Conformation-selective methylation of geminiviral DNA

Paprotka, T.¹, Deuschle, K., Metzler, V., Jeske, H*)

Biologisches Institut, Abteilung für Molekularbiologie und Virologie der Pflanzen,
Universität Stuttgart, Pfaffenwaldring 57, D-70550 Stuttgart, Germany

*)Author for correspondence:
E-mail: holger.jeske@bio.uni-stuttgart.de
Phone: +49-711-685-65070
Fax: +49-711-685-65096

¹ present address: Viral Mutation Section, HIV Drug Resistance Program, National Cancer Institute at Frederick, Frederick, MD 21702, USA

Running title: Bisulfite sequencing

Keywords: geminivirus; begomovirus; bisulfite sequencing; rolling circle amplification, methylation

Abstract: 200 words

Figures: 6
Tables: 0
Abstract

Geminiviruses with small circular single-stranded DNA genomes replicate in plant cell nuclei using various double-stranded DNA intermediates: distinct open circular and covalently closed circular as well as heterogeneous linear DNA. Their DNA may be methylated partially at cytosine residues, which has been detected previously by bisulfite sequencing and subsequent polymerase chain reaction. In order to determine the methylation patterns of the circular molecules, the DNA of tomato yellow leaf curl Sardinia virus and Abutilon mosaic virus were investigated utilizing bisulfite treatment followed by rolling circle amplification. Shotgun sequencing of the products yielded a randomly distributed fifty percent rate of C-maintenance after bisulfite reaction for both viruses. However, controls with unmethylated single-stranded bacteriophage DNA resulted in the same level of C-maintenance. Only one short DNA stretch within the C2/C3 promoter of TYLCSV showed a hyperprotection of C exceeding the threshold of the mean value plus one standard deviation. Similarly, the use of methylation-sensitive restriction enzymes suggested that geminiviruses escape silencing by methylation very efficiently either by rolling circle or by recombination-dependent replication modes. In contrast, attempts to detect methylated bases positively using methyl-cytosine-specific antibodies detected methylated DNA only in heterogeneous linear dsDNA and methylation-dependent restriction enzymes revealed that the viral heterogeneous linear dsDNA was methylated preferentially.
Introduction

Methylation of DNA is commonly investigated using methylation-sensitive restriction enzymes and bisulfite sequencing which converts unmethylated cytosines into uracil. The products are usually amplified with polymerase chain reaction (PCR) and sequenced after bacterial cloning (14, 23). The critical steps in various protocols are the complete conversion of unmethylated cytosines to uracil (26, 78) and the design of appropriate primers for the converted DNA (reviewed by 58, 85). Several bioinformatics tools have been developed to analyse the results (9, 46–48, 66, 71), including recent improvements for the analysis of plant DNA cytosine methylation (27, 35).

Eukaryotes methylate cytosine at C^5 by methyltransferases which differ between plants, fungi and mammals (reviewed by 25). In mammals, usually symmetric CpG sites are preferred, whereas nearly every cytosine residue in plant DNA can be methylated (reviewed by 5). As a result, only 2 to 8% of mammalian DNA is methylated in comparison to up to 50% in higher plants (reviewed by 84). The methylation status is highly dynamic. In plants, demethylation is mediated by DNA glycosylases and strand cleavage (84), a mechanism which would inactivate single-stranded (ss) DNA. In addition to C methylation, N^6-adenine methylation is well-known from bacteria, but described rarely for plants (19, 74).

The circular ssDNA-containing geminiviruses (reviewed by 37) are important plant pathogens causing severe damage in crop plants all over the world (49). Their DNA is packed into icosahedral twin particles (8, 83). Four geminivirus genera have been described, which differ in genome organization, host range and insect vector (67). After insect vectors have inoculated the plant, viral particles are disassembled (42) and the ssDNA is copied by complementary strand replication (CSR) to double-stranded (ds) covalently closed circular (ccc) DNA (31), which is packaged into nucleosomes to form minichromosomes (1, 55, 56). Further progeny DNA is replicated by rolling circle replication (RCR) as well as by
recombination-dependent replication (RDR, reviewed by 38), generating high amounts of heterogeneous linear dsDNA (hds).
Both geminiviral circular and linear as well as ssDNA and dsDNA may serve as template for PCR during bisulfite sequencing in total nucleic acids preventing the assignment of methylated nucleotides to a specific DNA conformation. Cloning, diagnosis, infection and direct sequencing of viruses with small circular DNA genomes has been considerably improved by rolling circle amplification (RCA) with bacteriophage Phi29 polymerase (29, 36, 39, 52-54, 82, and references therein). In combination with random hexamer primers, circular DNA can be amplified preferentially from plant samples without any prior knowledge of the sequence. Therefore, it was intriguing to test this technique for the identification of C-modifications by restriction analysis and bisulfite sequencing.
Geminiviruses induce small interfering RNAs (siRNA) (2, 7, 62, 73), which may lead to posttranscriptional (PTGS) as well as transcriptional (TGS) gene silencing (4, 65). This interplay of PTGS and TGS has been understood as part of the plant defense pathways (reviewed by 16, 64, 75, 76). Correspondingly, geminiviruses have developed potent suppressors of gene silencing (reviewed by 6, 61). Promoting PTGS and TGS in host plants against viruses to establish novel resistance strategies is promising, but also challenging (reviewed by 72). Moreover, both silencing routes may be used to knock-down host genes with the help of geminiviral vectors (virus-induced gene silencing, VIGS) (12, 13, 20, 43, 45, 51). In this context, the analysis of the methylation status of geminiviral episomal or integrated DNA has revealed important insights (4, 11, 28, 60, 62, 65).
Several lines of evidence for the relevance of C-methylation during geminiviral infection have been described: (1) Viral replication in protoplasts was inhibited if the transfected DNA was methylated before inoculation (10, 18). (2) Arabidopsis plants which were defective in the genes for key enzymes of the transcriptional gene silencing pathway including cytosine methyltransferases, methyl cycle enzymes, and Dicer-like proteins developed more severe
symptoms than wild-type (wt) plants (60). (3) Transgenes which were expressed under the control of geminiviral promoters were silenced and their cytosines methylated, if the plants were superinfected with the corresponding virus and, complementary, an increased level of C-methylation was observed for the replicated DNA (4, 65). (4) Recovery of plants was accompanied with elevated C-methylation of geminiviral DNA (28, 60, 62). (5) Geminiviral suppressors of silencing changed the global methylation profile of plants (11).

