



# Evaluation of the Properties of the DNA Methyltransferase from *Aeropyrum pernix* K1

Mao Hayashi,<sup>a</sup> Keisuke Sugahara,<sup>b</sup> Akira Yamamura,<sup>a,b</sup>  Yasuhiro Iida<sup>a,b</sup>

<sup>a</sup>Department of Applied Chemistry and Bioscience, Graduate School of Engineering, Kanagawa Institute of Technology, Atsugi, Kanagawa, Japan

<sup>b</sup>Department of Applied Bioscience, Faculty of Applied Bioscience, Kanagawa Institute of Technology, Atsugi, Kanagawa, Japan

**ABSTRACT** Little is known regarding the DNA methyltransferases (MTases) in hyperthermophilic archaea. In this study, we focus on an MTase from *Aeropyrum pernix* K1, a hyperthermophilic archaeon that is found in hydrothermal vents and whose optimum growth temperature is 90°C to 95°C. From genomic sequence analysis, *A. pernix* K1 has been predicted to have a restriction-modification system (R-M system). The restriction endonuclease from *A. pernix* K1 (known as ApeKI from New England Biolabs Inc. [catalog code R06435]) has been described previously, but the properties of the MTase from *A. pernix* K1 (M.ApeKI) have not yet been clarified. Thus, we demonstrated the properties of M.ApeKI. In this study, M.ApeKI was expressed in *Escherichia coli* strain JM109 and affinity purified using its His tag. The recognition sequence of M.ApeKI was determined by methylation activity and bisulfite sequencing (BS-seq). High-performance liquid chromatography (HPLC) was used to detect the position of the methyl group in methylated cytosine. As a result, it was clarified that M.ApeKI adds the methyl group at the C-5 position of the second cytosine in 5'-GCWGC-3'. Moreover, we also determined that the MTase optimum temperature was over 70°C and that it is strongly tolerant to high temperatures. M.ApeKI is the first highly thermostable DNA (cytosine-5)-methyltransferase to be evaluated by experimental evidence.

**IMPORTANCE** In general, thermophilic bacteria with optimum growth temperatures over or equal to 60°C have been predicted to include only N<sup>4</sup>-methylcytosine or N<sup>6</sup>-methyladenine as methylated bases in their DNA, because 5-methylcytosine is susceptible to deamination by heat. However, from this study, *A. pernix* K1, with an optimum growth temperature at 95°C, was demonstrated to produce a DNA (cytosine-5)-methyltransferase. Thus, *A. pernix* K1 presumably has 5-methylcytosine in its DNA and may produce an original repair system for the expected C-to-T mutations. M.ApeKI was demonstrated to be tolerant to high temperatures; thus, we expect that M.ApeKI may be valuable for the development of a novel analysis system or epigenetic editing tool.

**KEYWORDS** *Aeropyrum pernix*, archaea, DNA methylation, DNA methyltransferase, epigenetics

**D**NA methylation is an epigenetic modification that is found in eukaryotes, prokaryotes, and viruses (1). DNA methylation is involved in significant biological functions, such as X chromosome inactivation in females, genomic imprinting, and gene expression regulation in eukaryotic cells (2, 3). In prokaryotes, DNA methylation also has significant functions, including roles in restriction-modification (R-M) systems, indicating daughter strands for the mismatch repair system following DNA replication, and in cell cycle regulation (4, 5). Methyl groups are added by a DNA methyltransferase (MTase) from S-adenosylmethionine (SAM) as the donor molecule (6). MTases essentially have two types in mammals as follows: the maintenance MTases have a preference for hemimethylated DNA (7), and the *de novo* type MTases have no such preference (8). Conversely, this distinction is less clear in bacteria/archaea.

**Citation** Hayashi M, Sugahara K, Yamamura A, Iida Y. 2021. Evaluation of the properties of the DNA methyltransferase from *Aeropyrum pernix* K1. *Microbiol Spectr* 9:e00186-21. <https://doi.org/10.1128/Spectrum.00186-21>.

**Editor** Henning Seedorf, Temasek Life Sciences Laboratory

**Copyright** © 2021 Hayashi et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Yasuhiro Iida, [iida@bio.kanagawa-it.ac.jp](mailto:iida@bio.kanagawa-it.ac.jp).

**Received** 16 June 2021

**Accepted** 10 August 2021

**Published** 29 September 2021

In bacteria/archaea, N<sup>6</sup>-methyladenine (m6A), 5-methylcytosine (m5C), and N<sup>4</sup>-methylcytosine (m4C) have all been reported.

In prokaryotes, the functions of N<sup>6</sup>-methyladenine in the genomic DNA have been studied in DNA replication and repair, gene expression regulation, and immunity via the R-M system. (9). In eukaryotic cells, however, m6A had not been discovered until recently, when it was identified in *Caenorhabditis elegans* by methylated DNA immunoprecipitation sequencing (MeDIP-seq) (10). m6A has been identified in the DNA of many eukaryotes, including *Drosophila melanogaster*, *Bombyx mori*, *Arabidopsis thaliana*, *Mus musculus*, and *Homo sapiens*, and its function is suggested to be cell cycle control and gene expression regulation (11).

5-Methylcytosine has been widely studied in eukaryotic cells, where it acts in general terms as a transcription repressor. m5C is enriched in CpG islands in promoter regions, where DNA methylation inhibits gene expression due to the inhibition of transcription factor binding either directly or via mediation of methyl binding domain proteins that recruit histone deacetylase (HDAC) (12). Furthermore, m5C is related to development and differentiation (1). In *Escherichia coli*, DNA cytosine methylase (Dcm) is not part of an R-M system, and even in this mesophilic organism, the resulting m5C-to-T mutations are repaired by a “very short patch” repair system (13).

N<sup>4</sup>-methylcytosine is observed in thermophilic bacteria and archaea (14). The studies of Wang et al. and Ehrlich et al. showed that m5C is changed to thymine by demethylation under high temperatures (15, 16). The optimum growth temperature of thermophilic bacteria is over 60°C. Thus, thermophilic bacteria have been suggested to use m4C, which is resistant to heat, instead of m5C to minimize genome mutation (17).

