Research in Microbiology
May 2023, Volume 174, Pages 104044 (6p.)
https://doi.org/10.1016/j.resmic.2023.104044
https://archimer.ifremer.fr/doc/00822/93371/

Plasmid pMO1 from *Marinitoga okinawensis*, first non-cryptic plasmid reported within Thermotogota

Lossouarn Julien [1, *], Nesbø Camilla L. [2, 3], Bienvenu Nadege [4], Geslin Claire [4]

[1] Université Paris-Saclay, INRAE, AgroParisTech, Micalis Institute, 78350, Jouy-en-Josas, France
[2] Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada
[3] Biozone, Department of Chemical Engineering and Applied Chemistry and BioZone, University of Toronto, 200 College Street, Toronto, Ontario, Canada, M5S 3E5
[4] Univ Brest, Ifremer, CNRS, Unité Biologie et Ecologie des Ecosystèmes marins Profonds, F-29280 Plouzané, France

* Corresponding author : Julien Lossouarn, email address : julien.lossouarn@inrae.fr

nesbo@ualberta.ca ; nadege.bienvenu@univ-brest.fr ; claire.geslin@univ-brest.fr

Abstract :

Mobile genetic elements (MGEs), such as viruses and plasmids, drive the evolution and adaptation of their cellular hosts from all three domains of life. This includes microorganisms thriving in the most extreme environments, like deep-sea hydrothermal vents. However, our knowledge about MGEs still remains relatively sparse in these abyssal ecosystems. Here we report the isolation, sequencing, assembly, and functional annotation of pMO1, a 28.2 kbp plasmid associated with the reference strain Marinitoga okinawensis. Carrying restriction/modification and chemotaxis protein-encoding genes, pMO1 likely affects its host's phenotype and represents the first non-cryptic plasmid described among the phylum Thermotogota.

Keywords : deep-sea hydrothermal vent, thermophiles, Thermotogota, plasmid

Introduction

Mobile genetic elements (MGEs, collectively referred as the mobilome) impact on prokaryotic communities from all environments on Earth, including the most extreme such as deep-sea hydrothermal vents [1]. Deep-sea archaea from the phylum Euryarcheota were reported to harbor a panel of MGEs, including viruses and plasmids that are prone to interactions [2, 3]. Compared to the plethora of MGEs that have been characterized in Bacteria elsewhere, our knowledge about the deep marine vent bacterial mobilome is still limited [1]. A bacterial lineage broadly represented in hot environments, and deep-sea hydrothermal vents in particular, is the phylum Thermotogota, which is mostly composed of thermophilic, anaerobic, chemo-organotrophic and sulfur-reducing bacteria [4]. Within this phylum, evidence of lysogeny was originally reported in the order *Petrotogales* with the characterization of three *Marinitoga* inducible temperate siphoviruses [5, 6]. MPV1 lysogenises *Marinitoga piezophila* whereas MCV1 and MCV2 infect two *Marinitoga camini* strains. Even if the host organisms come from geographically distant deep-sea hydrothermal sites, MPV1 shares numerous sequence similarities with MCV1 and MCV2. Related proviral regions were recently reported in other species of *Marinitoga* [7, 8] as well as in deep-sea vent *Thermosipho* isolates from the order *Thermotogales* [8]. This may illustrate ancient virus host interactions within these deep marine vent bacteria.

First discovered among these (pro)viruses, MPV1 shares its host with a 13.3 kbp plasmid; pMP1. This plasmid acts as a potential bacteriovirus hijacker, being preferentially packaged into the viral capsids of MPV1 following its lytic cycle induction in *M. piezophila* [5]. pMP1 carries no recognizable genes for functions beside its replication and transfer, and is therefore likely a cryptic plasmid. Only four other cryptic miniplasmids were previously described within Thermotogota and none were associated with deep-sea hydrothermal vents.

These are the negatively supercoiled, and rolling circle replicated, small miniplasmid pRQ7 (846 bp) from *Thermotoga* sp. RQ7 and its very close relatives, pMC24 and pRKU-1, from *Thermotoga maritima* and Thermotoga *petrophila* [9-12], as well as a miniplasmid pTHEBA.01 (1.724 kbp) identified in *Mesotoga prima* [13]. Here we describe the isolation, sequencing, and annotation of a new 28.2 kbp-length plasmid isolated from the deep-sea vent reference strain *Marinitoga okinawensis* [14]. This plasmid, named pMO1, contains 21 coding sequences (CDSs) where 14 were assigned a putative function. Functions not directly associated with replication and transmission were identified, *e.g.* a restriction/modification (RM) system and chemotaxis proteins, making this the first non-cryptic plasmid reported within the phylum Thermotogota.

