

Table S1 Information of isolates.

General information		Antibiotic susceptibility (mg/L)											MLST-oxford typing							Genomic characteristics								
Isolate ID	Specimen type	Sampling time	Imipenem	Meropenem	Cloxacillin	Cefepime	Ceftazidime	Cefoperazone-sulbactam	Piperacillin-tazobactam	Amikacin	Ciprofloxacin	Levofloxacin	Minocycline	MLST	gltA	gtrB	gtdB	cpn60	recA	gpi	rpoD	Number of assembled contigs	Size (Mb)	GC content (%)	Nucleotide composition (CDS)	SR A_bio project	SRA bio sample	GenBank accession number
A1	Sputum suction with bronchoscope	2018/4/23	>32	>32	1	256	128	64	>256	>256	>64	64	16	2211	123	33	22	22	97	3	1	40376	3	80	92	PRJNA640471	SA MN 15778733	CP06 0732
A2	Sputum suction with bronchoscope	2018/4/28	>32	>32	1	256	128	64	>256	>256	>64	128	16	2211	123	33	22	22	97	3	19903	4	39	73	63	PRJNA640471	SA MN 15778734	JAC KXF0 0000 0000
A3	Sputum suction with bronchoscope	2018/4/30	>32	32	2	128	128	64	>256	>256	>64	64	16	2211	123	33	22	22	97	3	19103	4	39	74	61	PRJNA640471	SA MN 15778735	JAC KXG 0000 0000 0

A1	Blood	20 18 /5/ 22	16	16	2	64	12 8	32	25 6	>2 56	>6 4	32	16	2 2 1 1	1 2 3 2	3 3	2 2	2 9 7	3 3	16 6	0 2 6 4 2 3. 7	3 7 6 9 4 4	6 3 7 3	PRJ NA6 404 71	SA MN 157 787 36	JAC KXH 0000 0000 0	
C2	Sputum suction with bronchos cope	20 17 /7/ 18	>3 2	>3 2	0. 5	25 6	>6 4	64	>2 56	>2 56	64	16	32	4 6 9	1 1 2	3 3	2 2	2 0 3	1 0 3	13 1	7 6 7 4 2 3. 7	3 8. 9 4	5 5 2 9	PRJ NA6 404 71	SA MN 157 787 37	JAC KXI0 0000 0000	
C4	Blood	20 17 /7/ 21	32	>3 2	0. 5	25 6	>6 4	64	>2 56	>2 56	64	16	32	4 6 9	1 1 2	3 3	2 2	2 0 3	1 0 3	20 8	7 8 9 8 9	3 3 5 6	5 5 9	PRJ NA6 404 71	SA MN 157 787 38	JAC KXJ0 0000 0000	
C6	Blood	20 17 /7/ 25	>3 2	>3 2	0. 5	25 6	>6 4	64	>2 56	>2 56	64	16	32	4 6 9	1 1 2	3 3	2 2	2 0 3	1 0 3	15 0	7 5 5 1	3 8. 9 7	5 2 3	6 3	PRJ NA6 404 71	SA MN 157 787 39	JAC KXK 0000 0000 0
L6	Sputum suction with bronchos cope	20 15 /2/ 25	>3 2	>3 2	1	12 8	>2 56	128	>2 56	>2 56	64	8	8	1 9 5	1 3 3	3 3	2 2	2 9 6	3 3	74 0	3. 8 7 0 9	3 3 5 6	5 5 6	PRJ NA5 398 47	SA MN 115 136 59	JAC ATT0 0000 0000	
L7	Sputum suction with	20 15 /3/ 3	>3 2	>3 2	1	>2 56	>2 56	256	>2 56	>2 56	64	8	8	1 9 5	1 3 3	3 3	2 2	2 9 6	3 3	77 1	3. 8 7 1	3 8. 9 0	5 6 7 7	6 398 47	SA MN 115	JAC ATS0 0000 0000	

	bronchos																				0																136					
	cope																				4															60						
	Sputum	20																			3.																	SA	JAC			
	suction	15	>3	32	1	12	>2	>25	>2	>2	64	8	8	1							8	3	3			PRJ											MN	ATR				
L8	with	/3/	2			8	56	6	56	56				9	1	3	3	2	2	9	7	8.	5	6		NA5											115	0000				
	bronchos	11												5						6	0	9	5			47											136	0000				
	cope																			6																	61	0				
																				3.																						
		20												1							8		3	3			PRJ											SA	JAC			
L9	Blood	15	>3	32	1	12	>2	128	>2	>2	64	8	8	9	1	3	3	2	2	9	6	8.	5	7		NA5													115	0000		
		/3/	2			8	56		56	56				5						6	9	8.	8	0		398													136	0000		
		17																		4		9	5			47													62	0		
																				8																						
																				3.																						
L1	Blood	20	>3	32	1	12	>2	128	>2	>2	64	8	8	9	1	3	3	2	2	9	7	8.	5	6		PRJ														SA	JAC	
0		/3/	2			8	56		56	56				5						6	0	8.	6	8		NA5															115	ATP0
		17																		6		9	5			47															136	0000
																				6																					63	0000

GGAAACCTGATCCAGC
CCTTGTTGAGGCGTTTT
TAAAAGAATATCCCGAA
GCGCGTAAGTCTATTG
AGCAAAAGGCCAAGAT
GCCATTAAGTGGTAGTT
TCAGCGGTGCAGAATT
TTATGTCGTAAATGCCT
TTATATTGCGGGCAGC
CAATGGACAAGAACAG
CCAGTCCGTTGGTCTAT
GCGTCCTCATTCAAAT
TTATTAGTCTGACGAAA
GAACAACGTGATCAAG
CTGATCATGATTTCTTA
TTTAAAGACATCAAAAA
ACGTTTGGCTGAAGGG
CCGCTCTATTGGGACTT
AGTATTGCAACTTGCCG
AGCCAGGAGACCCAGT
GAATGACCCATCTCAG
CCTTGGCCAAAAGACC
GTAAAGAAGTTATAGCC
GGAACATTAGAAGTCAA
AAATGTAAGTATGATCAGG
TAAATGGCGCTTGTCTG
GATATTAAGTTTACCC
AACTTTGGTGCCACCC
GGAATTGAGTTATCAGA
TGACCCAGTACTCGCA
GCACGAGCGGCGATCT
ATTCACAGTCTTATAAT
GCTCGATTACGTGAAAT

A1

2159
603

T

TTA
A

disruptiv
e_infra
me_inse
rtion

A1.peg
.2105

360_362 dup
TAA

Phe120_Lys1
21 ins Asn

TGGATTTGGCAAAGCA
ACGGATGCCGTGGGCA
AATAG
ATGATAAAAAATAAATA
TATTCACATCCGAAGCA
TTGATTCTCGCCTTGCA
GCACGATCGAGCAGTG
CTCGATTA ACTCGCGC
CTCAGGATTTACTTTGG
TTGAATTGTTAGTCGCG
ATCGCTATCTTTGCGGT
TTTATCTTTATTGGGGT
GGAAAATTTTCGATTAC
TTGCTCAAAGTTCGTGA
TCGTAATGCCGAACAT
GAAGTGCATTTATTTGA
ATTACAAGATGCTTATC
AACAAATTCCTTCGTGAT
ACTTTGCAGATTATCCC
TTTATCCGCAAACCAAG
GTGGTCAACTTCATCCA
GCTTTAGAAATCGATAA
TCAAATTTTACGTTTTA
GCAAAGCAGGCGTAAC
GGATCCTTTAAAACAAG
GCTTATCACCATTTGAA
CGAATTGAATATCGTTA
TGATGCAGACCAAAAAA
AATTATATCGTCTTAAA
TATACAAATTTGAATAC
TTCGAACAGAGAGCAA
CCTTTATCGAGTACTTT
GTTAAGTCAAGTCGAC

MIKNKYIHIRSIDSRLAAR
SSSARLTRASGFTLVELL
VAIAIFAVLSLLGWKIFDY
LLKVRDRNAEHEVHLFE
LQDAYQQILRDTLQIPLS
ANQGGQLHPALEIDNQIL
RFSKAGVTDPLKQGLSP
FERIEYRYDADQKLYRL
KYTNLNTSNREQPLSST
LLSQVDQYQIMVLTPQE
VTKWPEVNIDPTKPNEL
KKLPKGIKIQLTVAGVNY
EWIYSLNQGDLSLSQEG
NS

0 0 0 1 / / / / / / / /

NODE
_1_len
gth_26
8445_
cov_1
6.2871
09

5301
7

G

GCG
TGG
ACG
ATC
TGA
GTT
GAA
CTC
GCG
ACG
CGG
ACG
ATC
TTC
ACC
ACC
GAA
GCT
ACG
TTTT
GGA

conserv
ative_inf
rame_in
sertion

C1.pe
g.1222

1375_1376
ins
TCGTGAAG
GTGGTTTT
GAAGACCG
TCCACGTC
GTAGCTTC
GGTGGTGA
AGATCGTC
CACGTCGC
GAGTTCAA
CTCAGACC
GTCCACGT
CGTGAAGG
TGGTTTTG
AAGACCGT
CCAAAACG
TAGCTTCG
GTGGTGAA
GATCGTCC
GCGTCGCG

His458_Ala45
9 ins
ValValLysVal
ValLeuLysThr
ValHisValVal
AlaSerValVal
LysIleValHisV
alAlaSerSerT
hrGlnThrValHi
sValValLysVal
ValLeuLysThr
ValGlnAsnVal
AlaSerValVal
LysIleValArgV
alAlaSerSerT
hrGlnIleValHis

CAATATCAGATTATGGT
TTTGACTIONCGCAAGAA
GTCACAAAATGGCCCG
AGGTTAATATTGATCCA
ACGAAACCTAATGAGTT
AAAAAAATTACCGAAAG
GGATAAAAATCCAACCTT
ACAGTCGCTGGTGTA
ATTATGAGTGGATTTAT
AGCTTAAATCAGGGTG
ACTTATCGCTTTTCGCAG
GAAGGTAATTCATAA
ATGAGCAAACTTTTGC
CGAATTTTCTTTGCACG
AAACCTTACAACAAGCT
CTTGAAGGTTTAGGTTT
TACTAACCCAACTCCTG
TGCAAGAACAATCTATT
CCTGCTGCTTTAGAAG
GTAAAGACCTTCTTGTT
TCTAGCCAGACTGGTT
CTGGTAAAACCTGCTGC
ATTCTTATTACCGACAT
TACATAATTTGGCAGGT
CAAGATACTTTTCGTTCC
GTTCAAAGAACGTATGA
AAGCAGTGACTCAGCC
TAGCATTTTAGTTTTAT
GCCCAACACGTGAATT
GGCTCAACAAGTTAGC
CAAGACGCAATTGCTTT
TGTACGTACATGAAAG
GTGTTTCGTATTGCTGCA

MSKTFAEFSLHETLQQA
LEGLGFTNPTPVQEQSIP
AALEGKDLLVSSQTGSG
KTA AFL LPTLHNL AGQDT
FVPFKERMKA V TQPSILV
LCPTRELAQQVSQDAIA
FVRHMKGVRIAAIMGGM
PFGKQIQQLKGAQVVVA
TPGRLLDLVNRRLKLD
KVDALIVDEADRMLDLG
FSEDLEAISDLAANRGQT
LMFSATFADRIIRLAERM
MNEPERIAIETGHSTNTD
ITQTLHWT DGF EHK KLL
THWLADETL DQAVVFAS
TQEDTDMLAEELAEAGH
SVVALHGAMPQTVRNR
RLRSIREGRAKILVATDV
AARGLDVPTISHVINFG
LPMKHEDYVHRIGRTGRA
GRTGQAITLATYRERGI

/ / / / 0 1 1 / / / /

CGG	AGTTCAAC	ATTATGGGTGGTATGC	RALEDYLEARLNVSEIEG
TCT	TCAGATCG	CTTTTGGTAAACAAATT	LEPSPPPARPRREGGR
TCA	TCCACG	CAACAGCTTAAAGGTG	GGNGGRDGRRFEFEGGN
AAA		CTCAAGTTGTAGTTGCG	FKRREGGRDDRPRRSF
CCA		ACTCCAGGTCGTTTATT	DDKPRGERPAFGGEDR
CCT		AGACTTAGTTAACCGTC	PRREFNSDRPRREGGF
TCA		GTCAATTAAGCTAGAC	NDKPRFDSNDDNRGNR
CGA		AAAGTTGATGCTTTAAT	VDYKPRRENGFGDRPQ
CGT		CGTGGACGAAGCTGAC	RSFGGEDRPRREGGF
GGA		CGTATGTTAGACCTTGG	DRPKRSFGGEERPRRA
CGG		TTTCTCTGAAGACTTAG	VREEHFNQESRGERR
TCT		AAGCGATCAGTGACTT	KFDR
GAG		GGCTGCGAATCGTGGT	
TTG		CAAACGTTAATGTTCTC	
AAC		TGCAACTTTTGCTGACC	
TCG		GTATTATTCGTCTTGCT	
CGA		GAGCGCATGATGAATG	
CGT		AACCTGAGCGTATTGC	
GGA		AATTGAAACTGGTCACA	
CGA		GCACCAATACTGACATC	
TCT		ACCCAAACTTTTACTG	
TCA		GACAGATGGTTTCGAA	
CCA		CACAAGAAGAACTTCT	
CCG		GACTCACTGGTTGGCT	
AAG		GATGAAACGCTTGACC	
CTA		AAGCTGTTGTATTGCG	
CGA		AGCACACAAGAAGATA	
CGT		CAGATATGTTGGCTGAA	
GGA		GAACTTGCAGAAGCAG	
CGG		GTCACTCAGTTGTTGCA	
TCT		CTTCACGGTGCAATGC	
TCA		CGCAAACAGTTCGTAA	
AAA		CCGTCGTTTACGTAGTA	

CCA
CCT
TCA
CGA

TCCGTGAAGGCCGTGC
GAAAATCTTAGTTGCTA
CTGACGTTGCTGCACG
TGGTCTTGACGTTCCAA
CAATTCACACGTTATT
AACTTCGGCTTACCAAT
GAAGCATGAAGATTATG
TACACCGTATTGGTCGT
ACAGGTCGTGCTGGTC
GTACAGGTCAAGCGAT
TACTTTAGCGACTTATC
GTGAACGCGGTAAAAT
CCGTGCGCTTGAAGAC
TACTTAGAAGCTCGTTT
AAATGTTTCTGAAATTG
AAGGTTTAGAGCCATCT
CCACCTCCAGCTCGTC
CACGTCGTGAAGGTGG
TCGTGGTGGTAACGGT
GGCCGTGATGGTCGTC
GTTTCGAGGGTGAAGG
TAACTTCAAACGTCGTG
AAGGTGGTCGTGATGA
TCGTCCACGTCGTAGC
TTCGATGATAAGCCTCG
TGGTGAGCGTCCTGCT
TTCGGTGGCGAAGATC
GTCCACGTCGCGAGTT
CAACTCAGATCGTCCA
CGCCGTGAAGGTGGTT
TTAACGATAAGCCTCGT
TTTGATTCAAATGATGA
CAATCGTGGTAACCGC

NODE
 _13_le
 ngth_1
 16163
 cov
 17.981
 428

1158
 35

G A

upstrea
 m_gene
 _variant

C1.pe
 g.559

-897C>T

GTAGACTATAAACACG
 TCGTGAAAACGGTTTC
 GGTGACCGTCCACAAC
 GTAGCTTTGGTGGTGA
 AGATCGTCCACGTCGT
 GAAGGCGGTTTTGGTG
 ATCGTCCAAAACGTAG
 CTTTGGTGGTGAAGAA
 CGTCCACGTCGCGCAG
 TTCGCGAAGAACACTTT
 AATCAAGAATCTCGTGG
 TGAGCGCCGTCGTAAA
 TTCGACCGTTAA
 ATGCAAATGATTAGCCC
 GACTCGACTTGCTTCTT
 TTAATGATATGCAAGAC
 AGTAATTTCTTTACTCA
 GTTTTTAAATATCTGTT
 GTGGAAAACCTGTACA
 ACCAACTATAAAGAAT
 ATGTGTCTTTACAACAT
 GCTCTTTATGAAGGGG
 ATATTGAAATGGATAAG
 GTGATTGATTGGGTTAT
 GCAAAGTCCCAAAGAT
 CACCGAATAATTTTTGA
 GAAGATATTATTTCAAG
 GCCGTGATGAATTCTCT
 AAGCCTATACCTACTGA
 ATTAGAAAATTTCTTTAA
 TTACATTGAACAAAAAC
 CTGATTGGTTAGACCAA
 CAACAAATTGATGAGG

MQMISPTRLASFNDMQD
 SNFFTQFLNICCGKPVQ
 PNYKEYVSLQHALYEGD
 IEMDKVIDWVMQSPKDH
 RIIFEKILFQGRDEFKPI
 PTELENFFNYIEQKPDW
 LDQQQIDEAVKFTHRLGI
 NNGFILRDLSLMAGYLYP
 GFNQPLILTGALKKEAGT
 RLAETTKWWVDITEPQG
 LFRLSAGFTSTIYVRFIHA
 LVRRQLKKSSRWSEIW
 GIPLNQFDLAMTNLAFSS
 VLLGIRALGIWPTKQEA
 KSFLHFWRYVGWLMGID
 EKWLIQSEPEGWRLLYW
 MQFAHPRSDHSSIELGL
 SLSKEPFERKYLHLRSL
 QQKLAYRQHLELTQFFI
 GKCRMKLLGLPQQSAS

/ / / / 0 0 1 / / / /

CTGTAAATTTACCCAT	WFAYYLIVRNLLLLYNGAK
CGTTTAGGAATAAATAA	LSPKVEKFLSKSGRNIQK
TGGTTTTATTCTTAGAG	LGLTLYQNQGKAKTLAS
ACTTATCCTTAATGGCG	MHQ
GGATATTTATATCCGGG	
GTTTAATCAACCCCTAA	
TTTTAACGGGTGCATTA	
AAGAAAGAAGCGGGAA	
CACGTTTAGCAGAAAC	
CACAAAATGGTGGGTG	
GATATTACTIONAACCTCA	
AGGGCTTTTTCGTTTGA	
GTGCAGGTTTTACGTCA	
ACGATTTATGTTGCTT	
TATTCATGCACTCGTTC	
GCCGGCAATTAATAAAA	
GTCTAGCCGTTGGGAT	
AGTGAGATTTGGGGAA	
TTCCTCTTAATCAATTT	
GATTTAGCAATGACTAA	
TTTGGCTTTTAGCAGTG	
TCGTTTTACTCGGTATT	
AGAGCTTTAGGTATATG	
GCCAACAAAACAAGAG	
GCAAAGAGTTTTTTTACA	
TTTCTGGCGTTATGTGCG	
GCTGGTTAATGGGTATT	
GATGAAAAATGGCTTAT	
TCAGTCTGAACCTGAG	
GGTTGGAGTTACTTTA	
CTGGATGCAGTTTGCT	
CATCCTCGCTCCGATC	
ACAGCAGTATTGAATTA	

GGATTGAGCTTGTCGA
 AAGAACCTTTTGAGAGA
 AAATATCTGCATTTACG
 CTCTTTACAGCAGAAAC
 TTGCTTATCGACAACAT
 TTAGAACTTACCCAATT
 TTTTATTGGTAAGAAAA
 GAATGAAACTACTCGG
 ATTGCCTCAGCAATCTG
 CCTCATGGTTTGCTTAT
 TATTTAATTGTGCGTAA
 TTTACTACTTTATAATG
 GCGCAAAGCTTTCACC
 AAAAGTTGAAAAGTTTT
 TATCTAAATCAGGCCGT
 AATATACAAAACCTTGG
 CTTAACGCTTTATCAGA
 ATCAAGGTAAGGCAAA
 AACCTTAGCAAGTATGC
 ATCAATAA

ATGTTCTTACTGAAAAA MFLKLNLVSSISKVTQDL
 TCTCGTCTCTTCAATAT GNIVSITPVVNTGSSVNV
 CAAAAGTTACTCAGGAC NVSDINIANVSTTGLLSN
 CTGGGTAATATTGTATC VISTVTDTVSHTTDLVS
 GATAACCCCTGTAGTAA NVVGTVTGTVGSTSPID
 ACACAGGTTTCTCTGTT TVTNIIGGVTGGVTGNPL
 AATGTAAATGTTTCTGA EVVTDIIGGVTGGVVGG
 TATCAACATTGCTAATG TSPISPVIDVVQGGIDILQ
 TTTCAACAACAGGCTTA GVESLKTEIINTGIETVAD
 CTTTCTAATGTGATTTT AVIGVLPQAEHPVSEIAD
 AACTGTAACCGATACAG LGTLTFETSRDTVNGAL
 TTTACATACAACCACA EVVSDLVGGDIQGATES
 GACTTAGTTTCTAATGT ATGILDTLITNGTTATGIV

/ / / / 0 1 1 / / / / /

NODE
 _33_le
 ngth_2
 3883_ 1720 G A
 cov_1
 5.8555
 83
 upstrea
 m_gene
 _variant
 C1.pe
 g.2207
 -206C>T

AGTTGGTACCGTAACA	TGVVGSVTDVIGGVTGG
GGTACTGTAGGTTCTAC	VDGNPLEVITDIIGGVTG
AAGCCCAATTGATACTG	GVDGNPLEVITDIIGGVT
TAACCAATATTATTGGC	GGVGGDNPLGVVTDIIG
GGCGTAACGGGTGGCG	GVTGGIIGGGTSPISPVID
TTACAGGCAACCCTCTT	VVQGGIDILQGVESLKTE
GAAGTTGTTACTGACAT	IINTGIDTVADTIIGVLPQA
TATCGGCGGAGTAACG	EHPVSEIADLGLTLFETS
GGTGGTGTGTCGGCG	RDTVNGALEVVS DLVGG
GAACTTCTCCAATCTCT	DIQGATENATGILGLTIQ
CCAGTAATCGATGTTGT	NGSTATGIVTDILGGVTG
TCAAGGCGGCATTGAT	TIGGVTGGIGGDNPLGIV
ATTCTTCAAGGCGTTGA	TDIIGGVTGGIGGDNPLG
AAGCTTAAAACTGAAA	VVTDIIGGVTGGIGGDNP
TTATTAATACAGGCATT	LGVVTDIIGGVSGGGGG
GAAACTGTAGCTGATG	EYD
CAGTTATTGGTGATTA	
CCTCAAGCTGAGCATC	
CAGTTTCTGAAATTGCA	
GACCTTGGCACATTAAC	
TTTTGAACTTCTCGTG	
ACACAGTCAATGGTGC	
TCTTGAAGTTGTTTCTG	
ATTTAGTCGGTGGTGAT	
ATTCAAGGCGCTACTG	
AATCAGCGACTGGTATT	
CTCGATACGCTTATTAC	
AAATGGCACAACCTGCTA	
CAGGTATTGTTACTGGT	
GTGGTTGGAAGTGTTA	
CTGATGTTATTGGTGGC	
GTAAGTGGTGGTGTG	
ATGGCAACCCACTTGA	

AGTGATTACAGATATTA
TCGGTGGTGTAAGTGG
CGGTGTTGATGGCAAC
CCACTTGAAGTGATTAC
AGATATTATTGGCGGTG
TAACTGGTGGTGTGG
GGGTGATAACCCTCTT
GGCGTAGTAACTGATAT
TATCGGCGGCGTAACC
GGTGGCATTATTGGTG
GCGGTACATCTCCTATT
TCTCCAGTTATTGATGT
TGTTCAAGGTGGCATC
GATATTCTTCAAGGCGT
TGAAAGCCTAAAACTG
AAATTATTAACACTGGT
ATTGATACAGTAGCCGA
CACTATTATTGGTGTAT
TACCTCAAGCTGAGCAT
CCAGTTTCTGAAATTGC
AGACCTTGGCACATTAA
CTTTCGAAACGTCTCGT
GACACGGTCAATGGTG
CTCTTGAAGTTGTTTCT
GATTTAGTCGGTGGCG
ATATTCAAGGTGCTACT
GAAAATGCAACTGGTAT
TTTAGGAACACTTATTC
AGAATGGTTCAACAGC
GACTGGAATTGTTACTG
ATATTCTTGGTGGCGTA
ACAGGTAATATCGGCG
GCGTAACTGGTGGTAT

CGGTGGTGACAATCCG
CTCGGTATAGTAACTGA
CATTATCGGCGGTGTTA
CTGGCGGTATCGGCGG
TGATAATCCGCTTGGTG
TAGTGA CTGACATTATC
GGTGGTGT TACTGGCG
GTATCGGCGGTGATAA
TCCGCTTGGTGTAGTG
ACTGACATTATCGGAG
GTGTTAGTGGGGGTGG
TGGTGGTGAGTATGAT

NODE
_46_le
ngth_5
019_c 1880 T
ov_13
0.4757
18
TAA
GAT
CTT
GGT
TTA
TTA
ACT intergeni
TCT c_region
GTG
ATT
TCA
TTA
TCA
CG

/ / / / 0 0 1 / / / / /

NODE
_46_le
ngth_5
019_c 1980 G
ov_13
0.4757
18
A intergeni
c_region

/ / / / 1 0 0 / / / / /

NODE _165_l ength_ 137_c ov_17. 61666 7	123	G	A	synony mous_v ariant	C1.pe g.794	15C>T	Asp5Asp	GAGTTAGGCACAGACG GCAGCTTCAATGCACA AGTGGCGTTGGGACCA GATGCGCTTGACGGCA CAGTGGTGAATGTTAAC GGTACTAACTACACCGT GACAGCAGCTGACCTT GCAAACGGCTATATCA CAGCA GAGTTAGGCACAGACG GCAGCTTCAATGCACA AGTGGCGTTGGGACCA GATGCGCTTGACGGCA CAGTGGTGAATGTTAAC GGTACTAACTACACCGT GACAGCAGCTGACCTT GCAAACGGCTATATCA CAGCA	ELGTDGSFNAQVALGPD ALDGTVVNVNGTNYTVT AADLANGYITA	/ / / / 1 0 0 / / / / /
NODE _165_l ength_ 137_c ov_17. 61666 7	127	G	T	missens e_varian t	C1.pe g.794	11C>A	Thr4Lys	GAGTTAGGCACAGACG GCAGCTTCAATGCACA AGTGGCGTTGGGACCA GATGCGCTTGACGGCA CAGTGGTGAATGTTAAC GGTACTAACTACACCGT GACAGCAGCTGACCTT GCAAACGGCTATATCA CAGCA CCTGAAGACCTAAATG GCGACGGTATCTTGAA TGCAGACGAGTTAGGC ACAGACGGCAGCTTCA ATGCACAAGTGGCATT GGACCGGATGCAGTTG ATGGCACAGTGGTGAA TGTTAACGGTACTAACT ACACCGTGACAGCAGC TGACCTTGCAAACGGC TATATCACAGCAACTCT GGATGCAACAGCAGCC GATCCTGTGACTGGTC AGATTGTGATTCATGCT GAGGCTGTAGATGCGC	ELGTDGSFNAQVALGPD ALDGTVVNVNGTNYTVT AADLANGYITA	/ / / / 1 0 0 / / / / /
NODE _63_le ngth_4 64_co v_45.4 85788	411	G	A	synony mous_v ariant	C1.pe g.3027	54C>T	Asp18Asp	GAGTTAGGCACAGACG GCAGCTTCAATGCACA AGTGGCGTTGGGACCA GATGCGCTTGACGGCA CAGTGGTGAATGTTAAC GGTACTAACTACACCGT GACAGCAGCTGACCTT GCAAACGGCTATATCA CAGCA CCTGAAGACCTAAATG GCGACGGTATCTTGAA TGCAGACGAGTTAGGC ACAGACGGCAGCTTCA ATGCACAAGTGGCATT GGACCGGATGCAGTTG ATGGCACAGTGGTGAA TGTTAACGGTACTAACT ACACCGTGACAGCAGC TGACCTTGCAAACGGC TATATCACAGCAACTCT GGATGCAACAGCAGCC GATCCTGTGACTGGTC AGATTGTGATTCATGCT GAGGCTGTAGATGCGC	PEDLNLDGILNADELGT DGSFNAQVALGPD GTVVNVNGTNYTVTAAD LANGYITATLDATAADPV TGQIVIHAEAVDAQGNV DVADADVTLTIDTTPQDL ITAITVPEDLNLDGILNAA ELGTDGSFNAQVALGPD AVDGTVVNVNGTNYT	/ / / / 0 1 1 / / / / /

AAGGTAACGTTGATGTT
 GCTGACGCAGATGTAA
 CCTTAACCATTGATACA
 ACTCCACAAGACCTGAT
 TACAGCAATTACGGTTC
 CTGAAGACCTAAATGG
 CGACGGTATCTTGAAT
 GCTGCAGAACTTGGTA
 CAGATGGCAGCTTCAA
 TGCACAAGTGGCATT
 GGACCGGATGCAGTTG
 ATGGCACAGTGGTGAA
 TGTTAACGGTACTAACT
 ACACC

NODE
 _66_le
 ngth_3
 47_co 270 A G intergeni
 v_102. c_region
 73703
 7

/ / / / 0 1 1 / / / /

NODE
 _155_l
 ength_
 3160_ 114 T C upstrea
 cov_7 m_gene L6.peg
 8.8401 _variant .478 -124T>C
 87

ATGCCTTCAATTGAGCC
 AGCTTCACCGACAATCA
 ATTTCAAATTGAAGTG
 ATTAATCAGGTGAGTG
 GGGCGACAGTTGAAGC
 CGAACAACTGGACGAG
 CAAACAGTCCGGATCA
 TTGTTAAAGATGAACTG
 GATAAGCAGCTTCCAA
 GAACGGTACCTAAGCT
 TGTAAGTGATCAAATCG
 CAAATCCAACTCAACC

MPSIEPASPTINFKIEVIN
 QVSGATVEAEQLDEQTV
 RIIVKDELDKQLPRTVPK / / / / / / / 1 0 1 0 1
 LVSDQIANPNSTISRSLT
 ENTTARRNRT

ATTAGTCGGTCTTTGAC
 TGAGAATACGACAGCG
 AGAAGAAATCGTACTTA
 A

NODE									
_30_le									
ngth_1									
922_c	158	A	G	intergeni					/ / / / / / / 1 0 1 1 1
ov_27				c_region					
4.8293									
46									
NODE									
_30_le									
ngth_1									
922_c	169	A	G	intergeni					/ / / / / / / 1 0 1 1 1
ov_27				c_region					
4.8293									
46									

NODE										
_50_le										
ngth_6										
99_co	564	G	A	missens	L6.peg	550G>A	Gly184Ser			/ / / / / / / 1 1 0 1 1
v_171.				e_varian	.2348					
09155				t						
3										

ATGAACCCATTCAAAGG	MNPFKGRHFQRDILWA
CCGGCATTTCAGCGT	VRWYCKYGISYRELQEM
GACATCATTCTGTGGG	LAERGVNVDHSTIYRWV
CCGTACGCTGGTACTG	QRYAPEMEKRLRWYWR
CAAATACGGCATCAGTT	NPSDLCPWHMDETYVK
ACCGTGAGCTGCAGGA	VNGRWAYLYRAVDSRG
GATGCTGGCTGAACGC	RTVDFYLSSRRNSKAAY
GGAGTGAATGTCGATC	RFLGKILNNVKKWQIPRF
ACTCCACGATTTACCGC	INTDKAPAYGRALALLKR
TGGGTTACGCGTTATG	EGRCPSDVEHRQIKYRN
CGCCTGAAATGGAAAA	NVIECDHGKLRIGATL
ACGGCTGCGCTGGTAC	GFKSMKTAYATIKGIEVM
TGGCGTAACCCTTCCG	RALRKGQASAFYYGDPL
ATCTTTGCCCGTGGCA	GEMRLVSRVFEM
CATGGATGAAACCTAC	

GTGAAGGTCAATGGCC
GCTGGGCGTATCTGTA
CCGGGCCGTCGACAGC
CGGGGCCGCACTGTCTG
ATTTTTATCTCTCCTCC
CGTCGTAACAGCAAAG
CTGCATACCGGTTTCTG
GGTAAAATCCTCAACAA
CGTGAAGAAGTGGCAG
ATCCCGCGATTCAACAA
CACGGATAAAGCGCCC
GCCTATGGTCGCGCGC
TTGCTCTGCTCAAACGC
GAAGGCCGGTGCCCGT
CTGACGTTGAACACCG
ACAGATTAAGTACCGG
AACAACGTGATTGAATG
CGATCATGGCAAACCTG
AAACGGATAATCGGCG
CCACGCTGGGATTTAA
ATCCATGAAGACGGCT
TACGCCACCATCAAAG
GTATTGAGGTGATGCG
TGCACTACGCAAAGGC
CAGGCCTCAGCATTTTA
TTATGGTGATCCCCTG
GGCGAAATGCGCCTGG
TAAGCAGAGTTTTTGAA
ATGTAA

NODE
_29_le
ngth_3 3312 T C intergeni
402_c c_region

/ / / / / / / 0 0 0 1 0

ov_28
1.5529
17

NODE
_32_le
ngth_6
5492_
cov_4
5.1184
27

G
A
A
A
T
T
G
C
A
T
G
T

upstrea
m_gene
_variant

L6.peg
.1770

-1595_
1585delACA
TGCAATTT

ATGGCAGAACAGCAAC
CACTGGTTTTGCTTGT
GACGATGAAGAAGATTT
GTGCCTTTTAATGCAAA
TGACACTTGCACGAAT
GGGGATTAACACACAT
CTTGCTTATCGGGTTGA
ACAGGCCAAACAACCTC
TTCACCTCAGTTTCATTA
CGATGCGTGTTTAACC
GACTTAAACCTACCCGA
TGGGAGCGGGATAGAT
TTAGTTAAACATGTCTC
TCAAATTATACTAATA
CTCCCATTGCCGTTTTA
ACCGCCTACGGCAATA
TGGATATTGCAATTGCA
GCATTAAGCAGGCG
CTTTTGATTTTGTAAGC
AAACCGGTCAACCAAG
TACATCTAGATCAATTA
TTACAAAAAGCCTTGAA
TCGGCAAAAAGTAGAG
CATGATGCTGCTGAGA
ATGCGTTAGAAAATAAA
TTATTAATCGGCCGTTT
TCTACCCATCCAGCAG
CTACGTATTGCAATTA
AAAAATTGCGCGCTCA
CAAGCACCTGTATTTGT

MAEQQLVLLVDDEEDL
CLLMQMTLARMGIKTHL
AYRVEQAKQLFTQFH
YD
ACLTDLNLPDGGSIDLVK
HVSQNYTNTPIAVLTAYG
NMDIAIAALKAGAFDFVS
KPVNQVHLDQLLQKALN
RQKVEHDAENALENKL
LIGRSLPIQQLRIAIIK
KIAR
SQAPVFVTGESGTGKEV
VANLVHRLSNRSEGP
FIA
INCGAIPTELMESE
LFGH
KKGSFTGATQDKQGLIL
SAHGGSLFLDEIAELPLS
MQVKLLRAVQEKKIRPV
GSDQEIDVDFRVISASH
QDLDLLVRQGKFRQDLF
FRIHVMDLILPPLRERGE
DVLLLANHFQKICMEWE
TPPKQLTEAAETYLLQQ
HFPGNVRELNRMIERAIT
LSEDTTIDVSHLHPAPLR
ANISNPFASAAQSIQTTV
AAPQVVKKLPSEGLERY
LENIEKDILLNALNMTHW
NRTLAAKKLGMFRSLR
YRLKKFGLDTETEQEV

/ / / / / / / 0 0 1 1 1

TACCGGTGAGTCTGGA
ACAGGTAAAGAAGTAG
TTGCTAACTTAGTTCAT
CGGCTGAGTAACCGTA
GTGAAGGCCCTTTTATT
GCGATTAAGTGCAGT
CTATTCCAACCGAACTC
ATGGAAAGTGAGCTTTT
TGGGCATAAAAAAGGG
AGCTTTACCGGAGCAA
CTCAAGATAACAAGG
CCTCATTATCGGCAC
ACGGTGGCAGTTTATTT
TTAGATGAAATTGCCGA
ATTACCTTTAAGTATGC
AGGTCAAAGTCTCCG
TGCAGTACAGGAAAAA
AAGATCCGACCCGTCG
GTTTCAGATCAGGAAATC
GATGTTGATTTCCGTGT
GATTAGTGCAAGTCATC
AAGATTTAGATTTATTG
GTTAGACAAGGTAATTT
CCGTCAAGATTTATTCT
TCCGTATTCATGTTATG
GACCTCATATTGCCACC
TCTACGTGAACGTGGC
GAAGATGTCCTTTTACT
AGCCAATCACTTTATT
AGAAAATTTGTATGGAG
TGGGAAACGCCACCTA
AACAATTAACAGAAGCG
GCCGAACTTATCTGCT

ACAGCAACACTTTCCG
 GGTAATGTTGCGGAATT
 AAGAAATATGATTGAGC
 GCGCAATTACCTTAAGT
 GAAGATACCACTATAGA
 TGTATCTCATTTACATC
 CTGCTCCACTAAGAGC
 AAATATTTCTAATCCTTT
 TGCTTCAGCTGCTCAA
 GCATACAAACCACTGT
 GGCAGCTCCTCAAGTC
 GTAAAAAATTACCAAG
 TGAAGGCTTAGAACGTT
 ATTTAGAAAATATTGAA
 AAAGATATTTTACTCAA
 TGCACTCAATATGACTC
 ATTGGAATCGTACCTTA
 GCAGCTAAAAAATTAGG
 AATGACTTTCCGCTCTT
 TACGCTATCGTCTGAAA
 AAATTTGGCTTAGATAC
 GGAGACAGAACAGGAA
 GTATAA

ATGAGCTACCAGAACT MSYQKLSRDQIAQLVAQ
 CAGCCGTGACCAGATT DIPDGAYVNLGIGLPTKIA
 GCGCAACTTGTGGCAC SYLPADKDVFLHSENGL
 AAGATATTCCCGATGGA LAFGPPPAAGEEDPELIN
 GCCTATGTGAACTTGG AGKEFVTMLEGGSFFHH
 GCATTGGCTTACCGAC GDSFAMMRGGHLDIAVL
 CAAGATTGCCAGCTACT GAFQVAANGDLANWHT
 TACCTGCTGACAAAGAT GAPDAIPAVGGAMD LAV
 GTCTTCCTACATTCTGA GAKKVFITTDHITKQGE P
 AAATGGCCTGTTGGCTT KIVAEITYPVTKHCVDR

/ / / / / / / 0 1 1 1 1

NODE
 _114_I
 ength_
 20019 1895 T G missens
 7_cov 99 e_varian L6.peg 271A>C lle91Leu
 _34.99 .365
 1173

TTGGCCCACCACCAGC IYDLCVIDVTKDGLKVIE
 TGCAGGTGAAGAAGAC KVEGLSFDELQALTGAT
 CCTGAACTGATTAATGC LIDATQG
 GGGTAAAGAGTTTGTG
 ACCATGCTTGAAGGCG
 GCAGCTTTTTCCACCAT
 GGCGATTCATTTGCCAT
 GATGCGTGGTGGTCAC
 CTTGATATTGCGGTGCT
 TGGGGCATTTCAGGTG
 GCAGCCAATGGTGA
 TGGCGAACTGGCACAC
 GGGTGCACCGGATGCG
 ATTCCTGCGGTAGGTG
 GTGCAATGGATTTGGC
 AGTTGGTGCCAAGAAA
 GTATTTATCACTACAGA
 CCACATCACCAAGCAA
 GGTGAGCCGAAGATTG
 TCGCTGAGCTCACTTAT
 CCGGTAACGGGTAAAC
 ACTGTGTTGACCGTATT
 TACACTGACCTGTGTGT
 GATTGATGTGACCAAA
 GACGGTTTAAAGGTGA
 TTGAGAAAGTGGAAAG
 TCTTAGCTTTGATGAGT
 TACAGGCTTTAACCGGT
 GCAACTTTGATTGATGC
 GACTCAAGGGTAA
 ATGAGCTACCAGAACT MSYQKLSRDQIAQLVAQ
 CAGCCGTGACCAGATT DIPDGAYVNLGIGLPTKIA / / / / / / / 0 1 1 1 1
 GCGCAACTTGTGGCAC SYLPADKDVFLHSENGL

NODE
 _114_I
 ength_

1896
 02

A G

missens
 e_varian
 t

L6.peg
 .365

268T>C Phe90Leu

/ / / / / / / 0 1 1 1 1

20019
7_cov
_34.99
1173

AAGATATCCCGATGGA	LAFGPPPAAGEEDPELIN
GCCTATGTGAACTTGG	AGKEFVTMLEGGSFFHH
GCATTGGCTTACCGAC	GDSFAMMRGGHLDIAVL
CAAGATTGCCAGCTACT	GAFQVAANGDLANWHT
TACCTGCTGACAAAGAT	GAPDAIPAVGGAMDLAV
GTCTTCCTACATTCTGA	GAKKVFITTDHITKQGEF
AAATGGCCTGTTGGCTT	KIVAELTYPVTGKHCVDR
TTGGCCCACCACCAGC	IYDLCVIDVTKDGLKVIE
TGCAGGTGAAGAAGAC	KVEGLSFDELQALTGAT
CCTGAACTGATTAATGC	LIDATQG
GGGTAAAGAGTTTGTG	
ACCATGCTTGAAGGCG	
GCAGCTTTTTCCACCAT	
GGCGATTCATTTGCCAT	
GATGCGTGGTGGTCAC	
CTTGATATTGCGGTGCT	
TGGGGCATTTCAGGTG	
GCAGCCAATGGTGACTION	
TGGCGAACTGGCACAC	
GGGTGCACCGGATGCG	
ATTCCTGCGGTAGGTG	
GTGCAATGGATTTGGC	
AGTTGGTGCCAAGAAA	
GTATTTATCACTACAGA	
CCACATCACCAAGCAA	
GGTGAGCCGAAGATTG	
TCGCTGAGCTCACTTAT	
CCGGTAACGGGTAAAC	
ACTGTGTTGACCGTATT	
TACACTGACCTGTGTGT	
GATTGATGTGACCAAAA	
GACGGTTTTAAAGGTGA	
TTGAGAAAGTGGAAGG	

NODE
_114_I
ength_
20019
7_cov
_34.99
1173

1896
05

G T

synony
mous_v
ariant

L6.peg
.365

265C>A

Arg89Arg

TCTTAGCTTTGATGAGT
TACAGGCTTTAACCGGT
GCAACTTTGATTGATGC
GACTCAAGGGTAA
ATGAGCTACCAGAACT
CAGCCGTGACCAGATT
GCGCAACTTGTGGCAC
AAGATATTCCCGATGGA
GCCTATGTGAACTTGG
GCATTGGCTTACCGAC
CAAGATTGCCAGCTACT
TACCTGCTGACAAAGAT
GTCTTCCTACATTCTGA
AAATGGCCTGTTGGCTT
TTGGCCCACCACCAGC
TGCAGGTGAAGAAGAC
CCTGAACTGATTAATGC
GGGTAAAGAGTTTGTG
ACCATGCTTGAAGGCG
GCAGCTTTTTCCACCAT
GGCGATTCATTTGCCAT
GATGCGTGGTGGTCAC
CTTGATATTGCGGTGCT
TGGGGCATTTCAGGTG
GCAGCCAATGGTGACT
TGGCGAACTGGCACAC
GGGTGCACCGGATGCG
ATTCCTGCGGTAGGTG
GTGCAATGGATTTGGC
AGTTGGTGCCAAGAAA
GTATTTATCACTACAGA
CCACATCACCAAGCAA
GGTGAGCCGAAGATTG

MSYQKLSRDQIAQLVAQ
DIPDGAYVNLGIGLPTKIA
SYLPADKDVFLHSENGL
LAFGPPPAAGEEDPELIN
AGKEFVTMLEGGSFFHH
GDSFAMMRGGHLDIAVL
GAFQVAANGDLANWHT
GAPDAIPAVGGAMD LAV
GAKKVFITTDHITKQGEP
KIVAELTYPVTGKHCVDR
IYDLCVIDVTKDGLKVIE
KVEGLSFDELQALTGAT
LIDATQG

/ / / / / / / 0 1 1 1 1

NODE
_31_le
ngth_3
210_c
ov_10
5.6009
37

912 C T

synony
mous_v
ariant

L6.peg
.1746

46C>T

Leu16Leu

TCGCTGAGCTCACTTAT
CCGGTAACGGGTAAAC
ACTGTGTTGACCGTATT
TACTGACCTGTGTGT
GATTGATGTGACCAA
GACGGTTTAAAGGTGA
TTGAGAAAGTGGAAGG
TCTTAGCTTTGATGAGT
TACAGGCTTTAACCGGT
GCAACTTTGATTGATGC
GACTCAAGGGTAA
ATGATTGAATCTAATGC
TGTC AATATCGATCGTA
GTCGTATCATTCTAAAA
CCCAATGTGGTTGTCG
GTTATGAGGGGCAACC
CTATAAAATTGTGAATG
TTTTAAATGCCAATGAC
ATTGTTATATCAAGTCT
TGATTCGGTACGTAGTC
TACAGGTCAATGCTAGA
AGTTTGACCATTTTTAA
TGGGGGAAATACGCTA
ACTGAGCGTTTGAATAA
GGGTGATCAGGATATT
AGTAATGAAGCTTGCC
AGATTGCTTTACAACGT
TATGAGATTATTAAGCC
TTTGATTGAGTATTCAA
CGACTGAACTGGTCGA
AAATAGGGCGAATGAA
TATGATGTGAATCGCTC
AACATTGTGGAAATGG

MIESNAVNIDRSRIILKPN
VVVGYEGQPYKIVNVLN
ANDIVISSLDSVRS LQVN
ARSLTIFNGGNTLTERLN
KGDQDISNEAWQIALQR
YEIIKPLIEYSTTELVENR
ANEYDVNRSTLWKWLK
DYRENNSLMALIPKCRG
WTTEKSRLSPQVSNIHKQ
AINDEYLNKKPSISK TIE
IVKAEC SRLQLEAPHENS
IRRRREALNDYQVTKARL
GSKAAIDKFKAAGSFP
NADYPLAYVQIDHTPLDI
EIVDDEFREAIGKPHLTL
AIDVFSRMIVGYYSLEA
PSTTSVAMCIASCILSKK
RKLIELDIDA EWQVEGIM
DSVHTDNGPDFRTNHIS
KACLKYGIHWEYRPIGG
ARFGGHIERMLGIVNLE
MHVLDGTTFSNVQQRG

/ / / / / / / 0 1 0 0 1

CTTAAAGATTATCGGGA	TYDSAKQACMTLKELEY
AAATAATTCTCTAATGG	YIVYWITKVYHQKHSAL
CATTGATTCCCAAAAAA	GTSPIVKWEEGVWGTKT
CGAGGTTGGACAACTG	TAGTGLKERVSEDTLFI
AAAAATCAAGGTTGTCA	DFLPEFEATIQRTGVQK
CCTCAAGTGAGCAATAT	DNLFYFADCLRQWVNSI
CATTAAACAAGCTATCA	DPEDNNRKRKRKFLFKR
ATGATGAATATTTGAAT	DPRDISQIWFYEPYSNTY
GCGAAGAAACCCAGTA	FKVPTAKREIPPISLFEYK
TTTCCAAAACGATTGAA	QVQNYLKSERQDVQNQ
ATTGTAAAAGCAGAGTG	DEIYKAILHLRDQLNQAR
TTCTCGGCTTCAGCTTG	SLTRKQRRSNQRKKENA
AAGCACCGCATGAAAA	KAITQLSEQNQSKKAVV
CTCTATTCGTAGAAGAA	SESLQTSDDLWSTPLTA
TCGAGGCTTTAAATGAT	FDDLRL
TACCAAGTAACGAAAG	
CACGTCTTGGATCTAAA	
GCTGCAATAGATAAATT	
TAAGGCTGCTGCTGGT	
TCATTCCCAAATGCTGA	
TTACCCTTTGGCATATG	
TGCAGATTGATCATACA	
CCTCTAGATATTGAAAT	
CGTGGATGACGAGTTT	
AGAGAAGCGATTGGAA	
AACCTCATTTAACCTTG	
GCAATCGATGTTTTTCAG	
TCGGATGATCGTAGGG	
TATTACCTATCACTAGA	
AGCGCCAAGTACGACT	
TCAGTGGCGATGTGTA	
TCGCATCTTGATTTTG	
TCAAAAAACGTAAGTT	

GATCGAACTCGATATTG
ATGCTGAATGGCAGGT
CGAAGGGATTATGGAC
TCGGTGCATACAGATAA
TGGCCCCGATTTTAGAA
CCAATCACATCAGTAAA
GCTTGTCTAAAGTATGG
GATTCATTGGGAATACC
GACCAATTGGTGGGGC
AAGATTTGGCGGCATA
TCGAGCGTATGCTTGG
TATCGTCAATCTTGAAA
TGCATGTCTTGGATGGT
ACAACTTTTTCAAATGT
ACAGCAGCGTGGGACT
TATGATTCCGCTAAACA
AGCCTGTATGACATTAA
AAGAGCTTGAATACTAT
ATCGTGTATTGGATTAC
TAAAGTTTATCACCAGA
AGAAACATTCCGCTTTA
GGAACATCCCCAATCG
TTAAATGGGAAGAAGG
TGTTTGGGGGACGAAG
ACGACAGCAGGTACAG
GTTTAAAAGAACGTGTA
TCAGATGAAGATACCTT
GTTTATTGATTTTCTAC
CTGAATTTGAAGCGACT
ATTCAGCGTACAGGGG
TGCAAAAAGATAACTTA
TTCTATTTTGCCGACTG
TTTAAGACAATGGGTTA

ATTCTATTGATCCAGAG
GACAATAATCGAAAGC
GCAAAAGAAAATTCTTG
TTTAAACGTGATCCGAG
GGATATTTACAGATTT
GGTTTTATGAGCCATAT
TCTAATACCTATTTTAAA
GTACCTACCGCAAAGC
GAGAGATACCACCGAT
TAGCTTATTTGAATATA
AGCAGGTTCAGAACTAT
CTCAAAAGTGAACGTCA
GGATGTTCAGAATCAG
GATGAGATTTACAAGG
CGATTTTGCATCTACGT
GATCAACTTAATCAGGC
TAGAAGCTTGACTCGAA
AGCAACGTCGCTCAAA
CCAAAGAAAAAAGGAA
AATGCAAAAGCCATTAC
TCAGCTTTCCGAACAGA
ATCAGTCTAAAAAGGCT
GTAGTTTCTGAAAGTCT
TCAGACAAGTGATGATT
TGTGGAGTACTCCTTTA
ACTGCATTTGACGATTT
AAGGTGA

Table S3 Summary of genes that gained or lost. The bloodstream specific genes were highlighted in red.

Gene	Non-unique Gene name	Annotation	Isolate ID of gene lost	COG functional categories	description	Nucleic acid sequence
<i>group_1297</i>		hypothetical protein	A1, A3 L6 C4, C6	/	transcriptional regulator	ATGGTCCGTATGGGAGAAACAGGGTTTATGGAATGGACGGAGCATAAAAG TTTGTCAATCACGAATAAATATATTAATTTGATAACACTATTCGTATTTTA TAACAATTGATGTGTTTGCCTAGAAAATGCATTAATAGGAGTGTGTGTGT GGATCTATTTTCATGCCATGA
<i>group_1301</i>		hypothetical protein	A1, A3 L6, L7, L8, L10 C4	/	membrane protein	ATGTCAGTAACAACCTCAAGAGGGTTGCCTGTAACGCCACCCGTTACGCC GCCAATAATATTGGTTACAGTATCAATTGGGCTTGTAACCTACAGTACC TGTTACGGTACCAACTACATTAGAACTAAGTCTGTGGTTGTATGTGAAAC TGTATCGGTTACAGTTGAAATCACATTAGAAAGTAAGCCTGTTGTTGAAAC ATTAGCAATGTTGATATCAGAAACATTTACATTAACAGAGGAACCTGTGTTT ACTACAGGGGTTATCGATACAATATTACCCAGGTCCTGA
<i>group_39</i>		hypothetical protein	A2, A3, A14 L9 C2, C6	UW	Hep Hag repeat protein	ATGTTCTTACTGAAAAATCTCGTCTCTTCAATATCAAAAGTTACTCAGGACC TGGGTAATATTGTATCGATAACCCCTGTAGTAAACACAGGTTCTCTGTTA ATGTAAATGTTTCTGATATCAACATTGCTAATGTTTCAACAACAGGCTTACT TTCTAATGTGATTTCAACTGTAACCGATACAGTTTCACATACAACCACAGA CTTAGTTTCTAATGTAGTTGGTACCGTAACAGGTTACTGTAGGTTCTACAAG CCCAATTGATACTGTAACCAATATTATTGGCGGCGTAACGGGTGGCGTTA CAGGCAACCCTCTGAAGTTGTTACTGACATTATCGGCGGAGTAACGGGT GGTGTGTCGGCGGAACCTTCTCCAATCTCTCCAGTAATCGATGTTGTTCAA GGCGGCATTGATATTCTTCAAGGCGTTGAAAGCTTAAAACTGAAATTATT AATACAGGCATTGAACTGTAGCTGATGCAGTTATTGGTGTATTACCTCAA GCTGAGCATCCAGTTTCTGAAATTGCAGACCTTGGCACATTAACTTTTGAA ACTTCTCGTGACACAGTCAATGGTGTCTTGAAGTTGTTTCTGATTTAGTC GGTGGTGTATTCAAGGCGCTACTGAATCAGCGACTGGTATTCTCGATAC GCTTATTACAAATGGCACAACTGCTACAGGTATTGTTACTGGTGTGGTTGG AAGTGTTACTGATGTTATTGGTGGCGTAACTGGTGGTGTGATGGCAACC CACTTGAAGTGATTACAGATATTATCGGTGGTGTAACTGGCGGTGTTGAT GGCAACCCACTTGAAGTGATTACAGATATTATTGGCGGTGTAACCTGGTGG

TGTTGGGGTGATAACCCTCTTGGCGTAGTAACTGATATTATCGGCGGCG
TAACCGGTGGCATTATTGGTGGCGGTACATCTCCTATTTCTCCAGTTATTG
ATGTTGTTCAAGGTGGCATCGATATTCTTCAAGGCGTTGAAAGCCTAAAAA
CTGAAATTATTAACACTGGTATTGATACAGTAGCCGACACTATTATTGGTG
TATTACCTCAAGCTGAGCATCCAGTTTCTGAAATTGCAGACCTTGGCACAT
TAACTTTGAAACGTCTCGTGACACGGTCAATGGTGCTCTTGAAGTTGTTT
CTGATTTAGTCGGTGGCGATATTCAAGGTGCTACTGAAATGCAACTGGT
ATTTTAGGAACACTTATTCAGAATGGTTCAACAGCGACTGGAATTGTTACT
GATATTCTTGGTGGCGTAACAGGTACTATCGGCGGCGTAACTGGTGGTAT
CGGTGGTGACAATCCGCTCGGTATAGTAACTGACATTATCGGCGGTGTTA
CTGGCGGTATCGGCGGTGATAATCCGCTTGGTGTAGTACTGACATTATC
GGTGGTGTACTGGCGGTATCGGCGGTGATAATCCGCTTGGTGTAGTGA
CTGACATTATCGGTGGTGTACTGGCGGTATCGGTGGTGACAATCCGCTT
GGTGTAGTAACTGATATTATCGGTGGTGTACTGGCGGTATTGGTGGTGA
CAATCCGCTTGGTGTAGTAACTGATATTATCGGTGGTGTACTGGCGGTAT
TGGTGGTGACAATCCGCTTGGTGTAGTAACTGACATTATCGGCGGCGTAA
CCGGTGGTATTATTGGTGGCGGTACATCTCCTATTTCTCCAGTAATCGATG
TTGTTCAAGGTGGCATCGATATTCTTCAAGGCGTTGAAAGCCTAAAACTG
AAATTATTAATACTGGTATTGATACAGTTGCGGATACAATTATCGGTATCCT
TCCACAAGCAGAACATCCAGTTTCTGAAATTGCAGACCTTGGCACATTAAC
TTTTGAAACGTCTCGTGACACGGTCAATGGTGCTCTTGAAGTTGTGTCTGA
TTTAGTCGGTGGTGATATTCAAGGCGTACTGAATCGGCTACCGGTATTC
TCGATACGCTTATTACAAATGGTACAACACTGCAACTGGTTTGTAACTGAAA
TCATTGGTGGAGTAACAGGAACTATTGGTGGCGTAACTGGCGGAGATAGC
CCACTAGGTTTAGTAACTGATTTACTTGGTGGTTTAACTGGTAGCGTAGGC
GGCGACAGTCCGCTTGGTGTAGTCACTGACCTTCTCGGTGGTTTAACTGG
TAGCGTAGGTGGCGACAGTCCGCTTGGCGTAGTCACTGATCTGCTCGGC
GGTTTAACTGGTGGTGTACTGGCGGTACAGACAACCCTATCGGTATCGT
AACGGATATCGTTGGAAGCTTAACGGGTGGTGTACTGGTGAAGGCAGTT
TAGATGTTATTTCTAATTTGCTAGGCGGATTGACTGGTGGTAGCTTACTCG
GTGGTGTAACTTCTACAGTATCTAGCGTAACAAATACAACGCATACGATCG
TACCGACATCATTACTTACTGATAACTTCTAGAAAACACTCATTCAATACTGT
TTAA

<i>group_4700</i>	hypothetical protein	A1 L7, L8, L9 C4	/	type I secretion C-terminal target domain-containing protein	ATGTCTGCTGAAACTGTGCCTGTACCTGGCAATGAAGTGTTCCTGCAGG GTCAGTTGCTGTTGCAGTCACTGTATCGCCATCCACCAGGTTACCTGGGT TTGGTACTGACCAGCTGCCATCTGTGCCTGCTACCACTGTTGCAGTTGTG CCATCAGGATAA
<i>group_839</i>	hypothetical protein	A14 L9, L10 C4 C4	/	putative membrane protein	ATGTTTATTTGCGTTGATGGTGAAGTCAATGGGCAAGTGATAGAAAAAAGG GGCGTTAAGAACAAGATGTATATAAATATTATAAACTCAGTAA ATGGCACTTAATGTTTTTAAATTCAAAAAATATGCAAGGATGTTACACTAT TAAATTTAATTTGCTTTTATCTATCTGGCTAGGTTATTTCTGAATATAGGT TTTTTAAAAAATCCATCAACTTACACCTTATAATGGTATTAAGTCAGTTCT TTTCTTAGGGGCGACATTAGTTATTTAATAGCGGCATATAATTTAATTTTT CAATTAATAAATTGGAAATGGACTGCCAAAATCTTTGCAATTTTATTGATAT TTATTGGTGGCTTTAGCTCTTATTTTGTAAACACATTGGGTGTCATTATTTTC ACCCGACCAAATTCAAAATATGGTGCAGACCGATGTTTCGGAATTTACCGA TCTAATCTCTTACGCTTTGTTTTATGGACAGTTTTTTTTGTTATTTGCCCA TTTTTTAATTACTCAAGTTAAATTTAAACAAGAAAAAGCATCACGGTTGTT ATTGAAGAAAGTATTCTCACTGGTAGCTTCATTTGCAGTGGTCGGTGTTTT ACTTTTTACTTACTATGTCGATTTTCGCTGCAATATTTTCGTGAGCATCGTGAT TTAAAAGGGATGATTTACCGCAAATAGTATTTTCATCGCTTATGTCCTTACT ATCATAAGAAGGCTCCGAAGAAAAATCTGCCTCTTGTGATATATGGACAAG ATGCTCATCAAGTTCAGCAAGTACAAAAGAACCCTCCCTAAGTTAATGATAC TTGTTGTAGGTGAAACGGCACGTGCCGAAAGTTTCTCTCTAAATGGGTAT GCAAAAAATACGAATCCGGAGCTTTCTAAACAAGATATTTCAACTTTTCG CAAGTGAGCTCATGCGGTACGGCGACAGCTGTTTCTGTGCCATGTATGTT CTCGGGTATGCCACGTGTAGATTATGATGAGCAATTAGCTAGTCACCGCG AAGGTTTATTAGATATTGCAAAACGTGCGGGTTACCAAGTGACTTGGATTG ATAATAACTCGGGTTGTAAAGGTGCATGTGATCGCGTTGAACAATACCAG ATTCCAGAAAATTTAAAGAAAAAATGGTGTAAAGATGGCGAATGTTATGAT GACATTCTCATTGACAGCTTAAAGCAGTATTTGGCTACTATTGCCAAAGAT GATGATCGTCCGCGTTTGATTGTTCTGCATCAAGTGGGTAGTCATGGACC TGCATATTACAAGCGTGCACCTGAGGCATATCAACCTTTTAAACCGACTTG TGATACTAATGCGATACAGGGCTGTTTCGCAAACTGAATTGCTAAATAGTTA TGATAATACAATTGTATATACAGACCATGTATTAAGCCAAATGATTAATACT
<i>eptA_2</i>	Phosphotransferase EptA	L9	S	Domain of unknown function (DUF1705)	

		A2, A3, A14			CTAAAAGAAATATCAAAGTATCAGACAGGTTTATGGTATTTATCTGATCATG GCGAATCAACCGGAGAACATGGTTTATATTTACATGGTTCACCTTATGCAA TCGCACCGAGCCAACAAACACATGTACCAATGATTATGTGGTTCTCTGAAA GTTGGAAACAACATAATCTTGCTCAAGTGAATTGTTTAAGCCAACAACTA AACAAAAGTTAAGTCAGGATAATTTATTCCCAAGTTTGTTAAGTTTGCTGGA TGTA AAAACTCAGGTAGTAAATAACAACTTGATATGTTGAGCCAATGTAA ATAA ATGAATACTCTTGCTGAACGACTTAGGTATGCCATGGAAGTTTTGCCACCT AAAAAGATTAAGGTGTTGAGCTTGCTCGTGCAAGTAGGAGTTAAACCTCCT TCTGTGAGTGATTGGCTGTCTGGAAAATCCAAAACAATGGAAGGTGAAAA TTTATTACGTGCCTCAAATTTTTGAATGTAAATCCTTCATGGCTTGCATCT GGCACGGGAGAGATTCAATCAAGCACGAGAGATAAATTTAAACAACCTGGA TATCGAAGAGTTCAAAAAGAAATACAACATTAGTGATAGTGATGAAGCTCT TTTATTTTCAACAATTATCGAAAAACCGTTTATCCCATCATCTAAGCGTTGG GTTCTGTTAAGGCTTACTCCAAGATGGGCATGGATGGCTATTTACAGAT ATGGGTTATGAAGGCAATGCTGGAGATGGGTATGTTCCAACCTCACTCAGC AGGACCAAGAGCCTATGGCATTAAAGGCACTGGCGACTCAATGTTTCCAG CAATTCGTAATGGCTGGTATGTTGTATGCGACCCTGATGCAGAGCTTGTG CCGAATGAGTTTGTTCAGGTGTGCTTGAAGGATGGAAGATGCACAATTAA AGAATTTGTGCGCATCAATGGTGGGGTTTTAAGTTTGCTTTCTGTGAATGG TGGTGAGCGATTTTTCTTTGAAATGGACGAGGTTGAAAGTATTACCGCTAT TACAGATATCGTGCCGCCAAGTCAGCACAGACAAGAACATCCTTATTCGC ATTAA ATGCTGGCTGAACGCGGAGTGAATGTCGATCACTCCACGATTTACCGCTG GGTTCAGCGTTATGCGCCTGAAATGGAAAAACGGCTGCGCTGGTACTGG CGTAACCCTTCCGATCTTTGCCCGTGGCACATGGATGAAACCTACGTGAA GGTCAATGGCCGCTGGGCGTATCTGTACCGGGCCGTCGACAGCCGGGG CCGCACTGTGATTTTTATCTCTCCTCCCGTCGTAACAGCAAAGCTGCATA CCGGTTTCTGGTAAAATCCTCAACAACGTGAAGAAGTGGCAGATCCCGC GATTCATCAACACGGATAAAGCGCCCGCTATGGTCGCGCGCTTGCTCTG CTCAAACGCGAAGGCCGGTGCCCGTCTGACGTTGAACACCGACAGATTA AGTACCGGAACAACGTGATTGAATGCGATCATGGCAAACCTGAAACGGATA ATCAACGCCACGCTGGGATTTAAATCCATGAAGACGGCTTACGCCACCAT
<i>group_107</i>	hypothetical protein	L6	K	Peptidase S24-like	
		A3			
<i>group_11</i>	IS6 family transposase IS15	L6, L9, L10	L	Transposase and inactivated derivatives	

