Supplemental material

Complete sequence analysis of two methanotroph-specific *repABC* plasmids from *Methylocystis* sp. strain SC2

Bomba Dam[1], Michael Kube[2, #], Somasri Dam[1], Richard Reinhardt[2] and
Werner Liesack[1,*]

Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Str. 10,
D-35043 Marburg, Germany[1]
Max Planck Institute for Molecular Genetics, Ihnestrasse 63,
D-14195 Berlin, Germany[2]

*Corresponding author
Mailing address: Max-Planck-Institut für terrestrische Mikrobiologie
Karl-von-Frisch-Str. 10, D-35043 Marburg, Germany
Phone: +49 (6421) 178 720 // Fax: +49 (6421) 178 809
Electronic mail address: liesack@mpi-marburg.mpg.de

[#]Present address:
Faculty of Agriculture and Horticulture, Humboldt-University of Berlin,
Lentzeallee 55/57, D-14195 Berlin, Germany

CONTENT:

Supplemental Methods

Strategy to construct singleton *pmoC* deletion mutants.

Attempts were made to create in-frame deletion mutants of all three singleton *pmoC* present in strain SC2. Primers used are listed in the Supplemental Material Table S1 and a schematic representation of the method is provided in the Supplemental Material Fig. S2. Briefly, PCR amplicons containing in-frame deletions were generated by fusion PCR and cloned in pJQ200SK. For production of these amplicons, four primers (A, B, C, and D) were designed to PCR-amplify an ~1,260-bp fragment carrying a *pmoC* in-frame deletion, using strain SC2 genomic DNA or the plasmids pBSC2-1 and pBSC2-2 as a template. Primers A and B were used to amplify the upstream flanking region of the singleton. Primer A contained a restriction site for cloning in pJQ200SK, and primer B contained a region complementary to the downstream flanking PCR fragment. Primers C and D were used to amplify the downstream flanking fragment of the singleton. Primer C contained a region complementary to the upstream flanking PCR fragment, and primer D contained a restriction site for cloning in pJQ200SK. The two flanking PCR fragments AB and CD were used as templates to generate the full-length in-frame deletion fragment in a second-round (fusion) PCR with primers A and D (see Supplementary Fig. S2 and Table S1 for the strategy to incorporate overlapping regions during PCR with primers B and C). The in-frame fragments were cloned in pJQ200SK and transformed into *E. coli* Top 10 [F^{-} *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *recA1 ara*Δ139 Δ(*ara-leu*)7697 *galU galK rpsL endA1 nupG*] and checked by sequencing. A kanamycin resistance gene was then inserted into all three constructs, as strain SC2 was found to be resistant to gentamicin (present in pJQ200SK). The Kan^{r} cassette was PCR amplified from pCR2.1-TOPO vector (Invitrogen) using primers that introduced a BamHI site at both termini. In pJQC1 and pJQC3, the cassette was cloned in the BglII site present within the gentamicin resistance gene in pJQ200SK.

In pJQC2, a BamHI site is present in the inserted AD PCR product; so the marker was cloned into a unique PstI site present at the extreme 3’ end of the AD PCR product. To achieve this, the marker was PCR amplified from a different source (pBJ114), because the Kan^r cassette amplified from pCR2.1-TOPO contains an internal PstI site. Note that the P15A origin of replication present in pJQ200SK will not be functional in strain SC2. Thus, the prepared constructs would function as suicide plasmids. They were introduced into strain SC2 by electroporation (using a 2-mm gap cuvette and Gene Pulser (Bio-Rad) set at 2.5 kV, 400 Ω, and 25 μF) and recombinants were selected on kanamycin-containing AMS plates (described in the main text). The Kan^r transformants from the single-crossover event also contained the *sacB* gene encoding levansucrase, which is lethal for many gram-negative bacteria when they are grown in presence of sucrose. This genotype allowed for a subsequent negative selection step, i.e., selection for the loss of integrated plasmid containing *sacB* during second recombination. Randomly selected first recombinants were then grown in AMS containing 5% sucrose for 2 days and spread on AMS plates containing 5% sucrose to select colonies that have lost the integrated plasmid through a second homologous crossover event. The recombinants should carry either the wild-type allele (reconstitution) or an in-frame deletion of the singleton *pmoC* (gene replacement). To check the second recombination event, at least 500 colonies (for each singleton knock-out experiment) were screened by colony PCR with primers E and F, which respectively bind upstream and downstream of primers A and D (see Supplementary Fig. S2).

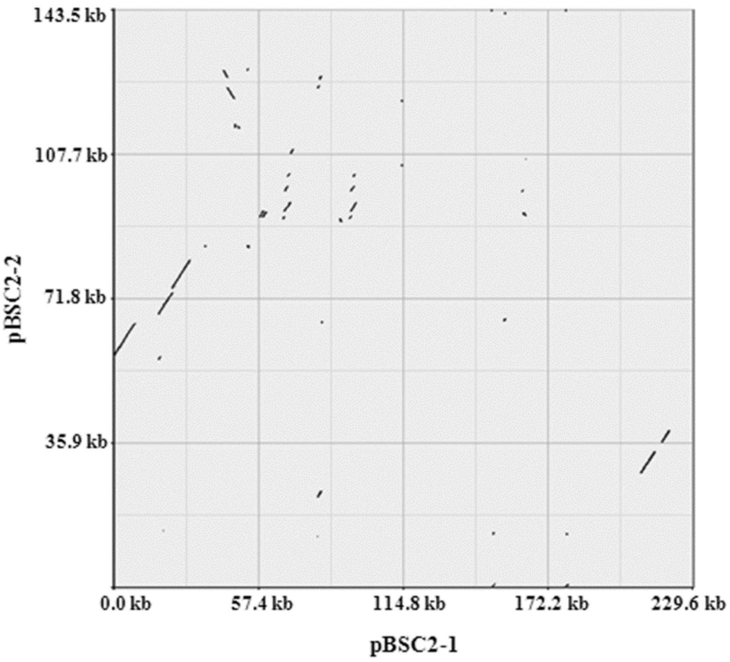

Figure S1. Similarity dot plot of plasmid sequences, pBSC2-1 vs pBSC2-2. The online tool zPicture (http://zpicture.dcode.org/), which aligns two sequences using BLAST engine for local alignment, was used to generate the dot plot. The pBSC2-1 and pBSC2-2 sequences are represented on the X- and Y-axis, respectively, and the numbers represent the base positions (note that the scale is different). Alignments are shown in the plot as lines. Plus strand matches are slanted from the bottom left to the upper right corner, minus strand matches are slanted from the upper left to the lower right. The number of lines shown in the plot is the same as the number of alignments found by BLAST. The plasmids have only a few conserved regions. Both plasmids share the same conserved replication module consisting of the *repABC* genes. Although both plasmids contain complete type IV secretion systems, they are not similar. Other common regions include genes encoding single-strand DNA binding protein, nitric oxide reductase and several small stretches of transposase and conserved hypothetical proteins.

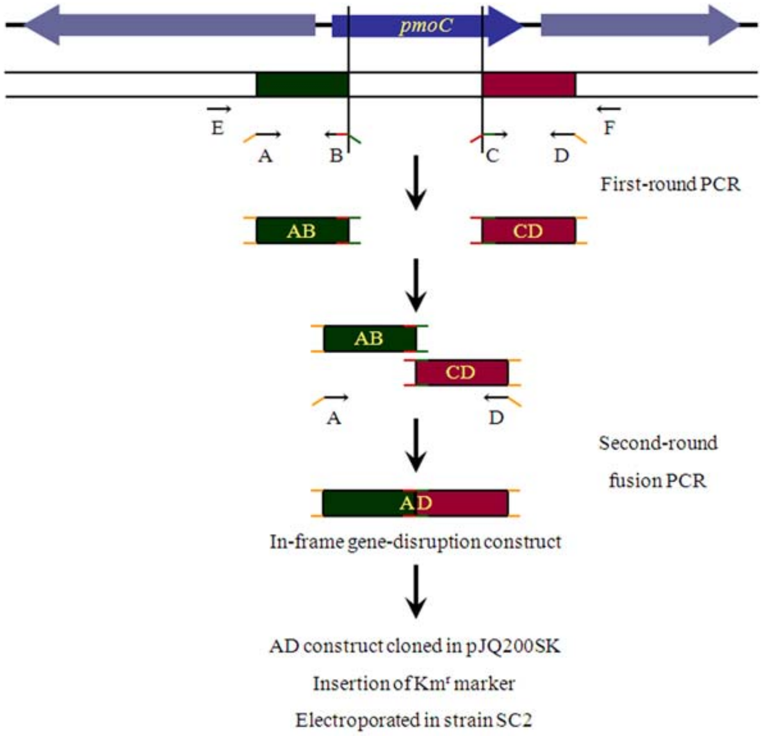

Figure S2. PCR-based strategy to construct suicide vectors for singleton *pmoC* in-frame deletion (see Supplementary Methods and Table S1for details)**.**

Table S1. Oligonucleotide primers and plasmids used in this study

Name of primer/ plasmid	Oligonucleotide sequence (5’ to 3’)^{a} Plasmid features	Restriction site	Source/ reference
Primers used to generate gene disruption constructs			
C1G_A	CGCGGATCCCGGCATCATGAAGATCAATCC	BamHI	-
C1G_B	CTGAACGGGAAGGAAGGCGTCGCCCTC	-	-
C1G_C	GCCTTCCTTCCCGTTCAGAACGGCCAG	-	-
C1G_D	CCGCTCGAGCGGCTATTCCCTCGGACGT	XhoI	-
C2G_A	TGCTCTAGATCTTTCTGACGCAGAGGCG	XbaI	-
C2G_B	GCCCGTTCCGATGCGCTGGAACAGCTCT	-	-
C2G_C	CAGCGCATCGGAACGGGCCTCTGCTTT	-	-
C2G_D	ACGCGTCGACACCTAAGTTCGGCGCTTGG	SalI	-
C3P_A	TGCTCTAGACCCTCAACCCAATCGACCTA	XbaI	-
C3P_B	AACGCCTTCGCCTACCGCGGATTGCGTA	-	-
C3P_C	GCGGTAGGCGAAGGCGTTGAATTGCTGACG	-	-
C3P_D	CGCGGATCCCCCGCACGAGCGCCAC	BamHI	-
C1G_E	CCGGGCGGAACCGCGG	-	-
C1G_F	CCAGCAATGCACCCGTGTGG	-	-
C2G_E	GCGGTCGTCGCCGATCTCG	-	-
C2G_F	GGCGCGAAGCGGCGTCG	-	-
C3P_E	CGTGATACGAGCGCTGTAC	-	-
C3P_F	CAGGAACTCTTCGCCGAC	-	-

KanF_Bam	GACGCGGATCCGATTTCGGCCTATTGGTTA	BamHI	-
KanR_Bam	GACGCGGATCCACCCAACTGATCTTCAGCA	BamHI	-
KanF_Pst	AAACTGCAGGGGTGCACGAGTGGGTTAC	PstI	-
KanR_Pst	AAACTGCAGGGGCGTCGCTTGGTCGGT	PstI	-
Primers used to generate probes for Southern hybridization and to check plasmid loss			
pmo2F[b]	CAACGCTTCGAAACGGAAAGCG	-	-
pmo2R[b]	CCTTGACCTGCCAATATCGGA	-	-
Pl-F[c]	CAACCCGTTGTTGATGACAG	-	-
Pl-R[c]	GCGCTCATCCATTTTCTGAC	-	-
P2-F[d]	GTGCATCTGGCCTGGTATCT	-	-
P2-R[d]	CGAAACGTGAAACCCAGTCT	-	-
Plasmids			
pJQ200SK	Suicide vector; P15A *ori sacB* Gm^{r}	-	ATCC
pCR2.1-TOPO	Source of Km^{r} gene casette	-	Invitrogen
pJQC1	1,260-bp fusion PCR obtained with C1G_A/C1G_D and cloned into BamHI-XhoI site of pJQ200SK	-	This study
pJQC2	1,260-bp fusion PCR obtained with C2G_A/C2G_D and cloned into XbaI-SalI site of pJQ200SK	-	This study
pJQC3	1,260-bp fusion PCR obtained with C3P_A/C3P_D and cloned into XbaI-BamHI site of pJQ200SK	-	This study
pJQC1-Km	BamHI-cut Km^{r}-gene inserted into BglII-cut pJQC1	-	This study
pJQC2-Km	PstI-cut Km^{r}-gene inserted into PstI-cut pJQC2	-	This study
pJQC3-Km	BamHI-cut Km^{r}-gene inserted into BglII-cut pJQC3	-	This study

[a] The green- and red-colored bases in primers B and C represent the complementary regions that were used for performing the second-round fusion PCR (see Supplementary Methods and Fig. S2 for details). Bases highlighted with yellow represent the restriction sites present in the respective primer (see column 3).