Material and methods

The circular covalently closed DNA (cccDNA) was isolated following the screening of Thermotogota strains for MGEs described in [5]. pMO1 was identified in the reference strain *M. okinawensis* TFS10-5^T, which was recovered from a deep-sea hydrothermal vent chimney sample located on the Yonaguni Knoll IV field in Southern Okinawa Trough (latitude, 24°50,938'N; longitude, 122°40,020'E; depth, 1365 m [14]). *M. okinawensis* was cultured at 55°C in 100 milliliters modified Ravot medium [15] with minor modifications as described in [5]. cccDNA was extracted from the culture in log growth phase by alkaline lysis method [16]. This extra-chromosomal element was sequenced on an Illumina Miseq sequencer as detailed in [6]. The reads were assembled using the CLC Genomics Workbench 6.5.1 (http://www.clcbio.com) with trimming default settings, and Spades v3.2.1 [17]. Assemblies were compared and consolidated in Geneious R10, which was also used to identify CDSs in combination with prokka 1.11 (VBC | Victorian Bioinformatics Consortium). Homologue of each predicted protein was extracted from the NCBI *nr* protein database using BlastP as in [5]. Matches with the highest identity and percent coverage were retained as best blast hit (BBH), while only considering matches spanning ≥ 50% cover and sharing ≥ 30% identity with queries. Additional analysis was performed on HHPred as described in [8] by only considering biological function predictions with a probability ≥ 98%. CDSs were finally assigned to gene superfamilies/domains using Superfamily and Conserved Domains databases as in [5]. The genomic data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB54960. The assembled and annotated pMO1 plasmid sequence was deposited under the accession number OX370180. pMO1 replication origin ori was predicted using the default parameters of Ori-Finder 2022, which uses repeat sequence analyses, homologous bacterial replication origin (oriC) searches, strand-biased

analyses, and regulatory protein annotation to predict oriC [18]. A map of pMO1 plasmid was drawn using mummer2circos (https://github.com/metagenlab/mummer2circos).

Results and discussion

Sequencing and assembly of cccDNA extracted from *M. okinawensis* resulted in a circular genome of 28.242 bp with an average GC content of 24.4%. The genome sequence has no detectable homologs in nr/nt database and contains 21 predicted CDSs (Figure 1). Fourteen *in silico* translated CDSs had homologues in Genbank and/or HHPred with a putative function, 12 of them could be assigned to gene superfamilies and/or conserved protein domains (Table 1). Based on these predictions, and given that no known structural genes involved in virion formation were identified, we concluded the cccDNA, pMO1, corresponded to a plasmid. pMO1 replication origin ori was predicted to be between position 28007 and 160, the start of the genome was therefore set between the two CDS in this region. The 396 bps-long ori was predicted to contain an AT-rich region and a dnaA box that are features found in many origins of theta-replicating plasmids [19]. A binding site for a factor for inversion stimulation (Fis), frequently found at the replication origin of plasmids using the theta mechanism [19], was also detected in the ori sequence. However, further investigations are needed to determine the specific mode of pMO1 replication.

Nine of the 14 protein-coding pMO1 genes with putative functions were predicted to be involved in DNA interactions (Table 1). PMO1_20 is homologous to ParA proteins and probably involved in plasmid partitioning, whereas PMO1_02, 07, 08, 12 and 19 could ensure diverse DNA regulation activities. PMO1_11 encodes a putative site-specific serine recombinase, which is homologous to *Steptomyces* phage phiC31 integrase following HHPred (Table 1). A CD-search confirmed that PMO1_11 protein had a resolvase/invertase catalytic domain in the N-terminal end, but apparently only a partial recombinase domain in the C-terminal end (Table 1). A possible function of the PMO1_11 encoded site-specific serine

recombinase in plasmid integration (as an integrase) or in its monomeric state maintenance (simply as a resolvase) remains to be investigated.

