG pattern at the position C73, while the same plasmid amplified in a dcm- strain INV110 (Invitrogen, USA) did not give the dcm methylation pattern at this position. However, at the position C18, both these two bacterial strains sometimes showed the dcm methylation pattern (Supplementary Figure 2E).

Since it is known that several transcription factors play an important role in activation of the CMV promoter, we have analyzed the methylation status of certain transcription factor recognition sequences. Four NFkB sites with the consensus sequences CGGGACTTTCC and GGGGATTTCC are present in the CMV promoter. We found no methylation in 3 of 4 NFkB sites, located close to the transcription start (C118, C102, C81-83) in the 5 analyzed clones. In 2 of 4 analyzed clones the distal NFkB site (C43-45) was partially methylated. The CMV promoter contains three Sp1 sites with the consensus sequences CCGCCC and CCCGCC. No methylation occurred at the site close to the TATAA box of the promoter (C125-130) and the other two sites, C20-C21 and C27 were only partially methylated (1/4 and 2/4 analyzed clones respectively) and all other Cs were non-methylated. The CMV promoter also contains three CREB/ATF binding sites (TGACGTCAA) and one AP-1 binding site (TGACTCA). Among CREB/ATF sites, one site (C66) was not methylated in any of six analyzed clones, and two others were only partially methylated (1/4 and 3/4 analyzed clones respectively).

### Transcription factors involved in the interaction with CCW(A/T)GG pentanucleotides in the eukaryotic promoters

Since methylation at the CCW(A/T)GG sequence exists in mammalians as well as in the eukaryotic viral genome (Supplementary Figure 3), methylation at this site regulates the gene expression through blockage of the interaction between this sequence and the transcription factors [25]. This raises an interesting question of how many transcription factors and genes that might be involved in this interaction and what is the significance of this interaction. We have searched the database for transcription factors [21,22] and we report 22 factors for the CCAGG sequence (Supplementary Figure 4) and 14 factors for CCTGG (Supplementary Figure 5). Since all transcription factors have not yet been included in the database, e.g., the early B cell factor, some additional transcription factors affected by methylation might be added to our list in the future.

## Discussion

The human cytomegalovirus (CMV) immediate early gene promoter has been extensively used in transfection systems to drive target gene expressions in mammalian cells. However, when this and other viral promoters are used in *in vivo* approaches for gene therapy, the transgene is silenced over time following the transfection [16]. DNA methylation of the CMV promoter has been shown to be the reason for a reduced promoter activity and silencing of the target gene [17-19,25-27]. When an anti-methylating agent such as 5'-azacytidine or 5'-deoxyazacytidine was used to treat the cells transfected with a target gene, driven by CMV promoter, the transcriptional expression of the target gene was reactivated and was maintained for longer time [26-28]. However, it is not known whether the anti-methylating agent directly affects the CMV promoter or indirectly affects the transcriptional expression of, for instance, the transcription factor genes that are required for the CMV promoter activation. It is of interest to know the methylation status of the CMV promoter. As we established an *in vitro* model system in which the human GBM cells were transfected with the pAdTrack-CMV-GFP gene construct, we have determined the methylation status of the CMV promoter region, which has been integrated from the vector into the GBM genome. Intriguingly, we found that the region from -302 to -532 nt showed hypermethylation, which is consistent with the early report [19], which showed that in transfected rat muscle a representative 50-bp region of the CMV-PE was extensively methylated 7 or 14 days after transfection. However, we also compared the methylation status in the 5' upstream region close to the transcription start (-1 to -301) between the rat muscle cells and the human GBM cells, and found that in the rat muscle 60-70% of the cytosine residues were methylated [19], but in the human GBM cells they were hypomethylated (Figure 3). Besides that, we also found two CCTGG sites at positions from -304 to -300 nt and -497 to -493 nt of the CMV promoter region, in which the internal C was methylated in all analyzed clones. Since we know that the CCW(A/T)GG methylation was initially discovered in the dcm+ *E. coli* strains such as DH5-α, it means that this type of DNA methylation in prokaryotes is somehow maintained in the integrated CMV promoter in the human genome.

Naturally occurring DNA methylation at the CCW(A/T)GG site has been sparsely found in eukaryotes and human genome [25,29-31]. For example, In the human myogenic gene, Myf-3, which is not targeted by the methylating system that methylates 5'-CG-3' dinucleotide, the cytosine methylation occurs within the 5'-CCTGG-3' pentanucleotides [29]. Another reported example comes from human primary effusion lymphoma (PEL), where the B cell-specific B29 gene is silenced. Bisulfite sequencing revealed two types of DNA methylation at a conventional CpG and at a CCW(A/T)GG site in the B29 promoter. Methylation of the CCW(A/T)GG site significantly repressed transcriptional activity *in vivo* and blocked the binding of early B cell factor [25]. CCW(A/T)GG methylation as well as normal CpG methylation has also been reported and implicated as a reason for silencing integrated provirus after retroviral infection of murine erythroleukemia (MEL) cells [25]. Information on the genomic distribution of CCWGG methylation is rarely addressed. Recently, it has been found that the ICSBP/IRF8 gene, a member of the interferon regulatory factor (IRF) family of transcriptional regulators expressed in monocytic and lymphocytic cells, contains a large number of the CCWGG sequences in the promoter region as well as in the coding region/exons. The CpG sites in the promoter region were hypermethylated, while no CCWGG sites were methylated and a change in histone modification was found [32,33]. Nevertheless, the heavily methylated gene was still expressed. This was reported to be due to a higher expression of H3K9-ac then the H3K9me3. These histone modifications can over-ride the silencing effect of DNA methylation at the promoter, thereby permitting transcription of the ICSBP/IRF8 gene [33].