					CAAAGGTATTGAGGTGATGCGTGCACTACGCAAAGGCCAGGCCTCAGCA TTTTATTATGGTGATCCCCTGGGCGAAATGCGCCTGGTAAGCAGAGTTTTT GAAATGTAA
		A1			ATGAGTACTCTTCAAGAGCGAATGTCTTTAGCTATAAAGCACTATGAATCT GTGACAGGTAAAAGATTCAAAAATACTGAGCTGGCTAGATTTGCAGGCGT GAGCAGGGCTAATGTTGGGTTGTGGGTAAATGGCCCAACCCAAGAGTTG GAGGGGTCAAATTTAGTTAAAGCGGCTGAGTTTTTTAGGAGTTTCTAAAGAT TGGCTAGCTGGACAAGGTAATAAAATGATTGCCACACAGCTGGATGGTGG TGGTGACAATTAATGTTCTTGATATTGAAGCCTTTAAGCAGAAGTACAA CATTCCAGATAGTGAAGATGCTGTTAAGTTTGTTCAAACATCAGATAAACC ATTCCCTATTCAAAAAGATACGTTCCCGTCAAAGCTTATTCAAAGATGGG TATGGATGGGTATTTACAGATATGGGATACGATGGAATGCTGGAGATG GGTATGTTCCAACACTCACTCAGCAGGACCAAGAGCCTATGGTATTAAGGC ACTGGCGACTCAATGTTTCCAGCTATCCGTAATGGATGGTATGTGGTTTTGT GATCCAGATGCGGAACTCGTGCCGAATGAGTTTGTTCAAGGATGCTTGAA GGATGGAAGATGCACAATTAAGAATTTGTTGGCATAAATGGCGGGGTTT TAAGCTTGCTTTCTGTGAATGGTGGTGAGCGATTTTTCTTTGAAATGGATG AGGTAGAAAGCATTACAGCTATTACTGACATCGTACCACCAAGTCAGCAC AGACAAGAACATCCTTATTCGCATTA
<i>group_110</i>	hypothetical protein	L10	K	Peptidase S24-like	
		A2, A3, A14			ATGAATATGCCAGTACAACAACACATTTTACAAGCGGTGCGATTGGTCTAGA TTTAGTTTTGAAGAGTGGTGTGCGCCAGCTTGGAGCTTGGCTAAACGGCGA TACCGAAACAATGGTCAAATTTGTTAAGACGATGCCAACAAAACGCATCAC TCAAAAACAAGAGAAAAATTAATAGCTATGTATATGAGCGATGAAAATCT AAAAGATCGTTTATGCATTCGCCGTAAGGGTACTTGCTGTGAGTTAAATGA CAATGAGGCACGTGCAATCCATAGATTGATTATTGATATTAATTAATTGAA GACCATATTTTACAAGAATGGATCTCAGCAATTTGGTCACATCATGTTATG GGCAATTCATTACGTGATATTGCTCAAAGTAATGACACTTCAGTTAATCAA ATCAGACAGGATTTAAAATGTGGTATGGCTTATATCAAAGTCGAAATCCG CATTTTCAGATTTGAACTTTTTGAAAAACCGCTTGA
<i>group_113</i>	hypothetical protein	L7, L9	/	hypothetical protein	
<i>group_1298</i>	<i>deaD</i> ATP-dependent RNA	A1, A2 C2, C4	L	Belongs to the DEAD box helicase family	ATGAGCAAACCTTTTGCCGAATTTCTTTGCACGAAACCTTACAACAAGCT CTTGAAGGTTTAGGTTTTACTAACCCAACTCCTGTGCAAGAACAATCTATT CCTGCTGCTTTAGAAGGTAAAGACCTTCTTGTCTTAGCCAGACTGGTTCT GGTAAAACCTGCTGCATTCTTATTACCGACATTACATAATTTGGCAGGTCAA

helicas
e
DeaD

GATACTTTCGTTCCGTTCAAAGAACGTATGAAAGCAGTGA CT CAGCCTAG
CATTTTAGTTTTATGCCCAACACGTGAATTGGCTCAACAAGTTAGCCAAGA
CGCAATTGCTTTTGTACGTCACATGAAAGGTGTTTCGTATTGCTGCAATTAT
GGTGGTATGCCTTTTGGTAAACAAATTCAACAGCTTAAAGGTGCTCAAGT
TG TAGTTGCGACTCCAGGTCGTTTATTAGACTTAGTTAACCGTCGTCAATT
AAAGCTAGACAAAGTTGATGCTTTAATCGTGGACGAAGCTGACCGTATGTT
AGACCTTGGTTTCTCTGAAGACTTAGAAGCGATCAGTGA CT TGGCTGCCA
ATCGTGGTCAAACGTTAATGTTCTCTGCAACTTTTGTGACCGTATTATTC
GTCTTGCTGAGCGCATGATGAATGAACCTGAGCGTATTGCAATTGAAACT
GGTCACAGCACCAATACTGACATCACCCAACTTTACACTGGACAGATGG
TTTCGAACACAAGAAGAACTTCTGACTCACTGGTTGGCTGATGAAACGCT
TGACCAAGCTGTTGTATTGCGGAGCACACAAGAAGATACAGATATGTTGG
CTGAAGA ACTTGCAGAAGCAGGTC ACTCAGTTGTTGCACTTCACGGTGCA
ATGCCGCAAACAGTTCGTAACCGTCGTTTACGTAGTATCCGTGAAGGCCG
TGCGAAAATCTTAGTTGCTACTGACGTTGCTGCACGTGGTCTTGACGTT
CAACAATTTACACGTTATTA ACTTCGGCTTACCAATGAAGCATGAAGATT
ATGTACACCGTATTGGTCGTACAGGTCGTGCTGGTCGTACAGGTCAAGCG
ATTACTTTAGCGACTTATCGTGAACGCGGTA AAATCCGTGCGCTTGAAGA
CTACTTAGAAGCTCGTTTAAATGTTTCTGAAATTGAAGGTTTAGAGCCATC
TCCACCTCCAGCTCGTCCACGTCGTGAAGGTGGTCGTGGTGGTAACGGT
GGCCGTGATGGTCGTGTTTTGAGGGTGAAGGTA ACTTCAAACGTCGTGA
AGGTGGTCGTGATGATCGTCCACGTCGTAGCTTCGATGATAAGCCTCGTG
GTGAGCGTCCTGCTTTTCGGTGGCGAAGATCGTCCACGTCGCGAGTTCAA
CTCAGATCGTCCACGTCGTGAAGGTGGTTTTGAAGACCGTCCAAAACGTA
GCTTCGGTGGTGAAGATCGTCCGCGTCGCGAGTTCAACTCAGATCGTCCA
CGCCGTGAAGGTGGTTTTAACGATAAGCCTCGTTTTGATTCAAATGATGAC
AATCGTGGTAACCGCGTAGACTATAAACACGTCGTGAAAACGGTTTTCGG
TGACCGTCCACAACGTAGCTTTGGTGGTGAAGATCGTCCACGTCGTGAAG
GCGTTTTTGGTATCGTCCAAAACGTAGCTTTGGTGGTGAAGAACGTCCA
CGTCGCGCAGTTGCGGAAGAACACTTTAATCAAGAATCTCGTGGTGAGCG
CCGTCGTAAATTCGACCGTTAA
ATGAGCAAACCTTTTGGCGAATTTCTTTGCACGAAACCTTACAACAAGCT
CTTGAAGGTTTAGGTTTTACTAACCCAACTCCTGTGCAAGAACAATCTATT

group_
1316 *deaD* ATP-
depen A1, A3,
A14 L Belongs to the DEAD box helicase family

dent
RNA
helicas
e
DeaD

C6

group_
1327

hfq

RNA-
bindin

A1, A2,
A14

/

hypothetical protein

CCTGCTGCTTTAGAAGGTAAAGACCTTCTTGTTTCTAGCCAGACTGGTTCT
GGTAAAAGTCTGCTGCATTCTTATTACCGACATTACATAATTTGGCAGGTCAA
GATACTTTCGTTCCGTTCAAAGAACGTATGAAAGCAGTGACTCAGCCTAG
CATTTTAGTTTTATGCCAACACGTGAATTGGCTCAACAAGTTAGCCAAGA
CGCAATTGCTTTTGTACGTCACATGAAAGGTGTTTCGTATTGCTGCAATTAT
GGGTGGTATGCCTTTTGGTAAACAAATTCAACAGCTTAAAGGTGCTCAAGT
TG TAGTTGCGACTCCAGGTCGTTTATTAGACTTAGTTAACCGTCGTC AATT
AAAGCTAGACAAAGTTGATGCTTTAATCGTGGACGAAGCTGACCGTATGTT
AGACCTTGGTTTCTCTGAAGACTTAGAAGCGATCAGTGACTTGGCTGCGA
ATCGTGGTCAAACGTTAATGTTCTCTGCAACTTTTGTGACCGTATTATTC
GTCTTGCTGAGCGCATGATGAATGAACCTGAGCGTATTGCAATTGAAACT
GGTCACAGCACCAATACTGACATCACCCAACTTTACTGGACAGATGG
TTTCGAACACAAGAAGAACTTCTGACTCACTGGTTGGCTGATGAAACGCT
TGACCAAGCTGTTGTATTTCGCGAGCACACAAGAAGATACAGATATGTTGG
CTGAAGA ACTTGCAGAAGCAGGTC ACTCAGTTGTTGCACTTCACGGTGCA
ATGCCGCAAACAGTTCGTAACCGTCGTTTACGTAGTATCCGTGAAGGCCG
TGCGAAAATCTTAGTTGCTACTGACGTTGCTGCACGTGGTCTTGACGTTT
CAACAATTTACACGTTATTA ACTTCGGCTTACCAATGAAGCATGAAGATT
ATGTACACCGTATTGGTCGTACAGGTCGTGCTGGTCGTACAGGTCAAGCG
ATTACTTTAGCGACTTATCGTGAACGCGGTAAAATCCGTGCGCTTGAAGA
CTACTTAGAAGCTCGTTTAAATGTTTCTGAAATTGAAGTTTAGAGCCATC
TCCACCTCCAGCTCGTCCACGTCGTGAAGGTGGTCGTGGTGGTAACGGT
GGCCGTGATGGTCGTGTTTTCGAGGGTGAAGGTA ACTTCAAACGTCGTGA
AGGTGGTCGTGATGATCGTCCACGTCGTAGCTTCGATGATAAGCCTCGTG
GTGAGCGTCCTGCTTTTCGGTGGCGAAGATCGTCCACGTCGCGAGTTCAA
CTCAGATCGTCCACGCCGTGAAGGTGGTTTTAACGATAAGCCTCGTTTTG
ATTCAAATGATGACAATCGTGGTAACCGCGTAGACTATAAACCACGTCGT
GAAAACGGTTTTCGGTGACCGTCCACAACGTAGCTTTGGTGGTGAAGATCG
TCCACGTCGTGAAGGCGGTTTTGGTGATCGTCCAAAACGTAGCTTTGGTG
GTGAAGAACGTCCACGTCGCGCAGTTTCGCGAAGAACA CTTTAATCAAGAA
TCTCGTGGTGAGCGCCGTCGTAAATTCGACCGTTAA
ATGTCTAAAGGTCAA ACTTTACAAGATCCGTTCTTAAATTCTCTCCGTAAAG
AACGCATCCCAGTTTCTATTTTCTTGTTAACGGTATTAATTACAAGGTCA

	g protein Hfq	C6			TATTGAATCTTTTGACCAATATGTTGTTTTATTA AAAAATACTGTAAGTCAAA TGGTTTACAAACACGCAATTTCTACAGTTGTTCCAGCTCGCAACCCACGTC CAGCAGGTGCACAAGGTGCAGGTTTCCAGCTCAGGGTGGTAGTCAAGG TGGCTTCGGTGGTCAAGGCGGCTTTGGTGGTCAAGGCGGCTTCGGCGGA CATCAAGGCGGTTTTGATAACGATTCTAAATTTGAAGATGGTCAAGACGAC GAAAACAATCGTTAA
<i>group_174</i>	hypothetical protein	A1, A2, A14 L6, L8	S	AAA domain	ATGCGACAAGCAATGTTAATGAGAGCTAAAGCATTGAATTGCACGTTTGT AAACAACGTGGCACTTGGATTAGTCCCCCTGAATTTAACGGCATCTCAGAT CAACAAAGAGACGAACCTTCAAACCTTTATTGCTGAACGTGGCCTAGACGTA AAAACAGTATGTGAGCACTTAGGTATCGATGCCCTTATTCAAATTGAAGCA GCAAACTTAAGGCAGTTAAACAAGACATTGAAACATTAGCTAAAACGGG GATGACAGCATGA
		A2, A3, A14			ATGAACCCATTCAAAGGCCGGCATTTCAGCGTGACATCATTCTGTGGGC CGTACGCTGGTACTGCAAATACGGCATCAGTTACCGTGAGCTGCAGGAGA TGCTGGCTGAACGCGGAGTGAATGTCGATCACTCCACGATTTACCGCTGG GTTTCAGCGTTATGCGCCTGAAATGGAAAAACGGCTGCGCTGGTACTGGC GTAACCCTTCCGATCTTTGCCCGTGGCACATGGATGAAACCTACGTGAAG GTCAATGGCCGCTGGGCGTATCTGTACCGGGCCGTCGACAGCCGGGGC CGCACTGTCGATTTTTATCTCTCCTCCCGTCGTAACAGCAAAGCTGCATAC CGGTTTCTGGGTAAAATCCTCAACAACGTGAAGAAGTGGCAGATCCCGCG ATTCATCAACACGGATAAAGCGCCCGCCTATGGTCGCGCGCTTGCTCTGC TCAAACGCGAAGGCCGGTGCCCGTCTGACGTTGAACACCGACAGATTAA GTACCGGAACAACGTGATTGAATGCGATCATGGCAAACCTGAAACGGATAA TCAACGCCACGCTGGGATTTAAATCCATGAAGACGGCTTACGCCACCATC AAAGGTATTGAGGTGATGCGTGCCTACGCAAAGGCCAGGCCTCAGCATT TTATTATGGTGATCCCCTGGGCGAAATGCGCCTGGTAAGCAGAGTTTTTG AAATGTAA
<i>group_18</i>	IS6 family transposase IS15	L6, L7, L8, L10	L	Transposase and inactivated derivatives	ATGAACCCATTCAAAGGCCGGCATTTCAGCGTGACATCATTCTGTGGGC CGTACGCTGGTACTGCAAATACGGCATCAGTTACCGTGAGCTGCAGGAGA TGCTGGCTGAACGCGGAGTGAATGTCGATCACTCCACGATTTACCGCTGG GTTTCAGCGTTATGCGCCTGAAATGGAAAAACGGCTGCGCTGGTACTGGC GTAACCCTTCCGATCTTTGCCCGTGGCACATGGATGAAACCTACGTGAAG GTCAATGGCCGCTGGGCGTATCTGTACCGGGCCGTCGACAGCCGGGGC CGCACTGTCGATTTTTATCTCTCCTCCCGTCGTAACAGCAAAGCTGCATAC CGGTTTCTGGGTAAAATCCTCAACAACGTGAAGAAGTGGCAGATCCCGCG ATTCATCAACACGGATAAAGCGCCCGCCTATGGTCGCGCGCTTGCTCTGC TCAAACGCGAAGGCCGGTGCCCGTCTGACGTTGAACACCGACAGATTAA GTACCGGAACAACGTGATTGAATGCGATCATGGCAAACCTGAAACGGATAA TCAACGCCACGCTGGGATTTAAATCCATGAAGACGGCTTACGCCACCATC AAAGGTATTGAGGTGATGCGTGCCTACGCAAAGGCCAGGCCTCAGCATT TTATTATGGTGATCCCCTGGGCGAAATGCGCCTGGTAAGCAGAGTTTTTG AAATGTAA
<i>group_20</i>	IS6 family transposase IS26	C2 L7, L8, L9	L	Transposase IS66 family	ATGAACCCATTCAAAGGCCGGCATTTCAGCGTGACATCATTCTGTGGGC CGTACGCTGGTACTGCAAATACGGCATCAGTTACCGTGAGCTGCAGGAGA TGCTGGCTGAACGCGGAGTGAATGTCGATCACTCCACGATTTACCGCTGG GTTTCAGCGTTATGCGCCTGAAATGGAAAAACGGCTGCGCTGGTACTGGC GTAACCCTTCCGATCTTTGCCCGTGGCACATGGATGAAACCTACGTGAAG GTCAATGGCCGCTGGGCGTATCTGTACCGGGCCGTCGACAGCCGGGGC

group_
21

Alpha-
ketogl
utaric
semial
dehyd
e
dehydr
ogena
se

A2, A3,
A14

L7, L9

C

Aldehyde dehydrogenase family

CGCACTGTCGATTTTTATCTCTCCTCCCGTCGTAACAGCAAAGCTGCATAC
CGGTTTCTGGGTAAAATCCTCAACAACGTGAAGAAGTGGCAGATCCCGCG
ATTCATCAACACGGATAAAGCGCCCGCCTATGGTCGCGCGCTTGCTCTGC
TCAAACGCGAAGGCCGGTGCCCGTCTGACGTTGAACACCGACAGATTAA
GTACCGGAACAACGTGATTGAATGCGATCATGGCAAACCTGAAACGGATAA
TCGGCGCCACGCTGGGATTTAAATCCATGAAGACGGCTTACGCCACCATC
AAAGGTATTGAGGTGATGCGTGCCTACGCAAAGGCCAGGCCTCAGCATT
TTATTATGGTGATCCCCTGGGCGAAATGCGCCTGGTAAGCAGAGTTTTTG
AAATGTAA

ATGACCATTATCGGACACAACCTTTATTGGTGGTTCACGCAGCGCACAAAGG
TACAACCTTATTA AAAAGTATTCAAGCTACGACTGGTGAAGCTTTGCCTTA
TGAGTTTCACCATGCAACCGAGCAAGAGATTAATCAGGCCTGTGAAGCAG
CCAGTCAAGCCTTTAAGACTTACCGCCATACCACACCTGAACAGCGTGCT
GTTTTTTTAGAAAATATTGCAGATGAACTCGATGCTTTGGGTACCGATTTTC
TAGAAATTGTCTCACAGGAAACTGCTTTGCCACTTGACGTTTACAAGGTG
AACGTGCCCGGACCAGTGGGCAAATGCGCCTGTTTGCTAAAGTGCTGCG
CCGTGGTGATTTTTTAGGTGCCCGTATTGATACCGCTTTGCCTGAGCGCC
AGCCTTTACCTCGCCCAGACTTGCGTCAGATTA AAAATCGGTGTAGGCCCT
GTTGCAGTCTTTGGGGCAAGTAACTTTCCTTTAGCATTCTCAACTGCGGGT
GGTGACACTGCTTCGGCACTGGCAGCAGTTGCTCTGTAGTGGTGAAG
CCCATAGTGGACACATGGCCACAGCAGACTTTGTAGCGCAGGCAATTGAA
CGTGCGGTAGAAAATCAAATATGCCAAAAGGGGTGTTTAATATGATCTAC
GGTAATGGCGTGGGTGAACCTTTAGTGAAGCATCCTTTAATTCAGGCTGT
GGGTTTTACCGTTTCATTGCGTGGTGGTTCGTGCTTTATGTGATATGGCCG
CAGCCCGTCCACAACCGATTCCAGTCTTTGCTGAAATGAGCAGTATTAAC
CCGATGTTGATGTTGCCCGAAGCCTTAAAAAACCGTGGTGACAAAATTGC
ACAGGACTTGGCTGATTCAGTGGTTTTAGGCTGTGGTCAGTTCTGTACCA
ATCCGGGTTTAATCTTGGGTATTAATCAGCTGAATTTAGCCAGCTTATTA
GCAACCTGACTGAAATTATGGGGGCTAAGCCTGCACAGACTATGCTCAAT
GCCGGCACCTTAAAAAGTTATACCGCTGGTCTTGAGCACTTAACCGAGCA
TCAAGGTATTAAGCATTGCGCAGGTCAAACCTCAACAAGGCAATCAGGCTC
AGCCACAACCTGTTTAAAGCTGATGTTGAGCTTTTACTGGCAGGTGACCAG
CTTTTACAGGAAGAAATCTTTGGGCCAACACAGTCATCATTGAAGTTGAA

<i>group_22</i>	Alpha-ketoglutaric semialdehyde dehydratase	L6, L10	A1	C	Aldehyde dehydrogenase family	<p>GATAAAGCCCAACTCATTGAGGCATTACAAAGCATGAATGGGCAACTGAC TGCGACTTTAATTGCCGATGAGGCTGACTTAACCGAGTTTGCAGAGGTAG TTCCTGTATTAGAAGAAAAAGCTGGGCGATTACTTATAAATGGCTATCCAA CAGGTGTTGAGGTGTGTGATGCCATGGTACACGGCGGGCCATACCCAGC CACTTCAGATGCACGAGGTA CTTGAGTTGGAACCTTGGCCATTGACCGTT ACTTGCGCCCTGTGTGTTACCAGAACTATCCGCAAAGTTTATTGCCGGAA GCCTTAAAAGACAGCAACCCGTTGCAGATTTAAGATTGGTGAATGGTGA GATGACCAGAGAAGCGATCTAA</p> <p>GTGGTCAAAGCCCATAGTGGACACATGGCCACAGCAGACTTTGTAGCGC AGGCAATTGAACGTGGCGTAGAAAAATCAAATATGCCAAAAGGTGTGTTTA ATATGATCTACGGTAATGGCGTGGGTGAACCTTTAGTGAAGCATCCTTTAA TTCAGGCTGTGGGTTTTACCGGTTTCATTGCGTGGTGGTTCGTCGTTTATGT GATATGGCCGCAGCCCGTCCACAACCGATTCCAGTCTTTGCTGAAATGAG CAGTATTAACCCGATGTTGATGTTGCCGAAGCCTTAAAAACCGTGGTG ACAAAATTGCACAGGACTTGGCTGATTCAGTGGTTTTAGGCTGTGGTCAG TTCTGTACCAATCCGGGTTTAATCTTGGGTATTAATCAGCTGAATTTAGC CAGCTTATTAGCAACCTGACTGAAATTATGGGGGCTAAGCCTGCACAGAC TATGCTCAATGCCGGCACCTTAAAAAGTTATACCGCTGGTCTTGAGCACTT AACCGAGCATCAAGGTATTAAGCATTTGGCAGGTCAAACCAACAAGGCA ATCAGGCTCAGCCACAACCTGTTTAAAGCTGATGTTGAGCTTTTACTGGCAG GTGACCAGCTTTTACAGGAAGAAATCTTTGGGCCAACACAGTCATCATT GAAGTTGAAGATAAAGCCCAACTCATTGAGGCATTACAAAGCATGAATGG GCAACTGACTGCGACTTTAATTGCCGATGAGGCTGACTTAACCGAGTTTG CAGAGGTAGTTCCTGTATTAGAAGAAAAAGCTGGGCGATTACTTATAAATG GCTATCCAACAGGTGTTGAGGTGTGTGATGCCATGGTACACGGCGGGCC ATACCCAGCCACTTCAGATGCACGAGGTA CTTGAGTTGGAACCTTGGCCA TTGACCGTTACTTGCGCCCTGTGTGTTACCAGAACTATCCGCAAAGTTTAT TGCCGGAAGCCTTAAAAGACAGCAACCCGTTGCAGATTTTAAAGATTGGTG AATGGTGAAGATGACCAGAGAAGCGATCTAA</p> <p>ATGGCACTTAATGTTTTTAAATTCAAAAAATATGCAAGGATGTTACACTAT TAAATTTAATTTGCTTTTATCTATCTGGCTAGGTTTATTTCTGAATATAGGT TTTTTAAAAAATCCATCAACTTACACCTTATAATGGTATTAAGTCAGTTCT TTTCTTAGGGGCGACATTAGTTATTTAATAGCGGCATATAATTTAATTTTT</p>
<i>group_30</i>	<i>eptA_2</i> Phosphoenolpyruvate carboxylase	L6, L8, L9, L10 C2, C6		S	Domain of unknown function (DUF1705)	

transfe
rase
EptA

CAATTAATAAATTGGAAATGGACTGCCAAAATCTTTGCAATTTTATTGATAT
TTATTGGTGGCTTTAGCTCTTATTTTGTAACACATTGGGTGTCATTATTTTC
ACCCGACCAAATTCAAAATATGGTGCAGACCGATGTTTCGGAATTTACCGA
TCTAATCTCTTTACGCTTTGTTTTATGGACAGTTTTTTTTGTTATTTGCCCA
TTTTTTAATTACTCAAGTTAAATTTAAACAAGAAAAAGCATCACGGTTGTT
ATTGAAGAAAGTATTCTCACTGGTAGCTTCATTTGCAGTGGTCGGTGTTTT
ACTTTTTACTTACTATGTCGATTTTCGCTGCAATATTTTCGTGAGCATCGTGAT
TTAAAAGGGATGATTTACCGCAAATAGTATTTTCATCGCTTATGTCTTACT
ATCATAAGAAGGCTCCGAAGAAAAATCTGCCTCTTGTGATATATGGACAAG
ATGCTCATCAAGTTCAGCAAGTACAAAAGAACCTCCCTAAGTTAATGATAC
TTGTTGTAGGTGAAACGGCACGTGCCGAAAGTTTCTCTCTAAATGGGTAT
GCAAAAATACGAATCCGGAGCTTTCTAAACAAGATATTTCAACTTTTCG
CAAGTGAGCTCATGCGGTACGGCGACAGCTGTTTCTGTGCCATGTATGTT
CTCGGGTATGCCACGTGTAGATTATGATGAGCAATTAGCTAGTCACCGCG
AAGGTTTATTAGATATTGCAAAACGTGCGGGTTACCAAGTGACTTGGATTG
ATAATAACTCGGGTTGTAAAGGTGCATGTGATCGCGTTGAACAATACCAG
ATTCCAGAAACTTAAAGAAAAAATGGTGTAAAGATGGCGAATGTTATGAT
GACATTCTCATTGACAGCTTAAAGCAGTATTTGGCTACTATTGCCAAAGAT
GATGATCGTCCGCGTTTTGATTGTTCTGCATCAAGTGGGTAGTCATGGACC
TGCATATTACAAGCGTGCACCTGAGGCATATCAACCTTTTAAACCGACTTG
TGATACTAATGCGATACAGGGCTGTTTCGCAAACCTGAATTGCTAAATAGTTA
TGATAATACAATTGTATATACAGACCATGTATTAAGCCAAATGATTAATACT
CTAAAAGAAATATCAAAGTATCAGACAGGTTTATGGTATTTATCTGATCATG
GCGAATCAACCGGAGAACATGGTTTATATTTACATGGTTCACCTTATGCAA
TCGCACCGAGCCAACAAACACATGTACCAATGATTATGTGGTTCTCTGAAA
GTTGGAAACAACATAATCTTGCTCAAGTGAATTGTTTAAAGCCAACAAACTA
AACAAAAGTTAAGTCAGGATAATTTATTCCCAAGTTTGTAAAGTTTGTCTGGA
TGTA AAAACTCAGGTAGTAAATAACAACTTGATATGTTGAGCCAATGTAA
ATAA
ATGGCTCTACCGATTACTGCTGACCAAACCTTTATTGGTTCAAGCAATT
ATTGTGTACCTATACGCTGATCCGGGTTTAGGTAAATCATCAATGGGCTTT
ACTGCGGAAAAAGCAATTTCTTTTGACTTTGACCGTGGTGCTCACCGTACT
GGTGAATTACGTCGTGGTGGGTTGTACAGTTCAACAATGGAGTGATGT

group_
3194

hypoth
etical
protein

A3, A14
L7, L8,
L9, L10

S

AAA domain

					TGCAAACCTTACTCCGCAGGACTTAGCACCATATAAAACCGTAGTCATTGA TACCGTGGGTGCAATGCTTGAATGCATTAACCCACCTGTTACTTACGG CAAATAACCGTCAAAAAGATGGTTCTTTAAAGTTAAAGGCTCAAGGATTAG CGAACCAAACGTTCAAGCAATACATCAATACTTTGATCAGTTTAGGTAAG ATGTTGTTTTATTGCACACGCATCAGAAGATCAAAACGGTGATCAAATTA TTTACCGCCCAGATCTAGGTGGTAAAAACCGTAACGAGCTTTACCGTATC GCAGATGTCATGGGTTATCTAACAACCTGTTACTACTGGTGAAGGTAATAA GCCCCGCTTATTAATTTCAAACCTTCGCCTACACATCATGCGAAAACTCA GGTGCTTTAGGCGGTGAAACCGGTGAAGTATGGGTACCTGATCTTAAAGC ACACCCTACTTTCTTGGCTGACCTGATTACTCAAGCTAAAGATCACATTAA CACCTTAACGCCTGCACAACCTTGCAGCAGCTAAAGCCCAAGAAGAGCTAG AAAACCTGGAAACAAAGCTGTGAAGAAGCTGAGCATGCAGGTGACCTTAAT CAATTAACCTGAGTCGCTTGATAAAGAACACATGTATTACCAGAACATGCGA CAAGCAATGTTAATGAGAGCTAAAGCATTGAATTGCAGTTTGATAAACAA CGTGGCACTTGGATTAGTCCCCCTGAATTTAACGGCATCTCAGATCAACA AAGAGACGAACTTCAAACTTTATTGCTGAACGTGGCCTAGACGTAAAAAC AGTATGTGAGCACTTAGGTATCGATGCCCTTATTCAAATTGAAGCAGCAAA ACTTAAGGCAGTTAAACAAGACATTGAAACATTAGCTAAAACGGGGATGAC AGCATGA
		A2, A3, A14			ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTATTGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT TTGATTCAACGCTTTATAGCCATTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG
<i>group_</i> 3291	IS4 family transp osase ISAb 33	L6, L7, L8, L9	L	Transposase DDE domain	
<i>group_</i> 3292	IS4 family	A1 L10	L	Transposase DDE domain	ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTTAAA

	transposase ISAb33					GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTATTTCGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTACCAATCA TTTAG
<i>group_3510</i>	hypothetical protein	L8 C2	/	hypothetical protein		ATGAGCAAAGTATTAATGAATTGCCTGCAAGCGCTAGCAATAACGAATCG CTCATATTGCAAGCACTTAACGCTAGCAATCAAAGACAAGTAGCAGAGAT GATAAATGTTCGATGCAAGCATCCTTTACGGATGAAAACAGAAAAGAAATC AAATGGATGGACTGAGATTGAGTTTATTAGCTTTTTGTTGACAGCCATTGG TTTGAAGGTTGTGCAAGAAAGTGATGTGTATTGCTCACCTGAAATTGCAGA AGCAACGCGAGTTTATTTAGCACATGCATTCACTTCACCTGAATACATGCG GATTTTATTCAAATAA
<i>group_40</i>	hypothetical protein	A1 L6, L7, L8, L10	S	Bacterial collagen, middle region		GTGTCTGATTTAGTCGGTGGTGATATTCAAGGCGCTACTGAATCGGCTAC CGGTATTCTCGATACGCTTATTACAAATGGTACAACCTGCAACTGGTTTAGT AACTGAAATCATTGGTGGAGTAACAGGAATATTGGTGGCGTAACTGGCG GAGATAGCCCACTAGGTTTAGTAACTGATTTACTTGGTGGTTTAACTGGTA GCGTAGGCGGCGACAGTCCGCTTGGTGTAGTCACTGACCTTCTCGGTGG TTTAACTGGTAGCGTAGGTGGCGACAGTCCGCTTGGCGTAGTCACTGATC TGCTCGGCGGTTTAACTGGTGGTGTACTGGCGGTACAGACAACCCTATC GGTATCGTAACGGATATCGTTGGAAGCTTAACGGGTGGTGTACTGGTGA AGGCAGTTTAGATGTTATTTCTAATTTGCTAGGCGGATTGACTGGTGGTAG CTTACTCGGTGGTGTAACTTCTACAGTATCTAGCGTAACAAATACAACGCA TACGATCGTACCGACATCATTACTTACTGATAACTTCTAGAAAACCTCATT AATACTGTTTAA
<i>group_53</i>	hypothetical protein	A3, A14 C4, C6	D	Phage-related minor tail protein		ATGGCACAAGAATCCCGTTTGGTCATTGTTATTGATTGCGAAAATGCTGAA CGTAATGCGCGTAATCTAGGCAATGAACTTGTTAGCATTGAACGTAAAGGT GAATTTGCATCTAAGTCTATGGACAGCTTGTCTGTAGCCACCAGAGCTTTA GCTGGACACATGGCTGGTTTATTAACAGTAGGTTTCAGCCATTTCAAAGATG

GATACATATACTGGATTACAAAATCGCCTTAAGTTAGTCACTAACAATCAA
GTTGAACTAAATAAAGCAACGGAAGACACTTTCCGAATTGCTCAAAAACC
TATTCAGCTTGGGATTCTGTGTTACAGGTTTACCAGCGTTTTAGTGATAAT
GCCAAAACTTTAAACCTCACAATGGATGACACAGCACGTTTAACTGAAACA
GTTTCTAAAGCTGTAGCAATTAGTGGTGCAAGCGCAGAAGCTGCTGATGC
AGCTTTAGTTCAGTTCGGGCAGGCCTTGGCTAGTGAACGTTGCGTGGA
GAAGAACCTAATTCTGTAATGGAGCAAACCCAGCACTAGCAAAGGCTATT
GCTAAAGGTATGGGTATTACTGTAGGTGAATTACGTTCACTAGCAGCTGA
AGGAAAATTACTTCACAAGAAATTGAAAAGCGCTTAGAAATGTAGAATC
TGATGTTGATGCTCTTTTTGCTAAAACAGATATCACAATCGGGCAGTCTCT
CACACTCCTAAACAACGAGATCACAAAATTTGTTGGCGAAGCAGGTAAGG
GAAGTGGTGCGGCACAGGTATTAGCTGGATCAGTTCAAACCTTTGCAAGT
AATTTAGATTTAATTGCTGATGGGGCTTTAGTAGTTGGTATTGGATATATCA
CTCGTGCAATTTTATGAAGAGCGCTGCTATTAAGAGGGAATGGCTTCA
ACTTTAGCGAGCCGCCAAGCATCTGTATTAATGCTCAAGCAGAATATGCA
GAAGCTACCGCTGCTTTGAATGCAGCAAAGCTCATCTCGCGAATGTGCG
AGCAACAAATGCAGAAACCAAGCTAAATTTGGCGCAACAGCGGCAGCAA
CTCGATACGCACAAGCACAGGCAGCAGTAACTGCTGCTACAAATGCACAA
ACAGCAGCTCAAATTAAGCTAAATACTGCAACTTCAATTGCAGGGAGACTA
GCTAAAGGGGCGTTTGGATTAATTGGTGGGTGGGCTGGAGTTGCAACATT
AGGAGTAATGGGATTAGCGGCAGCCTATTCTTATTTAATAATAAGGCAGA
GGAGGCAAAGCAAAGCTTGCTGAACAAGCTAAAGTTGCTGAGAAAGCTG
ATGAGGAGTTAAAAAATTAAGTGGCAATGATAAGGCTAAAGCAGTTAATG
ATTTAACTACTGCTTTTAAATGCACAAAATAAAGCATTAGAGAAATCATCGCG
TGCTGTAGGGTCTGCATTAATTGATATCGAGAATATGCACGAGGAAATA
GGGAGGTTGAAAAATTTCCAAGAAGCGAGAATGGAATATCAGCTAT
ACAGAAGCCATTGAACGTCTAAATAAAATTAAGTTGCCTACAGATCTATAT
GAAAATCTGAAAAACAGGCTGCGCAGTATGATGACAATGCATCTAAAGC
AAGTTTATCAGCTGAGAACTTAAATTATTAAGAGTTGAAGTGAACCTGG
AGGTAATGAAGCACAAAATGCGGCAATTCAGCATCAAAAACAAGCGGATG
CTTTAGGAAATACTGCTACTGAAGCAGAAAAGGCAACTAAGGCTTTGCAA
GATTATCAAGCCAAGCAAAGATAGCGTTATTGATTCAATCTATAAATCA
GGTTGGCTTGATAAAGGTTACACTGTTGCTCAAGCTAATGCCATTTTAGAA

CTGCAAAAAGCTAAAGGAATGAGTGCAATTTTGTCTAAAGATGAAATTGAT
AGCGCACTTAGAAATCTCAAGATCATCGAAGAACAACAGGAGCGAGAAGA
TAAATTAAGTGAAGCTAAAAGAAAGCAGACGCAGGAAATTGAAAAACAAGC
AAAACCTACTAAACGCTTGGTCGGTATTTCCGGTCAATCCGGTATTGGTAC
TGGTCCACATCTTGACGTCCGATATGGTGGCTCATTGTCAGGTCAGAAAG
TTTCTAATGAACATCTGGCTCGATTACAGGCGGGAGGAAAACCTTTAACTT
CCTACAAGATCAGTTCTAATTATGGTCCACGAAAAGCCCCAACTAAAGGG
GCTTCTTCATTTCATAAAGGTATTGATTTTTCAATGCCTGAAGGAACACCAA
TCACGACCAATGTTGCTGTGAAAGATATCAAGACATGGTATGACAGCAAG
GGAGGTGGTTATGTCAGTGAAGTATCTTTGAGGATGGAGTGTCTCTTAA
GCTTCTACATCAATCTCCCAAGATGCAGAGCAAGGTGAAAGGTGGTGCAA
GTAAAGGAAGTGATAAAGCAGCTGGTGATATTCAATCTCAACTTGAACGTC
AACAGGATTTGCAACGGTCACTTGAAAATGAGGTGGCTAGTGAAGTCGGA
CGGATTAACAATAATAGAAAGGCAAGACTGGAGGATGTTGATAAAGCAA
CTTTAGCCCGGAACGTAAGTGCAGAAATAAAGGCGGAAATAAATCGTCGTG
CAGATAATGATATTGCTATAGCCAAACAAGCCCTTAGAACGAAATTGGAAG
ACTATAAGGAGTTCCAGAAAACCGAGGAACAGTTACTAGAAGAGTCCTTTA
ACCGTAAAAAGTTCAATGCAGCTCATGACCTTGAATTAAGTAAGTTTGAGC
AGAAGCAAGCTGTTGAATTGCTGGAACAGCAAAAACAGCAAGAGTTAGGG
TTATTA AAACTAGCTCAGGAACAGCGGTTGTTTCAAGCCCGTTTATCTCTG
CTTTCTGAAACGCAAGCCATGCAGGAACGTTACAGACTCGAACGGGAGGA
AATTCTTAAGAATACCAAGCTTTCTATAGAAGAGCGGCAAAAGCTAATCGC
ATTATCTAAAGCCAATCAGGATAAAGAGACACGCGATAAAGTGAATAATGC
TGTTCAAAACTGGGGTGGTATCCAAGCGGATATGAATGGTACCGGAGAAT
TTTTCAGACAGGATCAGGAACGATTTAGCCGTTTAAATGCTGCAAATGATT
TAGCAGATAGTCAATTTGCTGCTACTGATCTTGATGAAAAAATGGTTTAG
ATGTTCTAAATGCACACATGGAAGCAGGACTCATCAAGCAACAGGACTTC
GAAAACCGGAAAACAGCTATCATTCAAGCTGCTCAGGACCAACGCAATCA
GATCGCTGCCGAATATGCTCAGAATGCTCAGGATATTGAAGATAAGTATCA
CCAAGATCGATTGAATGCTCAAATTGCTCTTGGTGGCCAAATGATGGGTT
CACTCACATCGATGTTTGGTTCAATGTTTGGCGAGCAATCAAAGCATACA
AGATCATGTTTGGCTGCTGATAAAGCTTATGCCATTGCAGCTGCAGGTATTT
CTATTCAGCAAAGTATTGCAAAGGCGGCTAGTGTTGGTTTTCCAGCAAATA

TCCCATTAATTGCAAGTGCTATTGCACAAGGTGCAAGCATCATTGCAAACA
TCCGGGCAATTAAGATCAAGGCTTTGCTGACGGTGGTTACTGATCA
GGTGGGAAATATCAGCCTGCTGGTATTGTCCATAAAGGAGAGGTGGTCTG
GTCCCAAGAAGATATTAACGCTGGGGCGGTGTTGGCTTAGTCGAGAAAA
TGCCTAAGAGTGCAAACCCTGAAGCTTTTCTCAATAACAATGCCTTGGCA
GATAGTGTCATGCGCCGTGCAATGATGAGCTCTAGTGCCTTTATAGAAAAG
CCAAAAGCAGGCTGACATCTTTAATCAACCGGTTCAAGATACTCAGATTAT
CTATAAGGGTAATAGAGACACACCTAAATTAGCTTCTTCTGGAAATTTAGA
CTTATTCCATGATGGCAAGGTCTACTTCTCATCCAATGGTTTAGTTCAGGA
TCGTTCAAATCTGGATGATGTTCCAGGACTTTACTTTAGGAAGTACTTCACG
CCCTCAAGCTGAGATGATGCCTTCAATTGAGCCAGCTTCACCGACAATCA
ATTTCAAATTGAAGTGATTAATCAGGTGAGTGGGGCGACAGTTGAAGCT
GAACAACCTGGATGAGCAAACAGTCCGGATCATTGTTAAAGATGAACTGGA
TAAGCAGCTTCCAAGAACGGTACCTAAGCTTGTAAAGTATCAAATCGCAA
TCCAAACTCAACCATTAGTCGGTCTTTGACTGAGAATACGACAGCGAGAA
GAAATCGTACTTAA

ATGATTAAGTTATTTATAAAAAAGACGCTTTGTCTGAAGAAAAGACAATTG
AGCAGGCTCAAACCATTGGGCAATGGCTCACTTCAAATATGAACATATG
CCTGAGCATGTCCGTATCTTTCATACTACAAGCAATATGGATCATGCCGAA
ATTTCAATTTGCGAATGAAGTCACACCAAAGAATGCATATGACTTAAAGCAG
CTTGATTTCTTACCGGGCACTTTTATCGTAGTTGAGAACCCTAAATGGGTC
GCGGCTATTGTTTCGATTGTGATTAGTATTGCGATCGCATTTTTAATGCCA
ACGCCATCAATAGCACAAACGACTCAAATACTAACCAGTCTTCTTCAGCA
AACAATGAACTTTCTAACCGGGAAAACAAGATCCGGGTGAATGGTCGTAT
TGCTGATAACTATGGAGCTGGGTGGAATACTCCCGACCTAATCGCAGTAC
CTTACAAGGTATATGAAAACAACGTTGAAGTTGAGCATGTAGTGGGCTGTA
TTGGGCGTGGACACTATAAAATCAATGGAGCTTATGACGGTGAACCAAT
ATTGTCGATATTGCTGGCGCATCGGTAGAAGTCTTTCGACCAGGTGTAGA
TATTGTTTCAGGTGAGCCATATTTCTCGCTTGGTACCGAAATTACCACGCC
GCCACTAACGGTTCAGCATCAAACCTTCTGTTAATGGCCAAGTTTTACGTCC
TGCTGATACACAATCTTTAGAAGGTACGAACTACCTTCATTTTGCATATCC
AAACGAGATTCTTCGGGCAACGGCAAACAACACAGATTTAACCCTAAGTT
TGTAAGTAATGACCGCGTAGAAATCACCAATGCCTCATTACGTTTAAATGG

A2, A3,
A14

group_
6

hypoth
etical
protein

L9

S

cellulase activity

CCAGACTTTTGATTTAAATGGTACTTATAGCGTTCTATCGGTAGCTGATGA
CCGTATGACGTTATCAAATCCGGCGGCCGTTAATGCTAACTGGTTAAAGC
TTAAAGAGTTAAATAACCAACAACTGCAGCTTTGTCACCAAGATCAGTT
CAATAGGTGAAAAATGGATTGGTCCATTCTGGACAATGTTGAACGTA
GCCGGGTGCTGTGTAATTTTGTGGCCACCAATGGACTTTATACCGTTTCTT
CAGGTGGGTATCAGGCCGCTGTTAATGTCACGATTGAAGTTGAAGTAACA
CCGGTAAATGAATCTGGTGCAGCGATTGGTAATCCGATGCTGAAGCAGAT
CATTTTGAAAGGTTCCGGCAAAGTCACGTCAGACCGTTGGTGCAACACTTG
ATATGGTCACGTTTCAGGGCGTTGTAGTGTCCGTGCACGCCGTTTAACT
CCGACTCCGACAGTCACAACAGTTGTTGATGAAGTAAAGTGGCAGGCGCT
TTACGGTGCTTATCCTTTACAAAGCACAGTGTATGAACATGAAACGGTTTT
TCGTGCGCGTACTTATGCAACCACTGGAGCTTTATCTGTTAAGTCCCGCA
AGATCAATTTTGATCTTCAGCGAATGTTGCCGACTTATAAAAACGGGGCAA
TGACAACAGAGCTATATCCAACGTCTAGCTTTGCTGATGCTTTGGTATCTA
TGGCACTCGATGACAAGATTGGCCGCCGTTTCGATCGATGAGATTGATCTT
GAAAACATCTATCGGACCTATAATGATGTAGTTGATTATTTTGGTACGCCG
CTAGCGGCTGAGTTCTGTACTACCATTGATGATACGAATCTATCTTTTGAA
GAGCTGGTTACCAATCTTTGTGATGCGGTGTTTTGTA CTGCATATCGGCAA
ACAATAAGCTCAAGCTTTATTTTGAACGGCCA ACTGATAACTCGGTAATG
CTGTTTAACTTCAGGAATATCATTCCGGATAGTTACAAGCATGACCTGACC
TTTGGCGTGATGGATGACTACGACGGACTGATCTATGAATACACGGATCC
GACCGACGATAGCCGTATCAATATCTATTTACCGGATAAAGGAGCCAAAA
ACCCAAAAGAGGTGAAATCTGTTGGTGTACGAAACAAGTGGCAAGCGCAT
TTCAATGCATACCGGATTTGGAACAAGATGCGCTTCCAGCGCAAATCCATT
ACCTTTGATGCGGCACCAGAATCAGAATTACTGGTTTTACGTGACCGGATT
GCTGTAGCGGATTATCGCAATGGTATTCATCAAAGCGGTGAGGTGGTACA
GCAAGAAGGTTTAGTCCTCACCTTAAGCCATGATGTAGATTTTCATTGCAGG
CAAGAGCTATGTGATTTATCTGCAAATGGCCGGATGGCACAGTGGACCTGA
TTCCTGTTACACCGGGTTCAGCCAAAAATAAGGTGGTTTTAGGGCGGTTG
CCGAACGGGGCCTTAAAGCTAAGTCCTGATGATTTTGTAAATACCATCTAT
ACAGTAGTTAATGATGATACTAAAGGTTTCATTGCCTTATCTGGTTGCAAAA
AGAGAACCGGTTGACCAGTTCTCTAATACTATTACAGCAATTAATTACGAT
GAGCGCTATTACCTCAACGATAAAGACTTTATTGACGTGCCGGTTGATGAT

<i>group_623</i>	<i>hfq</i>	RNA-binding protein Hfq	A1, A3 C2, C4	J	RNA chaperone that binds small regulatory RNA (sRNAs) and mRNAs to facilitate mRNA translational regulation in response to envelope stress, environmental stress and changes in metabolite concentrations. Also binds with high specificity to tRNAs	TCTCCAATTTACATTTCGATATGACCAGCTGGATATTAATCTGGCGCGTTTATATCAGATGCAAAGAGGGGATTTGCCAACGACTGGCGAAATCAGTTTTGTAGTTGAATCTGGTGCAGTACTAGTTTCTAGTTTCGAGTTCTTATCGACCGGAAACAGATTTGTCTATAAATTCGACTACAACCTCCAGTCCGCCGAAACAGGAATTATTGCCCTGCAGCGACTGAACTACCCGCCATTGATACTGGTGAGTTCCACCTGATCTGGTGGTAAATCTGACTATTAAGGTGCTGTTGTTGGACGTGGTGGAGATGGCGGGTTGCCACATTTGGCATTGTTGGTGCATGGTCTACCGATCCGGATTATAACTTTACTAAAACCCGCCGTGACGGTTTTTCAGGGAGCACCCGGTTTATTAACCGGCACAGTAACTAAACCTGATTATTGATGGTGGAACTCTGGCTCGAGGCGGCTCAGGTGGTGGAGCAACACCAAGCGGTATTTATACAGGATTATCGTATGGAGTTCAGGGTATTCCCGGTGGAGCTGGAGCACCTTTTGGTTCGGGTTATGACCGGACAACCTATAACGATTCACAAGACTGGCGTTGGTACTTAAATGGTGACTTTATGGTTGTCAAAGTAACCGATGCCGAAGCTTCGGTACCCGGTAAAGGTTACCGAACCCAAATGATCGATATGGATCTCCATTGTCTGGTGGTGGAGGTTGGGGCCAGCGCGGTACCAAGTCCACCAATGATGGAACATGGAACCTGGCAATACCATGGCACAACCTGAAGCCAGCCGGGGCCGGGTGGACCTGCAATTGTTGGGGTGGCACCACCTTACA
<i>group_835</i>		hypothetical protein	L8 C2	/	hypothetical protein	ACTCAATTGATCAATGGAGGGAAAATTCTACAAACACTTTAAATGTCTAAAGGTCAAACCTTTACAAGATCCGTTCTTAAATTCTCTCCGTAAGAACGCATCCCAGTTTCTATTTTCTTGTAAACGGTATTAATTACAAGGTCA
<i>ahpF_1</i>		Alkyl hydroperoxidase	L6, L7, L10	C	Alkyl hydroperoxide reductase	TATTGAATCTTTTGACCAATATGTTGTTTTATTAATAAATACTGTAAGTCAAAATGGTTTACAACACGCAATTTCTACAGTTGTTCCAGCTCGCAACCCACGTCAGCAGGTGCACAAGGTGCAGGTTTCCAGCTCAGGGTGGTAGTCAAGGTGGCTTCGGTGGTCAAGGCGCTGGCTTTGGTGGTCAAGGTGCTGGCTTCGGTGGTCAAGGCGGCTTTGGTGGTCAAGGCGGCTTCGGCGGACATCAAGGCGGTTTTGATAACGATTCTAAATTTGAAGATGGTCAAGACGACGAAAACAATCGTTAA
						ATGGATTTACTCAGATCAAAAAGCTTTAAAACAAAAACACTTTAGATAAAA CAAAAACAGATGTAGGAATGACTTTTTCTTTTGGGCCGTTTCAGCGACACAAT AA
						ATGTTAGATCAAAACATTAATAACTCAATTAAGCTTACCTAGAACGTTTAGAAAGTCCAATCGAATTAGTTGCTGCTTTAGATGAATCAGACAAAGCTGCTCAATTAAGAATTAGTAACTGAAATCGCTGAGCTTTCTGACAAGGTAACAG

e
reduct
ase
subuni
t F

CCCGTTTTGACGGTAACAATACACATCGTCCAAGCTTTGGTGTGGCTAAA
GCAGGTGAACAGCCTCGTGTGTTCTTTGCAGGCTTACCAATGGGCCATGA
GTTTACGTCTTTGATCTTGGCTTTGTTACAAGTGTCTGGCTATGCACCTAA
AGTGTCCGATGAAGTTTTAAATCAGATTAAGGCTTAAACCTCAAAGCTAA
CTTTGATGTATTTGTATCACTCAGCTGTCATAACTGTCCGGACGTGGTACA
GGCGCTTAACCTGATTGCCATTTATAACCCGAACACCACGGCAACCATGA
TTGATGGTGCATTCTTCCAAGACGAAGTGAACAACGAAAATCATGGCT
GTACCGATGGTATTCCAGGACAATGAACACATTGGTCAAGGTCGTATGAC
GCTTGAAGAGATTGTGGCAAACCTTGATACCAATTCAGCTGAAAAGACG
CAGCAGCACTCAATGCCAAAGATGCGTTTTGATGTATTGGTATTGGTGGT
GGACCTGCGGGTGCAACAGCTGCAATCTATGCGGCACGTAAAGGCATTA
ACACAGGTATCGTGGCTGAACGCTTTGGTGGTCAGGTCATGGATACCATG
GACATTGAAAACCTCACTTCTGTGCAAAAAACACAAGGTCCTAAGTTTGCT
GCTGAAATGGAAGCACATGTTTCGTGAATACGATGTGGATATCATGAACCT
GCAACGTGTCAGCAAAATCACTGGTGCAGACCAAACAGCAAATGGCTTGG
TTGAAGTTGAACTTGAACCGGTGCAAACTTGAATCGAAAACGTGATCC
TTTCAACAGGGGCACGTTGGAGAGAAATGAATGTACCCGGTGAAGCTGAA
TACCGTACCCGTGGTGTAGCGTACTGTCCACACTGTGATGGTCCATTGTT
CAAAGGCAAACGTGTTGCCGTGATTGGTGGCGGTAACCTCAGGTGTTGAAG
CGGCGATTGACCTTGCAGGTATTGTTGAGCATGTAACCTCTGGTTGAGTTT
GACACTAACTTCGTGCTGACCAAGTCTTGCAAAACAAGTTAAACAGCTTG
CCAAACACCACTGTGATTATGAATGCATTAAGCACAGAAGTATTGGGTGAC
GGTTCACAAGTAACCGGTCTTAAATATAAAGACCGTGCAACTGATGGAGA
GCATGTGGTAGAGCTTGCAGGGATCTTTGTACAGATTGGTTTACTTCCAAA
CACTGACTTCTTAAAAGACAGTGAGGTTGAGTTAACCAATCGTGGAGAGA
TCATTGTCAATGACCGCAACGAAACCAATGTAAGGTTGATTTGCTGCAG
GTGACTGTACCACTGTACCTTACAAGCAGATCATTATTGCTACAGGTGAAG
GCGCTAAAGCATCACTCTCTGCGTTTTGATTACATCATTCTGTTCTGGGCAGT
AG
ATGTTAGATCAAACATTAATAACTCAATTAAGCTTACCTAGAACGTTTAG
AAAGTCCAATCGAATTAGTTGCTGCTTTAGATGAATCAGATAAAGCTGCTC
AAATTAAGAATTAGTAACTGAAATCGCTGAGCTTTCTGACAAGGTAACAG
CCCGTTTTGACGGTAATAATACACGTCGTCCAAGCTTTGGTGTGGCTAAA

ahpF
2

Alkyl
hydrop
eroxid
e

L6, L7

C

Alkyl hydroperoxide reductase

reduct
ase
subuni
t F

GCAGGTGAACAGCCTCGTGTGTTCTTTGCAGGCTTACCAATGGGCCATGA
GTTTACCTCTTTGATCTTGGCTCTGTTACAAGTGTCTGGCTATGCACCTAA
AGTGTCCGGATGAAGTGCTTAACCAGATTAAGGCTTAAACCTCAAAGCTAA
CTTTGATGTGTTTGTTCGCTTAGCTGTCATAACTGTCCGGACGTGGTACA
GGCACTTAACCTGATTGCCATTTATAACCCGAACACCACGGCAACCATGA
TTGATGGTGCATTCTTCCAAGACGAAGTGGAAACAACGCAAATCATGGCT
GTACCGATGGTCTTCCAGGACAATGAACACATTGGTCAAGGTCGTATGAC
CCTTGAAGAGATTGTGGCAAACTTGATACCAATTCAGCTGAAAAAGACG
CAGCAGCACTCAATGCCAAAGATGCGTTTGTATGATTGGTGGTGGT
GGACCTGCGGGTGCAACAGCTGCAATCTATGCGGCACGTAAAGGCATTA
ACACAGGTATCGTGGCTGAACGCTTTGGTGGTCAGGTCATGGACACCATG
GATATTGAAACTTCACTTCTGTGCAAAAGACCCAAGGTCCTAAGTTTGCT
GCTGAAATGGAAGCACATGTTTCGTGAATACGATGTGGATATTATGAACCT
GCAACGTGTCAGCAAAATCACTGGTGCAGACCAAAACAGCAAAATGGCTTGG
TTGAAGTTGAACTTGAACCGGTGCAAACTTGAATCGAAAAGTGTGATCC
TTTCAACAGGGGCACGTTGGAGAGAAATGAATGTACCAGGTGAAGCTGAA
TACCGTACCCGTGGTGTAGCGTACTGTCCACACTGTGATGGTCCATCGTT
CAAAGGCAAACGTGTTGCTGTGATTGGTGGTGGTAACTCAGGTGTTGAAG
CGGCGATTGACCTTGCAGGTATTGTTGAGCATGTAACCTCTGGTTGAGTTC
GACACTAACTTCGTGCTGACCAAGTCTTGCAAAACAAGTTAAATAGCTTG
CCAAACACCACCGTGATTATGAATGCATTAAGCACAGAAGTATTAGGTGAC
GGTTCACAAGTAACCGGTCTTAAATATAAAGACCGTACAACCTGATGAAGAG
CATGTGGTAGAGCTTGCAGGGATCTTTGTACAGATTGGTTTACTTCCAAAC
ACTGACTTTTTAAAAGACAGTGAGGTTGAGTTAACCAACCGTGGTGGATC
ATTGTCAATGACCGCAACGAAACCAATGTAAGGTTGATTTGCTGCGGG
TGACTGTACCACTGTACCTTACAAGCAAATCATTATTGCTACAGGTGAAGG
CGCTAAAGCATCACTCTCTGCGTTTGATTACATCATCCGTTCTGGGCAATA
A
ATGAGGGTGCCTCAAGCAACAGTTTATATTGATCGCCAAGCGCTTCAATAT
AATTTAAACCGTGTCAAACAACCTCGCAGCAAATTCAAAATCGTAAGCATG
GTAAAGGCCAATGCATACGGACATGGAGTCAAAGACTGTTTAGCAGCCTT
AAATGCCTCAGATGCCTTTGGTGTAGCCTGCCTACAAGAAGGATTAGAAA
TCCGTGAACTCGGTTTGAACAACCAAGTTACTTTAATTGAAGGGGATTTT

alr_2

Alanin
e
racem
ase,

A2, A3,
A14

E

Catalyzes the interconversion of L-
alanine and D- alanine. May also act on
other amino acids

biosynt
hetic

CTGAAGATGAAATGCCCGTCGCAATTGAACAAAAATTTGAATGCGTCATT
ACCATCAACAACAGTTTGAATGGTTAATTAACACAAACAAGCATATATTG
CCCAAGGCTTAAAAGTATGGGTTAAATTAATAGCGGTATGAACCGTTT
GTTTTAAAGACCCAGAAATTATTGAAGTGATTAAGACATTAAGGTCTGA
ATGAGCAAACTTTTGCCGAATTTTCTTGCACGAAACCTTACAACAAGCT
CTTGAAGGTTTACTAACCCTCTGTGCAAGAACAATCTATT
CCTGCTGCTTTAGAAGGTAAAGACCTTCTGTTTCTAGCCAGACTGGTCT
GGTAAAACCTGCTGCATTCTTATTACCGACATTACATAATTTGGCAGGTCAA
GATACTTTCGTTCCGTTCAAAGAACGTATGAAAGCAGTGACTCAGCCTAG
CATTTTAGTTTTATGCCCAACACGTGAATTGGCTCAACAAGTTAGCCAAGA
CGCAATTGCTTTTGTACGTCACATGAAAGGTGTTTCGATTGCTGCAATTAT
GGTGGTATGCCTTTTGGTAAACAAATTCAACAGCTTAAAGGTGCTCAAGT
TGTAAGTTCGACTCCAGGTGCTTTATTAGACTTAGTTAACCGTCGTCATT
AAAGCTAGACAAAGTTGATGCTTTAATCGTGGACGAAGCTGACCGTATGTT
AGACCTTGGTTTCTCTGAAGACTTAGAAGCGATCAGTGACTTGGCTGCGA
ATCGTGGTCAAACGTTAATGTTCTCTGCAACTTTTCTGACCGTATTATTC
GTCTTGCTGAGCGCATGATGAATGAACCTGAGCGTATTGCAATTGAACT
GGTCACAGCACCAATACTGACATCACCCAACTTTACTGACAGATGG
TTTCGAACACAAGAAGAACTTCTGACTACTGGTTGGCTGATGAAACGCT
TGACCAAGCTGTTGTATTTCGCGAGCACACAAGAAGATACAGATATGTTGG
CTGAAGAACTTGCAGAAGCAGGTCACTCAGTTGTTGCACTTCACGGTGCA
ATGCCGAAACAGTTCGTAACCGTCGTTTACGTAGTATCCGTGAAGGCCG
TGCGAAAATCTTAGTTGCTACTGACGTTGCTGCACGTGGTCTTGACGTT
CAACAATTTACACGTTATTAACCTTCGGCTTACCAATGAAGCATGAAGATT
ATGTACACCGTATTGGTCGTACAGGTCGTGCTGGTCGTACAGGTCAAGCG
ATTACTTTAGCGACTTATCGTGAACGCGGTAATCCGTGCGCTTGAAGA
CTACTTAGAAGCTCGTTTAAATGTTTCTGAAATTGAAGGTTTGAAGCCATC
TCCACCTCCAGCTCGTCCACGTCGTGAAGGTGGTCGTGGTGGTAACGGT
GGCCGTGATGGTCGTGTTTTCGAGGGTGAAGGTAACCTCAAACGTCGTGA
AGGTGGTCGTGATGATCGTCCACGTCGTAGCTTCGATGATAAGCCTCGTG
GTGAGCGTCCTGCTTTTCGGTGGCGAAGATCGTCCACGTCGCGAGTTCAA
CTCAGATCGTCCACGTCGTGAAGGTGGTTTTGAAGACCGTCCACGTCGTA
GCTTCGGTGGTGAAGATCGTCCACGTCGCGAGTTCAACTCAGACCGTCCA

deaD

ATP-
depen
dent
RNA
helicas
e
DeaD

A2, A3,
A14

L

Belongs to the DEAD box helicase family

dhbF

Dimod
ular
nonrib
osoma
l
peptid
e
syntha
se

L9

/

amino acid adenylation domain-
containing protein

CGTCGTGAAGGTGGTTTTGAAGACCGTCCAAAACGTAGCTTCGGTGGTGA
AGATCGTCCGCGTCGCGAGTTCAACTCAGATCGTCCACGCCGTGAAGGT
GGTTTTAACGATAAGCCTCGTTTTGATTCAAATGATGACAATCGTGGTAAC
CGCGTAGACTATAAACCACGTCGTGAAAACGGTTTTCGGTGACCGTCCACA
ACGTAGCTTTGGTGGTGAAGATCGTCCACGTCGTGAAGGCGGTTTTGGTG
ATCGTCCAAAACGTAGCTTTGGTGGTGAAGAACGTCCACGTCGCGCAGTT
CGCGAAGAACAATTAATCAAGAATCTCGTGGTGAGCGCCGTCGTAAATT
CGACCGTTAA
ATGTTCTGGGAACGTCTGGGTGGTATGGATGAGTTTAAAGAACTACTCCT
ACCAATCCAGACTCCGGCCACTTTGGTTGGACAAGATTTGACGTTAGGAA
GTTTTCCGGTTCGCCAGCAAGAAGGACAGTTGGATATTACGCTGGAAATG
GGCGGTGAATATCAAGGTGAACTCGTTGGTGTATTGAAATACAACACTGA
TCTATTTTCTGCGCAGTCTGCCAAGAATATGGTGAACACTGCTGCAAGCTGT
GCTATCAGAGATGGTGGCTCACCTGAACGTAAGATCGTTGAGTTAGATA
TCGCACCTGATTATAAAGATAGTATCCAGTTCGAGGCCCTCCGTGGAAAG
GCGACTGATTATGCTCAGCATGATCTATTCGCAATGATTCTCAAGCAGATT
GATGAGCGTGGTGACAATCATGCTTTGACCTCAAACGATCATGCGGTTAG
CTACTGTGAGTTAGGTCAACATATTGCTGGTATTGCTGAATATTTGTGTGC
ACACGGCATTACACGCGGTGATCGTGTAGGTCTGATGTTAGACCGCAGG
GCGTTATTGCCAGCAGCAATCCTTGGGATCTGGGCGGCAGGTGCAGCTT
ATGTGCCACTGGATCCTAATTTCCCAACTGAACGTCTACAGAACATTATTG
AAGACGCTGAACCTAAAGTCATCTTGACTCAAACCGAGTTAATGGATGGT
CTTAACGTTTCTGTGCCTCGCTTAGATATCAATCAAGCTGGTGTGTGGCA
CTTGAACAAGTGCGCGAGACGCTTGCCTTTGGTGACATAGCTTATGTGAT
GTACACATCAGGCTCGACAGGTAAGCCTAAAGGTGTCCGATTGGGCATC
CTTCAATTATCAATTTCTTGCTATCTATGAATGATAGATTGCAAGTAACGAC
TGAAACTCAGTTGTTAGCAATCACACGATGCATTTGATATTTCAATTCTT
GAATTAATCATTCCGCTGATGTACGGCGGTGTGGTGCATGTTTGCCACG
TGAGGTATCACAAGATGGCAATCAATTGGTGGATTATCTAAATGCTAAGTC
TATCAATATTCTGCAGGCAACACCGGCTTCTTGAAAATGCTCTTAGACAG
CGAGTGGAGCGGCAATGCGGGATTGACAGCACTCTGTGGAGGCGAAGC
GTTAGATACTATTCTGGCAGAAAACTGCTGGGTAAGTTCGGTTGCTTATG
GAATGTCTATGGTCCGACCGAAACTACAGTGTGGTCTAGTGCGGCTCGTA