Interestingly, the best matches of PMO1_12 and 19 are DNA-binding proteins involved in lytic cycle anti-repression of temperate bacterioviruses. Comparisons using HHPred revealed that PMO1_12 is homologous to the Cox multifunctional protein from *E. coli* P2 virus, which acts as an excisionase and repressor of the lysogenic operon, as well as transcriptional activator of the satellite provirus P4 replication [20]. The same HHPred match was obtained for MARPI_RS10485 on pMP1 from *M. piezophila* (Genbank accession no, NC_016748). This may indicate that pMO1 is involved in a molecular piracy relationship similar to that observed for pMP1 and MPV1 [5]. The putative integrative capacity of pMO1, provided by PMO1_11, could be a common feature shared with the pirates pMP1 and plasmid/satellite virus P4, which both encode integrases from the site-specific tyrosine recombinase superfamily. However, we note that this activity has only been demonstrated for P4 to date [5, 20].

Finally, PMO1_04 and 05 were annotated as RM system subunits. RM systems confer prokaryotic immunity against exogenous foreign DNA via two major enzymatic activities [19]. The restriction endonuclease cleaves invading DNA at a specific recognized sequence and the methyltransferase prevents cleavage of host DNA by methylation, ensuring discrimination between self and non-self DNA. PMO1_04 encodes the Type I RM system, restriction subunit R (Table 1) while PMO1_05 putatively encodes a type I RM system methyltransferase subunit (Table 1). Despite their function in prokaryotic defense, RM systems tend also to be seen as mobile selfish elements ensuring their survival by propagating and increasing their frequency in host genomes [21, 22]. Several works have demonstrated that plasmid-encoded RM systems contribute to host cell addiction and in turn plasmid stabilization [21, 22]. This type of association could be advantageous for both RM system and plasmid, facilitating their maintenance and propagation into bacterial hosts exposed to invasive MGEs. Although

methylation has been reported for some *Thermosipho* genomes [23], no evidence has been provided concerning *Marinitoga* genus. Further investigations will be required to clarify if PMO1_04 and 05 indeed confer a line of defense for *M. okinawensis* and at the same time ensure the maintenance of pMO1. BBH for PMO1_04 and 05 were found in bacterial genomes of a *Candidatus latescibacteria* and an *Anaerolinae sp*, respectively. These genomes were obtained from deep-sea hydrothermal sediments [24] and a deep subsurface thermal aquifer [25] metagenome samples. Therefore, these genes might represent horizontal gene transfer (HGT) events between thermophilic bacteria belonging to distinct phyla; evolutionary studies suggest RM genes have indeed sustained extensive HGT between different groups of prokaryotes [22].

Contrary to pMP1, where all proteins with predicted functions and/or superfamily assignments (7/13), likely corresponded to DNA interaction proteins, pMO1 may impact its host phenotype more directly. In addition to the two RM subunit genes, five CDSs likely encode key proteins involved in the chemotaxis signal transduction pathway that allows prokaryotes to sense and navigate according to their chemical environment [26]. PMO1_14 and 15 were annotated as methyl-accepting chemotaxis proteins (MCPs) while PMO1_17 encodes a putative chemotaxis signal transduction protein CheW. MCPs are transmembrane receptors that monitor a concentration of attractant or repellent to direct the cells flagellar locomotion by regulating the histidine kinase CheA [26]. MCP and CheA form cytoplasmic signaling complexes with CheW that couples CheA activity to receptor control [26]. PMO1_13 and 16 were annotated as GGDEF and HD domain-containing diguanylate cyclase and phosphodiesterase, respectively (Table 1). These domains, often in the same protein, control the turnover of C-di-GMP by ensuring its synthesis and hydrolysis respectively [27]. C-di-GMP is a ubiquitous bacterial second messenger that regulates several functions including extracellular polysaccharides production, biofilm formation and motility, and is an important chemotaxis regulator [27]. Data

on chemotaxis remains scarce in *Thermotogota*, being mostly focused on phylogenetic analyses of the chemotaxis machineries across prokaryotes [28]. While *M. okinawensis* has been described as motile with a polar flagellum [14], we currently have no information about *M. okinawensis* genome and its chromosomal chemotaxis genes cluster(s). It is therefore difficult to assess the exact contribution of pMO1 chemotaxis genes and a *M. okinawensis* plasmidic-dependent motility. However, chemotaxis genes have been shown to be frequently acquired by HGT in other *Thermotogota* [13, 29], and finding these genes on a MGE may provide a vehicle for such transfer. In agreement with this, the pMO1 chemotaxis genes are homologous to chromosomal versions found in species of the genus *Marinitoga*, *Kosmotoga* or *Fervidobacterium* (Table 1).

