Nucleic Acids Research, 1997, Vol. 25, No. 12 **2409–2416**

Novel pattern of DNA methylation in *Neurospora crassa* transgenic for the foreign gene *hph*

Antonio C. Codón, Yeon-Su Lee and Vincenzo E. A. Russo*

MPI Molekulare Genetik, Ihnestrasse 73, D-14195 Berlin, Germany

Received February 25, 1997; Revised and Accepted April 21, 1997

ABSTRACT

It has previously been reported that multiple copies of the *hph* gene integrated into the genome of *Neurospora crassa* are methylated at *Hpa*II sites (CCGG) during the vegetative life cycle of the fungus, while *hph* genes integrated as single copies are not methylated. Furthermore, methylation is correlated with silencing of the gene. We report here the methylation state of cytosine residues of the major part of the promoter region of the *hph* gene integrated into the genome of the multiple copy strain HTA5.7 during the vegetative stage of the life cycle. Cytosine methylation is sequence dependent, but the sequence specificity is complex and is different from the sequence specificity known for mammals and plants (CpG and CpNpG). The pattern of DNA methylation reported here is very different from that measured after meiosis in *Neurospora* or in *Ascobulus*. After the sexual cycle in those two fungi all the cytosines of multiple stretches of DNA are heavily methylated. This indicates that the still unknown methyltransferase in *Neurospora* has a different specificity in the sexual and the vegetative stages of the life cycle or that there are different methyltransferases. The pattern of methylation reported here is also different from the pattern of cytosine methylation of transgenes of *Petunia*, the only pattern published until now in plants that has DNA methylation at cytosines which are not in the canonical sequences CpG and CpNpG.

INTRODUCTION

In prokaryotes the major methylated bases are 4-methylcytosine (m^4C), 5-methylcytosine (m^5C), 5-hydroxymethylcytosine (hm^5C) and 6-methyladenine (m^6A) (1). Hundreds of methyltransferases have been identified which methylate either cytosine or adenine at specific recognition sites. The recognition site for methylation can be as simple as the double base CpG or a much more complex sequence (up to 15 bases). On average it is 4–6 bases long (1). Over 320 restriction endonucleases sensitive to site-specific methylation are known today (1). Many biological functions have been attributed to m^6A generated by the Dam methyltransferase in *Escherichia coli*: induction of post-replicative mismatch repair, control of *E.coli* chromosome replication and segregation, control of plasmid segregation, regulation of transposition, gene expression and control of initation of phage P1 DNA packaging (2).

In contrast, in eukaryotes the base modification mostly studied is m^5C. The modified cytosine is only in the symmetrical sequence CpG or CpNpG in both mammals and plants (3). The frequency of m^5C is very high, from 4% in humans (4) to 36% in some higher plants (5). Recently it was found that in transgenic fungi after the meiotic phase cytosine methylation occurs at very high frequency (>50%) and at any cytosine with equal probability in *Neurospora* (6) and in *Ascobolus* (7). In *Neurospora*, DNA methylation was studied in correlation with RIP (repeat-induced point) mutations, which acts on sequence duplications and occurs only in special pre-meiotic cells containing haploid nuclei from both parents (6). In the ascomycete *Ascobolus immersus*, duplicated DNA segments are subject to the methylation-induced pre-meiotic (MIP) process (7).

The only pattern of m^5C in plants that is different from the simple one quoted above was found in transgenic *Petunia* (8). An indication that some non-symmetrical sequences could be methylated in plants was published in the same year by two different groups (9,10). Tasheva and Roufa (11,12) reported densely methylated DNA islands in mammalian chromosomal replication origins regardless of their dinucleotide composition. However, those results could not be reproduced by Rein *et al.* (13,14).

In eukaryotes, m^5C has a role in gene silencing, genetic imprinting, X-chromosome inactivation in mammals and paramutation in plants (15). In *Neurospora*, the only foreign gene that is known to be silenced is *hph*, a gene that confers resistance to the antibiotic hygromycin (16). Transgenic *Neurospora* with a single copy of the *hph* gene had conidia that were 100% resistant to the antibiotic, while in many transgenics with multiple copies the *hph* gene was silenced, so that only a small percentage of the conidia were resistant to hygromycin. The *hph* gene in the transgenic with a single copy was not methylated, while the genes in the multiple copies transgenic were heavily methylated (16). The degree of methylation of the *hph* gene was measured as resistance to restriction by the enzymes *Hpa*II and *Msp*I (restriction site CCGG). Further evidence that methylation could be the cause of silencing of the *hph* gene came from experiments with the demethylating agent 5-azacytidine, which converted silenced sensitive strains into resistant strains (16,17). All the studies on silencing and methylation of the *hph* gene were done

* To whom correspondence should be addressed. Tel: +49 30 8413 1264; Fax: +49 30 8413 1385; Email: russo@mpimg-berlin-dahlem.mpg.de

on mycelia in the vegetative stage of the life cycle, which never went through meiosis after transformation.

Here we present evidence that the *hph* genes of a transgenic strain of *Neurospora* with multiple copies have a pattern of cytosine methylation that is different from that of animals (3), higher plants (5) and transgenic *Petunia* (8).

MATERIALS AND METHODS

Neurospora crassa strains and culture conditions

The strains used (wild-type, HTA5.7 and HTA23) have been described previously (16). The culture conditions were standard ones (16). Strain HTA23 has only one copy of the *hph* gene inserted into the genome, while strain HTA5.7 has at least four copies (18).