TCACTGATGCAAAGTATATTGATTTAGGTGAACCACTGGCGAACACTCAGC
TTTATGTCCTTGATGAACAGCAGCGTTTAGTTCCACCAGGTGTTATGGGTG
AGCTTTGGATTGGTGGTGGTGGTCTGGCAGTTGATTACTGGCATCGACCA
GAACTGACAGATGCCCAATTTAGAACATTGCCATCTCTACCAAATGCCGGT
CGACTCTACCGTACAGGGGATAAAGTCTGCCTGCGTACAGATGGACGCTT
GACTCATCACGGACGTCTGGATTTTCAGGTCAAGATTTCGTGGGTTCCGTA
TTGAACTAGGTGAAATTGAGAATGTTCTCAAACAAATCGACGGTATTACAG
ACGCAGTAGTGCTAGTTAAACTACAGCTGACAACGACCAGAACTGGTG
GCGTATGTA ACTGGACAAGAGCTTGATATTGCTGGTTTGAAAAAGACCCTA
CAAATCCATCTTCCAGCTTATATGGTGCCAAGTGCCTTTATCCGATTGGAT
GAGTTTCCAATGACGGCCAACAATAAGTTGGATCGTAAGGCATTTCTCTGA
GCCGATCTTTGAGCAGAGCAATGATTATGTTGCACCGCGTGATCCGATTG
AAATCGAACTGTGTACTACCTTTGAACAAATCTTGTGCGTCAAGCGAGTTG
GTATACATGATGATTTCTTTGAATTAGGTGGGCATTCACTATTAGCTGTGA
AGCTTGTCAACCATCTTAAAAAGGCTTTTGGTACCGAACTGTCAGTTGCTT
TGTTGGCACAGTACTCCACTGTGGAACGTCTGGGTGAGATCATTTCGTGAA
AATAAAGAAATTAACCATCCATCGTAATTGAGCTGCGCCGTGGAACCTAT
GAACAACCACTGTGGCTGTTCCATCCAATTGGTGGCAGTACCTTCTGTTAT
ATGGAACTTTCGCGTCATCTTAATCCGAACCGTACCCTACGTGCTATTAG
TCTCCTGGACTGATTGAGGCCGATGCTGCGGAAGTGGCGATTGAAGAAAT
GGCGACGCTCTATATTGCTGAAATGCAAAGATGCAGCCTCAAGGACCTT
ATTTCTCGGTGGCTGGTGTTTCGGTGGTGGCAGTGCATACGAGATCAGT
CATCAGTTGCGTCAAATGGGACAGCAGGTACAGGGATTGTCATGATCGA
TACCCGTGCTCCAATTCCAGAGAATGTGCCTGAAGATGCCGACGATGCCA
TGCTTTTGTCTTGGTTCGCCCCTGATTTGGCTGCACCTTATGGCAAAAAGT
TGACGATCCCAGCGCAATATTTGCGTGAGTTGAGTCCAGATCAAATGTTT
GATCATGTACTTAAAGAAGCTAAAGCCATTAATGTTCTGCCTCTCGATGCA
GACCCATCAGATTTCCGACTTTACTTTGATACTTATTTGGCGAATGGCATC
GCACTCCAGACCTATTTCCCTGAACCTGAAGATTTTCCAATCCTTTTGGTC
AAGGCCAAAGATGAGCAGGAAGACTTTGGAGAAAGTCTAGGTTGGGATCA
ACTAGTCAAAGACACGTTGACACAAGTCGATTTGCCTGGTGACCACTCAT
CAATTATGTACGCTGAGAATGTCGTGGCTGTAGCCCAGACAATTGATCAG
ATGTACCCAATTCCTACATAA

eptA_1

Phosphatase
EptA

L9

S

Domain of unknown function (DUF1705)

ATGTTTAATCTCATTATAGCCATTTGGCTAGGTGCAATTTTAAATATTGGTT
TTTATCATCAAGTCCATACTTTAACTCCATATTTTGGAGTAAAGGCTATTTT
GTTTTAGCTGCTACTCTCATAATTCTTGTTGCTACCTATTATGCGGTTTTA
CAAATCTTAAATTGGAAATGGACTGCCAAAATCTTTGCAATTTTATTGATAT
TTATTGGTGGTTTTAGCTCTTATTTTGTAAACACATTAGGTGTCATTATTC
ACCCGACCAAATTCAAATATGGTGCAGACCGATGTTTCGGAAGTTACCG
ATCTAATCTCTTACGCTTTGTTTTATGGACAATTTTTTTTGTATTTTACCC
ATTTTTTTAATTACTCAAGTTAAATTTAAACAAGAAAAAGTATCACGGTTGTT
ATTGAAGAAAGTATTCTCACTGGTAGCTTCATTGGCCGTGGTCCGGTGT
ACTTTTTACCTACTATGTCGATTTTGTGCAATATTTTCGTGAGCATCGTGAT
TTAAAAGGGATGATTTACCGCAAATAGTATTTTCATCGCTTATGTCTTACT
ACCATAAGAAGGCTCCGAAGAAAATCTGCCTCTTGTGATATATGGACAA
GATGCTCATCAAGTTCAGCGCGTACAAAAGAATCTCCCTAAGTTAATGATA
CTTGTTGTGGGTGAAACGGCACGTGCCGAAAGTTTCTCTCTAAATGGGTA
TGCAAAAATACGAATCCGGAGCTTTCTAAACAAGATATTTTCAACTTCTCT
CAAGTGAGCTCATGCGGTACGGCAACAGCGGTTTCTGTGCCATGTATGTT
CTCGGGTATGCCACGTGTAGATTATGATGAGCAATTAGCCAGTCACCGCG
AAGGTTTACTAGATATTGCAAAACGTGCGGGTTACCAAGTAACTTGGATTG
ATAATAACTCGGGTTGTAAAGGTGCATGTGATCGCGTTGAGCAATACCAG
ATTCCAGAAAATTAAGAAAAAATGGTGTAAAGATGGCGAATGTTATGAT
GACATTCTCATTGACAGCTTAAAGCAGTATTTGTCTACTATTGCCAAAGAT
GATGATCGTCCACGTTTGATTGTTTTGCATCAGGTGGGTAGTCATGGACC
TGCATATTACAAGCGTGCGCCTGAAGCATATCAACCCTTTAAACCGACTTG
TGATACGAATGCGATACAGGGCTGTTTCGCAAACCGAATTGCTAAATAGTTA
TGATAATACAATCGTATATACAGACCATGTATTAAGCCAAATGATTAATACT
CTAAAAGAAATATCAAATATCAGACAGGTTTATGGTATTTATCTGATCATG
GCGAATCAACCGGAGAACATGGTTTATATTTACATGGTTTACCTTATGCAA
TCGCACCGAGCCAACAAACACATGTACCAATGATTATGTGGTTCTCTGAAA
GTTGGAAACAACATAATCTTGCTCAAGTGAATTGTTTAAAGCCAACAACTA
AACAAAAGTTAAGTCAGGATAATTTATTCCCAAGTTTGTAAAGTTTGTGGA
TGTAACAACCTCAGGTCATCAACCCTCAACTGGACATGTTGCACTCTTGTGC
CCATGTAAACTAA

<i>ermC'</i>	rRNA adenine N-6-methyltransferase	A1, A3, A14	J	Specifically dimethylates two adjacent adenosines (A1518 and A1519) in the loop of a conserved hairpin near the 3'-end of 16S rRNA in the 30S particle. May play a critical role in biogenesis of 30S subunits	GTGGAATACGGGTTTGCTAAAAGATTATTAATACAAAACGCTCATTGGCA TTACTTTTAATGGCAGAAGTTGATATTTCTATATTAAGTATGGTTCCAAGAG AATATTTTCATCCTAAACCTAAAGTGAATAGCTCACTTATCAGATTAATAG AAAAAATCAAGAATATCACACAAAGATAAACAGAAGTATAATTATTTGTT ATGAAATGGGTTAACAAAGAATACAAGAAAATATTTACAAAAATCAATTTA ACAATTCCTTAAAACATGCAGGAATTGACGATTTAAACAATATTAGCTTTGA ACAATTCCTTATCTCTTTTCAATAGCTATAAATTATTTAATAAGTAA ATGCCACCCTTAGCAAGTTTGCCTCATTCTGAGTTGGTTATTCATGTAGGT TCATTGTCTAAAGTTTTTGCACCCGGTATTCGGTTAGGTTACATCGTGGCA GCTCCAACATCATTCAAACCATTACTGAAGATATTTTATTGATTGATCGGC AGGGCAATAATATTACGGAACCTCGCTTTAGCCGATTTAATGCAACGAGGG GAAATCAAGCGACACATCCGTAAAATGAAAAGATTTATCAGTTACGTCGA GACCATGCGTTAGCTGAGTTTCAAAGAGTTTTTGTCTGAGAGTATTCAGATA CAACCGCCCGCGGGTGGAAATGGCTTTGTGGTAAAATTTCAGAAATCATT TACTCAAAATCAGCTCATAAAGCTAAAAGAACTCAACATTGATACAGAGCA GAAGTTTAAACAGAAAGAATATCCTAATTGTAGTATCCGCTTTGGTTTTGC AGCTTTATCAGAGAAAGAAATCACCTCTCTAATTGAGACCTTAAATGAAGC TCTTAATAAAAAGAACAGCGGACTTATAA
<i>gabR_3</i>	HTH-type transcriptional regulatory protein GabR	L6, L8, L9	K	helix_turn_helix gluconate operon transcriptional repressor	ATGCCCTGGCTGGAAGCGGCGGGGGTGGCGCTGGACAGTCAGGGGCGT ATTAAGCCGGAGTGGAGAGTCGTTACCGTTATCAGACCAGTCAGGAGAA GATCTTTGCCGGCGGCGATGCCGTGCGCGGCGCAGACCTTGTAGTGACG GCGATGGCGGAAGGCCGTCACGCGGCGCAGGGGATCCTGGACTATCTG GCGCGGAAAACGACGCCGCTTCACTGA
<i>gltD_2</i>	Glutamate synthase [NADPH] small chain	A1, A3, A14	C	Glutamate synthase	TTGCCACATTTGGCATTGGTGCATGGTCTACCGATCCGGATTATAACTTT ACTAAAACCCGCCGTGACGGTTTTTCAGGGAGCACCCGGTTTATTAACCG GCACAGTAAACTAAACCTGATTATTGATGGTGGAACTCTGGCTCGAGGCG GCTCAGGTGGTGGAGCAACACCAAGCGGTATTTATACAGGATTATCGTAT GGAGTTCAGGGTATTCCCGGTGGAGCTGGAGCACCTTTTGGTCGGGTTAT GACCGGACAACCTATTAACGATTCACAAGACTGGCGTTGGTACTTAAA TGGTGACTTTTATGGTTGTCAAAGTAACCGATGCCGAAGCTTCGGTACCCG
<i>group_10</i>	hypothetical protein	A1	/	hypothetical protein	

<i>group_1024</i>	hypothetical protein	L7, L10	/	hypothetical protein	<p>GTAAGGTTACCGAACCCAAAATGATCGATATGGATCTCCATTGTCTGGTG ATGGTGGAGGTTGGGGCCAGCGCGGTACCAAGTCCACCAATGATGGAAC ATGGAACGGCAATACCATGGCACAACCTGAAGGCCAGCCGGGGCCGGGT GGACCTGCAATTGTTGGGGTGGCACCCTTACAACCTCAATTGATCAATGG AGGGAAAATTCTACAAACACTTTAA ATGAAAAAGTTGGTTGCTGCACTTGTCTGTAATTGTTATTTAACTGGATGC GTTTACGACCCTGTTAATTACGACAAAATACACGACCAAGAGTTCCAAGAC CACCTCAGGCCAAAACGGTTAG ATGAGTACTCTTCAAGAGCGAATGTCTTTAGCTATAAAGCACTATGAATCT GTGACAGGTAAAAGATTCAAAAATACTGAGCTGGCTAGATTTGCAGGCGT GAGCAGGGCTAATGTTGGGTTGTGGGTAAATGGCCCAACCCAAGAGTTG GAGGGGTCAAATTTAGTTAAAGCGGCTGAGTTTTTAGGAGTTTCTAAAGAT TGGCTAGCTGGACAAGGTAATAAAATGATTGCCACACAGCTGGATGGTGG TGGTGACAATTAATGTTCTTGATATTGAAGCCTTTAAGCAGAAGTACAA CATTCCAGATAGTGAAGATGCTGTTAAGTTTGTTCAAACATCAGATAAACC ATCCCTATTCAAAAAGATACGTTCCCGTCAAAGCTTATTCAAAGATGGG TATGGATGGGTATTTACAGATATGGGATACGATGGAAATGCTGGAGATG GGTATGTTCCAACCTCACTCAGCAGGACCAAGAGCCTATGGTATTAAGGC ACTGGCGACTCAATGTTTCCAGCTATCCGTAATGGATGGTATGTGGTTTGT GATCCAGATGCGGAACTCGTGCCGAATGAGTTTGTTCAGGTATGCTTGAA GGATGGAAGATGCACAATTAAGAATTTGTTGGCATAAATGGCGGGGTTT TAAGCTTGCTTTCTGTGAATGGTGGTGAGCGATTTTTCTTTGAAATGGATG AGGTAGAAAGCATTACAGCTATTACTGACATCGTACCACCAAGTCAGCAC AGACAAGAACATCCTTATTCGCATTAA ATGTTTCCAGCAATTCGTAATGGCTGGTATGTTGTATGCGACCCTGATGCA GAGCTTGTGCCGAATGAGTTTGTTCAGGTGTGCTTGAAGGATGGAAGATG CACAATTAAGAATTTGTTCGGCATCAATGGTGGGGTTTTAAGTTTGCTTTC TGTGAATGGTGGTGAGCGATTTTTCTTTGAAATGGACGAGGTTGAAAGTAT TACCGCTATTACAGATATCGTGCCGCCAAGTCAGCACAGACAAGAACATC CTTATTCGCATTAA</p>
<i>group_108</i>	hypothetical protein	A2, A3, A14	K	Peptidase S24-like	<p>ATGTCATCAGTCAGCATTGCTGAATACCGCAAGTTATTTCCATAAAGAAA AATAAAAAGCGCGGTTTCCAGCAAAGCAAGTTGCCAGACAACCAAGTGTGGG TGAAGTGGTTCTGGCAACGCATTTAAGAGCATGCAAGATTGGTTTTGAACA</p>
<i>group_109</i>	hypothetical protein	L6, L7, L8, L9	S	Peptidase S24-like	<p>ATGTCATCAGTCAGCATTGCTGAATACCGCAAGTTATTTCCATAAAGAAA AATAAAAAGCGCGGTTTCCAGCAAAGCAAGTTGCCAGACAACCAAGTGTGGG TGAAGTGGTTCTGGCAACGCATTTAAGAGCATGCAAGATTGGTTTTGAACA</p>
<i>group_111</i>	hypothetical protein	A2, A3, A14	S	Protein conserved in bacteria	<p>ATGTCATCAGTCAGCATTGCTGAATACCGCAAGTTATTTCCATAAAGAAA AATAAAAAGCGCGGTTTCCAGCAAAGCAAGTTGCCAGACAACCAAGTGTGGG TGAAGTGGTTCTGGCAACGCATTTAAGAGCATGCAAGATTGGTTTTGAACA</p>

group_
115

hypothetical protein

A2, A3, A14

S

Domain of unknown function (DUF4055)

GGAATATAAGTTCCATCCTGAACGCAAATGGAGAGCAGATTTTTTAATAAAA
GGGTTCAAAGATTTTGATTGAGGTAGAAGGCGGGATCTGGAGCGGAGGC
CGTCACACAAGAGGTAAGGGCTATTTAGGGGATATGGAGAAATACAACCTC
CGCAGCAATGATGGGTTTTACAGTTTTACGGTTCAGCACAGAGCAAGTGA
AAGCAGGCGTGGCGATTAACAAATTGAGCAATTGGTGGGATGA
ATGAGTGATGTAACTTTTCAACATGCTGAATATGTTAAGAACTTGCCATACT
GGCAAAAACCTTGATGATGTTTGTGAAGGTGAAGATGCAGTTAAGGCTAAA
GGTGAAAAATATTTGCCGATGCCAAATGCACATGATAAATCACCTGCAAAT
AAAAGCGCTTATGAGGCTTATCTTACCCGTGCAGTCTTTTTATGAAGTAACA
GGGACTACATTAATAGTTTAGTTGGTGCAGCTTTTGAACCGATCCAAGT
TTAAATTTCTCCGGAACCTTGCTCATTTAGAACGTAATGCAAATGGTGCT
GATTTAAGTACTTATCAATTGGCTCAAAATGGAATTCGCCATTTATTGAAGC
ATTATCGTTGTGCTTTATATGTAGATTATCCTGATGTGCCGCCAGCTCGTA
ATCTAGCGGAATTTAAAGCACAAAAGCCTATCCGATGATTCATTTACTAA
ATGCCCTTGATGTAGTGAATTGGGATTCAGTAATGATCGATAACCAGAAAA
AGCTTTGCTTAGTGGTTATACGTGAATTTAAGTCTGAGCGCGGTGCTGAT
GGATTTAGTAAAACCGAACAAGAGCAATATCGTGTACTTTCGTTTAGAGCAA
GAGGGAAATGGGGAATATATTTATTCCGTTTCAGGTGTACACAAAGGGTGA
AAAGGGTAACTGGGTTGGCGGAGAGAAGAAGTTTCCAACAGATTACAACG
GGAATTTCTGGACCTATATACCTTTTACATTTGTAGGTGCAATTGATAATTC
AGAAGAGATTA AAAAGCCACCATTACTTCCTTTGGCTAATCTCAATTTAGC
CCATTACAGAGACAGTGCGGACTTTCAAGAGTCCGTTTTTTATATGGGGC
AACCTCAATATTATGCGAAGGGTGTTAATTGGGAGTGGTATGACCAAGCC
AAGAAACGTGGCATCTACATTAGAGCGAAAGTACTTTTTGCCTTTACCTGAA
AATGGTGGTTTAGGAATTGTACAAGCCGACCCTAATACTCTTGCCCGGGA
AGCGATGAAAGATAAGTGGGAAAAAATGAAGGAGATGGGGGCGCGTTTA
ATTGAGAAGGGCTCGGGAAGTAAAAGACCGCTACCGAAGCGAATAGTG
ATGACGCCGTTTCAGCATTACAGTTCTTTTCGCTCTGTGTCGTTAATATGAATG
AAGCCTTGTCAGCAGCATTACGATGGGCTGCTAAGTTTTAA
ATGTTCGAAATCAGTCAAGAATTTAACAACAGGGTTATTTAGCTGAGTTA
GCTCGACAGTTATTTGAAGCAGCTCTACAAGGCCGATCTTCATTTAAATCA
TGGTGGGAATACAACCAAACAGGTATGTTCCCTAAACAAAAATATGAAGAA

group_
117

hypothetical protein

A2, A3, A14

S

Domain of unknown function (DUF4055)

<i>group_12</i>	IS6 family transposase IS26	A2, A3, A14	L	Transposase IS66 family	GAGCTTCAGAATGTTGAAGCAGAGCAAGATGGGACTTTAAATCAAAAGGT AGAGTGA ATGAACCCATTCAAAGGCCGGCATTTCAGCGTGACATCATTCTGTGGGC CGTACGCTGGTACTGCAAATACGGCATCAGTTACCGTGAGCTGCAGGAGA TGCTGGCTGAACGCGGAGTGAATGTCGATCACTCCACGATTTACCGCTGG GTTTCAGCGTTATGCGCCTGAAATGGAAAAACGGCTGCGCTGGTACTGGC GTAACCCTTCCGATCTTTGCCCGTGGCACATGGATGAAACCTACGTGAAG GTCAATGGCCGCTGGGCGTATCTGTACCGGGCCGTCGACAGCCGGGGC CGCACTGTCGATTTTTATCTCTCCTCCCGTCGTAACAGCAAAGCTGCATAC CGGTTTCTGGGTAAAATCCTCAACAACGTGAAGAAGTGGCAGATCCCGCG ATTCATCAACACGGATAAAGCGCCCGCCTATGGTCGCGCGCTTGCTCTGC TCAAACGCGAAGGCCGGTGCCCGTCTGACGTTGAACACCGACAGATTAA GTACCGGAACAACGTGATTGAATGCGATCATGGCAAACCTGAAACGGATAA TCGGCGCCACGCTGGGATTTAAATCCATGAAGACGGCTTACGCCACCATC AAAGGTATTGAGGTGATGCGTGCACTACGCAAAGGCCAGGCCTCAGCATT TTATTATGGTGATCCCCTGGGCGAAATGCGCCTGGTAAGCAGAGTTTTTG AAATGTAA ATGACTCATTTCAAAAAGCACCCCGACGGCTACAAGTCATTTTTAGGCCGT GATGATAAGGGCCTCTACTCTGTTCTGATTGGCTGGCAAGTGTACGCATC TAATGCTAATGGCTCAGTTCTTTACAAAGTTAAAGACGGATTTAAGACGCC TTTAAATGTGTATAAGTTTCAAACCGACTATCCAAAAGTTTGAATGAACTC ACACAAGAAATCGACTTTCAACGCAGAAAGCAGCTCGCAATAAAACTGCG TGAAACAAACATTCTACTTATGACCGCAAAGCATATAAGCAAAAACGCGG CTTCACCGGCTCTAGATGA
<i>group_1212</i>	hypothetical protein	A3	/	hypothetical protein	ATGAGTTTATTTTCGCTATGGTTTAGACTATGTTCAAATGGCGATTGAGCGT TTAATTGGTTTTGGGAAAAAAGAAGAGTTTAAAGGAAATTTGGCAATTTTAA GAAGGCAGAACCCTGATAGGATAAGGGTTCTGTGA ATGCCTGAGATACAAATTATTGCCAAGGATAATCATAAAACCTTGGTAACC ACAGAGGGGAACTTCTGCAAACCTTTCAGAAGCGTCAGTTGTGTTGGTGAA AGTGGCTGCCAGCGATGTATTGGTAGTAAATAGGGAGGGTACCAATGCAG TCATACGTCTGAAAAATGGCGAGACTATTGTTATTGAAGGCTTCTTTAGTG GAACAGCTGAGCCTAAAGATAATAGTCTGGTTTTCCAAGATGAAAATGGAC AGTTGATTTGGGCAAATTCAAAGATGCTGAAAATGATGCAGATGCTGATA
<i>group_1216</i>	hypothetical protein	A2, A3, A14	L	Transposase DDE domain	ATGAGTTTATTTTCGCTATGGTTTAGACTATGTTCAAATGGCGATTGAGCGT TTAATTGGTTTTGGGAAAAAAGAAGAGTTTAAAGGAAATTTGGCAATTTTAA GAAGGCAGAACCCTGATAGGATAAGGGTTCTGTGA ATGCCTGAGATACAAATTATTGCCAAGGATAATCATAAAACCTTGGTAACC ACAGAGGGGAACTTCTGCAAACCTTTCAGAAGCGTCAGTTGTGTTGGTGAA AGTGGCTGCCAGCGATGTATTGGTAGTAAATAGGGAGGGTACCAATGCAG TCATACGTCTGAAAAATGGCGAGACTATTGTTATTGAAGGCTTCTTTAGTG GAACAGCTGAGCCTAAAGATAATAGTCTGGTTTTCCAAGATGAAAATGGAC AGTTGATTTGGGCAAATTCAAAGATGCTGAAAATGATGCAGATGCTGATA
<i>group_1256</i>	hypothetical protein	L7, L9	S	Protein conserved in bacteria	ATGAGTTTATTTTCGCTATGGTTTAGACTATGTTCAAATGGCGATTGAGCGT TTAATTGGTTTTGGGAAAAAAGAAGAGTTTAAAGGAAATTTGGCAATTTTAA GAAGGCAGAACCCTGATAGGATAAGGGTTCTGTGA ATGCCTGAGATACAAATTATTGCCAAGGATAATCATAAAACCTTGGTAACC ACAGAGGGGAACTTCTGCAAACCTTTCAGAAGCGTCAGTTGTGTTGGTGAA AGTGGCTGCCAGCGATGTATTGGTAGTAAATAGGGAGGGTACCAATGCAG TCATACGTCTGAAAAATGGCGAGACTATTGTTATTGAAGGCTTCTTTAGTG GAACAGCTGAGCCTAAAGATAATAGTCTGGTTTTCCAAGATGAAAATGGAC AGTTGATTTGGGCAAATTCAAAGATGCTGAAAATGATGCAGATGCTGATA

GTGATGCGGACGCAGATGCTGACAGTGACGTGCGAACCACAAGCACTTTTA
GGTGAAGATTTACCAGCAGCTCTACCAGCAGAAGCTCCACAAGAGTTGGT
ATCTGATGTTATTTACCAGCCTATTTCTTCAATTGAACCTCTACTTTATCAT
GATGCAGGCGTGAACCCATGGTTATGGGCTGCTATTCCACTAGTTGCAGG
CGGTATTATCGCTGCAGCATCAAATCATGATAGTGATGATGATTCTTCAGC
ACCTGTTGATAACCACTCCACCAAGTACCGATGGCGTAACTTTCTCGGTTG
ACCCTGTTACTTCTGATAATGTGATTAATGCATCAGAAGCATCCGGTAATG
TCACTATTACAGGTGTTTTGAAAAACATTCCAGCAGATGCGGCCAACACG
GCCGTTACAGTTGTTAATGGGGTAACTTATAATGCGACTGTAGATAAAA
GCAGCAGGTACATGGACAGTAAGTGTACCGGGCAGTGGCTTAGTTGCGG
ATGCAGATAAGACGATTGATGCCAAAGTAACGTTTACAGATGCAGCAGGT
AACAGCAGCAGTGTTAACGATACACAACTTATACACTTGATAACCACTGCA
CCAAATACACCAGTGATTGACCCAGTTAACGGGACAGACCCGATTACAGG
TACAGCAGAACCTGGTTCAACAGTGACTGTGACTTATCCGGACGGTAGTA
CGAAAACCTGTTGTTGCAGGACCAGATGGTACTTGGACTGTGCCAAACCCA
GGTTTGAATGATGGCGATGAAGTTACAGCTGTAGCAACGGACCCTGCTGG
CAATACATCAGGACCAGGAACAGCTATTGTTGATGCAGTTGGTCCAAATA
CGGATGGTGTTAATTTTGCTGTTGATTCCGTAACAGCTGATAATGTGATTA
ATGCATCAGAAGCGGCTGGCAATGTCACGATTACAGGTATATTGAAAAAC
GTTCCCTCAGATGCGGCTGCTACAGCCGTTACAATTGTTAATGGGGTA
ACTTATACTGCAACTGTAGACAGTGCAGCAGGTACATGGACAGTAAGTGT
ACCAGGTAGCGGTTTGGTTGCTGATACTGATAAAACGATTGATGCTAAAGT
AACCTTCACTGATGCAGCAGGTAATAGCAGCACAGTTAACGATACACAAA
CTTATACACTGGATACAGCTGCTCCTTCAGCGCCAGTAATTGACCCTGTTA
ACGGGACAGACCCGATTACAGGTACTGCAGAACCTGGTTCAACCGTAACA
GTAACCTATCCTAATGGTGACACAGCAACAGTTGTAGCAGGACCTGATGG
CAGCTGGTCAGTACCAAACCCAGTTTTAAATGATGGTGACGAAGTTGAGG
CAATTGCTACAGATCCAGCAGGCAACCCATCTTTACCAGGTACAGCTATT
GTTGATGCAGTTGGTCCAAATACCGATGGCGTTAACTTCACGGTTGATTCA
GTAACAGCTGACAATGTGATTAATGCATCAGAAGCGTCAGGCAACGTTAC
TGTTACAGGCGTATTGAAAAATGTTCCGGCAGATGCAGCAAATACAGTGG
TCACTGTTGTGATCAATGGCCAGACGTATACTGCAACTGTAGATAGCACA
GCGGGCACATGGACAGTAAGCGTACCAGGTAGTGACTTGACTGCGGATG

<i>group_1296</i>	hypothetical protein	A1	N	domain, Protein	CAGATAAGACAATTGATGCCAAAGTAACGTTTACAGATGCAGCAGGTAATA GCAGCAGTGTTAACGATACACAACTTATACACTTGATACCACTGCACCTG ATGCACCAGTAATTAACCCGGTTAACGGGACAGATCCAATTACAGGTA GCAGAGCCTGGTTCAACCGTAACAGTAACCTATCCTAATGGTGACACAGC AACAGTTGTAGCAGGACCTGATGGCAGCTGGTCAGTATGCACCCGGTGCTT GA ATGCGTGCTGAAGGTGCATTTAACTTAGATATTCTTGGTTATAGCTGGGAT ACATCAAAAGGTGAAAACCCTGACCTTACTTTACTTGGTTCAAGTGCCAAC TGGAAAAACATGCTACTAGCAACAAAATGACAGCAGGCACATTGCTTGAT CTATCTGGCACAACAACAACACTGGTTAA
<i>group_1299</i>	hypothetical protein	A1, A2	/	hypothetical protein	ATGACTGTATATTTTCCATCATTATCAGACTTTGCAGTTCCAATGATTTTT CAGATGAATCTTTAACGGTGACAGTTGAATTAGCCTCTGTCAAACCTGTAA TAGTTTTACCGTTAGAAGAAATAGCATGTGTTAAAGTAGGAGCTGTAATAT CAGAAGCGATAAGCTCTTGTTTAACTGTACTATTACCGGCAGTATCTTGA ATGAACCCATTCAAAGGCCGGCATTTCAGCGTGACATCATTCTGTGGGC CGTACGCTGGTACTGCAAATACGGCATCAGTTACCGTGAGCTGCAGGAGA TGCTGGCTGAACGCGGAGTGAATGTCGATCACTCCACGATTTACCGCTGG GTTTCAGCGTTATGCGCCTGAAATGAAAAACGGCTGCGCTGGTACTGGC GTAACCCTTCCGATCTTTGCCCGTGGCACATGGATGAAACCTACGTGAAG GTCAATGGCCGCTGGGCGTATCTGTACCGGGCCGTCGACAGCCGGGGC CGCACTGTCGATTTTTATCTCTCCTCCCGTCGTAACAGCAAAGCTGCATAC CGGTTTCTGGGTAAAATCCTCAACAACGTGAAGAAGTGGCAGATCCCGCG ATTCATCAACACGGATAAAGCGCCCGCCTATGGTCGCGCGCTTGCTCTGC TCAAACGCGAAGGCCGGTGCCCGTCTGACGTTGAACACCGACAGATTAA GTACCGGAACAACGTGATTGAATGCGATCATGGCAAACCTGAAACGGATAA TCAAACGCCACGCTGGGATTTAAATCCATGAAGACGGCTTACGCCACCATC AAAGGTATTGAGGTGATGCGTGCACTACGCAAAGGCCAGGCCTCAGCATT TTATTATGGTGATCCCCTGGGCGAAATGCGCCTGGTAAGCAGAGTTTTTG AAATGTAA
<i>group_13</i>	IS6 family transposase IS15	A2, A3, A14	L	Transposase and inactivated derivatives	GTGACATCGGCTGTACCTGTACCCACTACGGTTCCTACCAGTTCTTAACT GTTACTGTTGCATTGGCCTCAGTCACACCGGTAATTGCTTTACCTGCGGC ATCAATCGTTGCAGTCAGACTTGAAGGTGCTGTAATGTCTGGTGCAACTAT
<i>group_1300</i>	hypothetical protein	A1, A2	/	hypothetical protein	

<i>group_1302</i>	hypothetical protein	A1	S	Ribonuclease toxin, BrnT, of type II toxin-antitoxin system	TATTTTTGGTGCGGTAGCATTACCCGCTTTATCTGTCGCTGTTATACTTAA CTTTCGCCATTTAGATAA ATGGAACAGTATTTTGAATGGGATGAGGCTAAGAATCGAAAGAATCAGAA AAAGCATGATATCTCTTTGCAAAGCTGCAAGCCTTGTTTTTGAAGATCCATTA CGGATCTCAATCCAAGACAGATATACTAATGGTGAAGAACGTTGGCAAAC CATTGGACGAGTTAAAGGCGTACTAATGTTATTGGTGGCTCACACCATTTT TGATGAAGATGACTGTGAAATCATCCGAATCATTAGTGCAAGGCAAGTCA CTAAAGCGGAGCGAAACAAATATGAGCATGGTTAG ATGAGCATGGTTAGATACTCACACAAAGAGCTGAATGAGAAATTTGGTGA GAAGCAAGATGCTGAAATTCAACGCTTGCTTGCTAAAGGCACGGTTCCTG ATGATCAGCTAGATCTATCAGATATTCCTGAAATTACAGACTGGAGCAATG CTGTACGCCAAAATCAATTTTACCGCCCAGTAAAGCAACAAACTTCAATTC GGTTAGATGCTGATGTACTTGCTTGGTTCAAAGCACAAGGCAAAGGTTAT CAAACACGTATGAATGAAATTTTACGAGATGCTATGCTAAAAGAATTA AATCATCAATAA ATGCTAAATAAAAGTAAACTATTTTTAGCTTTAATTACTTTAGGGCCTCGA AAATTTTATTGGCAGCTGAAGGGCCTGTTACTACCTTAAATACAATAGTATT AACTGCACAATCTGATGAGTTAGGGTCTGAGCTGTTAGGCAAAAGTCTAA ATGTTAGTAATCAGTTTATTGATACTAGTAAACTTAAACAGCGTTCCGACAAC TCTAGGGGATGCCTTGGGTACGGAATTAGGAATTCATTCTAATCAATATGG AGGTGGAGCATCTACTCCTATTATTCGTGGGCAGGAAGGAAAACGTATTA AAGTTTTACAGAATAATGCTGACGTGTTGGATATGTCTAATATGTCTCCAG ATCACGCTGTCACAGTTGAACCAAGTTTAGCTAAGTCTATTGAAATTATAC GTGGTGCAAGTACCCTTTTGTATAGTTCAAATAGTGCGGCAGGGGTAGTC AATGTCATTGACTATAAAATCCAACCCAAATGCCTCAAGACGGATTAGAG GGTAATACTACATTACGTTTTAATACTGGCAGTAATGAGAAGTTGACCACA GCAGGTGTAACAGTTGGCTTAAGCCACATGTAGCTTTACGGGCAGAAGG GTTATATCGTAATGCTGGTAACTATAAACTCCTCATTACCAGTCTAGCTCT TATAATTCCTTTAGAAGATCTTGAAAACCAAACAGTATTTATAAAAAATTTAA ATATTTACCAGAAAGCTGGGCAGAGTCGAGAGTAGGTACGCTTGGATTAT CCTGGATAGACGATAATACTTATTTAGGAGTTTCTTATACACATCGTCATG ATGAATATGGTTTACCAGCCCATAGCCATTTATATGAAGGTTGTGGTGCCT CTGCAATCGGTATCGATTCACGAATTAGTGGATTAATAAATTATCTTTTATA
<i>group_1303</i>	hypothetical protein	A1	S	BrnA antitoxin of type II toxin-antitoxin system	
<i>group_1304</i>	putative TonB-dependent receptor	A1	P	Outer membrane receptor proteins mostly Fe transport	

group_
1305

hypothetical
protein

A1

/

hypothetical protein

TTATCCTCAATTGATGGATGAGCAAGATATTAATTACATAAATCCACGTCCT
GATTGTCATCAGCATAATCATATTCATGAACTAATTTCTCTCATAATGCTC
CTTATATTGATTTAAATACTCGACGTTATGATGTGCGTGGAGAATTCACTC
AACCATTTACCGGTATTGACAAAATTCGTACCTCATTGAGTTATATAGATTA
TTTTCACAATGAACTCGAGGGAGATAAGATTACTAACTTTTTTAAAAATACA
GGAAAAGTAGGCCGCATAGAGTTATCACATCAGCCATTAGGGGAATTAAC
CGGTATCTTGGGGCTACAATATTTGGAACAGGATAATAGTGCTTTATCACC
AGTCCATTCTCAAGAAGGCCATAACCACCTATCTAGATAATCAGCAGCTATT
AAATCGAAATGTCACTAAAACTTTAGTGTATTTGGCCTAGAAAAATATAAC
TGGAACGACTTTACTTTTGAAGTGGTCTCGTATAGAAAAGCAGAAAAGTT
AGTATGGATTATGATATAGAAAAATTAAGATTCAATGAAACCTTGGCCTA
ATAAATATAACAGTCCTTATGTAGAAAAAATAATAAAATCAGGGCTCAAAA
TTTTAAATCTATTTTAGAAGCTGTTCAACCTAACAAAGAAACAGCTTTTTCT
TATGCAGGTACAGTCCACTGGAGATTTGCTCCAAATTATATCTTATCATTAA
CAGGTACACATCAAGAAAGATTACCTAATGCACAAGAAATGTATACTCATG
GTATGCATCTCGCTACTAATTCTTTTGAATAGGTAACCGCTTTCTAAGAAA
AGAGAAATCCAATAATTTAGAAATCAGTCTGGCCTATAAGGATGATTTACT
TGATTATCAAATCAGTACTTACTATTATGATTTTGATAACTATATATATTTAC
AGACACTCAATGAAGTTTTAGGAACTACAAAAGTTAGAGATCAGCATACAC
TACGTATTAATCATTATAGCCAGTCTGCTGCTAATTTTTATGGATTAGAGG
GCAACATAGGGTATCAATTTAATTCTGTTTATCACGGCTCTCTTTTTGGAG
ATTATGTGAAAGGCCGTTTAAACAATTTACCTGATGCTGTAATTGCATACG
ATATCTGGAATAGAGAACCGACTTTAGCTCCCCAAAAGGATCGTTACACTC
CTAGATTACCTCCTGCTCGTTTGGGTACACGACTAAAAGCTGATTTTGATG
AATCGCTTAAAGGTGAAATAGAATATTATCGAGTCTTTAAACAAGATAATAT
TTCTAAATTTGAGCAGGTTACATCAGGTTATAATATGCTTAATATGACTTTA
GCATATAAAAATAAGCTATCTCACACAGAATATGATTTATTCTTTAAAGCGA
ATAATTTACTTGATCAGAAAGTCTATGCCCATGAAACATTTTTGCCATATAT
TCCCCAAATTGGAAGAAATTTAGTTTAGGCTTAAATCTTAATTTCTAG
ATGAAATTAGCAGAAGCTCTTACTTTCGAAGTGATCAACAAAAAAAATT
ATTTCACTAAAACAACGTATCAATGCAAATGTATTGGTACAAGATGGTGAT
CAACCATCTGAAGATCCAAATGAACTACTAAAACAAGTATTTTCTTTAATTC
AAGAATTTCAAAAATTAAGTTATGCAATTCATGAACTAATGCTTTAACAAA

<i>group_1306</i>	hypothetical protein	A1	/	hypothetical protein	<p>ATTAATGATGGCCGTTCTCTACTTGCTTTATTGACTTTGCGTGATGAATTT GTAGAACAGCACAAAACCCTAACTGCAGCAATAAGTAATACATCTCGTGAA TCTGATCGTTATAGTACGCGGGAAATAAAGTGGCATAAAGTTATCCAGTT TCATCATTACAAAACAAGCTGATGATATTAGTCTTAAATTACGTGATTTAA ACGTGTTAATTCAGTCGACTAATTGGAAAATTGATTTGATTGAAGCCTAA ATGTCTAATAATAAACCTTTGAAAAATAGTACTAAATTACTTGTTAACTTAG ACAAGTTTATATTTTTAGTCAATGCTGCCGATAGCCTAGAAGAAATTGAAAT TATTAGAGATTTATGCTGTGAATATTTTTACATTGTAAGAGGCCAAGCTAT TACATTGATATTTTTGACAATGCTTATTGGATAAAAATATTACGAAATAGCTT AA</p> <p>ATGGCTAAGAGCTTACAAGAATTATTGGCTAGTCGCTCTCCAGAGAGTCA AGCTCGTATTCAAAAATGGCTGATGAATTACTTTTAGAAAAGTCAGCTTCA CCTTATCCGAGAAGAGTTAGAAATTTCCAGAAAAGAACTAGCTGCAACCCT TGGAATTAACAGCCGTCATTATCAGCAATCGAAAATCGTGGTCATGATCT TAAAATTTCAACGATGAAAAAGTATGTTGAAGCAATGGGTGGAAAACCTTCG TATTGATGTAGAGCTTCCAACAGGAAAACACATAGGGTTCAATGTCTAA ATGAAAAGCCTTACAGTTCTTGAGCTTTCTAAGCTTTACAATATTAATCGAC AAACCATTTACAACAATATAAAGAAAGGTATTTTAAGTAAAAATTCTGACAA TAAGATTGATCTTGCAGAAGCGATCCGGGTATTTGGTGAGCCAGTACAAA AACAGAGTGTAAGAAAGAACTGTAAGAGTAGACAGTCCAACCTTCGGCAGAA GTTTTACTCCTTAGACAGCAGATAGACCTACTAAAAAATCAGCTAGATGAT GCAAAAAATAGAGAAAGCTTCTATCAAAATCAAATTGAAACTATGCAACGC CTATTGGAAGCTCCAAAAACCAATACAACACAATCTGATGATCCTGGATCA ACACAGGCAGAATCCACAGTAGAAGAAAAAGAGGACGTCATAGCTATGCC TCCTAAGGAGTCAGAGGTTATCGAGAATAAGCGTATTCCAGTCCCTGAAC ATATTGAGCCAGAACCAGAAAAAAGGGGCTTCTGGAGCCGTTTTTTCAAG CCCTATGGATGA</p> <p>ATGAGTGAATTAATCGTAAAGGATAATGCTCTAATTCAAGCAAGTTATACC TTAGATACAGTTGAACAAAGACTGATCCTATTAGCCATTGCTGAAGCTCGA GAAACTGGACATGGGATAACTGAAAATAGTCTTTTAGAAGTACATGCAAGT AGTTATATAAATACTTTTAATGTCGAGAAACATACCGCTTATACCGTACTCA GAGAGGCATCTAAAAGCTTATTTGATCGCTATGTCACATACCATGACATTA ATCCTAAGACAGGTAAGGATCGTAGCTTTCATTGTCGTTGGGTGACAAA</p>
<i>group_1307</i>	hypothetical protein	A1	K	Helix-turn-helix domain	<p>ATGGCTAAGAGCTTACAAGAATTATTGGCTAGTCGCTCTCCAGAGAGTCA AGCTCGTATTCAAAAATGGCTGATGAATTACTTTTAGAAAAGTCAGCTTCA CCTTATCCGAGAAGAGTTAGAAATTTCCAGAAAAGAACTAGCTGCAACCCT TGGAATTAACAGCCGTCATTATCAGCAATCGAAAATCGTGGTCATGATCT TAAAATTTCAACGATGAAAAAGTATGTTGAAGCAATGGGTGGAAAACCTTCG TATTGATGTAGAGCTTCCAACAGGAAAACACATAGGGTTCAATGTCTAA ATGAAAAGCCTTACAGTTCTTGAGCTTTCTAAGCTTTACAATATTAATCGAC AAACCATTTACAACAATATAAAGAAAGGTATTTTAAGTAAAAATTCTGACAA TAAGATTGATCTTGCAGAAGCGATCCGGGTATTTGGTGAGCCAGTACAAA AACAGAGTGTAAGAAAGAACTGTAAGAGTAGACAGTCCAACCTTCGGCAGAA GTTTTACTCCTTAGACAGCAGATAGACCTACTAAAAAATCAGCTAGATGAT GCAAAAAATAGAGAAAGCTTCTATCAAAATCAAATTGAAACTATGCAACGC CTATTGGAAGCTCCAAAAACCAATACAACACAATCTGATGATCCTGGATCA ACACAGGCAGAATCCACAGTAGAAGAAAAAGAGGACGTCATAGCTATGCC TCCTAAGGAGTCAGAGGTTATCGAGAATAAGCGTATTCCAGTCCCTGAAC ATATTGAGCCAGAACCAGAAAAAAGGGGCTTCTGGAGCCGTTTTTTCAAG CCCTATGGATGA</p> <p>ATGAGTGAATTAATCGTAAAGGATAATGCTCTAATTCAAGCAAGTTATACC TTAGATACAGTTGAACAAAGACTGATCCTATTAGCCATTGCTGAAGCTCGA GAAACTGGACATGGGATAACTGAAAATAGTCTTTTAGAAGTACATGCAAGT AGTTATATAAATACTTTTAATGTCGAGAAACATACCGCTTATACCGTACTCA GAGAGGCATCTAAAAGCTTATTTGATCGCTATGTCACATACCATGACATTA ATCCTAAGACAGGTAAGGATCGTAGCTTTCATTGTCGTTGGGTGACAAA</p>
<i>group_1309</i>	hypothetical protein	A1	/	hypothetical protein	<p>ATTAATGATGGCCGTTCTCTACTTGCTTTATTGACTTTGCGTGATGAATTT GTAGAACAGCACAAAACCCTAACTGCAGCAATAAGTAATACATCTCGTGAA TCTGATCGTTATAGTACGCGGGAAATAAAGTGGCATAAAGTTATCCAGTT TCATCATTACAAAACAAGCTGATGATATTAGTCTTAAATTACGTGATTTAA ACGTGTTAATTCAGTCGACTAATTGGAAAATTGATTTGATTGAAGCCTAA ATGTCTAATAATAAACCTTTGAAAAATAGTACTAAATTACTTGTTAACTTAG ACAAGTTTATATTTTTAGTCAATGCTGCCGATAGCCTAGAAGAAATTGAAAT TATTAGAGATTTATGCTGTGAATATTTTTACATTGTAAGAGGCCAAGCTAT TACATTGATATTTTTGACAATGCTTATTGGATAAAAATATTACGAAATAGCTT AA</p> <p>ATGGCTAAGAGCTTACAAGAATTATTGGCTAGTCGCTCTCCAGAGAGTCA AGCTCGTATTCAAAAATGGCTGATGAATTACTTTTAGAAAAGTCAGCTTCA CCTTATCCGAGAAGAGTTAGAAATTTCCAGAAAAGAACTAGCTGCAACCCT TGGAATTAACAGCCGTCATTATCAGCAATCGAAAATCGTGGTCATGATCT TAAAATTTCAACGATGAAAAAGTATGTTGAAGCAATGGGTGGAAAACCTTCG TATTGATGTAGAGCTTCCAACAGGAAAACACATAGGGTTCAATGTCTAA ATGAAAAGCCTTACAGTTCTTGAGCTTTCTAAGCTTTACAATATTAATCGAC AAACCATTTACAACAATATAAAGAAAGGTATTTTAAGTAAAAATTCTGACAA TAAGATTGATCTTGCAGAAGCGATCCGGGTATTTGGTGAGCCAGTACAAA AACAGAGTGTAAGAAAGAACTGTAAGAGTAGACAGTCCAACCTTCGGCAGAA GTTTTACTCCTTAGACAGCAGATAGACCTACTAAAAAATCAGCTAGATGAT GCAAAAAATAGAGAAAGCTTCTATCAAAATCAAATTGAAACTATGCAACGC CTATTGGAAGCTCCAAAAACCAATACAACACAATCTGATGATCCTGGATCA ACACAGGCAGAATCCACAGTAGAAGAAAAAGAGGACGTCATAGCTATGCC TCCTAAGGAGTCAGAGGTTATCGAGAATAAGCGTATTCCAGTCCCTGAAC ATATTGAGCCAGAACCAGAAAAAAGGGGCTTCTGGAGCCGTTTTTTCAAG CCCTATGGATGA</p> <p>ATGAGTGAATTAATCGTAAAGGATAATGCTCTAATTCAAGCAAGTTATACC TTAGATACAGTTGAACAAAGACTGATCCTATTAGCCATTGCTGAAGCTCGA GAAACTGGACATGGGATAACTGAAAATAGTCTTTTAGAAGTACATGCAAGT AGTTATATAAATACTTTTAATGTCGAGAAACATACCGCTTATACCGTACTCA GAGAGGCATCTAAAAGCTTATTTGATCGCTATGTCACATACCATGACATTA ATCCTAAGACAGGTAAGGATCGTAGCTTTCATTGTCGTTGGGTGACAAA</p>
<i>group_1310</i>	hypothetical protein	A1	L	Initiator Replication protein	<p>ATTAATGATGGCCGTTCTCTACTTGCTTTATTGACTTTGCGTGATGAATTT GTAGAACAGCACAAAACCCTAACTGCAGCAATAAGTAATACATCTCGTGAA TCTGATCGTTATAGTACGCGGGAAATAAAGTGGCATAAAGTTATCCAGTT TCATCATTACAAAACAAGCTGATGATATTAGTCTTAAATTACGTGATTTAA ACGTGTTAATTCAGTCGACTAATTGGAAAATTGATTTGATTGAAGCCTAA ATGTCTAATAATAAACCTTTGAAAAATAGTACTAAATTACTTGTTAACTTAG ACAAGTTTATATTTTTAGTCAATGCTGCCGATAGCCTAGAAGAAATTGAAAT TATTAGAGATTTATGCTGTGAATATTTTTACATTGTAAGAGGCCAAGCTAT TACATTGATATTTTTGACAATGCTTATTGGATAAAAATATTACGAAATAGCTT AA</p> <p>ATGGCTAAGAGCTTACAAGAATTATTGGCTAGTCGCTCTCCAGAGAGTCA AGCTCGTATTCAAAAATGGCTGATGAATTACTTTTAGAAAAGTCAGCTTCA CCTTATCCGAGAAGAGTTAGAAATTTCCAGAAAAGAACTAGCTGCAACCCT TGGAATTAACAGCCGTCATTATCAGCAATCGAAAATCGTGGTCATGATCT TAAAATTTCAACGATGAAAAAGTATGTTGAAGCAATGGGTGGAAAACCTTCG TATTGATGTAGAGCTTCCAACAGGAAAACACATAGGGTTCAATGTCTAA ATGAAAAGCCTTACAGTTCTTGAGCTTTCTAAGCTTTACAATATTAATCGAC AAACCATTTACAACAATATAAAGAAAGGTATTTTAAGTAAAAATTCTGACAA TAAGATTGATCTTGCAGAAGCGATCCGGGTATTTGGTGAGCCAGTACAAA AACAGAGTGTAAGAAAGAACTGTAAGAGTAGACAGTCCAACCTTCGGCAGAA GTTTTACTCCTTAGACAGCAGATAGACCTACTAAAAAATCAGCTAGATGAT GCAAAAAATAGAGAAAGCTTCTATCAAAATCAAATTGAAACTATGCAACGC CTATTGGAAGCTCCAAAAACCAATACAACACAATCTGATGATCCTGGATCA ACACAGGCAGAATCCACAGTAGAAGAAAAAGAGGACGTCATAGCTATGCC TCCTAAGGAGTCAGAGGTTATCGAGAATAAGCGTATTCCAGTCCCTGAAC ATATTGAGCCAGAACCAGAAAAAAGGGGCTTCTGGAGCCGTTTTTTCAAG CCCTATGGATGA</p> <p>ATGAGTGAATTAATCGTAAAGGATAATGCTCTAATTCAAGCAAGTTATACC TTAGATACAGTTGAACAAAGACTGATCCTATTAGCCATTGCTGAAGCTCGA GAAACTGGACATGGGATAACTGAAAATAGTCTTTTAGAAGTACATGCAAGT AGTTATATAAATACTTTTAATGTCGAGAAACATACCGCTTATACCGTACTCA GAGAGGCATCTAAAAGCTTATTTGATCGCTATGTCACATACCATGACATTA ATCCTAAGACAGGTAAGGATCGTAGCTTTCATTGTCGTTGGGTGACAAA</p>

group_
1311

hypoth
etical
protein

A1

U

MobA/MobL family

ATTGGGTATGAATCTCAATCAGGAATTATTTTCCTACGATTCACCCAAGAT
ATCGTTCCACTCATAACTCGACTTGAAGAAAATTTCACTAAATATGAATTGC
AGCAGGTTTCTAGGTTAAGTAGCTCATATGCTATTCGGCTATACGAGCTAT
TAATTC AATGGAGATCTGCCGGA AAAACGCCACTTTTTGATCTATCTATCT
TTAGACAACA ACTTGGTGTCAAACCTCATCAATACAAAACAATGAGTAACT
TTAAAACATATGTTTTAGATTTTGCTCTTAAGCAGGTA AATGAGTTAACCGA
TATAACAGCTAAATATGAGCAACATAAAAAAGGGCGTTCTATTT CAGGTTT
TTCATTCACTTTCAAACAGAAAAAATGAGTAATCTGCCAATAAAAAATAAG
CGTGACCCAGACACTATAGATATTTTCTCAAAAATGACAGATGCTCAACGC
CATCTGTTTTCCACAAA CTGTCAGAACTTCCTGAAATGGGAAAAGTATTCT
CATGGTACAGAAAGCTATCCGCAATTCGCTGTACGCATTGCAGAAATGCT
TCAAAACCCAGAAAAGTTTAAAGAACTCTATCCTTATCTCCAAAAGTAGG
CTTCAAAGCTGCATAG
ATGGCCAGCTACCACTTCAGCGTCAA AAGCAAAAATAAGGGCTATGCGAC
CGGCCATTATCTCTACATTTCCCGGTGATGCAGTATGAAGGCATCCGCA
AAACCAGCAATGAGACCATAGAACATATTGAACCTGGTCGATGTATGCCT
GCCTTTGTAAAAGATCCCATTTGATTTTTGGCAGGCAGCAGATACGTATGAA
CGAGCCAATGCCAAGGCCTACATGGAGTATGAAATTGCCCTACCGAATGA
GTTACAGT CAGAACAGCGTAAAACGCTCATAGAGACGTTTTT TAGACAAACA
TATAGTTCCACAGCAATATCCGCACTCATACGCCATACACAACGTAAATC
ACGTATCAGTGGCGAAGATCAGCCTCATTGTCATCTGATGTTCTCGCTAAA
AGCGGATGACGGCATTGAACGCAGCGCAGAACAGTACTTCAAACGCTATA
ACCCTCAAGATCCAGCCAAAGGCGGTGCCAAAAAATCCA ACTTCAAGAT
GGCCATGAAGACTATTCCACTTTTTT GATCTATATCCGAAAACAATGGGAA
AATCATTTGAATGATGCCCTTGACAGCACTGCCCCACCGTCACTACAC
ACTGGATGGCCAAGACATCACCATCAAAAATCAGGTTTCAGCAGACAGCT
ATGAAAAATATAACGAGATCCACAGCACGCTATACCTTCCAGAGCCAAAAC
TAGGCGTTGGACAACAAAATGCGACAGCGGAATATCTCGCCCAAATCCAA
AAAATCAGAGAACACAACCAGCAAGAACGAGA ACTCGAAGAACGCCAAAA
GCAATTAGATCAACAACATGAATATTTTTATAGTGCTGAAAACAACCCATTT
TCATCATTGGAAGTCATGTATTACCTGAATTACTTAATCAAGAATGCAACAA
ATAAGGACAAACTAAAAAATGATTTGAATGAAATTATCAAAAACCAAGGCC
CTATACTTAAACAAGAAGCTTGGTCCATCTACGAAATGAATGAAAACTAG