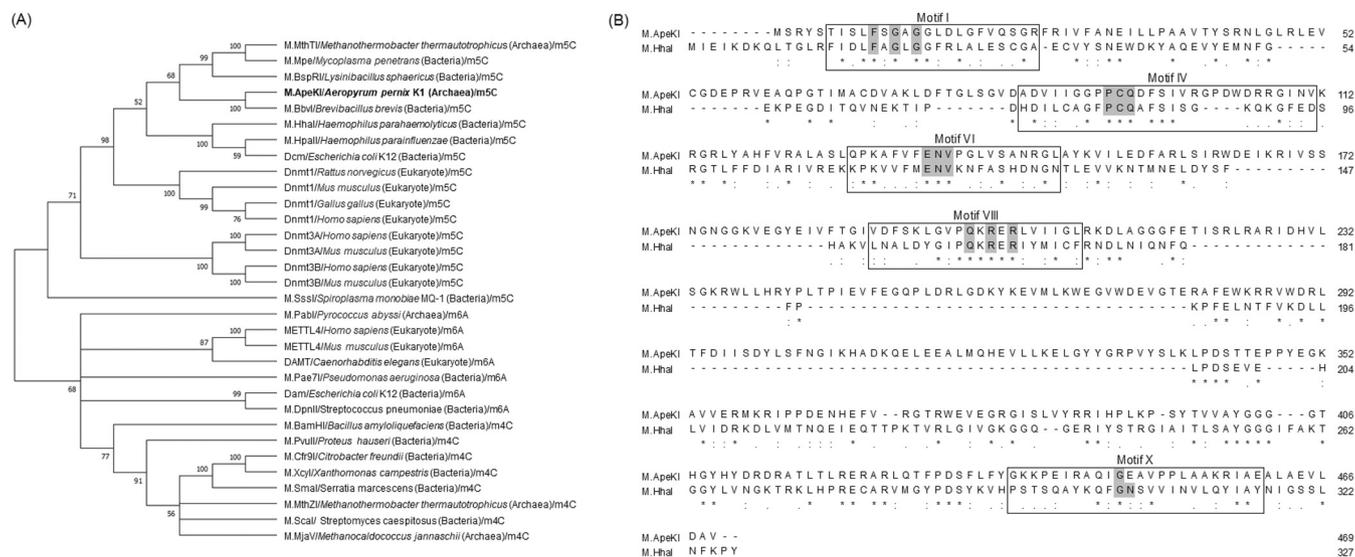
In particular, m6A, m5C, and m4C, in the context of R-M systems, have been well studied in bacteria. An *E. coli* R-M system, for example, was reported as an immunity system around 70 years ago (18). The R-M system constitutes two factors as follows: restriction endonucleases (REases), which cleave-specific sequences, and MTases, which add methyl groups within the specific sequence. This system, among other roles, protects host cells from phage infection. The invasive phage DNA is cleaved by the REase, whereas the genomic DNA is protected by the cognate DNA MTase (19).

Although MTases have been studied in various biological species, little is known regarding thermostable MTases. The study of Watanabe et al. showed that M.PabI, an MTase from the hyperthermophilic archaeon *Pyrococcus abyssi*, had the highest optimum temperature among then-known DNA (adenine-6)-methyltransferases (20). M.PabI is part of an R-M system and recognizes the sequence 5'-GTAC-3' (20). The optimum temperature of M.PabI is 85°C, and activity has been identified at temperatures higher than 95°C (it retains at least half of its activity after incubation at 95°C for 9 min) (20).

Conversely, a thermostable DNA (cytosine-5)-methyltransferase has not yet been reported. A thermostable DNA cytosine methyltransferase was reported by Nölling and de Vos (21). In the article, MTase from *Methanobacterium thermoformicum* THF (M.MthTI) was predicted to be DNA (cytosine-5)-methyltransferase just from amino acid alignment, and there are no experiments demonstrating the position at which the cytosine was methylated by M.MthTI.

In this study, we focus on an MTase of the hyperthermophilic archaeon, *Aeropyrum pernix* K1, whose genome analysis is complete (22). According to the National Institute of Technology and Evaluation, the optimum growth temperature of *A. pernix* K1 is 90°C to 95°C. The enzymes of *A. pernix* K1 are all presumed to be thermostable. Thus, they have been used in biotechnology applications. For example, uracil-DNA glycosylase from *A. pernix* K1 is used in hot-start PCR (23). *A. pernix* K1 is predicted to have an R-M system. The REase of *A. pernix* K1 is already known as ApeKI and recognizes the sequence 5'-GC(A/T)GC-3'. In contrast, the MTase of *A. pernix* K1, which was designated as M.ApeKI (24) (UniProt accession number [Q9YDP3](#)) has only been evaluated by homology analysis (25). Watanabe et al. mention that not only the activity of M.ApeKI but also the expression of the enzyme has not yet been reported (20).

In this study, we investigated the properties of M.ApeKI. The recognition sequence of M.ApeKI was evaluated by methylation activity and bisulfite sequencing (BS-seq),



**FIG 1** Prediction of modified DNA base by MTases. (A) The evolutionary history was inferred using the neighbor-joining method (34). The optimal tree is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates) is shown next to the branches (35). The evolutionary distances were computed using the p-distance method (36) and are in the units of the number of amino acid differences per site. This analysis involved 32 amino acid sequences. All positions containing gaps and missing data were eliminated (complete deletion option). There were a total of 169 positions in the final data set. Evolutionary analyses were conducted in MEGA X (37). (B) Amino acid alignment. Asterisk (\*), perfect alignment; colon (:), a site belonging to a set exhibiting strong similarity; period (.), a site belonging to a group exhibiting weak similarity. For the common motif of DNA (cytosine-5)-methyltransferase, we referred to the study of M.HhaI (28, 29). As a result, in the highly conserved motifs, FXGXG in motif I, PCQ in motif IV, ENV in motif VI, and QXRXR in motif VIII were observed; however, GN in motif X was not observed.

and we used high-performance liquid chromatography (HPLC) to determine the position of methyl groups in methylated cytosine. We demonstrated by experimental evidence that M.ApeKI is a DNA (cytosine-5)-methyltransferase and that it recognizes the sequence 5'-GCWGC-3', the same recognition sequence as the associated endonuclease ApeKI. Furthermore, we discovered that the optimum temperature of M.ApeKI is over 70°C. We suggest that M.ApeKI has strong thermostability, as activity was still observed after incubation at 90°C for 10 h. Therefore, we conclude that *A. pernix* K1 has an R-M system and that M.ApeKI is a thermostable DNA (cytosine-5)-methyltransferase.

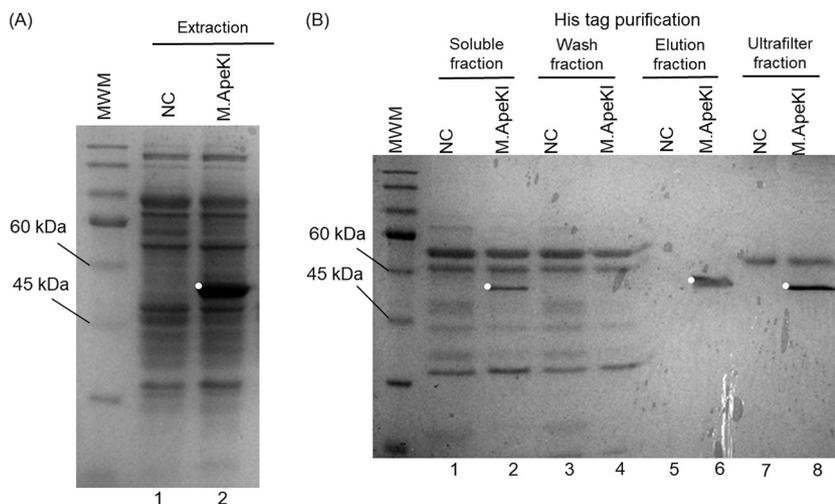
## RESULTS

**Phylogenetic tree construction and amino acid alignment.** We investigated whether M.ApeKI is likely to be a DNA (adenine-6)-methyltransferase, DNA (cytosine-5)-methyltransferase, or DNA (cytosine-4)-methyltransferase by phylogenetic tree construction (Fig. 1A) and amino acid alignment. In Fig. 1A, DNA (adenine-6)-methyltransferase, DNA (cytosine-5)-methyltransferase, or DNA (cytosine-4)-methyltransferase were used, and these have been reviewed in UnprotKB/Swiss-Prot. The results indicated that M.ApeKI and the DNA (cytosine-5)-methyltransferase share a recent common ancestor.