Plasmid pCSN44

Plasmid pCSN44 (19), containing the *hph* gene, was a generous gift of C.Yanofsky (Stanford University). For the Southern blot analysis the plasmid was grown in *E.coli* strain JM110 (Biolabs), which is *dam*$^-$ and *dcm*$^-$.

Southern blots

The Southern blots were hybridized using as probe a stretch of DNA that was labelled with digoxigenin according to the manufacturer (Boehringer Mannheim) (16).

Bisulphite reaction

The bisulphite reaction was performed according to Olek *et al.* (20), with the following modifications. The chromosomal DNA of *Neurospora* was digested with the restriction enzyme *Eco*RV, boiled for 10 min and quickly chilled on ice. A total of 5 µg DNA were mixed with 2 vol 2% LMP agarose (FMC Bio-products, Rockland, ME) dissolved in water to give a final volume of 70 µl. This mixture was directly pipetted into chilled mineral oil to form seven beads of 10 µl. In the reaction tubes the seven beads were covered with 1 ml bisulphite solution and the protocol of Olek *et al.* (20) was strictly followed.

Choice of oligonucleotide primers

We noticed that the oligonucleotides for amplification of the DNA after bisulphite reaction must be chosen after a very stringent and time consuming test in order to avoid amplification of non-representative DNA strands. When we amplified a stretch of the promoter region of the *hph* gene of the multiple copies insert HTA5.7 strain after bisulphite treatment we found that all the C residues must have been methylated. This result was found also in the case where the bisulphite-treated DNA came from strain HTA23, which contains only a single integrated copy. This was at variance with the results of the Southern blots, which indicated no methylation of the *hph* gene in this strain (data not shown). It seems that our oligonucleotide primers were amplifying rare DNA strands that did not represent the bulk of the bisulphite-treated DNA. We devised the following test to check if a pair of primers amplified DNA strands in a biased way (bias test). The pCSN44 plasmid, containing the *hph* gene, was treated with bisulphite. A pair of degenerated primers were used to PCR amplify a stretch of DNA containing the promoter of the *hph* gene, using oligonucleotides O1 and O2 and as template both the plasmid before and after bisulphite treatment (Fig. 1). Afterwards, we cloned the PCR products into the TA vector (Invitrogen) and sequenced them. We obtained in this way two different types of TA plasmids. One had the subregion of the *hph* gene with the original sequence (pTAC), the other had the same sequence but with all the C residues changed to T (pTAT). These two plasmids were used as template, mixed in equal amount, in a PCR reaction with the pair of primers to be tested, the rationale being that if there was no selection after PCR, cloning and sequencing, the number of clones with the pTAC sequence obtained in the second round of PCR should be equal to the number of clones with the pTAT sequence. Using in this bias test the pair of primers that gave conflicting results between the degree of methylation measured with the bisulphite reaction and that measured with the Southern blot, we obtained 30 clones with the pTAC sequence and zero with the pTAT sequence. This is a clear indication that this pair of primers made a strong selection among the two different plasmids. We then tested four more pairs of primers chosen at random and found in total two pairs of oligonucleotides that, in the bias test explained above, did not give any bias, while the other three did. One of these pairs of primers was used here and their sequences are shown in Figure 2.

PCR experiments

The PCR cocktail was assembled according to the recomendations of the manufacturer of the kit (Amplitaq DNA polymerase, Stoffel fragment; Perkin Elmer Cetus). Denaturation was at 96°C for 30 s and annealing at 50°C for 30 s. The extension step was at 72°C for 1 min. At the end of 40 cycles a delay of 4 min at 72°C was added.

Cloning and sequencing

The PCR products were ligated and cloned following the instructions for the TA cloning kit (Invitrogen). Sequence analysis of the cloned DNA fragments was performed by cycle sequencing using 25 ng DNA, 2 µl d/ddNTP reaction mixture (Promega Silver sequencing kit), 1.25 µl special buffer (50:1 mixture of 5× Promega Silver sequencing buffer and Triton X-100), 2.5 pmol fluorescently marked primer, 2 µl sterile water, 0.25 µl Taq polymerase (Promega Silver sequencing kit). The DNA was denatured at 93°C for 2 min. Amplification was for 36 cycles of 20 s denaturation at 94°C, 20 s annealing at 55°C and 40 s elongation at 72°C. The final elongation step proceeded for 10 min and the reaction mixture was then cooled and stored at 4°C after 3 µl sequencing stop buffer (Pharmacia) had been added. Sequences were read on an ALF automatic sequencer (Pharmacia). The following fluorescently marked primers were used for the sequencing: universal primer and reverse primer (Pharmacia).

RESULTS

The two methods most widely used today to determine the state of methylation of cytosine residues in eukaryotic genomes are: (i) Southern blots analysis of DNA restricted with methylation-sensitive enzymes (1); (ii) the bisulphite method (20).

