

## NOTE

# Characterization of Plasmid pSY3 in *Sphingobium chungbukense* DJ77

Sun-Mi Yeon and Young-Chang Kim\*

Department of Microbiology, Chungbuk National University, Cheongju, Chungbuk 361-763, Republic of Korea

(Received October 22, 2009 / Accepted December 15, 2009)

**This study determined the complete nucleotide sequence of the plasmid pSY3 from *Sphingobium chungbukense* DJ77. It was 35,735 bp long with a G+C content of 61.9%. Forty open reading frames (ORFs) were found. We predicted these ORFs would encode proteins associated with plasmid replication, conjugative transfer, transposition of genes, plasmid stability/partition, hypothetical protein, and some other functions. Genes for biodegradation were not found. No other plasmid homologous to pSY3 in the overall nucleotide sequence or gene organization could be found in the NCBI database.**

**Keywords:** plasmid pSY3, characterization, *Sphingobium chungbukense* DJ77

The applied microbiology field is becoming more interested in *Sphingomonas* due to its various member species that degrade xenobiotics and antagonize phytopathogenic fungi (Berg and Ballin, 1994). The members of the *Sphingomonadales* order have the ability to adapt more quickly and efficiently to degrade new compounds in their environment than do members of other bacterial genera (Basta *et al.*, 2004). *Sphingobium chungbukense* DJ77, an unusual and interesting bacterium, degrades monocyclic and polycyclic aromatic compounds (Shin *et al.*, 1997), and synthesizes cell envelope glycosphingolipids (Burenjargal *et al.*, 2007). The whole genome sequencing project for *Sphingobium chungbukense* DJ77 (Chungbuk National University, Korea) is currently in progress. While plasmid vectors are very useful tools for such genetic analysis, there has until now been no known vector plasmid for *Sphingomonadales*. In this study, we characterized the plasmid pSY3 in *Sphingobium chungbukense* DJ77. The Materials and Methods used in this study, including bacterial strains and cultures, preparation of genomic DNA, and sequence determination and analysis, can be found (Yeon *et al.*, 2008).

We identified the vegetative replication's origin (*oriV*) and terminus (*terV*) positions using the GenSkew program, on-line version (<http://mips.gsf.de/services/analysis/genskew>).

We obtained pSY3's complete nucleotide sequence from two contigs, 00087 (34,412 bp) and 00103 (1,323 bp), generated by Genome Sequencer FLX automatic sequencers (Goldberg *et al.*, 2006). By comparing these to shotgun sequencing data we found there were no gaps between these two contigs. Computational analyses indicated that pSY3 consisted of 35,735 bp encoding 40 putative open reading

frames (ORFs), which we predicted would function in replication, conjugative transfer, mobilization, plasmid partitioning, and some other functions, based on their functional similarity to particular proteins (Fig. 1 and Table 1). This plasmid contained 69.90% GC. The *oriV* was located (16,521 bp) on the intergenic region between TraG and the plasmid replication initiator protein, while the *terV* was located (35,701 bp) on the intergenic region between TonB and the transposase. A nucleotide (NCBI) data base search on the a total pSY3 sequence revealed that the 40 open reading frames could be categorized into 6 groups based on their functional similarity to previously identified proteins (Table 2).