Moreover, the amino acid sequence of M.ApeKI was aligned with the known m5C bacterial MTase M.HhaI. Figure 1B shows the results of the amino acid alignment. The alignment of the amino acid sequence was analyzed using multiple sequence comparison by log-expectation (MUSCLE) (26).

**Extraction and purification.** Figure 2 shows the results of SDS-PAGE for the extracted and purified M.ApeKI. The estimated molecular weight of M.ApeKI was 56.4 kDa. The negative control (NC) was from *E. coli*, including the pCold I vector without M.ApeKI, and it was treated in the same way as the M.ApeKI sample. Figure 2A shows the results of the extraction of the NC and M.ApeKI. The band near the estimated molecular weight of M.ApeKI was unobserved in the NC lane. Conversely, in the M.ApeKI lane, a dark band was observed in the estimated molecular weight of M.ApeKI.

Figure 2B shows the results of the purification and ultrafiltration of M.ApeKI. In the soluble fraction, a single band of M.ApeKI was observed near the target molecular weight,



**FIG 2** Extraction and purification of M.ApeKI. (A) The results of extraction. The molecular weight marker was protein marker PM2500 from Smobio. Negative control (NC) was not included. (B) Results of purification. The band near to the target molecular weight was observed in the soluble fraction, elution fraction, and ultrafilter fraction in the M.ApeKI lane (white circle). Polyacrylamide gel electrophoresis was performed in 12% separation gel in a buffer containing 25 mM Tris, 0.1% SDS, and 192 mM glycine at 30 mA. It was stained with EzStain AQua from CBB.

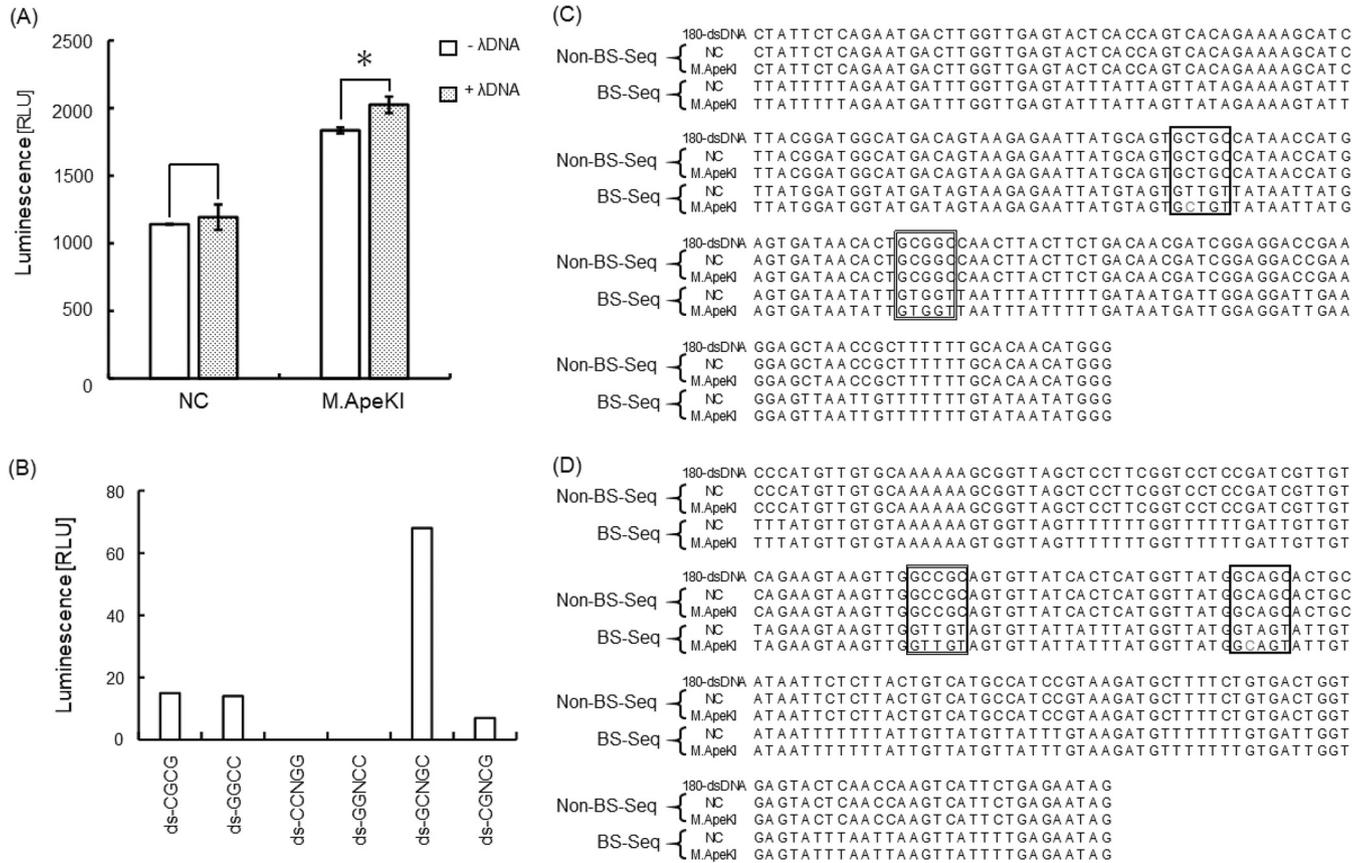
but that band was not observed in the NC lane (Fig. 2B, lanes 1 and 2). After the extraction, heat treatment was effective for a shift of M.ApeKI from the insoluble fraction to the soluble fraction. In the wash fraction, the band was unobserved (Fig. 2B, lanes 3 and 4). In the elution fraction, the band that is near the target molecular weight of M.ApeKI was not observed in the NC lane but was observed in the M.ApeKI lane (Fig. 2B, lanes 5 and 6). In addition, in the ultrafiltration fraction, a band was not observed near the target molecular weight of M.ApeKI in the NC lane, but a band was detected over 60 kDa in the M.ApeKI lane (Fig. 2B, lane 7). A single band was detected near the target molecular weight of M.ApeKI in the M.ApeKI lane, and a band was observed over 60 kDa, similar to that in the NC lane.

**Determination of recognition sequence.** The recognition sequence of M.ApeKI was characterized by methylation activity. Figure 3A shows the results of the confirmation of DNA methylation activity using double-stranded  $\lambda$  DNA as a substrate and a luminescence assay for the conversion of SAM (methyl donor) to its product S-adenosylhomocysteine (SAH). The results showed that in M.ApeKI, there was a significant difference between including  $\lambda$  DNA (+ $\lambda$  DNA) and excluding  $\lambda$  DNA ( $-\lambda$  DNA). In the NC, however, there was no significant difference. Thus, M.ApeKI has DNA methylation activity.

