SUPPLEMENTARY MATERIALS

MATERIALS AND METHODS IN DETAIL

Bacterial strains, bacteriophages, plasmids, and oligonucleotides

All *E. coli* strains and plasmids used here are listed in Table S1. The synthetic oligonucleotides used are listed in Tables S2 and S3.

Bacteriophage lambda *vir* and bacteriophage P1 *vir* from our laboratory collection were used for the restriction assay and construction of *E. coli* mutant strains, respectively. We used *E. coli* HST08 (TaKaRa) for standard DNA cloning. *E. coli* strains were grown in LB broth (10 g tryptone, 5 g yeast extract, 10 g NaCl per 1 L) at 37^{o}C. Antibiotics were used at the following concentrations: kanamycin (Km), 30 ug/mL; ampicillin (Ap), 50 ug/mL; chloramphenicol (Cm), 25 ug/mL. Agar plates were made by adding 1.5% agar to LB broth.

P1 phage-mediated transduction was carried out as previously described (1). The donor strains *E. coli* JW2146 (*Δnfo*::*kan*), JW1738 (*Δxth*::*kan*), JW0221 (*ΔdinB*::*kan*), JW0059 (*ΔpolB*::*kan*), JW2669 (*ΔrecA*::*kan*), and JW2788 (*ΔrecB*::*kan*) are from the KEIO collection (1), and the recipient strain was *E. coli* T7 Express *lysY/I^{q}* (New England Biolabs, Ipswich, MA, USA). After transduction, pFLP3 plasmid was introduced into the transductant strain to remove the kanamycin resistance gene (2). The plasmids used for expression of restriction enzymes and the DNA methyltransferase were introduced into these strains.

A substrate plasmid that carries a single GTAC site was constructed by two cycles of site-directed mutagenesis. First, plasmid pHSG398 was PCR-amplified using primer pairs m-phsg398-F/m-phsg398-R and KOD FX Neo polymerase (Toyobo), and the PCR product was

purified using a Gen-Elute PCR Clean-UpKit (Sigma). 10 uL (≈ 200 ng) PCR product was digested with 10 U DpnI (Fermentas) in 1×FastDigest reaction buffer, and the mixture was used for heat-shock transformation into *E. coli* HST08 competent cells, which were prepared with 0.1 M $CaCl_2$ and spread on an LB agar plate (Cm, 25 ug/mL) for overnight incubation. Plasmid from the resulting single colony was used as a template for PCR with another primer pair (m-phsg398-F-2/m-phsg398-R-2), as above. These two steps introduced two mutations in the *cat* gene of pHSG398 as confirmed with RsaI (New England Biolabs) and KpnI (Fermentas). The resulting construct was designated as pHSG398m.

The synthetic genes for R.PabI homologs, R.HpyAXII and R.CcoLI, were designed, considering the protein sequences in NCBI resources (WP_000052868.1, WP_002830209.1) and codon optimization for expression in *E. coli*, by Funakoshi. The plasmids carrying the synthetic genes were designated pTAKN-2-*ccoLI* and pTAKN-2-*hpyAXII*. The coding regions of the two homologs were PCR-amplified with primer pairs CcoLI-F/CcoLI-R for CcoLI and HpyAXII-F/HpyAXII-R for HpyAXII (Table S2) and then cloned into plasmid pET28a using an In-Fusion® HD Cloning Kit (TaKaRa). The resulting plasmids were designated as pET28a_*ccoLIR* and pET28a_*hpyAXIIR*. The mutation D214A of R.PabI coding region was amplified with the primer pairs P-D214A-R/P-D214A-F (carrying mutation site) and pET28a_*pabIR* plasmid as template, and then the coding region was cloned into pET28a.The resulting construct was designated as pET28a_*pabIRD214A*.

Restriction enzyme expression *in vivo*

pET28a::*pabIR,* pET28a::*ccoLIR,* or pET28a::*hpyAXIIR* was introduced into *E. coli* T7 Express *lysY/I*^{q} harboring pBAD30_*cviQIM* (3). Methyltransferase M.CviQI was always induced in *E. coli*

T7 Express *lysY/I*^{q} with 0.5% arabinose. If necessary, 0.5 mM IPTG and 0.5% glucose were added to control the expression level of the restriction enzyme.

Purification of restriction enzymes

E. coli T7 Express *lysY/I*^{q} harboring pBAD30_*cviQIM* and one of the restriction enzyme expression plasmids, pET28a:: *pabIR*, pET28a:: *ccoLIR*, or pET28a:: *hpyAXIIR*, was streaked on LB agar with 50 ug/mL Ap, 30 ug/mL Km, and 0.5% arabinose. After overnight incubation, a single colony was picked up and transferred to 5 mL LB medium containing the same concentrations of reagents as above. After overnight incubation, the 5-mL culture was transferred to 1 L LB medium containing the same concentrations of reagents. After 3 h of incubation with shaking, IPTG was added to a final concentration of 0.5 mM, and the culture was further incubated for 6 h. The cells were collected by centrifugation at 5 krpm for 10 min at 4^{o}C and resuspended in 15 mL of 20 mM Tris-HCl, 150 mM NaCl, pH 7.5. Then they were sonicated (Ultrasonic Disruptor UD-200, TOMY) and centrifuged at 7 krpm for 20 min. The supernatant containing R.PabI or R.PabI(D214A) was heated at 75^{o}C in a water bath for 120 min while the supernatant containing R.HpyAXII or R.CcoLI was left on ice. The supernatants were centrifuged at 7 krpm for 20 min again and filtered using a 0.45-um PDVF filter.

The filtrates were bound to 4 mL of Ni-NTA Agarose resin (QIAGEN, NO.1018244) equilibrated with a binding buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl), and washed with the same buffer, and eluted in 2 mL of elution buffer (20 mM Tris-HCl pH7.5, 100 mM NaCl, 300 mM imidazole) using a gravity flow column. The fused His-tag was removed by thrombin digestion using a Thrombin Cleavage Capture Kit (Novagen, No.69022-3FRZ) following the manufacturer's protocol. The eluate was reacted with 1uL (1U) thrombin at 25^{o}C overnight. The

agarose resin and His-tag were removed using a centrifuge column (Pierce Centrifuge Columns, Thermo, 0.5 mL). The flowthrough containing the restriction enzyme was directly loaded onto a Heparin HP column (HiTrap Heparin HP, GE healthcare Life Sciences, No. 17-0406-01) equilibrated with a Heparin binding buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl), and then the proteins were eluted in 2 mL elution buffer for heparin (20 mM Tris-HCl pH 7.5, 2 M NaCl). The proteins were concentrated using a centrifugal filter (Amicon Ultra-0.5mL, Ultracel-10K, UFC501096).

The buffer in the protein solution was replaced with 2x stock buffer (20 mM MES pH 6.0, 200 mM NaCl) using a centrifugal filter (Amicon Ultra-0.5mL, Ultracel-10K, UFC501096) and the protein was concentrated. Concentrated protein solution was mixed with appropriate volumes of glycerol, 100 mM EDTA, and 100 mM DTT to generate a final protein stock solution in 10 mM MES pH 6.0, 100 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 50% glycerol.

R.PabI interacts with a cognate DNA as a dimer in a cocrystal (5). In gel filtration (HiLoadSuperdex 200 prep grade, GE healthcare), its native form gave a peak consistent with its dimer form (4)(5).

DNA cleavage assay

An 861-bp linear double-stranded DNA fragment was used as a substrate for the cleavage assay (Figure S2 B).This substrate was generated by PCR-amplification of the ampicillin resistance gene (containing one GTAC site) of pBAD30 using primers pBAD30F and pBAD30R (Table S2).The cleavage reaction was performed in 10 uL with 0.1 M sodium phosphate buffer (pH 6.5), 2.66 pmol (266 nM) of purified enzyme, and 0.38 pmol (0.38 nM) of substrate at 85^{o}C (R.PabI/ R.PabI(D214A)) or 37^{o}C (R.HpyAXII/R.CcoLI) for 1 h. DNA was separated by agarose gel

electrophoresis and visualized with ethidium bromide and ultraviolet light.

DNA glycosylase assay

Each of the top and bottom strands of the 40-mer oligonucleotide with a single “GTAC” (Table S2) was labeled with γ ^{32}P-ATP (Perkin-Elmer) using T4 polynucleotide kinase (New England Biolabs), and annealed with its complementary oligonucleotide. 0.2 pmol of the 40-bp substrate was incubated with 0-1.4 pmol of purified enzyme (that is, with 4-fold dilution) in 20 uL of 0.1 M phosphate buffer (pH 6.5) at 60^{o}C (R.PabI) or 37^{o}C (R.HpyAXII, R.CcoLI) for 1 h. For R.CcoLI, we also incubated the 0.2 pmol 40-bp substrate with 0-0.2 pmol enzyme in 20uL of 20 mM MOPS-KOH (pH 7.0) for 1 h. Half the reaction mixture was treated with 0.1 M NaOH at 70^{o}C for 10 min to cleave DNA at the generated AP sites, and then neutralized with HCl. The products were separated by 18% denaturing PAGE. Gels were exposed to an imaging plate, and the plate image was detected using an FLA5100 scanner (Fujifilm).

AP lyase assay

To construct the AP site-containing 40-bp substrate, uracil-containing (5’-GTUC/3’-CUTG) double-stranded oligonucleotides (Table S2) were labeled and generated as described above, and then incubated with uracil N-glycosylase (UNG) (New England Biolabs) (3). Complete conversion to AP sites was confirmed by NaOH cleavage. An AP site is present instead of an adenine nucleotide in the recognition sequence (5’-GTAC/3’-CATG) in this substrate.

