

From sequence to activity: the Hgal-homologous restriction modification system RM.MhoVI of *Mycoplasma hominis*

Lars Vogelgsang, Manuel Dolgopelow-Schmidt, Azlan Nisar, Dana Bäcker, Alexander T. Dilthey & Birgit Henrich

Article - Version of Record



Suggested Citation:

Vogelsang, L., Dolgopelow-Schmidt, M., Nisar, A., Bäcker, D., Dilthey, A., & Henrich, B. (2025). From sequence to activity: the Hgal-homologous restriction modification system RM.MhoVI of *Mycoplasma hominis*. BMC Microbiology, 25, Article 543. <https://doi.org/10.1186/s12866-025-04270-3>

Wissen, wo das Wissen ist.

 UNIVERSITÄTS- UND
LANDESBIBLIOTHEK
DÜSSELDORF

This version is available at:

URN: <https://nbn-resolving.org/urn:nbn:de:hbz:061-20251103-113217-0>

Terms of Use:

This work is licensed under the Creative Commons Attribution 4.0 International License.

For more information see: <https://creativecommons.org/licenses/by/4.0>

RESEARCH

Open Access



From sequence to activity: the *Hgal*-homologous restriction modification system RM.*MhoVI* of *Mycoplasma hominis*

Lars Vogelgsang^{1*†}, Manuel Dolgopelow-Schmidt^{1†}, Azlan Nisar¹, Dana Bäcker¹, Alexander T. Dilthey¹ and Birgit Henrich¹

Abstract

Restriction-modification (RM) systems are widespread defense mechanisms in prokaryotes that protect the host from potentially harmful foreign DNA. They typically consist of a DNA methyltransferase (MTase), which methylates the host genome at an adenine (6 mA methylation) or cytosine (4mC or 5mC), and a restriction endonuclease (REase), which cleaves foreign, unmethylated DNA. In addition to the 2023 published family of 5mC-MTases, an *Hgal*-homolog RM system was detected in *Mycoplasma hominis* with the more rare constellation of two 5mC MTase genes, called RM.*MhoVI*. A qPCR screening of 239 randomly selected *M. hominis* isolates revealed a prevalence of the *MhoVI*-RM system of 12.97% ($n=31/239$). Notably, in all tested *MhoVI*-positive isolates, the *MhoVI*-RM cassette localized between MHO_3110 and MHO_3120 and comprised an XRE-family transcriptional regulator gene in addition to the RM genes. Intra-species conservation of the encoded *MhoVI*-enzymes was high (> 99% identities), and inter-species conservation was the lowest compared to the eponymous species *Haemophilus gallinarum* (46.6% M1.*MhoVI*; 48.1% M2.*MhoVI*; 27.4% R.*MhoVI*). A polycistronic organization of the *MhoVI*-genes was strongly suspected due to the discovery of gene-overlapping mRNA regions. The MTases activity was demonstrated in RM.*MhoVI* positive *M. hominis* isolates by protection of genomic DNA from cleavage by the methylation-sensitive endonuclease *Hgal*; and bioinformatics analysis using the Dorado basecaller on the Oxford Nanopore sequenced genomes revealed methylation rates of the respective motifs, 5'-GA^mCGC-3'/5'-G^mCGTC-3', above 95% in *MhoVI*-positives, with a higher methylation frequency of 5'-GA^mCGC-3' than 5'-G^mCGTC-3' in most isolates. A final proof of *MhoVI*-RM representing an *Hgal*-RM-like methylation activity was demonstrated through expression and analysis of recombinant rM2.*MhoVI* in *E. coli*.

Keywords *hgal*, *Mycoplasma hominis*, RM system, methylation-sensitive restriction, Modified bases, 5mC, Dorado, Nanopore sequencing

[†]Lars Vogelgsang and Manuel Dolgopelow-Schmidt shared first authorship.

*Correspondence:

Lars Vogelgsang
lars.vogelgsang@hhu.de

¹Institute of Med. Microbiology and Hospital Hygiene, Heinrich-Heine-University Duesseldorf, Duesseldorf, Germany



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Background

M. hominis is a facultative pathogen of the human urogenital tract associated with local infections, such as pelvic inflammatory disease or bacterial vaginosis, and disseminated infections, like bacterial arthritis or pre-term birth [1, 2]. The factors which render this cell wall-less mollicute from urogenital commensal to pathogen are not fully understood, but are hypothesized to result from its genetic heterogeneity, which derives from a strain-specific spectrum of variable lipoproteins, mobile genetic elements (MGE) and restriction modification (RM) systems [3].

In prokaryotes, RM systems function as the main defense mechanism against the invasion of foreign DNA by methylating specific sequence motifs of their own DNA by the DNA methyltransferase (MTase) and cutting the unmethylated motifs of the foreign DNA by the restriction endonuclease (REase) [4]. The acquisition of RM systems is mostly derived from horizontal gene transfer (HGT) as part of mobile genetic elements [5, 6]. Once integrated, they act as “selfish genetic elements” by increasing their prevalence in bacterial populations by exerting a form of genetic control through post segregational killing [7]. The loss of the MTase gene leads to cleavage of the unprotected host DNA by residual REase activity, thereby creating selective pressure to retain the RM system. After RM uptake, a controlled RM gene expression is essential to ensure that REase activity does not start earlier than methylation and thus lead to damage of the host DNA [8]. The transcriptional regulation of Type II RM systems, the most abundant subgroup of RM systems, is highly diverse. It ranges from polycistronic organization, where a single mRNA encodes all components (e.g., RM.*SalI* in *Streptomyces albus*), to monocistronic mRNAs regulating each gene individually (e.g., RM.*EcoRII* in *E. coli*) [9, 10]. Complex RM systems, such as RM.*Eco29kl* involve fine-tuned regulation by multiple intrinsic promoters, while in others, like RM.*AhdI* in *Aeromonas hydrophila*, the presence of small additional genes that encode transcriptional regulators has been described [11–14]. These regulators often belong to the XRE (xenobiotic response element) family, which is characterized by a DNA-binding helix-turn-helix (HTH) domain. The HTH domain is involved in promoter binding and regulation of respective gene expression and is also found in RM systems like RM.*FoKI* [15–17]. The controller (or C) proteins are the best-studied members of this family. As documented in REBASE, their genes are mostly positioned upstream of REase genes enabling the control of REase transcription, whereas the MTase genes are positioned on the complementary strand (see supplementary Tab. S2) [18–23].

RM systems are widely distributed among bacteria, with a vast diversity of types and specificities. Each RM

system targets a specific motif. The motifs of type II RM systems are typically 4–6 bp sequences and mainly but, not always palindromic [24–26]. The MTase adds a methyl-group to either adenine (6 mA-methylation) or cytosine (5mC- or 4mC-methylation) of the motif, thereby protecting the motif against endonucleolytic digest by the REase [27]. The identification of the RM recognition motif is the basis of further investigation of the enzymatic activity of MTase and REase and their possible role in the organism.

Methylated DNA bases were detected in the early 1990s by sequencing bisulfite-treated DNA [28]. Established methods, like methylation-sensitive-restriction (MSR), enable identification of methylated motifs to a certain degree [29, 30]. However, these methods fail to quantify methylation rates. By now, more flexible and universally applicable bioinformatical methods have been developed and allow identification of 4mC, 5mC and 6 mA modified bases with single-molecule real-time sequencing (SMRT) by Pacific Biosciences or Oxford Nanopore Sequencing Technology (ONT) [31, 32]. SMRT has long been considered the gold standard due to its relatively low error rate. However, recent advances in Oxford Nanopore Technology (ONT), including the introduction of R10 flow cells and updated chemistry, have significantly improved its accuracy [31].

In the REBASE database (<http://rebase.neb.com>), which is an extensive collection of information about restriction enzymes and DNA methyltransferases, three type II restriction enzymes have been deposited for *M. hominis*: the *AsuI*-homolog *MhoI*, the *HhaI*-homolog *Mho2965* and the *AluI* homolog *Mho2111I*. From our group, four additional type II RM systems of *M. hominis* have recently been described with homologies to RM.*DpnII* (named RM.*MhoII*; recognizing the motif 5'-GATC-3'), RM.*Sau3AI* (RM.*MhoIII*; 5'-GATC-3'), RM.*HhaI* (RM.*MhoIV*, which corresponds to *Mho2965*; 5'-GCGC-3') and RM.*Eco47II* (RM.*MhoV*; 5'-GGNCC-3'); with the methylated nucleotides underlined [30]. In this study, RM.*MhoVI* is characterized as a homolog of the *HgaI* RM system of *H. gallinarum*, which harbors two MTases - M1 and M2 - that methylate the complementary motif G^mCGTC and GA^mCGC, respectively [26].

Methods

Mycoplasma culturing and nucleic acid preparations

Clinical *M. hominis* strains were cultivated in arginine-medium and nucleic acids were prepared from *M. hominis* cultures in mid- to late logarithmic growth phase as described in detail previously [33]. Proteinase K digested cell lysate from 1 mL culture was used for qPCR analysis. Genomic DNA, isolated by the DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany) [34], was used for methylation-sensitive restriction (MSR-) analysis and nanopore

Table 1 Primers used

primer	sequence (5'–3')	position
hitA_F	TTGAGGCACAGCAATAGC	439 269–439 286 ¹
hitA_R	AAGGCTTAGGTAAGGAATGATTAG	439 204–439 227 ¹
gap_F	GCAGGCTCAATATTGACTCACT	660 140–660 162 ¹
gap_R	GATGATTCATTGTCGTATCATGC	660 212–660 234 ¹
lgt_F	TGAAATTGATTACGTCCAGGAA	420 553–420 574 ¹
lgt_R	CCGAAACGAATTATCCATAATAAAC	420 507–420 532 ¹
upXRE_F1	CGACATATTTTATTAAGACTTAAACTTATT	497–528 ²
upXRE_F2	GATATGCCATGGTATTATCAGTTAT	571–596 ²
XRE_F	GAGAATTGACCGATTAGGAACR	649–671 ²
XRE_R	CTCCACCACATCTTCTATTTTACAATTT	743–770 ²
M1_R1	CTATTAATAATATAATTAATCATTATCTCA	1224– 1251 ²
M1_F	GAGTTCACACTGCACAACATATCTTAGTT	1451– 1479 ²
M1_R2	TTTTTGGATAGAAGACAGGATTTTC	1513– 1537 ²
M1_F2	ATGAAATACTGGTGAAAGACTAAAAG	1538– 1564 ²
M2_R1	TCTTCAATAATATCCATTACATAAAAAAC	2156– 2184 ²
M2_F	GGATGGATCTCCAATTAAGGTTTT	2596– 2620 ²
M2_R2	CCACCCATATTGCCATTATTCA	2673– 2694 ²
R_F	TTCGCTTTAGTCAAATGTGCTTTT	3113– 3136 ²
R_R	AGGATGTGGATATTGCAAAGCAA	3245– 3267 ²
downR_F	GTATAGTTGTTGCTCCACGTTTCTCAC	4172– 4198 ²
downR_R1	GCTGTATTGATTATTGTTTTTTCAT	4448– 4473 ²
downR_R2	ACGACCTAACTCATAATTGAGCCAAG	4483– 4508 ²

¹Position in *M. hominis* PX1114 (acc.-no. CP032849.1)²Position relative to the *mhoVI*-gene cassette in PX1114 (pos. 391 805–386 982; see Fig. 1)

sequencing. RNA was prepared with the RNeasy Kit (Qiagen GmbH, Hilden, Germany) and cDNA was synthesized as published [35].