We found an ORF23.3-encoded 319-aa protein, which showed 61% identity with the previously identified plasmid replication initiator protein of *Caulobacter* sp. K31 (Table 1). This indicated that our plasmid possesses self-replication capacity. Using a NCBI blast homology search, we identified four ORFs whose conjugative function was identical to that of previously identified conjugative proteins (Table 2). ORF17.3, a 973-aa protein, showed 65% identity with the Ti-type conjugative transfer relaxase (TraA) protein of *Caulobacter* sp. K31, and ORF18.3 (150-aa) had 67% identity with conjugal transfer protein TraC. ORF19.3 (103-aa) encoded a protein that showed 71% identity with the TraD protein of *Oceanicaulis alexandrii* HTCC2633, and ORF22.3 encoded a 114-aa protein that was 70% identical with the TraG family protein of *Rhizobium leguminosarum* bv. *trifolii* WSM1325. An important feature of the plasmid pSY3 was the presence of a transposable element/mobile genetic element. It contained 11 ORFs encoding different lengths of proteins, including ISRm2011-2 transposase, mutator type transposase, the Tn3 family, the IS3/IS911 family, the IS66 family, resolvase, and site-specific recombinase. The ORF1.3-encoded 407-aa protein showed 99% identity with the mutator type transposase of *Sphingomonas wittichii* RW1. The ORF11.3-

\* To whom correspondence should be addressed.  
(Tel) 82-43-261-2302; (Fax) 82-43-264-9600  
(E-mail) youngkim@chungbuk.ac.kr

**Table 1.** Putative ORFs on pSY3 and their functions

ORF	ORF		Strand	Length of protein (aa)	Function, closest match	GenBank accession number	Alignment region of pSY3 ORF (aa)	Alignment region of closest match (aa)	% Identity
	Start	End							
1.3	1265	42	-	407	Transposase, mutator type [ <i>Sphingomonas wittichii</i> RW1]	YP_001260485.1	1-407	1-407(407)	99
2.3	1492	1779	+	95	ISRM2011-2 transposase protein [ <i>Sphingomonas</i> sp. SKA58]	ZP_01301879.1	40-95	186-241(241)	67
3.3	3219	1792	-	475	Major facilitator superfamily transporter [marine gamma Proteobacterium HTCC2143]	ZP_01615150.1	50-447	2-403(443)	35
4.3	3484	4263	+	259	Resolvase [ <i>Ochrobactrum tritici</i> ]	ABO70327.1	69-253	1-184(188)	62
5.3	4426	4776	+	116	Hypothetical protein CJA_1213 [ <i>Cellvibrio japonicus</i> Ueda107]	YP_317350.1	1-128	1-116(116)	57
6.3	4773	5111	+	112	Conserved hypothetical protein [ <i>Methylobacterium populi</i> BJ001]	ZP_02200554.1	5-99	3-97(100)	55
7.3	5268	5149	-	39	Predicted: zinc finger protein 7 (KOX 4, clone HF.16) [ <i>Macaca mulatta</i> ]	XP_001106092.1	40-10	67-97(838)	48
8.3	5317	6000	+	227	ParA-like protein [ <i>Gluconobacter oxydans</i> 621H]	YP_190422.1	1-219	9-221(225)	53
9.3	5997	6263	+	88	Hypothetical protein S7335_720 [ <i>Synechococcus</i> sp. PCC 7335]	YP_002710812.1	35-81	40-86(96)	48
10.3	6385	6576	+	63	Hypothetical protein S7335_720 [ <i>Synechococcus</i> sp. PCC 7335]	YP_002710812.1	25-74	33-81(96)	44
11.3	7510	6890	-	206	Transposase Tn3 [ <i>Mesorhizobium</i> sp. BNC1]	YP_665994.1	19-193	1-175(991)	39
12.3	8070	7465	-	201	Resolvase domain [ <i>Methylobacterium populi</i> BJ001]	YP_001919449.1	1-188	1-188(194)	73
13.3	8282	8488	+	68	conserved hypothetical protein [ <i>Octadecabacter antarcticus</i> 238]	YP_002629721.1	49-3	2-48(613)	48
14.3	8552	9172	+	206	putative resolvase [ <i>Sphingomonas</i> sp. SKA58]	ZP_01303723.1	1-201	1-201(206)	97
15.3	9173	9988	+	271	Hypothetical conserved protein [ <i>Rhizobium etli</i> CIAT 652]	YP_001977335.1	7-243	4-265(271)	25
16.3	10889	10182	-	235	hypothetical protein [ <i>Sphingobium</i> sp. MI1205]	BAF56669.1	1-211	1-213(240)	51
17.3	13817	10896	-	973	Ti-type conjugative transfer relaxase TraA [ <i>Caulobacter</i> sp. K31]	YP_001681961.1	1-971	1-950(952)	65
18.3	13847	14299	+	150	Putative conjugal transfer protein, TraC [ <i>Xanthobacter autotrophicus</i> Py2]	YP_001415186.1	46-123	2-78(112)	67
19.3	14244	14555	+	103	Probable conjugal transfer protein TraD [ <i>Oceanicaulis alexandrii</i> HTCC2633]	ZP_00953566.1	1-68	6-75(75)	71
20.3	14552	15055	+	167	Multi-sensor Hybrid Histidine Kinase [ <i>Stigmatella aurantiaca</i> DW4/3-1]	ZP_01459151.1	15-56	1895-1935 (1991)	38
21.3	15294	15926	+	210	Virulence virD4-like protein [ <i>Agrobacterium tumefaciens</i> str. C58]	NP_355804.1	8-178	220-393(554)	70
22.3	15970	16260	+	97	TraG family protein [ <i>Rhizobium leguminosarum</i> bv. <i>trifolii</i> WSM1325]	ZP_02296277.1	1-90	466-555(560)	70
23.3	18102	17143	-	319	Plasmid replication initiator protein-like protein [ <i>Caulobacter</i> sp. K31]	YP_001676602.1	5-315	5-314(322)	61
24.3	18443	18099	-	114	Hypothetical protein MchIDRAFT_0088 [ <i>Methylobacterium chloromethanicum</i> CM4]	YP_002424249.1	1-114	1-110(110)	60
25.3	19096	18443	-	217	ATPase involved in chromosome partitioning-like protein [ <i>Methylobacterium chloromethanicum</i> CM4]	ZP_02059571.1	1-217	1-217(217)	88
27.3	20799	20173	-	208	Conserved hypothetical protein [ <i>Xylella fastidiosa</i> Ann-1]	ZP_00684053.1	105-4	1-102(131)	62
28.3	21068	20931	-	45	Phosphoglycerate mutase [ <i>Sphingomonas wittichii</i> RW1]	YP_001259994.1	46-1	23-67(209)	62