The reaction was performed in 20 uL of 0.1 M phosphate buffer (pH 6.5) or 20 mM MOPS buffer (20 mM MOPS-KOH, pH 7.0, 50 mM NaCl, 1 mM EDTA, 1 mM DTT), containing 0.2 pmol

of AP site-containing substrate (^{32}P-labeled top or bottom strand) and 0-2 pmol of purified R.CcoLI at 37°C for 1h. Samples were separated by 18% denaturing PAGE.

$NaBH_4$ trapping

The method described (3) was modified as follows. Briefly, R.CcoLI (0-6 pmol, 0-300 nM) and a 40-bp substrate DNA (0.2 pmol, 10 nM), containing 5'-GT#C (# = AP site)/3'-C#TG and a 5'-^{32}P label on the top strand, were incubated in 20 uL of 20 mM MOPS buffer (20 mM MOPS-KOH, pH 7.0, 50 mM NaCl, 1 mM EDTA, 1 mM DTT) at 37°C for 20 min and then with 100 mM $NaBH_4$ at 25°C for 30 min. DNA-R.CcoLI complexes were denatured in gel loading buffer containing 3% SDS at 90°C for 10 min and separated through 10% SDS-PAGE.

For the control, a single-strand labeled oligonucleotide with a single "GTUC" was reacted with uracil N-glycosylase to generate a single AP site and annealed with a bottom strand with "GTAC". 20 units of Endo III (New England Biolabs) and 0.2 pmol (10 nM) of this P^{32}-5'-GT#C/3'-CATG substrate were incubated in 20 mM Tris-HCl (pH 8), 1 mM EDTA , 1 mM DTT (total 20 uL) at 37°C for 20 min followed by incubation with 100 mM $NaBH_4$, as above. DNA-Endo III complexes were analyzed as described for R.CcoLI.

Restriction of phage propagation

The wild-type or mutant *E. coli* strain harboring (or not harboring) an RM expression plasmid was grown at 37°C overnight in LB medium with 50 ug/mL Ap, 30 ug/mL Km, and 0.5% arabinose. Then the culture was diluted 100-fold and grown in tryptone broth (1% tryptone and 0.5% NaCl), supplemented with 0.2% maltose and 10 mM $MgSO_4$, at 37°C for 1h. IPTG was added to 0.1 mM,

and incubated at 37^{o}C for 3 h. 2 mL of culture was mixed with 2 mL top agar (1% polypeptone, 0.5% NaCl, 0.6% agar), vortexed for several seconds, and then poured onto a dried bottom agar plate (1% polypeptone, 0.5% NaCl, 1% agar) and allowed to solidify at room temperature for 10 min. Bacteriophage lambda *vir* (plaque forming units/mL≈ 1.2×10^{7}) was serially diluted and 5 uL of each dilution was spotted on the plates. The plates were incubated at 37^{o}C overnight for plaque counting and estimating plaque-forming units/mL.

Restriction of the chromosome *in vivo*

The wild-type or the mutant strain harboring (or not harboring) an RM expression plasmid was grown overnight at 37^{o}C in LB medium with 50 ug/mL Ap, 30 ug/mL Km, and 0.5% arabinose, and the culture was diluted 100-fold in LB medium with 50 ug/mL Ap, 30 ug/mL Km, and 0.5% arabinose and grown at 37^{o}C for 2-3 h until OD_{600nm} = 0.2 by a plate reader. IPTG was added to a final concentration of 0.1 mM prior to incubation at 37^{o}C for 0.5 h. The culture was chilled on ice. 1 mL of the culture was centrifuged at 5 krpm for 10 min at 4^{o}C, and the cells were resuspended and adjusted to OD_{600nm} = 0.2 with chilled saline. The suspension was diluted 10-fold with chilled saline and immediately spotted on the following plates: (i) LB agar with 0.5% arabinose, 50 ug/mL Ap, and 30 ug/mL Km; (ii) LB agar with 0.5% glucose, 50 ug/mL Ap, and 30 ug/mL Km. The plates were incubated overnight at 37^{o}C before colony counting and calculation of the colony-forming units/mL.

Restriction in transformation

Plasmid pBAD30_*cviQIM* (0.84 pmol) was treated with purified R.PabI (8.4 pmol) in 50 uL of 0.1

M sodium phosphate buffer (pH 6.5) at 37°C for 1 h as previously described (3) to generate AP sites, and the product was purified with a Gen-Elute PCR Clean-Up Kit (Sigma). A 200-ng equivalent of the AP site containing pBAD30_*cviQIM* was transferred into the original and the mutant T7 Express *lysY/I*^{q} *E. coli* strains by electroporation to count the colony-forming units/mL.

To evaluate the re-circularization efficiency of restriction glycosylase-treated plasmids, plasmid pHSG398m (2.1 pmol) was first treated with purified R.PabI (10.2 pmol), R.CcoLI (10.2 pmol), or KpnI (20 U) (Fermentas) in 50 uL at 37°C for 1 h. R.PabI and R.CcoLI reactions were conducted in 0.1 M sodium phosphate buffer (pH 6.5), while the KpnI reaction was in 1× KpnI buffer (Fermentas). The R.PabI products were further purified using a Gen-Elute PCR Clean-Up Kit and treated with Endonuclease IV (New England Biolabs) at 37°C for 1 h to cleave the DNA at an AP site. All linearized products were purified by agarose gel electrophoresis and with a NucleoSpin Extract II Kit (Macherey-Nagel). Then, the three types of linearized plasmid products were treated with T4 DNA ligase (New England Biolabs) at 16°C for 2 h. The products (100 ng each) were dissolved in 0.1 M $CaCl_2$ and transferred to *E. coli* HST08 competent cells by heat-shock transformation.

Motif frequency analysis

Expected motif frequency (E_{GTAC}) was defined following the Markov maximum order model (6): $E_{GTAC} = (N_{TAC} \times N_{GTA})/N_{TA}$, where N_{GTA}, N_{TAC}, and N_{TA} are the numbers of motifs GTA, TAC, and TA, respectively. Motif frequency R_{GTAC} (observed count/expected count) was defined as N_{GTAC}/E_{GTAC}. This calculation was performed using Bioconductor package Biostrings (7). Motif frequency was determined for complete chromosome sequences (Figure 6) of 22 prokaryotic species and complete sequences of *H. pylori* plasmids (Figure S9) retrieved from NCBI RefSeq

database.

Restriction enzyme expression *in vitro*

Two protein-coding regions of the *pabIR* homologs from plasmids pTAKN-2-*ccoLI* and pTAKN-2-*hpyAXII* were inserted into pEU3-NIIb, a plasmid for cell-free protein synthesis (8). The R.PabI homolog coding region was amplified by PCR with KOD-plus (Toyobo) from plasmids with a BamHI site attached at their 3' ends by using primers CF-CcoLI-F/CF-CcoLI-R and CF-HpyAXII-F/CF-HpyAXII-R (Table S2). The resulting plasmids were designated as pEU3-NIIb_*ccoLIR* and pEU3-NIIb_*hpyAXIIR*. This connected the coding region to an SP6 promoter (8). And a translation promotion signal, omega, at its 5' end. The amplified DNA fragment was inserted between the EcoRV and BamHI sites of pEU3-NIIb.

The putative restriction enzymes were expressed in a wheat-germ-based cell-free protein synthesis system (Protein Research Kit (S), CellFree Science). The resulting solutions were centrifuged at 12 krpm for 10 min at 4°C and soluble fractions were recovered.

To test DNA cleavage activity, 1 uL of the soluble fraction was incubated with 150 ng of lambda DNA (TaKaRa) in 1x NEB1 buffer (New England Biolabs) at 37°C for 1 h, which was followed by 0.8% agarose gel electrophoresis and detection with ethidium bromide or Gel Red (Biotium Inc.) by ultraviolet light irradiation.

PCR analysis of deletion mutant strains

Constructed mutant *E. coli* strains *nfo, xth, recA, recB, dinB* and *polB,* were confirmed by colony PCR with the corresponding primers, nfo-F/nfo-R, recA-F/recA-R, recB-F/recB-R, dinB-F/dinB-R

and polB-F/polB-R (Table S2). The wild type and mutant strains from KEIO-collection were as controls as well. The sample was followed by 0.8% agarose gel electrophoresis and detection with ethidium bromide by ultraviolet light irradiation.

Helicobacter pylori **strains for transcriptome analysis**

H. pylori strains were grown in Brucella Broth (BB) (BD Bioscience) supplemented with 10% fetal bovine serum (FBS) (Cell Culture Laboratories) with shaking at 37^{o}C in a multi-gas CO_2 incubator in the presence of 10% CO_2 and 5% O_2. Primers used for this experiment are listed in Table S3. *Helicobacter pylori* strain P12 derivatives carrying a deletion in the *pabIR* homolog (=HPP12_0511) alone (strains PIK65, PIK70) or both the *pabIR* homolog and the *pabIM* homolog (=HPP12_0510) (strain PIK69) were constructed using homologous recombination to avoid post-segregational killing (9).The approximately 800-bp regions flanking HPP12_0511, and the chloramphenicol resistance gene on plasmid pHEL2, were PCR-amplified using primer pairs HPP12_0511_Lf and HPP12_0511_Lr (for HPP12_0511 upstream), HPP12_0511_Rf and HPP12_0511_Rr (for HPP12_0511 downstream), and catf and catr for the chloramphenicol resistance gene with KOD FX Neo DNA polymerase (Toyobo). These three fragments were combined using PCR. Then the combined fragment was inserted into vector pUC19 using an In-Fusion HD Cloning Kit (TaKaRa). The resulting construct was designated pMZA3. Next, the insert in pMZA3 was PCR-amplified using the primer pair HPP12_0511_Lf and HPP12_0511_Rr and then electroporated into P12 cells washed with 300 mM sucrose. After overnight incubation on blood agar plates (BBL™ Trypticase™ Soy Agar with 5% Sheep Blood, BD Bioscience), the cell mixture was diluted in 0.85% NaCl and plated on Brucella broth containing 10% fetal bovine serum (FBS), 1.5% agar (BB-FBS agar), and 10 ug/mL chloramphenicol. The candidate

transformant colonies were streaked on a selective plate, and then one of the clones that appeared was designated as PIK65 (ΔHPP12_0511).