Concentration of isolated nucleic acids was measured by Qubit 4 Fluorometer (Invitrogen, Carlsbad, USA) and the quality was verified by NanoDrop 1000 Spectrophotometer (Thermo Fisher, Waltham, USA).

Table 2 PCR-products

PCR	forward primer	reverse primer	length [bp]
XRE	XRE_F	XRE_R	122
M1	M1_F	M1_R2	87
M2	M2_F	M2_R2	99
R	R_F	R_R	155
A	upXRE_F2	M1_R1	681
B	M1_F2	M2_R1	647
C	M2_F	R_R	672
PI	upXRE_F2	XRE_R	200
PII	upXRE_F1	XRE_R	274
TI	downR_F	downR_R1	302
TII	downR_F	downR_R2	337

qPCR

Oligonucleotides used in PCRs were designed using Profinder (Roche Applied Science) (<https://qpcr.probefinder.com>) or PrimerSelect of DNASTAR (Madison, WI, USA). Primer sequences are listed in Table 1.

The qPCR assays were carried out in a total volume of 25 µl consisting of 1 × MesaGreen MasterMix, 5 mM MgCl₂, Amperase, 300 nM of each primer and 2.5 µl of genomic DNA (1 ng/µl) or cDNA (0.8 ng/µl) solution, which was derived from 100 ng RNA [35]. Thermal cycling conditions were as follows: 1 cycle at 50 °C for 10 min, 1 cycle at 95 °C for 5 min followed by 35 cycles of 95 °C for 15 s and 60 °C for 1 min. PCR products are listed in Table 2.

PCR products were differentiated from primer dimers with melting curve analysis by subsequent heating from 65 °C to 95 °C with an increment of 0.5 °C/15 s. Cycling, fluorescent data collection and analysis were carried out on a CFX-Cycler of BioRad Laboratories (Munich, Germany) according to the manufacturer's instructions. Ct values of genomic DNA amplification were interpreted relative to the chromosomal *M. hominis*-specific *hitA* gene [36]. Transcript levels were calculated relative to the mean of two mycoplasma (housekeeping) genes, the glyceraldehyde-3-phosphate dehydrogenase (*gap*, also known as GAPDH) and the prolipoprotein diacylglycerol transferase (*lgt*), which had formerly been shown to be unregulated [3, 35]. Ct-values > 27 were interpreted as negative. Larger regions of the *MhoVI*-gene cassette (> 500 bp) were amplified in 3-step qPCR with 35 cycles of 95 °C for 30 s, 60 °C for 30 s and 72 °C for 2 min.

Cloning and expression of recombinant methyltransferases rM1 and rM2

The protein encoding regions of M1.*MhoVI* and M2.*MhoVI* of *Mycoplasma hominis* strain 16753VA were amplified by PCR using oligonucleotides (Metabion, Planegg, Germany) that change the mycoplasma TGA tryptophan codon to TGG (rM2) and add restriction sites (*Bam*HI and *Hin*DIII) up- and downstream of the open

reading frames, respectively. rM2 fragments were fused by SOE (splicing by overlap extension)-PCR [37]. Amplicons were ligated in-frame into the expression vector pQE30 (Qiagen, Hilden, Germany) and plasmids were propagated in *E. coli* DH5 α F'Iq. Sequence integrity of the inserts was analyzed by Sanger sequencing.

LB-medium (10 g/L bacto-tryptone, 5 g/L bacto-yeast extract, 5 g/L sodium chloride) containing 50 mg/L ampicillin and 0.02 mM IPTG (isopropylthio-P-D-galactoside) was inoculated with 1 mL of an overnight culture of the appropriate *E. coli* clone and incubated at 37 °C with vigorous shaking until an OD₆₀₀ of 1.8 was reached. The cells were harvested by centrifugation (4500 $\times$ g, 20 min, 4 °C) and the cell sediments stored at -20 °C until use.

Methylation sensitive restriction (MSR) analysis

Genomic DNA (0.75 μ g) was digested with 1 U restriction enzyme *Hga*I (GACGCN_{5/10}), *Dpn*I (G^mA/TC) and *Mbo*I (GATC/N) in 15 μ L for 1 h at 37 °C, followed by heat inactivation of the enzyme at 80 °C for 20 min and subsequent analysis on 0.8% agarose gels.

Sanger-sequencing

For Sanger-sequencing, PCR-products were purified with “NucleoSpin Gel and PCR Clean-up” Kit from Macherey-Nagel (Dueren, Germany) and sequenced at the “Genomics & Transcriptomics Laboratory” of the Biological Medical Research Centre of the Heinrich-Heine University of Duesseldorf.

Genome sequencing

Whole genome sequencing of clinical *M. hominis* strains and genetically modified *E. coli* was performed using Oxford Nanopore technology KIT (SQK-NBD114.24) according to the manufacturer's instructions. The DNA libraries, generated by following the barcode ligation protocol (SQK-LSK109), were loaded on PromethION R10.4.1 flow cells and sequenced on the PromethION PC24 with the MinKNOW 23.11.7 software. Assembly of the nanopore long reads was conducted via Canu 2.1 (<https://www.github.com/marbl/canu>). Normalization of the genome sequences was done with respect to gene order and position in type strain PG21 (Acc.-Nr. FP236530.1) using Geneious Pro 5.5.8. Genome sequences are deposited at <https://doi.org/10.17605/OSF.IO/3N4SY>.

Software based analyses

Geneious Pro 5.5.8 was used for multiple sequence alignments, generation of consensus sequences of overlapping PCR products and identification of open reading frames. BLASTP (<https://www.blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>) was used for protein prediction and analysis. Prokaryotic promoters were calculated by ProPr 2.0 and Sapphire [38, 39]; transcription

terminators by ARNold [40]. Multiple sequence alignments were calculated using Geneious Pro vers. 5.5.8 (Dotmatics, Boston, Madison, WI, USA). MegAlign 5.08 of the Lasergene software package (DNASTar, Madison, WI, USA) was used with default settings for phylogenetic tree construction.

Calculation of methylation scores

FAST5 files of the Oxford Nanopore-derived reads that contain the methylation data underwent processing with the pod5 package (ver. 0.3.10) within the Miniconda3 24.3.0 environment to convert them to POD5 files. Methylation scores for 5-methylcytosine (5mC) were computed via the basecaller Dorado (ver. 0.5.1; <https://github.com/nanoporetech/dorado>), employing the basecalling model dna_r10.4.1_e8.2_400bps_sup@v4.2.0_5mC@v2 and using the respective Canu-assembled *M. hominis* genomes as reference. Conversion of BAM output files to BED files was done using the modbam2bed tool (ver. 0.10.0) to calculate the methylation frequencies of each cytosine. Positions of the internal cytosines in the *Mho*VI motifs (5'-GACGC-3'/5'-GCGTC-3') were identified in genomes through JupyterLite (vers. 0.3.0; <https://github.com/jupyter/notebook>). The motif positions were then combined with the methylation frequencies to calculate the motif-specific methylation scores.

Results

By analyzing strain-specific regions in Oxford Nanopore-derived *M. hominis* genome sequences, a hitherto undiscovered RM system was detected in *M. hominis* strain 16753VA. It was homologous to the *Hga*I-RM system of *Haemophilus gallinarum* (synonym *Avibacterium volantium*) [41]. BLASTN analysis identified the presence of this RM system in the published genomes of *M. hominis* strains PX1114 (acc.-no. CP032849.1) and MIN-132 (CP086131.1). As schematically shown in Fig. 1, the 4.8 kb *mho*VI cassette of *M. hominis* was positioned in the genomic region between the core genes MHO_3110 and MHO_3120 with respect to type strain PG21 and was composed of the three *hga*I-homologous genes, encoding two methyltransferases (M1.*Mho*VI and M2.*Mho*VI), a restriction endonuclease (R.*Mho*VI) and a 70-amino-acid (AA) XRE family transcriptional regulator gene, positioned upstream of the M1.*Mho*VI gene (Fig. 1).

Multiple sequence alignment of the intergenic regions between MHO_3110 and MHO_3120 in PG21 and PX1114 revealed integration of the *mho*VI-cassette 54 bp upstream of the MHO_3120 3'-end (at position 369,286 of PG21) with duplication of the genomic PG21 region [369,084 to 369,286] upstream [386,670–386,982] and downstream [391,808–392,002] of the *Mho*VI-cassette of PX1114 (see Fig. 1, red line).

