

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

SUPPLEMENTARY MATERIALS

Characterization of three novel SXT/R391 integrating conjugative elements ICEMfuInd1a and ICEMfuInd1b, and ICEMprChn1 identified in the genomes of *Marinomonas fungiae* JCM 18476^T and *Marinomonas profundimaris* strain D104.

Journal: Frontiers in Microbiology

Authors: Jhasketan Badhai and Subrata K Das*

Addresses: Department of Biotechnology, Institute of Life Sciences, Nalco Square, Bhubaneswar, India.

***Correspondence:**

Institute of Life Sciences

Department of Biotechnology

Nalco Square, Bhubaneswar 751023 India

E-mail: subratkdas@hotmail.com; subrata@ils.res.in

Phone: (+91) 674 230 3342

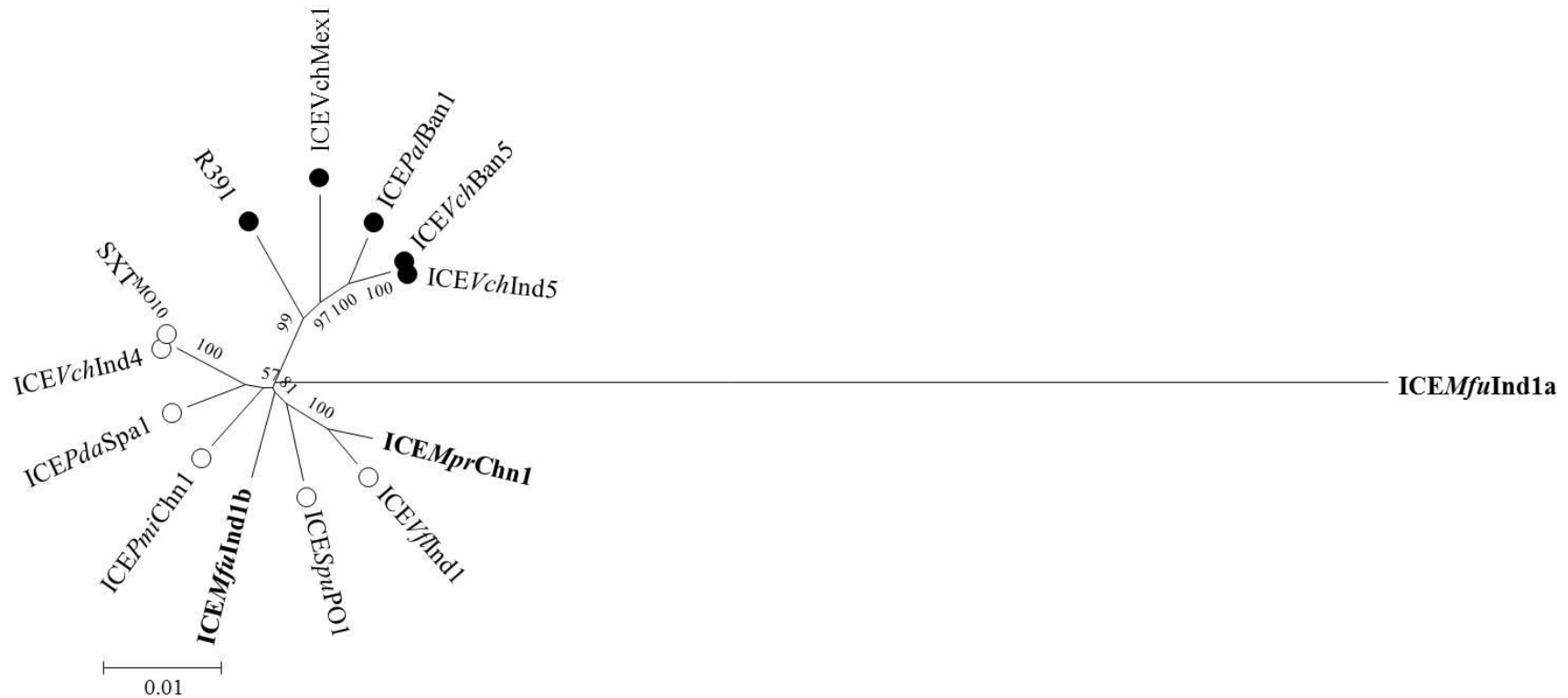
Fax: (+91) 674 230 0728

STXM010	GCT CTG TTT GGC GGC GGA TGA CCT AGT CAA AAA AAT CGA GAC GCG AAA CGT CGT TTG CAT TCT GGC CTG AAC TCG CCA
ICEVchInd4
ICEPdaSpa1
ICEPmiChn1T.G.....C.....T.....T.....
ICEMfuInd1bA.....C.....
ICESpuPO1	..G.....TG..C.....
ICEVfIInd1T.G.....C.....T.....
ICEMprChn1A.TG..C.....T.....
R391T.....TG..C.....
ICEVchMex1
ICEPaIBan1T.G.....C.....T.....T.....
ICEVchBan5T.G.....C.....T.....T.....
ICEVchInd5T.G.....C.....T.....T.....
ICEMfuInd1aT.....T.....
STXM010	AAC GGT TTG TAT CTT CAT GAC GAT ACG TCG TTT TAG GCG TTT TTA AGT GAA ATC GGG CTG TAT CCC TTG TCA GGT ATG
ICEVchInd4
ICEPdaSpa1
ICEPmiChn1G.....T.....T.....A.C.....G.....
ICEMfuInd1bT.....G.A.....T.....C.....C.....
ICESpuPO1T.....
ICEVfIInd1G.....T.....T.....A.C.....G.....
ICEMprChn1T.....T.....
R391G.....T.....C.....A.....A.....T.....
ICEVchMex1T.....T.....
ICEPaIBan1G.....T.....T.....A.C.....G.....
ICEVchBan5G.....T.....T.....A.C.....G.....
ICEVchInd5G.....T.....T.....A.C.....G.....
ICEMfuInd1aT.....G.....T.....C.....C.....
STXM010	GGA TTG CGC GAG TTG ATT TAT ATC GAG ACG CCA AAC AGT GAT TGT TAC GGC AGT TTT ACG TTT GGC GTT TCG ATC CAA
ICEVchInd4
ICEPdaSpa1
ICEPmiChn1A.....C.....G.....
ICEMfuInd1b
ICESpuPO1G.....T.....
ICEVfIInd1A.....C.....G.....C.....
ICEMprChn1	..G.....C.....
R391
ICEVchMex1
ICEPaIBan1A.....C.....G.....
ICEVchBan5A.....C.....G.....
ICEVchInd5A.....C.....G.....
ICEMfuInd1a
STXM010	AAG CCA AAC GGA TAG TGG TTT TGG CTT TTG GCG TTA ATT GGA TGG GGA AAT TGG TTT GGT AGA AA
ICEVchInd4
ICEPdaSpa1
ICEPmiChn1A.....
ICEMfuInd1bT.....
ICESpuPO1G..A.....
ICEVfIInd1A.....
ICEMprChn1
R391A.....T.....
ICEVchMex1
ICEPaIBan1A.....
ICEVchBan5A.....
ICEVchInd5A.....
ICEMfuInd1a