The chloramphenicol resistance gene region of PIK65 or the chloramphenicol resistance gene region together with its upstream gene HPP12_0510 (*pabIM* homolog) was further deleted from PIK65 using the same method. The plasmid constructs used in this experiment were pHY1278 and pHY1279. For construction of pHY1278, approximately 800-bp regions of HPP12_0510 upstream and HPP12_0511 downstream and the kanamycin resistance gene (*aphA-3*) on plasmid pHEL3 were PCR-amplified using primer pairs, HPP12_0510_Lf and HPP12_0510_Lf, HPP12_0511_Rf4 and HPP12_0511_Rr, and AphA3f and AphA3r, respectively. Then, the two fragments were combined by PCR using primers HPP12_0510_Lf and HPP12_0511_Rr. The combined fragment was inserted into pUC19. This gave rise to pHY1278. Similarly, pHY1279, which carries approximately 800-bp regions of HPP12_0511 upstream and HPP12_0511 downstream and the kanamycin resistance gene on pUC19, was constructed. The inserts on pHY1278 and pHY1279 were PCR-amplified, and then electroporated into competent cells of PIK65. Kanamycin-resistant clones were selected and purified on a BB-FBS ager plate with 5 ug/mL kanamycin. The obtained clones were designated as PIK70 (ΔHPP12_0511; $r_{HpyPXII}^{-}$ $m_{HpyPXII}^{+}$) and PIK69 (Δ(HPP12_0510-HPP12_0511; $r_{HpyPXII}^{-}$ $m_{HpyPXII}^{-}$).

Transcriptome analysis

Total RNA was extracted from two replicate exponential-phase cultures (OD_{600nm} = 0.4-0.5) of PIK69 and PIK70 using a PureLink RNA-min Kit (Thermo Fisher Scientific). Then, ribosomal RNA was removed using the Ribo-Zero rRNA Removal Kit for Gram-negative bacteria (Epicentre). The rRNA-depleted samples were used for cDNA library construction using a

SureSelect Strand Specific RNA Seq Library Construction Kit (Agilent Technologies). The cDNA library was sequenced on the HiSeq2500 platform. 100 base-long reads were mapped onto the strain P12 chromosome and plasmid pHPP12 using BWA, and then the number of reads mapped to the coding regions were counted using HTseq. To detect differentially expressed genes, read counts were normalized and compared between PIK70 (r^-m^+) and PIK69 (r^-m^-) using Bioconductor TCC, following TMM-normalization and the edgeR iteration protocol. Read depth per strand was calculated using BEDtool.

Accession number

NGS data was submitted under DRA accession no. DRA004356.

References

1. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L. and Mori, H. (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*, **2**, 2006 0008.

2. Hoang, T.T., Karkhoff-Schweizer, R.R., Kutchma, A.J. and Schweizer, H.P. (1998) A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked Pseudomonas aeruginosa mutants. *Gene*, **212**, 77-86.

3. Fukuyo, M., Nakano, T., Zhang, Y., Furuta, Y., Ishikawa, K., Watanabe-Matsui, M., Yano, H., Hamakawa, T., Ide, H. and Kobayashi, I. (2015) Restriction-modification system with methyl-inhibited base excision and abasic-site cleavage activities.*Nucleic Acids Res.*, **43**, 2841-2852.

4. Miyazono, K., Watanabe, M., Kosinski, J., Ishikawa, K., Kamo, M., Sawasaki, T., Nagata, K., Bujnicki, J.M., Endo, Y., Tanokura. And Kobayashi, I. (2007) Novel protein fold discovered in the PabI family of restriction enzymes. *Nucleic Acids Res.*,**35**, 1908-1918.

5. Miyazono, K., Furuta, Y., Watanabe-Matsui, M., Miyakawa, T., Ito, T., Kobayashi, I. and Tanokura, M. (2014) A sequence-specific DNA glycosylase mediates restriction-modification in Pyrococcus abyssi. *Nature Communications*, **5**, 3178.

6. Rocha, E.P., Viari, A. and Danchin, A. (1998) Oligonucleotide bias in Bacillus subtilis: general trends and taxonomic comparisons.*Nucleic Acids Res.*,**26**, 2971-2980.

7. Pagès, H., Aboyoun, P., Gentleman, R. and DebRoy, S. (2016) Biostrings: String objects representing biological sequences, and matching algorithms. *R package version* 2.40.1.

8. Ishikawa,K., Watanabe,M., Kuroita,T., Uchiyama,I., Bujnicki,J.M., Kawakami,B., Tanokura,M. and Kobayashi,I. (2005) Discovery of a novel restriction endonuclease by genome comparison and application of a wheat-germ-based cell-free translation assay: PabI (5prime;-GTA/C) from the hyperthermophilic archaeon Pyrococcus abyssi. *Nucleic Acids Res.*,**33**, 1–10.

9. Kusano,K., Naito,T., Handa,N. and Kobayashi,I. (1995) Restriction-modification systems as genomic parasites in competition for specific sequences. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 11095–11099.

Figure S1. Cleavage of lambda DNA with R.PabI homologs expressed *in vitro.*
R.CcoLI and R.HpyAXII were expressed *in vitro.* 10 ul reaction mixture with 1 ul of the soluble fraction was incubated with 150 ng of lambda DNA in 1 x NEB1 buffer at 37^{o}C for 1 hour, electrophoresed through agarose gel and visualized with ethidium bromide and ultraviolet light.

Figure S2. Purification of R.PabI, R.PabID214A, R.HpyAXII and R.CcoLI.
A. Purified R.PabI, R.PabID214A, R.HpyAXII and R.CcoLI were visualized in 12% SDS-PAGE. B. Upper: PCR product (861 bp) with a single GTAC (0.38 pmol, 200 ng). Lower: This substrate (s) was treated with R.PabI or R.PabID214A (2.66 pmol) in 100 mM phosphate buffer, pH 6.5, at 85 ^{o}C for 1h, electrophoresed through agarose gel and visualized with ethidium bromide and ultraviolet light.

Figure S3. Glycosylase and AP lyase activities of R.CcoLI under an optimaized condition.
The reaction condition was optimized for glycosylase activity in 20 mM MOPS buffer (50 mM NaCl, pH 7.0) at 37^{o}C. **A.** Glycosylase. A 40 mer oligo with single “GTAC” (0.2 pmol) (Top or bottom strand P^{32} labeled at 5’) was treated with 0 - 0.2 pmol R.CcoLI for 1h. The mixture was further treated with 0.1 M NaOH at 70^{o}C for 10 min, neutralized with HCl and electrophoresd through 18% denaturing PAGE. **B.** AP lyase. An oligo with 5’-GTUC/3’-CUTG was treated with Uracil DNA Glycosylase to generate the double AP sites. A 20 ul reaction mixture with 0.1 pmol of the AP-site substrate (P^{32}-labeled in the top or bottom strand at 5’ end) and 0 - 0.2 pmol R.CcoLI for 1 h. Samples were separated through 18% denaturing PAGE. n=3, error bar indicate standard deviation.

Figure S4. PCR analysis of deletion mutant strains.
The primers for the amplification of the genes *nfo, xth, polB, dinB, recA, recB* were listed in Table S2. The wild type *E. coli* T7 Express *lysY/I^{q}*, deletion mutants *E. coli* from KEIO-collection and the mutant strains of T7 Express *lysY/I^{q}* after P1 phage transduction were used for bacterial colonies PCR, electrophoresed through agarose gel and visualized with ethidium bromide and ultraviolet light.

Figure S5. Construction of a plasmid with a single PabI site (= KpnI site).

A. Plasmids. **B.** Confirmation of structure of pHSG398m by cleavage with restriction enzymes. R.RsaI recognizes 5'GTAC as R.PabI.

Figure S6. Cleavage of hemi-methylated DNA by R.CcoLI.

Hemi-methylated substrates were generated by annealing of different strands. The substrates were with P^{32} label at 5' end of top strand. A 10 ul reaction mixture with 0.1 pmol substrate and 0.7 pmol R.CcoLI in 20 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT was incubated at 37^{o}C for 1h, and samples were separated through 10% denaturing PAGE.

Figure S7. Cleavage of plasmid DNA by R.CcoLI

Plasmid pHSG398m with single "GTAC" site was treated with R. CcoLI at 37^{o}C.

A. 0.28 pmol pHSG398m plasmid DNA was treated with 2.8 pmol R.CcoLI in 100 mM phosphate buffer, pH 6.5, at 37^{o}C for 0-10min. **B.** 0.28 pmol pHSG398m plasmid DNA was treated with 0.028 pmol R.CcoLI in 100 mM phosphate buffer, pH 6.5, at 37^{o}C for 0-10 min. n=3, error bar indicate standard deviation.

Figure S8. Outcompetition of the AP lyase activity of R.CcoLI by a GTAC-containing duplex.

A 40-mer oligo with 5'-GT**#**C/3'-C#TG (0.2 pmol, # = AP site, top strand 32P-labeled at 5') was treated with 0.25 pmol R.CcoLI in the presence of the indicated concentrations of an unlabeled specific (GTAC40T/GTAC40B) or nonspecific (CTAC40T/CTAC40B, Table S2) competitor in 100 mM phosphate buffer, pH 6.5, 37^{o}C for 1 h. The samples were separated by 18% denaturing PAGE. After correction of nonspecific DNA cleavage during incubation, the percentage of nicked DNA (5'-GT#C) is plotted against the amount of competitors. Data are the means of two independent experiments.