The first method takes advantage of the fact that many enzymes do not digest DNA if one cytosine of the recognition/restriction site is methylated. The method is simple and reliable, but it has the disadvantage that only few cytosines can be probed, namely

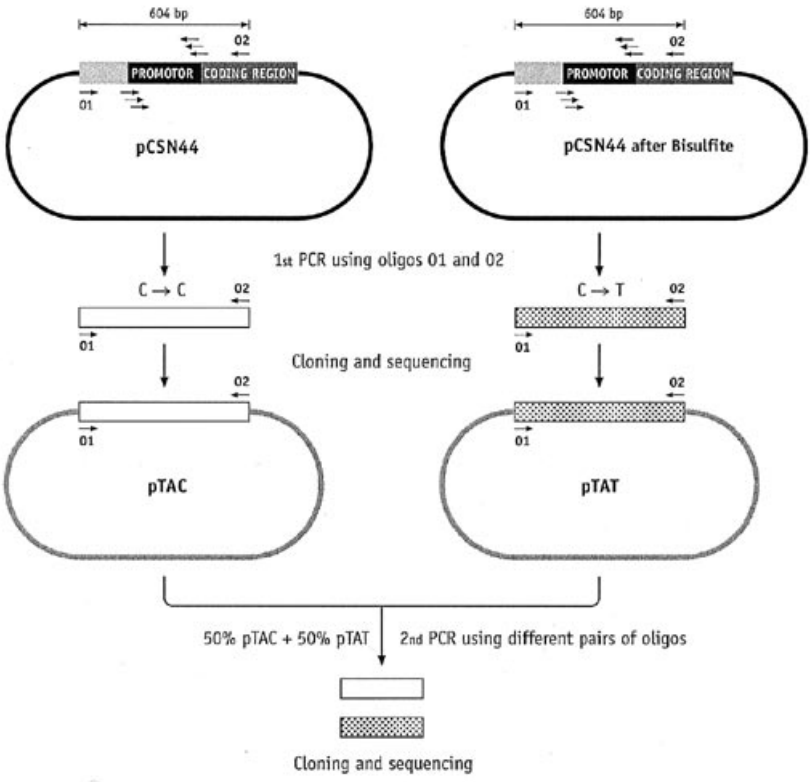

Figure 1. Experimental scheme of the bias test. In the top plasmids are shown the promoter and the coding region of the *hph* gene with the position of the O1 and O2 oligonucleotides used to amplify a stretch of DNA containing the promoter region, as well as the position of the primers tested in the bias test. In the first PCR a stretch of DNA containing the promoter region was amplified with the O1 and O2 oligonucleotides from a pCSN44 plasmid and cloned into a TA vector. The plasmid obtained (pTAC) contains the original sequence (white bar). With the same protocol the same region was amplified and cloned from previously bisulphite-treated pCSN44. The plasmid obtained (pTAT) contains the same insert as that in pTAC with the difference that all the C residues have been changed to T (hatched bar). All plasmids were checked by sequencing. The two plasmids were used for the actual bias test. They were used as template, in equal amount, in a second PCR using the pair of oligonucleotides under test. The amplified DNA was cloned into the TA vector and several clones were sequenced. We assumed that a pair of oligonucleotides display no preference, neither in PCR nor during cloning of the resulting PCR product, only when the number of clones with the white sequence was equal to the number of clones with the hatched sequence. In the case of the oligonucleotides used in the experiments reported in Figure 2 this number was seven to eight.

only those which are in the recognition/restriction site of the methylation-sensitive enzymes.

The bisulphite method takes advantage of the fact that bisulphite, under appropriate experimental conditions, will change *in vitro* any C of a single strand into U. m^5C is resistant to the action of bisulphite under the same experimental conditions. Bisulphite does not work on cytosine residues of double-stranded DNA. The bisulphite-treated DNA is subsequently amplified by PCR, cloned and sequenced. During PCR amplification the U is substituted by T, with the consequence that in the DNA sequence of the amplified stretch of DNA there will be a T for any C of the original sequence and a C for any m^5C. The advantage of this technique is that it can determine the methylation state of each single cytosine and the methylation pattern on each single filament of DNA. If in a tissue there are several identical genes with different patterns of methylation, that can be picked up. The drawback is that it is difficult to use because there are many potential artifacts that can give false results. Several papers have been published after the original one in order to improve this powerful technique (22–24). We used here another protocol for bisulphite treatment, developed in our institute, that minimizes reannealing of the DNA during the bisulphite treatment and permits amplification of long stretches of DNA (20). We have also found that a very important point is the choice of primers for the PCR, as discussed at length in Materials and Methods. For unknown reasons many pairs of primers can select rare DNA strands with a sequence that is not representative of the sequence of the bulk of the DNA strands present after bisulphite treatment.

DNA methylation of the transgene *hph* according to the bisulphite method

In *Neurospora*, multiple copies of the same stretch of DNA after meiosis contain RIP mutations and are heavily methylated, with methylation at any cytosine having equal probability (6). It was suggested that RIP mutations created a signal for cytosine methylation (6). We were interested to know if the methylation pattern is the same in the case of transformants which never went through the meiotic cycle. Therefore, we have analysed the pattern of methylation of the promoter region of the *hph* gene in the strain HTA5.7, which has at least four inserts of the *hph* gene, and in the strain HTA23, which has only one insert (18). It was reported previously that the *hph* transgene was methylated at the *Hpa*II sites in the HTA5.7 strain, while it was not methylated in the HTA23 strain (16,18). The stretch of DNA between positions –359 and –137 of the promoter region of the gene in the HTA5.7 strain was analysed. The pattern of methylation in the HTA5.7