<i>group_1312</i>	hypothetical protein	A1	/	hypothetical protein	<p>AAGACTCAAGACGTAACAGTACCCCATGTTTTGAGTTTCGAGGTGCATTCA CTTCCTGAATACGTATTACCGATACAAAATACGCCTCCCAAACGCACTCAA TCAGAACCAGTACAACCTAAAGCTGATGATTATGATTACGGTGGCCCTGG TTTTTAA</p> <p>ATGCCTGTATGGCTGAAAAATCCGGTTGAAGGGGGTCAAGGGGGAAGCG CCCCATTGCGTACTAGTTTTCTCCGATGAGGAGAACTGTCAAGTGCGAAT TATTTAATATAA</p> <p>ATGGAAATGGAAAAAGCAAAAAGCGTGGTCGACCTGCTCAGTTGTTACA GATTGCTGAGTTACATGCATTTGTAGAGTTTTTAGAACAGCAAGAACAGCT TTCTGATCTTCAATCTCAAGTGTTAAAGGCTTTAAATAGTGTTGATTGTAAT TTTGAGGGGCTAACTCAGACAGATCAGGTTCTGGTCAAAGAGGCTTTAAA ACCTTATCGTGAACATCTGAAGTTGAAATTGCTGTTTGAAGAGTTGAATAA TCTGCCTCTTAAAACCGAGTATGAACAGAAGTTTTTATAGATCTGTATGAACT TTTTCAAAAGAATGCACTAGATCAGATGGAACTAAATTTTTAAAACCTTG GCGACACGGTATCTGAATTTTAAAGCGCAAAGCTTGAATACAGTGATTTA GAATTGATTTGAGTCAGCTACAGAAGAAGGATGCTGGGAAAAACGCAA AGCTGAAAATCAACGTAAATTTGAATTAGGTGGTGCTGTTCTTGTGCGCATT TAAGAACTCAATATTGATATTAGCAATGATACTCCTCAACAGATTACCAAT CGGATTGTGAATACGACAAAGTTTTATAATGAAGTACGTAAATCTTTAATTT TTAAAGATGTTAAGACATATGAAAATGAATTTTTAAAGCAAATAAACTTTTT ATTCAGGTGTTAGAAGGATTACATACCTGGCAAAGGGCGGAGAATTATT GTCAGTAATTGAAATTTAAAAGGCCTTAGAAAAAGGCGAAGAGTAG ATGGACAAAAAGATAGAAGAGAAAATAAAGATATTAATGAAAAAATTCAA AAATTGGATTTAGCTATTCAAGAACTAATAGAACTACAGACAATTTTTTTA TTAATGATAAGTACTATGGAATTGAATGTGCTTTTTATTTTAGACCATGCACT AAATATTTAAAATGCACTTTTTCAACAAGTATCCAGCATTTGCACCACCTCAC TTACGAGGAGTAGAAAATTATGATCGAATTATACATATTAGAGAAAATAATC CTGAAGATAAGCTAGAACAATTTCTCCTCTTAAAGACTAA</p> <p>ATGACCCCAAATGAAAATGGACTCTCGGATCACCTCAGTGAAACAGAGCA GATGAATGCGAAGATGAAAGATAGAACCAGCGACCAAATGGTGCTCTAG AAAGCCAACAAGCAAATGGACAATGTGCCAAGAGTCTGAACCAGCAGGT AAGACTGACTGA</p>
<i>group_1313</i>	hypothetical protein	A1	/	hypothetical protein	<p>AAGACTCAAGACGTAACAGTACCCCATGTTTTGAGTTTCGAGGTGCATTCA CTTCCTGAATACGTATTACCGATACAAAATACGCCTCCCAAACGCACTCAA TCAGAACCAGTACAACCTAAAGCTGATGATTATGATTACGGTGGCCCTGG TTTTTAA</p> <p>ATGCCTGTATGGCTGAAAAATCCGGTTGAAGGGGGTCAAGGGGGAAGCG CCCCATTGCGTACTAGTTTTCTCCGATGAGGAGAACTGTCAAGTGCGAAT TATTTAATATAA</p> <p>ATGGAAATGGAAAAAGCAAAAAGCGTGGTCGACCTGCTCAGTTGTTACA GATTGCTGAGTTACATGCATTTGTAGAGTTTTTAGAACAGCAAGAACAGCT TTCTGATCTTCAATCTCAAGTGTTAAAGGCTTTAAATAGTGTTGATTGTAAT TTTGAGGGGCTAACTCAGACAGATCAGGTTCTGGTCAAAGAGGCTTTAAA ACCTTATCGTGAACATCTGAAGTTGAAATTGCTGTTTGAAGAGTTGAATAA TCTGCCTCTTAAAACCGAGTATGAACAGAAGTTTTTATAGATCTGTATGAACT TTTTCAAAAGAATGCACTAGATCAGATGGAACTAAATTTTTAAAACCTTG GCGACACGGTATCTGAATTTTAAAGCGCAAAGCTTGAATACAGTGATTTA GAATTGATTTGAGTCAGCTACAGAAGAAGGATGCTGGGAAAAACGCAA AGCTGAAAATCAACGTAAATTTGAATTAGGTGGTGCTGTTCTTGTGCGCATT TAAGAACTCAATATTGATATTAGCAATGATACTCCTCAACAGATTACCAAT CGGATTGTGAATACGACAAAGTTTTATAATGAAGTACGTAAATCTTTAATTT TTAAAGATGTTAAGACATATGAAAATGAATTTTTAAAGCAAATAAACTTTTT ATTCAGGTGTTAGAAGGATTACATACCTGGCAAAGGGCGGAGAATTATT GTCAGTAATTGAAATTTAAAAGGCCTTAGAAAAAGGCGAAGAGTAG ATGGACAAAAAGATAGAAGAGAAAATAAAGATATTAATGAAAAAATTCAA AAATTGGATTTAGCTATTCAAGAACTAATAGAACTACAGACAATTTTTTTA TTAATGATAAGTACTATGGAATTGAATGTGCTTTTTATTTTAGACCATGCACT AAATATTTAAAATGCACTTTTTCAACAAGTATCCAGCATTTGCACCACCTCAC TTACGAGGAGTAGAAAATTATGATCGAATTATACATATTAGAGAAAATAATC CTGAAGATAAGCTAGAACAATTTCTCCTCTTAAAGACTAA</p> <p>ATGACCCCAAATGAAAATGGACTCTCGGATCACCTCAGTGAAACAGAGCA GATGAATGCGAAGATGAAAGATAGAACCAGCGACCAAATGGTGCTCTAG AAAGCCAACAAGCAAATGGACAATGTGCCAAGAGTCTGAACCAGCAGGT AAGACTGACTGA</p>
<i>group_1314</i>	hypothetical protein	A1	/	hypothetical protein	<p>AAGACTCAAGACGTAACAGTACCCCATGTTTTGAGTTTCGAGGTGCATTCA CTTCCTGAATACGTATTACCGATACAAAATACGCCTCCCAAACGCACTCAA TCAGAACCAGTACAACCTAAAGCTGATGATTATGATTACGGTGGCCCTGG TTTTTAA</p> <p>ATGCCTGTATGGCTGAAAAATCCGGTTGAAGGGGGTCAAGGGGGAAGCG CCCCATTGCGTACTAGTTTTCTCCGATGAGGAGAACTGTCAAGTGCGAAT TATTTAATATAA</p> <p>ATGGAAATGGAAAAAGCAAAAAGCGTGGTCGACCTGCTCAGTTGTTACA GATTGCTGAGTTACATGCATTTGTAGAGTTTTTAGAACAGCAAGAACAGCT TTCTGATCTTCAATCTCAAGTGTTAAAGGCTTTAAATAGTGTTGATTGTAAT TTTGAGGGGCTAACTCAGACAGATCAGGTTCTGGTCAAAGAGGCTTTAAA ACCTTATCGTGAACATCTGAAGTTGAAATTGCTGTTTGAAGAGTTGAATAA TCTGCCTCTTAAAACCGAGTATGAACAGAAGTTTTTATAGATCTGTATGAACT TTTTCAAAAGAATGCACTAGATCAGATGGAACTAAATTTTTAAAACCTTG GCGACACGGTATCTGAATTTTAAAGCGCAAAGCTTGAATACAGTGATTTA GAATTGATTTGAGTCAGCTACAGAAGAAGGATGCTGGGAAAAACGCAA AGCTGAAAATCAACGTAAATTTGAATTAGGTGGTGCTGTTCTTGTGCGCATT TAAGAACTCAATATTGATATTAGCAATGATACTCCTCAACAGATTACCAAT CGGATTGTGAATACGACAAAGTTTTATAATGAAGTACGTAAATCTTTAATTT TTAAAGATGTTAAGACATATGAAAATGAATTTTTAAAGCAAATAAACTTTTT ATTCAGGTGTTAGAAGGATTACATACCTGGCAAAGGGCGGAGAATTATT GTCAGTAATTGAAATTTAAAAGGCCTTAGAAAAAGGCGAAGAGTAG ATGGACAAAAAGATAGAAGAGAAAATAAAGATATTAATGAAAAAATTCAA AAATTGGATTTAGCTATTCAAGAACTAATAGAACTACAGACAATTTTTTTA TTAATGATAAGTACTATGGAATTGAATGTGCTTTTTATTTTAGACCATGCACT AAATATTTAAAATGCACTTTTTCAACAAGTATCCAGCATTTGCACCACCTCAC TTACGAGGAGTAGAAAATTATGATCGAATTATACATATTAGAGAAAATAATC CTGAAGATAAGCTAGAACAATTTCTCCTCTTAAAGACTAA</p> <p>ATGACCCCAAATGAAAATGGACTCTCGGATCACCTCAGTGAAACAGAGCA GATGAATGCGAAGATGAAAGATAGAACCAGCGACCAAATGGTGCTCTAG AAAGCCAACAAGCAAATGGACAATGTGCCAAGAGTCTGAACCAGCAGGT AAGACTGACTGA</p>
<i>group_1315</i>	hypothetical protein	A1, A2, A3	O	Nuclear body protein	<p>AAGACTCAAGACGTAACAGTACCCCATGTTTTGAGTTTCGAGGTGCATTCA CTTCCTGAATACGTATTACCGATACAAAATACGCCTCCCAAACGCACTCAA TCAGAACCAGTACAACCTAAAGCTGATGATTATGATTACGGTGGCCCTGG TTTTTAA</p> <p>ATGCCTGTATGGCTGAAAAATCCGGTTGAAGGGGGTCAAGGGGGAAGCG CCCCATTGCGTACTAGTTTTCTCCGATGAGGAGAACTGTCAAGTGCGAAT TATTTAATATAA</p> <p>ATGGAAATGGAAAAAGCAAAAAGCGTGGTCGACCTGCTCAGTTGTTACA GATTGCTGAGTTACATGCATTTGTAGAGTTTTTAGAACAGCAAGAACAGCT TTCTGATCTTCAATCTCAAGTGTTAAAGGCTTTAAATAGTGTTGATTGTAAT TTTGAGGGGCTAACTCAGACAGATCAGGTTCTGGTCAAAGAGGCTTTAAA ACCTTATCGTGAACATCTGAAGTTGAAATTGCTGTTTGAAGAGTTGAATAA TCTGCCTCTTAAAACCGAGTATGAACAGAAGTTTTTATAGATCTGTATGAACT TTTTCAAAAGAATGCACTAGATCAGATGGAACTAAATTTTTAAAACCTTG GCGACACGGTATCTGAATTTTAAAGCGCAAAGCTTGAATACAGTGATTTA GAATTGATTTGAGTCAGCTACAGAAGAAGGATGCTGGGAAAAACGCAA AGCTGAAAATCAACGTAAATTTGAATTAGGTGGTGCTGTTCTTGTGCGCATT TAAGAACTCAATATTGATATTAGCAATGATACTCCTCAACAGATTACCAAT CGGATTGTGAATACGACAAAGTTTTATAATGAAGTACGTAAATCTTTAATTT TTAAAGATGTTAAGACATATGAAAATGAATTTTTAAAGCAAATAAACTTTTT ATTCAGGTGTTAGAAGGATTACATACCTGGCAAAGGGCGGAGAATTATT GTCAGTAATTGAAATTTAAAAGGCCTTAGAAAAAGGCGAAGAGTAG ATGGACAAAAAGATAGAAGAGAAAATAAAGATATTAATGAAAAAATTCAA AAATTGGATTTAGCTATTCAAGAACTAATAGAACTACAGACAATTTTTTTA TTAATGATAAGTACTATGGAATTGAATGTGCTTTTTATTTTAGACCATGCACT AAATATTTAAAATGCACTTTTTCAACAAGTATCCAGCATTTGCACCACCTCAC TTACGAGGAGTAGAAAATTATGATCGAATTATACATATTAGAGAAAATAATC CTGAAGATAAGCTAGAACAATTTCTCCTCTTAAAGACTAA</p> <p>ATGACCCCAAATGAAAATGGACTCTCGGATCACCTCAGTGAAACAGAGCA GATGAATGCGAAGATGAAAGATAGAACCAGCGACCAAATGGTGCTCTAG AAAGCCAACAAGCAAATGGACAATGTGCCAAGAGTCTGAACCAGCAGGT AAGACTGACTGA</p>

<i>group_1317</i>	hypothetical protein	A1, A3, A14	S	Plasmid replication protein	ATGGCAAATAAAGCATCACAAAGGTAATGTAAGAACTAGAAATTGGACGTTT GTAATTTATCCAGAAAGTGTCTGAAGGTTGGAGAGATATATTAGATAAT GAACATATACAATGGGTTGAATCTCCATTACATGATAAGGATACTAATCCA AATGGAGAGATAAAAAAGGCTCATAAGCATATATTGGTTATGTATGATGGG GTTAAAAGTTATAATCAGATATTGGAATTGACCGAAAGAATAAATGCGACT GTTCCACAAAGGTGTGGTTCTGCTAAAGGTCTTGTTAGATATATGTTGCAT ATGGATAATCCTGAGAAATATCAGTATGATCGTGAAGATATGATTGCTCAT GGTGGAGCAGATATTCTAGAGATGTTAAAACCGACGTCTGCAAGCAGATA TGAGATGTTTAAAGAAATGACTTCTTTTATAGTTGAAAATGACATAAGAGAA TACGAAGAACTATGGATTTATGCTATGGAGCATAGATTTGATGATTGGTTC CCTTTGTTAGCTGATAATGGTACTTTTGCTATAAATACGTTTATAAAATCTA GAAGACATAGAATAAAAGATAACAAATGA
<i>group_1318</i>	hypothetical protein	A1, A3, A14	/	hypothetical protein	ATGAGTAAAAGATTGCAAGTGACTTTGAAAGATGAGATGTCGGAAAAGTTA GATGAATTGGCGAAATATTTAGGAGTTTCAAAAAGTGCGGTAATTGCGTTG GCTTTGCAAGATTTTGATAAAAAGCATAAAAATAGATAA
<i>group_1319</i>	hypothetical protein	A1, A3, A14	L	Replication protein	ATGTTGAGTTATGGTGGTTTGCTTAAACAAAAACATAAAAATTTTAAATTTAG ACGATGCCGAAGAAGGTAATTTAATTCAAACGAACGATGAAGAAAAAACG ACCGAAGAAGAACAAAAAGCCCATTCAATTACGGCTATTTGGAATTTTGAA AAACAAAATTATTTCTTAAAGAATTTGTAA
<i>group_1320</i>	hypothetical protein	A1, A3, A14	L	COG2801 Transposase and inactivated derivatives	ATGGCTGAAATACGTGTGTTTTATAGAGATGAATTTCTTTGTA CTGCTATAT GTCCAGAAATAGCTGATTACTCCATAGATATAAAAGACATACAGCACGCTC GTAGTCAGAGAAGAAAACACTTAAAACAAAATATTGCTAGTCCAAGTACCA CGGAATTAATTAAGGAAGAAAAAATTGTGGTTATTCACCTCAAGAAACGA CTAAAAACGTCAAAAAATTAAGAGTTATCGTAATGACTAA
<i>group_1322</i>	hypothetical protein	A1, A3, A14	/	hypothetical protein	ATGTACAAAGAAAAATTGAAAACAGATGGTGCATTAGTAATATTTGTGTGG ATTATTTTCTTTATATTTAACATTTTGTATGATGATACAATGCTAAATACTATGCA TAATCTTCCAGATAAAATTTAAATATGAGGGCAAACCTTCAAATATGTTACCA ACATTTTCATGATTTATTTTATGGATTTATAACTTTTGATATTTTAACTTTTTTA TTGCTAATATTTATCACTATAGATTGGTATCTCAGAAAGAAGAAGTAA
<i>group_1324</i>	hypothetical protein	A1, A3, A14	/	hypothetical protein	ATGAACGATAACAAAATGACGCCAGGGGAACTGCGCGGACCTGGGGGT TTGGGGACAGTATTCTCCCTGCGCATGCTTGGCATGTTTATGGTCCTGCC GGTTCTGACCACGTATGGCATGGCGTTGCAGGGTGCAAGTGA

<i>group_1325</i>	hypothetical protein	A1, A3, A14	/	hypothetical protein	ATGAATACCCTGTTTGCCCTTGTTCATCAGCGTGTGTGCTCTCACTGGTGAA TGCTCTGATGTTCTGATCGGTGTTTATCCATCAGAGGCCAGTTGCAACAG CAACGCCGATGAACAAAAAGTACAGGGCCAGTGTCTCCCCTACCGAAATG CACAAAAATGGCTGACGACCAACGGCCTGCAGTGAGTTTTTGA ATGCAAATGCTGGTACATTATGATGTTCTGTTATCCAAAATCGTTTATAGCC GAGTTTGTCTAGCTAATTGTGCTAAATGAACAGTATCTTGTAAGCTTCTTC GGCAGATTTGCCTTCATCAATCACAGCGTAATCTAACACATTAA ATGATGAAGCGCAATATTCTGGCAGTGGTGATCCCTGCCCTGCTGGTAGC CGGTGCAGCCAACGCTGCAGAAATCTATAACAAAACGGCAACAAACTGGA CTTCTATGGAAAAATGGTTCGGCGAGCACGTCTGGACCACCAATGGCGACA CCAGCAGCGACGATACCACCTATGCCCGTATCGGCCTGA ATGCGCTTCCAGTTGCTGAATGGTTTTGCTGAGGGTGGCGTTGGGCATAT TCAGCGAGTCCGCGGCGCGGGTAAAGCTGCCGGCTTCGACAACGCGAGT GAAGGCTTTGAGAGCTGA ATGAACAATTCAAGTTCAAACGGTCAAGAGCTGTTTTTCAATTAGATGCG GCAAGTGCTTTGACACTTAATACCATCACGCTCATGACATTCATGGCTGCA TCTAGTGCGCCGACACCACTCTACCGTCTATACCAACAGCTTTGGCAATTT TCACCGTTACATTGACTTTGATTTTTGCCACCTATGCTTTTACGCTGCTT GGCTCATTACTCATTATTGGTTCTTTGTCTGATTATATTGGACGCCGCCA GTCATTATTGCAGCTATTTCTTTGCAAATTATTTCAATGAGCTTTTTCTTATT TGCTTCAGATGTTTCGATGTTATTCATTGCACGAGGCTTACAAGGAATCGC AACAGGGCTTGCTGTTTCAGCTATAGGCGCAGCAATTTTAGATTTAGTAA GCTTCATGGTTCACTAATTAATAGTATTGCACCAATGATAGGTATGGCTGT TGGGATTTTTCTGACTTGTTCTATTTTGAATTCTCTGCCACCCCTTACAG CTTGTTTTTGAGTTCTTATGTTTTCTTCTTATTTGTGAATCATTTTAAGTTT TCTTACTCCAGAGACCGCACAAAGAAGATCCGGTGCACCTGCTTCATTA ACCCAATATGGCCATTCCGCCACAGACAAAAGCGCTTTATTGAGTATTAG CCCGATTAACATTGCACTTTGGATGGTCAGTGGTTTCTCCTATCGCTTAT GCCATCTTACTGGCAAAAATTTCCACACGTATCGGCATGGCTAAACG GGATTATGTTTATGGCGCTCGCCCTTTCCGGTGCGGTAGGTATATTAACG CTTAGGAAGTCGACAAATTTCCGTATTTTATTAACCGGTACGCTCTCGATT GCTATTGGTGCCATTGTCTTATTCATCGCCATTAATCTAACCAATGCTGTA GTTCTATTTTTAGGTTCTATTACTGGAGTTGGTTTTGGTACAGCATTTA
<i>group_1326</i>	hypothetical protein	A1, A3, A14	/	hypothetical protein	
<i>group_1329</i>	hypothetical protein	A1, A2, A14	/	hypothetical protein	
<i>group_1330</i>	hypothetical protein	A1, A2, A14	/	hypothetical protein	
<i>group_14</i>	hypothetical protein	C4, C6	EGP	Sugar (and other) transporter	

<i>group_1495</i>	hypothetical protein	L6	G	Phosphoglucomutase/phosphomannomutase, alpha/beta/alpha domain I	<p>TGGGGGCTATTCGTAGTGTAAATGCCATTGGCACTTCCCGAAGAACGTGCG GGTCTTATGGCTGCTTTTTTTGTTGAAAGTTATTTGGCTTTTAGTATTCTG CAATTTTAGCTGGCTATTTTGTAGGTAATAATTGGGCTGATGAGTACGGCAA ATAGTTATATTAGTTTTATTATCTTATTGTCACTGGTCGCCCTTCTAATGATT ATAAAAATTTTAAAAATAAATAA</p> <p>ATGGTGCGCAAATATTTTCGGCACAGACGGTATTCGTGGCAAAGCCAACGA AGGCGCGATGACGGCGGAAACCGCCTTGCGCGTCGGCATGGCGGCTGG CCGTGTCTTTTCGTCGCGGTGACCACCGCCATCGTGTCTGTGATCGGCAAG GATACGCGCTGTCTGGGCTATATGCTTGAACCCGCGCTCACAGCCGGTTT CACCTCGATGGGCATGGACGTATTCCTTTTTGGCCCGCTGCCGACAACGT ATAGGAAGAATAAACGCCCTTTTACCCAAGTCCAACAGCTTTGGACCGC AGTTGA</p>
<i>group_15</i>	hypothetical protein	A2, A3, A14	L	Transposase and inactivated derivatives	<p>ATGAACCCATTCAAAGGCCGGCATTTCAGCGTGACATCATTCTGTGGGC CGTACGCTGGTACTGCAAATACGGCACTGTTGCAAAGTTAGCGATGAGGC AGCCTTTTGTCTTATTCAAAGGCCTTACATTTCAAAACTCTGCTTACCAG GCGCATTTTCGCCAGGGGATCACCATAATAAAATGCTGAGGCCTGGCCTT TGCGTAGTGCACGCATCACCTCAATACCTTTGA</p> <p>ATGAACCCATTCAAAGGCCGGCATTTCAGCGTGACATCATTCTGTGGGC CGTACGCTGGTACTGCAAATACGGCATCAGTTACCGTGAGCTGCAGGAGA TGCTGGCTGAACGCGGAGTGAATGTCGATCACTCCACGATTTACCGCTGG GTTTCAGCGTTATGCGCCTGAAATGGAAAAACGGCTGCGCTGGTACTGGC GTAACCCTTCCGATCTTTGCCCGTGGCACATGGATGAAACCTACGTGAAG GTCAATGGCCGCTGGGCGTATCTGTACCGGGCCGTCGACAGCCGGGGC CGCACTGTGATTTTTATCTCTCCTCCCGTCGTAACAGCAAAGCTGCATAC CGGTTTCTGGGTAAAATCCTCAACAACGTGAAGAAGTGGCAGATCCCGCG ATTCATCAACACGGATAAAGCGCCCGCCTATGGTCGCGCGCTTGCTCTGC TCAAACGCGAAGGCCGGTGCCCGTCTGACGTTGAACACCGACAGATTAA GTACCGGAACAACGTGATTGAATGCGATCATGGCAAACCTGAAACGGATAA TCAAACGCCACGCTGGGATTTAAATCCATGAAGACGGCTTACGCCACCATC AAAGTATTGAGGTGATGCGTGCCTACGCAAAGGCCAGGCCTCAGCATT TTATTATGGTGATCCCCTGGGCGAAATGCGCCTGGTAAGCAGAGTTTTTG AAATGTAA</p>
<i>group_16</i>	IS6 family transposase IS15	A2, A3, A14	L	Transposase and inactivated derivatives	

<i>group_17</i>		IS6 family transposase IS15	A2, A3, A14	L	Transposase and inactivated derivatives	<p>ATGAACCCATTCAAAGGCCGGCATTTCAGCGTGACATCATTCTGTGGGC CGTACGCTGGTACTGCAAATACGGCATCAGTTACCGTGAGCTGCAGGAGA TGCTGGCTGAACGCGGAGTGAATGTGATCACTCCACGATTTACCGCTGG GTTACGCGTTATGCGCCTGAAATGGAAAAACGGCTGCGCTGGTACTGGC GTAACCCTTCCGATCTTTGCCCGTGGCACATGGATGAAACCTACGTGAAG GTCAATGGCCGCTGGGCGTATCTGTACCGGGCCGTCGACAGCCGGGGC CGCACTGTGATTTTTATCTCTCCTCCCGTCGTAACAGCAAAGCTGCATAC CGGTTTCTGGGTAAAATCCTCAACAACGTGAAGAAGTGGCAGATCCCACG ATTCATCAACACGGATAAAGCGCCCGCCTATGGTCGCGCGCTTGCTCTGC TCAAACGCGAAGGCCGGTGCCCGTCTGACGTTGAACACCGACAGATTAA GTACCGGAACAACGTGATTGAATGCGATCATGGCAAACCTGAAACGGATAA TCAACGCCACGCTGGGATTTAAATCCATGAAGACGGCTTACGCCACCATC AAAGGTATTGAGGTGATGCGTGCACTACGCAAAGGCCAGGCCTCAGCATT TTATTATGGTGATCCCCTGGGCGAAATGCGCCTGGTAAGCAGAGTTTTTG AAATGTAA</p> <p>TTGATCAGTTTAGGTAAGACGTTGTTTTTCATTGCACACGCTTCAGAAGAT CAAAACGGTGATCAAATTATTTACCGACCAGATCTAGGTGGTAAAAACCGT AACGAGCTTTACCGTATCGCAGATGTGATGGGTTATCTAACAACCTGTTACT ACAGGTGAAGGTAAAAATGCCCGCGTTATTAATTTAAACCCTCGCCTACA CATCATGCGAAAAACTCAGGTGCTTTAGGTGGTGAACCTGGTGAAGTATG GGTACCAGATCTTAAAGCACATCCTACTTTCTTGGCTGACCTGATTACTCA AGCTAAAGATCACATTAACACCTTAACGCCTGCACAACCTTGACAGCAGCTAA AGCCCAAGAAGAGCTAGAAAACCTGGAACAAAGCTGTGAAGAAGCTGAGC ATGCAGGTGACCTTAATCAATTAACCTGAGTCGCTTGATAAAGAACACATGT ATTACCAGAACATGCGCCAAGCAATGTTAATGAGAGCTAAAGCATTGAATT GCACGTTTGATAAACAACGTGGCACTTGGATTAGTCCACCAGAATTTAACG GTATCTCAGATCAACAAAGAGATGAACTTCAAACTTTATTGCTGAACGTG GCCTCGATGTAAAAACAGTTTGTGAGCACTTCGGCATAGATGCCCTGATC CAAATTGAAGCGGCAAACTAACTGCAGTTAAACAAGAAATTGAAACCTTA GCGAAAACGGGGATGACAGCATGA</p>
<i>group_175</i>		hypothetical protein	L9, L10	S	AAA domain	<p>ATGAAAACCGCCACTGCGCCGTTACCACCGCTGCGTTCGGTCAAGGTTCT GGACCAGTTGCGTGAGCGCATACGCTACTTGCATTACAGCTTACCAACCG AACAGGCTTATGTCAACTGGGTTTCGTGCCTTCATCCGTTTCCACGGTGTG</p>
<i>group_19</i>	<i>xerC_1</i>	Tyrosine recom	A2, A3, A14	L	Phage integrase family	<p>ATGAAAACCGCCACTGCGCCGTTACCACCGCTGCGTTCGGTCAAGGTTCT GGACCAGTTGCGTGAGCGCATACGCTACTTGCATTACAGCTTACCAACCG AACAGGCTTATGTCAACTGGGTTTCGTGCCTTCATCCGTTTCCACGGTGTG</p>

binase
XerC

CGTCACCCGGCAACCTTGGGCAGCAGCGAAGTCGAGGCATTTCTGTCCT
GGCTGGCGAACGAGCGCAAGGTTTCGGTCTCCACGCATCGTCAGGCATT
GGCGGCCTTGCTGTTCTTCTACGGCAAGGTGCTGTGCACGGATCTGCC
TGGCTTCAGGAGATCGGAAGACCTCGGCCGTCGCGGCCGCTTGCCGGTG
GTGCTGACCCCGGATGAAGTGGTTCGCATCCTCGGTTTTCTGGAAGGCG
AGCATCGTTTTGTTTCGCCCAGCTTCTGTATGGAACGGGCATGCGGATCAGT
GAGGGTTTGCAACTGCGGGTCAAGGATCTGGATTTTCGATCACGGCACGAT
CATCGTGCGGGAGGGCAAGGGCTCCAAGGATCGGGCCTTGATGTTACCC
GAGAGCTTGGCACCCAGCCTGCGCGAGCAGCTGTGCGGTGCACGGGCA
TGGTGGCTGAAGGACCAGGCCGAGGGCCGAGCGGCGTTCGCTTCCC
GACGCCCTTGAGCGGAAGTATCCGCGCGCCGGGCATTCTGGCCGTGGT
TCTGGGTTTTTTCGCGCAGCACACGCATTTCGACCGATCCACGGAGCGGTGT
CGTGCGTCGCCATCACATGTATGACCAGACCTTTCAGCGCGCCTTCAAAC
GTGCCGTAGAAGGCACTGTTGCAAAGTTAGCGATGAGGCAGCCTTTTGT
TTATTCAAAGGCCTTACATTTCAAAACTCTGCTTACCAGGCGCATTTCGC
CCAGGGGATCACCATAATAAAATGCTGAGGCCTGGCCTTTGCGTAGTGCA
CGCATCACCTCAATACCTTTGA
ATGAAAAAATTGAGTGACGAATACTTACCTGTTAGAAAAGCTCAAACCTGTT
TATGGGAGTATTTTCGGGAAATTATGCCTTTAGAGGTGAGAAAACCATTTGG
TTTGAATCGACCCTTGAGCGTGATTTCACTTAAGCAAGAATTTAATAACA
ATGTGATTGATGTTATCGGTCAACCTGTAGTTATCCCCTACATCACTGAAT
TGGTAACCAATCAACCTTACTCCTGACTTTTTAGTGAGTTTTCAAGCT
CGAATTGTGATGATTTTGGAGATTTTCCTGTACCAATGCTCATCGAAATTA
AGCCTAAGAAAAAGCTTGTGAGGACTGGGATAAGTTAAAACCAAATTC
GAGCAGCCATCGTTTTGCTGCTGATAAAGGATGGAAGTTCAAATTTATCA
GTGAAACAAGATTATATGATCAGTATTGGGAAAATATTAATTTCTTAAAAA
ATTTAGACGCTCTCATGTAGGTGCCGTTGATCAACAACTATTACTCGATAC
CCTACAAGAGTTAGGTCAATCAACAATTAATCAATTACCTGCCCATATTTTT
AAAGATAAAACCAATGTTTTAAAAGGAATTCACCAAGTTTGGGCATTAGTG
GCAAAGCGCATGATTGCTTGTGACTTGTATTGCCCTTTAACGGCTGAAACA
GTGATATGGGTGAACCAAATGATGCATTTGCGAGGAATATCTAA

group_
192

hypoth
etical
protein

L7, L8,
L9

L

TnsA endonuclease N terminal

<i>group_193</i>	hypothetical protein	L7, L9	L	TnsA endonuclease N terminal	<p>ATGAAAAAATTGAGTGACGAATACTTACCTGTTAGAAAAGCTCAAAGCTGTT TATGGGAGTATTTTCAGGAAATTATGCTTTTAGAGGTGAAAAACCATATGG TTTGAATCGACCCTTGAGCGTGATTTTCATACTTAAGCAAGAATTTAATAACA ATGTGATTGATGTTATCGGTCAACCTGTAGTTATCCCCTACATCACTGAAT TGGGTAATCAATCAAACCTACACCCCTGACTTTTTGGTGCAGTTTTCAAGCT CGAATTGTGATGATTTTGAGGATTTCCCTGTACCTATGCTTATTGAAATTA GCCAAGAAAAAGCTTGCTGAGGACTGGGATAAGTTAAAACCAAATTTA GAGCAGCTCATCGTTTTGCTGCTGATAAAGGATGGAAGTTCAAATTATCA GTGAAACAAGATTATATGATCAGTATTGGGAAAATATTAATTTCTAAAAAA ATTTAGACGCTCTCATATAGATGCCGTTGATCAACAATTACTCGATAC CCTACAAAAGTTAGGTCAATCAACAATTAATCAATTACCTGCTCATCTTTTT AAAGATAAAACCAATGTTCTAAAAGGGATTCACCAAGTTTGGGCACTAGTG GCAAAGCGAATGATTGCTTGTGACTTGTATTGCCCTTAAACGGCTGAAACT GTGATATGGGTTAACCAAATGATGCGTTTGTGAGGAATATCTAA ATGATTAGTCTTATAACTGCCACATATAATAGAGAAAACTTCTCAAAAAGT TATATGAGAGTTTATGTGTACAAACCGTTAAAAAGTTTGAGTGGATTGTAAT TGATGATGGTAGTATAGATAAAACAGATGAGTTAATAAATAGTTTTAAATTA GATAATATTATTGATATTACCTATTTGAAAAAACAATGGTGGGAAGCATA CTGCTATGAATATTGGTGTGAAATAGCTAAGTACGCTTATGTTTTTTTTAT AGATAGTGATGATTTTCTTCCCTAACGATAGTATAGAAAAATTATCAATTAT ATTGATAAGGTTTCTGCTAGAGAAGACTATTCAGAAATATCGGGAGTCTGT GGTCTGATTGCTGACTTTCAAGGTAATTTAATCGGTAATAATTCAGAA AACTTATGCTGTAGCTATATAGATTATAGATATAAATATCATATTAAGGTG ATAAGGCTGAAATTTTTAAGAGAGAGGTTCTGCTAGAATATAAATTTCCAG TTATTGAAAAAGAAAAATTTTGCCAGAAAGCCTTAGTGTGGAATAGAATTT CCGATAAATATAAAATGTATTTTTTTAATGAAGTTATATATTATAGAGAGTAT TTGGAGGGTGGGTTATCTGATAGAAGTGTTGAAATAAGAAAAAAGGCTCC CATATCAACATTATTATATTATAAAGAATTATATTTAAATAAAAAATTTAAATT GGTATTATAGAATAAGAGCTTATATAAATTATATGAGATTTAAGTTTATTCT TAA</p>
<i>group_2056</i>	hypothetical protein	C2	M	Glycosyltransferase, group 2 family protein	<p>ATGGCGATGAATAAGACAATGGCACCAATAGCAATCGAGAGCGTACCGGT TAATAAAATACGGAAATTTGTGACTTCCCTAAGCGTTAATATACCTACCGC ACCGGAAAGGGCGAGCGCCATAAACATAATCCCGTTTAGCCATGCCGATG</p>
<i>group_2057</i>	hypothetical protein	C2	/	hypothetical protein	<p>ATGGCGATGAATAAGACAATGGCACCAATAGCAATCGAGAGCGTACCGGT TAATAAAATACGGAAATTTGTGACTTCCCTAAGCGTTAATATACCTACCGC ACCGGAAAGGGCGAGCGCCATAAACATAATCCCGTTTAGCCATGCCGATG</p>

<i>group_2058</i>	hypothetical protein	C2, C6	/	hypothetical protein	<p>ACGTGTGGAAAATTTTTGCCAGTAAAGATGGCATAAGCGATAGGAAGAA CCACTGACCATCCAAAGTGCAATGTTAATCGGGCTAATACTCAATAAAGCG CTTTTTGTCTGTGGCGGAATGGCCATATTGGGTTTTAATGAAGCAAGTGCA CCGGATCTTCTTTGTGCGGTCTCTGGAGTAAGAAAACCTTAAAATGAGTTCA CAAATAAGAAGAAAACATAAGAACTCAAAAACAAGCTGTAAGGGGTGGGC AGAGAATTGCAAAATAGAACAAGTCAGAAAAATCCCAACAGCCATACTAT CATTGGTGCAATACTATTAATTAGTGAACCATGA</p> <p>ATGTTGATCTTGATTGGTATGGCACTTGTAGCGTTATCGGTCTACCGTGGT ATAAGTTTAGGCAGAGATTTAGAAGGCTCGTCGGTAACCCAACCTATCCAAT AGCATGTATATTCAGGAAATGGTAATCGCGCAGATTAAGCAGGCTTGGTA TTGTATTGTCGGTTTTACCGGAGCAACTATGATTGCAATGGGGGTATGGC ATACCTGGCTATCTTAA</p>
<i>group_2059</i>	hypothetical protein	C2, C6	/	hypothetical protein	<p>ATGGCAGACGTGGGCGATGGCGGTCGCTTCCATTTCCACGGCGATCGCC TGCGGGAAATTATGGCGGATCTTCGCCAGACCGACAGAGCCGTTAATGAA GGCATCCCCGCTAACGATGAGGCCGCGAACGGCGTTGAGATCTAA GTGGCTTTAGCACCACCACCTACAGCAGCAGCATTTTCATTTGTAGCCTCA GCTCCTTTACCAATTGCTGCTGAATTATCACCTGTTGCCTGCGCACCACTA CCAATTGCTGCAGAACTGTTCCCAGTCGCTTTTGTCTCTTCACCCGGTTTA TTACCAGAGCCAATTGCTGTTGAATTTGTACCCGATGCCAATGAGTTAATA CCCATTGCCGATGCGCCATCACCTGTTGCTTGTGCAACAGCACCGAATGC AGCGGCACCTTTACCCGAAGCATTTGACATGACGCCAAAAGCAGAAGAAG CATCACCAGTTGCTTTAGAGGTCATACCAAAGCAGAAGAAAACGCTCCT GTACCTACAGCTTGA</p> <p>ATGAGTGAATCAAAAGTTAGACATTTGGTACTTAAAGAGTTTCAGATAAAT CTTCTCATCTTGCTCTTTGTGACGAGGAAACAGGTATTCCATTAGCTGGAT TAACCGCTGTAAAAATGAATTGTAGTGTTTTTTGAGGGTCCAGCGACTATCA CGGCAACATTTGATGTAGGTGGTCCTCAAGGCATCCGCTTAGTTGGTGAT GAACCTAGACAAAAGGTTTGGAGTGCAAAGGAAACGTAG</p>
<i>group_2063</i>	hypothetical protein	L7, L10	/	hypothetical protein	<p>ATGAGCACACCACAATATCAAACAATGAAAGAAAGTGAAGTTTGAATGCC ATCGGATGGGGGTTAATTGTTCTAGGTATTATATCTGGATTTATTTTTATAC TTGTGTTTGGCCGAGTTGAAGTTCCAAGAACTTATTATGGCACCGAGACC GTATGGTCAGGAATCATGGTTATTACAGGTATCGGGATAATCTTAAATGGA</p>
<i>group_2102</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGAGTGAATCAAAAGTTAGACATTTGGTACTTAAAGAGTTTCAGATAAAT CTTCTCATCTTGCTCTTTGTGACGAGGAAACAGGTATTCCATTAGCTGGAT TAACCGCTGTAAAAATGAATTGTAGTGTTTTTTGAGGGTCCAGCGACTATCA CGGCAACATTTGATGTAGGTGGTCCTCAAGGCATCCGCTTAGTTGGTGAT GAACCTAGACAAAAGGTTTGGAGTGCAAAGGAAACGTAG</p>
<i>group_2115</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGAGTGAATCAAAAGTTAGACATTTGGTACTTAAAGAGTTTCAGATAAAT CTTCTCATCTTGCTCTTTGTGACGAGGAAACAGGTATTCCATTAGCTGGAT TAACCGCTGTAAAAATGAATTGTAGTGTTTTTTGAGGGTCCAGCGACTATCA CGGCAACATTTGATGTAGGTGGTCCTCAAGGCATCCGCTTAGTTGGTGAT GAACCTAGACAAAAGGTTTGGAGTGCAAAGGAAACGTAG</p>

group_
2148

hypothetical
protein

A2, A3,
A14

/

hypothetical protein

TTCTTAGTGGGCTATCTGTTCCAAAAGGTTGCCAGCATATTGAGATATCAC
GAGAACAAGAGCGCATCTTAA
ATGACTTGGAGTGCACATGAGGTCTATGACAGCTTTCAGGTTATACCTGAT
AATGATTTAAAACCTCATTCAATTTTTTTACTGCGAATGCCATCCCGAATATG
TGGATGGCATTTTTTATTCATAACTCATTATTTATGGTAGAGAAGCGACTTAA
ATGAGTACTGATGCGGCTCGGGATAAAGCCATCCGGATCGAGGCGCAAG
AAGATTTATATTTCTTCACAAGGTACATGTTTAAAGGAGCGCCGTGGTTATA
AATGGATGCAAAATTGGCACCCTTAGAAATCTGCGAAGCTTTAATGAAAG
TTTATCGCGGAGAGATAAAGCGGTTAATTATTAACGTTCCACCACGATATT
CTAAAAGTAAATTGCTGTAATTAATTTTCATGGCTTGGTGTGTTTGGTAAGAA
TCCAGACTGTGAGTTTATTCATATCAGTTACTCGGCAATGCTTGCCGCAAA
TAATGCCTTCCAAATACGAACCTTGTACAAGAAGAGGCGTATAGAAAAGT
CTTTCCTGAGCTTACATTGCGTGATGATAGTAAGGCTAAAGACTTCTGGAG
AACTTCTCAAGGCGGTGTCTGCTATGCGACAGGTACAGGCGGTACGATTA
CTGGTTTTGGTGCAGGAAAACCTTCGTAAGGCTTTGGTGGCTGCATTATTA
TTGATGACCCACATAAAGCACATGAAGCTTCATTAAAAAGTATTCGAGAAG
GGGTAATTGATTGGTTTTCAGAACACACTCGAATCGCGTACTAACTCGCCA
GATACGCCGATCATTGTGATTATGCAGCGACTTCATGAAGATGATTTAGCT
GGATGGTTGCTAGGTGATAGAAAAGACGGCGTTCCCTGTAGCTGGTGGTAA
CGGTGAAGTGTGGGAGCATCTATGTCTTTCAGCTATTCAGGAAGACGGAT
CCGCACTGTGGCCAGCAAAACACAATATCCAAAATTGAGGCTAATGGAG
CAAGCAGCACCATATGTATTTGCCGGGCAGTACCGACAAATGCCATCACC
GCCAGCAGGCGGTTTTTTAAGCCCAGCAATATTCAAATTGTTGATGCTTT
GCCTGCGGATGTAGTGAAACAAGTTAGGGCTTGGGATTTTGGGGCTACC
GAAAATGAGGGCGACTTTACAGTAGGTGTGCGAGAAGCTCTAGGCGCAG
ATGGTTTTACTTACATTGTCGATGTAAGTACTAGAGGACAGCTTGGACCTGACA
ATGTGAATAAGCGCTTAGAACAAACAGCAAAAATAGATGGGAAAAAGTTT
CTGTGCGTCTACCACAAGATCCCGGTCAAGCTGGTAAATCACAAGCTAGT
TCATTTGTGAAGCTTCTTGCGGGTTATAGCGTGATAGCTAAGCCAATTTCA
GGTGACAAGCTTACACGTGCACAACCATTTGCGGCCCAAGTTAACGTAGG
AAATGTACGAATGCTCAAAGGTGAATGGAATAAGGATTTTATTGATGAGCT
TCGTCATTTTCTAATGGCACACATGACGACCAAGTGGATGCAGCTTCAG

group_
2153

hypothetical
protein

A2, A3,
A14

S

Terminase RNaseH-like domain

<i>group_2299</i>	hypothetical protein	A2, A3, A14	S	Uncharacterized conserved protein (DUF2280)	<p>ATGCGTTTAATGAATTACATGAAGGTTTTGAAGCCTTCTTTGCTGATATGG GATTTGCTCGATGA</p> <p>ATGGCGACACTAAAAGAGCCTGTGAAAATCTTTATAGTTCAGTCTCTTGCT TGTCGTGATACACCTCAAGAAGTGGCTGAACTCGTAAAACAAGAGTTTGG CGTTGATATAGATCGTGTTCAAGTTGCAACTTATGACCCTACAAAGGTTGC TGGTAAGAACTTAAGCAAAAAGTATGTGCGAACTATTTGAAAAAACCAGAGA TGAGTTTGATAAAGGCTTAATTGATATTCCAATTGCTAATAAGTACTACCGA TTGAAGCAATACCAAAGACAACCTTGAGAAGACTAGAAACGTCAAACAGC CTTAAAAATCTTGAGCAAGCCGCTAAAGACATTGGTGGTCAATTTACTAA TCGCCAAGAAATTACAGGCAAAGACGGCGGACCAGTCCAAACAGTTAATT CAGAAATTCCAGTTCCAATGGAAGATTACTTAAAAGCGCGGAGGGAAGTC TTAGATGAGTACTGA</p> <p>GTGGTCAAAGCCCATAGTGGACACATGGCGACAGCAGATTTTGTAGCGCA GGCAATTGAGCGTGCCGTAGAAAAATCAAATATGCCTAAAGGCGTATTTAA CATGATCTACGGTAATGGTGTGGGTGAACCACTGGTTAAGCATCCTTTAAT TCAAGCTGTGGGTTTTACAGGTTTATTGCGTGGTGGACGAGCTTTATGTG ATATGGCCGCAGCCCGTCCACAACCGATTCCGGTGTGGTGGTAAATGAGC AGTATTAACCCGATGCTGATGTTGCCAGAAGCCTTAAAAAACCCTGGTGA AAAAATTGCACAGGATTTGGCAGATTCAGTGGTTTTAGGCTGTGGTCAGTT CTGTACCAATCCGGGTTTAATCTTGGGAATTAATCAGCTGAATTTAGTCA GCTCATTAAACAACCTGACCGACATTATGGGGGCTAAGCCTGCGCAGACTA TGCTCAATGCCGGCACCTTAAAAAGTTACACCGCTGGTCTTGAGCACTTG ACCCAGCATCAAGGCATTAAGCATTGGGCTGGCAATACTCAACAAGGTAA TCAGGCGCAGCCACAACCTGCTTAAAGCTGATGTTGAGCTTTTACTGGCGG GTGACCAGCTTTTACAAGAAGAAATCTTTGGGCCAACACAGTCGTCATT GAAGTTGAAGATAAAGCCCAACTCATTGAGGCACTACAAAGTATGAATGG GCAATTGACTGCGACTTTAATTGCCGATGAAGCTGACTTAACCGAGTTTGC AGATGTAGTTCCAGTTTTAGAGGAAAAAGCCGGTCGATTACTTATAAATGG CTACCCAACCTGGTGTGAGGTCTGTGATGCCATGGTTCATGGTGGGCCAT ACCCAGCCACTTCAGATGCACGAGGTACTTCAGTTGGAACCTTGCCATT GACCGTTACTTGCGCCCTGTGTGTTACCAGAACTATCCACAGAGTCTATT GCCGGAAGCCTTAAAAGACAGCAACCCGTTGCAGATTTTAAAGATTGGTCA ATGGTGAGATGACCAGAGAAGCGATCTAA</p>
<i>group_23</i>	Alpha-ketoglutaric semialdehyde dehydrogenase	L8	C	Aldehyde dehydrogenase family	<p>ATGCGTTTAATGAATTACATGAAGGTTTTGAAGCCTTCTTTGCTGATATGG GATTTGCTCGATGA</p> <p>ATGGCGACACTAAAAGAGCCTGTGAAAATCTTTATAGTTCAGTCTCTTGCT TGTCGTGATACACCTCAAGAAGTGGCTGAACTCGTAAAACAAGAGTTTGG CGTTGATATAGATCGTGTTCAAGTTGCAACTTATGACCCTACAAAGGTTGC TGGTAAGAACTTAAGCAAAAAGTATGTGCGAACTATTTGAAAAAACCAGAGA TGAGTTTGATAAAGGCTTAATTGATATTCCAATTGCTAATAAGTACTACCGA TTGAAGCAATACCAAAGACAACCTTGAGAAGACTAGAAACGTCAAACAGC CTTAAAAATCTTGAGCAAGCCGCTAAAGACATTGGTGGTCAATTTACTAA TCGCCAAGAAATTACAGGCAAAGACGGCGGACCAGTCCAAACAGTTAATT CAGAAATTCCAGTTCCAATGGAAGATTACTTAAAAGCGCGGAGGGAAGTC TTAGATGAGTACTGA</p> <p>GTGGTCAAAGCCCATAGTGGACACATGGCGACAGCAGATTTTGTAGCGCA GGCAATTGAGCGTGCCGTAGAAAAATCAAATATGCCTAAAGGCGTATTTAA CATGATCTACGGTAATGGTGTGGGTGAACCACTGGTTAAGCATCCTTTAAT TCAAGCTGTGGGTTTTACAGGTTTATTGCGTGGTGGACGAGCTTTATGTG ATATGGCCGCAGCCCGTCCACAACCGATTCCGGTGTGGTGGTAAATGAGC AGTATTAACCCGATGCTGATGTTGCCAGAAGCCTTAAAAAACCCTGGTGA AAAAATTGCACAGGATTTGGCAGATTCAGTGGTTTTAGGCTGTGGTCAGTT CTGTACCAATCCGGGTTTAATCTTGGGAATTAATCAGCTGAATTTAGTCA GCTCATTAAACAACCTGACCGACATTATGGGGGCTAAGCCTGCGCAGACTA TGCTCAATGCCGGCACCTTAAAAAGTTACACCGCTGGTCTTGAGCACTTG ACCCAGCATCAAGGCATTAAGCATTGGGCTGGCAATACTCAACAAGGTAA TCAGGCGCAGCCACAACCTGCTTAAAGCTGATGTTGAGCTTTTACTGGCGG GTGACCAGCTTTTACAAGAAGAAATCTTTGGGCCAACACAGTCGTCATT GAAGTTGAAGATAAAGCCCAACTCATTGAGGCACTACAAAGTATGAATGG GCAATTGACTGCGACTTTAATTGCCGATGAAGCTGACTTAACCGAGTTTGC AGATGTAGTTCCAGTTTTAGAGGAAAAAGCCGGTCGATTACTTATAAATGG CTACCCAACCTGGTGTGAGGTCTGTGATGCCATGGTTCATGGTGGGCCAT ACCCAGCCACTTCAGATGCACGAGGTACTTCAGTTGGAACCTTGCCATT GACCGTTACTTGCGCCCTGTGTGTTACCAGAACTATCCACAGAGTCTATT GCCGGAAGCCTTAAAAGACAGCAACCCGTTGCAGATTTTAAAGATTGGTCA ATGGTGAGATGACCAGAGAAGCGATCTAA</p>

<i>group_2322</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGTTATATTTTTTAAAACATCAGAATTTATACAATATGAAAACAATAGCTTT TGTATGTCTAACCCTAATTTCCATCACTTGTTTAGCTGAACCAAGTCAAAAA TATCTTAAAGAATATGATCGATTGTCTGAAGCTTTGGAGTCAGCAATGGCA AATGCATATTCTTTTGATCCTGCAACTGGTCAAGTAAACAGGCTACTCAA GGTTTAGAAGCTAAAAATAATTTATGTAGAGCTGCCAGGCGAAACTAAAC CTCACCACGTTTTTAAAAGACAATTTAGAGGAATCTAAAGAGCTTTATAAAT CTATTGATGGTGCAGAGACTCTAGATAAAAATTATCTTAGTGGACAACAGC AGGAACAACAAAATCTCGTTTCAAATTTGAAAAAGACCTTGTTGGAAGT GATTTAACTGTGAGTAA</p> <p>ATGAAAAAGATTTTATTAGCGGGATTTCTTGATTGGGCTTAGCGGGGTGT GCGACAACCTCCCAACAACCCTCAGAGCCTGTAAAATTTGAAAAGGTTTAT CAAATTGATGGATTAACCAAGCACAGATTTATGATGGCGCTAGACAATGG TTCGCTGTAGCTTTTGCTTCTGCTAACGCAGTAATTC AATATGAAGATAAG GCATCAGGCACTATCATTGGAAAGGGCAATATGCGATACCTTGTTTCGGG CATGGAGTGCTTGGCAATGACAGGAAACGAACGTGTTGATTTTACTGTAA GAGTGGACACTAAGGATGGGAAAATGCGCGTGGGTTATGATGGTTTAAAC TATAGCGCTCCATCGCACATGAGTGCTGGAATAATGATGCCTGCACAAAA TTACCCTATAACTGAAAGTAGGAAGTCCACACCACTGATTATTAGTAAGAT TAATACTCTATCGGATGATATGGCTGAAAAGATTA AACTCAGCAGAAAGT AAATTCGAATTGGTAA</p> <p>ATGGATGATAAAGAGTACTTTTGGCTTACACAAAAAAAAGAGCTCAAACG AAACCCAAATCCAGACCACTGCCTAAAGCTAAAGAAAAATATCTCGAGGC CGAAGAAACCTTATTTCAAGAACTAGAAGAGCATCGAATTGGTTATAGAAG AAAATTTCAATTTGAATCAACAAAAAATTGGCGGTTTCGATTTTATATTGTG AAGTTGAATCTTCTTATAGAAATTGCTGGCAGTCCGTGGGCAGTTGGCCG AGGTGGCACAAAGATAGCAAATTCATTTAATAAGTATGATCTAGCACTAGA CCGAGGTTATGTATTTGAGCGTCTTGAGCCTCACCAAATTGAATCAGGTTA TGCAATCAACTGGATTA AAGCAAATTAGCGAGAATTGAAGATGGATCAGA TCAGACCATTTCTCCA ACTGA</p> <p>ATGAGAATCATATTTTCTTTAATTACGTTTGTCTTATTTTCATTTATTTCTTT ATCCTTTTAAGGAATAAATATATTGAGCCAAACCACTTCGTCATTTTGATAA TATTTTCTGCAATTGTATCCGCAATAATTGCATATTTTGTATGAGGTTCAAGA GCTATCTATTGGAGGCAATATCGTTAAACTAAAAGAAGCAAAAAAGGAGTT</p>
<i>group_2370</i>	hypothetical protein	A2, A3, A14	S	Domain of unknown function (DUF4468) with TBP-like fold	<p>ATGAAAAAGATTTTATTAGCGGGATTTCTTGATTGGGCTTAGCGGGGTGT GCGACAACCTCCCAACAACCCTCAGAGCCTGTAAAATTTGAAAAGGTTTAT CAAATTGATGGATTAACCAAGCACAGATTTATGATGGCGCTAGACAATGG TTCGCTGTAGCTTTTGCTTCTGCTAACGCAGTAATTC AATATGAAGATAAG GCATCAGGCACTATCATTGGAAAGGGCAATATGCGATACCTTGTTTCGGG CATGGAGTGCTTGGCAATGACAGGAAACGAACGTGTTGATTTTACTGTAA GAGTGGACACTAAGGATGGGAAAATGCGCGTGGGTTATGATGGTTTAAAC TATAGCGCTCCATCGCACATGAGTGCTGGAATAATGATGCCTGCACAAAA TTACCCTATAACTGAAAGTAGGAAGTCCACACCACTGATTATTAGTAAGAT TAATACTCTATCGGATGATATGGCTGAAAAGATTA AACTCAGCAGAAAGT AAATTCGAATTGGTAA</p> <p>ATGGATGATAAAGAGTACTTTTGGCTTACACAAAAAAAAGAGCTCAAACG AAACCCAAATCCAGACCACTGCCTAAAGCTAAAGAAAAATATCTCGAGGC CGAAGAAACCTTATTTCAAGAACTAGAAGAGCATCGAATTGGTTATAGAAG AAAATTTCAATTTGAATCAACAAAAAATTGGCGGTTTCGATTTTATATTGTG AAGTTGAATCTTCTTATAGAAATTGCTGGCAGTCCGTGGGCAGTTGGCCG AGGTGGCACAAAGATAGCAAATTCATTTAATAAGTATGATCTAGCACTAGA CCGAGGTTATGTATTTGAGCGTCTTGAGCCTCACCAAATTGAATCAGGTTA TGCAATCAACTGGATTA AAGCAAATTAGCGAGAATTGAAGATGGATCAGA TCAGACCATTTCTCCA ACTGA</p> <p>ATGAGAATCATATTTTCTTTAATTACGTTTGTCTTATTTTCATTTATTTCTTT ATCCTTTTAAGGAATAAATATATTGAGCCAAACCACTTCGTCATTTTGATAA TATTTTCTGCAATTGTATCCGCAATAATTGCATATTTTGTATGAGGTTCAAGA GCTATCTATTGGAGGCAATATCGTTAAACTAAAAGAAGCAAAAAAGGAGTT</p>
<i>group_2407</i>	hypothetical protein	A2, A3, A14	S	Protein conserved in bacteria	<p>ATGAAAAAGATTTTATTAGCGGGATTTCTTGATTGGGCTTAGCGGGGTGT GCGACAACCTCCCAACAACCCTCAGAGCCTGTAAAATTTGAAAAGGTTTAT CAAATTGATGGATTAACCAAGCACAGATTTATGATGGCGCTAGACAATGG TTCGCTGTAGCTTTTGCTTCTGCTAACGCAGTAATTC AATATGAAGATAAG GCATCAGGCACTATCATTGGAAAGGGCAATATGCGATACCTTGTTTCGGG CATGGAGTGCTTGGCAATGACAGGAAACGAACGTGTTGATTTTACTGTAA GAGTGGACACTAAGGATGGGAAAATGCGCGTGGGTTATGATGGTTTAAAC TATAGCGCTCCATCGCACATGAGTGCTGGAATAATGATGCCTGCACAAAA TTACCCTATAACTGAAAGTAGGAAGTCCACACCACTGATTATTAGTAAGAT TAATACTCTATCGGATGATATGGCTGAAAAGATTA AACTCAGCAGAAAGT AAATTCGAATTGGTAA</p> <p>ATGGATGATAAAGAGTACTTTTGGCTTACACAAAAAAAAGAGCTCAAACG AAACCCAAATCCAGACCACTGCCTAAAGCTAAAGAAAAATATCTCGAGGC CGAAGAAACCTTATTTCAAGAACTAGAAGAGCATCGAATTGGTTATAGAAG AAAATTTCAATTTGAATCAACAAAAAATTGGCGGTTTCGATTTTATATTGTG AAGTTGAATCTTCTTATAGAAATTGCTGGCAGTCCGTGGGCAGTTGGCCG AGGTGGCACAAAGATAGCAAATTCATTTAATAAGTATGATCTAGCACTAGA CCGAGGTTATGTATTTGAGCGTCTTGAGCCTCACCAAATTGAATCAGGTTA TGCAATCAACTGGATTA AAGCAAATTAGCGAGAATTGAAGATGGATCAGA TCAGACCATTTCTCCA ACTGA</p> <p>ATGAGAATCATATTTTCTTTAATTACGTTTGTCTTATTTTCATTTATTTCTTT ATCCTTTTAAGGAATAAATATATTGAGCCAAACCACTTCGTCATTTTGATAA TATTTTCTGCAATTGTATCCGCAATAATTGCATATTTTGTATGAGGTTCAAGA GCTATCTATTGGAGGCAATATCGTTAAACTAAAAGAAGCAAAAAAGGAGTT</p>
<i>group_2423</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGAAAAAGATTTTATTAGCGGGATTTCTTGATTGGGCTTAGCGGGGTGT GCGACAACCTCCCAACAACCCTCAGAGCCTGTAAAATTTGAAAAGGTTTAT CAAATTGATGGATTAACCAAGCACAGATTTATGATGGCGCTAGACAATGG TTCGCTGTAGCTTTTGCTTCTGCTAACGCAGTAATTC AATATGAAGATAAG GCATCAGGCACTATCATTGGAAAGGGCAATATGCGATACCTTGTTTCGGG CATGGAGTGCTTGGCAATGACAGGAAACGAACGTGTTGATTTTACTGTAA GAGTGGACACTAAGGATGGGAAAATGCGCGTGGGTTATGATGGTTTAAAC TATAGCGCTCCATCGCACATGAGTGCTGGAATAATGATGCCTGCACAAAA TTACCCTATAACTGAAAGTAGGAAGTCCACACCACTGATTATTAGTAAGAT TAATACTCTATCGGATGATATGGCTGAAAAGATTA AACTCAGCAGAAAGT AAATTCGAATTGGTAA</p> <p>ATGGATGATAAAGAGTACTTTTGGCTTACACAAAAAAAAGAGCTCAAACG AAACCCAAATCCAGACCACTGCCTAAAGCTAAAGAAAAATATCTCGAGGC CGAAGAAACCTTATTTCAAGAACTAGAAGAGCATCGAATTGGTTATAGAAG AAAATTTCAATTTGAATCAACAAAAAATTGGCGGTTTCGATTTTATATTGTG AAGTTGAATCTTCTTATAGAAATTGCTGGCAGTCCGTGGGCAGTTGGCCG AGGTGGCACAAAGATAGCAAATTCATTTAATAAGTATGATCTAGCACTAGA CCGAGGTTATGTATTTGAGCGTCTTGAGCCTCACCAAATTGAATCAGGTTA TGCAATCAACTGGATTA AAGCAAATTAGCGAGAATTGAAGATGGATCAGA TCAGACCATTTCTCCA ACTGA</p> <p>ATGAGAATCATATTTTCTTTAATTACGTTTGTCTTATTTTCATTTATTTCTTT ATCCTTTTAAGGAATAAATATATTGAGCCAAACCACTTCGTCATTTTGATAA TATTTTCTGCAATTGTATCCGCAATAATTGCATATTTTGTATGAGGTTCAAGA GCTATCTATTGGAGGCAATATCGTTAAACTAAAAGAAGCAAAAAAGGAGTT</p>

group_
2497

hypothetical
protein

A2, A3,
A14

/

hypothetical protein

ACAAGTAACAATAGATCAATTAAGTCAATTAAGTTTCAACATATCGGATG
TTACTTTTGAAAAGTTTACATTTTTCAGGTGTTTTTGGAAAGCAGCCATTTAG
TGGATAGTAGAGCAGAATATTTTTTTCCTCATCAATGAAATTAACAATC
GGATTGTTCAATGATCTTAAGTCTGAAATAAAAGTTCAATTAACAAGGTTG
TTAATTGATCAATTAATAAATTTTATCCTTTATTTTATGGCAAACAATTCAA
TGATAGCGATGAATCCCTAAATCTACGGTTTTTTATATCGAGTTGAAAGAT
GAGATTATTGATAAAGTTCATCAAAAACGGACACCTGTTATACCATTTGAT
CAAAAAAGCAGGAAATTGTTACAGCTATAGATAACTATGCAGCTTTGTAT
ATTTTATTTAAGAAGTTGAACAGTAG
ATGCAAGTTTTAATGGCTTTGGCATACCCTTTACAGGGCATGATATGGGC
TTACGTGACGAAATTCAGGCAGACATTGCTGAAGCATTTAATGATGATTTA
GCAGATGCCATTCATACCTTTACATGTGAGCGGATCTCTAAAACGAATTGG
GATCCTAAAACCTGAACTTATGTTGAAGTTAAAGAAAACCTATTCTGGCCGT
GGCGTTCTGTTTGGCTCATAACAGTCAATATGAGATCCAAACCTTGGAGTA
CTGGCCACGGATAAAAAGGCTACAGTGCTGCAGAATGAAGTTACCAAAGA
GCCAAAGATTGATGATGAGTGGTTAACAGCCTTAGGCTCATTCCGGGTAA
TTCATATTCAACAGGATCCAGCTTCTACTATTTGGAAATGTCAGTTGAGGA
AGGTTTAA
ATGATTGAATCTAATGCTGTCAATATCGATCGTAGTCGTATCATTCTAAAAC
CCAATGTGGTTGTCGGTTATGAGGGGCAACCCTATAAAATTGTGAATGTTT
TAAATGCCAATGACATTGTTATATCAAGTCTTGATTCCGGTACGTAGTCTAC
AGGTCAATGCTAGAAGTTTGACCATTTTTAATGGGGGAAATACGCTAACTG
AGCGTTTGAATAAGGGTGATCAGGATATTAGTAATGAAGCTTGGCAGATT
GCTTTACAACGTTATGAGATTATTAAGCCTTTGATTGAGTATTCAACGACT
GAACTGGTCGAAAATAGGGCGAATGAATATGATGTGAATCGCTCAACATT
GTGGAATGGCTTAAAGATTATCGGGAAAATAATTCTCTAATGGCATTGAT
TCCCAAAAACGAGGTTGGACAACCTGAAAATCAAGGTTGTCACCTCAAG
TGAGCAATATCATTAAACAAGCTATCAATGATGAATTTTGAATGCGAAGA
AACCCAGTATTTCCAAAACGATTGAAATTGAAAAGCAGAGTGTCTCGGC
TTCAGCTTGAAGCACC GCATGAAAACCTATTCGTAGAAGAATCGAGGCTT
TAAATGATTACCAAGTAACGAAAGCACGTCTTGGATCTAAAGCTGCAATAG
ATAAATTTAAGGCTGCTGCTGGTTCATTTCCAATGCTGATTACCTTTGG
CATATGTGCAGATTGATCATAACCTCTAGATATTGAAATCGTGGATGACG

group_
25

hypothetical
protein

L7, L8,
L9

L

Mu transposase, C-terminal

group_
2532

hypoth
etical
protein

A2, A3,
A14

/

hypothetical protein

AGTTTAGAGAAGCGATTGGAAAACCTCATTTAACCTTGGCAATCGATGTTT
TCAGTCGGATGATCGTAGGGTATTACCTATCACTAGAAGCGCCAAGTACG
ACTTCAGTGGCGATGTGTATCGCATCTTGTATTTTTGTCAAAAAACGTAAG
TTGATCGAACTCGATATTGATGCTGAATGGCAGGTGCAAGGGATTATGGA
CTCGGTGCATACAGATAATGGCCCCGATTTTAGAACCAATCACATCAGTAA
AGCTTGTCTAAAGTATGGGATTCATTGGGAATACCGACCAATTGGTGGGG
CAAGATTTGGCGGTCATATCGAGCGTATGCTTGGTATCGTCAATCTTGAAA
TGCATGTCTTGGATGGTACAACCTTTTTCAAATGTACAGCAGCGTGGGACTT
ATGATTCCGCTAAACAAGCCTGTATGACATTAAGAGCTTGAATACTATA
TCGTGTATTGGATTACTAAAGTTTATCACCAGAAGAAACATTCCGCTTTAG
GAACATCCCCAATCGTTAAATGGGAAGAAGGTGTTTGGGGGACGAAGAC
GACAGCAGGTACAGGTTTGAAAGAACGTGTATCAGATGAAGATACCTTGT
TTATTGATTTTCTACCTGAATTTGAAGCGACTATTCAGCGTACAGGGGTGC
AAAAAGATAACTTATTCTATTTTGCCGACTGTTTAAGACAATGGGTTAATTC
TATTGATCCAGAGGACAATAATCGAAAGCGCAAAAGAAAATTCTTGTTTAA
ACGTGATCCGAGGGATATTTACAGATTTGGTTTTATGAGCCATATTCTAA
TACCTATTTTAAAGTACCTACCGCAAAGCGAGAGATACCACCGATTAGCTT
ATTTGAATATAAGCAGGTTTCAGAACTATCTCAAAAGTGAACGTCAGGATGT
TCAGAATCAGGATGAGATTTACAAGGCGATTTTGCATCTACGTGATCAACT
TAATCAGGCTAGAAGCTTGACTCGAAAGCAACGTCGCTCAAACCAAAGAA
AAAAGGAAAATGCAAAAGCCATTACTCAGCTTTCCGAACAGAATCAGTCTA
AAAAGGCTGTAGTTTCTGAAAGTCTTCAGACAAGTGATGATTTGTGGAGTA
CTCCTTTAACTGCATTTGACGATTTAAGGTGA
ATGAAACTTAAAACAGTAACAATCGACGGTAAAGTTTATGCGGAAGTAGAC
GGTGATAAGCCGATCTATATTCATGATGACGGCAAAGAAAATGCCACATGA
TGCACCACACTCGGTAGCAACAATTGCACGCTTAAACAATGAAGCTAAAA
CACATCGTGAAGCCAAAGAAGCAGCCGAAAAAGCATTAAAAGCTTTTGAA
GGAATTGAAGACCCAGCGGCAGCTAAAAAGGCATTACAAACAATCCAAAA
TCTCGATGATAAAAAGCTGGTGGATGCCGGTGAAGTTGAGAAAAGTTAAAG
CTGAAGCTATCAAAGCAGTTGAGGAAAAATATGCCCCGATTGTTGCGCAA
CGTGATGCTCTAGAAGCCTCTTTACATAAAGAAGCTTATCGGCGGTGGTTTT
GCTCGTTCTAAGTACATTCAAGACAACATTGCAGTACCTGTGGACATGGTT
CAGGCAACCTTTGGTCATCACTTCAAAATCGAAGAAGGCAAGGTGGTTGC