**Table 1.** Continued

ORF	ORF		Strand	Length of protein (aa)	Function, closest match	GenBank accession number	Alignment region of pSY3 ORF (aa)	Alignment region of closest match (aa)	% Identity
	Start	End							
29.3	21512	21117	-	131	Phage integrase family protein [ <i>Sphingomonas wittichii</i> RW1]	YP_001259993.1	1-119	165-283(304)	87
30.3	24588	21586	-	1000	Probable transposase protein [ <i>Sinorhizobium meliloti</i> ]	YP_001965568.1	15-993	1-982(989)	61
31.3	24684	25256	+	190	Site-specific recombinases, DNA invertase Pin homologs [ <i>Magnetospirillum magnetotacticum</i> MS-1]	ZP_00053042.1	1-187	1-187(189)	79
32.3	25570	25851	+	93	Putative addiction module antidote protein, CopG/Arc/MetJ family [ <i>Gluconacetobacter diazotrophicus</i> PAI 5]	YP_002276116.1	1-53	1-53 (82)	67
33.3	25848	26144	+	98	Hypothetical protein msr9263 [ <i>Mesorhizobium loti</i> MAFF303099]	NP_085803.1	1-95	1-95(98)	49
34.3	26477	28165	+	552	Putative monooxygenase, FAD-binding [ <i>Bradyrhizobium</i> sp. BTAi1]	YP_001240048.1	4-549	17-562(578)	52
35.3	28195	30387	+	730	Hypothetical protein pNL1_p167 [ <i>Novosphingobium aromaticivorans</i> ]	NP_049215.1	20-730	22-726(726)	51
36.3	30532	30939	+	135	Transposase IS3/IS911 family protein [ <i>Caulobacter</i> sp. K31]	YP_001672184.1	36-113	9-110(113)	50
37.3	30936	31289	+	117	IS66 Orf2 family protein [ <i>Sinorhizobium medicae</i> WSM419]	YP_001314201.1	8-117	9-118(118)	60
38.3	31341	32888	+	515	IS66 family element, orf3 [ <i>Brucella abortus</i> biovar 1 str. 9-941]	YP_223439.1	32-504	32-511(523)	55
39.3	32885	33493	+	202	yecA family protein [ <i>Caulobacter</i> sp. K31]	YP_001672182.1	18-197	1-182(192)	39
40.3	35681	33666	-	671	TonB-dependent receptor [ <i>Novosphingobium aromaticivorans</i> DSM 12444]	YP_001165641.1	17-703	40-726(726)	49