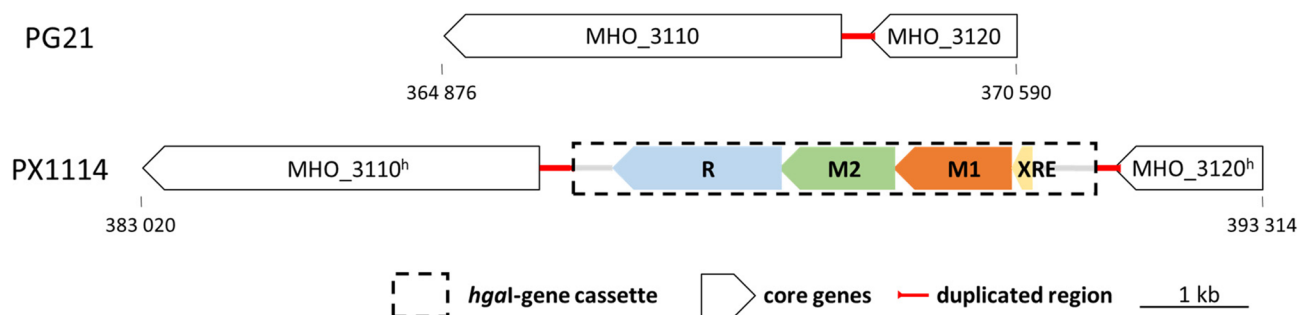


Fig. 1 Scheme of the genomic locus of the *mhoVI*-RM cassette. The *mhoVI*-RM gene cassette (dashed square), schematically shown for isolate PX1114 (acc.-no. CP032849.1), encodes an XRE-transcriptional regulator (yellow), two MTases (M1.*MhoVI* (orange) & M2.*MhoVI* (green)) and a REase (R.*MhoVI* (blue)) and is positioned at an intraspecies-conserved integration site between the MHO_3110- and MHO_3120-homolog (^h) core genes (with respect to the PG21 genome; white; acc.-no. FP236530.1) in reverse complementary direction. Integration caused a 202 bp duplication (red line). Arrow tips mark the 3'-end of the genes

Conservation of the *mhoVI*-RM cassette is high in *M. hominis*

Based on sequence information of the *mhoVI*-cassette in *M. hominis* strains 16753VA, PX1114 and MIN-132, qPCR primer pairs were designed for each of the *mhoVI*-genes (see Table 1) and subjected to 239 clinical isolates. 12.97% (31/239 isolates) tested positive for each of the four genes (see Supplementary Tab. S4), which is moderate compared to the known 5mC RM systems of *M. hominis* ranging from 2.39% (*Eco47II*-homolog RM.*MhoV*) to 27.2% (*HhaI*-homolog RM.*MhoIV*) [30]. Oxford Nanopore sequencing of the genomes of eight *mhoVI*-positive isolates was performed and confirmed the conserved architecture and genomic position of the *mhoVI*-genes cassette as shown for strain PX1114 (Fig. 1). To decipher error-free gene sequences of the *mhoVI*-cassette region of nine isolates, overlapping PCR products were sequenced according to Sanger [42]. Intra-species conservation of the total *mhoVI*-gene cassette was high (99.89% nucleotide identity) and homology of the deduced protein sequences was highest for the MTases (99.93% for M1.*MhoVI* and 99.96% for M2.*MhoVI*), followed by 99.83% for R.*MhoVI* and 99.65% for XRE (see Table 3, A and B).

XRE, the smallest protein consisting of 70 amino acids (aa) with a calculated molecular mass of 7.88 kDa, exhibited the highest intra-species conservation with no amino acid substitutions. The highest divergence of the RM genes was observed for the restriction endonuclease R.*MhoVI*, the largest protein with 520 aa and a molecular mass of 60.3 kDa, with a total of 5/520 aa changes. M1.*MhoVI* and M2.*MhoVI* were of similar size with 362 and 358 aa, respectively, and a calculated molecular mass of 41.3 kDa for both. Homology of *M. hominis* M1.*MhoVI* and M2.*MhoVI* was 49% throughout all tested *M. hominis* genomes (Supplementary. Tab. S1). Phylogenetic analysis of the amino acid sequence of RM.*MhoVI* MTases and RM.*HgaI* MTases, including

M1 and M2 homologs of *HgaI* in *M. lactucae*, *H. pylori*, *L. petauri*, that comprised the region from the first conserved domain (Motif I) to the last (Motif X), revealed a monophyletic M2 branch (see Supplementary Fig. S3A). In contrast, M1-MTases clustered paraphyletically with M1.*MhoVI*, showing a similar homology to M1.*HgaI* and M2.*HgaI*. Comparable results were obtained when analyzing the region between motif VIII and motif X, which includes the target recognition domain (TRD) responsible for DNA-binding and motif specificity (Fig. S3B) [43]. Due to the perfect clustering of M2-MTases in a common branch, it was assumed that M2.*MhoVI* methylates motif GA^mCGC and, even though M1.*HgaI*-homologs seem to be more heterogeneous, M1.*MhoVI* was assigned to the complementary sequence motif G^mCGTC.

XRE is only in *M. hominis* part of the *HgaI*-homologous RM cassette

Homologs of M1.*MhoVI*, M2.*MhoVI*, and R.*MhoVI* were also detected in other bacteria: *Mesoplasma lactucae* (acc.-no GCF_002441935.1), a mollicute isolated from the plant *Lactuca sativa*; the human pathogen *Helicobacter pylori* (GCF_002831845.1); the marsupial and fish-colonizing *Lactococcus petauri* (GCF_023499275.1), and the avian *H. gallinarum* which gives the *HgaI*-RM system its name. Thus, neither an evolutionary spread of the *MhoVI*-RM element in a common phylogenetic branch nor a common anatomic site of colonization is obvious. The analysis of inter-species conservation revealed varying degrees of similarity across the different enzymes (Table 4).

The amino acid conservation of the homologs ranged from 46.5 to 68.5% to M1.*MhoVI*, 48.1–71% to M2.*MhoVI*, and 27.4–49.5% to R.*MhoVI*, demonstrating that the REases were more divergent than MTases. The MTases of *Mycoplasma hominis* strain 16753VA, used as reference, exhibited the highest homology to *Mesoplasma lactucae* (M1.*MhoVI*: 68.5%;

Table 3 Intra-species homology (%) of the *MhoVI*-RM enzymes

A		M1.MhoVI										
M2.MhoVI	isolates	16753VA	2740	727J	1014VA	1021J	10428	1627	6148	9840		M1.MhoVI
	16753VA		99.7	100	100	100	100	100	100	100	16753VA	
	2740	100		99.7	99.7	99.7	99.7	99.7	99.7	99.7	2740	
	727J	100	100		100	100	100	100	100	100	727J	
	1014VA	99.7	99.7	99.7		100	100	100	100	100	1014VA	
	1021J	100	100	100	99.7		100	100	100	100	1021J	
	10428	100	100	100	99.7	100		100	100	100	10428	
	1627	100	100	100	99.7	100	100		100	100	1627	
	6148	100	100	100	99.7	100	100	100		100	6148	
	9840	100	100	100	99.7	100	100	100	100		9840	
	16753VA	2740	727J	1014VA	1021J	10428	1627	6148	9840	isolates		
M2.MhoVI												

B		R.MhoVI										
XRE	isolates	16753VA	2740	727J	1014VA	1021J	10428	1627	6148	9840		R.MhoVI
	16753VA		99.8	99.4	100	100	100	100	99.6	99.8	16753VA	
	2740	100		99.2	99.8	99.8	99.8	99.8	99.4	99.6	2740	
	727J	100	100		99.4	99.4	99.4	99.4	99.0	99.2	727J	
	1014VA	100	100	100		100	100	100	99.6	99.8	1014VA	
	1021J	100	100	100	100		100	100	99.6	99.8	1021J	
	10428	100	100	100	100	100		100	99.6	99.8	10428	
	1627	100	100	100	100	100	100		99.6	99.8	1627	
	6148	100	100	100	100	100	100	100		99.4	6148	
	9840	100	100	100	100	100	100	100	100		9840	
	16753VA	2740	727J	1014VA	1021J	10428	1627	6148	9840	isolates		
XRE												

Table 4 Inter-species homology (%) of the RM.*MhoVI* enzymes

RM. <i>MhoVI</i> -homolog in organism	RM. <i>MhoVI</i> of <i>Mycoplasma hominis</i> PX1114		
	M1	M2	R
<i>Mesoplasma lactucae</i>	68.5	71.0	49.5
<i>Helicobacter pylori</i>	58.0	66.7	52.0
<i>Lactococcus petauri</i>	46.5	50.9	26.8
<i>Haemophilus gallinarum</i>	46.6	48.1	27.4

M2.*MhoVI*: 71%), followed by *Helicobacter pylori* (M1.*MhoVI*: 58%; M2.*MhoVI*: 66.7%), *Lactococcus petauri* (M1.*MhoVI*: 46.5%; M2.*MhoVI*: 50.9%), and *H. gallinarum* (M1.*MhoVI*: 46.6%; M2.*MhoVI*: 48.1%). Notably, inter-species conservation of M2.*MhoVI* was on average 4.3% higher than of M1.*MhoVI*. Interestingly, XRE.*HgaI* sequence homologs were not found in RM.*HgaI* homologs of other procaryotes, but were associated with other putative RM systems. Analysis of XRE.*MhoVI* homologs revealed their affiliation with members of the YozG subfamily within the XRE family. Phylogenetic reconstruction further demonstrated that C proteins are closely related to XRE.*MhoVI* but form a distinct monophyletic clade *HipB*, clearly separated from the YozG-like XRE subfamily branch (Fig. 2).

The members of the two branches differ not only in sequence homology but also in their genomic organization regarding the relative position of the transcriptional regulator gene in the respective RM cassette. In the *HipB*-associated branch, the genes encoding controller proteins are located directly upstream of the REase genes, while the corresponding MTase genes are typically found on the opposite DNA strand (see Supplementary Tab. S2). In contrast, members of the YozG subfamily, such as XRE.*MhoVI*, exhibit a conserved gene arrangement, in which the transcriptional regulator is followed by the MTase gene and then the REase gene.

***MhoVI*-gene transcript levels support a polycistronic organization**

Expression of the *MhoVI*-RM system was investigated by the quantification of transcript amounts relative to the housekeeping genes *gap* and *lgt*; for each *mhoVI*-gene using primer pairs binding within each gene. Transcription of the *mhoVI*-genes was higher than that of the housekeeping genes in all RM.*MhoVI* positive isolates (Fig. 3).