<i>group_2537</i>	Beta-lactamase OXA-133	A2, A3, A14	M	Penicillin binding protein transpeptidase domain	<p>ATATGATCCGAACGGCGAAAAGATTTATTCACGTGTCCGCCCGGGTGAAC TTGCAAATGTTGATGAAGCTTTAGAGTCATTGGTTGGTGGATACCAGCATA AAGACTTAATTCTTAAAGGTGGTAAAGGAACTGGTGGCGGTTTTCAAGGT GGGGCAAAGGTGGAGCACCTACTGGAATGAAACGCAGTGAAATGTCTG TTTCTCAGAAAGCAGATTACATCAAAGAACATGGCAATGATGCCTTCCTAA AACTACCGAACTAA ATGAATAAATATTTTACTTGCTATGTGGTTGCTTCTCTTTTTCTTTCTGGTT GTACGGTTCAGCATAATTTAATAAATGAAACCCCGAGTCAGATTGTTCAAG GACATAATCAGGTGATTCATCAATACTTTGATGAAAAAACACCTCAGGTG TGCTGGTTATTCAAACAGATAAAAAAATTAATCTATATGGTAATGCTCTAAG CCGCGCAAATACAGAATATGTGCCAGCCTCTACATTTAAAATGTTGAATGC CCTGATCGGATTGGAGAACCAGAAAACGGATATTAATGAAATATTTAAATG GAAGGGCGAGAAAAGGTCATTTACCGCTTGGGAAAAAGACATGACACTAG GAGAAGCCATGAAGCTTTCTGCAGTCCCAGTCTATCAGGAACTTGCGCGA CGTATCGGTCTTGATCTCATGCAAAAAGAAGTAAAACGTATTGGTTTCGGT AATGCTGAAATTGGACAGCAGGTTGATAATTTCTGGTTGGTAGGACCATTA AAGGTTACGCCTATTCAAGAGGTAGAGTTTGTTCCTCAATTAGCACATACA CAGCTTCCATTTAGTGAAAAAGTGCAGGCTAATGTAAAAAATATGCTTCTT TTAGAAGAGAGTAATGGCTACAAAATTTTTGGAAAGACTGGTTGGGCAATG GATATAAAACCACAAGTGGGCTGGTTGACCGGCTGGGTTGAGCAGCCAG ATGGAAAAATTGTCGCTTTTGCATTAATATGGAAATGCGGTCAGAAATGC CGGCATCTATACGTAATGAATTATTGATGAAATCATTAAACAGCTGAATAT TATTTAA ATGTCTAATATAGCACAAATCAACGATACCAAATATCAATTGTTAACTTCA AATCTGTTCCAGTTGTTACTACAGCAATGCTTGCTGATTTCTATGGAACCG ATACAGACAACATCAAACAAAACCTATTCTCGAAATAAAGAGCGGTTTGTAG AAGGTAAACACTTCTTCAAATTTATTGGTGAAGAATTGAAAAATTTGTAGG TGACTTAAAGTCACTTGCAAATTTCCCTGCAATTTCAAATAAACTCGATCC CTTATCTTATGGACAGAACGCGGTGCTGCACGTATGCCAAGATGTTAGA CACAGACCAAGCATGGGAAGTTTTCGAGCAACTTGAGGATTGCTATTTTGT CCGTAAAGAGATTTTAGCCAAAACCCACAAATCAGAACGTGAACCCCTAA CCAATGCTGTAAATCTTCTTGTAGCTAAAATAAGCATTTGAATTACAGCG ATGCTTATAAATTAGTTCATCAGCGTTTCAATGTTTCAGCATATTGATGAAAT</p>
<i>group_2575</i>	hypothetical protein	A2, A3, A14	K	ORF6N domain	

group_
2587

hypothetical
protein

A2, A3,
A14

/

hypothetical protein

TCCATACGATGTAATACCTGTGGCTGTGGAGTATGTTCCACCACTTAATTGC
TATGTACAGCAAGGCTGAAAAACAAGGTTCTTTATTTGATGAAGATCAATT
TAAGCTGCTCAAGAACCTAATTGATGCAATTATTTCCCAAACCTTTGCGAC
TAGTCGAATCTATCGAGCAGTACATATGCTTAACAACGAGCAAGGACACTA
CTTAGCTGAATATGCTTTTAAACTAATATTGCAGTTCTAAACTTACTCGG
GCAATGGATTTAAGAGGGCCACTTAATAGAAAAATCATTAGTGATGATTTA
AAAACCATAAGCTACACAACAGGCAATCAACATTATAGCGACCGTTGGTTT
CATCCATTGATGGAATCGGGAATGCTAGCTGGTGCTTTGCGAATTTCTGG
TGGTTGGTAG

ATGAAGATTATCTATACACGCATTGCAGCAGCGGCTGCATTAGAGACAGG
CATTATTGCTAACCCCTGACTATTATGAAAACCCAAATTTGAAAGCAAAAGA
GGTAATTATTTACGGTAATTATCCAAAGATTCAAAGGATTATGAATCTTTG
GAAGTCCAGTTGAAGTTCGTAAGTTGGAAGTGCCACAAAAACGACTTT
GGCCACAGTAAATGTCGCAGTGGGAATTACCCCTGAACTTCAAGCTGTGA
TGGATGATGCAAAAGCTGAATGTGAAAAGGTAGTTGAAGAAAACACTCAG
CTTAAGCAGAAAATTGCCATCTTAGAGCAGGCCGGTGGTAACCAGTCAGA
GTTGTTATCTGAAAATTCACGATTAAGATGCAGCAGTCTTAGCAGATAA
AGCTCTCAAAGATGCTGAAGCTCAAGTGGTCCGTATAAAAACTGAATTTGA
AGCTTTTAAAAATGATATTCCTGCAATGCAGGCACGTATTGCTGAATTGGA
AGCTGGAAAAGCGGCAGAAAACCCAGCTACAGAAACGGCAGCTAATGATT
TTGAAAACCTGGTCAAATGATCAATTAAGAGTATTTGGCTAGTAAAAACAT
TGGTTACAAGCCGTCTGCAACAAAAGCAGAACTTCTTAAATTAATCCCGAA
GGAATAA

group_
26

tnsB_1

Transposon
Tn7
transposition
protein
TnsB

L6, L7,
L8

L

Mu transposase, C-terminal

ATGATTGAATCTAATGCTGTCAGTATTGATCGTAGTCGTATCACACTTAAA
CCCAATGTGGTTGTTGGTTATGAGGGAAACCCTTATAAAATTGTGAATGTT
TTAAATGCTAATGACATTGTTATATCAAGTCTTGATTCCGGTACGTAGTTTAC
AGGTCAATTCTAAAAGTTTTGACCGTTTTTGGAGGGAGAAAATACGCTAACTG
AAAACCTAAACAAAGGTGATAAGGATATCAGCAATGAAGCTTGGCAGATT
GCTTTACAACGTTATGAGATTATTAAGCCTTTGATTGAGTATTCAACCACTG
AACTGGTTGAAAATAGAGCGAATGAATATGATGTAAATCGTTCTACATTGT
GGAAGTGGCTTAGAGATTATCGTGAAAATAACTCTCTAATGACGTTGATTC
CCAAAAACGAGGTTGGACAACCTGAAAATCAAGGTTGTCACCTCAAGTG
AGCAATATCATTAAACAAGCTATCAATGATGAATATTTGAATGCGAAGAAA

group_
2626

hypoth
etical A2, A14 /
protein

hypothetical protein

CCTAGTATTTCCAAAACGATTGAAATTGTAAAAGCAGAGTGTTACGGCTT
CAGCTTGAAGCACCATGAAAACCTATTTCGTAGAAGAATCGAGGCTTTA
AATAATTACCAAGTAACGAAAGCACGTCTTGGATCTAAAGCTGCAATAGAT
AAATTTAAGGCTGCTGCGGGTTCATTCCCAAATGCTGATTACCCTTTGGCA
TATGTGCAGATTGATCATAACCTCTAGATATTGAAATCGTGGATGACGAG
TTCAGAGAAGCGATTGGAAAACCTCACTTAACCTTGGCAATCGATGTTTTT
AGTCGGATGATTGTTGGATATTACCTATCACTAGAAGCACCAAGTACGACT
TCAGTGGCGATGTGTATAGCATCTTGTATTTTGTCAAAAAACGTAAGTTG
ATCGAACTAAATATTGATGCTGAATGGCAGGTGCAAGGTATTATGGACTC
GGTGCATACAGATAATGGACCCGATTTTAGAACCAATCACATCAGTAAAGC
TTGCCTAAAATATGGGATTCATTGGGAATATAGACCAATTGGTGGGGCGA
GATTTGGCGGTCATATTGAACGTATGCTTGGTATTGTCAATCTTGAATGC
ATGTGCTGGATGGTACAACTTTTTCAAATGTGCAGCAGCGAGGAACTTAT
GATTCCGCTAAACAAGCCTGTATGACATTAAGGAACCTGAATACTATATC
GTTTATTGGATTACTAAAGTTTATCATCAGAAGAAGCATTACAGCTTTAGGAA
CATCACCTATTGTTAAATGGGAAGAAGGTGTTTGGGGAACGAAGACGACA
GCAGGAACAGGTTTGAAGAACGTGTATCAGATGAAGATACCTGTTTTATT
GATTTTCTACCTGAATTTGAAGCGACCATTACAGCGTACAGGGGTGCAGAA
AGATAACTTATTCTATTTTGCAGACTGTTTAAAGCAATGGGTTAATTCTATT
GATCCAGACGACAATAATCGAAAGCGCAAAAGAAAATTCTATTTAAGCGT
GATCCGAGGGATATTTACAGATTTGGTTTTATGAGCCATTTTCTAATACC
TATTTTAAAGTACCCACTGCAAAGCGAGAGATACCACCGATTAGCTTATTT
GAATATAGACAGGTACAGAACTATCTCAAGAATGGGCGTCAGGATGTTCA
GAATCAGGATGAGATTTACAAGGCGATTTTACATCTACGTGAGCAGCTTAA
TCAGGCTCGAAGCTTGACTCGAAAGCAACGTGCTCAAACCAAAGAAAA
AGGAAAATGAAAAAGCCATTACTCAGCTATCCGAACAGAATCAGTCTAAAA
AGGCTGTAGTTTCTGAAAGCCTTCAGACAAGTGATGATTTGTGGAGTACTC
CTTTAACTGCATTTGATGATTTAAGGTGA
ATGAAAGCATCTAAATTGATTAGAGATAAAGGACTGCAATACGCGAAGGAA
ATCGTAGATTCAGCACCCGATAACGCAACTGAATGGAACGAGGGTTATGA
GTTCCAATGTGGTCAAAGTGTAGAAATCAGCCCAGCAGATCGTGAGAAGT
ATTTTGTAGATTTGGTTGAGCTTAAACGTCTGGTGGAGTCTTTGAAAATCA
TCAGCGATTTAGGTGGAGTTGAGAAGCTAACGCCTGCATTCATTACGACA

group_
2627

hypothetical
protein

A3

/

hypothetical protein

GATAAGCATGTTGGTTACACGCATGTTCCGCATGGTGGGAAATGGGAGATT
GAGCTTTCTTGATGATTTTTGCGACTTCATTCCAGATGGTCCATTTCAATT
AAGCGTGTGATGACTGCTATCCGCGACCACGAATCAATATACGGAGGCG
GTGAATCTCATGCCAACTAG
ATGAAAGCATCTAAATTGATTAGAGATAAAGGACTGCAATACGCGAAGGAA
ATCGTAGATTCAGCACCCGATAACGCAACTGAATGGAACGAGGGTTATGA
GTTCCAATGTGGTCAAAGTGTAGAAATCAGCCCAGCAGATCGTGAGAAGT
ATTTTGTAGATTTGGTTGAGCTTAAACGTCTGGTGGAGTCTTTGAAAATCA
TCAGCGATTTAGGTGGAGTTGAGAAGCTAACGCCTGCATTCATTACGACA
GATAAGCATGTTGGTTACACGCATGTTCCGCATGGTGGGAAATGGGAGATT
GAGCTTTCTTGATGATTTTTGCGACTTCATTCCAGATGGTCCATTTCAATT
AAGCGTGTGATGACTGCTATCCGCGACCACGAATCAATATACGGAGGCG
GTGAATCTCATGCCAACTAG
ATGGACAATACCACGGCTCATATTGTTGCGATGACAGGTTCCATACTTCCGT
GGCGATAGCGTTCCTGTACTTACGCCTGAAGATGAAGCGAAGTTTGATAA
AGTTTCTTTAACTATTACGAACAACCTCAATGGGTATAGCTATTTAAAGTCA
TTAGGCATTGGTTATCACTTCTATAAAGGCCGCTACTTTGCAGAAGATGAT
GAAGGGGTTATCGCAATTGCTGAAGTTTTAGATACCGATAAAAAACCATA
ATTCACATTCCAAATGTAATTCAGGTGAATCAACCAAAGATAAATATGATG
AAGTGGACGCGATCCTCGAAATCATTGGTGAATTTGTCGCTAAAGATCCT
CATACTGGAATTATCACCGTCAAGCGTGCAGATGGTAAGATCTTAAAAGTA
GCAGACCTCGTTGAAGAAGATGGTTCGTGATAAAGTTGTTGAATATCTACGT
AACATTTAAACCGTAGATGATCTTGATCTCATCATTGCCCTAGGCATGGCG
AAAGAAGGCTTTGACTGGCAATTCTGTGAACATGCTTTAACGATTGGTTAT
CGTGGCTCTTTGACTGAAATTGTGCAGATTGTGGGACGTTGTACACGAGA
TAGCTCAAACAAATCACATGCGCAGTTTACCAATCTTATTGCTCAACCTGA
TGCACAGGATACTGAAGTTAAATTTGCAGTAAACAACATGCTGAAAGCAAT
TACTGCATCTTTGTTAATGGAACAAATACTTGCACCAAACCTCAAGTTTAAA
ACCAAACGTGATGATGAAGATATTGTTCCGCCAGGAACACTAACGATTTAAA
GGTTTTAAAGAACCAAGCACGGCAAAAACAAAGGCGATTATTGAAAATGAT
CTTGTGGATCTAAAAGCAGTGATTCTTCAAGATGCTAAAGTGATTGCAGGT
TCAGCAGGTAAGGTTGAGCCTGAAGTGGTGAATAAAATCCTTATTCCCAA
GATTATCCAAACAAAATATCCTGACTTAACGAAAGCAGAAGCTTGAAGAAGT

group_
266

hypothetical
protein

A2, A3,
A14

L

helicase

<i>group_2748</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>GAGTCAGTATTTAGTCGTGGACTCAGTGATTAGCACAGCAGAAGTCAAAG AAGTTGGCGATAAAAAGTTTATTCTGAATGGCAGATAAATTTATCAATATCG AAGACCTGAATATTGATTTAATTCGAATCTGTGAATCCTTTCCAAAAAGCGTT TGAAATCCTATCAAAAAGAAGTCACAGTCAATGTGCTCAAATTTGATTAAGGA AGCAATTGAAGCAACTAAGATTAGTATGACTCAAGAAGAAGCCATTGAGCT ATGGCCTAAGTTAAAAGCCTTCAAAGTTCAGTTTGGTTCGTGAGCCTAGTCT AGAGTCAAATAACTTAATTGAAAAACGATTAGCCGAATGCCTAATCTATAT CCGTGAGCAACGTAGACGAGCTGGAGTGTA ATGCGAACATTTATATACTTGGATGAAAGTGGTGATTTAGGTTGGAATATG GAAAAGCCTTATCAAAAGGGTGGTTCCAGTCGAATGCTTACGTTAGCAGC AATCTGTTTGCCTGAGAATAAGGTTAAGTATGTTTACGCGTATTGTAAGAGC ATTATATGAAAAAGAAAAAGACCTTTAAAAAATGAATTAATAATCAGTTGAT TTGAATCTAAAAGATAAAGAAATATTCGTCAAATTGACTGCGAACTTATCA AAGACCATCCAGATATACAACCTTCGCTCAATTACAGCAAATAAAGAATTTG TTAATGCAAGATTCAAGAACGACCCAAATGCTTTCTATAATTATATGGTGAA ACTTTTACTTCTTGGGACTATCTGCAAGCATAAATATGTAGATTTTATGCCT GACAGAAGAAGTGAGCGGGTTTCGTTGAAATGGAATATGGGTGAGTATTT AAAACAGATGGTTTTAGAGTGTGGCATTGAAAACCAAATTGTTAACCAGTC ATGCAATATTATGCCAATGGATAGCTCAAAGTGCCTTGAGCTACAATTTAT AGACTTCTATGCAGGTTTAGTCTGGTCGGCATATGAATTTAAGACATGAC TGCAAGAAAATTCATGGCAGAAAACCGAAATACCAACCATAAGCTTTTCTT TCCAAAAGAAGACAAAGTGGATAACATTGTTGATGAAGCTGTCTAA ATGCCAAGCCCTATTATCCAATATTTCCAATATGAACATTTACCTGAACATT TGCAGCAAGTTAGTAAGCCAATTGGTGATTTAGCTCGGCAAATGGATGAG CAACTTCTGACGGGCTGAAAAATCCACAGGATTAAGAAAGCTACTTGA AGCAAAAAGATGCATTTGTACGCCAAGCTTTAAGTAAATAA ATGGTCTCACGTTTAACTTTATTTCTCAAAGTACGGTGGCTAAAATGACA ATGACTTTAGAGCAGACAAGGCAAGCTATTATTGATCGCATGCAAAGCTTT ACAGGTATTACGCAAGACAGAATCCAGTATCCAAATTTACCAGGCTTTAAT GTACCTAAAGATGGTGTTTGGTGCCGCTTAACGATTGCAGGTGGTCCCAG TTTTACTTCTGGCATTGCAGATAAGCCATGTACTCGCCGTACCGGTAATAT CATGATTCAATGCTTTGCACGTCCCAATTCAGGAATAATTGAAATCACAAA ATTGAGTGATGCATTAATTTGCTCATTGCTCAATTTGCAATCGAACACTTA</p>
<i>group_2780</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	
<i>group_2784</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	

<i>group_2815</i>	hypothetical protein	A2, A3, A14	S	Phage Mu protein F like protein	<p>GAATGTTTGAATGGCCAATCTATTTATGCGGGTAAAGATGCTGACTTCATT CAATACAATGTATCAATAAGTTTTTTAGTTAACTAA ATGGCAACAGATATCAAAAACTATTTGAAGTACTCACTCAGCACCAGGCC TATCTTTATCGTGCTTCATCAAAAACGGTAAATGAGTTATTGGCTTTATTCA ATGATGATACGAGCAAGATGCTATCTAAGCTTCGGGATTTATTGGATGAGC TTAATGAGTCCGAGAAAGTTGCTTTAGCTGGTGGTAAATATACAACCTCAA ATTTAAGGGAAATTAGGGATTTGATTGCCCAATGGTTTGCCAGTGTTAATT TAGCATTACCTGAAGCTTTTGCCGTTTCTGCTACGGCGCTGGCTGTTTATG AGGCCAATTACGTAGCTAAGCTCTATGGAGCAAAAATTAATAAGCCTGATG GGGAAAACTATTCTTATCCGCTAAAAAAGTTCCGTTGGCAGGTGGCGCT CTTGTCGATGATCTGCTTTCAAGAATTGCTGAAAGTGCCCGTCAAAGGTT GAGTATGCAATTCGAGATGGTATTAATTCAGGCAAACTAACCAAGAAATT GTTCCAGCGTATTCGTGGTACCAAACGGCTTAACTATGAAGATGGGATCTTA AATGGTACCAAACTGATATTGAGCGAACGGTAAGAAGTGTGCGAAGTCA TGTAGCTAATCAAGCCTATCTAAATAGCTTCAACCAAATTGGCTTTGAATAT GTCCGATTTGTTAGCGTTTTAGATGGACGAACCTTCTAAGCTTTGCGCTTCA TTAGATGGTTCAGTGTGGGAAATAAATGATCCGGCAAAGCGAGTGCCGCC GTTACATCCTAACTGTGCGCAGTATCTTGTTCCGGTCGAGAAGGACGGTC AACTTGTGGCGAACGGCCATTTGTAATGGACGAACGTAGAGTTAAAGAC ATCCCCAAAGAAGAGCGAAGCCAGTTAATAGGACAGTTAGATGCAAACAC CACATTCAAAGAGTTCTTTAAGAAAACAGATGATTTCTTTCAAAGGGAGTG GCTAGGGCCAAAGCGCTTTAAGCTCTATAAAGATGGGAAATTTGATTTTGA TAAGTTCTTTGATCCTGAAGGCCGTTTCTATAGCTTAGATGATTTGAGAAA GTTGGATGAAAAAGCTTTTTAAAAAGTTGGGTCTGTAA ATGCCAACTAGATATAACACAGGCGAGTATAGCTACGATCTTGAATATCAC TATGGAGATATGTCAGCAAGCATGGAGATGCTTAGAGCACGTTTAATTGAA TTGTTGACTCCTCATCTGTCTGACCGTTATGTGAAATGGAGAGAAGCATAT TTCACATGGTTTACAAAGTGCGGCGGGGATTCGGGGTGGATGTTTTGTGT AGGTCCACACGAATTTTATATTGATGGGGCGTTAAGGCGCTATTACTCAG GTTCTATTGATATTACCTACAACCAGAAAGATCGATATTTCTTGGTGGGTG AGAAAAAGAAAGTCAAATGTAAGGCTTGTAAAGGGTTTGGCTTCATTCTGA GATGATGGGTGGGGGCATATAGATAAATGTGAAACGTGTGATGCAGAAAA AGGAGCCAGCCATGAGTGA</p>
<i>group_282</i>	hypothetical protein	A2, A14	/	hypothetical protein	<p>GAATGTTTGAATGGCCAATCTATTTATGCGGGTAAAGATGCTGACTTCATT CAATACAATGTATCAATAAGTTTTTTAGTTAACTAA ATGGCAACAGATATCAAAAACTATTTGAAGTACTCACTCAGCACCAGGCC TATCTTTATCGTGCTTCATCAAAAACGGTAAATGAGTTATTGGCTTTATTCA ATGATGATACGAGCAAGATGCTATCTAAGCTTCGGGATTTATTGGATGAGC TTAATGAGTCCGAGAAAGTTGCTTTAGCTGGTGGTAAATATACAACCTCAA ATTTAAGGGAAATTAGGGATTTGATTGCCCAATGGTTTGCCAGTGTTAATT TAGCATTACCTGAAGCTTTTGCCGTTTCTGCTACGGCGCTGGCTGTTTATG AGGCCAATTACGTAGCTAAGCTCTATGGAGCAAAAATTAATAAGCCTGATG GGGAAAACTATTCTTATCCGCTAAAAAAGTTCCGTTGGCAGGTGGCGCT CTTGTCGATGATCTGCTTTCAAGAATTGCTGAAAGTGCCCGTCAAAGGTT GAGTATGCAATTCGAGATGGTATTAATTCAGGCAAACTAACCAAGAAATT GTTCCAGCGTATTCGTGGTACCAAACGGCTTAACTATGAAGATGGGATCTTA AATGGTACCAAACTGATATTGAGCGAACGGTAAGAAGTGTGCGAAGTCA TGTAGCTAATCAAGCCTATCTAAATAGCTTCAACCAAATTGGCTTTGAATAT GTCCGATTTGTTAGCGTTTTAGATGGACGAACCTTCTAAGCTTTGCGCTTCA TTAGATGGTTCAGTGTGGGAAATAAATGATCCGGCAAAGCGAGTGCCGCC GTTACATCCTAACTGTGCGCAGTATCTTGTTCCGGTCGAGAAGGACGGTC AACTTGTGGCGAACGGCCATTTGTAATGGACGAACGTAGAGTTAAAGAC ATCCCCAAAGAAGAGCGAAGCCAGTTAATAGGACAGTTAGATGCAAACAC CACATTCAAAGAGTTCTTTAAGAAAACAGATGATTTCTTTCAAAGGGAGTG GCTAGGGCCAAAGCGCTTTAAGCTCTATAAAGATGGGAAATTTGATTTTGA TAAGTTCTTTGATCCTGAAGGCCGTTTCTATAGCTTAGATGATTTGAGAAA GTTGGATGAAAAAGCTTTTTAAAAAGTTGGGTCTGTAA ATGCCAACTAGATATAACACAGGCGAGTATAGCTACGATCTTGAATATCAC TATGGAGATATGTCAGCAAGCATGGAGATGCTTAGAGCACGTTTAATTGAA TTGTTGACTCCTCATCTGTCTGACCGTTATGTGAAATGGAGAGAAGCATAT TTCACATGGTTTACAAAGTGCGGCGGGGATTCGGGGTGGATGTTTTGTGT AGGTCCACACGAATTTTATATTGATGGGGCGTTAAGGCGCTATTACTCAG GTTCTATTGATATTACCTACAACCAGAAAGATCGATATTTCTTGGTGGGTG AGAAAAAGAAAGTCAAATGTAAGGCTTGTAAAGGGTTTGGCTTCATTCTGA GATGATGGGTGGGGGCATATAGATAAATGTGAAACGTGTGATGCAGAAAA AGGAGCCAGCCATGAGTGA</p>

<i>group_283</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGCCAACTAGATATAACACAGGCGAGTATAGCTACGATCTTGAATATCAC TATGGAGATATGTCAGCAAGCATGGAGATGCTTAGAGCACGTTTAATTGAA TTGTTGACTCCTCATCTGTCTGGCCGTTATGTGAAATGGAGAGAAGCATAT TTCACACGTTTTACAAAGTGCGGCGGGGATTCCGGGGTGGATGTTTTGTGT AGGTCCACACGAATTTTCATATTGATGGGGCGTTAAGGCGCTATTACTCAG GTTCTATTGATATTACCTACAACCAGAAAGATCGATATTTCTTGGTGGGTG AGAAAAAGAAAGTCAAATGTAAGGCTTGTAAGGGGTTTGGCTTCATTCTGA GATGATGGGTGGGGGCATATAGATAAATGTGAAATGTGTGATGCAGAAAA AGGAGCCAGCCATGAGTGA</p> <p>ATGCCAACTAGATATAACACAGGCGAGTATAGCTACGTTCTTGAATATCAC TATGGAGATATGTCAGCAAGCATGGAGATGCTTAGAGCACGTTTAATTGAA TTGTTGACTCCTCATCTGTCTGGCCGTTATGTGAAATGGAGAGAAGCATAT TTCACATGGTTTTACAAAGTGCGGCGGGGATTCCGGGGTGGATGTTTTGTGT AGGTCCACACGAATTTTCATATTGATGGGGCGTTAAGGCGCTATTACTCAG GTTCTATTGATATTACCTACAACCAGAAAGATCGATATTTCTTGGTGGGTG AGAAAAAGAAAGTCAAATGTAAGGCTTGTAAGGGGTTTGGCTTCATTCTGA GATGATGGGTGGGGGCATATAGATAAATGTGAAATGTGTGATGCAGAAAA AGGAGCCAGCCATGAGTGA</p>
<i>group_284</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGCCAACTAGATATAACACAGGCGAGTATAGCTACGTTCTTGAATATCAC TATGGAGATATGTCAGCAAGCATGGAGATGCTTAGAGCACGTTTAATTGAA TTGTTGACTCCTCATCTGTCTGGCCGTTATGTGAAATGGAGAGAAGCATAT TTCACATGGTTTTACAAAGTGCGGCGGGGATTCCGGGGTGGATGTTTTGTGT AGGTCCACACGAATTTTCATATTGATGGGGCGTTAAGGCGCTATTACTCAG GTTCTATTGATATTACCTACAACCAGAAAGATCGATATTTCTTGGTGGGTG AGAAAAAGAAAGTCAAATGTAAGGCTTGTAAGGGGTTTGGCTTCATTCTGA GATGATGGGTGGGGGCATATAGATAAATGTGAAATGTGTGATGCAGAAAA AGGAGCCAGCCATGAGTGA</p> <p>ATGCCAACTAGATATAACACAGGCGAGTATAGCTACGTTCTTGAATATCAC TATGGAGATATGTCAGCAAGCATGGAGATGCTTAGAGCACGTTTAATTGAA TTGTTGACTCCTCATCTGTCTGGCCGTTATGTGAAATGGAGAGAAGCATAT TTCACATGGTTTTACAAAGTGCGGCGGGGATTCCGGGGTGGATGTTTTGTGT AGGTCCACACGAATTTTCATATTGATGGGGCGTTAAGGCGCTATTACTCAG GTTCTATTGATATTACCTACAACCAGAAAGATCGATATTTCTTGGTGGGTG AGAAAAAGAAAGTCAAATGTAAGGCTTGTAAGGGGTTTGGCTTCATTCTGA GATGATGGGTGGGGGCATATAGATAAATGTGAAATGTGTGATGCAGAAAA AGGAGCCAGCCATGAGTGA</p>
<i>group_285</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGCCAACTAGATATAACACAGGCGAGTATAGCTACGTTCTTGAATATCAC TATGGAGATATGTCAGCAAGCATGGAGATGCTTAGAGCACGTTTAATTGAA TTGTTGACTCCTCATCTGTCTGGCCGTTATGTGAAATGGAGAGAAGCATAT TTCACATGGTTTTACAAAGTGCGGCGGGGATTCCGGGGTGGATGTTTTGTGT AGGTCCACACGAATTTTCATATTGATGGGGCGTTAAGGCGCTATTACTCAG GTTCTATTGATATTACCTACAACCAGAAAGATCGATATTTCTTGGTGGGTG AGAAAAAGAAAGTCAAATGTAAGGCTTGTAAGGGGTTTGGCTTCATTCTGA GATGATGGGTGGGGGCATATAGATAAATGTGAAATGTGTGATGCAGAAAA AGGAGCCAGCCATGAGTGA</p>
<i>group_288</i>	hypothetical protein	C4, C6	S	Phage Terminase	<p>ATGACTGCAAACTACCAGACTGGACTACAGCTTGCCAGACTGGGCGAC CCGTATTGTTTCTAAACAATCGTTAATGCCGTGTAAGCCATTATCCCAA AGTGGCTGACGTAGCGGAGCGTATCTTTAAAGAGTTAATTCTTGTTGATGT GATGGGTAGCCCTAAGATGGGTGATGTCACATTGGAATGGGTGATCGAGT TTGTTGTCGCAATCTTTGGCGCATATGATCCAAGCACAAAGCGCAGATTAA TTCGTGAATCTTTCTTTTGAATTCGAAGAAGAATACTAATCTACGATTGC</p>

group_
2899

hypoth
etical
protein

A2, A3,
A14

/

hypothetical protein

CGCCGGCATTATGCTTACTGCATTAATTCTTAATGATCGACAATCTGCCGA
ACTAATTATTCTTGCGCCTACTAAAGAAGTTGCTGATAACTCATTTAATCCA
ATCCGGGATTTTCATACGCGCAGATGAAGAATTAAGTGAAAAGATTTAATGTA
TCTGAGCACACAAAAACAGTTACGCATCTAGGTACCGGAGCAACACTTAA
AGTTATTGCAGCAGAATCTAACGCTGCAGCTGGTAAGAAAGCTTCAATCAT
TTTGATAGATGAGGTCTGGCTATTCGGGAAACGTGCCAACGCTGAATCAA
TGTTCCGTGAAGCAAAGGGTGGTTTAGCATCTCGTCCAGAAGGTTGTGTG
ATTTATCTGTCTACCATGTCCGATGAAGTGCCATGTGGAGTATTTAAGCAG
CTTTTAGATTATGCCAGAGATGTACGTGACGGAATTAAGTTGATAAAAGT
TTTCTACCACTTATTTATGAATCCCTAAGCATCTTGTAAGAAGCAGGCCGAA
CATTTAAACCTGAAAATTTCTACATCACAAACCCAACTTGGGTGCTTCG
GTTGATCTTGAATATCTGATTTCCGGAATTTAACAAAGTTAAGATGCTAGTG
AAGAATCTCTTAGAGACTTCTTGGCCAAACACTTAAACATTGAAATCGGCA
TGAACCTTCGTGCTAACCGGTGGGCGGGTGCAGAGTATTGGAATGCTCAA
GCTAAAGATATCCAAATCGACCAACTAATTGAGCTATCCGATGTCATTACT
TTGGGTATTGATGGCGGTGGTCTCGACGACTTACTTGGCTTCGCTGCTTT
AGGTCGTTTAACAGAAGATCCTCGTATCTGGTGGCTATGGAATCATGCAT
GGGCAAATAAGATTGCTTTAGAGCGCAGAAAAGAGAATGTGCCTAAGTAT
GAAGACTTCAAGTCTGAGGGTTCTCTAACTGTTGTTGACCGAATAGGCCGA
TGACATTGACCAACTCGCAGCAATTGCTAAGAAGTTTTATGACAGTGGAA
AGCTTAATAAGATCGGACTAGATCCATTGGGCTTAGGCGGTCTTTTAGATG
GCTTACTTGAGGCAGGAATCCAGAGGAAAGCATGTTTGCTGTGCCACAA
GGCTACAAACTCATGTCCTACATCCTTACTACTGAGCGCAAATTGGCAGAA
GGCAATCTGCACCATGCTGGACAACAGCTAATGACTTGGGCGGCAGGTA
ATGCCCGTGTCTGATGGTCCGCAATGGTATGCGAATAACCAAGCAAGAA
TCAGGTGTTGGGAAGATTGACCCATTGATTGCCACATTTAACGCAGTTGCT
TTGATGTCAAGCAATCCTGAGCCTGCCAATCGCGTTGATATTGACGAATAC
TTAGAGGATGTCGTGATAGCATGA
ATGGAAGTAAAGAATAATGTTGCTTGTGTTTGCCTGAAAAAGCAGGCTTAACG
GTTTATGAGCTATCAAAGCGGTGTGGTTTTGTTAGTGGTAGCAGAGTTCTA
TCAAACCTATGTGACAAGAGCCGAGCAGGGACATTCTGTCAAGATCGATAC
AGCCTTACTTATATATAAAGAAGTCAAAAAAGTAGGTGTATGTAAAAATTTT
GAGGATGTATTTTGGCTTGACCACATGGACTAG

<i>group_3</i>	hypothetical protein	L7, L8, L9, L10	I	3-oxoadipate CoA-transferase activity	<p>ATGATTGACAAAAGTAAGTCCTCACTAAGCGAGGTTTTATCGCAGATTA GATGGTGCAACCATTCTGATTGGTGGTTTTGGTACCGCAGGACAACCCGC TGAACCTATTGACGGTCTGATTGAACTTGGCGTTAAGGACTTGACCATTGT GAGCAACAATGCCGGTAATGGGGATTATGGTCTGGCTAACTGCTCAAAG CCGGTTCANNNNNNNNNCCACGTCAGTCTGACTCTTATGTGTTTGATG AACTGTACCGTGCCGAAAGGTTGAACTTGAAGTGGTACCACAAGGCAAT CTGGCGTGCCGTATTCAGGCAGCAGGTATGGGCCTTGGGGCGGTGTTTA CACCAACAGGCTTTGGAACACTTTTAG</p>
<i>group_3041</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGACTGAATTTCAAAAAATTACGAATGAGATTAGACAGCTTCAAATAGAG CTAAACCATTTGGGAAGTTGCAATACAAAAGGTTTAAATACAGAACAGATC GCTCACCTAGATGAGCGATTTTTTTTTGGCCATAGCAAAGCAACATAAATTA ATTGCTCGTCTCAACAGTAAGCCAGAGGGCTTTTTATAA</p>
<i>group_3063</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGAGCAATTTTGTTTTTAAGCGTGGTGACACATTCAACTTGAATTTGCAG CTGGTTGATATGGATGAAACCCTGCAGTATCCACCGGATGATGTTCTGTCG TGCAATTGATCTTACAGGTTATACCTTCACTTCACAGGTTAAAGCTCTGGC TGATGGCGCTGCTGTGGCTACCTTGACTTGCAGCAGCATTAAACCAGAGCA CACAGAAGGGATGGCTTAACGTTAAATCAGGTGCAAGTACAGCAGCTTGG CCTTTAGGTCTGTGCCAGATGGATATTAAGGCTGTTCGTGAATGGAGTTAC CCAGCATAACAGATACTTTGATTTTCCAAGTGATTGATGGGGTAACAGCATA A</p>
<i>group_3095</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGAGCTTTATTACTGTAGATGACGCAAATTCAATTTTGGGCAGCGATTTT GCACCAGACAGTGATAAGGCTCGTCTGGTGAAGCTGGCAAATGTGTGGAT GAAAAACAGAATTGGTTTTGTACCAGATCCAATTGATCCACTTCTTAAGGA CGCGGCTTGTGAAATTATCAAAGGAATTCTGGCCAAAGTAATTTATAACGG CAAAGAGCAGCAGTTGAAGCGTAAGAAAGTTAAGGCTGATTCTGTTGAGT CAGAAAAAGAATTTCAAGACGGATCTGAAGCAATCTCTAGCTTTGAACAGA TAGCAATTGATTTTATTGATTCACCTTGATTTGAAAGATCCAAATGCAAGTTT TAATGGCTTTGGCATACCACTTTACAGGGCATGA</p>
<i>group_3141</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGCTTAAAAAGTCATTACATGATCAAATTAATAAATTGGATTTTGGACCG TTGGTGGAGTTTTTTGGTATTTAGTTATAGCTTTTTTTCTTAAAAGTAAATAT CCAATTTTTGATTATAGCTTTAATCTAGAAATTGCATACGACGTCATAAAAG ATGCTTTAACTCTTGCAGCAAGCTTTTTAGCTCCAGTTGCAGCATTGTTTC TGTTTAGCGATTGGAGAGTTCAACATAAAGCTCTAAAAAATGAAAAGTTAA</p>

group_
3166

hypothetical
protein

A2, A3,
A14

/

hypothetical protein

GTGAAGATATCTTAAGAATCCTTAACACGGAATTATTATCCTTTTATAATTT
TAATCCTCGATCAAAATCAGATGTTGAAGATTTTAATAATCATCAAATGCAA
TTTCATAGGAACGTTGCAAATATTTATGTGATGTTAGATGAAATTGATGCAA
ATGAAGTGCAGGCAAACCACTTCATTGAAAATATTAATAAATAAGAGTTG
ATCTAGATGGTTTATACATGAGTATTTTTAAACAAATTGAAATTGTTATTGA
ACATGATGCGATTTCTGATTTTCTAGATACTCATTCAATGCGTAAAAAAGAA
ATATTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
AAATTTAATTAAGTAATTTTACAATTGAAACCTTTAAAGTTTAG
ATGGATCAGATCAGACCATTTCTCCAATTGATTTTATGGATCAGGCAGAA
GAAGAGGAAGCAATTCGTTTAATACCCGCTCCAGACCTAAAGAAATGGGT
TGTGGCCAACACTTAACGATAGGTGGACCTCTTTATAACCCTGACCATGA
CCATATTGCTGAGCTGCTTCACGATAATGAAGAATTTTAGCATTGCTTG
GGCCTCTTCTGCATATAAAAGTAAGCAAGCTATGGTGTAGGTCAGTGCG
AAAAAGTCATGTTCAATGTTGGTGGATGGCGTAAGGCCAGACAAGAGCAA
CAGATGCGTGATTGGTTTGGTTTTGTACCTACTTATTTAATAACTGTGCGAC
GCTTCTTCTGTGAGCGTGCAAACGATACAGAGTTCTGTTACTTACTTGAA
CATGAGCTTTACCACATTGGAGTGATGAGAGACGAGGACGGAGAAATTGT
TTATAGCGATAGTTCTGGTCTTCTAAGCACTATCTTGCTGGTCATGACGT
TGAAGAGTTTATTGGCGTAGTTAAACGTTATGGACCAAGCAAAAATGTTAA
GCGACTTATTGAAGTCGCAAAAATCCGCCGTTTGTTCGAATCTTGATAT
TTCAAAATGCTGCGGCAACTGTGTAATCAATTGA

ATGGCTCTACCGATTACTGCTGACCAAACCTTTATTAGTTCAAGCAATTA
TTGTGTACCTATACGCTGATCCGGGTTTAGGTAATCATCGATGGGCTTTA
CTGCGGAAAAAGCAATTTCTTTTGACTTTGACCGTGGTGCTCACCGTACTG
GTGAATTACGTCGAGGTGCGGTTGTACAGGTTCAACAATGGAGTGATGTT
GCAAACCTTACGCCGAGGACTTAGCACCTATAAAACCGTAGTCATTGA
TACCGTGGGTGCAATGCTTGAATGCATTAACCCACCTGTTACTTACGG
CAAATAACCGTCAAAAAGATGGTTCTTTAAAGTTAAAGGCTCAAGGATTAG
CGAACCAACGTTCAAGCAATACATCAATACTTTGATCAGTTTAGGTAAG
ACGTTGTTTTTATTGCACACGCTTCAGAAGATCAAAACGGTGATCAAATTA
TTTACCGACCAGATCTAGGTGGTAAAAACCGTAACGAGCTTTACCGTATC
GCAGATGTGATGGGTTATCTAACAACCTGTTACTACAGGTGAAGGTAATAA
GCCGCGTTATTAATTTAAACCCTCGCCTACACATCATGCGAAAACTCA

group_
3193

hypothetical
protein

L6, L7,
L8

S

AAA domain

group_3195

hypothetical protein

A3

S

AAA domain

GGTGCTTTAGGTGGTGAAACTGGTGAAGTATGGGTACCAGATCTTAAAGC
ACATCCTACTTTCTTGGCTGACCTGATTACTCAAGCTAAAGATCACATTAA
CACCTTAACGCCTGCACAACCTTGCAGCAGCTAAAGCCCAAGAAGAGCTAG
AAAAGTGGAAACAAAGCTGTGAAGAAGCTGAGCATGCAGGTGACCTTAAT
CAATTAAGTGAAGTCGCTTGATAAAGAACACATGTATTACCAGAACATGCGC
CAAGCAATGTTAATGAGAGCTAAAGCATTGAATTGCACGTTTGATAAACAA
CGTGGCACTTGGATTAGTCCACCAGAATTTAACGGTATCTCAGATCAACAA
AGAGATGAACTTCAAACTTTATTGCTGAACGTGGCCTCGATGTAAAAACA
GTTTGTGAGCACTTCGGCATAGATGCCCTGATCCAAATTGAAGCGGCAAA
ACTAACTGCAGTTAAACAAGAAATTGAAACCTTAGCGAAAACGGGGATGA
CAGCATGA

ATGGCTCTACCGATTACTGCTGACCAAACCTTTATTGGTTCAAGCAATT
ATTGTGTACCTATACGCGGATCCGGGTTTAGGTAAATCATCGATGGGCTTT
ACTGCGGAAAAAGCAATTTCTTTGACTTTGACCGTGGTGCTCACCGTACT
GGTGAATTACGTCGTGGTGCAGTTGTACAGGTTCAACAATGGAGTGATGT
TGCAAACCTTACTCCGCAGGACTTAGCACCATATAAACTGTTGTCATTGA
TACCGTGGGTGCAATGCTTGAATGCATTAACCCACCTGTTACTTACGG
CAAATAACCGTCAAAAAGATGGTTCTTTAAAGTTAAAGGCTCAAGGTTTAG
CGAACCAAACGTTCAAGCAATACATCAATACTTTGATCAGTTTAGGTAAAG
ATGTTGTTTTTCATTGCACACGCATCAGAAGATCAAAACGGTGATCAAATTA
TTTACCGACCAGATCTAGGTGGTAAAAACCGTAACGAGCTTTACCGTATC
GCAGATGTCATGGGTTATCTAACAACCTGTTACTACTGGTGAAGGTAAAAAT
GCCCGCGTTATTAATTTCAAACCTTCGCCTACACATCATGCGAAAACTCA
GGTGCTTTAGGCGGTGAAACCGGTGAAGTATGGGTACCTGATCTTAAAGC
ACACCCTACTTTCTTGGCTGACCTGATTACTCAAGCTAAAGATCACATTAA
CACCTTAACGCCTGCACAACCTTGCAGCAGCTAAAGCCCAAGAAGAGCTAG
AAAAGTGGAAACAAAGCTGTGAGGAAGCAGAGCATGCAGGTGACCTTAAT
CAATTAAGTGAAGTCGCTTGATAAAGAACACATGTATTACCAGAACATGCGT
CAAGCAATGTTAATGAGGGCTAAAGCATTGAATTGCACGTTTGATAAGCAA
CGTGGCACTTGGATTAGTCCACCAGAATTTAACGGTATCTCAGATCAACAA
AGAGACGAACTTCAAACTTTATTGCTGAACGTGGCCTAGACGTAAAAACA
GTATGTGAGCACTTAGGTATCGATGCCCTTATTCAAATTGAAGCAGCAAAA

<i>group_3226</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>CTTAAGGCAGTTAAACAAGACATTGAAACATTAGCTAAAACGGGGATGACAGCATGA</p> <p>ATGGAACCAGTTTCCACAAGCGGTTTAACAGCAATTTTAAAATTTTATGGA</p> <p>GCGGCAATCATGGTGACTTTAGCGGTTCGCTTTAGTTGCAGCAGTAGTATT</p> <p>AATGACACGTATGCCACGCTCACCACAAGAGTGGGCAGTGGGCTTGATCT</p> <p>GTACGGTTGTATCAAGTTTGGCTGGCGGCTCATTCAATTATTGTGAAGTGG</p> <p>GGGCTTCATGAATGGGTTACTGATGTATGGGGGATGATTGCTCTAGGTGG</p> <p>GTTCTTCTTTGTTTGTGGATTACCCGGTTGGGCTTTGGTCCGTTGGATTTT</p> <p>TAACTTCATAGATAAACAGGAAGGTAAAACGATTGTGGAAGTGATCAAAGA</p> <p>GTTTAAGAAAGCCAGAAAAGACATTGAAAACAGTTAA</p> <p>ATGGAACCAGTTTCCACAAGCGGTTTAACAGCAATTTTAAAATTTTATGGA</p> <p>GCGGCAATCATGGTGACTTTAGCGGTTCGCTTTAGTTGCAGCAGTAGTATT</p> <p>AATGACACGTATGCCACGCTCACCACAAGAGTGGGCAGTGGGCTTGATCT</p> <p>GTACGGTTGTATCAAGTTTGGCTGGCGGCTCATTCAATTATTGTGAAGTGG</p> <p>GGGCTTCATGAATGGGTTACTGATGTATGGGGGATGATTGCTCTAGGTGG</p> <p>GTTCTTCTTTGTTTGTGGATTACCCGGTTGGGCTTTGGTCCGTTGGATTTT</p> <p>TAACTTCATAGATAAACAGGAAGGTAAAACGATTGTGGAAGTGATCAAAGA</p> <p>GTTTAAGAAAGCCAGAAAAGACATTGAAAACAGTTAA</p> <p>ATGAAACGTAAAAGCTCCTCCTTTCAACTCATTGTCATTTGTCAGCATGGCT</p> <p>GCTCTTGCAATCTCTGGTGGTCTTTAGTTGCTTGCCAATTGCAACCAGCT</p> <p>TTCCAAACAAAAGAAGCCCCTTCTCTATTTACCCCTAAGACTCAACCAAGT</p> <p>ACTTACGGGGTTTTAACCGCAAAAATCACAGGTAAACATTCTGGCGTTGCT</p> <p>GTAATTAATTAGATAGCTTCCGTTTAAACGTTAGCTTTGATTTTGAAGCTC</p> <p>ATCCAGACAGCTACGGCGTTCCGGGTTCTGAATTTACCGCTGTTGATATTA</p> <p>CTCAACTCACGGTAAATGAAATTAAGTACATTAACGGTAAGTCATATAACG</p> <p>ATTTACCGAATTTGAAGACATCCGCAACATCAATGACCTTCTAAAAGGCT</p> <p>TCATCGAACGTAACAAGTTGGTGGAGGCTTAA</p> <p>ATGGGGTTCCCTATAGATGAAGGTGTTCTAGGATTTCAAAAATATTACGAT</p> <p>GCTGCTGCCGTAAGCCCAGCATACAAAATCTTTAGATTAAAAACGAGAAGTT</p> <p>AATGTAGAATATTTGGATTTGATTTGAGATCTAATTCTCTAAGAAAATAT</p> <p>ACAAAAGTAAAATGCAAGGCAGTGTAGAGAGACGACGCAGTATTCCTGAT</p> <p>GAAATGTTTTTGAATATTGAGATCCCGAATCCTCCTGAAGAGGTTAAAGAT</p> <p>CAAATAGTAAAACAACATAAACTAATAAAGGAAATTGAGAATAGTCTCAAG</p>
<i>group_3228</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGAAACGTAAAAGCTCCTCCTTTCAACTCATTGTCATTTGTCAGCATGGCT</p> <p>GCTCTTGCAATCTCTGGTGGTCTTTAGTTGCTTGCCAATTGCAACCAGCT</p> <p>TTCCAAACAAAAGAAGCCCCTTCTCTATTTACCCCTAAGACTCAACCAAGT</p> <p>ACTTACGGGGTTTTAACCGCAAAAATCACAGGTAAACATTCTGGCGTTGCT</p> <p>GTAATTAATTAGATAGCTTCCGTTTAAACGTTAGCTTTGATTTTGAAGCTC</p> <p>ATCCAGACAGCTACGGCGTTCCGGGTTCTGAATTTACCGCTGTTGATATTA</p> <p>CTCAACTCACGGTAAATGAAATTAAGTACATTAACGGTAAGTCATATAACG</p> <p>ATTTACCGAATTTGAAGACATCCGCAACATCAATGACCTTCTAAAAGGCT</p> <p>TCATCGAACGTAACAAGTTGGTGGAGGCTTAA</p> <p>ATGGGGTTCCCTATAGATGAAGGTGTTCTAGGATTTCAAAAATATTACGAT</p> <p>GCTGCTGCCGTAAGCCCAGCATACAAAATCTTTAGATTAAAAACGAGAAGTT</p> <p>AATGTAGAATATTTGGATTTGATTTGAGATCTAATTCTCTAAGAAAATAT</p> <p>ACAAAAGTAAAATGCAAGGCAGTGTAGAGAGACGACGCAGTATTCCTGAT</p> <p>GAAATGTTTTTGAATATTGAGATCCCGAATCCTCCTGAAGAGGTTAAAGAT</p> <p>CAAATAGTAAAACAACATAAACTAATAAAGGAAATTGAGAATAGTCTCAAG</p>
<i>group_323</i>	hypothetical protein	A3	/	hypothetical protein	<p>ATGAAACGTAAAAGCTCCTCCTTTCAACTCATTGTCATTTGTCAGCATGGCT</p> <p>GCTCTTGCAATCTCTGGTGGTCTTTAGTTGCTTGCCAATTGCAACCAGCT</p> <p>TTCCAAACAAAAGAAGCCCCTTCTCTATTTACCCCTAAGACTCAACCAAGT</p> <p>ACTTACGGGGTTTTAACCGCAAAAATCACAGGTAAACATTCTGGCGTTGCT</p> <p>GTAATTAATTAGATAGCTTCCGTTTAAACGTTAGCTTTGATTTTGAAGCTC</p> <p>ATCCAGACAGCTACGGCGTTCCGGGTTCTGAATTTACCGCTGTTGATATTA</p> <p>CTCAACTCACGGTAAATGAAATTAAGTACATTAACGGTAAGTCATATAACG</p> <p>ATTTACCGAATTTGAAGACATCCGCAACATCAATGACCTTCTAAAAGGCT</p> <p>TCATCGAACGTAACAAGTTGGTGGAGGCTTAA</p> <p>ATGGGGTTCCCTATAGATGAAGGTGTTCTAGGATTTCAAAAATATTACGAT</p> <p>GCTGCTGCCGTAAGCCCAGCATACAAAATCTTTAGATTAAAAACGAGAAGTT</p> <p>AATGTAGAATATTTGGATTTGATTTGAGATCTAATTCTCTAAGAAAATAT</p> <p>ACAAAAGTAAAATGCAAGGCAGTGTAGAGAGACGACGCAGTATTCCTGAT</p> <p>GAAATGTTTTTGAATATTGAGATCCCGAATCCTCCTGAAGAGGTTAAAGAT</p> <p>CAAATAGTAAAACAACATAAACTAATAAAGGAAATTGAGAATAGTCTCAAG</p>
<i>group_324</i>	hypothetical protein	A2, A3	V	N-6 DNA Methylase	<p>ATGAAACGTAAAAGCTCCTCCTTTCAACTCATTGTCATTTGTCAGCATGGCT</p> <p>GCTCTTGCAATCTCTGGTGGTCTTTAGTTGCTTGCCAATTGCAACCAGCT</p> <p>TTCCAAACAAAAGAAGCCCCTTCTCTATTTACCCCTAAGACTCAACCAAGT</p> <p>ACTTACGGGGTTTTAACCGCAAAAATCACAGGTAAACATTCTGGCGTTGCT</p> <p>GTAATTAATTAGATAGCTTCCGTTTAAACGTTAGCTTTGATTTTGAAGCTC</p> <p>ATCCAGACAGCTACGGCGTTCCGGGTTCTGAATTTACCGCTGTTGATATTA</p> <p>CTCAACTCACGGTAAATGAAATTAAGTACATTAACGGTAAGTCATATAACG</p> <p>ATTTACCGAATTTGAAGACATCCGCAACATCAATGACCTTCTAAAAGGCT</p> <p>TCATCGAACGTAACAAGTTGGTGGAGGCTTAA</p> <p>ATGGGGTTCCCTATAGATGAAGGTGTTCTAGGATTTCAAAAATATTACGAT</p> <p>GCTGCTGCCGTAAGCCCAGCATACAAAATCTTTAGATTAAAAACGAGAAGTT</p> <p>AATGTAGAATATTTGGATTTGATTTGAGATCTAATTCTCTAAGAAAATAT</p> <p>ACAAAAGTAAAATGCAAGGCAGTGTAGAGAGACGACGCAGTATTCCTGAT</p> <p>GAAATGTTTTTGAATATTGAGATCCCGAATCCTCCTGAAGAGGTTAAAGAT</p> <p>CAAATAGTAAAACAACATAAACTAATAAAGGAAATTGAGAATAGTCTCAAG</p>

group_
325

hypoth
etical
protein

A2, A14

V

N-6 DNA Methylase

GAAAATCAAAAAAATTGCGTCTAAAGACAGAAGCATTATGGGAACCTTCCT
CAAAATTACAATAA
ATGTCTAATATTGAGCAAGATACACGTTTTATTGTTAACAATAATTTGATTA
ACAAGGGCTGGATCTTGGACATTCAAGATCCAAACAAAAATGTCTTTTTT
AATCAGATATCTTAAGAATTGTTAATAATGAGTTTCTCAAGAAAAGTAAAA
AAGACCCGATTATGTTCTTTTCGATTCCAAAATAAGCGGCCAATCGGTGT
AATTGAAACGAAATCAGGTGGAAAAAGCTTAACAAAAGCACTGGATCAGG
CAACCGAATATGCTGAAATGCTTGATGCACCTTTGATATTTGCAATGAATA
ATGGTTTCTGCGAAACACGGCATTGTATACCCAAAAACCATTATTTATTG
ATGAAAATGAGGTTAATGAATTAATAAGAGTAAATGAAGCTAAAGAGTTCA
TATTGCAGGAAACAAATGGTATTTATATTACACCTAAAGAAATTTTAGTCTC
TCGCAAAGAGTTAATTAATGTTTTCAAGAAGTTAATAACTCACTAAGAGGT
GAAGGTTTAAGAGCTGGTATAGAAAGGCTTTCAGAATTTGCAAACATTCTT
TTTTTAAATTGTATACAGAGAATGCTAATACAGGTATTTGGAATTCTCTCA
AAAGTCTCGATAATGATTTGCTAATTAATACAATAACATACTACAAGA
TATTGATAGACAATATGGTGCTTCTGTTTTTACAAATTTACAGCTAACCAAC
CCTGTTGCTGTTAAAGAGATGATCAAAGAGTTGGATAAGTTAAAACCTCTCA
TCAATAGATACCGATATTAAGGAGATGCTTTTGAGTATTTCTTACAGCAA
GCTACAGCAACTAATAATGACTTAGGAGAATATTTACTCCACGTCACATA
ACTAAAACCATTGTTAACTTAGTCAACCCTAAATATGGTGAAAAGATCTATG
ACCCTTTTTGTGGGACAGGTGGTTTTTTAACAGAGGCATTTGATCATATAA
AAGATAACACTTTAATTGCAAACAATAGTAGTGAAGAAATCAAGCTTAAAC
ATAATACTATTTTTGGAAGAGAAATTACCTCAAATGCAAAACTCGCAAAAAT
GAATATGATTCTGCATGGGGATGGGCATAGTGGAAATTTGCCAGATAGACA
CACTTCAAACCCTATTGAATCTGAATATGATGTGGTTATAACCAACATGC
CATTTTCTCAAAAAACTTCTTATTCTCACTTATATGAGAATAAGTTAGCTAA
AAACGATGGTGATGGAGTATGTGTTCTACATTGCTTTAAAGCAACAAAAAA
AGGAGGGCGAATGGCATTAGTAGTACCTGAAGGCTTTCTTTTTAAAGCCG
CTTTAGCTCCAGTAAGGAAGTATTTATTTGAAAACGCCCAACTAAAAGCAG
TAGTTTCACTTCCAAAAGAAGTTTTTCTGCCATATGCAAAAAGTTAAACCAA
TATACTCTACTTTACCAACTGTCATAATGGTAGAACAAATTCTGACGTTTTT
TACTACAATGTGACAAATGATGGCCTAAGTTTAGATTCTTTCCGTAGAAAA
ATTGACGAAAATGATTTAAAAAATTTAGATTTTGCTGATTTAATAAGAGCG

group_
326

hypothetical
protein

A1, A3,
A14

V

N-6 DNA Methylase

group_
327

hypothetical
protein

A2, A3,
A14

/

hypothetical protein

group_
3271

IS4
family
transposase
ISAb
33

A2, A3,
A14

L

Transposase DDE domain

ACTTTGATAAATATTATAATGAATTAGGTTTCTTAAAAGTTAATCCAGAATTA
ATCAGAAGCAATGATTATATTTATAATTATGCTCACTATAGTAATTCACATA
TAAAATCAAATTCCTCAACTATAAACTAAAAGAACTCCTATCCTTGTCTGG
CAAAGTCAAAGTGGGAGAGGATACAAATATACCTATTATGAGTATCACTAT
GGAACATGGCTTAATTGATCAGCATGAGAAATTTAAAAACGAGTCGCAAG
TTCTGATATTTCTGGGTATAAAAAGATTTTTAAAATGAACTTGTAAATGGGG
TTCCCTATAGATGAAGGTGTTCTAGGATTTCAAAAATATTACGATGCTGCT
GCCGTAAGCCCAGCATACAAATCTTTAGATTA AAAACGAGAAGTTAATGTA
GAATATTTGGATTTGATTTGAGATCTAATTCTCTAAGAAAAATATACAAA
GTAAAATGCAAGGCAGTGTAGAGAGACGACGCAGTATTCCTGATGAAATG
TTTTTGAATATTGAGATCCCGAATCCTCCTGAAGAGGTTAAAGATCAAATA
GTAAAACAACATAAACTAATAAAGGAAATTGAGAATAGTCTCAAGGAAAT
CAAAAAAATTGCGTCTAAAGACAGAAGCATTATGGGAACTTCCTCAAAAT
TACA ACTAA
ATGGGGTTCCCTATAGATGAAGGTGTTCTAGGATTTCAAAAATATTACGAT
GCTGCTGCCGTAAGCCCAGCATACAAATCTTTAGATTA AAAACGAGAAGTT
AATGTAGAATATTTGGATTTGATTTGAGATCTAATTCTCTAAGAAAAATAT
ACAAAAGTAAAATGCAAGGCAGTGTAGAGAGACGACGCAGTATTCCTGAT
GAAATGTTTTTGAATATTGAGATCCCGAATCCTCCTGAAGAGGTTAAAGAT
CAAATAGTAAAACAACATAAACTAATAAAGGAAATTGAGAATAGTCTCAAG
GAAAATCAAAAAAATTGCGTCTAAAGACAGAAGCATTATGGGAACTTCCT
CAAAATTACA ACTAA
ATGGACA ACTCTAAACTACCAATCAACCAGATTATTGCTCGCATCAATGAT
GCTGCGAAACATGGTGAAGCTTTGGTGCTAACCCTGGAAGAAGTAAAGAT
TCTTTCTAAAGATATTGGCGACAAAGTCTTTATTCCTGTGCTTACTAATGAG
CAGGTCGTGCAGTTGGTAAAAGAAGGAAAGCTAGGCCAGAAAATTAATAA
CACAAAAGATTAA
ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA
ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA
GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT
ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT
TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTATTTCGACCAA
GTGCAACTGACTTTAGATAGAACCAATTGGAATGGGGAAAACGAAATATT

	osase					ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTATTTCGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGCATTGCT TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG
	ISAb					ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTATTTCGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG
	33					
<i>group_</i> 3275	IS4 family transp osase ISAb 33	A2, A3, A14	L	Transposase DDE domain		ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTATTTCGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG
<i>group_</i> 3276	IS4 family transp osase ISAb 33	A2, A3, A14	L	Transposase DDE domain		ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTATTTCGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG

<i>group_3280</i>	IS4 family transposase ISAb 33	A2, A3, A14	L	Transposase DDE domain	TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTCATTGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT TTGATTCAACGCTTTATAGCCATTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTCATTGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT TTGATTCAACGCTTTATAGCCATTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG
<i>group_3281</i>	IS4 family transposase ISAb 33	A2, A3, A14	L	Transposase DDE domain	TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTCATTGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT TTGATTCAACGCTTTATAGCCATTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATCTCTAAAATGGA ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTCATTGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT
<i>group_3282</i>	IS4 family transposase ISAb 33	A2, A3, A14	L	Transposase DDE domain	TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATCTCTAAAATGGA ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTCATTGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT

<i>group_3283</i>	IS4 family transposase ISAb 33	A2, A3, A14	L	Transposase DDE domain	<p>AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGCATTGCT TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG</p> <p>ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTCATTGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG</p> <p>ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTCATTGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG</p>
<i>group_3284</i>	IS4 family transposase ISAb 33	A2, A3, A14	L	Transposase DDE domain	<p>AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGCATTGCT TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG</p> <p>ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTCATTGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG</p>
<i>group_3285</i>	IS4 family transposase	A2, A3, A14	L	Transposase DDE domain	<p>ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT</p>

	osase					ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTATTTCGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTATTTCGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATCTCTAAAATGGA ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTATTTCGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGCATTGCT TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG
<i>group_</i> 3286	IS4 family transp osase ISAb 33	A2, A3, A14	L	Transposase DDE domain		
<i>group_</i> 3287	IS4 family transp osase ISAb 33	A2, A3, A14	L	Transposase DDE domain		