Figure S9. GTAC motif avoidance in *Helicobacter pylori* plasmids (N=41).

Some plasmids show GTAC motif frequency higher than chromosome average (0.10, red line). Note that some plasmids without microcin operon (*mcc*$^{-}$) also show relatively high GTAC motif frequency.

Figure S10. Effect of *hpyPIVM* knockout on *H. pylori* gene expression.

Upper panel: genetic map of the HpyPVI (PabI homolog) locus in the two isogenic strains PIK70 (up) and PIK69 (bottom). Lower panel: RNA-seq read coverage on the plasmid pHel12 indicating the reduced expression of the *mcc* operon in the R-M- strain PIK69. The *mccB* gene (red) is

differentially expressed. Vertical lines on the lower gene map indicate GTAC sites. TA, toxin-antitoxin. *aphA*: a kanamycin resistance gene.

Table S1. *E. coli* strains and plasmids used.

Table S2. Oligos or primers used in *E. coli* experiment.

Table S3. Primers used in *Helicobacter pylori* transcriptome analysis.

Table S4. Genes differentially expressed upon knockout of the methyltransferase gene.

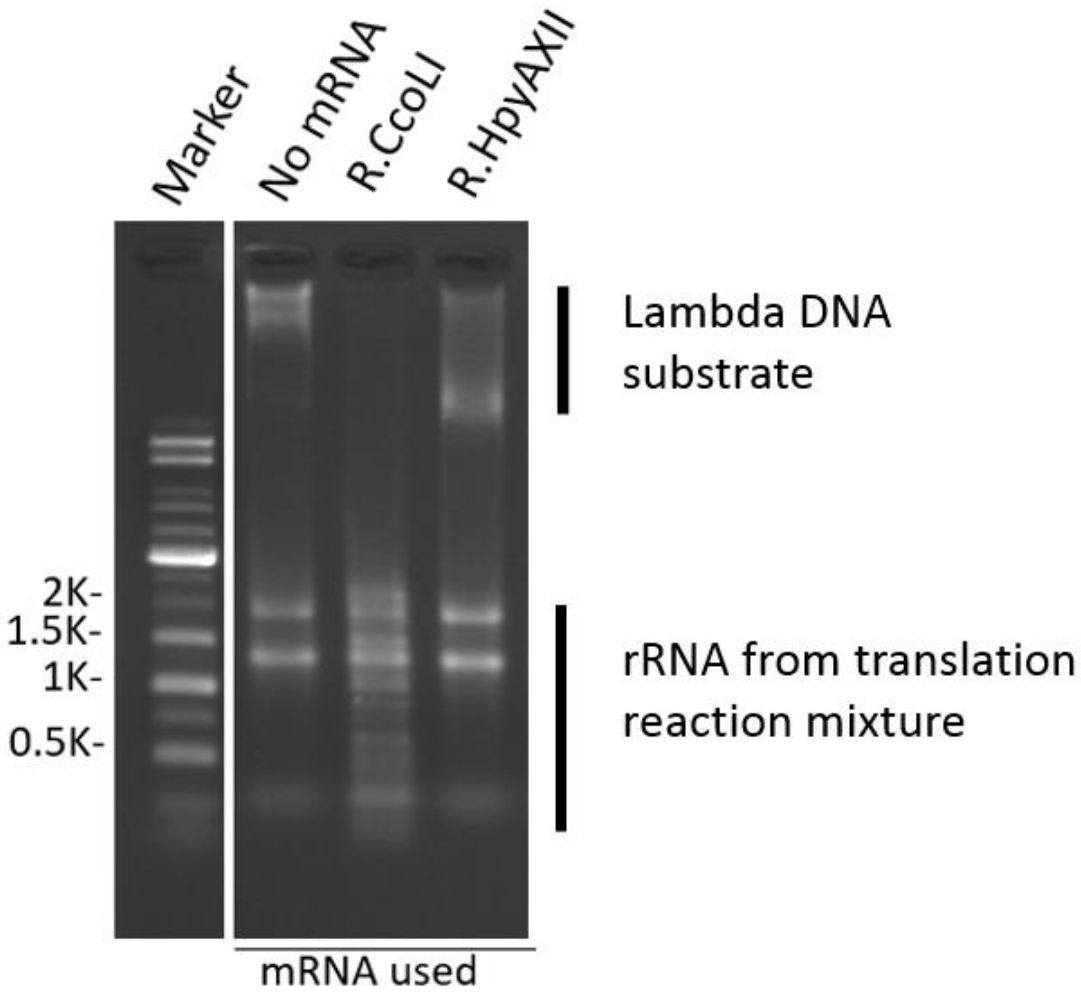

Figure S1. Cleavage of lambda DNA with R.PabI homologs expressed *in vitro*

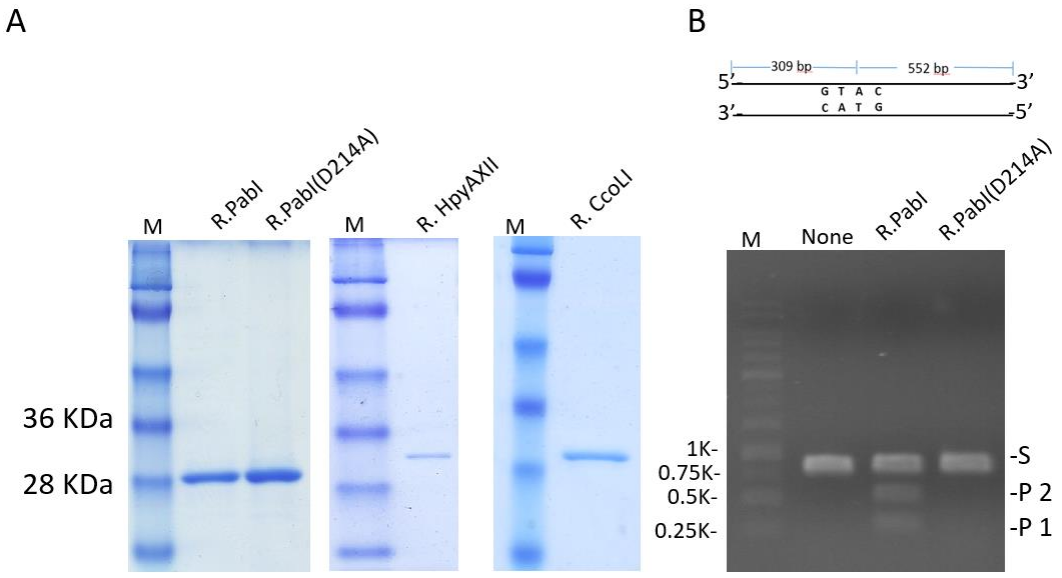

Figure S2. Purification of R.PabI, R.PabID214A, R.HpyAXII and R.CcoLI

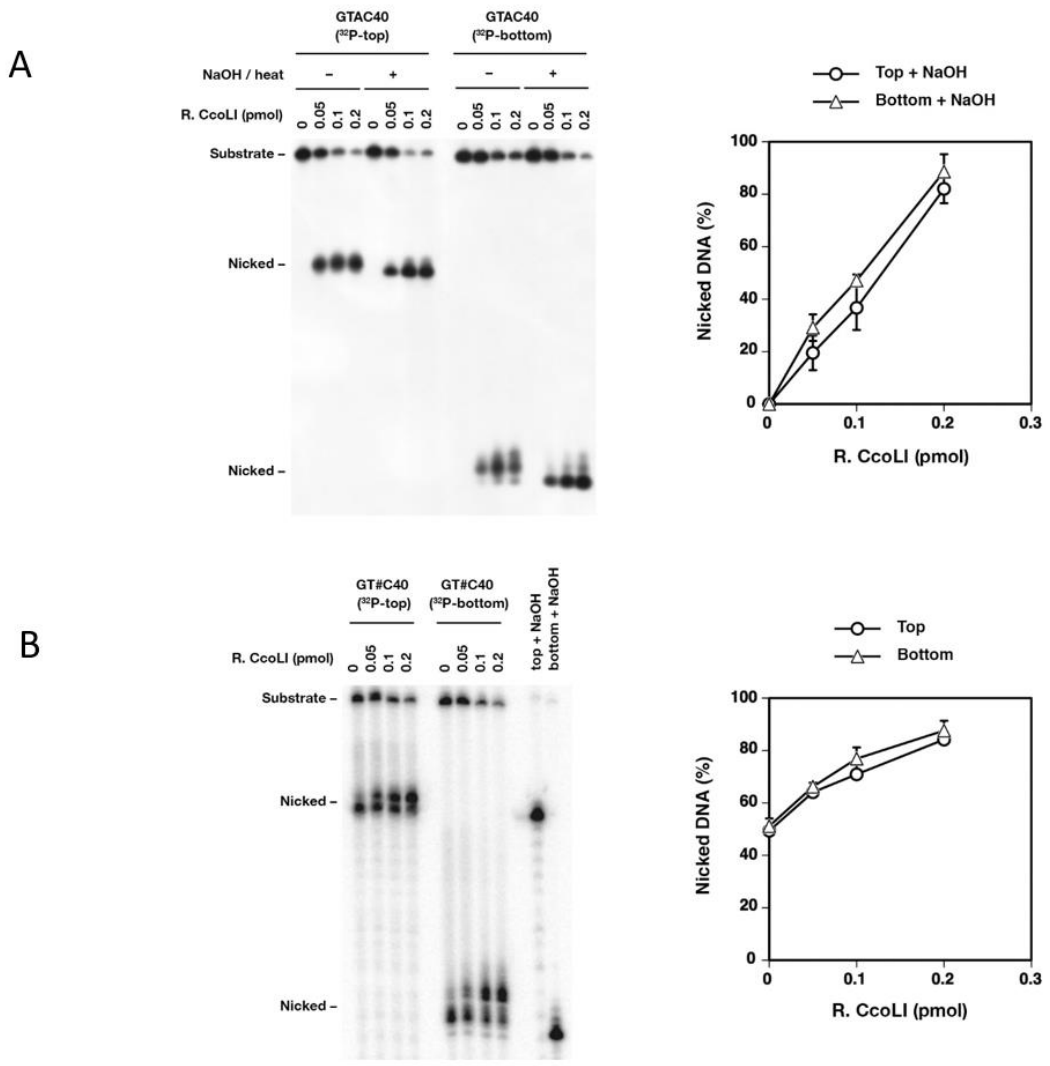

Figure S3. Glycosylase and AP lyase activities of R.CcoLI under an optimized condition