All investigations with respect to sequence-specific DNA methylation have been based on the PCR technique and were, thus, not able to discriminate between the multitude of geminiviral DNA forms described (3, 17, 40, 41, 59). In particular, it is interesting to determine whether there are differences in methylation between heterogeneous linear dsDNA and the circular dsDNA conformers. To this aim, various methods on the basis of RCA have been tested, in order to determine the contribution of C-methylation to the circular population of geminiviral DNAs. The results allowed, for the first time, a broad statistical approach to validate the relevance of C-methylation for complete genome components. Two distinct viruses in the genus Begomovirus were chosen as examples: an Old World and monopartite representative, tomato yellow leaf curl Sardinia virus (TYLCSV, 50) and a New World bipartite representative, Abutilon mosaic virus (AbMV, 22). Both viruses were investigated during the course of systemic infection in the experimental model plant Nicotiana benthamiana (79).

Materials and Methods

Microorganisms and plants

Agroinfectious clones of TYLCSV (L27708, 50) and AbMV (X15983, X15984, 22) were inoculated to the experimental host Nicotiana benthamiana Domin and grown in a greenhouse with supplementary lighting or a climate chamber as described (81), under containment according to gene technology licence S2. Escherichia coli bacteria and phages used have been
described (34). In addition, leaves of naturally occurring AbMV-infected ornamental Abutilon plants (80) were used.

**Total nucleic acid extraction**

Total nucleic acids were extracted from 0.1-0.2g systemically infected upper leaves that were harvested between 9 and 10 a.m. as described (29, 40). DNA amounts were quantified by the aid of 4’,6-diamidino-2-phenylindole (DAPI, 69) fluorescence compared to a standard of lambda DNA with defined concentrations.

**Methylation-sensitive restriction enzyme analysis**

Total nucleic acids were digested by either of the isoschizomer pairs MspI/HpaII or Sau3A/MboI, using 300ng total DNA per reaction supplemented with 300ng of λ DNA to control the completeness of the digestion. 10U each of the restriction enzymes, and RNase A (0.5mg/ml) were added and incubated overnight at 37°C. Samples were separated on 1.4% agarose gels (3h, 120V). As restriction fragment length polymorphism (RFLP) standards, 1µl of 1:50 diluted RCA products (TempliPhi kit; GE Healthcare, Munich) from total nucleic acids of AbMV- or TYLCSV-infected plants was digested in parallel. Southern blots were hybridized as described (81) using digoxigenin (DIG)-labelled DNA probes (Roche, Mannheim) from full-length viral fragments.

**Methylation-dependent restriction enzymes analysis**

Total nucleic acids were digested by McrBC (NEB, Frankfurt), using 200ng total DNA per reaction supplemented with 100µg/ml BSA, 1mM GTP and 15U of the restriction enzyme at 37°C overnight. Undigested samples were treated in the same way without adding the enzyme. The samples (60ng DNA each) were separated on 1.4% agarose gels containing 20µg/ml or 50µg/ml chloroquine (19h, 40V). As a loading control and to ensure the completeness of digestion, samples were separated in parallel on 1.4% agarose gels which
then were stained with ethidium bromide. Southern blots were hybridized as described above (81) using viral probes lacking the intergenic region.

Two-dimensional (2D) agarose gel electrophoresis (1st dimension with 0.3% SDS; 2nd dimension with 20µg/ml chloroquine) was performed with 100ng total DNA per sample as described (40).

**Separation of ds and ssDNA**

Single-stranded and double-stranded DNA intermediates were separated on benzoylated naphthoylated DEAE (BND) cellulose (SigmaB6385) as described (59). Total DNA (1.5µg) was loaded on the matrices and eluted with increasing salt concentrations (0.4–1.25M NaCl) according to the manufacturer’s recommendations. Nucleic acids were collected by ethanol precipitation (63) and dissolved in 20µl sterile water.

**Sample preparation for bisulfite conversion**

The wash fraction from BND chromatography, which was enriched in dsDNA, was used for bisulfite conversion. Based on previous reports (23, 26, 78), RNA- and protein-free samples were used to promote complete conversion of unmethylated cytosines. Samples (20µl) were treated with 1mg/ml RNase A for 2h 30min, followed by 1mg/ml proteinase K overnight. Proteins were removed by two phenol/chloroform/isoamylalcohol (25:24:1) extractions, one chloroform extraction, and ethanol precipitation (63). DNA was dissolved in 20µl nuclease-free water and converted using the EpiTect bisulfite conversion kit (Qiagen, Hilden) according to the manufacturer’s protocol. The conversion rate of cytosines was validated by two control experiments. The 5S rRNA gene was partially sequenced using the primer pair (5’-TAAGAAAATCTAGAGTGTAAGGAATGTTGGATCGATTATAT-3’; 5’-TTCAATATACAGCTTTACCAAAAAAAATAACAACCGAAA-3’) as previously described for the bisulfite converted top strand (4, 24). The bacterial plasmid pBluescript II
SK(-) (Stratagene, Amsterdam) with a similar size as geminviral DNA (2961 bp) was transreplicated with the help of phage (R408) to obtain artificial standard phage progeny lacking methylation as described (34). Phages corresponding to 10, 50, 100 or 1000 ng of ssDNA were mixed with leaf samples from uninfected plants and DNA extraction was performed as for virus-infected plant tissue.

One volume of converted DNA was added to ten volumes of an RCA reaction mixture prepared by use of a TempliPhi kit (GE Healthcare, Munich), and in order to preserve the labile converted DNA, the reaction was performed at 4°C for 72h, a critical prerequisite to obtain final products.