The expression ranged from 0.78 ± 0.5 (M2.*mhoVI* in isolate 9840U) to 8.65 ± 0.62 (R.*mhoVI* in isolate 1627). M1.*mhoVI* and M2.*mhoVI* were the least transcribed

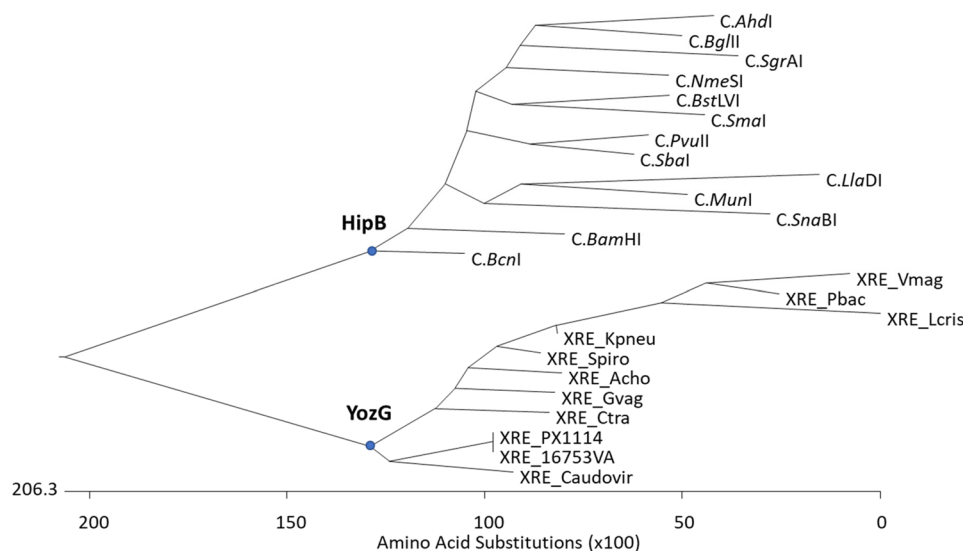


Fig. 2 Phylogenetic tree of selected XRE.MhoVI-homologs. The construction of the phylogenetic tree was performed with MegAlign 5.08; accession numbers and abbreviations of the proteins are deposited in Supplementary Table S2

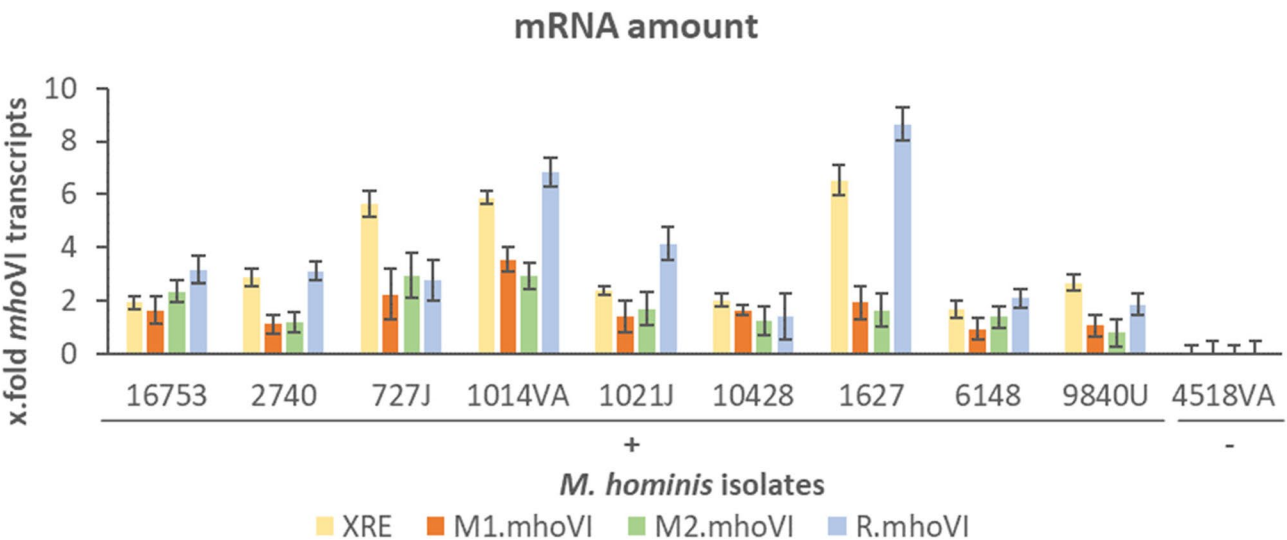


Fig. 3 Transcripts of *mhoVI*-genes in *M. hominis*. The mRNA levels of the *mhoVI*-genes were quantified relative to the mean transcript level of gap and lgt using the ΔCT method. Error bars representing standard deviations were calculated on the base of two biological and one technical replicate; each measured in duplicates. Displayed are the results from nine RM.mhoVI positive (+) and one RM.mhoVI negative isolate (-) for comparison

on average with 1.72 ± 1.61 and 1.80 ± 1.64 , respectively, while *R.mhoVI* showed the highest relative expression at 3.78 ± 1.75 . The differences in transcript levels of the *mhoVI*-genes calculated by one-way ANOVA test (p -value > 0.05) showed no overall statistical significance. However, looking at the transcription profile of the four genes in each isolate, the *M1.mhoVI* and *M2.mhoVI* transcript levels were approximately equal, and generally lower than those of *XRE* and *R.mhoVI*. As expected, no significant transcript levels of RM genes were detected in the nine RM.MhoVI negative isolates, as shown for isolate 4518VA as a representative (Fig. 3).

To test the hypothesis of a polycistronic mRNA, gene-overlapping RT-qPCRs were conducted using primers that were designed to produce PCR products that extend at least 300 nt across two adjacent genes (Fig. 4).

Gene-overspanning RT-qPCR products (A-C) were quantified using the ΔCT method. Because amplicon efficiency is affected by primer pair characteristics and the length of RT-PCR products, simultaneous amplification of genomic DNA was performed to normalize transcript levels. Amplicon (A) started 8 bp upstream of *XRE*, spanned the *XRE* gene completely and covered 474 nt of the *M1.mhoVI* gene. Amplicon (B) spanned 329 nt *M1.mhoVI* 3'-end and 323 nt *M2.mhoVI* 5'-end; and

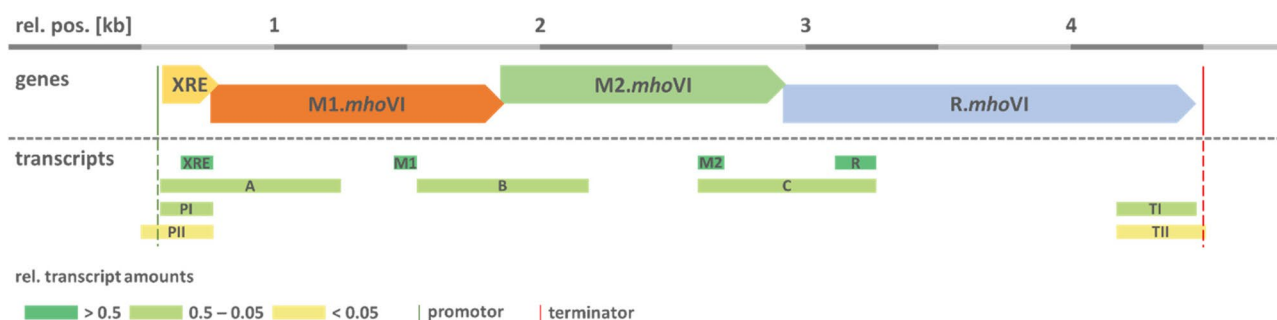


Fig. 4 RT-qPCR analysis of gene overlapping *MhoVI*-regions. Schematic representation of the 4.8 kb *MhoVI* gene locus with PCR products numbered according to Table 2. The levels of *MhoVI*-specific gene-, promoter- and terminator-overlapping transcripts were calculated relative to *gap* transcripts and rated to the nearest power of ten (<0.05, 0.05–0.5, >0.5). Bioinformatically predicted promoter (green) and terminator (red) positions are indicated by dashed lines. All positions (rel. pos.; 1 bp – 4.825 kb) correspond to the *mhoVI*-gene-cassette region from nt 391,807 – nt 386,982 in PX1114 (CP032849.1)

amplicon (C) covered 333 nt *M2.mhoVI* (3'-end) and 374 nt *R.mhoVI* (5'-end). Overlapping amplicons (A, B, C) were detected at the three gene junctions, supporting the thesis of a polycistronic *mhoVI*-cassette mRNA.

Two potential universal promoter regions were identified upstream of the XRE gene. The first was detected by Sapphire at the relative nucleotide position 531–559 with a *p*-value of 0.0002, the second by ProPr2.0 in an overlapping region at nt 496–567 with the highest possible score of 1. In RT-qPCR analysis, a 200 bp product was generated, starting upstream of the suspected promoter at position nt 571 (Fig. 4; PI), whereas a 274 bp product, starting downstream at position nt 491 (Fig. 4; PII), was barely detected (Δ CTs of 0.22 and 0.007, respectively), strongly supporting the transcription promoter site between position 497 and 571.

Using the browser tool ARNnold, a universal transcription terminator was detected downstream of *R.mhoVI* at nt 4470–4519, with a moderate Δ G-score of –7.3. qPCR analysis supported the assumption of a terminator region downstream of nt 4473, as a 302 bp PCR product (Fig. 4; TI) extending to nt 4473 was synthesized in significant amounts (Δ CT of 0.126), while a 337 bp product (Fig. 4; TII) extending to nt 4508 was synthesized in significantly lower amounts (Δ CT of 0.026; Fig. 4).

RM.*MhoVI* positive isolates always carry GA^mCGC/G^mCGTC-methylated DNA

To demonstrate the activity of the native *MhoVI*-MTases of *M. hominis*, the methylation state of the genomic DNA of *MhoVI*-RM positive and negative isolates was investigated in a methylation-sensitive restriction (MSR) analysis. The restriction enzyme *HgaI*, which restricts only unmethylated GACGC/GCGTC motifs was utilized to determine 5mC methylations within the presumed RM.*MhoVI* recognition motif (Fig. 5).

As shown in Fig. 5A, lack of DNA restriction by the endonuclease *HgaI* was observed in the *MhoVI*-RM-positive isolate 16753VA, indicating protection of the DNA

by 5mC methylation of the *MhoVI*-recognition motif GACGC/GCGTC and thus suggesting methylation activity of the native *MhoVI*-MTases. Restriction of the DNA of strain 4518VA by *HgaI* demonstrated a lack of 5mC methylation in the GACGC/GCGTC motif, which was consistent with the absence of the *MhoVI*-RM system (Fig. 5B). The restriction patterns of isolates 16753VA and 4518VA corresponded to those of the eight *MhoVI*-positive and -negative tested isolates, respectively (see supplementary Fig. S1 and supplementary Tab. S3). To confirm that the DNA is susceptible to nucleolytic digestion at all, the two restriction enzymes *MboI* and *DpnI* were used, as they restrict the GATC motif in 6 mA-methylated or -unmethylated form, respectively.

The GATC-specific restrictions seemed to fragment the DNA more extensively compared to *HgaI*. This could be attributed to a reduced number of GC-enriched motifs (such as GACGC) in the genome (Fig. 5C–D), which corresponds to the reduced GC content of mycoplasma genomes ranging from 21 to 40% [44]. As expected, the calculation of total DNA restriction produced approximately eight-fold more GATC-fragments with mean lengths of 600 bp than GACGC-fragments of meanly 4600 bp. Thus, it was hypothesized that the MTases of *MhoVI* methylated *each* recognition motif in the genome of *MhoVI*-positives. To prove this hypothesis, the Oxford Nanopore reads of the sequenced genomes of strains 16753VA and 4581VA were analyzed to calculate 5mC and 6 mA modifications across the genomes. As M1.*HgaI* of *H. gallinarum* is known to methylate G^mCGTC and M2.*HgaI* the complementary GA^mCGC motif, the methylation frequencies (MF) of both motifs were analyzed independently and that of G^mATC as a control (Fig. 6) [26].