Differently, in the MEL cells transfected with Moloney murine leukemia virus (M-MuLV) encoding GFP, dynamic analysis of proviral induction and *de novo* methylation demonstrated a histone deacetylase-independent, methylation density-dependent mechanism of transcriptional repression [31]. In both above-mentioned situations, the analyzed genes all contained CCW(A/T)GG sequences. The role of these sequences in the regulation of gene expression is not known. In contrast to the published results, our present study of the relationship between the CCWGG methylation and the gene expression shows that the GFP expression driven by the CMV promoter seemed not to be affected by the methylation of this site in the CMV promoter region.

Taking all above discussed issues together we can conclude that genes containing CCW(A/T)GG sequences in their promoter regions may be controlled by opposing mechanisms leading to varying effects on cellular functions. It is important to pay attention to the DNA methylation of the CCW(A/T)GG sequences when the expression and regulation of these genes are analyzed. Studies on interaction between transcription factors and the CCW(A/T)GG sequences will therefore be helpful to elucidate the mechanism of CCW(A/T)GG methylation and gene regulation in mammals. For therapeutic purposes, it may be wise to propagate the plasmid vector DNA in dcm- bacteria, avoiding the target gene silencing caused by methylation of the EcoRII site.

The methylation status of the transcription factor recognition sequences in the CMV promoter region was also analyzed in the present study. The human CMV promoter is rich in enhancer elements, which are recognized by transcription factors. It consists of at least four types of repetitive sequence elements, referred to as the 17-, 18-, 19- and 21-bp repeats, which are present three to five times within the promoter/enhancer region of the CMV promoter and that can form complexes with nuclear proteins [34,35]. The 18- and 19-bp repeats contain consensus binding sites for NFkB, CREB/ATF and Ap1, respectively, and were shown to mediate enhancement of the CMV promoter activity by these transcription factors [36]. The 17-bp repeat was suggested to bind to the transcription factor NF-1 [37,38]. The 21-bp repeat binds to a negative regulator, specific for undifferentiated cells as well as to the transcription factor YY1 and it was suggested to repress CMV promoter-dependent transcription [39]. Other factors, which bind to the CMV promoter, are SP1 and MDBP [40,41]. The recognition sites for transcription factors such as NFkB (5'-GGG G/A G/A/C/ TTTCC-), CREB/ATF (5'-GACGTCA-), and SP-1 (5'-CCCGCC-) in the CMV promoter region have been previously determined [36,38,40]. The transcription factor NFkB has been shown to play an important role in activation of the CMV promoter both in the mouse liver and in hepatocyte-derived cell lines *in vitro* [16]. Our methylation study on the CMV promoter shows no methylation at 3 of 4 NFkB sites (C118, C102, C81-83), that are close to the transcription start, in the 5 analyzed clones and only partial methylation at the distal NFkB site (C43-45) in the 2 of 4 analyzed clones (Figure 2B). Among three CREB/ATF sites, one (C66) was not methylated in any of six analyzed clones, and two other sites were partially methylated (1/4 and 3/4 analyzed clones respectively). The Ap1 site was methylated in only one of six clones. This implies that methylation might not prevent these sites from binding of the transcription factors; the promoter activity was therefore maintained. It has been known that NFkB is highly expressed in cancer cells. Misregulated NFkB activates genes involved in the cell proliferation and protects the cell from conditions that would otherwise cause cell apoptosis [42]. NFkB is therefore the subject of active research as a target for anti-cancer therapy [43].

The mechanism of lowering the *de novo* methylation level in the integrated CMV promoter is not known. One possibility is demethylation, which would be similar to that in pSV-CAT and SV40 early promoter/enhancer [44]. The SV40 promoter is frequently demethylated in the HeLa cell line upon stable transfection resulting in resistance to silencing. Stably transfected DNA might have more chromatin-like interactions with histone and non-histone DNA binding proteins than transiently transfected DNA.

## Acknowledgements

The work was supported by: The Hans and Märit Rausing Charitable Trust, Lilly and Sven Lawskis Foundation, The Sten Lexner's Fund, and The Swedish Childhood Cancer Foundation.