**Random cloning and sequencing of bisulfite treated DNA**

RCA products (15µl each) from bisulfite treated DNA were digested with 60U DraI (New England Biolabs, Frankfurt am Main) in a volume of 50µl, ethanol precipitated, and ligated (63) into the EcoRV site of pBluescript SKII(+) vector (50ng) in a volume of 10µl. The product was transformed into *E. coli* DH5α and plated on LB agar containing ampicillin, X-Gal and IPTG. Positive clones were selected by blue/white screening followed by colony RCA. The presence of inserts was checked by digestion with *Bam*HI and *Cla*I. Sequencing was performed at Macrogen (Korea) using universal M13 forward primers. Sequences were analyzed using NCBI BLASTN and sorted. Viral fragments were assembled with the corresponding genomes using the CAP contig assembly program included in the BioEdit 7 software (30) and the numbers of converted nucleotides were determined with the help of Microsoft Excel.
Identification of methylated DNA in Southern blots by comparing hybridisation and immunolabelling

Two-dimensional (2D) gel electrophoresis (1st dimension with SDS; 2nd dimension with chloroquine) was performed with 1µg total DNA per sample as described (40). One of two parallel gels was processed as described previously (40), the DNA of the second gel was blotted under neutral conditions (20x SSC; 3M NaCl, 0.3M sodium citrate, pH 7.0) onto nitrocellulose membrane (Protran; Whatman, Dassel) and UV-crosslinked (2 min, 700µJ/cm², 254 nm; Ultraviolet, Amersham). The second membrane was blocked for 1h in 1% Blocking Reagent (BR; Roche, Mannheim) in PI (0.1 M maleic acid, 150 mM NaCl, pH 7.5), incubated for 1h with a 5-methylcytosine-specific antibody (33D3, mouse; Aviva, San Diego) diluted 1:5,000 in 1% BR in PI, washed three times for 15min with PI supplied with 0.03% Tween-20 (Roth, Karlsruhe) and incubated with an anti-mouse-HRP antibody conjugate (Rockland, Gilbertsville; 1:10,000 in 1% BR in PI) for 45min. After three 15min washings in PI with 0.03% Tween-20, the membrane was incubated for 15min in PBS (0.137M NaCl, 6.5mM Na2HPO4, 2mM NaH2PO4, pH 7.4) and developed using a chemoluminescence reaction (ECL femto; Pierce, Rockford) according to the manufacturer’s protocol.

Results

Nucleotide methylation at restriction enzyme recognition sites

Nucleotide methylation can be inferred from isoschizomer pairs of restriction endonucleases which are sensitive or insensitive to methylated nucleotides (reviewed by 68, 85, 86). A previous study found no cytosine methylation with isoschizomer pairs for tomato golden mosaic geminivirus, but a decrease of infectivity for in vitro methylated DNA (10). To test the methylation at specific sites for TYLCSV and AbMV, total nucleic acid extracts were harvested during the time course of systemic infections at 14, 21 and 49 days post infection.
(dpi), digested with each of the isoschizomer pairs MspI/HpaII and Sau3AI/MboI, and analyzed by blot hybridisation (Fig. 1). RCA products from infected plants served as controls for completely unmethylated DNA in parallel (Fig. 1, lanes C). To test the completeness of the enzyme reaction under the chosen conditions, equal amounts of bacteriophage λ DNA were mixed to the target DNA and analyzed in ethidium bromide-stained gels (Fig. 1a,c). MspI and HpaII share the recognition sequence CCGG. Whereas HpaII is blocked if any one of the cytosines is methylated, MspI can cut, if the internal (but not the external) C is methylated (44). The Southern blots (Fig. 1b,d) show for both geminiviruses, that the majority of hybridisation signals for the cut viral dsDNAs co-migrated with the RCA fragments indicating that the restriction of most viral dsDNA molecules was not affected by methylation. Only minor bands (Fig. 1b,d; black asterisks) deviate from this general conclusion and might indicate base modifications. However, none of these bands differ between both isoschizomer treatments for all time points of harvest. Such a result may have been caused by general methylation of both or the external cytosines, or the bands may represent mutated molecules within the viral quasispecies population without any contribution of base modifications. They also may have originated from DNA molecules that have not been replicated completely during complementary strand synthesis in vivo or from incomplete digestion. Sau3AI and MboI share the recognition sequence GATC. According to the supplier's information (New England Biolabs), both enzymes are equally impaired by cytosine methylation at C in the recognition sequence and are only differentially active for N6-adenine. The plant DNA was cut differentially by the two enzymes (Fig. 1a,c), suggesting that adenines are modified in these plants (74). The interpretation of the restriction patterns on the corresponding Southern blots (Fig. 1b,d) was impaired partially by viral ssDNA for TYLCV fragment detection, in particular during late infection. ssDNA generally appears as more diffuse bands (Fig. 1b, ss), and interferes.
especially, when defective DNAs had accumulated (Fig. 1b, lanes for 49 dpi). If compared to the RCA product controls, only two minor differential bands (Fig. 1b, white asterisks) appeared after Sau3AI digestion, but not after MboI digestion. This result is unexpected and difficult to explain, since Sau3AI should be the more effective isoschizomer in this comparison.

These initial experiments indicated that methylated viral DNA, if present, was a minor fraction of the total viral dsDNA during the analyzed time course of infection in emerging leaves. However, we had screened only a small portion of the putatively methylated sequences. To explore this question further, bisulfite sequencing was employed, a technique which selects positively for those sequences with 5-mC.

**Bisulfite treated DNA can be amplified with rolling circle amplification**

The bisulfite reaction promotes deamination of C to U, which is read as T after amplification of the target sequence whereas 5-mC is protected from this conversion. However, the target DNA becomes instable during the reaction and strand breakage removes the majority of molecules from the DNA pool under investigation (78, and data not shown). This effect leads in the end to an enrichment of protected or non-converted Cs in the resulting sequencing data. High temperature treatment promotes this bias and the longer the target sequence, the stronger the relative selection process for non-converted cytosines. The resulting well-known problems, in particular with quantification in combination with PCR, have been reviewed carefully (78).