Figure 3B shows the results of the characterization of the recognition sequence by methylation activity using double-stranded oligonucleotides. When we used double-stranded GCNGC (ds-GCNGC) as a substrate, we observed the highest activity. The methylation activities of M.ApeKI on the other substrates showed little or no activity. Thus, M.ApeKI has predicted specificity for 5'-GCNGC-3'.

Figure 3C and D show the results of the determination of the base (N) in the 5'-GCNGC-3' sequence (N = A, T, C, G) that used 180-bp double-stranded DNA (180-dsDNA). In this experiment, methylcytosine was detected by BS-seq. The sequence of the substrate is shown in Table 1. We analyzed the sense strand (Fig. 3C) and the antisense strand (Fig. 3D) in 180-dsDNA. From the results of BS-seq, the second cytosine in 5'-GCTGC-3' and 5'-GCAGC-3' sequence was observed to have been modified to methylcytosine. Conversely, no cytosine residues in the 5'-GCGGC-3' and 5'-GCCGC-3' sequences were observed to have been methylated.

**Determination of the modified position in methyl cytosine.** We determined the modified position in methyl cytosine by HPLC. Figure 4A shows the results of the determination of the modified position in methyl cytosine by M.ApeKI, and the magnification view for the retention time of 2'-deoxy-5-methylcytidine (m5dC) is shown in Fig. 4B. Double-stranded GCWGC (ds-GCWGC) was used as the substrate (Table 1) and was detected by HPLC after



**FIG 3** Determination of recognition sequence. (A) Activity of M.ApeKI was measured by MTase-Glo methyltransferase assay. Statistical significance (*P* value) was determined by two-tailed Student's *t* test (*P* < 0.05). (B) To estimate the recognition sequence of M.ApeKI, MTase-Glo methyltransferase assay was used. Bar chart values indicate correction values, which subtracts -λ DNA from +λ DNA. BS-seq was used for the determination of the recognition sequence of M.ApeKI sense strand (C) and antisense strand (D). "Non-BS-seq" indicates 180-dsDNA not treated with bisulfite, and "BS-seq" indicates 180-dsDNA treated with bisulfite. "NC" indicates 180-dsDNA treated with the negative control, and "M.ApeKI" indicates 180-dsDNA treated with M.ApeKI. The single border indicates the 5'-GCTGC-3' and 5'-GCAGC-3' sequences, and the double border indicates 5'-GCGGC-3' and 5'-GCCGC-3'. Gray letters indicate methylcytosine in 5'-GCTGC-3' and 5'-GCAGC-3'.

digestion. Five deoxyribonucleosides (2'-deoxycytidine, dC; 2'-deoxy-4-methylcytidine, m4dC; 2'-deoxyadenosine monohydrate, dA; 2'-deoxyguanosine hydrate, dG; thymidine, T), except m5dC, were detected in the NC (Fig. 4B). The ds-GCWGC was treated with purified M.ApeKI (solid line in Fig. 4A) or M.SssI as the positive control (methylates CpG to m5CpG; data not shown), and six deoxyribonucleosides (dC, m4dC, m5dC, dA, dG, and T) were detected. The peak of m5dC in M.ApeKI was consistent with the peak of m5dC from the authentic sample and from M.SssI, which is known as DNA (cytosine-5)-methyltransferase (27).

**Optimum temperature and thermostability of M.ApeKI.** The optimum temperature and thermostability of M.ApeKI were determined by measurement of methylation activity via luminescent assay of the SAM product SAH. In this experiment, ds-GCWGC was used as a substrate (Table 1). The lowest activity was observed at 10°C (Fig. 5A), and at over 70°C, the highest luminescence value was detected. Figure 5B shows the results of the thermostability investigation. M.ApeKI was heated at 90°C for 10 min, 30 min, 60 min, and 10 h before the evaluation. As a result, the activity decreased by 40% when M.ApeKI was heated at 90°C for 30 min, and the activity decreased by 60% when M.ApeKI was heated at 90°C for 60 min. Furthermore, when M.ApeKI was heated at 90°C for 10 h, the activity decreased by 66% compared with that of the unheated sample.

**DISCUSSION**

In this study, we characterize the properties of M.ApeKI. The phylogenetic results suggest that M.ApeKI is more closely related to DNA (cytosine-5)-methyltransferases than any other MTase. DNA (cytosine-5)-methyltransferases have 10 conserved motifs,

**TABLE 1** Primers, double-stranded DNAs, and substrates used in this study

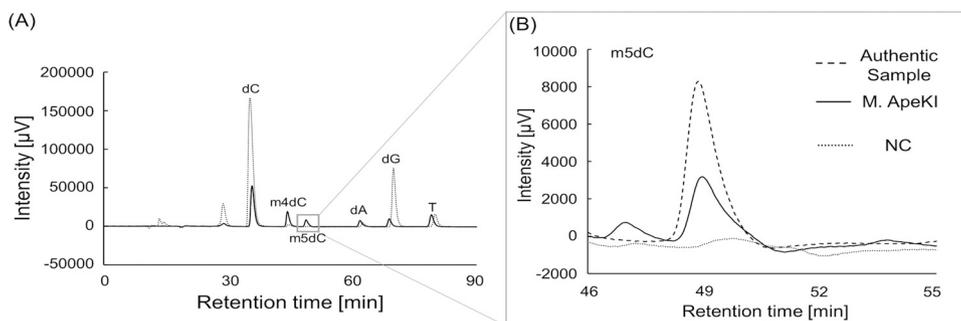
Expt	Name	Sequence (5' to 3')	bp	Tm [°C] <sup>a</sup>	
Constructed plasmid	APE ORF(EcoRI) Fw	GCGGAATTCATGTCACGATATAGTACCATTAGC	33	62	
	Dnmt1 (APE) PstI+fin Rv	CATCTGCAGTTATCACACAGCATCCGAACCTC	33	63	
Sequence analysis	pCold I_Universal_F	GTAAGGCAAGTCCCTTCAAGAG	22	57.7	
	pCold-R Primer	GCCAGGATCTTAGATTCTG	20	55.4	
	Dnmt1 (APE)primer walk Fw	AACGGGAACGCTGGTAATCATCG	24	61	
	Dnmt1 (APE)primer walk Rv	CCAAGACGATCCAACGGTTGAC	22	60	
Determination of recognition sequence	MTase-Glo methyltransferase assay	double-stranded CGCG (ds-CGCG)	CGCGCGCGCGCGCGG	16	75
		double-stranded GGCC (ds-GGCC)	GGCCGGCCGGCCGGCC	16	72
		double-stranded CCNGG (ds-CCNGG)	CCAGGCCGGCCGGCCCTGG	20	73
		double-stranded GGNC (ds-GGNC)	GGACCGGCCGGCCGGTCC	20	73
		double-stranded GCNGC (ds-GCNGC)	GCAGCGCCGGCCGGCGTGC	20	76
		double-stranded CGNCG (ds-CGNCG)	CGACGCGCCGGCCGGCGTGC	20	75
BS-seq	180-bp double-stranded DNA (180-dsDNA)	CTATTCTCAGAATGACTTGGTTGAGTACTCACC	180		
		AGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGA			
		GAATTATGCAGTCTGCCATAACCATGAGTGATAA			
		CACCTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAG			
		GAGCTAACCGCTTTTTTGCACAACATGGG			
BS-seq	BS-seq Fw	TTATTTTAGAATGATTTGGTTGAG	25	50	
		BS-seq Rv	CCCATATTATACAAAAACAATTA	25	48
		Non-Bs-seq Fw	CTATTCTCAGAATGACTTGGTTGAG	25	55
		Non-Bs-seq Rv	CCCATGTTGTCAAAAAAGCGGTTA	25	61
Determination of modified base	GCWGC Fw double-stranded GCWGC (ds-GCWGC)	GCAGCGCTGCGCAGCGCTGCGCAGCGTGCAGC	35	83	
		GCWGC Rv	GTCGCGCAGCGCTGCGCAGCGCTGCGCAGCGTGC	35	83

<sup>a</sup>The temperature of melting (Tm) is defined as the temperature at which one half of the DNA duplex will dissociate to become single stranded and indicates the duplex stability.