As shown in Fig. 6A, high methylation rates (>95%) of both, 5'-G^mCGTC-3' and 5'-GA^mCGC-3', were observed at almost every motif site in the RM.*MhoVI* positive isolate 16753VA, but below background in the RM.*MhoVI* free isolate 4518VA. In contrast, 5'-G^mATC-3'

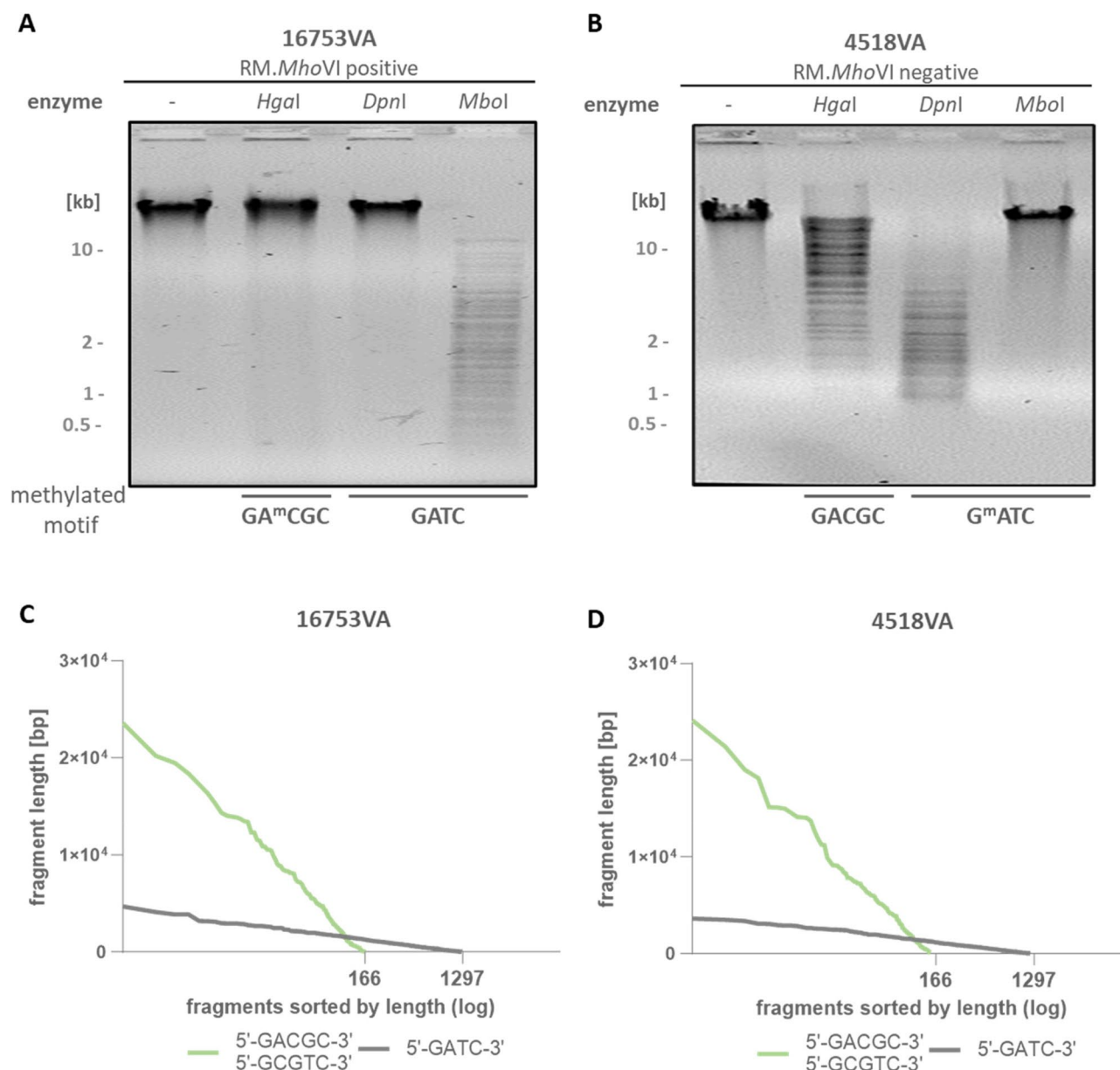


Fig. 5 Methylation-sensitive restriction analysis. DNA of a RM.MhoVI positive (**A**, 16753VA) and a negative strain (**B**, 4518VA) was restricted by methylation-sensitive restriction (MSR) enzymes *HgaI* (GACGC/GCGTC), *DpnI* (G^mATC) and *MboI* (GATC) and separated on agarose gels. ^m = methylated base (**C - D**) Simulation of total DNA restriction by *HgaI* (green line) or *DpnI*/*MboI* (grey line) of isolates 16753VA (**C**) and 4518VA (**D**). The number of fragments is plotted against fragment length

methylation was below background in isolate 16753VA, but positive in isolate 4518VA, which had recently been shown to express (in contrast to 16753 A) a G^mATC methylation by Dam1 (Fig. 6B) [30]. Interestingly, six positions were identified in isolate 16753VA with an apparent hemimethylation, as the methylation rate of GA^mCGC were lower than that of G^mCGTC-by up to 70% (Fig. 6A, green dots).

Methylation at GA^mCGC is generally higher than at G^mCGTC.

Average methylation rates of G^mACGC, G^mCGTC and G^mATC were then calculated for each *MhoVI*-positive isolate (Fig. 7A and Supplementary Tab. S3). A putative hemimethylation of the non-palindromic RM.MhoVI-motif, which was observed for six out of 166 motifs in isolate 16753VA, was not calculated for other isolates; However, analysis of eight *MhoVI*-positive isolates revealed differences in the average methylation rates of

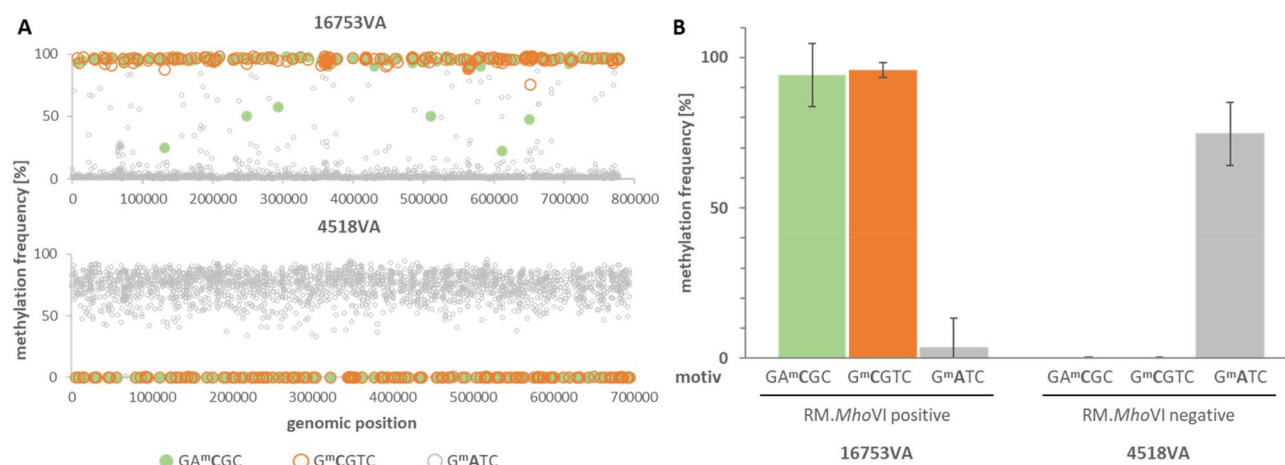


Fig. 6 Nanopore-derived methylation frequencies of GA^mCGC, G^mCGTC and G^mATC. **(A)** Dot plots show the percentage of 5mC modified bases (methylation frequency) in the RM.*Mho*VI recognition motifs GACGC (green) and GCGTC (orange) as well as 6 mA methylation frequency of GATC (gray) at their respective genomic positions in the RM.*Mho*VI-positive isolate 16753VA and the RM.*Mho*VI-negative isolate 4518VA. **(B)** Bar graphs show the average methylation frequency of all GA^mCGC (green), G^mCGTC (orange) and G^mATC (gray) which were identified in Dorado analysis of the Nanopore sequenced genomes of isolate 16753VA and 4518VA.

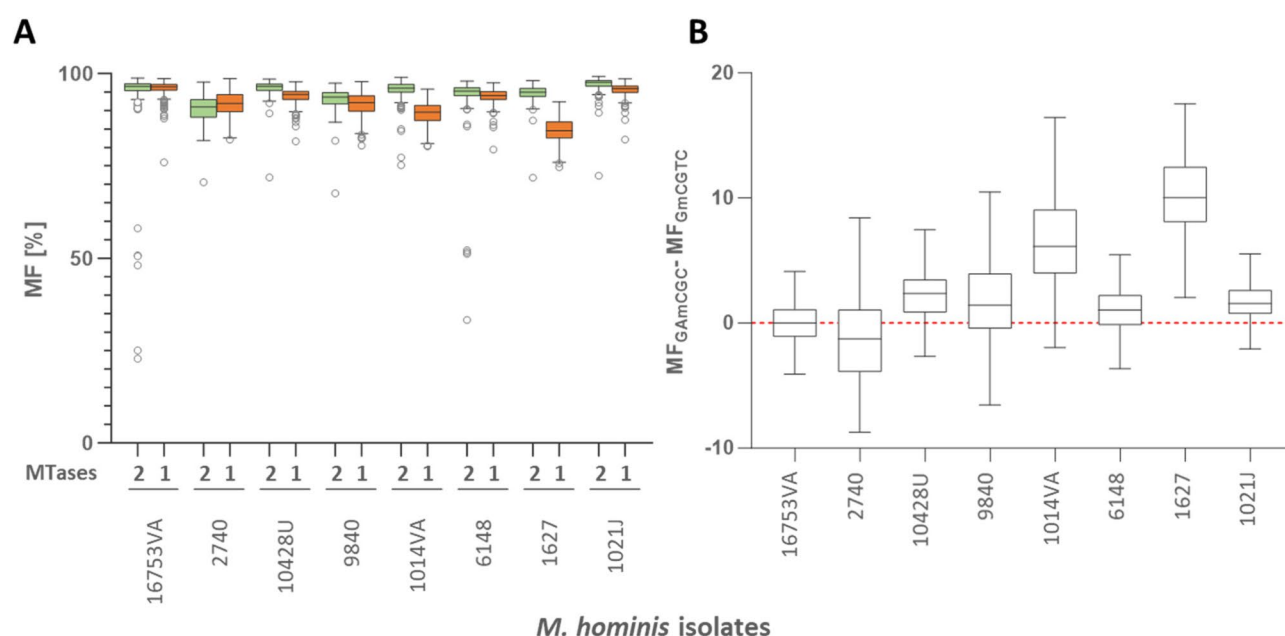


Fig. 7 Strain-specific GA^mCGC [2] and G^mCGTC [1] methylation frequencies. **(A)** Tukey box plots showing the proportion of methylated motifs (methylation frequency [% MF]) of GA^mCGC (2, green) and G^mCGTC (1, orange) for each of eight *Mho*VI-positive isolates. **(B)** Boxplots of the average MF differences between M2.*Mho*VI and M1.*Mho*VI at each motif site of the isolate's genome sequence. Wilcoxon signed-rank test calculated high significance ($p < 0.0001$) for all isolate but 16753VA ($p > 0.05$)

GA^mCGC and G^mCGTC in all but the reference isolate 16753VA (Fig. 7A).