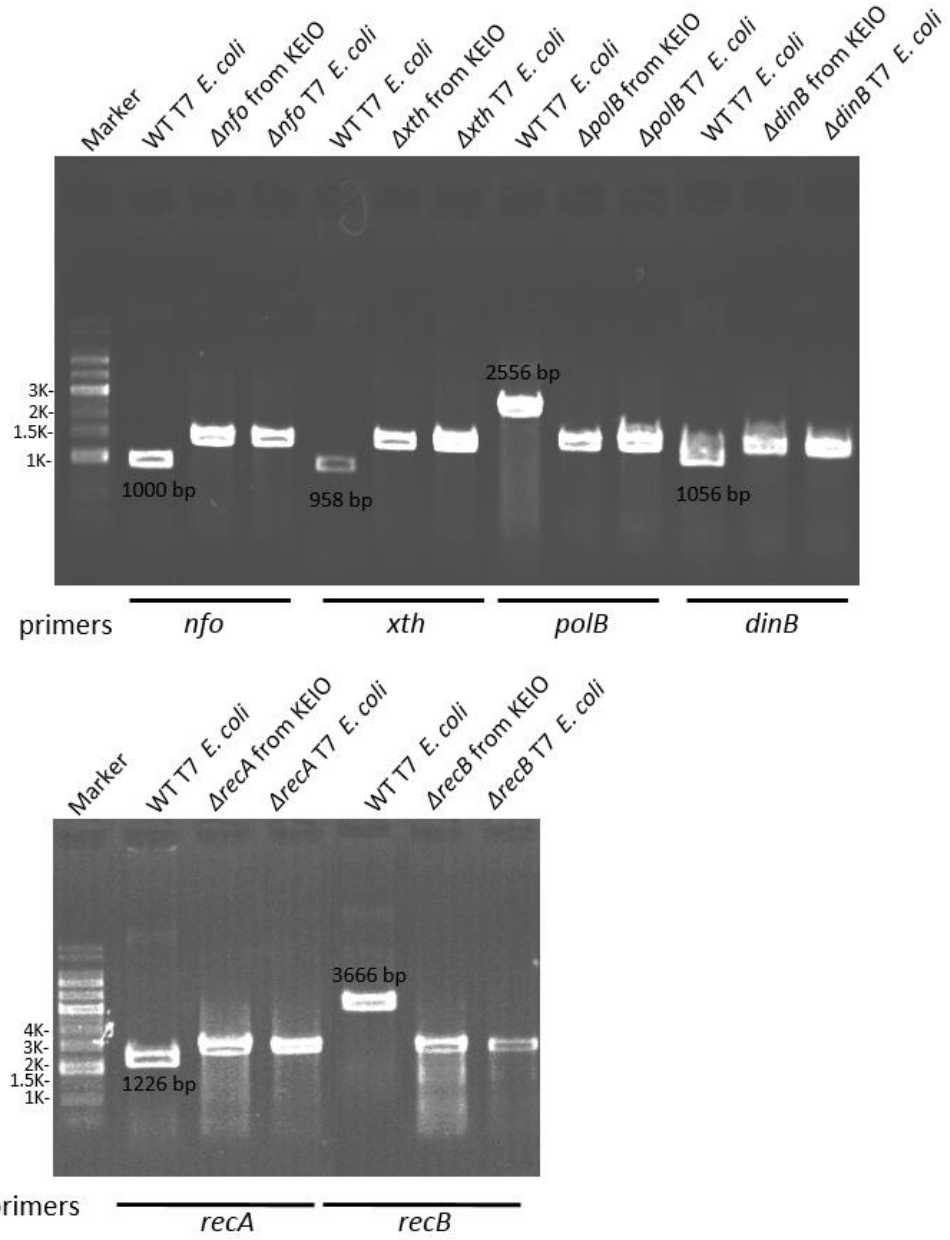

Figure S4. PCR analysis of deletion mutant strains.

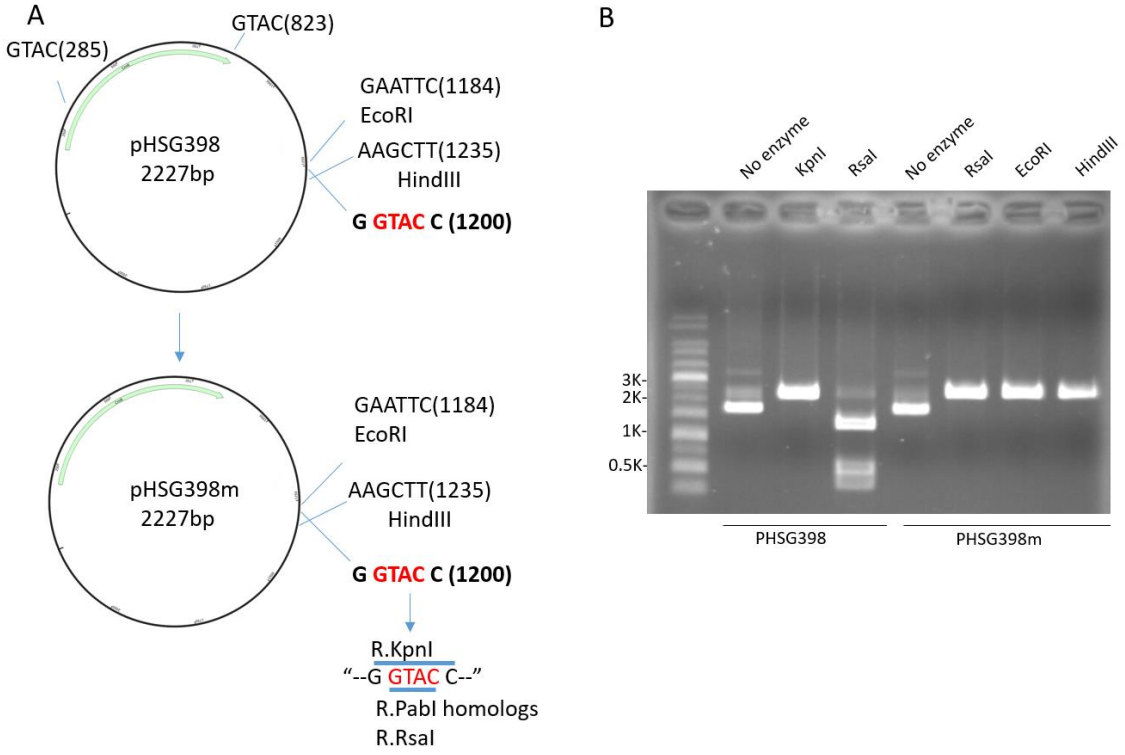

Figure S5. Construction of a plasmid with a single R.PabI site (= KpnI site)