[b] The primer pair pmo2F and pmo2R amplifies a 1629-bp product from the *pmoCAB2* operon present in the genome (BX649604).

[c] The primer pair Pl-F and Pl-R amplifies a 1074-bp product from *mobD-mobE* region of pBSC2-1 (P1_200 to P1_201).

[d] The primer pair P2-F and P2-R amplifies an 1168-bp product from *parA-parB* region of pBSC2-2 (P2_23 to P2_24).

Table S2. The ORFs present in pBSC2-1 and their annotations.

ORF identifier	Begin cds	End cds	Size (aa)	First codon	% G+C content	Product	Gene Name	Organism of protein to which most similar	Homolog protein	Similarity (*e*- value)
P1_1	40	2100	687	ATG	62.98	Putative DNA-binding protein	-	*Methylocystis* sp. strain Rockwell	ParB domain protein nuclease	65% identical over 665 aa (0.0)
P1_2	2278	2685	136	ATG	60.05	Conserved hypothetical protein	-	*Sphingomonas wittichii*	Uncharacterized protein	50% identical over 127 aa (5e-33)
P1_3	2833	7278	1482	ATG	64.44	Putative methylase/helicase	-	*Methylocystis* sp. strain Rockwell	Putative methylase/helicase	86% identical over 1471 aa (0.0)
P1_4	7298	8293	332	ATG	59.64	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_3583	69% identical over 323 aa (6e-150)
P1_5	8293	8925	211	ATG	60.35	Conserved hypothetical protein	-	*Sinorhizobium meliloti*	Uncharacterized protein	71% identical over 45 aa (2e-12)
P1_6	10573	9005	523	ATG	59.27	Conserved hypothetical protein	-	*Acinetobacter* sp. SH024	Predicted protein	29% identical over 248 aa (1e-23)
P1_7	13047	10582	822	ATG	64.44	Conserved hypothetical protein	-	*Oligotropha carboxidovorans* OM5	Hypothetical protein OCA5_pOC16701600	78% identical over 822 aa (0.0)
P1_8	14051	13068	238	ATG	62.3	AAA ATPase	-	*Oligotropha carboxidovorans* ATCC 49405	AAA ATPase central domain-containing protein	81% identical over 328 aa (0.0)
P1_9	16331	14898	478	ATG	62.48	Conserved hypothetical protein	-	*Rhodobacter sphaeroides* BisA53	Hypothetical protein RPE_2955	58% identical over 476 aa (0.0)
P1_10	17477	16419	353	ATG	61.1	Fic family protein (filamentation induced by cAMP)	-	*Polymorphum gilvum* SL003B-26A1	Bacterial regulatory protein, AsnC family	83% identical over 351 aa (0.0)

P1_11	18050	18985	312	ATG	60.79	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Protein of unknown function DUF2493	85% identical over 311 aa (0.0)
P1_12	19169	20374	402	GTG	59.78	Plasmid partitioning protein, RepA	-	*Rhizobium leguminosarum* bv. *viciae* 3841	Replication protein, RepA	73% identical over 388 aa (0.0)
P1_13	20380	21375	332	ATG	62.25	Plasmid partitioning protein, RepB	-	*Nitrobacter hamburgensis* X14	ParB family protein	53% identical over 326 aa (6e-109)
P1_14	21602	22969	439	ATG	64.54	Replication protein C	-	*Methylocystis* sp. strain Rockwell	Replication protein C	71% identical over 439 aa (0.0)
P1_15	23002	23256	85	ATG	60.39	Hypothetical protein	-	No significant homology	-	-
P1_16	24510	23272	413	ATG	65.86	Phage integrase	-	*Methylocystis* sp. strain Rockwell	Integrase domain protein SAM domain protein	71% identical over 387 aa (0.0)
P1_17	24603	26216	538	ATG	66.42	Transposase	-	*Mesorhizobium loti* MAFF303099	Transposase	47% identical over 488 aa (1e-140)
P1_18	26212	27090	293	ATG	64.85	TniB transposase	-	*Rhodobacter sphaeroides* ATCC 17025	Hypothetical protein Rsph17025_4235	67% identical over 293 aa (2e-136)
P1_19	27090	28130	347	TTG	69.45	Conserved hypothetical protein		*Rhodobacter sphaeroides* ATCC 17025	Hypothetical protein Rsph17025_4234	45% identical over 319 aa (3e-71)
P1_20	28199	28402	68	ATG	58.82	Hypothetical protein	-	No significant homology	-	-
P1_21	28456	29514	353	ATG	68.46	Conserved hypothetical protein		*Methylocystis* sp. strain Rockwell	Protein of unknown function DUF1403	52% identical over 353 aa (6e-77)
P1_22	29641	30345	235	ATG	65.96	Putative transcriptional regulator		*Nitrobacter hamburgensis* X14	Putative transcriptional regulator	60% identical over 201 aa (1e-76)
P1_23	31009	30536	158	ATG	65.4	Conserved hypothetical protein		*Methylocella silvestris* BL2	Hypothetical protein Msil_0068	41% identical over 78 aa (7e-04)
P1_24	31381	31091	97	ATG	65.64	Histidine kinase		*Trichophyton rubrum* CBS	Histidine kinase	36% identical over

								118892		57 aa (0.58)
P1_25	31567	31442	42	ATG	52.38	Hypothetical protein	-	No significant homology	-	-
P1_26	31650	31841	64	TTG	59.9	Tryptophanyl-tRNA synthetase		*Burkholderia gladioli* BSR3	Tryptophanyl-tRNA synthetase II	40% identical over 56 aa (0.35)
P1_27	32003	32356	118	GTG	58.19	NAD-dependent epimerase/dehydratase		*Nitrosococcus watsoni* C-113	NAD-dependent epimerase/dehydratase	43% identical over 111 aa (4e-15)
P1_28	33109	32852	86	GTG	60.47	Transposase		*Methylosinus trichosporium* OB3b	Transposase	62% identical over 50 aa (2e-14)
P1_29	33689	33126	188	ATG	60.46	Transposase		*Octadecabacter antarcticus* 238	Transposase	53% identical over 106 aa (8e-21)
P1_30	34180	34755	192	ATG	60.76	Cytochrome C family protein, di-heme cytochrome C (Class I)		*Oceanicola batsensis* HTCC2597	Di-heme cytochrome C (Class I)	53% identical over 174 aa (6e-66)
P1_31	34764	35210	149	ATG	56.82	Conserved hypothetical protein		*Pseudomonas stutzeri* A1501	Hypothetical protein PST_3374	75% identical over 138 aa (2e-74)
P1_32	35258	36277	340	ATG	59.51	V-type H(+)-translocating pyrophosphatase	-	*Spirosoma linguale* DSM 74	V-type H(+)-translocating pyrophosphatase	42% identical over 162 aa (1e-25)
P1_33	36619	36395	75	GTG	61.33	Transposase	-	*Acetobacter pasteurianus* IFO 3283-01	Transposase	89% identical over 38 aa (2e-15)
P1_34	36845	36648	66	ATG	59.09	Conserved hypothetical protein	-	*Desulfotomaculum acetoxidans* DSM 771	Conserved hypothetical protein	55% identical over 42 aa (2e-08)
P1_35	36938	37630	231	ATG	65.51	Transposase, mutator type	-	*Xanthobacter autotrophicus* Py2	Transposase mutator type	82% identical over 231 aa (3e-138)
P1_36	37690	38205	172	ATG	63.37	Transposase, mutator type	-	*Rhodopseudomonas palustris* BisA53	Transposase, mutator type	88% identical over 129 aa (1e-76)
P1_37	39514	38351	388	ATG	63.75	Transposase, IS116/IS110/IS902	-	*Nitrobacter hamburgensis*	Transposase, IS116/IS110/	62% identical over

						family protein		X14	IS902 family protein	380 aa (3e-164)
P1_38	39823	40275	151	ATG	59.6	Transposase, IS116/IS110/IS902 family protein	-	*Ochrobactrum anthropi* ATCC 49188	Transposase, IS116/IS110/ IS902 family protein	79% identical over 149 aa (9e-80)
P1_39	41475	40639	279	ATG	57.11	Conserved hypothetical protein	-	*Methylobacterium populi* BJ001	Hypothetical protein Mpop_2534	62% identical over 254 aa (1e-110)
P1_40	42835	41489	449	ATG	56.27	Conserved hypothetical protein	-	*Methylobacterium populi* BJ001	Hypothetical protein Mpop_2535	70% identical over 441 aa (0.0)
P1_41	43327	43049	93	TTG	57.35	Transcriptional regulator, LysR family protein	-	*Roseovarius* sp. TM1035	Transcriptional regulator, LysR family protein	60% identical over 70 aa (4e-18)
P1_42	43339	43536	66	GTG	61.11	Hypothetical protein	-	No significant homology	-	-
P1_43	43645	44853	403	ATG	65.43	NnrS protein	-	*Nitrobacter* sp. Nb-311A	NnrS family protein	47% identical over 393 aa (5e-98)
P1_44	45296	44901	132	ATG	60.35	Invasion gene expression up-regulator, SirB	-	*Oceanobacter* sp. RED65	Invasion gene expression up-regulator, SirB	34% identical over 131 aa (4e-17)
P1_45	47615	45312	768	GTG	62.54	Nitric-oxide reductase	-	*Parvibaculum lavamentivorans* DS-1	Putative nitric-oxide reductase	74% identical over 767 aa (0.0)
P1_46	47778	48065	96	GTG	55.56	Alanine racemase	-	*Treponema pallidum*	Alanine racemase	40% identical over 44 aa (0.025)
P1_47	48458	48649	64	ATG	66.67	Putative transposase IS1312, IS3 family	-	*Magnetospirillum gryphiswaldense* MSR-1	Putative transposase IS1312, IS3 family	71% identical over 59 aa (1e-23)
P1_48	48924	49355	144	ATG	60.65	Transposase (integrase)	-	*Oligotropha carboxidovorans* OM5	Integrase	72% identical over 141 aa (2e-67)
P1_49	49785	49366	140	ATG	56.19	GTP cyclohydrolase I	*folE2*	*Methylocystis* sp. strain Rockwell	GTP cyclohydrolase I	86% identical over 101 aa (9e-59)