<i>group_3288</i>	IS4 family transposase ISAbA33	A2, A3, A14	L	Transposase DDE domain	<p>ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATCTCTAAAATGGAACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTTACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTCATTGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGCATTGCT TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG</p> <p>ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGAACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTTACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTCATTGACCAA GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA TTTAG</p> <p>ATGACTACGACAGTTAATTCAGACATGATCATCTATAATCAATTGGCTCAA ACTGCTTATTTAGAGCGTTTGCAAGATAATTTGAATGTATTTAACCAAGCCT CTAATGGTGCAATTGTTTATCGCAATGAGATCATTGAAGGTGATTTCAATA AAGAAGCATTCTACAAAGTGGGCGGTAGCATCAAACATCGTGATGTGAAT TCAACCGCCAAAGTAGTTCCAGAGAAAATTGGTTCTGGTGAGTCTGTAGG CGTAAAAGTCCCATATAAATATGGTCCTTATGCATCAACTGAAGAGGCATT TAAGCGCCGTGCTCGTACACCAGAAGAATTTGCTATGGTTGTTGGTTACG ATCTTGCAGATGCATTGGTTGCAGGCCGATTAGAGTACAGTTTAGCTTCTT TAAAAGCTGCTATTTCTAGCAATCCAGACATGGTTGCAAAGGTAGTATCG</p>
<i>group_3289</i>	IS4 family transposase ISAbA33	A2, A3, A14	L	Transposase DDE domain	<p>ATGACTACGACAGTTAATTCAGACATGATCATCTATAATCAATTGGCTCAA ACTGCTTATTTAGAGCGTTTGCAAGATAATTTGAATGTATTTAACCAAGCCT CTAATGGTGCAATTGTTTATCGCAATGAGATCATTGAAGGTGATTTCAATA AAGAAGCATTCTACAAAGTGGGCGGTAGCATCAAACATCGTGATGTGAAT TCAACCGCCAAAGTAGTTCCAGAGAAAATTGGTTCTGGTGAGTCTGTAGG CGTAAAAGTCCCATATAAATATGGTCCTTATGCATCAACTGAAGAGGCATT TAAGCGCCGTGCTCGTACACCAGAAGAATTTGCTATGGTTGTTGGTTACG ATCTTGCAGATGCATTGGTTGCAGGCCGATTAGAGTACAGTTTAGCTTCTT TAAAAGCTGCTATTTCTAGCAATCCAGACATGGTTGCAAAGGTAGTATCG</p>
<i>group_329</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGACTACGACAGTTAATTCAGACATGATCATCTATAATCAATTGGCTCAA ACTGCTTATTTAGAGCGTTTGCAAGATAATTTGAATGTATTTAACCAAGCCT CTAATGGTGCAATTGTTTATCGCAATGAGATCATTGAAGGTGATTTCAATA AAGAAGCATTCTACAAAGTGGGCGGTAGCATCAAACATCGTGATGTGAAT TCAACCGCCAAAGTAGTTCCAGAGAAAATTGGTTCTGGTGAGTCTGTAGG CGTAAAAGTCCCATATAAATATGGTCCTTATGCATCAACTGAAGAGGCATT TAAGCGCCGTGCTCGTACACCAGAAGAATTTGCTATGGTTGTTGGTTACG ATCTTGCAGATGCATTGGTTGCAGGCCGATTAGAGTACAGTTTAGCTTCTT TAAAAGCTGCTATTTCTAGCAATCCAGACATGGTTGCAAAGGTAGTATCG</p>

group_3290
 IS4 family
 transposase
 ISAbal
 33

A2, A3, A14

L

Transposase DDE domain

group_33
eptA_2
 Phosphotransferase
 EptA

L7, L8, L9, L10

S

Domain of unknown function (DUF1705)

TTGTTGATGGCCGCAAAGCATTGACTCGTGGTATGCGAAAGTTTGGTGAT
 AAGTTTGGCCGCATTGGCTTATGGGTGATGAACTCAGATACATATTTTCGAT
 ATTGTCGATGATGCAATCACTAAGCAAATTTATGGTGAATCTGAAATCGTT
 ATCTACGGTGGTTTACCGGGAACCTTAGGAAAGCCGGTCTTGGTGACGGA
 CGCTGTAGGTGATAACGATGCTTTTGGTTTGCAGTATGGCGCTGTCACTG
 TAACTGAATCACAAGTACCGGGCTTCCGAGCTTATGACATCAATGATGAA
 GAAAACCTAGCAATCGGTATGCGTGCTGAAGGTGCATTTAACTTAGATATT
 CTTGGTTATAGTTGGGATACATCGAAAGGTGAAAATCCTGACCTTACATTA
 CTTGGTTCAAGCGCTAACTGGATCAAATATGCAACCAGCAACAAAATGACA
 GCAGGTACCTTACTTGATTTATCGGGTACAGCGACAACCTGGTTAA
 ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA
 ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTTAAA
 GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT
 ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT
 TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTCATTCGACCAA
 GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT
 AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT
 TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT
 TTGATTCAACGCTTTATAGCCATTTTGGTAAAGACCGTATTGTGAATGTGT
 TCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAATTGAA
 CAAGACATCAACTTCTGCATTCGTGTTAAAAAACTTCATTGTCACCAATCA
 TTTAG
 ATGTTTAATCTCATTATAGCCATTTGGCTAGGTGCAATTTTAAATATTGGTT
 TTTATCATCAAGTCCATACTTTAACTCCATATTTTGGAGTAAAGGCTATTTT
 GTTTTTAGCTGCTACTCTCATAATTCTTGTGCTACCTATTATGCGGTTTTA
 CAAATCTTAAATTGGAAATGGACTGCCAAAATCTTTGCAATTTTATTGATAT
 TTATTGGTGGTTTTAGCTCTTATTTTGTAAACACATTAGGTGTCATTATTTT
 ACCCGACCAAATTCAAAATATGGTGCAGACCGATGTTTCGGAAGTTACCG
 ATCTAATCTCTTTACGCTTTGTTTTATGGACAATTTTTTTTTGTTATTTTACCC
 ATTTTTTAACTCAAGTTAAATTTAAACAAGAAAAAGTATCACGGTTGTT
 ATTGAAGAAAGTATTCTCACTGGTAGCTTCATTGGCCGTGGTCCGGTGT
 ACTTTTTACCTACTATGTGATTTTGTGCAATATTTTCGTGAGCATCGTGAT
 TAAAAGGGATGATTTACCGCAAATAGTATTTTCATCGCTTATGTCTTACT

ACCATAAGAAGGCTCCGAAGAAAAATCTGCCTCTTGTGATATATGGACAA
GATGCTCATCAAGTTCAGCGGTACAAAAGAATCTCCCTAAGTTAATGATA
CTTGTTGTGGGTGAAACGGCACGTGCCGAAAGTTTCTCTCTAAATGGGTA
TGCAAAAATACGAATCCGGAGCTTTCTAAACAAGATATTTTCAACTTCTCT
CAAGTGAGCTCATGCGGTACGGCAACAGCGGTTTCTGTGCCATGTATGTT
CTCGGGTATGCCACGTGTAGATTATGATGAGCAATTAGCCAGTCACCGCG
AAGGTTTACTAGATATTGCAAAACGTGCGGGTTACCAAGTAACTTGGATTG
ATAATAACTCGGGTTGTAAAGGTGCATGTGATCGCGTTGAGCAATACCAG
ATTCCAGAAAACTTAAAGAAAAAATGGTGTAAAGATGGCGAATGTTATGAT
GACATTCTCATTGACAGCTTAAAGCAGTATTTGTCTACTATTGCCAAAGAT
GATGATCGTCCACGTTTGATTGTTTTGCATCAGGTGGGTAGTCATGGACC
TGATATTACAAGCGTGCGCCTGAAGCATATCAACCCTTAAACCGACTTG
TGATACGAATGCGATACAGGGCTGTTTCGCAAACCGAATTGCTAAATAGTTA
TGATAATACAATCGTATATACAGACCATGTATTAAGCCAAATGATTAATACT
CTAAAAGAAATATCAAAATATCAGACAGGTTTATGGTATTTATCTGATCATG
GCGAATCAACCGGAGAACATGGTTTATATTTACATGGTTCACCTTATGCAA
TCGCACCGAGCCAACAAACACATGTACCAATGATTATGTGGTTCTCTGAAA
GTTGGAAACAACATAATCTTGCTCAAGTGAATTGTTTAAGCCAACAACTA
AACAAAAGTTAAGTCAGGATAATTTATCCCAAGTTTGTAAAGTTTGTCTGGA
TGTAACAACCTCAGGTCATCAACCCTCAACTGGACATGTTGCACTCTTGTGC
CCATGTAAACTAA
ATGACTACAACAGTTAATTCAGACATGATCATCTACAACCAATTGGCTCAA
ACTGCTTATTTAGAGCGTTTGCAAGACAATTTGAACGTATTTAACCAGGCC
TCTAATGGTGCAATTGTATATCGCAATGAGATTATTGAAGGTGATTTCAATA
AAGAGGCATTTTACAAAGTGGGCGGTAGCATCAAGCATCGTGATGTGAAT
TCAACCGCCAAAGTAGTGCCTGAGAAGATTGGGTCTGGTGAATCTGTAGG
CGTAAAAGTCCCATATAAATATGGTCCTTATGCTTCAACCGAAGAGGCATT
CAAACGCCGTGCACGCACACCTGAAGAGTTTGCCATGATTCTTGGTTATG
ATTTAGCAGATGCATTAGTTGCAGGACGTTTACAGTACAGTTTAGCTTCTT
TAAAAGCTGCTATTTCTAGCAATCCAGACATGGTTGCAAAGGTAGTATCG
TTGTTGATGGCCGCAAAGCATTGACTCGTGGTATGCGAAAGTTTGGTGAT
AAGTTTGGCCGATTGGCTTATGGGTGATGAACTCAGATACATATTTGAT
ATTGTCGATGATGCAATCACTAAGCAAATTTATGGTGAATCTGAAATCGTT

group_
330

hypoth
etical
protein

A2, A3,
A14

N

domain, Protein

<i>group_3315</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATCTACGGTGGTTTACCGGGAACCTTAGGAAAGCCGGTCTTGGTGACGGA CGCTGTAGGTGATAACGATGCTTTTGGTTTGCAGTATGGCGCTGTCACTG TAACTGAATCACAAGTACCGGGCTTCCGAGCTTATGACATCAATGATGAA GAAAACCTTAGCAATCGGTATGCGTGCTGAAGGTGCATTTAACTTAGATATT CTTGGTTATAGCTGGGATACATCAAAAGGTGAAAACCCTGACCTTACTTTA CTTGGTTCAAGTGCCAACTGGAAAAAACATGCTACTAGCAACAAAATGACA GCAGGCACATTGCTTGATCTATCTGGCACAACAACAACTGGTTAA ATGAAAAGTCATGATCATTGCTCGATAGGCAATATGACGAGGATCACTAC AACTGTGTTCACTTTGTTTCATGAAGCTGCAATGGACCTATATGGCATAGAT CGGGCTGAAGCGCTTGAACTCTTTATGCAGCCTAAGGGCAAAAATTACTTTT TTATCTTCACGGTTAAACTTTTTAAATCCGCTACCCATGCCCAAGGAAGGC TGCATAGTCGCCTTCCATCCGAGACAAAGAAATAAGCCCCCGCATGTGGG GCTTTTTTCGTGGGCAAAAGATTCTTCACCTCATGGAAAGCGGAGTCACTTA TTTGCCTGAAGAGGTTGTGATGGAAATGGGGTTTAATCGGGTCAGTTATTA TGATTAA ATGAAAAGTCATGATCATTGCTCGATAGGCAATATGACGAGGATCACTAC AACTGTGTTCACTTTGTTTCATGAAGCTGCAATGGACCTATATGGCATAGAT CGGGCTGAAGCGCTTGAACTCTTTATGCAGCCTAAGGGCAAAAATTACTTTT TTATCTTCACGGTTAAACTTTTTAAATCCGCTACCCATGCCCAAGGAAGGC TGCATAGTCGCCTTCCATCCGAGACAAAGAAATAAGCCCCCGCATGTGGG GCTTTTTTCGTGGGCAAAAGATTCTTCACCTCATGGAAAGCGGAGTCACTTA TTTGCCTGAAGAGGTTGTGATGGAAATGGGGTTTAATCGGGTCAGTTATTA TGATTAA ATGGGATGGAAGGGGAAAAAGCCAACCTGAATTTAGTTTTGATGTGGCTAA AACAGCAGAAGACCATGTAAGAATATTGTCATGGATACCGTGCAATCCTT AGTTAATTTAAGTCCCGTCGATACTGGTGCATACCGTGCTTCACATATTGT CTCGATTAGATCTGCTGATTTAGGCGTGCGTGAACCTGAAACAAACCCTG TTAACGATGCAGCAATTCAAGCTGTAAAGATTAATTGGGCAATTTGGTCT ACATTCAGAATAACCAACCTTATGCTGAACGTTTAAAGAAACGGCTGGTCTG ATCAAGCACCACAAGGTATTTATGGTCTCACGTTTAACTTTATTTCTCAAAA GTACGGTGGCTAA ATGATTGACAAAAGTAAGTCCTCACTAAGCGAGGTTTTATCGCAAATTA GATGGTGCAACCATTCTGATTGGTGGTTTTGGTACCGCAGGACAACCCGC</p>	
<i>group_3316</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGGGATGGAAGGGGAAAAAGCCAACCTGAATTTAGTTTTGATGTGGCTAA AACAGCAGAAGACCATGTAAGAATATTGTCATGGATACCGTGCAATCCTT AGTTAATTTAAGTCCCGTCGATACTGGTGCATACCGTGCTTCACATATTGT CTCGATTAGATCTGCTGATTTAGGCGTGCGTGAACCTGAAACAAACCCTG TTAACGATGCAGCAATTCAAGCTGTAAAGATTAATTGGGCAATTTGGTCT ACATTCAGAATAACCAACCTTATGCTGAACGTTTAAAGAAACGGCTGGTCTG ATCAAGCACCACAAGGTATTTATGGTCTCACGTTTAACTTTATTTCTCAAAA GTACGGTGGCTAA ATGATTGACAAAAGTAAGTCCTCACTAAGCGAGGTTTTATCGCAAATTA GATGGTGCAACCATTCTGATTGGTGGTTTTGGTACCGCAGGACAACCCGC</p>	
<i>group_332</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGGGATGGAAGGGGAAAAAGCCAACCTGAATTTAGTTTTGATGTGGCTAA AACAGCAGAAGACCATGTAAGAATATTGTCATGGATACCGTGCAATCCTT AGTTAATTTAAGTCCCGTCGATACTGGTGCATACCGTGCTTCACATATTGT CTCGATTAGATCTGCTGATTTAGGCGTGCGTGAACCTGAAACAAACCCTG TTAACGATGCAGCAATTCAAGCTGTAAAGATTAATTGGGCAATTTGGTCT ACATTCAGAATAACCAACCTTATGCTGAACGTTTAAAGAAACGGCTGGTCTG ATCAAGCACCACAAGGTATTTATGGTCTCACGTTTAACTTTATTTCTCAAAA GTACGGTGGCTAA ATGATTGACAAAAGTAAGTCCTCACTAAGCGAGGTTTTATCGCAAATTA GATGGTGCAACCATTCTGATTGGTGGTTTTGGTACCGCAGGACAACCCGC</p>	
<i>group_3448</i>	<i>pcaI_1</i>	3-oxoadi	L6, L8, L9	I	3-oxoadipate CoA-transferase activity	<p>ATGATTGACAAAAGTAAGTCCTCACTAAGCGAGGTTTTATCGCAAATTA GATGGTGCAACCATTCTGATTGGTGGTTTTGGTACCGCAGGACAACCCGC</p>

pate
CoA-
transfe
rase
subuni
t A

TGAACTCATTGACGGTCTGATTGATCTTGGCGTTAAGGACCTGACCATTGT
GAGCAACAATGCCGGTAATGGCGATTATGGTCTGGCTAAACTGCTCAAAG
CCGGTTCTATTA AAAAAGTGATTGCTCTTTCCACGTCAGTCTGACTCTT
ATGTGTTTGATGAACTGTACCGTGCCGAAAGGTTGAACTTGAAGTGGTA
CCGCAAGGCAATCTGGCGTGCCGTATTCAGGCAGCAGGTATGGGCCTTG
GGGCGGTGTTTACCCCAACTGGCTTTGGAACACTTTTGGCCGAAGGCAAA
GAAACCCGTCATATTGATGGTAAAGATTACGTGCTTGAGTATCCGATCAAG
GCTGACTTTGCCCTGATTAAGCCTATAAGGGCGACCGCTGGGGAAATCT
GGTTTACCGTAAATCTGCACGTAACCTTTGGCCCGATCATGGCCATGGCTG
CTGATGTCACCATTGCTCAGGTTTCTGAAGTGGTTGAATTGGGTGGATTA
GACCCGGAACACATCATCACCCAGGTATTTTTGTACAGCATGTTGTACAA
GTACAGCCCGCACAGTAA
GTGGATACTACCGCGCCGCAAGCAGGTGAACTGACTTTATCTGACTTGAG
TGATACAGGTATTTTCAGCAACAGATCAGATCACGCAAGATAAAAACCTTCAA
TTTAAAGCTTGAAGGACAGGAAAGCGGCAGCCGAGTAACATATTTAGTTT
CTACCGATGAAGGTAAAACCTGGCAAGAGACGACATTAATCAAAAAGATT
TAGCTGACGGTATTTATCTTTATAAAGCTGTAGTGACAGACGCTGCTGGCA
ATACCTCGGAAACGGCTGTACAAAAAGTGGTTGTGGATACTACCACACCA
CACTCAGGTGAACTGACTCTATCTGACTTGAATGATACAGGCGTTTCAGTA
ACAGATCAGATCACGCAAGATAAAAACCTTCAATTTAAAGCTTGAAGGACAG
GAAACTGGTAGCCGAGTAACATATTTAGTTTCCACTGATGAAGGAAAAACT
TGGCAGGAAACCACGGTAGTCCAAAAGATTTGGCTGACGGTATTTATCA
ATATAAAGCTGTAGTGACAGACGCTGCTGGCAATACCTCGGAAACGGCTG
TACAAAAAGTGGTTGTGGATACTNNNNNNNNNAGCAGGTGAACTGA
ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA
ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA
GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT
ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT
TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTATTTCGACCAA
GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT
AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT
TGGACATTGCTTAATAAACGTGGAAATTAG

group_
35

hypoth
etical
protein

L8, L9,
L10

Q

von willebrand factor, type A

group_
350

IS4
family
transp
osase
ISAb
33

A2, A3,
A14

L

Transposase DDE domain

<i>group_351</i>	IS4 family transposase ISAb1 33	A2, A3, A14	L	Transposase DDE domain
------------------	---	----------------	---	------------------------

ATGACACATCTCAATGAGTTATATCTTATCTTAAACAAATATCTAAAATGGA
 ACAAGTCACATTTAAAGTGCTTTGCGCTCATCATGCTTGTGATTATTTAAA
 GCAAACATGTAATCTTTCTTCTGCATCTAAAGCCTTGCCCATCAAGTGCTT
 ACCACAATCATTTTATCGACGTATGCAGCGCTTCTTTGCAGGTCAGTATTT
 TGATTATCGTCAAATTTCTCAGTTGATTTTCAATATGTTTTATTTCGACCAA
 GTGCAACTGACTTTAGATAGAACCAATTGGAAATGGGGAAAACGAAATATT
 AATATCCTGATGCTCGCAATCGTTTATCGTGGAATAGCGATACCTATCCTT
 TGGACATTGCTTAATAAACGTGGAAATTCAGATACGAAAGAGCGTATTGCT
 TTGATTCAACGCTTTATAGCCATTTTTGGTAAAGACCGTATTGTGAATGTGT
 TCTCGCAGACAGAGAGTTTATCGGTGAGCAGTGGTTTACATGGTTAA
 ATGGCAAATCTTGTTTTTAAATTTAGTTGGGATCATCGGCCATTCCCGTAT
 AACTCGGCTCAGGGAAAACGGCAATTCATGCTGCCATTGCTTCAGGTAT
 TCCTAATCTGGCACCAAACTTTTCGCAGGTCCAAGGTAAGTCTGCTGCAGTCT
 CTCAAGGTGGTACTGGGGCGCAACTGCACTAGATGCTCGAAATAATCTC
 GGAGCAGCAGAAAAAGGGTGAATAGTGACATTACAGAGCTAAAAGGATT
 AACCAAGGCTATTGCAATTTCTCAAGGTGGTACCGGTGCAACAACCTCCAT
 CCGATGCTCGAACTAACTTAGGGCTTGGTAGTGCCGCAACTAGAAATGTT
 GGTACTACAGCTGGTAATTTGATAGAAGTTGGCAGCTTTGGAATTGGTGG
 AGTAGGCCAAACTTTTTGAAAGAAAATGATTACGGGAGTAAACCTAGATTCT
 TGTCGTTAGCTATGTATTGTTATTTCTTATTCTGTGAGCAGCTCACCCAAT
 CGAAACATGTTTGGTGAGCTAGTGTTCGAGGGGTGATTCAGGCTCAGC
 AAATCAACATTCGAGAATTTAGTATCAATTCAGCAAGCATATGATCGTGT
 TACAGCTCGGCTTATTAGTATTGGTGTAAAAACTCATATTTAGGTATGGC
 TGTAGTTAAATATCAAAATGTAGACTATGTTGCCATTGGAAGAACAGCAAG
 TTCTTCAACATCGGCATTTAGATATTTTTCCGGTATTTCCAATATTACATCT
 GATAATTATTTAGTTACTGTTTCATACAGATGACGTTGTTATTGTGAGTGA
 TACCTGTTGTAATTGAGCAGCTAAGAACATCTGCGAATACTTCTGTGGATT
 CCAACGGTTTTATAAAAGCAGCATCACCAAGTAGTTAAGCTATTTAACGACC
 ATATCGAGCTCAATAATGATGCAAAAAACAGCCGATTGAATTTAAGAGAA
 TTGATGTTGGTGATTATTTACTAGAAGGTTCTTTAGGCTTTGCTCAGGAAG
 GCTGGTATATCGAAGTACCGAAAGATGCAAGCGGCAACACAATCGTCGCA
 GTAGTGTATGACACCCTAGAAAATGGTGACATCTCAATTAACCTTACAAG
 CGTAAGTTTGATTTTGAACCTTGCTGCTGTTGTGGCAGATCACGAGAACCCA

<i>group_3514</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein
-------------------	-------------------------	----------------	---	----------------------

<i>group_3663</i>	hypothetical protein	A2, A3, A14	S	Domain of unknown function (DUF1833)	<p>ATGGACATTCCAGAAGGCCGCTGGATTGATATCCGTCTGCATGAAGAACC TGAACCAGAACCTGAGGTTGAAGAACTTTGACTGAAACACCAGTGGATT TCCAGCCGACTAATTATCTCAGGCAGTTGCTGCAGCAATGAATGGCGTG GAACCCCAGAAATCTCAGACACAGACGAAACACTTTAA ATGGATAACGAATATGCCAAGTTCTTTTTCAATCGGAAAGTTGATGTCTAT CAACTGGAGTGTATTGAGCTATCACATCCTTCTTTTATGAATACTTATCGG GTAGTCCGTAATGATGACCGAGGTGTCTATGTACAACATAAGGAAGGATC CGGTCAGGTCTATTATGAGTTCTTGCCAGTCTCTATACAAAGATCCGGAAT GCTTGGTGATCTGGACCAGACATTAACCGTTTCTATCTCTGGTCTAGGTGA TGTGATGCCTGATGAGTTTGAACGGGTAATCGAAGGGCAATATCCAGATG TAAAGCCAACCGTAAATTACCGGATTTACAGTTCAGACAATCTGAACTCTC CAATGTTTTATTTACTTGGACTGCAACTCTCAAGTGTCGCCATGAACCATA AAGCTGTGACATTCAAGGCTGAATCACCACGATTAATACCACTAAAACCTG GGGACATTTTTGCACTGGATCGCTTTAGTGGTTTGAAGGGGGCTATATGA ATGGATAACGAATATGCCAAGTTCTTTTTCAATCGGAAAGTTGATGTCTAT CAACTGGAGTGTATTGAGCTATCACATCCTTCTTTTATGAATACTTATCGG GTAGTCCGTAATGATGACCGAGGTGTCTATGTACAACATAAGGAAGGATC CGGTCAGGTCTATTATGAGTTCTTGCCAGTCTCTATACAAAGATCCGGAAT GCTTGGTGATCTGGACCAGACATTAACCGTTTCTATCTCTGGTCTAGGTGA TGTGATGCCTGATGAGTTTGAACGGGTAATCGAAGGGCAATATCCAGATG TAAAGCCAACCGTAAATTACCGGATTTACAGTTCAGACAATCTGAACTCTC CAATGTTTTATTTACTTGGACTGCAACTCTCAAGTGTCGCCATGAACCATA AAGCTGTGACATTCAAGGCTGAATCACCACGATTAATACCACTAAAACCTG GGGACATTTTTGCACTGGATCGCTTTAGTGGTTTGAAGGGGGCTATATGA GTGAAACTATATATAAGTGCACTACAGTTAGAAAATGGAGAGCTTTTACTC GTCGTTTCTCCTCAGTTTAATGCCAATGCTATTACAGGATTATGCATTACGC TGGGAAATTGAAACCTTATTAGTTGTCTCAAAGGACGCGGGTTTAATCTT GAAAATACGCGCTTGACAGACCCTAGACGAGTGAAAAAATTGATTGCGGT GTTAGCTATAAGCTTCTGTTGGTGTACTTAACGGGTGAATGGCAACATGA TCAAAAAAAGCGATAAAAATAAAGAAGCATGGACGACTCTCAATGAGTTT ATTTGCTATGGTTTACTATGTTCAAATGGCGATTACGCGTTTAATTGG TTTTGGGAAAAAAGAAGAGTTTAAGGAAATTTTGGCAATTTTAAGAAGGCA GAACCCTGATAGGATAAGGGTTCTGTGA</p>
<i>group_3664</i>	hypothetical protein	A2, A3, A14	S	Domain of unknown function (DUF1833)	<p>ATGGACATTCCAGAAGGCCGCTGGATTGATATCCGTCTGCATGAAGAACC TGAACCAGAACCTGAGGTTGAAGAACTTTGACTGAAACACCAGTGGATT TCCAGCCGACTAATTATCTCAGGCAGTTGCTGCAGCAATGAATGGCGTG GAACCCCAGAAATCTCAGACACAGACGAAACACTTTAA ATGGATAACGAATATGCCAAGTTCTTTTTCAATCGGAAAGTTGATGTCTAT CAACTGGAGTGTATTGAGCTATCACATCCTTCTTTTATGAATACTTATCGG GTAGTCCGTAATGATGACCGAGGTGTCTATGTACAACATAAGGAAGGATC CGGTCAGGTCTATTATGAGTTCTTGCCAGTCTCTATACAAAGATCCGGAAT GCTTGGTGATCTGGACCAGACATTAACCGTTTCTATCTCTGGTCTAGGTGA TGTGATGCCTGATGAGTTTGAACGGGTAATCGAAGGGCAATATCCAGATG TAAAGCCAACCGTAAATTACCGGATTTACAGTTCAGACAATCTGAACTCTC CAATGTTTTATTTACTTGGACTGCAACTCTCAAGTGTCGCCATGAACCATA AAGCTGTGACATTCAAGGCTGAATCACCACGATTAATACCACTAAAACCTG GGGACATTTTTGCACTGGATCGCTTTAGTGGTTTGAAGGGGGCTATATGA ATGGATAACGAATATGCCAAGTTCTTTTTCAATCGGAAAGTTGATGTCTAT CAACTGGAGTGTATTGAGCTATCACATCCTTCTTTTATGAATACTTATCGG GTAGTCCGTAATGATGACCGAGGTGTCTATGTACAACATAAGGAAGGATC CGGTCAGGTCTATTATGAGTTCTTGCCAGTCTCTATACAAAGATCCGGAAT GCTTGGTGATCTGGACCAGACATTAACCGTTTCTATCTCTGGTCTAGGTGA TGTGATGCCTGATGAGTTTGAACGGGTAATCGAAGGGCAATATCCAGATG TAAAGCCAACCGTAAATTACCGGATTTACAGTTCAGACAATCTGAACTCTC CAATGTTTTATTTACTTGGACTGCAACTCTCAAGTGTCGCCATGAACCATA AAGCTGTGACATTCAAGGCTGAATCACCACGATTAATACCACTAAAACCTG GGGACATTTTTGCACTGGATCGCTTTAGTGGTTTGAAGGGGGCTATATGA GTGAAACTATATATAAGTGCACTACAGTTAGAAAATGGAGAGCTTTTACTC GTCGTTTCTCCTCAGTTTAATGCCAATGCTATTACAGGATTATGCATTACGC TGGGAAATTGAAACCTTATTAGTTGTCTCAAAGGACGCGGGTTTAATCTT GAAAATACGCGCTTGACAGACCCTAGACGAGTGAAAAAATTGATTGCGGT GTTAGCTATAAGCTTCTGTTGGTGTACTTAACGGGTGAATGGCAACATGA TCAAAAAAAGCGATAAAAATAAAGAAGCATGGACGACTCTCAATGAGTTT ATTTGCTATGGTTTACTATGTTCAAATGGCGATTACGCGTTTAATTGG TTTTGGGAAAAAAGAAGAGTTTAAGGAAATTTTGGCAATTTTAAGAAGGCA GAACCCTGATAGGATAAGGGTTCTGTGA</p>
<i>group_3769</i>	IS4 family transposase ISAb1	A2, A3, A14	L	Transposase DDE domain	<p>ATGGACATTCCAGAAGGCCGCTGGATTGATATCCGTCTGCATGAAGAACC TGAACCAGAACCTGAGGTTGAAGAACTTTGACTGAAACACCAGTGGATT TCCAGCCGACTAATTATCTCAGGCAGTTGCTGCAGCAATGAATGGCGTG GAACCCCAGAAATCTCAGACACAGACGAAACACTTTAA ATGGATAACGAATATGCCAAGTTCTTTTTCAATCGGAAAGTTGATGTCTAT CAACTGGAGTGTATTGAGCTATCACATCCTTCTTTTATGAATACTTATCGG GTAGTCCGTAATGATGACCGAGGTGTCTATGTACAACATAAGGAAGGATC CGGTCAGGTCTATTATGAGTTCTTGCCAGTCTCTATACAAAGATCCGGAAT GCTTGGTGATCTGGACCAGACATTAACCGTTTCTATCTCTGGTCTAGGTGA TGTGATGCCTGATGAGTTTGAACGGGTAATCGAAGGGCAATATCCAGATG TAAAGCCAACCGTAAATTACCGGATTTACAGTTCAGACAATCTGAACTCTC CAATGTTTTATTTACTTGGACTGCAACTCTCAAGTGTCGCCATGAACCATA AAGCTGTGACATTCAAGGCTGAATCACCACGATTAATACCACTAAAACCTG GGGACATTTTTGCACTGGATCGCTTTAGTGGTTTGAAGGGGGCTATATGA ATGGATAACGAATATGCCAAGTTCTTTTTCAATCGGAAAGTTGATGTCTAT CAACTGGAGTGTATTGAGCTATCACATCCTTCTTTTATGAATACTTATCGG GTAGTCCGTAATGATGACCGAGGTGTCTATGTACAACATAAGGAAGGATC CGGTCAGGTCTATTATGAGTTCTTGCCAGTCTCTATACAAAGATCCGGAAT GCTTGGTGATCTGGACCAGACATTAACCGTTTCTATCTCTGGTCTAGGTGA TGTGATGCCTGATGAGTTTGAACGGGTAATCGAAGGGCAATATCCAGATG TAAAGCCAACCGTAAATTACCGGATTTACAGTTCAGACAATCTGAACTCTC CAATGTTTTATTTACTTGGACTGCAACTCTCAAGTGTCGCCATGAACCATA AAGCTGTGACATTCAAGGCTGAATCACCACGATTAATACCACTAAAACCTG GGGACATTTTTGCACTGGATCGCTTTAGTGGTTTGAAGGGGGCTATATGA GTGAAACTATATATAAGTGCACTACAGTTAGAAAATGGAGAGCTTTTACTC GTCGTTTCTCCTCAGTTTAATGCCAATGCTATTACAGGATTATGCATTACGC TGGGAAATTGAAACCTTATTAGTTGTCTCAAAGGACGCGGGTTTAATCTT GAAAATACGCGCTTGACAGACCCTAGACGAGTGAAAAAATTGATTGCGGT GTTAGCTATAAGCTTCTGTTGGTGTACTTAACGGGTGAATGGCAACATGA TCAAAAAAAGCGATAAAAATAAAGAAGCATGGACGACTCTCAATGAGTTT ATTTGCTATGGTTTACTATGTTCAAATGGCGATTACGCGTTTAATTGG TTTTGGGAAAAAAGAAGAGTTTAAGGAAATTTTGGCAATTTTAAGAAGGCA GAACCCTGATAGGATAAGGGTTCTGTGA</p>

	osase ISAb 1					GAAAATACGCGCTTGACAGACCCTAGACGAGTGAAAAAATTGATTGCGGT GTTAGCTATAAGCTTCTGTTGGTGTACTTAACGGGTGAATGGCAACATGA TCAAAAAAAGCGATAAAAATAAAGAAGCATGGACGACTCTCAATGAGTTT ATTTTCGCTATGGTTTACTATGTTCAAATGGCGATTCAGCGTTTAATTGG TTTTGGGAAAAAAGAAGAGTTTAAGGAAATTTTGGCAATTTTAAGAAGGCA GAACCCTGATAGGATAAAGGGTTCTGTGA GTGAAACTATATATAAGTGCCTACAGTTAGAAAATGGAGAGCTTTTACTC GTCGTTTCTCCTCAGTTTAATGCCAATGCTATTCAGGATTATGCATTACGC TGGGAAATTGAAACCTTATTAGTTGTCTCAAAGGACGCGGGTTTAATCTT GAAAATACGCGCTTGACAGACCCTAGACGAGTGAAAAAATTGATTGCGGT GTTAGCTATAAGCTTCTGTTGGTGTACTTAACGGGTGAATGGCAACATGA TCAAAAAAAGCGATAAAAATAAAGAAGCATGGACGACTCTCAATGAGTTT ATTTTCGCTATGGTTTACTATGTTCAAATGGCGATTCAGCGTTTAATTGG TTTTGGGAAAAAAGAAGAGTTTAAGGAAATTTTGGCAATTTTAAGAAGGCA GAACCCTGATAGGATAAAGGGTTCTGTGA GTGAAACTATATATAAGTGCCTACAGTTAGAAAATGGAGAGCTTTTACTC GTCGTTTCTCCTCAGTTTAATGCCAATGCTATTCAGGATTATGCATTACGC TGGGAAATTGAAACCTTATTAGTTGTCTCAAAGGACGCGGGTTTAATCTT GAAAATACGCGCTTGACAGACCCTAGACGAGTGAAAAAATTGATTGCGGT GTTAGCTATAAGCTTCTGTTGGTGTACTTAACGGGTGAATGGCAACATGA TCAAAAAAAGCGATAAAAATAAAGAAGCATGGACGACTCTCAATGAGTTT ATTTTCGCTATGGTTTACTATGTTCAAATGGCGATTCAGCGTTTAATTGG TTTTGGGAAAAAAGAAGAGTTTAAGGAAATTTTGGCAATTTTAAGAAGGCA GAACCCTGATAGGATAAAGGGTTCTGTGA GTGAAACTATATATAAGTGCCTACAGTTAGAAAATGGAGAGCTTTTACTC GTCGTTTCTCCTCAGTTTAATGCCAATGCTATTCAGGATTATGCATTACGC TGGGAAATTGAAACCTTATTAGTTGTCTCAAAGGACGCGGGTTTAATCTT GAAAATACGCGCTTGACAGACCCTAGACGAGTGAAAAAATTGATTGCGGT GTTAGCTATAAGCTTCTGTTGGTGTACTTAACGGGTGAATGGCAACATGA TCAAAAAAAGCGATAAAAATAAAGAAGCATGGACGACTCTCAATGAGTTT ATTTTCGCTATGGTTTACTATGTTCAAATGGCGATTCAGCGTTTAATTGG TTTTGGGAAAAAAGAAGAGTTTAAGGAAATTTTGGCAATTTTAAGAAGGCA GAACCCTGATAGGATAAAGGGTTCTGTGA GTGAAACTATATATAAGTGCCTACAGTTAGAAAATGGAGAGCTTTTACTC GTCGTTTCTCCTCAGTTTAATGCCAATGCTATTCAGGATTATGCATTACGC TGGGAAATTGAAACCTTATTAGTTGTCTCAAAGGACGCGGGTTTAATCTT GAAAATACGCGCTTGACAGACCCTAGACGAGTGAAAAAATTGATTGCGGT GTTAGCTATAAGCTTCTGTTGGTGTACTTAACGGGTGAATGGCAACATGA TCAAAAAAAGCGATAAAAATAAAGAAGCATGGACGACTCTCAATGAGTTT ATTTTCGCTATGGTTTACTATGTTCAAATGGCGATTCAGCGTTTAATTGG TTTTGGGAAAAAAGAAGAGTTTAAGGAAATTTTGGCAATTTTAAGAAGGCA GAACCCTGATAGGATAAAGGGTTCTGTGA
<i>group_3778</i>	IS4 family transp osase ISAb 1	A2, A3, A14	L	Transposase DDE domain		
<i>group_3779</i>	IS4 family transp osase ISAb 1	A2, A3, A14	L	Transposase DDE domain		
<i>group_3780</i>	IS4 family transp osase ISAb 1	A2, A3, A14	L	Transposase DDE domain		

	osase ISAb 1					GAAAATACGCGCTTGACAGACCCTAGACGAGTGAAAAAATTGATTGCGGT GTTAGCTATAAGCTTCTGTTGGTGTACTTAACGGGTGAATGGCAACATGA TCAAAAAAAGCGATAAAAATAAAGAAGCATGGACGACTCTCAATGAGTTT ATTTTCGCTATGGTTTACTATGTTCAAATGGCGATTCAGCGTTTAATTGG TTTTGGGAAAAAAGAAGAGTTTAAGGAAATTTTGGCAATTTTAAGAAGGCA GAACCCTGATAGGATAAAGGGTTCTGTGA GTGAAACTATATATAAGTGCCTACAGTTAGAAAATGGAGAGCTTTTACTC GTCGTTTCTCCTCAGTTTAATGCCAATGCTATTCAGGATTATGCATTACGC TGGGAAATTGAAACCTTATTAGTTGTCTCAAAGGACGCGGGTTTAATCTT GAAAATACGCGCTTGACAGACCCTAGACGAGTGAAAAAATTGATTGCGGT GTTAGCTATAAGCTTCTGTTGGTGTACTTAACGGGTGAATGGCAACATGA TCAAAAAAAGCGATAAAAATAAAGAAGCATGGACGACTCTCAATGAGTTT ATTTTCGCTATGGTTTACTATGTTCAAATGGCGATTCAGCGTTTAATTGG TTTTGGGAAAAAAGAAGAGTTTAAGGAAATTTTGGCAATTTTAAGAAGGCA GAACCCTGATAGGATAAAGGGTTCTGTGA GTGAAACTATATATAAGTGCCTACAGTTAGAAAATGGAGAGCTTTTACTC GTCGTTTCTCCTCAGTTTAATGCCAATGCTATTCAGGATTATGCATTACGC TGGGAAATTGAAACCTTATTAGTTGTCTCAAAGGACGCGGGTTTAATCTT GAAAATACGCGCTTGACAGACCCTAGACGAGTGAAAAAATTGATTGCGGT GTTAGCTATAAGCTTCTGTTGGTGTACTTAACGGGTGAATGGCAACATGA TCAAAAAAAGCGATAAAAATAAAGAAGCATGGACGACTCTCAATGAGTTT ATTTTCGCTATGGTTTACTATGTTCAAATGGCGATTCAGCGTTTAATTGG TTTTGGGAAAAAAGAAGAGTTTAAGGAAATTTTGGCAATTTTAAGAAGGCA GAACCCTGATAGGATAAAGGGTTCTGTGA GTGAAACTATATATAAGTGCCTACAGTTAGAAAATGGAGAGCTTTTACTC GTCGTTTCTCCTCAGTTTAATGCCAATGCTATTCAGGATTATGCATTACGC TGGGAAATTGAAACCTTATTAGTTGTCTCAAAGGACGCGGGTTTAATCTT GAAAATACGCGCTTGACAGACCCTAGACGAGTGAAAAAATTGATTGCGGT GTTAGCTATAAGCTTCTGTTGGTGTACTTAACGGGTGAATGGCAACATGA TCAAAAAAAGCGATAAAAATAAAGAAGCATGGACGACTCTCAATGAGTTT ATTTTCGCTATGGTTTACTATGTTCAAATGGCGATTCAGCGTTTAATTGG TTTTGGGAAAAAAGAAGAGTTTAAGGAAATTTTGGCAATTTTAAGAAGGCA GAACCCTGATAGGATAAAGGGTTCTGTGA ATGAAAAAATAATTGTAATTTCTACAACACTTTTAGGCCTTACGGGATGTG CCATTCCTGCGGTAATAATCTCGTAAGATCCACAAATATGTATCAGGATG ATGTTTCGGGAAATACTGCAAATTTAAGGGTTTATAGAAGTAATATCCCAT GGTGCAGTTCTATATTACTTATCAAATAATGAGGGTGAAAAATTTCAAAA AACCTAATAACTAAGCAGATTTCAAATAATTTAACAAAGTATGGCTCTATGC ATGAGCCCAAAAATTAATATGCCTAAACCCACAATCAGTTTAAATAATG GTGAAGAGTTTTTTGAGTTTAAAGTACCCGCAAATAAGAAGTTAACTTTCA GGCTTACTTCTGTTATTGGGTCACTACTATGTATAGTTGTGATGTAAAAAT GGACTATCAGTTGGAAAGAAATGGAAATTATGAATTGATCCGTTTAAACA
<i>group_3789</i>	IS4 family transp osase ISAb 1	A2, A3, A14	L	Transposase DDE domain		
<i>group_3790</i>	IS4 family transp osase ISAb 1	A2, A3, A14	L	Transposase DDE domain		
<i>group_3798</i>	hypoth etical protein	A2, A3, A14	/	hypothetical protein		

<i>group_3799</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>GATCAAAGATTTTGTGAATCCAGCTTTACTGACTGAACCATCTCAAGATGA AGCCTACTGCAAGTTTGTAGTAAAAGAGATTTTTGAAGATGGTAAAGAAAC TATTATTAATCGATTTCTTAA ATGAAAAAATAATTGTAATTTCTACAACACTTTTAGGCCTTACGGGATGTG CCATTCCTGCGGTAAATAATCTCGTAAGATCCACAAATATGTATCAGGATG ATGTTTCGGGAAATACTGCAAATTTAAGGGTTTATAGAAGTAATATCCCAT GGTGCAGTTCTATATTACTTATCAAATAATGAGGGTGAAAAATTTCAAAA AACCTAATAACTAAGCAGATTTCAAATAATTTAACAAAGTATGGCTCTATGC ATGAGCCCAAAAAATTAATATGCCTAAACCCACAATCAGTTTAAATAATG GTGAAGAGTTTTTTGAGTTTAAAGTACCCGCAAATAAGAAGTTAACTTTCA GGCTTACTTCTGTTATTGGGTCAACTACTATGTATAGTTGTGATGTAAAAAT GGACTATCAGTTGGAAAGAAATGGAAATTATGAATTGATCCGTTTAAACA GATCAAAGATTTTGTGAATCCAGCTTTACTGACTGAACCATCTCAAGATGA AGCCTACTGCAAGTTTGTAGTAAAAGAGATTTTTGAAGATGGTAAAGAAAC TATTATTAATCGATTTCTTAA ATGGGCTTTTTTATTTCTAAAATTTAGAGGTTGTTATGGCTTTAAAAGTCG GAATTATTAAGCTCGGACGTATCAAATGGTGTGAATACAAGGGGGTT GATGGCGAAGTACAGGCAGAGTTCAAAGTCCGTGGTATCGCCTATAAACC TTTTCAGGTAGCAATTGAACGAGCCGGAATCAGATTTTCAATCAAAGGCTA TGATGTGATGGTCAAAGATGAAAATGCCAAGCTTTACCATGAATTGTTAAT GGATGCGTGTGCTGCCCACTTAATTGAAGACTGGAAAGGTGTGGTATTTG CCGAAATCGTAGACGGTAAAACGGTTGAATCTGAAAAGCCATATACACCT GAGAATGCCTCAAAGCTTCTTAATCTTGGTGATATTGGTATTTCAATCTGG CTATTCATTAAGAAGCAGGCTCAGAAGATTCAGGAAGAAGCCGACAAGGA TAAGGCTTTAATTCTGGGAAAAGTCATCGAGCTCTACAAATACCAAAAAACG TATGCGTCGAAAACGCCGCATGAAATCGAACAAATCAAGTTCTTAG ATGGGCCTTGGGGCGGTGTTTACCCCAACTGGCTTTGGAACACTTTTGGC CGAAGGCAAAGAAACCCGTCATATTGATGGTAAAGATTACGTGCTTGAGT ATCCGATCAAGGCTGACTTTGCCCTGATTAAGCCTATAAGGGCGACCCG TGGGAAATCTGGTTTACCGTAAATCTGCACGTAACCTTTGGCCGATCAT GGCCATGGCTGCTGATGTCACCATTGCTCAGGTTTCTGAAGTGGTTGAAT TGGGTGGATTAGACCCGGAACACATCATCACCCAGGTATTTTTGTACAG CATGTTGTACAAGTACAGCCCGCACAGTAA</p>
<i>group_38</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>GATCAAAGATTTTGTGAATCCAGCTTTACTGACTGAACCATCTCAAGATGA AGCCTACTGCAAGTTTGTAGTAAAAGAGATTTTTGAAGATGGTAAAGAAAC TATTATTAATCGATTTCTTAA ATGAAAAAATAATTGTAATTTCTACAACACTTTTAGGCCTTACGGGATGTG CCATTCCTGCGGTAAATAATCTCGTAAGATCCACAAATATGTATCAGGATG ATGTTTCGGGAAATACTGCAAATTTAAGGGTTTATAGAAGTAATATCCCAT GGTGCAGTTCTATATTACTTATCAAATAATGAGGGTGAAAAATTTCAAAA AACCTAATAACTAAGCAGATTTCAAATAATTTAACAAAGTATGGCTCTATGC ATGAGCCCAAAAAATTAATATGCCTAAACCCACAATCAGTTTAAATAATG GTGAAGAGTTTTTTGAGTTTAAAGTACCCGCAAATAAGAAGTTAACTTTCA GGCTTACTTCTGTTATTGGGTCAACTACTATGTATAGTTGTGATGTAAAAAT GGACTATCAGTTGGAAAGAAATGGAAATTATGAATTGATCCGTTTAAACA GATCAAAGATTTTGTGAATCCAGCTTTACTGACTGAACCATCTCAAGATGA AGCCTACTGCAAGTTTGTAGTAAAAGAGATTTTTGAAGATGGTAAAGAAAC TATTATTAATCGATTTCTTAA ATGGGCTTTTTTATTTCTAAAATTTAGAGGTTGTTATGGCTTTAAAAGTCG GAATTATTAAGCTCGGACGTATCAAATGGTGTGAATACAAGGGGGTT GATGGCGAAGTACAGGCAGAGTTCAAAGTCCGTGGTATCGCCTATAAACC TTTTCAGGTAGCAATTGAACGAGCCGGAATCAGATTTTCAATCAAAGGCTA TGATGTGATGGTCAAAGATGAAAATGCCAAGCTTTACCATGAATTGTTAAT GGATGCGTGTGCTGCCCACTTAATTGAAGACTGGAAAGGTGTGGTATTTG CCGAAATCGTAGACGGTAAAACGGTTGAATCTGAAAAGCCATATACACCT GAGAATGCCTCAAAGCTTCTTAATCTTGGTGATATTGGTATTTCAATCTGG CTATTCATTAAGAAGCAGGCTCAGAAGATTCAGGAAGAAGCCGACAAGGA TAAGGCTTTAATTCTGGGAAAAGTCATCGAGCTCTACAAATACCAAAAAACG TATGCGTCGAAAACGCCGCATGAAATCGAACAAATCAAGTTCTTAG ATGGGCCTTGGGGCGGTGTTTACCCCAACTGGCTTTGGAACACTTTTGGC CGAAGGCAAAGAAACCCGTCATATTGATGGTAAAGATTACGTGCTTGAGT ATCCGATCAAGGCTGACTTTGCCCTGATTAAGCCTATAAGGGCGACCCG TGGGAAATCTGGTTTACCGTAAATCTGCACGTAACCTTTGGCCGATCAT GGCCATGGCTGCTGATGTCACCATTGCTCAGGTTTCTGAAGTGGTTGAAT TGGGTGGATTAGACCCGGAACACATCATCACCCAGGTATTTTTGTACAG CATGTTGTACAAGTACAGCCCGCACAGTAA</p>
<i>group_4</i>	<i>pcal_2</i> 3-oxoadipate CoA-transferase	L10	I	3-oxoadipate CoA-transferase activity	<p>GATCAAAGATTTTGTGAATCCAGCTTTACTGACTGAACCATCTCAAGATGA AGCCTACTGCAAGTTTGTAGTAAAAGAGATTTTTGAAGATGGTAAAGAAAC TATTATTAATCGATTTCTTAA ATGAAAAAATAATTGTAATTTCTACAACACTTTTAGGCCTTACGGGATGTG CCATTCCTGCGGTAAATAATCTCGTAAGATCCACAAATATGTATCAGGATG ATGTTTCGGGAAATACTGCAAATTTAAGGGTTTATAGAAGTAATATCCCAT GGTGCAGTTCTATATTACTTATCAAATAATGAGGGTGAAAAATTTCAAAA AACCTAATAACTAAGCAGATTTCAAATAATTTAACAAAGTATGGCTCTATGC ATGAGCCCAAAAAATTAATATGCCTAAACCCACAATCAGTTTAAATAATG GTGAAGAGTTTTTTGAGTTTAAAGTACCCGCAAATAAGAAGTTAACTTTCA GGCTTACTTCTGTTATTGGGTCAACTACTATGTATAGTTGTGATGTAAAAAT GGACTATCAGTTGGAAAGAAATGGAAATTATGAATTGATCCGTTTAAACA GATCAAAGATTTTGTGAATCCAGCTTTACTGACTGAACCATCTCAAGATGA AGCCTACTGCAAGTTTGTAGTAAAAGAGATTTTTGAAGATGGTAAAGAAAC TATTATTAATCGATTTCTTAA ATGGGCTTTTTTATTTCTAAAATTTAGAGGTTGTTATGGCTTTAAAAGTCG GAATTATTAAGCTCGGACGTATCAAATGGTGTGAATACAAGGGGGTT GATGGCGAAGTACAGGCAGAGTTCAAAGTCCGTGGTATCGCCTATAAACC TTTTCAGGTAGCAATTGAACGAGCCGGAATCAGATTTTCAATCAAAGGCTA TGATGTGATGGTCAAAGATGAAAATGCCAAGCTTTACCATGAATTGTTAAT GGATGCGTGTGCTGCCCACTTAATTGAAGACTGGAAAGGTGTGGTATTTG CCGAAATCGTAGACGGTAAAACGGTTGAATCTGAAAAGCCATATACACCT GAGAATGCCTCAAAGCTTCTTAATCTTGGTGATATTGGTATTTCAATCTGG CTATTCATTAAGAAGCAGGCTCAGAAGATTCAGGAAGAAGCCGACAAGGA TAAGGCTTTAATTCTGGGAAAAGTCATCGAGCTCTACAAATACCAAAAAACG TATGCGTCGAAAACGCCGCATGAAATCGAACAAATCAAGTTCTTAG ATGGGCCTTGGGGCGGTGTTTACCCCAACTGGCTTTGGAACACTTTTGGC CGAAGGCAAAGAAACCCGTCATATTGATGGTAAAGATTACGTGCTTGAGT ATCCGATCAAGGCTGACTTTGCCCTGATTAAGCCTATAAGGGCGACCCG TGGGAAATCTGGTTTACCGTAAATCTGCACGTAACCTTTGGCCGATCAT GGCCATGGCTGCTGATGTCACCATTGCTCAGGTTTCTGAAGTGGTTGAAT TGGGTGGATTAGACCCGGAACACATCATCACCCAGGTATTTTTGTACAG CATGTTGTACAAGTACAGCCCGCACAGTAA</p>

		subunit A					
<i>group_4003</i>		hypothetical protein	A2, A3, A14	/		hypothetical protein	<p>ATGGACAATTCAAAACGACCAATCAACCAGATTATTGCTCGCATCAATGATGCTGCGAAACATGGTGAAGCTTTGGTGCTAACCGCTGAAGAAGTAAAGATCTTTCTAAAGATATTGGCGACAAAGTCTTTATTCCTGTGCTTACTAATGAGCAGGTCGTGCAGTTGGTAAAAGAAGGAAAGCTAGGCCAGAAAATTAATAACACAAAAGATTAA</p> <p>ATGATTGATGGTGCATTCTTCCAAGACGAAGTGAACAACGCAAAATCATGGCTGTACCGATGGTATTCCAGGACAATGAACACATTGGTCAAGGTCGATGACGCTTGAAGAGATTGTGGCAAACTTGATACCAATTCAGCTGAAAAAGACGCAGCAGCACTCAATGCCAAAGATGCGTTTGATGTATTGGTGATTGGTGGTGGACCTGCGGGTGCAACAGCTGCAATCTATGCGGCACGTAAAGGCA TTAACACAGGTATCGTGGCTGAACGCTTTGGTGGTCAGGTCATGGACACCATGGATATTGAAACTTCACTTCTGTGCAAAAGACCCAAGGTCCTAAGTTTGCTGCTGAAATGGAAGCACATGTTTCGTGAATACGATGTGGATATTATGAACCTGCAACGTGTCAGCAAAATCACTGGTGCAGACCAAACAGCAAATGGCTGGTTGAAGTTGAACTTGA AAAACGGTGCAAACTTGAATCGAAA ACTGTGATCCTTTCAACAGGGGCACGTTGGAGAGAAATGAATGTACCAGGTGAAGCTGAATACCGTACCCGTGGTGTAGCGTACTGTCCACACTGTGATGGTCCATTGTTCAAAGGCAAACGTGTTGCTGTGATTGGTGGTGGTAACTCAGGTGTTGAAGCGCGATTGACCTTGCAGGTATTGTTGAGCATGTA ACTCTGGTTGAGTTCGACACTAACTTCGTGCTGACCAAGTCTTGCAAAACAAGTTAAATAGCTTGCCAAACACCACCGTGATTATGAATGCATTAAGCACAGAAGTATTAGGTGACGGTTCACAAGTAACCGGTCTTAAATATAAAGACCGTACA ACTGATGAAGAGCATGTGGTAGAGCTTGCAGGGATCTTTGTACAGATTGGTTTACTTCCAAACTGACTTTTTAAAAGACAGTGAGGTTGAGTTAACCAACCGTGGTGA GATCATTGTCAATGACCGCAACGAAACCAATGTA AAAAGGTGATTTGCTGCGGGTACTGTACCACTGTACCTTACAAGCAAATCATTATTGCTACAGGTGAAGGCGCTAAAGCATCACTCTCTGCGTTTGATTACATCATCCGTTCTGGGCAATAA</p> <p>TTGCAGAATCTGCCAGCCAATGATGTTCTCCTTACTGGTTCACGCGGCACAGGAAAATCTTCAATTGTCCGTGCCCTTTTAACCGAGTATAAAGATCAAGGATTACGTTTAATTGAAATTGAACGTGATGATCTGTCCGATCTTCTGAAATC</p>
<i>group_41</i>	<i>ahpF_1</i>	NADH dehydrogenase	L8, L9, L10	C		Alkyl hydroperoxide reductase	
<i>group_4127</i>		hypothetical protein	A2, A3, A14	S		Protein of unknown function (DUF815)	

		Alkyl			
		hydrop			
		eroxid			
<i>group_</i>	<i>ahpF_</i>	e	L8, L9	C	Alkyl hydroperoxide reductase
42	2	reduct			
		ase			
		subuni			
		t F			

CAGAAGCTCATCCAGAAGCGTCCGGAAAAGTTCATTGTCTACTGCGATGA
 TCTGGCCTTTAATGCTGAAGATGAAAACATCGCAGCTTAAAAAGTGTACT
 TGATGGCTCATTACAGTCAGGTTCAAGTAATTTTATTATTTATGCAACCAGC
 AACCGTCGACATCTGTTACCAGAATTTATGCATGAGAACACGCCAGTGAC
 TCGTGTAGATGTACCGCAATATACCGAGTTGCATCCACAAGAAGCCATTG
 AGGAAAAATTTCACTCTCGGATCGTTTTGGAATGTGGCTTTTCATTCTATC
 CAATGGATCAAACCTGTATTTAACCATTTGAGAACACTATCTTGCTAAAAC
 AGATATGCCAATGAATGACGAAGCCCATGCTGAAGCCCTACGCTGGTGTC
 AGGCACGTGGTCAACGGTCAGGACGTGCCGCCTATCAGTTCTCTAAACAC
 TGGATTGGATCACAACTAAAAGCACTGTAA
 GTGTCTGGCTATGCACCTAAAGTGTCCGGATGAAGTTTTAAATCAGATTA
 GGCTTAAACCTCAAAGCTAACTTTGATGTATTTGTATCACTCAGCTGTCAT
 AACTGTCCGGACGTGGTACAGGCGCTTAACCTGATTGCCATTTATAACCC
 GAACACCACGGCAACCATGATTGATGGTGCATTCTTCCAAGACGAAGTGG
 AACACGCAAAATCATGGCTGTACCGATGGTATTCCAGGACAATGAACAC
 ATTGGTCAAGGTCGTATGACGCTTGAAGAGATTGTGGCAAACTTGATAC
 CAATTCAGCTGAAAAGACGCAGCAGCACTCAATGCCAAAGATGCGTTTG
 ATGTATTGGTGATTGGTGGTGGACCTGCGGGTGCAACAGCTGCAATCTAT
 GCGGCACGTAAGGCATTAACACAGGTATCGTGGCTGAACGCTTTGGTG
 GTCAGGTCATGGATACCATGGACATTGAAAACCTTCACTTCTGTGCAAAAA
 CACAAGGTCCTAAGTTTGTGCTGCTGAAATGGAAGCACATGTTTCGTGAATAC
 GATGTGGATATCATGAACCTGCAACGTGTCAGCAAAATCACTGGTGCAGA
 CCAAACAGCAAATGGCTTGGTTGAAGTTGAACTTAAAACGGTGCAAAAC
 TTGAATCGAAAACCTGTGATCCTTTCAACAGGGGCACGTTGGAGAGAAATG
 AATGTACCAGGTGAAGCTGAATACCGTACCCGTGGTGTAGCGTACTGTCC
 ACACTGTGATGGTCCATTGTTCAAAGGCAAACGTGTTGCCGTGATTGGTG
 GCGGTAACCTCAGGTGTTGAAGCGGCGATTGACCTTGCAGGTATTGTTGAG
 CATGTAACCTCTGGTTGAGTTTGACACTAACTTCGTGCTGACCAAGTCTTG
 CAAAACAAGTTAAACAGCTTGCCAAACACCACTGTGATTATGAATGCATTA
 AGCACAGAAGTATTGGGTGACGGTTCACAAGTAACCGGTCTTAAATATAAA
 GACCGTGCAACTGATGGAGAGCATGTGGTAGAGCTTGCAGGGATCTTTGT
 ACAGATTGGTTTACTTCCAAACACTGACTTCTTAAAAGACAGTGAGGTTGA
 GTTAACCAATCGTGGAGAGATCATTGTCAATGACCGCAACGAAACCAATG

<i>group_4253</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>TAAAAGGTGATTTGCTGCAGGTGACTGTACCACTGTACCTTACAAGCAGATCATTATTGCTACAGGTGAAGGCGCTAAAGCATCACTCTCTGCGTTTGATTACATCATTCGTTCTGGGCAGTAG</p> <p>ATGTACAAGTTAAAGCTAAATCCTCAGACCAGCGGCTATGGCGTAACACCGGGTGATGATGTGAAACGTCAGCAGATGGATGGCGGTCGTGGTCGCTATTACATCGATGTGAAACGTAACAGCCACATTGTTGATGTGAACTGGAACTTAAGTAAAACCGATTTCAATAAAATGATGGCTTTCTGGCGGGTCTACCAGAACAACCAGCCTCATTTTATGCGGATTTGGTGATTGATCAGGGAACACGTCA</p> <p>GCAATATCTATGCAATTTCACTCCAAACTCGTTCAAGACCAATGAAGTCAACGGCAACTTATATCGTGTGAACGCGCAGCTCGAAGTTGTTCAAACCAGCCTAACCTGAATGCCGATATAGCATTAAATTAAGATTGGGAGGTCTAAATGTACAAGTTAAAGCTAAATCCTCAGACCAGCGGCTATGGCGTAACACCGGGTGATGATGTGAAACGTCAGCAGATGGATGGCGGTCGTGGTCGCTATTACATCGATGTGAAACGTAACAGCCACATTGTTGATGTGAACTGGAACTTAAGTAAAACCGATTTCAATAAAATGATGGCTTTCTGGCGGGTCTACCAGAACAACCAGCCTCATTTTATGCGGATTTGGTGATTGATCAGGGAACACGTCA</p> <p>GCAATATCTATGCAATTTCACTCCAAACTCGTTCAAGACCAATGAAGTCAACGGCAACTTATATCGTGTGAACGCGCAGCTCGAAGTTGTTCAAACCAGCCTAACCTGAATGCCGATATAGCATTAAATTAAGATTGGGAGGTCTAAATGAGTTCAGGTGCAAAAATTCGATTATATGCTTGTGAAGAAGCTGTTTTAGGAACTACTCCGGCAAATCCGGTCTGGTATACCGTTCGCCGTGTTACTGATAGTTTGACCGAAAATGTTACTACTGAAGATAGCAGTGAAGTAGTTGATCACGTTTTCGCCAAGGTGCTGTAGTAACGGAAGCTGAAGTAACTGGTCAGTTAGAATTTGAATTATCGCTAGGTACCTTTGATTTATTCTTAAATGTTCTCGCTTTCAATAACTGGGCTGCAAAATGCTTTAAGTTTTGGTGGTGGAGTCCGTAA</p> <p>GTCTCTTACCTTGGTAAAAGTCTTTGAAGATATTGGTCAAGTCTTTATTTATCGTGGTATTCAAGTGAATACAGGTGAAATGACAATCCAGACCACGGGGAAATCACTGGTAACTTTGGTTTAGTCGGTAGCTCATTTACACGACAGCAGGCAATCCTGTTACCAATCCTATTCCAGCATCGACTCGCCACTGGTGAGTATGCCAAACGTTGAAAACACTATTATTAATGGTCAATCAATTCAGGGTAAAGCTTGTCTGCAGACACTTACCATCAACTTTAGTAATAATCTGGAAGCGATCCGTTGTATCGGATCTGGTAAATACACTCCAGAGTTTTATTTAGAGAAGATGATGGATATCGAAGCGAATGCTTCATTCATGTTCTCGGCCACAGCTGCTGGTT</p>
<i>group_4254</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>TAAAAGGTGATTTGCTGCAGGTGACTGTACCACTGTACCTTACAAGCAGATCATTATTGCTACAGGTGAAGGCGCTAAAGCATCACTCTCTGCGTTTGATTACATCATTCGTTCTGGGCAGTAG</p> <p>ATGTACAAGTTAAAGCTAAATCCTCAGACCAGCGGCTATGGCGTAACACCGGGTGATGATGTGAAACGTCAGCAGATGGATGGCGGTCGTGGTCGCTATTACATCGATGTGAAACGTAACAGCCACATTGTTGATGTGAACTGGAACTTAAGTAAAACCGATTTCAATAAAATGATGGCTTTCTGGCGGGTCTACCAGAACAACCAGCCTCATTTTATGCGGATTTGGTGATTGATCAGGGAACACGTCA</p> <p>GCAATATCTATGCAATTTCACTCCAAACTCGTTCAAGACCAATGAAGTCAACGGCAACTTATATCGTGTGAACGCGCAGCTCGAAGTTGTTCAAACCAGCCTAACCTGAATGCCGATATAGCATTAAATTAAGATTGGGAGGTCTAAATGAGTTCAGGTGCAAAAATTCGATTATATGCTTGTGAAGAAGCTGTTTTAGGAACTACTCCGGCAAATCCGGTCTGGTATACCGTTCGCCGTGTTACTGATAGTTTGACCGAAAATGTTACTACTGAAGATAGCAGTGAAGTAGTTGATCACGTTTTCGCCAAGGTGCTGTAGTAACGGAAGCTGAAGTAACTGGTCAGTTAGAATTTGAATTATCGCTAGGTACCTTTGATTTATTCTTAAATGTTCTCGCTTTCAATAACTGGGCTGCAAAATGCTTTAAGTTTTGGTGGTGGAGTCCGTAA</p> <p>GTCTCTTACCTTGGTAAAAGTCTTTGAAGATATTGGTCAAGTCTTTATTTATCGTGGTATTCAAGTGAATACAGGTGAAATGACAATCCAGACCACGGGGAAATCACTGGTAACTTTGGTTTAGTCGGTAGCTCATTTACACGACAGCAGGCAATCCTGTTACCAATCCTATTCCAGCATCGACTCGCCACTGGTGAGTATGCCAAACGTTGAAAACACTATTATTAATGGTCAATCAATTCAGGGTAAAGCTTGTCTGCAGACACTTACCATCAACTTTAGTAATAATCTGGAAGCGATCCGTTGTATCGGATCTGGTAAATACACTCCAGAGTTTTATTTAGAGAAGATGATGGATATCGAAGCGAATGCTTCATTCATGTTCTCGGCCACAGCTGCTGGTT</p>
<i>group_4286</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>TAAAAGGTGATTTGCTGCAGGTGACTGTACCACTGTACCTTACAAGCAGATCATTATTGCTACAGGTGAAGGCGCTAAAGCATCACTCTCTGCGTTTGATTACATCATTCGTTCTGGGCAGTAG</p> <p>ATGTACAAGTTAAAGCTAAATCCTCAGACCAGCGGCTATGGCGTAACACCGGGTGATGATGTGAAACGTCAGCAGATGGATGGCGGTCGTGGTCGCTATTACATCGATGTGAAACGTAACAGCCACATTGTTGATGTGAACTGGAACTTAAGTAAAACCGATTTCAATAAAATGATGGCTTTCTGGCGGGTCTACCAGAACAACCAGCCTCATTTTATGCGGATTTGGTGATTGATCAGGGAACACGTCA</p> <p>GCAATATCTATGCAATTTCACTCCAAACTCGTTCAAGACCAATGAAGTCAACGGCAACTTATATCGTGTGAACGCGCAGCTCGAAGTTGTTCAAACCAGCCTAACCTGAATGCCGATATAGCATTAAATTAAGATTGGGAGGTCTAAATGAGTTCAGGTGCAAAAATTCGATTATATGCTTGTGAAGAAGCTGTTTTAGGAACTACTCCGGCAAATCCGGTCTGGTATACCGTTCGCCGTGTTACTGATAGTTTGACCGAAAATGTTACTACTGAAGATAGCAGTGAAGTAGTTGATCACGTTTTCGCCAAGGTGCTGTAGTAACGGAAGCTGAAGTAACTGGTCAGTTAGAATTTGAATTATCGCTAGGTACCTTTGATTTATTCTTAAATGTTCTCGCTTTCAATAACTGGGCTGCAAAATGCTTTAAGTTTTGGTGGTGGAGTCCGTAA</p> <p>GTCTCTTACCTTGGTAAAAGTCTTTGAAGATATTGGTCAAGTCTTTATTTATCGTGGTATTCAAGTGAATACAGGTGAAATGACAATCCAGACCACGGGGAAATCACTGGTAACTTTGGTTTAGTCGGTAGCTCATTTACACGACAGCAGGCAATCCTGTTACCAATCCTATTCCAGCATCGACTCGCCACTGGTGAGTATGCCAAACGTTGAAAACACTATTATTAATGGTCAATCAATTCAGGGTAAAGCTTGTCTGCAGACACTTACCATCAACTTTAGTAATAATCTGGAAGCGATCCGTTGTATCGGATCTGGTAAATACACTCCAGAGTTTTATTTAGAGAAGATGATGGATATCGAAGCGAATGCTTCATTCATGTTCTCGGCCACAGCTGCTGGTT</p>