With respect to the differences in methylation rates of GA^mCGC or G^mCGTC at each motif site, Wilcoxon signed-rank test was performed to verify statistical significance (Fig. 7B). Methylation discrepancies (MF_{GA^mCGC} - MF_{G^mCGTC}) were calculated in seven out of eight isolates were highly significant (p -values < 0.0001). A non-significant difference of 0.01 ($p > 0.05$) was only detected in

isolate 16753VA. In six of eight isolates, GA^mCGC methylation exceeded G^mCGTC methylation ranging from 1.05 in isolate 6148 to 10.03 in isolate 1627. Isolate 2740 exhibited a reverse pattern exhibiting higher G^mCGTC than GA^mCGC activity, suggesting isolate-specific factors affecting MTases' activity.

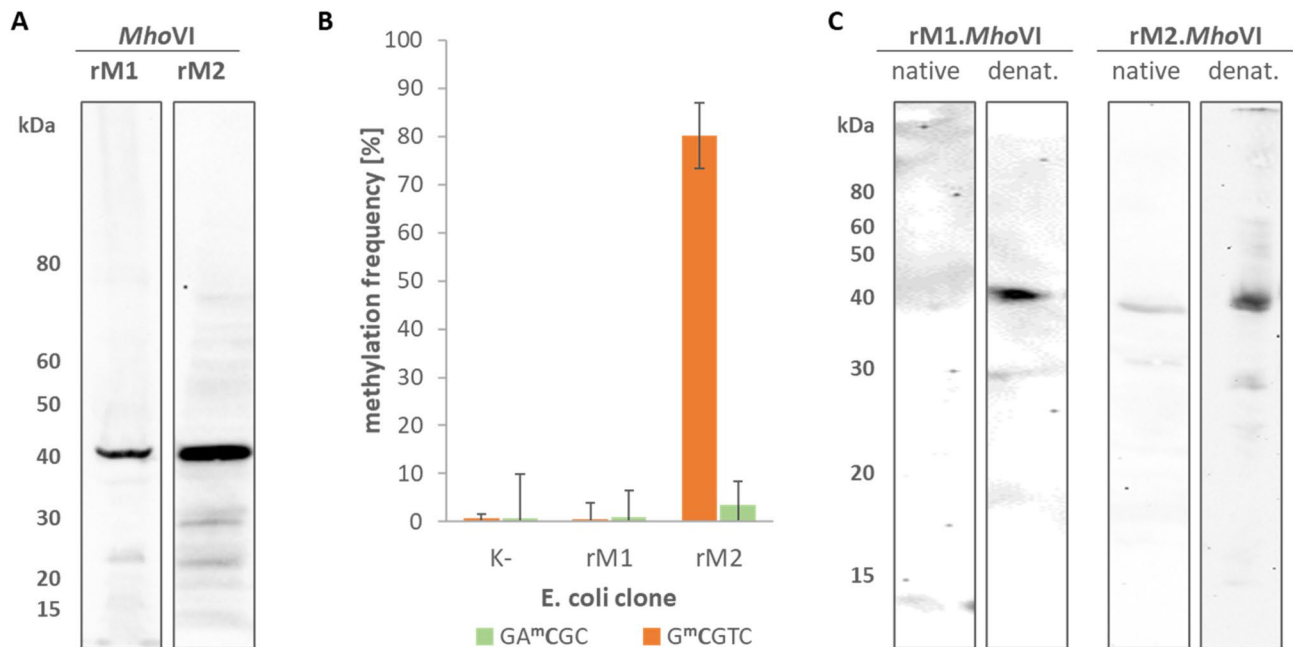


Fig. 8 Analysis of recombinant *MhoVI*-MTases. **(A)** Protein lysates of *E. coli* clones, expressing recombinant rM1.*MhoVI* and rM2.*MhoVI*, were separated on 4–12% gradient SDS-PAGE gels, blotted and visualized via immunostaining with His₆-antibody. **(B)** Mean methylation frequencies of GA^mCGC and G^mCGTC motifs in DNA of these *E. coli* clones and of an untransformed *E. coli* (K-) as a control, **(C)** *E. coli* cells expressing rM1.*MhoVI* or rM2.*MhoVI* were lysed under native or denaturing (denat.) conditions. The fraction of soluble proteins separated on 12% SDS-PAGE followed by western blotting and His₆-antibody staining

Recombinant M2.*MhoVI* exhibits G^mACGC methylation activity

To confirm the methylation activity and motif specificity of the RM.*MhoVI* MTases as *HgaI*-MTases homologs, the two MTases were cloned into expression vectors and propagated into *E. coli*. Sanger sequencing confirmed the presence and open reading frame of the inserts, but revealed mutations that led to amino acid changes. The recombinant rM1.*MhoVI* carried the one isofunctional substitution and (D304E), as well as two non-isofunctional substitution (R195G and R268G). Both proteins were heterologously expressed in *E. coli* by IPTG induction and detected in immunoblotting, with rM2 showing a higher expression level than rM1 (Fig. 8A).

Subsequent nanopore sequencing on the genomic DNA of these *E. coli* clones demonstrated significant GA^mCGC methylation of the rM2 clone DNA (MF 80.1%; Fig. 8B), and a minimal G^mCGTC methylation (MF = 3.4%). Both motifs, GACGC and GCGTC, were found to be unmethylated in the rM1 clone DNA (MF < 1%). As shown in Fig. 8C, both recombinant proteins were mainly soluble only under denaturing conditions. However, small amounts of rM2, which showed higher expression, were also detected in the soluble fraction following native lysis. This suggests that both rMTases are predominantly expressed in (methylation-inactive) inclusion bodies, with limited solubility of (presumably active) rM2 under native conditions.

Discussion

Sequence homology, conserved genomic organization of the *MhoVI*-RM system, and the detection of methylated DNA motifs in *M. hominis* consistent with those of the *HgaI* RM system in *H. gallinarum* collectively suggest that RM.*MhoVI* represents the *Mycoplasma hominis* homolog of RM.*HgaI*. At last, functional analysis of the recombinant methyltransferase rM2.*MhoVI* proved its target motif as GCGTC, which is identical to that of M2.*HgaI*, thereby confirming the proposed orthology.

In most type II RM systems, the common recognition motif of REase and MTase is palindromic and enables methylation on both DNA strands by a single MTase. However, there are examples, like the *HhaI*-homolog RM.*MhoIV*, that (like *HgaI*) harbors two 5mC MTases, but recognizes a palindromic sequence motif GCGC [30, 45]. RM systems like the *HgaI* RM system belong to the subclass of Type IIS RM systems, which is characterized by (1) recognition of a non-palindromic motif, mostly methylated by two individually acting MTases and (2) restriction of the DNA outside the recognition sequence by the endonuclease [45, 46]. As RM.*HgaI* homolog, RM.*MhoVI* of *M. hominis* was therefore assigned as Type IIS.

Analysis of heterologously expressed proteins rM1 and rM2 in *E. coli*, confirmed that the target motif of rM2.*MhoVI* is identical to that of M2.*HgaI*, GCGTC. No methylation activity of rM1.*MhoVI* could be detected

under the tested conditions. Notably, the low methylation frequency of GACGC (3.4%) observed in the rM2 clone suggests that this motif is not the primary target of rM2.MhoVI, but rather reflects a weak off-target activity, as the frequency slightly exceeded the background signal (~0.7%). Similar off-target methylation events have been reported for the GATC-specific 6 mA MTases *M.EcoK*Dam and MAGa2700 methylating GACC and AATC or YRATC and RGATC (Y = C or T; R = A or G), respectively [47, 48]. The results of this study demonstrate that rM2.MhoVI functions independently of rM1.MhoVI and suggest that M1.MhoVI is responsible for GACGC methylation in *M. hominis*, consistent to the function of M1.HgaI.

The absence of detectable rM1 activity may be unlikely attributed to the detected amino acid mutations, as secondary structure predictions did not indicate a misfolded protein (data not shown). Furthermore, the mutations were not located within known functional domains of Type II-DNA MTases (see supplementary Fig. S3A) [49]. It is more likely the result of a weak expression, poor solubility and aggregation into inclusion bodies in *E. coli*. In contrast, M2.MhoVI displayed methylation activity despite also forming inclusion bodies. This is likely due to its higher expression level leading to a significant proportion of soluble, functional protein. Contrarily to the traditional view that inclusion bodies consist solely of misfolded, inactive proteins, recent studies have shown that in some cases they can contain substantial amounts of correctly folded, active proteins [50, 51]. Moreover, the proportion of active protein within inclusion bodies can vary depending on factors such as cultivation temperature and bacterial growth phase [52, 53]. In accordance with this, but in contrast to the MhoVI-RM system, we found methylation activity of recombinant MTases of another RM system, despite their expression in inclusion bodies (manuscript in preparation).

Differences in expression levels observed for the recombinant proteins rM1.MhoVI and rM2.MhoVI in *E. coli* may reflect differential expression and activity in the native host. Since the MF of GA^mCGC was higher than the MF of G^mCGTC in six out of eight *M. hominis* isolates it can also be hypothesized that on average M2.MhoVI exhibits higher methylation activity than M1.MhoVI in their native environment. Analyses on protein level will help to elucidate the basis of MTase activity differences in future work.