We reasoned that processing the bisulfite-treated DNA at lower temperatures could reduce this problem. Therefore, PCR was replaced by RCA to amplify the target DNA after the conventional bisulfite reaction. In addition, this technique is not dependent on an elaborate choice of specific primers and, thus, reduces a further bias in the results. RCA selects...
positively for circular DNA molecules allowing a conformation-selective view on the replication intermediates.

In initial experiments, the denaturation step (95°C for 3 min) in the standard RCA protocol dramatically decreased the amount of RCA products obtained with bisulfite-treated templates (data not shown). Because the majority of bisulfite-treated DNA is single-stranded anyway (23), the denaturation step was omitted and the complete temperature regime was lowered to 4°C. To compensate for the lower activity of the Phi 29 polymerase at 4°C, the incubation time was extended to 72h which resulted in reasonable levels of RCA products (Fig. 2, 3). This outcome shows for the first time that a proofreading polymerase with high fidelity is able to replicate on U containing templates (78).

In order to monitor the efficiencies of conversion and subsequent amplification reactions, the final products were digested with two restriction enzymes which yielded complementary results in parallel: \(Hpa\)II (CCGG) with only Cs and Gs, or \(Dra\)I (TTTAAA) with only As and Ts in the recognition sequences (Fig. 2, 3). Upon RCA/RFLP analyses, \(Hpa\)II sites would survive the bisulfite treatment if their Cs were methylated and the resulting population would be detectable as remaining bands at the position of control DNA (Fig. 2, 3). Complementary, bisulfite treatment would generate seven new diagnostic \(Dra\)I sites (from CCCAAA, TCCAAA, TCTAAA, CCTAAA, CTTAAA, CCTAAA, TCTAAA) due to C to T exchange. Screening for such sites (genome maps in Figs. 2, 3) revealed seven in viral and eleven in complementary orientation for TYLCSV, one in viral orientation for AbMV DNA A, or two in viral and nine in complementary orientation for AbMV DNA B. In comparison to the control DNA (Fig. 2, 3, lanes C), remaining bands would indicate protection against the bisulfite reaction.

Systemically infected samples were analyzed at 9, 14, 20, 35 dpi (Fig. 2, 3; lanes 9-35). Fragment patterns changed and deviated from controls for both viruses. Prominent shifts to larger fragments than in control DNA upon \(Hpa\)II digestion (Fig. 2, 3, asterisks) indicated that...
several sites were no longer recognized by this enzyme. For \textit{DraI}, the appearance of smaller fragments documented the same effect. Similar results were obtained for the restriction enzymes \textit{SspI} (AATATT) or \textit{RsaI} (GTAC) (data not shown). Only minor variations were observed when band patterns from samples harvested at different time points after inoculation were compared. These results show that RCA/RFLP is an efficient novel tool to monitor restriction site changes after bisulfite treatment, providing the opportunity to monitor complete viral genome components without the bias of primer selection during PCR.

\textit{Random cloning of RCA products}

In addition, this procedure was ideal to identify exact sequence positions of converted cytosines. The RCA products of bisulfite-treated DNAs were cleaved with \textit{DraI} and inserted into a plasmid’s \textit{EcoRV} site by blunt end ligation, allowing a random cloning of the target DNA. The sequences determined for the resulting cloned DNAs were selected for those with similarity to the viral DNA components. For TYLCSV, 60\% of 25892 sequenced nucleotides, and for AbMV, 68\% of 17231 sequenced nucleotides were derived from viral DNAs. As expected from the few existing or potentially obtainable \textit{DraI} sites in AbMV DNA A (Fig. 3, genomic map), the AbMV DNA B-derived clones were strongly overrepresented. Due to this bias, only the complete sequences of TYLCSV and the AbMV DNA B component were covered by the sequencing and were analysed in further detail. Because no significant differences in the results were obtained for samples harvested at different time points and the presence of protected Cs appeared stochastically in all clones, the information of the pooled data sets for the two viruses was analysed. Classes of 100nts per genome segment were built and the percentages of protected Cs were averaged for each class (Fig. 4; %C). In addition, the distributions of the frequencies (Fig. 4c) were determined for these classes. For TYLCSV, 51 ± 11\% of 2249 analyzed cytosines were protected, with 53\% for CG, 50\% for CNG and 50\% for CHH sites. A single short sequence (nts 1800-2000) revealed a
considerable increase in the frequency of protected Cs (Fig. 4d; 90% of protected Cs on the viral, 83% on the complementary strand). It is located within the C1 ORF and the promoter for the C2/C3 transcript. For AbMV, 53 ± 13 % of 2042 analyzed cytosines were protected, with 50% for CG, 52% for CNG and 54% for CHH sites. These results indicate a stochastic appearance of protected Cs with no significant difference between the viruses and the potential methylation sites.

To judge these results, proper controls became essential. First, the frequently used control of sequencing PCR-amplified 5S rRNA gene segments (see Material and Methods) was applied and revealed a complete conversion of all Cs (data not shown). Subsequently, a novel control was developed and led to a totally different judgment of the results obtained. Circular single-stranded DNA from pBluescript-derived artificial bacteriophages was mixed with plant samples in amounts which were similar to those of viral DNAs in systemically infected tissues and was treated in the same way as infected samples. Cloning and sequencing the control DNAs revealed 50% protected Cs in 761 Cs analyzed. Thus, a more reliable threshold for the baseline of the experiments was determined here for the first time (Fig. 4b; dotted line). As the experimentally determined data for TYLCSV and AbMV DNA B followed mainly Gaussian distributions (Fig. 4c), only values with a significant increase above the threshold may indicate a true methylation as source for the protection of Cs. The C2/3 promoter region within ORF C1 of TYLCSV was the only genome portion of which the methylation rate significantly surpassed this threshold level in the circular geminiviral DNAs analysed (Fig. 4, c).