encoded 206-aa protein showed 39% identity with the Tn3 family transposase of *Mesorhizobium* sp. BNC1, whereas the

ORF2.3-encoded 95-aa protein showed 67% identity with the ISRM2011-2 transposase of *Sphingomonas* sp. SKA58.

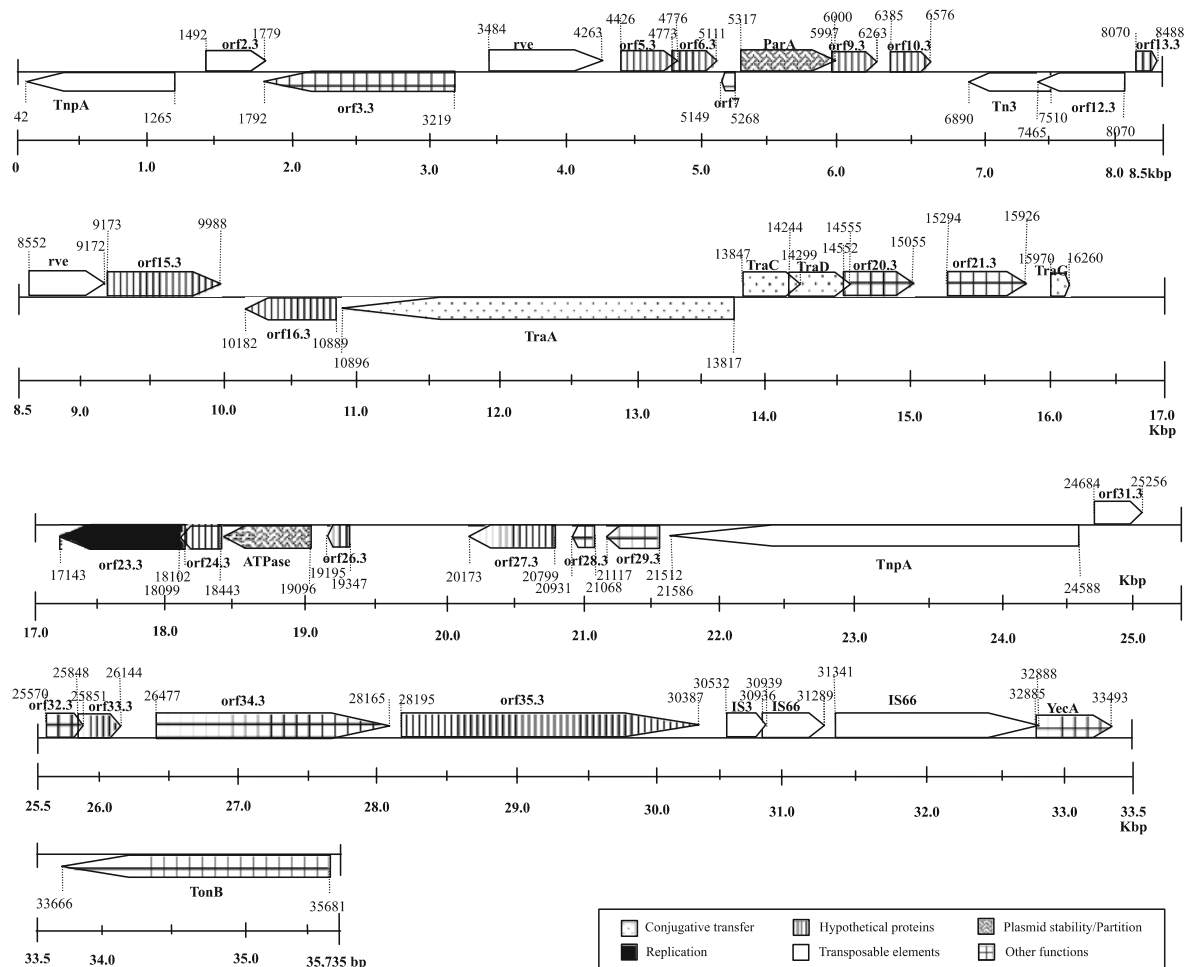
**Table 2.** Functional categories of genes in the plasmid pSY3