## References

1. Jeltsch A (2002) Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. Chembiochem 3: 274-293.
2. Worm J, Guldberg P (2002) DNA methylation: an epigenetic pathway to cancer and a promising target for anticancer therapy. J Oral Pathol Med 31: 443-449.
3. Doerfler W (2005) On the biological significance of DNA methylation. Biochemistry (Mosc) 70: 505-524.
4. Lingworth R, Kerr A, DeSousa D, Ellis P, Clee C, et al. (2008) A novel CpG island set identifies tissue-specific methylation at developmental gene loci. PLoS Biol 6: e22.
5. Stein R, Razin A, Cedar, H (1982) *In vitro* methylation of the hamster adenine phosphoribosyl transferase gene inhibits its expression in mouse L cells. Proc Natl Acad Sci USA 79: 3418-3422.
6. Herman JG, Latif F, Weng Y, Zbar B, Liu S, et al. (1994) Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. Proc Natl Acad Sci USA 91: 9700-9704.
7. Curradi M, Izzo A, Badaracco G, Landsberger N (2002) Molecular mechanisms of gene silencing mediated by DNA methylation. Mol Cell Biol 22: 3157-3173.
8. Baylin SB, Höppener JW, de Bustros A, Steenbergh PH, Lips CJ, et al. (1986) DNA methylation patterns of the calcitonin gene in human lung cancers and lymphomas. Cancer Res 46: 2917-2922.
9. Jones PA, Baylin SB (2002) The fundamental role of epigenetic events in cancer. Nat Rev Geneti 3: 415-428.
10. Feinberg AP, Vogelstein B (1983) Hypomethylation of ras oncogenes in primary human cancers. Biochem Biophys Res Commun 111: 47-54.
11. Tao L, Yang S, Xie M, Kramer PM, Pereira MA (2000) Hypomethylation and overexpression of c-jun and c-myc protooncogenes and increased DNA methyltransferase activity in dichloroacetic and trichloroacetic acid-promoted mouse liver tumors. Cancer Lett 158: 185-193.
12. Sakai T, Toguchida J, Ohtani N, Yandell DW, Rapaport JM, et al. (1991) Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. Am J Hum Genet 48: 880-888.
13. Dammann R, Li C, Yoon JH, Chin PL, Bates S, et al. (2000) Epigenetic inactivation of a RAS association domain family protein from the lung tumour suppressor locus 3p21.3. Nature Genetics 25: 315-319.
14. van Engeland M, Roemen GM, Brink M, Panchen MM, Arends JW, et al. (2002) K-ras mutations and RASSF1A promoter methylation in colorectal cancer. Oncogene 21: 3792-3795.
15. Feinberg AP, Ohlsson R, Henikoff S (2006) The epigenetic progenitor origin of human cancer. Nat Rev Genet 7: 21-33.
16. Löser P, Jennings GS, Strauss M, Sandig V (1998) Reactivation of the previouslysilenced cytomegalovirus major immediate-early promoter in the mouse liver: involvement of NFkB. J Virol 72: 180-190.
17. Prosch S, Stein J, Staak K, Liebenthal C, Volk HD, et al. (1996) Inactivation of the very strong HCMV immediate early promoter by DNA CpG methylation *in vitro*. Biol Chem Hoppe Seyler 377: 195-201.
18. Collas P (1998) Modulation of plasmid DNA methylation and expression in zebrafish embryos. Nucleic Acids Res 26: 4454-4461.
19. Brooks AR, Harkins RN, Wang P, Qian HS, Liu PX, et al. (2004) Transcriptional silencing is associated with extensive methylation of the CMV promoter following adenoviral gene delivery to muscle. The Journal of Gene Medicine 6: 395-404.
20. Herman JG, Graff JR, Myöhänen S, Nelkin BD, Baylin SB (1996) Methylation-specific PCR: A novel PCR for methylation status of CpG islands. Proc Natl Acad Sci USA 93: 9821-9826.
21. Jiang C, Xuan Z, Zhao F, Zhang MQ (2007) TRED: a transcriptional regulatory element database, new entries and other development. Nucleic Acids Res 35: D137-D140.
22. Zhang M (2008) Transcriptional regulatory element database. Cold Spring Harbor Laboratory (http://rulai.cshl.edu/TRED)
23. He TC, Zhou S, da Costa LT, Yu J, Kinzler KW, et al. (1998) A simple system for recombinant adenoviruses. Proc Natl Acad Sci USA 95: 2509-2514.

24. Salford LG, Siesjö P, Skagerberg G (2002) Search for effective therapy against glioblastoma multiforme - clinical immunisation with autologous glioma cells transduced with the human Interferon-gamma gene. Developments in Neuroscience, Excerpta Medica Int. Congr. Ser. 1247, Elsevier, Ed.K.Watanade before. pp 221-220.

25. Malone CS, Miner MD, Doerr JR (2001) CmC(A/T)GG DNA methylation in mature B cell lymphoma gene silencing. Proc Natl Acad Sci USA 98: 10404-10409.

26. Choi KH, Basma H, Singh J, Chen PW (2005) Activation of CMV promoter controlled glycosyltransferase and β -galactosidase glycogenes by butyrate, tricostatin A and 5-Aza-2′-deoxycytidine. Glycoconjugate Joul 22: 63-69.

27. Meilinger D, Fellinger K, Bultmann S, Rothbauer U, Spada F, et al. (2009) Np95 interacts with *de novo* DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells. EMBO Rep 10: 1259-1264.

28. Kong Q, Wu M, Wang Z, Li L, Liu X, et al. (2011) Effect of trichostatin A and 5-Aza-2'-deoxycytidine on transgene reactivation and epigenetic modification in transgenic pig fibroblast cells. Mol Cell Biochem 355: 157-165.

29. Franchina M, Kay PH (2000) Evidence that cytosine residues within 5'-CCTGG-3' pentanucleotides can be methylated in human DNA independent of the methylating system that modifies 5'-CG-3' dinucleotides. DNA Cell Biol 19: 521-526.

30. Watson B, Munson K, Clark J, Shevchuk T, Smith SS (2007) Distribution of CWG and CCWGG in the human genome. Epigenetics 2: 151-154.

31. Lorincz MC, Schubeler D, Gocke SC, Walters M, Groundine M, et al. (2000) Dynamic analysis of proviral induction and de novo methylation demonstrated a histone deacetylase-independent, methylation density-dependent mechanism of transcriptional repression. Mol Cell Biol 20: 842-850.

32. Tshuikina M, Jerberg-Wiklund H, Nilsson K, Öberg F (2008) Epigenetic silencing of the interferon regulatory factor ICSBP/IRF8 in human multiple myeloma. Experimental Hematology 36: 1673-1681.