Fig. S1

23 **Supplementary Figure S1.** ClustalW alignment on MEGA6 of the 299-bp long putative *oriT* DNA sequence located within the
24 intergenic region *srpM-mobI* in ICEMfuInd1a, ICEMfuInd1b and ICEMprChn1 with the corresponding sequences from 11 reference
25 SXT/R391 ICEs. Reference sequences were retrieved from the GenBank: SXT^{MO10} (*V. cholerae* O139; accession: AY055428),
26 ICEVchInd4 (*V. cholerae* O139; accession: GQ463141), ICEVchInd5 (*V. cholerae* O1; accession: GQ463142), ICEVchBan5 (*V.*
27 *cholerae* O1; accession: GQ463140), ICEVchMex1 (*V. cholerae* non O139; accession: GQ463143), R391 (*Providencia rettgeri*;
28 accession: AY090559), ICEPalBan1 (*Providencia alcalifaciens*; accession: GQ463139), ICEVflInd1 (*V. fluvialis*; accession:
29 GQ463144), ICEPdaSpa1 (*Photobacterium damsela*; accession: AJ870986), ICESpuPO1 (*Shewanella putrefaciens*; accession:
30 CP000503), and ICEPmiChn1 (*Proteus mirabilis*; accession: KT962845).

31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51



53

54

55 **Supplementary Figure S2.** Phylogenetic analysis of core ICE genes encoded proteins. The unrooted tree was constructed by
 56 applying the Maximum Likelihood method based on the Poisson correction model using the MEGA6. Bootstrap analysis with 1000
 57 replications was performed to test the reliability of the tree. (●), ICEs of the R391 exclusion group; (○), ICEs of the SXT exclusion
 58 group.

59

60

61

62

63 **Supplementary Table S1.** Predicted core genes or ORFs present in the SXT-like ICEs: **(A)**, ICEMfuInd1a and **(B)**, ICEMfuInd1b
 64 derived from the genome of *M. fungiae* JCM 18476^T; and **(C)**, ICEMprChn1 of *M. profundimaris* strain D104.

65

66 **(A)** Core genes in ICEMfuInd1a

ICEMfuInd1a gene/ORF id	Putative gene product name	SXT/R391 equivalent gene	GenBank accession no. of <i>M. fungiae</i> genomic scaffolds encoding the ICE
Ga0061065_11949	Transcriptional regulator, AlpA family	<i>xis</i>	LIQF01000019.1
Ga0061065_11948*	Integrase, N-terminal region	<i>int</i>	LIQF01000019.1
Ga0061065_11947*	Integrase	<i>int</i>	LIQF01000019.1
Ga0061065_11946	Hypothetical protein	<i>srpR</i>	LIQF01000019.1
Ga0061065_11945	Plasmid segregation protein	<i>srpM</i>	LIQF01000019.1
Ga0061065_11944	Hypothetical protein	<i>mobI</i>	LIQF01000019.1
Ga0061065_11943	DNA PolV/UmuCDNA repair protein	<i>rumB</i>	LIQF01000019.1
Ga0061065_11942	DNA PolV/UmuDDNA repair protein	<i>rumA</i>	LIQF01000019.1
Ga0061065_11941	DNA polymerase-3 subunit epsilon	<i>s024</i>	LIQF01000019.1
Ga0061065_11940	Hypothetical protein	<i>s025</i>	LIQF01000019.1
Ga0061065_11939	WYL domain containing protein	<i>s026</i>	LIQF01000019.1
Ga0061065_11932	Conjugal transfer pilus assembly protein	<i>traI</i>	LIQF01000019.1
Ga0061065_11931	Conjugal transfer pilus assembly protein	<i>traD</i>	LIQF01000019.1
Ga0061065_11930	Conjugal transfer pilus assembly protein	<i>s091</i>	LIQF01000019.1
Ga0061065_11929	Conjugal transfer pilus assembly protein	<i>s043/traJ</i>	LIQF01000019.1
Ga0061065_11926	Conjugal transfer pilus assembly protein	<i>traL</i>	LIQF01000019.1
Ga0061065_11925	Conjugal transfer pilus assembly protein	<i>traE</i>	LIQF01000019.1
Ga0061065_11924	Conjugal transfer pilus assembly protein	<i>traK</i>	LIQF01000019.1
Ga0061065_11923	Conjugal transfer pilus assembly protein	<i>traB</i>	LIQF01000019.1
Ga0061065_11922	Conjugal transfer pilus assembly protein	<i>traV</i>	LIQF01000019.1
Ga0061065_11921	Conjugal transfer pilus assembly protein	<i>traA</i>	LIQF01000019.1
Ga0061065_11914	Thiol:disulfide interchange protein	<i>s054/dsbC</i>	LIQF01000019.1
Ga0061065_11913	Type-IV secretion system protein	<i>traC</i>	LIQF01000019.1
Ga0061065_11912	Conserved hypothetical protein	<i>orf92</i>	LIQF01000019.1
Ga0061065_11911	Conjugal transfer pilin signal peptidase	<i>trhF</i>	LIQF01000019.1
Ga0061065_11910	Conjugal transfer pilus assembly protein	<i>traW</i>	LIQF01000019.1
Ga0061065_1199	Conjugal transfer pilus assembly protein	<i>traU</i>	LIQF01000019.1
Ga0061065_1198	Conjugal transfer pilus assembly protein	<i>traN</i>	LIQF01000019.1
Ga0061065_1196	Hypothetical protein	<i>s063</i>	LIQF01000019.1