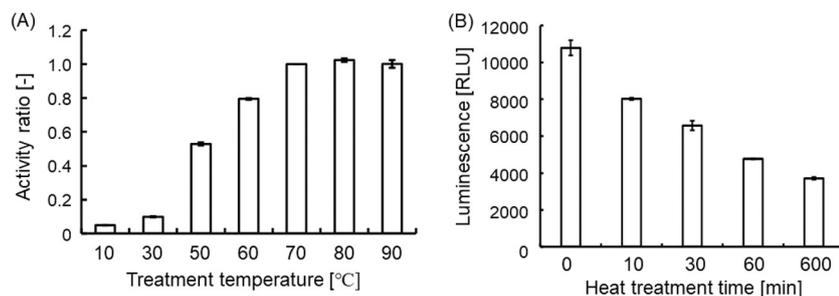
and motifs I, IV, VI, VIII, and X were reported to be especially highly conserved (28, 29). We focused on these motifs. As a result, conservation of amino acid motifs I, IV, VI, and VIII (but not X) was observed. The highly conserved motifs are reported to play significant roles (28, 30) as follows: motifs I and X are involved in SAM binding, motifs IV and VI are involved in catalysis, and motif VIII is involved in protonation of the target cytosine. The motifs involved in catalysis were observed in M.ApeKI; thus, it was predicted to be an active DNA (cytosine-5)-methyltransferase.

To determine the recognition sequence, ds-GCNGC showed the highest methylation activity. Moreover, we performed the determination of the base (N) in 5'-GCNGC-3' by BS-seq. As a result, methylcytosine was observed in the second cytosines of 5'-GCTGC-3' and 5'-GCAGC-3'. Conversely, the methylation activities of M.ApeKI were very low when ds-CGCG, ds-GGCC, or ds-CGNCG were used as substrates. However, from the results of BS-seq, a methylcytosine was not observed in 5'-GGCC-3'. Therefore, we suggested that, while 5'-CGCG-3' and 5'-CGNCG-3' were not assessed via BS-seq, these sequences were not methylated by M.ApeKI. We concluded that M.ApeKI acts on the sequence 5'-GC(A/T)GC-3'.

We also investigated the position where the methyl group was added to cytosine by M.ApeKI and found that a peak with the same retention time as that of m5dC was observed in the substrate treated with M.ApeKI. A peak was also observed at the same



**FIG 4** Determination of the modified position of methyl groups in methylcytosine. All nucleotides were detected at 287 nm, the absorption maximum wavelength of m5dC (data not shown). (A) The result of HPLC analysis. Six deoxyribonucleosides (dC, m4dC, m5dC, dA, dG, and T) as authentic sample, the retention times were as follows: dC = 35.8, m4dC = 44.4, m5dC = 48.9, dA = 61.9, dG = 68.9, T = 79.1 min. (B) Magnified view of the retention time of m5dC.



**FIG 5** Evaluation of the optimum temperature and investigation of thermostability. The activity of M.ApeKI was measured by MTase-Glo methyltransferase assay. (A) The detected optimum temperature. M.ApeKI was reacted with double-stranded DNA as a substrate. Activity ratio was based on the result of treatment at 70°C. (B) The evaluated thermostability. M.ApeKI was reacted with double-stranded DNA as a substrate (Table 1).

retention time of m5dC from the substrate treated with M.Sss1, which is a known DNA (cytosine-5)-methyltransferase (27), and a peak with the same retention time as m5dC was not observed in the NC. A peak with the same retention time as that of m4dC was observed in the substrate treated with M.ApeKI or M.Sss1; however, the NC, which was not treated with any enzyme, also showed a peak at m4dC. The intensity ratio of adenine to m4dC was the same as that of the substrate treated with M.ApeKI or M.Sss1 and the NC, which was not treated with any sample (data not shown). This result suggests that the use of the enzymes was not related to a peak at the same retention time as m4dC. The study of O’Brown et al. showed that enzymes for digesting DNA contain unmethylated DNA and methylated DNA (14). It is interpreted from this result that M.ApeKI is a DNA (cytosine-5)-methyltransferase. In general, thermophilic bacteria that modify cytosines have DNA (cytosine-4)-methyltransferases that allow them to avoid m5C-to-T mutation. Therefore, the details regarding m5C in thermophilic archaea require further elucidation.

The highest activity was observed after treatment with M.ApeKI at over 70°C, and the activities at 70°C, 80°C, and 90°C were not significantly different. However, DNA as a substrate and SAM as a methyl group donor were predicted to be unstable at high temperature (31, 32); thus, we were concerned that the accurate measurement of methylation activity is difficult at high temperature. In this study, we concluded that the optimum temperature of M.ApeKI is 70°C to 90°C. The novel method for the determination of the optimum temperature of thermostable MTase should be refined further. Moreover, the activity of M.ApeKI was still detected when it was heated at 90°C for 10 h, indicating that M.ApeKI is tolerant to high temperatures.

We concluded that M.ApeKI is a thermostable DNA (cytosine-5)-methyltransferase, which adds a methyl group to the second cytosine in 5’-GCWGC-3’. This recognition sequence corresponds to the recognition sequence of ApeKI. Thus, we suggest that *A. pernix* K1 has an R-M system. Further investigation into the construction of M.ApeKI and its functions will help in elucidating the thermostability mechanism.

## MATERIALS AND METHODS

**Materials and chemicals.** For all experiments, primers were purchased from Eurofins Genomics (Ebersberg, Germany). Sequence analysis was performed by Fasmac Co., Ltd. (Kanagawa, Japan) using the MUSCLE from the EMBL-EBI. The known methyltransferase data were from the UniProt database. The sequences for primers and substrates are shown in Table 1. In the substrate of the determination of recognition sequence by BS-seq, the underline shows the 5’-GC(T/A) GC-3’ sequence, and the bold shows the 5’-GC(G/C) GC-3’ sequence. These were predicted to be the recognition sequence of M.ApeKI.