**Methylation-dependent restriction analysis**

In contrast to the aforementioned experimental strategies which rely on an indirect detection of methylated cytosines, the following experiments allow their positive detection. Enzymes which cut DNA only if cytosines are methylated are excellent tools to monitor epigenetic
changes (15). After using one of them (McrBC) most of the plant DNA was digested (Fig. 5, loading controls). Viral DNA, either digested with the enzyme or treated similarly without enzyme (Fig. 5; E+, E-) was resolved in one dimensional chloroquine-containing agarose gels for optimal resolution of their different conformations as shown previously (17, 56). Under these conditions, the most prominent and reproducible effect was the reduction of multimeric DNA (Fig. 5; m) and the appearance of more linear DNA of genomic size (Fig. 5, lin). Open circular (oc) and single-stranded (ss) remained stable. Supercoiled DNA (ccc) was reduced to a certain extent in some but not in all samples indicating a differential and dynamic change of its modification during the course of infection, more at late sampling times but stochastic by nature. Four independent replications of the time-course experiments for each of the two viruses revealed no general trend for the diminishing of cccDNA after McrBC treatment. It remains to be determined, thus, which experimental factor enhances the putative methylation of cccDNA. As observed previously (17), ocDNA levels remained constant during the course of infection and can be used as internal controls to estimate the equal loading of the sample pairs. Although the reason for the constant amount of the ocDNA portion is unknown, this observation was confirmed by four independent sets of experiments (Fig. 5, and unpublished results).

**Immunolabelling of 5-methylcytosine**

To further explore the conformation-dependence of geminiviral DNA, AbMV DNA was separated on two-dimensional (2D) gels which enabled the discrimination of circular and linear, ds and ssDNA with high resolution (Fig. 6a) by hybridisation and in parallel by antibodies directed against methylated bases (Fig. 6b). Out of five, the blot with the maximal signal for the detection of 5-methylcytosine (5-mC) with antibodies is shown, the others varied from a spot only at position h' (Fig. 6b) or intermediates with growing arcs of heterogeneous linear double-stranded (hds) DNA. No signal was detectable at the positions of
viral circular DNA forms in any case. The 5-mC arc (Fig. 6b) may result from sheared host DNA as well as from viral hdsDNA which cannot be discriminated with this technique. It is difficult to estimate the detection level under these conditions and it remains, therefore, possible that some circular viral molecules have escaped detection. Therefore, McrBC digestion was applied as described above and the products were separated in 2D gels. As for the 1D gels (Fig. 5), the most prominent effects of the digestion was the reduction of heterogeneous linear dsDNA (h; Fig. 6c compared with d) and the increase of the lin dsDNA level (lin; Fig. 6c vs. d).

In summary, the results of all experiments suggest that only a minor portion of geminiviral DNA is methylated and that the heterogeneous linear dsDNA is the main target of this modification. The concentration of methylated cytosines in the DNA population of the investigated geminiviruses in the chosen host and under the respective environmental conditions remained low, even after prolonged infection periods. It is, for sure, far below the reported 50% level revealed by bisulfite sequencing.

Discussion

The results indicate a lower quantitative contribution of methylated nucleotides during the chosen geminivirus infection without rejecting the hypothesis that the concomitant transcriptional gene silencing is a host defense mechanism. DNA methylation may influence the race between virus and host in a stochastic and dynamic way, as extensively reviewed in great detail (61). However, our findings stress the importance of proper controls for bisulfite sequencing-based conclusions before assigning an unconverted C as potentially methylated as critically described before (78). Unfortunately, we cannot compare our data with those in the literature because the results on the reproducibility of the control experiments have not been reported in other publications (11, 60, 62).
The ability of the Phi29 polymerase to amplify bisulfite-converted DNA is demonstrated here for the first time. Bisulfite treated DNA is a challenging template for DNA polymerases because the presence of uracil needs to be tolerated (78). In our study, the amount of RCA products obtained with bisulfite-treated DNA was optimized by lowering the incubation temperature to 4°C and thereby reducing the polymerization speed. Using random hexamer primers, it was expected to obtain the complete pool of bisulfite-converted circular DNA genomes, concomitantly excluding linear DNAs and circumventing the complicated design of non-selective primers.

Usually 84-96% of the treated DNA is degraded during conventional bisulfite reactions, resulting in a high loss of template (26). The commercial kit (EpiTect; Qiagen, Hilden) used in the current study has been reported to generate conversion levels of over 99% with only modest DNA degradation (58). However, with increasing length of the target sequence, template loss due to single-strand breaks will enhance the selection for unmodified Cs. This is particularly the case for RCA, since a single break in the circular template completely abolishes the reaction. Therefore, a proper control template to assure complete conversion of unmethylated Cs should be similar in length and conformation to the DNA molecules being analyzed and treated in the same manner. The single-stranded bacteriophage DNA used in this study fulfilled most of these requirements and provoked a more cautious evaluation of the data set. DNA of filamentous phages has been used before as essentially non-methylated template (26). Low percentages of methylation of either adenine or cytosine have been reported for single-stranded DNA containing phages (21, 32, 33, 77).

Published reports have shown probabilities of 50% cytosine methylation for the intergenic region of cabbage leaf curl begomovirus DNA A and for the beet curly top curtovirus (BCTV), as determined by bisulfite conversion followed by PCR (60), or of 36% for symmetric and 46% for asymmetric sites for tomato leaf curl begomovirus C1 ORF (4). Cytosine methylation of the viral genome was found to be stimulated in plants transgenic for
the analyzed sequences, leading to 55% protected symmetric and 47.5% asymmetric sites (4).

Extreme methylation rates of up to 80% have been reported for BCTV constructs with mutated silencing suppressors in recovered *A. thaliana* *wt* tissues (60). In contrast, analyses on pepper golden mosaic virus indicated not more than about 10% methylated cytosines in the intergenic region and 2% in its coat protein region at 10dpi, and an increase up to 25% in the intergenic region and to 10% in the coat protein region after recovery of the plants at 15dpi, with similar proportions between symmetrical and asymmetrical sites in older and younger leaves (62). These results indicate significant variability between different virus-host combinations analysed, with reference to the virus and host, and, in particular, to recovery phenomena. AbMV induces mild symptoms in *N. benthamiana* during the observed period of infection and may, thus trigger less defense response by the plant. It would be extremely interesting to determine how non-methylated control DNAs behave in the environment of the specific combinations of plant tissues and viruses in the assay. Changing viral DNA concentrations in the particular sample should influence the result as discussed for repetitive DNA in dependence of *C* ot (78).