<i>group_4699</i>	hypothetical protein	C2, C6	/	hypothetical protein	<p>ATGTCTGCTGAAACTGTGCCTGTACCTGGCAATGAAGTGTTGCCTGCAGG GTCAGTTGCTGTTGCAGTCACTGTATCGCCATCCACCAGGTTACCTGGGT TTGGTACTGACCAGCTGCCATCTGTGCCTGCTACCACTGTTGCAGTTGTG CCATCAGGATAA</p> <p>ATGCCACCCTTAGCAAGTTTGCCTCATTCTGAGTTGGTTATTCATGTAGGT TCATTGTCTAAAGTTTTTGCACCCGGTATTCGGTTAGGTTACATCGTGGCA GCTCCAACATCATTCAAACCATTACTGAAGATATTTTATTGATTGATCGGC AGGGCAATAATATTACGGAACCTCGCTTTAGCCGATTTAATGCAACGAGGG GAAATCAAGCGACACATCCGTAAAATGAAAAGATTTATCAGTTACGTGCA GACCATGCGTTAGCTGAGTTTCAAAGAGTTTTTGTGAGAGTATTCAGATA CAACCGCCCGCGGGTGAATGGCTTTGTGGTAAAATTTCAGAAATCATT TACTCAAAATCAGCTCATAAAGCTAAAAGAACTCAACATTGATACAGAGCA GAAGTTTAAACAGAAAGAATATCCTAATTGTAGTATCCGCTTTGGTTTTGC AGCTTTATCAGAGAAAGAAATCACCTCTCTAATTGAGACCTTAAATGAAGC TCTTAATAAAAAGAACAGCGGACTTATAA</p>
<i>group_47</i>	<i>gabR_3</i> HTH-type transcriptional regulatory protein GabR	C4, C6	K	helix_turn_helix gluconate operon transcriptional repressor	<p>ATGTCTGCTGAAACTGTGCCTGTACCTGGCAATGAAGTGTTGCCTGCAGG GTCAGTCGCTGTTGCAGTCACTGTATCGCCATCCACCAGGTTACCTGGGT TTGGTACTGACCAGCTGCCATCTGTGCCTGCTACCACTGTTGCAGTTGTG CCATCAGGATAA</p> <p>ATGTCTGCTGAAACTGTGCCTGTACCTGGCAATGAAGTGTTGCCTGCAGG GTCAGTCGCTGTTGCAGTCACTGTATCGCCATCCACCAGGTTACCTGGGT TTGGTACTGACCAGCTGCCATCTGTGCCTGCTACCACTGTTGCAGTTGTG CCATCAGGATAA</p> <p>ATGCCACCCTTAGCAAGTTTGCCTCATTCTGAGTTGGTTATTCATGTAGGT TCATTGTCTAAAGTTTTTGCACCCGGTATTCGGTTAGGTTACATCGTGGCA GCTCCAACATCATTCAAACCATTACTGAAGATATTTTATTGATTGATCGGC AGGGCAATAATATTACGGAACCTCGCTTTAGCCGATTTAATGCAACGAGGG GAAATCAAGCGACACATCCGTAAAATGAAAAGATTTATCAGTTACGTGCA GACCATGCGTTAGCTGAGTTTCAAAGAGTTTTTGTGAGAGTATTCAGATA CAACCGCCCGCGGGTGAATGGCTTTGTGGTAAAATTTCAGAAATCATT TACTCAAAATCAGCTCATAAAGCTAAAAGAACTCAACATTGATACAGAGCA GAAGTTTAAACAGAAAGAATATCCTAATTGTAGTATCCGCTTTGGTTTTGC</p>
<i>group_4701</i>	hypothetical protein	L7, L8, L9, L10	/	hypothetical protein	<p>ATGTCTGCTGAAACTGTGCCTGTACCTGGCAATGAAGTGTTGCCTGCAGG GTCAGTCGCTGTTGCAGTCACTGTATCGCCATCCACCAGGTTACCTGGGT TTGGTACTGACCAGCTGCCATCTGTGCCTGCTACCACTGTTGCAGTTGTG CCATCAGGATAA</p> <p>ATGTCTGCTGAAACTGTGCCTGTACCTGGCAATGAAGTGTTGCCTGCAGG GTCAGTCGCTGTTGCAGTCACTGTATCGCCATCCACCAGGTTACCTGGGT TTGGTACTGACCAGCTGCCATCTGTGCCTGCTACCACTGTTGCAGTTGTG CCATCAGGATAA</p>
<i>group_4702</i>	hypothetical protein	L6, L8, L9, L10	/	hypothetical protein	<p>ATGTCTGCTGAAACTGTGCCTGTACCTGGCAATGAAGTGTTGCCTGCAGG GTCAGTCGCTGTTGCAGTCACTGTATCGCCATCCACCAGGTTACCTGGGT TTGGTACTGACCAGCTGCCATCTGTGCCTGCTACCACTGTTGCAGTTGTG CCATCAGGATAA</p> <p>ATGCCACCCTTAGCAAGTTTGCCTCATTCTGAGTTGGTTATTCATGTAGGT TCATTGTCTAAAGTTTTTGCACCCGGTATTCGGTTAGGTTACATCGTGGCA GCTCCAACATCATTCAAACCATTACTGAAGATATTTTATTGATTGATCGGC AGGGCAATAATATTACGGAACCTCGCTTTAGCCGATTTAATGCAACGAGGG GAAATCAAGCGACACATCCGTAAAATGAAAAGATTTATCAGTTACGTGCA GACCATGCGTTAGCTGAGTTTCAAAGAGTTTTTGTGAGAGTATTCAGATA CAACCGCCCGCGGGTGAATGGCTTTGTGGTAAAATTTCAGAAATCATT TACTCAAAATCAGCTCATAAAGCTAAAAGAACTCAACATTGATACAGAGCA GAAGTTTAAACAGAAAGAATATCCTAATTGTAGTATCCGCTTTGGTTTTGC</p>
<i>group_48</i>	<i>gabR_3</i> HTH-type transcriptional regulatory protein GabR	C2, C6	K	helix_turn_helix gluconate operon transcriptional repressor	<p>ATGTCTGCTGAAACTGTGCCTGTACCTGGCAATGAAGTGTTGCCTGCAGG GTCAGTCGCTGTTGCAGTCACTGTATCGCCATCCACCAGGTTACCTGGGT TTGGTACTGACCAGCTGCCATCTGTGCCTGCTACCACTGTTGCAGTTGTG CCATCAGGATAA</p> <p>ATGCCACCCTTAGCAAGTTTGCCTCATTCTGAGTTGGTTATTCATGTAGGT TCATTGTCTAAAGTTTTTGCACCCGGTATTCGGTTAGGTTACATCGTGGCA GCTCCAACATCATTCAAACCATTACTGAAGATATTTTATTGATTGATCGGC AGGGCAATAATATTACGGAACCTCGCTTTAGCCGATTTAATGCAACGAGGG GAAATCAAGCGACACATCCGTAAAATGAAAAGATTTATCAGTTACGTGCA GACCATGCGTTAGCTGAGTTTCAAAGAGTTTTTGTGAGAGTATTCAGATA CAACCGCCCGCGGGTGAATGGCTTTGTGGTAAAATTTCAGAAATCATT TACTCAAAATCAGCTCATAAAGCTAAAAGAACTCAACATTGATACAGAGCA GAAGTTTAAACAGAAAGAATATCCTAATTGTAGTATCCGCTTTGGTTTTGC</p>

group_49 *gabR_3* HTH-type
 optional regulatory protein
 GabR

L7, L8, L9, L10

K helix_turn_helix gluconate operon transcriptional repressor

group_5 *pcal_2* hypothetical protein

L6, L7, L10

I 3-oxoadipate CoA-transferase activity

group_50 *gabR_3* HTH-type
 optional regulatory protein
 GabR

L6, L7, L8, L10

K helix_turn_helix gluconate operon transcriptional repressor

AGCTTTATCAGAGAAAGAAATCACCTCTCTAATTGAGACCTTAAATGAAGC
 TCTTAATAAAAAGAACAGCGGACTTATAA
 ATGCCACCCTTAGCAAGTTTGCCTCATTCTGAGTTGGTTATTCATGTAGGT
 TCATTGTCTAAAGTTTTTGCACCCGGTATTCGGTTAGGTTACATCGTGGCA
 GCTCCAACATCATTCAAACCATTACTGAAGATATTTTATTGATTGATCGGC
 AGGGCAATAATATTACGGAACCTCGCTTTAGCCGATTTAATGCAACGAGGG
 GAAATCAAGCGACACATCCGTAAAATGAAAAGATTTATCAGTTACGTCTGA
 GACCATGCGTTAGCTGAGTTTCAAAGAGTTTTTGTCTGAGAGTATTCAGATA
 CAACCGCCCGCGGGTGAATGGCTTTGTGGGTAAAATTTTCAGAAATCATT
 TACTCAAAATCAGCTCATAAAGCTAAAAGAACTCAACATTGATACAGAGCA
 GAAGTTTAAACAGAAAGAATATCCTAATTGTAGTATCCGCTTTGGTTTTGC
 AGCTTTATCAGAGAAAGAAATCACCTCTCTAATTGAGACCTTAAATGAAGC
 TCTTAATAAAAAGAACAGCGGACTTATAA
 ATGATTGACAAAAGTAAGTCCTCACTAAGCGAGGTTTTATCGCAAATTTAAA
 GATGGTGCAACCATTCTGATTGGTGGTTTTGGTACCGCAGGACAACCCGC
 TGAATCATTGACGGTCTGATTGATCTTGGCGTTAAGGACCTGACCATTGT
 GAGCAACAATGCCGGTAATGGCGATTATGGTCTGGCTAACTGCTCAAAG
 CCGGTTCTATTAATAAAGTGATTTGCTCTTTCCACGTCAGTCTGACTCTT
 ATGTGTTTGTGAACTGTACCGTGCCGNNNNNNNNNGCCAATAAAGTTG
 TGTCCAATGACCATTTTTATATTCCTTAGAGAGTAA
 ATGCCACCCTTAGCAAGTTTGCCTCATTCTGAGTTGGTTATTCATGTAGGT
 TCATTGTCTAAAGTTTTTGCACCCGGTATTCGGTTAGGTTACATCGTGGCA
 GCTCCAACATCATTCAAACCATTACTGAAGATATTTTATTGATTGATCGGC
 AGGGCAATAATATTACGGAACCTCGCTTTAGCCGATTTAATGCAACGAGGG
 GAAATCAAGCGACACATCCGTAAAATGAAAAGATTTATCAGTTACGTCTGA
 GACCATGCGTTAGCTGAGTTTCAAAGAGTTTTTGTCTGAGAGTATTCAGATA
 CAACCGCCCGCGGGTGAATGGCTTTGTGGGTAAAATTTTCAGAAATCATT
 TACTCAAAATCAGCTCATAAAGCTAAAAGAACTCAACATTGATACAGAGCA
 GAAGTTTAAACAGAAAGAATATCCTAATTGTAGTATCCGCTTTGGTTTTGC
 AGCTTTATCAGAGAAAGAAATCACCTCTCTAATTGAGACCTTAAATGAAGC
 TCTTAATAAAAAGAACAGCGGACTTATAA

group_
54

hypothetical
protein

A2, A3,
A14

D

Phage-related minor tail protein

ATGGCACAAGAATCACGTCTCGTCATTGTAATTGATGCAAAAAATGCAGAA
CGAAATGCGCGCAATCTAGGCAATGAACTGGATAGTATTGAGCGTAAAGG
TGAGTTTGCATCTAAGTCTATGGACAACCTTATCTGTAGCTACGCGAGCACT
AGCTGGGTATATGGCTGGGCTAGTAACAGTAAGTTCTGCCATTTCAAAGA
TGGATACATATACTGGACTACAAAACCGTCTTAAGTTGGTCACTAATAATC
AAGTTGAACTAAATAAAGCAACGGAAGACACTTTCCGAATTGCTCAAAAA
CCTATTCGGCTTGGGATTCTGTGTTACAGGTCTACCAGCGTTTTAGTGATA
ATGCCAAAACCTTTAAACCTCACAATGGATGACACAGCACGTTTAACTGAAA
CAGTTTCTAAAGCTGTAGCAATTAGTGGTGCAAGTGCAAGAAGCAGCTGAT
GCAGCTTTAGTACAATTCGGGCAGGCTTTGGCAAGCGGTACATTACGTGG
TGAAGAACTCAACTCAGTTATGGAACAAACACCAGCTCTAGCAAAGGCTAT
TGCTCAAGGCATGGGGATCACCGTAGGAGAGTTGCGTTCAGTAGCAGCT
GAAGGAAAAATTACTTCACAAGAAATTGAAAAGCGCTTAGAAATGTAGAA
TCTGATGTTGATGCTCTTTTTGCTAAAACAGATATCACAATCGGGCAGTCT
CTCACACTCCTAAACAACGAGATCACAAAATTTGTTGGCGAAGCAGGTAA
GGGAAGTGGTGCGGCACAGGTATTAGCTGGATCAGTTCAAACCTTTGCAA
GTAATTTAGATTTAATTGCTGATGGGGCTTTAGTAGTTGGTATTGGATATAT
CACTCGTGCAATTTTATGATGAAGAGCGCTGCTATTAAGAGGGAATGGCTT
CAACTTTAGCGAGCCGCCAAGCATCTGTATTAATGCTCAAGCAGAATATG
CAGAAGCTACCGCTGCTTTGAATGCAGCAAAAGCTCATCTCGCGAATGTG
CGAGCAACAAATGCAGAAACCCAAGCTAAATTTGGCGAACAGCGGCAG
CAACTCGATACGCACAAGCACAGGCAGCAGTAACTGCTGCTACAAATGCA
CAAACAGCAGCTCAAATTAAGCTAAATACTGCAACTTCAATTGCAGGGAGA
CTAGCTAAAGGGGCGTTTGGATTAATTGGTGGGTGGGCTGGAGTTGCAAC
ATTAGGAGTAATGGGATTAGCGGCAGCCTATTCTTATTTTAATAATAAGGC
AGAGGAGGCAAAGCAAAAGCTTGCTGAACAAGCTAAAGTTGCTGAGAAAG
CTGATGAGGAGTTAAAAAATTAACCTGGCAATGATAAGGCTAAAGCAGTTA
ATGATTTAACTACTGCTTTTAAATGCACAAAATAAAGCATTAGAGAAATCATC
GCGTGCTGTAGGGTCTGCATTAATTGATATCGAGAACTATGCACGAGGAA
ATAGGGAGGTTGAAAAATTTCCCAAGAAGCGAGAAGTGGAACTATCAGC
TATACAGAAGCCATTGAACGTCTAAATAAAATTAAGTTGCCTACAGATCTAT
ATGAAAATCTGAAAAACAGGCTGCGCAGTATGATGACAATGCATCTAAAG
CAAGTTTATCAGCTGAGAACTTAAATTATTAAGAGTTGAAGTGAACTTG

GAGGTAATGAAGCACAAAATGCGGCAATTCAGCATCAAAAACAAGCGGAT
GCTTTAGGAAATACTGCTACTGAAGCAGAAAAGGCAACTAAGGCTTTGCA
AGATTATCAAGCCAAGCAAAAAGATAGCGTTATTGATTCAATCTATAAATCA
GGTTGGCTTGATAAAGGTTACTGTTGCTCAAGCTAATGCCATTTTAGAA
CTGCAAAAAGCTAAAGGAATGAGTGCAATTTTGTCTAAAGATGAAATTGAT
AGCGCACTTAGAAATCTCAAGATCATCGAAGAACAACAGGAGCGAGAAGA
TAAATTAAGTGAAGCTAAAAGAAAGCAGACGCAGGAAATTGAAAAACAAGC
AAAACCTACTAAACGCTTGGTCGGTATTTCCGGTCAATCCGGTATTGGTAC
TGGTCCACATCTTGACGTCGGATATGGTGGCTCATTGTCAGGTCAGAAAG
TTTCTAATGAACATCTGGCTCGATTACAGGCGGGAGGAAAACCTTTAACTT
CCTACAAGATCAGTTCTAATTATGGTCCACGAAAAGCCCCAACTAAAGGG
GCTTCTTCATTTATAAAGGTATTGATTTTTCAATGCCTGAAGGAACACCAA
TCACGACCAATGTTGCTGTGAAAGATATCAAGACATGGTATGACAGCAAG
GGAGGTGGTTATGTCAGTGAAGTATCTTTGAGGATGGAGTGTCTCTTAA
GCTTCTACATCAATCTCCAAGATGCAGAGCAAGGTGAAAGGTGGTGCAA
GTAAAGGAAGTGATAAAGCAGCTGGTATCAATCTCAACTTGAACGTC
AACAGGATTTGCAACGGTCACTTGAAAATGAGGTGGCTAGTGAAGTCGGA
CGGATTAACAATAATAGAAAGGCAAGACTGGAGGATGTTGATAAAGCAAA
CTTTAGCCCCGGAACGTAAGTGCAGAAATAAAGGCGGAAATAAATCGTCGTG
CAGATAATGATATTGCTATAGCCAAACAAGCCCTTAGAACGAAATTGGAAG
ACTATAAGGAGTTCCAGAAAACCGAGGAACAGTTACTAGAAGAGTCCTTTA
ACCGTAAAAAGTTCAATGCAGCTCATGACCTTGAATTAAGTAAGTTTGAGC
AGAAGCAAGCTGTTGAATTGCTGGAACAGCAAAAACAGCAAGAGTTAGGG
TTATTA AAACTAGCTCAGGAACAGCGGTTGTTTCAAGCCCGTTTATCTCTG
CTTTCTGAAACGCAAGCCATGCAGGAACGTTACAGACTCGAACGGGAGGA
AATTCTTAAGAATACCAAGCTTTCTATAGAAGAGCGGCAAAAAGCTAATCGC
ATTATCTAAAGCCAATCAGGATAAAGAGACACGCGATAAAGTGAATAATGC
TGCTCAAACTGGGGTGGTATCCAAGCGGATATGAATGGTACCGGAGAAT
TTTTCAGACAGGATCAGGAACGATTTAGCCGCTTAAATGCTGCAAATGATT
TAGCAGATAGTCAATTTGCTGCTACCGACCTGAATGAGCAAACTCTTTAG
ATGGTTTGAATGCTCAATTCGAAGCTGGACTAATTAAGCAGCAGGATTACG
AAAATCAAAAACGGCTATCATTCAAGCTGCTCAGGACCAACGTAATCAGA
TAGCTGCCGAATATGCTCAGAATGCTCAAGATATTGAAGATAAGTATCACC

group_
55

hypoth
etical
protein

A2, A3,
A14

D

Phage-related minor tail protein

AAGATCGATTGAATGCTCAAATTGCCCTTGGTGGCCAAATGATGGGTTCA
CTCACATCGATGTTTGGCTCAATGTTTGGAGAGCAATCTAAAGCCTACAAA
ATCATGTTTGGCTGCTGATAAAGCTTATGCCATTGCAGCTGCAGGTATTGCG
ATTCAGCAAATATTGCAGCAGCTTCAAAGCTGGTTTTCCACTTAACATT
CCATTAATTGCTGGGGCGGTTGCTCAAGGCGCTAGCATTATTGCAAACAT
CCGGGCAATTAAGGATCAAGGCTTTGCTGATGGTGGTTACACTGGTTCGAG
GTGGGAAATATGAAGTTGCTGGAGCTGTGCACAAAGGCGAGATTGTATGG
TCCCAAGAAGACATTAACGTTGGGGTGGTGGTTAGTTGAGAGAAT
GCGTAAGAGTGCAAACCCTGAAGCTTTTCTCAATAACAATGCCTCGACTG
ATAGTGTTCATGCGCCGTGCAATGATGAGCTCTAATGCCTTTATAGAAAGC
CAAAGCAGGCTGACATCTTTAATCAACCGGTTCAAGATACTCAGATTATC
TATAAGGGTAATAGAGACACACCTAAATTAGCTTCTTCTGGAAATTCAGAC
TTATTCCATGATGGCAAGGTCTACTTCTCATCCAATGGTTTAGTTCAGGAT
CGTTCAAATCTGGATGATGTTTCAGGACTTTACTTTAGGAAGTACTTCACGC
CCTCAAGCTGAGATTATGCCTTCAATTGAGCCTGCTTCACCGACAATCAAT
TTCAAATTGAAGTGATTAATCAGGTGAGTGGGGCGACAGTTGAAGCCGA
ACAACCTGGACGAGCAAACAGTCCGGATCATTGTTAAAGATGAACTGGATA
AGCAGCTTCCAAGAACGGTACCTAAGCTTGTAAGTGATCAAATCGCAAAT
CCAAACTCAACCATTAGTCGGTCTTTGACTGAGAATACGACAGCGAGAAG
AAATCGTACTTAA
ATGGCACAAGAATCACGTCTCGTCATTGTAATTGATGCAAAAAATGCAGAA
CGAAATGCGCGCAATCTAGGCAATGAACTGGATAGTATTGAGCGTAAAGG
TGAGTTTGCATCTAAGTCTATGGACAACCTTATCTGTAGCTACGCGAGCACT
AGCTGGGTATATGGCTGGGCTAGTAACAGTAAGTTCTGCCATTTCAAAGA
TGGATACATATACTGGACTACAAAACCGTCTTAAGTTGGTCACTAATAATC
AAGTTGAACTAAATAAAGCAACGGAAGACACTTTCCGAATTGCTCAAAAAA
CCTATTCGGCTTGGGATTCTGTGTTACAGGTCTACCAGCGTTTTAGTGATA
ATGCCAAAACCTTAAACCTCACAATGGATGACACAGCACGTTTAACTGAAA
CAGTTTCTAAAGCTGTAGCAATTAGTGGTGAAGTGCAGAAGCAGCTGAT
GCAGCTTTAGTACAATTCGGGCAGGCTTTGGCAAGCGGTACATTACGTGG
TGAAGAACTCAACTCAGTTATGGAACAAACACCAGCTCTAGCAAAGGCTAT
TGCTCAAGGCATGGGGATCACCGTAGGAGAGTTGCGTTTCAGTAGCAGCT
GAAGGAAAAATTACTTCACAAGAAATTGTAAGGCGCTTAGAAATGTAGAA

TCTGATGTTGATGCTCTTTTTGCTAAAACAGATATCACAATCGGGCAGTCT
CTCACACTCCTAAACAACGAGATCACAAAATTTGTTGGCGAAGCAGGTAA
GGGAAGTGGTGCGGCACAGGTATTAGCTGGATCAGTTCAAACTCTTGCAA
GTAATTTAGATTTAATTGCTGATGGGGCTTTAGTAGTTGGTATTGGATATAT
CACTCGTGCAATTTTATGAAGAGCGCTGCTATTAAGAGGGGAATGGCTT
CAACTTTAGCGAGCCGCCAAGCATCTGTATTAATGCTCAAGCAGAATATG
CAGAAGCTACCGCTGCTTTGAATGCAGCAAAGCTCATCTCGCGAATGTG
CGAGCAACAAATGCAGAAACCCAAGCTAAATTTGGCGCAACAGCGGCAG
CAACTCGATACGCACAAGCACAGGCAGCAGTAACTGCTGCTACAAATGCA
CAAACAGCAGCTCAAATTAAGCTAAATACTGCAACTTCAATTGCAGGGAGA
CTAGCTAAAGGGGCGTTTGGATTAATTGGTGGGTGGGCTGGAGTTGCAAC
ATTAGGAGTAATGGGATTAGCGGCAGCCTATTCTTATTTTAATAATAAGGC
AGAGGAGGCAAAGCAAAGCTTGCTGAACAAGCTAAAGTTGCTGAGAAAG
CTGATGAGGAGTTAAAAAATTAAGTGGCAATGATAAGGCTAAAGCAGTTA
ATGATTTAACTACTGCTTTTAATGCACAAAATAAAGCATTAGAGAAATCATC
GCGTGCTGTAGGGTCTGCATTAATTGATATCGAGAACTATGCACGAGGAA
ATAGGGAGGTTGAAAAATTTCCCAAGAAGCGAGAAGTGGAACTATCAGC
TATACAGAAGCCATTGAACGTCTAAATAAAATTAAGTTGCCTACAGATCTAT
ATGAAAATCTGAAAAACAGGCTGCGCAGTATGATGACAATGCATCTAAAG
CAAGTTTATCAGCTGAGAACTTAAATTATTAAGAGTTGAAGTGAAGCTTG
GAGGTAATGAAGCACAAAATGCGGCAATTCAGCATCAAAAACAAGCGGAT
GCTTTAGGAAATACTGCTACTGAAGCAGAAAAGGCAACTAAGGCTTTGCA
AGATTATCAAGCCAAGCAAAGATAGCGTTATTGATTCAATCTATAAATCA
GGTTGGCTTGATAAAGGTTACACTGTTGCTCAAGCTAATGCCATTTTAGAA
CTGCAAAAAGCTAAAGGAATGAGTGCAATTTTGTCTAAAGATGAAATTGAT
AGCGCACTTAGAAATCTCAAGATCATCGAAGAACAACAGGAGCGAGAAGA
TAAATTAAGTGAAGCTAAAAGAAAGCAGACGCAGGAAATTGAAAAACAAGC
AAAACCTACTAAACGCTTGGTCGGTATTTCCGGTCAATCCGGTATTGGTAC
TGGTCCACATCTTGACGTCCGATATGGTGGCTCATTGTCAGGTCAGAAAG
TTTCTAATGAACATCTGGCTCGATTACAGGCGGGAGGAAAACCTTTAACTT
CCTACAAGATCAGTTCTAATTATGGTCCACGAAAAGCCCCAACTAAAGGG
GCTTCTTCATTTCATAAAGGTATTGATTTTTCAATGCCTGAAGGAACACCAA
TCACGACCAATGTTGCTGTGAAAGATATCAAGACATGGTATGACAGCAA

GGAGGTGGTTATGTCAGTGAAGTGATCTTTGAGGATGGAGTGTCTCTTAA
GCTTCTACATCAATCTCCCAAGATGCAGAGCAAGGTGAAAGGTGGTGCAA
GTAAAGGAAGTGATAAAGCAGCTGGTGATATTCAATCTCAACTTGAACGTC
AACAGGATTTGCAACGGTCACCTGAAAATGAGGTGGCTAGTGAAGTCGGA
CGGATTAACAATAATAGAAAGGCAAGACTGGAGGATGTTGATAAAGCAAA
CTTTAGCCCGGAACGTA CTGCAGAAATAAAGGCGGAAATAAATCGTCGTG
CAGATAATGATATTGCTATAGCCAAACAAGCCCTTAGAACGAAATTGGAAG
ACTATAAGGAGTTCCAGAAAACCGAGGAACAGTTACTAGAAGAGTCCTTTA
ACCGTAAAAAGTTCAATGCAGCTCATGACCTTGAATTAAGTAAGTTTGAGC
AGAAGCAAGCTGTTGAATTGCTGGAACAGCAAAAACAGCAAGAGTTAGGG
TTATTA AAACTAGCTCAGGAACAGCGGTTGTTTCAAGCCCGTTTATCTCTG
CTTTCTGAAACGCAAGCCATGCAGGAACGTTACAGACTCGAACGGGAGGA
AATTCTTAAGAATACCAAGCTTTCTATAGAAGAGCGGCAAAAGCTAATCGC
ATTATCTAAAGCCAATCAGGATAAAGAGACACGCGATAAAGTGAATAATGC
TGCTCAAAACTGGGGTGGTATCCAAGCGGATATGAATGGTACCGGAGAAT
TTTTCAGACAGGATCAGGAACGATTTAGCCGCTTAAATGCTGCAAATGATT
TAGCAGATAGTCAATTTGCTGCTACCGACCTGAATGAGCAAAACTCTTTAG
ATGGTTTGAATGCTCAATTCGAAGCTGGACTAATTAAGCAGCAGGATTACG
AAAATCAAAAACCGGCTATCATTCAAGCTGCTCAGGACCAACGTAATCAGA
TAGCTGCCGAATATGCTCAGAATGCTCAAGATATTGAAGATAAGTATCACC
AAGATCGATTGAATGCTCAAATTGCCCTTGGTGGCCAAATGATGGGTTCA
CTCACATCGATGTTTGGCTCAATGTTTGGAGAGCAATCTAAAGCCTACAAA
ATCATGTTTGCTGCTGATAAAGCTTATGCCATTGCAGCTGCAGGTATTGCG
ATTCAGCAAAATATTGCAGCAGCTTCAAAAGCTGGTTTTCCACTTAACATT
CCATTAATTGCTGGGGCGGTTGCTCAAGGCGCTAGCATTATTGCAAACAT
CCGGGCAATTAAGGATCAAGGCTTTGCTGATGGTGGTTACACTGGTCGAG
GTGGGAAATATGAAGTTGCTGGAGCTGTGCACAAAGGCGAGATTGTATGG
TCCCAAGAAGACATTAACGTTGGGGTGGTGGTTAGTTGAGAGAAT
GCGTAAGAGTGCAAACCCTGAAGCTTTTCTCAATAACAATGCCTCGACTG
ATAGTGTCATGCGCCGTGCAATGATGAGCTCTAATGCCTTTATAGAAAGC
CAAAAGCAGGCTGACATCTTTAATCAACCGGTTCAAGATACTCAGATTATC
TATAAGGGTAATAGAGACACACCTAAATTAGCTTCTTCTGGAAATTCAGAC
TTATTCCATGATGGCAAGGTCTACTTCTCATCCAATGGTTTAGTTTCAGGAT

group_
57

hypothetical
protein

A2, A3,
A14

/

hypothetical protein

CGTTCAAATCTGGATGATGTTTCAGGACTTTACTTTAGGAAGTACTTCACGC
CCTCAAGCTGAGATTATGCCTTCAATTGAGCCTGCTTCACCGACAATCAAT
TTCAAATTGAAGTGATTAATCAGGTGAGTGGGGCGACAGTTGAAGCCGA
ACAACGGACGAGCAAACAGTCCGGATCATTGTTAAAGATGAACTGGATA
AGCAGCTTCCAAGAACGGTACCTAAGCTTGTAAGTGATCAAATCGCAAAT
CCAACTCAACCATTAGTCGGTCTTTGACTGAGAATACGACAGCGAGAAG
AAATCGTACTTAA

ATGAATAATCTAATCACTGCAGCTGAAGCATTGTCAGCTCTTCAAAAAGGT
AAAACGTCTTTGTCGTCTATTGGAGACATGTTGGACTTTTCTGACTTA
GATCAATCCCCGCTTCTGTTTTTGGTAAACCGGGTTTTGAATTCTGCATC
AAAATCGAACTATTGAGCTGGCTGGCATTACATTCACAAAGCCATTA
ATTGATGAGTATGAAGAGGGTCAGGAAGTTTATGTAATCAGTACATATTCA
CCTACGGTTTATGTTTAGATTTCAAACTAACGCATTAATTGATTCTATTA
ACAGTGGCTTCGTTCAACGTGATGCAGAAAACGCCAAGCTTCAATTA
GCACTGTCCAAAGCGTTAGGTTTTGAAGTTAATGATGACTTAAGTGTATT
CGTCTTGGTGAGGAACCTAAAAACAGAGAGGCCAAAAAATCAAAAGCTGA
AAAGCCTAGTGAAGTTATTTCTGCAGAACTCAACCAACAATTGTTATTAC
CGAAAAACAAATGTCATCACATCTGAAGATCTGTTAGTTCCAGAACTAA
CGAGCCTAAAGTAGATCCTGAATATCAAATAAACTTAATACCCTGCTGCA
AAGAGTTAAAGACTCAAAACACCAGATGAAGTAAATGCGGTTTATCGTTA
TACCCGCACATGGGATGACGAACAAATGAAGCCTATCCTTCTCGCCACTC
ACAAACGTCTTGAAGAGCTAGAAAAAAGGCATCTGCTAATGAGCCA
CCCTCTTAATGGTTCTAATCCAACTGCACCAGACCTTACAACGCTAGAT
GCTTTGAAATAGACGTGGCTGCACGAGATCCGCAGATTCAACCGAAGCT
AATGGGCTATATAAGAAAACGCCGCTATGAATTAGAAAATCCAGCAGTTTC
TCAACCAGAAGCAGAGCCTGATTTTCTATTAGGAGACGGTTTCTAG

group_
58

hypothetical
protein

A1

/

hypothetical protein

ATGTTGGACTTTTCTGACTTAGATCAATCCCCGCTTCTGTTTTTGGTAAAC
CGGGTTTTGAATTCTGCATCAAATCGAACTATTGAGCTGGCTGGCATT
CATTACAAAGCCATTA
ACTATTGATGAGTATGAAGAGGGTCAGGAAGTTT
ATGTAATCAGTACATATTACCTACGGTTTATGTTTTAGATTTCAAACTAA
CGCATTAAATTGATTCTATTAACAGTGGCTTCGTTCAACGTGATGCAGAAAA
CGCCAAGCTTCAATTA
AAAGCACTGTCCAAAGCGTTAGGTTTTGAAGTTAA
TGATGACTTAAGTGTATTTCGTCTTGGTGAGGAACCTAAAAACAGAGAGG

group_
59

hypothetical
protein

A2, A14 /

hypothetical protein

group_
60

hypothetical
protein

A3 /

hypothetical protein

CAAAAAATCAAAGCTGAAAAGCCTAGTGAAGTTATTTCTGCAGAACTCA
ACCAACAATTGTTATTACCGAAAAACAAATGTCATCACATCTGAAGATCT
GTTAGTTCAGAACTAACGAGCCTAAAGTAGATCCTGAATATCAAATAA
ACTTAATACCCTGCTGCAAAGAGTTAAAGACTCAAAAACACCAGATGAAGT
AAATGCGGTTTATCGTTATACCCGCACATGGGATGACGAACAAATGAAGC
CTATCCTTCTCGCCACTCACAAACGCTTTGAAGAGCTAGAAAAAAAAGG
CATCTGCTAATGAGCCACCCTCTTTAATGGTTCTAATCCAACTGCACCAG
ACCTTACAACGCTAGATGCTTTGGAAATAGACGTGGCTGCACGAGATCCG
CAGATTCAACCGAAGCTAATGGGCTATATAAGAAAACGCCGCTATGAATTA
GAAAATCCAGCAGTTTCTCAACCAGAAGCAGAGCCTGATTTTCTATTAGGA
GACGGTTTCTAG
ATGCATGAGATTACCTTGAACGAAGTGCCTCAATTAATCGCTTCTCTTCGC
ACTGTTTACGCTGCTCAGTTCAATAAGCAATTTCCAGCAACAGGCGAAAAGT
GCAATTCCTCTGTCAGTGGTTGAGCAAATCGCACTTAAACACTGGTTGG
CGTTCAACAAAACCAATTTAACACGCACTTGCTCGATTACTTACAGCAGG
TGGACGTTTTATGCCGTCATTTGCTGAGTTTCGCACCTGGTGTATTGGTGA
AAGTTGGATGTCTCCAGAGGAAGCTTGGTCACGTGCATGTAAGTTTACGA
CTGACAGTACCGTGGTTATTACACAAATTACAAAATATGCATTAGACGAAG
TGATGTATTTGATCGAAGCCGGCCAAATGCGAGCAGCTCAAGATAATTTCT
TCGGAACCTATAACGTGATGGTGGCTAAAGCTCAATTGAAAGGTCGTCAG
CAAGAGTTTTACGCTCCACCGCTACAACACTAGAACACAAAGAACCTAACAC
GTTCCCTGTGAGCAATGACGAGGCTCAAAGCATCTCAAATCATTGATGGA
AAGATTAATAATCAATGGTCGTAAACCTGCACCAGTTCAAAAACCTGAGGC
AAAAGAAAAGAGCCTGAGCTTGCAAAAAGAATTAG
ATGCATGAGATTACCTTGAACGAAGTGCCTCAATTAATCGCTTCTCTTCGC
ACTGTTTACGCTGCTCAGTTCAATAAGCAATTTCCAGCAACAGGCGAAAAGT
GCAATTCCTCTGTCAGTGGTTGAGCAAATCGCACTTAAACACTGGTTGG
CGTTCAACAAAACCAATTTAACACGCACTTGCTCGATTACTTACAGCAGG
TGGACGTTTTATGCCGTCATTTGCTGAGTTTCGCACCTGGTGTATTGGTGA
AAGTTGGATGTCTCCAGAGGAAGCTTGGTCACGTGCATGTAAGTTTACGA
CTGACAGTACCGTGGTTATTACACAAATTACAAAATATGCATTAGACGAAG
TGATGTATTTGATCGAAGCCGGCCAAATGCGAGCAGCTCAAGATAATTTCT
TCGGAACCTATAACGTGATGGTGGCTAAAGCTCAATTGAAAGGTCGTCAG

<i>group_622</i>	hypothetical protein	A1	S	Ankyrin repeats (many copies)	CAAGAGTTTTACGCTCCACCGCTACAAC TAGAACACAAAGAACCTAAACAC GTTCTGTGAGCAATGACGAGGCTCAA AAGCATCTCAAATCATTGATGGA AAGATTA AAAATCAATGGTCGTAAACCTGCACCAGTTCAAAAACCTTGAGGC AAAAGAAAAAGAGCCTGAGCTTGCAA AAGAATTAGGTCCAGATCCTTTTCG ACAATCCGCACGAATACGCTGAGATGTGCCGCCGTGAAGGTATGCCAATA CCTAGAAATATTCTTCAGCTAATTGATGGGGCGAATGCATGA GTGTTTTTTTATATGGGAAAGCAATCAATAAGAATTATTTAAATATAGAAC TACAAATCATTTTTGGATGCTTACTCGATACTATGGATATAATGAGGAGGT TTTTATTTTTGGTGTCCCCGTGAAAATTA ACTCTTTCTCTCCCATACCTGA AGACGATGAAGA ACTCTCTCTTCCAGAACTGATTTATTGGGCATCTTTAGG TGATCTAGAGCAAGTAAACAATTATTAATAGATGGAGAAGACCCTAATCA GACTGATGATGAAGGGTATAGTGCTTTGCAAGCAGCGGCTGAGAATGATC ATTTAGATATAGTGAAGCTATTAGTTGAAAAAGGGCGCATGTCCCTTATA AAAGCGAATATACAGCACTACAGCTTGCAGAAATGGCGGGTAATACTGAC ATAGTGAATTATTTAAAAAGTTTATAA ATGATGTTGTGGCTCTTAATGAAAAATGGCGCTTCTAAAGATGAGGTTAGA ATGATACAAGCTAAAGGTTATTCATATGCAGATGATAACTATAAGAGGGTA AAGTAG ATGGATAAACTGAACCCA ACTACAGTGATTAATATTCATATCTCATTGCTT TTGGACTATTCATTATTGCTGGTCTGGTGAGCTCAAATATTTCTTGTTAGG TCTTACTATCTTCTTAATTGGTGCTTTACTCGCTGGTGACAGTCTTCATTA CTGCCACTTGCAGCTATGTTCTATCCAGCAGTATGCCGCGCGGTTGGAGT TTCATGGATGCATGGTATCGGTCGTATCGGTGCAATTTTAGGCGCTTTCTT TGGTTCGCTCATTTTCACTTTCAACCTAAGCTTAAGCGGTATTTTCTTTATT CTAGCAATACCAACCTTCATCTCATTTATTGCACTTAGCCTGAAAGTAATTT ATGAGAAATCTAAAAACAACAAGTGTTAAAGCTAGAAGAAAGTCTATAA
<i>group_624</i>	hypothetical protein	A1, A2, A14	/	hypothetical protein	ATGGATAAACTGAACCCA ACTACAGTGATTAATATTCATATCTCATTGCTT TTGGACTATTCATTATTGCTGGTCTGGTGAGCTCAAATATTTCTTGTTAGG TCTTACTATCTTCTTAATTGGTGCTTTACTCGCTGGTGACAGTCTTCATTA CTGCCACTTGCAGCTATGTTCTATCCAGCAGTATGCCGCGCGGTTGGAGT TTCATGGATGCATGGTATCGGTCGTATCGGTGCAATTTTAGGCGCTTTCTT TGGTTCGCTCATTTTCACTTTCAACCTAAGCTTAAGCGGTATTTTCTTTATT CTAGCAATACCAACCTTCATCTCATTTATTGCACTTAGCCTGAAAGTAATTT ATGAGAAATCTAAAAACAACAAGTGTTAAAGCTAGAAGAAAGTCTATAA
<i>group_64</i>	<i>mhbT</i> 3-hydroxybenzoate transporter MhbT	L6, L8, L9, L10	EGP	Uncharacterised MFS-type transporter YbfB	ATGGATAAACTGAACCCA ACTACAGTGATTAATATTCATATCTCATTGCTT TTGGACTATTCATTATTGCTGGTCTGGTGAGCTCAAATATTTCTTGTTAGG TCTTACTATCTTCTTAATTGGTGCTTTACTCGCTGGTGACAGTCTTCATTA CTGCCACTTGCAGCTATGTTCTATCCAGCAGTATGCCGCGCGGTTGGAGT TTCATGGATGCATGGTATCGGTCGTATCGGTGCAATTTTAGGCGCTTTCTT TGGTTCGCTCATTTTCACTTTCAACCTAAGCTTAAGCGGTATTTTCTTTATT CTAGCAATACCAACCTTCATCTCATTTATTGCACTTAGCCTGAAAGTAATTT ATGAGAAATCTAAAAACAACAAGTGTTAAAGCTAGAAGAAAGTCTATAA
<i>group_66</i>	<i>mhbT</i> 3-hydroxybenzoate transporter MhbT	L7, L8, L9, L10	EGP	Uncharacterised MFS-type transporter YbfB	ATGGATAAACTGAACCCA ACTACAGTGATTAATATTCATATCTCATTGCTT TTGGACTATTCATTATTGCTGGTCTGGTGAGCTCAAATATTTCTTGTTAGG TCTTACTATCTTCTTAATTGGTGCTTTACTCGCTGGTGACAGTCTTCATTA CTGCCACTTGCAGCTATGTTCTATCCAGCAGTATGCCGCGCGGTTGGAGT TTCATGGATGCATGGTATCGGTCGTATCGGTGCAATTTTAGGCGCTTTCTT TGGTTCGCTCATTTTCACTTTCAACCTAAGCTTAAGCGGTATTTTCTTTATT

<i>group_664</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	CTAGCAATACCAACCTTCATCTCATTTATTGCACTTAGCCTGAAAGTAATTT ATGAGAAATCTAAAAACAAACAAGTGTTAAAGCTAGAAGAAAGTCTATAA ATGAGAAAACCTTGTTTCTATCATCATTTTGCTGGGCATCGCTTGGGTGGCT AAGCTCAGCTACGATATGTGGCAAATTTCTCGTACCGTCCCTGAACTTCAG CAAAGCTTGCTGCAATCTGAACAGCAATATGCCAATTTGAATGATCAATTG GTTGCTTTGCAAAGACAAATTCAAAATCAACCAAGTCATAATTCTAAAAACA ACGCCTTTAGCAACGACAGAAGTTGTTCAAACCGGGAATTGCACCGACCG TATTAA ATGGATAAACTGAACCCAACTACAGTGATTAATATTCATATCTCATTGCTT TTGGACTATTCATTATTGCTGGTCTGGTGAGCTCAAATATTTTCTTGTTAGG TCTTACTATCTTCTTAATTGGTGCTTTACTCGCTGGTGCACAGTCTTCATTA CTGCCACTTGACAGCTATGTTCTATCCAGCAGTATGCCGCGCGGTTGGAGT TTCATGGATGCATGGTATCGGTCGTATCGGTGCAATTTTAGGCGCTTTCTT TGGTTCGCTCATTTTCACTTTCAACCTAAGCTTAAGCGGTATTTTCTTTATT CTAGCAATACCAACCTTCATCTCATTTATTGCACTTAGCCTGAAAGTAATTT ATGAGAAATCTAAAAACAAACAAGTGTTAAAGCTAGAAGAAAGTCTATAA ATGGATAAACTGAACCCAACTACAGTGATTAATATTCATATCTCATTGCTT TTGGACTATTCATTATTGCTGGTCTGGTGAGCTCAAATATTTTCTTGTTAGG TCTTACTATCTTCTTAATTGGTGCTTTACTCGCTGGTGCACAGTCTTCATTA CTGCCACTTGACAGCTATGTTCTATCCAGCAGTATGCCGCGCGGTTGGAGT TTCATGGATGCATGGTATCGGTCGTATCGGTGCAATTTTAGGCGCTTTCTT TGGTTCGCTCATTTTCACTTTCAACCTAAGCTTAAGCGGTATTTTCTTTATT CTAGCAATACCAACCTTCATCTCATTTATTGCACTTAGCCTGAAAGTAATTT ATGAGAAATCTAAAAACAAACAAGTGTTAAAGCTAGAAGAAAGTCTATAA ATGACGTTATCAAATCCGGCGGCCGTTAATGCTAACTGGTTAAAGCTTAAA GAGTTAAATAACCAACAAACTGCAGCTTTGTCAACAAAGATCAGTTCAATA GGTGAAAATGGATTGGTCCATTCACTTCTGGACAATGTTGAACGTAGCCG GGTGCTGTGTAATTTTGTGGCCACCAATGGACTTTATACCGTTTCTTCAGG TGGGTATCAGGCCGCTGTTAATGTCACGATTGAAGTTGAAGTAACACCGG TAAATGAATCTGGTGCAGCGATTGGTAATCCGATGCTGAAGCAGATCATT TGAAAGGTTCCGGCAAAGTCACGTCAGACCGTTGGTGCAACACTTGATATG GTCACGTTTTCAGGGGCGTTGTAGTGTCCGTGCACGCCGTTTAACTCCGAC TCCGACAGTCACAACAGTTGTTGATGAAGTAAAGTGGCAGGCGCTTTACG
<i>group_67</i>	<i>mhbT</i> 3-hydroxybenzoate transporter MhbT	L6, L7, L8, L10	EGP	Uncharacterised MFS-type transporter YbfB	ATGGATAAACTGAACCCAACTACAGTGATTAATATTCATATCTCATTGCTT TTGGACTATTCATTATTGCTGGTCTGGTGAGCTCAAATATTTTCTTGTTAGG TCTTACTATCTTCTTAATTGGTGCTTTACTCGCTGGTGCACAGTCTTCATTA CTGCCACTTGACAGCTATGTTCTATCCAGCAGTATGCCGCGCGGTTGGAGT TTCATGGATGCATGGTATCGGTCGTATCGGTGCAATTTTAGGCGCTTTCTT TGGTTCGCTCATTTTCACTTTCAACCTAAGCTTAAGCGGTATTTTCTTTATT CTAGCAATACCAACCTTCATCTCATTTATTGCACTTAGCCTGAAAGTAATTT ATGAGAAATCTAAAAACAAACAAGTGTTAAAGCTAGAAGAAAGTCTATAA ATGGATAAACTGAACCCAACTACAGTGATTAATATTCATATCTCATTGCTT TTGGACTATTCATTATTGCTGGTCTGGTGAGCTCAAATATTTTCTTGTTAGG TCTTACTATCTTCTTAATTGGTGCTTTACTCGCTGGTGCACAGTCTTCATTA CTGCCACTTGACAGCTATGTTCTATCCAGCAGTATGCCGCGCGGTTGGAGT TTCATGGATGCATGGTATCGGTCGTATCGGTGCAATTTTAGGCGCTTTCTT TGGTTCGCTCATTTTCACTTTCAACCTAAGCTTAAGCGGTATTTTCTTTATT CTAGCAATACCAACCTTCATCTCATTTATTGCACTTAGCCTGAAAGTAATTT ATGAGAAATCTAAAAACAAACAAGTGTTAAAGCTAGAAGAAAGTCTATAA ATGACGTTATCAAATCCGGCGGCCGTTAATGCTAACTGGTTAAAGCTTAAA GAGTTAAATAACCAACAAACTGCAGCTTTGTCAACAAAGATCAGTTCAATA GGTGAAAATGGATTGGTCCATTCACTTCTGGACAATGTTGAACGTAGCCG GGTGCTGTGTAATTTTGTGGCCACCAATGGACTTTATACCGTTTCTTCAGG TGGGTATCAGGCCGCTGTTAATGTCACGATTGAAGTTGAAGTAACACCGG TAAATGAATCTGGTGCAGCGATTGGTAATCCGATGCTGAAGCAGATCATT TGAAAGGTTCCGGCAAAGTCACGTCAGACCGTTGGTGCAACACTTGATATG GTCACGTTTTCAGGGGCGTTGTAGTGTCCGTGCACGCCGTTTAACTCCGAC TCCGACAGTCACAACAGTTGTTGATGAAGTAAAGTGGCAGGCGCTTTACG
<i>group_68</i>	<i>mhbT</i> 3-hydroxybenzoate transporter MhbT	L6, L7, L8, L9	EGP	Uncharacterised MFS-type transporter YbfB	ATGGATAAACTGAACCCAACTACAGTGATTAATATTCATATCTCATTGCTT TTGGACTATTCATTATTGCTGGTCTGGTGAGCTCAAATATTTTCTTGTTAGG TCTTACTATCTTCTTAATTGGTGCTTTACTCGCTGGTGCACAGTCTTCATTA CTGCCACTTGACAGCTATGTTCTATCCAGCAGTATGCCGCGCGGTTGGAGT TTCATGGATGCATGGTATCGGTCGTATCGGTGCAATTTTAGGCGCTTTCTT TGGTTCGCTCATTTTCACTTTCAACCTAAGCTTAAGCGGTATTTTCTTTATT CTAGCAATACCAACCTTCATCTCATTTATTGCACTTAGCCTGAAAGTAATTT ATGAGAAATCTAAAAACAAACAAGTGTTAAAGCTAGAAGAAAGTCTATAA ATGACGTTATCAAATCCGGCGGCCGTTAATGCTAACTGGTTAAAGCTTAAA GAGTTAAATAACCAACAAACTGCAGCTTTGTCAACAAAGATCAGTTCAATA GGTGAAAATGGATTGGTCCATTCACTTCTGGACAATGTTGAACGTAGCCG GGTGCTGTGTAATTTTGTGGCCACCAATGGACTTTATACCGTTTCTTCAGG TGGGTATCAGGCCGCTGTTAATGTCACGATTGAAGTTGAAGTAACACCGG TAAATGAATCTGGTGCAGCGATTGGTAATCCGATGCTGAAGCAGATCATT TGAAAGGTTCCGGCAAAGTCACGTCAGACCGTTGGTGCAACACTTGATATG GTCACGTTTTCAGGGGCGTTGTAGTGTCCGTGCACGCCGTTTAACTCCGAC TCCGACAGTCACAACAGTTGTTGATGAAGTAAAGTGGCAGGCGCTTTACG
<i>group_7</i>	hypothetical protein	L6, L7, L8, L10	/	hypothetical protein	CTAGCAATACCAACCTTCATCTCATTTATTGCACTTAGCCTGAAAGTAATTT ATGAGAAATCTAAAAACAAACAAGTGTTAAAGCTAGAAGAAAGTCTATAA ATGAGAAAACCTTGTTTCTATCATCATTTTGCTGGGCATCGCTTGGGTGGCT AAGCTCAGCTACGATATGTGGCAAATTTCTCGTACCGTCCCTGAACTTCAG CAAAGCTTGCTGCAATCTGAACAGCAATATGCCAATTTGAATGATCAATTG GTTGCTTTGCAAAGACAAATTCAAAATCAACCAAGTCATAATTCTAAAAACA ACGCCTTTAGCAACGACAGAAGTTGTTCAAACCGGGAATTGCACCGACCG TATTAA ATGGATAAACTGAACCCAACTACAGTGATTAATATTCATATCTCATTGCTT TTGGACTATTCATTATTGCTGGTCTGGTGAGCTCAAATATTTTCTTGTTAGG TCTTACTATCTTCTTAATTGGTGCTTTACTCGCTGGTGCACAGTCTTCATTA CTGCCACTTGACAGCTATGTTCTATCCAGCAGTATGCCGCGCGGTTGGAGT TTCATGGATGCATGGTATCGGTCGTATCGGTGCAATTTTAGGCGCTTTCTT TGGTTCGCTCATTTTCACTTTCAACCTAAGCTTAAGCGGTATTTTCTTTATT CTAGCAATACCAACCTTCATCTCATTTATTGCACTTAGCCTGAAAGTAATTT ATGAGAAATCTAAAAACAAACAAGTGTTAAAGCTAGAAGAAAGTCTATAA ATGACGTTATCAAATCCGGCGGCCGTTAATGCTAACTGGTTAAAGCTTAAA GAGTTAAATAACCAACAAACTGCAGCTTTGTCAACAAAGATCAGTTCAATA GGTGAAAATGGATTGGTCCATTCACTTCTGGACAATGTTGAACGTAGCCG GGTGCTGTGTAATTTTGTGGCCACCAATGGACTTTATACCGTTTCTTCAGG TGGGTATCAGGCCGCTGTTAATGTCACGATTGAAGTTGAAGTAACACCGG TAAATGAATCTGGTGCAGCGATTGGTAATCCGATGCTGAAGCAGATCATT TGAAAGGTTCCGGCAAAGTCACGTCAGACCGTTGGTGCAACACTTGATATG GTCACGTTTTCAGGGGCGTTGTAGTGTCCGTGCACGCCGTTTAACTCCGAC TCCGACAGTCACAACAGTTGTTGATGAAGTAAAGTGGCAGGCGCTTTACG

GTGCTTATCCTTTACAAAGCACAGTGTATGAACATGAAACGGTTTTTCGTG
CGCGTACTTATGCAACCACTGGAGCTTTATCTGTTAAGTCCCGCAAGATCA
ATTTTGATCTTCAGCGAATGTTGCCGACTTATAAAAACGGGGCAATGACAA
CAGAGCTATATCCAACGTCTAGCTTTGCTGATGCTTTGGTATCTATGGCAC
TCGATGACAAGATTGGCCGCCGTTTCGATCGATGAGATTGATCTTGAAAAC
ATCTATCGGACCTATAATGATGTAGTTGATTATTTTGGTACGCCGCTAGCG
GCTGAGTTCTGTACTACCATTGATGATACGAATCTATCTTTTGAAGAGCTG
GTTACCAATCTTTGTGATGCGGTGTTTTGTA CTGCATATCGGCAAAACAAT
AAGCTCAAGCTTTATTTTGAACGGCCA ACTGATAACTCGGTAATGCTGTTT
AACTTCAGGAATATCATTCCGGATAGTTACAAGCATGACCTGACCTTTGGC
GTGATGGATGACTACGACGGACTGATCTATGAATACACGGATCCGACCGA
CGATAGCCGTATCAATATCTATTTACCGGATAAAGGAGCCAAAACCCAAA
AGAGGTGAAATCTGTTGGTGTACGAAACAAGTGGCAAGCGCATTTC AATG
CATACCGGATTTGGAACAAGATGCGCTTCCAGCGCAAATCCATTACCTTT
GATGCGGCACCAGAATCAGAATTACTGGTTTTACGTGACCGGATTGCTGT
AGCGGATTATCGCAATGGTATTCATCAAAGCGGTGAGGTGGTACAGCAAG
AAGGTTTAGTCCTCACCTTAAGCCATGATGTAGATTTTCATTGCAGGCAAGA
GCTATGTGATTTATCTGCAAATGGCGGATGGCACAGTGGACCTGATTCT
GTTACACCGGGTTCAGCCAAAATAAGGTGGTTTTAGGGCGGTTGCCGAA
CGGGGCCTTAAAGCTAAGTCCTGATGATTTTGTAAATACCATCTATACAGT
AGTTAATGATGATACTAAAGGTTCAATTGCCTTATCTGGTTGCAAAAAGAGA
ACCGGTTGACCAGTTCTCTAATACTATTACAGCAATTAATTACGATGAGCG
CTATTACCTCAACGATAAAGACTTTATTGACGTGCCGGTTGATGATTCTCC
AATTTACATTCGATATGACCAGCTGGATATTAATCTGGCGGTTTATATCA
GATGCAAAGAGGGGATTTGCCAACGACTGGCGAAATCAGTTTTGTAGTTG
AATCTGGTGCAGTATTTCTAGTTTCGAGTTCTTATCGACCGGAAACCAGAT
TTGTCTATAAATTCGACTACA ACTCCAGTCCGCCGAAACAGGAATTTATTG
CCCCTGCAGCGACTGAACTACCCGCCATTGATACTGGTGAGTTCCACCT
GATCTGGTGGTAAATCTGACTATTAAGGTGCTGTTGTTGGGCGTGGTGG
TGATGGCGGTTTACCTCATTTGGCATTGCGCATGGGAGTCCGATCCGG
ATTACA ACTTTACTAAAACCCGACGTGATGGGTTTCAGGGAGCACCCGGT
TTATTGAACCGGCACAGCAA ACTAAACCTGATTATTGATGGCGGCACTCTA
GCTCGAGGTGGATCTGGTGGTGGAGCAACACCAAGCGGTATTTACTG

group_
74

hypothetical
protein

L6, L7,
L8, L9

H

DNA polymerase III is a complex,
multichain enzyme responsible for most
of the replicative synthesis in bacteria.
This DNA polymerase also exhibits 3' to
5' exonuclease activity

group_
8

hypothetical
protein

A2, A3,
A14

S

cellulase activity

GGTTGTCTTATGGTGTTC AAGGTATTCCGGGTGGAGCTGGTGCACCATTT
GGACGGGTCATGACAGGCCAGCCAATTTCAAGCGACTCACAAGATTGGC
GCTGGTATTTTGAAGTTACTTCAATGTCTTAAAAATTACTGATGCCGAAG
CTTCGGTACCCGGAAAAGGCTACCGAACCCAAAATGACCGTTATGGATCC
CCATTATCAGGTGATGGCGGAAACTGGGGCGAACGTGGTACCAAGTCTA
CTAATGATGGAACATGGAAGTGGAAATACCATGGCACAACCTGAAGGTGAG
CCGGGGCCAGGTGGACCTGCAATTGTGGGAGTTGCACCACTGACAACCTC
AATTGATTAATGGAGGGAAAATCTTACAAACCCTTTAA
ATGGTATTTGACCCAAATCATCATGAGTTGATTGGACTTGAGTCAGCAGTT
GTTCAAGAGACTGTTAGTGTCTTGAGGAAGATTTTATTCCGGTACCAGAA
CAAAAGTTAGTACAAGTTCAGGCTGAAACCCAAAGTCAAACAATCGAACCT
GAACCAGCTTCTACTGCTGAGCCTATAGGTTTATTTGAAGCATCGAGTGCT
GAATTTTCACTTGCTCAAGATACGTCAGCTTATGACTTGGTTTCTGAACCT
GTTATTGAACAACAGTCATTGGTGCAAGTTGAAATTGTTGAAACAGTTGTG
GTGGTAAAAGAGCCTAACGCAACAGATAACAGCCAGCTGATGCCACAAGA
TATTTTAAAGTTGCCATCACAACCTTTGGAAGGTGAGTGGACACTTGAAAA
GTGGGAGTTTTGGTTTAGAAATAGCCAGCTTTCTCCAGCCGTTCAAGAACT
TGCTCAACATGGTGTGATGACTGGGGAAATTGGTGGCAACACGGTTTTTCC
ATATTCCTCAAGAATATGAAAATATGTTGACTCAGCTACAACAAGCCTTAAT
TGAAGCTTTAAAGCAGCAGTGGCCAAATACACAATTTACGGTTCGAGTATG
GTGCAGTCAATACTGCTACACCGTATGTTATGCAAAGCGTTCGTAAAGAAA
AAGCTTTCCATCGGGCTACAGAACTATTACAGCAACAACCCGTTATTA
GCCTCATAGAGACTTTTTGATGGTGGAGCTGCAAAAATTTCAATTA
AAACCTTAA
ATGATTAAAGTTATTTATAAAAAAGACGCTTTGTCTGAAGAAAAGACAATTG
AGCAGGCTCAAACCATTTGGGCAATGGCTCACTTCAAAAATATGAACATATG
CCTGAGCATGTCCGTATCTTTCATACTACAAGCAATATGGATCATGCCGAA
ATTTCAATTTGCGAATGAAGTCACACCAAGAATGCATATGACTTAAAGCAG
CTTGATTTCTTACCGGGCACTTTTATCGTAGTTGAGAACCCTAAATGGGTC
GCGGCTATTGTTTCGATTGTGATTAGTATTGCGATCGCATTTTTTAATGCCA
ACGCCATCAATAGCACAAACGACTCAAATACTAACCAGTCTTCTTCAGCA
AACAATGAACCTTTCTAACCGGGAAAACAAGATCCGGGTGAATGGTTCGTAT
TGCTGATAACTATGGAGCTGGGTGGAATACTCCCGACCTAATCGCAGTAC