33. Tshuikina M, Nilsson K, Öberg F (2008) Positive histone marks are associated with active transcription from a methylated ICSBP/IRF8 gene. Gene 410: 259-267.

34. Boshart M, Weber F, Jahn J, Dorsch-Häsler K, Fleckenstein B, et al. (1985) A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. Cell 41: 521-530.

35. Ghazal P, Lubon H, Fleckenstein B, Hennighausen L (1987) Binding of transcription factors and creation of a large nucleoprotein complex on the human cytomegalovirus enhancer. Proc Natl Acad Sci USA 84: 3658-3662.

36. Sambucetti LC, Cherrington JM, Wikinson GWG, Mocarski ES (1989) NF-kappa B activation of the cytomegalovirus enhancer is mediated by a viral transactivator and by T-cell stimulation. EMBO J 8: 4251-4258.

37. Niller HH, Hennighausen L (1991) Formation of several specific nucleoprotein complexes on the human cytomegalovirus immediate early enhancer. Nucleic Acids Res 19: 3715-3721.

38. Baskar JF, Smith PP, Nilaver G, Jupp RA, Hoffmann S, et al. (1996) The enhancer domain of the human cytomegalovirus major immediate-early promoter determines cell type-specific expression in transgenic mice. Journal of Virology 70: 3207-3214.

39. Kothari S, Baullie J, Sissons JGP, Sinclair JH (1991) The 21 bp repeat element of human cytomegalovirus major immediate early enhancer is a negative regulator of gene expression in undifferentiated cells. Nucleic Acids Res 29: 1767-1771.

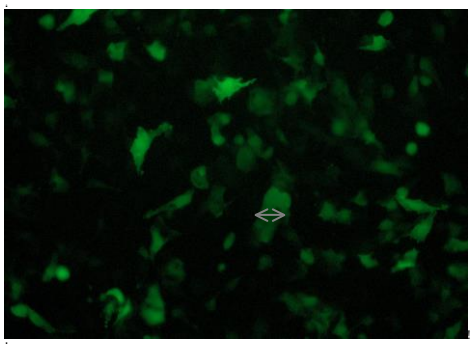
40. Lang D, Fickenscher H, Stamminger T (1992) Analysis of protein binding to the proximal promoter region of the human cytomegalovirus IE-1/2 enhancer/promoter reveals both consensus and aberrant recognition sequences for transcription factor Sp1 and CREB. Nucleic Acids Res 20: 3287-3295.

41. Zhang XY, Inamdar NM, Supakar PC, Wu K, Ehrlich KC, et al. (1991) Three MDBP sites in the immediate early enhancer-promoter region of human cytomegalovirus. Virology 182: 865-869.

42. Rayet B, Gelinas C (1999) Aberrant Rel/Nfkb genes and activity in human cancer. Oncogene 18: 6938-6947.

43. Escárcega RO, Fuentes-Alexandro S, García-Carrasco M, Gatica A, Zamora A (2007) The transcription factor nuclear factor-kappa B and and cancer. Clinical Oncology 19: 154-161.

44. Qu GZ, Ehrlich M (1999) Demethylation and expression of methylated plasmid DNA stably transfected in Hela cells. Nucl Acids Res 27: 2332-2338.



<-> 20 uM

**Figure 1A:** GFP positive cells among transfected GA49 cells under fluorescence light. The cells were co-transfected with Plasmid pAdTrack-CMV-GFP and pLXSN DNAs, and grown for 5 months in the selection medium.

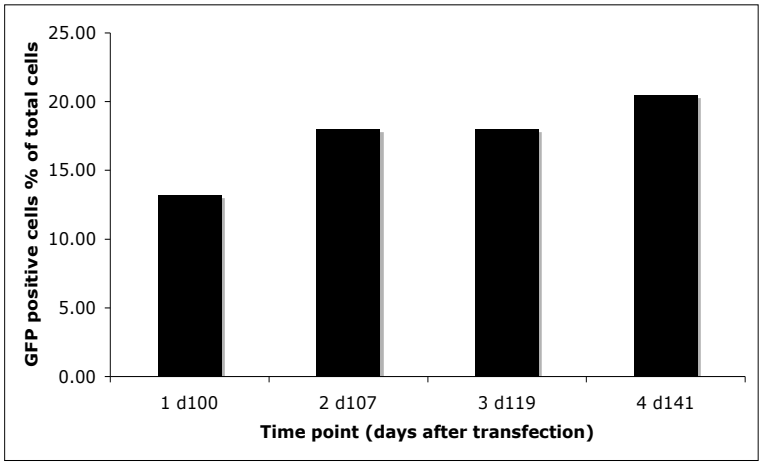



**Figure 1B:** Time curse of flow cytometric analysis of the transfected GA49 cells. The Cells were collected and used for flow cytometric analysis and DNA isolation.