Further analysis will be needed to establish the extent to which pMO1 provides traits (RM or chemotaxis components) to its host and thus participate in bacterial defense or motility mechanisms, as well as if this plasmid could drive HGT within the phylum Thermotogota. Extensive and comparative studies of a growing panel of abyssal MGEs, with their intricate network of interactions, will help us to better understand dynamics of the deep-sea vent prokaryotic communities.

Conflict of interest

The authors declare no conflict of interest.

274

275

276 **References**

277

278 [1] Lossouarn J, Dupont S, Gorlas A, Mercier C, Bienvenu N, Marguet E, et al. An abyssal
279 mobilome: viruses, plasmids and vesicles from deep-sea hydrothermal vents. Res Microbiol
280 2015;166:742–52. https://doi.org/10.1016/j.resmic.2015.04.001.

281

282 [2] Gorlas A, Krupovic M, Forterre P, Geslin C. Living side by side with a virus:
283 characterization of two novel plasmids from *Thermococcus prieurii*, a host for the spindle-
284 shaped virus TPV1. Appl Environ Microbiol 2013;79:3822–8.
285 https://doi.org/10.1128/AEM.00525-13.

286

287 [3] Thiroux S, Dupont S, Nesbø CL, Bienvenu N, Krupovic M, L'Haridon S, et al. The first
288 head-tailed virus, MFTV1, infecting hyperthermophilic methanogenic deep-sea archaea.
289 Environ Microbiol 2021;23:3614–26. https://doi.org/10.1111/1462-2920.15271.

290

291 [4] Pollo SMJ, Zhaxybayeva O, Nesbø CL. Insights into thermoadaptation and the evolution of
292 mesophily from the bacterial phylum Thermotogae. Can J Microbiol 2015;61:655–70.
293 https://doi.org/10.1139/cjm-2015-0073.

294

295 [5] Lossouarn J, Nesbø CL, Mercier C, Zhaxybayeva O, Johnson MS, Charchuck R, et al.
296 "Ménage à trois": a selfish genetic element uses a virus to propagate within *Thermotogales*.
297 Environ Microbiol 2015;17:3278–88. https://doi.org/10.1111/1462-2920.12783.

298

[6] Mercier C, Lossouarn J, Nesbø CL, Haverkamp THA, Baudoux AC, Jebbar M, et al. Two viruses, MCV1 and MCV2, which infect *Marinitoga* bacteria isolated from deep-sea hydrothermal vents: functional and genomic analysis. Environ Microbiol 2018;20:577–87. https://doi.org/10.1111/1462-2920.13967.

[7] L'Haridon S, Gouhier L, John ESt, Reysenbach A-L. *Marinitoga lauensis* sp. nov., a novel deep-sea hydrothermal vent thermophilic anaerobic heterotroph with a prophage. Syst Appl Microbiol 2019;42:343–7. https://doi.org/10.1016/j.syapm.2019.02.006.

[8] Haverkamp THA, Lossouarn J, Zhaxybayeva O, Lyu J, Bienvenu N, Geslin C, et al. Newly identified proviruses in Thermotogota suggest that viruses are the vehicles on the highways of interphylum gene sharing. Environ Microbiol 2021;23:7105–20. https://doi.org/10.1111/1462-2920.15723.

[9] Harriott OT, Huber R, Stetter KO, Betts PW, Noll KM. A cryptic miniplasmid from the hyperthermophilic bacterium *Thermotoga* sp. strain RQ7. J Bacteriol 1994;176:2759–62. https://doi.org/10.1128/jb.176.9.2759-2762.1994.

[10] Yu JS, Noll KM. Plasmid pRQ7 from the hyperthermophilic bacterium *Thermotoga* species strain RQ7 replicates by the rolling-circle mechanism. J Bacteriol 1997;179:7161–4. https://doi.org/10.1128/jb.179.22.7161-7164.1997.

[11] Akimkina T, Ivanov P, Kostrov S, Sokolova T, Bonch-Osmolovskaya E, Firman K, et al. A highly conserved plasmid from the extreme thermophile *Thermotoga maritima* MC24 is a

member of a family of plasmids distributed worldwide. Plasmid 1999;42:236–40. https://doi.org/10.1006/plas.1999.1429.