P1_50	50183	52876	898	ATG	59.69	Copper-translocating P-type ATPase	-	*Nitrosococcus halophilus* Nc4	Copper-translocating P-type ATPase	68% identical over 686 aa (0.0)
P1_51	53410	53051	120	GTG	54.44	Transposase	-	*Sinorhizobium meliloti*	Hypothetical transposase	64% identical over 118 aa (2e-45)
P1_52	55105	53990	372	TTG	63.17	Putative ABC transporter ATP binding protein/permease	-	*Bradyrhizobium* sp. BTAi1	Putative ABC transporter ATP binding protein/permease	40% identical over 333 aa (7e-75)
P1_53	55462	55211	84	GTG	57.14	RNA polymerase, sigma-24 subunit, ECF subfamily	-	*Methylocystis* sp. strain Rockwell	RNA polymerase, sigma-24 subunit, ECF subfamily	88% identical over 33 aa (6e-12)
P1_54	56337	55603	245	ATG	61.5	Conserved hypothetical protein	-	*Methylosinus trichosporium* OB3b	Conserved hypothetical protein	39% identical over 198 aa (3e-23)
P1_55	56699	56496	68	ATG	63.73	Copper-translocating P-type ATPase	-	*Cyclobacterium marinum* DSM 745	Copper-translocating P-type ATPase	77% identical over 26 aa (2e-06)
P1_56	57327	57872	182	ATG	59.34	Glucan 1,4-alpha-glucosidase	-	*Methylocystis* sp. strain Rockwell	Glucan 1,4-alpha-glucosidase	62% identical over 181 aa (3e-74)
P1_57	58018	58716	233	ATG	63.95	Transposase, IS116/IS110/ IS902 family protein	-	*Paracoccus denitrificans* PD1222	Transposase, IS116/IS110/ IS902 family protein	67% identical over 221 aa (3e-103)
P1_58	58706	59029	108	ATG	61.73	Transposase of ISMdi12, IS110 family	-	*Methylobacterium extorquens* DM4	Transposase of ISMdi12, IS110 family	71% identical over 103 aa (3e-46)
P1_59	59301	60392	364	ATG	58.24	Transposase, IS116/IS110/ IS902 family protein	-	*Nitrobacter hamburgensis* X14	Transposase, IS116/IS110/ IS902 family protein	72% identical over 343 aa (0.0)
P1_60	61104	60856	83	ATG	56.22	Hypothetical protein	-	No significant homology	-	-
P1_61	61694	61254	147	ATG	61.45	Conserved hypothetical protein	-	*Sulfitobacter* sp. NAS-14.1	Hypothetical protein	55% identical over

									NAS141_01721	40 aa (0.012)
P1_62	61847	62140	98	GTG	62.24	Hypothetical protein	-	No significant homology	-	-
P1_63	62390	63340	317	ATG	59.31	Phosphofructokinase	-	*Rhodopseudomonas palustris* DX-1	Phosphofructokinase	56% identical over 312 aa (1e-86)
P1_64	63606	64118	171	ATG	61.79	Phosphoenolpyruvate synthase	-	*Burkholderia* sp. CCGE1003	Phosphoenolpyruvate synthase	64% identical over 128 aa (5e-36)
P1_65	64135	65370	412	ATG	59.79	D-isomer specific 2-hydroxyacid dehydrogenase	-	*Methylocystis* sp. strain Rockwell	D-isomer specific 2-hydroxyacid dehydrogenase, NAD-binding	64% identical over 318 aa (8e-147)
P1_66	65411	66454	348	ATG	59.2	Glyceraldehyde-3-phosphate dehydrogenase	-	*Novosphingobium* sp. PP1Y	Glyceraldehyde-3-phosphate dehydrogenase	79% identical over 347 aa (0.0)
P1_67	66825	67232	136	TTG	62.01	Conserved hypothetical protein	-	*Rhodopseudomonas palustris* BisA53	Hypothetical protein RPE_1013	69% identical over 39 aa (3e-08)
P1_68	68941	67358	528	GTG	63.51	Transposase	-	*Methylocystis* sp. strain Rockwell	Transposase and inactivated derivative	89% identical over 523 aa (0.0)
P1_69	69332	68982	117	ATG	63.25	Transposase, IS66 Orf2 family protein	-	*Methylocystis* sp. strain Rockwell	Transposase, IS66 Orf2 family protein	95% identical over 117 aa (2e-75)
P1_70	69790	69335	152	ATG	67.54	Transposase, IS3/IS911 family protein	-	*Methylocystis* sp. strain Rockwell	Transposase, IS3/IS911 family protein	64% identical over 152 aa (4e-59)
P1_71	69978	71126	383	TTG	59.97	YVTN beta-propeller repeat-containing protein	-	*Polaromonas naphthalenivorans* CJ2	YVTN beta-propeller repeat-containing protein	32% identical over 263 aa (1e-34)
P1_72	71818	71486	111	ATG	66.97	TonB family protein	-	*Methylosinus trichosporium* OB3b	TonB family protein	44% identical over 77 aa (8e-09)
P1_73	72598	71885	238	ATG	66.67	Conserved hypothetical protein	-	*Methylocystis* sp. strain	Hypothetical protein	69% identical over

								Rockwell	Met49242DRAFT_0113	238 aa (1e-95)
P1_74	73116	72598	173	ATG	65.32	Biopolymer transport protein ExbD/TolR	-	*Methylocystis* sp. strain Rockwell	Biopolymer transport protein ExbD/TolR	78% identical over 147 aa (4e-70)
P1_75	73750	73175	192	ATG	69.97	MotA/TolQ/ExbB proton channel	-	*Methylocystis* sp. strain Rockwell	MotA/TolQ/ExbB proton channel	88% identical over 187 aa (6e-105)
P1_76	74283	75116	278	ATG	61.99	Glucose kinase	-	*Bacillus halodurans* C-125	Glucose kinase	31% identical over 224 aa (2e-26)
P1_77	76601	75138	488	ATG	59.29	Glucan 1,4-alpha-glucosidase	-	*Methylocystis* sp. strain Rockwell	Glucan 1,4-alpha-glucosidase	58% identical over 481 aa (0.0)
P1_78	76858	76670	63	GTG	55.56	Hypothetical protein	-	-	-	-
P1_79	77324	76926	133	ATG	58.4	ABC transporter domain-containing protein	-	*Methylocella silvestris* BL2	ABC transporter domain-containing protein	50% identical over 68 aa (9e-14)
P1_80	77527	77324	68	ATG	52.94	Hypothetical protein	-	-	-	-
P1_81	77564	79678	705	ATG	60.47	Glycoside hydrolase family 3 domain protein	-	*Methylocystis* sp. strain Rockwell	Glycoside hydrolase family 3 domain protein	64% identical over 698 aa (0.0)
P1_82	79908	80411	168	GTG	57.74	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Uncharacterized protein	73% identical over 44 aa (5e-12)
P1_83	81597	81043	185	GTG	60.36	Transposase	-	*Agrobacterium vitis* S4	Transposase	70% identical over 185 aa (5e-95)
P1_84	81648	82325	226	ATG	61.21	Hypothetical protein	-	-	-	-
P1_85	82455	82685	77	GTG	53.25	Hypothetical protein	-	-	-	-
P1_86	83057	83452	132	TTG	47.22	Nickel responsive transcriptional regulator, NikR	*nikR*	*Dickeya dadantii* Ech586	Nickel responsive transcriptional regulator NikR	42% identical over 129 aa (6e-26)
P1_87	83604	86069	822	ATG	57.91	TonB-dependent receptor	-	*Methylocystis* sp. strain	TonB-dependent receptor	65% identical over

								Rockwell		819 aa (0.0)
P1_88	86243	87643	467	ATG	61.03	Outer membrane efflux protein	-	*Methylocystis* sp. strain Rockwell	Outer membrane efflux protein	66% identical over 461 aa (0.0)
P1_89	87691	88938	416	ATG	58.01	Efflux transporter, RND family, MFP subunit	-	*Methylocystis* sp. strain Rockwell	Efflux transporter, RND family, MFP subunit	66% identical over 403 aa (0.0)
P1_90	89031	92213	1061	ATG	60.1	Heavy metal efflux pump, CzcA family	-	*Methylocystis* sp. strain Rockwell	Heavy metal efflux pump, CzcA family	88% identical over 1044 aa (0.0)
P1_91	92351	92692	114	ATG	52.34	Hypothetical protein	-	-	-	-
P1_92	93523	92921	201	GTG	61.69	Transposase, IS4 family protein	-	*Methylocystis* sp. strain Rockwell	Transposase, IS4 family protein	81% identical over 177 aa (4e-96)
P1_93	95116	93533	528	GTG	63.51	Transposase	-	*Methylocystis* sp. strain Rockwell	Transposase	89% identical over 523 aa (0.0)
P1_94	95507	95157	117	ATG	63.25	Transposase, IS66 Orf2 family protein	-	*Methylocystis* sp. strain Rockwell	Transposase, IS66 Orf2 family protein	95% identical over 117 aa (2e-75)
P1_95	95716	95510	69	GTG	67.63	Transposase, IS3/IS911 family protein	-	*Methylocystis* sp. strain Rockwell	Transposase, IS3/IS911 family protein	61% identical over 69 aa (7e-17)
P1_96	96825	95842	328	ATG	64.23	Transposase, IS4 family protein	-	*Methylocystis* sp. strain Rockwell	Transposase, IS4 family protein	81% identical over 278 aa (5e-159)
P1_97	97950	97105	282	GTG	64.3	H^+ transporting two-sector ATPase gamma subunit	-	*Methylocystis* sp. strain Rockwell	H^+ transporting two-sector ATPase gamma subunit	40% identical over 280 aa (8e-52)
P1_98	99491	97950	514	ATG	63.81	ATP synthase F1 sub-complex, alpha 2 subunit	*atpA2*	*Methylocystis* sp. strain Rockwell	ATP synthase F1, alpha subunit	77% identical over 500 aa (0.0)
P1_99	100222	99491	244	ATG	63.52	ATP synthase F0 sub-complex, B subunit	-	*Bradyrhizobium* sp. BTAi1	ATP synthase F0 subcomplex, B subunit	58% identical over 241 aa (3e-82)
P1_100	100464	100228	79	ATG	63.29	ATP synthase F0 sub-complex, C 2	*atpE2*	*Bradyrhizobium* sp. BTAi1	ATP synthase F0	88% identical over

						subunit			subcomplex, C subunit	78 aa (3e-34)
P1_101	101959	100478	494	ATG	59.11	ATP synthase F0 sub-complex, A 2 subunit	*atpB2*	*Methylocystis* sp. strain Rockwell	ATP synthase F0 subcomplex, A subunit	60% identical over 490 aa (0.0)
P1_102	102297	102013.	95	ATG	62.46	Hypothetical protein	-	-	-	-
P1_103	102578	102297	94	ATG	64.89	H^+ transporting ATP synthase protein 1	-	*Methylocystis* sp. strain Rockwell	F0F1-ATPase subunit, putative (H^+ transporting ATP synthase protein1)	67% identical over 89 aa (1e-29)
P1_104	103027	102578	150	ATG	60.22	ATP synthase F1 sub-complex, epsilon subunit	-	*Methylocystis* sp. strain Rockwell	ATP synthase F1 subcomplex, epsilon subunit	68% identical over 146 aa (2e-63)
P1_105	104475	103027	483	TTG	64.46	ATP synthase F1 sub-complex, beta 2 subunit	*atpD2*	*Methylocystis* sp. strain Rockwell	ATP synthase F1 subcomplex, beta subunit	83% identical over 464 aa (0.0)
P1_106	104678	105940	421	TTG	58.75	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_3945	80% identical over 421 aa (0.0)
P1_107	105948	106418	157	ATG	66.03	Phosphate acetyl/butaryl transferase	-	*Novosphingobium* sp. PP1Y	Phosphate acetyl/butaryl transferase	60% identical over 150 aa (1e-48)
P1_108	106427	108454	676	ATG	62.08	Oxidoreductase FAD/NAD(P)-binding domain protein		*Methylocystis* sp. strain Rockwell	Oxidoreductase FAD/NAD(P)-binding domain protein	69% identical over 560 aa (0.0)
P1_109	108461	109219	253	ATG	60.47	Conserved hypothetical protein		*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_3295	66% identical over 253 aa (4e-104)
P1_110	109219	109986	256	ATG	58.46	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_3294	63% identical over 256 aa (3e-107)
P1_111	110056	110400	115	ATG	57.1	Hypothetical protein	-	-	-	-
P1_112	110841	110653	63	GTG	61.38	Hypothetical protein	-	-	-	-