Here, we analyzed for the first time complete geminiviral genome components and were therefore able to judge the overall variation. Only in a short stretch of the TYLCSV DNA sequence, the amount of protected Cs exceeded the threshold value. This stretch is located within the ORF C1 in a highly conserved region, which was found to bind plant transcription factors promoting the transcription of a 0.7kb complementary sense mRNA of tomato golden mosaic virus (70). Methylation of the respective DNA region may, therefore, function in transcriptional control. Further studies on this locus of different geminiviruses are, however, needed before final conclusions can be drawn. For the rest of the sites that exceeded the threshold value, the statistical basis of the data is so far too low to judge the significance of the results.
All our sequencing data are consistent with the experimental results yielded by immunological detection of 5-mC (Fig. 6) and the restriction enzyme-based analysis (Fig. 1). Monoclonal antibodies directed against 5-mC (57) have confirmed induced demethylation in *Arabidopsis* (64). Here we combined blot immunolabelling with 2D gel electrophoresis (40) for the first time, in order to differentiate between linear and circular DNA in the analysis.

Methylated plant genomic DNA was most efficiently detected showing the feasibility of this approach. However, linear viral DNA can not be differentiated from the co-migrating excess of plant nucleic acid. If the circular virus DNA had been a target of methylation at all, its amount was below the detection limit of the technique. The data resulting from RFLP analyses of the source DNA before (Fig. 1) and after RCA (Fig. 2, 3) did not support a high level of methylation either. Fragments with molecular weights deviating from the expectations do not really prove methylation, because the sensitive detection technique used here also allowed to monitor mutational polymorphisms in the quasispecies population of viral DNA molecules, and ssDNA molecules with incomplete second strand synthesis can lead to the same result (see 39 for further discussion).

The application of methylation-dependent restriction enzymes allowed a more direct determination of methylated bases and the results support the conclusions of the aforementioned experiments. This approach reveals, for the first time in geminivirology, a more quantitative estimation of the relevance of methylation and a discrimination between differently affected DNA conformations, if combined with chloroquine-aided 1D and 2D electrophoresis. On this basis, it became clear that the heterogeneous lin dsDNA was the main target of methylation, whereas ocDNA, ssDNA or cccDNA were not or considerably less modified.

In summary, our results demonstrate that the major proportion of the investigated geminiviral DNA is not methylated. Nevertheless, they do not exclude a role of methylation during viral infection. Methylation may participate in plant defense to suppress viral multiplication. It
might be induced by overlapping viral mRNAs from sense and complementary sense
transcription and subsequent RNA-directed DNA methylation. In a productive infection,
however, methylated viral DNA will be overgrown by the unmethylated viral progeny that
has escaped plant defense. Obviously, the replication modes of geminiviruses are highly
efficient by using RCR and RDR which resurrect DNA molecules from methylation.

Acknowledgements

This work was supported by DFG Je 116/10-3, a trilateral ERA-PG/BMBF project (to H.J.,
BMBF 0313986) and COST FA0806. We like to thank Drs. M. Wassenegger for asking the
right questions, R. Burdick and C. Wege for critically reading the manuscript and the
gardeners D. Gotthardt and A. Allinger for taking care of the plants.

References

the Abutilon mosaic virus (geminivirus group) in a chromatin-like structure. Mol. Gen.
Genet. 212:252-258.
2. Akbergenov, R., A. Si-Ammour, T. Blevins, I. Amin, C. Kutter, H. Vanderschuren, P. Zhang, F. Meins, Jr., T. Hohn, and M. M.
Pooggin. 2006. Molecular characterization of geminivirus-derived small RNAs in
different plant species. Nucleic Acids Res. 34:462-471.
ToLCV and its satellite DNAs. Virology 331:441-448.
silencing escape of Tomato leaf curl virus: An evaluation of the role of DNA
344:158-168.
7. Blevins, T., R. Rajeswaran, P. V. Shivaprasad, D. Beknazarians, A. Si-Ammour,
H. S. Park, F. Vazquez, D. Robertson, F. Meins, Jr., T. Hohn, and M. M.
Pooggin. 2006. Four plant Dicers mediate viral small RNA biogenesis and DNA virus
induced silencing. Nucleic Acids Res. 34:6233-6246.
using the novel primer design program: MSPprimer provides accurate methylation
analysis of the ATM promoter. Oncogene 26:6229-6237.


54. Paprotka, T., V. Metzler, and H. Jeske. 2010. The first New World begomoviruses with DNA 1-like α-satellites are able to infect Arabidopsis thaliana. Virology in revision.
645


646

Sulewska, A., W. Niklinska, M. Kozlowski, L. Minarowski, W. Naumnik, J.

649

DAPI (4',6-diamidino-2-phenylindole) binds differently to DNA and RNA: Minor-
groove binding at AT sites and intercalation at AU sites. Biochemistry 31:3103-3112.

652

Tu, J., and G. Sunter. 2007. A conserved binding site within the Tomato golden
mosaic virus AL-1629 promoter is necessary for expression of viral genes important

655


657

Engineering resistance to geminiviruses--review and perspectives. Plant Biotechnol J.
5:207-220.

660


662

Vanyushin, B. F. 2005. Enzymatic DNA methylation is an epigenetic control for
genetic functions of the cell. Biochemistry (Moscow) 70:488-499.

664


666


669

Wang, R. Y.-H., S. Shenoy, and M. Ehrlich. 1984. DNA methylation inhibits the

670

2002. Identification and resolution of artifacts in bisulfite sequencing. Methods

673

Wege, C. 2007. Movement and localization of Tomato yellow leaf curl viruses in the
Springer, Dordrecht.