No.	Functional category	No. of ORFs	ORFs
1	DNA replication	1	ORF23.3 (plasmid replication initiator protein-like protein)
2	Conjugative function	4	ORF17.3 (Ti-type conjugative transfer relaxase TraA), ORF18.3 (putative conjugal transfer protein, TraC), ORF19.3 (probable conjugal transfer protein traD, ORF22.3 (TraG family protein)
3	Transposition	11	ORF1.3 (transposases mutator type), ORF2.3 (ISRM2011-2 transposase protein), ORF4.3 (resolvase), ORF11.3 (transposase Tn3), ORF12.3 (resolvase domain), ORF14.3 (Putative resolvase), ORF30.3 (probable transposase protein), ORF31.3 (Site-specific recombinases), ORF36.3 (transposase IS3/IS911 family protein), ORF37.3 (IS66 Orf2 family protein), ORF38.3 (IS66 family element)
4	Plasmid stability/Partition	2	ORF8.3 (ParA), ORF25.3 (ATPase involved in chromosome partitioning-like protein)
5	Hypothetical protein	12	ORF5.3 (hypothetical protein CJA_1213), ORF6.3(hypothetical protein MpopDRAFT_5182), ORF9.3 (hypothetical protein S7335_720), ORF10.3 (hypothetical protein Cpham1DRAFT_3276), ORF13.3 (conserved hypothetical protein), ORF15.3 (hypothetical conserved protein), ORF16.3 (conserved hypothetical protein), ORF24.3 (hypothetical protein MchDRAFT_0088), ORF26.3 (hypothetical protein Swit_5119), ORF27.3 (conserved hypothetical protein), ORF33.3(hypothetical protein msr9263), ORF35.3 (hypothetical protein pNL1_p167)
6	Others	10	ORF3.3 (major facilitator superfamily transporter), ORF7.3 (predicted zinc finger protein 7), ORF20.3 (Multi-sensor Hybrid Histidine Kinase), ORF21.3 (virulence virD4-like protein), ORF28.3 (phosphoglycerate mutase), ORF29.3 (phage integrase family protein), ORF32.3 (CopG/Arc/MetJ family), ORF34.3 (Putative monooxygenase, FAD-binding) ORF39.3 (yecA family protein), ORF40.3 (TonB-dependent receptor)