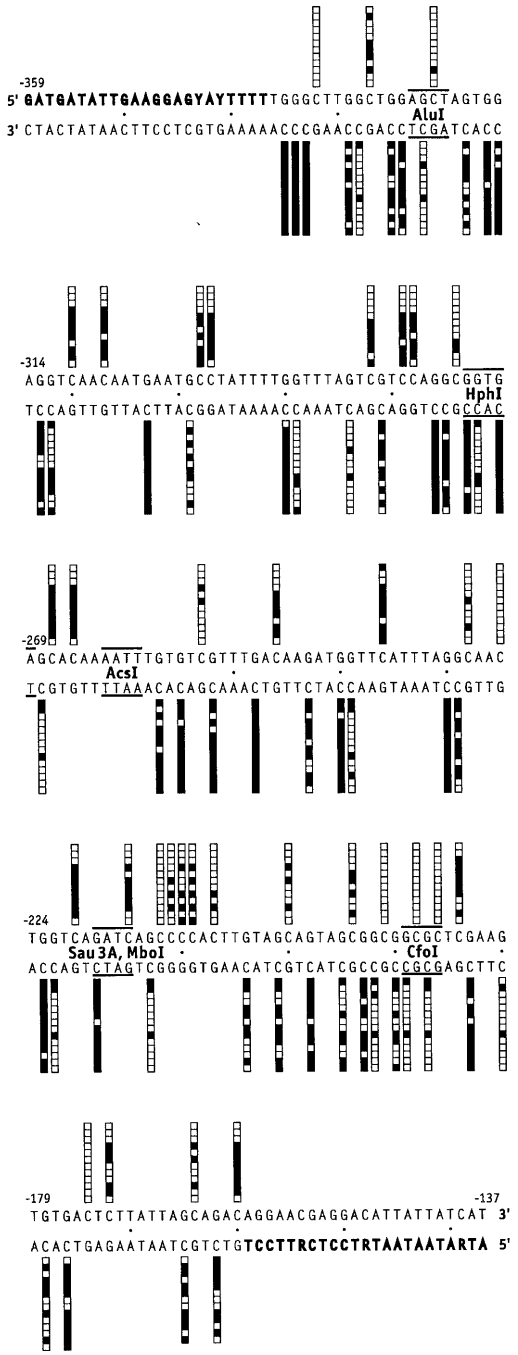

Figure 2. Methylated C residues determined by the bisulphite method. The –359 to –137 region upstream of the start codon (+1 is the first position of the ATG codon), containing most of the promoter region of the *hph* gene, was analyzed. Data from 12 DNA filaments cloned from the top strand and 14 clones from the bottom strand, from three different bisulphite reactions, have been compiled. Methylated C residues are labelled with filled symbols. The sequence of two oligonucleotides used for cloning of the PCR product after bisulphite treatment are highlighted. The restriction sites of enzymes used in the Southern blot analysis presented in Figure 3 are underlined. Y and R in the sequences of the two oligonucleotides indicate that in this position the bases C and T (Y) or A and G (R) have been incorporated with equal probability. The first seven strands of the top strand and the first nine of the bottom strand (counting from the top) were obtained in the first bisulphite treatment, the next three strands were obtained from the second bisulphite treatment and the last two strands were obtained from the third bisulphite treatment.

strain is quite complex, as can be seen in Figure 2. Twelve amplified DNA strands of the top strand and 14 of the bottom strand have been analysed, from a total of three independent bisulphite reactions.

It can be noticed that each DNA strand analysed has a different methylation pattern, but there is still a general trend. The general level of methylation is very high, both in the top strand (35% of all cytosines are methylated) as well as in the bottom strand (61%). This level is much higher than the average level of natural m^5C in *Neurospora*, which is 1.5% of all cytosines (25).

The same region of the *hph* gene shown in Figure 2 was analysed from the single insert strain HTA 23. Eight clones (seven from the bottom and one from the top strand) were sequenced from one bisulphite reaction. None of the cytosines analysed appeared to be methylated (data not shown).

An internal control was performed, to test the validity of our protocol of the bisulphite method. We also amplified a stretch of the promoter region of the *cpc* gene from the same bisulphite reactions used for the analysis in Figure 2. This gene is not expected to be methylated and was amplified using the same primers as in Selker *et al.* (6). For each bisulphite reaction, a stretch of the promoter region of the *cpc* gene was PCR amplified, cloned and five clones were sequenced. The sequence had all the C residues changed to T, evidence that the bisulphite reacted with 100% efficiency (data not shown). The PCR sequencing was very faithful, with only three mistakes in >5000 bases sequenced.

Confirmation of the methylation state of some cytosines by Southern blot analysis

Knowing the pattern of methylation of the promoter region of the *hph* gene in the HTA5.7 strain from the data of Figure 2, we were able to obtain a second and independent evaluation of the methylation state of some of the cytosines by Southern blot analysis. We digested the DNA with several restriction enzymes: two enzymes whose sites were not methylated (*Alu*I and *Cfo*I), two enzymes whose sites were methylated (*Hph*I and *Sau*3A), one enzyme that is insensitive to methylation (*Mbo*I) and one enzyme that cuts at a site that does not contain cytosines (*Acs*I). The last enzyme was used as a control for the quality of the DNA. The position of the restriction sites of these enzymes in the sequence of Figure 2 are: *Cfo*I, –186; *Sau*3A and *Mbo*I, –215; *Hph*I, –269; *Alu*I, –320; *Acs*I, –258.

Table 1 shows the restriction sites for these enzymes, the methylcytosines that inhibit their activity and the methylation state of those sites in the stretch of DNA analysed in Figure 2. The DNAs used in the Southern blots were those of HTA 5.7, the DNA used for the bisulphite analysis of Figure 2, HTA23, the single copy transformant which did not show any methylation according to the bisulphite method, and plasmid pCSN44. The plasmid DNA was used as a positive control of the activity of the enzymes.