For the construction of the plasmid, the M.ApeKI synthetic sequence was purchased from Eurofins Genomics. The *Ex Taq* hot-start PCR reagent and the restriction enzymes were purchased from TaKaRa Bio Inc. (Shiga, Japan). The vector, pCold I, used in this study was also purchased from TaKaRa. The gel purification system used was a FavorPrep GEL/PCR purification minikit from Favorgen Biotech Corporation (Ping-Tung, Taiwan). Ligation Mighty Mix from TaKaRa was used for ligation. The *E. coli* strain, JM109, was purchased from Toyobo Co., Ltd. (Osaka, Japan), and Luria-Bertani (LB) medium was purchased from Nacalai Tesque, Inc. (Kyoto, Japan).

The extraction of M.ApeKI was performed using protease inhibitor cocktail for general use (100×) from Nacalai Tesque, Inc., and M.ApeKI was extracted using an ultrasonic disrupter (UD-201; Tomy Digital Biology Co., Ltd., Tokyo, Japan). M.ApeKI was purified using a HisLink spin protein purification system (Promega Corporation, WI, USA), and liquid was replaced using a 3K Omega membrane (Pall Corporation, NY, USA). The electrophoresis reagent gel staining solution used after the SDS-PAGE was EzStain Aqua from ATTO Corporation (Tokyo, Japan).

**TABLE 2** DNA methyltransferases used in this phylogenetic tree

Type or no.	Accession no. in Uniprot	Protein name	Scientific name	Organism
DNA (cytosine-5)-methyltransferase				
1	P13906	M.BspRI	<i>Lysinibacillus sphaericus</i>	Bacteria
2	P05102	M.HhaI	<i>Haemophilus parahaemolyticus</i>	
3	P0AED9	Dcm	<i>Escherichia coli</i> K-12	
4	P15446	M.HpaII	<i>Haemophilus parainfluenzae</i>	
5	P15840	M.SssI	<i>Spiroplasma monobiae</i> MQ-1	
6	Q8EVR5	M.Mpe	<i>Mycoplasma penetrans</i>	
7	P34905	M.BbvI	<i>Brevibacillus brevis</i>	
8	P29567	M.MthTI	<i>Methanothermobacter thermautotrophicus</i>	Archaea
9	Q92072	Dnmt1	<i>Gallus gallus</i>	Eukaryote
10	P26358	Dnmt1	<i>Homo sapiens</i>	
11	Q9Z330	Dnmt1	<i>Rattus norvegicus</i>	
12	P13864	Dnmt1	<i>Mus musculus</i>	
13	Q9Y6K1	Dnmt3A	<i>Homo sapiens</i>	
14	O88508	Dnmt3A	<i>Mus musculus</i>	
15	Q9UBC3	Dnmt3B	<i>Homo sapiens</i>	
16	O88509	Dnmt3B	<i>Mus musculus</i>	
DNA (adenine-6)-methyltransferase				
1	P0AEE8	Dam	<i>Escherichia coli</i> K-12	Bacteria
2	P04043	M.DpnII	<i>Streptococcus pneumoniae</i>	
3	P05103	M.Pae7I	<i>Pseudomonas aeruginosa</i>	
4	Q9V2B5	M.PabI	<i>Pyrococcus abyssi</i>	Archaea
5	Q09956	DAMT	<i>Caenorhabditis elegans</i>	Eukaryote
6	Q8N3J2	METTL4	<i>Homo sapiens</i>	
7	Q3U034	METTL4	<i>Mus musculus</i>	
DNA (cytosine-4)-methyltransferase				
1	P11409	M.PvuII	<i>Proteus hauseri</i>	Bacteria
2	P23941	M.BamHI	<i>Bacillus amyloliquefaciens</i>	
3	P14243	M.Cfr9I	<i>Citrobacter freundii</i>	
4	P14230	M.SmaI	<i>Serratia marcescens</i>	
5	O52692	M.ScaI	<i>Streptomyces caespitosus</i>	
6	P30774	M.XcyI	<i>Xanthomonas campestris</i> pv. <i>cyanopsidis</i>	
7	Q58893	M.MjaV	<i>Methanocaldococcus jannaschii</i>	Archaea
8	P29568	M.MthZI	<i>Methanothermobacter thermautotrophicus</i>	

Bisulfite treatment was performed using an EpiTect Fast DNA bisulfite kit from Qiagen (Hilden, Germany), and the sample was amplified using KOD-Multi&Epi from Toyobo.

To measure methylation activity, we used an MTase-Glo methyltransferase assay kit from Promega and a plate reader from Mithras LB940 from Berthold Technologies (Baden-Württemberg, Germany). For the substrate, a double-stranded oligonucleotide was produced using thermal cycler FTGENE-Y2 from Techne (Staffordshire, UK).  $\lambda$  DNA from  $\lambda$  Hind III (TaKaRa) was used as the substrate for the confirmation of methylation activity.

The HPLC system comprised a three-line degasser (JASCO Corporation, Tokyo, Japan), a ternary gradient unit (JASCO), an intelligent column oven (JASCO), an intelligent HPLC insert pump (JASCO), an intelligent UV-visible (UV-Vis) detector and interface box (JASCO), and an InertSustain AQ-C<sub>18</sub> column (GL Sciences Inc., Tokyo, Japan). dC, dA, T, dG, and m5dC were purchased from Tokyo Chemical Industry (Japan). m4dC was purchased from Biosynth Carbosynth (Berkshire, UK). All samples were filtrated using a syringe-driven filter (0.20  $\mu$ m) from Merck Millipore (MA, USA) before injection to HPLC. The double-stranded oligonucleotide was digested using Nucleoside Digestion Mix from New England BioLabs Inc. (NEB, MA, USA).

**Phylogenetic tree construction and amino acid alignment.** For this analysis, M.ApeKI was compared with DNA (cytosine-5)-methyltransferase, DNA (cytosine-4)-methyltransferase, and DNA (adenine-6)-methyltransferase. Table 2 shows the MTases list for the construction of a phylogenetic tree. The phylogenetic tree (Fig. 1A) was constructed using MEGA-X software and the neighbor-joining method. For alignment analysis, MUSCLE was used.

**Construction of the plasmid.** The vector was constructed for the expression of M.ApeKI in *E. coli*. M.ApeKI was amplified by PCR using *Ex Taq* HS with a template based on the synthesized plasmid. The conditions of PCR were as follows: 98°C for 30 s, 98°C for 10 s, 58°C for 30 s, 72°C for 90 s, and 72°C for 90 s. Then, the sample was held at 4°C. The two primers that were used are the EcoRI site before the M.ApeKI sequence, and the other is the M.ApeKI sequence in addition to the stop codon and the PstI site (shown in Table 1).

After amplification, the PCR product was gel purified. The pCold I vector and the purified PCR product were reacted at 37°C for 2 h with PstI and EcoRI. After the reaction, the product was gel purified, and then, M.ApeKI was introduced into pCold I.

The recombinant plasmid was transformed into *E. coli* JM109 by heat shock treatment. The transformed *E. coli* was cultured at 37°C in 50 ml of LB medium with ampicillin. The plasmid was extracted

from the cultured *E. coli* by the alkaline lysis method. The extracted plasmid was gel purified, and the insert sequence was analyzed. The primers used for sequence analysis are shown in Table 1.