The ORF4.3-encoded 259-aa protein showed some similarity to *Ochrobactrum tritici* resolvase. The ORF11.3-encoded 206-aa protein showed 39% identity with the transposase Tn3 of *Mesorhizobium* sp. BNC1. The ORF12.3-encoded 201-aa protein showed 73% identity with the resolvase domain of *Methylobacterium populi* BJ001, while the ORF14.3-encoded a 206-aa protein that possessed a high similarity (97%) to the resolvase found in *Sphingomonas* sp. SKA58. ORF30.3 (a 1,000-aa protein) was the longest transposable element, showing 61% identity with *Sinorhizobium meliloti*. ORF36.3 (a 135-aa protein) was 50% identical to the transposase IS3/IS911 family protein of *Caulobacter* sp. K31. ORF37.3 (a 117-aa protein) was 60% identical with the IS66 family protein of *Sinorhizobium medicae* WSM419, while ORF38.3 (515-aa) was 55% identical to the IS66 family element of *Brucella abortus* biovar 1 str. 9-941. ORF31.3 (190-aa) was 79% identical to site-specific recombinase of the DNA invertase Pin homolog of *Magnetospirillum magnetotacticum* MS-1. Bacterial transposons play an important role in the degradation of xenobiotic compounds, due to their catabolic activities. Maet and Heinaru (1994) reported that the transposons Tn4652 and Tn3614 play a key role in

the evolution of *Pseudomonas* and spread of new catabolic plasmids in nature.

We found two ORFs for plasmid stability/partitioning (Table 2). ORF8.3 and ORF25.3 encode 227-aa and 217-aa proteins which showed 53% and 88% identity with the ParA-like protein of *Gluconobacter oxydans* 621H and the ATPase of *Methylobacterium chloromethanicum* CM4, respectively (Table 1). The pSY3 plasmid-encoded proteins ParA and ATPase are required for partitioning prior to cell division. These proteins provide the energy for actively partitioning plasmids into daughter cells upon cell division. These proteins provide the energy for actively partitioning plasmids into daughter cells upon cell division (Gerdes *et al.*, 2000). A plasmid's partition system is essential to its stability. Partitioning is due to the positioning reaction and ensures at least one copy of the plasmid is positioned within each new daughter cell (Funnell and Gagnier, 1995). Abeles *et al.* (1985) reported that the P1 partitioning system consists of two genes, *parA* and *parB*, and a site called *parS*. Therefore, as the pSY3 plasmid contains ParA and ATPase, it possesses stability during cell division, thus ensuring the delivery of at least one copy to each daughter cell plasmid. The ATPase activity of ParA is required only for autoregulation, and ATPase is also involved in some



**Fig. 1.** Genetic map of pSY3 (35,735 bp). The genetic map represented by the line is marked with kb coordinates. ORF running from left to right or from right to left are indicates by boxes above and below the line, respectively.

other important step in partition (Davis *et al.*, 1992). ParA has conserved sequence motifs characteristic of a diverse range of proteins that bind and hydrolyze ATP (Walker *et al.*, 1982). There were twelve hypothetical proteins on pSY3 which showed homology to previously identified hypothetical proteins of different organisms (Table 1 and 2).

We categorized ten ORFs as having other functions. ORF3.3 encoded a 475-aa protein, which showed 35% identity with the major facilitator superfamily transporter protein of marine gamma Proteobacterium HTCC2143. ORF7.3 (39-aa) possesses 48% similarity with zinc finger protein, which is extraordinarily diverse, functioning in DNA recognition, RNA packaging, transcriptional activation, regulation of apoptosis, protein folding and assembly, and lipid binding. ORF20.3 (167-aa) showed 38% identity with the multi-sensor hybrid histidine kinase of *Stigmatella aurantiaca* DW4/3-1. ORF21.3 encoded a 210-aa protein that had 70% identity with the virulence (VirD4) protein of *Agrobacterium tumefaciens* str. C58. ORF28.3 encoded a protein (45-aa) that showed 62% identity with the phosphoglycerate mutase of *Sphingomonas wittichii* RW1. ORF29.3 encoded a protein that was similar (87%) to the Phage integrase family protein of *Sphingomonas wittichii* RW1. ORF32.3 was 67% identical to the CopG/Arc/MetJ family of *Gluconacetobacter diazotrophicus* PAI 5. ORF34.3 encoded a 552-aa protein that possessed 52% identity with a monooxygenase, the FAD-binding protein of *Bradyrhizobium* sp. BTai1. ORF39.3 encoded a protein (202-aa) that showed 39% identity with the YecA family protein of *Caulobacter* sp. K31. ORF40.3 encoded a 671-aa protein that had 49% identity with the TonB-dependent receptor protein of *Novosphingobium aromaticivorans* DSM 12444.