ICEM _{fu} Ind1a gene/ORF id	Putative gene product name	SXT/R391 equivalent gene	GenBank accession no. of <i>M. fungiae</i> genomic scaffolds encoding the ICE
Ga0061065_1195	Hypothetical protein	<i>s089</i>	LIQF01000019.1
Ga0061065_1194	Single strand binding protein	<i>ssb</i>	LIQF01000019.1
Ga0061065_1193	Phage recombination protein	<i>bet</i>	LIQF01000019.1
Ga0061065_1192	Putative phage-type endonuclease	<i>exo</i>	LIQF01000019.1
Ga0061065_1191	AAA-ATPase domain, CobS	<i>s067</i>	LIQF01000019.1
Ga0061065_1331	Hypothetical protein	<i>s073</i>	LIQF01000033.1
Ga0061065_1333	Conjugal transfer pilus assembly protein	<i>traF</i>	LIQF01000033.1
Ga0061065_1334	Conjugal transfer pilus assembly protein	<i>traH</i>	LIQF01000033.1
Ga0061065_1335*	TraG-like protein, N-terminal region	<i>traG</i>	LIQF01000033.1
Ga0061065_11487*	Conjugal transfer mating pair stabilization protein	<i>traG</i>	LIQF01000033.1
Ga0061065_11486	Entry exclusion protein	<i>eex</i>	LIQF01000014.1
Ga0061065_11485	Flagellar transcriptional activator (FlhC)	<i>setC</i>	LIQF01000014.1
Ga0061065_11484	Flagellar transcriptional activator (FlhD)	<i>setD</i>	LIQF01000014.1
Ga0061065_11483	Transglycosylase SLT domain containing protein	<i>s082</i>	LIQF01000014.1
Ga0061065_11482	Hypothetical protein	<i>s083</i>	LIQF01000014.1
Ga0061065_11481	Hypothetical protein	<i>s084</i>	LIQF01000014.1
Ga0061065_11480	Putative antitoxin of toxin-antitoxin system	<i>croS/setQ</i>	LIQF01000014.1
Ga0061065_11479	Phage repressor protein	<i>setR</i>	LIQF01000014.1

67

68 (B) Core genes in ICEM_{fu}Ind1b

ICEM _{fu} Ind1b gene/ORF id	Putative gene product name	SXT/R391 equivalent gene	GenBank accession no. of <i>M. fungiae</i> genomic scaffolds encoding the ICEs
Ga0061065_12222	Integrase	<i>int</i>	LIQF01000022.1
Ga0061065_12221	Hypothetical protein	<i>srpR</i>	LIQF01000022.1
Ga0061065_12220	Plasmid segregation protein	<i>srpM</i>	LIQF01000022.1
Ga0061065_12219	Hypothetical protein	<i>mobI</i>	LIQF01000022.1
Ga0061065_12218	DNA PolV/UmuCDNA repair protein	<i>rumB</i>	LIQF01000022.1
Ga0061065_12217	DNA PolV/UmuDDNA repair protein	<i>rumA</i>	LIQF01000022.1
Ga0061065_12216	Hypothetical protein	<i>s024</i>	LIQF01000022.1
Ga0061065_1221*	Conjugal transfer pilus assembly protein	<i>traI</i>	LIQF01000022.1
Ga0061065_1231*	Conjugal transfer pilus assembly protein	<i>traE</i>	LIQF01000023.1
Ga0061065_1232	Conjugal transfer pilus assembly protein	<i>traK</i>	LIQF01000023.1
Ga0061065_1233	Conjugal transfer pilus assembly protein	<i>traB</i>	LIQF01000023.1
Ga0061065_1234	Conjugal transfer pilus assembly protein	<i>traV</i>	LIQF01000023.1

ICEM <i>fu</i> Ind1b gene/ORF id	Putative gene product name	SXT/R391 equivalent gene	GenBank accession no. of <i>M. fungiae</i> genomic scaffolds encoding the ICEs
Ga0061065_1235	Conjugal transfer pilus assembly protein	<i>traA</i>	LIQF01000023.1
Ga0061065_1238	Thiol:disulfide interchange protein	<i>s054/dsbC</i>	LIQF01000023.1
Ga0061065_1239	Type-IV secretion system protein	<i>traC</i>	LIQF01000023.1
Ga0061065_12310	Conserved hypothetical protein	<i>orf92</i>	LIQF01000023.1
Ga0061065_12311	Conjugal transfer pilin signal peptidase	<i>trhF</i>	LIQF01000023.1
Ga0061065_12312	Conjugal transfer pilus assembly protein	<i>traW</i>	LIQF01000023.1
Ga0061065_12313	Conjugal transfer pilus assembly protein	<i>traU</i>	LIQF01000023.1
Ga0061065_12314	Conjugal transfer pilus assembly protein	<i>traN</i>	LIQF01000023.1
Ga0061065_12321	Hypothetical protein	<i>s063</i>	LIQF01000023.1
Ga0061065_12322	Hypothetical protein	<i>s089</i>	LIQF01000023.1
Ga0061065_12323	Single strand binding protein	<i>ssb</i>	LIQF01000023.1
Ga0061065_12324	Phage recombination protein	<i>bet</i>	LIQF01000023.1
Ga0061065_12325	Hypothetical protein	<i>orfZ</i>	LIQF01000023.1
Ga0061065_12326	Putative phage-type endonuclease	<i>exo</i>	LIQF01000023.1
Ga0061065_12327	Cobaltochelataase, CobS	<i>s067</i>	LIQF01000023.1
Ga0061065_12328	Hypothetical protein	<i>s088</i>	LIQF01000023.1
Ga0061065_13011	von Willebrand factor type-A domain containing protein	<i>s070</i>	LIQF01000030.1
Ga0061065_13010	RadC-like JAB domain	<i>s071</i>	LIQF01000030.1
Ga0061065_1309	Hypothetical protein	<i>s092</i>	LIQF01000030.1
Ga0061065_1308	Uncharacterized domain associated with phage or plasmid primase	<i>s072</i>	LIQF01000030.1
Ga0061065_1307	Hypothetical protein	<i>s073</i>	LIQF01000030.1
Ga0061065_1303	Conjugal transfer pilus assembly protein	<i>traF</i>	LIQF01000030.1
Ga0061065_1302	Conjugal transfer pilus assembly protein	<i>traH</i>	LIQF01000030.1
Ga0061065_1301*	TraG-like protein, N-terminal region	<i>traG</i>	LIQF01000030.1
Ga0061065_109135*	Conjugal transfer mating pair stabilization protein	<i>traG</i>	LIQF01000009.1
Ga0061065_109134	Entry exclusion protein	<i>eex</i>	LIQF01000009.1
Ga0061065_109133	Flagellar transcriptional activator (FlhC)	<i>setC</i>	LIQF01000009.1
Ga0061065_109132	Flagellar transcriptional activator (FlhD)	<i>setD</i>	LIQF01000009.1
Ga0061065_109131	Transglycosylase SLT domain containing protein	<i>s082</i>	LIQF01000009.1
Ga0061065_109130	Hypothetical protein	<i>s083</i>	LIQF01000009.1
Ga0061065_109129	Hypothetical protein	<i>s084</i>	LIQF01000009.1
Ga0061065_109128	Putative antitoxin of toxin-antitoxin system	<i>croS/setQ</i>	LIQF01000009.1
Ga0061065_109127	Phage repressor protein	<i>setR</i>	LIQF01000009.1