**Extraction and purification of M.ApeKI.** M.ApeKI was extracted and purified for the evaluation of its properties. The transformed *E. coli* was cultured at 37°C overnight in 100 ml of LB medium, including ampicillin as a preculture. Then, it was cultured in water at 15°C for 30 min and incubated at 15°C for 24 h after the addition of isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM. All of the culture medium in the 50-ml tube was centrifuged at  $7,000 \times g$  for 5 min at 4°C. The supernatant was discarded, and cells were washed in 0.2 M phosphoric acid buffer (pH 7.4). Cells were suspended in 2 ml of 0.2 M phosphoric acid buffer and protease inhibitor cocktail for general use (100 $\times$ ) and were extracted using an ultrasonic disrupter. The extract was placed in a 1.5-ml tube, was heated at 90°C for 10 min, and then centrifuged at  $9,400 \times g$  for 30 min at 4°C. The supernatant was regarded as a soluble fraction and was purified. The elution buffer was replaced with 1 $\times$  reaction buffer (Promega; no. TM453) by ultrafiltration. In this study, M.ApeKI as used for evaluation contains 15 amino acids including a His tag at the N terminus.

**Measurement of methylation activity.** The recognition sequence and optimum temperature were determined by measurement of methylation activity. For the measurement of the DNA methylation activity of M.ApeKI, 0.39 mmol of substrate, 19  $\mu$ g of the purified M.ApeKI, 32 nmol of SAM, and 0.1 mg/ml of bovine serum albumin (BSA) were mixed. The reactant was evaluated using the activity, and the level of luminescence of the reaction solution was measured by using a plate reader. A double-stranded oligonucleotide as a substrate was produced using a thermal cycler that contained 45  $\mu$ M PCReady primer and annealing buffer (1 mM Tris-HCl, 100 mM NaCl, and 0.1 mM EDTA). The conditions after incubation were for 5 min at 95°C, and then the temperature was decreased by 1°C/min for 25 min. To confirm the methylation activity of M.ApeKI,  $\lambda$  HindIII digest was used, and the reactant was incubated at 40°C for 2 h. To estimate the optimum temperature of M.ApeKI, ds-GCWGC was used as the substrate, and the reactant was incubated at 40°C for 2 h. To evaluate the optimum temperature of M.ApeKI, the reactant was incubated at 10°C, 30°C, 50°C, 60°C, 70°C, 80°C, and 90°C for 2 h.

**Determination of the recognition sequence of M.ApeKI.** The recognition sequence of M.ApeKI was determined by BS-seq. The conditions of treatment were the same as those for the measurement of methylation activity. The substrate used was an amplicon that included a part of the  $\beta$ -lactamase sequence, and the reactant was incubated at 40°C for 24 h. The bisulfite treatment was performed using an EpiTect Fast DNA bisulfite kit. After the treatment, the sample was amplified by PCR using a KOD Multi & Epi, and the primers are shown in Table 1. The conditions for PCR were 40 cycles at 94°C for 2 min, 98°C for 10 s, 53°C for 30 s, and 68°C for 15 s, and then, the sample was held at 4°C. After gel purification, alignment analysis was performed.

**HPLC.** The modified positions of cytosine by M.ApeKI were determined by HPLC. For the conditions for HPLC, we referred to a previous report (33). This experiment used 0.1% formic acid in Milli-Q and 0.1% formic acid in methanol. The HPLC program was gradient elution. The 0.1% formic acid in methanol was increased to 0.15 ml/min from 0 to 22.5%, and 20  $\mu$ l of the sample was injected.

dC, dA, T, m5dC, and m4dC were dissolved in Milli-Q water, and dG was dissolved in methanol. After dissolution, all authentic samples were filtrated using a syringe-driven filter (0.20  $\mu$ m).

A ds-GCWGC as a substrate was produced using a thermal cycler that contained 23  $\mu$ M PCReady primer of each (forward and reverse) and annealing buffer (1 mM Tris-HCl, 100 mM NaCl, and 0.1 mM EDTA). After incubating for 5 min at 95°C, the temperature was decreased by 1°C/min to 25°C. The sequence of ds-GCWGC is shown in Table 1.

A 0.39-mmol aliquot of ds-GCWGC, 19  $\mu$ g of the purified M.ApeKI, 32 nmol of AdoMet, and 0.1 mg/ml of BSA were mixed and then reacted for 22 h at 40°C. The sample was digested from the double-stranded oligonucleotide to nucleoside overnight at 37°C using Nucleoside Digestion Mix and was then filtrated using a syringe-driven filter (0.20  $\mu$ m).

## ACKNOWLEDGMENTS

This work was supported by an MEXT-Supported Program for the Strategic Research Foundation at Private Universities 2015 to 2019 (S1511019L).

## REFERENCES

- Kumar S, Chinnusamy V, Mohapatra T. 2018. Epigenetics of modified DNA bases: 5-methylcytosine and beyond. *Front Genet* 9:640. <https://doi.org/10.3389/fgene.2018.00640>.
- Li E, Beard C, Jaenisch R. 1993. Role for DNA methylation in genomic imprinting. *Nature* 366:362–365. <https://doi.org/10.1038/366362a0>.
- Moore LD, Le T, Fan G. 2013. DNA methylation and its basic function. *Neuropsychopharmacology* 38:23–38. <https://doi.org/10.1038/npp.2012.112>.
- Jurkowska RZ, Jeltsch A (ed). 2016. Preface, p v–vi. *In* DNA methyltransferases—role and function. Springer, New York, NY.
- Bujnicki JM. 2001. Understanding the evolution of restriction–modification systems: clues from sequence and structure comparisons. *Acta Biochim Pol* 48: 935–967. [https://doi.org/10.18388/abp.2001\\_3859](https://doi.org/10.18388/abp.2001_3859).
- Cheng X, Roberts RJ. 2001. AdoMet–dependent methylation, DNA methyltransferases and base flipping. *Nucleic Acids Res* 29:3784–3795. <https://doi.org/10.1093/nar/29.18.3784>.
- Bestor TH. 2000. The DNA methyltransferases of mammals. *Hum Mol Genet* 9:2395–2402. <https://doi.org/10.1093/hmg/9.16.2395>.
- Yokochi T, Robertson KD. 2002. Preferential methylation of unmethylated DNA by mammalian de novo DNA methyltransferase Dnmt3a. *J Biol Chem* 277:11735–11745. <https://doi.org/10.1074/jbc.M106590200>.
- Vasu K, Nagaraja V. 2013. Diverse functions of restriction–modification systems in addition to cellular defense. *Microbiol Mol Biol Rev* 77:53–72. <https://doi.org/10.1128/MMBR.00044-12>.
- Greer EL, Blanco MA, Gu L, Sendinc E, Liu J, Aristizabal-Corrales D, Hsu CH, Aravind L, He C, Shi Y. 2015. DNA methylation on N<sup>6</sup>-adenine in *C. elegans*. *Cell* 161:868–878. <https://doi.org/10.1016/j.cell.2015.04.005>.
- Li Z, Zhao P, Xia Q. 2019. Epigenetic methylations on N<sup>6</sup>-adenine and N<sup>6</sup>-adenosine with the same input but different output. *Int J Mol Sci* 20:2931. <https://doi.org/10.3390/ijms20122931>.
- Deaton AM, Bird A. 2011. CpG islands and the regulation of transcription. *Genes Dev* 25:1010–1022. <https://doi.org/10.1101/gad.203751.1>.