[12] Nesbø CL, Dlutek M, Doolittle WF. Recombination in *Thermotoga*: implications for species concepts and biogeography. Genetics 2006;172:759–69. https://doi.org/10.1534/genetics.105.049312.

[13] Zhaxybayeva O, Swithers KS, Foght J, Green AG, Bruce D, Detter C, et al. Genome sequence of the mesophilic *Thermotogales* bacterium *Mesotoga prima* MesG1.Ag.4.2 reveals the largest *Thermotogales* genome to date. Genome Biol Evol 2012;4:700–8. https://doi.org/10.1093/gbe/evs059.

[14] Nunoura T, Oida H, Miyazaki M, Suzuki Y, Takai K, Horikoshi K. *Marinitoga okinawensis* sp. nov., a novel thermophilic and anaerobic heterotroph isolated from a deep-sea hydrothermal field, Southern Okinawa Trough. Int J Syst Evol Microbiol 2007;57:467–71. https://doi.org/10.1099/ijs.0.64640-0.

[15] Ravot G, Magot M, Fardeau ML, Patel BK, Prensier G, Egan A, et al. *Thermotoga elfii* sp. nov., a novel thermophilic bacterium from an African oil-producing well. Int J Syst Bacteriol 1995;45:308–14. https://doi.org/10.1099/00207713-45-2-308.

[16] Birnboim HC, Doly J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 1979;7:1513–23. https://doi.org/10.1093/nar/7.6.1513.

[17] Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 2012;19:455–77. https://doi.org/10.1089/cmb.2012.0021.

[18] Dong M-J, Luo H, Gao F. Ori-Finder 2022: A Comprehensive Web Server for Prediction and Analysis of Bacterial Replication Origins. Genomics Proteomics Bioinformatics 2022:S1672-0229(22)00130-9. https://doi.org/10.1016/j.gpb.2022.10.002.

[19] del Solar G, Giraldo R, Ruiz-Echevarría MJ, Espinosa M, Díaz-Orejas R. Replication and control of circular bacterial plasmids. Microbiol Mol Biol Rev 1998;62:434–64. https://doi.org/10.1128/MMBR.62.2.434-464.1998.

[20] Christie GE, Dokland T. Pirates of the *Caudovirales*. Virology 2012;434:210–21. https://doi.org/10.1016/j.virol.2012.10.028.

[21] Vasu K, Nagaraja V. Diverse functions of restriction-modification systems in addition to cellular defense. Microbiol Mol Biol Rev 2013;77:53–72. https://doi.org/10.1128/MMBR.00044-12.

[22] Furuta Y, Kobayashi I. Restriction-Modification Systems as Mobile Epigenetic Elements. In: Robert AP, Mullany P, editors. Bacterial Integrative Mobile Genetic Elements, Austin, TX : Landes Bioscience; 2011, p. 456–67.

[23] Haverkamp THA, Geslin C, Lossouarn J, Podosokorskaya OA, Kublanov I, Nesbø CL. *Thermosipho* spp. Immune System Differences Affect Variation in Genome Size and

372 Geographical Distributions. Genome Biol Evol 2018;10:2853–66.
373 https://doi.org/10.1093/gbe/evy202.
374
375 [24] Dombrowski N, Teske AP, Baker BJ. Expansive microbial metabolic versatility and
376 biodiversity in dynamic Guaymas Basin hydrothermal sediments. Nat Commun 2018;9:4999.
377 https://doi.org/10.1038/s41467-018-07418-0.
378
379 [25] Kadnikov VV, Mardanov AV, Beletsky AV, Banks D, Pimenov NV, Frank YA, et al. A
380 metagenomic window into the 2-km-deep terrestrial subsurface aquifer revealed multiple
381 pathways of organic matter decomposition. FEMS Microbiol Ecol 2018;94.
382 https://doi.org/10.1093/femsec/fiy152.
383
384 [26] Szurmant H, Ordal GW. Diversity in chemotaxis mechanisms among the bacteria and
385 archaea. Microbiol Mol Biol Rev 2004;68:301–19. https://doi.org/10.1128/MMBR.68.2.301-
386 319.2004.
387
388 [27] Römling U, Galperin MY, Gomelsky M. Cyclic di-GMP: the first 25 years of a universal
389 bacterial second messenger. Microbiol Mol Biol Rev 2013;77:1–52.
390 https://doi.org/10.1128/MMBR.00043-12.
391
392 [28] Briegel A, Ortega DR, Huang AN, Oikonomou CM, Gunsalus RP, Jensen GJ. Structural
393 conservation of chemotaxis machinery across Archaea and Bacteria. Environ Microbiol Rep
394 2015;7:414–9. https://doi.org/10.1111/1758-2229.12265.
395