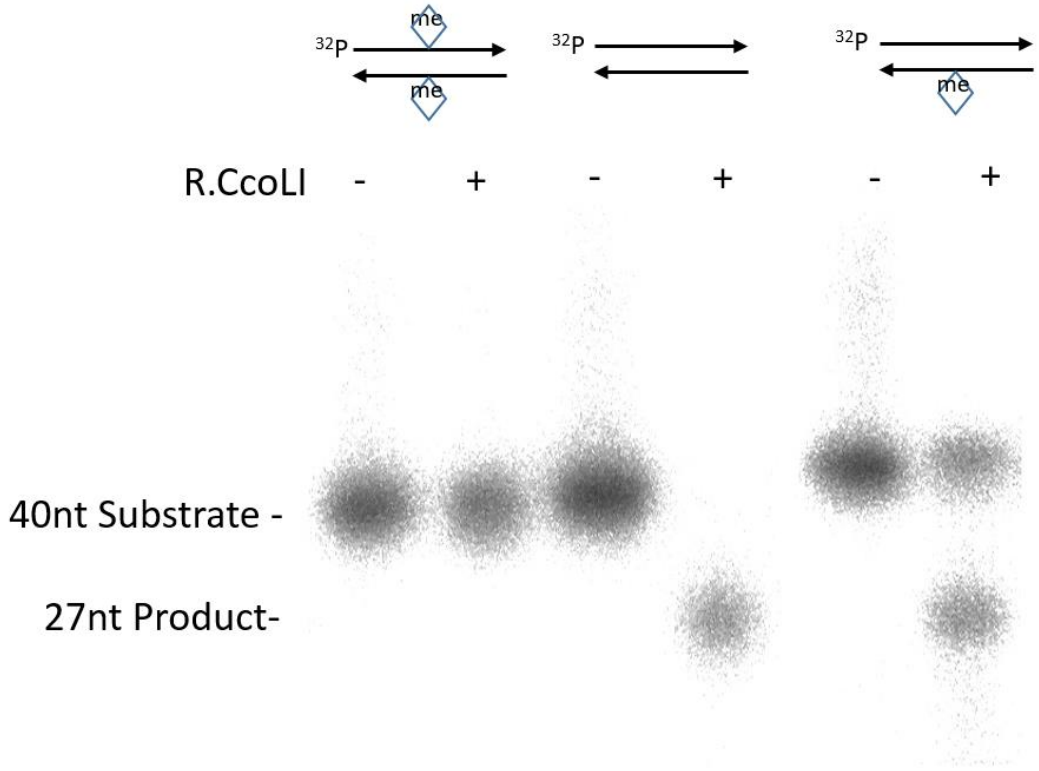

Figure S6. Cleavage of hemi-methylated DNA by R.CcoLI

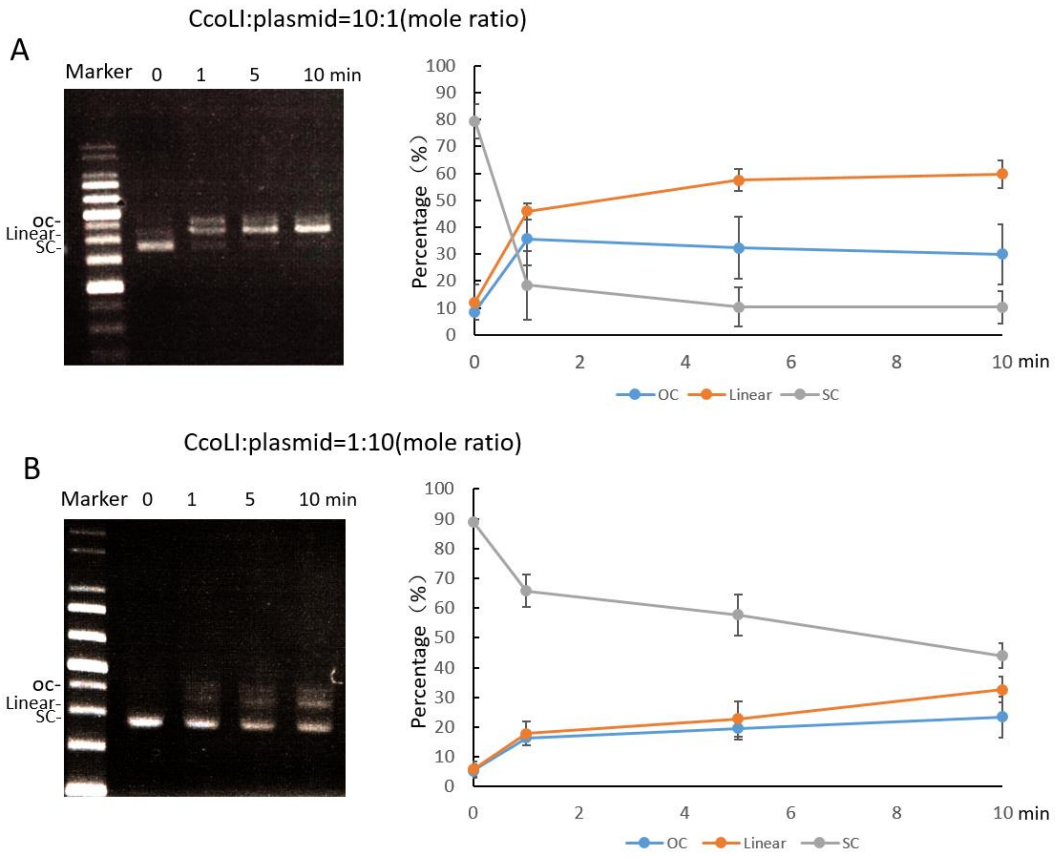

Figure S7. Cleavage of plasmid DNA by R.CcoLI

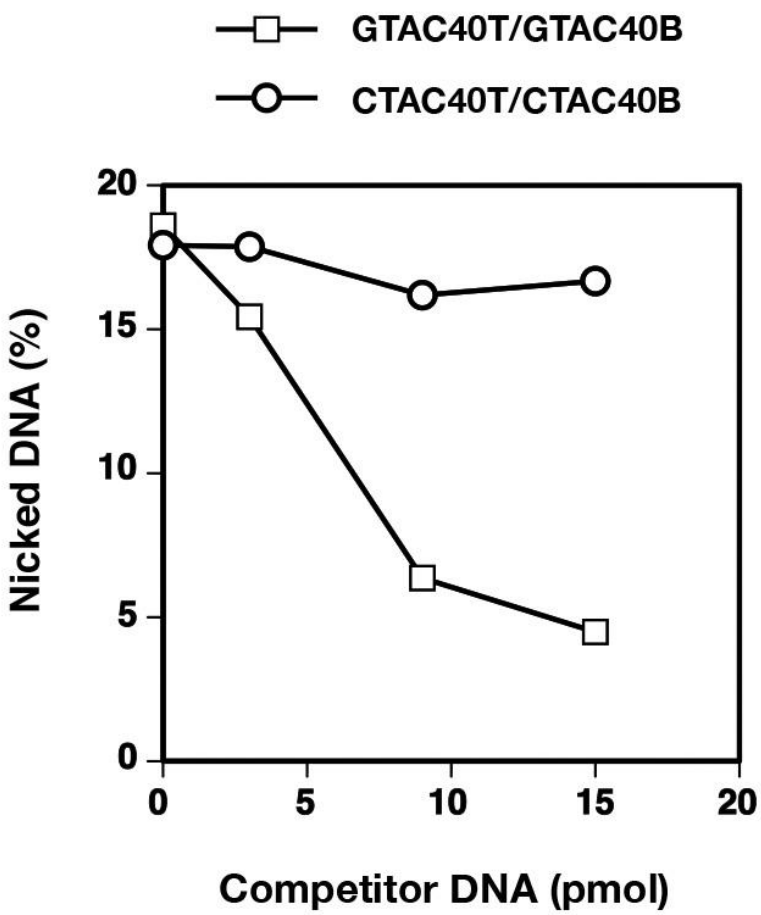

Figure S8. Outcompetition of the AP lyase activity of R.CcoLI by a GTAC-containing duplex.

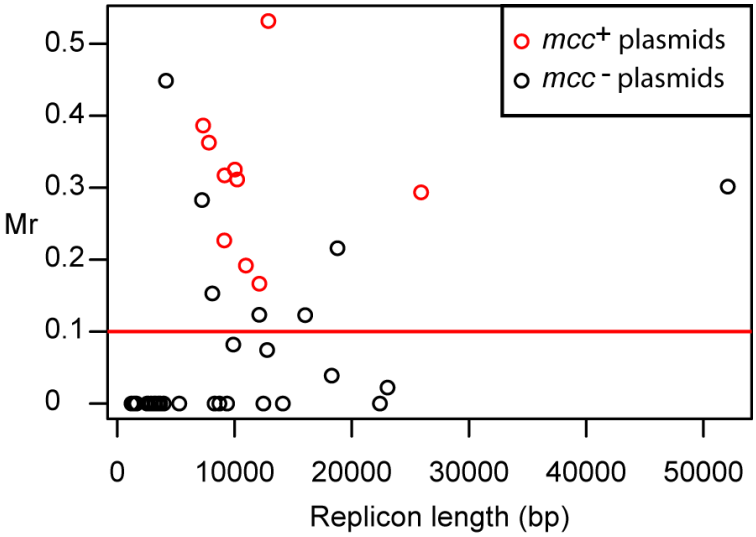

Figure S9. GTAC motif avoidance in *Helicobacter pylori* plasmids (N=41).