P1_113	111161	110841	107	TTG	65.11	Beta-glucosidase	-	*Methylocystis* sp. strain Rockwell	Beta-glucosidase	57% identical over 84 aa (1e-15)
P1_114	112329	111406	308	ATG	60.17	Fructose-bisphosphate aldolase, class I	*fbaB*	*Rhodospirillum centenum* SW	Fructose-bisphosphate aldolase, class I	74% identical over 307 aa (1e-171)
P1_115	112589	112783	65	GTG	53.85	Hypothetical protein	-	-	-	-
P1_116	114112	113732	127	ATG	59.58	Conserved hypothetical protein		*Methylosinus trichosporium* OB3b	Hypothetical protein MettrDRAFT_3909	48% identical over 71 aa (4e-15)
P1_117	114498	114289	70	GTG	60.0	Transposase	-	*Methylosinus trichosporium* OB3b	Transposase and inactivated derivatives-like protein	57% identical over 42 aa (1e-06)
P1_118	115733	115071	221	GTG	56.56	Conserved hypothetical protein	-	*Burkholderia vietnamiensis* G4	Conserved hypothetical protein	76% identical over 219 aa (2e-120)
P1_119	116052	115765	96	ATG	57.64	Conserved hypothetical protein	-	*Burkholderia vietnamiensis* G4	Conserved hypothetical protein	58% identical over 82 aa (3e-25)
P1_120	117156	116095	354	ATG	61.11	Copper-translocating P-type ATPase	-	*Thermaerobacter marianensis* DSM 12885	Copper-translocating P-type ATPase	65% identical over 120 aa (7e-53)
P1_121	119699	117156	848	TTG	61.4	Copper-translocating P-type ATPase	*actP*	*Thermaerobacter marianensis* DSM 12885	Copper-translocating P-type ATPase	64% identical over 825 aa (0.0)
P1_122	120245	119751	165	GTG	58.79	Predicted metal-binding protein	-	*Hahella chejuensis* KCTC 2396	Predicted metal-binding protein	47% identical over 147 aa (1e-48)
P1_123	120757	120398	120	ATG	61.67	Periplasmic copper-binding protein, CusF	-	*Methylocystis* sp. strain Rockwell	Periplasmic copper-binding protein	53% identical over 110 aa (2e-32)
P1_124	121057	120872	62	ATG	60.22	Heavy metal efflux pump, CzcA family	-	*Magnetospirillum magnetotacticum* MS-1	Heavy metal efflux pump, CzcA family	76% identical over 46 aa (4e-12)

P1_125	124601	121443	1053	GTG	60.34	Heavy metal efflux pump, CzcA family	-	*Methylocystis* sp. strain Rockwell	Heavy metal efflux pump, CzcA family	86% identical over 1040 aa (0.0)
P1_126	126019	124601	473	ATG	61.59	Efflux transporter, RND family, MFP subunit	-	*Methylocystis* sp. strain Rockwell	Efflux transporter, RND family, MFP subunit	67% identical over 473 aa (0.0)
P1_127	126903	126019	295	GTG	63.16	Outer membrane efflux protein	-	*Methylocystis* sp. strain Rockwell	Outer membrane efflux protein	73% identical over 154 aa (7e-70)
P1_128	127313	126909	135	ATG	59.26	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_3930	67% identical over 119 aa (1e-49)
P1_129	127512	127321	64	GTG	59.9	Hypothetical protein	-	-	-	-
P1_130	127994	127545	150	ATG	61.11	Putative periplasmic copper-binding protein, CusF	-	*Methylobacterium extorquens* DM4	Putative periplasmic copper-binding protein (cusF-like domain)	48% identical over 83 aa (8e-20)
P1_131	128602	128378	75	ATG	59.56	Hypothetical protein	-	-	-	-
P1_132	129085	129558	158	ATG	59.92	Methyltransferase type 11	-	*Methylocystis* sp. strain Rockwell	Methyltransferase type 11	81% identical over 149 aa (5e-84)
P1_133	129671	130885	405	GTG	53.5	Glycoside hydrolase family 1	-	*Methylocystis* sp. strain Rockwell	Glycoside hydrolase family 1	94% identical over 405 aa (0.0)
P1_134	130928	131725	266	GTG	53.26	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Protein of unknown function DUF344	80% identical over 251 aa (1e-156)
P1_135	133712	131796	639	ATG	59.1	Conserved hypothetical protein	-	*Alicycliphilus denitrificans* BC	Hypothetical protein Alide_2732	70% identical over 572 aa (0.0)
P1_136	134173	133712	154	ATG	56.71	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_4004	42% identical over 115 aa (6e-19)
P1_137	134853	134245	203	ATG	53.2	Cytochrome C class I protein	-	*Rhodopseudomonas palustris* TIE-1	Cytochrome C class I protein	50% identical over 154 aa (7e-48)

P1_138	135083	135466	128	ATG	61.46	Transposase, resolvase domain-containing protein	-	*Vibrio furnissii* CIP 102972	Resolvase domain-containing protein	42% identical over 62 aa (1e-07)
P1_139	136332	136096	79	GTG	60.76	Transposase, integrase catalytic subunit	-	*Geobacter uraniireducens* Rf4	Integrase catalytic subunit	51% identical over 42 aa (0.011)
P1_140	136664	136356	103	TTG	59.87	Glycoside hydrolase family 3 domain containing protein	-	*Methylocystis* sp. strain Rockwell	Glycoside hydrolase family 3 domain containing protein	75% identical over 77 aa (5e-29)
P1_141	137062	136664	133	ATG	58.9	Glycoside hydrolase family 3 domain containing protein	-	*Methylocystis* sp. strain Rockwell	Glycoside hydrolase family 3 domain containing protein	59% identical over 78 aa (5e-20)
P1_142	137361	139184	608	ATG	58.99	ABC transporter domain-containing protein	-	*Methylocystis* sp. strain Rockwell	ABC transporter domain-containing protein	72% identical over 593 aa (0.0)
P1_143	139230	140663	478	ATG	58.58	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_3939	73% identical over 465 aa (0.0)
P1_144	141980	140961	340	ATG	58.33	D-isomer specific 2-hydroxyacid dehydrogenase, NAD-binding catalytic domain	-	*Methylocystis* sp. strain Rockwell	D-isomer specific 2-hydroxyacid dehydrogenase, NAD-binding catalytic domain	64% identical over 332 aa (8e-159)
P1_145	144552	142150	801	ATG	60.84	Phosphoenolpyruvate synthase	-	*Methylosinus trichosporium* OB3b	Phosphoenolpyruvate synthase	81% identical over 787 aa (0.0)
P1_146	145068	144805	88	GTG	52.65	Conserved hypothetical protein	-	*Halothiobacillus neapolitanus* c2	Hypothetical protein Hneap_0972	61% identical over 87 aa (3e-27)
P1_147	145263	148550	1096	GTG	58.88	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_3848	68% identical over 1077 aa (0.0)
P1_148	148873	149061	63	TTG	58.73	Conserved hypothetical protein	-	*Agrobacterium rhizogenes*	Uncharacterized protein	59% identical over 61 aa (2e-16)

P1_149	150200	150583	128	ATG	59.9	Single-strand binding protein/primosomal replication protein n	-	*Nitrobacter hamburgensis* X14	Single-strand binding protein/primosomal replication protein n	49% identical over 104 aa (5e-33)
P1_150	150742	150996	85	TTG	58.43	Hypothetical protein	-	-	-	-
P1_151	151691	151278	138	GTG	63.04	Phosphoenolpyruvate synthase	-	*Methylosinus trichosporium* OB3b	Phosphoenolpyruvate synthase	80% identical over 117 aa (4e-59)
P1_152	152274	151714	187	ATG	61.5	Hypothetical protein	-	-	-	-
P1_153	153933	152281	551	ATG	65.64	Mercuric reductase, MerA	*merA*	*Salmonella enterica* subsp. *enterica* serovar Kentucky	Mercuric reductase, MerA	78% identical over 551 aa (0.0)
P1_154	154389	154114	92	ATG	57.61	Mercuric ion transport protein periplasmic component, MerP	*merP*	*Pseudomonas* sp.	Mercuric transport protein periplasmic component, MerP	73% identical over 92 aa (1e-41)
P1_155	154811	154410	134	ATG	66.17	Mercuric ion transport protein, MerT	*merT*	*Methylocystis* sp. strain Rockwell	Mercuric transport protein, MerT	74% identical over 134 aa (5e-65)
P1_156	155101	155577	159	ATG	59.96	Mercuric ion resistance operon regulatory protein, MerR	*merR*	*Methylocystis* sp. strain Rockwell	Mercuric resistance operon regulatory protein, MerR	82% identical over 153 aa (1e-85)
P1_157	155626	155904	93	ATG	60.93	Hypothetical protein	-	-	-	-
P1_158	155923	157029	369	ATG	58.81	Predicted aminopeptidase	-	NC10 bacterium 'Dutch sediment'	Predicted aminopeptidase	67% identical over 290 aa (3e-130)
P1_159	157412	157086	109	ATG	66.67	Putative pyruvate formate lyase-activating enzyme	-	*Endoriftia persephone*	Putative pyruvate formate lyase-activating enzyme	54% identical over 100 aa (1e-33)
P1_160	157617	159152	512	ATG	64.58	Transposase, integrase catalytic region	-	*Methylobacterium chloromethanicum* CM4	Integrase catalytic region	71% identical over 512 aa (0.0)
P1_161	159792	159196	199	ATG	65.66	Conserved hypothetical protein	-	*Bradyrhizobium* sp. BTAi1	Uncharacterized protein	44% identical over 195 aa (3e-42)