[29] Nesbø CL, Dlutek M, Zhaxybayeva O, Doolittle WF. Evidence for existence of "*mesotogas*," members of the order *Thermotogales* adapted to low-temperature environments. Appl Environ Microbiol 2006;72:5061–8. https://doi.org/10.1128/AEM.00342-06.

421 **Legend to figure and table**

422

423 Figure 1: Map of the pMO1 plasmid. Only informative CDS annotations are shown. The inner
424 blue and green circle shows the GC content variation compared to mean GC, the red and blue
425 circle shows the GC-skew of the plasmid. CDSs on the leading strand are shown in blue and on
426 the lagging strand in red. The figure was produced using mummer2circos
427 (https://github.com/metagenlab/mummer2circos).

428

429 Table 1**:** Functional prediction of pMO1 protein-encoding genes. pMO1 plasmid sequence is
430 available under the accession number OX370180. *translational initiation region where
431 putative manually predicted RBS sequences are underlined and the start codons are in bold.
432 AAA=ATPases associated with diverse cellular activities; HsdR, M, and S= type I restriction-
433 modification system restriction, methylase and specificity subunits; HNH =homing
434 endonuclease with a consensus sequence including two pairs of conserved His and one Asp;
435 DUF=domain of unknown function; _N=N-terminal domain, HTH=helix turn helix domain;
436 GGDEF= conserved domain named after its amino acid motif Gly-Gly-Asp-Glu-Phe;
437 MCP=methyl-accepting chemotaxis protein; MA= methyl-accepting chemotaxis-like domain;
438 HD=conserved domain named after its His and/or Asp (D) residues; c-di-GMP= Bis-(3′-5′)-
439 cyclic dimeric guanosine monophosphate; RpfG=response regulator c-di-GMP
440 phosphodiesterase; Bscq=cellulose biosynthesis protein.