**-652 GCATTCTAGT TGTGGTTTGT CCAAACTCAT CAATGTATCT TAACGCGGAG**

**-602 GTTTATCGAC GATCTGCTAG TGATTAATAG TAATCAATTA CGGGGTCATT**

**>-573**

**-552 AGTTCATAGC CCATATATGG AGTTCCGCGT TACATAACTT ACGGTAAATG**

**-502 GCCCGC*CTGG CTGACCGCCC AACGACCCCC GCCCATTGAC GTCAATAATG**

**-452 ACGTATGTTC CCATAGTAAC GCCAATAGGG ACTTTCCATT GACGTCAATG**

**-402 GGTGGAGTAT TTACGGTAAA CTGCCCACTT GGCAGTACAT CAAGTGTATC**

**-352 ATATGCCAAG TACGCCCCCT ATTGACGTCA ATGACGGTAA ATGGCCCGC*C**

**-302 TGGCATTATG CCCAGTACAT GACCTTATGG GACTTTCCTA CTTGGCAGTA**

**-252 CATCTACGTA TTAGTCATCG CTATTACCAT GGTGATGCGG TTTTGGCAGT**

**-202 ACATCAATGG GCGTGGATAG CGGTTTGACT CACGGGGATT TCCAAGTCTC**

**-152 CACCCCATTG ACGTCAATGG GAGTTTGTTT TGGCACCAAA ATCAACGGGA**

**-102 CTTTCCAAAA TGTCGTAACA ACTCCGCCCC ATTGACGCAA ATGGGCGGTA**

**-52 GGCGTGTACG GTGGGAGGTC TATATAAGCA GAGCTGGTTT AGTGAACCGT**

**-2 CAGATCCGCT AGCGCTACCG GTCGCCACCA TGGTGAGCAA GGGCGAGGAG**

**<-1 Transcription start->**

**+49 CTGTTCACCG GGGTGGTGCC CATCCTGGT**

**Figure 2A:** Nucleotide sequence of the CMV promoter region chosen for analysis of the methylation status in the transformed human cancer (GBM) cells. > < Sequence region for determination of the methylation status; start codon, TATA box, and CCTGG motif are underlined; methylated cytosine in the CCTGG.


**Figure 2B:** Methylation status in the CMV promoter region in the GA clones transformed with pAdtrack-CMV-GFP plasmid. ■ : CpG, C methylated; □ : CpG, C non-methylated; ■ : CpN (CpA, CpC, CpT), C methylated; □ : CpN (CpA, CpC, CpT), C non-methylated; ■ : CCTGG, second C methylated; □ : CCTGG, second C non-methylated.

```
BTC T1F1-5   -573                                AGTAATCAATTATGGGGTTATT
BTC T1F1-15                                      AGTAATCAATTACGGGGTCATT
BTC T1F3-2                                       AGTAATCAATTATGGGGTTATT
BTC T1F3-10                                      AGTAATTAATTATGGGGTTATT
BNT DNA      -601 GTTTATCGACGATCTGCTAGTGATTAATAGTAATCAATTACGGGGTCATT
BTC T1F1-5   -551 AGTTTATAGTTTATATATGGAGTTCCGCGTTACATAACTTACGGTAAATG
BTC T1F1-15       AGTTTATAGCCTATATATGGAGTTCCGCGTTATATAACTTACGGTAAATG
BTC T1F3-2        AGTTTATAGTTTATATATGGAGTTCCGCGTTACATAACTTACGGTAAATG
BTC T1F3-10       AGTTTATAGTTTATATATGGAGTTTTGTGTTATATAATTTATGGTAAATG
BNT DNA           AGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATG
BTC T1F1-5   -501 GTTTGTCTGGTTGATTGTTTAACGATTTTTGCCCATTGACGTCAATAATG
BTC T1F1-15       GTTTGTCTGGTTGATTGTTTAACGATCTTTGTTTATTGATGTTAATAATG
BTC T1F3-2        GTTTGTCTGGTTGATTGTTTAACGATTTTTGCCCATTGACGTCAATAATG
BTC T1F3-10       GTCCGCCTGGCTGACCGTTTAACGATCTTTGTTTATTGACGTCAATAATG
BNT DNA           GCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATG
BTC T1F1-5   -451 ACGTATGTTTCCATAGTAACGCCAATAGGGATTTTTTATTGATGTTAATG
BTC T1F1-15       ACGTATGTTTTCATAGTAACGTTAATAGGGACTTTCCATTGATGTCAATG
BTC T1F3-2        ACGTATGTTTCCATAGTAACGCCAATAGGGATTTTTTATTGATGTTAATG
BTC T1F3-10       ATGTATGTTTCCATAGTAACGCCAATAGGGATTTTCCATTGACGTCAATG
BNT DNA           ACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATG
BTC T1F1-5   -401 GGTGGAGTATTTATGGTAAATTGTTTATTTGGTAGTATATTAAGTGTATT