71 (C) Core genes in ICE*MprChn1*

ICEMprChn1 gene/ORF id	Putative gene product name	SXT/R391 equivalent gene	GenBank accession no. of <i>M. profundimaris</i> D104 genomic scaffolds encoding the ICEs
D104_12225	Cro/C1-type helix-turn-helix domain	<i>hipB</i>	AYOZ01000034.1
D104_12220	Serine/threonine-protein kinase HipA	<i>hipA</i>	AYOZ01000034.1
D104_12210	Transcriptional regulator, AlpA family	<i>xis</i>	AYOZ01000034.1
D104_12205	Integrase	<i>int</i>	AYOZ01000034.1
D104_12200	Hypothetical protein	<i>srpR</i>	AYOZ01000034.1
D104_12195	Plasmid segregation protein	<i>srpM</i>	AYOZ01000034.1
D104_12190	Hypothetical protein	<i>mobI</i>	AYOZ01000034.1
D104_12185	DNA PolV/UmuCDNA repair protein	<i>rumB</i>	AYOZ01000034.1
D104_12180	DNA PolV/UmuCDNA repair protein, N-terminal	<i>rumB</i>	AYOZ01000034.1
D104_12175	DNA PolV/UmuDDNA repair protein	<i>rumA</i>	AYOZ01000034.1
D104_09920	DNA polymerase-3 subunit epsilon	<i>s024</i>	AYOZ01000017.1
D104_09925	Hypothetical protein	<i>s025</i>	AYOZ01000017.1
D104_09930	WYL domain containing protein	<i>s026</i>	AYOZ01000017.1
D104_09960	Conjugal transfer pilus assembly protein	<i>traI</i>	AYOZ01000017.1
D104_09965	Conjugal transfer pilus assembly protein	<i>traD</i>	AYOZ01000017.1
D104_09970	Conjugal transfer pilus assembly protein	<i>s091</i>	AYOZ01000017.1
D104_09975	Conjugal transfer pilus assembly protein	<i>s043/traJ</i>	AYOZ01000017.1
D104_10460	Conjugal transfer pilus assembly protein	<i>traL</i>	AYOZ01000022.1
D104_10465	Conjugal transfer pilus assembly protein	<i>traE</i>	AYOZ01000022.1
D104_10470	Conjugal transfer pilus assembly protein	<i>traK</i>	AYOZ01000022.1
D104_10475	Conjugal transfer pilus assembly protein	<i>traB</i>	AYOZ01000022.1
D104_10480	Conjugal transfer pilus assembly protein	<i>traV</i>	AYOZ01000022.1
D104_10485	Conjugal transfer pilus assembly protein	<i>traA</i>	AYOZ01000022.1
D104_04230	Thiol:disulfide interchange protein	<i>s054/dsbC</i>	AYOZ01000004.1
D104_04225	Type-IV secretion system protein	<i>traC</i>	AYOZ01000004.1
D104_04220	Conserved hypothetical protein	<i>orf92</i>	AYOZ01000004.1
D104_04215	Conjugal transfer pilin signal peptidase	<i>trhF</i>	AYOZ01000004.1
D104_04210	Conjugal transfer pilus assembly protein	<i>traW</i>	AYOZ01000004.1
D104_04205	Conjugal transfer pilus assembly protein	<i>traU</i>	AYOZ01000004.1
D104_04200	Conjugal transfer pilus assembly protein	<i>traN</i>	AYOZ01000004.1
D104_04185	Hypothetical protein	<i>s063</i>	AYOZ01000004.1
D104_04180	Hypothetical protein	<i>s089</i>	AYOZ01000004.1
D104_04175	Single strand binding protein	<i>ssb</i>	AYOZ01000004.1

ICEM _{pr} Chn1 gene/ORF id	Putative gene product name	SXT/R391 equivalent gene	GenBank accession no. of <i>M. profundimaris</i> D104 genomic scaffolds encoding the ICEs
D104_04170	Phage recombination protein	<i>bet</i>	AYOZ01000004.1
D104_04165	Hypothetical protein	<i>orfZ</i>	AYOZ01000004.1
D104_04160	Putative phage-type endonuclease	<i>exo</i>	AYOZ01000004.1
D104_04155	Cobaltochelataase, CobS	<i>s067</i>	AYOZ01000004.1
D104_04150	Hypothetical protein	<i>s088</i>	AYOZ01000004.1
D104_04145	Protein of unknown function (DUF3150)	<i>s068</i>	AYOZ01000004.1
D104_04140	Hypothetical protein	<i>s069</i>	AYOZ01000004.1
D104_04135	von Willebrand factor type-A domain containing protein	<i>s070</i>	AYOZ01000004.1
D104_04130	RadC-like JAB domain	<i>s071</i>	AYOZ01000004.1
D104_04125	Hypothetical protein	<i>s092</i>	AYOZ01000004.1
D104_04120	Plasmid primase	<i>s072</i>	AYOZ01000004.1
D104_04110	Hypothetical protein	<i>s073</i>	AYOZ01000004.1
D104_04105	Conjugal transfer pilus assembly protein	<i>traF</i>	AYOZ01000004.1
D104_04100	Conjugal transfer pilus assembly protein	<i>traH</i>	AYOZ01000004.1
D104_04095	Conjugal transfer mating pair stabilization protein	<i>traG</i>	AYOZ01000004.1
D104_04090	Entry exclusion protein	<i>eex</i>	AYOZ01000004.1
D104_04085	Flagellar transcriptional activator (FlhC)	<i>setC</i>	AYOZ01000004.1
D104_04080	Flagellar transcriptional activator (FlhD)	<i>setD</i>	AYOZ01000004.1
D104_04075	Transglycosylase SLT domain containing protein	<i>s082</i>	AYOZ01000004.1
D104_04070	Hypothetical protein	<i>s083</i>	AYOZ01000004.1
D104_04065	Hypothetical protein	<i>s084</i>	AYOZ01000004.1
D104_04060	Putative antitoxin of toxin-antitoxin system	<i>croS/setQ</i>	AYOZ01000004.1
D104_04055	Phage repressor protein	<i>setR</i>	AYOZ01000004.1

72

73 SXT/R391 ICE equivalent gene are based on TBLASTN homology search of genes in the two ICEs of *M. fungiae* with the
74 corresponding proteins encoded by the SXT/R391 ICEs.