CDS	Lenght (bp); TIR*	Protein_id Predicted function	Protein_id Best blasP match in nr (% identity; % cover)	PDB_id Best HHPred homology (% probability)	Superfamily database *Conserved domain database
PMO1_01	267; AAATTGCAGGTGATAAT**GTG**	CAI4093953.1 Hypothetical protein			
PMO1_02	1533; CTCAAGATGGAGGCGAA**ATG**	CAI4093954.1 ATPase	WP_205100982.1 AAA-like domain-containing protein *Marinitoga litoralis* (81%; 99%)	2QEN Walker-type ATPase *Pyrococcus abyssi* (99.8%)	P-loop containing nucleoside triphosphate hydrolases
PMO1_03	177; AAATGTAACAATAAAAA**ATG**	CAI4093955.1 Hypothetical protein			
PMO1_04	2235; ATATAGGGGGATTAATT**ATG**	CAI4093956.1 Type I restriction modification system, restriction subunit R	RKY72093.1 Restriction endonuclease subunit R *Candidatus latescibacteria bacterium* (74%; 100%)	3H1T Type I restriction-modification system restriction subunit *Vibrio vulnificus* YJ016 (100%)	P-loop containing nucleoside triphosphate hydrolases *HSDR
PMO1_05	2685; TTAAGGAGGAATAATAT**ATG**	CAI4093957.1 Type I restriction modification system, DNA methyltransferase subunit M and specificity subunit S	MBW2979628.1 N-6 DNA methylase *Candidatus Woesearcheota archeon* (63%; 99%)	7EEW Type I restriction restriction-modification system methyltransferase subunit *Vibrio vulnificus* YJ016 (100%)	S-adenosyl-L-methionine-dependent methyltransferases and DNA methylase specificity domain * HSDM and HSDS
PMO1_06	1236; AAAAGAGGTGTTTGCGA**TTG**	CAI4093958.1 Hypothetical protein			
PMO1_07	192; AAATGGAGGTAAATATT**ATG**	CAI4093959.1 HNH endonuclease		1U3E HNH homing endonuclease *Bacillus* virus SPO1 (99.1%)	
PMO1_08	954; AAGGATGGGGTGACTCT**ATG**	CAI4093960.1 Putative endonuclease	WP_244859573.1 DUF4268 domain-containing protein *Tepid anaerobacter acetatoxydans* (68%; 85%)	2VLD Endonuclease NucS *Pyrococcus abyssi* (99.3%)	
PMO1_09	1014; TATAGGAGGATATTATT**ATG**	CAI4093961.1 Hypothetical protein			
PMO1_10	162; AAAACGTAATCTCTGAA**ATG**	CAI4093962.1 Hypothetical protein			
PMO1_11	678; TTTAAGGGGGATTCATT**ATG**	CAI4093963.1 Site-specific serine recombinase	MBU0581305.1 Recombinase family protein *Candidatus Margulisbacteria bacterium* (62%; 97%)	4BQQ Integrase, serine recombinase *Streptomyces* phage phiC31 (100%)	Resolvase-like *Resolvase_N *Recombinase (partial)
PMO1_12	258; AGGGAGTTGGCATTAAG**ATG**	CAI4093964.1 DNA-binding protein	WP_246051367.1 Helix-turn-helix domain-containing protein *Balnearium lithotrophicum* (41%; 64%)	4LHF Regulatory protein Cox Enterobacteria phage P2 (98.8%)	Putative DNA-binding domain *HTH
PMO1_13	1608; TTAGTGGAGAATTAAAT**ATG**	CAI4093965.1 GGDEF domain–containing diguanylate cyclase	WP_011993259.1 GGDEF domain-containing protein *Fervidobacterium nodosum* (51%; 96%)	3EZU GGDEF domain-containing protein *Geobacter sulfurreducens* (99.5%)	Nucleotidyl_cyclase *GGDEF
PMO1_14	1578; AAAAAGGGGGGAGTTCA**GTG**	CAI4093966.1 Methyl-accepting chemotaxis protein	WP_014296321.1 HAMP domain-containing methyl-accepting chemotaxis protein *Marinitoga piezophila* KA3 (54%; 82%)	6S1K Methyl-accepting chemotaxis protein *Escherichia coli* (99.9%)	MCP signaling domain *MA
PMO1_15	1980; ATAAGGGGGAAGGATAT**ATG**	CAI4093967.1 Methyl-accepting chemotaxis protein	WP_047266480.1 Methyl-accepting chemotaxis protein *Marinitoga* sp1155 (76%; 100%)	6S1K Methyl-accepting chemotaxis protein *Escherichia coli* (100%)	MCP signaling domain *MA
PMO1_16	1788; TATTAGGAGTTGAAAAA**ATG**	CAI4093968.1 HD domain-containing c-di-GMP phosphodiesterase	WP_063728003.1 HD domain-containing protein *Kosmotoga* sp DU53 (41%; 63%)	4MCW c-di-GMP phosphodiesterase *Persephonella marina* EX-H1 (99.9%)	HD-domain/PDEases-like *RpfG family
PMO1_17	1308; TGAGTTAGGTTGGTGAT**ATG**	CAI4093969.1 Chemotaxis signal transduction protein CheW	WP_047266840.1 Chemotaxis protein CheW *Marinitoga* sp1155 (63%; 99%)	2QDL Chemotaxis signal transduction protein *Thermoanaerobacter tengcongensis* (99.6%)	CheW_like * CheW
PMO1_18	852; GTTAAGGAGGGGATAAT**ATG**	CAI4093970.1 Hypothetical protein	WP_205097826.1 DUF4382 domain-containing protein *Marinitoga litoralis* (60%; 100%)		Carboxypeptidase regulatory domain-like
PMO1_19	297; AAAGAGATGGAGATATA**ATG**	CAI4093971.1 DNA-binding protein	MCI1750911.1 Phage antirepressor KilAC domain-containing protein *Megasphaera cerevisiae* (67%; 95%)		*BRO_N
PMO1_20	810; AAACAGGAGGGATAAAT**ATG**	CAI4093972.1 Plasmid partitioning protein ParA	MBP6820249.1 ParA family protein *Acidobacteria* bacterium (33%; 97%)	3EZ2 Plasmid partition protein ParA *Escherichia coli* (99.9%)	P-loop containing nucleoside triphosphate hydrolases *Bcsq/ParA_family
PMO1_21	1368; AGAAGGTGGCGTATAAA**ATG**	CAI4093973.1 Hypothetical protein