The results shown in Figure 3 indicate that the single copy insert in strain HTA23 had no methylation. The enzymes *Hph*I and *Sau*3A did not cut the DNA of the multicopy insert strain HTA 5.7, while the other enzymes did cut it, as expected from the data of Table 1. This is confirmation that for at least five different cytosines the methylation level was that measured by the bisulphite method: at positions –218 (*Sau*3A) and –270 (*Hph*I) the cytosines are methylated, while at positions –187 and –189 (*Cfo*I) and –222 (*Alu*I) the cytosines are not methylated.

Table 1. Methylation state of the cystosines present in the restriction sites of the enzymes used in Figure 3

Restriction enzyme	Sites cut	Sites not cut	Methylation state of the restriction site in the promotor region of the *hph* gene according to sequence data[a]		
			HTA5.7	HTA23	pCSN44
*Acs*I	AATT	–	–	–	–
*Alu*I	AGCT	AGm^{5}CT	AGCT	AGCT	AGCT
*Cfo*I	GCGC	Gm^{5}CGC GCGm^{5}C	GCGC	GCGC	GCGC
*Hph*I	TCACC TCACm^{5}C	Tm^{5}CACC TCAm^{5}CC	Tm^{5}CACm^{5}C	TCACC	TCACC
*Mbo*I	GATC GATm^{5}C	–	GATm^{5}C	GATC	GATC
*Sau*3A	GATC	GATm^{5}C	GATm^{5}C	GATC	GATC

[a]HTA5.7 according to the sequence of Figure 2, HTA23 and pCSN44 from sequences not shown.

Table 2. Frequency of methylation in the central C of all possible triplets of sequence 5′-NCN

Triplet	*Neurospora crassa*		*Petunia hybrida*	
	Top strand	Bottom strand	Top strand	Bottom strand
ACA	58, 67, 50, 67	43, 50, 93, 78	60	24
TCA	58, 75, 67, 50	78, 100, 100, 100	16	0, 0, 0, 92
CCA	50, 42	93, 93, 93, 93, 64, 71, 100	50, 0, 50, 100, 50, 0, 100, 0	76, 16, 76, 68
GCA	67, 17, 25, 25	43	50, 16, 0, 24	76, 60
ACT	0, 17, 0	14, 86, 14, 57	40	16
TCT	42	62, 93, 64	100, 0	100, 84, 8, 92
CCT	58	100, 100, 78	0, 76	0, 24, 76, 76
GCT	0, 41, 8, 0	62, 43, 64, 21, 21, 14	–	0, 16
ACC	–	21, 29, 21, 21, 43	0, 0, 0	76
TCC	58	86, 78	50, 16, 0, 0, 0	0, 0, 8, 60, 8, 68
CCC	25, 33	100	0, 0, 0, 16, 0, 8	0
GCC	50, 0	14, 21, 64, 78, 14, 100	16, 40	68
ACG	–	86, 64	100, 100, 100, 100, 100	100, 100, 100, 100, 92
TCG	50, 17, 58	86	100, 100, 100	92, 92, 86
CCG	–	43, 71, 93	100	–
GCG	17, 33, 0, 0	21	92	100, 92

Each entry represents the frequency of methylation of one triplet. The order of entries is 5′→3′ for each strand. The frequency of methylation is taken from the data presented in Figure 2 for *Neurospora crassa* and from Meyer *et al*. (8) for *Petunia hybrida*. For example, the first ACA triplet of the top strand in *Neurospora* is at position –307 and has a frequency of methylation of 58%, the next ACA triplet is at position –265 and has a frequency of methylation of 67%, etc.
The average percentage of m^{5}C is 35% in the top strand and 61% in the bottom strand of *Neurospora*; it is 25% in the top strand and 41% in the bottom strand of *Petunia*.

Are cytosine residues in the promoter region of the *hph* gene methylated at random?

In order to answer this question the overall level of m^{5}C in the top strand and in the bottom strand of the stretch of DNA shown in Figure 2 was analysed. In the top strand there were, on average, 4.2 m^{5}C at each cytosine position of the 12 DNA strands analysed (this number was obtained by counting the total number of m^{5}C in the 12 top strands divided by the number of cytosines between the two primers in the top strand, 35). If methylation were random, independent of the DNA sequence, then we would expect the distribution of m^{5}C to follow a Poisson distribution. A comparison between the distribution of the experimental data and the Poisson distribution is shown in Figure 4. It is clear that there is a discrepancy between the two curves at both low and high percentages of m^{5}C. Similar results were obtained on analysis of the distribution of m^{5}C in the bottom strand, which contains on average 8.5 m^{5}C at each cytosine position of the 14 DNA filaments analysed (Fig. 4). In both strands there are more cytosines with very high or very low levels of methylation than expected from a random process. Therefore, methylation must have some sequence preference, since cytosines are not methylated at random.