75 *, partial sequence.

76

77

78

79

80

Gene name	ICEM <i>fu</i> Ind1a gene locus id	Percentage homology of predicted proteins with SXT/R391 ICE proteins												
		ICEM <i>fu</i> Ind1b	ICE <i>Pa</i> Ban1	ICE <i>Pda</i> Spa1	ICE <i>Pmi</i> Chn1	R391	ICE <i>Spu</i> PO1	SXT ^{MO10}	ICE <i>Vch</i> Ban5	ICE <i>Vch</i> Ind4	ICE <i>Vch</i> Ind5	ICE <i>Vch</i> Mex1	ICE <i>Vf</i> Ind1	ICEM <i>pr</i> Chn1
<i>mfuind1a-14</i>	Ga0061065_11915	n.h	n.h	97.35	n.h	n.h	n.h	n.h	n.h	n.h	n.h	n.h	n.h	n.h
<i>dsbC</i>	Ga0061065_11914	95.22	95.22	95.22	95.65	95.65	95.65	95.65	95.22	95.65	95.22	95.22	95.65	95.65
<i>traC</i>	Ga0061065_11913	95.74	95.74	95.87	95.99	95.99	95.87	96.12	95.99	96.12	95.99	96.12	95.87	95.74
<i>orf92</i>	Ga0061065_11912	84.35	85.22	ab	ab	ab	85.22	ab	85.22	85.22	85.22	85.22	88.17	84.35
<i>trhF</i>	Ga0061065_11911	84.62	85.21	84.02	85.21	85.8	84.62	82.84	84.62	82.84	84.62	85.8	83.43	83.43
<i>traW</i>	Ga0061065_11910	89.97	91.03	89.45	90.77	90.24	89.71	90.5	n.a	90.5	n.a	n.a	91.03	91.03
<i>traU</i>	Ga0061065_1199	92.98	93.27	92.98	93.57	92.4	93.27	93.57	93.86	93.85	93.86	93.57	92.69	92.69
<i>traN</i>	Ga0061065_1198	73.02	73.02	73.02	73.27	72.94	72.86	73.35	72.53	73.02	72.53	72.94	73.02	73.11
<i>mfuind1a-15</i>	Ga0061065_1197	n.h	n.h	n.h	n.h	n.h	79.37	80.62	81.94	80.62	81.94	n.h	n.h	n.h
<i>s063</i>	Ga0061065_1196	89.73	89.19	89.19	89.73	88.11	90.22	89.19	88.11	89.19	88.11	88.65	87.03	87.03
<i>s089</i>	Ga0061065_1195	69.44	70.37	70.37	ab	70.37	69.44	ab	69.44	68.52	69.44	65.74	68.09	70.37
<i>ssb</i>	Ga0061065_1194	91.37	91.37	91.37	91.37	90.65	90.65	91.37	90.65	91.37	90.65	91.37	90.65	90.65
<i>bet</i>	Ga0061065_1193	94.85	94.49	94.85	94.85	94.85	94.85	94.49	94.85	94.49	94.85	94.85	94.49	94.85
<i>exo</i>	Ga0061065_1192	97.93	98.22	98.52	97.93	97.63	97.93	98.52	97.93	98.52	97.93	98.22	98.52	98.52
<i>s067</i>	Ga0061065_1191	97	98	97	97	97	97	96	97	96	97	97	98	98
<i>s073</i>	Ga0061065_1331	87.02	87.02	ab	75.78	87.79	87.02	89.31	87.02	89.31	87.02	87.79	87.79	75.78
<i>mfuind1a-16</i>	Ga0061065_1332	n.h	n.h	n.h	n.h	98.3	n.h	n.h	n.h	n.h	n.h	n.h	n.h	n.h
<i>traF</i>	Ga0061065_1333	93.31	95.54	94.9	92.9	94.59	93.63	94.9	95.22	94.9	95.22	97.45	93.63	93.63
<i>traH</i>	Ga0061065_1334	98.7	98.92	99.13	98.48	99.13	99.35	99.13	98.7	99.13	98.7	98.7	98.27	98.7
<i>traG</i>	Ga0061065_1335	98.61	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]
<i>traG</i>	Ga0061065_11487	100	98.3 [#]	98.47 [#]	98.56 [#]	97.54 [#]	98.56 [#]	98.56 [#]	98.3 [#]	98.56 [#]	98.3 [#]	98.47 [#]	98.3 [#]	98.98 [#]
<i>eex</i>	Ga0061065_11486	87.41	61.54	85.31	87.41	60.56	87.41	85.31	60.84	85.31	60.84	60.14	86.71	87.41
<i>setC</i>	Ga0061065_11485	85.06	85.06	85.06	85.06	85.06	85.06	85.06	85.06	85.63	85.06	85.06	85.06	85.06
<i>setD</i>	Ga0061065_11484	90.91	89.9	90.91	90.91	90.91	89.9	90.91	89.9	90.91	89.9	88.89	89.9	90.91
<i>s082</i>	Ga0061065_11483	75.12	75.27	ab	75.27	75.82	75.27	75.27	75.27	75.27	75.27	n.a	75.27	75.82
<i>s083</i>	Ga0061065_11482	81.91	80.79	80.54	72.38	80	80.79	72.38 [#]	80.79	80.79	80.79	n.a	80.3	80.3
<i>s083</i>	Ga0061065_11482	-	-	-	-	-	-	88.76 [#]	-	-	-	-	-	-
<i>s084</i>	Ga0061065_11481	86.97	87.68	88.03	88.03 [#]	88.38	86.97	83.93	87.68	87.32	87.68	88.03	88.03	87.68
<i>s084</i>	Ga0061065_11481	-	-	-	84.52 [#]	-	-	-	-	-	-	-	-	-
<i>croS/setQ</i>	Ga0061065_11480	75	76.32	76.32	76.32	n.h	75	76.32	76.32	76.32	76.32	75	76.32	76.32
<i>setR</i>	Ga0061065_11479	78.14	78.14	78.14	78.14	78.14	78.14	78.14	78.14	78.14	78.14	78.14	78.14	78.14