P1_162	160868	159768	367	ATG	59.58	Radical SAM domain-containing protein	-	*Methylobacterium nodulans* ORS 2060	Radical SAM domain-containing protein	75% identical over 257 aa (1e-143)
P1_163	160903	161892	330	ATG	62.32	Radical SAM domain-containing protein	-	*Chloroherpeton thalassium* ATCC 35110	Radical SAM domain-containing protein	58% identical over 293 aa (2e-124)
P1_164	161918	162310	131	ATG	58.78	Conserved hypothetical protein	-	*Xenorhabdus bovienii* SS-2004	Sea14	63% identical over 46 aa (4e-12)
P1_165	162614	162441	58	TTG	62.07	Transposase	-	*Magnetospirillum magnetotacticum* MS-1	Transposase and inactivated derivatives	59% identical over 51 aa (4e-13)
P1_166	163477	162638	280	ATG	59.17	Transposase, IS116/IS110 /IS902 family protein	-	*Acidiphilium cryptum* JF-5	Transposase, IS116/IS110 /IS902 family protein	66% identical over 236 aa (4e-110)
P1_167	164357	163905	151	ATG	59.38	Periplasmic copper-binding protein, CusF	-	*Methylobacterium extorquens* DM4	Periplasmic copper-binding protein	54% identical over 78 aa (1e-24)
P1_168	164488	164913	142	ATG	62.44	2-oxo-acid dehydrogenase E1 subunit, homodimeric type	-	*Hyphomicrobium denitrificans* ATCC 51888	2-oxo-acid dehydrogenase E1 subunit, homodimeric type	33% identical over 89 aa (0.48)
P1_169	165113	165739	209	ATG	63.0	Ubiquinol-cytochrome c reductase, iron-sulfur subunit	-	*Nitrobacter hamburgensis* X14	Ubiquinol-cytochrome c reductase, iron-sulfur subunit	59% identical over 202 aa (9e-91)
P1_170	165842	165991	50	ATG	60.0	Hypothetical protein	-	No significant homology	-	-
P1_171	166202	166480	93	ATG	62.72	Conserved hypothetical protein	-	*Afipia* sp. 1NLS2	Hypothetical protein AfiDRAFT	49% identical over 60 aa (2e-11)
P1_172	166639	167274	212	ATG	54.72	Conserved hypothetical protein	-	*Nitrobacter hamburgensis* X14	Conserved hypothetical protein	69% identical over 204 aa (1e-92)
P1_173	168828	167422	469	GTG	63.18	Type I secretion outer membrane	-	*Methylocystis* sp. strain	Type I secretion outer	72% identical over

						protein, TolC family		Rockwell	membrane protein, TolC family	465 aa (0.0)
P1_174	172055	168828	1076	ATG	62.48	Heavy metal efflux pump, CzcA family	-	*Methylocystis* sp. strain Rockwell	Heavy metal efflux pump, CzcA family	81% identical over 1076 aa (0.0)
P1_175	173369	172071	433	ATG	63.43	Efflux transporter, RND family, MFP subunit	-	*Methylocystis* sp. strain Rockwell	Efflux transporter, RND family, MFP subunit	65% identical over 397 aa (3e-171)
P1_176	173547	173873	109	ATG	59.02	Conserved hypothetical protein	-	*Alicycliphilus denitrificans* BC	Propeptide pepsy amd peptidase m4	37% identical over 104 aa (3e-15)
P1_177	173876	174535	220	ATG	58.79	Two component transcriptional regulator, winged helix family	-	*Methylobacterium nodulans* ORS 2060	Two component transcriptional regulator, winged helix family	66% identical over 216 aa (6e-105)
P1_178	174535	175917	461	ATG	60.45	Integral membrane sensor signal transduction histidine kinase	-	*Allochromatium vinosum* DSM 180	Integral membrane sensor signal transduction histidine kinase	45% identical over 434 aa (5e-108)
P1_179	178933	175982	984	ATG	64.06	Heavy metal translocating P-type ATPase	-	*Mesorhizobium* sp. BNC1	Heavy metal translocating P-type ATPase	60% identical over 729 aa (0.0)
P1_180	179571	179954	128	ATG	59.9	Single-strand binding protein/primosomal replication protein n		*Nitrobacter hamburgensis* X14	Single-strand binding protein/primosomal replication protein n	51% identical over 104 aa (3e-34)
P1_181	180915	180649	89	GTG	62.92	Efflux transporter, RND family, MFP subunit	-	*Methylocystis* sp. strain Rockwell	Efflux transporter, RND family, MFP subunit	40% identical over 51 aa (0.067)
P1_182	181087	181329	81	ATG	58.85	Transposase, mutator type	-	alphaproteobacterium BAL199	Transposase, mutator type	91% identical over 43 aa (5e-17)
P1_183	181761	181348	138	ATG	57.0	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_3306	44% identical over 93 aa (3e-21)

P1_184	183302	182229	358	TTG	63.69	Conserved hypothetical protein	-	*Mesorhizobium loti* MAFF303099	Hypothetical protein mll9238	36% identical over 246 aa (2e-38)
P1_185	183352	183537	62	GTG	56.45	Hypothetical protein	-	No significant homology	-	-
P1_186	183620	183967	116	ATG	54.6	Single-strand binding protein/primosomal replication protein n	-	*Methylocystis* sp. strain Rockwell	Single-strand binding protein/primosomal replication protein n	59% identical over 111 aa (9e-46)
P1_187	184014	184235	74	GTG	59.01	Hypothetical protein	-	No significant homology	-	-
P1_188	184578	184345	78	ATG	59.83	Hypothetical protein	-	No significant homology	-	-
P1_189	184797	185066	90	GTG	60.0	Conserved hypothetical protein	-	*Rhodopseudomonas palustris* DX-1	Hypothetical protein Rpdx1_3305	60% identical over 82 aa (1e-26)
P1_190	185321	185097	75	ATG	58.22	Conserved hypothetical protein	-	*Methylosinus trichosporium* OB3b	Hypothetical protein MettrDRAFT_3909	32% identical over 67 aa (1.0)
P1_191	185587	185321	89	GTG	61.42	Hypothetical protein	-	No significant homology	-	-
P1_192	185630	186373	248	GTG	61.56	Single-strand binding protein/primosomal replication protein n	-	*Methylocystis* sp. strain Rockwell	Single-strand binding protein/primosomal replication protein n	48% identical over 125 aa (9e-29)
P1_193	187190	186606	195	ATG	60.34	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_3576	66% identical over 188 aa (3e-85)
P1_194	187310	187522	71	ATG	56.34	Conserved hypothetical protein	-	*Caulobacter* sp. K31	Hypothetical protein Caul_2262	55% identical over 56 aa ()3e-13
P1_195	187518	188660	381	ATG	57.92	DNA-cytosine methyltransferase	-	*Methylocystis* sp. strain Rockwell	DNA-cytosine methyltransferase	82% identical over 381 aa (0.0)
P1_196	188660	189514	285	GTG	58.6	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_3578	78% identical over 283 aa (2e-167)
P1_197	189510	190241	244	ATG	57.24	Excinuclease ABC C subunit	-	*Methylocystis* sp. strain	Excinuclease ABC C	83% identical over

						domain protein		Rockwell	subunit domain protein	233 aa (3e-139)
P1_198	190275	190535	87	GTG	57.47	Hypothetical protein	-	-	-	-
P1_199	190529	191257	243	GTG	64.88	Phosphatidylserine/phosphatidylglycerophosphate/cardiolipin synthase-like protein	-	*Methylosinus trichosporium* OB3b	Phosphatidylserine/phosphatidylglycerophosphate/cardiolipin synthase-like protein	61% identical over 210 aa (2e-83)
P1_200	191762	191349	138	ATG	67.39	Mobilization protein, MobE	*mobE*	*Acidithiobacillus ferrooxidans*, plasmid pTF-FC2	Mobilization protein, MobE	48% identical over 122 aa (3e-23)
P1_201	192453	191770	228	ATG	59.21	Mobilization protein, MobD	*mobD*	*Pseudomonas aeruginosa*	Mobilization protein, MobD	64% identical over 225 aa (2e-100)
P1_202	192863	192468	132	ATG	59.09	Conserved hypothetical protein	-	*Desulfovibrio* sp. 3_1_syn3	Conserved hypothetical protein	29% identical over 124 aa (3e-06)
P1_203	193281	193685	135	ATG	60.99	Mobilization protein, MobC	*mobC*	*Methylocystis* sp. strain Rockwell	Mobilization protein, MobC	50% identical over 129 aa (2e-34)
P1_204	193685	196648	988	ATG	65.86	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_3318	50% identical over 987 aa (0.0)
P1_205	198012	196984	343	ATG	61.61	Transposase, IS116/IS110/IS902 family protein	-	*Nitrobacter hamburgensis* X14	Transposase, IS116/IS110/IS902 family protein	90% identical over 343 aa (0.0)
P1_206	198443	198721	93	ATG	62.72	Addiction module antidote protein, CopG/Arc/MetJ/RelE family	-	*Starkeya novella* DSM 506	Addiction module antidote protein, CopG/Arc/MetJ family	63% identical over 93 aa (8e-38)
P1_207	198721	199017	99	GTG	62.63	Toxin-antitoxin system, toxin component, RelE family	-	*Roseibium* sp. TrichSKD4	Toxin-antitoxin system, toxin component, RelE family	56% identical over 96 aa (2e-29)
P1_208	199458	199159	100	ATG	63.33	Conserved hypothetical protein	-	*Streptomyces cattleya*	Conserved exported protein	36% identical over

								NRRL 8057	of unknown function	82 aa (3e-08)
P1_209	199603	200106	168	ATG	54.76	Invasion associated locus B family protein	-	*Methylosinus trichosporium* OB3b	Invasion associated locus B family protein	36% identical over 150 aa (5e-28)
P1_210	200125	200802	226	ATG	60.62	Conserved hypothetical protein	-	*Methylosinus trichosporium* OB3b	Protein of unknown function DUF159	64% identical over 220 aa (1e-95)
P1_211	200923	201141	73	ATG	61.64	Hypothetical protein	-	No significant homology	-	-
P1_212	201922	201242	227	ATG	58.44	Conserved hypothetical protein	-	*Roseobacter* sp. MED193	Hypothetical protein MED193_00415	37% identical over 216 aa (1e-33)
P1_213	202880	201978	301	ATG	58.03	Putative transcriptional regulator	-	*Brevundimonas diminuta* ATCC 11568	Putative transcriptional regulator	64% identical over 297 aa (1e-131)
P1_214	203016	204161	382	ATG	63.53	Patatin-like phospholipase	-	*Gluconacetobacter diazotrophicus* PAl 5	Patatin-like phospholipase	87% identical over 382 aa (0.0)
P1_215	204168	205211	348	ATG	63.12	Conserved hypothetical protein	-	*Paracoccus denitrificans* SD1	Hypothetical protein PDI_0629	86% identical over 324 aa (0.0)
P1_216	205207	207042	612	ATG	65.47	Putative ubiquitin-activating protein	-	*Gluconacetobacter diazotrophicus* PAl 5	Putative ubiquitin-activating protein	79% identical over 585 aa (0.0)
P1_217	207071	207265	65	ATG	61.03	Hypothetical protein	-	No significant homology	-	-
P1_218	207499	208395	299	ATG	62.99	Peptidase M, neutral zinc metallopeptidase, zinc-binding site	-	*Polymorphum gilvum* SL003B-26A1	Peptidase M, neutral zinc metallopeptidase, zinc-binding site	59% identical over 289 aa (9e-120)
P1_219	210882	208471	804	ATG	63.14	Plasmid conjugal transfer protein, TraG family	*traG*	*Methylocystis* sp. strain Rockwell	TraG family protein	60% identical over 793 aa (0.0)
P1_220	211971	210901	357	ATG	62.0	P-type DNA transfer ATPase, VirB11	*virB11*	*Methylocystis* sp. strain Rockwell	P-type DNA transfer ATPase VirB11	77% identical over 351 aa (0.0)
P1_221	213190	211952	413	ATG	61.74	Plasmid conjugal transfer protein,	*virB10*	*Methylocystis* sp. strain	Conjugation TrbI family	64% identical over