Pattern of DNA methylation of the promoter region of the *hph* gene in the HTA5.7 transformant

We did not expect the recognition site of DNA methylation activity to be very complex, because ~50% of the cytosines are methylated. We first asked whether in *Neurospora* the configuration CpG and/or CpNpG is a recognition site for DNA methyltransferase activity, as it is for mammals and plants. The

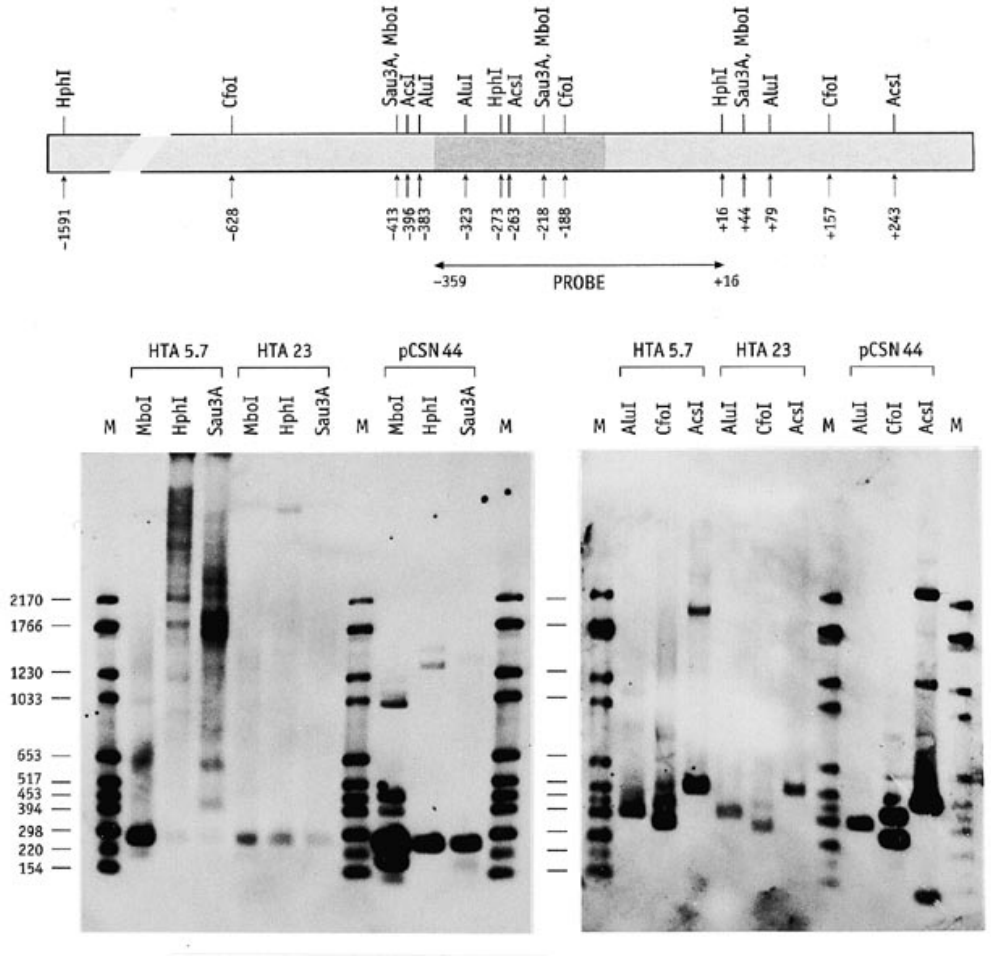

Figure 3. Methylation state of the promoter of the *hph* gene measured by Southern blot analysis. The DNAs have been cut with the following restriction enzymes: (i) *Mbo*I, which cuts at GATC independently of methylation of the C residue, giving DNA fragments of expected sizes 195 and 262 bp; (ii) *Sau*3A, which cuts the same sequence, GATC, if the C is not methylated; (iii) *Hph*I, which cuts TCACC if neither of the two internal C residues are methylated, giving DNA fragments of expected sizes 289 and 1318 bp; (iv) *Alu*I, which cuts AGCT if the C is not methylated, giving DNA fragments of expected sizes 60 and 402 bp; *Cfo*I cuts GCGC if none of the C residues are methylated, giving DNA fragments of expected sizes 345 and 440 bp; *Acs*I, which cuts AATT, giving DNA fragments of expected sizes 133 and 506 bp. In each lane labelled HTA5.7, 30 µg HTA5.7 DNA was loaded; in the lanes labelled HTA 23, 30 µg HTA23 was loaded; in the lanes labelled pCSN44, 30 µg wild-type DNA + 3 ng DNA from plasmid pCSN44 were loaded.

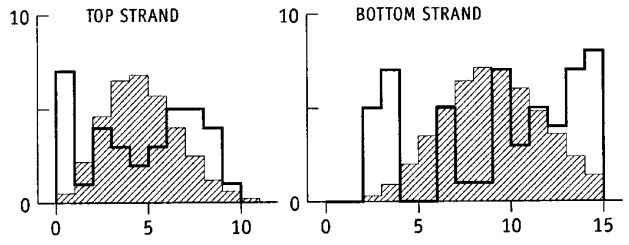

Figure 4. Distribution of the methylated C residues of the top and bottom strand compared with the Poisson distribution. The abscissa indicates the methylation degree of a given C residue, namely the number of DNA filaments that have a methylated C residue at a given position of the DNA sequence analysed. The methylation degree can be between 0 and 12 for the top strand and between 0 and 14 for the bottom strand. The ordinate indicates the number of C residues having a given methylation degree. The average value of the methylation degree is 4.2 for the top strand and 8.5 for the bottom strand. With these two values we calculated the Poisson distribution for each strand (hatched histogram). The histogram with a continuous line shows the experimental data obtained from the data presented in Figure 2.

distributions shown in Figure 5 indicate that they are not preferably methylated. It was of interest to compare our data with the data on *Petunia hybrida*, the only data published until now on the methylation pattern of a transgenic gene in plants (8). We analysed the data of Meyer *et al.* (8) in the same way as our own data and this analysis is shown in Figure 5. It is evident that the configurations CpG and/or CpNpG are strong signals for almost 100% methylation of cytosine, as known in plants for endogenous genes (5).