BTC T1F1-15       GGTGGAGTATTTACGGTAAATTGCCCACTTGGTAGTATATCAAGTGTATT
BTC T1F3-2        GGTGGAGTATTTATGGTAAATTGTTTATTTGGTAGTATATTAAGTGTATT
BTC T1F3-10       GGTGGAGTATTTATGGTAAACTGCCTACTTGGCAGTATATCAAGTGTATC
BNT DNA           GGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATC
BTC T1F1-5   -351 ATATGTTAAGTATGTTTTTTATTGATGTTAATGATGGTAAATGGTTTGTC
BTC T1F1-15       ATATGCTAAGTACGTTTTTTATTGATGTTAATGATGGTAAATGGTTTGTC
BTC T1F3-2        ATATGTTAAGTATGTTTTTTATTGATGTTAATGATGGTAAATGGTTTGTC
BTC T1F3-10       ATATGCCAAGTATGTTTTTTATTGATGTCAATGACGGTAAATGGCCCGTC
BNT DNA           ATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCC
BTC T1F1-5   -301 TGGTATTATGTTTAGTATATGATTTTATGGGATTTTTTTATTTGGTAGTA
BTC T1F1-15       TGGTATTATGTTTAGTATATGATTTTATGGGATTTTTTTATTTGGTAGTA
BTC T1F3-2        TGGTATTATGTTTAGTATATGATTTTATGGGATTTTTTTATTTGGTAGTA
BTC T1F3-10       TGGTATTATGTCCAGTATATGATTTTATGGGATTTTTTTACTTGGCAGTA
BNT DNA           TGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTA
BTC T1F1-5   -251 TATTTATGTATTAGTTATTGTTATTATTATGGTGATGTGGTTTTGGTAGT
```

```
BTC T1F1-15       TATTTACGTATTAGTTATTGTTATTATTATGGTGATGTGGTTTTGGTAGT

BTC T1F3-2        TATTTATGTATTAGTTATCGTTATTATTATGGTGATGTGGTTTTGGTAGT

BTC T1F3-10       CATTTATGTATTAGTTATTGTTATTATTATGGTGATGTGGTTTTGGTAGT

BNT DNA           CATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGT

BTC T1F1-5   -201 ATATTAATGGGTGTGGATAGTGGTTTGATTCATGGGGATTTTTAAGTTTT

BTC T1F1-15       ATATTAATGGGTGTGGATAGTGGTTTGATTCATGGGGATTTTTAAGTTTT

BTC T1F3-2        ATATTAATGGGTGTGGATAGTGGTTTGATTCATGGGGATTTTTAAGTTTT

BTC T1F3-10       ATATTAATGGGTGTGGATAGTGGTTTGATTCATGGGGATTTTTAAGTTTT

BNT DNA           ACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC

BTC T1F1-5   -151 TATTTTATTGATGTTAATGGGAGTTTGTTTTGGTATTAAAATTAATGGGA

BTC T1F1-15       TATTTTATTGATGTTAATGGGAGTTTGTTTTGGTATTAAAATTAATGGGA

BTC T1F3-2        TATTTTATTGATGTTAATGGGAGTTTGTTTTGGTATTAAAATTAATGGGA

BTC T1F3-10       TATTTTATTGATGTTAATGGGAGTTTGTTTTGGTATTAAAATTAATGGGA

BNT DNA           CACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGA

BTC T1F1-5   -101 TTTTTTAAAATGTTGTAATAATTTTGTTTTATTGATGTAAATGGGTGGTA

BTC T1F1-15       TTTTTTAAAATGTTGTAATAATTTTGTTTTATTGATGTAAATGGGTGGTA

BTC T1F3-2        TTTTTTAAAATGTTGTAATAATTTTGTTTTATTGATGTAAATGGGTGGTA

BTC T1F3-10       TTTTTTAAAATGTTGTAATAATTTTGTTTTATTGATGTAAATGGGTGGTA

BNT DNA           CTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTA

BTC T1F1-5   - 51 GGTGTGTATGGTGGGAGGTTTATATAAGTAGAGTTGGTTTAGTGAATTGT

BTC T1F1-15       GGTGTGTATGGTGGGAGGTTTATATAAGTAGAGTTGGTTTAGTGAATTGT

BTC T1F3-2        GGTGTGTATGGTGGGAGGTTTATATAAGTAGAGTTGGTTTAGTGAATTGT

BTC T1F3-10       GGTGTGTATGGTGGGAGGTTTATATAAGTAGAGTTGGTTTAGTGAATTGT

BNT DNA           GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGT

BTC T1F1-5    - 1 TAGATTTGTTAGTGTTATTGGTTGTTATTATGG

BTC T1F1-15       TAGATTTGTTAGTGTTATTGGTTGTTATTATGG

BTC T1F3-2        TAGATTTGTTAGTGTTATTGGTTGTTATTATGG

BTC T1F3-10       TAGATTTGTTAGTGTTATTGGTTGTTATTATGG

BNT DNA           CAGATCCGCTAGCGCTACCGGTCGCCACCATGG
```