675


678

Wege, C., and D. Siegmund. 2007. Synergism of a DNA and an RNA virus:
Enhanced tissue infiltration of the begomovirus Abutilon mosaic virus (AbMV)

681

Wyant, P. S., D. Gotthardt, B. Schäfer, B. Krenz, and H. Jeske. 2007. The
genomes of four novel begomovirus species and a new Sida micrantha mosaic virus

685

Zhang, W., N. H. Olson, T. S. Baker, L. Faulkner, M. Agbandje-McKenna, M.
Boulton, J. W. Davies, and R. McKenna. 2001. Structure of the maize streak virus

687

Zhu, J. K. 2009. Active DNA demethylation mediated by DNA glycosylases. Annu

689

Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an

692


694

**Figure legends**

**Figure 1:** Methylation-sensitive restriction enzyme digestion of viral DNA using isoschizomers. Total nucleic acids (300ng DNA each) from *N. benthamiana* plants systemically infected with AbMV or TYLCSV and harvested at 14, 21 or 49 dpi were digested with two sets of isoschizomers: *Msp*I (blocked by cytosine methylation of the external C) and *Hpa*II (blocked by each of both cytosine methylations) in CCGG, or *Sau*3AI (blocked by cytosine methylation, but not by adenosine methylation) and *Mbo*I (blocked by cytosine as well as adenine methylation) in GATC. Untreated samples (U) with the same amounts of DNA were loaded in parallel. The samples were electrophoresed on 1.4% agarose gels (3h; 120V), then stained with ethidium bromide (a, c), blotted to nylon membranes and detected with virus-specific full-length probes of TYLCSV (b) or AbMV A (d). RCA products (1µl; 1:50 diluted) from correspondingly infected plants were digested with the according enzymes and served as size markers for completely unmethylated DNA fragments (C). Hybridisation standards (M) were 1, 10 or 100pg of linearized full-length dsDNA fragments (lin) of AbMV or TYLCSV. In order to control the completeness of digestion, 300ng of lambda DNA were supplied to each enzyme reaction forming the band patterns (λp) in ethidium bromide-stained gels. Undigested λ DNA is expected to migrate to the same position as genomic plant DNA (λ/g). Geminiviral fragments which were not detected in the controls were marked by black or white asterisks. Expected fragment sizes in base pairs for the respective digestion products as calculated from the sequences are shown at the left and right of the panels. Note that ssDNA tends to smear and to create more diffuse bands in comparison to dsDNA as seen after Southern blot hybridisation. Bands in undigested samples of ethidium bromide-stained gels refer to the usual genomic DNA and host RNA species (not indicated).
Figure 2: RCA/RFLP analysis of bisulfite-treated DNA. Genome map of TYLCSV showing DraI sites which are present in the source DNA (underlined) or which may be created after bisulfite treatment for the viral (v) or complementary (c) strand. Total nucleic acids from systemically TYLCSV infected *N. benthamiana* plants at 9, 14, 20, 35dpi were converted by bisulfite treatment, and amplified by RCA. Three µl of these products were digested with HpaII or DraI and separated in 1.4% agarose gels (3h; 120V), then stained with ethidium bromide (a) and Southern blotted (b). As controls (C), similarly digested RCA products (1µl) of untreated samples were applied. The gels were blotted to nylon membranes and detected with a full-length virus DNA probe specific for TYLCSV. The remaining fragments of unconverted molecules are marked by asterisks. No RCA product was obtained in the sample at 9dpi.

Figure 3: Analysis of AbMV DNA as described in Fig. 2, but detected with a full-length virus DNA probe specific for AbMV DNA A.

Figure 4: Overall cytosine methylation of the TYLCSV viral strand and AbMV DNA B complementary strand. The distribution of fragments obtained after sequencing of DraI-digested randomly cloned RCA products is indicated (a). Positions of the ORFs are marked with arrows. Frequencies of protected Cs in classes of 100 nts (%C) are represented by bars and a threshold value for unmethylated phage DNA is shown as stippled line (b). The distribution of frequencies (f) for all class results is given in (c). An exceptionally protected stretch of TYLCSV is exemplified for the viral (v) and complementary (c) strand (d).

Figure 5: Methylation-dependent restriction enzyme digestion of AbMV (a) or TYLCSV (b) infected *N. benthamiana* plants (plant numbers, P#) were harvested at the indicated days post infection (dpi). In addition, individual leaves (starting with the youngest leaf as #1) of ornamental *Abutilon* plants which were naturally infected with AbMV were collected (c). Total nucleic acids (60ng DNA each) were digested with McrBC (E+), or not (E-), separated electrophoretically on 1.4% agarose gels containing chloroquine [50 µg/ml (a); 20µg/ml (b,
c) and analysed by Southern blot hybridisation against the respective probes (AbMV DNA A or TYLSCV lacking each the intergenic region). Viral DNA forms are indicated as in Fig. 1 and multimeric (m), open circular (oc), relaxed covalently closed circular (rccc). As a loading control and to ensure complete digestion, equal amounts of every sample were separated in parallel on 1.4% agarose gels, and stained with ethidium bromide to show the genomic plant DNA. Hybridisation standards (M) were 100pg (a) 1, 10, 100pg (b,c) of linearized full-length dsDNA fragments (lin) of AbMV DNA A or TYLCSV. For each time point, two plants out of five were selected randomly and used for digestion (a, b).