We determined thereafter whether the immediate neighbourhood of a cytosine influences the level of methylation. The data on the level of cytosine metylation of a stretch of the promoter region of the *hph* gene (Fig. 2) were classified into 16 different groups, depending on the bases that are 5′ and 3′ of the cytosine. The data are shown in Table 2. All the cytosine residues in the following triplet context were highly methylated (42–100%): A<u>C</u>A (63 ± 16), T<u>C</u>A (78 ± 20), C<u>C</u>A (77 ± 21), T<u>C</u>T (65 ± 21), C<u>C</u>T (84 ± 20), T<u>C</u>C (74 ± 14), A<u>C</u>G (75 ± 15), C<u>C</u>G (69 ± 25); the average percentage cytosine methylation ± SD is reported in

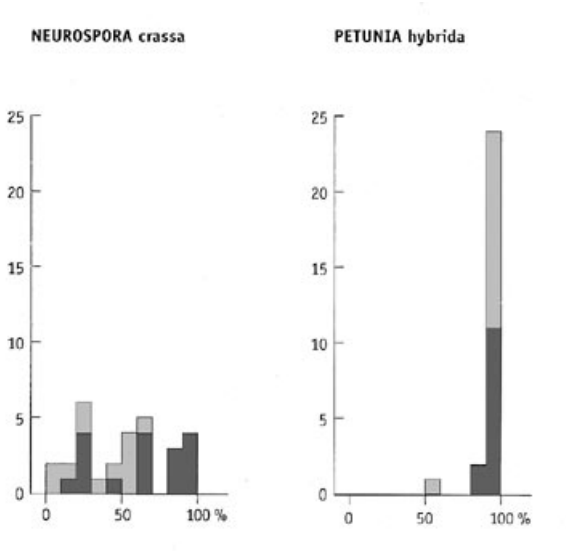

Figure 5. Distribution of the methylated C residues present in the CpG and CpNpG sequences. On the left are the data for *Neurospora* obtained by analysis of the data of Figure 2. On the right are the data for *Petunia* obtained from the same type of analysis using the published data of Meyer *et al.* (8). The abscissa shows the percentage of methylation of a given C residue and the ordinate indicates how many different DNA filaments have at any C residue with that particular percentage of methylation. The light boxes represent cytosines from the top strand, dark boxes cytosines from the bottom strand.

parentheses after each triplet. In contrast, the cytosines in the middle of the triplets ACC (27 ± 9) and GCG (14 ± 14) have <43% methylation, while the cytosines in the other six triplets have both high and low levels of methylation. In the case of *Petunia* it is not possible to find a triplet with a consistently high level of m^5C, except for the four triplets containing CpG (ACG, TCG, CCG, GCG); only in the case of the triplet CCC (3 ± 6) is the middle C almost free of methylation (Table 2).

This is evidence that methylation during the vegetative stage of the life cycle of *Neurospora* has some sequence specificity and that this is different from that of *Petunia*.

DISCUSSION

We have presented data on the pattern of cytosine methylation in the strain HTA5.7 containing at least four copies of the promoter of the *hph* transgene (16). Two different techniques have been used, the bisulphite method and Southern blot analysis of DNA digested by four methylation-sensitive restriction enzymes. The level of methylation is very high: 35% of all cytosines are methylated in the top strand and 61% in the bottom strand (Fig. 2). This level is much higher than the average level of cytosine methylation in wild-type *Neurospora* (1.5%) (25). We have reported that the strain HTA23 which has, after transformation, a single copy insert of the *hph* gene, has no m^5C, in accord with the Southern blot data published previously (16,18; data not shown). We do not know why there is a difference in the methylation level of the two strands. We can only note that the same difference exists in the data published on the pattern of methylation of a transgene in *Petunia* (8; Table 2). One possibility is that the bottom strand is more heavily methylated because it is transcribed.

The pattern of DNA methylation of HTA5.7 is complex. It is clear, however, that the methylation process is not a random one (Fig. 4). The cytosines in the middle of the triplets ACA, TCA, CCA, TCT, CCT, TCC, ACG and CCG have consistently high levels of methylation (Table 2). The cytosines in the middle of the triplets ACC and GCG have low levels of methylation (Table 2). The cytosines in the other six triplets have a wide range of methylation (Table 2).

It is clear from our data that cytosines in the CpG and CpNpG configurations are not a preferred target for cytosine methylation, as is the case in plants (5; Fig. 5).

The pattern of methylation reported here is very different from that reported earlier in *Neurospora* for the *am* gene after induction of RIP mutations, where 'most of the molecules assayed showed methylation at more than 80% of the cytosines and nearly half showed methylation at 95 to 100% of the cytosines' (6). It seems that the hypothetical methyltransferase of *Neurospora* has a different specificity in the sexual cycle than in the asexual one or that there is more than one methyltransferase, as suggested for *Ascobulus* (27).