						TrbI/VirB10		Rockwell	protein	413 aa (2e-176)
P1_222	214055	213210	846	ATG	60.4	Plasmid conjugal transfer protein, TrbG/VirB9	*virB9*	*Methylocystis* sp. strain Rockwell	Conjugal transfer protein TrbG/VirB9	63% identical over 265 aa (4e-111)
P1_223	214843	214055	263	ATG	61.6	Plasmid conjugal transfer protein, TrbF/VirB8	*virB8*	*Methylocystis* sp. strain Rockwell	Type IV secretion protein	71% identical over 248 aa (4e-116)
P1_224	215943	214852	364	ATG	61.72	Plasmid conjugal transfer protein, TrbL/VirB6	*virB6*	*Methylocystis* sp. strain Rockwell	TrbL/VirB6 plasmid conjugal transfer protein	63% identical over 361aa (7e-151)
P1_225	216858	216145	238	ATG	61.76	Plasmid conjugal transfer protein, TrbJ/VirB5	*virB5*	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_3564 (conjugal transfer protein)	76% identical over 218 aa (1e-70)
P1_226	217539	216874	222	ATG	62.76	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_3565	49% identical over 214 aa (1e-49)
P1_227	219964	217532	811	GTG	60.3	Plasmid conjugal transfer ATPase, TrbE/VirB4	*virB4*	*Methylocystis* sp. strain Rockwell	Type IV secretion/conjugal transfer ATPase, VirB4 family	82% identical over 811 aa (0.0)
P1_228	220264	219977	96	ATG	55.21	Plasmid conjugal transfer protein, TrbD/VirB3	*virB3*	*Methylocystis* sp. strain Rockwell	Type IV secretory pathway VirB3 family protein	86% identical over 92 aa (7e-54)
P1_229	220581	220267	105	ATG	62.22	Plasmid conjugal transfer protein, TrbC/VirB2	*virB2*	*Methylocystis* sp. strain Rockwell	Conjugal transfer protein TrbC	70% identical over 103 aa (4e-34)
P1_230	221422	220568	285	ATG	62.69	Plasmid conjugal transfer protein, TrbB/VirB1	*virB1*	*Methylocystis* sp. strain Rockwell	Lytic transglycosylase catalytic	57% identical over 194 aa (5e-65)
P1_231	221470	221898	143	TTG	55.94	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_3570	46% identical over 69 aa (1e-09)
P1_232	221898	222512	205	ATG	60.81	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_4335	60% identical over 197 aa (2e-78)

P1_233	222554	223126	191	ATG	61.43	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_3772	42% identical over 104 aa (3e-19)
P1_234	223129	223797	223	ATG	63.38	Transglycosylase SLT domain-containing protein	-	*Desulfovibrio alaskensis* G20	Transglycosylase SLT domain-containing protein	34% identical over 126 aa (5e-09)
P1_235	223803	224378	192	ATG	63.72	Succinoglycan biosynthesis protein	-	*Methylocystis* sp. strain Rockwell	Succinoglycan biosynthesis protein	63% identical over 185 aa (4e-73)
P1_236	224384	225289	302	ATG	65.56	Virulence-associated protein E	-	*Methylocystis* sp. strain Rockwell	Virulence-associated protein E	42% identical over 279 aa (1e-57)
P1_237	226951	225500	484	ATG	56.75	Conserved hypothetical protein	-	*Mesorhizobium loti* MAFF303099	Hypothetical protein mlr6204	95% identical over 484 aa (0.0)
P1_238	228273	226951	441	ATG	56.16	Conserved hypothetical protein	-	*Mesorhizobium loti* MAFF303099	Hypothetical protein mlr6203	91% identical over 441 aa (0.0)
P1_239	229120	228560	187	ATG	58.65	Putative transcriptional regulator	-	*Methylocystis* sp. strain Rockwell	Helix-turn-helix domain-containing protein	55% identical over 157 aa (1e-55)
P1_240	229314	229568	85	TTG	72.55	Conserved hypothetical protein	-	*Mesorhizobium* sp. BNC1	Hypothetical protein Meso_4593	57% identical over 67 aa (8e-15)

Table S3. The ORFs present in pBSC2-2 and their annotations.

ORF identifier	Begin cds	End cds	Size (aa)	First codon	% G+C content	Product	Gene Name	Organism of protein to which most similar	Homolog protein	Similarity (*e*- value)
P2_1	49	432	384	ATG	58.85	Single-strand DNA binding protein/primosomal replication protein n	-	*Nitrobacter hamburgensis* X14	Single-strand DNA binding protein/primosomal replication protein n	50% identical over 104 aa (3e-08)
P2_2	707	528	60	ATG	51.36	Hypothetical protein	-	No significant homology	-	-
P2_3	1230	964	89	ATG	61.8	Conserved hypothetical protein	-	*Yersinia enterocolitica*	Uncharacterized protein	37% identical over 59 aa (1.0)
P2_4	2345	1347	333	ATG	63.96	Nitrile hydratase	-	*Mycobacterium vanbaalenii* DSM 7251	Nitrile hydratase, beta subunit	60 % identical over 142 aa (22 - 163) (2e-38)
									Nitrile hydratase, beta subunit	72% identical over 83 aa (187 -269) (4e-25)
P2_5	2422	2622	67	GTG	59.2	Hypothetical protein	-	No significant homology	-	-
P2_6	3477	2758	240	ATG	61.81	RNA polymerase, sigma-24 subunit, ECF subfamily	-	*Variovorax paradoxus* S110	RNA polymerase, sigma-24 subunit, ECF subfamily	63% identical over 205 aa (1e-58)
P2_7	4363	3689	225	ATG	61.19	Conserved hypothetical protein	-	*Nocardioides* sp. BAA-499	Uncharacterized protein	53% identical over 172 aa (6e-41)
P2_8	4355	4582	76	GTG	55.26	Hypothetical protein	-	-	-	-
P2_9	5415	4891	175	GTG	58.29	Conserved hypothetical protein	-	*Burkholderia ambifaria* MEX-5	Uncharacterized protein	65% identical over 175 aa (2e-66)
P2_10	5637	5431	69	ATG	52.66	Conserved hypothetical protein	-	*Sphingomonas wittichii*	Uncharacterized protein	85% identical over

								RW1		69 aa (3e-20)
P2_11	5781	6743	321	ATG	56.8	Integrase catalytic region	-	*Rhodomicrobium vannielii* ATCC 17100	Integrase catalytic region	86% identical over 321 aa (0.0)
P2_12	6854	7048	65	GTG	62.05	Putative integrase	-	*Roseovarius* sp. 217	Putative integrase/recombinase protein	78% identical over 64 aa (1e-26)
P2_13	7057	8241	395	GTG	67.26	Putative transposase for insertion sequence element	-	alphaproteobacterium BAL199	Putative transposase for insertion sequence element	64% identical over 392 aa (e-128)
P2_14	8448	8242	69	GTG	63.29	Hypothetical protein	-	-	-	-
P2_15	8663	9205	181	ATG	64.83	Putative transposase	-	*Azorhizobium caulinodans* ATCC 43989	Putative transposase	84% identical over 161 aa (1e-69)
P2_16	9232	9624	131	GTG	60.56	Putative transposase	-	*Ochrobactrum intermedium* LMG 3301	Transposase-like protein	60% identical over 131 aa (9e-37)
P2_17	9719	9904	62	ATG	58.6	Putative transposase	-	*Sphingobium japonicum* NBRC 101211	Putative transposase	72% identical over 44 aa (1e-06)
P2_18	10364	11944	527	ATG	62.56	Drug resistance transporter, EmrB/ QacA subfamily	-	*Methylobacter tundripaludum* SV96	Drug resistance transporter, EmrB/QacA subfamily	50% identical over 522 aa (e-127)
P2_19	12016	12180	55	ATG	53.94	CheB methylesterase	-	*Sinorhizobium meliloti* BL225C	Chemotaxis-specific protein-glutamate methylesterase	65% identical over 55 aa (6e-13)
P2_20	13004	12582	141	ATG	57.92	Putative transposase	-	*Nitrobacter winogradskyi* Nb-255	IstB-like ATP-binding protein	78% identical over 71 aa (6e-25)
P2_21	13022	13345	108	ATG	60.49	Single-strand DNA-binding protein	-	*Sphingobium japonicum* NBRC 101211	Probable single-strand DNA-binding protein	49% identical over 51 aa (2e-04)
P2_22	13646	13404	81	TTG	60.91	Hypothetical protein	-	No significant homology	-	-
P2_23	14026	14766	247	ATG	63.56	Putative ParA-like ATPase, plasmid partition protein	*parA*	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Putative ParA-like (IncC) ATPase	74% identical over 245 aa (2e-97)

P2_24	14782	15750	323	ATG	59.86	Putative ParB, plasmid partition protein	*parB*	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Putative ParB-like (KorB) partition protein	49% identical over 320 aa (4e-65)
P2_25	15905	16420	172	ATG	63.18	Conserved hypothetical protein	-	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Uncharacterized protein	69% identical over 172 aa (4e-54)
P2_26	16420	17493	358	ATG	64.8	Conserved hypothetical protein	-	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Uncharacterized protein	75% identical over 345 aa (e-116)
P2_27	19529	17781	583	ATG	45.97	Conserved hypothetical protein	-	*Desulfovibrio aespoeensis* Aspo-2	Abortive phage infection protein	33% identical over 219 aa (8e-17)
P2_28	19857	20168	104	ATG	61.86	UvrD/REP helicase	-	*Rhodopseudomonas palustris* TIE-1	UvrD/REP helicase	51% identical over 60 aa (1e-10)
P2_29	20175	21302	376	ATG	58.69	DNA-directed DNA polymerase	*dinB*	*Methylocystis sp. strain Rockwell*	DNA-directed DNA polymerase	74% identical over 353 aa (0.0)
P2_30	21839	21360	160	ATG	54.17	Conserved hypothetical protein	-	*Octadecabacter antarcticus* 238	Uncharacterized protein	42% identical over 126 aa (4e-23)
P2_31	22106	22402	99	TTG	52.86	Hypothetical protein	-	No significant homology	-	-
P2_32	23051	22581	157	GTG	60.08	Putative transposase	-	*Rhodopseudomonas palustris* BisB18	Putative transposase	82% identical over 156 aa (2e-71)
P2_33	23647	23141	169	ATG	62.13	Putative transposase	-	*Rhodopseudomonas palustris* BisB18	Transposase and inactivated derivative	75% identical over 167 aa (1e-60)
P2_34	26090	23874	739	ATG	62.74	Helicase, RecD/TraA family protein	-	*Methylobacterium* sp. 4-46	Helicase, RecD/TraA family protein	70% identical over 724 aa (0.0)
P2_35	26637	26254	128	ATG	63.02	Hypothetical protein	-	No significant homology	-	-
P2_36	26803	27072	90	ATG	62.96	Transcriptional regulator	-	*Parvibaculum lavamentivorans* DS-1	Helix-turn-helix domain protein	57% identical over 80 aa (2e-14)
P2_37	28122	27100	341	ATG	64.91	Conserved hypothetical protein	-	*Bradyrhizobium* sp. BTAi1	Uncharacterized protein	36% identical over