### Acknowledgements

This work was supported by a research grant of the Chungbuk National University in 2007.

### References

- Abeles, A.L., S.A. Friedman, and S.J. Austin. 1985. Partition of unit-copy miniplasmids to daughter cells. III. The DNA sequence and functional organization of the P1 partition region. *J. Mol. Biol.* 185, 261-272.
- Basta, T., A. Keck, J. Klein, and A. Stolz. 2004. Detection and characterization of conjugative degradative plasmids in xenobiotic-degrading *Sphingomonas* strains. *J. Bacteriol.* 186, 3862-3872.
- Berg, G. and G. Ballin. 1994. Bacterial antagonists to *Verticillium dahliae* Kleb. *J. Phytopathol.* 141, 99-110.
- Burenjargal, M., Y.S. Lee, J.M. Yoo, Y.C. Kim, Y.M. Lee, S. Oh, Y.P. Yun, J.T. Hong, Y.B. Chung, D.C. Moon, and H.S. Yoo. 2007. Endogenous sphingolipid metabolites related to the growth in *Sphingomonas chungbukensis*. *Arch. Pharm. Res.* 30, 317-322.
- Davis, M.A., K.A. Martin, and S.J. Austin. 1992. Biochemical activities of the ParA partition protein of the P1 plasmid. *Mol. Microbiol.* 6, 1141-1147.
- Funnell, B.E. and L. Gagnier. 1995. Partition of P1 plasmids in *Escherichia coli* MukB chromosomal partition mutants. *J. Bacteriol.* 177, 2381-2386.
- Gerdes, K., J. Moller-Jensen, and R.B. Jensen. 2000. Plasmid and chromosome partitioning: surprises from phylogeny. *Mol. Microbiol.* 37, 455-466.
- Goldberg, S.M., J. Johnson, D. Busam, T. Feldblyum, S. Ferriera, R. Friedman, A. Halpern, H. Khouri, S.A. Kravitz, F.M. Lauro, K. Li, Y.H. Rogers, R. Strausberg, G. Sutton, L. Tallon, T. Thomas, E. Venter, M. Frazier, and J.C. Venter. 2006. A sanger/pyrosequencing hybrid approach for the generation of high-quality draft assemblies of marine microbial genomes. *Proc. Natl. Acad. Sci. USA* 103, 11240-11245.
- Maet, A.A. and A.L. Heinaru. 1994. Transposon-mediated mobilization of chromosomally located catabolic operons of the CAM plasmid by TOL plasmid transposon Tn4652 and CAM plasmid transposon Tn3614. *Microbiology* 140, 915-921.
- Shin, H.J., S.J. Kim, and Y.C. Kim. 1997. Sequence analysis of the *phnD* gene encoding 2-hydroxymuconic semialdehyde hydrolase in *Pseudomonas* sp. strain DJ77. *Biochem. Biophys. Res. Commun.* 232, 288-291.
- Walker, J.E., M. Saraste, M.J. Runswick, and N.J. Gay. 1982. Distantly related sequences in the  $\alpha$ - and  $\beta$ -subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* 1, 945-951.
- Yeon, S.M., B.S. Choi, and Y.C. Kim. 2008. Organization of three rRNA (*rrn*) operons from *Sphingobium chungbukense* DJ77. *J. Microbiol.* 46, 697-703.