90

91

92

93

94

95

96

97

98

99 (B) Homology of ICEMfuInd1b with other SXT/R391 ICEs

Gene name	ICEMfuInd1b gene locus id	Percentage homology of predicted proteins with SXT/R391 ICE proteins											
		ICEPalBan1	ICEPdaSpa1	ICEPmiChn1	R391	ICESpuPO1	SXT ^{MO10}	ICEVchBan5	ICEVchInd4	ICEVchInd5	ICEVchMex1	ICEVfInd1	ICEMprChn1
<i>int</i>	Ga0061065_12222	99.75	99.49	99.75	99.49	99.75	99.24	99.75	99.24	99.75	100	99.75	98.98
<i>s002</i>	Ga0061065_12221	95.51	97.75	98.88	96.63	95.51	98.88	92.98	98.88	95.51	93.26	95.51	96.63
<i>s003</i>	Ga0061065_12220	99.38	98.77	99.38	99.38	98.77	98.77	99.38	98.77	99.38	98.46	99.38	99.07
<i>mobI</i>	Ga0061065_12219	97.87	ab	ab	ab	99.29	ab	97.87	98.58	97.87	98.58	97.87	98.58
<i>rumB</i>	Ga0061065_12218	91.76 [#]	ab	91.76 [#]	97.01	96.92	91.49 [#]	91.76 [#]	91.49 [#]	91.76 [#]	96.81	91.76 [#]	97.07 [#]
<i>rumB</i>	Ga0061065_12218	97.44 [#]	-	95.6 [#]	-	-	95.79 [#]	98.72 [#]	98.72 [#]	98.72 [#]	-	98.72 [#]	98.7 [#]
<i>rumA</i>	Ga0061065_12217	97.99	97.32	97.32	98.66	99.33	98.66	98.66	98.66	98.66	98.66	98.66	98.66
<i>s024</i>	Ga0061065_12216	84.52	85.71	83.33	85.71	84.52	84.52	84.52	84.52	84.52	85.71	84.52	85.71
<i>Mfuind1b-11</i>	Ga0061065_1225	n.h	n.h	n.h	n.h	n.h	n.h	n.h	n.h	n.h	n.h	n.h	61.46
<i>Mfuind1b-11</i>	Ga0061065_1224	n.h	n.h	n.h	n.h	n.h	n.h	n.h	n.h	n.h	n.h	n.h	97.85
<i>traI</i>	Ga0061065_1221	93.92	92	93.44	92.96	93.44	91.68	92.8	91.68	92.8	92.64	n.a	92
<i>traE</i>	Ga0061065_1231	98.56	99.04	99.23	99.23	99.04	98.56	99.04	98.56	99.04	98.08	98.56	98.56
<i>traK</i>	Ga0061065_1232	95.97	96.32	95.32	95.65	96.32	95.32	95.3	95.3	95.3	94.63	96.64	96.66
<i>traB</i>	Ga0061065_1233	99.07	99.77	98.37	99.53	98.14	98.83	99.53	98.83	99.53	98.6	n.a	99.53
<i>traV</i>	Ga0061065_1234	96.84	96.84	97.37	97.89	97.37	96.84	97.89	96.84	97.89	97.89	96.24	96.32
<i>traA</i>	Ga0061065_1235	99.22	99.22	98.44	98.44	97.66	99.22	96.88	99.22	96.88	98.44	97.66	98.44
<i>mfuind1b-15</i>	Ga0061065_1236	n.h	n.h	n.h	96.23	n.h	n.h	n.h	n.h	n.h	92.86	n.h	n.h
<i>mfuind1b-16</i>	Ga0061065_1237	n.h	n.h	n.h	98.86	n.h	n.h	n.h	n.h	n.h	98.86	n.h	n.h
<i>dsbC</i>	Ga0061065_1238	99.13	99.13	99.57	99.57	99.57	99.57	99.13	99.57	99.13	99.13	99.57	99.57
<i>traC</i>	Ga0061065_1239	99	99.12	99.25	99.25	99.12	99.25	99.25	99.25	99.25	99.37	99.12	99
<i>orf92</i>	Ga0061065_12310	97.39	ab	ab	ab	98.26	ab	97.39	97.39	97.39	97.39	97.85	96.52
<i>trhF</i>	Ga0061065_12311	98.82	97.06	98.24	98.77	97.06	98.24	98.82	98.24	98.82	97.65	96.47	96.47
<i>traW</i>	Ga0061065_12312	95.99	95.72	97.06	96.52	94.92	96.26	n.a	96.26	n.a	n.a	95.72	95.72
<i>traU</i>	Ga0061065_12313	98.54	98.83	98.83	98.25	98.54	98.83	98.83	99.08	98.83	99.42	99.12	99.12
<i>traN</i>	Ga0061065_12314	98.13	98.29	98.13	97.89	97.89	97.91	97.56	97.8	97.56	96.67	98.46	98.54
<i>s063</i>	Ga0061065_12321	98	98.5	98	98	97.5	98.5	97.5	98.5	97.5	98.5	97	97
<i>s089</i>	Ga0061065_12322	99.07	95.37	ab	98.15	99.07	ab	92.59	97.22	92.59	93.52	98.94	99.07
<i>ssb</i>	Ga0061065_12323	97.84	100	100	97.84	99.28	96.4	97.12	96.4	97.12	98.56	99.28	99.28
<i>bet</i>	Ga0061065_12324	99.63	100	100	100	99.63	99.63	100	99.63	100	100	99.63	100
<i>orfZ</i>	Ga0061065_12325	97.92	-	-	-	-	-	96.53	99.31	96.53	98.61	97.92	93.62
<i>exo</i>	Ga0061065_12326	98.52	98.82	98.82	98.52	98.82	98.82	98.82	98.82	98.82	98.52	98.82	98.82
<i>s067</i>	Ga0061065_12327	99.69	99.37	99.69	99.06	99.37	99.69	99.37	99.69	99.37	99.69	99.37	99.69
<i>s088</i>	Ga0061065_12328	96.47	96.08	99.22	96.86	96.86	89.31	96.47	97.24	96.47	96.47	97.25	97.25
<i>s070</i>	Ga0061065_13011	97.56	96.34	97.56	100	100	98.78	98.78	98.78	98.78	97.56	97.56	97.56
<i>s071</i>	Ga0061065_13010	96.36	97.58	97.58	97.58	96.36	98.18	96.36	98.18	96.36	98.18	97.58	97.58
<i>s092</i>	Ga0061065_1309	95.58	ab	94.69	92.92	97.35	ab	94.69	97.35	94.69	96.46	94.69	94.69
<i>s072</i>	Ga0061065_1308	96.64	95.8	98.6	97.2	95.24	97.55	96.64	96.92	96.64	96.64	98.32	98.6
<i>s073</i>	Ga0061065_1307	89.79	ab	84.91	89.36	89.79	91.49	90.21	91.49	90.21	90.21	90.17	84.91
<i>traF</i>	Ga0061065_1303	95.54	95.22	93.87	97.77	96.82	95.86	95.22	95.86	95.22	93.31	96.5	94.59
<i>traH</i>	Ga0061065_1302	98.92	98.7	98.27	98.7	98.92	99.13	98.7	99.13	98.7	98.7	97.84	98.48
<i>traG</i>	Ga0061065_1301	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]	100 [#]
<i>traG</i>	Ga0061065_109135	98.89 [#]	99.17 [#]	99.03 [#]	97.5 [#]	98.89 [#]	98.89 [#]	98.89 [#]	98.89 [#]	98.89 [#]	99.03 [#]	98.75 [#]	98.89 [#]
<i>eex</i>	Ga0061065_109134	70.63	95.8	97.9	69.72	96.5	96.5	69.93	96.5	69.93	69.23	97.2	96.5
<i>setC</i>	Ga0061065_109133	99.44	100	99.44	99.44	100	98.87	99.44	99.44	99.44	100	99.44	100