**Figure 2C:** Sequence comparison of the bisulfite treated and non-treated samples in the CMV promoter region. BTC: Bisulfite treated clone DNA sample; BNT: Bisulfite non-treated DNA sample. C: methylated; C: CCTGG, second C methylated.

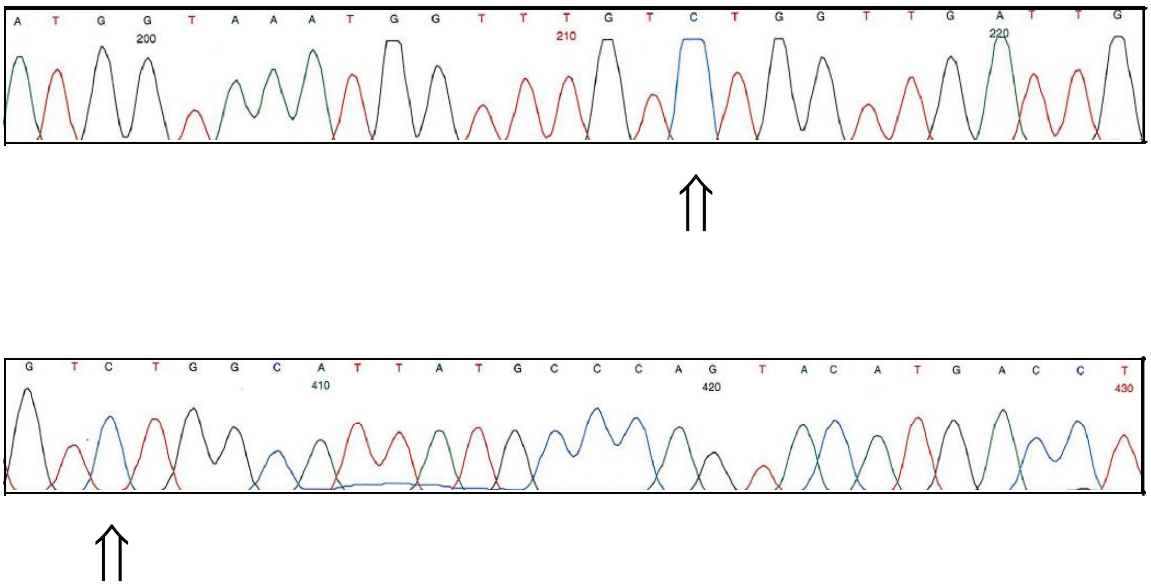



**Figure 2D:** Bisulfite sequencing of the representative clone T1F1-5 shows the dcm pattern ($C^{m}CTGG \rightarrow TCTGG$, 212-216n and 404-408n).


**Figure 2E:** Methylation status in the CMV promoter region in the *E.coli* DH5-α (dcm+) and *E.coli* INV110 (dcm-) clones transformed with pAdtrackcmv-GFP plasmid. ■□□ : CpG, C methylated; □□□ : CpG, C non-methylated; ■ : CpN (CpA, CpC, CpT), C methylated; □ : CpN (CpA, CpC, CpT), C non-methylated; ■□ : CCTGG, second C methylated; □□ : CCTGG, second C non-methylated.