CTTACAAGGTATATGAAAACAACGTTGAAGTTGAGCATGTAGTGGGCTGTA
TTGGGCGTGGACACTATAAAATCAATGGAGCTTATGACGGTCAAACCAAT
ATTGTCGATATTGCTGGCGCATCGGTAGAAGTCTTTCGACCAGGTGTAGA
TATTGTTTCAGGTGAGCCATATTTCTCGCTTGGTACCGAAATTACCACGCC
GCCACTAACGGTTCAGCATCAAACCTTCTGTTAATGGCCAAGTTTTACGTCC
TGCTGATACACAGTCTTTAGAAGGTACGAACTACCTTCATTTTGCATATCC
AAACGAGATCCTTCGGGCATCTGCAAACAATACGGATTTAACCACTAAGTT
TGTAAGTAATGACCGCGTAGAAATCACCAATGCCTCATTACGTTTAAACGG
CCAGACTTATGATTTAAACGGCACTTACAGCGTTCTATCGGTAGCAGATGA
TCGAATGACGTTATCAAATCCGGCGGCCGTTAATGCTAACTGGTTAAAGC
TTAAAGAGTTAATAACCAACAAACTGCAGCTTTGTCACCAAAGATCAGTT
CAATAGGTGAAAAATGGATTGGTCCATTCTGGACAATGTTGAACGTA
GCCGGGTGCTGTGTAATTTTGTGGCCACAAATGGACTTTATACCGTTTCTT
CAGGTGGGAATCAGGCCGCTGTTAACGTCACGATTGAAGTTGAAGTAACA
CCGGTAAATGAATCTGGTGCAGCGATTGGTAATCCGATGCTGAAGCAGAT
CATTTTGAAGGTTCCGGCAAAGTCGCGTCAAACCGTTGGCGCAACGCTGG
ATATGGTGACATTTCAAGGTCGCTGTAGTGTCCGTGCACGTCGTTTAAACA
CCAACACCGGCGGTTACAACGGTAGTAGATGAAGTAAAGTGGCAGGCGC
TTTATGGTGCTTATCCTTTGCAAAGCACAGTGTATGAACATGAAACGGTTT
TTCGTGCGCGCACTTATGCAACCACTGGAGCTTTATCTGTTAAGTCCCGC
AAGATCAATTTTGTCTTCAGCGGATGTTACCGACTTTTAAAAACGGCGCA
ATGACGACAGAGCTATTTCCAACATCAAGCTTTGCTGATGCATTGGTTTCA
ATGGCACTGGATGACAAGATAGGCCGCCGTACGATCGACGAAATAGATCT
GGAAAATATCTATCGGACTTATAACGATGTAGTTGATTATTTTGGTACACC
ACTTGCGGCTGAGTTCTGTACTACGATTGATGATACAAACCTGTCTTTTGA
AGAGCTGGTCACCAATCTTTGTGATGCCGTGTTTTGTACTGCATATCGTCA
AAATAATAAGCTCAAGCTTTATTTTGAACGTCCAACCTGATAACTCGGTAAT
GCTATTTAACTTCAGGAATATTATTCCTGATAGTTACAAGCATGATCTTACC
TTTGGCGTGATGGATGACTACGATGGACTGATCTATGAATACACGGATCC
GGCCGACGATAGTCGTATCAATATCTATCTACCGGATAAAGGGCCAAGA
ACCCCAAAGAGGTGAAATCTGTAGGTGTGCGTAACAAGTGGCAAGCTCAT
TTAATGCGTACCGGCTTTGGAACAAGCTTCGTTCCAGCGCAAATCCATT
ACCTTTGATGCGGCACCTGAGTCAGAATTACTGGTTTTACGTGACCGGAT

CGCTGTAGCTGATTATCGCAATGGTATTCATCAAAGCGGTGAGGTGGTAC
AGCAAGAAGGTTTAATTCTCACCCCTAAGCCATGATGTGCGATTTTCATTGCAG
GCAAGAGTTATGTGATTTATTTGCAAATGGGGGATGGTACCGTGGACCTG
ATTCCCGTTACGCCGGGTTTCAGCCAAGAACAAGTAGTTTTAGGGCGTTT
ACCGAACGGGGCCTTAAAGCTTAGTCCCGATGACTTTGTGAATACTATCTA
CACCGTAGTTAATGACGATACCAAAGGCTCACTGCCTTATCTGGTTGCAA
AAGAGAACCGGCTGACCAGTTCTCTAATACCATTACTGCAATTAATTACGA
TGAACGTTATTACCTCAATGACAAGGACTTTATTGATGTGCCGGTTGATGA
TTCACCGATTTACATTCGATATGACCAGCTGGATATTAATCTGGCACGTTT
ATATCAGATGCAAAGAGGGGATTTGCCAACGACTGGAGAAATCAGTTTTG
TAGTTGAAGCAGGTGCACTAGTTTTCAAGTTCAAGTTCTTATCGACCGGAAA
CCAGATTTGTCTATAAATTCGACTATAACTCTAGTCCTGCAAACGAGAGT
ATATCGTTCAGCTGCATCAGAATTACCTGCTATTGATACTGGTGAGTTCC
CACCTGATCTCGTGGTAAATTTGACTATTAAGGTGCTGTTGTTGGACGTG
GTGGAGATGGCGGGTTGCCACATTTGGCATTGGTGATGGTCTACCGAT
CCGGATTATAACTTTACTAAAACCCGCCGTGACGGTTTTTCAGGGAGCACC
CGGTTTATTAACCGGCACAGTAACTAAACCTGATTATTGATGGTGGAAC
TCTGGCTCGAGGCGGCTCAGGTGGTGGAGCAACACCAAGCGGTATTTAT
ACAGGATTATCGTATGGAGTTCAGGGTATTCCCGGTGGAGCTGGAGCACC
TTTTGGTCGGGTATGACCGGACAACCTATTACTAACGATTCACAAGACTG
GCGTTGGTACTTAAATGGTGACTTTATGGTTGTCAAAGTAACCGATGCCGA
AGCTTCGGTACCCGGTAAAGTTACCGAACCCAAAATGATCGATATGGAT
CTCCATTGTCTGGTATGGTGGAGGTTGGGGCCAGCGCGGTACCAAGTC
CACCAATGATGGAACATGGAAGTGGCAATACCATGGCACAACCTGAAGGCC
AGCCGGGGCCGGGTGGACCTGCAATTGTTGGGGTGGCACCACTTACAAC
TCAATTGATCAATGGAGGGAAAATTCTACAAACACTTTAA
ATGACTGGAAATAATACAAGCAACAAGTTTAAAGCTGCTAGAGAGAACAAG
GATTTTGTGCAAGACGTTAATTCAGATGCAGCAGATCCAATGCCTAAACAA
ATTACTGTTACTAGCAGAAATAAAAAGTCAAATAACCCTTGAGCAAAATACC
CTAAGTCAACGAGAGAGATCCAGAGTTAAAAATGGTAGGAACCTAACGATT
CCACTATTTGCAGAAGAACTTTTGATTATAGAACTGCTGTTGAAAGCCTT
TCTCAAAAAGAAGATGTTTCTATCAATAGCTTCATTAGAAGTATCGTTCCTG

group_
804

hypoth
etical
protein

A2, A3,
A14

/

hypothetical protein

group_
806

hypothetical
protein

A2, A3,
A14

D

AAA domain

ATAAGTGTAAAGAAGGTGATAGGAAAAGATGCCTTCGATAAATTAATGAAA
ACAAGCTAAATGTTGTTAAAGCAAAAAAGAAAAGTAG
ATGGCTGAAATAATTACTATTGCTAGTAGAAAGGGTGGGGTTGGAAAAAC
CACTATTGCAATAAATATGGTTGTTGCTTTAAGCAATAAAGGTAAAACTTTA
TTGGTAGATGCTGATGAACAACAATCTGCTATGAAATGGAATAAGCATCGG
AAAGAAAAGATAGATGCTATATCCATACATAAAGGATCTAATTGCTGAGTTG
GAAAACTAGATGGGGAGTATCAGTACATTCTCATAGATGTTGCTGGCCG
TGATTCTGAAATATTTAGAGAAGCACTTCTTATCAGTGACAAGCTAATCGT
ACCCACTCAAGCAAGTCTTCTTGATTTAGATGTAATCCCTTATCTAAATGAA
AAGGTTAAAACAGCTAAAAAACCAATCCAACTTGGTTTCAATGGTTGTG
ATTAATAGAGCATCAACCAATGCCAATCCAATGAAACAGAGCAGGCAAG
AGAATTTGTTGCTGATTATGAGGAATTCAAGTTATTGCAAACCACATTGCA
TGAACGAAAACAGTATAGAGATGCAGTTCTTGAATCAAAGAGTGTTCTTGA
GATGGCGGCATCTAAAGCTAAAGACGAATTAATCAATTTCTAATTGAGGT
GCTGTGA

group_
808

hypothetical
protein

A2, A3,
A14

P

T5orf172

ATGTCAGATATTAATCACTTGAAGATATTTTTAATGATGATGATCTAGGTT
TGCTCGATGATGGTAAAGATGGTGACATCTTTAAAATTAAGCATGTGCCTG
TCCAACGTGCAGCAGCAGATGAAGTTGCTAAACGTAAACGCTGTAAGGAC
TTTGATCAATTTGAGCATCTATTTAAAGATGTTCCACAGAGAGCTTCGAGAA
GGTAAGCGTAAGCTATTACCATTTTATGATAAAAAGTTATCAGCTCGAAAAA
GGCAGTTACTATGTTTTAAATGGTCTTTTGGTTATTTGGCAGATTTTACAT
TACTAGAGAAGATAAGACCGTTAATAAAGTCAGAACCAGAAAAGATGGTA
GAACTAGAACTATTTTTGAAAATGGTACTGAATCCACAATGCTCTATCGAT
CACTAGCTAAAGCATTATATAACGATGGGCGAATTGTCTCTGAACTAATG
AGCAAGCAAATTTCAATTTCTATACTAATTTTGGTGGGATTACAGAAAATG
ATACATCTACTGGTTACATCTACATCTTGAATCTCATAGTAAAGAACCAAA
AATTCAGTCACTTGAACATCTCTATAAAAATTGGATTTTCAACCACAACCTGTT
GAAAAACGTATTGCTAATGCAAAAAAGAACCGACTTATTTGATGGCTGAC
GTTGAAGTAATTGCTGAATATCAAGCATTCAATTAATCCACAGAAGTTTG
AATATTACTTACATGCTTTTCTTTGGTGAGGCTTGCCTTGATTTATTGGTCGC
AGACAAAGGTGGTAAAAATCATCAACCTAAAGAATGGTTTATTGCGCCTTT
AAATGTGATTCAACAAGCCGTTGAATTATTAGTAAATAGCCAAATTATGAAT
TACAAATACAATAAGCTCACAAAAGAGATAGAATTAATAATAG

<i>group_810</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>GTGGAAGATTTTAGCTGGCTTCAAGATATGATTGCTGATGATGACTTAGGT CTTTAAACGTAAGAAAAAGATGTCTGCTCTAACGGCTGATGAGCAATTA GTCTCTAAGTTTAATGAAATTAATAATTTTGTGACACAGCATGGGCGTGTG CCTACGAAGAATATGGCAAATATCACTGAATATATGCTCGCTAGCCGTCTA GAAGCAGTTCAAGCCAATCCAGAGCAATACATGGCTTTAGCAGATTATGAT GAACATAATTTGCTGCCACAACATCAGGATTAA ATGTTCTGGGAACGTCTGGGTGGTATGGATGAGTTTAAAGAACTACTCCT ACCAATCCAGACTCCGGCCACTTTGGTTGGACAAGATTTGACGTTAGGAA GTTTTCCGGTTCGCCAGCAAGAAGGACAGTTGGATATTACGCTGGAAATG GGCGGTGAATATCAAGGTGAACTCGTTGGTGTATTGAAATACAACACTGA TCTATTTTCTGCGCAGTCTGCCAAGAATATGGTGCAACTGCTGCAAGCTGT GCTATCAGAGATGGTGGCTCACCTGAACGTAAGATCGTTGAGTTAGATA TCGCACCTGATTATAAAGATAGTATCCAGTTCGAGGCCCTCCGTGAAAG GCGACTGATTATGCTCAGCATGATCTATTCGCAATGATTCTCAAGCAGATT GATGAGCGTGGTGACAATCATGCTTTGACCTCAAACGATCATGCGGTTAG CTACTGTGAGTTAGGTCAACATATTGCTGGTATTGCTGAATATTTGTGTGC ACACGGCATTACACGCGGTGATCGTGTAGGTCTGATGTTAGACCGCACG GCGTTATTGCCAGCAGCAATCCTTGGGATCTGGGCGGCAGGTGCAGCTT ATGTGCCACTGGATCCTAATTTCCCAACTGAACGTCTACAGAACATTATTG AAGACGCTGAACCTAAAGTCATCTTGACTCAAACCGAGTTAATGGATGGT CTTAACGTTTCTGTGCCTCGCTTAGATATCAATCAAGCTGGTGTGTGGCA CTTGAACAAGTGCGCGAGACGCTTGCCTTTGGTGACATAGCTTATGTGAT GTACACATCAGGCTCGACAGGTAAGCCTAAAGGTGTCCGTATTGGGCATC CTTCAATTATCAATTTCTTGCTATCTATGAATGATAGATTGCAAGTAACGAC TGAAACTCAGTTGTTAGCAATCACCACGTATGCATTTGATATTTCAATTCTT GAATTACTATTCCGCTGATGTACGGCGGTGTGGTGCATGTTTGGCCACG TGAGGTATCACAAGATGGCAATCAATTGGTGGATTATCTAAATGCTAAGTC TATCAATATTCTGCAGGCAACACCGGCTTCTTGGAAAATGCTCTTAGACAG CGAGTGGAGCGGCAATGCGGGATTGACAGCACTCTGTGGAGGCGAAGC GTTAGATACTATTCTGGCAGAAAACTGCTGGGTAAAGTCGTTGCTTATG GAATGTCTATGGTCCGACCGAACTACAGTGTGGTCTAGTGCGGCTCGTA TCACTGATGCAAAGTATATTGATTTAGGTGAACCACTGGCGAACACTCAGC TTTATGTCCTTGATGAACAGCAGCGTTTAGTTCCACCAGGTGTTATGGGTG</p>
<i>group_82</i>	<i>dhbF</i> Dimodular nonribosomal peptide synthase	L6, L7, L8, L10	Q	Pfam:HxxPF_rpt	

AGCTTTGGATTGGTGGTGATGGTCTGGCAGTTGATTACTGGCATCGACCA
GAACTGACAGATGCCCAATTTAGAACATTGCCATCTCTACCAAATGCCGGT
CGACTCTACCGTACAGGGGATAAAGTCTGCCTGCGTACAGATGGACGCTT
GACTCATCACGGACGTCTGGATTTTCAGGTCAAGATTCGTGGGTTCCGTA
TTGAACTAGGTGAAATTGAGAATGTTCTCAAACAAATCGACGGTATTACAG
ACGCAGTAGTGCTAGTTAAACTACAGCTGACAACGACCAGAAACTGGTG
GCGTATGTAAGTGGACAAGAGCTTGATATTGCTGGTTTTGAAAAGACCCTA
CAAATCCATCTTCCAGCTTATATGGTGCCAAGTGCCTTTATCCGATTGGAT
GAGTTTCCAATGACGGCCAACAATAAGTTGGATCGTAAGGCATTTCTGA
GCCGATCTTTGAGCAGAGCAATGATTATGTTGCACCGCGTGATCCGATTG
AAATCGAACTGTGTACTACCTTTGAACAAATCTTGTGCGTCAAGCGAGTTG
GTATACATGATGATTTCTTTGAATTAGGTGGGCATTCACTATTAGCTGTGA
AGCTTGTCAACCATCTTAAAAAGGCTTTTGGTACCGAACTGTCAGTTGCTT
TGTTGGCACAGTACTCCACTGTGGAACGTCTGGGTGAGATCATTGCGTGAA
AATAAAGAAATTAACCATCCATCGTAATTGAGCTGCGCCGTGGAACCTAT
GAACAACCACTGTGGCTGTTCCATCCAATTGGTGGCAGTACCTTCTGTTAT
ATGGAACCTTTCGCGTCATCTTAATCCGAACCGTACCCTACGTGCTATTCAG
TCTCCTGGACTGATTGAGGCCGATGCTGCGGAAGTGGCGATTGAAGAAAT
GGCGACGCTCTATATTGCTGAAATGCAAAGATGCAGCCTCAAGGACCTT
ATTTCTCGGTGGCTGGTGTTCGGTGGTGGCAGTGCATACGAGATCAGT
CATCAGTTGCGTCAAATGGGACAGCAGGTCACAGGGATTGTCATGATCGA
TACCCGTGCTCCAATTCCAGAGAATGTGCCTGAAGATGCCGACGATGCCA
TGCTTTTGTCTTGGTTCGCCCCTGATTTGGCTGCACCTTATGGCAAAAAGT
TGACGATCCCAGCGCAATATTTGCGTGAGTTGAGTCCAGATCAAATGTTT
GATCATGTACTTAAAGAAGCTAAAGCCATTAATGTTCTGCCTCTCGATGCA
GACCCATCAGATTTCCGACTTTACTTTGATACTTATTTGGCGAATGGCATC
GCACTCCAGACCTATTTCCCTGAACCTGAAGATTTTCCAATCCTTTTGGTC
AAGGCCAAAGATGAGCAGGAAGACTTTGGAGAAAGTCTAGGTTGGGATCA
ACTAGTCAAAGACACGTTGACACAAGTCGATTTGCCTGGTGACCACTCAT
CAATTATGTACGCTGAGAATGTCGTGGCTGTAGCCCAGACAATTGATCAG
ATGTACCCAATTCCTACATAA

<i>group_840</i>	hypothetical protein	L6, L10	/	hypothetical protein	<p>ATGGTAAATACAAACTACGTTCCCTTTGTGGCTTATCTCACCATTTTCAGCAT GTGCATTACACATTAGTTTCGAAATCAACTGCATATGGATTTGCTATTTGAG GACATGAATAAGGTTCGATCAATTCTTGTCTATTGAAGGGGCTGCAGCTCA GGTTGATTTCTATTCCGAAGGTGCATATGCAGTTGTTTCAGCTTGGTGATAC TTCAGAAAGAAATCAGATTGAAGTGTATGGATTGCTTTTACATGAAGCTGT TCATGTCTGGCAAAGATTA AAAAGCTCATGGGTGAACGAGAACCGAGCT CTGAGTTTGAAGCTTATTCAATTCAGGCGATCGCTCAGGATCTCTTTAAGA TGTATGAGGAAAGCGAGGTTAAAAGTCATGGGGTGAAGGGGAAAAAGC CGACTAG</p> <p>GTGACGCTACAGGCACCCGCACAAACCCAGAAGAACCCTAAACAAAAACAA TTGGGGTTCTGGTCGAGGTGGCCGTCCGTGGCGCCGTCTTAAAGCTAAG ATCCATTTACGTGATGAGTGGACCTGTCAATGTTGTGGCATCGTCACTAAA GACTTAGAGCTTGACCATATTGTGAATGTGGCAAGAGGTGGAACGGATGA TGAATCAAACCTCCAGTCTCTTTGTGTTCCATGCCATAAAAAGAAAACCCA ACAGGAGAGCCGGCAATGA</p> <p>ATGTTGATTATCGATGAAATTCATAATTGTCTGACAGGTTTCAGCCAAATTG CTTCCCGAAGTCATGAATACGCTCAAGAACCTCAGCAATGAATTGAGTCT GAATATCGTAGGTGTTGGTACTCGTGAGGCAATCACCATACTGCATACTG ATCCGCAATATGCCAGTCGCTTTGATGTGGTGAATTTACCTAAATGGGAAC TCAATCAAGATTTTTTAAGATTATTAGTCAGCTATGTTTCGACTATTGCCGTT AAAAAACAGTCTAATCTTGCATCCAAAGAGATTGCTACGTTGATTTTTGAA GTGTCAGGTGGAATTTAGGGGATTTAAATCGATTACTGGTGGAGTGTGC CAAAGAAGCCATTTTACAAGGTGATGAAGAAATTACCTATGACATCGTACA TAAGTTTAAATGGCTTAAACCTACGGAGGGTTTAAAGAAATATTAGAAATATA AATCTCAGCCTGAGTTAA</p> <p>ATGTTGATTATCGATGAAATTCATAATTGTCTGACAGGTTTCAGCCAAATTG CTTCCCGAAGTCATGAATACGCTCAAGAACCTCAGCAATGAATTGAGTCT GAATATCGTAGGTGTTGGTACTCGTGAGGCAATCACCATACTGCATACTG ATCCGCAATATGCCAGTCGCTTTGATGTGGTGAATTTACCTAAATGGGAAC TCAATCAAGATTTTTTAAGATTATTAGTCAGCTATGTTTCGACTATTGCCGTT AAAAAACAGTCTAATCTTGCATCCAAAGAGATTGCTACGTTGATTTTTGAA GTGTCAGGTGGAATTTAGGGGATTTAAATCGATTACTGGTGGAGTGTGC CAAAGAAGCCATTTTACAAGGTGATGAAGAAATTACCTATGACATCGTACA</p>
<i>group_863</i>	hypothetical protein	A2, A3, A14	L	HNH endonuclease	
<i>group_87</i>	hypothetical protein	C4, C6	S	Bacterial TniB protein	
<i>group_88</i>	hypothetical protein	C2	S	Bacterial TniB protein	

<i>group_928</i>	hypothetical protein	A2, A3, A14	S	Protein of unknown function (DUF1376)	<p>TAAGTTTAAATGGCTTAAACCTACGGAGGGTTTAAGAAATATTAGAAATATA AATCTCAGCCTGAGTTAA</p> <p>ATGAACTATTACCAACACCATATTGGTGACTTCAACAATGCGACTCGCCAC CTCAGTTTAATTGAGCGTGCGATTTACCGCGACTTATTAGATATGTATTAC GACACAGAAAAGGCGATTGATGCATCAAGCATTGATCGTCTAGCACGTGCG TTTGCAATGTACTACCGAAGAGCAAAAAGAAGCTCTCAAATATGTACTTGA TGAGTTTTTCATTCTTGAAGAAGGTGTTTATCGCAATAATCGTTGTGAACG AGAAATTGCTGAATATCACGGGAAAAAGAAACAAGCGAGTGAGGCTGGTA AGGCGTCTGCTGCAAAACGTGCAGCGAAAAAGAAAGGCTCGTCCAACAG TGATTCATCAAAGATGATCAAGCGTCTAACGAAAATTCAACGGTCGTTGA AAATCCGTTAAACGAAGAACAACGGATGTGCAACCAACCAATAACCATAA ACCATTAACCATAAACCAAGAACCAATTATTGATAGTAGTAATACGCG TGGAGAAAATTCGCAATTAACCTCAATTTGCTCAGTATCAGATCGA TGATCACAAACGCTATTCAATGCGTGAATTCATTTCTGAATACAGCGAGTT TCAATACGATTTCAATTCACCTTGCTCAACAAAGATTTGTTTCGGTACCTGAA ATCGACTTGAGAACCATGATTCAAATTTTCGGTGACTGGTACTTTGCAAAC GAATCAAGTTCGTTGAATACACCAAGCATCTGGTTGGTTAAGTGGTTCTCT TGGGTTCAAACAACGAGAAACAAGTCGCTGCTAACCGCAAGAAACAAGA GCAAATCAATTCAGCTGGTCAAAAACCACAAGAGTCCGGTTACTTCGCTA ATCTTTTTGAAGAACAGAGCGAATCTCAAATCGTGGATGTAACCCAGCAA AAAAGTTTCCAATGATTGAGGAGGTAGGTCATGCATGA</p> <p>ATGAGTGAGTTTGAGGGTAAATCTGGAAAGTGGGCTTGGGAGATTCAAAA AGAACAACAAGCGAAAGTGGAGGAGCTGCAAAGCGTTTAGATGGGGCA TAAAAGAGACTCAATATGCTTTGCAGTATGTTGAAGAAGACATGCGCGG CAATCATGAATTTCTACAAATGGCAATGATTTCGAACCCTTAAAGCTATAGA GCAAGTGCTCAAAGGTGGTGCTTGA</p> <p>ATGGTTAATTCTGATTATGTTCCCTGAATGGTATATCTTGCCATTCCAACATG TGCAGTACACGCTCGCTCGAAATCAACTACACATGGATTTGTTATTTGAAG ATATGGATAAGGCTGATCAATTTTTGGATATGGGAGCGGATGCACAGGTT AGTACTTTTTCTGATGGTGCATATGCAATCGTCCAAATTGGTGATACGGCG GATAAAGACCGAATTCAAGTTTATGGATTGCTTTTACATGAAGCTGTTTCAT GTCTGGCAAAGATTA AAAAGCTCATGGGTGAACGAGAACCGAGCTCTGA</p>
<i>group_929</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	<p>ATGAGTGAGTTTGAGGGTAAATCTGGAAAGTGGGCTTGGGAGATTCAAAA AGAACAACAAGCGAAAGTGGAGGAGCTGCAAAGCGTTTAGATGGGGCA TAAAAGAGACTCAATATGCTTTGCAGTATGTTGAAGAAGACATGCGCGG CAATCATGAATTTCTACAAATGGCAATGATTTCGAACCCTTAAAGCTATAGA GCAAGTGCTCAAAGGTGGTGCTTGA</p> <p>ATGGTTAATTCTGATTATGTTCCCTGAATGGTATATCTTGCCATTCCAACATG TGCAGTACACGCTCGCTCGAAATCAACTACACATGGATTTGTTATTTGAAG ATATGGATAAGGCTGATCAATTTTTGGATATGGGAGCGGATGCACAGGTT AGTACTTTTTCTGATGGTGCATATGCAATCGTCCAAATTGGTGATACGGCG GATAAAGACCGAATTCAAGTTTATGGATTGCTTTTACATGAAGCTGTTTCAT GTCTGGCAAAGATTA AAAAGCTCATGGGTGAACGAGAACCGAGCTCTGA</p>
<i>group_930</i>	hypothetical protein	L6, L7, L8, L9	/	hypothetical protein	<p>ATGAGTGAGTTTGAGGGTAAATCTGGAAAGTGGGCTTGGGAGATTCAAAA AGAACAACAAGCGAAAGTGGAGGAGCTGCAAAGCGTTTAGATGGGGCA TAAAAGAGACTCAATATGCTTTGCAGTATGTTGAAGAAGACATGCGCGG CAATCATGAATTTCTACAAATGGCAATGATTTCGAACCCTTAAAGCTATAGA GCAAGTGCTCAAAGGTGGTGCTTGA</p> <p>ATGGTTAATTCTGATTATGTTCCCTGAATGGTATATCTTGCCATTCCAACATG TGCAGTACACGCTCGCTCGAAATCAACTACACATGGATTTGTTATTTGAAG ATATGGATAAGGCTGATCAATTTTTGGATATGGGAGCGGATGCACAGGTT AGTACTTTTTCTGATGGTGCATATGCAATCGTCCAAATTGGTGATACGGCG GATAAAGACCGAATTCAAGTTTATGGATTGCTTTTACATGAAGCTGTTTCAT GTCTGGCAAAGATTA AAAAGCTCATGGGTGAACGAGAACCGAGCTCTGA</p>

<i>group_931</i>	hypothetical protein	A2, A3, A14	/	hypothetical protein	GTTTGAAGCTTATTCAATTCAGGCGATCGCTCAAGACCTTTTTAAAATGTAT GAAGAAAGCGAGGTGAATGATGGGATGGAAGGGGAAAAAGCCAACTGA ATGGTTAATTCTGATTATGTTCCCTGAATGGTATATCTTGCCATTCCAACATG TGCAGTACACGCTCGCTCGAAATCAACTACACATGGATTTGTTATTTGAAG ATATGGATAAGGCTGATCAATTTTTGGATATGGGAGCGGATGCACAGGTT AGTACTTTTTCTGATGGTGCATATGCAATCGTCCAAATTGGTGATACGGCG GATAAAGACCGAATTC AAGTTTATGGATTGCTTTTACATGAAGCTGTTTCAT GTCTGGCAAAAAGATTA AAAAGCTCATGGGTGAACGAGAACCGAGCTCTGA GTTTGAAGCTTATTCAATTCAGGCGATCGCTCAAGACCTTTTTAAAATGTAT GAAGAAAGCGAGGTGAATGATGGGATGGAAGGGGAAAAAGCCAACTGA ATGTTTCGTGATAAAGGGTATCAAGAACATGAAGAATCTAACAAAGAATTT AATGCGTCACAAAAAGTTTCCGATATACGGCAACCTGTTATCTCAACTTTTT GGTCAAAAAGAGAGGTTATCTAG ATGGACGATTGGCAATGTAAGTATTGTAATGGTTACATTATGGTTAATCATT CAAGGATAGAGGTAGGGGAAAAAGTTTACTTTTTGGTATATAAATTCGATG CAAAAAATGAAAGAAAAAACTATATAAAAAAGGAAGTGAATTGCACGTT GCGATAGTATTTTACATATCGAATCAAGAAAAAGACATATAAATAGAGG AGGCTAAAGTCTATCCTTTAGGTGCACCTATGCCATTTGTTTATAATATGTT CTGGATATGTGGTTGTGAAAGCAGGCCATAA ATGTCTAAAGGTCAAACCTTTACAAGATCCGTTCTTAAATTCTCTCCGTAAAG AACGCATCCCAGTTTCTATTTTCCTTGTTAACGGTATTAATTACAAGGTCA TATTGAATCTTTTGACCAATATGTTGTTTTATTA AAAAATACTGTAAGTCAA TGGTTTACAAACACGCAATTTCTACAGTTGTTCCAGCTCGCAACCCACGTC CAGCAGGTGCACAAGGTGCAGTTTCCAGCTCAGGGTGGTAGTCAAGG TGGCTTCGGTGGTCAAGGCGCTGGCTTTGGTGGTGGTCAAGGTGCTGGC TTCGGTGGTCAAGGTGGCTTTGGTGGTCAAGGCGGCTTCGGTGGTCAAG GTGGCTTCGGTGGTCAAGGTGGCTTCGGTGGTCAAGGTGGCTTCGGTGG TCAAGGCGGCTTTGGTGGTCAAGGCGGCTTCGGCGGACATCAAGGCGGT TTTGATAACGATTCTAAATTTGAAGATGGTCAAGACGACGAAAACAATCGT TAA ATGTGGACAGTCATTACAACGGATCTGTTTAATGAATGGCTTGAACAGCAA GATGAAGCAACACAAGAGAAAGTCTTAGCTGCATTGGTTGTTCTACAACA GCAAGGGCCTAGTTTAGGTTCGTCCTTTAGTAGATACCGTCTACGATTCTAA
<i>group_954</i>	hypothetical protein	L8	/	hypothetical protein	
<i>group_968</i>	hypothetical protein	L7	/	hypothetical protein	
<i>hfq</i>	RNA-binding protein Hfq	A2, A3, A14	J	RNA chaperone that binds small regulatory RNA (sRNAs) and mRNAs to facilitate mRNA translational regulation in response to envelope stress, environmental stress and changes in metabolite concentrations. Also binds with high specificity to tRNAs	
<i>higB2</i>	Putative toxin HigB2	A1	S	Phage derived protein Gp49-like (DUF891)	

<i>ispE</i>	4- diphos phocytidyl-2- C- methyl -D- erythritol kinase	C2, C6	F	Catalyzes the phosphorylation of the position 2 hydroxy group of 4-diphosphocytidyl-2C-methyl-D-erythritol	<p>ATTTACCAATATGAAGGAACTACGTGTTTCAGCATCGTGGTAAACCGTTAAG GGCATTTCATTCGATCCCCTAAGACAAGCGATTGTGCTGTGTATTGG TGATAAGAGTAATAAAAAGCGTTTTTATACAGAGATGCTAGCAATTGCAGA TGAGCAATACGCACTTCATCTCTCAACTTTAGGAGATCAATCTAATGGCTA A ATGATTTCGAGTACCTTCTCCAGCTAAGCTCAACCTGTTTTTGCATATCACA GGACGTCGTGAAAATGGTTACCATGAACTACAACTATTTTCCAACCTATT GATTTATATGACTGGATGACATTTACGCCTATATCTGAAGATGAGATTGAG ATTGAAGGTTTAGGTGAGGTTCAACTTGAGCAAAAATCTGATTTATCGGGCG GCACAGATATTACGCCCTCACGCTCAAATCCTTGCGGACTCCATATAAAA ATTGAAAAAATATTCCGATGGGTGCCGGTTTAGGTGGCGGTTCTTCCAAT GCAGCAACGACGCTTATTGTCTTAAATCAATTATGGCAATGTGGTCTAACT GAAGAACAATTAGCTCAGTTCGGTGTCAAGCTTGGTGCCGATGTTCCAAT TTTTATTTATGGGTTAAATGCATGGGCTGAAGGGATTGGTGAACATTTATC ATTCATAGACTTAGATCAAAAACAATTCATTGTTTTAAACCTGATTGTTTTA TCAGCACTCAACTGCTTTTTTCAAAAAACATTGACAAGAGATTCTAAGC CCACTACATTTTGCGCCTATCAGTTAGAACCTTCTAATTTTCGAAATAACTT TGAGCCTTTGGCTAGAGAGTTATACCCTGAAGTTGAAGAAGCAATGCAAT ATCTCGATCAGTTTGGTCATGCAAAGCTTACAGGTACAGGTGCTTGTGTTT TTGCTGAGGTAACCGATGAAATGAATGTTGATGACATTCTTAAACATGCAC CATGTAAAGCTTACTTGGTACATAGTTTAAAAGAATCTCCTTTACGTCATTT TAAAGTTGCAAGTTAG ATGGATAAACTGAACCCAACCTACAGTGATTAATATTCATATCTCATTGCTT TTGGACTATTCATTATTGCTGGTCTGGTGAGCTCAAATATTTTCTTGTTAGG TCTTACTATCTTCTAATTGGTGCTTTACTCGCTGGTGCACAGTCTTCATTA CTGCCACTTGCAGCTATGTTCTATCCAGCAGTATGCCGCGCGGTTGGAGT TTCATGGATGCATGGTATCGGTCGTATCGGTGCAATTTTAGGCGCTTTCTT TGGTTCGCTCATTTCACCTTCAACCTAAGCTTAAAGCGGATTTTCTTTATT CTAGCAATACCAACCTTCATCTCATTATTGCACTTAGCCTGAAAGTAATTT ATGAGAAATCTAAAAACAACAAGTGTTAAAGCTAGAAGAAAGTCTATAA ATGATTGACAAAAGTAAGTCCTCACTAAGCGAGGTTTTATCGCAGATTTAAA GATGGTGCAACCATTCTGATTGGTGGTTTTGGTACCGCAGGACAACCCGC TGAACCTATTGACGGTCTGATTGAACTTGGCGTTAAGGACTTGACCATTGT</p>
<i>mhbT</i>	3- hydrox ybenz oate transp orter MhbT	C2	EGP	Uncharacterised MFS-type transporter YbfB	<p>ATGGATAAACTGAACCCAACCTACAGTGATTAATATTCATATCTCATTGCTT TTGGACTATTCATTATTGCTGGTCTGGTGAGCTCAAATATTTTCTTGTTAGG TCTTACTATCTTCTAATTGGTGCTTTACTCGCTGGTGCACAGTCTTCATTA CTGCCACTTGCAGCTATGTTCTATCCAGCAGTATGCCGCGCGGTTGGAGT TTCATGGATGCATGGTATCGGTCGTATCGGTGCAATTTTAGGCGCTTTCTT TGGTTCGCTCATTTCACCTTCAACCTAAGCTTAAAGCGGATTTTCTTTATT CTAGCAATACCAACCTTCATCTCATTATTGCACTTAGCCTGAAAGTAATTT ATGAGAAATCTAAAAACAACAAGTGTTAAAGCTAGAAGAAAGTCTATAA ATGATTGACAAAAGTAAGTCCTCACTAAGCGAGGTTTTATCGCAGATTTAAA GATGGTGCAACCATTCTGATTGGTGGTTTTGGTACCGCAGGACAACCCGC TGAACCTATTGACGGTCTGATTGAACTTGGCGTTAAGGACTTGACCATTGT</p>
<i>pcaI_2</i>	3- oxoadi pate	L6, L7, L9, L10	I	3-oxoadipate CoA-transferase activity	<p>ATGGATAAACTGAACCCAACCTACAGTGATTAATATTCATATCTCATTGCTT TTGGACTATTCATTATTGCTGGTCTGGTGAGCTCAAATATTTTCTTGTTAGG TCTTACTATCTTCTAATTGGTGCTTTACTCGCTGGTGCACAGTCTTCATTA CTGCCACTTGCAGCTATGTTCTATCCAGCAGTATGCCGCGCGGTTGGAGT TTCATGGATGCATGGTATCGGTCGTATCGGTGCAATTTTAGGCGCTTTCTT TGGTTCGCTCATTTCACCTTCAACCTAAGCTTAAAGCGGATTTTCTTTATT CTAGCAATACCAACCTTCATCTCATTATTGCACTTAGCCTGAAAGTAATTT ATGAGAAATCTAAAAACAACAAGTGTTAAAGCTAGAAGAAAGTCTATAA ATGATTGACAAAAGTAAGTCCTCACTAAGCGAGGTTTTATCGCAGATTTAAA GATGGTGCAACCATTCTGATTGGTGGTTTTGGTACCGCAGGACAACCCGC TGAACCTATTGACGGTCTGATTGAACTTGGCGTTAAGGACTTGACCATTGT</p>

<i>pcaL_3</i>	CoA-transferrase subunit A	L7, L8	I	3-oxoadipate CoA-transferase activity	<p>GAGCAACAATGCCGGTAATGGGGATTATGGTCTGGCTAAACTGCTCAAAG CCGGTTCAATTAATAAAGTGATTTGCTCTTTCCCACGTCAGTCTGACTCTT ATGTGTTTGATGAACTGTACCGTGCCGGAAAGGTTGAACTTGAAGTGGTA CCACAAGGCAATCTGGCGTGCCGTATTCAGGCAGCAGGTATGGGCCTTG GGGCGGTGTTTACACCAACAGGCTTTGGAACACTTTTAGCTGAAGGCAAA GAAACCCGTCATATTGATGGTAAAGACTATGTGCTTGAGTATCCGATCAAG GCTGACTTTGCCCTGATTAAGCCTATAAGGGCGACCGCTGGGGCAATCT GGTTTACCGTAAATCTGCACGTAACCTTTGGCCCGATCATGGCCATGGCTG CTGATGTCACCATGCCCAGGTTTCTGAAGTGGTTGAACTGGGTGGATTA GACCCGGAACACATCATTACCCAGGTATTTTTGTACAGCACGTTGTACAA GTACAGCCCGCACAGTAA ATGATTGACAAAAGTAAGTCCTCACTAAGCGAGGTTTTATCGCAGATTA GATGGTGCAACCATTCTGATTGGTGGTTTTGGTACCGCAGGACAACCCGC TGAACCTATTGACGGTCTGATTGAACTTGGCGTTAAGGACTTGACCATTGT GAGCAACAATGCCGGTAATGGGGATTATGGTCTGGCTAAACTGCTCAAAG CCGGTTCANNNNNNNNNCTTTCCCACGTCAGTCTGACTCTTATGTGTTT GATGAACTGTACCGTGCCGGAAAGGTTGAACTTGAAGTGGTACCACAAGG CAATCTGGCGTGCCGTATTCAGGCAGCAGGTATGGGCCTTGGGGCGGTG TTTACACCAACAGGCTTTGGAACACTTTTAGCTGAAGGCAAAGAAACCCGT CATATTGATGGTAAAGACTATGTGCTTGAGTATCCGATCAAGGCTGACTTT GCCCTGATTAAGCCTATAAGGGCGACCGCTGGGGCAATCTGGTTTACCG TAAATCTGCACGTAACCTTTGGCCCGATCATGGCCATGGCTGCTGATGTCA CCATTGCCAGGTTTCTGAAGTGGTTGAACTGGGTGGATTAGACCCGGAA CACATCATTACCCAAGGTATTTTTGTACAGCACGTTGTACAAGTACAGCCC GCACAGTAA ATGAATACCTCTTTACGTTTAAACCACTATTGGATTACTTGTGCTGATGGTT TAGAAACCCTCTTACAAGAAGAAATAGAACAGCTTGGCACCAAAGTGACA GAACGTAAAGCAGGACGTTTAAATCATTGAAGGAACGCTTGAACATGCTTA CCGTATTTGCATGTGGTACGCTTTGCTTACGTTTACTTCCAATTCA TACTTATGAACTTGAACGTACCCACGATGCACGTGATGTAGCAGAAGAGC TTTATGAAGGTGCAATTAGTTTTGATTGGTCACTTATTTTTGCACCACAAAG TACCTTTGCAATTGCTTACACGCAGAGCGTGAATTAAGGTAATACGCA GTTTGCAACCCTTCGTGTAAAAGACGGTGTGTTGATTCCCTTTATGGAAGC</p>
<i>rimL_2</i>	3-oxoadipate CoA-transferrase subunit A	A2, A3, A14	J	Specifically methylates the guanine in position 2445 (m2G2445) and the guanine in position 2069 (m7G2069) of 23S rRNA	

	rase K/L					TGTAGGTCGCCGCCCAAGTATTGATACCAAACAACCGGAAATTACATTATA TGTGCTTGCGGGCAAACCGAGCATAACATATTGCCTAGATTTATCTGGTG ATTCATTACATAAACCGCGGCTACCGTCGTTTCATGACAGATGCGCCAATTA AAGAAAACCTCGCAGCAGCAATTTTACAAAAGCCAAGCTTAAGGAACGT AACCCGGAAATTGTACTIONGACCCAATGTGTGGTTCAGGCACATTTATTATT GAAGCATTAAATGATTTTAACTGATCGTGCGCCGGGTCTTGTACGTCGTTTT GGTTTTAATGGCTGGCATGGTCATGACCGTGAACCTTTGGCTTTCACTCAA GCCGAAGCAGCAGAGCGTCATGAAAAAGCACTTGAACAACCGTTACCTAA GTTTTATGCATACGATGCTGACTGGGAAGCTGTAAAAGCAACACGCGAAA AACATTATTGCAGCAGGTTTTGA ATGGATAAGAATGATGTTGTTAAGAAGATACTTGAATCAAAAAGTACGAA AACCTTGATTCAGATATTGTTGAAAAGGTTGTTCCATTTCTGAGAAGAAAT ATAAATTAAGGAAGTTGAGAATTATTCTAAAAAGAAATTGCATCAAATATG GGGTCTTACTATTCTGCCTATCCTAATTGGGATAAATTATTAAAAAGTAC AATCAGGGGCAGTTATCAATAGAAGATTTACTAAAGATTCATTCTTCGACG AATGAAAGAGTCGCAACATTAATGACTTTTACTTATGTATTTGGAAATA TCAAACATGTCTCATCTATTTTAGATTTTGGTTGTGGCTTCAATCCATTAGC TTTATACCAATGGAATGAAAATGAAAAATAATATATCATGCATACGATATT GATAGAGCTGAGATAGCTTTTTTTGAGTAGCATTATTGGGAAGTTAAAGACG ACGATAAAGTATAGGTTTTTTGAATAAAGAGAGTGATGTCTACAAAGGTACT TATGATGTAGTATTCCTTTTAAAGATGCTTCTGTGCTAAAACAGCAAGAT GTAAATATCTTGATTTTCTACAGCTTTTTCATACTCAAACCTTTGTAATAT CTTTTCCAATAAAGTCTTTATCTGGAAAGGAGAAGGGAATGGAAGAGAATT ACCAGCTATGGTTTGAATCTTTTACAAAAGGTTGGATAAAAATCCTTGATTC GAAGGTTATAGGGAATGAGTTAGTATATATTACTAGTGGATTTTCAGAAATA A ATGACGAAACAACAGAAAACGGCGCTGAACATGGCTAAATTCATCCAGAA TCAGTCTCTGCTCCTGCTGGAGAAGCTTAACGAGCTTGATCTTGATGCTG AGGCCGACCTTTGTGAGAACTCCACGATGATGCAGAGCATCTTTTCCGT ACACTCTCCTCGAGACTGGATTCATTACAGGATGGTAACTGA ATGATTGAATCTAATGCTGTCAGTATTGATCGTAGTCGTATCACACTTAAA CCCAATGTGGTTGTTGGTTATGAGGGAAACCCTTATAAAATTGTGAATGTT TTAAATGCTAATGACATTGTTATATCAAGTCTTGATTCCGGTACGTAGTTTAC
<i>rmtB</i>	16S rRNA (guanine(140 5)- N(7))- methyl transferase	L7	H	PFAM ribosomal RNA methyltransferase		
<i>rop</i>	Regulatory protein rop	A1, A2, A14	S	Rop protein		
<i>tnsB</i>	Transposon Tn7	L7, L9, L10	L	Mu transposase, C-terminal		

transp
osition
protein
TnsB

AGGTCAATTCTAAAAGTTTGACCGTTTTTGAGGGAGAAAATACGCTAACTG
AAAACCTAAACAAAGGTGATAAGGATATCAGCAATGAAGCTTGGCAGATT
GCTTTACAACGTTATGAGATTATTAAGCCTTTGATTGAGTATTCAACCACTG
AACTGGTTGAAAATAGAGCGAATGAATATGATGTAAATCGTTCTACATTGT
GGAAGTGGCTTAGAGATTATCGTGAAAATAACTCTCTAATGACGTTGATTC
CCAAAAACGAGGTTGGACAACCTGAAAATCAAGGTTGTACCTCAAGTG
AGCAATATCATTAAACAAGCTATCAATGATGAATATTTGAATGCGAAGAAA
CCTAGTATTTCCAAAACGATTGAAATTGTAAAAGCAGAGTGTTCCACGGCTT
CAGCTTGAAGCACCACATGAAAACCTATTTCGTAGAAGAATCGAGGCTTTA
AATAATTACCAAGTAACGAAAGCACGTCCTGGATCTAAAGCTGCAATAGAT
AAATTTAAGGCTGCTGCGGGTTCATTCCCAAATGCTGATTACCCTTTGGCA
TATGTGCAGATTGATCATACACCTCTAGATATTGAAATCGTGGATGACGAG
TTCAGAGAAGCGATTGGAAAACCTCACTTAACCTTGGCAATCGATGTTTTTC
AGTCGGATGATTGTTGGATATTACCTATCACTAGAAGCACCAAGTACGACT
TCAGTGGCGATGTGTATAGCATCTTGTATTTTGTCAAAAAACGTAAGTTG
ATCGAACTAAATATTGATGCTGAATGGCAGGTGCAAGGTATTATGGACTC
GGTGCATACAGATAATGGACCCGATTTTGAACCAATCACATCAGTAAAGC
TTGCCTAAAATATGGGATTCATTGGGAATATAGACCAATTGGTGGGGCGA
GATTTGGCGGTCATATTGAACGTATGCTTGGTATTGTCAATCTTGAATGC
ATGTGCTGGATGGTACAACTTTTTCAAATGTGCAGCAGCGAGGAACTTAT
GATTCCGCTAAACAAGCCTGTATGACATTAAGGAACTTGAATACTATATC
GTTTATTGGATTACTAAAGTTTATCATCAGAAGAAGCATTACAGCTTTAGGAA
CATCACCTATTGTTAAATGGGAAGAAGGTGTTTGGGGAACGAAGACGACA
GCAGGAACAGGTTTGAAGAACGTGTATCAGATGAAGATACCTTGTTTATT
GATTTTCTACCTGAATTTGAAGCGACCATTACAGGTACAGGGGTGCAGAA
AGATAACTTATTCTATTTTGCAGACTGTTTAAGGCAATGGGTTAATTCTATT
GATCCAGACGACAATAATCGAAAGCGCAAAAAGAAAATTCTTATTTAAGCGT
GATCCGAGGGATATTTACAGATTTGGTTTTATGAGCCATTTTCTAATACC
TATTTTAAAGTACCCACTGCAAAGCGAGAGATACCACCGATTAGCTTATTT
GAATATAGACAGGTACAGAACTATCTCAAGAATGGGCGTCAGGATGTTCA
GAATCAGGATGAGATTTACAAGGCGATTTTACATCTACGTGAGCAGCTTAA
TCAGGCTCGAAGCTTGACTCGAAAGCAACGTGCGTCAAACCAAAGAAAA
AGGAAAATGAAAAGCCATTACTCAGCTATCCGAACAGAATCAGTCTAAAA

tnsB_1

hypothetical
protein

L6, L9,
L10

L

Mu transposase, C-terminal

AGGCTGTAGTTTCTGAAAGCCTTCAGACAAGTGATGATTTGTGGAGTACTC
CTTTAACTGCATTTGATGATTTAAGGTGA
ATGATTGAATCTAATGCTGTCAATATCGATCGTAGTCGTATCATTCTAAAAC
CCAATGTGGTTGTCGGTTATGAGGGGCAACCCTATAAAATTGTGAATGTTT
TAAATGCCAATGACATTGTTATATCAAGTCTTGATTCCGGTACGTAGTCTAC
AGGTCAATGCTAGAAGTTTGACCATTTTTAATGGGGGAAATACGCTAACTG
AGCGTTTGAATAAGGGTGATCAGGATATTAGTAATGAAGCTTGGCAGATT
GCTTTACAACGTTATGAGATTATTAAGCCTTTGATTGAGTATTCAACGACT
GAACTGGTCGAAAATAGGGCGAATGAATATGATGTGAATCGCTCAACATT
GTGGAAATGGCTTAAAGATTATCGGGAAAATAATTCTCTAATGGCATTGAT
TCCCCAAAAACGAGGTTGGACAACGAAAATCAAGGTTGTCACCTCAAG
TGAGCAATATCATTAAACAAGCTATCAATGATGAATTTTGAATGCGAAGA
AACCCAGTATTTCCAAAACGATTGAAATTGAAAAGCAGAGTGTCTCGGC
TTCAGCTTGAAGCACCGCATGAAAACCTATTCGTAGAAGAATCGAGGCTT
TAAATGATTACCAAGTAACGAAAGCACGTCTTGGATCTAAAGCTGCAATAG
ATAAATTTAAGGCTGCTGCTGGTTCATTTCCAAATGCTGATTACCCTTTGG
CATATGTGCAGATTGATCATAACCTCTAGATATTGAAATCGTGGATGACG
AGTTTAGAGAAGCGATTGGAAAACCTCATTTAACCTTGGCAATCGATGTTT
TCAGTCGGATGATCGTAGGGTATTACCTATCACTAGAAGCGCCAAGTACG
ACTTCAGTGGCGATGTGTATCGCATCTTGTATTTTGTCAAAAAACGTAAG
TTGATCGAACTCGATATTGATGCTGAATGGCAGGTGCAAGGGATTATGGA
CTCGGTGCATACAGATAATGGCCCCGATTTTAGAACCAATCACATCAGTAA
AGCTTGTCTAAAGTATGGGATTCATTGGGAATACCGACCAATTGGTGGGG
CAAGATTTGGCGGTCATATCGAGCGTATGCTTGGTATCGTCAATCTTGAAA
TGCATGTCTTGGATGGTACAACCTTTTTCAAATGTACAGCAGCGTGGGACTT
ATGATTCCGCTAAACAAGCCTGTATGACATTAAGAGCTTGAATACTATA
TCGTGTATTGGATTACTAAAGTTTATCACCAGAAGAAACATTCCGCTTTAG
GAACATCCCCAATCGTTAAATGGGAAGAAGGTGTTTGGGGGACGAAGAC
GACAGCAGGTACAGGTTTGAAGAACGTGTATCAGATGAAGATACCTTGT
TTATTGATTTTCTACCTGAATTTGAAGCGACTATTCAGCGTACAGGGGTGC
AAAAAGATAACTTATTCTATTTTGCCGACTGTTTAAGACAATGGGTAAATTC
TATTGATCCAGAGGACAATAATCGAAAGCGCAAAAGAAAATTCTTGTTTAA
ACGTGATCCGAGGGATATTTACAGATTTGGTTTTATGAGCCATATTCTAA

tnsB_2

Transposon
Tn7
transposition
protein
TnsB

L8, L10

L

Mu transposase, C-terminal

TACCTATTTTAAAGTACCTACCGCAAAGCGAGAGATACCACCGATTAGCTT
ATTTGAATATAAGCAGGTTTCAGAACTATCTCAAAAGTGAACGTCAGGATGT
TCAGAATCAGGATGAGATTTACAAGGCGATTTTTGCATCTACGTGATCAACT
TAATCAGGCTAGAAGCTTGACTCGAAAGCAACGTCGCTCAAACCAAAGAA
AAAAGGAAAATGCAAAAGCCATTACTCAGCTTTCCGAACAGAATCAGTCTA
AAAAGGCTGTAGTTTCTGAAAGTCTTCAGACAAGTGATGATTTGTGGAGTA
CTCCTTTAACTGCATTTGACGATTTAAGGTGA
GTGAGCAATATCATTAAACAAGCTATCAATGATGAATATTTGAATGCGAAG
AAACCTAGTATTTCCAAAACGATTGAAATTGTAAAAGCAGAGTGTTACCGG
CTTCAGCTTGAAGCACCACATGAAAACCTCTATTCGTAGAGAATCGAGGCT
TTAAATAATTACCAAGTAACGAAAGCACGTCTTGGATCTAAAGCTGCAATA
GATAAATTTAAGGCTGCTGCGGGTTCATCCCAAATGCTGATTACCCTTTG
GCATATGTGCAGATTGATCATAACCTCTAGATATTGAAATCGTGGATGAC
GAGTTCAGAGAAGCGATTGGAAAACCTCACTTAACCTTGGCAATCGATGT
TTTCAGTCGGATGATTGTTGGATATTACCTATCACTAGAAGCACCAAGTAC
GACTTCAGTGGCGATGTGTATAGCATCTTGTATTTTGTCAAAAAACGTAA
GTTGATCGAACTAAATATTGATGCTGAATGGCAGGTGCGAAGGTATTATGGA
CTCGGTGCATACAGATAATGGACCCGATTTTAGAACCAATCACATCAGTAA
AGCTTGCCTAAAATATGGGATTCATTGGGAATATAGACCAATTGGTGGGG
CGAGATTTGGCGGTCATATTGAACGTATGCTTGGTATTGTCAATCTTGA
TGCATGTGCTGGATGGTACAACCTTTTCAAATGTGCAGCAGCGAGGAACT
TATGATTCCGCTAAACAAGCCTGTATGACATTAAGGAACTTGAATACTAT
ATCGTTTATTGGATTACTAAAGTTTATCATCAGAAGAAGCATTGAGCTTTAG
GAACATCACCTATTGTTAAATGGGAAGAAGGTGTTTGGGGAACGAAGACG
ACAGCAGGAACAGGTTTGAAGAACGTGTATCAGATGAAGATACCTTGTTT
ATTGATTTTCTACCTGAATTTGAAGCGACCATTGAGCGTACAGGGGTGCA
GAAAGATAACTTATTCTATTTTGCAGACTGTTTAAGGCAATGGGTTAATTCT
ATTGATCCAGACGACAATAATCGAAAGCGCAAAGAAAATTCTTATTTAAG
CGTGATCCGAGGGATATTTACAGATTTGGTTTTATGAGCCATTTTCTAAT
ACCTATTTTAAAGTACCCACTGCAAAGCGAGAGATACCACCGATTAGCTTA
TTTGAATATAGACAGGTACAGAACTATCTCAAGAATGGGCGTCAGGATGTT
CAGAATCAGGATGAGATTTACAAGGCGATTTTACATCTACGTGAGCAGCTT
AATCAGGCTCGAAGCTTGACTCGAAAGCAACGTCGCTCAAACCAAAGAAA

tuf1_1

Elonga
tion
factor
Tu
A2, A3,
A14

J

This protein promotes the GTP-
dependent binding of aminoacyl-tRNA to
the A-site of ribosomes during protein
biosynthesis

AAAGGAAAATGAAAAAGCCATTACTCAGCTATCCGAACAGAATCAGTCTAA
AAAGGCTGTAGTTTCTGAAAGCCTTCAGACAAGTGATGATTTGTGGAGTAC
TCCTTTAACTGCATTTGATGATTTAAGGTGA
ATGGCTAAAGCCAAGTTTGAACGTAATAAACCACACGTAAACGTGGGTAC
AATCGGTCACGTTGACCATGGTAAAACAACTTTAACTGCTGCGATTGCAAC
AATTTGTGCAAAAACCTTACGGCGGTGAAGCGAAAGATTACTCACAAATCGA
CTCAGCACCTGAAGAAAAAGCACGTGGTATTACAATTAATACATCACACGT
AGAATACGATTCTCCAACCTCGTCACTACGCACACGTTGACTGCCCAGGCC
ACGCCGACTACGTTAAAAACATGATCACTGGTGCTGCTCAGATGGACGGC
GCGATCCTTGTATGTGCTGCAACTGACGGTCCAATGCCACAAAACCTCGTGA
GCACATCCTTCTTTCTCGTCAGGTAGGTGTACCTTACATCATCGTATTCTT
AAACAAATGCGACCTTGTGGATGACGAAGAATTACTTGAATTAGTAGAAAT
GGAAGTACGTGAACTTCTTTCTACTTATGACTTCCCAGGTGATGACACTCC
AGTAATCCGTGGTTCAGCTCTTGCAGCGCTTAACGGTGAAGCTGGTCCCTT
ACGGTGAAGAATCAGTTCTTGCTCTTGTAGCAGCACTTGACTCTTATATCC
CAGAGCCAGAGCGTGCAATCGACAAAGCATTCTTGATGCCAATCGAAGAC
GTATTCTCAATTTCTGGTCGTGGTACAGTAGTAACAGGCCGTGTTGAAGCT
GGTATCATCAAAGTTGGTGAAGAAGTAGAGATCGTTGGTATTAAGATACA
GTTAAAACAACCTGTAACCTGGCGTAGAAATGTTCCGTAAACTTCTTGACGAA
GGCCGTGCAGGTGAGAACTGTGGTATCTTACTTCGTGGTACTAAGCGTGA
AGAAGTACAACGTGGTCAAGTACTTGCTAAACCAGGTACAATCAAGCCGC
ACACTAAATTCGACGCAGAAGTATACGTACTTTCTAAAGAAGAAGGTGGTC
GTCACACTCCATTCTTAAATGGTTACCGTCCACAGTTCTACTTCCGTACAA
CTGACGTAACCTGGTGCGATCCAGTTGAAAGAAGGCGTTGAAATGGTAATG
CCAGGTGACAACGTTGAAATGTCAGTAGAATTAATCCACCCAATCGCAAT
GGACCCAGGTCTACGTTTTGCGATCCGTGAAGGTGGTCTGACTGTAGGTG
CTGGTGTGTTGCTAAAGTAACTGCATAA

tuf1_2

Elonga
tion
factor
Tu
A2, A3,
A14

J

This protein promotes the GTP-
dependent binding of aminoacyl-tRNA to
the A-site of ribosomes during protein
biosynthesis

ATGGCTAAAGCCAAGTTTGAACGTAATAAACCACACGTAAACGTGGGTAC
AATCGGTCACGTTGACCATGGTAAAACAACTTTAACTGCTGCGATTGCAAC
AATTTGTGCAAAAACCTTACGGCGGTGAAGCGAAAGATTACTCACAAATCGA
CTCAGCACCTGAAGAAAAAGCACGTGGTATTACAATTAATACATCACACGT
AGAATACGATTCTCCAACCTCGTCACTACGCACACGTTGACTGCCCAGGCC
ACGCCGACTACGTTAAAAACATGATCACTGGTGCTGCTCAGATGGACGGC

GCGATCCTTGATGTGCTGCAACTGACGGTCCAATGCCACAACTCGTGA
GCACATCCTTCTTTCTCGTCAGGTAGGTGTACCTTACATCATCGTATTCTT
AAACAAATGCGACCTTGTGGATGACGAAGAATTACTTGAATTAGTAGAAAT
GGAAGTACGTGAACTTCTTTCTACTTATGACTTCCCAGGTGATGACTCC
AGTAATCCGTGGTTCAGCTCTTGCAGCGCTTAACGGTGAAGCTGGTCCTT
ACGGTGAAGAATCAGTTCTTGCTCTTGTAGCAGCACTTGACTIONTACATCC
CAGAGCCAGAGCGTGCAATCGACAAAGCATTCTTGATGCCAATCGAAGAC
GTATTCTCAATTTCTGGTCGTGGTACAGTAGTAACAGGCCGTGTTGAAGCT
GGTATCATCAAAGTTGGTGAAGAAGTAGAGATCGTTGGTATTAAGATACA
GTTAAAACAACCTGTAACGGCGTAGAAATGTTCCGTAACCTTCTTGACGAA
GGCCGTGCAGGTGAGAACTGTGGTATCTTACTTCGTGGTACTAAGCGTGA
AGAAGTACAACGTGGTCAAGTACTTGCTAAACCAGGTACAATCAAGCCGC
ACACTAAATTCGACGCAGAAGTATACGTACTTTCTAAAGAAGAAGGTGGTC
GTCACACTCCATTCTTAAATGGTTACCGTCCACAGTTCTACTTCCGTACAA
CTGACGTAACCTGGTGCATCCAGTTGAAAGAAGGCGTTGAAATGGTAATG
CCAGGTGACAACGTTGAAATGTCAGTAGAATTAATCCACCCAATCGCAAT
GGACCCAGGTCTACGTTTTGCGATCCGTGAGGGTGGTCGTACTIONTGTAGGT
GCTGGTGTGTTGCTAAAGTAACTGCATAA

<i>tufB_1</i>	Elongation factor Tu 2	A1	J	This protein promotes the GTP-dependent binding of aminoacyl-tRNA to the A-site of ribosomes during protein biosynthesis	ATGTCAGTAGAATTAATCCACCCAATCGCAATGGACCCAGGTCTACGTTTT GCGATCCGTGAAGGTGGTCGTACTIONTGTAGGTGCTGGTGTGTTGCTAAAGT AACTGCATAA
<i>tufB_2</i>	Elongation factor Tu 2	A1	J	This protein promotes the GTP-dependent binding of aminoacyl-tRNA to the A-site of ribosomes during protein biosynthesis	ATGTCAGTAGAATTAATCCACCCAATCGCAATGGACCCAGGTCTACGTTTT GCGATCCGTGAGGGTGGTCGTACTIONTGTAGGTGCTGGTGTGTTGCTAAAGT AACTGCATAA

Figure S1 Distribution of the antibiotic resistance genes and virulence genes. The blue and white colour represented presence and absence of the gene.