With determination of the *in vivo* pattern of DNA methylation of the promoter region of the foreign gene *hph* in the multiple copy strain HTA5.7 we are in the position to characterize the hypothetical methyltransferase(s), once it is isolated. Only an isolated methyltransferase(s) that produces an *in vitro* methylation pattern similar to that reported here can be involved in silencing of the *hph* gene. The question of whether methylation is essential for silencing of the *hph* gene is still open. In a similar system, quelling or silencing of the *al*-1 gene in *Neurospora*, it has been reported that DNA methylation is not essential for silencing (28). However, the two silencing processes are quite different from each other (17). Only a complete analysis of the system, after isolation and characterization of the DNA methyltransferase(s) and its gene(s), can elucidate the still obscure but fascinating phenomenon of silencing of foreign genes in *Neurospora*.

ACKNOWLEDGEMENTS

We are greatful to J.Alonso, R.Hakenbeck, S.Meier-Ewert, W.Messer, T.A.Trautner and Jörn Walter for very helpful discussions and critical reading of the manuscript. Dedicated to T.A.Trautner on the occasion of his 65th birthday.

REFERENCES

1 McClelland,M., Nelson,M. and Raschke,E. (1994) *Nucleic Acids Res.*, **22**, 3640–3659.
2 Noyer-Weidner,M. and Trautner,T.A. (1993) In Jost,J.P. and Saluz,H.P. (eds), *DNA Methylation: Molecular Biology and Biological Significance*. Birkäuser, Basel, Switzerland, pp. 39–108.
3 Holliday,R. (1996) In Russo,V.E.A., Martienssen,R.A. and Riggs,A.D. (eds), *Epigenetics Mechanisms of Gene Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 5–28.
4 Yang,A.S., Jones,P.A. and Shibata,A. (1996) In Russo,V.E.A., Martienssen,R.A. and Riggs,A.D. (eds), *Epigenetics Mechanisms of Gene Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 77–94.
5 Finnegan,E.J. (1996) In Russo,V.E.A., Martienssen,R.A. and Riggs,A.D. (eds), *Epigenetics Mechanisms of Gene Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 127–140.
6 Selker,E.U., Fritz,D.Y. and Singer,M.J. (1993) *Science*, **262**, 1724–1728.
7 Guyon,C., Nogueira,T.I.V. and Faugeron,G. (1994) *J. Mol. Biol.*, **240**, 42–51.
8 Meyer,P., Niedenhof,I. and ten Lohuis,M. (1994) *EMBO J.*, **13**, 2084–2088.

9 Ingelbrecht,I., Van Houdt,H., Van Montagu,M. and Depicker,A. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 10502–10506.
10 Martienssen,R. and Baron,A. (1994) *Genetics*, **136**, 1157–1170.
11 Tasheva,E.S. and Roufa,D.J. (1994) *Mol. Cell. Biol.*, **14**, 5636–5644.
12 Tasheva,E.S. and Roufa,D.J. (1995) *Cell. Mol. Genet.*, **19**, 369–383.
13 Rein,T., Zorbas,H. and DePamphilis,M.L. (1997) *Mol. Cell. Biol.*, **17**, 416–426.
14 Rein,T., Natale,D.A., Gärtner,U., Niggemann,M., DePamphilis,M.L. and Zorbas,H. (1997) *J. Biol. Chem.*, **272**, 10021–10029.
15 Russo,V.E.A., Martienssen,R.A. and Riggs,A.D. (eds) (1996) *Epigenetics Mechanisms of Gene Regulation.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
16 Pandit,N.N. and Russo,V.E.A. (1992) *Mol. Gen. Genet.*, **234**, 412–422.
17 Russo,V.E.A., Lee,Y.-S. and Codón,A.C. (1996) In Russo,V.E.A., Martienssen,R.A. and Riggs,A.D. (eds), *Epigenetics Mechanisms of Gene Regulation.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 345–360.
18 Lee,Y.-S. (1996) Doktorarbeit (PhD thesis), Free University Berlin.
19 Staben,C., Jensen,B., Singer,M., Pollock,J., Schechtman,M., Kinsey,J. and Selker,E. (1989) *Fungal Genet. Newsl.*, **36**, 79–81.
20 Olek,A., Oswald,J. and Walter,J. (1996) *Nucleic Acids Res.*, **24**, 5064–5066.
21 Frommer,M., McDonald,L.E., Millar,D.S., Collis,C.M., Watt,F., Grigg,G.W., Molloy,P.L. and Paul,C.L. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 1927–1831.
22 Clark,S.J., Harrison,J., Paul,C.L. and Frommer,M. (1994) *Nucleic Acids Res.*, **22**, 2990–2997.
23 Feil,R., Charlton,J., Bird,A.P., Walter,J. and Reik,W. (1994) *Nucleic Acids Res.*, **22**, 695–696.
24 Raizis,A.M., Schmitt,F. and Jost,J.P. (1995) *Anal. Biochem.*, **226**, 161–166.
25 Russell,P.J., Wagner,S., Rodland,K.D., Feinbaum,R.L., Russell,J.P., Bret-Harte,M.S., Free,S.J. and Metzerberger,R.L. (1984) *Mol. Gen. Genet.*, **196**, 275–282.
26 Bull,J.H. and Wootton,J.C. (1994) *Nature*, **310**, 701–704.
27 Goyon,C., Barry,C., Grégoire,A., Faugeron,G. and Rossignol,J.-L. (1996) *Mol. Cell. Biol.*, **16**, 3054–3065.
28 Cogoni,C., Irelan,J.T., Schumacher,M., Schmidhauser,T.J., Selker,E.U. and Macino,G. (1996) *EMBO J.*, **15**, 3153–3163.