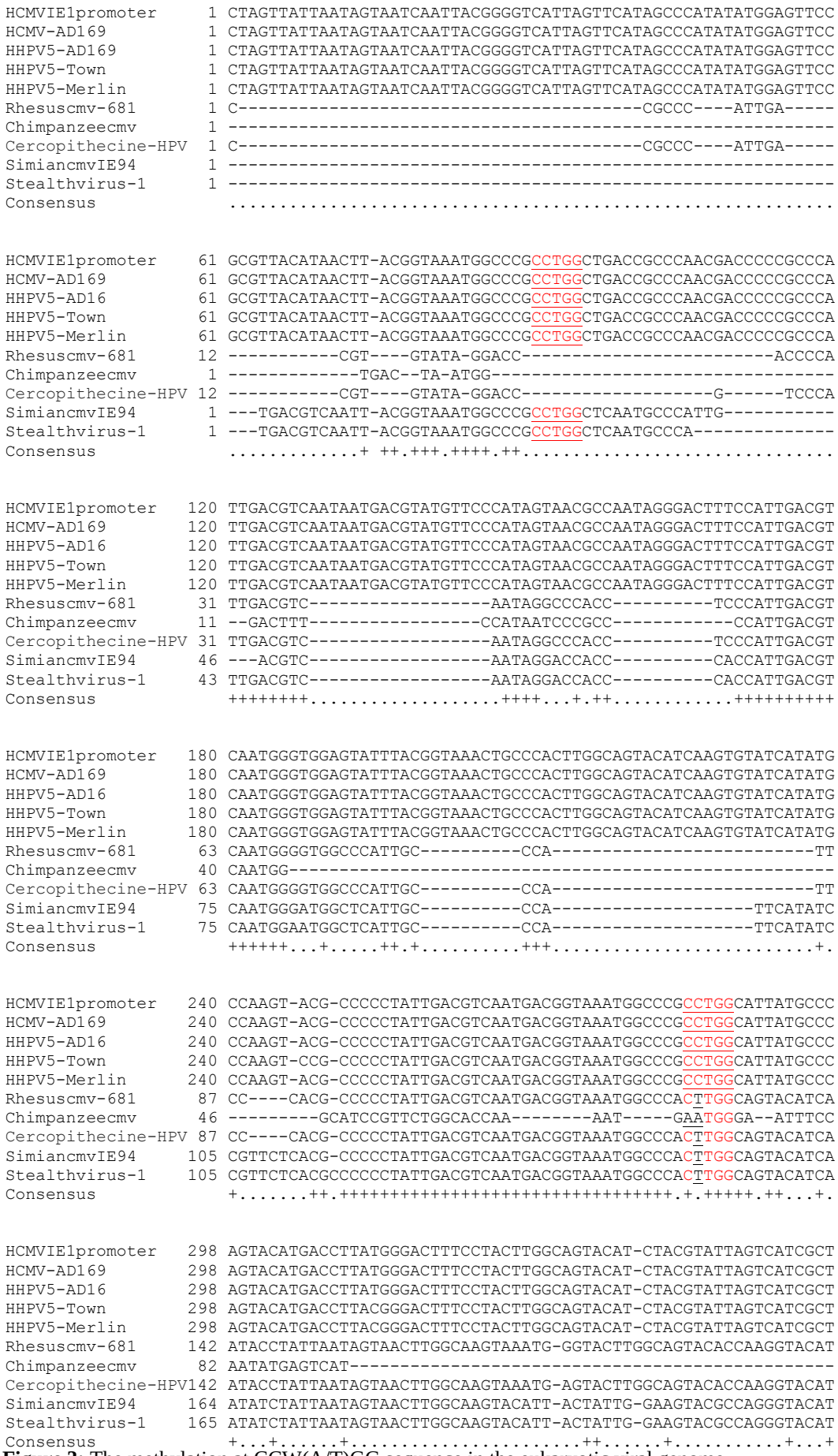



**Figure 3:** The methylation at CCW(A/T)GG sequence in the eukaryotic viral genome.

| Transcription Factor | Binding sequence | Regulated gene | Regulated gene on chromosome | Species |
|---|---|---|---|---|
| NFKB1 | AGTAGGGTTTTCCCCAGGA | Cxc19 | 5 | Mouse |
| BCL6 | CCAGGGAGTGACTTTCCGA-GGAAGGCATTTCGGAGAAG-ACGGGGGT | BCL2L1 | 20 | Human |
| POU2F1( OCT1, OTF1) | CGCCCTCCGGGCTCAATATG-CAAATCCGAGCACCAGGAA-GTAGCTGG | LYN | 8 | Human |
| E2F1 | CTCGTGGCGCCCCAGGG | CYC1 | 8 | Human |
| ESR1,ESR2 | CACCAGGTGGCCCTGACCC-TGGGA | C3 | 1 9 | Human |
| TFAP2A(AP-2, AP2TF, TFAP2)(Mus) | GAGGCAGACCACGTGAGAG-CCTGGCCAGGCCTTCC | CCNB1 | 5 | Human |
| Trp53(p53) (Mus) | CTCAGACATGTCTGGAGACC-CTAGGACGACAAGCCCAGG | Apaf-1,CED4 | 1 | Human |
| Sp1(Mus) | GATTCTTTCCCCGCCCTCCTC-TCGCCCCAGGGACA | TNF, DIF CACHECTIN | 6 | Human |
| TFAP2A, AP-2 | CCCCAGGC | CD7 | 17 | Human |
| Sp1 | TTCCCTTCCTCCGAAACT-TCCCAGGAGAAGGGCGG-AAGACCCCAGGGGAAGG-GGCGAGGCGAATCTTCGCG | HTR1A | 5 | Human |
| E2F-4 | CCCGCCAGG | TK1 | 17 | Human |
| HIF1A | CTTCACGTGCGGGGACCAG-GGACCGT | ALDOA | 16 | Human |
| Sp1 | GGAGAGGGCGGGACCAGG-ATGGGAGGTGTCTGTCACG- | ALdoc | 10 | Rat |

| | | | | |
|---|---|---|---|---|
| | CCCCCAGGGAG | | | |
| Amt (Mus) | CCTCCAGGCTCTTCTCACGC- | Cyp1a1 | 8 | Rat |
| | AACTC | | | |
| Pparg (Mus) | TCATCCCAGGGCAAAGTACA | Slc2a2 | 2 | Rat |
| Egr1 (Mus) | ACCAGTCCTGGGGAGAGGG- | ALdoc | 10 | Rat |
| | CGGGACCAGGA | | | |
| STAT5B | GGGCAGTTCCAGGAATCGG- | Scos3 | 11 | Mouse |
| | GGGGC | | | |
| STAT3 | TCTTGACGTCACGCACTGCC- | Stat3 | 11 | Mouse |
| | AGGAACT | | | |
| Nfic (Mus) | TTCCAAATTGGGGGCCGGG- | Colla1 | 11 | Mouse |
| | CCAGGCA | | | |
| Tcfap2a (Mus) | GGGCCAGGTGACCT | Hgf | 5 | Mouse |
| Nfic (Mus) | ATTGCAGCTGGCCTCGGG- | | | Mouse |
| Pparg, Usf1 | CCAGGTGACCTTT | Hgf | 5 | |
| Ets1, Tpl1, (Mus) | GCGTCCAGGAAGCCTG | Timp1 | X | Mouse |
| Ets-1, MGC18571, | | | | |
| p42Ets-1,p51Ets-1 | | | | |

**Figure 4:** Transcription factors involved in the interaction with CCAGG sequence and the corresponding regulated genes.

| Transcription Factor | Binding sequence | Regulated gene | Regulated gene on chromosome | Species |
|---|---|---|---|---|
| H1F1A | TGGCCAGACGTGCCTGG- | BHLHB2 | 3 | Human |
| | AGTCACAGGGTAG | | | |
| WT1 | GGAAGGTCCGCCCTCTC- | FOXD1 | 5 | Human |
| | CTGGACTC | | | |

| | | | | |
|---|---|---|---|---|
| TP53 | TCACAAGTTAGAGACAA-GCCTGGGCGTGGGCTAT-ATTG | BAX | 19 | Human |
| STAT1 | ACTTCCTGGAAAT | CCL2 | 17 | Human |
| Cebpa (Rat) | GACCTTTTGCAATCCTGG | APOB | 2 | Human |
| Egr1 (Mus) | TGCCTTCGCCCCCGCCTG-GCGG | SYN1 | X | Human |
| TFAP2A | GAGGCAGACCACGTGAG-AGCCTGGCCAGGCCTTCC | CCNB1 | 5 | Human |
| E2F-4 | CCTGGCGCCCAC | MTHFR | 1 | Human |
| Sp1 | CCTGGCTCCTCCCCTCCT-CCG | PSEN1 | 14 | Human |
| ESR1 | CACCAGGTGGCCCTGAC-CCTGGGA CACCAGGTG-GCCCTGACCCTGGGA | C3 | 19 | Human |
| RARA | CCAGGGCTTGCCCTGGG-TTAAGAGCC | SFTPB | 2 | Human |
| Sp1 (Mus) | TTCTGGGCGGAAACTTCC-TGGG | Tnfsf6 | 1 | Mouse |
| Egr1 | ACCAGTCCTGGGGAGAGG-GCGGGACCAGGA | Aldoc | 10 | Rat |
| CREB1 | GAGGGGCTTTGACGTCAG-CCTGGCC | Th | 1 | Rat |

**Figure 5:** Transcription factors involved in the interaction with CCTGG sequence and the corresponding regulated genes.