13. Lieb M, Bhagwat AS. 1996. Very short patch repair: reducing the cost of cytosine methylation. *Mol Microbiol* 20:467–473. <https://doi.org/10.1046/j.1365-2958.1996.5291066.x>.
14. O’Brown ZK, Boulias K, Wang J, Wang SY, O’Brown NM, Hao Z, Shibuya H, Fady P-E, Shi Y, He C, Megason SG, Liu T, Greer EL. 2019. Sources of artifact in measurements of 6mA and 4mC abundance in eukaryotic genomic DNA. *BMC Genomics* 20:445. <https://doi.org/10.1186/s12864-019-5754-6>.
15. Wang RY, Kuo KC, Gehrke CW, Huang LH, Ehrlich M. 1982. Heat- and alkali-induced deamination of 5-methylcytosine and cytosine residues in DNA. *Biochim Biophys Acta* 697:371–377. [https://doi.org/10.1016/0167-4781\(82\)90101-4](https://doi.org/10.1016/0167-4781(82)90101-4).
16. Ehrlich M, Norris KF, Wang RY, Kuo KC, Gehrke CW. 1986. DNA cytosine methylation and heat-induced deamination. *Biosci Rep* 6:387–393. <https://doi.org/10.1007/BF01116426>.
17. Ehrlich M, Gama-Sosa MA, Carreira LH, Ljungdahl LG, Kuo KC, Gehrke CW. 1985. DNA methylation in thermophilic bacteria: N<sup>4</sup>-methylcytosine, 5-methylcytosine, and N<sup>6</sup>-methyladenine. *Nucleic Acids Res* 13:1399–1412. <https://doi.org/10.1093/nar/13.4.1399>.
18. Bertani G, Weigle JJ. 1953. Host controlled variation in bacterial viruses. *J Bacteriol* 65:113–121. <https://doi.org/10.1128/jb.65.2.113-121.1953>.
19. Casadesus J, Low D. 2006. Epigenetic gene regulation in the bacterial world. *Microbiol Mol Biol Rev* 70:830–856. <https://doi.org/10.1128/MMBR.00016-06>.
20. Watanabe M, Yuzawa H, Handa N, Kobayashi I. 2006. Hyperthermophilic DNA methyltransferase M.PabI from the archaeon *Pyrococcus abyssi*. *Appl Environ Microbiol* 72:5367–5375. <https://doi.org/10.1128/AEM.00433-06>.
21. Nölling J, de Vos WM. 1992. Characterization of the archaeal, plasmid-encoded type II restriction-modification system MthII from *Methanobacterium thermoformicicum* THF: homology to the bacterial NgoPII system from *Neisseria gonorrhoeae*. *J Bacteriol* 174:5719–5726. <https://doi.org/10.1128/jb.174.17.5719-5726.1992>.
22. Kawarabayasi Y, Hino Y, Horikawa H, Yamazaki S, Haikawa Y, Jin-No K, Takahashi M, Sekine M, Baba S, Ankai A, Kosugi H, Hosoyama A, Fukui S, Nagai Y, Nishijima K, Nakazawa H, Takamiya M, Masuda S, Funahashi T, Tanaka T, Kudoh Y, Yamazaki J, Kushida N, Oguchi A, Kikuchi H. 1999. Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, *Aeropyrum pernix* Kl. *DNA Res* 6:83–101. <https://doi.org/10.1093/dnares/6.2.83>.
23. Liu XP, Liu JH. 2011. Characterization of family IV UDG from *Aeropyrum pernix* and its application in hot-start PCR by family B DNA polymerase. *PLoS One* 6:e27248. <https://doi.org/10.1371/journal.pone.0027248>.
24. Roberts RJ, Belfort M, Bestor T, Bhagwat AS, Bickle TA, Bitinaite J, Blumenthal RM, Degtyarev SK, Dryden DT, Dybvig K, Firman K, Gromova ES, Gumpport RI, Halford SE, Hattman S, Heitman J, Hornby DP, Janulaitis A, Jeltsch A, Josephsen J, Kiss A, Klaenhammer TR, Kobayashi I, Kong H, Kruger DH, Lacks S, Marinus MG, Miyahara M, Morgan RD, Murray NE, Nagaraja V, Piekarowicz A, Pingoud A, Raleigh E, Rao DN, Reich N, Repin VE, Selker EU, Shaw PC, Stein DC, Stoddard BL, Szybalski W, Trautner TA, Van Etten JL, Vitor JM, Wilson GG, Xu SY. 2003. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res* 31:1805–1812. <https://doi.org/10.1093/nar/gkg274>.
25. Roberts RJ, Vincze T, Posfai J, Macelis D. 2015. REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res* 43:D298–D299. <https://doi.org/10.1093/nar/gku1046>.
26. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* 47:W636–W641. <https://doi.org/10.1093/nar/gkz268>.
27. Renbaum P, Abrahamove D, Fainsod A, Wilson GG, Rottem S, Razin A. 1990. Cloning, characterization, and expression in *Escherichia coli* of the gene coding for the CpG DNA methylase from *Spiroplasma* sp. strain MQ1(M.Sssl). *Nucleic Acids Res* 18:1145–1152. <https://doi.org/10.1093/nar/18.5.1145>.
28. Gowher H, Jeltsch A. 2004. Mechanism of inhibition of DNA methyltransferases by cytidine analogs in cancer therapy. *Cancer Biol Ther* 3:1062–1068. <https://doi.org/10.4161/cbt.3.11.1308>.
29. Posfai J, Bhagwat AS, Posfai G, Roberts RJ. 1989. Predictive motifs derived from cytosine methyltransferases. *Nucleic Acids Res* 17:2421–2435. <https://doi.org/10.1093/nar/17.7.2421>.
30. Xu F, Mao C, Ding Y, Rui C, Wu L, Shi A, Zhang H, Zhang L, Xu Z. 2010. Molecular and enzymatic profiles of mammalian DNA methyltransferases: structures and targets for drugs. *Curr Med Chem* 17:4052–4071. <https://doi.org/10.2174/092986710793205372>.
31. Sugimoto N, Nakano S, Yoneyama M, Honda K. 1996. Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucleic Acids Res* 24:4501–4505. <https://doi.org/10.1093/nar/24.22.4501>.
32. Parks LW, Schlenk F. 1958. The stability and hydrolysis of S-adenosylmethionine; isolation of S-ribosylmethionine. *J Biol Chem* 230:295–305. [https://doi.org/10.1016/S0021-9258\(18\)70564-4](https://doi.org/10.1016/S0021-9258(18)70564-4).
33. Song L, James SR, Kazim L, Karpf AR. 2005. Specific method for the determination of genomic DNA methylation by liquid chromatography-electrospray ionization tandem mass spectrometry. *Anal Chem* 77:504–510. <https://doi.org/10.1021/ac0489420>.
34. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>.
35. Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791. <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x>.
36. Nei M, Kumar S. 2000. *Molecular evolution and phylogenetics*. Oxford University Press, New York, NY.
37. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 35:1547–1549. <https://doi.org/10.1093/molbev/msy096>.