								plasmid pBBta01		341 aa (9e-43)
P2_38	29953	28109	615	GTG	61.63	Type IV secretory pathway protein, VirD4	*virD4*	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Type IV secretory pathway protein VirD4	80% identical over 613 aa (0.0)
P2_39	31049	30021	343	GTG	64.82	Type IV secretory pathway protein, AvhB11	*avhB11*	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Type IV secretory pathway AvhB11 protein	81% identical over 341 aa (e-149)
P2_40	32188	31049	380	ATG	65.35	Type IV secretory pathway protein, AvhB10	*avhB10*	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Type IV secretory pathway AvhB10 protein	68% identical over 379 aa (e-125)
P2_41	33053	32202	284	GTG	64.44	Type IV secretory pathway protein, AvhB9	*avhB9*	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Type IV secretory pathway AvhB9 protein	77% identical over 274 aa (e-107)
P2_42	33721	33053	223	ATG	65.02	Type IV secretory pathway protein, AvhB8	*avhB8*	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Type IV secretory pathway AvhB8 protein	81% identical over 223 aa (5e-88)
P2_43	34050	33727	108	ATG	65.12	Type IV secretory pathway protein, AvhB7	*avhB7*	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Type IV secretory pathway AvhB7 protein	48% identical over 102 aa (2e-09)
P2_44	35067	34108	320	ATG	64.58	Type IV secretory pathway protein, AvhB6	*avhB6*	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Type IV secretory pathway AvhB6 protein	81% identical over 319 aa (e-101)
P2_45	35375	35082	98	ATG	62.59	Conserved hypothetical protein	-	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Uncharacterized protein	49% identical over 85 aa (2e-14)
P2_46	36088	35375	238	ATG	61.06	Type IV secretory pathway protein, AvhB5	*avhB5*	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Type IV secretory pathway AvhB5 protein	80% identical over 223 aa (4e-82)
P2_47	38451	36088	788	ATG	63.41	Type IV secretory pathway protein, AvhB4	*avhB4*	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Type IV secretory pathway AvhB4 protein	82% identical over 788 aa (0.0)
P2_48	38779	38447	111	ATG	62.16	Type IV secretory pathway protein, AvhB3	*avhB3*	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Type IV secretory pathway AvhB3 protein	79% identical over 111 aa (9e-45)
P2_49	39075	38785	97	ATG	58.08	Type IV secretory pathway protein, AvhB2	*avhB2*	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Type IV secretory pathway AvhB2 protein	80% identical over 95 aa (7e-26)

P2_50	39857	39153	235	TTG	62.13	Type IV secretory pathway protein, AvhB1	*avhB1*	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Type IV secretory pathway AvhB1 protein	53% identical over 188 aa (3e-39)
P2_51	40099	40824	242	ATG	56.34	Transcriptional regulator, LuxR family	*traR*	*Bradyrhizobium* sp. BTAi1 plasmid pBBta01	Transcriptional regulator, LuxR family	53% identical over 242 aa (1e-64)
P2_52	41358	41167	64	TTG	64.06	Hypothetical protein	-	No significant homology	-	-
P2_53	41805	42080	92	ATG	59.78	Hypothetical protein	-	No significant homology	-	-
P2_54	42088	42408	107	ATG	59.5	Putative integron gene cassette protein	-	Uncultured bacterium	Putative integron gene cassette protein	69% identical over 100 aa (8e-28)
P2_55	42456	43442	329	ATG	63.22	Transcriptional regulator, AraC family	-	*Methylocystis* sp. strain Rockwell	Transcriptional regulator, AraC family	64% identical over 317 aa (8e-149)
P2_56	43653	44423	257	ATG	58.37	Particulate methane monooxygenase subunit C	*pmoC*	*Methylocystis* sp. strain Rockwell	Particulate methane monooxygenase subunit C (singleton)	89% identical over 256 aa (3e-172)
P2_57	44564	45208	215	ATG	63.88	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Hypothetical protein	46% identical over 190 aa (2e-54)
P2_58	46232	45660	191	ATG	49.04	Conserved hypothetical protein	-	No significant homology	Hypothetical protein P3TCK_10093	26% identical over 163 aa (1e-07)
P2_59	46736	47029	98	ATG	51.02	Hypothetical protein	-	No significant homology	-	-
P2_60	47039	48496	486	ATG	51.78	Radical SAM domain protein	-	*Acidiphilium cryptum* JF-5	Radical SAM domain protein	49% identical over 483 aa (e-138)
P2_61	48499	49695	399	ATG	52.88	Radical SAM domain protein	-	*Acidiphilium cryptum* JF-5	Radical SAM domain protein	64% identical over 45 aa (2e-06)
P2_62	49712	50287	192	ATG	58.51	Conserved hypothetical protein	-	*Rhodobacter capsulatus* ATCC BAA-309	Uncharacterized protein	64% identical over 45 aa (2e-06)
P2_63	50403	50876	158	ATG	55.49	Putative nuclease	-	*Methylobacterium*	Nuclease (SNase domain	54% identical over

								extorquens PA1	protein)	155 aa (7e-50)
P2_64	51454	55023	1190	ATG	56.78	Superfamily I DNA/RNA helicase	-	*Caulobacter* sp. K31	Superfamily I DNA or RNA helicase or helicase subunit-like protein	40% identical over 1188 aa (0.0)
P2_65	55023	56780	586	ATG	57.34	Phosphatidylserine/phosphatidylglycerophosphate/ cardiolipin synthase-like protein	-	*Caulobacter* sp. K31	Phosphatidylserine/phosphatidylglycerophosphate/ cardiolipin synthase-like protein	37% identical over 585 aa (4e-99)
P2_66	57446	59515	690	ATG	61.48	ParB-like nuclease	-	*Methylocystis* sp. strain Rockwell	ParB domain protein nuclease	64% identical over 690 aa (0.0)
P2_67	59599	60015	139	ATG	61.87	Conserved hypothetical protein	-	*Mesorhizobium* sp. BNC1	Uncharacterized protein	50% identical over 114 aa (7e-24)
P2_68	60108	64562	1485	ATG	63.86	Methylase/helicase	-	*Cyanothece* sp. CCY0110	Probable methylase/helicase	46% identical over 1435 aa (0.0)
P2_69	64562	65575	338	TTG	61.64	Conserved hypothetical protein	-	*Rhodopseudomonas palustris* BisA53	Uncharacterized protein	51% identical over 332 aa (8e-76)
P2_70	66326	65901	142	ATG	67.14	Mercuric transport protein, MerT	*merT*	*Methylocystis* sp. strain Rockwell	Mercuric transport protein MerT	81% identical over 142 aa (8e-65)
P2_71	66379	66804	142	ATG	64.08	Transcriptional regulator, MerR family protein	*merR*	*Methylocystis* sp. strain Rockwell	Transcriptional regulator, MerR family protein	81% identical over 129 aa (6e-68)
P2_72	67605	66823	261	GTG	63.09	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Protein of unknown function DUF81	82% identical over 261 aa (4e-131)
P2_73	68153	69085	311	ATG	61.84	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Protein of unknown function DUF2493	86% identical over 294 aa (0.0)
P2_74	69197	70465	423	GTG	58.08	Plasmid partitioning protein,	*repA*	*Methylobacterium nodulans*	Plasmid partitioning protein	77% identical over

						RepA		ORS2060	RepA	391 aa (e-172)
P2_75	70465	71445	327	ATG	59.84	Plasmid partitioning protein, RepB	*repB*	*Methylobacterium nodulans* ORS2060	Plasmid partitioning protein RepB	50% identical over 323 aa (3e-66)
P2_76	71639	73006	456	ATG	63.96	Replication protein C	*repC*	*Methylocystis* sp. strain Rockwell	Replication protein C	74% identical over 456 aa (0.0)
P2_77	73006	73299	98	ATG	64.18	Hypothetical protein	-	No significant homology	-	-
P2_78	73609	73436	58	TTG	56.32	Conserved hypothetical protein	-	*Parvularcula bermudensis* HTCC2503	Hypothetical protein PB2503_12564	58% identical over 43 aa (0.001)
P2_79	73982	73638	115	ATG	64.64	Putative transcriptional regulator		*Bordetella parapertussis* 12822	Putative DNA-binding protein	44% identical over 90 aa (2e-15)
P2_80	74334	73990	115	ATG	60.0	Conserved hypothetical protein		*Methylobacterium extorquens* DM4	Uncharacterized protein	59% identical over 101 aa (2e-29)
P2_81	75674	74433	414	ATG	66.02	Integrase family protein	-	*Methylocystis* sp. strain Rockwell	Integrase family protein	72% identical over 388 aa (0.0)
P2_82	75767	77380	538	ATG	66.42	Plasmid replication initiator protein	-	*Rhodobacter sphaeroides* ATCC 17025	Plasmid replication initiator protein	63% identical over 530 aa (0.0)
P2_83	77376	78254	293	ATG	64.85	Conserved hypothetical protein	-	*Rhodobacter sphaeroides* ATCC 17025	Hypothetical protein Rsph17025_4235	67% identical over 293 aa (2e-136)
P2_84	78254	79294	347	TTG	69.45	Conserved hypothetical protein	-	*Rhodobacter sphaeroides* ATCC 17025	Hypothetical protein Rsph17025_4234	45% identical over 314 aa (3e-71)
P2_85	79363	79548	62	ATG	59.14	Hypothetical protein	-	No significant homology	-	-
P2_86	79583	80656	358	ATG	73.56	Conserved hypothetical protein	-	*Methylocystis* sp. strain Rockwell	Protein of unknown function DUF1403	62% identical over 358 aa (2e-103)
P2_87	80656	81348	231	ATG	69.41	Chromosome segregation and condensation protein, ScpB	-	*Methylocystis* sp. strain Rockwell	Chromosome segregation and condensation protein, ScpB	79% identical over 201 aa (7e-107)