Gene name	ICEMfuInd1b gene locus id	Percentage homology of predicted proteins with SXT/R391 ICE proteins											
		ICEPa/Ban1	ICEPdaSpa1	ICEPmiChn1	R391	ICESpuPO1	SXT ^{M010}	ICEVchBan5	ICEVchInd4	ICEVchInd5	ICEVchMex1	ICEVf/Ind1	ICEMprChn1
<i>s082</i>	Ga0061065_109131	100	ab	98.35	97.8	100	100	100	100	100	n.a	99.45	98.9
<i>s083</i>	Ga0061065_109130	99.09	98.51	97.54	98.02	99.09	97.54	99.09	99.09	99.09	n.a	98.18	98.64
<i>s083</i>	Ga0061065_109130	-	-	-	-	-	97.78 [#]	-	-	-	-	-	-
<i>s084</i>	Ga0061065_109129	98.62	98.27	98.62 [#]	98.27	99.65	99.42 [#]	98.62	98.96	98.62	98.62	98.62	96.89
<i>s084</i>	Ga0061065_109129	-	-	100 [#]	-	-	-	-	-	-	-	-	-
<i>croS/setQ</i>	Ga0061065_109128	97.59	97.59	96.39	n.h	100	97.59	97.59	97.59	97.59	96.39	97.59	96.39
<i>setR</i>	Ga0061065_109127	100	100	100	100	100	100	100	100	100	100	100	100
<i>prfC</i>	Ga0061065_109126	n.a	n.a	84.12	n.a	78.64	n.a	n.a	n.a	n.a	96.2	89.04	69.19

100

101 n.h, no homology based on BLASTP search with e-value cut-off of 1e-5, query coverage > 90 %, and identity > 40 %.

102 n.a, sequence not available.

103 [#], partial protein sequences.

104 ab, absent or deleted.

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126 **Supplementary Table S3.** Descriptions of genes or ORFs present in the hotspot regions of ICEMfuInd1a.
 127

Hotspot region	ICEMfuInd1a gene/ORF id	Gene name	Length (bp)	Hotspot gene product name	Accession no. of protein homolog in ICEberg database	% similarity with homolog
HS-5 (s026-traI)	Ga0061065_11938	mfuind1a-1	1560	Type-I restriction enzyme M subunit	<i>Vibrio cholerae</i> Mex1; ACV96414.1	96
	Ga0061065_11937	mfuind1a-2	1245	Type-I restriction enzyme S subunit	<i>Vibrio cholerae</i> Mex1; ACV96472.1	41
	Ga0061065_11936	mfuind1a-3	1176	-	-	n.h
	Ga0061065_11935	mfuind1a-4	3117	Type-I restriction enzyme R subunit	<i>Vibrio cholerae</i> Mex1; ACV96447.1	98
	Ga0061065_11934	mfuind1a-5	564	-	-	n.h
	Ga0061065_11933	mfuind1a-6	912	-	-	n.h
HS-1 (traJ-traL)	Ga0061065_11928	mfuind1a-7	588	Hypothetical protein	<i>Providencia rettgeri</i> ; AAM08045.1	96
	Ga0061065_11927	mfuind1a-8	240	Hypothetical protein	<i>Providencia rettgeri</i> ; AAM08045.1	96
HS-2 (traA-s054)	Ga0061065_11920	mfuind1a-9	507	Acetyltransferase (GNAT) family	<i>Vibrio cholerae</i> Mex1; ACV96420.1	96
	Ga0061065_11919	mfuind1a-10	267	Conserved hypothetical protein	<i>Vibrio cholerae</i> Mex1; ACV96425.1	95
	Ga0061065_11918	mfuind1a-11	2730	Hypothetical protein	<i>Vibrio cholerae</i> MJ-1236; ACQ61857.1	99
	Ga0061065_11917	mfuind1a-12	1110	Hypothetical protein	<i>Vibrio cholerae</i> MJ-1236; ACQ61856.1	98
	Ga0061065_11916	mfuind1a-13	1533	Hypothetical protein	<i>Vibrio cholerae</i> MJ-1236; ACQ61855.1	99
	Ga0061065_11915	mfuind1a-14	1446	Type 4 fimbriae expression regulatory protein pilR	<i>Vibrio cholerae</i> MJ-1236; ACQ61854.1	99
HS-4 (s073-traF)	Ga0061065_1197	mfuind1a-15	693	Endonuclease-1	<i>Vibrio cholerae</i> Ind5; ACV96406.1	82
HS-3 (traN-s063)	Ga0061065_1332	mfuind1a-16	1065	Hypothetical protein	<i>Providencia rettgeri</i> ; AAM08049.1	98

128
 129 n.h, No homology

130
 131
 132
 133
 134
 135
 136
 137

138 **Supplementary Table S4.** Descriptions of genes or ORFs present in the hotspot regions of ICE*Mfu*Ind1b.
 139