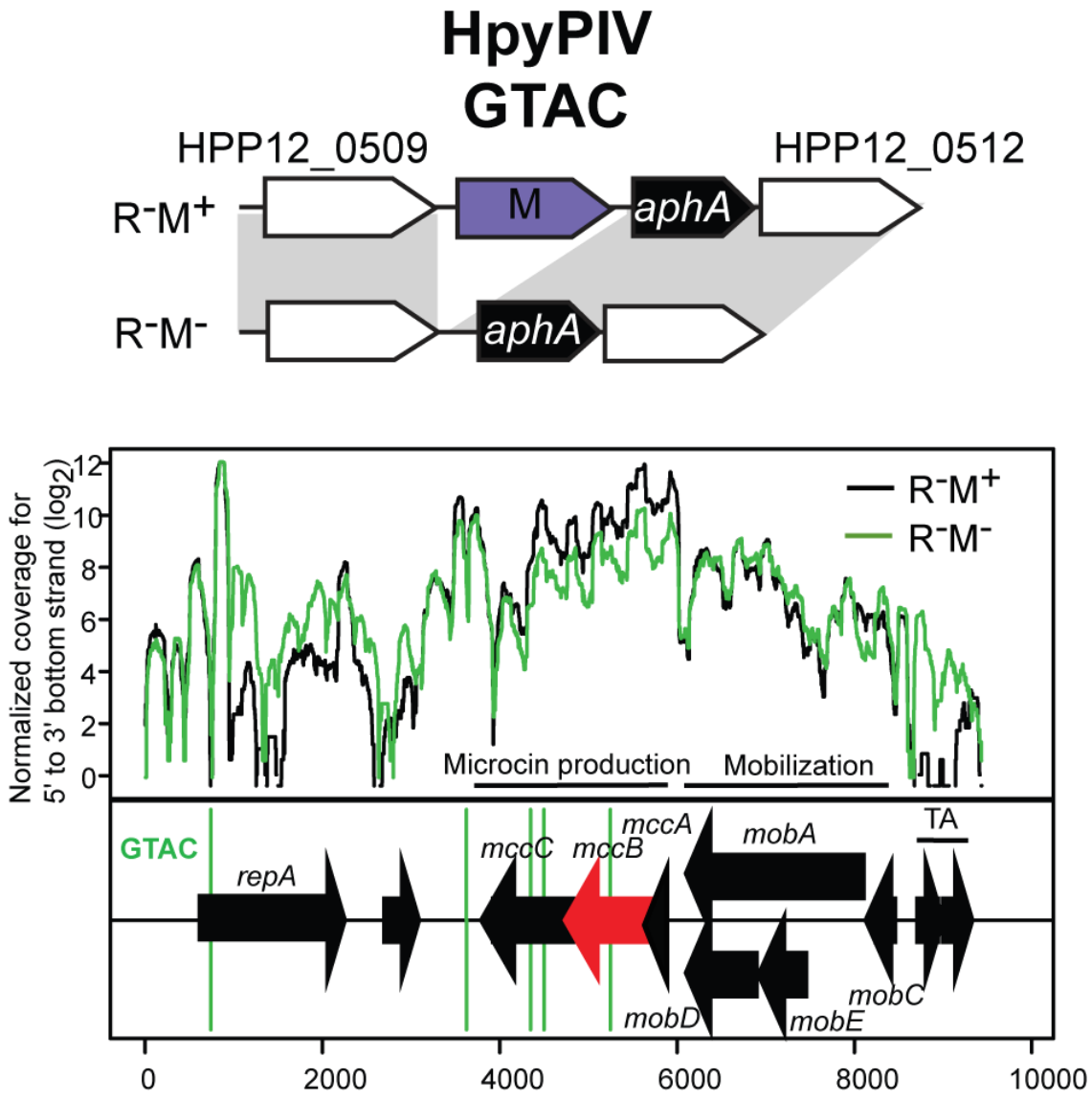

Figure S10. Effect of *hpyPIVM* knockout on *H.pylori* gene expression.

Table S1. *E. coli* strains and plasmids used

	Strains NO.	Strains name	Relevant Properties	Source/ Reference
E. coli	BIK35001	*E. coli* HST08	*F⁻, endA1, supE44, thi-1, recA1, relA1, gyrA96, phoA, Φ80dlacZΔM15, Δ(lacZYA-argF)U169, Δ(mrr-hsdRMS-mcrBC), ΔmcrA, λ⁻*	TaKaRa
	BIK35004	*E. coli* T7 Express *lysY/I^q*	*MiniF lysY lacI^q(CamR) / fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr-73::miniTn10--Tet^S)2 [dcm] R(zgb-210::Tn10--Tet^S) endA1 Δ(mcrC-mrr) 114::IS10*	New England Biolabs
	BIK35016	*E. coli* JW2146	*Δnfo*-786::kan	KEIO collection
	BIK35019	*E. coli* JW1738	*Δxth*::kan	KEIO collection
	BIK35020	*E. coli* JW0221	*dinB*749(del)::*aph*	KEIO collection
	BIK35022	*E. coli* JW0059	*polB*770(del)::*aph*	KEIO collection
	BIK35028	*E. coli* JW2669	*recA*774(del)::*aph*	KEIO collection
	BIK35029	*E. coli* JW2788	*recB*745(del)::*aph*	this work
	BIK35006	*E. coli* T7 Express *lysY/I^q* (M+PabI+)	BIK35004 (pBAD30_*cviQIM*)(pET28a_*pabIR*)	this work
	BIK35009	*E. coli* T7 Express *lysY/I^q* (M+PabID214A+)	BIK35004 (pBAD30_*cviQIM*)(pET28a_*pabIRD214A*)	this work
	BIK35010	*E. coli* T7 Express *lysY/I^q* (M+HpyAXII+)	BIK35004 (pBAD30_*cviQIM*) (pET28a_*hpyAXIIR*)	this work
	BIK35011	*E. coli* T7 Express *lysY/I^q* (M+CcoLI+)	BIK35004 (pBAD30_*cviQIM*) (pET28a_*ccoLIR*)	this work
	BIK35015	*E. coli* T7 Express *lysY/I^q* (M+)	BIK35004 (pBAD30_*cviQIM*) (pET28a)	this work

BIK35036	*nfo- E. coli* T7 Express *lysY/I*^{q}	*Δnfo*-786, P1 Transduction from BIK35016 to BIK35004	this work
BIK35037	*polB- E. coli* T7 Express *lysY/I*^{q}	*Δxth*, P1 Transduction from BIK35019 to BIK35004	this work
BIK35038	*xth- E. coli* T7 Express *lysY/I*^{q}	*ΔdinB*749, P1 Transduction from BIK35020 to BIK35004	this work
BIK35039	*dinB- E. coli* T7 Express *lysY/I*^{q}	*ΔpolB*770, P1 Transduction from BIK35022 to BIK35004	this work
BIK35040	*recA- E. coli* T7 Express *lysY/I*^{q}	*ΔrecA*774, P1 Transduction from BIK35028 to BIK35004	this work
BIK35041	*recB- E. coli* T7 Express *lysY/I*^{q}	*ΔrecB*745,P1 Transduction from BIK35029 to BIK35004	this work
BIK35048	*nfo- E. coli* T7 Express *lysY/I*^{q} (M+PabI+)	BIK35036 (pBAD30_*cviQIM*) (pET28a_*pabIR*)	this work
BIK35049	*polB- E. coli* T7 Express *lysY/I*^{q} (M+PabI+)	BIK35037 (pBAD30_*cviQIM*) (pET28a_*pabIR*)	this work
BIK35050	*xth- E. coli* T7 Express *lysY/I*^{q} (M+PabI+)	BIK35038 (pBAD30_*cviQIM*) (pET28a_*pabIR*)	this work
BIK35051	*dinB- E. coli* T7 Express *lysY/I*^{q} (M+PabI+)	BIK35039 (pBAD30_*cviQIM*) (pET28a_*pabIR*)	this work
BIK35052	*recA- E. coli* T7 Express *lysY/I*^{q} (M+PabI+)	BIK35040 (pBAD30_*cviQIM*) (pET28a_*pabIR*)	this work
BIK35053	*recB- E. coli* T7 Express *lysY/I*^{q} (M+PabI+)	BIK35041 (pBAD30_*cviQIM*) (pET28a_*pabIR*)	this work
BIK35054	*nfo- E. coli* T7 Express *lysY/I*^{q} (M+CcoLI+)	BIK35036 (pBAD30_*cviQIM*) (pET28a_*ccoLIR*)	this work
BIK35055	*polB- E. coli* T7 Express *lysY/I*^{q} (M+CcoLI+)	BIK35037 (pBAD30_*cviQIM*) (pET28a_*ccoLIR*)	this work
BIK35056	*xth- E. coli* T7 Express *lysY/I*^{q}	BIK35038 (pBAD30_*cviQIM*) (pET28a_*ccoLIR*)	this work

		(M+CcoLI+)		
	BIK35057	*dinB- E. coli* T7 Express *lysY/I*^{q} (M+CcoLI+)	BIK35039 (pBAD30_*cviQIM*) (pET28a_*ccoLIR*)	this work
	BIK35058	*recA- E. coli* T7 Express *lysY/I*^{q} (M+CcoLI+)	BIK35040 (pBAD30_*cviQIM*) (pET28a_*ccoLIR*)	this work
	BIK35059	*recB- E. coli* T7 Express *lysY/I*^{q} (M+CcoLI+)	BIK35041 (pBAD30_*cviQIM*) (pET28a_*ccoLIR*)	this work
Plasmids		pBAD30	P_{BAD}, Cml^{R}	[3]
		pBAD30_*cviQIM*	pBAD30 CviQIM+	[3]
		pET28a	P_{T7}, Kan^{R}, *lacI*^{Q}	[3]
		pET28a_*pabIR*	pET28a::*pabIR*	[3]
		pET28a_*pabIRD214A*	pET28a::*pabIR(D214A)*	this work, [5]
		pET28a_*ccoLIR*	pET28a::*ccoLIR*	this work
		pET28a_*hpyAXIIR*	pET28a::*hpyAXIIR*	this work
		pTAKN-2-*ccoLI*	pTAKN-2::*ccoLI*	Funakoshi
		pTAKN-2-*hpyAXII*	pTAKN-2::*hpyAXII*	Funakoshi
		pEU3-NIIb_*ccoLIR*	pEU3-NIIb::*ccoLIR*	this work
		pEU3-NIIb_*hpyAXII*	pEU3-NIIb::*hpyAXII*	this work
		PFLP3	Ap^{R}, Tc^{R}; Source of Flp recombinase	Addgene

Table S2 Oligos and primers used in *E. coli* experiment

Oligonucleotide name	Length (nt/ bp)	sequence	Source
GTAC40T	40	5'GGGGAGGCGCCGGCAGTGCGTCAGGTACTCCGCCACGTCC3'	Hokkaido System Science
GTAC40B	40	5'GGACGTGGCGGAGTACCTGACGCACTGCCGGCGCCTCCCC3'	Hokkaido System Science
GTAC40Tme	40	5'GGGGAGGCGCCGGCAGTGCGTCAGGTmACTCCGCCACGTCC3'	Hokkaido System Science
GTAC40Bme	40	5'GGACGTGGCGGAGTmACCTGACGCACTGCCGGCGCCTCCCC3'	Hokkaido System Science
GTUC40T	40	5‘GGGGAGGCGCCGGCAGTGCGTCAGGTUCTCCGCCACGTCC3’	Hokkaido System Science
GTUC40B	40	5‘GGACGTGGCGGAGTUCCTGACGCACTGCCGGCGCCTCCCC3’	Hokkaido System Science
GT#C40	40	5‘GGGGAGGCGCCGGCAGTGCGTCAGGT#CTCCGCCACGTCC3’ 5‘GGACGTGGCGGAGT#CCTGACGCACTGCCGGCGCCTCCCC3’	This work; GTUC40T/GTUC40B treated with UNG; # = AP site
CTAC40T	40	5'GGGGAGGCGCCGGCAGTGCGTCAGCTACTCCGCCACGTCC3'	Tsukuba Oligo Service