P2_88	81348	82673	442	GTG	62.59	Iron-containing alcohol dehydrogenase	-	*Methylocystis* sp. strain Rockwell	Iron-containing alcohol dehydrogenase	69% identical over 340 aa (2e-156)
P2_89	82757	82918	54	ATG	53.7	Hypothetical protein	-	No significant homology	-	-
P2_90	82918	83199	94	TTG	64.89	Hypothetical protein	-	No significant homology	-	-
P2_91	83213	84217	335	ATG	63.68	FAD dependent oxidoreductase	-	*Oceanibulbus indolifex* HEL-45	Dehydrogenase subunit	53% identical over 261 aa (2e-95)
P2_92	84217	84921	235	GTG	58.16	Transposase, IS204/IS1001/ IS1096/IS1165 family protein	-	*Beijerinckia indica* subsp. *indica* ATCC 9039	Transposase, IS204/IS1001/ IS1096/IS1165 family protein	56% identical over 182 aa (3e-62)
P2_93	84926	85741	272	GTG	57.48	Signal transduction histidine kinase with response regulator receiver domain	-	*Bradyrhizobium* sp. BTAi1	Signal transduction histidine kinase with response regulator receiver domain	48% identical over 240 aa (3e-61)
P2_94	87536	85965	524	ATG	52.61	RND efflux system, outer membrane lipoprotein, NodT family	-	*Methylosinus trichosporium* OB3b	NodT family RND efflux system outer membrane lipoprotein	63% identical over 501 aa (0.0)
P2_95	88879	87674	402	ATG	56.72	Efflux transporter, RND family, MFP subunit	-	*Methylosinus trichosporium* OB3b	Efflux transporter, RND family, MFP subunit	68% identical over 382 aa (0.0)
P2_96	91521	88882	880	GTG	58.07	Acriflavin resistance protein	-	*Methylosinus trichosporium* OB3b	Acriflavin resistance protein	76% identical over 872 aa (0.0)
P2_97	93029	91554	492	ATG	59.96	Transposase	-	*Methylocystis* sp. strain Rockwell	Transposase	94% identical over 184 aa (7e-114)
P2_98	94506	93130	459	ATG	65.07	Transposase	-	*Methylocystis* sp. strain Rockwell	Transposase	94% identical over 340 aa (0.0)
P2_99	94912	94562	117	ATG	64.67	Transposase, IS66 family protein	-	*Methylocystis* sp. strain Rockwell	IS66 family protein, Orf2	98% identical over 117 aa (4e-78)
P2_100	95361	94915	149	ATG	66.89	Transposase, IS3/IS911 family	-	*Methylocystis* sp. strain	Transposase, IS3/IS911	93% identical over

						protein		Rockwell	family protein	149 aa (8e-91)
P2_101	96134	95394	247	GTG	54.93	Acriflavin resistance protein	-	*Methylosinus trichosporium* OB3b	Acriflavin resistance protein	78% identical over 233 aa (8e-124)
P2_102	96567	96241	109	GTG	60.55	Hypothetical protein	-	No significant homology	-	-
P2_103	96840	96655	62	ATG	52.15	Hypothetical protein	-	No significant homology	-	-
P2_104	97087	98487	467	ATG	56.96	Transposase	-	*Azospirillum brasilense* Sp245	Transposase	35% identical over 450 aa (6e-77)
P2_105	99186	98383	268	TTG	64.55	Transposase	-	*Methylocystis* sp. strain Rockwell	Transposase	90% identical over 213 aa (1e-136)
P2_106	99487	99167	107	ATG	62.93	Transposase	-	*Methylocystis* sp. strain Rockwell	Transposase	89% identical over 94 aa (7e-48)
P2_107	99988	102075	696	ATG	64.32	Transposase	-	*Sinorhizobium fredii* NGR234	Transposase number 4 of insertion sequence NGRIS-4b	39% identical over 661 aa (2e-129)
P2_108	102075	102773	233	GTG	64.52	Conserved hypothetical protein	-	*Magnetospirillum gryphiswaldense* MSR-1	Hypothetical protein mgI544	38% identical over 161 aa (9e-08)
P2_109	104277	103474	268	ATG		Universal stress protein UspA-like nucleotide-binding protein	-	*Magnetospirillum magneticum* AMB-1	Universal stress protein UspA-like nucleotide-binding protein	33% identical over 267 aa (4e-38)
P2_110	105201	104731	157	ATG	63.48	Transposase	-	*Methylocystis* sp. strain Rockwell	Transposase and inactivated derivatives-like protein	87% identical over 157 aa (4e-97)
P2_111	105674	105291	128	ATG	66.41	Transposase of insertion sequence ISRm10-1	-	*Methylosinus trichosporium* OB3b	Transposase of insertion sequence ISRm10-1, orfA protein	87% identical over 127 aa (6e-75)
P2_112	106474	106235	80	ATG	61.67	Hypothetical protein	-	No significant homology	-	-
P2_113	107115	106570	182	ATG	53.11	Hypothetical protein	-	No significant homology	-	-

P2_114	107723	108862	380	ATG	60.61	YVTN beta-propeller repeat-containing protein	-	*Acidovorax* sp. JS42	YVTN beta-propeller repeat-containing protein	34% identical over 287 aa (1e-44)
P2_115	109117	109491	125	ATG	66.4	Transposase	-	*Rhizobium leguminosarum* bv. *viciae* 3841	Putative transposase	83% identical over 125 aa (5e-69)
P2_116	109491	109874	128	TTG	57.81	Transposase	-	*Rhizobium leguminosarum* bv. *viciae* 3841	Putative transposase	86% identical over 128 aa (7e-76)
P2_117	110014	110709	232	ATG	55.32	RNA polymerase, sigma-24 subunit, ECF subfamily	-	*Planctomyces brasiliensis* DSM 5305	RNA polymerase, sigma-24 subunit, ECF subfamily	36% identical over 202 aa (1e-23)
P2_118	110734	110976	81	ATG	51.03	Hypothetical protein	-	No significant homology	-	-
P2_119	111026	112066	347	TTG	58.5	multicopper oxidase type 3	-	*Methylomicrobium album* BG8	Multicopper oxidase type 3	69% identical over 311 aa (2e-164)
P2_120	112098	113936	613	ATG	59.49	Multicopper oxidase type 1	-	*Methylomicrobium album* BG8	Multicopper oxidase type 1	68% identical over 606 aa (0.0)
P2_121	114131	114733	201	ATG	58.71	GTP cyclohydrolase I	*folE2*	*Methylocystis* sp. strain Rockwell	GTP cyclohydrolase I	88% identical over 189 aa (5e-122)
P2_122	115114	116301	396	ATG	64.14	Secretion protein HlyD family protein	-	*Methylocystis* sp. strain Rockwell	Secretion protein HlyD family protein	71% identical over 396 aa (0.0)
P2_123	117856	116333	508	ATG	67.78	Putative RND efflux system, outer membrane chanel lipoprotein, NodT family	-	*Methylocystis* sp. strain Rockwell	RND efflux system, outer membrane lipoprotein, NodT family	68% identical over 501 aa (0.0)
P2_124	118123	119046	308	ATG	54.11	Transcriptional regulator, LysR family	-	*Paracoccus denitrificans* PD1222	LysR family transcriptional regulator	66% identical over 279 aa (1e-129)
P2_125	119305	119529	75	ATG	56.44	Hypothetical protein	-	No significant homology	-	-
P2_126	119893	120132	80	ATG	53.33	UspA domain-containing universal stress protein	-	*Methylocystis* sp. strain Rockwell	UspA domain-containing protein	53% identical over 54 aa (1e-09)

P2_127	120168	120551	128	ATG	66.67	Transposase of insertion sequence ISRm10-1	-	*Methylosinus trichosporium* OB3b	Transposase of insertion sequence ISRm10-1, orfA protein	87% identical over 127 aa (2e-74)
P2_128	120641	121111	157	ATG	63.27	Transposase	-	*Methylosinus trichosporium* OB3b	Transposase and inactivated derivatives-like protein	88% identical over 157 aa (2e-98)
P2_129	121145	121801	219	ATG	57.23	Putative universal stress protein	-	*Methylocystis* sp. strain Rockwell	UspA domain-containing protein	43% identical over 202 aa (1e-46)
P2_130	121889	124192	768	GTG	62.28	Nitric-oxide reductase	-	*Parvibaculum lavamentivorans* DS-1	Putative nitric-oxide reductase	74% identical over 767 aa (0.0)
P2_131	124509	124234	92	GTG	56.88	Transposase		*Methylocystis* sp. strain Rockwell	Putative transposase	84% identical over 92 aa (2e-50)
P2_132	124760	124413	116	TTG	61.21	Transposase	-	*Gluconacetobacter diazotrophicus* PAl 5	Transposase	81% identical over 73 aa (1e-36)
P2_133	125919	125050	290	ATG	63.33	Transposase	-	*Gluconacetobacter diazotrophicus* PAl 5	Transposase	65% identical over 284 aa (5e-118)
P2_134	126174	125965	70	ATG	58.1	Transposase	-	*Gluconacetobacter diazotrophicus* PAl 5	Transposase	41% identical over 58 aa (1e-04)
P2_135	126800	126294	169	ATG	61.54	Transposase	-	*Rhodopseudomonas palustris* BisB18	Transposase and inactivated derivative	74% identical over 167 aa (3e-84)
P2_136	126946	127269	108	ATG	59.26	Invasion gene expression up-regulator, SirB	-	*Methylophaga aminisulfidivorans* MP	Invasion gene expression up-regulator SirB	44% identical over 93 aa (6e-16)
P2_137	128522	127317	402	ATG	65.67	NnrS protein	-	*Nitrobacter* sp. Nb-311A	NnrS protein	47% identical over 393 aa (7e-100)

P2_138	128778	128605	58	ATG	54.02	Transposase		*Sinorhizobium fredii* NGR234	Transposase	71% identical over 56 aa (2e-19)
P2_139	129869	128886	328	ATG	62.3	ApbE family lipoprotein, putative NosX	*nosX*	*Oligotropha carboxidovorans* OM5	ApbE family lipoprotein	50% identical over 317 aa (2e-97)
P2_140	130709	129882	276	ATG	60.87	Nitrous-oxide reductase, NosY	*nosY*	*Bradyrhizobium* sp. BTAi1 (chromosome)	Copper ABC transporter permease NosY	60% identical over 275 aa (1e-95)
P2_141	131602	130709	298	ATG	62.53	Putative copper ABC transporter ATP-binding protein, NosF	*nosF*	*Sinorhizobium meliloti* BL225C	Putative copper ABC transporter ATP-binding protein NosF	54% identical over 292 aa (7e-99)
P2_142	132840	131602	413	ATG	60.21	Periplasmic copper-binding protein, NosD	*nosD*	*Rhodospirillum centenum* SW	Periplasmic copper-binding protein, NosD	60% identical over 370 aa (2e-164)
P2_143	134789	132864	642	ATG	59.19	Nitrous oxide reductase, NosZ	*nosZ*	*Rhodospirillum centenum* SW	Nitrous oxide reductase NosZ	72% identical over 632 aa (0.0)
P2_144	136936	134798	713	TTG	62.37	Nitrous oxide reductase expression regulator, NosR	*nosR*	*Azospirillum brasilense*	Nitrous oxide reductase transcription regulator, NosR	58% identical over 704 aa (0.0)
P2_145	137187	138041	285	ATG	61.64	Hypothetical protein	-	*Rhodopseudomonas palustris* DX-1	Hypothetical protein Rpdx1_3445	41% identical over 105 aa (9e-06)
P2_146	139019	138390	210	ATG	59.05	Conserved hypothetical protein	-	*Afipia* sp. 1NLS2	Protein of unknown function DUF2478	49% identical over 161 aa (7e-43)
P2_147	140062	140805	248	ATG	61.69	Hydroxylamine reductase	*hcp*	*Methylosinus trichosporium* OB3b	Hybrid cluster protein, hydroxylamine reductase	94% identical over 239 aa (9e-150)
P2_148	140805	141521	239	ATG	63.74	Hydroxylamine reductase	*hcp*	*Methylosinus trichosporium* OB3b	Hybrid cluster protein, hydroxylamine reductase	83% identical over 241 aa (5e-124)
P2_149	141596	141760	55	ATG	54.55	Oxidoreductase FAD/NAD(P)-	-	*Methylosinus trichosporium*	Oxidoreductase	84% identical over

						binding domain protein		OB3b	FAD/NAD(P)-binding domain protein	31 aa (4e-08)
P2_150	141949	142422	158	ATG	59.7	Conserved hypothetical protein		*Methylocystis* sp. strain Rockwell	Hypothetical protein Met49242DRAFT_2113	73% identical over 146 aa (5e-73)
P2_151	142595	142846	84	GTG	61.9	Hypothetical protein	-	No significant homology	-	-
P2_152	142872	143084	71	GTG	56.81	Hypothetical protein	-	No significant homology	-	-