Figure 6: Positive detection of methylated viral DNA after 2D agarose gel electrophoresis. Following Southern blotting, either AbMV DNA was probed with full-length DNA A (a), or methylated DNA was visualized with 5-mC antibodies (b). In an independent experiment, viral DNA (100ng total nucleic acids from AbMV infected *N. benthamiana* plants harvested at 21 dpi) was kept untreated (c) or was treated with McrBC (d) and analyzed by 2D agarose gel electrophoresis (first dimension 0.3% SDS, second dimension 20µg/ml chloroquine; 19h 45V). Southern blots were detected (c, d) with a probe of AbMV DNA A lacking the intergenic region. The most prominent viral DNA forms are described in Fig. 5 and in addition heterogeneous linear dsDNA (h, h'), recombination-dependent replication intermediates (RDR), complementary strand replication intermediates (CSR), dimeric forms (2x) are indicated.
Paprotka 2011-Bis Figure 1
Figure 1: Methylation-sensitive restriction enzyme digestion of viral DNA using isoschizomers. Total nucleic acids (300ng DNA each) from *N. benthamiana* plants systemically infected with AbMV or TYLCSV and harvested at 14, 21 or 49 dpi were digested with two sets of isoschizomers: *Msp*I (blocked by cytosine methylation of the external C) and *Hpa*II (blocked by each of both cytosine methylations) in CCGG, or *Sau*3AI (blocked by cytosine methylation, but not by adenosine methylation) and *Mbo*I (blocked by cytosine as well as adenine methylation) in GATC. Untreated samples (U) with the same amounts of DNA were loaded in parallel. The samples were electrophoresed on 1.4% agarose gels (3h; 120V), then stained with ethidium bromide (a, c), blotted to nylon membranes and detected with virus-specific full-length probes of TYLCSV (b) or AbMV A (d). RCA products (1µl; 1:50 diluted) from correspondingly infected plants were digested with the according enzymes and served as size markers for completely unmethylated DNA fragments (C). Hybridisation standards (M) were 1, 10 or 100pg of linearized full-length dsDNA fragments (lin) of AbMV or TYLCSV. In order to control the completeness of digestion, 300ng of lambda DNA were supplied to each enzyme reaction forming the band patterns (λp) in ethidium bromide-stained gels. Undigested λ DNA is expected to migrate to the same position as genomic plant DNA (λ/g). Geminiviral fragments which were not detected in the controls were marked by black or white asterisks. Expected fragment sizes in base pairs for the respective digestion products as calculated from the sequences are shown at the left and right of the panels. Note that ssDNA tends to smear and to create more diffuse bands in comparison to dsDNA as seen after Southern blot hybridisation. Bands in undigested samples of ethidium bromide-stained gels refer to the usual genomic DNA and host RNA species (not indicated).
**Figure 2:** RCA/RFLP analysis of bisulfite-treated DNA. Genome map of TYLCSV showing _DraI_ sites which are present in the source DNA (underlined) or which may be created after bisulfite treatment for the viral (v) or complementary (c) strand. Total nucleic acids from systemically TYLCSV infected _N. benthamiana_ plants at 9, 14, 20, 35dpi were converted by bisulfite treatment, and amplified by RCA. Three µl of these products were digested with _HpaII_ or _DraI_ and separated in 1.4% agarose gels (3h; 120V), then stained with ethidium bromide (a) and Southern blotted (b). As controls (C), similarly digested RCA products (1µl) of untreated samples were applied. The gels were blotted to nylon membranes and detected with a full-length virus DNA probe specific for TYLCSV. The remaining fragments of unconverted molecules are marked by asterisks. No RCA product was obtained in the sample at 9dpi.
Figure 3: Analysis of AbMV DNA as described in Fig. 2, but detected with a full-length virus DNA probe specific for AbMV DNA A.
Paprotka 2011-Bis Figure 4
Figure 4: Overall cytosine methylation of the TYLCSV viral strand and AbMV DNA B complementary strand. The distribution of fragments obtained after sequencing of DraI-digested randomly cloned RCA products is indicated (a). Positions of the ORFs are marked with arrows. Frequencies of protected Cs in classes of 100 nts (%CC) are represented by bars and a threshold value for unmethylated phage DNA is shown as stippled line (b). The distribution of frequencies (f) for all class results is given in (c). An exceptionally protected stretch of TYLCSV is exemplified for the viral (v) and complementary (c) strand (d).
Paprotka 2011-Bis Figure 5
**Figure 5:** Methylation-dependent restriction enzyme digestion of AbMV (a) or TYLCSV (b) infected *N. benthamiana* plants (plant numbers, P#) were harvested at the indicated days post infection (dpi). In addition, individual leaves (starting with the youngest leaf as #1) of ornamental *Abutilon* plants which were naturally infected with AbMV were collected (c). Total nucleic acids (60ng DNA each) were digested with McrBC (E+), or not (E-), separated electrophoretically on 1.4% agarose gels containing chloroquine [50µg/ml (a), 20µg/ml (b, c)] and analysed by Southern blot hybridisation against the respective probes (AbMV DNA A or TYLCSV lacking each the intergenic region). Viral DNA forms are indicated as in Fig. 1 and multimeric (m), open circular (oc), relaxed covalently closed circular (rccc). As a loading control and to ensure complete digestion, equal amounts of every sample were separated in parallel on 1.4% agarose gels, and stained with ethidium bromide to show the genomic plant DNA. Hybridisation standards (M) were 100pg (a) or 1, 10 and 100pg (b, c) of linearized full-length dsDNA fragments (lin) of AbMV DNA A or TYLCSV. For each time point, two plants out of five were selected randomly and used for digestion (a, b).
Paprotka 2011-Bis Figure 6
Figure 6: Positive detection of methylated viral DNA after 2D agarose gel electrophoresis. Following Southern blotting, either AbMV DNA was probed with full-length DNA A (a), or methylated DNA was visualized with 5-mC antibodies (b). In an independent experiment, viral DNA (100ng total nucleic acids from AbMV infected *N. benthamiana* plants harvested at 21 dpi) was kept untreated (c) or was treated with McrBC (d) and analyzed by 2D agarose gel electrophoresis (first dimension 0.3% SDS, second dimension 20µg/ml chloroquine; 19h 45V). Southern blots were detected (c, d) with a probe of AbMV DNA A lacking the intergenic region. The most prominent viral DNA forms are described in Fig. 5 and in addition heterogeneous linear dsDNA (h, h'), recombination-dependent replication intermediates (RDR), complementary strand replication intermediates (CSR), dimeric forms (2x) are indicated.