Interestingly, the number of MTases in homologous RM systems can vary from species to species: in contrast to MhoV- and MhoVI-, the HhaI-RM system of *H. haemolyticus* consists of just one MTase [54], and the HgaI-RM system, NmeBI, of *N. meningitis* still harbors a single M1.HgaI-homolog MTase [55], respectively. A hemimethylation of the DNA was shown to be sufficient

to maintain defense of the RM system: the data of Sugisaki et al. showed that methylation of just GA^mCGC or G^mCGTC retains protection against the REase [26], and a change of the specificity of the HhaI-MTase from GCGC to GCG revealed full protection of the DNA by the biotechnologically designed hemimethylase [56]. However, methylation of both strands is of utmost importance in DNA replication. In the case of hemimethylation, synthesis of the strand carrying the unmethylated motif leads to a completely unmethylated recognition motif downstream of the replication fork, which is then susceptible to double-strand breaks by the REase [8]. Therefore, a regulatory mechanism for balanced MTase and REase expression is beneficial to prevent DNA degradation, not only in the case of replication.

The regulation of RM systems is complex and can occur at both the transcriptional and translational level [11–14]. XRE may be a species-specific regulator for HgaI. RM expression. Average transcript levels of XRE and R.mhoVI exceeded those of M1.mhoVI and M2.mhoVI, suggesting that additional internal promoter and terminator sites might actively influence transcription. This points to a mixed transcriptional organization, with potentially monocistronic expression of XRE.mhoVI and R.mhoVI embedded within an otherwise polycistronic RM locus (see Supplementary Table S4), a structure also described for RM.Eco29kI [12].

Alternatively, post-transcriptional processes such as RNA degradation may account for the observed differences in transcript abundance. Similar regulatory mechanisms have been proposed for RM.MhoII and demonstrated for RM.EcoP1I and RM.EcoP15I, where REase expression is modulated in response to intracellular MTase levels [30, 57].

Transcription factors like the XRE elements are widespread among prokaryotes, where they are involved in stress response and virulence by either enhancing or inhibiting the transcription of downstream genes, e.g. in *Photobacterium luminescens* and *Streptococcus* sp [58–64]. They are also involved in the regulation of toxin-antitoxin (TA) systems in *Brucella abortus* by repressing the zinc-dependent metalloproteinase (ZnMP) toxin under normal conditions [65]. As RM systems function similar to TA systems, where restriction endonucleases act as toxins and MTases as antitoxins, XRE is also known to interact with the transcription of nuclease genes [66]. In *Caulobacter*, the RM system associated proteins CCNA_00744 and CCNA_01405, exhibit an 11- and 2-fold increase in XRE-deletion strains [67]. In the *Mycoplasma gallisepticum* Mgas6I RM system, the controller protein HsdC has been shown to contain an XRE-like domain and to bind promoter regions within the RM operon to modulate the expression of RM genes [68]. This serves as an example of RM system regulation by members of the XRE family

in *Mycoplasma* species. In *M. hominis*, the role of XRE as a potential regulator of RM.*MhoVI* remains unclear, but genomic features may offer insights into its regulatory function. XRE family members of the *HipB* branch are positioned directly upstream of the REase gene to regulate its expression, while the MTase gene is typically located on the opposite strand, suggesting separate regulation [18, 19]. In contrast, *YozG* branch members, including XRE.*MhoVI*, show a gene order of XRE-MTase-REase on the same strand. Although the latter group is barely characterized, the different arrangement implies a regulatory mechanism distinct from the *HipB* group [19]. To investigate the transcriptional regulation of the XRE element, a bacterial one-hybrid system could be used to determine whether XRE interacts with promoter regions of the RM system to regulate REase or MTase expression [62].

In this study, the clinical *M. hominis* strains originated all from arginine cultures in the mid- to late-logarithmic growth phase. Although it has been described that the RM systems are active at this stage, other factors may play a more important role in RM expression that has not yet been analyzed [30]. Further research should also examine RM expression under conditions that mimic the in vivo environment, such as in polymicrobial communities, variations of temperature and pH value or exposure to foreign DNA. Such studies could provide deeper insight into the postulated XRE-based regulation and function of RM.*MhoVI* in pathophysiology and defense mechanism of *M. hominis*.

Conclusions

A combination of modern ONT whole genome sequencing, bioinformatics and experimental molecular biology was utilized to identify RM.*MhoVI* as the *M. hominis* RM.*HgaI* homolog. The genomic arrangement and MTases activity in the native host resembled those of RM.*HgaI* and functional analysis of recombinant rM2.*MhoVI* confirmed its methylation activity. The detection of a *M. hominis*-specific XRE transcriptional regulator gene within the *MhoVI*-RM cassette suggests a regulatory mechanism in *M. hominis* that is absent in other *HgaI* homologous RM systems.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12866-025-04270-3>.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements

We thank Klaus Pfeffer, director of the Institute of Medical Microbiology and Hospital Hygiene, for scientific discussion and institutional funding, and the "Jürgen Manchot Foundation" for granting personal funds. Computational

support and infrastructure were provided by the "Centre for Information and Media Technology" (ZIM) at the University of Duesseldorf (Germany). Furthermore, we want to thank Sebastian Scharf for the extensive help with nanopore sequencing.

Authors' contributions

BH designed the study. LV, MDS and DB carried out the experimental work. LV, AN and AD performed the bioinformatical analysis, and LV and BH wrote the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. This work was supported by the Jürgen Manchot Foundation and departmental funds from the Institute of Medical Microbiology and Hospital Hygiene, Heinrich-Heine-University of Duesseldorf, Germany.

Data availability

Genome sequences of *M. hominis* strains PX1114 (acc.-no. CP032849.1) and MIN-132 (CP086131.1) as well as the genomes of *Mesoplasma lactucae* (acc.-no. GCF_002441935.1), *Helicobacter pylori* (acc.-no. GCF_002831845.1) and *Lactococcus petauri* (acc.-no. GCF_023499275.1) were downloaded from NCBI (<https://www.ncbi.nlm.nih.gov/nucleotide/>). Normalized genome sequences from Nanopore assemblies of genomes of clinical *M. hominis* strains are available at OSF (<https://doi.org/10.17605/OSF.IO/3N4SY>). All other data, including Nanopore raw data will be provided on request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 5 November 2024 / Accepted: 28 July 2025

Published online: 25 August 2025

References

1. Taylor-Robinson D. Thoughts about *Mycoplasma hominis*. Sex Transm Infect. 2020;96(7):492.
2. Madoff S, Hooper DC. Nongenitourinary infections caused by *Mycoplasma hominis* in adults. Rev Infect Dis. 1988;10(3):602–13.
3. Henrich B, Hammerlage S, Scharf S, Haberhausen D, Furnkranz U, Kohrer K, et al. Characterisation of mobile genetic elements in *Mycoplasma hominis* with the description of ICEHo-II, a variant *Mycoplasma* integrative and conjugative element. Mobile DNA. 2020;11(1): 30.
4. Seong HJ, Han S-W, Sul WJ. Prokaryotic DNA methylation and its functional roles. J Microbiol. 2021;59:242–8.
5. Kobayashi I. Behavior of restriction–modification systems as selfish mobile elements and their impact on genome evolution. Nucleic Acids Res. 2001;29(18):3742–56.
6. Kobayashi I, Nobusato A, Kobayashi-Takahashi N, Uchiyama I. Shaping the genome – restriction–modification systems as mobile genetic elements. Curr Opin Genet Dev. 1999;9(6):649–56.
7. Naito T, Kusano K, Kobayashi I. Selfish behavior of restriction-modification systems. Science. 1995;267(5199):897–9.
8. Morozova N, Sabantsev A, Bogdanova E, Fedorova Y, Maikova A, Vedyaykin A, et al. Temporal dynamics of methyltransferase and restriction endonuclease accumulation in individual cells after introducing a restriction-modification system. Nucleic Acids Res. 2016;44(2):790–800.
9. Alvarez MA, Chater KF, Rodicio MR. Complex transcription of an operon encoding the sail restriction-modification system of *Streptomyces albus* G. Mol Microbiol. 1993;8(2):243–52.
10. O'Connor C, Humphreys G. Expression of the EcoRI restriction-modification system and the construction of positive-selection cloning vectors. Gene. 1982;20(2):219–29.