Hotspot region	ICE <i>Mfu</i> Ind1b gene/ORF id	Gene name	Length (bp)	Hotspot gene product name	Accession no. of protein homolog in ICEberg database	% similarity with homolog
HS-5 (s024- <i>traI</i>)	Ga0061065_12215	mfuind1b-1	1038	Transcriptional regulator, AraC family	<i>Polaromonas naphthalenivorans</i> CJ2; ABM39468.1	25
	Ga0061065_12214	mfuind1b-2	1485	Putative gamma-glutamyl-gamma-aminobutyraldehyde dehydrogenase	<i>Klebsiella pneumoniae</i> 342; ACI07089.1	27
	Ga0061065_12213	mfuind1b-3	1779	-	-	n.h
	Ga0061065_12212	mfuind1b-4	900	DNA-binding transcriptional regulator, LysR	<i>Bordetella petrii</i> ; CAP41590.1	35
	Ga0061065_12211	mfuind1b-5	1512	Histidine ammonia lyase	<i>Klebsiella pneumoniae</i> 342; ACI08660.1	32
	Ga0061065_12210	mfuind1b-6	2013	-	-	n.h
	Ga0061065_1229	mfuind1b-7	1272	-	-	n.h
	Ga0061065_1228	mfuind1b-8	1050	-	-	n.h
	Ga0061065_1227	mfuind1b-9	240	-	-	n.h
	Ga0061065_1226	mfuind1b-10	1254	-	-	n.h
	Ga0061065_1225	mfuind1b-11	885	Type-III restriction system methylase	<i>Bordetella petrii</i> ; CAP44585.1	37
	Ga0061065_1224	mfuind1b-12	3078	Type-III restriction enzyme	<i>Bordetella petrii</i> ; CAP44586.1	36
	Ga0061065_1223	mfuind1b-13	282	-	-	n.h
	Ga0061065_1222	mfuind1b-14	474	-	-	n.h
HS-1 (<i>traJ-traL</i>)	Sequence not known	-	-	-	-	-
HS-2 (<i>traA-s054</i>)	Ga0061065_1236	mfuind1b-15	507	Acetyltransferase (GNAT) family	<i>Vibrio cholerae</i> Mex1; ACV96420.1	93
	Ga0061065_1237	mfuind1b-16	267	Conserved hypothetical protein	<i>Vibrio cholerae</i> Mex1; ACV96425.1	99
HS-4 (s073- <i>traF</i>)	Ga0061065_12315	mfuind1b-17	684	-	-	n.h
	Ga0061065_12316	mfuind1b-18	897	-	-	n.h
	Ga0061065_12317	mfuind1b-19	1131	-	-	n.h
	Ga0061065_12318	mfuind1b-20	531	-	-	n.h
	Ga0061065_12319	mfuind1b-21	1317	Conserved hypothetical protein	<i>Bordetella petrii</i> ; CAP41687.1	22
	Ga0061065_12320	mfuind1b-22	324	-	-	n.h
HS-3 (<i>traN-s063</i>)	Ga0061065_1306	mfuind1b-23	1782	Signal transduction GGDEF/EAL domain protein	<i>Burkholderia multivorans</i> ATCC 17616; BAG42888.1	33
	Ga0061065_1305	mfuind1b-24	2400	Probable signaling protein	<i>Bordetella petrii</i> ; CAP41698.1	31
	Ga0061065_1304	mfuind1b-25	1557	Hypothetical protein	<i>Bordetella petrii</i> ; CAP41755.1	37

140
 141 n.h, No homology
 142

143
 144
 145
 146
 147

148 **Supplementary Table S5.** Descriptions of genes or ORFs present in the hotspot regions of ICE*MprChn1*.
 149

Hotspot region	ICE <i>MprChn1</i> gene/ORF id	Gene name	Length (bp)	Hotspot gene product name	Accession no. of protein homolog in ICEberg database	% similarity with homolog
HS-5 (<i>s026-traI</i>)	D104_09935	mprchn1-1	3333	DNA-dependent ATPase, SNF2 family protein	<i>Bordetella petrii</i> ; CAP44583.1	41
	D104_09940	mprchn1-2	687	Hypothetical protein	<i>Bordetella petrii</i> ; CAP44584.1	26
	D104_09945	mprchn1-3	1947	Type-III restriction system methylase	<i>Bordetella petrii</i> ; CAP44585.1	36
	D104_09950	mprchn1-4	3078	Type III restriction enzyme R subunit	<i>Bordetella petrii</i> ; CAP44586.1	36
	D104_09955	mprchn1-5	855	-	-	n.h
HS-1 (<i>traJ-traL</i>)	D104_09980	mprchn1-6	1557	Transposase	<i>Vibrio fluvialis</i> Ind1; ACV96587.1	99
	D104_09985	mprchn1-7	96	-	-	n.h
	D104_09995	mprchn1-8	219	-	-	n.h
	D104_10000	mprchn1-9	243	-	-	n.h
	D104_10005	mprchn1-10	708	-	-	n.h
	D104_10010	mprchn1-11	372	-	-	n.h
	D104_10015	mprchn1-12	693	TetR family transcriptional regulator	<i>Burkholderia xenovorans</i> LB400; ABE31807.1	30
	D104_10020	mprchn1-13	1047	Hypothetical protein	<i>Bacteroides ovatus</i> ; CAD89793.1	28
	D104_10025	mprchn1-14	3042	Transporter, hydrophobe/amphiphile efflux-1 (HAE1) family	<i>Shewanella</i> sp. W3-18-1; ABM23977.1	25
	D104_10030	mprchn1-15	615	-	-	n.h
	D104_10035	mprchn1-16	234	-	-	n.h
	D104_10040	mprchn1-17	1305	Putative transposase	<i>Streptococcus suis</i> SC84; CAZ51643.1	22
	D104_10445	mprchn1-18	1173	Conserved hypothetical protein	<i>Vibrio fluvialis</i> Ind1; ACV96594.1	99
	D104_10450	mprchn1-19	696	Transposition helper protein	<i>Vibrio fluvialis</i> Ind1; ACV96542.1	100
D104_10455	mprchn1-20	573	Conserved hypothetical protein	<i>Vibrio fluvialis</i> Ind1; ACV96517.1	100	
HS-2 (<i>traA-s054</i>)	D104_10490	mprchn1-21	834	Conserved hypothetical protein	<i>Vibrio fluvialis</i> Ind1; ACV96573.1	100
	D104_10495	mprchn1-22	702	Hypothetical protein	<i>Vibrio cholera</i> ; AAL59716.1	99
	D104_10500	mprchn1-23	408	MerR family transcriptional regulator	<i>Burkholderia xenovorans</i> LB400; ABE31843.1	99
	D104_10505	mprchn1-24	885	Heavy metal efflux pump, CDF family	<i>Burkholderia xenovorans</i> LB400; ABE31842.1	98
HS-4 (<i>s073-traF</i>)	D104_04195	mprchn1-25	434	Hypothetical protein	<i>Vibrio cholerae</i> MJ-1236; ACQ61846.1	99
	D104_04190	mprchn1-26	663	Hypothetical protein	<i>Proteus mirabilis</i> ; BAI48432.1	100

150
 151 n.h, No homology

152
 153