CTAC40B	40	5'GGACGTGGCGGAGTAGCTGACGCACTGCCGGCGCCTCCCC3'	Tsukuba Oligo Service
R.pabI_R	44	5‘TGGTGGTGGTGGTGCTCGAGTTATGAAGTGCCGATAATACTCCT3’	Hokkaido System Science
R.pabI_F	36	5‘CGCGCGGCAGCCATATGATTCATTTGACTAGTGTAG3’	Hokkaido System Science
HpyAXII-F	39	5‘GTGCCGCGCGGCAGCCATATGTCCCTCATTCGCATCGAC3’	Hokkaido System Science
HpyAXII-R	45	5‘GTGGTGGTGGTGGTGCTCGAGTTAATTCTGCAGAATTTTCTCCAG3’	Hokkaido System Science
CcoLI-F	39	5‘GTGCCGCGCGGCAGCCATATGAAATTCAAAATTGACTAT3’	Hokkaido System Science
CcoLI-R	44	5‘GTGGTGGTGGTGGTGCTCGAGTTACTTACTGTTCAGAATGAACT3’	Hokkaido System Science
P-D214A-R	71	5'TGGTGGTGGTGGTGCTCGAGTTATGAAGTGCCGATAATACTCCTCAAAAATTTAACAATA**G**CATTCTTGTG3’	Hokkaido System Science
P-D214A-F	36	5‘CGCGCGGCAGCCATATGATTCATTTGACTAGTGTAG3’	Hokkaido System Science
m-phsg398-F	24	5‘TCAGTTGCTCAATGCACCTATAAC3’	Hokkaido System Science
m-phsg398-R	24	5‘CTCATCGCAGTATTGTTGTAATTC3’	Hokkaido System Science
m-phsg398-F-2	20	5‘GAATTACAACAATACTGCGATGAG3’	Hokkaido System

			Science
m-phsg398-R-2	20	5‘GTCTGGTTATAGGTGCATTGAG3’	Hokkaido System
			Science
pBAD30F	22	5‘ATGAGTATTCAACATTTCCGTG3’	Hokkaido System
			Science
pBAD30R	23	5‘TTACCAATGCTTAATCAGTGAGG3’	Hokkaido System
			Science
CF-CcoLI-F	37	5'ATCCATATGAAATTCAAAATTGACTATGAACTGCCGT3'	Hokkaido System
			Science
CF-CcoLI-R	44	5'CGGGATCCCGTTACTTACTGTTCAGAATGAACTCAATAATCTGA3'	Hokkaido System
			Science
CF-HpyAXII-F	28	5'ATCCATATGTCCCTCATTCGCATCGACA3'	Hokkaido System
			Science
CF-HpyAXII-R	43	5'CGGGATCCCGTTAATTCTGCAGAATTTTCTCCAGAATTTTGAG3'	Hokkaido System
			Science
nfo-F	20	CACTACATCTTGCTCCTGTT	Hokkaido System
			Science
nfo-R	20	CAATTTCGTTCTGCTGAATC	Hokkaido System
			Science
xthA-F	20	GACATCATTAACAACCATCG	Hokkaido System
			Science
xthA-R	20	CAAGGTTAATTCTCCTGACC	Hokkaido System
			Science

polB-F	19	TTACGGGCAGTAATGACTG	Hokkaido System Science
polB-R	22	CCGTGCTTATGAGGTAGTGGTG	Hokkaido System Science
dinB-f	19	GTGGTGCAGCCGCTGGTGC	Hokkaido System Science
dinB-r	20	CAGCGAGAATTCGATGCATG	Hokkaido System Science
recA-F	20	TGCTTCAACAGAACATATTG	Hokkaido System Science
recA-R	20	AGGCGAGCATATGCCGGGCG	Hokkaido System Science
recB-F	20	AGTGTGGGAGAACGTCAGCG	Hokkaido System Science
recB-R	20	AATTGCACATCCAGCGGGCG	Hokkaido System Science

Table S3 Primers used in *Helicobacter pylori* transcriptome analysis

Primers	Sequence	Use	Source
HPP12_0511_Lf	CTCTAGAGGATCCCCTTTAGTCAATCCTAAAAGAGAAAA	Construction of pMZA3, pHY1279	This study
HPP12_0511_Lr	ACACAATATGGCGGATTAATCAAATTCACAAGGGTGTC	Construction of pMZA3,	This study
HPP12_0511_Lr2	CTCAAATGGTTCGCTGGGTTTTAATCAAATTCACAAGGGTGTCC	Construction of pHY1279	This study
catf	TCCGCCATATTGTGTTGAAAC	Construction of pMZA3	This study
catr	GGGCACCAATAACTGCCTTA	Construction of pMZA3	This study
HPP12_0511_Rf	CAGTTATTGGTGCCCTAATGATACGAATTGTTAGAAAGG	Construction of pMZA3	This study
HPP12_0511_Rr	TCGAGCTCGGTACCCTTTTGATAGTCGGCTGGCCAA	Construction of pMZA3, pHY1279	This study
HPP12_0510_Lf	CTCTAGAGGATCCCCTGACTTTTTTCTATCGTTATTTACGCC	Construct ion of pHY1278	This study
HPP12_0510_Lr	CTCAAATGGTTCGCTGGGTTTTACACATCTAGCTGTTTCACGTCTTT	Construction of pHY1278	This study
HPP12_0511_Rf4	ATGAATTGTTTTAGTACCTAGATTTAGATGTCTAAAAATAATGATACGAATTGTTAGAAAGGGAATATC	Construction of pHY1278, pHY1279	This study
AphA3f	TTTTTAGACATCTAAATCTAGGTACTAAAAC	Construction of pHY1278, pHY1279	This study
AphA3r	AACCCAGCGAACCATTTGAG	Construction of pHY1278, pHY1279	This study

Table S4. Genes differentially expressed upon knockout of the methyltransferase gene.

temporal gene ID	product name*	position (left)*	position (right)*	strand	refseq_locus_tag	old_locus-tag	motif around the gene GTAC	M+ R- rep1	M+ R- rep2	M-R- rep1	M-R- rep2	a.value	fold change	p.value	q.value
gene_91	uncharacterized protein	96046	96222	-	HPP12_RS00485	HPP12_0093	N	208.56	187.82	22505.97	21837.98	11.03	6.81	2.14E-97	3.46E-94
gene_517	M.HpyPIV	535439	536431	+	HPP12_RS02620	HPP12_0510	N	1160.39	1395.52	0.00	2.05	5.18	-10.29	2.04E-77	1.65E-74
gene_954	uncharacterized protein	1002619	1002831	-	NA	NA	N	14.05	13.62	818.43	641.15	6.65	5.72	2.22E-45	1.19E-42
gene_1552	hypothetical protein	1624404	1624523	+	NA	NA	N	12.57	6.81	924.03	562.29	6.41	6.26	2.72E-40	1.10E-37
gene_136	(S)-2-hydroxy-acid oxidase domain protein	145905	146108	+	NA	NA	N	7.40	8.76	180.15	161.82	5.22	4.40	1.92E-23	6.20E-21
gene_890	hypothetical protein	929247	929552	-	NA	NA	N	11.83	10.70	184.81	171.04	5.48	3.98	1.32E-21	3.55E-19
gene_1526	**Type I RM system S protein (GRAN7TAYC)**	**1598091**	**1599287**	**-**	**HPP12_RS07730**	**HPP12_1508**	**N**	**149.39**	**146.95**	**841.72**	**1398.03**	**8.67**	**2.92**	**5.01E-19**	**1.16E-16**
gene_1170	uncharacterized protein	1239653	1239901	+	NA	NA	N	14.79	18.49	225.18	161.82	5.83	3.54	5.08E-18	1.02E-15

gene_318	**membrane protein**	**328157**	**330283**	**-**	**NA**	**NA**	**N**	**4749.53**	**3850.82**	**19712.14**	**21870.75**	**13.21**	**2.27**	**2.13E-17**	**3.82E-15**
gene_231	uncharacterized protein	235455	235634	+	NA	NA	N	8.14	10.70	88.52	100.37	4.90	3.33	1.64E-13	2.64E-11
gene_479	hypothetical protein	491929	492708	-	HPP12_RS02425	HPP12_0473	N	27.36	14.60	267.11	147.48	6.04	3.30	5.20E-13	7.63E-11
gene_1123	hypothetical protein	1182613	1183044	+	HPP12_RS05650	HPP12_1109	N	6.66	2.92	23.29	48.14	3.71	2.90	7.12E-07	9.59E-05
gene_883	alkylphosphonate uptake protein	921255	921584	+	HPP12_RS04430	HPP12_0872	N	485.90	464.20	978.39	1401.10	9.55	1.32	2.88E-06	3.41E-04
gene_1608	**microcin C biosynthesis protein**	**(plasmid pHel12) 4851**	**(plasmid pHel12) 5903**	**-**	**NA**	**NA**	**Y**	**4968.44**	**9146.78**	**2253.39**	**2797.09**	**12.04**	**-1.48**	**2.96E-06**	**3.41E-04**
gene_185	hypothetical protein	190315	190686	-	NA	NA	N	19.97	9.73	80.76	46.09	4.94	2.09	2.06E-05	2.22E-03

* Product names and gene positions are based on revised annotations by MiGAP (http://www.migap.org/index.php/en/aboutpipeline) for our resequenced strain P12.