11. Bogdanova E, Djordjevic M, Papapanagiotou I, Heyduk T, Kneale G, Severinov K. Transcription regulation of the type II restriction-modification system AhdI. *Nucleic Acids Res.* 2008;36(5):1429–42.
12. Nagornykh M, Zakharova M, Protsenko A, Bogdanova E, Solonin AS, Severinov K. Regulation of gene expression in restriction-modification system Eco29kI. *Nucleic Acids Res.* 2011;39(11):4653–63.
13. Semenova E, Minakhin L, Bogdanova E, Nagornykh M, Vasilov A, Heyduk T, et al. Transcription regulation of the EcoRV restriction-modification system. *Nucleic Acids Res.* 2005;33(21):6942–51.
14. Bogdanova E, Zakharova M, Streeter S, Taylor JE, Heyduk T, Kneale G, et al. Transcription regulation of restriction-modification system Esp1396I. *Nucleic Acids Res.* 2009;37(10):3354–66.
15. Aravind L, Anantharaman V, Balaji S, Babu MM, Iyer LM. The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev.* 2005;29(2):231–62.
16. Aravind L, Koonin EV. DNA-binding proteins and evolution of transcription regulation in the archaea. *Nucleic Acids Res.* 1999;27(23):4658–70.
17. Wah DA, Hirsch JA, Dorner LF, Schildkraut I, Aggarwal AK. Structure of the multimodular endonuclease Fok I bound to DNA. *Nature.* 1997;388(6637):97–100.
18. McGeehan J, Papapanagiotou I, Streeter S, Kneale G. Cooperative binding of the C. AhdI controller protein to the C/R promoter and its role in endonuclease gene expression. *J Mol Biol.* 2006;358(2):523–31.
19. McGeehan J, Streeter S, Papapanagiotou I, Fox G, Kneale G. High-resolution crystal structure of the restriction-modification controller protein C. AhdI from *Aeromonas hydrophila*. *J Mol Biol.* 2005;346(3):689–701.
20. Zheleznyaya L, Kainov D, Yunusova A, Matvienko N. Regulatory C protein of the eco RV modification–restriction system. *Biochemistry (Moscow).* 2003;68:105–10.
21. Vijesurier RM, Carlock L, Blumenthal RM, Dunbar JC. Role and mechanism of action of C. Pvu II, a regulatory protein conserved among restriction-modification systems. *J Bacteriol.* 2000;182(2):477–87.
22. Ives CL, Sohail A, Brooks JE. The regulatory C proteins from different restriction-modification systems can cross-complement. *J Bacteriol.* 1995;177(21):6313–5.
23. Roberts RJ, Vincze T, Posfai J, Macelis D. REBASE: a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.* 2023;51(D1):D629–30.
24. Ershova A, Rusinov I, Spirin S, Karyagina A, Alexeevski A. Role of restriction-modification systems in prokaryotic evolution and ecology. *Biochemistry (Moscow).* 2015;80:1373–86.
25. Ren J, Lee H-M, Shen J, Na D. Advanced biotechnology using methyltransferase and its applications in bacteria: a mini review. *Biotechnol Lett.* 2022;44(1):33–44.
26. Sugisaki H, Yamamoto K, Takanami M. The hgaI restriction-modification system contains two cytosine methylase genes responsible for modification of different DNA strands. *J Biol Chem.* 1991;266(21):13952–7.
27. Niewmierzycka A, Clarke S. S-adenosylmethionine-dependent methylation in *Saccharomyces cerevisiae*: identification of a novel protein arginine methyltransferase. *J Biol Chem.* 1999;274(2):814–24.
28. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A.* 1992;89(5):1827–31.
29. Hashimoto K, Kokubun S, Itoi E, Roach HI. Improved quantification of DNA methylation using methylation-sensitive restriction enzymes and real-time PCR. *Epigenetics.* 2007;2(2):86–91.
30. Vogelgsang L, Nisar A, Scharf SA, Rommerskirchen A, Belick D, Diltz A et al. Characterisation of type II DNA methyltransferases of *Metamycoplasma hominis*. *Microorganisms.* 2023;11(6):1591.
31. Liu Y, Roskiewicz W, Pan Z, Jillette N, Wang P, Taghbalout A, et al. DNA methylation-calling tools for Oxford nanopore sequencing: a survey and human epigenome-wide evaluation. *Genome Biol.* 2021;22:1–33.
32. Roberts RJ, Carneiro MO, Schatz MC. The advantages of SMRT sequencing. *Genome Biol.* 2013;14:1–4.
33. Henrich B, Feldmann RC, Hadding U. Cytoadhesins of *Mycoplasma hominis*. *Infect Immun.* 1993;61(7):2945–51.
34. QIAGEN. DNeasy Blood & Tissue Handbook. Germany: QIAGEN; 2023 p. 29–32.
35. Henrich B, Kretzmer F, Deenen R, Kohrer K. Validation of a novel Mho microarray for a comprehensive characterisation of the *Mycoplasma hominis* action in HeLa cell infection. *PLoS One.* 2017;12(7): e0181383.
36. Möbius N, Brenneisen W, Schaeffer A, Henrich B. Protocol for the rapid detection of the urogenital tract mollicutes and Chlamydia with concomitant LGV-(sub) typing. In: MacKenzie C, Henrich B, editors. *Diagnosis of Sexually Transmitted Diseases: Methods*. 2012;903:235–53. https://doi.org/10.1007/978-1-61779-937-2_15.
37. Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene.* 1989;77(1):61–8.
38. Coppens L, Lavigne R. SAPHIRE: a neural network based classifier for $\sigma 70$ promoter prediction in *Pseudomonas*. *BMC Bioinformatics.* 2020;21(1):415.
39. Knudsen S. Promoter2.0: for the recognition of polii promoter sequences. *Bioinformatics.* 1999;15(5):356–61.
40. Naville M, Ghuillot-Gaudeffroy A, Marchais A, Gautheret D. ARNold: a web tool for the prediction of Rho-independent transcription terminators. *RNA Biol.* 2011;8(1):11–3.
41. Blackall PJ, Christensen H, Beckenham T, Blackall LL, Bisgaard M. Reclassification [haemophilus]f [haemophilus]asteurella gallinarum, [Haemophilus] [haemophilus]aragallinarum, [haemophilus]asteurella [haemophilus]vium [haemophilus]nd [haemophilus]asteurella volantium [haemophilus]s Avibacterium gallinarum gen. Nov., comb. Nov., Avibacterium [haemophilus]aragallinarum comb. Nov., Avibacterium [haemophilus]vium comb. Nov. [haemophilus]nd Avibacterium volantium comb. Nov. *Int J Syst Evol Microbiol.* 2005;55(1):353–62.
42. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A.* 1977;74(12):5463–7.
43. Kumar S, Cheng X, Klimasauskas S, Mi S, Posfai J, Roberts RJ, et al. The DNA (cytosine-5) methyltransferases. *Nucleic Acids Res.* 1994;22(1):1.
44. Thompson CC, Vieira NM, Vicente ACP, Thompson FL. Towards a genome based taxonomy of *Mycoplasmas*. *Infect Genet Evol.* 2011;11(7):1798–804.
45. Sugisaki H, Kita K, Takanami M. The FokI restriction-modification system: II. Presence of two domains in FokI methylase responsible for modification of different DNA strands. *J Biol Chem.* 1989;264(10):5757–61.
46. Roberts RJ, Belfort M, Bestor T, Bhagwat AS, Bickle TA, Bitinaite J, et al. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.* 2003;31(7):1805–12.
47. Clark TA, Murray IA, Morgan RD, Kislyuk AO, Spittle KE, Boitano M, et al. Characterization of DNA methyltransferase specificities using single-molecule, real-time DNA sequencing. *Nucleic Acids Res.* 2012;40(4):e29–e.
48. Dordet-Frisoni E, Vandecasteele C, Contarin R, Sagné E, Baranowski E, Klopp C, et al. Impacts of *Mycoplasma agalactiae* restriction-modification systems on pan-epigenome dynamics and genome plasticity. *Microb Genom.* 2022;8(5): 000829.
49. Malone T, Blumenthal RM, Cheng X. Structure-guided analysis reveals nine sequence motifs conserved among DNA amino-methyl-transferases, and suggests a catalytic mechanism for these enzymes. *J Mol Biol.* 1995;253(4):618–32.
50. Hrabárová E, Achbergerová L, Nahálka J. Insoluble protein applications: the use of bacterial inclusion bodies as biocatalysts. In: García-Fruitós E, editor. *Insoluble proteins. Methods in Mol Biol.* 2014;1258:411–22. https://doi.org/10.1007/978-1-4939-2205-5_24.
51. Peternel Š, Grdadolnik J, Gaberc-Porekar V, Komel R. Engineering inclusion bodies for non denaturing extraction of functional proteins. *Microb Cell Fact.* 2008;7(1):34.
52. Worrall D, Goss N. The formation of biologically active beta-galactosidase inclusion bodies in *Escherichia coli*. *Aust J Biotechnol.* 1989;3(1):28–32.
53. Jevšvar S, Gaberc-Porekar V, Fonda I, Podobnik B, Grdadolnik J, Menart V. Production of nonclassical inclusion bodies from which correctly folded protein can be extracted. *Biotechnol Prog.* 2005;21(2):632–9.
54. Wu JC, Santi D. Kinetic and catalytic mechanism of HhaI methyltransferase. *J Biol Chem.* 1987;262(10):4778–86.
55. Claus H, Friedrich A, Frosch M, Vogel U. Differential distribution of novel restriction-modification systems in clonal lineages of *Neisseria meningitidis*. *J Bacteriol.* 2000;182(5):1296–303.
56. Gerasimaitė R, Vilkaitis G, Klimasauskas S. A directed evolution design of a GCG-specific DNA hemimethylase. *Nucleic Acids Res.* 2009;37(21):7332–41.
57. Redaschi N, Bickle TA. Posttranscriptional regulation of EcoPII and EcoPII restriction activity. *J Mol Biol.* 1996;257(4):790–803.
58. Eckstein S, Brehm J, Seidel M, Lechtenfeld M, Heermann R. Two novel XRE-like transcriptional regulators control phenotypic heterogeneity in *Photobacterium luminescens* cell populations. *BMC Microbiol.* 2021;21:1–16.

59. Zhang Y, Liang S, Pan Z, Yu Y, Yao H, Liu Y, et al. XRE family transcriptional regulator XtrSs modulates *Streptococcus suis* fitness under hydrogen peroxide stress. *Arch Microbiol.* 2022;204(5):244.
60. Liu G, Gao T, Zhong X, Ma J, Zhang Y, Zhang S, et al. The novel streptococcal transcriptional regulator XtgS negatively regulates bacterial virulence and directly represses PseP transcription. *Infect Immun.* 2020;88(10):00035–20. <https://doi.org/10.1128/iai>.
61. Hu Y, Hu Q, Wei R, Li R, Zhao D, Ge M et al. The XRE family transcriptional regulator SrtR in *Streptococcus suis* is involved in oxidant tolerance and virulence. *Front Cell Infect Microbiol.* 2019;8:452.
62. Ulrich LE, Koonin EV, Zhulin IB. One-component systems dominate signal transduction in prokaryotes. *Trends Microbiol.* 2005;13(2):52–6.
63. Trouillon J, Ragno M, Simon V, Attrée I, Elsen S. Transcription inhibitors with XRE DNA-binding and cupin signal-sensing domains drive metabolic diversification in *Pseudomonas*. *Msystems.* 2021;6(1):00753–20. <https://doi.org/10.1128/msystems>.
64. Lu H, Wang L, Li S, Pan C, Cheng K, Luo Y, et al. Structure and DNA damage-dependent derepression mechanism for the XRE family member DG-DdrO. *Nucleic Acids Res.* 2019;47(18):9925–33.
65. Gómez LA, Molina RE, Soto RI, Flores MR, Coloma-Rivero RF, Montero DA, et al. Unraveling the role of the zinc-dependent metalloproteinase/HTH-Xre toxin/antitoxin (TA) system of *Brucella abortus* in the oxidative stress response: insights into the stress response and virulence. *Toxins.* 2023;15(9):536.
66. Mruk I, Kobayashi I. To be or not to be: regulation of restriction–modification systems and other toxin–antitoxin systems. *Nucleic Acids Res.* 2014;42(1):70–86.
67. McLaughlin M, Fiebig A, Crosson S. XRE transcription factors conserved in *caulobacter* and ϕ CbK modulate adhesin development and phage production. *PLoS Genet.* 2023;19(11):e1011048.
68. Semashko TA, Arzamasov AA, Evsyutina DV, Garanina IA, Matyushkina DS, Ladygina VG, et al. Role of DNA modifications in *Mycoplasma gallisepticum*. *PLoS One.* 2022;17(11):e0277819.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.